

Connecting to Redshift

- Setup
 - Redshift Settings
 - Location
 - Create Cluster
 - Cluster Details
 - Node Configuration
 - Additional Configuration
 - Review & Launch
 - View Clusters
 - Cluster Name
 - Cluster Details
- Yellowfin
 - Login
 - Add Data Source
 - Data Source Details
 - Connection Details
- Section navigation
 - Current topic - Install in the Cloud
 - Install on Premises
 - Install in the Cloud
 - Install in a container
 - Deploy Yellowfin
 - Advanced Deployments

Setup

Redshift Settings

1. Log in to Amazon Web Services (AWS).

2. Open the **Services** menu at the top of the page, click on **Redshift** within the **Database** category.

The screenshot shows the AWS console 'Services' page. At the top, there's a navigation bar with the AWS logo, 'Services' dropdown, and 'Edit' dropdown. Below this, the 'Amazon Web Services' title is displayed. The services are organized into several categories:

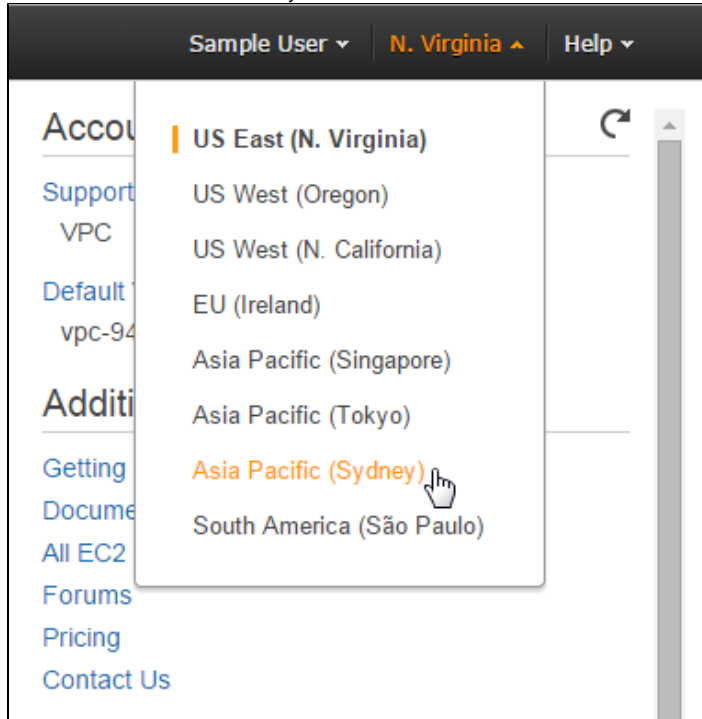
- Compute & Networking:** Direct Connect, EC2, Route 53, VPC.
- Storage & Content Delivery:** CloudFront, Glacier, S3, Storage Gateway.
- Database:** DynamoDB, ElastiCache, RDS, and **Redshift** (highlighted with a mouse cursor). Redshift is described as 'Managed Petabyte-Scale Data Warehouse Service'.
- Deployment & Management:** CloudFormation, CloudTrail, CloudWatch, Elastic Beanstalk, IAM, OpsWorks, and Trusted Advisor.
- Analytics:** Data Pipeline, Elastic MapReduce, and Kinesis.
- Mobile Services:** Cognito, Mobile Analytics, and SNS.

Location

While on the Amazon Redshift page it's important to select the closest location to your users. This will generally result in the fastest response times.

3. Click on the **Location** currently in use next to your name at the top right of the page.


4. Select the location closest to you.



Create Cluster

5. While still on the Amazon Redshift page, click on the **Launch Cluster** button.

From here you will be walked through a setup process where you define and configure the cluster.



Services ▾

Edit ▾

Amazon Redshift

Clusters


Snapshots

Security

Parameter Groups

Reserved Nodes

Events



Welcome to Amazon Redshift

You do not appear to have any clusters in the US East (N. Virginia) region.

Amazon Redshift is a fast and powerful, fully managed, petabyte-scale data ware query performance when analyzing virtually any size data set using the same SQL today. With a few clicks in the AWS Management Console, you can launch a Redshift cluster, scale it to a petabyte or more, for under \$1,000 per terabyte per year.

