

# FusionInsight HD V100R002C60U10 容量调整指导书

文档版本 05

发布日期 2017-10-30



## 版权所有 © 华为技术有限公司 2017。 保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传 播。

## 商标声明

HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标,由各自的所有人拥有。

## 注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束,本文档中描述的全部或部分产品、服务 或特性可能不在您的购买或使用范围之内。除非合同另有约定,华为公司对本文档内容不做任何明示或默示 的声明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约定,本文档仅作为使用指导,本 文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

# 华为技术有限公司

地址: 深圳市龙岗区坂田华为总部办公楼 邮编: 518129

网址: http://e.huawei.com

i

# 目录

1 扩容概述	1
1.1 扩容场景	1
1.2 何时进行服务扩容	2
1.2.1 何时扩容 HDFS DataNode	2
1.2.2 何时扩容 Yarn NodeManager	2
1.2.3 何时扩容 HBase RegionServer	2
1.2.4 何时扩容 Solr SolrServer	3
1.2.5 何时扩容 Kafka Broker	3
1.2.6 何时扩容 Streaming Supervisor	4
1.2.7 何时扩容 Redis	4
1.3 扩容前准备	4
1.3.1 扩容规划	4
1.3.2 环境检查	5
1.3.3 辅助参考文档	5
1.4 扩容约束	5
1.5 扩容影响	5
1.6 扩容期间禁止的操作	5
2 扩容操作指导	6
2.1 集群规模与组网规划	6
2.1.1 简介	6
2.1.2 节点规模扩张	6
2.1.3 添加组件服务	7
2.2 添加新节点至集群	7
2.2.1 使用模板添加新节点至集群	7
2.2.1.1 使用配置规划工具(LLD)创建扩容模板	7
2.2.1.2 上传配置文件到主 OMS 节点并进行预处理	9
2.2.1.3 执行 Preinstall	10
2.2.1.4 导入扩容模板文件	12
2.2.1.5 扩容验证	12
2.2.2 手动添加新节点至集群	13

2.2.2.1 环境准备	13
2.2.2.2 操作步骤	14
2.3 添加新节点主机名至客户端节点	15
2.4 添加组件实例	15
2.4.1 添加服务	15
2.4.2 添加 Name Service	16
2.4.3 添加实例	17
2.5 升级服务器	17
2.6 DataNode 扩容后 Balance	18
2.7 DataNode 扩容后 Colocation	19
2.7.1 应用场景	19
2.7.2 操作步骤	19
2.8 Solr 扩容	20
2.8.1 Solr 组件扩容后增加 Shard	20
2.8.2 Solr 组件扩容后进行 split 操作	21
2.8.3 Solr 组件扩容后新增 Collection	
2.9 Kafka 扩容	26
2.9.1 Kafka 磁盘扩容	26
2.9.2 Kafka 扩容后 Reassign	26
2.10 为 Redis 业务集群添加 Redis 实例	
3 扩容验证	30
4 减容概述	31
4.1 减容场景	31
4.2 何时进行服务减容	31
4.2.1 何时减容 HDFS DataNode	31
4.2.2 何时减容 Yarn NodeManager	32
4.2.3 何时减容 HBase RegionServer	32
4.2.4 何时减容 Solr SolrServer	32
4.2.5 何时减容 Redis	
4.2.6 何时减容 Kafka Broker	32
4.2.7 何时减容 Streaming Supervisor	33
4.3 减容前准备	33
4.3.1 减容规划	33
4.3.2 环境检查	33
4.3.2.1 检查服务状态	33
4.3.2.2 检查告警	33
4.3.2.3 检查 Kafka Topic 同步状态	
4.4 减容约束	35
N. N. H. J.	35

4.6 减容期间禁止的操作	35
5 DataNode 减容操作指导	36
5.1 DataNode 减容前退服	36
5.2 获取节点的 Instance	37
5.3 删除 Instance	37
5.4 减容节点	38
5.5 DataNode 减容后 Colocation	38
5.5.1 应用场景	38
5.5.2 操作步骤	38
6 其他组件减容操作指导	40
6.1 减容 Red is 集群	40
6.2 NodeManager 减容前退服	
6.3 Solr 减少组件实例	42
6.4 Kafka 减容后的 Reassign	43
7 减容验证	46

# **1** 扩容概述

- 1.1 扩容场景
- 1.2 何时进行服务扩容
- 1.3 扩容前准备
- 1.4 扩容约束
- 1.5 扩容影响
- 1.6 扩容期间禁止的操作

# 1.1 扩容场景

当前典型的扩容场景包括:

- 安装新服务
  - 集群在初始安装时某些组件未安装
  - FusionInsight 升级后支持更多的组件,需要在现有网络拓扑下安装新组件
- 扩容已有服务

安装完成后的集群某组件的服务能力无法满足业务需求,需要对不满足服务能力需求的组件进行扩容,支持业务节点扩容的组件有: DataNode、RegionServer、NodeManager、Solr 等。

- 在现有集群规模上进行服务扩容,即在当前集群中安装更多的业务组件实例,提升服务能力。
- 节点扩容,然后在新添节点上添加业务组件实例。

综上,从操作上可将扩容场景划分成两类:

- 扩容节点,即往集群中添加更多服务器。
- 添加服务实例,无论是全新添加服务还是为业务组件添加业务实例,在 FusionInsight Manager 上的操作均为添加服务实例。

# 1.2 何时进行服务扩容

随着业务量的增长,集群存储容量、计算能力或服务能力无法满足业务需求时,需要考虑对服务进行扩容。

## 1.2.1 何时扩容 HDFS DataNode

满足以下条件时,建议进行 DataNode 扩容:

- FusionInsight Manager 上已经正确设置了"DataNode 磁盘空间使用率超过阈值" 监控转告警(AlarmID: 14002)。具体设置方法请参考产品文档。
- 系统收到此告警且该告警的原因已经排除硬盘硬件故障。

#### 扩容详请如下:

● 扩容量

满足至少一年的数据存储需求。例如,未来业务模型不会增加或变更,近 6 个月的数据增长为 m/月,则需要扩容存储量为 12m,单个 DataNode 实例下所挂载磁盘容量为 n,则供需要扩容 DataNode 实例=12m/n。具体扩容量以客户需求规划为主。

● 扩容方式

节点扩容后在新增节点上添加 DataNode 实例。

- 关联扩容
  - 对于分析型集群,建议扩容 DataNode 的节点上同时扩容 NodeManager。
  - 对于实时数据存取型集群,建议扩容 DataNode 的节点上同时扩容 RegionServer。

# 1.2.2 何时扩容 Yarn NodeManager

满足以下条件时,建议进行 NodeManager 扩容:

- FusionInsight Manager 上已经正确设置了 MapReduce Job Execution Timeout (AlarmID: 18006) 监控转告警。具体设置方法请参考产品文档。
- 系统收到此告警且排除了 MapReduce 服务不可用和部分 NodeManager 异常的场景。

#### 扩容方式:

- 如果集群内仍有部分安装了 DataNode 实例的节点未安装 NodeManager,且在集群 角色部署规划中允许这些节点安装 NodeManager 实例,可直接在这些节点上添加 NodeManager 实例完成扩容。
- 节点扩容后在新增节点上添加 DataNode 实例与 NodeManager 实例。

# 1.2.3 何时扩容 HBase RegionServer

根据业务需求,理论分析当前 RegionServer 的读性能是否满足业务需求。假定网络状况足够好,最坏的情况下,对 HBase 的请求均为采用简单的 Get 接口,且每次 Get 请求所需数据均需从硬盘中读取,业务对集群的 IOPS 为 R,RegionServer 所在节点的单个硬盘的 IOPS 为 M,单节点挂载业务硬盘为 N 块,则至少需要 R/(M\*N)个

RegionServer 实例,如果当前实例个数无法满足此需要,建议进行 RegionServer 的扩容。

对 RegionServer 的运行状态进行分析,通过 FusionInsight Manager 获取 HBase 服务的 "RegionServer Queue Statistics"报表,如果在最近 1-3 个月内发现"Total Compaction Queue Size"和"Total Flush Queue Szie"在持续增长,且无明显的周期性,则说明 HBase 服务在写数据方面遇到性能瓶颈,需要进行扩容。

#### 扩容方式:

- 如果集群内仍有部分安装了 DataNode 实例的节点未安装 RegionServer,且在集群 角色部署规划中允许这些节点安装 RegionServer 实例,可直接在这些节点上添加 RegionServer 实例完成扩容。
- 节点扩容后在新增节点上添加 DataNode 实例与 RegionServer 实例。

## 1.2.4 何时扩容 Solr SolrServer

Solr 扩容的场景主要有两种:

- 物理资源消耗过大,即 Solr 的服务节点的 CPU、内存占用率过高、磁盘空间不足场景。具体操作步骤为:进入 FusionInsight Manager 后,选择"服务管理 > Solr > 实例",单击所要查看的实例,即可看到当前一段时间内,CPU、内存使用情况;同时,进入服务端后台查看磁盘使用情况。
- 服务扩容,即 Solr 单实例的索引数据太大,索引的数目达到 10 亿条或者是数据大小达到 1T。

#### 扩容方式:

- 增加 SolrServer 实例,可以直接在 FunsionInsight Manager 上添加 Solr 实例。
- 增加节点,当 Solr 出现物理资源消耗过大,或者是无法在已经有的节点上添加 Solr 实例,则需要在新的节点上增加实例。

