

HDFS安装部署

www.huawei.com





目标

- 学完本课程后，您将能够：
 - 掌握**HDFS**组件安装部署
 - 熟悉**HDFS**常用配置参数



目录

1. 部署规划
2. 部署方法
3. 常用参数配置
4. 健康检查关键指标

HDFS部署组网

组网原则	适用场景
管理节点、控制节点和数据节点分开部署	（推荐）节点数大于 200 节点的集群使用此场景，此方案至少需要 8 个节点
管理节点和控制节点合并部署，数据节点单独部署	（推荐）节点数大于等于 6 ，小于等于 200 节点的集群使用此场景
管理节点、控制节点和数据节点合并部署	节点数小于 6 的集群使用此场景，此方案至少需要 3 个节点

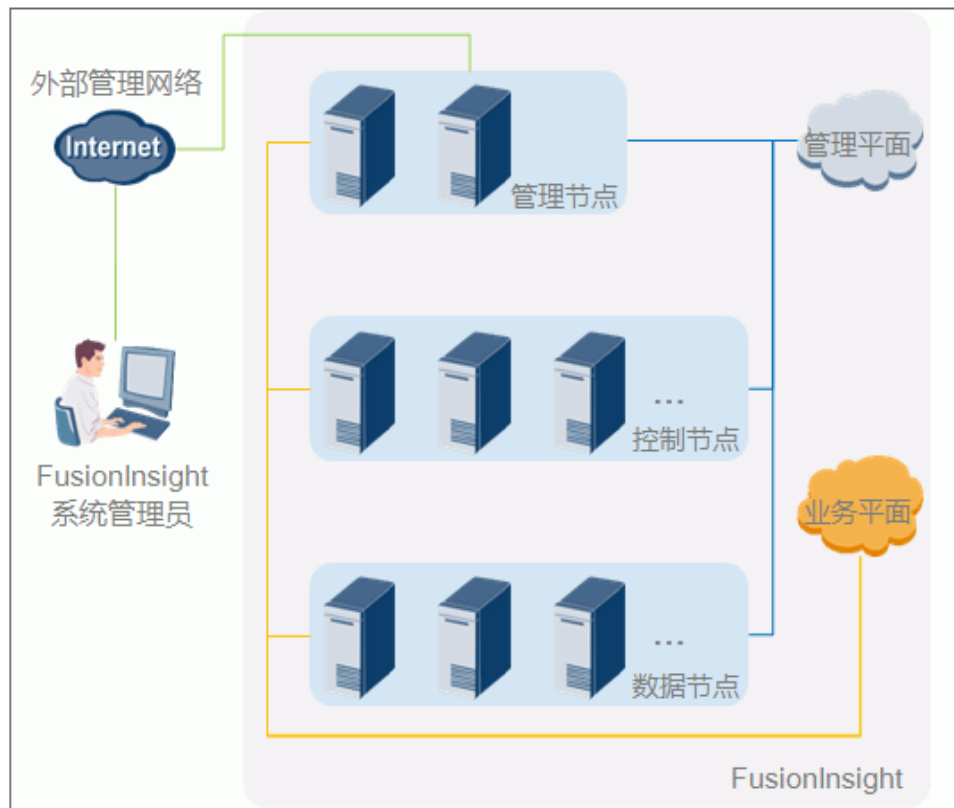
HDFS典型组网

FusionInsight HD整个系统网络划分为**2**个平面，即业务平面和管理平面，两个平面之间采用物理隔离的方式进行部署，保证业务、管理各自网络的安全性。

业务平面通过业务网络接入，为用户和上层用户提供业务通道，对外提供数据存取、任务提交及计算能力。

管理平面通过运维网络接入，提供系统管理和维护功能。用于集群的管理，对外提供集群监控、配置、审计、用户管理等服务。

在典型配置下，**FusionInsight HD** 集群采用双平面组网，如右图所示。



HDFS部署原则及约束

服务名称	角色名称	内存最小要求	依赖关系	角色业务部署原则
HDFS	NN (NameNode)	4GB	NameNode与ZKFC关联 依赖于ZooKeeper	分别部署在2个控制节点上，主备配置。
	ZKFC (ZooKeeper FailoverController)	1GB		分别部署在2个控制节点上，主备配置。
	JN(JournalNode)	4GB		至少部署3个在控制节点上，每个节点保留一份备份数据。如需保留超过三份以上备份，可部署多个在控制或数据节点上，请保持数量为奇数个。
	DN (DataNode)	4GB		至少部署3个，建议部署在数据节点上。



