

Release the potential of your R/3 System!!!

The SAS® Warehouse Solution to SAP AG

SAS Institute Inc.
August 12, 1998

Overview:

Many Organizations have made the strategic decision to replace existing production applications with package solutions designed to re-engineer their existing business process. Actually these applications are designed to do much more, including things like replacing outdated applications with ones that are Year 2000 compliant, cross business process integration, and more. 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 60% of the market share. SAP AG, as with many of the other ERP Vendors, 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.

The decisions to implement these solutions have proven to be successful in areas of integrating business processes for automated and optimal production system management. 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 do I turn this packaged application and its plethora 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 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 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 Warehousing Solution including some technical details on the architecture chosen to integrate R/3 with other sources of data within your organization.

SAS® Data Warehouse Solution:

The SAS Data Warehouse 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. As the world's largest privately held software company, SAS Institute has built and maintained its customer base of 3.5million users by building decision support systems for a diverse user community.

The components of the SAS Data Warehouse Architecture and required building blocks for a successful decision support environment include:

Extraction of data from operational and legacy environments. SAS® provides data access engines to 100's of data sources. These include native access engines to the more popular data sources across

multiple platforms. Included are DB2, Oracle, Informix, IMS, Sybase, SQL Server. In addition to these popular data stores, SAS access technology also provides the ability to integrate with data in ERP environments. In particular, SAS/Access® to R/3 allows an organization to access SAP data along with other source systems, providing a true enterprise view of information.

Adding value and enriching data through a transformation engine. The SAS transformation engine is a mature, platform independent, procedural language that provides facilities for integrating data from diverse operational environments, validating that data, summarizing, and applying advanced statistical functions in support of data restructuring for the decision support environment. This task alone in many situations can take up to 75% of the total time creating, maintaining and processing the warehouse. This phase of the warehousing process is where the prevention of the "Garbage In, Garbage Out" syndrome takes place.

Loading of the warehouse. Once transformed, SAS once again takes advantage of its access technology to populate warehouses and marts across required platforms. Again, the data is not necessarily stored in a single database, but may be distributed to various desired targets and formats. Storage possibilities include, but are not limited to: SAS, Oracle, Informix, Sybase, SQL Server, OLAP cubes (Relational, Multi-dimensional and Scalable), ODBC-compliant structures.

Management of the process. Because the solution to data warehousing and decision support encompasses a broad range of hardware platforms, operating systems, data sources, and data targets, managing this process can be quite challenging. SAS/Warehouse Administrator™ provides a single point of control for managing the process of moving data from the operational environment through to population of a distributed data warehouse environment. This is accomplished via a common metadata repository. 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.

SAP AG R/3 Integration:

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/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 impacts is to implement a warehouse or mart structure to support this environment. However, the question still remaining to be answered is what are my options to intelligently extract data from this system? The answer lies within 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 warehouse process (as described above). This approach will be outlined in more detail in sections to follow.

Before diving deeper into the process flow of the SAS/ACCESS to R/3 software solution we should take a close look at the components that make up this solution.

- Two template modules (an ABAP/4 Function and an ABAP/4 Report) are installed in the SAP R/3 system at the time SAS/ACCESS to R/3 is installed. The purpose of these modules is to resolve passed parameters at run-time to satisfy the particular data request. These modules communicate together using the CPIC protocol of SAP R/3, ensuring maximum performance.
- A RFC server is installed on the SAP R/3 machine or on another machine (without requiring the SAS® System). 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.

- 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.

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

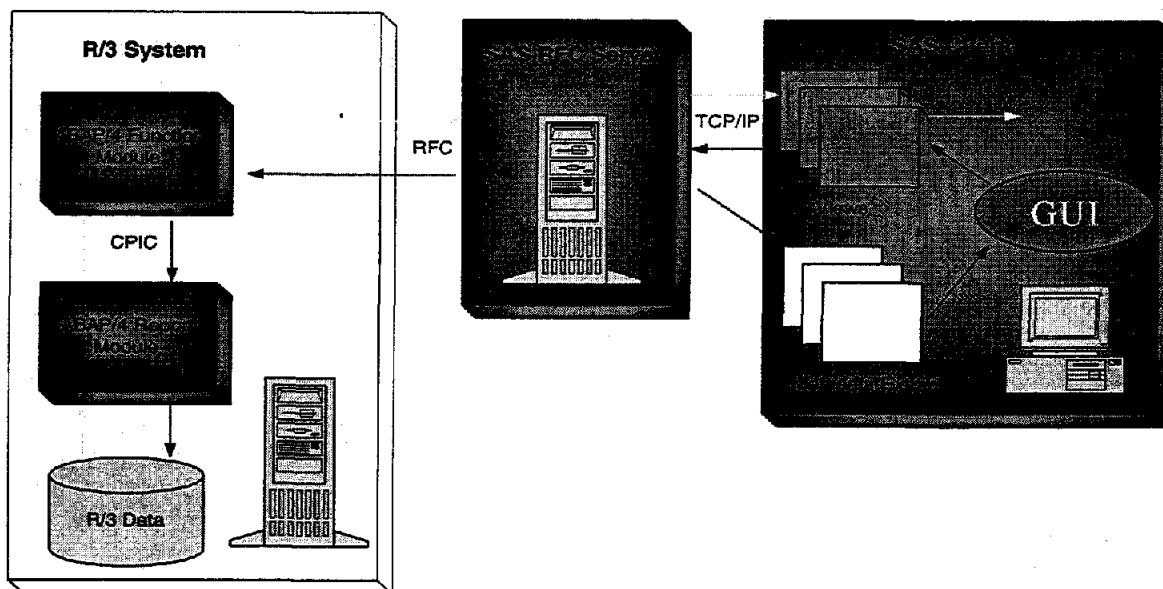


Fig2: SAS/Access to R/3 Framework

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.

