Unlocking SAP AG R/3 data with the SAS System!!!

The SAS® Intelligent Warehousing Solution for R/3

SAS Institute Inc.
SUGI 1999

Introduction

Thousands of Organizations have made the strategic decision to replace existing production application with package solutions designed to re-engineer their existing business processes. Actually these applications are designed to do much more, including things like replacing outdated applications with ones that are Year 2000 compliant, cross business process integration, and more. Integrating business processes allows organizations to streamline activities that span across operational functions. Since these applications typically affect many aspects of the enterprise they are commonly referred to as Enterprise Resource Planning or ERP Solutions. The market leader in the ERP Industry is clearly SAP AG's R/3 System, with approximately 40% of the market share. SAP AG, as with other ERP Vendors, such as JD Edwards, Baan, and PeopleSoft, offer increased efficiency in the day-to-day management of their production or operational infrastructure by providing integrated modules covering various functional areas. Such modules may include (but are not limited to) finance, sales and distribution, human resources, materials management and more. One benefit of these integrated solutions is that everything is linked together, so that if a customer places an order the company issues the required invoice. The invoice is not just registered in the financial module, but any other relevant area, like logistics, materials management, and others are updated appropriately.

The decisions to implement these solutions have proven to be successful in areas of integrating business processes for automated and optimal production system management bringing improved operational efficiency, better productivity and increased profitability. The ability to enter a simple transaction and automatically propagate downstream functions to occur without manual intervention is an obvious step in the right direction for reducing cost while increasing productivity. With these successes and benefits in hand, it now becomes a question of how to turn this packaged application and its massive amount of data into useful information for successful decision making? When evaluating this question it does not take long to determine that the existing functionality in the ERP applications fall short in the areas of Business Intelligence, Information Reporting, and Decision Support. Not to mention when looking to third party applications to satisfy this requirement, the complexity of these systems and the lack of robust/flexible tools limit the choices in solutions that can be successfully integrated with them.

The scope of this paper is to illustrate a framework designed to unlock the potential of data trapped in an SAP AG R/3 System and provide a vision that covers an enterprise wide view of an organizations Data Warehousing and Decision Support needs. It will cover various aspects of the SAS Intelligent Warehousing Solution illustrating the ability to turn an abundance of raw data into meaningful information. The delivery of this information now becomes the centerpiece for strategic decision making in an effort to increase customer profitability and maintain or achieve a competitive advantage, allowing your organization to continue to thrive into the new millennium.

SAS Intelligent Warehousing Solution

Think globally, act locally

The SAS Intelligent Warehousing Solution is architected to take on an enterprise wide focus, commonly referred to as "Think Globally, Act Locally". In other words, providing a structured framework for turning data into useful information – regardless of platform, source data, or target repository. This is accomplished by having a strategic vision in providing an enterprise wide view of corporate information (data and metadata) and delivering this information to the business community in the form of a "Dependent" Data Mart structure. Dependent Data Marts take advantage of an enterprise-wide view of the data observing a common set of business definitions or rules. Thus providing the ability for the entire organization to base decisions on the same, reliable information. Dependent Data Marts are much different than their evil sister, the "Independent" or "Stove-Pipe" Data Marts which are not reusable and do not provide a consolidate view of the enterprise. Independent Data Marts do NOT share metadata and business rules throughout the organization, thus introduction more islands of information or the legacy systems of the future. This in itself does not produce a repeatable process and can negatively affect the return on investment and time of implementation. Not to mention the likelihood of inconsistent results being available to the information consumers. Additional information on this topic and proven methodology to accomplish this structure is available in the SAS Institute White Paper: SAS® Rapid Warehousing Methodology, 1998.

The Architecture which supports the SAS Intelligent Warehousing Solution are as follows:

<u>Extraction</u> of data from operational and legacy environments

The SAS Intelligent Warehouse provides robust data access and acquisition technologies to virtually any data source. These include native access engines to the more popular data sources across multiple platforms, including DB2®, Oracle®, Informix®, IMS®, Sybase®, SQL Server® and more. In addition to these popular data stores, the SAS Intelligent Warehouse also provides the ability to integrate with data in ERP environments. In particular, SAS/Access® interface to SAP AG system R/3 enables organizations to access R/3 and integrate it with other source systems, providing a true enterprise view of information. There is no additional programming overhead involved: the use of metadata imported directly from R/3 greatly simplifies the process.

Adding value and enriching data through a <u>transformation</u> engine

The SAS software transformation technology is a mature, platform independent environment that provides facilities for integrating diverse data. This process enriches the data while protecting or ensuring it's integrity by validating that data, summarizing, and applying advanced statistical functions in support of data restructuring for the decision support environment. The SAS software transformation technology also includes a transformation library for reusability and distribution of pre-defined transformations and provides the ability to integrate external transformation programs or applications (COBOL, C, etc.) The transformation process can take up to 75% of the total time spent creating, maintaining and processing the warehouse. This phase of the warehousing process eliminates the risk of "Garbage In, Garbage Out" … your analysis is only as good as the information it is provided.

