



Intel® Expressway 分布式系统套件 运维手册

1. 简介.....	3
1.1. 文档目的.....	3
1.2. 软件要求.....	3
1.3. 名词解释:	4
2. 集群维护.....	4
2.1. 维护概述.....	4
2.2. 集群或单一组件的启动停止.....	4
2.3. 集群或单一组件的启动停止（控制台方式）	6
2.4. 配置更新.....	7
2.5. 集群许可证更新.....	9
3. 节点维护.....	11
3.1. 维护概述.....	11
3.2. 增加节点（在线扩容）	11
3.3. 删除节点.....	13
3.4. 重启节点.....	14
3.5. 重启节点服务（控制台方式）	15
3.6. 节点日志查看（控制台方式）	16
4. 状态监控.....	16
4.1. HDFS 状态查看（控制台方式）	16
4.2. HDFS 数据检查（控制台方式）	18
4.3. HBase 数据检查（控制台方式）	19
4.4. 集群节点状态检查.....	20

1. 简介

1.1. 文档目的

本文档用于描述如何验证，检查和维护 Intel® Expressway 分布式系统套件的工作情况。

1.2. 软件要求

集群中所有机器都需要安装 Intel® Expressway 分布式系统套件。并且确认集群组件至少满足以下最低配置要求：

HDFS 最低配置：

一台 Namenode，

一台或一台以上 Secondary Namenode，

至多一台 Standby Namenode，

至少大于文件备份数目的 Datanode；

MapReduce 最低配置：

一台 Jobtracker，

至多一台 Backup Jobtracker，

至少一台 Tasktracker；

HBase 最低配置：

至少一台 HBase Master，

三台或三台以上（需要为奇数）数量的 Zookeeper，

至少一台 HBase Region Server，

如需 Thrift 功能，则还需要配置至少一台的 HBase Thrift 服务器；

Hive 最低配置:

至少一台 Hive Server;

管理节点配置:

确保该节点设置为 Ganglia Master,

安装有 Intel® Software Cluster Management Suite 套件。

1.3. 名词解释:

管理节点: 安装有 Ganglia Master 以及 Intel® Software Cluster Management Suite 套件, 并确保 <http://【管理节点】/ganglia> 以及 <https://【管理节点】:9443> 能够访问。

中心节点: 作为集群 Namenode 和 Jobtracker 的服务器。

Zookeeper 节点: 作为集群中 Zookeeper 服务器的任意节点。

从节点: 作为集群 Datanode, Tasktracker, HBase Region Server 的任意节点。

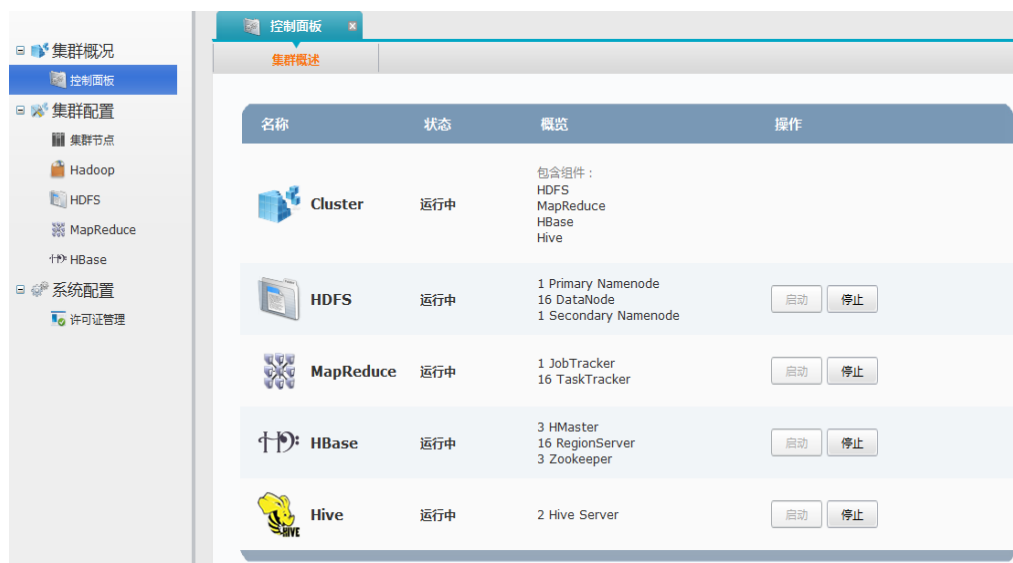
2. 集群维护

2.1. 维护概述

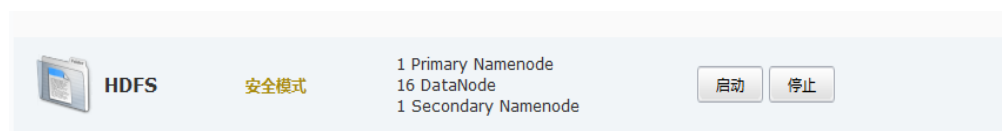
集群维护主要工作需要保证集群的正常运行, 包括启动/停止集群, 启动/停止单一集群, 配置升级, 集群许可证的更新。用于集群中某组件整体工作出现异常时的恢复。

