Unlocking DC Throughput Capacity through Improved Flow

By

Ryan Cannon Morrison

B.S. Metallurgical Engineering, University of Utah, 2011

Submitted to the MIT Sloan School of Management and the Department of Mechanical Engineering in partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration and Master of Science in Mechanical Engineering

In conjunction with the Leaders for Global Operations Program at the

Massachusetts Institute of Technology

June 2018

Signature of Author Signature redacted
MIT Sloan School of Management Department of Mechanical Engineering May 11, 2018

Certified by Signature redacted
Dr. Stephen Graves, Thesis Supervisor
Professor of Management Science

Certified by Signature redacted
Dr. David Simchi-Levi, Thesis Supervisor
Professor of Civil and Environmental Engineering

Certified by Signature redacted
Dr. Stanley Gershwin, Thesis Reader
Senior Research Scientist Department of Mechanical Engineering

Accepted by Signature redacted
Dr. Rohan Abeyarante
Chair of the Graduate Program Committee Professor of Mechanical Engineering

Accepted by Signature redacted
Maura Herson
Director of MBA Program MIT Sloan School of Management

This page has been intentionally left blank

Unlocking DC Throughput Capacity through Improved Flow

by

Ryan Cannon Morrison

Submitted to the MIT Sloan School of Management and the Department of Mechanical Engineering on May 11, 2018 in partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration and Master of Science in Mechanical Engineering

Abstract

The large multi-channel apparel distribution center was built as a purely wholesale building in the 1900s. With distribution becoming more complex due to e-commerce and smaller orders for more frequent replenishment, the requirements for the distribution center have changed significantly.

The manner in which work flows through the building is influenced by batch size and release logic, which have not evolved to keep up with changing demand. Under the current flow of work, associates cannot stay busy working at the same processing (packing) station. Unstaffed processing stations are used as buffers, and associates move to unstaffed processing stations when they run out of work. The current flow prevents staffing all of the processing stations which significantly reduces the long-term throughput capacity.

This thesis lays out a methodology to improve flow and unlock throughput capacity by changing batch size and control logic to meet future demand. The key enablers of this thesis were:

1) The collection of data to evaluate progress at key steps in the process
2) A holistic understanding of how the system functions, as well as the implications on the long-term throughput capacity

Thesis Supervisor: Stephen Graves
Title: Professor, Management Science

Thesis Supervisor: David Simchi-Levi
Title: Professor, Civil and Environmental Engineering

Thesis Reader: Stanley Gershwin
Title: Senior Research Scientist, Department of Mechanical Engineering

This page has been intentionally left blank

Acknowledgements

I would like to acknowledge and thank many of the individuals who contributed to this project.

I would like to thank those at Nike for the opportunity to work on this project. My managers Angharad Porteous and Brian Dillard provided great guidance and generosity of time. I would like to thank Reagan Griffin and Teresa Carter for their leadership and perspective while I was in Memphis. Thank you to Neal Sheridan for letting me dive into his area and was willingly sharing his experiences and outlook. I would like to thank Michael Terry for sharing first his friendship and then his knowledge of the processes I was studying. Thank you to Chris Vetoulis for welcoming me to the Memphis team and also Felicia Miller and Terrani Burke for never making me feel like an intern.

So many people helped with the all phases of this project. I would like to thank Seth Howard and Seth Spofford from Logistipoint for helping with the big ideas. Thank you to Stephanie Coates, Jesse Feuerborn, Carl Jasper, Michael Hall, and Naveen Nandigam for lending your technical expertise and talent. I also want to thank Nausheen Kaul for supporting the LGO program as well as Brandon Stoker, and Dana Honn and the rest of the Supply Chain Innovation team. Thank you to Hugo Mora for sharing his time and knowledge of so many useful skills and insights.

I would also like to acknowledge and share my deep appreciation for my advisors, Stephen Graves, David Simchi-Levi, as well as Stan Gershwin. Their encouragement and direction from afar through the process was invaluable. Thank you to the MIT Leaders for Global Operations Program for its support of this work.

I am grateful for the opportunity to share the LGO internship experience at Nike with Matt Wallach and Abhishek Tambat who's experience, listening ear, and wise counsel helped me to find the success I sought after.

I would like to thank my parents for always encouraging me to learn and believing I could accomplish whatever I set out to do. They laid the foundation of who I am and how to work with others. I am also grateful for the support of my in-laws as well as their generosity and insightful experiences.

Most of all I would like to thank my wife Sarah, who's support this past couple years has been nothing short of outstanding. It has been a wild journey and I can't think of anyone who could have handled it with such grace and love and care as she did. She has put up with more uncertainty than anyone ought to, and all so I could pursue this opportunity to learn and grow. My children have been the bright spot of each day. My daughter Amelia recently told me that "internship means long road trip" and is always game for the next "internship". My son Theo who's smile and energy is infectious and who always seems to know where a ball is to play catch after a long day at work.

This page has been intentionally left blank

Note on Nike Proprietary Information

In order to protect information that is proprietary to Nike, Inc., the data presented throughout this thesis has been modified and does not represent actual values. Data labels have been altered, converted or removed in order to protect confidential information, while still conveying the findings of this project.

Table of Contents

Table of Figures

1 Introduction

The purpose of this thesis is to describe the methodology and findings of a project focused on the development of increased throughput capacity to meet future demand in an apparel distribution center. Throughput capacity is defined as the quantity processed in a given time and a common representation of capacity is units per day or units per hour. The project explores the impact of legacy equipment and legacy warehouse control systems that were designed for a single channel wholesale distribution on the flow of work through the distribution center that is now a multi-channel distribution center. Additionally, the collection of new data and a framework to analyze the current state and impact of potential recommendations will be presented.

The thesis is structured as follows: first, this section provides an overview of the project and problem facing the distribution center. Section 2 Literature Review presents relevant background information. Section 3 Current State details the current processes, and methods in which the processes are controlled and. Section 4 Analysis and Implications of Current State identifies the problems and opportunities of the current state with existing parameters. Section 5 Recommendations and Evaluation lays out both recommendations and evaluates the impact of those recommendations. Finally the concluding section summarizes what has been presented as well as discussing future opportunities. This thesis is based upon the experience of the author while at Nike in Memphis, Tennessee, USA at the main North America apparel distribution center.

1.1 Company Overview

Nike is a one of the premier athletic footwear, apparel, and equipment manufacturing, along with one of the most recognizable brands in the world. From its early days when it was known as Blue Ribbon Sports and distributing Tiger brand shoes from Japan, it has experienced unprecedented growth and expansion. This success has led to the opportunity to expand into new countries, new sports, and new products. Innovation and the development of new technology has also pushed the growth of Nike into new arenas. The growth and expansion has created a large complex company with thousands of SKUs and an ever-changing portfolio of products. In Figure 1, Nike revenue for the past ten years shows continued growth[1]. True to its origins, the company continues to be headquartered in the Portland, Oregon region in Beaverton, Oregon.

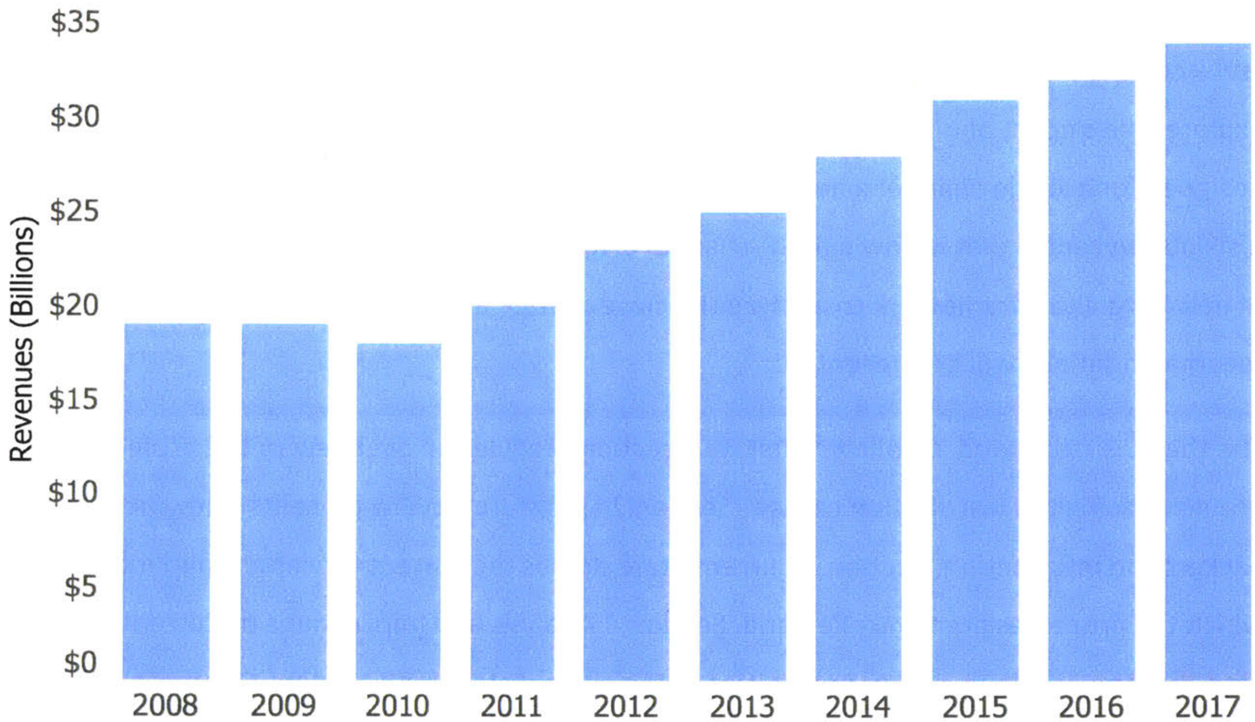


Figure 1: Nike Ten Year Revenues

As Nike grew, the majority of the products were sold through a futures wholesale model, meaning that the main demand signal were the orders placed by wholesale customers and major retailers over 6 months in advance. Often these were large orders going to large companies that would sell Nike products. Retail has evolved to include direct to consumer models through digital orders and retailers want the flexibility to quickly adjust to demand and consumer trends. This is often done through lower inventories, which require more frequent, smaller orders.

1.1.1 Nike Supply Chain Overview

The Nike Supply Chain is large and complex. Products are manufactured primarily in Asia and then shipped to various geographies around the world. Speaking to investors, COO Eric Sprunk shared that in fiscal year 2018, over 1.3 billion units will be shipped. The units will be manufactured in 566 factories, where more than 1 million workers are employed. The units will ship through 75 distribution centers to more than 30,000 retailers in more than 190 countries. The digital commerce business for both Nike owned and partners is expected to grow from less than 15% of Nike's revenue in fiscal year 2018 to over 30% over the next 5 years [2].

The North America supply chain, which includes the distribution center discussed in this thesis, is the largest by unit volume within the Nike global supply chain.

Figure 2 illustrates the Nike supply chain. Product flows from factories to consolidators, from consolidators to distribution centers, and from distribution centers to external customers (i.e. Dick's Sporting Goods or Footlocker), Nike Stores and Nike Factory Stores, or consumers (i.e. individuals ordering directly from Nike.com).

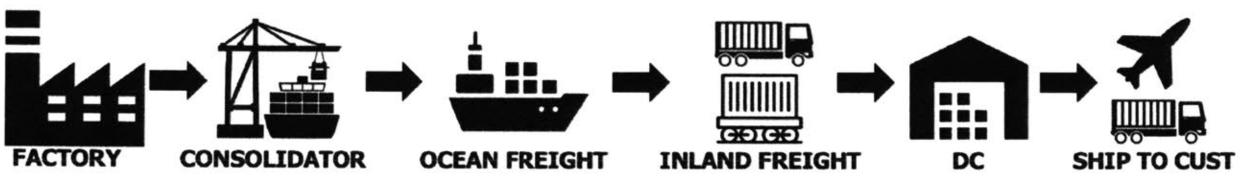


Figure 2: Nike Supply Chain

Nike has three product engines: footwear, apparel, and equipment. Nike also has three main product channels of selling products: wholesale, Nike owned stores, and Nike digital.

Nike has distribution centers in various locations throughout the United States, but the main distribution is centrally located in Memphis, TN with several distribution centers including its largest footwear distribution center. As the 2.8 million square foot facility opened in 2015, it has had an impact on the entire North America supply chain and particularly on the other distribution centers in the Memphis network. [3] As the new footwear DC was built, the company had to make a decision regarding which product engines and product flows would be absorbed by the new building and removed from one of the many other Nike distribution centers in Memphis including the apparel distribution center. Currently, the new footwear DC primarily distributes footwear and equipment. One of the other main distribution centers in the Nike Memphis network is the main apparel distribution center, where the author spent time during a 6-month internship.

