

Setting up Ampool Hub

This document describes how to set up Ampool Hub.

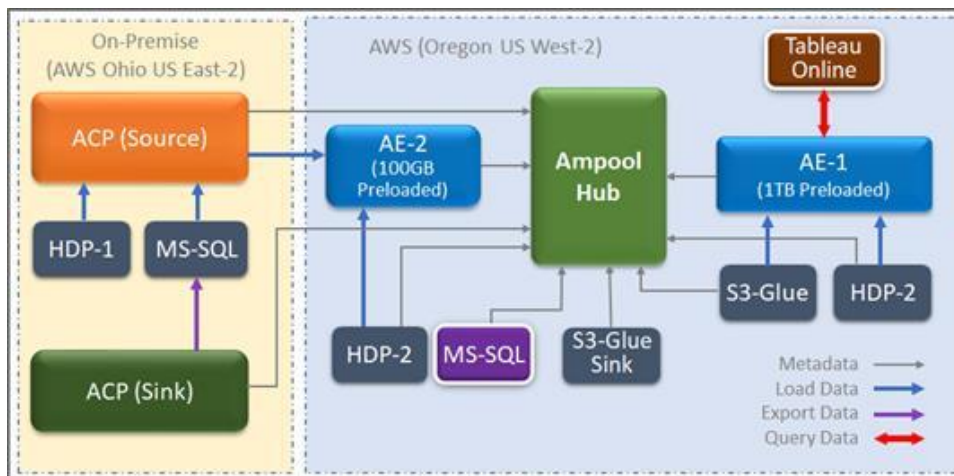
Overview

Ampool Hub provides a unified view across data systems such as data lake, data stream, data warehouse, relational database management sources, deployed in multiple cloud and on-premise environments.

Ampool Hub facilitates the central administration to onboard resources, assign resources, and map data. Ampool Engine can be launched and monitored with Ampool Hub.

Ampool Hub can be installed on cloud-based virtual machine (VM) instances. Amazon Web Services (AWS) and Microsoft Azure (Azure) are the supported cloud providers for Ampool Hub installation.

Following is a diagrammatic representation of Ampool Hub in conjunction with Ampool Engine (AE) and Ampool Cache Proxy (ACP).



Data engineers, business analysts, and data scientists can reduce time to business data insights from months to days across real-time and batch data, with faster performance and security.

Set up Ampool Hub

Setting up of Ampool Hub on AWS or Microsoft Azure cloud comprises the following steps.

1. Launch VM instance on AWS or Azure
2. Whitelist VM instance IP address
3. Install Ampool Hub
4. Register Ampool Hub
5. Create an Ampool Engine cluster from Ampool Hub

Let us look at the steps in detail.



Step 1: Launch VM Instance on cloud

Follow the steps in the respective sections below, depending on the cloud that you are using.

Launch Instance on AWS

You must launch a VM instance based on the custom AMI provided by Ampool for Ampool Hub installation.

Ensure that the following specifications are selected while creating the VM instance.

Field	Description
AMI	Red Hat Enterprise Linux (RHEL) 7.1(HVM)/ Ubuntu 18.04 LTS
Disk Size	50 GiB
IAM Role	Ampool-Admin

Note: IAM Role is required if you want to use IAM to authenticate to Ampool Hub and Ampool Engine.

Refer to <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/launching-instance.html> to launch.

Once the instance is launched, click the Instance ID to view instance details. Note down the public IP address and private IP address of the instance.

Launch Status

Your instances are now launching
The following instance launches have been initiated: i-0e42fb4279921be34 [View launch log](#)

Get notified of estimated charges
Create billing alerts to get an email notification when estimated charges on your AWS bill exceed an amount you define (for example, if you exceed the free usage tier).

How to connect to your instances

Your instances are launching, and it may take a few minutes until they are in the **running** state, when they will be ready for you to use. Usage hours on your new instances will start immediately and continue to accrue until you stop or terminate your instances.

Click **View Instances** to monitor your instances' status. Once your instances are in the **running** state, you can **connect** to them from the Instances screen. Find out [how to connect to your instances](#).

▼ Here are some helpful resources to get you started

- How to connect to your Linux instance
- Learn about AWS Free Usage Tier

- Amazon EC2: User Guide
- Amazon EC2: Discussion Forum

Launch Instance on Azure

Log in to Azure Dashboard and create a VM on Azure. Refer to <https://docs.microsoft.com/en-us/azure/virtual-machines/linux/quick-create-portal>.

