

Informatica Cloud & Redshift Getting Started User Guide

INFORMATICA CLOUD & REDSHIFT GETTING STARTED USER GUIDE	1
OVERVIEW	4
REDSHIFT CONNECTOR OVERVIEW	4
INFORMATICA CLOUD ARCHITECTURE	4
REDSHIFT CONNECTOR PREREQUISITES	5
DOWNLOADING AND INSTALLING THE VIBE SECURE AGENT	6
REDSHIFT CONNECTOR CONFIGURATION	7
GET YOUR AWS ACCOUNT SECRET KEY	7
GET YOUR REDSHIFT JDBC URL	9
CONFIGURE THE CONNECTOR PROPERTIES IN INFORMATICA CLOUD.....	10
USING THE DATA SYNCHRONIZATION WIZARD WITH REDSHIFT	12
CREATE YOUR DSS TASK.....	12
READING DATA FROM REDSHIFT	18
ODBC CONFIGURATION	19
SECURITY CONSIDERATIONS	19
CONFIGURING THE REDSHIFT CLUSTER VPC'S INBOUND IP SECURITY	19
CONFIGURING FOR REDSHIFT SSL.....	21
REDSHIFT CONNECTOR BEST PRACTICES.....	23



Overview

Amazon Web Services Redshift is a fast, fully managed, petabyte-scale data warehouse optimized for business intelligence. The Informatica Cloud Redshift Connector is a native, high-volume data connector enabling users to quickly and easily design petabyte-scale data integrations from any cloud or on premise sources to any number of Redshift nodes.

Getting started with Amazon Redshift is now easier than ever thanks to the Informatica Cloud 60 day trial for Amazon Redshift. Easily and quickly move data from all of your on premise and Cloud data sources, without writing a single line of code and without being a data integration expert. You can use our 6 step wizard to quickly replicate your data or use our intuitive web based designer to tackle more advanced use cases, such as combining multiple data sources into one Redshift table.

In this document we will cover all aspects of using Informatica Cloud to quickly get started with loading data into Redshift.

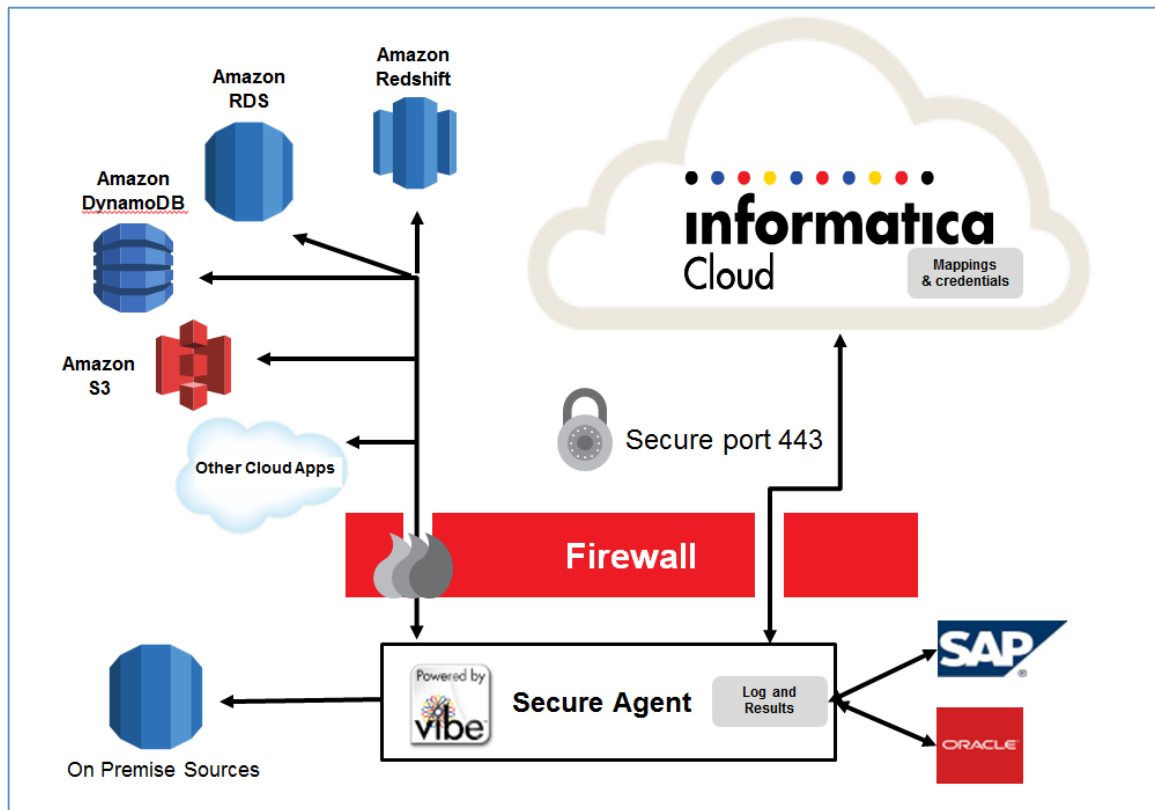
Redshift Connector Overview

The Redshift connector is a bulk-load type connector and allows you to perform inserts, deletes, and upserts (insert and/or update). Although Redshift does not natively support Upsert, the connector allows Upsert functionality by creating and loading a staging table first and then merging that with the existing table.

Access to Redshift data is available via ODBC or JDBC PostgreSQL drivers.

Informatica Cloud Architecture

The diagram below describes Informatica Cloud's high level architecture. It is important to note that none of your data flows through the cloud service; it all runs through the Vibe Secure Agent installed behind your firewall or on EC2.



Redshift Connector Prerequisites

Before using the Redshift connector you will need the following prerequisites:

- An Informatica Cloud user account. You can sign up for a trial here: <http://www.informaticacloud.com/>
- An Amazon Web Services (AWS) Account .You can sign up here: <http://aws.amazon.com/>
- If you are not familiar with Redshift, it is recommended to go through the Amazon Get Started Guide here: <http://docs.aws.amazon.com/redshift/latest/gsg/getting-started.html>
- A Redshift Cluster with a schema that your user has CREATE and USAGE privileges to. By default all users have those privileges with the “public” schema.
- The user name and password for your Redshift cluster. These are not the same as your AWS account user credentials.
- An S3 bucket in the same region as your Redshift cluster
- An Informatica cloud agent that has access to the Redshift cluster. **IMPORTANT!** The IP of your Informatica Cloud Secure agent will need to be in the access inbound list of the VPC for your Redshift cluster.
- If you are running the agent on Windows make sure the 2010 Visual Studio C++ redistributables are installed. Please see this link: <http://www.microsoft.com/en-us/download/details.aspx?id=5555>
- It is recommended you set up a new IAM user for using with your Redshift cluster and Informatica Cloud. More information about IAM can be found here: <http://aws.amazon.com/iam/>

Downloading and Installing the Vibe Secure Agent

To download and install the agent, follow these steps:

1. Click on the Configuration tab.
2. Click on Agents.
3. Click the Download Agent button:

Actions	Name	Host Name	Platform	Status	Version	Upgrade Status	Last Upgrade Check	Last S
	CAW178217	CAW178217	Windows	Inactive	17.3.12.26.0.0.0	Up-to-date	Jun 5, 2014 10:45 AM	Jun 5,
	CAW180967	CAW180967	Windows	Inactive	17.3.12.26.0.0.0	Up-to-date	Jun 23, 2014 1:00 PM	Jun 23
	ip-172-31-0-185	ip-172-31-0-185	Linux	Inactive	17.3.12.26.0.0.0	Up-to-date	Apr 4, 2014 1:27 PM	Apr 4,
	mxvminfa	mxvminfa	Windows	Inactive	17.3.12.26.0.0.0	Up-to-date	Apr 29, 2014 4:53 PM	Apr 29
	NICK-ICINQ1-MSSQL	WIN-RDVOFFCHU4G	Windows	Active	17.3.12.26.0.0.0	Up-to-date	Jun 27, 2014 1:55 PM	Jun 26

