ABSTRACT

MATHEWS, REENA. Simple Strategies to Improve Data Warehouse Performance. (Under the direction of Professor. Robert B. Handfield).

Data warehouse management is fast becoming one of the most popular and important topics in industries today. For business executives, it promises significant competitive advantage for their companies, while presenting the information system managers a way to overcome the obstructions in providing business information to managers and other users. Here the company is going through the problem of inefficient performance of its data warehouse. To find an appropriate solution to this problem we first try to understand the data warehouse concept and its basic architecture, followed by an in depth study of the company data warehouse and the various issues affecting it. We propose and evaluate a set of solutions including classification of suppliers, implementing corporate commodity classification and coding system, obtaining level three spend details for PCard purchases, etc. The experimental results show considerable improvement in the data quality and the data warehouse performance. We further support these recommendations by evaluating the return on investment for improved quality data. Lastly, we discuss the future scope and other possible improvement techniques for obtaining better results.

SIMPLE STRATEGIES TO IMPROVE DATA WAREHOUSE PERFORMANCE

By

REENA MATHEWS

A Thesis submitted to the graduate faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

OPERATIONS RESEARCH

Raleigh

2004

APPROVED BY

_____	_____
Dr. Yahya Fathi	Dr. Negash Medhin

_____

Dr. Robert B. Handfield

(Chair of Advisory Committee)

DEDICATION

To my best friend Amitabh Shukla.

To my loving parents.

BIOGRAPHY

Reena Mathews was born in the marble city of Jabalpur, situated at the heart of India. She received her Bachelor of Technology degree in Production Engineering from Calicut University. She started her career as a software engineer at Infosys Technologies Ltd, Bangalore, India before joining North Carolina State University to pursue an MS degree in Operations Research. During her graduate studies she undertook course work in Operations Research and Supply Chain Management. As part of her MS co-operative program, she worked as a 'Supply Chain Analyst' at Sonoco Products Company, Hartsville, South Carolina. Her research interests and experience include supply chain management, logistics, discrete optimization and software programming.

ACKNOWLEDGEMENT

First and foremost, I am grateful to my Lord, Jesus Christ, for being there for me at every step of my life. I am also thankful to my parents and my sister for their love and affection and for making my higher education dreams come true.

I would especially like to thank my advisor Dr. Robert Handfield for his suggestions, insights and constant support during this research. I am also thankful to my committee members Dr. Yahya Fathi and Dr. Negash Medhin for their support and guidance.

I also wish to thank Mr. Marc Ensign, Mr. John Boyette, Mr. Brent Snyder, and Mr. Dan Theismann at Sonoco for sharing their knowledge and guiding this study.

Finally, my special thanks to the following: Seem (for her friendship); Louie (for believing in me); Vel (for not letting me give up) and all my other friends for their support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Information is pivotal in the business environment. A company's success depends on its timely and decisive use. Today the major obstacle for industries is to obtain high quality data in the required format. This is fundamental to the development of the field of operations research as the accuracy of the data dictates various sensitivity analyses essential for making sound decisions. Businesses also require the ability to access and combine data from both internal and external sources, perform complex data analyses on operational and historical data, be able to drill down, roll up, slice and dice data to obtain the required information to help them stay ahead of their competitors. Data warehouse provides the means to do this.

1.1 DATA WAREHOUSE

A data warehouse can be defined as a single, complete, and consistent store of data obtained from a variety of sources and made available to end users in a way they can understand and use in a business context. A data warehouse is a copy of transaction data specifically structured for querying and reporting. It is a huge (sometimes terabytes of disk storage) database, which stores volumes of historical data for the company. The database can be of any form. It can be a relational database, multi dimensional database, flat file, hierarchical database, etc. A data warehouse organized for a single department or for a specific group of people is termed a data mart.

The concept of a data warehouse came into existence as a result of two different sets of requirements. First, the end users need to view and understand the company wide view of information and second, the information system (IS) department's need to manage the data for technological and economic reasons. Although, the two requirements may seem completely different from each other, the data warehouse gurus confirm that addressing any one of these requirements will make it easier to meet the other requirement and vice versa.

1.2 DATA WAREHOUSING

Data warehousing is a process that involves the physical separation of an organization's transaction operations from its decision support operations. It includes a set of methods, techniques and tools that are used to deliver data to the end users. Data warehousing involves modeling data for the warehouse based on the company's information needs, extracting data from the operational databases, cleansing data like eliminating negative inventory quantities or standardizing fields or attribute titles, transforming data into the warehouse model and loading data into the data warehouse database.

1.3 GOALS OF A DATA WAREHOUSE

1 Data warehouse helps in making an organizations' information accessible**.** This data is understandable, navigable and can be reached in zero wait time.

2 Data warehouse keeps the information well defined, consistent and non-volatile. This information is high quality information, which is accounted for and complete.

3 Data warehouse contributes to the decision-making process by providing the right data in large quantity to support data analysis, querying and reporting. It serves as an integral part of any company's decision support system.

4 Data warehouse aids in improving the efficiency of transaction processing systems by serving as a storage area for older data.

5 Data warehouse also serves as a repository for a variety of information, thereby fulfilling the different information needs of different users.

These are the fundamental goals of a data warehouse. Any system, which does not work towards these goals, is a compromise and needs to be improved.

1.4 COMPANY BACKGROUND

Sonoco Products Company (Sonoco) is one of the world's largest manufacturer of industrial and consumer packaging products and provider of packaging services. The company's industrial packaging segment serves the textile, paper and film industries and makes cores, cones, tubes, plastic products, wood and metal reels. Its consumer packaging segment makes composite cans, flexible packaging, can ends for plastic and metal containers, plastic grocery sacks and offers artwork and brand management

services through its UK-based subsidiary Sonoco Trident. Sonoco has approximately 300 locations in 32 countries and serves customers in 85 nations.

1.5 PROBLEM DEFINITION

Like any other big company Sonoco prospered through mergers and acquisitions but with this came the problem of managing the various Enterprise Resource Planning (ERP) systems that now became a part of the company. With the number of ERP and legacy systems on the increase, it became extremely difficult to obtain consolidated supplier or commodity spend details, this in turn propelled the need to have a common place to store data from other sources.

Over the past year and a half, Corporate Supply Management, Corporate IT, the Project Management Office and the Finance department together worked to develop a data warehouse, which includes all supplier spend information for US and Canada based plants that have supplier payments made through the PeopleSoft financial database system. The data warehouse provides significant benefits to the employees and to the company as a whole, by enabling users to retrieve information in large quantities for analysis purpose, the results of which are then utilized for negotiation with vendors, contract formation and other activities pertaining to the company's vendor development and purchasing strategies.

Here the matter of concern is the inefficient performance of the data warehouse in terms of poor quality data and time required to access information. Extensive research on the warehouse and inputs from supply management personnel revealed that the inefficiency is caused mainly due to technical and design issues. This problem is further aggravated because of the users' failure to understand the data warehouse. The goal of this study is to provide an overview of the data warehouse concept and architecture, conduct an in-depth study of the post-implementation stage issues that surround the warehouse and determine feasible solutions that can be implemented to improve its performance.

2 RELATED RESEARCH

2.1 BASIC ARCHITECTURE OF A DATA WAREHOUSE

For business users to access information anytime and anywhere, a robust data warehouse architecture is required that can accommodate data from both the external and internal sources. When deciding on architecture many factors like number of users, type of business, diversity of information, volume of data, data storage, data access, refresh cycle, etc need to be considered. Figure 1 depicts four discrete layers in the data warehouse architecture: source system, ETL, data repository and presentation.

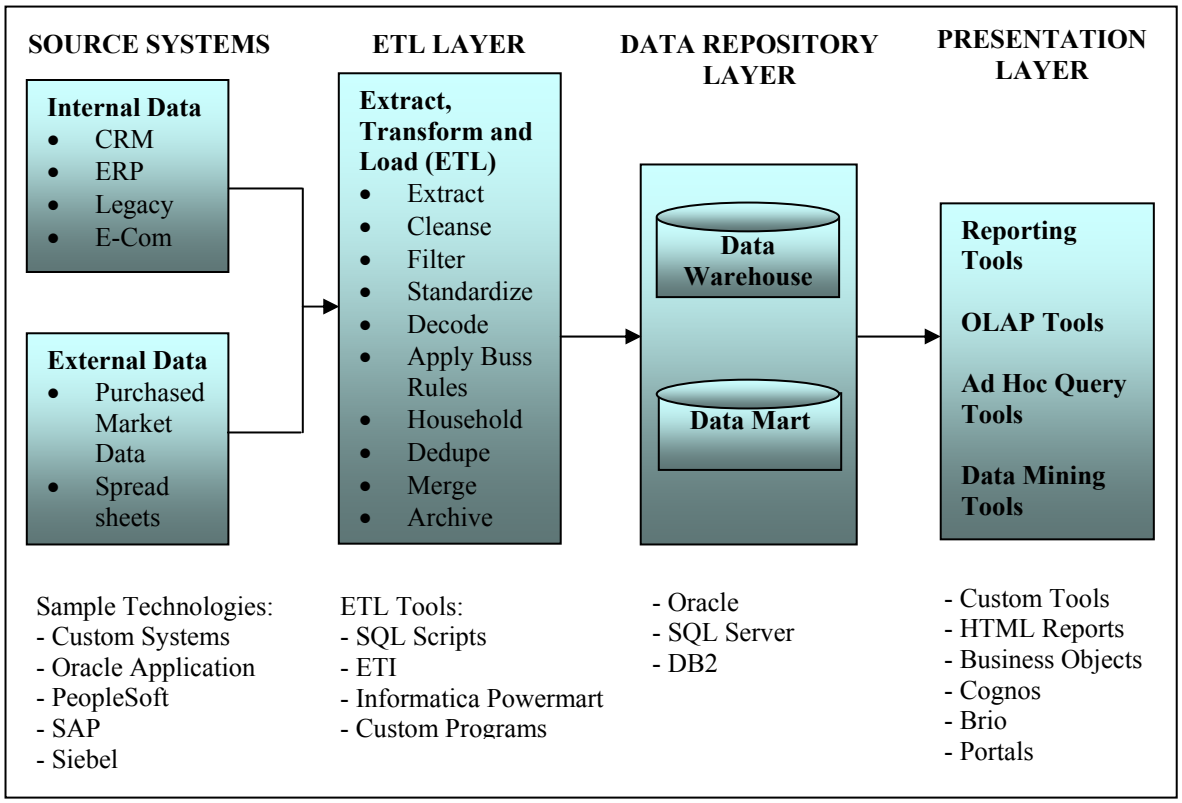

Figure 1: Basic Architecture of the Data Warehouse

SOURCE SYSTEMS

Source systems are transaction systems whose main function is to capture the day-to-day transaction activity details of the business. Source systems constitute of ERP systems, customer relationship management (CRM) systems, legacy systems and external data. These systems are different from the data warehouse database as they maintain little historical data and are not very user friendly for querying and reporting purpose.