## □ 说明

集群中安装 Solr 服务时,必须也只能有两个 Solr Server Admin 实例,且一个节点只能包含一个 Solr Server Admin 实例(这里假设在 A、B 节点上各有一个 Solr Server Admin 实例)。如果后续在 节点 A 上删除了 Solr Server Admin 实例,要添加 Solr Server Admin 实例,有两种方式:

- 1. 由于 HA 本身的限制,形成 HA 的两个 SolrServerAdmin 一旦选定主机节点后,就不能再单独更改一台主机,因为 HA 已经记住这两主机的 IP 信息。如果不对 B 节点上 SolrServerAdmin 实例做改动,则新添加的 SolrServerAdmin 实例只能在原来删除 SolrServerAdmin 实例的 A 节点上创建。
- 2. 删除 B 节点上的 Solr Server Admin,可以在 A 和 B 节点或者其他节点重新同时添加两个 Solr Server Admin 实例。

# 1.2.5 何时扩容 Kafka Broker

假定在 FusionInsight Manager 上已经正确设置了"Kafka 磁盘容量不足"监控转告警 (AlarmID: 38001),当收到此告警且该告警的原因已经排除硬盘硬件故障时,建议扩容 Broker。

扩容详请如下:

● 扩容量

满足用户所设置的最大数据保存周期的的存储需求。

例如,所设置的数据保存周期为7天,数据的增长为m/天,则需要扩容存储量为7m,单个Broker 实例下所挂载磁盘容量为n,则供需要扩容Broker 实例=7m/n。 具体扩容量以客户需求规划为主。

- 扩容方式
  - a. 节点扩容后在新增节点上添加 Broker 实例。
  - b. 增加已有 Broker 实例的磁盘空间。

# 1.2.6 何时扩容 Streaming Supervisor

假定在 FusionInsight Manager 上已经正确设置了"Streaming Slot 使用率超过阈值"监控转告警(AlarmID:26003)监控转告警。且排除了 Streaming 服务不可用和部分 Supervisor 异常的场景。

扩容方式:

• 节点扩容后在新增节点上添加 Supervisor 实例。

## 1.2.7 何时扩容 Redis

Redis 扩容的场景主要有两种:

- FusionInsight Manager 上已经正确设置了"CPU 使用率超过阈值"监控转告警(告警 ID: 12016),当系统收到此告警,并且告警详细信息中指明出现 CPU 过载的是Redis 节点。
- FusionInsight Manager 上已经正确设置了"磁盘容量不足"监控转告警(告警ID: 12017),当系统收到此告警,并且告警详细信息中指明出现磁盘容量不足的主机名和磁盘分区是 Redis 的磁盘分区。

#### 扩容方式:

 如果集群内仍有部分节点安装了 Redis 实例,但这些实例并未安装加入 Redis 业务 集群中,那么可以在"服务管理 > Redis > Redis 管理"页面找到需要扩容的 Redis 集群,并点击扩容按钮。

然后在集群扩容操作页面上,选择将这些未使用的 Redis 实例加入到该 Redis 集群中,即可完成扩容。

● 如果所有的 Redis 实例都已经使用了,那么就要进行节点扩容,并在新增节点上添加 Redis 服务和实例。

然后再进入 Redis 集群扩容页面将这些新增加的 Redis 实例,加入到 Redis 集群中,从而完成扩容。

# 1.3 扩容前准备

# 1.3.1 扩容规划

- 扩容分析,确定扩容场景:
  - 安装新组件
  - 现有集群中添加服务实例

- 添加节点后添加服务实例
- 根据扩容需求文档及《FusionInsight 配置规划工具》完成服务实例安装规划,该 文档汇总了集群中哪个节点上安装哪个服务实例,以及实例的相关参数配置等安 装信息。

## 1.3.2 环境检查

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"服务管理 > 更多操作 > 启动集群健康检查",执行集群的健康检查,确保所有已安装服务和节点的"健康状态"为"良好"。

----结束

## 1.3.3 辅助参考文档

- 《FusionInsight 配置规划工具》
- 《产品文档》中"软件安装 > 安装过程 > 配置并检查安装环境"
- 硬件自带的《硬件安装》

# 1.4 扩容约束

● 扩容的节点建议不要选择 ResourceManager、JobHistory 或者 NameNode 所在的节点。

在 ResourceManager 或者 JobHistory 上增加或删除 NodeManager 或 DataNode 会导致 ResourceManager 或者 JobHistory 出现"配置状态"处于"过期",此时需要重启服务才能恢复。

● 扩容的节点建议不要选择 HMaster 所在的节点。

在 HBase 上增加或者删除 HMaster 节点会出现"配置状态"处于"过期",此时需要重启服务才能恢复。

# 1.5 扩容影响

无。

# 1.6 扩容期间禁止的操作

为了避免数据丢失, Solr 组件扩容期间, 不要进行创建索引操作。

# **2** 扩容操作指导

- 2.1 集群规模与组网规划
- 2.2 添加新节点至集群
- 2.3 添加新节点主机名至客户端节点
- 2.4 添加组件实例
- 2.5 升级服务器
- 2.6 DataNode 扩容后 Balance
- 2.7 DataNode 扩容后 Colocation
- 2.8 Solr 扩容
- 2.9 Kafka 扩容
- 2.10 为 Redis 业务集群添加 Redis 实例

# 2.1 集群规模与组网规划

# 2.1.1 简介

目前的组网按规模分为三种基本布局,分别是节点数小于 6 的集群,200 节点内规模部署,200 节点以上集群大规模部署。针对这三种不同的规模部署,我们给出了典型的组网规划。详情请参见《硬件部署和组网方案指南》。

如果当前的组网不足以支撑业务需求就需要进行扩容,扩容包括两个方面,一个是节点规模的扩张,另一个是添加部署新的组件服务或者实例。

# 2.1.2 节点规模扩张

节点规模的扩张可先参见 2.2.1 使用模板添加新节点至集群完成物理节点的部署,然后可以根据《硬件部署和组网方案指南》和《集群组件迁移指导》进行节点实例的规划部署。

## 2.1.3 添加组件服务

添加组件服务操作可结合《硬件部署和组网方案指南》进行实例部署规划,并参见 2.4 添加组件实例完成。

# 2.2 添加新节点至集群

添加新节点至集群时,可以使用配置规划工具创建的扩容模板添加,也可以手动添加。对于节点较多的集群,为避免扩容出错,建议使用模板添加新节点。

## 2.2.1 使用模板添加新节点至集群

## 2.2.1.1 使用配置规划工具(LLD)创建扩容模板

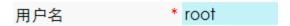
步骤 1 打开 LLD 工具, 启用宏:



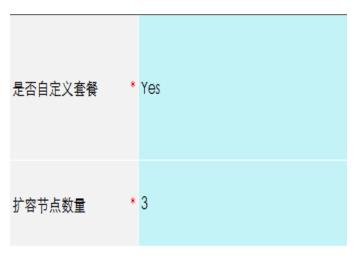
步骤 2 选择产品类型 "HD", 安装模式选择为"集群扩容", 认证模式和原集群保持一致:

配置	内容
产品类型 *	HD
集群名称 *	Test_01
版本号 *	FusionInsight V100R002C60U10
安装模式 *	集群扩容
认证模式 *	安全模式

步骤 3 选择用户名,如果集群安装时使用的是 **root** 用户,这里就要选择为"root",如果集群安装时使用的是非 **root** 用户,这里就要选择为"omm"。



步骤 4 输入扩容节点数量,如果需要手动修改实例部署,需要将"是否自定义套餐"修改为"Yes":



- 步骤 5 进入"选择服务"界面,选择服务,需要和原集群中安装的服务保持一致。
- 步骤 6 进入"IP 规划与进程部署"界面,配置"机架名称"、"管理 IP"、"业务 IP",如果扩容节点的默认部署不满足规划,可以修改扩容节点的实例部署。

类型	机架名称	管理IP	业务IP
DN	/default/rack0	192.168.2.101	10.168.2.101
DN	/default/rack0	192.168.2.102	10.168.2.102
DN	/default/rack0	192.168.2.103	10.168.2.103

- 步骤 7 进入"节点信息"界面,填写"CPU虚拟核数"、"内存"、"主机逻辑磁盘数量"、"最小数据盘容量"、"主机名"。
- 步骤 8 跳过"浮动 IP"界面,扩容场景下,该界面不需配置。
- 步骤 9 进入"磁盘配置"界面,对扩容节点的元数据盘数、数据盘数进行配置。
- 步骤 10 跳过"集群参数配置"界面,扩容场景下,该界面不需配置。

步骤 11 进入"实例参数配置"界面,该界面配置使用默认配置即可,如有特殊需求,请在 SA 指导下修改:

该界面默认无需配置,如需配置请在项目		IDFS	
SA/SE指导下进行。	DataNode		NameNode
管理IP	DotaNode的数据存放目录,自动生成。	DataNode运行的JVM GC参数配置。	NameNode运行的JVM GC参数配置。
	dfs.datanode.data.dir	GC_OPTS	GC_OPTS
192.168.2.101	/srv/BigData/hadoop/data1~15/dn	-Xms2G -Xmx4G	
192.168.2.102	/srv/BigData/hadoop/data1~15/dn	-Xms2G -Xmx4G	
192.168.2.103	/srv/BigData/hadoop/data1~15/dn	-Xms2G -Xmx4G	

