Prerequisites-Ampool Hub

This document describes the various prerequisites for your organization to be able to sign up for, set up, and use Ampool Hub and Ampool Engine.

Before you Begin

Ampool Hub must be installed on-premise, or on one of the supported cloud providers, Amazon Web Services (AWS) or Microsoft Azure (Azure).

While signing up for Ampool Hub, Ampool requires you to provide the following details of the superadmin user for Ampool.

- Organization name
- Email ID
- AWS Account ID, if you are installing Ampool Hub on AWS

The AWS account ID is essential to give access to Ampool Engine AMI. The AWS account ID must be subscribed to https://aws.amazon.com/marketplace/pp/B00O7WM7QW (CentOS 7image) or https://aws.amazon.com/marketplace/pp/B00VIMU19E?ref=cns_srchrow (RHEL 7 image) on AWS marketplace, as Ampool Engine custom AMI is based on the AWS marketplace image.

Note: Azure account details would not be required if you want to set up Ampool Hub on Microsoft Azure.

Once the organization name and the email ID are made available to Ampool. Ampool will provide, via email, the license key and detailed instructions to install Ampool Hub.

If you have already received the license key, proceed with the remaining instructions that follow in this document, to satisfy the prerequisites necessary to be able to install and use Ampool Hub and Ampool Engine.

User Role Requirements

User role requirements vary per public cloud provider.

AWS User Role Requirement

If you are installing Ampool Hub on AWS EC2 instance,you must create two separate roles, preferably with the role name, Ampool-Admin for Ampool Hub, and the role name, Ampool-User for Ampool Engine clusters. The user role requirement is specific to AWS.

Ampool-Admin Role: Ampool-Admin role is used to launch and manage Ampool Engine cluster. The role facilitates access to Amazon Glue for metadata browsing.

Ampool-User Role: Ampool-User role is used by Ampool Engine. Currently only Amazon Glue and Amazon S3 services make use of the Ampool-User role that is attached to Ampool Engine cluster instances.

Refer to https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_create_for-service.html to create the Ampool-Admin role for Ampool Hub, and the Ampool-User role for Ampool Engine.

Azure User Role Requirement

If you are installing Ampool Hub on Azure, you must assign Contributor role to the service principal that you are using.

Guideline for Ampool-Admin role

The following minimum permissions are required for Amazon EC2 and Amazon Glue.

```
{
  "Version": "2012-10-17",
  "Statement":
     [
      {
     "Sid": "VisualEditor0",
     "Effect": "Allow",
     "Action":
            [
                "ec2:AttachVolume",
                "ec2:DescribeInstances",
                "ec2:DescribeAddresses",
                "ec2:TerminateInstances",
                "ec2:DeletePlacementGroup",
                "ec2:DescribeTags",
                "ec2:CreateTags",
                "ec2:RunInstances",
                "ec2:StopInstances",
                "ec2:DescribeSecurityGroups",
                "ec2:CreateVolume",
                "ec2:DescribeImages",
                "ec2:DeleteVolume",
                "ec2:StartInstances",
                "ec2:DescribeSubnets",
                "ec2:AssociateAddress",
                "ec2:DescribeAvailabilityZones",
                "ec2:DescribeKeyPairs",
                "ec2:DescribePlacementGroups",
                "ec2:CreatePlacementGroup"
            ],
```

```
"Resource": "*"
},
{
"Sid": "VisualEditor1",
"Effect": "Allow",
    "Action": [
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetPartition",
        "glue:GetTableVersion",
        "glue:GetTables",
        "glue:GetTableVersions",
        "glue:GetPartitions",
        "glue:GetDatabases",
        "glue:GetTable"
    ],
    "Resource": "*"
 },
  {
   "Effect": "Allow",
   "Action":
     [
       "iam:PassRole",
       "iam:ListInstanceProfiles"
     ],
    "Resource": "*"
    }
  ]
  }
```

Guideline for Ampool-User IAM Role

The Ampool-User IAM role will need following minimum set of permissions for Amazon S3 and Amazon Glue.

```
{
    "Version": "2012-10-17",
     "Statement": [{

                "Sid": "VisualEditor0",
                "Effect": "Allow",

           "Action": [
                "glue:SearchTables",
                "glue:BatchCreatePartition",
                "glue:UpdateTable",
                "glue:GetTableVersion",
                "glue:GetTableVersions",
                "glue:GetPartitions",
                "glue:CreatePartition",
                "glue:DeleteTable",
                "glue:UpdatePartition",
                "glue:CreateTable",
                 "glue:GetTables",
                "glue:BatchGetPartition"
                "glue:GetDatabases",
                "glue:GetPartition",
                "glue:GetTable",
                 "glue:GetDatabase",
                "kms:ReEncryptFrom",
                "kms:GenerateDataKeyPair",
                "kms:GenerateDataKey",
                "kms:Decrypt",
```

```
                "kms:GenerateDataKeyWithoutPlaintext",
                "kms:GenerateDataKeyPairWithoutPlaintext",
                "kms:Encrypt",
                "kms:ReEncryptTo",
                "kms:DescribeKey",
                "s3:DeleteObjectVersion",
                "s3:ListBucketVersions",
                "s3:ListBucket",
                "s3:GetBucketPolicy",
                "s3:GetObjectAcl",
                "s3:GetEncryptionConfiguration",
                "s3:AbortMultipartUpload",
                "s3:PutObjectTagging",
                "s3:DeleteObject",
                "s3:HeadBucket",
                "s3:GetBucketPolicyStatus",
                "s3:GetBucketPublicAccessBlock",
                "s3:ListBucketMultipartUploads",
                "s3:ListAccessPoints",
                "s3:PutObjectVersionTagging",
                "s3:GetBucketAcl",
                "s3:PutObject",
                "s3:GetObject",
                "s3:ListAllMyBuckets",
                "s3:PutObjectRetention",
                "s3:GetBucketCORS",
                "s3:GetBucketLocation"
            ],
            "Resource": "*"
        }]
```

```
}
```

