SymmetricDS Pro 2.4.0

Quick Start Guide

© 2011 JumpMind, Inc.

SymmetricDS Synchronization Concepts

SymmetricDS is a change data capture, replication solution that can be used to synchronize databases in near real time. Synchronization can be bi-directional, scale to a very large number of databases, work across low bandwidth connections, and withstand periods of network outage. It is a general purpose tool that can be deployed and configured in many different ways to solve a wide variety of synchronization concerns.

In order to get the solution up and running it is important to understand some base terminology. We'll use a simple retail business scenario that synchronizes data (information about pricing, inventory, transactions, etc.) between a central office database and multiple store databases to help explain the terms.

Term	Description
Node	A single installed and running instance of SymmetricDS that is attached to, and has responsibility for, keeping a database instance in synch. In the example below, #1, #2, and #3 are all examples of nodes in a SymmetricDS implementation
Root Node	The master or primary Node in the SymmetricDS implementation that is responsible for storing and coordinating all configuration (metadata) information about the synchronization scenario. The root node is also responsible for hosting the SymmetricDS web management console that allows you to configure, manage and monitor your solution. In the example below, #1 is the root node.
Node Group	A logical grouping of Nodes used in defining your synchronization configuration. A Node Group can be a single node, or it can be thousands of nodes. Node Groups are based on which Nodes share common synchronization configuration or rules. For example, all stores need to get pricing information from, and send transaction information to the Central Office. We don't want to define that configuration for each store (Node), but instead want to be able to define it for the group of stores as a whole (Node Group). Each Node must be contained within a Node Group. In our example below, we have two Node Groups defined, "Corp" which is #4 and "Stores" which is #5. The Stores node group includes each Node for every store, and the Corp includes the single instance that is the Root Node.
Root Database	The jdbc compliant database attached to the Root Node. This database usually contains both data you want synchronized, as well as all of the SymmetricDS configuration data for your synchronization solution. In the example below, #6 is the Root Database
Client Database	The jdbc compliant database attached to one of the other SymmetricDS nodes. This database contains data you want synchronized as well as SymmetricDS configuration data sent to it by the Rood Node from the Root Database. In the example below, #7 is a Client Database

# SymmetricDS Pro 2.4 QuickStart Guide
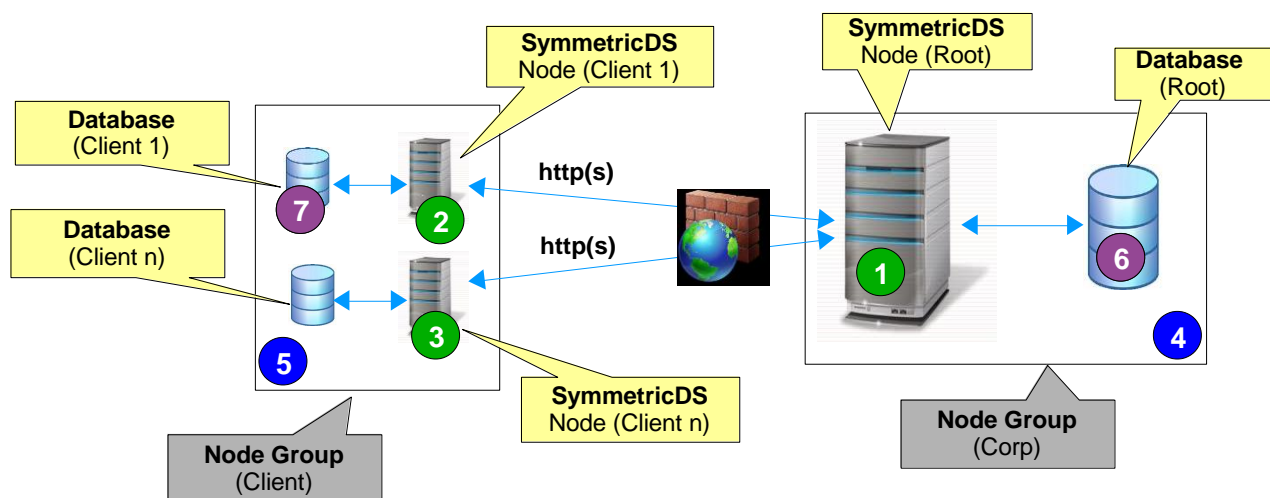
Installing the Root Node

To begin installation of the root node:

- Verify that you have the 1.5.x Java Runtime Environment (JRE) installed on the Root Node computer.
- Verify that the 1.5.x JRE bin directory is in your path (windows) or <what do I need to put here for Mac, Linux or Unix>
- Run the SymmetricDS installer by double clicking the symmetric-pro-2.4.x-install.jar
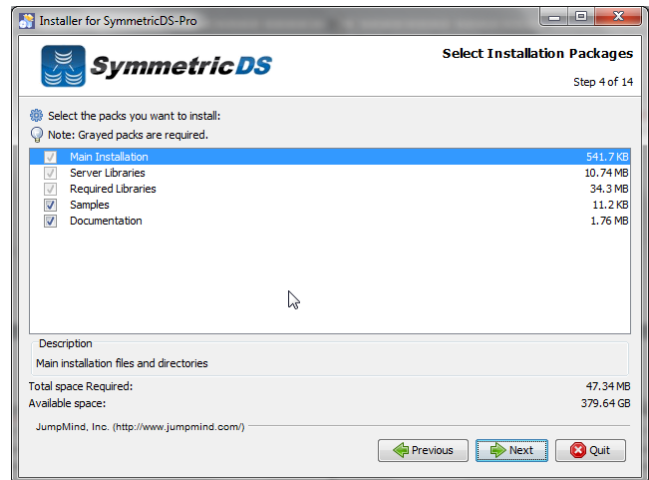- Click "Next" after reading the Welcome screen.
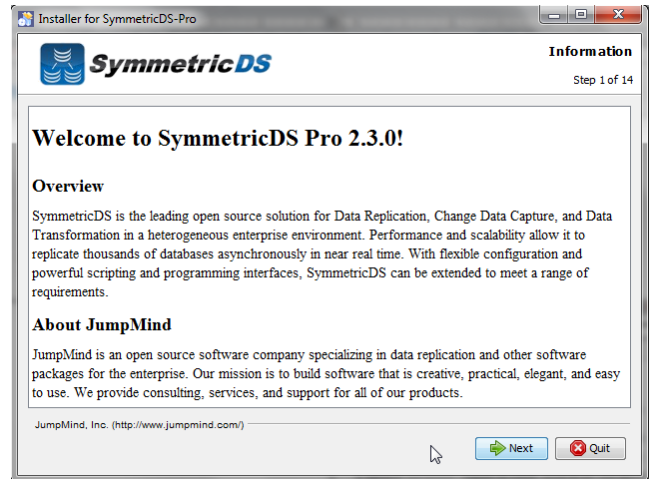
Review and accept the license agreement
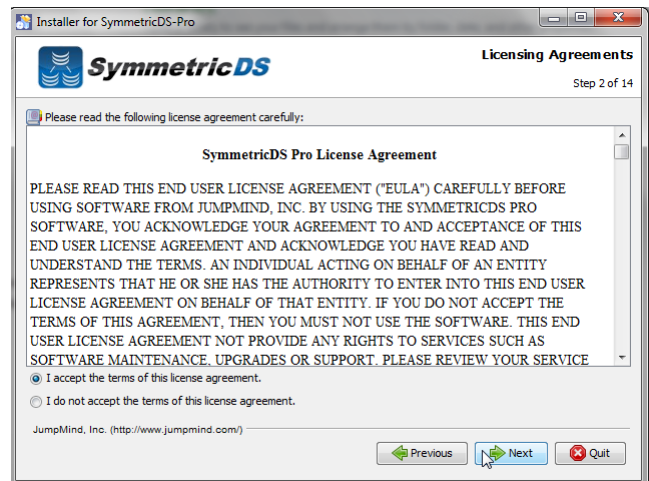
Select the installation directory where the SymmetricDS application should be installed
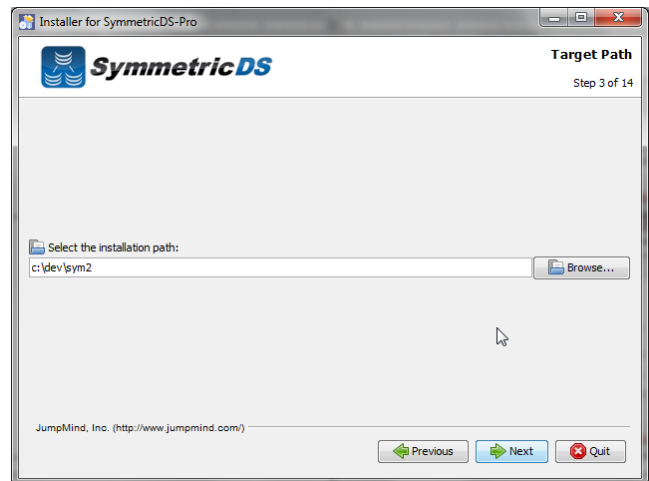
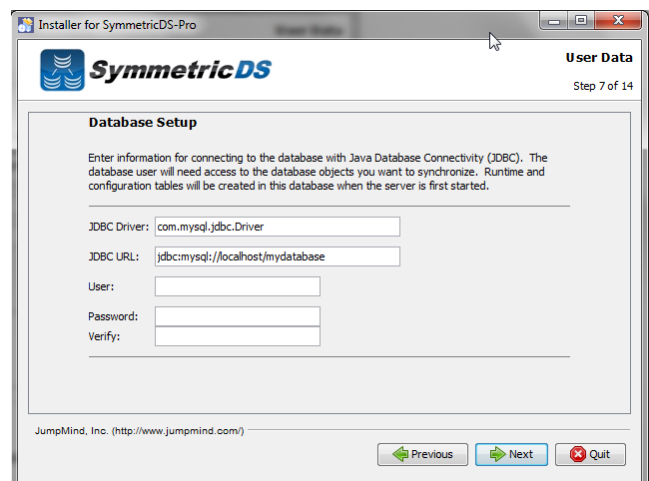Select the components you would like to install.