2.2. 集群或单一组件的启动停止

打开 <https://【管理节点】:9443>, 在左侧的下拉列表中选择“控制面板”选项, 打开控制面板页面:



如图所示，在运行中的四样组件分别为 HDFS, MapReduce, HBase 和 Hive。除 HDFS 外，所有组件都只有两种状态“运行中”和“未运行”，在界面的最右侧有操作列表，可以通过点击按钮来对集群单一组件进行“启动/停止”操作。对于 HDFS 而言，启动后会先进入“安全模式”，以浅黄色字样标识（下图），此状态下 HDFS 处于不可写状态，因此不可继续进行任何操作。所以请确保 HDFS 状态正常后方可再进行剩余操作。



如需启动集群，需要严格按照启动顺序：HDFS, MapReduce, HBase, Hive。

单一组件的启动顺序如下：

HDFS 不需要依赖另外组件；

MapReduce 启动之前，需要确保 HDFS 处于运行状态下；

HBase 启动之前，需要确保 HDFS 处于运行状态下；

Hive 启动之前，需要确保 HDFS, MapReduce 以及 HBase 处于运行状态下。

如需停止集群，需要严格按照停止顺序：Hive, HBase, MapReduce, HDFS。

单一组件的停止顺序如下：

HDFS 停止之前，需要确保 MapReduce, HBase, Hive 处于未运行状态下；

MapReduce 停止之前，需要确保 Hive 处于未运行状态下；

HBase 停止之前，需要确保 Hive 处于未运行状态下；

Hive 可以在没有任务执行时直接停止。

2.3. 集群或单一组件的启动停止（控制台方式）

控制台方式，没有提供整体集群的启动/停止方案。对于单一组件的启动/停止，首先登录【控制节点】，启动停止组件的依赖关系和相关顺序与界面方式相同。

HDFS 启动，在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/start-dfs.sh
```

HDFS 停止，在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/stop-dfs.sh
```

MapReduce 启动，在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/start-mapred.sh
```

MapReduce 停止，在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/stop-mapred.sh
```

HBase 启动分为两个步骤，第一个步骤是 Zookeeper 的启动，第二个步骤是 HBase Master 以及 HBase Region Server 的启动。在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/start-zookeeper.sh
```

确保 Zookeeper 服务启动正确完成后执行：

```
bash /usr/lib/hadoop/bin/start-hbase.sh
```

HBase 的停止分为两个步骤，第一个步骤是停止 HBase Master 以及 HBase Region Server，第二个步骤是停止 Zookeeper，在【控制节点】运行命令：

```
bash /usr/lib/hadoop/bin/stop-hbase.sh
```

确保 HBase Master 服务完全停止后执行：

```
bash /usr/lib/hadoop/bin/stop-zookeeper.sh
```

Hive 的启动分为三个阶段，首先需要启动 Derby 服务，在【中心节点】上运行

```
service derbyserver start
```

然后需要启动 Hive Meta Server，在【中心节点】上运行：

```
service hadoop-meta-server start
```

最后在每一台配置过 Hive Server 角色的机器上运行：

```
service hadoop-hive-server start
```

Hive 的停止分为三个阶段，首先停止 Hive Server，在每一台配置过 Hive Server 角色的机器上运行：

```
service hadoop-hive-server stop
```

然后再停止 Hive Meta Server，在【中心节点】上运行：

```
service hadoop-meta-server stop
```

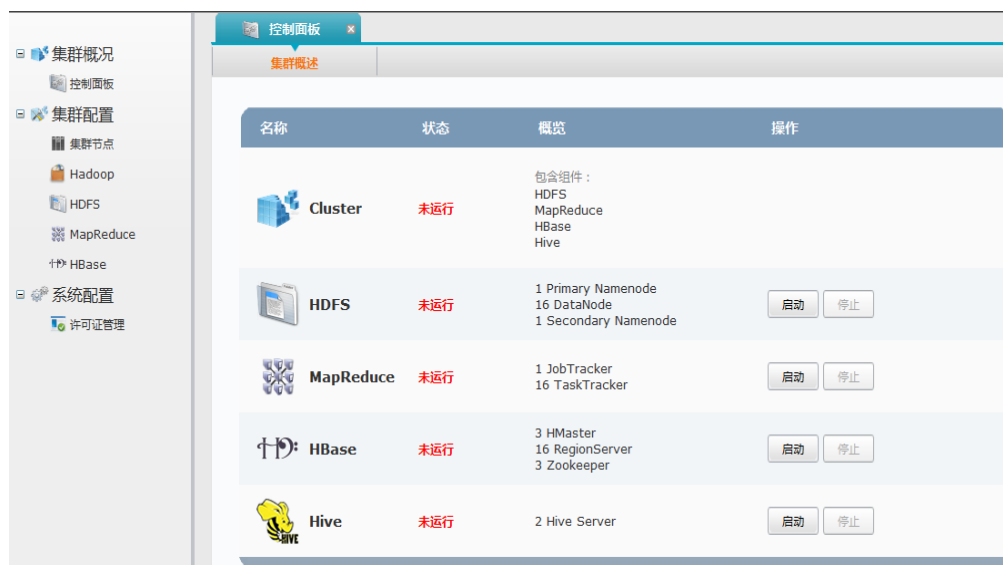
最后在【中心节点】上关闭 Derby 服务：