Ensure that the following specifications are selected while creating the VM instance.

Field	Description
Image	Ubuntu 18.04 LTS
Disk Size	50 GiB
Role	Contributor

Note down the public IP address and private IP address of the VM.

Setting up Ampool Hub

2



Step 2: Whitelist Public IP

You must allow inbound and outbound traffic from the public IP address of the instance created in step 1 above.

If you have created the instance on AWS, refer to

https://docs.aws.amazon.com/vpc/latest/userguide/VPC_SecurityGroups.html to create and configure security group.

If you have created the instance on Azure, refer to <https://docs.microsoft.com/en-us/azure/virtual-network/manage-network-security-group> to create and configure security group.

Step 3: Install Ampool Hub

Follow the steps given below to install Ampool Hub.

1. Login to Ampool Hub Instance with ssh.

```
ssh -i <pem_file_path> <username>@<Instance Public OR Private IP>
```

Note: *username* in the above ssh command should be replaced with *ec2-user* for RHEL VM instance and *ubuntu* for Ubuntu VM instance.

For example,

```
ssh -i user.pem ec2-user@42.200.216.210
```

2. Add the entry for Hashicorp Vault to /etc/hosts.

```
sudo vi /etc/hosts
```

```
# add the following entry for vault
127.0.0.1 localhost vault

# The following lines are required for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts
```

3. Download a latest Ampool-Hub repo. This will take around a few minutes, depending on your network speed.

```
wget https://s3-us-west-2.amazonaws.com/<ampool-hub-version-number>/<build-number>/ampoolhub.tar.gz
```

4. Unpack the ampoolhub.tar file.

```
tar -xvf ampoolhub.tar.gz
```



- Run prerequisites file provided to you by Ampool. This should take around 10 minutes to complete the run.

```
bash <prerequisites file>
```

Note: Refer to the following table for OS-specific prerequisite to run

Instance	Prerequisites file to use
Ubuntu	prerequisites.sh
RHEL	prerequisites_RHEL.sh

- Install Ampool Hub.

```
sudo python3.8 deploy_ampool_hub.py
```

During the installation, the script prompts you to enter details for certificates, database drivers, public IP, private IP, Kerberos realm settings. Provide private IP address of the virtual machine (VM) instance. The following table describes some of the important installation parameters.

Parameter	Description
Database password	MariaDB password
External Calls for Hub	Calls outside of Ampool Hub network
Oracle Drivers	URL of ODBC driver of the data source to connect to. If URL is not specified, Oracle data source will not work.
Kerberos	Kerberos KDC server realm

Note: If you are not using a public IP address, provide private IP address of the VM for both public and private IP addresses.

After installation, the message 'Ampool Hub started successfully' is displayed.

The installation will take a few minutes to complete.

- After successful installation, verify that the following Docker containers that are running.

```
sudo docker ps
```

```
ubuntu@rk-tmp-machine:~$ sudo docker ps
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS        NAMES
9f673d2aa7e    ampool/ampoolweb:2.3.85             "nginx -g 'daemon of..." 11 minutes ago Up 11 minutes        webContainer
eceb57911523    ampool/ampoolhub:2.3.85             "unicorn --worker-c..." 11 minutes ago Up 11 minutes        appContainer
bd5a451ec02b    mariadb:10                          "docker-entrypoint.s..." 11 minutes ago Up 11 minutes        dbContainer
ae987bf6f6d1    ampool/ampoolkerberos:2.3.85        "/sbin/init"            12 minutes ago Up 11 minutes        kerberosContainer
4f9d5b13be5f    ampool/ampooljava:2.3.85            "java -jar -Dlogging..." 12 minutes ago Up 11 minutes        javaContainer
74e2254798fd    ampool/ampoolvault:2.3.85           "vault server -confi..." 12 minutes ago Up 11 minutes        vaultContainer
ubuntu@rk-tmp-machine:~$
```

Container Name	Description
webContainer	Container for Ampool Hub UI
appContainer	Container for Ampool Hub Web service
dbContainer	Container for database

