IBM InfoSphere Optim for z/OS
Version 11 Release 3

Compare for IMS/VSAM/Sequential File Data

IBM InfoSphere Optim for z/OS
Version 11 Release 3

Compare for IMS/VSAM/Sequential File Data

Note
Before using this information and the product it supports, read the information in "Notices" on page 125.

Version 11 Release 3

This edition applies to version 11, release 3 of IBM InfoSphere® Optim for z/OS and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

About this publication

This document explains how to use Compare to compare relational sets of data from IMS tables, VSAM or sequential files.

Chapter 1. Introduction to Comparing IMS, VSAM and Sequential Files

Compare is a relational comparison facility used to compare the data from one set of source tables with another. This facility is unique in providing the ability to analyze multiple tables to identify not only changes to a pair of tables, but more importantly changes to related tables. By analyzing an entire set of data across multiple tables, users can pinpoint exactly where the changes occurred.

Depending on your site's license, the source of the data can be a set of rows that currently resides in a DB2® or IMS™ database, VSAM or sequential file, or a set of rows extracted previously from the database and stored.

The information in this document pertains to Compare for IMS, VSAM and sequential file data only. For complete details on Compare processing, refer to the *Compare User Manual*.

When comparing IMS, VSAM or sequential file data, you can compare one table or multiple tables. Any of these sources can be compared to any other: DB2 tables, IMS tables, VSAM files, sequential files. A single Extract File can contain data from any of those sources. Tables may be related.

IMS, VSAM and sequential file data is hereafter referred to as "I/V/S data".

Considerations and Restrictions

In planning for comparison of I/V/S data, be aware of the following:

- To compare an IMS table with a non-IMS table, the IMS table must be a uniquely keyed root segment in the IMS database.
- Optim™ Compare support for IMS database tables required changes to the Extract File to properly compare non-unique segments. As a result, if an Extract File was created in a prior release and includes IMS tables, it can be compared with I/V/S data only if the IMS segments are uniquely keyed.
- A table definition for IMS, VSAM or sequential file data is limited to a maximum of 2500 fields.
- An extract file on tape cannot be used in a compare process.

Processing Flow

To compare a set of related data, you specify the sources of the data to be compared. Compare then accesses the data, performs the comparison, and presents the results in a formatted, row-by-row manner either as a screen display, a printed report, or both, according to your specification.

You can compare individual tables or a set of related data residing in multiple tables. When the data resides in multiple tables, the changes are characterized as ***Direct*** changes (occurring in that table) and ***Related*** changes (occurring in a related subordinate table).

In general, Compare performs the following steps for each request:

- Identify the two sets of data to be compared.
- Compare the data in each pair of tables from the two sources and mark the changes.
- If the comparison involves data from multiple tables, chain rows from related tables and indicate related changes.
- Store the results of the comparison.

Identify the Data

To perform a Compare Process, you must identify the source of the data. You can specify:

- DB2 table data
- DB2 table data previously extracted and stored in an Extract or Archive File
- Extract file containing IMS table data, VSAM or sequential file
- IMS table, VSAM or sequential file

Compare Definition

The information specified for a Compare Process is called a Compare Definition. You can direct Compare to save the definition so it can be reused as needed, or you can use the definition for a single Compare Process and discard it.

Specify Data to be Compared

To specify the data to be compared, the following is required:

- The names of the tables from which data is to be selected.
- The relationships between the named tables that are to be used for the data selection.
- For each table, you can qualify which rows are to be selected by specifying selection criteria for any of the tables.

If a single table is the source, you need specify only the name of the table and any selection criteria for that table.

The specifications for the set of data can be used for a single Compare Process or stored in an Access Definition. Access Definitions are described in detail in the *Common Elements Manual*.

Map Tables and Columns

After specifying the two sets of data, you can map these sets of source tables. This is useful when the names of the tables are different. You can also map the columns from any pair of tables to map unlike column names or eliminate one or more columns from the comparison.

Compare Process Options

In addition to specifying the data to be compared, there are several processing options. Before execution, you are prompted for these Compare Process options. The options include whether the process is to be executed online or in batch, is to be limited to a specific number of rows, or is to create a report. If you specify that a report is to be created, you are prompted to specify what the report is to include. For example, you can request that all rows or only changed rows are included. In addition, if you perform the Compare Process online, the browse facilities can be invoked automatically to review the results when the process completes.

Terminology

This topic describes some common terms that are used when comparing IMS, VSAM, and sequential file data.

Access Definitions

The set of specifications for a set of related data to be extracted when the data to be compared resides in multiple tables is an Access Definition. All Optim components can use the same Access Definitions. The Access Definition contains a variety of information, including:

- Set of tables from which to extract data.
- Relationships used in the Extract Process and direction of traversal.
- Optionally, the order in which data is displayed or selection criteria for data in the listed tables.

The first table from which the data is extracted is the Start Table. All other listed tables are visited in logical sequence based on relationships and specifications in the Access Definition.

You can save an Access Definition for repeated future use. The saved definition can be modified and re-saved under the same or a new name.

The name of an Access Definition consists of three parts: *group.user.name*

The *group* and *user* portions are useful for organizing projects. For example, you can assign a unique group name to each project and, within each project, a unique user value for each person (user) in the group.

Archive Files

An Archive File is a sequential file that contains the archived data and information about the data characteristics. An Archive File provides safe, unmodified, storage of archived data. You can use an Archive File as a source file for Compare. See the *Archive User Manual* for more information.

Note: An Archive File can contain DB2 table data, but not data from an IMS table, VSAM or sequential file. I/V/S data can be extracted and stored in an Extract File, but not an Archive File.

Column Maps

A Column Map is a set of specifications used by Compare to determine which columns are to be included or omitted from the comparison and the correspondence of unlike-named columns. For example, a timestamp may be different in different executions of the application. If included in the Compare Process, every compared row would be identified as having changed.

The Column Map name consists of two parts: *mapid.mapname*. The *mapid* is frequently used to group the maps by user or project.

Column Maps can be stored in the Directory.

Compare Definitions

A Compare Definition contains information you specify to perform a Compare Process. This definition can be specified for a single Compare Process, or stored in the Optim Directory and reused. A Compare Definition includes:

- Specifications for the data sources. You can compare one table to another, or you can compare two sets of tables and specify the data sources as:
  - An Extract or Archive File containing the data.
  - A set of data defined by an Access Definition.
  - All rows from the tables specified for Source 1 (available only for Source 2).
  - Data from an IMS or DB2 table, VSAM or sequential file.
- Specifications (that is, a Table Map and, optionally, Column Maps) to map the two sets of data.
- A Match Key.

The name of a Compare Definition consists of three parts: *group.user.name*

The *group* and *user* portions are useful for organizing projects. For example, you can assign a unique group name to each project and, within each project, a unique user value for each person (user) in the group.

Compare Files

A Compare File is a sequential file created by Compare to store the results of the comparison. The contents of a Compare File can be browsed online or printed.

Direct Change

In a Compare Process, a direct change indicates that rows from Source 1 and Source 2 have the same match key value, but different values in one or more other columns. Direct changes are noted in the comparison results by a D marking the changed rows. The columns containing the changes are highlighted. A variety of commands are available to locate and scroll to the changed rows.

Extract Files

An Extract File contains the extracted data and information that defines the characteristics of the data. If an Extract File is used to define a source, the name of the Extract File is stored with the results of the comparison instead of the extracted data. Once created, Extract Files can be reused, providing a constant set of data with which to compare modified data. Extract Files are available to all users.

Note: An extract file on tape cannot be used in a compare process.

Match Keys

A match key is used by the Compare Process to match rows from one source with rows from the other source. When available, Compare uses a primary key from one of the source tables for the Compare Process. However, when a primary key is not defined in either the DB2 or IMS database or the Optim Directory for either table, you are prompted to create a match key. You can also define a match key for the current Compare Process or a Compare Definition.

Match keys are similar to primary keys, except that the column name in the key must exist in both source tables either directly, having the same base name, or indirectly, using a Column Map to match unlike names. Also, unlike the database primary keys, match keys do not have to be based on a unique index and, unlike Optim primary keys, match keys are not available to be used by any other user or process. Match keys are stored in the current Compare Definition only and not in the Optim Directory.

Primary Keys

A primary key is the column or set of columns that uniquely identifies each row in a table. For example, the CUSTOMERS table has a column, CUST_ID, that contains a unique value for each row in the table. CUST_ID is an acceptable primary key for the CUSTOMERS table.

Related Change

In a Compare Process, a related change indicates that dependent rows in a related table have direct changes or unmatched rows. Related changes are noted in the comparison results by an R marking the row in the parent table related to the changed rows in the dependent table. The R indicator is propagated through the levels such that the related row in the parent, grandparent, great-grandparent, and so on, also displays the indicator. The actual changed rows in the dependent table are identified as direct changes by a D indicator.

Relationships

A relationship determines how two tables are related. A relationship can be defined in the DB2 Catalog, IMS database, or the Optim Directory. Whenever a list of relationships is presented, the source is indicated. You can create or modify an Optim relationship using Compare, or you can browse DB2 relationships. You can also use DB2 relationships as a model for new Optim relationships.

In the DB2 Catalog, a relationship is defined by a primary key/foreign key pairing. The foreign key is the set of columns in a child table that describes the correspondence with the primary key columns in the parent table. However, Optim relationships do not require primary key/foreign key pairing. You can define Optim relationships that pair any compatible columns between two tables, and specify substring and concatenation functions for columns, and literal and constant values.

In an IMS database, tables can be related through physical or logical relationships. An IMS physical relationship is composed of the parent's concatenated key along with any sequence field defined in the child table. The Optim solution will display the concatenated key (along with additional suffix information for non-uniquely keyed segments) in the FOP_SURKEY column of the compare results.

Source 1 and Source 2

In a Compare File, Source 1 and Source 2 can be either:

- An Extract or Archive File containing DB2 table data
- An Extract File containing IMS table, VSAM or sequential file data
- An Access Definition that defines the data to be extracted for the comparison.
- An IMS table
- VSAM or sequential file

Note: For the Compare Process, it is irrelevant which data is identified as Source 1 and Source 2. However, when browsing the data online and in reports, Source 1 is used to determine table names and column headings, by default.

Surrogate Key

When comparing IMS tables, the Optim solution will automatically use the surrogate key, along with any other additional match keys to match rows from source 1 and source 2. The surrogate key is composed of the full IMS concatenated key of the segment along with additional suffix information that is added by the Optim solution for non-uniquely keyed segments.

Table Definitions – IMS, VSAM, Sequential Files

To compare I/V/S data, you need to create Table Definitions for each unique record type. When you create a Table Definition for data with multiple record layouts, you must specify selection criteria to uniquely identify each record layout. For IMS data, see Defining a Table for IMS Data on page 38 and for VSAM or sequential data, see Defining a Table for VSAM or sequential data.

Table Maps

A Table Map is a set of specifications used by Compare to match the sets of source tables when comparing data from multiple tables. The Table Map name consists of two parts: *mapid.mapname*. The *mapid* is frequently used to group the maps by user or project. Table Maps can be stored in the Optim Directory. (Table Maps can also be used and created by Move. Different rules are used to define Table Maps for Compare than for Move.)

Tables

Throughout this document, the word tables refers to tables, views, aliases, and synonyms, which are operated on in a similar manner. The differences in handling are noted where pertinent. Data from Materialized Query Tables (MQTs) can be browsed. You can edit data in a User-Maintained MQT. System-maintained MQTs are protected from modification.

Chapter 2. Session Overview

Assume you want to compare two sets of related data--the version before and the version after executing an application in development. This task comprises three basic steps: specify the two sets of source data to be compared, execute the Compare Process, and review the results.

Specify the Data to Compare

To begin the Compare Process, select Option 8 on the Main Menu, to display the following panel:

```
----------------------------- COMPARE Process --------------------------------
OPTION ===>
                                                        SQLID ===>
1 SINGLE   - Compare One Table to Another               SUBSYS ===>
2 MULTIPLE - Compare Two Sets of Tables LOCATION ===>
3 PERFORM  - Specify COMPARE Parameters and Perform COMPARE

B BROWSE   - Browse Results of Previous COMPARE
R REPORT   - Generate Report from Previous COMPARE

Type of Compare Definition to Use for COMPARE ===> T (P-Perm, T-Temp)

If Permanent, Specify New or Existing Compare Definition Name:
  GROUP ===>
  USER ===>
  NAME ===>

Use '_' for DB2 LIKE Characters ===> N (Y-Yes, N-No)
*******************************************************************************
```

First, you must choose an option to compare versions of a single table or versions of a set of tables. The single table option provides a quick path when you want to compare two versions of a single table. All you need to provide are the names of the two tables for the comparison and a Compare File in which to store the results of the comparison.

In contrast, several considerations are added when comparing data in multiple tables. If the data is to be extracted, you must provide the names of the tables and may need to select relationships and traversal paths to determine the set of data.

Once the data is defined and, if necessary, extracted, Compare identifies all changes and indicates direct and related changes. (Changes to a row are Direct; changes to a row or rows in a dependent table are Related.) When two versions of a single table are compared, there are no related changes.

In this sample, Option 2 is specified. The following panel is displayed:

```
-- Specify COMPARE Source Types: TEMPORARY CD --------------------------------
Command ===>

COMPARE can process data saved in an Extract File or in DB2, IMS, VSAM and
Sequential Tables.  Specify source types as follows:


Source 1 ===> 1  1 - Extract File
                 2 - Set of Data Defined by an Access Definition

Source 2 ===> 3  1 - Extract File
                 2 - Set of Data Defined by an Access Definition
                 3 - All Rows from Multiple Tables
```

The Specify COMPARE Source Types panel is used to specify the types of sources to be used. For both Source 1 and Source 2, you can use data previously extracted and stored in an Extract File on disk, a DB2 or IMS table, VSAM or sequential file, or you can define the data to be extracted for the comparison using an Access Definition.

An additional choice for Source 2, All Rows from Multiple Tables, is available to specify that all rows from the tables defined for Source 1 are to be extracted. This is useful when comparing before and after images of your test database. You can retain the Extract File used to create the test database, execute your application, and then compare the Extract File with the entire test database. This comparison will identify all inserted, deleted, and changed rows, and all orphan rows.

The options on the Specify COMPARE Sources panel are specified as an Extract File for Source 1, and All Rows from Multiple Tables for Source 2. When you press ENTER, Compare prompts for additional information about each source based on your response on this panel.

```
-- Specify COMPARE Sources: TEMPORARY CD -------------------------------------
Command ===>                                        SCROLL ===> PAGE

Source 1: Extract File
  DSN ===> 'DEMO.EXTRACT'

Source 2: All Rows from Multiple Tables
  COMPARE will initially use matching Table Names from the above
  Extract File for the Table Map
```

Source 1 An Extract File

An Extract File is used for Source 1, with the dataset name DEMO.EXTRACT

Note: An extract file on tape cannot be used in a compare process.

Source 2 All Rows

After the information needed for Source 1 has been supplied, specify Source 2. Since the data for Source 2 has been specified as All Rows from Multiple Tables, you are not prompted for additional specifications. In this case, Compare assumes that the same tables specified in Source 1 are to be used as Source 2.

Execute the Compare Process

After the sets of data to be compared are specified and the relationships to be used for the comparison are selected, the Specify Compare Parameters and Execute panel is displayed. (This panel is also displayed when Option 3 is selected on the Compare Process menu to re-execute an existing Compare Definition.)

```
------------------- Specify COMPARE Parameters and Execute --------------------

Command ===>                                                SCROLL ===> PAGE

Compare File DSN ===> 'FOPDEMO.SAMPLE.COMPARE'

Source 2 Extract Options:
  Limit Number of Extract Rows   ===>       (1-4294967295, Blank/SL)
  Extract Data using             ===> D     (D-DB2,I-IBM High Perf Unload)

Compare Options:
  Generate Report After Process  ===> N     (Y-Yes, N-No)
  Run Process in Batch or Online ===> O     (B-Batch, O-Online)
    If Online, Invoke Browse     ===> Y     (Y-Yes, N-No)
    If Batch, Review or Save JCL ===> R     (N-No, R,Review, S-Save)
```

This panel has the following fields:

Compare File DSN
: Specify the name of a sequential file that is to contain the results of the comparison. This is the Compare File. If the file you name does not exist, Compare will prompt for allocation information and automatically allocate the file for you. The Compare File for this sample is named FOPDEMO.SAMPLE.COMPARE.

Compare Options
: Several Compare Process options are provided. These include the ability to limit the number of rows of data to be extracted if one or both sources are database tables, to specify whether or not a report is generated, and to specify whether the process is executed online or in batch. If online, you can specify whether the results of the comparison are automatically displayed in a browse session when the Compare Process is completed.

  If the Compare Process must extract the data from one or both sources (in this example, Source 2 must be extracted), you are prompted to specify a maximum number of rows to be extracted for the source. You can specify any value from 1 through the site limit that is displayed on the panel or leave blank to automatically default to the site limit. (Additional prompts are displayed only when pertinent. For example, if an unload program is available, you can specify whether it is to be used. For details on all the prompts, see “Specify Compare Parameters and Execute Panel” on page 72.)

Generate Reports
: When you specify **Yes** to **Generate Reports,** the following panel is displayed to prompt for information required to generate the report before the Compare Process executes.

```
---------------------- Specify COMPARE Report Parameters -----------------------
Command ===>                                                 SCROLL ===> PAGE

Compare File DSN ===> FOPDEMO.SAMPLE.COMPARE
Report File DSN ===> 'FOPDEMO.SAMPLE.REPORT'

  Report Type    ===> D (S-Summary, D-Details)
  Lines Per Page ===> (0-No Titles, 1-99, Blank=57)
  Specify Table Name to Limit Report (Blank for ALL Tables in Compare)
   Table Name    ===>

If Detail Report Specify Format and Select Desired Row Types:
  Report Format          ===> C (C-Columnar, S-Sidelabels, E-External)
  Which Columns?         ===> A (A-All, D-Different and Key Columns Only)
  All Rows               ===> Y (Y-Yes, N-No)
  Or, if NO, Select One or More of the following Row Types:
    Direct Changes       ===> N (Y-Yes, N-No)
    Related Changes      ===> N (Y-Yes, N-No)
    Unmatched Rows       ===> N (Y-Yes, N-No)
    Orphan Rows          ===> N (Y-Yes, N-No)
    Duplicate Match Keys ===> N (Y-Yes, N-No)
  Wide Lines             ===> C (C-Change File, W-Wrap Data)
Display Unused Columns      : Y (Y-Yes, N-No)
```

The name of the Compare File is supplied. You should specify a sequential file to receive the output. If you specify a sequential file, you can use standard ISPF facilities to browse, edit, and print the file.

When multiple tables have been compared, you can request a report on only one of the tables by typing a table name at the prompt, **Provide Table Name to Report**.

The remaining prompts determine the contents of the report. **Report Type** prompts you to specify whether the report is to include the summary information, or both the detail and summary information.

The summary information includes the names of the tables from each source, the total number of rows for every type of detail that can be reported on and the total number of rows in the comparison from each table.

The details are the rows that have been compared. You can select which details are to be included based on their status. (The status is printed along with each row.) The available statuses are listed on the panel under the prompt **Select Rows to be Printed**. You can select all rows (specify Y for All Rows) or one or more of the following:

Direct Changes
: All rows from Source 1 and Source 2 that have the same match key value but different values in one or more other columns.

Related Changes
: All rows that have dependent rows with direct changes or unmatched rows.

Unmatched Rows
: All rows from Source 1 and Source 2 that do not have a match key value that matches a row from the other source.

Orphan Rows
: All rows from Source 1 and Source 2 that do not have a parent.

Duplicate Match Keys
: All rows from Source 1 and Source 2 that have duplicate match key values. Since the match key values are the same in multiple rows, Compare does not attempt to match the rows arbitrarily. Instead, the rows are unmatched and marked with a special flag.

  For details on all the fields on the Specify COMPARE Report Parameters panel, see “Generate Report” on page 84.

Display Unused Columns

This column is non-modifiable. Indicates whether unused columns are displayed in the Compare Process Report. Use the Compare Options panel to specify a value for this field.

Perform the Process

When you have finished indicating the report options, press ENTER. The Compare Process is performed. The steps in the Compare Process are as follows:

1. Get data for Source 1. If Source 1 is an Access Definition, the data is extracted. If Source 1 is an Extract File, the data has already been extracted and is available.
2. Get data for Source 2. If Source 2 is an Access Definition or All Tables, the data is extracted. If Source 2 is an Extract File, the data has already been extracted and is available.
3. Compare the data in each pair of tables to determine the equal rows, direct changes and unmatched rows.
4. Process the relationships to determine the related changes and orphan rows.

Status information is displayed as the process executes. This information notes which step is currently being performed and is updated as each table is extracted, as each pair of tables is compared, and as the selected relationships are traversed to chain the related rows that have been changed.

Table Map

After you have specified Source 1 and Source 2, Compare prompts for the Table Map. This provides an opportunity to specify tables that do not have exact name matches or omit tables from the comparison.

The following panel is displayed:

```
------------------------- COMPARE Process Table Map -----------------------------
Command ===>                                                 Scroll ===> PAGE

Commands: APPLY,SAVE,LIST,MAP,POPULATE,ACM,CLEAR,MKEY,MOPT  END when Done
Source 2 Tables must Exist in Corresponding Extract File
 Src 1 CID: SAMPLE                                          Column
 Src 2 CID: SAMPLE                                      >>  Map ID ===>

 Source 1 Table Name      Source 2 Table Name      Type   Column Map or LOCAL
-------------------- ------------------------>> ------- --------------------
************************************ TOP ************************************
SALEHDAM_SALES        SALEHDAM_SALES              EXISTS
SALEHDAM_CUST         SALEHDAM_CUST               EXISTS
SALEHDAM_SHIPT        SALEHDAM_SHIPT              EXISTS
SALEHDAM_SHIPI        SALEHDAM_SHIPI              EXISTS
********************************** BOTTOM ***********************************
```

Source 1

Different rules are used for populating the CID (default Creator ID) and Table Name values based on the source types. For Source 1, CID is populated with the Default Creator ID from the Extract File or Access Definition. In addition, all table names from the Extract File or Access Definition are filled in as the Source 1 tables. These values are protected.

Source 2

For Source 2, if the source type is an Extract File or Access Definition, CID is populated appropriately from the named source. The Table Name values are populated with only the table names from Source 2 that match table names in Source 1. The remaining values are blank.

If the source type is **All Rows in Multiple Tables** (as in this example), then the Src 2 CID and Table Name values are populated with the same values as their Source 1 counterparts.

Specifying CID

If Source 2 is **All Rows**, you can overtype the CID as desired. For this sample, assume the default Creator ID is FOPDEMO.

Specifying Table Names

You can edit the table names directly or request a selection list of available tables not currently mapped as Source 2. When editing names, you can use the CLEAR command to remove all Source 2 names before typing names on the panel. You can also prefix or suffix the Source 2 names with a string of your choice.

Use the LIST TABLES command to display a selection list of tables and automatically insert your selection as the Source 2 table name. The APPLY command overlays all or part of the displayed Table Map with the specifications from a stored Table Map. (For details about defining and storing Table Maps, see the *Common Elements Manual*.)

You can also replace the initial table names with an existing Table Map. In addition to Table Maps, you can specify Column Maps to map the columns for a pair of tables that do not have matching names provided they have compatible data types. You can also eliminate columns from the comparison. You can define a Column Map for the current Compare Definition only (referred to as LOCAL) or a Column Map that is stored in the Optim Directory and reusable. (For details about defining and storing Column Maps, see the *Common Elements Manual*.)

For this sample, the tables and column names are the same and all tables and columns are to be compared. Column Maps are not needed.

When the Table Map specifications are complete, use **END** to proceed.

Relationships

When comparing two sets of data from multiple tables, the relationships used to determine the related changes are important. Only one relationship may be selected between any two tables in a given direction (that is, for a specific parent and child), regardless of the relationships used to extract the data. In this example, only one relationship is defined between these tables; therefore, Compare will automatically use the appropriate relationships.

Review the Results

The Compare Process creates the Compare File containing the results and, if requested, a report that can be stored in a sequential file.

In the following panel, the specifications for this sample session establish that the process is executed online and a browse session is invoked when the process terminates. This browse session enables you to display the results of the comparison online and to scroll and view the changes as they relate to the other compared data. The source of each row is clearly identified and all changes are highlighted.

To begin the browse session, a panel providing summary information and a selection list of the pairs of tables involved in the Compare Process is displayed.

```
---------------------- Compare Summary Selection List -------------------------
Command ===>                                            Scroll ===> PAGE

Use 'S' to Select Browse Start Table, 'I' for Extended Table Information

Source 1: XF - Z13600MP.FOPDEMO.EXTRACT, SUBSYS: TDB2
Source 2: Database Tables, SUBSYS: TDB2
                                                                 Rows    Non-
                                                 Changes         with    Unique
                          Total   UnMatched  Equal (D)irect     Missing Match
Sel   Source:Table Name   Rows    Rows       Rows  (R)elated    Parents  Keys
--- --------------------->> -------- -------- -------- --------- -------- -------
*** ********************************* TOP ************************************
S__ 1 FOPDEMO.CUSTOMERS     703         1        690    D: 12     N/A      0
    2 FOPDEMO.CUSTOMERS     704         2               R: 25     N/A      0

___ 1 FOPDEMO.ORDERS       1712        10       1697    D: 5        0      0
    2 FOPDEMO.ORDERS       1709         7               R: 4        0      0

___ 1 FOPDEMO.SHIP_TO       503         9        490    D: 4        0      8
    2 FOPDEMO.SHIP_TO       526        32               R: N/A      6     25

___ 1 FOPDEMO.DETAILS      3591        11       3574    D: 6        0      0
    2 FOPDEMO.DETAILS      3596        16               R: N/A      0      0

___ 1 FOPDEMO.ITEMS         102         0        102    D: 0      N/A      0
    2 FOPDEMO.ITEMS         102         0               R: 0      N/A      0
*** ******************************** BOTTOM ***********************************
```

The summary information provides an overview of the comparison results. Each source is identified.

Note that Source 1 is the Extract File Z13600MP.FOPDEMO.EXTRACT, and Source 2 is a set of database tables. The source of both is the subsystem TDB2.

Statistics are included for the following items:

Total Rows
The total number of rows from each table.

UnMatched Rows
The number of rows from each table that contain a match key value that does not match a row in the table with which it was compared. This occurs when rows have been added to or deleted from one of the sources.

Equal Rows
The number of rows in a pair of compared tables that are the same. That is, every column included in the comparison in the row from Source 1 exactly matches the corresponding columns in Source 2.

Changes
The number of rows in each table where the match key value matched a row in the other source, but a value in another compared column did not match. Changes are identified as:

D - Direct
Number of rows that are different between the two named tables.

R - Related
Number of rows that have dependent rows with direct changes or unmatched rows.

Rows with Missing Parents
The number of rows in each dependent table that do not have a parent row in a source table. This can occur when the parent row was deleted from one source, but the children were not.

Non-Unique Match Keys
The number of rows that have duplicate match key values in each source table. (The Optim

Directory does not require that primary keys are based on unique indexes. Therefore, if a Directory primary key or an explicit match key is used for the Compare Process, non-unique match key values can be encountered.)

To display the comparison results, select any pair of tables as the starting point for the session by typing an S in **Sel** prior to that pair of tables. In this example, the pair of CUSTOMERS tables is selected. The following panel is displayed:

```
-------------------- Optim: Browse (Source 1 Names Shown) ---------------------
Command ===>                                                 Scroll ===> PAGE

Cmd Chg Src == Table: CUSTOMERS(T1) ======================== 1 OF 717 === MORE>>
            CUST_ID       CUSTNAME                ADDRESS           CITY
            ------- -------------------- ==================== ---------------
*** ************************************ TOP ********************************
___ DR  1   00001   Audio-Video World   593 West 37th Street   Brass Castle
___ DR   2  00001   Audio Video World   593 West 37th Street   Black Castle
___      12 00002   Select-A-Vision     5720 MacArthur Drive   Evening Shade
___  R   12 00003   Showplace           1 Ocean Parkway        Alto
___ DR  1   00004   Audio-Video World   593 West 37th Street   West Palm Beach
___ DR   2  00004   Audio-Video World   593 West 37th Street   Panacea
___      12 00005   Take Home Movies    Box 357                Fence Lake
___      12 00006   Main Street Video   Gateway Shopping Cen   Pumpkin Center
___  R   12 00007   Cinemagic           Pass-a-Grille Beach    Pass-a-Grille
___       2 00008   Director's Chair    347 Miners Row         Spuds
___  R   12 00009   Prime Time Video    64 Newberg Avenue      Loving
___  R   12 00010   Reely Great Videos  590 Frontage Rd        Christmas Valley
___      12 00011   Director's Chair    347 Miners Row         Kiester
___       2 00012   Main Street Video   Gateway Shopping Cen   Howey in Hills
___      12 00013   Front Row Video     U.S. Highway 130       Christmas
___      12 00014   Reely Great Videos  590 Frontage Rd        Economy
___      12 00015   Director's Chair    347 Miners Row         Happy Camp
___      12 00016   Movies-R-Us         1772 Bridge St         Bonanza
```

The data from both sources is displayed. The differences in the data are highlighted. The columns in the Match Key are listed first. On the display, the column heading and underline is highlighted to identify the Match Key column. In this example, CUST_ID is the only Match Key column.

In addition to the data, the display includes an information line that provides headings (Cmd, Chg, and Src), the table name, the number of rows and the relative position of the first displayed row, and a horizontal scroll indicator.

In this example, the table names from each source are the same, but that may not always be the case. By default, the Source 1 table name is displayed. The Creator ID, the portion of the name most likely to differ, is not displayed. (Similarly, the Source 1 column headings are displayed by default. If you prefer, you can display the Source 2 table names and column headings. Use the FLIP command to toggle the names.)

Compare assigns an identifier to the table name. This identifier provides shorthand notation for specifying the table as an operand on the many available primary commands. In the figure, T1 (Table 1), is assigned.

Src The rows from both sources are displayed. The source is identified in Src. You can readily identify rows that have changed by the value in Src.

Equal Rows
When the row in Source 1 exactly matches the row in Source 2, the row is displayed once and Src contains 12, indicating Source 1 and Source 2. (The customers with CUST_ID 00002 and 00003 are the same in both sources.)

Changed Rows
When the rows differ, the row from each source is displayed and Src contains either 1 or 2 to

indicate the source. (The rows for the customers with CUST_ID 00001 and 00004 have changed.) These rows are marked as having a Direct change. Data in the columns that do not match is highlighted.

One Source Only
When a row exists in only one source (it was added or deleted by the application), **Src** contains 1 or 2, as appropriate. (The customer Director's Chair, CUST_ID 00008, and Main Street Video, CUST_ID 00012, exist only in Source 2.) The Src value for these rows is highlighted.

Chg In addition to identifying the source of each row, the type of change, if any, is also marked. **Chg** identifies the rows with changes as **D** for Direct changes, and notes related rows with changes as **R** for Related changes. (In the figure, customers with CUST_ID 00001 and 00003, for example, have Related changes.)

Rows with no parent are identified by a U for Unusual row. The CUSTOMERS table is the parent, so this is not applicable for this table. The rows that have duplicate match keys are displayed and identified by brackets. None are displayed in this sample.

Display Related Changes

Initially, data from one pair of tables is displayed. To get a more complete view of what has changed, you must use the Join facility to display the related rows from other tables.

In this sample session, the CUSTOMERS table is directly related to the ORDERS table, which was included in the comparison. **Chg** indicates related changes for CUST_ID 00007, Cinemagic. Use the **J** line command or the **JOIN** primary command to display the related data. In this example, type **J** in the line command entry area to display the following panel.

```
-------------------- Optim: Browse (Source 1 Names Shown) ---------------------
Command ===>                                                 Scroll ===> PAGE

Cmd Chg Src == Table: CUSTOMERS(T1) ====================== 5 OF 503 === MORE>>
            CUST_ID       CUSTNAME          ADDRESS            CITY
            ------- -------------------- ================== ---------------
___ R   12  00007   Cinemagic            Pass-a-Grille Beach  Pass-a-Grille

Cmd Chg Src == Table: ORDERS(T2) ============================ 1 OF 3 === MORE>>
            ORDER_ID CUST_ID ORDER_DATE ORDER_TIME FREIGHT_CHARGES
            -------- ------- ---------- ---------- ---------------
*** ********************************* TOP ********************************
___ R   12  77784    00007   1998-01-29 11.30.11     52.00
___      2  77785    00007   1998-01-29 11.35.11     57.00
___     12  77786    00007   1998-01-29 14.25.41     15.25
*** ******************************** BOTTOM *******************************
```

An information line displays headings for the joined table. As with the first table, the table name is taken from Source 1 and an identifier assigned by Compare. Here, T2 (table 2) is assigned to the ORDERS table. The source and nature of changes is highlighted. In this example, a row in Source 2 does not exist in Source 1; ORDER_ID 77785 has been added to Source 2.

The DETAILS table, a dependent table related to ORDERS, is included in the Compare Process. Changes to related data in the DETAILS table are indicated by the R in the **Chg** column for ORDER_ID 77784 in the ORDERS table.

You can join to the other tables in the comparison and scroll the data.

Chapter 3. Data Migration

In addition to extracting sets of related data as part of the Compare Process when needed to obtain the source data, Compare enables you to extract the data as a separate, explicit process and store that data in an Extract File. When you extract the data as a separate process, you can compare data from different subsystems.

To specify the source you use an existing Access Definition or create a new Access Definition. The created Access Definition can be temporary, for a single use, or permanent, saved for repeated use. The Access Definition is used as input to the Extract Process. (See the *Common Elements Manual*, for a detailed discussion of creating and modifying Access Definitions.)

The Extract Process copies the specified data to an Extract File. The Extract File is saved and can be reused as needed.