```
service derbyserver stop
```

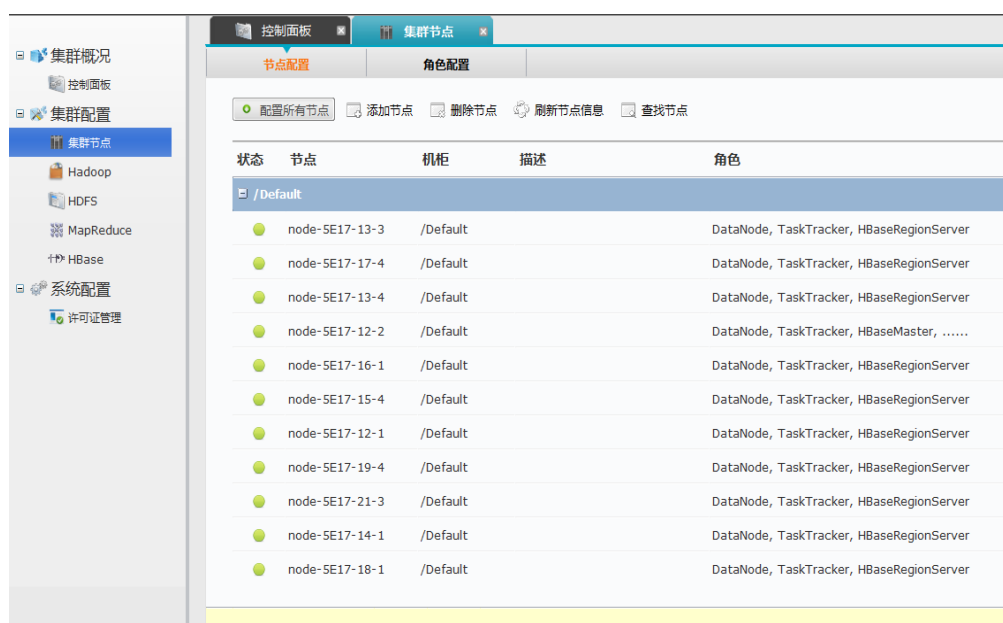
2.4. 配置更新

打开【管理节点】管理页面，地址是：<https://【管理节点】:9443>。

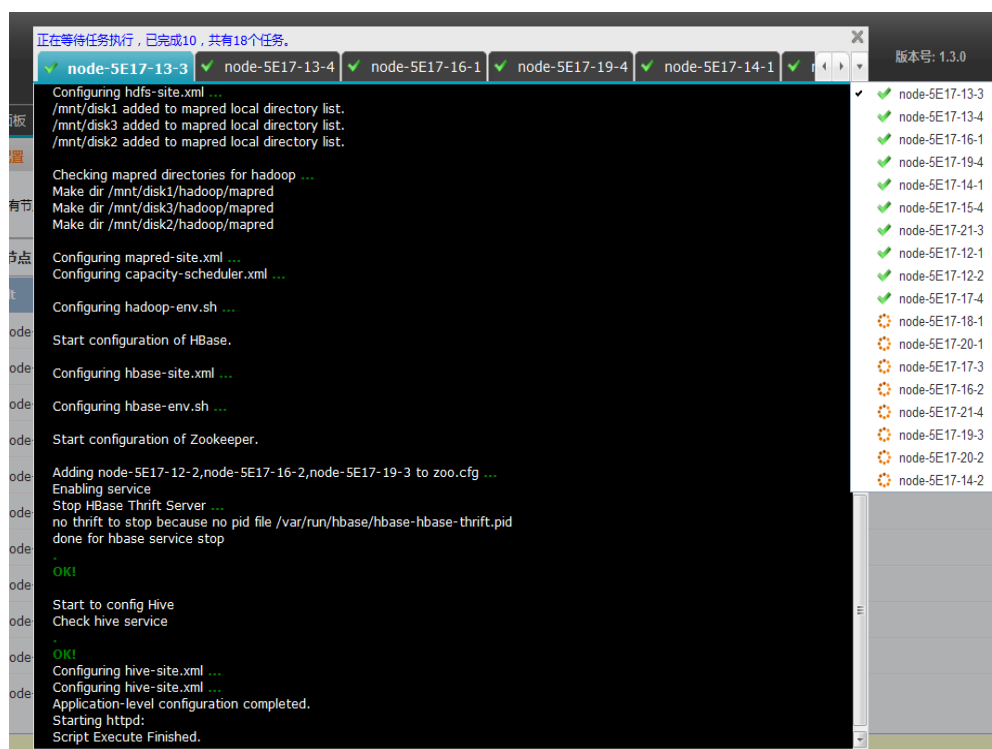
在配置更新前，请确保集群所有服务处于“未运行”状态，如图所示：



然后点击左侧工具栏“集群节点”项，进入集群节点配置界面。



在节点配置上方工具栏中选择“配置所有节点”项，点击后，集群开始重新配置。

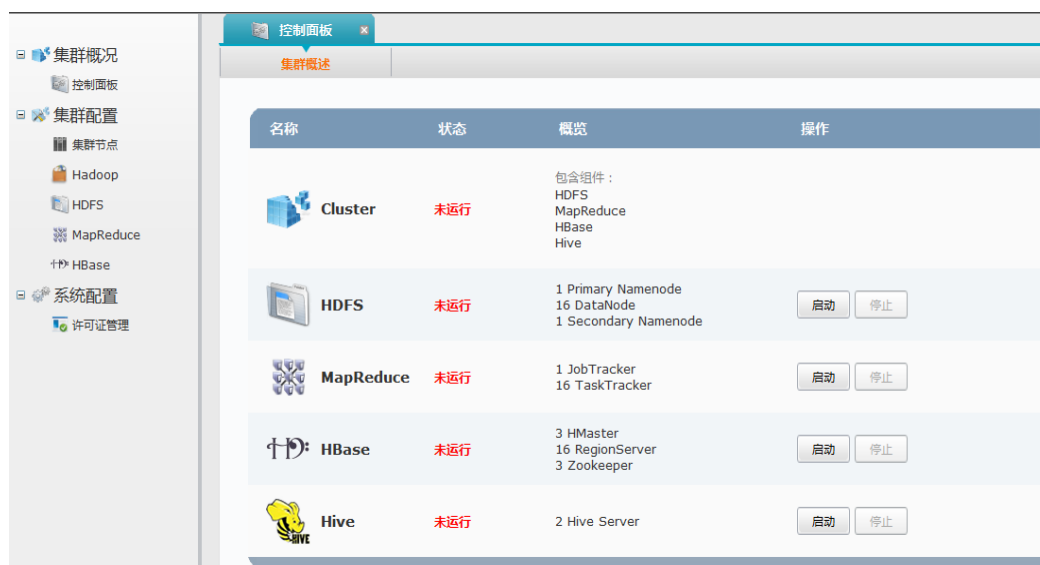


如图所示的配置界面，等到所有等待标志  变成成功标志  后，集群所有配置更新完毕。

2.5. 集群许可证更新

打开【管理节点】管理页面，地址是：<https://【管理节点】:9443>。

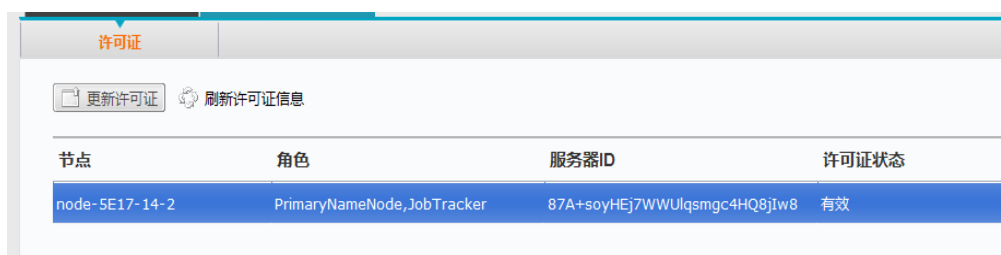
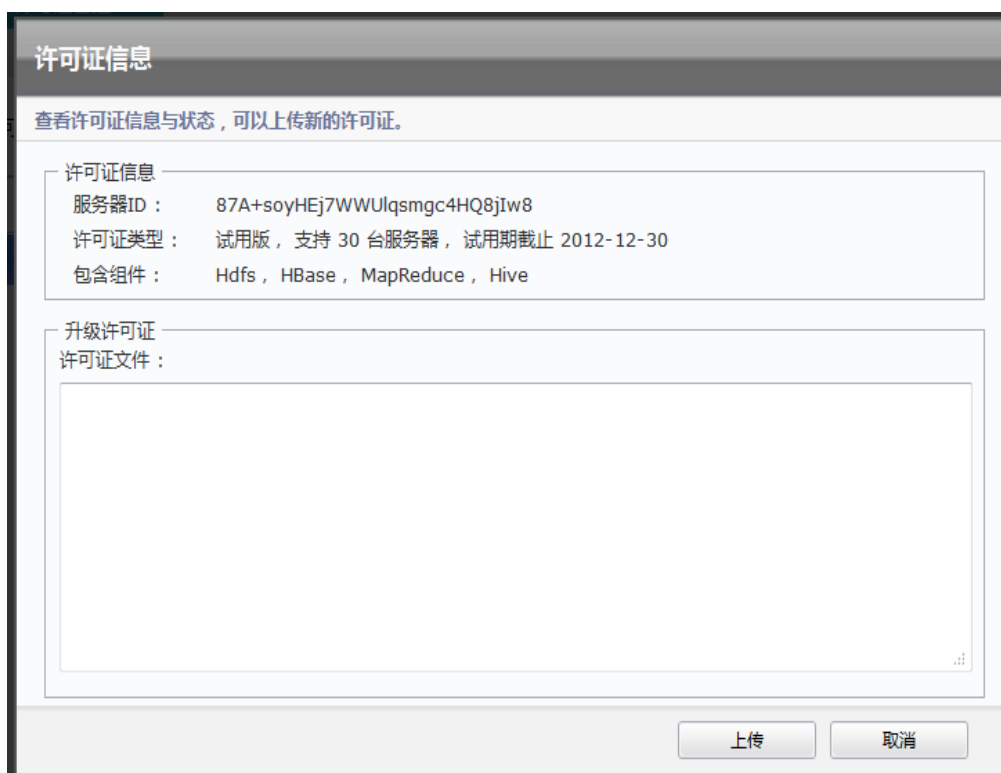
在配置更新前，请确保集群处于关闭状态，如图所示：



然后点击左侧工具栏“许可证管理”项，进入许可证管理界面。



许可证管理界面中列出了所有需要许可证的服务器，双击选择需要更新的服务器，在弹出的窗口中输入新的许可证。



输入许可证字符串后，点击“上传”按钮，上传成功后在页面上方的工具栏中选择“更新许可证”项，在集群中部署新的许可证信息。

3. 节点维护

3.1. 维护概述

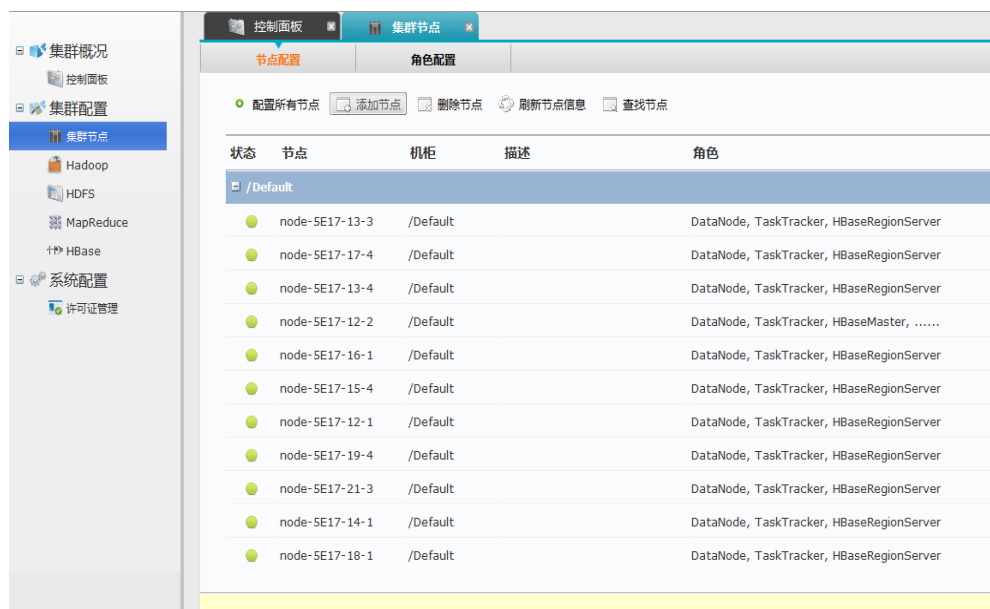
节点维护主要指对单一节点的维护，包括增加/删除节点，对节点上服务的启动/停止等常用功能，用于单一节点故障时的修复或对集群进行扩容等常见维护工作。

3.2. 增加节点（在线扩容）

增加节点操作仅仅限于对 Tasktracker/Datanode/HBase Region Server/Hive Server 等非中心主控节点生效，增加节点过程中不会影响现有组件的正常运行。

如果需要增加核心节点，则需要在添加节点后重启集群，请参考“集群维护”章节中“重新配置集群”操作。

打开【管理节点】管理页面，地址是：<https://【管理节点】:9443>。



状态	节点	机柜	描述	角色
	/Default			
	node-5E17-13-3	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-17-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-13-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-12-2	/Default		DataNode, TaskTracker, HBaseMaster,