Try Amazon Redshift for free! If you've never created an Amazon Redshift cluster node. You get 750 hours per month for free, enough hours to continuously run or can also build clusters with multiple nodes to test larger data sets, which will consume 750 hours or your usage exceeds 750 hours per month, you can shut down your cluster. For more information, please see the [free trial FAQ page](#).

Launch Cluster


Cluster Details

On this step you will need to define the following:

- Cluster Identifier
- Database Name
- Database Port
- Master User Name
- Master Password

Each option is documented on the page.

6. Configure the cluster details and click the **Continue** button.



Services ▾

Edit ▾

Amazon Redshift

Clusters

Snapshots

Security

Parameter Groups

Reserved Nodes

Events

CLUSTER DETAILS

NODE CONFIGURATION

ADDITIONAL CONFIGURATION

REVIEW

Provide the details of your cluster. Fields marked with * are required.

Cluster Identifier*

my-yf-aws-db-instance

This is the unique key that identifies a cluster. This parameter is stored as a lowercase string. (e.g. my-dw-instance)

Database Name

yellowfindb

Optional. A default database named dev is created for the cluster. Optionally, specify a custom database name (e.g. mydb) to create an additional database.

Database Port*

5439

Port number on which the database accepts connections.

Master User Name*

master_user

Name of master user for your cluster. (e.g. awsuser)

Master User Password*

.....

Password must contain 8 to 64 printable ASCII characters excluding: /, ", ', \, and @. It must contain 1 uppercase letter, 1 lowercase letter, and 1 number.

Confirm Password*

.....

Confirm Master User Password.

Cancel

Continue

Node Configuration

Here you will need to define the number and types of nodes. Each option is documented on the page.

7. Define the **Type** and **Number of Nodes**. Click **Continue**.

Services Edit

Amazon Redshift

Clusters

Snapshots

Security

Parameter Groups

Reserved Nodes

Events

CLUSTER DETAILS

NODE CONFIGURATION

ADDITIONAL CONFIGURATION

REVIEW

Choose a number of nodes and Node Type below. Number of Compute Nodes is required for multi-node clusters.

Node Type

dw2.large

Specifies the compute, memory, storage, and I/O capacity of the cluster's nodes.

CPU

7 EC2 Compute Units (2 virtual cores) per node

Memory

15 GiB per node

Storage

160GB SSD storage per node

I/O Performance

Moderate

Cluster Type

Single Node

Single Node clusters consist of a single node which performs both leader and compute functions.

Number of Compute Nodes*

1

Maximum

1

Minimum

1

Cancel

Previous

Continue

Additional Configuration

Finally, there are some additional configuration items you can define, depending on the your network and security requirements. Each option is documented on the page.

8. Click the **Continue** button.

Services Edit

Amazon Redshift

Clusters

Snapshots

Security

Parameter Groups

Reserved Nodes

Events

CLUSTER DETAILS

NODE CONFIGURATION

ADDITIONAL CONFIGURATION

REVIEW

Provide the optional additional configuration details below.

Cluster Parameter Group

A default parameter group will be associated with this cluster.

Encrypt Database

No

Select Yes to encrypt all data within the cluster and in backups at a small cost to performance.

Use HSM

No

You have not created any HSM Connections. You must create an HSM Connection to use HSM. You must also create at least one HSM Client Certificate.

Configure Networking Options:

Choose a VPC

Default VPC (vpc-94b90bf1)

The identifier of the VPC in which you want to create your cluster

Cluster Subnet Group

default

Selected Cluster Subnet Group may limit the choice of Availability Zones

Publicly Accessible

Yes

Select Yes if you want the cluster to be accessible from the public internet. Select No if you want it to be accessible only from within your private VPC network

Choose a Public IP Address

No

Select Yes if you want to select your own public IP address from a list of elastic IP (EIP) addresses that are already configured for your cluster's VPC. Select No if you want Amazon Redshift to provide an EIP for you instead.

Availability Zone

No Preference

The EC2 Availability Zone that the cluster will be created in.

Optionally, associate your cluster with one or more security groups.

VPC Security Groups

launch-wizard-2 (sg-b4eec...)

default (sg-162f0c73)

launch-wizard-1 (sg-4eb79...)

launch-wizard-3 (sg-01182...)