1.1.2 Apparel Distribution Center Overview

The main apparel distribution center opened in the early 1990s and is a Nike owned facility. The facility was retrofitted in 1997 to be 1.2 million square feet. The primary product engine shipped through this building is apparel, but there is also equipment. No footwear ships through the main apparel distribution center. It is the largest apparel distribution center in the Nike North American network. The apparel DC also processes orders for all three distribution channels: wholesale, Nike owned stores, and Nike digital. Work is allocated to the floor in the DC when the inventory required for an order is in the distribution center and the planned ship date is within a predetermined window. The work is divided into various batches that will be discussed subsequently and these batches are determined by warehouse management system (WMS) settings.

1.2 Project Summary

The project objective is to improve operations within the four walls of the main apparel distribution center such that existing capacity can both be more fully understood and unlocked. The main Nike Apparel distribution center, which was built in the early 1990s and retrofitted in 1997, was designed for wholesale orders. With large wholesale orders, the process design and logic of the building was sufficient. As the retail industry as a whole and Nike's customers have evolved, so too has the order size and profile at the apparel facility. Customers are ordering smaller, more frequent orders. Digital orders which didn't exist at the design of the building have grown both in volume as well as strategic importance to Nike. The existing infrastructure, warehouse management system, and warehouse control system(i.e., batch size, release logic) have remained largely unchanged. There have been some changes within the distribution center to try and adapt the building to the evolving customer demand such as adding extra steps, and adding distinct work types (normal orders and expedited orders).

1.2.1 Problem Statement

Nike wants to understand how to develop greater throughput capacity to meet future requirements with existing infrastructure in the apparel distribution center. Growth in volume is expected as well as increased customer expectations of speed. Developing greater capacity will allow the apparel distribution center the flexibility and capability required to meet future demand and prevents Nike from having to build additional distribution centers. Developing greater throughput capacity within the existing footprint and utilizing existing infrastructure are critical in developing a cost-effective and implementable solution.

1.2.2 Hypothesis

The limiting factor to developing additional throughput capacity is the flow of work through the building, particularly the main flow path of tote picking. By adjusting batch size and/or release logic, the throughput capacity will be increased significantly to meet future requirements while using the same footprint and existing infrastructure.

1.2.3 Methodology

The methodology of this research will be first an understanding of current operations, then an operational analysis of the current state. Additional required data was identified and collected that allows a robust analysis of individual parts of the total process. This new data set allowed for scenario analysis of recommendations. High level simulation and experimentation were also utilized to analyze recommendations.

2 Literature Review

This section will look at existing practices and literature with regards to distribution center design, evolution of distribution centers, and distribution center sortation.

2.1 Distribution Center Design

General guidelines to warehouse design and operation according to John Bartholdi and Steven Hackman, professors at Georgia Tech, are:

- 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.[4]

Howard Coleman of MCA Associates consulting firm, discusses the importance of collecting data regarding the mix of products and how they are moving through the building to understand how to design or redesign a distribution center. With that data one can "apply multiple technologies to optimize the efficiency within the different "bandwidths" of the Pareto Curve" or the product mix. [5]

2.2 Evolution of Distribution Centers and Order Profile

More is expected of distribution centers and operations are increasingly complex. One example is the trend where warehouses assume more value-added processes or services (VAP or VAS), this can include ticketing/labeling, repackaging, and kitting. These services are being pushed on distribution centers by upstream manufacturers who want to delay differentiation as well as downstream retailers who lack space and resources. [4]

Marty Weil, an award-winning business journalist, wrote about the increasing complexity of distribution, even before the e-commerce boom where "pallet shipments gave way to cases and split cases; customer requirements have reached unprecedented levels". [7] This complexity is driving growth in the WMS (warehouse management system) and WCS (warehouse control system). Bartholdi and Hackman highlight that "the control afforded by software systems such as WMS that the pace of the supply chain has accelerated so much during the last 20 years" [4] According to Clint Reiser, analyst at ARC Advisory Group. "The growth is likely to continue as facilities are reconfiguring and retooling for e-commerce fulfillment. It's almost the singular reason for current growth." [7]. Weil also identified an example where upgrading a WCS had a major impact on flow and throughput. The new WCS led to "Increasing throughput dramatically by doubling it in less time than it previously took to ship half the volume." [7] In 2017 there were over 300 WMS vendors in the US alone with no company having more than 20% of the market. [4]

"E-commerce has really become a game changer in warehouse operations," according to Don Derewecki, a senior consultant with St. Onge Company. [8] In the "2017 Warehouse and Distribution Center (DC) Operations Survey", e-commerce is affecting network and facility size, investment in technology, productivity measurement, and labor availability. [8]

Distribution centers are becoming an increasingly important part of the supply chain as e-commerce puts more pressure to be faster, and cheaper. Matthias Winkenbach, director of the Megacity Logistics Lab at MIT stated that e-commerce and the impact on last mile delivery "also increases the need for greater visibility, everything from inbound and outbound deliveries to where are we in the picking process, what's the estimated time for getting an order ready for shipping and then connecting all of that with the transportation management system to minimize any time lost in the warehouse." [9]

2.3 Distribution Center Buffer and Sortation

A buffer within a distribution center is planned for work-in-process (WIP). Conway, Maxwell, and McClain of Cornell University researched the role of WIP in serial production lines. While not explicitly about a buffer within a distribution center, the role of WIP applies regardless of the environment. Interesting insights from this paper include that WIP serves to give each step of a production system some degree of independent action. The location of WIP is as important as the quantity. For lines with identical workstations and equal capacity buffers, to achieve a target production capacity, the buffer capacity should be proportional to the coefficient of variation of the processing time. To recover 80 to 85% of the capacity lost due to variability, the buffer capacity should equal 10 times the coefficient of variation. [10]

Shelia Bragg in a previous academic research project in a fulfillment center environment studied sorting and batch techniques. She recognized that if batching orders for subsequent sortation "there is a waiting phenomenon that occurs. A completed and sorted order will have to wait for all the other orders in the batch to be fully sorted and completed before the sorted batch is passed to the downstream stacking or packing labor." The cost advantages to improved labor productivity with large batch size is considered against the negative impact on waiting time and customer responsiveness. [11]

There are many different types or sorters, and the type of material along with the speed required based upon the distribution center design. For soft materials items like apparel a tilt-tray sorter is often used. A tilt-tray sorter does not need to know the orientation of the item it carries. Tilt trays serve as both conveyor and sorter; but they must circulate and so must be built as loop. [4]

Scott Eshleman, project manager for warehousing and distribution planning and engineering firm St. Onge, Ruff & Associates in reference to a trend of upgrading to high speed sorters said "Speed isn't the only factor. You may not need a new sorter if you fix other things upstream". [12]

3 Current State

This section will first look at the overall operation and processes of the apparel distribution center, subsequently a detailed explanation of the tote picking flow path processes will be presented. Finally the way in which the tote picking flow path processes are controlled through batches and warehouse control system logic will be discussed.

To understand the current state of the system, the author spent time observing the system in the apparel distribution center and meeting with operations directors, operations managers, area managers, operators, and associates. Time observing the system was critical to understanding the status quo. Due to the physical location of equipment and machine handling equipment and distribution center design, it is difficult to see the flow of work. The buffer-sorter-processing system is particularly difficult to see the flow of work as all three systems are stacked upon each other vertically. Climbing up into the system with experienced operators to observe the process and arrival, release, and movement of totes is critical to begin to visualize the flow of work and understand what process gates exist and necessary to understand what data to seek after. Time meeting with and observing work within the Flow Center, an arm of operations that monitors the various processes within the building and controls the release of work into the distribution center, is also essential to gaining a picture of the status quo. Lastly, but critical to gaining an understanding of the status quo is meeting with the subject matter experts for both the warehouse control system (WCS) and warehouse management systems (WMS), and data management.

3.1 Alternative Apparel DC Process Flows

There are four main flows through the apparel distribution center which are illustrated in Figure 3. Three of the four process flows will be explained in this section. RSR references receiving, pallet or case reserve storage, and retrieval. Products arrive to the distribution center from vendors in cases that typically range from 6 units per case up to 50 units per case and are placed into either case or pallet storage. Fulfillment in apparel distribution center is reactive, meaning that when work is released, product is pulled from storage and routed to the pick locations according to the number of units released in a given batch of work. The pick locations are not restocked proactively with a minimum safety stock. This occurs because there are more stock keeping units (SKUs) than pick locations, which in part is driven by the release of new SKUs every three months in conjunction with Nike's seasonal release of products. The pick locations also act as a forward storage location. Each process flow is triggered by orders that are to be fulfilled. More detail of how orders are released into the distribution center is explained in section 3.3.1 Overview of DC Batch Types. The output of each flow path consists of a number of outbound cartons that are routed to shipping and sent to the customer. Outbound cartons are created by the warehouse management to fulfill the customer orders. Customer orders can consist of one or many outbound cartons.

The first and simplest flow path is full case to shipping. When a customer orders a quantity of the same stock keeping unit (SKU) that corresponds to the quantity in a full case and no value added service (VAS) is required, then full cases of the same SKU are pulled from storage and routed directly to shipping where a label is applied and the full cases are loaded onto a truck.

The second flow path is full case that requires VAS, such as hanging the apparel, applying special tickets, etc.. The full cases are pulled from storage and routed directly to the sorter, which acts as a delivery mechanism to a processing station where the full case is unpacked by an associate, the requisite value added service is performed to the product, and the product is repacked into an outbound carton.

The third flow path is referred to as pick pack. Inventory is pulled from storage and routed to pick locations specified as pick pack locations. At these pick locations, products are picked and placed directly into their outbound carton and sent directly to the shipping dock. This process was developed to alleviate capacity constraints in the fourth and main flow path, the tote picking flow path, which is explained in the following section.

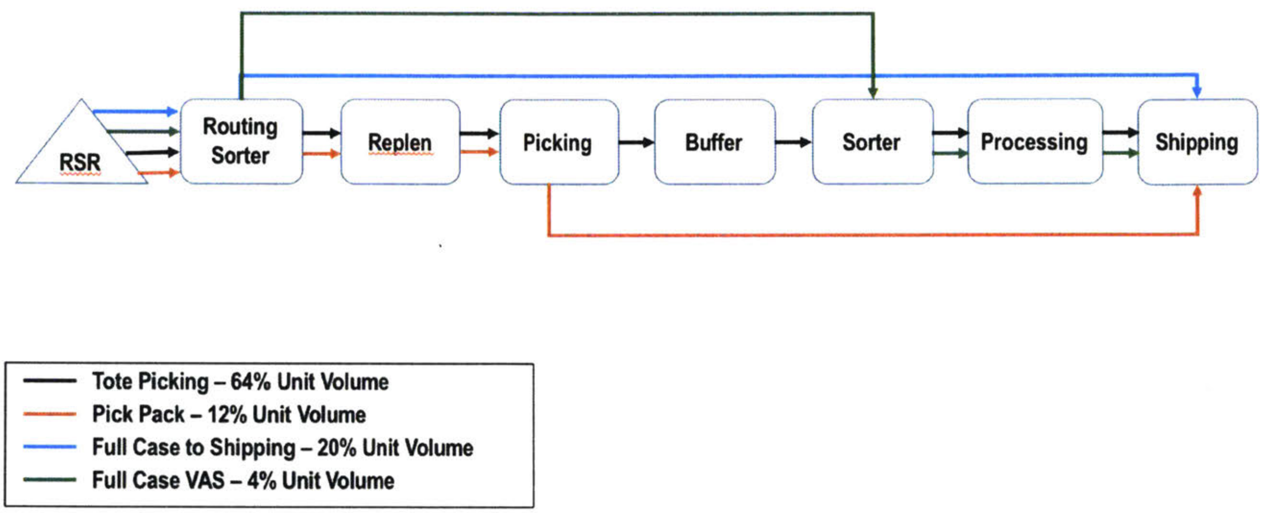


Figure 3: Apparel DC Process Flows

3.2 Tote Picking Process Flow

The fourth and main flow path is tote picking. This flow path is studied and discussed in detail through this thesis. As the largest and most complex of the flow paths in the apparel distribution center, it presents the greatest opportunity for improvement to efficiently adapt to the current customer demand. Similar to other flow paths, inventory is pulled from storage. After the routing sorter, the inbound cases are then sent to the pick locations. A detailed discussion of each process in the tote picking process flow path follows.

