DB2® Warehouse Manager for OS/390® and z/OS™ White Paper

By Sarah Ellis and Cathy Drummond

October 2001

DB2 Warehouse Manager for OS/390 and z/OS

This paper begins with an overview of DB2 Warehouse Manager for OS/390 and z/OS, and its components. It continues with information on how to use the product, and some practical advice.

It contains the following sections:

- What is DB2 Warehouse Manager for OS/390?
- What are the components of DB2 Warehouse Manager for OS/390?
- What are warehouse agents?
- How can you use the OS/390 agent?
- What are warehouse transformers?
- What skill set is required?
- What authorities are required?
- How to run multiple agent daemons
- How to run DB2 for OS/390 utilities
- How to unload data from DB2 for OS/390
- How to load the data
- Job scheduling tips
- Recommendations for creating tables
- Managing resources on the OS/390 and z/OS platforms
- How to determine and diagnose problems
- Data Warehouse Center return codes
- Where to find more information

What is DB2 Warehouse Manager for OS/390?

DB2 Warehouse Manager is the most effective way to build, manage, govern, and access DB2 data warehouses. You can use it to simplify and speed your warehouse prototyping, development, and deployment to production. With DB2 Warehouse Manager, you can pull source data from DB2 (on any platform), from relational databases such as Oracle or Sybase, and from VSAM and IMS^{TM} files that interface with tools such as DataJoiner®, Data Propagator^{TM}, and Classic Connect.

For OS/390 and z/OS platforms, DB2 Warehouse Manager is available as a feature of DB2 Universal Database for OS/390 and z/OS Version 7 (5675-DB2).

What are the components of the DB2 Warehouse Manager for OS/390?

The Warehouse Manager for OS/390 consists of the following components:

- DB2 warehouse server

  The warehouse server, the hub of the DB2 warehousing solution, controls the interactions between all of the DB2 warehousing components. The server polls the warehouse control database and sends messages to each defined agent site with instructions regarding the tasks to be carried out. The server runs on Windows NT® and Windows 2000® only.

- Data Warehouse Center

  The Data Warehouse Center is the user interface for warehouse administration. Warehouse administrators can use the Data Warehouse Center to specify the source and target databases, perform data extraction and transformation, and run population steps and create associated schedules. The Data Warehouse Center runs on DB2 for Windows NT, Windows 2000, Windows 98®, Windows 95®, AIX®, and Solaris Operating Environment platforms.

- Warehouse agent

  The warehouse agent for OS/390 runs z/OS and OS/390-based processes on behalf of the Data Warehouse Center so that data can be processed in your OS/390 environment without the need to export the data to an intermediate platform. With the warehouse agent, you can take full advantage of the power, security, reliability, and availability of DB2. The warehouse agent runs on OS/390 and z/OS UNIX® Systems Services, requires OS/390 and z/OS Version 2 Release 7 or later, and is compatible with DB2 for OS/390 Versions 5 and 6, and DB2 for OS/390 and z/OS Version 7.

  See "What are warehouse agents?" for more details about the OS/390 agent.

  Warehouse agents are also available for the following platforms: Windows NT, Windows 2000, AS/400®, OS/2®, AIX®, and Solaris Operating Environment.

- Warehouse transformers

  Warehouse transformers are stored procedures that provide commonly used data transformations for building data warehouses.

- The Information Catalog Manager

  Previously provided with Visual Warehouse under the name DataGuide®, the

Information Catalog Manager helps users find, understand, and access enterprise data.

- The QMF™ family of products

  These multipurpose query and reporting tools help you easily access and present business information. They allow data sharing across the enterprise, and provide server resource protection, robust application development, and native connectivity to all of the DB2 workstation platforms.

- OLAP Starter Kit

  The DB2 OLAP Starter Kit lets users of DB2 Universal Database explore and apply online analytical capabilities before acquiring the full product, DB2 OLAP Server, for more complex analysis or deployment to large groups of users.

The warehouse server, Data Warehouse Center and OLAP Starter Kit are included with DB2 Universal Database™ Version 7 Enterprise Edition. This edition of DB2 has a restricted use license to permit delivery and utilization of the Windows-based components of the DB2 Warehouse Manager for OS/390.

A Sourcing Agent for OS/390 and z/OS (5655-F36) is available as an optional feature of the DB2 Warehouse Manager for the UNIX, Windows, and OS/2® platforms. This agent enables easier access to source data managed by DB2 for OS/390 for customers who already have a license for DB2 Warehouse Manager on a different platform. This OS/390 Sourcing Agent feature is identical to the OS/390 agent component of the DB2 Warehouse Manager.

What are warehouse agents?

Warehouse agents run processes on various platforms on behalf of the Data Warehouse Center. Each agent enables data to be fetched from any DB2 database and inserted into the warehouse. The agents also enable the scheduling of DB2 utilities, OS/390 job streams, and FTP jobs. The agent consists of an agent daemon and the agent processes that are started by the daemon at the request of the warehouse server.