	node-5E17-16-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-15-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-12-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-19-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-21-3	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-14-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-18-1	/Default		DataNode, TaskTracker, HBaseRegionServer

在左侧列表中选择“集群节点”项，点击进入，在页面的上方工具条中选择“添加节点”按钮，点击进入“添加节点”页面。

添加新节点

添加新节点,如果集群只有DNS服务器,只需要输入主机名;否则还要输入IP地址。

节点IP地址 : 添加的节点的IP地址。

完成配置取消

点击“完成”按钮后机器被添加进入机器列表中。

双击新添加的机器，打开角色设置页面。

配置节点 node-5E17-13-3

配置节点的信息，如角色，机架等。

一般集群是由中心节点（一台），主要节点（三或五台）和一般节点（多台）构成。一般来说中心节点可靠性要求最高，主要节点需要分布在不同机架上。

机架：

模板：

描述：

HDFS

☐ PrimaryNameNode
 ☐ StandbyNameNode
 ☐ SecondaryNameNode
 ☒ DataNode

Map/Reduce

☐ JobTracker
 ☐ BackupJobTracker
 ☒ TaskTracker

HBase

☐ HBaseMaster
 ☒ HBaseRegionServer
 ☐ Zookeeper
 ☐ HBaseThrift

Hive

☐ HiveServer

Monitor

☐ GangliaServer

完成

取消

选择所需要配置的角色，这里仅仅限于 Tasktracker, HBase Region Server, Datanode, Hive Server 这几个角色。选择合适的角色后，点击“完成”按钮后完成角色配置。

右键点击需要配置的节点，选择“配置节点”项，单独配置一台机器，配置完成后同样再右键选择该机器，选择“启动服务”，确认服务启动成功后即可完成在线扩容。

