Contents

Preface

0.1 Why this book .................. i
0.2 Organization .................. i
0.3 Resources .................. ii
0.4 But first.................. ii

1 Introduction 1

2 Material flow 5

2.1 The warehouse as a queueing system ................ 6
2.1.1 Extensions .................. 8
2.2 Questions .................. 9

3 Warehouse operations 11

3.1 Receiving .................. 12
3.2 Put-away .................. 12
3.3 Process customer orders .................. 13
3.4 Order-picking .................. 13
3.4.1 Sharing the work of order-picking .................. 14
3.5 Checking and packing .................. 16
3.6 Shipping .................. 16
3.7 Summary .................. 16
3.8 More .................. 17
3.9 Questions .................. 18

4 Storage and handling equipment 19

4.1 Storage equipment .................. 19
4.1.1 Pallet storage .................. 20
4.1.2 Bin-shelving or static rack .................. 23
4.1.3 Gravity flow rack .................. 23
4.1.4 Carousels .................. 26
4.2 Conveyors .................. 27
4.3 Sortation equipment .................. 27
4.4 Summary .................. 27
4.5 On the lighter side .................................................. 28
4.6 Questions ................................................................. 29

5 Pallet operations ......................................................... 31
5.1 Labor ................................................................. 31
    5.1.1 Storage locations ............................................. 32
5.2 Location of receiving and shipping .......................... 33
5.3 Space ................................................................. 36
    5.3.1 Stack height .................................................. 36
    5.3.2 Lane depth .................................................. 37
5.4 Summary .............................................................. 41
5.5 More ................................................................. 42
5.6 Questions .............................................................. 43

6 Case-picking ............................................................. 47
6.1 Summary .............................................................. 47

7 Design of a fast pick area ........................................... 49
7.1 What is a fast-pick area? .......................................... 49
7.2 Estimating restocks ................................................. 50
7.3 Storing optimal amounts ......................................... 50
7.4 Allocating space in the fast-pick area ...................... 52
    7.4.1 Two commonly-used storage strategies ................. 54
    7.4.2 Minimum and maximum allocations ..................... 56
7.5 Which sku’s go into the fast-pick area? ...................... 57
    7.5.1 Managing storage slots ..................................... 60
    7.5.2 Priority of claim for storage in the fast-pick area .... 62
7.6 Additional issues .................................................. 64
    7.6.1 Storage by family ........................................... 64
    7.6.2 Reorder points .............................................. 64
    7.6.3 Limits on capacity ......................................... 65
    7.6.4 Accounting for on-hand inventory levels ............... 65
    7.6.5 Setup costs .................................................. 66
7.7 Limitations of the fluid model .................................. 66
7.8 Size of the fast-pick area ........................................ 67
    7.8.1 How large should the fast-pick area be? ............... 67
    7.8.2 How can the fast-pick area be made larger? ........... 69
7.9 Multiple fast-pick areas ........................................... 70
7.10 A fast-pick area in pallet storage .......................... 71
7.11 On the lighter side ................................................. 73
7.12 Summary .............................................................. 73
7.13 Questions .............................................................. 75
8 Detailed slotting 81
  8.1 Case orientation and stack level ........................... 81
  8.2 Measuring the effectiveness of a slotting .................. 83
  8.3 Packing algorithms ........................................... 84
  8.4 Questions ....................................................... 85

9 Order-Picking by ‘bucket brigade’ 87
  9.1 Introduction ..................................................... 87
  9.2 Order-assembly by bucket brigade ................................
      9.2.1 A model .................................................. 88
      9.2.2 Improvements that are not ................................ 92
      9.2.3 Some advantages of bucket brigades .................... 93
  9.3 Bucket brigades in the warehouse .............................. 94
  9.4 Summary ......................................................... 96
  9.5 Questions ....................................................... 98

10 Warehouse activity profiling 99
  10.1 Basics .......................................................... 99
  10.2 A closer look: ABC analysis ................................... 100
  10.3 Universal warehouse descriptors ............................... 102
  10.4 A more detailed look .......................................... 103
      10.4.1 Data sources ............................................. 103
      10.4.2 Where is the work? ...................................... 106
      10.4.3 What are the patterns of work? .......................... 109
  10.5 Doing it ........................................................ 111
      10.5.1 Getting the data ........................................ 111
      10.5.2 Data-mining .............................................. 112
      10.5.3 Discrepancies in the data ................................ 113
      10.5.4 Viewing the data ......................................... 114
  10.6 Summary ........................................................ 114
  10.7 On the lighter side ............................................ 114
  10.8 Questions ....................................................... 117

11 Benchmarking warehouse performance 119

12 Crossdocking 121
  12.1 Why have a crossdock? ....................................... 121
  12.2 Operations ..................................................... 122
  12.3 Freight flow .................................................... 123
      12.3.1 Congestion ............................................... 123
  12.4 Design .......................................................... 124
      12.4.1 Size ...................................................... 124
      12.4.2 Geometry ................................................ 125
  12.5 Trailer management ........................................... 126
  12.6 Resources ....................................................... 126
  12.7 Questions ....................................................... 126
13 Trends 129
List of Figures

2.1 A product is generally handled in smaller units as it moves down the supply chain. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17). ................................. 7


4.4 Profiles of three styles of carton flow rack. Each successive style takes more space but makes the cartons more readily accessible. ........................................ 25


5.1 The economic contours of a warehouse floor in which receiving is at the middle bottom and shipping at the middle top. The pallet positions colored most darkly are the most convenient. ........................................ 33

5.2 The economic contours of a warehouse floor in which receiving and shipping are both located at the middle bottom. The pallet positions colored most darkly are the most convenient. ........................................ 34

5.3 Flow-through and U-shaped layouts, with double-deep storage. When pallets are stored in lanes, as here, the first pallet position in a lane is more convenient than subsequent pallet positions in that lane. ........................................ 35

5.4 The footprint of a 2-deep lane. The footprint includes all the floor space required to support that lane, including the gap between lanes (to the left of the pallet) and one-half the aisle width (in front of the two pallet positions). ........................................ 38
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>Layout of a distributor of spare parts.</td>
<td>44</td>
</tr>
<tr>
<td>7.1</td>
<td>For the simplest case, assume that every sku is represented in the fast-pick area.</td>
<td>51</td>
</tr>
<tr>
<td>7.2</td>
<td>In general, some sku’s should not be represented in the fast-pick area.</td>
<td>57</td>
</tr>
<tr>
<td>7.3</td>
<td>The net benefit $c_i(v)$ realized by storing a sku as a function of the quantity stored. The net benefit is zero if sku $i$ is not stored in the forward area; but if too little is stored, restock costs consume any pick savings.</td>
<td>58</td>
</tr>
<tr>
<td>7.4</td>
<td>The majorizing function $\bar{c}_i(v)$ is linear in the interval $[0, 2r_i f_i / \mu]$.</td>
<td>59</td>
</tr>
<tr>
<td>7.5</td>
<td>Optimal Allocations have less variability in volume than Equal Time Allocations and less variability in number of restocks than Equal Space Allocations.</td>
<td>62</td>
</tr>
<tr>
<td>7.6</td>
<td>The net benefit realized from the fast-pick area in the warehouse of a major telecommunications company. The three graphs show the net benefit realized by storing the top $j$ sku’s in the forward area for three different ways of allocating space among the sku’s. The greatest benefit is realized by ranking sku’s according to $p_i / \sqrt{f_i}$ and allocating optimal amounts of space (the solid line on the graph); next most benefit is realized by ranking according to $p_i / v_i$ and allocating equal time supplies (the dotted dark gray line); and least net benefit is realized in sorting by $p_i / v_i$ and allocating equal amounts of space (the medium gray line).</td>
<td>63</td>
</tr>
<tr>
<td>7.7</td>
<td>Example of slotting that accounts for the geometry of the sku’s and the storage mode. This pattern of storage minimizes total pick and restock costs for this set of sku’s and this arrangement of shelves. Numbers on the right report the percentage of picks from this bay on each shelf. (This was produced by software written by the authors.)</td>
<td>68</td>
</tr>
<tr>
<td>7.8</td>
<td>Multiple fast-pick areas, such as flow rack and carousel, each with different economics</td>
<td>70</td>
</tr>
<tr>
<td>7.9</td>
<td>Net benefit of storing various full-pallet quantities of a sku in the fast-pick area.</td>
<td>72</td>
</tr>
<tr>
<td>8.1</td>
<td>Storing an item so that it blocks access to another creates useless work and increases the chance of losing product.</td>
<td>82</td>
</tr>
<tr>
<td>8.2</td>
<td>Each storage unit may be placed in any of up to six orientations; and each lane might as well be extended to the fullest depth possible and to the fullest height.</td>
<td>82</td>
</tr>
<tr>
<td>8.3</td>
<td>Each lane should contain only a single sku, with each case oriented identically; and the lane should be extended as deep and as high as the shelf allows.</td>
<td>83</td>
</tr>
<tr>
<td>9.1</td>
<td>A simple flow line in which each item requires processing on the same sequence of work stations.</td>
<td>88</td>
</tr>
<tr>
<td>9.2</td>
<td>Positions of the workers after having completed $k$ products.</td>
<td>90</td>
</tr>
</tbody>
</table>
9.3 A time-expanded view of a bucket brigade production line with three workers sequenced from slowest to fastest. The solid horizontal line represents the total work content of the product and the solid circles represent the initial positions of the workers. The zigzag vertical lines show how these positions change over time and the rightmost spikes correspond to completed items. The system quickly stabilized so that each worker repeatedly executes the same portion of work content of the product. ................................ 91

9.4 Average pick rate as a fraction of the work-standard. Zone-picking was replaced by bucket brigade in week 12. (The solid lines represent best fits to weekly average pick rates before and after introduction of the bucket brigade protocol.) ....................... 95

9.5 Distribution of average pick rate, measured in 2-hour intervals before (clear bars) and after bucket brigades (shaded bars). Under bucket brigades production increased 20% and standard deviation decreased 20%.) ................................ 96

10.1 The horizontal axis lists the SKU’s by frequency of requests (picks). The top curve shows the percentage of picks within the most popular SKU’s and the lower curve shows the percentage of orders completed within the most popular SKU’s. ........................................ 107

10.2 A bird’s eye view of a warehouse, with each section of shelf colored in proportion to the frequency of requests for the SKU’s stored therein. (The arrows indicate the default path followed by the order-pickers.) 108

10.3 Very few customers order more than ten line-items (left), but this accounts for the bulk of the picking (right). ................ 109

12.1 A typical LTL freight terminal. Shaded rectangles represent trailers to be unloaded; clear rectangles represent destination trailers to be filled. 122

12.2 There is less floor space per door at outside corners and therefore more likely to be congestion that retards movement of freight. .......... 124

12.3 For data representative of The Home Depot an L-shaped dock was never preferable to an I-shaped dock, no matter the size. A T-shaped dock begins to dominate for docks of about 170 doors and a +-shaped dock around 220 doors. ...................... 127
List of Tables

7.1 Pick density and its surrogate under three different schemes for allocating space ........................................ 63

10.1 Top ten items of a chain of retail drug stores, as measured in number of cases moved .................................. 101
10.2 Top ten items of a chain of retail drug stores, as measured by the number of customer requests (picks) .......... 101
10.3 Top ten items of a chain of retail drug stores, as measured by the number of pieces sold. ............................ 101
10.4 Top ten office products measured by customer requests ............................................................ 102
10.5 Top ten office products by weight, a measure of shipping costs ......................................................... 102
10.6 Top ten items moving from a grocery warehouse, as measured by number of cases .................................... 115
Preface

0.1 Why this book

The focus of this book is the science of how to design, operate and manage warehouse facilities. We say “science” because we emphasize the building of detailed mathematical and computer models that capture the detailed economics of the management of space and labor. The goal of our approach is to give managers and engineers the tools to make their operations successful.

0.2 Organization

- We begin with a brief discussion of material flow and provide an simple, aggregate way of viewing it. This “fluid model” provides useful insights in the large.

- Next we give an overview of warehouse operations: Typical kinds of warehouses; how they contribute to the operations of a business; what types of problems a warehouse faces; what resources a warehouse can mobilize to solve those problems; and some simple tools for analysis.

- We discuss the essential statistics by which the operations of a warehouse are to be understood.

- We survey typical equipment used in a warehouse and discuss the particular advantages and disadvantages of each.

- We then look at a particularly simple type of warehouse: A “unit-load” warehouse in which the sku’s arrive on pallets and leave on pallets. In such warehouses, the storage area is the same as the picking area.

- We move to more complicated warehouses in which most sku’s arrive packaged as pallets and leave as cases. In such operations the storage area is frequently the same as the picking area.

- Next we examine high-volume, labor-intensive warehouses, such as those supporting retail stores: Sku’s leave as pieces that are generally part of large orders.
and orders are assembled as on an assembly line. It is frequently the case that there is a separate picking area that is distinct from the bulk storage area.

- This chapter may be seen as a companion to the preceding one. We explain a new style of order-picking that has many advantages in a high-volume operation. It is unusual in that it balances itself to eliminate bottlenecks and so increase throughput.

- We study operations of a crossdock, which is a kind of high-speed warehouse in which product flows with such velocity that there is no point in bothering to put it on a shelf. Occupancy times are measured in hours rather than days or weeks.

- We show how to compare the performance of two different warehouses, even if they are of different sizes or serve different industries.

- Finally, we close with a collection of case studies to show the application of warehouse science in real life.

### 0.3 Resources

We support this book through our web site [www.warehouse-science.com](http://www.warehouse-science.com), from which you may retrieve the latest copy — at the time of this writing we are revising the book about four times a year — and find supporting materials, such as photographs and data sets, to which we are adding continuously.

The photographs and data sets are from our consulting and from generous companies, whom we thank. At their request, some company identities have been disguised and/or some aspect of the data cloaked. (For example, in some cases we have given the products synthetic names to hide their identities.)

### 0.4 But first...

A few friends and colleagues have contributed so much to the work on which this book is based that they deserve special mention. We particularly thank Don Eisenstein of the University of Chicago (Chapter 9), Paul Griffin of Georgia Tech (Chapter 11), Kevin Gue of the US Naval Postgraduate School (Chapter 12), Ury Passy of The Technion (Chapter 7), and Loren “Dr. Dunk” Platzman of On Technology (Chapter 7).
Chapter 1

Introduction

Why have a warehouse at all? A warehouse requires labor, capital (land and storage-and-handling equipment) and information systems, all of which are expensive. Is there some way to avoid the expense? For most operations the answer is no. Warehouses, or their various cousins, provide useful services that are unlikely to vanish under current economic scene. Here are some of their uses:

To consolidate product to reduce transportation costs and to provide customer service.

There is a fixed cost any time product is transported. This is especially high when the carrier is ship or plane or train; and to amortize this fixed cost it is necessary to fill the carrier to capacity. Consequently, a distributor may consolidate shipments from vendors into large shipments for downstream customers. Similarly, when shipments are consolidated, then it is easier to receive downstream. Trucks can be scheduled into a limited number of dock doors and so drivers do not have to wait. The results are savings for everyone.

Consider, for example, Home Depot, where a typical store might carry product from hundreds of vendors but has only three receiving docks. A store receives shipments at least weekly and so for many products, the quantities are insufficient to fill a truck. This means that most shipments from a vendor to a store must be by less-than-truck-load (LTL) carrier, which is considerably more expensive than shipping a full truck load. But while most stores cannot fill a truck from a vendor, all the Home Depot stores in aggregate certainly can, and several times over. The savings in transportation costs alone is sufficient to justify shipping full truck loads from vendors to a few consolidation centers, which then ship full truck loads to many, many stores.

To realize economies of scale in manufacturing or purchasing.

Vendors may give a price break to bulk purchases and the savings may offset the expense of storing the product. Similarly, the economics of manufacturing may dictate large batch sizes to amortize large setup costs, so that excess product must be stored.
To provide value-added processing: Increasingly, warehouses are being forced to assume value-added processing such as light assembly. This is a result of manufacturing firms adopting a policy of postponement of product differentiation, in which the final product is configured as close to the customer as possible. Manufacturers of personal computers have become especially adept at this. Generic parts, such as keyboards, disk drives, and so on, are shipped to a common destination and assembled on the way, as they pass through a warehouse or the sortation center of a package carrier. This enables the manufacturer to satisfy many types of customer demand from a limited set of generic items, which therefore experience a greater aggregate demand, which can be forecast more accurately. Consequently safety stocks can be lower. In addition, overall inventory levels are lower because each item moves faster.

Another example is pricing and labeling. The state of New York requires that all drug stores label each individual item with a price. It is more economical to do this in a few warehouses, where the product must be handled anyway, than in a thousand retail stores, where this could distract the workers from serving the customer.

To reduce response time: For example, seasonalties strain the capacity of a supply chain. Retail stores in particular face seasonalties that are so severe that it would be impossible to respond without having stockpiled product. For example, Toys R Us does, by far, most of its business in November and December. During this time, their warehouses ship product at a prodigious rate (some conveyors within their new warehouses move at up to 35 miles per hour!). After the selling season their warehouses spend most of their time building inventory again for the following year.

Response-time may also be a problem when transportation is unreliable. In many parts of the world, the transportation infrastructure is relatively undeveloped or congested. Imagine, for example, shipping sub-assemblies to a factory in Ulan Bator, in the interior of Asia. That product must be unloaded at a busy port, pass through customs, and then travel by rail, and subsequently by truck. At each stage the schedule may be delayed by congestion, bureaucracy, weather, road conditions, and so on. The result is that lead time is long and variable. If product could be warehoused in Shanghai, it could be shipped more quickly, with less variance in lead time, and so provide better customer service.

While there are many reasons why warehouse facilities provide value in the supply chain, one of the main themes of this book is that there is a systematic way to think about a warehouse system regardless of the industry in which it operates. As we shall show, the selection of equipment and the organization of material flow are largely determined by

- Inventory characteristics, such as the number of products, their sizes, and turn rates;
- Throughput requirements, including the number of lines and orders shipped per day;
• The footprint of the building.


Chapter 2

Material flow

The “supply chain” is the sequence of processes through which product moves from its origin toward the customer. In our metaphor of fluid flow we may say that warehouses represent storage tanks along the pipeline.

The analogy with fluid flows can also convey more substantial insight. For example, consider a set of pipe segments of different diameters that have been joined in one long run. We know from elementary fluid dynamics that an incompressible fluid will flow faster in the narrower segments of pipe than in the wider segments. This has meaning for the flow of product: The wider segments of pipe may be imagined to be parts of the supply chain with large amounts of inventory. On average then, an item will move more slowly through the region with large inventory than it will through a region with little inventory.

The fluid model immediately suggests other general guidelines to warehouse design and operation, such as:

- Keep the product moving; avoid starts and stops, which mean extra handling and additional space requirements.
- Avoid layouts that impede smooth flow.
- Identify and resolve bottlenecks to flow.

Later we shall rely on the fluid model to reveal more profound insights.

It is worth remarking that the movement to “just-in-time” logistics is roughly equivalent to reducing the diameter of the pipe, which means product flows more quickly and so flow time and in-transit inventory are reduced.

Even though it is a frequently useful metaphor, most products do not, of course, flow like incompressible fluids. Instead, they flow more like sand or gravel or rocks or even boulders. In other words, the product is not infinitely divisible but rather is granular.

A “stock keeping unit” is the smallest physical unit of a product that is tracked by an organization. For example, this might be a box of 100 Gem Clip brand paper clips. In this case the final customer will use a still smaller unit (individual paper clips), but the supply chain never handles the product at that tiny scale.
Upstream in the supply chain, product generally flows in larger units, such as pallets; and is successively broken down into smaller units as it moves downstream, as suggested in Figure 2.1. Thus a product might move out of the factory and to regional distribution centers in pallet-loads; and then to local warehouses in cases; and finally to retail stores in inner-packs or even individual pieces, which are the smallest units offered to the consumer. This means that our fluid model will be most accurate downstream, where smaller units are moved.

2.1 The warehouse as a queueing system

A queueing system is a model of the following structure: Customers arrive and join a queue to await service by any of several servers. After receiving service the customers depart the system.

A fundamental result of queueing theory is known as Little’s Law, after the man who provided the first formal proof of a well-known piece of folk-wisdom.

**Theorem 2.1 (Little’s Law)** For a queueing system in steady state the average length of the queue equals the average arrival rate times the average waiting time. More succinctly:

\[ L = \lambda W. \]

A warehouse may be roughly modeled as a queueing system in which sku’s are the customers that arrive at the receiving dock, where they join a queue (that is, are stored in the warehouse) to wait for service (shipping). If the warehouse is at steady state the product will be shipped at the same average rate at which it arrives. Then Little’s Law applies and the average amount of product in the warehouse equals the arrival rate of product multiplied by the average time product is resident in our warehouse.

Here is an example of how we can use Little’s Law to tease out information that might not be immediately apparent. Consider a warehouse with about 10,000 pallets in residence and that turn an average of about 4 times a year. Is the labor force sufficient to support this? By Little’s Law:

\[ 10,000 \text{ pallets} = \lambda (1/4 \text{ year}). \]

so that

\[ \lambda \approx 40,000 \text{ pallets/year}. \]

Assuming one 8-hour shift per day and about 250 working days per year, there are about 2,000 working hours per year, which means that

\[ \lambda \approx 20 \text{ pallets/hour}. \]

Notice what we have just done! From a simple count of pallets together with an estimate of the number of inventory turns per year we estimated the labor requirements.
2.1. THE WAREHOUSE AS A QUEUEING SYSTEM

Figure 2.1: A product is generally handled in smaller units as it moves down the supply chain. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).
2.1.1 Extensions

What makes Little’s Law so useful is that it continues to hold even when there are many types of customers, with each type characterized by its own arrival rate $\lambda_i$, waiting time $W_i$, and queue length $L_i$. Therefore the law may be applied to a single SKU, to a family of SKU’s, to an area within a warehouse, or to an entire warehouse.

One way of understanding Little’s Law is as a simple identity of accounting. Divide a fixed period of time into $n$ equally spaced intervals. Let $A(n)$ denote the total number of arrivals during this period of time, and let $T_j$ denote the time in the system of the $j^{th}$ arrival (assumed to be an integer number of periods). Arrivals occur only at the beginning of a period. Let $I_i$ denote the inventory in the system at time $i$. Assume for now that $I_0 = I_n = 0$. If each arrival (customer) must pay 1 dollar per period as rent at the end of each period of stay, how much money does the system collect? On the one hand customer $j$ pays $T_j$ dollars, and so the answer is $\sum_{j=1}^{A(n)} T_j$. On the other hand, the system collects $I_i$ dollars each period, and so the answer must also be $\sum_{i=1}^{n} I_i$. Therefore,

$$\sum_{i=1}^{n} I_i = \sum_{j=1}^{A(n)} T_j,$$

or, equivalently,

$$\frac{\sum_{i=1}^{n} I_i}{n} = \left(\frac{\sum_{j=1}^{A(n)} T_j}{A(n)}\right) \left(\frac{A(n)}{n}\right).$$

Equation (2.2) may be interpreted as Little’s Law with arrival rate $A(n)/n$ and average time in the system $\sum_{j=1}^{A(n)} T_j/A(n)$. In a real system we cannot rely on $I_0 = I_n = 0$ and the true amount of money collected would have to be adjusted by adding the amount of money collected from the initial customers in the system at time 0, and substracting out the amount of money we have yet to collect from those new arrivals who have not yet finished their service by time $n$. These adjustments will be divided by $A(n)$ in (2.2), and so if bounded will go to zero as $n$ and therefore $A(n)$ get large.
2.2 Questions

**Question 2.1** What are the five typical physical units-of-measure in which product is handled in a warehouse? For each unit-of-measure, state whether there are any standardized dimensions and, if so, identify them.

**Question 2.2** Your third-party warehouse has space available for 10,000 pallets and you have 20 forklift operators per 8-hour day for 250 working days a year. If the average trip from receiving to storage to shipping is 10 minutes, how many inventory turns a year could you support for a full warehouse?

**Question 2.3** Your third-party warehouse is bidding for a contract to store widgets as they are manufactured. However, widgets are perishable and should be turned an average of six times per year. The manufacturer produces at an average rate of 32 pallets per day. How many pallet positions should you devote to widgets to ensure that widgets turn as required.

**Question 2.4** Imagine that you are a consultant who tours a pallet storage facility. You estimate that 10,000 pallets are in storage, serviced by 7 forklift operators per day. The average forklift travel time from receiving to a storage location and then to shipping is about 6 minutes. Estimate the inventory turns per year.
Chapter 3

Warehouse operations

A warehouse reorganizes and repackages product. Product typically arrives packaged on a larger scale and leaves packaged on a smaller scale. In other words, an important function of this warehouse is to break down large chunks of product and redistribute it in smaller quantities. For example, some sku’s may arrive from the vendor or manufacturer in pallet quantities but be shipped out to customers in case quantities; other sku’s may arrive as cases but be shipped out as eaches; and some very fast-moving sku’s may arrive as pallets and be shipped out as eaches.

In such an environment the downstream warehouse operations are generally more labor-intensive.

This is still more true when product is handled as eaches. In general, the smaller the handling unit, the greater the handling cost. Think of it: Moving 10,000 boxes of paper clips is terribly expensive when each box must be picked separately, as they may when, for example, supplying retail stores. Much less labor is required to move those 10,000 boxes if they are packaged into cases of 48 boxes; and still less labor if those cases are stacked 24 to a pallet.

Even though warehouses can serve quite different ends, most share the same general pattern of material flow. Essentially, they receive bulk shipments, stage them for quick retrieval; then, in response to customer requests, retrieve and sort sku’s, and ship them out to customers.

The reorganization of product takes place through the following processes.

- Inbound processes
  - Receiving
  - Put-away

- Outbound processes
  - Processing customer orders
  - Order-picking
  - Checking
- Packing
- Shipping

A general rule is that product should, as much as possible, flow continuously through this sequence of processes. Each time it is put down means that it must be picked up again sometime later, which is double-handling. When such double-handling is summed over all the tens-of-thousands of sku’s and hundreds-of-thousands of pieces and/or cases in a warehouse, the cost can be considerable.

Another rule is that product should be scanned at all key decision points to give “total visibility of assets”, which enables quick and accurate response to customer demand.

3.1 Receiving

Receiving may begin with advance notification of the arrival of goods. This allows the warehouse to schedule receipt and unloading to coordinate efficiently with other activities within the warehouse. It is not unusual for warehouses to schedule trucks to within 30-minute time windows.

