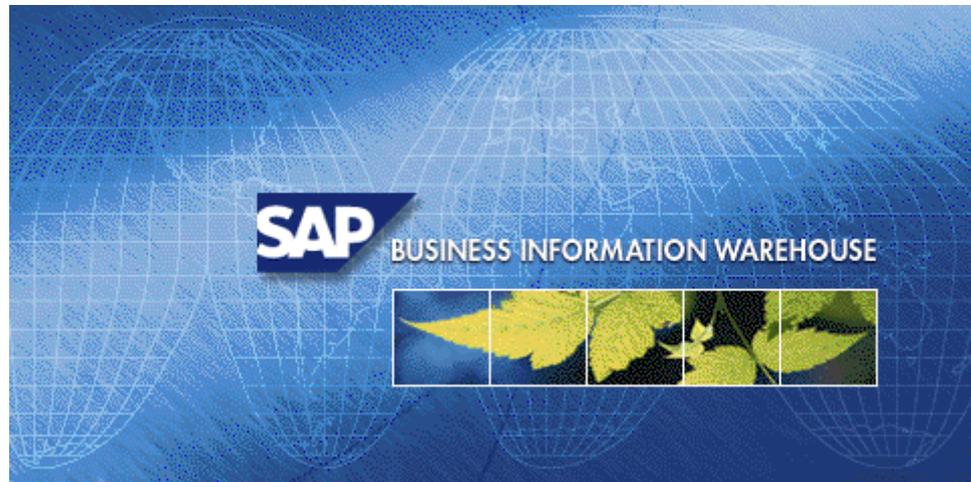


Staging Scenarios

Business Information Warehouse 2.0B



Version 1.0

Last Changed: 25.09.2000

SAP AG assumes no responsibility for errors or omissions in these materials.

These materials are provided "as is" without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement.

SAP shall not be liable for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials.

SAP does not warrant the accuracy or completeness of the information, text, graphics, links or other items contained within these materials. SAP has no control over the information that you may access through the use of hot links contained in these materials and does not endorse your use of third party web pages nor provide any warranty whatsoever relating to third party web pages.



TABLE OF CONTENTS

1	INTRODUCTION.....	1
2	DEFINITION.....	1
3	DATA STORAGE FEATURES	1
3.1	Persistent Staging Area	1
3.2	ODS Objects	2
3.3	InfoCubes.....	2
3.4	SAP RemoteCubes.....	2
3.5	BAPI RemoteCubes	3
3.6	MultiCubes	3
4	STAGING SCENARIOS WITH PERSISTENT DATA STORAGE	4
4.1	Standard Scenario	4
4.2	Staging with ODS Objects without InfoCubes	4
4.3	Staging without ODS Objects with InfoCubes	4
4.4	Further Processing from InfoCube to InfoCube.....	5
4.5	ODS Server and BW Server.....	5
5	STAGING SCENARIOS WITHOUT PERSISTENT DATA STORAGE.....	6
5.1	Staging with SAP RemoteCubes.....	6
5.2	Data Mart Scenarios with SAP RemoteCubes	7

1 Introduction

This document describes staging scenarios that you can carry out with BW 2.0B. The elements considered are the location of the data, in which structures, and how it gets into the reporting-compatible BW structures. The special features of each scenario then determine which scenario is most appropriate for use.

2 Definition

Staging: Process of data retrieval in a Data Warehouse.
The following differentiates between staging scenarios with persistent data storages and staging scenarios without persistent data storages. Persistent means that data, which is loaded into the BW System from the source system, is saved by using a transaction. Not persistent means that the data is always recreated and only lasts for the length of transaction in the BW System.

The extraction methods used are BW Content Extractors, customer-defined extractors (for COPA, LIS) or generic extractors. With the DB link, you can extract the data directly from the database level and therefore also extract it from non-SAP Systems. See also [Data Upload via a Database Link](#).

3 Data Storage Features

This unit describes the different data storages and their tasks in the staging process.

3.1 Persistent Staging Area

This Persistent Staging Area (PSA) is the initial storage of requested transaction data, master data attributes and texts from different source systems within the Business Information Warehouse.

The request data is stored unmodified in the form of the transfer structure in transparent, relational database tables and could therefore contain errors if it already contained errors in the source system. When you load flat files, the data is not entirely untouched, since it may have been adjusted by conversion routines (for example, the data format 31.12.1999 becomes 19991231, to ensure the data is consistent). The requests can now be checked for significance, sequence, and completeness.

At the same time, the PSA delivers the backup status for the ODS (until the entire staging process is activated), and improves the loading performance by detaching the loading process from the subsequent processing in BW. This does not slow down the operative system if errors only occur during the subsequent processing of the data.

If you update the data directly in the InfoCube, the length of the data storage in the PSA is mid-term, since it can still be used for reconstruction. On the other hand, if you update the data into ODS objects, it is only stored on a short-term basis.

