Contents

Preface

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

I Issues, equipment, processes .......................... 1

1 Warehouse rationale ............................... 3

1.1 Why have a warehouse? .......................... 3
1.2 Types of warehouses ............................. 5

2 Material flow ............................................. 7

2.1 The fluid model of product flow ..................... 7
2.2 Units of handling ..................................... 8
2.3 Storage: “Dedicated” versus “random” ............. 8
2.4 The warehouse as a queueing system .............. 12
   2.4.1 Extensions ..................................... 12
2.5 Questions ............................................. 14

3 Warehouse operations .................................. 15

3.1 Receiving ................................................ 16
3.2 Put-away ................................................. 16
3.3 Process customer orders ............................. 17
3.4 Order-picking ......................................... 17
   3.4.1 Sharing the work of order-picking .......... 19
3.5 Checking and packing ............................... 20
3.6 Shipping ............................................... 20
3.7 Summary ............................................... 21
3.8 More .................................................. 21
3.9 Questions ............................................. 23
## 8 Pieces: Design of a fast pick area

8.1 What is a fast-pick area? ............................................ 71
8.2 Estimating restocks .................................................. 72
8.3 How much of each sku to store in the fast-pick area? ........... 73
8.4 Allocating space in the fast-pick area ............................ 73
  8.4.1 Two commonly-used storage strategies ....................... 75
  8.4.2 Managerial implications ...................................... 80
  8.4.3 Minimum and maximum allocations ............................ 80
8.5 Which skus go into the fast-pick area? ........................... 81
  8.5.1 Priority of claim for storage in the fast-pick area ........... 86
8.6 Additional issues .................................................... 89
  8.6.1 Storage by family ............................................. 89
  8.6.2 Accounting for safety stock .................................. 90
  8.6.3 Limits on capacity ............................................. 90
  8.6.4 Accounting for on-hand inventory levels .................... 91
  8.6.5 Setup costs ..................................................... 91
8.7 Limitations of the fluid model .................................... 92
8.8 Size of the fast-pick area ......................................... 93
  8.8.1 How large should the fast-pick area be? ....................... 93
  8.8.2 How can the fast-pick area be made larger? .................. 95
8.9 Multiple fast-pick areas ............................................ 96
8.10 On the lighter side .................................................. 97
8.11 Summary ............................................................. 97
8.12 Questions ............................................................ 99

## 9 Pieces: Detailed slotting

9.1 Case orientation and stack level .................................... 107
9.2 Packing algorithms .................................................. 108
  9.2.1 Next Fit ......................................................... 110
  9.2.2 First Fit ......................................................... 111
  9.2.3 Other issues .................................................... 111
9.3 Questions ............................................................ 113

## III Order-picking

### 10 Piece-picking by ‘bucket brigade’

10.1 Introduction ......................................................... 117
10.2 Order-assembly by bucket brigade .................................. 118
  10.2.1 A model ........................................................ 118
  10.2.2 Improvements that are not .................................. 122
  10.2.3 Some advantages of bucket brigades ......................... 124
10.3 Bucket brigades in the warehouse ................................ 125
10.4 Summary ............................................................. 128
10.5 Questions ............................................................ 129
CONTENTS

11 Pick-paths 131
  11.1 The problem of pick-path optimization 131
  11.2 Heuristic methods of generating short pick paths 132
    11.2.1 Path outlines 134
    11.2.2 Product placement 136
  11.3 Pick-path optimization 137
    11.3.1 How to take advantage of optimization 143
    11.3.2 How much is optimization worth? 144
  11.4 Summary 144
  11.5 Questions 146

IV Special topics 149

12 Crossdocking 151
  12.1 Why crossdock? 151
  12.2 Operations 152
  12.3 Freight flow 153
    12.3.1 Congestion 153
  12.4 Design 155
    12.4.1 Size 155
    12.4.2 Geometry 156
  12.5 Trailer management 158
  12.6 Resources 160
  12.7 Questions 160

V Measuring warehouse efficiency 163

13 Activity profiling 165
  13.1 Basics 165
  13.2 Warehouse activity profiling 166
    13.2.1 ABC analysis 166
    13.2.2 Statistical analysis 169
    13.2.3 Doing it 178
    13.2.4 Visualization 183
  13.3 Benchmarking 184
  13.4 Summary 184
  13.5 On the lighter side 184
  13.6 Questions 186

14 Benchmarking 187
VI Miscellaneous

15 Warehousing around the world
   15.1 North America .............................................. 191
   15.2 East Asia .................................................. 191
      15.2.1 India ................................................... 192
      15.2.2 China .................................................. 193
      15.2.3 Singapore, Hong Kong, Japan ......................... 193
   15.3 South and Central America .................................. 196
   15.4 Europe ..................................................... 196

16 In preparation .................................................. 197
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>If two pipes have the same rates of flow, the narrower pipe holds less fluid. In the same way, faster flow of inventory means less inventory in the pipeline and so reduced inventory costs.</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>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)</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>An idealization of how the inventory level at a location changes over time.</td>
<td>10</td>
</tr>
<tr>
<td>5.1</td>
<td>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)</td>
<td>31</td>
</tr>
<tr>
<td>5.2</td>
<td>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)</td>
<td>34</td>
</tr>
<tr>
<td>5.4</td>
<td>Profiles of three styles of carton flow rack. Each successive style takes more space but makes the cartons more readily accessible.</td>
<td>36</td>
</tr>
<tr>
<td>5.5</td>
<td>One of a “pod” of carousels. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17)</td>
<td>37</td>
</tr>
<tr>
<td>6.1</td>
<td>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.</td>
<td>45</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Pallets that require extra labor to retrieve.</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>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).</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Layout of a distributor of spare parts.</td>
<td></td>
</tr>
<tr>
<td>6.7</td>
<td>Pallet location A is more convenient than locations B or C.</td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>Cartons are picked from the bottom, most convenient, level; when the bottom position has been emptied, a restocker fills it by dropping a pallet from above. Newly arriving pallets are inserted into the (high) overstock. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).</td>
<td></td>
</tr>
<tr>
<td>7.2</td>
<td>Cartons are picked from pallet flow rack onto a conveyor. The flow rack is replenished from behind. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).</td>
<td></td>
</tr>
<tr>
<td>7.3</td>
<td>Net benefit of storing various full-pallet quantities of a sku in the fast-pick area.</td>
<td></td>
</tr>
<tr>
<td>8.1</td>
<td>For the simplest case, assume that every sku is represented in the fast-pick area.</td>
<td></td>
</tr>
<tr>
<td>8.2</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></td>
</tr>
<tr>
<td>8.3</td>
<td>Especially large or slow-moving skus are better picked from the reserve area. This allows the space they would otherwise occupy to be devoted to more popular skus.</td>
<td></td>
</tr>
<tr>
<td>8.4</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></td>
</tr>
<tr>
<td>8.5</td>
<td>The majorizing function $\bar{c}_i(v)$ is linear in the interval $[0, 2c_i f_i/(p_i s)]$.</td>
<td></td>
</tr>
</tbody>
</table>
8.6 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 \( j \) skus in the forward area for three different ways of allocating space among the skus. The greatest benefit is realized by ranking skus 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).

8.7 This sku, which is a plastic liner for a truck body, nests and so can be stored in not much more space than that required by a single piece.

8.8 Example of slotting that accounts for the geometry of the skus and the storage mode. This pattern of storage minimizes total pick and restock costs for this set of skus 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.)

8.9 Multiple fast-pick areas, such as flow rack and carousel, each with different economics.

9.1 Storing an item so that it blocks access to another creates useless work and increases the chance of losing product.

9.2 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.

9.3 Once quantities and orientations of skus have been established there remains the challenge of fitting them all in as few shelves as possible.

10.1 A simple flow line in which each item requires processing on the same sequence of work stations.

10.2 Positions of the worker 2 immediately after having completed the \( k \)-th order and walked back to take over the order of worker 1 (who has walked back to the start of the line to begin a new customer order).

10.3 Because the slope is of absolute value less than 1, the dynamics function converges to a globally attracting fixed point. In other words, the assembly line balances itself.

10.4 Positions of the workers after having completed \( k \) products.
10.5 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.

10.6 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.)

10.7 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%.

10.8 Would bucket brigades be a good way of coordinating order-pickers in this warehouse?

11.1 An order picker has only local information, such as a sequence of locations and the view from an aisle (a), from which to determine a pick path that is globally efficient. It is hard for the order picker to know or account for the global layout of the pick area (b).

11.2 This picking area contains only popular sku’s and so every order is likely to require walking down the aisle. Thus an efficient route of travel is known in advance.

11.3 A serpentine pick path can result in unnecessary travel (in this case along aisles 3 and 6).

11.4 A modified path outline that sorts aisles 4–6 but not their locations.

11.5 Example of branch-and-pick.

11.6 Left: Travel path generated by detouring into and back out of aisles. Right: Travel can be reduced by allowing picker to determine entry and exit from each aisle.

11.7 A much shorter travel path is possible if popular items are stored close to path outline, as on the right.

11.8 To visit the locations on the left, imagine a decision point at the end of each aisle. Because backtracking is forbidden, the order-picker must visit aisle $i$ before visiting aisle $i + 1$.

11.9 Enumeration and summary of paths from Aisle 1 to Aisle 2. Each candidate path is represented by an edge of the same length in the graph.

11.10 Enumeration and summary of paths from Aisle 2 to Aisle 3.

11.11 Enumeration and summary of paths from Aisle 3 to Aisle 4.

11.12 Enumeration and summary of paths from Aisle 4 to completion.

11.13 The shortest path on the associated graph gives an efficient pick path in warehouse.
11.14 Only the lengths of the edges along each side of the graph need be updated to reflect new customer orders. .................................................. 143
11.15 A pick list with travel directions suitable for use on a branch-and-pick pick-path outline ................................................................. 144
11.16 An order-picker must start at the leftmost dot, visit the shaded locations by traveling the aisles of the warehouse, and finish at the right-most dot. ................................................................. 147

12.1 A typical LTL freight terminal. Shaded rectangles represent trailers to be unloaded; clear rectangles represent destination trailers to be filled. ................................................................. 152
12.2 There is less floor space per door at outside corners and therefore more likely to be congestion that retards movement of freight. ........ 154
12.3 A typical crossdock is built in the shape of the letter I, so that freight can flow across from incoming trailers to outgoing trailers. 156
12.4 Crossdocks have been built in a variety of shapes. Clockwise from upper left: An L-shaped terminal of Yellow Transport; a U-shaped terminal of Consolidated Freightways; a T-shaped terminal of American Freightways; an H-shaped terminal of Central Freight. ................................................................. 157
12.5 Outside corners are vulnerable to congestion; and inside corners forfeit door positions. ................................................................. 158
12.6 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 X-shaped dock around 220 doors. ................................................................. 159

13.1 How picking is distributed over the skus. The horizontal axis lists the skus by frequency of requests (picks). By showing both actual values and cumulative fractions it is easy to see how concentrated the picking is amongst the most popular skus. ..................... 174
13.2 A bird’s eye view of a warehouse, with each section of shelf colored in proportion to the frequency of requests for the skus stored therein. ................................................................. 175
13.3 About two-thirds of the orders are for a single line but these account for only about one-third of the picks. ............................... 177
13.4 Why are there so many skus with exactly a half-year supply? .... 182

15.1 This Amazon.com distribution center is typical of the large, high-volume DCs in North America. ................................................................. 192
15.2 The relatively low cost of labor, high cost of capital, and artificially small market mean that this warehouse in India may be economically efficient. (Photo courtesy of Rohan Reddy) .......... 193
15.3 Multi-storey warehouses are common in Singapore, Hong Kong, and Japan where land is expensive. ................................................................. 194
15.4 Multi-storey warehouse with no automation but accessibility provided by a spiral truck ramp. 195
List of Tables

8.1 Pick density and its surrogate under three different schemes for allocating space ........................................ 87
9.1 Local space efficiencies of each orientation of a $1 \times 2 \times 3$ carton stored in a shelf opening of height 4 and depth 10. ($H =$ Height, $D =$ Depth, $W =$ Width) .......................................................... 108
13.1 Top ten items of a chain of retail drug stores, as measured in number of cases moved .................................. 167
13.2 Top ten items of a chain of retail drug stores, as measured by the number of customer requests (picks) ................. 167
13.3 Top ten items of a chain of retail drug stores, as measured by the number of pieces sold .................................. 168
13.4 Top ten office products measured by customer requests ............ 168
13.5 Top ten office products by weight, a measure of shipping costs ..... 169
13.6 Top ten items moving from a grocery warehouse, as measured by number of cases ........................................... 185
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 fre-
quently 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 ware-
house in which product flows with such velocity that there is not 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 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, data sets, links to other resource. We are constantly
adding to this collection.

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 10), Kevin Gue of the
Auburn University (Chapter 12), Ury Passy of The Technion (Chapter 8), and
Loren “Dr. Dunk” Platzman of On Technology (Chapter 8).
Part I

Issues, equipment, processes
Chapter 1

Warehouse rationale

1.1 Why have a warehouse?

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.
CHAPTER 1. WAREHOUSE RATIONALE

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 items 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, seasonalities strain the capacity of a supply chain. Retail stores in particular face seasonalities 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.
1.2 Types of warehouses

Warehouses may be categorized by type, which is primarily defined by the customers they serve. Here are some of the more important distinctions:

A retail distribution center typically supplies product to retail stores, such as Wal-Mart or Target. The immediate customer of the distribution center is a retail store, which is likely to be a regular or even captive customer, receiving shipments on regularly scheduled days. A typical order might comprise hundreds of items; and because the distribution center might serve hundreds of stores, the flow of product is huge. The suite of products changes with customer tastes and marketing plans.

A service parts distribution center is among the most challenging of facilities to optimize. They typically hold spare parts for expensive capital equipment, such as automobiles, airplanes, computer systems, or medical equipment. Consequently, a typical facility contains a huge investment in inventory: tens of thousands of parts, some very expensive. (A typical automobile contains almost 10,000 parts.) Because of the large number of parts, total activity in the DC may be statistically predictable, but the demand for any particular part is relatively small and therefore hard to predict. This means that the variance in orders is large and large quantities of safety stock must be held. Furthermore, when a part is requested, it is generally very urgent, because an important piece of capital equipment might be unusable, such as a truck or a MRI device.

A catalogue fulfillment or e-commerce distribution center typically receives small orders from individuals by phone, fax, or the internet. Orders as typically small, for only 1–3 items but there may be many such orders.

A 3PL warehouse is one to which a company might outsource its warehousing operations. The 3PL provider might service multiple customers from one facility, thereby gaining economies of scale that the customers would be unable to achieve on their own.

While there are many types of warehouses 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

Here we briefly discuss a few issues that help lay the foundations for warehouse analysis.

2.1 The fluid model of product 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 (Figure 2.1).
CHAPTER 2. MATERIAL FLOW

Figure 2.1: If two pipes have the same rates of flow, the narrower pipe holds less fluid. In the same way, faster flow of inventory means less inventory in the pipeline and so reduced inventory costs.

2.2 Units of handling

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

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 2.2. 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.

2.3 Storage: “Dedicated” versus “random”

Each storage location in a warehouse is assigned a unique address. This includes both fixed storage locations, such as a portion of a shelf and mobile locations such as the forks of a lift truck. Storage locations are expensive because they
Figure 2.2: 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).
represent space, with consequent costs of rent, heating and/or air-conditioning, security, and so on. In addition, storage locations are typically within specialized equipment, such as shelving, which are a capital cost. These costs impel us to use storage space as efficiently as possible.

There are two main strategies used in storing product. The simplest is dedicated storage, in which each location is reserved for an assigned product and only that product may be stored there. This has the advantage that the locations of products does not change and so workers can learn where various products are located.

The problem with dedicated storage is that it does not use space efficiently. This can be seen by tracking the amount of inventory in a given location. If we plot the inventory level, measured for example by volume, we would see a sawtooth shape such as in Figure 2.3 (which represents an idealization of the inventory process.) In one cycle, the storage location is initially filled but empties as product is withdrawn to send to customers. As a result, on average our storage capacity is only about 50% utilized.

A warehouse may have thousands or tens-of-thousands of storage locations. If using dedicated storage, each will have an assigned product. Each product may have a different replenishment cycle and so, upon entering such a warehouse, one expects to see many storage locations that are nearly empty, many that are half-full, and many that are nearly full. On average our storage capacity is only about 50% utilized.

To improve on this, one can use a strategy of random storage. The idea here
is to assign a product to more than one storage location. When one location becomes empty, it is available for reassignment, perhaps to a different product. This space then can be filled again, rather than waiting until the original product is replenished (presumably when the last of the warehouse supply has been exhausted). The more storage locations over which a product is distributed, the less product in each location, and so the sooner one of those locations is emptied and the sooner that space is recycled. Therefore we expect better utilization of space when random storage is used.

Unfortunately, random storage also has some disadvantages. Most immediately, the locations of products will change over time as locations are emptied and restocked with other products. This means that workers cannot learn locations and so must be directed to locations by a warehouse management (software) system. Another disadvantage is that it becomes more time-consuming to put away newly received product because it has to be taken to more locations. There can be other, social complications as well. For example, imagine an order picker who has been directed to the other side of the warehouse to pull a product for a customer. That order picker may be tempted to pick the product from a more convenient location, thus creating discrepancies between book and physical inventory at two locations. For these reasons, random storage requires greater software support and also more disciplined warehouse processes.

Random storage is generally more complicated to manage because it introduces many possible tradeoffs. In particular, one can manage the tradeoff between space and time (labor) on an activity-by-activity basis. For example, one can retrieve product from the least-filled location (to empty and recycle that location as soon as possible) or from the most convenient location (to save labor). Similarly, one can replenish product to the most nearly empty location to fill that empty space or to the most convenient location to save labor time.

