

**Multi Echelon Supply Chain Design for Amazon Private Brands**

by  
**Shouvik Das**

B Tech, Electrical Engineering and M Tech, Instrumentation Engineering  
Indian Institute of Technology, Kharagpur, 2011

Submitted to the MIT Sloan School of Management and the MIT Department of Civil and Environmental  
Engineering in partial fulfillment of the requirements for the degrees of

**MASTER OF BUSINESS ADMINISTRATION  
AND  
MASTER OF SCIENCE IN CIVIL AND ENVIRONMENTAL ENGINEERING**

in conjunction with the  
**LEADERS FOR GLOBAL OPERATIONS PROGRAM**

at the

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY**  
May 2020

© 2020 Shouvik Das. All rights reserved

The author hereby grants MIT permission to reproduce and to distribute publicly paper and electronic  
copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author.....  
MIT Sloan School of Management  
MIT Department of Civil and Environmental Engineering  
May 8, 2020

Certified by.....  
Dr. Stephen C. Graves, Thesis Supervisor  
Abraham J. Siegel Professor of Management Science, MIT Sloan School of Management

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

Accepted by.....  
Colette L. Heald, Professor of Civil and Environmental Engineering  
Chair, Graduate Program Committee

Accepted by.....  
Maura Herson, Assistant Dean, MBA Program  
MIT Sloan School of Management

THIS PAGE IS INTENTIONALLY LEFT BLANK

# **Multi Echelon Supply Chain Design for Amazon Private Brands**

by

Shouvik Das

B Tech, Electrical Engineering and M Tech, Instrumentation Engineering  
Indian Institute of Technology, Kharagpur, 2011

Submitted to the MIT Sloan School of Management and the MIT Department of Civil and Environmental Engineering on May 8, 2020, in partial fulfillment of the requirements for the degrees of

Master of Business Administration

and

Master of Science in Civil and Environmental Engineering

## **Abstract**

Retailers across the globe continue to grow their private label portfolio to offer customers an alternative to existing brands. Typically, retailers source private label products directly from manufacturers to remove middlemen from the value chain, thereby capturing greater value and subsequently passing it on to customers. Combined with the growth of e-commerce as the primary method for consumers to shop for products, expanding private label portfolio has made e-retailers to re-think their supply chain.

Amazon began its journey in Private-Label Brands (PB) in 2009 with the launch of Amazon Basics. Since then, it has expanded its presence across multiple categories. The majority of these products are imported from Asia-Pacific region (APAC) and require sourcing larger quantities to account for long-lead time between production runs and high variability in demand to maintain competitive costs. These factors result in PB inventory dwelling for a long period at the Amazon Robotic Fulfillment Centers (FCs), reducing the turns-ratio of expensive storage bins there, which could otherwise be utilized for storing high-velocity products.

The growth of PB products raises the need to build more storage space, which is expensive in highly automated robotic FCs. Additionally, since fixed storage cost is proportional to the space occupied in FCs, high 'dwell time' translates to high storage cost. To increase utilization of FC storage bins, the Inbound Supply Chain Team plans to build a low-cost upstream storage (LCS1) to supply the FCs and store excess PB inventory there. Alternatively, Amazon can also use its third party storage center in APAC, another low-cost storage node (LCS2), after sourcing PB products from manufacturers in Asia before shipping to regional markets in US, EU, Japan etc. This could provide an opportunity for inventory savings from risk pooling by optimizing inventory storage across various nodes in the supply chain.

Using multi-echelon inventory optimization techniques, this thesis explores the tradeoffs between using low-cost storage node close to end customers in the US (LCS1) versus that close to manufacturing source in APAC (LCS2). The objective of the thesis is to find the optimal inventory placement strategy across three storage points – FCs, LCS1 in US, and LCS2 in APAC - to achieve the best-in-class customer experience (InStock availability) at minimal inventory storage cost.

**Thesis Supervisor:** Dr. Stephen Graves

**Title:** Abraham J. Siegel Professor of Management

**Thesis Supervisor:** Dr. David Simchi-Levi

**Title:** Professor, Department of Civil and Environmental Engineering

THIS PAGE IS INTENTIONALLY LEFT BLANK

## Acknowledgements

I would like to thank those who contributed significantly towards making this thesis a reality.

I am grateful to my advisors, Dr. Stephen Graves and Dr. David Simchi-Levi. I thank Dr. Graves for his constant feedback throughout the project, from formulation of the problem to development of the solution. I acknowledge Dr. David Simchi-Levi for his timely inputs on the project.

I am extremely thankful to my manager at Amazon, Siddhartha Trivedy, who always provided guidance, support, and resources throughout the project. I would like to appreciate Amazon's entire Supply Chain Innovation team that backed me during my internship. I recognize Sameer Devji's valuable assistance throughout the project; I am forever indebted to him for his constant support.

Additionally, I would like to recognize the contribution of following Amazonians from different business functions, whose thoughtful suggestions guided my project through various stages.

Salal Humair  
Ali Sadighian  
Alvaro Maggiar  
Muhong Zhang  
Maro Enoka  
Chris Lawson  
Harsha Murthy

Bryan Lunak  
Michael Winters  
Jake Jorgensen  
Ted Seo  
Hillary Zak  
Ameya Margaj  
Taylor Hearn

Brent Yoder  
Joe Case  
Kolli Black  
Avi Nagavalli  
Shamiron Thomas  
Mariel Lisud  
Jason Cheng

A special gratitude to Carl Berger for championing my project and for providing mentorship. His leadership ensured a highly conducive work environment that helped me thrive in the organization. I also thank Amazon for its continued support to the Leaders for Global Operations program at MIT. This dream experience for me would not have been possible without this long-standing partnership.

I would like to acknowledge MIT's Leaders for Global Operations program for providing me the opportunity to carry out this research work. I thank my classmates, my family away from home, who have made the past two years a once-in-a-lifetime experience for me.

Finally, I would like to thank my family for their unconditional love, support, and patience that made my dream of studying at MIT a reality. I thank my parents, Basabi Das and Susanta Kumar Das, for instilling in me a passion for learning and the value of perseverance. I thank my wife, Surabhi, for being a pillar of support along the way.

THIS PAGE IS INTENTIONALLY LEFT BLANK

## Table of Contents

	Abstract	3
	Acknowledgements	5
	List of Figures	9
	List of Tables	10
	Commentary on Amazon.com Proprietary Information	11
1	Introduction	12
1.1	Amazon.com Overview	12
1.2	Project Motivation	14
1.3	Problem Statement	16
1.4	Project Hypothesis	16
1.5	Project Methodology	17
2	Background	18
2.1	Emergence of private-label brands in Retail	18
2.2	Supply chain challenges in private-label products business	19
2.3	Amazon Supply Chain Network	20
2.4	Supply Chain Planning at Amazon	21
2.5	Amazon moved to 1-day delivery for its Prime members	22
2.6	Pilot study in Amazon, US for low-cost storage node	22
3	Problem Formulation	23
3.1	Identification of improvement areas	23
3.2	Selection of a subset of private-label portfolio for detailed study	23
3.3	Formulation of the optimization problem	25
4	Literature Review	26
4.1	Multi-Echelon Inventory Optimization (MEIO)	26
4.2	Guaranteed Service Model (GSM)	27
4.3	Strategic Inventory Placement Model (SIP)	30
5	Research Methodology	32
5.1	Metrics to measure the success of the study	32
5.2	Building the private-brand supply chain model	33
5.2.1	Current PB supply chain	33
5.2.2	Network modeling to meet 1-day delivery promise	33
5.2.3	Network modeling to reduce inventory storage cost	34
5.3	Rationale behind using GSM to model private-label supply chain	36
5.4	Model Setup	38
5.4.1	Selection of location for LCS1 (in US) and LCS2 (in APAC)	38
5.4.2	Building SIP model using LCS1 and LCS2	39
5.4.3	Data sources for estimating the model parameters	41
5.4.4	Parameters of a <i>Typical</i> SKU for comparing the scenarios in SIP model	43
5.4.5	Working formulae used in the SIP model	44

6	Results and Discussion	45
6.1	Optimal Inventory Storage Strategy	45
6.1.1	Effect on the target network inventory	47
6.1.2	Effect on the per-unit cost of inventory storage	48
6.2	Comparison between the current FC-only storage and FC + LCS1 storage	48
6.3	Sensitivity Analysis	49
6.4	Limitations of the model	50
7	Recommendations and Future Work	52
7.1	Final recommendation	52
7.2	Implementation Roadmap	52
	References	54
	Appendices	55

## List of Figures

1	Amazon's financial performance over last 20 years	14
2	Amazon's employee base and revenue-per-employee over last 20 years	14
3	Maximum selling product categories on Amazon in 2019	15
4	Growth of Amazon Prime accounts in the US and difference in average spending between Prime accounts and Non-prime accounts	16
5	Annual sales growth of Amazon private label categories in the US in 2018	16
6	Private-label as %share of category-sales in the world (left) and in the US (right)	19
7	Share of US consumers (by generation) willing to buy Private Labels	20
8	Traditional costing mechanisms distort product and customer costs	21
9	Current Supply Chain for Amazon private-label brands	22
10	Breakup of imported Amazon Private-label by Business line (left) and by product-size (right)	25
11	Breakup of imported SKUs of BL1 by demand characteristics	25
12	Different network topologies for Multi-Echelon Supply Chain systems	27
13	Schematic to formulate Guaranteed Service Model	29
14	Schematic of current Amazon private-label supply chain for imported products	34
15	Representative schematic of private-label supply chain to meet 1-day delivery promise	35
16	Representative schematic showing LT benefits of introducing a LCS node on target inventory	36
17	Representative schematic of private-label supply chain after introducing low cost storage nodes (LCS1 in the US and LCS2 in APAC)	37
18	Distribution of manufacturing lead time for private-label products over 20 months	38
19	Distribution of private-label SKUs' demand (left) and service level (right)	39
20	Maximum private-label receipt in US (left) and maximum dispatch in APAC (right)	40
21	Normalized supply chain costs, inventory storage costs, and opportunity cost of capital aggregated over 763 SKUs used in the study	46
22	Comparative performance of four storage scenarios (in SIP model) on Typical PB SKU	47
23	Breakdown of storage costs contributed by safety stock and cycle stock at different nodes across four scenarios	48
24	Key drivers for reduction in network inventory and storage cost in FC + LCS1 storage	49
25	Key drivers for reduction in inventory storage costs in FC + LCS1 storage strategy	50
26	Sensitivity analysis of the inventory storage cost with SKU size	50
27	Sensitivity analysis of the best solution as holding cost rates in low cost nodes change	51

## List of Tables

1	Comparison between two approaches – SSM and GSM – to Multi Echelon system design	28
2	Comparison of LCS1 and LCS2 storage nodes	37
3	List of data required as input parameters for the SIP model	42
4	Parameters of a Typical private-label SKU received from the Category team	44
5	Normalized inventory storage details for <i>Typical</i> SKU in four scenarios	47

## **Commentary on Amazon.com Proprietary Information**

In order to protect information that is proprietary to Amazon.com, 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 competitive information, while still conveying the findings of this project.

## 1. Introduction

This thesis seeks to highlight the positive business impacts of designing and implementing multi-echelon inventory planning in online retailing supply chains. Today's consumers increasingly demand low cost, better quality, and greater convenience (faster delivery, easy payment etc.). This study is based on a six-month research work carried out at Amazon, Seattle as part of MIT's Leaders for Global Operations (MIT LGO) curriculum. This chapter focuses on presenting an overview of the problem statement and the motivation to address it. In addition, the author also presents the research methodology used in the study.

### 1.1 Amazon.com Overview

The Amazon story started in 1994, when a 30-year old Vice-President at D.E. Shaw & Co., Jeff Bezos was given the charge of exploring opportunities for commercialization of upcoming internet technology. Brainstorming with D. E. Shaw, and others at the company for over months, Jeff came up with three nascent ideas – 1) e-mail services (Gmail, Yahoo Mail of today), 2) online financial services (E-Trade of today), and 3) *the everything store* (E-commerce giants Amazon, Alibaba of today). [1]. The key idea behind e-commerce was disintermediation of a product's supply chain between manufacturers and customers. Diving deep into this idea, Jeff learnt from *Matrix News*, a technology newsletter then, that data sharing over internet grew by roughly a factor of 2300 in 1993 alone. This unusual growth of internet spurred Jeff's imagination and strengthened his conviction to give up his high-paying hedge fund job and found Amazon.com in 1994 with the initial vision of building "Earth's largest bookstore". [2]

Though Jeff chose books as the first category to test his internet-based retailing idea, since Amazon's early days, Jeff envisioned selling all possible products on internet someday. His ultimate goal was that a customer should *click* on a product catalog on Amazon.com, and the product should mysteriously show up at the customer's doorstep in a few hours. Contrasting the grand vision, the initial days of Amazon were filled with challenges, the most difficult of which was convincing customers, investors, and employees that the internet-business would work. Amazon had to prove that e-commerce was attractive, secure, and easy for the first-time online buyers. With a firm belief that long-term focus on 'superior customer experience' will be a key differentiator, Jeff came up with Amazon's *flywheel* model, built on three pillars - 1) low cost, 2) unlimited selection, and 3) convenience and reliability of delivery. [3]

Looking back after 25 years from its launch in 1995, the world agrees that Amazon has had an immensely successful run, completing 2019 with a revenue of USD 280.52 billion with an annual growth of 20%. [4] [5]. Despite taking nine years to register a positive net income for the first time in 2003, Amazon has enjoyed exponential compound annual growth rate of 53% from 1996 till 2019. (Figure 1). Its employee

base, characterized by some of the best talents across the world, has grown to 798,000 in 2019. (Figure 2) [6]. Indeed, Amazon is the second-most valuable company in the world after Apple and ahead of Alphabet (Google’s parent company). [7]

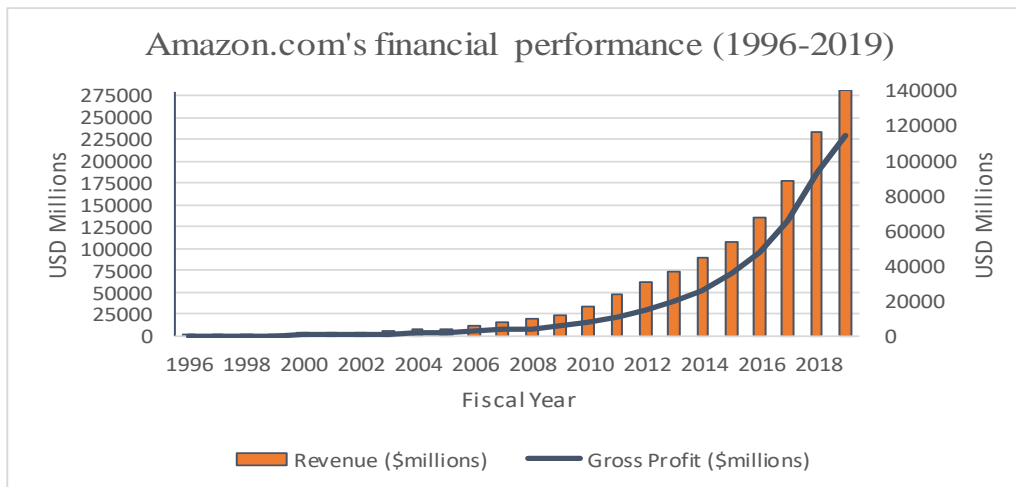


Figure 1 – Amazon’s financial performance over last 20 years (Source: Capital IQ)

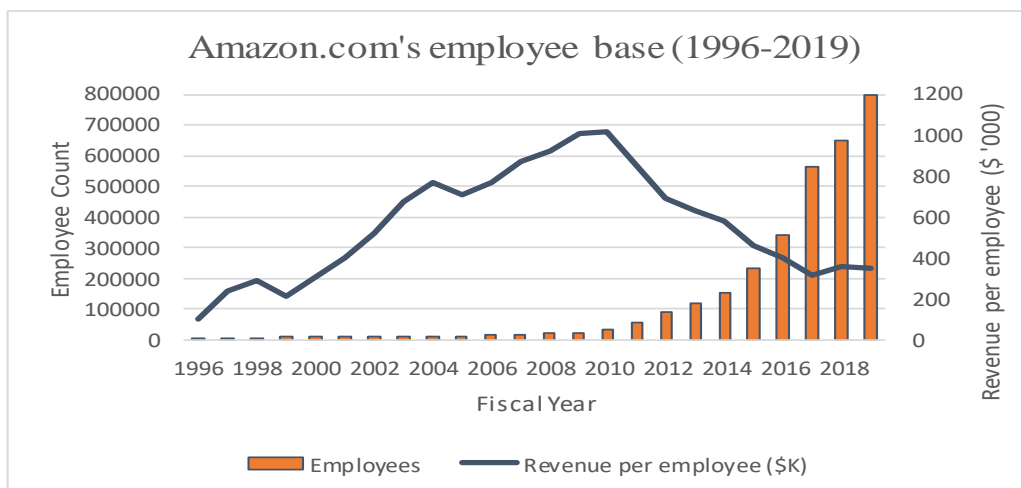


Figure 2 – Amazon’s employee base and revenue-per-employee over last 20 years (Source: Harvard Business School Case 9-716-402, Capital IQ)

How did Amazon achieve this success? Through its exponential growth over the last 25 years, Amazon has relentlessly focused on the three pillars of its flywheel model. After strengthening its foothold in book category, Amazon expanded into online video retailing (acquisition of IMDB in 1998), toys and electronics, home improvement and tools, videogames and software, furniture and kitchenware (in 1999), apparel (2002), sports and outdoor (2003), and health and personal care (2003).

