

翼 MapReduce

用户操作指南

天翼云科技有限公司

目录

1. 准备用户	5
1.1. 创建翼 MapReduce 操作用户	5
1.2. IAM 用户同步	7
2. 配置集群	8
2.1. 创建方式简介	8
2.2. 快速创建集群	9
2.2.1. 快速创建数据分析集群	9
2.2.2. 快速创建数据湖集群	11
2.2.3. 快速创建云搜索集群	14
2.2.4. 快速创建数据服务集群	17
2.3. 管理数据连接	22
2.3.1. 配置数据连接	22
2.3.2. 配置 Ranger 数据连接	25
2.3.3. 配置 Hive 数据连接	29
3. 管理集群	31
3.1. 登录集群	31
3.1.1. 翼 MR 集群节点简介	31
3.1.2. 登录集群节点	32
3.2. 集群概览	35
3.2.1. 集群列表简介	35
3.2.2. 查看集群状态	35
3.2.3. 查看集群基本信息	36
3.2.4. 查看集群监控指标	38
3.2.5. 管理组件和主机监控	41
3.3. 健康检查	44
3.3.1. 健康检查	44
3.3.2. 查看翼 MR 服务操作日志	46
3.4. 节点管理	48
3.4.1. 管理节点操作	48
3.4.2. 配置升级	49
3.4.3. 远程连接	51
3.4.4. 绑定/解绑弹性 IP	52
3.5. 组件管理	54
3.5.1. 对象管理简介	54
3.5.2. 查看配置	55
3.5.3. 管理服务操作	55
3.5.4. 配置服务参数	56
3.5.5. 同步服务配置	57
3.5.6. 管理角色实例操作	58

3.5.7.	启动及停止集群.....	58
3.5.8.	下载集群的配置数据.....	59
3.5.9.	支持滚动重启.....	59
3.6.	作业管理.....	61
3.6.1.	运行 MapReduce 作业.....	61
3.6.2.	运行 SparkSubmit 作业.....	62
3.6.3.	运行 HiveSQL 作业.....	63
3.6.4.	运行 SparkSQL 作业.....	64
3.6.5.	运行 Kafka 作业.....	65
3.6.6.	查看主机和组件日志.....	66
3.7.	告警管理.....	66
4.	使用翼 MR 客户端.....	67
4.1.	安装客户端.....	67
4.2.	各组件客户端使用实践.....	70
4.2.1.	使用 YARN 客户端.....	70
4.2.2.	使用 Flume 客户端.....	71
4.2.3.	使用 HBase 客户端.....	74
4.2.4.	使用 HDFS 客户端.....	76
4.2.5.	使用 Hive 客户端.....	77
4.2.6.	使用 Kafka 客户端.....	79
4.2.7.	使用 Trino 客户端.....	81
4.2.8.	使用 ZooKeeper 客户端.....	82
4.2.9.	使用 Kyuubi 客户端.....	83
4.2.10.	使用 Spark 客户端.....	84
5.	访问翼 MR 集群上托管的开源组件 WEB 页面.....	85
5.1.	开源组件 WEB 站点.....	85
5.2.	开源组件端口列表.....	86
5.3.	通过弹性公网 IP 访问.....	98
6.	MANAGER 操作指导.....	100
6.1.	从这里开始.....	100
6.1.1.	Manager 入门指导.....	100
6.1.2.	查询 Manager 版本号.....	102
6.2.	首页概述.....	102
6.3.	资源概览.....	103
6.4.	集群服务.....	104
6.4.1.	集群服务管理概述.....	104
6.4.2.	集群服务详情页面概述.....	104
6.4.3.	下载客户端.....	105
6.4.4.	启动、停止所有集群服务.....	106
6.4.5.	启动、停止单集群服务.....	107
6.4.6.	滚动重启单个集群服务.....	109
6.4.7.	启动、停止、重启实例.....	112
6.4.8.	滚动停止、滚动重启实例.....	114
6.4.9.	实例置维护/取消维护.....	117

6.4.10.	查看集群服务配置.....	119
6.4.11.	新增配置组.....	119
6.4.12.	查看配置组详情.....	120
6.4.13.	删除配置组.....	121
6.4.14.	新增配置.....	122
6.4.15.	查看配置详情.....	129
6.4.16.	编辑配置.....	130
6.4.17.	配置项对比.....	132
6.4.18.	配置文件预览.....	135
6.4.19.	下载配置文件.....	137
6.4.20.	配置文件历史对比.....	138
6.4.21.	查看配置历史版本.....	140
6.4.22.	删除配置.....	141
6.4.23.	配置同步.....	142
6.4.24.	查看配置状态.....	144
6.4.25.	重新同步.....	146
6.4.26.	配置回滚.....	146
6.4.27.	查看运维操作历史.....	148
6.4.28.	查看告警历史.....	149
6.4.29.	查看集群服务大屏.....	150
6.5.	主机.....	150
6.5.1.	主机管理页面概述.....	150
6.5.2.	主机详情页面概述.....	151
6.5.3.	主机置维护、取消维护.....	151
6.5.4.	查看主机状态.....	153
6.5.5.	查看主机实例列表.....	154
6.5.6.	查看主机告警历史.....	154
6.6.	租户与资源.....	155
6.7.	监控与告警.....	158
6.7.1.	指标查询-指标查询概述.....	158
6.7.2.	告警历史-告警历史概述.....	160
6.8.	运维与配置.....	161
6.8.1.	流水线历史-流水线历史概述.....	161
6.8.2.	流水线历史-查看流水线运行详情.....	161
6.8.3.	配置管理-配置管理概述.....	162
6.8.4.	配置管理-新增配置组.....	162
6.8.5.	配置管理-查看配置组详情.....	163
6.8.6.	配置管理-删除配置组.....	164
6.8.7.	配置管理-新增配置.....	165
6.8.8.	配置管理-查看配置详情.....	166
6.8.9.	配置管理-编辑配置.....	167
6.8.10.	配置管理-配置项对比.....	169
6.8.11.	配置管理-配置文件预览.....	172
6.8.12.	配置管理-下载配置文件.....	174
6.8.13.	配置管理-配置文件历史对比.....	175
6.8.14.	配置管理-查看配置历史版本.....	176

6.8.15.	配置管理-删除配置.....	177
6.8.16.	配置管理-配置同步.....	178
6.8.17.	配置管理-查看配置状态.....	181
6.8.18.	配置管理-重新同步.....	182
6.8.19.	配置管理-配置回滚.....	183
6.8.20.	记登历史-配置历史概述.....	184
6.8.21.	配置历史-查询配置历史.....	185
6.8.22.	配置历史-查看配置详情.....	186
6.8.23.	配置历史-查看配置状态.....	187
6.8.24.	配置历史-查看配置对比.....	188
6.8.25.	配置历史-重新同步.....	189
6.8.26.	配置同步历史-配置同步历史概述.....	190
6.8.27.	配置同步历史-查看配置同步状态.....	190
7.	安全认证原理和认证机制.....	192
8.	高危操作一览表.....	198
9.	附录.....	210

1. 准备用户

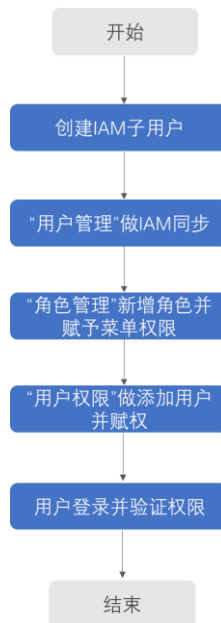
1.1. 创建翼 MapReduce 操作用户

本章节主要介绍如何创建翼 MapReduce 服务的操作用户。

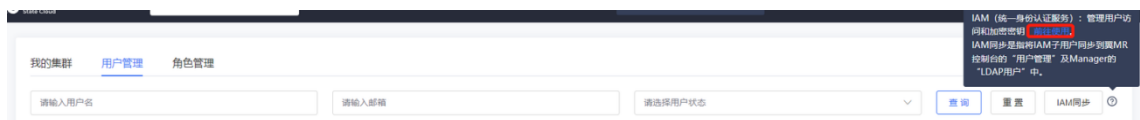
如果您需要对您所拥有的翼 MapReduce 服务进行精细的权限管理，您可以使用统一身份认证服务（Identity and Access Management，简称 IAM），创建 IAM 子用户，并在翼 MR 控制台给予用户赋予不同的角色权限。

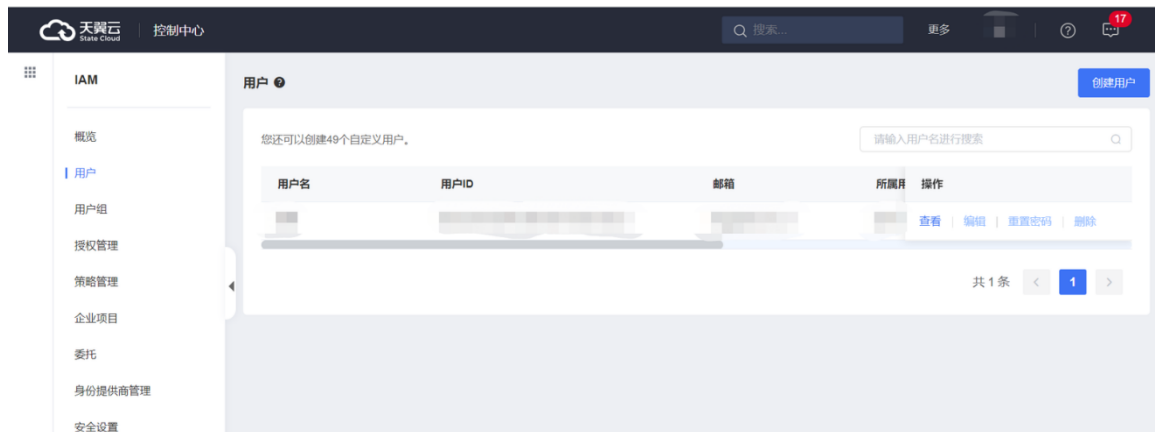
如果云帐号已经能满足您的要求，不需要创建独立的 IAM 用户，您可以跳过本章节，不影响您使用翼 MR 服务的其它功能。

本章节为您介绍对用户授权的方法，操作流程如下图所示。



(1) 创建 IAM 子用户。进入翼 MR 控制台的“用户管理”页面，点击“IAM 同步”按钮后面“？”中的“前往使用”，即可进入到 IAM 用户页面。点击页面右上角的“创建用户”，按照页面提示即可完成子用户创建。

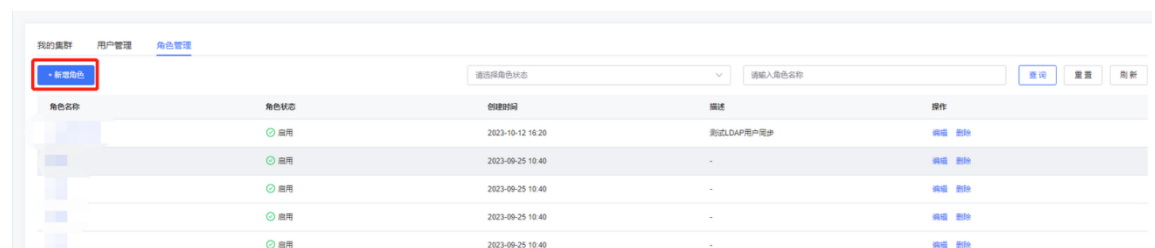




(2) 在翼 MR 控制台的“用户管理”页面点击 IAM 同步，将 IAM 子用户信息同步到翼 MR 控制台和 Manager 的“LDAP 用户”中。



(3) 在翼 MR 控制台的“角色管理”页面新增角色，并给角色赋予相应的菜单功能权限。



(4) 在翼 MR 控制台的“用户权限”页面添加用户，并给用户赋予对应的角色权限。



(5) IAM 子用户登录翼 MR 控制台，访问至授权集群，验证权限。

说明：

默认情况下，管理员创建的 IAM 用户没有任何权限，需要将其赋予角色权限后，才能使得用户获得对应集群的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