To migrate data, select Option 7 MIGRATION on the Main Menu. (If Move is not installed, this option is SNAPSHOT.) The following panel is displayed when Move and Archive are also installed. (When only Compare is installed, the options to EXTRACT and BROWSE are available.)

```
------------------------------ Data Migration --------------------------------
OPTION ===>
                                                                SQLID  ===> FOPDEMO
                                                                SUBSYS ===> TDB2
1 EXTRACT - Extract Data from Source Tables                     LOCATION ===>
2 INSERT  - Insert Data into Destination Tables
3 LOAD    - Create Load Files and Perform Load
4 CREATE  - Create Tables and Related Object Definitions
5 CONVERT - Convert Extract File using Table and Column Maps
6 LIST    - List Extract Files in Directory
7 IMPORT  - Import Extract File and Populate Directory

R RETRY/RESTART  - Retry/Restart an Insert Process
B BROWSE         - Browse Content of Extract File or Control File
```

The following options are available on the Data Migration menu:

1 - EXTRACT

Specify the set of data to be extracted. After the set of source data is specified, this option extracts the data and stores it in an Extract File. (If Move is installed, the extracted data can include the rows from the tables and the object definitions for those tables. If Move is not installed, only the data is extracted.)

The specifications for the extracted data can be defined in an Access Definition and stored for repeated use or defined as temporary for one-time use.

If an unload program is available, it can be used to extract the data from Image Copy files or directly from the DB2 VSAM files.

2 - INSERT

This option is only available when Move is installed. For detail, see the *Move User Manual*.

3 - LOAD

This option is only available when Move is installed. For detail, see the *Move User Manual*.

4 - DELETE

This option is only available when Archive is installed. For detail, see the *Archive User Manual.*

5 - CREATE

This option is only available when Move or Archive is installed. For detail, see the *Move User Manual* or the *Archive User Manual.*

6 - CONVERT

This option is only available when Move or Archive is installed. For details, see the *Common Elements Manual*.

7 - LIST

Display a list of Extract files that match selection criteria you specify. The List process allows you to manage extract files registered in the Optim Directory. With List you can display, delete, browse, generate a report or see extended information for an extract file. You can also select a file to use in an Insert process. This option is only available when Move is installed. Refer to the *Move User Manual*.

8 - IMPORT

Create entries to register extract files in the current Optim directory. An extract file stored on tape must be registered in the Optim Directory before it can be used in a Convert, Create, Insert, or Report process. This option is only available when Move is installed. Refer to the *Move User Manual*.

R RETRY / RESTART

This option is only available when Move or Archive is installed. For details, see the *Common Elements Manual*.

B BROWSE

Display an Extract File or Archive File to view the contents of that file or a Control File to examine extracted data or identify rows in error.

Data Migration menu prompts

The following prompts are available on Data Migration menu:

SQLID

The current SQLID. Modify this value to connect using a different SQLID.

SUBSYS

The current DB2 subsystem. Modify this value to connect to a different DB2 subsystem. When connecting to a remote subsystem, this value should be the local subsystem where the remote location is defined.

LOCATION

The remote location. This prompt is displayed if remote access is available. Specify a value to connect to a remote DB2 subsystem. You can use a percent sign (%) to obtain a selection list of available locations. If the connection fails, the session is restarted and the Main Menu is redisplayed. If you leave this prompt blank, the local subsystem is assumed.

Extracting Data

The Extract Process is used to create an Extract File. An Extract File contains the selected set of related rows from one or more tables. An Extract File on disk can be used as one or both sources for a Compare Process. The Extract File can be used repeatedly and simultaneously by many users.

Note: An extract file on tape cannot be used in a compare process.

An Extract File is created by traversing a set of tables and extracting specific data from those tables. The tables and the relationships to use to traverse those tables are specified in an Access Definition. You can use specifications from an existing Access Definition, create a new Access Definition, or specify temporary definitions.

The Access Definition also includes other specifications, such as:

- Manual selection of specific rows in the Start Table. This selection process is referred to as Point-and-Shoot.
- Selection criteria for one or more tables.
- A maximum number of rows to extract for one or more tables.
- A numeric value used to determine random selection, such as select every twentieth row.

Extract Process Menu

When you select Option 1 EXTRACT on the Data Migration menu to perform the Extract Process, the following menu is displayed.

```
------------------------------- EXTRACT Process -----------------------------
OPTION ===>                                                  SCROLL ===> PAGE

1 TABLES              - Specify Set of Tables and Selection Criteria
2 PATHS               - Specify Traversal Paths via Relationship List
3 OBJECTS             - Specify Object Definitions to Extract
4 PERFORM             - Specify EXTRACT Parameters and Perform EXTRACT

Type of Access Definition to Use for EXTRACT ===> T (P-Perm, T-Temp)

If Permanent, Specify New or Existing Access Definition Name
  Group ===>
  User  ===>
  Name  ===>

Use '_' for DB2 LIKE Character ===> N (Y-Yes, N-No)
```

The Select Extract Process menu includes the following options:

1 TABLES
Define or modify the set of tables to be used for the Extract Process. The Select Tables/Views for AD panel is used to specify the names of the tables to be included in the extract. This is called the Table List.

This panel also displays the type of selection criteria, if any, that has been defined for each table. You can specify a random factor and row limits for each table on this panel. You can use commands to display additional panels for defining selection criteria, an SQL WHERE clause, and substitution variables. From the Table List, you can also invoke the Point-and-Shoot facility to select rows from the Start Table.

For detailed information on the Select Tables/Views for AD panel, see **Access Definitions** in the *Common Elements Manual*.

2 PATHS
Display and modify the relationship list. The Specify Relationship Usage panel is used to select the relationships to be traversed when extracting the data. For detailed information on how to specify the relationships and accompanying parameters, see the *Common Elements Manual.*

3 OBJECTS
This option is only available when Move is installed.

4 PERFORM
Display the Specify EXTRACT Parameters and Execute panel to specify the parameters and invoke the Extract Process.

Temporary or Permanent
Indicates whether the specifications defined for the extract are temporary, discarded after the extract, or permanent, saved in the Access Definition. For temporary, specify T for TEMP at the prompt, **Type of Access Definition to Use for EXTRACT**. For permanent, specify P. You must specify the name of an Access Definition for permanent specifications.

Access Definition Name
If you decide that the specifications are permanent, specify the name of the Access Definition to be used for the process. The prompts on the panel correspond to the three parts of the Access Definition name: GROUP, USER, and NAME.

Create a New Access Definition

If the name of the Access Definition you specify does not exist, Compare automatically prompts you to create a new Access Definition. The Select Tables/Views for AD panel is displayed. See the *Common Elements Manual* for details on how to define an Access Definition.

You can leave the prompts blank or use DB2 LIKE syntax to display a selection list of available Access Definitions. Use the Select line command, S, to select an Access Definition from the list, and then use ENTER to redisplay the EXTRACT Process menu with the name of the selected Access Definition displayed in Access Definition Name.

You can define or modify the Access Definition specifications, whether temporary or permanent, by selecting Options 1 or 2 again.

Once the source data has been specified, use the PERFORM Option to perform the extract.

Perform the Extract Process

When you select Option 4 PERFORM from the EXTRACT Process menu to perform the extract, the Specify EXTRACT Parameters and Execute panel is displayed.

```
---------------------- Specify EXTRACT Parameters and Execute ------------------
Command ===>

Current AD Name     : FOPDEMO.EXTRACT.SAMPLE
Extract File DSN ===>
Extract          ===> B (D-Data,
                         O-Object Definitions,
                         B-Both)
If Extracting Data:
  Limit Number of Extract Rows ===>        (1-4294967295, Blank/SL)
  Extract Data using           ===> D      (D-DB2, B-BMC UnloadPlus)
  Extract Data to Tape         ===>        (Y-Yes, N-No)

Perform Convert with Extract   ===> N      (Y-Yes, N-No)

Extract with Uncommitted Reads ===>        (Y-Yes, N-No)

Run Process in Batch or Online ===> O      (B-Batch, O-Online)
  If Batch, Review or Save JCL ===> S      (N-No, R-Review, S-Save)

Process Report Type            ===> S      (D-Detailed, S-Summary)
```

The Specify EXTRACT Parameters and Execute panel includes the following prompts:

Current AD Name
Name of the currently active Access Definition. This read-only value is provided by the system.

Extract File DSN
Name of the Extract File that is to contain the extracted data. This file must be a sequential data set. The Extract File name can be specified explicitly by enclosing it in quotes; otherwise, the default prefix as specified on the User Options panel is automatically prepended to the name.

When the Extract Process begins, Compare searches for the named data set.

- If the data set exists, Compare checks to see if it is suitable for an Extract File. If it is, the current Extract Process overlays the data.

- If it is not an Extract File, Compare does not perform the extract and prompts you for a new data set name.
- If the data set does not exist, Compare prompts for the necessary information to allocate the file. See the *Common Elements Manual*, section on Allocating External Files for a description of the allocation prompts.

You can obtain a selection list of data sets using either of the wild card characters, % or *, in the last position of the name. Use the Select line command, S, on the selection list to select the file.

Extract
This prompt is only displayed when Move is also installed.

Limit Number of Extract File Rows
Maximum number of rows of data that can be extracted. The Extract is terminated if the number of extracted rows exceeds this limit. Specify:

value 1 – 4,294,967,295

blank The site-defined limit, which is set on the Site Options panel.

Extract Data using
If you are extracting DB2 table data this prompt is displayed, offering two options for extracting the data: DB2 and the utility specified in the Site Options panel. Specify:

D Use DB2

B Use BMC UNLOAD PLUS to access the data in batch.

C Use CDB Auto-Unload (formerly known as SuperUnload) to access the data in batch.

O Use CDB Auto-Online Unload (formerly called RW-Unload) to access the data in batch.

I Use IBM® High Performance Unload to access the data in batch.

Extract Data to Tape
An extract file on tape cannot be used in a compare process.

Perform Convert with Extract
This option is only available when Move is installed. For details, see the *Move User Manual.*

Run Process in Batch or Online
Indicates whether the Extract Process is executed in batch or online. Specify:

B Batch

O Online

If site management has established a maximum number of rows for online processing and this request exceeds that limit, this option is forced to Batch and cannot be changed. Consult site management for guidelines.

If Batch, Review or Save JCL
For batch execution, indicate whether the JCL and Batch Utility control statements should be submitted, reviewed prior to job submission or saved for submission at a later time. Since the JCL and control statements are displayed in the ISPF editor, you can modify them for the current request and save them to submit later. Specify:

N Submit job, do not display or save the JCL and control statements.

R Display the JCL and control statements for review prior to job submission.

S Save the JCL and control statements. Prompts are provided for you to specify the name of a file in which to store the JCL and control statements.

Process Report Type
Indicator to include additional information in the Extract Process Report. Specify:

D Display detailed information in the Extract Process Report.

S Display summarized information in the Extract Process Report.

Perform the Extract

After you have completed the Specify EXTRACT Parameters and Execute panel, press ENTER. Compare evaluates the Access Definition to be used and your entries on the Specify EXTRACT Parameters and Execute panel. The Extract proceeds as follows:

Online Processing

- The Access Definition is evaluated. If the Access Definition contains an invalid entry, an appropriate error or warning is displayed.
- The Extract File is located and the contents evaluated to ensure that the named file is an Extract File. If it is not an Extract File, an error message is displayed and you are re-prompted to specify the file name. If the Extract File does not exist, you are prompted for allocation information and Compare creates the file.
- The Extract Process is executed online. A status report is displayed and updated periodically during the processing.
- An Extract Process Report is generated and displayed for browsing.

Batch Processing

- The Access Definition is evaluated. If the Access Definition contains an invalid entry, an appropriate error or warning is displayed.
- The Extract File is located and its attributes are evaluated to ensure that the named file is suitable as an Extract File. If it is not suitable, the job terminates and an error message is written to the job output file.
- The JCL and Batch Utility control statements are built.
- The Extract Process is executed as a batch job. Note the Extract File must be located again when the process is executed.
- An Extract Process Report is generated and stored in the default output file specified in the JCL.

If one or more errors or warnings are encountered, the EXTRACT Errors & Warnings panel is displayed. Any error prevents the Extract Process from proceeding. The Extract Process can continue despite warnings.

You may use the SHOW STEPS command to display additional information about how the Extract Process will proceed.

Chapter 4. Compare Process

Compare provides a powerful facility for comparing two sets of related data. This topic details the steps used to perform the Compare Process.

The following steps are performed when sets of data are compared:

- Define sources. That is, specify the two sets of data to be compared.
- Match tables using Table Maps and, if desired, Column Maps.
- Specify Match Keys. These are the sets of columns used to match rows in the corresponding tables and are usually determined automatically by Compare.
- Select relationships. These are the relationships used to mark related changes and find orphan rows. These are usually determined automatically by Compare.
- Perform Compare Process.
- Browse results using the relational facilities available for a Browse session or reviewing the generated report.

Define the Sources

Many options are provided for defining the sources of the data. You can use a new or existing Access Definition and Compare will extract the data as part of the process. Alternatively, you can use the data previously extracted and stored in an Extract File on disk. The data to be compared can reside in a single table or in multiple related tables.

To compare data from IMS, VSAM or sequential files, you must create Table Definitions. For IMS data, see "Defining a Table for IMS Data" on page 32 and for VSAM or sequential data, see "Defining a table for VSAM or Sequential File Data" on page 41.

Store and Browse the Results

The results of the comparison are stored in a Compare File. You can browse the results online using Compare, or generate a report and print or store the report in a sequential file. A variety of report options enable you to control the contents.

Processing Flow

The Compare Process performs a series of steps based on your specifications. The following documents the steps that are performed when multiple tables are compared. A single table comparison only performs the first two steps--retrieving the data and comparing a pair of tables.

Retrieve the Data

Based on your specifications, Compare will obtain the data for each source from an Extract or Archive File, IMS or DB2 tables, VSAM or sequential file.

Compare Each Pair of Tables using the Match Key

The rows in each pair of tables are compared independently. Compare uses the "match key" to determine the row from Source 1 that is compared to a row from Source 2. The match key consists of one or more corresponding columns from each source. These match key columns are used to determine if two rows are the same or different. That is, when the values in the match key columns are the same, the remaining columns in the pair of rows are compared.

Compare attempts to use the primary key defined for a source table as the match key, evaluating the primary keys for each pair of tables, as follows:

- If both tables have the same primary key, the primary key is used.
- If both tables have a primary key and one primary key consists of columns common to both, the common primary key is used.
- If both tables have a primary key and either is valid as the match key, you are prompted to select one.

- If neither table has a primary key or if a Column Map does not include a primary key column, you are prompted to define the match key, which is stored in the Compare Definition. The steps used to define a match key are the same as those used to define a primary key.

You can also define a match key that you can use with the current Compare Process or save in the Compare Definition. For a single table comparison, use the Specify COMPARE Sources panel to indicate a user-defined match key will be used. For a multiple table comparison, use the **MKEY** command in the **COMPARE Process Table Map** to indicate the tables that will use a user-defined match key. After completing each panel, you will be prompted to define the match key.

The columns in a match key can correspond directly by name and compatible attributes (for example, CUST_ID in both tables) or the correspondence can be established using a Column Map. (When a Column Map is used, the names may be different but the attributes must be compatible.)

Processing Fixed Arrays

Special considerations apply when you compare data that includes fixed array elements. Be aware of these limitations. Mapped array columns must have:

- the same number of elements
- the same number of levels (dimensions)
- compatible data types and lengths

When you display the results of a comparison that includes mapped fixed array columns, Optim lists each array element in a separate column. Column headings have the following format: *name_n_n...* where:

name_n_n
: is the name of the fixed array column as defined in the column map. The name is followed by numerics for the occurrence of the element in the array, and the table dimension. Underscores (_) separate the name from the numerics.

In this example, a Legacy Table is defined with a 2-dimension fixed array.

```
Command ===>

User defined I/O Exit   ===>                                        Row 1 of 5
Cmd Level/Field Name              Type  Len  Occur Column Name
--- ----------------------------- --- ----- ----- ---------------------------
*** ******************************** TOP ************************************
___  1 POLICY                          70        POLICY
___   5 POLICY-ID                  CHR    2        POLICY_ID
___   5 POLICY-INFO                      17     4  POLICY_INFO
___    10 POLICY-NAME              CHR    8        POLICY_NAME
___    10 POLICY-RATE              CHR    3     3  POLICY_RATE
*** ***************************** BOTTOM ***********************************
```

When you use this table in a comparison and display the results, the column headings will appear as shown in the following example:

```
------------------- Optim: Browse (Source 1 Names Shown)  ----------------------
Command ===>
                                                        SUBSYS: DD9F
Cmd Chg Src == Table: TESTNEST(T1) ================================================
            POLICY_ID POLICY_NAME_1 POLICY_RATE_1_1 POLICY_RATE_2_1 POLICY_RATE_3_1
            --CH(2)-- ---CH(8)----- -----CH(3)----- -----CH(3)----- -----CH(3)-----
*** ********************************************************************************
___       12    01       POLICY A          100             200             300
*** ********************************************************************************
```

In the example, the column heading displayed as POLICY_RATE_2_1 is for the column defined in the Legacy Table as POLICY_RATE, which is the second element of the first table dimension.

You can use the generated column name with functions that accept a column name as input, such as FIND or LOCK.

Primary Keys and Match Keys

You can create an Optim primary key for a VSAM or sequential file as though the file were a DB2 table. Fields in a record are treated as columns for the purpose of creating a primary key. You can select and browse IMS primary keys, but IMS primary keys cannot be modified.

When a primary key has not been defined for either table in a pair of tables to be compared, you are prompted to define a match key.

Match keys are used to match rows from one source with rows from the other source for the comparison. When available, Compare uses a primary key from a source table for the Compare Process. However, when a usable primary key is not defined in IMS (for an IMS table) or in the Optim Directory for either table, you are prompted to create a match key.

For the most part, the prompts to define match keys are the same as those documented in this section for defining primary keys. When defining a match key, the column names in both sources must match either directly, having the same base name and attributes, or indirectly, using a Column Map to match.

Also note that, unlike database primary keys, match keys need not be based on a unique index and, unlike OPT primary keys, match keys are not available to other users or processes. Match keys reside in the Compare Definition only.

Create or Select a Primary Key

To edit or create a primary key, select Option 6.1 Primary Keys from the Main Menu. You can also select Option 6 DEFINITIONS to display the Choose a Definition Option menu and select Option 1 PRIMARY KEYS. In either case, the Choose a Primary Key panel is displayed.

```
----------------------------- Choose a Primary Key -----------------------
COMMAND ===>

Primary Key:
  Creator ID ===> FOPDEMO                                    >
  Table Name ===> CUSTOMERS                                  >

Primary Key Type                    ===> OPT (P|O-OPTIM, D-DB2, I-IMS, A-All)

Use '_' for DB2 LIKE character      ===> NO (Y-Yes, N-No)
```

This panel includes the following items:

Primary Key
The Creator ID and Table Name for the primary key. You can enter an explicit value, DB2 LIKE syntax, or blanks for these prompts in any combination.

Creator ID
The Creator ID for the table for which the primary key is being created or modified. The default is the previously entered value or the TSO ID of the current user if a Creator ID has never been specified. Specify 1 to 128 characters. To create or edit a generic primary key, type an asterisk at this prompt.

Table Name
The name of the table for which the primary key is defined or modified. Specify 1 to 128 characters.

Primary Key Type

Indicate the type of primary keys to include in the selection list. This value has no effect if you indicate an explicit Creator ID and Table Name. Specify:

P or O
Include primary keys defined in the Optim Directory.

D Include primary keys defined to DB2.

I Include primary keys defined to IMS.

A Include all types.

Use '_' for DB2 LIKE character

Use of the underscore (_) character. Specify Y if the underscore is used as a DB2 LIKE character, or specify N if it is used literally as part of the name.

For example, depending on the use of the underscore character, A_B is a three-character name containing the characters A_B, as entered, or a three-character name that begins with A and ends with B with any valid character in the middle. The default is No, which means that the underscore is not handled as a DB2 LIKE character.

When you supply an explicit **Creator ID** and **Table Name** and the table does not exist, an error message is displayed on the Choose a Primary Key panel. If the supplied name is for a DB2 table, the Optim solution checks the DB2 Catalog to determine if a primary key is defined for the table. If the name references an IMS Legacy Table, the solution checks the IMS DBD to determine if a key is defined for the associated IMS segment.

- If a DB2 or z/OS® primary key for the table exists, the primary key information is displayed. This information cannot be modified.
- If a DBMS primary key for the table does not exist, the solution checks the Optim Directory. If a primary key is defined in the Optim Directory, the information is displayed and can be modified.
- If the database or Legacy Table exists and no primary key information is available, you are prompted to define a new primary key for the table. The new primary key is stored in the Optim Directory.

Requesting a Selection List

A selection list is requested using DB2 LIKE syntax or by leaving one or both of the Primary Key prompts blank (that is, Creator ID and Table Name).

- DB2 tables, IMS Tables, or all are listed, depending on the value specified for Primary Key Type.
- If no tables satisfy the selection list criteria, a message is displayed on the Choose a Primary Key panel.

Primary Key Selection List

This topic describes how to use the Primary Keys selection list.

Here is an example of the primary key selection list.

```
---------------------------- Select Primary Keys -----------------------------
Command ===>                                                 Scroll ===> PAGE

    Line Cmds: S-Select, D-Delete, C-Copy, R-Rename, AT-Attr, I-Info     1 OF 4

Cmd Creator    Table            Type  By       Date
--- -------- ------------------ ---- -------- -------------------
****************************** TOP ************************************
___ *        ORDERS             OPT   DB2MGR   2005-09-03-08.45.50
___ FOPDEMO  CUSTOMERS          DB2   DB2MGR   2005-08-26-16.06.30
___ FOPDEMO  DETAILS            DB2   DB2MGR   2005-07-09-11.55.47
___ FOPDEMO  SHIP_INSTR         OPT   DB2MGR   2005-07-09-11.55.47
****************************** BOTTOM *********************************
```

Tables for which a primary key is defined are listed. The **Creator ID** and table name are displayed. An asterisk (*) in **Creator** indicates a generic primary key. **Type** indicates the source of the primary key, as defined in the Optim Directory (OPT), the DB2 Catalog (DB2), or in an IMS DBD (IMS).

A User Option setting determines whether each primary key's description is displayed. See "User Options" in the *Common Elements Manual* for information about the Selection List Format option.

The Select Primary Keys panel includes the following line commands:

Cmd The line command entry area. Valid line commands are:

- **S** Select a primary key.
- **D** Delete an Optim primary key.
- **C** Copy an Optim primary key.
- **R** Rename an Optim primary key.
- **AT** Modify attributes of an Optim primary key.
- **I** Display information about an Optim primary key

The following messages might be displayed when you enter a line command for a primary key.

- The message *BUSY is displayed when a primary key is being modified by another user.
- The message *NO COLUMN MATCH is displayed when there are no Table columns that match an IMS primary key.

Edit or Browse a Primary Key

When you enter the fully qualified name of a table on the Choose a Primary Key panel, the panel that displays next depends on whether or not a primary key exists for the table, and the type of primary key, if one does exist.

If the table:

- Has a DB2 or IMS primary key, the Browse DB2 Primary Key or Browse IMS Primary Key panel is displayed to allow you to view information about the primary key.
- Has no primary key, the Define Optim Primary Key panel is displayed to allow you to create the primary key.
- Has only an Optim primary key, the Modify Optim Primary Key panel is displayed.

The appearance of these panels is identical. The operation of the Define Optim Primary Key and the Modify Optim Primary Key panels is identical and is described in this section. However, the Browse Primary Key panel is read-only.

Here is an example of the Modify Optim Primary Key panel.

```
--------------------------- Modify OPTIM Primary Key --------------------------
Command ===>                                                 Scroll ===> PAGE

Modify OPTIM Primary Key for FOPDEMO.SHIP_INSTR
Limit Key to 64 Columns and 2000 Bytes

Cmd           Column Name            Data Type
--- ---------------------------- ---------------
*** ********************** TOP **********************
___ SHIP_INSTR_ID                 DECIMAL(5,0)
*** ********************* BOTTOM *********************
```

The Primary Key panel shows the name of the table for the primary key, as well as the following items:

Cmd The line command entry area. Valid line commands on the Define/Modify Optim Primary Key panel are:

D, Dn, or DD
Delete one or more column names. DD is the block form of the command.

I or In Insert one or more blank lines.

M, Mn, or MM
Move one or more column names. MM is the block form of the command.

A or B Destination (After or Before) for the Move line command.

Column Name
The name of each column in the primary key. When editing an Optim primary key, you can specify a maximum of 64 columns with a total of 2000 bytes.

Data Type
The data type for each column in the primary key. The Optim solution automatically displays the data type. This value cannot be modified.

Selection List of Columns

Use the LIST COLUMNS command to display the columns available to be included in the primary key. Columns currently included in the primary key are not listed. For example, type the LIST COLUMNS command to display the Select One or More Columns pop-up for the SHIP_INSTR table.

```
--------------------------- Modify OPTIM Primary Key --------------------------
Command ===>                                                 Scroll ===> PAGE

Modify OPTIM Primary Key for FOPDEMO.SHIP_INSTR
Limit Key to 64 Columns and 2000 Bytes

Cmd         Column Name                    Data Type             1 OF 1
--- -------
*** ******* +----------------Select One or More Columns-----------------+
___ SHIP_IN | Cmd Column Name Data Type 1 OF 3                         |
*** ******* | --- ------------------------------------------- |       |
            | ********************** TOP ********************** |       |
            | ___ SHIP_ID                  SMALLINT             |       |
            | ___ ORDER_SHIP_INSTR         VARCHAR(254)         |       |
            | ___ SHIP_UPDATED             TIMESTAMP            |       |
            | ********************* BOTTOM ********************** |      |
            +-----------------------------------------------------------+
```

For each available column, the name and data type are listed. Use the S line command to select columns for the primary key. When finished, press ENTER or use END to redisplay the Modify Optim Primary Key panel. The selected column names are inserted in the list of columns comprising the key.

Press ENTER or use END to end the Modify Optim Primary Key panel processing. Primary key information is stored in the Optim Directory and used the next time the table is accessed. Any change to the information affects all users.

Match Key Processing

To demonstrate the Compare processing, assume the match key for the pair of CUSTOMERS tables contains only the CUST_ID column as the match key column, and the following data is contained in that column in each source.

```
Source 1              Source 2
00001                 00001
00002                 00002
00003                 00004
```

```
00004                 00005
00006                 00006
00006                 ...
...
```

Matched equal rows

When reviewing the compared data, Compare provides two areas to indicate the results. **Chg** defines the type of change and **Src** identifies the source of the displayed row. The following examples discuss the possible values for **Chg** and **Src** and the comparison results that generate them.

In the following example, the rows containing 00001, 00002, and 00004 in both sources of the CUSTOMERS table contain the same value. These rows are identified as "matched" rows and the remaining columns are compared. When all of the remaining columns match, the rows are "equal". These equal rows are indicated in **Src** as "12," but there is no value for **Chg**.

```
Chg         Src               CUST_ID

            1 2               00001

            1 2               00002

            ...               00004

            1 2
```

Matched rows with a Direct change

When one or more of the remaining columns do not match, the row is marked as having a "direct" change and the row from each source is displayed. For these rows, the source is indicated in **Src** as 1 or 2, and a D for Direct change is inserted in **Chg**, as shown in the following example.

```
Chg         Src               CUST_ID

D            1                00002

D            2                00002
```

Unmatched rows

When one or more of the column values in the match key are not the same, the row is identified as "unmatched." In the following example, the rows containing 00003 in Source 1 and 00005 in Source 2 do not match a row in the other source. These rows are "unmatched" and identified as such by the value in **Src**; the remaining columns are not compared.

```
Chg         Src               CUST_ID

             1                00003

             ...              00005

             2
```

Duplicate Match Key values

In the following example, there are two rows with a CUST_ID of 00006 in Source 1. These are identified as duplicate match key rows. Although Source 2 may have one or more corresponding rows with identical match key column values, these rows are not compared because it is impossible to determine which rows are to be paired. Duplicate rows are identified by brackets.

```
Chg         Src               CUST_ID

 /           1                00006
```

```
|            1                00006
\            2                00006
```

Comparing Data that is Uniquely Keyed

Compare processes uniquely keyed tables as described:

- If rows exist in source 1 and source 2, with keys that match and identical contents, they are reported as “Equal”.
- If rows exist in source 1 and source 2, with keys that match but different contents, they are reported as "Changed".
- If rows exist only in source 1 or source 2, with keys that do not match, they are reported as "Unmatched".
- If multiple rows exist in either source 1 or source 2, with the same key they are reported as “non-unique Match Keys” and no comparison is made.

Comparing Data that is Not Uniquely Keyed - Soft Match

When comparing data that is not uniquely keyed or non-keyed, you have the option to use soft match for the comparison. Soft match means you decide the criteria to determine whether rows match. The Optim solution will attempt to match rows that have the same key or are non-keyed and do not have identical values. Soft match cannot be used for uniquely-keyed data.

For rows to be considered matched, you can determine the minimum columns required. That is, you can specify the percentage of column values that must be equal (in each source) for the rows to be considered a match. The Optim solution calculates the percentage from the number of columns that are not keyed. A row is matched to the row that has the highest matching percentage.

You can limit the soft match processing in two ways:

- Specify a minimum percentage of columns that can be considered a match.
- Set a maximum number of comparisons that can be performed. When this value is reached, the rows compared will be matched and those not compared will be considered unmatched.

See “Match Options Panel” on page 53for details on soft match processing.

Compare Process Menu

When **Option 8 COMPARE** is selected from the Main Menu, the following panel is displayed to begin prompting for the information necessary to perform the Compare Process.

```
----------------------------- COMPARE Process ---------------------------------
OPTION ===>
                                                                    SQLID ===>
1 SINGLE   - Compare One Table to Another                           SUBSYS ===>
2 MULTIPLE - Compare Two Sets of Tables                             LOCATION ===>
3 PERFORM  - Specify COMPARE Parameters and Perform COMPARE

B BROWSE   - Browse Results of Previous COMPARE
R REPORT   - Generate Report from Previous COMPARE

Type of Compare Definition to Use for COMPARE ===> T (P-Perm, T-Temp)

If Permanent, Specify New or Existing Compare Definition Name:
  GROUP ===>
  USER ===>
  NAME ===>

Use '_' for DB2 LIKE Characters ===> N (Y-Yes, N-No)
```

The following options are available on the COMPARE Process panel:

1 - SINGLE
Compare one pair of tables. The source for each table can be an Extract File, DB2 or IMS table, VSAM or sequential file.

2 - MULTIPLE
Compare a set of related data that resides in multiple tables. The source for each set of data can be an Extract File, DB2 or IMS table, VSAM or sequential file. The data from the tables can be defined by an Access Definition or, for Source 2, based on the specifications from Source 1.

3 - PERFORM
Perform the Compare Process using an existing Compare Definition.

B -BROWSE
Browse an existing Compare File. Special facilities are available when browsing to highlight the differences, to scroll, and to join to related data in other tables.

R - REPORT
Generate and print a report based on the contents of the Compare File. This report can include summary information, details, or both. Summary information includes statistics about the Compare Process. The detail report includes the data rows from each table from each source and the status of those rows.

In addition to these options, this panel prompts for the following:

SQLID
The current SQLID. Modify this value to connect using a different SQLID.

SUBSYS
The current DB2 subsystem. Modify this value to connect to a different DB2 subsystem. When connecting to a remote subsystem, this value should be the local subsystem where the remote location is defined.

LOCATION
The remote location. This prompt is displayed if remote access is available. Specify a value to connect to a remote DB2 subsystem. You can use a percent sign (%) to obtain a selection list of available locations. If the connection fails, the session is restarted and the Main Menu is redisplayed. If you leave this prompt blank, the local subsystem is assumed.

Type of Compare Definition to Use
Specify whether the Compare Definition is to be used for the current Compare Process only or is to be stored in the Directory. If stored, it is available for reuse and to other users. Specify:

P PERManent. Compare Definition is to be stored in the Directory.

T TEMPorary. Compare Definition is not saved.

If Permanent, Specify New or Existing Compare Definition Name
Specify the name of the Compare Definition to be used. The name is composed of three parts:

GROUP
The 1- to 8-character group identifier. The default is the previously entered value or the TSO ID of the current user if a GROUP has never been specified.

USER The 1- to 8-character user name. The default is the previously entered value or the user's DB2 SQLID if a USER has never been specified.

NAME
The 1- to 12-character name of the Compare Definition. You can use DB2 LIKE syntax or leave one or more parts blank to obtain a selection list.

Use '_' for DB2 LIKE Character
Specifies whether or not the underscore ('_') is to be used as a DB2 LIKE character or used

literally as part of the name. For example, 'A_B' could be assumed to be a three-character name containing the characters A_B as entered or a three-character name that begins with A, ends with B, and has any valid character in the middle. The default is NO which means that the underscore is not handled as a DB2 LIKE character.

Using an Existing Compare Definition

To use an existing Compare Definition, specify **P** for permanent in **Type of Compare Definition to Use for COMPARE** and the name of the Compare Definition. For example, during the iterative testing process you need to compare the same sets of data each time you execute your revised application. Rather than redefine the two sources, just specify the same Compare Definition each time you execute the Compare Process.

Subsequent Use and Reuse

The **GROUP**, **USER**, and **NAME** information specified on this panel is stored in your profile. The next time you display this panel, those values are automatically provided. You may change any of these values. If you press ENTER without supplying a fully qualified name, a selection list of all Compare Definitions for the named **GROUP**, **USER**, and **NAME** are displayed. If there are no Compare Definitions to satisfy these values, a message is displayed prompting you to re-specify.

Primary Commands

The END command can be used to return to the Main Menu. Any values specified on this panel are retained in your profile. The CANCEL command returns to the previous menu but your profile is not updated. The OPTIONS command can be used to display the User Options panel or, if authorized, the Site Options panel.

Defining a Table for IMS Data

The Optim solution uses two definitions to process IMS data. These definitions are **IMS Environment Definitions** and **IMS Retrieval Definitions**. An IMS Environment Definition for each IMS database is required in order to create a Table with IMS data. An IMS Retrieval Definition for each DBD within the IMS database provides defaults for processing IMS data.