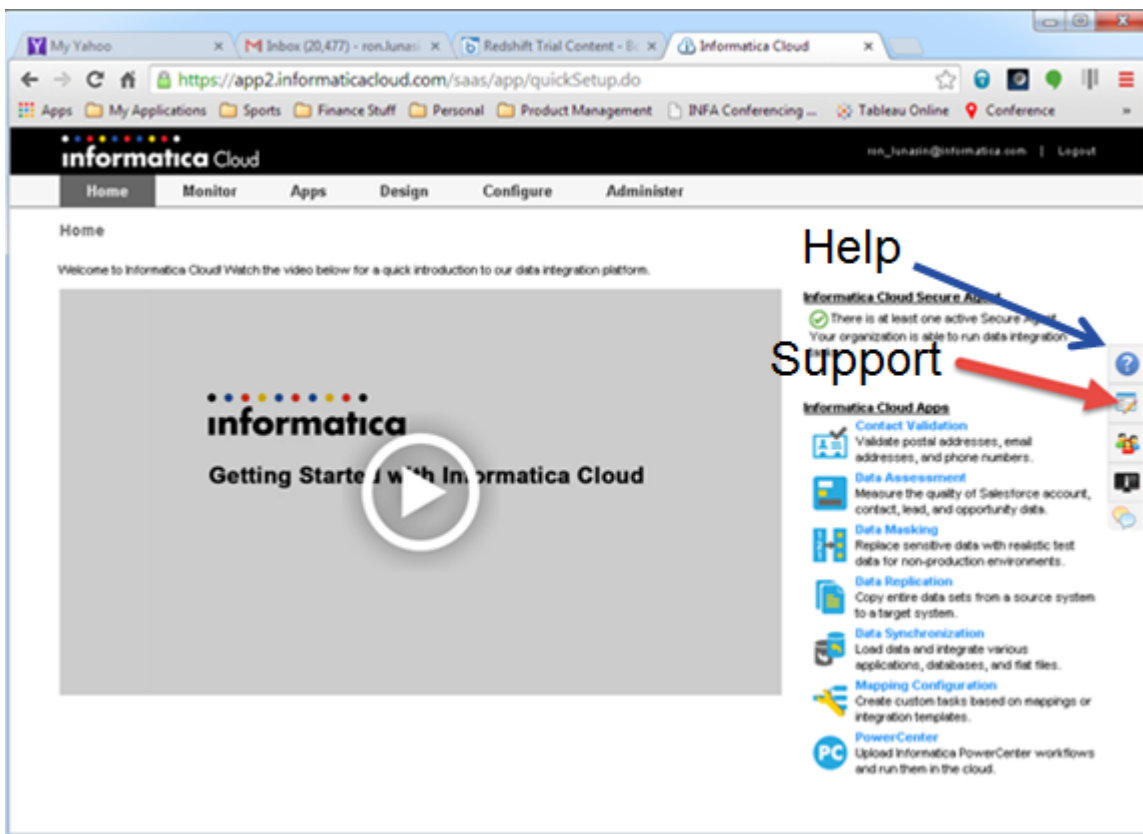
- 4.
5. Select your operating system (in this guide we use Windows), and click the Download button:
6. Click the Save button to save the installer to your local machine:
7. You will be prompted to select a location for the file on your local machine.
8. Select a location and click the Save button:
9. When the installer has finished downloading, locate the file - named "agent_install.exe" - and double-click it to start the installation.
10. Click the Run button and follow the remaining steps in the installation wizard.
 - a. We recommend that you accept the installation default values.
11. A registration page appears.
12. Enter your Informatica Cloud user name and password and click Register
13. The Secure Agent starts.
14. The Informatica Cloud Secure Agent window displays the status of the Secure Agent. You can restart, stop, and configure the Secure Agent proxy in this window. You can

close the window at any time. The Secure Agent continues to run as a service until stopped.

15. The Secure Agent Manager minimizes to the Windows taskbar notification area. Closing the Secure Agent Manager does not affect the Secure Agent status.

Help & Support

Informatica Cloud provides a number of getting started videos which are available in the Home tab of your Informatica Cloud Account. You can also click on the Help icon (see blue arrow below) from any page to access the online help documentation. If you need further assistance, click on the Support icon shown below (shown with the red arrow).

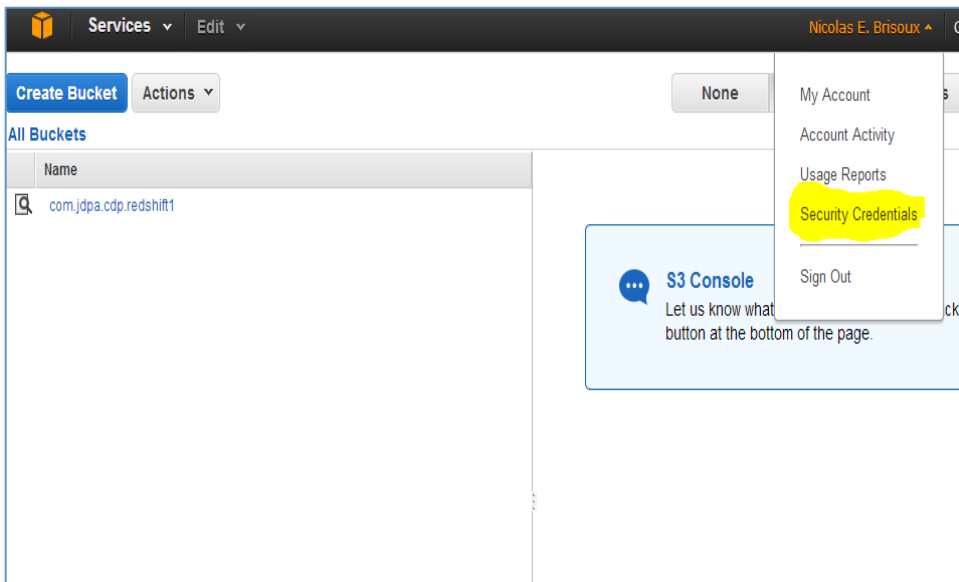


Redshift Connector Configuration

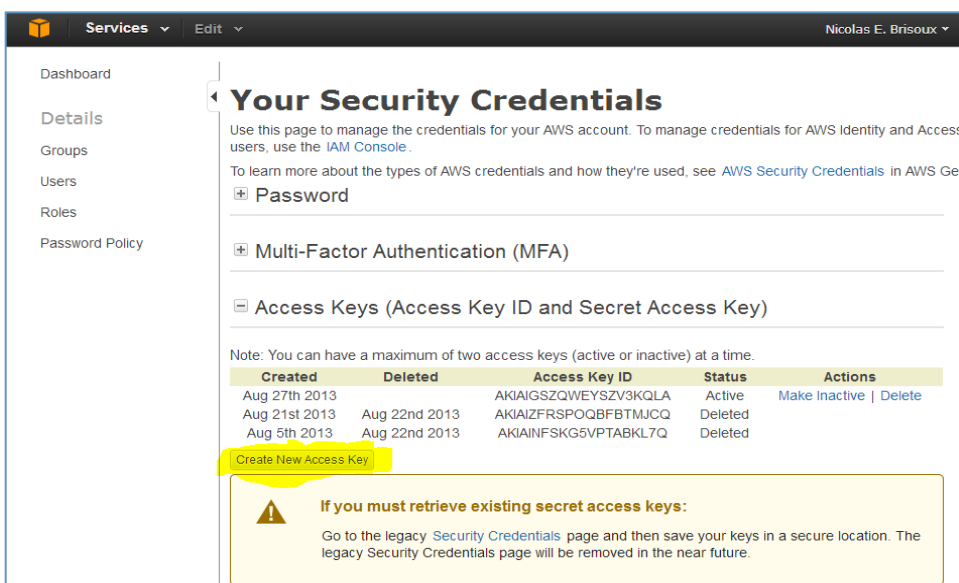
In order to configure the Redshift connector you will need to follow the steps below.