目录

1. 部署规划

2. 部署方法

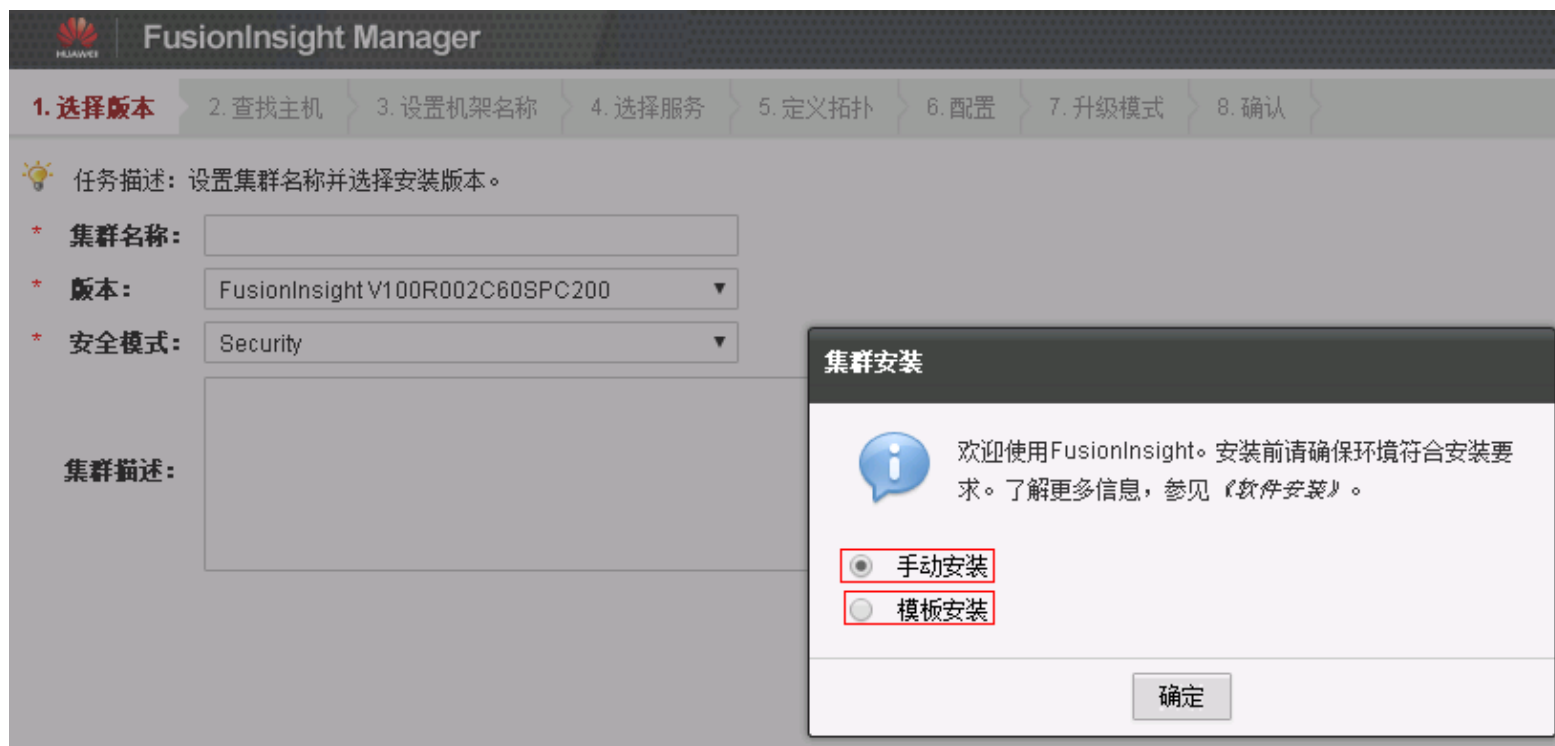
- 新建部署
- 扩容部署
- 减容部署

3. 常用参数配置

4. 健康检查关键指标


新建部署

- 1、进入安装页面，选择“手动安装”（也可选择“模板安装”，这里以手动安装为例），点击确定。



新建部署

2、输入主机IP以及主机root用户密码，点击“发现主机”，选择发现成功的节点，点击“下一步”。

 FusionInsight Manager

1. 选择版本

2. 查找主机

3. 设置机架名称


4. 选择服务


5. 定义拓扑


6. 配置

7. 升级模式

8. 确认

 任务描述：基于主机匹配表达式搜索主机，输入的用户名和密码用于连接目的主机，要求所有主机配置相同的用户名和密码。

 帮助：关于主机匹配表达式

 输入所有节点（包括管理节点、控制节点和数据节点）的管理平面地址。

* 匹配表达式：

192.168.1.81

* 用户名：


root

* 密码：

查找主机

停止查找

清除结果

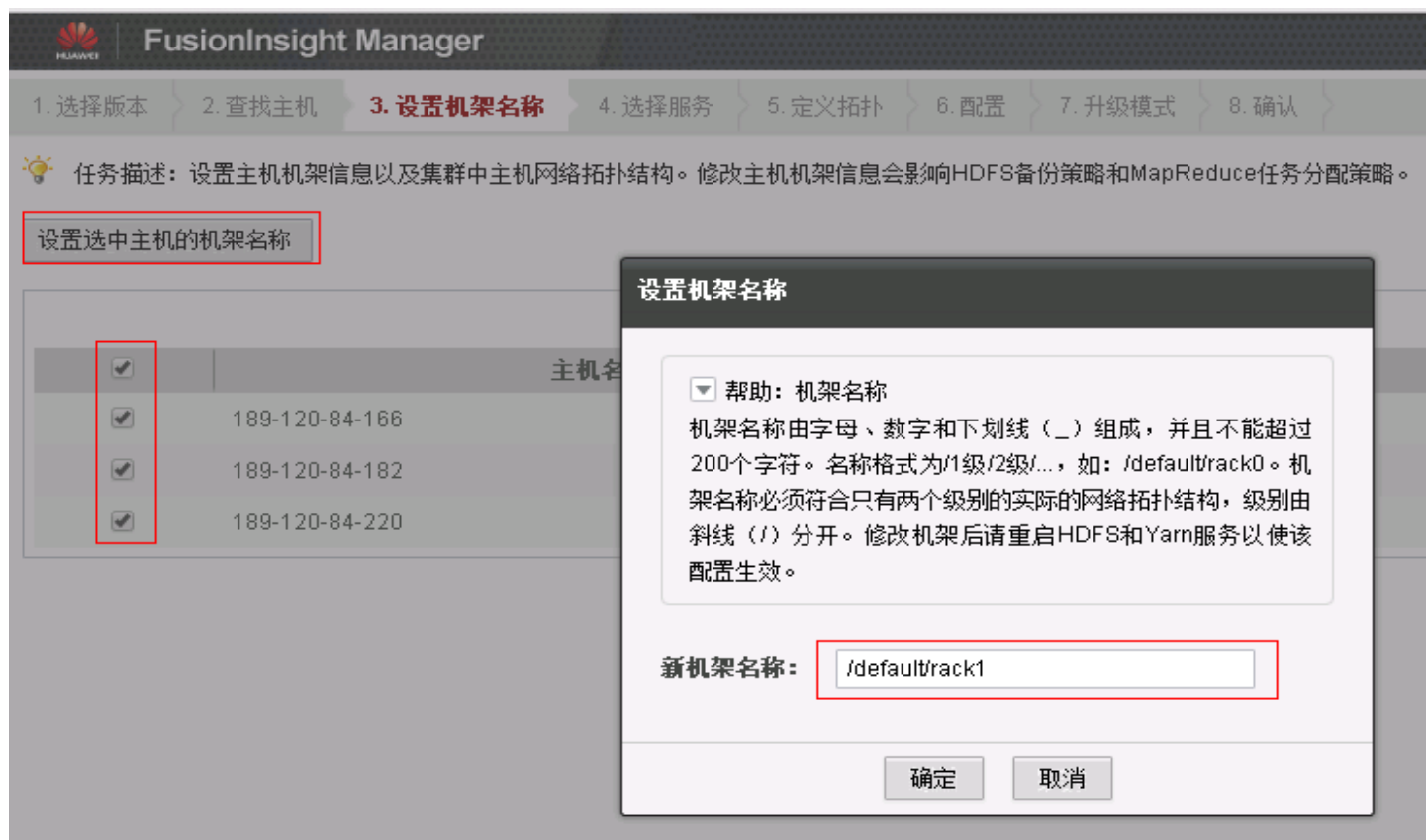
 选择所有节点（包括管理节点、控制节点和数据节点）。

发现主机：3个主机。（主机名为绿色表示此节点为OMS节点。请选择要添加的节点，然后单击下一步。）

<input checked="" type="checkbox"/>	主机名	管理IP	业务IP
<input checked="" type="checkbox"/>	189-120-84-166		
<input checked="" type="checkbox"/>	189-120-84-182		
<input checked="" type="checkbox"/>	189-120-84-220		

新建部署

- 3、选中若干主机，点击“设置选中主机的机架名称”，在弹出的对话框中设置“新机架名称”，点击“确定”，然后选中所有主机，点击“下一步”。



新建部署

4、选择安装服务

根据业务需要选择待安装的服务。当选择**HDFS**服务时，系统会自动为其选择依赖的**ZooKeeper**服务。

<input type="checkbox"/>	Streaming	Streaming是一个基于Apache Storm的分布式实时计算系统。
<input type="checkbox"/>	Kafka	Kafka是一个分布式消息发布订阅系统。
<input checked="" type="checkbox"/>	HDFS	用于海量数据存储的Hadoop分布式文件系统。
<input checked="" type="checkbox"/>	ZooKeeper	Zookeeper是一个集中的服务，它用于维护配置信息、命名、提供分布式协调。
<input type="checkbox"/>	Redis	Redis是基于BSD开源协议开发的高性能Key-Value存储系统。它通常用于缓存。
<input type="checkbox"/>	Metadata	元数据管理系统，用于查询和管理集群中产生的所有元数据。
<input type="checkbox"/>	Hue	Apache Hadoop的UI界面。
