ClustrixDB AWS Installation Guide

This document provides step-by-step instructions on how to install and configure ClustrixDB in AWS using the ClustrixDB AMI. To proceed with this installation, you will also need a ClustrixDB license. If you do not yet have a ClustrixDB license, please contact Clustrix Sales. If you encounter any problems or issues while using these instructions, please contact Clustrix Support.

These instructions are a customized version of the steps found in the following ClustrixDB Online Documentation:

- Best Practices for ClustrixDB Platform Configuration
- ClustrixDB Installation Guide Bare OS Instructions

Table of Contents

Recommended Server Specifications

ClustrixDB AMI comes preinstalled with CentOS 7, ClustrixDB, and a few additional linux tools such as ntpd, htop, etc. For recommended server specifications for your particular deployment, please refer to ClustrixDB Reference Server Configurations.

Overview of Setup Steps

The installation of ClustrixDB follows these high-level phases:

1. Creating Placement and Security Groups
2. Launching of EC2 instances using the correct parameters
3. Formation of the ClustrixDB Cluster
4. Initialization of the ClustrixGUI
5. Creating the Elastic Load Balancer

Creating Placement and Security Groups

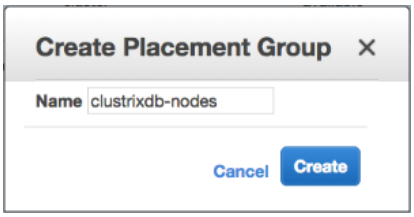
Log into the AWS Console, and select the EC2 service.

Creating a Placement Group

A Placement Group is used to ensure low latency and high throughput between nodes. It is a logical grouping of instances within an Availability Zone (see Regions and Availability Zones) which affects how close together the instances are created, from a networking standpoint. By placing all ClustrixDB instances in the same Placement Group you ensure the network latency is as low as possible between the instances. If you are using ClustrixDB Zones you will want to create one Placement Group for each Zone. See the Zones documentation for more information.

To create a Placement Group, select the “Placement Groups” link under Network & Security on the left-side menu.

Then click the “Create Placement Group” button and enter a name for the Placement Group. The name must be a name that has not been previously used by another Placement Group.



Click the “Create” button.

Creating a Security Group

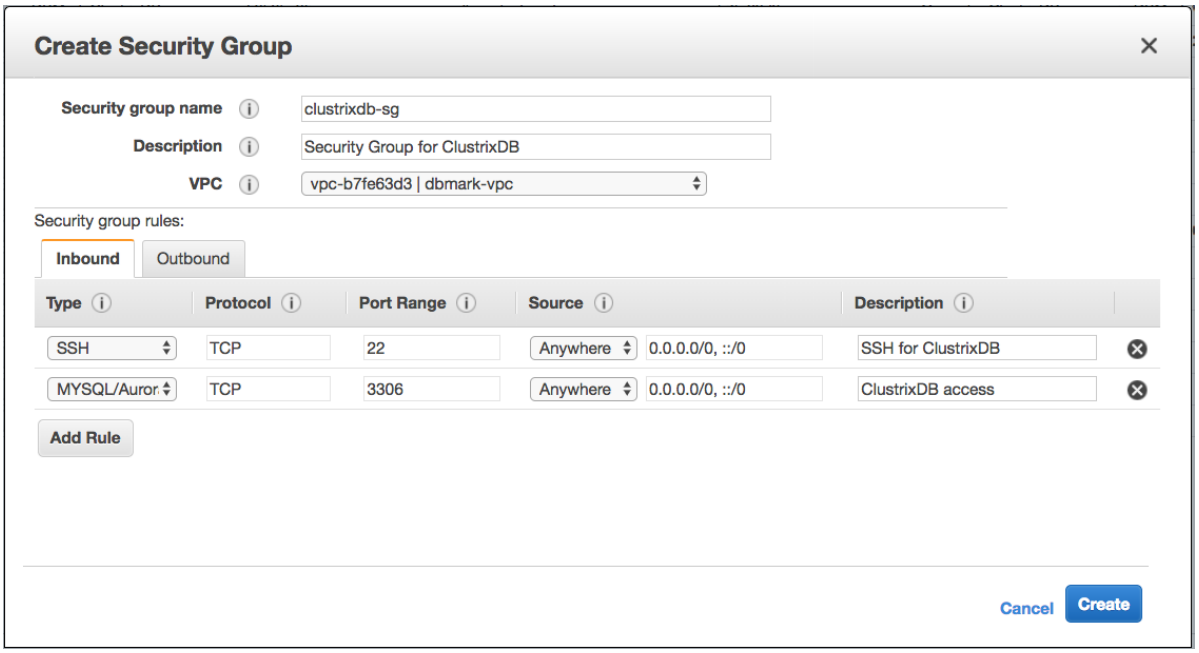
A security group, like a firewall, is used to control incoming and outgoing network traffic to and from your AWS machine instance. Select the “Security Groups” link under Network & Security on the left-side menu.