List of VPC Security Groups to associate with this cluster.

Optionally, create a basic alarm for this cluster.

Create CloudWatch Alarm

Yes

No

Create a CloudWatch alarm to monitor the disk usage of your cluster.

Cancel

Previous

Continue

Review & Launch

On the final step you are able to review all the configuration options you've applied in the previous steps.

9. Click the **Launch Cluster** option to finish.

You are about to launch a cluster with following the following specifications:

Cluster Properties

These attributes specify the name of your cluster, what type of virtual hardware it will run on, how many nodes it will contain, and the availability zone in which it will be located.

Cluster Identifier: my-yf-aws-db-instance

Node Type: dw2.large

Number of Compute Nodes: 1 (leader and compute run on a single node)

Availability Zone: No Preference

Database Configuration

These properties specify the database name, port, and username you will use to connect to the database. The parameter group contains configuration values used by the database.

Database Name: yellowfindb

Database Port: 5439

Master User Name: master_user

Cluster Parameter Group: A default parameter group will be created when the cluster is launched.

Security, Access, and Encryption

These settings control whether your cluster will be created in an existing VPC to allow for simpler integration with other AWS Services, and the security groups which define access rules to your cluster.

Virtual Private Cloud: vpc-94b90bf1

Cluster Subnet Group:

Publicly Accessible: Yes

Elastic IP: Not used

VPC Security Groups: sg-b4eec4d1

Encrypt Database: No

Use HSM: No

CloudWatch Alarms

CloudWatch alarms are used to notify if metrics for your cluster are within a certain threshold. All recipients under the SNS topic specified for your alarm will receive notifications once an alarm is triggered.

Basic alarms will not be created for this cluster.



Unless you are eligible for the free trial, you will start accruing charges as soon as your cluster is active.

Applicable charges:

The on-demand hourly rate for this cluster will be **\$0.25**, or **\$0.25/node**. If you have purchased reserved nodes in this region for this node type that are active, your costs will be discounted. Additional nodes will be billed at the on-demand rate.

If you are eligible for a free trial, you will receive 750 hours of free usage for each month of the trial, applied across all running dw2.large nodes across all regions. Regardless of when you start your trial, you will receive two full months of free usage. Once your trial expires or your usage exceeds 750 hours/month, you can shut down your cluster, avoiding any charges, or keep it running at our standard [On-Demand Rate](#).

For more information, see [Amazon Redshift Free Trial FAQ](#), [Amazon Redshift Pricing](#), and [Reserved Nodes Documentation](#).