<u>Loading</u> of the warehouse

Once the data has been transformed, the next phase is to populate the various data warehouses and data marts across the required platforms. The data is not necessarily stored in a single database, but may be distributed to various desired targets and formats. SAS software provides a variety of storage options that are optimized for business intelligence operations; SAS Tables, a SAS/MDDB® (multidimensional database … ROLAP, MOLAP, and HOLAP), and the SAS Scaleable Performance Data Server®. The SAS

Intelligent Warehouse also supports other storage options if these are required (for example, Relational Databases based on corporate standards).

<u>Management</u> of the process

The SAS Intelligent Warehouse is a robust and flexible solution: it supports a broad range of hardware platforms, operating systems, data sources, and data targets. To make the task of managing the data warehouse process as simple as possible, SAS Institute therefore offers a single point of control through the SAS/Warehouse Administrator® software. SAS/Warehouse Administrator uses a common metadata repository to automate the process of extracting, transformation and loading (ETL) the operational data into a distributed data warehouse environment. The integration of this metadata through each of the many phases of the warehouse process is one of the keys to success in any warehouse or mart implementation.

<u>Exploitation</u> of the Warehouse

The final phase of the warehousing process is to deliver the information for exploitation. SAS Institute provides a robust and flexible variety of business intelligence software for mining, analyzing, visualizing and reporting on warehoused information. This includes functional, cross-industry and industry-specific business solutions covering all aspects of an organization's Knowledge Management requirements.

<u>Extracting data from R/3</u>

Intelligent Data Access

As mentioned in the overview section above, a problem faced by both IT and Business professionals when implementing an SAP AG System is the lack of robust and flexible tools that can integrate data in R/3 and other sources with decision support applications. The best approach, or the one with the least impact on the operational systems, is to implement a data warehouse or data mart structure to support this environment. To do this effectively you must answer the question; what are my options to intelligently extract data from this system?

The answer lies with the SAS/Access to R/3 technology. The interface provides an easy-to-use, reliable and intelligent solution to access your SAP data. The R/3 Interface is SAP metadata-aware; meaning the extraction process is based upon metadata definitions that are taken directly from the R/3 System.   This metadata is actually integrated directly with the SAS/Warehouse Administrator to be utilized throughout the entire data warehousing process.

SAS/Access to R/3: Components

SAS/Access to R/3 includes the following components (Figure 1):

- Two template modules; an ABAP/4 Function and an ABAP/4 Report
  The template modules are installed in the SAP R/3 system (application server) at the time SAS/ACCESS to R/3 is installed.   These modules resolve passed parameters at run-time to satisfy the particular data request (table(s), columns, subsetting, etc.).   These modules communicate together using R/3's CPIC protocol, ensuring maximum performance.

- A Remote Function Call (RFC) Server
  A RFC server is installed on the SAP R/3 machine or on another machine (does note require SAS System software). The purpose of the RFC server is to handle user authentication and processing the RFC parameters required by the template modules. This SAS RFC server communicates with the ABAP/4 function module installed in the SAP R/3 system using the RFC protocol (Remote Function Call) provided by SAP.   RFCs are the underlying technology that most of the SAP internal and external data movement and interface facilities or protocols are based on.

- SAS/ACCESS Graphical User Interface (GUI)
  The SAS/ACCESS graphical user interface (GUI) is used to navigate and/or search through the SAP metadata structure (snapshot) to drive the extraction process of the SAP tables.   The GUI component lives on any PC or Unix based platform and communicates with the SAS RFC server using TCP/IP communication protocol.

