Adequately Account for Business Requirements and Avoid Archiving Anarchy

Rolf Gersbacher

Dr. Rolf Gersbacher joined SAP in 1994, and since 1998 has been a member of SAP's Performance, Benchmark and Data Archiving department, where he is responsible for conducting extensive analysis, and has co-written projects in the area of data archiving. Together with Helmut Stefani, Rolf developed the best practices guide "Managing SAP Data Archiving Projects."

(complete bio appears on page 60)

The use of SAP systems by many users — thousands in some cases — leads to the creation of enormous volumes of data. At times, data volumes can even fall in the terabytes range. In order to guarantee efficient use of storage devices, CPU, memory, and other resources, and, more important, to ensure acceptable levels of system availability and throughput, data archiving is performed to relocate data objects from the online database to external storage media, all the while enabling read access to that relocated or "archived" data.

SAP has developed two tools to provide users with read access to archived data: the Document Relationship Browser (DRB) and the Archive Information System (AS).[1] While the technical facets of setting up and using these tools are easy and straightforward,[2] the planning you must do ahead of time is not.

The archiving decisions involved in the planning process are not black-and-white technical decisions. An "old" document from the perspective of the technical team may not be an "old" document in the eyes of a user who can't bear to see that document relocated from the online system to an archive file. Decisions like this require tradeoffs, because once data is archived, it cannot be changed. Technical teams

[1] The DRB is a tool used to display linked data objects and the entire business process of which they are a part, whether they are archived or in the live database. The AS is a tool used to display and efficiently search archived data. These key tools can be used in a complementary fashion or independently. The DRB is part of the standard R/3 system since 4.6C and the AS since 4.5A. For installation on earlier releases, see SAP Notes 567195 and 99388.

[2] See the article "Mastering the Administrative and Development Tasks Required to Put Archived Data in Easy Reach of Every User" (July/August 2001) for more on configuring the DRB and the AS.

Figure 1 Examples of Business, Legal, Auditing, and Technical Considerations

Business Considerations	Possible Actions
Is it necessary to reconstruct the business process context for the affected data objects?	If yes, configure the AS for all involved archiving objects. If no, configure the AS only for the relevant archiving objects.
What search criteria are needed?	Ask end users which search fields are required for accessing archived data objects, and define archive information structures (infostructures*) using those search fields.
What data (fields) must be accessed?	Evaluate whether this information is displayed in business views. If not, show end users that the information can still be accessed through technical views within the AS. If neither is possible, modify existing business views by copying them into the customer namespace.
Is reusability a requirement? Will it be necessary to create a new document based on a template in the archive? (Note that only very few objects support the "create with reference" functionality.)	If yes, archive only data objects for which this requirement can be ruled out by setting a high residence time (the length of time data stays resident in the online system before it can be archived).
Will it be necessary to carry out statistical analysis on the archived data?	If yes, carry out the analysis before archiving, while data is still in the database. If BW is in use, transfer data prior to archiving.**

* A database table that contains a pointer to the position of an archived data object in the archive file system.

** At present, SAP provides extractors to enable the BW to process archived data. The relevant functions for loading archived OLTP data into the BW will be generally available soon, enabling archived data to be loaded into the BW at a later point in time (for example, if a BW is implemented and used sometime in the future).

must therefore be prudent about the data objects they choose to archive. Just because it's technically feasible to archive an object doesn't mean it's wise to do so.

Also, don't let the technical ease with which data can be archived lull you into a false sense of security — it's not enough to use the SAP analysis tools to locate large or fast-growing tables, run transaction DB15 to identify the archiving objects, and then run the archiving programs. This is *not* a sound archive plan! These actions only touch the surface of the broader technical issues involved, including:

- Which data objects are "old" (i.e., rarely accessed in display mode) and can be relocated
- The business processes that require access to the data objects you plan to archive (even though

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

Figure 1 (continued)

Legal and Auditing Considerations	Possible Actions
Do internal auditing or controlling departments require access to archived legacy data?	If yes, investigate which fields are affected and demonstrate the possibilities for accessing this information (business views, technical views, AS archive information structures).
Which fields and what information must remain accessible for reasons of product liability?	Evaluate whether these fields should act as search criteria. If yes, include the fields in the AS infostructure. Configure the infostructure with the appropriate fields for emergency cases in order to activate and fill them.
What country-specific regulations require data to be stored accessibly? (In this scenario, accuracy, completeness, and consistency play an important role.)	Legal requirements vary from country to country, so it's best to ask and incorporate the information from the subsidiaries in the affected countries.
Technical Considerations	**Possible Actions**
What is the best storage medium for archive files?	When making this decision, take into account the existing hardware environment, scalability, integration into existing storage media, etc.
Is a content management system (CMS) in use? Is it necessary to integrate the archiving project with any "imaging" projects?	Investigate whether a content management system (CMS) should be used for storing archive files and for storing scanned-image documents. If yes, store archive files together with scanned documents, print lists, etc. in the CMS (for central storage of files, easy administration, etc.).
How can the required access time be realized?	For fast access (within a few seconds), configure the AS infostructures and choose appropriate storage media (for example, magnetic disks, network equipment, etc.).

we're talking about data objects belonging to closed business processes,[3] this is still a key consideration)

- Whether there are previous or subsequent documents referenced by the data objects to be archived
- How the archived data is to be accessed

These user-focused technical considerations, along with business, legal, and auditing requirements, are what make or break an archiving project. **Figure 1** shows some examples of these considerations and the possible ways to address them.

To help guarantee the success of an archiving

[3] A business process is considered "closed" when the last business transaction has taken place.

An Overview of SAP's Five-Phase Model for Data Archiving Projects

Project Initiation: Here you identify table sizes and growth rates, and the corresponding archiving objects. You also contact the appropriate user departments. At this stage of the project, mainly system administration personnel are involved.

Analysis (Business Blueprint): This phase is concerned with examining and documenting the potential effects that relocating data may have on your operational business. Special care must be taken with application objects that are incorporated into cross-application processes. All those responsible for these processes must be consulted and requirements must be set accordingly.

Design and Conception (Realization): During this phase, the business, legal, auditing, and technical requirements identified during the Analysis phase are implemented by the project team and evaluated by the user departments. This includes, for example, implementing a business concept or specifying the archiving sequence.

Final Preparation: During this phase, testing is carried out and customizing settings and programs are transported from the test system to the production system.

Going Live: Real archiving is then carried out — that is, the data is relocated to the archive

project, SAP has established a five-phase *Best Practices Model*, which is summarized in the sidebar above. First is the Project Initiation phase, where you determine from a purely technical perspective which data objects are good candidates for archiving. Next is the Analysis phase, where you decide whether it is wise from a *business* perspective to relocate those data objects. Next is the Design and Conception phase, where you design your archiving solution. Then comes the Final Preparation phase, where testing is carried out and customizing settings and programs are transported from the test system to the production system. In the last phase, Going Live, data is relocated from the production system's online database to the archive system.

Without question, the Analysis phase, the focus of this article, is the most important.

The best practices associated with the Analysis phase were designed to help you get a handle on all the technical, business, legal, and auditing issues of an archiving project.[4] This article shows archiving teams, using the central Logistics sales order process as an example, how to apply these best practices in your own environment.

[4] The full Best Practices Model is available at **http://service.sap.com/data-archiving**. From there, choose *Media Center → Literature*.

The Steps of the Analysis Phase

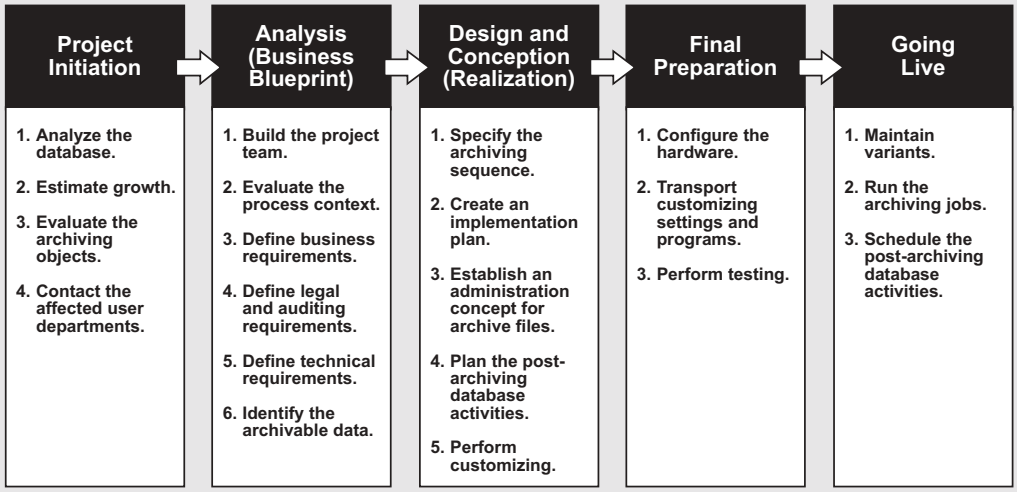
The Analysis phase includes the following steps:

1. Build the project team.
2. Evaluate the process context.
3. Define business requirements.
4. Define legal and auditing requirements.
5. Define technical requirements.
6. Identify the archivable data.

The order in which the steps are performed is important: notice that the definition of business, legal, and auditing requirements precedes the definition of technical requirements. Leading with business requirements is often neglected in archiving projects and, as a result, errors nearly always ensue.

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

system and then deleted from the online database. Post-archiving activities, such as database reorganization, should also be scheduled during this last phase.



For example, in April 2001, I assisted the archiving team at a large corporation that had archived some closed sales documents created during the current fiscal year. Even though the status of these archived sales documents was "closed," they contained cost-relevant data still needed for the current year-end closing, and since archived data cannot be processed, these documents needed to be in the live database (not in the archive system). The archiving team's error was choosing too short a residence time[5] for the cost-relevant data. Fortunately, there is a program (reload archive) that can reload parts of archived sales documents into the database, so the problem was easily solved. But this function is not available for all archiving objects, so this solution won't always work. The lesson here is that if the archiving team had started with the business requirements, the accounting team could have helped define the appropriate residence time for the cost-relevant data, saving valuable time and resources.

Let's now examine the steps of the Analysis phase using the elements involved in a sales order process, considered to be *the* central process in Logistics. The sales order process centers on the areas of procurement, production, and sales, and involves data objects such as sales documents, shipment documents, shipment cost documents, deliveries, billing documents, production orders, material documents, and, in certain situations, backflushing documents.

The sales order process usually starts with an inquiry in which the customer requests more specific information about a product (its price, quantity, any discounts, etc.). The customer then receives a quotation, based on the inquiry, that contains the corresponding conditions of sale. If the customer accepts the quotation and submits an order, a production order is generated for products that are manufactured in-house. After the product is finished, it is shipped and delivered to the customer. The associated shipment document might also include additional information about the shipment, such as the shipping route, forwarding agent, means of transport, etc. The process ends with the creation of an invoice, which is sent to the customer and includes the forwarding agent's bill, product costs, etc.

