CHAPTER 2

The Process View of the Organization

Matching supply and demand would be easy if business processes would be instantaneous and could immediately create any amount of supply to meet demand. Understanding the questions of "Why are business processes not instantaneous?" and "What constrains processes from creating more supply?" is thereby at the heart of operations management. To answer these questions, we need to take a detailed look at how business processes actually work. In this chapter, we introduce some concepts fundamental to process analysis. The key idea of the chapter is that it is not sufficient for a firm to create great products and services; the firm also must design and improve its business processes that supply its products and services.

To get more familiar with the process view of a firm, we now take a detailed look behind the scenes of a particular operation, namely the Department of Interventional Radiology at Presbyterian Hospital in Philadelphia.

2.1 Presbyterian Hospital in Philadelphia

LO 2-1

Explain the process view of a firm.

Interventional radiology is a subspecialty field of radiology that uses advanced imaging techniques such as real-time X-rays, ultrasound, computed tomography, and magnetic resonance imaging to perform minimally invasive procedures.

Over the past decades, interventional radiology procedures have begun to replace an increasing number of standard "open surgical procedures" for a number of reasons. Instead of being performed in an operating room, interventional radiology procedures are performed in an angiography suite (see Figure 2.1). Although highly specialized, these rooms are less expensive to operate than conventional operating rooms. Interventional procedures are often safer and have dramatically shorter recovery times compared to traditional surgery. Also, an interventional radiologist is often able to treat diseases such as advanced liver cancer that cannot be helped by standard surgery.

FIGURE 2.1 Example of a Procedure in an Interventional Radiology Unit

U.S. Air Force photo/Staff Sgt. Robert Barnett

Note: If this image is not available, please replace it with a similar image from the DAL.

Although we may not have been in the interventional radiology unit, many, if not most, of us have been in a radiology department of a hospital at some point in our life. From the perspective of the patient, the following steps need to take place before the patient can go home or return to his or her hospital unit. In process analysis, we refer to these steps as *activities*:

- Registration of the patient.
- Initial consultation with a doctor; signature of the consent form.
- Preparation for the procedure.
- The actual procedure.
- Removal of all equipment.
- Recovery in an area outside the angiography suite.
- Consultation with the doctor.

Figure 2.2 includes a graphical representation of these steps, called a ***Gantt diagram*** (named after the 19th-century industrialist Henry Gantt). It provides several useful pieces of information.

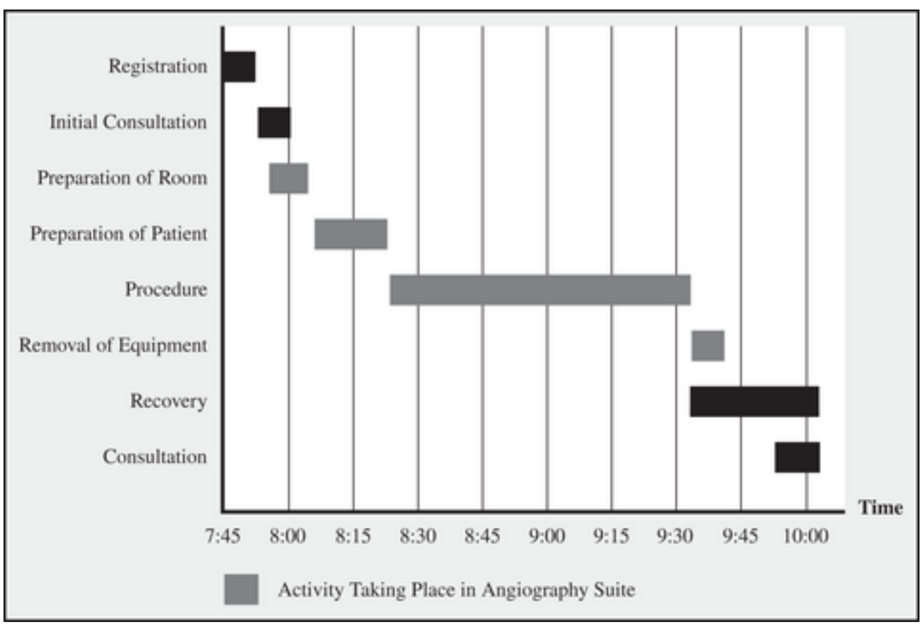


FIGURE 2.2 Gantt Chart Summarizing the Activities for Interventional Radiology

First, the Gantt chart allows us to see the process steps and their durations, which are also called ***activity times*** or ***processing times***. The duration simply corresponds to the length of the corresponding bars. Second, the Gantt diagram also illustrates the dependence between the various process activities. For example, the consultation with the doctor can only occur once the patient has arrived and been registered. In contrast, the preparation of the angiography suite can proceed in parallel to the initial consultation.

You might have come across Gantt charts in the context of project management. Unlike process analysis, project management is typically concerned with the completion of one single project. (See **Chapter 12** for more details on project management.) The most well-known concept of project management is the ***critical path***. The critical path is composed of all those activities that—if delayed—would lead to a delay in the overall completion time of the project, or—in this case—the time the patient has completed his or her stay in the radiology unit.

In addition to the 8 steps described in the Gantt chart of **Figure 2.2**, most of us associate another activity with hospital care: waiting. Strictly speaking, waiting is not really an activity, as it does not add any value to the process. However, waiting is nevertheless relevant. It is annoying for the patient and can complicate matters for the hospital unit. For this reason, waiting times take an important role in operations management. **Figure 2.3** shows the actual durations of the activities for a patient arriving at 12:30, as well as the time the patient needs to wait before being moved to the angiography suite.

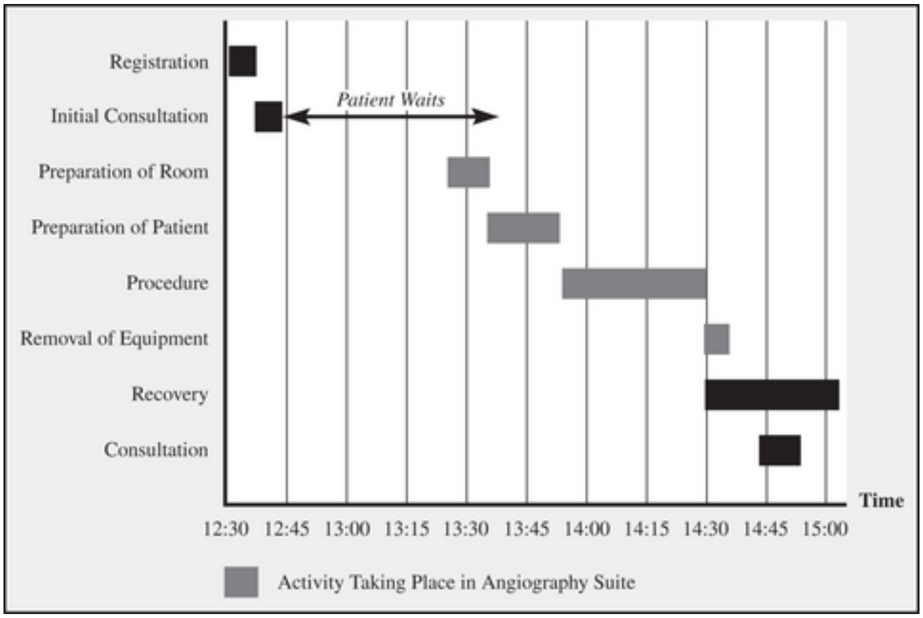


FIGURE 2.3 Gantt Chart Summarizing the Activities for a Patient Arriving at 12:30

Page 12

>But why is there waiting time? Waiting is—to stay in the medical language for the moment—a symptom of supply–demand mismatch. If supply would be unlimited, our visit to the hospital would be reduced to the duration of the activities outlined in **Figure 2.2** (the critical path). Imagine visiting a hospital in which all the nurses, technicians, doctors, and hospital administrators would just care for you!

Given that few of us are in a position to receive the undivided attention of an entire hospital unit, it is important that we not only take the egocentric perspective of the patient but look at the hospital operations more broadly. From the perspective of the hospital, there are many patients "flowing" through the process.

Page 13

The people and the equipment necessary to support the interventional radiology process deal with many patients, not just one. We refer to these elements of the process as the ***process resources***. Consider, for example, the perspective of the nurse and how she/he spends her/his time in the department of interventional radiology. Obviously, radiology from the viewpoint of the nurse is not an exceptional event, but a rather repetitive endeavor. Some of the nurse's work involves direct interaction with the patient; other work—while required for the patient—is invisible to the patient. This includes the preparation of the angiography suite and various aspects of medical record keeping.

Given this repetitive nature of work, the nurse as well as the doctors, technicians, and hospital administrators think of interventional radiology as a process, not a project. Over the course of the day, they see many patients come and go. Many hospitals, including Presbyterian Hospital in Philadelphia, have a "patient log" that summarizes at what times patients arrive at the unit. This patient log provides a picture of demand on the corresponding day. The patient log for December 2 is summarized in **Table 2.1**.

TABLE 2.1 Patient Log on December 2

Number	Patient Name	Arrival Time	Room Assignment
1		7:35	Main room
2		7:45	
3		8:10	
4		9:30	Main room
5		10:15	Main room
6		10:30	Main room
7		11:05	
8		12:35	Main room
9		14:30	Main room
10		14:35	
11		14:40	

Many of these arrivals were probably scheduled some time in advance. Our analysis here focuses on what happens to the patient once he/she has arrived in the interventional radiology unit. A separate analysis could be performed, looking at the process starting with a request for diagnostics up to the arrival of the patient.

Given that the resources in the interventional radiology unit have to care for 11 patients on December 2, they basically need to complete the work according to 11 Gantt charts of the type outlined in **Figure 2.2**. This—in turn—can lead to waiting times. Waiting times arise when several patients are "competing" for the same limited resource, which is illustrated by the following two examples.

First, observe that the critical path for a typical patient takes about 2 hours. Note further that we want to care for 11 patients over a 10-hour workday. Consequently, we will have to take care of several patients at once. This would not be a problem if we had unlimited resources, nurses, doctors, space in the angiography suites, and so forth. However, given the resources that we have, if the Gantt charts of 2 patients are requesting the same resource simultaneously, waiting times result. For example, the second patient might require the initial consultation with the doctor at a time when the doctor is in the middle of the procedure for patient 1. Note also that patients 1, 4, 5, 6, 8, and 9 are assigned to the same room (the unit has a main room and a second room used for simpler cases), and thus they are also potentially competing for the same resource.

A second source of waiting time lies in the unpredictable nature of many of the activities. Some patients will take much longer in the actual procedure than others. For example, patient 1 spent 1:50 hours in the procedure, while patient 9 was in the procedure for 2:30 hours (see **Figure 2.4**). As an extreme case, consider patient 5, who refused to sign the consent form and left the process after only 15 minutes.

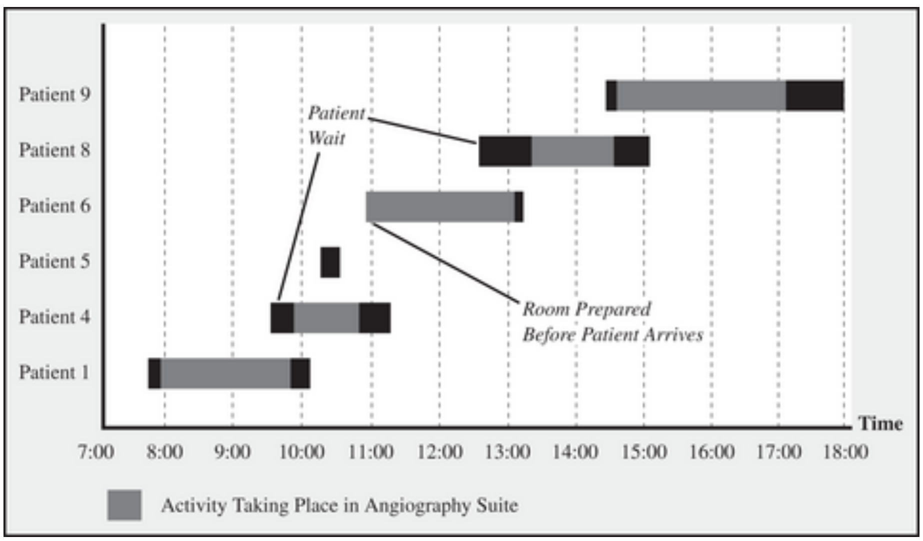


FIGURE 2.4 Time Patient Spent in the Interventional Radiology Unit (for Patients Treated in Main Room Only), Including Room Preparation Time

Such uncertainty is undesirable for resources, as it leaves them “flooded” with work at some moments in the day and “starved” for work at other moments. **Figure 2.5** summarizes at what moments in time the angiography suite was used on December 2.

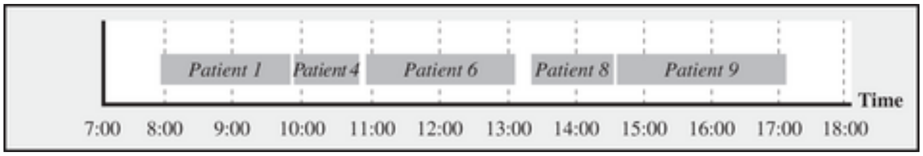


FIGURE 2.5 Usage of the Main Room

By now, we have established two views on the interventional radiology:

- The view of the patient for whom the idealized stay is summarized in **Figure 2.2**. Mismatches between supply and demand from the patient's perspective mean having a unit of demand (i.e., the patient) wait for a unit of supply (a resource).
- The view of the resources (summarized in **Figure 2.5**), which experience demand–supply mismatches when they are sometimes "flooded" with work, followed by periods of no work.

As these two perspectives are ultimately two sides of the same coin, we are interested in bringing these two views together. This is the fundamental idea of process analysis.

2.2 Three Measures of Process Performance

LO 2-2

Describe the three basic measures of process performance: flow rate, inventory, and flow time.

At the most aggregate level, a process can be thought of as a "black box" that uses ***resources*** (labor and capital) to transform *inputs* (undiagnosed patients, raw materials, unserved customers) into *outputs* (diagnosed patients, finished goods, served customers). This is shown in **Figure 2.6**. **Chapter 3** explains the details of constructing figures like **Figure 2.6**, which are called ***process flow diagrams***. When analyzing the processes that lead to the supply of goods and services, we first define our unit of analysis.

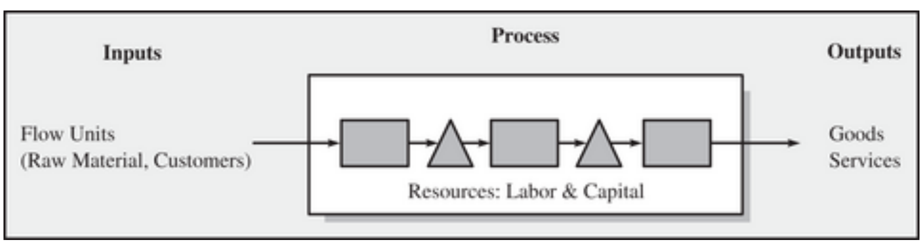


FIGURE 2.6 **The Process View of an Organization**

In the case of the interventional radiology unit, we choose patients as our ***flow unit***. Choosing the flow unit is typically determined by the type of product or service the supply process is dealing with; for example, vehicles in an auto plant, travelers for an airline, or gallons of beer in a brewery.

As suggested by the term, flow units flow through the process, starting as input and later leaving the process as output. With the appropriate flow unit defined, we next can evaluate a process based on three fundamental process performance measures:

- The number of flow units contained within the process is called the ***inventory*** (in a production setting, it is referred to as ***work-in-process, WIP***). Given that our focus is not only on production processes, inventory could take the form of the number of insurance claims or the number of tax returns at the IRS. There are various reasons why we find inventory in processes, which we discuss in greater detail below. While many of us might initially feel uncomfortable with the wording, the inventory in the case of the interventional radiology unit is a group of patients.
- The time it takes a flow unit to get through the process is called the ***flow time***. The flow time takes into account that the item (flow unit) may have to wait to be processed because there are other flow units (inventory) in the process potentially competing for the same resources. Flow time is an especially important performance metric in service environments or in other business situations that are sensitive to delays, such as make-to-order production, where the production of the process only begins upon the arrival of the customer order. In a radiology unit, flow time is something that patients are likely to care about: it measures the time from their arrival at the interventional radiology unit to the time patients can go home or return to their hospital unit.
- Finally, the rate at which the process is delivering output (measured in [flow units/unit of time], e.g., units per day) is called the ***flow rate*** or the *throughput rate*. The maximum rate with which the process can generate supply is called the ***process capacity***. On December 2, the throughput of the interventional radiology unit was 11 patients per day.

Table 2.2 provides several examples of processes and their corresponding flow rates, inventory levels, and flow times.

