United States Department of Agriculture

Feasibility Study of Capturing Supplemental Nutrition Assistance Program (SNAP) Purchases at the Point of Sale

Final Report

Nutrition Assistance Program Report **November 2016**
Food and Nutrition Service
Office of Policy Support

United States Department of Agriculture

Food and Nutrition Service, Office of Policy Support *November 2016*

Feasibility Study of Capturing Supplemental Nutrition Assistance Program (SNAP) Purchases at the Point of Sale
Final Report

Authors:
Steven Garasky
Slava Katz
Kassim Mbwana
Samuel Ampaabeng
Joseph Llobrerra

Submitted by:
IMPAQ International, LLC
10420 Little Patuxent Parkway, Suite 310
Columbia, MD 21044
www.impaqint.com

Project Director:
Steven Garasky

Project Manager:
Slava Katz

Submitted to:
Office of Policy Support
Food and Nutrition Service
3101 Park Center Drive
Alexandria, VA 22302-1500

Project Officer:
Barbara Murphy

This study was conducted under Contract number AG-3198-C-12-0013 with the Food and Nutrition Service, United States Department of Agriculture.

This report is available on the Food and Nutrition Service website:

http://www.fns.usda.gov/research-and-analysis

Suggested Citation:

Garasky, Steve, Slava Katz, Kassim Mbwana, Samuel Ampaabeng, Joseph Llobrerra. *Feasibility Study of Capturing Supplemental Nutrition Assistance Program (SNAP) Purchases at the Point of Sale – Final Report*. Prepared by IMPAQ International, LLC for the U.S. Department of Agriculture, Food and Nutrition Service, November 2016.

Acknowledgments

A number of individuals contributed to the successful completion of this report. Barbara Murphy of the Food and Nutrition Service (FNS) Office of Policy Support (OPS) served as our Contracting Officer's Representative and helped ensure that our work proceeded in a timely fashion, facilitated our communication with staff at FNS, and generally helped guide the development of this report. At OPS, Kathryn Law and Kelly Kinnison (currently at the Health and Human Services); at Special Nutrition Programs, Erin McBride; and at Retailer Policy and Management Division, Kathy Ottobre and Andrea Gold, provided timely feedback and guidance to help us understand relevant issues and refine our research. We are also grateful to the Office of Information Technology for helping us understand some of the technical issues pertinent to FNS. Within the IMPAQ team, our colleagues Vadi Ranganathan, Beatrice Brogan, and Sooraz Mantena at Ventera; and Len Fuller and Glen Glasscock at Truno (formerly Cash Register Services Inc.) all made important contributions to the research. IMPAQ's research support staff helped prepare the final document. Any remaining errors are our own.

Table of Contents

Table of Exhibits

Glossary

ACS	Advanced Checkout Solution
ALERT	Anti-Fraud Locator Using Electronic Benefits Transfer Retailer Transactions
ANSI	American National Standards Institute
BIN	Bank Identification Number
CES	Consumer Expenditure Survey
CPCM	Cost Per Case Month
EAN-13	European Article Number
EBT	Electronic Benefits Transfer
ECR	Electronic Cash Register
ERS	Economic Research Service
FNS	Food and Nutrition Service
GMT	Greenwich Mean Time
ICC	Integrated Circuit Card
IECR	Integrated Electronic Cash Register
ISO	International Organization for Standardization
MAC	Message Authentication Code
NCR	NCR Corporation
NITC	National Information Technology Center
OIT	Office of Information Technology
OPS	Office of Policy Support
ORA	Office of Research and Analysis
PAN	Primary Account Number
PCISSC	Payment Card Industry Security Standards Council
PIN	Personal Identification Number
PLU	Price Look-up Code
POC	Proof of Concept
POS	Point of Sale
RPMD	Retailer Policy and Management Division
SFPD	Supplemental Food Programs Division
SFTP	Secure File Transfer Protocol
SNAP	Supplemental Nutrition Assistance Program
STARS	Store Tracking and Redemption System
TPP	Third-Party Processor
TWILD	Transactions with Item-Level Data
UPC	Universal Product Code
USDA	United States Department of Agriculture
WIC	Special Supplemental Nutrition Program for Women, Infants, and Children

Executive Summary

This study explored the feasibility of creating a data collection system capable of providing item-level data to the U.S. Department of Agriculture (USDA) Food and Nutrition Service (FNS) about food purchases made by Supplemental Nutrition Assistance Program (SNAP) households. The data would be captured at the point of sale (POS) for purchases made by SNAP participants using their Electronic Benefit Transfer (EBT) cards. Several technical solutions were identified to capture the data at the point of sale. This report presents findings from a proof of concept (POC) and cost model assessing the technical feasibility and estimated cost of potential technical solutions for capturing item-level data on SNAP EBT purchases.

1. BACKGROUND

To support its mission of providing access to food and healthful diets for low-income families, FNS needs a variety of data that describe diet and food purchasing patterns of SNAP households. However, the legislation authorizing SNAP, the Food and Nutrition Act of 2008, does not currently require retailers to collect or transmit item-level transaction data to FNS. As a result, FNS must rely on consumer-reported data, transaction data from the Anti-Fraud Locator for EBT Redemption Transaction (ALERT) system, and data from retailers, third-party companies, and other sources. These sources present a variety of challenges, including data quality and reliability, that limit their usefulness. In consequence, FNS is investigating the feasibility of creating a data collection system to automatically deliver, on an ongoing basis, item-level data on purchases made by SNAP households.

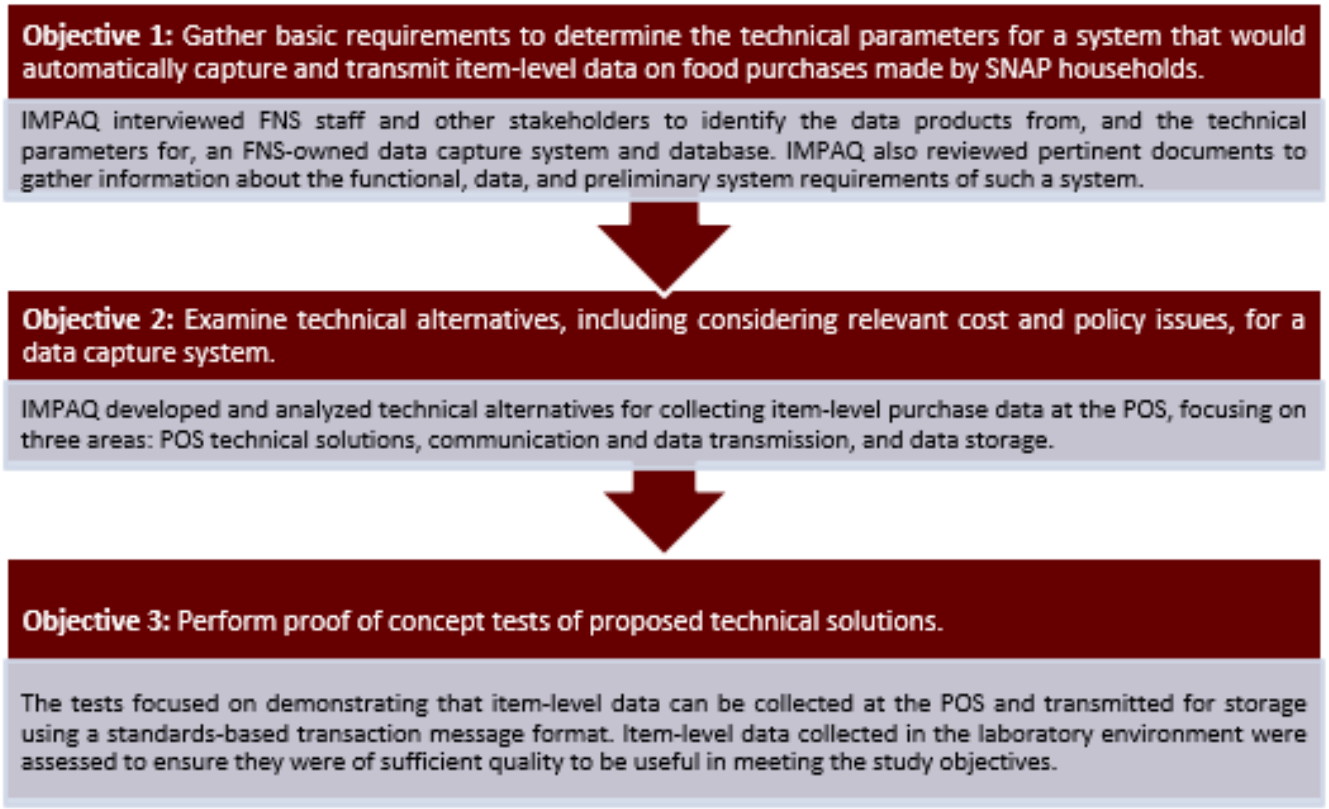
2. STUDY DESIGN AND METHODOLOGY

Exhibit E.1 outlines the three primary objectives of this study and their associated activities. The first phase of this study focused on Objectives 1 and 2. During this phase, IMPAQ identified the desired data products resulting from, and technical parameters for, an FNS-owned data capture system and database for storing SNAP purchase data. Using that information, the project team then developed and analyzed an array of technical alternatives for collecting item-level purchase data at the POS. After carefully weighing the advantages and disadvantages of each potential solution, FNS and IMPAQ identified four technical solution options as the most

promising for further examination. Each of these solutions would operate in a store environment where retailers use Integrated Electronic Cash Register (IECR) systems.[1]

This report focuses on Objective 3, performing limited POC testing on those technical solutions. IMPAQ designed the testing to demonstrate that useful data of sufficient quality can be collected using the proposed technical solutions. This testing also served to inform an overall assessment of technical feasibility and potential costs of the four identified technical solutions.

Exhibit E.1. Research Objectives and Associated Tasks



[1] Garasky, S., S. Katz, K. Mbwana, S. Aampaabeng, Z. Miller, M. Roy. "Feasibility Study of Capturing Supplemental Nutrition Assistance Programs (SNAP) Purchases at the POS – Final Technical Solutions Report". Prepared for the U.S. Department of Agriculture, FNS – Office of Policy Support. May 2014.

The POC needed to have three characteristics to produce maximally useful results. It needed to:

- **Be based on real-world retail data** to give insight into the challenges retailers would face in extracting item-level data on SNAP purchases. IMPAQ acquired actual POS data from large national retailers and regional retailers.

- **Be implemented in a real-world retail environment,** using actual SNAP EBT cards in a shopping lane to complete real transactions. Because project time constraints prevented full achievement of this goal, IMPAQ conducted the POC in a laboratory environment, simulating data capture in register lanes and transmission to EBT processors.

- **Address multiple POS software and hardware configurations.** The five major POS hardware providers support a large number of POS software applications. Testing all software and hardware configurations was beyond the scope of this POC. To approach generalizability as much as possible, IMPAQ ran the POC on a Toshiba/IBM configuration and tested data captured by *QuickBooks*™ POS software configurations. The IBM/Toshiba configuration is among the most widely used in the industry, while *QuickBooks*™ is one of the likely configurations used by smaller low-volume retailers.

3. PROPOSED TECHNICAL SOLUTIONS AND PROOF OF CONCEPT IMPLEMENTATION

IMPAQ's researchers addressed four technical solutions during the POC:

1. Technical Solution 1: Existing transmission infrastructure with real-time data transmission
2. Technical Solution 2: Existing infrastructure with batch transmission
3. Technical Solution 3: New transmission infrastructure using electronic messaging
4. Technical Solution 4: New transmission infrastructure using batch transmission

The four proposed technical solutions presented four key system design considerations:

1. **Transmission paths:** Using new routes ***versus*** existing routes through third-party processors and EBT processors
2. **Transmission frequency:** Transmitting data in real time (via electronic messaging) ***versus*** in batch mode

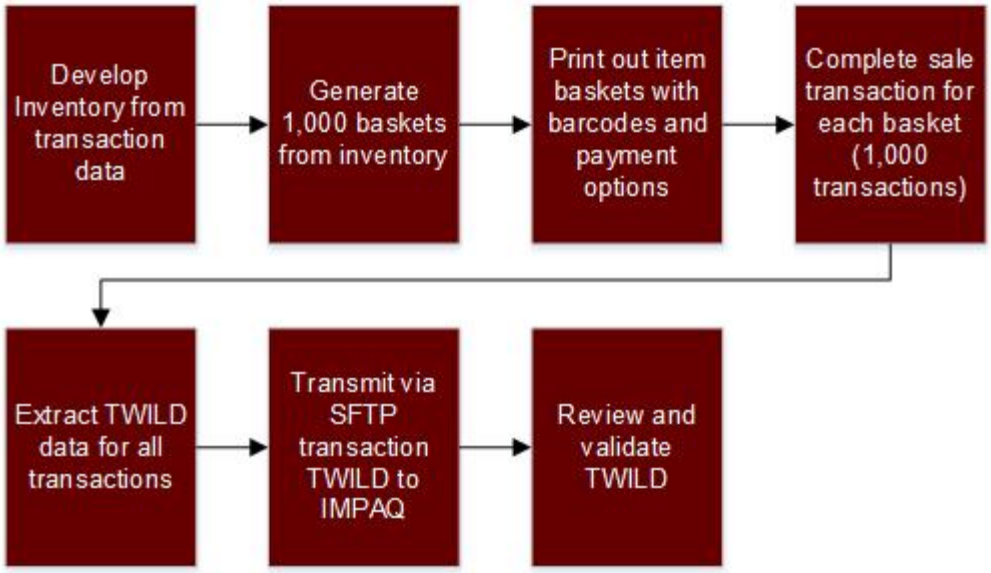
3. **Encryption and Masking:** while transaction PIN numbers are encrypted, the middle digits of 16 or 19 digit EBT card numbers are masked. Only the first six and last four digits are typically displayed.
4. **Storage:** Processing and storing data on FNS computers, ***versus*** on contractor computers, ***versus*** in cloud storage owned by FNS or a contractor

Proof of Concept Laboratory Testing

The POC focused on implementing the proposed technical solutions in an IECR environment, because IECR transactions represent more than 80 percent of all EBT redemptions. IECR systems are technologically sophisticated enough to capture SNAP item-level data, and can leverage existing data transmission infrastructures from retailers to processors and then to FNS.

The study team designed and tested a POC to assess the four key system design considerations. As shown in Exhibit E.2, the POC simulated transactions from use of an EBT card in a retail outlet, through capture of item-level data, to transmission of those data and storage in a national database. Important to note is that the POC was generic enough to facilitate understanding of the four design considerations without requiring separate simulations for each proposed technical solution.

Exhibit E.2. POC Testing Laboratory Process



Proof of Concept Analyses

The study team concluded the POC by verifying the integrity of actual item-level data on SNAP EBT transactions and matching those data to corresponding transactions in the ALERT data.

The transaction data verification determined the completeness and usefulness of the item descriptions contained in the item-level data. All of the retailer data included a Universal Product Code (UPC) making identification with a unique identifier possible. However, few of the retailer UPCs were available in national UPC databases. Similarly, all of the retailer data included item descriptions, but descriptions were inconsistent across items. More research needs to be done to find an approach to standardize item descriptions provided by retailers. Cross-referencing retailer UPCs to a national UPC database would have enabled the research team to use standardized item descriptions from the national UPC databases in the POC.

Because retailers comply with industry standards by masking the middle digits of EBT card numbers, testing was necessary to determine whether it is possible to reliably match encrypted transactions to ALERT data if other information is included in the transaction record. The study team achieved high matching rates: 99.6 percent for the small regional retailer and 99.3 percent for the large national retailer.

Cost Model

To assess the key design considerations, the study team developed a flexible cost model that could be applied across all four technical solutions. This cost model followed the key implementation and operational components of Technical Solution 4:

1. Extraction of item-level SNAP transaction data by retailers and transmission to the state-level contractor (an EBT processor was used in the POC for illustrative purposes) through a batch process
2. Consolidation of item-level data from all retailers by the state contractor/EBT processor
3. Transmission of consolidated data by the state contractor/EBT processor to FNS for storage and analysis

The study team estimated implementation and operational costs for small, medium, and large firms in Texas, assuming all stores use IECRs and transmit their data separately to an EBT processor. Adjusting the underlying assumptions of the model enabled IMPAQ to extrapolate cost estimates across these retailer-size categories to estimate state and national costs.

4. FINDINGS AND CONCLUSIONS

Technical Feasibility of Proposed Technical Solutions

Findings from the POC and key stakeholder interviews suggest that Technical Solutions 1 and 3—solutions relying on real-time transmission of item-level data—are not technically feasible. However, Technical Solutions 2 and 4 are technically feasible. Both of these solutions transmit item-level data in batch mode; they vary only in whether they use existing or new transmission pathways.

Technical Solution 1, which uses existing infrastructure to transmit Transactions with Item-Level Data (TWILD)[2] through EBT processors and third-party processors (TPPs) in real time, is not technically feasible. Although the idea of real-time data collection is intriguing, it poses several challenges that could potentially be a burden to retailers, TPPs, EBT processors, and FNS or its data storage contractor. The current infrastructure is not optimized to handle transmission of large volumes of data from retailers to TPPs and EBT processors, especially not while SNAP EBT transaction authorization is occurring simultaneously. Burdening the existing infrastructure used for transaction authorization with simultaneous transmission of large volumes of POS data might dramatically reduce transaction approval times.

Similarly, Technical Solution 3, which proposes using standardized email to transmit TWILD, is also not technically feasible. While this solution eliminates the challenge associated with real-time data transmission overburdening transmission lines with large data loads during transaction authorization, it poses the new challenge of consolidating a vast volume of email messages in one data center.

Technical Solution 4 is the preferred solution, as establishment of new transmission paths enables efficiencies to be introduced, such as sending TWILD directly to a contractor rather than through TPPs. Technical Solution 4 is preferred over Technical Solution 2 as existing infrastructure has been built to transmit fixed-size messages, but TWILD messages would likely be larger and vary in size. Therefore, current bandwidth would have to be expanded. Using existing channels for TWILD would incur significant POS software modification costs. In addition, using existing communication infrastructure would also require FNS to obtain legal authority to require EBT processors and TPPs to collect and transmit TWILD.

---

[2] *TWILD* refers to data captured at the POS when a SNAP EBT card is used. The data would be transmitted and stored in a proposed national FNS-owned database. See Section 4.4 for additional details.

Cost Estimate

The cost model and its underlying assumptions were applied to Technical Solution 4. Exhibit E.3 summarizes total implementation and operational costs of Technical Solution 4 for small, medium, and large firms; new state-level contractors (state/EBT processor); and FNS, for a single state, for example Texas. Implementation costs are one-time costs while operational costs are ongoing and estimated annually.

Exhibit E.3. Estimated Implementation and Operational Costs for a Single State

Stakeholder	Number of Stores	Number of Firms	Implementation Costs	Annual Operational Costs	Implementation and One Year of Operational Costs	Annual Operational Cost per Store
Firms Large	11,079	128	$10,931,200	$12,351,488	$23,282,688	$1,115
Firms Medium	627	93	$4,960,620	$3,946,920	$8,907,540	$6,295
Firms Small	3,431	3,247	$11,738,736	$27,666,400	$39,405,136	$8,066
State Level Contractor	-	-	$670,800	$470,880	$1,141,680	-
FNS	-	-	$1,710,000	$2,820,480	$4,530,480	-
Total	**15,136**	**3,468**	**$29,523,276**	**$47,256,168**	**$76,779,444**	**$3,122**

Exhibit E.4 provides a summary of all national costs to extract TWILD. These costs build on applying the single-state cost model to all states, based on store and firm counts for each state. As can be seen, the cost model estimates that it will cost all stakeholders approximately $419 million to implement the infrastructure necessary for TWILD collection. Additionally, each year, stakeholders are estimated to spend $598 million to operate the TWILD process.

Exhibit E.4. Summary of National Costs for Extraction of TWILD

Stakeholder	Number of Stores	Number of Firms	Implementation Costs	Annual Operational Costs	Implementation and One Year of Operational Costs	Annual Operational Cost per Store
Firms Large	147,774	1,708	$145,842,104	$164,791,332	$310,633,436	$1,115
Firms Medium	8,365	1,237	$65,970,725	$52,489,644	$118,460,369	$6,275
Firms Small	45,747	43,302	$156,537,393	$368,934,604	$525,471,997	$8,065
State Level Contractor *	-	-	$5,048,400	$4,429,437	$9,477,837	-
FNS	-	-	$3,919,800	$2,820,480	$6,740,280	-
Total	**201,886**	**46,247**	**$418,572,742**	**$597,659,749**	**$970,783,919**	**$2,960**

* Assuming 4 national state-level processors providing services to all SNAP-authorized stores nationally. This assumption introduces economies of scale as each state may not contract a unique processor.

5. RECOMMENDATIONS FOR NEXT STEPS

Technical Solutions 2 and 4 are technically feasible, as assessed on a limited scale using narrow parameters. Based on this preliminary analysis, Technical Solution 4 is preferred, as the introduction of new transmission pathways potentially eliminates the need for data transmission through TPPs.

The cost model showed that the effort by small firms to collect and format the data would be disproportionate when compared to large and medium sized firms. Small firms are estimated to represent between 5-10% of all SNAP redemption dollars per year, but they would account for 44% of total firm development costs and 63% of total firm operational costs. A seamless data collection strategy targeted for small firms would minimize the burden and bring down the overall cost for small firms.

An important next step for this research would be to validate the technical and cost implications presented in this report on a larger scale, such as a pilot test in a small geographic area. Specifically, a two-stage pilot of Technical Solution 4 would provide invaluable information. The first stage of the pilot would involve a number of selected retailers charged with the responsibility of piloting TWILD extraction and submission in a real-world environment. A second stage of the pilot would add more retailers across the state or FNS region to participate in modifying their POS systems to capture TWILD. The pilot's primary purpose would be to refine the technical solution and cost implications of a national data collection effort. The pilot could also serve as an opportunity to test emerging big data technology and data science approaches associated with data collection such as machine learning. Approaches being developed in these areas may prove useful for finetuning the technical solution for TWILD collection and may provide insight for other existing FNS data management and analytic activities.