Then click the “Create Security Group” button.

Enter a name and description for the Security Group, and select the VPC (Best Practices For AWS Security Groups) in which the ClustrixDB instances will run. Then add the following 2 rules:

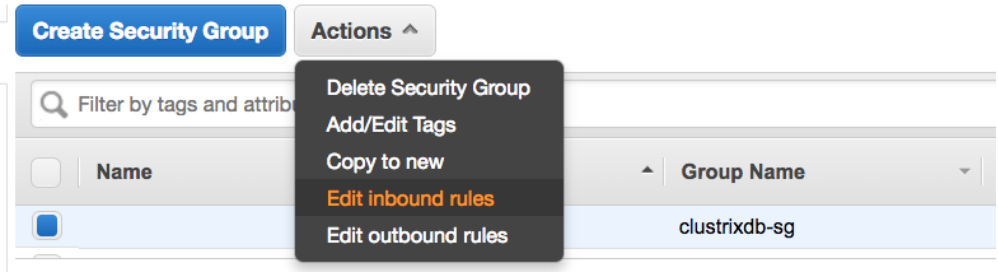
- Type: **SSH** (this will be used to install/configure/administer the instances)
- Type: **MYSQL/Aurora** (this is the port on which ClustrixDB listens for DB connections)

You may customize the Source of each rule. Ensure that the rule allows proper access from SSH and MySQL clients. We recommend starting with the Source set to “Anywhere” as the rules can be configured to be more restrictive after the installation is complete.



Click the “Create” button.

Now we must add three more rules that were not easy to add on the previous screen. Select the Security Group created from the list of groups, and choose Actions Edit inbound rules.

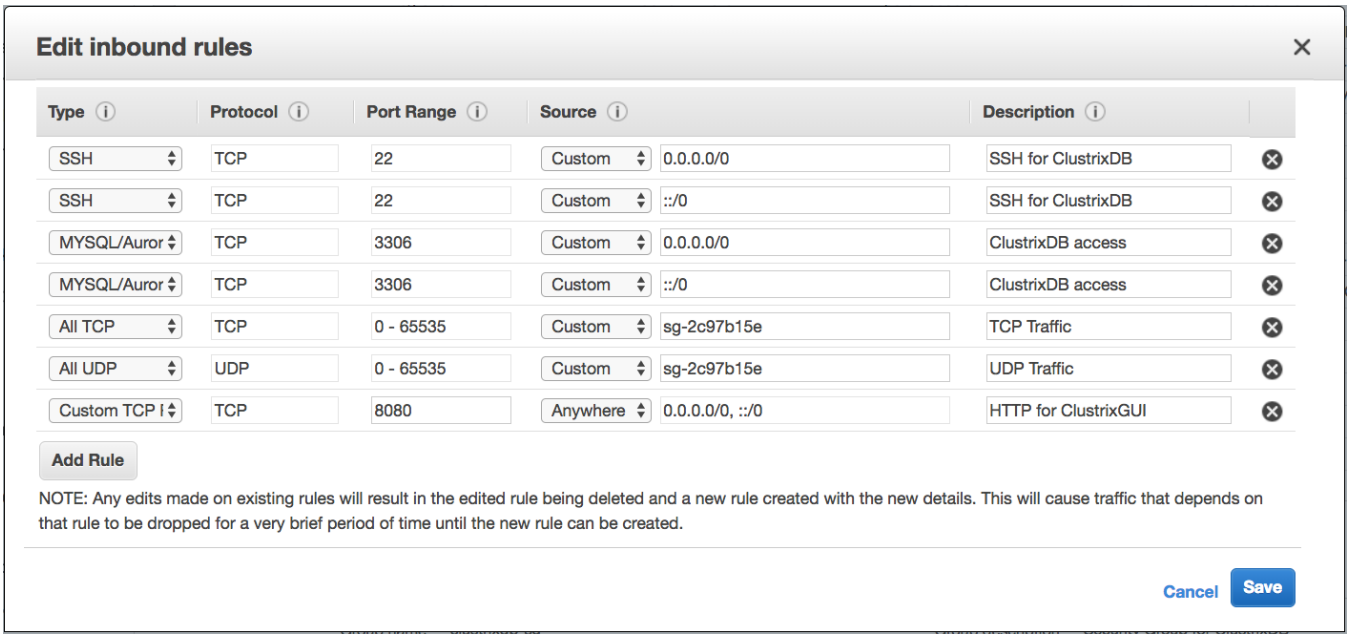


Add the following rules:

- “All TCP” and an “All UDP” rule. For both of these rules, the Source should be “Custom” and the name should be that of the current Security Group. These rules will allow all TCP and UDP traffic ONLY within this Security Group.
- "Custom TCP Rule" to set the HTTP rule for port 8080. Set the Port Range to 8080 and the Source to Anywhere. This will be used to access the ClustrixGUI.