The installation process creates the Root Node Group and the Root Node configuration definitions based on the entries in this screen.

- **Node Group Id** –This is the unique Identifier for the Node Group that will contain the Root Node. Typically there is only one Node in the Root Node Group.
- **External Id** – The unique identifier for the Root Node itself (each Node needs to have a unique identifier that allows other nodes in the synchronization scenario to identify it)
- **Node Type** – Type type of Node that is being installed (either Root Node or Client Node). As this is the Root Node, leave "Root Node" selected.
- **Database** – This is the type of database that will be used for the Root Database. Select your Root Node database type from the list.
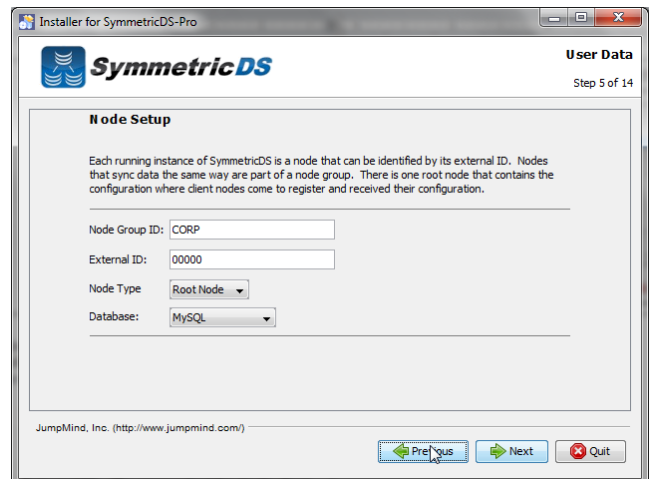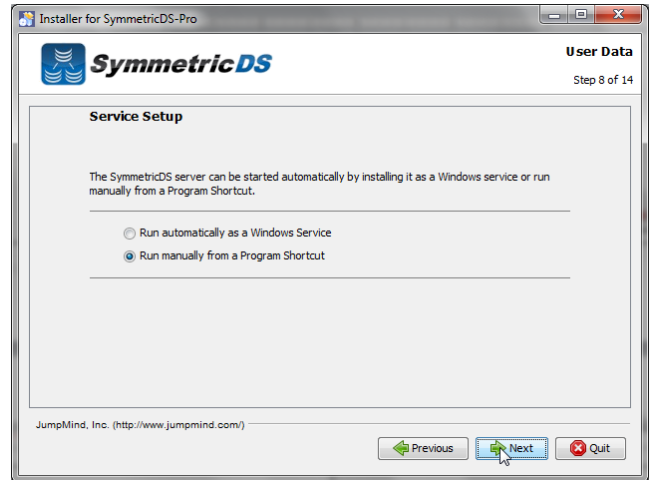
The next screen allows you to specify the connection information for the Root Database.
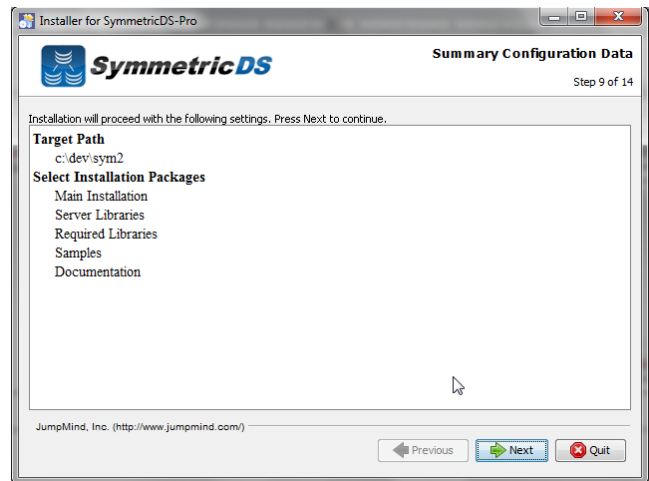
- **JDBC Driver** – The class name for your JDBC driver for your Root Database
- **JDBC URL** – The URL for your Root Database
- **User Id** – The user id for your Root Database. This Id will need Create, Read, Update, Delete (CRUD) privileges for tables
- **Password** – Associated password for the User ID for the Root Database
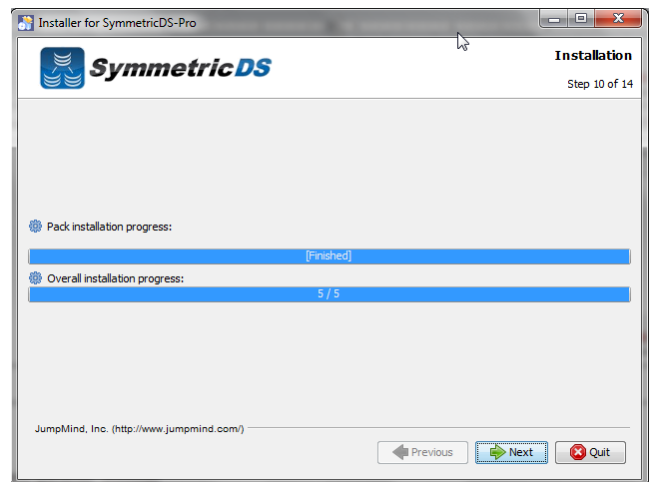- **Verify** – Verification of the password