Get your AWS account secret key

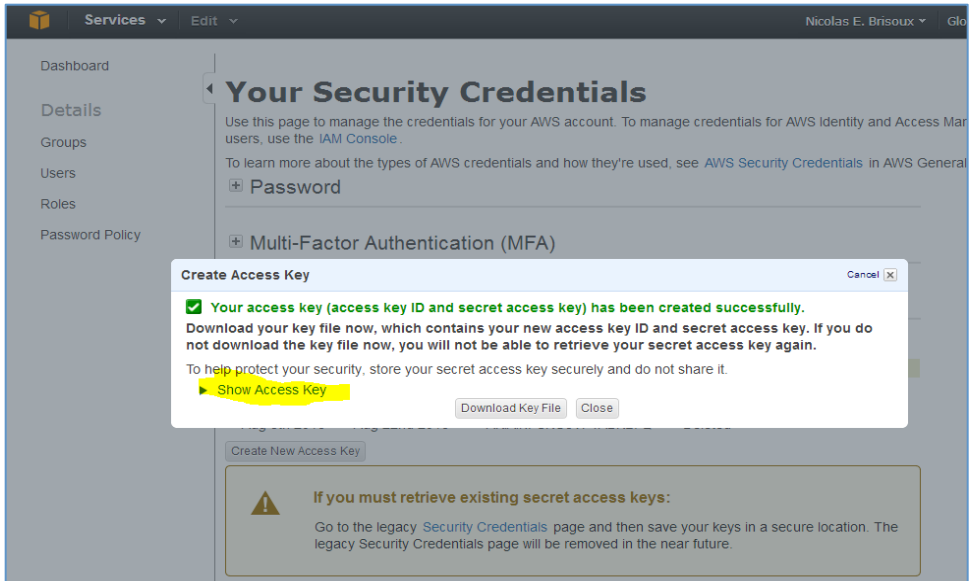
1. Go to your AWS account Security Credentials console as shown below:



- 2.
3. Click on the Continue to Security Credentials button in the next dialog
4. Once in the console, expand the Access Keys section, and click on the Create New Access Key button:

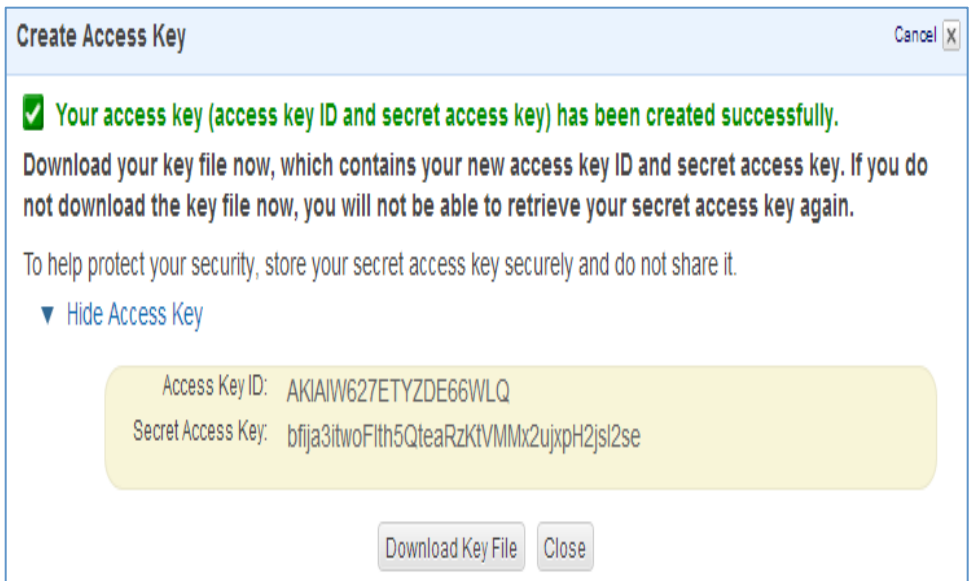


- 5.
6. The following screen will appear. Click on the Show Access Key link



7.

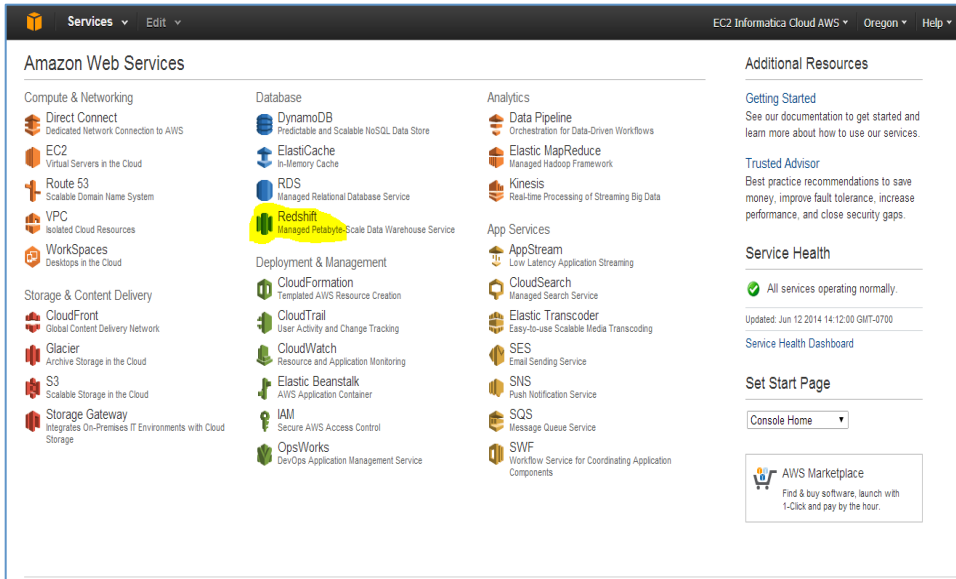
8. Note down your Access Key ID and the Secret Access Key:



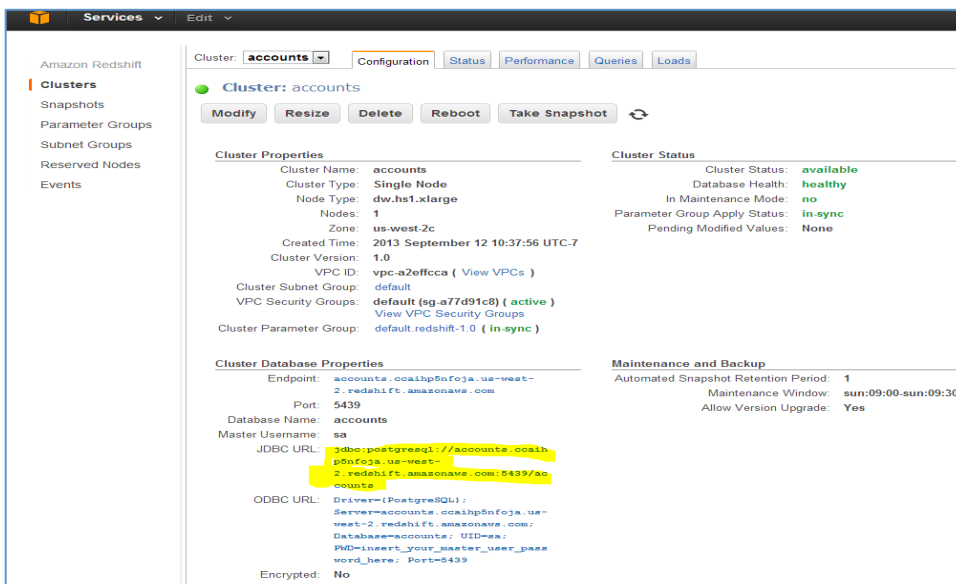
9.

Get your Redshift JDBC URL

1. Go to the AWS management console: <https://console.aws.amazon.com/console/home> and from there go to the Redshift management page.



- 2.
3. Bring up your cluster properties.
4. Note down the JDBC URL as shown below:



- 5.