Picking

After full cases from case or pallet storage arrive in the picking areas on conveyor belt, associates replenish full cases to the assigned picking locations. After replenishment, the picking associates will pick units into plastic totes. All products picked into the same tote must be for the same outbound carton. A tote can consist of one item or up to as many items that can fit into one tote. The warehouse control system (WCS) indicates to picking associates when the tote is complete and the tote is placed onto a conveyor belt and sent to the buffer system.

Buffer

The buffer has multiple lanes where totes are diverted and held. The warehouse control system knows where each tote will be routed as it is inducted into the buffer system. Once the requirements for a buffer lane to be released (discussed in detail in subsequent sections), the buffer lane then releases a group of totes to the sorter. The buffer is essentially what it is named. It is there to buffer the upstream variability of totes coming from picking against the downstream variability of processing.

Sorter

The sorter acts as the delivery mechanism from the buffer to the processing chutes and stations. As totes from the same buffer lane are released, they travel down a conveyor system to one of five sorter induction points, where they are inducted onto a closed loop tilt tray system. All totes for the same outbound carton are assigned to a specific processing station chute and are tilted and delivered to that station as the tray passes the chute.

Processing

After totes are delivered to a processing chute, they remain in that processing chute until an associate pulls them off the bottom of the chute onto the processing station. The processing chute holds multiple totes and can hold totes that belong to multiple outbound cartons. As an illustrative example, there could be 10 totes in a processing chute belonging to 10 distinct outbound cartons or 10 totes in a processing chute belonging to 2 distinct outbound cartons. Processing chutes also are not necessarily full. A processing chute could be 20% full or 80% full and still contain totes corresponding to multiple outbound cartons or 1 outbound carton.

An associate will be given the instructions of what needs to be packed in the outbound carton as they grab the nearest tote from the processing chute and scan it. There are two main types of work performed in processing. The first is when some value added service (VAS) is required such as placing the apparel on hangers or adding special stickers. The second type of work is referred to as straight pack and the units are simply placed from the totes into an outbound carton. The processing step is what allows units that were picked into multiple totes from various pick locations in the building to be consolidated into one outbound carton. Ultimately the warehouse management system determines what units will go into each outbound carton. Once the outbound cartons are complete, they are then sent to shipping to be loaded into trucks.

3.3 Tote Picking Flow Path - Controls

Now with a basic understanding of the tote picking process, this section will address how those processes are controlled. At a high level, the processes are controlled through various batches and logic that controls how those batches move within each process as well as move to the next process. First, there will be an overview of the various batches within the system, and then a detailed view of the logic that controls each process. The logic that controls the buffer and release of work from the buffer is the most complex and discussed more in depth.

3.3.1 Overview of DC Batch Types

This section discusses the various batch types and provides the background information necessary to understand the logic that controls the processing of picking, the buffer, the sorter, and processing steps.

<u>Dispatch</u>

A dispatch is the largest batch of work. The process of dispatching or creating a dispatch is when the planning group identifies a set of customer orders that is able to be released and groups it in a way that will meet the required service level as well keep the work in progress (WIP) within the desired operational range. Dispatches can be created as frequently as every hour for small dispatches or once a day. Dispatches can range in size from tens of units to tens of thousands of units that are to be picked, boxed and shipped. Creating a dispatch does not actually release work to the distribution center to begin processing, but it determines the pool of work that can be released. The group controlling the release of work can release a portion of a dispatch at a time. Each portion is known as a wave and is discussed below.

At the dispatching process the warehouse management system generates a series of smaller batches and corresponding IDs. It is important to note that these smaller batches are created and fixed prior to release of the first batch (or wave). This distinction is important because by creating these batches upfront, there is less flexibility in the system. Batches and IDs for waves, picking assignment, buffer lane batch, totes and outbound carton are generated by the warehouse management system.

Various parameters have been set in the warehouse management system (WMS) that typically determine the maximum size of various batches in the system. Batches can be smaller if the planning group identifies the need to release smaller batches of work to either expedite through the system or supplement what currently exists in the system.

<u>Wave</u>

A wave is the smallest division of work that can be released to begin processing at one time, although multiple waves can be released simultaneously. The maximum number of totes in a wave is in the thousands and the maximum wave size has been agreed upon by planning and operations. This setting can be increased or decreased in the warehouse management system. If a large dispatch is allocated to the floor, the warehouse management system, based on the content of the work and various parameters will divide the dispatch into multiple waves where each wave corresponds to the maximum wave size setting. All work within a single wave corresponds to the same dispatch. Multiple waves can belong to one dispatch, but multiple dispatches cannot belong within one wave.

<u>Tote</u>

At the dispatching process each unit to be picked and processed is assigned to a tote ID. The units in a tote can vary from 1 unit per tote up to the cubic capacity of the tote which may fit 5-7 bulky items like sweatshirts or 30-40 small items like t-shirts. The items that are picked into a tote are all destined for a single outbound carton. For each outbound carton, there may be multiple totes. However multiple outbound cartons cannot belong within one tote. Totes are the building block of the other batches

Picking Assignment

Within a wave each tote is assigned to a picking assignment. A picking assignment is a group of totes that are assigned to a single associate for picking. All totes within the same picking assignment must belong to the same wave.

Buffer Lane Batch

A buffer lane batch is a group of totes that will be collected at a buffer lane and then released from the buffer lane to be delivered to the processing stations. This batch contains totes from the same wave. Totes belonging to the same outbound carton will belong to the same buffer lane batch. However a buffer lane batch can include multiple outbound cartons. The buffer lane batch size historically has been set at the max number of totes that can physically fit in a buffer lane.

Outbound Carton

Outbound cartons are made up of the units that the warehouse management system (WMS) designates to be placed in the same box and shipped to the customer. An outbound carton can consist of one or multiple totes and similarly could be one or multiple units.

An illustrative example of the batches created by the dispatching process is below in Figure 4. This example shows one wave, two distinct buffer lane batches, 28 distinct totes, 5 distinct outbound cartons. The 28 totes are coming from 8 distinct picking assignments. Totes with the same carton ID, buffer lane batch ID, and pick assignment are highlighted by the same color. This visually illustrates how the batches are nested relative to each other. Totes are unique and have varying number of units in them. Totes make up a buffer lane batch size. Within a buffer lane batch there can be multiple outbound cartons and picking assignments.

Tote ID	Units Per Tote	Wave Number	Carton ID	Buffer Lane Batch ID	Pick Assignment
11	4	1	101	1001	10007
12	4	1	101	1001	10001
13	7	1	101	1001	10001
14	1	1	102	1001	10007
15	1	1	102	1001	10001
16	1	1	102	1001	10008
17	1	1	102	1001	10006
18	1	1	102	1001	10005
19	1	1	102	1001	10001
20	1	1	102	1001	10006
21	1	1	102	1001	10008
22	1	1	102	1001	10005
23	1	1	102	1001	10001
24	1	1	102	1001	10002
25	11	1	103	1002	10005
26	6	1	103	1002	10008
27	1	1	103	1002	10003
28	4	1	103	1002	10007
29	2	1	103	1002	10006
30	2	1	103	1002	10002
31	5	1	103	1002	10003
32	5	1	103	1002	10004
33	1	1	104	1002	10001
34	9	1	104	1002	10001
35	7	1	105	1002	10002
36	5	1	105	1002	10008
37	2	1	105	1002	10004
38	5	1	105	1002	10003

Figure 4: Example of Batch ID Creation

3.3.2 Process Logic

This section details the logic that controls each step of the process within the tote picking process flow path. The key process is the logic that controls the release of totes from the buffer.

<u>Picking</u>

As a wave is released to be processed, all of the associated picking assignments are assigned to available associates in the picking department. Each associate can only be assigned one picking assignment and conversely each picking assignment can only be assigned to one associate and cannot be divided. Picking assignments can vary significantly in size from a few totes to 40+ totes depending on the criteria in the warehouse management system, the customer order profile, and where individual SKUs are slotted. For totes to be in the same picking assignment, they must be from the same wave and the same pick zone. The pick zone size can be configured within the WMS. A picking assignment is complete when all of the totes have been picked and placed on the conveyor to the buffer. Ideally all picking assignments within a wave would complete at nearly the same time and meet up in the buffer.

Buffer

After the units assigned to a tote are picked, the tote is put onto a conveyor belt, and moved to the buffer. Each tote is identified by its buffer lane batch, which was created before any work was performed during the dispatching process by the warehouse management system. Each buffer lane batch will be assigned to a unique buffer lane and all the totes from the same buffer lane batch will be diverted to that buffer lane where they are held until that buffer lane is released.

The logic that controls the release of a buffer lane is complex despite the fact that there are two seemingly simple requirements for a buffer lane to release.

Requirements for a buffer lane batch to be released from the buffer to the sorter:

- Requirement 1: All totes from a buffer lane batch must have arrived in the assigned buffer lane
- Requirement 2: All cartons from a buffer lane batch must be assigned to a unique, available station

Requirement 1

Requirement 1 is straightforward, but its impact is much more complex. Within a buffer lane batch, there can be totes corresponding to multiple outbound cartons. Most buffer lane batches have multiple outbound cartons. Consider an illustrative example buffer lane in Figure 5: each color represents a unique carton and each square represents a tote, and there are 5 unique outbound cartons and 14 distinct totes. The first tote to arrive is the red tote on the left. Although the two green totes complete one unique outbound carton, they are stuck in the buffer lane until the last gray tote arrives. When the 14th and final tote arrives, only then is the buffer lane eligible to release. At this point Requirement 1 is satisfied.

Figure 5: Example Buffer Lane Completion

To draw an analogy to a process which is familiar to most, consider a grocery store shopping experience. Gathering items in a shopping cart is analogous to the picking process. Before the shopping experience begins, five people are randomly put into a "shopping group" (buffer lane batch). Once a shopper has all of their items gathered in the cart, rather than proceed to a checkout stand, they go to a "shopper buffer" (buffer). Each shopper must wait until all five shoppers from the "shopping group" arrive. When all five shoppers arrive, then the "shopping group" then becomes "checkout eligible" (releasable). The group would then satisfy Requirement 1, but could not yet travel to the cash registers without first satisfying Requirement 2. This would impede flow significantly at a grocery store where as soon as an individual gathers all of their groceries, they immediately get in line rather than waiting for other shoppers.

Requirement 2: All cartons from a buffer lane batch must be assigned to a unique, available station

Requirement 2 is more complex and it is critical to understand what is considered a unique and available station. A unique station means that from one buffer lane batch, only the totes from one outbound carton can go to the same station. Continuing the above illustrative example of a buffer lane batch with 14 total totes and 5 unique cartons indicated by color, the 5 unique cartons will need to go to 5 unique processing stations, see Figure 6. In Figure 6, the color boxes correspond to totes that belong to the buffer lane batch introduced in Figure 5. The light gray boxes in each process station chute represent totes that had previously arrived at process station. The totes corresponding to the green and red outbound cartons cannot be sent to the same station. The need for each outbound carton to have a unique station is due to the fact the totes arrive to the buffer in a random order from picking (see Figure 5) and the totes are only permitted to travel on the sorter one full circuit at most on the closed loop sorter. If the totes were permitted to travel around the closed loop sorter more than one revolution, then it would be possible for the totes of two outbound cartons to sort to the same station and still be sequenced together. For example the green totes would be tilted and delivered to the process station chute during the first revolution and then during the second revolution the red totes would be tilted and delivered to the same process station chute.