SymmetricDS can be started manually from a Program Shortcut or as a Windows Service. Select one of the two options and click "Next."
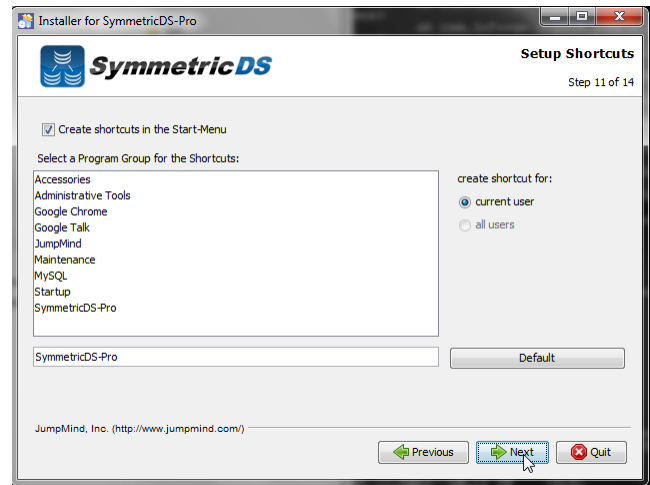
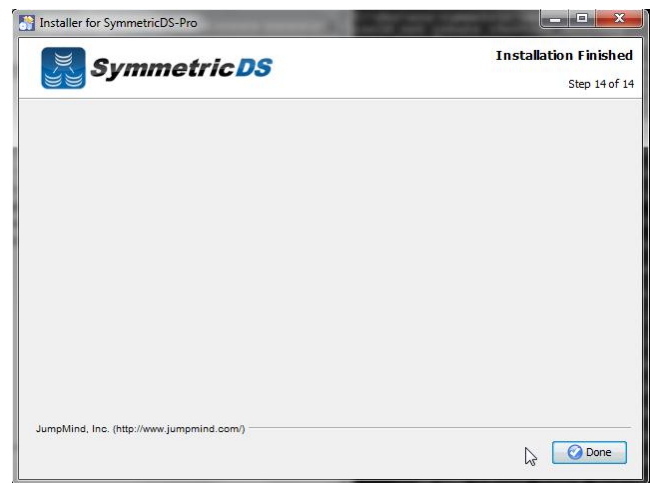The Summary Configuration screen allows you to review your installation choices. Review and click "Next"

Once the application files have been copied to the installation directory, click the "Next" button.

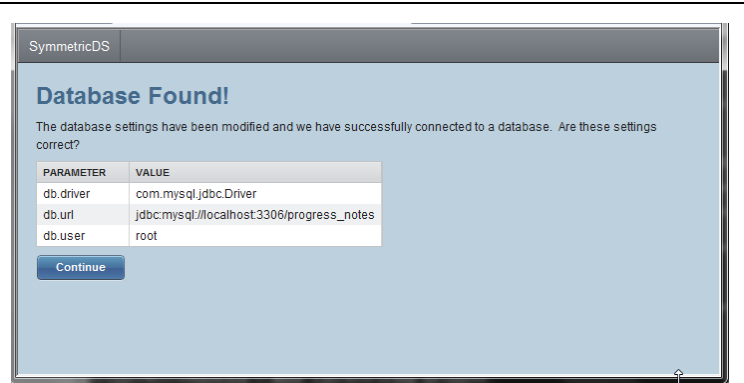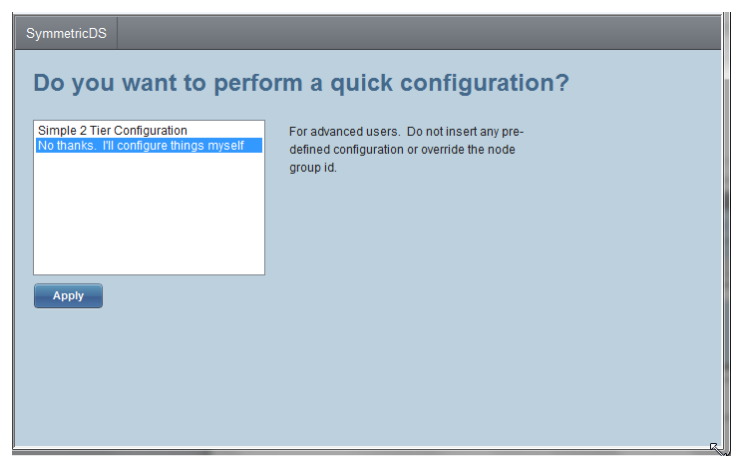|	
The Setup Shortcuts screen allows you to specify where you would like shortcuts created for the Root SymmetricDS instance. Select a group and click the "Next" button.	
Installation of the Root Node is now complete. Click the "Done" button.	
To start the SymmetricDS Pro Software, either start the service (if you selected to run as a service), or start manually from the start menu.	
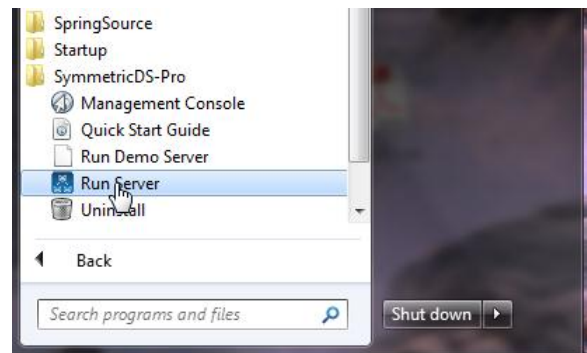
Once you have started the Root Node, open a web browser and go to the following url, replacing machine_name with the name of the computer on which the Root Node is installed

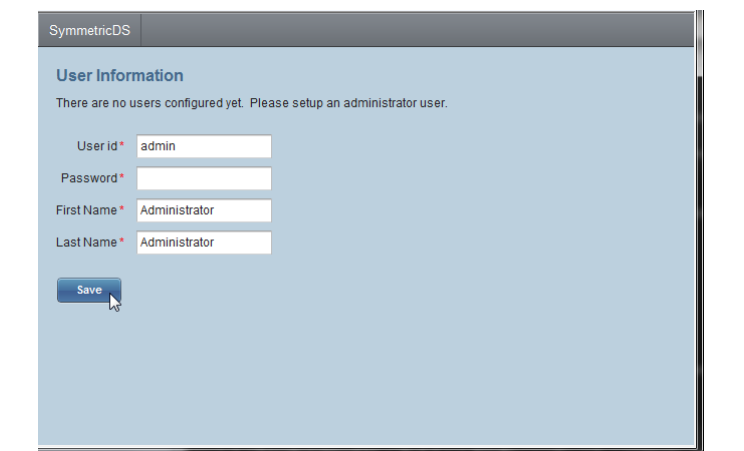http://<machine_name>:31415

The first thing the Root Node will do is let you know that it found the Root Database you specified during installation. Click the Continue button.

SymmetricDS

Database Found!