Configure the connector properties in Informatica Cloud

1. Log in to your Informatica Cloud account and go to your Connections page and click on New.

The screenshot shows the Informatica Cloud interface with the 'Configure' tab selected. Under the 'Connections' section, a 'New' button is highlighted. Below it is a table listing existing connections:

Actions	Name▲	Type	Service URL
	aimsaddressapi	Webservices (Informatica Cloud Labs)	
	Amazon Redshift	AmazonRedshift (Informatica Cloud)	
	Amazon Redshift ADV2012	AmazonRedshift (Informatica Cloud)	
	Amazon Redshift Cloud Agent	AmazonRedshift (Informatica Cloud)	
	Amazon Redshift UBM	AmazonRedshift (Informatica Cloud)	
	Amazon S3	Amazon S3 (Informatica Cloud)	
	AWS RDS - Adventure Works 2012	SqlServer 2008	advworks2012.cktm8yelczaw.us-we
	aws rds ms sql	SqlServer 2008	infaawrdsmsql.cbxs97dlyix9.us-we

- 2.
3. Select Amazon Redshift as your connection type

The 'New Connection' dialog box is shown with the following fields and options:

- Buttons:** OK, Cancel, Test
- Connection Details:**
 - Connection Name: *
 - Description:
 - Type: * AmazonRedshift (Informatica Cloud) ▼
- AmazonRedshift Connection Properties:**
 - Secure Agent: * ?
 - Username: *
 - Password: *
 - Schema: * public
 - AWS Access Key ID: *
 - AWS Secret Access Key: *
 - Cluster Node Type: * XL ▼
 - Number of Nodes in the Cluster: *
 - Jdbc URL: *

- 4.
5. Enter the Redshift cluster username and password
6. Enter the schema name. If you did not create a specific schema for your cluster, you can use the "public" one.
7. Enter the cluster type, number of nodes, and the JDBC URL. See below for an example.

Edit Connection

You can edit the connection here. Use a connection to access a database, file or Salesforce. Er

OK **Cancel** **Test**

Connection Details

Connection Name:* awsredshift

Description: amazon redshift

Type:* RedShift (Evangelists)

RedShift Connection Properties

Secure Agent:* CAW180967

Username:* sa

Password:*

Connection URL:* jdbc:postgresql://redshiftwebinar.ccaihp5nfoja.u

Schema:* public

AWS Access Key ID:* AKIAIN*****

AWS Secret Access Key:* 8K6tUa1wuscAwQR*****

Cluster Node Type:* XL

Number of Nodes in the Cluster:* 1

8. .
9. Click on the Test button to make sure you can connect to the Redshift Cluster.

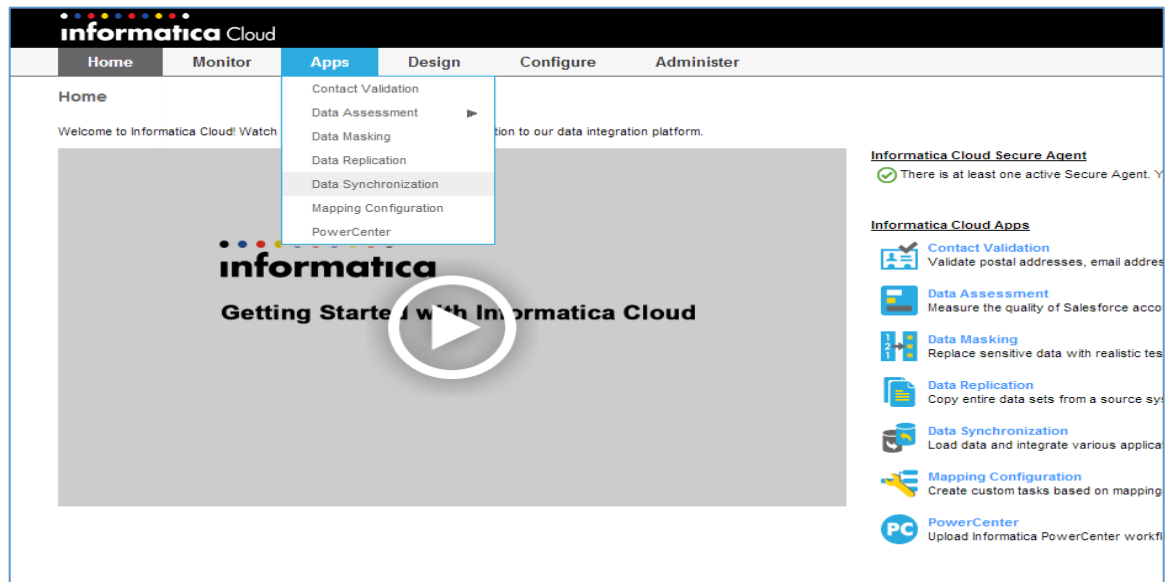
Using The Data Synchronization Wizard With Redshift

The Informatica Cloud data synchronization service (DSS) application delivers all of the key bi-directional synchronize data integration functions you need – and all through an intuitive web-based wizard. You can perform data transformation through a drag and drop web interface, perform lookups, as well as automate the running of your jobs on an hourly or to the minute schedule.

The guide below will show how to configure your first DSS task to load data into Redshift.

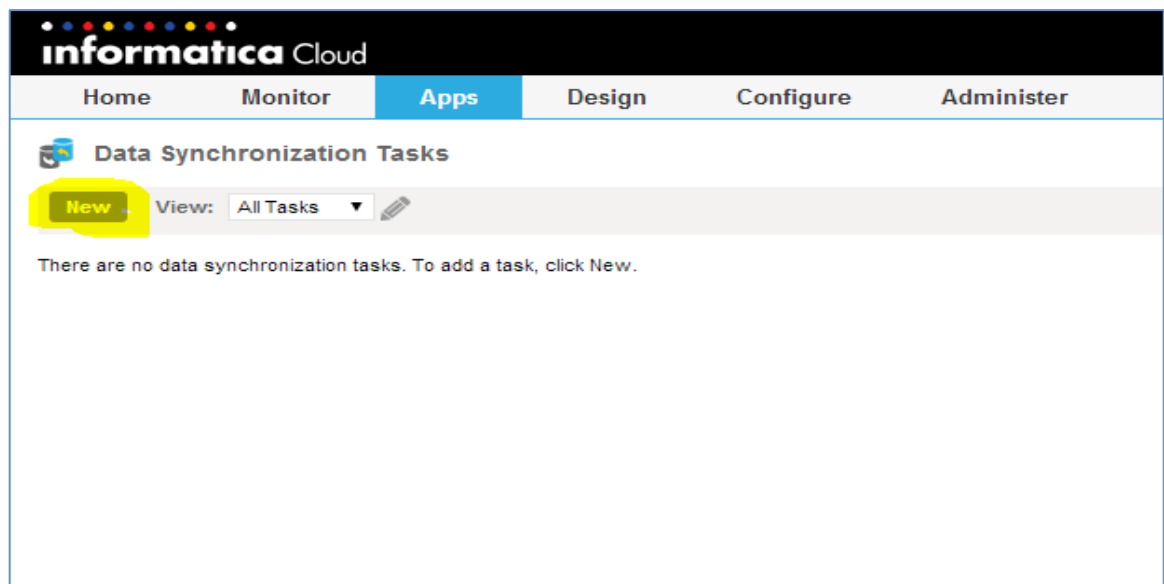
Create Your DSS Task

1. Go to the Apps menu and select the Data Synchronization application



2.

3. Click the "New" button.



4.

5. Choose a name for your task and from the Task Operation drop down selection box and choose "Insert"

informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard

1 Definition 2 Source 3 Target 4 Data Filters 5 Field Mapping 6 Schedule

< Previous Next > Save Cancel

Task Details

Task Name:* mytask ?

Description: ?

Task Operation:* Insert ?