步骤 12 进入"生成配置文件"界面,点击"生成配置文件"按钮,生成如下文件:扩容模板(expandTemplet.xml)、preinstall 配置文件、precheck 配置文件。

----结束

## 2.2.1.2 上传配置文件到主 OMS 节点并进行预处理

- 步骤 1 使用 "PuTTY"工具,以 root 用户登录主管理节点。
- 步骤 2 执行以下命令, 防止 "PuTTY" 超时退出。

#### TMOUT=0

- 步骤 3 检查 OMS 上的安装包解压目录是否存在,如果该目录已经被删除,则需要重新下载 FusionInsight HD 软件包,并上传到主 OMS 节点,并进行校验和解压操作。可参考 "《软件安装》的"检验软件包"章节。
- 步骤 4 检查扩容节点 OS 的编码格式是否符合要求。

使用"PuTTY"工具,以 root 用户登录任意节点。执行 locale 命令,查看 OS 的编码格式是否为"en\_US.UTF-8"或"POSIX"。

- 是, 执行步骤 6。
- 否, 执行步骤 5。
- 步骤 5 使用"PuTTY"工具,以 **root** 用户登录所有要扩容的节点,更改 OS 编码格式。以更 改为"en\_US.UTF-8"格式为例进行说明。

Red Hat 系统下执行 vi /etc/sysconfig/i18n 命令,按"Insert"键开始编辑。将"LANG"的值修改为"en\_US.UTF-8","SYSFONT"的值修改为"latarcyrhebsun16",按"Esc"键退出编辑,并输入:wq 保存退出,重启操作系统后配置生效。

SUSE 系统下执行 vi /etc/sysconfig/language,将 "RC\_LANG"的值修改为 "en US.UTF-8",并保存退出,重启操作系统后配置生效。

步骤 6 是否使用 omm 用户进行扩容。

- 是,执行步骤 7。
- 否,执行步骤 9。

步骤7 使用"PuTTY"工具,以 root 用户登录主管理节点。

步骤 8 将主管理节点上的 "/opt/FusionInsight/software/preset" 文件夹拷贝至所有要扩容节点的 "/opt" 目录下并在这些节点上执行 **sh/opt/preset/preset.sh** 命令。

在集群内多个节点上执行同一命令或者存取文件,可以使用软件安装包中的脚本工具快速完成,详细操作请参考"管理员指南 > 附录 > FAQ > Manager > 如何在集群内多个节点上执行命令或者存取文件"章节。

步骤 9 使用"WinSCP"工具,把生成的配置文件目录中的整个 software 目录覆盖主 OMS 节 点解压后软件中的相应目录"/opt/FusionInsight/software"。

----结束

## 2.2.1.3 执行 Preinstall

- 步骤 1 使用"WinSCP"工具,以 root 用户登录主管理节点。将操作系统镜像文件上传至主管理节点目录下,例如"/opt"。
- 步骤 2 使用 PuTTY 工具,以 root 用户登录主管理节点,执行以下命令,防止"PuTTY"超时退出。

TMOUT=0

步骤 3 执行以下命令进行挂载:

mount 虚拟光盘文件名 /media/ -o loop

例如 Red Hat 系统,假设虚拟光盘文件在"/opt"下,执行: mount /opt/rhel-server-6.4-x86\_64-dvd.iso /media/ -o loop

步骤 4 进入解压目录,例如 "/opt/FusionInsight/software/preinstall" ,检查配置规划工具生成的 preinstall.ini 是否已上传到此目录。请确保"g\_parted\_conf"节点中要格式化的分区中没有数据。

cd /opt/FusionInsight/software/preinstall

cat preinstall.ini

```
prinstall.ini (举例)
g_hosts="192.168.10.[10-20]"
g_user_name="root"
g_port=22
g parted=2
g_parted_conf="192.168.10.10:host0.ini;192.168.10.11:host1.ini;
192.168.10.12:host2.ini;192.168.10.13:host3.ini;192.168.10.14:host4.ini;
192.168.10.15:host5.ini;192.168.10.16:host6.ini;192.168.10.17:host7.ini;
192.168.10.18:host8.ini;192.168.10.19:host9.ini;192.168.10.20:host10.ini;"
g_add_pkg=1
g pkgs dir="suse-11.1:/media/"
g_log_file="/tmp/fi-preinstall.log"
g debug=0
g_hostname_conf="192.168.10.10:192.168.20.10:host0;192.168.10.11:192.
168.20.11:host1;192.168.10.12:192.168.20.12:host2;192.168.10.13:192.16
8.20.13:host3;192.168.10.14:192.168.20.14:host4;192.168.10.15:192.168.2
0.15:host5;192.168.10.16;192.168.20.16:host6;192.168.10.17;192.168.20.1
7:host7;192.168.10.18;192.168.20.18;host8;192.168.10.19;192.168.20.19;h
ost9;192.168.10.20:192.168.20.20:host10;"
g_swap_off=1
g feed random setup=2
```

步骤 5 执行安装前配置命令。输入 root 用户的密码,等待执行完毕。

#### cd /opt/FusionInsight/software/preinstall

#### ./preinstall.sh

## 步骤 6 "preinstall"过程结束后,默认会自动继续进行"precheck"过程。如下所示:

```
[INFO] start checking each hosts.
[INFO] localhost: start parsing the configuration file.
[INFO] localhost: parse the configuration file success.
...
```

## 步骤 7 如下例所示,若执行结果没有"ERROR"提示,则检查通过。

```
=====Summary Output======
Environment check success.
```

#### ----结束

## 2.2.1.4 导入扩容模板文件

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"主机管理"。单击"添加主机"。



- 步骤 3 选择"模板添加主机",选择生成的扩容模板,并点击"确定"。
- 步骤 4 填写各节点操作系统 root 用户(或 omm 用户)的密码,点击查找主机:



- 步骤 5 导入模板后会自动跳转至"确认"页面,此时若发现配置规划数据有误,可单击"上一步"回到各配置项检查或更改参数值。
- 步骤 6 点击"提交",进入扩容进度界面,等待扩容完成。

#### ----结束

## 2.2.1.5 扩容验证

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"主机管理",查看新增节点的"健康状态"是否为"良好"。
- **步骤** 3 单击"服务管理",查看"HDFS"、"HBase"、"MapReduce"等已安装服务的状态。

请确保已安装服务"健康状态"为"良好"。

请确保已安装服务"配置状态"为"已同步"。

- 步骤 4 单击某一服务,例如"HDFS"。假设新增 DataNode,单击"DataNode",查看新增 DataNode 的"健康状态"为"良好"。
- 步骤 5 单击"告警管理",查看是否存在新的"HDFS"、"HBase"、"MapReduce"等组件的相关告警。

若有,请参照《故障管理》中"告警处理"相关章节的步骤清除告警。

----结束

## 2.2.2 手动添加新节点至集群

## 2.2.2.1 环境准备

- 步骤 1 新节点服务器已上电,并已接入集群所在网络。
- **步骤 2** 新节点满足《FusionInsight HD V100R002C60U10 产品文档》中"软件安装 > 安装过程 > 配置并检查安装环境"要求
  - □ 说明

在准备新节点执行 preinstall.sh 的 preinstall.ini 配置文件中, g\_hosts 和 g\_parted\_conf 配置项只能填写要添加节点的 ip 和配置文件,不能包含集群已经存在的节点,防止破坏已存在的集群节点。

步骤 3 检查扩容节点 OS 的编码格式是否符合要求。

使用"PuTTY"工具,以 root 用户登录任意节点。执行 locale 命令,查看 OS 的编码格式是否为"en\_US.UTF-8"或"POSIX"。

- 是,执行步骤 5。
- 否,执行步骤 4。
- 步骤 4 使用"PuTTY"工具,以 root 用户登录所有要扩容的节点,更改 OS 编码格式。以更改为"en\_US.UTF-8"格式为例进行说明。

Red Hat 系统下执行 vi /etc/sysconfig/i18n 命令,按"Insert"键开始编辑。将"LANG"的值修改为"en\_US.UTF-8","SYSFONT"的值修改为"latarcyrhebsun16",按"Esc"键退出编辑,并输入:wq 保存退出,重启操作系统后配置生效。

SUSE 系统下执行 vi /etc/sysconfig/language,将 "RC\_LANG"的值修改为 "en\_US.UTF-8",并保存退出,重启操作系统后配置生效。

步骤 5 是否使用 omm 用户进行扩容。

- 是,执行步骤 6。
- 否,结束操作。
- 步骤 6 使用"PuTTY"工具,以 root 用户登录主管理节点。
- 步骤 7 将主管理节点上的 "/opt/FusionInsight/software/preset" 文件夹拷贝至所有要扩容节点的 "/opt" 目录下并在这些节点上执行 **sh /opt/preset/preset.sh** 命令。

在集群内多个节点上执行同一命令或者存取文件,可以使用软件安装包中的脚本工具快速完成,详细操作请参考"管理员指南 > 附录 > FAQ > Manager > 如何在集群内多个节点上执行命令或者存取文件"章节。