TABLE 2.2 Examples of Flow Rates, Inventories, and Flow Times

	Immigration Authorities	Champagne Industry	MBA Program	Large PC Manufacturer
Flow unit	Application for immigration benefit	Bottle of champagne	MBA student	Computer
Flow rate/throughput	Approved or rejected visa cases: 6.3 million per year	260 million bottles per year	600 students per year	5,000 units per day
Flow time	Average processing time: 7.6 months	Average time in cellar: 3.46 years	2 years	10 days

	Immigration Authorities	Champagne Industry	MBA Program	Large PC Manufacturer
Inventory	Pending cases: 4.0 million cases	900 million bottles	1,200 students	50,000 computers

You might be somewhat irritated that we have moved away from the idea of supply and demand mismatch for a moment. Moreover, we have not talked about profits so far. However, note that increasing the maximum flow rate (capacity) avoids situations where we have insufficient supply to match demand. From a profit perspective, a higher flow rate translates directly into more revenues (you can produce a unit faster and thus can produce more units), assuming your process is currently ***capacity-constrained***, that is, there is sufficient demand that you could sell any additional output you make.

Shorter flow times reduce the time delay between the occurrence of demand and its fulfillment in the form of supply. Shorter flow times therefore also typically help to reduce demand–supply mismatches. In many industries, shorter flow times also result in additional unit sales and/or higher prices, which makes them interesting also from a broader management perspective.

Lower inventory results in lower working capital requirements as well as many quality advantages that we explore later in this book. A higher inventory also is directly related to longer flow times (explained below). Thus, a reduction in inventory also yields a reduction in flow time. As inventory is the most visible indication of a mismatch between supply and demand, we will now discuss it in greater detail.

2.3 Little’s Law

LO 2-3

Explain Little’s Law and when it is useful.

Accountants view inventory as an asset, but from an operations perspective, inventory often should be viewed as a liability. This is not a snub on accountants; inventory *should* be an asset on a balance sheet, given how accountants define an asset. But in common speech, the word *asset* means “desirable thing to have” and the dictionary defines *liability* as “something that works to one’s disadvantage.” In this sense, inventory can clearly be a liability. This is most visible in a service process such as a hospital unit, where patients in the waiting room obviously cannot be counted toward the assets of the health care system.

Let’s take another visit to the interventional radiology unit. Even without much medical expertise, we can quickly find out which of the patients are currently undergoing care from some resource and which are waiting for a resource to take care of them. Similarly, if we took a quick walk through a factory, we could identify which parts of the inventory serve as raw materials, which ones are work-in-process, and which ones have completed the production process and now take the form of finished goods inventory.

However, taking a single walk through the process—dishwasher factory or interventional radiology unit—will not leave us with a good understanding of the underlying operations. All it will give us is a snapshot of what the process looked like at one single moment in time. Unfortunately, it is this same snapshot approach that underlies most management (accounting) reports: balance sheets itemize inventory into three categories (raw materials, WIP, and finished goods); hospital administrators typically distinguish between pre- and postoperative patients. But such snapshots do not tell us *why* these inventories exist in the first place! Thus, a static, snapshot approach neither helps us to analyze business processes (why is there inventory?) nor helps us to improve them (is this the right amount of inventory?).

Page 17

Now, imagine that instead of our single visit to the hospital unit, we would be willing to stay for some longer period of time. We arrive early in the morning and make ourselves comfortable at the entrance of the unit. Knowing that there are no patients in the interventional radiology unit overnight, we then start recording any arrival or departure of patients. In other words, we collect data concerning the patient inflow and outflow.

At the end of our stay, we can plot a graph similar to **Figure 2.7**. The upper of the two curves illustrates the cumulative number of patients who have entered the unit. The curve begins at time zero (7:00) and with zero patients. If we had done the same exercise in a unit with overnight patients, we would have recorded our initial patient count there. The lower of the two curves indicates the cumulative number of patients who have left the unit. **Figure 2.7** shows us that by noon, 7 patients have arrived, of which 5 have left the unit again.

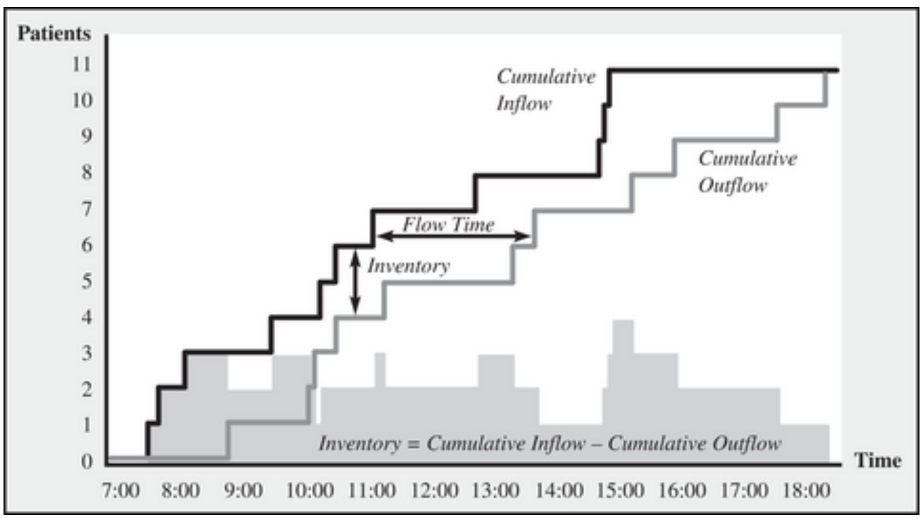


FIGURE 2.7 Cumulative Inflow and Outflow

At any given moment in time, the *vertical distance* between the upper curve and the lower curve corresponds to the number of patients in the interventional radiology unit, or—abstractly speaking—the inventory level. Thus, although we have not been inside the interventional radiology unit this day, we are able to keep track of the inventory level by comparing the cumulative inflow and outflow. For example, the inventory at noon consisted of 2 patients.

We also can look at the *horizontal distance* between the two lines. If the patients leave the unit in the same order they entered it, the horizontal gap would measure the exact amount of time each patient spent in the interventional radiology unit. More generally, given that the length of stay might vary across patients and patients do not necessarily leave the unit exactly in the same sequence in which they entered it, the average gap between the two lines provides the average length of stay.

Thus, **Figure 2.7** includes all three of the basic process performance measures we discussed on the previous page: flow rate (the slope of the two graphs), inventory (the vertical distance between the two graphs), and flow time (the horizontal distance between the two graphs).

Based on either the graph or the patient log, we can now compute these performance measures for December 2. We already know that the flow rate was 11 patients/day.

Next, consider inventory. Inventory changes throughout the day, reflecting the differences between inflow and outflow of patients. A "brute force" approach to compute average inventory is to count the inventory at every moment in time throughout the day, say every five minutes, and then take the average. For December 2, this computation yields an average inventory of 2.076 patients.

Next, consider the flow time, the time a patient spends in the unit. To compute that information, we need to add to the patient log, **Table 2.1**, the time each patient left the interventional radiology unit. The difference between arrival time and departure time would be the flow time for a given patient, which in turn would allow us to compute the average flow time across patients. This is shown in **Table 2.3** and is in many ways similar to the two graphs in **Figure 2.7**. We can easily compute that on December 2, the average flow time was 2 hours, 4 minutes, and 33 seconds, or 2.076 hours.

TABLE 2.3 **Calculation of Average Flow Time**

Number	Patient Name	Arrival Time	Departure Time	Flow Time
1		7:35	8:50	1:15
2		7:45	10:05	2:20
3		8:10	10:10	2:00
4		9:30	11:15	1:45
5		10:15	10:30	0:15
6		10:30	13:35	3:05
7		11:05	13:15	2:10
8		12:35	15:05	2:30
9		14:30	18:10	3:40

Number	Patient Name	Arrival Time	Departure Time	Flow Time
10		14:35	15:45	1:10
11		14:40	17:20	2:40
			Average	2:04:33

At this point, you might ask: "Does the average inventory always come out the same as the average flow time?" The answer to this question is a resounding *no*. However, the fact that the average inventory was 2.076 patients and the average flow time was 2.076 hours is no coincidence either.

To see how inventory and flow time relate to each other, let us review the three performance measures, flow rate, flow time, and inventory:

- Flow rate = 11 patients per day, which is equal to 1 patient per hour.
- Flow time = 2.076 hours.
- Inventory = 2.076 patients.

Thus, while inventory and flow time do not have to—and, in fact, rarely are—equal, they are linked in another form. We will now introduce this relationship as **Little's Law** (named after John D. C. Little).

$$\text{Average inventory} = \text{Average flow rate} \times \text{Average flow time} \qquad \text{(Little's Law)}$$

Many people think of this relationship as trivial. However, it is not. Its proof is rather complex for the general case (which includes—among other nasty things—variability) and by mathematical standards is very recent.

Little's Law is useful in finding the third performance measure when the other two are known. For example, if you want to find out how long patients in a radiology unit spend waiting for their chest X-ray, you could do the following:

1. Observe the inventory of patients at a couple of random points during the day, giving you an average inventory. Let's say this number is 7 patients: four in the waiting room, two already changed and waiting in front of the procedure room, and one in the procedure room.
2. Count the procedure slips or any other records showing how many patients were treated that day. This is the day's output. Let's say there were 60 patients over a period of 8 hours; we could say that we have a flow rate of 60/8 = 7.5 patients/hour.

3. Use Little's Law to compute Flow time = Inventory/Flow rate = 7/7.5 = 0.933 hour = 56 minutes. This tells us that, on average, it takes 56 minutes from the time a patient enters the radiology unit to the time his or her chest X-ray is completed. Note that this information would otherwise have to be computed by collecting additional data (e.g., see **Table 2.3**).

When does Little's Law hold? The short answer is *always*. For example, Little's Law does not depend on the sequence in which the flow units (e.g., patients) are served (remember FIFO and LIFO from your accounting class?). (However, the sequence could influence the flow time of a particular flow unit, e.g., the patient arriving first in the morning, but not the average flow time across all flow units.) Furthermore, Little's Law does not depend on randomness: it does not matter if there is variability in the number of patients or in how long treatment takes for each patient; all that matters is the average flow rate of patients and the average flow time.

In addition to the direct application of Little's Law, for example, in the computation of flow time, Little's Law is also underlying the computation of inventory costs as well as a concept known as inventory turns. This is discussed in the following section.

2.4 Inventory Turns and Inventory Costs

LO 2-4

Use Little's Law to calculate inventory turns and inventory costs.

Using physical units as flow units (and, hence, as the inventory measure) is probably the most intuitive way to measure inventory. This could be vehicles at an auto retailer, patients in the hospital, or tons of oil in a refinery.

However, working with physical units is not necessarily the best method for obtaining an aggregate measure of inventory across different products: there is little value to saying you have 2,000 units of inventory if 1,000 of them are paper clips and the remaining 1,000 are computers. In such applications, inventory is often measured in some monetary unit, for example, $5 million worth of inventory.

Measuring inventory in a common monetary unit facilitates the aggregation of inventory across different products. This is why total U.S. inventory is reported in dollars. To illustrate the notion of monetary flow units, consider Kohl's Corp, a large U.S. retailer. Instead of thinking of Kohl's stores as sodas, toys, clothes, and bathroom tissues (physical units), we can think of its stores as processes transforming goods valued in monetary units into sales, which also can be evaluated in the form of monetary units.

As can easily be seen from Kohl's balance sheet, on January 30, 2021, the company held an inventory valued at $2.590 billion (see Table 2.4). Given that our flow unit now is the "individual dollar bill," we want to measure the flow rate through Kohl's operation.

TABLE 2.4 Excerpts from Financial Statements of Kohl's and Walmart (All Numbers in Millions)

	2021
Kohl's	
Revenue	$ 15,955
Cost of Goods Sold (COGS)	$ 10,360
Inventory	$ 2,590

	2021
Gross Profit (Revenue - COGS)	$ 5,595
Walmart	
Revenue	$ 559,151
Cost of Goods Sold (COGS)	$ 420,315
Inventory	$ 44,949
Gross Profit (Revenue - COGS)	$ 138,836

Source: Taken from Nasdaq.com.

The direct approach would be to take "revenue" as the resulting flow. Yet, this measure is inflated by Kohl's gross profit margin; that is, a dollar of sales is measured in revenue dollars, while a dollar of inventory is measured, given the present accounting practice, in a cost dollar. Thus, the appropriate measure for flow rate is the cost of goods sold, or COGS for short.

With these two measures—flow rate and inventory—we can apply Little's Law to compute what initially might seem a rather artificial measure: How long does the average flow unit (dollar bill) spend within the Kohl's system before being turned into sales, at which point the flow units will trigger a profit intake? This corresponds to the definition of flow time.

$$\text{Flow rate} = \text{Cost of good sold} = \$10,360 \text{ million/year}$$
$$\text{Inventory} = \$2,590 \text{ million}$$

Hence, we can compute flow time via Little's Law as

$$\text{Flow time} = \frac{\text{Inventory}}{\text{Flow rate}}$$

$$= \$2,590 \text{ million} / \$10,360 \text{ million} / \text{year} = 0.25 \text{ year} = 91.25 \text{ days}$$

Thus, we find that it takes Kohl's—on average— 91.25 days to translate a dollar investment into a dollar of—hopefully profitable—revenues.

This calculation underlies the definition of another way of measuring inventory, namely in terms of *days of supply*. We could say that Kohl's has 91.25 days of inventory in their process. In other words, the average item we find at Kohl's spends 91.25 days in Kohl's supply chain.

Alternatively, we could say that Kohl's turns over its inventory 365 days/year/91.25 days = 4 times per year. This measure is called ***inventory turns***. Inventory turns are the common benchmark in the retailing environment and other supply chain operations:

$$\text{Inventory turns} = \frac{1}{\text{Flow time}}$$

To illustrate this application of Little's Law further, consider Walmart, one of Kohl's competitors. Repeating the same calculations as outlined on the previous page, we find the following data about Walmart:

$$\begin{array}{ll}
\text{Cost of goods sold} & = \$420,315\ \text{million / year} \\
\text{Inventory} & = \$44,949\ \text{million} \\
\text{Flow}\ \ \text{time} & = \$44,949\ \text{million} / \$420,315\ \text{million / year} \\
 & = 0.1069\ \text{year} = 39.03\ \text{days} \\
\text{Inventory turns} & = 1/\ (39.03\ \text{day}\,s) \\
 & = 365\ \text{days /year} \times 1/(39.03\ \text{day}\,s) = 9.35\ \text{turns per year}
\end{array}$$

Thus, we find that Walmart is able to achieve substantially higher inventory turns than Kohl's. **Table 2.5** summarizes inventory turn data for various segments of the retailing industry. **Table 2.5** also provides information about gross margins in various retail settings (keep them in mind the next time you haggle for a new pair of jeans or watch!).

A word of caution. Usually, Kohl's inventory has been very stable from year to year. For example, Kohl's inventory was $3,537 Million at the beginning of 2020, $3,475 Million at the beginning of 2019, and $3,542 Million at the beginning of 2018. In our calculation above, however, we have used the inventory at the beginning of 2021, which was $2,590 Million. This lower inventory is likely a result of the Covid-19 pandemic which apparently impacted Kohl's more profoundly than it did impact Walmart.

Page 21

TABLE 2.5 Inventory Turns and Margins for Selected Retail Segments

Retail Segment	Company	Annual Inventory Turns	Gross Margin
Apparel and accessory	GAP	3.9	37.0%

Retail Segment	Company	Annual Inventory Turns	Gross Margin
Department stores	Macy's	2.9	38.0%
Food stores	Kroger	13.4	23.2%
Home furniture/equipment	Bed Bath & Beyond	2.63	31.7%
Jewelry	Tiffany	0.68	61.0%
Consumer electronics	Best Buy	6.35	23.0%
Variety stores	Walmart	9.35	24.8%

Inventory requires substantial financial investments. Moreover, the inventory holding cost is substantially higher than the mere financial holding cost for a number of reasons:

- Inventory might become obsolete (think of the annual holding cost of a microprocessor).
- Inventory might physically perish (you don't want to think of the cost of holding fresh roses for a year).
- Inventory might disappear (also known as theft or shrinkage).
- Inventory requires storage space and other overhead costs (insurance, security, real estate).
- There are other less tangible costs of inventory that result from increased wait times (because of Little's Law, to be discussed in **Chapter 9**) and lower quality (to be discussed in **Chapter 7**).

Given an annual cost of inventory (e.g., 20 percent per year) and the inventory turn information as computed above, we can compute the per-unit inventory cost that a process (or a supply chain) incurs. To do this, we take the annual holding cost and divide it by the number of times the inventory turns in a year:

$$\text{Per-unit inventory costs} = \frac{\text{Annual inventory cost}}{\text{Annual inventory turns}}$$

For example, a company that works based on a 20 percent annual inventory cost and that turns its inventory 6 times per year incurs per-unit inventory costs of

$$\frac{20\% \text{ per year}}{6 \text{ turns per year}} = 3.33\%$$

In the case of Kohl's (we earlier computed that the inventory turns 3.15 times per year), and assuming annual holding costs of 20 percent per year, this translates to inventory costs of about 6.35 percent of the cost of goods sold (20%/3.15 = 6.35). The calculations to obtain per-unit inventory costs are summarized in **Exhibit 2.1**.