集群概况
 控制面板
 集群配置
 集群节点
 Hadoop
 HDFS
 MapReduce
 HBase
 系统配置
 许可证管理

控制面板

集群节点

节点配置

角色配置

配置所有节点
 添加节点
 删除节点
 刷新节点信息
 查找节点

状态	节点	机架	描述	角色
	/Default			
●	node-5E17-13-3	/Default	添加节点 网络配置 配置节点 启动服务 删除节点 重启节点 格式化NameNode 格式化Datanode	DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-17-4	/Default		DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-13-4	/Default		DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-12-2	/Default		DataNode, TaskTracker, HBaseMaster,
●	node-5E17-16-1	/Default		DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-15-4	/Default		DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-12-1	/Default		DataNode, TaskTracker, HBaseRegionServer
●	node-5E17-19-4	/Default		DataNode, TaskTracker, HBaseRegionServer

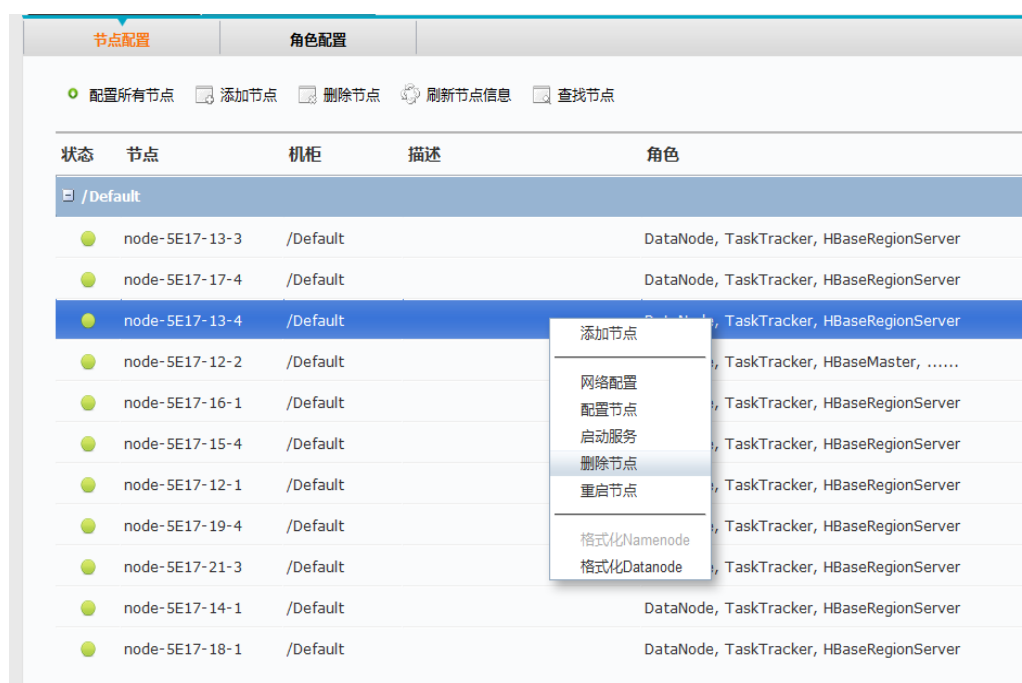
3.3. 删除节点

删除节点用于单一节点的退役/移除/硬件故障排查等等，需要注意的是，单一节点的删除操作仅仅限于 Tasktracker/Datanode/HBase Region Server/Hive Server 等非中心节点。对于 Datanode，请勿一次移除大于等于复制份数的机器，否则会造成数据丢失等不可控制事件的发生。

如需删除核心节点，则在删除完成后需要重启集群，请参考“集群维护”章节中“重新配置集群”操作。

打开【管理节点】管理页面，地址是：<https://【管理节点】:9443>。

在左侧列表中选择“集群节点”项，点击进入，选择需要删除的节点，鼠标右键点击，并在弹出的下拉菜单中选择“删除节点”项。



3.4. 重启节点