How much improvement in space utilization is possible with random storage? Consider a product that is requested at a constant rate, as in our idealization of Figure 2.3. Suppose we hold two weeks supply of this product. If we store it in two locations of equal size and direct all order-picking to only one location then after one week, the first location will be emptied and available for reassignment. After the second week the second location will be emptied and available for reassignment. During the first week, the first location was half full on average and the second location was completely full. During the second week the second location was half-full on average. Thus, on average, we used two of the three location-weeks of storage space assigned to this product for an average utilization of 66%, which is better than the 50% estimated for dedicated storage. More improvement is possible if a product is stored in more locations, though the improvement diminishes and, moreover, the practical problems of management increase.

In practice, a strategy of random storage is typically used in the bulk storage areas, where most of the volume of product is held on pallets. Dedicated storage may be used in the most active picking areas, which are much smaller than the bulk storage. Thus one gets efficient use of most of the space (bulk storage) with labor benefits where it matters most (active picking areas).
CHAPTER 2. MATERIAL FLOW

2.4 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 [18].

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 skus 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. What labor force is necessary 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.

2.4.1 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 skus, 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 subtracting 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.
2.5 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 In what ways has the inventory process depicted in Figure 2.3 been idealized.

Question 2.3 In real life a certain amount of safety stock may be held in a storage location to guard against stockout while awaiting replenishment to that location (which would interrupt order-picking). How would this affect the average utilization of storage space?

Question 2.4 What is the average space utilization one would expect from random storage of a product among n locations of identical size? Assume an idealized inventory process such as depicted in Figure 2.3 and assume that all picks are directed to the location that is most nearly empty.

Question 2.5 Why does the model of Question 2.4 break down (that is, lose meaning) for large values of n?

Question 2.6 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 2.7 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 2.8 A pallet storage facility holds about 10,000 pallets in storage. Arrivals and departures are handled by 7 forklift operators and 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. Assume each driver works 8 hours per day for 250 days of the 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
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 order 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:

<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 \textit{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 \textit{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. ORDER-PICKING

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 theSKU'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 cor-
CHAPTER 3. WAREHOUSE OPERATIONS

responding to work-stations, 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 than 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.
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 is the receiving staff in a typical warehouse much smaller than the staff of order-pickers?

**Question 3.8** Explain the difference between:

- A pick-line and a sku
- An order and a batch
- Popularity and demand
Chapter 4

Warehouse management systems

A warehouse management system (WMS) is basically just software to track and manage warehouse activities. It is generally built around an industrial strength relational database product such as Oracle, Informix, Sybase, DB2, or other. At a minimum, the database tracks all product arriving and all product shipped out. (This is required to track fundamental financial transactions.)

4.1 Receiving and shipping

The most fundamental capability of a WMS is to record receipt of inventory into the warehouse and to register its shipment out. This is fundamental because it drives essential financial transactions: receipt drives the paying of bills to suppliers upstream; and shipping drives the sending of invoices downstream to the consignee. This is the base from which modern, complex WMS's have grown.

4.2 Stock locator system

The next significant increase in functionality is to add a stock locator system. This is essentially the ability to manage an inventory of storage locations in addition to an inventory of product. With this capability a software system can do more than trigger financial transactions but can finally support warehouse operations by directing warehouse activities to/from storage locations.

In addition, a WMS should also track the inventory of storage locations in the warehouse. A good WMS will track every place that product can be, down to and including the forks of individual forklift trucks. The ability to manage the inventory of storage locations makes possible the most fundamental capability
of a WMS, which is the stock locator system, which supports directed putaway and directed picking.

To track warehouse activities in real-time, the database must support transaction processing, which means that the database can maintain its integrity even though being updated simultaneously from multiple sources (purchasing, receiving, picking, shipping, etc.).

4.3 Menu of features

• Basic features of most WMS’s include tools to support
  – Appointment scheduling
  – Receiving
  – Quality assurance
  – Putaway
  – Location tracking
  – Work-order management
  – Picking
  – Packing and consolidation
  – Shipping

• High-end features include support for
  – RF-directed operation
  – Cycle counting
  – Carton manifesting
  – Replenishment
  – Value-added services
  – Vendor/carrier compliance
  – Trailer manifesting
  – Configurability
  – Returns
  – Pick/put to light
  – Yard management
  – Wave management
  – Labor management
  – Task interleaving
  – Flow-through processing

• Advanced features include support for
4.4 The market

At the time of this writing, there are over 300 WMS vendors in the US alone. The largest companies are Manhattan Associates, which grew out of service to apparel distribution; and EXE Technologies, whose strengths lie in pallet and case movement, especially in support of the grocery industry.

4.5 Supply Chain Execution Systems

Warehouse Management Systems are extending their functionality out along the supply chain, both upstream and downstream, to include features that support collaboration. In this they are increasingly competing with enterprise systems such as SAP, which are trying to build specialization in warehouse management. WMS’s have an advantage in that they are already connected to financial systems and already hold information that is important to supply chain visibility and execution systems. The enterprise systems have advantages where they can grow out of a manufacturing enterprise, especially if the manufactured product is an important, high-value item.

As WMS’s grow out along the supply chain it is natural that the WMS providers become global, pulled by the supply chains they hope to manage. The global WMS providers have a big advantage when selling to multinational companies, who can then standardize their WMS operations around the world.

At the time of this writing, there are hundreds if not thousands of WMS vendors in the world but only a few companies with significant global presence:

- EXE Technologies
- Logistics and Internet Services (LIS)
- Manhattan Associates
- MARC Global Systems
- Swisslog Software
4.6 Summary

The core of a Warehouse Management System (WMS) is a database of sku’s and a stock locator system so that one can manage both the inventory of sku’s and the inventory of storage locations.

There are significant opportunities to save labor when the system is expanded to include systems to direct receiving/putaway and/or order-picking. Additional systems are available to control ever finer detail within the warehouse as well as events ever farther up/downstream in the supply chain.

4.6.1 The ugly

The working life of a warehouse management system is generally greater than that of the computer language in which it was written. Consequently, most WMS’s in current use are an accretion of many different computer languages, including COBOL, PL1, Fortran, C, C++, SQL, and others. This can make them hard to maintain or customize.

Most WMS’s do not currently “optimize” anything. Instead, they are extremely configurable and let the user choose from among various heuristics to guide decision-making. But the client — or more likely a consultant — must choose the logic from among the modules provided, or else have the software customized.

Most vendors will customize their WMS for the right price; and in fact, some derive the bulk of their revenue from customization. But it is typical that the vendor owns the intellectual property inherent in the customization.

4.7 More

See [www.mywms.de](http://www.mywms.de) for an interesting open source project to write a warehouse management system in Java. Better yet: Help build one!
Chapter 5

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.

5.1 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.
5.1.1 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 (1.22 meters by 1.02 meters) — but, as with all standards, there are several: most standard pallets are 48 inches along one dimension (1.22 meters) but the other may be 32 inches (0.81 meters) for pallets that go directly from manufacturer onto retail display, such as in a Home Depot; 42 inches (1.07 meters), or 48 inches (1.22 meters), such as those for transport of 55-gallon steel drums.

A 2-way pallet allows forks from a standard forklift or pallet jack to be inserted on either of the 40 inch sides. A 4-way pallet also has narrow slots on the 48 inch sides by which it can be lifted by fork lift. The 4-way pallets are slightly more expensive.

There is no standard height to which a pallet may be loaded.

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 5.1). 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 5.1. 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
Figure 5.1: 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).
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.

**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 lined 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.

Except for automated storage-and-retrieval systems (AS/RS), some type of lift truck is required to access the loads in pallet rack; 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 (3.7–4.6 meters), its lift height is limited to 20–22 feet (6.1–6.7 meters), and it travels at about 70 feet/minute (21.3 meters/minute). The stand-up version requires an aisle width of 10–12 feet (3.1–3.7 meters), its lift height is limited to 20 feet (6.1 meters), and it travels at about 65 feet/minute (19.8 meters/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
5.1. STORAGE EQUIPMENT

truck is required to access the rear positions in double deep rack storage. Each truck requires an aisle width of 7–9 feet (2.1–2.7 meters), their lift height is limited to 30 feet (9.1 meters), and they travel at about 50 feet/minute (15.2 meters/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 (approximately 10 centimeters) 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 (1.5–2.1 meters) is required, its lift height is limited to 40–45 feet (12.2–13.7 meters), and it travels at about 75 feet/minute (22.9 meters/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 (30.5 meters). Roof or floor-mounted tracks are used to guide the crane. The aisle width is about 6–8 inches (0.15–0.20 meters) 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.

5.1.2 Bin-shelving or static rack

Simple shelving is the most basic storage mode and the least expensive (Figure 5.2). The shelves are shallow: 18 or 24 inches (0.46 or 0.61 meters) are typical, for example, but 36 inch (0.91 meter) 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.

5.1.3 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 5.3). The shelves may be 3–10 feet
Figure 5.2: 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).
5.1. STORAGE EQUIPMENT

Figure 5.3: 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).

deep (0.91–3.0 meters). 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.

There are several subtypes of carton flow rack, as shown in Figure 5.4.

- 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.)
CHAPTER 5. STORAGE AND HANDLING EQUIPMENT

Figure 5.4: Profiles of three styles of carton flow rack. Each successive style takes more space but makes the cartons more readily accessible.

- 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.

5.1.4 Carousels

Carousels are motorized, computer-controlled, independently-rotating aisles of shelving (see Figure 5.5). 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.

5.2 Conveyors

Main points:

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

Figure 5.5: One of a “pod” of carousels. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).
• 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).

5.3 Sortation equipment

• Push

• Tilt-tray

5.4 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.

5.5 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!
5.6 Questions

Question 5.1 Which types of pallet storage generally provide the most efficient use of floor space: floor storage or pallet rack? Explain.

Question 5.2 What are the relative advantages and disadvantages of push-back rack compared to pallet rack? How do both compare to gravity flow rack?

Question 5.3 Which type of storage generally makes pallets more accessible: Pallet flow rack or drive-in rack? Explain.

Question 5.4 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 5.5 (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 5.6 Conveyable sku’s should be stored far from shipping; why?

Question 5.7 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 5.8 What are the dimensions of a standard pallet? Go to the web and see how many other “standards” you can find.

Question 5.9 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 5.10 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
Part II

Layout
Chapter 6

Pallets

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. An example of unit-load is a 3rd party transshipment warehouse that receives, stores, and forwards pallets. Because pallets are (mostly) standardized and are (mostly) handled singly, both space and labor requirements scale: It takes about $n$ times the space to store $n$ pallets as for one; and it takes about $n$ times the labor to handle $n$ pallets as for one.

The 3rd party warehouse 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. In this chapter we study unit-load issues in the context of such a warehouse.

6.1 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 skus.)

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.

6.1.1 Storage locations

Where should each sku be stored? It is intuitive that the “most active” skus 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: Travel from receiving dock to location; and travel 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

$$\sum_i d_i n_i.$$  \hfill (6.1)

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 6.1 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 $d_i$.

Fast-moving skus versus fast-moving pallets

From the expression 6.1 for the labor cost of a location we prefer that skus with a lot of movement per storage location be stored in the best locations. In other words, we want to identify those skus 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.
6.2 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 6.1 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.

Figure 6.1: 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.

- Rank all skus from most to least turns.
- Move down the list, assigning the pallets of the next fastest-turning skus 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 that the busiest skus belong in the most convenient locations; only now we must adopt a more detailed view that considers the rate at which individual pallets turn (not just skus). Now we choose that particular pallet that will be leaving soonest to be stored in the best location. So to minimize labor costs:

- Rank all the pallet positions of the warehouse from least cost $c_i$ to greatest cost.
- Rank all pallets from soonest departure to latest departure.
- Move down the list, assigning the next pallet to the next location.
Figure 6.2: 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.

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 6.2, where 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 skus passing through the warehouse. If there will be a small amount of very fast-moving skus, 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
6.3. SPACE

- 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 skus 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.
- 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 [6.3] 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.

6.3 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.
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.

6.3.1 Stack height

Pallets that can be stacked high and allow many pallet positions per square foot of floor space. Conversely, pallets that are unusually heavy or fragile or that have uneven top surfaces cannot be stacked very high and so render unusable all the space above. This waste may be avoided by installing pallet rack, so that pallets may be stored independently of each other.

How much pallet rack should be purchased and what should be stored therein? The value of the rack to the user depends on the sizes and movement patterns of the particular skus that visit the warehouse. We require that each sku present an economic argument for why it should be stored in pallet rack rather than floor storage.

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 only 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 skus in rack?

Here are some possible benefits of moving a sku from floor storage into pallet rack.

- It might reduce labor by making product easier to store and retrieve. For example, consider the pallet stack in Figure 6.4 where the uneven top pallet cannot be easily picked up by forklift. In this warehouse, a second forklift was needed to straighten the stack so that the first forklift could insert its forks under the pallet. This means that at least twice the labor
was required to retrieve each pallet above the ground floor. This savings can be estimated easily and could be realized as increased throughput or reduced labor requirements.

- It might create additional pallet positions. For example, storing the pallets of Figure 6.4 in 4-high pallet rack would create an additional pallet position, which is potentially revenue-generating. The value of this additional pallet position depends on how easily it can be rented and so the extent of savings depends on the market for space.

Some skus might not generate any new pallet positions if put in rack, and, in fact, might even lose pallet positions. For example, consider 3 pallets, each 4 feet (1.2 meters) high, of a sku that stack 3-high to come within a two feet (0.6 meters) of the ceiling. Moving these pallets to rack will not create additional pallet locations. In fact, it might not be possible to install 3-high pallet rack to hold these pallets because of the additional height required to accommodate the cross-beams of the rack and the space above each pallet in rack.

- It will help protect product from damage, for example, by forklifts. This savings is hard to quantify except by comparison to past experience.

- It might help provide a safer work environment by avoiding unstable pallet stacks. This savings is also hard to quantify.

Each sku builds its economic argument by estimating the savings for each of these categories and then sums the total savings. The result is a number of dollars, which represents the value of storing that sku in pallet rack. Now one can compare that the cost of pallet rack and decide, on a sku-by-sku basis, whether the pallet rack is economically justified.

Finally, to compare different rack configurations, repeat the process on each alternative and choose the one that is of greatest value.
6.3.2 Lane depth

Space for aisles cannot be used for storage and so is not directly revenue-generating. Consequently we prefer to reduce aisle space to the minimum necessary to provide adequate accessibility. For this the aisles must be at least wide enough for a forklift to insert or extract a pallet. Aisle space is then a fixed cost to provide immediate access to the pallet positions immediately adjacent to the aisle.

By storing product in lanes, additional pallet positions can share the same aisle space and so amortize that cost. 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 6.3 fits about 41% more pallet positions in the same floor area than does the single-deep layout of Figure 6.2 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.

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 6.5).

Let $w$ and $d$ be the (standard) pallet width and depth respectively; let the lanes be $x$ 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 footprint of the lane is given by

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

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 skus of our warehouse.

One strategy is to force the entire census of skus to jointly choose a single lane depth. We can do this by the following simple model. Let there be $n$ skus, 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 6.1 (Floor storage)** To minimize the average space consumed per pallet, floor storage should be configured with lane depth of approximately

$$\sqrt{\frac{a}{2dn} \sum_{i=1}^{n} \frac{q_i}{z_i}}. \quad (6.2)$$

**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 $[q_i/(z_i x)]$ 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

$$\left\lceil \frac{q_i}{2z_i x} \right\rceil$$

lanes. \quad (6.3)

Note that we are justified in rounding up in Expression 6.3 because an entire lane is rendered unusable by other skus if any portion of it is occupied. Now we replace Expression 6.3 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 skus:

$$\sum_{i=1}^{n} \left( \frac{q_i}{2z_i x} + \frac{1}{2} \right) (g + w) (dx + a/2). \quad (6.4)$$

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

This theorem gives an *ideal* lane depth for the collection of skus. It is an ideal because it ignores the space constraints imposed by the physical layout of the warehouse, so it should be taken as advisory, as the preferred single lane depth of the community of skus.

**Example 6.1** Consider the following population of skus:

<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 (about 4.6 meters) and the pallets are 48 inches deep and 42 inches wide (1.22 meters by 1.07 meters)? Use a gap between aisles of 1 foot (0.3 meters).

Substituting into Expression 6.2 gives

$$0.79 \sqrt{\frac{50}{3} + \frac{40}{4} + \frac{36}{2}} \approx 5.28 \quad (6.5)$$

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

$$\sqrt{\frac{a}{dn} \sum_{i=1}^{n} q_i}. \quad (6.6)$$
Proof Left as an exercise. (In addition to the one-half of the aisle in front, to enable access, be sure to charge each lane of storage for one-half of the aisle in back, to enable restocking. This accounts for the term $a/dn$ rather than $a/2dn$.)

In general 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 forwardmost pallet to be retrieved and, in addition, the forwardmost 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 6.1 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 6.6 is never less than Expression 6.2 by using deeper lanes.

To return to floor storage, note that Theorem 6.1 also allows us to determine the preferred lane depth of each individual sku:

Corollary 6.1 The preferred lane depth of a single sku, with $q$ pallets stackable $z$ high, is

$$
\sqrt{\frac{a}{2d}q}
$$

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 skus. Nevertheless, it is helpful to know the preferred depth of each sku because we can group skus with similar storage preferences together, so that, for example, skus that prefer 3-deep storage might be grouped with skus 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 $n$ skus 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 6.4, 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 skus in increasing size of their ideal lane depths. Set $d_0 = 0$ and $d_{n+1}$ to an integer larger than $x^*_n$. Then the following result says roughly that assigned lane depths should respect the order of ideal lane depths.

**Theorem 6.3** 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, $b_j/a_j \leq d_{k_j}$. As $d_{k_j} = d_{k_i}$ 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.

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 skus, such as food products. In addition, deeper lanes can increase insert/extract times.

### 6.4 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 skus, 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 skus 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 6.1 and 6.2.

- 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.
6.5 More

Computer distributors tend to have pallets that are either high and light (a pallet of printers can be 7 feet, or 2.1 meters, high); or low and heavy (such as a pallet of software).

The management of pallets is a constant problem: They flow downstream in the supply chain and so must be either replaced or recirculated; they become damaged; they eventually have to be replaced. Some companies address these problems this by maintaining a communal pool of high-quality pallets that they recirculate among their clients, each of whom pays rent for the pallets they use. (See Chep for example.)
6.6 Questions

Question 6.1 What is a unit-load warehouse and why is it easier to layout than others?

