

CHAPTER

3

Using MDP Extensions

<i>What Is the Multidimensional Data Provider?</i>	29
<i>Data Requirements</i>	30
<i>Setting Up the MDP Metadata</i>	30
<i>Data Groups</i>	31
<i>Servers</i>	34
<i>EIS Registrations</i>	37
<i>Using MDP in EIS without Defining MDP Metadata</i>	38

What Is the Multidimensional Data Provider?

SAS/EIS software enables users to build applications based on high-level components, such as viewers (pre-built reports), data models, and metadata. The Multidimensional Data Provider (MDP) is a data model that accesses a variety of data structures in a distributed environment. These data structures include SAS data sets, multidimensional databases (MDDBs), views on external relational database management system (RDBMS) files, star schemas, and output data sets from PROC SUMMARY. The MDP runs from multiple data sources. It allows you to split multidimensional databases both horizontally (storing different aggregation levels in different locations) and vertically (storing new data in separate files).

You can use the MDP as an extension to the SAS/EIS multidimensional data model, thus enabling SAS/EIS OLAP applications to access transparently any data that has been distributed optimally. The MDP increases the power and efficiency of SAS/EIS multidimensional applications that process large amounts of data, and it also adds scalability and full compute server processing.

You should consider using MDP if any of the following situations exists:

- if the amount of data is too large to be handled within one MDDB
- if there is a performance bottleneck when accessing data on a server
- if you want to make use of existing summarized data sources.
- if you want to secure your investment in OLAP applications by
 - providing for scalability as data volumes grow
 - maintaining flexibility in your choice of data storage format and location.

To take full advantage of MDP, you should have a moderate to advanced knowledge of the development features of the SAS System.

Data Requirements

The Multidimensional Data Provider allows you to treat data in different formats and different locations as one logical unit. Single aggregation levels of the same data can be distributed over different locations and stored in different formats, allowing you to optimize resource use and performance. You can also use existing presummarized data without needing to convert the data into MDDBs or to copy the data onto client machines. You can split aggregation levels by class data value, which enables you to add data to your multidimensional database as it is generated, without having to re-create already existing presummarized units to integrate the new data.

Data in the following formats are supported by default:

- MDDBs
- PROC SUMMARY output data sets
- nonsummarized data sets
- star schemas
- SQL views
- RDBMS views.

Data sources for one logical multidimensional unit, the data group, need to share the names for class and analysis columns. Each data source needs to contain only those columns that are necessary for the aggregation levels that the data source represents. Create your distributed multidimensional data sources using the SAS data management tools, or take advantage of the menu-driven HOLAP generator in SAS/Warehouse Administrator software.

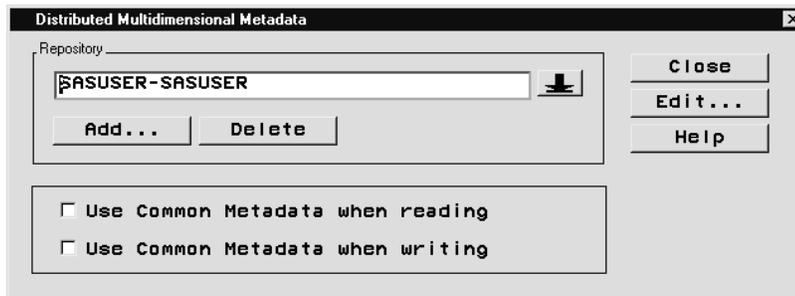
Setting Up the MDP Metadata

The Multidimensional Data Provider uses a set of metadata that describes where each component of a logical data group is located and what it contains. You use the Distributed Multidimensional Metadata facility to define data groups and servers.

To invoke the Distributed Multidimensional Metadata facility, invoke SAS/EIS and select **Metabase** from the EIS Main Menu. When the Metabase window appears, select from the menu bar



or type the **MDMDB** command to open the Distributed Multidimensional Metadata window.