1.2. IAM 用户同步

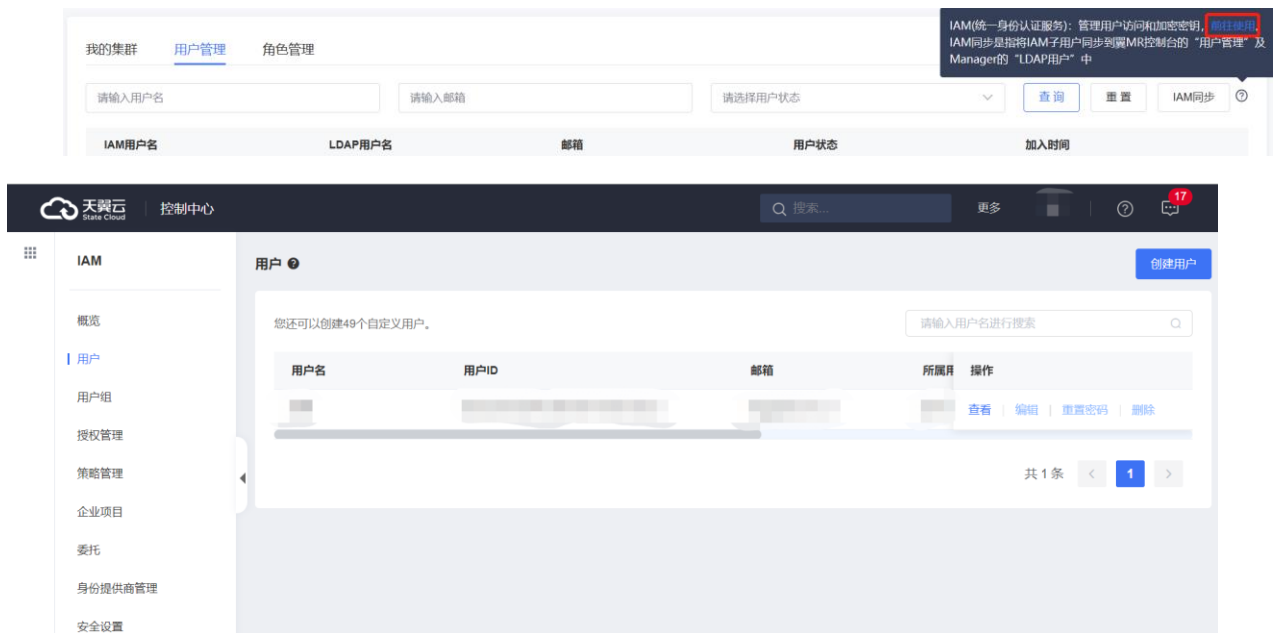
本章节主要介绍翼 MapReduce 服务的 IAM 用户同步。

IAM 用户同步是指将创建的 IAM 子用户同步至翼 MapReduce 系统中：一个是将 IAM 用户信息同步到“用户管理”服务，包括 IAM 用户名、邮箱、用户状态、加入时间等信息；另一个是将 IAM 用

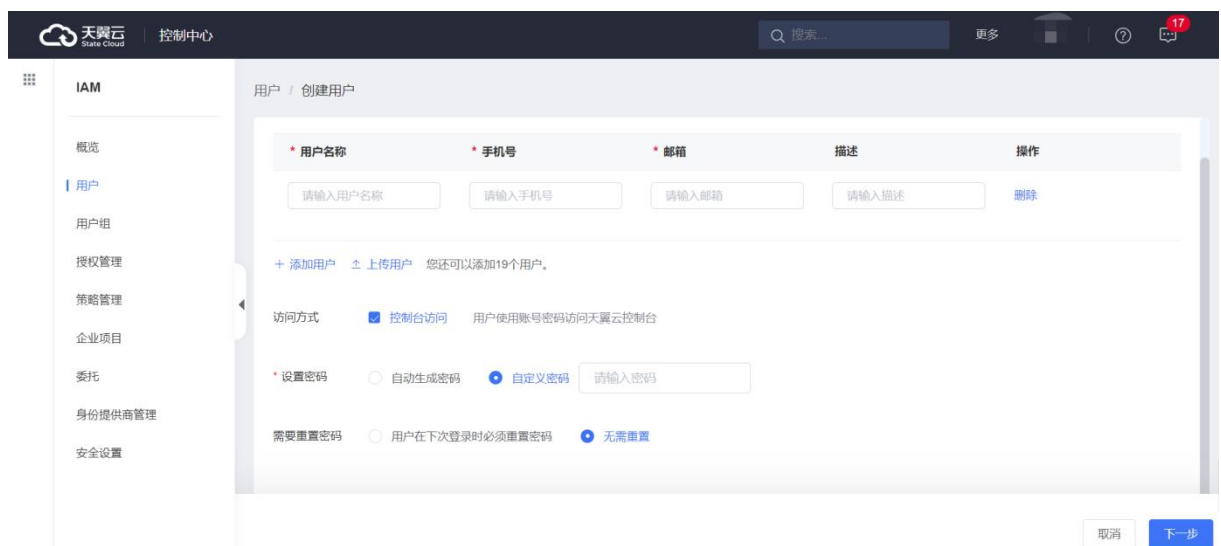
户同步到翼 MR Manager 中的“LDAP 用户”服务中，LDAP 用户名以“emr_邮箱前缀”表示，但是第二种情况的前提是：该集群下需要完成 OpenLDAP 集群服务的部署、配置以及启动，否则无法同步到“LDAP 用户”服务中。

操作步骤

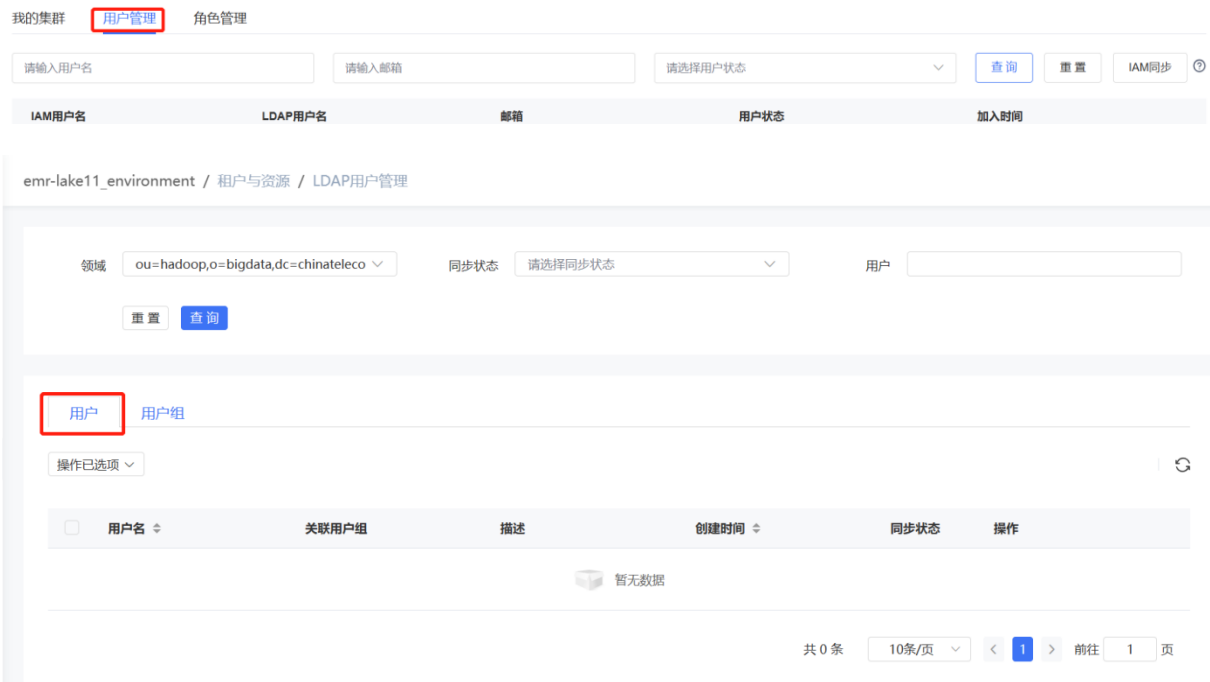
(1) 点击“用户管理”服务页面的“IAM 同步”的“?”按钮，点击“前往使用”，进入 IAM 的“主子账号及授权中心”页面。



(2) 点击“创建用户”，根据页面提示输入信息，完成 IAM 用户创建。



(3) 创建完成后，返回翼 MR 控制台的“用户管理”服务页面，点击“IAM 同步”按钮，就能将 IAM 用户信息同步到“用户管理”服务页面，同时也将 IAM 用户同步到翼 MR Manager 中的“LDAP 用户”服务中。



2. 配置集群

2.1. 创建方式简介

本章节主要介绍创建翼 MapReduce 服务的方式。

- **快速创建数据分析集群：**快速创建数据分析集群可以在短时间内快速完成数据分析集群的创建，使用开源的 MPP 架构的 OLAP 分析引擎，支持亚秒级的数据查询和多表 join，帮助您更加快捷的进行数据分析工作。
- **快速创建数据湖集群：**快速创建数据湖集群可以在短时间内快速完成数据湖集群的创建，数据湖集群提供更高效、灵活的集群管理，更快地运行大数据的计算引擎，更好地提供数据分析能力，帮助您更加方便快捷的对数据湖的计算、存储资源进行构建和管理。
- **快速创建云搜索集群：**快速创建云搜索集群可以在短时间内快速完成云搜索集群的创建，通过分布式计算和存储系统，为结构化/非结构化数据提供低成本、高性能及可靠性的检索、分析服务能力，帮助您更方便的检索大规模数据集。
- **快速创建数据服务集群：**快速创建数据服务集群可以在短时间内快速完成数据服务集群的创建，数据服务集群提供高度可靠性、可扩展性的数据服务以满足您的数据需求。

2.2. 快速创建集群

2.2.1. 快速创建数据分析集群

本章节主要介绍如何快速创建一个数据分析集群。

数据分析集群使用 Apache Doris，Apache Doris 是开源的 MPP 架构的 OLAP 分析引擎，支持亚秒级的数据查询和多表 join。在创建数据分析集群前，需要先创建虚拟私有云。

快速创建数据分析集群步骤

- (1) 登录翼 MapReduce 控制台。
- (2) 在“我的集群”页面，单击“创建集群”，进入“创建集群”页面。
- (3) 在创建集群页面进行集群相关配置。
- (4) 软件配置页面如下图所示，参数说明如下：



软件配置 | 2 硬件配置 | 3 基础配置 | 4 确认订单

* 区域集群①: 华东1

* 可用区②: 可用区1 | 可用区2 | 可用区3

* 业务场景: 数据湖 | **数据分析** | 数据服务 | 云搜索

数据分析
—Apache Doris开源的MPP架构的OLAP分析引擎，支持亚秒级的数据查询和多表join。

* 产品版本: **MR-2.12.0**

* 服务高可用: **开启**
翼MR默认启用服务高可用，master节点将分布在不同的底层硬件上以降低故障风险。

* 可选服务③: Doris(1.2.6)

[下一步: 硬件配置](#)

- a. 区域集群：集群节点 ECS 实例所在的物理位置，根据需要选择区域及可用区，也可以使用默认值。
 - b. 业务场景：此处选择“数据分析”。
 - c. 产品版本：选择使用的产品版本，默认值即可。
 - d. 服务高可用：翼 MapReduce 默认启用服务高可用且不可关闭高可用模式。
 - e. 可选服务：由可选组件和必选组件组成，根据业务场景而定。您可根据自身业务场景对可选组件进行选择。数据分析集群场景下默认只有 Doris 一个必选组件。
 - f. 软件配置选择完成后，点击“下一步”进入硬件配置页面。
- (5) 硬件配置页面如下图所示，参数说明如下：