EXTRACT, TRANSFORM, LOAD

The ETL layer poses the biggest challenge in implementing a data warehouse and comprises a set of processes that perform gathering, refining, cleansing and aggregating of data from the source systems. The primary responsibility of this layer is data consolidation and integration; hence the software to be used for this purpose should be designed and developed in a generic way.

DATA REPOSITORY LAYER

This layer acts as the storage area for the prepared data from the ETL layer. The data in this layer is stored in a specific format and at a very granular level so as to provide ultimate flexibility to the users by satisfying their different requirements.

PRESENTATION LAYER

This layer is the primary point of contact between the end user and the data warehouse. It serves as a user interface and supports a set of presentation and analytical tools for accessing information. It also gives a common look and feel to the reports and queries made available to the users. In short, this layer provides the users with the end product, information.

If any of the above functions are missing or not properly performed, the data warehouse cannot work towards fulfilling its goals for which it was designed.

2.2 DATA FLOW

The main priority for most companies is uptime and availability of source systems. The high performance of these systems is maintained by moving most of the historical data from their database to the data warehouse database. The arrows mark the flow of data across the system in figure 1. The ETL layer serves as an interface between the source systems and data warehouse. It converts the format of the data from the source to the format and model used by the warehousing system and also keeps track of changes occurring in the source data. When new or modified data is added, the ETL layer extracts the required information from the source, processes it, which may involve filtering the information, summarizing it or merging it with information from other sources and loads

it into the data warehouse database. The propagation of information from the source system to the warehouse is mostly performed as a batch process. The presentation layer enables the users to directly query the data warehouse database and retrieve information in the required format. Although figure 1 shows a single, centralized system, the warehouse can also be implemented as a distributed system, which sometimes may be necessary for the desired performance.

2.3 SONOCO DATA WAREHOUSE ARCHITECTURE

Figure 2 illustrates the company's data warehouse architecture. This architecture is based on the basic data warehouse system architecture explained in section 2.1. Source system layer comprises supplier maintenance application system, purchasing application systems, Ariba e-procurement tool and external data in the form of bank statements. Batch processes extract the required data from the source systems and transform it into the required format before loading it into the PeopleSoft Financial database. A different set of processes move the spend data from PeopleSoft and Ariba to the data warehouse database. The final layer houses the user interface.

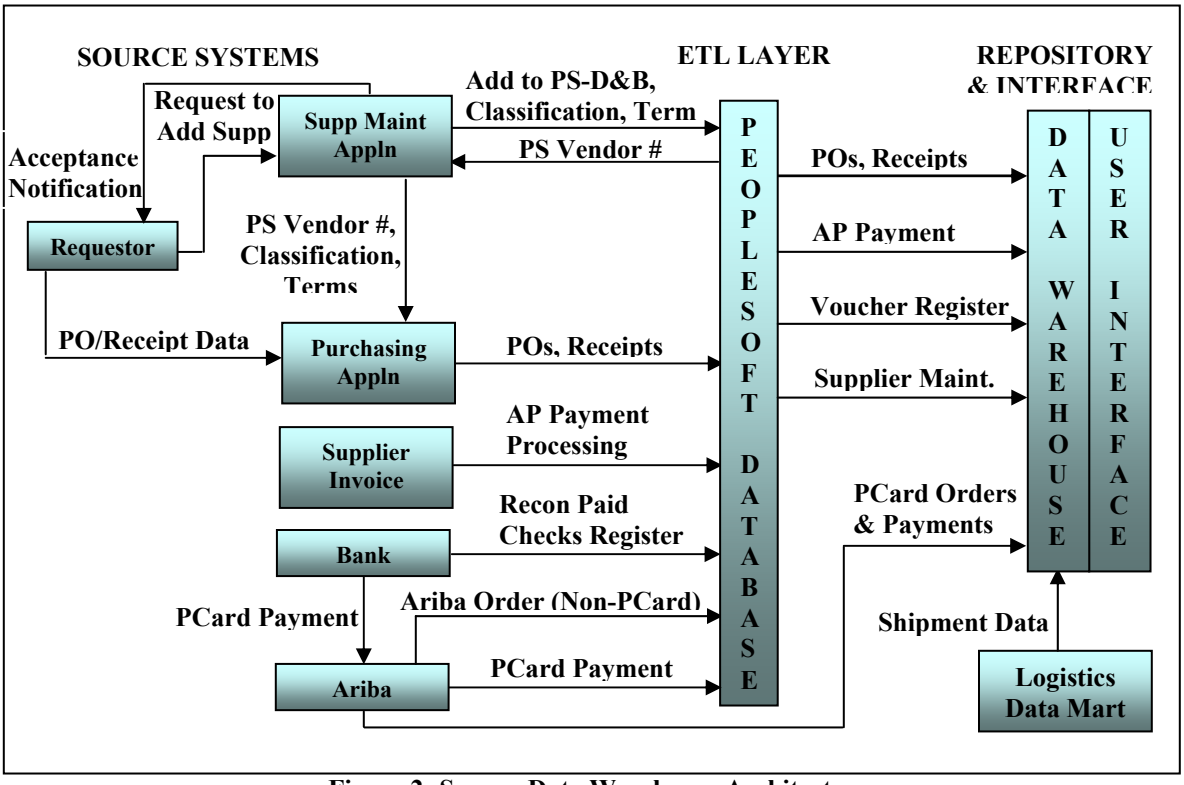

Figure 2: Sonoco Data Warehouse Architecture

2.3.1 SOURCE SYSTEMS

Supplier Maintenance Application System: This system is responsible for adding, deleting or modifying any information pertaining to new or existing suppliers. This system is also responsible for classifying suppliers according to their area of business and assigning them a suitable code. Supplier classification is performed to manage the company's large supplier base. Under this system, each supplier is assigned a four-digit alphanumeric code, which identifies the supplier's primary area of business. Each character of this four-digit code represents four distinct levels of classification.

Level 1: The first character of the code represents the major business sector the supplier belongs to. The five business sectors identified for Sonoco are: Direct Material (1000), Indirect Material (2000), Capital Expenditures (3000), Nonsourceable (4000) and Uncategorized (5000).
Level 2: The second character of the code represents the supplier's business group.
Level 3: The third character of the code represents the supplier's product/service group.
Level 4: The fourth character of the code represents the supplier's product or service.
If a supplier is assigned a subclass code[1] 2231 then it can be read as

	Supplier Classification Code	Supplier Classification Description
Major business sector	2000	Indirect Materials
Business group	2200	Operating Supplies and Services
Product/Service group	2230	MRO
Product/Service	2231	Belts

Table 1: Supplier Classification and Coding System[2]

A supplier may provide multiple products or services but the classification code assigned is based on the maximum percentage of spend in a particular product or service. Also, multiple suppliers providing same or similar products are assigned the same code as shown in table 2. To protect the privacy of the company the suppliers are named arbitrarily in the table below and throughout this report where required.

[1] Shortened for supplier classification code
[2] For a complete list of supplier classification for indirect materials, refer to Appendix I

S.No	Supplier Name	Product/Service Description	Percentage Spend (%)	Subclass Code
1	Supplier A	Belts	80	2231 (Belts)
		Gaskets	20	
2	Supplier B	Belts	70	2231 (Belts)
		Gaskets	20	
		Bushings	10	

Table 2: Supplier Classification Based on Highest Percentage Spend

When a new supplier is added, in addition to name, identification number and address, the supplier classification code is also entered into the PeopleSoft database.

Purchasing Application System: These systems like Baan, TQMP, Globetech, etc, are used for daily transaction processes involving online creation of purchase orders and entry of receipt and invoice details. Some of the important information captured by these systems include purchase order number, supplier name, supplier id, division commodity code, division commodity description, receipt and invoice number, quantity purchased, unit price and total amount paid to the supplier. The company has close to twenty-five purchasing application systems but at present only nineteen of them are directly connected to the data warehouse through PeopleSoft while the data from the rest of the systems is manually entered.

Ariba: This is an e-procurement system, which helps in streamlining the requisition process, purchase order transmission and management and supplier collaboration. It also provides visibility into all spend categories by keeping track of all purchase order details and identifies opportunities for supplier negotiation and contract. There are two types of transactions that can be performed through Ariba – 1.) Traditional transactions and 2.) Purchasing card (PCard) transactions. Traditional transactions are similar to transactions made online through purchasing application systems with payment being made as per the agreement between supplier and the company. On the other hand, PCard transactions through Ariba also involve online creation of purchase orders but the payment is

instantaneously made using PCards, which are visa credit cards and issued specifically for a company.

PCard transactions can also be made by placing the order to the supplier through phone or in person; such transactions do not result in the generation of a purchase order and are termed as PCard transactions outside Ariba.