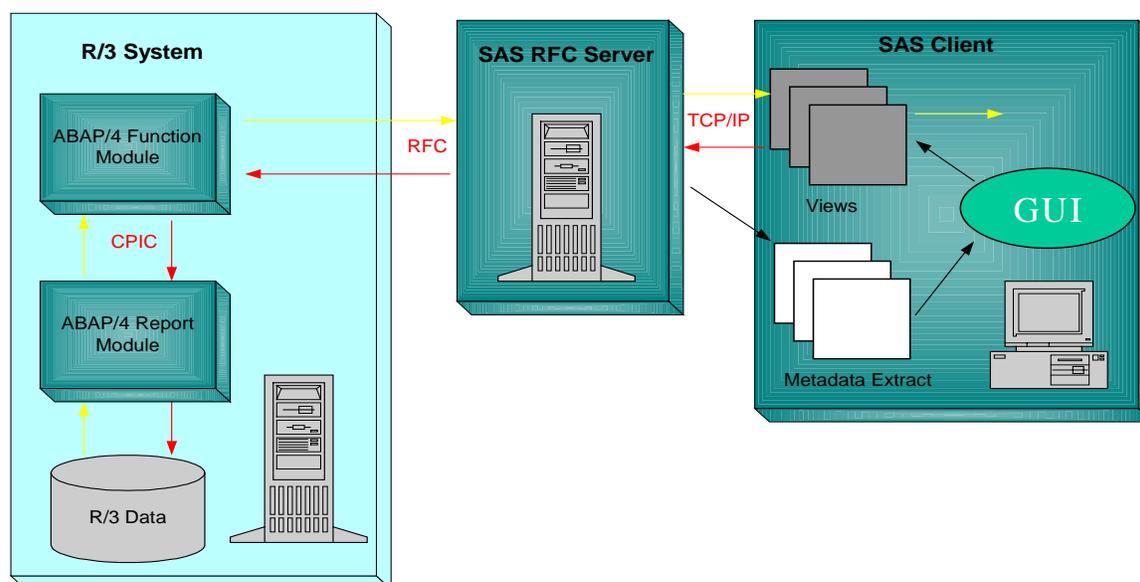The following picture shows these three elements for a better understanding.

Figure 1: SAS/Access to R/3 Framework

The illustration in Figure 1 demonstrates a robust, flexible and optimal framework for surfacing data within an R/3 System. The connectivity built into the solution allows enormous flexibility and the ability to support an environment distributed across many platforms. The Snapshot of the SAP Metadata (Metadata Extract) is used to navigate through the operational metadata without impacting the production application. By providing a business content metadata structure and robust search capabilities, much of the complexity behind the system is removed or hidden.   The SAS/Access interface to R/3 is also fully integrated with the SAS/Warehouse Administrator product so the metadata content can be used to directly populate the metadata repository and be used through-out the entire warehousing process.

<u>The Nuts and Bolts:</u>

So far we have laid out the components for a warehousing solution off of R/3 data and the framework for successful integration. Now we will take a closer look at what makes up the interface, the steps required for seamless access and the benefits of the solution.

Before moving along, we should identify a few problems that may be encountered when trying to extract data from R/3 or attempting to make it available for business user access.

The first issue to consider is that users of the R/3 System itself are typically not the same as the users requiring R/3 data to make critical business decisions. They often need information already consolidated, sorted or even enriched with data coming from the other legacy systems of their company.   SAP Systems contain detail-level data that is transaction-oriented in order to ensure the best performance for the SAP

end-users. This specific architecture of SAP systems associated to the fact that people needing the data have little SAP and even less ABAP/4 (SAP's programming language) knowledge make it usually the responsibility of the IT department to create (lots of) ABAP/4 reports. These ABAP/4 reports usually lack of flexibility and will confine the IT department in the role of writing and modifying ABAP/4 reports to answer ad-hoc business user requests. The complexity of this reporting environment and the low availability of trained resources drives the cost of customization through the roof. It is not uncommon to hear customizations for a single report costing upwards of $10,000 (as described by one customer currently using the SAS Intelligent Warehousing solution to eliminate this excessive cost).

While the IT department is kept busy writing ABAP/4 reports they are facing yet another problem: where do they find the information? SAP systems indeed make the creation of entities, data domains, etc, an easy task at the logical level of the SAP data model. But IT users usually do not know where information is physically stored.

The RFC Server

The RFC Server is an extremely important component of this framework. It performs several roles in integrating with the R/3 System. First, it is the control point for authentication into the SAP System. The interface is designed to leverage the robust security model built into the R/3 system and requires each request to be checked for appropriate authority.   The RFC server also provides the controls for communicating with multiple R/3 instances. *This is important in situations where the R/3 infrastructure includes multiple instances, possibly broken up by geographical regions or business functions like HR, Finance, etc.* This is also important when testing the extraction process off the Development or Test instance before rolling into production. The RFC Server contains parameters that provide control over what instance it communicates with.

Navigating the R/3 data dictionary

The SAS/Access Interface to R/3 is much more than a simple extraction engine. The interface provides the ability to navigate through the R/3 dictionary and illustrate the relationships between the various elements and entities of the system.   Understanding the relationship between the 20,000 + tables and their associated fields can drastically reduce the pain many IT professionals are currently faced with when trying to pull data out of an R/3 System. One customer was recently quoted saying "Understanding what is in an R/3 system is about as painful as having a root-canal without the novocaine".   Another added benefit available with the data dictionary navigation is the ability to intelligently search through the metadata using common business terminology or descriptions.