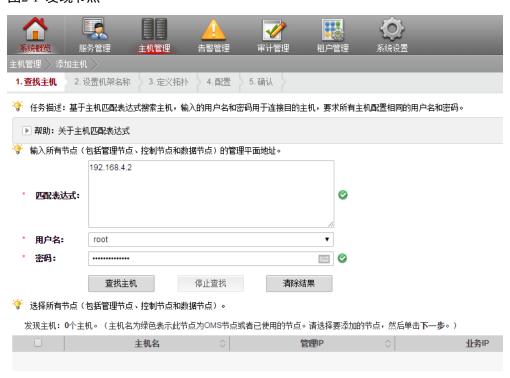
#### ----结束

## 2.2.2.2 操作步骤

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"主机管理"。单击"添加主机 > 手动添加主机"。
- 步骤3 发现节点。

输入节点的 IP 地址和用户名、密码,如果集群安装时使用的是非 root 安装方式,则用户名选择 omm,如果集群安装时使用的是 root 安装方式,则用户名选择 root,然后单击"查找主机",勾选已发现的节点,如图 2-1 所示。单击"下一步"。

#### 图2-1 发现节点



## □ 说明

如果集群中以 HDFS 为存储系统,则检查下当前集群中各个机架上的 Datanode 数据节点个数,尽量保证各个机架上的 DataNode 节点个数大致相等。

## 步骤 4 设置机架号。

勾选节点,单击"下一步"。

步骤 5 根据扩容规划,将需要扩容的服务(仅限已经安装的服务,并且没有超过要求的最大实例数目)映射到节点上,如图 2-2 所示。单击"下一步"。

#### 图2-2 定义拓扑

	Metadata		HBase	)	Hi	ve	FI	ume	FTP-Server
hostName	MS	НМ	□ TS	<b></b> RS	MS	HS	FM	<b>₹</b> FS	FTPS
	(1)	(2)	(3)	(4)	(2)	(2)	(2)	(4)	(3)
192-168-4-2				•				•	
192-168-4-3		<b>₩</b>	4	•	<b>4</b>	4	4	4	<b>₽</b>
192-168-4-4	✓	4	4	4	4	4	4	4	€
192-168-4-5			<b></b>	•				4	€

步骤 6 根据扩容规划,配置需要扩容服务的参数。单击"下一步"。

步骤7 对照扩容规划内容,确认信息无误,单击"提交"。

**步骤** 8 添加新节点(包含实例)后,添加完成后部分服务的配置状态可能由于新加节点配置 而出现配置状态处于过期状态。此时需要重启相关服务,进行配置同步生效。

----结束

# 2.3 添加新节点主机名至客户端节点

当客户端应用所在的主机不是集群中的节点时,需要在客户端所在节点的本地 hosts 文件中设置新节点的主机名和 IP 地址映射,主机名和 IP 地址请保持一一对应。

本地 hosts 文件存放路径举例:

- Windows: "C:\WINDOWS\system32\drivers\etc\hosts"
- Linux: "/etc/hosts"

例如,集群新增了 3 个节点 192.168.2.1、192.168.2.2、192.168.2.3,对应的主机名是 192-168-2-1、192-168-2-2、192-168-2-3,则需要在 hosts 文件中追加以下内容:

192.168.2.1 192-168-2-1 192.168.2.2 192-168-2-2 192.168.2.3 192-168-2-3

# 2.4 添加组件实例

# 2.4.1 添加服务

## 操作场景

根据业务需要添加系统中尚未安装的服务。每次只能添加一个服务。

## 操作步骤

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"服务管理"。单击"添加服务"。
- 步骤3 将服务添加到已有节点。

选择"仅添加服务",如图 2-3 所示。单击"下一步"。

## 图2-3 添加服务



- 步骤 4 根据扩容规划,选择需要添加的服务。单击"下一步"。
- 步骤 5 根据扩容规划,勾选节点。单击"下一步"。
- 步骤 6 根据扩容规划,配置需要扩容服务的参数。单击"下一步"。
- 步骤7 对照扩容规划内容,确认信息无误,单击"提交"。

### ----结束

## 2.4.2 添加 NameService

### 操作场景

为 HDFS 添加新的 NameService。单个集群最多可添加 10 个 NameService。

### 操作步骤

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"服务管理",单击"HDFS"服务。
- 步骤 3 单击"管理 NameService"。单击"添加 NameService"。
- 步骤 4 在 "NameService"输入框内输入新增 NameService 的名称。

- 步骤 5 根据扩容规划,勾选 NameService 需要的实例(NameNode 和 Zkfc)。单击"下一步"。
- 步骤 6 根据扩容规划,配置添加的实例的参数。单击"下一步"。
- 步骤7 对照扩容规划内容,确认信息无误,单击"提交"。
- 步骤 8 添加实例后,部分服务的配置状态可能由于新加实例配置而出现"配置状态"处于"过期"状态。此时需要重启相关配置"过期"状态服务,进行配置同步生效。

#### ----结束

## 2.4.3 添加实例

## 操作场景

为指定的服务添加新的角色实例。

## 操作步骤

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"服务管理"。单击某一服务,例如"HDFS"。
- 步骤 3 单击"实例"。单击"添加实例"。
- 步骤 4 根据扩容规划,勾选需要增加的实例。单击"下一步"。
- 步骤 5 根据扩容规划,配置需要扩容实例的参数。单击"下一步"。

## □ 说明

如果扩容的实例有数据目录的参数,该类参数必须要以真实的数据目录进行配置。否则,如果在实例添加后修改其数据目录参数,将可能造成该实例的数据丢失。主要有以下实例及其数据目录参数如下:

- DataNode 的"dfs.datanode.data.dir"参数。
- JournalNode 的"dfs.journalnode.edits.dir"参数。
- NameNode 的"dfs.namenode.name.dir"和"dfs.namenode.edits.dir"参数。
- NodeManager 的 "yarn.nodemanager.local-dirs" 和 "yarn.nodemanager.log-dirs" 参数。
- QuorumPeer 的"dataDir"参数。
- Broker 的"log.dirs"参数。
- 步骤 6 对照扩容规划内容,确认信息无误,单击"提交"。
- 步骤 7 添加实例后,部分服务的配置状态可能由于新加实例配置而出现"服务状态"处于"过期"状态。此时需要重启相关配置"过期"状态服务,进行配置同步生效。

#### ----结束

# 2.5 升级服务器

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"主机管理"。单击将要升级服务器对应的节点,例如"Hostname1"。

步骤3 单击"停止所有角色"。

步骤 4 关闭该节点对应的服务器。

步骤 5 升级硬件。

#### □ 说服

升级后的硬件不影响该节点的 IP 地址、主机名称、用户名和密码等, 否则请参照《软件安装》设置该节点参数。

步骤 6 接通电源并开启服务器。

步骤 7 在 FusionInsight Manager 页面,单击"主机管理",单击"Hostname1",单击"启动所有角色"。

等待一段时间,检查该节点上所有实例的"健康状态"为"良好"。

----结束

# 2.6 DataNode 扩容后 Balance

#### 建议:

- 业务相对空闲时执行,因为 Balance 操作会占用一定的带宽资源。
- 建议采用默认带宽控制,如果更改该值需要重启 HDFS 服务,会对正在运行的业务产生一定的影响。

例如客户端安装到指定目录"/opt/hadoopclient"。

步骤1 调整带宽控制(可选)。

- 1. 登录 FusionInsight Manager。
- 2. 选择"服务管理>HDFS>服务配置"。
- 3. 将"参数类别"选择为"全部配置",搜索"bandwidthPerSec",修改"dfs.datanode.balance.bandwidthPerSec"参数值。



默认为 20M/s,如现网集群中的网络为万兆,且带业务执行,该值比较合适;如有足够的业务空闲时间窗来专门做 Balance,可适当放大该值以缩短 Balance 时间。

步骤 2 使用 hdfs 用户执行 kinit。

首先到客户端的安装目录下,执行 source bigdata env。

source 成功后,执行kinit hdfs 进行初始化。

根据提示输入正确的用户密码。

#### 步骤 3 执行负载均衡操作命令。

• 全部 Datanode 节点执行负载均衡

#### sh /opt/client/HDFS/hadoop/sbin/start-balancer.sh -threshold 偏差值

例如: sh/opt/client/HDFS/hadoop/sbin/start-balancer.sh-threshold 10

这个命令中-threshold 参数值是 HDFS 达到平衡状态的磁盘使用率偏差值。如果各节点间磁盘使用率与所有节点磁盘使用率的平均值偏差小于 10%,则 HDFS 集群已经达到了平衡的状态。

• 部分 Datanode 节点执行负载均衡

# sh /opt/client/HDFS/hadoop/sbin/start-balancer.sh -include -f 文件路径 -threshold 偏差值

例如: sh /opt/client/HDFS/hadoop/sbin/start-balancer.sh -include -f /tmp/includeHost -threshold 10

includeHost 文件内容:

host1

host2

host3

这个命令中-include -f 参数值是指定了一个文件路径,该文件内容为需要进行负载 均衡的 Datanode 节点,文件中每一行为一个节点主机名,执行负载均衡只在指定 的节点中进行。

#### ----结束

# 2.7 DataNode 扩容后 Colocation

# 2.7.1 应用场景

扩容后加入了新的数据存储节点,但并未关联到 Colocation 的 locators 中。

为了平衡地使用所有的数据节点,使新的数据节点的分配频率与旧的数据节点趋于一致,建议扩容后,创建新的 locators,并把新数据导入到新的 locators 中。或者扩容后,删除旧的 locators,并重新创建 locators,使得数据节点在 locators 中均匀的关联。

