Participant Monthly DVD User Guide

An introductory guide

User's Guide for version 1.0 of the Participant Monthly DVD

Version 5.10 released on 01 Jul 2009

Prelude

In this document, references to the 'software' are to the Participant Monthly DVD software, unless specifically otherwise. References to associated materials include this document, *Participant Monthly DVD User Guide*.

Disclaimer

1. This software and the associated materials are provided by AEMO to users free of charge to interface to the AEMO MMS.

2. A user runs the software or uses the materials at its own risk. AEMO does not represent or warrant that the software and materials are complete, free from errors, reliable and fit for the purpose set out above. Nor does AEMO represent or warrant that the software is free from viruses and other programs which may affect other software or systems. The software and associated materials are provided on the basis that users proposing to use or rely on the software or associated materials undertake responsibility for independently verifying and checking the accuracy, completeness, reliability or suitability of the software or associated materials.

3. Subject to paragraph 4, AEMO and its employees and contractors are not liable for any loss, damage, costs or expenses suffered or incurred by users in connection with the software or materials or their use of them (including loss, damage, costs or expenses caused by negligence).

4. The Trade Practices Act and some other laws imply terms into contracts for the supply of goods or services ("Prescribed Terms"). To the extent permitted by law, the liability of AEMO in respect of a breach of a Prescribed Term relating to the supply of the software or the materials or any associated service is limited at AEMO's sole discretion to:

   - in the case of a good, any one or more of the following:
     - its replacement or the supply of an equivalent good or payment of the cost of such replacement or supply; or
     - its repair or payment of the cost of its repair; and
   - in the case of a service, the re-supply of the service or payment of the cost of doing so.

5. Except as outlined in this disclaimer, to the extent permitted by law, all conditions, warranties and terms (including Prescribed Terms) relating to the software and materials (whether written, oral or implied by law, custom or usage) are excluded.

Copyright

Distribution

Commercial-in-confidence with participants

Document Meta-information

Title: Participant Monthly DVD User Guide

Document ID:

Version: 5.10

Responsible Department: MMS

Notes: First publication

Version History

Ver.Rel	Date	Author	Notes
1.00	1Jul05	MMS	First Release
2.00	11Aug06	MMS	Change date format to ENDDATE
3.00	3Sep06	MMS	Change date format to STARTDATE
4.00	20Sep07	MMS	Add BCP support
5.00	06Feb08	MMS	Add BCPT utility
5.10	01Jul09	MMS	AEMO branding and updated to remove Replica Server references

Contents

Participant Monthly DVD User Guide   Page 5
Version 5.10
Participant_Monthly_DVD.doc (saved on 17 July 2009)

Participant Monthly DVD User Guide   Page 6
Version 5.10
Participant_Monthly_DVD.doc (saved on 17 July 2009)

Introduction

Overview

AEMO distributes a monthly archive DVD to provide participants with an historical source of a subset of the MMS Data Model data.

Who is this document for?

This document (*Participant Monthly DVD User Guide*) is for participants in the National Electricity Market who use the DVD to load historical data, use the monthly DVD as an archive of data or use this data as part of their disaster recovery process.

What this user manual covers

This document (*Participant Monthly DVD User Guide*) describes the directory layout, describes the file formats, and provides a description of the steps needed to load the data on the DVD into an Oracle database.

This document starts with a description of the contents of the DVD followed by file structures and formats, how to load the data into the target database and a description of the utilities provided.

Assumed Knowledge

Please note that any loading of data involves risks, including the potential for performance impact on yopur data interchange systems. Any loading of data is best done by an experienced DBA with the necessary business knowledge of your system.

Glossary

Term	Explanation
CSV	Comma separated values
DBA	Data Base Administrator; a knowledgeable person for maintaining your database system
Helpdesk	AEMO's Helpdesk; phone 1300 300 295
MMS	Market Management System; software, hardware, network and related processes to implement the National Electricity Market (NEM); an AEMO department responsible for maintaining the system
NEMNET	a private computer network containing participants and AEMO
NER	National Electricity Rules; sometimes just called the Rules
PDR Batcher /Loader	A pair of AEMO application providing an interface to MMS to collect CSV files published to a participant's directory on the participant fileserver and then parse those files into a local MMS Data Model database instance
Rules	National Electricity Rules (NER)
ZIP	A compressed file, containing one or more files

Table 1: Glossary of terms

Requirements

A MMS Data Model compliant database is a key requirement. An Oracle database can receive all the data on the DVD, and all other databases can receive most of the data.

Resources

The following resources may be useful in addition to the information contained in this manual:

MMS Data Model Report	Describes the table layout, which matches the data file format on the DVD
Replication Manager	An AEMO-supplied user interface to manage the replication of data to your MMS Data Model database
PDR batcher and Loader	AEMO-supplied products for replicating data using CSV files

Outline of functions

Due to size and performance considerations, AEMO regularly purges older data from the online data interfaces with participants. The monthly DVD provides a mechanism to transfer large volumes of historical AEMO data (at least, the public data).

If Participants allow their database to grow too large, it complicates day-to-day management, affects performance and increases recovery times. The monthly DVD provides a longer-term storage of AEMO public data.

The Participant monthly DVD contains data extracts of the previous month's data for a subset of tables. The Index file on the DVD provides a list of the tables included.

DVD Contents

Directories

The files that AEMO provides can include confidential as well as public information. For ease of generation and management, the top level directory includes the visibility as part of its name. The Visibility is either PUBLIC or a specific Participant Id.