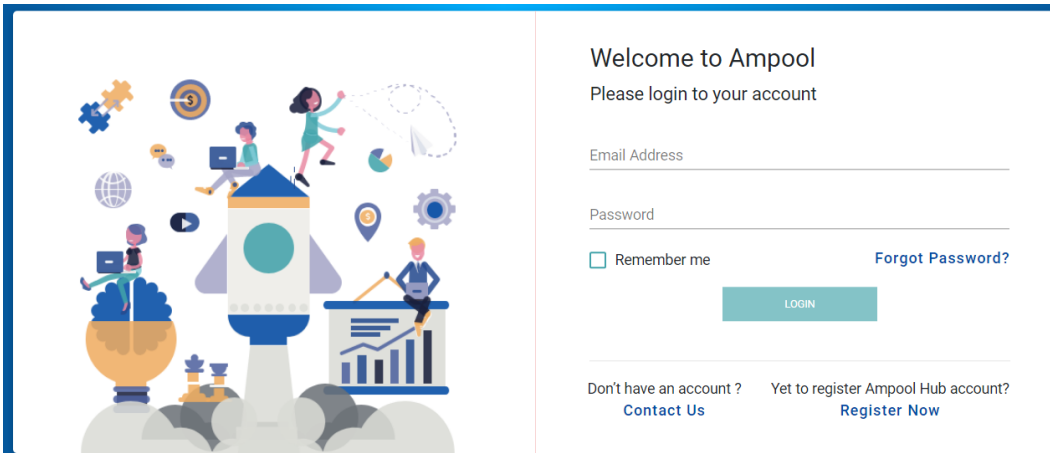
kerberosContainer	Container for Kerberos Key Distribution Center
javaContainer	Container for Java-based Web service
vaultContainer	Container for Hashicorp Vault

Step 4: Register Hub

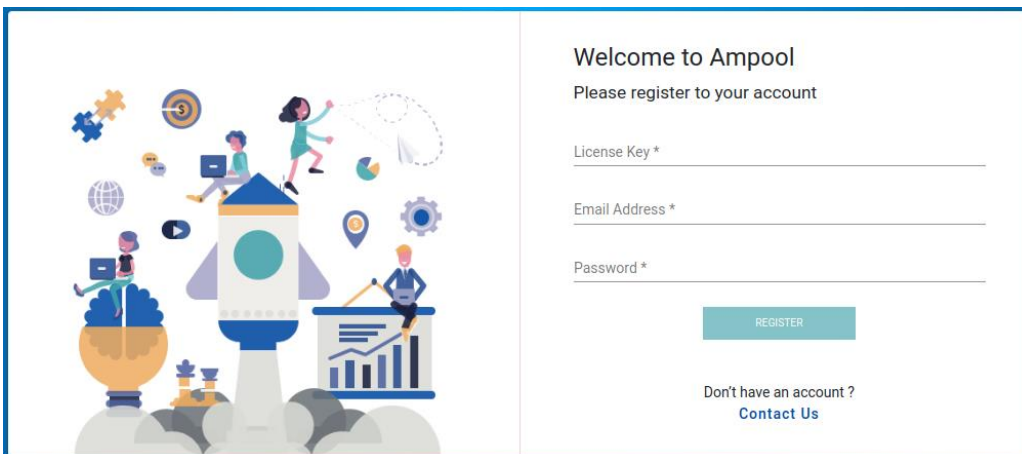
Once Ampool Hub installation is complete, register your Ampool Hub account.

Follow the steps given below to register the Ampool Hub account.

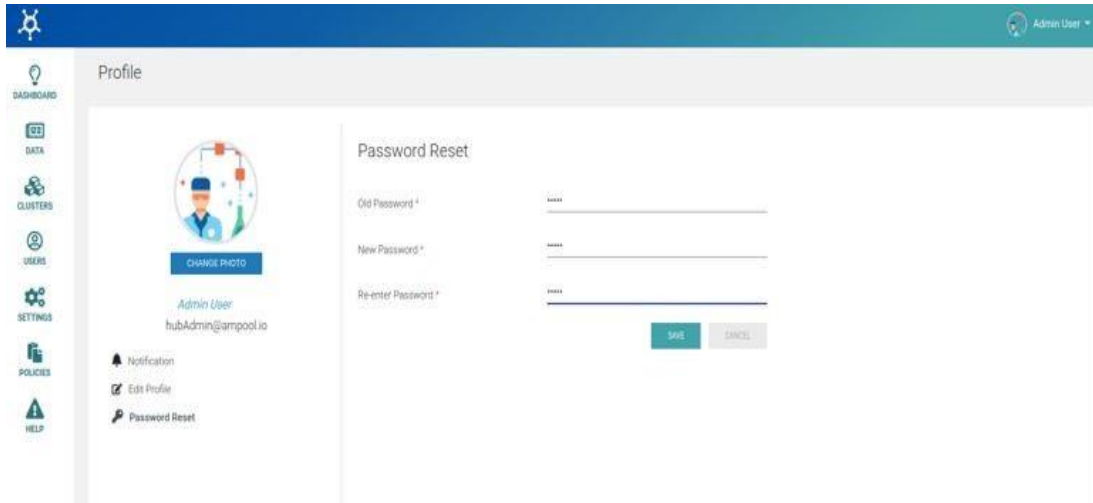
1. Navigate to https://<hub_ip>/ and click the **Register Now** link on the Login page to register Ampool Hub account.