<input checked="" type="checkbox"/>	KrbServer	KrbServer是密钥的管理中心，并负责票据的分发。
<input checked="" type="checkbox"/>	LdapServer	LdapServer是用户和群组的管理中心。

新建部署

5、拓扑设置

根据拓扑规划，将组件部署到指定节点上。完成设置后点击“下一步”。



 **任务描述：**按照服务部署的要求，设置服务和主机之间的映射关系。（可以将鼠标移动到角色名上查看部署建议。）

主机名	HDFS			
	<input checked="" type="checkbox"/> DN (3)	JN (3)	zkfc (2)	NN (2)
189-120-84-166	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
189-120-84-182	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
189-120-84-220	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

新建部署

6、配置

为Ladp配置一个未被占用的管理平面IP地址作为浮动IP，并配置管理平面网关。然后点击“下一步”。

The screenshot shows the 'FusionInsight Manager' configuration interface. At the top, there is a progress bar with steps: 1. 选择版本, 2. 查找主机, 3. 设置机架名称, 4. 选择服务, 5. 定义拓扑, 6. 配置 (highlighted in red), 7. 升级模式, and 8. 确认. Below the progress bar, a task description states: '任务描述: 选择服务、角色和主机, 查看和修改参数配置。'. There are two buttons: '导入集群配置' and '参数类别: 基础配置'. To the right, the '服务:' dropdown is set to 'LdapServer'. The main configuration area is titled 'LdapServer->SlapdServer' and contains the following parameters:

参数	配置
LDAP_ENABLE_HA	<input checked="" type="radio"/> true <input type="radio"/> false
* LDAP_FLOAT_IP	192.168.2.43
* LDAP_MEDIATOR_IP	192.168.2.66
* LDAP_SERVER_PORT	21780


At the bottom, there is a pagination bar showing '当前页: 1 总页数: 1' and a '跳转到' field with a '确定' button.

