Sahara MapR Plugin Documentation

Release 10.1.0.dev5

Sahara team

Nov 02, 2023

CONTENTS

USER GUIDE

1.1 MapR Distribution Plugin

The MapR Sahara plugin allows to provision MapR clusters on OpenStack in an easy way and do it, quickly, conveniently and simply.

1.1.1 Operation

The MapR Plugin performs the following four primary functions during cluster creation:

1. MapR components deployment - the plugin manages the deployment of the required software to the target VMs

2. Services Installation - MapR services are installed according to provided roles list

3. Services Configuration - the plugin combines default settings with user provided settings

4. Services Start - the plugin starts appropriate services according to specified roles

1.1.2 Images

The Sahara MapR plugin can make use of either minimal (operating system only) images or pre-populated MapR images. The base requirement for both is that the image is cloud-init enabled and contains a supported operating system (see http://maprdocs.mapr.com/home/InteropMatrix/r_os_matrix.html).

The advantage of a pre-populated image is that provisioning time is reduced, as packages do not need to be downloaded which make up the majority of the time spent in the provisioning cycle. In addition, provisioning large clusters will put a burden on the network as packages for all nodes need to be downloaded from the package repository.

Table 1: Support matrix for the *mapr* plugin

Version (image tag)	Distribution	Build method	Version (build parameter)	Notes
5.2.0.mrv2	Ubuntu 14.04, CentOS 7	sahara-image-pack	5.2.0.mrv2	
5.2.0.mrv2	Ubuntu 14.04, CentOS 7	sahara-image-create	5.2.0	

For more information about building image, refer to Sahara documentation.

MapR plugin needs an image to be tagged in Sahara Image Registry with two tags: 'mapr' and '<MapR version>' (e.g. '5.2.0.mrv2').

The default username specified for these images is different for each distribution. For more information, refer to the registering image section of the Sahara documentation.

1.1.3 Hadoop Version Support

The MapR plugin currently supports Hadoop 2.7.0 (5.2.0.mrv2).

1.1.4 Cluster Validation

When the user creates or scales a Hadoop cluster using a mapr plugin, the cluster topology requested by the user is verified for consistency.

Every MapR cluster must contain:

- at least 1 *CLDB* process
- exactly 1 *Webserver* process
- odd number of *ZooKeeper* processes but not less than 1
- *FileServer* process on every node
- at least 1 ephemeral drive (then you need to specify the ephemeral drive in the flavor not on the node group template creation) or 1 Cinder volume per instance

Every Hadoop cluster must contain exactly 1 *Oozie* process

Every MapReduce v1 cluster must contain:

- at least 1 *JobTracker* process
- at least 1 *TaskTracker* process

Every MapReduce v2 cluster must contain:

- exactly 1 *ResourceManager* process
- exactly 1 *HistoryServer* process
- at least 1 *NodeManager* process

Every Spark cluster must contain:

- exactly 1 *Spark Master* process
- exactly 1 *Spark HistoryServer* process
- at least 1 *Spark Slave* (worker) process

HBase service is considered valid if:

- cluster has at least 1 *HBase-Master* process
- cluster has at least 1 *HBase-RegionServer* process

Hive service is considered valid if:

- cluster has exactly 1 *HiveMetastore* process

- cluster has exactly 1 *HiveServer2* process

Hue service is considered valid if:

- cluster has exactly 1 *Hue* process

- *Hue* process resides on the same node as *HttpFS* process

HttpFS service is considered valid if cluster has exactly 1 *HttpFS* process

Sqoop service is considered valid if cluster has exactly 1 *Sqoop2-Server* process

1.1.5 The MapR Plugin

For more information, please contact MapR.

CONTRIBUTOR GUIDE

2.1 So You Want to Contribute...

For general information on contributing to OpenStack, please check out the contributor guide to get started. It covers all the basics that are common to all OpenStack projects: the accounts you need, the basics of interacting with our Gerrit review system, how we communicate as a community, etc.

sahara-plugin-mapr is maintained by the OpenStack Sahara project. To understand our development process and how you can contribute to it, please look at the Sahara project's general contributor's page: http://docs.openstack.org/sahara/latest/contributor/contributing.html