The topmost directory on the DVD has a pattern based on the visibility and the period of the data, being *<VISIBILITY ID>_<YYYMMY>_archive* (e.g. *PUBLIC_200604_archive*). Within the topmost directory are the following directories:

Directory name	Contents
DATA	Zipped CSV files for the month *<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat* script Oracle export file for the MARKETNOTICEDATA tables
CTL	SQL*Loader control files
BCP_DATA	*<VISIBILITYID>_CONVERT_BCP_<YYYYMM>.bat* script *<VISIBILITYID>_RUN_BCP_<YYYYMM>.bat* script
BCP_FMT	SQL Server BCP format files
INDEX	the *<VISIBILITYID>_monthly_cd_index_<YYYYMM>.txt* file
DOCUMENTATION	All participant documentation
LOGS	Empty, this directory is for the SQL*Loader log files when the *<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat* script is run And the SQL Server BCP log files when the *<VISIBILITYID>_RUN_BCP_<YYYYMM>.bat* script is run
UTILITIES	The *CRLF.txt* and *CRLF.exe* files for the conversion of CSV files to a UNIX format The *BCPT.EXE* and *BCPT.INI* files for the conversion of CSV files to BCP loadable files

DATA files

The data files are CSV (comma separated values) files.

File Naming Convention

The name of each CSV file contains the VISIBILITYID and a FILEID plus a date in the form of YYYYMMHHMM. The VISIBILITYID represents the confidentiality of the data on the DVD and will be either PUBLIC or a specific Participant Id. The FILEID is the main identifier of each datafile and is the MMS Data Model table name unless there is more than one data file for a table.

The monthly report is generated using a start and end date with the start date inclusive and the end date exclusive. The YYYYMMHHMM component of the CSV filename reflects the start date and time used for the data extract. Therefore, the CSV files in the APR2005 directory will have a date format of 200504010000 in the file name.

Some tables have more than one data file. There are many reasons for multiple files for a table but the most common reasons are:

- The data set is too large for one file (e.g. PREDISPATCHLOAD1 contains data for the first half of the month and PREDISPATCHLOAD2 contains data for the last half of the month)

- The data is not the complete data set from the source table (e.g. PREDISPATCHLOAD_D contains summary data from the PREDISPATCHLOAD table)

When tables have more than one data file, the FILEID is as informative as possible as to the contents of the file.

The file naming convention is *<VISIBILITYID>_DVD_<FILEID>_ <YYYYMMHHMM>.zip*.

An example of a file name is PUBLIC_*DVD_BIDDAYOFFER_200504010000.zip*.

CSV file Format

Record Types

The input files contain three record types:

- C: A comment record

- I: An information record defining the column names of data to follow

- D: A data record containing data values matching the column names in the preceding "I" record

The key data in the first "C" record includes the AEMO Source Database, the File Creation Date and the File Creation Time.

The key data contained in the "I" and "D" records includes the Report Name, the Report Sub-Type, the Report Version and the Data Column Names.

MMS Data Model compatibility

The columns extracted in the CSV file and the syntax in the supporting SQL*Loader control file are specific to the version of the MMS Data Model as it was on the last day of the month of the data extract (e.g. \PUBLIC_200401_archive data corresponds to the MMS Data Model as it was on 31st January 2004).

The loading process on the DVD ignores the lines in the CSV file starting with C and I, and ignores the first 3 columns of the D line. The loading process assumes that the table you are loading conforms to the correct MMS Data Model structure. Participants using any other loading method (either for Oracle or for another database) must include column 4 and the following columns only from D records when loading the data.

Local Database

TIP

The DVD is created in a Windows environment. CSV file conversion may be necessary if loading on a UNIX environment (for more details, refer to Utility Files *on page 22).*

CTL Files

The SQL*Loader control files are named according to the CSV files they support with the same VISIBILITYID and FILEID as those files using the format
<VISIBILITYID>_DVD_<FILEID>_<YYYYMM>.ctl

An example SQL*Loader control file name is PUBLIC_DVD_BIDDAYOFFER_200412.ctl.

For an example of the contents of a supplied control file, refer to Appendix_1_Example

Oracle Export File (MARKETNOTICEDATA)

Due to problems with extracting the MARKETNOTICEDATA table into a CSV format, the table data has been extracted using the Oracle Export Utility.

Users of non-Oracle systems who wish to retrieve this data can contact the AEMO Helpdesk on 1300 300 295.

BCP_DATA Files

The BCP_DATA directory does not contain any datafiles. The files are produced from the CSV files in the DATA directory when the participant runs the *VISIBILITYID>_CONVERT_<YYYYMM>.bat* file. Once the CSV files are converted into the BCP_DATA directory, the data files are separated with ~|.

File Naming Convention

The name of each CSV file contains the VISIBILITYID and a FILEID plus a date in the form of YYYYMMHHMM. The VISIBILITYID represents the confidentiality of the data on the DVD and will be either PUBLIC or a specific Participant Id. The FILEID is the main identifier of each datafile and is the MMS Data Model table name unless there is more than one data file for a table.

The monthly report is generated using a start and end date with the start date inclusive and the end date exclusive. The YYYYMMHHMM component of the BCP filename reflects the start date and time used for the data extract. Therefore, the BCP files in the APR2005 directory will have a date format of 200504010000 in the file name.