Amazon expanded geographically to Japan, Europe, and Emerging Markets. It entered new businesses such as supply chain as a service (Amazon Marketplace), e-reader hardware (Kindle), online streaming (Prime Video), web services (AWS), and voice-command device (Amazon Echo). While this *expanded selections* made Amazon a household name for consumers [8], the large scale also provided Amazon competitive advantage in form of economic value - *low cost* of serving customers. (Figure 3).

As per Jeff’s long-term thinking, Amazon continued to invest capital (either raised or realized profits) into its operations to enhance *convenience* for customers. It built an in-house fleet of surface and air carriers to transport its packages uninterrupted and comply with its promised delivery time, developed an innovative Fulfilment Center (FC) design that suits its high-velocity product flow, and assembled a team of researchers, engineers, and product managers that re-invented the domain of tech-enabled supply chain design. Through these investments, Amazon constantly raised the bar on order-to-delivery lead time in the retail industry.



Figure 3 – Maximum selling product categories on Amazon in 2019 (Source: Statista)

## 1.2 Project Motivation

In 2005, Amazon launched Amazon Prime, a two-day delivery service for 1 million eligible products for an annual fee. It added other products over time, increasing the list of eligible products to 30 million items. In 2014, Prime Now was launched in New York, offering Prime members 2-hour delivery for 25,000 daily essentials. Subsequent rollouts in other US cities have tripled Prime memberships since then. In addition, a Prime member spends \$1,400 annually, twice that of a non-Prime member (Figure 4). Unsurprisingly, Amazon wants to sustain such supply chain innovations going forward.

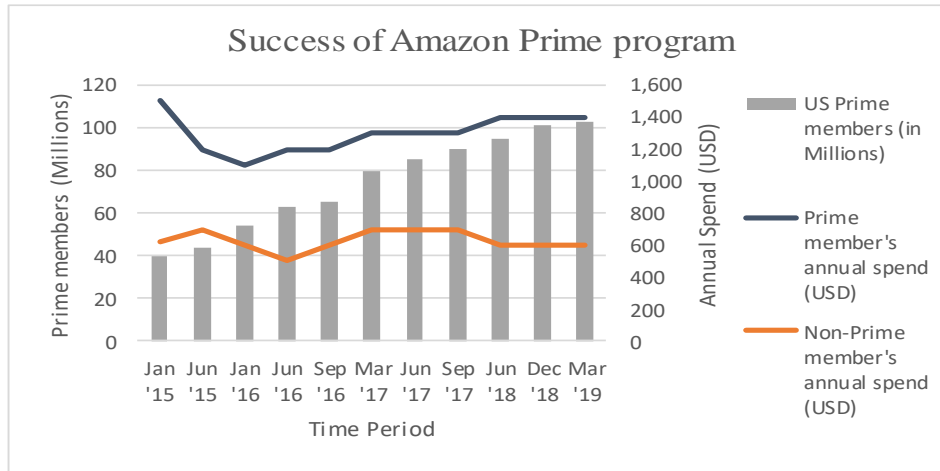


Figure 4 – Growth of Amazon Prime accounts in the US and difference in average spending between Prime accounts and Non-prime accounts (Source: Statista)

Like major retailers across the globe, Amazon continues to grow its private-label products (Appendix 1) to offer customers an alternative to existing brands. In this model, retailers source products directly from manufacturers to remove middlemen from the value chain, thereby capturing greater value and then passing it on to customers. Amazon’s brands have seen growing demand, which is forcing Amazon to re-think their supply chain strategy. (Figure 5).

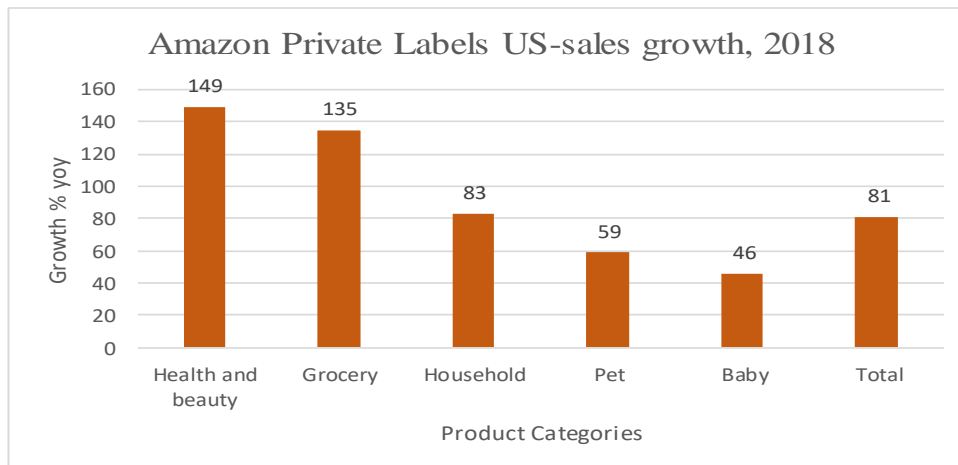


Figure 5 – Annual sales growth of Amazon private label categories in the US in 2018 (Source: Statista)

Unlike in traditional ecommerce model, which has negative cash cycle, private-label products require e-tailers (like Amazon) to own the inventory, adding burden on firm’s working capital, and hence, generates the need for optimal inventory planning and storing. This situation presents another inflection point in Amazon’s history of supply chain innovations, where the traditional e-commerce model needs to be optimally modified to service the demand for private-label products cost-effectively.

This thesis proposes a multi-echelon inventory model to effectively manage private-label products supply chain for e-tailers. A clear understanding of how supply chain innovations (such as Amazon Prime) have contributed to the success of e-tailers (such as Amazon) motivates this thesis work.

### **1.3 Problem Statement**

Amazon imports the majority of its private-label products from Asia; these products require sourcing larger quantities to account for long-lead time between production runs and high variability in demand to maintain competitive costs. Amazon's current supply chain is a single-echelon network of FCs, meaning all sourced products flows directly from the vendor's warehouse through cross-docks to the Amazon's FC network in the US, from which customer demand is served. The unique characteristics of private-label products' supply chain (mentioned earlier) result in their inventory dwelling for a long period at Amazon FCs, reducing the turns-ratio of expensive storage bins there, which could otherwise be utilized for storing high-velocity products.

Though private-label product was only 1% of Amazon's total sales in 2017 [9], the anticipated growth of private-label products raises the need to build more storage space, which is very expensive in its FCs, especially in highly-automated robotic FCs. Additionally, since fixed storage cost is proportional to the space occupied in FCs, high 'dwell time' translates into higher storage cost.

Independently, Amazon continues to invest in customer experience by reducing the delivery time to customer, resulting in the need to store more inventory close to customers, and in turn the need for more storage space in its FCs. In April 2019, Amazon announced 1-day free shipping for its Prime customers. With less time available to move products from FCs to customers, this change requires more inventory that is optimally spread in the FCs.

All the three factors call for higher storage in FCs, and storage space in FCs is expensive. Hence, lowering storage cost through better supply chain management is of utmost importance to keep Amazon's *flywheel* rotating.

### **1.4 Project Hypothesis**

Mathematically, inventory storage costs can be lowered by either (1) reducing the amount of inventory held in storage (without compromising on customer experience) or (2) reducing per-unit storage cost.

The author's hypothesis in this project is that intermediate echelon, low-cost storage nodes that are downstream of vendors in Asia and upstream of the FC network can serve both purposes.

Today, FCs need to hold large safety stocks that can cover demand fluctuations over the 18-week lead time from vendors in Asia to the US. The low-cost, upstream storage nodes will act as a storage buffer for the entire FC network. Demand aggregation at this echelon will result in the ‘pooling effect’ over the FC-cluster level demand variations, reducing the required safety stock, and hence, the overall target inventory in the network.

Additionally, the reduction in lead time to FCs from this echelon enables the placement of inventory in the FCs based on more accurate demand forecasts closer to when the demand actually materializes. This optimal placement decreases the transshipment cost by reducing the need for rebalancing transfers within the network, adding to cost savings beyond what was hypothesized earlier in this section.

### **1.5 Project Methodology**

The introduction of an intermediate low-cost storage echelon in Amazon’s current supply chain will convert its essentially single-echelon network into a multi-echelon network. For such a network, research shows that a multi-echelon inventory optimization (MEIO) model can generate a more cost-effective network design.

This thesis focuses on developing a MEIO model for Amazon’s private-label products supply chain, and then performing an in-depth study with the model for a suitable subset of the private-label products portfolio. By applying the MEIO model on a sample of private-label products, this thesis aims to unravel the key drivers of cost reduction among multiple possible network configurations, and recommend the one that incurs least annualized operations cost for the same customer experience.

## 2. Background

This chapter provides the relevant background on the importance of private-label products to retailers, the unique operational challenges that a retailer needs to overcome while dealing with private-label products, an overview of Amazon’s current supply chain network, and the other factors that significantly affect this study.

### 2.1 Emergence of private-label brands in Retail

The twentieth century experienced dominance of manufacturer brands as consumers moved away from inconsistent quality of local products to branded products from global manufacturers (such as P&G). These manufacturers used emerging media to highlight their product’s quality, focusing on trustworthiness and consistency. The supply chain in between manufacturers and consumers included wholesalers and retailers, who were small, powerless price-takers. But, in 1970s, retailers began to both consolidate and expand globally [10]. This trend continued to the point that retailers (such as Walmart) became larger than global manufacturers, changing the power dynamics in favor of retailers.

Taking advantage of their global knowledge in marketing, store operations, and merchandize planning, these large retailers introduced private-label brands into their stores and captured a larger share of the value created. Over time, the share of private-label products in category sales or channel sales has grown, both globally and in the US. (Figure 6). The charge is led by categories such as refrigerated foods, groceries, apparel, and accessories. Not just in volume, PB is competing with manufacturer brands in quality too.

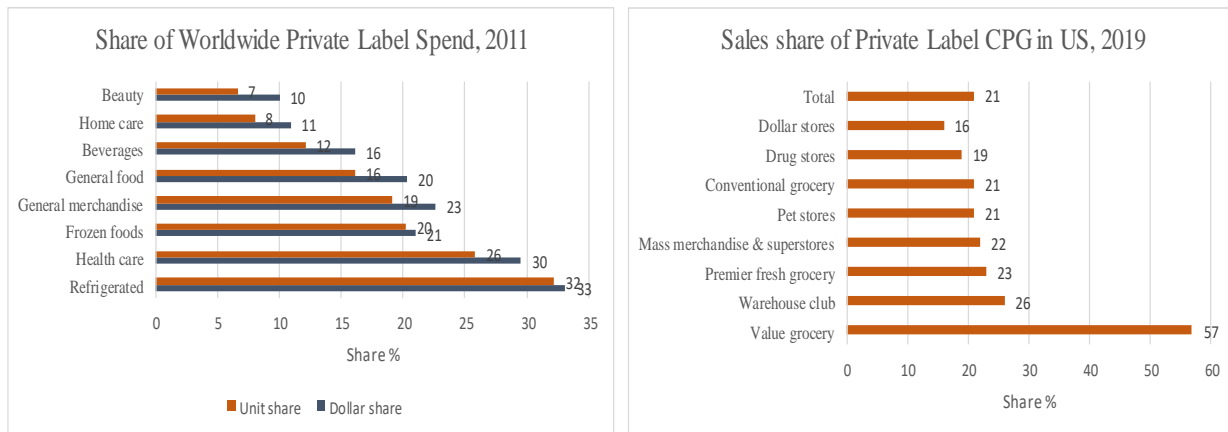


Figure 6 – Private-label as % of category-sales in the world (left) and in the US (right) (Source: Statista)

With the emergence of the internet, the same push towards private-label products is seen among e-tailers. Amazon has added 26 private-label brands in 2016 and 2017 combined. [11] Interestingly, the younger generations have limited fascination with manufacturer brands, and are more open to consuming private-

label products. (Figure 7). This is encouraging for retailers, as they continue to invest heavily in building their private-label portfolio.

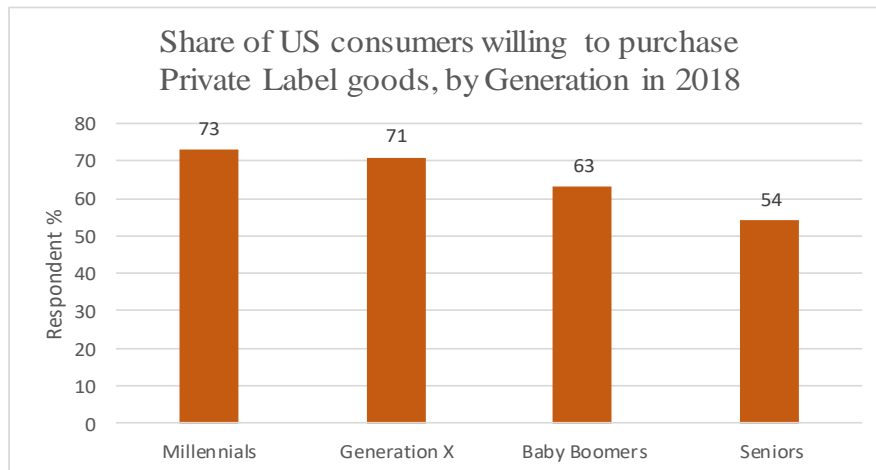


Figure 7 – Share of US consumers (by Generation) willing to buy Private Labels (Source: Statista)

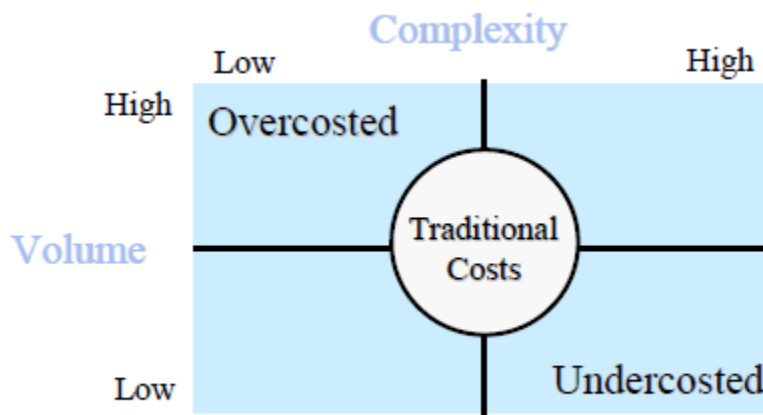
## 2.2 Supply chain challenges in private-label products business

Most retailers lack in-house manufacturing capability for their private-label products. They outsource the production of these products to manufacturing partners, who enjoy economies of scale by producing similarly designed items for different retailers, except for stamping the retailer-specific private label. Strategically, rather than competing directly with larger, more resourceful national brands and incurring business-related overhead expenses (such as sales, and advertising), these manufacturers choose to grow by marketing their product development and supply chain expertise to several retailers. These manufacturing partners are equally thrilled with the growth of retailer’s private-label portfolio, since their production runs lengthen as the required product volume increases and they realize further reduction in their per-unit manufacturing cost.

This symbiotic relationship between a retailer and its private-label manufacturing partner brings its own challenges for the retailer. Private-label product manufacturing has unique supply chain challenges, some of the risks of which the manufacturing partner wants the retailer to share. Once a manufacturer stamps a retailer’s private label (say, Amazon Basics) onto the same base product (say, a mobile charger), it can be sold only to that retailer (Amazon). Despite producing mobile chargers in bulk to meet aggregated demand from multiple retailers, this manufacturer has short production runs, one for each retailer, as it needs to change private label in the manufacturing process. Human resources to effect frequent changeovers, and machine downtime due to setup are sources of inefficiency and value wastage. To overcome this loss, the manufacturer generally imposes a minimum order quantity for each order on the retailer.

Most retailers have only one private brand for one product category in a particular consumer segment. Any product quality challenge can cause a big hit to the retailer’s reputation and threaten its play in that category. This is a huge risk outsourced by the retailer to its manufacturing partner. As such, the retailer requires its manufacturing partner to take utmost care to prevent mixing of products across retailers manufactured in the same shop floor. In addition, each retailer has a different quality check procedure, which the manufacturer has to adhere to.