[5] The length of time data stays resident in the online system before it can be archived.

The Most Important Logistics Archiving Objects

- ☑ **SD_VBAK (for sales documents):** Customer inquiries, quotations, sales orders regarding product delivery, and outline agreements are various types of sales documents. These documents use a common data model that includes approximately 150 tables in the database. All these document types have the same basic structure and underlying database tables. SD_VBAK archives the document header data from table VBAK, the document line items from table VBAP, and, for each line item, any schedule line data from table VBEP.* The document header contains information that is valid for all line items, such as the document number, name of the customer, name of the ship-to party, and so on. The line data lists the products or services that are to be shipped, and contains the material number, quantity, plant details, and so on. If the ordered product is divided into separate partial deliveries, there may be schedule line data for each line item. Addresses, change documents, classification data, configuration data, and texts are also written to the archive files with the relevant archiving classes.
- ☑ **SD_VTTK (for shipment documents):** In the case of shipments, incoming and outgoing goods flows are documented and controlled. In this process, outgoing shipments on the basis of sales orders and deliveries are especially meaningful. Shipment documents contain information on the means of transport, relevant haulage companies, itineraries, and deadlines. SD_VTTK archives the header data from table VTTK and the line item data from table VTTP.
- ☑ **SD_VFKK (for shipment cost documents):** Shipment costs play an important role in determining the price of a product. The shipment cost document summarizes cost-relevant data for a particular shipment. SD_VFKK archives the header data from table VFKK and the line item data from table VFKN. Corresponding change documents, address data, and texts are also included in the archive.
- ☑ **RV_LIKP (for deliveries):** The deliveries data object documents and controls all delivery activities, including the commissioning information (the details on the composition of the product according to order specifications), the packing information, and the goods exit, including transfer to transport. RV_LIKP archives the header data (such as the shipping-point delivery date and ship-to party) from

* With transaction DB15 you can show the tables from which any data is archived, along with their associated archiving objects.

Regardless of how your company has implemented this process, anyone working on a Logistics-related archiving project will be using a finite set of SAP archiving objects.[6] The most important ones are listed in the sidebar above, along with each object's associated data object (the application-specific object type). The data objects archived via these archiving objects access customer master and material master data, bills of material, and work schedules.

> ✓ ***Tip***
>
> *Tables that store transaction data generally grow much faster than those that store master data, so your focus should be on the former. These tables can be archived independently from the master data to which they refer. The archiving process does not check for master data.*

[6] An archiving object is the smallest entity with business meaning that can be archived as a whole. It ensures that all table entries characterizing a business object are written to an external archive file and deleted from the database. SAP provides archiving objects for all fast-growing standard tables. Archive Development Kit (ADK) archiving routines can be used to develop customer-specific archiving objects. For more on the ADK, see "Data Archiving Essentials — What Every Administrator Needs to Know" (March/April 2001).

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

table LIKP and the item data (such as material number, quantities, and prices) from table LIPS. Furthermore, address data, change documents, and texts are included in the archive file.

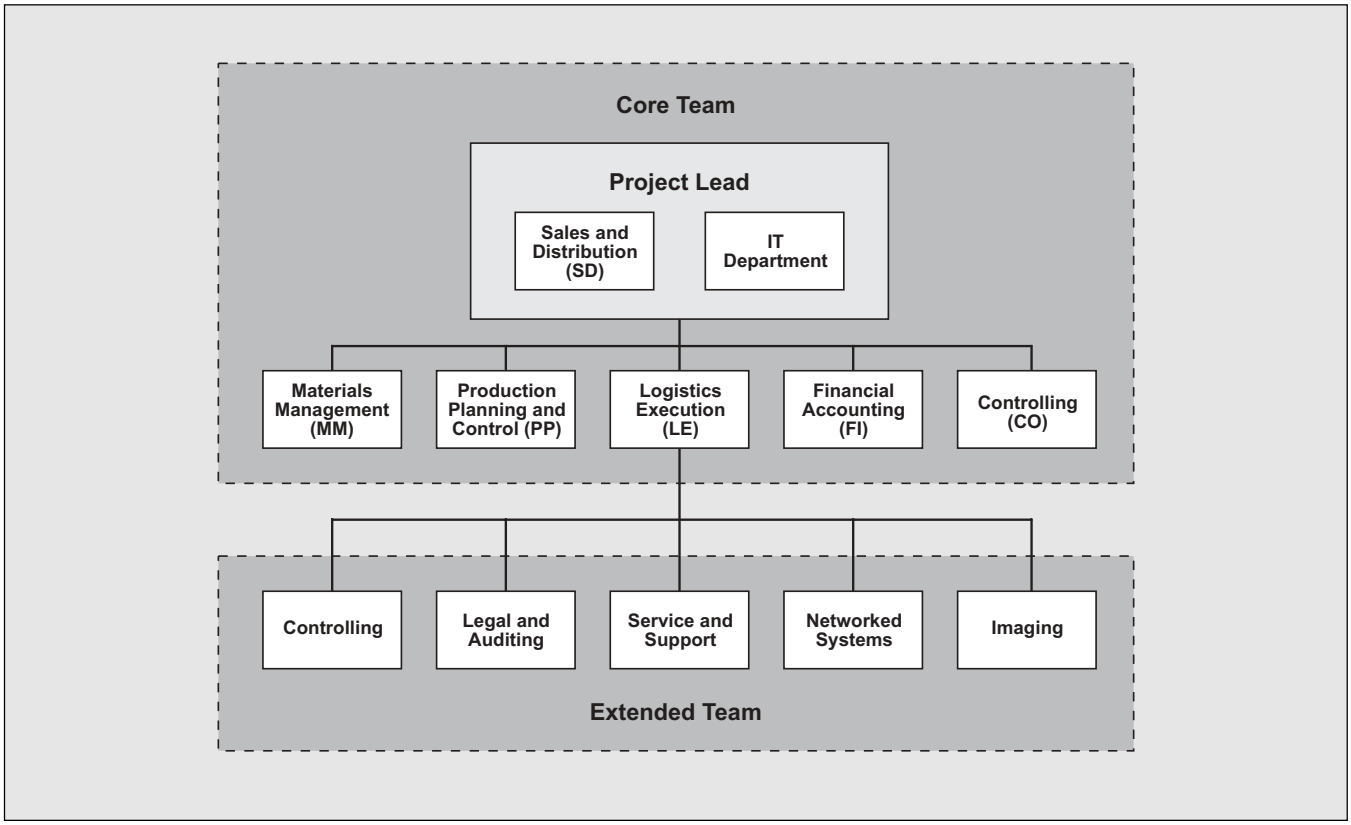
- ☑ **SD_VBRK (for billing documents):** A billing document signals the end of the sales process and controls the following functions: creation of invoices triggered by deliveries, creation of credit and debit memos in case of incorrect deliveries, cancellations, and so on. SD_VBRK archives the header data from table VBRK and the line item data from table VBRP, which contains the material number, billing quantity, price of individual items, number of the underlying delivery, etc.
- ☑ **PP_ORDER (for production orders):** The production order triggers and controls the production of a product. In addition, it controls the bills of material, the materials, the production operations and resources to be scheduled, and the dates to be observed. Archiving object PP_ORDER writes configuration data, settled (posted) costs from Financial Accounting (FI), used materials, confirmation of completed operations (processing steps), texts, and change documents to the archive file. PP_ORDER archives the header data from table AFKO and the line item data from table AFPO.
- ☑ **MM_MATBEL (for material documents):** Every goods movement triggers the creation of a material document. The material document header contains general data on the goods movement transaction, such as the posting date, transaction type, and document number. The items describe the individual material movements. Material movements include goods entrance, goods exit, stock transfer, and transfer posting. MM_MATBEL archives the header data from table MKPF and the line item data from table MSEG.
- ☑ **PP_BKFLUSH (for backflushing documents):** A backflushing document is a document log based on "confirmations," which document mass-produced volumes. A typical characteristic of mass production is that the same product is repeatedly produced over a long period of time. The confirmation of production takes place once production is complete, as recorded by postings from the goods entrance and goods exit. Additionally, the production costs are determined and statistical data is updated. The corresponding data is identified and saved to several tables. PP_BKFLUSH archives the header data from table BLPK and the line item data from table BLPP. (Note that PP_BKFLUSH is only used in repetitive manufacturing.)

The objects listed in the sidebar above are some of the most important archiving objects in Logistics in that they are typically associated with the fastest-growing tables. Note that since many billing documents come into play, you will also need to account for any associated Financial Accounting (FI) and Controlling (CO) data objects and their corresponding archiving objects (e.g., FI_DOCUMNT and CO_ITEM).

As I mentioned at the outset of this article, identifying fast-growing tables and data objects for relocation is a relatively easy task. Once you complete this task, though, the ones that follow require a lot more finesse! These tasks are the ones I will help you with here. So, after you have identified the data objects you want to archive in the Project Initiation phase, it's time to start the Analysis phase, where for starters you will assemble your archiving project team and begin working with the affected user departments. *All* departments with a vested interest in the data you have earmarked for archiving must be consulted.

That said, let's now examine the six steps of the Analysis phase, from the perspective of our Logistics sales order process example.

Figure 2 The Responsibilities of the Core Team and the Extended Team



Step 1: Build the Project Team

The archiving objects listed in the sidebar on pages 36-37 relocate data from the various business components in Logistics — Sales and Distribution (SD), Materials Management (MM), Production Planning and Control (PP), and Logistics Execution (LE). Any relevant table entries from Financial Accounting (FI) and Controlling (CO) are archived too, since cost-relevant data is transferred directly from sales and production to finance. The majority of archiving objects you will work with, however, and the ones we will focus on here, belong to SD and LE.

To make the best use of each project team member's time, and to keep the number of project team members to a minimum, it is advisable to divide the group into two different teams (see **Figure 2**):

- The *core team* should be composed of the project leader, who is responsible for the main module (SD) and for interfacing with IT, and those responsible for the remaining modules (MM, PP, and LE), along with any relevant FI and CO components. The core team provides direction for the duration of the project, including during the Analysis, Design and Conception, and Going Live phases. Core team members are responsible for discussing the requirements with end users and representing the end users' interests.

> ✓ ***Tip***
>
> *Since the majority of the archiving objects come from SD, one high-level representative from sales and one system administrator should head the project, so that both the business and technical sides of the operation are considered.*

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

- The *extended team* should be composed of employees from the controlling, legal, customer service, and technical support departments, and, when it makes sense, external consultants. The extended team should be active mainly in the Analysis phase and, in some circumstances, in the Design and Conception phase.

