Runtime Cluster Hosts

and Role Assignments

Cluster hosts can be broadly described as master hosts, utility hosts, gateway hosts, or worker hosts.

- Master hosts run Hadoop master processes such as the HDFS NameNode and YARN Resource Manager.
- Utility hosts run other cluster processes that are not master processes such as Cloudera Manager and the Hive Metastore.
- Gateway hosts are client access points for launching jobs in the cluster. The number of gateway hosts required varies depending on the type and size of the workloads.
- Worker hosts primarily run DataNodes and other distributed processes such as Impalad.

Important

Cloudera recommends that you always enable high availability when Runtime is used in a production environment.

The following tables describe the recommended role allocations for different cluster sizes:

3 - 10 Worker Hosts without High Availability

Master Hosts	Utility Hosts	Gate
Master Host 1: NameNode YARN ResourceManager JobHistory Server ZooKeeper Kudu master Spark History Server	One host for all Utility and Gateway roles: Secondary NameNode Cloudera Manager Cloudera Manager Management Service Hive Metastore HiveServer2 Impala Catalog Server Impala StateStore Hue Oozie	

Flume	
Gateway configuration	

3 - 20 Worker Hosts with High Availability

Master Hosts	Utility Hosts	Gate
Master Host 1:	Utility Host 1:	One
NameNode	Cloudera Manager	Gate
JournalNode	Cloudera Manager Management Service	Hue
FailoverController	Hive Metastore	Hive
YARN ResourceManager	Impala Catalog Server	Flum
ZooKeeper	Impala StateStore	Gate
JobHistory Server	Oozie	confi
Spark History Server	ZooKeeper (requires dedicated disk)	
Kudu master	JournalNode (requires dedicated disk)	
Master Host 2:		
NameNode		
JournalNode		
FailoverController		
YARN ResourceManager		
ZooKeeper		
Kudu master		
Master Host 3:		
Kudu master (Kudu requires an odd		
number of masters for HA.)		

20 - 80 Worker Hosts with High Availability

Master Hosts	Utility Hosts	Gat
Master Host 1:	Utility Host 1:	One
NameNode	Cloudera Manager	Gat
JournalNode	Utility Host 2:	Hue
FailoverController	Cloudera Manager Management Service	Hive
YARN ResourceManager	Hive Metastore	Flur
ZooKeeper	Impala Catalog Server	Gat
Kudu master	Oozie	con
Master Host 2:		
NameNode		
JournalNode		
FailoverController		
YARN ResourceManager		
ZooKeeper		
Kudu master		
Master Host 3:		
ZooKeeper		
JournalNode		
JobHistory Server		
Spark History Server		
Kudu master		

_

80 - 200 Worker Hosts with High Availability

Master Hosts	Utility Hosts	Gat
Master Host 1:	Utility Host 1:	On
NameNode	Cloudera Manager	Gat
JournalNode	Utility Host 2:	Hu
FailoverController	Hive Metastore	Hiv
YARN ResourceManager	Impala Catalog Server	Flu
ZooKeeper	Impala StateStore	Gat
Kudu master	Oozie	cor
Master Host 2:	Utility Host 3:	
NameNode	Activity Monitor	
JournalNode	Utility Host 4:	
FailoverController	Host Monitor	
YARN ResourceManager	Utility Host 5:	
ZooKeeper	Navigator Audit Server	
Kudu master	Utility Host 6:	
Master Host 3:	Navigator Metadata Server	
ZooKeeper	Utility Host 7:	
JournalNode	Reports Manager	
JobHistory Server	Utility Host 8:	
Spark History Server	Service Monitor	
Kudu master		

200 - 500 Worker Hosts with High Availability

Master Hosts	Utility Hosts	Ga
Master Host 1: NameNode JournalNode FailoverController ZooKeeper Kudu master Master Host 2: NameNode JournalNode FailoverController ZooKeeper Kudu master Master Host 3: YARN ResourceManager ZooKeeper JournalNode Kudu master	Utility Host 1: Cloudera Manager Utility Host 2: Hive Metastore Impala Catalog Server Impala StateStore Oozie Utility Host 3: Activity Monitor Utility Host 4: Host Monitor Utility Host 5: Navigator Audit Server Utility Host 6: Navigator Metadata Server Utility Host 7: Reports Manager	Ga Or Ga Hu Hiv Flu Ga co
Master Host 4: YARN ResourceManager ZooKeeper JournalNode Master Host 5:	Utility Host 8: Service Monitor	

JobHistory Server	
Spark History Server	
ZooKeeper	
JournalNode	
We recommend no more than three Kudu	
masters.	
4	

500 -1000 Worker Hosts with High Availability

Master Hosts	Utility Hosts	Gat
Master Host 1:	Utility Host 1:	On
NameNode	Cloudera Manager	Ga
JournalNode	Utility Host 2:	Hu
FailoverController	Hive Metastore	Hiv
ZooKeeper	Impala Catalog Server	Flu
Kudu master	Impala StateStore	Ga
Master Host 2:	Oozie	cor
NameNode	Utility Host 3:	
JournalNode	Activity Monitor	
FailoverController	Utility Host 4:	
ZooKeeper	Host Monitor	
Kudu master	Utility Host 5:	
Master Host 3:	Navigator Audit Server	
YARN ResourceManager	Utility Host 6:	
ZooKeeper	Navigator Metadata Server	
JournalNode	Utility Host 7:	
Kudu master	Reports Manager	
Master Host 4:	Utility Host 8:	
YARN ResourceManager	Service Monitor	
ZooKeeper		
JournalNode		
Master Host 5:		
JobHistory Server		
Spark History Server		
ZooKeeper		
JournalNode		
We recommend no more than three Kudu		
masters.		
1		>