Some tables have more than one data file. There are many reasons for multiple files for a table but the most common reasons are:

- The data set is too large for one file (e.g. PREDISPATCHLOAD1 contains data for the first half of the month and PREDISPATCHLOAD2 contains data for the last half of the month)

- The data is not the complete data set from the source table (e.g. PREDISPATCHLOAD_D contains summary data from the PREDISPATCHLOAD table)

When tables have more than one data file, the FILEID is as informative as possible as to the contents of the file.

The file naming convention is *<VISIBILITYID>_DVD_<FILEID>_<YYYYMMHHMM>.zip*.

An example of a file name is PUBLIC_*DVD_BIDDAYOFFER_200504010000.zip*.

BCP file Format

Record Types

The input files contain three record types:

- C: A comment record

- I: An information record defining the column names of data to follow

- D: A data record containing data values matching the column names in the preceding "I" record

The key data in the first "C" record includes the AEMO Source Database, the File Creation Date and the File Creation Time.

The key data contained in the "I" and "D" records includes the Report Name, the Report Sub-Type, the Report Version and the Data Column Names.

MMS Data Model compatibility

The columns extracted in the CSV file and the syntax in the supporting SQL Server BCP format file are specific to the version of the MMS Data Model as it was on the last day of the month of the data extract (e.g. *\PUBLIC_200401_archive* data corresponds to the MMS Data Model as it was on 31st January 2004).

The loading process on the DVD ignores the lines in the CSV file starting with C and I, and ignores the first 3 columns of the D line. The loading process assumes that the table you are loading conforms to the correct MMS Data Model structure. Participants using any other loading method (either for Oracle or for another database) must include column 4 and the following columns only from D records when loading the data.

The *BCPT.INI* file contains any rounding directives required for the *BCPT.EXE* utility to convert the files into the BCP_DATA directory. The contents of the BCPT.INI file are specific to the MMS Data Model version at the time the DVD was produced.

BCP_FMT Files

The BCP format files are named according to the BCP files they support (BCP_DATA) with the same VISIBILITYID and FILEID as those files using the format *<VISIBILITYID>*_DVD_*<FILEID>*_*<YYYMM>*.fmt

An example BCP format file name is PUBLIC_DVD_BIDDAYOFFER_200412.fmt

For an example of the contents of a supplied control file, refer to Appendix_2_Example

Documentation Files

Any documentation files relating to the Monthly DVD release (e.g. *Participant_Monthly_DVD.doc*).

Index Files

There is one Index file, named *<VISIBILITYID>_monthly_dvd_index_<YYYYMM>.txt*.

Utility Files

Any utility files relating to the Monthly DVD release (e.g. *Crlf.exe, bcpt.exe*).

Loading the Data

There are three types of extracts provided on the DVD.

The files in the DATA directory are in CSV format using the ',' separator and can be loaded into any database that supports the loading of CSV files.

The files in the DATA directory can be converted with the *BCPT.EXE* utility and placed in the BCP_DATA directory with the separator '~|' and can be loaded into any database that supports the loading of CSV files where a separator of 2 characters is supported. The BCPT.EXE utility also removes any quotation marks around text fields.

A small number of tables are currently only provided in an Oracle export file. These tables are only available for loading into an Oracle database as the file format can only be interpreted by the Oracle import utility.

Oracle SQL*Loader

This document is not a substitute for Oracle SQL*Loader documentation and it is intended that this document should always be read in conjunction Oracle's official documentation specific to your version of the database.

> **TIP**
>
> *Whilst this document aims to assist in the loading of data into your database, it is always recommended that any data load be performed by an experienced DBA as any problems can have on-going ramifications.*

SQL*Loader requires a CTL file in order to load the data into an Oracle database. The monthly DVD supplies a CTL file to support the loading of each datafile via SQL*Loader.

The CTL file appends data into the destination table and loads only rows beginning with D (data rows) from field 4 onwards (the first 3 fields are AEMO control fields). An example of a supplied CTL file is in *Appendix 1: Example SQL*Loader CTL file* on page 25.

> **TIP**
>
> *The CTL files only support loading data into a table you own. The table name is not prefixed with any Oracle schema name. This is because the destination schema name is not known. It is therefore necessary to login as the table owner when using these CTL files.*

This section covers two methods for loading the data, being

- using the supplied SQL*Loader Wrapper script (refer to *Using the supplied SQL*Loader Wrapper script* below)

- manually loading the data via SQL*Loader using either conventional path load or direct path load (refer to *SQL*Loader command line* on page 16)

Using the supplied SQL*Loader Wrapper script

AEMO provides the SQL*Loader wrapper script on the Participant Monthly DVD. Please note the following when using this script:

- The SQL*Loader wrapper script is intended to aid the loading of all of the supplied data files. If you require a subset of the data, either edit the wrapper script to remove the entries for table you do not wish to load or load each table manually (For manual loading, refer to *SQL*Loader command line* on page 16).

- The wrapper script defaults to loading the tables via the conventional path (DIRECTFLG=N). It is possible to use DIRECT load by setting the DIRECTFLG=Y.

- The login details you supply to this script must be the details of the destination table owner. The CTL files do not contain a schema name prefix since the schema name is not known (i.e. it could be different from one database to another).

- The script loads the tables in the order they appear in the wrapper script. It is therefore necessary to disable any Foreign Keys on the destination tables/schema and re-enable them after the data is loaded.