IMS Environment Definitions

An IMS Environment Definition identifies the DBD, PSB, and IMS Load Libraries to be used during processing and the information needed to process the data when online to IMS. In fact, Optim requires an Environment Definition as a prerequisite to creating both a Table Definition for an IMS segment and an IMS Retrieval Definition.

A Table Definition name is in two parts, envdef.tablename. The first portion of the name references the Environment Definition used with the Table in processing. If you attempt to create an IMS Table Definition, but an Environment Definition with a matching name does not exist, The Optim solution prompts you to create the needed Environment Definition before it displays the Table Definition editor.

To create an Environment Definition from the Specify Copybooks for Table panel, use the **ENVIRONMENT** primary command to display the Define IMS Environment panel. The first segment of the Table name is used as the new Environment Definition name.

Create an IMS Environment Definition

To create or edit an Environment Definition using the DEFINITIONS option on the Main Menu, select **Option 7 IMS ENVIRONMENT** on the Choose a Definition Option menu.

The Choose an IMS Environment panel is displayed and prompts for name of the Environment Definition.

```
--------------------------- Choose an IMS Environment -------------------------
Command ===>

IMS Environment:
  Environment Name ===>

Use '_' for DB2 LIKE character ===> NO     Y-Yes, N-No
```

This panel includes the following items:

Environment Name
Specify the 1- to 8-character name of the IMS Environment Definition. Use DB2 LIKE syntax to display a selection list. Leave blank or enter "%" to display a list of all IMS Environments.

Use '_' for DB2 LIKE character
Indicate whether the underscore (_) is used as a DB2 LIKE character or literally, as part of the name.

Note: The Environment Definition name must match the name of the Creator ID used to create the Table referencing IMS data.

If you specify an explicit value for the **Environment Name**, and:

- The Environment Definition exists, Compare displays the named IMS Environment on the Modify IMS Environment panel.
- The Environment Definition does not exist, Compare displays the Define IMS Environment panel with all fields blank.

Selection List of IMS Environment Definitions

When you use DB2 LIKE syntax in **Environment Name** or leave it blank, a selection list is displayed.

```
---------------------- Choose An Environment Definition ----------------------
Command ===>                                                 Scroll ===> PAGE

     Line Cmds: S-Select D-Delete C-Copy R-Rename, AT-Attribute 1 OF 24

              ------- Last Modified -------
Cmd Env Name   By     Date
--- -------- -------- -------------------
****************************** TOP ***********************************
___ FOPIMS    FOPLEG 2000-05-07 11.48.21
___ FOPIMS01  FOPLEG 2000-11-13 12.30.15
___ FOPIMSO2  FOPLEG 2001-02-05 10.03.55
___ FOPIMS03  FOPLEG 2001-02-06 09.16.41
****************************** BOTTOM ********************************
```

The panel includes the following items:

Cmd Line command area. Valid line commands are:

S Select an Environment Definition.

D Delete an Environment Definition. After deleting, the message *DELETED is displayed

C Copy an Environment Definition to create a new one. The Copy Environment panel prompts for the name of the new Environment Definition. After copying, the message *COPIED is displayed.

R Rename an Environment Definition. The Rename Environment panel prompts for the new name of the selected Environment Definition. After renaming, the message *RENAMED is displayed.

AT Modify attributes of an Environment Definition. The Object Attributes panel allows you to edit the Environment Definition description and security status (if authorized).

Env Name
The 1- to 8-character name of the Environment Definition.

By The TSO ID of the user that last modified the Environment Definition.

Date The date and time the Environment Definition was last modified.

Select an Environment Definition

On the selection list, type the S line command to display the Modify IMS Environment panel and select an Environment Definition to modify.

Copy an Environment

To copy an Environment Definition, type the C line command next to the name of the source Environment Definition to display the Copy Environment panel.

```
+-------Copy Environment-------+
|                               |
| Existing Name: FOPIMS         |
|                               |
| New Name:                     |
|                               |
| Environment ===> FOPIMSO1     |
+-------------------------------+
```

Rename the Environment

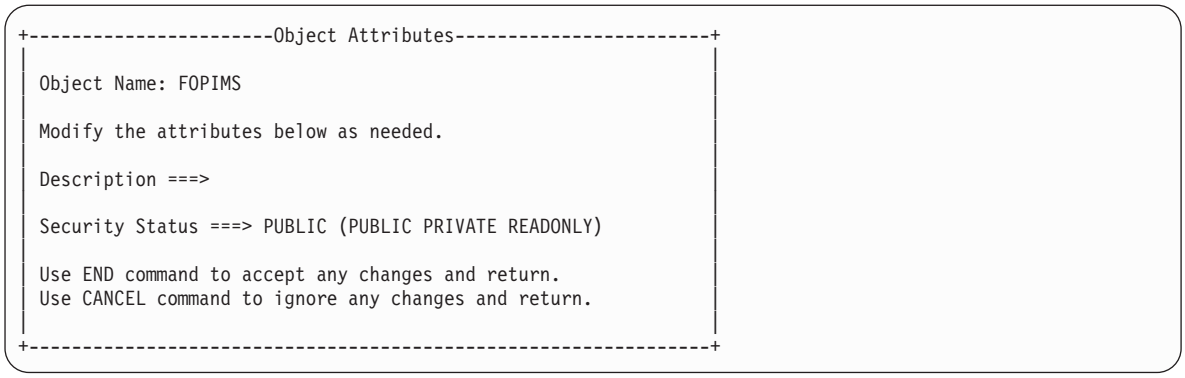
To rename an Environment Definition, type the R line command next to the name of the Environment Definition to display the Rename Environment panel.

The Rename Environment panel displays the current name of the Environment Definition and prompts for a new name.

```
+------Rename Environment------+
|                               |
| Existing Name: FOPIMS         |
|                               |
| New Name:                     |
|                               |
| Environment ===> FOPIMS02     |
+-------------------------------+
```

Modify the Object Attributes

To modify the description and security status attributes of an Environment Definition, type the AT line command next to the name of the Environment Definition. The description and security status are specified on the Object Attributes panel.



The Object Attributes panel provides a 40-character area to display and edit the **Description**. An 8-character area is also available to specify an appropriate **Security Status**, as follows:

Public Anyone can edit and use.

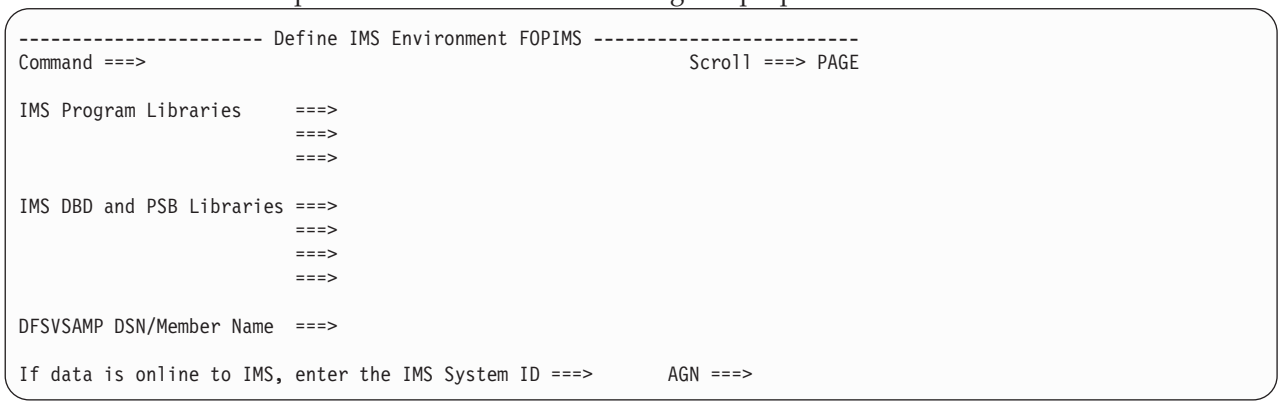
Private
Only the owner can edit and use.

Readonly
Anyone can use, only the owner can edit.

A site option determines whether Security Status is available. If it is not available, Security Status is not displayed. For additional information about the Object Attributes panel, see the *Common Elements Manual*.

Defining an IMS Environment

The Define IMS Environment panel allows you to specify the IMS Program Libraries, the IMS DBD and PSB Libraries, and the IMS System ID and AGN when data is online to a control region. Specifying this information allows Compare to access IMS data utilizing the proper libraries.



This panel includes the following items:

IMS Program Libraries
Enter the names of the IMS Program Libraries to include in the Environment Definition. These libraries define and load the libraries required to run IMS. You may enter up to three library names. You may also enter wildcards to select from a list.

Note:

- All data sets defined as IMS program libraries must be APF-authorized.
- If you have authorized the Optim SFOPLLIB data set, add it to this list.

- If your site uses IMS dynamic allocation, include the name of the data set containing the dynamic allocation load modules.
- If the IMS Program Libraries are already available (e.g., by Linklist or other method), you do not need to enter the library names.

IMS DBD and PSB Libraries
Enter the names of up to four IMS DBD and PSB Libraries to include in the Environment Definition. You must enter at least one library name. You may also enter wildcards in order to select from a list.

The DBD and PSBs within the libraries define the structure of the IMS database as well as the method Compare uses to access the data.

DFSVSAMP DSN/Member Name
Specify the name of the DFSVSAMP data set member that contains the IMS buffer parameters to be used when running in batch mode.

Enter the DFSVSAMP data set name to display the Member Selection List panel and select the DFSVSAMP data set member.

You may also enter wildcard characters to access the Select PDS Library pop-up window and select the DFSVSAMP data set from a list.

IMS System ID
Specify the IMS System ID required to access the IMS data when allocated to a control region (that is, the data is online to IMS).

AGN Specify the IMS Application Group Name required to access the IMS data when allocated to a control region, if required by your IMS database.

When you have completed the Environment Definition, use END to save it. The panel from which the Modify/Define IMS Environment panel was invoked is redisplayed. If any referenced libraries cannot be found, an error is displayed.

IMS Retrieval Definition

An IMS Retrieval Definition allows you to optionally define appropriate default settings for processing IMS data. By creating a Retrieval Definition, you establish the PSB and PCB defaults used to access the specified DBD, and you associate the segments within the DBD with default database data sets to be referenced during processing.

The Retrieval Definition name is in two parts (*envdef.dbdname*), the name of the Environment Definition (which is also the name of the Creator ID of the Table Definition) and the name of an IMS DBD referenced by the Environment Definition.

The Optim solution uses the information in the Retrieval Definition to provide default values when the Table is specified in an Access Definition or a Table Map. If you have not created a Retrieval Definition, you must manually enter the information when you create the Access Definition.

To create a Retrieval Definition directly from the Modify Table panel, use the DATASOURCE primary command to display the Provide Retrieval Definition for DBD panel using the Creator ID of the Table and the specified DBD as the two-part Retrieval Definition name.

Create an IMS Retrieval Definition

To create or edit a Retrieval Definition using the DEFINITIONS option on the Main Menu, select **Option 8 IMS RETRIEVAL** on the Choose a Definition Option menu. The Choose an IMS Retrieval Definition panel is displayed, and prompts for the Environment Name and the IMS DBD Name.

```
---------------------- Choose an IMS Retrieval Definition --------------------
Command ===>
IMS Retrieval Definition:
  Environment Name ===> %
  IMS DBD Name ===>

Use '_' for DB2 LIKE character ===> NO Y-Yes, N-No
```

This panel includes the following items:

Environment Name
Specify the 1- to 8- character name of the IMS Environment Definition that references the specified IMS DBD. Use DB2 LIKE syntax to display a selection list. Leave blank or enter "%" to display a list of Retrieval Definitions for all Environment Definitions with the specified DBD name.

IMS DBD Name
Specify the 1- to 8- character name of the desired IMS DBD in a DBD library that is referenced in the specified IMS Environment Definition. Use DB2 LIKE syntax to display a selection list. Leave blank or enter "%" to display a list of Retrieval Definitions for all IMS DBDs with the specified Environment Name.

Use '_' for DB2 LIKE character
Indicate whether the underscore, '_', is used as a DB2 LIKE character or literally, as part of the name.

If explicit values are specified for **Environment Name** and IMS **DBD Name** and

- The Retrieval Definition exists, Compare displays the named Retrieval Definition on the Provide Retrieval Information for DBD panel.
- The Retrieval Definition does not exist, but the specified IMS DBD is referenced in the Environment Definition, Compare displays the Provide Retrieval Information for DBD panel with all prompts blank.
- If the Retrieval Definition does not exist, and the specified IMS DBD is not referenced in the Environment Definition, an error is displayed.

Selection List of Retrieval Definitions

Use DB2 LIKE syntax or leave a prompt blank to request a selection list. Compare displays a list of Retrieval Definitions on the Choose a Retrieval Definition panel.

```
---------------------- Choose A Retrieval Definition -------------------------
Command ===>                                                Scroll ===> PAGE

    Line Cmds: S-Select, D-Delete, C-Copy, R-Rename, AT-Attribute   1 OF 2

    -- Retrieval Def --         ------- Last Modified -------
Cmd Env Name    DBD Name        By         Date
--- --------    --------     --------     -------------------
****************************** TOP ***********************************
___ FOPDEMO     FOPDEPDB      FOPNA       2002-11-19 14.34.13
___ FOPDEMO     FOPJOBDB      FOPNA       2002-11-19 14.34.13
***************************** BOTTOM *********************************
```

If no Retrieval Definitions satisfy the criteria, a message is displayed on the Choose an IMS Retrieval Definition panel.

This panel includes the following items:

Cmd Line command area. Valid line commands are:

S Select a Retrieval Definition.

D Delete a Retrieval Definition. After deleting, the message *DELETED is displayed.

C Copy a Retrieval Definition to create a new one. The Copy Retrieval Definition panel prompts for the name of the new Retrieval Definition. After copying, the message *COPIED is displayed.

R Rename a Retrieval Definition. The Rename Retrieval Definition panel prompts for the new name of the selected Retrieval Definition. After renaming, the message *RENAMED is displayed.

AT Modify attributes of a Retrieval Definition. The Object Attributes panel allows you to edit the Retrieval Definition description and security status (if authorized).

Env Name
The 1- to 8- character name of the Environment Definition.

DBD Name
The 1- to 8- character name of a DBD referenced in the Environment Definition.

By The TSO ID of the user that last modified the Retrieval Definition.

Date The date and time the Retrieval Definition was last modified.

Select a Retrieval Definition

On the selection list, type the S line command to select a Retrieval Definition. The Provide Retrieval Information for DBD panel displays. See “Provide Retrieval Information for DBD” for more information.

Copy a Retrieval Definition

On the selection list, type the C line command to display the Copy Retrieval Definition panel to copy a Retrieval Definition.

The Copy Retrieval Definition panel displays the name of the original Retrieval Definition, Environment Definition, and DBD name, and prompts for a new Environment Definition and DBD name.

```
+--------Copy Retrieval Definition-------+
|                                         |
|                                         |
| Existing Name: FOPJS.FOPDEPDB           |
|                                         |
| New Name:                               |
|                                         |
| Environment ===> FOPJS                  |
| DBD Name    ===> FOPDEPDB               |
+-----------------------------------------+
```

Rename the Retrieval Definition

On the selection list, type the R line command to display the Rename Retrieval Definition panel to rename a Retrieval Definition.

The Rename Retrieval Definition panel displays the current name of the Retrieval Definition, Environment Definition, and DBD name, and prompts for a new Environment and DBD name.

```
+--Rename Retrieval Definition--+
|                               |
| Existing Name: FOPJS.FOPDEPDB |
|                               |
| New Name:                     |
|                               |
| Environment ===> OPTJS        |
| DBD Name ===> OPTDEPDB        |
+-------------------------------+
```

Modify Object Attributes

To modify the description and security status attributes of a Retrieval Definition, type AT in **Cmd** next to the name of the Retrieval Definition. The description and security status are specified on the Object Attributes panel.

```
+-----------------------Object Attributes-----------------------+
|                                                               |
| Object Name: FOPDEMO.FOPDEPDB                                 |
|                                                               |
| Modify the attributes below as needed.                        |
|                                                               |
| Description ===>                                              |
|                                                               |
| Security Status ===> PUBLIC    (PUBLIC PRIVATE READONLY)      |
|                                                               |
| Use END command to accept any changes and return.             |
| Use CANCEL command to ignore any changes and return.          |
|                                                               |
+---------------------------------------------------------------+
```

The Object Attributes panel provides a 40-character area to display and edit the **Description**. An 8-character area is also available to specify an appropriate **Security Status**, as follows:

Public Anyone can edit and use.

Private
Only the owner can edit and use.

Readonly
Anyone can use, only the owner can edit.

A site option determines whether Security Status is available. If it is not available, Security Status is not displayed.

Provide Retrieval Information for DBD

The Provide Retrieval Information for DBD panel allows you to provide the information necessary to access the IMS data by specifying the PSB and PCB to be used to access the referenced DBD and associate each DD with an IMS Database Dataset Name. It also allows you to use DBRC and IMS logging for processing in DL/I mode.

```
-------------- Provide Retrieval Information for DBD: FOPDEPDB ----------------
Command ===>                                                 Scroll ===> PAGE

                                                                      1 of 4
Default PSB Name       ===> FOPDEPPA              PCB Number     ===> 1
Default Dataset Prefix ===>                       Default IMS ID ===> IMSA
Use DBRC               ===> N                     (Y or N)
IMS Log Dataset Name   ===>

Cmd  Segment   DD Name   Associated IMS Database Dataset Name
---  --------  --------  ------------------------------------------------------
*** ********************************* TOP ************************************
___ DEPT       FOPDEPDB 'FOPQA.RT.FOPDEMO.FOPDEPDB'
___ EMPLOYEE
___ POSITION
___ -INDEX     FOPDEPIX 'FOPQA.RT.FOPDEMO.FOPDEPIX'
*** ******************************** BOTTOM ***********************************
```

This panel includes the following items:

Default PSB Name

Specify the 1- to 8-character name of the default PSB in a PSB library referenced in the associated Environment Definition. The PSB provides access to the IMS services required to access the database records.

Specify an asterisk to generate a selection list that includes PSBs in the associated Environment Definition.

You can override the default PSB when creating an Access Definition or Table Map that references the Retrieval Definition.

PCB Number

Specify the relative number of the database PCB within the specified PSB that grants the authorization to manipulate the data. Specify an asterisk to generate a selection list of PCBs in the specified PSB. For an overview of this selection list, see "PCB Selection List."

You can override the default PCB number when creating an Access Definition or Table Map that references the Retrieval Definition.

Default Dataset Prefix

Optional 1- to 8-character prefix used when specifying Associated IMS Database Dataset Names.

Default IMS ID

Specify the default IMS System ID used to access the IMS data during processing.

Use DBRC

This entry is valid only for IMS processing in DL/I mode (i.e, when an IMS ID is not specified). If appropriate, enter Y for yes to use Database Recovery Control (DBRC) to control logging and perform database recovery, otherwise enter N for no. IMS uses the online log datasets (OLDS) if the database is accessed in BMP or DBB mode.

The default for a HALDB (High Availability Large DataBase) is Y, and that entry cannot be changed.

DBRC use is optional for a non-HALDB, such as HIDAM, HDAM, HISAM, and so on. Thus, you may specify **Y** for a non-HALDB, but it is not required.

IMS Log Dataset Name

If appropriate, specify the dataset name for the IMS Log used to perform database recovery. This dataset name is used to dynamically allocate the IEFRDER dataset. It is recommended that you provide a GDG dataset for the IMS log because the Optim solution may make multiple calls to IMS while traversing through the Tables defined in an Access Definition or Table Map. This would result in the Log dataset being overwritten if a sequential dataset is used. If a GDG is used, the Optim solution will allocate one GDS for each invocation of IMS, which will prevent the IMS log from being overwritten.

If a PSB with a Processing Option (PROCOPT) other than G (for Get) is used while accessing a HALDB in DL/I mode that is, an IMS region name is not specified), you must specify the name of the dataset to be allocated for DD Name IEFRDER.

If you specify an IMS Log Dataset Name, when you exit the Provide Retrieval Information for DBD panel, the Allocate Dataset panel automatically displays. You must provide sufficient Primary and Secondary space units on that panel to allocate the IEFRDER dataset. Failing to do so will cause IMS to abort processing and lock the database from further updates until a recover/rollback is done.

Segment

Lists the segments within the specified DBD. This column cannot be modified.

DD Name

Lists the names of each DD or Data Definition (that is, the physical data sets) associated with each segment. This column cannot be modified.

Associated IMS Database Dataset Name

Enter the location of the IMS Database Dataset associated with each DD Name in the DBD. This data is then associated with the named Table during processing.

A Site Option (Require IMS Data Set Names) determines whether you can omit the data set name to allow IMS to dynamically allocate the data set. All users can specify '$MDA' as the data set name to choose dynamic allocation, regardless of this Site Option.

If you are including a data set from an IMS database with multiple data set groups, for each database you must either:

- specify all the data set names, or
- enable dynamic allocation by leaving the data set names blank (if your site option `Require IMS Data Set Names` is set to `N`), or
- enable dynamic allocation by specifying `'$MDA'` for all the data set names

You must use the same option for all the data set groups in a database.

You can override the default IMS Database Dataset name when creating an Access Definition or Table Map that references the Retrieval Definition.

Note: You do not have to specify a dataset name for a DEDB or HALDB because the appropriate dataset name will already be known to the IMS subsystem.

PCB Selection List

When you enter an asterisk in the PCB Number prompt, the PCB Selection List is displayed, allowing you to select a PCB for the specified PSB. The selection list displays the available PCB numbers, the number of segments in each PCB, and the first six segment names in each PCB.

```
+-------------------PCB Selection List for PSB SALHIDMA--------------------+
| Cmd PCB Seg Segment Names (First 6 segment names listed)          1 OF 2 |
| --- ------------------------------------------------------------------- |
| ****************************** TOP ********************************** |
| ___ 3     6 SALES   CUST   SHIPT   SHIPI   ORDERS   DETAILS             |
| ___ 5     4 SALES   CUST   ORDERS  DETAILS                             |
| *************************** BOTTOM ********************************** |
+-------------------------------------------------------------------------+
```

When you have completed defining the Retrieval Definition, use END to save it and redisplay the panel from which the Provide Retrieval Information for DBD panel was invoked. If the PSB name, PCB number, or Associated IMS Database Dataset cannot be found, an error is displayed.

Defining a table for VSAM or Sequential File Data

An IMS table, VSAM or sequential file must be associated with a specific data source (that is, the location of the data and how the data is accessed) in order to be used in a Compare Process. For VSAM or sequential files, you must provide the data set name. For IMS, you must provide the PSB and PCB for the source table's DBD, as well as the data set name.

The Associate Tables to Data Sources panel contents vary, depending on whether the tables reference IMS data, VSAM or sequential files, or both.

The Associate Tables to Data Sources panel shows the default data set name (if any) for the VSAM Table.

Choose a Table

To create or edit an I/V/S Table using the **DEFINITIONS** option on the Main Menu, select **Option 6** LEGACY TABLES on the Choose a Definition Option menu. The Choose a Legacy Table panel is displayed and prompts for the Creator ID and Table Name.

```
----------------------------- Choose a Legacy Table --------------------------
Command ===>

Legacy Table:
  Creator ID ===>
  Table Name ===>

Use '_' for DB2 LIKE character ===> NO Y-Yes, N-No
```

This panel includes the following items:

Creator ID
Specify the Creator ID. Use DB2 LIKE syntax or leave Creator ID blank to display a selection list.

Each IMS Legacy Table must have a matching Environment Definition. The name of the matching Environment Definition and the Creator ID are the same. The Legacy Table cannot be defined without a matching Environment Definition. If the matching Environment Definition has not been predefined, you can create it by using the ENV command from the Choose a Definition Option menu. See "Defining a Table for IMS Data" on page 32 for information on creating an Environment Definition.

Table Name
Specify the name of the Legacy Table to define or modify. Use DB2 LIKE syntax or leave **Table Name** blank to display a selection list.

Use '_' for DB2 LIKE character
Indicate whether the underscore (_) is used as a DB2 LIKE character or literally, as part of the name.

If you specify explicit values for **Creator ID** and **Table Name** and

- The Legacy Table exists, Compare displays the named Legacy Table on the Modify Legacy Table panel.
- The Legacy Table does not exist, Compare prompts for information needed to create a new Legacy Table on the Specify Copybooks for Legacy Table panel. For more information, see "Specifying Legacy Table Copybooks" on page 44.

Selection List of Legacy Tables

When you use DB2 LIKE syntax or leave Creator ID or Table Name blank, a selection list is displayed.

```
----------------------------- Select Legacy Table------------------------------
Command ===>                                                  Scroll ===> PAGE

    Line Cmds: S-Select, D-Delete, C-Copy, R-Rename, AT-Attr         1 OF 4

    ------ Legacy Table -------  ------ Last Modified -------
Cmd Creator      Table Name        By             Date
--- -------- ------------------ -------- -------------------
***************************** TOP ********************************
___ FOPLEG    CUSTOMERS          FOPDEMO 2000-11-13 10.57.15
___ FOPLEG    ITEMS              FOPDEMO 2000-11-13 10.09.35
___ FOPLEG    ITEMSX             FOPDEMO 2000-11-13 10.57.32
___ FOPLEG    ORDERS             FOPDEMO 2000-11-13 10.09.07
************************** BOTTOM ********************************
```

This panel includes the following items:

Cmd Line command area. The possible line commands are:

S Select a Legacy Table.

D Delete a Legacy Table. After deleting, the message *DELETED is displayed.

C Copy a Legacy Table to create a new one. The Copy Legacy Table panel prompts for the name of the new Legacy Table. After copying, the message *COPIED is displayed.

R Rename a Legacy Table. The Rename Legacy Table panel prompts for the new name of the selected Legacy Table. After renaming, the message *RENAMED is displayed.

AT Modify attributes of a Legacy Table. The Object Attributes panel allows you to edit the Legacy Table description and security status (if authorized).

Creator
The Creator ID that qualifies the Table Name.

Table Name
The name of the Legacy Table.

By The TSO ID for the user that last modified the Legacy Table.

Date The date and time the Legacy Table was last modified.

Select a Legacy Table

On the selection list, use the S line command to select a Legacy Table to modify. The Modify Legacy Table panel is displayed. For more information about modifying the Legacy Table, see "Define Legacy Tables Panel" on page 45.

Copy a Legacy Table

To copy a Legacy Table, type C in Cmd next to the name of the source Legacy Table. The following figure shows the Copy Legacy Table panel.

```
+-----------Copy Legacy Table-----------+
|                                       |
| Existing Name: FOPLEG.CUSTOMERS       |
|                                       |
| New Name:                             |
|                                       |
|    Creator ID ===> FOPLEG             |
|    Table Name ===> CUSTOMERS          |
+---------------------------------------+
```

The Copy Legacy Table panel displays the name of the original Legacy Table and prompts for a new Creator ID and Table Name.

Rename the Legacy Table

To rename a Legacy Table, type R in Cmd next to the name of the Legacy Table to display the Rename Legacy Table panel.

The Rename Legacy Table panel displays the current name of the Legacy Table and prompts for a new Creator ID and Table Name.

```
+----------Rename Legacy Table----------+
|                                       |
| Existing Name: FOPLEG.CUSTOMERS       |
|                                       |
| New Name:                             |
|                                       |
|    Creator ID ===> FOPLEG             |
|    Table Name ===> CUSTOMERS          |
+---------------------------------------+
```

Specifying Legacy Table Copybooks

The Specify Copybooks for Legacy Table panel prompts you for the copybook(s) used to define the Legacy Table. Compare parses the copybooks to populate the Legacy Table editor. If you do not specify any copybooks, you must define the Legacy Table manually.

Here is an example of the Specify Copybooks for Legacy Table panel.

```
----------------Specify Copybooks for Legacy Table: FOPLEG.ITEMS ---------------
 Command ===>

 INSTRUCTIONS:  Specify the name(s) of the Copybook members containing the record
                description to be used for this Legacy Table.  All members will be
                concatenated before being parsed.

  Legacy Table Type ===> I    I-IMS F-File (VSAM or Sequential)

  Copybook Language ===> C    C-COBOL P-PL/I

  Copybook(s) ===>
              ===>
              ===>
              ===>
              ===>
```

The Specify Copybooks for Legacy Table panel includes the following options:

Legacy Table Type
The type of data to be used with the Legacy Table. Specify:

I Data in an IMS database.

F Data in sequential or VSAM files.

Copybook Language
The language used to create the copybooks. Specify **C** for COBOL or **P** for PL/I.

Copybook(s)
The names of copybooks used as a source for the Legacy Table. You can enter up to five copybook names. You can use wildcards to select one or more copybooks from a list.

Note: Compare remembers your entries and presents the same information whenever you invoke the Specify Copybooks for Legacy Table panel and subsequent panels.

About Copybooks

Copybooks processed by Optim Legacy must be valid COBOL or PL/I definitions. If an error is detected, a message is issued identifying the specific problem and relevant line number. Optim Legacy supports field redefinitions, and Occurs and Occurs Depending On (ODO) clauses, including nested ODOs. Level 66 (RENAMES) are not supported

To create a Legacy Table manually, leave all Copybook entries blank. Compare displays the Define Legacy Table panel in Edit mode, without populating any Field names. This allows you to manually create the Legacy Table.

You can enter as many as five copybook names on the Specify Copybooks for Legacy Table panel. If a desired copybook is not in the default copybook library, you must provide both the copybook data set and member names.

- If the copybook member does not exist in the default or specified copybook library, an error message is displayed.
- If you do not use single quotes around the copybook library (and member) name, the Data Set Prefix, specified as a User Option, is used as the prefix.

Select Copybook Library

Enter a percent (%) wildcard to display the Select Copybook Library panel, allowing you to select a copybook library other than the Default Copybook Library, specified in Legacy Options.

(For more information, see the *Common Elements Manual*.) You may also place a wildcard at the end of a fully qualified data set name to create a more limited selection list. For example, entering 'FOPDEMO.RT.%' might return the following selection list.

```
+-----------------Select Copybook Library------------------+
|                                                          |
| Cmd      Data Set Name                           1 OF 4  |
| --- ------------------------------------------           |
| ********************** TOP **********************       |
| ___ FOPDEMO.RT.LOCAL.CLIST                               |
| ___ FOPDEMO.RT.LOCAL.EXEC                                |
| ___ FOPDEMO.RT.TEST.CLIST                                |
| ********************* BOTTOM **********************      |
+----------------------------------------------------------+
```

Select Record Name

A Legacy Table can include only one 01-level field. If the copybook text (possibly concatenated from multiple copybook members) contains multiple 01-level record definitions, the Select Record Name Used To Create Legacy Table panel is displayed.

```
+---Select Record Name Used To Create Legacy Table----+
|                                                     |
| Cmd  Record Name                    Lang    1 OF 2  |
| --- ------------------------------------            |
| ******************* TOP *****************           |
| ___ SALES                            C              |
| ___ ORDERS                           C              |
| ****************** BOTTOM ***************           |
+-----------------------------------------------------+
```

This panel allows you to select the record name used to populate the Legacy Table Editor. The resulting Legacy Table is displayed on the Define Legacy Table panel. If you select more than one copybook with a 01-level, only the information from the first selected copybook is displayed on the Define Legacy Table panel.

Define Legacy Tables Panel

The Define Legacy Table panel allows you to create or edit a Legacy Table. If a copybook member is specified on the Specify Copybooks for Legacy Table panel, Compare uses the copybook field names to populate the panel. If no copybook member is specified, then the Define Legacy Table panel is blank, allowing you to manually create the Legacy Table.

```
--------------------- Define Legacy Table: FOPLEG.ITEMS -----------------------
Command ===>                                                 Scroll ===> PAGE

Associated IMS DBD Name ===>            Segment ===>
                                                                 Row 1 of 7
Cmd  Level/Field Name              Type Len  Occur  Column Name
--- ------------------------------ --- ----- ----- ------------------
************************************ TOP ************************************
___ 1  ITEMS                            102         ITEMS
___  5 ITEM_ID                    CHR    5         ITEM_ID
___  5 ITEM_DESCRIPTION           CHR   72         ITEM_DESCRIPTION
___  5 CATEGORY                   CHR   14         CATEGORY
___  5 RATING                     CHR    4         RATING
___  5 UNIT_PRICE                 DEC    3         UNIT_PRICE
___  5 ON_HAND_INVENTORY          BIN    4         ON_HAND_INVENTORY
********************************** BOTTOM **********************************
```

The Define Legacy Table panel includes:

Associated IMS DBD Name
Specify the name of the DBD with which the Legacy Table is associated. The DBD must be in a DBD Library referenced in the Environment Definition. (The name of the Environment Definition determines the association. An associated Environment Definition has the same name as the Creator ID for the Legacy Table.) Specify an asterisk to generate a selection list that includes DBDs in the referenced Environment Definition.

This prompt is displayed only if the Legacy Table type is IMS.

Segment
Specify the name of the segment in the associated DBD. Specify an asterisk to generate a selection list of segments in the specified DBD.

This prompt is displayed only if the Legacy Table type is IMS.

Cmd Line command area. Valid line commands are:

C Copy the selected line to create a new one.

D Delete a line.

I Insert a line.

M Move a line.

R Repeat a line.

Z Zoom the field.

For Copy and Move, use A or B to indicate the destination. You can use the block form (for example, CC) or a repetition factor (for example, C4) for these commands. All line commands, other than the Z line command, are enabled only in full edit mode.

Level / Field Name
Level and name of the field. The level number indicates the hierarchical order of the fields described in the Legacy Table. For example, 01 indicates the highest level for a field in a Legacy Table. Entries are indented one space for each logical level.

Type The data type for the field. Possible data types are:

- BIN - Binary
- CHR - Character
- DEC - Decimal
- FLT - Floating
- NUM - Other numeric
- OTH - Other PL/I field types (Bit, Pointer, etc.)
- GRP - Group fields

Len Length of the field. If a field is of varying size (contains Occurs Depending On), length is listed as VAR.

The length of a 01-level field or a group field is the sum of all subordinate field lengths. If you change subordinate field lengths, use the VERIFY command to update the length of the higher-level field.

Occur Number of times a field occurs. Fields that occur a variable number of times are indicated by a D after the numeric value.

