Unloading and Loading Database Content using Oracle Data Pump

Purpose

This module describes how you can unload and load data and metadata.

Topics

This module will discuss the following topics:

Move your mouse over this icon to show all screenshots. You can also move your mouse over each individual icon to see only the screenshot associated with it.

Overview

Back to List

Oracle Database 10 *g* offers a variety of methods for unloading and loading data. In this module, you will use the Oracle Data Pump Export utility to unload data then use the Oracle Data Pump Import utility to load data. You will also watch a demonstration of an alternative method of quickly loading large amounts of data with cross-platform transportable tablespaces.

What is Oracle Data Pump?

Oracle Data Pump is a new feature of Oracle Database 10 *g* that provides high speed, parallel, bulk data and metadata movement of Oracle database contents. A new public interface PL/SQL package, DBMS_DATAPUMP, provides a server-side infrastructure for fast data and metadata movement. In Oracle Database 10 *g* , new Export (expdp) and Import (impdp) clients that use this interface have been provided. The new Data Pump Export and Import tools have vastly improved performance and greatly enhanced functionality, such as restartability, flexible object selection, and better monitoring and control of export and import jobs. Because of these valuable improvements, Oracle recommends that you use these new Data Pump Export and Import clients rather than the original Export (exp) and Import (imp) clients.

Lesson Overview

MyCompany is evaluating its product portfolio to determine which products are most profitable to its bottom line. To accomplish this, MyCompany is using Data Pump Export and Import to unload and load various database objects that they will later analyze. The Sales History (SH) schema has several tables which need to be unloaded then loaded into a different schema for analysis.

Prerequisites

Back to List

Data Pump is server-based, rather than client-based. Dump files, log files, and SQL files are accessed relative to server-based directory paths, so that appropriate file security can be enforced. Data Pump requires you to specify directory paths as directory objects. A directory object maps a name to a directory name on the file system.

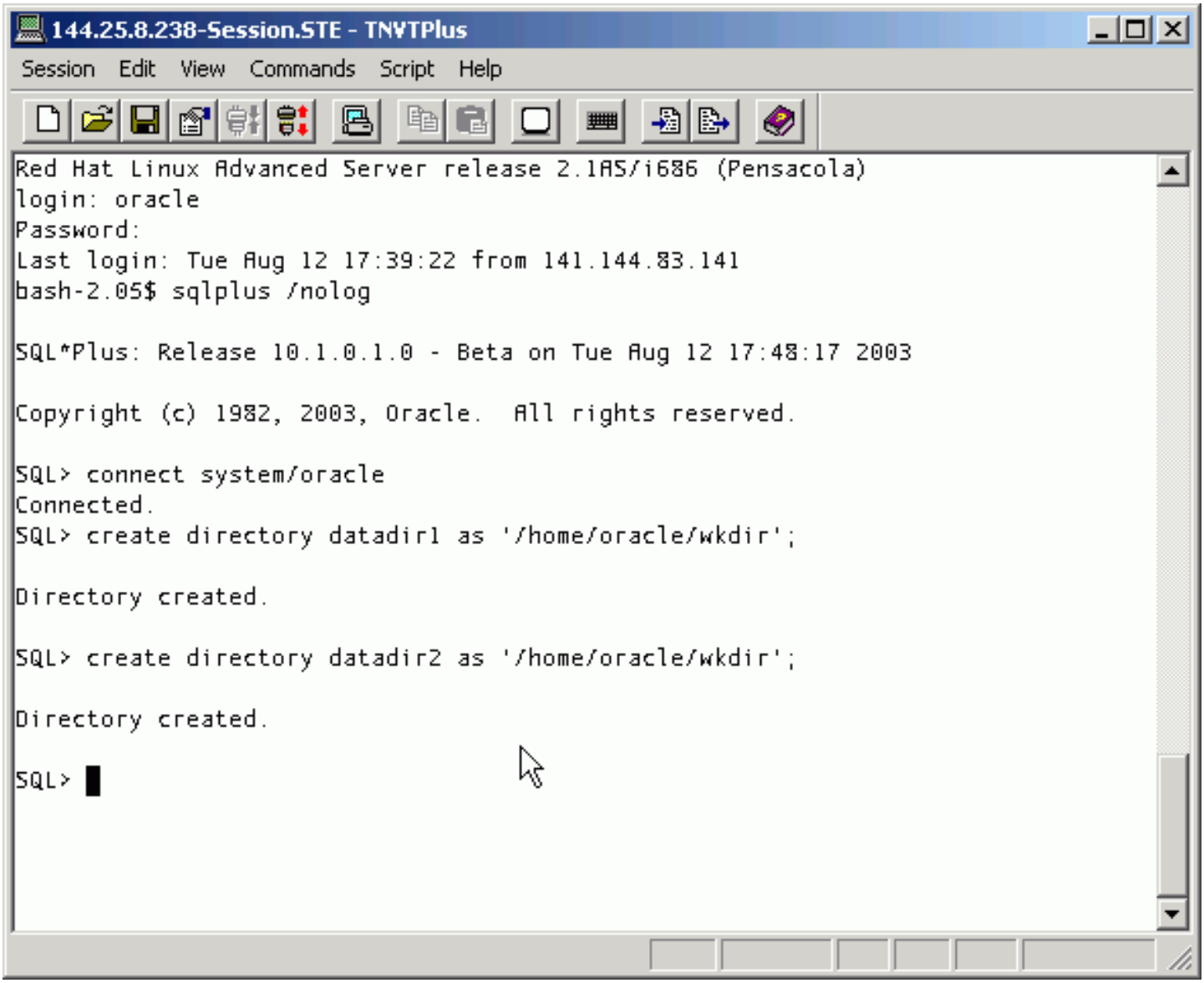
Before you can run Data Pump Export or Data Pump Import, a directory object must be created by a DBA or by any user with CREATE ANY DIRECTORY privilege. Then, when you are using Export or Import, you specify the directory object with the DIRECTORY parameter.

Creating Directory Objects

To create directory objects, perform the following:

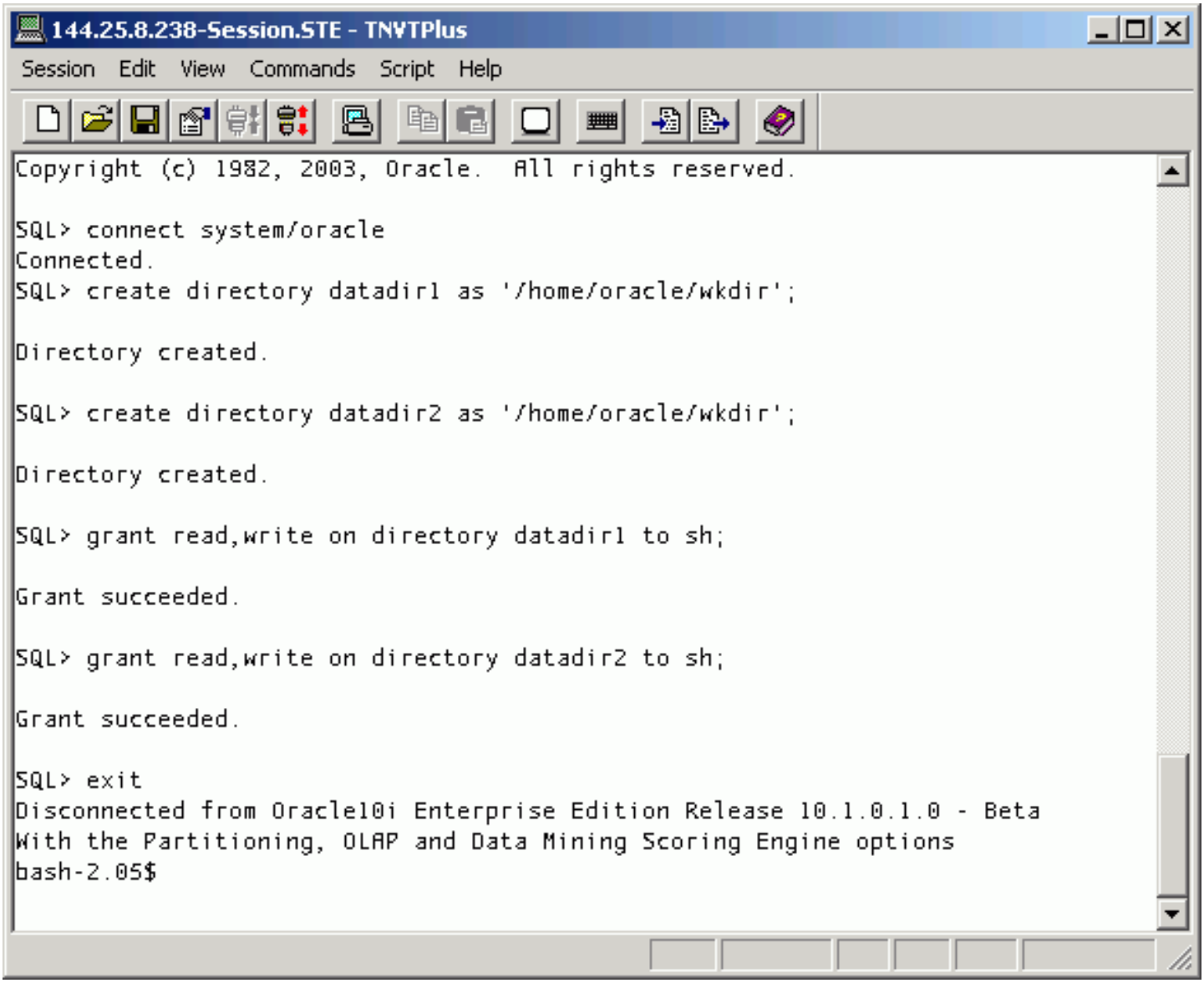
☐ **1.** Login to **SQL*Plus** as **system** and enter the following SQL commands to create two directories. Note the directory path that you specify will depend on your system. This is only an example. You will need to determine what directories on your system will be used to locate the files created by the Data Pump.

```
CREATE DIRECTORY datadir1 AS '/home/oracle/wkdir';
CREATE DIRECTORY datadir2 AS '/home/oracle/wkdir';
```



☐ **2.** After a directory is created, the user creating the directory object needs to grant READ and WRITE permission on the directory to other users. To allow the Oracle database to read and to write to files on behalf of user SH in the directories named by datadir1 and datadir2, execute the following command:

```
GRANT READ,WRITE ON DIRECTORY datadir1 TO sh;

GRANT READ,WRITE ON DIRECTORY datadir2 TO sh;
exit
```



Determining Table Dependencies

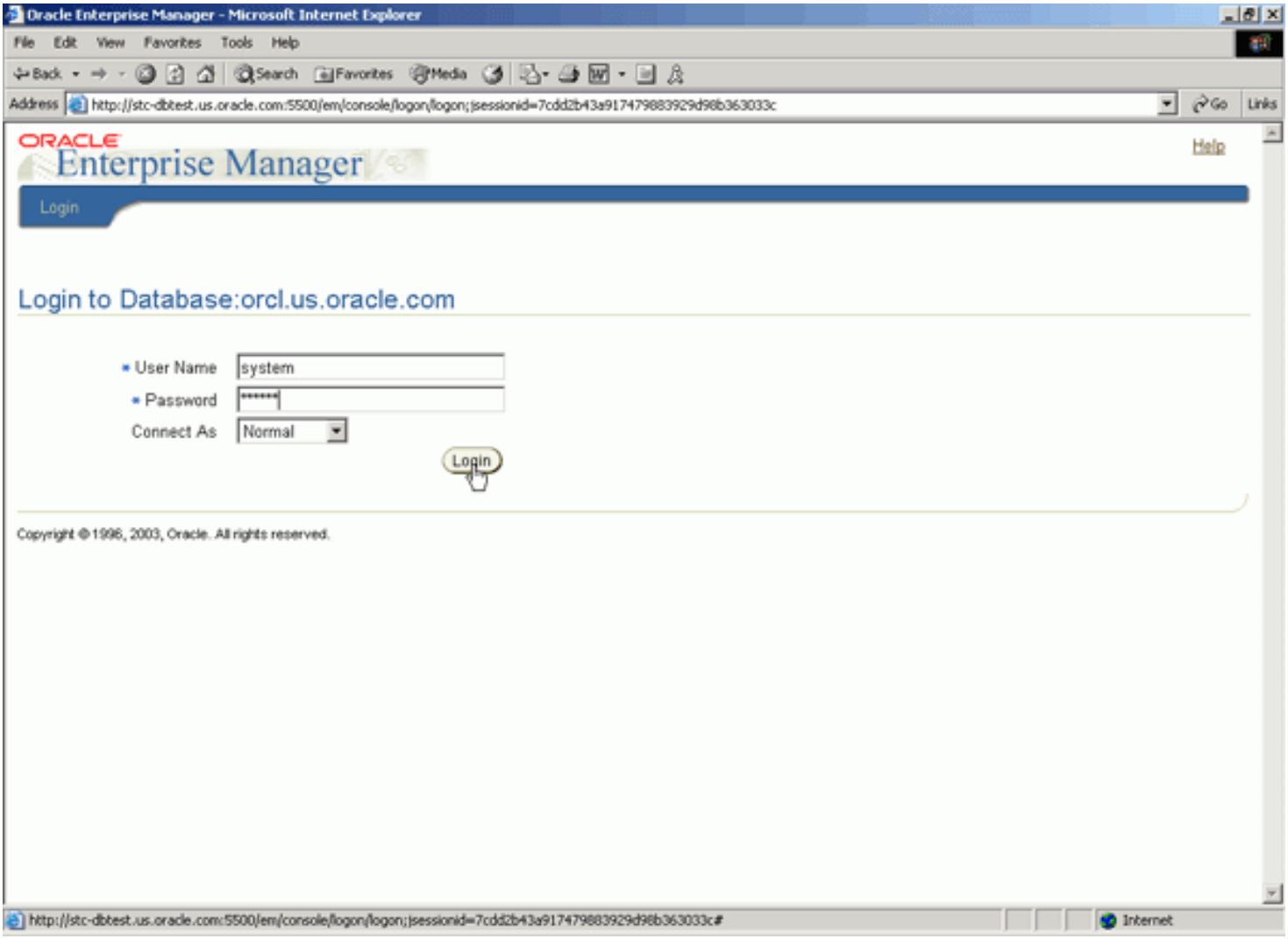
Back to List

You need to unload the SALES, PRODUCTS, and COSTS tables. To determine if these tables have any dependencies on other tables (which you may also want to unload), perform the following:

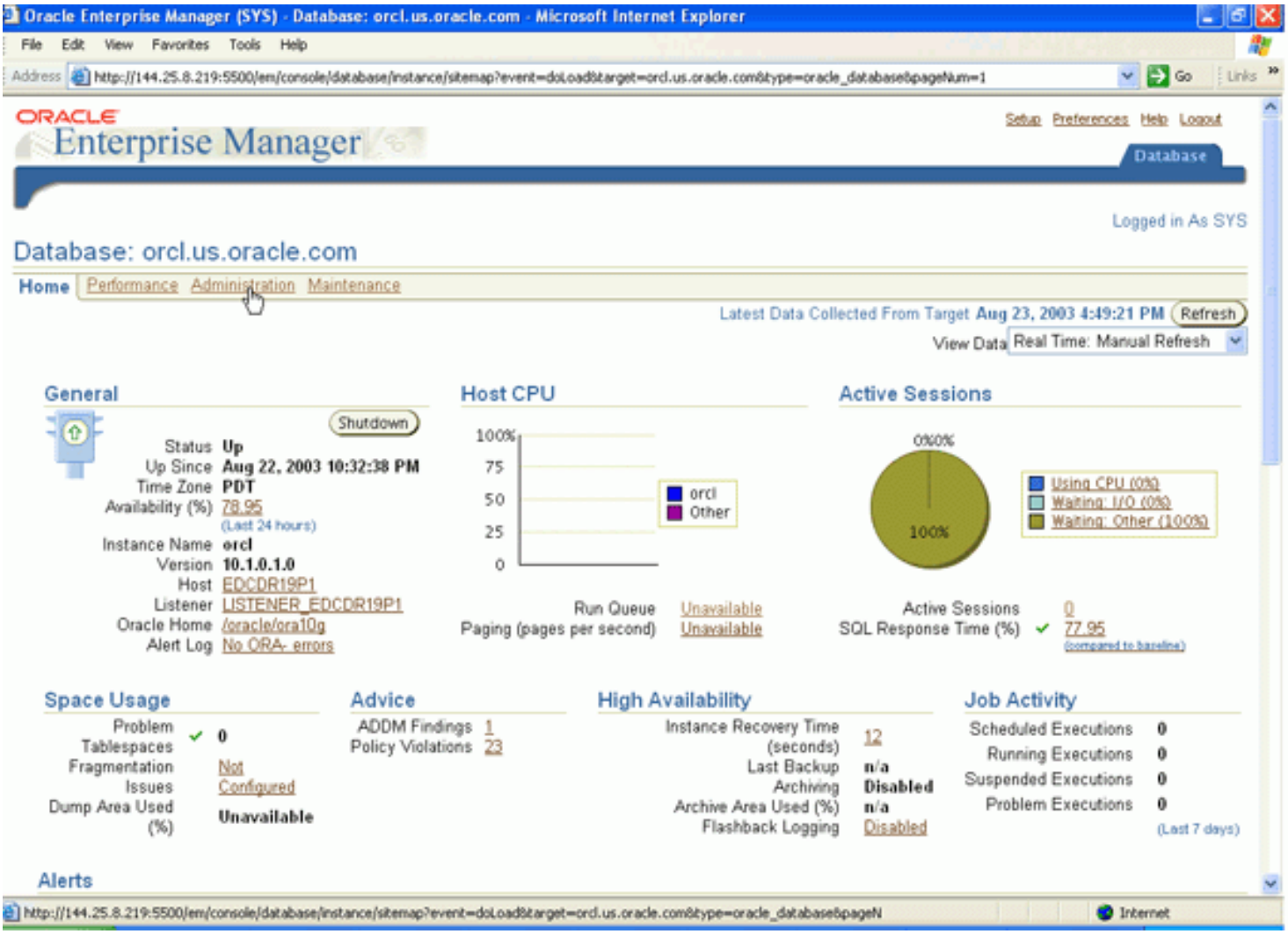
☐ **1.** Open a browser and enter the following URL:

`http://<hostname>:5500/em`

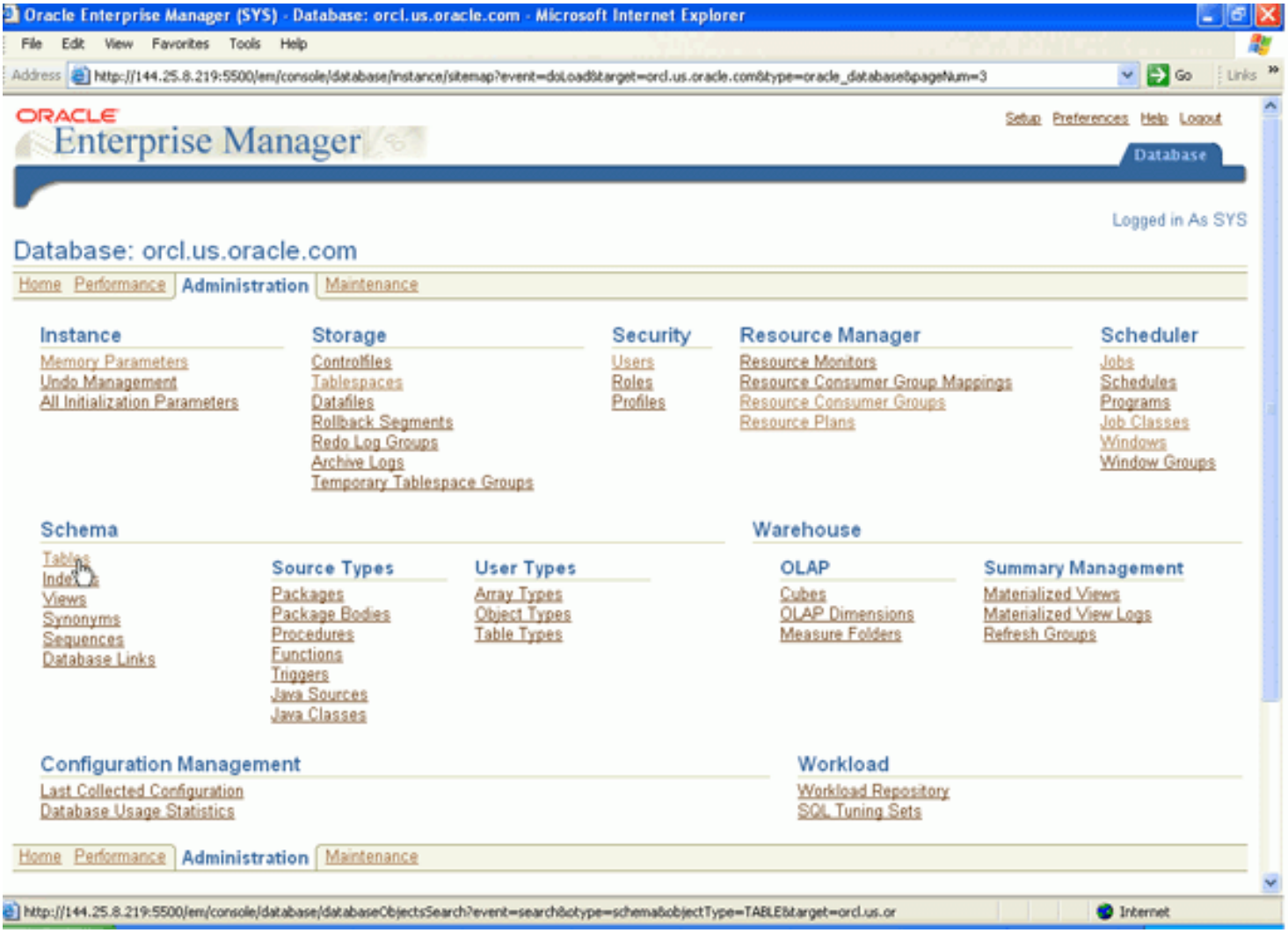
Login as **system/<password>** then click **Login** .



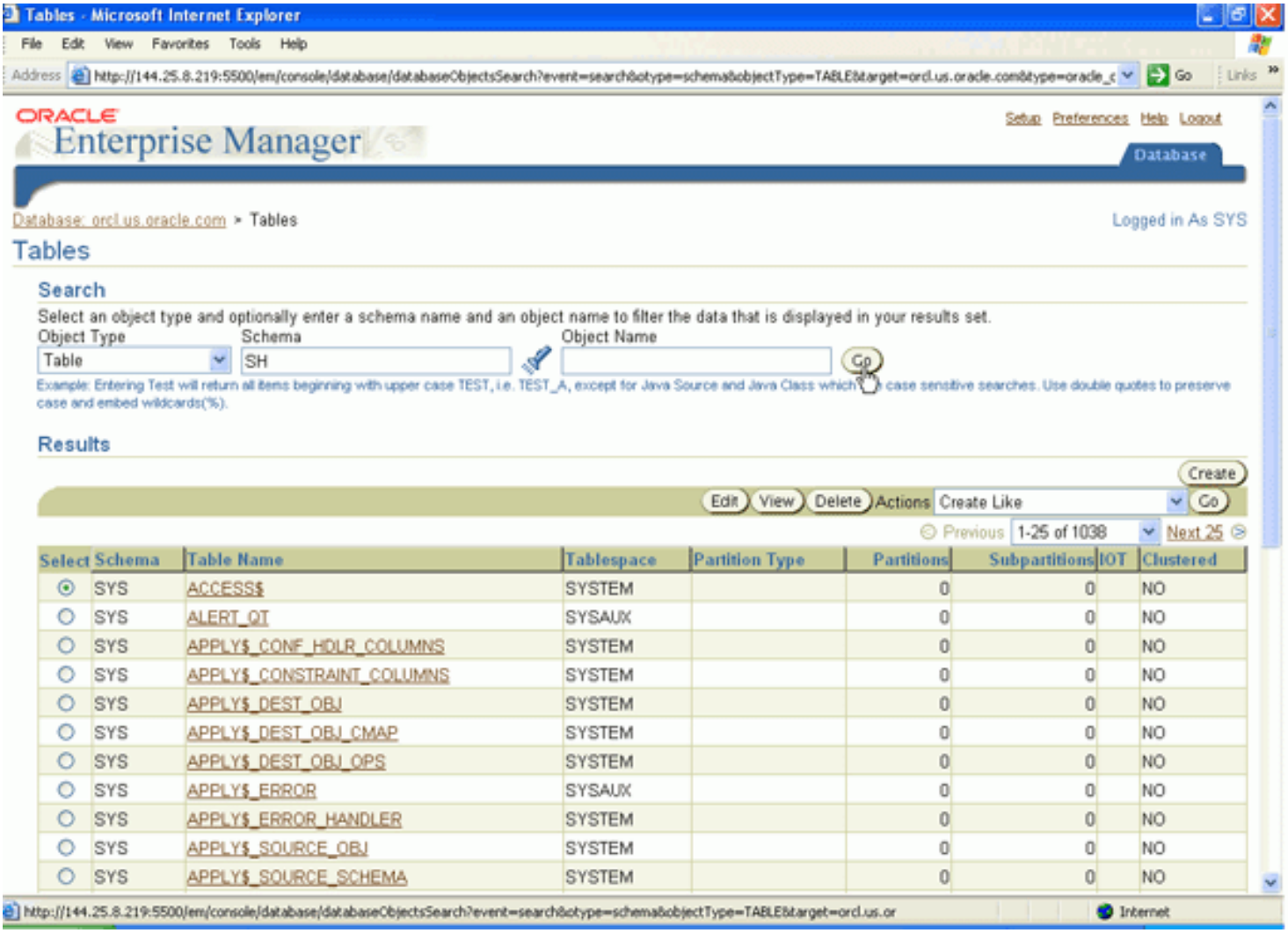
☐ **2.** Click on the **Administration** tab.