External Source: The bank provides monthly PCard statement for all PCard transactions through and outside Ariba. This statement does not include any drilled down expense report but mainly contains information pertaining to the supplier and total spend amount.

2.3.2 EXTRACT, TRANSFORM AND LOAD

These functions are performed by different set of batch processes. First, all purchase order, receipt and invoice details are transferred to PeopleSoft for financial purpose then this data along with any payment details is moved to the data warehouse. PCard transaction details are uploaded separately into the system once the statement is received from the bank.

2.3.3 DATA REPOSITORY

The figure below is not a true replica of the data warehouse database schema but has been included to provide an overview of the major tables in the database, the way they are linked to each other and the kind of information that is available to the users.

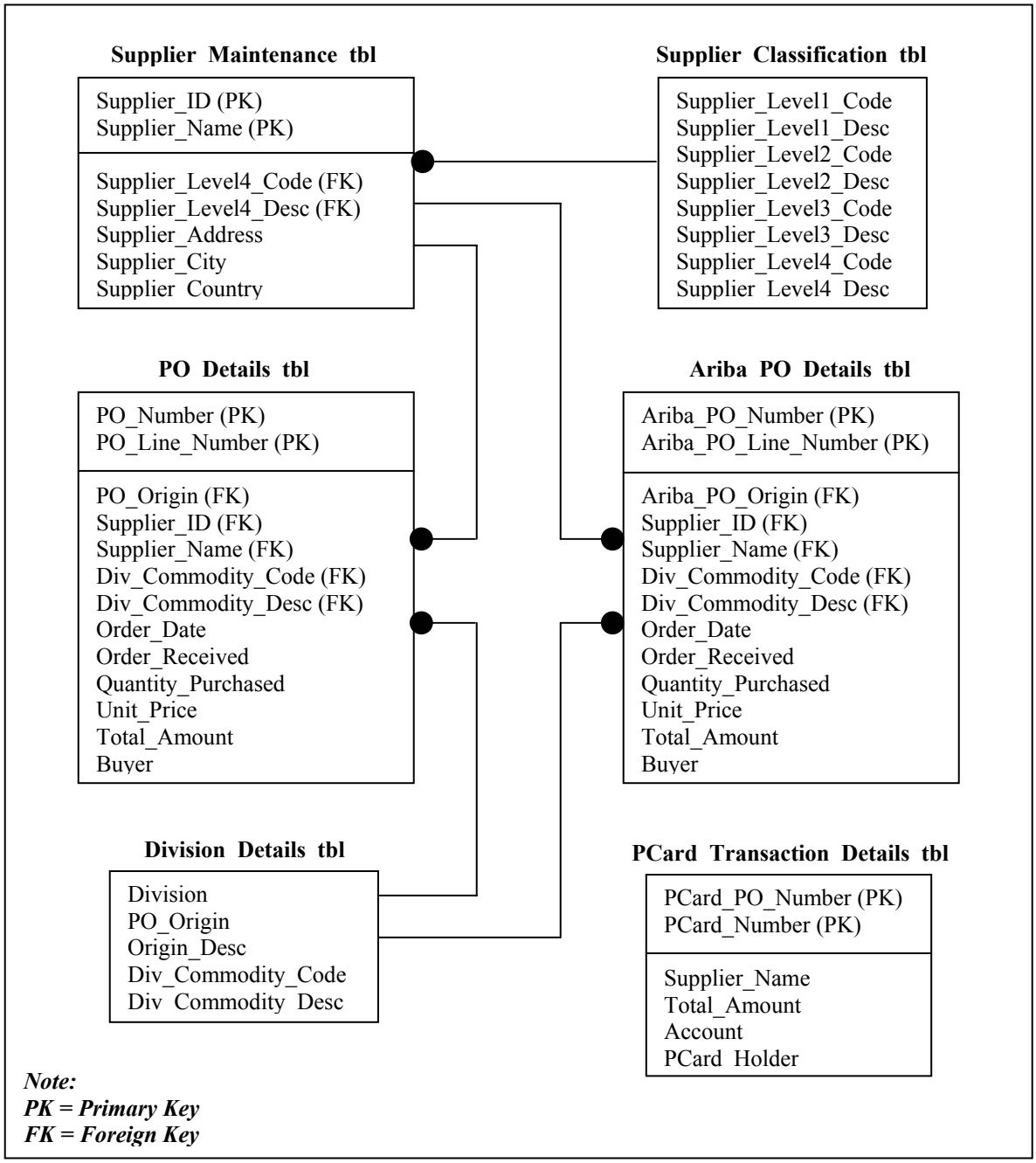

Figure 3: Data Repository Layer

2.3.4 DATA PRESENTATION

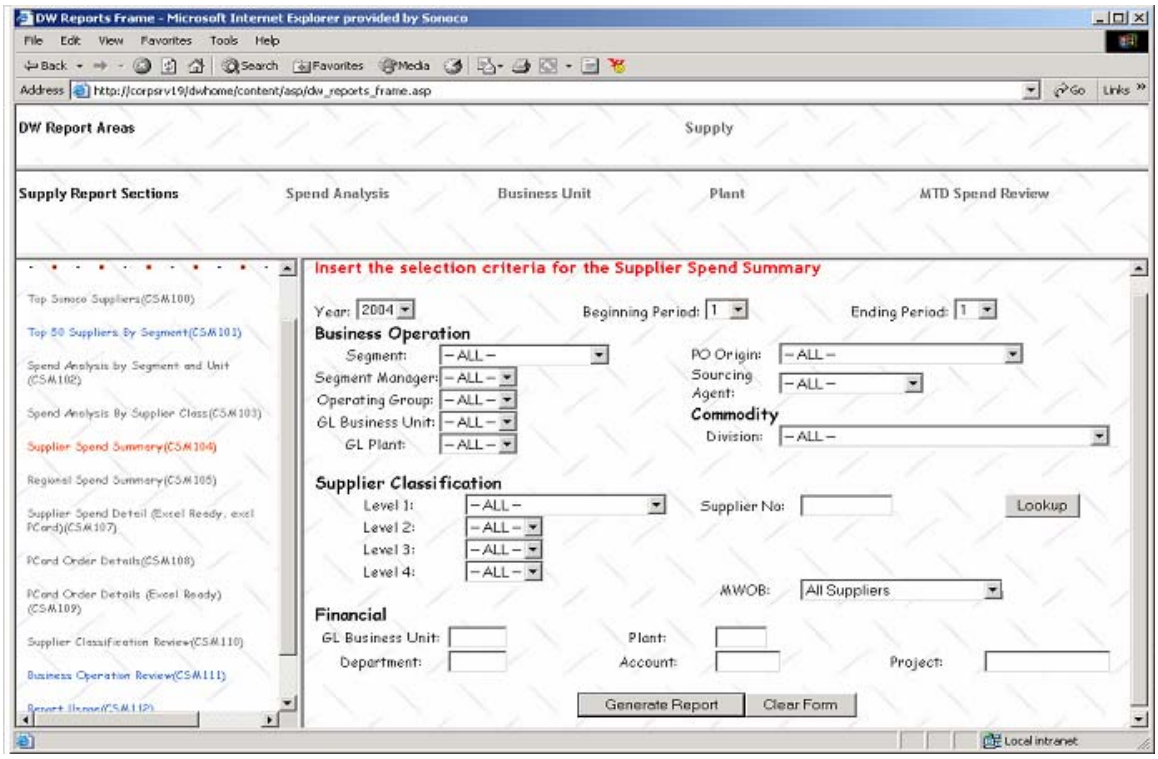

Figure 4: Data Warehouse User Interface

The user interface is divided into two vertical columns as shown in the above figure. The left column lists the various reports that can be generated and the right column houses the parameters essential for querying. Some of the most frequently used parameters are:

Supplier Classification: Provides spend details for all suppliers who are classified under the selected category.

PO Origin: Provides details of all transactions made through a particular application system like Ariba, TQMP, etc.

Division: Allows users to get commodity spend for a particular division.

Supplier No: Provides spend details for the selected supplier

A user is required to select one or more of these parameters to get the required data. Each additional parameter selected further refines the search and provides the user with more specific information. Thus, the interface provides flexibility in terms of obtaining drilled down or rolled up spend details.

3 ISSUES

In this chapter we try to understand the common problems that frequently surface once the data warehouse is established. A good understanding of these issues is critical in order to make effective recommendations.

INABILITY TO CAPTURE CORPORATE SPEND FOR A COMMODITY:

Division commodity codes are specific for each division. This results in different codes and descriptions referring to the same or similar product. For example, the table below shows six different division commodity codes for paper:

Division	Division Comm. Code	Division Comm. Code Description
Ariba	14120	Paper - Industrial
Consumer Products Division	PPA	Paper
Consumer Products Division	BST	Paper BTM Stock
Specialty Packaging	R28	Paper
Industrial Products Division	PA	Sonoco Paper
Industrial Products Division	PP	Paper Outside Purchased

Table 3: Division Commodity Codes

To obtain the total spend for any commodity, the current scenario requires the user to know all the different division commodity codes, generate as many spend reports as the number of commodities and consolidate the spend amount to obtain the final value. This process is tedious, time consuming and error prone as the user may not be aware of all the existing division commodity codes leading to erroneous analysis of incomplete data, which eventually impacts the decision making process.

HIGH LEVEL SUPPLIER CLASSIFICATION:

The existing supplier classification and coding system follows a very high level classification of suppliers, consequently, users end up with not-so accurate supplier spend information. As shown in table below, both suppliers A and B are classified as 'MRO Other' suppliers although they provide different products. This broad way of classification does not provide any valuable insights to the suppliers' spend information or area of business.