- 6.
7. Click the "Next" button.
8. Choose your source connection for the data you will be loading into Redshift. Below is an example.
9. Pick your RS connection as the connection type and click on the "Create Target" button.

informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard (mytask)

1 Definition 2 Source 3 Target 4 Data Filters 5 Field Mapping 6 Schedule

< Previous Next > Save Cancel

Target Details

Connection:* Amazon Redshift EDW View... New... ?

Target Object:* Select... Show Data Preview Create Target... ?

Child Object: Show Data Preview ?

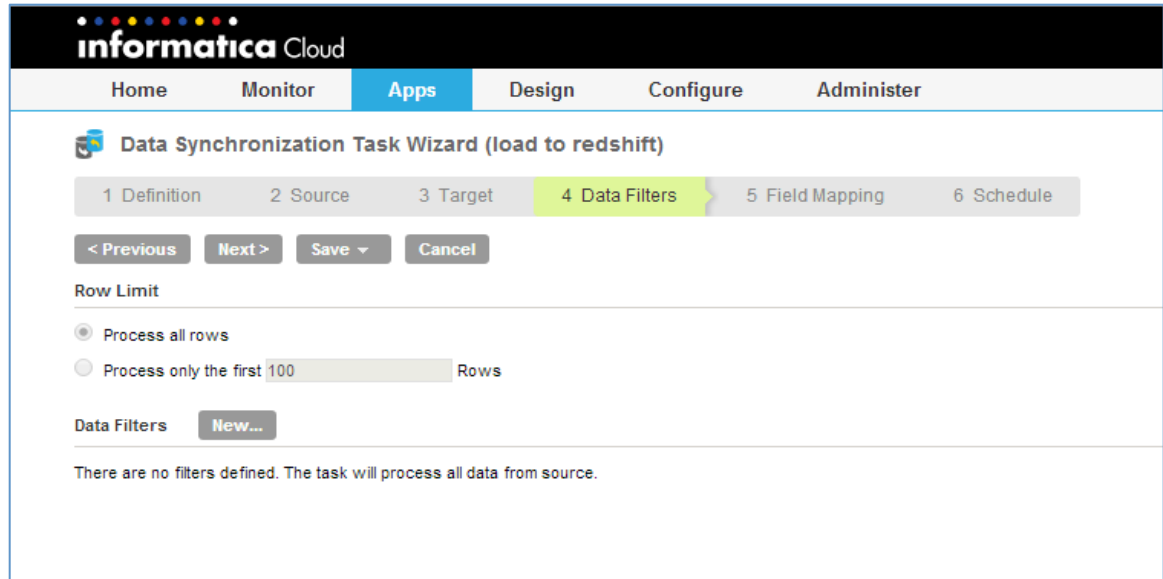
☒ Display technical names instead of labels

☐ Display target fields in alphabetical order

Data Preview

Select an object to preview its data.

- 10.
11. In Step 4 you can specify a source filter. This is optional. Click on the "Next" button.



informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard (load to redshift)

1 Definition 2 Source 3 Target **4 Data Filters** 5 Field Mapping 6 Schedule

< Previous Next > Save Cancel

Row Limit

☒ Process all rows

☐ Process only the first 100 Rows

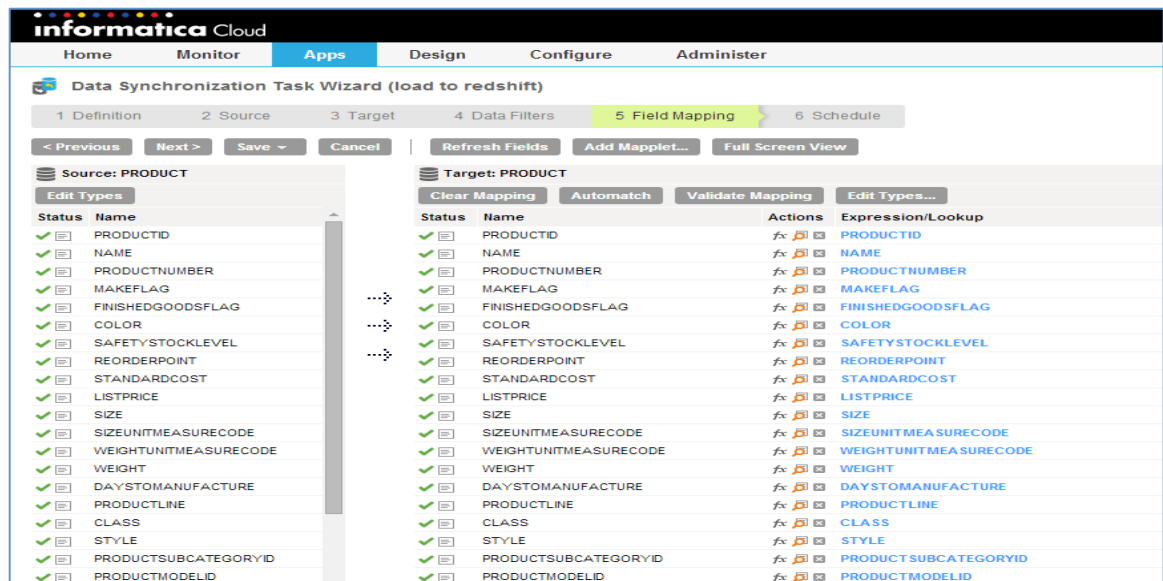
Data Filters New...

There are no filters defined. The task will process all data from source.

12.

13. In Step 5, shown below, you specify the mapping via the drag and drop interface or by using the "Automatch" feature. You can also apply transformations or do lookups. You can get more information on how to do this by taking a look at the following training video:

<http://asdasd.asdasd.com>



informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard (load to redshift)

1 Definition 2 Source 3 Target 4 Data Filters **5 Field Mapping** 6 Schedule

< Previous Next > Save Cancel Refresh Fields Add Maplet... Full Screen View

Source: PRODUCT

Status	Name
<input checked="" type="checkbox"/>	PRODUCTID
<input checked="" type="checkbox"/>	NAME
<input checked="" type="checkbox"/>	PRODUCTNUMBER
<input checked="" type="checkbox"/>	MAKEFLAG
<input checked="" type="checkbox"/>	FINISHEDGOODSFLAG
<input checked="" type="checkbox"/>	COLOR
<input checked="" type="checkbox"/>	SAFETYSTOCKLEVEL
<input checked="" type="checkbox"/>	REORDERPOINT
<input checked="" type="checkbox"/>	STANDARDCOST
<input checked="" type="checkbox"/>	LISTPRICE
<input checked="" type="checkbox"/>	SIZE
<input checked="" type="checkbox"/>	SIZEUNITMEASURECODE
<input checked="" type="checkbox"/>	WEIGHTUNITMEASURECODE
<input checked="" type="checkbox"/>	WEIGHT
<input checked="" type="checkbox"/>	DAYSTOMANUFACTURE
<input checked="" type="checkbox"/>	PRODUCTLINE
<input checked="" type="checkbox"/>	CLASS
<input checked="" type="checkbox"/>	STYLE
<input checked="" type="checkbox"/>	PRODUCTSUBCATEGORYID
<input checked="" type="checkbox"/>	PRODUCTMODELID

Target: PRODUCT

Status	Name	Actions	Expression/Lookup
<input checked="" type="checkbox"/>	PRODUCTID	fx	PRODUCTID
<input checked="" type="checkbox"/>	NAME	fx	NAME
<input checked="" type="checkbox"/>	PRODUCTNUMBER	fx	PRODUCTNUMBER
<input checked="" type="checkbox"/>	MAKEFLAG	fx	MAKEFLAG
<input checked="" type="checkbox"/>	FINISHEDGOODSFLAG	fx	FINISHEDGOODSFLAG
<input checked="" type="checkbox"/>	COLOR	fx	COLOR
<input checked="" type="checkbox"/>	SAFETYSTOCKLEVEL	fx	SAFETYSTOCKLEVEL
<input checked="" type="checkbox"/>	REORDERPOINT	fx	REORDERPOINT
<input checked="" type="checkbox"/>	STANDARDCOST	fx	STANDARDCOST
<input checked="" type="checkbox"/>	LISTPRICE	fx	LISTPRICE
<input checked="" type="checkbox"/>	SIZE	fx	SIZE
<input checked="" type="checkbox"/>	SIZEUNITMEASURECODE	fx	SIZEUNITMEASURECODE
<input checked="" type="checkbox"/>	WEIGHTUNITMEASURECODE	fx	WEIGHTUNITMEASURECODE
<input checked="" type="checkbox"/>	WEIGHT	fx	WEIGHT
<input checked="" type="checkbox"/>	DAYSTOMANUFACTURE	fx	DAYSTOMANUFACTURE
<input checked="" type="checkbox"/>	PRODUCTLINE	fx	PRODUCTLINE
<input checked="" type="checkbox"/>	CLASS	fx	CLASS
<input checked="" type="checkbox"/>	STYLE	fx	STYLE
<input checked="" type="checkbox"/>	PRODUCTSUBCATEGORYID	fx	PRODUCTSUBCATEGORYID
<input checked="" type="checkbox"/>	PRODUCTMODELID	fx	PRODUCTMODELID

14.