The database settings have been modified and we have successfully connected to a database. Are these settings correct?

| PARAMETER	VALUE
db.driver	com.mysql.jdbc.Driver
db.url	jdbc:mysql://localhost:3306/progress_notes
db.user	root

Continue

SymmetricDS Pro has the concept of pre-configured synchronization profiles. These profiles are out of the box synchronization configurations for off the shelf software programs. To set up a sample synchronization from scratch, select "No Thanks, I'll configure things myself."
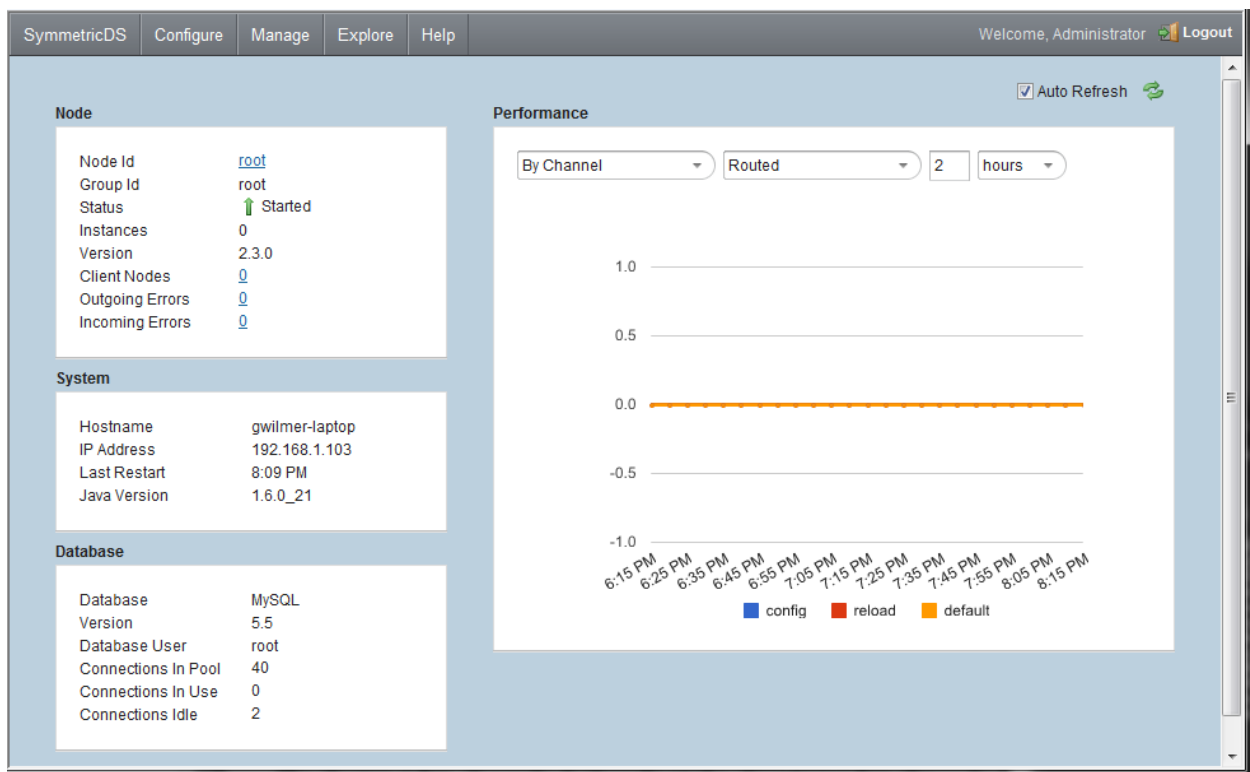
SymmetricDS

Do you want to perform a quick configuration?

Simple 2 Tier Configuration
No thanks. I'll configure things myself

For advanced users. Do not insert any pre-defined configuration or override the node group id.

Apply

Next, you must specify a user id and password for the administrative console. Select a userid, password, and click the Save button.

SymmetricDS

User Information

There are no users configured yet. Please setup an administrator user.

User id* admin
Password*
First Name* Administrator
Last Name* Administrator

Save

After entering your user id and password, the console will make you log in with your credentials before proceeding further.
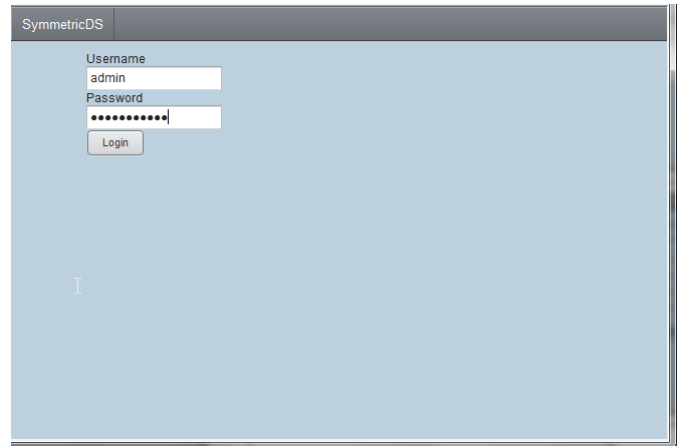
The next screen (shown below) is the main screen and dashboard for SymmetricDS Pro. From this screen we can configure, manage and explore our synchronization scenario. The Configure menu allows us to set up our scenario. The Manage menu allows us to manage the day to day operations of that scenario, and the Explore menu allows us to look at data flowing through the system. We'll start with the Configure menu option and configuring your scenario, but first, let's talk about some additional concepts.
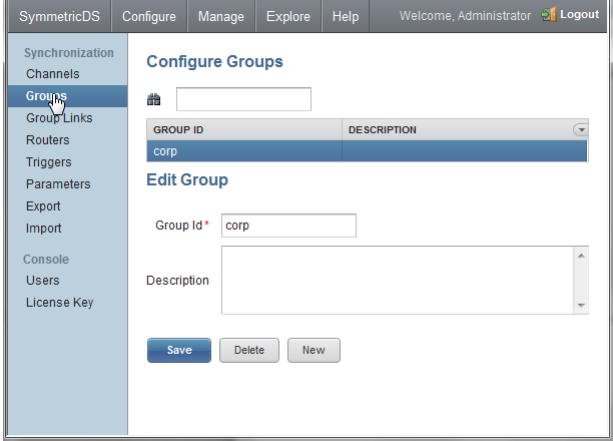
SymmetricDS Configuration Concepts

We've already talked about Nodes and Node Groups. Now we'll define terms that are needed to configure the data that is captured, how that data is logically grouped, and where the data should be sent.
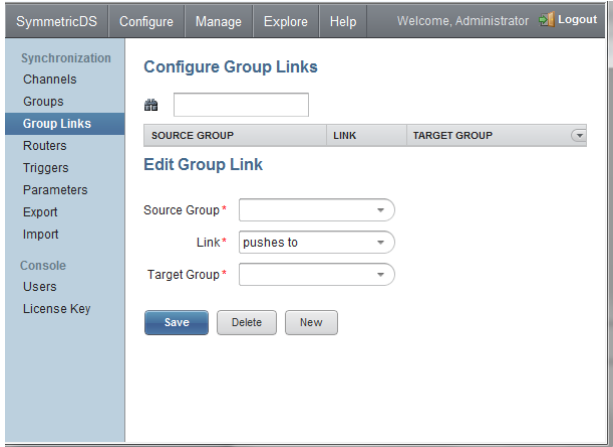
Term	Description
Node Group	A logical grouping of Nodes used in defining your synchronization configuration. See SymmetricDS Synchronization Concepts section earlier in this document for additional details.
Node Group Link	A Node Group Link defines how data flows between each of the Node Groups in your synchronization scenario. Each Node Group Link allows you to define a Source Node Group, a Target Node Group, and whether data is pulled or pushed between those Node Groups.
Channel	SymmetricDS supports the concept of Channels of data. Data synchronization is defined at the table (or table subset) level, and each managed table can be assigned to a *channel* that helps control the flow of data. A channel is a category of data that can be enabled, prioritized and synchronized independently of other channels. For example, in a retail environment, users may be waiting for inventory documents to update while a promotional sale event updates a large number of items. If processed in order, the item updates would delay the inventory updates even though the data is unrelated. By assigning changes to the item tables to an *item* channel and inventory tables' changes to an *inventory* channel, the changes are processed independently so inventory can get through despite the large amount of item data.
Trigger	SymmetricDS uses triggers to capture changes that need to be routed to other Nodes. SymmetricDS Triggers are standard database triggers supported by your database platform that are automatically created based on your configuration data.
Router	Routers are responsible for taking changes that are captured by Triggers and determining to which Node Groups those changes are delivered.