## 2.7.2 操作步骤

例如客户端安装到指定目录"/opt/hadoopclient"。

步骤 1 使用 hdfs 用户执行 kinit。

首先到**客户端的安装目录**目录下,执行 source bigdata\_env

然后执行**客户端的安装目录/ HDFS/hadoop/kinit hdfs** 进行初始化

例如: /opt/hadoopclient / HDFS/hadoop/kinit hdfs

根据提示输入正确的用户密码。

步骤 2 执行删除旧的 locators 方法(可选)。

# 客户端的安装目录/ HDFS/hadoop/bin/hdfs colocationadmin -deleteGroup *Colocation 组ID*

例如: /opt/hadoopclient/HDFS/hadoop/bin/ hdfs colocationadmin -deleteGroup 2015

hdfs colocationadmin -deleteGroup 2015: 这个命令中-deleteGroup 表示删除某个组,后面跟组的 ID(如需查询组 ID,可使用命令: hdfs colocationadmin -listGroups),执行以后,就会删除该组下面的 locators 与数据节点的对应关系。

步骤 3 执行创建新的 locators 方法(可选)

客户端的安装目录/ HDFS/hadoop/bin/hdfs colocationadmin -createGroup -groupId Colocation 组 ID -locatorIds 该组下的 locators

例如:/opt/hadoopclient/HDFS/hadoop/bin/hdfs colocationadmin -createGroup - groupId 2015 -locatorIds 01,02,03,04,05,06,07,08,09,10,11,12

hdfs colocationadmin -createGroup -groupId 2015 -locatorIds 01,02,03,04,05,06,07,08,09,10,11,12: 这个命令中-groupId 表示创建一个叫 2015 的组,-locatorIds 指定在该组下创建 01 到 12,共 12 个 locators,后续就可以把数据按照月份存储到各自的 locators 中。

----结束

## 2.8 Solr 扩容

- 1. 在原有的 Collection 上面进行扩容
  - 根据 Collection 创建时选择的路由方式的差别,其扩容的方式也不一样。对采用 implicit 路由方式创建的 collection,需要通过增加 shard 方式进行扩容;对于采用 compsiteID 路由方式(默认)创建的 collection,则需要进行 split 后进行迁移扩容。
- 2. 新建 Collection 到新增 SolrServer 实例上进行业务迁移扩容 创建 Collection 到新增的 SolrServer 实例上,数据写入到该 Collection,建立索引。为避免上层业务变更,可通过 CREAT EALIAS 操作将多个 Collection 通过别名关联到一起,在查询时指定别名,可以查询该别名下所有 Collection。

# 2.8.1 Solr 组件扩容后增加 Shard

□ 说明

业务相对空闲时执行。

如果需要对该 Collection 上添加新的 shard 分片, 其创建 Collection 的路由方式需要为 implicit。 具体看 Solr 的业务操作指南。

在多副本情况下,新建 shard 的副本不支持自动跨主机,存在单点故障风险。可以在新建 shard 时,通过 createNodeSet 指定不同主机实例,使其副本创建到不同的主机上。

在添加组件实例后,如果需要对已经创建的 Collection 添加新的分片,可以通过下面的操作步骤完成。

步骤 1 打开 Chrome 或者其他浏览器,输入下面 URL 进行增加分片。注意,如果系统为安全版,则需要发送 https 消息;如果为非安全版,则为 http 消息。下面给出例子为 https 消息。

https://SolrIp:SolrPort/solr/admin/collections?action=CREATESHARD&shard=shard1&collection=test1&createNodeSet=SolrIp:SolrPort\_solr

#### 参数说明:

- SolrIp: 新增的 SolrServer 的所在节点 IP
- SolrPort: 新增 SolrServer 的对应的 https 业务端口
- collection: 所需要进行添加分片的 Collection 名称
- shard: 需要增加的 shard 的名称,注意这里 shard 不要与已经创建的 shard 同名
- 多副本情况下,createNodeSet 指定多个(副本个数)"SolrIp: SolrPort \_solr",以逗号分隔,可以使其副本创建到不同主机上。

步骤 2 在 Solr Admin 页面的 Cloud->Graph,可以看到该 Collection 已经添加一个 shard。

----结束

# 2.8.2 Solr 组件扩容后进行 split 操作

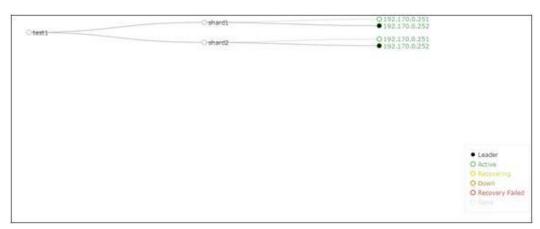
□ 说明

业务相对空闲时执行。

采用 compositeID 路由方式进行创建的 collection,每个 shard 分片的哈希范围在创建 collection 时已分配好,无法通过直接增加 shard 方式进行扩容。一般需要对 shard 进行 split 操作后进行迁移来完成扩容。进行 split 操作相当于在原有的哈希范围上再次进行划分,索引数据需要重新计算哈希值,进行 split 操作在索引数据比较大时非常耗时。

在多副本情况下。新建 shard 的副本不支持自动跨主机,存在单点故障风险。

该任务有一个名称为 test1 的索引集, test1 有两个 shard, 每个 shard 分片都有两个 replica。该 collection 的 shard1 和 shard2 分布在 fi01 和 fi02 的机器上, 扩容添加的节点为 fi03,并在 fi03 上创建了一个 SolrServer1 的实例。现需要将 shard1 进行 split 操作,并进行迁移。

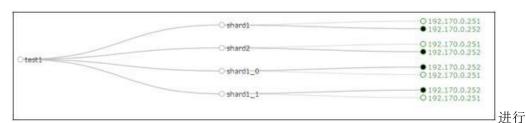


步骤 1 打开 Chrome 或者其他浏览器,输入下面 URL 对 shard 1 进行 split 操作。注意,如果系统为安全版,则需要发送 https 消息;如果为非安全版,则为 http 消息。下面给出例子为 https 消息。

 $https://SolrIp:SolrPort/solr/admin/collections?action=SPLITSHARD\&collection=collection\_n ame\&shard=shardId\\$ 

#### 参数说明:

- SolrIp: 任选一个集群中 solr 实例的 ip 地址
- SolrPort:对应选择的 solr 实例的 https 业务端口
- collection: 所需要进行 split 操作的 collection 名称
- shard: 所需要进行 split 操作的 shard 名称



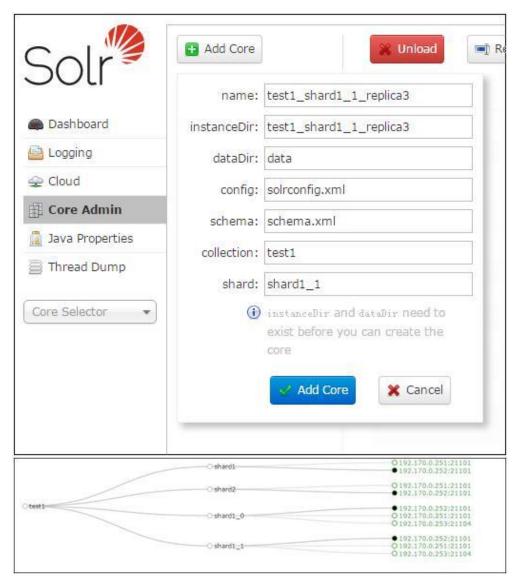
split 操作后,test1 的视图如上所示,shard1 分裂成为 shard1\_0 和 shard1\_1。以后如果有索引数据进入 shard1, 就会往 shard1\_0 和 shard1\_1 进行分发。

- 步骤 2 登录 SolrServer1 实例的实例配置页面。首先在 FusionInsight Manager 的主界面上,单击 "服务管理"中的"Solr"服务,进入后单击"操作状态和健康状态"中的"SolrServer1",进入 SolrServer1 角色的实例界面。点击主机名 fi03 对应的SolrServer1 实例,点击"实例配置",查看其对应端口,得出对应 Solr Admin UI 的URL,并在浏览器中输入 URL,进入 Solr Server1 实例的 Solr Admin UI 界面。
  - 安全模式下,其端口为 SOLR\_SEC\_PORT 配置项值 21104,则其 SolrServer1 Solr Admin UI 的 URL 为
    - https://SolrServer1\_IP:21104/solr
  - 普通模式下,其端口为 SOLR\_PORT 配置项值 21103,则其 SolrServer1 Solr Admin UI 的 URL 为

http://SolrServer1\_IP:21103/solr

步骤 3 在 fi03 的新创建的 SolrServer1 实例上,添加 shard1\_0 和 shard1\_1 的 replica。



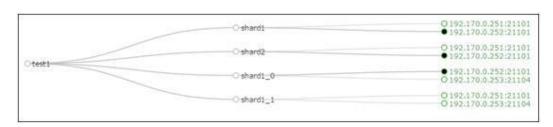


步骤 4 在分片 shard1\_0 和 shard1\_1 上,把处于已有机器上的副本删掉一个来减少索引数据的资源占用。可以对应实例界面 Core Admin,选择相应的 shard,如这里选择 shard1\_0 上的 192.170.0.251:21101 和 shard1\_1 分片上 192.170.0.252:21101,点击 unload 进行删除操作。也可以通过 rest 接口或 DELETEREPLICA 进行操作。以下给出例子为 https 消息。

https://SolrIp:SolrPort/solr/admin/collections? action = DELETEREPLICA & collection = test 1 & shard = shard 1 & replica = replica name