Exhibit 2.1

CALCULATING INVENTORY TURNS AND PER-UNIT INVENTORY COSTS

1. Look up the value of inventory from the balance sheet.
2. Look up the cost of goods sold (COGS) from the earnings statement; do *not* use sales!
3. Compute inventory turns as

$$\text{Inventory turns} = \frac{\text{COGS}}{\text{Inventory}}$$

4. Compute per-unit inventory costs as

$$\text{Per-unit inventory costs} = \frac{\text{Annual inventory costs}}{\text{Inventor turns}}$$

Note: The annual inventory cost needs to account for the cost of financing the inventory, the cost of depreciation, and other inventory-related costs the firm considers relevant (e.g., storage, theft).

Take another look at Table 2.5. What is the relationship between gross margins and inventory turns? A casual reading of the data shown in the table suggests that higher inventory turns will lead to lower margins. After all, Kroger has the highest inventory turns and the lowest gross margins. In contrast, Tiffany turns its inventory really slowly and has high gross margins. With this in mind, shouldn't a retailer attempt to *reduce* its inventory turns?

Be careful to not jump to a conclusion too quickly. The table is based on different retail segments. Tiffany & Co turns its inventory 0.66 times per year. In other words, a random product in their store will be sold after 553 days (365 days per year/0.66 turns per year). Tiffany is in the business of finding necklaces or rings that are just perfect for a given customer. For that perfect match to happen, they need to hold a lot of inventory and that is bad for their inventory turns. Walmart, in contrast, is turning its inventory much more quickly. As we have seen above, items only spend an average of 39 days in Walmart's inventory. And Kroger, as a grocery retailer is turning its inventory even faster (inventory turns of 13.4 translate to 27.2 days of supply).

Instead of comparing "apples with oranges," we should ask ourselves how much more profitable a retailer would be if we would increase its inventory turns. Let's consider the case of BestBuy. BestBuy turns its inventory a little more than 6 times per year (6.35, to be exact). Thus, it keeps an average item for about 57 days.

Now, what would happen to the financial performance of BestBuy if it could cut this number by a week, that is, go from 57 days of supply to 50 days of supply? This acceleration in turning the inventory would lead to a new inventory turns the number of 7.3 (365/50).

Electronic retailers face high annual inventory costs, reflecting high storage and substantial obsolescence costs. Say the annual inventory costs for BestBuy are 30% and consider an item that costs BestBuy $400 to source (COGS). If BestBuy held the item for an entire year, it would thus incur inventory costs of $400 * 30% per year = $120 per year. Note that these are not out-of-pocket expenses that you could find on the company's financial statements. Rather, it is a combination of expenses related to mark-downs and storage as well as an opportunity cost of capital.

But, of course, BestBuy is not holding an item for an entire year. It turns its inventory 6.36 times per year. With 6.36 turns per year, BestBuy would incur a cost of inventory of 30%/6.36 * 400 = $18.87 per year for every time it sells our imaginary $400 item. Now, assume that we would reduce the days of supply by 1 week and thereby improve the inventory turns to 7.3. The costs associated with the same item would now go down to 30%/7.3 * $400=$16.44 for every unit of sale of this item. Is this cost saving significant? When you compare it to the slim net margins of BestBuy, it certainly is!

2.5 Five Reasons to Hold Inventory

LO 2-5

Identify five reasons for holding inventory.

While Little's Law allows us to compute the average inventory in the process (as long as we know flow time and flow rate), it offers no help in answering the question we raised previously: Why is there inventory in the process in the first place? To understand the need for inventory, we can no longer afford to take the black-box perspective and look at processes from the outside. Instead, we have to look at the process in much more detail.

As we saw from **Figure 2.7**, inventory reflected a deviation between the inflow into a process and its outflow. Ideally, from an operations perspective, we would like **Figure 2.7** to take the shape of two identical, straight lines, representing process inflow and outflow. Unfortunately, such straight lines with zero distance between them rarely exist in the real world. De Groote (1994) discusses five reasons for holding inventory, that is, for having the inflow line differ from the outflow line: (1) the time a flow unit spends in the process, (2) seasonal demand, (3) economies of scale, (4) separation of steps in a process, and (5) stochastic demand.

We will explain each of these five reasons with an example further below. Depending on the reason for holding inventory, inventories are given different names: pipeline inventory, seasonal inventory, cycle inventory, decoupling inventory/ buffers, and safety inventory. It should be noted that these five reasons are not necessarily mutually exclusive and that, in practice, there typically exist more than one reason for holding inventory.

Pipeline Inventory

This first reason for inventory reflects the time a flow unit has to spend in the process in order to be transformed from input to output. Even with unlimited resources, patients still need to spend time in the interventional radiology unit; their flow time would be the length of the critical path. We refer to this basic inventory on which the process operates as ***pipelin e inventory***.

For the sake of simplicity, let's assume that every patient would have to spend exactly 1.5 hours in the interventional radiology unit, as opposed to waiting for a resource to become available, and that we have 1 patient arrive every hour. How do we find the pipeline inventory in this case?

The answer is obtained through an application of Little's Law. Because we know two of the three performance measures, flow time and flow rate, we can figure out the third, in this case inventory: with a flow rate of 1 patient per hour and a flow time of 1.5 hours, the average inventory is

$$\text{Inventory} = 1[\text{patient/hour}] \times 1.5[\text{hours}] = 1.5 \text{ patients}$$

which is the number of patients undergoing some value-adding activity. This is illustrated in **Figure 2.8**.

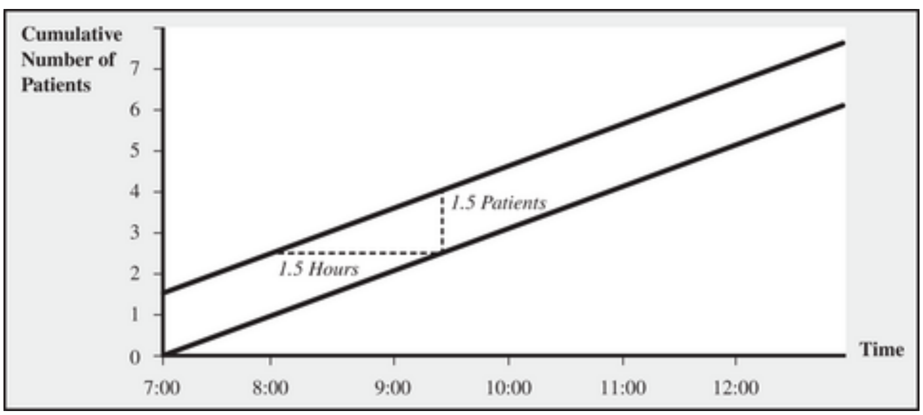


FIGURE 2.8 Pipeline Inventory

In certain environments, you might hear managers make statements of the type "we need to achieve zero inventory in our process." If we substitute Inventory = 0 into Little's Law, the immediate result is that a process with zero inventory is also a process with zero flow rate (unless we have zero flow time, which means that the process does not do anything to the flow unit). Thus, as long as it takes an operation even a minimum amount of time to work on a flow unit, the process will always exhibit pipeline inventory. There can be no hospital without patients and no factory can operate without some work in process!

Little's Law also points us toward the best way to reduce pipeline inventory. As reducing flow rate (and with it demand and profit) is typically not a desirable option, the *only* other way to reduce pipeline inventory is by reducing flow time.

Seasonal Inventory

Seasonal inventory occurs when capacity is rigid and demand is variable. Two examples illustrate this second reason for inventory. Campbell's Soup sells more chicken noodle soup in January than in any other month of the year—not primarily because of cold weather, but because Campbell's discounts chicken noodle soup in January. June is the next biggest sales month, because Campbell's increases its price in July.

So much soup is sold in January that Campbell's starts production several months in advance and builds inventory in anticipation of January sales. Campbell's could wait longer to start production and thereby not build as much inventory, but it would be too costly to assemble the needed capacity (equipment and labor) in the winter only to dismantle that capacity at the end of January when it is no longer needed.

In other words, as long as it is costly to add and subtract capacity, firms will desire to smooth production relative to sales, thereby creating the need for seasonal inventory.

An extreme case of seasonal inventory can be found in the agricultural and food processing sector. Due to the nature of the harvesting season, the sugar farmers in the U.S. Midwest collect all raw materials for their sugar production over a period of 6 weeks. At the end of the harvesting season, a large producer has accumulated—in the very meaning of the word—a pile of sugar beets, about 1 million tons, taking the form of a 67-acre sugar beets pile.

Given that food processing is a very capital-intense operation, the process is sized such that the 1.325 million tons of beets received and the almost 1 million tons of inventory that is built allow for a nonstop operation of the production plant until the beginning of the next harvesting season. Thus, as illustrated by **Figure 2.9**, the production, and hence the product outflow, is close to constant, while the product inflow is zero except for the harvesting season.

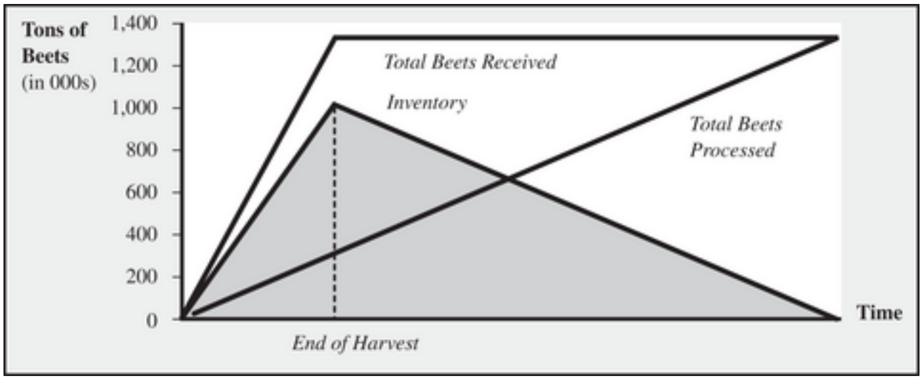


FIGURE 2.9 **Seasonal Inventory—Sugar**

Cycle Inventory

Throughout this book, we will encounter many situations in which it is economical to process several flow units collectively at a given moment in time to take advantage of scale economies in operations.

The scale economics in transportation processes provide a good example for the third reason for inventory. Whether a truck is dispatched empty or full, the driver is paid a fixed amount and a sizeable portion of the wear and tear on the truck depends on the mileage driven, not on the load carried. In other words, each truck shipment incurs a fixed cost that is independent of the amount shipped. To mitigate the sting of that fixed cost, it is tempting to load the truck completely, thereby dividing the fixed cost across the largest number of units.

In many cases, this indeed may be a wise decision. But a truck often carries more product than can be immediately sold. Hence, it takes some time to sell off the entire truck delivery. During that interval of time, there will be inventory. This inventory is labeled ***cycle inventory*** as it reflects that the transportation process follows a certain shipment cycle (e.g., a shipment every week).

Figure 2.10 plots the inventory level of a simple tray that is required during the operation in the interventional radiology unit. As we can see, there exists a "lumpy" inflow of units, while the outflow is relatively smooth. The reason for this is that—due to the administrative efforts related to placing orders for the trays—the hospital places only one order per week.

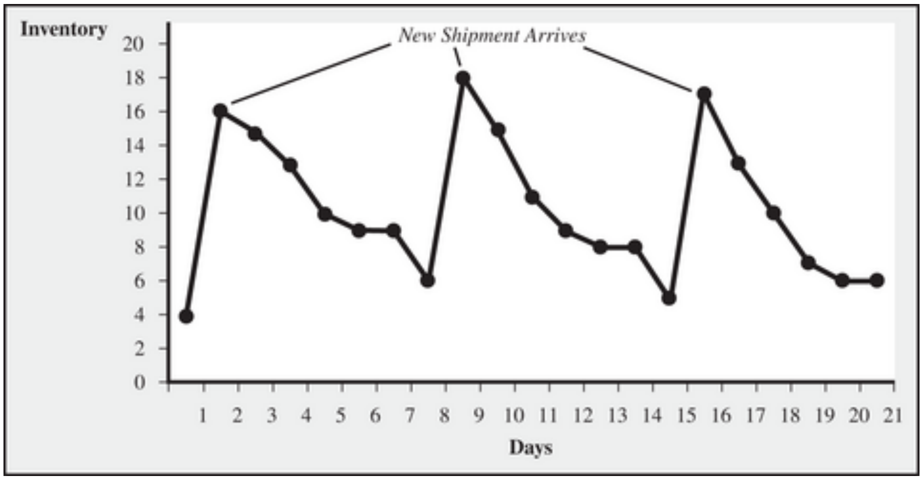


FIGURE 2.10 **Cycle Inventory**

The major difference between cycle inventory and seasonal inventory is that seasonal inventory is due to temporary imbalances in supply and demand due to variable demand (soup) or variable supply (beets) while cycle inventory is created due to a cost motivation.

Decoupling Inventory/Buffers

Inventory between process steps can serve as buffers. An inventory buffer allows management to operate steps independently from each other. For example, consider two workers in a garment factory. Suppose the first worker sews the collar onto a shirt and the second sews the buttons. A buffer between them is a pile of shirts with collars but no buttons. Because of that buffer, the first worker can stop working (e.g., to take a break, repair the sewing machine, or change thread color) while the second worker keeps working. In other words, buffers can absorb variations in flow rates by acting as a source of supply for a downstream process step, even if the previous operation itself might not be able to create this supply at the given moment in time.

An automotive assembly line is another example of a production process that uses buffers to decouple the various stations involved with producing the vehicle. In the absence of such buffers, a disruption at any one station would lead to a disruption of all the other stations, upstream and downstream. Think of a bucket brigade to fight a fire: there are no buffers between firefighters in a bucket brigade, so nobody can take a break without stopping the entire process.

Safety Inventory

The final reason for inventory is probably the most obvious, but also the most challenging: stochastic demand. Stochastic demand refers to the fact that we face uncertainty in the demand. Using terminology from Statistics, we can think of demand as a random variable to be realized from an underlying distribution. Like all distributions in Statistics, there will be a mean demand, which captures the effect of demand seasonality discussed in conjunction with Campbell's Soup above, and a standard deviation of demand.

So, the fact that we see a spike of Campbell’s chicken noodle soup in January is a matter of seasonality. For January 2022, Campbell knows that it will likely sell more than in other months. But, how much more? How much variation will there be relative to the mean of the demand distribution (which likely serves as Campbell's demand forecast; see Chapter 13 on the topic of forecasting)? That is a matter of stochastic demand.

Stochastic demand is an especially significant problem in retailing environments or at the finished goods level of manufacturers. Take a book retailer that must decide how many books to order of a given title. The book retailer has a forecast for demand, but forecasts are (at best) correct on average. Order too many books and the retailer is faced with leftover inventory. Order too few and valuable sales are lost. This trade-off can be managed, as we will discover in Chapter 14, but not eliminated (unless there are zero forecast errors).

The resulting inventory thereby can be seen as a way to hedge against the underlying demand uncertainty. It might reflect a one-shot decision, for example, in the case of a book retailer selling short-life-cycle products such as newspapers or magazines. If we consider a title with a longer product life cycle (e.g., children’s books), the book retailer will be able to replenish books more or less continuously over time.

Figure 2.11 shows the example of the blood bank in Presbyterian Hospital in Philadelphia. While the detailed inflow and consumption of blood units vary over the course of the month, the hospital always has a couple of days of blood in inventory. Given that blood perishes quickly, the hospital wants to keep only a small inventory at its facility, which it replenishes from the regional blood bank operated by the Red Cross.

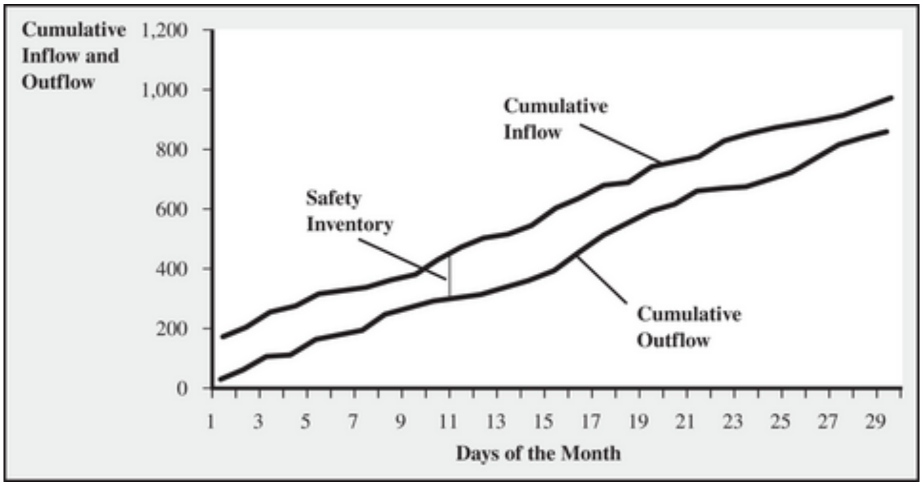


FIGURE 2.11 Safety Inventory at a Blood Bank

2.6 The Product–Process Matrix

LO 2-6

Explain how the product-process matrix works.