Column Name
Name by which the field and its data are referenced in Compare processes. Compare automatically generates a column name derived from the field name. The column name must be compatible with DB2 requirements in order to be used in Compare processes.

Compare column names must be unique and can be no more than 18 characters, while COBOL allows a maximum of 30 characters for field names, and PL/I allows 31. Thus, Compare may truncate COBOL and PL/I field names to use them as column names. Truncated names may require editing to be useful. In addition, the Compare-generated names do not include dashes; any dashes in a field name are translated to underscore characters. You can delete and type new column names or overtype the generated names.

Specify Data Source Information

An I/V/S table must be associated with a specific data source. To specify default data source information associated with the table, you can use the DATASOURCE primary command.

- For an IMS table, the Provide Retrieval Definition for DBD panel is displayed, using the Creator ID and the Associated IMS DBD Name as the two-part name of the Retrieval Definition. The Provide Retrieval Definition for DBD panel allows you to create or, if it exists, edit the Retrieval Definition. For more information about the Retrieval Definition see "IMS Retrieval Definition" on page 36.
- For a VSAM or sequential file table, the following pop-up is displayed.

```
+---------------------Specify Data Source Information--------------------+
|                                                                         |
|   Please specify the default Data Source information for this object.   |
|   Enter END when complete.                                              |
|                                                                         |
|     Data Source Type     : Sequential or VSAM Dataset                   |
|     Data Source DSN  ===> 'FOPRT.LEGACY.*'                              |
|                                                                         |
+-------------------------------------------------------------------------+
```

Enter the data set name explicitly, by enclosing it within single quotes, or enter part of a name with wildcards to obtain a selection list. On the Select Data Set panel, select the desired Data Set Name and press ENTER to insert the name in the Specify Data Source Information panel. Press END to continue.

Note: If a default data source has not been associated with the Legacy Table, you must define it when the Legacy Table is referenced in an Access Definition or Table Map.

Editing Commands

CRITERIA Command

Use criteria to match the Legacy Table to specific records in the source file. For example, a source file might contain three types of records of the same length, but different format. Each record type is distinguished by a code in the first field.

Criteria are created using an input format similar to that of an SQL WHERE clause. Use the CRITERIA primary command to specify criteria for a Legacy Table. Here is an example of a completed criteria statement.

```
------------- Modify Criteria For Legacy Table: FOPLEG.ITEMS ------------------
Command ===>                                                 Scroll ===> PAGE

                                                                   Row 1 of 3
Cmd Selection Criteria
--- ---------------------------------------------------------------------------
********************************** TOP ***********************************
___ RATING <>'X' AND CATEGORY = 'DRAMA' AND UNIT_PRICE < '12.00' AND
___ ITEM_ID > 'DR001' AND
___ ITEM_ID < 'DR060'
___ -- Sample comment text.
******************************** BOTTOM **********************************
Line Commands: (I)nsert (D)elete (R)epeat (M)ove (C)opy
Use the LIST COLUMNS command to add column names, if needed
```

A maximum of 200 lines is available for specifying criteria. Standard DB2 conventions apply to comments; each line must begin with two hyphens (--).

The criteria area is scrollable. A counter indicates the total number of lines and the relative position of the first displayed line.

To obtain a selection list of columns, enter LIST COLUMNS at the command line to display a column list, as shown in this example:

```
------------- Modify Criteria For Legacy Table: FOPLEG.ITEMS ------------------
Command ===>                                                  Scroll ===> PAGE

                                                                    Row 1 of 1
Cmd Selection Criteria
--- ------------------------------------------------------------------------
*** ********************************* TOP ************************************
...
*** ******* +---------------------Select One Column---------------------+ *****
Line Comman | Cmd           Column Name        Data Type         1 OF 7 |
Use the LIS | --- ------------------------------------------------      |
            | ********************** TOP ***********************        |
            | ___ ITEMS                      CHAR(102)                   |
            | ___ ITEM_ID                    CHAR(5)                     |
            | ___ ITEM_DESCRIPTION           CHAR(72)                    |
            | ___ CATEGORY                   CHAR(14)                    |
            | ___ RATING                     CHAR(4)                     |
            | ___ UNIT_PRICE                 DECIMAL(52)                 |
            | ___ ON_HAND_INVENTORY          INTEGER                     |
            | ********************** BOTTOM **********************        |
           +-------------------------------------------------------------+
```

Type S next to the desired Column Name and press ENTER to insert the column name in the panel. Format the criteria as you would an SQL WHERE clause. You may specify a combination of OR and AND logical operators as needed.

Several functions can be performed using line commands. The functions and the line commands are:

- Copy C*n*, CC
- Delete D*n*, DD
- Insert I*n*
- Move M*n*, MM
- Repeat R*n*, RR

For Copy and Move, use A or B to indicate the destination.

All leading and trailing spaces on each line are maintained; only the trailing spaces at the end of the last line are deleted.

Use END when finished. The criteria are saved with the Legacy Table.

EDIT Command

Use the EDIT command to toggle the Legacy Table editor between limited editing and full editing modes.

- Limited edit mode allows you to edit only the Column Name field of the Legacy Table editor.
- Full editing mode allows you to edit all fields in the Legacy Table editor. In addition, all line commands (that is, Copy, Delete, Insert, Move, and Repeat) are enabled.

Note: If no copybooks are specified on the Specify Copybooks for Legacy Table panel, then the Legacy Table editor is automatically in full editing mode when the Modify Legacy Table panel is displayed.

FIND Command

Use the FIND command to locate names that include a specific character string. The syntax is:

FIND *string*

where “string” is the search value. Do not delimit the string. Compare locates the first occurrence of the string beginning with the first displayed line of data. Use RFIND to continue the search from the current cursor location. If the string is not found, FIND goes to the beginning of the list and continues the search. If no match is found after the entire list is searched, the search terminates.

ONLY Command

The ONLY command is useful when you are working with a large Legacy Table and need to focus on a subset of the listed fields. ONLY displays the names of fields that satisfy command parameters. The syntax is:

ONLY *string*

where *string* is the exclusive value. Do not delimit the string. For example, to locate and display every field name that contains SHIPPING_ , specify:**ONLY SHIPPING_**

SHOW Command

Use the SHOW command to redisplay all field names after using ONLY.

VERIFY Command

Use the VERIFY command to check for errors without terminating the session or saving the data.

If a field length is blank when VERIFY is executed, a warning message is displayed. Group field lengths are automatically recalculated when you use VERIFY.

After making a correction, press ENTER to move to the next error.

Use END to verify the current specifications and terminate the edit session.

SAVE Command

Verify data in the current session and save it. The session is not terminated. If errors are detected, you are prompted to correct them as described for the VERIFY command.

Z (Zoom) Command

Use the Z (zoom) line command for a field listed on the Legacy Table editor to display field specifications on the Field Details panel.

```
------------------------------ Field Details ---------------------------------
Command ===>                                                 SCROLL ===> PAGE


Field Level   ===> 5
Field Name    ===> RATING
Column Name   ===> RATING
Type          ===> CHR       G-GRP, C-CHR, N-NUM, D-DEC, B-BIN, F-FLT, O-OTH
Length        ===> 4         Precision ===>      Scale ===>
Occurs From   ===>           Occurs To ===>
Depending On  ===>
Sync          ===>           Boundary ===>       Picture ===> X(4)
Is Redefined    :            Redefines ===>
Value         ===>
```

This panel includes the following items:

Field Level
Hierarchical level of the field within the Legacy Table.

Field Name
Name of the field.

Column Name
Name of the column for the field.

Type The data type for the field. Possible data types are:

- BIN - Binary
- CHR - Character
- DEC - Decimal
- FLT - Floating
- NUM - Other Numeric
- OTH - Other PL/I field types (Bit, Pointer, etc.)
- GRP - Group fields

Length
Length of the field. For a single field, the length is displayed. For a group, the combined length of subordinate fields is displayed.

Precision
Number of digits in a numeric field.

Scale Number of decimal positions in a numeric field.

Occurs From
Minimum number of times the field occurs.

Occurs To
Maximum number of times the field occurs.

Depending On
The number of times the field occurs (as in the COBOL parameter "Occurs Depending On"). If this value is inserted from working storage during program execution, you can use MAX_OCCURS to direct Compare to calculate the number of occurrences based on remaining space in the record. The DO field must be at the end of the record.

Sync Indicate how the field is aligned. Specify:

Y Explicitly aligned.

N Not explicitly aligned.

blank Default alignment. (Alignment inherited from higher level field or, if unspecified for PL/I, default determined by data type.)

Boundary
Alignment if field type is Other (applies to fields unique to PL/I). Specified as 1, 2, 4, or 8.

Picture
Picture of the field as defined in the source. For information only.

Is Redefined
Indicator for redefined field. This value is set on entry to the editor and with the VERIFY command.

Redefines
Name of the field that the current field redefines.

Value Value indicating the definition to use for redefined fields.

Use END or CANCEL to return to the Modify Legacy Table panel.

Use END twice to return to the Choose a Definition Option menu to configure Primary Keys, Relationships, or Column Maps, if necessary.

Saving the Legacy Table

After a Legacy Table is defined, it can be saved in the Optim Directory. Once saved, the Legacy Table is available to other users.

Use the END command to exit the Legacy Table editor. Modifications to the Legacy Table are automatically saved. By saving the current Legacy Table definition under a new name, you can use an existing Legacy Table as a prototype and retain the original version.

When defining a Legacy Table as part of a process, you can save the Legacy Table by specifying a fully qualified name. If you do not explicitly save the Legacy Table, it is not stored in the Directory and is available only for the current process.

Comparing One Pair of Tables

At times you may need to compare the “before” and “after” versions of a single table or the contents of one pair of tables. For example, after a reference table is updated, you can confirm the revisions by comparing the original version with the updated version. Although important to the data in other tables, the reference table is frequently maintained independently.

Select Option 1 on the COMPARE Process menu when you want to compare one pair of tables or two versions of a single table. The Specify COMPARE Sources panel is displayed. If an existing Compare Definition is to be used, the panel is populated with information from that definition. If not, the profiled values are displayed. The first time this panel is displayed, there are no profiled values. Only **Modify Sel. Criteria** and **Status** contain the default values, as in the following example.

```
-- Specify COMPARE Sources: RDEFAL1.COMPARE.SINGLE ----------------------------
 Command ===>

 Table 1 (Use LIKE Syntax for Lists):
   Creator ID             ===> PSTPROD         >>
   Table Name             ===> CUSTOMERS                          >>
   If Data from Extract File:
     Extract File DSN     ===> 'PSTRT.REL610.PSTDEMO.EXTRACT'
   If Data from Table:
     Modify Sel. Criteria ===> N  (S-SEL, Q-SQL, N-NO, A-ALL) Status: NONE

 Table 2 (Use LIKE Syntax for Lists):
   Creator ID             ===> BYXBEES         >>
   Table Name             ===> CUSTOMERS                          >>
   If Data from Extract File:
     Extract File DSN     ===>
   If Data from Table:
     Modify Sel. Criteria ===> N  (S-SEL, Q-SQL, N-NO, A-ALL)   Status: NONE

 Column Map (Optional) - Use MAP command to edit existing Column Map:
   MapID.MapName ===>                       (Use LOCAL for Internal/Unnamed Map)
 Define New Match Key              ===> N  (Y-Yes, N-No)
 Define Compare Match Options      ===> N  (Y-Yes, N-No)
```

The prompts on the panel for Table 1 and Table 2 are the same, as follows:

Creator ID
An up to 128-character creator ID. Use DB2 LIKE syntax to display a selection list of tables.

Table Name
Name of the table containing the data to be compared. Use DB2 LIKE syntax to display a selection list of tables.

Extract File DSN
If the data has been previously extracted and stored in an Extract File, specify the data set name of the Extract File containing the data for the named table. If an Extract File data set name is not

specified, Compare assumes the data is to be extracted from the database as part of the current Compare Process. Use an * as the last character in the data set name to display a selection list of Extract Files.

If Data from Table Modify Sel. Criteria
If the data is to be extracted, specify whether you want to define or modify selection criteria for the data. Specify:

S SEL. Display the Specify Selection Criteria panel.

Q SQL. Display the SQL WHERE Clause panel.

N NO. The criteria is not to be modified.

A ALL. Delete any selection criteria and extract all rows.

Selection criteria cannot be specified when the source is an Extract File.

Status Indicates whether selection criteria has been specified when the data is to be extracted for the Compare Process. Status cannot be modified. Valid values are:

NONE
Criteria has not been specified.

SEL Selection criteria has been specified.

SQL An SQL WHERE clause has been specified.

SEL/SQL
Both selection criteria and an SQL WHERE clause have been specified and will be logically ANDed.

MapID.MapName
Specify the name of a Column Map to map Table 1 columns with Table 2. You can also exclude columns from the comparison. Specify:

Name Name of the Column Map to use. Use DB2 LIKE syntax or the LIST MAPS command to display a selection list of Column Maps. Use the MAP command to modify the specified Column Map.

LOCAL
Parameter identifying the Column Map as specific to the current Compare Process and not available for any other process.

To create a new Column Map, specify the name of a Column Map to be saved, or specify LOCAL to define a Column Map for the current Compare Process only. If you specify the name of a Column Map that does not exist, a confirmation prompt is displayed to confirm that a new map is to be created. Note that if the names of the tables in the existing Column Map conflict with Table 1 and Table 2, a confirmation prompt is displayed to allow you to cancel editing the Column Map or to change the table names in the Column Map. Column Maps are discussed later in this section.

Define New Match Key
Indicate if a user-defined match key will be used for the Compare Process. Specify:

Y A user-defined match key will be used. After completing the panel, you will be prompted to define a match key.

The steps used to define a match key are the same as those used to define a primary key. See “Primary Keys and Match Keys” on page 25 for details about creating a primary key.

N Do not use a user-defined match key.

Define Compare Match Options
Indicate whether or not you want to modify the compare match options. Specify:

Y Change the current options. The current options are displayed on the panel. Any changes you make will be in effect for this comparison and will be saved with the Compare Definition.

N Use the current options.

Match Options Panel

The Compare Match Options panel allows you to specify match options that control the method used to compare rows, the accuracy with which non-uniquely keyed tables are matched, and limits on the processing to be performed.

If you specify Y for Define Compare Match Options on the Specify Compare Sources panel, the Compare Match Options panel is displayed.

```
--Compare Match Options-------------------------------------------------------
Command ===>

Commands: MKEY, CANCEL    END when Done

Source 1: RDLC.FILE10B
Source 2: RDLC.FILE10B

Match Key Property   ===> Non-Unique  (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method         ===> ANY

Soft Match Parameters (if not unique)
  Enabled            ===> Y      (Y-Yes, N-No)
  Percent Required   ===> 50     Percentage of columns that must match
  Percent Acceptable ===> 100    Processing stops at this percentage match
  Maximum Attempts   ===> 0      Limits number of comparisons per row
```

The Compare Match Options panel includes the following fields:

Source 1
Source 2
These are the two sources that are to be compared. These values are not modifiable.

Match Key Property
Describes the match key used for the sources to be compared. The value in this field is automatically populated, based upon the known attributes of the data. If you change this value to Non-keyed any match key previously defined for these sources is deleted. If this value is Unique, the remaining fields on this panel are non-modifiable. Allowable values are:

Non-Keyed
Data has no match key defined. All rows will be matched using the specified **Match Method**.

Non-Unique
A non-unique match key has been defined between source 1 and source 2. Unique rows will be matched using only the match key. Non-unique rows will be matched using the specified **Match Method**.

Unique
A unique match key has been defined between source 1 and source 2. Rows will be matched using only the match key. If duplicate keys are found, they will be reported as duplicates and not compared.

Note: If you select Unique and the data is not uniquely keyed, the results of the comparison may not be as expected. That is, the comparison may produce a large number of unmatched rows.

Unknown
The match key property is not known and will be determined by the compare process. Unknown is the default value when either of the sources is not an extract file. If you specify Match Method and Soft Match parameters, they will be used only if it is determined that the match key property is not **Unique**.

Match Method
Controls the processing for non-uniquely keyed and non-keyed sources. This field is not modifiable and displays the value ANY. Data that is non-keyed or non-uniquely keyed is compared using this method: all of the rows in Source 1 with the same key are compared with the rows in Source 2 with the same key. Equal rows are matched without any consideration for their relative sequence. Unequal rows are either considered unmatched or are matched according to the **Soft Match Parameters**.

Soft Match Parameters (if not unique)

The following are the soft match parameters.

Enabled
Controls whether soft matching is performed. Specify:

Y Enables soft matching. Unequal rows are compared and the best matches are paired and considered changed.

N Disables soft matching. All unequal rows are considered changed. Unequal rows are considered to be source 1 or source 2 only. If you specify N, the other Soft Match parameters are ignored.

If the match key property is **Unique**, soft matching cannot be used.

Percent Required
Specifies the percentage of similarity required to consider two unequal rows matched. This value applies only to the columns that are not defined as part of the match key. Allowable values range from 1 through 100. The default is 50 percent.

For example, if there are 10 columns that are not part of the match key and you specify a value of 80 for **Percent Required**, 8 columns or more must be equal for the rows to be paired. For this reason, if there are columns that must match, ensure that you define them as part of the match key.

Percent Acceptable
This parameter limits processing by reducing the precision of the comparison. Normally, all of the rows in source 1 are compared to all of the rows in source 2 and the best matches that meet the **Percent Required** are selected. If you specify a value less than 100 for **Percent Acceptable**, match processing for a row will stop when a row comparison meets or exceeds this percentage. The default is 100 percent.

Maximum Attempts
Limits the number of compares performed for any set of unequal rows. The value 0 disables this limit. Specify a value in the range 1 to 999,999. The default is 0.

Display Current Key
This prompt is displayed if the Compare Definition has been saved and contains a match key. Indicate to display the current match key. Specify:

Y Display the match key.

N Do not display the match key.

Match Option Examples

This topic contains several examples of different matching options used for the segment SHIPI, which is non-uniquely keyed. The Match Key used for all examples includes FOP_SURKEY(1:15) and SHIPI_KEY.

Here is the Source 1 example.

```
FOP_SURKEY(1:15)          SHIPI_KEY CLASS TYPE    DATE_UPDATED OWNER
++++++++++++++++++++++++ +++++++++ ----- ------- ------------ -----
SALES1CUST10001SHIPI30000 SHIPI3    1st   LETTER  12.04.27     BOB
SALES1CUST10001SHIPI40000 SHIPI4    2nd   PACKAGE 12.03.02     JOHN
SALES1CUST10001SHIPI40001 SHIPI4    1st   PACKAGE 09.04.27     PHIL
SALES1CUST10001SHIPI40002 SHIPI4    1st   PACKAGE 08.04.27     MIKE
```

Here is the Source 2 example.

```
FOP_SURKEY(1:15)          SHIPI_KEY CLASS TYPE    DATE_UPDATED OWNER
++++++++++++++++++++++++ +++++++++ ----- ------- ------------ -----
SALES1CUST10001SHIPI30000 SHIPI3    1st   PACKAGE 12.04.27     PETER
SALES1CUST10001SHIPI40000 SHIPI4    1st   PACKAGE 08.04.27     FRANK
SALES1CUST10001SHIPI40001 SHIPI4    2nd   PACKAGE 12.03.02     JOHN
SALES1CUST10001SHIPI40002 SHIPI4    1st   LETTER  11.04.29     BOB
```

Example 1

In this example, the Match Options are specified as shown in this example:

```
Match Key Property    ===> Non-Unique  (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method          ===> ANY

Soft Match Parameters (if Non-Keyed or Non-Unique)
  Enabled             ===> N       (Y-Yes, N-No)
  Percent Required    ===>         Percentage of columns that must match
  Percent Acceptable  ===>         Processing stops at this percentage match
  Maximum Attempts    ===>         Limits number of comparisons
```

The following example shows the comparison results:

```
Cmd Chg Src == Table: SALEHDAM SHIPI (T1) ==================================

                FOP_SURKEY       SHIPI_KEY   CLASS     TYPE     DATE_UPDATED   OWNER
*** ***************************************************************************
___  D   1  SALES1CUST100010000  SHIPI3      1st     LETTER     12.04.27     BOB
___  D    2 SALES1CUST100010000  SHIPI3      1st     PACKAGE    12.04.27     PETER
___       1  SALES1CUST100010001  SHIPI4      1st     PACKAGE    09.04.27     PHIL
          1  SALES1CUST100010002  SHIPI4      1st     PACKAGE    08.04.27     MIKE
          12 SALES1CUST100010000  SHIPI4      2nd     PACKAGE    12.03.02     JOHN
___        2 SALES1CUST100010000  SHIPI4      1st     PACKAGE    08.04.27     FRANK
___        2 SALES1CUST100010002  SHIPI4      1st     LETTER     11.04.29     BOB
*** ***************************************************************************
```

In this example, Soft Match is disabled and Match Key Property is Non-Unique. The Optim solution attempts to find and match any identical unique keys found in source 1 and source 2. Rows with identical, unique keys **SALES1CUST10001** and **SHIPI3** are reported as matched, but changed.

Example 2

In this example, the Optim solution attempts to match all rows with non-unique keys. Only non-unique rows with identical contents will be reported as matched (Src 12). All other rows will be reported as existing only in source 1 or source 2.

If the Match Options are specified as shown in the following example:

```
Match Key Property    ===> Non-Unique  (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method          ===> ANY

Soft Match Parameters (if Non-Keyed or Non-Unique)
  Enabled             ===> Y       (Y-Yes, N-No)
  Percent Required    ===> 75      Percentage of columns that must match
  Percent Acceptable  ===> 100     Processing stops at this percentage match
  Maximum Attempts    ===> 0       Limits number of comparisons
```

The comparison results would be as shown in this example:

```
Cmd Chg Src == Table: SALEHDAM SHIPI (T1) ==================================

              FOP SURKEY      SHIPI KEY   CLASS     TYPE     DATE UPDATED   OWNER
*** ************************************************************************
___       1  SALES1CUST100010000  SHIPI3      1st     LETTER     12.04.27     BOB
___       2  SALES1CUST100010000  SHIPI3      1st     PACKAGE    12.04.27     PETER
___      12  SALES1CUST100010000  SHIPI4      2nd     PACKAGE    12.03.02     JOHN
    D     1  SALES1CUST100010002  SHIPI4      1st     PACKAGE    08.04.27     MIKE
___ D     2  SALES1CUST100010000  SHIPI4      1st     PACKAGE    08.04.27     FRANK
___       1  SALES1CUST100010001  SHIPI4      1st     PACKAGE    09.04.27     PHIL
___       2  SALES1CUST100010002  SHIPI4      1st     LETTER     11.04.29     BOB
*** ************************************************************************
```

In this example, Soft Match is enabled and Percent Required is 75. The Optim solution attempts to match all rows with identical keys, unique or non-unique, that have at least 75% of identical column data. Two matches were found: one has identical contents, and one has different OWNER contents. All other rows are reported as existing only in source 1 or source 2.

Column Maps

A Column Map provides specifications needed to match pairs of columns from separate tables for a Compare Process, or exclude columns from processing. You can use a Column Map for a pair of tables regardless of Creator ID or table name. For example, a Column Map defined to map columns in a table named FOPDEMO.CUSTOMERS to those in a table named TEST.CUSTOMERS can also be used to map the identical tables PROD.CUSTOMER1 and QA.CUSTOMERS.

A multi-table Compare Process requires a Table Map, which may reference one or more Column Maps. A Compare Process for a single table comparison may reference a Column Map directly. Column Maps stored in the Optim Directory are available for reuse or sharing with other users. A local Column Map is stored as part of a Table Map or Compare Request and is otherwise not available.

A Column Map must be used when column names or attributes do not match. Also, you can match dissimilarly named columns from two source tables for a Match Key definition.

Column Maps provide great control and flexibility for Compare Processes. However, Column Maps that modify data cannot be used with Compare.

Select a Column Map

You can enter the name of a new Column Map on a Table Map panel or the Specify COMPARE Sources panel to display the Column Map editor or select **Option 6.3 Column Maps** from the Main Menu.

You can also select **Option 6 DEFINITIONS** to display the Choose a Definition Option menu and select **Option 3 COLUMN MAPS**. If using **Option 6**, the Choose a Column Map panel is displayed. Use this panel to name a new Column Map or to select an existing Column Map to modify or delete.

```
------------------------------ Choose a Column Map ----------------------------
Command ===>

Column Map:
  Map ID ===>
  Map Name ===>

Use '_' for DB2 LIKE character    ===> YES    (Y-Yes, N-No)
Rules used to validate Column Map ===> C      (M-Move/Archive, C-Compare)
```

The Choose a Column Map panel includes the following items:

Column Map:
The Map ID and Map Name for the Column Map. You can enter an explicit value, DB2 LIKE syntax, or blanks for these prompts in any combination.

Map ID
The Map ID for the Column Map that is being created or modified. The default is the previously entered value. Specify 1 to 8 characters.

Map Name
The name of the map that is being defined or modified. Specify 1 to 12 characters.

Note: You can enter the name of an existing Column Map to use as a model for the Column Map you wish to create. After editing the Column Map, save it under a different name.

Use '_' for DB2 LIKE character
Use of the underscore (_) character. Specify Y if the underscore is used as a DB2 LIKE character or N if it is used literally as part of the name.

For example, depending upon the use of the underscore character, A_B is a three-character name containing the characters A_B, as entered, or a three-character name that begins with A and ends with B with any valid character in the middle. The default is N, which means that underscore (_) is not recognized as a DB2 LIKE character.

Rules used to validate Column Map
Setting to apply Archive or Move validation rules or Compare validation rules. Specify:

M Comply with rules for Archive or Move.

C Comply with rules for Compare.

Archive and Move support mapping columns with compatible attributes, eliminating columns from processing, and modifying data. Compare only allows mapping of columns with compatible attributes and elimination of columns from processing.

Specifying Explicit Names

When you supply an explicit **Map ID** and **Map Name** and the Column Map does not exist, the Specify Column Map Tables panel is displayed, allowing you to specify the names of tables with columns that will be mapped. If the Column Map exists, the Modify Column Map panel is displayed.

Requesting a Selection List

A selection list is requested by using DB2 LIKE syntax or leaving one or both prompts blank as criteria. Column Maps that match the criteria are listed on the Select Column Maps panel. If no Column Maps satisfy the criteria, a message is displayed.

Validation Rules

When creating a Column Map, you must consider the purpose of the Column Map and the Optim component with which it is used. A Column Map defined for a Compare Process can be used in Archive or Move processing; however, the Archive and Move functions are not available to Compare, and a Column Map defined for Compare processing may not provide functions needed for Archive and Move. Therefore, you may want to use naming conventions or other means to segregate Column Maps defined for Compare processing.

Column Map Selection List

The setting for **Rules used to validate Column Map** on the Choose a Column Map panel determines the format of the selection list and the rules applied to a selected Column Map. If you select Compare rules, the selection list identifies the Source 1 and Source 2 tables and Compare rules apply to the selected Column Map, as shown in the following figure.

```
--------------------------- Select Column Maps --------------------------------
Command ===>                                                 Scroll ===> PAGE

    Line Cmds: S-Select, D-Delete, C-Copy, R-Rename, AR-Attr, I-Info 1 of 2

-------- Map -------- -------- Source 1--------- ------ Source 2 ------
Cmd    ID      Name       Creator    Table         Creator     Table
--- -------- ----------—— ------>> -------------->> ------>> -------------->>
****************************** TOP ***********************************
___ CMP1997   MAIN25      FOPDEMO    ORDERS        FOPDEMO2      ORDERS
___ CMP1998   MARCH       FOPDEMO    CUSTOMERS     FOPDEMO2      CUSTOMERS
```

Use the **S** line command or the **SELECT** primary command to select a Column Map to display in the Modify Column Map panel. If the source table or file for the Column Map does not exist, the Specify Column Map Tables panel is displayed, prompting for source table names. After you provide the table names, the Define Column Map panel is displayed, allowing you to define a Column Map. (The Modify Column Map panel is the same as the Define Column Map panel.)

You can also use the **SELECT** primary command to display a Column Map that is not included on the selection list. For example, to select the Column Map named FOPDEMO.TESTMAP, enter **SELECT FOPDEMO.TESTMAP**

If the Column Map named on the **SELECT** primary command does not exist, the Specify Column Map Tables panel is displayed to allow you to create a new Column Map.

Tables for a Column Map

Creating a Column Map requires information about the tables that are mapped. The Specify Column Map Tables panel prompts for the names of Source 1 and Source 2 tables. These tables provide information (column names and data types) used to populate the Column Map editor.

You must provide information about the destination, but can create a Column Map without identifying a Source 1 table, manually entering the source information instead. The Specify Column Map Tables panel is displayed when you provide the name of a new Column Map or specify LOCAL while editing a Table Map or if you provide the name of a new Column Map on the Choose a Column Map panel. This panel is also displayed if, from the Select Column Maps panel, you select a Column Map for which a source file or table does not exist.

```
---------------------------- Specify Column Map Tables --------------------------
Command ===>                                                SCROLL ===> PAGE

Specify Source 2 Table (Required) and Source 1 Table (Optional)
for Column Map PSTDEMO.ORDMAP

Use Source 2 Columns From a Table:
  Creator ID ===>                                           >>
  Table Name ===>                                           >>

Use Source 1 Columns From:
  An Extract File:
    DSN of Extract File ===>
    Extract Table       ===>                                        >>

Or, an Existing Table (Use LIKE Syntax for List):
  Source 1 Creator ID ===>                                          >>
  Source 2 Table Name ===>                                          >>
```

The prompts for Source 2 table include the following:

Use Source 2 Columns From a Table:
The Creator ID and Table Name for the Source 2 table.You can enter an explicit value, DB2 LIKE syntax, or blanks for these prompts in any combination.

Creator ID
The Creator ID for the Source 2 table. The default is the SQL ID for the current user. Specify 1 to 128 characters.

Table Name
The name of the Source 2 table. Specify 1 to 128 characters.

The prompts for Source 1 table include the following:

Use Source 1 Columns From:
The Source 1 table for Column Name and Data Type information used to populate the Column Map editor. Optionally, specify a table in an Extract File, a DB2 table, or I/V/S data.

Note: Only one type of source entry is allowed. You cannot specify both file and table entries.

An Extract File: DSN of Extract File
The name of an Extract File for the Source 1 table. Specify the name of a file (on disk, not tape) explicitly by enclosing it in quotes; otherwise the default Data Set Prefix, specified on the User Options panel, is prefixed to the name. You can obtain a selection list of data sets by using a wildcard character in the last position.

Extract Table
The fully qualified name of the Source 1 table in the file (for example, FOPDEMO.ORDERS). Use DB2 LIKE syntax or leave blank to select from a list of tables in the file.

Or, An Existing Table: Source Creator ID
The Creator ID for the Source 1 table. Specify 1 to 128 characters.

Source Table Name
The name of the Source or Source 1 table. Specify 1 to 128 characters.

You can enter an explicit value or DB2 LIKE syntax for both prompts in any combination, or for a single prompt with blanks for the other.

Source 1 Table

The Column Map must reside in the Optim Directory for the Source 2 subsystem. In most cases, you are connected to the Source 2 subsystem when defining a Column Map and would typically select a Source 1 table from an Archive or Extract File used in the process for which the Column

Map is defined. If a file is not available, you can select a database or Table Definition as the source for column names. If you are processing data across DB2 subsystems, however, only a source table in an Archive or Extract File is available.

No Source Table

In certain circumstances, you may not have access to either the source table or the Archive or Extract File. If so, you can create a Column Map without identifying a source table and manually enter the column names that map to the Source 2 columns. Since the Optim solution cannot reference definitions for the manually entered source columns, validation is performed during processing.

Leave the **Use Source Columns From:** prompts on the Specify Column Map Tables panel blank to indicate there is no source table. A prompt is displayed to confirm that a source table is not desired. Press Enter to proceed with the Column Map definition. Use END to return to the Specify Column Map Tables panel, where you can identify a source table.

Multiple Source Tables

When creating a Column Map manually, you can sometimes avoid errors and labor from typing source column names by using the **CHANGE TABLES** command to return to the Specify Column Map Tables panel. Enter the name of a table that has needed source column names and press Enter to return to the Column Map editor. Any previous Source Column entries are retained and you can use the **LIST UNUSED** or **LIST ALL** commands to select names of columns. You can use **CHANGE TABLE** and **LIST** several times to collect source column names from accessible tables to avoid typing each source column name.

For example, assume you do not have access to the source 1 table and the source 2 table contains 100 columns. One method of inserting the source column names in the Column Map is to type the 100 entries. However, if an accessible table contains columns with 50 of the needed names, you can use **CHANGE TABLE** and specify this table as the source for the Column Map. Then, you can use **LIST** to insert the 50 names rather than type them. Also, if a second table includes another 25 columns with needed names, you can reuse **CHANGE TABLE** and **LIST** to include those names. In fact, since the source 2 table may include many of the needed source column names, you can specify it as the source in order to list those names in the Column Map.

Using **CHANGE TABLE** is a convenient way to avoid typing in names of source columns that are present in other tables. Note that you can turn validation off in order to save a Column Map created from several sources.

Column Map Editor

After you identify the tables on the Specify Column Map Tables panel and press Enter, the Column Map editor is displayed, listing the Source 2 columns and data types with any matching source columns. The Column Map editor is also displayed from the Table Map editor or from the Specify COMPARE Sources panel, when you enter the name of a new Column Map or use the MAP command.