Cancel

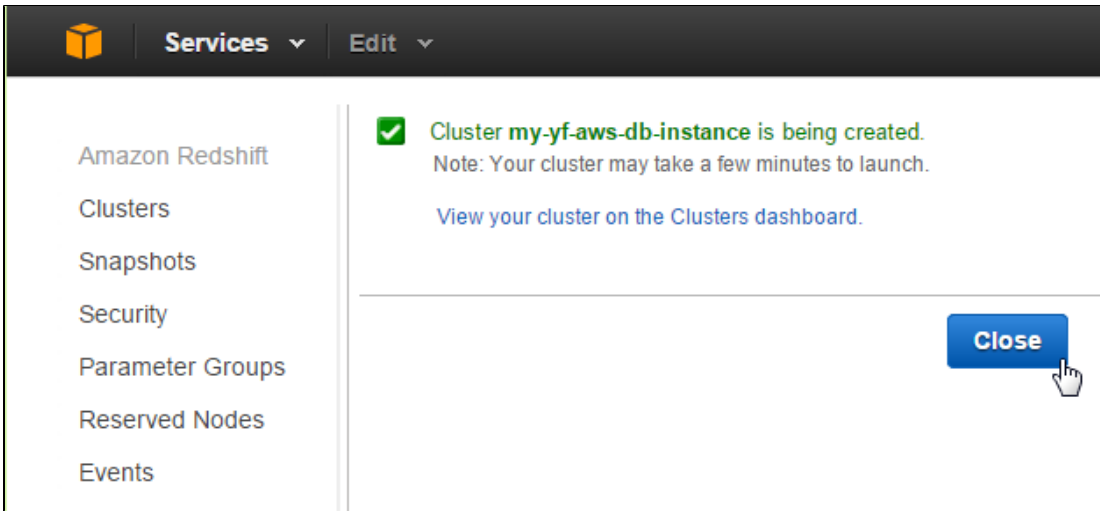
Previous

Launch Cluster

View Clusters

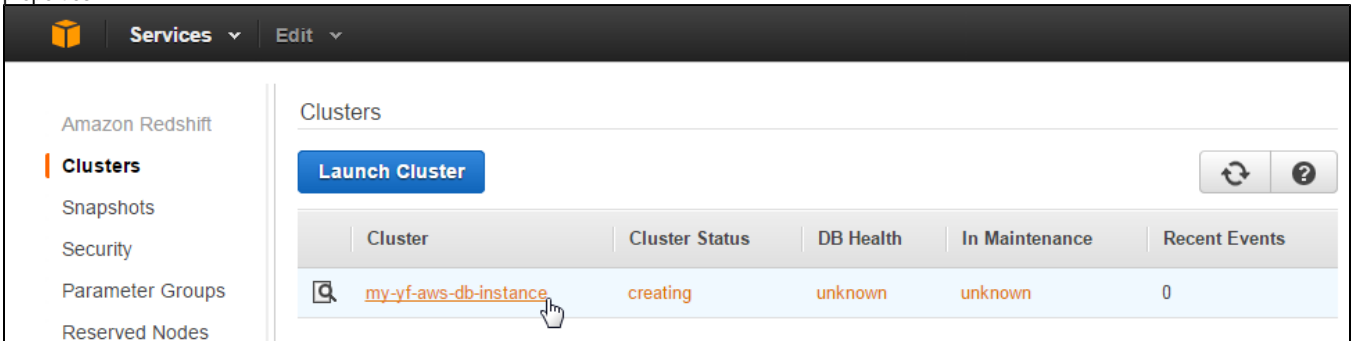
You will now be presented with a message informing you that your cluster is being created.

10. Click the **Close** button to return to a list of available clusters.



Cluster Name

Click on the name of your cluster in order to view its properties.



Cluster Details

Here you will be able to view all the details related to your cluster, including items important for Yellowfin use:

- Cluster Name
- End Point
- Port
- Database Name
- Master User Name
- Public IP

Make a note of the above values.

Services

Edit

Amazon Redshift

Clusters

Snapshots

Security

Parameter Groups

Reserved Nodes

Events

Cluster: yf

Configuration

Status

Performance

Queries

Loads

Cluster: yf

Cluster

Database

Backup

Cluster Properties

Cluster Name: yf

Cluster Type: Single Node

Node Type: dw2.large

Nodes: 1

Zone: ap-southeast-2a

Created Time: October 7, 2014 4:37:29 PM UTC+11

Cluster Version: 1.0.829

VPC ID: vpc-63e90206 (View VPCs)

Cluster Subnet Group: default

VPC Security Groups: default (sg-55e52530) (active)

Cluster Parameter Group: default.redshift-1.0 (in-sync)

Cluster Status

Cluster Status: available

Database Health: healthy

In Maintenance Mode: no

Parameter Group Apply Status: in-sync

Pending Modified Values: None

Cluster Database Properties

Endpoint: yf.ccjapojk5x0n.ap-southeast-2.redshift.amazonaws.com

Port: 5439

Publicly Accessible: Yes

Database Name: yellowfin

Master Username: master_user

Encrypted: No

JDBC URL: jdbc:postgresql://my-yf-aws-db-instance.ccjapojk5x0n.ap-southeast-2.redshift.amazonaws.com:5439/yf?tcpKeepAlive=true

ODBC URL: Driver={PostgreSQL}; Server=my-yf-aws-db-instance.ccjapojk5x0n.ap-southeast-2.redshift.amazonaws.com; Database=yf; UID=yf; PWD=insert_your_master_user_password_here; Port=5439