新建部署

7、升级模式

了解“普通模式”和“高级模式”的描述信息后，选中“普通模式”，点击“下一步”。



 任务描述：选择升级模式。

☒ 普通模式（推荐）

在普通模式下，你只需要在Manager主备管理节点使用root用户，其它节点使用运行用户，即可完成FusionInsight的升级安装操作。

☐ 高级模式（谨慎使用）

在高级模式下，用户必须使用root用户进行升级安装操作，此时必须保证集群内所有节点均可通过root用户远程登录，且密码一致。

新建部署

8、提交安装

完成配置操作后，进入确认页面，确认页面包含前面步骤的概要信息。然后点击“提交”。

新建部署

9、部署完成

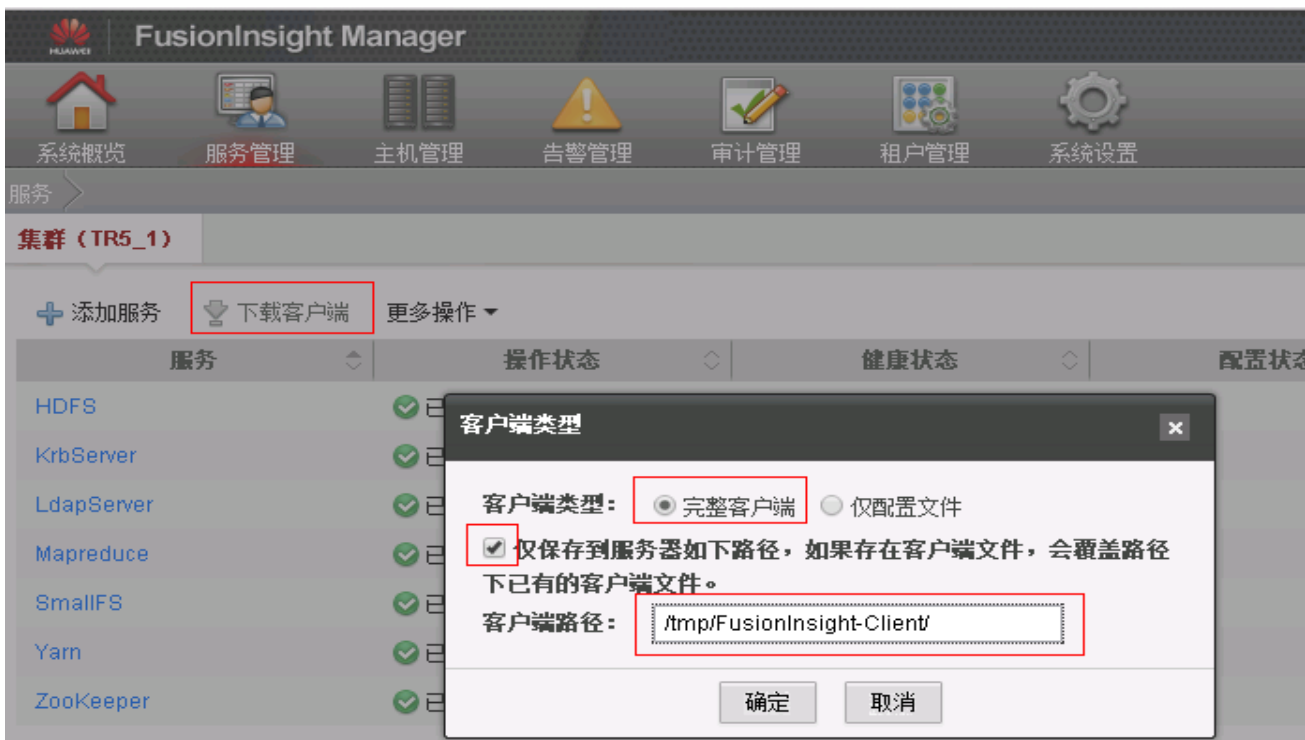
提交集群安装请求后，界面显示安装集群进度，全部步骤完成后点击“完成”。

安装集群				
	步骤	开始时间	进度	结束时间
+	1. 校验请求参数	2016-02-05 18:49:33	100%	2016-02-05 18:49:43
+	2. 初始化系统环境	2016-02-05 18:49:43	100%	2016-02-05 18:49:55
+	3. 配置系统环境	2016-02-05 18:49:55	100%	2016-02-05 18:49:59
+	4. 下载文件	2016-02-05 18:49:59	100%	2016-02-05 18:50:35
+	5. 安装节点	2016-02-05 18:50:36	100%	2016-02-05 18:52:46
+	6. 安装组件	2016-02-05 18:52:46	100%	2016-02-05 18:53:25
+	7. 配置集群	2016-02-05 18:53:25	100%	2016-02-05 18:54:15
+	8. 初始化集群	2016-02-05 18:54:15	100%	2016-02-05 18:56:44
+	9. 启动集群	2016-02-05 18:57:10	100%	2016-02-05 19:03:35
+	10. 持久化集群配置	2016-02-05 19:03:35	100%	2016-02-05 19:03:37
<p>✓ 操作成功。 你已经成功安装了集群，请单击完成继续。</p>				
完成				

新建部署

10、客户端下载

在FusionInsight Manager的“服务管理”界面，单击“下载客户端”，下载客户端到主管理节点的指定路径。



新建部署

11、客户端部署

- 使用**Xshell**登录客户端所在节点，解压前一步获取的客户端安装包。
- 进入解压文件夹，安装客户端：**sh install.sh /opt/client**
- 进入客户端安装目录**/opt/client**，初始化环境变量：

source bigdata_env

- 以**hdfs**用户登录认证：**kinit <hdfs用户>**

执行查看命令检查客户端是否部署成功：**hdfs dfs -ls /**

- 显示以下结果，说明部署成功。

```
[root@189-120-85-207 client]# hdfs dfs -ls /
16/04/27 17:44:23 INFO hdfs.PeerCache: SocketCache disabled.
Found 7 items
drwxrwx---+ - hive   hive           0 2016-04-25 20:21 /apps
drwxr-x---+ - flume  hadoop        0 2016-04-25 19:34 /flume
drwxrwxrwx+ - mapred hadoop        0 2016-04-25 19:35 /mr-history
drwxrwxrwx+ - spark  supergroup    0 2016-04-25 20:30 /sparkJobHistory
drwxrwx--x+ - admin  supergroup    0 2016-04-25 20:24 /tenant
drwxrwxrwx+ - hdfs   hadoop        0 2016-04-25 20:29 /tmp
drwxrwxrwx+ - hdfs   hadoop        0 2016-04-25 20:30 /user
```



目录

1. 部署规划

2. 部署方法

- 新建部署
- 扩容部署
- 减容部署

3. 常用参数配置

4. 健康检查关键指标

扩容部署

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，如**DataNode**。

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，无论全新添加服务还是为业务组件添加业务实例，在**FusionInsight Manager**上操作均为添加服务实例。

扩容部署

- 1、在安装了**HDFS**服务的集群管理界面，进入“服务管理”界面，点击**HDFS**服务，进入**HDFS**服务界面



The screenshot displays the FusionInsight Manager web interface. At the top, there's a navigation bar with icons for System Overview, Service Management (highlighted), Host Management, Alarm Management, Audit Management, Tenant Management, and System Settings. Below this, a 'Services' section shows a cluster named '集群 (C60)'. Underneath, there are buttons for 'Add Service', 'Download Client', and 'More Operations'. A table lists the services and their status.

