DESIGNING A MULTIDIMENSIONAL DATA WAREHOUSE FOR PROCUREMENT PROCESSES ANALYSIS USING BUSINESS DIMENSIONAL LIFECYCLE METHOD (CASE STUDY ON PT. ABC)

Ari Yanuar Ridwan

Industrial Engineeing, School of Industrial and System Engineering, Telkom University
ariyanuar@telkomuniversity.ac.id

ABSTRACT

Procurement Processes Analysis is needed to make significant cost savings opportunities those including by reducing the number of suppliers, negotiating agreements with preferred suppliers, supplier performance analysis, etc. Data warehosue is the foundation for business intelligence system that can provide PT. ABC the visibility and insight needed to improve procurement performance to further streamline the analytical needs of operational data so that it can be used as a support tool for decision making. Dimensional modeling is a design technique for databases intended to support end-user queries in a data warehouse. It is oriented around understandability and performance.

Keywords*: Dimensional Model, Data Warehouse, Procurement Analysis.*

1. INTRODUCTION

Procurement is a critical business activity. Effective procurement of products at the right price for resale is obviously important to retailers and distributors. Procurement also has strong bottom line implications for PT. ABC that buys products from various suppliers for manufacturing and trading purposes. PT. ABC is Bekasi-Indonesia based compay engaged in manufacturing, trading and distribution of various products of automotive spare parts. Procurement Processes Analysis is needed to make significant cost savings opportunities those including by reducing the number of suppliers, negotiating agreements with preferred suppliers, etc. In PT. ABC, Procurement involves a wide range of activities from issuing purchase requisitions to making payment to vendors.

To support the data analysis needs of the procurement derpartement and of senior management, PT. ABC takes transactional data from the purchasing database, and non-transactional information such as sale quote from spreadsheets, and consolidates this information into the database. However, the current operational relational database reporting presents the following challenges: (1) Reports are static. Users have no way to interactively explore the data in the reports to obtain more detailed information, such as they could do with a Microsoft Office Excel pivot table. Although the existing set of predefined reports is sufficient for many users, more advanced users need direct query access to the database for interactive queries and specialized reports. However, because of the complexity of the purchasing database, too much time is needed for such users to master how to create effective queries. (2) Query performance is widely variable. For example, some queries return results very quickly, in only a few seconds, while other queries take several minutes to return. (3) Complex calculation logic is buried in report definitions and is difficult to share between reports. Because this business logic is generated separately for each report, summary information sometimes is different between reports. Therefore, management has limited confidence in the operational reports.(4) Users in different business units are interested in different views of the data. Each group is distracted and confused by data elements that are irrelevant to them.

Data warehosue can provide PT. ABC the visibility and insight needed to improve procurement performance to further streamline the analytical needs of operational data so that it can be used as a support tool for decision making. Therefore,

PT. ABC needs an analysis and design of the data warehouse as the foundation for business intelligence system to be able to meet those needs.

2. THEORETICAL BACKGROUND

Data warehouse definition from Bill Inmon (2005) is a subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decision. Data warehouse definition from Ralph Kimball (2004) is a system that extracts, cleans, conforms, and deliver source data into a dimensional data store and then support and implements querying and analysis for the purpose of decision making. Both of them aggree that a data warehouse integrates data form various operational source systems. In Inmon's approach, the data warehouse is physically implemented as a normalized data store. In Kimball's approach, the data warehouse is physically implemented in a dimensional data store. a dimensional data store is better format to store data in the warehouse of querying and analyzing the data, compared to a normalized data store. A Normalized data store is a better format to integrate data from various source system (Rainardi, 2008).