 These benefits are made possible by the ability to navigate the SAP R/3 dictionary and understanding/illustrating the relationships between the various elements of this structure (Figure 2). SAS/ACCESS to R/3 does not just read some tables and fields names from SAP, it also links the selected SAP table to all related tables, extracts the metadata stored in SAP and allows for intelligent searching through the dictionary.
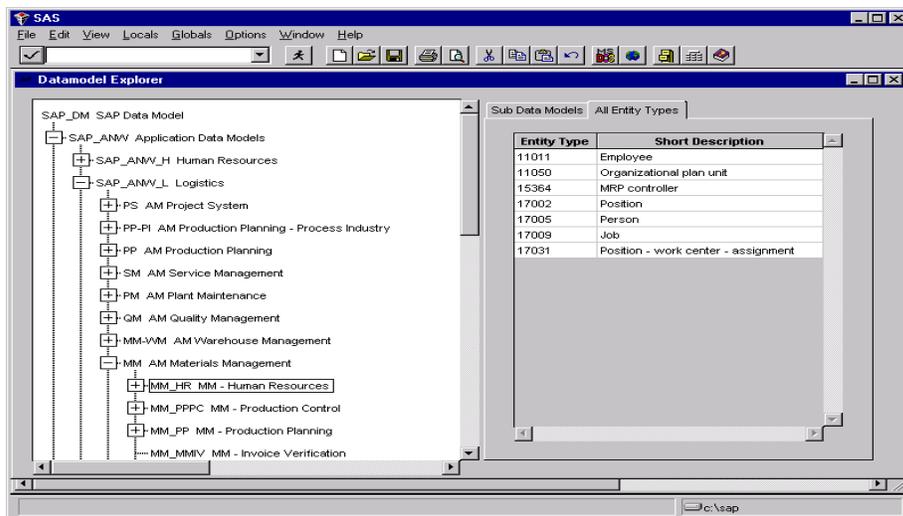
Extracting the data

Figure 2: Using the R/3 Data Dictionary

When the desired table(s) have been found, a few steps are to be accomplished in order to extract the data. If a connection between the SAS System and the SAP R/3 system has not been established, the connection window will allow the user to determine a connection identifier for the new extraction process (Figure 3). This connection ID can be saved for future use of the SAS/ACCESS to R/3 interface. A valid SAP Client, User and Password has to be provided. The connection ID contains information about the desired SAP instance, where the extracted data will be stored and some specific communication parameters (TCP/IP port, etc.) if one wishes to bypass the defaults proposed by the application.
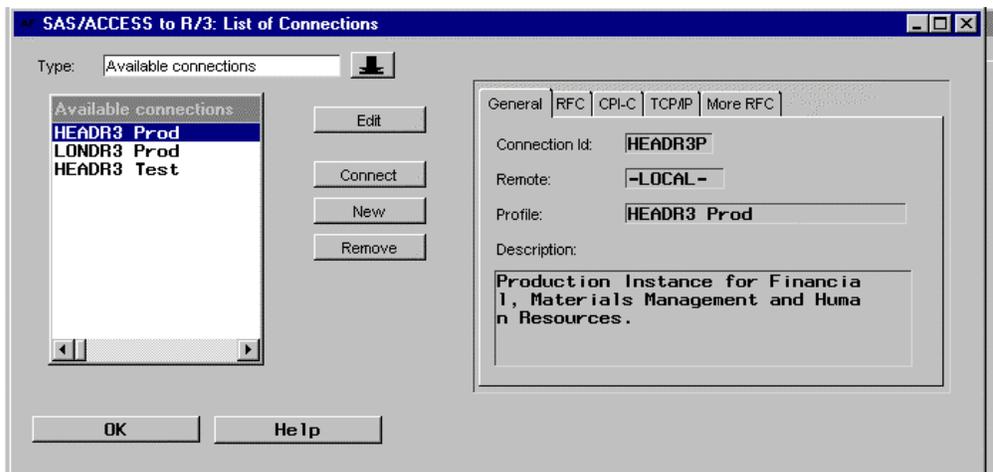

Figure   3:   Specifying   a   Connection

The next step of the process is to select the table to extract, specify the field(s) to be chosen from this table and define any subsetting criteria (select only information updated after a certain date, etc.). This step is important in that it allows the data being pulled back from R/3 to be limited to only that which is desired.