服务	操作状态	健康状态	配置状态
HDFS	✓ 已启动	✓ 良好	✓ 已同步
KrbServer	✓ 已启动	✓ 良好	✓ 已同步
LdapServer	✓ 已启动	✓ 良好	✓ 已同步
ZooKeeper	✓ 已启动	✓ 良好	✓ 已同步

扩容部署

2、在HDFS服务界面，点击“实例”页签，进入HDFS实例界面

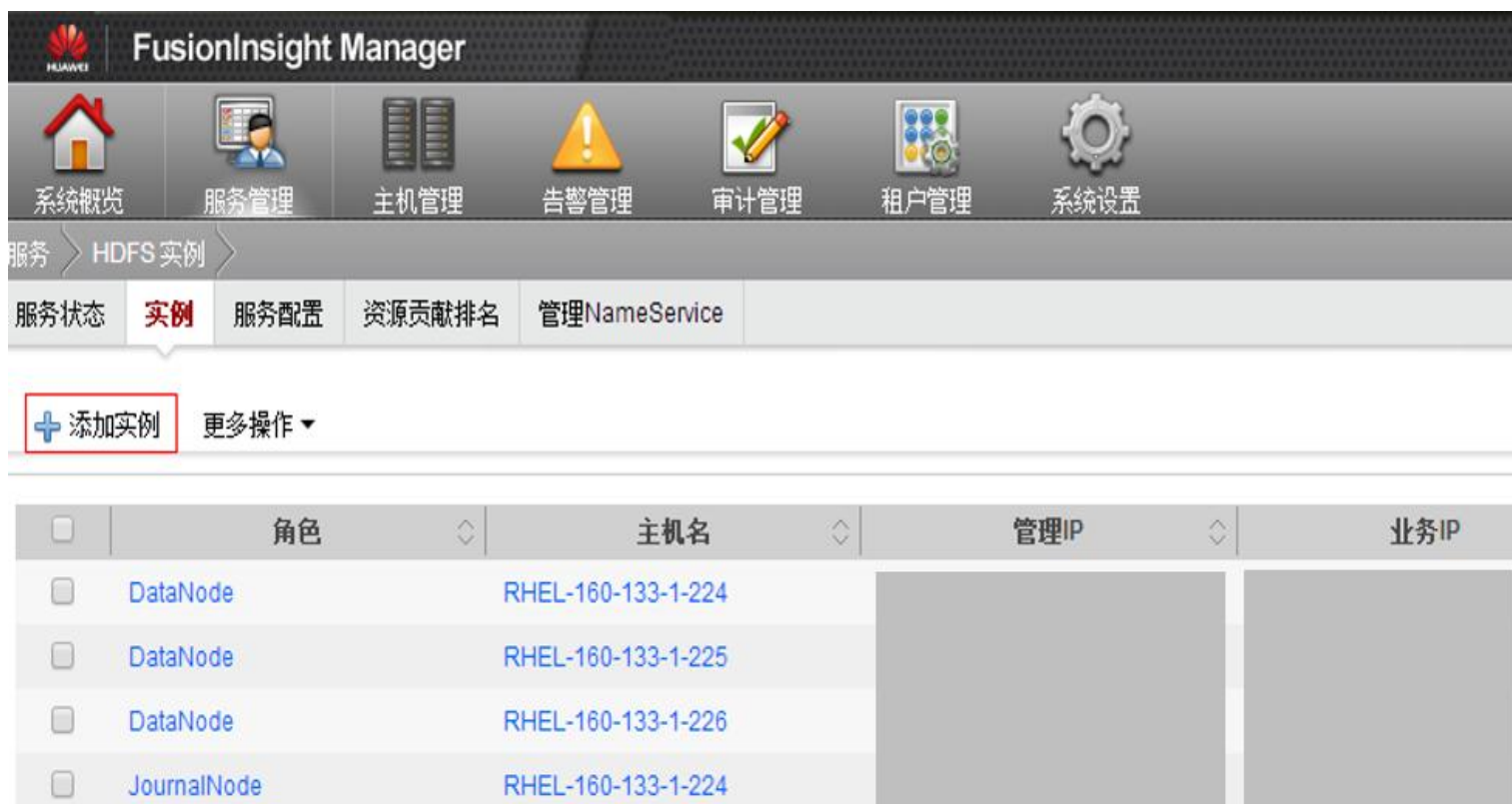
The screenshot shows the FusionInsight Manager web interface. At the top, there's a navigation bar with icons for System Overview, Service Management (selected), Host Management, Alarm Management, Audit Management, Tenant Management, and System Settings. Below this, a breadcrumb shows 'Service > HDFS Service Status'. A sub-navigation bar includes 'Service Status', 'Instance' (highlighted with a red box), 'Service Configuration', 'Resource Contribution Ranking', and 'Manage NameService'. Action buttons for 'Start Service', 'Stop Service', 'Download Client', and 'More Operations' are present. The main content area is divided into two panels. The left panel, titled 'HDFS Overview', lists various metrics: Health Status (Good), Configuration Status (Synchronized), Version (2.7.2), Read Rate (0.0 Byte), Write Rate (0.0 Byte, highlighted in blue), Security Mode (OFF), and Disk Space (24.35% 84.73 GB / 347.98 GB). The right panel, titled 'Real-time', shows a chart for 'HDFS configured total storage space' with a unit of GB. The chart has a y-axis ranging from 348.00 to 351.00 GB.