Chapter 1. Introduction

This report is the result of a United States Department of Agriculture (USDA) Food and Nutrition Service (FNS) study led by IMPAQ International, LLC. The study explored the feasibility of creating a data collection system capable of directly and automatically providing FNS with item-level data on food purchases made by Supplemental Nutrition Assistance Program (SNAP) households. These data would be captured from purchases made by SNAP recipients using their Electronic Benefit Transfer (EBT) cards; plastic cards, similar to bank debit cards, used at retailers authorized to accept SNAP payments. In addition to exploring the feasibility of creating and operating such a data capture system, the study identified and examined the possible technical and cost parameters of such a system.

This report builds on earlier work to: 1) determine the basic research and technical requirements to develop an item-level data capture solution[3]; and 2) design potential technical solutions to capture these data and transmit them to FNS on a regular basis for research.[4] This report focuses on findings from development of a proof of concept and cost model assessing potential technical solutions for capturing item-level data from SNAP EBT purchases.

1.1 BACKGROUND

To support its mission of providing access for needy families to food and healthful diets, FNS needs a variety of data on the food purchasing patterns of SNAP households. FNS currently does not have the infrastructure to capture item-level food purchases made by SNAP households using their EBT cards. Furthermore, SNAP authorized retailers are not required by statute to report item-level data to FNS. Therefore, FNS must rely on consumer-reported data from several sources to gain insight into SNAP households' food purchasing behaviors. These sources include large datasets collected by other federal entities, such as the Flexible

[3] Garasky, S., S. Katz, K. Mbwana, M. Roy, A. Romualdo. "Feasibility Study of Capturing Supplemental Nutrition Assistance Programs (SNAP) Purchases at the POS – Final Basic Requirements Gathering Document". Prepared for the U.S. Department of Agriculture, Food and Nutrition Service – Office of Policy Support. February 2013.

[4] Garasky, S., S. Katz, K. Mbwana, S. Aampaabeng, Z. Miller, M. Roy. "Feasibility Study of Capturing Supplemental Nutrition Assistance Programs (SNAP) Purchases at the POS – Final Technical Solutions Report". Prepared for the U.S. Department of Agriculture, Food and Nutrition Service – Office of Policy Support. May 2014.

Consumer Behavior Survey module of the National Health and Nutrition Examination Survey and the Bureau of Labor Statistics' Consumer Expenditure Survey. While these two datasets provide some insights on foods available in SNAP homes and consumer purchases, respectively, they do not provide adequate data on actual purchases using SNAP EBT. FNS also conducts, or collaborates with, other new data collections involving participants in SNAP and other programs, including the Healthy Incentives Pilot Evaluation, School Nutrition Dietary Assessment Study, and National Household Food Acquisition and Purchase Survey. However, while these data are useful for research, they are limited in that they are cross-sectional surveys and rely on self-reported dietary intake and purchasing behaviors. They also tend to cover only 1-day to 1-week time periods.

Insight into SNAP participant shopping patterns also comes from analysis of data from the SNAP ALERT system,[5] which records payment information for every EBT transaction for every SNAP household. Each ALERT record includes store and household identifiers; EBT card number; date, time, and type of transaction; total amount of transaction; and account balance. However, because ALERT was designed to monitor electronic transaction activity and identify suspicious stores for analysis and investigation, it does not capture item-level purchase data.

For marketing and other purposes, many retailers and third-party companies (for example, Catalina Marketing, LoyaltyOne, and Nielsen) regularly capture, store, and analyze data on individual food items purchased by a subset of customers. Unfortunately, they do not produce sufficiently accurate item-level data on a large enough sample of identifiable SNAP households to allow for statistically reliable research analyses of food purchases. For all of these reasons, FNS is investigating the feasibility of creating a data collection system that will automatically deliver item-level data on purchases made by SNAP households on an ongoing basis.

1.2 RESEARCH OBJECTIVES AND STUDY DESIGN

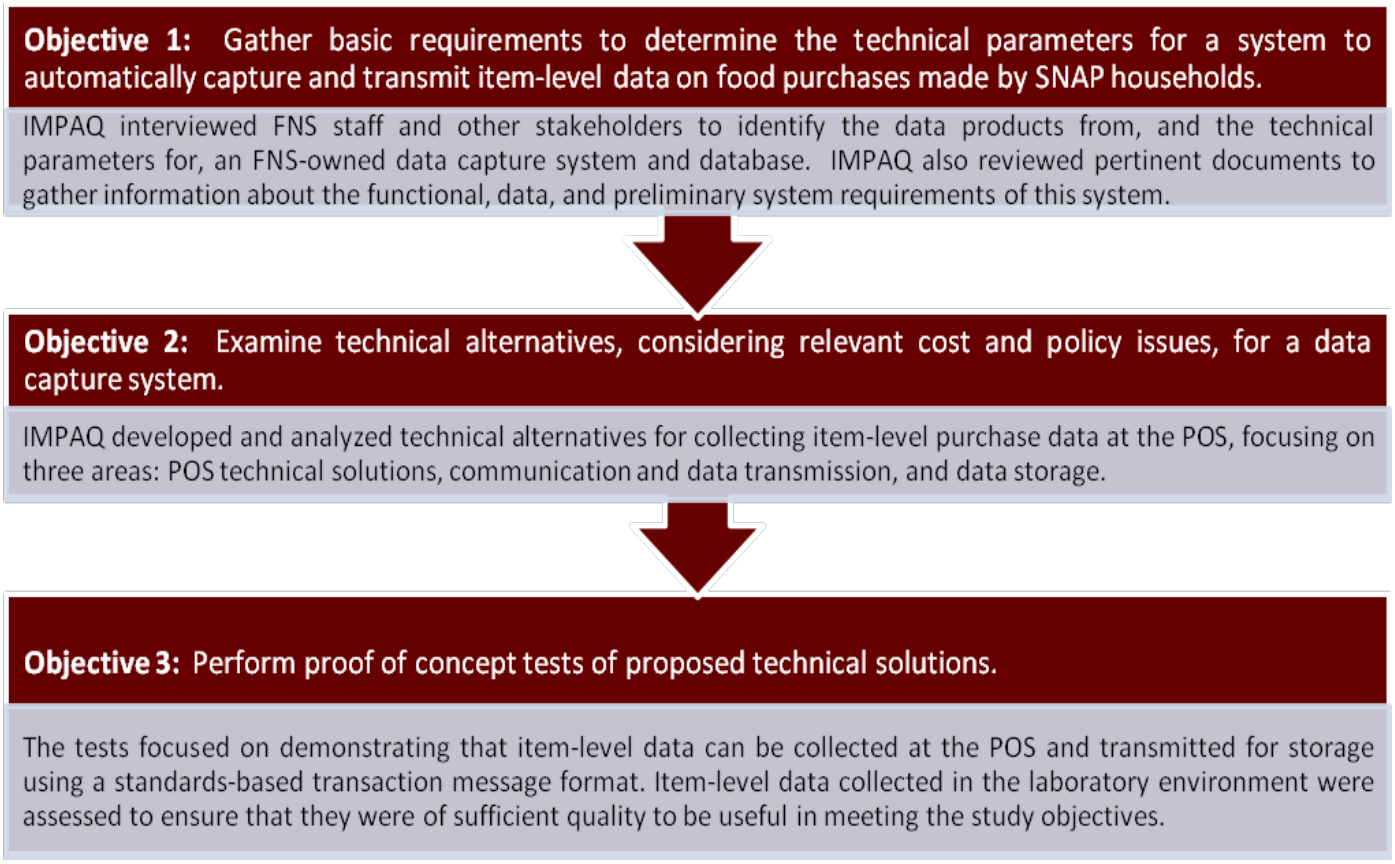
Exhibit 1.1 outlines the three primary objectives of this study. For the first objective—gathering basic requirements—IMPAQ identified the data products from, and technical parameters for, an FNS-owned data capture system and a database for storing SNAP purchase data. To do so, IMPAQ interviewed FNS staff and a variety of stakeholders. The study team also reviewed pertinent documents to gather information about the functional, data, and preliminary system requirements of a data collection and database storage system. Appendix A presents key findings from this phase of the project, along with a listing of key staff interviewed.

[5] The Anti-Fraud Locator Using Electronic Benefits Transfer Retailer Transactions

For the second objective—examining technical alternatives—IMPAQ developed and analyzed an array of technical alternatives for collecting item-level purchase data at the point of sale (POS). IMPAQ focused on three solution areas: POS technical solutions, communication and data transmission, and data storage. *POS technical solutions* refers to proposed data collection options that depend on POS hardware and software technology. *Communication and data transmission* refers to proposed options to address item-level data transmission pathways and frequency. *Data storage* refers to options for data processing and storage location.

For the third objective—POC testing—IMPAQ conducted preliminary limited testing of proposed solution areas to identify and better understand the technical solutions, communication and data transmission, and data storage areas by focusing on POC testing in a laboratory environment.

Exhibit 1.1. Research Objectives



This report focuses on the findings from the POC testing in the laboratory environment, as summarized in Chapter 2. The tests focused on demonstrating that Transactions with Item-Level Data (TWILD) can be collected at the POS and transmitted for storage using an American

National Standards Institute (ANSI) X9.58-based[6] transaction message format, as discussed in Appendix B. Additional testing ensured that the data that are collected would be of sufficient quality to be useful in meeting the study objectives.

1.3 PROOF OF CONCEPT CONSTRAINTS

Based on preliminary work,[7] the POC needed to meet three key challenges to ensure that findings were applicable to the prevailing retail environment:

- Obtain real-world retail data
- Implement the POC in a real-world retail environment
- Address multiple POS software and hardware configurations

IMPAQ conducted the POC using real-world retail data: SNAP EBT transactions drawn from actual operational retailers. This procedure enabled IMPAQ and FNS to gain insight into the challenges retailers would face when trying to capture and extract item-level data during their transaction processes. Moreover, using real-world data enabled the research team to better understand the nuances that could affect adoption and the cost of proposed technical solutions. Finally, working with real-world retail data revealed the challenges associated with standardizing item-level data from varied retailers. To do this, the study team acquired real-world, item-level data from a major national retail chain and a small (three-store) regional retailer.

Implementing the POC in a real-world retail environment was not possible, however, due to time constraints associated with negotiating access to retail systems and data. Using actual SNAP EBT cards in a shopping lane to complete real transactions would have offered several advantages. It would have given FNS firsthand experience with the benefits and challenges of each proposed technical solution, offered insight into the process necessary to obtain data use agreements, addressed data security compliance issues, and provided an opportunity for key

---

[6] The ANSI X9.58 standard provides all parties involved in EBT SNAP transactions with technical specifications for exchanging transaction messages between the retailer and an EBT processor. It specifies message structure and data elements used in SNAP transactions. It is based on the International Organization for Standardization (ISO) 8583 interchange specifications for messages originated by financial transaction cards; the X9.58 standards are tailored specifically for EBT transactions.

[7] Garasky, S., S. Katz, K. Mbwana, M. Roy, A. Romualdo. "Feasibility Study of Capturing Supplemental Nutrition Assistance Programs (SNAP) Purchases at the POS – Final Basic Requirements Gathering Document". Prepared for the U.S. Department of Agriculture, Food and Nutrition Service – Office of Policy Support. February 2013.

stakeholders such as EBT processors to be involved in the testing. To approximate these advantages to maximum extent, IMPAQ conducted the POC in a laboratory environment managed by a national provider of POS software and hardware solutions to retailers. The lab simulated an actual shopping lane complete with POS software and hardware, and simulated transmission of data to an EBT processor.

Testing all POS software and hardware configurations was also beyond the scope of the project. As outlined in Appendix C, five major POS hardware providers—Toshiba[8], NCR Corporation (NCR), Casio, Hewlett Packard (HP), and Dell—each supports a variety of POS software applications. Many software packages are cross-compatible because they use Windows-based operating systems. However, the software packages are also sufficiently differentiated that extraction of item-level data would potentially require customized solutions. Instead, IMPAQ conducted the POS testing using two configurations. The first, an IBM-based POS software configuration from a dominant market provider of POS software solutions that is used in customized versions by many national retail grocery chains. The second configuration used data generated from a *QuickBooks*™ POS application; more typical of the POS software solutions expected in small- and medium-sized retailers.

1.4 REMAINDER OF THE REPORT

Chapter 2 summarizes findings from earlier phases of the project to address Objective 2, designing technical alternatives to capture item-level data from SNAP EBT purchases at the POS. Chapter 3 discusses the design, development, and execution of the POC that includes an overview of the retail data and laboratory equipment acquired and the process used to assess the key design elements of the four proposed technical solutions outlined in Chapter 2. Chapter 4 assesses the steps required to ensure the data captured in the POC were complete and accurate and could potentially be linked to ALERT transactions. Specifically, Chapter 4 outlines the process and outcomes of matching item-level data captured in the testing environment to the ALERT database. Chapter 5 presents a cost model built on the POC findings that assessed the key development, implementation, and operational elements of the proposed technical solutions. Chapter 6 provides an overall feasibility assessment of the four proposed technical solutions using findings from the POC and the cost considerations, and provides recommendations for next steps.

[8] On April 17, 2012, Toshiba TEC acquired IBM's Retail Store Solutions business. http://www-03.ibm.com/products/retail/migration.html Retrieved June 16, 2015

Chapter 2. Technical Solutions

To prioritize options for further research and development, IMPAQ and FNS carefully weighed the advantages and disadvantages of the POS technical solutions in the areas of communication, data transmission, and data storage. The most promising technical solutions selected for further study were designed to work in an IECR environment. An IECR environment is ideal because:

1. IECR transactions represent more than 80 percent of all EBT redemptions.
2. IECR systems are technologically sophisticated enough to adapt to fulfilling FNS data needs. Retailers can use their existing ECR hardware with POS software modifications.
3. IECR systems include dynamic inventory databases that contain the item-level data FNS desires.
4. IECR-based solutions can leverage existing transmission infrastructures between retailers and processors.

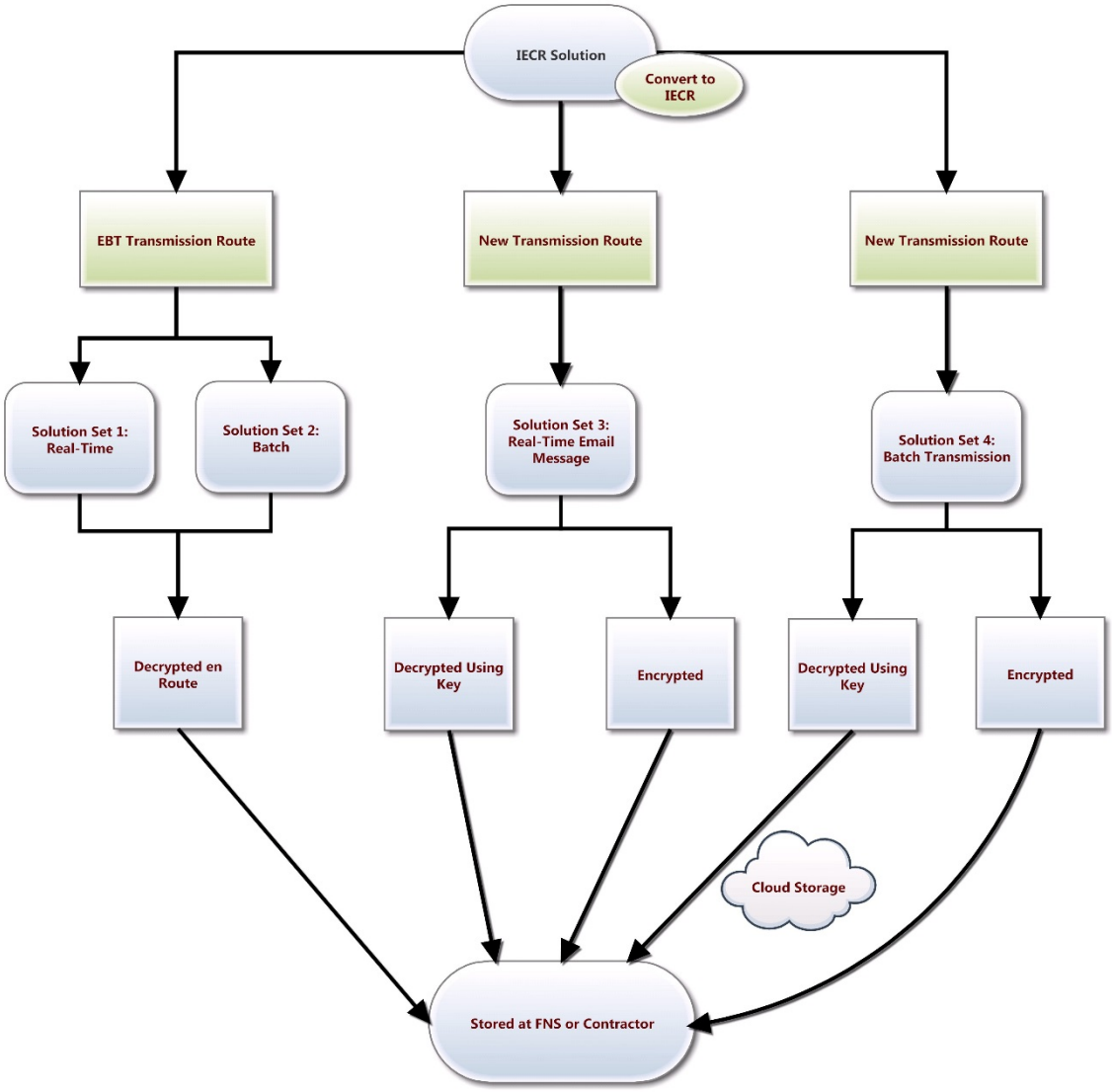
Exhibit 2.1 outlines the four technical solutions selected for further study. A summary of each potential solution follows. Each requires using a uniform data capture and transmission standard, either based on existing X9.58 standards or as part of a new standardized message format. As seen in the exhibit, these solutions vary in data collection and packaging, transmission route, and transmission frequency.

A critical element in SNAP transactions, regardless of the technical solution, is encrypting personally identifiable POS data to meet Payment Card Industry Security Standards Council (PCISSC) standards.[9] FNS does not mandate compliance with PCISSC standards for EBT transactions. Nonetheless, most retailers apply the standards to EBT transactions because EBT transactions integrate with other transaction methods using the same IECR, transmission, storage facilities, and PCISSC compliant transactions (e.g., credit cards). Retailers that accept EBT must encrypt the Personal Identification Number (PIN). These retailers also typically partially mask the primary account number—the EBT card number—displaying only parts of the

[9] According to PCISSC, cardholder data include the primary account number, PIN, cardholder name, card expiration date, and card service code. PCISSC standards have been widely adopted in the retail industry and require all these cardholder data elements to be encrypted.

number as necessary. Notably, retailers are already transmitting data via encrypted transmission channels and do not store EBT card numbers along with transaction data.

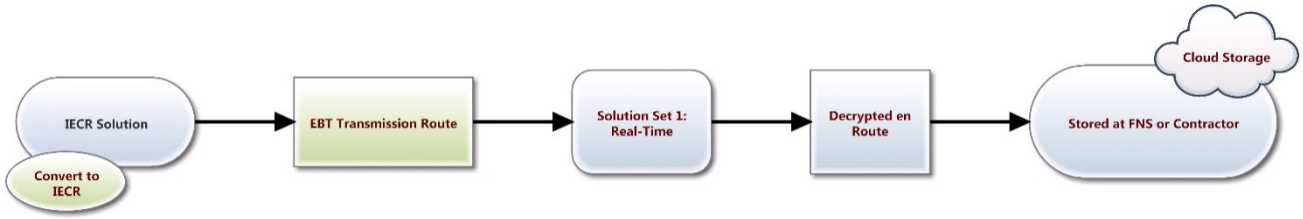
Exhibit 2.1. Technical Solutions Overview



2.1 TECHNICAL SOLUTION 1: EXISTING TRANSMISSION INFRASTRUCTURE WITH REAL-TIME DATA TRANSMISSION

Under Technical Solution 1, item-level data are captured and transmitted in real time using the current communications infrastructure through TPPs and EBT processors. EBT processors then transmit the data to FNS.

Exhibit 2.2. Technical Solution 1



2.2 TECHNICAL SOLUTION 2: EXISTING INFRASTRUCTURE WITH BATCH TRANSMISSION

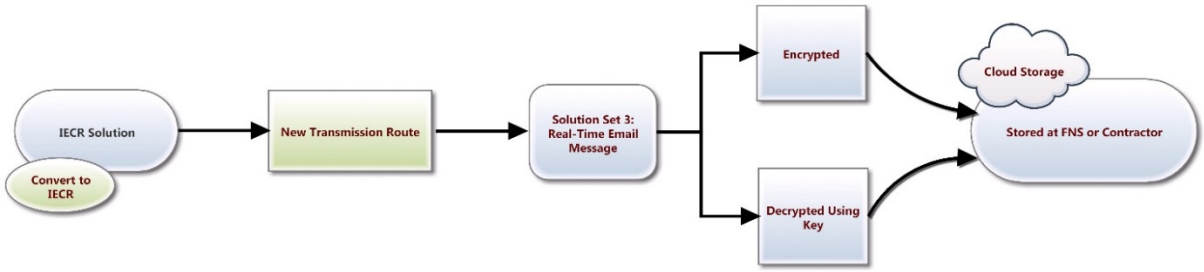
Technical Solution 2 is similar to Technical Solution 1, except that item-level data would be transmitted in batch mode, limiting the real-time capacity burden on all stakeholders. However, local and processor storage capacity would need to increase to accommodate data accumulation prior to batch transmission.

Exhibit 2.3. Technical Solution 2

2.3 TECHNICAL SOLUTION 3: NEW TRANSMISSION INFRASTRUCTURE USING ELECTRONIC MESSAGING

In Technical Solution 3, a new transmission process would generate standardized electronic email messages to capture and transmit item-level data from retailers, potentially avoiding the need to involve TPPs and EBT processors. These messages would be sent directly to FNS in real time.