DB2 Warehouse Manager for OS/390 and z/OS includes an OS/390 (zSeries™) warehouse agent. You can use this agent to communicate between DB2 Universal Database for OS/390 and z/OS and other databases, including DB2 databases on other platforms and non-DB2 databases. The DB2 for OS/390 agent runs its processes in UNIX Systems Services and can communicate with supported data sources that use an ODBC connection. The OS/390 agent also uses the cross loader functionality, new in DB2 Version 7 with APARs PQ45268 & PQ46759, to load from remote platforms directly into DB2 for OS/390.

The warehouse agent gives warehouse users the following advantages:

- The ability to move data between DB2 for OS/390 systems, avoiding triangular moves

- The ability to perform local inserts where OS/390 is the target, optimizing performance

- The ability to transform data while it is still on the OS/390 system, before moving it to other systems

- Access to IMS and VSAM from the OS/390 agent, via Classic Connect, without leaving the platform

If you are already running DB2 Warehouse Manager for UNIX, Windows, or OS/2, you
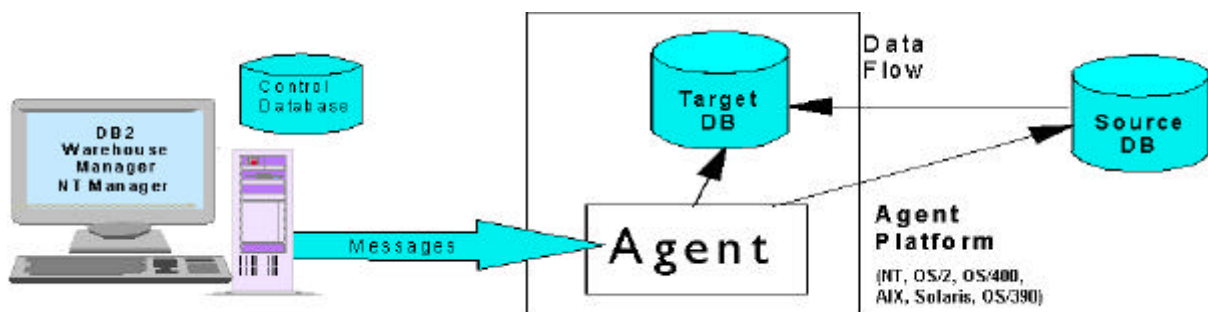
may benefit from additional agents in your warehouse system if data must be moved from multiple sources and targets. For example, you can add the OS/390 agent in order to be able to run the transformers on OS/390, particularly if you are aggregating data before moving it to another platform.

How can you use the OS/390 agent?

The OS/390 agent can improve performance in situations where the source data and target data are on OS/390. Using the OS/390 agent avoids the need to move the data to other platforms when copying from OS/390 to OS/390.

You can build warehouses from production data, produce data marts from warehouses, and build OLAP cubes using the OS/390 agent.

You will still need to make decisions regarding how to move the data to meet your currency and availability needs. However, after these decisions are made, you can use the DB2 Warehouse Manager to integrate and schedule your processes.

What are warehouse transformers?

Warehouse transformers are Java^{TM} stored procedures, provided with the product, to assist with commonly used transformations. These transformers provide complex transformations used in warehouse developments, such as data manipulation, data cleaning, key generation, and statistical calculations. Transformers add to the number of ways that you can transform data by using SQL built-in functions such as string manipulation, Boolean operations, data type conversions, and utility invocations.

What skill set is required?

The Warehouse Manager for OS/390 user should have DB2 for OS/390 skills. The DB2 systems programmer who installs the agent will require OS/390, UNIX Systems Services, Java, and Workload Manager (WLM)-managed stored procedure skills, as well as OS/390 data base administrator skills.

What authorities are required?

All users need to be valid UNIX Systems Services users and have a home directory.

The permission bits for the executable programs in the /usr/lpp/DWC directory need to be set to at least 755 so that all users can read and execute warehouse tasks. The

agent includes these permissions but the permissions can be reset by various actions, for example, by using FTP to move the agent to a different location.

All warehouse users require:
- Execute authority on CLI plan
- Write access to logging and trace directories, if agent logging or CLI trace is active
- SELECT authority on the source tables involved
- Execute authority on DSNUTILs and associated utilities, if applicable
- Authority to run the methods by which the data is moved, such as Data Propagator
- Authority to run JCL job streams and the appropriate authorities on the target tables

How to run multiple agent daemons

Multiple agent daemons might be required in testing environments or in situations where there are multiple warehouses or data marts on different DB2 subsystems. Each agent daemon will have its own dsnaoini file pointing to different DB2 subsystems or CLI plans.