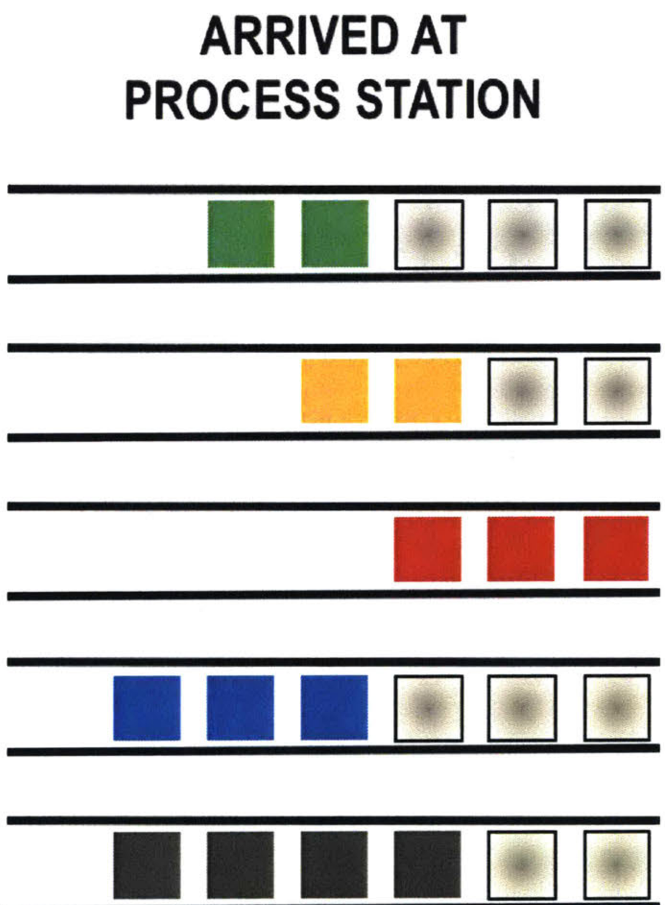
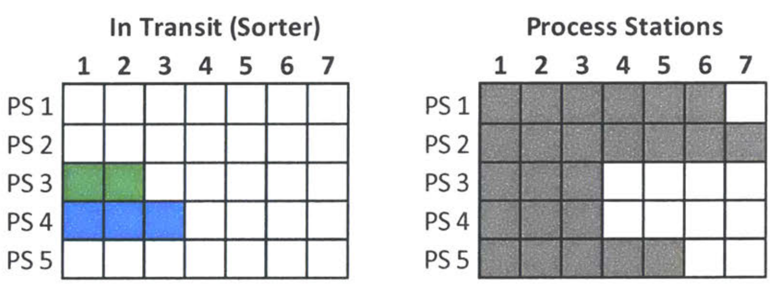


Figure 6: Unique Process Stations Required for Release

In addition to requiring a unique processing station, it also requires an available station. An available station consists of two criteria. First, there must be no other totes previously released in transit to the processing station, and second, there must be capacity in the processing chute to receive the number of totes for each outbound carton in the buffer lane batch that is to be released. This is more clearly illustrated visually in Figure 7. In Figure 7, we see an illustrative example where buffer lane batches are 5 totes and process station chute capacity is 7 totes. Only Process Station 1 and Process Station 5 are available by meeting both criteria of no totes being in transit and having capacity in its respective chute. Process Station 2 is unavailable because it has no capacity in its process station chute. Process Station 3 and Process Station 4 are unavailable because there are totes in transit. It is worth noting that between the two available process stations there is only capacity for a total of 3 totes which is less than a complete 5-tote buffer lane batch and the system would not be able to release any more buffer lanes until either the in transit totes arrived or totes are packed into an outbound carton by an associate and removed from the Process Station 2 chute.



In Transit	Capacity	Available?
No	1 Tote	Available
No	0 Totes	NA
Yes	4 Totes	NA
Yes	4 Totes	NA
No	2 Totes	Available

Figure 7: Simplified Example - Available Stations

Returning again to the grocery store analogy, and the scenario where all five shoppers are in the "shopper buffer" (buffer) and are checkout eligible. Before the shoppers can be released, all five shoppers must be assigned to five distinct cash registers or no shoppers can be released to the cash registers. For a shopper to be assigned to a cash register, there can be no other shoppers in transit to the cash register. Also, before a shopper is assigned, the potential cash register must have enough room on the conveyor belt (processing station chute) after considering all of the shoppers who are currently at the register.

Sorter

The travel time from the buffer to the processing chutes is relatively consistent if no jams occur and no totes get stuck. It can vary depending on the origin in the buffer and the final destination, and any congestion on the conveyor system before being inducted onto the sorter trays. This is an automated delivery system and is controlled by a specialized warehouse control system that controls equipment and movement of totes through a series of PLCs. There is individual monitoring the system for jams or issues and can intervene manually if required. While a specific processing station has totes in transit on the sorter, it is marked as unavailable in the system and that status is visible to the individual operator monitoring the system. The operator monitoring the system can also see which buffer lanes are empty, partially full, complete and releasable and actually releasing.

Processing

The controls that exist within processing is the ability to turn off processing station or specialize them to only receive a specific type of work, for example VAS only or straight pack only. Outside of these controls, the associate simply packs what is in front of them. They can be directed or assigned to specific areas by management. As associates pack down all of the totes at a given processing station, they routinely run out of work and have to move to another processing station to continue working. The system operator and management have ability to monitor the percentage of each wave that has been processed and sent to shipping.

4 Analysis and Implications of Current State

The analysis section is considering the data made available in the current state analysis and then further identifying possible changes and scenario analysis of potential changes to the system to develop recommendations.

4.1 Data Collection

The granular data used for both the analysis of current state as well as evaluation of recommendations are a series of snapshots of the live system. For additional detail regarding the data preparation please see Appendix A: Distribution Center Data.

The data captured for each tote includes tote ID, buffer lane batch ID, wave ID, carton ID, tote pick time, units per tote, carton pack time. Data could be grouped by wave, buffer lane batch, and carton to generate the insights highlighted in the analysis below.

4.2 Assumptions

Due to the data available, various assumptions are required for the analysis. The first is that the data collected is accurate and data integrity is maintained. The data collected during this analysis was validated by cross-validating with other more established data sources, and direct observation

Travel time on the conveyor system from picking to the buffer is assumed to be constant and can be ignored. This assumption is necessary due to the data available to be gathered. All totes are coming from the picking while one tote may have to travel from different areas in picking to a different area in the buffer, at an aggregate level the average travel from picking to the buffer can be assumed to be constant

For example, if data of when a specific tote arrived at a buffer lane was unavailable, the tote pick time was used as a proxy with the assumption that travel time on the conveyor system from the pick location is approximately constant.

4.3 Current State Analysis

The current state analysis section is structured as follows. First, an examination of the order profile, and second, an analysis of median process times for segments of the tote picking flow path.

4.3.1 Order Profile

The order profile is defined as the relationship between totes, units, outbound cartons, and buffer lane batches. The order profile is a function of the customer orders and how the warehouse management system generates the batches of work to meet the demand. It includes understanding units per tote, totes per carton, and finally cartons per buffer lane batch.

Units per Tote

The units per tote can vary from 1 unit per tote up to the cubic capacity of the tote which may fit 5-7 bulky items or 30-40 small items. The average units per tote is around the 25th percentile of the distribution of units per tote in Figure 8. Figure 8 shows the distribution of units per tote. The 0 percentile on the x-axis represents the minimum units per tote and the 100 percentile represents the maximum units per tote. It is especially noteworthy 38% of cartons are at the minimum end of the units per tote distribution. This is indicative of smaller orders processing through the building.

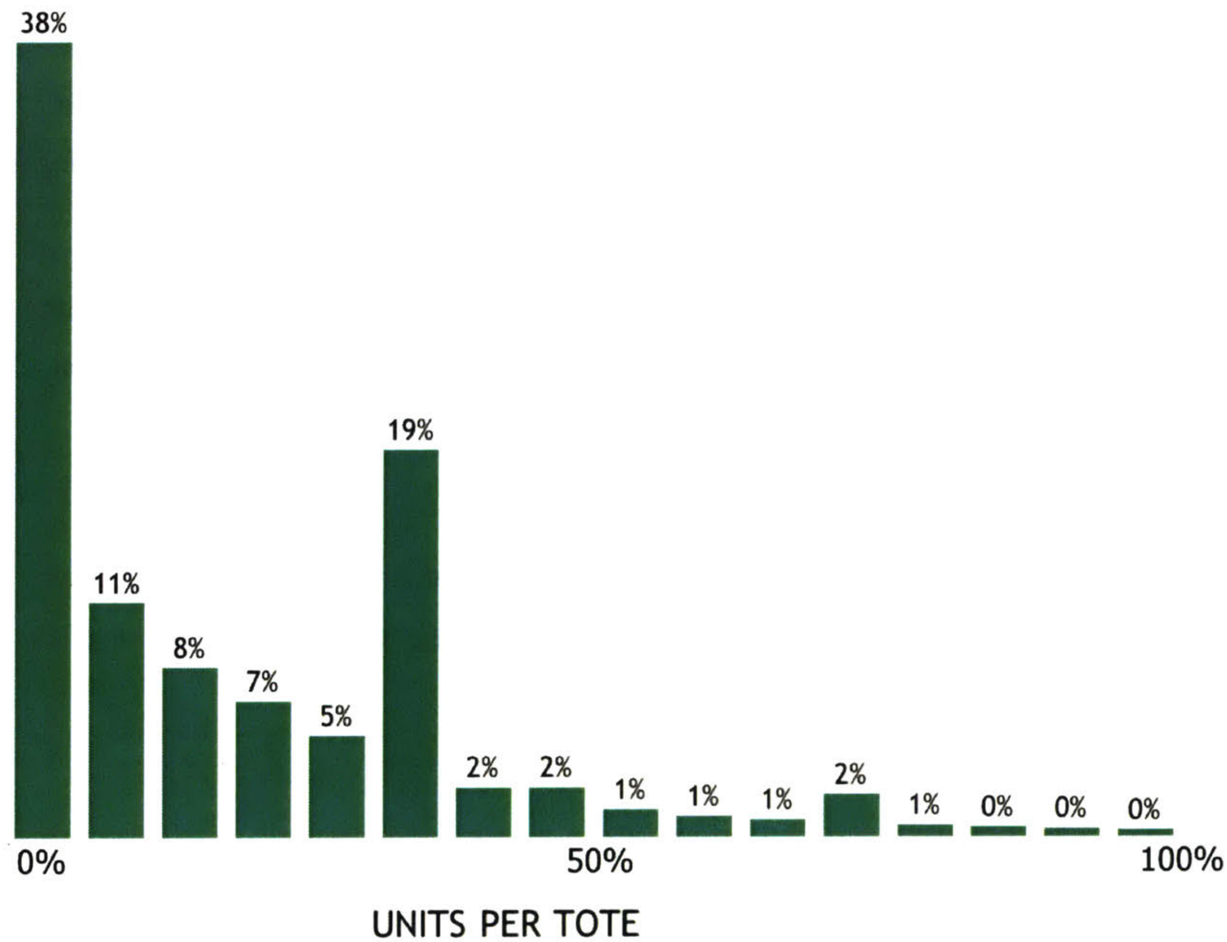


Figure 8: Units per Tote Histogram

Totes per Outbound Carton

Totes per outbound carton can vary from 1 tote per outbound carton up to 30 or more totes per outbound carton. The average is around the 15th percentile in the distribution of totes per outbound carton as seen in Figure 9. Figure 9 shows the distribution of totes per carton. The 0 percentile on the x-axis represents the minimum totes per carton and the 100 percentile represents the maximum totes per carton. In Figure 9 the pattern of small order types is seen in that 46% of all cartons are at the minimum end of the totes per carton distribution. The significance of cartons with only 1 tote, is that the outbound carton is complete at the point of picking, and the tote does not need to be consolidated with other totes at the buffer. Currently these totes are sent to the buffer anyway as part of a buffer lane batch.