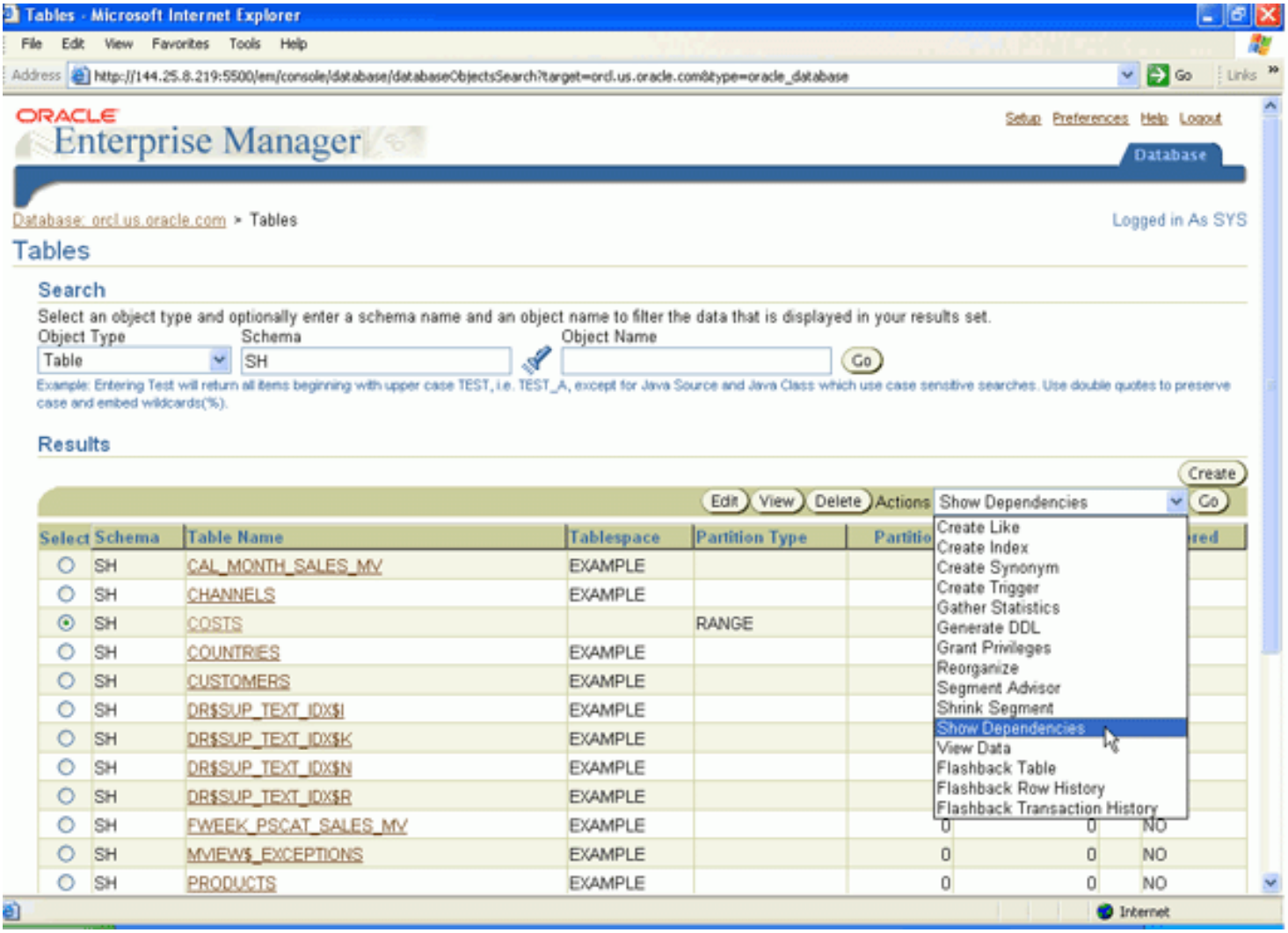
☐ **3.** Click on the **Tables** link.



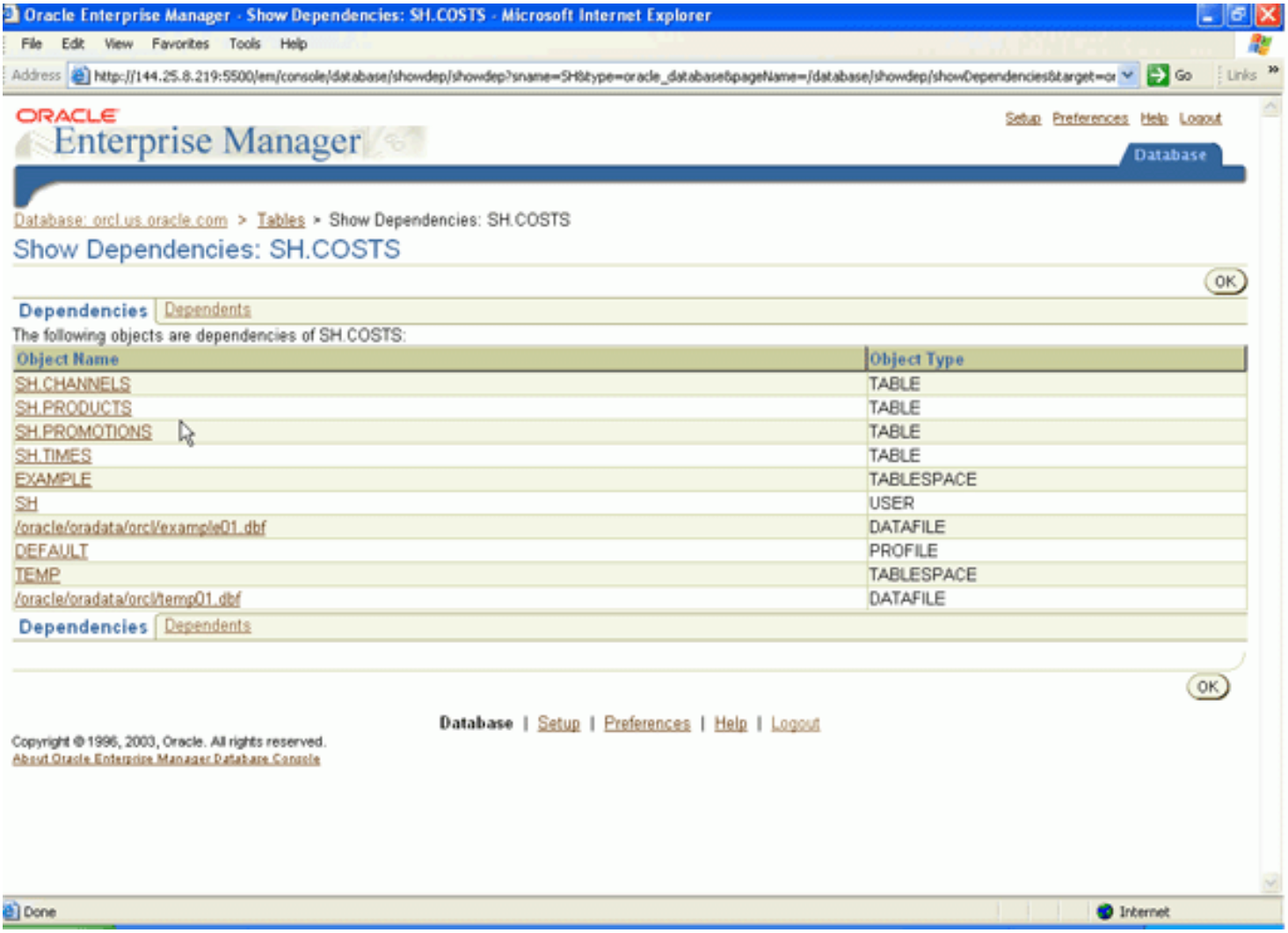
☐ **4.** Enter **SH** in the Schema field then click **Go** .



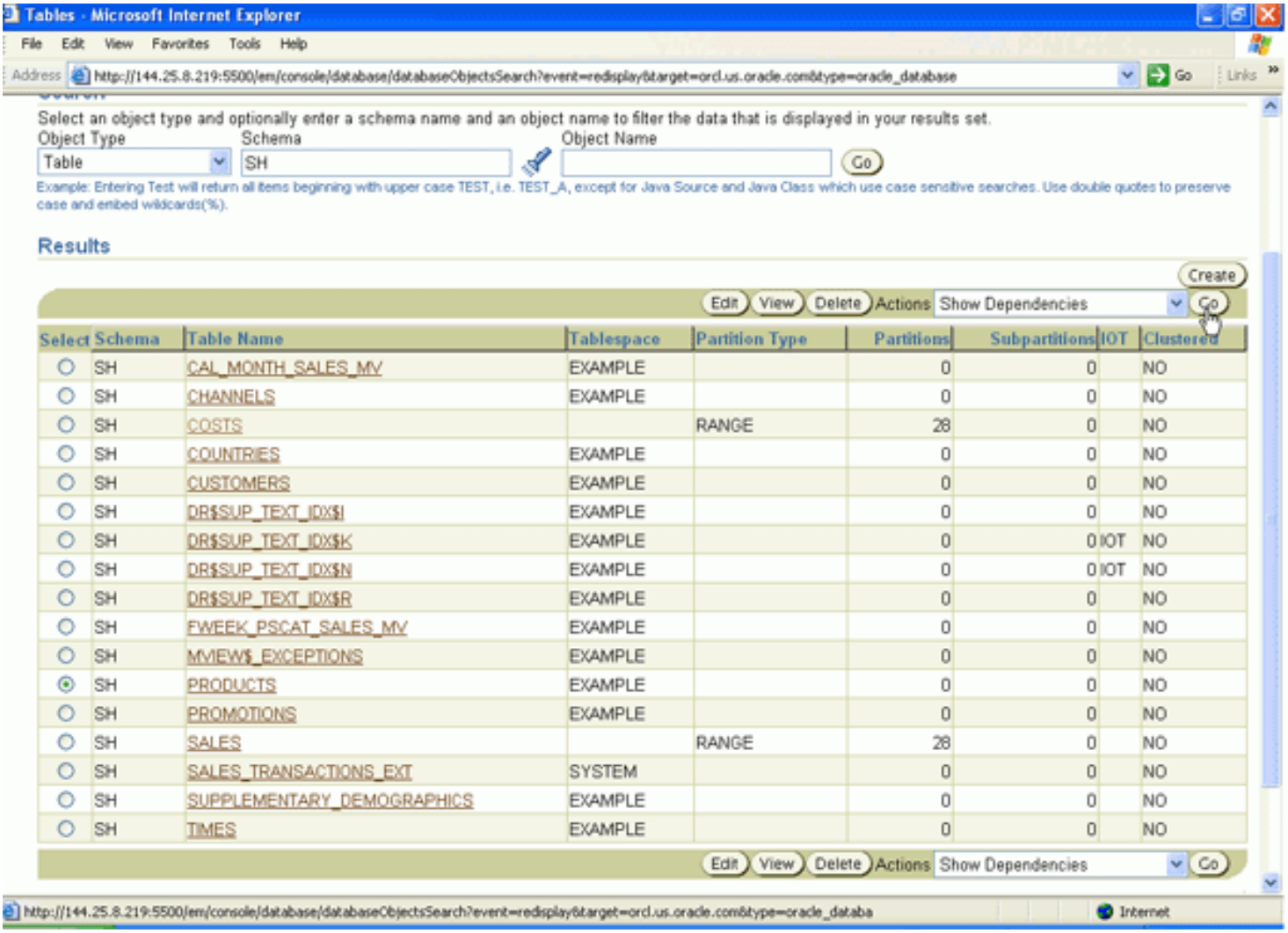
☐ **5.** Select **Costs** then select **Show Dependencies** from the drop-down list. Click **Go** .



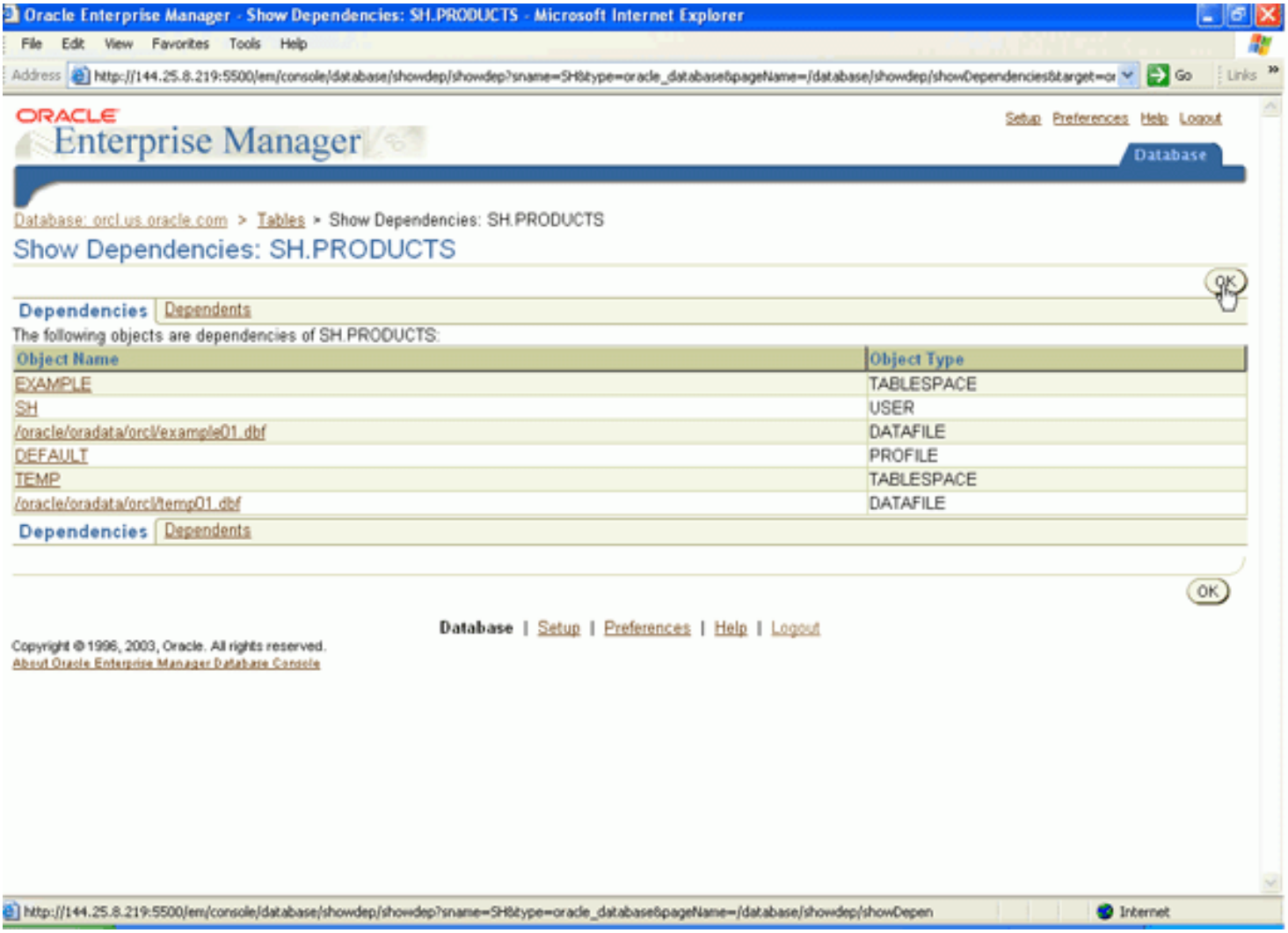
☐ **6.** Notice that the Costs table is dependent on four other tables: Channels, Products, Promotions, and Times. Other than Products, none of the other tables are needed for the analysis. Click **OK** .



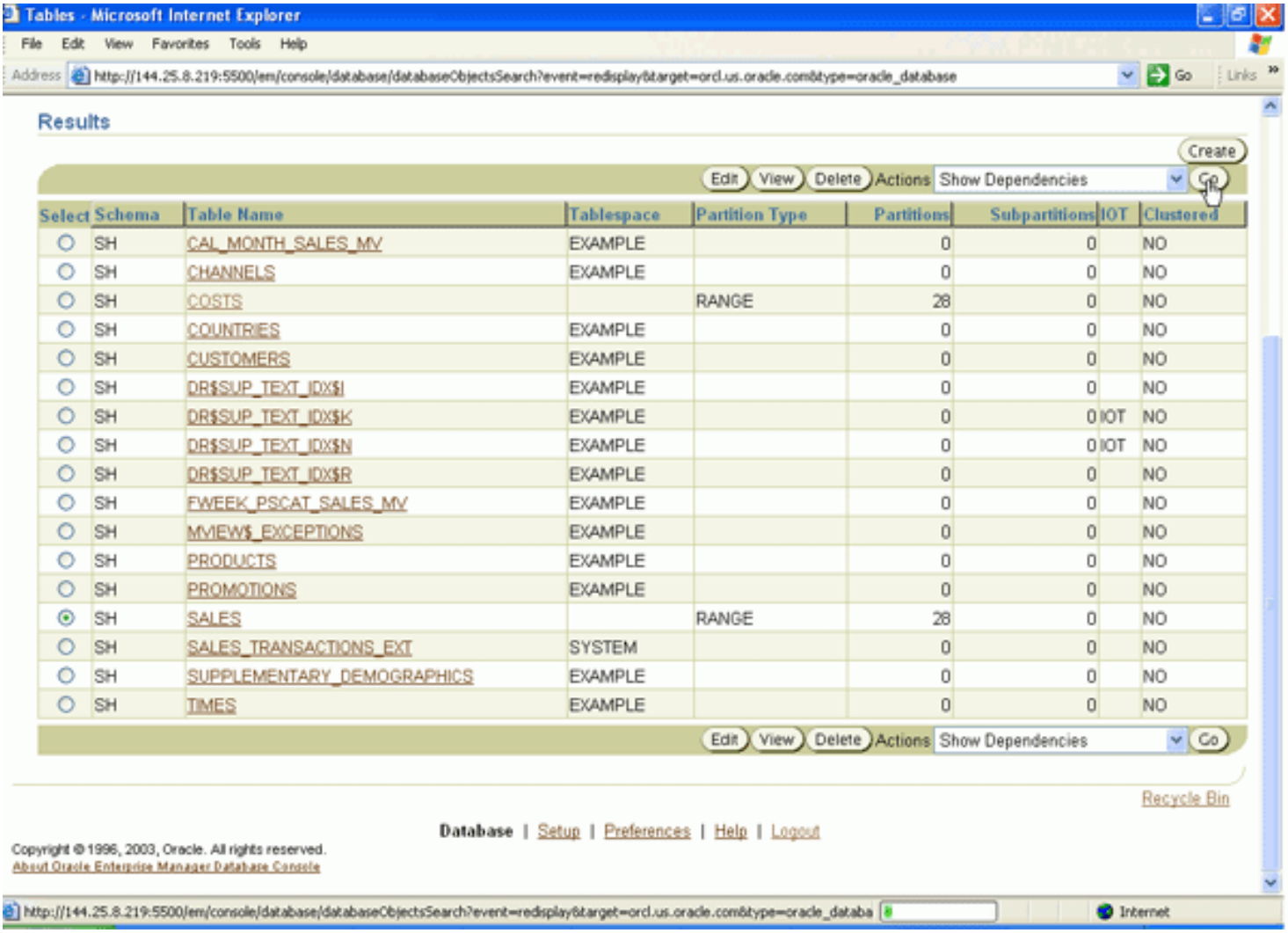
☐ **7.** Select **Products** then select **Show Dependencies** from the drop-down list. Click **Go** .



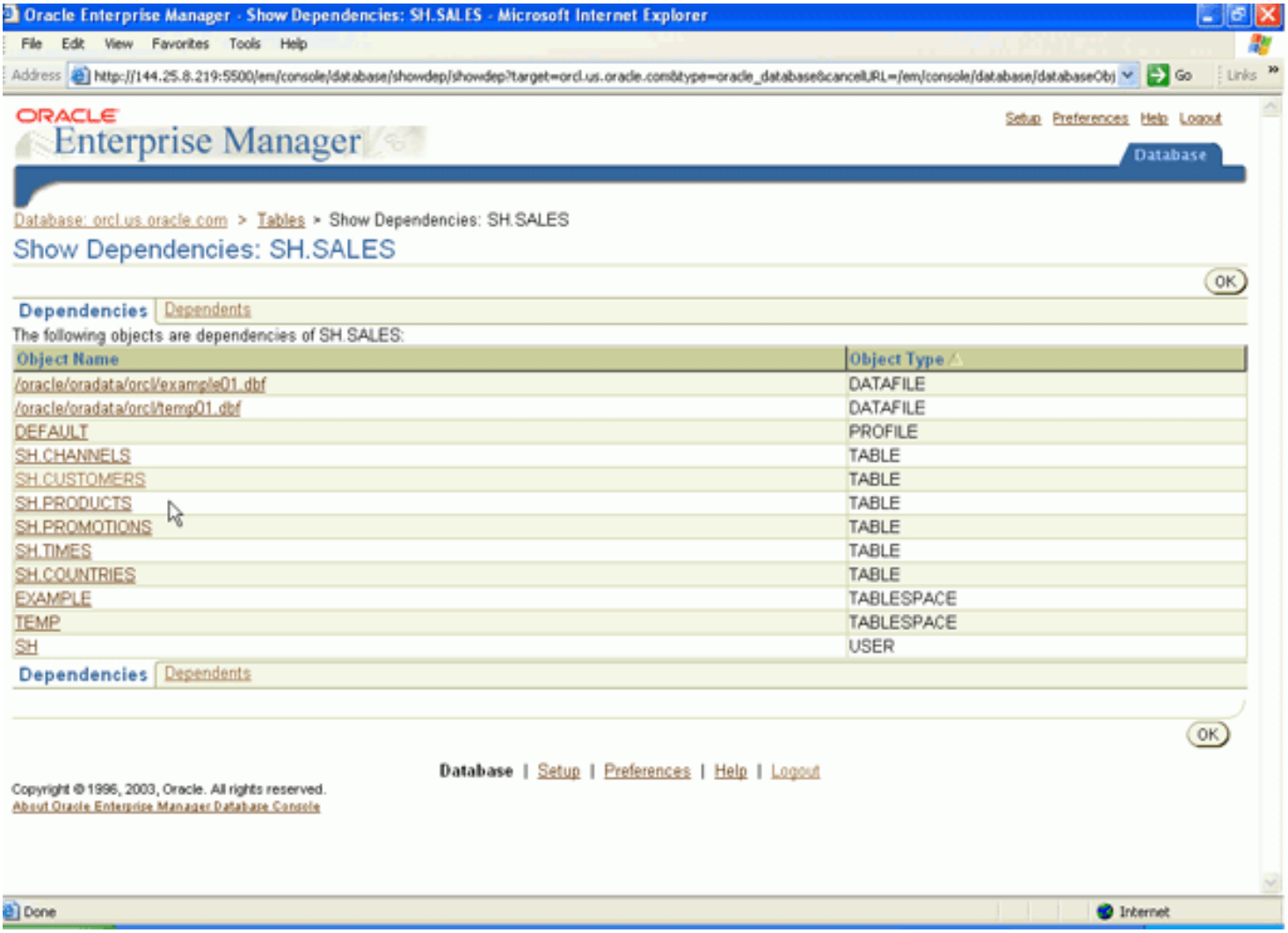
☐ **8.** The Products table is not dependent on any other table. Click **OK** .