The RFC Server provides 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. Another situation where this is important is when desiring to test the extraction process off the Development or Test instance. The RFC Server contains parameters that provide control over what instance it communicates with.

The Nuts and Bolts:

So far we have laid out the components for a warehousing solution off of SAP AG 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 at hand is that the users of the SAP System are typically not the same as those requiring it's data to make critical business decisions. Furthermore 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 personal 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.

While the IT department is kept busy writing ABAP/4 reports they are facing one new 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 never know where information is physically stored.

The SAS solution for solving these problems in accessing SAP AG's data is obviously attractive by the following two statements:

- No need to know the physical name of the table in the underlying database of the SAP System where information is stored.
- No need to know ABAP/4, the interface will guide you through the complete process of selection and extraction of the information.

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. 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 or allows for intelligent search in the dictionary.

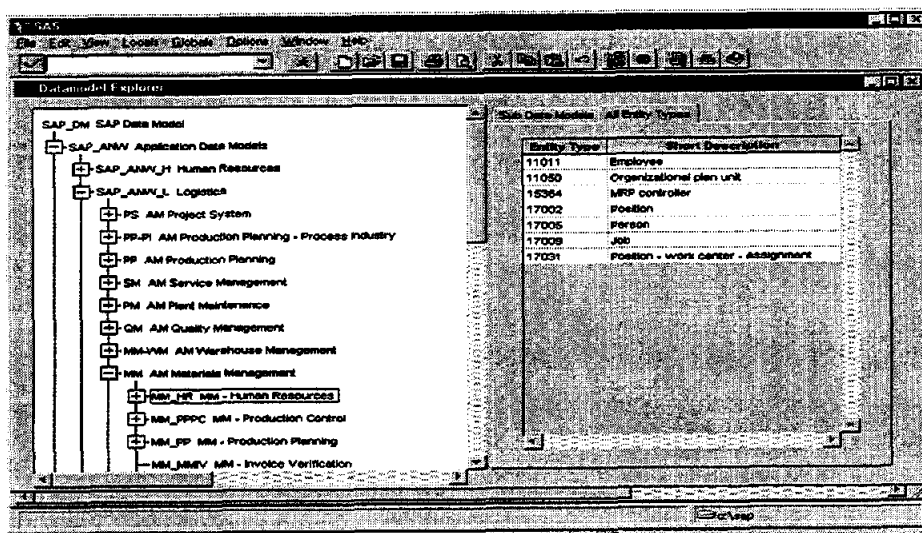
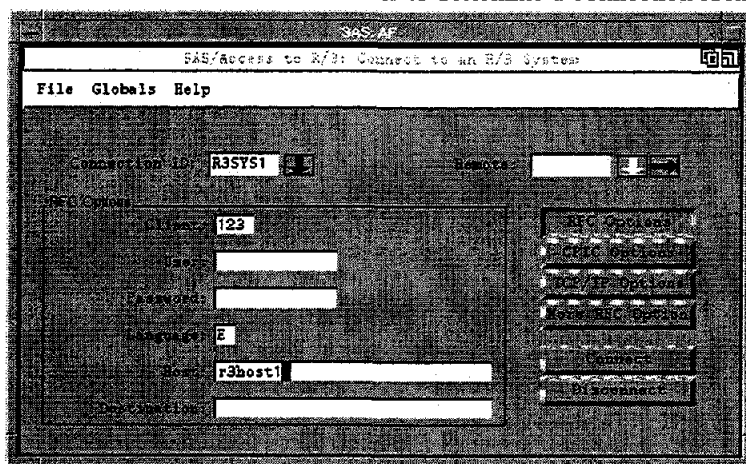


Fig2: Using SAP R/3 Data Dictionary

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.



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 where the extracted data has to be stored and some specific communication parameters (TCP/IP port, etc.) if one wishes to bypass the defaults proposed by the application.

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.

Fig3: Specifying a Connection Identifier

When selecting the SAP table to extract, a where clause may be entered that will be executed directly in the SAP system (i.e. the where clause is passed through to the SAP System). Note that this where clause is generated in a point and click environment and is checked automatically by the application as being a valid ABAP/4 where clause thus avoiding required knowledge of how to write it in ABAP/4.