The extended and core teams must also take into account any distributed or networked systems. For example:

- In an ALE scenario that uses multiple SAP systems (e.g., one central Financial Accounting system with multiple Logistics systems), it first should be determined whether or not predecessor and successor documents linked to the object targeted for archiving (more on this in Step 2) reside in the local system, and whether the data objects are replicated.

  Depending on the distribution scenario, the write programs of the corresponding archiving object must be modified in order to carry out business checks on associated data objects that reside in the remote system. Archiving object SD_VBAK, for example, offers a user exit to implement these checks. (More on write programs, business checks, and user exits in Step 3.)

- If a Business Information Warehouse (BW) system is in operation, the relevant data objects should be forwarded to the BW system before archiving.

As far as the storage of archive files is concerned, a strategy should be discussed with all those involved in any imaging projects, print lists, and so on, and all relevant users should be included in the project's extended team.

For additional pointers and tips for building a successful archiving project team, see the sidebar to the right.

The Keys to Building a Successful Archiving Project Team

Keep the number of people involved to a minimum. Include only as many as necessary. The core team should not include more than four to eight people. A similar number should be included in the extended team.

The **project leader** must:

- ☑ Evaluate and review the whole business process.
- ☑ Mediate between IT and other business departments.
- ☑ Evaluate dependencies between data objects.
- ☑ Establish a process-oriented view.
- ☑ Integrate business, legal, and technical requirements.

Project members of the **core team** and **extended team** must:

- ☑ Represent requirements from the business departments involved.
- ☑ Identify archivable data objects.
- ☑ Define access and search criteria.
- ☑ Define display functionality.
- ☑ Describe interdependencies between related process steps.
- ☑ Determine data residence times.
- ☑ Establish business checks that specify when a data object can be relocated.

Pitfalls When Building the Project Team

☑ **Inadequate involvement of those responsible for processes and components:** Often, the project group is composed of employees from the systems administration department whose mission in life is to guarantee system availability for the operation of an organization's business processes. An easy way to achieve this goal is to relocate lots of data objects from the online database. But this is not something that should be done indiscriminately! System administrators will need to turn to business specialists for direction and must receive adequate support when they do. The identification of data objects for archiving must be done in close cooperation with those who have a detailed knowledge of the company-specific business processes. Failure to include those responsible for the business processes will lead to the failure of the entire data archiving project.

☑ **Failure to involve *all* parties involved in the processes:** In many companies, business processes are spread over numerous organizational levels and across several national boundaries. It's a daunting task to identify the right people to serve on the project team, since the whole process is not usually fully documented. For instance, representatives from accounting and controlling should also be involved, as data is often required by those groups for tasks related to period and year-end closing. Mark my words, this is a monumental effort! Don't underestimate the time it will take to bring all relevant parties into the fold of your archive planning.

☑ **Involvement of external consultants:** The analysis and documentation of the relevant processes play a central role in the project and require expertise in the company's internal processes. Particularly in the Analysis phase, the best practices and experiences of external consultants — as derived from their various consulting projects — can help reduce the time needed for the process analysis. However, it is almost impossible to carry out the analysis with *only* external consultants, since in-house employees know the company-specific processes best. Complement external knowledge with internal knowledge.

Step 2: Evaluate the Process Context

It is crucial for archiving teams to understand that there is a high degree of linkage and interdependency among the Logistics data objects, and with any relevant Financial Accounting (FI) and Controlling (CO) data objects. Relocating one data object to an archive system can have an impact on an entire business process.

> ✓ ***Tip***
>
> *Archiving teams must also examine any interfaces to other processes and systems. Is the data distributed to other systems by means of ALE? Is a BW system in place to which the data is to be forwarded? In such cases, the project team's coordination requirements increase.*

The sales order process usually starts with an inquiry and ends with a customer's payment (see the sidebar on the next page for details on what constitutes a business process flow). The whole business process is mapped and controlled in the system by means of electronic documents (i.e., data objects) that are created automatically.[7] The data from the documents is copied implicitly from the previous document and updated accordingly — thus the system automatically generates a production order that extracts its data directly from the sales order. At the same time, the linkage between these documents is created in the form of previous and subsequent relationships, which are stored in a special relationship table called VBFA. In addition, relevant data from these documents might also be transferred to the Logistics Information System (LIS), where data can be aggregated and prepared for statistical evaluation. (More on this in a moment.)

[7] Automatic updating of data means that the same data does not have to be entered again and again, and that the relevant data is copied from the previous documents in the document chain.

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

What Makes Up a Business Process Flow?

Evaluating the business process context is the project team's primary task. But before the team can start evaluating the different elements of the process context, it first must understand the chain of processing steps, which is based on the use of common data structures.

A typical sales order business process proceeds as follows:

1. A customer inquiry triggers a quotation from sales. The inquiry and the quotation are both recorded in documents in the system. In this case, the required information is taken directly from the customer master and the material master data.

2. If the customer accepts the quotation, a sales order is created. The relevant data is copied from the quotation to the order. Additional data, such as prices and dispatch dates, are determined and added to the sales order. If the order is created without reference to a quotation or an inquiry, the relevant data is taken from the corresponding master data in the database. Obviously, the master data (material master, customer master, etc.) must not be archived as long as this data is required for transactions.

3. If the product is to be manufactured in-house, the next step triggers the timely completion of the finished product. A production order is created that determines how and when the end product is produced and with what resources. The user can navigate from the sales order to the production order and back by double-clicking the production order object number on the original sales order display screen (transaction VA03). (Note that this direct navigation capability is lost if the production order or the sales order is archived.) The project team must document this relationship and communicate it to end users so they are aware of this dependency when deciding what to archive. During production, the required material components are taken from the warehouse via goods exit.

   The delivery of the finished product to the warehouse is controlled via goods receipt. Both types of goods movement create material documents in the system and document the relevant transactions. The user can navigate from the display transaction of a production order to the relevant material documents and back — as long as they have not been archived.

(continued on next page)

All the data objects — including inquiries, quotations, sales orders, billing documents, deliveries, shipment documents, and shipment cost documents, along with the corresponding goods movement, FI, and CO documents — are implicitly included in the *document flow*, which is an SD tool for displaying the business process context from a sales point of view. In the document flow, related documents are displayed showing the relationships with previous and subsequent documents. The corresponding relationship entries are stored persistently in table VBFA, which is updated each time a new document is created. The document flow evaluates these entries and displays the documents hierarchically — i.e., according to the specified business process under consideration. The document flow is updated automatically at both the header and line item level. For each document in the document flow, the overall processing status is also displayed.

To display the document flow for a sales document, for example, start the sales order display transaction (VA03) and select *Environment → Display document flow* from the SAP standard menu.

(continued from previous page)

For repetitive manufacturing, remember that backflushing documents are used, in which case the user can navigate from the document log to the relevant material documents and back.

4. Once the goods have been produced and are ready for further processing in the warehouse, all activities relevant to shipping, such as commissioning and packing, are triggered and the product is released to the customer. These activities result in a deliveries data object that records all of these transactions. At this point, the system copies the relevant data from the sales order to the deliveries data object. When the product is delivered to the customer, a material document is created that documents the goods exit. The accompanying material document can be displayed in the document flow (an SD tool for displaying the process context of a business scenario from a sales perspective).
5. A shipment document is used to group deliveries with the same shipment conditions, such as haulage company, destination, itinerary, and means of transport. The shipment function is optional and can be configured in line with company-specific requirements in customizing for SD. The system uses copy routines to copy the data from the sales order to the deliveries data object, and from the deliveries data object to the shipment document.
6. The shipment cost document settles haulage company bills and forwards the data to accounting.
7. The billing document charges the customer for goods received. In this step, the costs incurred by

Figure 3 shows an example document flow for a simple sales order 6368.

As you can see, the system creates a deliveries data object (Delivery 80006080), then a shipment document (Shipment 1134), and finally posts a material document (GD goods issue 4900000031). Subsequently, a billing document is created (Invoice 90015531), which is passed on to accounting and appears in the document flow as Accounting document 100000028. You can also see the status for each document ("Completed," "Not cleared," etc.). The dots at the beginning of each line indicate the predecessor/successor relationships (e.g., Shipment 1134, WMS transfer order 462, GD goods issue 4900000031, and Invoice 90015531 are direct successors of Delivery 80006080).

You can only display the entire document flow if all the data objects involved have not been archived — that is, as long as there are still entries in table VBFA. If one of the data objects involved in the process is archived, the corresponding entries in table VBFA are written to the archive and deleted. No subsequent documents can be displayed, so the document flow is truncated.

In this case, you can use the Document Relationship Browser (DRB), which is able to display archived data in terms of the business process context. To enable this functionality, the accompanying archive information structures (infostructures)[8] of the Archive Information System (AS) must be configured according to the DRB fields.[9]

[8] Database tables that contain a pointer to the position of an archived data object in the archive file system.

[9] The 4.6C delivery of the DRB does not include all the data objects described here — i.e., production orders and shipment cost documents are not yet included. Consequently, documents linked as successors to production orders or shipment cost documents are not shown in the DRB. You can use the technical view of the AS to retrieve the object numbers of linked objects. With these numbers, you can search for production orders in either the archive (using the AS) or the online database.

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

the haulage company can be passed on to the customer and transferred to the billing document. When the billing document is created, the data is automatically forwarded to accounting for further processing.