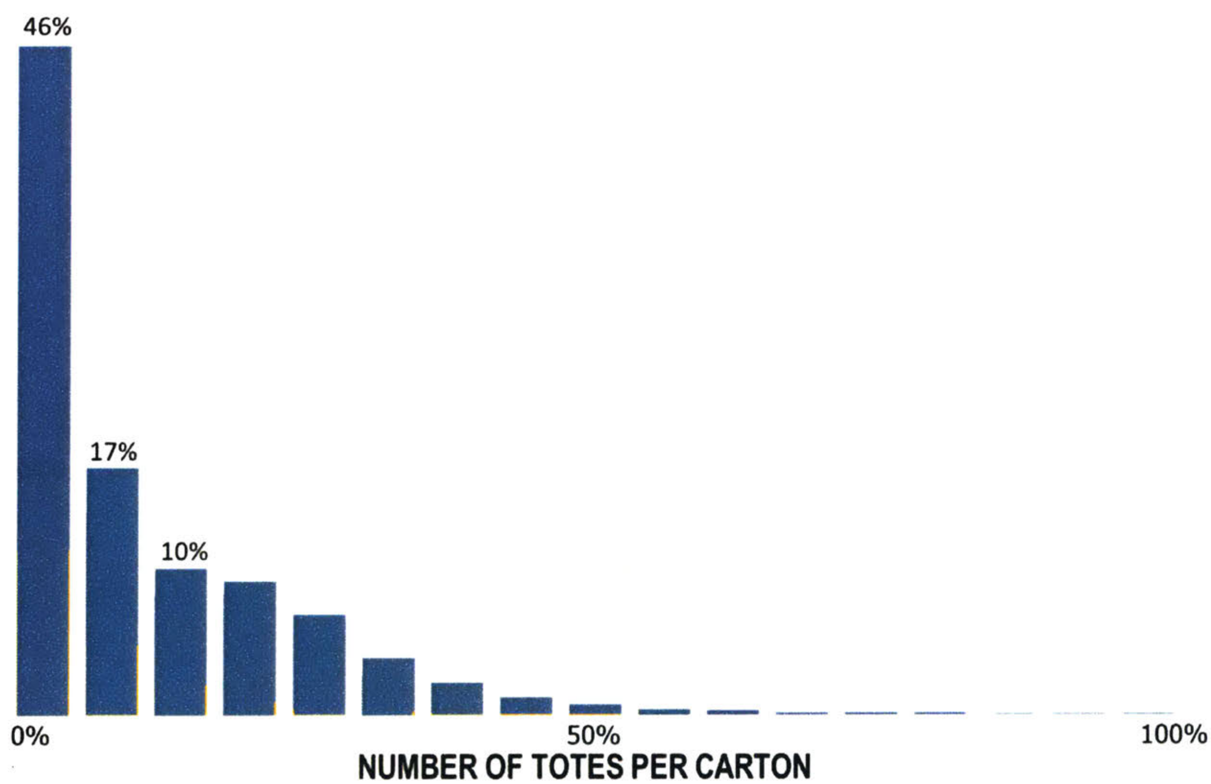


Figure 9: Totes per Carton Histogram

Outbound Cartons per Buffer Lane Batch

The ratio of outbound cartons per buffer lane batch is approximately 5 to 1. The average number of cartons per buffer lane batch is approximately at the 40 percentile in Figure 10. The distribution of cartons per buffer lane batch can be seen in Figure 10 where the 0% on the x-axis corresponds to the minimum number of cartons per buffer lane batch and the 100% corresponds to the maximum number of cartons per buffer lane batch. The 15% of buffer lane batches at the max have the maximum number of cartons because they are made up of all single tote cartons. The implications of a ratio of approximately 5 cartons per 1 buffer lane batch will be discussed in subsequent sections.

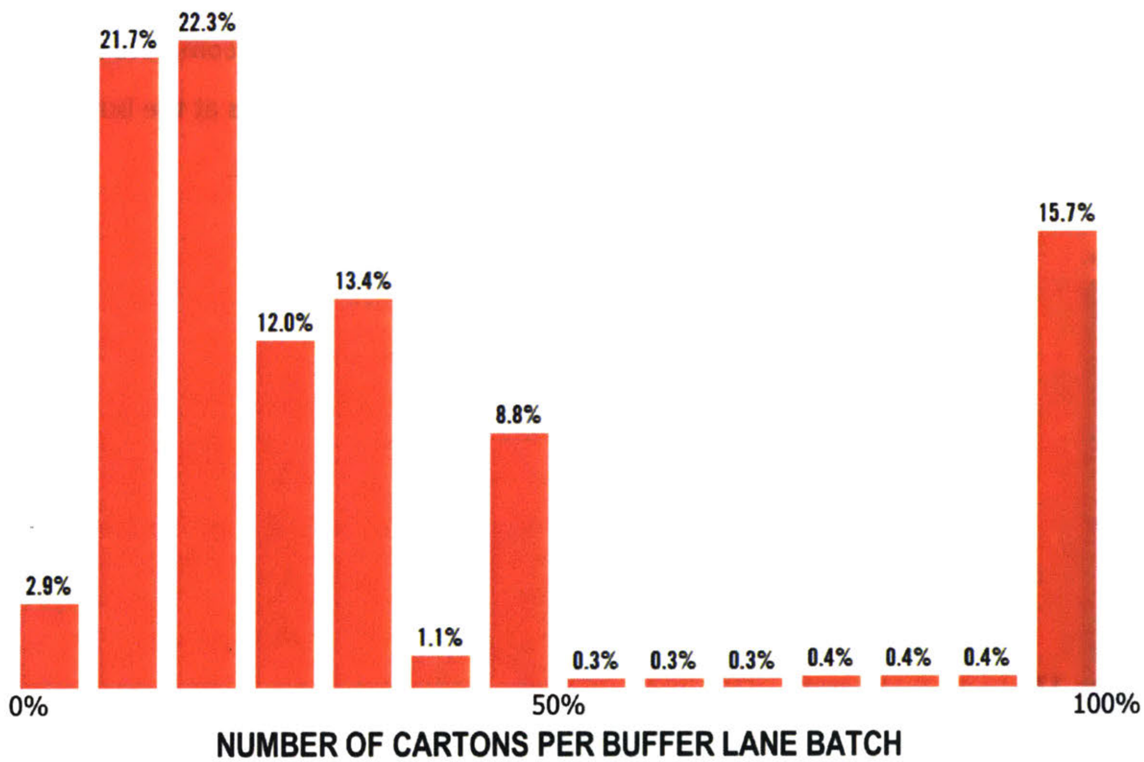


Figure 10: Cartons per Buffer Lane Batch Histogram

4.3.2 Buffer Lane Batch Cycle Times

Buffer lane batch cycle times are critical because the totes move through the system in buffer lane batches, not as totes and not as cartons. Cycle time is defined as the total time from the beginning to the end of your process.

The total pick-buffer-process cycle time of a buffer lane batch within the tote picking flow path is from the time the first tote is picked until the last carton is packed from a buffer lane batch. I had cycle time data from over 28,000 buffer lane batches. For each batch I had observations of three components of the total cycle time, namely the cycle time for picking, the cycle time for batch release, and the cycle time for processing. For each component I then computed the median cycle time, and summed these median cycle times to get a measure that I term to be the "total cycle time." For confidentiality reasons I cannot report the actual cycle times, but can express the median for each component as a percentage of the total cycle time (defined as the sum of the medians). This is shown in Figure 11. The following section contains a more detailed explanation of the calculations for picking, buffer release, and processing represented in Figure 11.

While work is released as waves, multiple waves will be ongoing in each step of the flow path and does not prevent other parts of the wave from moving forward in the system. A buffer lane batch is the largest batch type that has to be complete for the other parts of the batch to get all the way to the processing station chutes. The median time is used to depress the effect that major outliers, typically caused by a quality issue such as a missing tote or missing unit, have on the process times. This gives a more representative view of how work is flowing through the distribution center.

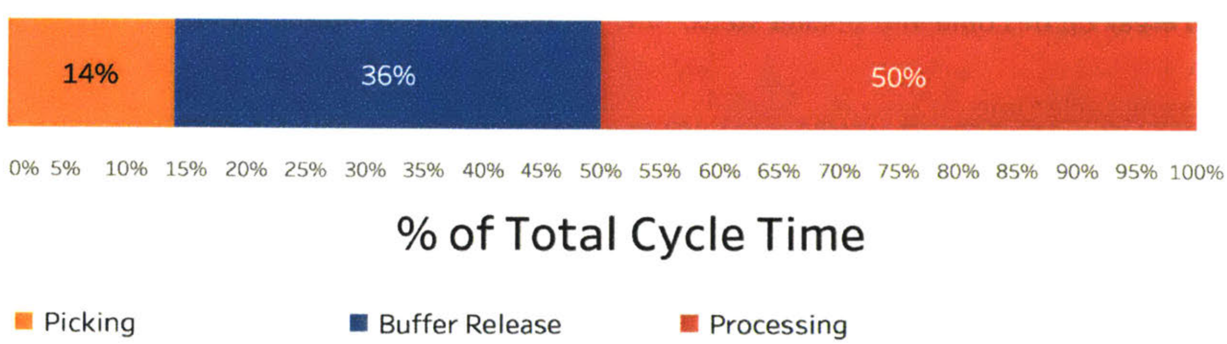


Figure 11: Buffer Lane Batch Cycle Time - Percent of Total Pick-Buffer-Process Cycle Time

Picking Cycle Time

The picking cycle time for a buffer lane batch is calculated from the first tote pick time to the last tote pick time within a buffer lane batch. The median picking cycle time for a buffer lane batch is 14% of the total cycle time as defined in prior section. Picking a buffer lane batch is synonymous with "completing" a buffer lane batch which puts it in a releasable state in the buffer.

Buffer Release Cycle Time

The buffer release cycle time is calculated from when the pick time of the last tote of a buffer lane batch to the pack time of the first carton in that same buffer lane batch. The median of this cycle time is 36% of the total cycle time as indicated in Figure 11. The 36% includes the time while the buffer lane batch complete and is waiting to be released, and also the time after the buffer lane batch is released and the totes travel down to a processing station chute and then are packed by an associate. Unfortunately data of the time the buffer lane batch is released is unavailable given the systems within the distribution center. It is known that there is a significant amount of time between when a buffer lane batch is complete and in a releasable state, and when it is actually released. At any given time, roughly 30% of the buffer lanes are in a releasable state but have not yet been released. Under normal operation, there are always completed buffer lane batches that have not released. Ideally, more granular data would break up this time into smaller steps.

Processing Cycle Time

The processing cycle time is calculated from the first carton pack time to the last carton pack time. Totes are released to unstaffed processing stations. The cause of this will be further discussed in the subsequent implications section. Due to the fact that totes are released to unstaffed processing stations, totes from the same buffer lane batch are not processed according to any methodology like First In First Out (FIFO). the median of the processing cycle time is 50% of the total cycle time.

4.3.3 Flow by Wave

The key takeaway in these cycle times is that it is quicker for a buffer lane batch to be picked and be complete in the buffer. The flow up to the buffer is relatively good. However after a buffer lane batch is complete, it takes more time for that buffer lane batch to be released and for the first carton of the buffer lane batch to be packed, and it takes even more time for the rest of the buffer lane batch to be packed. The flow after the buffer is relatively poor. Figure 12 below is a visual representation of the same key takeaway. The top chart of Figure 12 shows the time at which the last tote is picked for each buffer lane batch within a wave. The completion of a wave in picking is relatively tight. The bottom chart of Figure 12 shows the time at which the last carton is packed for each buffer lane batch within a wave. There is a tremendous amount of dispersion as buffer lane batches wait in the buffer as well as wait at processing station chutes to be packed. It is clear that work after the buffer is not being processing by a First In First Out (FIFO) methodology. The implications will detail why this is happening and why it is a problem that affects the apparel distribution center's ability to develop or unlock additional throughput capacity to meet future requirements.