#### 参数说明:

- SolrIp: 需要删除 replica 所在的 solr 实例的 ip 地址
- SolrPort:对应选择的 solr 实例的 https 业务端口
- shard: 所需要进行删除操作的 shard 名称
- replica: 将要删除的 replica 名称



步骤 5 登录 FunsionInsight Manager 管理页面,选择"服务管理 > Solr",确认 Solr 的实例全部正常工作。点击"Solr WebUI"的"SolrServerAdmin"(两个任选一个),进入Solr Admin 的页面。shard1 分片后,通过 Solr Admin 界面进入-> Cloud -> Tree -> /collections/test1/state.json 查看 replica 名为 shard1\_0 和 shard1\_1 两个分片会达到 active 状态,而 shard1 会变成 inactive 状态,可以通过 rest 接口或用 DELETESHARD 操作删除 shard1。下面给出例子为 https 消息。

https://SolrIp:SolrPort/solr/admin/collections?action=DELETESHARD&collection=test1&shard=shard1

#### 参数说明:

- SolrIp: 需要删除 shard1 所在的 solr 实例的 ip 地址
- SolrPort:对应选择的 solr 实例的 https 业务端口
- collection: 所需要进行删除操作的 collection 名称
- shard: 所需要进行删除操作的 shard 名称

#### ----结束

## 2.8.3 Solr 组件扩容后新增 Collection

## □ 说明

现有机器不能够满足现有业务需求,可以通过扩充服务器新增 SolrServer 实例来解决。

创建 Collection 到新增的 Solr Server 实例上,数据写入到该 Collection,建立索引,用别名关联多个 Collection。

关于添加主机和增加角色实例的操作,具体请参见 FusionInsight HD\_CPI 文档->管理员指南->对象管理->主机->添加主机和 FusionInsight HD\_CPI 文档->管理员指南->对象管理->角色实例->添加角色实例。

步骤 1 将新 Collection 创建到新增加的 SolrServer 实例上。在创建 Collection 时通过 creteNodeSet 为新 Collection 指定节点。下面给出例子为 https 消息。

 $https://SolrIp:SolrPort/solr/admin/collections?action=CREATE\&name=newCollection\&create NodeSet=newIp:newPort\_solr,newIp1:newPort1\_solr,newIp2:newPort2\_solr$ 

- SolrIp: 已有 SolrServer 实例所在的 ip 地址
- SolrPort: 对应选择的 solr 实例的 https 业务端口
- newIp,newIp1,newIp2: 新增 SolrServer 实例对应主机的 ip 地址
- newPort.newPort1.newPort2: 新增 SolrServer 实例对应主机的业务端口
- 步骤 2 用 CREATEALIAS 命令,通过 rest 接口进行操作,将多个 Collection 通过别名关联到一起。例如创建 testalias 别名将名称为 oldCollection 和 newCollection 新旧两个 Collection 关联起来。在 testalias 上面进行查询,可以达到同时查询 oldCollection 和 newCollection 的效果。下面给出例子为 https 消息。

https://SolrIp:SolrPort/solr/admin/collections?action=CREATEALIAS&name=testalias&collections=oldCollection, newCollection

- oldCollection: 旧机器上已有的 Collection
- newCollection: 新增 SolrServer 实例上创建的 Collection

#### ----结束

# 2.9 Kafka 扩容

## 2.9.1 Kafka 磁盘扩容

Kafka 采用增加磁盘的方式扩容后需要执行以下操作:

- 步骤 1 在 Manager 界面选择"服务管理 > Kafka > 服务配置", "参数类别"选择"全部配置", 在搜索框中搜索配置项"log.dirs", 将该配置项的值修改为新规划的目录。
- **步骤 2** 在 Manager 界面选择"服务管理 > Kafka",单击"更多操作 > 重启服务",重启 Kafka 服务。

----结束

# 2.9.2 Kafka 扩容后 Reassign

Kafka 扩容后,已有的 Topic Partitions 数据不会自动迁移到新的 Broker 实例或磁盘目录下,除非用户创建新的 Topic 或者是用户主动触发 Reassign。为了保证数据分布均衡,在新增 Broker Server 后,一般都会执行 Reassign 将已有数据迁移到新节点上。

#### 建议:

- 业务相对空闲时执行,因为 Reassign 操作会占用一定的带宽资源。
- 执行 Reassign 操作会使得数据进行重分布,执行过程中会影响当前业务。
- 数据迁移一旦开始无法停止,不要强行停止集群,造成数据不一致。
- 数据迁移是先拷贝,后删除,请确保足够空间。

例如客户端安装到指定目录"/opt/streamclient"。

- 步骤 1 在客户端的安装目录下,执行 source bigdata\_env 进行初始化。
- 步骤 2 (普通模式请跳过该步骤) 执行 kinit admin,使用 admin 用户进行安全登录,根据提示输入密码。
- 步骤 3 执行 kafka-broker-info.sh --zookeeper {zk\_host}:{port}/kafka,, 获取每个节点的 Broker\_ID。

Broker ID	IP Address
1	192.168.0.10
2	192.168.0.11
3	192.168.0.12

步骤 4 执行 kafka-topics.sh --list --zookeeper {zk\_host}:{port}/kafka, 查询当前 Topic 信息。

步骤 5 创建 Json 文件,内容为需要移动的 topic 信息,文件保存为 topics-to-move.json。

例如:

```
{ "topics": [{"topic":"topic1"}],
"version":1
```

步骤 6 指定移动 Topic 和 Broker 列表, 生成重分布信息。

#### kafka-reassign-partitions.sh

- --zookeeper {zk\_host}:{port}/kafka
- --topics-to-move-json-file {topics to reassign json file path}
- --broker-list {brokerlist}
- --generate

例如:

kafka-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a
- --topics-to-move-json-file topics-to-move.json
- --broker-list "23,24,25"
- --generate

命令会打印 Topic Partitions 原始分布情况,以及 Kafka 生成的重分布后的 Partitions 分布情况。用户可以将原始分布信息保存为文件,以便进行回滚;同时可以对重分布后的 Partitions 分布情况进行手动修改,将新的 Partitions 分布信息保存为 Json 文件。例如:expand-cluster-reassignment.json。

如果有磁盘容量告警,获取该磁盘目录下 Topic Partitions 作为 Reassign Partitions。

步骤 7 根据新的 Partitions 分布情况执行重分布。

#### kafka-reassign-partitions.sh

- --zook eeper {zk\_host}:{port}/kafka
- --reassignment-json-file {manual assignment json file path}
- --execute

例如:

/opt/streamclient/Kafka/kafka/bin/kafka-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/kafka
- --reassignment-json-file expand-cluster-reassignment.json
- --execute

命令会打印 Topic Partitions 原始分布情况和新的分布情况。

步骤 8 确认 Partitions 重分布执行结果。

#### kafka-reassign-partitions.sh

- --zookeeper {zk\_host}:{port}/kafka
- --reassignment-json-file {manual assignment json file path}
- --verify

例如:

/opt/streamclient/Kafka/kafka/bin/kafka-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a
- --reassignment-json-file expand-cluster-reassignment.json
- --verify

命令会打印 Topic Partitions 分配状态信息。例如:

Status of partition reassignment:

Reassignment of partition [topic1,1] completed successfully

Reassignment of partition [topic1,0] completed successfully

Reassignment of partition [topic1,2] completed successfully

----结束

# 2.10 为 Redis 业务集群添加 Redis 实例

- **步骤** 1 单击"服务管理> Redis > Redis 管理"。
- 步骤 2 单击要操作的 Redis 集群名称后的扩容按钮。
- 步骤 3 在"选择实例"页面中选择需要加入 Redis 集群的 Redis 实例, 然后单击"下一步"。



步骤 4 系统会生成新加入的 Redis 实例的主从关系,单击"提交"。

----结束

# **3** 扩容验证

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"服务管理 > 更多操作 > 启动集群健康检查",执行集群的健康检查,确保扩容后所有已安装服务的"服务状态"为"良好"。

----结束

# **4** 减容概述

- 4.1 减容场景
- 4.2 何时进行服务减容
- 4.3 减容前准备
- 4.4 减容约束
- 4.5 减容影响
- 4.6 减容期间禁止的操作

# 4.1 减容场景

当前典型的减容场景包括:

- 节点可能存在无法修复的硬件错误
- 节点需要重新安装操作系统

# 4.2 何时进行服务减容

# 4.2.1 何时减容 HDFS DataNode

DataNode 减容的场景主要有两种:

- 节点需要重新安装操作系统。
- 当 HDFS 的 DataNode 减容后,也满足至少一年的数据存储需求。
   具体减容量以客户需求规划为主。

## 减容方式

- 1. 在 FunsionInsight Manager 页面退服 DataNode 实例。
- 2. 在 FunsionInsight Manager 页面删除 DataNode 实例。

关联减容

建议减容 DataNode 的节点上同时减容 NodeManager 和 RegionServer。

## 4.2.2 何时减容 Yarn NodeManager

NodeManager 减容的场景主要有:

节点需要重新安装操作系统

减容方式:

- 1. 在 FunsionInsight Manager 页面退服 NodeManager 实例。
- 2. 在 FunsionInsight Manager 页面删除 NodeManager 实例。

## 4.2.3 何时减容 HBase RegionServer

RegionServer 减容的场景主要有:

- 节点需要重新安装操作系统。
- 节点较少的情况下将 HMaster 和 RegionServer 部署在同一个节点上,需要分开部署。

减容方式:

• 在 FunsionInsight Manager 上删除 RegionServer 实例。

## 4.2.4 何时减容 Solr SolrServer

Solr 减容的场景主要有两种:

- 节点需要重新安装操作系统
- 当 Solr 的创建的 shard 下的 replica 过多,需要减少 replica 数目。

减容方式:

在 FunsionInsight Manager 上删除 solr 实例。

## 4.2.5 何时减容 Redis

Redis 减容的场景主要有两种:

- 节点需要重新安装操作系统。
- 当 Redis 集群减容后,也能满足客户的数据缓存需求。
   具体减容量以客户需求规划为主。

减容方式:

在 FunsionInsight Manager 的"服务管理 > Redis > Redis 管理"页面找到需要减容的 Redis 集群,并点击减容按钮。

## 4.2.6 何时减容 Kafka Broker

Kafka Broker 减容的场景主要有两种:

• 节点需要重新安装操作系统。

当 Broker 减容后,也能满足客户的数据存储需求。
 具体减容量以客户需求规划为主。

减容方式:

在 FunsionInsight Manager 上删除 Broker 实例。

## 4.2.7 何时减容 Streaming Supervisor

Streaming Supervisor 减容的场景主要有两种:

- 节点需要重新安装操作系统。
- 当 Supervisor 减容后,也能满足客户的业务计算需求。
   具体减容量以客户需求规划为主。

减容方式:

在 FunsionInsight Manager 上删除 Supervisor 实例。

## 4.3 减容前准备

## 4.3.1 减容规划

- 减容只能减少数据节点,不能减少控制节点。
- 当 DataNode 数量少于或等于 HDFS 的副本数时,不能执行退服操作。若 HDFS 副本数为 3 时,则系统中少于 4 个 DataNode,将无法执行退服,Manager 在执行退服操作时会等待 30 分钟后报错并退出执行。

## 4.3.2 环境检查

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"服务管理 > 更多操作 > 启动集群健康检查",执行集群的健康检查,确保所有已安装服务的"服务状态"为"良好"。

----结束

## 4.3.2.1 检查服务状态

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"服务管理",查看"HDFS"、"Solr"的状态。

请确保 "HDFS"、"Solr"的"健康状态"为"良好"。

----结束

### 4.3.2.2 检查告警

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"告警管理",查看是否存在"HDFS"、"Solr"的相关告警。

若有,请参照故障处理步骤清除告警。

----结束

## 4.3.2.3 检查 Kafka Topic 同步状态

步骤 1 通过 PuTTY 使用 root 账户登录 FusionInsight Stream Client。例如客户端安装到指定目录"/opt/streamclient"。

步骤 2 使用 admin 用户执行 kinit。

首先到**客户端的安装目录**目录下,执行 source bigdata\_env 进行初始化。

然后执行**客户端的安装目录/KrbClient/kerberos/bin/kinit admin** 进行安全认证。

例如: /opt/streamclient/KrbClient/kerberos/bin/kinit admin

根据提示输入正确的用户密码。

步骤 3 执行 ZooKeeper 相关命令,获取当前集群信息。

客户端的安装目录/Zookeeper/zookeeper/bin/zkCli.sh - server zk\_host:port/chroot

例如:/opt/streamclient/Zookeeper/zookeeper/bin/zkCli.sh - server 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/kafka

进入 ZooKeeper 客户端命令行界面命令,查询相关信息。

1. 查询 Broker 信息, 25 为减容 Broker。

ls /brokers/ids

[23, 24,25]

2. 查询 Topic 信息

ls /brokers/topics

[topic1]

步骤 4 执行 Kafka 相关命令,获取 Topic 分布信息和副本同步信息,观察返回结果。

#### 客户端的安装目录/Kafka/kafka/bin/kafka-topics.sh

- --describe
- --zookeeper <zk\_host:port/chroot>

例如: /opt/streamclient/Zookeeper/Kafka/kafka/bin/kafka-topics.sh --describe

--zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a

Topic:topic1 PartitionCount:2 ReplicationFactor:2 Configs:

Topic: topic1 Partition: 0 Leader: 23 Replicas: 23,25 Isr: 23,25

Topic: topic1 Partition: 1 Leader: 24 Replicas: 24,23 Isr: 24,23

其中, Replicas 对应副本分布信息, Isr 对应副本同步信息。

- 步骤 5 如果 Topic Partition 对应的 Isr(同步副本)中列表为空,或者仅为待减容 Broker 的 id,则减容可能会导致数据丢失。
- 步骤 6 如果 Topic 的所有 Partition 对应的 Isr (同步副本) 列表,包括多个 Broker 的 id, 可进行下一步操作。

----结束

## 4.4 减容约束

减容的节点建议不要选择 ResourceManager、JobHistory 或者 NameNode 所在的节点。

在 ResourceManager 或者 JobHistory 上增加或删除 NodeManager 或 DataNode 会导致 ResourceManager 或者 JobHistory 出现"配置状态"处于"过期",此时需要重启服务才能恢复。

#### □ 说明

如果配置 HDFS Federation 并减容 NameService, 请检查 YARN 的"dfs.namenode.rpc-address.haclusterX.remotenn1"和"dfs.namenode.rpc-address.haclusterX.remotenn2"是否需要修改,同时检查该 NameService 是否存在对应的 NameNode 备份任务需删除。

- 减容的节点建议不要选择 HMaster 所在的节点。
   在 HBase 上增加或者删除 HMaster 节点会出现"配置状态"处于"过期",此时需要重启服务才能恢复。
- DataNode 减容前需要执行退服操作,并请保证退服后各机架 DataNode 节点数大 致相等。
- 如果 DataNode 减容前多于一个机架,保证减容后也多于一个机架(不然后期做退服一定失败)。

## 4.5 减容影响

- HDFS 存储空间减少。
- MapReduce 任务执行时间可能会更长。
- Yarn 可调度的集群 CPU 和内存资料会减少。
- HBase 可存储的数据量会减少,读写性能会降低
- Kafka 数据的保存周期可能会减少。
- Streaming 部署的业务性能会根据具体业务模型受到不同程度的影响。
- Solr 查询并发降低,本地索引存储空间减少。
- Redis TPS 降低,内存存储空间减少。

## 4.6 减容期间禁止的操作

为了避免数据丢失, Solr 组件减容期间, 不要进行创建索引操作。

# 5 DataNode 减容操作指导

- 5.1 DataNode 减容前退服
- 5.2 获取节点的 Instance
- 5.3 删除 Instance
- 5.4 减容节点
- 5.5 DataNode 减容后 Colocation

## 5.1 DataNode 减容前退服

#### 操作步骤

#### 1世 1世 田

- 业务相对空闲时执行,因为退服操作会占用一定的带宽资源。
- 建议采用默认带宽控制,如果更改该值需要重启 HDFS 服务,会对正在运行的业务产生一定的影响。
- HDFS DataNode 和 Yarn NodeManager 减容前需要执行退服操作。
- 步骤 1 登陆 FusionInsight Manager。
- 步骤 2 进入"服务管理>HDFS",点击"实例"进入 HDFS 实例页面。
- 步骤3 在 HDFS 实例页面,选择,要退服的 DataNode, 然后点击"更多操作"中的"退服"进行退服操作。



## □ 说明

- 如果退服前,DataNode 节点所在超过 1 个机架,请保证 DataNode 退服后,集群中 Datanode 分布的机架数仍然多余 1 个机架,否则会因为副本无法迁移到相同机架导致退服失败。
- 请尽量保证退服后,各个机架的 Datanode 节点数目大致相等。
- 步骤 4 DataNode 退服的时候,由于会将数据移动到其他节点,因此耗时很长。可以先关闭页面。如果 DataNode 退服完成,实例的状态会变为"退服"。



### □ 说明

只能对退服后的 DataNode 实例做删除实例减容的操作。

#### ----结束

## 5.2 获取节点的 Instance

- 步骤 1 登录 FusionInsight Manager。
- 步骤 2 单击"主机管理",再点击对应的需要减容的节点。
- 步骤3 然后记录该节点所安装的实例。
- 步骤 4 查询之后对每一个安装的实例进行删除操作。
  - □ 说明

为确保不丢失数据, 部分组件的实例不能直接删除, 例如 DataNode 需要首先进行退服操作。

#### ----结束

## 5.3 删除 Instance

- 步骤1 单击"服务管理"。
- 步骤 2 单击服务列表中的服务,例如 HDFS。
- 步骤3 单击"实例"页签。
- 步骤 4 勾选角色实例名称前的复选框。
- 步骤 5 选择"更多操作 > 删除实例",执行相应的操作。

#### ----结束

## 5.4 减容节点

步骤1 单击"主机管理"。

步骤 2 勾选需要减容节点前的复选框。

步骤 3 选择"更多操作 > 删除主机"执行相应的操作。

○ 说明 如果"删除主机"失败,请选择"强制删除主机"。

----结束

## 5.5 DataNode 减容后 Colocation

## 5.5.1 应用场景

减容后删除了一些数据存储节点, Colocation 中的 locators 需要重新关联数据节点。

为了平衡地使用所有的数据节点,使减容后各节点的数据分配频率趋于一致,建议减容后,创建新的 locators,并把新数据导入到新的 locators 中。或者减容后,删除旧的 locators,并重新创建 locators,使得数据节点在 locators 中均匀的关联。