✓ 软件配置 | **2 硬件配置** | 3 基础配置 | 4 确认订单

* 计费模式①: **包年包月**

* 购买时长②: 1 个月
1个月 | 2个月 | 3个月 | 4个月 | 5个月 | 6个月 | 7个月

* 虚拟私有云③: 请选择VPC | [创建虚拟私有云](#)
请在当前地域下创建专有网络。

* 子网: 请选择子网 | [创建子网](#)

* 安全组④: 请选择安全组 | [创建安全组](#)
创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务。安全组创建请参考[如何使用自定义安全组创建翼MR集群](#)。

* 节点组
(节点组类型、选型配置、云盘参数和性能)

节点组类型/节点组名称	规格	系统盘/数据盘	节点数量
* MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
* CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	1

- a. 计费模式：默认为包年包月（目前在公测期间，可购买时长会随着公测截止日期变化，具

体以订购页面实际可购买时长为准）。

- 购买时长：拖动按钮可以选择购买时长。
- 虚拟私有云：不同虚拟私有云（VPC）网络之间的逻辑彻底隔离。根据需要选择自己的 VPC。如果目前没有 VPC 可以点击“创建虚拟私有云”跳转到虚拟私有云页面创建。
- 子网：选择虚拟私有云后，子网可以根据需要进行选择。
- 安全组：设置集群内 ECS 实例的网络访问控制。创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务，详情请参见常见问题“[如何使用自定义安全组创建翼 MR 集群](#)”。
- 节点组：根据您自身需要选择集群节点规格及数量，包括对节点组类型、选项配置、云盘参数和性能的选择。
- 硬件配置选择完成后，左下角会计算当前配置需要费用。点击右下角“下一步”按钮进入基础配置页面。

(6) 基础配置页面如下图所示，参数说明如下：

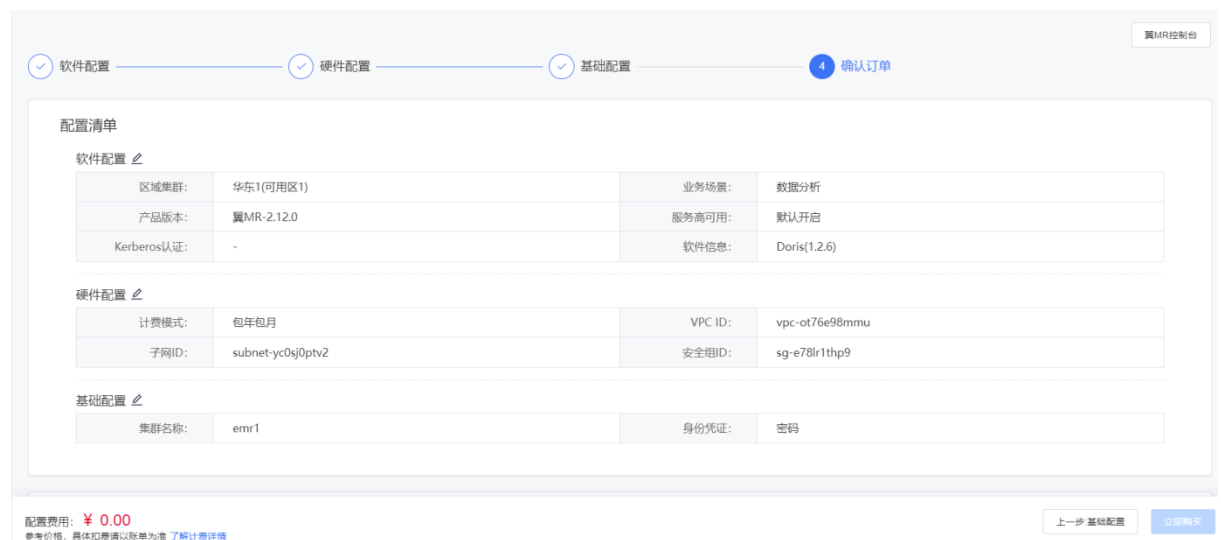


基础配置页面包含以下配置项：

- 集群名称：请输入集群名称(示例: emr1)
- 身份凭证：密码
- 登录账号：root
- 登录密码：请输入密码
- 确认密码：请确认密码

- 集群名称：集群名称可由大写字母、小写字母、数字及特殊符号组成，最大 32 字符。
- 身份凭证：默认密码。
- 登录账号：默认为 root。
- 登录密码：根据内容提示，创建符合要求的登录密码。
- 确认密码：与登录密码相同。
- 基础配置填写完成后，点击右下角“下一步”按钮，进入确认订单页面。

(7) 确认订单页面如下图所示，配置清单说明如下：



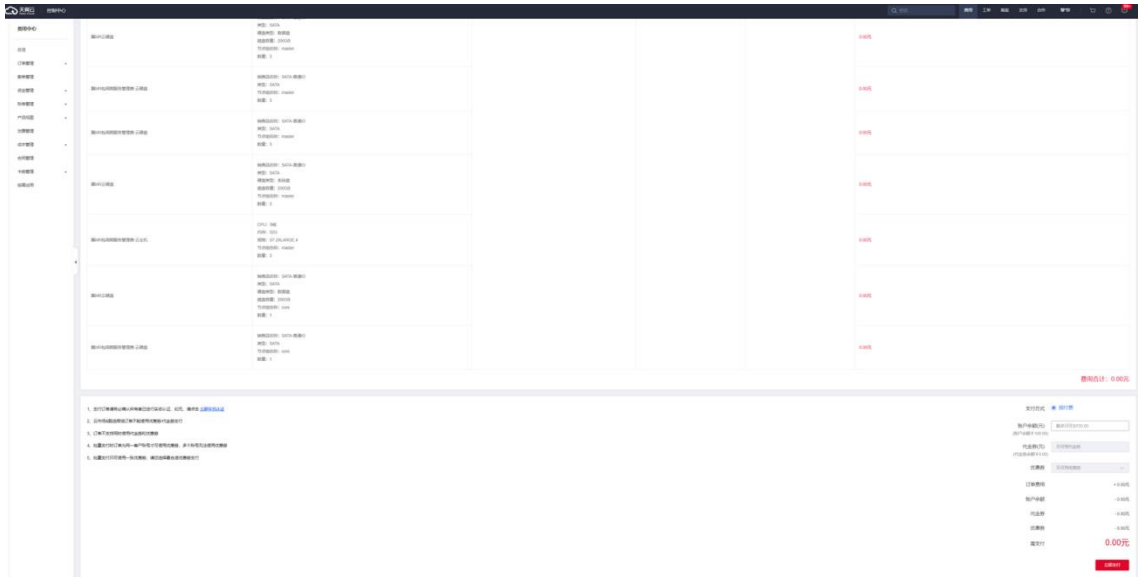
确认订单页面显示配置清单：

软件配置		硬件配置		基础配置	
区域集群:	华东1(可用区1)	业务场景:	数据分析	集群名称:	emr1
产品版本:	翼MR-2.12.0	服务高可用:	默认开启	身份凭证:	密码
Kerberos认证:	-	软件信息:	Doris(1.2.6)		
计费模式:	包年包月	VPC ID:	vpc-ot76e98mmu		
子网ID:	subnet-yc0sj0ptv2	安全组ID:	sg-e78lr1thp9		

配置费用: ¥ 0.00
参考价格，具体扣费请以账单为准 [了解计费详情](#)

上一步 基础配置 立即购买

- 软件配置：确认内容是否有误，点击“编辑”按钮，即可回退到“软件配置”面，修改集群信息。
- 硬件配置：确认内容是否有误，点击“编辑”按钮，即可回退到“硬件配置”页面，修改集群信息。
- 基础配置：确认内容是否有误，点击“编辑”按钮，即可回退到“基础配置”页面，修改集群信息。
- 节点组清单：确认内容是否有误，点击“编辑”按钮，即可回退到“硬件配置”页面，修改集群信息。
- 勾选“我已阅读并同意相关协议”，“立即购买”按钮变成高亮显示，点击进行支付，购买后即可租用当前配置的集群。



2.2.2. 快速创建数据湖集群

本章节主要介绍如何快速创建一个数据湖集群。

数据湖场景提供更高效、灵活的管理集群，更快地运行大数据的计算引擎，更好地提供数据分析能力。在创建数据湖集群前，需要先创建虚拟私有云。

快速创建数据湖集群步骤

- 登录翼 MapReduce 控制台。
- 单击“创建集群”，进入“创建集群”页面。
- 在创建集群页面进行集群相关配置。
- 软件配置页面如下图所示，参数说明如下：

1 软件配置
2 硬件配置
3 基础配置
4 确认订单

区域集群
华东1

可用区
可用区1
可用区2
可用区3

业务场景
数据湖
数据分析
数据服务
云搜索

数据湖
提供高效、灵活的管理集群，更快地运行大数据的计算引擎，更好地提供数据分析能力。

产品版本
EMR-2.12.0

服务高可用
开启

可选服务
OpenLDAP(2.4.50)
Kerberos(1.18.2)
ZooKeeper(3.7.1)
HDFS(3.3.3)
YARN(3.3.3)
Hive(3.1.2)
Spark(3.2.2)
HBase(2.4.12)
Trino(406)
Kyuubi(1.6.0)
Doris(1.2.6)
Ranger(2.2.0)
Kafka(2.8.1)

Hive元数据
使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

hive-site.xml
数据库URL

请按照以下示例格式输入：jdbc:mysql://192.168.0.1:13049/hive_db_name
只能通过内网地址进行数据库连接，部署Hive的节点和数据库需要在同一个VPC下且网络互通。

数据库用户名

数据库密码

Kerberos认证
开启

下一步 硬件配置

- 区域集群：集群节点 ECS 实例所在的物理位置，根据需要选择区域及可用区，也可以使用默认值。
- 业务场景：此处选择“数据湖”。
- 产品版本：选择使用的产品版本，默认值即可。
- 服务高可用：翼 MapReduce 默认启用服务高可用且不可关闭高可用模式。
- 可选服务：由可选组件和必选组件组成，根据业务场景而定。您可根据自身业务场景对可选组件进行选择。
- Hive 元数据：如您选择“使用已有 CT-RDS for MySQL”需要配置此项。配置内容参考如下图所示对应提示。

* Hive元数据 ☒ 使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

- Hive 元数据配置：默认选择“使用已有 CT-RDS for MySQL”。选择“使用已有 CT-RDS for MySQL”需要提前开通关系型数据库 MySQL，可点击链接前往开通，并填写如下信息。

Hive元数据配置

hive-site.xml

数据库URL

请按照以下示例格式输入：jdbc:mysql://192.168.0.1:13049/hive_db_name
只能通过内网地址进行数据库连接，部署Hive的节点和数据库需要在同一个VPC下且网络互通。

数据库用户名

数据库密码

- h. Kerberos 身份认证：集群中的软件以 Kerberos 安全模式启动，此选项不可关闭。
 - i. 软件配置选择完成后，点击“下一步”进入硬件配置页面。
- (5) 硬件配置页面如下图所示，参数说明如下：

软件配置

2 硬件配置

3 基础配置

4 确认订单

计费模式

包年包月

购买时长

1 个月

虚拟私有云

请选择VPC

创建虚拟私有云

子网

请选择子网

安全组

请选择安全组

创建安全组

节点组

节点组类型/节点组名称	规格	系统盘/数据盘	节点数量
MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	1

- a. 计费模式：默认为包年包月（目前在公测期间，可购买时长会随着公测截止日期变化，具体以订购页面实际可购买时长为准）。
 - b. 购买时长：拖动按钮可以选择购买时长。
 - c. 虚拟私有云：不同虚拟私有云（VPC）网络之间的逻辑彻底隔离。根据需要选择自己的VPC。如果目前没有VPC可以点击“创建虚拟私有云”跳转到虚拟私有云页面创建。
 - d. 子网：选择虚拟私有云后，子网可以根据需要进行选择。
 - e. 安全组：设置集群内ECS实例的网络访问控制。创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务，详情请参见常见问题“[如何使用自定义安全组创建翼MR集群](#)”。
 - f. 节点组：根据您自身需要选择集群节点规格及数量，包括对节点组类型、选项配置、云盘参数和性能的选择，可根据需要对task节点进行增加/删除。
 - g. 硬件配置选择完成后，左下角会计算当前配置需要费用。点击右下角“下一步”按钮进入基础配置页面。
- (6) 基础配置页面如下图所示，参数说明如下：