HDFS 概述	
健康状态	良好
配置状态	已同步
版本	2.7.2
读速率	0.0 Byte
写速率	0.0 Byte
安全模式	OFF
磁盘空间	24.35% 84.73 GB / 347.98 GB

HDFS配置的总存储空间	
单位：GB	
351.00	
350.00	
349.00	
348.00	

扩容部署

3、在HDFS实例界面，点击“添加实例”操作



FusionInsight Manager

系统概览 服务管理 主机管理 告警管理 审计管理 租户管理 系统设置

服务 > HDFS 实例

服务状态 **实例** 服务配置 资源贡献排名 管理NameService

+ 添加实例 更多操作 ▾

<input type="checkbox"/>	角色	主机名	管理IP	业务IP
<input type="checkbox"/>	DataNode	RHEL-160-133-1-224		
<input type="checkbox"/>	DataNode	RHEL-160-133-1-225		
<input type="checkbox"/>	DataNode	RHEL-160-133-1-226		
<input type="checkbox"/>	JournalNode	RHEL-160-133-1-224		

扩容部署

4、若扩容主机不在主机列表里，需点击“添加主机”，输入主机IP以及主机root用户密码，点击“发现主机”，选择发现成功的节点，点击“下一步”。

1. 查找主机

2. 设置机架

3. 定义拓扑

4. 配置

5. 确认

基于主机匹配表达式搜索主机，输入的用户名和密码用于连接目的主机，要求所有主机配置相同的用户名和密码。

▶ 帮助：关于主机匹配表达式

输入所有节点（包括管理节点、控制节点和数据节点）的管理平面地址。

192.168.1.81

* 匹配表达式：

✓

* 用户名：

root

* 密码：

.....

输入密码。

查找主机

停止查找


清除结果

选择所有节点（包括管理节点、控制节点和数据节点）。

发现主机： 1 个主机。（主机名为绿色表示此节点为OMS节点或者已安装的节点。请选择要添加的节点，然后单击 下一步。）

Copyright © 2016 Huawei Technologies Co., Ltd. All rights reserved.

Page 24

 HUAWEI

扩容部署

5、在扩容的节点上规划HDFS的角色，如DN，点击“下一步”



任务描述：按照服务部署的要求，设置服务和主机之间的映射关系。（可以将鼠标移动到角色名上查看部署建议。）

主机名	HDFS		ZooKeeper
	<input checked="" type="checkbox"/> DN (4)	JN (3)	<input type="checkbox"/> QP (3)
RHEL-160-163-0-231	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
RHEL-160-133-1-226	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
RHEL-160-133-1-225	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
RHEL-160-133-1-224	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

扩容部署

6、根据扩容规划，配置需要扩容服务的参数。点击“下一步”，对照扩容规划内容，确认信息无误，点击“提交”。

FusionInsight Manager

系统概览 服务管理 **主机管理** 告警管理 审计管理 租户管理 系统设置

主机管理 > 添加主机 >

1. 查找主机 2. 设置机架 3. 定义拓扑 **4. 配置** 5. 确认

任务描述：选择服务、角色和主机，查看和修改参数配置。

导入集群配置 参数类别：基础配置 服务：HDFS 角色：DataNode

参数	值
HDFS->DataNode	
* dfs.datanode.data.dir	/srv/BigData/hadoop/data1/dn
* dfs.datanode.http.port	25010
* dfs.datanode.https.port	25011
* dfs.datanode.ipc.port	25008
* dfs.datanode.port	25009
HDFS->DataNode->160.163.0.231	
* dfs.datanode.data.dir	/srv/BigData/hadoop/data1/dn

扩容部署

7、实例扩容成功。

添加主机

	步骤	开始时间	进度	结束时间
+	1. 校验请求参数	2016-02-05 19:27:30	100%	2016-02-05 19:27:39
+	2. 初始化系统环境	2016-02-05 19:27:39	100%	2016-02-05 19:27:47
+	3. 配置系统环境	2016-02-05 19:27:47	100%	2016-02-05 19:27:52
+	4. 下载文件	2016-02-05 19:27:52	100%	2016-02-05 19:28:15
+	5. 安装节点	2016-02-05 19:28:15	100%	2016-02-05 19:30:01
+	6. 安装组件	2016-02-05 19:30:01	100%	2016-02-05 19:30:28
+	7. 配置集群	2016-02-05 19:30:28	100%	2016-02-05 19:31:09
+	8. 初始化服务和实例	2016-02-05 19:31:09	100%	2016-02-05 19:31:58
+	9. 启动服务和实例	2016-02-05 19:31:58	100%	2016-02-05 19:32:19
+	10. 持久化集群配置	2016-02-05 19:32:19	100%	2016-02-05 19:32:21

✓ 操作成功。

单击完成继续。

完成

扩容部署

8、查看HDFS实例，成功扩容添加DataNode实例。

服务状态	实例	服务配置	资源贡献排名	管理NameService
+ 添加实例 更多操作 ▾				
<input type="checkbox"/>	角色	主机名	管理IP	业务IP
<input type="checkbox"/>	DataNode	RHEL-160-133-1-224		
<input type="checkbox"/>	DataNode	RHEL-160-133-1-225		
<input type="checkbox"/>	DataNode	RHEL-160-133-1-226		
<input type="checkbox"/>	DataNode	RHEL-160-163-0-231		
<input type="checkbox"/>	JournalNode	RHEL-160-133-1-224		