When selecting the SAP table to extract, a "where" clause may be entered that will be executed directly in the R/3 (Figure 4). The "where" clause is generated in a point and click environment and is checked automatically by the application the ABAP/4 syntax, thus avoiding required knowledge of how to actually write it in ABAP/4.
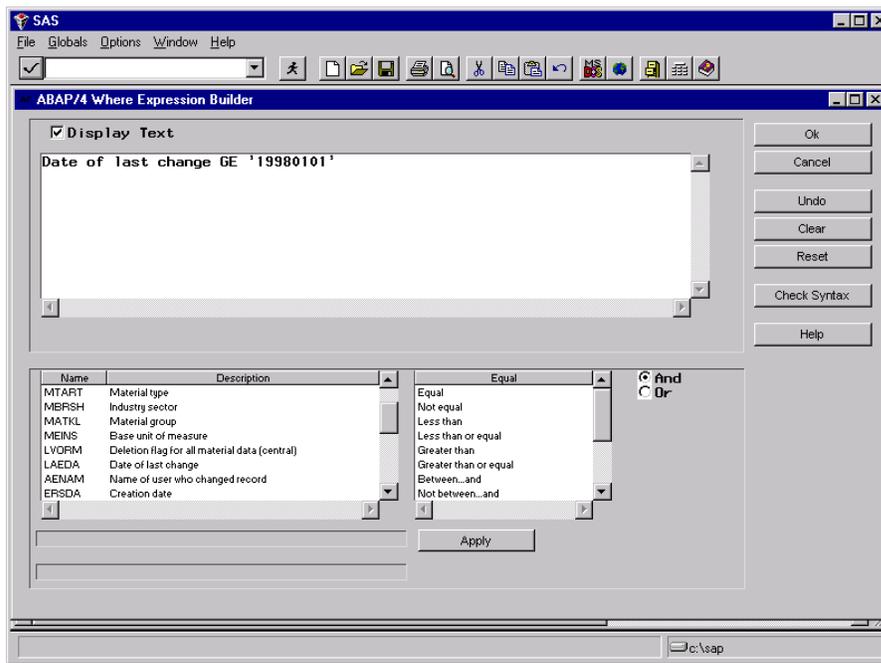

Fig4: Selecting specific information

Warehouse Administration

Once the data extraction requirements are defined and a view (logical mapping to the SAP structure) is created it is now time to begin the warehouse process. As described above, simply extracting that data from the operational source is only one piece of the puzzle. The remainder of the process must manage and execute the transformation (cleaning and scrubbing of the data), the data movement to the desire platform and the loading into the target storage structure and format. The SAS/Warehouse Administrator provides a single point of control for this entire process. Once the R/3 metadata is pushed into the SAS/Warehouse Administrator metadata repository, the information and knowledge about the SAP environment become part of the warehouse process and are treated simply as just another data source.

Full exploitation of the R/3 dictionary allows incremental updating of the data warehouse, which is critically important given the high volumes of data typically involved in R/3 systems.

The following illustration (figure 5) shows how the SAS/Warehouse Administrator can graphical display the various steps of the warehouse process including combing data from multiple data sources (including SAP R/3). The Process Editor provides a simple view of a complex environment while linking metadata related to each phase of the warehouse process.
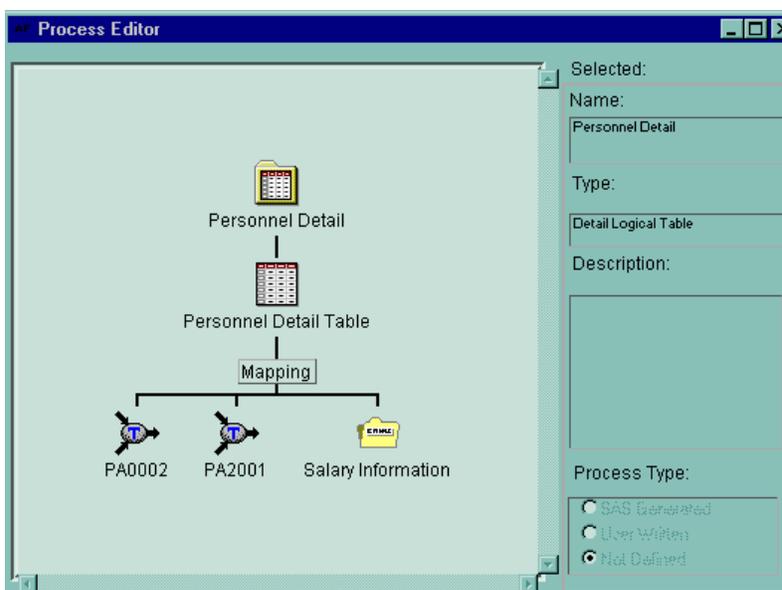

Figure 5: The Warehouse Process

Benefits of SAS/Access to R/3

The SAS solution for solving these problems in accessing R/3 data is obviously attractive. The benefits included (but are by no means limited to) the following advantages:

- There is <u>no need</u> to know the physical name of the table in the underlying database of the SAP System where information is stored.
- There is <u>no need</u> to know ABAP/4 since the interface will guide you through the complete process of selection and extraction of the information.
- It allows access to all the different data types contained in R/3 – Transparent, Pooled, Clustered and SAP Views.
- The solution is independent of the underlying database structure.
- Leverage the inherent security embedded within R/3
- You have the ability to navigate and understand the SAP Data Dictionary or metadata structure *without requiring extensive SAP specific knowledge* (this capability alone has attracted SAP support staff – giving them a view and understanding of business logic and relationships that was not previously available … at least not easily accessible).