Processes leading to the supply of goods or services can take many different forms. Some processes are highly automated, while others are largely manual. Some processes resemble the legendary Ford assembly line, while others resemble more the workshop in your local bike store. Empirical research in operations management, which has looked at thousands of processes, has identified five "clusters" or types of processes. Within each of the five clusters, processes are very similar concerning variables such as the number of different product variants they offer or the production volume they provide. **Table 2.6** describes these different types of processes.

TABLE 2.6 Process Types and Their Characteristics

	Examples	Number of Different Product Variants	Product Volume (Units/Year)
Job shop	• Design company • Commercial printer • Formula 1 race car	High (100+)	Low (1–100)
Batch process	• Apparel sewing • Bakery • Semiconductor wafers	Medium (10–100)	Medium (100–100k)
Worker-paced line flow	• Auto assembly • Computer assembly	Medium (1–50)	High (10k–1M)
Machine-paced line flow	• Large auto assembly	Low (1–10)	High (10k–1M)
Continuous process	• Paper mill	Low (1–10)	Very high

	Examples	Number of Different Product Variants	Product Volume (Units/Year)
	• Oil refinery • Food processing		

By looking at the evolution of a number of industries, Hayes and Wheelwright (1979) observed an interesting pattern, which they referred to as the product–process matrix (see **Figure 2.12**). The product–process matrix stipulates that over its life cycle, a product typically is initially produced in a job shop process. As the production volume of the product increases, the production process for the product moves from the upper left of the matrix to the lower right.

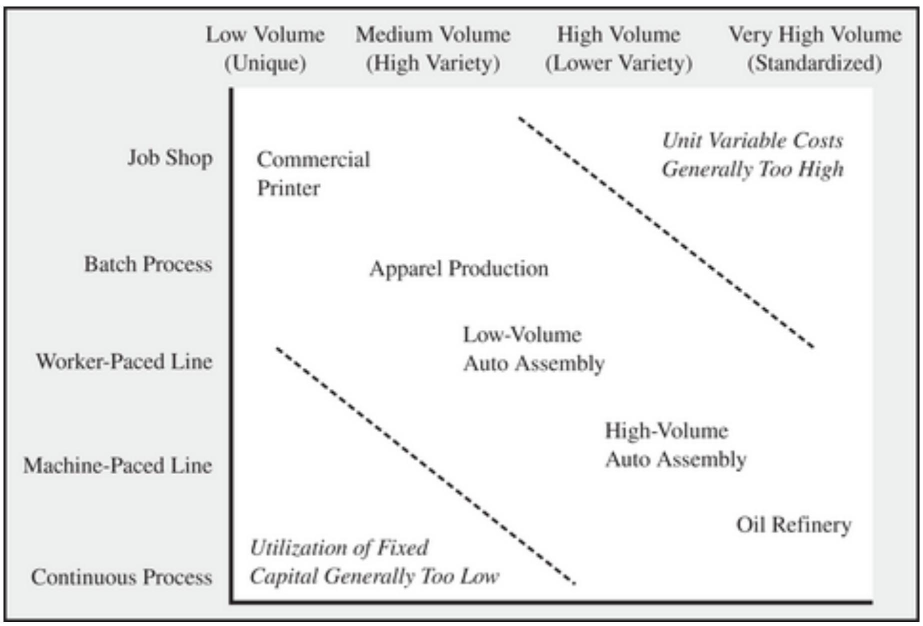


FIGURE 2.12 Product–Process Matrix

Source: Hayes and Wheelwright (1979).

For example, the first automobiles were produced using job shops, typically creating one product at a time. Most automobiles were unique; not only did they have different colors or add-ons, but they differed in size, geometry of the body, and many other aspects. Henry Ford’s introduction of the assembly line corresponded to a major shift along the diagonal of the product–process matrix. Rather than producing a couple of products in a job shop, Ford produced thousands of vehicles on an assembly line.

Note that the "off-diagonals" in the product–process matrix (the lower left and the upper right) are empty. This reflects that neither it is economical to produce very high volumes in a job shop (imagine if all of the millions of new vehicles sold in the United States every year were handcrafted in the same manner as Gottlieb Daimler created the first automobile) nor does it make sense to use an assembly line in order to produce only a handful of products a year.

We have to admit that few companies—if any—would be foolish enough to produce a high-volume product in a job shop. However, identifying a process type and looking at the product–process matrix is more than an academic exercise in industrial history. The usefulness of the product–process matrix lies in two different points:

1. Similar process types tend to have similar problems. For example, as we will discuss in **Chapter 4**, assembly lines tend to have the problem of line balancing (some workers work harder than others). Batch-flow processes tend to be slow in responding to customer demand (see **Chapter 5**). Thus, once you know a process type, you can quickly determine what type of problems the process is likely to face and what solution methods are most appropriate.
2. The "natural drift" of industries toward the lower right of **Figure 2.12** enables you to predict how processes are likely to evolve in a particular industry. Consider, for example, the case of eye surgery. Up until the 1980s, corrective eye surgery was done in large hospitals. There, doctors would perform a large variety of very different eye-related cases. Fifteen years later, this situation had changed dramatically. Many highly specialized eye clinics have opened, most of them focusing on a limited set of procedures. These clinics achieve high volume and, because of the high volume and the lower variety of cases, can operate at much higher levels of efficiency.

2.7 Summary

In this chapter, we emphasized the importance of looking at the operations of a firm not just in terms of the products that the firm supplies, but also at the processes that generate the supply. Looking at processes is especially important with respect to demand–supply mismatches. From the perspective of the product, such mismatches take the form of waiting times; from the perspective of the process, they take the form of inventory.

For any process, we can define three fundamental performance measures: inventory, flow time, and flow rate. The three measures are related by Little's Law, which states that the average inventory is equal to the average flow time multiplied by the average flow rate.

Little's Law can be used to find any of the three performance measures, as long as the other two measures are known. This is specifically important with respect to flow time, which is in practice frequently difficult to observe directly.

A measure related to flow time is inventory turns. Inventory turns, measured by 1/(flow time), capture how fast the flow units are transformed from input to output. It is an important benchmark in many industries, especially retailing. Inventory turns are also the basis of computing the inventory costs associated with 1 unit of supply.

2.8 Practice Problems and Selected Solutions

The following questions will help in testing your understanding of this chapter. After each question, we show the relevant section in parentheses [Section x].

Solutions to problems marked with an "*" appear at the end of this section. Video solutions to select problems are available in Connect.

Q2.1* **(Dell)** What percentage of cost of a Dell computer reflects inventory costs? Assume Dell's yearly inventory cost is 40 percent to account for the cost of capital for financing the inventory, the warehouse space, and the cost of obsolescence. In other words, Dell incurs a cost of $40 for a $100 component that is in the company's inventory for 1 entire year. In 2001, Dell's 10-k reports showed that the company had $400 million in inventory and COGS of $26,442 million. [**2.4**]

Q2.2 **(Airline)** Consider the baggage check-in of a small airline. Check-in data indicate that from 9 a.m. to 10 a.m., 255 passengers checked in. Moreover, based on counting the number of passengers waiting in line, airport management found that the average number of passengers waiting for check-in was 35. How long did the average passenger have to wait in line? [**2.3**]

Q2.3 **(Inventory Cost)** A manufacturing company producing medical devices reported $60,000,000 in sales over the last year. At the end of the same year, the company had $20,000,000 worth of inventory of ready-to-ship devices.

a. Assuming that units in inventory are valued (based on COGS) at $1,000 per unit and are sold for $2,000 per unit, how fast does the company turn its inventory? The company uses a 25 percent per year cost of inventory. That is, for the hypothetical case that 1 unit of $1,000 would sit exactly 1 year in inventory, the company charges its operations division a $250 inventory cost. [**2.4**]

b. What—in absolute terms—is the per-unit inventory cost for a product that costs $1,000? [**2.4**]

Q2.4 **(Apparel Retailing)** A large catalog retailer of fashion apparel reported $100,000,000 in revenues over the last year. On average, over the same year, the company had $5,000,000 worth of inventory in its warehouses. Assume that units in inventory are valued based on cost of goods sold (COGS) and that the retailer has a 100 percent markup on all products.

a. How many times each year does the retailer turn its inventory? [2.4]

b. The company uses a 40 percent per year cost of inventory. That is, for the hypothetical case that 1 item of $100 COGS would sit exactly 1 year in inventory, the company charges itself a $40 inventory cost. What is the inventory cost for a $30 (COGS) item? You may assume that inventory turns are independent of the price. [2.4]

Q2.5 **(LaVilla)** LaVilla is a village in the Italian Alps. Given its enormous popularity among Swiss, German, Austrian, and Italian skiers, all of its beds are always booked in the winter season and there are, on average, 1,200 skiers in the village. On average, skiers stay in LaVilla for 10 days.

a. How many new skiers are arriving—on average—in LaVilla every day? [2.3]

b. A study done by the largest hotel in the village has shown that skiers spend on average $50 per person on the first day and $30 per person on each additional day in local restaurants. The study also forecasts that—due to increased hotel prices—the average length of stay for the 2003/2004 season will be reduced to 5 days. What will be the percentage change in revenues of local restaurants compared to last year (when skiers still stayed for 10 days)? Assume that hotels continue to be fully booked! [2.3]

Q2.6 **(Highway)** While driving home for the holidays, you can't seem to get Little's Law out of your mind. You note that your average speed of travel is about 60 miles per hour. Moreover, the traffic report from the WXPN traffic chopper states that there is an average of 24 cars going in your direction on a 1-quarter mile part of the highway. What is the flow rate of the highway (going in your direction) in cars per hour? [2.3]

Q2.7 **(Industrial Baking Process)** Strohrmann, a large-scale bakery in Pennsylvania, is laying out a new production process for their packaged bread, which they sell to several grocery chains. It takes 12 minutes to bake the bread. How large an oven is required so that the company is able to produce 4,000 units of bread per hour (measured in the number of units that can be baked simultaneously)? [2.3]

Page 30

Q2.8 **(Mt. Kinley Consulting)** Mt. Kinley is a strategy consulting firm that divides its consultants into three classes: associates, managers, and partners. The firm has been stable in size for the last 20 years, ignoring growth opportunities in the 90s, but also not suffering from a need to downsize in the recession at the beginning of the 21st century. Specifically, there have been—and are expected to be—200 associates, 60 managers, and 20 partners.

The work environment at Mt. Kinley is rather competitive. After 4 years of working as an associate, a consultant goes “either up or out,” that is, becomes a manager or is dismissed from the company. Similarly, after 6 years, a manager either becomes a partner or is dismissed. The company recruits MBAs as associate consultants; no hires are made at the manager or partner level. A partner stays with the company for another 10 years (a total of 20 years with the company).

a. How many new MBA graduates does Mt. Kinley have to hire every year? [2.3]

b. What are the odds that a new hire at Mt. Kinley will become partner (as opposed to being dismissed after 4 years or 10 years)? [2.3]

Q2.9 **(Major U.S. Retailers)** The following table shows financial data for Costco Wholesale and Walmart, two major U.S. retailers.

	Costco	Walmart
	($ Millions)	($ Millions)
Inventories	$ 3,643	$ 29,447
Sales (net)	$ 48,106	$ 286,103
COGS	$ 41,651	$ 215,493

Source: Compustat, WRDS.

Assume that both companies have an average annual holding cost rate of 30 percent (i.e., it costs both retailers $3 to hold an item that they procured for $10 for 1 entire year).

a. How many days, on average, does a product stay in Costco’s inventory before it is sold? Assume that stores are operated 365 days a year. [2.4]

b. How much lower is, on average, the inventory cost for Costco compared to Walmart of a household cleaner valued at $5 COGS? Assume that the unit cost of the household cleaner is the same for both companies and that the price and the inventory turns of an item are independent. [2.4]

Q2.10 **(McDonald’s)** The following figures are taken from the financial statements of McDonald’s and Wendy’s.[1] Figures are in million dollars.

	McDonald’s	Wendy’s
Inventory	$ 129.4	$ 54.4
Revenue	17,140.5	3,148.9

	McDonald's	Wendy's
Cost of goods sold	11,943.7	1,634.6
Gross profit	5,196.8	1,514.4

a. What were McDonald's inventory turns in this year? What were Wendy's inventory turns? [**2.4**]

b. Suppose it costs both McDonald's and Wendy's $3 (COGS) per their value meal offerings, each sold at the same price of $4. Assume that the cost of inventory for both companies is 30 percent a year. Approximately, how much does McDonald's save in inventory cost *per value meal* compared to that of Wendy's? You may assume the inventory turns are independent of the price. [**2.4**]

Q2.11 **(BCH)** BCH, a large consulting firm in the United Kingdom, has a consulting staff consisting of 400 consultants at the rank of "associate." On average, a consultant remains at the associate level for 2 years. After this time, 30 percent of the consultants are promoted to the rank of "engagement manager," and the other 70 percent have to leave the company.

In order to maintain the consulting staff at an average level of 400 associates, how many new consultants does BCH have to hire each year at the associate level? [**2.3**]

Q2.12 **(Kroger)** The following provides financial information for Kroger (in million $):

	Kroger
Inventory	$ 6,244
Revenue	$ 95,751
Cost of goods sold	$76,858

a. What were Kroger's inventory turns in this year? [**2.4**]

If you would like to test your understanding of a specific section, here are the questions organized by section:

Section 2.3: Q2.2, Q2.5, Q2,6, Q2.7, Q2.8, and Q2.11

Section 2.4: Q2.1, Q2.3, Q2.4, Q2.9, Q2.10, and Q2.12

Selected Solutions

Q2.1 (Dell)

The following steps refer directly to **Exhibit 2.1**.

Step 1. For 2001, we find in Dell's 10-k: Inventory = $400 (in millions)

Step 2. For 2001, we find in Dell's 10-k: COGS = $26,442 (in millions)

Step 3. $\text{Inventory turns } = \dfrac{\$26,442/\text{Year}}{\$400} = 66.105 \text{ turns per year}$

Step 4. $\text{Per-unit inventory cost} = \dfrac{40\% \text{ per year}}{66.105 \text{ per year}} = 0.605 \text{ percent per unit}$

CHAPTER 3

Understanding the Supply Process: Evaluating Process Capacity

In the attempt to match supply with demand, an important measure is the maximum amount that a process can produce in a given unit of time, a measure referred to as the ***process capacity***. To determine the process capacity of an operation, we need to analyze the operation in much greater detail compared to the previous chapter. Specifically, we need to understand the various activities involved in the operation and how these activities contribute toward fulfilling the overall demand.

In this chapter, you will learn how to perform a process analysis. Unlike **Chapter 2**, where we felt it was sufficient to treat the details of the operation as a black box and merely focus on the performance measures inventory, flow time, and flow rate, we now will focus on the underlying process in great detail.

Despite this increase in detail, this chapter (and this book) is not taking the perspective of an engineer. In fact, in this chapter, you will learn how to take a fairly technical and complex operation and simplify it to a level suitable for managerial analysis. This includes preparing a process flow diagram, finding the capacity and the bottleneck of the process, computing the utilization of various process steps, and computing a couple of other performance measures.

We will illustrate this new material with the Circored plant, a joint venture between the German engineering company Lurgi AG and the U.S. iron ore producer Cleveland Cliffs. The Circored plant converts iron ore (in the form of iron ore fines) into direct reduced iron (DRI) briquettes. Iron ore fines are shipped to the plant from mines in South America; the briquettes the process produces are shipped to various steel mills in the United States.

The example of the Circored process is particularly useful for our purposes in this chapter. The underlying process is complex and in many ways a masterpiece of process engineering (see **Terwiesch and Loch [2002]** for further details). At first sight, the process is so complex that it seems impossible to understand the underlying process behavior without a detailed background in engineering and metallurgy. This challenging setting allows us to demonstrate how process analysis can be used to “tame the beast” and create a managerially useful view of the process, avoiding any unnecessary technical details.

3.1 How to Draw a Process Flow Diagram

LO 3-1

Identify the information necessary to create a process flow diagram.

The best way to begin any analysis of an operation is by drawing a ***process flow diagram***. A process flow diagram is a graphical way to describe the process and it will help us to structure the information that we collect during the case analysis or process improvement project. Before we turn to the question of how to draw a process flow diagram, first consider alternative approaches to how we could capture the relevant information about a process.

Looking at the plant from above (literally), we get a picture as is depicted in **Figure 3.1**. At the aggregate level, the plant consists of a large inventory of iron ore (input), the plant itself (the resource), and a large inventory of finished briquettes (output). In many ways, this corresponds to the black box approach to operations taken by economists and many other managerial disciplines.

FIGURE 3.1 Photo of the Circored Plant

Christian Terwiesch

In an attempt to understand the details of the underlying process, we could turn to the engineering specifications of the plant. Engineers are interested in a detailed description of the various steps involved in the overall process and how these steps are functioning. Such descriptions, typically referred to as specifications, were used in the actual construction of the plant. **Figure 3.2** provides one of the numerous specification drawings for the Circored process.