Product complexity arising from aforementioned challenges such as low production batch-size and frequent set-up cycles, and customer complexity arising from unpredictable orders and special tests and requirements pose unique challenges to manufacturers. Harvard professor Robert Kaplan’s research [13] shows that such items are generally under-costed using traditional costing mechanisms. (Figure 8).



*Figure 8 – Traditional costing mechanisms distort product and customer costs*

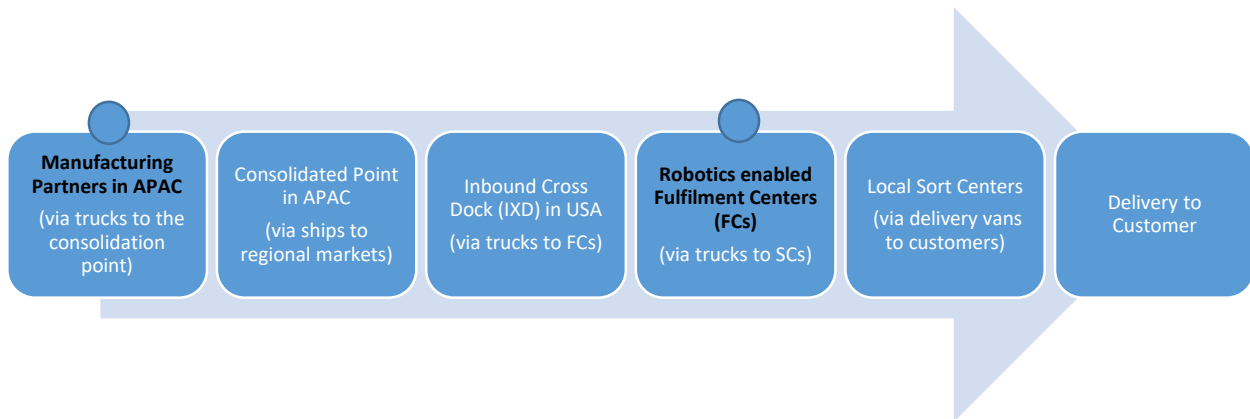
*(Source: HBS Note 9-197-076 on Introduction to Activity-Based Costing by Robert S. Kaplan)*

Hence, to avoid incurring losses, the manufacturing partner includes binding constraints, such as a fixed payment for allocating production capacity, minimum order quantity, minimum delivery lead time, advance payments, and handover inventory responsibility beyond manufacturer’s warehouse, on the retailer as part of their contract. Overcoming the private-label product’s supply chain challenges requires tight retailer-manufacturer partnership in real-time data sharing to master critical supply chain processes of (1) demand planning, (2) inventory optimization, and (3) flexibility to align supply with fluctuating demand. [14].

### **2.3 Amazon Supply Chain Network**

Amazon sources a portion of its private-label products from its manufacturing partners (vendors) in APAC. Like all other Amazon products, this supply chain has no intermediate storage node between manufacturer’s

warehouse in APAC and Amazon FCs in individual regions, say the US (highlighted in Figure 9). There is a consolidation point in APAC, where items from different vendors are received, sorted as per regional requirements, and shipped to individual regions. In individual regions, the shipped items are received at inbound cross docks, where items are de-consolidated and shipped to different FCs (based on local geographical demand forecasts).



*Figure 9 - Current Supply Chain for Amazon private-label brands*

This simple network offers high throughput for fast moving items, an ideal scenario for traditional e-commerce business, where e-tailers prefer not to own inventories and act as shippers between manufacturers and customers. However, in PB portfolio, e-tailers own inventories once it leaves the manufacturer's warehouse, thereby requiring cost-effective storage nodes along the supply chain network.

## **2.4 Supply Chain planning at Amazon**

Amazon follows the planning processes in a standard Sales and Operations Planning (S&OP) cycle. It starts with the forecasting team generating demand forecasts for each SKU for different planning horizons within a range (most items have weekly frequency). Then, vendor selection team generates a list of preferred vendors for each SKU with their expected delivery lead times. Based on the demand forecasts and on-hand inventory, the ordering team may either place one purchase order (PO) for the entire lot on one vendor or divide the PO quantity among multiple vendors. Based on the agreed upon lead time, each vendor delivers items at their warehouse, when the supply network mentioned earlier takes over logistics.

While the processes are standard, Amazon is unique in how it uses technology to run the processes. Most decisions are automated, facilitated through optimization algorithms running at the back-end. A team of research scientists and programmers constantly tweak these algorithms to improve their performance by incorporating advanced algorithms or new programming tools.

## **2.5 Amazon moved to 1-day delivery for its Prime members**

In April 2019, Amazon reduced its standard free delivery time for its Prime members from 2 days to 1 day. While this announcement boosted Amazon's position in the *convenience* pillar of its flywheel model, it has created new operational challenges for the company. Since the available shipping time from FCs to customers has been reduced from two to one day, Amazon has to restructure its FC network. To assure 1-day delivery, Amazon needs to redesign how it does inventory placement, so as to assure that inventory is locally available in all regions of the country. It has enhanced the need for (1) optimal spread of inventory across US FCs and (2) higher inventory holdup at each FC to supply local demand. These requirements hold true for private-label products too, so this study included 1-day delivery as a design criterion.

## **2.6 Pilot study in Amazon, US for low-cost storage node**

Amazon's traditional supply network is designed for productivity gains, a use case that is typically tailored to fast turning, high volume SKUs. However, not all private-label products fit this mold. Specifically, imported private-label products with attributes such as long-lead sourcing horizons, large minimum order quantities, and new item demand variability lead to high inventory in the network. The economics of private-label portfolio in the network requires a low fixed-cost structure to support future investment and expand selection.

Amazon's Inbound supply chain team piloted the concept of 'an intermediate low-cost storage node' that enabled bulk storage for selected private-label products. This node was designed to maximize storage space utilization. The node was located upstream of the cross-dock network and acted as the first point of receive for new vendor freight. Imported inventory arrived into this node, and got routed into FCs on a just-in-time basis according to localized demand patterns, transit time, storage cost, and network topology. The pilot showed promise for lower per-unit supply chain cost. This is another input for this study to explore further.

### 3. Problem Formulation

We have discussed the need to modify the traditional high-speed supply chain of an e-tailer to tackle the nuances of private-label portfolio in a cost-effective manner. Considering Amazon's supply chain as a representative e-tailer supply chain, this chapter identifies the scope for improvement in its network to benefit private-label portfolio, narrows down on a subset of its private-label products to explore in depth throughout the course of this study, and formulates the optimization problem this study aims to solve.

#### 3.1 Identification of improvement areas

In section 2.2, we identified three major capabilities e-tailers must develop to effectively manage private-label products supply chains. Amazon has a state-of-the-art demand planning process that considers demand-related, vendor-related, cost-related factors to forecast demand and determine purchase orders on most suitable vendors.

There is scope for improvement in inventory optimization, which if done well, in turn also improves the flexibility of matching supply and demand at the point of sale. As explained previously in section 1.3, Amazon's supply chain has operated as one giant FC network with a single stage (or *echelon*), meaning the system does not differentiate between the exact FC location where a product is placed as long as it is placed in some FC. The only inventory-related question that it had to answer was how much quantity of an SKU to buy for anticipated demand over a certain review period. Solution to the classic newsvendor problem in inventory management theory succinctly answered this question.

However, the new requirement of inventory placement in FCs close to the demand generation point to fulfill 1-day delivery promise (section 2.5) and that of low-cost storage among private-label products (section 2.6) necessitate consideration of a multi-echelon supply chain as an alternative. But, with each new echelon, there is an additional touch point, increasing handling costs.

As a result, this study aims to address these new inventory-related questions such as - a) which stages should store inventory, and b) how much inventory of each SKU should be stored in those stage.

#### 3.2 Selection of a subset of private-label portfolio for detailed study

Amazon's private-label product basket runs into tens of thousands SKUs. They vary in their demand patterns, product size, unit economics, delivery lead times, etc. To deliver maximum benefits to Amazon, this study focused on a subset of its private-label portfolio. A sensitivity analysis on the results of the *low-cost storage pilot* (explained in section 2.6) revealed that a minimum *dwell time* at the low-cost node was essential to compensate for this node's operational cost, an additional cost introduced into the system. This resulted pointed to the *imported product basket* as the prime beneficiary of the low-cost storage.

The SKU list obtained from applying the *import* filter on Amazon’s private-label portfolio was still in tens of thousands, making it ineffective to work with. To reduce the target item list to a manageable few hundreds, the study applied several other filtering lenses – 1) business lines that constituted the most imported products, 2) average product volume (physical size), and 3) non-hazardous product class. These filters revealed that most private-label products imported into US (both by units and by value) 1) belonged to *Business Line 1* (BL1) and 2) were *small* in size. (Figure 10). We further analyzed this filtered list.

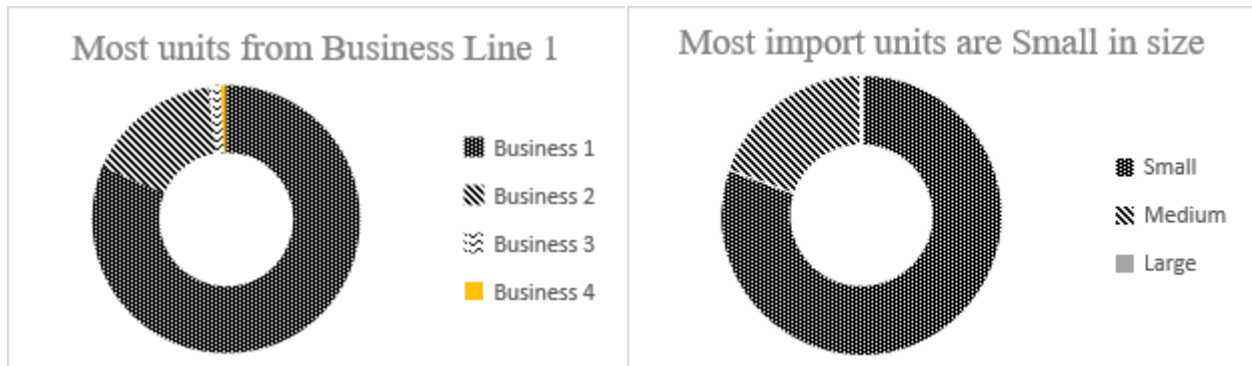


Figure 10 – Breakup of imported Amazon Private-label by Business line (left) and by product-size (right)

Analyzing small-sized, imported SKUs of BL1, the study found that new products and seasonal products account for only 20% of sales volume and 15% of sales value. The SKUs in these categories were excluded from the study. The study focused on *High Priority* and *Regular* SKUs that occupy a large share of storage space in Amazon FCs, and account for 80% sales volume and 85% of sales value (Figure 11). These 763 SKUs have regular customer demand and hence, need to be maintained in stock.

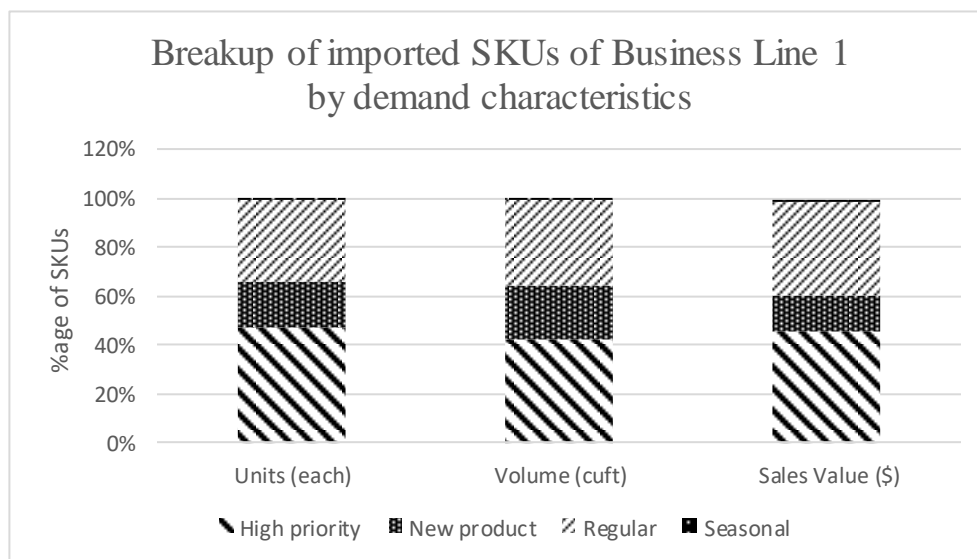


Figure 11 – Breakup of imported SKUs of BL1 by demand characteristics

### **3.3 Formulation of the optimization problem**

At this point, it was clear that a low-cost storage node between the manufacturer's warehouse in APAC and FC network in US (upstream of cross docks) would be a potential solution to better manage private-label product's supply chain. However, it was not clear what the optimal geographical location would be.

At a broad level, the nodes could be built either (1) close to the customers in the US or 2) close to the suppliers in APAC. After one of these broader geographies is finalized, Amazon needed to also decide the specific geographical locations for these nodes. These are the additional questions (beyond those in 3.1) that this study attempts to address; specifically, where should the first full-fledged 'low-cost node' be built. These strategic decisions will affect the total private-label supply chain costs.

Thus, the objective is to achieve the minimum annual private-label products supply chain costs while meeting the target customer service level and within 1-day delivery period. The study is scoped for 763 private-label products SKUs identified by the filtering process mentioned in section 3.2.

#### 4. Literature Review

Supply Chain Management (SCM) is ‘the backbone of modern economy’. In a way, SCM aids realization of the advantages from global trade theorized by economist David Ricardo in his 1817 classical theory of comparative advantage by connecting the most cost-effective suppliers of a product at one part of the world with consumers across the globe. It has been a dynamic field of research since the rise of globalization in 1960s, especially after enabling technological inventions such as personal computer and internet in 1990s. To manage today’s complex, multi-stage, global supply chains, like the one we have in our study, SCM researchers have found useful Multi-Echelon Inventory Optimization (MEIO) models that optimize inventory decisions for the network as a whole.

#### 4.1 Multi-Echelon Inventory Optimization (MEIO)

In a multi-echelon supply chain, each stage represents one of several processes of the value chain that transforms raw materials at their origins into finished commercial product in a consumer’s hand. Based on the interconnection between different stages, multi-echelon systems can be classified as a) serial, b) assembly, c) distribution, d) acyclic, and e) cyclic based on their structures. (Figure 12). Each stage is a potential location to store safety stocks of the intermediate product produced at that stage. MEIO looks at inventory across all stages in the network, considers the impact of inventories at a given stage on other stages, and provides the optimal safety stock levels for each stage.

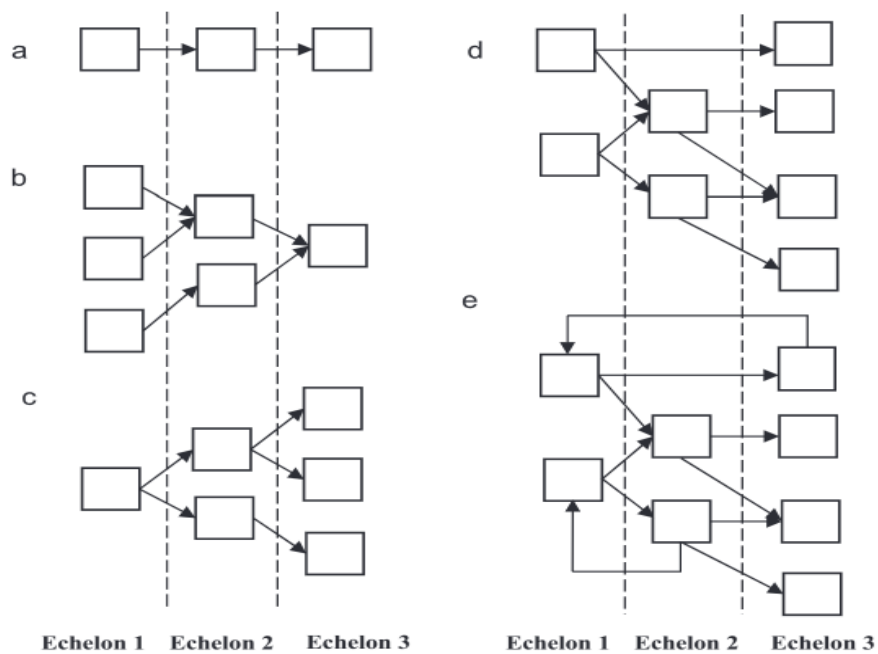


Figure 12 – Different network topologies for Multi-Echelon Supply Chain systems

MEIO research literature show two main approaches – 1) stochastic service model (SSM) introduced by Clark and Scarf in 1960, and 2) guaranteed service model (GSM) introduced by Simpson in 1958 [15]. The two approaches differ in the assumptions they make on how orders from a given (customer) stage are replenished by its (supplier) stage. Table 1 summarizes the key differences between the two. [16] The stochastic nature of the service time in SSM, its exact characterization becomes challenging, especially for large and complex supply chains. On the other hand, given its assumption that extraordinary demand situations are addressed with other tactics, GSM can determine the inventory needed for guaranteed service in complex supply chains as long as demand is within predetermined bounds.