Question 6.2 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 6.3 Suppose that we want to store all the skus of each customer together in a pallet-in, pallet-out warehouse. The skus 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 6.4 Why is it generally better to centrally locate receiving and shipping doors.

Question 6.5 Plot the distribution of travel distances associated with different arrangements of receiving and shipping doors.

Question 6.6 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.
Figure 6.6: Layout of a distributor of spare parts.

Question 6.7 Figure 6.6 shows the layout of a distributor of spare parts. The left side of the warehouse is pallet rack and cantilever rack. Receiving is on the bottom left and shipping is on the bottom center of the figure. Critique this layout.

Question 6.8 Under what circumstances are deep lanes appropriate? What advantages and disadvantages accrue for deeper lanes?

Question 6.9 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 6.10** Suppose that you have planned a layout for pallet rack with one global, optimal lane depth for your skus. Subsequently you learn that inventory levels will be half of what you had expected; how does this change the optimal lane depth?

**Question 6.11** 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 6.12** Why do you suppose single-deep rack also called “selective rack”?

**Question 6.13** Which method of storage makes empty pallet positions more quickly available, thereby increasing storage capacity: floor storage or rack storage? Explain.

**Question 6.14** Explain why a warehouse with only one or two pallets of each of many skus is likely to prefer rack to floor storage.

**Question 6.15** 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 6.16** In Section 6.1.1 we estimated the “convenience” of a storage location by the forklift travel from receiving to the location plus that from the location to shipping. We did not explicitly account the unloaded travel (“dead-heading”) that the forklift must make after storing a pallet in the location or delivering the pallet to shipping. How might this additional travel affect our analysis? Assume that we load and ship full trailers in the mornings and that we receive and unload full trailers in the afternoons.

**Question 6.17** Consider a toy warehouse with 3 pallet locations A, B, C. The total time to travel from receiving to the location to shipping is 1, 2 and 2 minutes, respectively, as shown in Figure 6.7.

Through this warehouse flow two skus, x and y, which you must manage in a strict FIFO manner. Every three days one pallet of sku x is shipped out in the morning and one pallet arrives in the afternoon and is put away. Every two days one pallet of sku y is shipped out in the morning and one pallet arrives and is put away. You maintain a constant inventory of one pallet of sku x and two pallets of sku y on-hand at all times.

A. If you are using a policy of dedicated storage, what is the assignment of skus to storage locations that minimizes average labor? (Justify your answer.) What is the corresponding average minutes per day spent moving these products?
6.6. QUESTIONS

Figure 6.7: Pallet location A is more convenient than locations B or C.

B. What is the smallest sustainable value of average minutes per day spent moving these skus if you adopt a policy of random storage? Does it make any difference where the skus are stored initially?

C. What is the best way of assigning pallets to storage locations if you are permitted to drop the FIFO requirement? Would this save any labor? If so, how much; if not, why not?

D. How would your previous answers change if the pallets of sku y arrive in batches of two every four days? Assume the pattern of shipping is unchanged.

Question 6.18 (Research question) Consider a unit-load warehouse with regular, patterned input stream of arriving pallets and output stream of departing pallets, as in the previous question. Explore how, under a policy of random storage, the disposition of skus in the warehouse organizes itself.

Question 6.19 (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 skus?

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

- Generalize to three or more lane depths.

Question 6.20 (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 7

Cartons

The handling unit “carton” or “case” is not standardized, but generally it refers to a rectangular box that:

- is between about 5 and 50 pounds (2.27 and 22.68 kilograms);
- is palletizable and conveyable (which depends on the type of conveyor);
- can be handled by one person;
- is generally stored on a pallet.

Cartons (or, equivalently, cases) are stored on pallets and so restocking is typically a unit-load process; but picking is not. Picking is a challenge because of the challenges of handling the cartons and of accumulating an order.

7.1 Some strategies for carton-picking

Because cartons are relatively large and stored on pallets they consequently do not achieve much sku-density and therefore do not achieve much pick-density. This means that there can be typically significant travel incurred in order-picking. Here are some typical strategies to reduce this travel, while also achieving suitably dense storage.

7.1.1 Pick from floor stack

Pick from pallets on floor onto pallet truck or pallet jack. Use this when product is very fast-moving and so there is no point in putting it in pallet rack, as it will be withdrawn soon afterwards.
Figure 7.1: Cartons are picked from the bottom, most convenient, level; when the bottom position has been emptied, a restocker refills it by dropping a pallet from above. Newly arriving pallets are inserted into the (high) overstock. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).

7.1.2 Simple pick from pallet rack

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.

Use this when there are many palletized skus but only a few pallets of each. Note that in this case, the problem of reducing travel is very much like a “Traveling Salesman Problem”, about which more in Chapter 11.

7.1.3 Forward- or fast-pick areas

Two typical configurations of fast-pick areas in warehouse are:

- Pick from pallets at bottom of rack and replenish from above

  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.)

- Pick from carton flow rack to conveyor

  Pick from pallet flow rack onto conveyor, which makes many storage locations convenient, and thence to sortation. This is possible when cases are sufficiently uniform.

Both of these configurations share the same organization: a small number of locations from which it is convenient to pick from are restocked from a bulk
7.1. SOME STRATEGIES FOR CARTON-PICKING

Figure 7.2: Cartons are picked from pallet flow rack onto a conveyor. The flow rack is replenished from behind. (Adapted from “Warehouse Modernization and Layout Planning Guide”, Department of the Navy, Naval Supply Systems Command, NAVSUP Publication 529, March 1985, p 8–17).
or reserve area. In the first case, the forward area is the ground floor pallet positions and the reserve includes all higher locations. In the second case, the forward pick area includes all pallet locations in the pallet flow rack and the reserve is typically 6-high pallet rack (not shown in the figure).

Because it is the more typical, we shall discuss the first case, picking from ground level, although everything we develop applies equally to picking from pallet flow rack.

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, we estimate the number of restocks (pallet drops) as simply the total number of pallets moved through the fast-pick area.

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.

The main insight here is that once one has decided to store a product in the forward pick area, giving it additional storage locations, beyond the minimum required, conveys no benefit: It does not increase the number of picks from the forward area, nor does it reduce the number of restocks (because restocking is a unit-load process the number of restocks is always equal to the number of pallets sold). There is additional savings only when one has put every pallet of the sku in the forward area so that no restocking at all is required. Therefore the only amounts to consider storing are: no pallets, one pallet, or all the pallets. We formalize this as:

**Theorem 7.1 (Law of None, One, or All)**  Any sku that is picked from pallets should either not be in the fast-pick area at all; or it should have one pallet; or it should have all of its on-hand inventory in the fast-pick area.

In real life it might not be desirable to store only one pallet as this might not be sufficient to protect against stock out during replenishment. To guard against this, there may be additional pallets stored as safety stock. In this case, we can reinterpret the result as claiming that any sku in the fast-pick area should either not be in the fast-pick area at all; or it should be stored in the minimum possible quantity (safety stock plus one pallet); or all of it should be in the fast-pick area.

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 let
7.1. SOME STRATEGIES FOR CARTON-PICKING

sku’s compete for each pallet location based on the marginal rate of return each sku offers for that space. In the following analysis we assume that each pick is either for only cartons or else for only full pallets. (The more general case is left as a homework question.)

Let $p_i$ be the number of less-than-pallet picks, $d_i$ the number of pallets moved by such 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

$$
\text{net benefit} = \begin{cases} 
0 & \text{if } x = 0; \\
sp_i - cr_i d_i & \text{if } 0 < x < u_i; \\
s (p_i + D_i) & \text{if } x = u_i
\end{cases} \quad (7.1)
$$

This function is represented in Figure 7.3.

![Figure 7.3: Net benefit of storing various full-pallet quantities of a sku in the fast-pick area](image)

Now we can allocate the space by a sort of auction amongst the skus, each of which may be imagined to hold two bidding paddles with the allowable values of its bids imprinted thereon. Bid #1 is

$$
(sp_i - cr_i d_i) / l_i, \quad (7.2)
$$

which is the value per pallet position sku $i$ can offer for the first few pallet positions. If sku $i$ wins with this bid, it is awarded $l_i$ pallet positions and its paddle #1 is discarded. Thereafter, sku $i$ can bid only with its second paddle, for which the bid is

$$
(sD_i + cr_i d_i) / (u_i - l_i), \quad (7.3)
$$

which is the value per pallet position sku $i$ can offer for additional space (above $l_i$). If a sku wins with its second bid, it receives sufficiently many positions to hold all its pallets and then retires from subsequent bidding.
During the auction, each pallet position will have been auctioned off for the best “price”, which means that it will generate the greatest possible labor savings. However, this might not be true of the last few pallet positions: Eventually a sku \( j \) will win the auction and there will not be enough available pallet positions to hold all the required pallets, so these last few pallet positions may not be allocated optimally. In general this discrepancy will be quite small compared to the total value of the fast-pick area.

**Building a new fast-pick area**

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.

### 7.2 Sortation

Sortation is used mostly when picking cartons because they tend to be uniform than eaches.

- 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.

### 7.3 Miscellaneous tidbits

- When picking cartons from pallet rack It is important to configure the rack to leave ample headroom so that the order-picker does not hit his head on a crossbeam. Suggested height is around 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.

- 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.

- Picking cases to pallet can be challenging in the same way as 3-dimensional Tetris. Cartons of many shapes, sizes, weights, and fragilities must be packed tightly together and quickly.

There are several goals in building pallets but the most important is achieving a high, tight, and stable load. As in Tetris, it is a good strategy to try to build full layers of even height, which is easier if you store similar kinds of product together. It is also useful to think in terms of building pairs of pallets, with one to be loaded on top of the other in the trailer. You should not split a sku among pallets, but if it is unavoidable, try to keep the sku on the same pair of pallets.

Large, heavy items should be on the bottom of each pallet 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.

### 7.4 Summary

- When skus are stored as full pallets in a 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.
7.5 Questions

Question 7.1 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.)

<table>
<thead>
<tr>
<th>Sku</th>
<th>Case picks</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>600</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>200</td>
<td>2</td>
</tr>
</tbody>
</table>

Question 7.2 Answer the previous question with the following, additional information that tells when to trigger replenishment of the fast-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>600</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>

How many pallet positions could these sku’s usefully occupy in the forward pick area?

Question 7.3 Compute and rank all positive bids for forward pallet positions offered by the following skus. Use this ranking to auction off 50 pallet positions on the first level of single-deep pallet rack. Which skus appear only in the forward area? Which in both? Which only in reserve? How many picks come from the first level? How many from higher levels? How many restocks are incurred? How many pallets are shipped from each area? How much are pallet positions #48, 49, and 50 worth to you?

Assume that on average it requires 1 minute to pick a carton from the first level, 2 minutes to pick a carton from higher levels, and 3 minutes to drop a pallet to the first level. In the table below, all demand is given in pallets, as are the minimum and maximum amounts to be stored.
7.5. QUESTIONS

<table>
<thead>
<tr>
<th>Sku</th>
<th>Carton Picks</th>
<th>Resultant Demand</th>
<th>Full pallet Demand</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>18.4</td>
<td>4.4</td>
<td>1.4</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>17.0</td>
<td>4.0</td>
<td>1.2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>C</td>
<td>18.3</td>
<td>4.8</td>
<td>10.3</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>D</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>4.2</td>
<td>0.6</td>
<td>3.2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>F</td>
<td>2.0</td>
<td>0.0</td>
<td>1.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>12.7</td>
<td>1.3</td>
<td>10.7</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>H</td>
<td>17.3</td>
<td>3.6</td>
<td>4.2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>I</td>
<td>4.2</td>
<td>1.0</td>
<td>5.2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>J</td>
<td>1.8</td>
<td>0.0</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Question 7.4** Suppose you are allocating pallet positions in a fast-pick area where cartons are picked from pallet flow rack and each lane holds two pallets. Suppose that, if sku G is allocated space in the pallet flow rack, at least 3 pallets worth must be stored to guarantee timely replenishment. How do your answers to Question 7.3 change?

**Question 7.5** Suppose that, to better protect against stockouts in the fast-pick area, the minimum allowable storage amount of a sku has recently increased. Does this increase or decrease its claim to space in the fast-pick area? Explain.

**Question 7.6** Suppose that during the last week a sku was requested in full pallet quantities 23 times, which accounted for a total demand of 40 pallets. How many times was this sku picked? Explain.

**Question 7.7** Generalize Expression 7.1 for pick-savings to include the case in which some picks might be for both cartons and pallets. (For example, a pick might be for 4 cartons and 1 pallet.)
Chapter 8

Pieces: 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.

8.1 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 (skus) 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 skus 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 skus 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 (skus); and we will simply measure the cubic
feet or meters 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 [3] 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.

8.2 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 SKUs 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 8.1 (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

\[
\frac{f_i}{v_i} \text{ restocks per year.} \tag{8.1}
\]
8.3. HOW MUCH OF EACH SKU TO STORE IN THE FAST-PICK AREA?

Figure 8.1: For the simplest case, assume that every sku is represented in the fast-pick area.)

8.3 How much of each sku to store in the fast-pick area?

Suppose that we have decided exactly which skus will be stored in a fast-pick area of volume $V$, as suggested by Figure 8.1. How much space should be allocated 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 available to other skus. 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 skus, 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.

8.4 Allocating space in the fast-pick area

We want to store just the right amount of every sku so that total restock costs are minimized. 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 8.1** To minimize total restocks over all skus \( j = 1, \ldots, n \), each sku \( i \) should be stored in the amount

\[ v_i^* = \left( \frac{\sqrt{f_i}}{\sum_{j=1}^{n} \sqrt{f_j}} \right) V, \quad (8.2) \]

**Proof** This may be seen as an instance of the problem studied in [19]. First we rewrite Problem 8.2 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

\[ \min \left( \sum_{i=1}^{n} f_i / v_i \right) + \lambda \left( \sum_{i=1}^{n} v_i - \lambda V \right) \]
\[ v_i \geq 0 \]

which is equivalent to

\[ \min \sum_{i=1}^{n} (f_i / v_i + \lambda v_i) \]
\[ v_i \geq 0 \]

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:

\[ v_i^* = \frac{\sqrt{f_i}}{\lambda}. \]

Then, setting \( \sum_i v_i^* = V \) 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. First, note that the solution is not a simplistic rule; instead, the amount of space awarded to each sku depends on
that awarded to all the other skus, as it should. Unfortunately, the warehouse industry resorts all too often to “80/20 rules” or “ABC rules”, which treat large classes of skus 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 skus, the total can be significant; and it can be avoided by using optimization, which accounts for all differences among skus.

It is also important to note that this result gives an “ideal” amount in which each sku should be stored and this might not be precisely 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 skus that are stored in amounts that are far from optimum.

Theorem 8.1 has several useful practical implications:

**Corollary 8.1** The fraction of available storage space that should be devoted to sku \(i\) is

\[
\left(\frac{\sqrt{f_i}}{\sum_j \sqrt{f_j}}\right).
\]

### 8.4.1 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:** Allocates the same amount of space to each sku, so that, if \(V\) cubic feet are available, \(v_i = V/n\) and sku \(i\) is restocked \(nf_i/V\) times a year.

**Equal Time Allocation:** stores an equal time supply of each sku, so that

\[
v_i/f_i = v_j/f_j.
\]

Substituting this in the expression \(\sum_i v_i = V\) yields

\[
v_i = \left(\frac{f_i}{\sum_j f_j}\right)V,
\]

from which it follows that each sku \(i\) is restocked \((\sum_j f_j)/V\) times a year.

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

**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 skus 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 8.2** For a given set of skus that are allocated space in Equal Time amounts, the total number of restocks required is the same as if the skus had been allocated Equal Space amounts.

**Proof** By simple algebra:

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

**Example 8.1** Consider two skus 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></th>
<th>sku A</th>
<th>sku B</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>flow</td>
<td>16</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Equal Space Allocations</td>
<td>1/2</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>Restocks</td>
<td>32</td>
<td>2</td>
<td>34</td>
</tr>
<tr>
<td>Equal Time Allocations</td>
<td>16/17</td>
<td>1/17</td>
<td>1</td>
</tr>
<tr>
<td>Restocks</td>
<td>17</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>Optimum allocations</td>
<td>4/5</td>
<td>1/5</td>
<td>1</td>
</tr>
<tr>
<td>Restocks</td>
<td>20</td>
<td>5</td>
<td>25</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 $EQT/OPT$ 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

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

How large can this be? In other words, how much restocking labor might we save by stocking Optimal Allocations? 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):

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

from which we conclude that:

**Observation 8.1** The more diverse the rates of flow of the skus, the more important it is to allocate space optimally, rather than by simple rules such as Equal Space or Equal Time.

**Corollary 8.2** If the \( \sqrt{f_i} \) are all independently and identically distributed exponential random variables then \( CV = 1 \) and \( \frac{\text{EQT}}{\text{OPT}} \approx 2 \), which means both the Equal Space and the Equal Time Allocations incur restocks at twice the minimum rate possible.

It is standard practice in the warehousing industry to store Equal Time amounts, typically after an ABC analysis. For example, Frazelle, in [23], 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 skus 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 skus, the total can be significant when compared to the optimal storage amounts.

To estimate this error we computed the ratio \( \frac{\text{EQT}}{\text{OPT}} \) for 6,000 fast-moving skus 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 skus of a telecommunications company we computed a ratio of 2.44, which means that storing skus in Equal Time Allocations incurred more than twice as many restocks as necessary.

**Managing storage, restocking**

Equal Space Allocations and Equal Time Allocations each offer a type of uniformity that can simplify warehouse management. For example, under Equal Space Allocations the uniformity of storage can simplify space management, especially when old skus are being phased out and new skus introduced. Because all storage slots are the same size, a newly-arrived sku always fits into a space in the fast-pick area. However, there is a cost to this; namely that the Equal Space Allocations will require highly non-uniform frequencies of restocking, which might make the restocking process more difficult to manage.