软件配置 硬件配置 3 基础配置

* 集群名称: 请输入集群名称(示例: emr1)

* 身份凭证: 密码

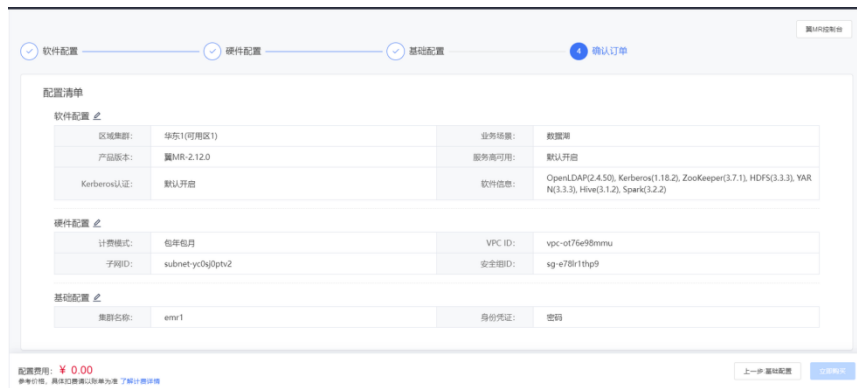
* 登录账号: root

* 登录密码: 请输入密码

* 确认密码: 请确认密码

- 集群名称: 集群名称可由大写字母、小写字母、数字及特殊符号组成, 最大 32 字符。
- 身份凭证: 默认密码。
- 登录账号: 默认为 root。
- 登录密码: 根据内容提示, 创建符合要求的登录密码。
- 确认密码: 与登录密码相同。
- 基础配置填写完成后, 点击右下角“下一步”按钮, 进入确认订单页面。

(7) 确认订单页面如下图所示, 配置清单说明如下:



软件配置 硬件配置 基础配置 4 确认订单

配置清单

软件配置		硬件配置		基础配置	
区域集群:	华东1(可用区1)	计费模式:	包年包月	集群名称:	emr1
产品版本:	EMR-2.12.0	子网ID:	subnet-yc9j0pvt2	身份凭证:	密码
Kerberos认证:	默认开启	VPC ID:	vpc-ot76e98mmu		
业务场景:	数据湖	安全组ID:	sg-e78r1thg9		
服务端可用:	默认开启				
软件信息:	OpenLDAP(2.4.50), Kerberos(1.18.2), ZooKeeper(3.7.1), HDFS(3.3.3), YARN(3.3.3), Hive(3.1.2), Spark(3.2.2)				

配置费用: ¥ 0.00

立即购买

- 软件配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“软件配置”面, 修改集群信息。
- 硬件配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“硬件配置”页面, 修改集群信息。
- 基础配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“基础配置”页面, 修改集群信息。
- 节点组清单: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“硬件配置”页面, 修改集群信息。
- 勾选“我已阅读并同意相关协议”, “立即购买”按钮变成高亮显示, 进入支付页面, 购买后即可租用当前配置的集群。

2.2.3. 快速创建云搜索集群

本章节主要介绍如何快速创建一个云搜索集群。

云搜索集群为结构化/非结构化数据提供低成本、高性能及可靠性的检索、分析服务能力。在创建云搜索集群前, 需要先创建虚拟私有云。

快速创建云搜索集群步骤

- (1) 登录翼 MapReduce 控制台。

- (2) 单击“创建集群”，进入“创建集群”页面。
- (3) 在创建集群页面进行集群相关配置。
- (4) 软件配置页面如下图所示，参数说明如下：



The screenshot shows the 'Software Configuration' step in a cluster creation wizard. It includes settings for Region (华东1), Availability Zone (可用区1), Business Scenario (云搜索), Product Version (默认 2.12.0), Service High Availability (默认), and Optional Services (ElasticSearch(7.10.2), Kibana(7.10.2)). A 'Next Step' button is at the bottom right.

- a. 区域集群：集群节点 ECS 实例所在的物理位置，根据需要选择区域及可用区，也可以使用默认值。
- b. 业务场景：此处选择“云搜索”。
- c. 产品版本：选择使用的产品版本，默认值即可。
- d. 服务高可用：翼 MapReduce 默认启用服务高可用模式，且无法关闭。
- e. 可选服务：由可选组件和必选组件组成，根据业务场景而定。您可根据自身业务场景对可选组件进行选择。云搜索集群场景下默认可选组件如下图所示，灰色表示必选。



This section shows the 'Optional Services' configuration. It includes a dropdown for 'Optional Services' and two buttons for 'ElasticSearch(7.10.2)' and 'Kibana(7.10.2)', both of which are highlighted in grey to indicate they are required.

- f. 软件配置选择完成后，点击“下一步”进入硬件配置页面。
- (5) 硬件配置页面如下图所示，参数说明如下：



The screenshot shows the 'Hardware Configuration' step. It includes settings for Billing Mode (包年包月), Purchase Duration (1 month), Virtual Private Cloud (选择VPC), Subnet (选择子网), and Security Group (选择安全组). Below these is a table for Node Groups.

节点组	节点组类型/节点名称	规格	系统盘/数据盘	节点数量
MASTER	MASTER	通用型 (x7.2xlarge.4)	通用型SSD云硬盘 200GB*1	3
	master	8 vCPU, 32GB	通用型SSD云硬盘 200GB*2	
CORE	CORE	通用型 (x7.2xlarge.4)	通用型SSD云硬盘 200GB*1	1
	core	8 vCPU, 32GB	通用型SSD云硬盘 200GB*2	

- a. 计费模式：默认为包年包月（目前在公测期间，可购买时长会随着公测截止日期变化，具体以订购页面实际可购买时长为准）。
- b. 购买时长：拖动按钮可以选择购买时长。
- c. 虚拟私有云：不同虚拟私有云（VPC）网络之间的逻辑彻底隔离。根据需要选择自己的VPC。如果目前没有VPC可以点击“创建虚拟私有云”跳转到虚拟私有云页面创建。
- d. 子网：选择虚拟私有云后，子网可以根据需要进行选择。

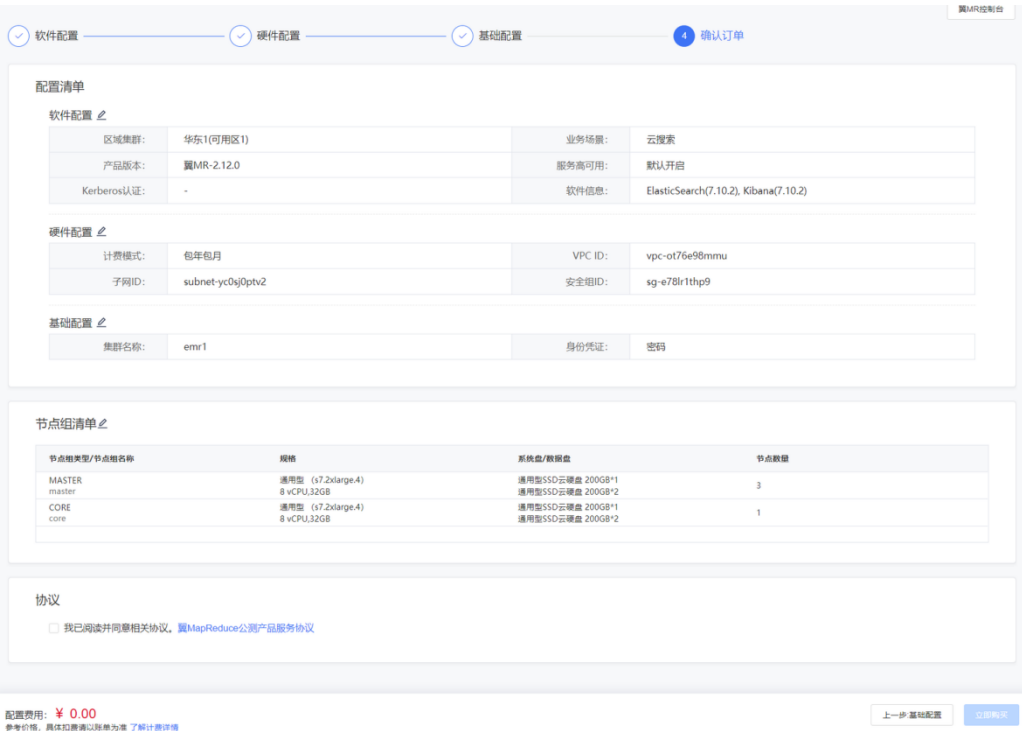
- e. 安全组：设置集群内 ECS 实例的网络访问控制。创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务，详情请参见常见问题“[如何使用自定义安全组创建翼 MR 集群](#)”。
 - f. 节点组：根据您自身需要选择集群节点规格及数量，包括对节点组类型、选项配置、云盘参数和性能的选择。
 - g. 硬件配置选择完成后，左下角会计算当前配置需要费用。点击右下角“下一步”按钮进入基础配置页面。
- (6) 基础配置页面如下图所示，参数说明如下：



基础配置页面包含以下配置项：

- 集群名称：请输入集群名称(示例: emr1)
- 身份凭证：密码
- 登录账号：root
- 登录密码：请输入密码
- 确认密码：请确认密码

- a. 集群名称：集群名称可由大写字母、小写字母、数字及特殊符号组成，最大 32 字符。
 - b. 身份凭证：默认密码。
 - c. 登录账号：默认为 root。
 - d. 登录密码：根据内容提示，创建符合要求的登录密码。
 - e. 确认密码：与登录密码相同。
 - f. 基础配置填写完成后，点击右下角“下一步”按钮，进入确认订单页面。
- (7) 确认订单页面如下图所示，配置清单说明如下：



确认订单页面包含以下配置清单：

配置清单

软件配置		硬件配置		基础配置	
区域集群：	华东1(可用区1)	业务场景：	云搜索	集群名称：	emr1
产品版本：	翼MR-2.12.0	服务高可用：	默认开启	身份凭证：	密码
Kerberos认证：	-	软件信息：	ElasticSearch(7.10.2), Kibana(7.10.2)		
计费模式：	包年包月	VPC ID：	vpc-ot76e98mmu		
子网ID：	subnet-yc0j0ptv2	安全组ID：	sg-e78lr1thp9		

节点组清单

节点组类型/节点组名称	规格	系统盘/数据盘	节点数量
MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	1

协议

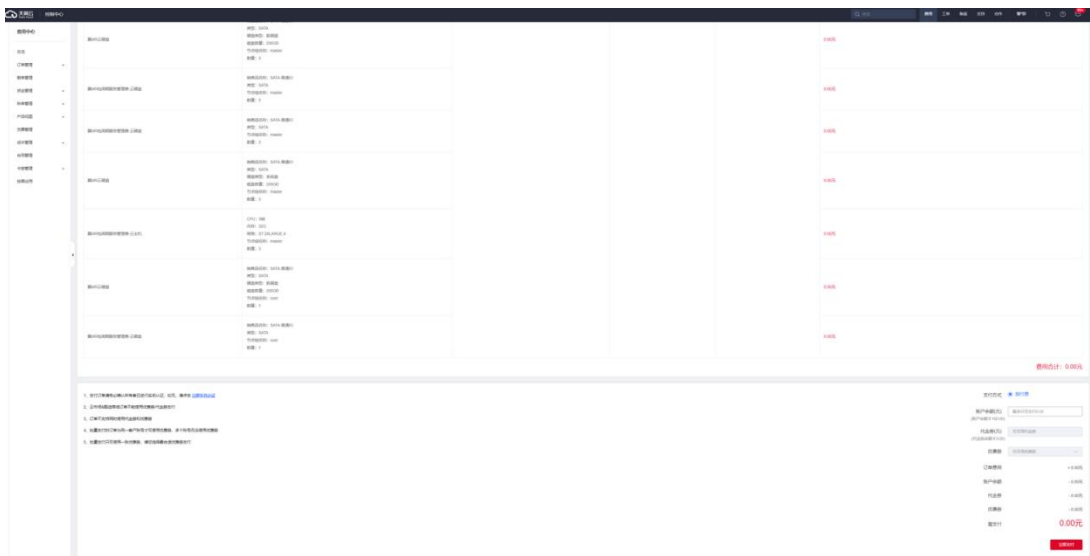
☐ 我已阅读并同意相关协议。 [翼MR-2.12.0产品服务协议](#)