*Table 1- Comparison between two approaches – SSM and GSM – to Multi Echelon system design*

Comparison between SSM and GSM approaches to MEIO			
	Attribute	SSM	GSM
<b>Demand Process</b>	Process	A stage places order on its preceding stage	A stage places order on its preceding stage
	Demand Quantity	Any quantity can be demanded	Demanded quantity is bounded
<b>Service Process</b>	Process	A stage services all connected downstream stages	A stage services all connected downstream stages
	Quantity Serviced	Depends on availability	100% if within demand bound
<b>Service Time</b>	Process	If items available, immediate replenishment If not, backorder and wait for next supply	If demand within bound, 100% within agreed Service Time If not, demand bound serviced within agreed Service Time; Assumes other tactical responses will be pursued
	Nature	Stochastic	Deterministic
<b>Best for</b>		Serial, Assembly, Distribution supply chain structure	Any general supply chain architecture
<b>Challenges</b>		Complex to solve, especially for general architecture	Does not solve for extraordinary demand cases

We shall use GSM to model the supply chain in our study to ensure that the same decision-making fundamentals work even when Amazon’s PB supply chain becomes complex in future.

#### 4.2 Guaranteed Service Model (GSM)

For modeling purposes, we assume decentralized control at each stage in the supply chain; namely, we assume that each stage manages its inventory with a simple base stock control policy that takes inputs from adjacent upstream and downstream stages. However, for practical implementation, the final recommendations of the model with regard to safety stocks can be translated into the system-driven control policies that govern decisions across the entire supply chain.

To introduce the assumptions for the guaranteed-service model (GSM), we will refer to the Figure 13, in which a generic stage  $k$  of the supply chain is depicted. This stage serves a customer (stage  $k-1$ ) and has a supplier (stage  $k+1$ ) that serves it. As shown, stage  $k$  takes an input from its supplier and processes this input into its final product that it can hold in inventory and that it supplies to its customer.

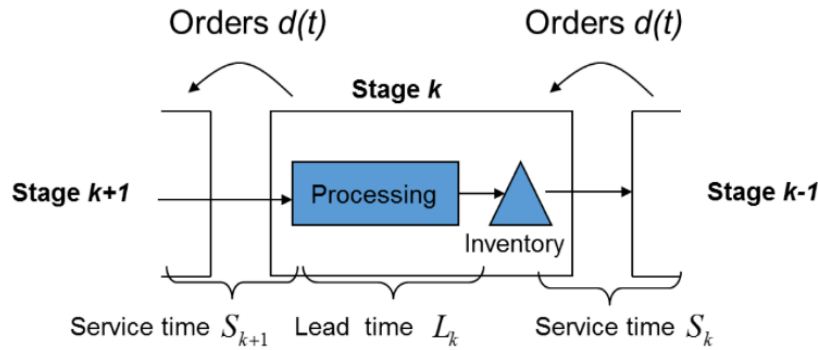


Figure 13 – Schematic to formulate Guaranteed Service Model

Key assumptions for GSM are:

- Periodic review, base stock control policy: Each stage operates under a periodic review policy with the review period of one-time period. In each time period  $t$ , stage  $k-1$  places an order,  $d(t)$ . Stage  $k$  controls its inventory with a base stock policy, meaning it will place the same order  $d(t)$  on its supplier to initiate replenishment of its inventory. Thus, the downstream demand from the customer is passed upstream to the supplier immediately.
- Deterministic Lead Time: the time to replenish the inventory in the stage is deterministic and given by  $L_k$ , assuming the inputs from stage  $k+1$  are available.
- Stationary demand: We assume stationary demand, implying the demand has a constant mean ( $\mu$ ).

$$\text{Expectation Value } [d(t)] = \mu$$

- Demand Bound: For setting the safety stock, we assume an upper bound on the demand at each stage over any interval of time. In particular, we are given a demand bound  $D(s)$ , which is defined as the maximum demand over any time interval of  $s$  periods.

$$D(s) \geq d(t-s+1) + \dots + d(t), \forall t$$

- Guaranteed service time: we assume that each stage quotes a service time by which it will satisfy any demand from its customer. In Figure 13, stage  $k$  quotes a service time  $S_k$  to its customer, stage  $k-1$ ; the implication is that the order that is placed at time  $t$  is to be fulfilled by time  $t + S_k$ . Similarly, stage  $k+1$  quotes a service time  $S_{k+1}$  to stage  $k$ , with the same interpretation.
- Integral parameters: we assume that the lead times  $L_k$  and service times  $S_k$  are non-negative integers, which implies that they are each multiples of the review period.

### Mathematical formulation

To model the inventory at stage  $k$ , we make observations on the inventory dynamics in each time period  $t$ :

- We assume that the replenishment time for stage  $k$  ( $S_{k+1} + L_k$ ) is greater than or equal to the service time  $S_k$ . If this were not true, then the stage can operate with a make-to-order policy with zero inventory, as its replenishment time is less than the service time it quotes to its customer.

- In period  $t$ , we observe the demand  $d(t)$  and place an order for this amount on our supplier, stage  $k+1$ . This order is received at time  $t + S_{k+1}$  and then enters processing at the stage  $k$ , which has a lead-time of  $L_k$ . Thus at time  $t + S_{k+1} + L_k$ , the order of  $d(t)$  units is available in inventory at the stage, and the replenishment lead-time is  $S_{k+1} + L_k$  periods.
- In period  $t$ , we observe the demand  $d(t)$  and commit to satisfy by our service time; that is the demand will be served at time  $t + S_k$ .

With this understanding of the inventory dynamics in each period, we can model the inventory at stage  $k$  at time  $t$  with the inventory balance equation:

$$I_k(t) = I_k(t-1) - d(t - S_k) + d(t - S_{k+1} - L_k)$$

This inventory is equal to the inventory in prior period i.e.  $I_k(t-1)$ , minus the demand served in period  $t$  i.e.  $d(t-S_k)$ , plus the replenishment in period  $t$  i.e.  $d(t-S_{k+1}-L_k)$ . With the assumption that the system started at  $t=0$  with  $I_k(0) = B_k$  (base stock), we can use backward substitution to obtain:

$$I_k(t) = B_k - d(t - S_{k+1} - L_k + 1) - d(t - S_{k+1} - L_k + 2) - d(t - S_{k+1} - L_k + 3) - \dots - d(t - S_k)$$

Thus, the inventory is a base stock level minus the demand that has been satisfied but that has not yet been replenished. This is the demand over a time interval of length equal to the net replenishment time, where the net replenishment time ( $T_k$ ) is the replenishment time for stage  $k$  minus the service time promised by stage  $k$ . The mathematical formula is as given below:

$$T_k = S_{k+1} + L_k - S_k$$

To assure guaranteed service, we need to have  $I_k(t) \geq 0$ , for which, we need to set the base stock such that:

$$B_k \geq \max[d(t - S_{k+1} - L_k + 1) + d(t - S_{k+1} - L_k + 2) + d(t - S_{k+1} - L_k + 3) + \dots + d(t - S_k)]$$

Using the assumption of bounded demand over a given period, we set base stock equal to the maximum demand over the net replenishment time  $B_k = D(T_k)$ . The expected value of inventory at stage  $k$  at any period  $t$  is the expected inventory at stage  $k$  at any time to get guaranteed service, given by the following equation.

$$E[I_k(t)] = E[B_k - d(t - S_{k+1} - L_k + 1) - d(t - S_{k+1} - L_k + 2) - \dots - d(t - S_k)] = D(T_k) - \mu T_k$$

This quantity depends on the net replenishment time, the average demand rate, and the demand bound. The net replenishment time depends on a) inbound service time, b) stage's processing time, and c) outbound service time. While the service times (a and c) are control variables, each stage's processing time is agreed upon as a service-level agreement between different sub-teams of a firm's supply chain.

The demand bound is a forecast of the range of demand variability for which we want to have sufficient safety stock. We chose the most common way to compute the demand bound:  $D(s) = \mu s + z\sigma(\sqrt{s})$ , where  $\mu, \sigma$  represent the mean and standard deviation in daily demand and  $z$  is the safety factor. Thus, the base stock level  $B_k$  is calculated as:  $B_k = D(T_k) = \mu T_k + z\sigma\sqrt{T_k}$

### 4.3 Strategic Inventory Placement Model (SIP)

Strategic Inventory Placement (SIP) model is a simple GSM-based inventory optimization model that finds the optimal placement and quantity of inventory across supply chain nodes [17]. For modeling purposes, we consider a serial supply chain of  $k$  stages. Each stage operates as per GSM explained earlier.

SIP model is based on the following parameters:

$L_k$ : the process lead time for stage  $k$ ,  $k = 1, 2, \dots, K$  (time periods).

$h_k$ : the holding cost for inventory held at stage  $k$ ,  $k = 1, 2, 3$  (\$ per unit per unit time)

$D(s)$ : demand bound, equal to the maximum demand from a customer over interval of length  $s$  periods.

$\mu$ : the average demand per period from the external customer

$S_1$ : guaranteed service time by which a customer should be serviced; typically, it is set to zero, implying instantaneous service

$S_{k+1}$ : the guaranteed service time for stage  $K$  to receive its ordered stocks

### Mathematical formulation

Given these inputs, the SIP model determines where to position safety stocks, and how much to store, to assure guaranteed service with the least amount of inventory holding cost.

The decision variables in the SIP model: Service times for each stage, namely  $S_k$

We now use the results from the guaranteed service model to represent the inventory at each stage:

$$E[I_k(t)] = D(T_k) - \mu T_k$$

Using the holding cost rate at each stage, we formulate the optimization problem as follows:

$$\text{Min} \sum_{k=1}^K h_k x (D(T_k) - \mu T_k) = \text{Min} \sum_{k=1}^K h_k x (z\sigma\sqrt{S_{k+1} + L_k - S_k})$$

Subject to

$$T_k = S_{k+1} + L_k - S_k \geq 0, k = 1, 2, 3, \dots, K; S_k \geq 0$$

$$S_1 = 0; S_{K+1} = \text{SKU specific vendor delivery Lead Time}$$

The solution of this optimization problem is actually easy, due to an interesting property of the optimal solution. Simpson proved that for a serial supply chain the optimal service times satisfy an extreme point property where the service time at a stage will equal either zero or its replenishment time [18].

That is, for each stage we have:

$$T_k = S_{k+1} + L_k - S_k = 0 \text{ or } S_k = 0$$

When  $S_k = 0$ , the stage needs to hold enough inventory so as to provide immediate service; as such its safety stock will need to be sized to cover any possible demand over its replenishment time. In this case, there is a decoupling point at stage  $k$ . The downstream stages can operate independent of the upstream stages, given that there is a large safety stock at stage  $k$  that is sufficient to guarantee a service time of zero.

When  $S_{k+1} + L_k - S_k = 0$ , the service time at stage  $k$  equals its replenishment time and this stage need not carry any safety stock.

This result helps formulate a methodology that facilitates the determination of the optimal solution. Simpson proves that each stage has two possibilities for optimal inventory storage – all or nothing. This means that there are  $2^{(K-1)}$  configurations of inventory storage. If  $K$  is not very large, one can enumerate each possible configuration and note the total supply chain costs. The configuration with the minimum storage cost is the optimal solution.

## **5. Research Methodology**

This chapter highlights the approach taken to build the strategic inventory placement model of Amazon's supply chain at the right level of abstraction, analysis of which can answer the key question identified in section 3.3. Following sections outline the identified success metrics, the network modelling strategy, the data collection methods, the plan for sensitivity analysis, and the working formulae used in the study.

### **5.1 Metrics to measure the success of the study**

In a company's financials, the supply chain function is typically a cost center. Thus, the true measure of success for the application of this study was agreed upon to be the reduction in 'total supply chain costs' measured annually. Measuring over a year not only normalizes for seasonal variations of demand, and variations in ordering frequency in a given year, but also aligns well with company's financial planning cycle. Total supply chain cost for a given SKU A is comprised of 1) inventory storage cost, and 2) opportunity cost of Amazon's working capital tied up in SKU A's inventory in the system, namely the purchase cost and all operational costs incurred in shipping. Other components of supply chain costs such as transportation cost and holding cost for pipeline inventory are not included in the metrics since these costs will be same across various inventory storage strategies. These cost terms are explained below.

Inventory storage cost includes the sum of the costs apportioned for the operational resources (storage space, and material handling) allocated for SKU A inside the walls of the storage nodes across the supply network. A larger SKU uses more storage space and material handling resources. Thus, inventory storage cost is generally proportional to the physical volume of a SKU. Moreover, the cost of these resources varies based on the geographical location of the storage node; the same resource is more expensive in the US than in the APAC region. Each node in the supply chain has an estimated storage cost, per-unit volume, per-time period. Simply multiplying this rate with the SKU physical volume estimates the storage cost per unit per-time period.

Opportunity cost is a fraction (called, opportunity cost rate) of the net working capital tied up in SKU A's inventory in the system, calculated as the sum of the costs incurred in SKU A's supply chain outside the walls of the storage nodes - purchase, transport, customs clearance, value adds, etc. The working capital tied to one unit of SKU A will vary based on where the inventory is stored in the supply chain, and will be higher in a downstream node closer to the customer and lower in an upstream node closer to the supplier. A unit of SKU A in the FCs will already have incurred the purchase, transportation, and handling costs in the supply chain, whereas a unit of SKU A in vendor's warehouse in APAC will only have incurred the purchase cost. For modeling purposes, we used 25% as the opportunity cost rate.

## 5.2 Building the private-label supply chain model

We began our analysis with the current private-label supply network and modified the network, identified design constraints, and made assumptions to build from ground-up the possible configurations that we would comparatively evaluate using the SIP model (section 4.3) for minimum ‘total supply chain costs’.

### 5.2.1 Current private-label supply chain

As mentioned in section 2.3, Amazon sources a lot of its private-label products from vendors in APAC. Shipment is inter-modal, trucks for surface transport and ships for the ocean way. There are no intermediate storage points; once a consignment leaves the vendor’s warehouse, it is either in transit or awaiting processing at ports.

A representative delivery lead time is 18 weeks from ‘raising PO’ to ‘receiving products in FCs’. (Figure 14). The breakdown of this lead time is: 1) 10 weeks of manufacturing lead time, i.e from raising purchase order to vendor manufacturing the product, 2) 1 week of surface transport from vendor’s warehouse to consolidation point in APAC, 3) 2 weeks of processing time at the consolidation point, 4) 4 weeks of ocean transport from APAC to US, and 5) 1 week of surface transport from US port to inbound dock of destination FC.

All FCs in the US are considered part of a single, nation-wide FC network, meaning that it is assumed possible to service most demand originating at any US location from any FC in the country within two-day delivery period. In a way, the SKU inventory is managed as if it were centrally controlled. The right level of abstraction to model this inventory management principle is one node representing all FCs.

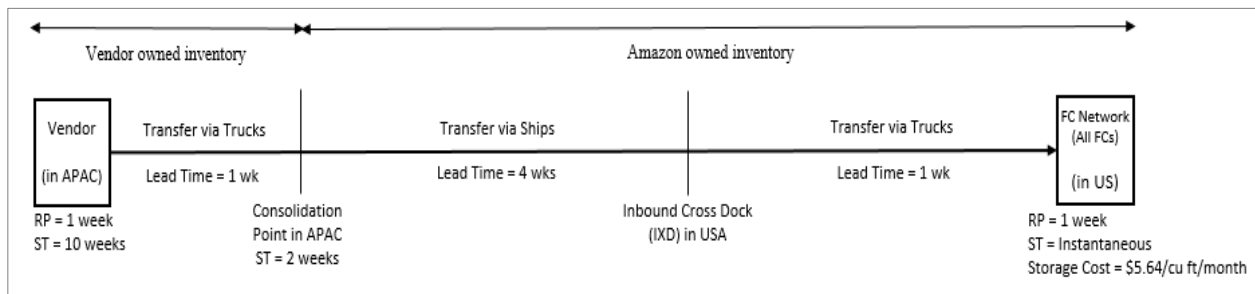


Figure 14 – Schematic of current Amazon private-label supply chain for imported products

### 5.2.2 Network modeling to meet 1-day delivery promise

Amazon’s April-2019 announcement of 1-day delivery promise to its Prime members implied that its last-mile delivery team will have less time available to move products from FC network to customers.

