

# HBase系统搭建与部署

主讲人： Louis



# Hbase集群部署



1) 开源社区版本下载

—— <http://hbase.apache.org/>

2) Cloudera发行版下载

—— <http://archive-primary.cloudera.com/cdh5/cdh/5/>

如果选择社区版本，注意jar包的替换

# Hbase集群部署

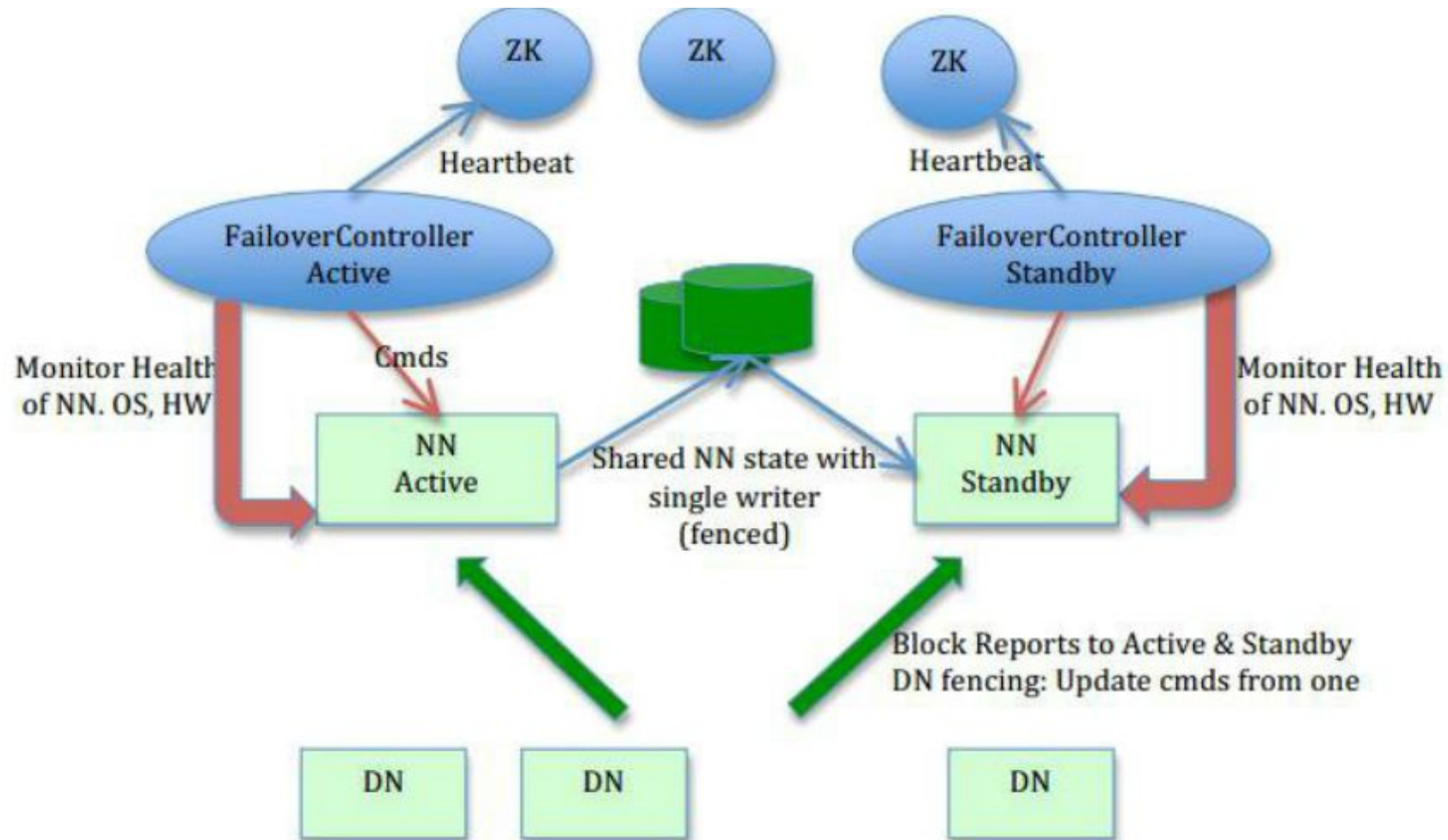
集群规模，一台master，三台slave

节点	运行进程
master(teksource-1)	Active NameNode, ResourceManager HMaster, HQuorumPeer
slave1(teksource-2)	Standby NameNode, DataNode, NodeManager, HRegionServer, JournalNode
slave2(teksource-3)	DataNode, NodeManager, HRegionServer, HQuorumPeer, JournalNode
slave3(teksource-4)	DataNode, NodeManager, HRegionServer, HQuorumPeer, JournalNode