For each agent daemon that you want to run, you need to update the etc/services file on OS/390 and use the vwd Port Name environment variable to define the port to use. For example, if the Port Name is `vwd2`, and the Port Number is `11014`, follow these steps to run multiple daemons:

1. Add a new entry in the etc/services file on OS/390:

   ```
   vwd2 11014/tcp
   ```

2. In the .profile file used to start the daemon, add:

   ```
   export vwdPortName=vwd2
   ```

3. On the workstation, change the vwd entry in c:\winnt\system32\drivers\etc\services to match the new port number:

   ```
   vwd 11014/tcp
   ```

How to run DB2 for OS/390 utilities

You can start DB2 for OS/390 utilities from the DB2 Warehouse Manager. The parameters can be entered through Data Warehouse Center windows and are run using the DB2-supplied stored procedure DSNUTILS.

You might want to run utilities by submitting OS/390 JCL streams. This enables you to run existing JCL, view the output within SDSF, and manage utility execution with job schedulers.

How to unload data from DB2 for OS/390

For DB2 for OS/390 Version 7, use the new UNLOAD utility to extract data from production tables. This utility provides additional field selection and filtering options and is a more natural approach than using the REORG UNLOAD utility.

Use the DB2 for OS/390 Utility step to run an UNLOAD. The following parameters are a minimum that should be specified:

```
UID
RESTART
RECDSN
WORKDSN1
WORKDEVT1
WORKSPACE1
WORKDSN2
WORKDEVT2
WORKSPACE2
UTSTMT
UTILITY
```

For DB2 for OS/390 Version 5 or Version 6, use the REORG UNLOAD utility to unload DB2 data.

Both methods produce load cards (SYSPUNCH), which can be used as input to the load utility.

How to load the data

Run a DB2 load utility to load the data. For complex load statements, use the DB2 for OS/390 Utility step with the :FILE option. Because the DB2 for OS/390 Load step requires field positioning to be entered, you cannot specify the SYSPUNCH data set that is produced from the unload procedure. In Version 7.1 the DB2 for OS/390 Utility step is the same as the REORG step. In Version 7.2 there is a separate DB2 for OS/390 Utility step.

The SYSPUNCH that is generated by an unload procedure will contain statements to load the table that was unloaded. The DB2 Warehouse Manager eliminates the need to manually edit the data set to specify the warehouse table name. It also allows you to specify the file containing the SYSPUNCH data set in the utility.

For example, you can specify the UTSTMT for a LOAD utility as:

```
:FILE:syspunch filename:TABLE:tablename to be loaded
```

This statement loads the table that is specified using the SYSPUNCH generated from an unload of a different table.

When you run the DSNUTILS utility to load data from a non-OS/390 platform, you might need to FTP one or both of the two files generated by the unload. The SYSPUNCH load cards must be sent using FTP to the site where the agent is running. The RECDSN parameter points to the data set that contains the load data, which needs to be on the server where the load is to run.

More detailed information on using this process to copy data from one table to another is in the README file.

Edit the load cards to include NOCOPYPEND or COPY YES depending on whether a backup is required. If you specify STATISTICS, you do not need to run an extra RUNSTATS step to update the catalog statistics.

The remaining sections, "Job scheduling tips," Recommendations for creating tables," Managing resources on the OS/390 and z/oS platforms," How to determine and diagnose problems," "Data Warehouse Center return codes," and "Where to find more information," provide some practical advice for DB2 Warehouse Manager users.

Job scheduling tips

You do not need to use dynamic SQL or the DSNUTILS stored procedure to unload production data with the DB2 Warehouse Manager. You might prefer to unload production data using the standard OS/390 job scheduling system rather than the DB2 Warehouse Manager scheduler, which runs on Windows NT. Doing this will enable normal controls to be in place for accessing operational data. Users who choose this approach can define the unloaded flat files as sources within DB2 Warehouse Manager, which can then be used as input to subsequent processes.

Whether you decide to run the jobs to build your warehouse from a Windows NT server will depend on the nature of the warehousing system. Some DB2 users consider their warehouse system as production and will want to run all jobs with an OS/390 job scheduler.

Currently the DB2 Warehouse Manager cannot integrate with OPCA, CA7, or other job schedulers because the warehouse server and metadata do not both exist on OS/390.

One way to alleviate the job scheduler requirement is to use the XTClient trigger facility, which enables the DB2 Warehouse Manager steps to be submitted from OS/390.

Recommendations for creating tables

The DB2 Warehouse Manager can create the target tables as part of the processes. However it is more likely that you will choose to define the tables in advance. This allows you to predefine the required table spaces (either partitioned or segmented), create indexes, and grant authorities.

If you decide to let the DB2 Warehouse Manager create large tables, do not create them in the default database, which uses minimal sizings. In addition, if multiple tables are placed in the same segmented tablespace this could be an issue during loading, as load replace works at a tablespace rather than a table level.