2. Enter license key, email and temporary password that has been received on your registered email ID from Ampool.

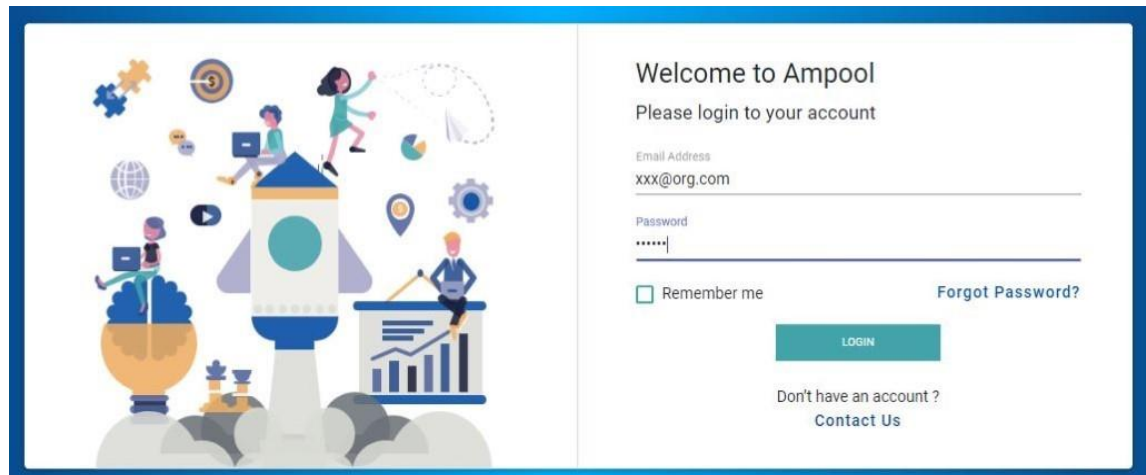


3. Change the temporary password.



The screenshot shows the 'Profile' page of the Ampool Hub. On the left, there is a sidebar with navigation icons for Dashboard, Data, Clusters, Users, Settings, Policies, and Help. The main content area is titled 'Profile' and contains a user profile card for 'Admin User' with the email 'hubAdmin@ampool.io'. Below the profile card are links for 'Notification', 'Edit Profile', and 'Password Reset'. To the right of the profile card is a 'Password Reset' form with three input fields: 'Old Password *', 'New Password *', and 'Re-enter Password *'. Each field has a 'PASSWORD' label above it. At the bottom of the form are 'SAVE' and 'CANCEL' buttons.

4. Login back to Ampool Hub with your new password as your session expires after changing the temporary password.



The screenshot shows the login page of the Ampool Hub. On the left, there is a large illustration featuring a rocket ship, a bar chart, a globe, and several people working on laptops and tablets. On the right, the text 'Welcome to Ampool' is displayed, followed by 'Please login to your account'. Below this, there are two input fields: 'Email Address' with the value 'xxx@org.com' and 'Password' with masked characters '*****'. There is a 'Remember me' checkbox and a 'Forgot Password?' link. A 'LOGIN' button is located below the password field. At the bottom, there is a link for 'Don't have an account ? Contact Us'.

5. Provide Organization details

Initial Setup

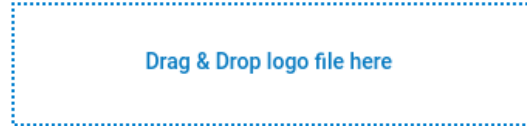
- ☒ Organization
- ☐ Settings
- ☐ Cloud Credentials
- ☐ Sources & Sinks
- ☐ Groups
- ☐ Users

Welcome to Ampool!

Begin the setup process by setting up your first organization.

Name *

Logo



OR

Enter Public URL and press upload button



6. Add SMTP settings on the SMTP tab. Enter the SMTP server details, Username, Password, Host Address, and SMTP Email. This is for Ampool Hub to send email related to user account creation, cluster status. click SAVE, and then click NEXT.