> **TIP**
>
> *The MMS Data Model does not have any foreign keys defined. If your systems have not implemented any foreign keys, you can omit any steps referencing foreign keys.*

Instructions

1. Load the monthly DVD to a drive or copy the files to a directory.

2. Unzip the data files.

3. If you have implemented any foreign keys, disable the Foreign Key constraints in the target schema (an example script to disable all foreign keys for a schema is in *Appendix 2: Example Foreign key disable/enable script* on page 27).

> **TIP**
>
> *Generate the enable Foreign Keys script in the same step as the disable so you can make sure to only enable the constraints originally enabled in your environment.*

4. Edit the *<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat* script, saving the script to a directory with write access in order to do this.

5. Locate the following section of the script, change where necessary and save the file:

```
REM * Variables default to current directory, change if necessary
REM * Syntax of running file: <PATH>\DATA\<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat username
password dbname TOP_INFO_ARCHIVEdir
REM * e.g.: <PATH>\DATA\PUBLIC_RUN_SQLLDR_200412.bat infoserver infoserver partdb
e:\dvd\test\info_archive
SET CTLDIR=%TOP_INFO_ARCHIVE%\<VISIBILITYID>_DEC2004_archive\CTL
SET LOGDIR=%TOP_INFO_ARCHIVE%\<VISIBILITYID>_DEC2004_archive\LOGS
SET DISDIR=%TOP_INFO_ARCHIVE%\<VISIBILITYID>_DEC2004_archive\LOGS
SET BADDIR=%TOP_INFO_ARCHIVE%\<VISIBILITYID>_DEC2004_archive\LOGS
SET BINDSIZE=1000000
SET READSIZE=1000000
SET DIRECTFLG=N
```

The default settings for the locations can be seen above. If they are left unaltered, the script assumes the directories in question are under the *%TOP_INFO_ARCHIVE%* directory (passed as a parameter on the command line when running the script). The script also assumes that it is possible to write files to the *LOGDIR*, *DISDIR* and *BADDIR* directories.

Change the above values to whatever you wish, but make sure that you pre-create the directories if they do not already exist.

6. On the command line, set the working directory (using CD) to the directory containing the *<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat* script and run the script.

For example, if the files were saved from the DVD to the *c:\monthlydvd* directory and the month being loaded was December 2004 with a visibility id of PUBLIC:

```
CMD> cd c:\monthlydvd
CMD> cd PUBLIC_DEC2004_archive\data
```

```
CMD> public_run_sqlldr_200412.bat username password tnsalias c:\monthlydvd
```

7. Check the logfiles produced in the location specified by the LOGDIR entry as set in the *<VISIBILITYID>_RUN_SQLLDR_YYYYMM.bat* for any errors.

8. If you disabled any foreign keys in earlier steps, enable the Foreign Keys in the target schema (an example script to enable the foreign keys for a schema is in *Appendix 2: Example Foreign key disable/enable script* on page 27).

> **TIP**
>
> *As in the example, it is better to generate the enable script at the time you disable the constraints to ensure you do not enable any constraints supposed to stay disabled in your environment.*

SQL*Loader command line

Conventional Path

The conventional path load reads the CSV file, creates an insert statement, then applies the insert to the target database. Conventional path loads are slower than direct path loads but have the advantage of applying all triggers and constraints on the target table. Conventional path loading also allows concurrent access to the target object by other users and applications. Conventional path loading is the default for SQL*Loader.

The following instructions are to load the CSV files into an Oracle database on a Windows platform:

1. Unzip the zipped CSV file

2. Run the SQL*Loader utility as follows:

```
sqlldr <username>/<password>@<tnsalias> control=<controlfilename> log=<logfilename>
bad=<badfilename> discard=<disfilename>

e.g. sqlldr scott/tiger@ORCL control=<PATH>\<VISIBILITYID>_DVD_BIDDAYOFFER_200412.ctl
log=<PATH>\<VISIBILITYID>_dvd_biddayoffer.log
bad=<PATH>\<VISIBILITYID>_dvd_biddayoffer.bad
discard=<PATH>\<VISIBILITYID>_dvd_biddayoffer.dis
```

3. Check the log files for any errors

4. Check the discard file for any records discarded due to failing the WHEN clause. Each discard file is expected to have 3 rows beginning with C, I and C respectively – these are comment rows and need to be discarded. Any rows beginning with D are data rows and are expected to load successfully

5. Check the bad file for any rows not able to be loaded. Unsuccessful loading of a row is due usually to a column data type mismatch or number of columns mismatch between the data in the CSV file and the structure of the table in your target database

Direct Path

A faster way to load large amounts of data is to use a direct path load. Please note that there are many considerations to take into account before choosing to use direct path load.

> **TIP**
>
> *Refer to the Oracle documentation for full instructions and advice for using direct path loads.*

SQL*Loader hints

It can be faster to load data if you drop the indexes before the load and re-create them after the load has finished.

Faster index creation can be achieved by creating indexes NOLOGGING and with the PARALLEL clause.

Be aware of your rollback segment size and the BUFFERSIZE value you use. It is possible to change the BUFFERSIZE and READSIZE values in the *<VISIBILITYID>_RUN_SQLLDR_<YYYYMM>.bat* script to improve performance according to your environment.

Oracle Import Utility

The table MARKETNOTICEDATA is provided in an Oracle export file.

These files can be imported using the following commands:

```
imp <username>/<passwd> FILE=<VISIBILITYID>_MARKETNOTICEDATA_<YYYYMM>.dmp
```

SQL Server BCP

This document is not a substitute for BCP documentation and it is intended that this document should always be read in conjunction with Microsoft's official documentation specific to your version of the database.