Assuming that last-mile delivery methods remain same, this change requires 1) more inventory that is 2) optimally spread in the network. For optimal spread, the Inventory Placement Team must stage safety stocks required to fulfill aggregate customer demand at a more granular level than the previous national level; that is the safety stocks are planned and placed (say) at a state or a regional level, rather than at a national level.

The decision should be based on the last-mile team’s ease of covering all geographical locations, at least concentrated demand-generation points in a given region from nearby FCs within 1-day. This new level of abstraction in the supply chain network would ensure that Amazon meets its 1-day delivery promise. For modeling purposes, we divided the US into 17 regional clusters (Appendix 2). In this structure, the inventory planned for a cluster (say, cluster #9) would be shipped from a cross dock to one or more of the FCs in cluster 9 (Figure 15).

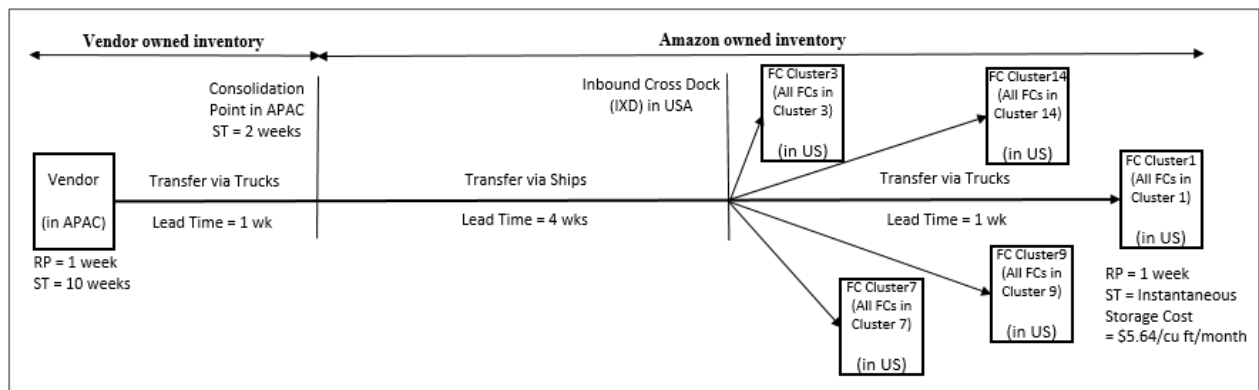


Figure 15 – Representative schematic of private-label supply chain to meet 1-day delivery promise

### 5.2.3 Network modeling to reduce inventory storage cost

Inventory storage costs can be lowered by either (1) reducing the target inventory (without compromising on customer experience) or (2) reducing inventory storage cost per-unit (explained in section 5.1). A low-cost storage node (LCS) upstream of the cross dock can also help reduce the system inventory, beyond the obvious contribution of lowering the inventory storage rate.

The LCS achieves this result by unlocking risk pooling benefits. It allows the ‘pooling’ of customer demands across multiple FC clusters, resulting in less variable aggregate demand at the LCS. As a consequence, the LCS can hold less safety stock while meeting the same customer level, and hence, reduce the overall inventory in the network. (illustrated in Figure 16).

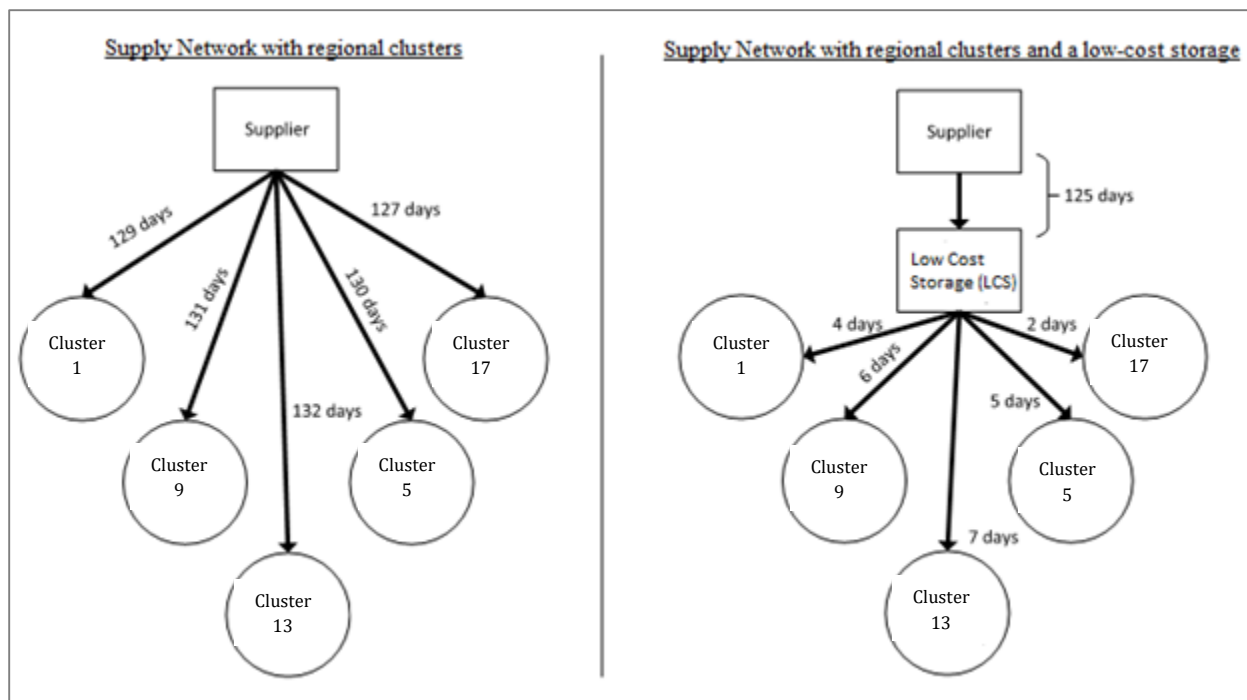


Figure 16 – Representative schematic showing LT benefits of introducing a LCS node on target inventory

As per 5.2.2, FCs in a cluster need to hold large safety stocks that can cover demand fluctuations over the 18-week lead time from vendors in APAC. However, the proposed LCS node will act as an upstream storage buffer with just 1-week delivery lead time for the FC clusters. By replenishing FC inventory quickly, the LCS allows each FC cluster to hold less safety stock, freeing up precious FC storage capacity. This reduces the need for capital investment in building new FCs.

In addition, by shortening the LCS-to-FC cluster lead time, the LCS allows the inventory placement decision for each FC cluster to be made closer to the point in time when the demand actually materializes. This significantly improves the demand-supply matches at each cluster as inventory is sent to the FCs based on more accurate demand forecasts, reducing the need for rebalancing transfers within the network, and avoiding increases to the outbound shipping costs (which is the major share of retail logistics).

While the benefits from a low-cost storage node are evident, the optimal location to build this node (in APAC or in the US) is less obvious. Figure 17 shows the two possible nodes (LCS1 in the US and LCS2 in APAC) incorporated in the supply network from 5.2.2. The total lead time from the vendor to the FCs remains the same, but there are intermediate storage nodes in the network now that can hold excess inventory in them at a cheaper location, and ship the required inventory to its upstream node when demanded. LCS2 meets demand from LCS1, and LCS1 meets demand from FCs. This abstraction of

private-label's supply network closely resembles the SIP model (explained in section 4.3) with 3 storage nodes - FC, LCS1, and LCS2. Due to the service requirements we will always hold inventory at the FCs. Hence, there are four [  $2^{(3-1)} = 2^{(2)}$  ] inventory storage configurations that are plausible to evaluate, namely (1) neither LCS1, nor LCS2, (2) LCS1 only, (3) LCS2 only, and (4) LCS1 and LCS2.

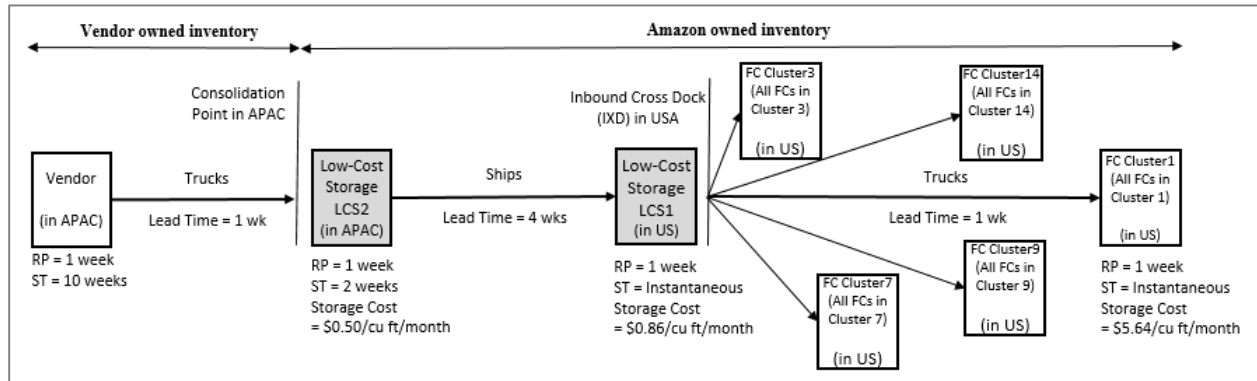


Figure 17 – Representative schematic of private-label supply chain after introducing low-cost storage nodes (LCS1 in the US and LCS2 in APAC)

Theoretically, we identified the cost-benefits between LCS1 and LCS2 as being due to: (1) shorter lead-time to the FCs, (2) lower storage costs, and (3) greater pooling over demand variations. Table 2 provides a comparison between the two storage options. There are certain pros and cons to both. While LCS1 provides a shorter and more reliable lead-time to the FCs, LCS2 provides lower storage costs. To resolve this puzzle, we needed to quantify the financial benefits of each supply chain; we do this by using the SIP model to evaluate a set of supply chain scenarios that use LCS1 or LCS2 or both.

Table 2- Comparison of LCS1 and LCS2 storage nodes

Parameter	Low Cost Storage node in US (LCS1)	Low Cost Storage node in APAC (LCS2)	Preferred
Description	Storage facility in the US, between import port and IXDs	Storage facility in APAC, receiving supply from PB vendors	N/A
Storage cost rate	\$0.86/cuft/month, 85% improvement over FC-storage rate	\$0.50/cuft/month, 91% improvement over FC-storage rate	LCS2
Lead time to FCs	1 week	5 weeks	LCS1
Demand Pooling benefits	High	Low	LCS1

### 5.3 Rationale behind using GSM to model private-label supply chain

Before proceeding with the SIP model, we verified whether it is reasonable to model each node of the private-label supply chain using GSM, the basic building block of a SIP model. To do that, we cross check each condition from section 4.2 with real data for 763 SKUs used in the study.

- Periodic review, base stock control policy: Inventory planning at Amazon FCs followed a 1-week base stock review policy, arguably an industry standard. This study continued to assume the same for FCs, and correspondingly added the review period of the new nodes – LCS1 and LCS2. For LCS2, the review period was set at 1-week. However, to maximize benefits to the network from ‘pooling’, for the scenarios with LCS1, we assume that the FCs operate with a 1-day review period,

an assumption that made sense since in reality, a cross dock regularly sends trucks to the individual clusters on a daily basis.

- Deterministic Lead Time: The lead time (LT) between vendors in APAC and FCs in the US is divided into two parts, 1) Vendor-owned Manufacturing LT, and 2) Amazon-owned logistics LT. Since inter-departmental service level agreements (SLAs) ensure greater control over internal processes, we can assume the latter to be deterministic. As for the former, statistical analysis of this data for different SKUs over a period of 20 months revealed a peak at 9 weeks. (Figure 18). The outliers in the data represented cases of *staggered deliveries*, meaning deliveries received from a vendor in different lots over time for the same purchase order. These inputs proved that using a higher manufacturing LT data (mean LT + some buffer LT) for each SKU would be a reasonable assumption, to allow the LT to be a deterministic variable.

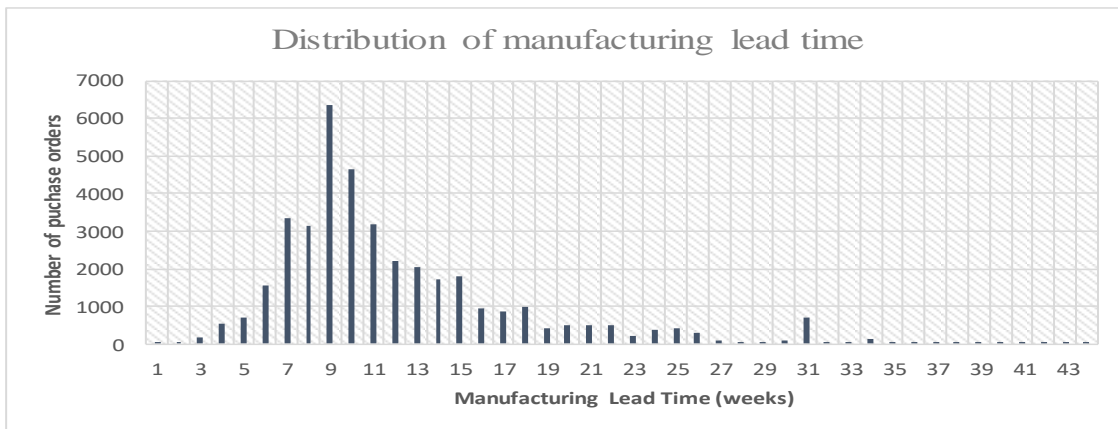


Figure 18 – Distribution of manufacturing lead time for private-label products over 20 months

- Stationary and Bounded Demand: If we are able to come up with a relatively stable demand bound, this assumption will be met. We use the lens of classic newsvendor problem formulation to arrive at a demand bound. For an SKU, the demand bound for a given horizon is stable if it has (1) a stationary demand distribution (constant mean and standard deviation) and (2) a stable target (type-1) customer service level that informs how much of the demand distribution a company aims to cover (as defined in the classic newsvendor problem). Both of these inputs influence the company’s inventory stocking norm, and thus the demand bound.

Given the evergreen nature (regular use) of products in the private-label portfolio, customer demand inherently for any SKU A is either stable or at best, predictable within a range. Even for products showing some seasonal fluctuations, we can consider the demand to be month-wise or quarter-wise stable. To validate this hypothesis, we analyzed the weekly demand for 763 private-

label SKUs over 20 months, Figure 19 (left) shows the cumulative share of these SKUs arranged in ascending order of their coefficient of variation (CV) of their weekly demand (standard deviation / mean of weekly demand). We find that 95% of these SKUs have demand distribution with a CV for weekly demand that is less than 1, showing predictable variation in weekly demand.



Figure 19 – Distribution of selected private-label SKUs’ demand (left) and service level (right)

A similar analysis of target (type 1) weekly service level for these SKUs over the same period shows that 81% of the SKUs have coefficient of variation of their target weekly service level within 10% of the mean weekly service level. (right graph in Figure 19). We can consider this variation negligible over a long horizon and thus, approximate a stable target service level.

- **Guaranteed Service Time:** The assumption made on manufacturing LT (above) and strict inter-departmental SLAs within Amazon logistics team ensured validity of this assumption.
- **Integral parameters:** Based on real data, as observed in periodic review policy, this was an easy assumption to make.

## 5.4 Model setup

In order to estimate the numerical parameters required to build SIP model (section 4.3) of the supply network, this study assumed that there would be at most one node each of LCS1 (in US) and LCS2 (in APAC) that would cater to all the 763 SKUs in the study. That is, each facility has sufficient storage space, so that there will not be a storage space constraint. This seemed reasonable given the intent to check the feasibility and financial viability of the idea to build low-cost storage nodes.

### 5.4.1 Selection of location for LCS1 (in US) and LCS2 (in APAC)

The operational costs related to storage and inter-nodal transportation depend on the geographical locations of the nodes. Analysis of the imported private-label products over 2018-19 revealed that most

consignments were received at one of two ports in the US. This study will assume that the single LCS1 node is located close to the port with maximum private-label receipt (Port1). From the same analysis, we found that most private-label products are shipped out of one country in APAC (Country1). (Figure 20) This study assumes that the single LCS2 node is located close to the port in Country1 that exports maximum private-label volume to US.