SAVE

The Column Map editor is labeled as the Define Column Map panel when you provide the name of a new Column Map, and the Modify Column Map panel when displayed from the Table Map editor or you enter the name of an existing Column Map. In the initial display for a new Column Map, source 2 columns are automatically paired with any source 1 columns with compatible data types that have the same name. Source 1 and Source 2 columns are paired automatically if both name and attributes (data type and precision) match exactly.

```
-- Define Column Map: PSTCBR.NEWCOMP ------------------------------------------
Command ===>                                                  Scroll ===> PAGE

Corresponding Columns MUST Have Compatible Data Types
Use LIST UNUSED Command for List of Unused Source Columns
Use LIST ALL Command for List of All Source Columns
                                                                       COMPARE
                                                                        1 OF 9
    --------PSTCBR.CUSTOMERS----------    -------------FOPDEMO.CUSTOMERS-----------
Cmd  Source 1 Column    Data Type Num         Source 2 Column   Data Type Status
--- ----------------->> ---------- ---     ------------------ ---------- -------
*** ********************************** TOP ***************************************
___ CUST_ID             CH(5)     1         CUST_ID            CH(5)     EQUAL
___ CUSTNAME            CH(20)    2         CUSTNAME           CH(2)     EQUAL
___ ADDRESS             VCH(50)   3         ADDRESS            VCH(50)   EQUAL
___ CITY                VCH(15)   4         CITY               VCH(15)   EQUAL
___ STATE               CH(2)     5         STATE              CH(2)     EQUAL
___ ZIP                 CH(5)     6         ZIP                CH(5)     EQUAL
___ YTD_SALES           DEC(7,2)  7         YTD_SALES          DEC(7,2)  EQUAL
___ SALESMAN_ID         CH(6)     8         SALESMAN_ID        CH(6)     EQUAL
___ PHONE_NUMBER        CH(10)    9         PHONE_NUMBER       CH(10)    EQUAL
*** ********************************* BOTTOM ***********************************
```

The Column Map editor includes the following items:

Cmd The line command area. Valid line commands on the Define/Modify Column Map panel are:

- **CLR** Clear Source 1 Column entry.
- **D** Delete entries with UNKNOWN status.
- **EXP** Display Source 1 Column entry in a 75-character pop-up window to specify a column value.
- **I** Display status information about a source and destination pairing. Useful for *ERROR* status.
- **SRC** Replace Source 1 Column entry with name of Source Column.

Source 1 Column

The fully qualified name of the Source 1 table is displayed on the line preceding this heading.

The value that is mapped to the Source 2 Column is listed on the same line. Specify a source column name or blanks.

Note: If you eliminate a foreign key column from a Compare Process by eliminating the Source 1 value, the relationship cannot be used to determine whether related changes have occurred. Also, to use a Match Key in a Compare Process, all Match Key columns must be processed. If you omit a primary key column from the Compare Process, you are prompted to specify a Match Key from columns in the Column Map.

Data Type

The data type for the source. Generally, **Data Type** is blank if the Source Column value is not the name of a column in the current source table.

The data types for source 1 and source 2 must be compatible. See "Column Map Compatibility," for more information.

The data type is automatically supplied and cannot be modified. Due to space limitations, the following data types are abbreviated. Data types are shown, followed by the abbreviations:

CHAR or CHAR FOR MIXED DATA – CH

VARCHAR – VCH

LONG VARCHAR – LVR

DECIMAL – DEC

FLOAT 4-byte length – SNGL FLOAT
FLOAT 8-byte length – DBL FLOAT
BLOB – BL
CLOB – CL
CLOB FOR MIXED DATA – CLOB
GRAPHIC – GR
VARGRAPHIC or LONG VARGRAPHIC – VGR
BINARY – BIN
VARBINARY – VARBIN

Source 2
The fully qualified name of the Source 2 table is displayed on the line preceding this heading.

Num A sequential number assigned, in ascending order, to each Source 2 column. Use these values with source column selection lists or the EXPAND command. Num is populated for columns in the current source 2 table only.

Column
The read-only names of the Source 2 columns listed in the order defined in the table.

Data Type
The data type for the Source 2 column. Data Type is blank if the column is not in the current source table. Due to space limitations, some data types are abbreviated. For more information, see the Source 1 Data Type description.

Status The status of each column pairing.

EQUAL
The columns have the same name, data type, and length.

ERROR
The specified mapping is invalid. This status is possible if the data type of a column in one of the sources has been changed since the Column Map was created or the column you selected does not match. A Column Map cannot be saved with columns in*ERROR* status. Use the I line command to determine the reason for the error.

MAPPED
The columns do not have identical data types or length but are compatible. This status also applies to mapped columns that are not validated.

NOTUSED
The column does not participate in a Compare Process.

UNKNOWN
The named source 2 column does not exist. This occurs if the column has been dropped from the Source 2 table since the map was created or if a different Source 2 table was specified. You can delete these unknown columns using the Delete line command, D, or the DELete UNKnown primary command. The line is deleted and the Source 2 column count is adjusted.

When you have completed editing the Column Map, regardless of whether it is LOCAL, existing or new, you can save the map under a new name using the **SAVE** command. Use the keyword LOCAL as the name to designate that the Column Map is not stored in the Directory but is to be used only with the current Compare Definition. When editing an existing map, the original version is unchanged if you explicitly save it under another name. (This is a handy way to model a new or a LOCAL Column Map after an existing Column Map.)

If you have edited the Column Map but do not use the **SAVE** command, the map is automatically saved when you use **END**. Use **CANCEL** to abandon your editing.

Both **END** and **CANCEL** terminate Column Map editing and redisplay the Specify COMPARE Sources panel. When you return to the Specify COMPARE Sources panel, MapID.MapName is automatically updated as appropriate to display the name under which the map was saved.

Specifications Complete

When you press **ENTER**, Compare evaluates your entries.

If a primary key is not specified for either of the tables, you are prompted to create a match key to be used and stored in the Compare Definition. (For more information on match keys, see "Processing Flow." For information about creating a primary key, see the *Common Elements Manual*.)

The Specify COMPARE Parameters and Execute panel is displayed. Since this is the same panel that is displayed when **Option 3 PERFORM** is selected on the **COMPARE Process** menu, see "Perform Process" on for information about this panel.

Comparing Two Sets of Data

Frequently you need to compare the "before" and "after" versions of related data that resides in multiple tables. For example, after an application in development is executed, you can verify that it has processed correctly by comparing the "before" and "after" images of the data manipulated by the application.

Since most applications address a set of related data that spans multiple tables, you need to evaluate not only the data from each table independently, but also the impact of the changes with respect to the set of related data. This option allows you to compare two sets of related data and identifies the changes in the rows in each table and propagates the change status to parent tables.

Specify the Source Types

Select **Option 2** on the COMPARE Process menu when you want to compare one set of related data with another. You are prompted to specify the type of source for each set of data. The following panel is displayed:

```
-- Specify COMPARE Source Types: TEMPORARY CD ---------------------------------
Command ===>

COMPARE can process data saved in an Extract File or in Tables.
Specify source types as follows:

Source 1 ===> 1 1 - Extract File
                2 - Set of Data Defined by an Access Definition

Source 2 ===> 3 1 - Extract File
                2 - Set of Data Defined by an Access Definition
                3 - All Rows from Multiple Tables
```

The possible types of sources are:

1 - Extract File
The source data has been previously extracted and stored in an Extract File on disk. This provides a constant, unchanging "snap shot" of one version of the data.

Note: An extract file on tape cannot be used in a compare process.

2 - Set of Data Defined by an Access Definition
The source data, as defined by an Access Definition, is to be selected as part of the Compare Process. Each time the Compare Process is executed using this Compare Definition, the data is extracted and, therefore, reflects the current state of the database.

3 - All Rows from Multiple Tables
This option is only available for Source 2. The source data is to be extracted for the Compare Process and consists of all the data from all the tables specified for Source 1. As with Option 2, the data is extracted each time a Compare Process is performed and, therefore, reflects the current state of the database.

This option is useful when:

1. Source 1 uses a temporary Access Definition, an Extract File, or resides on a different subsystem. You do not need to respecify the set of data using this option. Typically, when using a test database you will want to include all rows.
2. When the Source 2 tables do not have DB2 relationships, orphan rows can exist. Using this option, all rows are extracted for Source 2; therefore, when compared to the relational extract provided by Source 1, these orphans are identified.

Your selections are profiled. These profiled values are displayed whenever you are specifying a temporary Compare Definition. When you use an existing permanent Compare Definition, the values contained in that definition are displayed.

Any combination is valid. The panel that is displayed to prompt for more information about each source is determined by your selections. The following describes the prompts for each option.

Extract File

If the source is an Extract File, the following prompt is displayed:

```
-- Specify COMPARE Source Types: TEMPORARY CD ---------------------------------
                                                                  Command ===>
Source n: Extract File
   DSN ===>
```

where n in this panel segment is either 1 or 2 indicating the source. Specify:

DSN The data set name of the Extract or Archive File. You can display a selection list by typing * as the last character in the name.

Access Definition

If an Access Definition is selected, the following prompts are displayed:

```
-- Specify COMPARE Sources: FOPDEMO.SAMPLE.AD ---------------------------------
Command ===>                                                 SCROLL ===> PAGE

Source n: Relational Extract from Tables in Access Definition

    Type of Access Definition to Use ===> P (P-PERM, L-LOCAL)

    If Permanent, Specify Access Definition Name:

      GROUP ===>
      USER ===>
      NAME ===>
Modify Access Definition ===> N (Y-YES, N-NO)
```

where *n* in this panel segment is either 1 or 2 indicating the source. Specify:

Type of Access Definition to Use
Whether the Access Definition is stored in the Directory or is to be defined only for the current Compare Definition. Specify:

P Access Definition is stored permanently.

L Access Definition is LOCAL. It is stored in and used by the current Compare Definition only.

If Permanent, Specify Access Definition Name
The name of the Access Definition to be used. The name is composed of three parts:

GROUP
The 1- to 8-character group ID. The default is the previously entered value.

USER The 1- to 8-character user name. The default is the previously entered value.

NAME
The 1- to 12-character name of the Access Definition. You can use DB2 LIKE syntax in any of the prompts or leave one or more prompts blank to obtain a selection list.

Modify Access Definition
Used to indicate whether you want to edit the existing Access Definition. Specify:

Y Display Select Tables/Views for AD panel to edit the selected Access Definition

N Access Definition is not to be edited.

Editing an Access Definition

When you specify that the Access Definition is LOCAL, you must define the Access Definition. Therefore, Compare automatically displays the Select Tables/Views for AD panel for you to specify the set of data. If the Access Definition is PERMANENT and you specify the name of a new Access Definition, Compare also displays the Select Tables/Views for AD panel. (See the *Common Elements Manual* for details about defining and editing an Access Definition.)

After the initial editing of a LOCAL or new permanent Access Definition, you must specify **Y** to **Modify Access Definition** on the Specify COMPARE Sources panel if you want to re-edit the definition. Also, you must specify **Y** to edit an existing Access Definition.

Compare analyzes the steps involved to extract the data specified by the named Access Definition. Any error or warning conditions that may be encountered are displayed at this time.

When these errors and warnings are displayed, you can use the **SHOW STEPS** command to display information about how the extract would proceed. Operands on the command enable you to specify for which source you want the information. For example, to display the SHOW STEPS information for Source 2 enter: **SHOW STEPS 2**.

You can return to the Specify COMPARE Sources panel and specify **Y** to **Modify Access Definition** to correct the errors, warnings, or both; or you can specify a different source.

All Rows

If All Rows is selected for Source 2 and Source 1 is an Extract or Archive File, the following is displayed:

```
Source 2: All Rows from Multiple Tables
COMPARE will initially use matching Table Names from the above
Extract File for the Table Map
```

If All Rows is selected for Source 2 and Source 1 is an Access Definition, the following is displayed:

```
Source 2: All Rows from Multiple Tables
COMPARE will initially use matching Table Names from the above
Access Definition for the Table Map
```

As the text for this option states, you can modify the names of the tables and add and delete tables for Source 2 on the COMPARE Process Table Map panel.

Specifications Complete

After you specify the sources for the comparison, use **ENTER** to proceed. The COMPARE Process Table Map panel is displayed. Use **CANCEL** to return to the Specify COMPARE Sources panel.

Compare Process Table Map

The COMPARE Process Table Map panel enables you to match unlike named tables and to eliminate one or more tables from the comparison.

The tables from each source are listed under the appropriate Source 1 or Source 2 heading. Each table in Source 1 is compared to the corresponding table in Source 2 displayed on the same line. You may also specify the names of any Column Maps that are to be used. (Column Maps enable you to map unlike named columns and eliminate individual columns from the Compare Process.)

When this panel is initially displayed:

- Source 1 contains the Creator ID and table names defined in the Extract File or Access Definition that is used. The order in which the tables are listed is the same as in the Extract File or Access Definition. These values are protected.
- Source 2 contains the Creator ID and matching table names defined in the Extract File or Access Definition, if one is used. However, if **All Rows** is specified, these values are populated with the information from Source 1.

When Source 2 is specified as an Extract File or an Access Definition, you can use only the tables included in the Extract File or Access Definition. However, if Source 2 is **All Rows**, you can use any database table, view, alias, or synonym.

Assuming an Extract File is named for Source 1 and All Rows is specified for Source 2, the following panel is displayed:

```
------------------------ COMPARE Process Table Map ----------------------------
Command ===>                                                  Scroll ===> CSR

Commands: APPLY,SAVE,LIST,MAP,POPULATE,ACM,CLEAR,MKEY,MOPT  END when Done
Source 2 Tables must Exist in Corresponding Extract File
 Src 1 CID: RDLC                                        Column
 Src 2 CID: RDLC                                     >>  Map ID ===>

Source 1 Table Name      Source 2 Table Name        Type    Column Map or LOCAL
-------------------- ------------------------->> ------- --------------------
********************************** TOP ************************************
FILE10A              FILE10A                    EXISTS
FILE10B              FILE10B                    EXISTS
FILE10C              FILE10C                    EXISTS
FILE10D              FILE10D                    EXISTS
******************************** BOTTOM ***********************************
```

This panel includes the following items:

Src 1 CID

The default Creator ID for the Source 1 tables, as defined in the Access Definition, Extract or Archive File being used.

Src 2 CID

The default Creator ID for the Source 2 tables. The initial display is based on the source type.

If Source 2 is an Extract File or an Access Definition, the default Creator ID is displayed and cannot be modified.

If Source 2 is **All Rows** (Option 3 on the Specify COMPARE Source Types panel), Src 2 CID is unprotected. If Source 1 is an Extract File, the default Creator ID from that file is displayed in Src 2 CID. If Source 1 is an Access Definition, Src 2 CID is blank.

Column Map ID
The default qualifier for Column Maps.

Source 1 Table Name
The list of tables to be compared from Source 1. The names of the tables are taken from the Access Definition or Extract or Archive File specified for the source.

Source 2 Table Name
The list of tables to be compared from Source 2. If an Access Definition, Extract or Archive File is specified as the source, Compare attempts to match same name tables. Non-matching names are not displayed. If **All Rows** is the source, the Source 1 table names are repeated. To edit the list:

- Replace a name by typing over it or selecting a new table from a list. You can request a selection list using the **LIST TABLES** primary command.
- Clear all names using the **CLEAR** primary command.
- Prefix all names with a string using the **PREFIX** primary command.
- Append a string to all names using the **SUFFIX** primary command.

Type The type of the object named in Source 2 Table Name. Compare supplies this value and it is not modifiable. Possible values include:

TABLE
Table

S-MQT
System-maintained Materialized Query Table

U-MQT
User-maintained Materialized Query Table

VIEW View

UNKNOWN
Non-existent table or no value in Dest CID

EXISTS
Exists

TEMPTBL
Temporary table

A-TABLE
Alias of a table

S-TABLE
Synonym of a table

A-VIEW
Alias of a view

S-VIEW
Synonym of a view

UNUSED
Unused

When a Source 2 table is not mapped to a Source 1 table, the type is UNUSED. When a Source 2 table does not exist, the type is UNKNOWN. (This must be resolved before saving or using the Compare Definition.)

When Source 2 is a table and the type value is EXISTS, one of the following values may also be displayed:

- An asterisk (*) indicates a user-defined match key will be applied to the tables.
- A dash (-) indicates user-specified match options will be used.
- A plus sign (+) indicates both user-defined match key and user-specified match options.

When Source 2 is an Extract File, the original source type may be unavailable; therefore, the type is EXISTS.

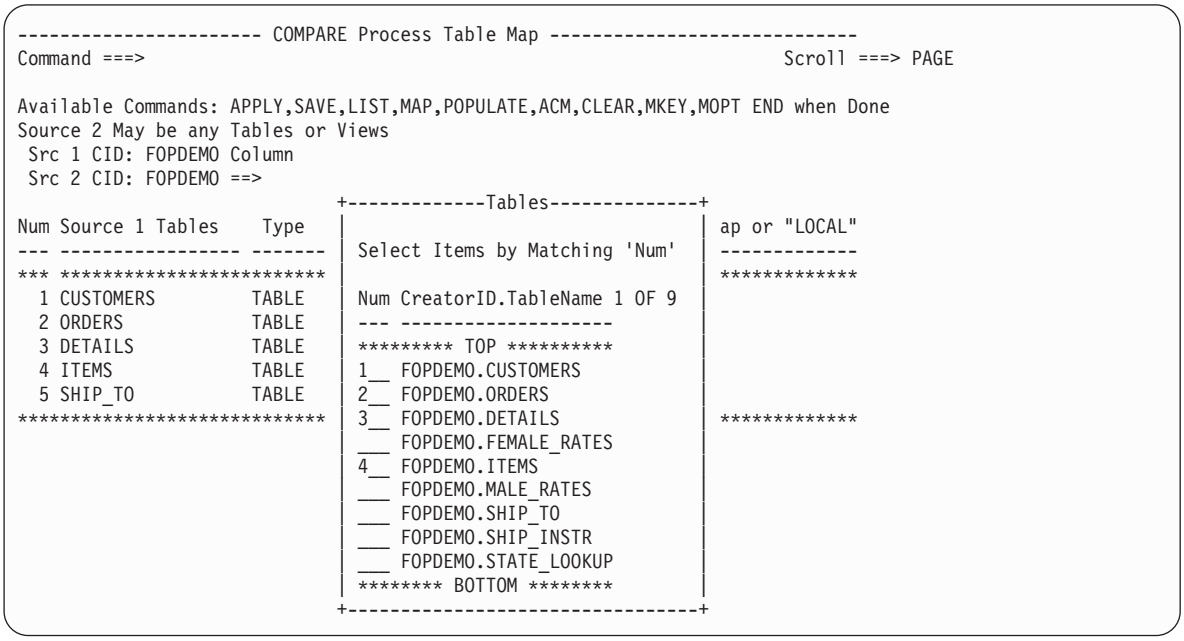
Column Map

The name of the Column Map used to compare the two tables.

Modify Source 2 Table Names

You can modify the Source 2 Creator ID when **All Rows** is the source. You can modify any of the table names specified for Source 2. However, if the source is specified as an **Access Definition, Extract or Archive File**, you can only specify table names included in the specified source. If the source is **All Rows**, you can specify any database table, view, alias, or synonym.

When Source 2 is **All Rows**, you can display a selection list of tables for Source 2, using the **LIST TABLES** command to replace or insert a table name. This command presents a list of the available tables and assigns a number to the tables listed for Source 1. This number is used to select a table from the list of Source 2 tables. The following is a sample of the LIST TABLES display. In the following figure, **LIST TABLES FOPDEMO.%** has been entered. Note there are a few database tables not used in the Extract File for Source 1, but included on the list for Source 2 tables because they are in the database.



Use **Num** on the Tables pop-up window to specify the number of the Source 1 table to match to the selected Source 2 table. In this example, five tables are mapped to the five Source 1 tables. The selected table names are automatically inserted under **Source 2 Table Name**. The **LIST** command can be used to display a selection list of tables, views, aliases, and synonyms.

When Source 2 is an Extract File, Archive File, or an Access Definition, the **LIST TABLES** command lists all of the objects currently not matched to Source 1 tables. Use **LIST ALL** to list all available objects in the Extract File, Archive File, or Access Definition. (See the *Common Elements Manual* for details about using **LIST** to map source tables.)

Use Existing Map

You can use the **APPLY** command to populate the Table Map with the specifications from a previously defined Table Map. If the Source 1 tables in the Compare Process match the Source 1 tables in the Table Map, Compare will populate the Source 2 tables from the existing Table Map, provided they exist.

User-Defined Match Key

You can use the **MKEY** command to create a match key. To specify the table pair to which you want to apply the match key, type one of the table names with the command (for example, **MKEY creatorid.tablename**), or type the command and place the cursor on the row with the tables. After pressing **ENTER**, an asterisk is displayed next to the Type for the tables. After completing the Table Map, you will be prompted to define a match key, which is saved with the Compare Definition. The steps used to define a match key are the same as those used to define a primary key.

Display Match Key

After a Compare Definition is saved, you can display a match key for a table pair using the **MKEY DISPLAY** command. To display a match key, type the command, place the cursor on the row with the table pair, and press **ENTER**.

MOPT Command (Match Options)

For a multiple compare, match options can be defined for each pair of tables by typing **MOPT** (Match Options) on the command line, placing the cursor on the source 2 table name, and then pressing **Enter**. The Optim solution displays the Compare Match Options panel, as described in"Match Options Panel" on page 53.

ACM Command

If necessary, you can use the **ACM** command to open the Choose Access Method pop-up dialog, allowing you to override the default method (scan or key lookup) for accessing the parent or child table for each relationship. A scan reads all rows in a table at one time; whereas a key lookup locates rows using a WHERE clause to search for primary or foreign key values.

Saving the Map

You can use the **SAVE** command to save this Table Map in the Optim Directory. Then, the Table Map is available to other users and can be used for other Compare Definitions when specified as the operand of the **APPLY** command.

If you do not explicitly save the Table Map, it is stored only in the current Compare Definition and is not available to other users or processes.

Column Maps

By default, columns whose names match and attributes are compatible are compared. When this default is inadequate, you can specify Column Maps for any pair of compared tables. These maps are used to match unlike named columns and eliminate columns from the comparison. However, the following restrictions apply:

- Mapped columns must be defined with compatible data types.
- You cannot eliminate a column that is part of the match key.
- If you eliminate a column that is part of a relationship, that relationship cannot be used to mark related changes or find orphans.

For details, see "Column Maps" on page 56.

Specify the Column Map Name

You can type the Column Map name in the provided area, request a selection list of Column Maps using the **LIST MAPS** command, or use the **POPULATE** command to automatically insert the Column Map names. You can create a new Column Map by specifying a new name. (Compare

will display a confirmation prompt and, if the name is correct, invoke the Column Map editor.) You can edit an existing Column Map using the **MAP** command.

You can define a Column Map to be used by the current Compare Definition by specifying the Column Map name as LOCAL. Compare assumes that you are going to create the Column Map; therefore, the Column Map editor panel is automatically displayed and populated with the names of the columns that match in name and have compatible data types from each source table. You can edit appropriately.

POPULATE command

You can use the **POPULATE** command to automatically insert the names of Column Maps stored in the Directory for each pair of tables to be compared. If the columns in the Column Map are not present in the paired tables, a diagnostic is displayed.

Specifications Complete

When you have completed specifying the Table Map information, use **END** to proceed. Compare analyzes the available relationships on the two sources. Typically all relationships can be used automatically and the Specify COMPARE Parameters and Execute panel is displayed. This is the same panel that is displayed when **Option 3 PERFORM** is selected on the COMPARE Process menu, see "Specify Compare Parameters and Execute Panel" on page 72 for information.

Select Relationships

After the Table Map is completed, Compare checks the relationships between the tables in each source. Regardless of how many relationships are traversed between tables to extract the data, only one relationship can be used for each pair of tables in each direction (parent to child and child to parent) to propagate related changes. Compare prompts when more than one relationship is available. That is, if the relationships from each source are not the same and/or more than one relationship exists in either or both sources, Compare displays a list of the available relationships from which you must select one for each pair of tables.

The list of relationships is obtained from the source: DB2 Catalog and Optim Directory if DB2 tables are the source, IMS database if IMS segments, the Optim Directory if VSAM or sequential files, and the Extract File when that is the source. The relationships from the Extract File are those that were available when the Extract File was created. (The list of relationships is displayed on the Specify Relationship Usage panel when defining the data to be extracted.)

The list of relationships on the Select Relationship to Use panel are grouped by parent and child table names.

```
------------------------- Select Relationship to Use -------------------------
Command ===>                                                  Scroll ===> PAGE

Select One Relationship from Each Set, Use END to Exit when Complete

Cmd Sel Src Parent Table               Child Table              Relation Type
--- --- --- -------------------------- ------------------------ -------- ----
*** ********************************* TOP ************************************
___ S     1 CUSTOMERS                  ORDERS                   RCO      OPT
___       2 CUSTOMERS                  ORDERS                   RCO      DB2
    ------------------------------------------------------------------------
___ S     1 ORDERS                     DETAILS                 *ROD      DB2
___       2 ORDERS                     DETAILS                  RODB     OPT
    ------------------------------------------------------------------------
___       1 DETAILS                    ITEMS                    RDI      DB2
___ S     2 DETAILS                    ITEMS                    DRI      DB2
___       2 DETAILS                    ITEMS                    RDI2     DB2
```

The Select Relationship to Use panel includes the following items:

Cmd The line command entry area. Specify:

S Select relationship.

I Display information about a relationship.

U Unselect relationship. Only one relationship can be selected for each pair of tables.

Sel Indicates whether the relationship has been selected. If selected, **Sel** contains an S, otherwise it is blank.

Src Indicates the source of the relationship as either 1 for Source 1 or 2 for Source 2.

Parent Table
Name of the parent table in the relationship.

Child Table
Name of the child table in the relationship.

Relation
Name of the relationship.

Type Indicates the type of the relationship as defined in the DB2 Catalog (DB2), the Optim Directory (OPT), or in the IMS database as a Physical Relationship (IMS) or Logical Relationship (IML).

Duplicate Relationships and Selecting Relationships

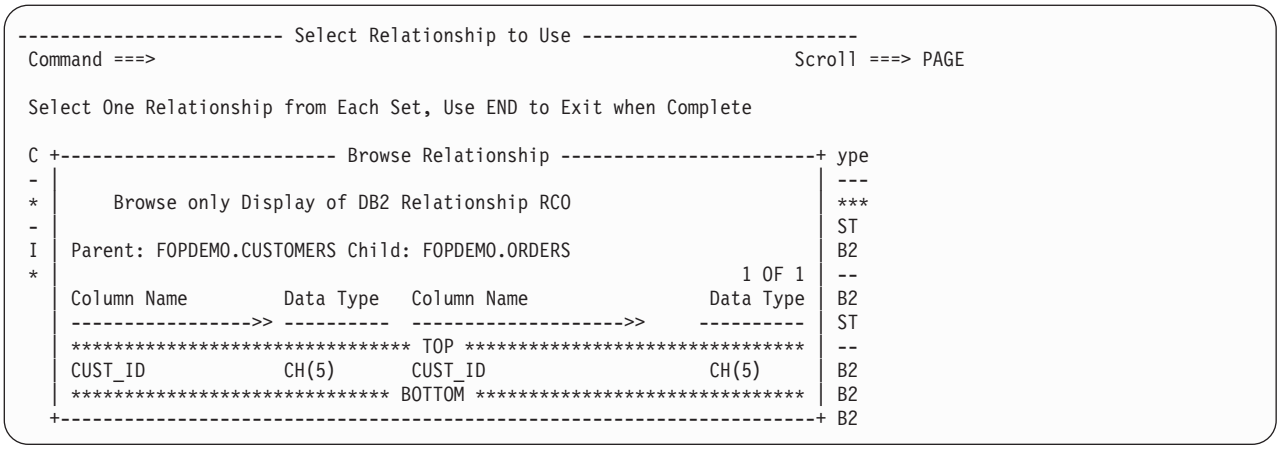
Sometimes relationships are encountered that, although named differently, are actually identical. That is, the names of the parent and child tables are the same and the columns are the same, but the relationship name is different. (A relationship is not identical if the assignment of parent/child table names is reversed such as the parent and child in one relationship are the child and parent in another.)

Rather than list these duplicate relationships individually, one of the relationships is listed preceded by an asterisk. The asterisk indicates that there are one or more other relationships identical to the listed relationship but these other relationships have different names. (On the panel in the figure, an asterisk precedes the relationship ROD for ORDERS and DETAILS indicating that at least one more identical relationship exists.)

Use the **Select** line command to select the relationship. In this example, one relationship is selected for each pair of related tables. Use the **Unselect** line command to remove a relationship from the select status. Also, if you select a relationship, any previously selected relationship for that pair of tables is automatically unselected because only one relationship can be selected for each pair of tables.

If you do not select a relationship for each pair of tables, the Compare Process does not identify the related changes between these tables. When you do not select a relationship for one or more pairs of tables, Compare displays a confirmation prompt when you attempt to leave the panel. You can proceed with the Compare Process by pressing **ENTER** or redisplay the Select Relationship to Use panel by using **END**.

The **Info** line command is available to aid in selecting a relationship. This line command displays the details of the relationship for the relationship selected by the command. Assume the **I** line command is entered for the relationship RCO between the CUSTOMERS and ORDERS tables. The following information is displayed:



The columns participating in the relationship are displayed along with their data type and length. If the list of columns does not fit in its entirety on the screen, you can scroll the display. When you have completed reviewing the relationship information, use **END** to return to the Select Relationship to Use panel.

The relationships included on this selection list are retrieved from the Extract File, DB2 Catalog, and Optim Directory. Only relationships where the table names and the column names in the relationship match or have been mapped are included.

If a relationship is encountered that does not meet these requirements, but a comparable relationship exists, the unsuitable relationship is not used. However, if a comparable relationship does not exist, a message is displayed. You can continue with the Compare Process; however, the related changes are not propagated for the tables that do not have a suitable relationship.

Perform the Process

After the sets of data to be compared are specified and the relationships to be used for the comparison are selected, the Specify COMPARE Parameters and Execute panel is displayed. (This panel is also displayed when **Option 3** is selected on the Compare Process menu to perform a compare using an existing Compare Definition.) The prompts are determined by the sources.

- When data is to be extracted from the database (Source 1 is an Access Definition and when Source 2 is an Access Definition or All Rows), prompts for Extract Options for the appropriate sources are displayed.
- When the data is from an Extract File, prompts for Extract Options for the appropriate sources are not displayed.
- The prompts for the Compare File DSN and Compare Options are always included.

Specify Compare Parameters and Execute Panel

After the sets of data to be compared are specified and the relationships to be used for the comparison are selected, the Specify COMPARE Parameters and Execute panel is displayed. The prompts are determined by the sources. In the following figure, both Source 1 and Source 2 request data to be extracted from the DB2 tables. Prompts for Extract Options are displayed for both sources.

```
------------------- Specify COMPARE Parameters and Execute --------------------
Command ===>                                                  SCROLL ===> PAGE

Compare File DSN ===> 'FOPDEMO.SAMPLE.COMPARE'

Source 1 Extract Options:
  Selection Criteria and Row List Specified for Start Table:
    Select Start Table Rows by    ===> R     (R-Row List Only, B-Both)
  Limit Number of Extract Rows    ===>       (1-4294967295, Blank/SL)
  Extract Data using              ===> D     (D-DB2, B-BMC Unload Program)

Source 2 Extract Options:
  Limit Number of Extract Rows    ===>       (1-4294967295, Blank/SL)
  Extract Data using              ===> D     (D-DB2, B-BMC Unload Program)

Compare Options:
  Generate Report After Process   ===> Y     (Y-Yes, N-No)
  Run Process in Batch or Online  ===> O     (B-Batch, O-Online)
    If Online, Invoke Browse      ===> Y     (Y-Yes, N-No)
    If Batch, Review or Save JCL  ===> R     (N-No, R,Review, S-Save)
```

The Specify COMPARE Parameters and Execute panel includes the following items:

Compare File DSN

Name of the Compare File that is to contain the results of the comparison. This file must be a sequential data set. The file name can be specified explicitly by enclosing it in quotes; otherwise, the default prefix as specified on the User Options panel is automatically prefixed to the name.

When the Compare Process begins, Compare searches for the named data set:

- If the data set exists, Compare checks to see if it is suitable for a Compare File. If it is, the current Compare Process overlays the data. If it is not, Compare prompts you to re-specify the name.
- If the data set does not exist, Compare prompts for the necessary information to allocate the file. See the *Common Elements Manual* for a description of the allocation prompts.

You can obtain a selection list of data sets using the wild card character, *, in the last position of the name. Use the **Select** line command, **S**, on the selection list to select the file name.

Extract Options

The following prompts for Extract Options are displayed for Source 1 or Source 2 or both only when the data for the source(s) must be extracted from the DB2 tables for the Compare Process.

Select Start Table Rows by

This prompt is only displayed when both a row list and selection criteria have been specified. Specify whether a row list or selection criteria is to be used for the Start Table when extracting the data. Specify:

R Only the row list values are used.

B Both the row list and selection criteria are used.

Limit Number of Extract Rows

Maximum number of rows of data that can be extracted for the source. The extract is terminated if the number of extracted rows exceeds this limit. (If the extract is terminated, the Compare Process is not performed.) Specify:

value 0-4294967295

blank The site-defined limit, which is set on the Site Options panel.

Extract Data using
Method for extracting the data. This prompt is displayed if a DB2 table is a source for the comparison. You will have two options: DB2 and a utility specified in the Site Options panel. Specify:

D DB2 is used.

B UNLOAD PLUS is used.

C Auto-Unload or Auto-Online Unload is used.

I High Performance Unload is used.

Compare Options

Generate Report After Process
Specify whether a report is to be generated after the Compare Process completes. The report can contain a statistical summary, details, or both. You can store the report in a file and print it. Specify:

Y Display the prompts for report options and generate a report.

N Do not generate a report. This is the default.

Whether or not a report is generated, you can browse the results of the Compare Process stored in the Compare File. Details about the prompts and the format of the report are discussed in "Generate Report" on page 84.

Run Process in Batch or Online
Specify whether execution of the Compare Process is batch or online. Specify:

B Batch

O Online

When using an unload utility, or if this request exceeds the site limit for the maximum number of rows for online processing, this option is set to Batch and cannot be modified.

If Online, Invoke Browse
Indicate whether the Compare Browse facility should be invoked automatically when the Compare Process completes. Specify:

Y Invoke Browse.

N Do not invoke Browse. The Compare Process Report is displayed.

This is specified for online execution only. For details about browsing the results, see "Browse a Compare File" on page 80.

If Batch, Review or Save JCL
For batch execution, indicate whether the JCL and control statements should be submitted, reviewed prior to job submission or saved for submission at a later time. Since the JCL and control statements are displayed in the ISPF editor, you can modify them for the current request and save them to submit later. Specify:

N Submit job, do not display or save the JCL and control statements.