Note: Alternatively, you can use the experimental MDP Metadata Batch Utility to define or to modify your multidimensional metadata, as described in Technical Support

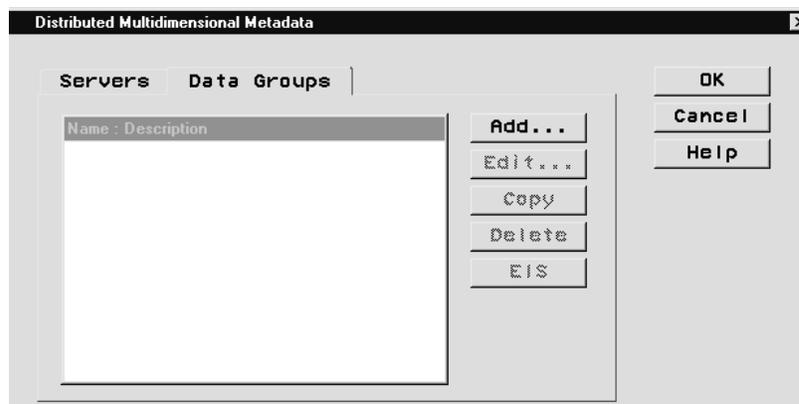
Document TS596. For details on using this utility, download the document from the SAS Institute Technical Support home page. [△](#)

Store your MDP metadata in a repository. To manage your repositories, select **Add** and **Delete** in the Distributed Multidimensional Metadata window. To define the MDP metadata for a given repository, select the repository and then select **Edit**. For more detailed information on managing your repositories, see online Help for the Repository Manager command (**REPOSMGR**).

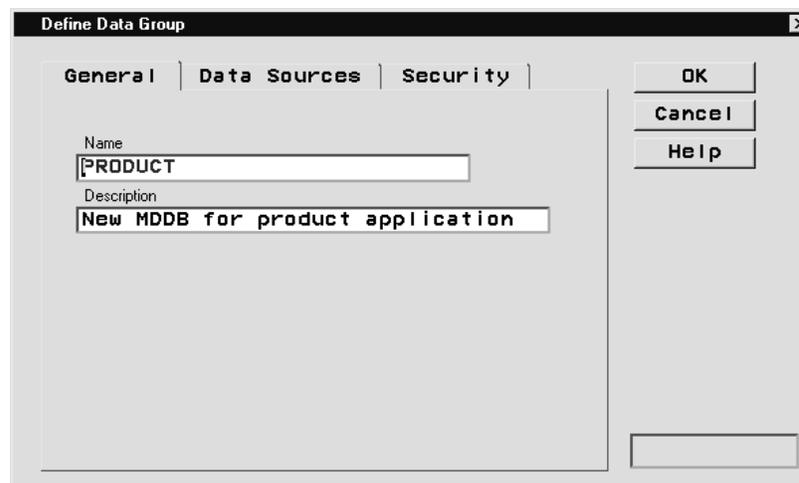
Data Groups

The following steps describe the process that you use to create and modify data groups.

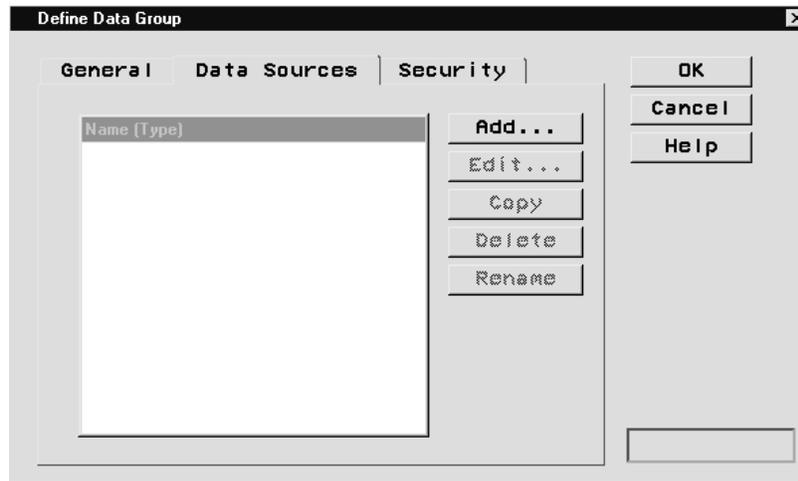
- 1 To define a data group, open the Distributed Multidimensional Metadata window, select a repository, and then select **Edit**.