配置费用：¥ 0.00
参考价格，具体扣费请以账单为准。 [了解详情](#)

上一步：基础配置 立即购买

- a. 软件配置：确认内容是否有误，点击“编辑”按钮，即可回退到“软件配置”面，修改集群信息。

- 硬件配置：确认内容是否有误，点击“编辑”按钮，即可回退到“硬件配置”页面，修改集群信息。
- 基础配置：确认内容是否有误，点击“编辑”按钮，即可回退到“基础配置”页面，修改集群信息。
- 节点组清单：确认内容是否有误，点击“编辑”按钮，即可回退到“硬件配置”页面，修改集群信息。
- 勾选“我已阅读并同意相关协议”，“立即购买”按钮变成高亮显示，进入支付页面，购买后即可租用当前配置的集群。



2.2.4. 快速创建数据服务集群

本章节主要介绍如何快速创建一个数据服务集群。

数据服务集群提供更灵活、可靠、高效的数据服务。在创建数据服务集群前，需要先创建虚拟私有云。

快速创建数据服务集群步骤

- 登录翼 MapReduce 控制台。
- 单击“创建集群”，进入“创建集群”页面。
- 在创建集群页面进行集群相关配置。
- 软件配置页面如下图所示，参数说明如下：



The screenshot shows the 'Software Configuration' (软件配置) step in a four-step process. It includes dropdowns for 'Region Group' (区域集群) and 'Available Zone' (可用区), buttons for 'Business Scenario' (业务场景) like 'Data Service' (数据服务), and a 'Product Version' (产品版本) dropdown. A 'Service High Availability' (服务高可用) toggle is set to 'On'. A 'Selectable Services' (可选服务) section shows various components like OpenLDAP, Kerberos, ZooKeeper, HDFS, YARN, HBase, and Ranger, with Kerberos and Ranger highlighted. A 'Kerberos Authentication' (Kerberos认证) toggle is also set to 'On'.

- 区域集群：**集群节点 ECS 实例所在的物理位置，根据需要选择区域及可用区，也可以使用默认值。
- 业务场景：**此处选择“数据服务”。
- 产品版本：**选择使用的产品版本，默认值即可。
- 服务高可用：**翼 MapReduce 默认启用服务高可用且不可关闭高可用模式。
- 可选服务：**由可选组件和必选组件组成，根据业务场景而定。您可根据自身业务场景对可选组件进行选择。数据服务集群场景下默认可选组件如下图所示，灰色表示必选，高亮的 Ranger 组件为可选项，您可根据业务场景进行选择。



A horizontal bar showing selectable services: OpenLDAP(2.4.50), Kerberos(1.18.2), ZooKeeper(3.7.1), HDFS(3.3.3), YARN(3.3.3), HBase(2.4.12), and Ranger(2.2.0). The Ranger component is highlighted in blue, indicating it is a selectable option.

当您选择添加 Ranger 组件后，需要为 Ranger 组件配置元数据，配置参数内容可参考相应的提示内容。



This is another screenshot of the 'Software Configuration' page, identical to the one above, showing the same settings and selectable services.

- Kerberos 身份认证：**集群中的软件以 Kerberos 安全模式启动，此选项不可关闭。
- 软件配置选择完成后，点击“下一步”进入硬件配置页面。

(5) 硬件配置页面如下图所示，参数说明如下：

软件配置

硬件配置

基础配置

确认订单

计费模式

包年包月

购买时长

0 1 个月

1个月 2个月 3个月 4个月 5个月 6个月 7个月

虚拟私有云

请选择VPC

创建虚拟私有云

请在当前地域下创建专有网络。

子网

请选择子网

安全组

请选择安全组

创建安全组

创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务。安全组创建请参考[如何使用自定义安全组创建MR集群](#)。

节点组

(节点组类型、选项配置、云盘参数和性能)

节点组类型/节点组名称	规格	系统盘/数据盘	节点数量
MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3

配置费用: ¥ 0.00

参考价格，具体扣费请以账单为准 [了解计费详情](#)

上一步 软件配置

下一步 基础配置

- 计费模式：默认为包年包月（目前在公测期间，可购买时长会随着公测截止日期变化，具体以订购页面实际可购买时长为准）。
- 购买时长：拖动按钮可以选择购买时长。
- 虚拟私有云：不同虚拟私有云（VPC）网络之间的逻辑彻底隔离。根据需要选择自己的VPC。如果目前没有VPC可以点击“创建虚拟私有云”跳转到虚拟私有云页面创建。

虚拟私有云

请选择VPC

创建虚拟私有云

请在当前地域下创建专有网络。

- 子网：选择虚拟私有云后，子网可以根据需要进行选择。
- 安全组：设置集群内ECS实例的网络访问控制。创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务，详情请参见常见问题“[如何使用自定义安全组创建翼MR集群](#)”。

安全组

请选择安全组

创建安全组

创建安全组时，出、入方向规则中的端口需要放开，授权策略不能选择“拒绝”，否则会导致不能部署集群和拉起服务。安全组创建请参考[如何使用自定义安全组创建MR集群](#)。

- 节点组：根据您自身需要选择集群节点规格及数量，包括对节点组类型、选项配置、云盘参数和性能的选择。

* 节点组 (节点组类型、选型配置、云盘参数和性能)

节点组类型/节点组名称	规格	系统盘/数据盘	节点数量
MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3

实例规格

通用型 通用计算增强型 内存优化型

请选择	产品规格	CPU	内存(GB)
<input checked="" type="radio"/>	s7.2xlarge.4	8	32
<input type="radio"/>	s7.2xlarge.4.selfComponent	8	32
<input type="radio"/>	s7.2xlarge.4.selfManager	8	32
<input type="radio"/>	s7.4xlarge.2	16	32
<input type="radio"/>	s7.4xlarge.4	16	64
<input type="radio"/>	s7.8xlarge.2	32	64
<input type="radio"/>	s7.8xlarge.4	32	128

反亲和性云主机组 ☒ 开启

MASTER节点组默认开启反亲和性云主机组。

系统盘 通用型SSD云硬盘 200 GB*1块

数据盘 通用型SSD云硬盘 200 GB*2块

实例数量 3 台

CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
-----------	------------------------------------	--	---

g. 硬件配置选择完成后，左下角会计算当前配置需要费用。点击右下角“下一步”按钮进入基础配置页面。

(6) 基础配置页面如下图所示，参数说明如下：

软件配置 硬件配置 3 基础配置

* 集群名称: 请输入集群名称(示例: emr1)

* 身份凭证: 密码

* 登录账号: root

* 登录密码: 请输入密码

* 确认密码: 请确认密码

- a. 集群名称: 集群名称可由大写字母、小写字母、数字及特殊符号组成，最大 32 字符。
- b. 身份凭证: 默认密码。
- c. 登录账号: 默认为 root。
- d. 登录密码: 根据内容提示，创建符合要求的登录密码。
- e. 确认密码: 与登录密码相同。
- f. 基础配置填写完成后，点击右下角“下一步”按钮，进入确认订单页面。

(7) 确认订单页面如下图所示，配置清单说明如下：

翼MR控制台

✓ 软件配置
✓ 硬件配置
✓ 基础配置
4 确认订单

配置清单

软件配置

区域集群:	华东1(可用区1)	业务场景:	数据服务
产品版本:	翼MR-2.12.0	服务商可用:	默认开启
Kerberos认证:	默认开启	软件信息:	OpenLDAP(2.4.50), Kerberos(1.18.2), ZooKeeper(3.7.1), HDFS(3.3.3), YARN(3.3.3), HBase(2.4.12)

硬件配置

计费模式:	包年包月	VPC ID:	vpc-ot76e98mmu
子网ID:	subnet-yc0sj0ptv2	安全组ID:	sg-e78lr1thp9

基础配置

集群名称:	emr1
身份凭证:	密码

节点组清单

节点组类型/节点名称	规格	系统盘/数据盘	节点数量
MASTER master	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3
CORE core	通用型 (s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB*1 通用型SSD云硬盘 200GB*2	3

协议

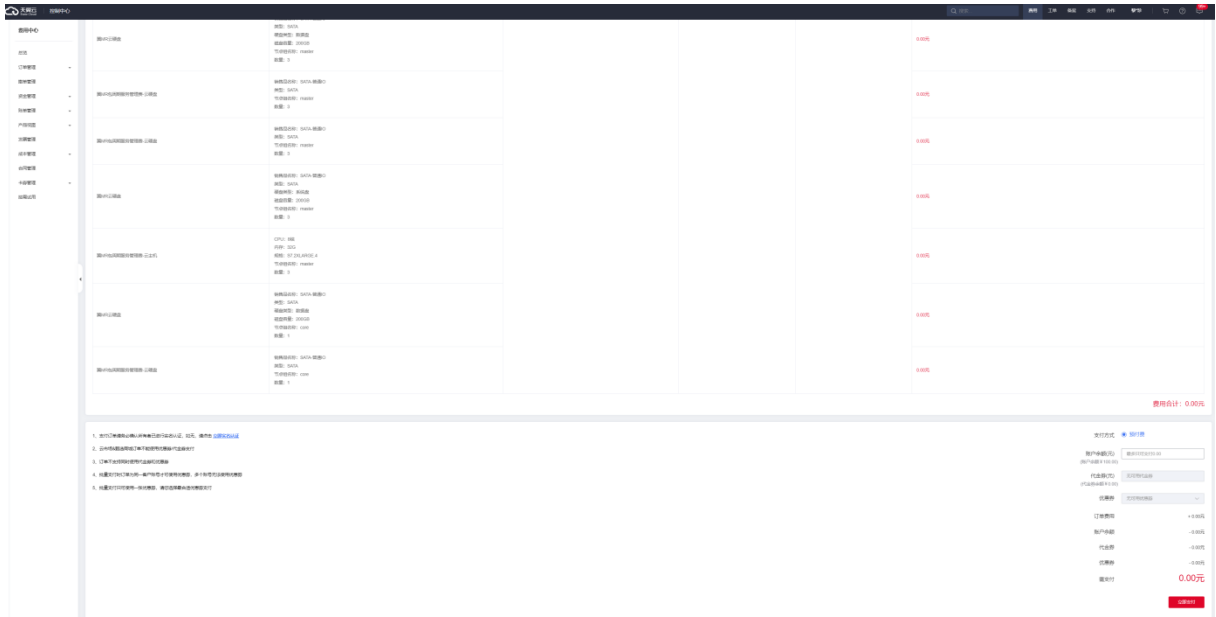
☐ 我已阅读并同意相关协议。 [翼MapReduce公测产品服务协议](#)

配置费用: **¥ 0.00**

参考价格, 具体计费请以账单为准 [了解计费详情](#)

上一步 基础配置
立即购买

- a. 软件配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“软件配置”面, 修改集群信息。
- b. 硬件配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“硬件配置”页面, 修改集群信息。
- c. 基础配置: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“基础配置”页面, 修改集群信息。
- d. 节点组清单: 确认内容是否有误, 点击“编辑”按钮, 即可回退到“硬件配置”页面, 修改集群信息。
- e. 勾选“我已阅读并同意相关协议”, “立即购买”按钮变成高亮显示, 进入支付页面, 购买后即可租用当前配置的集群。



2.3. 管理数据连接

2.3.1. 配置数据连接

本章节主要介绍数据连接的创建过程以及参数说明。

配置数据连接

翼 MR 的数据连接是用来管理集群中组件使用的外部源连接，如 Hive 和 Ranger 的元数据使用外部的关系型数据库，可以通过数据连接来关联 Hive 或者 Ranger 组件实现。