☐ **9.** Select **Sales** then select **Show Dependencies** from the drop-down list. Click **Go** .



☐ **10.** The Sales table is dependent on several other tables: Channels, Countries, Customers, Products, Promotions, and Times. Other than Products, none of the other tables are needed for the analysis. Click **OK** .



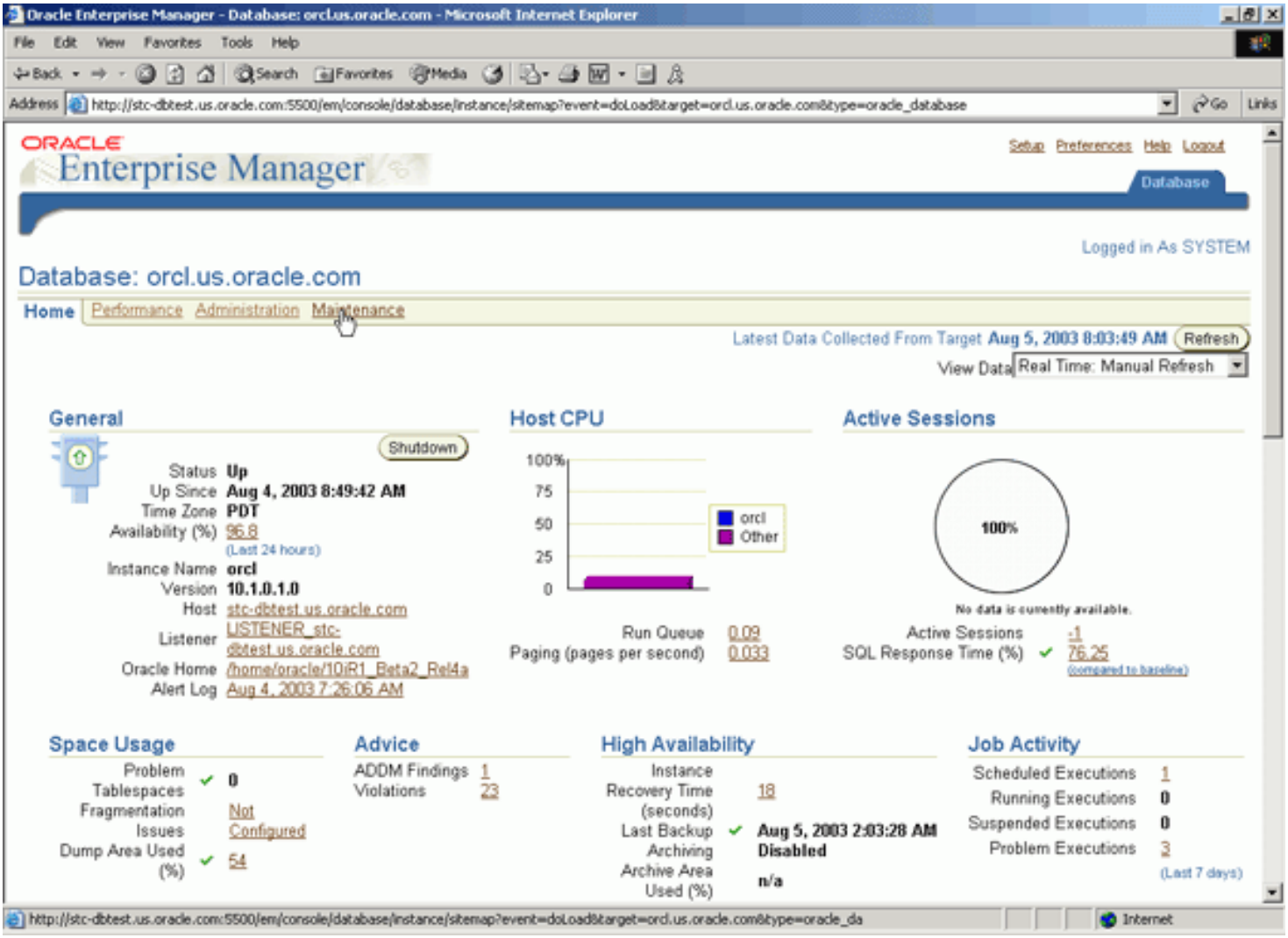
Unloading Data

Back to List

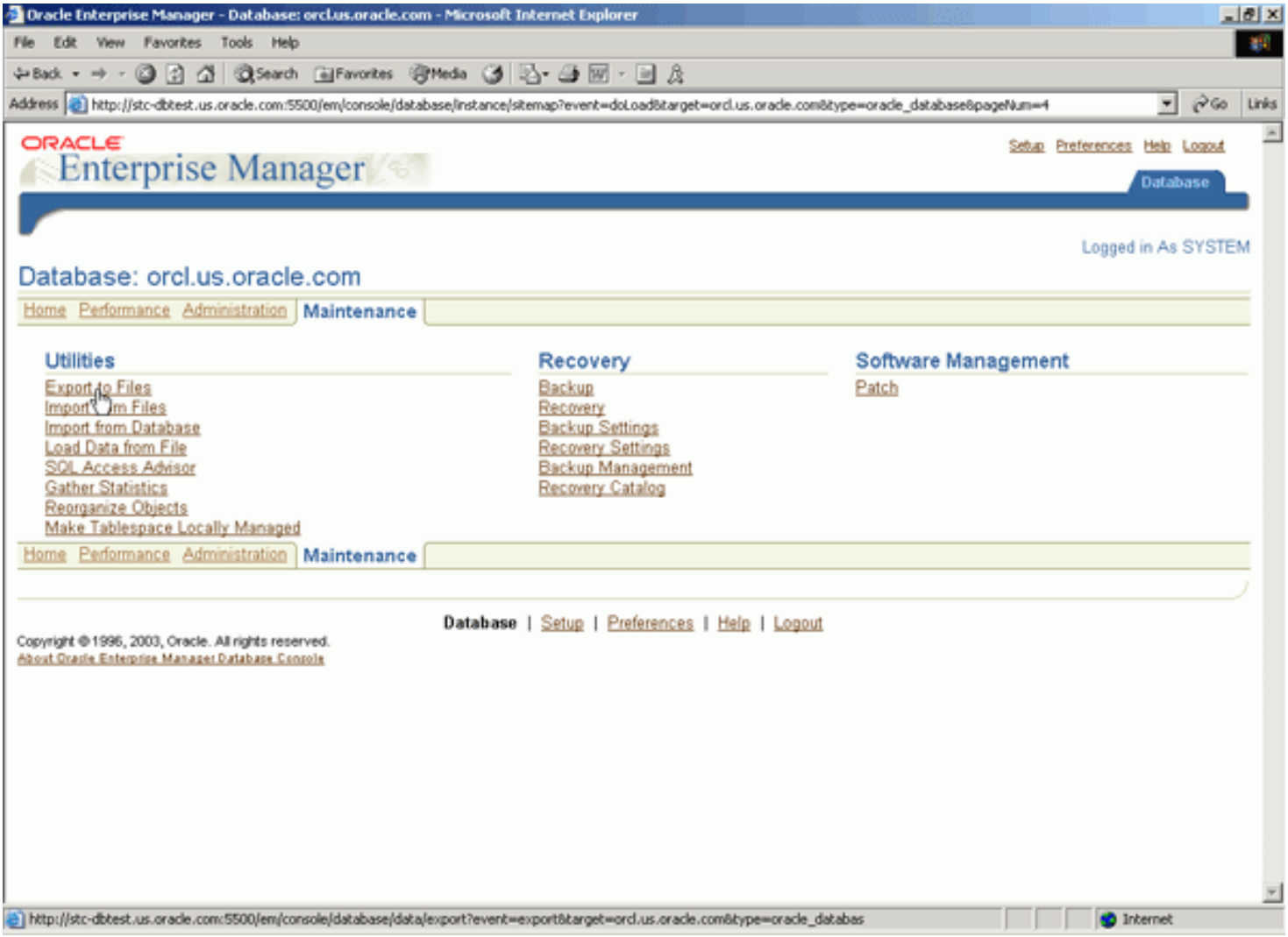
Data Pump Export is a utility for unloading data and metadata into a set of operating system files called a dump file set. The dump file set can be copied to another system and loaded by the Data Pump Import utility. The dump file set is made up of one or more disk files that contain table data, database object metadata, and control information. The files are written in a proprietary, binary format. During an import operation, the Data Pump Import utility uses these files to locate each database object in the dump file set. Data Pump Export allows you to specify that a job should move a subset of the data and metadata, as determined by the export mode. This is done using data filters and metadata filters, which are implemented through Export parameters.

Oracle Data Pump Export can be accessed through Enterprise Manager. To unload your company's sales data, perform the following:

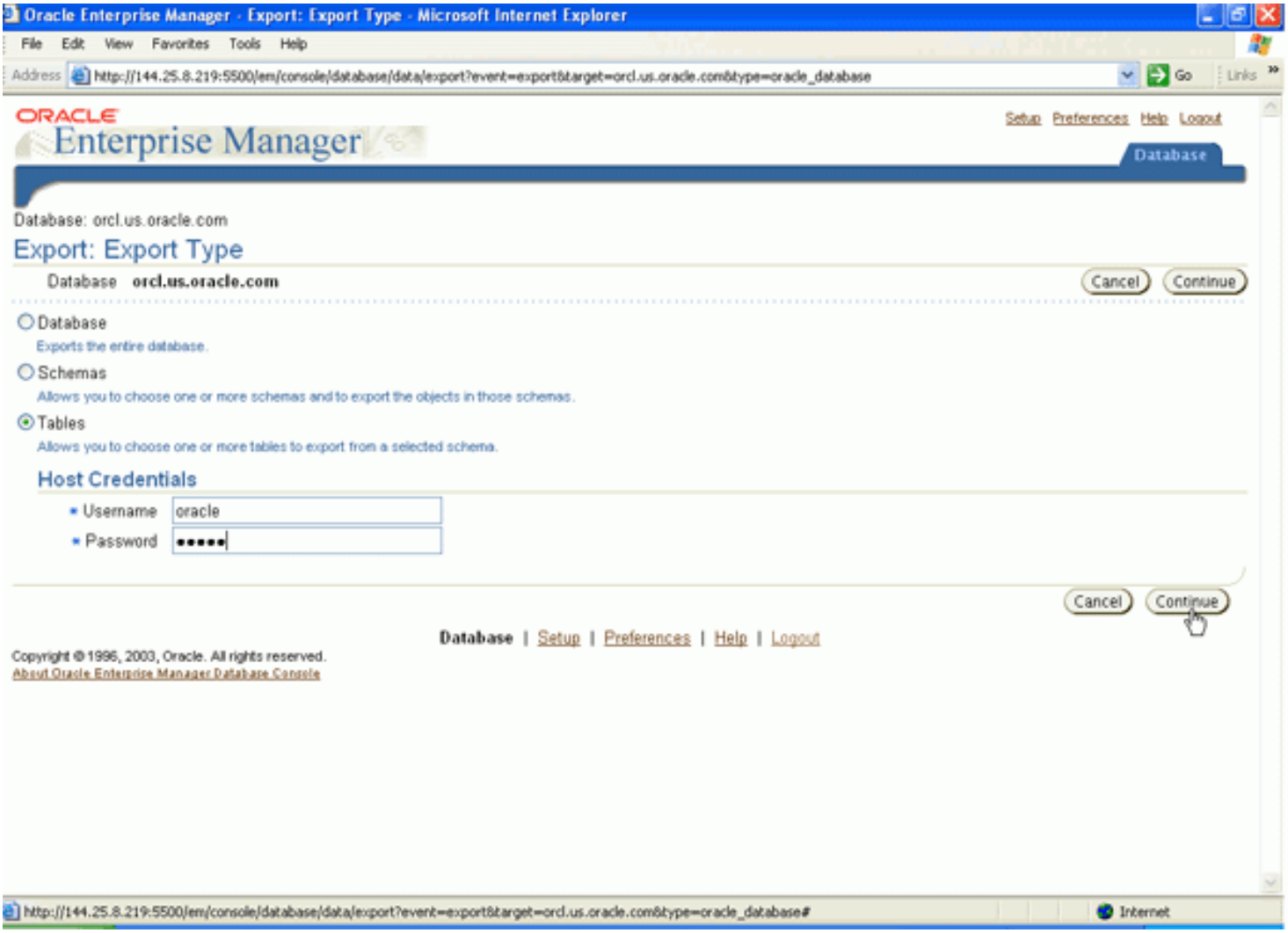
☐ **1.** Click on the **Maintenance** link.



☐ **2.** Click on **Export to Files** link.

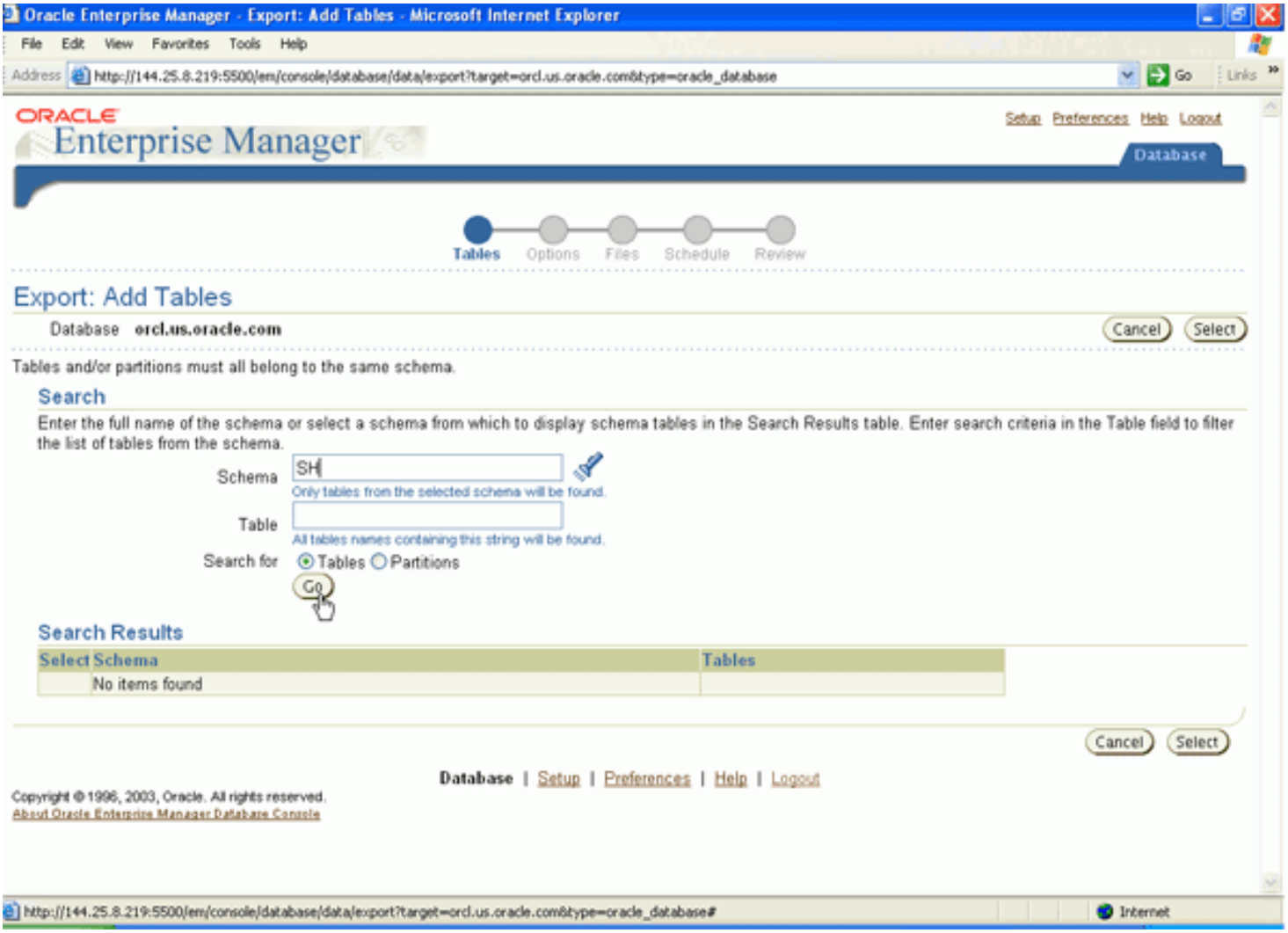


☐ **3.** To export tables, select the **Tables** button and enter your OS username and password then click on **Continue** button.

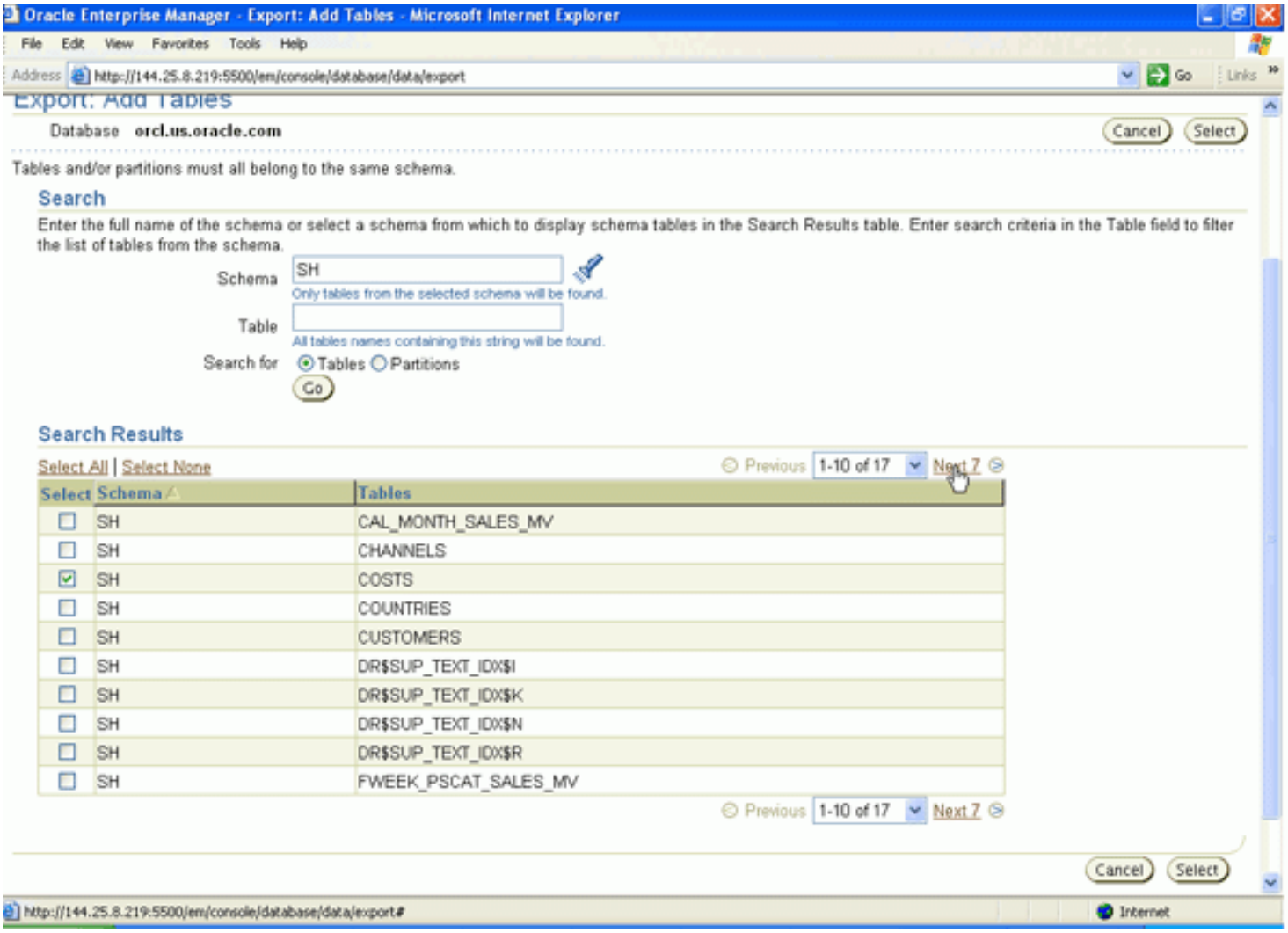


☐ **4.** Select the **Add** button to see the tables available for export.

☐ **5.** Enter **SH** in the Schema field and click **Go** .

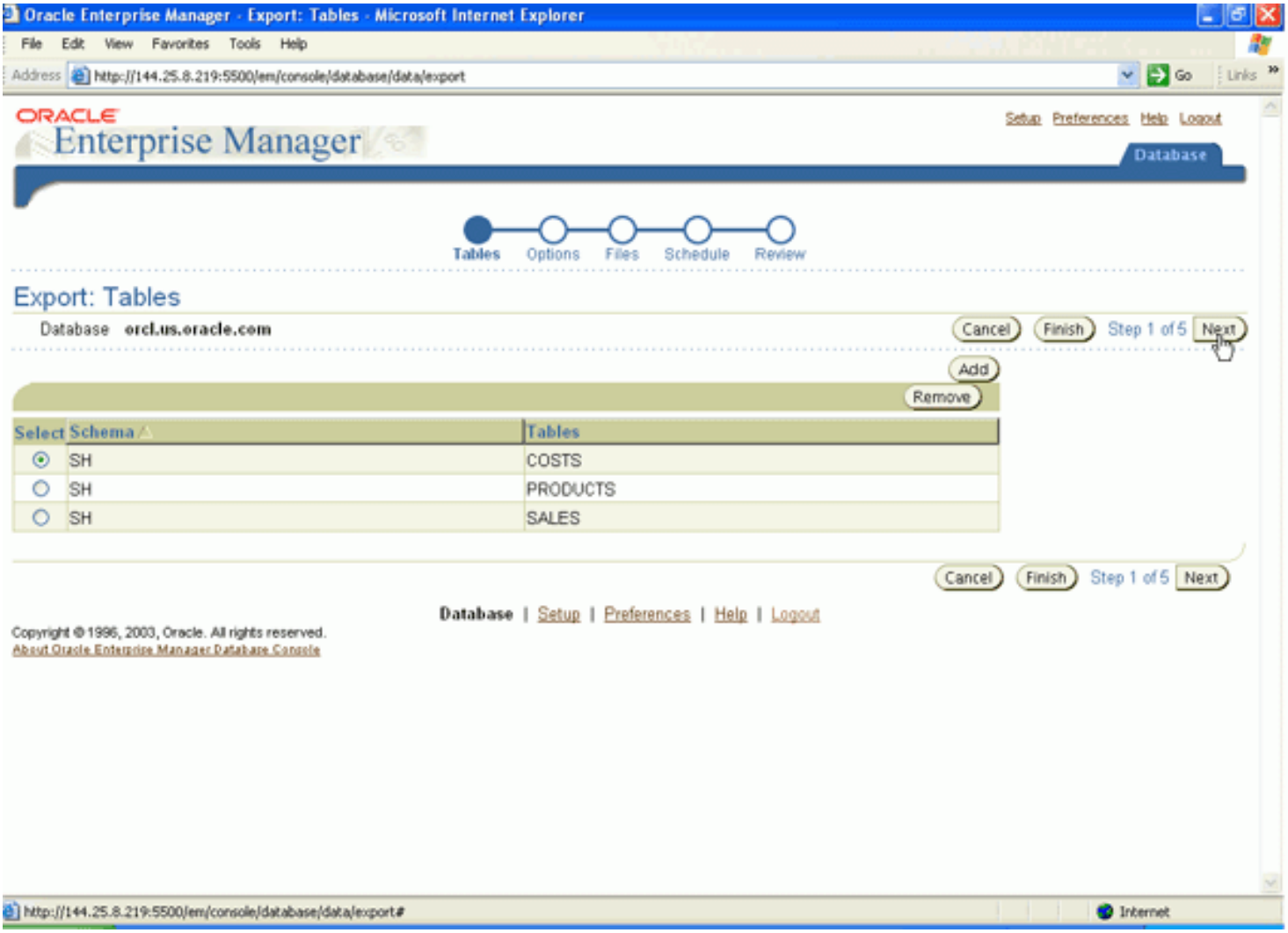


☐ **6.** Select the checkboxes for the **Costs, Products,** and **Sales** tables and click **Select** .


