Configure High Availability for Oozie

To configure HA for Oozie on your cluster, the cluster must meet the following prerequisites:

- Your cluster must have access to a database with support for multiple concurrent connections. To prevent this database from becoming a single point of failure, the database must support HA. HA for the Oozie service works regardless of the database's HA status. See MySQL Data Store for Oozie (MySQLDataStoreforOozie.html#MySQLDataStoreforOozie).
- To prevent the ZooKeeper service from becoming a single point of failure, your cluster must have at least 3 ZooKeeper nodes. You can still configure HA for Oozie on clusters with a single ZK node.
- Multiple nodes on the cluster must have Oozie installed. For installation instructions, see the Installing MapR and/or MapR Ecosystem Components (../AdvancedInstallation/InstallationGuide.html#InstallationGuide).

  Note: For greater consistency of behavior on your cluster, verify that all of the Oozie servers have the same configuration.

- A load balancer, virtual IP, or round-robin DNS set up, such as HAProxy. To prevent the load balancer from becoming a single point of failure, the load balancer must support HA.

1. Verify that the Oozie servers are all configured to connect to the same database. Do not start Oozie.
2. On each Oozie node, edit the `oozie-site.xml` file to add the following section, which changes the results in Oozie using the Zookeeper's version of the services, overriding the default implementations:

   ```
   <property>
       <name>oozie.services.ext</name>
       <value>
           org.apache.oozie.service.ZKLocksService,
           org.apache.oozie.service.ZKXLogStreamingService,
           org.apache.oozie.service.ZKJobsConcurrencyService,
           org.apache.oozie.service.ZKUUIDService
       </value>
   </property>
   ```

3. On each Oozie node, edit the `oozie-site.xml` file to include a comma-separated list of the host names and ports for the ZooKeeper servers. For example:

   ```
   <property>
        <name>oozie.zookeeper.connection.string</name>
        <value>zk1:5181,zk2:5181,zk3:5181</value>
   </property>
   ```

4. On each Oozie node, edit the `oozie-site.xml` file to specify the namespace. Each Oozie server that communicates to other Oozie servers must use the same namespace:

   ```
   <property>
        <name>oozie.zookeeper.namespace</name>
        <value>oozie</value>
   </property>
   ```

5. On each Oozie node, change the value of the `OOZIE_BASE_URL` property in the `oozie-site.xml` file to point to the load balancer or virtual IP.

   ```
   <property>
       <name>oozie.base.url</name>

   <value>http://my.loadbalancer.hostname:<oozie_port_number>/oozie</value>
   </property>
   ```

   The `<oozie_port_number>` (../AdvancedInstallation/InstallOozie.html#InstallOozie__step_djg_dp4_qbb) depends on whether your cluster is secure.

6. *(Optional)* On each Oozie node, change the value of the OOZIE_INSTANCE_ID property in the `oozie-env.sh` file to make the instance ID of each Oozie server unique. The default value for this property is ${OOZIE_HTTP_HOSTNAME}.

   ```
   export OOZIE_INSTANCE_ID="${OOZIE_HTTP_HOSTNAME}"
   ```

7. On all nodes, update the services line in warden.oozie.conf (stored in `/opt/mapr/conf/conf.d`) from:

   ```
   services=oozie:1:cldb
   ```

   to

   ```
   services=oozie:all:cldb
   ```

8. Start the Oozie servers. See Starting and Stopping Oozie Services (ManageOozieServicesSrtStp.html).
9. Run Oozie share lib update command to make sure that all Oozie services use the latest and the same version:

   ```
   {OOZIE_HOME}/bin/oozie admin -sharelibupdate
   ```