Once the product has arrived, it is unloaded and possibly staged for putaway. It is likely to be scanned to register its arrival so that ownership is assumed and so that it is known to be available to fulfill customer demand. Product will be inspected and any exceptions noted, such as damage, incorrect counts, wrong descriptions, and so on.

Product typically arrives in larger units, such as pallets, from upstream and so labor requirements are not usually great. Accordingly, this accounts for only about 10% of operating costs.

3.2 Put-away

Before product can be put away, an appropriate storage location must be determined. This is very important because where you store the product determines to a large extent how quickly and at what cost you later retrieve it for a customer. This requires managing a second inventory, not of product, but of storage locations. You must know at all times what storage locations are available, how large they are, how much weight they can bear, and so on.

When product is put away, the storage location should also be scanned to record where the product has been placed. This information will subsequently be used to construct efficient pick-lists to guide the order-pickers in retrieving the product for customers.

Put-away can require a fair amount of labor because product may need to be moved considerable distance to its storage location. Put-away typically accounts for about 15% of warehouse operating expenses.
3.3 Process customer orders

On receipt of customer orders, the warehouse must perform checks such as to verify that inventory is available to ship; and it may need to coordinate order fulfillment with other sites. Then the warehouse must produce pick lists to guide the order-picking. Finally, it must produce any necessary shipping documentation and it must schedule the order-picking and shipping.

These activities are typically accomplished by a warehouse management system, a large software system that coordinates the activities of the warehouse.

3.4 Order-picking

Order-picking typically accounts for about 55% of warehouse operating costs; and order-picking itself may be further broken like this [9]:

<table>
<thead>
<tr>
<th>Activity</th>
<th>% Order-picking time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traveling</td>
<td>55%</td>
</tr>
<tr>
<td>Searching</td>
<td>15%</td>
</tr>
<tr>
<td>Extracting</td>
<td>10%</td>
</tr>
<tr>
<td>Paperwork and other activities</td>
<td>20%</td>
</tr>
</tbody>
</table>

Notice that traveling comprises the greatest part of the expense of order-picking, which is itself the most expensive part of warehouse operating expenses. Much of the design of the order-picking process is directed to reducing this unproductive time.

In manual order-picking each worker is given a pick sheet, which lists the sku’s to be picked, in what amounts, and where they are to be found. The sku’s are listed in the order in which they will normally be encountered as the picker moves through the warehouse.

Each entry on the pick sheet is referred to as a line or line-item or pick-line (because it corresponds to one printed line of the sheet). Alternative terminology includes pick or visit or request. Note that a pick (line) may require more than one grab if, for example, several items of a sku are to be retrieved for an order.

Broken-case picking is labor-intensive and resistant to automation because of the variety of sku’s to be handled. In contrast, case-picking can sometimes be automated because of the relative uniformity of cases, which are almost always rectangular. The labor cost to retrieve sku’s is composed of time to travel to the storage location plus time spent retrieving the sku (reach/grab/place). Travel time generally represents over half of the labor in order-picking.

The pick face is that 2-dimensional surface, the front of storage, from which sku’s are extracted. This is how the sku’s are presented to the order picker. In general, the more different sku’s presented per area of the pick face, the less travel required per pick.

The pick density (#-picks per foot of travel) of an order tells us roughly how efficiently that order can be retrieved. An order that requires many picks per foot of aisle is relatively economical to retrieve: we are paying only for the actual cost of retrieval
and not for walking. On the other hand, small orders that are widely dispersed may be expensive to retrieve because there is more walking per pick.

Pick density depends on the orders and so we cannot know it precisely in advance of order receipt. However, it is generally true that pick density can be improved by ensuring high sku density, which is number of sku’s per foot of travel.

Pick density can be increased, at least locally, by storing the most popular sku’s together. Then order-pickers can make more picks in a small area, which means less walking.

Another way to increase the pick density is to batch orders; that is, have each worker retrieve many orders in one trip. However, this requires that the items be sorted into orders either while picking or else downstream. In the first case, the pickers are slowed down because they must carry a container for each order and they must sort the items as they pick, which is time-consuming and can lead to errors. If the items are sorted downstream, space and labor must be devoted to this additional process. In both cases even more work and space may be required if, in addition, the orders themselves must be sorted to arrive at the trailer in reverse sequence of delivery.

It is generally economic to batch single-line orders. These orders are easy to manage since there is no need to sort while picking and they can frequently be picked directly into a shipping container.

Very large orders can offer similar economies, at least if the sku’s are small enough so that a single picker can accumulate everything requested. A single worker can pick that order with little walking per pick and with no sortation.

The challenge is to economically pick the orders of intermediate size; that is, more than two pick-lines but too few to sufficiently amortize the cost of walking. Roughly speaking, it is better to batch orders when the costs of work to separate the orders and the costs of additional space are less than the extra walking incurred if orders are not batched. It is almost always better to batch single-line orders because no sortation is required. Very large orders do not need to be batched because they will have sufficient pick-density on their own. The problem then is with orders of medium-size.

To sustain order-picking product must also be replenished. Restockers move larger quantities of sku and so a few restockers can keep many pickers supplied. A rule of thumb is one restocker to every five pickers; but this will depend on the particular patterns of flow.

A restock is more expensive than a pick because the restocker must generally retrieve product from bulk storage and then prepare each pallet or case for picking. For example, he may remove shrink-wrap from a pallet so individual cases can be retrieved; or he may cut individual cases open so individual pieces can be retrieved.

### 3.4.1 Sharing the work of order-picking

A customer order may be picked entirely by one worker; or by many workers but only one at a time; or by many at once. The appropriate strategy depends on many things, but one of the most important is how quickly must orders flow through the process. For example, if all the orders are known before beginning to pick, then we can plan efficient picking strategies in advance. If, on the other hand, orders arrive in real time
and must be picked in time to meet shipping schedules then we have little or no time in which to seek efficiencies.

A general decision to be made is whether a typical order should be picked in serial (by a single worker at a time) or in parallel (by multiple workers at a time). The general trade-off is that picking serially can take longer to complete an order but avoids the complications of coordinating multiple pickers and consolidating their work.

A key statistic is flow time: how much time elapses from the arrival of an order into our system until it is loaded onto a truck for shipping? In general, it is good to reduce flow time because that means that orders move quickly through our hands to the customer, which means increased service and responsiveness.

A rough estimate of the total work in an order is the following. Most warehouses track picker productivity and so can report the average picks per person-hour. The inverse of this is the average person-hours per pick and the average work per order is then the average number of pick lines per order times the average person-hours per pick. A rough estimate of the total work to pick the sku’s for a truck is the sum of the work-contents of all the orders to go on the truck. This estimate now helps determine our design: How should this work be shared?

- If the total work to pick and load a truck is small enough, then one picker may be devoted to an entire truck. This would be a rather low level of activity for a commercial warehouse.

- If the total work to pick and load an order is small enough, then we might repeatedly assign the next available picker to the next waiting order.

- If the orders are large or span distant regions of the warehouse or must flow through the system very quickly we may have to share the work of each order with several, perhaps many, pickers. This ensures that each order is picked quickly; but there is a cost to this: Customers typically insist on shipment integrity, which means that they want everything they ordered in as few packages as possible, to reduce their shipping costs and the handling costs they incur on receipt of the product. Consequently, we have to assemble the various pieces of the order that have been picked by different people in different areas of the warehouse; and this additional process is labor-intensive and slow or else automated.

- For warehouses that move a lot of small product for each of many customers, such as those supporting retail stores, order-picking may be organized as an assembly-line: The warehouse is partitioned into zones corresponding to workstations, pickers are assigned to zones, and workers progressively assemble each order, passing it along from zone to zone.

  Advantages include that the orders emerge in the same sequence they were released, which means you make truck-loading easier by releasing orders in reverse order of delivery. Also, order-pickers tend to concentrate in one part of the warehouse and so are able to take advantage of the learning curve.

  The problem with zone-picking is that it requires all the work of balancing an assembly line: A work-content model and a partition of that work. Typically this is done by an industrial engineer.
Real warehouses tend to use combinations of several of these approaches.

### 3.5 Checking and packing

Packing can be labor-intensive because each piece of a customer order must be handled; but there is little walking. And because each piece will be handled, this is a convenient time to check that the customer order is complete and accurate. Order accuracy is a key measure of service to the customer, which is, in turn, that on which most businesses compete.

Inaccurate orders not only annoy customers by disrupting their operations, they also generate returns; and returns are expensive to handle (up to ten times the cost of shipping the product out).

One complication of packing is that customers generally prefer to receive all the parts of their order in as few containers as possible because this reduces shipping and handling charges. This means that care must be taken to try to get all the parts of an order to arrive at packing together. Otherwise partial shipments must be staged, waiting completion before packing, or else partial orders must be packaged and sent.

Amazon, the web-based merchant, will likely ship separate packages if you order two books fifteen minutes apart. For them rapid response is essential and so product is never staged. They can ship separate packages because their customers do not mind and Amazon is willing to pay the additional shipping as part of customer service.

Packed product may be scanned to register the availability of a customer order for shipping. This also begins the tracking of the individual containers that are about to leave the warehouse and enter the system of a shipper.

### 3.6 Shipping

Shipping generally handles larger units that picking, because packing has consolidated the items into fewer containers (cases, pallets). Consequently, there is still less labor here. There may be some walking if product is staged before being loaded into freight carriers.

Product is likely to be staged if it must be loaded in reverse order of delivery or if shipping long distances, when one must work hard to completely fill each trailer. Staging freight creates more work because staged freight must be double-handled.

The trailer is likely to be scanned here to register its departure from the warehouse. In addition, an inventory update may be sent to the customer.

### 3.7 Summary

Most of the expense in a typical warehouse is in labor; most of that is in order-picking; and most of that is in travel.
3.8  More

Many warehouses also must handle returns, which run about 5% in retail. This will become a major function within any warehouse supporting e-commerce, where returns run 25–30%, comparable to those supporting catalogue sales.

Another trend is for warehouses to assume more value-added processing (VAP, which is additional work beyond that of building and shipping customer orders. Typical value-added processing includes the following:

- Ticketing or labeling (For example, New York state requires all items in a pharmacy to be price-labeled and many distributors do this while picking the items in the warehouse.)
- Monogramming or alterations (For example, these services are offered by Lands End, a catalogue and e-mail merchant of clothing)
- Repackaging
- Kitting (repackaging items to form a new item)
- Postponement of final assembly, OEM labeling (For example, many manufacturers of computer equipment complete assembly and packaging in the warehouse, as the product is being packaged and shipped.)
- Invoicing

Such work may be pushed on warehouses by manufacturers upstream who want to postpone product differentiation. By postponing product differentiation, upstream distributors, in effect, see more aggregate demand for their (undifferentiated) product. For example, a manufacturer can concentrate on laptop computers rather than on multiple smaller markets, such as laptop computers configured for an English-speaking market and running Windows 2000, those for a German-speaking market and running Linux, and so on. This aggregate demand is easier to forecast because it has less variance (recall the Law of Large Numbers!), which means that less safety stock is required to guarantee service levels.

At the same time value-added processing is pushed back onto the warehouse from retail stores, where it is just too expensive to do. Both land and labor are typically more expensive at the retail outlet and it is preferable to have staff there concentrate on dealing with the customer.
3.9 Questions

Question 3.1 What are the basic inbound operations of a warehouse? What are the outbound operations? Which is likely to be most labor intensive and why?

Question 3.2 At what points in the path of product through a warehouse is scanning likely to be used and why? What is scanned and why?

Question 3.3 What is “batch-picking” and what are its costs and benefits?

Question 3.4 What are the issues involved in determining an appropriate batch-size?

Question 3.5 What are the costs and benefits of ensuring that orders arrive at the trailers in reverse sequence of delivery?

Question 3.6 Explain the economic forces that are pushing more value-added processing onto the warehouses.

Question 3.7 Why the receiving staff in most warehouses much smaller than the staff of order-pickers?
Chapter 4

Storage and handling equipment

There are many types of special equipment that have been designed to reduce labor costs and/or increase space utilization.

Equipment can reduce labor costs by

- Allowing many sku’s to be on the pick face, which increases pick density and so reduces walking per pick, which means more picks per person-hour
- Facilitating efficient picking and/or restocking by making the product easier to handle (for example, by presenting it at a convenient height and orientation).
- Moving product from receiving to storage; or from storage to shipping.

Equipment can increase space utilization by:

- Partitioning space into subregions (bays, shelves) that can be loaded with similarly-sized sku’s. This enables denser packing and helps make material-handling processes uniform.
- Making it possible to store product high, where space is relatively inexpensive.

4.1 Storage equipment

By “storage mode” we mean a region of storage or a piece of equipment for which the costs to pick from any location are all approximately equal and the costs to restock any location are all approximately equal.

Common storage modes include pallet rack for bulk storage, flow rack for high-volume picking, bin-shelving for slower picking, and carousels for specialized applications.
4.1.1 Pallet storage

On the large scale, there are standard sizes of packaging. Within the warehouse the largest unit of material is generally the pallet, which is a wood or plastic base that is 48 inches by 40 inches — but, as with all standards, there are several: most standard pallets are 48 inches along one dimension but the other may be 32 inches (for pallets that go directly from manufacturer onto retail display, such as in a Home Depot), 42 inches, or 48 inches (such as those for transport of 55-gallon steel drums).

A pallet allows forks from a standard forklift or pallet jack to be inserted on either of the 40 inch sides. Many pallets have more narrow “slots” on the 48 inch sides and can be lifted there only by fork lift. There is no standard height to a pallet.

The simplest way of storing pallets is floor storage, which is typically arranged in lanes. The depth of a lane is the number of pallets stored back-to-back away from the pick aisle. The height of a lane is normally measured as the maximum number of pallets that can be stacked one on top of each other, which is determined by pallet weight, fragility, number of cartons per pallet, and so on. Note that the entire footprint of a lane is reserved for a sku if any part of the lane is currently storing a pallet. This rule is almost always applied, since if more than one sku was stored in a lane, some pallets may be double-handled during retrieval, which could offset any space savings. Also, it becomes harder to keep track of where product is stored. For similar reasons, each column is devoted to a single sku.

This loss of space is called honeycombing.

Pallet rack is used for bulk storage and to support full-case picking (Figure 4.1). Pallet length and width are reasonably uniform and pallet rack provides appropriately-sized slots. The height of slots can be adjusted, however, as pallet loads can vary in height.

The advantage of rack storage is that each level of the rack is independently supported, thus providing much greater access to the loads, and possibly permitting greater stack height that might be possible in floor storage.

The most common types of rack storage are:

**Selective rack or single-deep rack** stores pallets one deep, as in Figure 4.1. Due to rack supports each pallet is independently accessible, and so any sku can be retrieved from any pallet location at any level of the rack. This gives complete freedom to retrieve any individual pallet but requires relatively more aisle space to access the pallets.

**Double-deep rack** essentially consists of two single-deep racks placed one behind the other, and so pallets are stored two deep. Due to rack supports each 2-deep lane is independently accessible, and so any sku can be stored in any lane at any level of the rack. To avoid double-handling, it is usual that each lane be filled with a single sku, which means that some pallet locations will be unoccupied whenever some sku is present in an odd number of pallets. Another disadvantage of deep lanes is that slightly more work is required to store and retrieve product. However, deep lanes have the advantage of requiring fewer aisles to access the pallets, which means that the warehouse can hold more product. A special truck is required to reach past the first pallet position.
Figure 4.1: Simple pallet rack. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).
Push-back rack. This may be imagined to be an extension of double deep rack to 3–5 pallet positions, but to make the interior positions accessible, the rack in each lane pulls out like a drawer. This means that each lane (at any level) is independently accessible.

Drive-In or drive-through rack allows a lift truck to drive within the rack frame to access the interior loads; but, again to avoid double-handling, all the levels of each lane must be devoted to a single sku. With drive-in rack the putaway and retrieval functions are performed from the same aisle. With drive-through rack the pallets enter from one end and of the lane and leave from the other, so that product can be moved according to a policy of First-In-First-Out (FIFO). Drive-in/through rack may be thought of as floor-storage for product that is not otherwise stackable. It does not enable the flexibility of access that other types of pallet rack achieve. In addition, there are some concerns; for example, in this rack each pallet is supported only by the edges, which requires that the pallets be strong. In addition, there is increased chance of accidents as a forklift driver goes deeper into the rack.

Pallet flow rack is deep lane rack in which the shelving is slanted and lines with rollers, so that when a pallet is removed, gravity pulls the remainder to the front. This enables pallets to be putaway at one side and retrieved from the other, which prevents storage and retrieval operations from interfering with each other. Because of weight considerations, storage depth is usually limited to about eight pallets. This type of rack is appropriate for high-throughput facilities.

To access the loads in rack storage (other than an AS/RS) some type of lift truck is required; and specialized racks may require specialized trucks. The most common type of lift trucks are:

Counterbalance lift truck is the most versatile type of lift truck. The sit-down version requires an aisle width of 12–15 feet, its lift height is limited to 20–22 feet, and it travels at about 70 feet/minute. The stand-up version requires an aisle width of 10–12 feet, its lift height is limited to 20 feet, and it travels at about 65 feet/minute.

Reach and double-reach lift truck is equipped with a reach mechanism that allows its forks to extend to store and retrieve a pallet. The double-reach truck is required to access the rear positions in double deep rack storage. Each truck requires an aisle width of 7–9 feet, their lift height is limited to 30 feet, and they travel at about 50 feet/minute. A reach lift truck is generally supported by “outriggers” that extend forward under the forks. To accommodate these outriggers, the bottom level of deep pallet rack is generally raised a few inches off the ground so that the outriggers can pass under.

Turret Truck uses a turret that turns 90 degrees, left or right, to putaway and retrieve loads. Since the truck itself does not turn within the aisle, an aisle width of only 5–7 feet is required, its lift height is limited to 40–45 feet, and it travels at about 75 feet/minute. Because this truck allows such narrow aisle, some kind of
guidance device, such as rails, wire, or tape, is usually required. It only operates within single deep rack and super flat floors are required, which adds to the expense of the facility. This type of truck is not easily maneuverable outside the rack.

**Stacker crane within an AS/RS** is the handling component of a unit-load AS/RS, and so it is designed to handle loads up to 100 feet high. Roof or floor-mounted tracks are used to guide the crane. The aisle width is about 6–8 inches wider than the unit load. Often, each crane is restricted to a single lane, though there are, at extra expense, mechanisms to move the crane from one aisle to another.

**4.1.2 Bin-shelving or static rack**

Simple shelving is the most basic storage mode and the least expensive (Figure 4.2). The shelves are shallow: 18 or 24 inches are typical, for example, but 36 inch deep shelf is sometimes used for larger cartons. Because the shelves are shallow, any significant quantity of a sku must spread out along the pick-face. This reduces sku-density and therefore tends to reduce pick density, increase travel time, and reduce picks/person-hour.

Sku’s which occupy more than one shelf of bin-shelving are candidates for storage in another mode that will enable greater sku-density.

A typical pick rate from bin-shelving is 50–100 picks/person-hour. (Of course this and the pick rates for any equipment depends on which sku’s are stored there.)

With bin-shelving, both picking and restocking must be done from the pick-face and so, to avoid interference, must be scheduled at different times. This can mean working an additional shift.

**4.1.3 Gravity flow rack**

Flow rack is a special type of shelving with shelves that are tilted, with rollers, to bring cases forward for picking (Figure 4.3). The shelves may be 3–10 feet deep. This means that only one case of a product need be on the pick face, which means that many sku’s can be available in a small area of pick-face. This means high sku-density, which tends to increase the pick-density, decrease travel, and increase picks/person-hour.

Frequently the picking from flow rack is accelerated by supporting technology such as a pick-to-light system, by which a centralized computer lights up signals at every location within a bay from which product is to be picked. After the worker picks the appropriate quantity, who pushes a button to signal the computer. There are several benefits of this: The order-picker is freed from handling a paper pick-list; he does not have to search for the next storage location; and because picking is guided, it is more accurate.

A typical pick rate from flow rack is about 150–500 picks/person-hour, but this varies widely.

Flow rack is restocked from the back, independently of picking, and so restocking never interferes with picking, as is the case for static shelving, for which picking and restocking must alternate.
Figure 4.2: Bin-shelving, or static rack. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).
4.1. STORAGE EQUIPMENT

Figure 4.3: Gravity flow rack. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8-17).

Figure 4.4: Profiles of three styles of carton flow rack. Each successive style takes more space but makes the cartons more readily accessible.
There are several subtypes of carton flow rack, as shown in Figure 4.4.

- Square front, vertical frame flow rack is suited to picking full cases, such as canned goods. (This is a specialized use, however, because it requires full-cases in and full-cases out, which suggests excessive handling.)

- Layback frame with straight shelves is suited to picking from open cases when those cases vary in size, such as health and beauty aids. This style of rack takes up more space than the vertical frame rack but it makes the cartons more readily accessible.

- Layback frame with front-tilted shelves is suited to picking from open cases when those cases are similar in size, such as liquor or books. This style of flow rack makes the cartons most accessible.

### 4.1.4 Carousels

Carousels are motorized, computer-controlled, independently-rotating aisles of shelving (see Figure 4.5). Because they carry product to the picker there is no need for the picker to walk along an aisle and so carousels can be packed tightly, which increases space utilization and also provides security for the product.
Carousels are almost always set up in pods of two or three, so that a picker can, in effect, walk down multiple aisles simultaneously. (A single carousel would convey no advantage because a worker could walk as fast down an aisle of shelving.)

Pick-rates vary from 80–200 picks/person-hour. The effective pick rate may be somewhat slower than one might expect because restocking must be interleaved with picking. This means that care must be taken in choosing the sku’s to be stored on carousels, because one must account not only for the rate at which they will be picked but also for the rate at which they must be restocked.

A disadvantage of carousels is that the pick rate cannot be significantly increased by adding more people, because generally only one picker can use the carousel at a time. This can reduce the ability of the warehouse to respond to surges in demand.

4.2 Conveyors

Main points:

- Conveyors change the economics of travel: Storage locations close to the conveyor are, in terms of labor, close to shipping.

- Conveyors partition the warehouse into zones. The track restricts movement of workers and product because it is hard to cross; and so create problems of balancing work among zones. To alleviate this, conveyors are run up high whenever possible.

- Issues: How many products are conveyable? What is capacity, especially surge capacity, of conveyor?

- Guidelines for layout: Store conveyable product far from shipping because it can travel “for free”. Reserve locations that are physically close to shipping for non-conveyables because they will have to be carried (for example, fork lift).

4.3 Sortation equipment

- Push

- Tilt-tray

4.4 Summary

The most common equipment is pallet rack, with corresponding trucks, for pallets; bin-shelving for slower-moving and/or small items; and gravity flow rack for cases of faster-moving items.
4.5 On the lighter side

When space is expensive economics says to store product high, which is cheaper than expanding the warehouse out. This has led to some pallet storage at dizzying heights. For example, three stories is not unusual. Automated storage and retrieval devices are generally used for dangerous heights but not always. We heard of an IBM site that relies on intrepid workers on person-aboard trucks that lift them high into the air. They are safely linked to their truck and so cannot fall; but the problem is how to get them down in case of accident. IBM requires that all order-pickers in this part of the warehouse be trained in rappelling!

We have heard some amusing stories about setting up carousels. The problem has usually been carelessness. The most dramatic one is of a site that was loading some thirty carousels with auto parts, which are high value, slow moving, and heavy. Especially heavy. As we heard it, each carousel was loaded one location at a time, top to bottom, then rotated one position, and so on. Of course this meant that, at some point, one long side of the current carousel was fully loaded and the other was completely empty. Inevitably, carousel number twenty-nine tipped, crashing into fully-loaded number 28, and so on in majestic, slow-motion disaster.

A final carousel story is from the designer of a software control system for a very fine operation. He was concerned that the hardware move quickly enough to keep order-pickers fully occupied and so rotated each carousel at the high end of the recommended velocity range. During trial runs he noticed some empty storage slots, which was not unusual before restocking; but he became alarmed when the empty slots began to increase quickly. It seems the boxes in which product was stored were just slick enough that, as the carousel rotated a shelf around the end, it (the box) might shoot off the carousel and go skidding across the warehouse floor!
4.6 Questions

**Question 4.1** Which types of pallet storage generally provide the most efficient use of floor space: floor storage or pallet rack? Explain.

**Question 4.2** What are the relative advantages and disadvantages of push-back rack compared to pallet rack? How do both compare to gravity flow rack?

**Question 4.3** Which type of storage generally makes pallets more accessible: Pallet flow rack or drive-in rack? Explain.

**Question 4.4** Consider two sku’s: One a small, slow-moving sku of which the warehouse has only a small amount; and the other a small, fast-moving sku of which the warehouse has considerable quantity. Which is a candidate for shelving and which for flow rack? Explain.

**Question 4.5 (Open-ended)** Roughly speaking, what types of sku’s should be stored in carousels: Large ones or small ones? Ones that move few cases or ones that move many cases? Ones that are infrequently requested or ones that are frequently requested?

**Question 4.6** Conveyable sku’s should be stored far from shipping; why?

**Question 4.7** Suppose you have many slow-moving sku’s in less-than-pallet quantities and these are shipped as cases. What arguments are there for storing them in carton flow rack?

**Question 4.8** What are the dimensions of a standard pallet? Go to the web and see how many other “standards” you can find.

**Question 4.9** Under what conditions might you prefer to store pallets with the shorter side on the pick face (that is, along the aisle)? When might you want to store pallets with the longer side facing the aisle?

**Question 4.10** In each case explain why the warehouse action described below is probably unwise.

- Picking product from a single carousel
- Storing cases in a gravity flow rack that is 0.5 meter deep
- Picking fast-moving product from static shelving
- Storing pallets in an aisle of single-deep rack that is free-standing (there is an aisle on either side of it)
- Storing product as free-standing eaches, with every piece lined up neatly
Chapter 5

Pallet operations

The simplest type of warehouse is a unit-load warehouse, which means that only a single, common “unit” of material is handled at a time. Such warehouses typically move only pallet quantities: pallets in and pallets out.

A typical unit-load warehouse is a 3rd party warehouse, which is a subcontractor to others for warehouse services. A 3rd party warehouse typically charges its customers for each pallet handled (received and later shipped); and rent for space occupied.

5.1 Labor

When a pallet arrives at the receiving dock, it is driven by a forklift driver to a storage location, where it resides until requested by the customer. Then a forklift driver moves it to a trailer on the shipping dock.

The warehouse pays its forklift drivers for person-hours but it bills its customers for two handles for each pallet (in/out); therefore the warehouse wants many handles/person-hour.

