



OFFICE OF
**INSPECTOR
GENERAL**
UNITED STATES POSTAL SERVICE

**Enterprise Data Warehouse
Cost of Use**

Audit Report

August 5, 2013

Report Number DP-AR-13-009



HIGHLIGHTS

August 5, 2013

Enterprise Data Warehouse Cost of Use

Report Number DP-AR-13-009

BACKGROUND:

The U.S. Postal Service created the Enterprise Data Warehouse (EDW) as the main source for storing data across functional areas throughout the organization. The EDW stores data from about 100 systems or applications. The process for loading data from these applications into the EDW is referred to as an extract, transform, and load process. There are about 14,000 individual programs needed to perform this process.

The extract, transform, and load process requires a series of tasks that are accomplished with contractor support. Additional contractor support has been required because of the growth in the number and complexity of manually developed extract, transform, and load scripts. Replacing the current method of coding and maintaining the extract, transform, and load scripts with a tool that automates the initial build of the scripts would reduce both current and future maintenance and development costs.

Our objective was to evaluate whether the Postal Service could reduce EDW costs by implementing technological advances available in the marketplace

WHAT THE OIG FOUND:

Opportunities exist for the Postal Service to reduce EDW costs by implementing technological advances

available in the marketplace.

Commercial software is available to automate the interface process and replace the current manual process. As a result, the Postal Service spent \$500,000 in labor costs that it could have avoided in fiscal years 2012 and 2013. In addition, the Postal Service could save about \$1.9 million annually in future years. Further, standardizing the process would eliminate the potential for manual programming errors and allow administrators easier access to data from the various functional areas.

WHAT THE OIG RECOMMENDED:

We recommended the Postal Service acquire and implement a tool that automates the scripts used to perform the extract, transform, and load process.

[*Link to review the entire report*](#)



August 5, 2013

MEMORANDUM FOR: JOHN T. EDGAR
VICE PRESIDENT, INFORMATION TECHNOLOGY

E-Signed by Inspector General
VERIFY authenticity with eSign Desktop

FROM: Darrell E. Benjamin, Jr.
Deputy Assistant Inspector General
for Revenue and Performance

SUBJECT: Audit Report – Enterprise Data Warehouse Cost of Use
(Report Number DP-AR-13-009)

This report presents the results of our audit of Enterprise Data Warehouse Cost of Use (Project Number 13BG001FF000).

We appreciate the cooperation and courtesies provided by your staff. If you have any questions or need additional information, please contact Kevin H. Ellenberger, director, Data Analysis and Performance, or me at 703-248-2100.

Attachment

cc: Ellis A. Burgoyne
Corporate Audit and Response Management

TABLE OF CONTENTS

Introduction 1

Conclusion 1

Current Interface Process 2

 Available Alternatives to the Current Process 2

 Technological Advances 2

 Cost Savings..... 3

 Operational Benefits..... 3

Recommendation 3

Management’s Comments 3

Evaluation of Management’s Comments 4

Appendix A: Additional Information 5

 Background..... 5

 Objective, Scope, and Methodology..... 5

 Prior Audit Coverage..... 6

Appendix B: Monetary Impacts 7

Appendix C: Management's Comments 9

Introduction

This report presents the results of our audit of Enterprise Data Warehouse (EDW) Cost of Use (Project Number 13BG001FF000). Our objective was to evaluate whether the U.S. Postal Service could reduce EDW costs by implementing technological advances available in the marketplace. This self-initiated audit addresses the strategic and operational risks associated with costs of storing and reporting data. See [Appendix A](#) for additional information about this audit.

The Postal Service created the EDW as the main source for storing data across functional areas throughout the organization. The EDW now stores data from about 100 systems or applications. The process for integrating data from these applications into EDW is referred to as an extract, transform, and load (ETL) process and there are now about 14,000 individual programs needed to integrate data into the EDW. Further, the number of programs continues to increase with each request to add new data to the EDW or make changes to programs or applications already storing data in the EDW. Replacing the current manual method of performing these procedures with software to automate the process would reduce both current and future maintenance and development costs for the EDW.

Information technology (IT) personnel use a combination of Postal Service employees and contractor personnel to perform ETL processes. The current method of ETL processing requires significant manual labor. For example, data coming in from an application that is not in an acceptable format for EDW causes the program to reject the entire file. This prevents loading data into the EDW until it is manually examined and corrected to meet the standards for EDW. Additional contractor support has been required because of the growth in the number and complexity of manually developed ETL scripts. Replacing the current method of coding and maintaining the ETL scripts¹ with a tool that automates the initial build of the scripts would reduce both current and future maintenance and development costs.

Conclusion

Opportunities exist for the Postal Service to reduce EDW costs by implementing technological advances available in the marketplace. Specifically, software is available to automate the manual process used to integrate data from source systems into the EDW. As a result, the Postal Service could save about \$1.9 million annually² and obtain additional operational benefits associated with automating the EDW interface.

We consider these savings as funds put to better use. See [Appendix B](#) for the calculation of monetary impact.