> **TIP**
>
> *Whilst this document aims to assist in the loading of data into your database, it is always recommended that any data load be performed by an experienced DBA as any problems can have on-going ramifications.*

BCP requires a format file in order to load the data into a SQL Server database. The monthly DVD supplies a format file to support the loading of each datafile via BCP.

The BCP file appends data into the destination table and loads all except the first two rows from field 4 onwards (the first 3 fields are AEMO control fields). An example of a supplied BCVP file is in *Appendix 1: Example SQL*Loader CTL file*.

The AEMO supplied CSV files require converting into a different format for loading with the BCP utility. The *BCPT.EXE* utility is provided to convert the files.

This section covers two methods for loading the data, being

- using the supplied BCP Wrapper scripts (refer to Using the supplied BCP Wrapper scripts below)

- manually loading the data via BCP command line

Using the supplied BCP Wrapper scripts

AEMO provides the BCP wrapper scripts on the Participant Monthly DVD. Please note the following when using these scripts:

- The BCP wrapper scripts are intended to aid the loading of all of the supplied data files. If you require a subset of the data, either edit the wrapper script to remove the entries for tables you do not wish to load or load each table manually (For manual loading, refer to *BCP command Line*).

- The login details you supply to the load script must be the details of the destination table owner. The BCP wrapper file does not contain a schema name prefix since the schema name is not known (i.e. it could be different from one database to another).

- The script loads the tables in the order they appear in the wrapper script. It is therefore necessary to disable any Foreign Keys on the destination tables/schema and re-enable them after the data is loaded.

Instructions

1. Load the monthly DVD to a drive or copy the files to a directory.

2. If you have implemented any foreign keys, disable the Foreign Key constraints in the target schema

> **TIP**
>
> *Generate the enable Foreign Keys script in the same step as the disable so you can make sure to only enable the constraints originally enabled in your environment.*

3. Unzip the CSV files in the DATA directory.

4. Edit the *<VISIBILITYID>_RUN_CONVERT_<YYYYMM>.bat* script, saving the script to a directory with write access in order to do this.

5. Locate the following section of the script, change where necessary and save the file:

```
REM * Variables default to current directory, change if necessary
REM * Syntax of running file: <PATH>\BCP_DATA\PUBLIC_CONVERT_BCP_200708.bat
TOP_INFO_ARCHIVE dir
REM * e.g.:<PATH>\BCP_DATA\PUBLIC_200708_archive\BCP_DATA\PUBLIC_CONVERT_BCP_200712.bat
h:\DVD\INFO_ARCHIVE
SET CSVDATADIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\DATA
SET BCPDATADIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\BCP_DATA
SET LOGDIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\LOGS
SET BCPEXEDIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\UTILITIES
```

The default settings for the locations can be seen above. If they are left unaltered, the script assumes the directories in question are under the *%TOP_INFO_ARCHIVE%* directory (passed as a parameter on the command line when running the script). The script also assumes that it is possible to write files to the *LOGDIR* and *BCPDATADIR* directories.

Change the above values to whatever you wish, but make sure that you pre-create the directories if they do not already exist.

N.B. Please note that the *bcpt.ini* file must be in the same directory as the *bcpt.exe* file (setting BCPEXEDIR)

6. On the command line, set the working directory (using CD) to the directory containing the *<VISIBILITYID>_CONVERT_BCP_<YYYYMM>.bat* script and run the script.

For example, if the files were saved from the DVD to the *c:\monthlydvd* directory and the month being loaded was December 2007 with a visibility id of PUBLIC:

```
CMD> cd c:\monthlydvd
CMD> cd PUBLIC_DEC2007_archive\bcp_data
CMD> public_convert_bcp_200712.bat c:\monthlydvd
```

7. Check the logfiles produced in the location specified by the LOGDIR entry as set in the *<VISIBILITYID>_CONVERT_BCP_YYYYMM.bat* for any file conversion errors.

8. Edit the *<VISIBILITYID>_RUN_BCP_<YYYYMM>.bat* script, saving the script to a directory with write access in order to do this.

9. Locate the following section of the script, change where necessary and save the file:

```
REM * Variables default to current directory, change if necessary
REM * Syntax of running file: <PATH>\DATA\<VISIBILITYID>_RUN_BCP_<YYYYMM>.bat username
password dbname servername TOP_INFO_ARCHIVEdir
REM * e.g.:<PATH>\BCP_DATA\RUN_BCP_200708.bat infoserver infoserver partdb host1
h:\DVD\INFO_ARCHIVE
SET BCPDATADIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\BCP_DATA
SET BCPFMTDIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\BCP_FMT
SET LOGDIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\LOGS
SET ERRDIR=%TOP_INFO_ARCHIVE%\PUBLIC_<YYYYMM>_archive\LOGS
SET MAXERRORS=1000
SET BATCHSIZE=10000
```

The default settings for the locations can be seen above. If they are left unaltered, the script assumes the directories in question are under the *%TOP_INFO_ARCHIVE%* directory (passed as a parameter on the command line when running the script). The script also assumes that it is possible to write files to the *LOGDIR* and *ERRDIR* directories.

Change the above values to whatever you wish, but make sure that you pre-create the directories if they do not already exist.

10. On the command line, set the working directory (using CD) to the directory containing the *<VISIBILITYID>_RUN_BCP_<YYYYMM>.bat* script and run the script.