SUPPLIER CLASSIFICATION AND CODING				PRODUCTS/SERVICE	'MRO OTHER' SUPPLIER
Level 1	**Level 2**	**Level 3**	**Level 4**		
Indirect (2000)	Operating Supplies & Services (2200)	MRO (2230)	MRO Other (2230)	Non Production Steel	Supplier A
				Plastics	Supplier B

Table 4: High Level Supplier Classification

UNAVAILABILITY OF LEVEL 3 SPEND DETAILS FOR PCARD PURCHASES

The most important characteristic of PCard transactions is that they do not involve creation of purchase order or handling of receipts and invoices. On one hand this adds efficiency to the purchasing process as non-value added tasks are eliminated, on the other it makes it very difficult to record and keep track of transaction details. As a result, the company is left with little understanding of the items bought or the price distribution.

UNCATEGORIZED SUPPLIERS

From a total of 56,846 suppliers that are present in the company's supplier maintenance system today, 16,378 suppliers remain unclassified, mainly, due to lack of supplier information, unavailability of purchase order details and sometimes even negligence. Together they are responsible for over $ 100,000,000.00 annually. This amount remains unaccounted for and adds no value to the users' decision-making process.

MULTIPLE ENTRIES OF SAME VENDOR

Many instances like the ones shown in table 5 have been found where the same supplier has been added into the system more than once. Such multiple entries bring up inaccurate spend details, cause confusion and are an overload to the system.

CAUSE OF MULTIPLE ENTRIES	EXAMPLE	
Use of Abbreviation	WALMART	WM
Incorrect Spelling	US POSTAL SERVICE	US POASTAL SERVICE
Lack of Standards	THE STEEL CO.	STEEL CO., THE

Table 5: Examples of Multiple Entries

4 SOLUTIONS AND RECOMMENDATIONS

4.1 CORPORATE COMMODITY CLASSIFICATION SYSTEM

We implemented a corporate commodity classification system to allow users to obtain consolidated spend for a given commodity. The steps taken to set up this system are explained in detail below:

Identifying Corporate Commodities

A consolidated list of all division commodities was created, analyzed and a new list of distinct commodities was prepared. This list was communicated to the commodity managers and buyers requesting them for their inputs with regards to adding more new commodities or making changes to the existing ones. Their suggestions were included and the final list was developed.

Classification Strategy

A system similar to the supplier classification was adapted under which each corporate commodity was assigned a five digit alphanumeric code. Each character of this code represents the five levels of the corporate commodity classification system.

Level 1: The first character of the code represents the major business sector the commodity belongs to.

Level 2: The second character of the code represents the business group

Level 3: The third character of the code represents the product or service group

Level 4: The fourth character of the code represents the product/service

Level 5: The fifth character of the code further classifies the product/service

The first four levels of this classification system are designed similar to the supplier classification system as users are quite familiar with the existing structure plus this has already proven to work well with Sonoco's requirement. The corporate commodities identified above form the next level and were mapped to the corresponding level-4 commodities and assigned a proper code to complete this hierarchical structure. The table below shows further classification of code 2335 to accommodate five more categories.

Level 1	Level 2	Level 3	Level 4	Level 5	
				Commodity Description	Corporate Comm. Code
Indirect (2000)	SGA (2300)	Facilities Management (2330)	Maintenance (2335)	Maintenance	2335A
				Cooling	2335B
				Heating	2335C
				Humidity Control	2335D
				Buildings & Grounds - Services	2335E

Table 6: Further Classification of Level-4 Corporate Commodity Code

Mapping Division Commodities

Once the commodity classification system was set in place, division commodity codes which represent same or similar products, were grouped and mapped to their respective parent level-5 commodity code as shown in table 7 (for more examples, see Appendix II).

Level 1	Level 2	Level 3	Level 4	Level 5 Corporate Comm. Code	PO Origin	Div. Comm. Code	Division Comm. Code Description
Direct (1000)	Fiber (1100)	Finished Paper (1110)	Finished Paper Other (1110)	Paper (1110C)	ANP	14120	Paper - Industrial
					APC	14120	Paper - Industrial
					ARI	14120	Paper - Industrial
					CPD	PPA	Paper
					IPD	PA	Paper
					BAN	R28	Paper
					IPD	PA	Sonoco Paper
					IPD	PP	Paper Outside Purchased
					CPD	BST	Paper BTM Stock

Table 7: Mapping Division Commodity Codes to level – 5 Corporate Commodity Code

Changes to Data Warehouse Database

Although the supplier classification and commodity classification system are similar in structure, the two systems are completely independent of each other and are required to maintain different sets of codes. Consequently, few changes were made to the database – a new table describing the commodity classification structure was added and division commodity code table was updated to include the mapping.

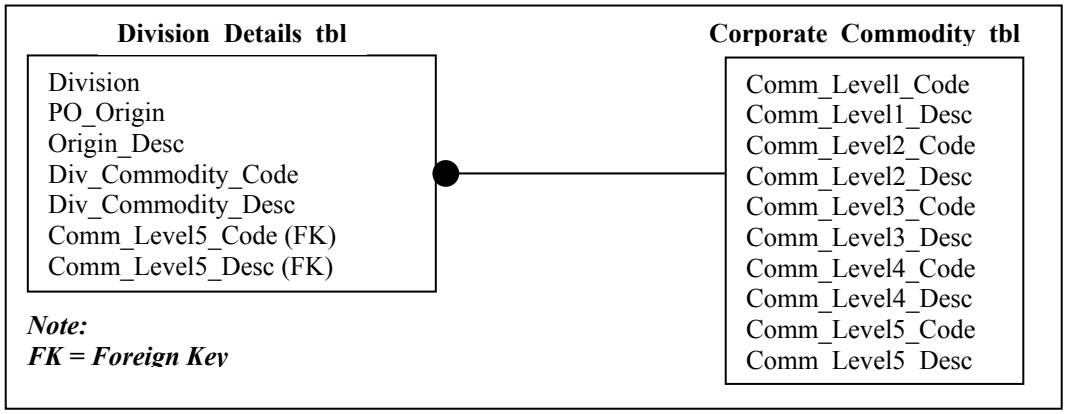

Figure 5: Changes to DW Database to Implement Corporate Commodity Classification

Changes to User Interface

Changes were also made to the user interface to reflect the mapping between the division and corporate commodity codes. A parameter called 'Corporate' was added which houses all the level-5 corporate commodity codes and provides the corporate commodity spend on querying. 'Division' parameter provides the division commodity spend as before.

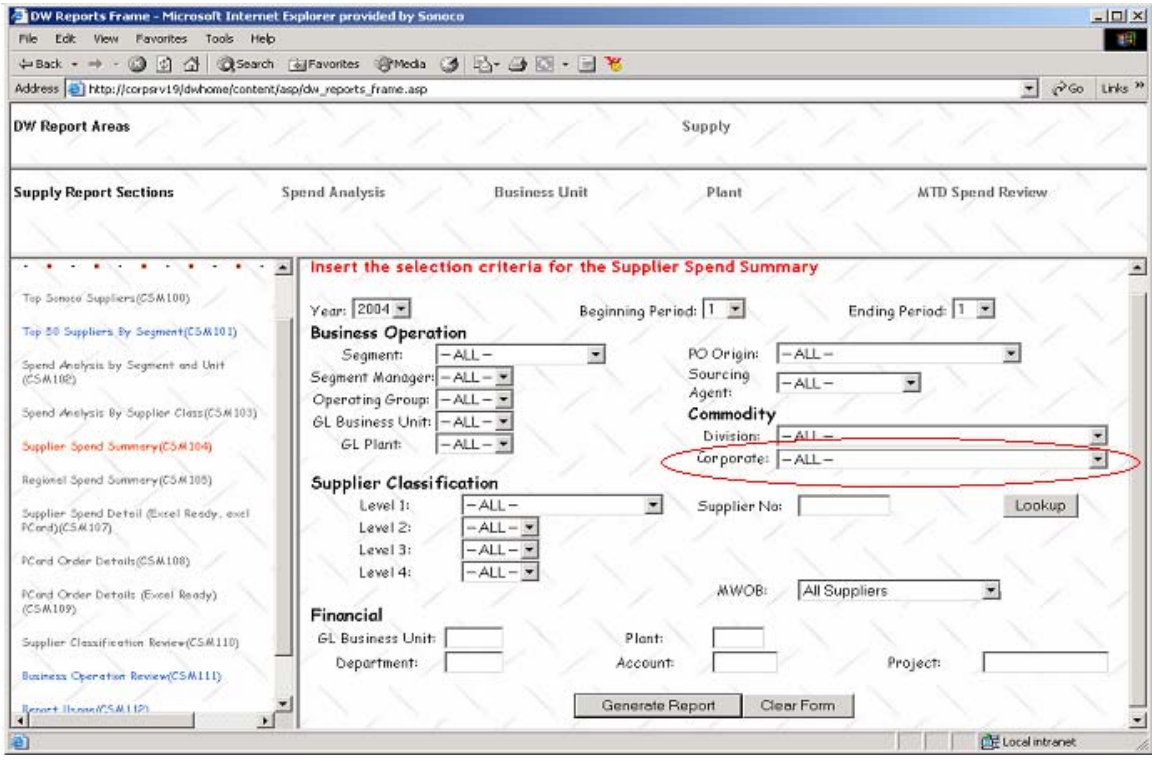

Figure 6: Corporate Commodity Parameter Added to the User Interface

4.2 MODIFICATION OF SUPPLIER CLASSIFICATION SYSTEM

To provide users with more credible supplier spend estimates new business categories were added to the existing list of supplier classification and in the process, suppliers who were found to be in the wrong category were reclassified. The process detail is described below:

Data Collection

Spend detail reports were generated for each subclass code and analyzed in detail. Similar commodities were grouped together and their total spend was calculated. If the spend amount was found to form a major proportion of the total subclass code spend, the commodity group was decided to be added as a new subclass code and the corresponding suppliers were reclassified. For commodities where spend details were not available we consulted with the commodity managers to decide on the new categories.

Few of the classifications added to 2330 – Facilities Management (FM) is shown in table 8. Refer to Appendix I for a complete list of classifications for indirect materials with the new categories highlighted in bold.