After selecting “Custom” for the Source, in the field to the right start typing this Security Group’s name (e.g. in our example: “clustrixdb-sg”). As you type, a selection menu will appear and you can click the Security Group’s name to cause the Security Group's ID to be populated (shown below).

Example:

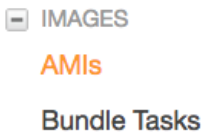


Please double-check your Security Group settings. Errors in the Security Group are the most common misconfiguration we see with ClustrixDB in AWS.

Click the “Save” button.

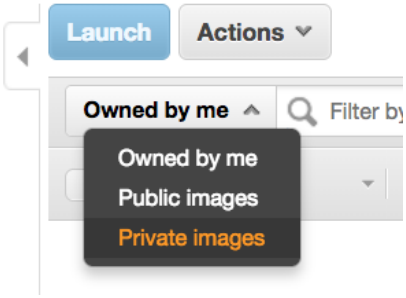
Launching the EC2 instances

Still, within the AWS Console, select the EC2 service and then select AMIs from the left-hand menu under the IMAGES category.



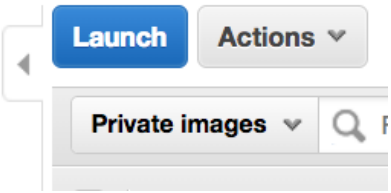
Step 1: Choose an Amazon Machine Image (AMI)

Then in the top-left corner of the list of AMIs, change the search drop-down box to “Private images”.



Identify and select the private AMI using the AMI ID provided by your Clustrix representative. If your list of private AMIs is long, you may need to enter the intended AMI ID in the search box and press enter.

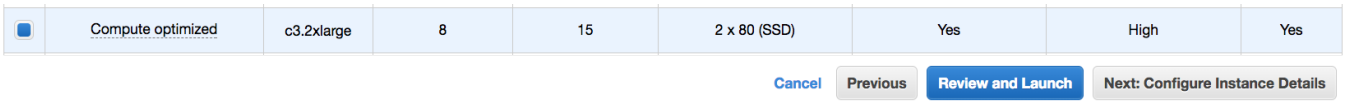
After selecting the correct AMI from the list of private AMIs, click the Launch button to move on to Step 2.



Step 2: Choose an Instance Type

Select one of the following instance types:

- C3: 2xlarge, 4xlarge, or 8xlarge
- I2: 2xlarge, 4xlarge, or 8xlarge
- I3: 2xlarge, 4xlarge, or 8xlarge



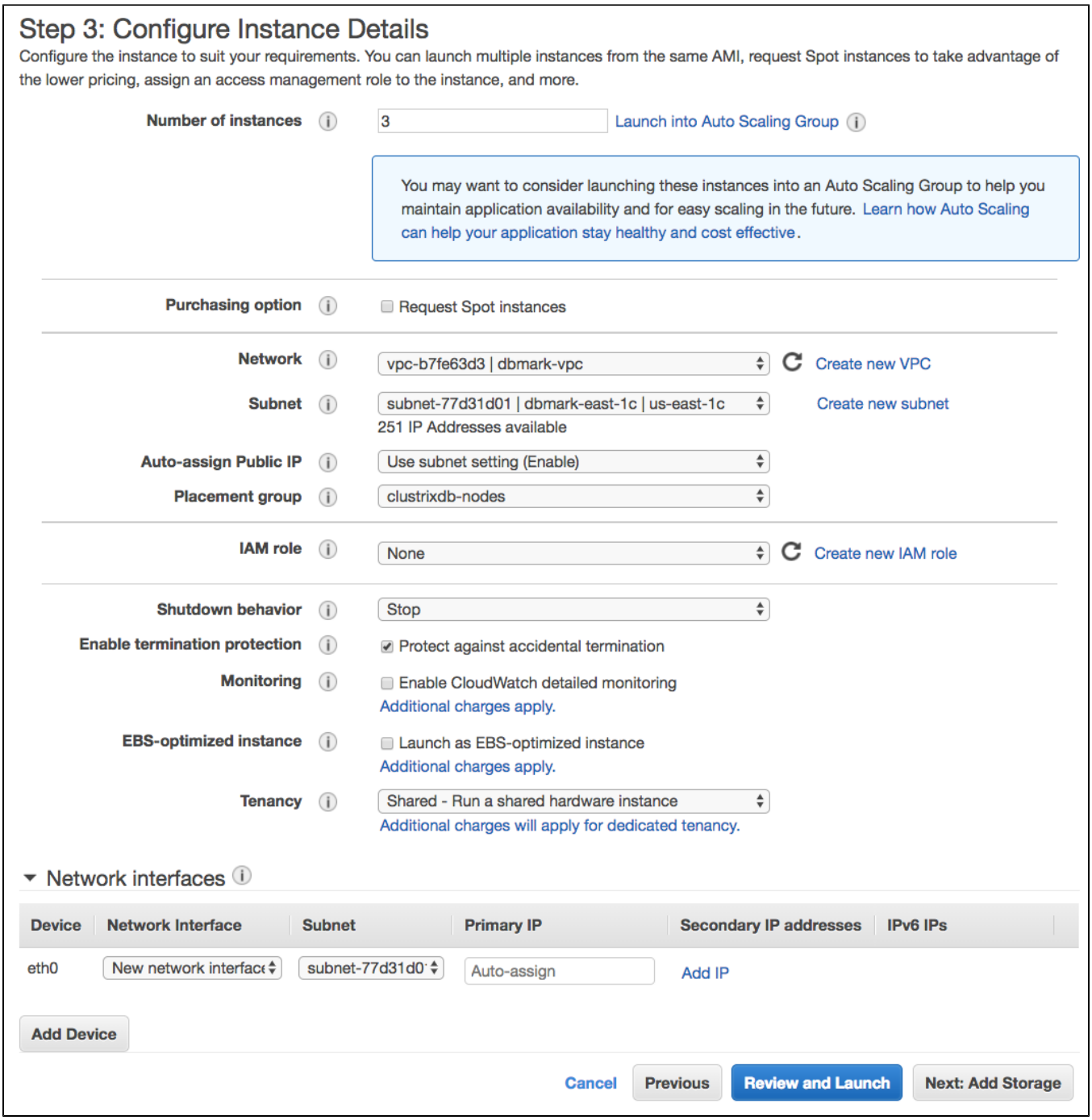
Click the “Next: Configure Instance Details” button to move to Step 3.