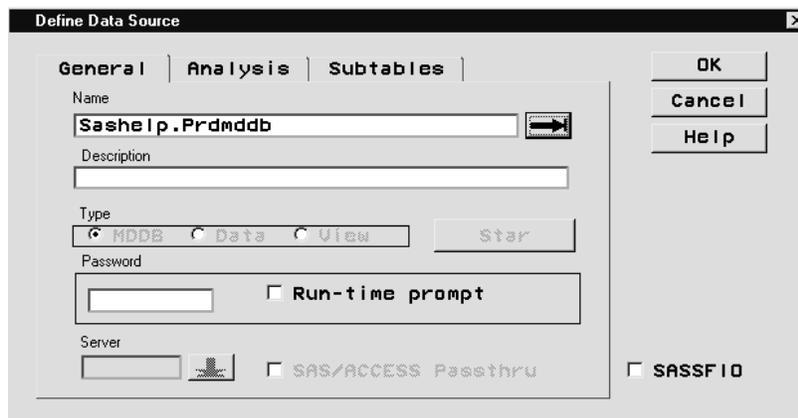
- 2 Next, select **Add** in the **Data Groups** tab.



- 3 In the **General** tab, type a name and a description for the data group.
- 4 Use the **Security** tab to provide read and write passwords to protect your metadata against unauthorized access.
- 5 To define all the data sources for your data group, select the **Data Sources** tab.



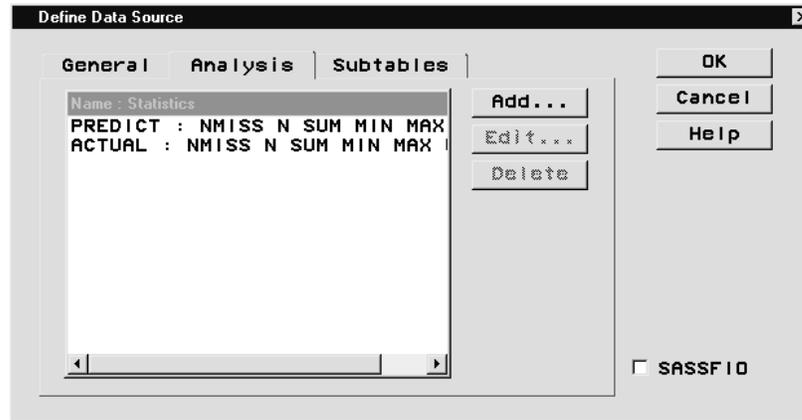
- 6 In the Data Sources window, select **Add** to open the Define Data Sources window.



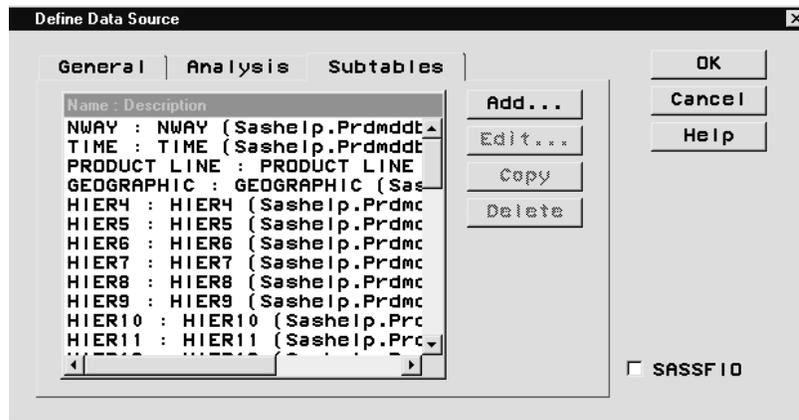
- Select or specify the SAS filename of your data source. If the data source already exists, the Type is automatically determined.
- If the data source is password-protected, you can specify the password in this window to avoid prompting the user when the data is accessed at application runtime.
- If it is a non-local data source, select **Server** to open a window in which you can specify an MDP server. See “Servers” on page 34 for details on how to set up an MDP Server definition.
- For existing data sources, the information about aggregation levels and analysis columns/statistics is automatically filled in for you.
- For non-MDDB data sources with analysis columns that represent multiple statistics per original analysis column and that use the MDDB SASSFIO prefixing mechanism to identify the analysis column name and statistic, select the SASSFIO check box before you enter the data source name.

