Paper 46-25

From the Analysts to the Executive -- Business Visualization and Enterprise Reporting with the SAS® System

Joe Costanzo, SAS Institute Inc.
Ben Zenick, SAS Institute Inc.

Over the past decade, many large organizations have deployed data warehousing technology to furnish their end user community, customers, and partners with the data necessary to improve the way they make business decisions. A key component of a successful enterprise data warehouse strategy is the Enterprise Reporting System (ERS). Most organizations are seeing enterprise level reporting systems as not only necessary but strategically imperative to bring these dysfunctional, department level data stores and reporting applications under control. This paper will explore options for expounding upon existing deployments as well as coordinating new technologies with existing infrastructure.

Enterprise reporting systems (ERS) are systematically designed to meet the data access and information discovery needs of the majority of end users within an organization. These systems should combine the ease of use and simplicity that users have come to demand, while maintaining the ability to create complex multi-variable reports from any data source available to the system. In short, enterprise reporting should connect users to the business by allowing for fast and easy access to daily corporate knowledge.

Several factors have contributed to limiting the design and deployment of reporting systems off of the typical data stores. During the early 1990's, the need to do enhanced reporting increased substantially, creating a massive overload and backlog of reports that most IT shops were not staffed to handle. As a result, many business units adopted new GUI reporting tools and began to implement reporting systems on an application by application basis. Although this approach proved successful in reducing the information access time and report generation time, many business units inadvertently created inconsistencies between their own reporting applications. Key metrics such as revenue, and sales and margin, were defined differently across business units. These inconsistencies created a multitude of problems as organizations began to implement development efforts on enterprise level reporting applications. The fact that different reporting applications used different tools also had a huge impact and intensified the problem. For the typical IT department, the ability to deal with these multiple data stores and applications was hindered by the backlog of reporting requests from the user community.

In order to address these problems, there are numerous initiatives that must be considered when designing any corporate-wide infrastructure. Covering all of these considerations in detail is too broad a scope for one paper. Yet, there are several key areas one must address when designing an enterprise wide reporting system.

- The system must be able to collect and manage the data from any source that meets the business needs of its end users. No matter how complex or large the data is, any solid ERS must be able to access and combine data from a multitude of sources.

- Distribution of data is a fundamental requirement as any enterprise system must be adept at pushing information out to potentially thousands of users that make up the extended enterprise. Thus, these systems must be designed to be scalable and capable of using the latest technology available to them such as Publish & Subscribe, Internet/Web based reporting, and message queuing.

- Because an end user community is typically made up of users with very different skills and reporting needs, any solid reporting infrastructure must simplify the viewing of the data. These systems must be able to select, combine, and compare the most complex of data yet provide a simple means of visualizing the information

The remainder of this paper will focus on these topics, as well as the tools and solutions available in the SAS® System for delivering Enterprise level reporting systems.

Data Collection and Management

Enterprise reporting requires an architecture design capable of integrating many different data sources. Most transactional systems were never designed for decision- support type reporting. They were designed to store day-to-day operational data. Building reporting systems that do not impact our daily transactional applications has been the main push behind all data warehousing activities over the last decade. For reasons mentioned above, these data stores whether marts or warehouses, have yielded a variety of problems when attempting to combine them under an enterprise structure. Thus, any design or architecture must be created as an integral part of the reporting system, not separately from it.

In order to provide an architecture that is capable of delivering the diverse needs of an organization's end- user community, IT must be able to build subject-specific data marts. Ideally, these data marts should be constructed using consistent business dimensions that can be shared across the organization's data structures. Some of the most typical dimensions are Time, Customer, and Product. One can ensure that all users are accessing and defining data in the same way, as well as measuring business the same. However, this may require a change in corporate culture by sharing these key dimensions across an organization's business units and departments. By building a measure or dimension once and enforcing its use across all corporate data stores, IT can reduce implementation times and cost while providing coordinated reporting and better business decisions.

Data access has long been one of the Institute's unique strengths. The SAS System treats enterprise data as a generalized and available resource by providing transparent access to popular database management systems, flat files, system specific host files, other historical or legacy data types, as well as the SAS System's own data structures. These time proven access strategies are grouped under MEA™ or MultiEngine Architecture™:

- MEA stands for Multiple Engine Architecture, the core technology within the SAS System to provide direct and transparent access to external data sources. In the SAS System, all data is accessed via engines, regardless of type. An engine is an internal group of instructions that allow SAS to access, read, and write data stored in a variety of data structures. MEA provides the frameworks for translating SAS read and write requests into the appropriate DBMS or file structure calls. The engine initiates and monitors communication between itself and the DBMS or file.