15. In the last step, Step 6, you can choose to run the task immediately or run it on a schedule.

informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard (load to redshift)

1 Definition 2 Source 3 Target 4 Data Filters 5 Field Mapping 6 Schedule

< Previous Next > Save Cancel

Schedule Details

☒ Do not run this task on a schedule ?

☐ Run this task on schedule: [dropdown] Now... ?

☐ Run this task in real-time upon receiving an outbound message from Salesforce ?
You must select a single Salesforce source object in Step 2 to use this option

Email Notification Options

☒ Use the default email notification options for my organization

☐ Use custom email notification options for this task:

Failure Email Notification: [text area]

Warning Email Notification: [text area]

Success Email Notification: [text area]

16.

17. Before we run the task however, we need to enter some additional information specific to Redshift. Under the "Advanced Options" enter the S3 bucket name and the folder location for the Secure Agent to use to stage the files it will upload to S3.

Advanced Options

Preprocessing Commands: [text area] ?

Postprocessing Commands: [text area] ?

Parameter File Name: [text area] ?

Advanced Target Properties

S3 Bucket Name: infarsbucket ?

Enable Compression: ☒ ?

Staging Directory Location: c:\temp ?

Batch Size: 10000 ?

Max Redshift Errors per Upload Batch for INSERT: 1 ?

Truncate Target Table Before Data Load: ☐ ?

Null value for CHAR and VARCHAR data types: [text area] ?

Wait time in seconds for file consistency on S3: 5 ?

18.

19. You can now run the task by selecting the "Save and Run" menu option from the "Save" menu.

informatica Cloud

Home Monitor **Apps** Design Configure Administer

Data Synchronization Task Wizard (load to redshift)

1 Definition 2 Source 3 Target 4 Data Filters 5 Field Mapping **6 Schedule**

< Previous Next > Save Save and Close Save and Continue Save and Run Cancel

Schedule Details

☒ Do not run this task on a schedule

☐ Run this task on schedule: [dropdown] New... ?

☐ Run this task in real-time upon receiving an outbound message from Salesforce ?
You must select a single Salesforce source object in Step 2 to use this option

Email Notification Options

☒ Use the default email notification options for my organization

☐ Use custom email notification options for this task:

Failure Email Notification: [text area]

20.

21. You will now be shown the Activity Monitor where you can see the running status of your task.

informatica Cloud

Home **Monitor** Apps Design Configure Administer

Activity Monitor

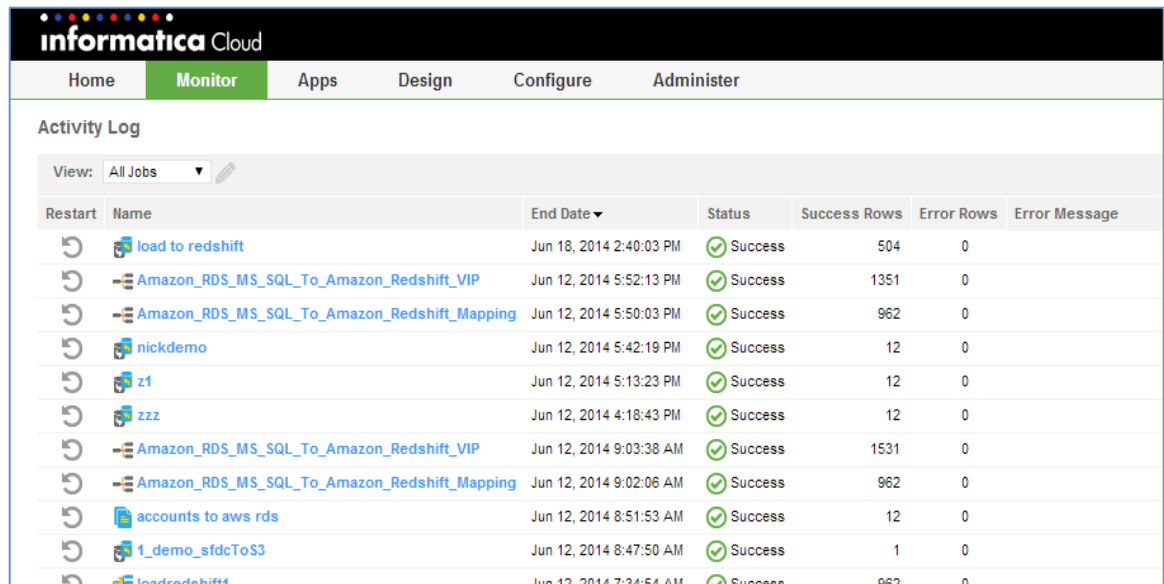
Refresh View: [list icon] [table icon]

Actions	Name	Type	Start Time
	load to redshift	Data Synchronization	Jun 18, 2014 2:38:53 PM

22.

23. Once the tasks complete you will be shown the Activity Log. Click on your task to get detailed information about the task results as well view the session log.

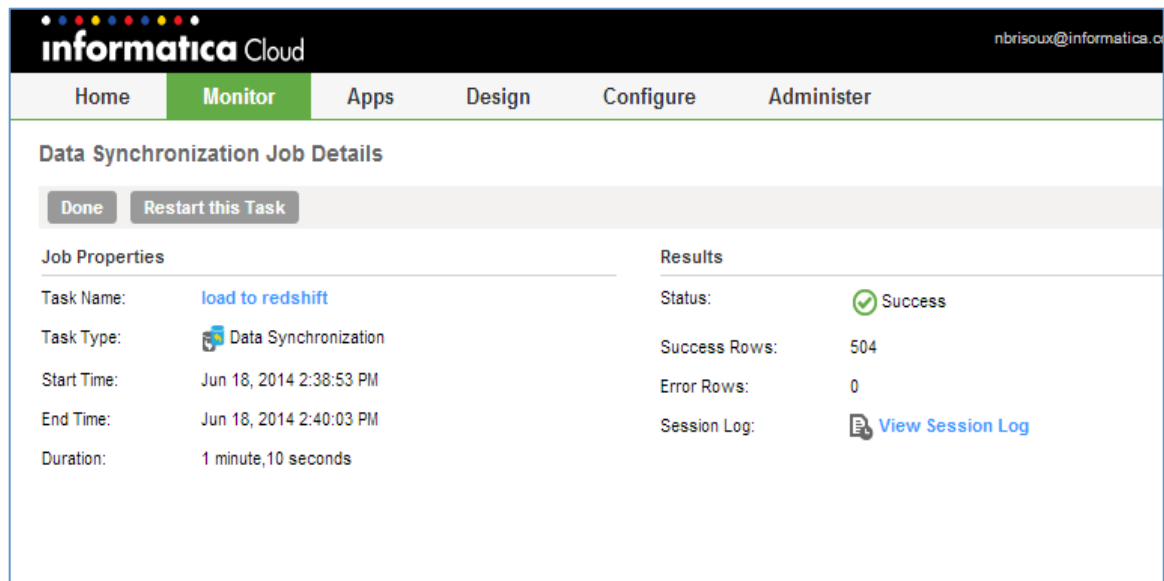
24.