Step 3: Configure Instance Details

Make the following selections:

- **Number of Instances**: Total # of ClustrixDB instances to launch
  - Note: It is best to launch all instances at the same time, instead of one by one. This will ensure that they all launch in the same Placement Group
  - If you are configuring ClustrixDB without zones (default), the minimum is 3 instances
  - If you are configuring ClustrixDB using zones, you must have at least 3 instances across 3 zones (one per zone). For additional information, see Zones
- **Network**: Choose a VPC network
  - Do not launch ClustrixDB into EC2 Classic since Enhanced Networking is not available in EC2 Classic
  - All the nodes should all be in the same VPC, including when ClustrixDB is configured using Zones.
- **Subnet**: Select the desired subnet in the VPC
  - If you are unsure which subnet to use, please consult with your AWS admin, or you can ask for guidance from Clustrix
  - If you are configuring ClustrixDB using Zones, the nodes in one zone should all be in the same subnet.
- **Auto-Assign Public IP**: Normally production databases are not assigned public IPs. But if you prefer this for testing, the default is Enable
- **Placement Group**: Select the Placement Group you created in the previous steps.
  - If you are configuring ClustrixDB using Zones nodes placed in the same zone should be in the same Placement Group.
- **IAM Role**: Leave at the default, unless your company requires a different setting
- **Shutdown behavior**: Leave at the default, Stop (choices are: Terminate or Stop)
- **Enable termination protection**: Select this option
- **Monitoring**: Leave at default, or select if your company wants to use CloudWatch with ClustrixDB servers
- **EBS-Optimized Instance**: Leave at the default (unchecked)
- **Tenancy**: Leave at the default (Shared)
- **Network Interfaces**: Leave at the default

Example:



Click the “Next: Add Storage” button

Step 4: Add Storage

The AMI is pre-configured to leverage the ephemeral SSD disks, therefore there is no action required to configure the /data/clustrix/ partition.

We recommend configuring an EBS volume for logging. More detail on that can be found here: How to use an EBS volume for ClustrixDB logs

Click the “Next: Add Tags” button

Step 5: Add Tags

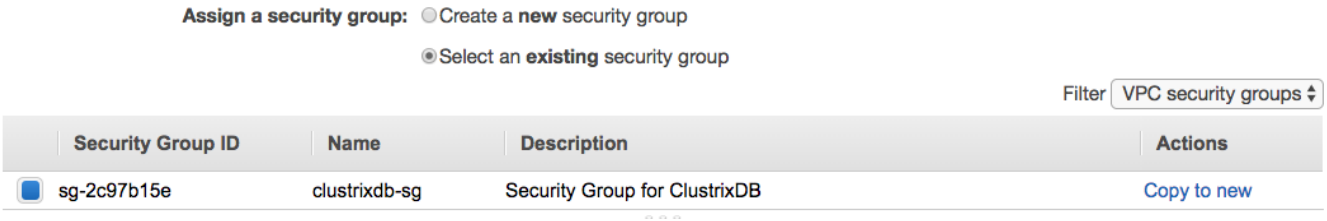
ClustrixDB does not require any additional instance tags.