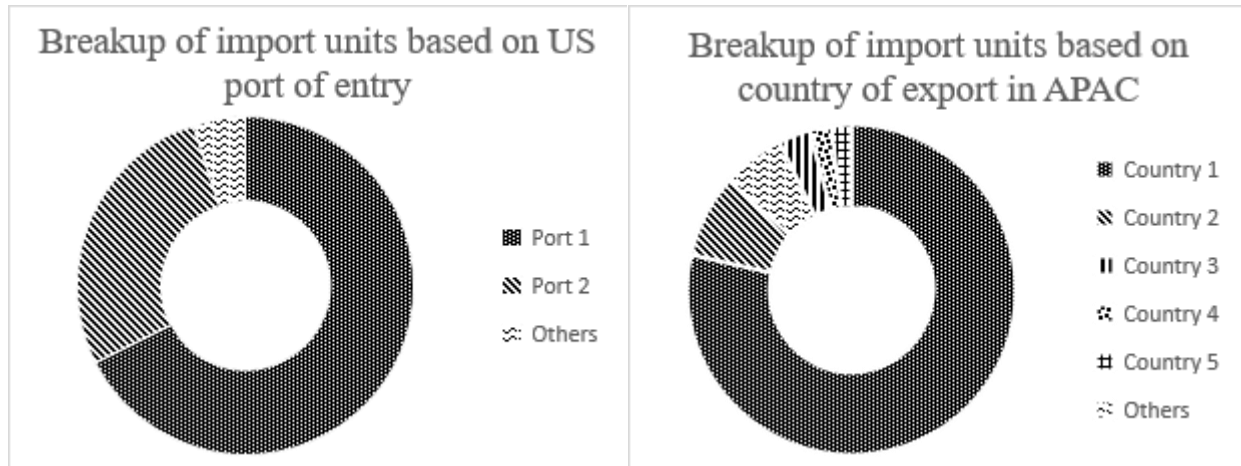


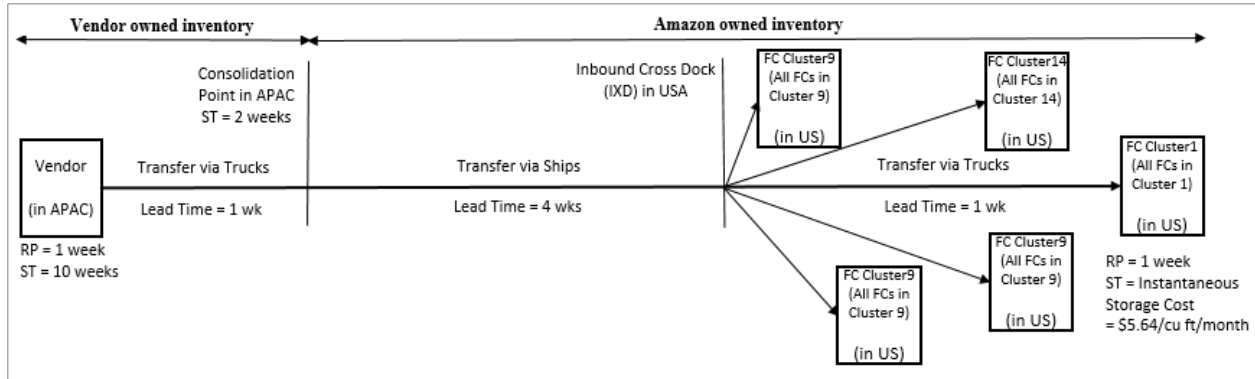
Figure 20 – Maximum private-label receipt in US (left) and maximum dispatch in APAC (right)

#### 5.4.2 Building SIP model using LCS1 and LCS2

We will use the SIP model of private-label import supply chain to compare the total costs of operations across four inventory storage scenarios: (1) inventory placed only at FCs (current scenario/ baseline), (2) inventory placed at FC + LCS1, but not at LCS2, (3) inventory placed at FC + LCS2, but not at LCS1, and (4) inventory placed at all nodes - FC + LCS1 + LCS2. (Figures below). In each scenario, we model the FCs at the cluster-level, i.e. placement of products in any FC of a cluster is acceptable to meet the target customer service level.

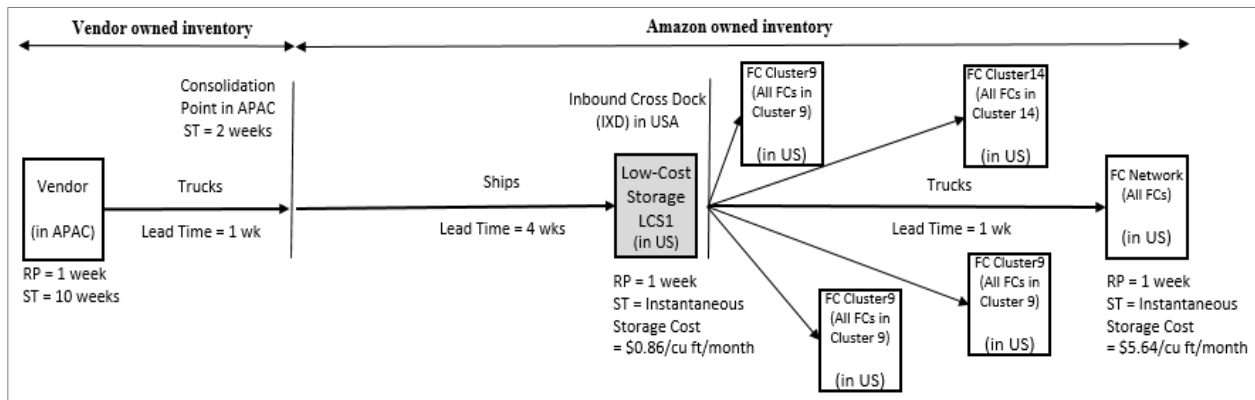
##### Scenario 1 – Inventory placed in FC clusters (Baseline)

This is the baseline case, where there is no storage node (LCS1 or LCS2) in the network. The delivery lead time between vendor and FCs is 18 weeks, consisting of the vendor lead time and review period, plus three legs of transportation. Each FC cluster has to hold necessary safety stock to meet demand fluctuations over 18 weeks of lead time.



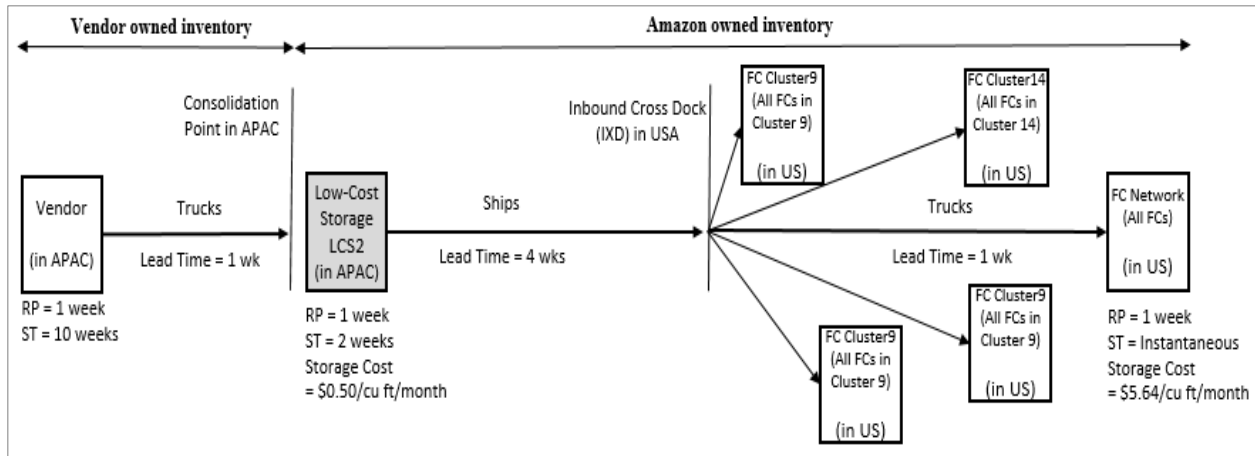
**Scenario 2 – Inventory placed in FC clusters and LCS1**

In this scenario, the study assumes that there is one low-cost storage node in the US (LCS1). The delivery lead time between LCS1 and FCs is 1 week. As a result, the safety stock stored in expensive FC storage bins is sized to meet demand fluctuations over 1 week, and is much less than for Scenario 1. It is expected to reduce the network inventory. In addition, the safety stock that LCS1 needs to store for the aggregated national demand (that it needs to store inventory for) benefits from risk pooling as discussed earlier.



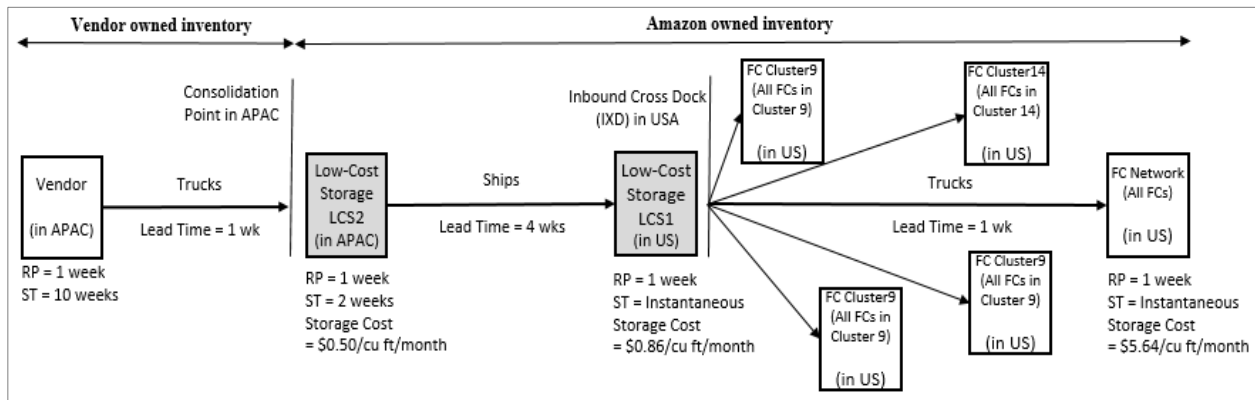
**Scenario 3 – Inventory placed in FC clusters and LCS2**

This scenario is very similar to scenario 2, where there is one low-cost node. However, this node is in APAC (LCS2). The delivery lead time between LCS2 and FCs is 7 weeks (2 weeks of LCS2 process time, 4 weeks of ocean travel time from APAC to the US, and 1 week of surface transport to FCs within the US). Due to the higher lead time (compared to scenario 2), the safety stock stored in expensive FC storage bins will be higher, increasing the network inventory. However, the per-unit cost of the storage space in LCS2 is lower than LCS1. This tradeoff between higher inventory and lower storage cost, when analyzed using SIP model, will determine the better option.



**Scenario 4 – Inventory placed in FC clusters, LCS1, and LCS2**

This is the most investment-intensive scenario, where there are two low-cost storage nodes, one in the US (LCS1) and other in APAC (LCS2). The delivery lead time between vendor and FCs is now broken into four parts – 1) 11 weeks of process time from vendor to LCS2, 2) 2 weeks of processing time at LCS2, 3) 4 weeks of transit time from LCS2 to LCS1, and 4) 1 week of delivery lead time from LCS1 to FCs. Both LCS1 and LCS2 add value to the network in unique ways. While LCS1 reduces the target inventory in the network, LCS2 provides the cheapest location for storage of inventory procured in excess of the optimal quantity needed for the network.



**5.4.3 Data Sources for estimating the model parameters**

The SIP model requires a large set of data points as inputs to the model. All relevant data that contribute to evaluating (1) the performance of different inventory strategies (scenarios) to match supply to demand, and (2) the total cost incurred in the process are required. In this study, the data was either received from periodically published metrics from different Amazon teams or estimated to the best accuracy possible

based on the tracked metrics. The estimated data was then validated by relevant Amazonian personnel. Table 3 highlights the high-level sources of the most important parameters used in the model.

Table 3 – List of data required as input parameters for the SIP model

Parameter	Why is it important?	How is it calculated/estimated?
Customer Demand	<p>-Demand profiles inform how much inventory needs to be kept on-hand</p> <p>-Weekly demand data matches the ordering cycle</p>	<p>Mean and Std. Deviation (SD) of weekly national demand for different SKUs calculated from actual weekly customer orders over a 20-month period</p> <p>Cluster level demand is estimated by breaking the aggregated national demand by regional demand factors that are based on historic data</p>
Target customer service level	<p>-Target customer service level also informs how much inventory needs should be kept on-hand, and thus, dictates the demand bound</p>	<p>For frequently ordered items (ordered in <math>\geq 80\%</math> of weeks), target service level is estimated at <math>(\text{Mean} + \text{Std. Dev.}/2)</math> of the weekly service level, capped at 0.99</p> <p>For less frequently ordered items (ordered in <math>&lt; 80\%</math> of weeks), target service level is estimated at Mean of the weekly service level</p> <p>Demand Bound = Mean Demand + <math>k \cdot (\text{SD of demand})</math>, where <math>k</math> is the inverse of the standard normal cumulative distribution function, evaluated at the target service level (calc. above)</p>
Lead Time (LT)	<p>Lead times affect the optimal inventory split between expensive FC bins and low-cost storage bins</p>	<p>Lead times are calculated for different legs of the supply chain starting from vendor in APAC to FC-clusters in US.</p> <ol style="list-style-type: none"> <li>1) <u>Manufacturing Lead Time</u>: PO date -&gt; Delivery by vendor in APAC; an average of 11 weeks is assumed.</li> <li>2) <u>Processing time at consolidation center</u>: Inbound time at the consolidation center in APAC -&gt; Ready to ship to US; Assumed 2 weeks</li> <li>3) <u>Ocean way LT</u>: Shipped from APAC -&gt; Clearance from customs at US port. Assumed 4 weeks</li> </ol>

		<p>4) <u>LCS1 processing time</u>: FC request for inventory -&gt; LCS1 Gate Out; Assumed 1 day in the model</p> <p>5) <u>LCS1 to FC LT</u>: Comparable cross dock -&gt; FC LT data is used to calculate the average LT to each FC-cluster from LCS1 site at Port1 (section 5.3.1)</p>
Operational costs	Costs affect the optimal inventory split between expensive FC bins and low-cost storage bins	<p>Costs are either directly obtained from suitable databases or estimated using first principles:</p> <ol style="list-style-type: none"> <li>1) Purchase cost: \$/unit paid to vendor <ul style="list-style-type: none"> <li>- Available from past Purchase Orders</li> </ul> </li> <li>2) LCS2 Processing Cost: \$/unit <ul style="list-style-type: none"> <li>- Benchmarked against similar operations in the system</li> </ul> </li> <li>3) Ocean-way shipment cost: \$/unit specific to an SKU <ul style="list-style-type: none"> <li>- Available from cost allocation of ocean freight</li> </ul> </li> <li>4) LCS1 Processing Cost: \$/unit <ul style="list-style-type: none"> <li>- From pilot operations of low-cost storage node</li> </ul> </li> <li>5) Inventory storage costs at all nodes: \$/cu-ft/month <ul style="list-style-type: none"> <li>- For FCs, available from internal database</li> <li>- For LCS1 and LCS2, estimated based on operations</li> </ul> </li> <li>6) Cross Dock-FC Freight: \$/cu-ft <ul style="list-style-type: none"> <li>- Average estimated for each cluster based on actual costs in the past</li> </ul> </li> <li>7) FC/Cross dock handling costs: \$/unit <ul style="list-style-type: none"> <li>- Available from internal database</li> </ul> </li> </ol>

**5.4.4 Parameters of a Typical SKU for comparing the scenarios in SIP model**

As evident from 5.4.3, the parameters for each SKU are unique, and depend on its demand pattern, size, and target customer service level. These differ widely for different SKUs. Feeding these parameters into the SIP model, we calculate the cost benefits of each alternate inventory storage strategy (scenario 2, 3, or 4) over the current system (baseline; scenario 1) by adding the benefits across all 763 SKUs in the study. The strategy with maximum cost benefit will be recommended for implementation.

However, it is important to check the robustness of the outcome of SIP model by carrying out sensitivity analysis on these input parameters to the model. Carrying out sensitivity analysis at an SKU-level is very

difficult, given the large number of SKUs. So, working with the product category team, we arrived at a basket of limited SKUs that was representative of the entire category. The weighted average SKU parameters of this representative basket is used to conduct the sensitivity analysis on the optimal solution. Table 4 lists the set of parameters for the representative imaginary product, referred to as *Typical SKU*.

*Table 4 - Parameters of a Typical private-label SKU received from the Category team*

Mean weekly demand (units)	3578
Standard Deviation of Weekly Demand (Units)	532
Target customer service level	0.95
Average Size (cubic feet / unit)	0.21
Purchase price (\$/unit)	13.47
Vendor Manufacturing Lead Time (Weeks)	11

#### **5.4.5 Working Formulae used in the SIP model**

The list of final working formulae used extensively to calculate the cycle, safety, and pipeline stocks for each storage node – FC, LCS1, LCS2 – of the SIP model in four scenarios are listed below.

$$1. \text{ Cycle Stock} = (\text{Average Weekly Demand} * \text{Review Period}) / 2$$

$$2. \text{ Safety Stock} = \text{Standard Deviation } (\sigma) * \sqrt{\text{Net Replenishment Time}} * \text{Safety Factor}$$

$$3. \text{ Pipeline Stock} = \text{Average Weekly Demand} * \text{Net Replenishment Time}$$

Cycle stock of a SKU for a given node varies with its review period. Safety stock of a SKU for a given node varies with its standard deviation of demand at that node and the net replenishment time of its order from its supplier node. These stocks for a node vary across scenarios as the network parameters change.