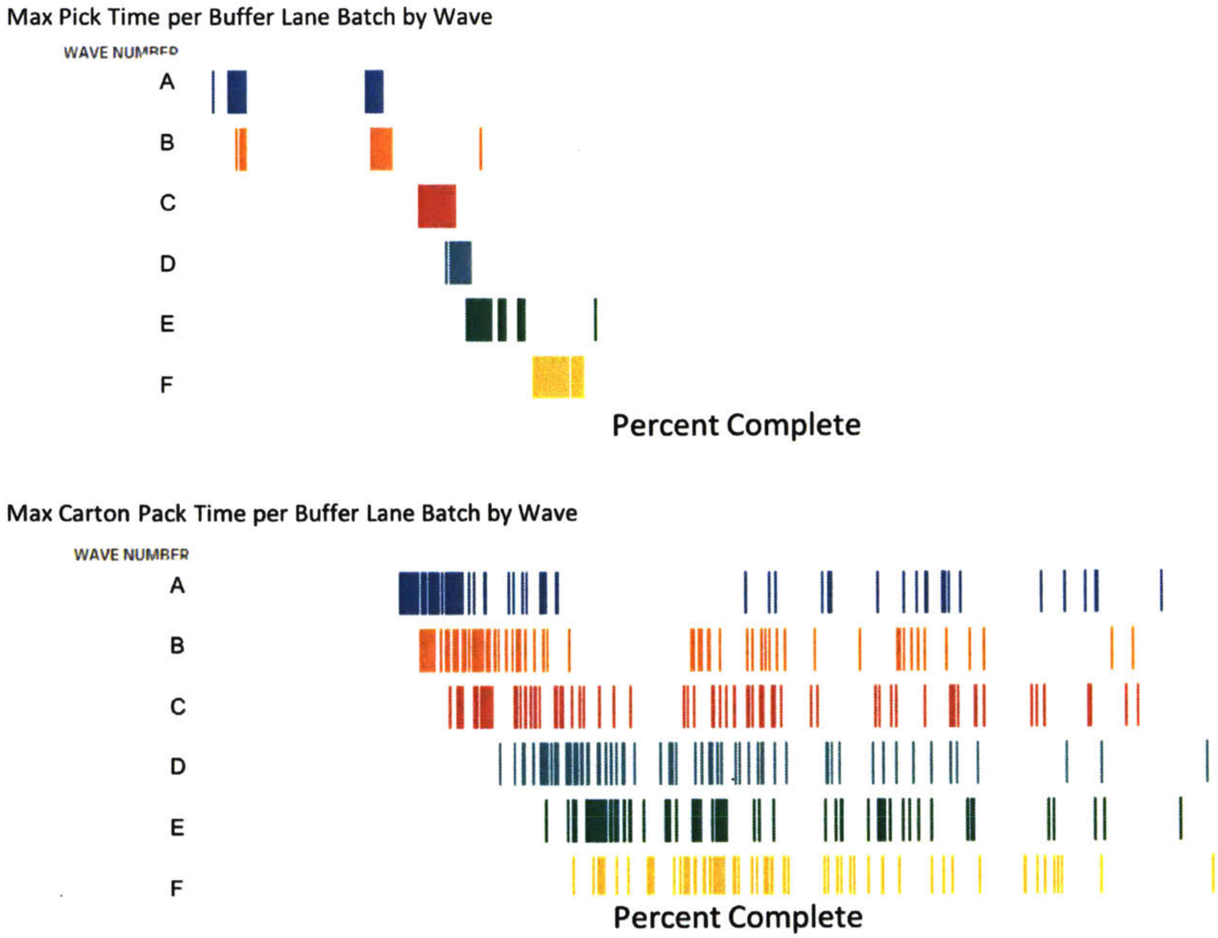


Figure 12: Flow by Wave - Picking and Processing

4.4 Implications

The implications section will first discuss how the buffer release logic is identified as the bottleneck of developing increased throughput capacity to meet future demand. Next there will be a discussion regarding the implications of flow on throughput capacity, with flow being impacted by current batch size and release logic.

4.4.1 Identification of Bottleneck for Long-Term Throughput Capacity

This research and thesis are concerned with the long-term bottleneck, not the current daily local bottlenecks. Nike is a growth company, and one of the key questions is when does the main apparel distribution center run out of capacity and what levers can be pulled to delay that from happening. Identifying the bottleneck that will constrain throughput capacity to meet future demand requires consideration to what each process looks like when fully staffed and identify where the buildup of WIP would occur and which area would be starved for work. Processes upstream from picking have the capability of being scaled up with more staffing. Picking can also be scaled up through staffing. The buffer-sorter-processing area is fixed. It is a monument so to speak and the future bottleneck if the building was running well would be the processing stations staffed at 100%.

In the current state and analysis sections, we learn that the buffer always has completed buffer lane batches ready to release, but the buffer release logic does not allow the work to be released down to the processing stations in such a manner that associates can stay at the same processing station and not run out of work. If associates have to move stations to stay productive, then the processing stations cannot be staffed at 100%. The issue is not the fact that associates moving from station to station takes a lot of time (although the movement is waste and lost capacity), the issue is that if an associate has to move stations it means that another associate cannot be added to the system to utilize the existing infrastructure.

To unlock throughput capacity to meet future demand within the apparel distribution center, the system must allow the existing processing station infrastructure to be utilized. The bottleneck of the system preventing the addition of associates to the processing stations is the ability of the buffer to release totes to the processing stations. This conclusion will be demonstrated by looking at the accumulation of WIP, the time it takes to deliver work vs. the time it takes to complete work at processing stations, the inability to increase staffing levels at processing stations. The buffer release logic along with the current order profile is causing this.

Accumulation of WIP and Staffing

One of the telltale signs of a bottleneck is the accumulation of work in process (WIP). This is a little counter-intuitive at a process like the buffer, where one of the main purposes of the buffer is to accumulate WIP, however WIP is not able to be released at a sufficient enough rate to keep downstream processes busy. The first indication is that associates run out of work if they remain at the same processing stations. Typically, one third of the total buffer lanes are in a complete and releasable state and the complete buffer lane batches are spending over 30% of the total pick to pack cycle time waiting in a releasable state. Despite the near constant availability of work to release in the buffer, through the course of a shift, an associate may move processing stations multiple times. The subsequent section helps explain why associates have to move processing stations.

With regards to staffing, historically at busier times associates would be added into processing increasing the total hours in processing, but productivity, measured by the hours it should take given established standards remained the same time. As a simplified example of what was happening, 5 people and a total of 40 hours produce 10 widgets. Each widget should take the same amount of time based upon standard practice and 10 widgets should take 30 hours to make. The total hours are 40 and the productive hours are 30. If 3 additional people were added for a total 8 people and 64 hours and they produced the same 10 widgets then the total hours is 64 and the productive hours are still 30. This phenomena was happening at the processing stations. Adding people and increasing total hours did not increase the productive hours. This was happening at around 60% of the processing stations being staffed. As associates would run out of work, there were not enough unstaffed processing stations that had totes waiting to be packed and therefore a portion of the associates could not be productive and were idle. At the same time that associates could not find work, there was still work in the buffer that was complete and ready to be released to the processing stations, but it could not be released at a rate high enough to keep associates productive. This is supported both by the performance management system data as well as first hand observation.

Delivery Time and Process Time

For each station, by operating rule, totes for only one outbound carton can be in transit at any one point in time. The transit time to an individual packing station from the buffer to processing stations is 40% longer than the average processing time. The system can release 6 cartons in a standard period of time to a station. The average process rate is 8.5 cartons in the same standard period of time. Given that the input rate of 6 cartons in a standard period of time is less than the process rate of 8.5 cartons, processing stations eventually run out of work.

4.4.2 Impact of Flow on Throughput Capacity

In earlier sections, an opportunity to improve the flow from the buffer onward was identified. There is a lot of unnecessary waiting. The unnecessary waiting or poor flow is due to the order profile along with the buffer release logic. The fact that the average ratio of the number of outbound cartons per buffer lane batch is around 5 causes unnecessary waiting in multiple different ways. Using a buffer lane batch with 5 distinct outbound cartons as an illustrative example, first, there is unnecessary waiting for the buffer lane batch to be complete due to the fact the totes corresponding to 4 of the 5 outbound cartons are held at least until the totes corresponding to the fifth and final outbound carton arrive. Once all the totes have arrived, there is unnecessary waiting for 5 unique processing stations to become available, meaning sufficient space and nothing in transit.

A mathematical approach to see the impact of poor flow and excess waiting on throughput capacity is through Little's Law. Little's Law states the number of customers in a system L is equal to the long-term effective arrival rate λ multiplied by the average time W that a customer spends in the system.

$$L = \lambda W$$

The maximum number of allowable totes that can be stored in the system is fixed as to how many totes can physically fit in the buffer, sorter, and processing station chutes. As changes are made to the system that increase the release of totes from the buffer and improve the overall flow of totes, the time spent in the system (W) is reduced. This in turn reduces the total number of totes (L) in the system. As the totes in the system (L) decrease and there is excess physical storage capacity, the arrival rate from picking can be increased as associates are added into the system up to the point that the total number of totes (L) is within the maximum number of allowable totes that fit in the system.

5 Recommendations and Evaluation

This section initially examines the methodology used to evaluate results. Then recommendations related to batch size and buffer release logic are presented and evaluated.

5.1 Evaluation Methodology

There are two different methodologies that are employed to evaluate recommendations. The first and most commonly used methodology is scenario analysis. Due to the difficulty and cost of experimentation in a production environment, scenario analysis using existing data under different batch size parameters can indicate potential impact both directionally and order of magnitude. As an example of scenario analysis, with the existing tote level data, I could create a scenario with a smaller buffer lane batch size and see how what changes or improvements would have occurred with the smaller buffer lane batch size example. The second evaluation methodology is experimentation in the distribution center. Due to the complexity of the system that is affected by the buffer release logic and batch size, to fully understand the flow improvements, experimentation is the preferred methodology and only way to fully capture the full impact. Unfortunately experimentation is difficult to implement and takes time for changes to take effect. Scenario analysis was the dominate evaluation method and is the evaluation method used in all sections with the exception of the section regarding experimentation. Additional detail is located in Appendix A: Distribution Center Data.

5.2 Batch Size

There are two approaches to adjust the buffer lane batch size. The first is to reduce to a fixed batch size (measured in totes) which is possible with the legacy warehouse management system. The second is to move from a fixed buffer lane batch to a dynamic buffer lane batch where the buffer lane batch size is equal to the number of totes. A dynamic buffer lane batch size would require an upgraded or new warehouse management system.

<u>Reduce buffer lane batch to a fixed batch size</u>

Based on the distribution of the number of totes per carton noted in Figure 9: Totes per Carton and average totes per carton, reducing the buffer lane batch size to be less than the current setting is a simple change. The logic in setting the buffer lane batch size to the current and max setting is for the purpose of having high buffer lane storage utilization, which is unnecessary and in this case counterproductive. High buffer lane utilization is unnecessary because there is a large number of buffer lanes that are never full and therefore do not block the upstream system in picking. There are two primary benefits to reducing the buffer lane batch size. The first benefit is that it will take less time for a smaller batch to complete. The second benefit is that with a smaller batch size, there will be fewer outbound cartons in a buffer lane batch, and thus fewer processing stations need to be available for a buffer lane to release. Returning again to the grocery shopping example. If the shopper group was only 3 shoppers, instead of 5 shoppers. The three shoppers finishing to meet in the buffer will on average be quicker than 5 shoppers. Also, it will be easier for the group of 3 shoppers to be released than the group of 5 shoppers, as the group of 3 shoppers requires 2 fewer cash registers to be available.

With the data collected, we are able to calculate both benefits of a smaller buffer lane batch size. This analysis shows the change from reducing the buffer lane batch size to be approximately 30% fewer totes, but a similar analysis could have been done to reduce the buffer lane batch size to 15% or 50% or even 75% from the current settings. There are risks in reducing the buffer lane batch size to be too small, the key risk being that the effective buffer capacity is the number of total lanes multiplied by the buffer lane batch size; making the effective buffer lane capacity too small increases the likelihood that the lanes fill up and block the upstream picking process. That is, when the lanes are fully occupied, it will be unable to receive totes from picking and also unable to have completed work to release to the downstream processing.

By reducing buffer lane batch size to be 30% smaller than the current setting, the median time for a buffer lane batch to complete in the buffer is reduced by 17%. This estimate used scenario analysis with the existing data. The scenario analysis utilized the existing tote level data to create smaller buffer lane batches. With the smaller buffer lane batches, I was able to measure the cycle times in picking, in the buffer and in processing and hence determine how long it would take for smaller buffer lane batches to be completed. More critical is the impact on the buffer release process by reducing the number of cartons per buffer lane batch. At the original buffer lane batch size, over 50% of all buffer lane batches require 4 or more processing stations to release a buffer lane. By reducing the buffer lane batch size to be 30% smaller, now over 50% require 2 or fewer processing stations for a buffer lane to release. Figure 13 shows the distribution of the cartons per buffer lane batch for a reduced buffer lane batch size of 30% fewer totes per buffer lane batch. The average number of cartons per buffer lane batch reduces from a ratio of over 5 to 4 cartons per 1 buffer lane batch by reducing the number of totes per buffer lane batch size by 30%. This again was calculated through scenario analysis and the smaller buffer lane batches that I was able to generate from the tote level data.