Click the “Next: Configure Security Group” button

Step 6: Configure Security Group

Choose the option “Select an existing security group” and select the Security Group you created in the previous steps.

Example:



Click the “Review and Launch” button.

Step 7: Review Instance Launch

You may review the selections, then click the “Launch” button.

You are prompted to select an existing key pair or create a new key pair. Make the appropriate selection and click the “Launch Instances” button.

Once the EC2 instance is running, the ClustrixDB software is already running on the instance. However, the instance is not yet part of a cluster.

Gather Private IP Addresses

From the AWS Console, gather the private IP address of each EC2 instance that will run the ClustrixDB software. You will need these IP addresses during the installation.

Form the ClustrixDB Cluster

To join the ClustrixDB servers into a single unified cluster, connect to one ClustrixDB server and run a few SQL commands.

```
# set permissions for the pem file
shell> chmod 600 path_to_ec2_private_key.pem
# SSH into one of the ClustrixDB servers
shell> ssh –i path_to_ec2_private_key.pem
clustrix@ip_of_one_clustrixdb_server
# Launch the mysql client
shell> mysql -u root
```

The ClustrixDB AMI uses the clxd OS user to run the database process. An additional management (clustrix) is also configured. See Overview of OS Users and Managing Users and Privileges for more information.

Run the commands below where **ip** is the internal/private ip address of each node (10.0.x.y).

```
-- Set the cluster license
sql> SET GLOBAL license = 'license_key';
-- Create a cluster with the other nodes
sql> ALTER CLUSTER ADD 'ip', 'ip';
```

The ALTER CLUSTER command will initiate a short group change , after which the cluster is formed.

Run the clx command-line utility to view the status of the ClustrixDB nodes. This command can be run on any of the ClustrixDB servers as the **clustrix** user.

```
shell> /opt/clustrix/bin/clx stat
Cluster Name:    claef30606ca58824f
Cluster Version: 5.0.45-clustrix-9.0.4
Cluster Status:  OK
Cluster Size:    3 nodes - 16 CPUs per Node
Current Node:    ip-10-76-3-58 - nid 1

nid |     Hostname     | Status |  IP Address  | TPS |      Used      |  Total
----+------------------+--------+--------------+-----+----------------+--------
  1 |  ip-10-76-3-58   |     OK |  10.76.3.58  | 0   | 10.6M (0.01%)  | 119.0G
  2 |  ip-10-76-3-105  |     OK | 10.76.3.105  | 0   |  5.9M (0.00%)  | 119.0G
  3 |  ip-10-76-3-232  |     OK | 10.76.3.232  | 0   |  5.9M (0.00%)  | 119.0G
----+------------------+--------+--------------+-----+----------------+--------
                                               0   | 22.4M (0.01%)  | 357.0G
```

Post Installation Setup

Configure passwordless SSH Authentication for the clustrix user (clustrix user has the role of clxm user) for all nodes: Configure SSH Authentication.

Create a user to be used during the import. This should be done while logged into ClustrixDB as the root DB user.

```
-- Create a user for importing
sql> grant all on *.* to 'impo
rtuser'@'%' identified by 'imp
ortuserpasswd';
```

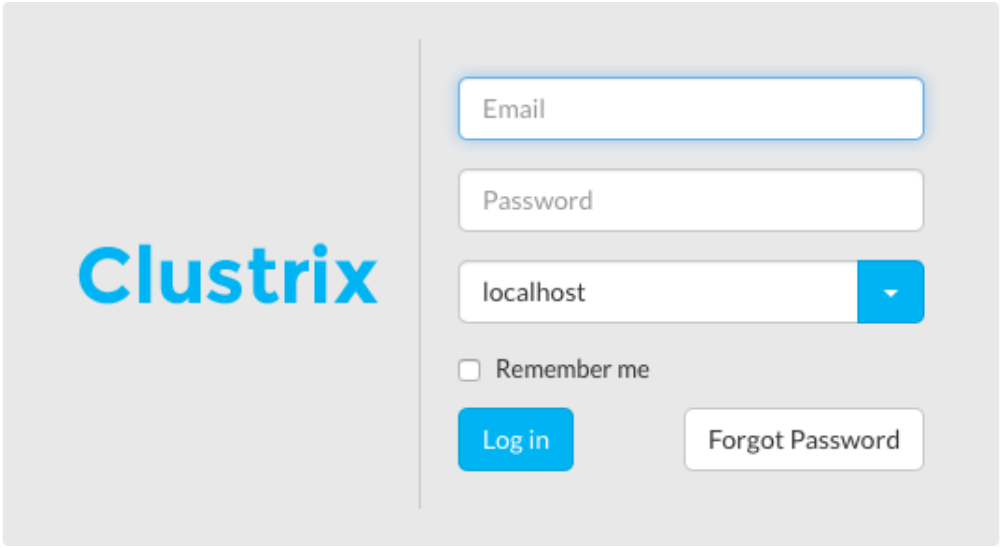
For more information see: clustrix_import

ClustrixGUI

The ClustrixGUI must be started with the following commands before the first time you access it.

```
shell> /opt/clustrix/bin/clx nanny stop_job
clxdbi
shell> /opt/clustrix/bin/clx nanny
start_job clxdbi
```

Access the ClustrixGUI by typing the ip or hostname of one of your nodes into a web browser, include port :8080. Chrome is the recommended web browser.