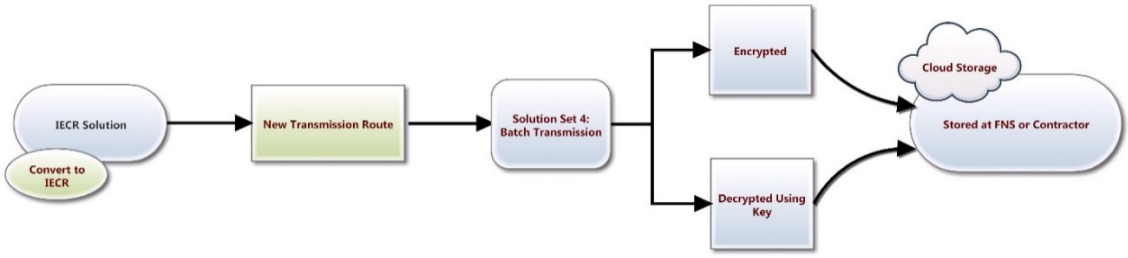
Exhibit 2.4. Technical Solution 3



2.4 TECHNICAL SOLUTION 4: NEW TRANSMISSION INFRASTRUCTURE USING BATCH TRANSMISSION

Technical Solution 4 is similar to Technical Solution 3, using new transmission routes, bypassing TPPs and potentially EBT processors. The new transmission paths call for a state-level contractor to receive TWILD from the retailers, and then consolidate, process, and pass them along to FNS. New transmission paths proposed in this solution would be very similar to the current pathways via TPPs and EBT processors: High-speed Internet transmission lines. Rather than burdening TPPs and EBT processors by adding to their current operations, this solution introduces new contractors specifically to manage TWILD receipt, processing, and transmission to FNS.

Exhibit 2.5. Technical Solution 4



These new contractors (possibly the current EBT processors) would need to develop direct transmission paths with retailers. Notably, the new transmission routes would be required to accommodate large batch files.

As demonstrated in the following chapters, the study team executed the POC and developed cost models on the assumption that a new state-level contractor(s) could enter the market and provide TWILD data consolidation and transmission services similar to current EBT processors providing services to states.

Chapter 3. Proof of Concept Lab Testing

The POC laboratory testing aimed to determine the feasibility of, and technical parameters for, an FNS-owned data capture system to collect item-level data from purchases made using SNAP EBT. The long-term goal of FNS is to create a data capture system that transmits POS item-level data from retailers to FNS. The IMPAQ team's POC approach tested the process of capturing POS item-level data from retailer systems, matched POS item-level data from retailers to ALERT data, assessed POS item-level data integrity, and estimated costs associate with implementing and operating the system.

The four proposed technical solutions discussed in Chapter 2 presented four key system design considerations:

1. **Transmission paths:** using new ***versus*** existing routes through TPPs and EBT processors.
2. **Transmission frequency:** transmitting data in real time ***versus*** in batch processing.
3. **Encryption and Masking:** while transaction PIN numbers are encrypted (as discussed in Chapter 2), the middle digits of 16 or 19 digit EBT card numbers are masked. Only the first six and last four digits are typically displayed.
4. **Storage:** processing and storing data on FNS computers, ***versus*** on contractor computers, ***versus*** in cloud storage owned by FNS or a contractor.

IMPAQ developed the POC process and lab environment on the basis of these design considerations. The study team designed the laboratory environment to simulate a single retailer lane using an IECR system and developed a process that simulated transactions from end to end: from using an EBT card at a checkout lane register to capture of TWILD to transmission and storage in a final national dataset. The POC design was also generic enough to provide information on all four design considerations encompassed in each of the four proposed technical solutions. Specifically, the IMPAQ team used the POC laboratory environment to assess the capture and extraction of TWILD from the retail environment and simulate transmission to a third-party, such as a state-level contractor.

This chapter begins by describing the retailer data used in the POC laboratory test and related analyses (Section 3.1). Exhibit 3.1 outlines the steps that the study team took to simulate a

SNAP transaction in a register lane for the POC. Beginning with Section 3.2, the chapter describes each step in detail, specifically:

- **Generating SNAP shopping baskets:** The study team used item-level data from retailers to create SNAP consumer shopping baskets. The team created UPC barcodes for each item. (Section 3.2)

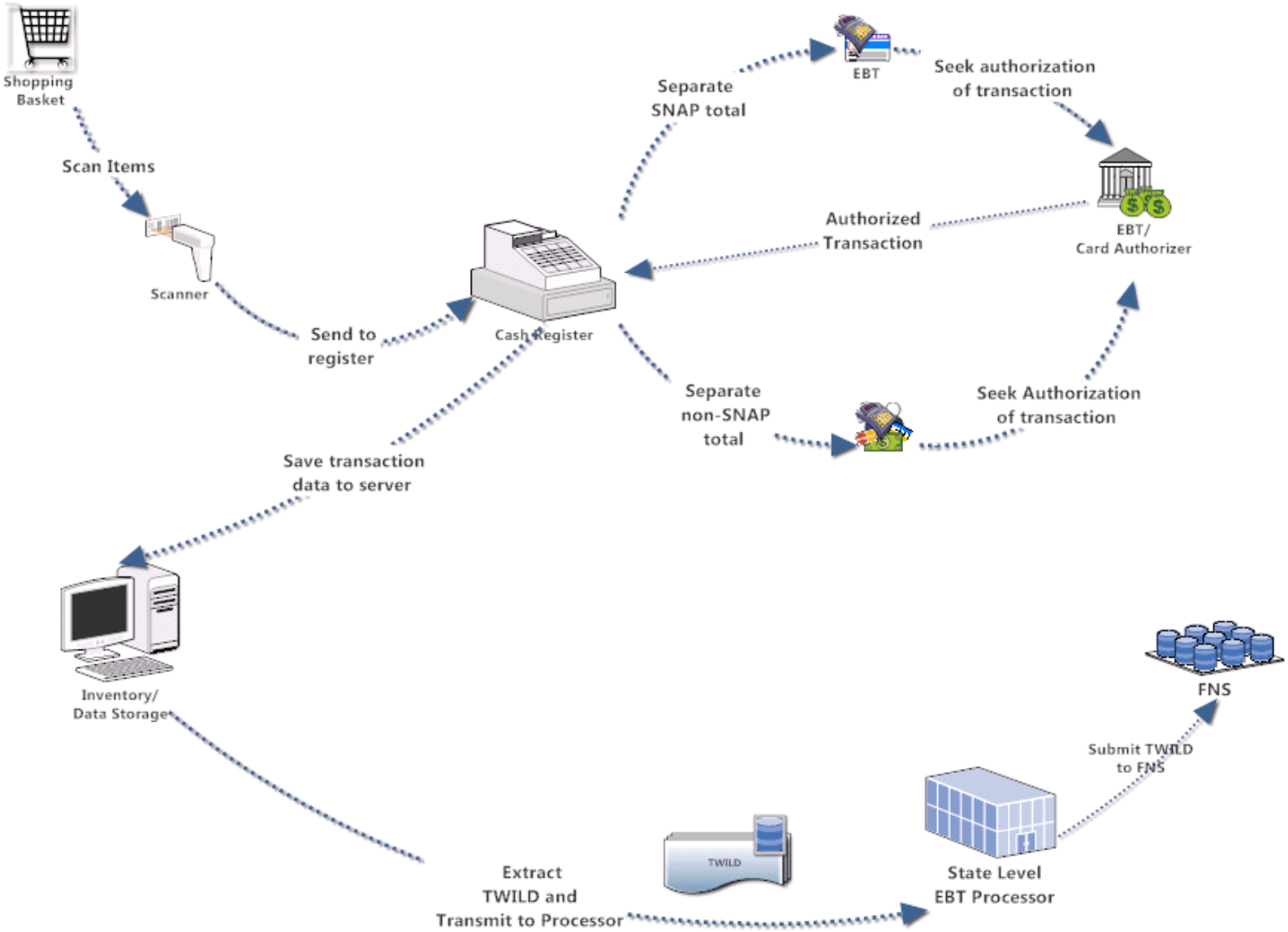
- **Selecting and testing cash register equipment:** The study team set up an IBM/Toshiba IECR.Notably, the team used data captured by *QuickBooks*™ POS software. (Section 3.3)

- **Simulating SNAP transactions and capturing TWILD:** Once items were scanned at the IECR, SNAP and non-SNAP transaction totals were sent to a simulated EBT processor and credit card processor for authorization, respectively. Retailer transaction records were stored locally. The study team then simulated transmission of the SNAP EBT transaction records to a state-level processor. (Section 3.4)

- **Extracting and storing TWILD data:** The study team developed software code to extract TWILD. The simulated state-level contractor then extracted the TWILD and simulated transmission of these data to FNS for storage. (Section 3.5)

The chapter concludes with Section 3.6, a summary of the various outputs and objectives achieved through the POC laboratory testing.

Exhibit 3.1. POC Testing Process – A Simulated Register Lane Capturing TWILD



3.1 ACQUIRING RETAILER ITEM-LEVEL DATA FOR THE POC

The IMPAQ team acquired item-level data for SNAP purchases from a large national retailer and three stores associated with a smaller regional retailer. The national retailer provided data in *Excel* file format that were already extracted from the native register format. These data were ready for further processing and analysis in software packages such as *Stata* and *SAS*. The regional retailer provided data that were still in the native register format. As a result, IMPAQ had to develop a procedure and computer program to extract the item-level data from the native format before importing it into software packages for analysis. Because larger national retailers are more likely than smaller retailers to have already extracted and processed the item-level data for internal marketing analysis, the size of the retailer will likely have implications for the cost of processing and consolidating TWILD (see further discussion in Chapter 5).

Key Outputs:

- *Inventory of items purchased using SNAP EBT, the corresponding UPC, and barcodes for these items.*
- *Retailer item-level data used to generate shopping baskets for testing in the POC laboratory environment.*

Item-Level Data from the National Retailer

The national retailer provided item-level transaction data from 14 retail store locations across 7 states (1 state from each of the 7 FNS regions) for May through July 2014. The 14 store locations averaged about 200 SNAP transactions per day per store. The national retailer extracted key variables (including EBT identifiers) needed to generate the TWILD and provided the data for each store location in *Excel* format, with separate tabs for item-level, basket- or transaction-level, and visit-level data.[10] The study team used these data to test algorithms for matching TWILD with the ALERT data, as described in Chapter 4.

Item-Level Data from the Regional Retailer

The regional retailer provided item-level transaction data from November 2013 to November 2014. Collectively, the regional retailer's three stores averaged about 200 SNAP transactions per day. Each of the stores provided transaction log files extracted from *QuickBooks* POS software. The IMPAQ team extracted the item- and transaction-level data (including EBT

[10] Visit-level data included information about store location.

identifiers) from the native register format (*QuickBooks* QDX file) into a single text file for each store.

Inventory of Items Purchased using SNAP

To develop an inventory of items purchased using SNAP, the IMPAQ team consolidated the item-level data provided by the regional retailer and identified all of the unique UPCs included in transactions involving SNAP EBT. The study team used this inventory to identify the UPCs comprising typical SNAP consumer baskets and used *JasperReports*®[11] to generate barcodes corresponding to the UPC code for each item. The study team then scanned these UPCs to simulate the transaction process and the capture of item-level data at the point-of-sale.

3.2 GENERATE SNAP SHOPPING BASKETS

This section describes the process of creating shopping baskets to simulate SNAP transactions using item-level data provided by the regional retailer. The study team used retail item-level data (detailed in section 3.1) to generate UPCs for items that comprise typical shopping baskets purchased using SNAP benefits.

IMPAQ generated 1,000 shopping baskets as described above. The items (UPCs) for 500 of the baskets were generated using a sampling method that selected UPCs in proportion to the frequency with which they were purchased in the data provided by the regional retailer. The study team sampled UPCs to match the overall distribution of the items in shopping baskets purchased with SNAP EBT, the retail department of the item (such as produce, cheese, or bakery), and the number of items (per UPC) purchased.[12] For instance, if 20

- *Printed PDFs for 1,000 shopping baskets. They included the UPCs, bar code, item description, item price, and purchase quantity of each item in the basket. The printed PDFs include the UPC bar code, 12-digit UPC-A number, UPC description, item price, and purchase quantity of each UPC in the basket.*
- *Appendix D provides an example of a PDF sheet.*

[11] The JasperReports® Library is the world's most popular open source reporting engine. It is able to use data coming from any source and produce pixel-perfect documents that can be viewed, printed or exported in a variety of document formats including HTML, PDF, Excel, OpenOffice and Word. http://community.jaspersoft.com/project/jasperreports-library Retrieved on November 11, 2015.

[12] IMPAQ used Stata's "sample" command to sample records to create the simulated shopping baskets. This command draws a random sample from the source dataset, while maintaining group proportions for specific variables. See http://www.stata.com/manuals13/dsample.pdf.

percent of all UPCs purchased using SNAP EBT in the retailer data belonged to the produce department, then approximately 20 percent of all UPCs sampled for inclusion in the 500 shopping baskets also belonged to the produce department.

By selecting items/UPCs in this manner, the study team created a sample of shopping baskets that preserved the distribution of UPC attributes described above and included a mix of UPCs representative of overall SNAP purchasing patterns at the regional retailer. The data provided by the regional retailer included all records that involved EBT as a tender type, including those involving multiple tender types (such as cash or credit card). As such, sampled UPCs could be SNAP-eligible or not; the original data included both. The data excluded returned purchases or voided items. Sampled UPCs were then combined to create the simulated baskets.[13]

In addition to the 500 shopping baskets generated through this sampling approach, IMPAQ randomly selected 500 actual shopping baskets from the regional retailer's item-level store data. Actual shopping baskets corresponded to transactions (and the corresponding items purchased) that actually occurred at the regional retailer's stores. The study team included these baskets to capture any purchasing patterns that may not have been emulated by the sampling method.

Assign SNAP EBT Card Numbers to Shopping Baskets

The actual item-level transaction data from the regional retailer were received with masked EBT card numbers, where all but the first six digits (the Bank Identification Number, BIN) and last four digits were replaced with zeroes. For the laboratory testing, the actual first six and last four digits of EBT card numbers were retained, while the masked EBT card numbers were replaced with randomly generated digits.[14] These numbers were attached to each of the 1,000 shopping baskets and keyed in during each POC transaction in the laboratory.

[13] The number of unique UPCs in the simulated shopping baskets was similar to the overall distribution of UPCs purchased based on the data provided by the regional retailer. For instance, if 6 percent of all shopping baskets included just 2 unique UPCs, then 30 of the 500 simulated baskets included 2 unique UPCs. Note that a shopping basket with just one unique UPC can contain more than one item for that UPC (such as a two-item basket containing two Red Delicious apples). The IMPAQ team also created simulated baskets with the number of UPCs equal to the number of UPCs in the 98th percentile, 99th percentile, and largest basket sizes based on the actual store data. This process resulted in TWILD files that varied in size (number of unique UPCs) and matched the distribution of basket sizes in the actual store data. Because UPCs in the simulated shopping baskets may have been drawn from one or more actual shopping baskets, the payment (tender) type of the first UPC was assigned to the remaining UPCs in the simulated basket.

[14] The middle digits were randomly generated using *Excel*'s RANDBETWEEN function.

Create and Print UPC Bar Codes for Scanning

After identifying the shopping basket items, IMPAQ developed an automated basket generation report using the following steps:

1. Convert UPC codes with 4, 5, and 10 digits to a standard 12-digit UPC-A code.
2. Convert 12-digit UPC-A codes into European Article Number (EAN-13) barcodes.[15]
3. Calculate the total dollar value for the basket.
4. Format the basket data to include basket number, barcodes, item names, item count, item prices, EBT card number, tender types, and basket total.
5. Create and print PDF files containing the shopping basket data.

3.3 SELECTING AND TESTING CASH REGISTER EQUIPMENT

In a retail environment, the IECR connects the electronic cash register, a local hard drive storing a dynamic version of the retailer inventory database and transaction records, a scanning device, a printer, and the POS terminal. These hardware components communicate with one another within an operating system. While not all IECRs function the same way, most cash register hardware also run a POS software application that facilitates tasks such as scanning purchases, applying discounts and coupons, and sending transaction totals to the POS terminal.

Key Outputs:

- *IECR POS system simulating a retailer checkout lane.*
- *Hardware and software components to be used to scan shopping baskets and store data in the local POS system.*

Exhibit 3.2 provides an overview of the hardware, operating systems, and software options generally available to retailers. The most critical hardware components affecting collection of POS item-level data are the IECR and the POS terminal.[16] Although there are a number of POS hardware manufacturers, most POS hardware can run Windows-based operating systems, which, in turn, can support a variety of POS software applications. Because most hardware run most POS software applications, POS hardware configurations are the major element to consider in modifying POS systems.

[15] This refers to the barcode symbology typically seen on retail items. This standard is intended to be worldwide and is backwards compatible with UPC barcodes. http://www.computalabel.com/aboutean.htm Retrieved on June 17, 2015.

[16] Appendix A provides additional detail on the major IECR hardware and software options in the retail environment.

Exhibit 3.2. POS Hardware, Operating Systems, and Software

	Hardware Manufacturer				
	IBM/Toshiba	NCR	Casio	HP	Dell
Compatible Operating Systems					
IBM/Toshiba-developed OS	✓				
Windows-based OS	✓	✓	✓	✓	✓
Linux-based OS		✓			
Supported POS Software Applications					
Toshiba SurePOS ACE for 4690 OS	✓				
Microsoft Dynamics RMS (Windows-based)	✓	✓	✓	✓	✓
Microsoft Dynamics POS 2009	✓	✓	✓	✓	✓
StoreNext (Retalix)	✓	✓	✓	✓	✓
Retalix10	✓	✓	✓	✓	✓
StorePoint (Retalix)	✓	✓	✓	✓	✓
StoreLine (Retalix)	✓	✓	✓	✓	✓
JPMA POS Designer	✓	✓	✓	✓	✓
Aurora	✓	✓	✓	✓	✓
RORCvIPOS V6	✓	✓	✓	✓	✓
Retail Professional	✓	✓	✓	✓	✓
NCR Advanced Checkout Solution (ACS)		✓			
QuickBooks Point of Sale					✓

IMPAQ selected IBM/Toshiba IECR hardware running IBM/Toshiba operating systems for POC testing because market reports indicate that IBM/Toshiba dominates the POS market in hardware and software.[17] Exhibit 3.3 summarizes the specifications of the hardware and software that were used, which required no modifications for the POC testing.

[17] http://newsroom.chainstoreguide.com/2014/02/pos-manufacturers-market-share-report/. Retrieved 7/20/2015.

Exhibit 3.3. POC Lab Equipment Specifications

Hardware	Specifications
Scanner equipment	DL Gryphon GD4430 2-D Imager
ECR printer	Toshiba 4610-2CR
Processor	Server Intel Celeron G540, Terminal Intel Celeron T3100
EBT Card Terminal	Verifone MX915
Data storage hardware	SATA hard drive
Software	**Specifications**
POS software vendor	Toshiba Global Commerce Solutions
POS software version number	Version 7 Release 4

3.4 SIMULATING SNAP TRANSACTIONS AND CAPTURING POS DATA

Steps for simulating transactions and capturing POS data in the laboratory environment follow below. This environment simulated an IECR lane capturing TWILD data.

1. **Scan item barcodes.** IMPAQ team members used a handheld barcode scanner to scan each item in the 1,000 shopping baskets. Scanning the barcodes attached product information to the item automatically. Barcodes that were not found in the laboratory inventory database required manually identifying the UPC number in the laboratory inventory database.
2. **Calculate total purchase amount for transaction**. After scanning every item in a given basket, pressing the "Cash/Tend" button on the register generated the total purchase amount due for the transaction.
3. **Pay with identified payment (tender) method**. The applicable payment type and EBT card number for each shopping basket was indicated at the bottom of each PDF sheet.
    a. **Manually enter EBT card number**. All transactions involved EBT as a tender type.

Key Outputs:

- *Verification of the process for capturing and storing item-level data for SNAP shopping baskets and transmitting to a third party.*
- *Verification of the process for masking middle digits of EBT card numbers.*
- *Transaction log files in native register format with item-level information for the SNAP shopping baskets.*

The 16-digit EBT card number was keyed in manually for each basket.[18]

b. **Apply complementary tender types, if applicable**. For cash tenders, an amount representing the cash tender amount was entered manually. For credit card tenders, a test credit card was swiped manually.

4. **Simulate EBT and credit card processor approval**. Approval of EBT and credit card tenders was automatically simulated.
5. **Capture POS data and store on the local POS system**. POS data capture and storage on the local POS system was PCISSC-compliant. The system automatically masked the middle digits of the EBT card number, storing them as zeroes.
6. **Print invoice and attach to shopping basket PDF sheet**. After each transaction, the POS system printed a paper receipt that was taped to a sheet of letter paper and then stapled to the printout of the PDF sheet containing the shopping basket barcodes.
7. **Transmit transaction log file from local POS system to state-level processor (IMPAQ)**. At the end of each day of testing, the transaction log file (in the native register format) with item-level data was transmitted to IMPAQ via an SFTP site to simulate transmission to a third-party.

3.5 EXTRACTING AND STORING TWILD DATA

After completing the POC testing in the laboratory environment, the study team retrieved the transaction log files posted to the SFTP site. The study team developed a program to extract TWILD from store transaction data and convert the data from the native format used by the POS software. First, the team developed a code to extract the data from the native ".qdx" format and saved the transaction data in a fixed position file layout format ("text"). The code was written using the Basic programming language. Next, the team used *SAS* to extract the transactions from the text format into SAS format for matching analysis.

Key Outputs:

- *A program to extract and convert TWILD from the native register format.*
- *Item-level data based on simulated SNAP transactions that can be used for further analysis, such as matching with ALERT data and validating data integrity.*

[18] As detailed above, the masked middle digits of actual EBT card numbers were replaced with randomly generated digits for POC testing. Because these randomly generated EBT card numbers did not correspond to actual account numbers, a card swipe could not be used.

3.6 SUMMARY OF PROOF OF CONCEPT LABORATORY TEST

The POC testing in the laboratory environment demonstrated the feasibility of capturing and storing item-level data from an IECR POS system and transmitting data from the local POS system to a third-party. These data were based on transactions that simulated SNAP EBT purchases at a retailer checkout register. The IMPAQ team developed a program and process for extracting and converting transaction data from native register format into data containing the specified TWILD data elements. The converted data were stored such that they could be analyzed using standard statistical software packages such as *SAS* or *Stata*. During the POC testing, the study team verified the masking process implemented by the local POS system, which converted the middle digits of EBT account numbers into zeroes. After extracting the TWILD and converting the data from the native POS format, the next steps of POC testing involved validating the integrity of the TWILD and matching them to ALERT data. Chapter 4 discusses these data validation and data matching activities.