<u>Exploiting R/3 data for a competitive advantage</u>

Accessing R/3 data and using it to populate a SAS Intelligent Warehouse is, of course, only the first part of the story. The profitable part is yet to come. After R/3 data has been accessed, combined, massaged and stored in a SAS Intelligent Warehouse, business user can start asking their own questions and getting their own answers. They can get the data they need, drill down, perform multidimensional analysis, and turn their data all around to examine and explore it.

SAS Institute offers a wide range of analytic applications (such as Online Analytical Processing (OLAP) for multidimensional analysis, data mining, query and reporting). They form part of a complete data warehousing system: from back-end (accessing the data from R/3 and other operational or external sources) to front-end applications (analytical, etc.). Moreover, these SAS solutions are fully Web-enabled, giving business users at R/3 sites the opportunity to access, manipulate and report on information via an ordinary Web browser.

Typical applications for these technologies include balanced scorecard reporting, customer relationship management, supply chain management, procurement, human resources management and financial risk controlling. For additional details on these front-end applications, please refer to the relevant SAS Institute literature.

Rapid Return on Investment

The Rapid Warehousing Methodology

Rapid deployment to business users, rapid return on investment and demonstrated benefits are key to success in data warehousing. They ensure that high-level sponsorship of the project is maintained and that it gets the resources required.

Therefore, SAS Institute has developed the Rapid Warehousing Methodology, which is designed to secure quantifiable benefits within a period of not more than 90 days, and a return on investment typically within one year.

This methodology is based on years of experience with implementation at hundreds of customer sites Worldwide. Its main features are:

* "Think globally, act locally"
* focus on the most pressing business objectives first
* repeatable process
* iterative development
* scalable approach to warehouse architecture

For further details please refer to the SAS Institute White Paper, SAS Rapid Warehousing Methodology.

Rapid Results through Business Templates

In the case of R/3, SAS Institute provides a set of pre-defined decision support solutions or business templates. The business templates are designed to address the common denominator of specific business problems. The business problems are either broken down by Functional area like: supply-chain management, sales and distribution, quality management or Departmentally like: finance, human resources, customer relationship management, procurement, etc.

The process behind a business template is to define a set of business questions that pose an immediate pain within the organization. These questions may be answered in the form of a report or multidimensional database or they may require some analytic capabilities. An example of such questions, say for Cost Center Analysis, may include:

* What is my organizational cost structure?
* What is the variance for actual and budgeted cost for a given organizational unit?
* For a given unit compare the variance of actual and budgeted cost by fiscal years and months.
    - Are costs within the budgeted amounts?
    - Which units are over budget?
    - Which Months?

– Which Years?

Based on this common set of business questions, the appropriate decision support model or structure is created to support the process. The next piece of the business template solution is probably one of the most important or at least the most beneficial. That is the R/3 content knowledge. Business Templates will provide a warehouse environment that contains pre-defined extractions of the R/3 data as well as all the transformation and loading requirements to populate the desire data model. This content knowledge takes the navigation capabilities (as described above) one step further, it provides the intelligence on what table(s) and their relationships are required to answer the defined business problems. These extractions now become the starting point of the warehouse process.

It should be noted that Business Templates are not the end all data warehouse solution. However, they are an excellent solution to getting on the way to solving more strategic and complex business problems. Business Templates are built on the foundation of the SAS Intelligent Warehousing Solution with this integrated capability (especially with the SAS/Warehouse Administrator), Business Templates are robust and easily extendable to meet what ever the individual business needs are. They also act as the stepping stone into more strategic business intelligence or decision support applications such as Customer Relationship Management, Balanced Scorecard, and more.