R Display the JCL and control statements for review prior to job submission.

S Save the JCL and control statements. Prompts are provided for you to specify the name of a file in which to store the JCL and control statements.

Unload Programs

If an unload program is available and the response to **Extract Data using** is an unload program for Source 1, Source 2, or both, you are prompted to define unload program parameters for each source. In the next figure, a sample prompt is displayed for Source 1.

```
------------------- Specify COMPARE Parameters and Execute --------------------
 Command ===>

 +--------------Specify Unload Program Parameters for Source 1--------------+
 |                                                                           |
 | Source 1 File Type    ===> I (I-IMAGE COPY, D-DB FILES)                   |
 |                                                                           |
 | If using an Image Copy, specify which Image Copy datasets should be used |
 |   Image Copy Criteria ===> L (A-First On or After Date/Time,              |
 |                               B-First On or Before Date/Time,             |
 |                               L-Latest Image Copy,                        |
 |                               S-Specific Image Copy DSN)                  |
 |                                                                           |
 | If selecting an Image Copy by Date and Time:                              |
 |   Date (YYYY-MM-DD)   ===>                                                |
 |   Time (HH.MM.SS)     ===>                                                |
 |                                                                           |
 | If selecting an Image Copy by data set name:                              |
 | Image Copy DSN        ===>                                                |
 |                                                                           |
 | If Start Table is partitioned, you may use a subset of the partitions     |
 | Use Subset            ===> N (Y-Yes, N-No)                                |
 +---------------------------------------------------------------------------+
```

If you use an unload program for both sources, the prompts appear consecutively. You are prompted to specify the **Source File Type** as either the Image Copy or the database VSAM files.

If you specify an Image Copy, you can specify the date and time of the file.

Generate Report

> If you have specified Yes to **Generate Report**, a panel prompting for information to generate the report is displayed prior to performing the Compare Process. This is the same panel that is displayed when option **R** is selected from the COMPARE Process menu and is discussed in "Generate Report" on page 84.

Image Copy data sets on the same tape volume

To extract or compare data from DB2 image copy data sets in multiple partitions of the same tablespace stored on the same tape volume, you must manually edit the JCL to allocate the data sets. Multiple image copy data sets cataloged on the same tape volume can not be allocated using dynamic allocation. This is a z/OS limitation. If you attempt to use dynamic allocation, extract or compare process fails with a dynamic allocation error. Refer to the *Common Elements Manual*, section on Allocating External Files, for details.

Batch Execution

If you specify batch execution, Compare builds the necessary JCL and Batch Utility control statements. The JOB card information is taken from the JCL specified on the Job Card and Print Options panel. If you enter **YES** at the **Prompt for Changes Before Job Submission** prompt on the Job Card and Print Options panel, the default Job card, as indicated on that panel, is displayed prior to job submission. You can edit the Job card and print options and specify whether your changes apply to the current job only or are applied permanently.

The information on the Job Card and Print Options panel is used, together with the Compare Process parameters, to build the JCL and control statements required to perform the process. If you enter **Review**

to **If Batch, Review or Save JCL** on the Specify COMPARE Parameters and Execute panel, the entire JCL and control statements are displayed in the ISPF editor. The JCL and control statements can be edited and saved. (See the Batch Utility Compare statement parameters and values in Chapter 5, "Batch Utility," on page 95.

Use **END** to return from the ISPF editor to Compare. However, your response to the **Submit Jobs with END** prompt on the User Options panel determines whether the job is automatically submitted. If you enter **NO** at the prompt, you must explicitly submit the job from the ISPF editor, using the **SUBMIT** command.

If you enter **YES**, the job is automatically submitted. Use the **CANCEL** command to return to the Specify COMPARE Parameters and Execute panel without submitting the job. You can modify the specifications or cancel the request from this panel.

(See the *Common Elements Manual* for more information on establishing whether jobs are automatically submitted when **END** is used.)

If you submit the job and an error is encountered in the Job card, a message is displayed. You can review the Job card and correct the error or terminate the Compare Process.

Batch Overrides

If you save the generated batch job to a data set, you can submit the job directly from the ISPF editor rather than from within an online session. Submitting the job directly is especially convenient when you want to compare different Extract Files to a set of tables or to another file, using common Compare Process JCL. When you submit the batch job directly, you can specify overrides to the Source 1, Source 2, or both data set names. Also, if an unload program is used to extract data for Source 1, Source 2, or both, you can override the Image Copy data set name, date, and time parameters for each source.

Note: With Release 5.5, a generated batch job executes the Batch Utility to perform the specified function. The batch job includes a series of control statements defining the function to be performed. You can edit these control statements directly as an alternative to providing batch overrides. If batch overrides are not available, you must edit the control statements directly. The *Batch Utilities Guide* describes the Batch Utility control statements. (All batch overrides that were valid prior to Release 5.5 will continue to be valid.)

Use the PSDFOVRD DD statement in the JCL to provide the desired overrides.

COMPARE_DSN_SRC1

To specify a new data set name for the Source 1 file, specify:
COMPARE_DSN_SRC1 *data.set.name*

This override allows you to use one set of saved JCL to compare files regardless of the names.

COMPARE_DSN_SRC2

To specify a new data set name for the Source 2 file, specify:
COMPARE_DSN_SRC2 *data.set.name*

This override allows you to use one set of saved JCL to compare files regardless of the names.

SOURCE_CID_OVERRIDE

To override the default Creator ID for Source 1 or Source 2 of the Compare Definition, specify:
SOURCE_CID_OVERRIDE_ *cid*

This override allows you to use one set of saved JCL to compare files regardless of the default Creator ID. To override the Source 1 specification, use SOURCE_CID_OVERRIDE_1. Use SOURCE_CID_OVERRIDE_2 to override the Source 2 specification.

cid Creator ID to override the default Creator ID for the source.

If the source is an Extract File, you can use the wildcard (%) to use the default Creator ID defined in the Extract File.

This Creator ID applies to all tables that were defined with the initial default Creator ID in the Compare Definition. Tables in the Compare Definition not defined with the initial default Creator ID will not be altered by the specified Creator ID override. This override also changes the names of the tables in the relationships processed by Compare to use the override default Creator ID. If a relationship is not found for the updated table name, an error occurs when the compare is performed. If the source is an extract file, and a table initially defined in the Compare definition with default Creator ID is not in the extract file, the table will be skipped, a warning message will be produced, but the Compare job will continue.

UNL_IMAGECOPY _DSN

To override the Image Copy DSN parameter for an unload program used to extract data for Source 1, specify:
UNL_IMAGECOPY_DSN *image.file.dsn*

UNL_IMAGECOPY _DATE

To override the Image Copy Date parameter for an unload program used to extract data for Source 1, specify:
UNL_IMAGECOPY_DATE *yyyy-mm-dd*

UNL_IMAGECOPY _TIME

To override the Image Copy Time parameter for an unload program used to extract data for Source 1, specify:
UNL_IMAGECOPY_TIME *hh.mm.ss*

UNL_IMAGECOPY _SELECT

To override the Image Copy Criteria parameter for an unload program used to extract data for Source 1, specify:
UNL_IMAGECOPY_SELECT { A | B | L | S }

A First Image Copy file created on or after the specified Date and Time.

B First Image Copy File created on or before the specified Date and Time.

L Latest Image Copy file. Any Date and Time values are ignored.

S Image Copy file. The name is provided as the Image Copy DSN parameter.

UNL_IMAGECOPY _DSN2

To override the Image Copy DSN parameter for an unload program used to extract data for Source 2, specify:
UNL_IMAGECOPY_DSN2 *image.file.dsn*

UNL_IMAGECOPY _DATE2

To override the Image Copy Date parameter for an unload program used to extract data for Source 2, specify:
UNL_IMAGECOPY_DATE2 *yyyy-mm-dd*

UNL_IMAGECOPY _TIME2

To override the Image Copy Time parameter for an unload program used to extract data for Source 2, specify:
UNL_IMAGECOPY_TIME2 *hh.mm.ss*

UNL_IMAGECOPY _SELECT2

To override the Image Copy Criteria parameter for an unload program used to extract data for Source 2, specify:
UNL_IMAGECOPY_SELECT2 { A | B | L | S }

A First Image Copy file created on or after the specified Date and Time.

B First Image Copy File created on or before the specified Date and Time.

L Latest Image Copy file. Any Date and Time values are ignored.

S Image Copy file. The name is provided as the Image Copy DSN parameter.

Save JCL

You can save the JCL and Batch Utility control statements, modify them and execute the process without re-invoking Compare. Specify **S** to the **If Batch, Review or Save JCL** prompt. The following window will then prompt for the information to save the JCL and control statements.

```
.---------------------------- Save JCL Parameters --------------------------.
|                                                                            |
| DSN to Save JCL to     ===>                                                |
| Member (if PDS)        ===>                                                |
| Replace Existing Data  ===> Y-Yes, N-NO                                    |
|                                                                            |
| DSN to Hold SYSIN Data ===>                                                |
| Member (if PDS)        ===>                                                |
| Replace Existing Dat   ===> Y-Yes, N-NO                                    |
|                                                                            |
| Submit JCL, or Review? ===> (S-Submit, R-Review, N-Neither)                |
.----------------------------------------------------------------------------.
```

The Save JCL Parameters window includes the following prompts:

DSN to Save JCL to
Name of the sequential file or partitioned data set to receive the JCL and control statements. If you specify a partitioned data set, specify the member name at the **Member** prompt.

Member (if PDS)
Name of the member in the partitioned data set specified for the DSN prompt. If a sequential file is specified and you specify a member name, an error message displays.

Replace Existing Data?
Specify whether the generated JCL and control statements replace existing data in the specified file.

DSN to Hold SYSIN Data
Name of the sequential file or partitioned data set to hold SYSIN data. If you specify a partitioned data set, specify the member name at the **Member** prompt.

Member (if PDS)
Name of the member in the partitioned data set specified for the DSN prompt. If a sequential file is specified and you specify a member name, an error message displays.

Replace Existing Data?
Specify whether the generated JCL and control statements replace existing data in the specified file.

Submit JCL or Review?
Specify whether the JCL and control statements are saved and submitted, displayed for review, or neither.

- If you select **Submit**, the JCL and control statements are saved and the job is submitted.
- If you select **Review**, use ISPF facilities to save or submit the JCL and control statements.
- If you select **Neither**, the JCL and control statements are saved, but not submitted or displayed for review.

Online Execution

When the Compare Process is executed online, a panel is displayed noting the progress of the process. If one or both of the sources must be extracted, the progress of the extract is shown. The next step, comparing the tables, is then noted, followed by the step in which related change indicators are propagated to related tables.

Here is an example of a Compare Process Status display.

```
+--------------------COMPARE Process Status--------------------+
|                                                              |
| COMPARE Process in Progress                                  |
|                                                              |
| Extracting Source 1 : COMPLETE                               |
|      Total Extracted: 703                                    |
|                                                              |
| Extracting Source 2 : COMPLETE                               |
|      Total Extracted: 704                                    |
|                                                              |
| Comparing Tables    : IN PROGRESS         Table Pair: 1 of 5 |
| Source 1: FOPDEMO.CUSTOMERS                                  |
| Source 2: FOPDEMO.CUSTOMERS                                  |
|                                                              |
| Joining Related Data: Pending                                |
|                                                              |
+--------------------------------------------------------------+
```

As shown on the panel:

- The data is extracted, if necessary. Any errors and warnings were evaluated when the Access Definition was specified on the Specify COMPARE Sources panel. If the process is being performed for an existing Compare Definition through option 3 (PERFORM) directly, the source is evaluated before the Compare Process is performed.
- The tables are compared to identify direct changes and duplicate match key rows.
- The table relationships are evaluated to identify related changes and unusual rows.

Several warnings and error conditions can occur during each step of the Compare Process. If an error is encountered when the job is executed in batch, the job terminates and an error message is written to the output file. If the job is executed online, the job is terminated and a message is displayed on the screen.

If an error is encountered when the data is extracted for either source and an unload program is used, the process is terminated.

Compare Process Report

If you do not invoke the browse when the online Compare Process completes, Compare displays the Compare Process Report.

Here is an example of the Compare Process Report.

```
------------------------- COMPARE Process Report -----------------------
COMMAND ===>                                           SCROLL ===> PAGE
                                                       ROW 0 OF 39
********************************* Top of Data *********************************

                             COMPARE Process Report

Compare File      : FOPDEMO.COMP.FILE
Report File       : FOPDEMO.COMP.RPT
User ID           : FOPDEMO
Time Started      : 2001-06-21 15.37.03
Time Finished     : 2001-06-21 15.37.10

Source 1 - Extract File    : FOPDEMO.EXTRACT
 Total Number of Tables    : 6
 Total Number of Rows      : 0

Source 2 - DB2 Subsystem   : DSNC
 Total Number of Tables    : 6
 Total Number of Rows      : 6608


                                                                  Rows     Non-
COMPARE Results                                        Changes    with    Unique
                             Total    UnMatched Equal    (D)irect  Missing   Match
Source:Table Name            Rows     Rows      Rows     (R)elated Parents   Keys
---------------------------- -------- -------- -------- ---------- -------- ------
1 FOPDEMO.CUSTOMERS             0            0     0      D:    0        0        0
2 FOPDEMO.CUSTOMERS           705          705            R:    0        1        0

1 FOPDEMO.FEMALE_RATES          0            0     0      D:    0        0        0
2 FOPDEMO.FEMALE_RATES         63           63            R:  N/A       56        0

1 FOPDEMO.ORDERS                0            0     0      D:    0        0        0
2 FOPDEMO.ORDERS             1709         1709            R:    0        0        0

1 FOPDEMO.DETAILS               0            0     0      D:    0        0        0
2 FOPDEMO.DETAILS            3596         3596            R:  N/A        0        0

1 FOPDEMO.SALES                 0            0     0      D:    0      N/A        0
2 FOPDEMO.SALES                22           22            R:    0      N/A        0

1 FOPDEMO.SHIP_TO               0            0     0      D:    0        0        0
2 FOPDEMO.SHIP_TO             513          513            R:  N/A        3        0
******************************** Bottom of Data **********************************
```

The report's header information includes the name of the Compare File, the user that generated the report and the time, and the number of tables and rows in the process. The sources are listed along with any pertinent information about each source. The Compare Results section contains the names of the tables from each source, the total number of rows for every type of detail that can be reported on and the total number of rows in the comparison from each table.

Browse a Compare File

The Compare File contains the results of the Compare Process. Using Compare browse, you can examine the results of the Compare Process as a set of related data and focus on the changes.

The status information in the Compare File combined with the powerful capacity to display related data from multiple tables on a single screen, allows you to:

- Exclude rows from the display based on their change status or source.

- Join to other tables to locate related changes.
- Simultaneously scroll from any level of a displayed table.

You can browse the Compare results online by specifying **Y** at the **If Online, Invoke Browse** prompt on the Specify COMPARE Parameters and Execute panel. You can also browse the contents by selecting Option **B BROWSE** on the COMPARE Process menu.

When you select **Option B** from the COMPARE Process menu, the following panel is displayed.

```
---------------------- Specify COMPARE Browse Parameters ---------------------
Command ===>                                                 SCROLL ===> PAGE

Compare File DSN ===>

Browse Options:
  Start Browse with ===> S (S-Start Table, L-Table Selection List)
```

The Specify COMPARE Browse Parameters panel includes the following prompts:

Compare File DSN
Data set name of the Compare File containing the data to be browsed. Compare File DSN initially contains the name of the current or profiled Compare File. If there is no current or profiled Compare File, it is blank. You can overtype the value as desired. To obtain a selection list of Compare Files, specify an asterisk (*) as the last character in the string used as the search value for the data set name.

Start Browse with
Table used to begin the browse session. Specify:

S Begin with Start Table.

L Begin with Compare Summary Selection List panel, listing pairs of tables from which you can select the table to browse. The value you specify is profiled.

When you have completed your responses, use **ENTER** to continue. If you specify **S** for Start Browse with, the Compare browse session is invoked. If you specify **L**, select the desired pair of starting tables from the Compare Summary Selection List panel.

Use **END** or **CANCEL** to terminate the Browse request and return to the COMPARE Process menu.

Compare Summary Selection List

When you select **L** for **Start Browse with** on the Specify COMPARE Browse Parameters panel, the Compare Summary Selection List panel is displayed. You can select a pair of tables to start the browse session. The Compare Summary Select List panel also displays when a browse session is invoked automatically after online execution of the Compare Process.

Here is an example of the Compare Summary Selection List panel.

```
---------------------- Compare Summary Selection List ------------------------
Command ===>                                                    Scroll ===> PAGE

Use 'S' to Select Browse Start Table, 'I' for Extended Table Information

Source 1: XF - Z13600MP.FOPDEMO.EXTRACT, SUBSYS: TDB2
Source 2: DB2 Tables, SUBSYS: TDB2

                                                                  Rows    Non-
                                                        Changes   with   Unique
                           Total     UnMatched Equal    (D)irect  Missing  Match
Sel Source:Table Name      Rows        Rows    Rows     (R)elated  Parents   Keys
--- -------------------->> --------  --------  -------- --------- -------- -------
*** ********************************* TOP *************************************
S__ 1 FOPDEMO.CUSTOMERS    703         1       690      D: 12      N/A        0
    2 FOPDEMO.CUSTOMERS    704         2                R: 25      N/A        0

___ 1 FOPDEMO.ORDERS      1712        10      1697      D:  5       0         0
    2 FOPDEMO.ORDERS      1709         7                R:  4       0         0

___ 1 FOPDEMO.SHIP_TO      503         9       490      D:  4       0         8
    2 FOPDEMO.SHIP_TO      526        32                R: N/A      6        25

___ 1 FOPDEMO.DETAILS     3591        11      3574      D:  6       0         0
    2 FOPDEMO.DETAILS     3596        16                R: N/A      0         0

___ 1 FOPDEMO.ITEMS        102         0       102      D:  0      N/A        0
    2 FOPDEMO.ITEMS        102         0                R:  0      N/A        0
*** ******************************* BOTTOM ***********************************
```

The Compare Summary Selection List panel includes the following items:

Source 1
: Identifies the first source of data and displays the name of the source. The DB2 subsystem from which the data was extracted is also displayed. Source 1 is an Extract File named Z13600MP.FOPDEMO.EXTRACT in the DB2 subsystem TDB2 in the figure.

Source 2
: Identifies the second source of data: an Extract File, Access Definition, or all rows from tables in Source 1, and the DB2 subsystem where the data was compared is also displayed. If Source 2 is all rows from tables in Source 1, table names and subsystems display. In the figure, Source 2 is a set of tables in the DB2 subsystem, TDB2.

I
: Display information about a pair of tables. See "Display Information" for details about the information displayed.

Source:Table Name
: Source identifier followed by the name of the table. For example, the first table in Source 1 is identified as: 1:FOPDEMO.CUSTOMERS

Total Rows
: Number of rows in each table included in the Compare Process.

UnMatched Rows
: Number of rows in each table for which the match key value of that row does not have a corresponding value in a row from the other source table with which it was compared.

Equal Rows
: Number of rows in each pair of tables for which the match key values and all other column values included in the Compare Process match exactly.

Changes
: Number of rows that generated flags for:

  D Direct changes. This is the number of rows in which the match key values match, but one or more of the other column values do not match.

R Related changes. This is the number of rows in dependent related tables that contain changes. This value does not apply to a comparison of a single pair of tables, any table that does not have dependent tables in the Compare Process, or reference tables.

Rows with Missing Parents
Number of rows that are orphans. This value does not apply to a comparison of a single pair of tables or any table that does not have parent tables in the Compare Process.

Non-Unique Match Keys
Number of rows in each table that contain non-unique values in the match key columns and, therefore, are not compared.

Display Information from Compared Tables

You can display information about any pair of compared tables using the **Information** line command. Type **I** in **Sel** for the desired pair. The following is displayed when the CUSTOMERS tables are selected.

```
+-------------------------Extended Compare Table Information--------------------------+
|                                                                              1 OF 25 |
| ------------------------------------------------------------------------------       |
| ********************************** TOP ***********************************          |
|         Source 1 Table Name                  Source 2 Table Name                     |
| ----------------------------------  ----------------------------------               |
| OPTUSER.DEPARTMENT_IMS               OPTUSER.DEPARTMENT_IMS                          |
|                                                                                      |
| Source                         Source Description                                    |
| ------ ---------------------------------------------------------------------         |
|    1   IMS Table, XF: OPTUSER.TEST.EXTRACT                                           |
|    2   IMS Table, DSN: 'OPTUSER.RFM0011.FOPDEMO2.FOPDEPDB'                           |
|                                                                                      |
|      Source 1 Column Name          Source 2 Column Name       Status  Attr          |
| ------------------------------ ------------------------------ ------- ----          |
|                                                                                      |
| DEPTCODE                        DEPTCODE                      MATCH   SAME          |
| DEPARTMT                        DEPARTMT                      COMPARE SAME          |
| DEPTNAME                        DEPTNAME                      COMPARE SAME          |
| LOCATION                        LOCATION                      COMPARE SAME          |
| LOCCODE                         LOCCODE                       COMPARE SAME          |
| FOP_SURKEY                      FOP_SURKEY                    COMPARE SAME          |
|                                                                         Rel          |
|            Related Tables            Type    Name    From Status       Src          |
| ---------------------------------- ------ -------- ---- ----------   ---          |
| EMPLOYEE_IMS                        CHILD   I1_2    IMS  SELECTED      1          |
|                                                                                      |
| --- Compare Match Options ---                                                        |
| Match Key Property   : Unique                                                        |
| ******************************** BOTTOM *********************************          |
+--------------------------------------------------------------------------------------+
```

The information displayed on this panel includes the names of the tables, the columns in each table, the status of the columns, and the names of the related tables included in the Compare Process.

Note: You can also display this information using the **INFO** primary command when browsing the Compare File.

Source Description
For the sources compared, indicates the name of the Extract File, data set name, or subsystem, depending on the source type. For each source, the column names are listed with each column's Status, using one of these values:

MATCH
The columns were included in the match key.

COMPARE
The columns were compared when the values in the paired match key columns were the same.

NOTUSED
The column was excluded from the Compare Process because it was found in one table only or was specifically excluded.

The Attributes column displays next:

Attr Indicates whether the attributes for each pair of compared columns are the same, **SAME**, or different, **DIFF**. Columns with different attributes can be compared when the attributes are compatible. (See the section Chapter 6, "Compare Column Map Compatibility," on page 105 for information.)

Related Tables
The Extended Compare Table Information panel also lists tables directly related to the pair of tables for which column information is provided. For each table, the following is displayed:

Type Indicates whether the related table is the parent or child table in the relationship.

Name The name of the relationship.

From The source of the relationship definition as either DB2 (DB2 Catalog), OPT (Optim Directory), IMS (IMS physical relationship), or IML (IMS logical relationship).

Status Indicates whether the relationship is selected or unselected.

Rel Src
The source of the relationship as either 1 for Source 1 or 2 for Source 2. When the number of lines required for the information exceeds the size of the screen, you can scroll the display. A count is provided to indicate the total number of rows and the relative position in that total of the first displayed row.

Generate Report

When you select Option **R** from the COMPARE Process menu or specify **Yes** to **Generate Reports** on the Specify COMPARE Parameters and Execute panel, you are prompted for the information required to generate the report. You can specify that the contents of the report should include a statistical summary, detail information, or both.

Specify Report Parameters

The summary information includes the names of the tables from each source, the total number of rows for every type of detail that can be reported, and the total number of rows in the comparison from each table.

The details are the rows that have been compared. You can select which details are to be included based on their status. (The status is printed along with each row.)

When you request a report, the following panel is displayed:

```
---------------------- Specify COMPARE Report Parameters -----------------------
Command ===>                                                  SCROLL ===> PAGE

Compare File DSN ===> 'FOPDEMO.TEST.COMPARE.NEW'

Report File DSN  ===> 'FOPDEMO.TEST.REPORT'

  Report Type    ===> D              (S-Summary, D-Detail)
  Lines Per Page ===>                (0-No Titles, 1-99, Blank=57)
  Specify Table Name to Limit Report (Blank for ALL Tables in Compare)
    Table Name ===>

If Detail Report Specify Format and Select Desired Row Types:
  Report Format          ===> C (C-Columnar, S-Sidelabels, E-External)
    If S, Which Columns? ===> A (A-All, D-Different and Key Columns Only)
  All Rows               ===> Y (Y-Yes, N-No)
  Or, if NO, Select One or More of the following Row Types:
    Direct Changes       ===> Y (Y-Yes, N-No)
    Related Changes      ===> Y (Y-Yes, N-No)
    Unmatched Rows       ===> Y (Y-Yes, N-No)
    Orphan Rows          ===> Y (Y-Yes, N-No)
    Duplicate Match Keys ===> Y (Y-Yes, N-No)
  Wide Lines             ===> C (C-Change File, W-Wrap Data)
Display Unused Columns      : Y (Y-Yes, N-No)
```

The Specify COMPARE Report Parameters panel includes the following prompts:

Compare File DSN
Name of the Compare File from which the report is to be generated. This value cannot be modified when this panel is displayed from Options 1, 2 or 3 on the Compare Process menu. This value can be modified when Option R is selected.

Report File DSN
Name of a sequential file to receive the report. You can use the standard ISPF facilities to browse and print the file.

Report Type
Specify whether the report is to include only the summary information, or both the detail and summary information:

S Summary information only.

D Detail and Summary information.

Lines Per Page
Specify the number of lines per page as:

0 Suppress titles.

1-99 Absolute number of lines per page.

blank Use the default, 57 lines per page.

Table Name
Specify the name of a table for which the report is to be generated when multiple tables have been compared. Leave blank to include all tables.

Report Format
Specify whether the report is to be printed in columnar, sidelabels, or external file format. An example of each is provided later in this section. Specify:

C Print in columnar format.

S Print in sidelabels format.

E Print in external (CSV) format. If you choose this value, the External Report Format panel displays, as shown in “External (CSV) Report Format” on page 86.

Which Columns?
Specify which columns are to be included:

A All columns are included.

D Only columns containing changes and the Match Key columns are included.

Select Row Types
You can select all rows by specifying **Y** for All Rows, or select one or more of the following:

Direct Changes
All rows from Source 1 and Source 2 that have the same match key value, but different values in one or more other columns.

Related Changes
All rows from Source 1 and Source 2 that are marked as having related changes. This only applies to multiple table comparisons and is not shown for single table comparisons.

Unmatched Rows
All rows from Source 1 and Source 2 that do not have a match key value that matches a row from the other source.

Orphan Rows
All rows from Source 1 and Source 2 that do not have a parent. This only applies to multiple table comparisons and is not shown for single table comparisons.

Duplicate Match Keys
All rows from Source 1 and Source 2 that have duplicate Match Key values.

Wide Lines
Specify whether the data is to wrap or the file attributes are to be changed when the length of a row exceeds the current file width. Specify:

C Change the file width to fit the longest record in the report file.

W Wrap the report data.

Display Unused Columns
Indicates if unused columns are displayed in the report. This option is specified on the Compare Options panel. For more information about this panel, see the *Common Elements Manual.*

Note: Unused columns are specified in a Column Map used with the Compare Process.

Y Unused columns are displayed.

N Unused columns are not displayed.

The report specifications have no impact on the Compare Process and the contents of the Compare File. They only define the data written to the report. You can use the ISPF facilities to browse and print the reports.

External (CSV) Report Format

When you generate a report, you can choose a Comma Separated Value (CSV) format and specify delimiters to separate the data and various options. To create a CSV report, select option **E-External** on the Compare Process Report Parameters panel.

Here is an example of the External Format Specification panel used to specify this information.

```
----------------------- External Format Specification --------------------------
Command ===>                                                 Scroll ===> PAGE

Compare File DSN            : FOPDEMO.TEST.COMPARE.NEW
Output Format               : Delimited
Generate Header          ===> Y (Y-Yes, N-No)
  Beginning Label        ===> $table,
  Ending Label           ===>
  Header Delimiter       ===> ,
  Use Column Labels      ===> Y (Y-Yes, N-No)
  Match Key Label        ===> M
  Relationship Key Label ===> R
Field Delimiter          ===> ,
String Delimiter         ===> '
  Delimiter Escape Char  ===> '
Use Data Labels          ===> Y (Y-Yes, N-No)
  Changed Data Label     ===> >
  Equal Data Label       ===> =
```

The External Format Specification panel include the following items:

Compare File DSN
Fully-qualified data set name of the Compare file for which the report will be generated. This field is automatically populated and cannot be modified.

Output Format
Format for the output file. This value cannot be modified.

Generate Header
Specify to generate headers using the column names. Allowable values are:

Y Generate headers

N Do not generate headers. This is the default.

Beginning Label
A label placed before the first column in the table. Specify **$table** to include the table name.

Ending Label
A label placed after the last column in table.

Header Delimiter
The character used to separate column headings. A comma (,) is the default.

Use Column Labels
Indicator for using DB2 column labels or column names in the header, if a header is generated. Specify:

Y Use column labels.

N Do not use column labels. Use column names. This is the default.

Match Key Label
Specify a single character to be placed before each column that was used as part of the Match Key between source 1 and source 2 during Compare processing.

Relationship Key Label
Specify a single character to be placed before each column that was part of a relationship when the Optim solution joined tables during Compare processing.

Field Delimiter
Specify a single character used to separate values in a row. A comma (,) is the default.

String Delimiter
Specify a single character used to separate character literal values.

Delimiter Escape Character
Specify a single character used to generate the value of a character normally used as a delimiter.

Use Data Labels
Specify whether the Optim solution places a single character before the data of each column to indicate if the column data in source 1 and source 2 is equal or unequal.

Y Use data labels.

N Do not use data labels. This is the default.

Changed Data Label
Specify a single character to be placed before column data that differs between source 1 and source 2.

Equal Data Label
Specify a single character to be placed before column data that is equal between source 1 and source 2.

When you press **Enter** or use the **End** command to exit the External Format Specification panel, the Optim solution validates your entries. Pressing **Enter** again displays the following panel:

```
---------- External Format Destination Data Set Specification ---------
Command ===>                                          Scroll===> PAGE

     Optionally specify the name of a destination data set where
     The external format report should be written for the associated
     table or view.

Default Destination DSN : FOPDEMO.TEST.REPORT

Table/View                  Destination Data Set
--------------------------- ----------------------------------------
******************************** TOP***********************************
FOPDEMO.ORDERS
FOPDEMO.DETAILS
******************************** BOTTOM********************************
```

This panel includes the following item:

Default Destination DSN
Name of the default output report data set that was specified on the Specify COMPARE Report Parameters panel. This field is automatically populated and cannot be modified.

You have the option of writing report information for a specific table or view to a destination file. In the Table/View list, supply a data set name next to the name of the table or view. The report information for that table will be written to the specified file. If you do not specify a data set name, report information for all tables in the Compare file is written to the default destination data set.

Sample Reports

Sample reports are provided to facilitate describing the contents of the reports. Samples of each report type, columnar, sidelabels, and External format (CSV), are shown in this section.

Sample Report - Columnar Format

The following is a sample of a Compare File report in columnar format. For this sample, the data has been truncated.

Consider a scenario comparing different versions (source 1 and source 2) of the related ORDERS (parent) and DETAILS (child) demo tables. In this example, one direct change exists in each of the table pairs. A columnar report for this comparison that displays only the direct and related changes would look as shown:

```
Chg  Src ORDER_ID   CUST_ID   ORDER_DATE   ORDER_TIME   FREIGHT_CHARGES
---- --- ++++++++++ --------- ------------ ------------ -----------------
 DR  1           11 00026     1998-01-26   14.22.31 >              9.22
 DR   2          11 00026     1998-01-26   14.22.31 >              9.23

Chg  Src ORDER_ID   ITEM_ID   ITEM_QUANTITY   DETAIL_UNIT_PRICE
---- --- ++++++++++ +++++++++ --------------- -------------------
 D   1           11 DR011                   1>              20.00
 D    2          11 DR011                   1>              30.00
```

Although it is not shown in the previous example, a columnar format report includes header information that shows the name of the Compare File, the user that generated the report, the time, and the number of tables in the process. The sources are listed along with any pertinent information about each source. This is followed by the requested report information for each table. In this example, the information is provided for the ORDERS table, the first table listed in the Compare Process.

The **Summary Information** is the same as that provided on the Compare Summary Selection List panel. If you request only summary information, the information for each table is provided one after the other in the order in which the tables were specified for Source 1. Only the pertinent summary information is included. This report documents a Compare Process that included multiple tables; however, if only one pair of tables or a reference table is involved, statistics about related rows and orphan rows would not be included. (For an explanation of this information see "Compare Summary Selection List".)

The **Details**, rows from the compared tables, are included after the summary information. This portion of the report is divided into the following:

Chg Identifies the rows containing changes, the rows for which related rows contain changes, orphan rows and duplicate match key rows. Symbols are used as the identifiers.

- **D** One or more columns, other than the match key columns, are different.
- **R** One or more columns, other than the match key columns, in a related row are different.
- **U** Identifies rows that are orphans and the pairs of compared rows that result in a different set of related data.

  Orphan rows in a table can only be determined when data from multiple tables is compared and the parent table is included.

  Different sets of related data can result when matched rows contain data in a relationship column that results in joining to a different child.

  **/ | **
  Identifies rows with duplicate match keys.

  If only two rows contain duplicate values:

  ```
  / 00001 Video Magic
  \ 00001 Video Mania
  ```

  If three or more rows contain duplicate values:

  ```
  / 00001 Video Magic
  | 00001 Video Mania
  \ 00001 Video Maniac
  ```

Src Identifies the source of each row as:

- **1** Source 1
- **2** Source 2

12 Source 1 and Source 2. The row is identical; therefore, included once.

The data is presented prior to this information. The match key columns are presented first and plus signs (+) are used to separate the data from the column name. The compared columns are listed next, followed by the columns unique to Source 1 and Source 2. The column headings defined for Source 1 are used for the compared columns. The column headings for columns unique to a source are prefixed with the source identifier 1: or 2:, as appropriate. (For example, the heading for a column named CUST unique to Source 1 is displayed as 1:CUST.)