Configuring The Synchronization Solution

Node Groups - The first step in configuring your synchronization scenario is to set up your Node Groups (see Node Groups from the Synchronization Concepts page. The Root Node Group has already been set up for you based on the information you supplied to the installer when installing the Root Node. To create additional Node Groups, click the "New" button, fill in the Group Id (unique Id that identifies the Node Group) and a description and click "Save." You must have at least two Node Groups to continue the configuration of your scenario.
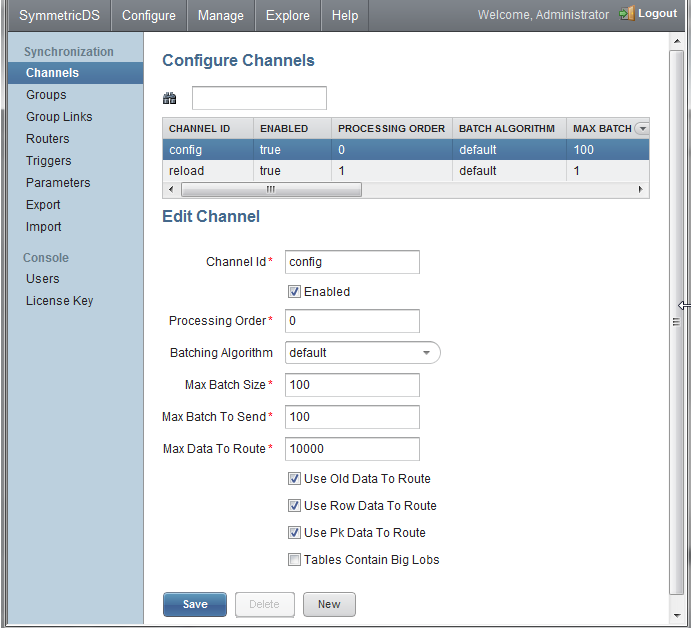
Node Group Links - The next step in configuration is setting up Node Group Links. The purpose of the links is to describe how data flows between the Node Groups. They define a source Node Group, a target Node Group, and a link type, namely whether the data changes are *pushed* or *pulled*. The push method causes the source Node Group to connect to the target Node Group and push data changes to it, while the pull method causes the source Node group to wait for the target Node Group to connect to it and pull the data changes that have been recorded on the source.

When you save a new "Node Group Link" you will be asked if you want to create a default router for the link. We'll describe this in more detail in the Router section, but for now, click "Yes" to create the Router.

Data Channels - The next step is to configure Data Channels. In a nutshell, Data Channels allo you to logically group data that is being synchronized within your scenario. When SymmetricDS captures data changes in the database, the changes are captured in the order in which they occur. In addition, that order is preserved when synchronizing the data to other nodes. Frequently, however, you will have cases where you have different "types" of data with differing priorities. Some data might, for example, need priority for synchronization despite the normal order of events. For example, in a retail environment, users may be waiting for inventory documents to update while a
promotional sale event updates a large number of items. SymmetricDS supports this by allowing tables being synchronized to be grouped together into Channels. A number of controls to the synchronization behavior of SymmetricDS are controlled at the Channel level. For example, Channels provide a processing order when synchronizing, a limit on the amount of data that will be batched together, and isolation from errors in other channels.

Triggers - SymmetricDS uses database triggers as the capture mechanism to record changes to be synchronized to other nodes. Based on the configuration you provide, SymmetricDS creates the needed database triggers automatically for you. Each trigger you define is for a particular table that needs to be sychronized. There is a great deal of flexibility in terms of defining the exact conditions under which a data change is captured by the trigger. For each trigger you can also specify:

- Whether to install a trigger for updates, inserts, and/or deletes
- Conditions on which an insert, update, and/or delete fires
- A list of columns that should not be synchronized from this table
- A SQL select statement that can be used to hold data needed for routing (known as External Data)

You can create triggers one at a time by clicking the "New" button, filling in the appropriate details for the trigger, and clicking the "Save" button, or you can "Auto Create" them. The "Auto Create" button is a feature that allows you to browse your database table definitions and select tables for which you would like triggers created. You can select multiple tables, and then simply click the "Apply" button which will automatically create triggers for each table selected. Once triggers have been created with the "Auto Create" mechanism you can update each trigger's attributes indivdually by selecting them from the list, editing them, and saving them.

Routers - The triggers that have been defined in the previous section only define *when* data changes are to be captured for synchronization. They do not define *where* the data changes are to be sent. Routers, plus a mapping between Triggers and Routers (Trigger/Router), define the process for determining which Nodes receive the data changes.

When you created your Node Group Links, a default Router for each Node Group Link was automatically created for you. This Router defines the data route between your Node Groups (i.e. Node Group 1 pushes data to Node Group 2). All that is remaining to do is to define which triggers are associated with which Router. You can complete this mapping between Triggers and Routers on either the Routers page or the Triggers page. From the Triggers page, simply select a Trigger from the list, and check the "Linked Router" box to have that trigger's captured data be routed by the given router. From the Routers page, select a Router from the list, and then select all of the "Linked Triggers" that you want routed via that Router.

After your Triggers and Routers have been configured, and the association between the two are mapped, you have completed your initial synchronization configuration.
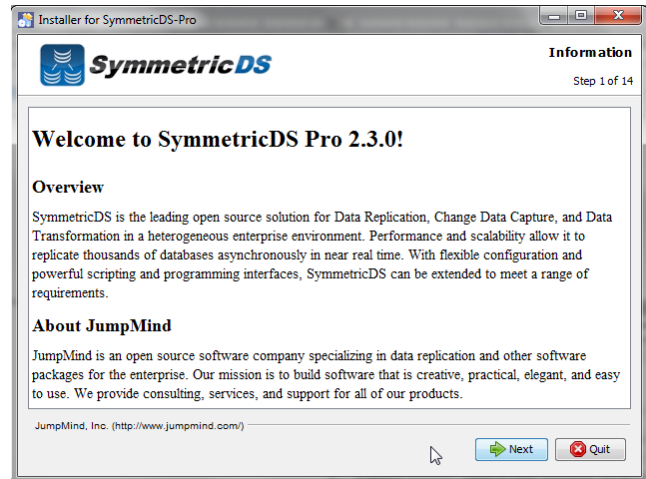
Client Node Concepts

Now that we have our root instance installed and our synchronization scenario configured, the next steps are to install our client instance. The following are some terms that will help in understanding setting up the client.

| Term	Description
Client Node	A SymmetricDS instance (other than the Rood Node) that is participating in the synchronization scenario
Node Registration	For a Node to participate in the synchronization it needs to ask for permission from the Registration Server. Node registration is the process of a Client Node requesting to be added to the synchronization scenario.
Registration Server	The Node (the Root Node) that is responsible for accepting Node Registration requests
Registration URL	The URL of the Root Node which is acting as the Registration Server. The URL includes the Name (or IP Address) and port that the SymmetricDS instance is running on
Initial Load	An initial load is the process of loading an initial set of data on a Client Node from the Root Node