On the other hand, it is not immediately obvious how much work will be required to maintain the fast-pick area by restocking. Similarly, under Equal
Time Allocations, each sku is restocked at the same frequency and so it is easier to estimate the restocking labor required to maintain the fast-pick area. For example, if each sku in the fast-pick area is stocked with 3-weeks supply then about one-third of the skus must be restocked each week. In some situations this can enable savings because it can allow restocks to be batched. However, Equal Time Allocations vary greatly in the amount of space allocated to the skus and so 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 as 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

\textbf{Theorem 8.3} Optimal Allocations vary less than those of Equal Time Allocations and so are easier to maintain in the warehouse. Similarly, under Optimal Allocations the frequencies of restocking skus vary less than under Equal Space Allocations.

There are several ways to formalize the notion of “vary less”. One way is to compare the span of the Equal Space Allocations with that of the Equal Time Allocations, where by “span” we mean the difference between the largest and smallest values. This is left as an exercise for the reader and we shall prove something stronger.

\textbf{Theorem 8.4} For a given set of skus, the sample variance of the Optimal Allocations is never greater than for the Equal Time Allocations.

\textbf{Proof} This idea behind this result is that the Equal Time Allocation for each sku \(i\) depends directly on the value \(f_i\)’s, while the Optimal Allocation depends on the value \(\sqrt{f_i}\); and if \(f_i\) is very large then \(\sqrt{f_i}\) is smaller and if \(f_i\) is very small then \(\sqrt{f_i}\) is larger. Thus the Optimal Allocations tend to avoid the extreme values that might be assumed by the Equal Time Allocations.

We prove the result by showing that, for an arbitrary set of skus, the vector of Optimal Allocations is \textit{stochastically dominated} by the vector of Equal Time Allocations [20]. This formalizes the notion that the Optimal Allocations are “less disordered” than the Equal Time Allocations and by known consequences of stochastic dominance the result follows.

To prove stochastic dominance, begin by defining

\[
q(k) = \left\{ \frac{\sqrt{f_i}}{\sum_{j=1}^{k} \sqrt{f_j}}, \frac{\sqrt{f_2}}{\sum_{j=1}^{k} \sqrt{f_j}}, \ldots, \frac{\sqrt{f_k}}{\sum_{j=1}^{k} \sqrt{f_j}}, 0, \ldots, 0 \right\},
\]
where the last \( n - k \) entries are 0. We claim that \( q(n) \) is stochastically dominated by \( q(k) \) (written \( q(n) \prec q(k) \)) for all \( k = 1, 2, \ldots, n \). Equivalently, we claim that

\[
\sum_{i=1}^{m} \left( \frac{\sqrt{f_i}}{\sum_{j=1}^{n} \sqrt{f_j}} \right) \leq \sum_{i=1}^{m} \left( \frac{\sqrt{T_i}}{\sum_{j=1}^{k} \sqrt{T_j}} \right) \tag{8.3}
\]

**Case 1:** \( m \geq k \) Inequality (8.3) holds because then the right-hand side sums to 1, which is an upper bound on the value of the left-hand side.

**Case 2:** \( m < k \) The inequality holds because the \( i \)-th summand on the left-hand side is never greater than the \( i \)-th summand on the right-hand side.

Now we imagine \( q(k) \) to be a probability distribution over the \( \sqrt{T_i} \), whence by the stochastic dominance of \( q(k) \) it follows that

\[
E_{q(n)} \left[ \sqrt{T_i} \right] \leq E_{q(k)} \left[ \sqrt{T_i} \right]. \tag{8.4}
\]

Writing this out and rearranging terms gives

\[
\sum_{j=1}^{n} \sqrt{T_j} \left( \frac{\sqrt{T_j}}{\sum_{i=1}^{n} \sqrt{f_i}} \right) = \sum_{j=1}^{k} \sqrt{T_j} \left( \frac{\sqrt{T_j}}{\sum_{i=1}^{k} \sqrt{f_i}} \right) \text{ for all } k = 1, 2, \ldots, n
\]

\[
\frac{\sum_{j=1}^{n} f_j}{\sum_{i=1}^{n} \sqrt{T_i}} = \frac{\sum_{j=1}^{k} f_j}{\sum_{i=1}^{k} \sqrt{T_i}} \text{ for all } k = 1, 2, \ldots, n
\]

\[
\frac{\sum_{j=1}^{k} \sqrt{T_j}}{\sum_{i=1}^{n} \sqrt{T_i}} = \frac{\sum_{j=1}^{k} \sqrt{T_j}}{\sum_{i=1}^{k} \sqrt{T_i}} \text{ for all } k = 1, 2, \ldots, n
\]

\[
\sum_{j=1}^{k} \left( \frac{\sqrt{T_j}}{\sum_{i=1}^{n} \sqrt{T_i}} \right) = \sum_{j=1}^{k} \left( \frac{f_j}{\sum_{i=1}^{n} f_i} \right) \text{ for all } k = 1, 2, \ldots, n
\]

The last line states that the Optimal Allocations are stochastically dominated by the Equal Time Allocations as claimed and so the result follows.

A similar argument proves that Equal Space Allocations induce less variable frequencies of restocking. The argument is similar: Under Equal Space Allocations, sku \( i \) is restocked a number of times that is proportional to \( f_i \); while under Optimal Allocations, it is restocked a number of times that is proportional to \( \sqrt{T_i} \) times.

Figure 8.2 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 skus of a retail chain. We see in each case that Optimal Allocations have significantly less variability.
8.4.2 Managerial implications

Equal Space Allocations result in restocking each sku \(i\) at a frequency proportional to its flow \(f_i\). Equal Time Allocations result in restocking each sku at the same frequency.

**Theorem 8.5 (“Law of Uniform Restocking”)** Under Optimal Allocations each unit of storage space will be restocked at the same frequency.

**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 8.2 for \(v_i^*\) gives

\[
\left( \sum_j \sqrt{f_j} \right)^2,
\]

which is independent of \(i\).

This means that restocks should be distributed *uniformly* throughout the fast-pick area. Roughly speaking, each aisle should be restocked at the same frequency and each section of shelving at the same frequency. (Note the difference from Equal Time Allocations, in which each sku is restocked at the same frequency.)

The Law of Uniform Restocking 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 skus in the frequently-restocked bays and less space to the skus in the infrequently-restocked bays.

8.4.3 Minimum and maximum allocations

Sometimes the fluid model will suggest storing impractically tiny amounts of large, slow-moving skus. This can cause problems when there are certain minimum amounts of the sku that must be stored. For example, we cannot allocate...
8.5. Which Skus Go Into the Fast-Pick Area?

Any warehouse has some skus 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 an 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 that picking from the fast-pick area. Therefore the economics become those of Figure 8.3.

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.

Let the cost-per-pick from the active pick area be \( c_1 \) and from reserve (or some alternative storage area) be \( c_2 \). Let our decision variables be

\[
x_i = \begin{cases} 
1 & \text{if sku } i \text{ is stored in and picked from the fast-pick area;} \\
0 & \text{otherwise.}
\end{cases}
\]

The total cost to manage sku \( i \) is then the cost of picking it plus, if it is stored in the fast-pick area, the cost of restocking it to the fast-pick area. Then we can formalize the problem minimizing total labor cost by choice of which skus to store in and pick from the active pick area. Let \( p_i \) be the number of picks
forecast for SKU $i$ during the planning horizon:

$$\min \sum_{i}^{n} (c_1 p_i + c_r f_i / v_i) x_i + c_2 p_i (1 - x_i)$$

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

$$v_i \geq 0$$

$$x_i \in \{0, 1\}$$

Frequently it will be convenient to rewrite this as an equivalent problem of maximizing the net benefit of the fast-pick area; that is, how much total labor does it save compared to the default of storing all SKUs in the reserve area and picking from there. Let $s$ be the savings realized when a pick is from the forward area rather than reserve.

$$\arg \min \sum_{i}^{n} (c_1 p_i + c_r f_i / v_i) x_i + c_2 p_i (1 - x_i)$$

$$= \arg \min \sum_{i}^{n} ((c_1 - c_2) p_i + c_r f_i / v_i) x_i + c_2 p_i$$

$$= \arg \max \sum_{i}^{n} ((c_2 - c_1) p_i - c_r f_i / v_i) x_i$$
8.5. WHICH SKUS GO INTO THE FAST-PICK AREA?

Net benefit

\[
\text{Volume} \quad \text{Net benefit}
\]

Figure 8.4: 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.

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

The net benefit of storing sku \( i \) forward in amount \( v \) is given by \( c_i \) (Figure 8.4): \[
c_i(v) = \begin{cases} 
0 & \text{if } v = 0 \\
sp_i - cr f_i/v & \text{if } v > 0
\end{cases}
\]

\[
(8.5)
\]

In choosing skus to put in the fast-pick area, the expression \( p_i/\sqrt{f_i} \) is so important that we give it a name: the labor efficiency of sku \( i \), because it represents, roughly speaking, the marginal net benefit, measured in total labor (picking plus restocking) of flowing the sku through the fast-pick area.

We shall show that

**Theorem 8.6** The skus that have strongest claim to the fast-pick area are precisely those offering the greatest labor efficiency.

The problem of deciding exactly which skus belong in the fast-pick area is now solvable. Instead of searching over all \( O(2^n) \) subsets of the \( n \) skus, we need consider only the \( O(n) \) ways of partitioning our ranked list of \( n \) skus 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 \( 10^5 \).
CHAPTER 8. PIECES: DESIGN OF A FAST PICK AREA

Figure 8.5: The majorizing function $\bar{c}_i(v)$ is linear in the interval $[0, 2c_r f_i/(p_i s)]$.

Here, then, is the procedure, first presented in [13], to decide what goes into the fast-pick area and in what amounts.

- Sort all skus from most labor efficient to least.

- Successively evaluate the total net cost of putting no skus in the fast-pick area; putting only the first sku in the fast-pick area; only the first two skus; only the first three; and so on. Choose the strategy that minimizes net cost.

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 8.7** Choosing skus based on labor efficiency 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 skus 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 8.6 is replaced by the smallest function $\bar{c}_i$ that majorizes it, as in Figure 8.5.

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

(8.7)

The derived problem 8.7 has two important properties:

---

1Technical Note: This process can be sped up significantly by specialized search methods such as Fibonacci search because the cost is unimodal [13].
• At optimality there cannot be two skus assigned values strictly within their intervals of linearity \((0, 2c_r f_i / (p_i s))\). (If there existed two such skus, 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 8.7 has the property that all skus chosen (that is, all with \(v^*_i > 0\) have an initial rate of return that is strictly greater than those not chosen \((v^*_i = 0)\). (Otherwise, one could improve the objective function by by giving more space to the sku with greater initial rate-of-return.)

Now, by these properties there exists an optimal solution to problem 8.7 of the form

\[
\tilde{z}^* = \left( \sum_{i=1}^{j} c_i(v^*_i) \right) + c_{j+1}(v^*_{j+1})
\]  

(8.8)

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

\[
\left( \sum_{i=1}^{j} c_i(v^*_i) \right) \leq z^* \leq \tilde{z}^* = \left( \sum_{i=1}^{j} c_i(v^*_i) \right) + c_{j+1}(v^*_{j+1})
\]  

(8.9)

and so

\[
0 \leq z^* - \left( \sum_{i=1}^{j} c_i(v^*_i) \right) \leq \bar{c}_{j+1}(v^*_{j+1}),
\]  

(8.10)

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 8.8 (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 \(s p_i - c_r f_i / v_i = 0\) to see what value of \(v_i\) results in a net-benefit of 0.
8.5.1 Priority of claim for storage in the fast-pick area

In interviews we have learned that many people recommend choosing skus for storage in the fast-pick area based on measures other than labor efficiency. 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 skus with many picks for the space they occupy. (This idea is sometimes expressed in equivalent form as the priority of skus with small value of cube per order index, or space occupied per pick [15, 16].)

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 skus 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 skus 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 skus chosen). The pick-density of a chosen sku would be $p_i/v_i = np_i/V$. Now here is the key observation: We can rank the skus 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 skus based only on values of $p_i$.

Reasoning similarly, we get other simplified measures of priority for storage in the fast-pick area that give the same rankings as pick density. These are summarized in Table 8.1: If using Equal Time Allocations one should select the skus with greatest value of $p_i/f_i$; and if using Optimal Allocations one should give priority to those skus with greatest value of $p_i/\sqrt{f_i}$.

We tested different versions of the generic solution procedure on the warehouse of a major telecommunications company for whom we designed a fast-pick
8.5. WHICH SKUS 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>$p_i/\left(\left(\sum_j f_j\right) V\right)$</td>
<td>$p_i/f_i$</td>
</tr>
<tr>
<td>Optimal</td>
<td>$p_i/\left(\left(\sqrt{\sum_j s_j}\right) V\right)$</td>
<td>$p_i/\sqrt{f_i}$</td>
</tr>
</tbody>
</table>

Table 8.1: Pick density and its surrogate under three different schemes for allocating space

area. In each scenario we ranked the skus by pick-density and then successively computed the net benefit from storing the top $k$ skus in the fast-pick area. Figure 8.6 shows the results: Sloting optimal amounts delivered greatest net benefit from the fast-pick area. Equal Time Allocations was next best; and Equal Space Allocations performed least well.

Furthermore, the graph of Figure 8.6 is representative of our experience: the net benefit of Equal Space Allocations peaks soonest (that is, with fewest number of selected skus); followed by the Optimal Allocations and then by Equal Time Allocations. Of course the value of the peak net benefit due to the Optimal Allocations is greatest.

Relation to the cube-per-order index

Let $\text{cpo}$ represent the average cube-per-order, measured in physical volume such as cubic meters or cubic feet. We may rewrite the expression for labor efficiency as:

$$\frac{p_i}{\sqrt{f_i}} = \frac{p_i}{\sqrt{p_i \cdot \text{cpo}}} = \sqrt{\frac{p_i}{\text{cpo}}}.$$

We may equally well express labor efficiency as

$$\frac{p_i}{\text{cpo}},$$

because it gives the same ranking of skus. In this form it is clear that the most suitable skus for the forward pick are those that are picked frequently but in small amounts.

The original work on cube-per-order argued that one should store the skus with smallest value of $\text{cpo}$ in the most convenient locations. However, this is based on an incomplete analysis as it ignores the costs of restocking. In our terms, the cube-per-order index is a measure of pick efficiency. The more general concept of labor efficiency reduces to pick efficiency when there is little difference in pick savings among skus (either all $p_i$ are nearly the same or else $s$ is small).
Figure 8.6: 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 $j$ skus in the forward area for three different ways of allocating space among the skus. The greatest benefit is realized by ranking skus 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).
8.6 Additional issues

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

8.6.1 Storage by family

Sometimes it is advantageous, or even required, to store related skus together. Reasons include the following.

- To get good space utilization, product of similar sizes or shapes might be stored together, such as rope, pipes, buttons, or refrigerators.
- To simplify put-away, product might be stored by vendor so that all product arriving on a truck goes to the same region of the warehouse.
- To simplify put-away at the downstream customer, product might be stored to reflect the layout of the customer’s facility. For example, a DC supporting a chain of retail drugstores might store all hair-care products together in the warehouse. Then the hair-care products will be picked and packed together and so will be unpacked together at the retail store, which means they can be put away with less walking.
- To reduce the need for specialized equipment, such as freezers, grocery distributors store product by temperature zone. The standard zones are frozen, refrigerated, and ambient.
- Chemicals typically must be stored by requirement for special handling or storage. Categories include hazardous, flammable, or aerosol.
- Product may be grouped by security requirements. For example, small, high-value items might be stored behind a fence with controlled access.

Call a group of related skus that must be stored together a **product family**. Now we must decide which families to store in the fast-pick area: Either all the skus of a family will be stored in the forward area or else none of them. 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 8.1, the total space for family $j$ should be $v_j^* = \sum_i v_{ij}^* = \left( \sum_i \sqrt{f_{ij}} / \sum_i \sqrt{f_{ij}} \right)$. But then, in a similar derivation, we should give priority to those families with greatest labor efficiency, now generalized to be

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

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 labor efficient 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.

8.6.2 Accounting for safety stock

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 want to be careful to avoid stockouts, which can disrupt order-picking, and so may prefer to carry some safety stock for each sku. If \( ss_i \) is the volume of safety stock carried for sku \( i \) if it goes in the fast-pick area then we can estimate the number of restocks of sku \( i \) as \( f_i/(v_i - ss_i) \). Our previous results (ideal amounts of storage, pick density, labor efficiency) are then a little more complicated but essentially unchanged. For example, the optimal amount of sku \( i \) to store, formerly given by Expression 8.2, becomes:

\[
v_i^* = ss_i + \left( \frac{\sqrt{f_i}}{\sum_j \sqrt{f_j}} \right) \left( V - \sum ss_j \right),
\]

In other words, each sku \( i \) in the fast-pick area is guaranteed space \( ss_i \); then any remaining space \( V - \sum ss_j \) is partitioned among the skus in the same proportions as we have already seen, that is, by the square roots of their flows.

Storing product in the optimal amounts reduces the number of restocks; and this tends to reduce lead time to restock, which reduces the required levels of safety stock — which further reduces the number of restocks or allows more skus into the fast-pick area or both.

8.6.3 Limits on capacity

Consider the problem of choosing skus to go into an fast-pick area: As previously described, we sort the skus by labor efficiency and then successively evaluate the total operating costs as more skus are added to the active pick area. However, as we add more skus 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 labor efficient skus are chosen for the fast-pick area, then we know that we can restrict our search to the \((k - 1)\) most labor efficient skus.
8.6.4 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 SKUs, 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 SKUs should not be restocked to the fast-pick area? The following theorem tells us.

**Theorem 8.9** If SKU $i$ goes into the fast pick area at all, put all of it if the maximum on-hand volume of SKU $i$ is no greater than

$$2 \left( \frac{c_r f_i}{sp_i} \right).$$

The term above will be familiar from Theorem 8.8, which gives us the following interpretation: There should be no separate reserve storage for any SKU 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 8.9 in slightly different terms.

**Corollary 8.3 (High-turnover SKUs 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 8.9 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.

8.6.5 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 8.5, giving the net benefit of storing $v$ cubic feet of SKU $i$ forward, is revised to

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

(8.12)
and with this enrichment Equation 8.9 becomes

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

and Equation 8.10 becomes

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

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.