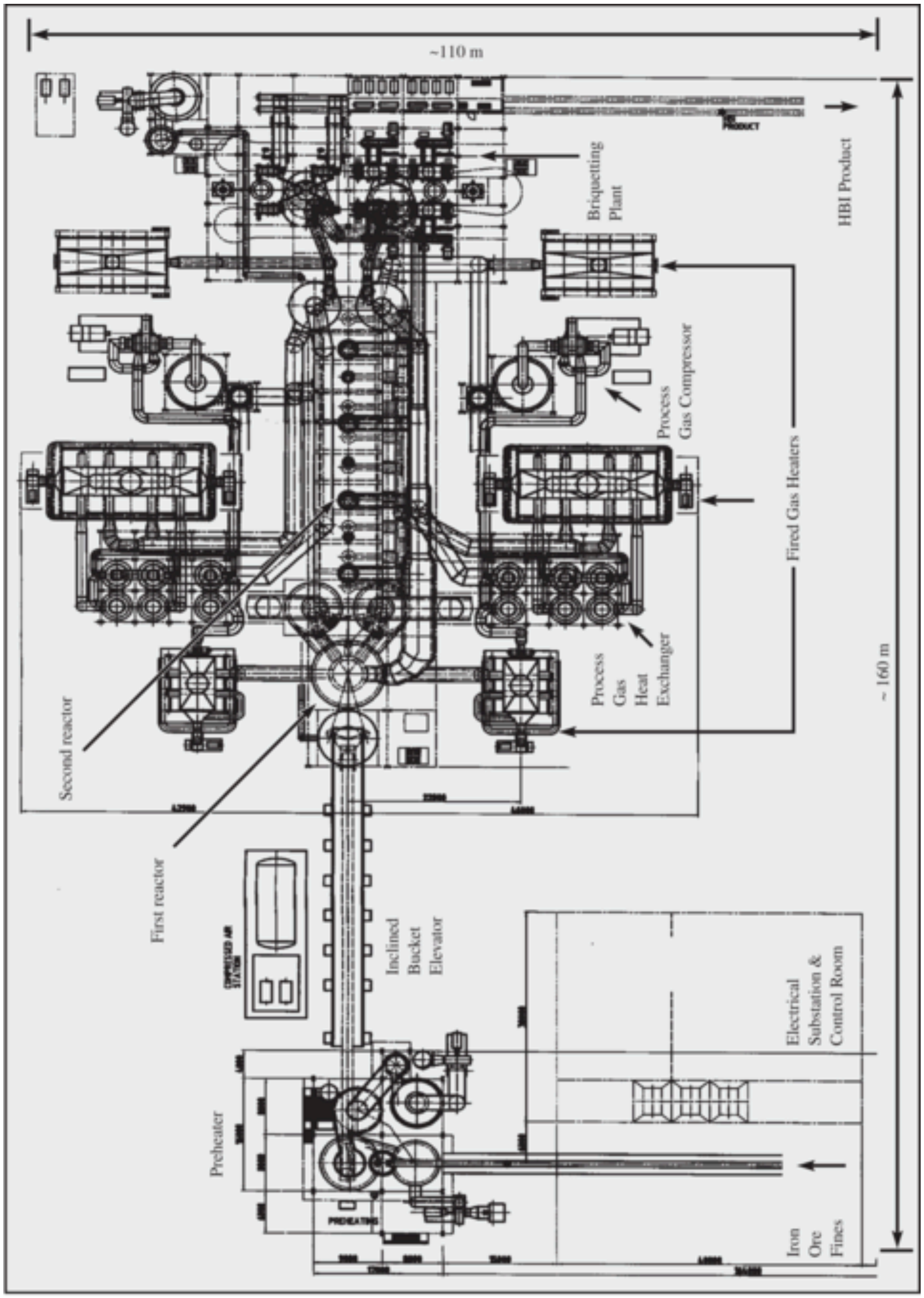


FIGURE 3.2 Engineering Drawing

Christian Terwiesch

Unfortunately, this attempt to increase our understanding of the Circored process is also only marginally successful. Like the photograph, this view of the process is also a rather static one: it emphasizes the equipment, yet provides us with little understanding of how the iron ore moves through the process. In many ways, this view of a process is similar to taking the architectural drawings of a hospital and hoping that this would lead to insights about what happens to the patients in this hospital.

In a third—and final—attempt to get our hands around this complex process, we change our perspective from the one of the visitor to the plant (photo in **Figure 3.1**) or the engineers who built the plant (drawing in **Figure 3.2**) to the perspective of the iron ore itself and how it flows through the process. Thus, we define a unit of iron ore—a ton, a pound, or a molecule—as our flow unit and "attach" ourselves to this flow unit as it makes its journey through the process. This is similar to taking the perspective of the patient in a hospital, as opposed to taking the perspective of the hospital resources. For concreteness, we will define our flow unit to be a ton of iron ore.

To draw a process flow diagram, we first need to focus on a part of the process that we want to analyze in greater detail; that is, we need to define the *process boundaries* and an appropriate level of detail. The placement of the process boundaries will depend on the project we are working on. For example, in the operation of a hospital, one project concerned with patient waiting time might look at what happens to the patient waiting for a lab test (e.g., check-in, waiting time, encounter with the nurse). In this project, the encounter with the doctor who requested the lab test would be outside the boundaries of the analysis. Another project related to the quality of surgery, however, might look at the encounter with the doctor in great detail, while either ignoring the lab or treating it with less detail.

A process operates on flow units, which are the entities flowing through the process (e.g., patients in a hospital, cars in an auto plant, insurance claims at an insurance company). A process flow diagram is a collection of boxes, triangles, and arrows (see **Figure 3.3**). Boxes stand for process activities, where the operation adds value to the flow unit. Depending on the level of detail we choose, a process step (a box) can itself be a process.

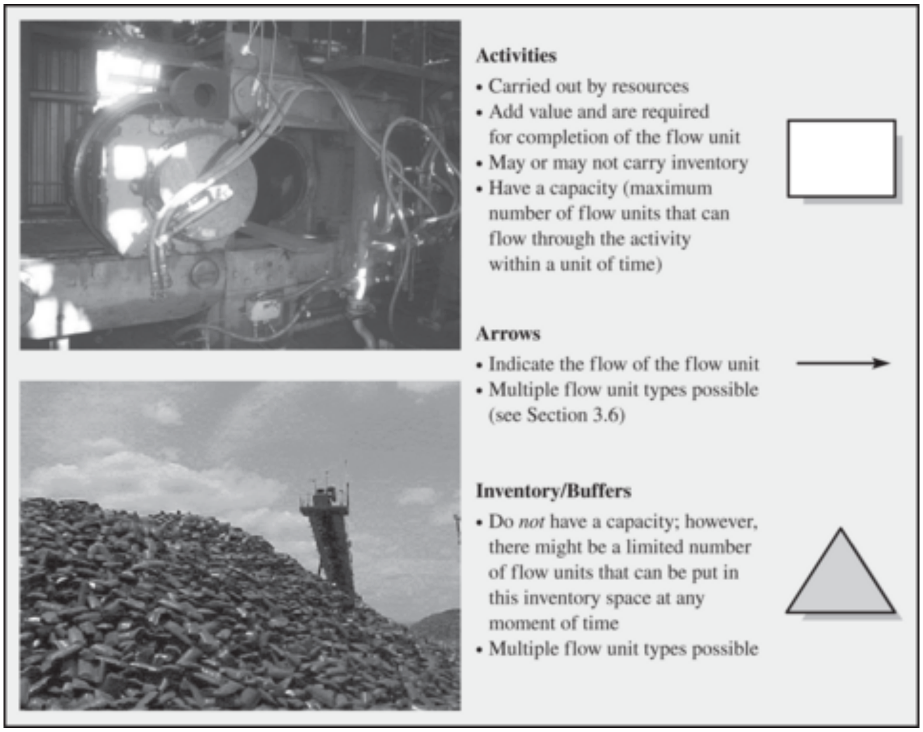


FIGURE 3.3 Elements of a Process

Christian Terwiesch

Triangles represent waiting areas or ***buffers*** holding inventory. In contrast to a process step, inventories do not add value; thus, a flow unit does not have to spend time in them. However, as discussed in the previous chapter, there are numerous reasons why the flow unit might spend time in inventory even if it will not be augmented to a higher value there.

The arrows between boxes and triangles represent the route the flow unit takes through the process. If there are different flow units that take different routes through the process, it can be helpful to use different colors for the different routes. An example of this is given at the end of this chapter.

Page 36

In the Circored plant, the first step the flow unit encounters in the process is the *preheater*, where the iron ore fines (which have a texture like large-grained sand) are dried and heated. The heating is achieved through an inflow of high-pressured air, which is blown into the preheater from the bottom. The high-speed air flow “fluidizes” the ore, meaning that the mixed air–ore mass (a “sandstorm”) circulates through the system as if it was a fluid, while being heated to a temperature of approximately 850–900°C.

However, from a managerial perspective, we are not really concerned with the temperature in the preheater or the chemical reactions happening therein. For us, the preheater is a resource that receives iron ore from the initial inventory and processes it. In an attempt to take record of what the flow unit has experienced up to this point, we create a diagram similar to **Figure 3.4**.

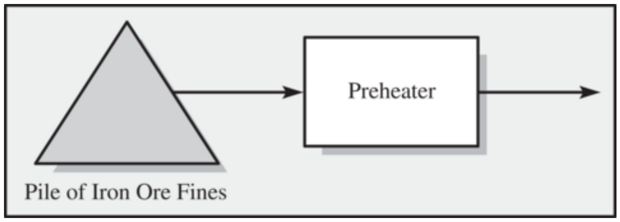


FIGURE 3.4 Process Flow Diagram, First Step

From the preheater, a large bucket elevator transports the ore to the second process step, the *lock hoppers.* The lock hoppers consist of three large containers, separated by sets of double isolation valves. Their role is to allow the ore to transition from an oxygen-rich environment to a hydrogen atmosphere.

Following the lock hoppers, the ore enters the *circulating fluid bed reactor*, or *first reactor,* where the actual reduction process begins. The reduction process requires the ore to be in the reactor for 15 minutes, and the reactor can hold up to 28 tons of ore.

After this first reduction, the material flows into the *stationary fluid bed reactor*, or *second reactor.* This second reaction takes about four hours. The reactor is the size of a medium two-family home and contains 400 tons of the hot iron ore at any given moment in time. In the meantime, our diagram from **Figure 3.4**. has extended to something similar to **Figure 3.5**.

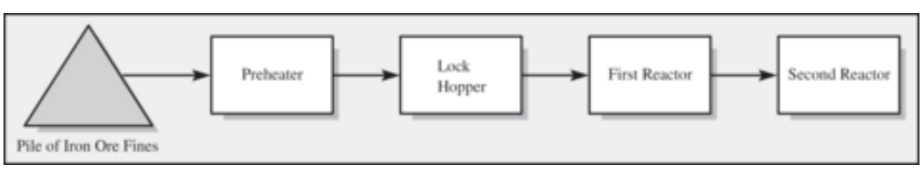


FIGURE 3.5 Process Flow Diagram (to Be Continued)

A couple of things are worth noting at this point:

- When creating **Figure 3.5**, we decided to omit the bucket elevator. There is no clear rule on when it is appropriate to omit a small step and when a step would have to be included in the process flow diagram. A reasonably good rule of thumb is to only include those process steps that are likely to affect the process flow or the economics of the process. The bucket elevator is cheap, the flow units spend little time on it, and this transportation step never becomes a constraint for the process. So it is not included in our process flow diagram.

- The reaction steps are boxes, not triangles, although there is a substantial amount of ore in them, that is, they do hold inventory. The reduction steps are necessary, value-adding steps. No flow unit could ever leave the system without spending time in the reactors. This is why we have chosen boxes over triangles here.

Following the second reactor, the reduced iron enters the *flash heater*, in which a stream of high-velocity hydrogen carries the DRI to the top of the plant while simultaneously reheating it to a temperature of 685°C.

After the flash heater, the DRI enters the *pressure let-down system (discharger).* As the material passes through the discharger, the hydrogen atmosphere is gradually replaced by inert nitrogen gas. Pressure and hydrogen are removed in a reversal of the lock hoppers at the beginning. Hydrogen gas sensors assure that material leaving this step is free of hydrogen gas and, hence, safe for briquetting.

Each of the three *briquetting* machines contains two wheels that turn against each other, each wheel having the negative of one-half of a briquette on its face. The DRI is poured onto the wheels from the top and is pressed into briquettes, or iron bars, which are then moved to a large pile of finished goods inventory.

This completes our journey of the flow unit through the plant. The resulting process flow diagram that captures what the flow unit has experienced in the process is summarized in **Figure 3.6**.

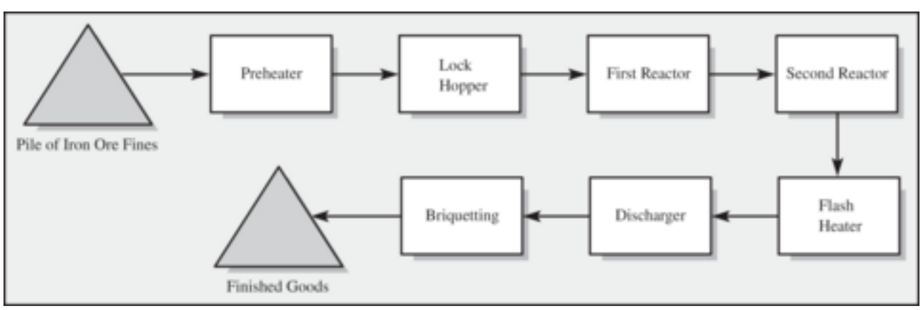


FIGURE 3.6 Completed Process Flow Diagram for the Circored Process

When drawing a process flow diagram, the sizes and the exact locations of the arrows, boxes, and triangles do not carry any special meaning. For example, in the context of **Figure 3.6**, we chose a "U-shaped" layout of the process flow diagram, as otherwise we would have had to publish this book in a larger format.

In the absence of any space constraints, the simplest way to draw a process flow diagram for a process such as Circored's is just as one long line. However, we should keep in mind that there are more complex processes; for example, a process with multiple flow units or a flow unit that visits one and the same resource multiple times. This will be discussed further at the end of the chapter.

Another alternative in drawing the process flow diagram is to stay much closer to the physical layout of the process. This way, the process flow diagram will look familiar for engineers and operators who typically work off the specification drawings (see **Figure 3.2**) and it might help you to find your way around when you are visiting the "real" process. Such an approach is illustrated in **Figure 3.7**.

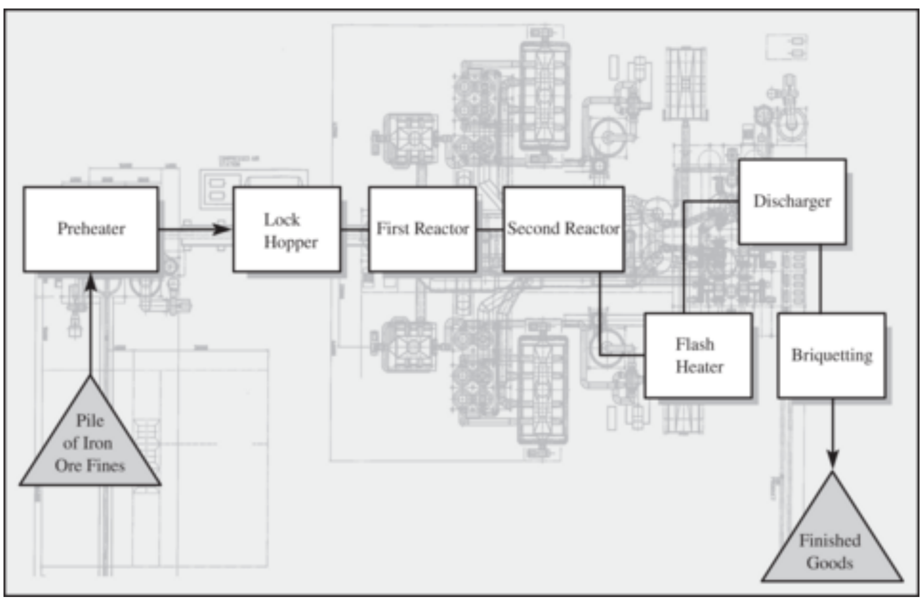


FIGURE 3.7 Completed Process Flow Diagram for the Circored Process

3.2 Process Capacity, Bottleneck, and Flow Rate (Throughput)

LO 3-2

Explain how to find the bottleneck in a process and how to determine process capacity and flow rate.

From a supply perspective, the most important question that arises is how much direct reduced iron the Circored process can supply in a given unit of time, say one day. This is the ***process capacity***, which we defined in the previous chapter as the maximum flow rate the process can supply.

Not only can capacity be measured at the level of the overall process, it also can be measured at the level of the individual resources that constitute the process. Just as we defined the process capacity, we define the **capacity** of a resource as the maximum amount the resource can produce in a given time unit.

Note that the process capacity measures how much the process *can* produce, opposed to how much the process actually *does* produce. For example, consider a day where—due to a breakdown or another external event—the process does not operate at all. Its capacity would be unaffected by this, yet the flow rate would reduce to zero. This is similar to your car, which might be able to drive at 130 miles per hour (capacity), but typically—or better, hopefully—only drives at 65 miles per hour (flow rate).