This supply chain has 18-weeks equivalent pipeline stock; however, pipeline stock is same for all the four scenarios since the overall lead time between vendor in APAC and FCs in US remains the same.

## 6. Results and Discussion

This chapter presents the results from the study, comparing the financial benefits in different scenarios. First, we highlight the best storage strategy based on the total supply chain costs for the 763 SKUs explored in the study. In order to understand the tradeoffs between alternate scenarios, we step away from aggregate data and analyze one SKU in depth. Specifically, we analyze the result for a *Typical* private-label SKU (explained in section 5.4.4), knowing that its result will be representative of the effects that each individual SKU (of this product category) will observe. We then, dive deep into the key drivers of financial benefits in the best storage strategy for the *Typical* SKU. Finally, we comment on the robustness of the solution by presenting the sensitivity analysis and highlight the limitations of the model.

### 6.1 Optimal Inventory Storage Strategy

Scenario 2, namely inventory storage in FC + LCS1 is the best storage strategy among the four scenarios, each of which maintains the same level of customer experience. This scenario incurs the minimum annualized supply chain costs. (Figure 21). Overall, the supply chain costs for the 763 SKUs in the study will be reduced by 51% in scenario 2 over the baseline (scenario 1).

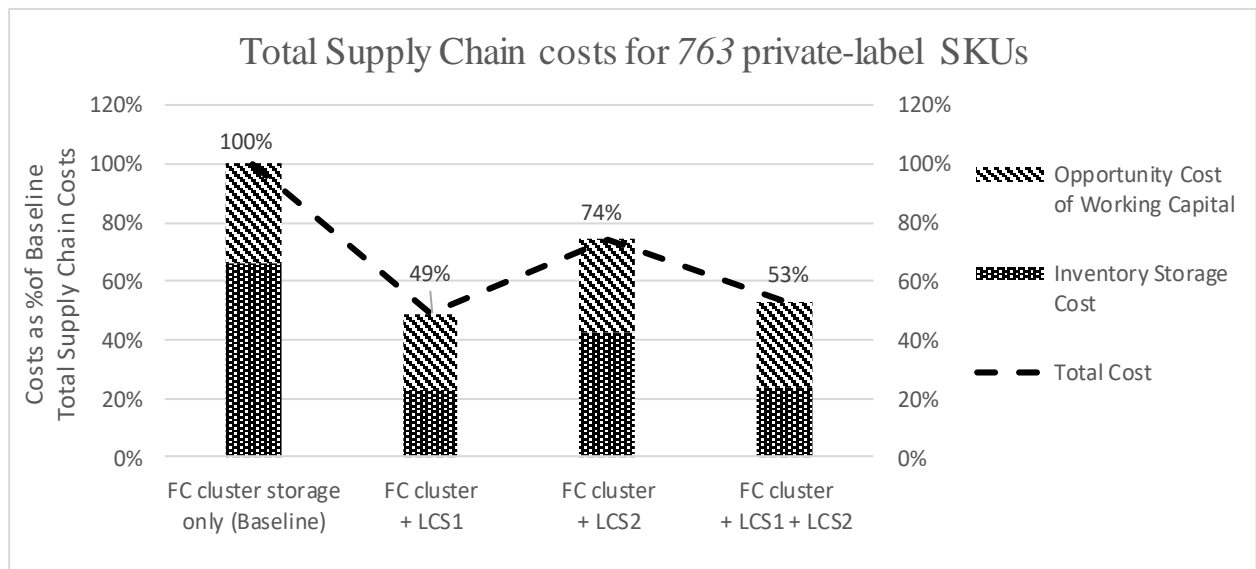


Figure 21 – Normalized supply chain costs, inventory storage costs, and opportunity cost of capital aggregated over 763-SKUs used in the study

The cost benefit for each individual SKU will vary based on its unique parameters (input to the SIP model). For instance, the corresponding normalized total costs for the *Typical* SKU across the scenarios are 62% (scenario 2), 85% (scenario 3), and 69% (scenario 4) of the total absolute costs incurred in the current network (baseline scenario 1).

Since opportunity cost of capital is a function of the unit purchase price of the SKU, its relative contribution to the total supply chain cost will be SKU-dependent. However, the differences in inventory storage costs does not depend on the SKU price, but rather depends on 1) network inventory units and 2) per-unit storage costs. Thus, hereafter, we carry out an in-depth analysis of the total inventory storage costs alone to uncover the features of the supply network that drive such large cost differences among the four scenarios.

Figure 22 shows the differences in the annualized inventory storage costs for the *Typical SKU*. Clearly, the total storage cost strongly correlates with the total inventory in the network, and particularly the inventory stored in the most expensive storage bins at FCs. Table 5 lists the normalized inventory storage needs, both the target network inventory and the inventory at FCs. It also highlights the average per-unit storage cost, calculated by dividing total storage cost by total inventory units in the network, to quantify the impact of each storage scenario using a common metric.

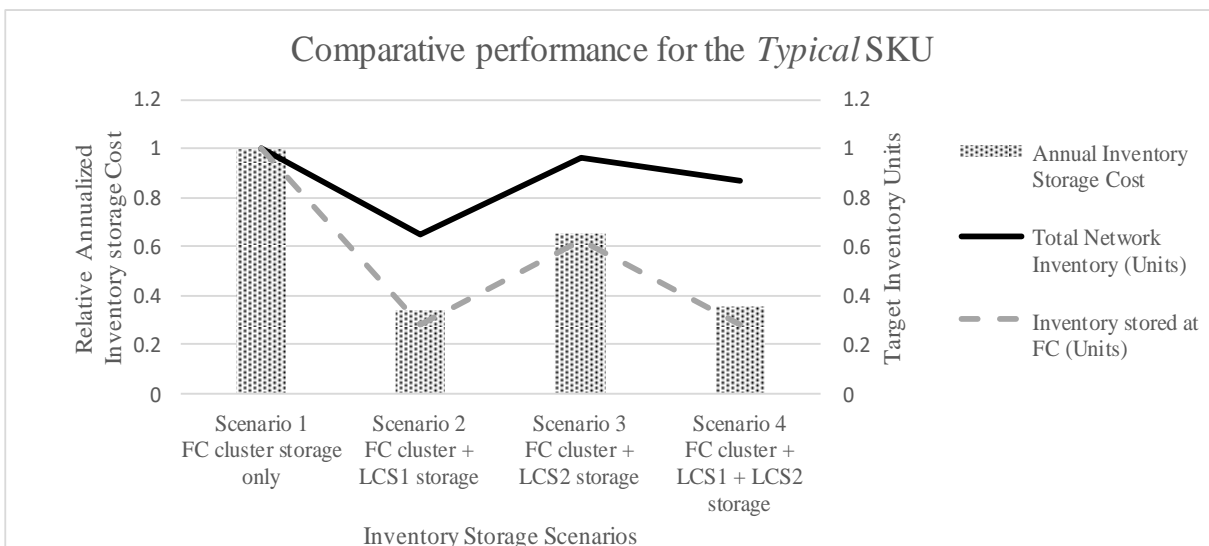


Figure 22 – Comparative performance of four storage scenarios (in SIP model) on the Typical SKU

Table 5 – Normalized inventory storage details for Typical SKU in four scenarios

Scenario	Inventory Storage Strategy	Annual Total Inventory Storage Cost (\$)	Network Inventory (Units)	Inventory in FCs (Units)	Per-unit storage (\$/unit)
1	FC cluster (Baseline)	100%	100%	100%	100%
2	FC cluster + LCS1	34%	65%	28%	52%
3	FC cluster + LCS2	65%	96%	62%	68%
4	FC cluster + LCS1 + LCS2	35%	87%	28%	40%

### 6.1.1 Effect on the target network inventory

With the minimum delivery lead time to FCs (1 week) among all supplying nodes, LCS1 can meet demand fluctuations from FCs quickly and reliably, thereby reducing the need for expensive FC inventory to 28% of baseline and releasing 72% of previously occupied FC storage space.

Comparing scenario 2 with scenario 3 (Table 5, or Figure 23), we find a difference of 34% (62% vs 28%) in FC-inventory requirements, confirming that the shorter delivery lead time from supplying node to FCs is a key differentiator. Due to this reason, the total network inventory in scenario 3 is 96% of baseline inventory, suggesting minimal benefits to the network from following FC + LCS2 storage strategy.

Additionally, the ‘pooling effect’ due to aggregation of demand over the vendor lead time is greater in scenario 2 than in scenario 4, leading to 22% lower total network inventory in scenario 2 (65% of baseline inventory in scenario 2 vs 87% in scenario 4). It must be noted, however, that despite a higher inventory unit in scenario 4, there is a lower inventory storage cost at LCS2 and this closes the gap in total storage costs between the two scenarios to a mere 1% (34% in scenario 2 vs 35% in scenario 4).

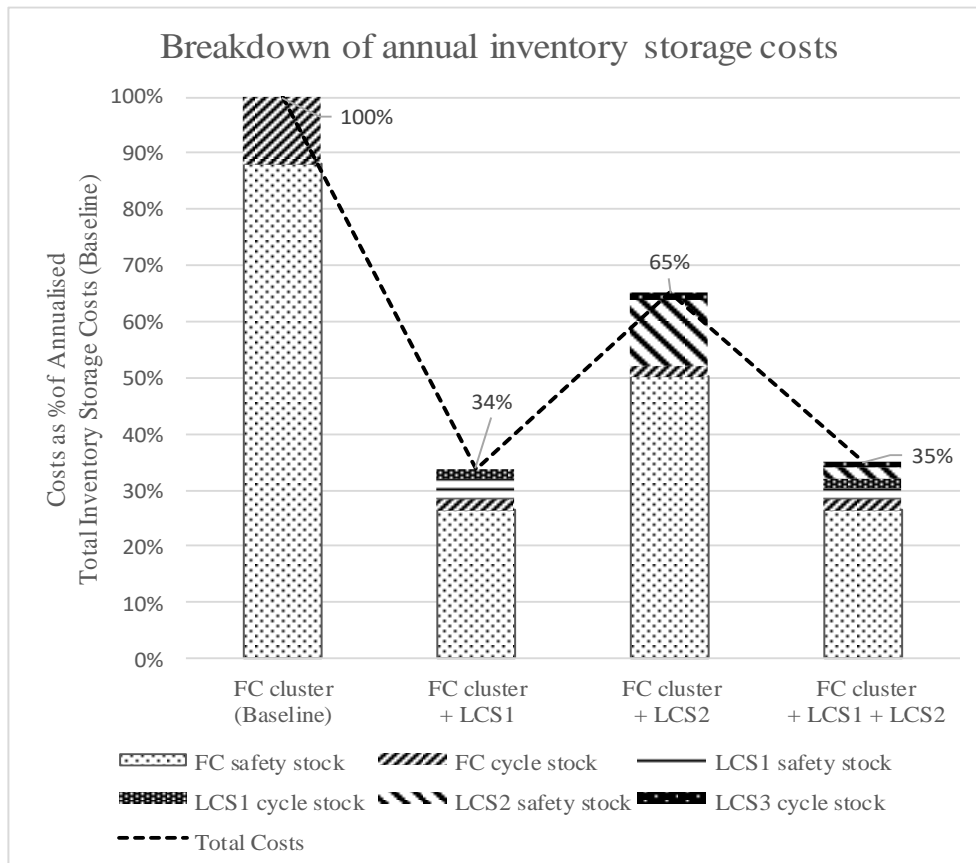


Figure 23 – Breakdown of storage costs contributed by safety stock and cycle stock at different nodes across four scenarios

### 6.1.2 Effect on the per-unit cost of inventory storage

The second factor determining the inventory storage cost, after total inventory in the system, is the per-unit storage cost. It represents how well inventory is distributed across high-cost and low-cost storage nodes in the network, such that total cost is minimized without losing on customer service level.

Mathematically, it is calculated by dividing the total annualized storage cost for all inventory of an SKU across the network by total inventory units of that SKU in the network.

The introduction of low-cost storage nodes in the network reduces the overall per-unit storage cost, as seen in Table 5. Relative to the per-unit storage cost in baseline scenario, the per-unit storage cost for the *Typical* SKU in scenario 2 is 52% of the baseline, and the per-unit storage cost in scenario 4 is 40% of the baseline. This result is particularly relevant for SKUs that have very large Minimum Order Quantity (MoQ) of purchase, when any quantity bought in excess of what is required at FCs and LCS1 to meet the service level must be stored in LCS2.

In either scenario 2 or scenario 4, we see that introduction of low-cost storage node will save private-label portfolio a large portion of its current (baseline) storage costs, which can be passed on to customers.

### 6.2 Comparison between the current FC-only storage and FC + LCS1 storage

Next, we evaluate the item-wise breakup of inventory units and storage costs in FCs and LCS1 between scenario 1 and scenario 2 to identify the key drivers of cost reduction in scenario 2. In line with the expected benefits from multi-echelon inventory storage, FC + LCS1 (1) significantly reduces the safety stock in the network by 37% (from 88% to 51%) while (2) marginally increasing the cycle stock in the network by 2% (from 12% to 14%). (Figure 24).

Normalized inventory unit and storage cost (FC + LCS1 storage vs FC-only storage)					
	Component	Inventory Units (%of baseline)		Inventory Storage Cost (%of baseline)	
		FC storage	FC + LCS1 storage	FC storage	FC + LCS1 storage
Node-wise Contributors	FC safety stock	88%	27%	88%	27%
	FC cycle stock	12%	2%	12%	2%
	LCS1 safety stock		25%		4%
	LCS1 cycle stock		12%		2%
Aggregated across network	Safety stock	88%	51%	88%	30%
	Cycle stock	12%	14%	12%	4%
	<b>Total</b>	<b>100%</b>	<b>65%</b>	<b>100%</b>	<b>34%</b>

Figure 24 – Key drivers for reduction in network inventory and storage cost in FC + LCS1 storage

Additionally, the larger portion of the inventory is now stored at an inexpensive LCS1 storage facility, reducing total storage cost by almost 2x the reduction in total inventory (66% reduction in total cost vs 35% reduction in total inventory). Figure 25 compares the item-wise storage cost between the two scenarios.

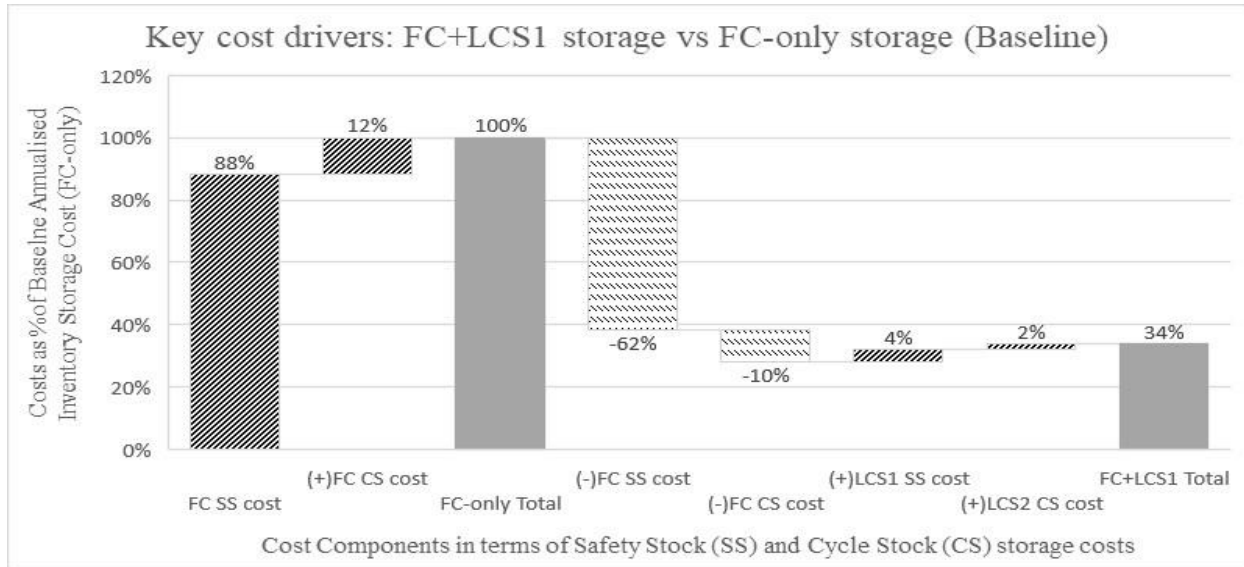


Figure 25 – Key drivers for reduction in inventory storage costs in FC + LCS1 storage strategy

### 6.3 Sensitivity Analysis

The output of the SIP model for *Typical* SKU is sensitive to the values of the input parameters – SKU size (cubic feet/unit), purchase price (\$/unit), demand pattern (mean and standard deviation of weekly demand), replenishment lead-time, and holding cost rate. However, we found that the superiority of scenario 2 (FC + LCS1 storage) relative to the other scenarios was very robust on all occasions.