本地元数据：元数据存储于集群内的本地中，当集群删除时元数据同时被删除，如需保存元数据，需提前前往数据库手动保存元数据。

数据连接：可选择关联与当前集群同一虚拟私有云和子网的 RDS 服务中的 PostgresDB 或 MySQL 数据库，元数据将存储于关联的数据库中，不会随当前集群的删除而删除，多个翼 MR 集群可共享同一份元数据。

数据连接前置操作

- (1) 登录 RDS 管理控制台。
- (2) 选择“实例管理”，选择翼 MR 数据连接使用的 RDS 实例所在行的操作列的“更多 > 登录”，以 root 用户登录该实例。

<input type="checkbox"/>	实例名称/ID	实例备注	实例类型	数据库引擎/版本	运行状态	计费模式	内网地址	企业项目	操作
<input type="checkbox"/>	c427d2b176b14611aa4218b7d15b9a2fin01	--	单机 2 vCPUs 8 GB	MySQL 5.7.37	正常	按需计费 2022/06/14 14:24...		default	查看监控指标 转包周期 更多 >

- (3) 在实例“首页”即可单击“新建数据库”创建新的数据库。

- (4) 在页面顶部选择“账号管理 > 用户管理”。
- (5) 单击“新建用户”，创建一个非 root 用户。
- (6) 在页面顶部选择“SQL 操作 > SQL 查询”，在“库名”处切换对应数据库，然后执行如下 SQL 命令为该数据库用户进行赋权，其中 {db_name} 与 {db_user} 为翼 MR 待连接的数据库名和新建的用户名。

```
grant SELECT, INSERT on mysql.* to '${db_user}'@'%' with grant option;  
  
grant all privileges on ${db_name}.* to '${db_user}'@'%' with grant option;  
  
grant reload on *.* to '${db_user}'@'%' with grant option;  
  
flush privileges;
```

说明

当用户选择的数据连接为“RDS 服务 MySQL 数据库”时，请确保使用的数据库用户为 root 用户。如果为非 root 用户，请参考步骤 4-步骤 6 操作。

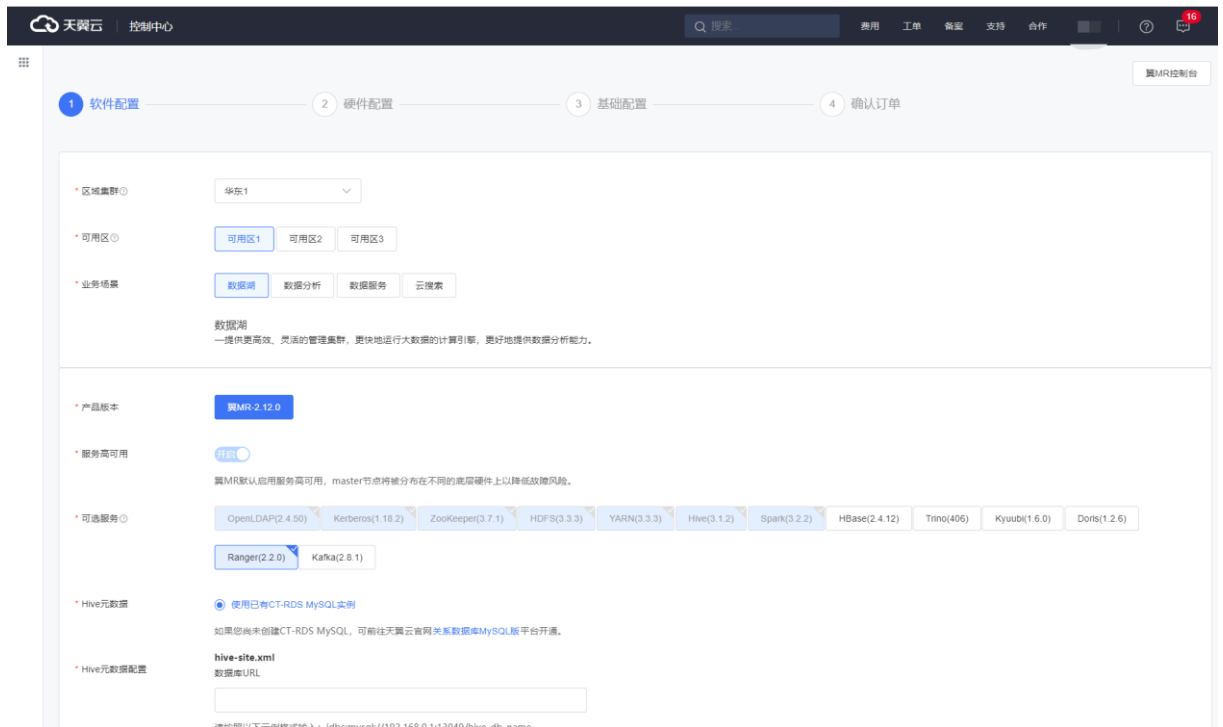
创建集群时配置数据连接

- (1) 登录翼 MR 管理控制台。
- (2) 进入“我的集群”列表页，点击列表上方“创建集群”按钮；或者通过天翼云官网选择“产品—大数据—翼 MapReduce”，进入订购页面，资源池选择“华东 1”。





(3) 进入集群创建参数配置页面“软件配置”。



(4) 在软件配置中，填写外部数据源的具体参数。创建集群时配置所需的数据连接参数详见下图：

• 可选服务②

OpenLDAP(2.4.50)

Kerberos(1.18.2)

ZooKeeper(3.7.1)

HDFS(3.3.3)

YARN(3.3.3)

Hive(3.1.2)

Spark(3.2.2)

HBase(2.4.12)

Trino(406)

Kyubid(1.6.0)

Doris(1.2.6)

Ranger(2.2.0)

Kafka(2.8.1)

• Hive元数据

• 使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

• Hive元数据配置

hive-site.xml

数据库URL

请按照以下示例格式输入：jdbc:mysql://192.168.0.1:13049/hive_db_name
只能通过内网地址进行数据库连接，部署Hive的节点和数据库需要在同一个VPC下且网络互通。

数据库用户名

数据库密码

• Ranger元数据

• 使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

• Ranger元数据配置

install.properties

数据库主机

请按照以下示例格式输入：192.168.0.1
只能通过内网地址进行数据库连接，部署Ranger的节点和数据库需要在同一个VPC下且网络互通

数据库端口

13049

输入内容提示：默认13049

数据库名称

请输入

数据库用户名

请输入

数据库密码

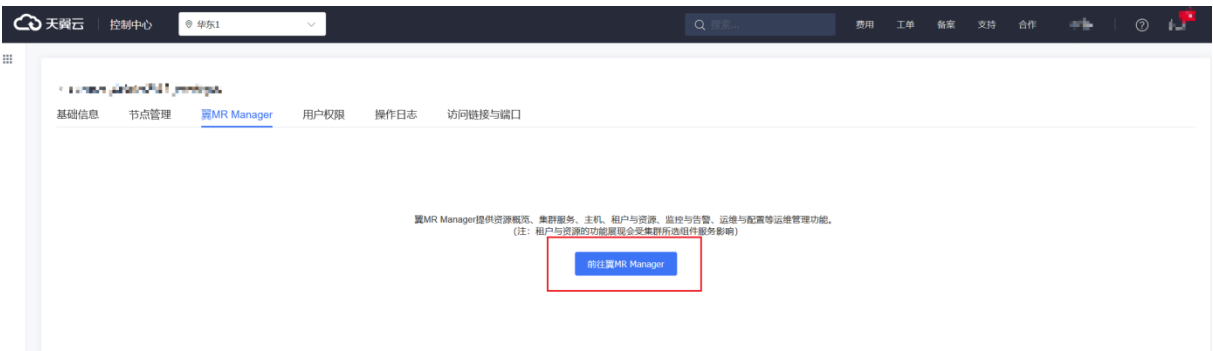
2.3.2. 配置 Ranger 数据连接

本章节主要介绍如何配置 Ranger 数据连接以及配置参数详情。

本章节主要是指导用户将现有集群的 Ranger 元数据切换为 RDS 数据库中存储的元数据。该操作可以使多个翼 MR 集群共用同一份元数据，且元数据不随集群的删除而删除。也能够避免集群迁移时 Ranger 元数据的迁移。

查看 Ranger 服务页面详情

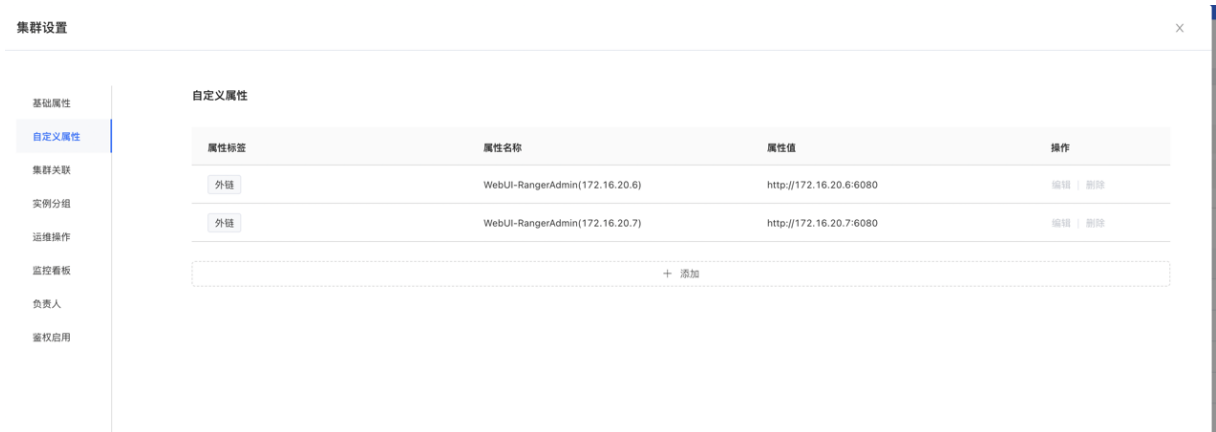
(1) 登录翼 MR Manager 页面。



(2) 选择“集群 > 服务 > Ranger”，进入 Ranger 服务概览页面。



(3) 单击“集群服务详情”页面中的“集群设置”弹框下的“自定义属性”，先进行所在集群的 Ranger Web UI 外链的配置。

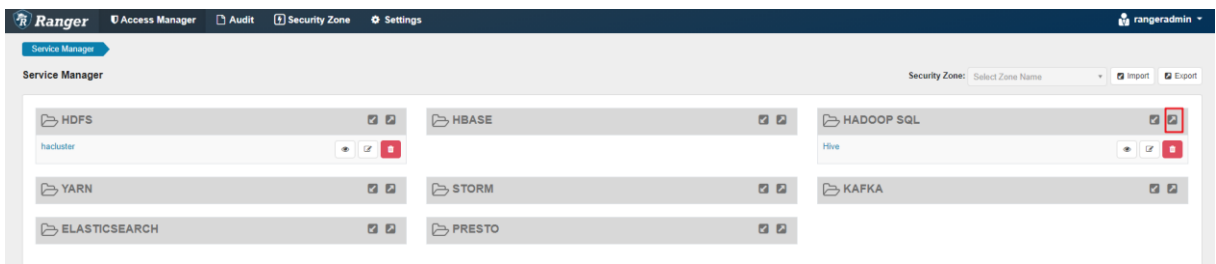


(4) 配置好外链之后，关闭集群设置弹框即返回至集群服务页面，可看到配置好的外链直接显示在集群详情的 tab 页，点击对应的链接即可跳转 Ranger 开源 UI 界面。

admin 用户在 Ranger 中的用户类型为“User”，如需查看所有管理页面，可点击右上角用户名后，选择“Log Out”，退出当前用户。

使用 rangeradmin 用户（默认密码为 Rangeradmin@123）或者其他具有 Ranger 管理员权限用户重新登录。

导出鉴权策略详见下图：

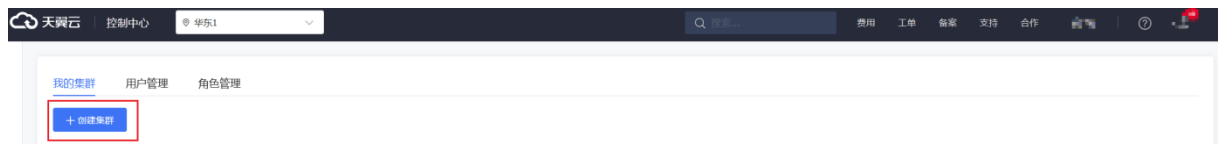


(5) 单击“Export”，导出后会生成本地的 json 格式的策略文件。导出 Hive 鉴权策略详见下图：

翼 MR 集群配置数据连接

- (1) 登录翼 MR 控制台。
- (2) 单击集群名称进入集群详情页面。

(3) 我的集群，点击我的集群列表页面左上方“创建集群”，进入数据连接配置界面。



(4) 单击“创建集群”，并配置相关参数。

- a. 组件名称：Ranger
- b. 模块类型：Ranger 元数据
- c. 连接类型：RDS 服务 MySQL 数据库
- d. 连接实例：请选择已创建的到 RDS 服务 MySQL 数据库的实例，如需创建新的数据连接，请参考[配置数据连接](#)。配置参数参考下图：