Note: Refer to online Help for the `_write` method of `sashelp.mddb.mddb_m.class` for details on the SASSFIO naming conventions. △

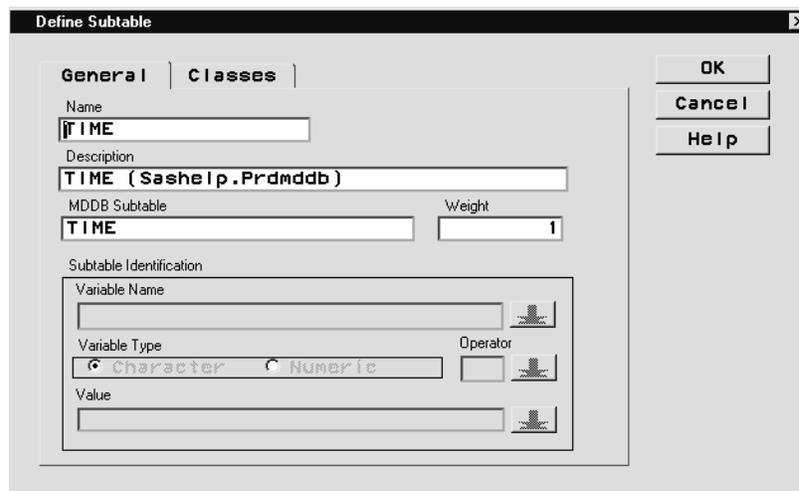
- If you have non-MDDB data sources that do not follow the SASSFIO naming convention, use the **Analysis** tab to specify which statistic each column represents (the default is SUM).



- 7 After you specify general information about your data source, select the **Subtables** tab to specify or to verify the aggregation levels that your data source contains.

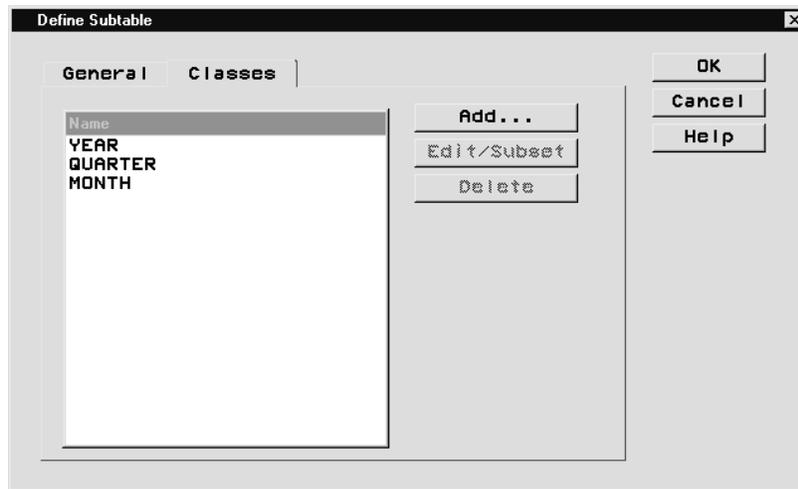


A *subtable* (or aggregation level) is defined by the combination of class variables it represents. Select a subtable in the list and then select **Edit**.