The main difference between the approach of Kimball et al. and that of Inmon is that Kimball's conformed dimensions are de-normalized, whereas Inmon uses a highly normalized central database model. Inmon's data marts stores a second copy of the data from the centralized data warehouse tables, whereas the dimensions of Kimball used in the data marts, are not copies of the conformed dimensions, but the dimension table themselves. Kimball et al refers to the set of conformed dimensions as the data warehouse bus. There is no right or wrong between these two ideas, as they represent different data warehousing philosophies. In reality, the data warehouse in most enterprises is closer to Ralph Kimball's idea. This is because most data warehouses started out as a departmental effort, and hence they originated as a data mart. Only when more data marts are built later do they evolve into a data warehouse.

The Lifecycle diagram depicts the sequence of high level tasks required for effective Data Warehouse design, development, and deployment. Several approaches have been developed that attempt to describe the data warehose lifecycle (Kimball et. Al., 2008). Kimball et al (2008) describe kimball dimensional lifecycle method that has a seven stage approach: project planning, business requirement definition, dimensional modelling, physical design, data staging design an developemnt, deployment, and maintenance and grawth. All approaches to the data warehouse lifecylce follow the typical software development lifecycle where the requirement are firtsly specified, followed by various design levels, implementation, testing, and maintenance.

.

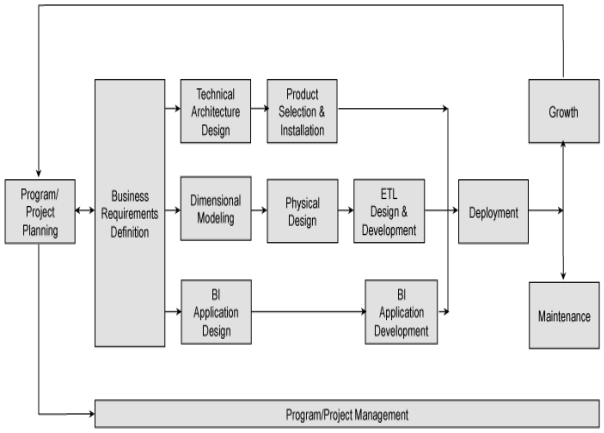

Figure 1. Business Dimensional Lifecylce (Kimball et.al, 2008)

3. RESULT AND DISCUSSION

3.1. Business Requirement Definition

In general, requirements are partitioned into functional requirements and non-functional requirements. Functional requirements are associated with specific functions, tasks or behaviors the system must support, it can be in any format but has to be in line with the business requirements. At this stage a questionnaire designed and sent to the some of the managers and users in each department of the company. Some interviews are held to find more about what they do, how they do it and in what way the project can be of help to them. Table 1. Shows the functional requirements and Table 2. Shows the non-functional requirements for procurement prosesses

analysis. As this is an academic project, the requirements are to be lists out but they did not intend to be achieved all due to time and resources required. Some of the data used in this project are secondary data as some of the analysis has been carried out and going through all the process again would be too big for this project in the context of academic.

Table 1. Functional Requirements

Functional Requirement	Priority
1. The business users need to be able to analyze what materials has the procurement organization purchased this period. (FR-1)	High
2. The business users need to be able to analyze suppliers performance. (FR-02)	High
3. The business users need to be able to analyze How many items on material list. (FR-03)	High
4. The business users need to be able to analyze How many transactions are performed for various stages in the procurement cycle. (FR-04)	High
5. The business users need to be able to analyze How long does it take to move from one stage to the next in the procurement cycle. (FR-05)	

Table 2. Non-Functional Requirements

Non-Functional Requirement	Priority
1. The company preference is to use Microsoft SQL Server to build the data warehouse from end to end, including the ETL tool, reporting, and OLAP. Specific front-end BI applications based on Microsoft.NET framework. The data warehouse should be upgradable to SQL Server future versions. (NFR-01)	High
2. The data warehouse needs to be flexible so we can enhance it easily and adapt to changes that happen in the transaction systems. In particular, it needs to include enhancements such as bringing new pieces of data into the warehouse, adding new reports/cubes/data quality rules, or modifying existing ones. (NFR-01)	HIgh