Oracle Enterprise Manager - Export: Add Tables - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://144.25.8.219:5500/em/console/database/data/export
Go
Links
Export: Add Tables
Database orcl.us.oracle.com
Cancel
Select
Tables and/or partitions must all belong to the same schema.
Search
Enter the full name of the schema or select a schema from which to display schema tables in the Search Results table. Enter search criteria in the Table field to filter the list of tables from the schema.
Schema
SH
Only tables from the selected schema will be found.
Table
All tables names containing this string will be found.
Search for
Tables
Partitions
Go
Search Results
Select All | Select None
Previous 1-10 of 17 Next 7
Select
Schema
Tables
SH CAL_MONTH_SALES_MV
SH CHANNELS
SH COSTS
SH COUNTRIES
SH CUSTOMERS
SH DRSUP_TEXT_IDXI
SH DRSUP_TEXT_IDXK
SH DRSUP_TEXT_IDXN
SH DRSUP_TEXT_IDXR
SH FWEEK_PSCAT_SALES_MV
Previous 1-10 of 17 Next 7
Cancel
Select
http://144.25.8.219:5500/em/console/database/data/export#
Internet


Oracle Enterprise Manager - Export: Add Tables - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://144.25.8.219:5500/em/console/database/data/export
Go
Links
ORACLE
Enterprise Manager
Setup Preferences Help Logout
Database
Tables Options Files Schedule Review
Export: Add Tables
Database orcl.us.oracle.com
Cancel Select
Tables and/or partitions must all belong to the same schema.
Search
Enter the full name of the schema or select a schema from which to display schema tables in the Search Results table. Enter search criteria in the Table field to filter the list of tables from the schema.
Schema SH
Only tables from the selected schema will be found.
Table
All tables names containing this string will be found.
Search for Tables Partitions
Go
Search Results
Select All | Select None
Previous 10 11-17 of 17 Next
Select Schema Tables
SH MVIEW$_EXCEPTIONS
SH PRODUCTS
SH PROMOTIONS
SH SALES
SH SALES_TRANSACTIONS_EXT
SH SUPPLEMENTARY_DEMOGRAPHICS
http://144.25.8.219:5500/em/console/database/data/export#
Internet

☐ **7.** Click **Next** .

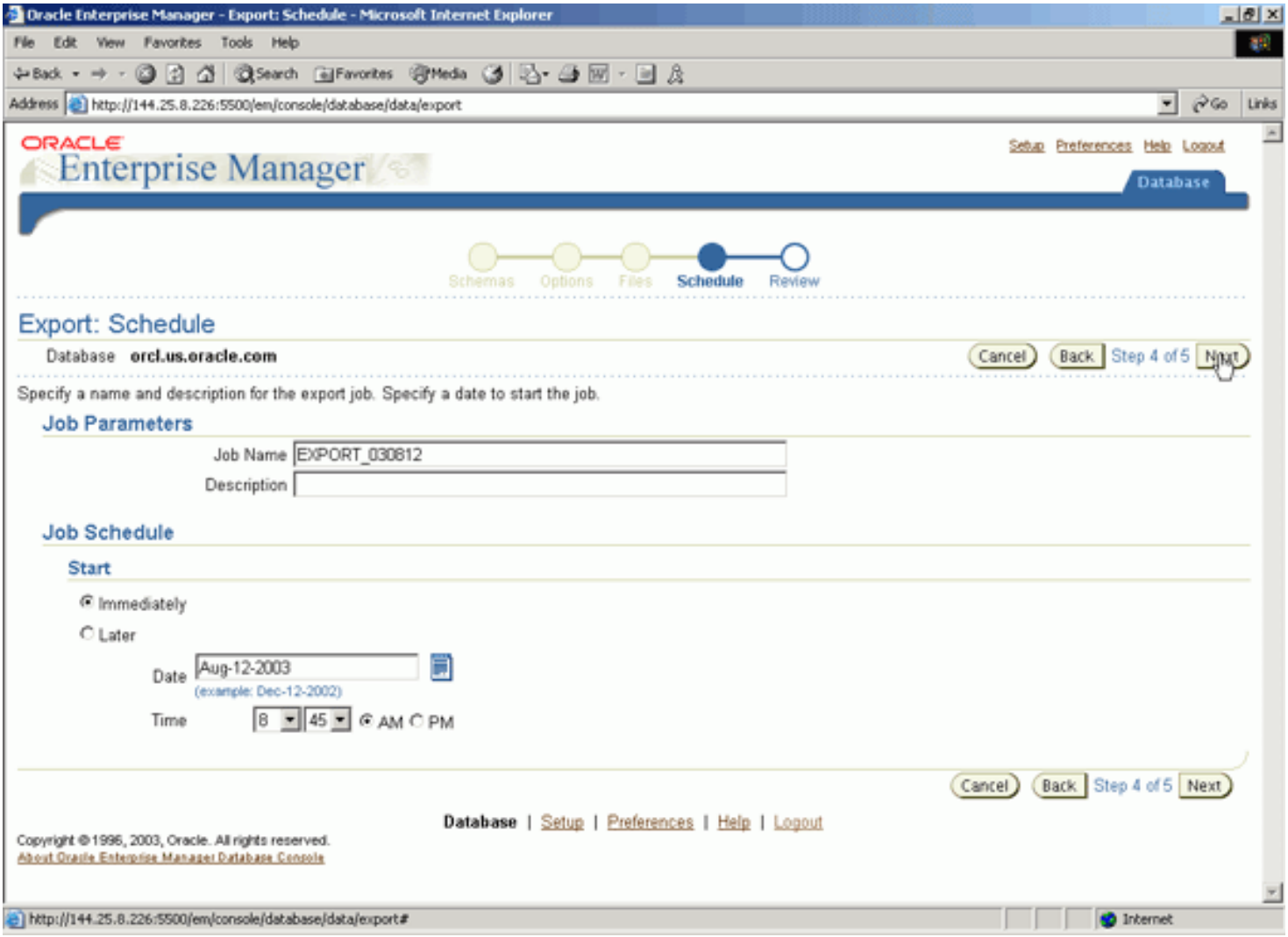


☐ **8.** Enter **4** for the Maximum Number of Threads in Export Job and select **DATADIR1** for the Directory Object. Change the name of the export log file to **EXPORT_<today's date in mmddyy format>.log** and click **Next** .

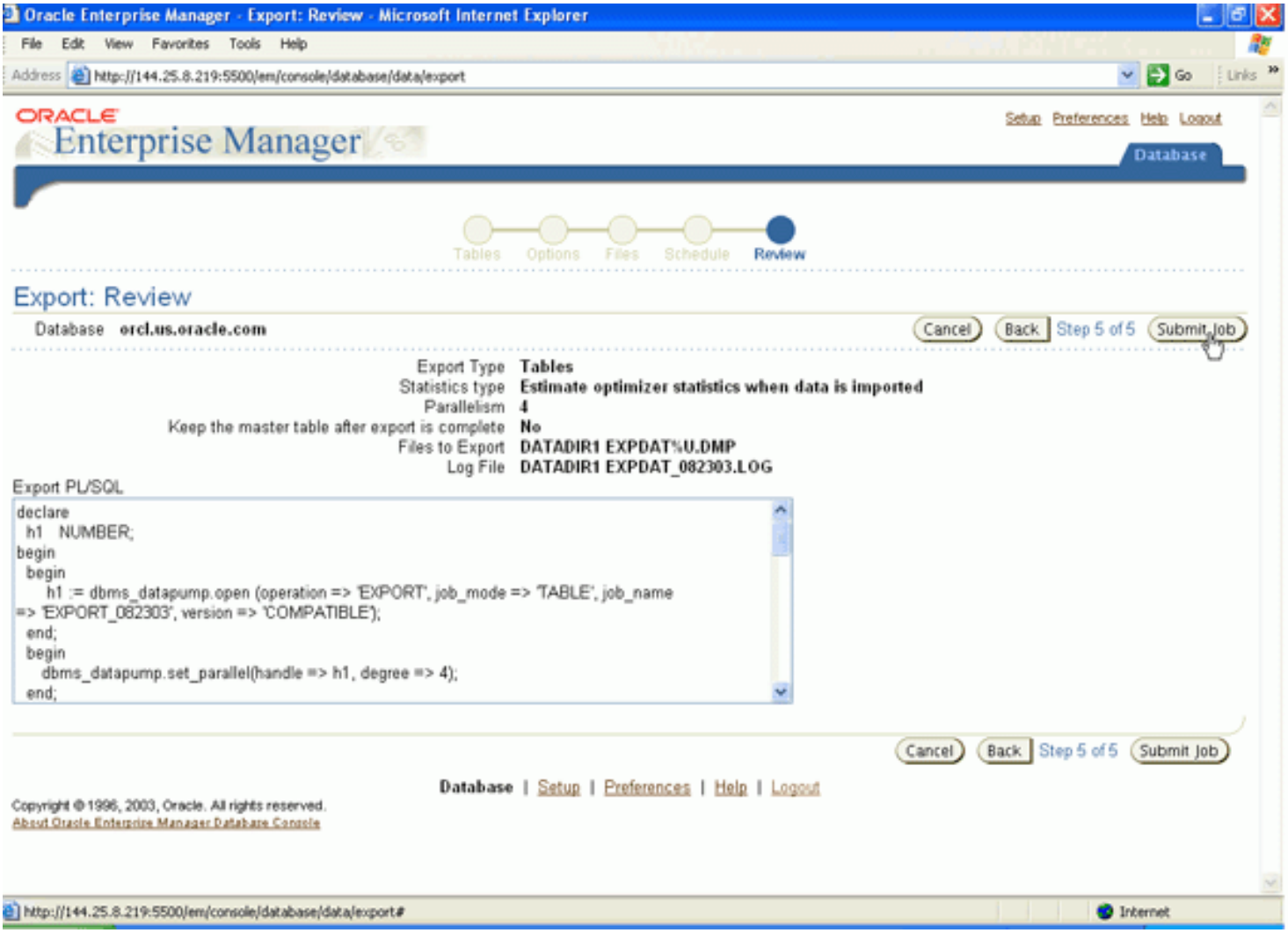

Oracle Enterprise Manager - Export: Options - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://144.25.8.219:5500/em/console/database/data/export
Go
Links
ORACLE
Enterprise Manager
Setup Preferences Help Logout
Database
Tables Options Files Schedule Review
Export: Options
Database orcl.us.oracle.com
Cancel Finish Back Step 2 of 5 Next
Maximum Number of Threads in Export Job 4
This option allows you to make tradeoffs between resource consumption and elapsed time.
Keep the master table after export is complete
Estimate Disk Space
Calculates an estimate of how much disk space the export job will consume (in bytes). The estimate is for table row data only and does not include metadata.
Blocks
Estimate will be calculated by multiplying the number of database blocks used by the target objects times the appropriate block sizes. This method will provide the quickest rough estimate.
Statistics
Estimate will be calculated using per-table statistics. This method will provide the most accuracy if all target tables have been recently analyzed.
Sampling
Estimate will be made sampling a fixed number of rows per table and extrapolating.
Estimate Disk Space Now
Calculate the estimate of space that will be consumed without actually performing the export operation. This may take a few minutes.
Optional Files
Generate Log File
Directory Object DATADIR1 Create Directory Object
Log File EXPDAT_082303.LOG
Show Advanced Options
Cancel Finish Back Step 2 of 5 Next
http://144.25.8.219:5500/em/console/database/data/export#
Internet

☐ **9.** Click **Next** .

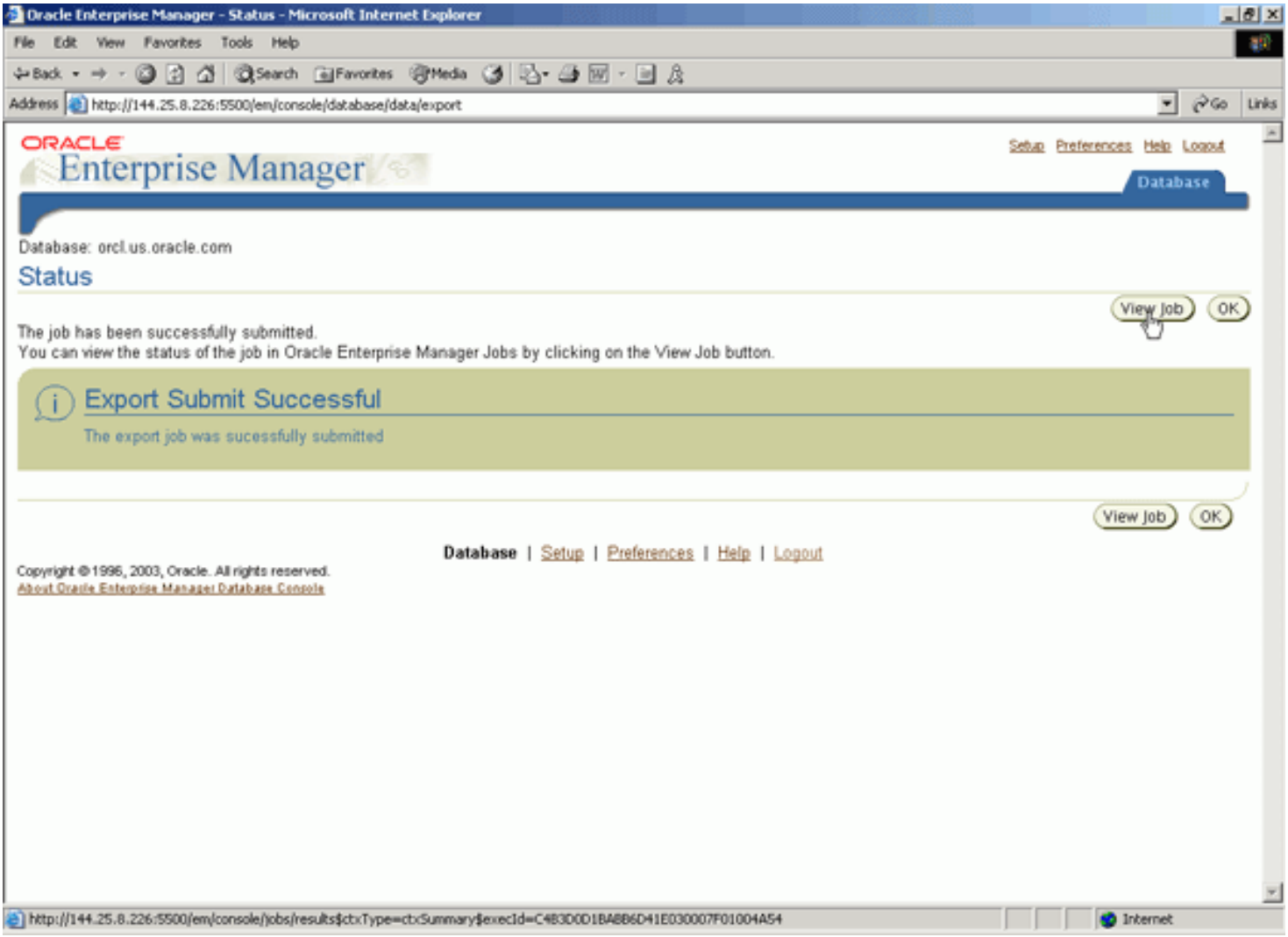
☐ **10.** Enter a job name of **EXPORT<Today's Date>** and make sure Start is set to **Immediate** and click **Next** .



11. ☐ Select the **Submit Job** button to submit the export job.

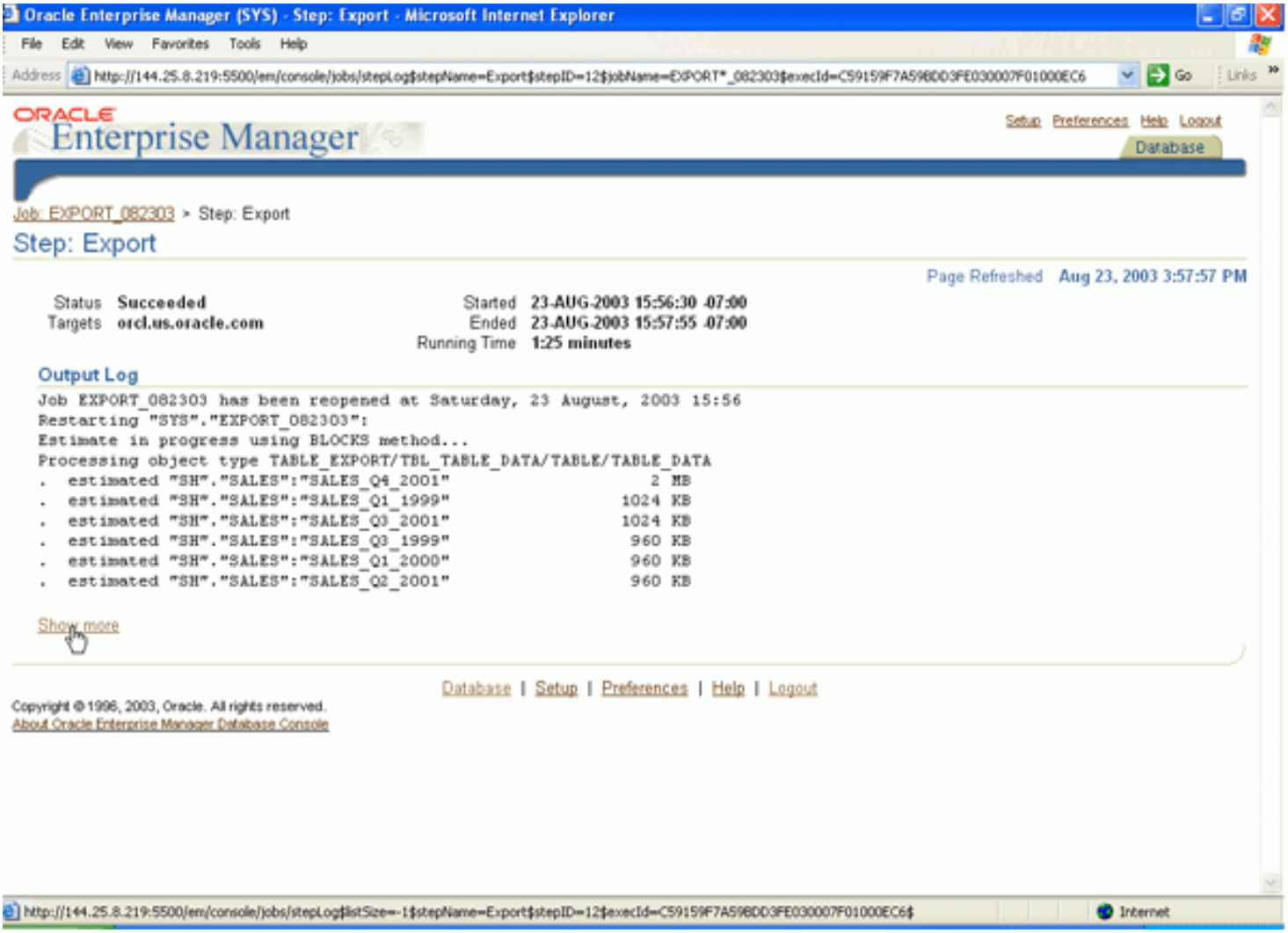


12. ☐ Your export job was successfully submitted. Click **View Job** .

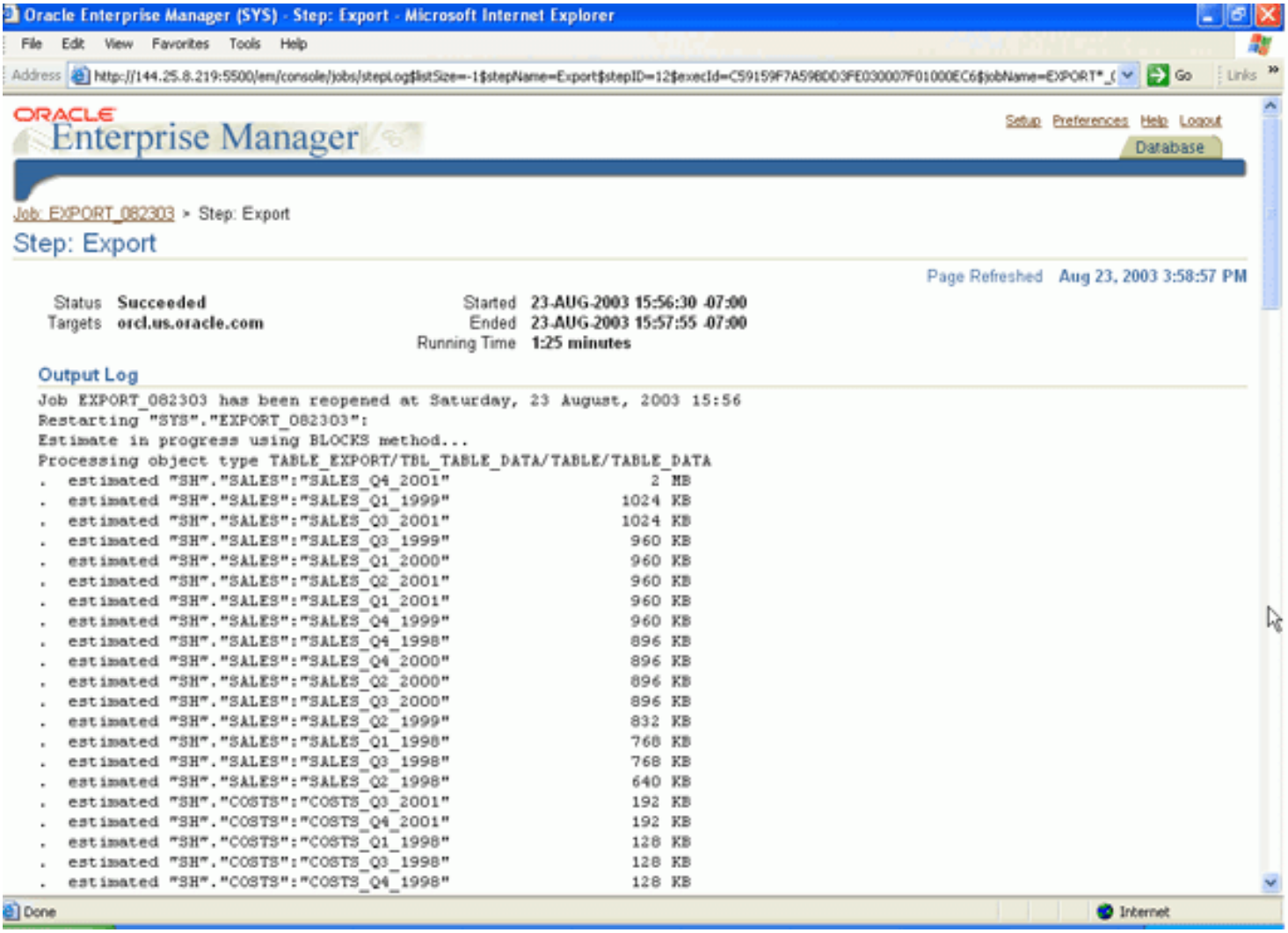


13. ☐ Select the **Export** link to see the Export log status.