Business Templates will be provided through a Rapid Results service offering. This service will include on-site consulting for installation, implementation and knowledge transfer of the Business Template. Upon completion of this service the ability to turn your R/3 data into meaningful information for analysis and exploration will be a reality. An example of the reporting environment may look like on of the following Java Applets (Figure 7 and 8):
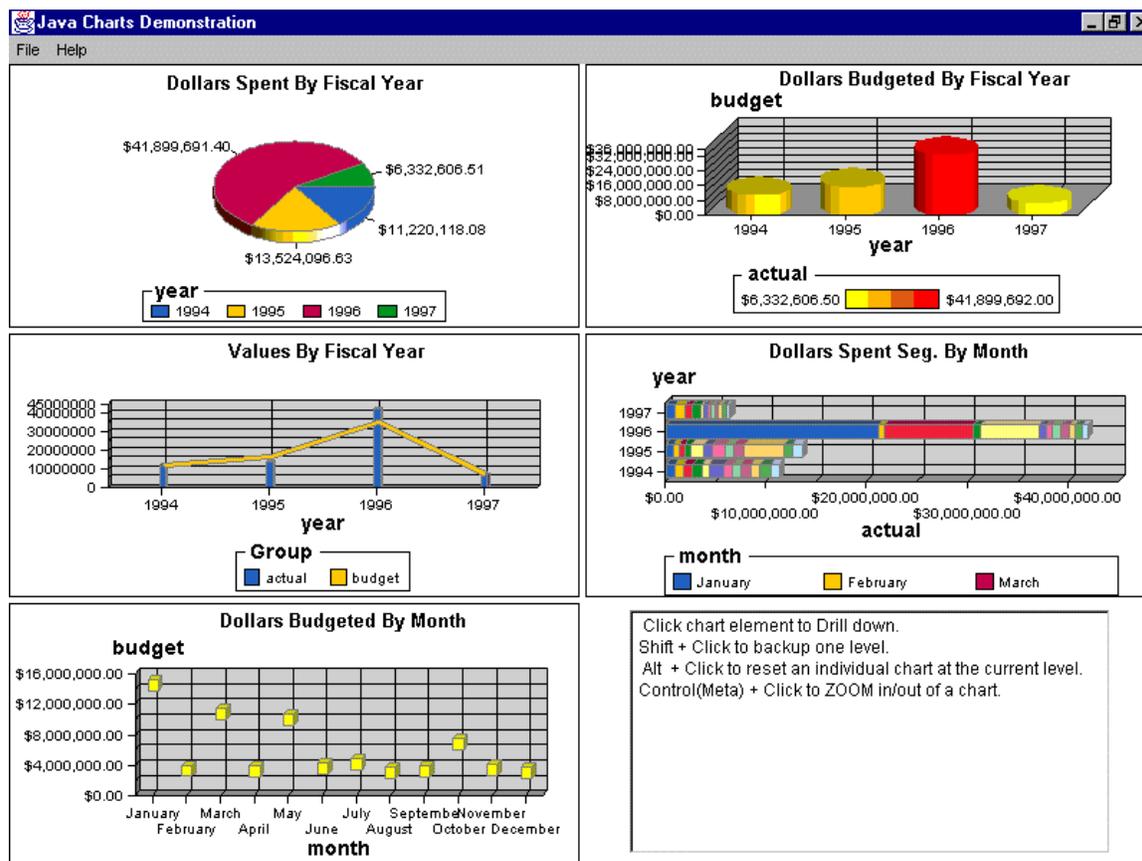

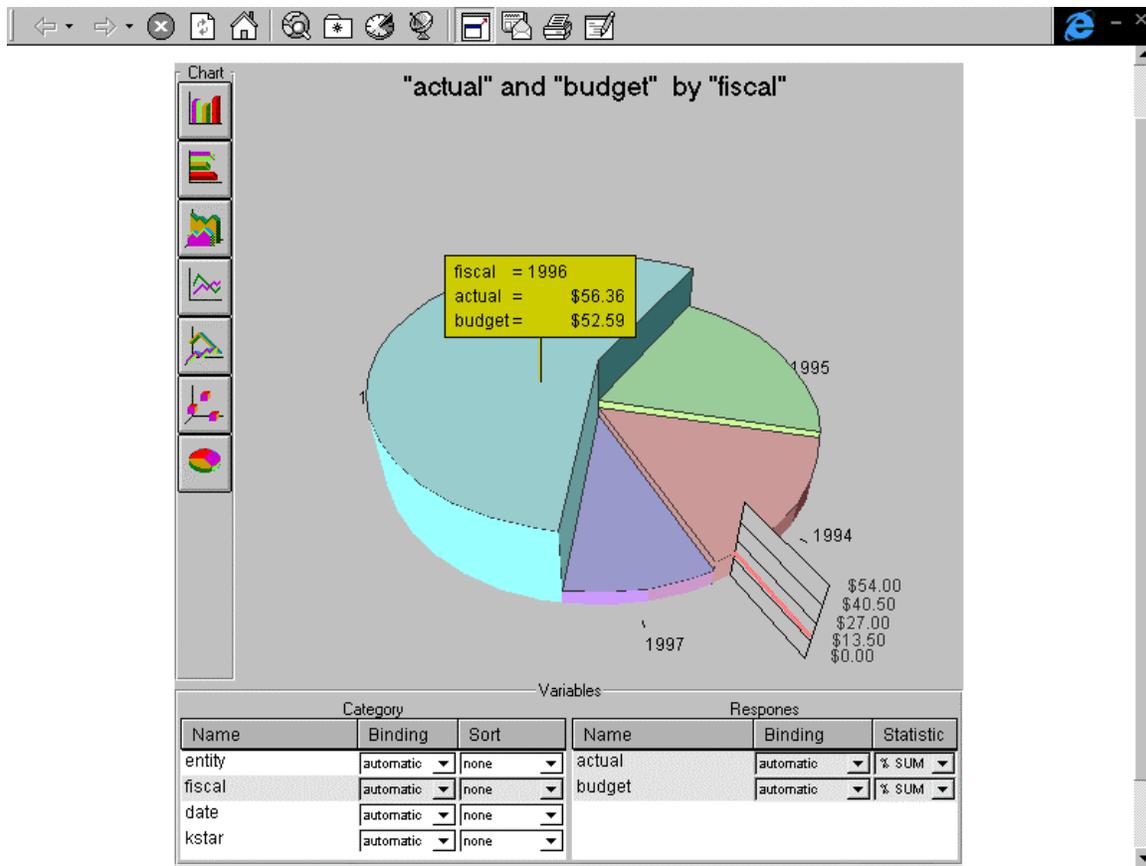Figure 6: Multiple views of Actual and Budgeted Costs through a Java Application