Unchanged rows (rows in which all compared column values match) are included once. Both the Source 1 and Source 2 versions of changed rows are included. Any rows unique to one source are included and the source is identified.

Within each pair of changed rows, the changed columns are identified by the symbol, ">". For example, in the following pair of compared rows, the second and third columns are different.

```
Chg       Src       CUST_ID          CUSTNAME             ADDRESS
---       ----      ++++++++      -------------       --------------
 D          1          00100      >CinemaMagic        >726 West State Street
 D          2          00100      >CineMagic          >123 East State Street
```

This format simplifies interpreting the results of the comparison.

Sample Reports - Sidelabels Format

The sidelabels Compare File print report in the following figure is especially useful for focusing directly on the changes. Therefore, the report parameters limit this sidelabels report to include only unmatched rows and rows with direct and related changes. For unmatched rows, all columns are included. For those rows with direct and related changes, only the columns containing differences are included. The Match Key columns are always included for every row in the report.

```
6/22/2012                Optim - Compare File Report            Page:1

Compare File           : OPTUSER.TEST.COMPARE2
Created by             : Job OPTUSER using SQLID OPTUSER on 22 Jun 2012 AT 11:25 am
Number of Tables       : 3

Source 1               : Extract File - OPTUSER.TEST.EXTRACT
   Created by          : Job OPTUSER using SQLID OPTUSER
   Created on          : 22 Jun 2012 AT 11:11 am from DB2 Subsystem DD8F

Source 2               : Access Definition - OPTUSER.RFM0097.IMS
   Created by          : Job OPTUSER using SQLID OPTUSER
   Created on          : 22 Jun 2012 AT 11:25 am from DB2 Subsystem DD8F

Table Statistics:      1:OPTUSER.DEPARTMENT_IMS 2:OPTUSER.DEPARTMENT_IMS

   1 - IMS Table, XF: OPTUSER.TEST.EXTRACT
   2 - IMS Table, DSN: 'OPTUSER.RFM0011.FOPDEMO2.FOPDEPDB'

   Total Number of Merged Rows                 :       5
   Total Number of Rows from Source 1          :       5
   Total Number of Rows from Source 2          :       5
   Number of Unmatched Rows from Source 1      :       0
Number of Unmatched Rows from Source 2         :       0
Total Number of Equal Rows                     :       4
Total Number of Rows with Direct Changes       :       1
Total Number of Rows with Related Changes      :       0
Total Number of Orphan Rows on Source 1        :     N/A
Total Number of Orphan Rows on Source 2        :     N/A
Sets of Rows with Non-Unique Keys              :       0
Total Number of S1 Rows with Non-Unique Keys   :       0
Total Number of S2 Rows with Non-Unique Keys   :       0

Compare Match Options:

  Match Key Property      : Unique

Column Name                    Src Data
------------------------------ --- ------------------------------Source:12 Changes:
DEPTCODE                       KEY S03
------------------------------ --- ------------------------------Source:12 Changes:
DEPTCODE                       KEY S05
------------------------------ --- ------------------------------Source:12 Changes:
DEPTCODE                       KEY S12
------------------------------ --- ------------------------------Source:1/2 Changes:
DEPTCODE                       KEY S13
DEPARTMT                       1   S13Sales - Europe               Pari , France
PAR
                                2  S13Sales - Europe               Paris, France
PAR
LOCATION                       1   Pari , France
                                2  Paris, France
------------------------------ --- ------------------------------Source:12 Changes:
DEPTCODE                       KEY S21
--------------------------------------------------------------------------------
```

Only the detail information is affected by the columnar and sidelabels format. Therefore, for details about the report headers and summary information, see the discussion provided for the columnar format sample.

The following items are included on the report:

Details

The details are presented after the summary information.

Headings are provided for the column information: **Column Name**, **Src**, and **Data**.

Column Name
The column name is displayed only once, regardless of whether the data is unique, common, or different between sources.

Src Indicates the source of the column data. Possible values are:

1 Source 1 only.

2 Source 2 only.

12 Common to Source 1 and Source 2.

KEY Match Key column.

Data The data is presented in 50-byte segments. That is, if the data exceeds 50 characters, the next 50 characters are placed on the next line. As many lines as are needed to include all of the data in the column are used.

When all columns are included, the data in the common columns are displayed once and Src contains 12.

The Match Key columns are listed first in the report and identified with KEY in **Src**.

A separator line is included after these headings and after each set of columns from a single row. The following row information is then provided:

Source
Indicates the row source. Possible values are:

1 Row is from Source 1. There is no matching row in Source 2.

2 Row is from Source 2. There is no matching row in Source 1.

12 Row is a common row. All compared columns in Source 1 and Source 2 match.

1 / 2 Compared rows have differences in columns, other than Match Key columns. Use the information in **Src** to determine the source of any unmatched column values.

Changes
Indicates the type of change. Possible values are:

blank Row is unique to one source.

D One or more columns, other than the match key columns, are different.

R One or more columns, other than the match key columns, in a related row are different.

U Identifies rows that are orphans and the pairs of compared rows that result in a different set of related data. Orphan rows in a table can only be determined when data from multiple tables is compared and the parent table is included.

/
|
\

Identifies rows with duplicate Match Keys.

Sample Reports - External (CSV) Format

In this example, different versions of the ORDERS and DETAILS tables are compared. When browsed, a portion of the External Format report produced for this comparison would look as shown in the following example:

```
FOPDEMO.ORDERS,'Chg','Src',MR'ORDER_ID','CUST_ID','ORDER_DATE','ORDER_TIME','FREIGHT_CHARGES'
' DR ','1 ',=11,='00026',='1998-01-26',='14.22.31',>9.22
' DR ',' 2',=11,='00026',='1998-01-26',='14.22.31',>9.23
FOPDEMO.DETAILS,'Chg','Src',MR'ORDER_ID',M'ITEM_ID','ITEM_QUANTITY','DETAIL_UNIT_PRICE'
' D ','1 ',=11,='DR011',=1,>20.00
' D ',' 2',=11,='DR011',=1,>30.00
```

This report was produced using the following options:

- The table name is included in the header because **$table** was specified as the **Beginning Label** parameter (this is similar to the Extract Convert function).
- The header includes the **Chg** and **Src** (Change and Source) compare columns and the column names. Column names are delimited by a comma (,), as specified in the **Header Delimiter** parameter.
- In the header, the primary key columns (**ORDER_ID** for the ORDERS table, and **ORDER_ID** and **ITEM_ID** for the DETAILS table) are preceded by **M**. This is the specified value for **Match Key Label**.
- The column **ORDER_ID** in the header is preceded by **R**. This is the specified value for **Relationship Key Label**, and **ORDER_ID** is the key column used in the ORDERS and DETAILS relationship.
- All data columns are separated by the specified **Field Delimiter** parameter value of a comma (,).
- Data columns that have direct changes are preceded by the specified **Changed Data Label** value of >.
- Data columns that are equal are preceded by the equal sign (=), the specified **Equal Data Label**.
- Summary information is not included, and there is no graphical data, such as dashes to separate column headers.

Chapter 5. Batch Utility

Most options that may be specified using the online Compare panel are also available using the COMPARE batch processing statement. Use a COMPARE statement to compare two sets of related data. You can compare a set of rows that resides in database tables, VSAM or sequential files, or rows extracted previously and stored in an Extract or Archive File. Results of a comparison are stored in a Compare File or can be used to create a report.

COMPARE Batch Utility Statement

```
COMPARE
  COMPARE_FILE ( File Allocation Parameters )
   {COMPARE_DEFINITION_DEFINE  (parameters) ; |
    COMPARE_DEFINITION  group.user.name
     [ SOURCE_FILE_OVERRIDE dsname ]
     [ SOURCE_CID OVERRIDE { 1 | 2 } { cid | % }
     [MAX_EXTRACT_ROWS  nn ]
     [ UNLOAD_UTILITY { BMC | IBM }
         [ IMAGE_COPY
             [ MODE { L | A | B | D }
             [ DATE yyyy-mm-dd ]
             [ TIME hh.mm.ss ]
             [ DSNAME dsname ] ) ]
     [ PARTITIONS_n ( part1,...partn ) ] ]
        [OTHER_PARTITIONS_SAME {YES | NO }]]
     [ MATCH_KEY ( creatorid.tablename,oldkeycol,newkeycol ) ]
     [ REPORT_MODE { YES | NO | ONLY }
         REPORT_FILE ( File Allocation Parameters )
        [ REPORT_LINES nn ]
        [REPORT_TABLE  [ creatorid. ] tablename]
        [REPORT_TYPE  { SUMMARY | DETAIL | BOTH } ]
        [ ROW_DISPLAY (DIRECT |RELATED |UNMATCHED |ORPHAN |DUPLICATE |UNUSED ) ]
        [ WRAP_LINES { EXTEND | WRAP } ] ]
        [ REPORT_FORMAT { COLUMN | SIDE_DIFF | SIDE_ALL | EXTERNAL} ]
            [ TABLE (cid.tablename,dsname)]
               [ STRING_DELIMITER 'c']
               [ ESCAPE_CHARACTER 'c']
               [ FIELD_DELIMITER ',']
               [ GENERATE_HEADER {YES | NO}]
                  [ BEGIN_LABEL label ]
                  [ END_LABEL label ]
                  [ HEADER_DELIMITER  ',']
                  [ MATCH_KEY_LABEL  'c' ]
                  [ REL_KEY_LABEL 'c' ]
                  [ USE_COLUMN_LABELS {YES | NO}]
               [ USE_DATA_LABELS {YES | NO}
               [ CHANGE_DATA_LABEL 'c' ]
               [ EQUAL_DATA_LABEL 'c'] }
```

COMPARE_FILE

Name of the Compare File that is to contain the results of the comparison. Use the keywords in File Allocation Parameters to name and allocate the Compare File. For details, refer to the File Allocation Parameters section in the *Batch Utility Guide*.

If you use the REPORT_MODE ONLY keyword, you need only specify the COMPARE_FILE dataset name.

Batch Utility Compare Definition Parameters

Use the following keywords to identify the Compare Definition and override or augment various parameters in it.

COMPARE_DEFINITION_DEFINE

The Optim online process generates this keyword when it creates a COMPARE job for batch execution.

When you create a COMPARE job outside of the Optim online process, the best practice is to use the COMPARE_DEFINITION keyword to refer to a named Compare Definition in the Optim Directory. Place the COMPARE_DEFINITION_DEFINE parameters within parentheses, with a semicolon after the last parenthesis.

COMPARE_DEFINITION

The name of the Compare Definition. COMPARE_DEFINITION must be included in the COMPARE statement and must precede any keywords that override parameters in the Compare Definition (for example, SOURCE_FILE_OVERRIDE_1).

group.user.name

The three-part Compare Definition name.

SOURCE_FILE_OVERRIDE

Override for the name of a file specified as a source in the Compare Definition. To override the Source 1 specification, use SOURCE_FILE_OVERRIDE_1. Use SOURCE_FILE_OVERRIDE_2 to override the Source 2 specification.

dsname

The fully qualified name of the Archive File, Extract File or Access Definition.

SOURCE_CID_OVERRIDE_{ 1 | 2 }

Override for the default creator ID for a source in the Compare Definition. To override the Source 1 specification, use SOURCE_CID_OVERRIDE_1. Use SOURCE_CID_OVERRIDE_2 to override the Source 2 specification.

cid Creator ID to override the default Creator ID for the source.

If the source is an Extract File, you can use the wildcard (%) to use the default Creator ID defined in the Extract File.

This Creator ID applies to all tables that were defined with the initial default Creator ID in the Compare Definition. Tables in the Compare Definition not defined with the initial default Creator ID will not be altered by the specified Creator ID override. This override also changes the names of the tables in the relationships processed by Compare to use the override default Creator ID. If a relationship is not found for the updated table name, an error occurs when the compare is performed. If the source is an extract file, and a table initially defined in the Compare definition with default Creator ID is not in the extract file, the table will be skipped, a warning message will be produced, but the Compare job will continue.

MAX_EXTRACT_ROWS

The maximum number of rows that can be extracted from the source table. If this value is exceeded, the Compare process terminates. The default is the site option value for Maximum Extract Rows. The Compare Definition must specify a table or an Access Definition as a source. Use MAX_EXTRACT_ROWS_1 for Source 1 and MAX_EXTRACT_ROWS_2 for Source 2.

nn The maximum number of rows to extract during the Compare process.

Use these parameters only for a Compare Definition that specifies a DB2 table or an Access Definition that includes a DB2 table as a source.

UNLOAD_UTILITY

The name of the unload utility used to retrieve the data. If omitted, DB2 is used to retrieve the data. This keyword is required to extract data from an image copy.

Use these keywords only for a Compare Definition that specifies a DB2 table or an Access Definition that includes a table as a source. Use UNLOAD_UTILITY_1 for Source 1 and UNLOAD_UTILITY_2 for Source 2.

BMC Use the UNLOAD PLUS utility.

IBM Use the High Performance Unload utility.

Note: You must add the appropriate utility DD statements. Refer to the utility documentation for the required DD statements.

IMAGE_COPY
The image copy file(s) to be used as input for a source. Use this parameter only if a DB2 table or an Access Definition that includes a DB2 table has been specified as a source, and the UNLOAD_UTILITY parameter has been specified. If omitted, the DB2 database is used as input. IMAGE_COPY_1 is allowed only if UNLOAD_UTILITY_1 is used; IMAGE_COPY_2 is allowed only if UNLOAD_UTILITY_2 is used.

MODE
The image copy file to use:

L Use the latest file.

A Use the first file created on or after the specified DATE and TIME.

B Use the first file created on or before the specified DATE and TIME.

D Use the file named in DSNAME.

DATE The date criteria for the image copy file in this format: *yyyy-mm-dd*.

TIME The time criteria for the image copy file in this format: *hh.mm.ss*. If omitted, a value of 00.00.01 is assumed.

DSNAME
The fully qualified name of the image copy file.

PARTITIONS
The partitions to be processed by an unload utility. Valid only if the Start Table for the source is a partitioned DB2 table. Use PARTITIONS_1 for Source 1 and PARTITIONS_2 for Source 2. Omit this keyword to process all partitions.

(*part1,...partn***)**
The partition numbers. Partitions that are not specified are not compared.

OTHER_PARTITIONS_SAME
Use this operand to indicate that the partitions specified in the PARTITIONS operand for the Start Table should be applied to all tables in the Compare Definition.

YES The partitions specified in the PARTITIONS operand for the Start Table will be applied to all tables in the Compare Definition.

NO The partitions specified in the PARTITIONS operand will apply only to the Start Table. All partitions for all other tables in the Compare Definition will be processed. This is the default.

MATCH_KEY
Replace a key column in the Match Key for the specified table.

(*creatorid.tablename,oldkeycol,newkeycol***)**
The fully qualified tablename, old key column name, new key column name.

REPORT_MODE
Specify whether a report is produced.

YES Produce a report as part of the Compare process.

NO Do not produce a report (default).

ONLY Produce a report from a Compare file created previously and specified in the COMPARE_FILE keyword.

REPORT_FILE
Name of the file for the report. Use the parameters in File Allocation Parameters to name and allocate the file. This keyword is required if REPORT_MODE is YES or ONLY.

REPORT_LINES
The number of lines per page for the report.

nn Number of lines per page. Valid values are 0-99. The default is 57. Specify 0 to suppress all page breaks and title lines.

REPORT_TABLE
The name of a table for which the report is generated when multiple tables have been compared. Omit this keyword to include all tables.

[creatorid .]tablename
The name of the table for which the report is generated. If the creator ID is omitted, all tables with the specified name are included, regardless of the creator ID.

REPORT_TYPE
The type of report to be produced.

SUMMARY
Produce a report containing summary information.

DETAIL
Produce a report of detailed information.

BOTH Produce a report containing both summary and detailed information (default).

ROW_DISPLAY
Rows that appear on a detailed report. If you omit this keyword, all rows are included. Operands must be enclosed in parentheses, separated by commas, and may appear in any order.

DIRECT
Include rows with the same Match Key value that have different values in one or more other columns.

RELATED
Include rows marked as having related changes.

UNMATCHED
Include rows with match key value that does not match any row from the other source.

ORPHAN
Include rows that do not have a parent.

DUPLICATE
Include rows that have duplicate match key values.

UNUSED
Include unused columns in the report.

WRAP_LINES
Option for rows that exceed the line width of the report file for a detailed report.

WRAP
Wrap the report data (default).

EXTEND
Increase the record length of the report file to fit the row.

REPORT_FORMAT
The format of a detailed report.

COLUMN
Report is in columnar format (default).

SIDE_ALL
Report is in sidelabels format and includes all columns.

SIDE_DIFF
Report is in sidelabels format and includes only changed columns and Match Key columns.

EXTERNAL
Report is in CSV (Comma Separated Values) format.

The following parameters are used only when REPORT_FORMAT=EXTERNAL is specified.

TABLE
Write the output for a specific table to an external file other than the file specified in the REPORT_FILE keyword. All file allocation parameters for the REPORT_FILE parameter, except dsname, apply to this file. You can use multiple TABLE parameters, provided each refers to a different table.

(*cid.tablename,dsname***)**
The fully qualified table name and data set name, enclosed in parentheses and separated by a comma.

STRING_DELIMITER
Specify a single character used to separate character literal values.

ESCAPE_CHARACTER
Specify a single character used to generate the value of a character normally used as a delimiter.

FIELD_DELIMITER
Specify a single character used to separate values in a row. A comma (“,”) is the default.

GENERATE_HEADER
Generate headers using the column names.

Y Generate headers

N Do not generate headers. This is the default.

BEGIN_LABEL
Place a label before the first column in the table. This parameter applies only when GENERATE_HEADER=Y. Specify “$table” to include the table name in the label.

END_LABEL
Specify a label to be placed after the last column in table. This parameter applies only when GENERATE_HEADER=Y. There is no default.

HEADER_DELIMITER
Character used to separate column headings. This parameter applies only when GENERATE_HEADER=Y. A comma (“,”) is the default.

MATCH_KEY_LABEL
Specify a single character to be placed before each column that was used as part of the Match Key between Source 1 and Source 2 during compare processing.. This parameter applies only when GENERATE_HEADER=Y.

REL_KEY_LABEL
Specify a single character to be placed before each column that was part of a relationship when tables were joined during compare processing. This parameter applies only when GENERATE_HEADER=Y.

USE_COLUMN_LABELS
Indicator for using DB2 column labels or column names in the header, if a header is generated. Specify:

Y Use column labels.

N Do not use column labels. Use column names. This is the default.

USE_DATA_LABELS
Specify whether to place a single character before the data of each column to indicate if the column data in Source 1 and Source 2 is equal or unequal.

Y Use data labels.

N Do not use data labels. This is the default.

CHANGE_DATA_LABEL
Specify a single character to be placed before column data that differs between Source 1 and Source 2. This parameter applies only when USE_DATA_LABELS=Y.

EQUAL_DATA_LABEL
Specify a single character to be placed before column data that is equal between Source 1 and Source 2. This parameter applies only when USE_DATA_LABELS=Y.

This example compares two sets of data defined by the Compare Definition PSTUSER.INV.CDQ4. The Extract File PSTUSER.ITEMS is used in place of the Extract File specified for Source 1 in the Compare Definition:

```
COMPARE_FILE (DSNAME PSTUSER.INVCOMP30)
   COMPARE_DEFINITION PSTUSER.INV.CDQ4
   SOURCE_FILE_OVERRIDE_1 PSTUSER.ITEMS
```

Batch Utility Statement - Table Map Define Parameters

You can use the TABLE_MAP_DEFINE parameters in the COMPARE Batch Utility control statement. The parameters must be enclosed within parentheses, with a semicolon after the close parenthesis.

```
[ DESCRIPTION //description// ]
[ SECURITY { PUBLIC | READONLY | PRIVATE } ]
OWNER ownerid
MODIFIED yyyy-mm-dd-hh.mm.ss
VALRULES { C | M }
SRCCID sourceid DESTCID destid [COLMAPID colmapid]
 { SRCEXT dsname | SRCAD ad }
 { DESTEXT dsname | DESTAD ad }
 [ SRCTYPE { X | A | T } ] [ DESTTYPE { X | A | T } ]
     ( sourcetablename = desttablename
        [ { CM mapname | LOCALCM (mapdef ) } ]
 [ LEGTYPE { D | I } ] [ LEGDSN dsname ]
 [ LEGPSB psbname ] [ LEGPCB pcbnum ]
 [ LEGSEG segname ] [LEGEXIT  exitname ]
[ PROCESSMODE { I | U | B } ] [ DELETEFLAG { Y | N } ]
[ MATCH_KEY name] [ USER_MATCH_KEY name]
[ ACCESS_METHOD { K | S } [KEY_LIMIT  n]
[ MATCH_KEY_PROPERTY value ] [ MATCH_METHOD value]
[ SOFT_MATCH_ENABLED { Y | N } ][ PERCENT_REQUIRED percentage ]
[ PERCENT_ACCEPTABLE percentage] [ MAXIMUM_ATTEMPTS n]
```

DESCRIPTION
The optional description. This up to 40-character description must be enclosed in double slashes.

SECURITY
The security assigned to the object, specified as:

PUBLIC
All can use and modify (default).

READONLY
All can use, but only owner can modify.

PRIVATE
Only owner can use and modify.

OWNER
The owner of the object.

MODIFIED
The date and time the object was last modified.

VALRULES
The validation rules for the Table Map, specified as:

M Move or Archive rules.

C Compare rules.

SRCCID
The default Creator ID for the source tables.

DESTCID
The default Creator ID for the destination tables.

COLMAPID
The optional default Column Map ID.

The following parameters are included, depending on the source and destination types, as well as whether the Table Map is for Move, Archive, or Compare.

SRCEXT
The name of the data set containing an Extract or Archive File used as the source.

SRCAD
The fully qualified name of the Access Definition used as the source.

DESTEXT
The name of the data set containing an Extract or Archive File used as the destination.

DESTAD
The fully qualified name of the Access Definition used as the destination.

SRCTYPE
The source type as:

X Extract or Archive File

A Access Definition

T Table

DESTTYPE
The destination type as:

X Extract or Archive File

A Access Definition

T Table

Table Mapping

The following source table to destination table mapping information is provided for each pair of tables referenced in the Table Map. At least one pair of tables must be specified.

sourcetablename
The source table name. The Creator ID is included if it differs from the SRCCID.

desttablename
The destination table name. The Creator ID is included if it differs from the DESTCID. The words NOT-SPECIFIED are inserted when the destination table in a pair is omitted.

CM The name of the Column Map for the pair of tables. The Map ID is included if it differs from the COLMAPID.

Note: CM or LOCALCM is included if a Column Map has been specified for a pair of tables.

LOCALCM
The local Column Map definition, enclosed in parentheses. Only the source and destination mapping (*source-expr* = *dest-column*) is included here. For details, see Export and Import Optim Objects in the *Common Elements Manual.*

The following parameters, with names prefixed by "LEG" apply only to Legacy Tables referenced in the Table Map. The type of legacy source file determines which parameters are included. (Compare for I/V/S only)

LEGTYPE
The type of legacy source file to be used with the Legacy Table:

D VSAM or Sequential file

I IMS™ file

LEGDSN
The name of the legacy source data set.

LEGPSB
The name of the PSB that provides access to the IMS services that Compare for I/V/S requires to access the database records. (Included only if LEGTYPE=I)

LEGPCB
The relative number of the database PCB within the specified PSB that grants Compare for I/V/S the authorization to manipulate the data.

LEGSEG
The names of the segments within the specified DBD. (Included only if LEGTYPE=I)

Note: Additionally, a Table Map referencing one or more Legacy Tables for IMS data includes parameters for any Retrieval Definitions that are referenced by the Legacy Table. (For details, see Export and Import Optim Objects in the *Common Elements Manual*.)

LEGEXIT
The name of the I/O Exit Load module (if different from the default value). This exit will be called only if the appropriate **Enable I/O Exit** site option is **U**. Otherwise it will be ignored.

PROCESSMODE
The processing mode for the table.

I Insert only.

U Update only.

B Both insert and update.

DELETEFLAG
Delete parameter for the table.

Y Delete all rows in the destination table prior to inserting data.

N Do not delete.

MATCH_KEY
For a Compare process only. Name of the match key used for this comparison.

USER_MATCH_KEY
For a Compare process only. Name of the user-defined match key for this comparison.

ACCESS_METHOD
(This parameter is not used for Convert processes.) The method used to access the parent or child table for each relationship. Specify:

K Key lookup locates rows using a WHERE clause to search for primary or foreign key values. Additionally, you can change the maximum number of key lookups performed at one time for a table.

S A scan reads every row in a table at one time.

KEY_LIMIT
Value for the maximum number of key lookups performed at one time for a table. This value is used whenever a key lookup is the access method used to scan a table and no specific key lookup limit is specified. Specify a value in the range 1 to 100.

Note: If used, the following parameters must be specified in this order:
MATCH_KEY_PROPERTY
MATCH_METHOD
SOFT_MATCH_ENABLED
PERCENT_REQUIRED
PERCENT_ACCEPTABLE
MAXIMUM_ATTEMPTS

MATCH_KEY_PROPERTY
Indicates the presence of a match key for the sources to be compared. The value in this field is automatically populated, based upon the known attributes of the data. If this value is Unique, the remaining fields on this panel are non-modifiable. Specify one of these values:

U Unique

M Multiple, non-unique.

N Not keyed. Data has no key defined.

MATCH_METHOD
Controls the processing for non-uniquely keyed and non-keyed sources. This field is not modifiable and displays the value ANY. For details see Soft match options in “Match Options Panel” on page 53.

SOFT_MATCH_ENABLED
(For Compare only.) Controls whether soft matching is performed. Specify:

Y Enables soft matching. Unequal rows are compared and the best matches are paired and considered "changed".

N Disables soft matching. All unequal rows are considered “changed”. Unequal rows are considered to be source 1 or source 2 only. If you specify N, the other Soft Match parameters are ignored.

Note: If the match key property is Unique, soft matching cannot be used.

PERCENT_REQUIRED

(For Compare only) This parameter is valid only if SOFT_MATCH_ENABLED is Y. Specifies the percentage of similarity required to consider two unequal rows matched. This value applies only to the columns that are not defined as part of the match key. Allowable values range from 1 through 100. The default is 50 percent.

PERCENT_ACCEPTABLE

This parameter is valid only if SOFT_MATCH_ENABLED is Y. This parameter limits processing by reducing the precision of the comparison. Normally, all of the rows in source 1 are compared to all of the rows in source 2 and the best matches that meet the Percent Required are selected. If you specify a value less than 100 for Percent Acceptable, match processing for a row will stop when a row comparison meets or exceeds this percentage. Allowable minimum value is the value specified for PERCENT_REQUIRED; allowable maximum is 100. The default is 100.

MAXIMUM_ATTEMPTS

This parameter is valid only if SOFT_MATCH_ENABLED is Y. Limits the number of compares performed for any set of unequal rows. The value 0 disables this limit. Specify a value in the range 1 to 999,999. The default is 0.

Chapter 6. Compare Column Map Compatibility

The Compare rules for Column Maps allow you to map columns with dissimilar names and columns with compatible data types. Compatible data types are determined as described in this topic.

Compatible data types are determined as follows:

- DATE, TIME, TIMESTAMP, or TIMESTAMP WITH TIME ZONE data type columns can be compared only with a column of the same data type.
- CHAR, VARCHAR, LONGVARCHAR data types can be compared.
- GRAPHIC and VARGRAPHIC data types can be compared
- Comparisons cannot be performed on LOB or XML data.
- Except for the binary data types, the shorter value is padded with blanks before the comparison.
- BINARY and VARBINARY can be compared. Binary values are padded with binary zeroes. When browsing a Compare File, the longer data type is assumed.
- INTEGERS, SMALLINT, and BIGINT columns can be compared. When browsing the Compare File, the data is displayed as INTEGER.
- DECIMALS of any scale and precision can be compared. When comparing DECIMALS of different scale, precision, or both, the larger in each dimension is used and displayed when browsing the Compare File. The resulting number of digits cannot exceed 31. For example, if comparing a column DEC(4,2) with a column DEC(4,4), the resulting dimension and precision is DEC(6,4).
- SINGLE PRECISION FLOATING POINT and DOUBLE PRECISION FLOATING POINT columns can be compared. When browsing the Compare File, the values are displayed as DOUBLE PRECISION FLOATING POINT.

Compatibility Rules

When you define a relationship or a Column Map, corresponding columns must be compatible. For relationships, column compatibility rules are the same for all Optim components. There are differences in the compatibility rules for Column Maps, depending on the Optim component. Included here are the compatibility rules for Column Maps used with Compare.

Relationship Column Compatibility

The following list shows compatibility for each column data type when you are creating or modifying a relationship. The columns listed are compatible and may be defined as corresponding columns. Here, the term “column” refers to actual columns, as well as expressions containing concatenation, substrings, and literals.

Numeric

- Numeric
- CHAR [2]
- VARCHAR [2]
- Numeric Constant [3]

CHAR

- Numeric [2]
- CHAR [4]
- VARCHAR [4]
- String Literal

- String Expression [1]
- CLOB [4]
- TIMESTAMP WITH TIME ZONE

VARCHAR

- Numeric [2]
- CHAR [4]
- VARCHAR [4]
- String Literal
- String Expression [1]
- CLOB [4]
- TIMESTAMP WITH TIME ZONE

DATE

- DATE
- TIMESTAMP WITH TIME ZONE

TIME

- TIME
- TIMESTAMP WITH TIME ZONE

TIMESTAMP

- CHAR
- VARCHAR
- DATE
- TIME
- TIMESTAMP
- TIMESTAMP WITH TIME ZONE

TIMESTAMP WITH TIME ZONE

- CHAR
- VARCHAR
- DATE
- TIME
- TIMESTAMP
- TIMESTAMP WITH TIME ZONE

StringLiteral

- CHAR
- VARCHAR
- String Expression [1]

Graphic Literal

- DBCLOB
- VARGRAPHIC

Numeric Constant

- Numeric [3]

String Expression [1]

- CHAR
- VARCHAR

- String Literal
- String Expression [1]

Binary Expression

- BINARY
- VARBINARY

CLOB

- CHAR [4]
- VARCHAR [4]
- CLOB [4]

DBCLOB

- Graphic Literal
- DBCLOB
- GRAPHIC
- VARGRAPHIC

GRAPHIC

- Graphic Literal
- DBCLOB
- GRAPHIC
- VARGRAPHIC

VARGRAPHIC

- Graphic Literal
- DBCLOB
- GRAPHIC
- VARGRAPHIC

BINARY

- Binary Literal
- BINARY
- VARBINARY

VARBINARY

- Binary Literal
- BINARY
- VARBINARY

Note:

[1] "String Expression" is any comparand containing a concatenation operator or a substring function.

[2] The "Numeric" column must be an INTEGER, SMALLINT, BIGINT column or a DECIMAL column with a scale of 0 (the numeric value cannot have digits after the decimal point).

[3] The "Numeric" column must be an INTEGER, SMALLINT, BIGINT, or DECIMAL column.

[4] If the parent column is for mixed data, the child column must also be for mixed data.

Verification

When you press **ENTER** on the panel to define or modify the relationship, the Optim solution verifies that corresponding columns are compatible. If they are not, an error message is displayed. You must correct the incompatibility before saving the relationship definition.

Comparing Character to Numeric Columns

Character columns are compatible with numeric columns when defined in an Optim relationship. To compare character data to numeric data, it may be necessary to convert the numeric data to character format. This may involve padding the data prior to the comparison. The Optim solution pads the data with zeros so the two strings are equal in length. This operation is performed internally and does not affect your DB2 data. To demonstrate, assume the following columns are defined in a relationship.

```
COL1 CHAR(5) COL2 DECIMAL(2)
```

If the data in COL2 is 43, the solution pads the value with three leading zeros to form 00043, prior to comparing that entry to the data in COL1, which is five characters long. This automatic padding may affect how the solution retrieves related rows of data.

Character-to-numeric conversion will occur only if the character column contains all digits.

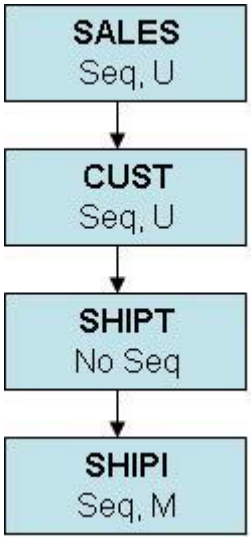
Comparing Different Length Columns

Corresponding CHAR and VARCHAR columns do not require the same length attribute. However, if the length of one column exceeds the length of its corresponding column, the Optim solution truncates any trailing blanks until the column is equal in length to the corresponding column. If the data in the longer column does not have enough trailing blanks to truncate, that data is not considered related and is not retrieved. Truncation also occurs when the length of a concatenated expression exceeds the length of its corresponding column.

Appendix A. IMS Database Support

The Optim Compare facility supports comparing IMS segments. This section describes some features of Optim Compare that are included to support IMS databases. Note that some of the features described, such as the Soft Match options, can also be applied to other data sources such as DB2, VSAM, and sequential files.

The following sample IMS database is used throughout the remainder of this section to illustrate the Compare support for IMS databases.



As input to the Compare process, you can specify the name of a previously created Extract file or one or more table names. If you specify table names, the Optim solution extracts the data prior to compare processing. Extract processing now includes the following additional information to allow for proper matching and hierarchal relationship processing of IMS segments:

- Optim-generated unique suffix values for all non-uniquely keyed and unkeyed segments
- Flags indicating the uniqueness of the IMS segments and whether Optim's extract processing allowed it to accurately generate the necessary unique suffixes

All Optim solution extract files produced for IMS segments (through either the Move or Compare interfaces) always include this additional information; no further action is required by the user.

RESTRICTION: The Optim solution does not support comparison of IMS databases from extract files created prior to this enhancement (as they do not contain the additional information described in this section).

IMS - Unkeyed and non-uniquely keyed segment support

Optim Extract processing is updated to produce and store a unique suffix, in addition to the IMS concatenated key, for all unkeyed and non-uniquely keyed IMS segments. This unique suffix is required by Optim Compare processing as show in this section.