Because a forklift handles one pallet at a time, the variable labor cost it incurs can be estimated fairly accurately as the time it takes to drive a forklift from receiving to the storage location to shipping. (Differences in insertion/extraction times are generally small in comparison to differences in travel costs; therefore we treat insertion/extraction times as fixed costs, which may be ignored in deciding where to place sku’s.)

There are other elements of work that might justly be charged to this pallet but they are harder to know, such as the time to drive to the receiving location (deadheading, which is travel with empty forks). We do not know how long this will take because we do not know in advance where the forklift will be when requested. (Of course we can learn this after the fact by recording all forklift travel.)

In summary, we can increase our handles/person-hour by reducing the travel time from receiving to storage location to shipping; and we can do this by careful choice of storage location.
5.1.1 Storage locations

Where should each sku be stored? It is intuitive that the “most active” sku’s should be in the most convenient locations. Let’s see what that means. First: What is a “convenient” location?

“Convenient” locations

Each location will generate the following variable labor costs: Drive from receiving dock to location; and drive from location to shipping dock. Therefore to each location we can associate a total travel time, which can serve as a model of labor cost $c_i$ incurred by storing a sku at location $i$. In fact, we can approximate this by the distance $d_i$ from receiving to location to shipping. This cost is independent of what is stored in other locations and so, if location $i$ is visited $n_i$ times during the year, the annual labor cost will be proportional to

$$\sum_i d_i n_i. \quad (5.1)$$

Distances $d_i$ are determined by the layout of the warehouse and frequencies of visit $n_i$ are determined by the customer. But we can choose what to store where to minimize annual travel costs. Expression 5.1 is minimized by pairing the largest $n_i$ with the smallest $d_i$, which means we would like to visit most frequently those locations with smallest total travel $d_i$.

“Active” sku’s

From the expression 5.1 for the labor cost of a location we prefer that sku’s with a lot of movement per storage location be stored in the best locations. In other words, we want to identify those sku’s that generate the most frequent visits per storage location. In steady state,

$$\text{average visits per storage location} = \frac{\text{number of units shipped}}{\text{number of units in storage}}$$

Thus, for example, a sku that is stored in average quantity 5 and which moved 20 units last year would have generated about $20/5 = 4$ visits per storage location, which is more than a sku that moved 100 units but was stored in quantity 50.

So to minimize labor costs:

• Rank all the available pallet positions of the warehouse from least cost $c_i$ to greatest cost.

• Rank all sku’s from most to least turns.

• Move down the list, assigning the pallets of the next fastest-turning sku’s to the next best locations.

This analysis is based on a warehouse that is operating at approximately steady state. What if the system is far from steady state? We can still approximate our intuition
5.2 Location of receiving and shipping

The layout of the warehouse determines the cost associated with each storage location. For example, consider the layout of Figure 5.1, in which the receiving and shipping docks are located in the middle of opposite sides of the warehouse. Every pallet must travel across the warehouse, from receiving to storage to shipping; and there are many locations tied for best. In fact, all the storage locations along one side of an aisle are equally good.

Now imagine how the convenience of the storage locations changes if the shipping and receiving doors were both moved to the right. Then storage locations to the left would become less convenient and the locations on the right more convenient; but the quality of the very best locations would not improve, while the quality of the very worst locations would becomes strictly worse! The result is that the layout would be absolutely worse than if the shipping and receiving areas are located in the middle.

Another alternative is to move both receiving and shipping to the middle of the same side. This would induce the economic terrain shown in Figure 5.2, where the
Figure 5.2: The economic contours of a warehouse floor in which receiving and shipping are both located at the middle bottom. The pallet positions colored most darkly are the most convenient.

The figure on the left shows the economic contours of all space within the warehouse and the figure on the right shows how the pallet storage positions inherit those economics. Now the best storage locations (near (19,0)) are very convenient indeed because a location that is close to receiving would also be close to shipping. However, there are relatively few such prime locations; moreover, there are more inconvenient locations than before and the least convenient locations are even less convenient (near (0,38) and (38,38)).

Which layout is better? In this case, as with so many of the design decisions, it depends on the populations of sku’s passing through the warehouse. If there will be a small amount of very fast-moving sku’s, it may be more efficient to put receiving and shipping on the same side of the facility, because the savings from the few very convenient locations may offset any loss due to the greater number of less convenient locations.

Here are some characteristics of each type of layout:

- U-shaped or cross-docking configuration
  - Receiving and shipping located on same side of the warehouse
  - Makes the most convenient locations still more convenient, less convenient locations even worse.
  - Appropriate when product movement has strong ABC skew (that is, when very few sku’s account for most of the activity)
  - Provides dock flexibility for both shipping, receiving: If one experiences a surge of activity, can make use of additional doors from other function.
  - Permits more efficient use of fork lifts: When a forklift reports for an assignment, he may be given a putaway and a retrieval, matched to reduce deadheading.
5.2. LOCATION OF RECEIVING AND SHIPPING

Figure 5.3: Flow-through and U-shaped layouts, with double-deep storage. When pallets are stored in lanes, as here, the first pallet position in a lane is more convenient than subsequent pallet positions in that lane.

- Minimizes truck apron and roadway
- Allows expansion along other three sides of warehouse.

- Flow-through configuration
  - Receiving and shipping on opposite sides of the warehouse
  - Makes many storage locations of equal convenience.
  - Conservative design: More reasonably convenient storage locations but fewer that are very convenient.
  - More appropriate for extremely high volume.
  - Preferable when building is long and narrow
  - Limits opportunity for efficiencies by dual transactions

Another factor that affects the economic terrain is the depth of the storage lanes. For example, consider the layouts of a warehouse in which pallets are stored in “double-deep” lanes. Figure 5.3 shows flow-through and U-shaped layouts, but with pallets stored in lanes that can be accessed only from the aisle-end. This restriction on the accessibility of the pallets changes the economic contours because the first position within an aisle is (slightly) more convenient that subsequent positions in that lane. (For example, notice that the aisle-front positions at $x = 7$ are more convenient than those at $x = 6$; and those at $x = 14$ are more convenient than those at $x = 13$). This effect is still more significant when product is stored in deeper lanes, which has the effect of making the deeper pallet positions less convenient.

Most warehouses have parallel aisles aligned with the receiving and shipping docks.
5.3 Space

Recall that the second revenue source of the 3rd party warehouse is to charge rent by the pallet. But because the warehouse typically bills its own expenses by the square-foot (for example, rent of the building, climate control, cleaning, and so on), the warehouse naturally wants many pallet-positions per square-foot. It can achieve this in two ways: by taking advantage of vertical space and by using deep lanes.

5.3.1 Stack height

Two distinguishing characteristics of a sku are its height of its pallet and the level to which its pallets can be stacked. Pallets that are low, heavy, and fragile cannot be stacked high or at all and so, when it is placed on the floor, renders unusable all the space above. This may be avoided by installing pallet rack, so that each pallet occupies one pallet position independently of all other pallets, which may occupy positions above or below it.

Conversely, pallets that are low, light, and sturdy can be stacked high and so allow many pallet positions per square foot of floor space.

How can you decide how much pallet rack to purchase and what to store in it? The value of the rack to you will depend on the sizes and movement patterns of the particular sku’s that visit your warehouse. (You will recognize this as a common theme to much that we will discuss.)

For the moment, imagine that we are considering purchase of a particular type of pallet rack. (All dimensions are fixed; we are just trying to decide how much of it to purchase.) Suppose that we have $n_i$ pallets of sku $i$ and, for simplicity, these populations are fairly stable.

The idea is that sku $i$ should go into rack if it is beneficial to do so. We just have to be careful to say exactly what we mean by “beneficial”. Intuitively, we want the benefits we get from putting a sku in rack to justify the cost of the rack. What sorts of benefits might we get from putting sku’s in rack? Notice that for each sku, one of the following holds: Putting this sku in the rack

- would create additional pallet positions the value of these additional positions would justify the rack;
- would create additional positions but of insufficient value to justify the rack;
- would not create additional pallet positions;
- would result in a net loss of pallet positions.

Let us see how these situations can happen. Imagine pallet in floor storage being considered for movement to 3-high pallet rack. First, consider four pallets of a sku that cannot be stacked at all. If left in floor storage these must occupy four pallet positions, which render all the space above them unusable. If 3-high rack was installed in the same space, this would create $3 \times 4 = 12$ pallet positions, of which only 4 would be occupied by this sku, for a net gain of 8 pallet positions; and these positions are potentially revenue-generating.
5.3. SPACE

Now consider 30 pallets of a sku that stack 3-high to come within a four feet of the ceiling. Moving these pallets to rack will not create additional pallet locations; and in fact, the cross-beams of the rack and space above the pallets are likely to consume the four feet of empty space.

Finally, consider 12 pallets a sku that stack 4-high to come to within two feet of the ceiling. These can be stored in three stacks on the floor; but would require four floor positions if put in the rack. This sku would lose pallet positions if put in rack.

You can see photographs of each of these situations at http://www.isye.gatech.edu/~jjb/wh/book/pallets/projects/projects.html.

Notice that our analysis allows you to decide exactly how much rack is justified:

- For each sku, compute how many pallet locations would be created by moving it into rack of a given configuration.
- For each sku, compute the value of the created pallet locations.
- Move a sku into rack if the value it creates is sufficient to justify the rack.

Now you know how to evaluate a given rack configuration to determine the value to you. You can repeat the process on each of several alternative configurations and choose the one that is of greatest net value.

It is worth remarking that we determined value solely by space utilization; there may be other issues. For example, product is less vulnerable to damage in rack. We have seen pallets loaded with bags of industrial dyes that cost several hundreds of dollars each; and each bag is easily slit by a wayward forklift blade. This in itself may be sufficient risk to justify storage in rack.

5.3.2 Lane depth

Space for aisles cannot be used for storage and so is not directly revenue-generating. Consequently we prefer to reduce aisle space, so the width of an aisle is the smallest space sufficient for a forklift to insert or extract a pallet. This is because the space of an aisle is not directly productive; that is, it is not rent-producing because nothing can be stored there. Therefore, aisle widths are pre-determined but the layout of storage space between the aisles — that is, the lane depth — is a design issue. Should lanes be four pallets deep? Six? Ten? There are many issues to consider, but the most important one is effective utilization of space. For example, the double-deep layout of Figure 5.3 fits about 41% more pallet positions in the same floor area than does the single-deep layout of Figure 5.2; but is it a better layout? Are enough of the additional pallet positions usefully engaged? It is clear that there is a trade-off: The single-deep layout has eight aisles and provides 196 pallet storage locations, all of which are directly accessible, which means that they are available for reassignment as soon as the current pallet is shipped out. In contrast, the double-deep layout has only six aisles and provides 280 pallet storage locations — but only 140 of them are directly accessible. Moreover, the 140 that are directly accessible are not available for reuse until the interior pallet location in the same lane becomes available. So: Deeper lanes produce more pallet storage locations but they are possibly of diminishing value.
Figure 5.4: The footprint of a 2-deep lane. The footprint includes all the floor space required to support that lane, including the gap between lanes (to the left of the pallet) and one-half the aisle width (in front of the two pallet positions).

The first step to quantifying the tradeoff is to measure the footprint of a lane; that is, all the space that may fairly be charged to that lane. This includes not only the area of the lane immediately occupied by pallets but also the space to one side that separates this lane from its neighbor; and also the space in front of the lane, up to one-half of the aisle (Figure 5.4).

Let $w$ and $d$ be the (standard) pallet width and depth respectively; let the lanes be $c$ pallets deep, with gap $g$ between adjacent lanes; and $a$ is the distance from the top pallet position in the lane to the top pallet position of the lane on the opposite side of the aisle. The the footprint of the lane is given by

$$(w + g)(dx + a/2).$$

This footprint is entirely dedicated to a single sku to avoid double-handling pallets.

A deeper lane requires less aisle space per pallet location, but on average does not get as much use out of each pallet position. To see this, imagine a lane of four pallets, each filled with the same sku, which moves at the rate of one pallet per week. Consider the deepest pallet location in the lane; it will always be either occupied or else available for use. The penultimate pallet location will mostly be either occupied or available; but for one week it may be both unoccupied and unavailable. That is, pallets 1–3 will have been removed but the last pallet will remain and so the lane will not be generally available for another week. In general, the pallet positions closer to the aisle are occupied less and less, on average; and the front pallet position, the one with the most convenient location of any in the lane, will be occupied only about $1/x$ of the time, on average (assuming the product is withdrawn at a constant rate).

In contrast, a lane that is but one pallet deep becomes immediately available for reuse when that pallet departs.

To get maximum space efficiency, we should choose lane depths to minimize the time-averaged floor space required to store the sku's of our warehouse.

One strategy is to force the entire census of sku's to jointly choose a single lane depth. We can do this by the following simple model. Let there be $n$ sku’s, with $q_i$ pallets of sku $i$, stackable in columns $z_i$ high. Again assume that each product moves out of the warehouse at a constant rate.
Theorem 5.1 (Floor storage) To minimize the average space consumed per pallet, floor storage should be configured with lane depth of approximately

\[
\sqrt{\frac{a}{2dn} \sum_{i=1}^{n} \frac{q_i}{z_i}}
\]

(5.2)

Proof Let \( x \) be the lane-depth measured in number of pallet positions. Then in floor storage sku \( i \) is stored with \( z_i \) \( x \) pallets per lane and so occupies \( \lceil q_i / (z_i x) \rceil \) lanes. By assumption, each product moves out of the warehouse at a constant rate and so on average about half of each sku \( i \) is present in the warehouse, or about

\[
\left\lfloor \frac{q_i}{2z_i x} \right\rfloor \text{ lanes.}
\]

(5.3)

Note that we are justified in rounding up in Expression 5.3 because an entire lane is rendered unusable by other sku’s if any portion of it is occupied. Now we replace Expression 5.3 with the following approximation, which minimizes the maximum error:

\[
\frac{q_i}{2z_i x} + \frac{1}{2}.
\]

Multiplying by the footprint of a lane gives the average floor area occupied by the population of sku’s:

\[
\sum_{i=1}^{n} \left( \frac{q_i}{2z_i x} + \frac{1}{2} \right) (g + w) (dx + a / 2).
\]

(5.4)

The result then follows from setting the derivative of average floor space to zero and solving for optimal lane depth \( x \).

Example 5.1 Consider the following population of sku’s:

<table>
<thead>
<tr>
<th>Sku</th>
<th>( q_i )</th>
<th>( z_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>36</td>
<td>2</td>
</tr>
</tbody>
</table>

What is the optimal lane depth in floor storage if aisles are 15 feet across and the pallets are 48 inches deep and 42 inches wide? Use a gap between aisles of 1 foot.

Substituting into Expression 5.2 gives

\[
0.79 \sqrt{\frac{50}{3} + \frac{40}{4} + \frac{36}{2}} \approx 5.28
\]

(5.5)

so lanes should be about five deep.
We can perform similar analysis for some types of rack storage. Consider, for example, pallet flow rack configured to a common height of \( z \) pallet openings. The distinctive feature of pallet flow rack is that all the forwardmost pallets of the \( z \) levels are independently accessible. Suppose such a rack is configured to be uniformly \( z \) pallet openings high and that each opening holds at most one pallet of any sku. Then:

**Theorem 5.2 (Pallet flow rack)** To minimize the average space consumed per pallet, pallet flow rack should be configured with lane depth of approximately

\[
\sqrt{\frac{a}{2dh_i} \sum_{i=1}^{n} q_i}.
\] (5.6)

**Proof** Left as an exercise.

Note that the optimum lane depth of pallet flow rack is greater than that for floor storage. The reason is that pallets in flow rack are more accessible: The forward pallet at any level may be retrieved without having to move another pallet. In contrast, only the topmost pallet of the forwardmost column may be removed from floor storage.

This lack of independent retrieval means that floor storage wastes more unoccupied storage positions over time and so cannot sustain lanes that are as deep and therefore space-efficient as flow rack.

This realization also helps us understand the other storage modes, such as drive-in pallet rack. This allows the forward-most pallet to be retrieved and, in addition, the forward-most pallet below that may be retrievable as well. Thus drive-in rack provides greater accessibility than floor storage but not as great as flow rack and so we expect that its optimum lane depth is between those of floor storage and flow rack. More generally we may observe the following.

**Observation 5.1** The greater the accessibility of pallets, the deeper and more space-efficient may be the lanes.

Note that rack storage can improve space utilization in two ways compared to floor storage: First by allowing not-readily-stackable products to use overhead space; and second, because Expression 5.6 is never less than Expression 5.2, by using deeper lanes.

To return to floor storage, note that Theorem 5.1 also allows us to determine the preferred lane depth of each individual sku:

**Corollary 5.1** The preferred lane depth of an single sku, with \( q \) pallets stackable \( z \) high, is

\[
\sqrt{\frac{a \cdot q}{2d \cdot z}}.
\] (5.7)

It is, of course, impractical to configure storage so that every sku has a different lane depth; but this may make sense for a few very populous sku’s. Nevertheless, it is helpful to know the preferred depth of each sku because we can group sku’s with similar storage preferences together, so that, for example, sku’s that prefer 3-deep storage
might be grouped with sku’s that prefer 5-deep, and the whole lot of them put in 4-deep storage. We shall re-visit this with greater care shortly.

Now imagine that our warehouse has \( m \) regions, each with a fixed lane depth \( d_1 < d_2 < \cdots < d_m \). Each of \( n \) sku’s must select one lane depth. The cost function for each item \( i \) as a function of lane depth \( x \) may be expressed, as shown in Expression 5.4, in the general form \( f_i(x) = a_i x + b_i / x + c_i \), the parameters of which are all positive. For each sku \( i \) let \( x_i^* = \sqrt{b_i / a_i} \) denote the ideal lane depth. For simplicity in what follows we shall assume that the global minimum is never an integer. Relabel the sku’s in increasing size of their ideal lane depths. Set \( d_0 = 0 \) and \( d_{n+1} \) to an integer larger than \( x_i^* \). Then the following result says roughly that assigned lane depths should respect the order of ideal lane depths.

**Theorem 5.3** There exists an optimal solution \( \{d^*_i\} \) for which \( d^*_i \leq d^*_j \) if \( i \leq j \).

**Proof** For each sku \( i \) find the unique integer \( k_i \in \{0, 1, \ldots, n\} \) for which \( x_i^* \in I_i^* \equiv [d_{k_i}, d_{k_i+1}] \). Due to the convexity of each cost function \( d_i^* \) must coincide with one of the endpoints of \( I_i^* \). Now pick items \( i \) and \( j \) for which \( i < j \). Since \( x_i^* \leq x_j^* \), if \( I_i^* \neq I_j^* \), the result immediately follows. Consider now the case that \( I_i^* = I_j^* \). If \( d_j^* = d_{k_j+1} \), then the result obviously holds. If, on the other hand \( d_j^* = d_{k_j} \), then \( f_j(d_{k_j}) \leq f_j(d_{k_j+1}) \) or, equivalently, \( b_j / a_j \leq d_{k_j} / d_{k_j+1} \). As \( d_{k_j} = d_{k_j} \) and \( b_i / a_i \leq b_j / a_j \) we then have that \( f_i(d_{k_i}) \leq f_i(d_{k_i+1}) \) and the result follows in this case as well. \( \blacksquare \)

We mentioned that there are other issues to consider besides space utilization. An important one is the observation of the rule of “First In First Out” (FIFO). Some special types of rack, such as drive-through rack or pallet flow rack support FIFO; but otherwise FIFO can be guaranteed only to within the lane depth. Making lanes deeper might give better space utilization; but it reduces compliance with FIFO, which could be a problem for some sku’s, such as food products. In addition, deeper lanes can increase insert/extract times.

### 5.4 Summary

- Location of shipping and receiving and layout/orientation of storage helps determine which pallet positions are convenient and which are not.
- Arrange your warehouse so that the convenience of storage positions matches the velocity of the products. For example, when activity is concentrated within a few sku’s, it is better to put receiving and storage near each other, which concentrates convenience in a few storage positions.
- Once the layout is determined, store the fastest-moving sku’s in the most convenient positions.
- Deep-lane storage reduces aisle space but loses use of pallet positions due to honeycombing. The optimal lane depth balances these two losses to get the most (time-averaged) storage locations per square foot of floor space. Optimal lane depth are given by Theorems 5.1 and 5.2.
• It is possible to use detailed histories or forecasts of product movement to exactly optimize equipment layout and configuration. Consider doing this when space is very expensive.

5.5 More

Computer distributors tend to have pallets that are either high and light (a pallet of printers can be 7 feet high); or low and heavy (such as a pallet of software).
5.6 Questions

Question 5.1 What is a unit-load warehouse and why is it easier to layout than others?

Question 5.2 For each of the following, explain whether the description makes the sku more or less appropriate to be stored in pallet rack rather than stacked on the floor, and why.

1. A pallet of this sku is fragile.
2. A pallet of this sku is heavy.
3. Pallets of this sku can be stacked safely almost to the ceiling.
4. A pallet of this sku is strong and square.
5. There are never more than two pallets of this sku in the warehouse.
6. Each case of this sku is dense and so, to keep the pallet from being too heavy it is loaded only one meter high.

Question 5.3 Suppose that we want to store all the sku’s of each customer together in a pallet-in, pallet-out warehouse. The sku’s of which customer are candidates for the most convenient locations? Explain why your selection is correct.

- That customer who has the most pallets in the warehouse
- That customer whose pallets we receive and ship in the greatest quantities
- That customer whose product represents the greatest dollar-volume when we account for the value of the product
- That customer whose shipments into the warehouse are the largest
- That customer whose shipments out of the warehouse are the largest
- It is impossible to tell
- None of the above

Question 5.4 Why is it generally better to centrally locate receiving and shipping doors.

Question 5.5 Plot the distribution of travel distances associated with different arrangements of receiving and shipping doors.

Question 5.6 Use the approach of the previous two questions to decide whether it is better for aisles to run perpendicular to receiving and shipping or parallel to them.

Question 5.7 Critique the layout of Figure 5.5.
Question 5.8 Under what circumstances are deep lanes appropriate? What advantages and disadvantages accrue for deeper lanes?

Question 5.9 Which of the following are true, and why? For a given set of pallets,

1. The optimal lane depth of floor storage is no greater than that for storage in pallet flow rack.

2. The optimal lane depth of floor storage is no less than that for pallet flow rack.

3. The optimal lane depths cannot be compared without knowing more about the stack heights of the pallets.

4. The optimal lane depths are the same for floor storage and for storage in pallet flow rack if the pallets cannot be stacked.

Question 5.10 Suppose that you have planned a layout for pallet rack with one global, optimal lane depth for your sku’s. Subsequently you learn that inventory levels will be half of what you had expected; how does this change the optimal lane depth?

Question 5.11 Consider a sku with pallets so short that a stack of two can be stored within a single pallet opening in rack. How would you estimate the average visits per storage location generated by such a sku? How can you continue to use the formula for optimal lane depth for such a sku?

Question 5.12 Why do you suppose single-deep rack also called “selective rack”?
Question 5.13  Which method of storage makes empty pallet positions more quickly available, thereby increasing storage capacity: floor storage or rack storage? Explain.

Question 5.14 (Project)  How deep should the lanes be if you are laying out a warehouse for the SKU's described in http://www.isye.gatech.edu/~jjb/wb/book/pallets/projects/projects.html? Answer the question for floor storage first; and then for 3-high pallet flow rack. (Hint: for pallet rack, begin by identifying those SKU's that ought to be in rack.)

Question 5.15  Explain why a warehouse with only one or two pallets of each of many SKU's is likely to prefer rack to floor storage.

Question 5.16  A distributor of major appliances, such as washers, dryers, dishwashers, and refrigerators, handles product by means of a “clamping forklift”, which picks up individual items by squeezing them from the side. All product is stored by stacking it on the floor.

Is this a “unit-load” warehouse? Why or why not?

Question 5.17 (Harder)  Suppose you are free to choose two lane depths for floor storage: One region of the warehouse will be devoted to deep lane storage and one to shallower storage.

- Prove the following: When the number of columns of each SKU greatly exceeds any candidate lane depth then, if a particular SKU belongs in the deeper lane storage, any SKU with more columns also belongs in the deeper lane storage.

- How can you use this fact to decide where to store the SKU's?

- How can you use this fact to choose the best two values for lane depth?

- Generalize to three or more lane depths.

Question 5.18 (Open-ended)  Build a simulation model to study the effect of lane depth on warehouse performance. Measure especially the effect on space utilization, labor, and observance of FIFO.
Chapter 6

Case-picking

6.1 Summary

- What is a “case”? Generally:
  - Between about 5 and 50 pounds
  - Palletizable and conveyable
  - Can be handled by a person
  - Storage unit is generally a pallet. Pallets are large and so there is not very much sku-density in pallet storage. This means there is more travel per pick. Also, large orders can be difficult to accumulate.

- Some order-picking strategies
  - Pick from pallets on floor onto pallet truck or pallet jack.
  - Use person-aboard truck to pick from pallet rack onto a pallet. Unfortunately, the horizontal speed decreases significantly when the cab is high. In addition, the truck can block other pickers, especially in narrow aisles.
  - Pick from pallets on the first level and replenish by dropping overstock pallets from above. (Note that the total number of pallets that must be dropped is (approximately) the total number of pallets sold.)
  - It is important to configure pallet rack to leave ample headroom for case-picking so that the order-picker does not hit his head on a crossbeam. Suggested height= 7 feet, which means diminished space utilization. However, it is not necessary to leave headroom for pallet openings above the first level because these will be accessed by person-aboard truck and the driver can adjust his height.
  - If sufficiently few sku’s going to lots of customers and orders are known in advance, fetch pallets of sku’s to be picked, put them in convenient spot and pull from them. Build a new pick area every day. In such a situation, it can be okay to store multiple sku’s within one opening or lane. This may increase the cost of extraction, but is a one-time cost.
– Pick from pallet flow rack onto conveyor, which makes many storage locations convenient, and hence to sortation. This is possible when cases are sufficiently uniform.

  * Sortation is inflexible and expensive, justified only by high volume.
  * Design issues: capacity, ability to handle surges. What happens if it breaks down? How many spurs are required?
  * How should recirculation be managed?
    - Exit the sorter into an accumulation lane for subsequent manual handling.
    - Recirculate back to the sorter induction point for re-consideration.
    - Divert into a separate recirculation lane
  * Operational issues: How fast? How should orders be assigned to spurs?
  * Belt versus tilt-tray sortation: Tilt-tray must circulate while conveyor need not and so can be cheaper. Tilt tray does not need to know precise orientation, size of each case, while sliding shoe sorter needs this information.