(5) Ranger 参数配置完成后确认订单信息，连接测试则在订单完成集群部署时自动进行验证，无需用户手动操作，后台验证结果会返回至集群服务列表状态栏，具体状态可登录翼 MR Manager 进行查看。

(6) 登录翼 MR Manager 页面。

(7) 选择“集群 > 服务 > Ranger”，进入 Ranger 服务概览页面。

(8) 单击“运维操作”> 重启集群服务。

重启服务会造成业务中断，滚动重启可以尽量减少或者不影响业务运行。

重启 Ranger 组件会影响所有受 Ranger 控制组件的权限，可能影响业务的正常运行，请在集群空闲或业务量较少时时执行重启。重启 Ranger 组件前，Ranger 中的策略依然生效。

重启服务操作如图：



(9) 登录 Ranger WebUI 界面，单击 Hive 组件对应行的导入按钮。



(10) 配置导入参数。



(2) 点击集群列表左上方“创建集群”。



(3) 在“创建集群”订购页面中，用户可选择对应的组件，点击元数据按钮，在元数据下方会显示 Hive 元数据的具体配置。

* Hive元数据 ☒ 使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

hive-site.xml

* Hive元数据配置

数据库URL

请按照以下示例格式输入：jdbc:mysql://192.168.0.1:13049/hive_db_name
只能通过内网地址进行数据库连接，部署Hive的节点和数据库需要在同一个VPC下且网络互通。

数据库用户名

数据库密码

(4) 元数据群存储可使用已有 CT-RDS MySQL，若用户 CT-RDS MySQL，可点击链接前往天翼云官网进行开通。

* Hive元数据 ☒ 使用已有CT-RDS MySQL实例

如果您尚未创建CT-RDS MySQL，可前往天翼云官网[关系数据库MySQL版](#)平台开通。

配置 Hive 数据连接：

参数	说明
数据库 URL	请按照以下示例格式输入：jdbc:mysql://192.168.0.1:13049/hive_db_name，只能通过内网地址进行数据库连接，部署 Hive 的节点和数据库需要在同一个 VPC 且网络互通。
数据库用户名	用户的数据库用户名。

数据库密码	用户的数据库密码。
-------	-----------

(5) 配置完成确认订单之后会在集群部署期间验证数据连接的连通性，可在集群列表中查看部署状态。

3. 管理集群

3.1. 登录集群

3.1.1. 翼 MR 集群节点简介

远程登录

翼 MR 集群节点支持用户远程登录，远程登录包含界面登录和 SSH 登录两种方式：

- 界面登录：直接通过翼 MR 管理控制台->节点管理页面中提供的远程登录功能，登录到集群中各个节点的 Linux 界面。
- SSH 登录：仅适用于 Linux 弹性云主机。您可以使用远程登录工具（例如 PuTTY、XSELL），登录弹性云主机。此登录方式，需要前置依赖该弹性云主机绑定弹性 IP 地址，详细步骤请参见[绑定/解绑弹性 IP](#)。

说明

当您使用密码方式访问集群节点，详细步骤请参见[登录集群节点](#)（SSH 密码方式）。

节点类型和功能

翼 MR 集群中每个节点即为一台弹性云主机，节点类型及节点功能如下表所示：

节点类型	功能
Master 节点	翼 MR 集群管理节点，负责管理和监控集群。在翼 MR 管理控制台选择“我的集群”，选中一个运行中的集群并单击集群名，进入集群信息页面。在“节点管理”中查看节点信息。节点可以通过点击“远程登录”选择界面的 VNC 方式登录，也可以通过 SSH 方式登录。
Core 节点	翼 MR 集群工作节点，负责处理和分析数据，并存储过程数据。

3.1.2. 登录集群节点

本章节主要介绍如何登录翼 MR 集群节点。

本章节介绍如何使用翼 MR 管理控制台上提供的节点远程连接（VNC 方式）和如何使用密码方式（SSH 方式）登录翼 MR 集群中的节点，远程连接主要用于紧急运维场景，远程登录弹性云主机进行相关维护操作。其他场景下，优先推荐用户使用 SSH 方式登录集群节点。

说明

如果需要使用 SSH 方式登录集群节点，需要在集群的安全组规则中手动添加入方向规则：其中源地址为“客户端公网 IPV4 地址/32”，端口为 22，具体请参见“帮助中心>虚拟私有云>安全组>[添加安全组规则](#)”。

登录弹性云主机（VNC 方式）

- (1) 登录翼 MapReduce 服务管理控制台。
- (2) 选择“集群列表 > 我的集群”，选中一个运行中的集群并单击集群名，进入集群基本信息页面。
- (3) 选择“节点管理”，单击展开节点组，显示节点列表。
- (4) 在目标节点的右侧“操作”列中，单击“远程连接”。
- (5) 在 VNC 方式新标签页 Web 页面中，可以看到服务器的命令行界面，类似于 Linux 服务器登录模式，输入 root 账号，并输入密码。

说明

创建集群时，登录方式选择了“密码”。此时，你需要输入的密码就是创建集群时设置的密码。

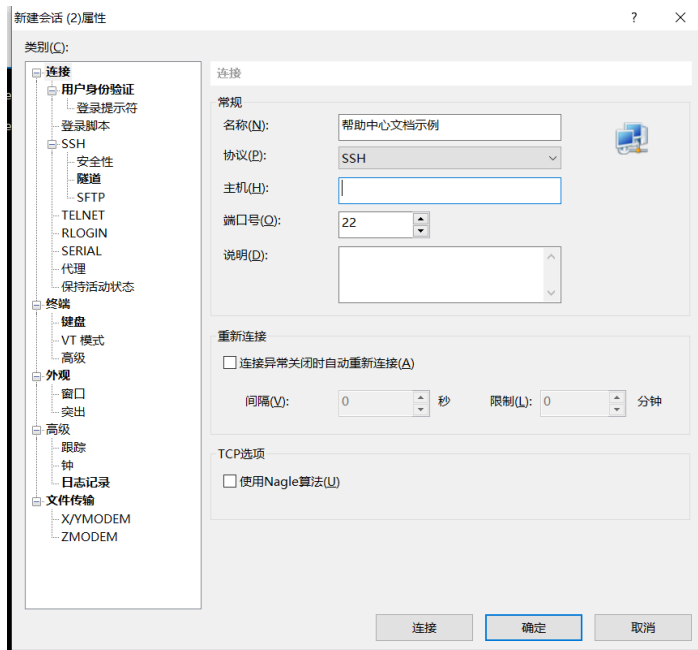
登录弹性云主机（SSH 密钥方式）

本地使用 Windows 操作系统

如果您本地使用 Windows 操作系统登录 Linux 弹性云主机，可以按照下面方式登录弹性云主机。下面步骤以 Xshell 为例。

- (1) 登录翼 MR 服务管理控制台。
- (2) 选择“集群列表 > 我的集群”，选中一个运行中的集群并单击集群名，进入集群基本信息页面。
- (3) 选择“节点管理”，单击展开节点组，显示节点列表。
- (4) 在目标节点的右侧“操作”列中，单击“更多”，选择“绑定弹性 IP”，可以为该节点选择一个已存在且未绑定使用的弹性公网 IP 或者可以单击“+创建弹性公网 IP”跳转至弹性 IP 控制台完成弹性 IP 购买后，再完成弹性公网 IP 绑定操作，若已绑定弹性公网 IP 请跳过该步骤。
- (5) 运行 Xshell。
- (6) 选择“新建会话”。
- (7) 名称：选择你认为合适的会话名称，用于方便管理。
- (8) 协议：使用默认的“SSH”。
- (9) 主机：输入弹性云主机所绑定的弹性公网 IP。

(10) 端口号：使用默认的“22”，详见下图： 单击“Session”。



(11) 单击“连接”，按照提示步骤，分别完成登录用户（默认为 root）、密码的输入并保存，即可完成远程登录操作。

本地使用 Linux 操作系统

如果您本地使用 Linux 操作系统登录 Linux 弹性云主机，可以按照下面方式登录。

执行如下命令，登录弹性云主机。

```
ssh 默认用户名@弹性公网 IP
```

假设 Linux 弹性云主机的默认用户名是 root，弹性公网 IP 为 123.123.123.123，则命令如下：

```
ssh root@123.123.123.123
```

3.2. 集群概览

3.2.1. 集群列表简介

本章节主要介绍集群列表。

登录翼 MapReduce 控制台后，可查看账号下拥有的集群列表和当前的集群状态。

集群状态

集群包含的状态如下表所示。

状态	说明
启动中	集群正在创建中，可以查看具体的进程。
运行中	集群创建成功且集群中组件状态处于正常运行中，集群正在正常运行。
异常终止	集群启动失败或因为异常原因而终止，可以查看失败原因。
释放中	表示集群正在退订。
已终止	集群已经到期 15 天以上或者集群已经完成退订。
已冻结	已经到期的集群，资源会保留 15 天，此时状态为已冻结。

3.2.2. 查看集群状态

本章节主要介绍查看集群状态。

集群创建完成后，登录翼 MapReduce 控制台后，可查看账号下所有集群列表和集群状态。

在“我的集群”中，显示所有您创建过的集群，默认页面上最多显示 10 条集群信息，当您集群数量较多时，可以修改页面右下角当前页面最多显示的集群条数，也可以翻页进行查看。

集群列表参数说明：

参数	参数说明
集群名称	用户自己定义的集群名称。
集群类型	用户创建集群时选择的集群类型。
付费类型	默认为包年包月。
状态	集群当前的状态信息。 启动中：集群正在创建中，通过“查看进程”查看具体的进程。 运行中：集群创建成功且集群中组件状态处于正常运行中，集群正在正常运行。 异常终止：集群启动失败或因为异常原因而终止，通过“失败原因”查看终

	<p>止原因。</p> <p>释放中：集群正在退订。</p> <p>已终止：集群已经到期 15 天以上或者集群已经完成退订。</p> <p>已冻结：已经到期的集群，资源会保留 15 天，此时状态为已冻结。</p>
创建时间	集群创建的时间。

3.2.3. 查看集群基本信息

本章节主要介绍查看集群基本信息。

登录翼 MapReduce 控制台后，进入集群列表页面，点击需要查看集群的名称，即进入集群基本信息界面。

集群基本信息参数主要包括集群信息、软件信息和网络硬件。

集群信息

参数	说明
集群名称	集群的名称，单击“编辑”可对集群名称进行修改。
创建时间	集群创建的时间。
运行时间	集群运行的时间。
到期时间	集群到期时间。
高可用	默认开启。
付费类型	默认包年包月。
Kerberos 认证	根据业务场景需要，默认开启。

软件信息

参数	说明
产品版本	集群使用的翼 MapReduce 版本。
软件信息	集群创建时使用的组件及其版本号。单击需要查看的组件名称，可以进入该组件详情信息。
业务场景	集群创建时选择的业务场景。