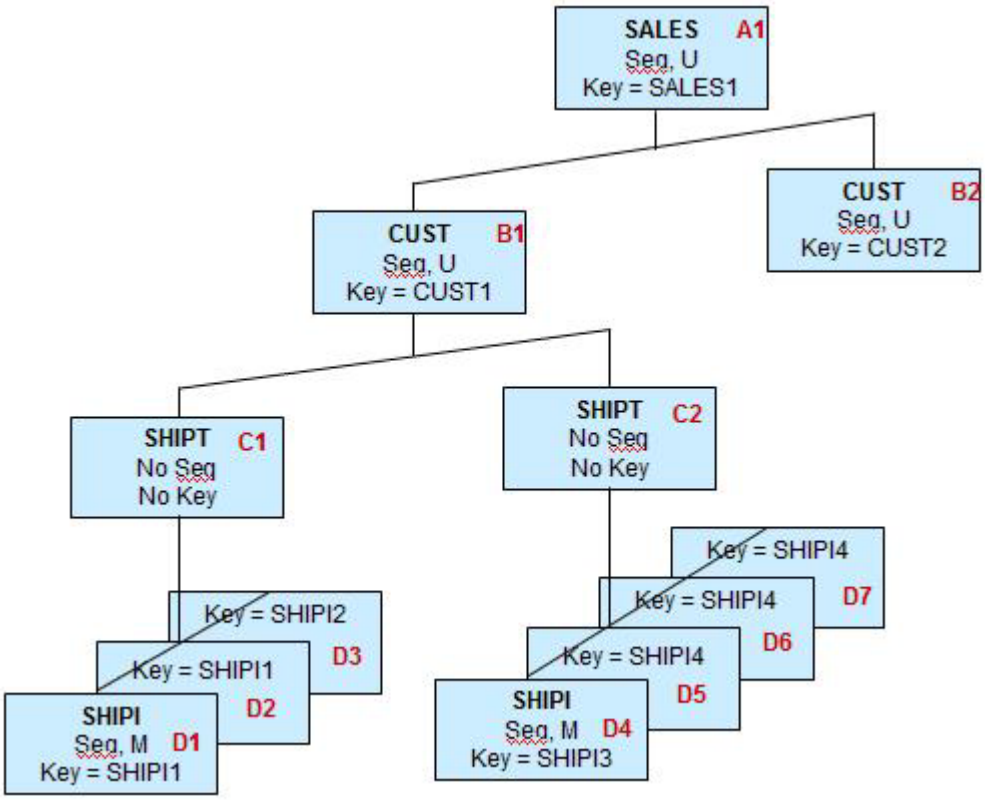
Optim Compare processing requires a unique suffix of IMS segments with non-unique and unkeyed parents to:

- relate the child segments to their unkeyed and non-uniquely keyed parents (and show related changes)
- match the correct child segments of matched parent segments

The unique suffix is a 4-byte field with an initial value of 0000. It is incremented for each unkeyed and non-uniquely keyed segment encountered within a twin-chain while reading the IMS database in physical hierarchical order. The generated unique suffix digits use the following order of precedence (from low to high):

```
0123456789ABCDEFGHIJKLMNPQRSTUVWXYZ
```

The unique suffix is reset to 0000 for the each unique occurrence within a twin chain. For example, consider the following sample IMS database:



Shown here are the SHIPT and SHIPI segments, followed by their generated unique suffix values:

C1 0000
C2 0001
D1 0000
D2 0001
D3 0000
D4 0000
D5 0000
D6 0001
D7 0002

IMS - FOP_SURKEY

For each segment, Optim processes the IMS concatenated key along with any unique suffix values produced for the segment or segment's parents as an entity called the FOP_SURKEY (surrogate key). Using the same sample database provided in the preceding section, the associated FOP_SURKEY for each segment is as follows:

Segment A1

Key SALES1
Unique Suffix n/a
FOP_SURKEY SALES1

Segment B1

Key CUST1
Unique Suffix n/a
FOP_SURKEY SALES1CUST1

Segment B2

Key CUST2
Unique Suffix n/a
FOP_SURKEY SALES1CUST2

Segment C1

Key n/a
Unique Suffix 0000
FOP_SURKEY SALES1CUST10000

Segment C2

Key n/a
Unique Suffix 0001
FOP_SURKEY SALES1CUST10001

Segment D1

Key SHIPI1
Unique Suffix 0000
FOP_SURKEY SALES1CUST10000SHIPI10000

Segment D2

Key SHIPI1
Unique Suffix 0001
FOP_SURKEY SALES1CUST10000SHIPI10001

Segment D3

Key SHIPI2
Unique Suffix 0000
FOP_SURKEY SALES1CUST10000SHIPI20000

Segment D4

Key SHIPI3
Unique Suffix 0000
FOP_SURKEY SALES1CUST10001SHIPI30000

Segment D5

Key SHIPI4

Unique Suffix 0000
FOP_SURKEY SALES1CUST10001SHIPI40000

Segment D6

Key SHIPI4
Unique Suffix 0001
FOP_SURKEY SALES1CUST10001SHIPI40001

Segment D7

Key SHIPI4
Unique Suffix 0002
FOP_SURKEY SALES1CUST10001SHIPI40002

This is a further breakdown of FOP_SURKEY for sample segment D7:

SALES1
Unique key of the parent SALES segment

CUST1
Unique key of the parent CUST segment

0001 Unique suffix generated for the unkeyed parent SHIPT segment

SHIPI4
Non-unique key of this SHIPI segment

0002 Unique suffix generated for this SHIPI segment

IMS - FOP_SURKEY Relationship Processing

When two IMS tables are related through a physical IMS relationship and participate in a Compare process, Optim uses FOP_SURKEY to form the relationship. The full FOP_SURKEY of the parent table is related to the parent's portion of FOP_SURKEY in the child table. This processing occurs automatically when an IMS physical relationship is used during Compare processing without any intervention from the user.

The FOP_SURKEY contents displays in the produced Compare results as FOP_SURKEY(*x:y*), where:

x is the start position of FOP_SURKEY that was used to relate this table to its parent table according to an IMS physical relationship (always 1)

y is the length of FOP_SURKEY that was used to relate this table to its parent table according to an IMS physical relationship

For example:

```
------------------- Optim: Browse (Source 1 Names Shown) ---------------------
Command ===>                                                 Scroll ===> PAGE

Cmd Chg Src == Table: SAMPLE_SHIPT(T1) ==================== 1 OF 2 ============
            FOP_SURKEY(1:10) CUST_ID  SHIP_ID   ADDRESS_LEN
            ----CH(15)------ -CH(5)- -SMALLINT- -SMALLINT--
*** ********************************* TOP ***********************************
___      12 SALES1CUST10000   00412        610          20
___      12 SALES1CUST10001   00416        545          24
*** ******************************** BOTTOM *********************************
```

In this example, FOP_SURKEY(1:10) indicates that characters 1 through 10 of the FOP_SURKEY displayed for the SHIPT segment were used to relate this table to its parent table CUST.

IMS - FOP_SURKEY match key processing

The default Compare match key for an IMS segment is the segment's sequence field. When two non-root IMS tables are compared, FOP_SURKEY(*x*:*y*) is automatically prepended to the match key. Only the parent's portion of the surrogate key, as specified by start position (*x*) and length (*y*) of FOP_SURKEY, is used as the match key. Other columns can be added to the match key and the sequence field can be removed from the match key. However, you cannot remove FOP_SURKEY from the match key as it is required for the Optim solution to properly compare child segments of related/matched parents.

As previously described, segments with unkeyed and non-uniquely keyed parents have generated unique suffix values imbedded within the FOP_SURKEY column. These suffix values are included in the match key.

If the segment itself is unkeyed or non-uniquely keyed, a generated unique suffix value is included as the last four bytes of FOP_SURKEY. This value is not included in the match key and it does not participate in the segment data comparisons.

The Optim solution displays FOP_SURKEY as a match key column, such as:

```
Chg  Src FOP_SURKEY(1:10) CUST_ID   SHIP_ID   ADDRESS_LEN
---- --- ++++++++++++++++ --------- --------- ------------
     12  SALES1CUST10000  00412           610           20
     12  SALES1CUST10001  00416           545           24
```

In this example of the SHIPT segment, FOP_SURKEY (start position 1, length 10) is displayed as the match key, as indicated by the plus signs that underscore the heading.

IMS - Soft Match

The Compare facility also allows users to control how the Optim solution matches unkeyed and non-uniquely keyed tables. Users are able to specify the key property along with matching properties that can be applied to any data source (VSAM/SEQ, DB2, and IMS) through the Match Options interface.

Soft match invocation

Soft match is active by default for all unkeyed and non-uniquely keyed tables. For a single compare, the match options can be defined on the Specify COMPARE Sources panel by specifying '**Y**' in the **Define Compare Match Options** field, as shown in the following example:

```
-- Specify COMPARE Sources: SAMPLE.COMPARE.USECASE ---------------------------
Command ===>

Table 1 (Use LIKE Syntax for Lists):
  Creator ID              ===>                    >>
  Table Name              ===>                                        >>
  If Data from Extract File:
    Extract File DSN      ===>
  If Data from Table:
    Modify Sel. Criteria ===> N  (S-SEL, Q-SQL, N-NO, A-ALL)   Status: NONE

Table 2 (Use LIKE Syntax for Lists):
  Creator ID              ===>                    >>
  Table Name              ===>                                        >>
  If Data from Extract File:
    Extract File DSN      ===>
  If Data from Table:
    Modify Sel. Criteria ===> N  (S-SEL, Q-SQL, N-NO, A-ALL)   Status: NONE

Column Map (Optional) - Use MAP command to edit existing Column Map:
  MapID.MapName ===>                       (Use LOCAL for Internal/Unnamed Map)
Define New Match Key            ===> N   (Y-Yes, N-No)
Define Compare Match Options    ===> Y   (Y-Yes, N-No)
```

For a multiple compare, the match options can be defined for each pair of tables by typing **MOPT** (Match Options) on the command line, placing the cursor on the source 2 table name, and then pressing **Enter**, as shown in the following example:

```
------------------------- COMPARE Process Table Map ----------------------------
Command ===> MOPT                                             Scroll ===> PAGE

Commands: APPLY,SAVE,LIST,MAP,POPULATE,ACM,CLEAR,MKEY,MOPT  END when Done
Source 2 Tables must Exist in Corresponding Extract File
 Src 1 CID: SAMPLE                                     Column
 Src 2 CID: SAMPLE                                  >> Map ID ===>

 Source 1 Table Name      Source 2 Table Name      Type   Column Map or LOCAL
--------------------- -------------------------->> ------- ---------------------
************************************ TOP ***************************************
SALEHDAM_SALES        SALEHDAM_SALES             EXISTS
SALEHDAM_CUST         SALEHDAM_CUST              EXISTS
SALEHDAM_SHIPT        SALEHDAM_SHIPT             EXISTS
SALEHDAM_SHIPI        SALEHDAM_SHIPI             EXISTS
********************************** BOTTOM **************************************
```

Each table pair for which you specify **MOPT** will be flagged with a - in the **Type** column indicating that **MOPT** is pending, as shown in the following example:

```
------------------------- COMPARE Process Table Map ------------- MOPT PENDING
Command ===>                                                  Scroll ===> PAGE

Commands: APPLY,SAVE,LIST,MAP,POPULATE,ACM,CLEAR,MKEY,MOPT  END when Done
Source 2 Tables must Exist in Corresponding Extract File
 Src 1 CID: SAMPLE                                     Column
 Src 2 CID: SAMPLE                                  >> Map ID ===>

 Source 1 Table Name      Source 2 Table Name      Type   Column Map or LOCAL
--------------------- -------------------------->> ------- ---------------------
************************************ TOP ***************************************
SALEHDAM_SALES        SALEHDAM_SALES             EXISTS
SALEHDAM_CUST         SALEHDAM_CUST              EXISTS
SALEHDAM_SHIPT        SALEHDAM_SHIPT             EXISTS-
SALEHDAM_SHIPI        SALEHDAM_SHIPI             EXISTS-
********************************** BOTTOM **************************************
```

You can also specify **MKEY** (Match Key) for a table pair to define a match key. If both **MKEY** and **MOPT** have been specified for the same table pair, it is flagged with a plus sign (+) to indicate both **MKEY** and **MOPT** are pending. If only **MKEY** has been specified, the table pair is flagged with an asterisk (*) to indicate only **MKEY** is pending.

Pressing **End (F3)** will then take you to the **MKEY** and **MOPT** dialogs for each table pair that has been flagged.

Soft Match Options

The Compare Match Options panel allows you to specify options that control the method used to match rows, the accuracy with which non-uniquely keyed tables are matched, and limitations to the processing that will be performed.

Here is an example of the Compare Match Options panel.

```
--Compare Match Options-------------------------------------------------------
Command ===>                                              SCROLL ===> PAGE

Commands: MKEY, CANCEL    END when Done

Source 1: SAMPLE.SALEHDAM_SHIPT
Source 2: SAMPLE.SALEHDAM_SHIPT

Match Key Property   ===> Unknown     (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method         ===> ANY

Soft Match Parameters (if Non-Keyed or Non-Unique)
  Enabled            ===> Y       (Y-Yes, N-No)
  Percent Required   ===> 50      Percentage of columns that must match
  Percent Acceptable ===> 100     Processing stops at this percentage match
  Maximum Attempts   ===> 0       Limits number of comparisons
```

The following commands are available on the Compare Match Options panel:

- **MKEY** - Displays a dialog to redefine the match key.
- **CANCEL** - Discards the changes you made to the compare match options.
- **END** - Accepts the changes you made to the compare match options and exits the panel.

The Compare Match Options panel includes the following fields:

Match Key Property

Describes the key used to match records in the source 1 and source 2 tables. The default value reflects the known attributes of the key (if one exists). This parameter can be set to one of the following values:

Non-Keyed

A Match Key has not been defined between the source 1 and source 2 tables. All rows will be matched using the specified **Match Method**.

Non-Unique

A non-unique Match Key has been defined between the source 1 and source 2 tables. Rows will be matched using the specified **Match Method**.

Unique

A unique Match Key has been defined between the source 1 and source 2 tables. Rows will be matched using only the match key. If duplicate keys are found, they will be reported as duplicates and not compared.

Unknown

The match key property is not known and is to be determined by the compare process

when it opens both sources. This is the default value when either of the sources is not an extract file. You may specify match method and soft match parameters but they will only be used if the match key property that is determined is not unique.

Note: If you change the **Match Key Property** value to **Non-keyed**, any match key previously defined for this table pair is deleted. If you then want to change it to another value, you will first need to use the **MKEY** function to define the match key.

Match Method

Controls the processing for non-unique and non-keyed tables. There currently is only one choice, **ANY**. With this method, all of the rows in source 1 with the same key are compared with all of the rows with the same key in source 2. Equal rows are matched without any consideration for their relative sequence. Unequal rows are either considered unmatched or are matched on a best-fit basis. This is controlled by the **Soft Match Parameters**.

Soft Match Parameters

Enabled

Controls whether soft matching will be performed on the unequal rows. Specify:

Y Unequal rows will be compared and the best matches will be paired and considered "changed."

N All of the unequal rows will be considered to be source 1 or source 2 only and the other parameters in this section will be ignored. Soft matching does not apply when the **Match Key Property** is **Unique**.

Percent Required

Specifies degree of similarity required to consider two unequal rows matched. This value applies to the non-match key columns. Specify a percentage in the range from 1 through 100. The default is 50 percent.

For example, if there are 10 non-match key columns and you specify 80 percent, 8 columns must be equal for the rows to be paired. (If there are columns that "must match," add them to the match key.)

Percent Acceptable

This parameter limits processing. Normally, all of the rows in source 1 are compared to all of the rows in source 2 and the best matches that meet the **Percent Required** are selected. If a value less than 100 is specified for **Percent Acceptable**, match processing for a row will stop when a row comparison meets or exceeds this percentage. The effect is to reduce processing cost at the expense of accuracy. The default is 100 percent.

Maximum Attempts

This value is a governor that limits the number of compares that will be performed for any set of unequal rows. When this value is reached, the results of the comparisons are evaluated for matching. Any rows that were not compared will be considered source 1 or source 2 only rows. The effect is to reduce processing cost at the expense of accuracy. The value 0 disables this limit. The default is 0.

Soft Match Examples

This section provides examples of Compare results when various Match Options are used. Sample data of the non-uniquely keyed SHIPI segment is used throughout the examples.

Here are examples of Source 1.

```
FOP_SURKEY(1:15)           SHIPI_KEY CLASS TYPE    DATE_UPDATED OWNER
+++++++++++++++++++++++++ +++++++++ ----- ------- ------------ -----
SALES1CUST10001SHIPI30000 SHIPI3    1st   LETTER  12.04.27     BOB
SALES1CUST10001SHIPI40000 SHIPI4    2nd   PACKAGE 12.03.02     JOHN
SALES1CUST10001SHIPI40001 SHIPI4    1st   PACKAGE 09.04.27     PHIL
SALES1CUST10001SHIPI40002 SHIPI4    1st   PACKAGE 08.04.27     MIKE
```

Here are examples of Source 2.

```
FOP_SURKEY(1:15)           SHIPI_KEY CLASS TYPE    DATE_UPDATED OWNER
+++++++++++++++++++++++++ +++++++++ ----- ------- ------------ -----
SALES1CUST10001SHIPI30000 SHIPI3    1st   PACKAGE 12.04.27     PETER
SALES1CUST10001SHIPI40000 SHIPI4    1st   PACKAGE 08.04.27     FRANK
SALES1CUST10001SHIPI40001 SHIPI4    2nd   PACKAGE 12.03.02     JOHN
SALES1CUST10001SHIPI40002 SHIPI4    1st   LETTER  11.04.29     BOB
```

The Match Key used for all samples that follow include FOP_SURKEY(1:15) and SHIPI_KEY.

Soft match sample 1 (Soft Match Disabled)

Using the Match Options in the following example:

```
Match Key Property    ===> Non-Unique  (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method          ===> ANY

Soft Match Parameters (if Non-Keyed or Non-Unique)
  Enabled             ===> N       (Y-Yes, N-No)
  Percent Required    ===>         Percentage of columns that must match
  Percent Acceptable  ===>         Processing stops at this percentage match
  Maximum Attempts    ===>         Limits number of comparisons
```

Compare would produce the results shown in the following example:

```
Cmd Chg Src == Table: SALEHDAM_SHIPI (T1) ==================================
               FOP_SURKEY      SHIPI_KEY   CLASS     TYPE    DATE_UPDATED   OWNER
             -----CH(15):N------ -CH(6):N- -CH(3):N- -CH(7):N- --CH(8):N--- -CH(5):N-
*** ********************************************************************************
___  D   1  SALES1CUST100010000  SHIPI3      1st     LETTER     12.04.27     BOB
___  D    2 SALES1CUST100010000  SHIPI3      1st     PACKAGE    12.04.27     PETER
___       1  SALES1CUST100010001  SHIPI4      1st     PACKAGE    09.04.27     PHIL
          1  SALES1CUST100010002  SHIPI4      1st     PACKAGE    08.04.27     MIKE
___      12  SALES1CUST100010000  SHIPI4      2nd     PACKAGE    12.03.02     JOHN
___       2  SALES1CUST100010000  SHIPI4      1st     PACKAGE    08.04.27     FRANK
          2  SALES1CUST100010002  SHIPI4      1st     LETTER     11.04.29     BOB
*** ********************************************************************************
```

When Soft Match is disabled and the Match Key Property is defined as Non-Unique:

- The Optim solution attempts to find and match any identical unique keys found in source 1 and source 2. Therefore, the rows with identical, unique keys SALES1CUST10001 and SHIPI3 are reported as matched, but changed.

- The Optim solution then will attempt to match all rows with non-unique keys. Only non-unique rows with 100% identical contents will be reported as matched (Src 12). Therefore, the rows with SALES1CUST10001, SHIPI4, 2nd, and so on, are matched. All other rows will be reported as existing only in source 1 or source 2.

Soft match sample 2 (Soft Match Enabled)

Using the Match Options in the following example:

```
Match Key Property    ===> Non-Unique  (Non-Keyed, Non-Unique, Unique, Unknown)

Match Method          ===> ANY

Soft Match Parameters (if Non-Keyed or Non-Unique)
  Enabled             ===> Y       (Y-Yes, N-No)
  Percent Required    ===> 75      Percentage of columns that must match
  Percent Acceptable  ===> 100     Processing stops at this percentage match
  Maximum Attempts    ===> 0       Limits number of comparisons
```

Compare would produce the results shown in the following example:

```
Cmd Chg Src == Table: SALEHDAM SHIPI (T1) ==================================

              POP SURKEY      SHIPI KEY   CLASS     TYPE     DATE UPDATED   OWNER
*** ************************************************************************
___       1  SALES1CUST100010000  SHIPI3      1st     LETTER     12.04.27     BOB
___       2  SALES1CUST100010000  SHIPI3      1st     PACKAGE    12.04.27     PETER
___      12  SALES1CUST100010000  SHIPI4      2nd     PACKAGE    12.03.02     JOHN
    D     1  SALES1CUST100010002  SHIPI4      1st     PACKAGE    08.04.27     MIKE
___ D     2  SALES1CUST100010000  SHIPI4      1st     PACKAGE    08.04.27     FRANK
___       1  SALES1CUST100010001  SHIPI4      1st     PACKAGE    09.04.27     PHIL
___       2  SALES1CUST100010002  SHIPI4      1st     LETTER     11.04.29     BOB
*** ************************************************************************
```

When **Soft Match** is enabled and **Percent Required** is specified as **75**:

- The Optim solution attempts to match all rows with identical keys that have at least 75% of identical column data. In this scenario, two matches were found, one with identical contents and one with different OWNER contents. All other rows are reported as existing only in source 1 or source 2.

Appendix B. Comparing hierarchically related data from VSAM or Seq file

One possible use case of Optim Compare may be to compare data that has been unloaded from an IMS database into a file. By defining Optim Legacy Tables, Primary Keys, and Relationships that map the unloaded data, the Optim solution is able to process (extract, compare) the sequential file of unloaded IMS data. However, when the unloaded IMS data contains one or more unkeyed or non-uniquely keyed segments, additional processing of the data must be performed before it can be used properly with the Optim Compare facility, as follows.

FOP_UNIQUE_SUFFIX

When comparing related tables, the Optim solution implements a hierarchical parent correlation process. In short, the Optim solution must compare only the related children of the related parents. This is similar to an IMS paradigm. The Optim solution applies this to all hierarchical relationships.

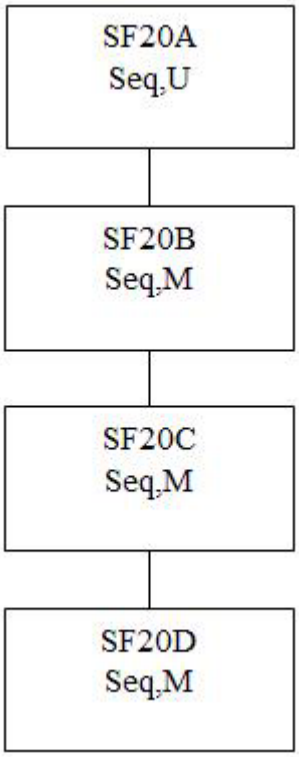
An additional "surrogate" key is required for a multiple table comparison when a segment with unkeyed or non-uniquely keyed records has dependant records. This surrogate key provides:

- Uniqueness in the concatenated key for the table's parents
- A unique key to resolve parent-child relationships

The surrogate key is comprised of the defined key and a uniqueness suffix. For native IMS database support, the solution automatically generates and uses the uniqueness suffix value when reading the data without any intervention from the user. However, for sequential files, such as an unloaded IMS database, the uniqueness suffix must be provided by the user in a special column called FOP_UNIQUE_SUFFIX. The column defined as FOP_UNIQUE_SUFFIX in a Legacy Table definition:

- is automatically excluded from comparison
- is expected to be inserted appropriately into the match key of its dependents
- is expected to be defined as part of the relationship keys

As an example, consider the following scenario where an IMS database (4 segment types, 3 of which are keyed multiple) is unloaded from IMS into a sequential file.



In this example, the user must add a unique suffix value to segments SF20B and SF20C in the sequential file because:

- The segment contains unkeyed or non-uniquely keyed records
- The segment has dependant segments

The provided suffix ideally is concatenated to the existing key in the records. The suffix should uniquely identify the segment within the twin chain. In this example, a 2-byte CHAR value has been appended to the concatenated key for segments SF20B and SF20C.

Corresponding definitions for these sample segments are shown in the next topic.

Sample Legacy Table Definitions

The previous topic shows examples of data comparisons, and the results obtained when different options for soft matching are used. Definitions for the Legacy tables, primary keys and relationships for that data is included in this topic.

```
---------------------- Modify Legacy Table: RDLC.SF20A ------------------------
Command ===>

User defined I/O Exit   ===>                                        Row 1 of 8
Cmd Level/Field Name              Type  Len  Occur Column Name
--- ----------------------------- --- ----- ----- ----------------------------
*** ********************************* TOP ************************************
___  1 RECORD-FILE20A                    10        RECORD_FILE20A
___   2 CRITERIA                   CHR    4        CRITERIA
___   2 KEY-AREA                          2        KEY_AREA
___    3 KV-A                      CHR    2        KV_A
___   2 RECORD-DATA                       4        RECORD_DATA
___    3 CHAR-D                    CHR    1        CHAR_D
___    3 DV                        CHR    3        DV
*** ******************************** BOTTOM **********************************
```

```
---------------------- Modify Legacy Table: RDLC.SF20B ------------------------
Command ===>

User defined I/O Exit   ===>                                       Row 1 of 10
Cmd Level/Field Name              Type  Len  Occur Column Name
--- ----------------------------- --- ----- ----- ----------------------------
*** ********************************* TOP ************************************
___  1 RECORD-FILE20B                    16        RECORD_FILE20B
___   2 CRITERIA                   CHR    4        CRITERIA
___   2 KEY-AREA                          4        KEY_AREA
___    3 KV-A                      CHR    2        KV_A
___    3 KV-B                      CHR    2        KV_B
___   2 FOP-UNIQUE-SUFFIX          CHR    4        FOP_UNIQUE_SUFFIX
___   2 RECORD-DATA                       4        RECORD_DATA
___    3 CHAR-D                    CHR    1        CHAR_D
___    3 DV                        CHR    3        DV
*** ******************************** BOTTOM **********************************
```

```
---------------------- Modify Legacy Table: RDLC.SF20C ------------------------
Command ===>

User defined I/O Exit   ===>                                       Row 1 of 12
Cmd Level/Field Name              Type  Len  Occur Column Name
--- ----------------------------- --- ----- ----- ----------------------------
*** ********************************* TOP ************************************
___  1 RECORD-FILE20C                    22        RECORD_FILE20C
___   2 CRITERIA                   CHR    4        CRITERIA
___   2 KEY-AREA                         10        KEY_AREA
___    3 KV-A                      CHR    2        KV_A
___    3 KV-B                      CHR    2        KV_B
___    3 US-B                      CHR    4        US_B
___    3 KV-C                      CHR    2        KV_C
___   2 FOP-UNIQUE-SUFFIX          CHR    4        FOP_UNIQUE_SUFFIX
___   2 RECORD-DATA                       4        RECORD_DATA
___    3 CHAR-D                    CHR    1        CHAR_D
___    3 DV                        CHR    3        DV
*** ******************************** BOTTOM **********************************
```

```
---------------------- Modify Legacy Table: RDLC.SF20D -----------------------
Command ===>

User defined I/O Exit   ===>                                      Row 1 of 13
Cmd Level/Field Name              Type  Len  Occur Column Name
--- ------------------------------ --- ----- ----- ------------------------------
*** ********************************* TOP ************************************
___   1 RECORD-FILE20D                  24        RECORD_FILE20D
___    2 CRITERIA                  CHR   4        CRITERIA
___    2 KEY-AREA                        16        KEY_AREA
___     3 KV-A                     CHR   2        KV_A
___     3 KV-B                     CHR   2        KV_B
___     3 US-B                     CHR   4        US_B
___     3 KV-C                     CHR   2        KV_C
___     3 US-C                     CHR   4        US_C
___     3 KV-D                     CHR   2        KV_D
___    2 RECORD-DATA                     4        RECORD_DATA
___     3 CHAR-D                   CHR   1        CHAR_D
___     3 DV                       CHR   3        DV
*** ******************************** BOTTOM **********************************
```

In segments SF20B and SF20C, the FOP_UNIQUE_SUFFIX column identifies the unique suffix that was added to the segments. Also note that FOP_UNIQUE_SUFFIX column defined in that parent segments (such as SF20B) is defined as part of the concatenated key in the child segments (such as US-B in segments SF20C and SF20D).

As previously mentioned, any column defined as FOP_UNIQUE_SUFFIX in the Legacy Table will not participate in the actual comparison. In addition, by default any column named FILLER or any group column will not be compared. For group columns, the individual elements will be compared, unless explicitly excluded in the column map.

```
------------------------- Define OPTIM Primary K
Command ===>

Define OPTIM  Primary Key for RDLC.SF20A
Limit Key to 64 Columns and 2000 Bytes

Cmd          Column Name                  Data Type
--- -------------------------------  ----------
*** ************************ TOP ****************
___ KV_A                                 CHAR(2)
*** ********************** BOTTOM ***************
```

```
------------------------- Define OPTIM Primary Ke
Command ===>

Define OPTIM  Primary Key for RDLC.SF20B
Limit Key to 64 Columns and 2000 Bytes

Cmd          Column Name                  Data Type
--- -------------------------------  ----------
*** ************************ TOP ****************
___ KV_A                                 CHAR(2)
___ KV_B                                 CHAR(2)
*** ********************** BOTTOM ***************
```

```
------------------------- Define OPTIM Primary Ke
Command ===>

Define OPTIM  Primary Key for RDLC.SF20C
Limit Key to 64 Columns and 2000 Bytes

Cmd          Column Name                  Data Type
--- -------------------------------  ----------
*** ************************ TOP ****************
___ KV_A                                 CHAR(2)
___ KV_B                                 CHAR(2)
___ US_B                                 CHAR(4)
___ KV_C                                 CHAR(2)
*** ********************** BOTTOM ***************
```

```
------------------------- Define OPTIM Primary Key
Command ===>

Define OPTIM  Primary Key for RDLC.SF20D
Limit Key to 64 Columns and 2000 Bytes

Cmd          Column Name                  Data Type
--- -------------------------------  ----------
*** ************************ TOP ****************
___ KV_A                                 CHAR(2)
___ KV_B                                 CHAR(2)
___ US_B                                 CHAR(4)
___ KV_C                                 CHAR(2)
___ US_C                                 CHAR(4)
___ KV_D                                 CHAR(2)
*** ********************** BOTTOM ***************
```

The FOP_UNIQUE_SUFFIX column is not defined in the Primary Key of the segment it occurs in (parent segment), such as SF20B. However, it is defined in the Primary Key of the child segments, such as US_B in SF20C and SF20D.

Here are the Sample Relationship Definitions:

```
---------------------------- Define Relationship ----------------------------
Command ===>                                                 Scroll ===> CS

                  Define OPTIM Relationship A_TO_B
        Special Commands: LIST COLUMNS, EXPAND, GENERIC, MODEL

    Parent: RDLC.SF20A                  Child: RDLC.SF20B
                                                                  1 OF 2
Cmd      Column Name        Data Type        Column Name        Data Type
--- ---------------------->> ----------  ---------------------->> ----------
*** ********************************* TOP ***********************************
___ KV_A                    CH(2)       KV_A                    CH(2)
*** ******************************** BOTTOM **********************************
```

```
---------------------------- Define Relationship ----------------------------
Command ===>                                                 Scroll ===> CSR

                  Define OPTIM Relationship B_TO_C
        Special Commands: LIST COLUMNS, EXPAND, GENERIC, MODEL

    Parent: RDLC.SF20B                  Child: RDLC.SF20C
                                                                  1 OF 4
Cmd      Column Name        Data Type        Column Name        Data Type
--- ---------------------->> ----------  ---------------------->> ----------
*** ********************************* TOP ***********************************
___ KV_A                    CH(2)       KV_A                    CH(2)
___ KV_B                    CH(2)       KV_B                    CH(2)
___ FOP_UNIQUE_SUFFIX       CH(4)       US_B                    CH(4)
*** ******************************** BOTTOM **********************************
```

```
---------------------------- Define Relationship ----------------------------
Command ===>                                                 Scroll ===> CS

                  Define OPTIM Relationship C_TO_D
        Special Commands: LIST COLUMNS, EXPAND, GENERIC, MODEL

    Parent: RDLC.SF20C                  Child: RDLC.SF20D
                                                                  1 OF 6
Cmd      Column Name        Data Type        Column Name        Data Type
--- ---------------------->> ----------  ---------------------->> ----------
*** ********************************* TOP ***********************************
___ KV_A                    CH(2)       KV_A                    CH(2)
___ KV_B                    CH(2)       KV_B                    CH(2)
___ US_B                    CH(4)       US_B                    CH(4)
___ KV_C                    CH(2)       KV_C                    CH(2)
___ FOP_UNIQUE_SUFFIX       CH(4)       US_C                    CH(4)
*** ******************************** BOTTOM **********************************
```

In the relationship definitions for these segments, FOP_UNIQUE_SUFFIX in the parent segment is mapped to the corresponding concatenated key value in the child segment (such as FOP_UNIQUE_SUFFIX in SF20B to US_B in SF20C).

FOP_UNIQUE_SUFFIX in the child segment is part of the concatenated key, and therefore must be mapped between the child segment and its child segment (such as US_B in SF20C to US_B in SF20D).

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
Software Interoperability Coordinator
Director of Engineering, Information Management
111 Campus Drive
Princeton, NJ 08540
USA

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this information and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement, or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml.

Java™ is a trademark of Sun Microsystems, Inc.

Linux is a registered trademark of Linus Torvalds.

UNIX is a registered trademark of The Open Group.

Microsoft and Windows are registered trademarks of Microsoft Corporation.

Other company, product, or service names may be trademarks or service marks of others.

Index

Printed in USA