目录

1. 部署规划

2. 部署方法

- 新建部署
- 扩容部署
- 减容部署

3. 常用参数配置

4. 健康检查关键指标

减容部署

典型的减容场景

1. 节点可能存在无法修复的硬件错误。
2. 节点需要重新安装操作系统。

减容约束

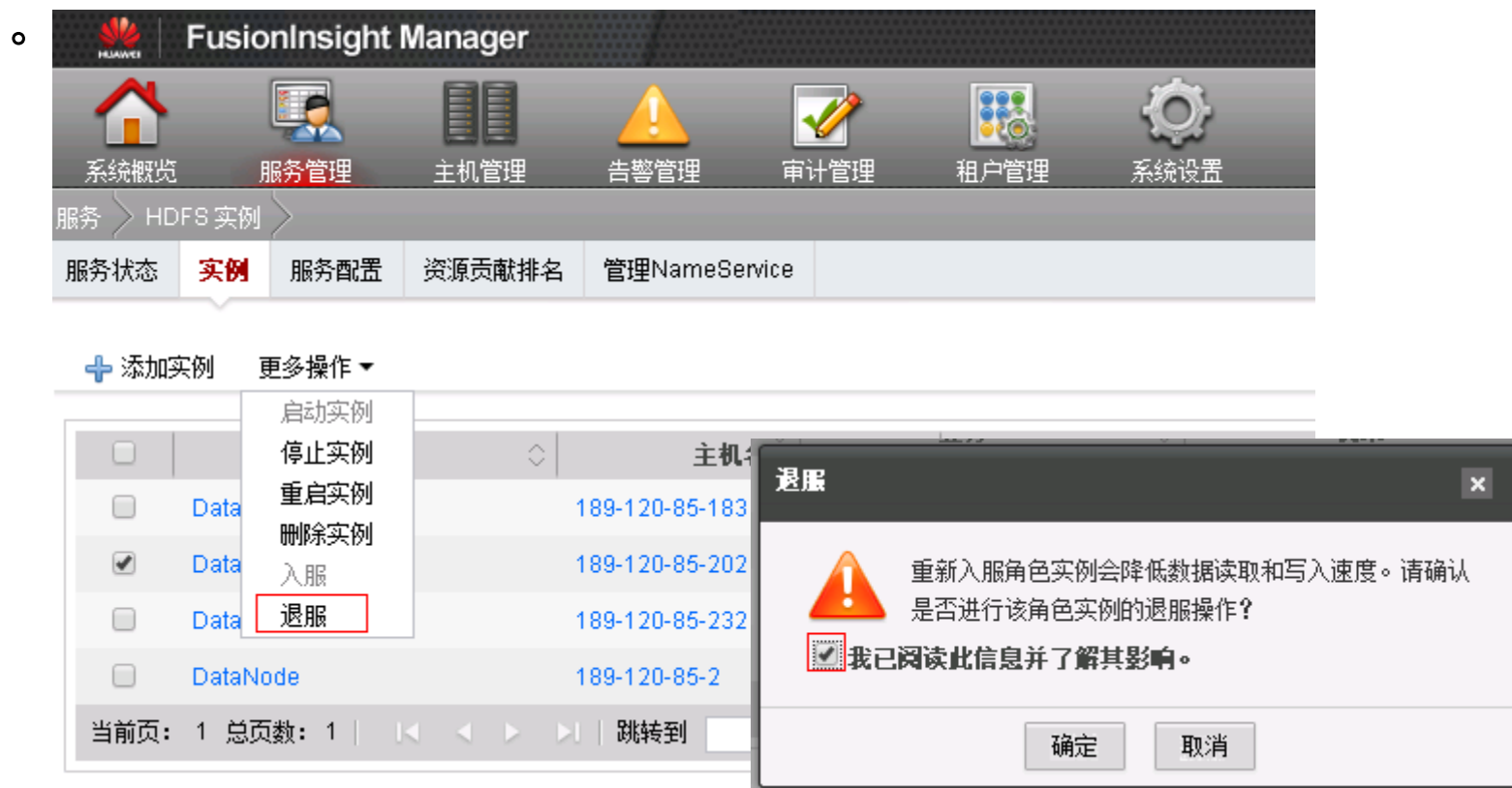
DataNode减容前需要执行退服操作。

减容影响

HDFS存储空间减少。

减容部署

- 1、点击“服务管理 > **HDFS** > 实例”，进入**HDFS**实例界面，选择待减容的实例，点击“更多操作”中的“退服”，开始退服实例。



减容部署

2、退服成功后，勾选减容实例节点，点击“更多操作”中的“删除实例”操作。

服务 > HDFS 实例

服务状态 **实例** 服务配置 资源贡献排名 管理NameService

+ 添加实例 更多操作 ▼

		主机名	管理IP	业务IP	机架	操作状态
<input type="checkbox"/>	DataNode	189-120-85-183			/default/rack0	✓ 已启动
<input type="checkbox"/>	DataNode	189-120-85-202			/default/rack0	✗ 已退服
<input checked="" type="checkbox"/>	DataNode	189-120-85-232			/default/rack0	✓ 已启动
<input type="checkbox"/>	DataNode	189-120-85-2			/default/rack0	✓ 已启动