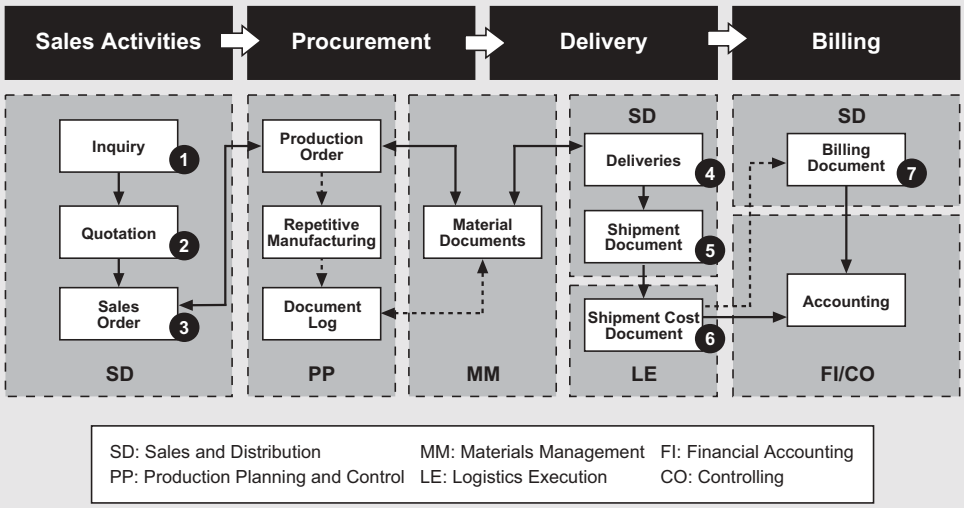
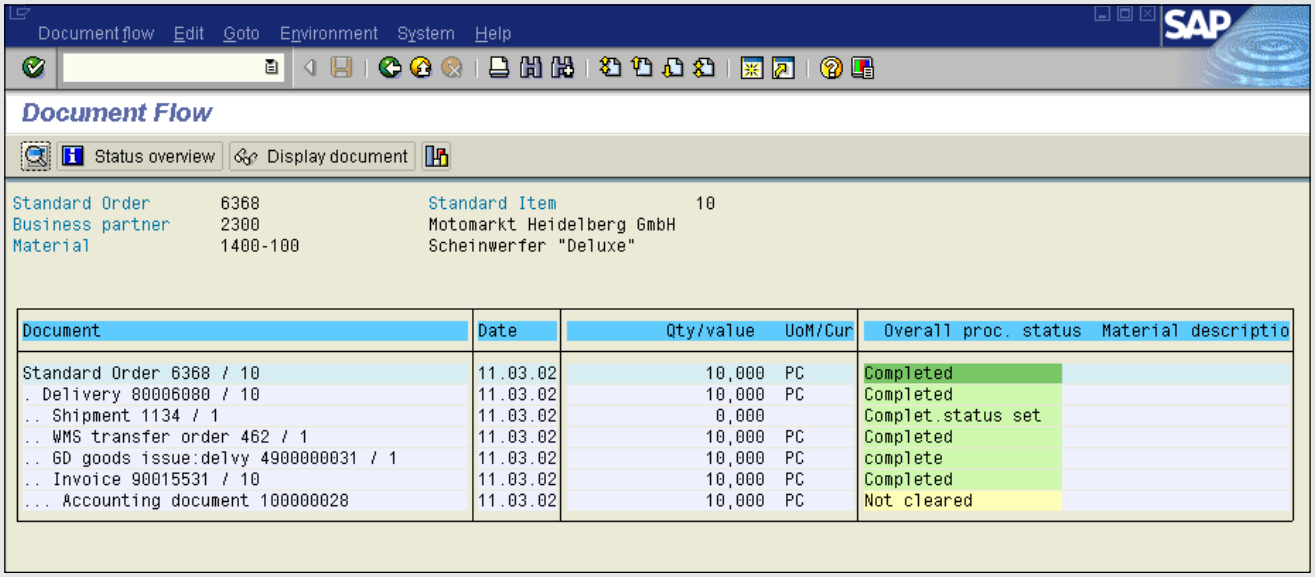


Figure 3 *An Example Sales Document Flow*



Document	Date	Qty/value	UoM/Cur	Overall proc. status	Material descriptio
Standard Order 6368 / 10	11.03.02	10,000	PC	Completed	
. Delivery 80006080 / 10	11.03.02	10,000	PC	Completed	
.. Shipment 1134 / 1	11.03.02	0,000		Complet.status set	
.. WMS transfer order 462 / 1	11.03.02	10,000	PC	Completed	
.. GD goods issue:delvy 4900000031 / 1	11.03.02	10,000	PC	complete	
.. Invoice 90015531 / 10	11.03.02	10,000	PC	Completed	
... Accounting document 100000028	11.03.02	10,000	PC	Not cleared	

Figure 4 *A Sample Process Flow Overview*

Archiving Object	Process Context	Description of Dependencies
SD_VBAK (for sales documents)	A sales document records customer requisitions for goods and services. It includes inquiries, quotations, and sales orders.	Integrated into SD document flow. Navigation from the sales document to the linked production order from transaction VA03 is not possible if the production order is archived.
SD_VTTK (for shipment documents)	A shipment document is a central element of shipment planning and completion. It includes: • Assignment of deliveries to shipments • Specification of shipment stages • Assignment of goods to handling units • Specification of shipment deadlines	Integrated into SD document flow.
SD_VFKK (for shipment cost documents)	A shipment cost document contains data related to shipment cost.	Integrated into SD document flow. Shipment cost documents are not displayed in the DRB as of 4.6C and can be accessed directly with the AS.
RV_LIKP (for deliveries)	Deliveries data objects record all shipping activities (picking, packing, goods issue, etc.).	Integrated into SD document flow.
SD_VBRK (for billing documents)	The term “billing documents” includes invoices, credit memos, and cancellation documents. It is created with a reference to the preceding document.	Integrated into SD document flow.

After analyzing the process flow, you should clarify which data objects are connected to each other within the system, so that you can evaluate how certain objects will be affected by archiving others. Toward this end, the core project team should create an overview, including which navigation options are possible, so that you can also evaluate how navigation dependencies will be affected. Such an overview can be provided in list form, like the example shown in **Figure 4**.

As I mentioned earlier, in addition to the relationships between the Logistics and FI/CO data objects, relevant data might also be transferred to the Logistics Information System (LIS), which aggregates data and prepares it for statistical evaluations.

Let’s take a quick look at the role of the LIS in the process context and some of the considerations you will need to keep in mind if it is in use.

The Role of the Logistics Information System (LIS)

The Logistics Information System (LIS) collects,

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

Figure 4 (continued)

Archiving Object	Process Context	Description of Dependencies
PP_ORDER (for production orders)	A production order specifies which material is to be produced, where it is to be produced, which operations are required, and on which date production is to take place. When creating a sales order, the system can automatically generate an assembly order. Here it concerns the automatic generation of a planned order or production order to cover the planned independent requirements.	Direct navigation from the production order to the sales document is not possible if the sales document is archived. Navigation to linked material documents is possible only if the material documents are still in the database. Production orders are not displayed in the DRB in 4.6C and can be accessed directly with the AS.
MM_MATBEL (for material documents)	Material documents support transactions resulting in a change in stock, including goods receipt, goods issue, stock transfer, and transfer posting.	Integration into SD document flow. Only possible to branch to material master (not depicted in the sidebar diagram on page 43), accounting document (also not shown in the sidebar diagram), production order, or deliveries if the data objects reside in the database.
PP_BKFLUSH (for backflushing documents)	A backflushing document is a document log for series production (i.e., repetitive manufacturing). It posts goods receipts for the finished products, goods issues for the components, and data on production costs.	Cross-referenced to material documents. Direct navigation to the display transaction of material documents.

summarizes, and evaluates data from the operational areas of purchasing, sales, production, and inventory.

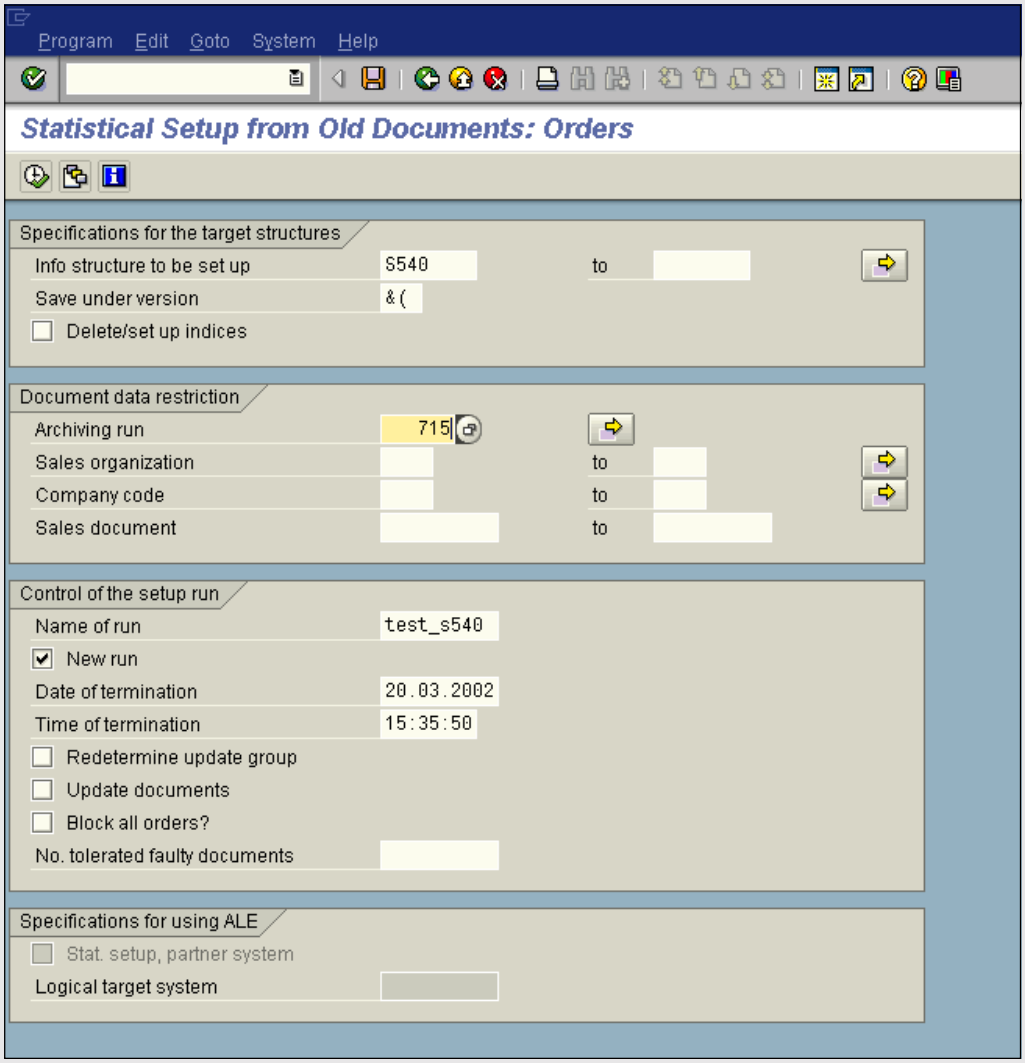
The LIS is based on infostructures that are created as transparent tables in the database (not to be confused with AS infostructures, which contain pointers to archived data objects in the archive file system). The system reads data for statistical evaluation from data objects and writes that data to the LIS infostructures at predefined times.[10] In this way, the relevant LIS infostructures are automatically updated when sales orders or production orders are created or changed.

SAP provides predefined LIS infostructures. Customers can also create their own infostructures in customizing and specify how this data is to be forwarded and updated in the LIS infostructures. Depending on how the updating rules are defined (daily, weekly, or monthly), these LIS infostructures can become very large.

Data from these infostructures is not deleted when the source data objects are deleted. Instead,

[10] The updating rules are defined in the Implementation Guide (IMG).

Figure 5 *Creating an LIS Infostructure for a Sales Order*



this data is retained in the system and can still be accessed using the LIS, so statistical analysis can be carried out even if the original source data objects have been archived.