3.2. Technical Architecture Design

Where the business requirements answer the question "What do we need to do?" the architecture answers the question "How will we do it?" The technical architecture is the overall plan for what you want the data warehouse system to be when it's ready for serious use. It describes the flow of data from the source systems to the decision makers and the transformations and data stores that data goes through along the way. It also specifies the tools, techniques, utilities, and platforms needed to make that flow happen (Kimball et.al, 2008). Figure 2 provides proposed system architecture for the production environment

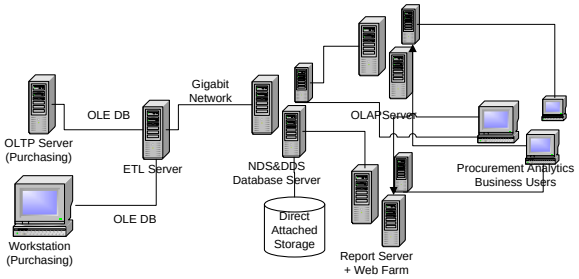

Figure 2. Proposed System Architecture For Data Warehouse

the system architecture consists of four servers: ETL server, database server, OLAP Server, and Report Server. The source system is an Procurement system that records data in the purchasing departement. Table 3. shows the product selection and installation of system architecture for data warehouse. The company preference is to use Microsoft SQL Server to build the data warehouse from end to end, including the ETL tool, reporting, and OLAP. Specific front-end BI applications based on Microsoft.NET framework. The data warehouse should be upgradable to SQL Server future versions (NFR-01).

Table 3. System Architecture fro the production environment for PT. ABC

Components	Spesification
Data Source	A custom-developed.NET Procurement Application, Win Server 2003 R2 EE, SQL server 2008 R2
ETL Server	SQL server 2008 R2 Integration Services, Win Server 2003 R2 EE, 8 Processor 16 GB RAM
OLAP Server	SQL Server 2008 R2 Analysis Services, Win Server 2003 R2 EE, 8 Processor 16GB RAM
Report Server	SQL Server 2008 R2 Reporting Services, Win 2003 EE, 2 Processor 4 GB RAM.
Storage Area Network (SAN)	6TB Raw Capacity
Client	A custom-developed.NET Application for procurement processes analysis, Win XP, Win 7, Win 8.

PT. ABC takes transactional data from the purchasing database, and non-transactional information from spreadsheets, and consolidates this information into the

database. Table 4. Shows several tables from purchasing database as data sources for data warehouse system.

Table 4. Source Tables (Purchasing OLTP database)

Table	Field (Partially)
Product	Product Key (PK), SKU Number, Product Description,
Vendor	Vendor Key (PK), Vendor Name, Vendor Address, Vendor Status
PurchaseOrderHeader	PurchaseOrderID (PK), EmployeeID, VendorID,
PurchaseOrderDetail	PurchaseOrderID (PK), PurchaseOrderDetailID (PK), OrderQty, UnitPrice
Employee	BusinessEntityID, NationalIDNumber, LoginID, OrganizationNode, OrganizationLevel
Warehouse	Warehouse Key (PK), Warehouse Number, Warehouse Name, Warehouse Name, Warehouse Address, Warehouse Description, Warehouse zone, Warehouse area
PurchaseRequistion Header	PurchaseReqID (PK), EmployeeID (FK), VendorID (FK), ShipMethodID,
RFQHeader	RFQId (PK), EmployeeID, VendorID, hipMethodID, OrderDate,
GoodsReceiptHeader	Goods Receipt time Key (FK), Requested By time Key (FK), Product Key (FK), Vendor Key (FK), Warehouse Key (FK), shipment Key (FK), Employee Ordered By key (FK), Employee Received by Key (FK),
VendorInvoiceHeader	Vendor Invoice time Key (FK), Product Key (FK), Vendor Key (FK), Warehouse Key (FK), Contract Terms Key (FK),
VendorPaymentHeader	Vendor Payment ID (PK), Product Key (FK), Vendor Key (FK), Contract Terms Key (FK),
Contract Terms	Contract Terms Key (PK), Contract Terms Description, Contract Terms Type
shipment	shipment Key (PK), shipment Name, shipment Type, shipment size, shipment Location