### 8.7 Limitations of the fluid model

In some special situations the fluid model may not be as accurate as desired.

**Subadditivity of space**

Some skus 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. Another example may be seen in Figure 8.7. This sort of complication is more likely to occur with *pieces* (individual units of use) rather than with the containers in which the items are shipped or stored.

**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 (approximately 3 meters). 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 skus, with the result
Figure 8.7: This sku, which is a plastic liner for a truck body, nests and so can be stored in not much more space than that required by a single piece.

that required storage in one warehouse was reduced from 325 to 285 bays of flow rack, with no increase in total restocks [8].

Figure 8.8 shows typical results of the more powerful model, applied in this example to slot skus 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.

8.8 Size of the fast-pick area

8.8.1 How large should the fast-pick area be?

As the fast-pick area becomes larger, we can fit more skus 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 8.8: Example of slotting that accounts for the geometry of the skus and the storage mode. This pattern of storage minimizes total pick and restock costs for this set of skus 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.)
Theorem 8.10 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 labor efficient skus, 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 skus from most to least labor efficient and repeatedly compute allocations for the \( k \) most labor efficient skus 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 skus and order history.

8.8.2 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 skus 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 skus \( 1 \ldots n \) in the fast-pick area:

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

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

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

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.
CHAPTER 8. PIECES: DESIGN OF A FAST PICK AREA

Figure 8.9: Multiple fast-pick areas, such as flow rack and carousel, each with different economics

**Theorem 8.11 ("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 $\alpha$.

8.9 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 8.9. Now the question becomes which skus go to which fast-pick areas and which to reserve. Each of the $n$ skus 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 skus 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 8.12** Skus of greatest labor efficiency should be stored in the fast-pick areas with greatest savings per pick.

In other words, some optimal stocking strategy has the following structure: If all skus are ranked by labor efficiency then the top $k_1$ skus 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 \( n \) skus among \( m \) storage modes may be found by the following procedure.

1. Sort skus from most to least labor efficient

2. Search for the partition into \( m \) sets of contiguous skus 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 8.13** The total net benefit over all storage modes is sequentially unimodal. That is, if we fix all \( k_i \) but one then the total net benefit is unimodal in that unfixed \( k_i \).

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

### 8.10 On the lighter side

More than 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. Each manager had 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!

### 8.11 Summary

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

- The configuration of a warehouse can be optimized based on physical size of the skus 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 labor efficiency of a sku \( i \) that is stored in less-than-pallet quantities is \( p_i/\sqrt{f_i} \), which measures the work required to pull a given amount of physical volume through your warehouse. The most labor efficient skus
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 labor efficient skus into the fast-pick area. You must search to determine exactly how many to put in.

• For multiple fast-pick areas, the most labor efficient skus should be stored in the fast-pick areas with greatest pick-savings.

• For those skus stored in less than pallet quantities in the fast-pick area, the optimal amounts, measured in cubic feet, are
  \[ v_i^* = \left( \sqrt{T_i} / \sum_j f_j \right) V. \]

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 skus 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.)
8.12 Questions

Question 8.1 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 8.2 Why is shallow storage generally preferable for smaller, slower-moving skus? For example, why might you prefer to put such skus in static rack (bin-shelving), which might be only 2.5 feet deep (approximately 0.76 meters), rather than in flow rack that is 10 feet deep (approximately 3.0 meters)?

Question 8.3 Show that the Equal Time Allocation assigns volume

$$v_i = \left( \frac{f_i}{\sum_j f_j} \right) V$$

to sku $i$.

Question 8.4 Let $v^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$ be the vector of Optimal Allocations for skus $1, 2, \ldots, n$ and let $v_{EQT} = \{v_1, v_2, \ldots, v_n\}$ be the vector of Equal Time Allocations for the same set of skus. Prove that the span of $v^*$ is no greater than the span of $v_{EQT}$. In other words, prove that $\max_i v_i^* - \min_i v_i^* \leq \max_i v_i - \min_i v_i$.

Question 8.5 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 8.6 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 8.7 A warehouse distributes two types of pen that are identical except for color. Customers prefer the black pen: It was picked 1,500 times last year and averaged 4 eaches per pick. In comparison, the green pen was picked 800 times per year and each pick was for a single each. Which pen has stronger claim to storage in a fast-pick area and why?

Question 8.8 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 8.9 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 8.10 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 8.11 Suppose that a marketing campaign has just begun for one of the more popular skus. 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 8.12 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 skus would leave the fast-pick area and the remaining skus would get more space.
• The skus and amounts stored in the fast-pick area would remain unchanged.
• New skus would join those already present and each of the skus in the fast-pick area would get less space.
• A completely new mixture of skus may be selected.

Question 8.13 Suppose you enlarge your fast-pick area slightly. Assuming the average cost-per-pick has not changed significantly, which of the following will you find after recomputing the optimal population of skus for the fast-pick area?

1. Some of the skus currently in the forward area may be moved out and replaced by an assortment of other skus.
2. All the skus currently in the forward area will remain but they will be joined by additional skus.
3. No new skus will be moved into the forward area; instead, the most labor efficient will remain and each will get more space.
4. It is impossible to tell.

Question 8.14 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 8.15** The following is a miniature of a real problem: In what amounts should I store skus to minimize labor?

Suppose you have 10 cubic feet available in flow rack, which is restocked from a distant reserve area, and you have just added three skus, 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 skus 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 8.16** A. Reconsider Question 8.15 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 8.17** A. Reconsider Question 8.15 supposing you may choose none or any or all of the three skus 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 8.18** Reconsider the preceding question, supposing that skus 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 skus A and B as one product family and sku C as another.)

**Question 8.19** 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
• \( c_r f_i / (sp_i) \)
• \( sp_i^2 / (4c_r f_i) \)

All of the sku that is stored in the warehouse

**Question 8.20** In Section 8.6.1 it was claimed that the correct value for the labor efficiency of family \( j \) is computed from the skus \( i \) 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 8.21** Follow the suggestion of Section 8.6.2 and derive the formulae for optimum allocations when each sku \( i \) has a non-zero reorder point \( ss_i \). Do the same for labor efficiency.

**Question 8.22** 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 skus remains unchanged.)

**Question 8.23** 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 8.24** 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” skus should be stored here? How might the high cost of restocking affect the assignment of skus to modes?

**Question 8.25** Consider a warehouse with multiple fast-pick areas. Which fast-pick area will get the most labor efficient skus?

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 8.26** The savings in restocks realized by storing skus 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 8.27** Make a numerical example that shows that sequencing skus by labor efficiency is not the same as sequencing by picks divided by flow. In other words, find values for which $p_i / \sqrt{f_i} < p_j / \sqrt{f_j}$ but $p_i / f_i > p_j / f_j$.

**Question 8.28** Consider three ways of stocking the fast-pick area: Equal Space Allocations, Equal Time Allocations, and Optimal Allocations. Which method results in

- Rank the methods according to variability among the volumes allocated.
- Rank the methods according to variability among the numbers of restocks required for each sku.

**Question 8.29** Consider a collection of skus 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 a particular sku $i$ we know the number $p_i$ of piece picks, the piece flow $f_i$, the number $p'_i$ of full-carton picks, and the carton flow $f'_i$. (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 pick full-cartons of sku $i$ 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 $i$ should be stored in the fast-pick area and if so, how much space it should be allocated and whether carton picks of sku $i$ 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.)

**Question 8.30** A. Suppose that the fast-pick area is supported by two bulk reserves. One of the reserves is off-site and both picks and restocks from this site cost more than restocks from the other. Each sku has been assigned to a single reserve and you must choose the skus and quantities to be stored in the forward area. How can you adapt the model and solution?

B. How can you determine which skus ought to be assigned to which bulk reserve?
CHAPTER 8. PIECES: DESIGN OF A FAST PICK AREA

Question 8.31 Suppose that the bulk storage area is connected to the fast-pick area by a conveyor. How can you adapt the model to choose the best set of skus and quantities for the fast-pick area while accounting for the maximum throughput of the conveyor?

Question 8.32 The following set of small parts are stored and picked as eaches.

<table>
<thead>
<tr>
<th>Sku</th>
<th>Picks</th>
<th>Qty</th>
<th>Length</th>
<th>Width</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2500</td>
<td>4500</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>1234</td>
<td>0.25</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>C</td>
<td>500</td>
<td>750</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>D</td>
<td>50</td>
<td>90</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>E</td>
<td>250</td>
<td>1000</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>F</td>
<td>100</td>
<td>250</td>
<td>2.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>G</td>
<td>800</td>
<td>850</td>
<td>1.0</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>H</td>
<td>400</td>
<td>800</td>
<td>0.4</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>I</td>
<td>200</td>
<td>400</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Using the fluid model, rank the skus according to the strength of their claim for storage in a forward pick area. Give the value of the labor efficiency of each.

Question 8.33 Suppose you are considering buying flow rack to hold the skus of Question 8.32. The labor economics are estimated to be as follows.

<table>
<thead>
<tr>
<th>Storage mode</th>
<th>Physical volume</th>
<th>Cost/pick</th>
<th>Cost/restock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow rack</td>
<td>1.0</td>
<td>1.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Reserve</td>
<td>Infinity</td>
<td>8.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Use the fluid model to answer the following. What is the total operating cost if the best k skus are stored in flow rack \( k = 0, \ldots, 9 \)? Which value of \( k \) gives the minimum? Which skus should go into which storage modes and in what quantities? What are the number of picks expected from each storage mode? How many times is it expected that the flow rack will be restocked? What is the total expected operating cost (picking costs plus restocking costs)? What would be the net labor savings expected of the flow rack?

Question 8.34 Continuing the analysis of Question 8.33, use the fluid model to evaluate shelving as an alternative to flow rack. Shelving is cheaper than flow rack and so you can afford twice the storage volume; but the labor economics are not as attractive.

<table>
<thead>
<tr>
<th>Storage mode</th>
<th>Physical volume</th>
<th>Cost/pick</th>
<th>Cost/restock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shelving</td>
<td>2.0</td>
<td>2.0</td>
<td>8.0</td>
</tr>
<tr>
<td>Reserve</td>
<td>Infinity</td>
<td>8.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
What is the total operating cost if the best k skus are stored in shelving ($k = 0, \ldots, 9$)? Which value of $k$ gives the minimum? Which skus should go into which storage modes and in what quantities? What are the number of picks expected from each storage mode? How many times is it expected that the shelving will be restocked? What is the total expected operating cost (picking costs plus restocking costs)? What would be the net labor savings expected of the shelving? How much more or less than flow rack is it worth to you?

Question 8.35 Continuing the analysis of Question 8.34, use the fluid model to evaluate the option to purchase both shelving and flow rack. Which skus would go in each and in what quantities? What would be the net benefit of having both types of storage mode? If you had already decided to buy the flow rack, how much additional value would you expect to get from the shelving?
Chapter 9

Pieces: 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 (the “golden zone”, from which it is easiest to pick).

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 requires 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.

9.1 Case orientation and stack level

Space issues are most significant for skus that are stored in less-than-pallet quantities. The storage unit of many such skus is a carton or 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 9.1

Consequently one might as well fill up that space as much as possible with additional cases of the same sku, as in Figure 9.2 We refer to such arrangements as “maximal” because
Figure 9.1: Storing an item so that it blocks access to another creates useless work and increases the chance of losing product.

\[
\begin{array}{cccccc}
\text{height, depth, width} & \{1,2,3\} & \{1,3,2\} & \{2,1,3\} & \{2,3,1\} & \{3,1,2\} & \{3,2,1\} \\
\text{Efficiency} & 1 & 0.9 & 1 & 0.9 & 0.75 & 0.75 \\
\end{array}
\]

Table 9.1: Local space efficiencies of each orientation of a $1 \times 2 \times 3$ carton stored in a shelf opening of height 4 and depth 10. ($H =$ Height, $D =$ Depth, $W =$ Width)

they attempt to fill all the space above and behind that is otherwise rendered unusable by other skus.

Note that alternative arrangements of the same sku may differ in the amount of space wasted above or behind. We can evaluate the local space efficiency of an orientation by the ratio of space actually occupied to the space rendered unusable by other skus. For example, consider a sku stored in cartons of dimensions $1 \times 2 \times 3$, which are to be stored on a shelf with opening of height 4 and depth 10. Then the (local) space efficiencies of the orientations are as given in Table 9.1

9.2 Packing algorithms

Once the skus to be slotted are known, together with the quantity of each, then there remains the problem of packing the cartons onto the shelves. This is typically done by simple packing algorithms. These most commonly-used
9.2. PACKING ALGORITHMS

packing algorithms all have the same general logic: With a sorted list of the skus, repeat the following. Take the next sku from the list and pack it onto the shelf most suitable for it. Note that once a sku has been placed on a shelf by this type of algorithm, that placement is never reconsidered.

As we will see, the simplicity of this family of packing algorithms make them easy to adapt to special circumstances of each warehouse. For example, one can affect the behavior of the heuristic by choosing how to sort the skus of the list. However, each of these packing algorithms is a heuristic and cannot be guaranteed to produce an optimal packing. But they do come with some performance guarantees and they work well in practice.

To maintain focus on the essential ideas of packing, we make a few simplifying assumptions. First, assume that we are slotting $n$ skus and that the quantity and orientation of each has been decided. These skus are to be slotted into shelf openings, all of identical and fixed dimensions and no sku has been allocated more than a single shelf. Our goal is to use the fewest shelf openings possible to hold all the skus.

Later we will discuss how to relax most of these assumptions.

Because orientation has already been decided and the shelf dimensions are fixed, our main concern is with the horizontal space along the shelf and so each shelf becomes essentially 1-dimensional. For us, a shelf is filled if the width of cartons assigned to that shelf fill the width of the shelf opening, as shown in Figure 9.3. Note that the width of a sku allocation is the width of all its cartons plus whatever gap might required between adjacent lanes of cartons.
CHAPTER 9. PIECES: DETAILED SLOTTING

Figure 9.3: Once quantities and orientations of skus have been established there remains the challenge of fitting them all in as few shelves as possible.

When all of the sku allocations have been packed into shelves, then the shelves may be arranged into bays.

9.2.1 Next Fit

The algorithm Next Fit directs its attention to one shelf at a time. It removes the next sku from the sorted list and tries to fit it onto the current shelf. If the sku fits, it is added and the process continues; otherwise the current shelf is deemed full and then closed and never reconsidered. A new, initially empty shelf becomes the current shelf and the process continues.

This heuristic obviously wastes some space because it never reconsiders a shelf after a sku has failed to fit. Fortunately it can not waste “too much” space.

Theorem 9.1 Slotting skus by Next Fit never require more than twice as many shelves as the minimum possible.

Proof Consider the list of shelves after the heuristic has packed the skus: Any two shelves that are adjacent in the list must contain skus that fill or overflow a single shelf.

To see how Next Fit can be fooled, consider a list of $2n$ skus with allocations of width 0.5 and together with another $2n$ of width $\epsilon < 1/(2n)$, to be packed onto shelves of width 1. Packing via Next Fit from the list $\{0.5, \epsilon, 0.5, \epsilon, \ldots, 0.5, \epsilon\}$ requires $2n$ shelves; but the packing that pairs skus with allocations of width 0.5 and places all skus with allocations of width $\epsilon$ on one shelf requires only $n + 1$ shelves.

The Next Fit heuristic produces a packing in which the skus appear in the shelves in exactly the same sequence as in the sorted list, which is sometimes useful. But the main use of this result is as a sort of worst-case: All the
remaining packing algorithms are more careful about wasting space and so pack at least as well.

\subsection{First Fit}

The heuristic \textit{First Fit} achieves more space-efficient packing by keeping every partially-loaded shelf open and available to receive more skus. \textit{Next Fit} removes the next sku from the sorted list and tries to fit it on a shelf, but here is the enhancement: \textit{First Fit} tries the sku on \textit{each} partially-loaded shelf, in order, and puts it on the first shelf on which it fits. If it does not fit on any open shelf, \textit{First Fit} opens a new, initially empty shelf, puts the sku there and continues.

Under \textit{First Fit}, the first shelves to be opened are considered for every subsequent sku; consequently the first shelves tend to be packed very tightly and the last shelves to be opened are packed less well.

\textbf{Theorem 9.2} Slotted skus by \textit{First Fit} never requires more than \([17/10]\) times as many shelves as the minimum possible.

The proof of this is too long to include here.

In practice \textit{First Fit} performs significantly better than \textit{Next Fit}. Furthermore, it may be made to pack with still greater space-efficiency by sorting the skus from widest to narrowest allocation. This follows the general rule about achieving space-efficiency: Pack the larger items first and fill in the remaining space with small things. With this addition the heuristic is known as \textit{First Fit Decreasing} and it is guaranteed to be within about 22\% of the maximum space efficiency.

\textbf{Theorem 9.3} Slotted skus by \textit{First Fit Decreasing} never requires more than four shelves plus \([11/9]\) times as many shelves as the minimum.

The proof of this is too long to include here.

\subsection{Other issues}

It is possible to strengthen \textit{First Fit}. For example the heuristic \textit{Best Fit} slightly extends \textit{First Fit} by storing each sku on the fullest shelf on which it fits. \textit{Best Fit} seems better in principle than \textit{First Fit} but has the same worst-case bound and, moreover, is not observed to perform any better on average. A theoretically interesting variation is \textit{Worst Fit}, which stores the next sku on the emptiest shelf on which it fits.

More sophisticated extensions have been made by tailoring the algorithm to the distribution of sku widths. Generally, such extensions are expected to provide tighter packings but are more complex and special purpose. No one knows an algorithm that is both fast (runs in polynomial time) and guaranteed to use the fewest shelves possible. Of course one can try every possible way of storing the skus but this would require hugely impractical amounts of time.
The bounds given by the theorems are worst-case bounds. In practice the heuristics perform much better. These heuristics may also be expected to pack skus efficiently because the typically large number of skus ensures that heuristics such as First Fit have many opportunities to top off any partially-packed shelves.