Subclass Code & Description	Subclass Spend (2003)	Commodity/ Service Group	Spend for Comm. Group ($) (2003)	% Of FM Spend	New Subclass Code
2330 'Facilities Management'	35,370,284	Roof (Repair and Replace)	527,890.00	1.5	2339
		HVAC	2,812,627.00	8	233I
		General Construction	812,096.00	2.3	233N

Table 8: New Supplier Categories

4.3 LEVEL 3 SPEND DETAILS FOR PCARD PURCHASES OUTSIDE ARIBA

Banks are the only source for obtaining complete spend details for PCard purchases but this functionality comes with an additional cost which is found to be quite high considering that for Sonoco, annual PCard spend forms only about 3 – 4 % of the total annual spend. To overcome this problem an alternate but efficient method of moving Non Ariba PCard suppliers to Ariba was considered and implemented. Moving the suppliers

to Ariba was advantageous because it not only included the flexibility of PCard usage but also provided transaction details to the end users and management.

To create the list of desired Non Ariba PCard suppliers we performed spend analysis on the following criteria:

4.3.1 SPEND RANGE

Table 9 lists the number of Non Ariba PCard suppliers in their respective spend ranges. As can be seen, out of a total of 13629 Non Ariba PCard suppliers, 37 suppliers fall in the very high spend range. This is highlighted in bold in the table below. These suppliers are recommended to be added to Ariba.

SPEND RANGE ($)	NUMBER OF NON ARIBA SUPPLIERS
0 – 2,000	10381
2,000 – 5,000	1466
5,000 – 10,000	775
10,000 – 25,000	602
25,000 – 50,000	237
50,000 – 80,000	78
80,000 – 150,000	53
150,000 and above	**37**
Total	13629

Table 9: Suppliers Classified by Annual Spend Range

4.3.2 HIGH DOLLAR SPEND

This category lists suppliers in the decreasing order of their annual spend with the cut off criterion set as $ 150,000.00. The top ten suppliers are listed in table 10. Suppliers with a high annual spend combined with a moderate number of transactions (800 and above) are the ones recommended to be moved to Ariba. These suppliers are highlighted in bold below. For more details refer to Appendix III.

S.NO	SUPPLIER ID	SUPPLIER NAME	ANNUAL SPEND ($) (2003)	NO. OF TRANSACTIONS
1	NPC0000119	SUPPLIER 1	1,659,039	434
2	NPC0000011	**SUPPLIER 2**	612,800	948

3	NPC0019419	**SUPPLIER 3**	556,370	1,578
4	NPC0000047	**SUPPLIER 4**	504,114	3,822
5	NPC0011396	SUPPLIER 5	431,779	148
6	NPC0000065	**SUPPLIER 6**	409,704	2,310
7	NPC0005987	**SUPPLIER 7**	407,647	2,742
8	NPC0017032	SUPPLIER 8	402,740	136
9	NPC0011891	SUPPLIER 9	365,969	793
10	NPC0012168	**SUPPLIER 10**	322,412	828

Table 10: List of Top Ten High Dollar Spend Suppliers

4.3.3 HIGH FREQUENCY SPEND

This category lists suppliers in decreasing order of their annual frequency[3] with the cut off parameter set as 900 transactions. The top ten suppliers are listed in table 11. Suppliers who not only have high frequency but high annual spend ($ 100,000.00 and above) as well are recommended as definite candidates to be moved to Ariba. These suppliers are shown in bold below. For more details refer to Appendix IV.

S.NO	SUPPLIER ID	SUPPLIER NAME	ANNUAL SPEND (2003)	FREQUENCY
1	NPC0000047	**SUPPLIER 1**	504,114	3,822
2	NPC0005987	**SUPPLIER 2**	407,647	2,742
3	NPC0000065	**SUPPLIER 3**	409,704	2,310
4	NPC0022292	**SUPPLIER 4**	301,887	1,687
5	NPC0019419	**SUPPLIER 5**	556,370	1,578
6	NPC0000056	**SUPPLIER 6**	156,598	996
7	NPC0006042	**SUPPLIER 7**	108,454	996
8	NPC0000011	**SUPPLIER 8**	612,800	948
9	NPC0014475	**SUPPLIER 9**	201,485	928
10	NPC0010702	SUPPLIER 10	76,672	916

Table 11: List of Top Ten High Frequency Suppliers

4.3.4 HIGH AVERAGE SPEND PER TRANSACTION

This category lists suppliers in decreasing order of their average spend per transaction with the cut off criterion set as $ 4000.00 per transaction. The top ten suppliers are listed

[3] Number of transactions

in table 12. Suppliers who not only have high average spend per transaction but also high annual spend ($ 100,000.00 and above) and moderate number of transactions (15 and above) are recommended to be moved to Ariba. These suppliers are highlighted in bold below. For more details refer to Appendix V.

S.NO	SUPPLIER ID	SUPPLIER NAME	ANNUAL SPEND (2003)	FREQUENCY	AVERAGE SPEND[4] PER TRANSACTION
1	NPC0016890	**SUPPLIER 1**	110,186	15	7345.74
2	NPC0022486	**SUPPLIER 2**	158,653	22	7211.49
3	NPC0020162	**SUPPLIER 3**	108,155	15	7210.35
4	NPC0017347	**SUPPLIER 4**	150,375	21	7160.71
5	NPC0000908	**SUPPLIER 5**	180,550	28	6448.23
6	NPC0013540	**SUPPLIER 6**	135,091	24	5628.79
7	NPC0014228	SUPPLIER 7	99,534	18	5529.67
8	NPC0018698	**SUPPLIER 8**	250,936	48	5227.84
9	NPC0021771	**SUPPLIER 9**	254,095	49	5185.62
10	NPC0007301	**SUPPLIER 10**	87,742	18	4874.58

Table 12: List of Top Ten Average Spend/Transaction Suppliers

4.3.5 80%-20% ANALYSIS

This analysis is performed to determine the total number of suppliers who are responsible for 80 % of spend. From our analysis we found that a total of 1461 suppliers (10.7 % of total Non Ariba PCard suppliers) together make for 80 % of the Non Ariba PCard spend. These suppliers are recommended to be moved to Ariba. For details and methodology, refer to Appendix VI.

4.4 CLASSIFYING UNCATEGORIZED SUPPLIERS

As mentioned in the previous section, a total of 16,378 suppliers are unclassified. Together they are responsible for about 7% of the annual spend. Transaction details also revealed that this list comprises of two distinct types of suppliers – Non Ariba PCard (NPC) suppliers and other suppliers. For 80% – 20 % analysis these suppliers were considered separately.

[4] = Annual Spend/Frequency

	NUMBER OF	ANNUAL SPEND (2003)	PERCENTAGE SPEND (%)
TOTAL SUPPLIERS	56,846	1,468,832,678.00	-
UNCATEGORIZED SUPPLIERS	16,378	100,370,479.00	7

Table 13: Annual Spend by Uncategorized Suppliers

	NUMBER OF	ANNUAL SPEND (2003)	PERCENTAGE SPEND (%)
UNCATEGORIZED SUPPLIERS	16,378	100,370,479.00	-
UNCATEGORIZED – NON ARIBA PCARD SUPPLIERS	7,810	17,780,805.00	18
UNCATEGORIZED - OTHER SUPPLIERS	8,568	82,589,674.00	82

Table 14: Annual Spend by Uncategorized NPC and Other Suppliers

4.4.1 80% - 20% ANALYSIS

Lack of purchase order details for uncategorized - NPC suppliers made it difficult to classify them. It was also noted that 80 % of uncategorized – NPC spend made only for 14 %[5] of the total uncategorized spend. Hence, trying to classify these suppliers at this stage was not considered very contributive to the whole task.

For uncategorized other suppliers it was found that a total of 1507 suppliers were responsible for 80 % of the total uncategorized other spend. Further analysis revealed that a total of 2585 uncategorized other suppliers were responsible for 90 % of the total uncategorized other spend which in turn made for 75 % of total uncategorized spend.

	NUMBER OF	TOTAL AMOUNT ($)	% SPEND[6]
UNCATEGORIZED OTHER SUPPLIERS	1507	66,071,739.00	66 %
	2585	74,331,574.00	75 %

Table 15: Uncategorized Suppliers

[5] 0.8 * 17.7 M ≅ 0.14 * 100 M
[6] (66 M/100M) * 100 and (74 M/100M) * 100

4.5 REMOVING MULTIPLE ENTRIES FOR VENDORS

The company's supplier maintenance database has close to 60,000 suppliers. Searching for every instance of multiple entries from this huge stack of records is not a practical solution nor does it provide great benefits for one's efforts, as many of these suppliers are inactive. But there are procedures that if strictly followed will reduce this problem to a great extent.

1. Deactivate suppliers: Suppliers who have not been used for a long time (18 months or more) should be deactivated in the system. Anytime a buyer wants to activate a supplier, the database should be thoroughly searched (refer to step 4: Standard Procedure) for that supplier and all multiple entries should be deactivated, thus restricting the search to as-and-when-it-is-required basis.

2. Search for Preferred Supplier: Preferred suppliers form just 2% of the total suppliers. This list can be further broken down into much smaller groups as – most preferred suppliers, moderately preferred suppliers and least preferred suppliers. Multiple entries for each of these groups should be searched (refer to step 4: Standard Procedure) on priority basis and corresponding multiple entries should be deactivated.

3. User Access: This problem predominantly occurs when multiple users have access to enter or modify supplier details in the system. This results in lot of variances and increases the chances of creating multiple entries. In view of this it is recommended that only a single user (Supply Maintenance Application Manager) be given complete access to change any information in the system and he should maintain record of any deviation from the standard procedure.

4. Standard Procedure

   Supplier maintenance application manager should follow the steps given below to add or edit any information in the system:

Step 1: Request buyer/commodity manager/supply management personnel to complete the form (template given in table 16) to request any type of change in the system.