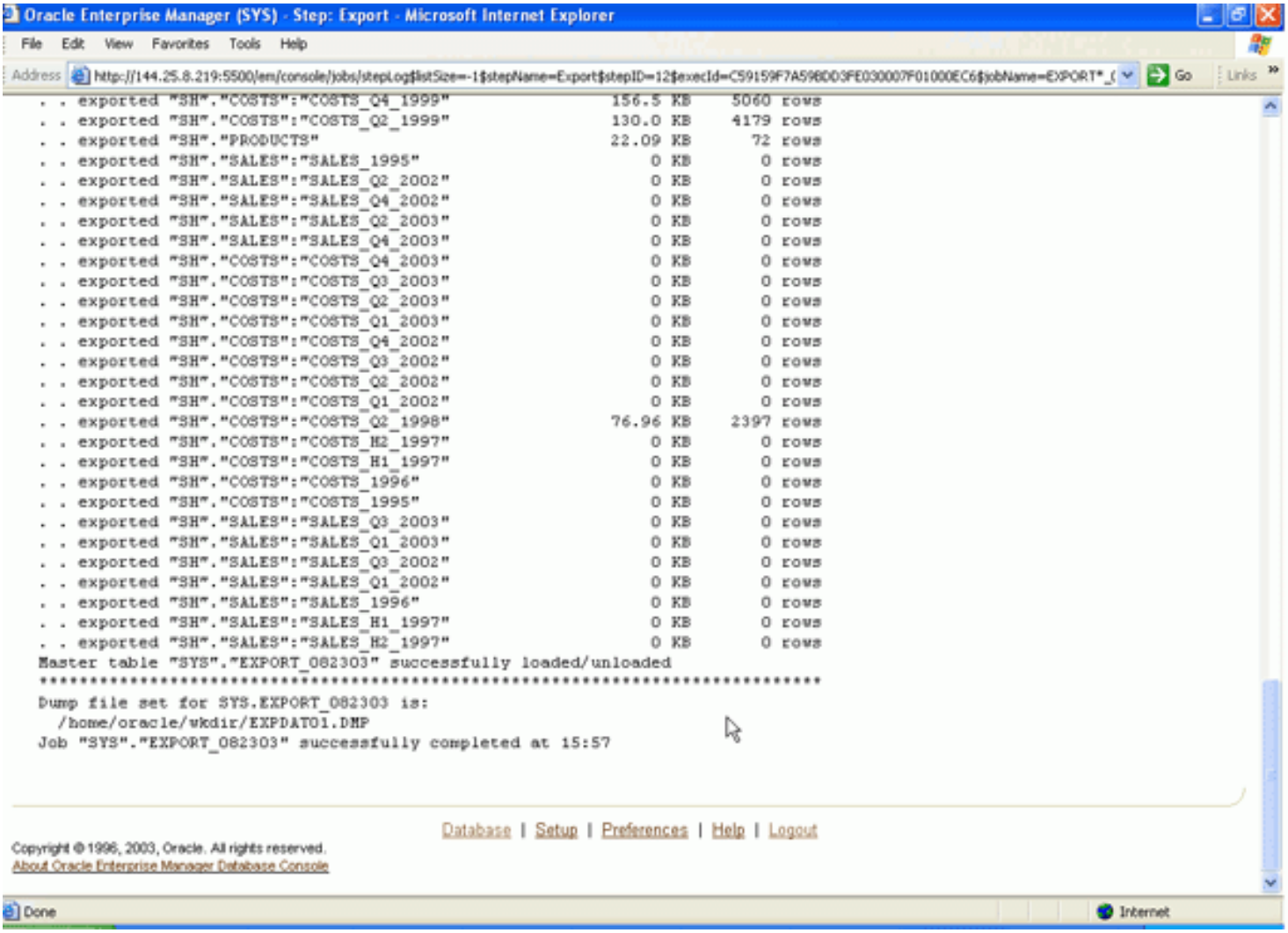
☐ **14.** The job is still running. Click **Show more** to see more of the log. If Show more does not appear, click Reload in your browser window.



15. ☐ The job has finished. Scroll down to the bottom to see all the messages in the log.



16. ☐ Your export has completed successfully. Note the name of the dump file for use later in the lesson.



You will perform the following examples of using Data Pump Export command line interface:

- Performing a Table Mode Export
- Estimating how much disk space will be consumed in a schema mode export
- Performing a schema mode export
- Performing a full database export using four parallel processes
- Attaching to and stopping an existing job
- Attaching to an restarting a stopped job

Performing a Table Mode Export

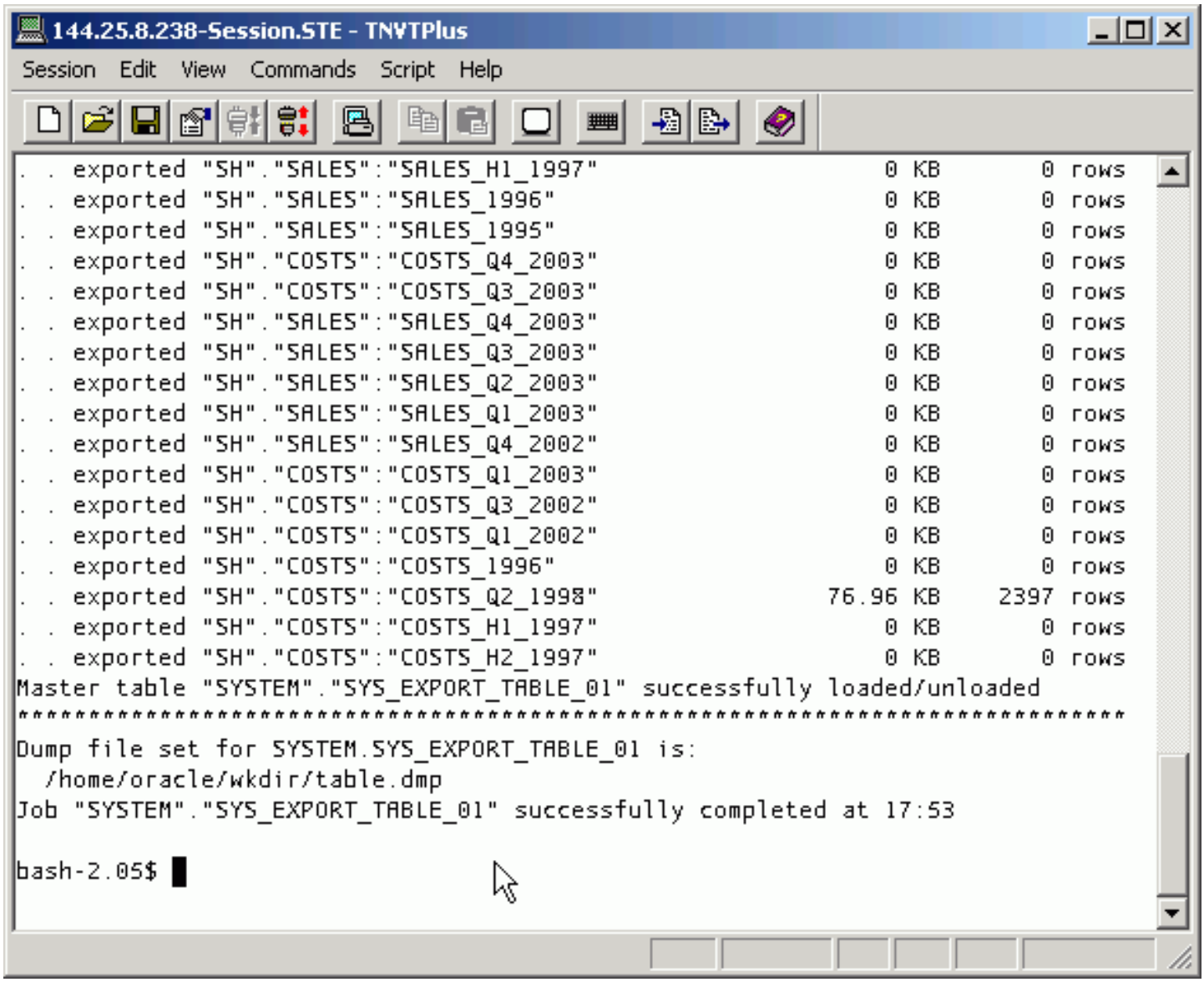
Back to List

A table export is specified using the TABLES parameter. In the following example, the NOLOGFILE parameter indicates that an Export logfile of the operation will not be generated.

Issue the following export command to perform a table export of table Costs and table Sales:

☐ **1.** Open a terminal window and execute the following:

```
expdp system/<password> \
TABLES=sh.costs,sh.sales \
DUMPFILE=datadir2:table.dmp \
NOLOGFILE=y
```



Estimating How Much Disk Space Will Be Consumed in a Schema Mode Export

Back to List

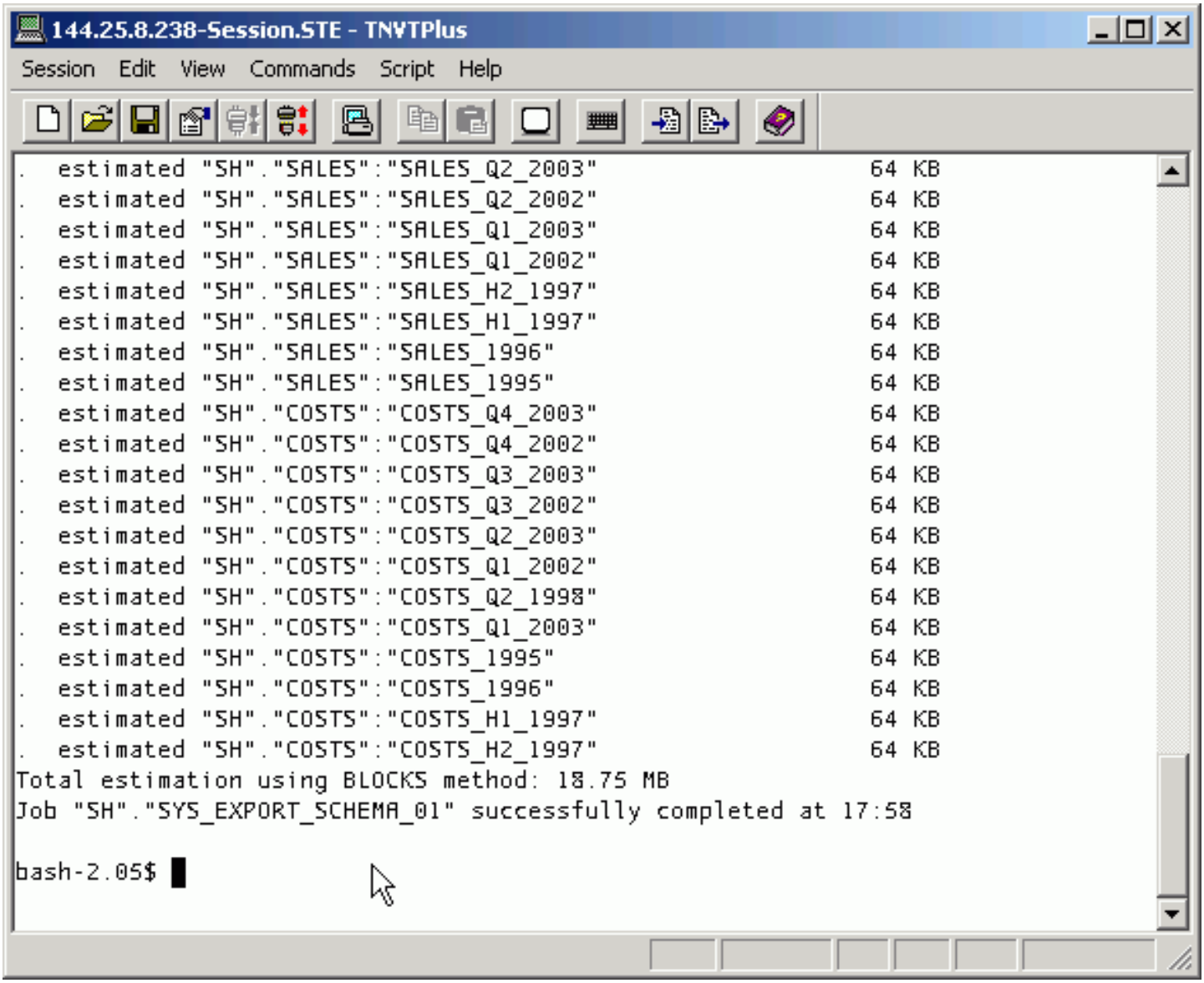
The ESTIMATE_ONLY parameter estimates the space that would be consumed in a schema export, but stops without actually performing the export operation. The estimate is printed in the log file and displayed on the client's standard output device. The estimate is for table row data only; it does not include metadata.

The INCLUDE parameter allows you to filter the metadata that is exported by specifying objects and object types for the current export mode. The specified objects and all their dependent objects are exported. Grants on these objects are also exported.

Perform the following:

☐ **1.** From your terminal window, issue the following command to estimate the amount of blocks required to export the data in the three tables: Sales, Products and Costs, from the Sales History (SH) schema. Use a backslash (\) as an escape character before a special character, such as a parenthesis, so that the character is not treated as a special character by the operating system.

```
expdp sh/sh \
INCLUDE=table:\"IN \( \'SALES\',\'PRODUCTS\',\'COSTS\'\) \" \
DIRECTORY=datadir2 \
ESTIMATE_ONLY=y
```



Performing a Schema Mode Export

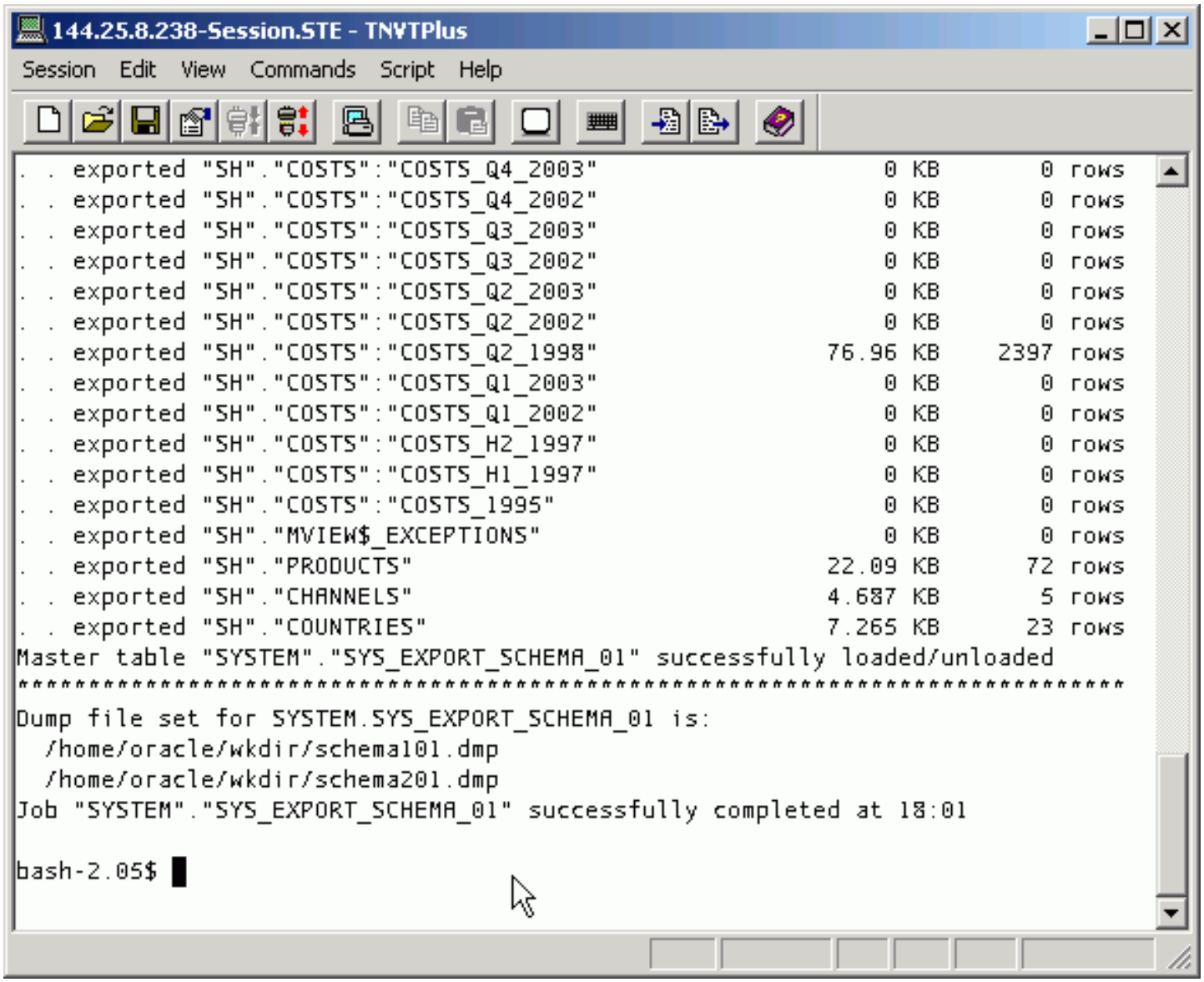
Back to List

A schema export is specified using the SCHEMAS parameter. In a schema export, only objects belonging to the corresponding schemas are unloaded. This is the default mode. If you have the EXP_FULL_DATABASE role, then a list of schemas can be specified, and the schema definitions themselves are included, as well as system privilege grants to those schemas. In the following example, the file names contain a substitution variable (%U), which implies that multiple files may be generated by export.

Perform the following:

☐ **1.** From your terminal window, issue the following export command to perform a schema export:

```
expdp system/<password> \

SCHEMAS=sh \

DUMPFILE=datadir1:schema1%U.dmp,datadir2:schema2%U.dmp \

LOGFILE=datadir1:expschema.log
```



Performing a Parallel Full Database Export

The FULL parameter indicates that the export is a full database mode export. All data and metadata in the database are exported.

The PARALLEL parameter specifies the maximum number of threads of active execution operating on behalf of the export job. This parameter allows you to make trade-offs between resource consumption and elapsed time. For best performance, the value specified for PARALLEL should be at least as large as the number of output files specified with the DUMPFILE parameter. Each Data Pump execution thread writes exclusively to one file at a time.

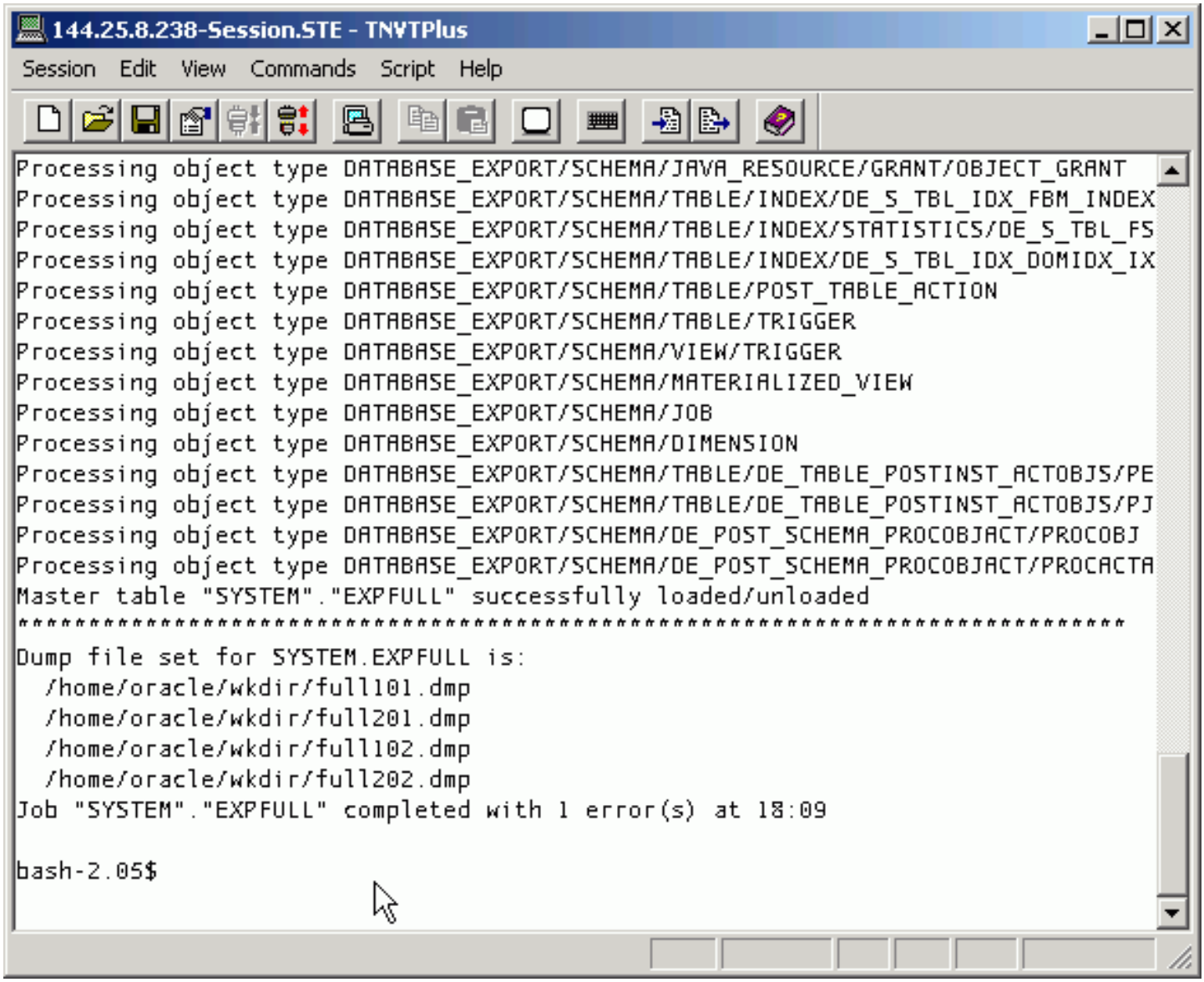
The PARALLEL parameter is only valid in the Enterprise Edition of the Oracle database. To increase or decrease the value of PARALLEL during job execution, use interactive-command mode that is described in the example below.

The FILESIZE parameter will limit the maximum size of each dump file to 2 Gigabytes.

Perform the following:

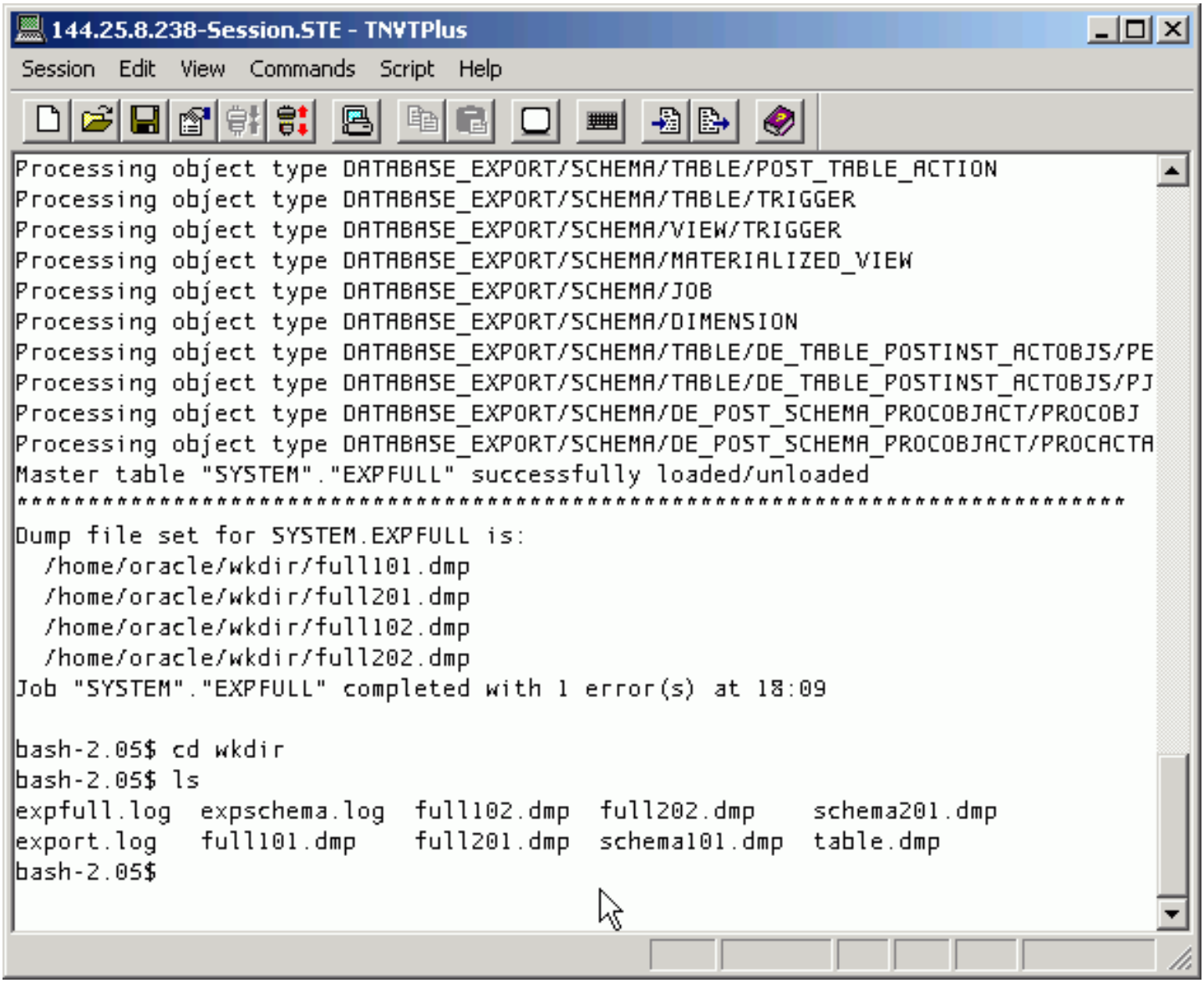
☐ **1.** From your terminal window, issue the following command to perform a full export using the PARALLEL parameter:

```
expdp system/<password> \
FULL=y \

DUMPFILE=datadir1:full1%U.dmp,datadir2:full2%U.dmp \

FILESIZE=2g \

PARALLEL=4 \

LOGFILE=datadir1:expfull.log \

JOB_NAME=expfull
```



Note: The export may complete with an expected error due to not supporting a certain type of table.