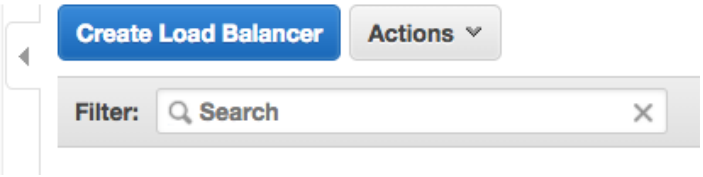
Provide the following credentials:

- Email: noreply@clustrix.com
- Password: clustrix
- Target: localhost

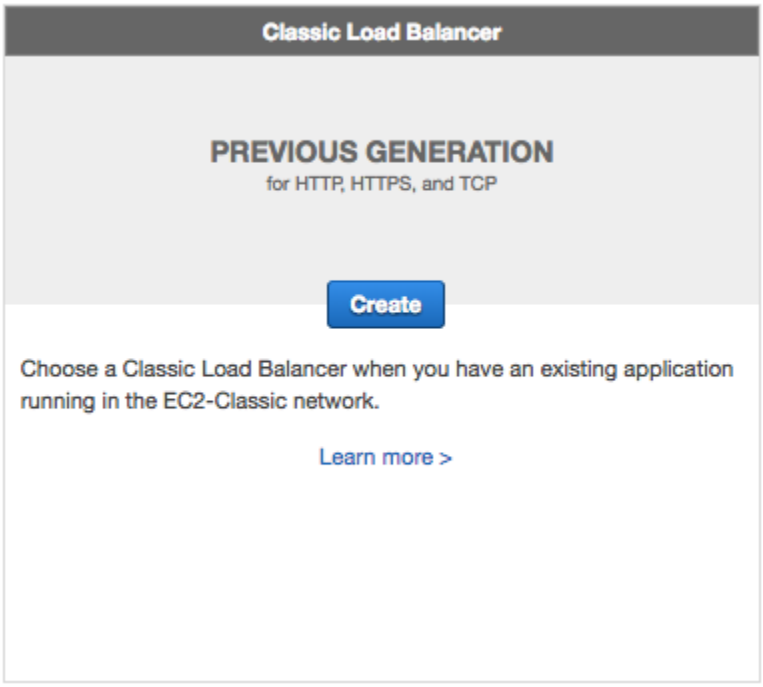
Setting up a Load Balancer

Select the “Load Balancers” link on the left-side menu.

Then click the “Create Load Balancer” button.



Choose the Classic Load Balancer, and click the Continue button.



Step 1: Define Load Balancer

Enter the Basic Configuration as follows.

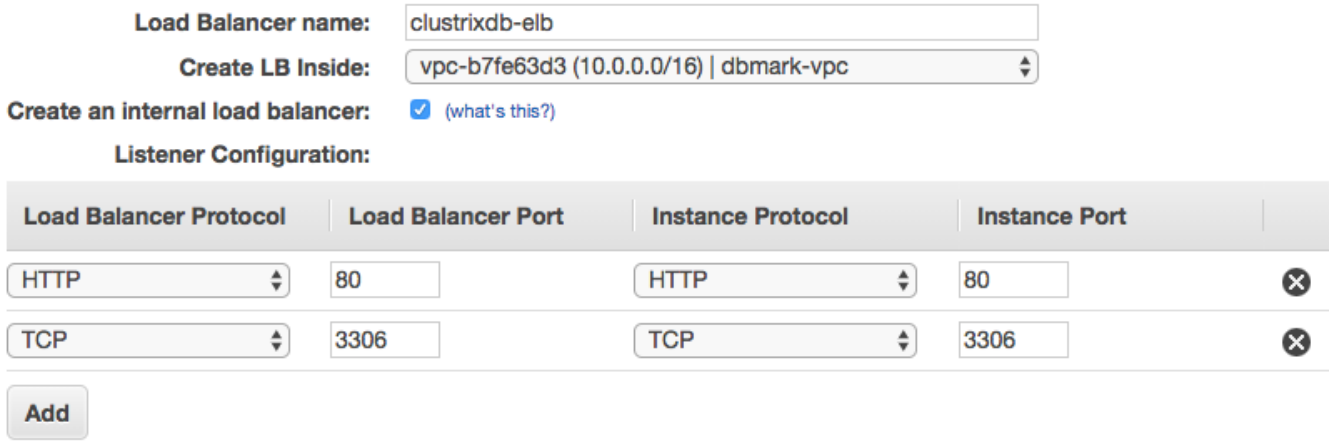
- Enter a name for the load balancer (any un-used name).
- For “Create LB Inside”, select the same VPC that you used above.
- Check the box labeled “Create an internal load balancer” so the ELB is fully contained inside of your VPC.