Settings

Begin the setup process by setting up your SMTP and HERE Details.

SMTP Settings

HERE Settings

Username *

Password * 

Host Address *

SMTP Email *

7. Add HERE map settings on the HERE tab. When a data source is added, HERE map helps with viewing the data source location on the map. click SAVE, and then click NEXT.

Settings

Begin the setup process by setting up your SMTP and HERE Details.

SMTP Settings

HERE Settings

☐ Disable HERE Settings

Here App Id *

.....



Here App Code *

.....



8. Add Cloud Credentials.

If you are using AWS, enter Access Method, Region, Availability Zone, Subnet Id and Security group where you wish to install or deploy Ampool Engine. Static Key means authentication to AWS with an access key and secret key provided by AWS. IAM means authentication to AWS with Identity Access Management role. You could use public IP or private IP for launched cluster.





Cloud Credential Setup

Add the cloud credential you will want to use

aws

Azure

AWS Cloud Credentials

Cloud Credential Name *	AutoAWSCred1
Access Method *	STATIC_KEYS
IAM Role To Attach	IAM Role To Attach
Access Key 
Secret Key 
Description	Auto cloud credential
Pem Key Name *	ampool-qa
Pem File *	File uploaded  
Assign Public IP *	<input checked="" type="radio"/> Yes <input type="radio"/> No
OS Type *	centos
Region *	US West (Oregon)

BACK

UPDATE



CANCEL

NEXT









If you are using Azure, enter the cloud credentials for Azure, along with other details such as Subscription ID, client ID, Secret, Tenant, Public Key, Private Key, Location, Resource Group, Storage Account, Security Group, Virtual Network, Subnet.





Cloud Credential Setup

Add the cloud credential you will want to use

AZURE Cloud Credentials

Cloud Credential Name *	AutoAzureCred1	
Subscription ID *	
Client ID *	
Secret *	
Tenant *	
Description	Auto Generated Cloud Credential	
public Key Name *	id_rsa.pub	
Public Key *	File uploaded  	
private Key Name *	id_rsa.pem	
Private Key *	File uploaded  	
Assign Public IP *	<input checked="" type="radio"/> Yes <input type="radio"/> No	

 Back
 Update
 Cancel
 Next

After entering the details, click TEST CONNECTION. Once the connection test is successful, click ADD, and then click NEXT.

Note: It is mandatory to test the connection to the cloud after entering the cloud credentials. Without testing the connection, you will not be able to save the cloud credentials.

- Add data sources and data sinks. Before adding a data system, ensure that the data system and its respective services are running. After adding the details, you can click **TEST CONNECTION** to test the connection to the data system, or you can directly click ADD. If you click ADD, Ampool Hub tests the connection to the data system. If Ampool Hub can connect to the data system successfully, the data system is added to Ampool Hub.

Refer to the following screenshots for addition of Glue, Redshift, Snowflake, and MS SQL Server.



Adding Redshift Data System

Add Data System

SOURCE

SINK

☒ Cloud

☐ On-Premise

Data System Type*

DATA_LAKE

Data System*

AWS_Redshift

Name of data system instance* ⓘ

AWS_Redshift

Vendor of data systems* ⓘ

Amazon Redshift

Version of data systems* ⓘ

14

Redshift user name* ⓘ

admin

Redshift password* ⓘ

.....

Redshift cluster endpoint* ⓘ

localhost

Redshift port* ⓘ

5439

Redshift database* ⓘ

SDI_DATA

Data Operations Engine* ⓘ

Spark

CANCEL

TEST

ADD

Adding Snowflake Data System

Add Data System

SOURCE

SINK

☒ Cloud ☐ On-Premise

Data System Type* DATA_LAKE

Data System* Snowflake

Name of data system instance* i Snowflake

Vendor of data systems* i Snowflake

Version of data systems* i 3.27

Snowflake JDBC URL* i jdbc:snowflake://<snowflake account>/

Snowflake database* i demo

Snowflake authentication type* i BASIC

Snowflake user name* i admin

Snowflake password i

Snowflake privatekey i

CANCEL

TEST

ADD



Adding Amazon Glue Data System

Add Data System

SOURCE

SINK

☒ Cloud

☐ On-Premise

Data System Type*

DATA_LAKE

Data System*

AWSGlue

Name of data system instance* ⓘ

AWSGlue

Vendor of data systems* ⓘ

Amazon Glue

Version of data systems* ⓘ

1.11

Amazon AWS Region Code* ⓘ

us-west-2

Access Method* ⓘ

STATIC_KEYS

Amazon User Access Key ⓘ

<enter access key>

Amazon User Secret Key ⓘ

.....