• Pallets

  – Pallets can be 2-way or 4-way. A 4-way pallet can be picked up from any side (and is more expensive); a 2-way pallet can be picked up from only two sides.

  – Picking cases to pallet can be challenging in the same way as 3-dimensional Tetris because you may not know exactly what will be requested next. In general, pallets should be built with heavy, large items on bottom and light, small items on top. This can be made easier to achieve if product is stored from heaviest to lightest along the pick path. This way the order-puller can build a stable pallet without undue travel.

  – Generally it is preferable to store pallets with the 40” side on the pickface because this means that more sku’s can be presented within a given length of aisle. But when storing 4-way pallets to support case-picking, sometimes it is preferable to orient them with the 48” side on pickface: For example, if the pallet has many small cases, the pallet is more shallow and so the order-picker does not have to reach as far.

• Sortation
Chapter 7

Design of a fast pick area

One of the first efficiencies a warehouse should consider is to separate the storage and the picking activities. A separate picking area, sometimes called a fast-pick or forward pick or primary pick area, is a sub-region of the warehouse in which one concentrates picks and orders within a small physical space. This can have many benefits, including reduced pick costs and increased responsiveness to customer demand. However, there is a science to configuring the fast-pick area.

7.1 What is a fast-pick area?

The fast-pick area of a warehouse functions as a “warehouse within the warehouse”: Many of the most popular stock keeping units (sku’s) are stored there in relatively small amounts, so that most picking can be accomplished within a relatively small area. This means that pickers do less unproductive walking and may be more easily supervised. The trade-off is that the fast-pick area must be replenished from bulk storage, or reserve.

The basic issues in the design of an fast-pick area are

- Which sku’s to store in the fast-pick area? And
- How much of each sku to store.

The answers to these questions determine the value of the fast-pick area, for if sku’s are stored there in insufficient amounts, the cost of restocking them can outweigh any savings in pick costs.

Initially, we will answer these questions by a fluid model that treats each sku as an incompressible, continuously divisible fluid; that is, we will ignore the fact that a sku actually comes in discrete “chunks” of space such as pallets or cases or individual units. In this simplest model, we can imagine the warehouse as a bucket holding various fluids (sku’s); and we will simply measure the cubic feet of storage space to be devoted to each sku. During the discussion we will point out times when this point of view can lead to inaccuracies.
A more detailed product layout, usually called a “slotting”, will explicitly account for the geometry of storage and tell exactly where each sku should be located (for example, on the third shelf of the second section of aisle 2, oriented with case width to the front). Such a plan can be constructed but it is beyond the scope of this discussion. See [7] for details and a case study.

Nevertheless, the fluid model has the advantage that it can be realized easily, for example, on a spreadsheet; and its answers are benchmarks or goals because they represent the ideal.

7.2 Estimating restocks

Since a fast-pick area is maintained by restocking, we must first estimate the cost of restocking. The cost of restocking a sku depends on the particulars of the warehouse but may include any of the following.

- The number of times the sku requires replenishment.
- The number of storage units to be replenished.
- When the restock occurs (during picking or on another shift, when timing might be less critical)

To be both usefully general and simple, we shall develop a theory in which the cost of restocking is based mostly on the number of restocks required. The first observation is that the number of restocks depends on the type of storage unit: In particular, if the sku’s are stored as pallets then each pallet will require separate handling. On the other hand, if the sku is stored in smaller containers, such as cases, one can estimate the number of restocks by a fluid model. Consider sku \( i \) of which volume \( v_i \) cubic-feet is stored in the fast-pick area. How often must we restock sku \( i \)? That depends on its rate of flow \( f_i \) through the warehouse. Flow is measured in cubic feet per year and may be determined from warehouse data as follows:

\[
\text{flow, in cubic feet/year} = \left( \frac{\# \text{ items/year}}{\# \text{ items/case}} \right) \text{ cubic feet/case}.
\]

**Estimate 7.1 (Fluid model (for small parts))** If sku \( i \) flows through the warehouse at rate \( f_i \) cubic feet per year then we can estimate that sku \( i \) will require about

\[
\frac{f_i}{v_i} \text{ restocks per year.}
\]

7.3 Storing optimal amounts

Suppose that we have decided that every sku will be represented in an fast-pick area of volume \( V \), as suggested by Figure 7.1. How would we decide how much space to allocate to each sku? The (variable) cost of storing \( v_i \) cubic feet of sku \( i \) is the cost of restocking it: The more we store, the less often we must restock but the less space
Figure 7.1: For the simplest case, assume that every SKU is represented in the fast-pick area.)
available to other sku’s. If the cost of each restock is $c_r$, then the cost per year of storing
$v_i$ cubic feet of sku $i$ may be estimated as $c_r f_i/v_i$.
Note that this contains some important assumptions:

- We restock only after exhausting our supply $v_i$.
- The variable cost of each restock is independent of the quantity restocked. This is
more likely to be true of small sku’s, which will be restocked in case quantities. It
will not hold if sku $i$ is restocked in pallet quantities, for then one trip is required
for each pallet consumed and so in this case the restock cost depends on the
quantity stored.

We will defer handling these complications for now.

7.4 Allocating space in the fast-pick area

We want to store just the right amount of every sku so that total restock costs are mini-
mized. To formalize this, let $V$ be the physical volume of available storage (measured
in, for example, cubic feet).

$$\min \sum_{i=1}^{n} c_r f_i/v_i$$

$$\sum_{i=1}^{n} v_i \leq V$$

$$v_i \geq 0$$

**Theorem 7.1** To minimize total restocks over all sku’s $j = 1, \ldots, n$, each sku $i$ should
be stored in the amount

$$v_i^* = \left( \frac{\sqrt{f_i}}{\sum_{j=1}^{n} \sqrt{f_j}} \right) V; \quad (7.2)$$

**Proof** This may be seen as an instance of the problem studied in [13]. First we rewrite
Problem 7.2 in its “Lagrangean” form, replacing the space constraint with a penalty $\lambda$
in the objective function for using too much space:

$$\min \left( \sum_{i=1}^{n} f_i/v_i \right) + \lambda \left( \sum_{i=1}^{n} v_i - V \right)$$

$$v_i \geq 0$$

On rearranging terms, this gives

$$\min \left( \sum_{i=1}^{n} f_i/v_i \right) + \lambda \left( \sum_{i=1}^{n} v_i \right) - \lambda V$$

$$v_i \geq 0$$
which is equivalent to
\[ \min \sum_{i}^{n} \left( \frac{f_i}{v_i} + \lambda v_i \right) \]
\[ v_i \geq 0 \]

But this decomposes into a collection of independent optimization problems in which each sku \( i \) solves for its optimal allocation in terms of the Lagrangean variable \( \lambda \) to get:
\[ v_i^* = \sqrt{\frac{f_i}{\lambda}}. \]

Then, setting \( \sum v_i^* = V \) we get that \( \lambda^* = \left( \sum_i \sqrt{f_i}/V \right)^2 \). Finally, substituting this expression back into that for \( v_i \) gives the theorem.

Note that the Lagrangean variable \( \lambda \) may be interpreted as the “rent” charged to each sku for storage space.

This result deserves several comments:

- The solution is not a simplistic rule. Unfortunately, the warehouse industry resorts all too often to “80/20 rules” or “ABC rules”, which treat large classes of sku’s as if they were identical. This is almost always wrong and results in small amounts of error for each sku. When this error is accumulated over tens of thousands of sku’s, the total can be significant; and it can be avoided by using optimization, which accounts for all differences among sku’s.

- This gives an “ideal” amount in which each sku should be stored. This might not be realizable in practice. For example, in flow rack one has to give each sku at least an entire lane. Nevertheless, this computation can help you identify sku’s that are stored in amounts that are far from optimum.

Theorem 7.1 has several useful practical implications:

**Corollary 7.1** The fraction of available storage space that should be devoted to sku \( i \) is
\[ \left( \frac{\sqrt{f_i}}{\sum_j \sqrt{f_j}} \right)^2. \]

**Corollary 7.2** (“Law of Uniform Restocking”) Each bay of storage (section of shelf) should be restocked at the same rate.

**Proof** In an optimal configuration of the fast-pick area each sku \( i \) will be stored in volume \( v_i^* \) and restocked \( f_i/v_i^* \) times per year. This means that each sku \( i \) will be restocked \( f_i/(v_i^*)^2 \) times per year per cubic foot stored. Substituting Expression 7.2 for \( v_i^* \) gives
\[ \left( \frac{\sum_j \sqrt{f_j}}{V} \right)^2, \]
which is independent of \( i \).
This means that restocks should be distributed uniformly throughout a storage mode, with no hot or cold spots. This provides a useful way to benchmark a fast-pick area without any measurements whatsoever. Simply ask restockers whether they tend to visit some parts of the fast-pick area more often than others; if so, then the storage policy is out of balance and there is excessive restocking. You should give more space to the sku’s in the frequently-restocked bays and less space to the sku’s in the infrequently-restocked bays.

7.4.1 Two commonly-used storage strategies

In the preceding section we derived the optimum storage policy. This result is new and not generally known to industry beyond our consulting clients. What do warehouses actually do then?

We have asked hundreds of our students in industry short courses how, in their experience, storage quantities are determined. In addition, we have interviewed warehouse managers, vendors of warehouse management systems software, systems integrators, and equipment manufacturers. The answer has always been one of the two following ways.

Equal Space Allocation: Assign each sku the same amount of space, so that, if $V$ cubic feet are available, $v_i = V/n$ and sku $i$ is restocked $n f_i/V$ times a year.

Equal Time Allocation: Assign each sku an equal time supply, so that $v_i = \left( f_i/(\sum_j f_j) \right) V$ and sku $i$ is restocked $(\sum_j f_j)/V$ times a year. (Note the similarity to the form of the optimal allocation.)

Let us study these two schemes and compare them to the optimal.

The first observation to be made is that the two schemes each offer a type of uniformity that can simplify warehouse management: Equal Space Allocation imposes a uniformity of storage that can simplify space management when old sku’s are being phased out and new sku’s introduced. Because all storage slots are the same size, a newly-arrived sku always fits into a space in the fast-pick area.

On the other hand, it is not immediately obvious how much work will be required to maintain the fast-pick area by restocking. For Equal Time Allocation, however, this work is much easier to estimate because each sku is restocked at the same frequency. For example, if each sku in the fast-pick area is stocked with 3-weeks supply then about one-third of the sku’s must be restocked each week.

Labor to maintain the fast-pick area

The fast-pick area must be maintained by restocking it. It seems obvious to most people that the Equal Space Allocation might not be effective in managing labor because it ignores labor implications in allocating space. Indeed, it treats all sku’s as if they were identical when they are manifestly not. On the other hand, the Equal Time Allocation seems to be an attractive improvement because a busier sku will get more space, which seems to make sense. This observation is folk wisdom in the industry — but, surprisingly, it is wrong:
Theorem 7.2 For a given set of sku’s, the Equal Time Allocation performs no better than the Equal Space Allocation when measured by the total labor required to restock the fast-pick area.

Proof By simple algebra:

<table>
<thead>
<tr>
<th>Equal Space</th>
<th>Equal Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allocation for sku $i$</td>
<td>$V/n \left( \frac{f_i}{\sum_j f_j} \right)$ $V$</td>
</tr>
<tr>
<td>Restocks for sku $i = \frac{f_i}{v_i}$</td>
<td>$n \frac{f_i}{V} \left( \frac{\sum_j f_j}{V} \right)$</td>
</tr>
<tr>
<td>Total restocks</td>
<td>$n \left( \frac{\sum_i f_i}{V} \right) \frac{1}{V}$</td>
</tr>
</tbody>
</table>

Example 7.1 Consider two sku’s with flows of 16 and 1 units/year respectively, which are to share 1 unit of storage. The different allocation strategies would result in the following:

<table>
<thead>
<tr>
<th>skus</th>
<th>flow</th>
<th>Equal Space Allocations</th>
<th>Restocks</th>
<th>Equal Time Allocations</th>
<th>Restocks</th>
<th>Optimum allocations</th>
<th>Restocks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/2 1/2</td>
<td>32 2</td>
<td>16/17 1/17</td>
<td>17 17</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>34</td>
<td>4/5 1/5</td>
<td>20 5</td>
</tr>
</tbody>
</table>

and the optimal allocation results in almost 30% fewer restocks.

The Equal Space/Time Allocations incur more restocks than necessary; but how severe can we expect the waste to be? Is it large enough to matter? Or are the Equal Space/Time Allocations “good enough”? We can estimate this by studying the ratio $\frac{EQT_{OPT}}{OPT}$ of the number of restocks under the Equal Space/Time Allocations to the number incurred under the optimum allocation.

As we have just shown, the Equal Space/Time Allocations incur $n \sum f_i/V$ restocks/year. In comparison, the Optimal Allocation incurs $(\sum \sqrt{f_i})^2 / V$ restocks/year. The ratio is

$$\frac{EQT}{OPT} = \frac{n \sum f_i}{(\sum \sqrt{f_i})^2} = \frac{\sum f_i/n}{(\sum \sqrt{f_i/n})^2}.$$  

Now we will estimate the magnitude of this ratio, which gives the relative numbers of restocks. Let $y_i = \sqrt{f_i}$ and consider each $y_i$ to be an independent sample of a random variable $Y$ with mean $\mu$, variance $\sigma^2$, and coefficient of variation $CV$. Then, for large $n$ (as would be expected in a warehouse):

$$\frac{EQT}{OPT} \approx \frac{\mu^2 + \sigma^2}{\mu^2} = 1 + CV^2,$$

from which we conclude that:
Observation 7.1 The more diverse the rates of flow of the sku’s, the more important it is to allocate space optimally, rather than by simple rules such as Equal Space or Equal Time.

Corollary 7.3 If the $\sqrt{f_i}$ are all independently and identically distributed exponential random variables then $CV = 1$ and $EQT/OPT \approx 2$, which means both the Equal Space and the Equal Time Allocations incur restocks at twice the minimum rate possible.

It is worth emphasizing the importance of making the optimal allocations of space. It is standard practice in the warehousing industry to store Equal Time amounts, typically after an ABC analysis. For example, Frazelle, in [14], suggests “an arbitrary allocation of space . . . or space for a quantity sufficient to satisfy the expected weekly or monthly demand”. Needless to say, an “arbitrary allocation of space” is suboptimal; and, as we have shown, an Equal Time Allocation is no better than an Equal Space Allocation, both of which treat large classes of sku’s as if they were identical. This is almost always wrong and results in small amounts of error incurred for each sku. When this error is accumulated over tens of thousands of sku’s, the total can be significant when compared to the optimal storage amounts.

To estimate this error we computed the ratio $EQT/OPT$ for 6,000 fast-moving sku’s of a major drug store chain; result: 1.45, so that Equal Time Allocations require almost 50% more restocks than optimum. With over 100 order-pickers and 20 restockers, this suggests that only about 14 restockers were necessary, which means a significant direct savings in labor.

Similarly, for the 4,000 highly diverse sku’s of a telecommunications company we computed a ratio of 2.44, which means that storing sku’s in Equal Time Allocations incurred more than twice as many restocks as necessary.

7.4.2 Minimum and maximum allocations

Sometimes the fluid model will suggest storing impractically tiny amounts of large, slow-moving sku’s. This can cause problems when there are certain minimum amounts of the sku that must be stored. For example, we cannot allocate less space for a sku than that occupied by a single unit of that sku; and sometimes we want to store at least one case of the sku or one “lane” (line of cases) of the sku.

Fortunately the fluid model can be extended to account for this. For example, to ensure that no sku receives less than its specified minimum, first allocate space according to Expression 7.2; then repeat the following steps until either all sku’s have received sufficient space or else there is no more space remaining:

- Identify those sku’s that received less than their minimum required space. If there are any such sku’s then require all other sku’s (those that received sufficient space) to return their space to be reallocated.

- Increase the allocations of the deficient sku’s to their minimum requirement. Remove them and their allocated space from the problem.

- Reallocation any remaining space among the remaining sku’s.
7.5. WHICH SKU’S GO INTO THE FAST-PICK AREA?

Figure 7.2: In general, some sku’s should not be represented in the fast-pick area.

Similarly, we may need to enforce an upper bound on how much space is allocated to any single sku. For example, we would not want to allocate space for a sku beyond that required for its maximum on-hand inventory. A procedure similar to that shown above can enforce any upper bounds on how much space individual sku’s can receive (for example, [13]).

7.5 Which sku’s go into the fast-pick area?

Any warehouse has some sku’s that are quite slow-moving by comparison to the others. It does not make sense to store such a sku in the prime real estate of a fast-pick area. Far better to store more of a popular sku so that we can defer restocking it. This will reduce our restocks but at a cost of occasionally having to pick the slow-moving sku from reserve, deep in the warehouse, which is more expensive than picking from the fast-pick area. Therefore the economics become those of Figure 7.2.

To better concentrate on the fast-pick area, let us assume for the moment that the rest of the warehouse, the reserve, is “sufficiently large” that space is not an issue there.
CHAPTER 7. DESIGN OF A FAST PICK AREA

Net benefit

![Graph](image)

Figure 7.3: The net benefit $c_i(v)$ realized by storing a SKU as a function of the quantity stored. The net benefit is zero if SKU $i$ is not stored in the forward area; but if too little is stored, restock costs consume any pick savings.

Let $s$ be the savings realized when a pick is from the forward area rather than reserve. Let $p_i$ be the number of picks forecast for SKU $i$ during the planning horizon. Then the net benefit of storing SKU $i$ forward in amount $v$ is given by $c_i$ (Figure 7.3):

$$
c_i(v) = \begin{cases} 
0 & \text{if } v = 0 \\
sp_i - c_r f_i / v & \text{if } v > 0 
\end{cases} \quad (7.3)$$

$$
\max \sum_{i=1}^{n} c_i(v_i) \quad \text{st} \quad \sum_{i=1}^{n} v_i \leq V, v_i \geq 0 \quad (7.4)
$$

In choosing SKU’s to put in the fast-pick area, the expression $p_i / \sqrt{f_i}$ is so important that we give it a name: the viscosity of SKU $i$, because it represents the effort (labor) required to move a given flow through the warehouse.

We shall show that

**Theorem 7.3** The SKU’s that have strongest claim to the fast-pick area are precisely those of greatest viscosity.

The problem of deciding exactly which SKU’s belong in the fast-pick area is now solvable. Instead of searching over all $O(2^n)$ subsets of the $n$ SKU’s, we need consider only the $O(n)$ ways of partitioning our ranked list of $n$ SKU’s into two pieces, those that go in the fast-pick area and those that do not. The difference in effort is enormous for a typical warehouse, for which $n$ may be on the order of $10^4$ or $10^5$.

Here, then, is the procedure, first presented in [10], to decide what goes into the fast-pick area and in what amounts.

- Sort all SKU’s from most viscous to least.
7.5. WHICH SKU’S GO INTO THE FAST-PICK AREA?

Figure 7.4: The majorizing function $\tilde{c}_i(v)$ is linear in the interval $[0, 2c_r f_i / (p_i s)]$.

- Successively evaluate the total net cost of putting no sku’s in the fast-pick area; putting only the first sku in the fast-pick area; only the first two sku’s; only the first three; and so on. Choose the strategy that minimizes net cost.\(^1\)

To evaluate the net cost: Charge each sku for each of its $p_i$ picks and for each of its $f_i/v_i$ restocks.

**Theorem 7.4** Choosing sku’s based on viscosity will result in a fast-pick area of total net-benefit that is no farther from optimum than the net-benefit of a single sku.

Since there are typically thousands or tens-of-thousands of sku’s considered for forward storage, the worst-case error of this heuristic is negligible. In other words, for all practical purposes, this procedure solves the problem of stocking the fast-pick area so as to realize the greatest possible net benefit.

**Proof** Consider the following problem in which each cost function $c_i$ from Problem 7.4 is replaced by the smallest function $\tilde{c}_i$ that majorizes it, as in Figure 7.4.

$$
\max \sum_{i=1}^{n} \tilde{c}_i(v_i) \quad \text{st} \quad \sum_{i=1}^{n} v_i \leq V, v_i \geq 0 \quad (7.5)
$$

The derived problem 7.5 has two important properties:

- At optimality there cannot be two sku’s assigned values strictly within their intervals of linearity $(0, 2c_r f_i / (p_i s))$. (If there existed two such sku’s, the objective function could be increased by reducing the allocation of the sku with the smaller initial rate-of-return and increasing that of the other sku.)

- An optimal solution to Problem 7.5 has the property that all sku’s chosen (that is, all with $v_i > 0$ have an initial rate of return that is strictly greater than those

---

\(^1\)Technical Note: This process can be sped up significantly by specialized search methods such as Fibonacci search because the cost is unimodal [10].
not chosen \((v_i^* = 0)\). (Otherwise, one could improve the objective function by giving more space to the sku with greater initial rate-of-return.)

Now, by these properties there exists an optimal solution to problem 7.5 of the form

\[
\mathbb{Z}^* = \left( \sum_{i=1}^{j} c_i(v_i^*) \right) + c_{j+1}(v_{j+1}^*)
\]

(7.6)

where the sku’s are numbered from greatest to least value of initial rate-of-return. Therefore,

\[
\left( \sum_{i=1}^{j} c_i(v_i^*) \right) \leq z^* \leq \mathbb{Z}^* = \left( \sum_{i=1}^{j} c_i(v_i^*) \right) + e_{j+1}(v_{j+1}^*)
\]

(7.7)

and so

\[
0 \leq z^* - \left( \sum_{i=1}^{j} c_i(v_i^*) \right) \leq e_{j+1}(v_{j+1}^*)
\]

(7.8)

and so the difference from the optimal net-benefit is no more than the contribution of the single sku \(j + 1\).

Finally, the following useful managerial guide says that there is a minimum sensible amount of each sku to store in the fast-pick area. Let \(s\) be the savings per pick achieved by storing a sku in the forward area. Then

**Theorem 7.5 (Minimum sensible storage)** If sku \(i\) goes into the fast-pick area at all, put at least volume

\[
\left( \frac{c_r f_i}{s p_i} \right).
\]

**Proof** This follows by solving \(sp_i - c_r f_i/v = 0\) to see what value of \(v\) results in a net-benefit of 0.

### 7.5.1 Managing storage slots

Equal Time Allocations are uniform in frequency of restocking; but the size of the allocations can vary greatly among sku’s. This means that the shelving of a fast pick area stocked under this policy can have many different sizes of slot (space allocated to a sku). This can make it hard to maintain the slotting when old products are discontinued and new ones introduced. It is unlikely that the newly-arrived sku would have a value of flow identical to that of the departing sku. If it is larger then the new sku requires more space than left by the old one; and if stored in this smaller space the new sku will have to be restocked more often than the intended frequency. In this way the uniformity of restocking frequency, which presumably is one of the attractions of the Equal Time Allocation, degrades over time.

Here we show that
7.5. WHICH SKU’S GO INTO THE FAST-PICK AREA?

**Observation 7.2** Optimal Allocations vary less in space allocated and so are easier to maintain than those of Equal Time.

We argue this by the following technical result.

**Lemma 7.1** For a given set of sku’s in the fast-pick area, the smallest optimal allocation for any single sku is no smaller than it would be under a Equal Time Allocation. Similarly the largest optimal allocation is no greater.

**Proof** Let the sku’s be numbered so that \( f_{[i]} \geq f_{[i+1]} \); therefore \( f_{[1]} = f_{\text{max}} \) and \( f_{[n]} = f_{\text{min}} \). We claim that

\[
\frac{f_{[k]}}{\sum_{i=1}^{k} f_{[i]}} \leq \frac{\sqrt{f_{[k]}}}{\sum_{i=1}^{k} \sqrt{f_{[i]}}}
\]

This follows by induction from the following. Because \( f_{[k]} \leq f_{[k-1]} \),

\[
\sqrt{f_{[k]}} \left( \sum_{i=1}^{k-1} \sqrt{f_{[i]}} \right) \leq \left( \sum_{i=1}^{k-1} \sqrt{f_{[i]}} \sqrt{f_{[i]}} \right) \leq \sum_{i=1}^{k-1} f_{[i]}
\]

Therefore,

\[
\frac{1}{\sqrt{f_{[k]}}} = \frac{\sqrt{f_{[k]}}}{f_{[k]}} \geq \frac{\sum_{i=1}^{k-1} \sqrt{f_{[i]}}}{\sum_{i=1}^{k-1} f_{[i]}}
\]

from which it follows that

\[
\frac{\sum_{i=1}^{k} \sqrt{f_{[i]}}}{\sum_{i=1}^{k-1} f_{[i]}} \leq \frac{\sqrt{f_{[k]}}}{f_{[k]}},
\]

from which the result follows.

A similar argument establishes that

\[
\frac{\sqrt{f_{\text{max}}}}{\sum_i \sqrt{f_i}} \leq \frac{f_{\text{max}}}{\sum_i f_i}
\]

Similarly it can be shown that

**Observation 7.3** The frequency of restocking each sku under the Optimum Allocation is less variable than under the Equal Space Allocation.

Figure 7.5 shows the variability in space under Equal Time Allocations compared to those of Optimal Allocations; and the variability in number of restocks per sku under Equal Space Allocations compared to those of Optimal Allocations. These comparisons were generated for 45 randomly-selected sku’s of a retail chain. We see in each case that Optimal Allocations have significantly less variability.
7.5.2 Priority of claim for storage in the fast-pick area

In our interviews we have learned that many people recommend choosing sku’s for storage in the fast-pick area based on measures other than viscosity. The measures reported to us have been either total picks $p_i$ or picks-per-flow $p_i/f_i$. Here we explain why these measures are popular and how they are mistaken.

As we have previously observed, the decision of how much to store in the fast-pick area seems to be ignored in practice and defaults to Equal Space Allocations or else Equal Time Allocations. Therefore the warehouse manager forfeits control over the labor to maintain the fast-pick area. In this case the only remaining issue is picks: How to get lots of picks out of the fast-pick area and so save labor that might otherwise be required to pick out of less efficient, alternative areas?

Note that this is a simpler question than we have been asking. We have been asking how to minimize total labor by getting lots of picks but not too many restocks.

The answer to this simpler question has long been known to industry through the intuition that one generally wants to store in convenient locations those sku’s with many picks for the space they occupy. (This idea is sometimes expressed in equivalent form as the priority of sku’s with small value of cube per order index, or space occupied per pick [11, 12].)