3.3. Dimensional Modelling

Dimensional modeling is a logical design technique for structuring data so that it's intuitive to business users and delivers fast query performance. Dimensional models stored in a relational database platform are typically referred to as star schemas; dimensional models stored in multidimensional online analytical processing (OLAP) structures are called cubes (Kimball et.al, 2008). A conceptual data model was prepared which laid out a high level structure of the entities and the relationships between them after identifying separate business needs. Facts and dimensions were identied. Grain of the tables was determined. Conformed dimensions are identied using the business matrix. A data mart may have multiple star schemas but dimensions can be shared between different fact tables or business processes.

The business matrix Figure is tabular representation of the business process versus the dimension tables. Table 5 shows the business matrix for procurement (procure-to-pay) business process.

Table 5. Business Matrix

	Business Processes (Procure-To-Pay)	Dimension						
		Vendor	Product	Time	Employee	Warehouse	Shipment	Contract/Outline Agreement
1	Purchase Requisitions	x	x	x	x			
2	Request For Quotation		x	x				x
3	Vendor Selection	x	x	x				x
4	Purchase Orders	x	x	x	x	x	x	x
5	Shipping Notifications	x	x	x	x	x	x	
6	Goods Receipt	x	x	x	x	x	x	
7	Vendor Invoices	x	x	x	x	x		x
8	Vendor Payments	x	x	x		x		x

Table 6. Granularity of Fact Tables

	Fact Tables	Atomic Granularity	Metrics
1	Purchase_Requisitions_Fact	1 row per requisition	PReq. Quantity, PReq. Amount
2	RFQ_Fact	1 row per RFQ	RFQ Quantity, RFQ Amount
3	Vendor_Performance_Fact	1 row per Quotation	Ordered Quantity, Ordered Amount
4	Purchase_Order_Fact	1 row per PO	PO Quantity, PO Amount
5	Shipping_Notifications_Fact	1 row per Shippping notice	Shipped Quantity
6	Goods_Receipt_Fact	1 row per GR	GR Quantity
7	Vendor_Invoices_Fact	1 row per IV	Invoice Quantity, Inv. Amount
8	Vendor_Payments_Fact	1 row per Payment	Payment Amount, Discount Amount, tax Amount

Table 7. Fact Tables

	Fact Tables	Fact Tables Fields (Partial)	Dimension Table
1	Purchase_Requisitions_Fact	Purchase Requisition time Key (FK), Product Key (FK), Vendor Key (FK), Contract Terms Key (FK), Employee Requested By Key (FK), Contract Number (DD), Purchase Requsition Number (DD), Purchase Requisition Quantity, Purchase Requisition Amount	Product, Contract Terms, Warehouse, time,Vendor, Employee
2	Vendor_Performance_Fact	Vendor Performance time Key (FK), Product Key (FK), Vendor Key (FK), Contract Terms Key (FK), Ordered Quantity, Returns Quantity, Ordered Value, Return Value, Reject Quantity, Reject Value, Total Spend, Stock Outage, Average Lead Time	Product, Contract Terms, time,Vendor, shipment
3	Purchase_Order_Fact	Purchase Order time Key (FK), Requested By time Key (FK), Product Key (FK), Vendor Key (FK), Contract Term Key (FK), Warehouse key (FK), shipment Key (FK), Employee Ordered By Key (FK), Employee Purchase Agent Key (FK), Contract Number (DD), Purchase Requisition Number (DD), Purchase Order Number (DD), Purchase Order Quantity , Purchase Order Amount	Product, Contract Terms, Warehouse,time,Vendor, Employee, shipment
4	Shipping_Notifications_Fact	Shipment Notification time Key (FK), Requested By time Key (FK), Product Key (FK), Vendor Key (FK), Contract Term Key (FK), Warehouse key (FK), shipment Key (FK), Employee Ordered By Key (FK), Employee Purchase Agent Key (FK), Contract Number (DD), Purchase Requisition Number (DD), Purchase Order Number (DD), Purchase Order Quantity , Purchase Order Amount	Product , Warehouse, time, Vendor, Employee, shipment
5	Goods_Receipt_Fact	Goods Receipt time Key (FK), Requested By time Key (FK), Product Key (FK), Vendor Key (FK), Warehouse Key (FK), shipment Key (FK), Employee Ordered By key (FK), Employee Received by Key (FK), Purchase Order Number (DD), Shipping Notification Number (DD), Warehouse Receipt Number (DD), Received Quantity	Product ,Warehouse, time,Vendor, Employee,shipment
6	Vendor_Invoices_Fact	Vendor Invoice time Key (FK), Product Key (FK), Vendor Key (FK), Warehouse Key (FK), Contract Terms Key (FK), Contract Number (DD), Payment Check Number (DD), Vendor Invoce Amount, Vendor Discount Amount , Vendor Net Payment Amount , Vendor Tax amount	Product , Contract Terms, Warehouse, time, Vendor, Employee, shipment
7	Vendor_Payments_Fact	Vendor Payment time Key (FK), Product Key (FK), Vendor Key (FK), Warehouse Key (FK), Contract Terms Key (FK), Contract Number (DD), Payment Check Number (DD), Vendor Invoce Amount, Vendor Discount Amount , Vendor Net Payment Amount , Vendor Tax amount	Product , Contract Terms, Warehouse, time, Vendor, Contract