As the completion of a flow unit requires the flow unit to visit every one of the resources in the process, the overall process capacity is determined by the resource with the smallest capacity. We refer to that resource as the ***bottleneck***. It provides the weakest link in the overall process chain, and, as we know, a chain is only as strong as its weakest link.

With the concept of the bottleneck in mind, we can now write the process capacity as

$$\text{Process capacity} = \text{Minimum}\{\text{Capacity of resource } 1, \ldots, \text{Capacity of resource n}\}$$

where there are a total of n resources. How much the process actually does produce will depend not only on its capability to create supply (process capacity) but also on the demand for its output as well as the availability of its input. As with capacity, demand and the available input should be measured as rates, that is, as flow units per unit of time. For this process, our flow unit is one ton of ore, so we could define the available input and the demand in terms of tons of ore per hour.

The combination of available input, demand, and process capacity yields the rate at which our flow unit actually flows through the process, called the ***flow rate***:

$$\text{Flow rate} = \text{Minimum}\{\text{Available input, Demand, Process capacity}\}$$

If demand is lower than supply (i.e., there is sufficient input available and the process has enough capacity), the process would produce at the rate of demand, independent of the process capacity. We refer to this case as ***demand-constrained.*** Note that in this definition demand also includes any potential requests for the accumulation of inventory. For example, while the demand for Campbell's chicken noodle soup might be lower than process capacity for the month of November, the process would not be demand-constrained if management decided to accumulate finished goods inventory in preparation for the high sales in the month of January. Thus, demand in our analysis refers to everything that is demanded from the process at a given time.

If demand exceeds supply, the process is ***supply-constrained.*** Depending on what limits product supply, the process is either input-constrained or capacity-constrained.

Figure 3.8 summarizes the concepts of process capacity and flow rate, together with the notion of demand- versus supply-constrained processes. In the case of the supply-constrained operation, there is sufficient input; thus, the supply constraint reflects a capacity constraint.

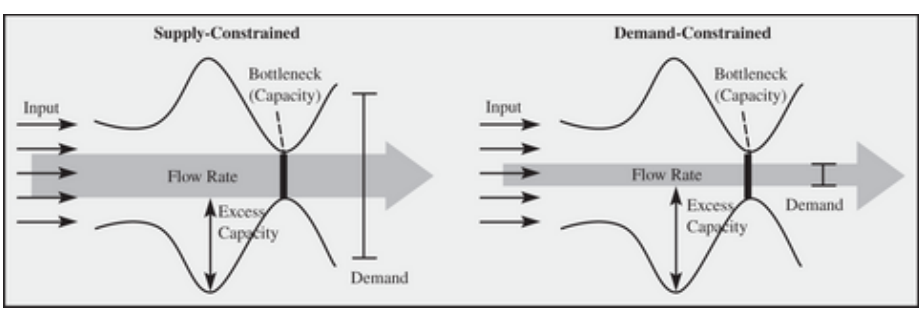


FIGURE 3.8 Supply-Constrained (left) and Demand-Constrained (right) Processes

To understand how to find the bottleneck in a process and thereby determine the process capacity, consider each of the Circored resources. Note that all numbers are referring to tons of process output. The actual, physical weight of the flow unit might change over the course of the process.

Finding the bottleneck in many ways resembles the job of a detective in a crime story; each activity is a "suspect," in the sense that it could potentially constrain the overall supply of the process:

- The preheater can process 120 tons per hour.
- The lock hoppers can process 110 tons per hour.
- The analysis of the reaction steps is somewhat more complicated. We first observe that at any given moment of time, there can be, at maximum, 28 tons in the first reactor. Given that the iron ore needs to spend 15 minutes in the reactor, we can use Little's Law (see **Chapter 2**) to see that the maximum amount of ore that can flow through the reactor—and spend 15 minutes in the reactor—is

$$28 \text{ tons} = \text{Flow rate} \times 0.25 \text{ hour} => \text{ Flow rate} = 112 \text{ tons/hour}$$

Page 40

Thus, the capacity of the first reactor is 112 tons per hour. Note that a shorter reaction time in this case would translate to a higher capacity.

- We can apply a similar logic for the second reactor, which can hold up to 400 tons:

$$400 \text{ tons} = \text{Flow rate} \times 4 \text{ hours} => \text{Flow rate} = 100 \text{ tons/hour}$$

Thus, the capacity (the maximum possible flow rate through the resource) of the second reactor is 100 tons per hour.

- The flash heater can process 135 tons per hour.
- The discharger has a capacity of 118 tons per hour.
- Each of the three briquetting machines has a capacity of 55 tons per hour. As the briquetting machines collectively form one resource, the capacity of the briquetting machines is simply 3 × 55 tons per hour = 165 tons per hour.

The capacity of each process step is summarized in **Table 3.1**.

TABLE 3.1 Capacity Calculation

Process Step	Calculations	Capacity
Preheater		120 tons per hour
Lock hoppers		110 tons per hour
First reactor	Little's Law: Flow rate = 28 tons/0.25 hour	112 tons per hour
Second reactor	Little's Law: Flow rate = 400 tons/4 hours	100 tons per hour
Flash heater		135 tons per hour
Discharger		118 tons per hour
Briquetting machine	Consists of three machines: 3 × 55 tons per hour	165 tons per hour
Total process	Based on bottleneck, which is the second reactor	**100 tons per hour**

Following the logic outlined above, we can now identify the second reactor as the bottleneck of the Circored process. The overall process capacity is computed as the minimum of the capacities of each resource (all units are in tons per hour):

$$\text{Process capacity} = \text{Minimum}\{120, 110, 112, 100, 135, 118, 165\} = 100$$

3.3 How Long Does It Take to Produce a Certain Amount of Supply?

LO 3-3

Explain how long it takes to produce a certain amount of supply.

There are many situations where we need to compute the amount of time required to create a certain amount of supply. For example, in the Circored case, we might ask, “How long does it take for the plant to produce 10,000 tons?” Once we have determined the flow rate of the process, this calculation is fairly straightforward. Let X be the amount of supply we want to fulfill. Then,

$$\text{Time to fulfill } X \text{ units} = \frac{\text{X}}{\text{Flow rate}}$$

To answer our question,

$$\text{Time to produce 10,000 tons} = \frac{\text{10,000 tons}}{\text{100 tons / hour}} = \text{100 hours}$$

Note that this calculation assumes the process is already producing output, that is, the first unit in our 10,000 tons flows out of the process immediately. If the process started empty, it would take the first flow unit time to flow through the process. **Chapter 4** provides the calculations for that case.

Note that in the previous equation we use flow rate, which in our case is capacity because the system is supply-constrained. However, if our system were demand-constrained, then the flow rate would equal the demand rate.

3.4 Process Utilization and Capacity Utilization

LO 3-4

Define utilization, how it is measured, and why it should be handled carefully.

Given the first-of-its-kind nature of the Circored process, the first year of its operation proved to be extremely difficult. In addition to various technical difficulties, demand for the product (reduced iron) was not as high as it could be, as the plant's customers (steel mills) had to be convinced that the output created by the Circored process would be of the high quality required by the steel mills.

While abstracting from details such as scheduled maintenance and inspection times, the plant was designed to achieve a process capacity of 876,000 tons per year (100 tons per hour × 24 hours/day × 365 days/year, see above), the demand for iron ore briquettes was only 657,000 tons per year. Thus, there existed a mismatch between demand and potential supply (process capacity).

A common measure of performance that quantifies this mismatch is utilization. We define the ***utilization*** of a process as

$$\text{Utilization} = \frac{\text{Flow rate}}{\text{Capacity}}$$

Utilization is a measure of how much the process *actually produces* relative to how much it *could produce* if it were running at full speed (i.e., its capacity). This is in line with the example of a car driving at 65 miles per hour (flow rate), despite being able to drive at 130 miles per hour (capacity): the car utilizes 65/130 = 50 percent of its potential.

Utilization, just like capacity, can be defined at the process level or the resource level. For example, the utilization of the process is the flow rate divided by the capacity of the process. The utilization of a particular resource is the flow rate divided by that resource's capacity.

For the Circored case, the resulting utilization is

$$\text{Utilization} = \frac{657{,}000 \text{ tons per year}}{876{,}000 \text{ tons per year}} = 0.75 = 75\%$$

In general, there are several reasons why a process might not produce at 100 percent utilization:

- If demand is less than supply, the process typically will not run at full capacity, but only produce at the rate of demand.
- If there is an insufficient supply of the input of a process, the process will not be able to operate at capacity.
- If one or several process steps only have a limited availability (e.g., maintenance and breakdowns), the process might operate at full capacity while it is running, but then go into periods of not producing any output while it is not running.

Given that the bottleneck is the resource with the lowest capacity and that the flow rate through all resources is identical, the bottleneck is the resource with the highest utilization.

In the case of the Circored plant, the corresponding utilizations are provided in **Table 3.2**. Note that all resources in a process with only one flow unit have the same flow rate, which is equal to the overall process flow rate. In this case, this is a flow rate of 657,000 tons per year.

TABLE 3.2 **Utilization of the Circored Process Steps Including Downtime**

Process Step	Calculations	Utilization
Preheater	657,000 tons/year/[120 tons/hour × 8,760 hours/year]	62.5%
Lock hoppers	657,000 tons/year/[110 tons/hour × 8,760 hours/year]	68.2%
First reactor	657,000 tons/year/[112 tons/hour × 8,760 hours/year]	66.9%
Second reactor	657,000 tons/year/[100 tons/hour × 8,760 hours/year]	75.0%
Flash heater	657,000 tons/year/[135 tons/hour × 8,760 hours/year]	55.6%
Discharger	657,000 tons/year/[118 tons/hour × 8,760 hours/year]	63.6%
Briquetting	657,000 tons/year/[165 tons/hour × 8,760 hours/year]	45.5%
Total process	657,000 tons/year/[100 tons/hour × 8,760 hours/year]	**75%**

Measuring the utilization of equipment is particularly common in capital-intensive industries, such as power plants, chemical plants, or air travel. Given limited demand and availability problems, the bottleneck in the Circored process did not operate at 100 percent utilization. We can summarize our computations graphically, by drawing a utilization profile. This is illustrated in **Figure 3.9**.

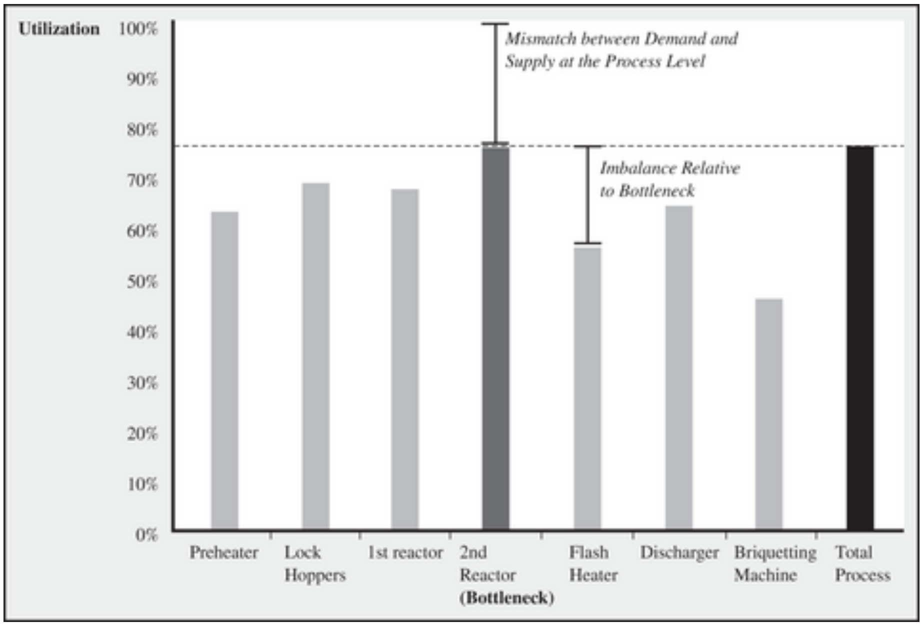


FIGURE 3.9 Utilization Profile

Although utilization is commonly tracked, it is a performance measure that should be handled with some care. Specifically, it should be emphasized that the objective of most businesses is to maximize profit, not to maximize utilization. As can be seen in **Figure 3.9**, there are two reasons in the Circored case for why an individual resource might not achieve 100 percent utilization, thus exhibiting excess capacity.

- First, given that no resource can achieve a higher utilization than the bottleneck, every process step other than the bottleneck will have a utilization gap relative to the bottleneck.
- Second, given that the process might not always be capacity-constrained, but rather be input- or demand-constrained, even the bottleneck might not be 100 percent utilized. In this case, every resource in the process has a "base level" of excess capacity, corresponding to the difference between the flow rate and the bottleneck capacity.

Note that the second reason disappears if there is sufficient market demand and full resource availability. In this case, only the bottleneck achieves a 100 percent utilization level. If the bottleneck in the Circored plant were utilized 100 percent, we would obtain an overall flow rate of 876,000 tons per year, or equivalently 100 tons per hour. The resulting utilization levels in that case are summarized in **Table 3.3**.

TABLE 3.3 **Utilization of the Circored Process Steps Assuming Unlimited Demand and No Downtime**

Process Step	Calculations	Utilization
Preheater	100/120	83.3%
Lock hoppers	100/110	90.9%
First reactor	100/112	89.3%
Second reactor	100/100	100.0%
Flash heater	100/135	74.1%
Discharger	100/118	84.7%
Briquetting machine	100/165	60.6%
Total process	**100/100**	**100%**

3.5 Workload and Implied Utilization

LO 3-5

Explain the difference between utilization and implied utilization.

Given the way we defined utilization (the ratio between flow rate and capacity), utilization can never exceed 100 percent. Thus, utilization only carries information about excess capacity, in which case utilization is strictly less than 100 percent. In contrast, we cannot infer from utilization by how much demand exceeds the capacity of the process. This is why we need to introduce an additional measure.

We define the ***implied utilization*** of a resource as

$$\text{Implied utilization} = \frac{\text{Demand}}{\text{Capacity}}$$

The implied utilization captures the mismatch between what could flow through the resource (demand) and what the resource can provide (capacity). Sometimes the "demand that could flow through a resource" is called the ***workload***. So you can also say that the implied utilization of a resource equals its workload divided by its capacity.

Assume that demand for the Circored ore would increase to 1,095,000 tons per year (125 tons per hour). **Table 3.4** calculates the resulting levels of implied utilization for the Circored resources.

TABLE 3.4 **Implied Utilization of the Circored Process Steps Assuming a Demand of 125 Tons per Hour and No Downtime**

Process Step	Calculations	Implied Utilization	Utilization
Preheater	125/120	104.2%	83.3%
Lock hoppers	125/110	113.6%	90.9%
First reactor	125/112	111.6%	89.3%

Process Step	Calculations	Implied Utilization	Utilization
Second reactor	125/100	125%	100.0%
Flash heater	125/135	92.6%	74.1%
Discharger	125/118	105.9%	84.7%
Briquetting machine	125/165	75.8%	60.6%
Total process	**125/100**	**125%**	**100%**

Several points in the table deserve further discussion:

- Unlike utilization, implied utilization can exceed 100 percent. Any excess over 100 percent reflects that a resource does not have the capacity available to meet demand.
- The fact that a resource has an implied utilization above 100 percent does not make it the bottleneck. As we see in **Table 3.4**, it is possible to have several resources with an implied utilization above 100 percent. However, there is only one bottleneck in the process! This is the resource where the implied utilization is the highest. In the Circored case, this is—not surprisingly—the second reactor. Would it make sense to say that the process has several bottlenecks? No! Given that we can only operate the Circored process at a rate of 100 tons per hour (the capacity of the first reactor), we have ore flow through every resource of the process at a rate of 100 tons per hour. Thus, while several resources have an implied utilization above 100 percent, all resources other than the second reactor have excess capacity (their utilizations in **Table 3.4** are below 100 percent). That is why we should not refer to them as bottlenecks.
- Having said this, it is important to keep in mind that in the case of a capacity expansion of the process, it might be worthwhile to add capacity to these other resources as well, not just to the bottleneck. In fact, depending on the margins we make and the cost of installing capacity, we could make a case to install additional capacity for all resources with an implied utilization above 100 percent. In other words, once we add capacity to the current bottleneck, our new process (with a new bottleneck) could still be capacity-constrained, justifying additional capacity to other resources.

Page 44

3.6 Multiple Types of Flow Units

LO 3-6

Explain why it is important to choose the correct flow unit when preparing a process flow diagram.

Choosing an appropriate flow unit is an essential step when preparing a process flow diagram. While, for the examples we have discussed so far, this looked relatively straightforward, there are many situations that you will encounter where this choice requires more care. The two most common complications are

- The flow of the unit moving through the process breaks up into multiple flows. For example, in an assembly environment, following an inspection step, good units continue to the next processing step, while bad units require rework.
- There are multiple types of flow units, representing, for example, different customer types. In an emergency room, life-threatening cases follow a different flow than less complicated cases.