Then add the following Load Balancer Protocols:

- HTTP port 80
    - Load Balancer Protocol: HTTP
    - Load Balancer Port: 80
    - Instance Protocol: HTTP
    - Instance Port: 80
- MySQL Port 3306
    - Load Balancer Protocol: TCP
    - Load Balancer Port: 3306
    - Instance Protocol: TCP
    - Instance Port: 3306

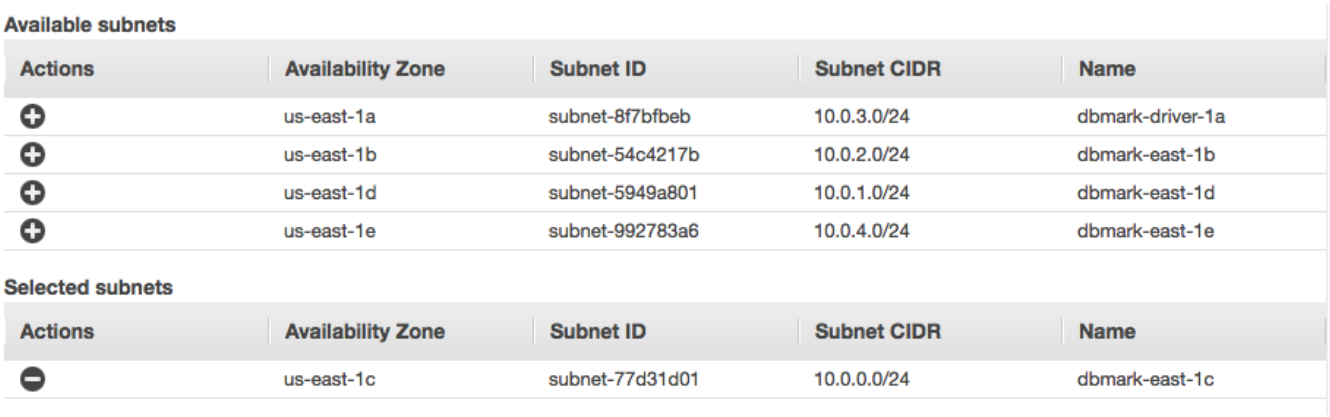
Then select the same subnet that you used above in “Step 3: Configure Instance Details” when launching the EC2 instances.

Example:



Select Subnets

Select the subnet your Cluster is in (see Configure Instance Details, above). If you are using ClustrixDB Zones, you will also need to select the subnet for each zone where you wish traffic to be routed by your load balancer:

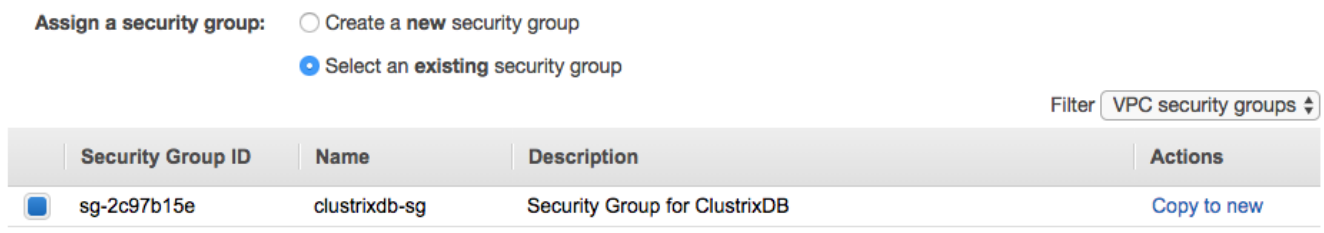


Click the “Next: Assign Security Groups” button.

Step 2: Assign Security Groups

Select the same Security Group that you created previously for the ClustrixDB instances. Choose the option “Select an existing security group”, and then check the proper security group.

Example:



Click the “Next: Configure Security Settings” button.