Data Operations Engine ⓘ

Presto

Register as presto catalog* ⓘ

true

CANCEL

TEST

ADD



Adding MS SQL Server Data System

Add Data System

SOURCE

SINK

☒ Cloud

☐ On-Premise

Data System Type*

DATA_LAKE

Data System*

SQLServer

Name of data system instance* ⓘ

SQLServer

Vendor of data systems* ⓘ

Microsoft SQLServer

Version of data systems* ⓘ

14

MS SQL server user name* ⓘ

admin

MS SQL server password* ⓘ

.....

MS SQL Server location* ⓘ

localhost

MS SQL Server port* ⓘ

1433

MS SQL Server Database* ⓘ

TPCH_PARTITIONED_ORC_2

Data Operations Engine* ⓘ

Spark

CANCEL

TEST

ADD

Adding Teradata Data System

Add Data System

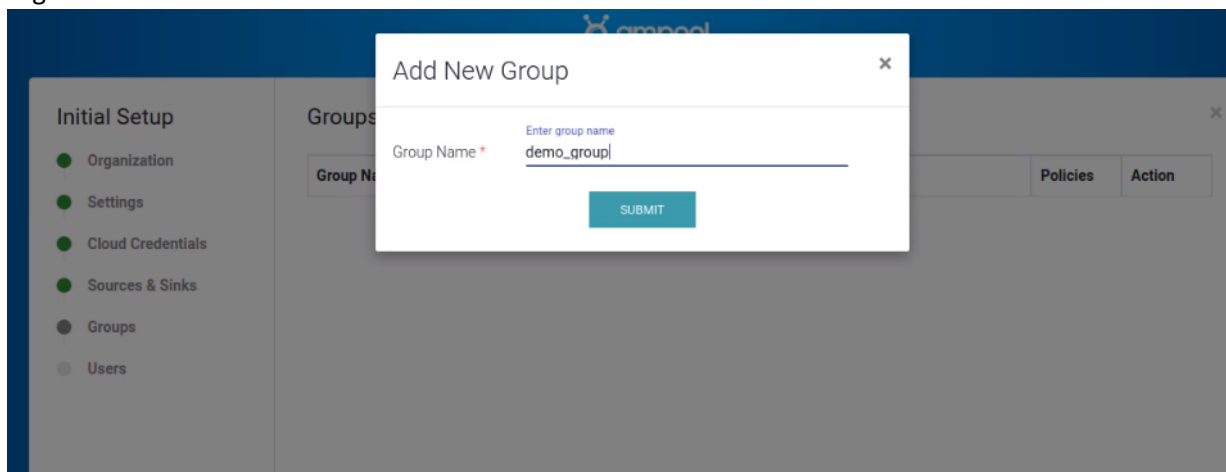
SOURCE
SINK

☒ Cloud
 ☐ On-Premise

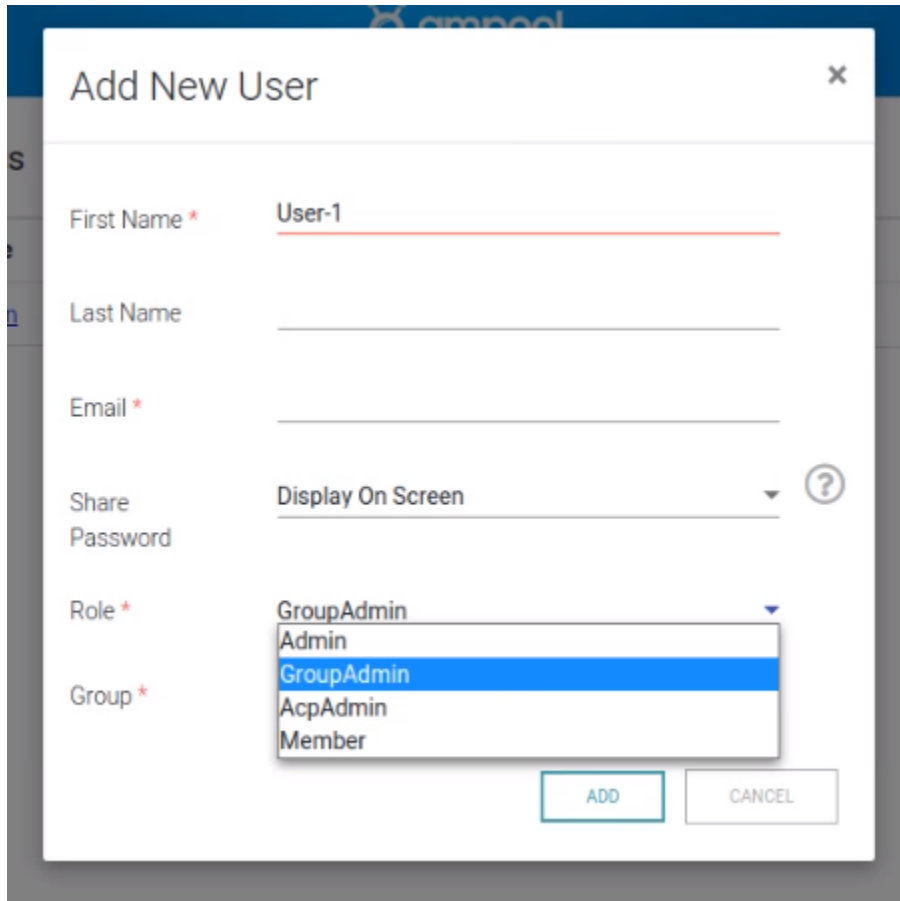
Data System Type*	DATA_LAKE
Data System*	Teradata
Name of data system instance* ⓘ	Teradata
Vendor of data systems* ⓘ	Teradata
Version of data systems* ⓘ	16
Teradata user name* ⓘ	admin
Teradata password* ⓘ	*****
Teradata Server location* ⓘ	localhost
Teradata Database* ⓘ	TERA_DB2
Data Operations Engine* ⓘ	Presto
Register as presto catalog* ⓘ	true

CANCEL
TEST
ADD

10. Add user groups under Groups. Groups represents user groups such as a department or a function in an organization. Click SUBMIT.



11. Create one or more users with different roles. The available roles are Admin, GroupAdmin, AcpAdmin and Member. An Admin is a superadmin user. A GroupAdmin is a group administrator that manages Members of the group. AcpAdmin is an administrator for the Ampool Cache Proxy. Click ADD after entering the new user details.



The screenshot shows a modal window titled "Add New User" with a close button (X) in the top right corner. The form contains the following fields and options:

- First Name ***: Text input field containing "User-1".
- Last Name**: Text input field.
- Email ***: Text input field.
- Share Password**: Text input field containing "Display On Screen". A help icon (?) is visible to the right of the field.
- Role ***: A dropdown menu with the following options: GroupAdmin, Admin, GroupAdmin (highlighted in blue), AcpAdmin, and Member.
- Group ***: Text input field.

At the bottom right of the modal, there are two buttons: "ADD" and "CANCEL".

Add New User

First Name *

John

Last Name

Doe

Email *

john.doe@example.com

Share Password

Display On Screen

?

Role *

GroupAdmin

Group *

demo_group x

12. After adding users click COMPLETE SETUP.

You are directed to the Ampool Hub dashboard after the setup is complete.

Step 5: Create Ampool Engine Cluster from Ampool Hub

You can create an Ampool Engine cluster from Ampool Hub. Username field is non-editable.

To create an Ampool Engine Cluster perform following steps.

1. Navigate to **CLUSTERS**.
2. Click on **+ NEW CLUSTER**.
3. Enter a cluster name.
4. Select **Worker Count** for cluster. Minimum value is 1.
5. Select **Cloud Provider** as AWS or Azure.
6. Select **Worker Type** – **Small, Medium, Large**. Worker Type is cloud provider specific.
7. Choose between Yes/No for HA Cluster.
8. Choose between Yes/No for Kerberised Cluster.
9. Enter a password for Cluster Login. Username is admin, by default, and is a read-only field.
10. If you want to associate one or more data systems registered on Hub with the cluster while launching it, select the required data systems.
11. Click **CREATE** to create Ampool Engine cluster.