The critical issue in choosing the flow unit is that you must be able to express all demands and capacities in terms of the chosen flow unit. For example, in the Circored process, we chose one ton of ore to be the flow unit. Thus, we had to express each resource's capacity and the demand in terms of tons of ore. Given that the process only makes ore, the choice of the flow unit was straightforward. However, consider the following example involving multiple product or customer types. An employment verification agency receives resumés from consulting firms and law firms with the request to validate information provided by their job candidates.

Figure 3.10 shows the process flow diagram for this agency. Note that while the three customer types share the first step and the last step in the process (filing and sending confirmation letter), they differ with respect to other steps:

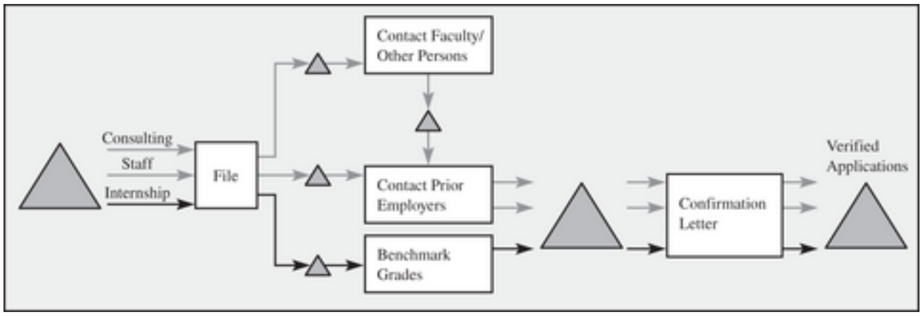


FIGURE 3.10 **Process Flow Diagram with Multiple Product Types**

- For internship positions, the agency provides information about the law school/business school the candidate is currently enrolled in as well as previous institutions of higher education and, to the extent possible, provides information about the applicant's course choices and honors.
- For staff positions, the agency contacts previous employers and analyzes the letters of recommendation from those employers.
- For consulting/lawyer positions, the agency attempts to call former supervisors and/or colleagues in addition to contacting the previous employers and analyzes the letters of recommendation from those employers.

As far as demand, this process receives 3 consulting, 11 staff, and 4 internship applications per hour. **Table 3.5** also provides the capacities of each activity, in applications per hour. Given that the workload on each activity as well as all of the capacities can be expressed in terms of "applications per hour," we can choose "one application" as our flow unit, despite the fact that there are multiple types of applications.

TABLE 3.5 **Finding the Bottleneck in the Multiproduct Case**

				Workload [Minutes/Hour]			
	Processing Time	**Number of Workers**	**Capacity**	**Consulting**	**Staff**	**Interns**	**Total**
File	3 [min./appl.]	1	1/3 [appl./min.] = 20 [appl./hour]	3	11	4	18
Contact persons	20 [min./appl.]	2	2/20 [appl./min.] = 6 [appl./hour]	3	0	0	3

				Workload [Minutes/Hour]			
Contact employers	15 [min./appl.]	3	3/15 [appl./min.] = 12 [appl./hour]	3	11	0	14
Grade/school analysis	8 [min./appl.]	2	2/8 [appl./min.] = 15 [appl./hour]	0	0	4	4
Confirmation letter	2 [min./appl.]	1	1/2 [appl./min.] = 30 [appl./hour]	3	11	4	18

The next step in our process analysis is to find the bottleneck. In this setting this is complicated by the *product mix* (different types of customers flowing through one process). For example, the process step "contact persons" might have a very long processing time, resulting in a low capacity for this activity. However, if the workload on this activity (applications per hour) is also very low, then maybe this low capacity is not an issue.

To find the bottleneck and to determine capacity in a multiproduct situation, we need to compare each activity's capacity with its demand. The analysis is given in **Table 3.5**.

To compute the demand on a given activity as shown in **Table 3.5**, it is important to remember that some activities (e.g., filing the applications) are requested by all product types, whereas others (e.g., contacting faculty and former colleagues) are requested by one product type. This is (hopefully) clear by looking at the process flow diagram.

To complete our analysis, divide each activity's demand by its capacity to yield each activity's implied utilization. This allows us to find the busiest resource. In this case, it is "contact prior employers," so this is our bottleneck. As the implied utilization is above 100 percent, the process is capacity-constrained.

The flow unit "one application" allowed us to evaluate the implied utilization of each activity in this process, but it is not the only approach. Alternatively, we could define the flow unit as "one minute of work." This might seem like an odd flow unit, but it has an advantage over "one application." Before explaining its advantage, let's figure out how to replicate our analysis of implied utilization with this new flow unit.

As before, we need to define our demands and our capacities in terms of our flow unit. In the case of capacity, each worker has "60 minutes of work" available per hour. (By definition, we all do!) So the capacity of an activity is (Number of workers) × 60 [minutes/hour]. For example, "contact persons" has two workers. So its capacity is 2 × 60 =

120 minutes of work per hour. Each worker has 60 “minutes of work” available per hour, so two of them can deliver 120 minutes of work.

Now turn to the demands. There are 11 staff applications to be processed each hour and each takes 3 minutes. So the demand for staff applications is 11 × 3 = 33 minutes per hour. Now that we know how to express the demands and the capacities in terms of the “minutes of work,” the implied utilization of each activity is again the ratio of the amount demanded from the activity to the activity’s capacity. **Table 3.6** summarizes these calculations. As we would expect, this method yields the same implied utilizations as the “one application” flow unit approach.

TABLE 3.6 Using “One Minute of Work” as the Flow Unit to Find the Bottleneck in the Multiproduct Case

				Workload [Minutes/Hour]			
	Processing Time	**Number of Workers**	**Capacity**	**Consulting**	**Staff**	**Interns**	**Total**
File	3 [min./appl.]	1	60 [min./hour]	3 × 3	11 × 3	4 × 3	54
Contact persons	20 [min./appl.]	2	120 [min./hour]	3 × 20	0	0	60
Contact employers	15 [min./appl.]	3	180 [min./hour]	3 × 15	11 × 15	0	210
Grade/school analysis	8 [min./appl.]	2	120 [min./hour]	0	0	4 × 8	32
Confirmation letter	2 [min./appl.]	1	60 [min./hour]	3 × 2	11 × 2	4 × 2	36

So if “one application” and “one minute of work” give us the same answer, how should we choose between these approaches? In this situation, you would work with the approach that you find most intuitive (which is probably “one application,” at least initially) because they both allow us to evaluate the implied utilizations. However, the “one minute of work” approach is more robust. To explain why, suppose it took 3 minutes to file a staff application, 5 minutes to file a consulting application, and 2 minutes to file an internship application. In this case, we get into trouble if we define the flow unit to be “one application”—with that flow unit, we cannot express the capacity of the file activity! If we receive only internship applications, then filing could process 60/2 = 30 applications per hour. However, if we receive only consulting applications, then filing can only process 60/5

= 12 applications per hour. The number of applications per hour that filing can process depends on the mix of applications! The "minute of work" flow unit completely solves that problem—no matter what mix of applications is sent to filing, with one worker, filing has 60 minutes of work available per hour. Similarly, for a given mix of applications, we can also evaluate the workload on filing in terms of minutes of work (just as is done in **Table 3.6**).

To summarize, choose a flow unit that allows you to express all demands and capacities in terms of that flow unit. An advantage of the "minute of work" (or "hour of work," "day of work," etc.) approach is that it is possible to do this even if there are multiple types of products or customers flowing through the process.

So what is the next step in our process analysis? We have concluded that it is capacity-constrained because the implied utilization of "contact employers" is greater than 100 percent—it is the bottleneck. Given that it is the only activity with an implied utilization greater than 100 percent, if we are going to add capacity to this process, "contact employers" should be the first candidate—in the current situation, they simply do not have enough capacity to handle the current mix of customers. Notice, if the mix of customers changes, this situation might change. For example, if we started to receive fewer staff applications (which have to flow through "contact employers") and more internship applications (which do not flow through "contact employers"), then the workload on "contact employers" would decline, causing its implied utilization to fall as well. Naturally, shifts in the demands requested from a process can alter which resource in the process is the bottleneck.

Although we have been able to conclude something useful with our analysis, one should be cautious to not conclude too much when dealing with multiple types of products or customers. To illustrate some potential complications, consider the following example. At the international arrival area of a major U.S. airport, 15 passengers arrive per minute, 10 of whom are U.S. citizens or permanent residents and 5 are visitors.

The immigration process is organized as follows. Passengers disembark their aircraft and use escalators to arrive at the main immigration hall. The escalators can transport up to 100 passengers per minute. Following the escalators, passengers have to go through immigration. There exist separate immigration resources for U.S. citizens and permanent residents (they can handle 10 passengers per minute) and visitors (which can handle 3 visitors per minute). After immigration, all passengers pick up their luggage. Luggage handling (starting with getting the luggage off the plane and ending with moving the luggage onto the conveyor belts) has a capacity of 10 passengers per minute. Finally, all passengers go through customs, which has a capacity of 20 passengers per minute.

We calculate the implied utilization levels in Table 3.7. Notice when evaluating implied utilization, we assume the demand on luggage handling is 10 U.S. citizens and 5 visitors even though we know (or discover via our calculations) that it is not possible for 15 passengers to arrive to luggage handling per minute (there is not enough capacity in immigration). We do this because we want to compare the potential demand on each resource with its capacity to assess its implied utilization. Consequently, we can evaluate each resource's implied utilization in isolation from the other resources.

TABLE 3.7 Calculating Implied Utilization in Airport Example

Resource	Demand for U.S. Citizens and Permanent Residents [Pass./Min.]	Demand for Visitors [Pass./Min.]	Capacity [Pass./Min.]	Implied Utilization
Escalator	10	5	100	15/100 = 15%
Immigration—U.S. residents	10	0	10	10/10 = 100%
Immigration—visitors	0	5	3	5/3 = 167%
Luggage handling	10	5	10	15/10 = 150%
Customs	10	5	20	15/20 = 75%

Based on the values in Table 3.7, the bottleneck is immigration for visitors because it has the highest implied utilization. Furthermore, because its implied utilization is greater than 100 percent, the process is supply-constrained. Given that there is too little supply, we can expect queues to form. Eventually, those queues will clear because the demand rate of arriving passengers will at some point fall below capacity (otherwise, the queues will just continue to grow, which we know will not happen indefinitely at an airport). But during the times in which the arrival rates of passengers is higher than our capacity, where will the queues form? The answer to this question depends on how we prioritize work.

The escalator has plenty of capacity, so no priority decision needs to be made there. At immigration, there is enough capacity for 10 U.S. citizens and 3 visitors. So 13 passengers may be passed on to luggage handling, but luggage handling can accommodate only 10 passengers. Suppose we give priority to U.S. citizens. In that case, all of the U.S. citizens proceed through luggage handling without interruption, and a queue of visitors will form at the rate of 3 per minute. Of course, there will also be a queue of visitors in front of immigration, as it can handle only 3 per minute while 5 arrive per minute. With this priority scheme, the outflow from this process will be 10 U.S. citizens per minute. However, if we give visitors full priority at luggage handling, then a similar analysis reveals that a queue of U.S. citizens forms in front of luggage handling, and a queue of visitors forms in front of immigration. The outflow is 7 U.S. citizens and 3 visitors.

The operator of the process may complain that the ratio of U.S. citizens to visitors in the outflow (7 to 3) does not match the inflow ratio (2 to 1), even though visitors are given full priority. If we were to insist that those ratios match, then the best we could do is have an outflow of 6 U.S. citizens and 3 visitors—we cannot produce more than 3 visitors per minute given the capacity of immigration, so the 2 to 1 constraint implies that we can "produce" no more than 6 U.S. citizens per minute. Equity surely has a price in this case—we could have an output of 10 passengers per minute, but the equity constraint would limit us to 9 passengers per minute. To improve upon this output while maintaining the equity constraint, we should add more capacity at the bottleneck—immigration for visitors.

3.7 Summary

Figure 3.11 is a summary of the major steps graphically. Exhibits 3.1 and 3.2 summarize the steps required to do the corresponding calculations for a single flow unit and multiple flow units, respectively.

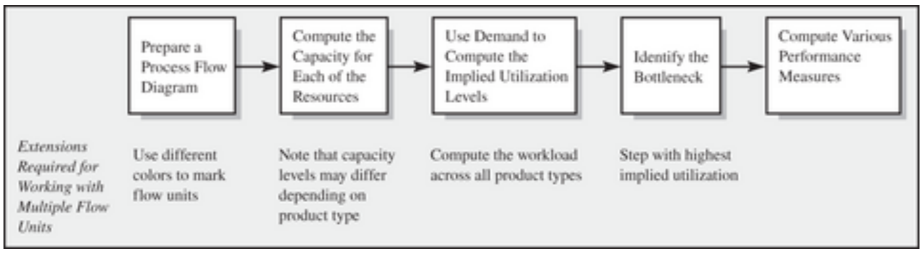


FIGURE 3.11 Summary of Process Analysis

Exhibit 3.1

STEPS FOR BASIC PROCESS ANALYSIS WITH ONE TYPE OF FLOW UNIT

1. Find the capacity of every resource; if there are multiple resources performing the same activity, add their capacities together.
2. The resource with the lowest capacity is called the ***bottleneck***. Its capacity determines the capacity of the entire process (***process capacity***).
3. The flow rate is found based on

$$\text{Flow rate} = \text{Minimum}\{\text{Available input, Demand, Process capacity}\}$$

4. We find the utilization of the process as

$$\text{Utilization} = \frac{\text{Flow rate}}{\text{Capacity}}$$

The utilization of each resource can be found similarly.

Exhibit 3.2

STEPS FOR BASIC PROCESS ANALYSIS WITH MULTIPLE TYPES OF FLOW UNITS

1. For each resource, compute the number of minutes that the resource can produce; this is 60 [min./hour] × Number of resources within the resource pool.
2. Create a process flow diagram, indicating how the flow units go through the process; use multiple colors to indicate the flow of the different flow units.
3. Create a table indicating how much workload each flow unit is consuming at each resource:
   - The rows of the table correspond to the resources in the process.
   - The columns of the table correspond to the different types of flow units.
   - Each cell of the table should contain one of the following:

   If flow unit does not visit the corresponding resource, 0; otherwise, demand per hour of the corresponding flow unit × processing time.
4. Add up the workload of each resource across all flow units.
5. Compute the implied utilization of each resource as

$$\text{Implied utilization} = \frac{\text{Result of step 4}}{\text{Result of step 1}}$$

The resource with the highest implied utilization is the bottleneck.

The preceding approach is based on Table 3.6; that is, the flow unit is “one minute of work.”

Any process analysis should begin with the creation of a process flow diagram. This is especially important for the case of multiple flow units, as their flows are typically more complex.

Next, we need to identify the bottleneck of the process. As long as there exists only one type of flow unit, this is simply the resource with the lowest capacity. However, for more general cases, we need to perform some extra analysis. Specifically, if there is a product mix, we have to compute the requested capacity (workload) at each resource and then compare it to the available capacity. This corresponds to computing the implied utilization, and we identify the bottleneck as the resource with the highest implied utilization.

Finally, once we have found the bottleneck, we can compute a variety of performance measures. As in the previous chapter, we are interested in finding the flow rate. The flow rate also allows us to compute the process utilization as well as the utilization profile across resources. Utilizations, while not necessarily a business goal by themselves, are important measures in many industries, especially capital-intensive industries.

3.8 Practice Problems and Selected Solutions

The following questions will help in testing your understanding of this chapter. After each question, we show the relevant section in parentheses [Section x].

Solutions to problems marked with an "*" appear at the end of this section. Video solutions to select problems are available in Connect.

Q3.1* (**Process Analysis with One Flow Unit**) Consider a process consisting of three resources:

Resource	Processing Time [Min./Unit]	Number of Workers
1	10	2
2	6	1
3	16	3

What is the bottleneck? What is the process capacity? What is the flow rate if demand is eight units per hour? [**3.2**] What is the utilization of each resource if demand is eight units per hour? [**3.4**]

Q3.2* (**Process Analysis with Multiple Flow Units**) Consider a process consisting of five resources that are operated eight hours per day. The process works on three different products, A, B, and C:

Resource	Number of Workers	Processing Time for A [Min./Unit]	Processing Time for B [Min./Unit]	Processing Time for C [Min./Unit]
1	2	5	5	5
2	2	3	4	5
3	1	15	0	0
4	1	0	3	3
5	2	6	6	6

Demand for the three different products is as follows: product A, 40 units per day; product B, 50 units per day; and product C, 60 units per day.