Figure 7: Actual and Budgeted Cost through a Java Applet

Conclusion:

Over the past few years, thousands of the world's top organizations have invested in ERP Systems to improve their critical business processes. For the most part the experience has been good in terms of the automation of operational functions. The experience has not been as good in terms of providing information to key decision-makers. It usually takes about 1 year in the implementation for this to become apparent. ERP Systems typically remain popular with supervisory staff and middle managers (after any initial implementation problems have been resolved), but less so with business analysts and upper-level managers who have longer time horizons and whose decisions depend on having cross-functional information.

At the very senior levels, questions are inevitably asked about the return on investment in ERP systems. In short, ERP Systems, including SAP AG's System R/3, have generally not delivered on the promise to make information available to the upper levels of management, and have completely failed when it comes to advance analytical techniques such as data mining.

It is therefore advisable to follow a "core competency" strategy by choosing the leading vendor for ERP and the leading vendor for Information Delivery. SAS Institute is committed to the end-to-end process of extracting information from heterogeneous environments and making it available for business intelligence purposes.   The SAS Intelligent Warehousing Solution is the optimum solution wherever R/3 users organizations have more than one R/3 instance, need to combine R/3 data with other sources, have legacy data or there is a possibility of organizational changes (such as merger or acquisitions).

The key enabling technology for accessing and navigating the R/3 data and populated the SAS Intelligent Warehousing Solution is the SAS/Access Interface to R/3. In closing, here are a few key features and benefits of the SAS Intelligent Warehousing Solution to R/3, including the SAS/Warehouse Administrator and SAS/Access to R/3:

- Access to a R/3 data is done through the metadata, not through the underlying database. This approach to data extraction allows access to all the different kinds of SAP tables: transparency, pooled, clustered and SAP views. Another important resulting benefit is the fact that the solution is independent on the underlying database used by SAP (i.e. Oracle, Informix, etc.).

- The full exploitation of the SAP metadata allows incremental update of the data warehouse, which is of critical importance given the huge amount of data usually considered in SAP Systems.

- The security put in place for the SAP System is exploited by the SAS/ACCESS to R/3 solution. No need to redefine specific security for the extraction process: SAP Client, user ID and Password have to be provided and validated before access is granted.

- The solution proposed is SAP hardware independent and allows multiple configurations of the different components of the solution, whatever the method of extraction (online or batch). This ensures hardware resources and network traffic optimization.

- Using Remote Function Call and the CPIC internal protocol of SAP provide maximum performance for the extraction process.

- SAS/ACCESS to R/3 has a point and click interface and does not require any SAP or ABAP/4 knowledge. This results in significant time and cost savings as well as increased productivity of both the IT and Business Communities.

- Rapid Results with Business Templates provides the framework to begin solving business problems. Built with the Rapid Warehousing Methodology ("think globally, act locally") in mind, business templates will act as the stepping stone to more strategic applications.

- Finally, the integration with the entire suite of SAS products and solutions including the SAS/Warehouse Administrator illustrates why the SAS Intelligent Warehousing Solution is a key component to your strategic vision for leveraging your SAP investment and the goal of achieving or maintaining your competitive advantage.

For more information on how SAS Solutions can leverage your SAP (ERP) investments, visit our Web Site at: www.sas.com Learn how you can take advantage of industry leading solutions to deliver competitive advantage through Customer Relationship Management, Data Mining, Balanced Scorecard, Risk Management/Forecasting, Financial Consolidation and Reporting, IT Service Management and much more.

Trademarks:

SAS, SAS/Access, SAS/MDDB, Scaleable Performance Data Server, SAS/Warehouse Administrator, SAS Rapid Warehouse Methodology are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand or product names are registered trademarks or trademarks of their respective companies.

References:

SAS Institute White Paper: SAS Rapid Warehousing Methodology, 1998

Author Information:

Eric Hunley,
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Marketing Product Manager, ERP Warehousing Solutions
919/677-8000 X6501,
sasexh@wnt.sas.com

Craig Rubendall,
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
Development Manager, Data Warehousing Technologies
919/677-8000
sascxr@wnt.sas.com