## 5.5.2 操作步骤

例如客户端安装到指定目录"/opt/hadoopclient"。

步骤 1 使用 hdfs 用户执行 kinit。

首先到**客户端的安装目录**目录下,执行 source bigdata\_env

然后执行**客户端的安装目录** HDFS/hadoop/kinit hdfs 进行初始化

例如: /opt/hadoopclient / HDFS/hadoop/kinit hdfs

根据提示输入正确的用户密码。

步骤 2 执行删除旧的 locators 方法(可选)。

客户端的安装目录/ HDFS/hadoop/bin/hdfs colocationadmin -deleteGroup *Colocation 组ID* 

例如: /opt/hadoopclient/HDFS/hadoop/bin/ hdfs colocationadmin -deleteGroup 2015

hdfs colocationadmin -deleteGroup 2015: 这个命令中-deleteGroup 表示删除某个组,后面跟组的 ID(如需查询组 ID,可使用命令: hdfs colocationadmin -listGroups),执行以后,就会删除该组下面的 locators 与数据节点的对应关系。

步骤 3 执行创建新的 locators 方法(可选)

## 客户端的安装目录/ HDFS/hadoop/bin/hdfs colocationadmin -createGroup -groupId Colocation 组 ID -locatorIds 该组下的 locators

例如: /opt/hadoopclient/HDFS/hadoop/bin/ hdfs colocationadmin -createGroup - groupId 2015 -locatorIds 01,02,03,04,05,06,07,08,09,10,11,12

hdfs colocationadmin -createGroup -groupId 2015 -locatorIds 01,02,03,04,05,06,07,08,09,10,11,12: 这个命令中-groupId 表示创建一个叫 2015 的组,-locatorIds 指定在该组下创建 01 到 12,共 12 个 locators,后续就可以把数据按照月份存储到各自的 locators 中。

----结束

# 6 其他组件减容操作指导

- 6.1 减容 Redis 集群
- 6.2 NodeManager 减容前退服
- 6.3 Solr 减少组件实例
- 6.4 Kafka 减容后的 Reassign

## 6.1 减容 Redis 集群

- 步骤 1 登录 FusionInsight Manager,单击"服务管理"。
- 步骤 2 单击服务列表中的 Redis。
- 步骤 3 单击 "Redis 管理"页签。
- 步骤 4 选择 Redis 集群列表中的一行,点击"操作"列的减容按钮进行 Redis 集群减容操作。

#### 图6-1 减容操作



步骤 5 选择要减容的 Redis 实例,并点击下一步。

#### □ 说明

为保证集群数据的高可靠性,系统须确保 Redis 集群内的实例满足以下条件,否则无法正常分配主从关系。

- 一个 Redis 集群内的 Redis 实例必须分布到至少三台节点中(至少 6 个 Redis 实例)。
- 一个 Redis 集群内的 Redis 实例总数为偶数,且同一组主从实例(例如 M1 与 S1)不能在一台节点上。
- 任意一台节点上的主 Redis 实例必须小于整个 Redis 集群主实例个数的一半。

#### 图6-2 选择实例



#### 步骤 6 确认后点击提交。

#### 图6-3 提交任务



#### ----结束

## 6.2 NodeManager 减容前退服

- □ 说明
  - 业务相对空闲时执行, 因为退服操作会占用一定的带宽资源。
- 步骤 1 登陆 FusionInsight Manager。
- 步骤 2 进入服务管理 > Yarn,点击"实例"进入 Yarn 实例页面。
- 步骤 3 在 Yarn 实例页面,选择,要退服的 NodeManager,然后点击"更多操作"中的"退服"进行退服操作。



步骤 4 如果 NodeManager 退服完成,实例的状态会变为已退服。



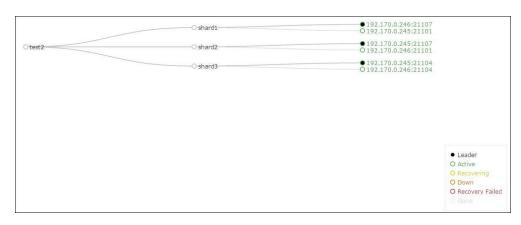
□ 说明

只能对退服后的 NodeManager 实例做删除实例减容的操作。

#### ----结束

## 6.3 Solr 减少组件实例

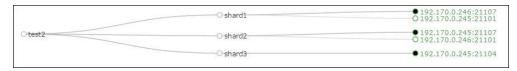
- 步骤 1 登陆 FusionInsight Manager。
- 步骤 2 进入服务管理 > Solr > 服务状态,点击 Solr WebUI 的右边的 SolrServerAdmin,跳转进入 Solr Admin 页面。
- 步骤 3 在 Solr Admin 的页面,进入 Cloud > Graph,查看 Collection 的图,确保要删除的实例上的 Core 所属的 shard 下的 replica 在其他的实例存在。如下图中,要删除192.170.0.246:21104 这个 SolrServer 的实例,确保其所在 shard3 下还有其他 replica。



步骤 4 进入服务管理 > Solr > 实例中,选择要删除的实例删除实例。



步骤 5 进入服务管理 > Solr > 服务状态,点击 Solr WebUI 的右边的 SolrServerAdmin,跳转进入 Solr Admin 页面。在 Solr Admin 的页面,进入 Cloud > Graph,查看 Collection 的图,查看 Solr 实例正常,则删除 SolrServer 实例成功。



#### □ 说明

集群中删除 Solr Server 实例时,如果删除实例为 Solr Server Admin,由于 Solr Server Admin 的双活机制,需要对 Solr Server Admin 实例进行迁移。具体迁移步骤请参考《Fusion Insight HD V100R002C60U10 集群组件迁移指导》中 2.4.2 Solr Server Admin 迁移章节。

#### ----结束

## 6.4 Kafka 减容后的 Reassign

Kafka 减容后,Topic Partitions 不会进行 Reassign。为了保证数据分布均衡,在减容 Broker Server 后,一般都会执行 Reassign 将已有数据进行重分布。

#### 建议:

- 业务相对空闲时执行,因为 Reassign 操作会占用一定的带宽资源。
- 执行 Reassign 操作会使得数据进行重分布,执行过程中会影响当前业务。
- 数据迁移一旦开始无法停止,不要强行停止集群,造成数据不一致。
- 数据迁移是先拷贝,后删除,请确保足够空间。

例如客户端安装到指定目录"/opt/streamclient"。

- 步骤 1 在客户端的安装目录下,执行 source bigdata\_env 进行初始化。
- 步骤 2 (普通模式请跳过该步骤) 执行 kinit admin,使用 admin 用户进行安全登录,根据提示输入密码。
- 步骤 3 执行 kafka-broker-info.sh --zookeeper {zk\_host}:{port}/kafka,, 获取每个节点的 Broker\_ID。

- 步骤 4 执行 kafka-topics.sh --list --zookeeper {zk\_host}:{port}/kafka, 查询当前 Topic 信息。
- 步骤 5 创建 Json 文件,内容为需要移动的 topic 信息,文件保存为 topics-to-move.json。

例如:

```
{ ''topics'': [{"topic":"topic1"}],
''version'':1
}
```

步骤 6 指定移动 Topic 和 Broker 列表,生成重分布信息。

#### kafka-reassign-partitions.sh

- --zookeeper {zk\_host}:{port}/kafka
- --topics-to-move-json-file {topics to reassign json file path}
- --broker-list {brokerlist}
- --generate

例如:

#### kafka-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a
- --topics-to-move-json-file topics-to-move.json
- --broker-list "23,24,25"
- --generate

命令会打印 Topic Partitions 原始分布情况,以及 Kafka 生成的重分布后的 Partitions 分布情况。用户可以将原始分布信息保存为文件,以便进行回滚;同时可以对重分布后的 Partitions 分布情况进行手动修改,将新的 Partitions 分布信息保存为 Json 文件。例如:expand-cluster-reassignment.json。

如果有磁盘容量告警,获取该磁盘目录下 Topic Partitions 作为 Reassign Partitions。

步骤 7 根据新的 Partitions 分布情况执行重分布。

#### kafka-reassign-partitions.sh

- --zookeeper {zk\_host}:{port}/kafka
- --reassignment-json-file {manual assignment json file path}
- --execute

例如:

/opt/streamclient/Kafka/kafka/bin/kafka-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a
- --reassignment-json-file expand-cluster-reassignment.json
- --execute

命令会打印 Topic Partitions 原始分布情况和新的分布情况。

步骤 8 确认 Partitions 重分布执行结果。

### kafka-reassign-partitions.sh

- --zook eeper {zk\_host}:{port}/kafka
- --reassignment-json-file {manual assignment json file path}
- --verify

例如:

/opt/streamclient/Kafk a/k afk a/bin/k afk a-reassign-partitions.sh

- --zook eeper 10.149.0.90:24002,10.149.0.91:24002,10.149.0.92:24002/k afk a
- --reassignment-json-file expand-cluster-reassignment.json
- --verify

命令会打印 Topic Partitions 分配状态信息。例如:

Status of partition reassignment:

Reassignment of partition [topic 1,1] completed successfully

Reassignment of partition [topic 1,0] completed successfully

Reassignment of partition [topic1,2] completed successfully

----结束

## **7** 减容验证

步骤 1 登录 FusionInsight Manager。

步骤 2 单击"服务管理 > 更多操作 > 启动集群健康检查",执行集群的健康检查,确保减容后所有已安装服务的"服务状态"为"良好"。

----结束