Step 3: Configure Security Settings

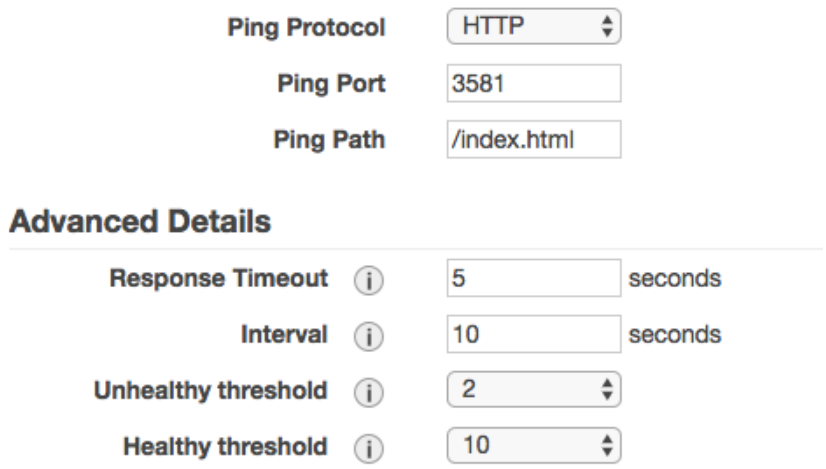
Click the “Next: Configure Health Check” button

Step 4: Configure Health Check

Configure the health check using the following values:

- Ping Protocol: HTTP
- Ping Port: 3581
- Ping Path: /index.html
- Advanced Details:
    - Response Timeout: 5 seconds
    - Interval: 10 seconds
    - Unhealthy threshold: 2
    - Healthy threshold: 10

Example:



Click the “Next: Add EC2 Instances” button

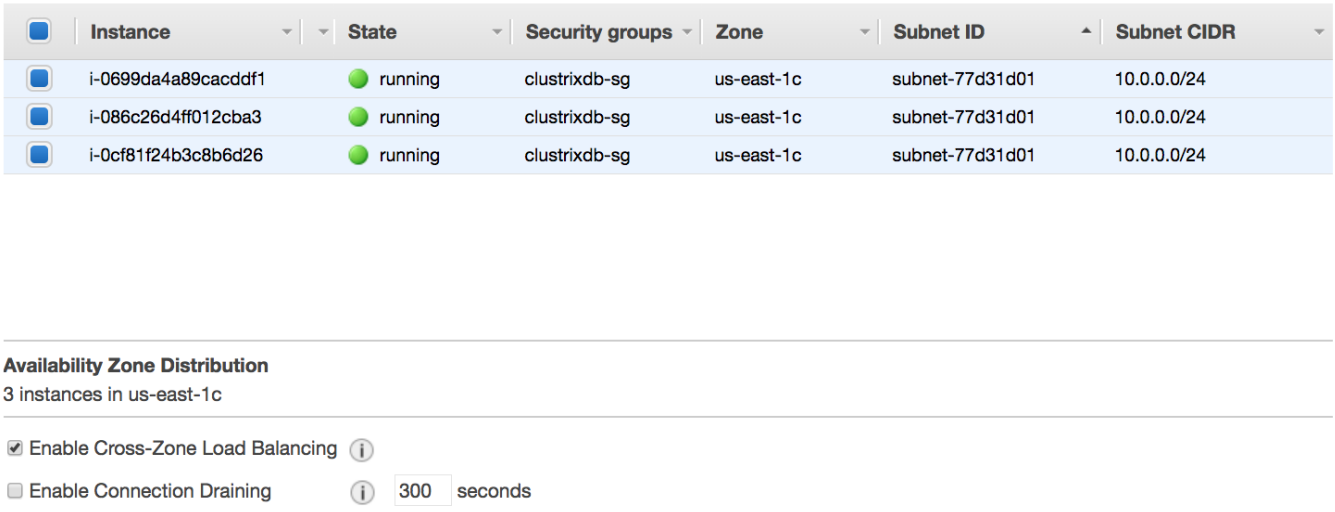
Step 5: Add EC2 Instances

Select from the list the ClustrixDB EC2 instances you launched previously.

Additionally, check the following options:

- Enable Cross-Zone Load Balancing
    - Unselected - if not using ClustrixDB Zones
    - Selected - if using ClustrixDB Zones
- Enable Connection Draining - Unselected

Example:



Click the “Next: Add Tags” button.

Step 6: Add Tags

ClustrixDB does not require any additional tags.

Click the “Review and Create” button.

Step 7: Review

Review the settings, then click the “Create” button.

Increase Idle Timeout

We want the database to handle terminating idle connections instead of the ELB. The default ELB Idle Timeout value is 60 seconds, which is too short for long-running queries. Therefore, we recommend setting the ELB Idle Timeout to its maximum value of 3600 seconds.

Return to the main Load Balancers page by selecting the “Load Balancers” link on the left-side menu.

Then in the list of load balancers, select the load balancer you just created. In the Description tab of the details pane below the list of load balancers, find the Attributes section near the bottom:

Click the "Edit idle timeout" button, then enter 3600, and click the Save button.