Ampool supports authentication through the following two ways.

- Static credentials such as access key and secret access key in case of AWS
- IAM role attached to instances for deployment of Ampool Engine clusters

Note: A user using static credentials must have the same permissions as mentioned in the JSON for the both roles (Ampool-Admin IAM role and Ampool-User IAM role above).

Configure Resource Requirements

The section describes the resource requirements for AWS and Azure. Refer to the relevant section for details.

AWS Resource Requirements

Log in to AWS and create an instance for Ampool Hub as specified in the following table, per your Hub Instance Type.

Hub Instance Type	Instance Specification
Standard Ubuntu 18.04 LTS	t2.medium (2 vCPUs, 4 GiB memory) EBS Volume size min 30 GiB
Standard RHEL 7	t2.medium (2 vCPUs, 4 GiB memory) EBS Volume size min 30 GiB

You can create the Ampool Engine cluster instance type per your requirement on the aforementioned Ampool Hub instance, by referring to the following table.

Ampool Engine Instance Type	Instance Specification
Small (CentOS-based/RHEL- based custom AMI)	c5d.9xlarge (32 vCPUs, 72 GiB memory)
Medium (CentOS-based/RHEL- based custom AMI)	m5d.8xlarge (32 vCPUs, 128 GiB memory)
Large (CentOS-based/RHEL- based custom AMI)	r5d.8xlarge (32 vCPUs, 256 GiB memory)

Azure Resource Requirements

Log in to Azure and create an image for Ampool Hub as specified below.

Hub Instance Type	Instance Specification
Standard Ubuntu 18.04 LTS	Standard D4s v3 (4 vCPUs, 16 GiB memory)

You can create the Ampool Engine cluster instance type per your requirement on the aforementioned Ampool Hub instance, by referring to the following table.

Ampool Engine Instance Type	Instance Specification
Small (CentOS-based)	Standard F32s_v2 (32 vCPUs, 72 GiB memory)
Medium (CentOS-based)	Standard D32s_v3 (32 vCPUs, 128 GiB memory)
Large (CentOS-based)	Standard E32s_v3 (32 vCPUs, 256 GiB memory)

Cloud-specific Prerequisites

Refer to the following cloud specific prerequisites, depending on the cloud provider you are using to set up Ampool Hub.

Azure-related Prerequisites

It is assumed that details about VPC, Subnet, storage, network security group is available which is needed when installing Ampool Hub and Ampool Engine.

AWS-related Prerequisites

It is assumed that details about VPC, Subnet is available which is needed when installing Ampool Hub and Ampool Engine.

Network Access Requirements

Ampool Engine cluster creation requires outbound network access to install standard third-party packages such as python-pip, and Ampool Engine packages.

Ampool Hub registration with Ampool Admin Hub is done over https post installation.

Clients accessing Ampool Hub and Ampool Engine should be able to accept self-signed certificates through browsers.

Data Source Related Prerequisites

The following table describes the data source related prerequisites.

Note: The data source must be accessible from Ampool Engine.

Data Source	Accessibility	User Permissions
Amazon Glue with Amazon S3	AWS access key and secret are required, or the Ampool-User role is required.	Key should have Read/Write/List access for Glue. Same key/role should have S3 object read/list permission.
Amazon Redshift	JDBC user must access to necessary databases and tables from Redshift. Redshift cluster endpoint is required to connect through Ampool	Read and Write permission is mandatory for the same user.
Snowflake	A JDBC user must have access to necessary databases and tables from Snowflake.	Read and Write permission is mandatory for the same user.
Teradata	A JDBC user must have necessary privileges to access required databases and tables from Teradata.	Read and Write permission is mandatory for the same user.
MSSQL Server	A JDBC user must have necessary privileges to access required databases and tables from MSSQL Server.	Read and Write permission is mandatory for the same user.

Once these prerequisites are met, you can go ahead and set up Ampool Hub. Refer to the *Setting up Ampool Hub* document for details.