Configuring MapR Security

Security features for the MapR distribution for Hadoop are disabled by default. You can enable security features at any time, but additional configuration is required for the individual components to work with security enabled, particularly elements that use Kerberos for user authentication. This section discusses initial configuration of a secure cluster and guides you through securing individual aspects of MapR along with open source components.

The following access control elements are active whether or not your cluster's security features are enabled:

Security Feature	Description
Access Control Lists (ACLs)	For the cluster, the volumes in the cluster, and the MapReduce job queue
Access Control Expressions (ACEs)	Control user permissions for MapR-DB tables that are stored natively
File permissions	For objects in the MapR-FS layer
Subnet whitelisting	Restricts access to the cluster's FileServer service

Once security features are enabled, these elements benefit from encrypted traffic within the cluster and strong authentication to the cluster.

On clusters with security features enabled, several open-source components require additional configuration:

- Hive functionality has different security requirements depending on the interaction between the HiveServer2 component, the Hive command-line interface, and the Hive metastore.
- HBase functionality requires configuration to use Kerberos for securing the HBase RPCs.
- Flume functionality can be configured to use either MAPRSASL or Kerberos. Kerberos configuration is only required if Flume is loading HBase logs.
- Oozie clients can communicate with the Oozie server over HTTPS secured by SPNEGO with Kerberos authentication. Java clients can authenticate with MapR tickets.

To enable security features on a cluster running version 3.1 or later of the MapR distribution for Hadoop, see Enabling and Disabling Security Features on Your Cluster.

Java Applications and MapR Security

A secure computing environment places additional requirements on the Java Virtual Machine (JVM) properties of Java clients. The JVMs launched by MapR with scripts, such as those used by the `maprcli`, `hadoop`, or `hbase` commands, have those properties automatically set by the MapR software. The MapR software attempts to set useful values for these properties when you use a JVM you are launching directly, such as when you write a stand-alone Java program. Be aware that existing Java code that sets values for these properties may cause trouble on your cluster.

Property	Default Value	Description
java.security.auth.login.config	`/opt/mapr/conf/mapr.login.conf`	Path to the file that specifies JAAS configurations used by MapR.
javax.net.ssl.trustStore	`/opt/mapr/conf/ssl_truststore`	Controls the truststore used by MapR clients for HTTPS connections
http.auth.preference	`basic`	The default setting disables JVM's default handling of SPNEGO, enabling MapR's Hadoop code to handle SPNEGO authentication.
zookeeper.saslprovider	`com.mapr.security.maprsasl.MaprSaslProvider`	Enables ZooKeeper security.
hadoop.login	`hadoop_default`	Controls the JAAS configuration used by MapR security.

Ports Used by Web Interfaces for Secure Hadoop 2 Services

The MapR Distribution for Hadoop supports Hadoop 2 services built on the YARN framework. While Hadoop 1 services such as JobTracker and TaskTracker use the same ports for the HTTP and HTTPS protocols, the Hadoop 2 services use different ports, listed in the following table:

Service	HTTP (unsecure) port	HTTPS (secure) port

NodeManager	8042	8044
ResourceManager	8088	8090
HistoryServer	19888	19890

Subnet Whitelisting

To provide additional cluster security, you can limit cluster data access to a whitelist of trusted subnets. The `mfs.subnets.whitelist` parameter in `mfs.conf` accepts a comma-separated list of subnets in CIDR notation. If this parameter is set, the FileServer service only accepts requests from the specified subnets.