These heuristics are simple and fast and produce efficient packings; but sometimes they can behave in ways that are counterintuitive. For example, one would expect these heuristics to slot with greater space-efficiency when working from a list of skus sorted from allocation of greatest to least width. This is generally true, but not always! Furthermore it can happen that removing a sku forces the heuristic to use more shelves.

Much of the early work on packing was done by R. Graham in the context of machine scheduling; a good summary of his work can be found in [12]. More technical details, extensions, and pointers to related work can be found in [11].

The simplicity of these list-processing heuristics make it easy to adapt them to handle other complications. For example:

- If you want to use as few shelves as possible: Pack from a list in which the skus have been sorted from greatest width of allocation to least.
- If you wish to concentrate picking: Pack from a list in which skus have been sorted from most picks per width of allocation to least.
- If there are shelves of different heights: Put each sku on the least-high shelf on which it will fit.
- If skus must be stored in predefined groups (such as all hair care products together): When a new shelf is opened, designate it to hold only those skus in the same group as that sku for which the shelf was opened. Thereafter place the next sku in the appropriate shelf from among those designated for its group.
- If the orientation of the skus has not been decided in advance: One can adapt a variant such as Best Fit to, for example, choose the orientation and the open shelf that leave the least shelf space remaining.
9.3 Questions

Question 9.1 Can it happen that increasing the height of a shelf renders a given orientation of a sku less space efficient? Explain.

Question 9.2 Can it happen that increasing the height of a shelf renders every orientation of a sku less space efficient? Explain.

Question 9.3 Suppose the carton of dimensions $1 \times 2 \times 3$ from Table 9.1 was to be stored in a shelf of height 3; how would the values of space efficiency change? Compute the space efficiencies on shelves of height 2 and of height 1.

Question 9.4 Consider a carton of dimensions $1.25 \times 2.4 \times 2.75$ that is to be stored in a shelf of depth 9; Compute the local space efficiencies on shelves of height 1.5, 2, 2.5.

Question 9.5 How many shelves of width 10 are required to hold sku allocations of widths 5, 6, 7, 4, 8, 2, 1? Use the Next Fit heuristic to pack from the list as given. Do the same with First Fit. Use each heuristic to pack the skus from the list sorted in decreasing order of width.

Question 9.6 Repeat the previous exercise using the Best Fit and Worst Fit heuristics.

Question 9.7 Use the First Fit heuristic to slot the following skus onto shelves of width 10 so as to produce a packing with high pick-density (many picks per unit width of shelf). Justify your choice of list from which to pack.

<table>
<thead>
<tr>
<th>Width</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picks</td>
<td>50</td>
<td>200</td>
<td>100</td>
<td>350</td>
<td>120</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>
Part III

Order-picking
Chapter 10

Piece-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.

10.1 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 \[9\].

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.

Exploring these simple ideas has led to some practical applications within management science/industrial engineering. Here is one in warehousing.
10.2 Order-assembly by bucket brigade

“Bucket brigades” are a way of coördinating workers who are progressively assembling product along a flow line in which there are fewer workers than stations (work stations in the context of manufacturing; storage locations in the context of order-picking). 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 (or customer order), 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 (customer order) 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.

10.2.1 A model

Consider a flow line in which each of a set of items (customer orders) requires processing on the same sequence of $m$ work stations (storage locations), as in Figure 10.1. A station can process at most one item at a time, and exactly one worker is required to accomplish the processing.

The Normative Model, suggested in [4], 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 10.1 (Insignificant Walkback Time) The total time to assem-
ble a product is significantly greater than the total time for the workers to handoff their work and walk back to get more work.

Assumption 10.2 (Total Ordering Of Workers By Velocity) Each worker \( i \) can be characterized by a work velocity \( v_i \).

Assumption 10.3 (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 [4]).

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 [4]. 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 [4 2]. Their main results, slightly simplified, are as follows.

**Theorem 10.1** No matter where a given set of workers start,

- There is a unique balanced partition of the effort wherein worker \( i \) performs the interval of work:

\[
\text{from } \frac{\sum_{j=1}^{i-1} v_j}{\sum_{j=1}^{n} v_j} \text{ to } \frac{\sum_{j=1}^{i} v_j}{\sum_{j=1}^{n} v_j},
\]

(10.1)

so that each worker invests the same clock time in each item produced.
CHAPTER 10. PIECE-PICKING BY ‘BUCKET BRIGADE’

Figure 10.2: Positions of the worker 2 immediately after having completed the $k$-th order and walked back to take over the order of worker 1 (who has walked back to the start of the line to begin a new customer order).

- 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.

Before proving the result in general, we first argue that it is true for the case with two workers. 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 $x$ be the percent completion of the product held by worker 2 in the $k$-th photograph (that is, after a total of $k$ items have been completed), as in Figure 10.2. Then the next order will be completed after an elapsed time of $t = (1 - x)/v_2$. During that time, worker 1 will have traveled forward distance $v_1t$ and so the next handoff will occur at position $(v_1/v_2)(1 - x)$. We can therefore summarize the changes in the locations of the handoffs as the following dynamics function:

$$f(x) = (v_1/v_2)(1 - x).$$

This function is linear with negative slope, as illustrated in Figure 10.3. Furthermore — and importantly — because the workers have been sequenced from slower to faster, $v_1 < v_2$ and so $v_1/v_2 < 1$. Thus the slope of the dynamics function is of absolute value less than one and so is a contraction map. Figure 10.3 traces a sequence of handoffs at positions $x, f(x), f(f(x)), \ldots$, which is seen to converge to the fixed point, which is the intersection of the dynamics function with the identity, where $f(x) = x$.

Here is a proof of the general result for an arbitrary number $n$ workers.

**Proof** As before, let $x_i^{(k)}$ be the percent completion of the order held by worker $i$ immediately after completion of order $k$ and having walked back to start the next order. (See Figure 10.4).
Figure 10.3: Because the slope is of absolute value less than 1, the dynamics function converges to a globally attracting fixed point. In other words, the assembly line balances itself.

Figure 10.4: Positions of the workers after having completed $k$ products.
Then the clock time separating workers $i$ and $i + 1$ is
\[ t_i^{(k)} = \frac{x_{i+1}^{(k)} - x_i^{(k)}}{v_i}; \]
and the next item will be completed after time
\[ t_n^{(k)} = \frac{1 - x_n^{(k)}}{v_n}. \]

In the next, $k + 1$-st photograph, the clock-time separating workers $i$ and $i + 1$ becomes
\[
\begin{align*}
    t_i^{(k+1)} &= \frac{x_{i+1}^{(k+1)} - x_i^{(k+1)}}{v_i} \\
    &= \left( \frac{x_i^{(k)} + v_i t_n^{(k)}}{v_i} \right) - \left( \frac{x_{i-1}^{(k)} + v_{i-1} t_n^{(k)}}{v_{i-1}} \right) \\
    &= \left( \frac{v_i - 1}{v_i} \right) t_{i-1}^{(k)} + \left( 1 - \frac{v_i - 1}{v_i} \right) t_n^{(k)}. 
\end{align*}
\]

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 10.5 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 $v = (1, 2, 3)$.

### 10.2.2 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.
Figure 10.5: 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.
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 preempt 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.

10.2.3 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.
10.3. **BUCKET BRIGADES IN THE WAREHOUSE**

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.

### 10.3 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.

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% [5], while reducing need for management intervention (Figure 10.6). 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.
10.3. BUCKET BRIGADES IN THE WAREHOUSE

Figure 10.6: 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.)

- 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.

- 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 10.7 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 variance of pick rate decreased by 90%; thus bucket brigades were both more productive and more predictable, which made it easier to staff.
CHAPTER 10. PIECE-PICKING BY ‘BUCKET BRIGADE’

Figure 10.7: 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%.

10.4 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.

The ideas behind bucket brigades are simple:

- Abolish rigid assignments of work, which prevent people from going to where the work is.
- 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.
10.5 Questions

Question 10.1 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 10.2 How does low pick density affect the effectiveness of bucket brigades in order-picking?

Question 10.3 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 10.4 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 10.5 How does variability of work content from order-to-order affect the performance of a bucket brigade?

Question 10.6 What is the throughput of a bucket brigade if the workers are sequenced other than slowest-to-fastest?

Question 10.7 Would bucket brigades be appropriate for the warehouse shown in Figure 10.8? Assume any one sku has a quite small probability of being requested and workers pick fewer than five pick-lines on each trip.
Figure 10.8: Would bucket brigades be a good way of coördinating order-pickers in this warehouse?
Chapter 11

Pick-paths

Travel time to retrieve an order is a direct expense. In fact, it is the largest component of labor in a typical distribution center. Furthermore, travel time is waste: It costs labor hours but does not add value.

Travel time matters also because it affects customer service. The faster an order can be retrieved, the sooner it is available for shipping to the customer.

For example, McMaster-Carr is a distributor of hardware and tools. They distribute over 450,000 sku's from four North American distribution centers. Frequently a customer will order tools or parts that are required for a current project, such as building construction. Fast service is important because the project might be waiting for the part or tool. To ensure best service, McMaster-Carr almost always has the complete order packed and ready to ship within 30 minutes of order receipt.

For unit-load warehouses it is easy to estimate the work to retrieve orders, because each pallet corresponds to an independent trip to the storage location. This is not the case when the items are picked in less-than-unit-load quantities because the picker can accumulate items; and so, once in the warehouse, items, it makes sense to pick additional items so that total travel is reduced. This means that the travel time to the next storage location depends on which storage location was most recently visited; in other words, the sequence of picks now matters. The problem of order retrieval has become a problem of finding a short route amongst the items to be retrieved.

11.1 The problem of pick-path optimization

The problem of visiting a given set of locations as quickly as possible has been nicknamed the “Travelling Salesman Problem” (TSP) and has been much studied [17]. In general, the TSP is difficult in several senses:

- There is no known fast solution technique that works in general.
Randomly-generated instances, even small ones, can be surprisingly time-consuming to solve.

Optimum, or even good solutions can be complex and hard to describe.

Order-retrieval in a warehouse presents a special case of the TSP in which travel is constrained by aisles and this special structure makes it possible to find optimal solutions quickly by computer. However, despite marketing claims, most warehouse management systems do not support pick-path optimization. There are several reasons for this. The most important is that any optimum-finding algorithm must know the geometric layout of the warehouse, including distances between all pairs of storage locations; and most WMS's do not maintain this level of information. Such detailed information would not only be time-consuming to gather but would have to be specialized to every site and updated after any change in physical layout.

Finally, even if the WMS does support some kind of pick-path awareness, there remains the problem of communicating the path to the picker. A high-quality path is not useful if the order picker does not follow it. Typically the WMS tells the picker only the sequence of locations, not the actual path to follow. The picker must figure out the shortest path from location to location; and this can be hard to do because order pickers work under pressure and with only local information. Figure 11.1 shows the difficulty.

Incidentally, in this regard it can be more effective to pick from a paper pick list than from an RF device. With paper, the order picker can see at a glance the next few locations to be visited and can use his knowledge of the warehouse to plan his path. On the other hand, a typical RF device displays only the very next location to be visited, which makes it impossible for the order picker to improve the picking sequence. This situation may change soon as advanced telecommunications enables the WMS to pass the order pickers actual maps of pick paths to be followed.

11.2 Heuristic methods of generating short pick paths

How can we generate short travel paths that are realizable by an order picker who has no detailed map of the warehouse?

Imagine that a picker must visit all the storage locations of a warehouse; and suppose further that we can find an efficient global path to visit all these locations. We have to compute this efficient path only once and then we can use it many times: When a picker travels to retrieve the items of an order, we require that he simply visit the required locations in the same sequence as does the efficient global path. Thus the global path imposes a sequence that will be respected by all travel. When we receive a customer order the WMS simply sorts the pick lines by storage location so that they appear in the same sequence as the efficient global path. The idea is that if the global path is efficient the sub-path induced on each customer order is likely to be efficient as well.
Figure 11.1: An order picker has only local information, such as a sequence of locations and the view from an aisle (a), from which to determine a pick path that is globally efficient. It is hard for the order picker to know or account for the global layout of the pick area (b).
The problem of finding a good global path through the storage locations is known as the “Probabilistic Traveling Salesman Problem” or PTSP and there is a large literature \[.\] For the PTSP problem the main issue is length of the induced sub-paths. Within the context of order-picking there are additional issues.

### 11.2.1 Path outlines

A good global path should not only induce short sub-paths on the customer orders, it should also help the picker visualize where the next location and how to travel there most directly. We want a path outline that will induce a short pick path for most orders and yet is simple in structure so that order-pickers can understand it. An effective path outline will account for the physical layout of rack, where the most popular items are stored, and what a typical order looks like. In addition, management may devise simple rules by which the path outline can be adapted for the particular customer orders. By providing the order-pickers with a set of rules to adapt the path, they leverage the intelligence of the work force, rather than embedding the decision-making in the WMS software.

The simplest path outline is that along a single aisle, as shown in Figure 11.2. Such a path outline has the desirable property that any optimal path is consistent with this ordering. This configuration is typically found in a fast-pick module of a distribution center.

A commonly found path outline through static shelving is the so-called serpentine pick path illustrated in Figure 11.3. The path induces 1-way direction of travel in each aisle, which means that it may be possible to conserve floor space by using narrow aisles. However, location sequence might not give optimal travel paths, as in this instance where the picker would have to travel needlessly along the lengths of aisles 3 and 6. Unless a typical customer order visits every aisle, such a path can result in wasted travel.

One way of ameliorating this problem is to specify only a partial ordering among the storage locations. For example, Figure 11.4 shows an incompletely-specified serpentine outline that sequences the locations of the first few aisles from the left but thereafter sequences only the aisles themselves and not locations within the aisles. This imposes less *a priori* structure on the eventual pick path and relies on the intelligence of the order picker to adapt it appropriately. This example could be an effective path outline if the early (leftmost) aisles contain the more popular products. In this case the picker can construct a more efficient pick path by skipping aisles 3 and 6.

Because this path outline cannot guarantee in advance in which direction the picker may travel the later (rightmost) aisles, the right-side of the warehouse must allow 2-way travel for passing and so may need wide aisles.

Another common type of pick path is the *branch-and-pick*, which sequences only the aisles and not the locations within an aisle. The pick path passes a reduced set of locations (endcaps), which are where fastest-moving items should be stored (Figure 11.2.1) and the picker detours as necessary into the aisles. This
11.2. HEURISTIC METHODS OF GENERATING SHORT PICK PATHS

Figure 11.2: This picking area contains only popular sku’s and so every order is likely to require walking down the aisle. Thus an efficient route of travel is known in advance.

Figure 11.3: A serpentine pick path can result in unnecessary travel (in this case along aisles 3 and 6).
11.2.2 Product placement

To help pickers guess the best way to travel, one must make the geometry and addresses work together. For example, if locations that are close also have similar addresses, and *vice versa*, then the pick list also tells picker something about *where* the next address is in the warehouse. The global (layout of the warehouse) is embedded in the local (list of addresses).

In addition, it is generally better to store the most popular sku’s close to
11.3 Pick-path optimization

To compute optimal pick paths requires that the computer know the shortest distance between any two locations in the warehouse (and the corresponding route of travel). As of this writing, no warehouse management systems that we know of manage an explicit geometric model of the layout of the warehouse. Therefore true pick-path optimization is not currently done.
Eventually warehouse management systems will have geometric models of their warehouses; and advances in telecommunications will make it easy and economical to give picker precise travel instructions. Then there will be no reason not to take advantage of pick-path optimization.

The fundamental result in pick-path optimization is due to Ratliff and Rosenthal [21], who gave an algorithm for quickly finding the shortest tour of a set of locations in a warehouse. We will illustrate their ideas by giving a simplified version of their algorithm, which will generate near-optimal pick paths. The simplification is to restrict slightly the allowable patterns of travel so that the picker is forbidden to revisit a previously-visited aisle. In other words, we will find the shortest pick path subject to the constraint that the aisles cannot be visited out of their natural sequence.

Because of this restriction the suggested path may be slightly longer than the unconstrained optimum; but

- This algorithm is in the same spirit as the optimum-finding algorithm, but is much simpler and so is easier to program and to explain.
- Any unnecessary travel required because of the no-backtracking restriction is generally small.
- The generated path is simpler in structure than an optimum path and so easier for an order picker to understand and follow.

Following [21] we generate a pick-path by dynamic programming. This takes advantage of the fact that an optimum path can assume only a limited number of patterns at each end of an aisle. Our restriction that the picker can never revisit an aisle reduces the number of possible patterns to only two so that each order-picker can be imagined to follow this rule.

Pick all the required items from the current aisle and travel forward to the next aisle.

Note that there are two ways of realizing this rule: An order-picker can either

- Travel all the way through the current aisle, picking as necessary; or else
- Enter the aisle only as far as necessary to pick all required items therein and then return to the same end of the aisle from which it was entered.

The decision points are at the ends of each aisle, as suggested in Figure 11.8. Because there is no backtracking the picker must travel left to right across the warehouse and so we can consider sequentially the decisions to be made at each aisle. We can graphically summarize the sequence of decisions to be made as in the series of figures beginning with Figure 11.9.

Figure 11.13 shows the final graph summarizing the sequence of decisions to be made by the order picker. The shortest path in this graph corresponds to an efficient pick path through the warehouse. (The shortest path can be found by elementary algorithms from graph theory or discrete mathematics.)
11.3. PICK-PATH OPTIMIZATION

Figure 11.8: To visit the locations on the left, imagine a decision point at the end of each aisle. Because backtracking is forbidden, the order-picker must visit aisle $i$ before visiting aisle $i + 1$.

Figure 11.9: Enumeration and summary of paths from Aisle 1 to Aisle 2. Each candidate path is represented by an edge of the same length in the graph.
Figure 11.10: Enumeration and summary of paths from Aisle 2 to Aisle 3
11.3. PICK-PATH OPTIMIZATION

Figure 11.11: Enumeration and summary of paths from Aisle 3 to Aisle 4
Figure 11.12: Enumeration and summary of paths from Aisle 4 to completion.

Figure 11.13: The shortest path on the associated graph gives an efficient pick path in warehouse.
Figure 11.14: Only the lengths of the edges along each side of the graph need be updated to reflect new customer orders.