Installing the Client Node(s)

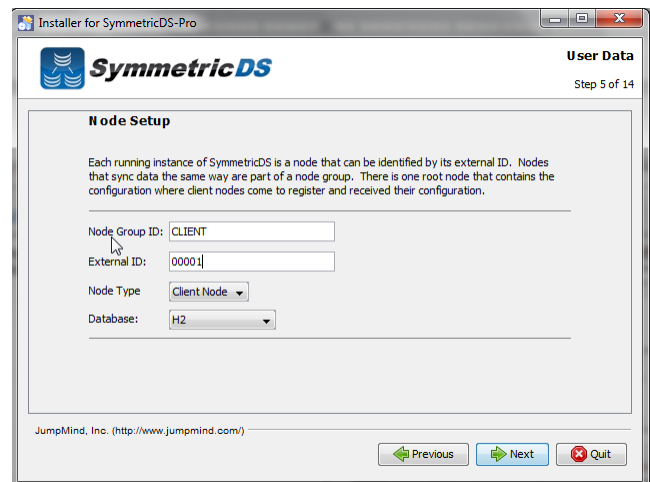|	
The steps for installing Client Nodes start exactly the same as installing the Root Node. • Verify that you have the 1.5.x Java Runtime Environment (JRE) installed on the Root Node computer. • Verify that the 1.5.x JRE bin directory is in your path (windows) or <what do I need to put here for Mac, Linux or Unix> • Run the SymmetricDS installer by double clicking the symmetric-pro-2.4.x-install.jar	
Review and accept the license agreement	
Select the installation directory where the SymmetricDS application should be installed	
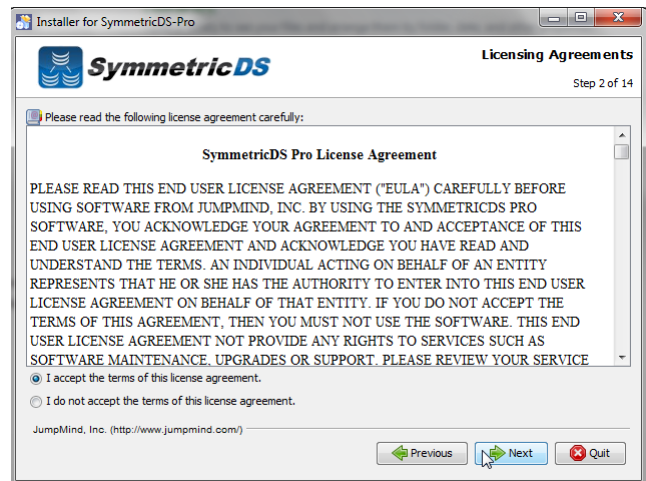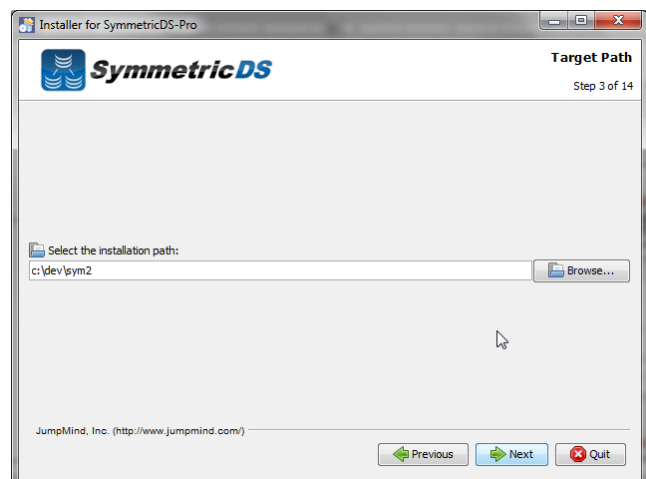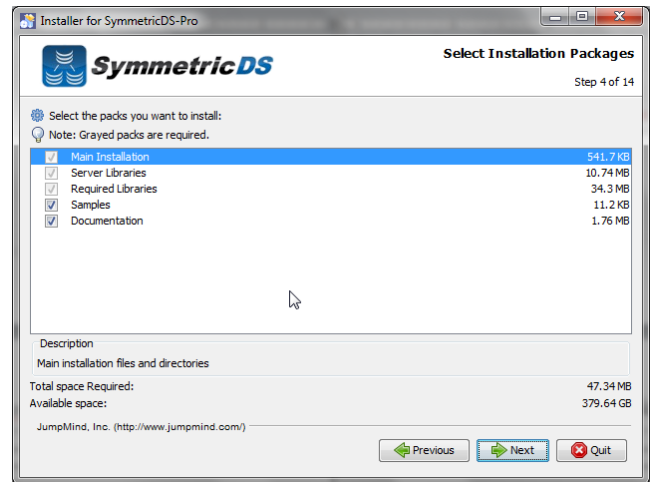
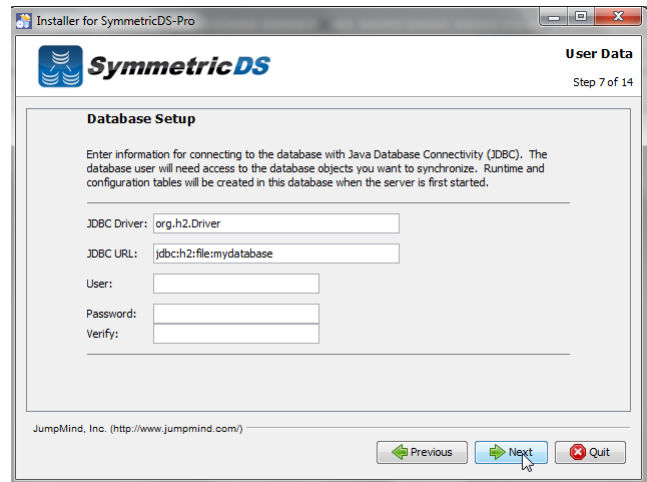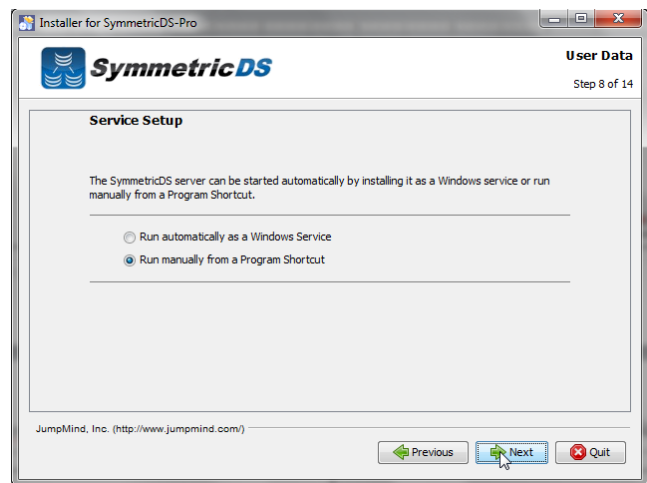|	
Select the components you would like to install.	
At this point, the installation steps are slightly different for the Client Node. • **Node Group Id** –This is the unique Identifier for the Node Group that will contain the Client Node. This is the Node Group that you defined in your synchronization scenario for your Client Nodes. • **External Id** – The unique identifier for this Client Node itself (each Node needs to have a unique identifier that allows other nodes in the synchronization scenario to identify it) • **Node Type** – Type type of Node that is being installed (either Root Node or Client Node). Select "Client Node." • **Database** – This is the type of database that will be used for the Client Database. Select your Client Node database type from the list	
The next step is to specify the Registration URL. This is the URL that the Root Node Registration Server is listening on for registration requests. This parameter will allow the Client Node to request permission to participate in the synchronization.	
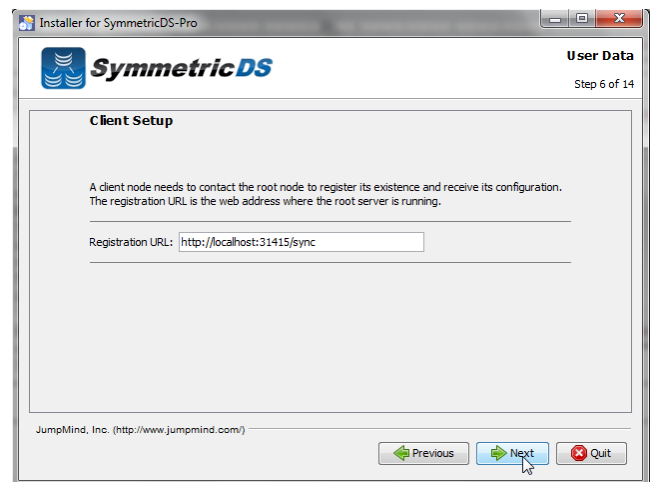
The next screen allows you to specify the connection information for the Client Database.
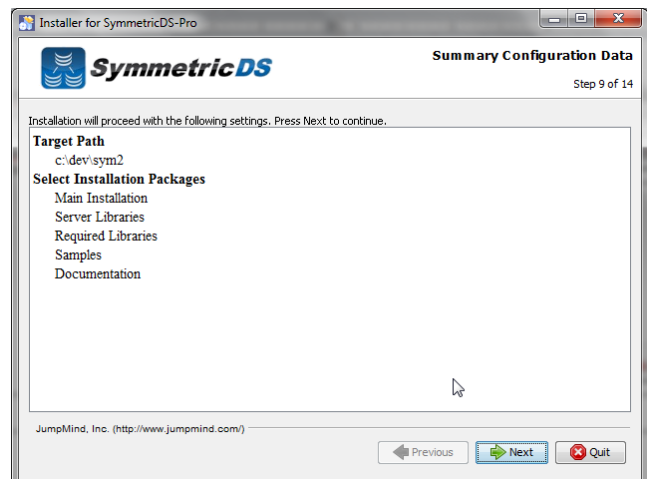
- **JDBC Driver** – The class name for your JDBC driver for your Root Database
- **JDBC URL** – The URL for your Root Database
- **User Id** – The user id for your Root Database. This Id will need Create, Read, Update, Delete (CRUD) privileges for tables
- **Password** – Associated password for the User ID for the Root Database
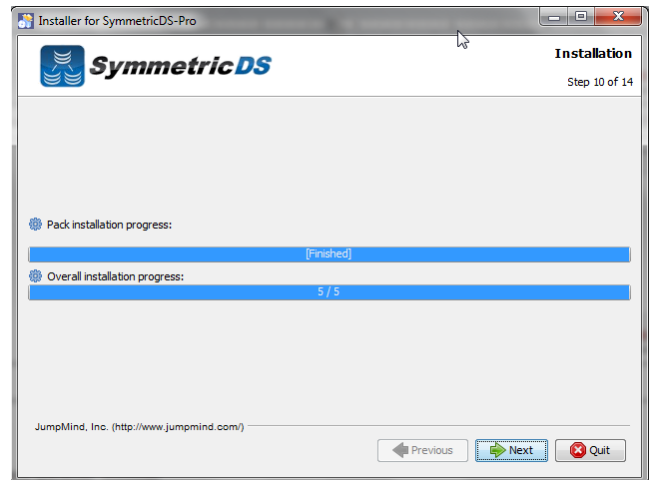- **Verify** – Verification of the password