Chapter 4. Proof of Concept Analysis

The POC laboratory testing demonstrated the feasibility of capturing, storing, and transmitting TWILD generated from simulated SNAP EBT transactions. This chapter begins by describing the analysis performed by the study team to validate the data captured in the laboratory environment, checking for completeness and accuracy of the data with respect to TWILD specifications. The chapter then describes the process for matching the simulated transaction TWILD files that included partially masked EBT card numbers to ALERT transaction files, and assesses the results of this matching process. Validating TWILD and matching it to ALERT data in this way would facilitate linking transaction data for research purposes, for example, by linking all purchase information for a given household in a given month to examine purchase patterns.

4.1 ASSESSMENT OF THE COMPLETENESS AND QUALITY OF TWILD

Completeness and Quality of UPCs/PLUs and Item Descriptions

As outlined in Chapter 3, the study team received item-level data from a national and a regional retailer. The data from the regional retailer was used to generate the SNAP shopping baskets used in the POC laboratory environment. Regarding the item-level data, the critical data fields for identification of an item are its UPC or Price Look-up Code (PLU) and item description. The study team analyzed the data to determine the completeness and quality of these data fields.

As shown in Exhibit 4.1, the national retailer data inventory contained 142,863 unique items. The data included a UPC or PLU for all items. Similarly, the regional retailer data inventories contained 17,406 unique items, all containing UPC or PLU codes. In addition, both the national and regional retailer data contained information describing the item and the store department that housed the item. However, the *item descriptions* contained varying amounts of detail. For example, some descriptions contained size and quantity information. Overall, it appears that national retailers deploy more sophisticated inventory systems and can maintain more complete item-level records compared to regional retailers.

Exhibit 4.1. Characteristics of UPCs/PLUs in the National and Regional Retailer Inventory

Retailer Category	Total Number of Transactions	Number of Unique Items in all Transactions	Ratio of Unique Items to Transactions	Percentage of Unique Items with UPC/PLU	Percentage of Unique Items with Item Descriptions
National	260,033	142,863	54%	100%	100%
Regional	76,302	17,406	22%	100%	100%

UPC and PLU codes are determined by industry coding conventions to ensure the codes are universal and recognizable to POS software. UPCs are typically 12 or 13 characters long while PLUs are 4 or 5 characters. Visual inspection of UPCs and PLUs determined that the national retailer inventories were of higher quality, having a lower rate of invalid UPCs or PLUs.

4.2 RATIONALE FOR MATCHING TWILD TO ALERT DATA RECORDS

The goal of the matching process was to take a transaction in the retailer data and find a corresponding transaction in the ALERT data. Matching retailer transaction files to ALERT transaction files required identifying data elements contained in both datasets. During the first phase of this study, FNS identified several key data elements for inclusion in TWILD data. Exhibit 4.2 lists relevant data elements across the three data sources, together with the elements FNS requires in the TWILD.

The matching process for the national retailer differed from that for the regional retailer because of differences in the data that were provided.

Exhibit 4.2. Summary of Retailer Data, ALERT, and TWILD

Level	Data Element	Retailer Type		FNS	Desired in TWILD
		National	Regional	ALERT	
Store	FNS Number	✓	✓	✓	✓
	Store Number	✓	✓		
	Store Address	✓			
	Store State	✓			
	Store City	✓			
	Store Zip Code	✓			
Transaction	Register/Terminal Number	✓			
	Transaction Date	✓	✓	✓	✓
	Transaction Time	✓	✓	✓	
	Transaction Number	✓	✓		
	Number of Items	✓	✓	✓	✓
	Visit ID		✓		
	Tender Type	✓			✓
	Tender Amount	✓			
	Transaction Amount	✓	✓	✓	✓
	SNAP EBT Basket Amount	✓		✓	✓
	Tax Amount		✓	✓	✓
	EBT Card Number			✓	✓
Item	SNAP Eligible	✓	✓		
	UPC/PLU	✓	✓		✓
	UPC Description	✓	✓		✓
	Retail Department	✓	✓		
	Price Paid	✓	✓		✓
	Net expenditure	✓	✓		✓
	Quantity Purchased	✓	✓		✓
	Item Weight	✓			✓
	Item Weight Unit	✓			✓

National Retailer Transaction Records

IMPAQ obtained transaction records from 14 stores of mixed sizes from the national retailer. Seven were large-volume stores with 250 to 300 daily SNAP EBT redemptions; the other 7 were low- to medium-volume stores with 50 to 70 daily EBT redemptions. Data elements were provided in three categories: store-level variables, transaction-level variables, and item-level

variables.[19] Exhibit 4.3 shows the data elements received from the national retailer and the use of each variable in the matching process.

Exhibit 4.3. National Retailer Data Elements

Variable	Description	Use in Matching
Store-Level Variables		
Store State	2-digit state code	Not used
Store City	City name	Not used
Store Address	Full address of store	Not used
Store Number	Internal store identifier	Match to FNS store number
Transaction-Level Variables		
Local Date	Date of transaction in store	Match to ALERT data
Local Time	Hour, minute, second of transaction	Not used
Register Number	Identifier for POS register	Check for duplicates or unmatched records
Transaction Number	Store-generated transaction identifier	Check for duplicates or unmatched records
Visit ID	Store-generated visit number	Check for duplicates or unmatched records
Number of Items	Number of items bought, including non-SNAP eligible	Check for duplicates or unmatched records
Tender Type	The type of payment	Help identify SNAP total
Tender Amount	Total dollar amount for tender type	Match to ALERT data
Transaction Amount	Total dollar amount	Not used in matching
Total Tax Amount	Total tax on transaction	Help identify SNAP total if necessary
EBT Number	Full EBT number (stored separately from other personal information)	Match to ALERT data
Item-Level Variables		
Item Description	Description of the item purchased	Not used in matching (but in TWILD)
SNAP Eligible	Indicator for SNAP-eligible item	Identify SNAP-eligible items
UPC	12- or 13-digit universal product code	Not used (but in TWILD)
PLU	4- or 5-digit price/product code	Not used (but in TWILD)
Retail Department	Department description	Not used (but in TWILD)
Retail Price	Price of item	Verify amounts, using SNAP-eligible variable
Total Retail Price	Total price paid (retail price times quantity purchased)	Verify amounts
Quantity Purchased	Units of item purchased	Not used (but in TWILD)
Item Weight	Weight of item	Not used (but in TWILD)
Item Weight Unit	Unit of measure, e.g., pounds, each	Not used (but in TWILD)

[19] Except for the store number, IMPAQ did not use the store-level information for matching or the analyses.

One important feature of the data from the national retailer is the inclusion of transaction dollar amounts by tender or payment type: EBT card, cash, check, debit or credit card, coupons, and gift cards. Identification of tender type means easy identification of SNAP transactions and accurate determination of the SNAP transaction total submitted to the ALERT system. In contrast, the regional retail data used for the POC did not break out transaction dollar amounts by tender type. However, the final TWILD specifications should include such a breakdown.

Regional Retailer Transaction Records

IMPAQ received 76,302 transaction records from the regional retailer, reflecting approximately 70 SNAP EBT redemptions per day per store for the 3 stores over one year. Exhibit 4.4 shows the data elements contained in the retailer data, grouped into three categories: store-level, transaction-level, and item-level variables. The store-level variables were needed to match to the ALERT data.[20] The variables of interest for the TWILD are the item-level variables.

[20] Other variables not useful for matching were useful to check the validity of the matching and to diagnose potential problems. For example, variables such as *transaction number* have no corresponding elements in the ALERT data, but were used for checking duplicate records. Note that the data include payment types, but not the amount charged to each tender type.

Exhibit 4.4. Regional Retailer Data Elements

Variable	Description	Use in Matching
Store-Level Variable		
Store Number	Internal store identifier	Match to FNS store number
Transaction-Level Variables		
Register Number	Store identifier for POS register	Check for duplicates or unmatched records
Local Visit Date	Date of transaction in store	Match to ALERT data
Local Visit Time	Hour, minute, second	Not used
Transaction Number	Store-generated transaction identifier	Check for duplicates or unmatched records
Item Count	Number of items in transaction	Not used (but in TWILD)
SNAP Basket Total	Total dollar amount	Match to ALERT data
Complementary Payment Type	Other tender type(s) used for payment	Not used (but in TWILD)
EBT Number (Masked)	EBT card number with only BIN and last four digits	Match to ALERT data
Item-Level Variables		
Item Description	Description of the item purchased	Not used (but in TWILD)
SNAP Eligible	Indicator for SNAP-eligible item	Identify SNAP-eligible items
UPC	12- or 13-digit universal product code	Not used (but in TWILD)
PLU	4- or 5-digit price/product code	Not used (but in TWILD)
Retail Department	Department description	Not used (but in TWILD)
Retail Price	Price of item	Verify amounts, using SNAP-eligible variable
Total Retail Price	Total price paid (retail price times quantity purchased)	Verify amounts
Quantity Purchased	Units of item purchased	Not used (but in TWILD)

4.3 DETAILED STEPS TO MATCHING RETAILER DATA TO ALERT RECORDS

Matching the retailer data to the ALERT data involved four steps: 1) identify keys, or common variables, in the retailer and ALERT data; 2) process data to ensure that formats are the same across all datasets, particularly for the keys; 3) match the data using the keys; and 4) analyze the matching results.

Step 1: Identify Keys

To match the retailer data to the ALERT data, IMPAQ identified data variables in the datasets to serve as keys (i.e., common data elements). Although the national and regional retailers collected similar transaction data, the final data elements differed significantly.

The keys in both retailers' datasets that corresponded to the ALERT data were:

1. FNS store number
2. Local transaction date
3. Transaction amount
4. EBT card BIN (first six digits of the EBT card)
5. Last four digits of EBT card

Step 2: Process the Data

With the full EBT card number available in the ALERT data, IMPAQ created the last two variables for the ALERT data—EBT card BIN and the last four digits of the EBT card number—to ensure that ALERT and retailer data had the same variables for matching.

Step 3: Match the Data Using the Keys

The matching algorithm used ANSI SQL query statements executed in the statistical software package *SAS*. The matching algorithm took each combination of the five keys for each transaction from the retailer data and looked for an exact corresponding combination within the ALERT data.

Step 4: Analyzing the Matching Results

As indicated earlier, the goal of the matching was to link the retailer data to the ALERT data to create the TWILD. Analyzing the matching results therefore involved comparing the number of transaction records in retailer data that could be identified in the ALERT data.

The matching generated three possible outcomes:

1) Exact match: If the keys from the retailer data had exact matches in the ALERT data. Simply, each key should match to no more than one record in the ALERT data.
2) Duplicate match: If the combination of the keys from the retailer data had more than one record in the ALERT data.
3) No match: If there was not a corresponding record in the ALERT data for the keys from the retailer transaction record.

Exhibit 4.5 shows the matching results. Row 1 shows the results for the regional retailer while row 2 shows the results for the national retailer. The regional retailer data included 23,049 EBT transactions; the national retailer included 266,383 EBT transactions. The match rate was defined as the number of exact matches divided by the total number of retailer transactions. For the regional retailer, 99.6% of the records from the transaction data achieved exact matches. This compares to 99.3% for the national retailer transaction data. The few remaining retailer transactions resulted in either duplicate matches in the ALERT data or no match.

Exhibit 4.5. Retailer Data to ALERT Matching Results

Retailer	No. of Transactions	Matching Results			
		Exact Match	Match Rate (%)	Duplicate Match	No Match
Regional	23,049	22,967	99.6%	26	56
National	266,383	264,545	99.3%	1,678	160

4.4 SUMMARY OF POC ANALYSES

The IMPAQ team assessed the quality of the item-level data captured at the POS through analyses of data received from the national and regional retailers and data captured through the POC test in the laboratory environment. The data received from the retailers met the specifications for TWILD, with all records containing a UPC or PLU and an item description, as well as information about the store department that housed the item. The study team developed a method for matching retailer transaction records containing partially masked EBT card numbers to ALERT transaction files, successfully matching more than 99 percent of retailer transactions to ALERT transactions. Validating the quality of the data for the TWILD and demonstrating a high degree of reliability in matching to ALERT data demonstrated the feasibility of using the data for research purposes.

Additional Analysis

IMPAQ conducted additional analysis to assess availability of TWILD UPCs in commercial UPC databases and the degree of standardization of item descriptions in these databases. Using UPCs to identify items from a variety of retailers poses challenges. Retailers routinely add new products to their shelves and consequently generate new UPCs. Third-party databases that track UPCs are inherently unable to keep pace with new UPCs that retailers generate. In addition, large retailers often carry generic store-brand products. UPCs for these products are

often unique to the retailer and may not be easily accessible in national databases, especially for new products. Our analysis showed that only 5 to 6 percent of the UPCs and PLUs from the national retailer data were identified in two commercial UPC databases, Nutritionix and SimpleUPC, while 10 to 16 percent of the UPCs and PLUs from the regional retailer were identified in these databases.

Lack of standardization in data fields across retailers also poses a challenge. While UPC and PLU codes are largely universal (standardized), item descriptions are unique to each retailer. Retailers often use abbreviations and retain different item description lengths, leading to an assortment of item descriptions for the same product across retailers. However, national UPC databases such as Nutritionix or SimpleUPC generate their own internal item descriptions from each retailer's inventory, leading to a unified item description for the same product across retailers. Using existing national UPC databases such as Nutritionix and SimpleUPC could simplify the process of standardizing the item descriptions for an FNS SNAP item-level EBT purchases database.

Another option that can help with the UPC description standardization is using machine learning technologies to process the data after collection. Over time as more UPC data is collected from retailers, computer algorithms leveraging machine learning principles can produce standard item descpriptions and other detail for each UPC based on available data to date. The standardized UPC database would then be developed over time as more data is collected and processed.

Finally, although the POC analysis reliably matched retailer POS data to ALERT transactions even in situations where the middle digits of the EBT number were masked, the potential movement toward using tokenization to secure data in electronic transactions could affect matching retailer and ALERT data. Although it is beyond the scope of this study, further testing and analysis around tokenization should be considered to ensure that item-level data from retailers can be reliably matched to ALERT transactions in the future.

Chapter 5. Technical Solution Cost Model

This chapter develops a model to estimate the costs to implement and operate the proposed technical solutions. Findings from the POC laboratory testing, along with other evidence, offer guidance on the relative feasibility of the four proposed technical solutions for additional research and development (i.e., for piloting). As discussed above, Technical Solution 4 (New Transmission Infrastructure Using Batch Transmission) was deemed most feasible. The TWILD collection cost model discussed in this chapter is based on Technical Solution 4 and is designed to facilitate estimation of single store, retail firm, regional, and national costs. The model estimates the cost of a single state--Texas--and extrapolates to the national level using Texas' retailer characteristics. The cost model is only applicable to stores that use IECR cash register technology as this was a key premise on which the technical solutions were based. The cost model, with slight modifications, can be applied to the other technical solutions, as well.

Stakeholders in Technical Solution 4

In accordance with Technical Solution 4, the TWILD process will involve three main stakeholders: 1) SNAP-authorized stores, 2) State-level TWILD processors, and 3) FNS.

1. *SNAP-Authorized Stores:* For simplicity and to factor in economies of scale, the cost model focuses on retail firms rather than individual SNAP-authorized stores. These firms are expected to centralize TWILD operations and therefore reduce the cost per store.

2. *State-level TWILD Processors:* Technical Solution 4 involves stores extracting TWILD and transmitting the data to FNS in batch mode via a new transmission pathway. For the POC, the new pathway is assumed to be a state-level processer ("processers"). The Processers will establish secured network connections between Firms and FNS.

3. *FNS*: FNS will be the main driver of the TWILD collection process by establishing the legal framework within which to collect the data. It will coordinate all TWILD collection activities between firms and processors. In addition, FNS will receive the TWILD from processors and match to its ALERT data.

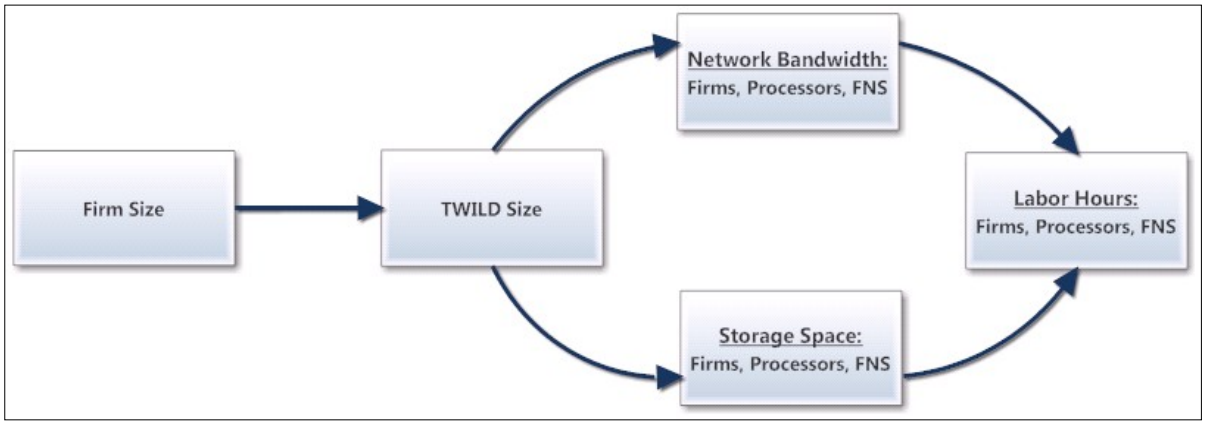
The TWILD collection process will involve establishing the infrastructure and processes to extract the TWILD from store transactions, consolidate the TWILD in a central location, and

submit to FNS for further processing and storage. Each stakeholder will have a separate cost model, taking into consideration their unique processes and implementation activities.

Cost Considerations

The costs for implementing and operating a TWILD process will be influenced primarily by the **size of firms** and the **number of SNAP transactions** within a specified period. Exhibit 5.1 shows the main factors influencing the cost of implementing Technical Solution 4. Firm size will determine the size (in kilobytes) of SNAP EBT transactions, hence the TWILD. The TWILD size will then determine the network bandwidth to transmit the data from firms to FNS and the required storage capacity. Finally, the size of the required network and storage infrastructure will determine the labor hours needed to establish and operate the TWILD transmission process for all stakeholders. As explained in section 5.3, implementation and operational costs will vary depending on firm size. Larger firms with multiple SNAP-authorized stores are assumed to benefit from economies of scale and are therefore able to reduce the cost per store.

Exhibit 5.1. Cost Considerations



Firms with IECR systems store data differently. From the POC, the research team learned that large and medium-sized firms typically transmit and store all detailed transaction data at a centralized location. Thus, they already possess the basic elements of the TWILD in some format for their own internal use. Smaller firms, however, do not necessarily store their transaction data in such manner. Rather, the transaction data may remain on a server, either central to all operations or unique to each POS terminal, and the data may remain in a native format. These differences mean that the processes for extracting transaction data will differ by firm size or configuration.

The cost of operationalizing the TWILD has two main components: 1) development and implementation ("implementation"), and 2) operational.

Implementation Costs

All stakeholders will have to undertake significant implementation activities to operationalize the TWILD collection. Exhibit 5.2 summarizes the main activities of each stakeholder in the implementation process. The activities center around four main components of the TWILD operations: extraction, transformation, transmission, and storage. Each stakeholder performs implementation activities related to each of the four components and their activities serve as inputs into the TWILD implementation costs.

Exhibit 5.2. Summary of Implementation Activities

<table>
<tr><th>Process</th><th>FNS</th><th>Processors</th><th>Firms</th></tr>
<tr><td rowspan="2">Extraction</td><td>Develop legislation authorizing TWILD collection</td><td rowspan="2">-</td><td rowspan="2">Develop process to consolidate data from terminals</td></tr>
<tr><td>Develop TWILD specifications</td></tr>
<tr><td rowspan="2">Transformation</td><td>Develop prototype TWILD extraction software</td><td rowspan="2">Develop software to verify, process, and consolidate TWILD from retailers</td><td rowspan="2">Develop/modify prototype TWILD software</td></tr>
<tr><td>Develop matching algorithm</td></tr>
<tr><td rowspan="2">Transmission</td><td>Develop transmission standards</td><td colspan="2">Develop transmission infrastructure</td></tr>
<tr><td>Develop transmission infrastructure</td><td colspan="2">Develop transmission protocol</td></tr>
<tr><td>Storage</td><td>Store data in cloud</td><td>Implement temporary storage</td><td>Implement temporary storage</td></tr>
</table>

Operational Costs

Similar to the implementation process, Exhibits 5.3 and 5.4 outline the roles and activities of the three key stakeholders regarding the key operational components of Technical Solution 4:

1. **Extraction by Firms.** Firms with IECRs will extract SNAP EBT TWILD from their transaction data to conform to the specifications of FNS. Firms could use either the TWILD extraction software provided by FNS or modified to meet their unique configurations. Also, firms differ in the availability of transaction data. While large firms typically store data in a warehouse, smaller firms using basic POS technology may store data only on the hard drives within their POS. Exhibit 5.3 shows these two different sources of transaction data in the right panel.