The screenshot shows the Informatica Cloud 'Monitor' tab with the 'Activity Log' section. A dropdown menu is set to 'All Jobs'. Below is a table listing various data synchronization jobs.

Restart	Name	End Date	Status	Success Rows	Error Rows	Error Message
	load to redshift	Jun 18, 2014 2:40:03 PM	Success	504	0	
	Amazon_RDS_MS_SQL_To_Amazon_Redshift_VIP	Jun 12, 2014 5:52:13 PM	Success	1351	0	
	Amazon_RDS_MS_SQL_To_Amazon_Redshift_Mapping	Jun 12, 2014 5:50:03 PM	Success	962	0	
	nickdemo	Jun 12, 2014 5:42:19 PM	Success	12	0	
	z1	Jun 12, 2014 5:13:23 PM	Success	12	0	
	zzz	Jun 12, 2014 4:18:43 PM	Success	12	0	
	Amazon_RDS_MS_SQL_To_Amazon_Redshift_VIP	Jun 12, 2014 9:03:38 AM	Success	1531	0	
	Amazon_RDS_MS_SQL_To_Amazon_Redshift_Mapping	Jun 12, 2014 9:02:06 AM	Success	962	0	
	accounts to aws rds	Jun 12, 2014 8:51:53 AM	Success	12	0	
	1_demo_sfdcToS3	Jun 12, 2014 8:47:50 AM	Success	1	0	
	loadredshift4	Jun 12, 2014 7:34:54 AM	Success	962	0	

25.



The screenshot shows the 'Data Synchronization Job Details' page for the 'load to redshift' task. It includes buttons for 'Done' and 'Restart this Task'. The page is divided into 'Job Properties' and 'Results' sections.

Job Properties		Results	
Task Name:	load to redshift	Status:	Success
Task Type:	Data Synchronization	Success Rows:	504
Start Time:	Jun 18, 2014 2:38:53 PM	Error Rows:	0
End Time:	Jun 18, 2014 2:40:03 PM	Session Log:	View Session Log
Duration:	1 minute, 10 seconds		

Reading Data From Redshift

You can read data from using PostgreSQL JDBC or ODBC drivers (see the following Amazon documentation for detailed information: <http://docs.aws.amazon.com/redshift/latest/mgmt/configuring-connections.html>) In this section we will explain how to configure ODBC to work with Informatica Cloud. In these examples we will be using Windows. Refer to the PostgreSQL website (<http://www.postgresql.org/>) for how to configure these drivers for Linux.

ODBC Configuration

Security Considerations

Configuring The Redshift Cluster VPC's Inbound IP Security

1. Go to the Redshift cluster you will be using with the Informatica Cloud Agent.
2. From the Redshift cluster management panel click on the name of your redshift cluster.
3. You can go through the next steps even if your cluster isn't active yet
4. On the following screen, click on **View VPC Security Groups**

The screenshot displays the Amazon Redshift console interface. On the left, a navigation menu includes 'Amazon Redshift', 'Clusters', 'Snapshots', 'Parameter Groups', 'Subnet Groups', 'Reserved Nodes', and 'Events'. The 'Clusters' section is selected. The main panel shows the configuration for a cluster named 'accounts'. At the top, there are tabs for 'Configuration', 'Status', 'Performance', 'Queries', and 'Loads'. Below these are buttons for 'Modify', 'Resize', 'Delete', 'Reboot', and 'Take Snapshot'. The 'Cluster Properties' section lists: Cluster Name: accounts, Cluster Type: Single Node, Node Type: dw.hs1.xlarge, Nodes: 1, Zone: us-west-2c, Created Time: 2013 September 12 10:37:56 UTC-7, Cluster Version: 1.0, VPC ID: vpc-a2effcca (View VPCs), Cluster Subnet Group: default, VPC Security Groups: default (sg-a77d91c8) (active) (View VPC Security Groups), and Cluster Parameter Group: default.redshift-1.0 (in-sync). The 'Cluster Status' section shows: Cluster Status: available, Database Health: healthy, In Maintenance Mode: no, Parameter Group Apply Status: in-sync, and Pending Modified Values: None. The 'Cluster Database Properties' section shows the Endpoint: accounts.ccaihp5nfoja.us-west-. The 'Maintenance and Backup' section shows the Automated Snapshot Retention Period: 1.

a.

- b. You should see your default VPC group listed
5. Select the default VPC group, and a panel will appear as below

Services

Edit

Nicolas E. B

EC2 Dashboard

Events

Tags

INSTANCES

Instances

Spot Requests

Reserved Instances

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Snapshots

NETWORK & SECURITY

Security Groups

Elastic IPs

Placement Groups

Load Balancers

Key Pairs

Network Interfaces

Create Security Group

Delete

Viewing:

All Security Groups

Search

	Group ID	Name	VPC ID	Description
<input checked="" type="checkbox"/>	sg-a77d91c8	default	vpc-a2effcca	default VPC security group

1 Security Group selected

Security Group: default

Details

Inbound

Outbound

Group Name:

default

Group ID:

sg-a77d91c8

Group Description:

default VPC security group

VPC ID:

vpc-a2effcca

- a.
- b. You will need to add any IP you are going to run the Cloud Agent from from to the Inbound list. In the example below, we use Informatica HQ's external IP.

Services

Edit

Nicolas E. Brisoux

Oregon

EC2 Dashboard

Events

Tags

INSTANCES

Instances

Spot Requests

Reserved Instances

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Snapshots

NETWORK & SECURITY

Security Groups

Elastic IPs

Placement Groups

Load Balancers

Key Pairs

Network Interfaces

Create Security Group

Delete

Viewing:

All Security Groups

Search

	Group ID	Name	VPC ID	Description
<input checked="" type="checkbox"/>	sg-a77d91c8	default	vpc-a2effcca	default VPC security group

1 Security Group selected

Security Group: default

Details

Inbound

Outbound

Create a new rule:

All Traffic

Source:

12.108.188.134/32

(e.g., 192.168.2.0/24, sg-47ad482e, or 1234567890/default)

Add Rule

Apply Rule Changes

ALL Port (Service)	Source	Action
ALL	12.108.188.134/32	Delete
ALL	12.108.188.0/32	Delete
ALL	24.130.95.56/32	Delete
ALL	sg-a77d91c8	Delete