重启节点用于单一节点的硬件故障排查或应用其他的系统设置，或当操作系统本身出现故障时对于操作系统本身的故障排查处理。需要注意的是，单一节点的重启仅仅限于 Tasktracker/Datanode/HBase Region Server/Hive Server 等非中心节点。对于 Datanode，请勿一次重启大于等于复制份数的机器，否则会造成数据丢失等不确定事件的发生。

如需重启核心节点，则需要在集群完全停止工作后进行。

打开【管理节点】管理页面，地址是：<https://【管理节点】:9443>。

在左侧列表中选择“集群节点”项，点击进入，选择需要删除的节点，鼠标右键点击，并在弹出的下拉菜单中选择“重启节点”项。

节点配置

角色配置

配置所有节点

添加节点

删除节点

刷新节点信息

查找节点

状态	节点	机柜	描述	角色
/Default				
	node-5E17-13-3	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-17-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-13-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-12-2	/Default		DataNode, TaskTracker, HBaseMaster,BaseRegionServer
	node-5E17-16-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-15-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-12-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-19-4	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-21-3	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-14-1	/Default		DataNode, TaskTracker, HBaseRegionServer
	node-5E17-18-1	/Default		DataNode, TaskTracker, HBaseRegionServer

3.5. 重启节点服务（控制台方式）

重启单一节点的服务主要应用于对于单一节点，单一组件的故障修复和排查。该操作可以通过控制台进行。

需要注意的是，该操作仅能对 Tasktracker/Datanode/HBase Region Server/Hive Server 等非中心节点进行。如果需要对中心节点进行操作，请参考“集群维护”中集群或单一组件的维护的相关章节。

Datanode 的重启：

用控制台方式，登录进入需要重启的节点中，运行命令