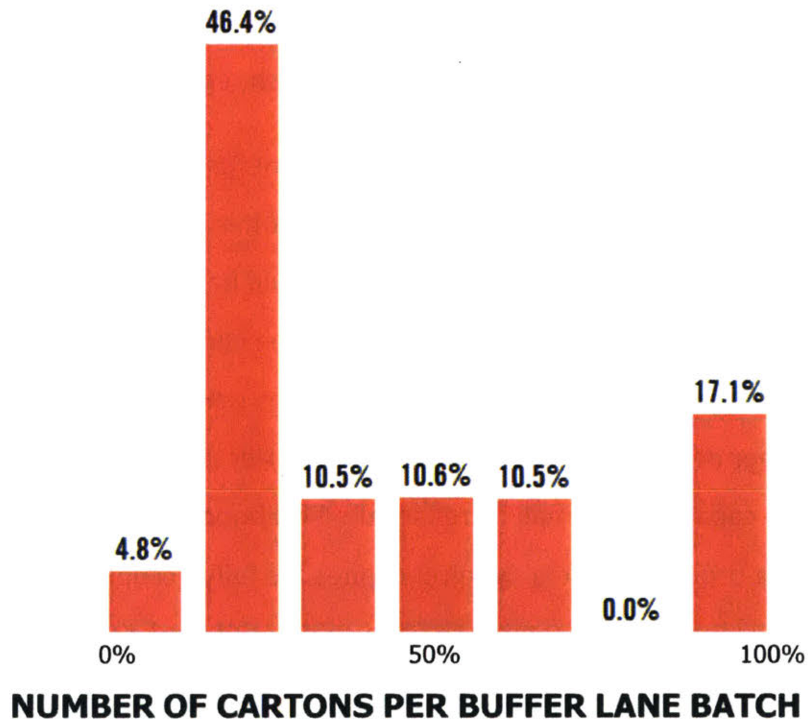


Figure 13: Reduced Buffer Lane Batch - Cartons per Buffer Lane Batch

Reducing the buffer lane batch size will effectively reduce the total number of totes the buffer could hold. If effective buffer lane capacity were reduced, this may require both upstream processes such as picking, and downstream processes such as processing to be more disciplined and maintain wave order integrity. Buffer lanes would turn more often with the smaller size that would likely offset any chance of the buffer running out of space; however this highlights the importance to test changes through actual experimentation in a complex system such as this.

Dynamic buffer lane batch size

Ultimately, the buffer lane batch size should be dynamic such that the buffer lane batch size corresponds to the number of totes for each outbound carton. As an illustrative example, if an outbound carton consisted of 7 totes, then the corresponding buffer lane batch for that batch should be 7 totes, corresponding to a single carton. This approach would allow for 1 outbound carton for each buffer lane batch as opposed to the current average of over 5 cartons per buffer lane batch. The significance is that as soon as all of the totes for each carton arrive, the buffer lane would immediately be able to be released, and only one processing station would need to be available for the actual buffer lane to release. Returning once again to the grocery shopping example, this represents how grocery shopping is done. The goods for each shopper represent an outbound carton, and regardless of the order size (one item or 100 items), once each shopper has gathered the required items, they are able to get in line and are not required to wait for any other shoppers. This dramatically improves flow and reducing time in the system.

Using scenario analysis, when comparing the current state to a dynamic batch size, the median cycle time for a buffer lane to complete and become releasable is reduced by over 85%. This includes all of the 1-tote cartons that are complete as soon as the single tote is picked and arrives in the buffer. If the single tote cartons are removed from the system via a bypass (which is part of a subsequent recommendation), then the median cycle time for a buffer lane batch to be complete is reduced by over 40%.

With dynamic batch size, cartons would be able to be released as soon as they were complete as opposed to the inflexible buffer lane batches which require that they wait for totes corresponding to other outbound cartons. If no other changes were made, but work flowed through as cartons as opposed to as buffer lane batches, the median throughput time would decrease by 43% relative to the current baseline. In reality the improvement could be even greater as the reduced cycle time still includes the time which cartons are waiting in the buffer lane batches and at unstaffed processing stations. Actual data and the scenario analysis method yielded these results.

5.3 Buffer Release Logic

There are multiple recommendations for improvements to the buffer release logic. The potential recommendations are categorized as new warehouse control systems and buffer release logic, or as modifications to existing warehouse control system and buffer release logic. Capital budget, company strategy, and costs to either update or get a new WCS will dictate the tradeoff of what recommendations are most desirable.

New WCS and Buffer Release Logic

1) Release buffer lanes without an assignment to a processing stations. Currently a buffer lane cannot be released until all totes that correspond to each outbound carton are assigned. There is no reason the assignment of a carton to a processing station needs to be made prior to being released and inducted onto the sorter. The assignment could be delayed until the totes are on the sorter which is the last location prior to being diverted and tilted to a processing station chute. The sorter is a closed loop circulating conveyor system with many tilt trays and could hold more totes that were ready for delivery. The buffer lanes would still be needed to absorb the variation of totes arriving from picking. Each tilt tray can hold a tote. A percentage of the processing sorter trays perhaps 20 to 40% can hold totes with no assignment and as a processing station becomes available, then a carton is assigned from the processing sorter trays as opposed to the buffer lanes. This decouples cartons from buffer lane batches, and one carton can be assigned at the same time as opposed to all cartons in the same buffer lane batch. The system would also decrease delivery time and act as a pull system. The delivery time would be from the sorter as opposed to the buffer and would decrease by 60% relative to the current delivery time. A prioritization system of which buffer lanes to release and which induct to use could be determined. This system would allow the number of stations open to match the number of associates which allows full capacity utilization of 100% of the processing stations when required by demand. In addition, matching the number of open stations to the number of associates dramatically improves throughput time and allows a true FIFO system. This is a major change to the existing buffer release logic, but would have the most positive impact on both long-term throughput capacity and throughput time.

Modifications to Existing WCS and Buffer Release Logic

2) Allow totes to ride around the processing sorter two or three times so that multiple cartons from the same buffer lane batch could be sent to the same processing station; this would reduce the number of stations required to release a buffer lane, and the number of totes to each processing station could increase. However, this is a band aid and each trip around the sorter takes a fixed amount of time. It could be a very simple programming change however.

3) Improve release algorithm so that when assignments are made, the system accounts for the capacity of each processing station chute. As an illustrative example, currently a processing station chute with capacity for 10 additional totes could be assigned a carton with 2 totes while at the same time there was a processing station chute with capacity for 4 totes and a carton with 8 totes that now cannot be released. If the algorithm considered processing chute capacity and sent the 8 tote carton to the station chute with capacity for 10 totes and the 2 tote carton to the station chute with capacity for 4 totes, then both cartons could be released simultaneously. The analogy is to fill a jar with big rocks first, then with gravel, and then with sand.

4) Make a station "available" when all the totes assigned to a processing station are on the sorter as opposed to having to wait until all the totes have been diverted/tilted to the station as is currently required. Further analysis and modification to the algorithm may be required to keep totes from the second carton from passing and mixing in with totes from the first carton. This may be a relatively simple way to shave 40% off the delivery time, which would be significant.

5.4 Processing Station Management and Monitoring

The analysis and recommendations require an increased management and monitoring of processing stations. The current condition of the buffer – sorter – processing system is for operations managers or even area managers to open or close a certain number of stations based upon preference and past experience, with the default to have the open number of stations far exceed the number of associates. This approach currently is required due to the poor flow and uses the unstaffed stations as an extra buffer of work so that overall the processing stations will not run out of work at typical operating levels. This leads to two significant negative systemic impacts. The first, as has been previously discussed, the fact that totes are released to unstaffed processing stations receiving work, cartons from the same buffer lane batch are waiting to be packed accounts for 50% of the total average buffer lane batch cycle time. There can be totes corresponding to multiple outbound cartons sitting in the same processing station chute, especially if that processing station is unstaffed but has space. This occurs because no FIFO-like system is followed. The second negative impact is that a large number of unstaffed processing stations leads to associates having to move around often as opposed to the work being forced to fewer staffed stations. Again, this is in reaction to the poor flow of work being released from the buffer

To get the true and expected benefit from any changes to either the buffer lane batch size or buffer release logic, the number of stations needs to be restricted and closely monitored. Initially until the system is fixed, there will still need to be more processing stations open to receiving totes than the number of associates working. However if 35% of the stations are staffed by associates as compared to 60% of the processing stations being staffed, the number of processing stations open to receiving work should not be 100% in each case.

Having a standardized procedure for the number of processing stations will improve the value of data collected and allow for more accurate experimentation and system understanding. Below in Figure 14 is an example of how to initially standardize the percent of stations open depending on the percent of stations staffed (number of associates working) using a fixed ratio of processing stations per associate. This is a simple and implementable approach. As the flow of work in the distribution center improves, the ratio could be reduced. Ultimately, if sufficient improvements to the system are implemented that work can flow to a staffed processing station and not run out of work at the processing station, the target ratio of 1 processing station per associate could be utilized. The target ratio of 1 processing station per associate is what will enable the maximum number of associates working in processing which will unlock the maximum number of throughput capacity to meet future demand as Nike continues to grow.

% of Stations Staffed	% of Stations Open
<30%	47%
35%	52%
38%	57%
42%	63%
45%	68%
49%	73%
52%	78%
56%	84%
59%	89%
63%	94%
66%	99%
>70%	100%

Figure 14: Example of Processing Station Management

5.5 Experimentation

Despite the difficulty of making system changes and experimenting within a production environment, the key stakeholders agreed an experiment should be designed and carried out based upon the newly available data and corresponding understanding of the system and implications of the current mix of work with the legacy batch size.

An experiment reducing the buffer lane batch size from the original setting to a setting approximately 15% smaller was carried out and preliminary results analyzed. The intention was to reduce buffer lane batch size several times in small increments, ensuring no unforeseen issues or impact to production. Unfortunately due to production requirements and unknown warehouse control system (WCS) issues, the experiment was cut short and it was not permitted to run long enough for all of the work in the distribution center to transition over to the reduced batch size.

Baseline data were collected prior to beginning the experiment. Efforts were made to find data from comparable weeks. Unusual weeks that included holidays, quarterly load-ins, etc. were excluded from the analysis. No day or week is identical, so when appropriate the data were normalized. An example of this is that weekly and even daily basis, the number of associates in processing fluctuates. The number of associates will impact the rate at which buffer lanes are released because the buffer releases buffer lane batches in response to space becoming available at processing stations. Key metrics were also identified and an anticipated response for each key metric was hypothesized. Two of the key metrics identified were processing associate performance and average buffer lane releases per hour per associate headcount. Processing associate performance is the target time (for the actual mix of work) divided by the actual time to process a set amount of work and is portrayed as a percent. Note, values have been disguised in the figures below to protect confidential information. Buffer lane releases per hour per associate headcount was calculated by extracting how often a buffer lane releases and averages it across a shift for all the buffer lanes and then normalizing it for the total hours and the headcount in processing. Headcount and hours are expected to be tightly correlated in processing, but throughout any given day, it is possible for associates to be moved in or out of processing. Both metrics were expected to increase during the experiment to reduce the buffer lane batch size. Performance was expected to increase because better flow would allow associates to stay at the same processing station and the work would come to the associates as opposed to the associates having to seek out the work. Buffer lane releases were expected to increase because smaller buffer lane batches would reach completion quicker and then be released more often because they required fewer unique processing stations to be available.

Despite the short duration of the experiment, positive results were observed as expected. Associate performance or speed improved as can be seen in Figure 15, the release of buffer lanes, normalized on a labor headcount basis, increased in Figure 16 (best indicator of flow is how often buffer lanes are released). Additional experimentation would be required to have greater certainty that these changes were significant. Unfortunately unanticipated WCS issues precluded the opportunity to run additional experiments. The average as well as +/- 3 standard deviations was plotted for both the experiment and the baseline to gauge the statistical significance of any change.