For example, if the files were saved from the DVD to the *c:\monthlydvd* directory and the month being loaded was December 2004 with a visibility id of PUBLIC:

```
CMD> cd c:\monthlydvd
CMD> cd PUBLIC_DEC2004_archive\bcp_data
CMD> public_run_bcp_200412.bat username password databasename servername c:\monthlydvd
```

11. Check the logfiles produced in the location specified by the LOGDIR entry as set in the *<VISIBILITYID>_RUN_BCP_YYYYMM.bat* for any errors.

12. If you disabled any foreign keys in earlier steps, enable the Foreign Keys in the target schema

> **TIP**
>
> *As in the example, it is better to generate the enable script at the time you disable the constraints to ensure you do not enable any constraints supposed to stay disabled in your environment.*

BCP command line

The following instructions are to load the CSV files into an MS SQL Server database on a Windows platform:

1. Unzip the required CSV files in the DATA directory.

2. Run the *BCPT.EXE* file conversion utility as follows:

```
call bcpt.exe <CSVDATADIR>\PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000.csv
<BCPDATADIR>\PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000.csv
<LOGDIR>\bcpt_convert_PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712.log

e.g. call bcpt.exe
<TOP_INFO_ARCHIVE>\DATA\PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000.csv
<TOP_INFO_ARCHIVE>\BCP_DATA\PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000.csv
<TOP_INFO_ARCHIVE>\LOGS\bcpt_convert_PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT.log
```

3. Check the log files and error files for any errors

4. Run the SQL Server BCP utility as follows:

```
bcp <username>.<password>.<tablename> in <csvfilename> -f<formatfilename> -S<servername>
-U<username> -P<password> -F2 -
o%LOGDIR%\PUBLIC_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000_bcp.lo<logfilename> -
e<errorfilename> -m<maxerrors> -b<batchsize>

e.g. bcp scott.tiger in
<PATH>\<VISIBILITYID>_DVD_ANCILLARY_RECOVERY_SPLIT_200712010000.csv -S<servername> -
U<username> -P<password> -F2 -
f<PATH>\<VISIBILITYID>_DVD_ANCILLARY_RECOVERY_SPLIT_200712.fmt –
o<PATH>\<VISIBILITYID>_dvd_ ANCILLARY_RECOVERY_SPLIT.log bad=<PATH>\<VISIBILITYID>_dvd_
ANCILLARY_RECOVERY_SPLIT.bad -e<PATH>\<VISIBILITYID>_dvd_ANCILLARY_RECOVERY_SPLIT.err
```

5. Check the log files and error files for any errors

6. Check the error file for any rows not able to be loaded. Unsuccessful loading of a row is due usually to a column data type mismatch or number of columns mismatch between the data in the CSV file and the structure of the table in your target database

BCP Load hints

Load errors which may be encountered:

1. Primary key violation when loading into the PARTICIPANTCLASS table:

   > Error = [Microsoft][SQL Native Client][SQL Server]Violation of PRIMARY KEY constraint 'PARTCLASS_PK

Some SQL Server configurations may have the collation set to case insensitive. To load the data into this table, modify the PARTICIPANTCLASSID column properties and set the collation to 'case sensitive'

2. Data Type error for rows one or two of the BCP CSV file:

   An error in the log file and an entry in the error file for Row 1 or Row – these errors can be ignored as the first two rows the BCP CSV file are the "I" and "D" records which includes the Report Name, the Report Sub-Type, the Report Version and the Data Column Names.

3. Load errors sue to differences in table structure

   The BCP format files supplied are very specific to the current MMS Data Model version. If a column is specified in the BCP format file and it is not in the destination table, an error will be generated and the rows will not be loaded.

   BCP can handle trailing null columns (e.g. if a column has been added in the most recent MMS Data Model but the destination table does not yet have the column) as long as the extra column is at the end of the table and is not specified in the format file.

4. Errors due to differences in source file and destination table scale for number columns

   If the source file scale is greater than the destination table, a BCP load error may be generated.

   The BCPT.INI file in the UTILITIES directory contains a list of any columns in the DATA*.CSV files which need to be rounded to the correct scale by the BCPT utility before they can be loaded into the MMS Data Model tables.

   The contents of the BCPT.INI file supplied are very specific to the CSV files provided on the DVD and the current MMS Data Model version. If a table affected is not listed in the BCPT.INI file, the column will not be rounded and an error may be generated when loading the rows.

5. Errors due to using a field delimiter of ','

   If the source file uses ',' as a field delimiter, a BCP load error may be generated as some data does contain the ',' character.

   The BCPT.EXE utility in the UTILITIES directory converts the provided CSV files into a file using the '~|' delimiter and removes the optional " text delimiter. Refer to the "Using the supplied BCP Wrapper Scripts" section of this document for instructions on converting the supplied CSV files.

6. Errors due to using an optional text delimiter of '"'

   If the source file uses '"' as an optional text delimited, a BCP load error may be generated as the provided BCP format files do not support it.

   The BCPT.EXE utility in the UTILITIES directory converts the provided CSV files into a file using the '~|' delimiter and removes the optional " text delimiter. Refer to the "Using the supplied BCP Wrapper Scripts" section of this document for instructions on converting the supplied CSV files.

7. "Corrupt zip file" message in the BCPT convert log file

It is possible to pass in a zip file name to the *BCPT.exe* utility. The utility will unzip the source file, convert it and zip it up in the output file location.