☐ **2.** Dump files full101.dmp, full201.dmp, full102.dmp, full202.dmp, and so on will be created in a round-robin fashion in the directories pointed to by datadir1 and datadir2. For best performance, these directories should be on separate I/O channels. Each dump file will be limited to 2 gigabytes in size. The job name will be expfull. The export log file will be written to expfull.log in the directory datadir1.



Attaching to and Stopping an Existing Job

Back to List

The ATTACH command attaches the client session to an existing export job and automatically places you in the interactive-command interface. Export displays a description of the job to which you are attached and also displays the export prompt. A job name does not have to be specified if there is only one export job that is associated with your schema. The job you attach to can be either currently executing or stopped.

In the following example, interactive mode will be run on the same terminal on which the export job is running. A user could also use interactive mode from a terminal other than the one on which the job is run; in this case, the expdp system/<password> ATTACH is required.

Perform the following:

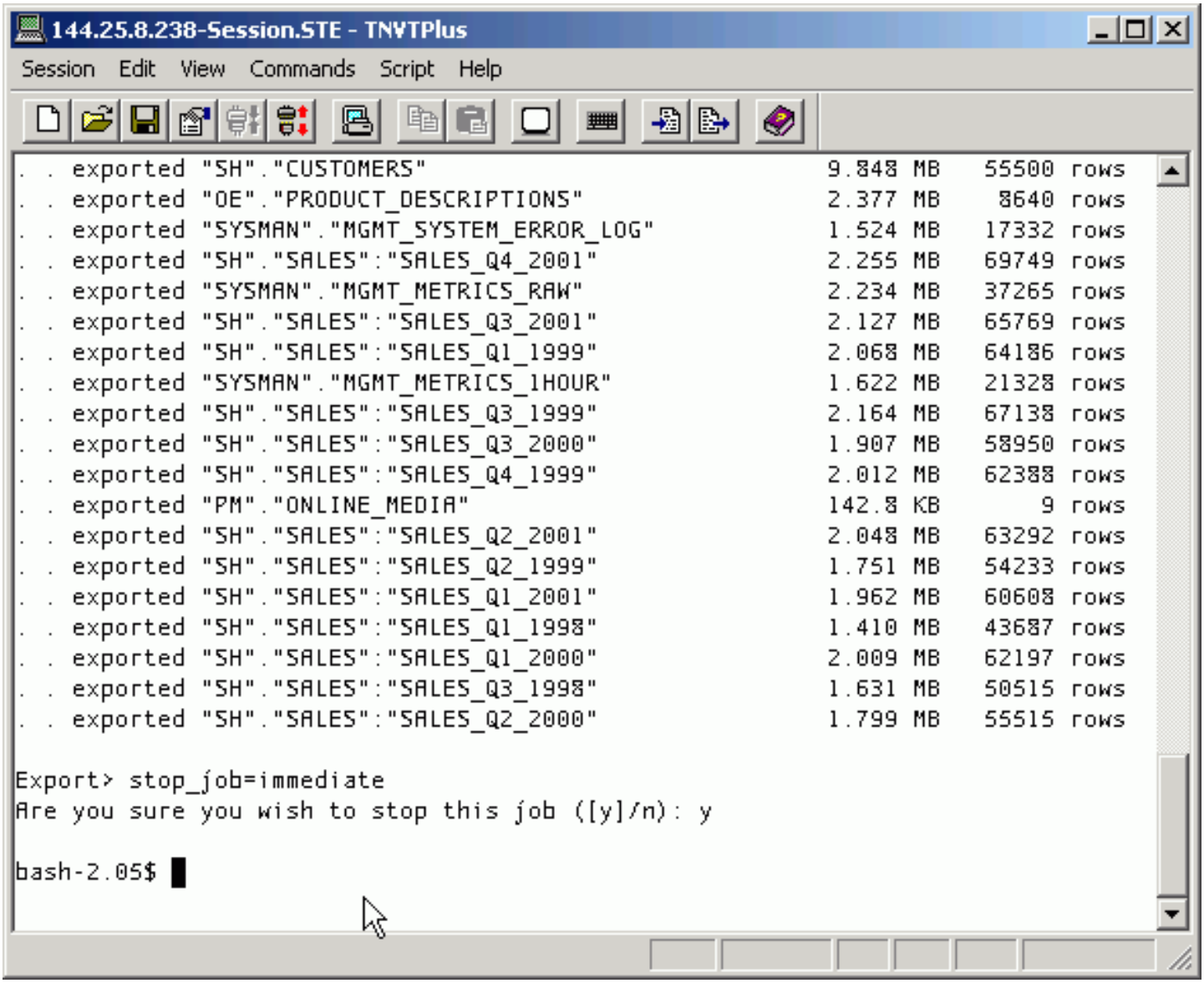
☐ **1.** Run the Full Export again. While the export is running, press **Ctrl+C** , to connect to the interactive-command interface, which is required for the next example. The interactive-command interface stops logging to the terminal and displays the Export prompt, from which you can enter various commands, some of which are specific to interactive mode.

```
expdp system/<password> \
FULL=y \

DUMPFILE=datadir1:full3%U.dmp,datadir2:full4%U.dmp \

FILESIZE=2g \

PARALLEL=4 \

LOGFILE=datadir1:expfull2.log \

JOB_NAME=expfull2
```

Note: Due to the worker processes running in parallel, the display of metadata objects processed during the export may not be sequential.

☐ **2.** From a terminal window, issue the following command to stop the job:

```
Export> STOP_JOB=immediate

Are you sure you wish to stop this job (y/n): y
```



Attaching to and Restarting a Stopped Job

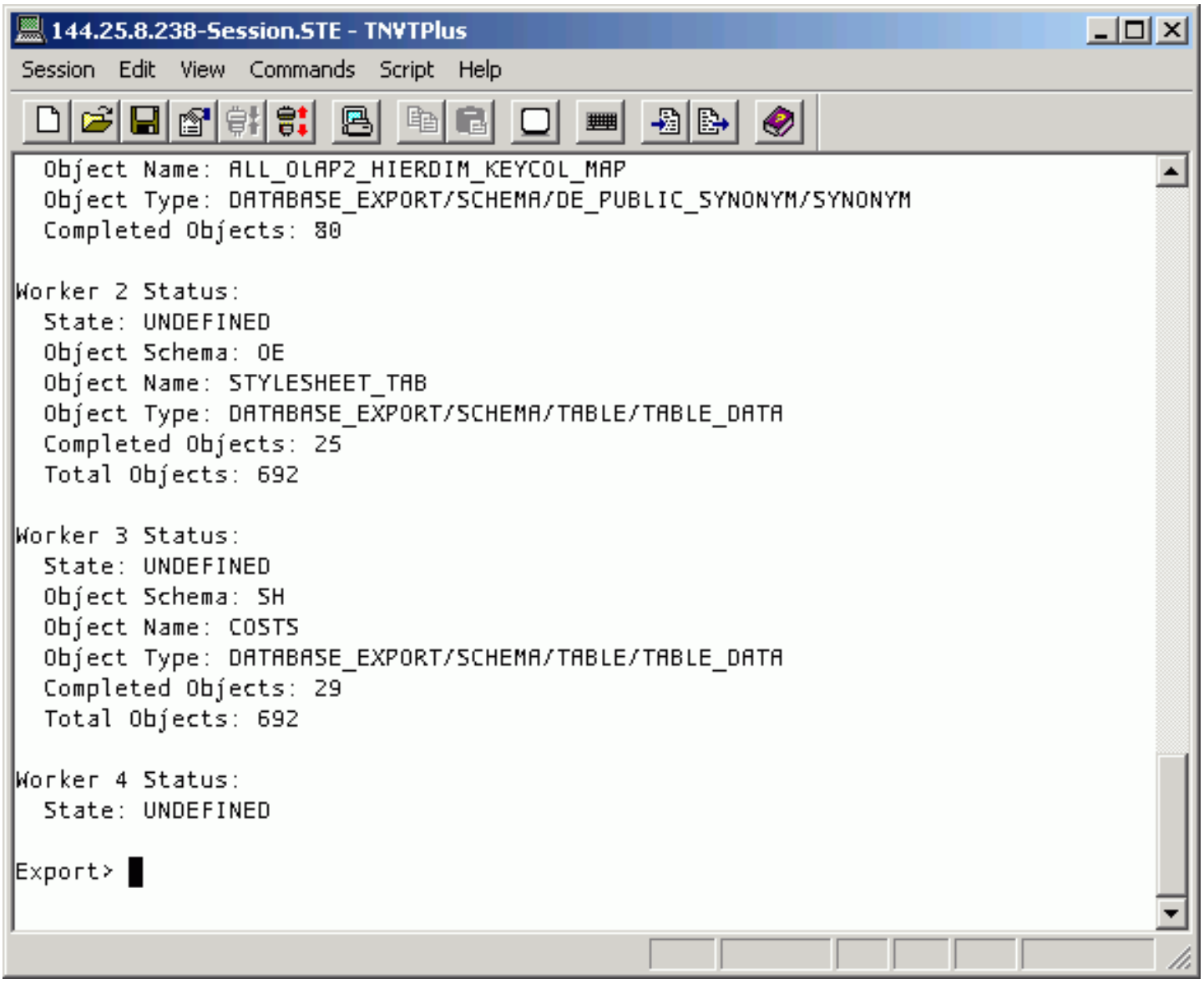
Back to List

The system manager restarts the job (perhaps during off hours) with a higher degree of parallelism. Note that a job name must be supplied in this case because the job was previously stopped. The job name is required in order to find the master table for the job. The system manager would also like a cumulative status of the job to be displayed, along with a description of the current operation. The system manager specifies how frequently, in seconds, this status should be displayed. This status information is written only to the standard output device, not to the log file.

Perform the following:

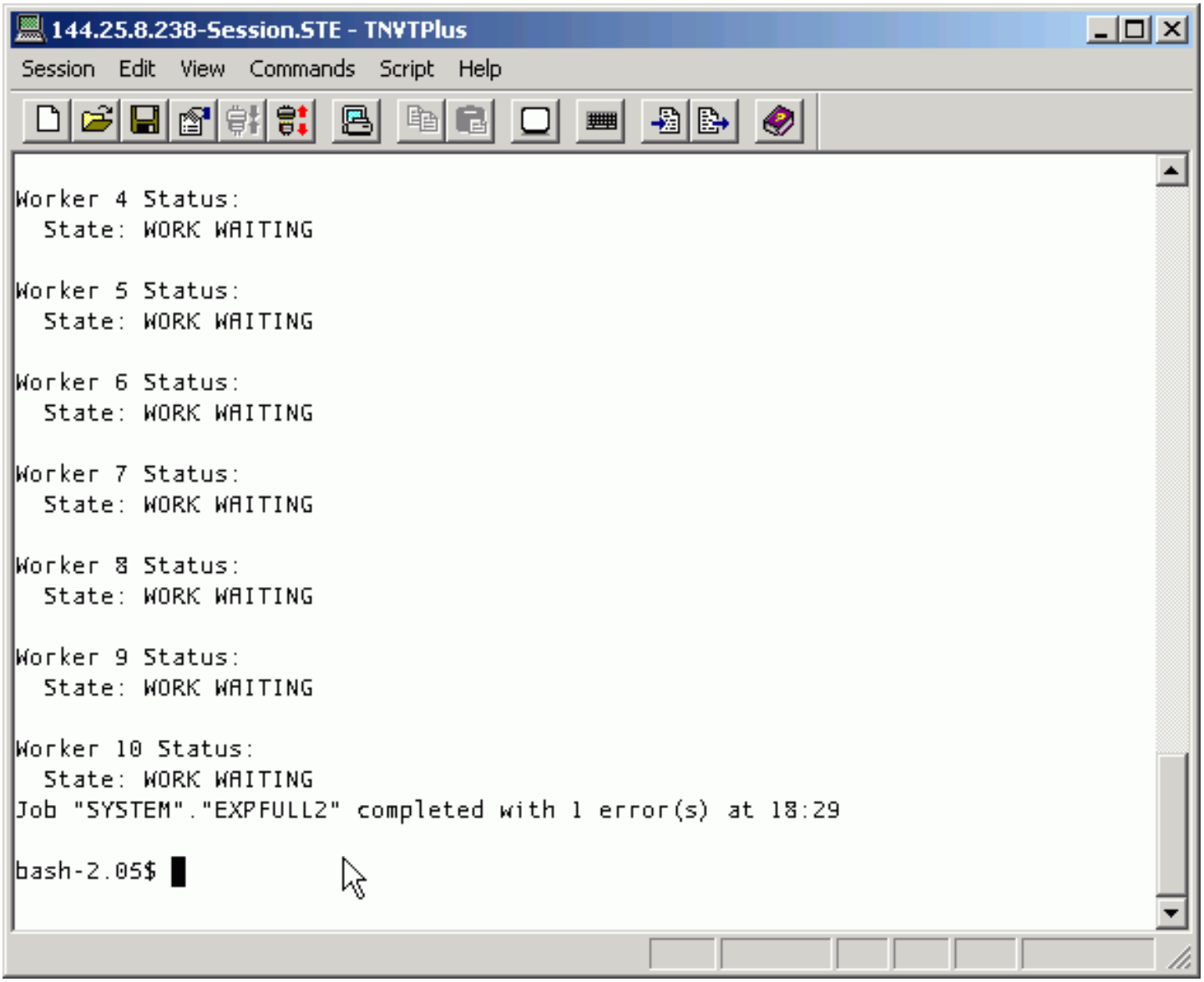
☐ **1.** From your terminal window, issue the following command:

```
expdp system/<password> ATTACH=expfull2
```



☐ **2.** After all the job statuses are displayed, issue the following interactive mode commands:

```
Export> PARALLEL=10
Export> START_JOB
Export> STATUS=600
Export> CONTINUE_CLIENT
```



Logging mode is entered, in which job status is continually output to the terminal every 10 minutes.

Note: The export may complete with an expected error due to stopping the job.

Loading Data

Back to List

Data Pump Import is a utility for loading an export dump file set into a target system. The dump file set is made up of one or more disk files that contain table data, database object metadata, and control information. The files are written by the Data Pump Export utility in a proprietary, binary format. During an import operation, the Data Pump Import utility uses these files to locate each database object in the dump file set.

Import can also be used to load a target database directly from a source database with no intervening files, which allows export and import operations to run concurrently. This avoids the creation of dump files on the file system, and may also minimize the total elapsed time for the entire export and import operation. This is known as network import.

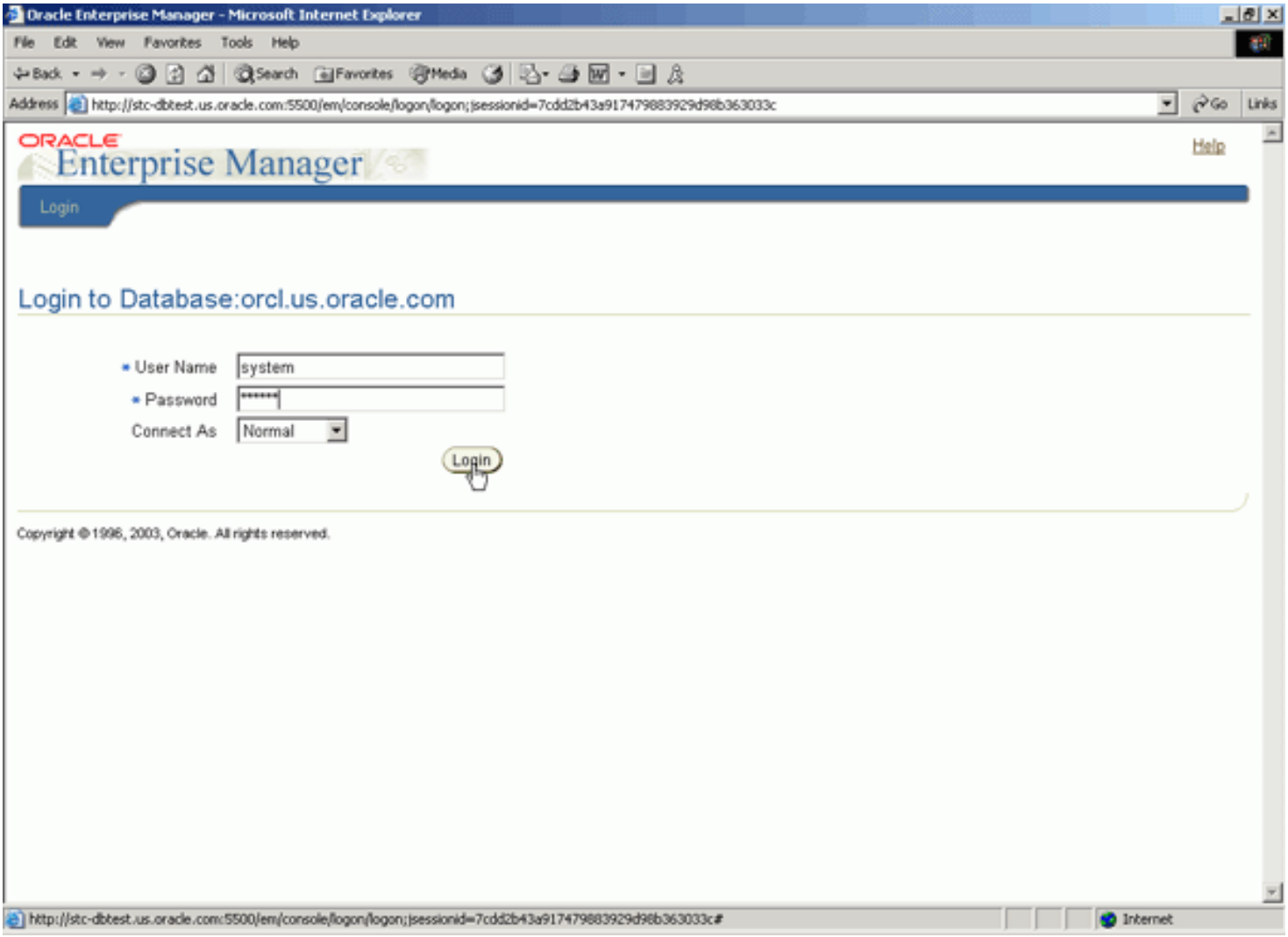
Data Pump Import allows you to specify whether a job should move a subset of the data and metadata, as determined by the import mode. This is done using data filters and metadata filters, which are implemented through Import parameters.

Oracle Data Pump Import can be accessed through Enterprise Manager. To import a schema through Enterprise Manager, perform the following:

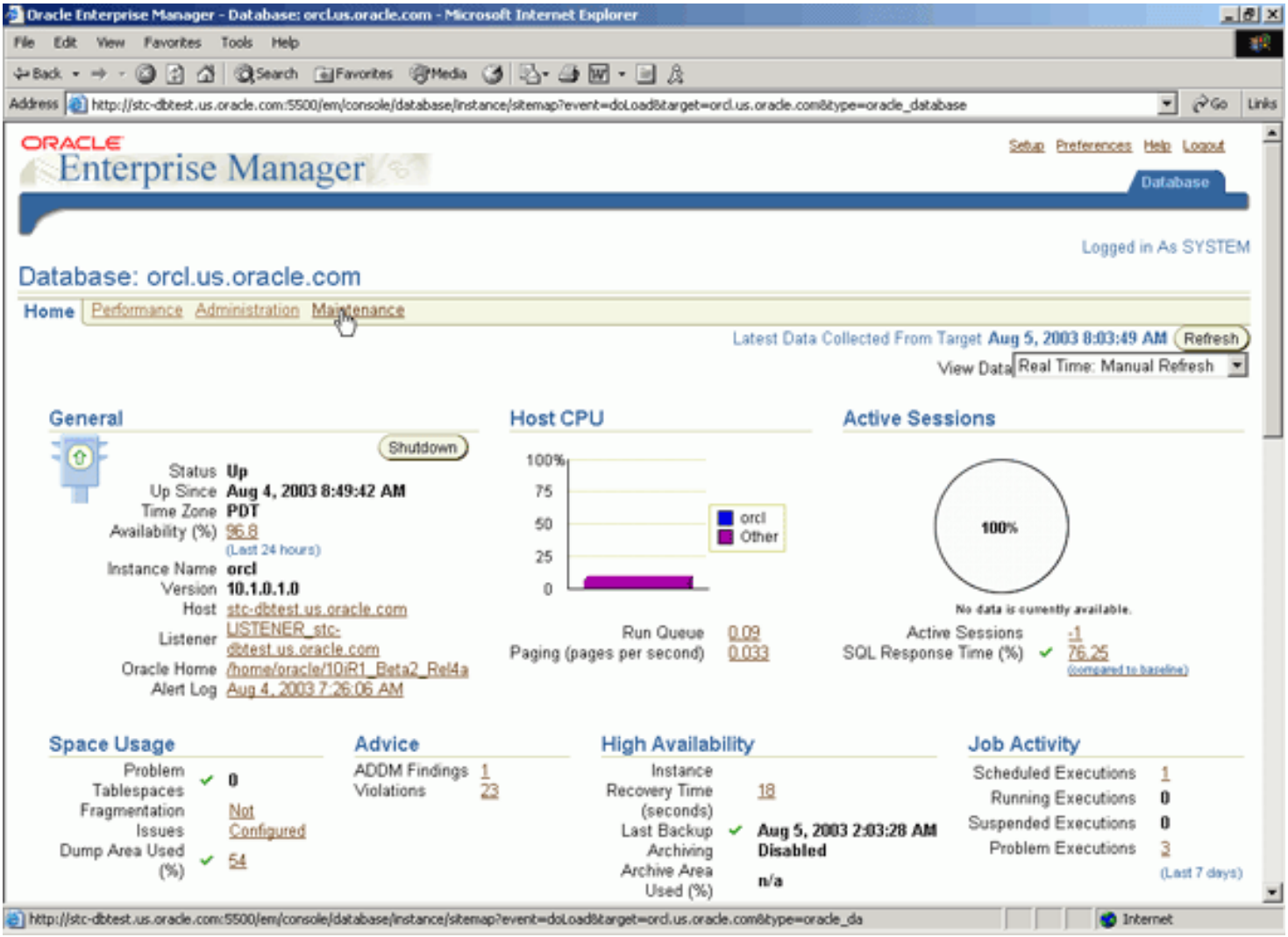
☐ **1.** Open a browser and enter the following URL:

`http://<hostname>:5500/em`

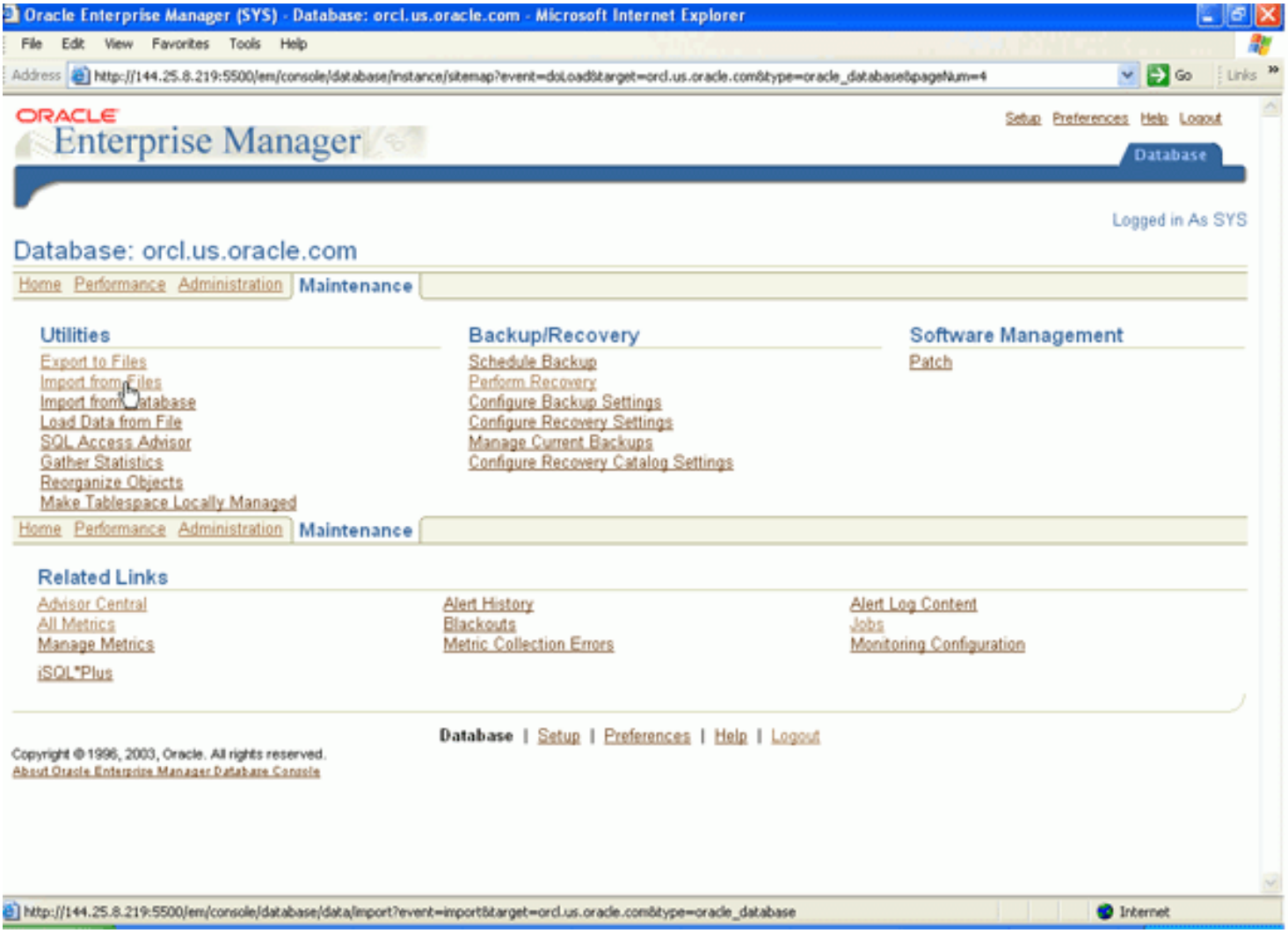
Login as **system/<password>** then click **Login** .