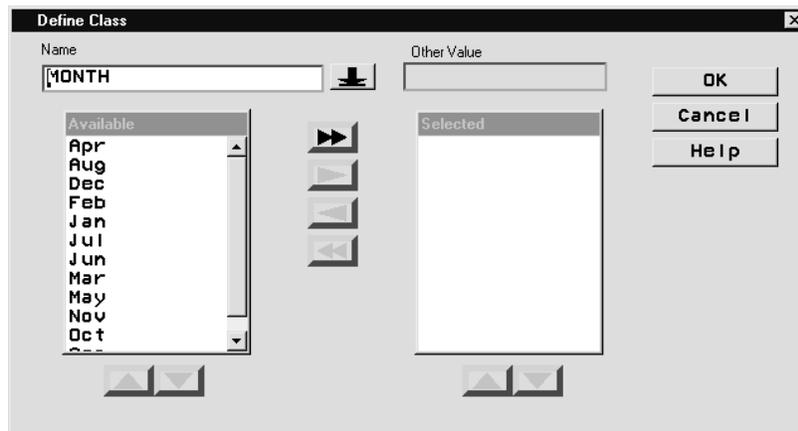


If a subtable exists in more than one data source, use the Weight attribute to establish which subtable should be used. The lowest value takes precedence. For

non-MDDB data sources with multiple subtables, each subtable is identified by the value of an identification variable (for example, `_TYPE_` for PROC SUMMARY output data sets). Select the **Classes** tab to define or verify the combination of class variables for the subtable.



When you add or edit a Class, you can specify a subset for that class. For example, you can select MONTH and then select **Edit/Subset** to display the Define Class window.



You can subset each class variable by value, which enables you to split summary pyramids vertically (for example, you can store the summarized data for each year in a separate file). Note that any subset specified is automatically applied to every occurrence of the given class variable in every subtable of that data source.

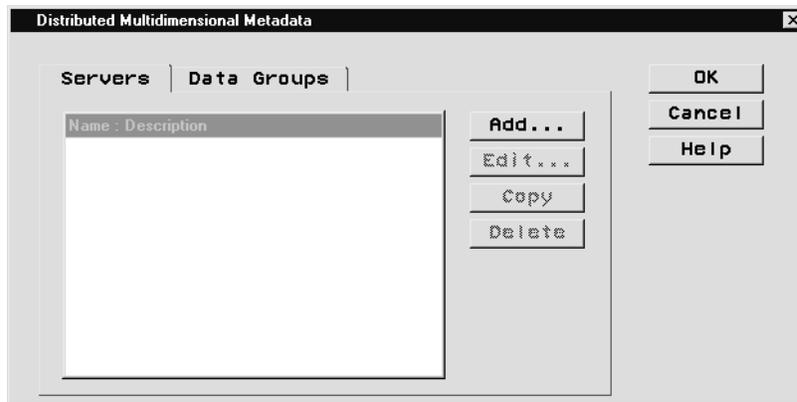
Servers

MDP server definitions are referenced by individual data sources. An MDP server definition is required if data sources reside on non-local machines. Please note that data sources can only use MDP server definitions that reside within the same repository.

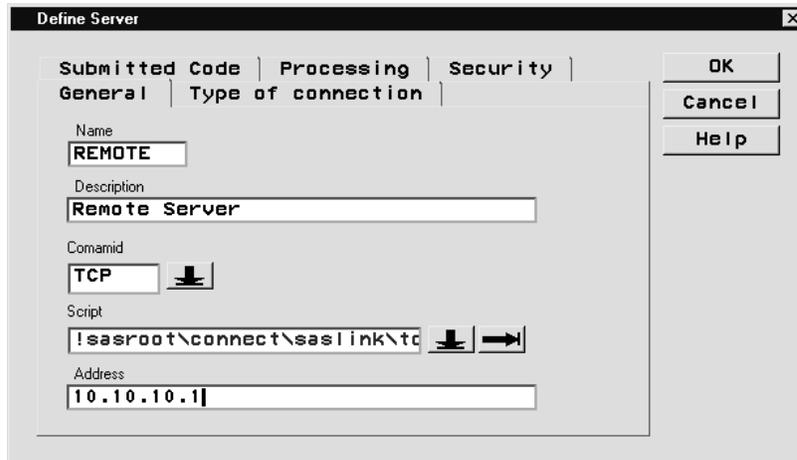
In addition to the SAS/CONNECT required information on a server, MDP server definitions define how to distribute processing load and enable you to specify additional

processing instructions. You can define more than one MDP server on the same SAS/CONNECT server. The following steps describe how to define a server.

- 1 Open the Distributed Multidimensional Metadata window, select a repository, and then select **Edit**.