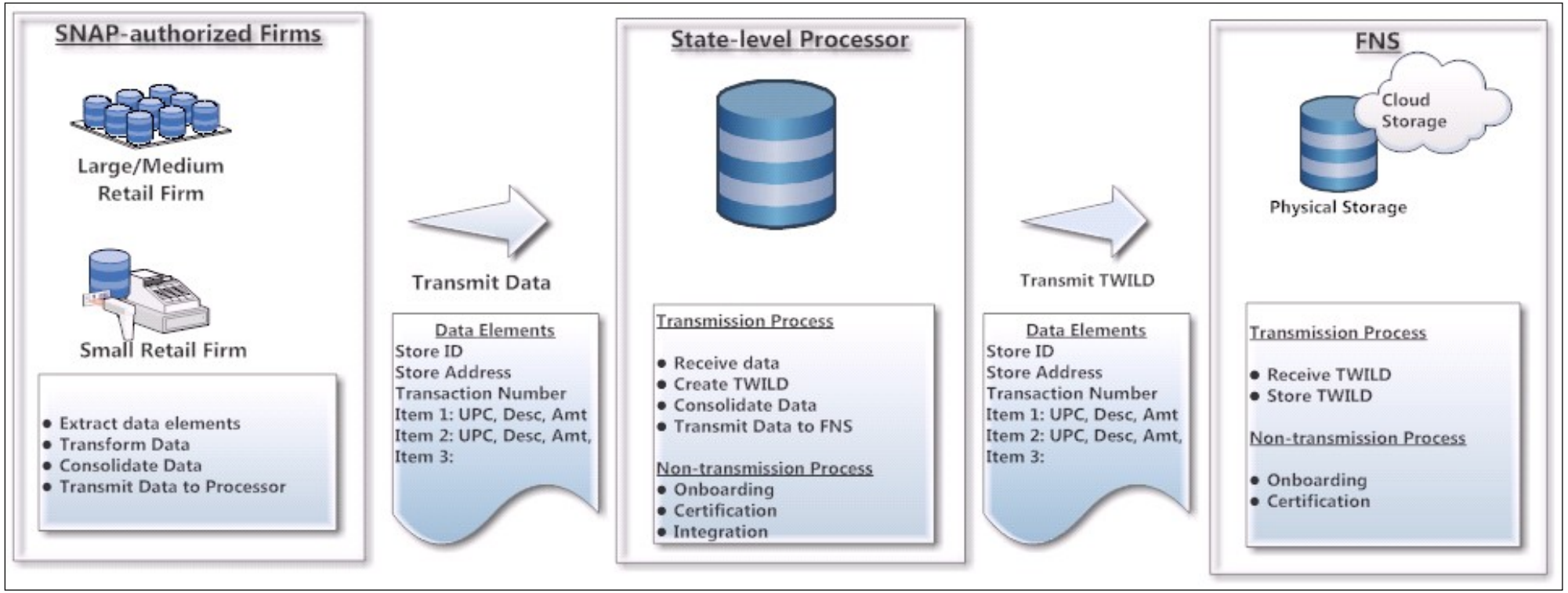
2. **Transmission of TWILD by retailers to processors.** Retail firms transmit TWILD to a processor in a batch process. **For this cost model, the research team assumes a daily batch process**. Batch transmission of TWILD requires no modification of the POS software. TWILD extraction occurs outside the normal EBT transaction process.
3. **Validation and consolidation of retailer TWILD by processors and transmission to FNS.** The processors will be responsible for validating and consolidating the TWILD they receive from firms. They will then transfer the consolidated TWILD to FNS for storage and analysis.
4. **Matching and FNS Storage of TWILD.** FNS will store and maintain the consolidated TWILD for analysis. As discussed in Chapter 4, the TWILD will be completed by matching with ALERT data to provide information that is more useful to researchers. To do so, FNS will implement a matching algorithm similar to that developed during the POC and discussed in Chapter 4.

The remainder of this chapter discusses implementation and operationalization of each of the three model components. The chapter is organized into: 1) Firm categories, 2) Cost inputs, 3) Development and implementation cost model, 4) TWILD operations cost model, 5) Application: Single state, and 6) National cost of TWILD collection.

Exhibit 5.3. Summary of TWILD Operational Activities

Summary of TWILD Operational Activities			
Process	**FNS**	**Processors**	**Firms**
Extract	--	--	Consolidate data from terminals/stores
			Extract required data elements
Transform	Maintain TWILD extraction software	Develop software to process data into consolidated TWILD file	Verify and consolidate TWILD
	Develop matching algorithm		
Transmit	Transmit data to Processor	Transmit data to FNS	Transmit to state-level Processor
Store	Match data to ALERT	Temporary store data for onward transmission to FNS	Temporarily store data for onward transmission to Processors
	Store data in cloud, warehouse		

Exhibit 5.4. Operations Cost Model



5.1 FIRM CATEGORIES

The cost model considers the possibility of economies of scale among retail firms. Larger firms, with more established processes and more SNAP-authorized stores, will have a different cost structure than smaller retail firms. Thus, while implementation costs for the firm as a whole may be similar between large and small retail firms, the cost per store will be lower for larger retail chains. Currently, FNS authorizes individual stores to accept SNAP EBT regardless of the firm. For example, for a large retail chain which has numerous individual stores, each of the individual stores have to be authorized to accept SNAP. However, due to the nature of the TWILD process, this large retail chain would benefit immensely from organizing at the national/regional level rather than at the store level because of economies of scale associated with operating the data system. For this reason, the cost model takes into consideration the number of stores of a particular firm. In this report, the number of stores will indicate firm size: small, medium, or large.

Large, medium, and some small firms likely have IECR cash register technology. As shown in Exhibit 5.5, retailers vary by type, SNAP redemption volume, and cash register technology. Per Technical Solution 4, the cost model focuses on retailers using IECR technology, which collectively account for 92 percent of all redemptions. IMPAQ estimates, using the most recent data (Exhibit 5.5), that 47 percent of all firms classified as small have the IECR technology required to implement the TWILD collection process.[21]

Exhibit 5.5. Benefit Redemptions by Retailer Type

Firm Types	Number of Authorized Stores	% of Total Redemptions	Cash Register Technology	Firm Size
IECR/Combination				
Super Store	18,942	48.66%	IECR	Large
Supermarket	18,594	33.02%	IECR	Medium
Combination Grocery/Other	68,283	6.84%	IECR	Large
Medium Grocery Store	11,746	2.07%	IECR	Medium
Large Grocery Store	3,827	1.55%	IECR	Large
Total IECR	**121,392**	**92%**	**--**	**--**
Non-IECR Technology (Excluded from Cost Model)				
Meat/Poultry Specialty	3,128	0.59%	Standalone	Small
Convenience Store	105,742	4.95%	IECR/Standalone	Small

[21] Share of total redemptions by Small Grocery Stores divided by the total non-IECR cash register usage.

Firm Types	Number of Authorized Stores	% of Total Redemptions	Cash Register Technology	Firm Size
IECR/Combination				
Small Grocery Store	13,845	1.16%	IECR/Standalone	Small
Seafood Specialty	1,634	0.23%	Standalone	Small
Bakery Specialty	2,732	0.18%	Standalone	Small
Delivery Route	1,080	0.16%	Standalone	Small
Military Commissary	190	0.12%	Standalone	Small
Fruits/Vegetable Specialty	1,337	0.10%	Standalone	Small
Non-profit Food Buying Co-op	403	0.04%	Standalone	Small
Farmers' Market	2,866	0.02%	Standalone	Small
Wholesaler	12	0.01%	Standalone	Small
Direct Marketing Farmer	2,309	0.01%	Standalone	Small
Total Non-IECR	**135,278**	**8%**	**--**	**--**
Total	**256,670**	**100%**		

Source: Information in columns 1 to 3 are from the SNAP Benefit Redemption Division 2014 Annual Report. The study team used their collective experience, subject matter expertise, and key informant interviews to estimate the column 4.

As noted earlier, the cost model is based on SNAP-authorized stores in Texas. A list of SNAP-authorized stores drawn from Store Tracking and Redemption System (STARS)[22] was used to construct three categories of stores: large, medium, and small. The steps taken to construct each category and descriptions of each category follow:

1. *Identify SNAP-authorized stores in Texas*. Identified the exact number of stores from STARS. Each store was counted independently, even if the store was part of a multi-store firm.
2. *Identify stores belonging to the same firm*. Store names from the STARS data were examined using statistical coding and visual inspection to identify stores belonging to the same retail firm.
    a. Multi-store retailers with at least five stores were consolidated into one record for the firm.
    b. Consolidating multi-store retailers reduced the list of Texas stores to 7,130 firms (approximately 40 percent of the initial STARS store count).

[22] http://www.fns.usda.gov/snap/retailerlocator

3. *Categorize firms into Large, Medium, and Small based on their store count*. A key cost assumption is that cost structures (inputs) are similar within a firm size category, but that cost structures vary across categories due to economies of scale, IECR sophistication, and resource availability. With this in mind, firms were categorized as follows:

    a. Large firms were defined as more than 20 stores.

    b. Medium firms were defined as 5 to 20 stores.

    c. Small firms were defined as 4 or fewer stores.

Exhibit 5.6 summarizes the distribution of SNAP-authorized firms in Texas by the aforementioned categories, with 128 large, 93 medium, and 6,909 small firms, for a total of 7,130 firms (compared to 19,004 stores). Adjusting for only firms with IECR cash register technology (Columns 5 and 6); there are 3,247 small firms. As discussed in Sections 5.5 and 5.6, information in columns 5 through 8 was used to extrapolate costs at both state and national levels.

Exhibit 5.6. Distribution of SNAP-Authorized Retailers in Texas by Firm Size

Firm Size (1)	Number of Firms (2)	Percentage of all Firms (3)	Number of Stores (4)	Estimated Firms with IECR Cash Registers (5)	Estimated Stores with IECR Cash Registers (6)	No. of Stores per Firm (7)	Percentage of Stores (8)
Large	128	2%	11,076	128	11,076	86.7	58%
Medium	93	1%	629	93	629	6.8	4%
Small	6,909	97%	7,299	3,247	3,431	1.1	38%
Total	**7,130**	**100%**	**19,004**		**15,136**	**--**	**100%**

5.2 COST INPUTS

The cost model assess the level of effort required by each stakeholder to implement and operate the TWILD process. Exhibit 5.7 shows implementation and operation costs for firms. The exhibit reports differences in labor rates across the three firm sizes. Column 1 shows the three categories of labor that a retail firm typically uses in implementation and operation activities. Column 2 provides the distribution of workload across these categories. The last columns show the blended labor rate, the average hourly rate assuming the workload distribution in Column 2. These rates were used in the cost model to estimate labor costs of all activities.

Exhibit 5.7. Labor Rates

Labor Category (1)	Ratio of Workload (2)	Firm Labor Cost per Hour			FNS Labor Cost per Hour	Processor Labor Cost per Hour
		Large	Medium	Small		
Implementation Labor Rates						
Senior Developer/Analyst	0.1	$150	$130	$80	$200	$200
Developer	0.3	$100	$90	$65	$150	$150
Junior Analyst	0.6	$75	$65	$55	$100	$100
Blended Rate	**--**	**$90**	**$79**	**$61**	**$135**	**$135**
Operations Labor Rates						
Senior Systems Analyst	0.2	$120	$100	$45	$200	$150
Systems Analyst	0.3	$100	$70	$30	$150	$100
Junior Analyst	0.5	$85	$50.0	$20	$100	$75
Blended Rate	**--**	**$93**	**$73**	**$26**	**$135**	**$90**

NB: These rates are based on Federal rates (Source: http://www.fedjobs.com/pay/washington.html) and typical contractor industry rates.

Exhibit 5.8. Blended Labor Rates for Stakeholders (Cost/hr)

Stakeholder	Implementation	Operation
FNS	$135.00	$135.00
State-level Processors	$135.00	$90.00
Retail Firms	$76.50	$52.83

Model Assumptions

The research team used information from the POC for the cost model. As mentioned previously, the main components of the cost model are driven by the size of SNAP EBT TWILD and the infrastructure and labor required to implement and operate the process.

Unit of Analysis: The research team estimated the cost model for a single state, Texas, over a single month, and then extrapolated to a single year and subsequently to the national level. Labor units required for all implementation and operational activities were blended labor units. Within the state, the unit of analysis is assumed to be firms. The firms will centralize activities and thus spread the total cost over the number of SNAP-authorized stores it owns.

Data Size: From the POC, the average TWILD size for a single transaction is 4.3 kilobytes (kb). Exhibit 5.9 shows the breakdown of the TWILD size by firm size. For the model state, using the total number of SNAP-authorized stores, the average TWILD size per month for the entire state

will be 314 gigabytes (GB). This final size of the collected TWILD will determine the inputs required to develop and operate the TWILD process. Exhibit 5.9 shows the storage and network bandwidth requirements for collecting TWILD. The data storage requirements include provision for backup and redundancy in accordance with industry standards. The cost model assumes all stakeholders will require storage space that is two times the TWILD size they manage.

Exhibit 5.9. Storage and Network Requirements for TWILD Collection

Firm Size	Number of Transactions	TWILD Size (GB)		Infrastructure Requirements		
		Texas	National	Network Bandwidth	Texas Storage	National Storage
Large	6,000	288	3,842	128	576	7,684
Medium	3,570	9	125	128	18	251
Small	1,260	17	217	64	34	433
Total	**10,830**	**314**	**4,184**	**--**	**628**	**8,368**

*The required bandwidth estimate assumes a maximum transfer time of 30 minutes.

**The required storage assumes provision for 40% redundancies and backups to four different sites. Storage measured in GB.

TWILD Process: Firms will extract TWILD in batch mode using a software prototype provided by FNS. For this cost model, IMPAQ assumed firms will submit TWILD daily. Firms will transmit the TWILD to designated processors who will verify and consolidate all TWILD from its firms. Processors will use a secure network to transmit the consolidated TWILD to FNS for final processing and storage. For the cost model, it was assumed that FNS will also perform the matching with ALERT before storage. It was assumed that FNS can then securely store the data in the cloud using FedRAMP or in a data warehouse.

5.3 DEVELOPMENT AND IMPLEMENTATION COSTS

All three stakeholders will perform activities required to setup and operate the TWILD process. The cost model assumes FNS will coordinate and lead the process. Some of the activities required of FNS, such as establishing legislative authority, are required before the retail firms and processors can begin their activities.

FNS Development and Implementation Cost Model

Before retail firms and state-level processors begin the process of collecting and processing TWILD, FNS must establish the legal basis for collecting TWILD, develop TWILD specifications and format, develop a TWILD extraction software prototype, develop a matching algorithm, and develop onboarding, certification standards, and operating manuals.

TWILD Extraction Legislation. Presently, retail firms are not obliged to extract and send detailed SNAP EBT transaction data to FNS, while current regulations allow FNS to collect only transaction amounts and not details of the purchase. As part of the implementation, FNS will have to propose new legislation or interpret existing regulations to allow collecting TWILD from SNAP-authorized retail firms.

Inputs and Resource Costs: The cost model uses the **labor hours** FNS will expend to formulate new legislation or interpret existing ones as main inputs for this sub-activity. The research team estimated that FNS will need 1,880 labor hours at the blended rate to establish the legal basis to collect TWILD from firms.

TWILD Specification Standards. With the legislative authority, FNS can define the technical and procedural specifications of the TWILD. The standard development will include working with ANSI to develop and implement new standards. The standards development process takes considerable effort and involvement of all stakeholders. Exhibit 4.2 from Chapter 4 provides a template of the required TWILD specifications and data elements.

Inputs and Resource Costs: Similar to the legislative authority, the standards specifications will require **labor hours** as the main inputs. Following past efforts at developing standards, the cost model assumes FNS will require 3,760 hours of labor to develop the TWILD specifications.

TWILD Extraction Software Prototype. The cost model assumes that FNS will develop prototype TWILD extraction software for use by retail firms, particularly smaller firms. Participating retail firms will either use the software directly or modify it to meet their specific POS configurations. Presently, there are 6 major POS environments serving large firms and some medium firms. Small retail firms typically use off-the-shelf products with minimal customizations. Large and medium firms often modify the POS software to meet their specific needs. However, as discussed in Chapter 3, the transaction data output from these system are similar. With small retail firms, the dominant POS software produces transaction log files in the proprietary "qdx" format. The cost model assumes that TWILD extraction software development will proceed along the lines of full software development methodology such as the waterfall development.[23]

[23] The Waterfall model is one of several software development methods. It typically involves 1) requirements gathering, 2) software design, 3) implementation, 4) testing, and 5) deployment.

Inputs and Resource Costs: The research team estimated the cost of extracting TWILD from the "qdx" files. Similar software will be required for the other file types. From the POC, the level of effort required to develop prototype TWILD extraction software for the 6 major POS systems is estimated to be 2,000 labor hours.[24]

Develop Matching Algorithm. As an option, FNS could decide to implement a matching algorithm similar to the discussion in Chapter 4.

Inputs and Resource Costs: From the POC, the research team estimates it could take up to 200 hours to develop and implement a complete matching algorithm.

Develop an Onboarding and Certification Process. The TWILD process will require FNS and the state-level processors to bring SNAP-authorized stores into the process. This process is known as onboarding. In addition to the onboarding process, FNS, in concert with the state-level processors, will develop and implement a process of certifying retail firms and/or stores.

Inputs and Resource Costs: This process will largely involve 400 labor hours on the part of FNS.

Implement Storage and Network Infrastructure. FNS will store and manage the consolidated TWILD. Additionally, FNS will be responsible for developing TWILD specifications and working with processors to integrate the TWILD certification process into the current SNAP retailer authorization process. The cost model assumes that FNS will store the data and host a custom application in a FedRAMP-approved cloud environment.[25] FNS will analyze the requirements for TWILD processing and storage. The requirements will reflect the national volume of EBT transactions of SNAP retailers. FNS will be responsible for installation and management of hardware and software for this process.

Inputs and Resource Costs: Exhibit 5.10 presents the implementation cost for FNS to store and manage the TWILD. Storing GBs of data requires an enterprise data storage solution (or data center) that includes provisions for redundancy and backup and recovery. The market rate for this storage solution in 2015 is $30,000 for 4TB of data. Similarly, FNS will need to implement

---

[24] There are 6 major POS systems serving the large firms and some medium firms. There are another 50-60 systems that primarily serve medium firms and a few small firms.

[25] FedRAMP, the Federal Risk and Authorization Management Program, is the result of close collaboration among cybersecurity and cloud experts from the General Services Administration, National Institute of Standards and Technology, Department of Homeland Security, Department of Defense, National Security Agency, Office of Management and Budget, and the federal Chief Information Officers Council and its working groups.

secure high-speed Internet connection to receive the data from processors. The implementation of the infrastructure (storage and network) will require 3,760 blended labor hours.

Exhibit 5.10. Implemenation Costs for FNS Storage of TWILD for a Single State

Step	Activities	Costs			
		Resource	Unit Cost	Units	Total Cost
1	Develop legislation and regulations	Labor Hours	$135	1,880	$253,800
2	Develop TWILD Specifications	Labor Hours	$135	3,760	$507,600
3	Develop TWILD extraction software	Labor Hours	$135	2,000	$270,000
4	Implement Storage Infrastructure	Secure Transmission Network	$30,000	1	$30,000
		Storage	$30,000	2	$60,000
		Labor Hours	$135	3,760	$507,600
5	Develop matching algorithm	Labor Hours	$135	200	$27,000
6	Develop onboarding and certification process	Labor Hours	$135	400	$54,000
Total FNS Implementation Cost					$1,710,000

Processor Development and Implementation Costs

Processors will receive TWILD data from firms, verify the data to ensure compliance with established standards, consolidate the data, and transmit it to FNS. In addition, processors will work with FNS to onboard and certify SNAP-authorized stores to extract and submit TWILD. To perform these operational functions, it is assumed that processors will build network and storage infrastructure, develop software to validate and consolidate TWILD data, and develop processes to onboard and certify SNAP-authorized stores/retail Firms.

Develop Secure Infrastructure. Processors will implement two main types of infrastructure--secure transmission network and storage--to support their TWILD operations. Retail firms will transmit the TWILD they extract to processors by secure FTP (SFTP). Similarly, processors will purchase and implement an SFTP connection with FNS to transmit the final TWILD. The entire transmission infrastructure must be highly secure given the sensitive nature of the data.

Inputs and Resource Costs: In addition to the SFTP connection, processors will need significant storage space to receive, process, and store the TWILD. The cost model estimates the storage space (including redundancies) that will be required by all processors. Using market rates in 2015, setting up network and storage infrastructure will require 880 labor hours, and 12TB of storage space.

Software Development. Processors will develop software to verify and consolidate the data they receive from retail firms. The verification process will ensure that the TWILD they receive from the retail firms conforms to the specified standards. From the POC, the research team estimates that it will require 2,400 labor hours to develop software or an algorithm to validate and consolidate the TWILD from firms.

Onboarding and Certification Process. Working with FNS, processors will develop a process to onboard and certify SNAP-authorized stores (and their retail firms) to be part of the TWILD process. The process will lead to a well-documented manual that processors follow to bring new firms into the TWILD process. Relying on lessons from prior studies, developing the onboarding and certification process will require 800 labor hours.

Exhibit 5.11 summarizes implementation costs for a state-level EBT processor to receive TWILD from firms and transmit them to FNS or an FNS-authorized third party. Notably, the cost model assumes these processes are similar for all states and apply to all firm categories. It is assumed that a state-level EBT processor would need to undertake the same processes whether they serve one firm or multiple firms. Moreover, TWILD processing would be the same regardless of firm size.

Exhibit 5.11. Implementation Cost for State-Level EBT Processor

Step	Activities	Costs			
		Resource	Unit Cost	Units	Total Cost
1	Implement secure infrastructure	Secure transmission network	$30,000	1	$30,000
		Storage space	$30,000	3	$90,000
		Labor Hours	$135	880	$118,800
2	Develop TWILD verification and consolidation software	Labor Hours	$135	2,400	$324,000
3	Develop onboarding and certification standards and process	Labor Hours	$135	800	$108,000
Total Processer Implementation Cost					**$670,800**

Retail Firms Implementation Cost Model

For the cost models, it is assumed that firms are responsible for identification and extraction of TWILD from the EBT transactions processed at store locations. Under this assumption, the firm will be responsible for transmitting data to their assigned processor. It is further assumed that FNS will specify data requirements and provide documentation of the transmission requirements. In addition, FNS will provide prototype TWILD extraction software that firms can use, with or without modification, to extract the TWILD data elements from their transaction

data. Each firm will be wholly responsible for development and implementation of TWILD extraction under these requirements.

Develop a process to consolidate transactions from the registers from each store. Firms will develop a process for consolidating the transactions from all of the registers from all of their stores. Implementation of this process will enable centralization in the next step. The process of consolidation will likely vary by firm size. It is assumed that large and medium sized firms have a process in place to consolidate transaction data, as the research team learned during the POC. Small retail firms, however, usually with one or just a few POS terminals, will need to consolidate data in a central location to extract the TWILD and submit it to the assigned processor.

Inputs and Resource Costs: The main inputs are storage space, network bandwidth, and labor hours to implement the process. From the POC, it was learned that SNAP-authorized firms will generate between 50 and 200 SNAP EBT transactions daily, leading to approximately 278GB of data per month. Infrastructure requirements for consolidating transactions will depend on the size of the firm and its current POS and transaction data configuration. Some firms not currently consolidating their transaction data will expend more resources to establish a consolidation process. Other firms with the infrastructure in place will likely only need to make necessary modifications to meet the performance requirements of the TWILD submission process.

Develop software to extract transactions from native register format into TWILD. Firms will use or modify the prototype TWILD extraction software they receive from FNS to extract TWILD from the store transaction data. The software to extract data from a native format will need customization for each POS software application.