Let us see how applying the logic of pick-density (or cube-per-order index) to each of the Equal Space, Equal Time, and Optimal Allocations yields in each case a measure of the strength of the claim by a sku for storage in the fast-pick area. We shall not, of course, be able to express pick-density exactly because the space allocated depends on what other sku’s are also residing the fast-pick area. But it is possible to accomplish nearly the same thing; that is, we can write a statistic that will give the same ranking of sku’s as would pick-density.

Consider, for example, Equal Space Allocations, under which each sku $i$ selected for storage in the fast-pick area would be allocated space $V/n$ (where $n$ is the total number of sku’s chosen). The pick-density of a chosen sku would be $p_i/v_i = n p_i / V$. Now here is the key observation: We can rank the sku’s by pick density even though we do not know the values of pick density. The reason is that $V$ is a constant and the eventual value of $n$ does not change the ranking. Therefore, under Equal Space Allocations we will get the same ranking by sorting sku’s based only on values of $p_i$.

In other words, if using Equal Space Allocations then it makes sense to select the most popular sku’s (because the sku’s with most picks will, under Equal Space Allocations, have the highest values of pick-density).

Reasoning similarly, we get other simplified measures of priority for storage in the
7.5. WHICH SKU’S GO INTO THE FAST-PICK AREA?

<table>
<thead>
<tr>
<th>Storage strategy</th>
<th>Pick density $p_i/v_i$</th>
<th>Proportional to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal Space</td>
<td>$p_i/(V/n)$</td>
<td>$p_i$</td>
</tr>
<tr>
<td>Equal Time</td>
<td>$p_i/\left(\frac{f_i}{\sum_j f_j} V\right)$</td>
<td>$p_i/f_i$</td>
</tr>
<tr>
<td>Optimal</td>
<td>$p_i/\left(\sqrt{\frac{f_i}{\sum_j \sqrt{f_j}} V}\right)$</td>
<td>$p_i/\sqrt{f_i}$</td>
</tr>
</tbody>
</table>

Table 7.1: Pick density and its surrogate under three different schemes for allocating space

Figure 7.6: The net benefit realized from the fast-pick area in the warehouse of a major telecommunications company. The three graphs show the net benefit realized by storing the top $j$ sku’s in the forward area for three different ways of allocating space among the sku’s. The greatest benefit is realized by ranking sku’s according to $p_i/\sqrt{f_i}$ and allocating optimal amounts of space (the solid line on the graph); next most benefit is realized by ranking according to $p_i/v_i$ and allocating equal time supplies (the dotted dark gray line); and least net benefit is realized in sorting by $p_i/v_i$ and allocating equal amounts of space (the medium gray line).

fast-pick area that give the same rankings as pick density. These are summarized in Table ??: If using Equal Time Allocations one should select the sku’s with greatest value of $p_i/f_i$; and if using Optimal Allocations one should give priority to those sku’s with greatest value of $p_i/\sqrt{f_i}$.

We tested different versions of of the generic solution procedure on the warehouse of a major telecommunications company for whom we designed a fast-pick area. In each scenario we ranked the sku’s by pick-density and then successively computed the net benefit from storing the top $k$ sku’s in the fast-pick area. Figure 7.6 shows the results: Slotting optimal amounts delivered greatest net benefit from the fast-pick area. Equal Time Allocations was next best; and Equal Time Allocations performed least well.

Furthermore, the graph of Figure 7.6 is representative of our experience: the net benefit of Equal Time Allocations peaks soonest (that is, with fewest number of se-
lected sku’s); followed by the Optimal Allocations and then by Equal Time Alloca-
tions. Of course the value of the peak net benefit due to the Optimal Allocations is
greatest.

7.6  Additional issues

Here we extend the basic model in several directions; but the main idea remains un-
changed: There is a single number that summarizes the relative claim of each sku or
family of sku’s to prime real estate. Populate that real estate by starting at the top of
the list of most viscous sku’s and add them until the net benefit is maximum.

7.6.1  Storage by family

Some warehouses store related sku’s together. For example, if toothbrushes and tooth-
paste are displayed near each other in a retail store then put-away time can be reduced
if those sku’s are picked and shipped together; and this is more likely if they are stored
together.

Call the groups of related sku’s families. Let \( p_{ij} \) be the number of picks per year
of the \( i \)-th sku of family \( j \) and so the total picks per year of family \( j \) is \( \sum_i p_{ij} \). By
reasoning similar to that of Theorem 7.1 the total space for family \( j \) should be \( v_j^* =
\sum_i v_{ij}^* = \left( \sum_i \sqrt{f_{ij}} / \sum_{ij} \sqrt{f_{ij}} \right) \). But then, in a similar derivation, we should give
priority to those families with greatest viscosity, now generalized to be

\[
\frac{\sum_j p_{ij}}{\sum_i \sqrt{f_{ij}}}
\]

Now we can search over all partitions of families, as before. For example, to decide
between storage in the fast-pick area or the reserve areas:

- Sort families from most viscous to least.
- Successively evaluate the net cost of putting no families in the fast-pick area;
  putting only the first family in the fast-pick area; only the first two families; and
  so on. Choose the strategy that minimizes net cost.

7.6.2  Reorder points

We have assumed that \( f_i/v_i \) is an adequate estimate of the number of restocks, which
implicitly assumes that we replenish only after a stock out. In practice we prefer to
issue a restock request sufficiently in advance to avoid stockouts, which would inter-
fer with picking. We can account for this by requiring that each sku \( i \) be restocked
whenever inventory in the fast-pick area reaches a preset reorder point \( \text{rop}_i \), so that
the number of restocks of sku \( i \) is \( f_i/(v_i - \text{rop}_i) \). Our previous results (ideal amounts
of storage, pick density, viscosity) are then a little more complicated but essentially
unchanged. For example, the optimal amount of sku $i$ to store, formerly given by Expression 7.2, becomes:

$$v_i^* = \text{rop}_i + \left( \frac{\sqrt{f_i}}{\sum_j \sqrt{f_j}} \right) \left( V - \sum_j \text{rop}_j \right), \quad (7.9)$$

In other words, each sku $i$ is guaranteed space $\text{rop}_i$; then any remaining space $V - \sum_j \text{rop}_j$ is partitioned among the sku’s in the same proportions as we have already seen, that is, by the square roots of their flows.

It is important to note that when there is a long lead time until restock then the reorder points of the sku’s must be high, which means that there will be more restocks. Conversely, if product is stored in the optimal amounts, this will reduce the number of restocks, which would be expected to reduce lead times to restock. This allows reorder points to be reduced — which further reduces restocks or allows more sku’s into the fast-pick area or both.

### 7.6.3 Limits on capacity

Consider the problem of choosing sku’s to go into an fast-pick area: As previously described, we sort the sku’s by viscosity and then successively evaluate the total operating costs as more sku’s are added to the active pick area. However, as we add more sku’s to the fast-pick area, the total picks from and total restocks to the fast-pick area increase. This can be a problem if there are a priori limits on total picks or total restocks, such as when the workforce cannot be increased (by labor policy; or when picks or restocks are done robotically; or when access to the pick area is limited, such as with carousel conveyors).

Our solution procedure proceeds as before, but if either the pick rate or restock rate exceeds capacity when the $k$ most viscous sku’s are chosen for the fast-pick area, then we know that we can restrict our search to the $(k - 1)$ most viscous sku’s.

### 7.6.4 Accounting for on-hand inventory levels

Imagine some tiny sku, the total supply of which occupies only a small portion of one shelf. It would likely be wasteful to store it in two separate locations, fast-pick and reserve, because it would take so little space to put it all in the fast-pick area and so avoid restock costs. Thus, from among our total population of sku’s, there will be some that will be stored only in reserve, some that will be stored in both fast-pick and in reserve, and some that will be stored only in the fast-pick area.

How can we tell which sku’s should not be restocked to the fast-pick area? The following theorem tells us.

**Theorem 7.6** If sku $i$ goes into the fast pick area at all, put all of it if the maximum on-hand volume of of sku $i$ is no greater than

$$2 \left( \frac{c_r f_i}{s p_i} \right).$$
The term above will be familiar from Theorem 7.5, which gives us the following interpretation: There should be no separate reserve storage for anySKU in the fast-pick area for which maximum on-hand inventory takes no more space than twice its minimum sensible storage amount.

Another use of this result is to quantify the intuition that one should avoid restocking a product that moves swiftly enough through the warehouse. The following is a restatement of Theorem 7.6 in slightly different terms.

**Corollary 7.4 (High-turnover SKU's should not be internally restocked)** Any SKU that requires fewer than \(2c_r/s\) customer orders on average to turn the warehouse supply should not be restocked; that is, it should be stored either entirely or else not at all in the fast-pick area.

**Proof** Let \(v\) be the maximum volume of SKU \(i\) typically held by the warehouse. If the annual flow of this SKU is \(f_i\) then the inventory of this SKU will turn over \(f_i/v\) times a year and will require \(p_i v / f_i\) customer orders per inventory turn. Combining this with Theorem 7.6 gives the result.

Thus, if every SKU turns fast-enough, such as in a high-turnover, “Just-in-Time” distribution center then there should be no internal restocking at all.

### 7.6.5 Setup costs

If SKU \(i\) is already in the forward area, it costs some amount \(m_i\) to move it back to reserve; if it is not in the forward area it costs \(M_i\) to move it forward. Now Expression 7.3, giving the net benefit of storing \(v\) cubic feet of SKU \(i\) forward, is revised to

\[
c_i(v) = \begin{cases} 
  m_i & \text{if } v = 0 \\
  s p_i - c_r f_i/v - M_i & \text{if } v > 0 
\end{cases}
\]  

and with this enrichment Equation 7.7 becomes

\[
\left( \sum_{i=1}^{j-1} c_i(v_{i}^*) \right) + \left( \sum_{i=j+1}^{n} c_i(0) \right) \leq z^* \leq Z^* = \left( \sum_{i=1}^{j-1} c_i(v_{i}^*) \right) + \tilde{c}_j(v_j^*) + \left( \sum_{i=j+1}^{n} c_i(0) \right)
\]

and Equation 7.8 becomes

\[
0 \leq z^* - \left( \sum_{i=1}^{j-1} c_i(v_{i}^*) + \sum_{i=j+1}^{n} c_i(0) \right) \leq \tilde{c}_j(v_j^*) - c_j(0),
\]

and the heuristic solution never exceeds (an upper bound on) optimal by more than the net benefit that could be contributed by a single SKU.

### 7.7 Limitations of the fluid model

In some special situations the fluid model may not be as accurate as desired.
7.8. SIZE OF THE FAST-PICK AREA

Subadditivity of space

Some sku’s resist approximation by the fluid model. For example, auto glass for windshields is curved and therefore is stored in nested fashion. As a result, two windshields occupy only a little more space than a single windshield, not twice the space as implicit in the fluid model.

Granularity of space

The fluid model becomes less accurate when the units of storage are large with respect to the size of the shelves, such as when storing pallets in pallet rack. Another example may be found in flow rack, where each sku must occupy an entire lane, which can be ten feet deep. In such instances the results of the fluid model may not be directly realizable, but will have to be rounded to the closest allowable amount.

Where space is critical, one must explicitly account for the geometry of storage by considering every reasonable way of storing each sku: which orientation of the case, how high to stack them, and how many lanes to devote. Now, instead of checking fit by simply summing cubic feet, as we have done in the fluid model, we must check fit by checking whether the spatial arrangements of the cases fit on the shelves, for every case, every shelf, and each dimension. This requires a much more detailed model and requires vastly more computation; but it can nevertheless be done with great precision. For example, Bartholdi and Hackman describe such a project for a major chain retailer in which space was saved, a quarter of an inch here and there, over 10,000 sku’s, with the result that required storage in one warehouse was reduced from 325 to 285 bays of flow rack, with no increase in total restocks [7].

Figure 7.7 shows typical results of the more powerful model, applied in this example to slot sku’s in a bay of flow rack. Notice that picks are concentrated in the “golden zone” (waist-high shelves) and that the program has determined the exact orientation of each sku, the number of lanes, and how high to stack the cases.

7.8 Size of the fast-pick area

7.8.1 How large should the fast-pick area be?

As the fast-pick area becomes larger, we can fit more sku’s in, which means more pick savings, or larger amounts, which means less restocking; but we get less savings per pick because of the additional walking.

It is possible to build explicit models of exactly how the pick savings diminishes with increased size of the fast-pick area. For example, suppose we are configuring an aisle of flow rack as our fast-pick area and are undecided about how many bays it should extend. With each additional bay, the pick savings decreases approximately linearly. The rate at which it decreases depends on the economics of each particular warehouse and must be estimated, for example, by time-motion studies. Let us assume that we have determined it decreases at rate $S$. Then we have the following.
Figure 7.7: Example of slotting that accounts for the geometry of the sku’s and the storage mode. This pattern of storage minimizes total pick and restock costs for this set of sku’s and this arrangement of shelves. Numbers on the right report the percentage of picks from this bay on each shelf. (This was produced by software written by the authors.)
7.8. SIZE OF THE FAST-PICK AREA

**Theorem 7.7** For linear model of storage (for example, adding bays to an aisle of flow rack), the optimum size of the fast pick area is given by

\[
V^* = \frac{\sum_{i=1}^{k} \sqrt{f_i}}{\sqrt{S \sum_{i=1}^{k} p_i}}
\]

for some number \(k\) of the most viscous sku’s, and where \(S\) is the rate of pick savings as a function of the size of the fast-pick area.

Again we recognize a familiar theme: Sort sku’s from most to least viscous and repeatedly compute allocations for the \(k\) most viscous sku’s together with the resultant net benefit (pick savings minus restock costs). Choose that value of \(k\) for which net benefit is maximized.

In the same way, you can use the discrete model to evaluate different configurations of equipment. For example, should racks have four, five, or six shelves per bay? With the discrete model you can compare the the net cost of each alternative configurations to see which is best for your particular population of sku’s and order history.

7.8.2 How can the fast-pick area be made larger?

Almost every warehouse manager would like to increase the size of his or her fast-pick area but is unable to because of space constraints. Enlargement is further unlikely because of the cost of specialized equipment. (For example, ten new bays of flow rack equipped with pick-to-light locations for 10 sku’s per bay could cost around ten thousand dollars for the pick-to-light fixture plus another the same amount for the flow rack.) Fortunately, one can realize all the benefits of a larger fast-pick area simply by reducing restock costs.

Consider the total net benefit of having stored sku’s \(1 \ldots n\) in the fast-pick area:

\[
\sum_{i=1}^{n} s p_i - c_r f_i / v_i,
\]

which, substituting \(v_i = \left(\sqrt{f_i / \sum_{i} \sqrt{f_j}}\right) V\), may be written as

\[
\sum_{i=1}^{n} s p_i - \left(\frac{c_r}{V}\right) \left(\sqrt{f_i \sum_{i} \sqrt{f_j}}\right).
\]

But notice that changing restock costs by a factor of \(\alpha\) changes the term \((c_r/V)\) to \((\alpha c_r/V)\) = \((c_r/(V/\alpha))\). In other words, the total net benefit of the fast-pick area is exactly the same, whether the cost-per-restock is halved or the space is doubled. As a practical matter it is better to reduce restock costs because increasing the size of the fast-pick area will in general increase the average cost per pick there and so reduce the savings per pick. Therefore we observe the following.

**Theorem 7.8 (“Law of virtual space”)** Changing the cost \(c_r\) by a factor of \(\alpha\) is economically equivalent to changing the space \(V\) available in the fast-pick area by a factor of \(\alpha\).
7.9 Multiple fast-pick areas

Suppose there are $m$ fast-pick areas, each with its own economics: savings per pick (compared to from reserve) and restock costs, as suggested by Figure 7.8. Now the question becomes which SKU’s go to which fast-pick areas and which to reserve. Each of the $n$ SKU’s can be picked from at most one area.

Brute-force search can find the best stocking strategy within $O(m^n)$ steps by enumerating all possible assignments of SKU’s to storage modes and then evaluating the result. However, this is completely impractical for realistic values ($m = 2, 3$ and $n$ on the order of tens of thousands).

It can be proven, but is beyond the scope of this book, that

**Theorem 7.9** *SKU’s of greatest viscosity go into fast-pick area with greatest savings per pick.*
In other words, some optimal stocking strategy has the following structure: If all sku’s are ranked by viscosity then the top \( k_1 \) sku’s would be assigned to the storage mode with largest savings per pick, the next \( k_2 \) would be assigned to the next best storage mode, and so on.

Note that “best storage” depends only on savings per pick and is independent of restock costs.

This theorem suggests that an optimal stocking strategy for \( n \) sku’s among \( m \) storage modes may be found by the following procedure.

1. Sort sku’s from most to least viscous
2. Search for the partition into \( m \) sets of contiguous sku’s that maximizes the net benefit.

This allows us to find an optimal stocking strategy within \( O(n^m) \) steps. But the following result, again beyond the scope of this book, allows us to further reduce the time to find an optimal strategy and so finally achieve practical solutions to realistic problems.

**Theorem 7.10** The total net benefit over all storage modes is sequentially unimodal. That is, if we fix all \( k_1 \) but one then the total net benefit is unimodal in that unfixed \( k_1 \).

Sequential unimodality thus allows us to further speed up our search, reducing it to only \( O((\log n)^m) \) steps.

### 7.10 A fast-pick area in pallet storage

A typical fast-pick area in pallet storage is when workers pick cases from pallets on the ground level and overhead storage is used as reserve. When a ground level pallet location is emptied then a forklift drops a pallet from above to the now empty ground location. In this case we consider the restock to be the pallet drop.

Because pallet drops are unit-load moves, it is no longer accurate to estimate the number of restocks of sku \( i \) as \( f_i / v_i \): Because each pallet must be handled separately, a better estimate of the number of restocks (pallet drops) is simply the total quantity moved through the fast-pick area measured in pallets.

Consider the following typical operating policy in a case-pick-from-pallet operation.

- If **no** pallets are in the fast-pick area, then all picks are from bulk storage.
- If **some but not all** pallets are in the fast-pick area, then all picks for less-than-pallet quantities are from the fast-pick area and all picks for full-pallet quantities are from the bulk storage area.
- If **all** the pallets are in the fast-pick area, then all picks, both for less-than-pallet quantities and for full-pallet quantities, are from the fast-pick area.
Let \( p_i \) be the number of less-than-pallet picks, \( d_i \) the number of pallets moved by such picks, \( P_i \) the number of full-pallet picks, and \( D_i \) the number of pallets moved by such picks. Let \( l_i \) be the minimum number of pallets to be stored in the fast-pick area and and \( u_i \) be the maximum on-hand inventory. Then the net-benefit of storing \( x \) pallets of sku \( i \) in the fast-pick area is

\[
\text{net benefit} = \begin{cases} 
0 & \text{if } x = 0; \\
sp_i - c_r d_i & \text{if } 0 < x < u_i; \\
sp_i + D_i & \text{if } x = u_i.
\end{cases}
\]  

(7.11)

This function is represented in Figure 7.9. An argument similar to that which established Theorem 7.4, gives

**Theorem 7.11 (“Law of All or Nothing… or Minimum Required”)** Any sku that is picked from pallets should either not be in the fast-pick area at all; or it should have the minimum required amount (that is, the reorder point plus one pallet); or it should have all of its on-hand inventory in the fast-pick area.

![Figure 7.9: Net benefit of storing various full-pallet quantities of a sku in the fast-pick area](image)

**Proof sketch** As before, replace the optimization problem with one in which the net benefit of each sku has been replaced with its majorization (the dotted lines in Figure 7.9). By the same reasoning, we can construct a solution that is no farther from optimum than the net benefit of a single sku. Such a solution has the property that if sku \( i \) is stored in the fast-pick area then any sku with greater marginal rate of return \( (sp_i - c_r d_i) / lb_i \) must also be in the fast-pick area. Furthermore, note that the net-benefit function offers an incentive for storing the minimum allowable amount; but there is no additional incentive to increase the amount stored unless the entire on-hand inventory of that sku can be stored in the fast-pick area.  

[\( \blacksquare \)]
7.11. ON THE LIGHTER SIDE

There are many ways to exploit this result to compute exactly which sku’s go into a pallet fast-pick area and in what amounts. One such way is to rank sku’s by the marginal rate of return of increasing its presence in the fast-pick area from zero to lb_i, which is the initial slope \((sp_i - c_i d_i)/l_i\). Then repeatedly choose that sku with greatest slope and increase the amount stored from 0 to its minimum allowable amount l_i, recompute its marginal rate of return as \((sD_i + c_i d_i)/(u_i - l_i)\), which is the incentive to increase its presence in the fast-pick area from l_i to u_i and reinsert into the sorted list. (See Homework Question 7.21.)

7.11 On the lighter side

One logistics manager told us that he too had been concerned that Equal Space Allocation, which was enforced by his warehouse management system, required too much labor to restock the fast-pick area. The manager paid a large sum of money to have his warehouse management system revised to support Equal Time Allocations — which, as we now know from this chapter, made no difference at all!

7.12 Summary

- Concentrate activity in a small footprint to reduce picking costs, increase responsiveness, and free up space to deal with growth, seasonalties, and other fluctuations.

- The configuration of a warehouse can be optimized based on physical size of the sku’s and a history of customer orders. To do this you must know the physical dimensions of the storage units and the number of selling units per storage unit.

- Key statistics for each sku are its picks \(p_i\) measured in pick-lines per year and its flow \(f_i\), measured in cubic-feet per year. These statistics can be forecasts or historical data.

- The viscosity of a sku \(i\) that is stored in less-than-pallet quantities is \(p_i/\sqrt{\sum f_i}\), which measures the work required to pull a given amount of physical volume through your warehouse. The most viscous sku’s are the most suitable for the best storage locations because they generate the largest net benefit (pick savings minus restock costs) for the space they consume.

- You should put the most viscous sku’s into the fast-pick area. You must search to determine exactly how many to put in.

- For multiple fast-pick areas, the most viscous sku’s should be stored in the fast-pick areas with greatest pick-savings.

- For those sku’s stored in less than pallet quantities in the fast-pick area, the optimal amounts, measured in cubic feet, are \(v_i^* = \left(\sqrt{T_i/\sum f_j}\right)V\).
This can be significantly better than the common practice of storing “equal-time” quantities of each sku (which requires the same total labor to restock the fast-pick area as does storing equal quantities).

Storing sku’s in the fast-pick area in the optimal quantities will reduce restocks, which can reduce lead time to restock, which allows reorder points to be reduced — which further reduces restocks! In other words, there are secondary savings that accrue automatically to correcting the amounts stored.

A test of the storage policy at a warehouse is this: At optimality each bay (section, cabinet) of shelving in the forward pick area should be restocked at the same rate. (One can ask the restockers whether they are visiting any part of the fast-pick area especially often or especially rarely.)

- When sku’s are stored as full pallets in the fast-pick area then each sku is either:
  - not in the fast-pick area at all; or
  - is in at its minimum allowable amount; or else
  - every pallet of that sku is in the fast-pick area.
7.13 Questions

**Question 7.1** The estimate of the number of restocks as $f_i/v_i$ in the fluid model is only an estimate. Explain how it can be wrong.

**Question 7.2** Why is shallow storage generally preferable for smaller, slower-moving sku’s? For example, why might you prefer to put such sku’s in static rack (bin-shelving), which might be only 2.5 feet deep, rather than in flow rack that is 10 feet deep?

**Question 7.3** Sku A is requested ten times as often as sku B but has one-half the flow. Assuming both go into the fast-pick area, what relative amounts of space should be allocated to each?

**Question 7.4** Consider sku A, which was picked 10,000 times last year, and sku B, which was picked 100 times. Which has greater claim to storage in the fast-pick area or is it impossible to tell? Explain.

**Question 7.5** Sku i had annual picks and flow of $p_i$ and $f_i$ respectively. Sku j has been in the distribution system for only 3 months, during which it was picked $p_j$ times, with flow of $f_j$ cubic feet. Do $p_i/\sqrt{f_i}$ and $p_j/\sqrt{f_j}$ accurately reflect the relative claims of sku i and j to storage in the fast-pick area? Explain.

**Question 7.6** Suppose that a sku is repackaged into smaller cases that hold 100 units, rather than the 250 units that the previous, larger cases held. Has the suitability of this sku for storage in the forward area increased or decreased or remained the same or is it impossible to tell? Explain.

**Question 7.7** Each drugstore of a retail chain is assigned one day in the week at which to submit its order to the regional distribution center. Each store typically includes a request for at least several bottles of popular brand of shampoo. Suppose that this shampoo, currently in the forward pick area, is replaced by a concentrated version of the same product, packaged in exactly the same way but now each bottle lasts twice as long as before. Has this sku become more or less suitable for storage in the fast-pick area or remained the same or is it impossible to tell? Explain.

**Question 7.8** Suppose that a marketing campaign has just begun for one of the more popular sku’s. This campaign does not attract any new customers but it is successful in getting regular customers to buy in greater quantities than previously. Has this sku become more or less suitable for storage in the fast-pick area or remained the same or is it impossible to tell? Explain.

**Question 7.9** Assume that you have set up a fast-pick area in flow rack and stocked it optimally. Later you add pick-to-light capability, which increases the pick rates. (Everything else, including restock costs, remain unchanged.) If you were to re-compute the optimal slotting you would find which of the following? Explain.

- Some sku’s would leave the fast-pick area and the remaining sku’s would get more space.
• The sku’s and amounts stored in the fast-pick area would remain unchanged.

• New sku’s would join those already present and each of the sku’s in the fast-pick area would get less space.

• A completely new mixture of sku’s may be selected.

**Question 7.10** Suppose you enlarge your fast-pick area (but not enough to change the savings per pick significantly). On recomputing the optimal population of sku’s, you may find which of the following results?

1. Some of the sku’s currently in the forward area may be moved out and replaced by an assortment of other sku’s.

2. All the sku’s currently in the forward area will remain but they will be joined by additional sku’s.

3. No new sku’s will be moved into the forward area; instead, the most viscous will remain and each will get more space.

4. It is impossible to tell.

**Question 7.11** We have implicitly assumed that each restock costs the same. Is this correct? Discuss whether and how restock costs might depend on

• Location of sku in fast-pick area

• Location of sku in reserve area

• Amount to be restocked

**Question 7.12** The following is a miniature of a real problem: In what amounts should I store sku’s to minimize labor?

Suppose you have 10 cubic feet available in flow rack, which is restocked from a distant reserve area, and you have have just added three sku’s, with projected activity as follows.