```
service hadoop-0.20-datanode restart
```

Tasktracker 的重启：

用控制台方式，登录进入需要重启的节点中，运行命令

```
service hadoop-0.20-tasktracker restart
```

HBase Region Server 的重启：

用控制台方式，登录进入需要重启的节点中，运行命令

```
service hadoop-hbase-regionserver restart
```

Hive Server 的重启：

用控制台方式，登录进入需要重启

hadoop-hadoop-datanode-node-5E17-21-4.log 的节点中，运行命令

```
service hadoop-0.20-tasktracker restart
```

3.6. 节点日志查看（控制台方式）

查看日志主要用于分析和发现集群组件运行的问题，不同组件的日志是完全分离开的。

HDFS, MapReduce 日志目录：/var/log/hadoop/

HBase 日志目录：/var/log/hbase/

Zookeeper 日志目录：/var/log/zookeeper/

操作命令：

例如想查看机器 “node-5E17-21-4” 上 Datanode 的日志：

首先通过控制台登录到服务器 “node-5E17-21-4” 上，然后运行

```
cat /var/log/hadoop/hadoop-hadoop-datanode-node-5E17-21-4.log
```

以此类推，可以在不同的机器上可以查看不同的组件角色的日志。

4. 状态监控

4.1. HDFS 状态查看（控制台方式）

登录到集群中任意节点，运行以下命令：

```
sudo -u hdfs hadoop dfsadmin -report
```

在输出的前段有如下的内容：

Configured Capacity: 64264650747904 (58.45 TB)

Present Capacity: 60503753035398 (55.03 TB)

DFS Remaining: 52110522388480 (47.39 TB)

DFS Used: 8393230646918 (7.63 TB)

DFS Used%: 13.87%

Under replicated blocks: 2

Blocks with corrupt replicas: 0

Missing blocks: 2

Datanodes available: 16 (16 total, 0 dead)

同时改命令还能用于检查当前活跃的 Datanode 节点数量:

例如以下输出:

Name: 192.168.1.112:50010

Decommission Status : Normal

Configured Capacity: 1768683216896 (1.61 TB)

DFS Used: 544110604288 (506.74 GB)

Non DFS Used: 90259705856 (84.06 GB)

DFS Remaining: 1134312906752(1.03 TB)

DFS Used%: 30.76%

DFS Remaining%: 64.13%

Last contact: Wed Mar 28 17:44:16 CST 2012



Name: 192.168.1.120:50010

Decommission Status : Normal

Configured Capacity: 3938013904896 (3.58 TB)

DFS Used: 504170627072 (469.55 GB)

Non DFS Used: 200449388544 (186.68 GB)

DFS Remaining: 3233393889280(2.94 TB)

DFS Used%: 12.8%

DFS Remaining%: 82.11%

Last contact: Wed Mar 28 17:44:15 CST 2012

Name: 192.168.1.116:50010

Decommission Status : Normal

Configured Capacity: 4781493559296 (4.35 TB)

DFS Used: 565249060864 (526.43 GB)

Non DFS Used: 243505713152 (226.78 GB)

DFS Remaining: 3972738785280(3.61 TB)

DFS Used%: 11.82%

DFS Remaining%: 83.09%

Last contact: Wed Mar 28 17:44:17 CST 2012

4.2. HDFS 数据检查（控制台方式）

登录到集群中任意节点，运行以下命令：

```
sudo -u hdfs hadoop fsck /
```

可以得到类似一下的输出，阅读输出可以得知 HDFS 运行的状态。需要注意，该命令在 HDFS 储存内容增多后运行时间会增长，并且对 HDFS 系统会造成少量压力，所以运行次数不能过于频繁。

```
FSCK started by webuser from /192.168.2.1 for path / at Wed Mar 28
17:06:32 CST 2012
```