当前页: 1 总页数: 1 | < < > > | 跳转到 确定

启动实例

停止实例

重启实例

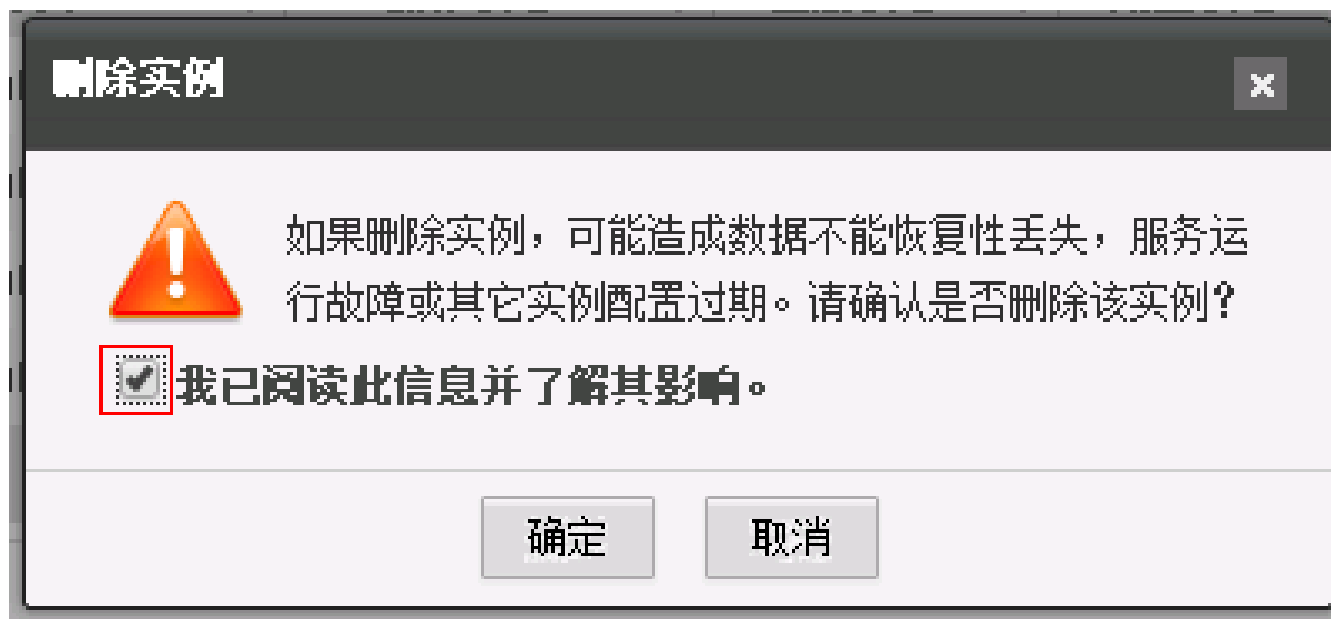
删除实例

入服

退服

减容部署

3、在确认减容相关影响提示后，点击“确定”，开始删除实例。



减容部署

4、删除完成：确认**DataNode**实例删除操作请求后，界面显示删除实例进度，全部完成后点击“完成”。

减容部署

5、查看HDFS实例，成功减容删除DataNode实例。

服务 > HDFS 实例

服务状态

实例

服务配置

资源贡献排名

管理NameService

+ 添加实例

更多操作 ▾

<input type="checkbox"/>	角色 ▾	主机名 ▾	管理IP ▾
<input type="checkbox"/>	DataNode	189-120-85-183	
<input type="checkbox"/>	DataNode	189-120-85-232	
<input type="checkbox"/>	DataNode	189-120-85-2	

当前页: 1 总页数: 1

⏪

⏩

⏴

⏵

跳转到 确定



目录

1. 部署规划
2. 部署方法
3. 常用参数配置
4. 健康检查关键指标

常用参数说明及配置建议-HDFS

参数	建议值	说明
<code>dfs.replication</code>	3	HDFS 存储数据的副本数。建议为 3 份，修改后， HDFS 将全局刷新副本数
<code>dfs.client.block.write.replace-datanode-on-failure.replication</code>	2	客户端在写入数据时，会强制要求所有数据写成功，否则就写入失败。当数据节点数为 3 台时，建议修改为 2 ，以防止数据节点复位，写入失败。
<code>dfs.namenode.rpc.port</code>	8020	HDFS 内部 RPC 的通讯端口
<code>dfs.namenode.http.port</code>	25000	HDFS HTTP 服务的端口。

常用参数说明及配置建议-NameNode

参数	建议值	说明
dfs.namenode.image.backup.enable	true	NameNode 元数据是否周期备份到 OMS 服务器上。备份文件在“系统设置”->“备份管理”中可以看到。
dfs.namenode.image.backup.num	48	元数据备份个数，元数据备份为每 30 分钟备份一次，一天可以产生 48 个元数据备份。
dfs.blocksize	134217728	数据块大小，默认为 128M ，单位为字节。系统安装后，不建议修改。
dfs.namenode.handler.count	64	NameNode 并行处理任务数，调大可以提高性能。
fs.trash.checkpoint.interval	60	HDFS 回收站扫描周期，单位为分钟。
fs.trash.interval	1440	HDFS 回收站文件删除时间，单位为分钟，默认 1 天。
dfs.permissions.enabled	true	HDFS 文件系统是否启用 ACL 安全管理，建议打开。

常用参数说明及配置建议-DataNode

参数	建议值	说明
dfs.blocksize	134217728	数据块大小，默认为 128M ，单位为字节。系统安装后，不建议修改。
dfs.datanode.handler.count	8	DataNode 并行处理任务数，调大可以提高性能。
dfs.datanode.directoryscan.interval	21600	DataNode 数据块校验周期，此功能用于检查损坏的数据块。默认为 21 天。参数单位为小时。
dfs.datanode.balance.bandwidthPerSec	20971520	HDFS Balancer 功能的最高流量，调大后能缩短 Balancer 时间，但会占用较多网络带宽。
dfs.datanode.data.dir.perm	700	DataNode 创建存储目录的默认权限，能防止数据被非 root 账号非法读取。



目录

1. 部署规划
2. 部署方法
3. 常用参数配置
4. 健康检查关键指标

健康检查关键指标

关键指标项	描述
发送包的平均时间统计	x.x ns 小于 1000000.0 ns
服务状态	健康
告警信息	无告警信息

健康检查关键指标操作指导

在FusionInsight Manager界面，单击“服务管理 > HDFS > 更多操作 > 启动服务健康检查”，即可查看健康检查结果，如下图。



总结

- 部署规划
- 部署方法
- 组件配置
- 健康检查关键指标



思考题

1. **HDFS** 有哪些角色，所依赖的服务有哪些？
2. 如何安装**HDFS**客户端？
3. 如何查看**HDFS**健康检查关键指标？



习题

1. **NameNode**、**ZKFC**、**journalNode**的依赖启动顺序是（ ）

- A. **NameNode** -> **ZKFC** -> **journalNode**
- B. **ZKFC** -> **journalNode** -> **NameNode**
- C. **journalNode**-> **ZKFC** -> **NameNode**
- D. **ZKFC** -> **NameNode** -> **journalNode**

2. **HDFS**中的**block**默认保存几个副本？

- A. 3
- B. 2
- C. 1
- D. 不确定

3. （判断题） 集群升级过程中，在观察期时可以进行扩容减容操作。

Thank you

www.huawei.com