There are a variety of tools available in the SAS System to manage this data collection process. SAS/Warehouse Administrator™ provides a framework for effective warehouse management through a metadata-driven architecture. It facilitates business subject definition, consolidation of business rules, scheduling of processes for warehouse maintenance, and integration with decision support tools for effective warehouse exploitation and creation of an enterprise-reporting environment.

With the capabilities of the SAS System in regards to MEA, it is imperative that the Software can access these disparate data sources no matter what platform they reside on. It is also important from a time consuming perspective; not to have to learn a different language based off of what platform the data does reside on. This is where SAS' Multi Vendor Architecture™ (MVA™) comes in to play. No matter what platform you are on, the software components act exactly the same. This is one of the advantages of being a 4GL language. How does MVA assist us in the area of data collection? Since the data supported by the SAS System typically does not reside on the same platform, the MVA allows us to read and write data the exact same way on essentially any platform. The Warehouse Administrator provides the framework to easily manage the reading, writing, and processing of the data, from one central point of control. Connecting to these different platforms is accomplished by using taking advantage of middleware. SAS has its own middleware, but the system integrates with 3rd party middleware as well. The usage of this

middleware will allow us to run processes on these different platforms without having to physically be on that machine.

Report and data distribution

Some of the latest technology, mainly the Internet, has unlimited potential for ERS. Any enterprise level system brings with it a tremendous amount of data. One must be careful not to overwhelm the end user with information overload. Using push technology, an ERS can broadcast information and reports only to those users who need the information and can act on it. Through the use of this technology, the system can offer a series of triggers and subscription lists to determine which reports are to be pushed to whom. The simplest of jobs or triggers would be to take quarterly batch reports and push them to a central location for retrieval or viewing. The end-users would receive e-mail notifying them that a report is ready for viewing and point them to a specific location. In this type of system, the ERS administrator should have the ability to set both event and manual triggers. Event reports would be kicked off if a critical success factor, such as sales volume, falls within a predetermined range.

This technology also allows the end-user community to decide for themselves what they want to receive from the ERS. Through this technology, the user will have the ability to subscribe to channels of information containing reports centered on a specific topic, such as product. Basically these channels organize push reports into groupings that users can subscribe to based on their interest or needs. It is still necessary to provide users with a short description of the actual report. That way they can decide for themselves if they want to view the actual report. It is quite possible that different channels will broadcast reports at different frequencies. As mentioned above, some reports may only be produced when a certain event occurs. An ERS should have the capability to provide and alert the users subscribed to particular channel when new reports are available. With this method, a user will always know when the reports and or data has been updated.

During the creation of the Enterprise Data Warehouse that is used to feed the ERS, the processes will be put into the Warehouse Administrator to publish packages to the channels that are specified to a job. These packages can contain information such as; static reports (html, pdf, etc), Job Completion/Failure notification, Exception Reports, as well as data needed by the analysts for things such as data mining, or other types of data discovery. These packages will be published through email, or to an LDAP server URL, based off of specific occurrences of events, defined by the business rules, during a job. This information can be published to individuals, groups, or servers. The servers can include the Enterprise Reporter Web Server (Gallery), Collaborative Business Intelligence (CBI), or a variety of other servers that are used for distribution within the corporation. The Enterprise Reporter Web Gallery can be used to publish reports over the web in either html or pdf format. Archiving of all of this published information can be done as well by using CBI. This will allow users to easily track changes in the data, as well as retrieve any packages that may have been lost.

Business Visualization

As we have mentioned, universal report viewing is essential to a complete ERS. Most organizations have several different user groups, each requiring a different view of the information contained in the enterprise data store. For example, a senior level manager may want to start with a scorecard type of report that contains the organization's critical success factors or CSFs. This manager might find that they want to further explore a specific CSF. They would then be able to drill down and explore the measures that make the CSF in detail. They may decide that they even want to go into the actual transactional level detail. All of this should be accomplished in one straightforward process. So whether the user is a CEO, Regional manager, or business analyst, the reporting system should provide an intuitive interface. It should allow the user to quickly get a "Big Picture" view of summary level information and then allow the user to drill into greater levels of detail, as deemed necessary for their analysis needs.