¹ Scripts are programs written for a special [environment](#) that can interpret and [automate](#) the [execution](#) of tasks which could alternatively be executed one-by-one by a human operator.

² The projected annual savings of \$1.9 million will occur after the first full year of implementation.

Current Interface Process

The current method of moving data from source systems to the EDW is accomplished by 13 full-time contractors and 12 full-time Postal Service employees. The ETL processes also require specialized IT skills often performed by contractors. As the volume of data in the EDW increases, more resources are needed to perform these processes to validate, test, and correct data before it is ready for input into the EDW. The Postal Service has required additional contractor support because of the growth in the number and complexity of manually developed ETL scripts.

Available Alternatives to the Current Process

Since 2009, industry experts in data warehousing technology have recommended that the Postal Service acquire and implement a tool to replace the manual programming process.³ Subsequently, managers from the EDW Solutions team began conducting research in May 2011 to identify products and vendors to automate the ETL process; however, management did not have a specific time line in place to purchase and implement software to automate the process. As such, headquarters officials have not replaced the manual ETL process for two reasons: the Postal Service's economic conditions caused concern that funding for this investment was not available; and they want more time to study, consider, and identify the most economically feasible product. IT personnel are still using manual procedures to develop and maintain internal ETL programs.

Technological Advances

Our research indicates that there are at least five vendors who offer commercial off-the-shelf software that may be capable of performing the ETL processes to integrate data from various source systems into the EDW. The software investment would replace the skilled labor requirements currently obtained from contractors and reduce the number of hours required to perform these processes.

We estimate the Postal Service could realize savings of about \$1.9 million annually during the first full year after the software is implemented and operational. Our research also indicates the Postal Service will realize additional operational benefits by automating this process.

³ Based on findings from an EDW assessment conducted from 2009 through 2010 by Clarview, a division of Teradata Corporation.

Cost Savings

The Postal Service could realize savings from reducing contractor workhours by investing in ETL software. We estimate an annual savings of about \$1.9 million after the software is implemented and fully operational.

To conduct our analysis, we obtained the number of contractor workhours used to support the ETL processes from October 2011 through September 2012. Contractor costs for [REDACTED] workhours to program the interface processes were about \$ [REDACTED]

Using the workhour data from 2012, our estimates show a monetary savings of about \$500,000 after deducting the initial investment plus other fees (about \$3 million) from the workhour savings (about \$3.5 million) over the first 2 years. After the Postal Service recoups the initial investment costs, the savings in years 3 and 4 increase to about \$1.9 million annually. See [Appendix B](#) for our calculation of monetary impact.

Operational Benefits

In addition to cost savings, there are additional benefits or improvements the Postal Service would realize by automating the ETL process. For example, because data from multiple sources are stored in a single warehouse, automation will make integrating data from various systems into the EDW much faster. Also, the software's graphical interface will provide database administrators with better access to data when consolidating different types of data from various functional units within the Postal Service. Software would also standardize the process and eliminate the potential for manual programming errors. Increasing the volume of data requires the ability to upgrade data integration to real-time processing, which is an option with the interface tool. Each improvement to the warehousing process will ultimately provide users with better access for analyzing and reporting all the data available, no matter the source or application.

Recommendation

We recommend the vice president, Information Technology:

1. Acquire and implement a tool that automates the scripts used to perform the extract, transform, and load process.

Management's Comments

Management agreed with the finding, recommendation, and monetary impact. Management stated the acquisition and implementation of a tool for managing the code in the ETL process would recognize the financial benefits stated in the audit report given

that the tool can provide a rapid conversion of the current ETL code. Management has identified the optimal ETL tool and is in the process of acquiring this tool by September 30, 2013. Targeted implementation date is May 31, 2014. See [Appendix C](#) for management's comments in their entirety.

[Evaluation of Management's Comments](#)

The U.S. Postal Service Office of Inspector General (OIG) considers management's comments responsive to the recommendation and corrective actions should resolve the issues identified in the report.

Appendix A: Additional Information

Background

The EDW, which became available for use in 2001, serves as the primary reporting system for the Postal Service. The EDW provides a single source of data to a variety of users across functional areas. About 100 applications or systems belonging to functional areas such as Retail, Network Operations, and Finance feed data into the EDW. The EDW has more than 40,000 registered users and provides business insight to more than 6,000 Postal Service employees per week. The EDW is the source for more than one million reports per month and is also the primary financial reporting system for the Postal Service.

The Postal Service uses a combination of Postal Service employees and contractor personnel to develop programs that transfer data from the applications or feeder systems into the EDW. This is referred to as an ETL process and it is, in essence, an interface between source systems and the EDW. There is an EDW Solutions team in Raleigh, NC, that has primary oversight and management of the ETL processing teams. Since the EDW's origin — and as its scope has increased — interface processing has grown to about 14,000 individual programs and continues to grow with every new release and source system added.