Inputs and Resource Costs: The main inputs for software development are labor hours. The cost model assumes that small firms will make few modifications to the extraction software provided by FNS. However, medium and large firms with customized or proprietary systems may have to make significant customizations to the prototype TWILD extraction software.

Develop infrastructure for additional processing and storage. Firms will need to develop and implement the storage space and transmission infrastructure required to process and transmit the data to their processors. The firms and their processors will define a secure method for batch delivery of extracted TWILD to the processor. This process will have to be separate from the SNAP transaction authorization process to limit interference with authorization.

The firm will implement the SFTP transmission of TWILD. For the cost model, the research team assumed that small firms will contract this process to a TPP to achieve some economies of scale. The firms will assess their TWILD processing and storage requirements, assumed to

reflect the EBT transaction volumes of each store within a retailer. The firm will be responsible for installation and management of hardware and software.

Inputs and Resource Costs: This activity will require the purchase of storage space, secure network bandwidth, and labor hours to implement the storage and network. The expected storage size and network bandwidth will depend on the size of the firm and the number of SNAP EBT transactions. Firms will have to certify that their TWILD extraction is accurate and that the data transmission is secure. FNS will administer the certification process. For cost modeling, it was assumed that FNS will contract certification to EBT processors or TPPs.

After firms have taken the steps to develop an TWILD extraction process, they will be responsible for implementing the extraction system. Implementation activities will differ based on firm size. The remainder of this section discusses implementation activities for different size firms.

Large Firm Implementation Cost Model. The researchers assumed that large firms have more than 20 stores and the most sophisticated IECR systems that operate at a national/corporate level. Moreover, the research team established during the POC that these firms need to have technical resources available to customize POS systems, as well as to manage inventory and other business activities. They currently consolidate transaction data for their management and marketing needs, a subset of which could be repackaged as TWILD. In essence, large firms likely face minimal effort to extract TWILD. Exhibit 5.12 lists the proposed TWILD extraction processes applicable to **each** large firm and their respective costs.

Input and Resource Costs: For the cost model, the research team assumes the NetApp FS800 4TB storage system will be used by large firms. Transferring 288 GB of data will require an extra network bandwidth of 128 megabits per second (Mbs). To implement the storage and network, large firms will require 260 labor hours. For TWILD software development and/or modification, large firms will require 220 labor hours. The research team estimated the required hours based on the extraction software developed during the POC. Finally, large firms will require 260 labor hours to develop/modify and implement the data consolidation process.

Exhibit 5.12. Implementation Cost Model per Large Firm

Step	Description	Resource	Unit Cost	Units	Total Cost
1	Modify process to consolidate data	Labor Hours	$90	80	$7,200
2	Develop/Modify TWILD extraction software	Labor Hours	$90	220	$19,800
3	Implement infrastructure	Secure network connection	$5,000	1	$5,000
		Labor Hours	$90	260	$23,400
		Storage	$30,000	1	$30,000
Total Large Firm Cost per Large Firm					**$85,400**
Total Cost per Store					**$987**

The estimated cost to develop and implement the TWILD extraction process per large firm is $85,400. Spreading this total cost over the average number of stores each large Texas firm operates yields an estimated cost per store of $987.

Medium Firm Implementation Cost Model. Medium firms, as noted, are assumed to operate between 5 and 20 stores. Similar to large firms, medium firms also have fairly sophisticated IECR systems that already aggregate transaction data. While large firms typically aggregate these data at a national level, medium firms aggregate at a regional level. This implies that they will require only moderate modifications to aggregate TWILD regionally and ultimately transmit them to a state-level EBT contractor. Similarly, inventory data are managed at a regional level. Hence, similar to large firms, efficiencies are gained over small firms by being able to deploy changes to POS software from regional headquarters to individual stores. In essence, extraction of TWILD would occur at the regional level. Moreover, again similar to large firms, it is anticipated that medium firms have in-house resources (such as programmers) that can make modifications to POS processes to extract TWILD under guidance from FNS or FNS-authorized third parties.

Input and Resource Costs: Based on findings from the POC (and estimates shown in Exhibit 5.9), medium-size firms are estimated to require 10 GB of data storage. For the cost model, the research team assumes the NetApp FS800 2TB storage system. Transferring 15MB of TWILD data per month will require an extra network bandwidth of 10Mbs to transfer TWILD efficiently to their assigned processors. To implement the storage and network, medium firms will require 80 labor hours. For TWILD software development and/or modification, medium firms will require

220 labor hours. The research team estimated the required hours based on the extraction software developed during the POC. Finally, medium-sized firms will require 160 labor hours to develop/modify and implement the data consolidation process.

Labor is the primary difference in costs between medium and large firms. As shown in Exhibit 5.7 in section 5.2, the cost model assumes that medium-sized firms have lower labor rates than large-sized firms and require fewer labor hours to complete implementation tasks. Exhibit 5.13 lists the proposed TWILD extraction processes applicable to **each** medium firm and their respective costs.

Exhibit 5.13. Implementation Cost Model per Medium Firm

Step		Resource	Unit Cost	Units	Total Cost
1	Develop process to consolidate	Labor Hours	$79	160	$12,640
2	Modify TWILD extraction	Labor Hours	$79	220	$17,380
3	Implement infrastructure	Secure network connection	$2,000	1	$2,000
		Labor Hours	$79	80	$6,320
		Storage	$15,000	1	$15,000
Total Medium Firm Cost					**$53,340**
Total Cost per Store					**$7,887**

The estimated cost to develop and implement the TWILD extraction process for a medium firm is $53,340, over $30,000 less than the cost to a large firm. Spreading the total over the average number of stores each medium Texas firm operates yields an estimated cost per store of $7,887, over 7 times the per store cost of large firms, which are able to spread costs across a much larger number of stores.

Small Firm Implementation Cost Model. Small firms operate fewer than five stores, and most have just one store. These firms likely have less sophisticated IECR systems integrating their inventory data, and are expected to have relatively fewer resources to deploy for extracting TWILD relative to medium and large firms. Furthermore, it is assumed that they will not have the capacity to develop software to extract TWILD from their POS systems. The cost model for small firms therefore assumes that: 1) retailers will receive extraction software at no cost from either FNS or a third-party authorized by FNS; and 2) the software will be implemented by POS vendors.

Input and Resource Costs: A typical small firm will only require resources to manage 5MB of TWILD per month. From the POC, the research team assumed a basic hard drive storage

system of 180GB. The transfer of 5MB of data will require a network bandwidth of 1Mbs. To implement the storage and network, small firms will require 10 labor hours. The research team estimated the required hours based on the extraction software developed during the POC. Finally, small firms will require 20 labor hours to develop/modify and implement the data consolidation process.

Notably, costs for TWILD data storage and transmission to FNS are also assumed to be incurred by third parties working with several small firms to introduce economies of scale, for example TPPs. Exhibit 5.14 lists the proposed TWILD extraction processes applicable to each small firm and their respective costs. The model assumes small firms will incur a one-time service fee to the intermediary processor.

Exhibit 5.14. Implementation Cost Model per Small Firm

Step	Description	Resource	Unit Cost	Units	Total Cost
1	Develop process to consolidate	Labor Hours	$61	20	$1,210
2	Implement TWILD software	Service Fee	$1,500	1	$1,500
3	Implement infrastructure	Secure network connection	$200	1	$200
		Labor Hours	$61	10	$605
		Storage	$100	1	$100
Total Small Firm Cost					**$3,615**
Total Cost per Store					**$3,422**

The estimated cost to develop and implement the TWILD extraction process per small firm is $3,615, considerably less than the estimated cost to large and medium firms. However, because small firms in Texas average 1.1 stores (see Exhibit 5.6), the cost per store, $3,422, remains nearly the same as the cost per small firm.

5.4 TWILD OPERATIONS COSTS

Retail Firms Operations Cost Model

Exhibits 5.15 – 5.17 summarize monthly operational cost estimates to transmit TWILD data for each firm type. Notably, it is assumed that large and medium firms achieve economies of scale as the respective monthly resources expended are applied across more stores. Firms of different sizes, however, are assumed to face different labor costs and have differing labor inputs, and labor rates for operational costs are assumed to be lower than development rates. Exhibit 5.15 shows estimated operational costs for large firms.

Inputs and Resource Costs: Large- and medium-sized firms will primarily incur labor costs to extract, process, and submit their TWILD to their designated processors. In addition, they will incur regular maintenance costs. The maintenance activities include paying a monthly subscription for a secure network connection and replacement costs for storage space. The model assumes that firms will replace storage media every 36 months. Thus, the cost of replacement is spread over 36 months. Using market rates in 2015, replacing storage media will cost a large firm $833 per month and maintaining a secure network connection to processors will cost another $2,000 per month. As Exhibit 5.15 shows, to extract 25MB of data from their systems, large firms will require 24 labor hours per month, similar hours to process the TWILD, and 8 hours to transmit the data to the designated processor. The cost to medium firms will be lower because their labor costs and data volume are lower.

Exhibit 5.15. Operational Costs per Large Firm

Step	Description	Input/Resource	Unit Cost	Units	Total Cost
1	Collect data from storage	Labor Hours	$93	24	$2,232
2	Run TWILD extraction process/software	Labor Hours	$93	24	$2,232
3	Submit TWILD	Labor Hours	$93	8	$744
4	System maintenance (Hardware upgrades, backup, etc)	Secure connection	$2,000	1	$2,000
		Storage	$833	1	$833
		Labor Hours	$93	16	$1,488
Total Monthly Cost					**$8,041**
Annual Cost per Firm					**$96,496**
Annual Cost per Store					**$1,115**

Exhibit 5.16. Operational Costs per Medium Firm

Step	Description	Resource	Unit Cost	Units	Total Cost
1	Collect data from storage	Labor Hours	$73	16	$1,168
2	Run TWILD extraction process/software	Labor Hours	$73	16	$1,168
3	Submit TWILD	Labor Hours	$73	8	$584
4	System maintenance (Hardware upgrades, backup, etc)	Secure connection	$200	1	$200
		Storage	$417	1	$417
		Labor Hours	$73	8	$584
Monthly Cost per Firm					**$3,537**
Annual Cost per Firm					**$42,440**
Annual Cost per Store					**$6,275**

Monthly operational cost estimates for small firms are limited to collecting, extracting, and submitting the data to EBT processors. Small (and some medium) firms can achieve economies of scale by contracting with a third party authorized by FNS to collect, extract, and submit TWILD to state-level contractors. TPPs already handling upstream functions (receiving, validating, and consolidating TWILD) may decide to offer services to small firms to collect, extract, and submit TWILD, creating a streamlined and vertically integrated system for handling TWILD. For example, FNS lists 30 TPPs that can provide EBT equipment and services.[26] These TPPs are in the best position to provide such services, given their familiarity with FNS and POS systems. Other likely candidates to provide these operational functions include retail technology solutions companies that provide POS solutions and technical assistance to stores.

The cost model assumes small firms will incur negligible costs in maintaining storage space. However, at 2015 market rates, small firms will need to maintain a reliable secure network connection to transmit the data to their TPPs. In addition, FNS (or an authorized third party) is assumed responsible for maintaining the hardware and software provided to small firms for extracting TWILD. As shown in Exhibit 5.17, annual operational costs for a small firm are estimated at $8,520.

Exhibit 5.17. Operational Costs for Small Firms

Step	Description	Resource	Unit Cost	Units	Total Cost
1	Collect data from storage	Labor Hours	$26	6	$153
2	Run TWILD extraction process/software	Labor Hours	$26	6	$153
3	Submit TWILD	Labor Hours	$26	8	$204
4	System maintenance (Hardware upgrades, backup, etc)	Secure connection	$200	1	$200
		Storage	-	-	$0
		Labor Hours	-	-	$0
				Monthly Cost per Firm	**$710**
				Annual Cost per Firm	**$8,520**
				Annual Cost per Store	**$8,065**

Processor Operations Cost Model

State-level processors will receive and validate the TWILD, and then transmit it to FNS. Exhibit 5.18 outlines monthly operational costs for an EBT processor. Operational activities include: 1)

[26] http://www.fns.usda.gov/sites/default/files/snap/SNAP-EBT-Third-Party-Processer-List.pdf

receiving, validating and consolidating TWILD from firms; 2) submitting TWILD to FNS; 3) maintaining an infrastructure; and 4) providing technical support to firms.

Receive, validate, and consolidate TWILD: Processors will receive TWILD from firms at regularly scheduled intervals. Upon receiving the TWILD, they will execute the validation software or process to ensure the integrity of the data. Validation also includes formatting the data to ensure compliance with FNS-defined TWILD specifications.

Inputs and Resource Costs: Processors will primarily incur labor, storage replacement, and secure network connection costs in processing the TWILD. Assuming daily batch data transfer within a state, processors will receive about 314 GB of data per month. The cost model provides for 80 hours of labor to monitor the transfer process and validate and consolidate the data. In addition, another 24 hours will be required to submit the data.

it is also assumed that processors will provide support to firms. The model assumes that on average, processors will provide one-tenth of all firms with remote support. Processors will also work with FNS to onboard, certify, and integrate new firms into the TWILD process.

Submit TWILD to FNS. Processors will transmit the data to FNS in batch mode periodically. The cost model assumes that an EBT processor will consolidated TWILD from 14,953 Texas authorized stores.

Inputs and Resource Costs: The main input for the submission to FNS will be 24 labor hours.

Maintenance and support:

Exhibit 5.18. Monthly Operational Costs per State-Level EBT Processor

Step	Description	Costs			
		Resource	Unit Cost	Labor Units	Total Cost
1	Receive, validate, and consolidate TWILD	Labor Hours	$90	80	$7,200
2	Submit TWILD to FNS	Labor Hours	$90	24	$2,160
3	Maintenance	Secure transmission network	$4,000	1	$4,000
		Storage	$5,000	1	$5,000
		Labor Hours	$ 90	32	$2,880
4	Provide ongoing support to firms; onboarding, certification, integration, help desk.	Labor Hours	$90	200	$18,000
				Total Monthly Cost	**$39,240**
				Total Annual Cost	**$470,880**

FNS Operations Cost Model:

Exhibit 5.19. Operational Costs for FNS Storage of TWILD for a Single State (Texas)

Step	Description	Costs			
		Resources	Unit Costs	Units	Total Cost
1	Maintain and update TWILD specification	Labor Hours	$135	24	$3,240
2	Maintain small firm TWILD software	Labor Hours	$135	40	$5,400
3	Maintain storage and environment	Secure transmission network	$2,000	1	$2,000
		Storage	$2,000	1	$2,000
		Labor Hours	$135	160	$21,600
4	Onboarding, certification, and integration	Labor Hours	$372	124	$46,128
				Total Monthly Cost	**$80,368**
				Total Annual Cost	**$964,416**

5.5 APPLICATION: SINGLE STATE COSTS

Exhibit 5.20 provides a summary of all costs associated with extracting, transmitting, and storing TWILD data for a single state (Texas) for each stakeholder in the first year of operations. Note that firm costs are associated with setting up and operating the TWILD process for all firms in a single state in the relevant firm size category.

Exhibit 5.20. Implementation and Operational Costs for a Single State (Texas)

Stakeholder	Implementation	First Year Operational Costs	Total First Year Costs	Share of Total First Year Costs
Large Firms	$10,931,200	$12,351,488	$23,282,688	27%
Medium Firms	$4,960,620	$3,946,920	$8,907,540	10%
Small Firms	$24,976,035	$25,151,272	$50,127,307	58%
Processors	$670,800	$470,880	$1,141,680	1%
FNS	$563,400	$2,844,480	$3,407,880	4%
Total	**$42,102,055**	**$44,765,040**	**$86,867,095**	**100%**

5.6 NATIONAL COSTS OF TWILD COLLECTION

This section extrapolates the POC cost models to the national level. The extrapolation uses the following information from the POC: the estimated number of daily SNAP transactions and SNAP redemptions, and the advantages of economies of scale that retailers, processors, and FNS can exploit.

Daily SNAP Transactions: From the POC, IMPAQ estimated that large retailers (such as the national retailer that provided data for the POC) generate an average of 361 SNAP EBT transactions each day. Using the TWILD specifications from Chapter 4 and results from the POC, these 361 transactions require 1.5MB of storage space. Exhibit [5-20] presents the estimated monthly storage requirements for all stakeholders. Firms will only need just the amount of storage sufficient to process the monthly TWILD. Processers will need two times the storage space as the Firms while FNS, who will be storing the TWILD, will require 100TB of storage per month. The storage requirements for FNS includes considerations for backup and redundancies as described earlier.

Exhibit 5.21. Average Daily Transactions and Storage File Sizes

Firm Size	Monthly TWILD Size (GB)	Required Storage	
		Processers (GB)	FNS (TB)
Large	3,842	7,684	92
Medium	125	251	3
Small	217	433	5
All Firms	4,184	8,368	100

As shown, extrapolating the daily transactions to the national level using the total number of SNAP authorized stores indicates that 519kb of data per firm will lead to three terabytes of data (4TB) each month, requiring 100TBs of storage. The cost of extracting, processing, transmitting, and storing TWILD data of this magnitude forms the basis of the national cost extrapolation.

Economies of Scale: Currently, of the more than 250,000 SNAP-authorized stores, 201,000 have IECR cash register technology. However, a significant portion of these are part of large firms (see Exhibit 5.2). From IMPAQ's experience, these retailers can spread their implementation and operational costs across all their stores, reducing the cost per store for developing and operating the TWILD transmission process. Similarly, national EBT processors will benefit from economies of scale, as currently happens with the EBT ALERT process executed by four national companies.

SNAP Redemptions: SNAP redemptions serve as a proxy for the anticipated volume of TWILD data generated by each state. A key cost assumption is that state-level EBT processors will negotiate a similar SNAP cost per case month (CPCM) with states in a competitive market, with SNAP redemptions serving as a sufficient proxy for TWILD volume and cost per case. While the number of SNAP households is an important metric, TWILD extraction is largely driven by actual redemptions (transactions), not number of households.

The remainder of this section presents national cost estimates for all three stakeholders in the TWILD process – Firms, TWILD EBT processors, and FNS.

National Firms Costs for Extracting TWILD

Implementation Costs

Exhibit 5.22 details estimated implementation costs by category for each state. The total number of SNAP-authorized stores in each state is drawn from STARS. In Chapter 5, IMPAQ determined that the distribution of large, medium, and small firms in Texas was 4, 3, and 93 percent, respectively. These proportions were applied to the total number of SNAP-authorized stores in each state, assuming distributions of stores by firm size are similar nationwide. Chapter 5 also includes implementation costs for Texas SNAP-authorized retailers by retailer category. These costs are applied to the number of retailers in each state to estimate total retailer costs nationwide. Overall, we estimate it will cost firms $377 million to develop and implement the infrastructure and processes to collect TWILD from SNAP-authorized stores.

Exhibit 5.22. National Firm Implementation Costs

Firm Size	No. of Firms	Cost per Firm	Implementation Costs
Large	1,708	$85,400	$145,842,104
Medium	1,237	$53,340	$65,970,725
Small	43,302	$3,615	$156,537,393
Total	**46,247**	**$7,966**	**$368,350,223**

National Annual Operational Costs

Annual operational costs were estimated to average $96,456 per large and $42,444 per medium firm in Texas. Small firm annual costs were estimated to be $8,520.[27] These costs are estimated at $1,115, $6,295, and $8,066 per large, medium and small store respectively. Costs by firm size are assumed to be similar for all states. Exhibit 5.23 summarizes national annual operational costs by firm size category. Total operating costs simply multiply the number of firms by the per firm cost estimates for Texas, assuming the distribution of firms in Texas is representative of the entire U.S.

Exhibit 5.23. Annual Operational Costs by Firm Size

Stakeholder	Number of Firms	Number of Stores	Monthly Operational Costs	Annual Operational Costs	Annual Cost Per Store
Stores – Large	1,708	147,774	$13,732,611	$164,791,332	$1,115
Stores – Medium	1,237	8,365	$4,374,137	$52,489,644	$6,275
Stores – Small	43,302	45,747	$30,744,550	$368,934,604	$8,065
Total	**46,247**	**201,886**	**$48,851,298**	**$586,215,579**	**$2,904**

Firms are estimated to spend $586 million annually to operate the TWILD collection process. Each store, on average, will spend $2,904 per year to extract and transmit TWILD to FNS. However, the bulk of the costs on both a per-store and total basis, would accrue to small firms.

Processer Costs to Receive and Transmit TWILD

Exhibit 5.24 details EBT processor costs to receive and transmit TWILD. These costs are provided for the development, implementation, and operation of necessary systems by the state. Estimated EBT processor implementation costs for a single state (Texas) were $670,800.

[27] The cost model assumes that small retailers will bear the cost of implementing an IECR environment as required by the 2014 Farm Bill.

Annual operational costs were estimated using a similar formula and Texas' monthly operational costs of $39,240. The cost model presents four scenarios, each representing different levels of economies of scale. With a single national processor, the cost model implies that implantation and operational costs could be spread over all firms, thereby significantly reducing overall costs. These advantages diminish as the number of processors increases. Exhibit 5.24 shows that with only one national processor, the annual cost of operations will be just over $4.4 million, rising to $6.4 million with 10 processors.

Exhibit 5.24. National Processor Operational Costs

Number of Stores	Data Size (GB) Per Month	1 Processor	4 Processors	10 Processors
46,247	**8,368**	**$4,429,437**	**$5,062,213**	**$6,442,817**

National FNS TWILD Costs

Exhibit 5.25 shows the costs to FNS to implement the TWILD collection process nationally, handling 4TBs of data each month, which requires 100TBs of storage (see Exhibit 5.20). The national implementation costs are primarily influenced by the storage costs.

Exhibit 5.25. National FNS Implementation Costs

Step	Activities	Costs			
		Resource	Unit Cost	Units	Total Cost
1	Develop legislation and regulations	Labor Hours	$135.00	1,880	$253,800
2	Develop TWILD Specifications	Labor Hours	$135.00	3,760	$507,600
3	Develop TWILD extraction software	Labor Hours	$135.00	2,000	$270,000
4	Implement Storage Infrastructure	Secure Transmission Network	$30,000.00	1	$30,000
		Storage	$30,000.00	24	$720,000
		Labor Hours	$135.00	7,920	$1,069,200
5	Develop matching algorithm	Labor Hours	$135.00	3,960	$534,600
6	Develop onboarding and certification process	Labor Hours	$135.00	3,960	$534,600
Total FNS Implementation Cost					$3,919,800

The national operations cost model assumes FNS, represented by a single contractor, will manage the TWILD storage. Therefore, FNS will benefit from the same level of economies of scale as in the case of a single national processor. Using this assumption, the cost model estimates that it will cost FNS $2.8 million each year to operate the TWILD storage. Exhibit 5.26 shows the national operations cost to FNS.