If the source zip file is very large, the *BCPT.EXE* utility may generate an error "Corrupt zip file". If this occurs, unzip the source file manually and pass in the unzipped file name to the *BCPT.EXE* utility

The BCP convert wrapper script requires the files to be unzipped manually and passes in the unzipped file name to avoid encountering this error.

Non-Supported Databases

AEMO supports the loading of the CSV files from the DVD into a non-Oracle database by the Parser application.

If you choose to load using a different utility, please refer to your product's documentation specific to loading CSV files. As part of this process, you need to exclude certain rows and fields from the CSV files supplied on the DVD. For more details regarding the format of the CSV files, refer to *CSV file Format* on page 10.

The .DMP file present in the *DATA* directory can only be loaded into an Oracle database, being for the table MARKETNOTICEDATA.

Non-Oracle users wishing to retrieve this data can contact AEMO's Helpdesk on 1300 300 295.

Utility Files

CRLF Utility

The CRLF utility converts text files between <CR><LF> format and <LF> format. MS-DOS (Windows) uses a file format that marks the end of a text line with two characters, <CR><LF>. UNIX uses a format with only <LF>.

If you experience problems loading the CSV files in a UNIX environment, you can use the CRLF utility to convert the CSV file into the UNIX format and see if that solves the problem.

For further instructions, please refer to the *CRLF.txt* file provided in the *\UTILITIES* directory of the DVD.

BCPT Utility

The *BCPT.EXE* utility converts the supplied CSV files for loading into SQL Server. The SQL Server BCP utility does not support the use of optional text delimiters and skipping any rows which are not at the beginning of the CSV file.

The *BCPT.EXE* utility does the following:

- Converts the field delimiter from ',' to '~|' to allow the loading of text fields which contain the ',' character as BCP does not support the use of optional text delimiters.

- Removes the last line which contains the row count for the file, BCP inly supports the skipping of rows at the beginning of the CSV file.

- Removes the optional text delimiter of ' " ' as BCP does not support the use of optional text delimiters.

- Rounds any number columns specified in the *BCPT.INI* file. The BCPT.INI file contains a list of filename, column and the precision and scale for columns which require rounding.

e.g.

DVD_BILLINGREGIONFIGURES.ENERGYOUT=16,6

The utility is run from the command line and takes up to 4 parameters (minimum 2), the parameters are:

1. **Input file** - Relative or absolute path. The file can be a ZIP file or a CSV file. If the input file is a zip file the zip is extracted to C:\temp and the file in the zip transformed. The temp files are then deleted after the transform is complete. If the input file is a CSV file the utility just performs the transform.   Please note that for very large source zip files, it is recommended to unzip the file manually before passing it into the *BCPT.EXE* utility. If the error "Corrupt zip file" is generated, the source zip file is too large for decompression by BCPT.EXE.

2. **Output file** - Relative or absolute path. The filename should be the non-zipped filename. If the inout file is a .zip file, the utility will zip the output file

3. **Optional log file path** - Relative or absolute path. If not included the default is the output filename with .log appended.

4. SQL Server - If not included the default is "~|". You must have parameter 3 to have parameter 4.

bcpt <input file>   <output file>   <logfile>   <field delimiter>

eg.

        bcpt c:\test\inputfile.zip   c:\test\outputfile.txt   c:\test\logs\logfile.log

        bcpt c:\test\inputfile.csv   c:\test\outputfile.txt   c:\test\logs\logfile.log

The supplied *<VISIBILITYID>_CONVERT_BCP_<YYYYMM>.bat* is a wrapper script to which contains the commands to convert of all the supplied CSV files for the month.

Participant Monthly DVD User Guide   Page 23
Version 5.10
Participant_Monthly_DVD.doc (saved on 17 July 2009)

Changes to the monthly DVD

The following is a summary of changes to the monthly DVD.

- The BCP_DATA directory no longer contains converted CSV files

- The BCPT utility is supplied to convert CSV files in the DATA directory into the format required for loading into MS SQL Server using the BCP utility

- The *<VISIBILITYID>_CONVERT_BCP_<YYYYMM>.bat* wrapper script assists in the converting of CSV files

Appendix 1: Example SQL*Loader CTL file

```
-- **************************************************************
-- Title:    PUBLIC_DVD_BIDDAYOFFER_200412.ctl
-- **************************************************************
-- Project : AEMO MONTHLY DVD RELEASE
-- Purpose : Control File to load monthly csv file
-- Target Application  : Monthly DVD Release
-- Application Version : 1
-- Visibility Id : PUBLIC
--
-- **************************************************************
-- Special Instructions : 1. USER TO CONNECT AS BIDDAYOFFER owner
--
-- **************************************************************
-- History :
--      MMS Data Model Version    Template Generated On
--      =====================     ======================
--      v4.4                      01JUL2005
-- **************************************************************
LOAD DATA
INFILE PUBLIC_DVD_BIDDAYOFFER_200412010000.csv
APPEND INTO TABLE BIDDAYOFFER
WHEN (1:1) = 'D'
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY '"'
TRAILING NULLCOLS
(row_type FILLER,
report_type FILLER,
report_subtype FILLER,
report_version FILLER,
DUID CHAR(10),
BIDTYPE CHAR(10),
SETTLEMENTDATE DATE "yyyy/mm/dd hh24:mi:ss",
OFFERDATE DATE "yyyy/mm/dd hh24:mi:ss",
VERSIONNO FLOAT EXTERNAL,
PARTICIPANTID CHAR(10),
DAILYENERGYCONSTRAINT FLOAT EXTERNAL,
REBIDEXPLANATION CHAR(64),
PRICEBAND1 FLOAT EXTERNAL,
PRICEBAND2 FLOAT EXTERNAL,
PRICEBAND3 FLOAT EXTERNAL,
PRICEBAND4 FLOAT EXTERNAL,
PRICEBAND5 FLOAT EXTERNAL,
PRICEBAND6 FLOAT EXTERNAL,
PRICEBAND7 FLOAT EXTERNAL,
PRICEBAND8 FLOAT EXTERNAL,
```

```
PRICEBAND9 FLOAT EXTERNAL,
PRICEBAND10 FLOAT EXTERNAL,
MINIMUMLOAD FLOAT EXTERNAL,
T1 FLOAT EXTERNAL,
T2 FLOAT EXTERNAL,
T3 FLOAT EXTERNAL,
T4 FLOAT EXTERNAL,
NORMALSTATUS CHAR(3),
LASTCHANGED DATE "yyyy/mm/dd hh24:mi:ss",
MR_FACTOR FLOAT EXTERNAL,
ENTRYTYPE CHAR(20))
-- End of Script
```