Objective, Scope, and Methodology

Our objective was to evaluate whether the Postal Service could reduce EDW costs by implementing technological advances available in the marketplace. We conducted this audit from October 2012 through August 2013, in accordance with generally accepted government auditing standards and included such tests of internal controls as we considered necessary under the circumstances. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objective. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objective. We discussed our observations and conclusions with management on June 6, 2013, and included their comments where appropriate.

We did not test the validity of controls over the Postal Service systems. We verified the accuracy of the data by interviewing agency officials and tracing data to source documents, such as time cards. We determined the data were sufficiently reliable for the purposes of this report.

Prior Audit Coverage

The OIG did not identify any prior audits or reviews related to the objective of this audit conducted during the past 3 years.

Appendix B: Monetary Impacts

Recommendation	Impact Category	Amount
1	Questioned Costs ⁴	\$483,808
1	Funds Put to Better Use ⁵	3,800,989
Total		\$4,284,797

The summary below shows how we computed the monetary benefits associated with converting to an automated interface process.

Category of Investment/Saving	Amount
Cost of software and licenses, software maintenance, vendor implementation support, and Postal Service implementation ⁶	\$3,030,420
Total Investment	\$3,030,420
Savings in years 1 and 2 from reduction in labor costs	\$3,514,228
Savings after initial investment (questioned costs)	483,808
Projected savings in years 3 and 4 from reduction in labor costs (funds put to better use)	3,800,989
Total Savings	\$4,284,797
Monetary Benefit to Postal Service	\$4,284,797

Source: Technology Investment Program Approval and Workhours from Product Cost Tracking System.

The initial cost of the investment based on FY 2012 costs was \$3,030,420, which included costs for vendor migration support, software licenses, software maintenance, and Postal Service implementation support. The software licensing and maintenance support costs for the following years were offset by a reduction in software licensing fees that would be replaced with the investment. Therefore, after deducting investment costs from the first 2 years of savings from reduced labor costs, we calculated about \$500,000 in questioned costs that could have been avoided if management had pursued the initiative.

Savings from reduced labor workhours over the next 2 years amounted to \$3,800,989. We consider these funds that could have been put to better use.

⁴ Questioned Costs - unnecessary, unreasonable, unsupported, or an alleged violation of law, regulation, contract, etc. May be recoverable or unrecoverable. Usually a result of historical events.

⁵ Funds Put to Better Use - funds that could be used more efficiently by implementing recommended actions.

⁶ Investment costs based on fiscal year (FY) 2012 Postal Service support implementation costs for a 6-month period.

Overall, the Postal Service could have saved about \$4.3 million if management had purchased the software needed to replace the programs and the manual processes used to ETL data from source systems into the EDW.

Appendix C: Management's Comments

JOHN T. EDGAR
VICE PRESIDENT
INFORMATION TECHNOLOGY



July 26, 2013

JUDITH LEONHARDT
DIRECTOR, AUDIT OPERATIONS

SUBJECT: Response to Draft Audit Report – Enterprise Data Warehouse Cost of Use (FF-AR-13-DRAFT)

Management Response: Management agrees that the acquisition and implementation of a tool for managing the code in the Extract, Transform, and Load (ETL) process would recognize the financial benefits given that the tool selected can provide a rapid conversion of the current ETL code into the new tool. Prior to this audit, the Postal Service had been actively pursuing the identification and verification of an ETL tool best suited for the Postal Service.

Based on market research, we have identified one (1) ETL tool that can automatically and accurately convert the existing code base. It is important to note that the potential savings of \$1.9 million reported by the OIG would only be achievable if the tool acquired rapidly converts the current ETL code base to the new tool's format.

Recommendation 1: Acquire and implement a tool that automates the scripts used to perform the extract, transform, and load process.

Management Response/Action Plan: Management agrees that the acquisition and implementation of a tool for managing the code in our ETL process would recognize the financial benefits stated in the report, given that the tool selected can provide a rapid conversion of the current ETL code into the new tool. Management has identified the optimal ETL tool which will replace the current ETL tool and as of today, the Postal Service is in the process of acquiring this tool. Management to provide a project plan of the new ETL implementation by September 30, 2013.

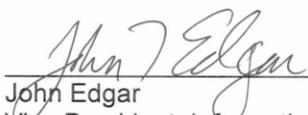
Target Implementation Date: May 31, 2014

475 L'ENFANT PLAZA SW
WASHINGTON DC 20260-2100
202-268-3977
FAX: 202-268-4492
JOHN.T.EDGAR@USPS.GOV
WWW.USPS.COM

- 2 -

Responsible Official: Mark A. Mittelman, Manager, Solutions Development & Support

This report and management's response contain information which management believes may contain proprietary or other business information that may be exempt from disclosure under the Freedom of Information Act (FOIA). The VP Information Technology, requests that sections Highlights, Technological Advances, Cost Savings, and Appendix B: Monetary Impact of the report should be redacted.



John Edgar
Vice President, Information Technology

Cc:
Mark A. Mittelman
Ellis Burgoyne
Corporate Audit and Response Management