Notice that the connections of the network need be built only once; and for subsequent use only the lengths of the edges need be updated, as shown in Figure 11.14.

This approach can be extended naturally to handle the addition of cross aisles, although the work to solve increases quickly [21].

11.3.1 How to take advantage of optimization

For the effort of optimization to be worthwhile, it must be easy to implement a computed solution. But it can be hard for an order picker to know how to travel in the shortest way if they are told only the sequence of locations and not the paths from location to location. In fact, the shorter the pick path, the less likely it is to make sense to the order picker, who has only local information.

Nevertheless there are some simple ways to help the picker follow optimized pick paths. For example, the WMS could include instructions giving the direction the order-picker should travel after making each pick. Figure 11.15 shows an example pick list that would give sufficient information to specify routes, not just sequences of picks, for a pick path that followed a branch-and-pick outline (no backtracking to a previously passed aisle). Of course the effectiveness of this depends on how disciplined the pickers are. There is a trade off here: The more complex paths allowed, the more information that must be passed to the order pickers, which consumes communications bandwidth (and is therefore expensive) and which demands more of the order pickers.
11.3.2 How much is optimization worth?

Finally, it must be asked how much pick-path optimization matters. The answer is that it depends. There are some situations in which pick path improvement does not matter much at all. Both are extreme cases: First, if the locations of the most popular items are concentrated in a small area then the length of any reasonable pick path could not be much longer than the shortest possible pick path. And if each order requires visits to very many locations then the optimal pick path must traverse most aisles and could not be much shorter than any reasonable pick path. This means that one can reduce any need for pick-path improvement by batching orders or by placing product so that the most frequently requested items are stored near each other.

On the other hand, the warehouses that are most likely to benefit from pick-path optimization are those that have many items, most of which are slow-moving. Examples include warehouses distributing hardware, building supplies or aftermarket auto parts.

11.4 Summary

Despite advertising claims, most WMS’s do not support pick-path optimization. Instead they simply sort each pick list according to storage address (path outline). To make this method work best:

- Define path outlines that will generate short, understandable routes.
- Give pickers local rules to help them adapt the path outline.
- Place product to work with the path outline.

Pick-path optimization is in principle doable now but is rarely implemented for several reasons.

- The WMS must have a geometric model of the warehouse layout to compute shortest paths between pairs of locations.
• Limited communications bandwidth makes possible to tell picker where to go next but not the route by which to travel there.

• It does not always generate significant savings.
11.5 Questions

Question 11.1 Consider the problem of finding the shortest pick path through a warehouse with parallel aisles but no cross aisle. Give an example of when a shorter path is possible if backtracking is allowed.

Question 11.2 Devise a worst-case example to show how much longer a picker might be required to travel if backtracking is forbidden in a warehouse with parallel aisles (no cross aisle).

Question 11.3 The word “detour” generally connotes the requirement to go out of your way. In what sense is it a detour to enter a side aisle to pick an item as part of a larger order in a branch-and-pick system?

Question 11.4 In which of the following scenarios might pick-path optimization be economically justified?

- A very busy unit-load warehouse
- A warehouse in which orders arrive intermittently and each is picked immediately (no batching). A typical order is for 1–2 sku’s.
- A distributor of recorded music, with only a very few, very popular sku’s.
- A warehouse that does most of its picking from a few single-aisle pick modules onto conveyor.
- All of the above
- None of the above

Question 11.5 Which of the following best describes the effect of adding crossover aisles to a warehouse? (That is, aisles that run orthogonally across the main direction of aisles.)

- It reduces travel because it creates shortcuts.
- It complicates travel-planning because it creates more possible paths.
- It reduces storage capacity by taking aisle space.
- None of the above
- All of the above

Question 11.6 To pick a customer order requires that the order-picker visit the set of locations shown in Figure 11.16. Label the corresponding network with the appropriate distances so that solving the shortest path problem on the network generates a shortest tour visiting all the locations.
Figure 11.16: An order-picker must start at the leftmost dot, visit the shaded locations by traveling the aisles of the warehouse, and finish at the right-most dot.

**Question 11.7** Generalize the algorithm to find shortest pick-paths to account for back-tracking.

**Question 11.8** Generalize the algorithm to find shortest pick-paths to account for a single “cross-aisle” running horizontally through the middle of the warehouse. How does worst-case computational effort depend on the number of cross-aisles?
Part IV

Special topics
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 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 incoming 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 manoeuver.)

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. In particular, the layout can be changed simply by changing the labels on the doors of the crossdock.

There are two kinds of doors on a typical crossdock: Those reserved for outgoing trailers (for example, the “Miami trailer”) and those reserved for incoming trailers. The outbound doors are reserved for specific destinations but the incoming doors are not so specific and may be used by any incoming trailer (because, while departures are scheduled to specific destinations, we do not have full control over arrivals).

12.3.1 Congestion

As more freight flows across a dock, congestion increases, which interferes with the flow.
CHAPTER 12. CROSSDOCKING

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 shipings 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.

The need to dock freight suggests that busy outgoing trailers be parked away from the corners of the dock.

**Interference among fork lifts:** Despite the intention of moving freight simply “across the dock”, most doors will be to the left or right of a door with an incoming trailer and so a significant amount of freight must travel along the length of the dock. Most crossdock set up two forklift “highways”, one along each long side of the dock. (It is a good idea to set up two so that, when one is blocked, some freight can still flow.) However, the flow of forklifts back and forth along the length of the dock may be interrupted by forklifts making left hand turns into doors with outgoing
trailers. This effect can be reduced by parking busy outgoing trailers away from the very middle of the dock (which is also the most convenient location). Note that this works opposite to convenience, which tends to push busy outgoing doors towards the middle of the dock.

**Competition for drag line capacity:** Each door receiving arriving trailers will need empty carts from the dragline and, after loading a cart, will need empty cart positions on the dragline. This means that there will be diminished dragline capacity downstream of this door. If the door is far from a busy outgoing door then the region of diminished capacity can be large. This creates an incentive to intersperse incoming doors with outgoing doors. In particular, this suggests that current practice, which is to create large banks of incoming doors, reduces the capacity of the dragline.

### 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.

![Figure 12.3: A typical crossdock is built in the shape of the letter I, so that freight can flow across from incoming trailers to outgoing trailers.](image)

A typical dock, such as illustrated in Figure 12.3 is generally around 120 feet wide (36.6 meters). This is to allow freight to be staged on the floor. A standard (large) trailer is 48 or 53 feet long (14.6 or 16.2 meters) and a “pup” is 28 feet long (8.5 meters); all are 9 feet wide (2.7 meters). 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, or 30.8 meters) plus allow space for travel along the length of the dock (for example, two aisles, each 10 feet wide, or about 3.0 meters). We have seen docks as narrow as 80 feet (24.4 meters), 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.

Figure 12.4: Crossdocks have been built in a variety of shapes. Clockwise from upper left: An L-shaped terminal of Yellow Transport; a U-shaped terminal of Consolidated Freightways; a T-shaped terminal of American Freightways; an H-shaped terminal of Central Freight

A dock does not have to be shaped like the letter I. For example, Figure 12.4 shows docks in the shapes of an L, U, T, and H. But, as illustrated in Figure 12.5, 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. (This can be ameliorated by putting doors there that have less need to dock freight, such as those that move full trailers of all pallets to a single location.)

- On the inside of a corner, you lose door positions because trailers will
CHAPTER 12. CROSSDOCKING

Bottleneck

Dead space

Figure 12.5: Outside corners are vulnerable to congestion; and inside corners forfeit door positions.

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 —-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.

An X-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 I-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 12.6 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 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 X-shaped dock around 220 doors.
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 [6, 7].

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?

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.
Part V

Measuring warehouse efficiency
Chapter 13

Activity profiling

A warehouse is a complicated and busy place and it can be hard to get an 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.

Warehouse benchmarking is the systematic comparison with other warehouses to understand strengths and weaknesses.

13.1 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 skus (it is difficult to maintain a rational storage policy when the population of skus changes quickly)
- Average number of skus 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 (skus) 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

13.2 Warehouse activity profiling

13.2.1 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 skus account for 80 percent of the activity”); or in ABC analysis, which simply classifies skus as A (the small fraction of skus that account for most of the activity), B (moderately important), or C (the bulk of the skus but only a small portion of the activity).

One of the first things to know about any warehouse is what skus matter. This is usually a simple matter of ranking the skus 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 skus 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 skus 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 13.1 gives the ten most important skus by number of cases moved. Note that skus 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 skus by the number of times they were picked during some recent interval, such as in Table 13.2.
### Table 13.1: 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 UL SLIMFAST BONUS CHOC ROYALE</td>
<td>6</td>
<td>3,085</td>
<td>514.17</td>
<td>525</td>
</tr>
<tr>
<td>2 BANDAID FAMILY TWIN PACK</td>
<td>12</td>
<td>4,488</td>
<td>374.00</td>
<td>374</td>
</tr>
<tr>
<td>3 SATHERS PIXY STIX</td>
<td>12</td>
<td>4,320</td>
<td>360.00</td>
<td>267</td>
</tr>
<tr>
<td>4 GEMINI VIDEO TAPE T-120</td>
<td>24</td>
<td>7,260</td>
<td>302.50</td>
<td>471</td>
</tr>
<tr>
<td>5 HOUSE BRAND ASPIRIN 5 GR.</td>
<td>12</td>
<td>3,144</td>
<td>262.00</td>
<td>255</td>
</tr>
<tr>
<td>6 HOUSE BRAND COMPLETE ALLERGY CAPS</td>
<td>24</td>
<td>5,850</td>
<td>243.75</td>
<td>538</td>
</tr>
<tr>
<td>7 ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
<td>806</td>
</tr>
<tr>
<td>8 HOUSE BRAND PAIN REL CAPLETS 500MG</td>
<td>24</td>
<td>5,604</td>
<td>233.50</td>
<td>569</td>
</tr>
<tr>
<td>9 HOUSE BRAND GESIC</td>
<td>24</td>
<td>5,562</td>
<td>231.75</td>
<td>485</td>
</tr>
<tr>
<td>10 SATHERS S/F ASST SOUR MIX</td>
<td>12</td>
<td>2,520</td>
<td>210.00</td>
<td>206</td>
</tr>
</tbody>
</table>

### Table 13.2: 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 ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
<td>806</td>
</tr>
<tr>
<td>2 BEACH BAG SET</td>
<td>6</td>
<td>815</td>
<td>135.83</td>
<td>781</td>
</tr>
<tr>
<td>3 ACT II MICRO LITE BUTTER</td>
<td>144</td>
<td>21,276</td>
<td>147.75</td>
<td>570</td>
</tr>
<tr>
<td>4 HOUSE BRAND PAIN REL CAPLETS 500MG</td>
<td>24</td>
<td>5,604</td>
<td>233.50</td>
<td>569</td>
</tr>
<tr>
<td>5 ACT II MICRO WHITE CHEDDAR</td>
<td>120</td>
<td>15,870</td>
<td>132.25</td>
<td>553</td>
</tr>
<tr>
<td>6 HOUSE BRAND COMPLETE ALLERGY CAPS</td>
<td>24</td>
<td>5,850</td>
<td>243.75</td>
<td>538</td>
</tr>
<tr>
<td>7 HOUSE BRAND OINTMENT TRIPLE ANTIBIO</td>
<td>144</td>
<td>4,776</td>
<td>33.17</td>
<td>534</td>
</tr>
<tr>
<td>8 WRIGLEY PLEN-T-PAK BIG RED</td>
<td>192</td>
<td>12,792</td>
<td>66.62</td>
<td>530</td>
</tr>
<tr>
<td>9 WRIGLEY PLEN-T-PAK DOUBLEMINT</td>
<td>192</td>
<td>14,736</td>
<td>76.75</td>
<td>526</td>
</tr>
<tr>
<td>10 UL SLIMFAST BONUS CHOC ROYALE</td>
<td>6</td>
<td>3,085</td>
<td>514.17</td>
<td>525</td>
</tr>
</tbody>
</table>
CHAPTER 13. ACTIVITY PROFILING

Table 13.3: 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>UPPER DECK BASEBALL LOW #1992</td>
<td>432</td>
<td>70,524</td>
<td>163.25</td>
<td>423</td>
</tr>
<tr>
<td>ACT II MICRO BUTTER</td>
<td>144</td>
<td>34,362</td>
<td>238.62</td>
<td>806</td>
</tr>
<tr>
<td>SCORE 92 BASEBALL SERIES II</td>
<td>720</td>
<td>25,344</td>
<td>35.20</td>
<td>348</td>
</tr>
<tr>
<td>ACT II MICRO LITE BUTTER</td>
<td>144</td>
<td>21,276</td>
<td>147.75</td>
<td>570</td>
</tr>
<tr>
<td>TOPPS 92 WAX PACK BASEBALL</td>
<td>720</td>
<td>18,684</td>
<td>217.25</td>
<td>570</td>
</tr>
<tr>
<td>ACT II MICRO WHITE CHEDDAR</td>
<td>120</td>
<td>15,870</td>
<td>32.25</td>
<td>353</td>
</tr>
<tr>
<td>WRIGLEY PLEN-T-PAK DOUBLEMINT</td>
<td>192</td>
<td>14,736</td>
<td>76.75</td>
<td>526</td>
</tr>
<tr>
<td>ACT II MICRO NATURAL</td>
<td>144</td>
<td>13,284</td>
<td>92.25</td>
<td>377</td>
</tr>
<tr>
<td>WRIGLEY PLEN-T-PAK BIG RED</td>
<td>192</td>
<td>12,792</td>
<td>66.62</td>
<td>530</td>
</tr>
<tr>
<td>HERSHEY REESE PEANUT BUTTER CP</td>
<td>432</td>
<td>12,708</td>
<td>29.42</td>
<td>310</td>
</tr>
</tbody>
</table>

Table 13.4: Top ten office products measured by customer requests

<table>
<thead>
<tr>
<th>SKU</th>
<th>Picks</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAPE,TRANS,MAGIC,3/4&quot;W,1&quot;CO</td>
<td>2,225</td>
</tr>
<tr>
<td>CLIP,BINDER,SMALL</td>
<td>2,171</td>
</tr>
<tr>
<td>FOLDER,FILE,LETTER,1/3,MAN</td>
<td>2,163</td>
</tr>
<tr>
<td>CRTDG,INK,DESKJT,BK</td>
<td>2,157</td>
</tr>
<tr>
<td>DISK,3.5,DS-HD,IBM FRMT</td>
<td>2,097</td>
</tr>
<tr>
<td>MARKER,SHARPIE,FN,PERM,BLCK</td>
<td>2,075</td>
</tr>
<tr>
<td>NOTE,HIGHLAND,3X3,YELLOW</td>
<td>2,062</td>
</tr>
<tr>
<td>CLIP,GEM,SIZE 1,REGULAR</td>
<td>2,049</td>
</tr>
<tr>
<td>PAD,LEGAL,LTR SIZE,WHITE</td>
<td>2,009</td>
</tr>
<tr>
<td>PEN,BALL PT,MED,STICK,BK</td>
<td>2,008</td>
</tr>
</tbody>
</table>

Finally, consider the number of pieces sold of each (Table 13.3). 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 skus 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.

We find similar surprises in examining activity at a wholesale distributor of office products, for whom the ten most frequently requested skus were as shown in Table 13.4.

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,
13.2. WAREHOUSE ACTIVITY PROFILING

<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,11.5GL,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 13.5: Top ten office products by weight, a measure of shipping costs

we get a clue as to which skus account for most of our shipping costs, which are based most strongly on weight (Table [13.5]).

13.2.2 Statistical analysis

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.

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 skus, 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 skus 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 2.2). 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
13.2. WAREHOUSE ACTIVITY PROFILING

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 $100 \text{ bytes} \times 10^4 \text{ skus}$ can be described in about $1\text{MB}$.

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.
Keep in mind that this is primarily financial information. This is good in that it is likely to be very accurate; but it can also be misleading because the transactions it represents are financial and not necessarily operational. For example, it can happen that a sku is shown to have been requested in a negative amount; but this generally means something like the item was returned and restocked to the shelf.

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. 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 obtaining only 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.

**Warehouse layout and location addresses** A map of the warehouse allows us to see where each sku is stored. We can infer that an order-picker had to travel to this location to retrieve the product; and from the map we can infer something about the required travel. This will enable us to evaluate alternative layouts and warehouse designs.

This type of information is generally least standardized and may be found in the form of blueprints, sketches, CAD files such as **.dwg** format, and so on.

**Where and when 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

- Skus
- Product families
13.2. WAREHOUSE ACTIVITY PROFILING

- 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 SKUs 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 SKUs 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 SKUs 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 SKUs. For example, Figure 13.1 shows data for a warehouse in which the 5,000 busiest SKUs account for over 75% of the picks. This suggests that it might make sense to layout the warehouse to support a fast-pick area (discussed in detail in Chapter 8).

Another useful Pareto curve can be generated by examining the broken-case or full-case picks separately. Similarly, curves can be generated by examining
Figure 13.1: How picking is distributed over the skus. The horizontal axis lists the skus by frequency of requests (picks). By showing both actual values and cumulative fractions it is easy to see how concentrated the picking is amongst the most popular skus.

key subsets of skus, such as those from one region of the warehouse, or with common seasonality, or from the same product family. Analyzing subsets of skus 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 13.2 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 skus 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 seasonalities: 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 technology, 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
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 skus. 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 skus 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 skus, and so on, as in Figure 13.3. 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.

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 skus 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, skus 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
13.2. WAREHOUSE ACTIVITY PROFILING

Figure 13.3: About two-thirds of the orders are for a single line but these account for only about one-third of the picks.

<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>

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 skus 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 13.2.2).

With five product families A–E there are ten family pairs, with the following frequency among orders:

<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.
CHAPTER 13. ACTIVITY PROFILING

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 skus 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.

13.2.3 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.

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.
13.2. WAREHOUSE ACTIVITY PROFILING

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 skus 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.