Objects for archiving LIS infostructures are automatically generated within the LIS (transaction MCSX) and can be accessed in transaction SARA.

While re-creating infostructures is sometimes necessary — for example, when updating rules are changed — be sure to check whether you need old archived data objects to rebuild any of the infostructures. Only some LIS components, such as the Sales Information System, can process archived data, so it is important to examine the extent to which archived data is required as a basis for the infostructure. **Figure 5** shows the entry screen in the Sales Information System for setting up an LIS infostructure (S540) for a sales order. The infostructure's corresponding program can process archived sales documents (indicated by the field

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

"Archiving run" in the "Document data restriction" frame). Other LIS components, however, cannot, so it is a good idea to note these dependencies in your process context overview as well.

Pitfalls When Evaluating the Process Context

☑ **Process documentation is not (fully) available:** As the process analysis showed, data objects from several different components (e.g., SD, MM, LE, FI/CO) are relocated in a sales order process. Very few companies have any process documentation, and if they do, it is usually minimal. Process knowledge is usually distributed across numerous people and recorded in various media, and, as demonstrated in Figure 4, must be carefully collected during the Analysis phase. Although collecting this knowledge is a particularly time-consuming activity for the project team, it is an essential key to determining archivable data.

☑ **Overly large infostructures:** As a rule, it is not advisable to rebuild infostructures based on line items, since several entries will be written per data object. This can lead to very large infostructures in the case of documents with many line items.

Step 3: Define Business Requirements

In this step of the Analysis phase, you work with end users and component managers for each of the Logistics and FI/CO components to determine the most important business parameters. In each case, this involves setting residence times,[11] activating business checks in the write program,[12] and defining search and display criteria.

Let's first take a look at the details of setting residence times and activating business checks for sales documents, shipment and shipment cost documents, deliveries, and billing documents.

Setting Residence Times and Activating Business Checks: Sales Documents, Shipment and Shipment Cost Documents, Deliveries, and Billing Documents

The archiving of sales documents, shipment and shipment cost documents, deliveries, and billing documents is based on the same residence time concept. In this case, residence times are fixed according to organizational structures and document types that take the date of creation or last change as a reference point, in line with the settings made during customizing. Generally the residence time states how long a document must stay in the database until it can be transferred to an external archive file system and deleted in the database.

Let's illustrate the residence time concept with an example. Suppose a sales order was created on 01.03.2001. The residence time, which is maintained in the application-specific customizing transaction SARA, is set to 20 days, so at the earliest, the sales order can be archived on 21.03.2001 — if it also has the status "Completed" and all business checks are performed successfully. Matching the residence time is only one criterion that a document has to fulfill in order to be archivable. The system still carries out business checks on the document and any predecessor or successor objects. Getting a consensus on residence times from all those involved also ensures that data objects created in the future can be archived smoothly without further process analysis.

I'll use archiving object SD_VBAK as an example. SD_VBAK archives sales documents and is the most important archiving object in Logistics. Similar considerations apply to shipment and shipment cost documents, billing documents, and deliveries. For any of these objects, you will need to set residence times and activate business checks (except for shipment cost documents, which have no business checks):

[11] The length of time data stays resident in the online system before it is archived. This can be defined in the customizing settings.

[12] Note that not all archiving objects offer business checks — i.e., business checks are not available for shipment cost documents, production orders, material documents, and backflushing documents.

- **Residence times:** In customizing for sales documents, each sales organization can specify a residence time in days for each sales document type. The sales document types control the various business events in sales. For shipment and shipment cost documents, duration can also be specified for each organizational unit. The settings are made in application-specific customizing of transaction SARA. By default, the residence time is based on the creation date of the sales document. The residence time specifies the time period that must elapse between the creation date and the start date for the write program that triggers archiving. **Figure 6** shows some residence times for several sales documents.

- **Business checks:** Internally, the write program determines the immediately previous and immediately subsequent documents for each sales document and checks whether they have the status "Completed." This check is carried out for all selected sales documents and cannot be deactivated. The following business checks can be activated individually for each archiving session. The project team should discuss these checks with the user departments during the Analysis phase.[13] The selected sales documents can then be archived successfully once the selected checks have been carried out:

  - *Create archive* — If this flag is deactivated, the write program runs in test mode and no archive files are created.

  - *Deletion program in test mode* — If you do not want the archived data to be automatically deleted from the database after the archive files are created, activate this flag. This option is not available if "Create archive" has been deactivated.

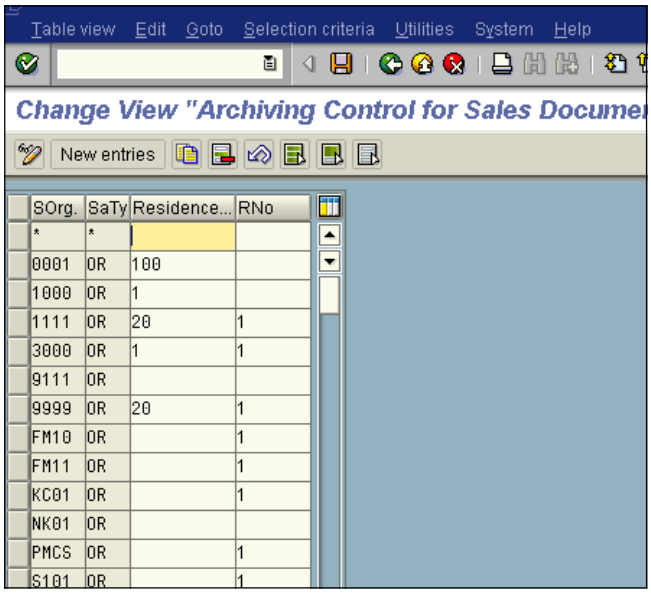
  - *Check purchase order* — If the product is being procured externally, rather than manufactured within the company, sales documents can be used to generate purchase orders or purchase requisitions that relate to the sales documents. If this check is activated, the system checks for the existence of linked purchase orders or purchase requisitions. If any are found in the database, the current sales document cannot be archived.

  - *Check FI document* — If this field is selected, the system checks whether the assigned billing document has been settled in Financial Accounting (FI).

  - *Check flow documents residence* — The whole document flow is evaluated and the creation or change date of the last document in the flow is used as the key date for the calculation of the residence time.

  - *Alternative DB access* — If this field is selected, the data is first placed in internal tables on the application server and then processed there to ease the burden on the online database server. If this field is not

Figure 6 Example Residence Times

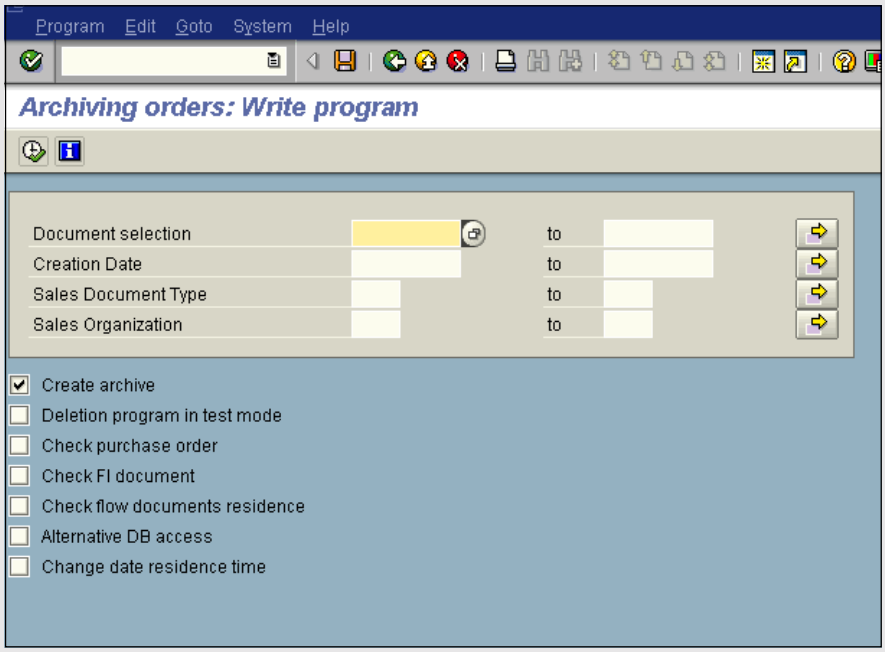


SOrg.	SaTy	Residence...	RNo
*	*		
0001	OR	100	
1000	OR	1	
1111	OR	20	1
3000	OR	1	1
9111	OR		
9999	OR	20	1
FM10	OR		1
FM11	OR		1
KC01	OR		1
NK01	OR		
PMCS	OR		1
S101	OR		1

[13] To enhance the business entry parameter selection in the write program, see SAP Note 314330, which explains how to make the parameters "Creation date of the sales document," "Sales document type," and "Sales organization" available.

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

Figure 7 *The Business Check Options in the Write Program of SD_VBAK*



selected, the data is obtained using the "open cursor technique," where a pointer is aimed at a "hit list" in the database and the data is requested from the database as "packages."[14]

- *Change date residence time* — This option checks the residence time against the last change date of the sales document or the last change date in the last subsequent document (instead of checking the residence time against the creation date).

Obviously, the more business checks you activate, the longer the runtime will last. You can activate these checks in the write program of archiving object SD_VBAK. **Figure 7** shows the entry screen with the selection parameters, which can be set within transaction SARA.

Additionally, user exits for customer-specific checks can be defined in archiving-object-specific customizing. The user exit concept is identical for sales orders, shipment and billing documents, and deliveries. It is triggered at the write program's runtime for each document that is to be archived. A form routine, which the customer must provide, is then called, in which additional customer-specific checks may be implemented.

After the business-relevant checks have been analyzed and discussed with end users, the retrieval functionality must be evaluated and defined. In the next section, I briefly explain the options available with 4.6C for defining search and display criteria for sales documents, deliveries, and billing documents.

[14] This technique is useful in Oracle databases in response to the long-runtime error message ORA1555, since Oracle tries by means of a "consistent read" to reconstruct the condition at the time of opening the cursor.