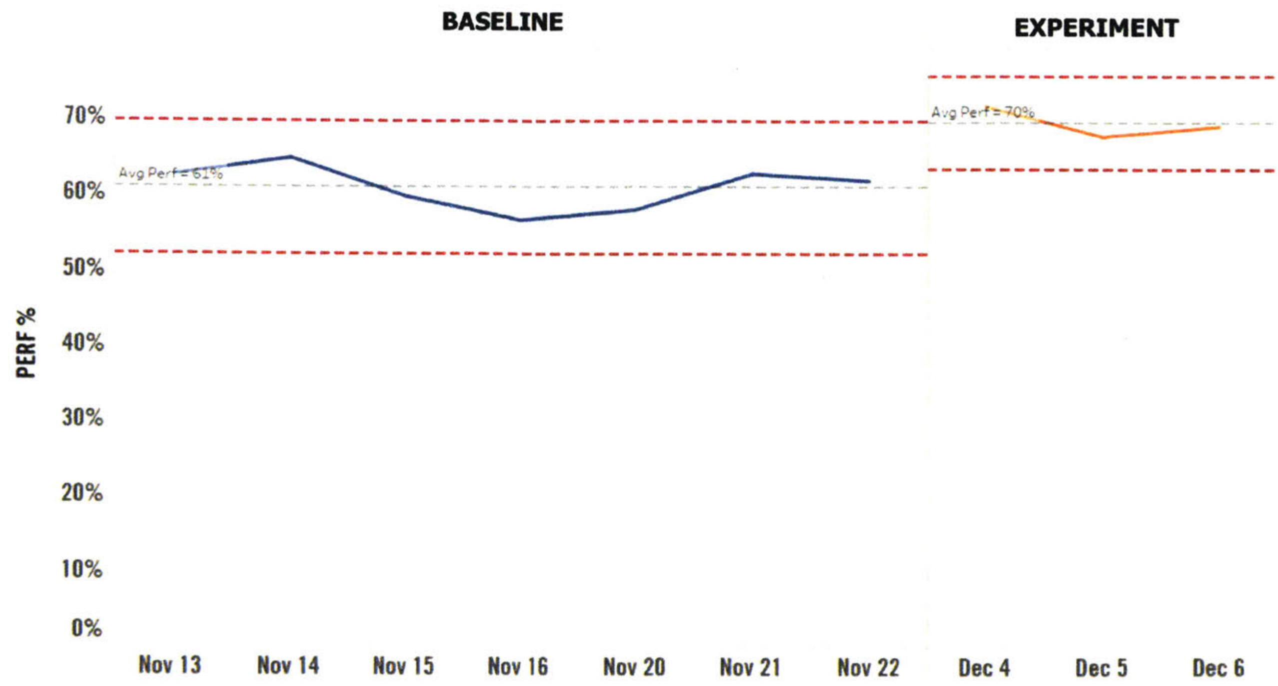


Figure 15: Associate Performance - Batch Size Experiment

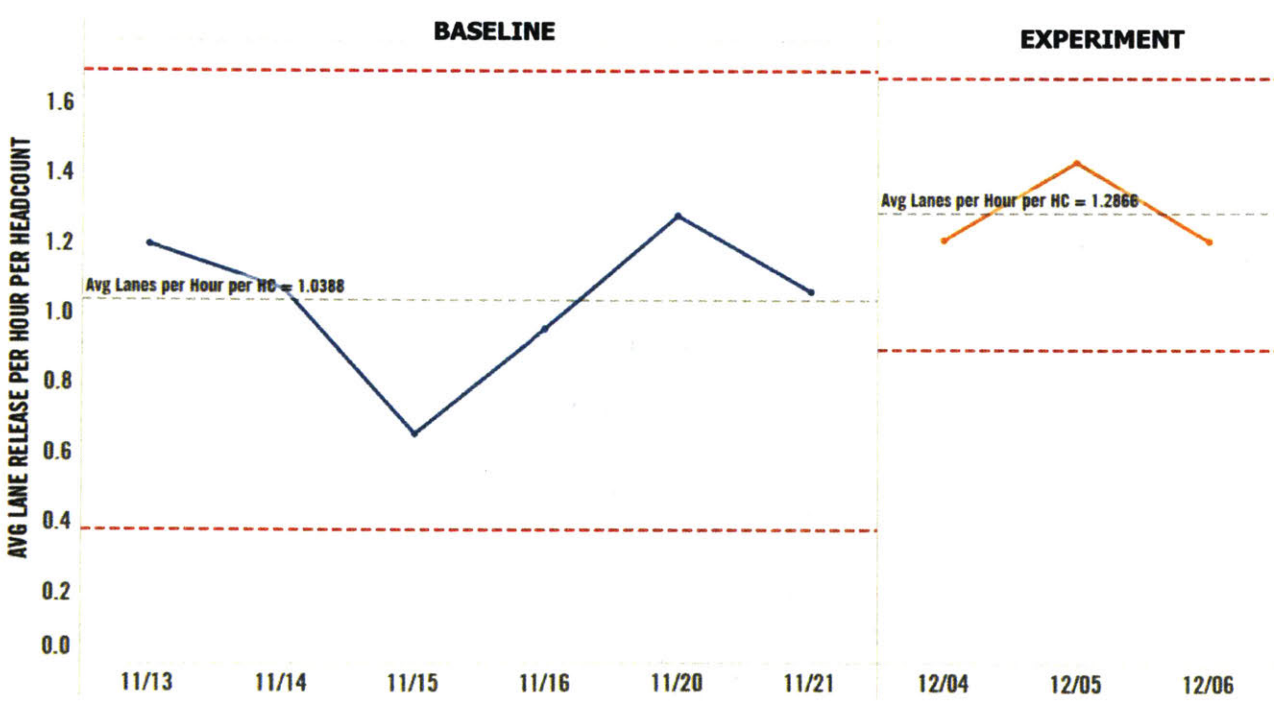


Figure 16: Buffer Lane Releases - Batch Size Experiment

6 Conclusion

The system is not optimally aligned to the current demand placed on it and there is opportunity for improvement. While there are several paths forward, the status quo is not sustainable as Nike's main apparel distribution center experiences both growth and increasing pressure to deliver smaller orders to both wholesale and digital customers. While the logic and batch size may have functioned sufficiently in the past as strictly a wholesale building servicing large customer orders, that is no longer reality. When such a disruptive change such as e-commerce has affected the business and changed the order profile, it becomes critical to examine the current system and implications of the change and collect sufficient data to properly analyze and develop recommendations. It is critical to understand that the actual work being performed is still the same and therefore the required infrastructure need not change. Any changes to the system would have to be validated through additional analysis and testing to prove that the recommendations below yield the positive expected results with regards to throughput and cycle time. This is a complex system and testing the buffer release under various conditions is necessary for a more complete understanding.

The recommendations discussed in depth in section 4 can be summarized as the need to:

1) Reduce buffer lane batch size to simplify both the completion and release of buffer lane batches
2) Modify or reinvent the buffer lane release logic to allow completed buffer lane batches the opportunity to flow to processing stations such that fewer unstaffed processing stations are required to be used as an extra buffer

Both recommendations should positively affect both facility throughput rate (increase) and cycle time to complete an order (reduce):

1) Reducing buffer lane batch size increases throughput capacity. Smaller buffer lane batch size allows for a simpler release of buffer lane batches from the buffer. It reduces cycle time by requiring fewer totes to be in the buffer plane prior to becoming releasable.

2) Modifying the buffer lane release logic will increase throughput capacity by allowing buffer lane batches to be released more frequently, which leads to improved flow, and allows the capability of placing more associates in processing which in turn unlocks capacity. The cycle time is reduced by reducing the time a completed buffer lane batch is waiting to be released and the improved flow allows totes to be sent to a higher percentage of staffed processing stations than unstaffed processing stations than the current state. This leads to less time waiting in unstaffed processing chutes and leads to a system that adheres more towards FIFO, reducing cycle time.

The key enablers of this thesis were

- The collection of data to evaluate progress at key steps in the process
- A holistic understanding of how the system functions, as well as the implications on the long-term throughput capacity

6.1 Impact of Improved Flow on Throughput Capacity

The impact of improved flow on long-term throughput capacity is the ability to staff 100% of processing stations, irrespective of the order size (totes per outbound carton). This unlocks over 40% of the current effective capacity. Long-term throughput capacity is critical for the apparel distribution center as Nike expects the recent growth trajectory to continue. The processing stations should be the bottleneck, because processing stations can be flexed up or down through headcount as opposed to fixed monuments like the buffer, that currently cannot flex up past a certain level, which is both less than the infrastructure in the processing stations allows as well projected demand according to Nike's growth projections. Understanding how to maximize the output of current assets is important for a company like Nike in a complex system with changing business demands.

6.2 Impact of Improved Flow on Throughput Time

The impact of improved flow on throughput time is significant as it allows service sensitive orders such as digital orders to move through the distribution center at a higher rate. The reduction in throughput time is due to the fact that smaller orders can be decoupled and allowed to complete each step of the process and move onto the next process as opposed to the start and stop flow that currently exists. Another key benefit of improved flow is the ability to close off unstaffed processing stations that are currently required to remain open as an extra buffer of work to offset the less efficient flow of work. A first in first out system will allow more systematic control of which work flows through the distribution center in which order and will compress the time it takes for waves of work to move through the system. Another major impact to throughput time could be the design and development of a single tote carton bypass. Outbound cartons that consist of single totes are complete after picking and there is no need to have them sit in the buffer or at unstaffed stations. If routed directly from picking to a processing station to be packed, up to 80% of the throughput time could be reduced simply by cutting out the unnecessary waiting.

6.3 Future Work

Additional opportunities exist in the apparel distribution center based upon the work and research comprised in this thesis.

1) Make in process data more accessible to identify how long work has been in each step of the process
2) Develop and implement standardization as to how employees move from one processing station to another until the need for associate movement is eliminated
3) Reduce the buffer lane batch size based upon previous analysis and recommendation on an experimental basis for long periods of time to identify optimal batch size range and monitor for additional effects of reducing batch size
4) Evaluate the cost and time of implementation for various buffer release logic changes
5) Evaluate cost and time of implementation of a single tote carton bypass

6.4 Other Applications

Other applications where this methodology to evaluate and improve flow to unlock throughput capacity are manufacturing and distribution processes where work is batched through systems and release logic is embedded both in the release of large batches as well as work moving between processes. Especially older systems that were designed for a legacy type of work that has had any type of disruption to order patterns (frequency, order size, etc.) are excellent candidates for this type of investigation and analysis.

A specific example is a small metals manufacturer that releases batches of work to travel through rolling, annealing, and slitting processes. By tracking time spent waiting for each process as well as for setups will start to highlight the inefficiency of large batch sizes. This leads to large build-up of WIP and a hurry up and wait approach and inability to control the order in which work is completed.

Each system will have its unique intricacies and design, however measuring flow and reducing the waste of waiting, particularly batches that are connected unnecessarily will improve flow and allow more throughput capacity at constant arrival rates by reducing the time spent in the system.

This page has been intentionally left blank

7 References

[1] "Select Financials - NIKE FY2017 Annual Report." [Online]. Available: http://s1.q4cdn.com/806093406/files/doc_financials/2017/ar/select_financials.html. [Accessed: 02-Mar-2018].

[2] "Nike Internal / 2017 NIKE, Inc. Investor Meeting." [Online]. Available: https://nike.brand.live/c/2017-investor-meeting?wmode=transparent&ub=00a4e4&lc=00a4e4&oc=ffffff&uc=ffffff&autoplay=true. [Accessed: 02-Mar-2018].

[3] "Nike Opens Its Largest Distribution Center Worldwide In Tennessee," *Nike News*. [Online]. Available: https://news.nike.com/news/nike-opens-its-largest-distribution-center-worldwide-in-tennessee. [Accessed: 29-Mar-2018].

[4] J. J. Bartholdi III and S. T. Hackman, *WAREHOUSE & DISTRIBUTION SCIENCE*, Release 0.98. 2017.

[5] H. Coleman, "Doing Distribution Center Design Right," *Electr. Wholes.*, vol. 96, no. 12, pp. 32–24, Dec. 2015.

[6] M. Weil, "Warehouse Controls: From Bridge to Balancing Act," *Mater. Handl. Logist.*, vol. 70, no. 4, pp. 13–14, Apr. 2015.

[7] R. Michel, "2017 Warehouse/DC Operations Survey: In the thick of e-commerce adjustments," *Supply Chain Manag. Rev.*, vol. 21, no. 6, pp. S52–S58, Nov. 2017.

[8] B. Trebilcock, "Supply Chain Basics for Materials Handlers," *Mod. Mater. Handl.*, vol. 72, no. 3, pp. 90–93, Mar. 2017.

[9] R. Conway, W. Maxwell, J. O. McClain, and L. J. Thomas, "The Role of Work-in-Progress Inventory in Serial Production Lines," *Oper. Res.*, vol. 36, no. 2, p. 229, Apr. 1988.

[10] S. J. (Sheila J. Bragg, "Analysis of sorting techniques in customer fulfillment centers," Thesis, Massachusetts Institute of Technology, 2003.

[11] M. Aichlmayr, "Sorting's Need for Speed," *Transp. Distrib.*, vol. 42, no. 8, p. 50, Aug. 2001.

Appendix

Appendix A: Distribution Center Data

Data Preparation

For both the status quo and analysis sections, snapshot data was utilized. The snapshot data populates an SQL table capturing real-time snapshots of the buffer-sorter-processing system at a regular intervals. The data analysis Alteryx was used to prepare the data.

The procedure to utilize the day is as follows:

- Filter for date range
- Ensure the key fields are not repeated, if so then create unique fields by joining fields such as "Tote + Buffer Lane Batch"
- Remove duplicates for records of picks showing up in multiple snapshots keeping only record of the previously identified record
- Determine which data can be considered complete,
- Summarize by buffer lane batch, carton, wave, picking assignment, etc. to create distributions and insights
- Join with archived pack history table
- With tote pick times and carton pack times, calculate data around buffer lane batch completion (first tote picked to last tote picked), buffer lane batch release to pack (last tote picked to first carton packed for the same buffer lane batch),
- There are numerous ways to groups the data and potential questions that can be answered with this data analyzed differently