SUPPLIER MAINTENANCE REQUEST FORM	
Requester:	
Requester Phone:	
Requester Department:	
Type of Change	• Add a new supplier • Edit an existing supplier • Deactivate an existing supplier
Purchasing Application System	
Authorized by: (Dept Head /Commodity Mgr)	
TO ADD A NEW SUPPLIER	
Supplier Name:	
Supplier Phone:	
Supplier Fax:	
Remit to Address:	
City:	
State:	
Postal Code:	
Buy from Address *(If different from remit to address)*:	
City:	
State:	
Postal Code:	
Parent Company:	
Contact At Company:	
Contact Phone:	
Supplier Classification:	
Items Being Purchased:	
Is this supplier a preferred supplier?	
EDIT AN EXISTING SUPPLIER	
Reason for Edit Request:	
Supplier ID:	
Supplier Name:	
Supplier Phone:	
Describe in detail the information that needs to be edited:	

DEACTIVATE AN EXISTING SUPPLIER	
Reason for Deactivate Request:	
Supplier Name:	
Supplier ID:	
Supplier Phone:	

Table 16: Supplier Maintenance Request Form

Step 2: Search the database for the requested supplier by

a. Name: e.g. Wal-Mart
b. Part of the name: e.g. Wal-M, W-Mart, Wmart, etc.
c. Abbreviation: e.g. WM
d. Address: e.g. North 2nd Street
e. Part of the address: e.g. N. 2nd Street, North 2nd St., North II Street, etc.
f. Telephone number
g. City
h. State
i. Postal Code

Step 3: Compare the results found at each search from step 2 with the requested supplier.

Step 4:

Case I: If no match is found

- For add request
    - Add the requested supplier into the system.

Case II: If a single match is found

- For add request
    - Do not add the requested supplier
    - Edit information if needed
    - Convey changes to requester
- For edit request
    - Edit the information as requested

- For deactivate request
    - Deactivate the supplier as requested

Case III: If multiple matches are found

- For add request
    - Do not add the requested supplier
    - Deactivate all but one entry
    - If needed, edit information to the active entry
    - Convey changes to requester, PeopleSoft application manager and data warehouse application manager
- For edit request
    - Deactivate all but one entry
    - Edit the information as requested to the active entry
    - Convey changes to requester, PeopleSoft application manager and data warehouse application manager
- For deactivate request
    - Deactivate all entries

5 DATA WAREHOUSE RETURN ON INVESTMENT

In the previous chapters we emphasized on improving the quality of data without quantifying its advantages. In this chapter we evaluate the cost verses benefits of undertaking these changes. It is important to understand that measurements can be predicted for only those benefits that are tactical and tangible such as dollars saved or hours reduced. Several financial measures can be applied such as the internal rate of return (IRR), net present value (NPV), payback period and return on investment (ROI). While each has its benefits, for our purpose we have employed ROI.

With this financial measure several assumptions were made with regards to estimating the cost of the data warehouse project, the previous reporting environment and payback period on the basis of information available from the internal groups within the organization.

S.No	ESTIMATES FOR CALCULATING RETURN ON INVESTMENT	
CASE I: Cost Estimate of Data Collection and Reporting Prior to Data Warehouse Implementation		
1	Number of hours/month per person spent on reporting	60
2	Number of persons	4
3	Average hourly labor rate	$ 30.00
CASE II: Cost Estimate of Data Warehouse Implementation		
DIRECT COST		
1	Hardware	Existing Server
2	Software	Existing License
3	Internal labor	$ 150,000.00
4	External labor	0
5	Data Warehouse developer training	$ 1000.00
INDIRECT COST		
6	Upgrade of client machine	0
7	Upgrade of network communication	0
8	Upgrade of support software	0
9	Additional internal support	0
10	End user training	$ 1,000.00

ONGOING SUPPORT		Number of Persons	% Time
1	Operational support personnel	1	50
2	Database administrator	1	10
3	System administrator	1	10
4	Network administrator	1	10
5	Trainer	1	10
6	Software Maintenance Cost		$ 10,000.00/year
OTHER DATA			
1	Number of Work Hours per Individual per year		2000
2	Average Hourly Labor Rate		$ 30.00
3	Time Horizon or Payback Period		5 years
4	Investment Yield		9 %
5	Additional Number of Hours/Month of Support because of Existing Problems in Data Warehouse		10
CASE III: Cost Estimate of Additional Effort to Implement Suggested Changes in Data Warehouse			
1	Number of Hours		400
2	Average Hourly Labor Rate		$ 30.00
3	Additional Number of Hours/Month of Support		2

Table 17: Cost Estimate for Calculating Return On Investment

COST CALCULATIONS

Case I: Cost of Data Collection and Reporting Prior to Data Warehouse Implementation						
	Year 0 ($)	**Year 1 ($)**	**Year 2 ($)**	**Year 3 ($)**	**Year 4 ($)**	**Year 5 ($)**
Total Cost	86,400.00	86,400.00	86,400.00	86,400.00	86,400.00	86,400.00

Table 18: Cost Calculation for Data Collection and Reporting Without Data Warehouse Application

Case II: Cost Calculation for Reporting Using Existing Data Warehouse						
	Year 0 ($)	**Year 1 ($)**	**Year 2 ($)**	**Year 3 ($)**	**Year 4 ($)**	**Year 5 ($)**
	150,000.00	10,000.00	10,000.00	10,000.00	10,000.00	10,000.00
	1,000.00	30,000.00[7]	30,000.00	30,000.00	30,000.00	30,000.00
	1,000.00	24,000.00[8]	24,000.00	24,000.00	24,000.00	24,000.00
		3,600.00	3,600.00	3,600.00	3,600.00	3,600.00
Total Cost	152,000.00	67,600.00	67,600.00	67,600.00	67,600.00	67,600.00

Table 19: Cost Calculation for Reporting Using Existing Data Warehouse

[7] =2000*30*0.5
[8] =2000*30*0.1*4

Case III: Cost Calculation for Reporting After Changes are Incorporated in Data Warehouse						
	Year 0 ($)	**Year 1 ($)**	**Year 2 ($)**	**Year 3 ($)**	**Year 4 ($)**	**Year 5 ($)**
	150,000.00	10,000.00	10,000.00	10,000.00	10,000.00	10,000.00
	1,000.00	30,000.00	30,000.00	30,000.00	30,000.00	30,000.00
	1,000.00	24,000.00	24,000.00	24,000.00	24,000.00	24,000.00
	12,000.00	720.00	720.00	720.00	720.00	720.00
Total Cost	164,000.00	64,720.00	64,720.00	64,720.00	64,720.00	64,720.00

Table 20: Cost Calculation for Reporting Using Improved Quality Data

NET SAVINGS						
	Year 0 ($)	**Year 1 ($)**	**Year 2 ($)**	**Year 3 ($)**	**Year 4 ($)**	**Year 5 ($)**
From Case II wrt Case I	N/A	18,800.00	18,800.00	18,800.00	18,800.00	18,800.00
Discounted Net Savings at 9%	N/A	17,247.12	15,823.96	14,517.36	13,317.92	12,218.12
TOTAL NET SAVINGS AT 9 %: $ 73,124.48						
From Case III wrt Case I	N/A	21,680.00	21,680.00	21,680.00	21,680.00	21,680.00
Discounted Net Savings at 9%	N/A	19,889.23	18,248.06	16,741.30	15,358.11	14,089.83
TOTAL NET SAVINGS AT 9 %: $ 84,326.53						
From Case III wrt Case II	N/A	2,880.00	2,880.00	2,880.00	2,880.00	2,880.00
Discounted Net Savings at 9%	N/A	2,642.11	2,424.10	2,223.94	2,040.19	1,871.71
TOTAL NET SAVINGS AT 9 %: $ 11,202.05						

Table 21: Net Savings and Discounted Net Savings

Return on Investment = Total Discounted Net Savings/ Investment

ROI for Implementing Data Warehouse

= (73,124.48/152,000.00) * 100 = **48 %**

ROI for Implementing Data Warehouse with Additional Effort for Better Quality Data

= (84,326.53/164,000.00) * 100 = **51.5 %**

ROI for Implementing Changes to Data Warehouse

= (11,202.05/12,000.00)*100 = **93.3 %**

6 RESULTS AND CONCLUSION

Corporate Commodity Classification System

This system was implemented keeping in mind the following criteria set forth by the company:

1. The classification must follow a simple structure.
2. The company must be able to implement it in a short period of time and
3. The system must be scalable.

In the process to set up the system we identified 361 distinct commodities and mapped 1185 division commodities to their corresponding corporate commodities. Simple changes were also made to the user interface (figure 4) to enable the user to obtain corporate commodity spend. Further, the hierarchical structure is designed to provide maximum flexibility in terms of adding new or modifying the existing commodities.

Level 3 Spend Details For PCard Purchases Outside Ariba

The 80 % - 20 % spend analysis resulted in highlighting 1461 suppliers responsible for 80 % of PCard spend. This list included the suppliers who were short listed as a result of other spend analysis mainly, Spend range, High annual spend, High frequency and High average spend per transaction. These suppliers are recommended to be added to Ariba. Although adding suppliers to Ariba will not completely rectify the problem of obtaining level 3 spend details but in the current state when company is looking for ways to reduce cost this strategy seems most suitable. Some other added advantages are in the area of sourcing and procurement and in terms of leveraging the entire spend lifecycle from analysis, sourcing, purchasing, tracking, and receiving to reporting.

Classification of Uncategorized Suppliers

Analysis on uncategorized suppliers spend led us to slightly deviate from the 80 % - 20 % criterion. We conclude that classifying 2585 suppliers (15.5 % of total uncategorized suppliers) will help us to capture 75 % of the uncategorized spend. Currently, efforts are

in progress to classify these suppliers on the basis of purchase order details, specifically commodity details, information available on the Internet and input from sourcing agents.