Backup, Audit Logging, and Maintenance

Automated Snapshot Retention Period: 1

Cross-Region Snapshots Enabled: No

Audit Logging Enabled: No

Maintenance Window: fri:14:30-fri:15:00

Allow Version Upgrade: Yes

Capacity Details

Current Node Type: dw2.large

CPU: 7 EC2 Compute Units (2 virtual cores) per node

Memory: 15 GiB per node

Storage: 160GB SSD storage per node

I/O Performance: Moderate

Platform: 64-bit

SSH Ingestion Settings

Cluster Public Key:

ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAQDnu7VqEBecf8xawRLP1i
nuqWSoM2dnNh4FE+5d1WTVJ6coA//vRodPIuf4XE27P4VYWGCF
ZsoFW609E8LHDq5G9kUpL0DuHbT3ZuxIMk8TSq8wV5NE8qhGIP
0dQWmF8UA0Ddo04jpfQDkuFQhC5hcrR2CdQudPLPBtjQf7mkXE
uM0hecBFSaxbj52vnCf1FACQLJGgQ5iwOGAkPDDk2P4CP+F6F1
hdINDq+pdAtK2Vqy2q+ht9hL3KBXGTQMxVcIc5k7EtMH9RhPm
Hsfn7NYeX/buCyLpeNalKm1T57Lho4jdLh50MyI1+0KNpgcA01
agir6RimzLLO2xMfLNckD3 Amazon-Redshift

Node IP Addresses:

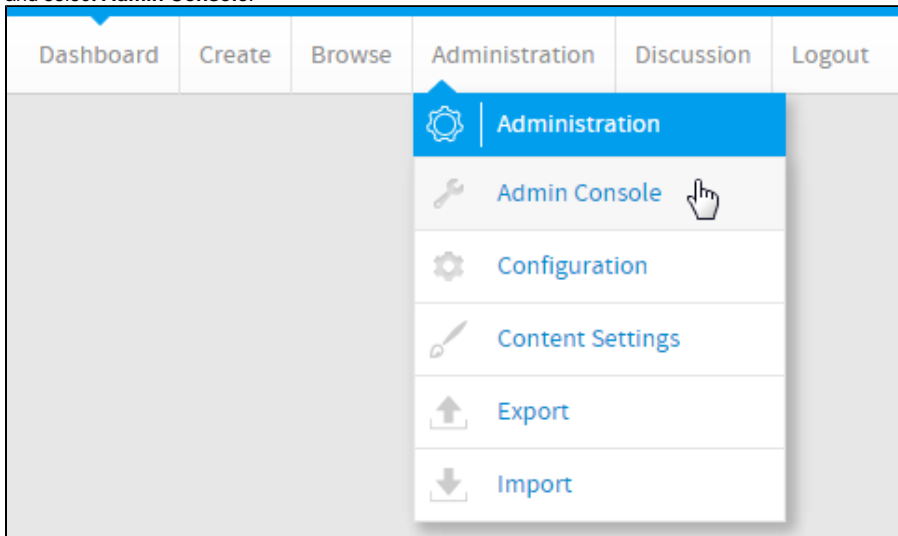
Node	Public IP	Private IP
Shared	54.66.138.222	172.31.22.15

Yellowfin

[Connecting to Redshift#top](#)

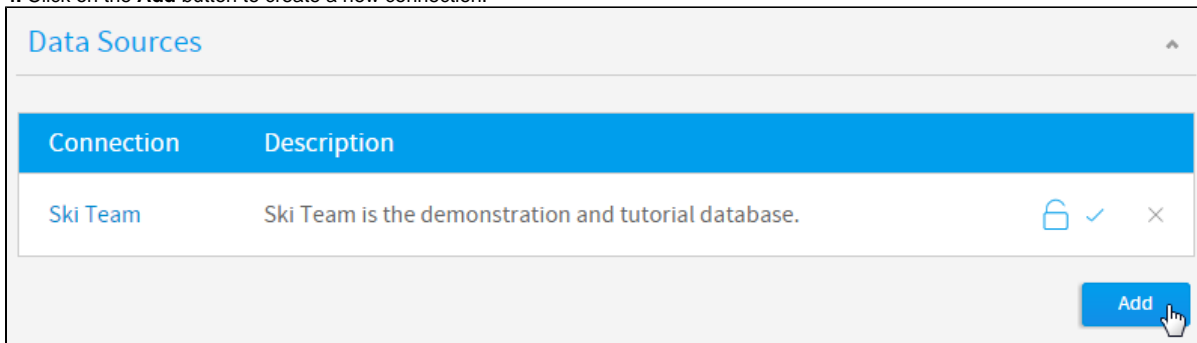
Login

1. Login to your Yellowfin instance.
2. Click on the **Administration** link in the main navigation bar and select **Admin Console**.