> ✓ ***Note!***
>
> *If the previous and subsequent documents for the process context are contained in other systems, these documents are not checked by the system. The document flow and the process context are only checked locally. If, for example, the accounting document is in a different system (a central Financial Accounting system or a decentralized Logistics system), and it is necessary to evaluate whether the accounting document is cleared, a customer-specific user exit must be programmed. For this, it is necessary to program and call an RFC-enabled function module that carries out the checks on accounting documents in the target system and returns the results. Knowledge of the relevant data model is required.*

Defining Search and Display Criteria: Sales Documents, Deliveries, and Billing Documents

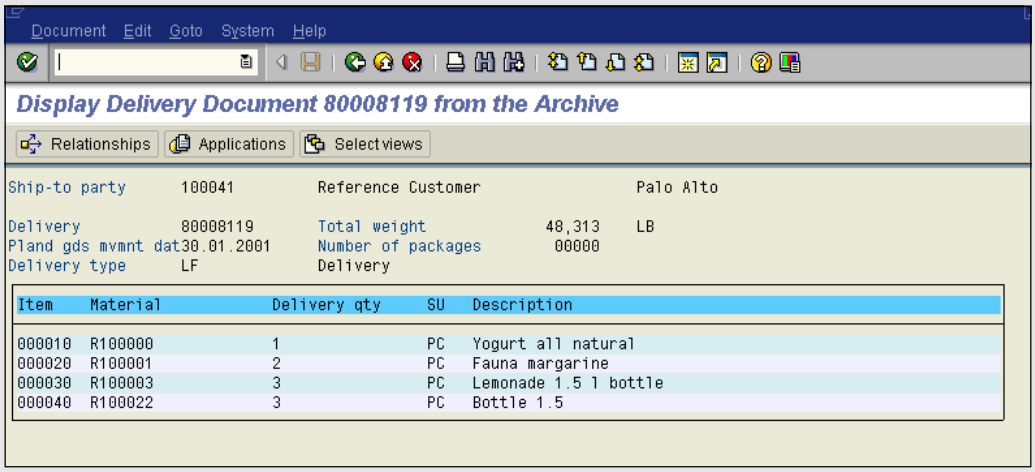
For the sales document, billing document, and deliveries archiving objects, there are sequential read programs that use the document number as search criteria. For each of these three archiving objects, there are two standard field catalogs[15]: one field catalog contains fields for the document header and the other contains fields for the line items. This enables direct access via search criteria that can be parameterized using the fields in the corresponding infostructure. The DRB also provides a combined feature that searches across infostructures and data objects in the database, although the search fields in the DRB cannot be modified.

Several business views are available for sales documents, billing documents,[16] and deliveries. The data is mostly displayed in list form, and only the most important data is shown at the header and line item level. For sales documents, billing documents, and deliveries it is also possible to access linked original documents in an external archive system. **Figure 8** shows an example of an archived deliveries object displayed in list form.

[15] A field catalog is a superset of fields taken from the data model, which can be used to configure AS infostructures.

[16] In R/3 Enterprise, SAP uses the original display transaction to display archived billing documents within the DRB (see SAP Note 518331).

Figure 8 The List Display of an Example Archived Deliveries Object



 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

Defining search and display criteria is a little bit different for shipment and shipment cost documents. Let's take a look at the considerations in this case.

Defining Search and Display Criteria: Shipment and Shipment Cost Documents

There are no sequential read programs for shipment and shipment cost documents. These objects are accessed using the AS. As I mentioned, SAP provides two field catalogs for each object type: one with fields for the document header and one with fields for the line items. The search criteria can then be set specifically according to the field catalogs. Shipment documents are implicitly considered when evaluating the business process context in the DRB, and cannot be selected specifically as a start object in the initial screen of transaction ALO1. Shipment cost documents are not yet considered by the DRB.

For shipment and shipment cost documents, a technical view is available in the AS. This should be communicated clearly to end users during the Analysis phase. In some cases, this technically oriented display functionality may require a longer residence time for shipment documents.

Setting Residence Times and Defining Search and Display Criteria: Material Documents

In application-specific customizing, under the residence time settings, it is possible to specify the minimum time that must elapse before material documents can be archived. The residence time can be set per document according to plant and goods movement type. The residence time always refers to the posting date in the material header.

In the write program (**Figure 9**) the criteria for archiving material documents can be more precisely restricted according to document number, material document year, plant, posting date, and transaction/event type. Since material document numbers are year-dependent, and since a material document number can occur several times in various years, it is a good idea to select an interval for the posting date.

Figure 9 Restrict Search Criteria for Material Documents with the Write Program

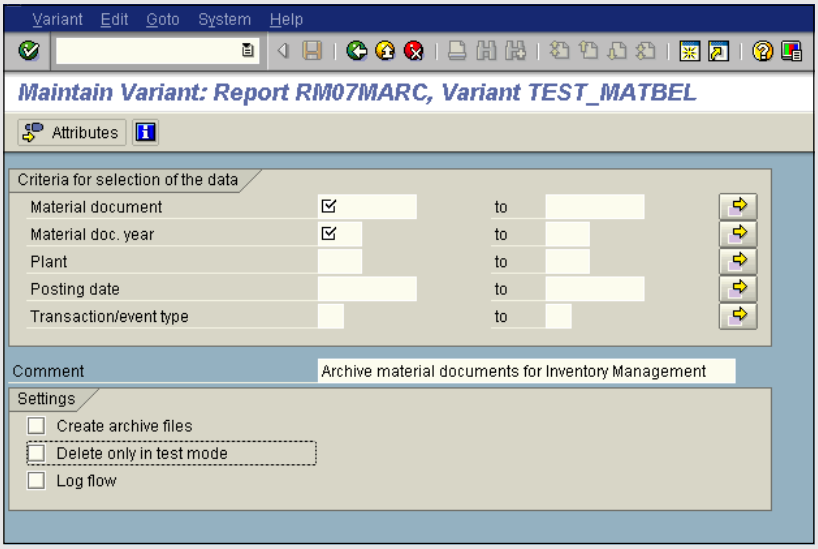
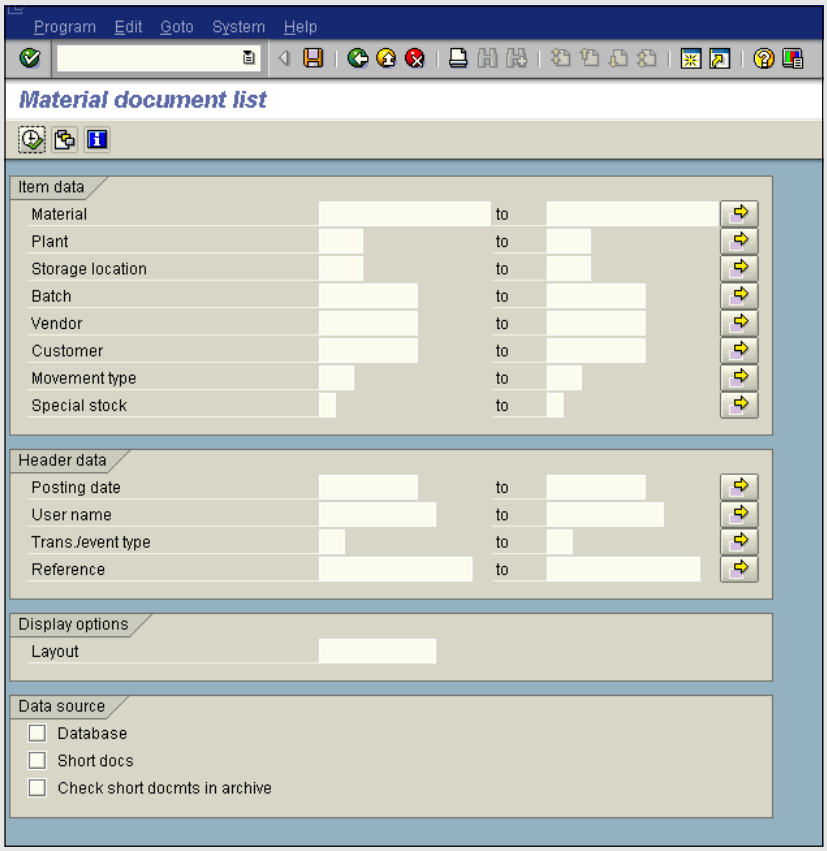


Figure 10 Select Criteria in Transaction MB51 to Display Material Documents



Archived material documents can be read using both sequential access and direct access. There are currently two index solutions for providing direct access to archived material documents:

1. An index and a short document can be created using MM-specific indexing programs. Using an index build program that can be started from transaction SARA, short documents are created in table MARI, while the archive file name (including the path name and the position of the material document within the archive file) is created in table MKPF-ARIDX. These tables must be built before it is possible to access archived material documents using the material document list (transaction MB51). Entries in these tables that are older than a specific reference date can be deleted using an index delete program.

2. Material documents are connected to the AS via the field catalogs SAP_MM_MATBEL01 and SAP_MM_MATBEL02.

   It is also possible to use the DRB to access archived material documents and data objects that belong to the same business process. A

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

prerequisite for this capability is that the archive infostructures in the AS must be filled.[17]

In addition to the technical view in the AS, the following options are available for displaying archived material documents:

- A sequential read program can be run from transaction SARA or MBAL, which displays the archived data in list format.
- Individual material documents can be viewed from the AS in the original display transaction MB03. (In future releases, MB03 will also be available in the DRB.) Consequently, it is possible to branch directly from the material document to the corresponding material master or accounting document, as long as these are still available in the database.
- A business view that is integrated into the AS and DRB can display the material documents in list form and also enables linked documents to be displayed from an external document system.
- Transaction MB51 can be used to display material documents in list form by selecting the desired criteria, which are shown in **Figure 10**.

Production orders follow a slightly different model for residence times. In the next section, I explain how this works, and also look at the considerations involved in defining search and display criteria for production orders.

Setting Residence Times and Defining Search and Display Criteria: Production Orders

For archiving, production orders use a combination of two deletion indicators and two residence times.

First, the order is prepared for archiving via a delete flag. A deletion flag can be set if certain business conditions have been fulfilled. If the order is cost-relevant, the order must first be settled — that is, the balance must be "0." (For more information on the deletion flag, refer to the online documentation at **http://help.sap.com**.)

The deletion flag can be set manually in the production order or as a background job in the preprocessing program. Once the first residence time has elapsed, the actual deletion indicator can be set. The second residence time is the time between the activation of the deletion indicator and the time from which the order can be archived and deleted from the database.

If the duration of the first residence time is "0," activating the deletion flag simultaneously triggers activation of the deletion indicator. Both residence times can be specified for individual production orders in customizing (transaction OPJH).