<table>
<thead>
<tr>
<th>sku</th>
<th>picks/month</th>
<th>units/month</th>
<th>units/case</th>
<th>ft³/case</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1000</td>
<td>2000</td>
<td>200</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>300</td>
<td>1200</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>250</td>
<td>4000</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

A. Suppose you have decided to put all three sku’s in flow rack. How much space should be allocated to each sku?
B. How often must each sku be restocked?
C. How many restocks would each sku incur if allocated an equal share of the space?
D. How many restocks would each sku incur if allocated equal time supplies?
Question 7.13  A. Reconsider Question 7.12 supposing that no more than 5 cubic feet may be allocated to any single sku. How would your answer change?
B. Generalize your answer to part A: Devise a procedure to handle arbitrary upper bounds on allocations. Do the same for lower bounds.

Question 7.14  A. Reconsider Question 7.12 supposing you may choose none or any or all of the three sku’s to put in the flow rack. Which should they be? What will be the net cost?
Assume that it costs an average of $0.15 per pick from flow rack but costs about $1/restock. The alternative is to pick from reserve, where each pick costs about $0.25.
B. Suppose that you must rent the space in flow rack from another department within your organization. How much rent would you be willing to pay for the use of this area?
C. How much would it be worth to you if sku B was reformulated as a “concentrate” so that exactly the same amount of product could be fit into a much smaller case, occupying only 1 cubic foot?

Question 7.15  Reconsider the preceding question, supposing that sku’s A and B must be stored together; that is, either they both must be stored in the forward area or else they both must remain in reserve. Now what is the best storage plan? (Hint: Treat sku’s A and B as one product family and sku C as another.)

Question 7.16  In the basic fluid model, if a sku is stored in the fast pick area, economics suggests that it ought to be stored in at least what amount (volume)? Choose the best answer and explain why it is so.

- $0$
- $c_r f_j / (s p_i)$
- $s p_i^2 / (4 c_r f_i)$
- All of the sku that is stored in the warehouse

Question 7.17  In Section 7.6.1 it was claimed that the correct value for the viscosity of family $j$ is computed from the sku’s $i$ within that family as follows:

$$\frac{\sum_i p_{ij}}{\sum_i \sqrt{f_{ij}}}.$$

Prove that this is correct. In particular, explain why it would be incorrect to use

$$\frac{\sum_i p_{ij}}{\sqrt{\sum_i f_{ij}}}.$$

Question 7.18  Follow the suggestion of Section 7.6.2 and derive the formulae for optimum allocations when each sku $i$ has a non-zero reorder point $r_{oi}$. Do the same for viscosity.
Question 7.19 Is the following true, false, or impossible to determine? (Explain your answer.) Suppose the reorder point of a single sku in the fast-pick area is increased. Then, in the optimal allocation of space to minimize restocks in the fast-pick area, every sku must be restocked more frequently. (Assume that the population of sku’s remains unchanged.)

Question 7.20 Suppose that the reorder point of a sku (the level at which a request for restocking is issued) has been increased due to variability in customer demand. Has that sku a greater or lesser claim to storage in the fast-pick area, or neither? Explain.

Question 7.21 We have argued that the “best” storage modes are those with the smallest pick costs. What if the “best” mode has been placed far from reserve and so each restock is quite expensive? How can it make sense that the “best” sku’s should be stored here? How might the high cost of restocking affect the assignment of sku’s to modes?

Question 7.22 Consider a warehouse with multiple fast-pick areas. Which fast-pick area will get the most viscous sku’s?

1. The fast-pick area with the greatest savings-per-pick.
2. The fast-pick area with the least cost-per-restock.
3. The fast-pick area with the largest ratio of savings-per-pick to cost-per-restock.
4. The fast-pick area with the greatest volume.
5. The fast-pick area with the least volume.

Question 7.23 The savings in restocks realized by storing sku’s in their optimal amounts (rather than Equal Space or Equal Time Allocations) is likely to be most significant when space is tight. Why is this?

Question 7.24 Make a numerical example that shows that sequencing sku’s by viscosity is not the same as sequencing by picks divided by flow. In other words, find values for which \( p_i / \sqrt{T_i} < p_j / \sqrt{T_j} \) but \( p_i / f_i > p_j / f_j \).

Question 7.25 Consider three ways of stocking the fast-pick area: Equal Space Allocations, Equal Time Allocations, and Optimal Allocations. Which method results in

- Least variability among the quantities allocated? Most variability?
- Least variability among the numbers of restocks required for each sku? Most variability?

Question 7.26 Consider a fast-pick area where cases are picked from pallets. (All full-pallet picks are from reserve storage and may be ignored.) Storing a sku in the fast-pick area realizes a savings of 1 minute per pick; but each restock requires about 3 minutes. Which of the following sku’s has greatest claim to storage in the fast-pick area? (Both “Demand” and “Reorder point” are given in numbers of pallets. Because
of volatility of purchasing, management is unable to specify a reasonable upper bound on how many pallets of each SKU may be expected in the warehouse.

How many pallet positions could these SKU's usefully occupy in the forward pick area?

<table>
<thead>
<tr>
<th>SKU</th>
<th>Case picks</th>
<th>Demand</th>
<th>Reorder Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>800</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>200</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Question 7.27 Consider a collection of SKU's that are candidates for storage in a fast-pick area composed of carton flow rack. Anything not stored in the fast-pick area will be picked from pallets in reserve.

Assume that for each SKU we know the number of piece picks, the piece flow $f_i$, the number of full-carton picks, and the carton flow $f_i^c$. (Piece flow is the total volume of product that was picked as individual pieces; carton flow is the total volume of product that was picked as full-cartons.)

A. Explain how to use the methods of this chapter to decide whether it is better to do all picking of full-cartons from flow rack or from reserve. (Ignore consolidation costs.)

B. Explain how to use the methods of this chapter to decide on a SKU-by-SKU basis whether SKU should be stored in the fast-pick area and if so, how much space it should be allocated and whether carton picks of SKU should be from its forward location or from its reserve location. (Thus each SKU may be stored in one of three configurations: No storage in the fast-pick area and all picks are from reserve; or: storage in the fast-pick area but only piece picks are from fast-pick and carton picks are from reserve; or: storage in the fast-pick area and both piece picks and carton picks are from fast-pick.)

Chapter 8

Detailed slotting

Slotting refers to the careful placement of individual cases within the warehouse. It may be thought of as layout in the small.

The most immediate goals in slotting a warehouse are the following.

- Squeeze more product into available space; and
- Achieve ergonomic efficiency by putting popular and/or heavy items at waist level.

At the same time, one wants to avoid creating congestion by concentrating popular items too much.

There are additional concerns that vary from warehouse to warehouse. For example, some distribution centers supporting retail drug stores prefer to store similar-looking medicines apart to reduce the chance of a picking error; but they store non-drug items in the same product family so that, for example, all the hair care products will tend to be picked together. This means that, when the retail store opens the shipping container, they will find it packed with items that are to be stored on the same aisle, so that putaway required less labor. Storing products in the warehouse by product family may cost the warehouse more in space and labor but it may save putaway labor in thousands of retail stores.

8.1 Case orientation and stack level

Space issues are most significant for sku’s that are stored in less-than-pallet quantities. The storage unit of many such sku’s is a cardboard case, which is a rectangular solid. Such a shape may be placed on a shelf in any of up to six orientations. Furthermore, once an orientation has been selected, the shelf space above and behind a case of that item are rendered unusable by other items, because to store another sku there would create unnecessary work to retrieve the sku behind, as shown in Figure 8.1. Consequently one might as well fill up that space as much as possible with additional cases of the same sku, as in Figure 8.2.
Figure 8.1: Storing an item so that it blocks access to another creates useless work and increases the chance of losing product.

Figure 8.2: Each storage unit may be placed in any of up to six orientations; and each lane might as well be extended to the fullest depth possible and to the fullest height.
8.2. MEASURING THE EFFECTIVENESS OF A SLOTTING

How can we tell if a slotting is effective? Let us concentrate on the main issue, which is to reduce labor and use space effectively. Thus ideally we would store each sku so that its slotting supports many picks in a small amount of space. We want many eaches so that restocking is infrequent; and we want them to occupy little of the pick face so that we can present many sku’s to a picker within a small area, thus reducing travel.

It is a simple matter to count the eaches in a slotting, but it is less clear how to charge for space consumed. The simplest measure of space consumed is just the physical volume of the sku, but this omits any consideration of geometry or of resultant effect on labor costs.

A more accurate measure would be the amount of storage space rendered unusable by other sku’s, which would include the physical volume of space occupied by this sku

Figure 8.3: Each lane should contain only a single sku, with each case oriented identically; and the lane should be extended as deep and as high as the shelf allows.

Note that each orientation results in different numbers of cases stored in a lane; and different amounts of space consumed.

Example 8.1

We shall assume that product is stored so that each lane contains only a single sku, with each case oriented identically; and the lane is extended as deep and as high as the shelf allows.
plus any clear height above it plus any empty space at the back of each lane.

We can base our evaluation of a slotting still more strongly on resultant labor costs by measuring only the space consumed on the pick face, which is where the labor takes place. On the pick face, horizontal space is generally more valuable than vertical space because most labor is devoted to travelling horizontally. (Vertical movement is avoided where possible because it is much slower than horizontal movement. For example, it is generally person-aboard trucks and there are safety concerns.) Consequently, a good slotting will fit many eaches within a small width of pick face. This leaves room for other sku’s on the shelf and so allows the pick face to present more sku’s per unit distance of travel along the floor.

8.3 Packing algorithms
8.4 Questions

Question 8.1 Assume that the case in Figure 8.2 is $1 \times 2 \times 3$. What are the storage efficiencies of each of the six orientations shown in a shelf of height 3?
Chapter 9

Order-Picking by ‘bucket brigade’

Self-organizing systems do not require a centralized authority to manage them. Instead, they achieve global coordination spontaneously through the interaction of many simple components.

When workers are organized into “bucket brigades” they can function as a self-organizing system that spontaneously achieves its own optimum configuration without conscious intention of the workers, without guidance from management, without any model of work content, indeed without any data at all. The system in effect acts as its own computer.

9.1 Introduction

A self-organizing system is one in which global organization spontaneously evolves from myriad local interactions of the pieces. Here is an example: Consider a hive of honeybees. Each day they face a logistics problem of how to coordinate their efforts to harvest nectar. The measure of success is a social one: the good of the colony. But bees have no blueprint, no mechanism of central planning. Instead, each bee follows a simple “algorithm” that determines what she does next; and when many bees follow the same algorithm, an allocation of foragers evolves that is close to the best imaginable. In effect the colony acts as a computer that finds the (nearly) optimal allocation of effort [8].

Among the advantages of this self-organization are that:

- It requires no central planning or higher organizational entity. There is no management function because each entity simply follows a local rule.

- It is adaptive: It will spontaneously reallocate effort in response to changes in the environment.
9.2 Order-assembly by bucket brigade

"Bucket brigades" are a way of coordinating workers who are progressively assembling product along a flow line in which there are fewer workers than stations. Each worker follows this simple rule: "Carry work forward, from station to station, until someone takes over your work; then go back for more". When the last worker completes a product, he walks back upstream and takes over the work of his predecessor, who then walks back and takes over the work of his predecessor, and so on, until the first worker begins a new product at the start of the line. No unattended work-in-process is allowed in the system.

Note that workers are not restricted to any subset of stations; rather they are to carry each product as far toward completion as possible. Note also that a worker might catch up to his successor and be blocked from proceeding; the bucket brigade rule requires that the blocked worker remain idle until the station is available.

The final requirement of bucket brigades is that the workers be sequenced from slowest to fastest along the direction of material flow. These protocols, taken together, make the bucket brigade line a perfect pull system.

9.2.1 A model

Call each instance of the product an item and consider a flow line in which each of a set of items requires processing on the same sequence of \( m \) work stations, as in Figure 9.1. A station can process at most one item at a time, and exactly one worker is required to accomplish the processing.

All items are identical and so each requires the same total processing time according to some work standard, which we normalize to one "time unit".

The Normative Model, suggested in [3], is given in the following assumptions. We call this model "normative" because it represents the ideal conditions for bucket brigades to work well. However, it is not necessary that these assumptions hold exactly: The behavior of a bucket brigade will resemble that predicted by the Normative Model to the degree that the assumptions of the Normative Model hold. Accordingly implementations should try to make these conditions hold as much as possible — but it is not necessary that they hold exactly, or even to any great extent.

The assumptions are:
Assumption 9.1 (Insignificant Walkback Time) The total time to assemble a product is significantly greater than the total time for the workers to handoff their work and walk back to get more work.

Assumption 9.2 (Total Ordering Of Workers By Velocity) Each worker $i$ can be characterized by a work velocity $v_i$.

Assumption 9.3 (Smoothness, Predictability Of Work) The work-content of the product is spread continuously and uniformly along the flow line (the length of which we normalize to 1).

The assumption of Insignificant Walkback Time is uncontroversial; it claims simply that it takes longer to assemble a product than it does to walk the line; and, furthermore, it is easy to handoff work.

The assumption of Total Ordering Of Workers By Velocity is likely to hold in a mass-production environment, where work has been “de-skilled” so that velocity is based on a single dimension, such as motivation or eye-hand coordination. (This point is more fully documented in [3]).

There is clearly some license in the assumption of Smoothness And Predictability Of Work; nevertheless, this assumption is reasonable in many instances, detailed by us elsewhere [3]. Suffice it to remind the reader that management and engineering strive to remove variance from work and eliminate bottlenecks, a result of which is to move practice closer to the Normative Model. Still, this assumption is at least less clear than the others and accounting for this is part of the art of implementing bucket brigades.

To what extent do the conclusions of the Normative Model hold when there is variation in the work-content? In short, the behavior of a bucket brigade remains qualitatively similar to behavior predicted by the Normative Model, with this caveat: the faithfulness of the replication depends on the degree of randomness. This means that, except in degenerate cases, it remains preferable to sequence the workers from slowest to fastest and one can expect a high production rate from bucket brigades.

Bartholdi and Eisenstein (1996a) have described the behavior of bucket brigade production lines under the Normative Model [3, 1]. Their main results, slightly simplified, are as follows.

**Theorem 9.1**  
There is a unique balanced partition of the effort wherein worker $i$ performs the interval of work:

$$
\text{from } \frac{\sum_{j=1}^{i-1} v_j}{\sum_{j=1}^{n} v_j} \text{ to } \frac{\sum_{j=1}^{i} v_j}{\sum_{j=1}^{n} v_j}, \quad (9.1)
$$

so that each worker invests the same clock time in each item produced.

If the workers are sequenced from slowest to fastest then, during the normal operation of the line, work is spontaneously and constantly reallocated to reach this balance; and the production rate converges to

$$
\sum_{i=1}^{n} v_i \text{ items per unit time},
$$
which is the maximum possible for the given set of workers.

- If the workers are not sequenced from slowest to fastest, then the line will “sputter”: that is, it will produce erratically and at suboptimal rate. Furthermore, the line can behave in counterintuitive ways, such as production rate decreasing when a worker increases his velocity.

Here is the proof that bucket brigades are self-balancing under the Normative Model.

**Proof** Imagine that we are taking a series of photographs of the line at those times when the workers have just made their handoffs and the first, slowest worker is beginning a new product. We will study how these photographs change.

Let \( x_i^{(k)} \) be the percent completion of the product held by worker \( i \) in the \( k \)-th photograph (that is, after a total of \( k \) items have been completed). (See Figure 9.2).

Then the clock time separating workers \( i \) and \( i + 1 \) is

\[
    t_i^{(k)} = \frac{x_i^{(k)} - x_{i+1}^{(k)}}{v_i}
\]

and the next item will be completed after time

\[
    t_i^{(k)} = \frac{1 - x_i^{(k)}}{v_i}.
\]

In the next, \( k + 1 \)-st photograph, the clock-time separating workers \( i \) and \( i + 1 \) becomes

\[
    t_i^{(k+1)} = \frac{x_i^{(k+1)} - x_{i+1}^{(k+1)}}{v_i} = \frac{(x_i^{(k)} + v_i t_i^{(k)}) - (x_i^{(k)} + v_{i-1} t_{i-1}^{(k)})}{v_i} = \left( \frac{v_{i-1}}{v_i} \right) t_i^{(k)} + \left( \frac{1 - v_{i-1}}{v_i} \right) t_i^{(k)},
\]

Because the workers are sequenced from slowest-to-fastest \( (v_{i-1} / v_i) < 1 \), and so we may interpret these equations as describing a finite state Markov Chain that is irreducible and aperiodic. By the Markov Chain Theorem the \( t_i^{(k)} \) and therefore the \( x_i^{(k)} \) converge; and the specific claims follow by simple algebra.

Figure 9.3 shows an example of how the movement of the workers stabilizes, with the faster workers eventually allocated more work. This figure was generated by a simulation of three workers of velocities \( v = (1, 2, 3) \).
Figure 9.3: A time-expanded view of a bucket brigade production line with three workers sequenced from slowest to fastest. The solid horizontal line represents the total work content of the product and the solid circles represent the initial positions of the workers. The zigzag vertical lines show how these positions change over time and the rightmost spikes correspond to completed items. The system quickly stabilized so that each worker repeatedly executes the same portion of work content of the product.
This analysis suggests an effective way to partition a work force into teams. Current practice in the apparel industry is to base the pay of each individual on the production of her team. Consequently, the fastest workers prefer to work with other fast workers. This is unattractive, however, from the perspective of management because it does not help integrate newer, slower workers into the work force. If slower workers form teams to themselves there can be morale problems. In addition, the newer workers will not be in a position to learn directly from more experienced workers.

It seems better for management to put very different workers on the same team, sequenced from slowest to fastest; then each production line will be self-balancing and will achieve the maximum production rate. Furthermore, the greater the range of velocities on a team, the more powerfully the line will be drawn to balance. Finally, when there are large differences in the velocities of team members then the system will remain self-balancing even allowing for the inevitable variations in the velocities of the team members.

### 9.2.2 Improvements that are not

It is tempting to try to improve the performance of bucket brigade lines by modifying the protocol; however, the variants that come first to mind actually perform worse. For example, an appealing but flawed variation of the bucket brigade protocol is to allow any worker, when blocked, to leave his partially-completed item in a buffer before the busy station and walk back to take over the work of his predecessor. This variant protocol will increase work-in-process inventory and can even reduce the production rate! This can be seen in simulations, where workers tend to collect in the region of the line preceding any station that is frequently busy. This increases the production rate of the preceding segment of the line, which only accelerates the accumulation of in-process inventory immediately preceding the highly-utilized station. This, in turn, decreases overall production rate of the line for two reasons:

- Fewer workers remain to staff the final segment of the line so each tends to assume a larger share of work and the time between product completions increases.

- Because no one waits in front of the frequently busy station, it is idle every time a worker leaves it, which is contrary to the principal of keeping bottleneck stations constantly busy.

Eschewing buffers seems to contradict conventional wisdom that it is important to have buffers near a bottleneck — until one realizes that in bucket brigade production one must buffer both work-in-process and a worker, which is done by requiring the blocked worker to remain waiting at the bottleneck station.

One might also think that the bucket brigade protocol could be improved by requiring the workers to circle through the work stations. This avoids any delay in handing off work but it requires that every worker perform every task. There are several objections to be made to this. First, when real workers are finally assigned to the line they will not be of identical skill levels and so the production rate will eventually be determined by that of the slowest worker, behind whom all the others will accumulate. The production rate will remain suboptimal even if faster workers are allowed to preëempt
9.2 ORDER-ASSEMBLY BY BUCKET BRIGADE

A slower worker and pass him: The slower worker would have to remain idle until his work station became free again and so the line could not keep all workers busy. Moreover, when workers are asked to perform every task on the line then the learning effect and so realized production rate will be reduced.

9.2.3 Some advantages of bucket brigades

Bucket brigade manufacturing has many attractive properties, including:

- It is a pure pull system, so work-in-process inventory is strictly controlled.
- It requires no special material handling system because the workers themselves carry the items from station to station.
- Because the line can be made self-balancing, it does not require accurate measurement of task times and so can avoid some of the expense of time-motion studies.
- It is consistent with other trends in manufacturing: For example, it exploits the advantages of work teams and the grouping of technology into cells.
- The protocol is simple and identical for each worker: Workers are not confused as to what task to perform next and management need not intervene to keep work flow balanced and production rate high.

Bucket brigade manufacturing seems most appropriate when:

- All the work is based on a single skill. This ensures that workers can move among the stations to where the work is, without worrying about whether they can do the work there. It also allows workers to be ranked by a single score, their velocity along the production line, so that the line can be made self-balancing. Economic forces ensure tend to move production lines in this direction, in which the primary worker skills are simple dexterity and enthusiasm.

- A worker can move easily among stations and can easily take over work in process. This ensures that the bucket brigade protocol does not introduce additional wasted time to pass work.

- Demand for the products varies significantly. Bucket brigade manufacturing can more easily track changeable demand because cross-training of workers and low work-in-process inventory mean flexibility of configuration, and short production lead times. In addition, a bucket brigade line can be configured quickly: The assignment of tasks to stations need not be carefully balanced because the movement of the workers balances the line; this reduces the time required to lay out a new line and so shortens changeovers. Finally, because the line is self-balancing, production rates are easily adjustable by simply adding or removing workers from a team.
9.3 Bucket brigades in the warehouse

In many high-volume distribution warehouses, fast moving items are picked from cases stored in a type of shelving called *flow rack*. Within each bay (section of storage) are shelves with rollers and the shelves are tilted to bring the cases forward.

The bays of flow rack are arranged in aisles and a conveyor system runs down each aisle. The *start of an aisle* is the end that is upstream with respect to the movement of the conveyor. For clarity we will describe a single-aisle of flow rack. (Even when there are multiple aisles of flow rack, each aisle is generally operated as an independent module within the warehouse)

An *order* is a list of items for a single customer together with quantities to be picked. It is typical that orders are released in a batch each day to the picking operation. Then each order is picked by “progressive assembly”: The order is picked by no more than one person at a time and the items are accumulated as the order is picked (rather than picking all orders simultaneously and sorting the items afterward).

Paperwork describing orders to be picked waits at the start of the aisle. Each order sheet lists the items and quantities to be picked in the sequence in which items will be encountered along the aisle. The first picker takes the next order sheet, opens a cardboard carton, and slides it along the passive lane of the conveyor as he moves down the aisle picking the items for that order. At some point the second picker takes over and continues picking that order while the first picker returns to the start to begin the next order. When the order is complete the carton(s) are pushed onto the powered portion of the conveyor, which takes them to the packing and shipping department.

There are several ways of coordinating the pickers. One way is to divide the bays into regions and to ask each picker to work within an assigned region: Worker 1 is responsible for picking all items lying within bays $b_1, \ldots, b_1$; worker 2 is responsible for picking all items lying within bays $b_1 + 1, \ldots, b_2$; and so on.

In designing such order-picking systems managers try to balance the expected work among the pickers during the each picking period. The trouble with this is that it balances the work only *on the average over the picking period*, which means only that everyone will have performed the same total number of picks — yet the line can have been significantly out of balance from order to order!

The order-picking system will constantly seek balance if configured as a bucket-brigade with pickers sequenced from slowest to fastest. However, there is an important difference here: Unlike manufacturing the “items” produced on this line (that is, orders picked) are *not identical* and in fact are best modeled as “random”. For example, one might think of each sku $i$ in the warehouse as being in the next order with probability $p_i$ independently of all other sku’s. Because of this, the system converges to a state of balance in a stochastic sense. This is still an improvement over a static balance because:

- It constantly seeks balance from order to order and so will be out of balance much less often and therefore it will be more productive.
- It spontaneously adapts to disruptions and seasonalities.
- It does not require anyone to compute a balance.
Figure 9.4: Average pick rate as a fraction of the work-standard. Zone-picking was replaced by bucket brigade in week 12. (The solid lines represent best fits to weekly average pick rates before and after introduction of the bucket brigade protocol.)

These advantages have been dramatically illustrated in the national distribution center of a major chain retailer for whom we implemented a bucket brigade style of order-picking. After changing to the bucket brigade protocol, their productivity, measured in average number of picks per person-hour, increased over 30% [4], while reducing need for management intervention (Figure 9.4). This was achieved at essentially no cost, and in particular, with no change to the product layout, equipment, or control system (except to render parts of the latter unnecessary).

Previously, work on this line had been assigned by a computer-based model of work content that was run each night preceding picking. Such a model cannot be accurate because

- It cannot economically account for all the relevant detail that determines work content, such as:
  - location, which might be at waist level or on an inconveniently high shelf.
  - shape and weight, which might make an item easy to grab or hard to handle.
  - velocities of the workers, who can range from 50–150% of standard.
  - distribution of locations: One worker might have her picks distributed over three bays while another has as many picks distributed over five bays.
  - additional work such as disposing of empty containers, sealing a full tote and opening another, prepping an SKU, reaching to pull additional stock to the front of the flow rack, and so on.
  - economies of scale: most SKU’s picking two units is less than twice the work of picking one unit.
Figure 9.5: Distribution of average pick rate, measured in 2-hour intervals before (clear bars) and after bucket brigades (shaded bars). Under bucket brigades production increased 20% and standard deviation decreased 20%.)

- Even though it might appear balanced on average, the allocation of work can nevertheless be quite unbalanced for every order.

- A static balance cannot adjust to unforeseen events such as equipment malfunction, employee absences, and so on.

Because the model of work content was inaccurate, as all such must be, considerable management time was devoted to adjusting the allocation of work during the day. (In fact, the retailer dedicated a manager to this.) The bucket brigade protocol has made this centralized managerial effort unnecessary — yet still results in better performance.

Figure 9.5 shows average pick rates at 2-hour intervals at a major US distributor of recorded music. After conversion to bucket brigades the average pick rate increased by 20% and the standard deviation of pick rate decreased by 20%; thus bucket brigades were both more productive and more predictable, which made it easier to staff.

9.4 Summary

One of the advantages of bucket brigades is that they are so flexible. As long as you are careful not to destroy the self-balancing mechanism they can be adapted to many different situations [2].

The ideas behind bucket brigades are simple:

- Abolish rigid assignments of work, which prevent people from going to where the work is.
9.4. **SUMMARY**

- Sequence the workers from slowest to fastest to make a “pull system” that is self-organizing.

- Amplify the technical improvements of bucket brigades by emphasizing teamwork.