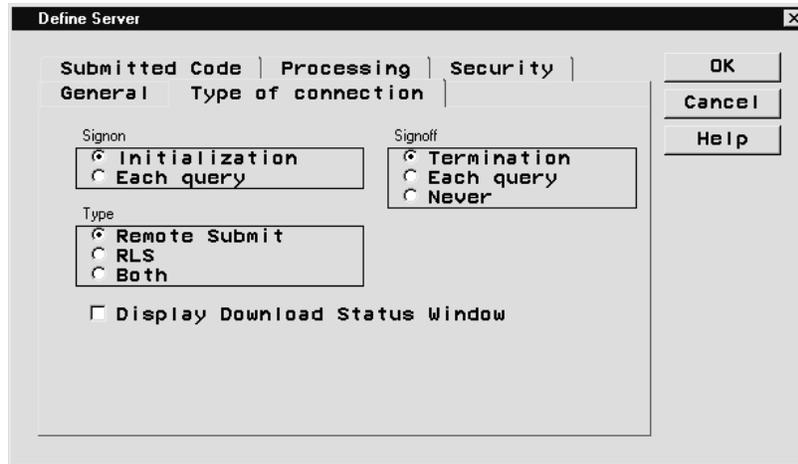


- 2 In the **Servers** tab, select **Add** to define a new server.



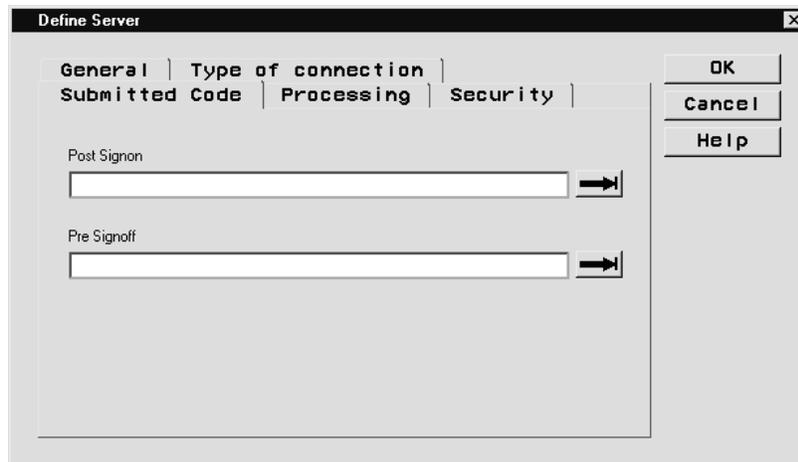
In the **General** tab, specify the items required to establish a SAS/CONNECT connection. Select a name for your server, the communications access protocol, and a SAS/CONNECT script file. If your server name is not a known network node name, you also need to specify an address. The form of the address depends upon the protocol used.

- 3 Select the **Type of Connection** tab to specify server invocation and processing load distribution options.



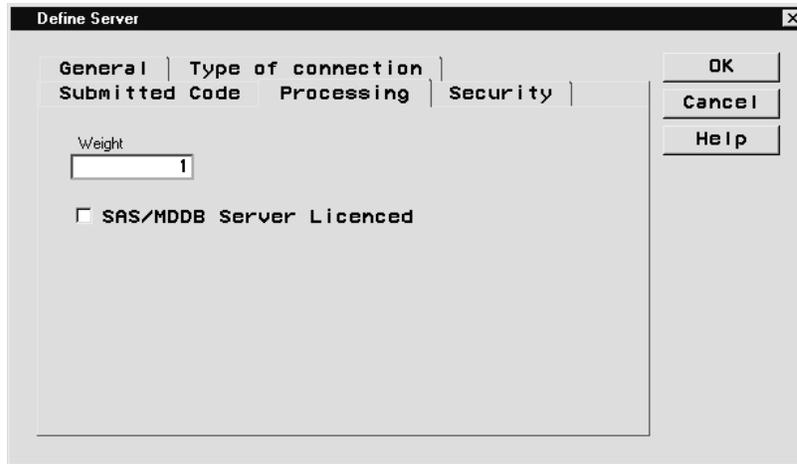
Specify when to sign on and sign off and whether to use Remote Library Services (RLS) or Remote Submit (or let MDP decide). RLS is recommended if the requested subtables exist on the server. If summarization on the fly is required, Remote Submit is usually the better choice because the summarization occurs on the server and only the data that is actually needed for display is downloaded.