Appendix 2: Example Foreign key disable/enable script

The intention for the foreign key disable/enable script is for the script to be run before loading any data. The foreign key disable/enable script generates disable and enable scripts only for the foreign key constraints currently enabled in your database. By limiting the generation of scripts to the current situation in your local database means you can ensure you do not accidentally enable a normally disabled constraint in your environment.

The foreign key disable/enable script creates scripts in the current directory.

```
set echo off
set feedback off
set linesize 300
set verify off
set pagesize 0


ACCEPT owner CHAR prompt 'Enter the schema owner: '


REM ***************************
REM Creating Disable script
REM **********************


spool disable_&owner._fk.sql


select 'set echo on feedback on' from dual;
select 'spool disable_&owner._fk.log'||chr(10) from dual;


select 'alter table '||owner||'.'||table_name||' disable constraint
'||constraint_name||';'
from dba_constraints
where owner = UPPER('&owner.')
and constraint_type='R'
and status='ENABLED';


select 'spool off'||chr(10) from dual;


spool off


REM *********************
REM Creating Enable script
REM *********************


Spool enable_&owner._fk.sql
select 'set echo on feedback on' from dual;
select 'spool enable_&owner._fk.log'||chr(10) from dual;


select 'alter table '||owner||'.'||table_name||' enable constraint
'||constraint_name||';'
```

```
from dba_constraints
where owner = UPPER('&owner')
and constraint_type='R'
and status='ENABLED';


select 'spool off'||chr(10) from dual;
spool off
```

Appendix 3: Example SQL Server BCP format file

```
8.0
31
1 SQLCHAR       0    0    "~|"     0        row_type    Latin1_General_CI_AS
2 SQLCHAR       0    0    "~|"     0        report_type Latin1_General_CI_AS
3 SQLCHAR       0    0    "~|"     0        report_subtype    Latin1_General_CI_AS
4 SQLCHAR       0    0    "~|"     0        report_version    Latin1_General_CI_AS
5 SQLCHAR       0    0    "~|"      1     DUID     Latin1_General_CI_AS
6 SQLCHAR       0    0    "~|"      2     BIDTYPE     Latin1_General_CI_AS
7 SQLCHAR       0    0    "~|"      3     SETTLEMENTDATE     ""
8 SQLCHAR       0    0    "~|"      4     OFFERDATE      ""
9 SQLCHAR       0    0    "~|"      5     VERSIONNO       ""
10 SQLCHAR    0    0    "~|"    6     PARTICIPANTID Latin1_General_CI_AS
11 SQLCHAR    0    0    "~|"    7     DAILYENERGYCONSTRAINT      ""
12 SQLCHAR    0    0    "~|"    8     REBIDEXPLANATION      Latin1_General_CI_AS
13 SQLCHAR    0    0    "~|"    9     PRICEBAND1      ""
14 SQLCHAR    0    0    "~|"    10    PRICEBAND2      ""
15 SQLCHAR    0    0    "~|"    11    PRICEBAND3      ""
16 SQLCHAR    0    0    "~|"    12    PRICEBAND4      ""
17 SQLCHAR    0    0    "~|"    13    PRICEBAND5      ""
18 SQLCHAR    0    0    "~|"    14    PRICEBAND6      ""
19 SQLCHAR    0    0    "~|"    15    PRICEBAND7      ""
20 SQLCHAR    0    0    "~|"    16    PRICEBAND8      ""
21 SQLCHAR    0    0    "~|"    17    PRICEBAND9      ""
22 SQLCHAR    0    0    "~|"    18    PRICEBAND10      ""
23 SQLCHAR    0    0    "~|"    19    MINIMUMLOAD      ""
24 SQLCHAR    0    0    "~|"    20    T1     ""
25 SQLCHAR    0    0    "~|"    21    T2     ""
26 SQLCHAR    0    0    "~|"    22    T3     ""
27 SQLCHAR    0    0    "~|"    23    T4     ""
28 SQLCHAR    0    0    "~|"    24    NORMALSTATUS Latin1_General_CI_AS
29 SQLCHAR    0    0    "~|"    25    LASTCHANGED      ""
30 SQLCHAR    0    0    "~|"    26    MR_FACTOR      "" > %BCPFILE%
31 SQLCHAR    0    0    "\r\n"   27   ENTRYTYPE     Latin1_General_CI_AS
```