The result is to make the assembly line into an analog computer. We program this computer by sequencing the workers from slowest-to-fastest. There is no need to input data to this computer because the task times are “read” by doing them. The output is the optimal allocation of work.
9.5 Questions

Question 9.1  Consider a bucket brigade line that assembles orders along an aisle of flow rack. Each order is distributed fairly evenly along the aisle. The picking is to be done by three workers that have the following work rates: worker $A$ can pick a typical order entirely by himself in twelve minutes; worker $B$ in ten minutes; and worker $C$ in 5 minutes.

- What is the production rate of the line, measured in orders per hour, if the workers are sequenced from slowest to fastest?
- What fraction of each order will be produced by each of the workers?
- What is the production rate if the workers are sequenced from faster to slower?

Question 9.2  How does low pick density affect the effectiveness of bucket brigades in order-picking?

Question 9.3  How are the operation and throughput of a bucket brigade affected if the production line requires a mix of skills, rather than a single skill?

The remaining questions are rather more difficult.

Question 9.4  Suppose the fastest-moving sku’s in an aisle of flow rack are concentrated at the beginning of the aisle (with respect to material flow). How might this affect the operation and throughput of order-picking by bucket brigade? What if the fastest-moving sku’s are at the end of the aisle?

Question 9.5  How does variability of work content from order-to-order affect the performance of a bucket brigade?

Question 9.6  What is the throughput of a bucket brigade if the workers are sequenced other than slowest-to-fastest?
Chapter 10

Warehouse activity profiling

A warehouse is a complicated and busy place and it can be hard to get a accurate sense of what is happening. Warehouse activity profiling is the careful measurement and statistical analysis of warehouse activity. This is a necessary first step to almost any significant warehouse project: Understand the customer orders, which drive the system.

10.1 Basics

There are several simple statistics that are the first things to learn about a warehouse. Each gives some hint as to the economics of that warehouse; but these are to be treated carefully because many are simple averages and so can be misleading. Their primary advantage is to summarize the warehouse environment succinctly (but at the cost of hiding much complexity).

The key facts to learn include the following.

- Area of warehouse (a larger warehouse will require either more labor or more equipment to move product)
- Average number of shipments received in a day (more shipments mean a larger receiving dock and/or more labor)
- Average rate of introduction of new sku’s (it is difficult to maintain a rational storage policy when the population of sku’s changes quickly)
- Average number of sku’s in the warehouse (a rough indicator of complexity of work)
- Average number of orders shipped in a day (more shipments mean a larger shipping dock and/or more labor)
- Average number of lines (sku’s) per order (very large orders can be picked more efficiently)
Average number of units (pieces, cases) per line

Number of order-pickers devoted to pallet movement, to case-picking, and to broken-case picking (suggests where to look for opportunities to reduce operating expenses, which are primarily due to order-picking)

Seasonalities

10.2 A closer look: ABC analysis

It is a truism of engineering that only a few things within any operation account for most of the activity. This is encoded in folklore by various rules-of-thumb, such as 80-20 rules (for example, “Twenty percent of the sku’s account for 80 percent of the activity”); or in ABC analysis, which simply classifies sku’s as A (the small fraction of sku’s that account for most of the activity), B (moderately important), or C (the bulk of the sku’s but only a small portion of the activity).

One of the first things to know about any warehouse is what sku’s matter. This is usually a simple matter of ranking the sku’s by various criteria. This helps reveal the contours of the economic terrain within the warehouse.

It is a popular misconception that an ABC analysis refers exclusively to the ranking of sku’s by dollar-volume, which is dollars/year in sales of each sku. This is merely one of many useful ways of looking at the activity of a warehouse. In fact, dollar-volume will not be of much interest to us because it is a financial perspective, while we are interested mainly in efficient warehouse operations. Consequently we will want to see the extent each sku consumes resources such as labor and space.

Frequently, an ABC analysis yields surprising results. For example, here are three different views of the activity at the national distribution center of a large retail drugstore chain. First, let us see which sku’s accounted for the most cases moving through the warehouse. This would be of interest to the receiving, put-away, and restocking operations because each case must be handled separately to put it on a shelf. It also might reveal what is flowing in greatest quantity along a conveyor in the warehouse. Table 10.1 gives the ten most important sku’s by number of cases moved. Note that sku’s with relatively few pieces per case, such as the number 1 item, can appear on this list even though its total sales (pieces) are only moderate. Effects like this sometimes make the results of ABC analysis surprising.

Most of the labor in a warehouse is devoted to order-picking and so it is useful to rank sku’s by the number of times they were picked during some recent interval, such as in Table 10.2:

Finally, consider the number of pieces sold of each (Table 10.3). This is of interest because each piece must be handled by a sales clerk ringing up merchandise in a retail store. Surprisingly, the ten busiest sku’s with respect to pieces sold are almost all baseball cards and microwave popcorn. It seems that much retail labor is devoted to handling these.

It is difficult to avoid making up stories to explain such surprising lists.
### 10.2. A CLOSER LOOK: ABC ANALYSIS

#### Table 10.1: Top ten items of a chain of retail drug stores, as measured in number of cases moved

<table>
<thead>
<tr>
<th>SKU</th>
<th>Pieces/Case</th>
<th>Pieces</th>
<th>Cases</th>
<th>Picks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UL SLIMFAST BONUS CHOC ROYALE</td>
<td>6</td>
<td>3,085</td>
<td>514.17</td>
</tr>
<tr>
<td>2</td>
<td>BANDAID FAMILY TWIN PACK</td>
<td>12</td>
<td>4,488</td>
<td>374.00</td>
</tr>
<tr>
<td>3</td>
<td>SATHERS PIXY STIX</td>
<td>12</td>
<td>4,320</td>
<td>360.00</td>
</tr>
<tr>
<td>4</td>
<td>GEMINI VIDEO TAPE T-120</td>
<td>24</td>
<td>7,260</td>
<td>302.50</td>
</tr>
<tr>
<td>5</td>
<td>HOUSE BRAND ASPIRIN 5 GR.</td>
<td>12</td>
<td>3,144</td>
<td>262.00</td>
</tr>
<tr>
<td>6</td>
<td>HOUSE BRAND COMPLETE ALLERGY CAPS</td>
<td>24</td>
<td>5,850</td>
<td>243.75</td>
</tr>
<tr>
<td>7</td>
<td>ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
</tr>
<tr>
<td>8</td>
<td>HOUSE BRAND PAIN REL CAPLETS 500MG</td>
<td>24</td>
<td>5,604</td>
<td>233.50</td>
</tr>
<tr>
<td>9</td>
<td>HOUSE BRAND GESIC</td>
<td>24</td>
<td>5,562</td>
<td>231.75</td>
</tr>
<tr>
<td>10</td>
<td>SATHERS S/F ASST SOUR MIX</td>
<td>12</td>
<td>2,520</td>
<td>210.00</td>
</tr>
</tbody>
</table>

#### Table 10.2: Top ten items of a chain of retail drug stores, as measured by the number of customer requests (picks)

<table>
<thead>
<tr>
<th>SKU</th>
<th>Pieces/Case</th>
<th>Pieces</th>
<th>Cases</th>
<th>Picks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
</tr>
<tr>
<td>2</td>
<td>BEACH BAG SET</td>
<td>6</td>
<td>815</td>
<td>135.83</td>
</tr>
<tr>
<td>3</td>
<td>ACT II MICRO LITE BUTTER</td>
<td>144</td>
<td>21,276</td>
<td>147.75</td>
</tr>
<tr>
<td>4</td>
<td>HOUSE BRAND PAIN REL CAPLETS 500MG</td>
<td>24</td>
<td>5,604</td>
<td>233.50</td>
</tr>
<tr>
<td>5</td>
<td>ACT II MICRO WHITE CHEDDAR</td>
<td>120</td>
<td>15,870</td>
<td>132.25</td>
</tr>
<tr>
<td>6</td>
<td>HOUSE BRAND COMPLETE ALLERGY CAPS</td>
<td>24</td>
<td>5,850</td>
<td>243.75</td>
</tr>
<tr>
<td>7</td>
<td>HOUSE BRAND OINTMENT TRIPLE ANTIBIO</td>
<td>144</td>
<td>4,776</td>
<td>33.17</td>
</tr>
<tr>
<td>8</td>
<td>WRIGLEY PLEN-T-PAK BIG RED</td>
<td>192</td>
<td>12,792</td>
<td>66.62</td>
</tr>
<tr>
<td>9</td>
<td>WRIGLEY PLEN-T-PAK DOUBLEMINT</td>
<td>192</td>
<td>14,736</td>
<td>76.75</td>
</tr>
<tr>
<td>10</td>
<td>UL SLIMFAST BONUS CHOC ROYALE</td>
<td>6</td>
<td>3,085</td>
<td>514.17</td>
</tr>
</tbody>
</table>

#### Table 10.3: Top ten items of a chain of retail drug stores, as measured by the number of pieces sold.

<table>
<thead>
<tr>
<th>SKU</th>
<th>Pieces/Case</th>
<th>Pieces</th>
<th>Cases</th>
<th>Picks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UPPER DECK BASEBALL LOW#1992</td>
<td>432</td>
<td>70,524</td>
<td>163.25</td>
</tr>
<tr>
<td>2</td>
<td>ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
</tr>
<tr>
<td>3</td>
<td>SCORE 92 BASEBALL SERIES II</td>
<td>720</td>
<td>25,344</td>
<td>35.20</td>
</tr>
<tr>
<td>4</td>
<td>ACT II MICRO LITE BUTTER</td>
<td>144</td>
<td>21,276</td>
<td>147.75</td>
</tr>
<tr>
<td>5</td>
<td>TOPPS 92 WAX PACK BASEBALL</td>
<td>720</td>
<td>18,684</td>
<td>25.95</td>
</tr>
<tr>
<td>6</td>
<td>ACT II MICRO WHITE CHEDDAR</td>
<td>120</td>
<td>15,870</td>
<td>132.25</td>
</tr>
<tr>
<td>7</td>
<td>WRIGLEY PLEN-T-PAK DOUBLEMINT</td>
<td>192</td>
<td>14,736</td>
<td>76.75</td>
</tr>
<tr>
<td>8</td>
<td>ACT II MICRO NATURAL</td>
<td>144</td>
<td>13,284</td>
<td>92.25</td>
</tr>
<tr>
<td>9</td>
<td>WRIGLEY PLEN-T-PAK BIG RED</td>
<td>192</td>
<td>12,792</td>
<td>66.62</td>
</tr>
<tr>
<td>10</td>
<td>HERSHEY REESE PEANUT BUTTER CP</td>
<td>432</td>
<td>12,708</td>
<td>29.42</td>
</tr>
</tbody>
</table>
CHAPTER 10. WAREHOUSE ACTIVITY PROFILING

Table 10.4: Top ten office products measured by customer requests

<table>
<thead>
<tr>
<th>SKU</th>
<th>Wt</th>
<th>Pieces</th>
<th>Total Wt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CRTDG,TONER,3035,4045,BK</td>
<td>874.81</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>FLDR,LT,11PT,SGL,1/3MA10330</td>
<td>6.38</td>
<td>5812</td>
</tr>
<tr>
<td>3</td>
<td>PPR,TW,25%RAG,8.5X11,20#,WE</td>
<td>5.27</td>
<td>5350</td>
</tr>
<tr>
<td>4</td>
<td>CARD,INDEX,CNT,3X5,5C/PK</td>
<td>137.25</td>
<td>156</td>
</tr>
<tr>
<td>5</td>
<td>POCKET,FLE,9.5X14.75,3.5,RR</td>
<td>5.78</td>
<td>3534</td>
</tr>
<tr>
<td>6</td>
<td>FLDR,LT W/2B FST/150L-13</td>
<td>5.06</td>
<td>3930</td>
</tr>
<tr>
<td>7</td>
<td>FLDR,LG,11,SGL,1/3MA 15330</td>
<td>8.14</td>
<td>1994</td>
</tr>
<tr>
<td>8</td>
<td>PROTECTOR,SURGE,6OUT,6’,PTY</td>
<td>87.04</td>
<td>156</td>
</tr>
<tr>
<td>9</td>
<td>FOLDER,LTR,2 PLI,STRT,24110</td>
<td>8.12</td>
<td>1662</td>
</tr>
<tr>
<td>10</td>
<td>FASTENER,P/S,2/68220</td>
<td>4.85</td>
<td>2662</td>
</tr>
</tbody>
</table>

Table 10.5: Top ten office products by weight, a measure of shipping costs

We find similar surprises in examining activity at a wholesale distributor of office products, for whom the ten most frequently requested sku’s were as shown in Table 10.4.

Notice that the ABC distribution for office products is not strongly skewed (that is, the number of picks falls off relatively slowly as you move down the list). This is a reflection of the maturity of the product and is typical of product movement in hardware and staples. In contrast, the ABC analysis of fashion products can be extraordinarily skewed; for example, the top-selling 100 music CD’s from a population of 100,000+ may account for 25% of all sales.

If we examine the same population of office products by total weight sold, we get a clue as to which sku’s account for most of our shipping costs, which are based most strongly on weight (Table 10.5).

10.3 Universal warehouse descriptors

You might have noticed that all the data presented in this chapter shares the general feature that activity falls off quite rapidly after the first, busiest sku’s. Is there some
universal law underlying the distributions that give rise to the idea of ABC analysis?  

10.4  A more detailed look

To design a new warehouse, retrofit an existing warehouse, or improve warehouse operations requires detailed understanding of the workload in the facility. One must analyze the patterns of customer orders and how this determines the workload within the facility.

10.4.1  Data sources

There are three main types of data required to support profiling: data pertaining to each sku, data pertaining to customer orders, and data pertaining to locations within the warehouse.

Sku data

Useful information to gather about each sku include

- A unique ID that distinguishes it from all other sku’s, which allows us to connect this data with that from other sources
- A short text description, which is useful in validation and error checking
- Product family, which may have implications for storage and/or handling. These tend to be particular to an industry and so require knowledge of the context. For example, product families for a drug store chain might include hair care products, dental products, shaving products and so on, which are displayed together at the retail store. For a grocery distributor product families might include dry goods, dairy, produce, refrigerated, frozen and so on. For a candy distributor product families might include chocolate (sensitive to heat), mint-flavored candies (odiferous), and marshmallow (light and tends to absorb the smells of its neighbors), and so on. For an apparel distributor product families might include garment type, mill, style, color, or size. Note that a sku might be in more than one product family.
- Addresses of storage locations within the warehouse. This might include zone, aisle, section, shelf and position on the shelf.
- For each location at which this sku is stored
  - Scale of the storage unit, such as pallets or cases. This is useful in validation and error checking
  - Physical dimensions of the storage unit (length, width, height, weight), which are useful in understanding space requirements.
– Scale of the selling unit, such as cases or pieces, which is useful for validation and error-checking
– Number of selling units per storage unit. This could be 1.

• Date introduced, which helps identify sku’s that may be underrepresented in activity because newly introduced

• Maximum inventory levels by month or week, which helps determine how much space must be provided for this sku

It is particularly important to understand the conventions the warehouse uses to distinguish among different types of storage units and selling units. For example, the word “case” is often called by other names (“carton”, “box”) and, depending on its use, can have substantially different meanings for a warehouse. For example, a vendor may ship a case that contains several inner packs each of which contains several boxes each of which contains pieces (Figure 2.1). A standard example is an office products distributor supplying a standard type of ball point pen. The manufacturer may supply the product 12 pieces to a box (as you find it in the store), 12 boxes to an inner pack (stored in a thin carton container), and 4 inner packs to a case for a grand total of 576 pens in the vendor’s case or shipping unit. While each of the terms each, box, inner pack, case, shipping unit are commonly used, there is no convention as to which level of packing they apply.

It is important to understand how this packing data is stored in the database. Often, the retail customer (of the distributor) is required to purchase an integer multiple of a selling unit. For the pen the selling unit may be an each, which means that a customer can actually order less than a full box of 12 pens. In the database this information may be stored as either the number “1” or by the symbol “EA” for each. If the customer is required to purchase boxes, then the selling unit may be listed as “12” or “box”. Now suppose the database records a customer purchase of 12, which appears on the order picker’s pick ticket? What exactly does this mean: 12 pens (1 box) or 12 boxes? If you think it means 12 pens when in fact it means 12 boxes, you would be underestimating demand by a factor of 12. If the manufacturer sells the pens as eaches, accounting will record transactions in units of a single pen, notwithstanding the restriction on the outbound side on how it is to be resold. To ensure consistency, one could always record demand in the smallest physical unit, but this would require an order picker to know that 156 = 13 boxes, for example. To facilitate picking accuracy, the demand may be recorded as 13, and the order picker would know that this means 13 boxes. (Of course, if another pen is resold as eaches, this would lead to confusion.) To avoid this confusion, many facilities separate how demand is recorded for accounting purposes from how it is presented on a pick ticket, e.g., the pick ticket should say something like demand = 156 = 13 boxes of 12 each.

There is another reason why such packing data is important. It is always more efficient to store and handle product in some kind of easily-held container as opposed to loose pieces. In the case of the pen this handling unit may be the box of 12 or the inner pack of 12 boxes. For purposes of restocking a shelf of product it would be much easier to restock in units of the inner pack. For purposes of space efficiency on
the shelf and order picking efficiency, it would be better to store the product as boxes neatly stacked.

The sku data may reside in different databases within a company and so it can present a challenge to collect it all. As a general rule, if you think there is the smallest chance that some data may be relevant, collect it! The resulting database is easily manageable: As a rough estimate figure $100 \text{ bytes} \times \text{ the number of sku's}$, which means that $10^4 \text{ sku's}$ can be described in about 1MB.

**Order history**

The order history is simply a concatenation of all the shopping lists submitted by all the customers during the preceding year. It contains the following information.

- Unique ID of this order, to distinguish it from the shopping lists of other customers and from the shopping list of the same customer on another day or later on the same day
- Unique ID of sku, which allows us to look up the sku to see where it is stored
- Customer
- Special handling
- Date/time order picked
- Quantity shipped

For analyzing warehouse operations you have to be careful where you obtain this data. Often this data is from a sales transaction database, which tracks financial rather than warehouse events. Consequently, the date recorded may represent the date the order was placed or when it was printed, not when it was processed in the facility. Similarly, the order that appears in the sales transaction database may actually have been processed at another facility. Generally, though, the information is available somewhere because each day the order pickers had to know what to pick.

There is a simple check as to whether the order data received is approximately correct. Most companies keep track of the lines shipped each day/week/month/year. As a very first validation check, count the number of lines in the database by time period. These numbers should closely match what is recorded. (If you are only obtaining what personnel believe has been shipped, do not be surprised if the numbers obtained through careful processing of a database are substantially different.)

The order data will be the largest file you must manage. As a rough estimate expect about 50 bytes per line. The number of lines can range from 2,000–8,000 lines per day (0.5–2 million lines per year) for a moderately active facility (for example, office product, fine paper, telecommunications distribution) to 10,000–40,000 lines per day (2.5–10 million lines per year) for an extremely active facility (for example, service parts, retail drug) to more than 80,000 lines per day (20 million lines per year) for the most active facilities (for example, pharmaceutical or catalog distribution). Consequently, a year’s order data could exceed 100 megabytes.
10.4.2 Where is the work?

How can we estimate the work in a warehouse? Work is generated by the customer orders; each customer order is a shopping list comprised of “pick lines”; and each pick line generates travel to the appropriate storage location and subsequent picking, checking, packing and shipping the product. Pick lines are then a strong indicator of work; and fortunately there is almost always a historical record of them because they correspond to entries in a sales invoice, which is one of the first pieces of information to be computerized.

We use this information to infer where the work is; that is, how it is distributed among

- Sku’s
- Product families
- Storage locations
- Zones of the warehouse
- Time (time of day, days of the week, weeks of the year, and so on)

Sometimes this is referred to as activity analysis because we examine the activity of each sku, in particular, how many times was it requested; and how much of the sku was sold? Notice that these are two different questions: The first asks “on how many customer orders did this sku appear?”; and the second asks “How many pieces, cases or pallets moved through the warehouse?”.

If a customer requests a quantity that is less than a full case, this is termed a broken-case pick. A broken-case pick can be further classified as to as an inner-pack pick, and so on, depending on how the product is packaged. If a customer requests a quantity that is an integer multiple of a case quantity but less than a pallet (unit) load, this is termed a full-case pick. A pallet pick represents an order quantity that is a multiple of a pallet load quantity. It is not uncommon for a customer request quantity to involve a mixed pick; that is, a pick involving both a broken- and full-case quantity or both a full-case and pallet-load quantity. Broken-case picking requires more time to process than a full-case pick, which takes more time to process than a pallet pick, when normalized by the quantity handled. It is therefore desirable to know how much of each activity is taking place each period.

Here is an application. In fine paper distribution, many sku’s have a considerable amount of both broken and full case picking. Many of the cases are also quite heavy. In one facility, order pickers were making circuitous, inefficient routes so that they could first store the case quantities on the bottom of their cart, and then store the broken case quantities loose on top so as to not crush or damage the loose quantities of paper. It was decided that there should be separate broken and full case picking zones to reduce this inefficiency. A mixture of shelving and flow rack was decided upon for storing the case quantities from which to execute the broken case picking activity. Pallet rack, as before, was to be used to store the full or partial pallet loads from which to execute the full case picking, and from which to restock the broken case picking area. To decide
the appropriate amounts of space (slot types) to each sku in each zone would require a breakdown of each sku’s broken versus full-case picking activity, both by picks and by demand. Those sku’s that had only a very small portion of its activity of one type may not have been assigned to both zones.

Once such activity has been calculated, a variety of Pareto curves can be generated. For example, one can rank the sku’s by popularity (number of requests, picks), sort the list, and then produce a graph that shows the percentage of all picks among the most popular sku’s. For example, Figure 10.1 shows data for a warehouse in which the 5,000 busiest sku’s account for over 75% of the picks and, moreover, complete about 50% of the orders. This means that it would make sense to layout the warehouse to support a fast-pick area (discussed in detail in Chapter 7).

Another useful Pareto curve can be generated by examining the broken-case or full-case picks separately. Similarly, curves can be generated by examining key subsets of sku’s, such as those from one region of the warehouse, or with common seasonality, or from the same product family. Analyzing subsets of sku’s in this way may suggest which areas of the warehouse or which product families stand to have the most opportunity for improvement, or which customers are responsible for the most workload. Finally, one could replace the word “picks” with any type of activity, such as demand. If demand is to be analyzed, it must be normalized into cases or pallets so that there is a common basis for comparison.

There are other distributions that can reveal patterns in the way product moves through the warehouse. For example, Figure 10.2 shows a bird’s eye view of a warehouse in which a darker shading indicates more frequent visits by order-pickers. It is clear from this that popular sku’s are stored all throughout the warehouse and so management can expect greater labor costs and reduced responsiveness because of all the walking necessary to retrieve a typical order.

**Seasonalities**

It is important to understand how the intensity of work varies over time. Most products have some natural “cycle” that repeats over the year or quarter or month or week. The
manager of Allied Foods in Atlanta, Georgia tells us that even dog food has season- 

talities: Demand increases slightly but dependably over the end-of-the-year holidays. 
Particularly in the US, the annual holiday season of roughly November–December is 
by far the busiest time for retail sales and this determines the timing of product flow 
upstream.

One sku with easily predictable seasonality is AA batteries. This is a mature tech-
nology, not subject to fashion or obsolescence, and so demand is fairly steady for most of 
the year . . . except of course for the month of December, when demand almost doubles. 
In fact, most of the demand in December comes on Christmas day and it is concentrated 
mostly at convenience stores, such as 7-11 or Quik-Trip.

Other seasonalities include:

• Office products sell most heavily on Mondays and Fridays, in January and in 
August. Among these, calendars sell most briskly in January, with sales dropping 
until June, when it disappears.

• The two fastest-moving items at Home Depot at Father’s Day are (barbecue) 
grills and (electric) drills. Barbecue grills sell in the spring up to July 4, when 
sales plummet.

Finally, see whether you can guess the seasonalities of the following. (Answers at 
the end of the chapter.)

• Sales of large screen color televisions
• Sales of CD’s and other recorded music
• Consumption of avocados in the US
• Sales of disposable diapers
• Rental of tuxedoes
• Sales of belts
10.4. A MORE DETAILED LOOK

Figure 10.3: Very few customers order more than ten line-items (left), but this accounts for the bulk of the picking (right).

10.4.3 What are the patterns of work?

Here we want to go beyond measuring the quantities of work to understand the patterns of work generated by the customer orders.

If no customer ordered more than one sku, then the preceding activity analysis gives a sufficient view of the warehouse; but this is rarely the case and customers order multiple sku’s. It is then important to understand the patterns in the customer orders.

For example, one indication of inherent work is the average lines per order. When this number is small, say no more than 2, there is an opportunity to batch orders and assign one picker to each batch. When the number of lines per order is in the range of 5–15, some form of zone picking will normally be required, but it may be possible to progressively assemble the order through various zones. As this number gets higher, zone picking will be required, and some type of accumulation will be required, too. If no accumulation occurs, then the customer must be willing to accept multiple packages (or totes) from the distributor. For example, the customer of a retail distribution center will be a retail store that may order small amounts of thousands of sku’s per week.

But, as always, one must beware of averages. It is always more informative to examine the complete distribution of lines per order. It shows the fraction of orders for a single sku, for exactly two sku’s, and so on, as in Figure 10.3. In this example, most orders are for a single line, which suggests some opportunities for efficient handling. For example, if these are mostly back-orders, then they could be cross-docked. If they are rush orders, they could be grouped together into a single batch and then picked in storage sequence.

A related graph is the distribution of picks by order-size. That is, it depicts the fraction of all picks that come from single-line orders, two-line orders, and so on. Because picks are a good indication of work, this shows which types of orders, small or large, contain the most aggregate work.

Here is an application. At a telecommunications facility, workers pushed order-picking carts through a series of zones to progressively assemble orders. Due to space limitations on each cart, the transfer batch was small. An analysis of orders showed that about 10% of the orders were relatively large, for more than 100 lines each, whereas the remaining 90% averaged less than 2 lines per order. It was decided to assign one worker to pick each extremely large order, one at a time, so that the remaining orders could be picked and transferred in larger batches, thereby increasing order-picking efficiency.
110 CHAPTER 10. WAREHOUSE ACTIVITY PROFILING

<table>
<thead>
<tr>
<th>Order ID</th>
<th>Families represented</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>A, B, C</td>
</tr>
<tr>
<td>200</td>
<td>A, B</td>
</tr>
<tr>
<td>300</td>
<td>C, D, E</td>
</tr>
<tr>
<td>400</td>
<td>B, D, E</td>
</tr>
<tr>
<td>500</td>
<td>D, E</td>
</tr>
<tr>
<td>600</td>
<td>A, D, E</td>
</tr>
<tr>
<td>700</td>
<td>B, D, E</td>
</tr>
</tbody>
</table>

It is frequently also useful to generate the distribution of families per order; that is, the fraction of all orders that involve exactly one product family, two product families, and so on. Here is an application. Generally it is helpful to locate sku’s near one another if they tend to be requested together (for example, running shoes and socks; flashlights and batteries). Such product assignment can reduce travel time during order-picking. But which products tend to be picked together? In one paper distribution facility, sku’s were classified into three categories. It was readily verified that customers rarely ordered across categories, so, the warehouse could be divided into three zones, each of which functioned as a smaller and more efficient warehouse. Further inspection of orders revealed that there were typically many lines per order; and that most orders requested the products of no more than two vendors. From this information, a vendor-based stock assignment plan made the most sense: store all sku’s within a vendor in the same area of the warehouse.