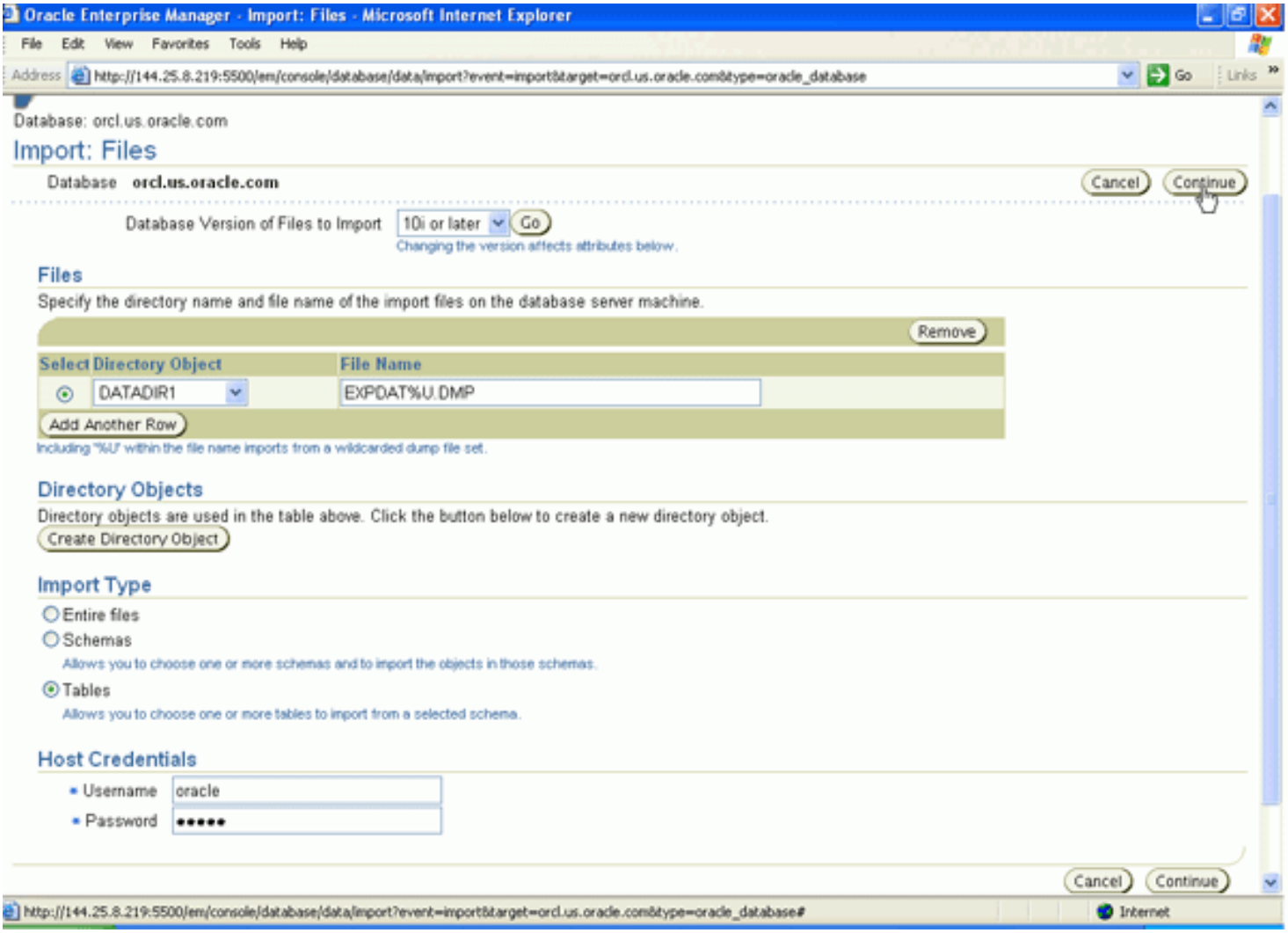
☐ **2.** Click on the **Maintenance** link.



☐ **3.** Click on **Import from Files** link.



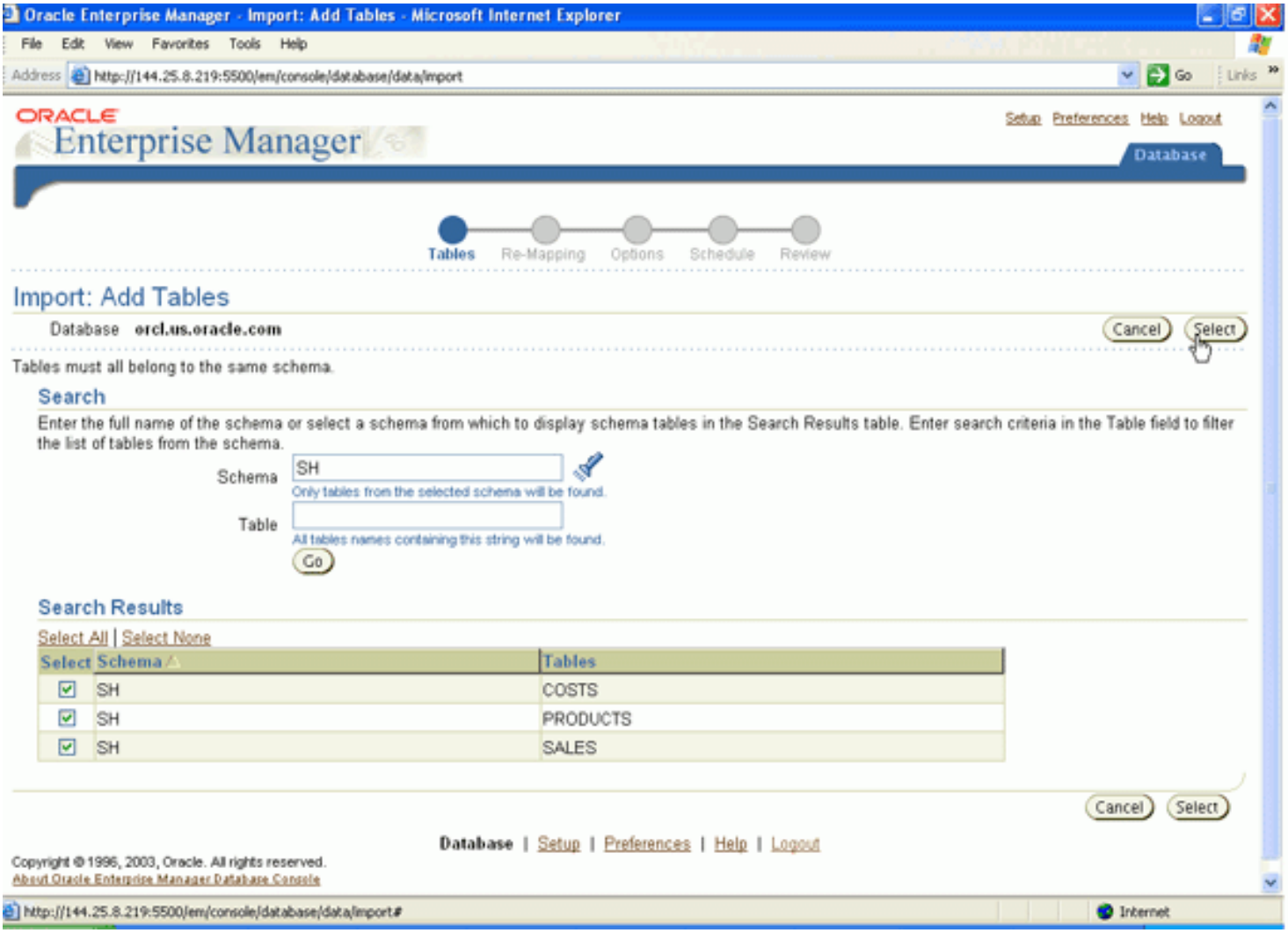
☐ **4.** Set the **Import Type** to **Tables** and enter the Host username and password, then click **Continue** .



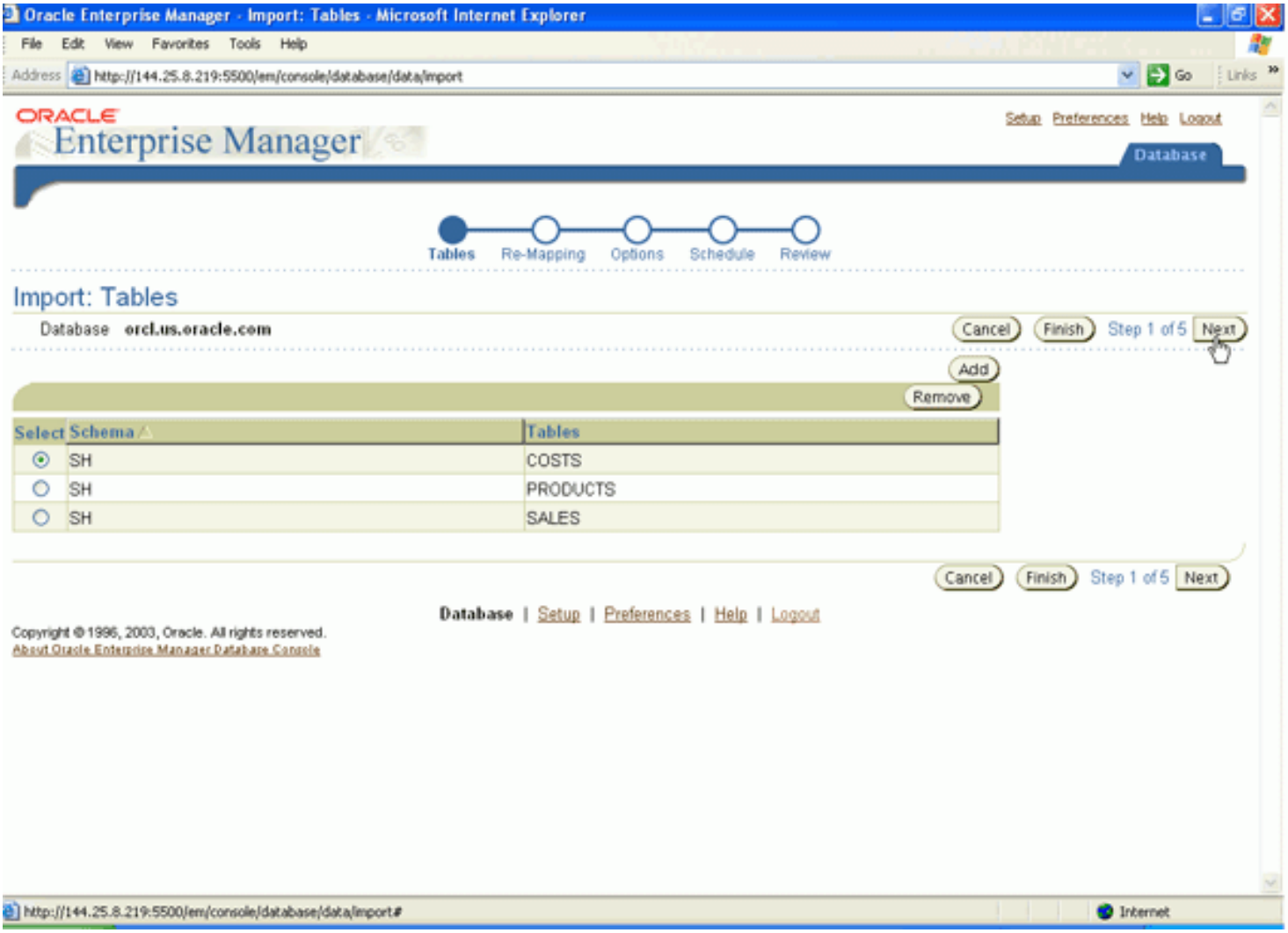
☐ **5.** Select the **Add** button to see the tables available for import.

☐ **6.** Enter **SH** in the Schema field and click **Go** .

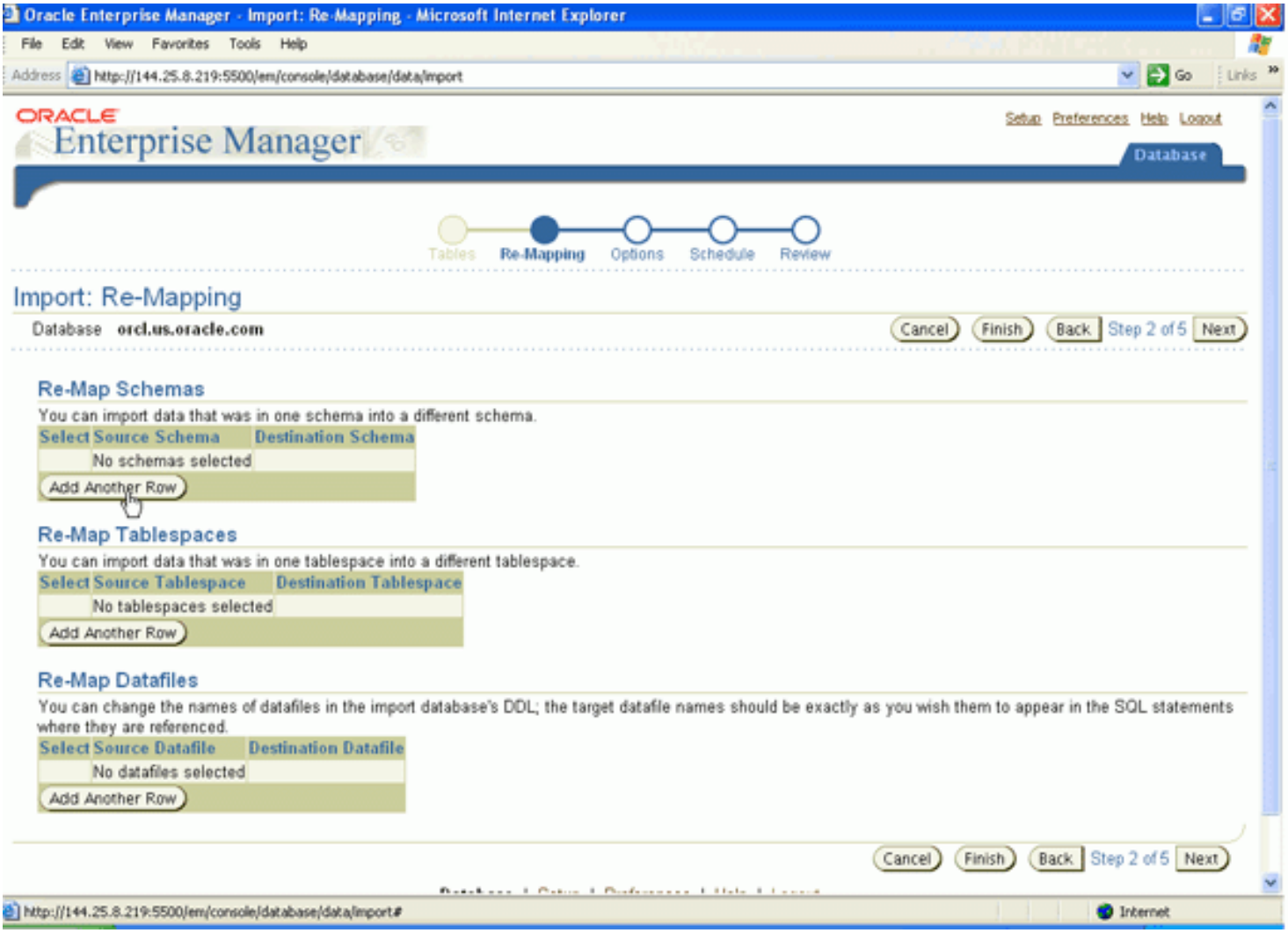
☐ **7.** Select the checkboxes for **Costs, Products,** and **Sales** then click **Select** .


Oracle Enterprise Manager - Import: Add Tables - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address http://144.25.8.219:5500/em/console/database/data/import
Go
Links
ORACLE
Enterprise Manager
Setup Preferences Help Logout
Database
Tables
Re-Mapping
Options
Schedule
Review
Import: Add Tables
Database orcl.us.oracle.com
Cancel
Select
Tables must all belong to the same schema.
Search
Enter the full name of the schema or select a schema from which to display schema tables in the Search Results table. Enter search criteria in the Table field to filter the list of tables from the schema.
Schema
SH
Only tables from the selected schema will be found.
Table
All tables names containing this string will be found.
Go
Search Results
Select All | Select None
Select
Schema
Tables
SH
COSTS
SH
PRODUCTS
SH
SALES
Cancel
Select
Database | Setup | Preferences | Help | Logout
Copyright © 1996, 2003, Oracle. All rights reserved.
About Oracle Enterprise Manager Database Console
http://144.25.8.219:5500/em/console/database/data/import#
Internet

☐ **8.** Click **Next** .



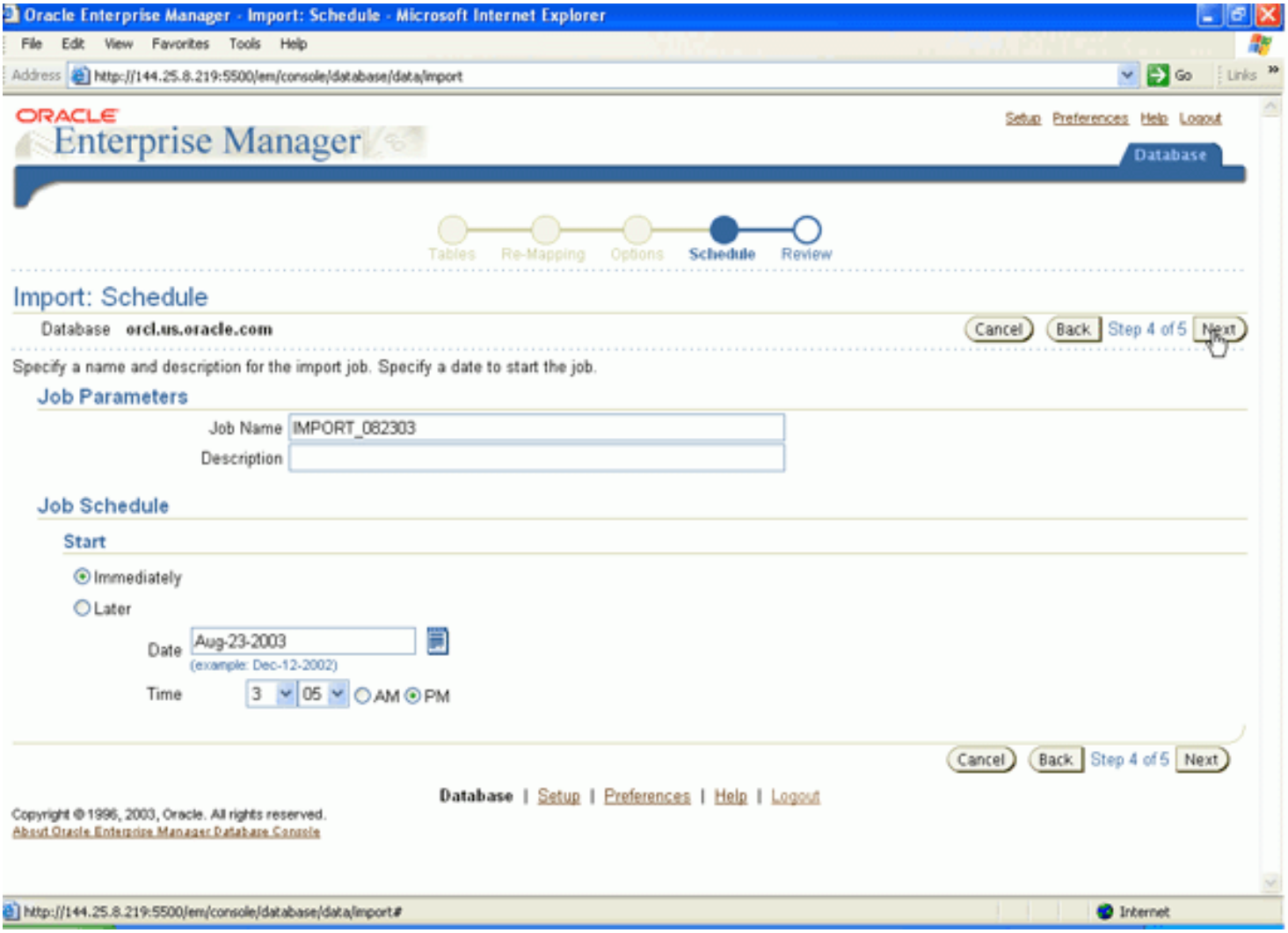
☐ **9.** Click the **Add Another Row** button under the **Re-Map Schemas** section.