3.2 ODS Objects

An ODS object stores consolidated and cleansed transaction data on document level. It consists of a consolidated dataset made up of one or more InfoSources.

You can report on this dataset with a BEx Query or an InfoSet Query. An ODS object contains key fields (for example, document number and item) as well as data fields that can also contain character fields (for example, order status, customer). The data of an ODS object can be updated with a delta update into InfoCubes and/or other ODS objects in the same system or other systems.

Contrary to multi-dimensional data storages with InfoCubes, ODS object data is stored in transparent, flat database tables. How detailed they are depends on how detailed the PSA is. This type of storage also applies to line items, which often need to be overwritten.

With the help of the change log, you can also use ODS objects for making DataSources delta-compatible. Likewise, you can bring data together from different processes, for example, order and delivery.

Contrary to the PSA, you can store data long-term in a cleansed and transformed form.

See also [ODS White Paper, Unit 4](#)

3.3 InfoCubes

An InfoCube (from the view of reporting) is a self-contained dataset, for example, of a business-oriented area, and is the central object upon which reports and analyses are based. The dataset can be reported on with a BEx Query.

An InfoCube can be:

- A BasisCube, presenting a set of relational tables in a star schema,
- A RemoteCube, with transaction data that is not managed in the Business Information Warehouse, but externally, or
- A MultiCube, presenting a view of data from several BasicCubes or RemoteCubes.

InfoCubes support the aggregated storage of loaded and consolidated data. They are made from a read-optimized data handling in fact tables and dimension tables (star schema), to enable performing, multi-dimensional reporting (OLAP).

In contrast to ODS objects, the data fields in the BasicCube fact table cannot contain fields of the type CHAR as key figures. Fields of the type CHAR are interpreted as keys. Another difference to ODS objects is that the data here is not volatile.

3.4 SAP RemoteCubes

An SAP RemoteCube is a specific RemoteCube that allows queries to be defined with direct access to transaction data in other SAP systems.

In contrast to BAPI RemoteCubes, you do not need to program interfaces in the source system. You use the same extractors to select data in the source system, as you use for replicating data into BW. When you execute a query, every navigation step sends a request to the extractors of the assigned source systems. Furthermore, the characteristic selection is transformed together with the selection conditions for these characteristics according to the transfer rules for the transfer structure fields. They are passed on to the extractor in this form. The delivered data records go through the transfer rules in BW and are then filtered into the query again.

Since master data and hierarchies are not read directly by the source system, they need to be in BW already when you execute a query.

3.5 BAPI RemoteCubes

A BAPI RemoteCube is an InfoCube with transaction data that is managed externally. Only the structure of the RemoteCube is defined in BW. The data is read for reporting using a BAPI from another system. With the help of a RemoteCube, you can carry out reporting using data in external systems without having to physically store transaction data in BW.

A BAPI RemoteCube can be used in special cases, for example, in market research data. When you start a query with a RemoteCube, you trigger a data request with characteristic selections. The source structure is dynamic and is determined by the selections. The external system transfers the requested data to the OLAP processor with the help of the BAPI.

A BAPI RemoteCube gives you the option of attaching external systems, structures in particular, which are not relational (hierarchical databases). With this, you can use any reading tool that supports the interface of an external system.

3.6 MultiCubes

A MultiCube is a superior InfoCube that combines data from several BasicCubes/RemoteCubes, and brings it together into one context. MultiCubes only exist as a logical definition; it does not contain any data itself. Their data comes exclusively from the basic InfoCubes. A query generated from a MultiCube is divided into several queries using the InfoCubes in question, which can be worked on in parallel.

MultiCubes are used to increase the transparency and performance and to enable more flexibility through more simple maintenance.

Example: Instead of creating plan and actual figures for an InfoCube, you create one InfoCube with plan figures and one with actual figures. This allows the integration of BW and APO.

You have to select a key figure, contained in a MultiCube, from at least one of the InfoCubes involved. In general, exactly one of the InfoCubes provides the key figure. However, there are cases where it is better for the user to select from more than one InfoCube. If the key figure 0SALES is in an InfoCube as an actual value, for example, and saved in another as a plan value, which avoids overlapping in the data records (in other words, Sales is divided disjunctively between several InfoCubes), then it makes sense to select several InfoCubes.

You cannot have joins on single record level when you use MultiCubes, since the MultiCube only brings together data from several InfoCubes.



4 Staging Scenarios with Persistent Data Storage

This unit introduces staging scenarios, in which the data is stored on a long-term basis.

4.1 Standard Scenario

Data flow:

Source	Initial storage	Long-term data storage	Reporting storage
Source System	→ PSA	→ ODS Objects	→ InfoCube

Usage

The process of using the standard scenario corresponds to the basic recommended procedure for storing data in BW. In special cases, you can use the scenarios as described in the following.