New Cluster

Cluster Name *

democluster

*Allowed characters 'a-z', '0-9' of length 3-20

Workers Count *

1

Cloud Provider *

AWS

Cloud Credentials *

ampool_aws

Worker Type *

Small

Type

c5d.9xlarge

Memory

68 GB

Cores

36

Medium

Type

m5d.8xlarge

Memory

120 GB

Cores

32

Large

Type

r5d.8xlarge

Memory

248 GB

Cores

32

Enable HA *

☐ Yes

☒ No

Enable Kerberos *

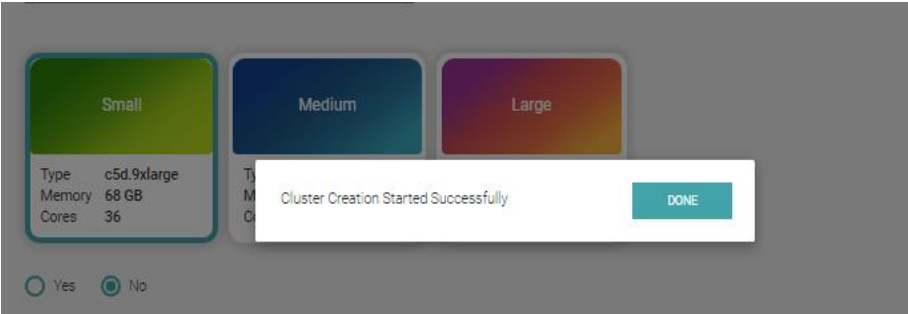
☐ Yes

☒ No

Username *

admin

Password *



12. Find the respective cluster card and check the **Status**.

DASHBOARD

DATA

CLUSTERS

USERS

SETTINGS

POLICIES

HELP

Clusters [1]

ALL

LARGE MEDIUM SMALL ALL

aws democluster SMALL

Workers	1
HA	NO
Kerberos	NO
Cloud Region	
Memory	68 GB
Cluster IP	
Cluster AE URL	
Dashboard	↗
Status	Installing ↗

DELETE

aws dytest SMALL

Workers	1
HA	NO
Kerberos	NO
Cloud Region	us-west-2
Memory	68 GB
Cluster IP	10.0.42.93
Cluster AE URL	https://10.0.42.120/ae/#/?id=1
Dashboard	↗
Status	Installing ↗ ↶ Click here to see ongoing status of installation

Wait till Installing turns to Running state

DELETE

13.

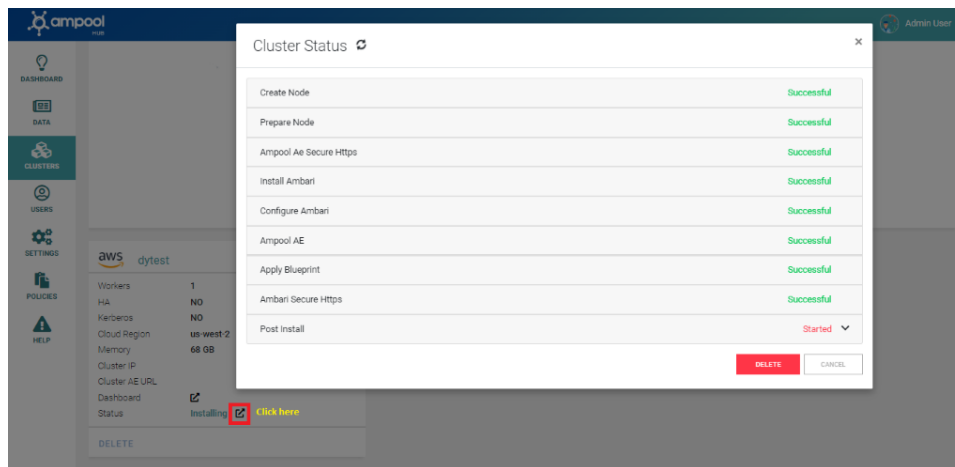
aws dytest

SMALL

Workers	1
HA	NO
Kerberos	NO
Cloud Region	us-west-2
Memory	68 GB
Cluster IP	10.0.42.93
Cluster AE URL	https://10.0.42.120/ae/#/?id=1
Dashboard	🔗
Status	Running 🔗

[EDIT](#) [DELETE](#)

The following screenshot shows the ongoing installation process.



Step	Status
Create Node	Successful
Prepare Node	Successful
Ampool Ae Secure Https	Successful
Install Ambari	Successful
Configure Ambari	Successful
Ampool AE	Successful
Apply Blueprint	Successful
Ambari Secure Https	Successful
Post Install	Started

Once the cluster creation is complete, the **Status** changes to **Running** from **Installing**.

You can click the Cluster URL to go to the Ampool Engine cluster login page, login to Ampool Engine cluster, and start working with Ampool Engine.