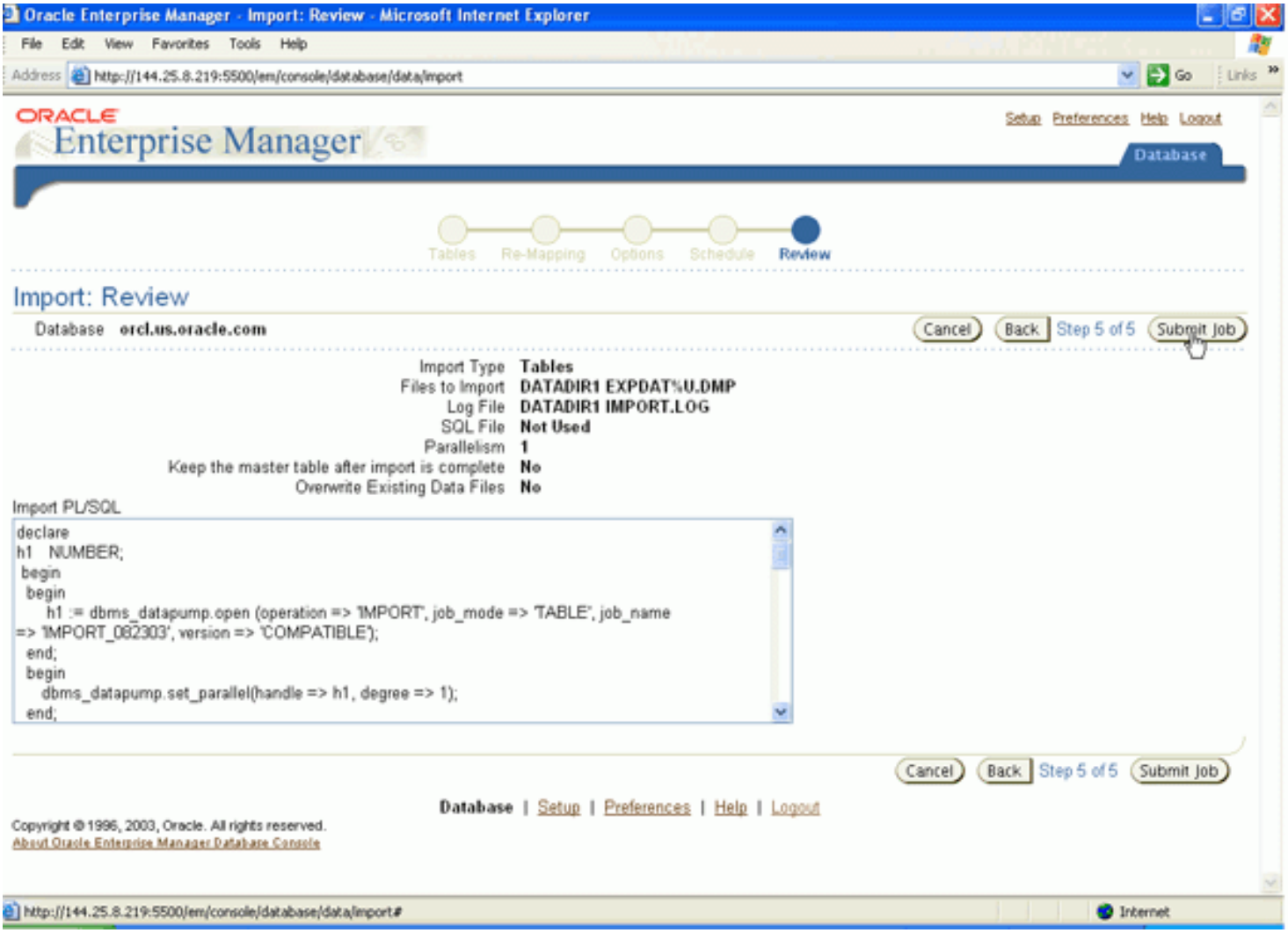
10. ☐ Under the Destination Schema column, select **SCOTT** then click **Next** .

11. ☐ Click **Next** .

12. ☐ Enter **IMPORT_<Today's Date>** in the Job field then click **Next** .



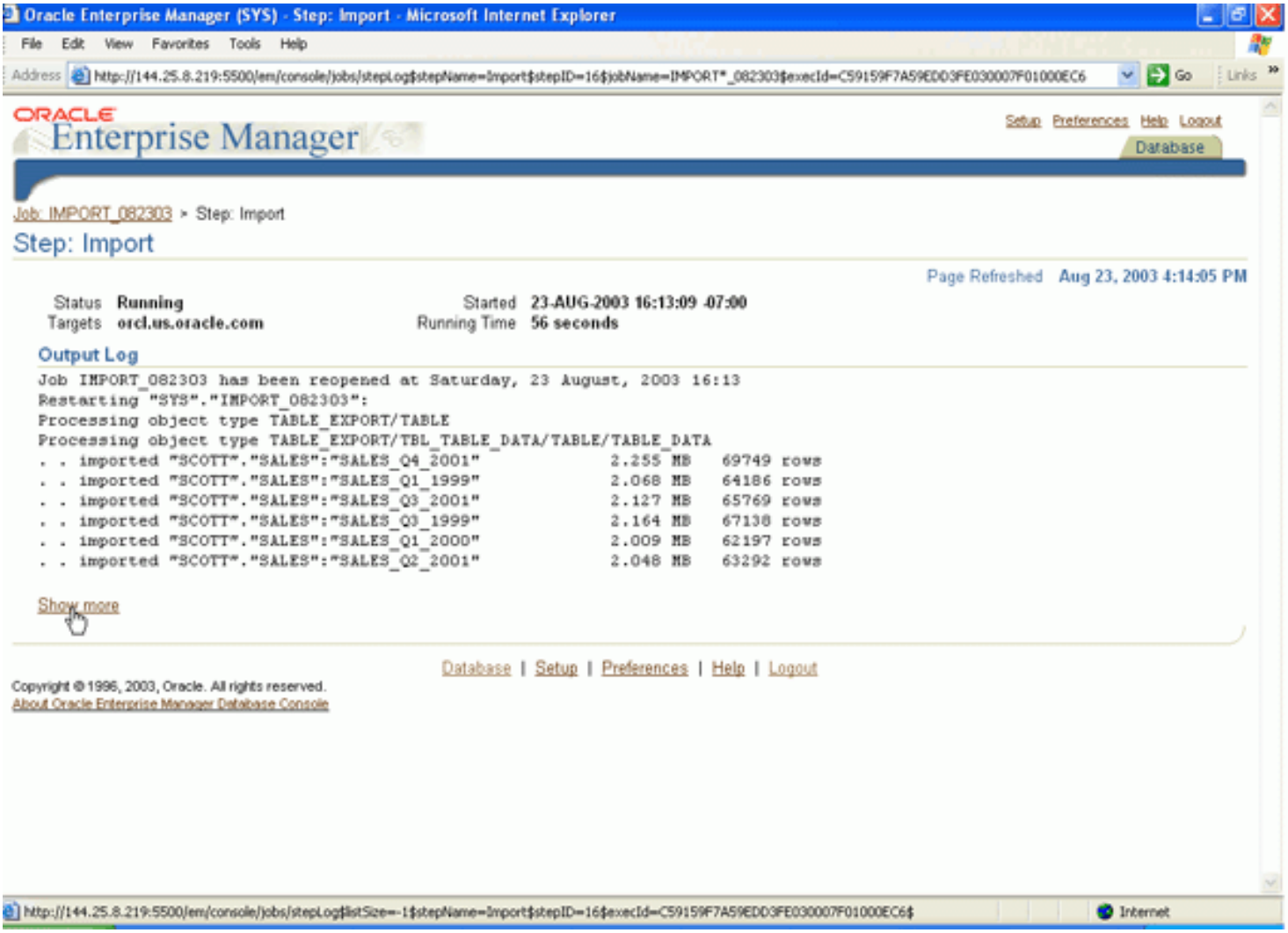
13. ☐ Click **Submit Job** .



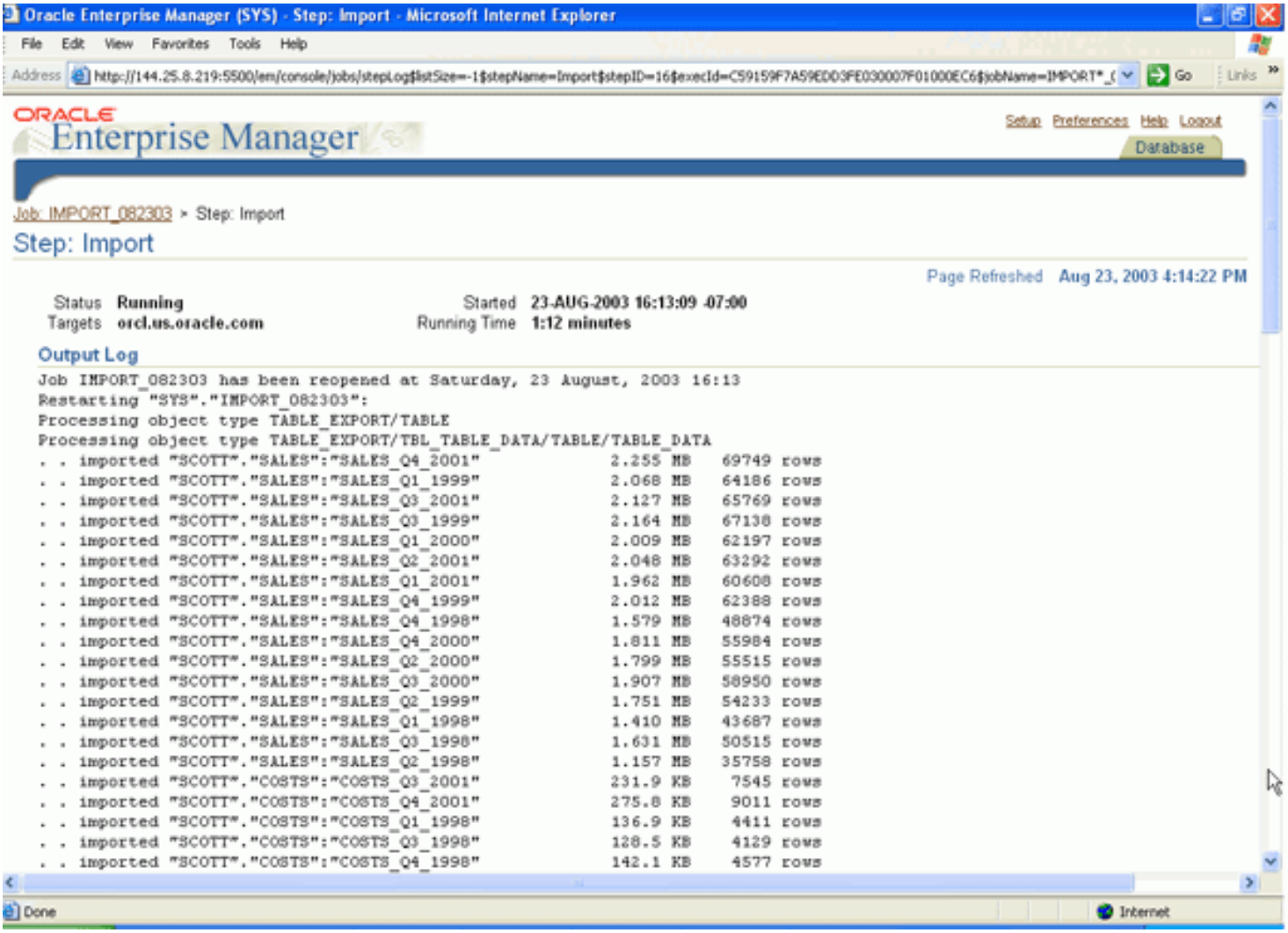
14. ☐ Click **View Job** .

15. ☐ Select the **Import** link to see the Import log status.

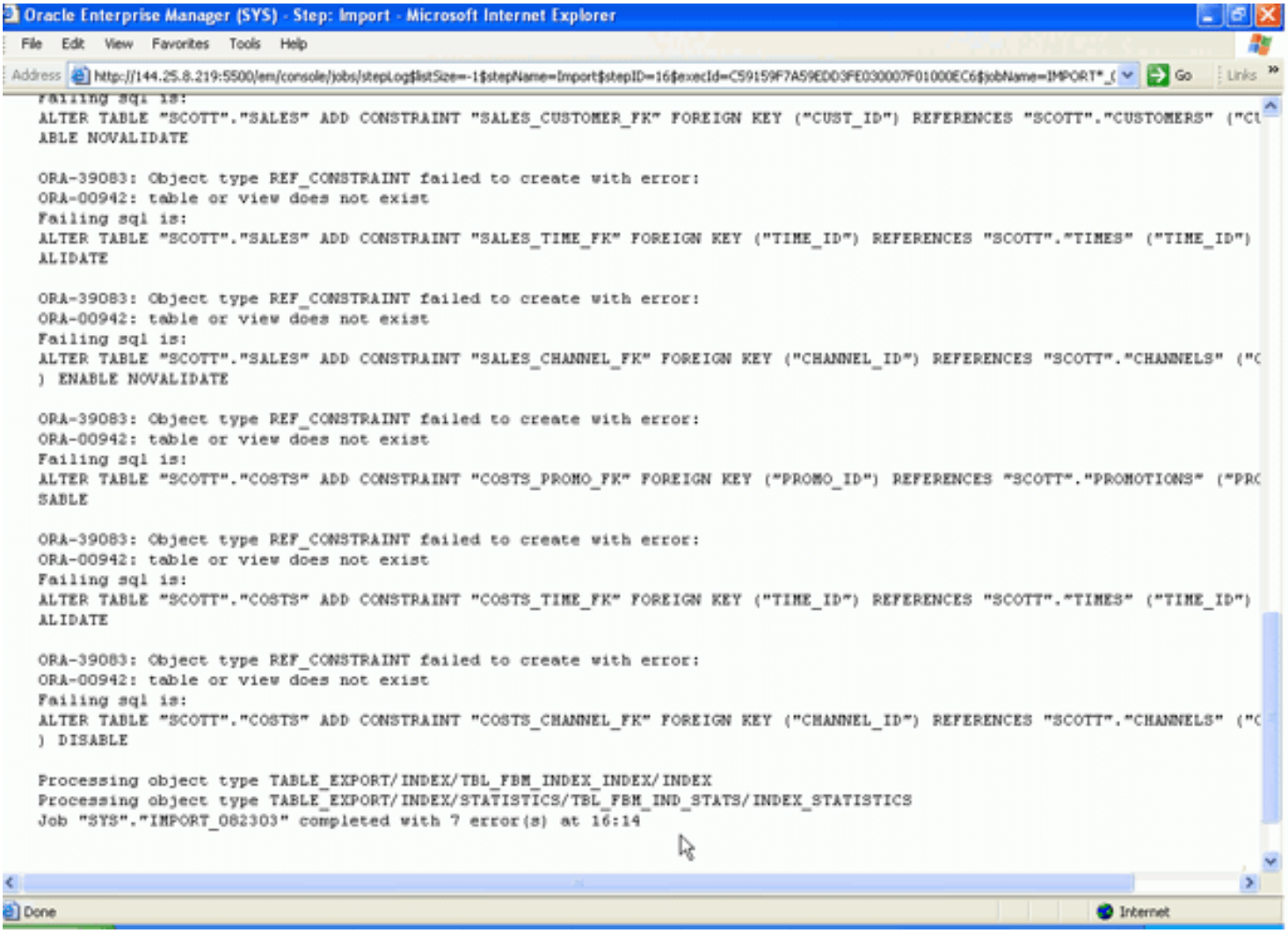
☐ **16.** The job is still running. Click **Show more** to see more of the log. If Show more does not appear, click Reload in your browser window.



17. ☐ The job has finished. Scroll down to the bottom to see all the messages in the log.



18. Your import has completed successfully even though the log file displayed some errors. These errors were generated because the **Sales** and **Costs** tables are dependent on several tables which were not included in the export and thus were not imported. In this case, the omission was intentional and you can ignore the errors in the log file.



You will go through the following examples of using Data Pump Import command line interface:

- Performing a data only table mode import
- Performing a schema mode import

Performing a Data Only Table Mode Import

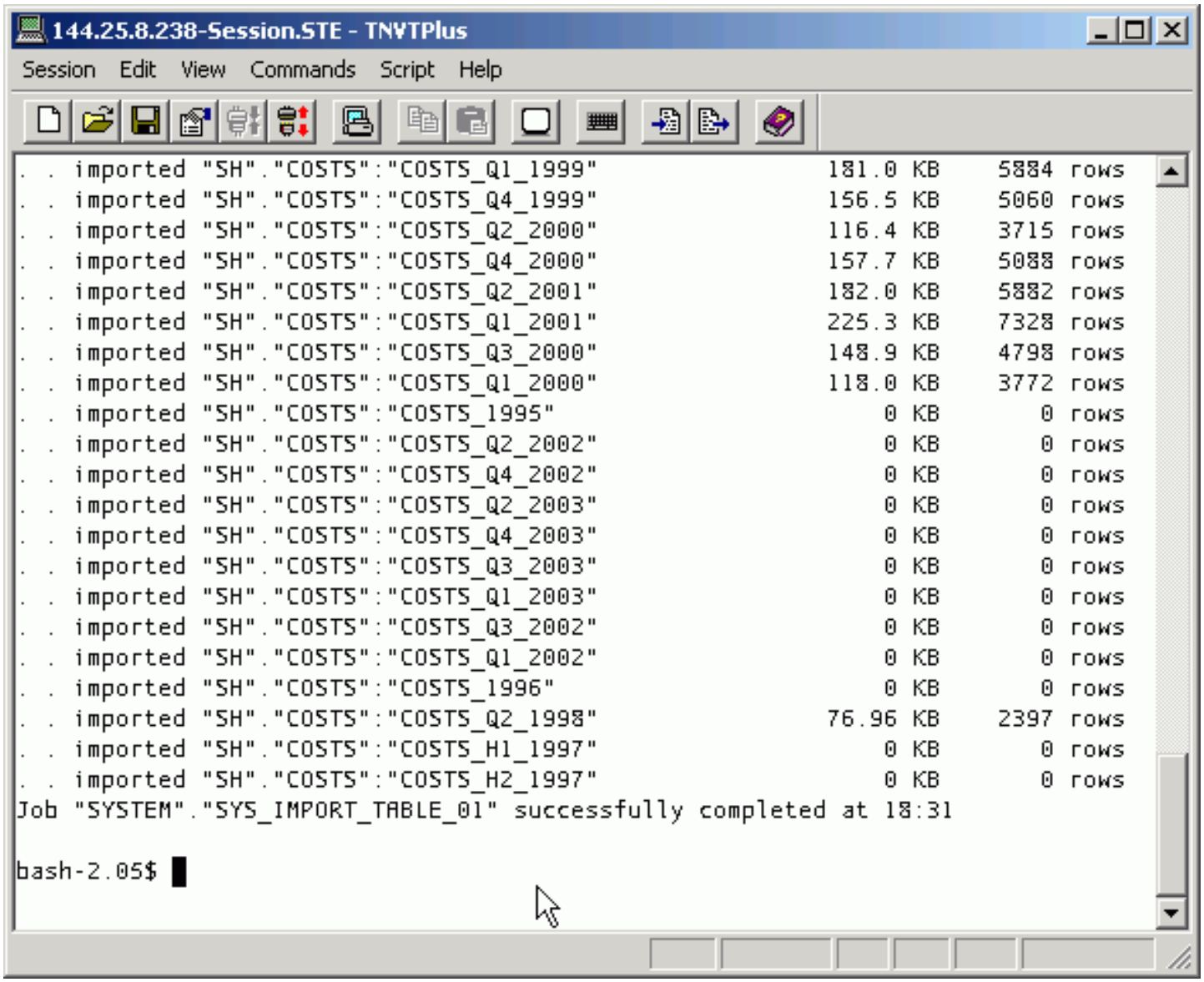
Back to List

The CONTENT parameter allows you to filter the data and metadata that Import loads. The DATA_ONLY value loads only table row data; no database object definitions (metadata) are recreated.

Perform the following:

☐ **1.** From a terminal window, issue the following IMPORT command to perform a table data only import of table Costs using the dump file created previously in the Export section of this lesson.

```
impdp system/<password> \

TABLES=sh.costs \

CONTENT=data_only \

DUMPFILE=datadir2:table.dmp \

NOLOGFILE=y
```



Performing a Schema Mode Import

Back to List

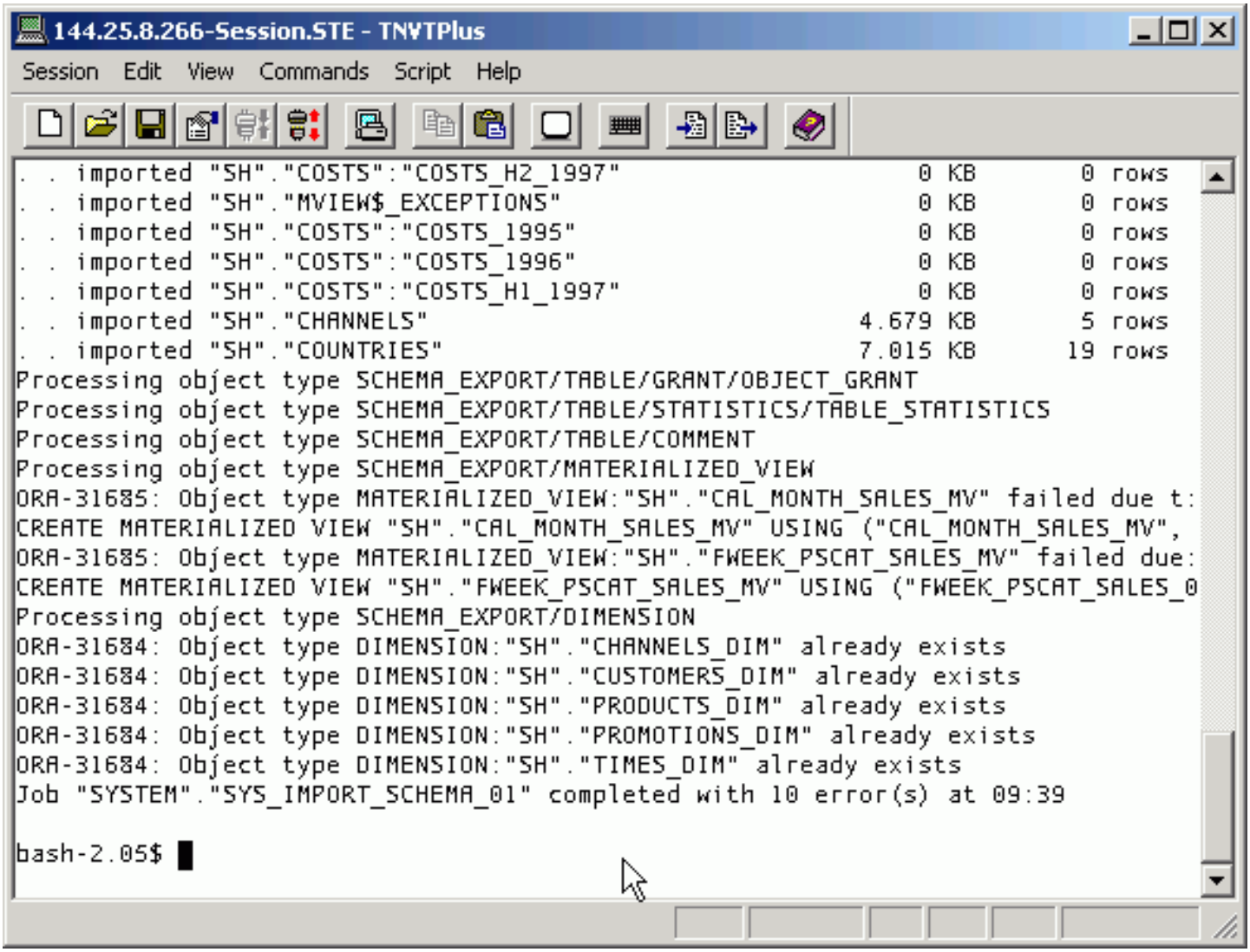
The EXCLUDE parameter allows you to filter the metadata that is imported by specifying database objects that you want to exclude from the import job. For the given mode of import, all the objects contained within the source, and all their dependent objects, are included except those specified in an EXCLUDE statement. If an object is excluded, all of its dependent objects are also excluded.

TABLE_EXISTS_ACTION instructs import on what to do if the table it is trying to create already exists. When TABLE_EXISTS_ACTION=REPLACE is specified, the import drops the existing table and then recreates and loads it using the source database contents.

Perform the following:

☐ **1.** From your terminal window, issue the following import command to perform a schema import that excludes constraints, referential constraints, indexes and materialized views using the dump file set created by the schema mode export in the Export section.

```
impdp system/<password> \
SCHEMAS=sh \
REMAP_SCHEMA=sh:sh2 \
DUMPFILE=datadir1:schema1%U.dmp,datadir2:schema2%U.dmp \

EXCLUDE=constraint, ref_constraint, index,materialized_view \
TABLE_EXISTS_ACTION=replace \

logfile=datadir1:impschema.log
```



Move your mouse over this icon to hide all screenshot