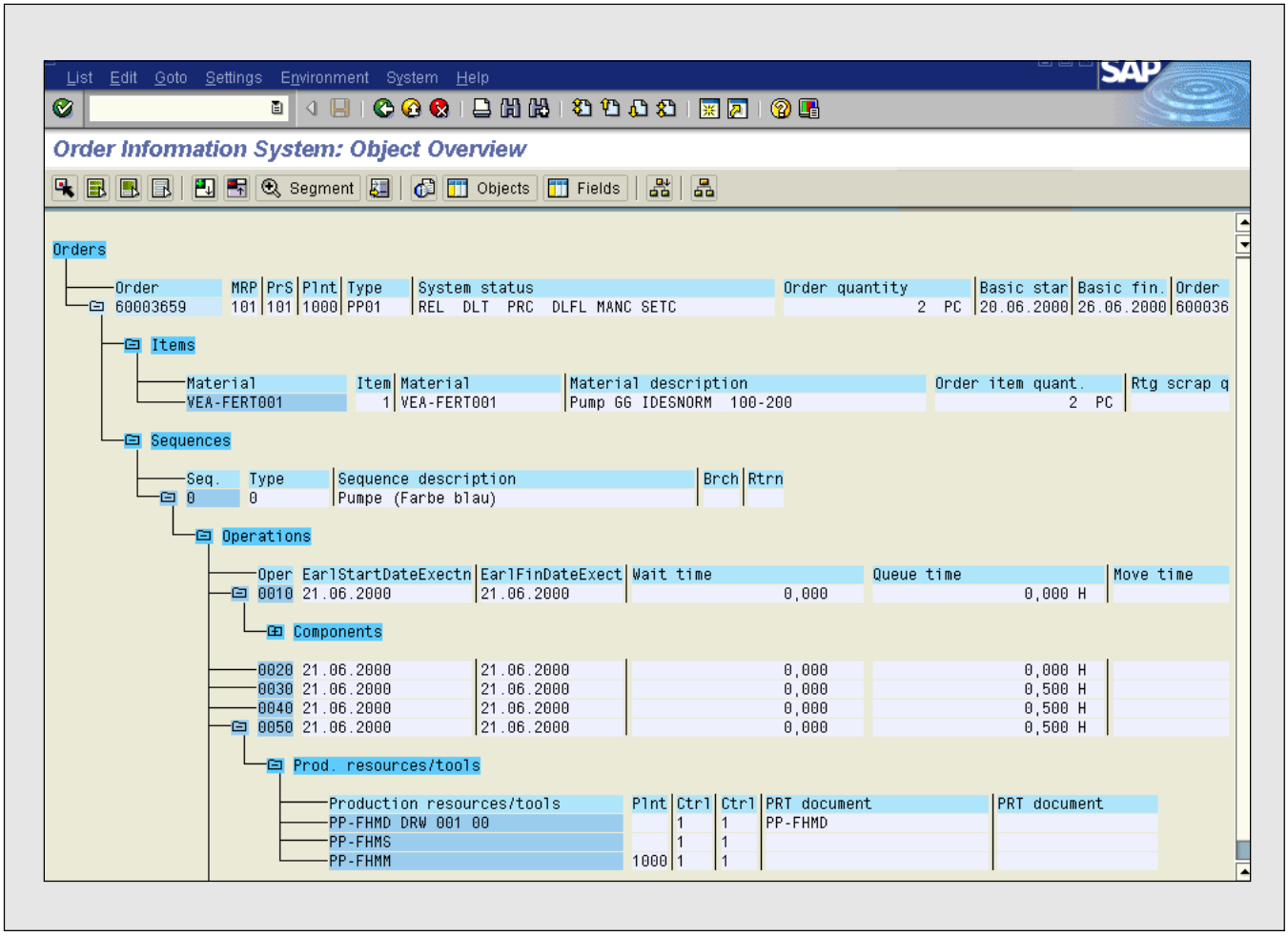
Production orders can be archived selectively using the following input parameters: order number, order type, plant, material number, MRP controller, production scheduler, and basic start date. These parameters are implemented in the write program. The project team must coordinate these input parameters with the production departments, including the definition of the residence times. As a rule, orders can only be archived if the deletion indicator is set and if the orders have fulfilled the second residence time.

Production orders are connected to the AS via field catalog SAP_PP_ORDER001. Furthermore, a sequential read program is available from transaction SARA. In Release 4.6C, production orders are not yet connected to the DRB.

In addition to the technical view in the AS, production orders are displayed using display programs in the Order Information System. Data is read from the archive and sent directly to the Order Information System, which then formats the data for display. The display includes the materials and

[17] In the future, SAP hopes to supply only the AS solution, so there is a standard procedure. There are currently two solutions for historical reasons. The index solution by means of tables MARI and MKPF-ARIDX was originally available before the AS solution.

Figure 11 Displaying Archived Production Orders Using the Order Information System



production tools used, and the completed operations (see **Figure 11**).

Setting Residence Times and Defining Search and Display Criteria: Backflushing Documents

While there is no connection to the AS available for backflushing documents during mass production, there is a sequential read program that can be called from transaction SARA in order to display data from backflushing documents in list form.

Note that there is no residence time concept for backflushing documents.

Pitfalls When Defining Business Requirements

☑ **Incorrect results from business checks:** It is often forgotten that the write program includes business checks that evaluate the status of subsequent documents. These business checks are only carried out in the local system. If the subsequent documents are in another system, these checks will not give the correct results. To remedy this, it is necessary to modify the write program logic after analyzing the system distribution scenario.

☑ **Conflicts between search functionality and storage capacity:** During discussions on infostructures, the project team should aim to find a

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

compromise between acceptable business search functionality and storage capacity in the database. Such a compromise should satisfy the requirements of both end users and system administrators. Moreover, the project team members should, as early as possible, become acquainted with the characteristics and functionality of the archiving objects involved so that they can make recommendations to the user departments and suggest possible alternatives.

☑ **Failure to consider enhancements to SAP tables:** Nearly all archiving objects also take into account extensions (includes, appends, etc.) of SAP tables in their write programs — that is, the contents of fields added by customers are written into the archive files and deleted along with the contents of the SAP table. In any case, the business display programs must be enhanced to display the contents of these fields.

Step 4: Define Legal and Auditing Requirements

The legal and auditing requirements for archived data vary from country to country, and from archiving object to archiving object, and are highly dependent on a company's particular situation, so I won't go into great detail on this subject.

There are two tools available from SAP to help you meet auditing requirements: the Data Retention Tool (DART) and the Audit Information System (AIS).

The US Internal Revenue Service (IRS) was involved at an early stage in the conception of DART to ensure that their needs were addressed. DART is able to create data extracts from online data *and* archived data. These data extracts are created outside the database so they can be evaluated using external tools. For the Logistics process, accounting-relevant material, sales documents, deliveries, and billing documents can be included in DART data extracts.

In Europe, the AIS is the preferred auditing tool, though the DART tool is gaining in popularity because it fits closer with the needs of tax authorities. Unlike DART, the AIS does not create a separate data extract. Instead, it prepares a hierarchically structured collection of reports that operate largely on data objects in the database. A few reports can, however, also evaluate archived data.

Beyond auditing requirements, further legal considerations in the Logistics process are product liability and product guarantees. Product liability means that the manufacturer is liable for damage that results from errors in the product. Product guarantees ensure that the manufacturer is liable for defects in the product itself. Both are, in various forms, established statutory requirements in every country. To meet these requirements, the availability of the relevant product information must, in some cases, be ensured for the duration of the product's lifecycle. It must also be possible to trace whether any damages resulted from production or construction defects, and, if necessary, for the manufacturer to take proactive corrective action. Frequent product recalls issued in recent years by the automobile industry illustrate this requirement clearly.

When configuring the AS to meet the archiving requirements imposed by product guarantees and product liability, answers to the following questions are crucial:

- What components, materials, and raw materials were used?
- What goods movements took place?
- Who were the suppliers?
- What operations took place during production?
- What was supplied?
- How were shipments organized? What route was taken?
- Who received the product?

Pitfalls When Defining Legal and Auditing Requirements

☑ **Non-compliance with country-specific regulations:** Legal requirements vary enormously from country to country, so be sure to check with subsidiaries and/or auditing companies in each affected country. Accuracy is crucial.

☑ **Loss of performance:** If DART is in use, it is advisable, for performance reasons, to archive data *after* the data extracts are created to avoid reloading data objects temporarily.

☑ **Inability to quickly access needed information:** In the context of product guarantees and liability, it is advisable to identify the relevant infostructures in advance and to maintain them in the system so that they can be activated swiftly if need be. The DRB can also be used to track the history of a process context for individual business incidents.

Step 5: Define Technical Requirements

Technical requirements include the access to archive files from all SAP application servers. Additionally, a business concept must be created that ensures the backup of newly created archive files. Since changes are not usually made to archive files, it is generally sufficient to make a single backup. In this section, I point out a few considerations regarding the choice of tertiary storage media (these are of interest primarily to IT project team members):

- As far as access times are concerned, there are no significant differences between the three storage options: file system, hierarchical storage management (HSM) system, or content management system (CMS).
- An HSM system is beneficial if the volumes of data involved are of terabyte proportions and if, in addition to archive files, further data volumes must be administered centrally. An ArchiveLink connection is not required.
- If a CMS is used for the administration and storage of documents (such as scanned-image documents, outgoing documents within the SAP system, print lists, etc.), it can also be used for the storage of archive files. In the case of Logistics, the CMS may be ideal for storage of the many shipping and billing papers that are involved. Invoices are created, printed, and sent to customers. In shipment processing, shipping papers are created and printed. The storage of large quantities of these documents is usually only possible outside the database in an external CMS. In some exceptional cases, these documents can also be maintained in the file system. If a CMS is used for storing the archive files, communication with the Archive Development Kit (ADK) is supplied by the ArchiveLink interface.

Step 6: Identify the Archivable Data

After evaluating the business process context and defining the requirements, the next and final step of the Analysis phase is to identify the data to be archived.

Toward that end, you should start by roughly classifying the data objects according to organizational criteria. Subsequently you can start to identify the concrete document numbers in accordance with the selection criteria of the write programs.

For example, let's say we want to identify sales documents for archiving. The most important organizational elements in sales are company code and sales organization, so we should include these in our classification criteria. We will also want to classify the documents according to year of creation (to determine how old the document is) and order type (to specify whether it is an inquiry, a quotation, a standard order, a repair order, etc.).

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

You can further classify sales documents by document number, and by the end user associated with the particular business process — it's a good idea to record this information so that you can check with users on how often certain documents are still accessed. Generally you will find large differences here — requests for a quotation are accessed less frequently than sales orders, for example, so they can be archived much earlier than sales orders.

Your classification criteria might look something like this in the end:

Criteria	Details
Company code	…
Sales organization	…
Order type	…
Year of creation	…
Document number	…
Business process	…
Created by	…
Approved by	…

The criteria you decide to use will depend on the particular data object you want to archive and your own unique environment. Regardless, you should have a well-thought-out, analogous plan for all archiving objects.

Next Step: The Design and Conception Phase

During this phase, the objective is to convert the requirements defined in the Analysis phase into a homogeneous archiving concept, beginning with specifying the archiving sequence. Here, I will provide some detail on this first step of the next phase.