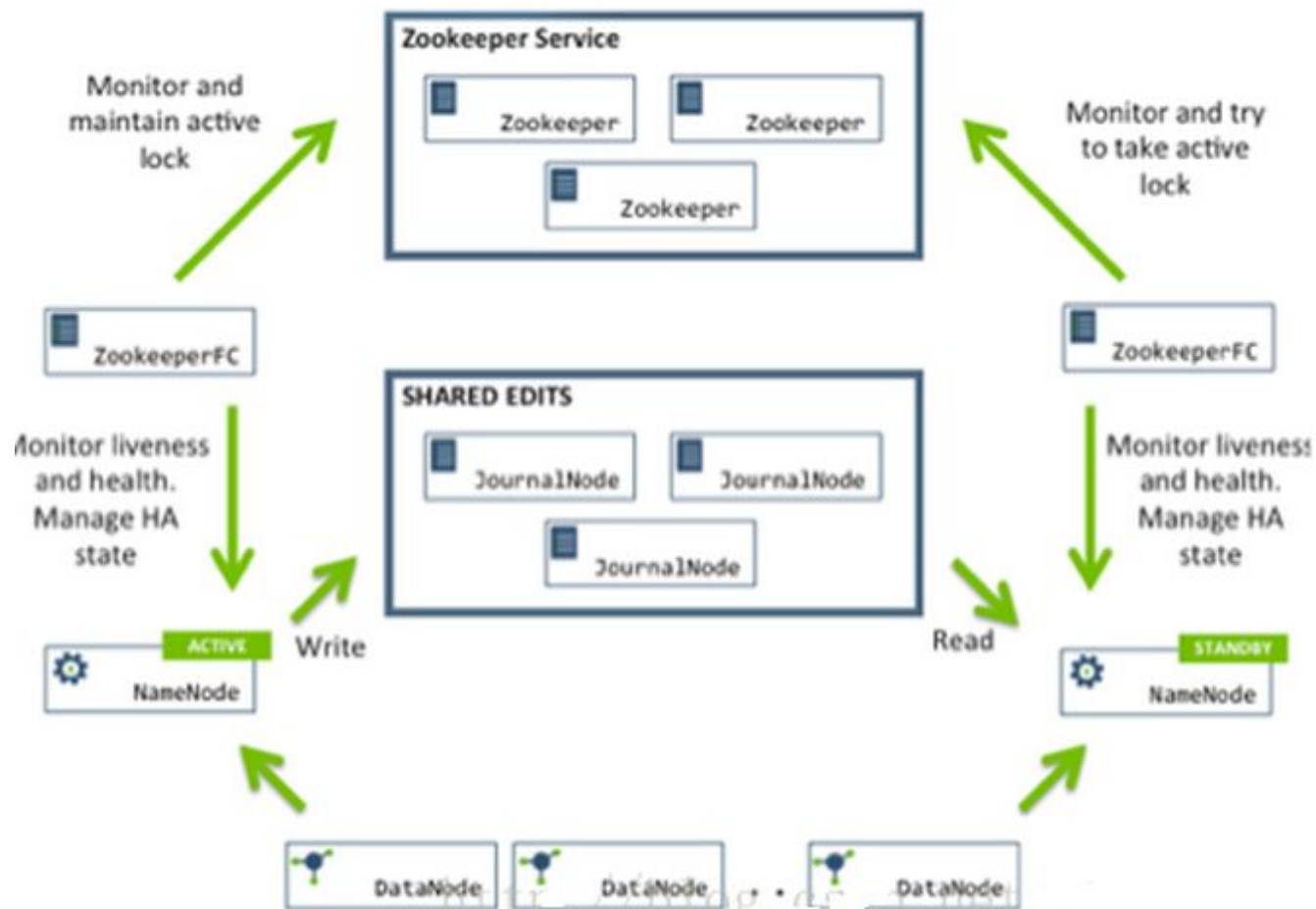
# Hadoop2.0部署

- 1) 配置hosts，确保涉及的主机名均可以解析
- 2) 编辑hadoop-env.sh,mapred-env.sh,yarn-env.sh
- 3) 编辑core-site.xml,hdfs-site.xml,yarn-site.xml
- 5) 编辑slaves文件
- 6) 把Hadoop复制到其他节点
- 7) 启动Hadoop
- 8) 验证启动

# Hadoop2.0 HA部署



# QJM方案



# Hadoop HA特点

## 共享存储

- 解决单点故障问题
- 利用NFS或QJM来保存Editlog
- 保证数据的一致性

## FailoverController

- 独立短小的watchdog
- 避免NN GC时的心跳暂停
- ZKFC控制

## Fencing

- 防止闹裂

# Hadoop2.0 HA部署

## 采用QJM方案

### 1. 启动JN

```
./sbin/hadoop-daemon.sh start journalnode  
./sbin/hadoop-daemons.sh start journalnode  
hdfs namenode -initializeSharedEdits
```

### 2. 启动Active NN

```
hdfs namenode -format  
bin/hadoop-daemon.sh start namenode
```

### 3. 启动Standby NN

```
hdfs namenode -bootstrapStandby  
./sbin/hadoop-daemon.sh start namenode
```

### 4. 启动Automatic Failover

```
hdfs zkfc -formatZK  
./sbin/hadoop-daemon.sh start zkfc
```



# Hbase部署

- 1) 配置hosts，确保涉及的主机名均可以解析
- 2) 编辑hbase-env.sh
- 3) 编辑hbase-site.xml（关联hadoop）
- 4) 覆盖Hadoop核心jar包
- 5) 编辑regionservers文件
- 6) 把Hbase复制到其他节点
- 7) 启动Hbase
- 8) 验证启动