SymmetricDS can be started manually from a Program Shortcut or as a Windows Service. Select one of the two options and click "Next."
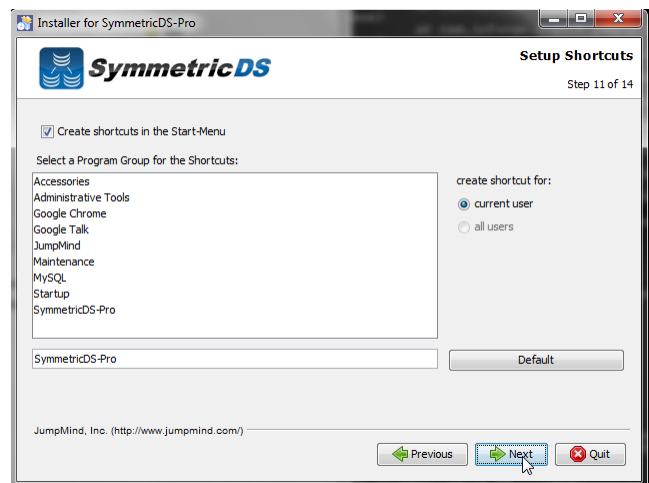
The Summary Configuration screen allows you to review your installation choices. Review and click "Next"

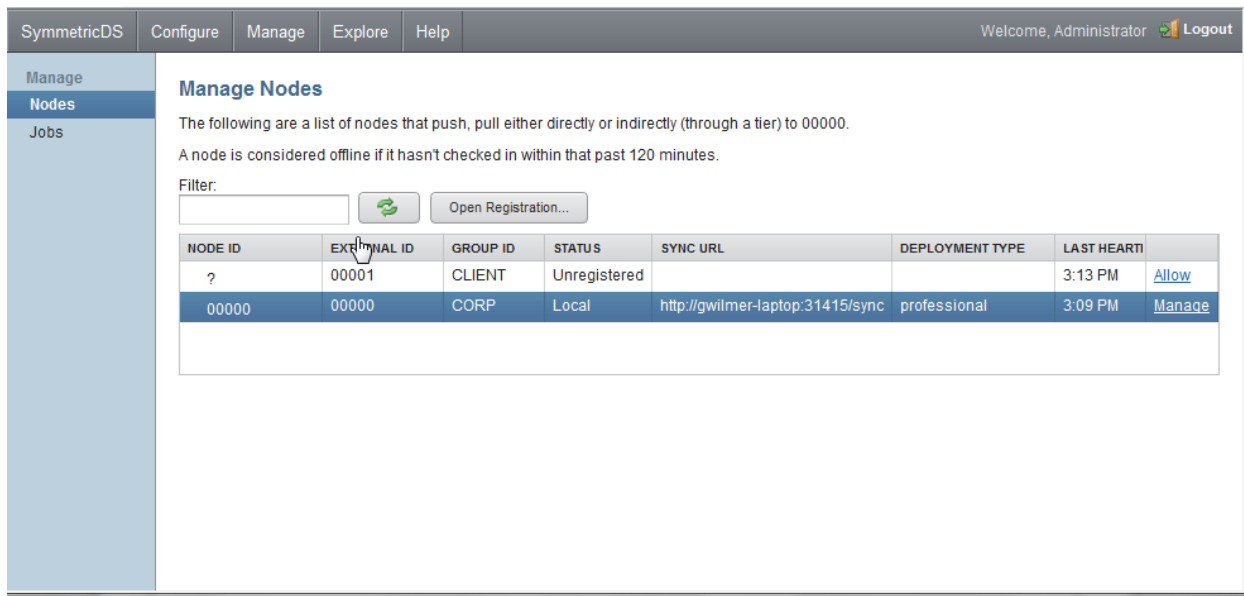|	
Once the application files have been copied to the installation directory, click the "Next" button.	
The Setup Shortcuts screen allows you to specify where you would like shortcuts created for the Client SymmetricDS instance. Select a group and click the "Next" button.	
Installation of the Client Node is now complete. Click the "Done" button.	
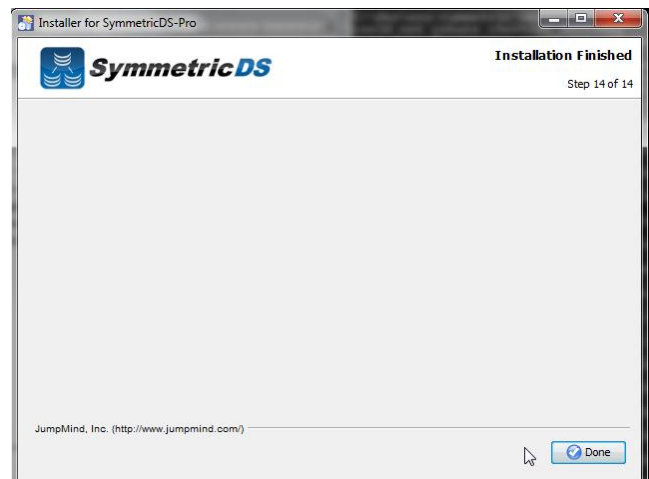
To start Client, either start the service (if you selected to run as a service), or start manually from the start menu.
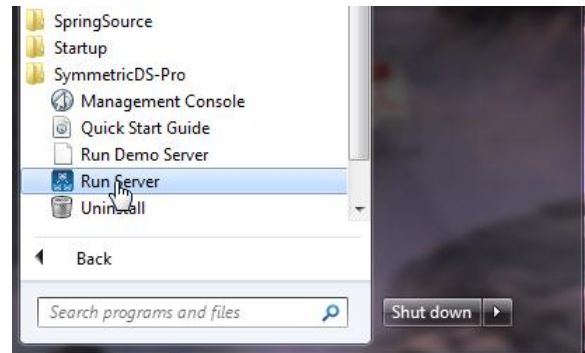
Registering the Client Node and Completing an Initial Load