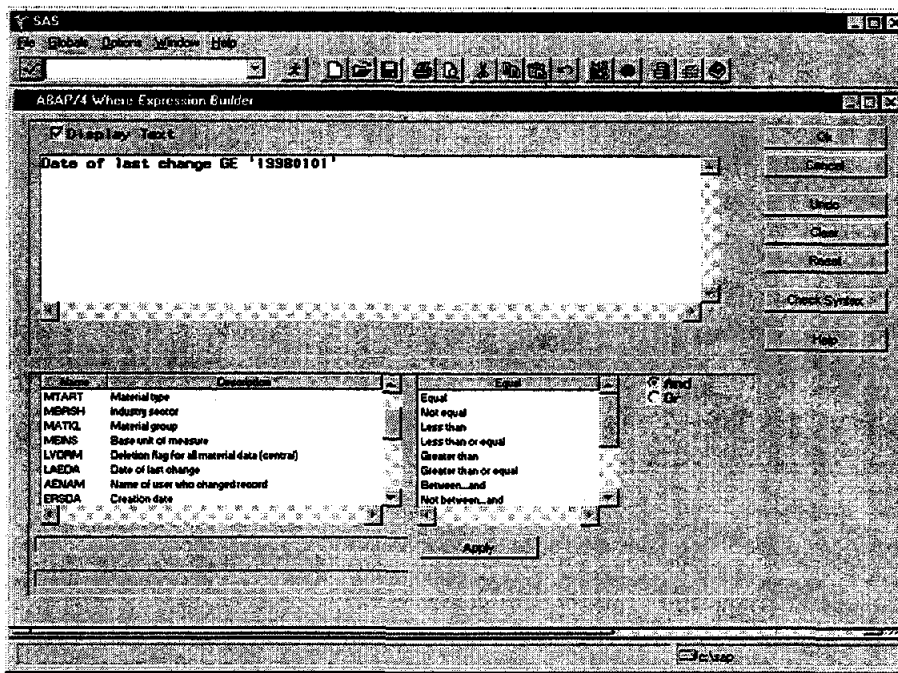


Fig4: Selecting specific information

Once the data extraction requirements are defined and a view (logical mapping to the SAP structure) is created it is now time to turn things over to 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 desired platform and the loading into the target storage structure and format. As mentioned above the SAS/Warehouse Administrator provides a single point of control for this entire process. Once the SAP Metadata is pushed into the SAS/Warehouse Administrator metadata repository the information and knowledge about the SAP environment are now part of the warehouse process and are treated simply as just another data source.

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

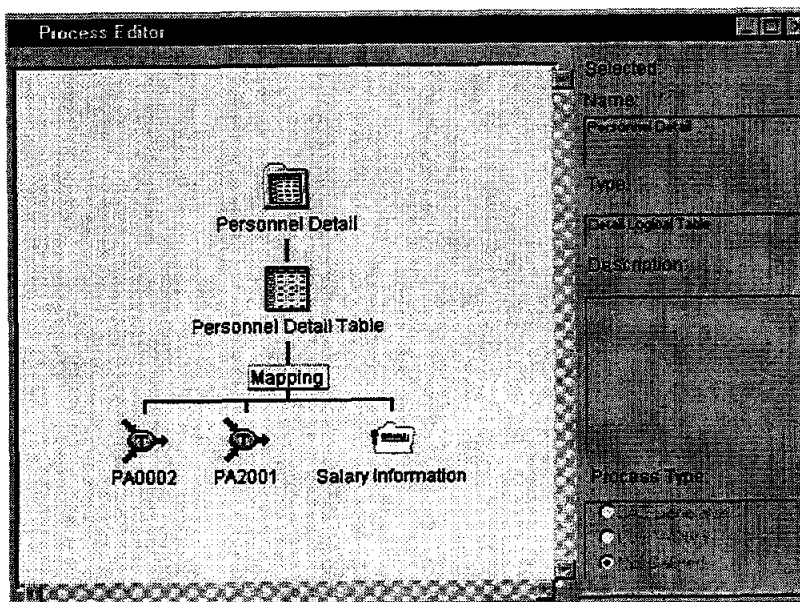


Fig4: The Warehouse Process

Conclusion:

To conclude we would like to point out some major benefits of using the SAS/ACCESS to R/3 solution to extract data from any SAP AG's Operational System.

- Access to a SAP System is done through the metadata, not through the underlying database. This makes sure we will get all the data contained in the SAP system at a given point of time. By taking this approach to data extraction the interface can access all the different kinds of SAP tables. These tables include "TRANSP", "CLUSTER" and "POOL" type of tables and especially SAP VIEWS can be read without any specific effort. 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 fully used 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 checked as valid 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. There is no need to create a temporary table containing the extracted data but rather, it starts flowing immediately whatever the format of the extracted information is (SAS table or view).
- SAS/ACCESS to R/3 has a point and click interface and does not require any SAP or ABAP/4 knowledge. It nevertheless opens the potential of the SAP data to any authorized user, for its exploration, analysis and exploitation using the powerful tools of the SAS System.

For more details on the SAS Warehousing Solution for SAP AG, please feel free to contact:

Eric Hunley,
SAS Institute Inc.,
Marketing Product Manager, ERP Warehousing Solutions,
919/677-8000 X6501,
SASEXH@wnt.sas.com