```
.....
.....
.....
```

4.3. HBase 数据检查（控制台方式）

登录到集群任意节点上，运行命令：

```
hbase hbck
```

在输出中可以找到类似一下内容。

```
Number of Tables: 22
```

```
Number of live region servers: 16
```

```
Number of dead region servers: 0
```

```
.Number of empty REGIONINFO_QUALIFIER rows in .META.: 0
```

```
Summary:
```

```
-ROOT- is okay.
```

```
Number of regions: 1
```

```
Deployed on: node-5E17-14-1:60020
```

```
.META. is okay.
```

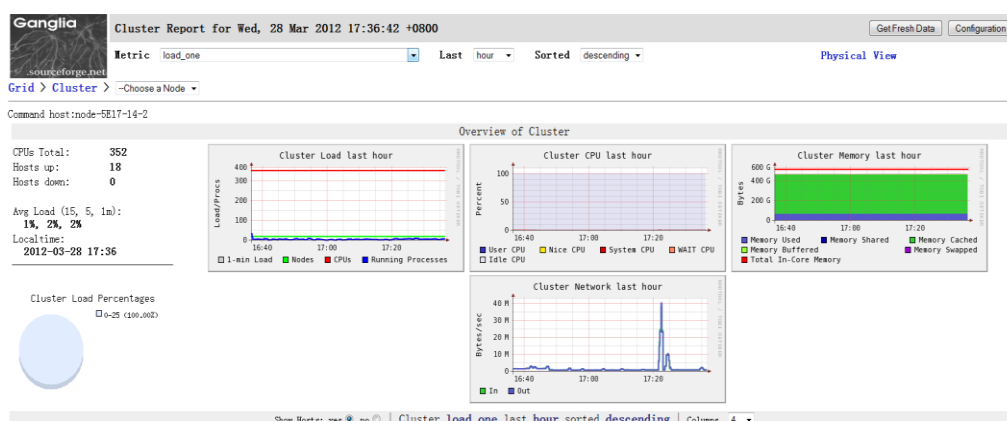
```
Number of regions: 1
```

```
Deployed on: node-5E17-20-1:60020
```

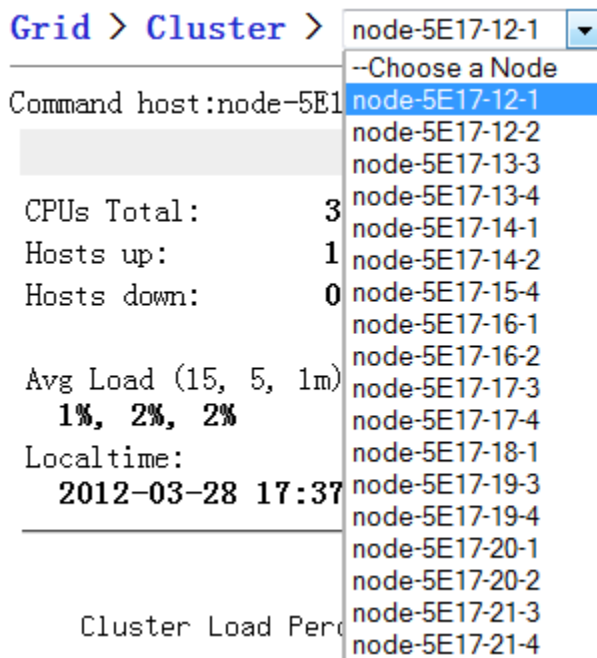
可以发现到目前 HBase 处于一个稳定的状态，-ROOT-表和.META.表都是一个 okay 的状态，证明这个集群处于一个正确的工作状态。

4.4. 集群节点状态检查

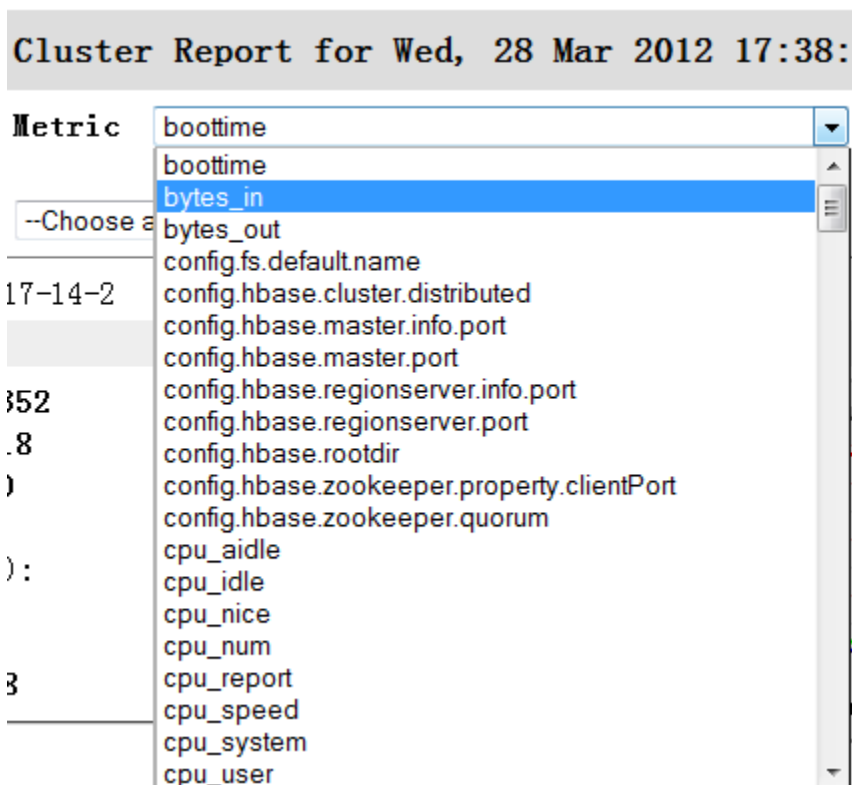
访问【控制节点】Ganglia 页面，在浏览器中输入地址：<http://【控制节点】/ganglia> 进入状态监控界面。



在“Choose a Node”下拉单中可以选择查看单一节点：



在“Metric”下拉列表中选择需要查看的具体表项，比如 CPU 利用率，内存利用率，硬盘空间利用率，以及 Hadoop 集群状态参数等内容查看。如下图所示：



选择需要监控的系统参数之后，就可以浏览相应的 **Metric** 数据。