Add Data Source

3. Expand the **Data Sources** list in the centre panel of the Admin Console.
4. Click on the **Add** button to create a new connection.



Data Source Details

Here you will need to provide a range of information to tell Yellowfin what the connection is, and how to access the database.

5. Define the **Data Source Details**, **Security**, and **Connection Pool** options based on your requirements.

See [Understanding Connection Parameters](#) for more information.

Data Source Details	
Source Name:	Redshift
Description:	Redshift Connection
Max Rows Returned:	<input type="radio"/> Unlimited <input checked="" type="radio"/> 10000
Writable:	<input type="radio"/> Yes <input checked="" type="radio"/> No
Region:	Australia ▼
Time Zone:	Sydney ▼
Security	
Access Level:	<input checked="" type="radio"/> Unsecure <input type="radio"/> Secure
Permissions:	<input checked="" type="checkbox"/> Broadcast <input checked="" type="checkbox"/> Subscribe
Source Filters:	<input type="radio"/> Yes <input checked="" type="radio"/> No

Primary Connection Pool Management	
Min Connections:	1
Max Connections:	5
Refresh Time:	3 hours
Timeout:	180 seconds
Use secondary pool:	<input type="radio"/> Yes <input checked="" type="radio"/> No
Availability	
Availability:	<input checked="" type="checkbox"/> Available. The connection to this Data Source was successful. [Set Unavailable]
<div> <div>Save</div> <div>Cancel</div> </div>	

Connection Details

6. In order to connection to your Redshift cluster, define the following:

- **Connection Method:** JDBC
- **Authentication:** Standard Authentication
- **Database:** Amazon Redshift
- **Include schema in SQL:** (ticked)
- **Host:** *this is the **Public IP** or **Endpoint** information (either can be used, depending on the security settings you applied) found on your Cluster Details page earlier.*

- **Port:** this is the **Port** found on your Cluster Details page earlier.
- **Database:** this is the Database Name found on your Cluster Details page earlier.
- **User Name:** this is the **Master User Name** found on your Cluster Details page earlier.
- **Password:** this is the password you entered while configuring your cluster earlier.
- **Schema:** None

7. Test the connection and click **Save**.

Connection

Connection Method:

JDBC

Authentication:

Standard Authentication

Database:

Amazon Redshift

Include schema in SQL:

☒

Host:

yf.ccjapojk5x0n.a

Port:

5439

Database:

yf

User Name:

yf

Password:

.....

Schema:

None

Connection Succeeded

Database:

PostgreSQL

Product Version:

8.0.2

Driver:

PostgreSQL Native Driver

Driver Version:

PostgreSQL 9.3 JDBC4 (build 1101)

Database Version:

8.0

Source Name:

Amazon Redshift

Click here to test the connection again.

Database Schema:

No tables found.

You are now ready to add data to your Cluster and build reports.

[top](#)

Section navigation

Current topic - Install in the Cloud

This page is part of the [Install And Deploy Yellowfin](#) section of the wiki, which has these topics:

Install on Premises

[Connecting to Redshift](#)

- [Installation Steps](#)

Install in the Cloud

[Install in the Cloud](#)

- [Yellowfin for AWS](#)
- [Yellowfin for Azure](#)
- [Yellowfin for Google Cloud Platform](#)

Install in a container

[Install in a Container](#)

- [Docker](#)
- [Kubernetes](#)
- [Upgrading Yellowfin Container Deployment](#)

Deploy Yellowfin

[Deploy Yellowfin](#)

- [Logs and Logging](#)
- [Yellowfin Directory Structure](#)
- [User Welcome](#)

Advanced Deployments

[Advanced Deployments](#)

- [Clustering Guide](#)
- [Yellowfin Server Specification](#)
- [Automate Yellowfin Deployment on Linux](#)
- [SAML Bridge](#)
- [Standalone Configuration Tools](#)

[top](#)