Exhibit 5.26. Summary of National Costs for Extraction of TWILD

Step	Description	Costs			
		Resources	Unit Costs	Units	Total Cost
1	Maintain and update TWILD specification	Labor Hours	$135	288	$38,880
2	Maintain small stores TWILD software	Labor Hours	$135	480	$64,800
3	Maintain storage and environment	Secure transmission network	$2,000	2.4	$4,800
		Storage	$30,000	4	$120,000
		Labor Hours	$135	15360	$2,073,600
4	Onboarding, Certification, and Integration	Labor Hours	$135	3840	$518,400
				Total Annual Cost	**$2,820,480**

Total National Costs

Exhibit 5.27 summarizes the national costs from all three stakeholders using information from Exhibits 5.21 through 5.26. Overall, the cost model estimates that it will cost all stakeholders $377 million to implement the infrastructure necessary for TWILD collection. Additionally, in the first year, stakeholders will spend a combined $970 million to operate the TWILD process. The exhibit also shows the share of total first year of operations. Collectively, small firms will bear the largest share (54%) of the implementation and first year operations cost. The research team expects the annual operational cost to reduce significantly after the second year of operations due to retailers becoming more efficient and experienced.

Exhibit 5.27. Summary of National Costs for Extraction of TWILD

Stakeholder	Implementation Costs	First Year Annual Operational Costs	Total First Year Costs	Share of Total Costs
Retailers – Large	$145,842,104	$164,791,332	$310,633,436	32%
Retailers – Medium	$65,970,725	$52,489,644	$118,460,369	12%
Retailers – Small	$ 156,537,393	$368,934,604	$525,471,997	54%
Processers	$5,048,400	$4,429,437.23	$9,477,837	1%
FNS	$3,919,800	$6,421,254.57	$10,341,055	1%
Total	**$377,318,423**	**$597,066,271**	**$974,384,694**	**100%**

Chapter 6. Study Findings and Recommendations

This chapter applies the lessons learned during the POC (as described in Chapters 3 and 4) to assess the relative feasibility of the four proposed technical solutions (as outlined in Chapter 2). Chapter 6 also concludes the report with recommendations for next steps.

6.1 FEASIBILITY OF THE TECHNICAL SOLUTIONS

This section presents an assessment of each proposed technical solution on the merits of both infrastructure and technology.

Technical Feasibility of Solutions 1 and 2: Transmission Using Existing Infrastructure

Technical Solution 1 uses existing infrastructure to transmit item-level data through EBT processors and TPPs in real time. These processors, or an approved contractor, could decrypt transaction or item-level data from retailers for FNS, if necessary. Discussions with industry stakeholders during the POC phase of the study established that decryption is unlikely to be necessary, as most retailers protect SNAP EBT card identifiers by masking the middle six digits, rather than encrypting the transaction. This approach is considered PCISSC-compliant.

Technical Solution 2 also uses existing infrastructure, but transmits item-level data in batch mode, rather than in real time. Advantages and disadvantages of real-time versus batch transmission include:

- Real-time transmission of data to EBT processors (Technical Solution 1) places an additional data load burden on retailers and contractors. Each entity would need to increase data storage and transmission capacity to accommodate TWILD, resulting in files larger than typical SNAP EBT authorization message packets—even if the storage is only temporary. In addition, item-level transmissions in real time are necessarily transmitted at the same time as EBT authorization and other transactions. This is likely to have a major effect on data transmission speed, leading to slower authorization feedback.
- Both real-time and batch data transmission will necessitate software modifications on the part of retailers, TPPs, and EBT processors. Retailers need to modify their POS software to transmit TWILD to TPPs, which in turn need to modify their software to

process and transmit TWILD to EBT processors. While software modification is necessary for both real-time and batch processing, separating TWILD transmission from EBT authorization in batch transmission reduces the data load burdens on the infrastructure, ensuring greater authorization efficiency.

- Batch transmission options require all stakeholders to increase storage capacity to accommodate storing item-level transaction data over extended periods. Having retailers store transaction data along with EBT card numbers and other identifiers may add privacy and fraud concerns, although storing partially masked EBT card numbers mitigates this concern.
- The POC informed several of these potential issues:
  - Large retailers are well equipped to manage large data files.
  - The predominant method of storing SNAP EBT card identifiers to comply with PCISSC standards is to mask the middle six digits rather than to encrypt the data.
  - Retailers that supplied data for the POC did not store SNAP EBT identifiers along with transaction records. They combined these data only for the purposes of this study.
  - Retailers completed transmission of the combined data using encrypted SFTP servers, a process that can be recommended as a matter of policy.

Both Technical Solution 1 and Technical Solution 2 have the advantage of using existing infrastructure for data transmission, thus offering potential cost and time savings as compared to solutions that require new infrastructure. However, this analysis reveals that batch transmission is the preferable approach. Real-time transmission is not technically feasible given the current technology, infrastructure, and industry expectations for transaction authorization speed and efficiency. Moreover, Technical Solution 2 implies the participation of TPPs, requiring them to have redundant infrastructure and services to support the pass through transmission of TWILD from retailers to EBT processors. While TPPs offer some value in providing retail management services, that service may not warrant the investment necessary to support TWILD collection.

Feasibility of Technical Solution 3: New Transmission Infrastructure Using Electronic Messaging

Technical Solution 3 proposes a new transmission route for TWILD. Specifically, standardized email messages containing TWILD from retailers sent directly to FNS, avoiding the need to

involve TPPs and EBT processors. This solution also eliminates the challenge associated with real-time data transmission overburdening transmission lines with large data loads during transaction authorization.

On the surface, this technical solution appears technically feasible. However, challenges unique to this solution render it infeasible given current technology, infrastructure, and industry preferences. In current electronic message applications such as Square[28], or at retailer register lanes, retailers often offer customers the option of receiving their receipt via email. This process requires converting the transaction log details for that particular transaction to an email message, accounting for data packet size limitations, security, and processing efficiency. The email is delivered to the individual customer's email inbox. Some retailers appear to have the capacity to generate an email message for each transaction.

Receipt emails, however, get distributed to individual email inboxes. To illustrate: 1,000 transactions made by 1,000 unique customers would generate 1,000 emails delivered to 1,000 inboxes. In the TWILD collection scenario, those 1,000 email messages would be sent to a single location for processing. It is estimated that Texas had more than 180 million EBT transactions in 2014.[29] An email message produced for each of those transactions would result in approximately 500,000 emails per day from Texas alone. Each email would contain additional metadata (e.g., email header) that would wrap around the TWILD data—which would increase the quantity of the data (size of the message) when compared to the same record inside a batch file. That overhead in each email message is significant when multiplied by the number of transactions generated nationally on a daily basis. While it is feasible to build a mechanism to collect and process a large number of email messages, this solution is unlikely the most efficient to implement on a national scale.

In conclusion, although intriguing and potentially feasible on a small scale, this technical solution is not optimal as the primary TWILD capture mechanism on a national scale, although additional research could be performed to identify appropriate scenarios for this solution.

Feasibility of Technical Solution 4: New Transmission Infrastructure Using Batch Processing

Technical Solution 4 is similar to Technical Solution 2 in that it uses batch processing, and to Technical Solution 3 in that it uses a new transmission route. New transmission paths proposed in this solution would be high-speed Internet transmission lines, similar to the current pathways

[28] https://squareup.com/

[29] 2014 STARS data

via TPPs and EBT processors. Rather than burdening TPPs by adding to their current operations, this solution introduces potentially new contractors at the state, regional, or national level to manage TWILD receipt, processing, and transmission to FNS. Existing state contractors (such as the EBT processors) would not be excluded in providing the TWILD collection services in addition to their current responsibilities. They would likely have existing capabilities (e.g., ALERT data submission experience) and resources that could be leveraged for this work. However, due to the additional operational requirements, the EBT processors would likely need to enhance their infrastructure and develop new data collection capabilities. Other contractors could possibly compete in this space, which would offer more competition for the Government's TWILD collection services.

Overall, Technical Solution 4 is the most technically feasible because it offers the advantages of batch transmission outlined under Technical Solution 2, while bypassing the technical disadvantages associated with email described under Technical Solution 3.

Summary of Technical Feasibility

Although real-time data transmission is an intriguing idea, it poses several challenges that could potentially burden retailers, TPPs, and the custodian of the stored data (whether FNS or an FNS-contracted vendor). The current infrastructure is not optimized to handle large volumes of data from the retailers to EBT processors, especially not as part of the SNAP EBT transaction authorization process. This makes batch processing and transmission preferable to real-time transmission.

In addition, implicit in this conclusion is that PCISSC compliance would be limited to masking the SNAP EBT card number to facilitate matching with ALERT data (as discussed in Chapter 4). If retailers encrypt SNAP EBT transactions using encryption keys, EBT processors would have to transmit TWILD because they would hold the decryption keys.

In conclusion, there are two technically feasible data transmission solutions: Technical Solutions 2 and 4. These solutions transmit TWILD in batch mode and vary only in whether they use existing or new transmission pathways. Technical Solution 4 is preferrable over Technical Solution 2 because it is deemed better to establish new transmission paths than to over-burdening existing paths.

6.3 RECOMMENDATIONS FOR NEXT STEPS

An important next step would be to validate the technical and cost implications of Technical Solution 4 on a larger scale via a pilot project. The pilot could also explore emerging technologies as a means to lower the burden on firms to collect the TWILD.

Two-Stage Pilot

A two-stage pilot of Technical Solution 4 would provide invaluable information. During the first stage, FNS could commission a pilot of this solution using several representative IECR stores and a limited number of small firms. Specifically, FNS might seek the cooperation of a national retailer, asking it to recruit stores to pilot over a set period of time. Similarly, FNS could seek participation of smaller IECR retailers selected for their use of a variety of POS hardware and software configurations. These smaller retailers would provide insight into the challenges smaller retailers would have in extracting TWILD. In essence, the first stage of the pilot would involve a number of selected retailers charged with the responsibility of piloting TWILD extraction and submission in a real-world environment.

The second stage of the pilot would build upon the first by adding more retailers and engaging with a state-level EBT processor or other third-party to transmit the TWILD to FNS for final storage. This stage's primary purpose would be to refine the technical and cost implications of a national data collection effort.

Emerging Technologies

The pilot could also serve as an opportunity to test emerging big data technology and data science approaches associated with data collection such as machine learning. Approaches being developed in these areas may prove useful for finetuning the technical solution for TWILD collection and may provide insight for other existing FNS data management and analytic activities.

Collecting item-level data on food purchases made by SNAP households is a true "Big Data" problem. Big data refers to data that are challenging to capture, store, manage, and analyze due to its volume, velocity, and variety. Eight million EBT transactions are processed daily nationally. If item-level data are captured from all EBT transactions, this activity would generate up to 500 gigabytes (GB) of new data per day or approximately 100 terabytes (TB) per year.

Over 250,000 stores are authorized by FNS to accept SNAP via EBT. The stores have different POS and inventory systems which house the transaction data in a range of formats. The biggest variety of data formats would be generated by small firms as they represent 94% of all firms with

SNAP authorization. While small firms, as defined for this study, are estimated to represent between 5-10% of all SNAP redemption dollars, as indicated by the national cost estimates, small firms would account for 44% of total firm implementation costs and 63% of total firm operational costs. The effort to collect and format the TWILD by small firms would be disproportionate when compared to large and medium sized firms. A mixed approach is recommend for small retailers for the pilot in which big data technologies and machine learning methods are explored to lower the data collection burden.

Appendix A. Basic Requirements Gathering Process

The purpose of this project was to assess the feasibility of collecting POS data on food items purchased by SNAP recipients. The first step was to identify the basic requirements for a FNS-owned data capture system and database for storing SNAP purchase data.

The gathering process was informed through interviews with FNS personnel and other important stakeholders and reviews of documentation related to the current EBT environment. Personnel from FNS Office of Policy Support (OPS), Retailer Policy and Management Division (RPMD), Supplemental Food Programs Division (SFPD), and Office of Information Technology (OIT), as well as the USDA Economic Research Service (ERS), participated in interviews and shared relevant documentation. As part of this effort, we also interviewed personnel from EBT and TPPs. This appendix outlines the data and preliminary system requirements established during the gathering process.

DATA REQUIREMENTS

FNS identified three major expected uses of SNAP purchase data collected at the POS: policy analysis and program development, research and reporting, and fraud and abuse detection.

Policy Analysis and Program Development

How SNAP households use their SNAP benefits has important policy implications, and understanding their use could influence administration of the SNAP program and distribution of SNAP benefits. Specifically, POS data could be used to gain insight into household expenditures on SNAP-eligible items; household weekly, monthly, and annual purchase patterns; what foods are purchased together during a transaction; and how SNAP recipients use multiple forms of tender to purchase SNAP-eligible commodities. These insights, in turn, could be used to improve SNAP program administration.

Research and Reporting

In addition to informing policy analysis and program development, POS data on purchases using SNAP benefits could be used to conduct research on numerous topics, including access to nutritious food, dietary intake by food classifications and subgroups (such as those in the

USDA Dietary Guidelines), and seasonal trends in food consumption. The POS data elements used to inform research and report functional requirements are generally the same as those for policy analysis and program development with one notable difference, observed by stakeholders from the Center for Nutrition Policy and Promotion: the research need to link POS data to existing nutritional information databases. Key to enabling this link would be the UPC, UPC description, and PLU data elements. In addition, linking the data to state FNS agency databases would enhance state-level analyses to reveal geographic variation in purchase patterns of SNAP households.

Fraud and Abuse

A key FNS responsibility pertaining to SNAP is to ensure that taxpayer dollars are spent wisely. To that end, FNS and the states undertake a variety of activities to combat fraud and abuse by retailers and recipients. Fraud involves misrepresenting circumstances to receive benefits or to become an authorized retailer; abuse involves the misuse of benefits, often by selling SNAP benefits for cash (a process known as trafficking). FNS currently uses the ALERT dataset to help meet this functional requirement. ALERT contains data on every transaction received by EBT processor systems: debits and purchases, refunds, voids, rejections due to insufficient funds, and balance inquiries. Each transaction record contains variables tracking the retailer, the EBT beneficiary, and transaction monetary details. However, the ALERT system does not collect detailed data on the individual items purchased—just transaction amounts. Detailed POS data, linked to the ALERT system using the EBT ID, could be used to enhance the algorithms developed to detect fraudulent patterns.

PRELIMINARY SYSTEM REQUIREMENTS

The preliminary system requirements for the FNS-owned data capture system and database for storing SNAP purchase data were collected through a set of interviews with FNS Office of Research and Analysis (ORA) and OIT staff. See Exhibit A.1. This section expands on the data requirements of this feasibility study by including the architectural requirements, operational requirements, operational scenarios, assumptions, and constraints as interpreted from conversations with OPS and OIT staff. We outline considerations that need to be addressed in designing and building a SNAP data collection mechanism on a national scale. The scope of the system requirements focused around execution of the project's tactical objectives to develop technical solutions.

Exhibit A.1. Basic Requirements Gathering Interviewees

Name	Agency or Organization	Position
Kelly Kinnison*	USDA FNS OPS	Branch Chief, SNAP Analysis Branch
Kathy Ottobre	USDA FNS RPMD	Program Analyst, Retailer Management and Issuance Branch
Erin McBride	USDA FNS SFPD	Technology Specialist
David Smallwood**	USDA ERS	Branch Chief, Food Assistance Branch
Shelly Pierce	USDA FNS RPMD	Branch Chief, Retailer Operations Branch
Allison Wilcox	USDA FNS OIT	Branch Chief, Program Management Branch
Kevin Russ	USDA FNS OIT	Portfolio Management Division
Joe Koss	USDA FNS OIT	Branch Chief, Applications Development Branch
John Coulter	USDA FNS OIT	Program Manager
Cary Jeffers	FIS Global	Division Product Manager
Ed Koslow	JP Morgan Chase	Account Manager
George Gogol	First Data	Product Director
Kelli Statler	First Data	EBT Compliance Manager, First Data Payments Compliance

* Meeting with Kelly Kinnison included Anita Singh, Kathryn Law, and Barbara Murphy.

** David Smallwood organized a panel interview with the staff of the ERS Food Assistance Branch.

Exhibit A.2 contains a framework for defining specific system requirements for the SNAP data collection system. The description column for each category includes baseline requirements.

Exhibit A.2. Draft System Requirements for SNAP Data Collection System

System Requirement Category	System Requirement Description
Application Architecture	The SNAP purchase data collection system needs to integrate into existing stakeholder systems and processes. The architecture should demonstrate both scalability to support the processing of large volumes of transactional data and flexibility to be deployed in different environments. The architecture will incorporate data capture solutions and data transmission solutions for the POS environment, communication environment, and data storage and retrieval environment.
Administration, Configuration, and Account Management	The system should be administered by operation staff through intuitive routines that do not require extensive programming or changes to the system. Data load schedules should be configurable by production support personnel. The system should accommodate new user requests and support changes to existing accounts, including user deactivation.
Audit	The data captured from the transaction (date and time, SNAP beneficiary ID, and SNAP store ID) will provide visibility about the transaction. The system should capture basic information about the sender. For example, if TPPs or EBT processors are involved in the transmission process, the system should capture basic information about them.
Availability	ALERT data are currently available in the database the morning after they are submitted by the EBT processors. The submission schedule will be a factor in determining how soon data are available to users. Reporting cycles can follow timeframes similar to those of ALERT.
Capacity and Performance	The performance requirements of the POS component to transmit the data should be in line with SNAP and Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) standards, if a real-time option is selected. The system should be able to store all data submitted across multiple years for all SNAP retailers selected to participate in the data collection process. The ALERT system can be used to project capacity and performance requirements. The ALERT system, for example, has around 12 terabytes of physical storage attached to a ProLiant DL980 G7 8 processor server with almost 400 MB of memory.
Communications	The data collection system should leverage existing EBT communication infrastructure. EBT message standards govern the communication between retailers and EBT Processors. The data need to be collected in a format that is consistent with existing messaging standards. The message layout may need modification to accommodate item-level detail. The MOVEit Managed File Transfer tool is currently used by Processors to transmit ALERT files on a nightly basis. The solution to transmit the item-level data from the POS systems to a storage system should outline the responsibility and burden of each stakeholder in the EBT process. Existing FNS technology should be reused when possible.
Data Archive	Data archival policy should follow best practices defined by FNS for other systems.

System Requirement Category	System Requirement Description
Infrastructure	Microsoft SQL Server is the enterprise relational database platform currently used by enterprise applications like ALERT and STARS. The National Information Technology Center (NITC) is the preferred data center provider for any new USDA enterprise production system. Waivers can be granted if the NITC infrastructure cannot support system demands. The development environment can be hosted by OIT at FNS headquarters and would need to be accessed through the FNS network on Government-owned equipment.
Integration	The technical solution should be integrated into a POS environment, making collection of item-level data is transparent to the participant. The solution can leverage existing integration platforms between retailers and TPPs and EBT processors by establishing standardized messaging protocols to transfer the data payload. The solution should support linking to current FNS systems such as ALERT and STARS.
Monitoring	Recommendations for monitoring a production data collection system can be included in the technical solutions report.
Policy	The solution needs to comply with federal and state laws, regulations, and policies. Recommendations to modify existing policy (if necessary) need to be provided in the technical solutions report.
Security	ALERT is a *moderate* level system as defined in FIPS 199 Standards for Security Categorization of Federal Information Systems. Develop the technical solution design under the same security standards if sensitive information such as Social Security numbers or date of birth is collected. Some states use Social Security numbers for SNAP case numbers.

Appendix B. ANSI X9.58 2007 Message Standards

The data interchange that takes place between different stakeholders involved in authorizing a SNAP EBT transaction needs to follow standard formats for integration, exchange, and interoperability. The ANSI X9.58 standard provides technical specifications for exchanging transaction messages between the retailer and an EBT processor. The ANSI standard, which specifies message structure and data elements used in SNAP transactions, is based on the International Organization for Standardization (ISO) 8583 interchange specifications for messages originated by financial transaction cards; the X9.58 standards are tailored specifically for EBT transactions. Exhibit B.1 outlines the data elements regulated by these standards.