What is the bottleneck? What is the flow rate for each flow unit assuming that demand must be served in the mix described above (i.e., for every four units of A, there are five units of B and six units of C)? [**3.6**]

Q3.3 **(Cranberries)** International Cranberry Uncooperative (ICU) is a competitor to the National Cranberry Cooperative (NCC). At ICU, barrels of cranberries arrive on trucks at a rate of 150 barrels per hour and are processed continuously at a rate of 100 barrels per hour. Trucks arrive at a uniform rate over eight hours, from 6:00 a.m. until 2:00 p.m. Assume the trucks are sufficiently small so that the delivery of cranberries can be treated as a continuous inflow. The first truck arrives at 6:00 a.m. and unloads immediately, so processing begins at 6:00 a.m. The bins at ICU can hold up to 200 barrels of cranberries before overflowing. If a truck arrives and the bins are full, the truck must wait until there is room in the bins.

a. What is the maximum number of barrels of cranberries that are waiting on the trucks at any given time? [**3.3**]
b. At what time do the trucks stop waiting? [**3.3**]
c. At what time do the bins become empty? [**3.3**]
d. ICU is considering using seasonal workers in addition to their regular workforce to help with the processing of cranberries. When the seasonal workers are working, the processing rate increases to 125 barrels per hour. The seasonal workers would start working at 10:00 a.m. and finish working when the trucks stop waiting. At what time would ICU finish processing the cranberries using these seasonal workers? [**3.3**]

Q3.4 **(Western Pennsylvania Milk Company)** The Western Pennsylvania Milk Company is producing milk at a fixed rate of 5,000 gallons/hour. The company's clients request 100,000 gallons of milk over the course of one day. This demand is spread out uniformly from 8 a.m. to 6 p.m. If there is no milk available, clients will wait until enough is produced to satisfy their requests.

The company starts producing at 8 a.m. with 25,000 gallons in finished goods inventory. At the end of the day, after all demand has been fulfilled, the plant keeps on producing until the finished goods inventory has been restored to 25,000 gallons.

When answering the following questions, treat trucks/milk as a continuous flow process. Begin by drawing a graph indicating how much milk is in inventory and how much milk is "back-ordered" over the course of the day.

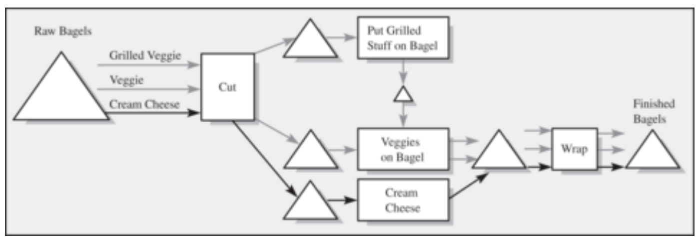
a. At what time during the day will the clients have to start waiting for their requests to be filled? [3.3]

b. At what time will clients stop waiting? [3.3]

c. Assume that the milk is picked up in trucks that hold 1,250 gallons each. What is the maximum number of trucks that are waiting? [3.3]

d. Assume the plant is charged $50 per hour per waiting truck. What are the total waiting time charges on a day? [3.3]

Q3.5** **(Bagel Store)** Consider a bagel store selling three types of bagels that are produced accor process flow diagram outlined below. We assume the demand is 180 bagels a day, of whicl 30 grilled veggie, 110 veggie only, and 40 cream cheese. Assume that the workday is 10 ho and each resource is staffed with one worker.



Moreover, we assume the following processing times:

	Cut	Grilled Stuff	Veggies	Cream Cheese	
Processing time	3 [min./bagel]	10 [min./bagel]	5 [min./bagel]	4 [min./bagel]	2 [m

Processing times are independent of which bagel type is processed at a resource (e.g., cutting a bagel same time for a cream cheese bagel as for a veggie bagel).

a. Where in the process is the bottleneck? [3.6]

b. How many units can the process produce within one hour, assuming the product mix has to rema constant? [3.6]

Q3.6 **(Valley Forge Income Tax Advice)** VF is a small accounting firm supporting wealthy individuals in their preparation of annual income tax statements. Every December, VF sends out a short survey to their customers, asking for the information required for preparing the tax statements. Based on 24 years of experience, VF categorizes their cases into the following groups:

- Group 1 (new customers, easy): 15 percent of cases

- Group 2 (new customers, complex): 5 percent of cases
- Group 3 (repeat customers, easy): 50 percent of cases
- Group 4 (repeat customers, complex): 30 percent of cases

Here, “easy” versus “complex” refers to the complexity of the customer’s earning situation.

In order to prepare the income tax statement, VF needs to complete the following set of activities. Processing times (and even which activities need to be carried out) depend on which group a tax statement falls into. All of the following processing times are expressed in minutes per income tax statement.

Group	Filing	Initial Meeting	Preparation	Review by Senior Accountant	Writing
1	20	30	120	20	50
2	40	90	300	60	80
3	20	No meeting	80	5	30
4	40	No meeting	200	30	60

The activities are carried out by the following three persons:

- Administrative support person: filing and writing.
- Senior accountant (who is also the owner): initial meeting, review by senior accountant.
- Junior accountant: preparation.

Assume that all three persons work eight hours per day and 20 days a month. For the following questions, assume the product mix as described above. Assume that there are 50 income tax statements arriving each month.

a. Which of the three persons is the bottleneck? [3.6]

b. What is the (implied) utilization of the senior accountant? The junior accountant? The administrative support person? [3.6]

c. You have been asked to analyze which of the four product groups is the most profitable. Which factors would influence the answer to this? [3.6]

d. How would the process capacity of VF change if a new word processing system would reduce the time to write the income tax statements by 50 percent? [3.6]

Q3.7 **(Car Wash Supply Process)** CC Car Wash specializes in car cleaning services. The services offered by the company, the exact service time, and the resources needed for each of them are described in the following table:

Service	Description	Processing Time	Resource Used
A. Wash	Exterior car washing and drying	10 min.	1 automated washing machine
B. Wax	Exterior car waxing	10 min.	1 automated waxing machine
C. Wheel cleaning	Detailed cleaning of all wheels	7 min.	1 employee
D. Interior cleaning	Detailed cleaning inside the car	20 min.	1 employee

The company offers the following packages to its customers:

- Package 1: Includes only car wash (service A).
- Package 2: Includes car wash and waxing (services A and B).
- Package 3: Car wash, waxing, and wheel cleaning (services A, B, and C).
- Package 4: All four services (A, B, C, and D).

Customers of CC Car Wash visit the station at a constant rate (you can ignore any effects of variability) of 40 customers per day. Of these customers, 40 percent buy Package 1, 15 percent buy Package 2, 15 percent buy Package 3, and 30 percent buy Package 4. The mix does not change over the course of the day. The store operates 12 hours a day.

a. What is the implied utilization of the employee doing the wheel cleaning service? [3.6]

b. Which resource has the highest implied utilization? [3.6] For the next summer, CC Car Wash anticipates an increase in the demand to 80 customers per day. Together with this demand increase, there is expected to be a change in the mix of packages demanded: 30 percent of the customers ask for Package 1, 10 percent for Package 2, 10 percent for Package 3, and 50 percent for Package 4. The company will install an additional washing machine to do service A.

c. What will be the new bottleneck in the process? [**3.6**]

d. How many customers a day will not be served? Which customers are going to wait? Explain your reasoning! [**3.6**]

Q3.8 **(Starbucks)** After an "all night" study session the day before their last final exam, four students decide to stop for some much-needed coffee at the campus Starbucks. They arrive at 8:30 a.m. and are dismayed to find a rather long line.

Fortunately for the students, a Starbucks executive happens to be in line directly in front of them. From her, they learn the following facts about this Starbucks location:

I. There are three employee types:

- There is a single cashier who takes all orders, prepares nonbeverage food items, grinds coffee, and pours drip coffee.
- There is a single frozen drink maker who prepares blended and iced drinks.
- There is a single espresso drink maker who prepares espressos, lattes, and steamed drinks.

II. There are typically four types of customers:

- Drip coffee customers order only drip coffee. This requires 20 seconds of the cashier's time to pour the coffee.
- Blended and iced drink customers order a drink that requires the use of the blender. These drinks take on average two minutes of work of the frozen drink maker.
- Espresso drink customers order a beverage that uses espresso and/or steamed milk. On average, these drinks require one minute of work of the espresso drink maker.
- Ground coffee customers buy one of Starbucks' many varieties of whole bean coffee and have it ground to their specification at the store. This requires a total of one minute of the cashier's time (20 seconds to pour the coffee and 40 seconds to grind the whole bean coffee).

III. The customers arrive uniformly at the following rates from 7 a.m. (when the store opens) until 10 a.m. (when the morning rush is over), with no customers arriving after 10 a.m.:

- Drip coffee customers: 25 per hour.

- Blended and iced drink customers: 20 per hour.
- Espresso drink customers: 70 per hour.
- Ground coffee customers: 5 per hour.

IV. Each customer spends, on average, 20 seconds with the cashier to order and pay.

V. Approximately 25 percent of all customers order food, which requires an additional 20 seconds of the cashier's time per transaction.

While waiting in line, the students reflect on these facts and they answer the following questions:

a. What is the implied utilization of the frozen drink maker? [3.6]

b. Which resource has the highest implied utilization? [3.6] From their conversation with the executive, the students learn that Starbucks is considering a promotion on all scones (half price!), which marketing surveys predict will increase the percentage of customers ordering food to 30 percent (the overall arrival rates of customers will *not* change). However, the executive is worried about how this will affect the waiting times for customers.

c. How do the levels of implied utilization change as a response to this promotion? [3.6]

Q3.9 **(Paris Airport)** Kim Opim, an enthusiastic student, is on her flight over from Philadelphia (PHL) to Paris. Kim reflects upon how her educational experiences from her operations courses could help explain the long wait time that she experienced before she could enter the departure area of Terminal A at PHL. As an airline representative explained to Kim, there are four types of travelers in Terminal A:

- Experienced short-distance (short-distance international travel destinations are Mexico and various islands in the Atlantic) travelers: these passengers check in online and do not speak with any agent nor do they take any time at the kiosks.
- Experienced long-distance travelers: these passengers spend three minutes with an agent.
- Inexperienced short-distance travelers: these passengers spend two minutes at a kiosk; however, they do not require the attention of an agent.
- Inexperienced long-distance travelers: these passengers need to talk five minutes with an agent.

After a passenger checks in online, or talks with an agent, or uses a kiosk, the passenger must pass through security, where they need 0.5 minute independent of their type. From historical data, the airport is able to estimate the arrival rates of the different customer types at Terminal A of Philadelphia International:

- Experienced short-distance travelers: 100 per hour
- Experienced long-distance travelers: 80 per hour
- Inexperienced short-distance travelers: 80 per hour
- Inexperienced long-distance travelers: 40 per hour

At this terminal, there are four security check stations, six agents, and three electronic kiosks. Passengers arrive uniformly from 4 p.m. to 8 p.m., with the entire system empty prior to 4 p.m. (the "midafternoon lull") and no customers arrive after 8 p.m. All workers must stay on duty until the last passenger is entirely through the system (e.g., has passed through security).

a. What are the levels of implied utilization at each resource? [**3.6**]

b. At what time has the last passenger gone through the system? Note: If passengers of one type have to wait for a resource, passengers that do not require service at the resource can pass by the waiting passengers! [**3.6**]

c. Kim, an experienced long-distance traveler, arrived at 6 p.m. at the airport and attempted to move through the check-in process as quickly as she could. How long did she have to wait before she was checked at security? [**3.6**]

d. The airline considers showing an educational program that would provide information about the airport's check-in procedures. Passenger surveys indicate that 80 percent of the inexperienced passengers (short or long distance) would subsequently act as experienced passengers (i.e., the new arrival rates would be 164 experienced short-distance, 112 experienced long-distance, 16 inexperienced short-distance, and 8 inexperienced long-distance [passengers/hour]). At what time has the last passenger gone through the system? [**3.6**]

If you would like to test your understanding of a specific section, here are the questions organized by section:

Section 3.2: Q1

Section 3.3: Q3, Q4

Section 3.4: Q1

Section 3.6: Q2, Q5, Q6, Q7, Q8, Q9

Selected Solutions

Q3.1 (Process Analysis with One Flow Unit)

The following steps refer directly to **Exhibit 3.1**.

Step 1. We first compute the capacity of the three resources:

$$\text{Resource 1:}\quad \frac{2}{10}\,\text{unit per minute} = 0.2\,\text{unit per minute}$$

$$\text{Resource 2:}\quad \frac{1}{6}\,\text{unit per minute} = 0.1666\,\text{unit per minute}$$

$$\text{Resource 3:}\quad \frac{3}{16}\,\text{unit per minute} = 0.1875\,\text{unit per minute}$$

Step 2. Resource 2 has the lowest capacity; process capacity therefore is 0.1666 unit per minute, which is equal to 10 units per hour.

Step 3.

$$\begin{aligned}\text{Flow rate} &= \text{Min}\{\text{Process capacity, Demand}\}\\ &= \text{Min}\{8 \text{ units per hour, } 10 \text{ units per hour}\} = 8 \text{ units per hour}\end{aligned}$$

This is equal to 0.1333 unit per minute.

Step 4. We find the utilizations of the three resources as

Resource 1: 0.1333 unit per minute/0.2 unit per minute = 66.66 percent

Resource 2: 0.1333 unit per minute/0.1666 unit per minute = 80 percent

Resource 3: 0.1333 unit per minute/0.1875 unit per minute = 71.11 percent

Page 55

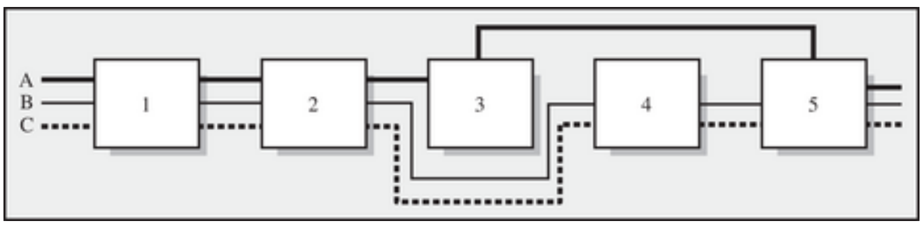
Q3.2 (Process Analysis with Multiple Flow Units)

The following steps refer directly to **Exhibit 3.2**.

Step 1. Each resource can contribute the following capacity (in minutes of work per day):

Resource	Number of Workers	Minutes per Day
1	2	2 × 8 × 60 = 960
2	2	2 × 8 × 60 = 960
3	1	1 × 8 × 60 = 480
4	1	1 × 8 × 60 = 480
5	2	2 × 8 × 60 = 960

Step 2. Process flow diagram:



Step 3. We create a table indicating how much capacity will be consumed by the three products at the resources.

Resource	Capacity Requirement from A	Capacity Requirement from B	Capacity Requirement from C
1	5 × 40 = 200	5 × 50 = 250	5 × 60 = 300
2	3 × 40 = 120	4 × 50 = 200	5 × 60 = 300
3	15 × 40 = 600	0 × 50 = 0	0 × 60 = 0
4	0 × 40 = 0	3 × 50 = 150	3 × 60 = 180
5	6 × 40 = 240	6 × 50 = 300	6 × 60 = 360

Step 4. Add up the rows to get the workload for each resource:

Workload for resource 1: 200 + 250 + 300 = 750

Workload for resource 2: 120 + 200 + 300 = 620

Workload for resource 3: 600 + 0 + 0 = 600

Workload for resource 4: 0 + 150 + 180 = 330

Workload for resource 5: 240 + 300 + 360 = 900

Resource	Minutes per Day (see Step 1)	Workload per Day (see Step 4)	Implied Utilization (Step 4/Step 1)
1	960	750	0.78
2	960	620	0.65
3	480	600	1.25
4	480	330	0.69
5	960	900	0.94

Step 5. Compute implied utilization levels. Hence, resource 3 is the bottleneck. Thus, we cannot produce units A at a rate of 40 units per day. Since we are overutilized by 25 percent, we can produce units A at a rate of 32 units per day (four units per hour). Assuming the ratio between A, B, and C is constant (40:50:60), we will produce B at five units per hour and C at six units per hour. If the ratio between A, B, and C is *not* constant, this answer changes. In this case, we would produce 32 units of A and produce products B and C at the rate of demand (50 and 60 units per day, respectively).

CHAPTER 4

Estimating and Reducing Labor Costs

The objective of any process should be to create value (make profits), not to maximize the utilization of every resource involved in the process. In other words, we should not attempt to produce more than what is demanded from the market, or from the resource downstream in the process, just to increase the utilization measure. Yet, the underutilization of a resource, human labor or capital equipment alike, provides opportunities to improve the process. This improvement can take several forms, including

- If we can reduce the excess capacity at some process step, the overall process becomes more efficient (lower cost for the same output).
- If we can use capacity from underutilized process steps to increase the capacity at the bottleneck step, the overall process capacity increases. If the process is capacity-constrained, this leads to a higher flow rate.

In this chapter, we discuss how to achieve such process improvements. Specifically, we discuss the concept of line balancing, which strives to avoid mismatches between what is supplied by one process step and what is demanded from the following process step (referred to as the process step downstream). In this sense, line balancing attempts to match supply and demand within the process itself.

We use Novacruz Inc. to illustrate the concept of line balancing and to introduce a number of more general terms of process analysis. Novacruz is the producer of a high-end kick scooter, known as the Xootr (pronounced "zooter"), displayed in **Figure 4.1**.