4.2 Staging with ODS Objects without InfoCubes

Data flow:

Source	Initial storage	Long-term data storage	Reporting storage
Source System	→ PSA	→ ODS Objects	[→] ODS Object

Usage

Use the staging process for storing data in an ODS object without InfoCubes, when the transaction data is highly volatile and when there is a high demand for the changes made to reporting data, which was already loaded, to be permanently available. The status of documents must always be kept up-to-date, for example. In turn, this requires the data, which is supported by the ODS object, to be overwritten.

This solution also makes it possible to use a second ODS object as storage for reporting.

As the ODS object data cannot be aggregated, this is not a good solution for reporting with mass data.

4.3 Staging without ODS Objects with InfoCubes

Data flow:

Source	Initial storage	Long-term data storage	Reporting storage
Source System	→ PSA	→ InfoCube	InfoCube

Usage

Use a staging scenario without ODS levels if you can work without the advantages of an ODS object (cleansing, transforming, consolidating, integrating data from different processes, level of detail of information, durability of the data). This applies in particular if the system already guarantees the integrity of the data (CO-PA, General Ledger) and therefore works without saving the consolidated data explicitly in BW, or if the data is extracted from a single source system.

4.4 Further Processing from InfoCube to InfoCube

Data flow:

Source	Long-term data storage	Reporting storage
InfoCube	→ InfoCube	InfoCube

Usage

This scenario is also known as a data mart scenario and is characterized by the data flow from InfoCube to InfoCube, which is possible within a BW System as well as between several systems. Data marts created like this offer better overview and maintenance of a dataset (snapshot) and allow you to separate the data according to task area.

You can also use this scenario if you want to change an InfoCube after it has already been loaded with data.

You are recommended to use a data mart scenario between two different BW Systems, if you are processing large amount of data. On the one hand it gives you better performance, on the other, the BWs remain more visible. In this way, you can separate the data according to task area. You also have the option of creating hub and spoke scenarios, which have a central BW where the data from distributed systems runs together and is standardized. This is where the data is made available for data marts.

4.5 ODS Server and BW Server

Data flow:

BW 1	BW 2
ODS Object	→ InfoCube

Usage

You have the option of using a BW server as the source BW and a BW server (target BW) for reporting. For this scenario, you need to use a BW server as the source BW, where the data is cleansed and updated in a complex process.

The source BW collects the data from the OLTP. It also cleanses and consolidates the data when it is updated into ODS objects. There is a central data entry here, which is also an ODS server. As a further process, the data is then replicated into the target BW in InfoCubes through the data mart interface. Reporting can take place once the master data has been uploaded.



5 Staging Scenarios without Persistent Data Storage

This unit introduces staging scenarios, in which the data is not stored on a long-term basis.

5.1 Staging with SAP RemoteCubes

Data flow:

Source	Initial storage	Long-term data storage	Reporting storage
Source System	→ none	→ none	→ SAP RemoteCube

Usage

SAP RemoteCubes are a suitable alternative to standard scenarios, if

- you need up-to-date availability of data from the SAP R/3 System
- if you only want to access small datasets sporadically, and replicating the data into BW would be too time-consuming in comparison
- only a few users execute queries simultaneously on the requested dataset.

As long as a suitable extractor is available (you can check this in the tables RSOLTPSOURCE or ROOSOURCE), an SAP RemoteCube query can replace a drill through on an SAP source system with the help of the report/report interface. The data in both cases is real-time. However, a drill through uses an OLTP report to present the data, whereas an SAP RemoteCube uses a BW report.

Restrictions:

- The source system must always be available.
- BW reporting on SAP RemoteCubes is an extra burden on the performance of the OLTP system. Performance is reduced even further when there are more users working on BW queries with SAP RemoteCubes and when the queries have a high level of aggregation.
- SAP RemoteCubes do not use update rules. The data to be presented in the query is taken directly from the communication structure. You cannot add Business Logic with update rules.
- SAP RemoteCubes do not support navigation attributes.
- Master data and hierarchies must be replicated in BW.

You can also use RemoteCubes with a prototyping process with the generic extractor. You can also access external systems quickly and easily with the help of the DB link.

Since SAP RemoteCubes do not require persistent data storage in BW, changes to structures (extract structure, transfer structure, transfer rules) in the entire staging process of the prototype phase can be performed quickly.

5.2 Data Mart Scenarios with SAP RemoteCubes

Data flow:

BW1	BW2
InfoCube/ODS Object	→ SAP RemoteCube

Usage

When you create a system landscape with several BW Systems, you have two alternatives for distributing the transaction data: You can replicate the data beforehand, or only read it once it is requested from the source system. You can use the SAP RemoteCubes for the second option. The restrictions that apply when you use RemoteCubes with an R/3 System are of little importance when you join two BW Systems. This can avoid the reduction of the source system (BW) performance, because suitable aggregates are used. The restrictions to system availability, navigation attributes, and master data still apply.

When you use a data mart scenario, it does not usually make sense to replicate data on document level in all the connected BWs. Since users only rarely use the data on the connected BWs in detail, this purpose is well suited to the use of SAP RemoteCubes.