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 skus 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”:

```
SELECT COUNT(*) FROM Lines
```
and to show all lines shipped on January 20, 2000 the command is

\[
\text{SELECT * FROM Lines}
\]

\[
\text{WHERE DateShipped = 2000-01-20.}
\]

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

\[
\text{SELECT SkuID, SUM(QtyShipped) FROM Lines}
\]

\[
\text{GROUP BY SkuID}
\]

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

\[
\text{SELECT SkuID, COUNT(*) FROM Lines}
\]

\[
\text{GROUP BY SkuID}
\]

To return a list of orders and the zone from which each line was requested, the command is

\[
\text{SELECT Lines.OrderID, Skus.Zone FROM Lines, Skus}
\]

\[
\text{WHERE Lines.SkuID = Skus.SkuID}
\]

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.

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 skus 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 skus, 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 skus may represent a snapshot of one instant. Consequently, you are likely to find skus 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 (16.15 meter) long trailer holds about 2,000 cubic feet (56.6 cubic meters) of product, which is about 20 pallets.
- A person walks about 4 miles (6.44 kilometers) per hour.

**Interpreting patterns**

It is easy to make mistakes in interpreting the data. Here are some to avoid.

**Beware of small numbers** If skus are unusually slow-moving, such as is typical at a service parts distribution center, then many skus will sell only tiny quantities during a year. A statistical fluctuation in demand that is small in absolute terms might represent a huge percentage increase and so mislead analysis.

At one site we estimated the time supply represented by the on-hand inventory of each sku so we could evaluate the appropriateness of inventory levels. This warehouse provided service parts to the US military and held hundreds of thousands of skus. Like most service parts warehouses, most skus were requested infrequently.

We received a snapshot of the inventory levels at an arbitrary date along with a sales history of the previous year. We estimated time supplies as follows: If a sku sold $Q$ units last year and had $q$ units currently on hand then we estimated there was enough to last for $q/Q$ years. Then we counted the number of skus with 1 year supply, 2 years supply, and so on. Here is the result.

The pattern is astonishing! There are three things to explain.

- Why are so many items grouped by time supply?
- Why are the most common values of time supply exactly six months, one year, and so on?
Why should we have observed exactly those values on an arbitrary date?

At first the pattern puzzled us, the client’s inventory managers, and the client’s purchasing department. Some suggested that the pattern reflected purchasing patterns; others thought that there might be requirements to hold, for example, a year’s supply of certain skus; and some thought that the inventory levels might have been set indirectly by the military training schedules that would consume the supplies. All these may be true to some extent and seem to explain why there might be peaks in the time-supplies, and while some can even explain why those peaks might occur at natural intervals such as one-half or one year, none explain why those natural intervals were observed on the arbitrary day for which we examined the data. Even if a particular sku is required to be held in quantities of at least one-year’s supply, why should we have found the inventories at exactly one-year’s supply? Why not at 1.2 years?

In fact, the explanation is quite different from any of these and it does not involve logistics at all: The observations can be best explained as a purely numerical artifact!

Recall that the distribution center holds very many skus that are very slow moving. There are tens of thousands of skus of which only one or two were shipped last year. Because these are such slow movers, the DC may hold only a few in stock. For example, there are thousands of skus that sold only a single unit last year and for which there are exactly two currently in stock, or two-year’s supply. For all these skus with few units sold and few units in stock, there are only a few possible values of estimated days-of-supply (units in stock divided by last year’s sales). For example, all the skus that sold 1 or 2 units last year and that have 1–3 units in stock now may be considered to have either 1/2

Figure 13.4: Why are there so many skus with exactly a half-year supply?
or 1 or 1.5 or 2 or 3 years supply on-hand. In other words, this is an artifact of small numbers.

Notice that because these are slow movers this condition persists. A sku that sold one unit last year and has two in stock now has an estimated two year supply; and, furthermore, this inventory level would be observed almost any time this sku was examined.

**Beware of sampling paradoxes**  If you want to understand how long skus are resident in the warehouse you will have to measure or estimate the time between arrival and departure and this can depend heavily on where you sample the skus. For example, if you observe the shipping department you can examine each pallet. Its date of departure is the current date; and you may find its arrival date in the inventory management system; and the time in residence is the difference in the current and arrival dates. However, such measurements will oversample those pallets that spend little time in the warehouse because those are the pallets most frequently on the shipping dock.

On the other hand, you might walk through the warehouse and estimate departure times. If you assume that demand for each sku follows a Poisson distribution then the time between departures is exponentially distributed. Thus if a pallet of the sku has been in the warehouse for $d$ days we expect it to depart after another $d$ days and so its estimated time in residence is $2d$ days. The problem with this estimate, however, is that it oversamples the slow-moving pallets, since those are the ones more likely to be found sitting in the warehouse.

The biases in these two estimates are illustrated in Question 13.4 at the end of this chapter.

### 13.2.4 Visualization

Readers are strongly encouraged to study E. Tufte’s “The Visual Display of Quantitative Information” and the companion books [24, 25, 26].

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.

- The defaults on standard office software, such as MS Excel, are generally poor. If you use these tools, you will need to intervene to make a quality graph.
13.3 Benchmarking

13.4 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 skus 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.

13.5 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 skus, as measured by number of cases purchased, reveals where order-picking labor is concentrated (Table 13.6).

It is not surprising that skus 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!
13.5. ON THE LIGHTER SIDE

<table>
<thead>
<tr>
<th>SKU</th>
<th>Pieces/Case</th>
<th>Cases Moved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mayonnaise</td>
<td>12</td>
<td>62,788</td>
</tr>
<tr>
<td>2 Charcoal</td>
<td>1</td>
<td>48,446</td>
</tr>
<tr>
<td>3 Sani-cat litter</td>
<td>1</td>
<td>46,632</td>
</tr>
<tr>
<td>4 2-liter gingerale</td>
<td>6</td>
<td>41,596</td>
</tr>
<tr>
<td>5 Charcoal briquets</td>
<td>1</td>
<td>40,660</td>
</tr>
<tr>
<td>6 2-liter cola</td>
<td>6</td>
<td>37,719</td>
</tr>
<tr>
<td>7 Apple juice</td>
<td>6</td>
<td>34,764</td>
</tr>
<tr>
<td>8 Granulated sugar</td>
<td>12</td>
<td>28,251</td>
</tr>
<tr>
<td>9 2-liter orange drink</td>
<td>6</td>
<td>27,301</td>
</tr>
<tr>
<td>10 Charmin bathroom tissue</td>
<td>16</td>
<td>24,379</td>
</tr>
</tbody>
</table>

Table 13.6: Top ten items moving from a grocery warehouse, as measured by number of 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 men’s belts: Easter in March/April; Father’s Day in June; back-to-school in August; Christmas. (Source: Italian Design Group, Atlanta, Georgia)
13.6 Questions

Question 13.1 Why is number of pick-lines generally a better indicator of labor than quantity-ordered or quantity-shipped?

Question 13.2 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 13.3 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 13.4 Suppose that a (in unrealistically tiny) warehouse has two pallet positions. You visit this warehouse on some randomly chosen day to study how long inventory stays in the DC. Unbeknownst to you, one pallet position has been occupied by the same pallet for the past 100 days; meanwhile a new pallet moves through the other position every day.

Suppose you take a census at the shipping dock and query each pallet that departs the warehouse over the course of the week and record its duration of stay in the warehouse. What value are you likely to estimate for the average time in the warehouse of a pallet? Suppose you stroll through the warehouse and check the time of arrival of each of the two pallets you find; what value are you likely to estimate for the average time in the warehouse of a pallet? Explain the biases built into each of these ways of estimating duration of stay in the warehouse. Are either of these estimates correct?
Chapter 14

Benchmarking
Part VI

Miscellaneous
Chapter 15

Warehousing around the world

Here is a brief survey of how warehousing issues differ around the world.

15.1 North America

North America is driven by mass consumption. Think WalMart. This enables huge economies of scale and, indeed, the trend has been for ever larger distribution centers and ever accelerating rates of product flow. As telecommunications enables better coordination along the supply chain, the uniformity of market and of distribution infrastructure allows fewer, more centralized and therefore larger distribution centers. The Amazon.com distribution center shown in Figure [15.1] is typical: One level, with conveyors and sortation equipment but little other significant automation. Such warehouses are generally built in the countryside surrounding major metropolitan areas, so that land is cheap but there is still ready access to large markets.

The fairly high costs of labor are held down somewhat by constant immigration into the US and Canada.

15.2 East Asia

Business in Asia has traditionally been based on personal relationships and less on computational models. Because of this tradition, data is not robust and not widely available; consequently the opportunities to improve operations by science are not fully developed at present.

In general, the most active economic areas are separated by lots of water, which means lots of product conveyed by air (for high-value or time-sensitive products) or ship (for bulky items or commodities). For both air and sea cargo, the large fixed costs increase incentives to consolidate freight. Consequently
one expects to see the emergence of strong regional hubs, such as Singapore and Hong Kong, to support this consolidation.

\section*{15.2.1 India}

India, like many developing countries, has both inexpensive land and low labor costs. Capital costs are relatively high in relation to the cost of labor and so there is less pressure to install specialized storage, even pallet rack. And because labor costs are low there is less incentive to increase efficiency. For example, it is not an attractive proposition to reduce labor costs by picking from flow rack: The labor savings cannot justify the cost of the rack.

In addition, warehouses in India distribute mainly to the local economy and so supply a market that is not wealthy. Consequently, the sku’s are not likely to be high cost items and so there is not much savings available from reducing inventories by precise timing. Consequently information technologies cannot generate much savings.

Finally, inefficiencies in transport make India in effect a collection of smaller markets. These inefficiencies include the physical, such as roads in less-than-ideal condition, as well as the administrative. For example, each state within India levies customs duties on freight transported across the border. This slows interstate commerce and increases the expense. Such factors increase the costs of transportation and so favor a strategy of having more, smaller distribution centers rather than fewer, larger ones, where the volume of activity could better justify capital investment.
15.2. EAST ASIA

15.2.2 China

A distinctive feature of the logistics scene in China is the seemingly boundless supply of very low cost of labor together with relatively inexpensive land. Consequently warehouses tend to be large, low buildings as in North America; but with some striking differences. For example, it is not unusual as of this writing to find a warehouse of 250,000 square feet with a single fork lift truck. The reason is that equipment is expensive but labor is cheap.

Despite cheap labor, China does have some capital-intensive warehouses, with the latest information technology and storage equipment. Such warehouses are most likely devoted to the distribution of high-value goods for export. Because such goods, such as consumer electronics, have high-value and short life-cycle, the warehouses can justify their equipment by substantial reductions in inventory costs.

The very different costs in the US and China sometimes leads to behavior that makes sense locally but may make the supply chain inefficient. For example, The Home Depot receives some Chinese-built product at its Import Distribution Center in Savannah, Georgia, USA. The shipping department in the warehouse in China de-palletizes freight in order to pack each trailer as tightly as possible for the drive to the sea port. Thus an expenditure of relatively cheap labor will reduce the relatively significant costs of equipment and transportation. But this means the product arrives in the US as loose cartons in containers and so The Home Depot must re-palletize the cartons before storage in deep, drive-in pallet rack. And most of it will, shortly after, be de-palletized once more when it is picked as cartons for shipment to stores.

15.2.3 Singapore, Hong Kong, Japan

Some economic powers such as Singapore, Japan, and Hong Kong suffer from limited space so land is much more expensive than elsewhere. Consequently, many of the warehouses are high-rise, such as shown in Figure 15.3. In addition, as first-world economies, labor in these places is expensive and so ware-
houses here are more likely to be automated. Freight elevators are likely to be bottlenecks to material flow in these facilities.

Space constraints have led to an interesting type of warehouse in Hong Kong and in Singapore: A multi-floor facility with no automation or elevators, but, instead, a spiral truck ramp so that trailers may be docked at any floor (Figure 15.4). In effect, each floor becomes a ground floor — but the cost is that significant land area, determined in part by the turning radius of a truck, is lost to the ramp and unavailable for storage. That storage space must be reclaimed up above.

This design is a clever and efficient way of using space if each floor is occupied by an independent tenant. But one must be careful if some tenant occupies multiple floors for then it may become necessary to shuttle trucks among floors. For example, a multi-storey warehouse with spiral ramp would be unsuitable as a local distribution center: Trailers departing with small shipments for many customers must be loaded in reverse sequence of delivery to avoid double-handling. But since the load may not match the layout of product amongst the floors, this could require much shuttling of the trailer among floors to load.

Similarly, this might be an inefficient warehouse in which to receive shipments of diverse items that may be stored on different floors: Because rear-entry trailers restrict the sequence in which freight can be accessed, a trailer may have to shuttle among the floors to deliver all the freight to the appropriate places. Alternatively, the trailer would have to be loaded to match the allocation of product among floors at the warehouse. In either case, extra work is required.

Sembawang Kimtrans claims to be among the first with a spiral truck ramp.
The main function of their Singapore facility is storage: It receives truckloads of product from the sea port or from factories in Malaysia and stores it for subsequent distribution to factories within Singapore. Thus a typical trailer-load, either in-bound or out-bound, is full of pallets of a single sku and so routing between floors is not an issue.

Over 50% of the product movement through this facility is by container.

The warehouse is 5 stories tall, with the first four floors used for warehousing and the fifth floor occupied by administration. Some of the floors are leased out to other companies and trucks to those warehouses are unlikely to visit other floors.

Each floor has four docks, each of which can accommodate a 45-foot trailer. The facility typically handles about 100 trailers in an 8-hour day.

About 30% of the space of the facility is consumed by the spiral ramp. Some of the ramp space is reclaimed for warehouse use as the hollow core of the ramp is used for miscellaneous handling of product.

In effect, the trailers and spiral ramp function as an AS/RS (automated storage and retrieval system). The point of this design is to get space utilization by building high storage; but to avoid installing automation, which can be a risky investment because inflexible. It also enables the owner to rent floors to different tenants.

Even simple automation, such as a freight elevator (lift), can be problematic because it is subject to queues and congestion, at least when deliveries are unscheduled, as they would be if different tenants occupy the facility.

The spiral ramp can simultaneously support more than 4 trucks travelling
in each direction. It would take a bank of at least 8 elevators to accomplish the same.

In addition, an elevator requires maintenance, which is not the case for the spiral ramp.

Finally, it must be observed that the most significant savings may be the reduction in land requirements. For example, if one imagines that the Sembawang Kimtrans facility in Singapore puts three warehouses in the space of 1.5, the land savings may be enough to pay for the cost of the building.

Incidentally, Hong Kong and Singapore make an interesting comparison. Both are premier logistics hubs due to splendid harbors, airports, and IT infrastructure; but they serve different purposes: Singapore is a point of transshipment for much manufacturing leaving Thailand, Malaysia, Indonesia, and other ASEAN nations and so it must be highly attuned to the challenges of handling international freight: Relatively little of it stays in Singapore; instead freight is likely to arrive from one country, receive value-added processing, and be forwarded to another country.

On the other hand, Hong Kong acts primarily as a logistics hub for manufactured goods leaving China. Consequently it receives goods from the same country, and frequently by truck or train; but it dispatches them overseas by ship or air.

15.3 South and Central America

- Not data robust
- Relatively low labor costs, though higher than in China and SE Asia.
- Developing markets, especially in Brazil. Despite competition from Asia, Mexico can get product quickly to market in the US and so it continues to grow as a manufacturer.

15.4 Europe

Warehouses in Europe, especially in Germany and France, are shaped by the relatively high labor costs and inflexibility of the workforce. These facts push designers to find engineering solutions rather than social solutions to logistics challenges. For example, there is a greater inclination to use automation than in comparable facilities in North America.

In the past, the economies of Europe were separate. More recently the economies are integrating into a common market, which will create economies of scale, which will likely lead to larger warehouses. However, urban areas, many of which have grown out of ancient towns, will still present challenges to the efficient flow of product.
Chapter 16

In preparation

The following topics are in various stages of preparation and will be available soon:

- Automation (carousels, miniloads): Throughput, control algorithms
- Cycle-counting: What to count and when
- Staff-scheduling: How many workers to hire and when/where should they work to meet shipping schedules
- Warehouse location: Where to locate that next warehouse or crossdock

Any other topics you would like to see? If so, please contact the authors (john.bartholdi@isye.gatech.edu).
Bibliography


Index

3PL
distribution center, 5
3rd party warehouse, 43

activity analysis, 173

batch, 18
benchmarking, 165
branch-and-pick, 134
broken-case pick, 173

carton, 61
case, 61
catalogue fulfillment
distribution center, 5
crossdock
defined, 151
cube-per-order, 87
cube-per-order index, 86
deadheading, 43, 58
dedicated storage, 10
distribution center
3PL, 5
catalogue fulfillment, 5
e-commerce, 5
total, 3
service parts, 3
third party logistics, 5
e-commerce
distribution center, 5
Equal Space Allocation, 75
Equal Time Allocation, 75
family
product family, 89
family-pairs analysis, 177

fast-pick area, 71
floor storage
pallets, 30
flow time
cycle time, 19
footprint, 50
forward pick area, 71
full-case pick, 173

grab, 17

honey-combing
in pallets, 30

join, in a relational database, 179

labor efficiency of a sku, 83
lane
storage lane, 50
lanes
pallet, 30
lift truck, 32
line, pick-line, 17
Little’s Law, 12
LTL
Less-than-Truck-Load, 151

mixed pick, 173

packing algorithms, 108
pallet
2-say, 30
4-way, 30
defined, 30
pallet pick, 173
pick density, 17
pick face, 17
pick-path optimization
pick path, 132
primary pick area, 71
Probabilistic Traveling Salesman Problem, 134
queueing system, warehouse as a, 12
random storage, 10
reserve storage, 71
retail
distribution center, 5
serpentine (pick path), 134
service parts
distribution center, 5
shipment integrity, 19
sku density, 18
slot, 78
Slotting, 107
SQL, Structured Query Language, 179
storage
dedicated, 10
random, 10
third party logistics
distribution center, 5
TL
Truck Load, 151
transfer batch, 176
transshipment warehouse, 43
Travelling Salesman Problem
TSP, 131
value-added processing
VAP, 21
warehouse activity profiling, 165
warehouse management system
WMS, 17, 25