Multiple Entries for Same Vendor

For issues concerning multiple entries, trying to completely eliminate this problem is not considered a feasible solution. Instead emphasis is on preventing any more occurrences of multiple entries by following a standard procedure (section 4.5) for performing any changes in the system.

Data Warehouse ROI

Return on investment calculations yielded ROI that are significantly greater than 0 percent. This helps us to conclude that the data warehouse application is financially beneficial for the company. While the ROI is one measure to evaluate the Data Warehouse project, there are other intangible benefits that should also be addressed. These include improved information access, improved information dissemination, improved relationship with suppliers, better relationship between IT, users and management, propagation of knowledge about the organization through training and use of data warehouse application, reputation of the organization in terms of managing its data, major selling point to the potential employee, etc. Striking a balance between the financial measurement and the strategic benefits is crucial for the ultimate success of the data warehouse.

7 FUTURE SCOPE OF WORK

Over the coming years, the growth of data in the warehouse is going to be enormous with new product and supplier data being frequently added to the system. In order to get the most out of the system, it is important that the warehouse planners and developers should look for further scope of improvement to the data warehouse.

To begin with, the recommendations mainly, corporate commodity classification and supplier classification system studied in this report are open-ended solutions. For their successful implementation, these systems need to be regularly updated with new product or supplier information.

With the rise in PCard usage for purchasing of small value items, additional research in terms of cost analysis need to be performed to consider obtaining level 3 spend details directly from the banks. It was also noted that the travel and entertainment expense data does not currently reside in the data warehouse. Additional benefits can be obtained by the inclusion of these two distinct types of information.

A feature of the data warehouse is that it stores historical data even when it is not stored in the source systems. In many cases it is not required to store data forever. Techniques should be developed to differentiate between old and obsolete data and ensure that outdated information is automatically and efficiently purged from the data warehouse.

8 REFERENCES

1. Barry Devlin "*Data Warehouse from Architecture to Implementation*", Addison Wesley, (1997), p. 7, 20 – 21
2. Ralph Kimball, Laura Reeves, Margy Ross and Warren Thornthwaite "*The Data Warehouse Lifecycle Toolkit*", Wiley Computer publishing, (1998), p. 10,14 – 16
3. Joyce Bischoff and Ted Alexander "*Data Warehouse: Practical Advice from the Experts*", Prentice hall, (1997)
4. Jennifer Widom "*Research Problems in Data Warehousing*", Association for Computing Machinery, CIKM (Nov. 1995)
5. Larry P. English "*Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*", Wiley Computer Publishing, (1999), p.3 – 13
6. Larry Greenfield "*Getting Started with Data Warehousing*", Data Warehousing Information Center, (1995)
7. Benjamin Taub "Data Warehouse Return on Investment: Myths and Mistakes", DM Direct Newsletter, (1999), p. 3 – 4
8. Jonathan Wu "Business Intelligence: Calculating the ROI for Business Intelligence Projects", DMReview.com, (July 13, 2000)
9. http://www.mlb.ilstu.edu/ressubj/subject/business/siccode.htm
10. http://www.dmreview.com/article_sub.cfm?articleId=5865

9 APPENDICES

APPENDIX I

SUPPLIER CLASSIFICATION FOR INDIRECT MATERIALS New classifications are highlighted in bold							
Level 1	**Level 2**	**Level 3**	**Level 4**	**Subclass Code**			
Indirect				2000			
	Logistics				2100		
		Fleet				2110	
			Fleet Other				2110
			Aircraft				2111
			Leased Vehicles				2112
			Trucks				2113
			Contract Labor (Drivers)				2114
		Freight				2120	
			Freight Other				2120
			Air				2121
			Ocean				2122
			Intermodal				2123
			Rail				2124
			Common Carriers				2125
			Dedicated Freight				2126
			Small Package				2127
			Land Transportation Europe				2128
		Third Party Logistics				2130	
			Third Party Logistics Other				2130
			Domestic Logistics				2131
			International Logistics				2132
		Warehousing				2140	
			Warehousing Other				2140
		Custom Brokers				2150	
			Custom Brokers Other				2150
		Fuel				2160	
			Fuel Other				2160
	Operating Supplies and Services				2200		
		Forklift Trucks				2210	
			Forklift Trucks Other				2210
			Forklift Maintenance				2211
			Owned Forklifts				2212
			Leased Forklifts				2213
		Maintenance Services - Process Equipment				2220	
			Maintenance Services Other				2220
			Maintenance Services - Electrical				2221

			Maintenance Services - Mechanical				2222
		MRO				2230	
			MRO Other				2230
			Belts				2231
			Electrical/Motors				2232
			Pipes/Valves/Fittings				2234
			Pumps				2235
			Safety Supplies				2236
			Cutting Tools				2237
			Bearings and Power Transmission				2238
			Lubricants/Gases/Oils				2239
			General Mill Supplies				223A
			Fasteners				223B
			OEM Parts				223C
			Industrial Chemicals				223D
			Fluid Power /Hydraulic - Pneumatic				223E
			Repair/Service				223F
			Other				223G
			Blade				223H
			Material Handling				223I
			Cutting Welding Supplies				223J
			Machine Shop				223K
			Plastics				223L
			Hardware Stores				223M
			Auto Supply				223N
			Non Production Steel				223O
			Uniforms				223Q
		Cleaning Supplies and Toiletries				2240	
			Cleaning Supplies and Toiletries Other				2240
		Fasteners				2250	
			Fasteners Other				2250
		Felts				2260	
			Felts Other				2260
	SGA				2300		
		Administrative Services				2310	
			Administrative Services Other				2310
			Dues and Membership				2311
			Mail/Dup/Repro				2312
			Postage				2313
			Temporary Labor				2314
		Benefits				2320	
			Benefits Other				2320
			Health				2321
			Life Insurance				2322
			Pension				2323

			Employee Programs				2324
			Benefits Administrator				2325
			Elective Benefits				2326
			Officer Life Insurance				2327
		Facilities Management				2330	
			Facilities Management Other				2330
			Real Estate				2331
			Janitorial				2332
			Security				2333
			Food Service				2334
			Maintenance				2335
			Waste Removal				2336
			Non-Hazardous Waste Removal				2337
			Waste Water Removal				2338
			Roof (Repair and Replace)				2339
			Dock Systems				233A
			Parking and Paving				233B
			Floor and Foundation				233C
			Fire Protection				233D
			Contract Painting				233E
			Equipment Relocation				233F
			Lawn and Snow				233G
			Pest Management				233H
			HVAC				233I
			Chiller				233J
			Boiler Systems				233K
			Lighting				233L
			Compressed Air Systems				233M
			General Construction				233N
			Plumbing				233O
			Electrical				233P
			Environmental Services				233Q
		SGA IT				2340	
			SGA IT Other				2340
			Software				2341
			Hardware				2342
			Leased Hardware				2343
			Outsourced IT Services				2344
			IT Maintenance				2345
		Office Supplies and Equipment				2350	
			Office Supplies and Equipment Other				2350
			Office Supplies				2351
			Office Equipment				2352
			Office Equipment Maintenance Other				2353
			Office Furniture				2354

		Professional Services				2360	
			Professional Services Other				2360
			Insurance				2361
			Legal/Management Consulting Services				2362
			Sales and Marketing				2363
			Training				2364
			Recruiting				2365
			Environmental				2366
			Consulting Services - IT				2367
			Advertising				2368
			New Employee Screening				2369
			Relocation				236A
			Patent Attorneys				236B
			Consulting Services - Strategy				236C
			Consulting Services - Operations				236D
			Consulting Services - Financial				236E
			Consulting Services - HR				236F
			Consulting Services - Other				236G
		Telecom				2370	
			Telecom Other				2370
			Cell Phone Hardware				2371
			Cell Phone Fees				2372
		Travel and Entertainment				2380	
			Travel and Entertainment Other				2380
			Travel Agency				2381
			Employee Reimbursements				2382
		Employee Recognition				2390	
			Employee Recognition Other				2390
	Utilities				2400		
		Energy				2410	
			Energy Other				2410
			Electricity				2411
			Natural Gas				2412
			Coal				2413
			Steam				2414
			Propane				2415
			Water				2416
	Services				2500		
		Print Design				2510	
			Print Design Other				2510

APPENDIX II

CORPORATE COMMODITIES AND CORRESPONDING DIVISION COMMODITY MAPPING							
Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5** **Corporate Comm. Code**	**PO Origin**	**Div. Comm. Code**	**Division Comm. Code Description**
Direct (1000)							
	Fiber (1100)						
		Finished Paper (1110)					
			Paper Board (1112)				
				Paper Board (1121A)	IPDBAAN	CAE	Paperboard
					IPDBAAN	CAG	Specialty Paperboard
					IPDBAAN	CAM	Unbleached Paperboard
					IPDBAAN	CAO	Paperboard - Used
					SPSBAAN	1176	PMP: Paperboard
		Wood and Wood Shavings (1140)					
			Dry Lumber (1141)				
				Dry Lumber (1141A)	IPD	LU	Lumber/Wood
					BAKER	1143	Southern Yellow Pine
					BAKER	1144	White Pine
					BAKER	1146	Russian Plywood
					BAKER	1147	Plywood
					BAKER	1148	Misc. Plywood Species
					BAKER	1149	Staves
					SPSBAAN	9919	SPS Lumber and Boards
					SPSBAAN	9926	SPS Reclaimed Lumber & Boards