There can be a significant problem encountered with this type of reporting system. This is information overload or too much data to effectively analyze. This is a common trend that

one must safeguard against when designing any enterprise level system. Too much data can often be as bad as insufficient data. If your user community feels that there is too much information to review, they will be hesitant to use the system. Thus ERS must move from the typical GUI format to a more interactive visual approach. So, what is the difference between a visual interface and a typical GUI?

When users describe a particular “Look and Feel” they are usually referring to the physical appearance of the system and the way they interact with the system, typically via a mouse or some other pointing device. Thus, a GUI tends to make an application look attractive and easier to use. Yet the GUI rarely changes the way the end-user analyzes the data that feeds the system. Business visualization can best be described as a way to provide the user with the ability to take on a more active role in data analysis via a series of highly interactive objects. These objects allow a user to interact directly with the objects on the screen while limiting the use of standard GUI controls such as pull down menus, list boxes and dialog boxes. This makes the on screen object seem almost “live”. The actual object is connected to the supporting data structures and the business rules that mange that data. Thus the user is able to get an immediate response from the object without the need to go to any additional windows. These objects and techniques can increase the user’s interaction with the data, thus allowing for better business decisions. It is important to be able to export their findings back into more traditional reporting formats. The typical user of this technology is an analyst who needs to get the uncovered information into the hands of the management chain that can make business decisions based on the results of the investigation.

There is no "One" solution in the SAS System for Business Visualization. Instead it is through a combination of SAS products that we can provide this technology. Our data mining solution incorporates several components from our visualization tool set to allow our customers to not only analyze large amounts of data, but to also see and present that data visually. SAS/GRAPH® software provides tools for creating and presenting business and scientific graphics in a wide variety of graphical formats. SAS/SPECTRAVIEW® software provides our customers with a tool for visually analyzing large and small data files in a highly interactive 3D environment. With recent advances in the software this technology can be used as a powerful tool to help illuminate patterns and trends in data. SAS/INSIGHT® software provides our customers with a very interactive tool for not only visualizing data, but for doing statistical analysis and number crunching. SAS/GIS® software is a desktop mapping package that offers a variety of mapping tools and map manipulation capabilities which allow users to organize, analyze, and visualize business data in its geographical context. SAS/GIS is a Geographic Information System (GIS) application that provides tools for organizing, analyzing, and displaying data that can be referenced spatially. Many types of data have a spatial aspect-that is they can be tied to a physical location. Here is an example of how we can increase the value of the data by combining several of the SAS Systems tools.

We have information about each store in our company by the actual location and measuring Gross Sales. This is information that can be easily be analyzed via our OLAP technology. We might see that we have a series of stores that are way off their forecasted sales mark. Upon further analysis and drill down we see that these stores are all located in Florida and we can see what the actual cites are by performing additional drill downs. In a pure OLAP environment this would be as far as we could go on our analysis. Why are these stores all in Florida? Unless the analysts is very good at geography they may miss a most obvious piece of information that could only be derived via Business Visualization techniques. In this instance what if we were to combine an OLAP viewer and a GIS object on the same screen? When the analyst drilled down to the actual store level a GIS map could be displayed on the same screen showing the actual Latitude and Longitude, or location, of each store. Upon viewing this analysts may see that all the stores are clustered within a 20-mile radius on the coast of Florida. They may also choose to see a layer that represents any recent weather activity for the same time last year. At this point the analysts might see that there was a hurricane that came through the area being analyzed during that time frame last year. This view on the map causes the analysts to reach through to the detail data in the OLAP viewer and they see that there were zero sales for that time period. By visualizing the data on a map and in an OLAP environment analysts

can easily conclude that the forecast was off because of a natural disaster that caused no sales for the flagged stores.

Currently the Enterprise Reporting market is relatively new with a small number of vendors. These vendors are not necessarily well versed in providing solution that can help an organization navigate their current legacy and complex ERP systems. It is important to partner with a vendor that is able to leverage your investment in your current reporting systems and has the technology to take those systems to the next level. Based on the reasons that this paper has discussed we feel that SAS Institute provides the flexibility and time proven technology to meet that challenge.