If you create indexes on large tables after they are loaded, large amounts of temporary storage are required. It can be more efficient to create indexes with the DEFER NO option and then run the REBUILD INDEX utility.

Managing resources on the OS/390 and z/OS platforms

You can use the OS/390 features and tools to manage your resources efficiently:

- Set up the Workload Manager to ensure that the warehouse jobs and queries are given the appropriate level of resources so that they do not restrict production work.

- Use the Resource Limit Facility and Predictive Governor to prevent runaway queries.

- Use the SQL Performance Analyzer tool to determine how long queries will take and optionally, to prevent running queries that would exceed a user-defined time limit.

For more information about DB2 controls, see "Chapter 17, Monitoring and controlling DB2 and its connections," in the *DB2 Universal Database for OS/390 and z/OS Version 7 Administration Guide*, SC26-9931.

How to determine and diagnose problems

You can monitor the progress of warehouse steps from the Data Warehouse Center Work in Progress windows. These windows show the schedule of steps and whether the steps ran successfully, failed, or are still to be scheduled. When a job fails,

information is displayed in a pop-up window. The diagnostic information is available from the Work in Progress window by clicking on the failing job name.

- For SQL and Classic Connect, the error information displayed is the formatted SQLCA. Use this information to investigate and correct the error, then restart the step.

- For utilities, the output from DSNUTILS that was written to SYSIBM.SYSPRINT is retrieved and made available in the Work In Progress window unless the return code is 0 or 4. If the return code is 0 or 4, the utility is deemed successful and this information is not retrieved.

  If agent tracing is 4, the output will also be available in the agent log regardless of the return code.

- For Visual Warehouse programs or user-defined programs, output is written to the location defined by the VWP_LOG environment variable.

- For transformers, the SQLCA is formatted and returned in a message window in the Data Warehouse Center. If further debugging is required, a JDBC trace should be started and used with output from the Java stored procedure startup procedure.

Another source of information on errors is the daemon output on the OS/390 system. If you start your daemon in the background, you can send the output to a file. For example, to route the daemon messages to /u/userid/logs/vwd.log, enter the command:

```
vwd > /u/userid/logs/vwd.log &
```

Data Warehouse Center return codes

You can look up the return codes (RC) that are shown in error messages from the Data Warehouse Center in the *DB2 Universal Database Message Reference*, GC09-2978 and GC09-2979. The four-digit return codes are prefixed with DWC0; for example, return code 8510 is listed under DWC08510.

Where to find more information

- The README file is a key place for information. In addition to providing installation information, it also describes how to use some of the optional components such as utilities, XtClient, and FTP. It contains more detailed OS/390-specific information than can be found in the *Data Warehouse Center Administration Guide*. You can find the file at: www.ibm.com/software/data/db2/warehouse/agenthelp.pdf

- *DB2 Warehouse Manager Installation Guide*, GC26-9998-02. New to the DB2 library, this guide is available in a PDF format at:

  ftp://ftp.software.ibm.com/ps/products/db2/info/vr7/pdf/letter/db2ide72.pdf

  and in HTML, PDF, and zipped HTML formats under **Installation and Configuration** at:

  http://www.ibm.com/cgi-bin/db2www/data/db2/udb/winos2unix/support/v7pubs.d2w/en_main

- For additional installation information, including using USS and installing Java, please refer to the Warehouse Knowledge Base documents available at:

  http://www.ibm.com/software/data/db2/datawarehouse/support.html

- Information about running utilities with DSNUTILS is in Appendix B of the *DB2 Universal Database for OS/390 and z/OS Version 7 Utility Guide and Reference*, SC26-9945. Instructions for setting up the DSNUTILS stored procedure, including the WLM-managed stored procedures, are in the DB2 Management Tools redbook *DB2 for z/OS and OS/390 Version 7 Using the Utilities Suite*, SG24-6289-00.

- The online help in the Data Warehouse Center Version 7.2. provide information about how to define agents, sources, targets, and processes, as well as information about the user interface and general use of the warehouse.

About the authors

Sarah Ellis is a member of the Product Introduction Centre for IBM
Europe. She has extensive DB2 experience covering many areas, including
design, performance, and database administration. She coauthored the DB2
V6 Technical Update redbook and is currently responsible for managing the
Early Support Programmes for DB2 UDB for OS/390 within Europe. She can be
reached at: sellis@uk.ibm.com.

Cathy Drummond is a developer of the OS/390 agent for the Data Warehouse Center
at IBM Silicon Valley Laboratory. She has over twelve years of experience developing
various components of DB2 for OS/390, such as CICS® attach, IFC, and utilities. She is
a regular speaker at DB2 conferences. She can be reached at: cdrummo@us.ibm.com.