- 4 Select the **Submitted Code** tab to specify code locations for additional processing.



Specify the name of a SOURCE catalog entry with code to be submitted immediately after logging on or before logging on. Use this code to provide for data allocations and environment setups. Note that the code is submitted locally. If you submit any code remotely, enclose it in an **RSUBMIT; ENDRSUBMIT;** block.

- 5 Select the **Processing** tab for further processing and load optimization.



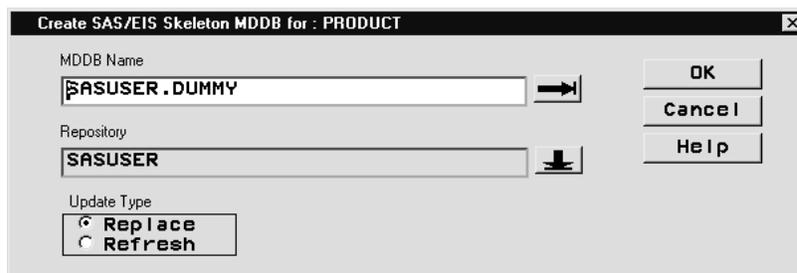
Specify whether SAS/MDDB Server software is licensed in your remote session. This makes it easier for MDP to decide whether remote processing is feasible on the server. You can also specify a weight for your server. This attribute is used in case a requested subtable is present on more than one server. The lowest value takes precedence.

- 6 Use the **Security** tab to provide read and write passwords to protect your metadata against unauthorized access.

EIS Registrations

To be able to use your data groups in SAS/EIS multidimensional applications, you need to register each data group in the SAS/EIS Metabase. To do that, use the following steps:

- 1 Open the Distributed Multidimensional Metadata window, select a repository, and then select **Edit**.
- 2 Select the Data Groups tab, and then select an existing data group.
- 3 Select **EIS**.



- 4 Specify the name of a skeleton MDDB, which is an MDDB with 0 cells that has all the class variables and analysis/statistic columns of your data group. Both the metabase registration and the EIS data model use this skeleton MDDB and send only the actual data request to MDP. The name of the skeleton MDDB cannot be the name of an existing data source.
- 5 Select a repository in which to store the Metabase registration.
- 6 MDP data groups are registered with a minimal set of table and column attributes. Therefore, you need to add any additional attributes (for example, the

HIERARCH table attribute, the DATE column attributes, and so on) after the automatic registration of your data group. To prevent the loss of existing additional attributes, use **Update Type: Refresh** when you update your Data Group registration.

- 7 Select **OK** to create the skeleton MDDB and to automatically register the data group as an EIS registration in the specified repository.

Using MDP in EIS without Defining MDP Metadata

When you use a single remote MDDB in your EIS application, you might want to take advantage of the remote computing capabilities offered by using the multidimensional data provider. You can expect to see performance gains through remote computing, as opposed to standard RLS processing, if you do not have a very fast server connection or if your remote MDDB is not highly aggregated. Also, especially with large subtables, the multidimensional data provider's caching facility will speed up repeated access to data and reduce network traffic.

You do not need to define an MDP data group and server when you use MDP for a single remote MDDB. Simply add the COMPSRV table attribute to your EIS metabase registration. Also, before you run any reports, set up your remote connection and define a libname in the remote session using the same name and path that you used in the metabase registration.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/EIS[®] Software: Administrator's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999.

SAS/EIS[®] Software: Administrator's Guide, Version 8

Copyright © 1999 SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-506-X

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, by any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute, Inc.

U.S. Government Restricted Rights Notice. Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999

SAS[®] and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. [®] indicates USA registration.

IBM, ACF/VTAM, AIX, APPN, MVS/ESA, OS/2, OS/390, VM/ESA, and VTAM are registered trademarks or trademarks of International Business Machines Corporation. [®] indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.