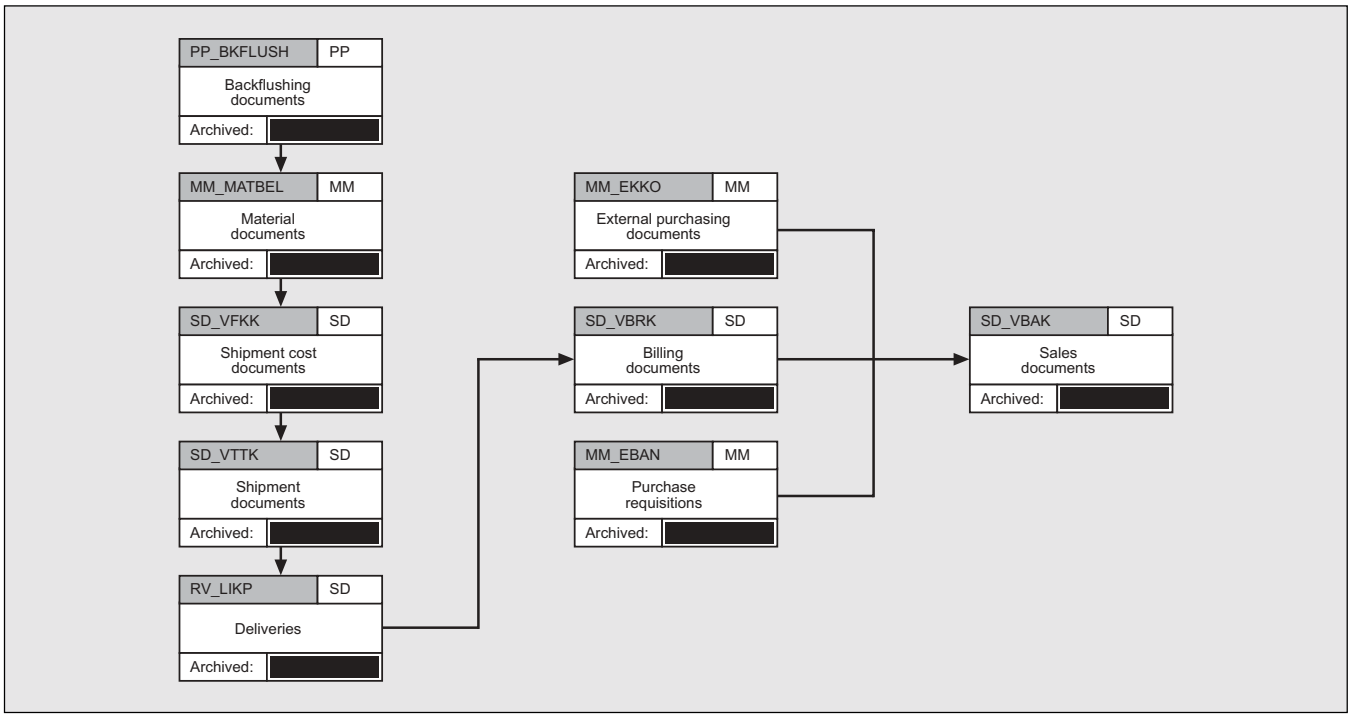
During the process flow (which was described earlier in the sidebar on pages 41-43), numerous data objects are created that characterize and document the individual process steps. These data objects are linked to each other in a chain. When these documents are archived, the status of the linked documents is checked. It is therefore essential that the data objects remain in the database at this stage. This results in a special *archiving sequence*, which is explained below in more detail for our central Logistics process.

The archiving object PP_BKFLUSH also writes data from tables MKPF and MSEG to the archive object that identifies material documents (MM_MATBEL). Backflushing documents from mass production should therefore be archived before the material documents, since the corresponding table entries are deleted from the database when the material documents are archived.

When shipment cost documents are archived by archiving object SD_VFKK, the system implicitly checks whether the shipment document to which the shipment cost document refers has been completely settled. Accompanying shipment documents must therefore still be available in the database when this check is being made — that is, shipment cost documents should be archived before shipment documents.

When shipment documents are archived, the planning status of the related deliveries is checked. Shipment documents should therefore be archived (SD_VTTK) before the accompanying deliveries (RV_LIKP). In principle, sales documents, deliveries, and billing documents can be archived independently from other data objects. However, in many cases accounting also requires sales documents and invoices to better control any credit and debit memos. After archiving deliveries, the following sequence should therefore be adhered to: FI documents should be archived first using FI_DOCUMNT, followed by billing documents using SD_VBRK (along with any purchasing documents), and finally the sales documents using SD_VBAK. The archiving

Figure 12 *Archiving Sequence*



sequence is shown schematically in **Figure 12**. (Note that production orders are omitted here because they have no direct dependencies to other objects.)

Helpful Hints

☑ A common misconception about data archiving is that it will always bring a significant improvement in response times. This is not always the case. A significant improvement in response times occurs only when the following points are true:

- Response time behavior is critical when accessing data via an optimal database index.
- The response time of the database is the main reason for overall poor response times.
- The underlying tables are very big.

☑ Data archiving *does* typically *prevent* drastic performance degradation in an SAP system.

☑ Be sure to allow sufficient time for the Analysis phase. During the Analysis phase, the business context and the processes that relate to the data objects for archiving must be *thoroughly* researched. Oversights and mistakes that are made at this stage, such as the removal of data that is still required for operational business, can rarely be rectified at a later date.

☑ When going live with a new component or process, establish database volume content guidelines. This will help ensure that as data volumes rise, you have ample time to execute your archiving project.

☑ You will find it particularly useful to have a standardized initial form for archiving, where the individual departments give their written agreement. It should contain an accurate definition of the classification criteria that identify the data objects to be archived, like the simple example shown in Step 6.

☑ The user departments must decide which search criteria are necessary and what display and process context display functions are necessary. This requires a clear strategy during the evaluation with end users. Unfortunately, communication with end users is often a point of weakness in an archiving project — questions designed to establish end user requirements are

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.

sometimes poorly formulated. Testing archiving during the Analysis phase using customer-specific infostructures is extremely helpful for determining users' needs. The project team should meet with end users to discuss the search functions available in the standard business views and the possibilities for evaluating the process context. In these meetings, the project team should emphasize to end users the effects of relocating and accessing data objects on the current business processes. Subsequently, the requirements should be recorded in a requirements log and sent to all those involved.

☑ Customer-specific tables are often used to store additional data associated with data objects. In such cases, the corresponding table entries must also be written to the archive file, requiring enhancements to the write, delete, and display programs. Unfortunately, these tables are frequently neglected when it comes to archiving. (Often, the archiving team does not even know of the existence of these tables.) It is a good idea for the IT specialists on the archiving team to draw up a list of all customer-specific tables before meeting with user departments, and note the components and processes to which the tables are assigned. Next, the project team must communicate with end users and agree on whether access to these table entries should be required after archiving, and if so, how.[18]

☑ After each archiving session, the users involved should be informed which data was archived and how to access those data objects in the future. This information should be recorded centrally and in writing.

☑ No data objects should be archived that belong to current business processes or will be required for year-end and month-end closing. To avoid this, the residence times should be set adequately high and the archiving project should be coordinated with representatives from accounting and controlling.

☑ Data archiving cannot be used to delete company-code-dependent data when companies are sold or restructured, so you cannot delete all data associated with a specific company code in the database (and then delete the created archive files).

☑ The DRB offers hard-coded search fields for precisely delimiting the initial object in a business process. If archived data objects are part of the search, the corresponding fields must also be available in the corresponding AS infostructures. For sales document infostructures, the DRB allows you to use the material type used in the process as a search criterion, but since the material is saved at the line item level, the infostructures can become quite large (several entries might be created for each sales document). It is essential to discuss which search fields are absolutely necessary with end users. If it is possible to search without referring to line item data, field catalog SAP_SD_VBAK_001 should be used for the configuration of the infostructure. Otherwise SAP_SD_VBAK_002 must be used.[19]

Conclusion

The first priority of an archiving team is to understand the business, legal, and auditing context of the data targeted for archiving. A thorough understanding of the business context to which the data objects relate is essential. It is not enough to simply determine viable data objects and their associated archiving objects. You must ensure that you have properly accounted for:

- The archiving objects' linked data objects
- All the points in the business process where the data objects are accessed
- All the users who access the data objects
- Customer-specific tables that must be archived
- Customer-specific enhancements to SAP tables
- Any current accessing of "old" data objects (including service and support requirements)

[18] If customer-specific tables are to be written to the archive, these entries are automatically displayed in the technical view in the AS; the same applies to enhancements to SAP tables.

[19] For more details on configuring the DRB and AS, see the article "Mastering the Administrative and Development Tasks Required to Put Archived Data in Easy Reach of Every User" (July/August 2001).

A Preview of What's to Come in the ADK and the DRB

Here's a preview of upcoming enhancements to the Archive Development Kit (ADK) and the Document Relationship Browser (DRB):

- ☑ To seamlessly integrate data archiving into the production system, system administrators will be able to interrupt an archiving session during the write phase and continue it later, enabling administrators to react, during archiving, to specific time constraints or storage space shortages. They can continue and complete interrupted sessions when there is more time or more storage space.
- ☑ When writing, deleting, or reloading, statistical data on each archiving session (such as the storage space that has been freed in the database by deletions or the number of deleted data objects) is automatically generated and persistently stored in the database. Data archiving administrators can analyze these figures so that they can better plan future archiving projects and request the necessary resources. The ADK collects the statistics.
- ☑ The most important recent development in the DRB is the ability to access and display data objects that are in remote systems. In the future, this will enable cross-system business processes, such as those in ALE scenarios, to be evaluated and displayed.
- ☑ Further data objects are to be integrated into the DRB, such as production orders. A personalization concept will control which object types an end user is entitled to see in the hit list.

- Search fields that need to access archived data
- The display functionality required by users
- Adequate residence times and business checks
- The selection criteria for archiving according to selection fields in the write program
- All relevant legal and auditing requirements that apply to each individual data object

Archiving Logistics objects makes considerable demands on those tasked with analyzing the process documentation. Internal company knowledge of the individual business steps and the people involved is critical. All relevant user departments must be involved, and their requirements for accessing and displaying archived data must be considered carefully. This requires a cautious and diligent approach when determining residence times, business checks, selection criteria, and requirements for the display programs. To guarantee a successful project, a thorough knowledge of the individual archiving objects and the Archive Information System is necessary. Toward this end, the project team must take an interdisciplinary approach and ensure that it includes representatives from both the business perspective and the IT side of operations.

Dr. Rolf Gersbacher joined SAP in 1994, where he was engaged in the implementation of business process reengineering projects using workflow technology. Since 1998 he has been a member of SAP's Performance, Benchmark and Data Archiving (PBA) department, where he is responsible for conducting extensive analysis and has co-written projects in the area of data archiving. Together with Helmut Stefani, Rolf developed the best practices guide "Managing SAP Data Archiving Projects," which has become the vade mecum *for data archiving issues. He can be reached at rolf.gersbacher@sap.com.*

 ©2003 SAP Professional Journal. Reproduction prohibited. All rights reserved.