APPENDIX III

HIGH DOLLAR SPEND ANALYSIS (2003) Top 30 suppliers listed in the decreasing order of their annual spend			
SUPPLIER ID	**NAME**	**AMOUNT ($)**	**NO. OF TRANSACTIONS**
NPC0000119	Supplier 1	1,659,039.00	434
NPC0000011	Supplier 2	612,800.00	948
NPC0019419	Supplier 3	556,370.00	1,578
NPC0000047	Supplier 4	504,114.00	3,822
NPC0011396	Supplier 5	431,779.00	148
NPC0000065	Supplier 6	409,704.00	2,310
NPC0005987	Supplier 7	407,647.00	2,742
NPC0017032	Supplier 8	402,740.00	136
NPC0011891	Supplier 9	365,969.00	793
NPC0012168	Supplier 10	322,412.00	828
NPC0022292	Supplier 11	301,887.00	1,687
NPC0000010	Supplier 12	283,160.00	585
NPC0004828	Supplier 13	279,352.00	250
NPC0018755	Supplier 14	271,780.00	184
NPC0012109	Supplier 15	271,436.00	149
NPC0015959	Supplier 16	267,809.00	90
NPC0014602	Supplier 17	257,283.00	100
NPC0021771	Supplier 18	254,095.00	49
NPC0018698	Supplier 19	250,936.00	48
NPC0011671	Supplier 20	235,783.00	66
NPC0014475	Supplier 21	201,485.00	928
NPC0012883	Supplier 22	200,711.00	322
NPC0007300	Supplier 23	200,001.00	158
NPC0011464	Supplier 24	190,294.00	419
NPC0003042	Supplier 25	183,473.00	85
NPC0000908	Supplier 26	180,550.00	28
NPC0010578	Supplier 27	178,949.00	180
NPC0000262	Supplier 28	165,946.00	216
NPC0011090	Supplier 29	163,191.00	384
NPC0000360	Supplier 30	161,708.00	96

APPENDIX IV

HIGH FREQUENCY SPEND ANALYSIS (2003) Top 30 suppliers listed in the decreasing order of their annual number of transactions			
SUPPLIER ID	**NAME**	**AMOUNT ($)**	**NO. OF TRANSACTIONS**
NPC0000047	Supplier 1	504,114.00	3,822
NPC0005987	Supplier 2	407,647.00	2,742
NPC0000065	Supplier 3	409,704.00	2,310
NPC0022292	Supplier 4	301,887.00	1,687
NPC0019419	Supplier 5	556,370.00	1,578
NPC0000056	Supplier 6	156,598.00	996
NPC0006042	Supplier 7	108,454.00	996
NPC0000011	Supplier 8	612,800.00	948
NPC0014475	Supplier 9	201,485.00	928
NPC0010702	Supplier 10	76,672.00	916
NPC0012168	Supplier 11	322,412.00	828
NPC0011891	Supplier 12	365,969.00	793
NPC0012020	Supplier 13	28,066.00	751
NPC0011285	Supplier 14	41,850.00	742
NPC0010060	Supplier 15	119,293.00	586
NPC0000010	Supplier 16	283,160.00	585
NPC0018674	Supplier 17	99,280.00	542
NPC0010098	Supplier 18	97,771.00	450
NPC0010136	Supplier 19	155,085.00	437
NPC0000119	Supplier 20	1,659,039.00	434
NPC0011933	Supplier 21	42,728.00	428
NPC0010795	Supplier 22	95,941.00	421
NPC0011464	Supplier 23	190,294.00	419
NPC0000245	Supplier 24	66,694.00	419
NPC0000446	Supplier 25	96,589.00	400
NPC0001062	Supplier 26	9,327.00	400
NPC0016744	Supplier 27	152,794.00	393
NPC0016142	Supplier 28	90,026.00	393
NPC0011090	Supplier 29	163,191.00	384
NPC0000242	Supplier 30	16,313.00	381

APPENDIX V

HIGH AVERAGE SPEND PER TRANSACTION ANALYSIS (2003) Recommended suppliers are highlighted in bold				
SUPPLIER ID	**NAME**	**AMOUNT ($)**	**NO. OF TRANSACTIONS**	**AVERAGE SPEND/ TRANSACTION**
NPC0016890	**Supplier 1**	110,186.00	15	7345.74
NPC0010527	Supplier 2	29,312.00	4	7328.00
NPC0022486	**Supplier 3**	158,653.00	22	7211.49
NPC0020162	**Supplier 4**	108,155.00	15	7210.35
NPC0017347	**Supplier 5**	150,375.00	21	7160.71
NPC0012721	Supplier 6	13,210.00	2	6605.23
NPC0015558	Supplier 7	26,395.00	4	6598.80
NPC0017486	Supplier 8	19,740.00	3	6580.00
NPC0000908	**Supplier 9**	180,550.00	28	6448.23
NPC0021216	Supplier 10	6,417.00	1	6416.90
NPC0022065	Supplier 11	31,702.00	5	6340.45
NPC0017772	Supplier 12	25,337.00	4	6334.14
NPC0021224	Supplier 13	6,050.00	1	6050.00
NPC0024654	Supplier 14	5,921.00	1	5920.53
NPC0021241	Supplier 15	5,713.00	1	5712.65
NPC0025118	Supplier 16	5,710.00	1	5710.00
NPC0024063	Supplier 17	5,700.00	1	5700.00
NPC0003598	Supplier 18	11,291.00	2	5645.66
NPC0022740	Supplier 19	5,631.00	1	5630.70
NPC0013540	**Supplier 20**	135,091.00	24	5628.79
NPC0014228	**Supplier 21**	99,534.00	18	5529.67
NPC0024009	Supplier 22	38,055.00	7	5436.38
NPC0024618	Supplier 23	5,254.00	1	5254.34
NPC0018698	**Supplier 24**	250,936.00	48	5227.84
NPC0023179	Supplier 25	5,200.00	1	5200.00
NPC0020523	Supplier 26	5,198.00	1	5197.50
NPC0021771	**Supplier 27**	254,095.00	49	5185.62
NPC0022582	Supplier 28	5,056.00	1	5056.25
NPC0023119	Supplier 29	5,040.00	1	5040.02
NPC0007301	**Supplier 30**	87,742.00	18	4874.58

APPENDIX VI

Methodology:

- Annual spend report for Non Ariba PCard suppliers was sorted in decreasing order of spend.
- Total Non Ariba PCard spend for 2003 was found to be 53. 55 M
- Total Non Ariba PCard Suppliers = 13629
- % Spend was calculated using the formula
  = (Amount/Total Non Ariba PCard Spend)*100
- Cumulative % spend was calculated.
- % Number of suppliers was calculated using the formula
  = (1/13629)*100
- Cumulative % of number of suppliers was calculated

Note: Due to space constraint only the first twenty and the last ten suppliers from the 80%-20% analysis have been included in the table below:

80% - 20% SPEND ANALYSIS (2003)								
S.No	**Supplier**	**Name**	**Amount**	**Frequency**	**% Spend**	**Cumulative % Spend**	**% No Of Suppliers**	**Cumulative % (%No of Supp)**
1	NPC0000119	Supplier 1	1659039.00	434	3.0978	3.0978	0.0073	0.0073
2	NPC0000011	Supplier 2	612800.00	948	1.1442	4.2420	0.0073	0.0147
3	NPC0019419	Supplier 3	556370.00	1578	1.0389	5.2809	0.0073	0.0220
4	NPC0000047	Supplier 4	504114.00	3822	0.9413	6.2221	0.0073	0.0293
5	NPC0011396	Supplier 5	431779.00	148	0.8062	7.0284	0.0073	0.0367
6	NPC0000065	Supplier 6	409704.00	2310	0.7650	7.7934	0.0073	0.0440
7	NPC0005987	Supplier 7	407647.00	2742	0.7612	8.5545	0.0073	0.0514
8	NPC0017032	Supplier 8	402740.00	136	0.7520	9.3065	0.0073	0.0587
9	NPC0011891	Supplier 9	365969.00	793	0.6833	9.9899	0.0073	0.0660
10	NPC0012168	Supplier 10	322412.00	828	0.6020	10.5919	0.0073	0.0734
11	NPC0022292	Supplier 11	301887.00	1687	0.5637	11.1556	0.0073	0.0807
12	NPC0000010	Supplier 12	283160.00	585	0.5287	11.6843	0.0073	0.0880
13	NPC0004828	Supplier 13	279352.00	250	0.5216	12.2059	0.0073	0.0954
14	NPC0018755	Supplier 14	271780.00	184	0.5075	12.7134	0.0073	0.1027
15	NPC0012109	Supplier 15	271436.00	149	0.5068	13.2202	0.0073	0.1101
16	NPC0015959	Supplier 16	267809.00	90	0.5001	13.7203	0.0073	0.1174
17	NPC0014602	Supplier 17	257283.00	100	0.4804	14.2007	0.0073	0.1247
18	NPC0021771	Supplier 18	254095.00	49	0.4744	14.6751	0.0073	0.1321
19	NPC0018698	Supplier 19	250936.00	48	0.4686	15.1437	0.0073	0.1394
20	NPC0011671	Supplier 20	235783.00	66	0.44	15.58	0.0073	0.1467
1452	NPC0007973	Supplier 21	6563.00	8	0.0123	79.8928	0.0073	10.6538
1453	NPC0022012	Supplier 22	6558.00	9	0.0122	79.9051	0.0073	10.6611
1454	NPC0017382	Supplier 23	6543.00	9	0.0122	79.9173	0.0073	10.6684
1455	NPC0016252	Supplier 24	6537.00	7	0.0122	79.9295	0.0073	10.6758
1456	NPC0013292	Supplier 25	6530.00	160	0.0122	79.9417	0.0073	10.6831
1457	NPC0011972	Supplier 26	6530.00	39	0.0122	79.9539	0.0073	10.6904
1458	NPC0018555	Supplier 27	6530.00	10	0.0122	79.9661	0.0073	10.6978
1459	NPC0004638	Supplier 28	6521.00	8	0.0122	79.9783	0.0073	10.7051
1460	NPC0009689	Supplier 29	6520.00	181	0.0122	79.9904	0.0073	10.7125
1461	NPC0008482	Supplier 30	6511.00	8	0.0122	**80.0026**	0.0073	**10.7198**