Once the client has been started, it will automatically request registration from the registration server that you specified during the client install. With the default settings, permission to participate in the synchronization must be granted to the Client Node. Go to the Root Node web management console (http://<server_name>:<port>). Once logged into the console, click on the "Manage" tab.

In the screenshot above, you will notice two Nodes in the "Manage Nodes" listbox. The first row is the Client Node requesting registration to participate in the synchronization. The second row is the Root Server. To allow the Client Node to participate in the synchronization, click the "Allow" link on the Client Node row.

After clicking the "Allow" link, a dialog box will be displayed that asks you to confirm the registration request for the Client Node. You will see the Node Group (CLIENT) and the External ID (00001) for the Node that is requesting permission to register.
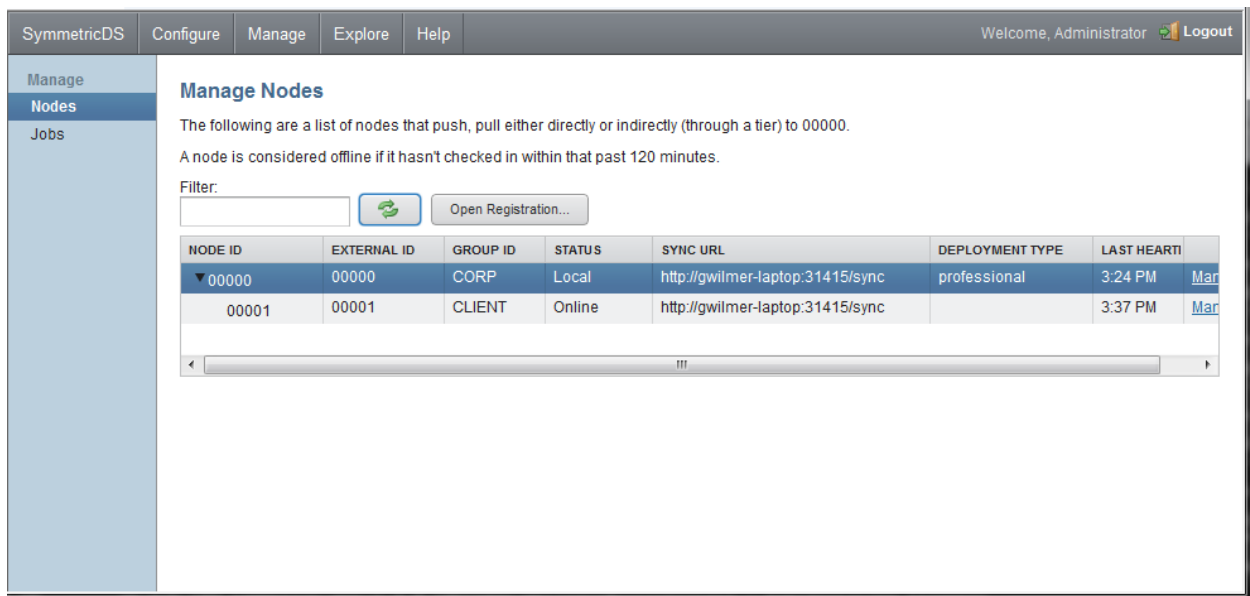
There is also an "Initial Load?" checkbox that allows you to determine whether this Client Node should get an initial load of data from the Root. By checking this box, any data captured by triggers since the install of SymmetricDS Pro on the Root Node and Root Database will be synchronized to the Client Node.

Once the Client Node has been successfully registered, the "Manage Nodes" screen will look as follows:

Common Questions

- **Installing JDBC Drivers** – SymmetricDS Pro ships with most JDBC drivers included in the install program. The Oracle JDBC driver is NOT included with the SymmetricDS Pro distribution. If you want to use Oracle as your Root or Client database, you must install the Oracle JDBC driver manually by copying the ojdbc<version>.jar file to the lib directory of your SymmetricDS Pro install directory.

- **Running Root and Client on the Same Computer** – For trial purposes, you may want to run both a Root Node and Client Node on the same computer. There are a couple of different ways to accomplish this. The first is to actually install the SymmetricDS Pro software twice, in two separate directories (one for the root, and one for the client). In this case, you will need to adjust the port number on either the Root Node or the Client Node as both can't be running at the same time on the same machine. In addition, if you install twice, only one shortcut will be created in your start menu, and it will reference the directory of the last install. SymmetricDS Pro can be run from the command line and started with different port numbers as follows:

  From the conf directory of the Root Node:
  ..\bin\sym.bat –p symmetric.properties –server

  From the conf directory of the Client Node:
  ..\bin\sym.bat –p symmetric.properties –P 31417 –server

  The –p argument specifies the properties file, and the –P property specifies the port number. Ensure that your properties file specifies the same port number that used when starting up the Nodes.

  The second option for running Root and Client on the same computer is to install once, and then create two separate properties files, one for the Root and another for the Client. When using this method, you can start the first instance from the start menu, but the second will need to be started from the command line or via a manually created option in the start menu. Creation of the Client Node properties file can be completed by copying the Root Properties file and making changes to the following configuration lines:

  group.id
  external.id
  registration.url
  sync.url
  db.driver
  db.url
  db.user
  db.password

  Remember to specify a different port on the Client Node.

- **Foreign Keys on Initial Loads** – For databases with complex relationships between tables, you need to ensure that tables that depend on each other are both in the same channel as well as are ordered properly such that data will be

inserted correctly such that key constraints are not violated. For example, if you have a header and line item table, you need to ensure that the header rows get inserted before the line items. This can be accomplished by specifying the proper initial load order on the "Configure", "Triggers" page for each table / trigger.