Exhibit B.1. ANSI X9.58 2007 Message Standards

Name	Description	Format	Representation	Notes	Bit
Account type	Code used to describe the accounts affected; contained in processing code and amounts, additional		an 2	see Annex B.1	3-2 3-3 54-1
Acquiring institution identification code	Code identifying the acquirer	LLVAR	n..11	Shall remain unchanged for all messages in a transaction, see C.6.5	32
Acquirer trace data	Data provided by the acquirer to be returned unaltered in the response message	LLLVAR	ans..100		127
Actual amount, transaction	Completed transaction; contained in replacement amounts		n 12	see 6.2.8	95-1

Name	Description	Format	Representation	Notes	Bit
Actual amount, settlement	Completed settlement amount; Contained in replacement amounts		n 12	see 6.2.8	95-2
Actual amount, transaction fee	Completed transaction fee; contained in Replacement amounts		x + n 8	see 6.2.1, 6.2.8	95-3
Actual amount, settlement fee	Completed settlement fee; contained in Replacement amounts		x + n 8	see 6.2.1, 6.2.8	95-4
Additional data	Data indicated by the Additional data dataset identifier in Additional data private, acquirer		an…17	see 6.2.2, B.12	111-3
Additional data dataset identifier	Code identifying the type of data contained in Additional data private, acquirer		a 2	see 6.2.2, B.12	111-1
Additional data length	Value indicating the length of the data contained in Additional data private, acquirer		n 3	see 6.2.2, B.12	111-2
Additional data private, acquirer	Additional information for EBT Food stamp transactions	LLVAR	ans...255	see 6.2.2, B.12	111
Additional response data	Other data required in response to an authorization or other transaction request	LLVAR	an...25	see 6.2.3	44

Name	Description	Format	Representation	Notes	Bit
Additional response data additional data	Any other additional data sent in a response in the Additional response data element		an...22	see 6.2.3	44-3
Additional response data bit error code	When Response code is equal to "30" identifies the bit position of the data element in error		n 3	see 6.2.3	44-1
Additional response data referral phone number	Phone number to be contacted for a voice authorization referral		n 10	see 6.2.3	44-2
Advice/reversal reason code	Codes identifying the reason for a Reversal or an advice		an 6	see Annex B.2	60
Advice/reversal reason code byte map	Code indicating which sub-elements of Advice/reversal reason code are present		an 2	see Annex B.2	60-1
Advice/reversal reason code reversal reason	Reason a Reversal message was sent		an 2	see Annex B.2	60-2
Advice/reversal reason code advice reason	Reason an advice message was sent		an 2	see Annex B.2	60-3
Amount, amounts additional	Amount contained in Amounts, additional data element		x + n 12	see 6.2.4	54-3

Name	Description	Format	Representation	Notes	Bit
Amount type, amounts additional	Identification of kind of amount contained in Amounts, additional data element and Processing code		an 2	see 6.2.4 Annex B.3	3-2 54-2
Amount, transaction	Funds requested by the cardholder in the local currency of the acquirer or source location of the transaction		n 12		4
Amount, transaction fee	Fee charge for transaction activity in the currency of the transaction amount		x + n 8	see 6.2.1	28
Amount, transaction processing fee	Fee charged for handling and routing of messages in the currency of the transaction amount		x + n 8	see 6.2.1	30
Amounts, additional	Information on up to six amounts and related account data for which specific data elements are not defined	LLLVAR	ans..120	see6.2.1 and 6.2.4	54
Approval code	Code assigned by the authorizing institution indicating approval		an 6		38

Name	Description	Format	Representation	Notes	Bit
Authorization life cycle	Value in calendar days, hours or minutes that defines the time period for which the acquirer is requesting guarantee of funds, or for which the card issuer shall guarantee funds for a Financial Presentment transaction which may follow		n 3	see Annex B.4	57
Authorizing agent institution identification code	Code identifying the authorizing agent institution	LLVAR	n..11		113
Bit map, primary	Series of 64 bits used identifying the presence (denoted by 1) or absence (denoted by 0) of data elements 1 through 64		b 8	see 5.2	n/a
Bit map, secondary	Series of 64 bits used to identify the presence (denoted by 1) or absence (denoted by 0) of data elements 65 through 128		b 8	see 5.2	1
Card acceptor address	Address of the card acceptor allocated by the acquirer, sufficient to allow the cardholder to recognize the location		ans 23	see 6.2.5	43-1

Name	Description	Format	Representation	Notes	Bit
Card acceptor city	City of the card acceptor allocated by the acquirer, sufficient to allow the Cardholder to recognize the location from his statement		ans 13	see 6.2.5	43-2
Card acceptor country code	Country code of the card acceptor as known to the cardholder	US = United States	a 2	see ISO 3166 and 6.2.5	43-4
Card acceptor identification code	Code identifying the card acceptor that defines the point of the transaction in both local and interchange environments		an 15	Shall remain unchanged for all messages in a transaction, see C.6.5	42
Card acceptor name/location	Name and location of the card acceptor as known to the cardholder		ans 40	see 6.2.5	43
Card acceptor state	State of the card acceptor allocated by the acquirer, sufficient to allow the cardholder to recognize the location		a 2	see U.S. Post Office Publication 65 and 6.2.5	43-3
Card acceptor terminal identification	Unique value identifying an individual POS terminal at the card acceptor location		ans 8	Shall remain unchanged for all messages in a transaction, see C.6.5	41

Name	Description	Format	Representation	Notes	Bit
Card reading method used at POS	Two digits to indicate the method by which the Primary Account Number (PAN) was entered into the system		n 2	see Annex B.8.1	22-1
Cardholder verification method used at POS	One digit to indicate method used to verify the cardholder		n 1	see Annex B.8.2	22-2
Card sequence number	Number distinguishing between separate cards with the same Primary account number		n 3		23
Currency code, amounts additional	Code identifying the currency of the amount contained in Amounts, additional data elements		n 3	see ISO 4217 and 6.2.4	54-4
Currency code, transaction	Local currency of the acquirer or source location of the transaction.	840 = US	n 3	see ISO 4217	49
Date and time, transmission	Date and time the message entered into the data interchange system, expressed in Greenwich Mean Time (GMT)	MMDD hh mmss	n 10		7
Date, capture	Month and day the transaction was processed by the acquirer	MMDD	n 4		17
Date, expiration	Year and month after which the card expires	YYMM	n 4		14

Name	Description	Format	Representation	Notes	Bit
Date, local transaction	Local month and day on which the transaction takes place at the card acceptor location	MMDD	n 4	Shall remain unchanged for all messages in a transaction, see C.6.5	13
Date, settlement	Month and day funds transferred between the acquirer and the EBT card issuer processor	MMDD	n 4		15
FNS number	Number assigned by the USDA FNS to identify an authorized retail location for the Food Stamp program		n 7	see 6.2.2, B.12	111-EB
Forwarding institution identification code	Identity of the institution forwarding the request or advice message in an interchange system	LLVAR	n..11		33
Info, text	Additional information for Administrative and Network management transactions	LLLVAR	ans...255		124
Institution/merchant name	Name of the financial institution that owns the ATM or name of the merchant where the POS terminal is located	LLVAR	ans...25		48

Name	Description	Format	Representation	Notes	Bit
Issuer trace data	Tracing information from the EBT card issuer processor	LLVAR	ans...100		126
Integrated circuit card (ICC) system related data	Data related to integrated circuit card systems. The structure of this data element is described in ISO 9992.	LLLVAR	b...255	see ISO 9992	55
Key management data	Verification between an acquirer and EBT card issuer processor; contains data related to key management		an 8	see ISO 11568, ISO 13492 and ANS X9.24	96
Merchant type	Classification of the merchant's type of business product or service		n 4	see ANS X9.10 for values	18
Message authentication code (MAC)	Used to validate the source and the text of the message between the sender and receiver		b 4	see ISO 16609, ISO 9797	64 or 128
Message type	Identification of the version, message class, message function and transaction originator		n 4	see 5.1.1 and Annex B.6	n/a
National point of service condition code	Series of codes intended to identify terminal class, presentation data, security condition and terminal type		an 10	see Annex B.6	58

Name	Description	Format	Representation	Notes	Bit
National point of service county code	County where the point of service device is physically located		n 3	see 6.2.10, INCITS 31	59-2
National point of service country code	Country where the point of service device is physically located	840 = US	n 3	see ISO 3166 and 6.2.10	59-4
National point of service geographic data	Series of codes to identify the state, county, postal service code and country code where the point of service device is physically located	LLLVAR	an...17	see 6.2.10	59
National point of service postal service code	Code allocated by postal authority of the card acceptor to identify its physical location		an 5 or 9	see U.S. Postal Service publication 65 and 6.2.10	59-3
National point of service presentation data	Code indicating whether the card and cardholder are present, device card capture capability and presentment type		an 4	see Annex B.6.2	58-2
National point of service security condition	Code indicating the security issue, if any, at the point of service		an 1	see Annex B.6.3	58-3
National point of service state code	State code allocated by the acquirer to indicate the physical location of the point of service device		n 2	see U.S. Postal Service publication 65 and 6.2.10	59-1

Name	Description	Format	Representation	Notes	Bit
National point of service terminal class	Code indicating the location of the point of sale terminal		an 3	see Annex B.6.1	58-1
National point of service terminal type	Code indicating the type of point of sale terminal at the point of sale		an 2	see Annex B.6.4	58-4
Network management information code	Code indicating the network status		n 3	see Annex B.7	70
Original data elements	Data elements contained in the original message	LLVAR	n...42	see 6.2.7	90
Original message type	Message type identifier of the original transaction		n 4	see 6.2.7	90-1
Original system trace audit number	System trace audit number of the original transaction		n 6	see 6.2.7	90-2
Original transmission date and time	Date and time of transmission of the original transaction		n 10	see 6.2.7	90-3
Original acquiring institution identification code	Acquiring institution identification code of the original Financial Presentment transaction		n...11	see 6.2.7	90-4
Original forwarding institution identification code	Forwarding institution identification code of the original Financial Presentment transaction		n...11	see 6.2.7	90-5

Name	Description	Format	Representation	Notes	Bit
PIN data	Number assigned to a cardholder intended to uniquely identify that cardholder at the point of service		b 8	see ANS X9.8	52
Point of service entry mode	Two digits to indicate the method by which the PAN was entered into the system and one digit to indicate cardholder verification method		n 3	see Annex B.8	22
Point of service PIN capture code	Code indicating the technique and/or maximum number of PIN characters accepted by the POS device used to construct the PIN data		n 2	see Annex B.9	26
PAN	Series of digits used to identify a customer account or relationship	LLVAR	n...19		2
Processing code	Series of digits used to describe the effect of a transaction on the customer account and identify the accounts affected		n 6	see Annex B.10; 6.2.6	3
Replacement amounts	New actual amount data elements, necessary to perform a partial reversal. The original transaction amount for a store and forward partial approval		an 42	see 6.2.1, 6.2.8 and C.5	95

Name	Description	Format	Representation	Notes	Bit
Response code	Code that defines the disposition of a message		an 2	see Annex B.11	39
Retrieval reference number	Document reference supplied by the system retaining the original source document used in locating the document or a copy of the document		an 12	shall remain unchanged for all messages in a transaction, see C.6.5	37
Retrieval reference number Julian date	4-position Julian date of the document reference number assigned to this transaction as contained in an optional format of retrieval reference number	YDDD	an 4	shall remain unchanged for all messages in a transaction, see C.6.5	37-1
Retrieval reference number terminal sequence number	8-position sequential number assigned to the transaction by the terminal at the POS as contained in an optional format of retrieval reference number		n 8	shall remain unchanged for all messages in a transaction, see C.6.5	37-2
Security related control information	Identification of security management information used in the current transaction or specification of security management information to be used in future transactions		b...48	see ISO 13492 and ANS X9.24	53

Name	Description	Format	Representation	Notes	Bit
Special program	Indicator of a service or program offered by a state or WIC authority (e.g., the California Healthy Food Purchase program)		n 7	contains FNS number associated with the program, see 6.2.2, B.12	111-SP
System trace audit number	Number assigned by the message initiator to uniquely identify a transaction		n 6	shall remain unchanged for all messages in a transaction, see C.6.5	11
Time code	Type of time interval in effect for a Pre-authorization transaction; contained in Authorization life cycle		n 1	see Annex B.4	57-1
Time interval	Value of the time interval in effect for a Pre-authorization; contained in authorization life cycle		n 2	see Annex B.4	57-2
Time, local transaction	Local time at which the transaction took place at the point of the card acceptor location	hhmmss	n 6		12
Track 2 data	Information encoded on track 2 of the magnetic stripe as defined in ISO 7813, excluding beginning and end sentinels and LRC characters as defined therein	LLVAR	z...37	see ISO 7813 and 6.2.9	35

Name	Description	Format	Representation	Notes	Bit
Transaction type	First two positions of the processing code that indicate the kind of transaction performed		n 2	see B.10.1	3-1
Voucher number	Preprinted number from a paper sales draft		an 15	see 6.2.2, B.12	111-VN

Source : American National Standards for Financial Services ANS X9.58-2007. Financial Transaction Messages – EBT – Food Stamps

Exhibit B.2. Key to Abbreviations Exhibit B.1

Abbreviation	Definition
17	Variable length up to maximum 17 characters. All variable length fields shall in addition contain two, three, or four positions at the beginning of the data element to identify the number of positions following to the end of that data element.
3	Fixed length of three characters
a	Alphabetical characters A through Z and a through z
an	Alphabetic and numeric characters
anp	Alphabetic, numeric, and space (pad) characters
ans	Alphabetic, numeric, and special characters
ansb	Alphabetic, numeric, special characters and binary representation of data
as	Alphabetic and special characters
b	Binary representation of data
CCYY	Century and year, 0000 through 9999 (see ISO 8601)
DD	Day, 01 through 31
hh	Hour, 00 through 23
LL	Length of variable data element that follows, 01 through 99
LLL	Length of variable data element that follows, 001 through 999
LLLL	Length of variable data element that follows, 0001 through 9999
MM	Month, 01 through 12
mm	Minute, 00 through 59
n	Numeric digits, 0 through 9
ns	Numeric and special characters
p	Pad character (e.g., space)
s	Special characters
ss	Second, 00 through 59
VAR	Variable length data element
xn	Numeric data with a preceding sign of "c" for credit, "d" for debit (e.g. x + n 8 in amount), transaction fee means prefix "c" or "d" and 7 digits of amount
YY	Year, 00 through 99
z	Tracks 2 and 3 code set as defined in ISO 4909, ISO 7811-2 and ISO 7813

Source : American National Standards for Financial Services ANS X9.58-2007. Financial Transaction Messages – Electronic Benefits Transfer (EBT) – Food Stamps

Appendix C. Integrated Electronic Cash Register Hardware and Software Technology

As can be seen in Exhibit C.1, IECRs connect the ECR, a local hard drive storing a dynamic version of the retailer inventory database and transaction records, a scanning device, a printer, and the integrated POS terminal. The IECR also connects either directly or through the POS terminal to a central processor or mainframe that transmits the transactions to a TPP. These hardware components, available from many different vendors, communicate with one another within an *operating system.*

While not all IECRs function the same way, cashiers typically use cash register hardware that runs a *POS software application,* which facilitates tasks such as:

- Scanning purchased items
- Looking up items in the retailer's inventory system
- Applying discounts and coupons
- Sending transaction totals to the POS terminal
- Communicating directly or indirectly with TPP or EBT processors to validate an EBT card or check availability of funds[30]
- Preparing receipts
- Reconciling the retailer's inventory database at both the ECR and the local server

The most critical hardware components that could affect potential solutions for collecting POS item-level data are the ECR, the POS terminal, and external servers. Given the variety of ECR hardware and software applications in the retail environment, it is important to understand the market penetration of each hardware and software technology, because market penetration informs the proportion of retailers affected by potential changes based on the technology they

[30] Indirect communication, used by large retail chains, involves connection to a centralized transaction processing system.

use. Because most ECR hardware will run most POS software applications, the primary consideration is the POS software application.

Exhibit C.1. Hardware Configuration in an IECR System

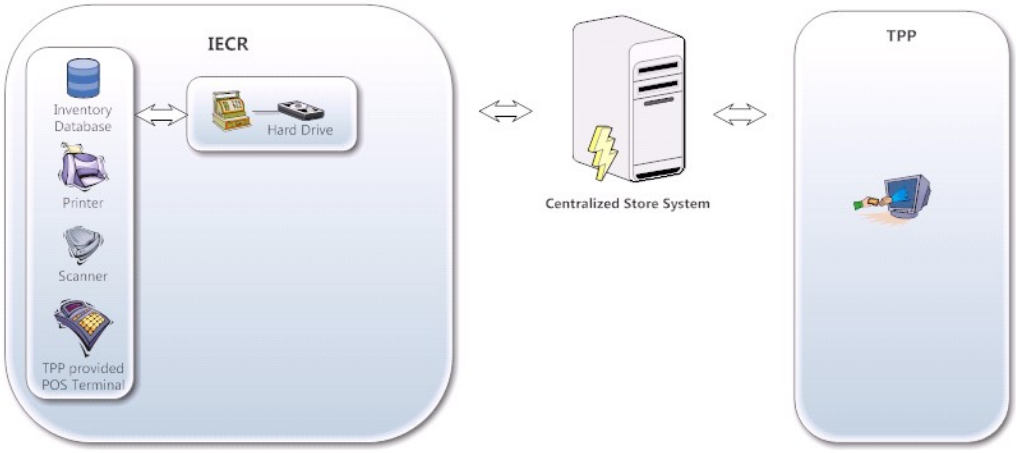


Exhibit C.2 outlines the major hardware and software technologies in the IECR environment. It provides an overview of the hardware, operating system, and software options available to retailers. Although there are a number of POS hardware manufacturers, most POS hardware can run Windows-based operating systems, which, in turn, can support a variety of POS software applications. Because most hardware runs most POS software applications, they are the major element to consider in modifying POS systems.

Exhibit C.2. Point-of-Sale Hardware Manufacturers, Operating Systems, and POS Software Applications

Primary Manufacturer	Operating Systems	Supported POS Software Applications
IBM/Toshiba	Proprietary OS ▪ Toshiba 4690 OS, Version 6.3	SurePOS ACE for 4690 OS
	Windows-based OS ▪ Windows Embedded POSReady ▪ Other	▪ Microsoft Dynamics RMS (Windows-based) ▪ Microsoft Dynamics POS 2009 ▪ StoreNext (Retalix) ▪ Retalix10 ▪ StorePoint (Retalix) ▪ StoreLine (Retalix) ▪ JPMA POS Designer ▪ Aurora ▪ RORCvIPOS V6 ▪ Retail Professional

Primary Manufacturer	Operating Systems	Supported POS Software Applications
NCR	Windows-based OS ▪ 82XRT – Windows 7 Professional ▪ 70XRT – Windows XP Professional ▪ Windows Embedded POSReady	▪ NCR ACS ▪ Microsoft Dynamics RMS (Windows-based) ▪ Microsoft Dynamics POS 2009 ▪ StoreNext (Retalix) ▪ Retalix10 ▪ StorePoint (Retalix) ▪ StoreLine (Retalix) ▪ JPMA POS Designer ▪ Aurora ▪ RORCvIPOS V6 ▪ Retail Professional
	Linux-based OS ▪ SUSE Linux Enterprise for Point-of-Service (SLEPOS) 25, 40, 50, 60	▪ NCR ACS ▪ Aurora
Casio	Window-based OS ▪ Windows Embedded POSReady ▪ Other	▪ Microsoft Dynamics RMS (Windows-based) ▪ Microsoft Dynamics POS 2009 ▪ StoreNext (Retalix) ▪ Retalix10 ▪ StorePoint (Retalix) ▪ StoreLine (Retalix) ▪ JPMA POS Designer ▪ Aurora ▪ RORCvIPOS V6 ▪ Retail Professional
HP	Windows-based OS ▪ Windows Embedded POSReady ▪ FreeDOS	▪ Microsoft Dynamics RMS (Windows-based) ▪ Microsoft Dynamics POS 2009 ▪ StoreNext (Retalix) ▪ Retalix10 ▪ StorePoint (Retalix) ▪ StoreLine (Retalix) ▪ JPMA POS Designer ▪ Aurora ▪ RORCvIPOS V6 ▪ Retail Professional
Dell	Windows-based OS ▪ Windows Embedded POSReady ▪ FreeDOS	▪ *QuickBooks* Point of Sale ▪ Microsoft Dynamics RMS (Windows-based) ▪ Microsoft Dynamics POS 2009 ▪ StoreNext (Retalix) ▪ Retalix10 ▪ StorePoint (Retalix) ▪ StoreLine (Retalix) ▪ JPMA POS Designer ▪ Aurora ▪ RORCvIPOS V6 ▪ Retail Professional

Appendix D. Extract of Transaction Basket Barcodes Generated

Exhibit D.1 presents a sample transaction basket generated for the proof of concept. Each row represents an EAN-13 barcode, as well as the associated UPC, UPC description, item price, and quantity to be purchased. For the POC, the register lane attendant used the IECR scanner and scanned each of these barcodes to simulate an actual transaction. Several items constituted the transaction basket (this example it is basket #251). A receipt and TWILD receipt were then generated for this basket.

Exhibit D.1. Basket Barcodes and Data Elements

Basket # 251

Barcode	UPC	UPC Description	Item Price	Quantity
	070043015866	OM LCHNB SMT	$ 3.59	1
	027815300538	EGGS LARGE	$ 2.69	1
	048564067008	OM FUNPK PEP	$ 5.00	2
	043000953686	HUNT GEL ST	$ 5.00	5
	015400837491	CHICKEN THIG	$ 6.96	1
	071899211693	SF SPAN OLIV	$ 1.59	1
	015900000975	FAMILY STEAK	$ 3.17	1

TRUNO

retail technology solutions

Truno
13912 FM 1730
Lubbock, TEXAS
(800) 555-1212

Your Checker today is Johnny Appleseed

1 @ 2/6.00	
OM LCHBL W SMTHIE	3.00 F
SF 18PK EGGS	3.09 F
1 @ 2/5.00	
OM FNPK LCHBL PEPP	2.50 F
1 @ 2/5.00	
OM FNPK LCHBL PEPP	2.50 F
1 @ 4/5.00	
HUNTS SNCK GEL STR	1.25 F
1 @ 4/5.00	
HUNTS SNCK GEL STR	1.25 F
1 @ 4/5.00	
HUNTS SNCK GEL STR	1.25 F
1 @ 4/5.00	
HUNTS SNCK GEL STR	1.25 F
1 @ 4/5.00	
HUNTS SNCK GEL STR	1.25 F
CHICKEN THIG	6.96 F
SF ST SPANIS	1.59 F
FAMILY SYEAK	3.17 F
CAYENNE PEPP	1.39 F
YOPLAIT STRAWBERY	0.48 F
YOPLAIT STRAWBERY	0.48 F
TAX	0.00
**** BALANCE	31.41
FS BAL DUE	31.41

LOWES'S BIG 8 #1
1 1 28

Date: 04/15/15 Time: 09:17am
Amount $20.00
FS End Bal $
Cash End Bal $
**********4205
1458 00 APPROVED
EBT Food Stamps Purchase

EBT Food Stamps	20.00
Cash	11.41
FS CHG	0.00
CHANGE	0.00

TOTAL NUMBER OF ITEMS SOLD = 15
04/15/15 09:17am 1 1 28 1
****
====Thank you for shopping with us=
****

LOWES'S BIG 8 #1
1 1 28

Date: 04/15/15 Time: 09:17am
Amount $20.00
FS End Bal $
Cash End Bal $
**********4205
1458 00 APPROVED
EBT Food Stamps Purchase

Appendix E. TWILD Extraction Software Specifications

Unified TWILD extraction software, which would bring both efficiencies and standardization, does not currently exist. IMPAQ proposes that FNS develop a template for this software for retailers or third-party vendors to use and customize their specific IECR and POS software applications. The software will have to be designed and developed to extract the TWILD from the POS native file format. To meet the needs of retailers identified in this scenario, we have identified the following initial TWILD extraction software specifications:

1. Extract TWILD from the native register log files and inventory data.
2. Extract only transactions with EBT payments.
3. Extract mixed tender transactions with EBT payments.
4. Enable the TWILD export location to be specified.
5. Enable TWILD file name format to be specified.
6. The software would not need to manage data transmission.