Table 8. Dimension Tables

Dimension Table	Field (Partial)
Product	Product Key (PK), SKU Number, Product Description, Brand Name, Category Name, Make Flag, Safety stock level, Reorder Point, Standar Cost, List Price, Size, Days to Manufacture, Product Line, Dealer Price , Model Name, Status, Start Date, End Date
Contract Terms	Contract Terms Key (PK), Contract Terms Description, Contract Terms Type
Warehouse	Warehouse Key (PK), Warehouse Number, Warehouse Name, Warehouse Name, Warehouse Address, Warehouse Description, Warehouse zone, Warehouse area
Vendor	Vendor Key (PK), Vendor Name, Vendor Address, Vendor Status
Employee	Employee Key (PK), Employee Name, Emplooye Position. Employee Departement, Employee Type
shipment	shipment Key (PK), shipment Name, shipment Type, shipment size, shipment Location

3.4. ETL Design And Development

ETL (Extract, Transform and Load) is the process of retrieving and transforming data from the source system and putting it into the data warehouse. architecture of the data warehouse system designed consists of four servers: ETL server, database server, OLAP Server, and Report Server. The source system is an Procurement system that records data in the purchasing department. System is designed to execute the ETL processes in a separate ETL server that sits between the source system and the data warehouse server. This approach provides

the highest performance. The ETL runs on its own server, so it does not use the resources of the data warehouse server or the source system server at all. Logical ETL Metadata uses to describe source and target system. Table 9 Lists product table metadata needed to get data into a data-staging area and prepare it for loading into data marts.

Table 9. ETL Metadata (ex. Product)