#### Sensitivity to varying SKU size

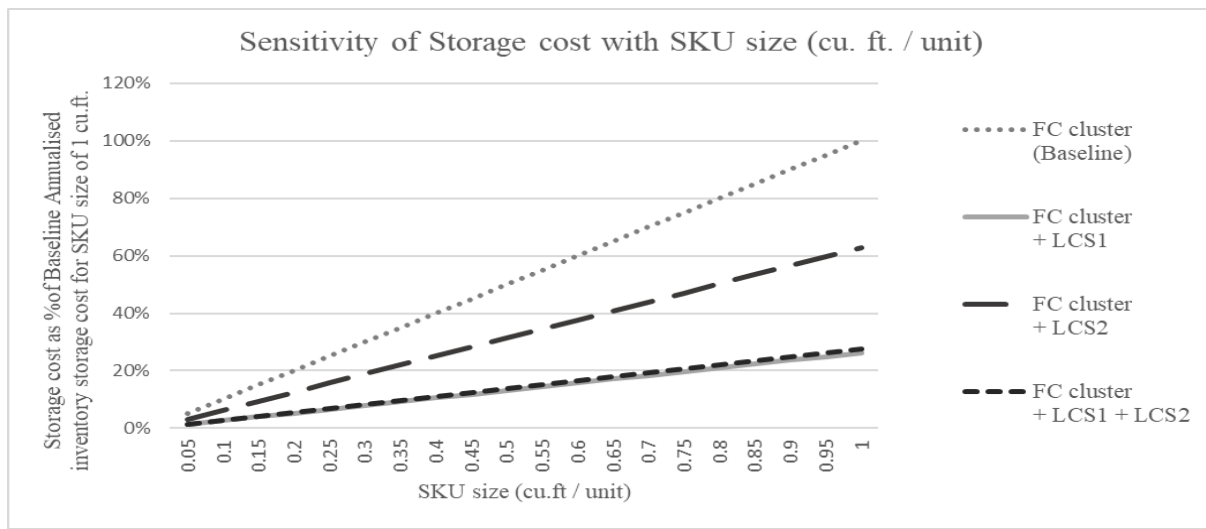


Figure 26 – Sensitivity analysis of the inventory storage cost with SKU size

Sensitivity to varying LCS1 – holding cost rate vs LCS2 – holding cost rate

At this point, we have established that LCS1 is the preferred location for building low-cost storage node; however, we have not completely negated the potential benefits of LCS2 (at least for SKUs with high MoQ). The following sensitivity analysis, studying the effect on optimality by simultaneously varying holding cost rates at LCS1 and LCS2, will explore this aspect. This analysis would reveal the necessary operational metrics that need to be true for LCS2 to become part of the best solution (i.e. scenario 3 or 4).

Figure 27 shows the ‘best storage option’ for different combinations of LCS1 and LCS2 holding cost rates. At current normalized rates (marked in yellow in both axes), scenario 2 is the best solution. The best scenario of FC + LCS1 does not change unless a) the holding cost rate at LCS1 increases to 142% of its current value (42% worse) and b) the holding cost rate at LCS2 reduces to 68% of its current value (32% improvement), both of which seem realistically unlikely. This further proves the robustness of scenario 2 as the best solution.

		LCS2 Holding Cost Rate (\$/cu-ft/month)														
		44%	52%	60%	68%	76%	84%	92%	100%	108%	116%	124%	132%	140%	148%	156%
LCS1 Holding Cost Rate (\$/cu-ft/month)	53%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	58%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	63%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	67%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	72%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	77%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	81%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	86%	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	91%	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	95%	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	100%	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	105%	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	109%	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2
	114%	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2
	119%	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2
	123%	4	4	4	2	2	2	2	2	2	2	2	2	2	2	2
	128%	4	4	4	2	2	2	2	2	2	2	2	2	2	2	2
	133%	4	4	4	2	2	2	2	2	2	2	2	2	2	2	2
	137%	4	4	4	2	2	2	2	2	2	2	2	2	2	2	2
	142%	4	4	4	4	2	2	2	2	2	2	2	2	2	2	2
147%	4	4	4	4	2	2	2	2	2	2	2	2	2	2	2	

Legend	Scenario 2 Storage	Legend	Scenario 4 Storage
2	FC cluster + LCS1	4	FC cluster + LCS1 + LCS2

Figure 27 – Sensitivity analysis of the best solution as holding cost rate in low cost nodes change

**6.4 Limitations of the model**

As with any modeling, some of the real-world complexities have been simplified in this analysis. These deviations listed below should be taken into consideration before translating this model into production.

- The model assumes stationary demand (constant mean and standard deviation) that is based on historic customer demand. During implementation, the data from the forecasting team should be used to calculate demand.
- SKU-specific critical ratio should be updated as an input to the model as and when unit economics for each SKU changes.
- The model assumes deterministic lead-time, as variations in the lead time were primarily due to the staggered delivery agreement.
- This model takes an unconstrained view of the supply chain and optimizes storage for a single SKU at a single point in time. However, real-world constraints such as freight container space, multi-SKU packaging, storage space at FCs/LCS1, loading/ unloading infrastructure capacity at different buildings etc. would require optimization across multiple SKUs.
- The assumption on ‘committed service time’ for APAC-LCS1 and LCS1-FC transfers seems achievable as these transfers are company-owned; however, the behavioral shift among supply chain teams to ensure strict compliance to these metrics is necessary for the system to deliver results.
- In unforeseen circumstances when customer demand spikes (example – celebrity endorsing a product, viral videos of specific product, results from the world of sports etc.) beyond the expected demand bound, manual interventions such as expedition of orders need to kick in to match supply and demand.
- The scope of this analysis included US-only private-label products, for which LCS2 storage does not seem beneficial due to the long lead times between LCS2 and the US. The benefits are marginal and only after realizing the bulk of the benefits from LCS1 storage. However, for products that are truly global or that have extremely large minimum order quantities, LCS2 storage could provide potential benefits by risk pooling value from facility in APAC.

## **7. Recommendations and Future Work**

The application of strategic inventory placement model in this study establishes the benefits that e-tailers can derive from adopting multi-echelon inventory optimization processes in their supply chains. This chapter provides a summary of the principal findings and recommends an implementation roadmap for Amazon. This suggested action plan is generalizable to other e-tailers as well.

### **7.1 Final recommendation**

The inventory model described in this thesis provides a quantified justification for adopting FC + LCS1 storage as the best network storage strategy for the selected product category - imported, small-sized, Business Line 1, private-label SKUs (section 3.2).

Based on the analysis of the *Typical SKU* used in the study, FC + LCS1 storage offers significant savings potential derived from (a) reduction in FC safety stocks by 70% from baseline (from 88% in scenario 1 to 27% in scenario 2) due to reduced replenishment lead time from LCS1 to FCs and reduced order review period at FCs, (b) pooling of demand variability across FCs, thereby reducing system-wide inventory requirements by 35% from baseline (from 100% in scenario 1 to 65% in scenario 2), and (c) lowering of network-wide average unit storage cost by 48% (from 100% in scenario 1 to 52% in scenario 2). These improvements can be realized by each SKU in this product category, but the specific numerical benefit will depend on each SKU's unique parameters.

LCS1 storage can be extended to all imported, slow-moving SKUs, preferably with high demand variability. The biggest benefits of this inventory storage strategy comes from 1) lower replenishment lead time to FC clusters, leading to lower safety stocks occupying expensive FC bins (relevant for imported SKUs with large lead time), 2) risk pooling over variability of demand in individual FC clusters, leading to lower safety stocks (relevant for SKUs with large swing in weekly demand), and 3) lower per-cuft cost of storage for excess system inventory (relevant for SKUs with large MoQ).

As seen from the model formulation, the inventory model does not make strict assumptions about SKU-specific characteristics. Thus, over time, as the operations team builds up infrastructural capabilities to handle SKUs of different characteristics, this model can benefit a larger proportion of Amazon's selection.

### **7.2 Implementation Roadmap**

This study proves that the introduction of low-cost nodes should immensely benefit private-label brands. To fast-track the implementation, a cross-functional supply chain team should work on two broad

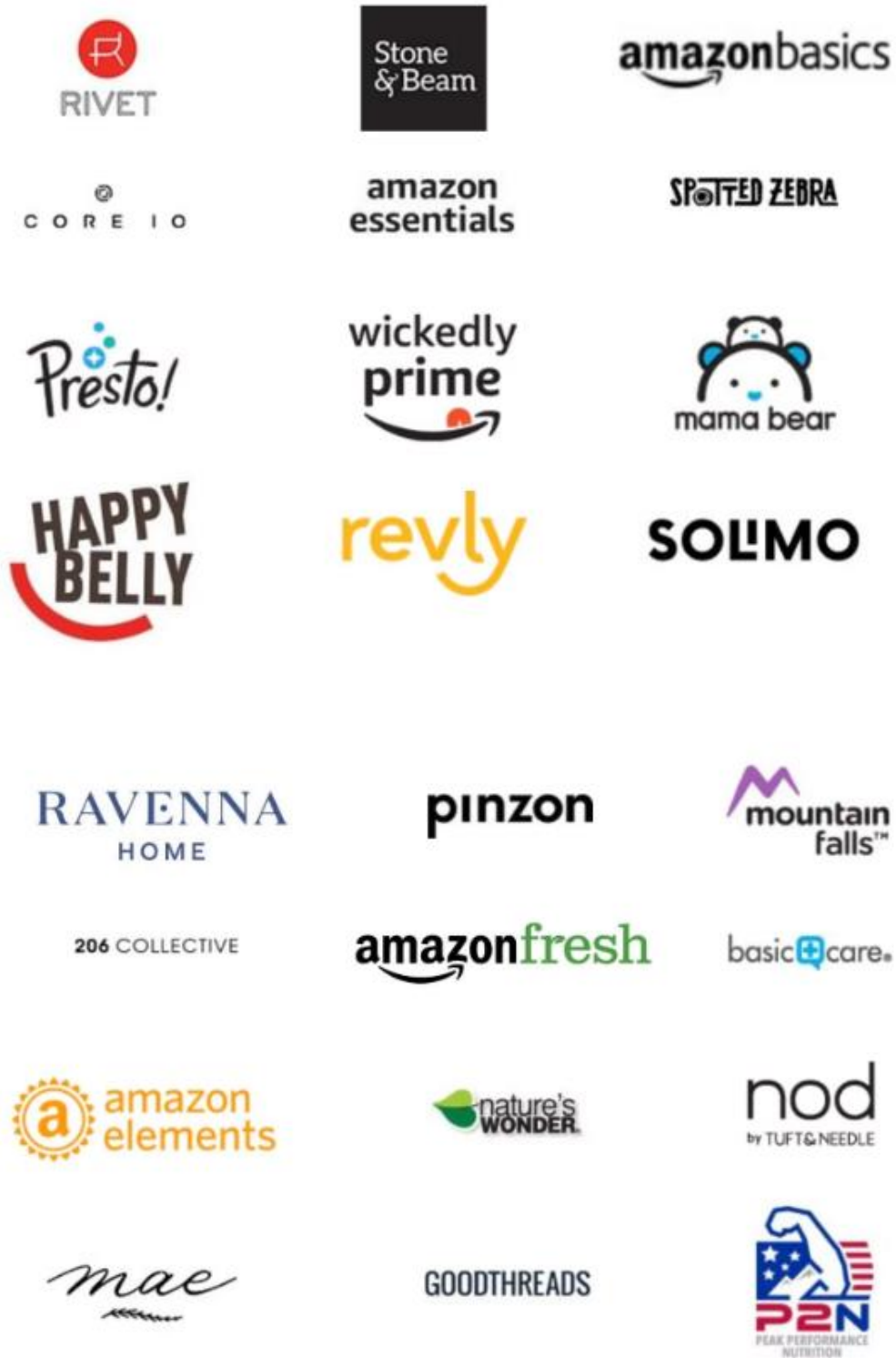
workstreams. First, the research scientists and operations teams should include real-world complexities into this model to arrive at an algorithm that can be directly implemented in the Amazon system. Second, a pilot study on a subset of private-label SKUs at the first LCS1 site should be carried out to test the model and validate the proposed financial entitlement. After completion of these validation steps, the model can be extended to various product categories as and when Amazon expands its storage and handling capabilities. Some of the high-level workstreams are listed below.

1. The supply chain research team to incorporate real-world constraints into the model, and update the LCS1-eligibility criteria for private-label SKUs
2. The Category team to analyze other product categories for inclusion in this model and revise the total entitlement
3. External Fulfillment team to pilot the model on an identified subset of private-label SKUs in LCS1 and monitor the savings vs current FC-storage
4. The network topology team should identify best places to build LCS1(s) so as to minimize the weighted average ship distance between LCS1 nodes and FCs (to reduce shipment costs) and aid the spread of products across FC clusters
5. The model assumes one LCS1 site as of now. As the need for low-cost storage space grows and multiple LCS1 nodes are built, operations research scientists must optimize inbound product placement for LCS1 network.

## References

- [1] B. Stone, *The Everything Store*. Back Bay Books. 2013
- [2] M. Roser, H. Ritchie, and E. Ortiz-Ospina, “Our World in Data,” *Our World in Data*, Jul. 2015.
- [3] *Jeff Bezos - Charlie Rose Interview*. 2010.
- [4] J. Wells, G. Danskin, and G. Ellsworth, “Amazon.com,2018.” Harvard Business School Publishing, 24-May-2018.
- [5] “Amazon.com, Inc. Form 10-K, the year ended December 31, 2018.” [Online]. Available: <https://www.sec.gov/Archives/edgar/data/1018724/000101872419000004/amzn-20181231x10k.htm>. [Accessed: 15-Feb-2020].
- [6] “Amazon.com, Inc (NasdaqGS:AMZN).” S&P Global Market Intelligence, 15-Feb-2020.
- [7] A. Levy, “Amazon just passed Alphabet to become the world’s second most valuable company,” *CNBC*, 20-Mar-2018. [Online]. Available: <https://www.cnbc.com/2018/03/20/amazon-just-passed-alphabet-to-become-the-worlds-second-most-valuable-company.html>. [Accessed: 16-Feb-2020].
- [8] “Leading U.S. Amazon product categories 2019,” *Statista*. [Online]. Available: <https://www.statista.com/statistics/639155/popular-amazoncom-sales-by-category/>. [Accessed: 16-Feb-2020].
- [9] “The Rise of Amazon’s Private Label Brands,” *Seattle Business Magazine*, 04-Jan-2018. [Online]. Available: <https://seattlebusinessmag.com/business-operations/rise-amazons-private-label-brands>. [Accessed: 16-Feb-2020].
- [10] N. Kumar, and Jan-Benedict Steenkamp, *Private Label Strategy*. Harvard Business School Press. 2007
- [11] “Amazon Private Label Brands,” *Marketplace Pulse*, 18-Mar-2019. [Online]. Available: <https://www.marketplacepulse.com/amazon-private-label-brands>. [Accessed: 16-Feb-2020].
- [12] “The Rise of Amazon’s Private Label Brands,” *Seattle Business Magazine*, 04-Jan-2018. [Online]. Available: <https://seattlebusinessmag.com/business-operations/rise-amazons-private-label-brands>. [Accessed: 16-Feb-2020].
- [13] R. S. Kaplan, “Introduction to Activity-Based Costing,” p. 14.
- [14] “Supply Chain Components for Private Label Success,” *Logility*, 2017.
- [15] A. S. Eruguz, E. Sahin, Z. Jemai, and Y. Dallery, “A comprehensive survey of guaranteed-service models for multi-echelon inventory optimization | Elsevier Enhanced Reader,” *International Journal Production Economics*, vol. 172, pp. 110–125, Feb. 2016, doi: 10.1016/j.ijpe.2015.11.017.
- [16] S. Graves and S. Willems, “Supply Chain Design: Safety Stock Placement and Supply Chain Configuration,” in *Handbooks in OR & MS*, vol. 11, 2003, pp. 95–132.
- [17] S. C. Graves and S. P. Willems, “Optimizing Strategic Safety Stock Placement in Supply Chains,” *M&SOM*, vol. 2, no. 1, pp. 68–83, Jan. 2000, doi: 10.1287/msom.2.1.68.23267.
- [18] K. F. Simpson, “In-Process Inventories,” *Operations Research*, vol. 6, no. 6, pp. 863–873, Dec. 1958, doi: <https://doi.org/10.1287/opre.6.6.863>.

Appendix 1 – A few common Amazon Private-label Brands



Source: <https://www.marketplacepulse.com/amazon-private-label-brands>

**Appendix 2 - Representative mapping of US into 17 zones (each with one FC cluster)**