Of course, the next question might be: which vendors should be located near one another? This brings us to the concept of family pairs-analysis. For example, consider the following list of multi-family orders and the families involved in each order (Table 10.4.3).

<table>
<thead>
<tr>
<th>Family pairs</th>
<th>AB</th>
<th>AC</th>
<th>AD</th>
<th>AE</th>
<th>BC</th>
<th>BD</th>
<th>BE</th>
<th>CD</th>
<th>CE</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

and it is clear that any order requesting a sku from family D is likely to request a sku from family E as well and so one should consider storing these families near one another.

It is possible to consider triples of product families and beyond, but the possibilities increase exponentially and so the work required quickly outstrips any computing power brought to bear on the analysis.

In one distribution center two product families were moderately correlated but on closer inspection it was realized that one of the product families was very active, while the other product family much less so. However, there was a greater than 98% chance that if an order requested a sku from the less active family it would also request a sku from the more active family. Since the less active family consumed little space, it made sense to store this family next to the larger, more active one.

A whole new picture takes place if one aggregates sku’s by location or zone within a facility. That is, A–E above could represent a zone in the warehouse. The distributions
so obtained represent “order-crossings” and reflect the degree of order-accumulation across zones.

One distribution center progressively assembled orders by picking pieces. The product was in cases stored in shelving across five zones of about three aisles each. A non-powered roller conveyor was used to move the totes from one zone to the next. The conveyor jammed frequently because too many totes were being staged. One way to reduce this congestion is to relocate the stock so that many more orders could be completed entirely within one zone or within neighboring zones. This may be achieved by seeing which families tend to be ordered together.

Finally, it should be noted that when speaking of orders, it is possible to partition the order file into groups of sub-orders. For example, one can consider only that portion of the order pertaining to a zone within the facility. This is especially helpful and required if picking policies are different in different zones.

Redo this section from more abstract point of view of projecting orders onto storage locations, product families, etc.

10.5 Doing it

How do you actually do profiling? In most cases a personal computer is adequate to the task, but it is messy business. The first problem is getting the data.

10.5.1 Getting the data

You will probably be working off-line, with a snapshots from corporate databases. The reason is that you will interfere with daily operations when your queries slow the database that supports the warehouse.

Budget several weeks to complete activity profiling. Most of this time will be spent checking, validating, and purifying the data.

As of this writing, there are hundreds, possibly thousands, of different warehouse management or record-keeping software systems, many of them written in-house. You must appeal to the IT department, which like all IT departments, is overworked and behind schedule. Running the queries to get your data may steal cpu cycles from the computers that are managing the business. Do not expect to be welcomed back, so you might as well ask for copies of every database and sort it out later, on your own.

Time permitting, make a sample request for a small piece of the data (say, one day’s worth) and check that it makes sense. Review the meaning of every data field with people from both operations and from IT. Show them summary statistics you extract from that day of data to see whether they seem reasonable. Check the text description of the biggest, busiest sku’s to see that there are no suspicious results.

Only after successfully sampling the data should you submit your request for the full data dump.

The preferable way to receive the data is on CD-ROM, which means you do not need unusual equipment to read it and you cannot accidentally erase it. Alternative methods include via ftp or on high-volume removable disks.
We find it best to receive the data in some simple, neutral format, such as tab-delimited ASCII (successive fields in a line of data are separated by a tab character). This keeps options open because the data is easily transferable to a range of software tools.

10.5.2 Data-mining

There is no packaged software that is universally applicable to extract important patterns from the data. The essential functionality you will need are the abilities to

- Sort the rows of a table;
- Select a subset of rows of a table, such as all the order-lines from a single day.
- Count distinct entries in a table, such as the number of times each sku appears in the order history;
- Connect the row of one table with a corresponding rows in another table (for which the database jargon is join). An example is connect order-lines (rows in the lines table), through the unique sku identifier, with additional information about each sku (rows in the sku’s table), such as where it is stored.
- Graph results.

It is a fact of life that corporate data resides today mostly in relational databases and so some facility in manipulating them is essential. Many commercial databases support some form of “data-mining” by providing an intuitive querying capability. This is a front end that produces the nearly universal language of relational databases, SQL (Structured Query Language). However, these front ends are proprietary and vary considerably from product to product. In any event, you are likely to have to write some SQL; fortunately it is fairly straightforward. For example, the following command returns the number of lines in the table named “Lines”:

```sql
SELECT COUNT(*) FROM Lines
```

and to show all lines shipped on January 20, 2000 the command is

```sql
SELECT * FROM Lines WHERE DateShipped = '2000-01-20'.
```

To return a view of the total units shipped for each sku, the command is

```sql
SELECT SkuID, SUM(QtyShipped) FROM Lines GROUP BY SkuID
```

and to return the total number of requests (picks) for each sku,

```sql
SELECT SkuID, COUNT(*) FROM Lines GROUP BY SkuID
```

To return a list of of orders and the zone from which each line was requested, the command is
SELECT Lines.OrderID, Skus.Zone FROM Lines, Skus
WHERE Lines.SkuID = Skus.SkuID

An alternative to using a database and writing SQL is to directly program, in some lower-level computer language, the ability to query databases. This is not hard and may be worthwhile for the savings in disk space and running time. Importing ASCII data into a commercial database may inflate its size by a factor of five to ten.

The main thing to look for is a set of tools that are flexible, because each warehouse is so different that it does not seem possible to standardize on a general set of queries. Plan on building or buying a library of queries that you can revise and adapt.

10.5.3 Discrepancies in the data

There will be discrepancies in your data. Expect them and have a strategy for dealing with them. First find them and then document their severity. For example, exactly what percentage of sku’s that appear in customer orders do not appear in the sku database? If it is a small problem, it can be ignored for purposes of warehouse profiling; but the warehouse IT staff should be informed so they can fix this.

Be sure to check with special care the busiest sku’s, locations, times, and so on. Errors tend to appear at the extremes of the ABC distributions. Here are two examples that happened to us. In one case we discovered that a particular style of paper accounted for a huge fraction of all the cubic feet shipped from the warehouse. This surprised our client. On checking further we learned that sales of this paper had been recorded by the sheet, not by the case, as we had assumed. Similarly we were surprised to find that a popular writing pad accounted for exactly zero cubic feet of product shipped. The pad had been measured on an automated device that captured its dimensions as 8.5 inches × 11 inches × 0 inches = 0 cubic inches.

Many problems arise because the order history spans an interval of time; while the file of sku’s may represent a snapshot of one instant. Consequently, you are likely to find sku’s appearing in one database but not in the other.

The importance of cross-checking

You can reduce problems by energetically cross-checking all data and analysis. Get in the habit of asking the same question in different ways or addressing it to different sources (people, databases). Check that the answers are consistent. For example, how many lines does the warehouse ship per day? How many orders? What is the average number of lines per order? Does the math check?

A few handy tools make it easier to cross-check.

- On average the total flow into a warehouse is the same as the total flow out.
- Little’s Law
- Approximation: One tightly-packed 53 foot long trailer holds about 2,000 cubic feet of product, which is about 20 pallets.

Add more
10.5.4 Viewing the data

Readers are strongly encouraged to study E. Tufte’s “The Visual Display of Quantitative Information” and the companion books [15].

Some general principles:

- Use graphs to show large scale patterns, especially when you want to compare with other patterns. And when you want graphs to be compared, draw them to the same scale.

- Use tables for closer looks when the actual numbers are important.

- Scale the data to make it easier to understand. For example, it is better to report average lines per day rather than total lines over the period of study. Most people will find it easier to understand the implications of the numbers.

10.6 Summary

An activity profile is essential to really understand what matters in a warehouse. You can build this from data about the physical layout of the warehouse, the sku’s stored therein, and the patterns of your customer orders. The activity profile will enable you to understand, manage, and improve use of labor, space, and equipment.

Warehouse activity profiling is a special case of “data-mining”, which is simply the rummaging through databases to look for opportunities to improve operations. As in mining for minerals, success depends on having good tools to support the search and on knowing where to look. Therefore you must be comfortable with SQL and databases and you must understand warehouse operations.

10.7 On the lighter side

Everyone who has profiled a warehouse has stories about how hard it was to get the data. One consultant described the following experience. Without checking in advance, a client attempted to e-mail him an enormous data file for analysis. When his mail server rejected it as too large, the client broke the file into 25 (still large) data files and sent them along. Meanwhile, the consultant was working on-site elsewhere and was expecting an urgent e-mail; but whenever he connected to his mail server the large files began to download over his slow phone connection. Calls to the mail service provider did not help. He finally resigned himself to sacrificing his machine for a few days while it received the data files, while he paid the connection charges. To add insult to injury, his mail server had accepted only 15 of the data files before overflowing his disk allotment, so it was all for naught.

An ABC analysis of a grocery distribution center provides an interesting glimpse into US eating habits. Most product moves out of a grocery warehouse in cases and the ten most popular sku’s, as measured by number of cases purchased, reveals where order-picking labor is concentrated (Table 10.6).
10.7. ON THE LIGHTER SIDE

Table 10.6: Top ten items moving from a grocery warehouse, as measured by number of cases

<table>
<thead>
<tr>
<th>SKU</th>
<th>Pieces/Case</th>
<th>Cases Moved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mayonnaise</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>Charcoal</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Sani-cat litter</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2-liter gingerale</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Charcoal briquets (10 lbs.)</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>2-liter cola</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Apple juice</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>Granulated sugar</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>2-liter orange drink</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>Charmin</td>
<td>16</td>
</tr>
</tbody>
</table>

It is not surprising that sku's that are relatively large are over-represented (charcoal, cat litter, 2-liter drinks); but mayonnaise moves half again as many cases as the sku in second place and every case contains twelve jars! Who is buying all that mayonnaise? What are they doing with it? The only clue we could find was that canned tuna was number 11, having moved 22,336 cases, with each case holding 48 cans. Even if all that was turned into tuna fish salad, that is still over 8 (large) jars of mayonnaise per (little) can of tuna. The situation becomes all the more mysterious when you learn that the twelfth most active sku was Miracle Whip, which moved 20,515 12-jar cases!

There are some items that are seasonal but still hard to predict. For example, in Mexico a sweet bread called marinela is eaten when it rains.

The peak seasons are as follows.

- Sales of large screen color televisions: During the two weeks preceding Superbowl. (Source: Mitsubishi Consumer Electronics, Atlanta, Georgia)
- Sales of CD’s and other recorded music: During the two weeks following Christmas, when presumably people received CD or DVD players as gifts. The most popular first purchases are “golden oldies”. (Source: SuperClub Music and Video, Atlanta, Georgia)
- Consumption of avocados in the US: The day at which consumption is highest is Cinco de Mayo, followed by the day of the Superbowl, when presumably it is consumed as guacamole. (Source: Wall Street Journal, January 26, 1999). At neither of these times is consumption high in Mexico.
- Sales of disposable diapers: No reliable seasonalities, but interestingly the rates differ by region of the world. For example, Jim Apple of The Progress Group says that the rate of use in the US is about seven per day per baby but it is only five per day in Europe.
- Rental of tuxedoes: Proms in May and weddings in June. (Source: Mitchell Tuxedo, Atlanta, Georgia)
Sales of belts: Easter in March/April; Father’s Day in June; back-to-school in August; Christmas. (Source: Italian Design Group, Atlanta, Georgia)
10.8 Questions

Question 10.1 Why is number of pick-lines generally a better indicator of labor than quantity-ordered or quantity-shipped?

Question 10.2 In a warehouse activity profile, is it possible for a single product group (for example, mill, style, color, etcetera) to complete more of the orders than any other but yet not have very many picks? Explain.

Question 10.3 Consider a lines-per-order distribution in which 30% of the orders are for single lines; 25% for 2-lines; 20% for 3-lines; 12% for 4-lines; 10% for 5-lines and the remaining 3% for 6-lines. Historically there have never been orders for 7 or more lines. What size order accounts for most of the picking in the warehouse?

Question 10.4 (Project) Describe the warehouse and its activity given by the data set at http://www.isye.gatech.edu/~jjb/wh/book/profile/projects/projects.html. Look for interesting patterns that reveal something potentially useful or interesting about the warehouse. Think especially about how to display the results in succinct, elegant graphical format, as suggested by [15, 16, 17]. 

Amidst your explorations answer the following in particular.

- Describe the sku’s. For example: Where are they? Which are most popular? On what percentage of orders do they appear? How many sku’s are never requested? How does the population of sku’s change over time?

- Describe the work. How much is there and where is it? For example, how is it distributed among sku’s, among zones, among aisles, among bays, among orders, among vendors, among days, etc.?

- Describe the orders. For example: What percentage of orders are for a single line? What percentage are for more than 20 lines? What is the distribution of lines per order? What is the distribution of lines per order within each zone? What percentage of the labor do single-line orders represent? Orders with more than 20 lines? What percentage of orders touch the mezzanine (zones A, B)? Bin-shelving on the floor (zones C, D)? The security zone (E)? Pallet rack (G)? What percentage of orders are completed entirely on the mezzanine (that is, within zones A and B)? Entirely within floor shelving (that is, within zones C and D)? Entirely within the high security area, for valuable items (zone E)? Entirely within pallet rack (zone G)? How many orders are completed with each of the possible pairs of zones?

- What percentage of orders are marred by shipping discrepancies (quantity shipped < quantity ordered)? Where are the discrepancies: In what sku’s, zones, days, and so on?
Chapter 11

Benchmarking warehouse performance
Chapter 12

Crossdocking

Crossdocks are high-speed warehouses.

If an arriving item has already been requested by a customer there is no need to store it as anticipation inventory; instead, the item can move directly from receiving to shipping, without intermediate storage and retrieval. Thus the item can move much more quickly through the facility and the most costly part of warehouse labor can be avoided.

In a high-volume crossdock the turnover times may be measured in hours. To support this velocity of movement, a crossdock may be nothing more than a slab of concrete with a roof and walls punctuated with doors for trailers. Freight is pulled off arriving trailers, sorted and loaded onto departing trailers without intermediate storage.

There is little or no storage provided in a crossdock because items do not stay long enough; but there is generally a lot of material-handling equipment, such as forklifts and pallet jacks, to move freight. Labor is frequently the main cost and it is devoted to unloading incoming trailers, moving the freight to the appropriate outgoing trailers, and loading.

12.1 Why have a crossdock?

The biggest reason to have a crossdock is to reduce transportation costs. This can be achieved by consolidating multiple shipments so that full truck loads can be sent.

The Home Depot is a major retailer and the largest user of Less-than-Truck-Load (LTL) shipping in North America. (LTL means sending shipments that do not fill a trailer and so are not economical to send by themselves. Instead, an LTL freight company consolidates many such shipments and so achieves efficiencies.) At the present writing, LTL costs about twice the cost of Truck Load (TL) shipping, so there is a strong incentive to fill trailers. The Home Depot has begun doing this by having vendors ship full trailers to its crossdock. (The trailers are full because they hold product for many stores.) At the crossdock the product is sorted out for individual stores and consolidated with product from other vendors bound for the same store. The result is that each store has enough freight that it or it and a few close neighbors generate a full truck load.
from the crossdock. The result can be considerable savings.

Additional benefits include less inventory (because all product flows right through) and less labor (because product does not have to be put away and later retrieved).

12.2 Operations

Most crossdocking freight terminals are laid out as long, narrow warehouses with doors around the perimeter. Figure 12.1 illustrates a typical terminal, where the small shaded rectangles represent incoming trailers with freight to be unloaded, and small clear rectangles represent (empty) outgoing trailers. Terminals range in size from fewer than 10 doors to more than 500 doors.

Inside a terminal, a variety of material handling methods is used to transport freight. Forklifts and palletjacks carry heavy or bulky items, and carts transport smaller items. In addition, large terminals may have draglines, which circulate carts around the inside perimeter of the dock.

There are two types of doors in a terminal: receiving, or strip, doors, where full trailers are parked to be unloaded, and shipping, or stack, doors, where empty trailers are put to collect freight for specific destinations. Once established, the designations of these doors do not change, although the trailers parked at them will. A shipping door always receives freight for the same destination. A receiving door may be occupied by any incoming trailer, regardless of its origin or contents.

Arriving trucks may deliver their trailers directly to an unoccupied receiving door; or, if none is available, they may place them in a queue. After the trailer is backed into a receiving door, a worker unloads the freight. After unloading all the items of a shipment onto a cart, the worker walks to the destination trailer and loads the items into that trailer; or he places the cart on the dragline, if the terminal is so equipped. To handle pallet loads, the worker uses a palletjack, or hails a forklift driver, or finds a forklift and delivers the load himself, if union rules permit.

After a trailer has been completely stripped, a driver replaces it with another incom-
ing trailer from the queue of trailers waiting to be stripped. After an outgoing trailer has been filled, a driver replaces it with an empty trailer to be filled with freight for the same destination.

12.3 Freight flow

The patterns of freight flow within a terminal — and therefore the work — are determined by:

**Layout** by which we mean the specification of doors as either receiving or shipping doors and the assignment of destinations to the shipping doors.

**Geometry** The shape of a terminal determines the travel distances between doors and the susceptibility to congestion. (For example, narrow docks tend to be more congested because workers have less room to maneuver.)

**Material handling systems** For example, palletjacks are slower than forklifts, but they may be more available; draglines reduce walking time, but can impede forklift travel.

** Freight mix** For example, terminals having a higher mix of pallet freight require more forklift travel than those receiving a majority of carton freight.

** Scheduling** In real time, the dock supervisor determines freight flow patterns by assigning incoming trailers to receiving doors.

Changing the geometry or material handling systems of a terminal is expensive; changing the freight mix is a marketing decision with implications outside the terminal. The two remaining ways to take work out of the system — change the layout or change the scheduling — are inexpensive.

12.3.1 Congestion

As more freight flows across a dock, congestion increases, which interferes with the flow.

There are several distinct types of congestion on a crossdock:

**Competition for floor space:** Freight may be docked outside a receiving door if, for example, it consists of many unpalletized cartons going to the same shipping door. Then there is an incentive to accumulate it all so that fewer carts must travel to the destination door. On the other hand, freight is very likely to be docked outside a shipping door while the loader figures out how to pack the trailer tightly. When several nearby doors compete for space to dock freight, some invariably interferes with other traffic. At the very least, it takes longer for a worker to maneuver through the shippings of freight.

The effects of docked freight are most severe near the inside corners of the dock, where there is less space per door, as shown in Figure 12.2.
12.4 Design

12.4.1 Size

The first decision in designing a crossdock is "how many doors?".

Generally doors are devoted to one of two types of trailers:

- Incoming, from which freight must be removed; and
- Outgoing, in which freight must be loaded

It is easier to unload than to load. A loader must try to get a tight pack and so may have to dock freight and this double-handling slows him down. A good rule of thumb is that it takes twice as much work to load a trailer as to unload one.

To achieve frictionless flow, you must match the flow into the dock with the flow out of the dock. You can have twice as many outgoing doors as incoming doors; or you can assign pairs of workers to load each trailer and so have equal numbers of incoming and outgoing doors.

Note, however, that crossdocks with many doors are generally less efficient than crossdocks with fewer doors. The reasons are as follows. A door can only have a few near neighbors on a dock and so a dock with more doors means that each door is likely
to have few more near neighbors but many more distant neighbors. This means that in general freight must move farther across a large dock. Consequently, labor costs are generally higher at larger docks.

An additional factor is that on larger docks more freight flows past the central doors, which are the most important because they tend to be close to many doors. In fact, the total flow of freight past a centrally-located door tends to be proportional to the square of the total number of doors. Therefore a dock with twice the doors tends to have 4 times the congestion in front of its central doors, which diminishes their value.

This follows from the following simple model: Imagine a rectilinear dock as a line with \( 2n \) doors (numbered from left to right), and assume that equal amounts of freight move between every pair of doors. Then the flow into any door is of intensity \( O(n) \). But the total flow passing the area between door \( i \) and \( i + 1 \) is \( i(2n - i) \), which means that the greatest total flow passes by the middle of the dock, door \( n \), past which flows \( O(n^2) \) units. But these central doors are exactly those that are nearest to most other doors and therefore are the best locations! Thus, as a dock design grows in length, the lengthwise traffic past the central doors increases rapidly while traffic directly across the dock remains unchanged. Increased traffic means congestion, which helps explain why docks can lose their efficiency as they grow. There are few docks larger than about 200 doors. Most are 80–120 doors long.

Do not forget to allow enough parking space in the yard for two trailers for every door. This means that for each origin or destination you can have a trailer at the door plus one full and one empty in the yard. This helps you handle surges in freight flow.

### 12.4.2 Geometry

What is a good shape for a crossdock? In general, one wants to enable efficient flow of freight from incoming trailers to outgoing trailers.

Typically, a crossdock is a long rectangle, with doors for trailers around it. The capacity of a dock is increased if it has many doors, but without being too close together so that trailers (outside) or freight (inside) interfere with one another.

A typical dock is generally around 120 feet wide. This is to allow freight to be staged on the floor. A standard (large) trailer is 48 or 53 feet long and a “pup” is 28 feet long; all are 9 feet wide. The width of the dock should include enough space for the trailer on each side of the dock to stage its freight (about 100 feet total) plus allow space for travel along the length of the dock (for example, two aisles, each 10 feet wide). We have seen docks as narrow as 80 feet, but this is practical only when it is possible to avoid staging most freight, such as when the material is palletized and also easily stackable and may be loaded in any order. If a dock is much wider than this, it just adds to the travel time to move the product from incoming trailer to outgoing trailer.

A dock does not have to be a rectangle. We have seen docks in the shapes of an L, U, T, E, and H. But every corner in a dock reduces effective capacity:

- On the outside of a corner you lose floor space per door on which to dock freight. This increases congestion on the dock, which interferes with the flow of freight.
12. CROSSDOCKING

On the inside of a corner, you lose door positions because trailers will interfere with each other in the yard. Because doors are lost, the dock must be longer to accommodate a given number of doors, which means that on average freight will have to travel farther to cross the dock. Thus, for example, freight has to travel farther to cross an H-shaped dock, with four inside corners, than to cross an L-shaped dock. (Because the door positions will be lost anyway, inside corners are a good place to locate administrative spaces or hazardous materials storage.)

It is hard to make generalizations independent of specific bills of lading; but in general an L-shaped crossdock is inferior: It incurs the costs of one inside and one outside corner but without getting anything in return. The result is that freight must travel farther because the dock must be longer from end to end to make up for lost doors at the inside corner. Furthermore, there is congestion at the outside corner.

The same observations hold even more strongly for a U-shaped dock.

A +−-shaped or a T-shaped dock also incur corner costs but they have a compensating benefit: The longest distance from door-to-door is less than that for an L-shaped or L-shaped dock with the same number of doors.

One can perform this comparison methodically, based not just on distances across the dock but on intensities of freight flow. For example, the comparisons of Figure 12.3 were generated based on aggregate freight flows for The Home Depot.

12.5 Trailer management

One can reduce labor costs in a crossdocking freight terminal by parking incoming and outgoing trailers so that freight can be efficiently moved across the dock. For example, if much of the freight flowing through the terminal is bound for Miami, the Miami trailers should probably be parked in a convenient location. The challenge is to formalize the notion of “convenient”; then labor-reducing door assignments can be made with optimization models based on the geometry of the terminal, the material handling systems within, and the mix of freight passing through.

12.6 Resources

More detailed technical treatment of these issues can be found in [5, 6].

K. R. Gue maintains website devoted to crossdock design and operation: http://web.nps.navy.mil/~krgue/Crossdocking/crossdocking.html.

12.7 Questions

Question 12.1 Crossdocks remove two of the fundamental warehouse activities: What are they and what enables crossdocks to omit these activities?
Figure 12.3: For data representative of The Home Depot an L-shaped dock was never preferable to an I-shaped dock, no matter the size. A T-shaped dock begins to dominate for docks of about 170 doors and a +-shaped dock around 220 doors.
Question 12.2 For what reasons is freight likely to be docked at the door of an outgoing trailer? An incoming trailer?

Question 12.3 Explain the cost of a corner in a crossdock.

Question 12.4 Where are the most convenient doors in a rectangular crossdock?

Question 12.5 What is a typical width for a crossdock? Explain the logic behind this width?

Question 12.6 Give two reasons why an I-shaped dock is almost certainly more efficient than an L-shaped dock for the same number of doors.

Question 12.7 Crossdocks are generally set up so that outside (tractor) traffic circulates counterclockwise. Why?

Question 12.8 The following figure shows the layout and flows on a crossdock on which freight moves mostly by forklift truck and by a dragline running counterclockwise. Each solid mark indicates a door reserved for arriving trailers and each bar represents the amount of freight bound for the destination for which that door is reserved. Describe three problems the following crossdock will likely experience and explain why.
Chapter 13

Trends
Bibliography


Index

activity analysis, 106
batch, 14
broken-case pick, 106
case, 47
crossdock
defined, 121
cube-per-order index, 62
deadheading, 31
Equal Space Allocation, 54
Equal Time Allocation, 54
family-pairs analysis, 110
fast-pick area, 49
floor storage
pallets, 20
flow time
cycle time, 15
footprint, 38
forward pick area, 49
full-case pick, 106
grab, 13
honey-combing
in pallets, 20
join, in a relational database, 112
lanes
pallet, 20
line, pick-line, 13
LTL, 121
mixed pick, 106

pallet
defined, 20
pallet pick, 106
pick density, 13
pick face, 13
primary pick area, 49
reserve storage, 49
shipment integrity, 15
sku density, 14
slot, 60
Slotting, 81
SQL, Structured Query Language, 112
TL, 121
transfer batch, 109
value-added processing
VAP, 17
viscosity of a sku, 58
Warehouse Management System
WMS, 13