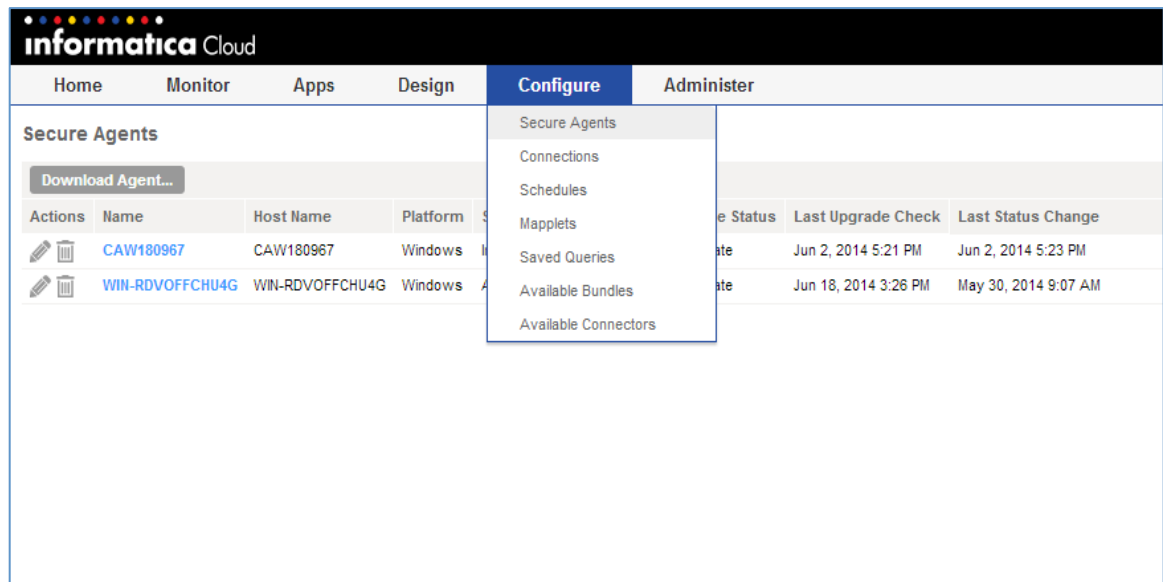
- i.
- c. Apply the rule changes

20

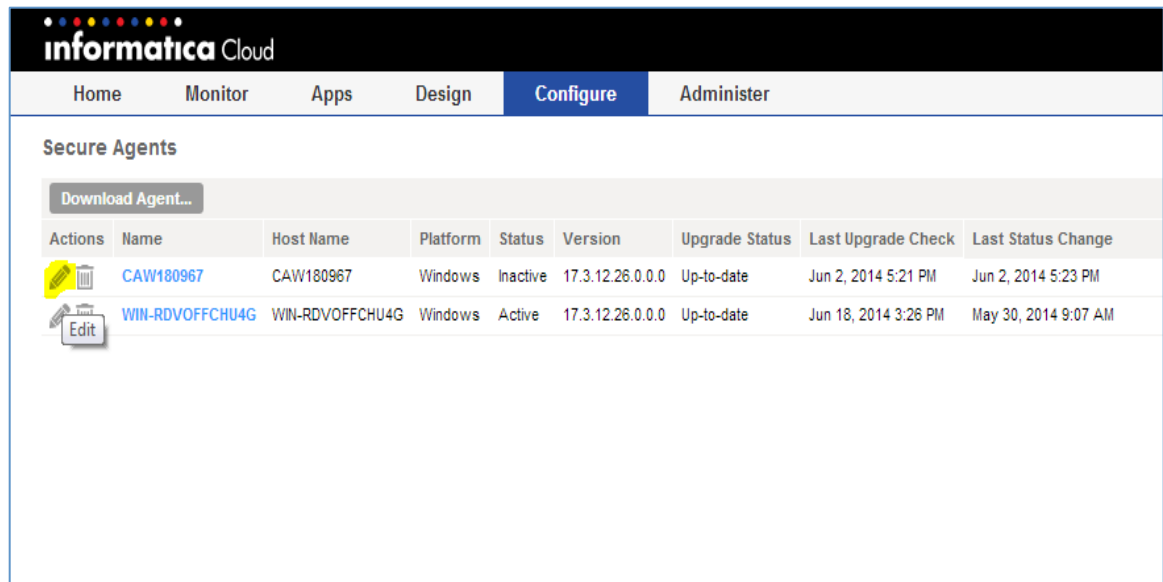
Configuring For Redshift SSL

The Secure Agent can be configured to support an SSL connection to Redshift. We recommend consulting the Amazon Redshift documentation on this topic (see <http://docs.aws.amazon.com/redshift/latest/mgmt/connecting-ssl-support.html#connecting-ssl-support-java>). The following steps outline how to configure your Secure Agent to run with an SSL connection.

1. First you will need to add the Amazon Redshift certificate to the Java system truststore. Download the certificate from <https://s3.amazonaws.com/redshift-downloads/redshift-ssl-ca-cert.pem>
2. Add the certificate to the key store by executing the following command:
`${JAVA_HOME}/bin/keytool -keystore {JAVA_HOME}/lib/security/cacerts -import -alias <alias> -file <certificate_filename>` Where <alias> is any user-provided string value and <certificate_filename> is the full path to the certificate file that you downloaded in Step 1.
3. You need to change the Secure Agent JVM startup properties to specify the keystore and password.
4. Go to your Secure Agents configuration page.

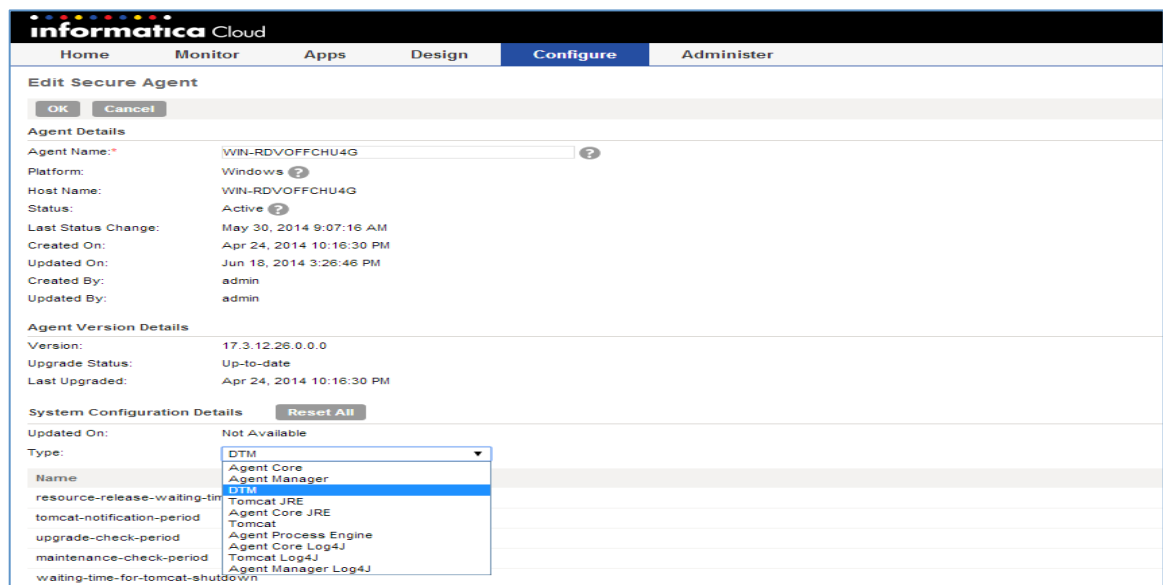


- 5.
6. Next, click on the edit button to left of your Secure Agent.



7.

8. In the System Configuration Details drop down box, change the Type to DTM



9.

10. Add the following to JVMOption1 and JVMOption2: -
 Djavax.net.ssl.trustStore=<keystore_name> and -
 Djavax.net.ssl.trustStorePassword=<password>. Here <keystore_name> is
 cacerts or the keystore you have created manually.

System Configuration Details Reset All

Updated On: Not Available

Type: DTM

Name	Value	
OptimizeODBCWrite	No	
__PMOV_FFW_ESCAPE_QUOTE	Yes	
RecordSessStatInRepo	No	
JVMOption2	-Djavax.net.ssl.trustStore	
JVMOption1	-Djavax.net.ssl.trustStore	
RepositoryName	XMLRepository	
PMPassword	rDTMPass	
PMUser	rDTMUser	
JVMClassPath	pmserversdk.jar;	
AgentConnectionTimeout	5	
SalesForceConnectionTimeout	300	
RepositoryDomain	rDTMDomain	

11.

12. Lastly, add a parameter to the JDBC URL you specified in your Redshift Connection properties, "ssl=true". See example below:
 jdbc:postgresql://mycluster.xyz789.us-west-2.redshift.amazonaws.com:5439/dev?ssl=true

Redshift Connector Best Practices

When working with the Redshift connector we recommend the following best practices.

1. Follow Amazon's best practices when designing your tables:
http://docs.aws.amazon.com/redshift/latest/dg/c_designing-tables-best-practices.html
2. Choose a batch size where the number of batches matches the number of slices in your cluster. Each XL node has 2 slices, each 8XL node has 16. If you have a 2 node XL cluster and 40,000 rows of data, choose a batch size of 10,000. The Informatica Cloud Redshift connector can maximize Amazon's parallel processing capabilities this way.
3. Only use the "upsert" when you know you will be updating rows. Otherwise use the "insert" capability as it will load the data more efficiently.