Database	: Purchasing						
Table	: Product						
Description	: Product sold or sued in the manufacturing of sold products						
SOURCE TARGET			DATA SOURCE			LOAD	
COLUMN	TYPE/LENGTH	DESCRIPTION	TABLE	FIELD NAME	TYPE/LENTH	STATUS	
ProductID	int	Primary key for product records	DIM_Product	ProductID	int	COPY	
SKUNumber	int	SKU Number	DIM_Product	SKUNumber	int	COPY	
ProductDesc	nvarchar(50)	Description of the product	DIM_Product	ProductDesc	nvarchar(50)	COPY	
BrandName	nvarchar(25)	Brand of the Product	DIM_Product	BrandName	nvarchar(25)	COPY	
CatName		Category of The Product					
MakeFlag	Flag(bit)	0 = product is purchased, 1 = Product is manufacture in-house	DIM_Product	MakeFlag	Flag(bit)	COPY	
FinishedGoodsFlag	Flag(bit)	0 = Product is Not a salable item. 1 = Product is salabel	DIM_Product	FinishedGoodsFlag	Flag(bit)	COPY	
SafetyStockLevel	smallint	Minimum inventory quantity	DIM_Product	SafetyStockLevel	smallint	COPY	
ReorderPoint	smallint	Inventory level that triggers a purchase order or work order	DIM_Product	ReorderPoint	smallint	COPY	
StandardCost	money	standars cost of the product	DIM_Product	StandardCost	money	COPY	
ListPrice	money	Selling price	DIM_Product	ListPrice	money	COPY	
size	nvarchar(5)	product size	DIM_Product	size	nvarchar(5)	COPY	
SizeUnitMeasureCode	nchar(3)	Unit of measure for size column	DIM_Product	SizeUnitMeasureCode	nchar(3)	COPY	
WeightUnitMeasureCode	decimal(8,2)	Unit of measure for weight column	DIM_Product	WeightUnitMeasureCode	decimal(8,2)	COPY	
DaysToManufacture	int	Number of days required to manufacture the product	DIM_Product	DaysToManufacture	int	COPY	

3.5. Front End Application Development

to facilitate access for users is developed based front-end application .net framework. applications developed so has the features to be able to access the data which is attached to the datawarehouse and allow a user to perform an analysis of the procurement process. application also has a number of Key Performance Iindicator (KPIs) dashboard that displays the procurement process as shown in Figure 3.

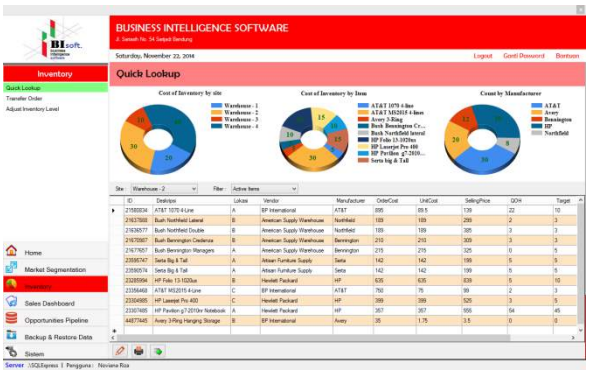

Figure 3. Dashboard for Procurement Analysis

4. CONCLUSION

Procurement Processes Analysis is needed to make significant cost savings opportunities those including by reducing the number of suppliers, negotiating agreements with preferred suppliers, etc

Data warehosue is the foundation for business intelligence system that can provide PT. ABC the visibility and insight needed to improve procurement performance to further streamline the analytical needs of operational data so that it can be used as a support tool for decision making.

5. REFERENCES

(a) Inmon, W.H., 2005, “Building the Data Warehouse”, Wiley
(b) Kimbal, R., Ross, M., Mundy, J., 2008, “The Data Warehouse Lifecycle Toolkit 2nd Edition”, Wiley.
(c) Kimball R., Ross, M., 2010 “The Kimball Group Reader”, Wiley.
(d) Kimball, R., Ross, M., 2013, “The Data Warehouse Toolkit 3rd Edition”, Wiley.
(e) Rainardi, V., 2008, “Building A Data Warehouse with examples in SQL Server”, Apress.
(f) Kimball R., Caserta, J., 2008,“The Kimball Warehouse ETL Toolkit”, Wiley.

AUTHOR BIOGRAPHIES

Ari Yanuar Ridwan is a lecturer in Department of Industrial Engineering, School of Industrial and System Engineering, Telkom University, Bandung. He received his Master of Informatics Engineering in 2005 and Master of Industrial Engineering in 2011 both from Bandung Institute of Technology. Her research interests are in the area of business intelligence, ERP System, IT For Logistics and Supply Chain management. His email address is <ariyanuar@telkomuniversity.ac.id>