网络硬件

参数	说明
区域集群	集群所在的区域。
硬件配置	集群所有节点 MASTER、CORE、TASK（如有）的硬件信息。
VPC ID	集群所使用 VPC 的 ID，单击  复制 VPC ID 内容，单击 VPC ID 内容可以跳转并查看该 VPC 详细信息。
安全组 ID	集群所使用安全组的 ID，单击  复制安全组 ID 内容，单击安全组 ID 内容可以跳转并查看安全组详细信息。

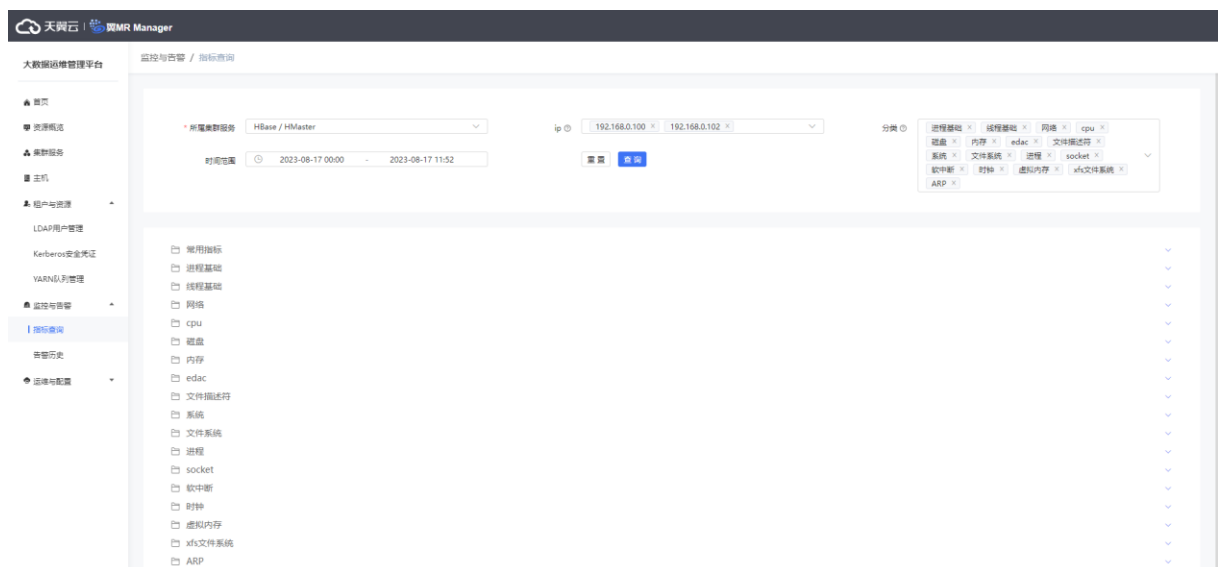
3.2.4. 查看集群监控指标

本章节主要介绍查看集群监控指标。

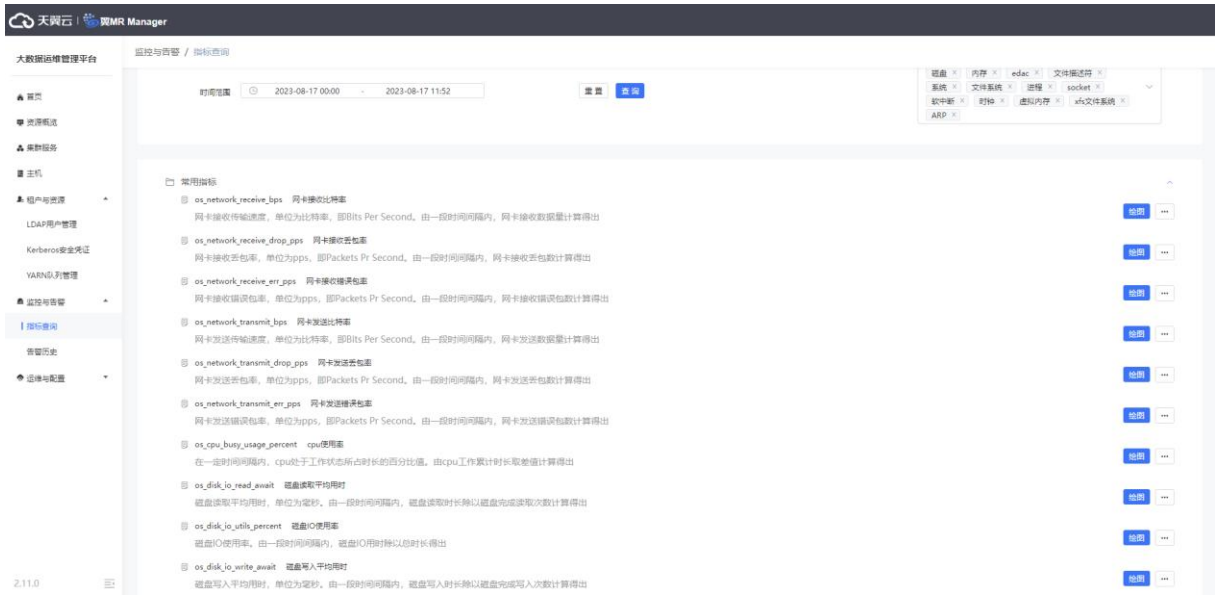
操作步骤

- (1) 登录翼 MR 控制台。
- (2) 选择“我的集群”，单击集群名称进入集群详情页面。
- (3) 在集群详情页，选择“翼 MR Manager”点击“前往翼 MR Manager”进入运维管理平台。
- (4) 选择“监控与告警 > 指标查询”，即可查看集群监控指标。

- (5) 在所属“集群服务”选择需要查看的集群。
- (6) 在“IP”处选择部署第 5 步已选择的对应集群节点 IP，可复选。
- (7) 在“分类处”选择需要设置的指标分类。默认勾选常用指标分类，可自主再添加其他选项，可供选择的选项如下：进程基础
 - a. 线程基础
 - b. 网络
 - c. CPU
 - d. 磁盘
 - e. 内存
 - f. edac
 - g. 文件描述符
 - h. 系统
 - i. 文件系统
 - j. 进程
 - k. socket
 - l. 软中断
 - m. 时钟
 - n. 虚拟内存
 - o. xfs 文件系统
 - p. ARP
- (8) 在“时间范围”选择需要查看监控数据的时间段。可供选择的选项如下：昨天
 - a. 上周
 - b. 本周
 - c. 上月
 - d. 本月
 - e. 去年
 - f. 自定义：在时间范围内自行选择需要查看的时间。
- (9) 点击“查询”并显示所选指标分类。



- (10) 选择任意一种指标分类，点击右侧的展开，显示某一指标分类下的所有指标。

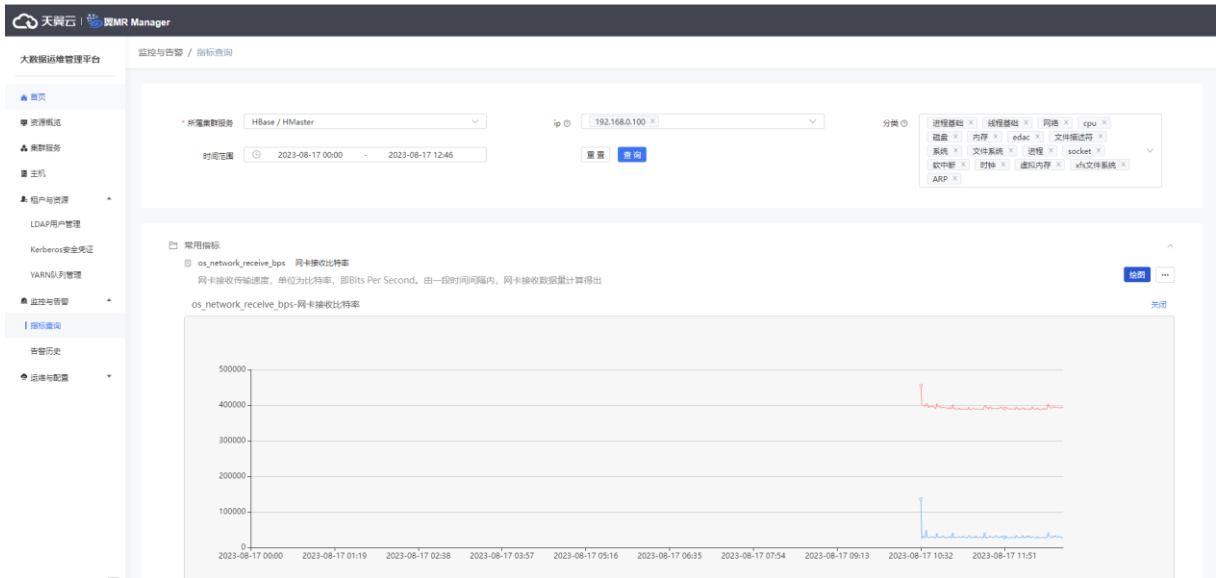
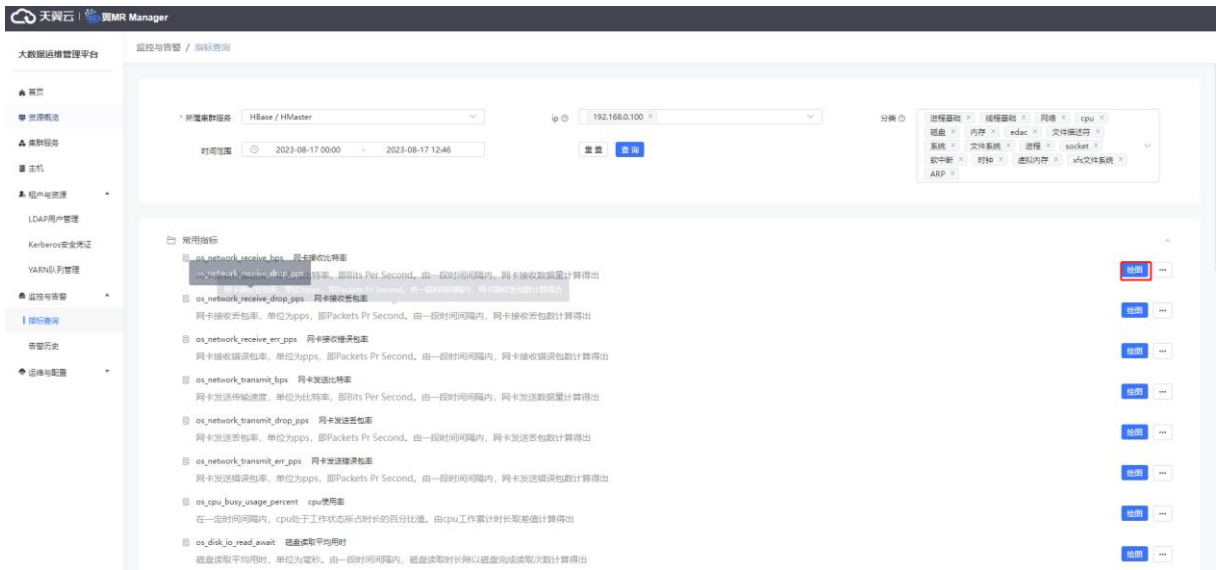


说明

常用指标类别共计 18 项，包含如下指标：

- 网卡接收比特率
- 网卡接收丢包率
- 网卡接收错误包率
- 网卡发送比特率
- 网卡发送丢包率
- 网卡发送错误包率
- CPU 使用率
- 磁盘读取平均用时
- 磁盘 IO 使用率
- 磁盘写入平均用时
- 磁盘 IO 操作平均用时
- 内存空闲率
- 1 分钟平均负载
- 5 分钟平均负载
- 15 分钟平均负载
- 磁盘 inode 空闲率
- 磁盘容量空闲率
- Blocked 状态进程数

11. 选择任一类别下的某一指标，点击右侧的“绘图”即查看对应集群的监控指标。



3.2.5. 管理组件和主机监控

本章节主要介绍管理组件和主机监控的操作。

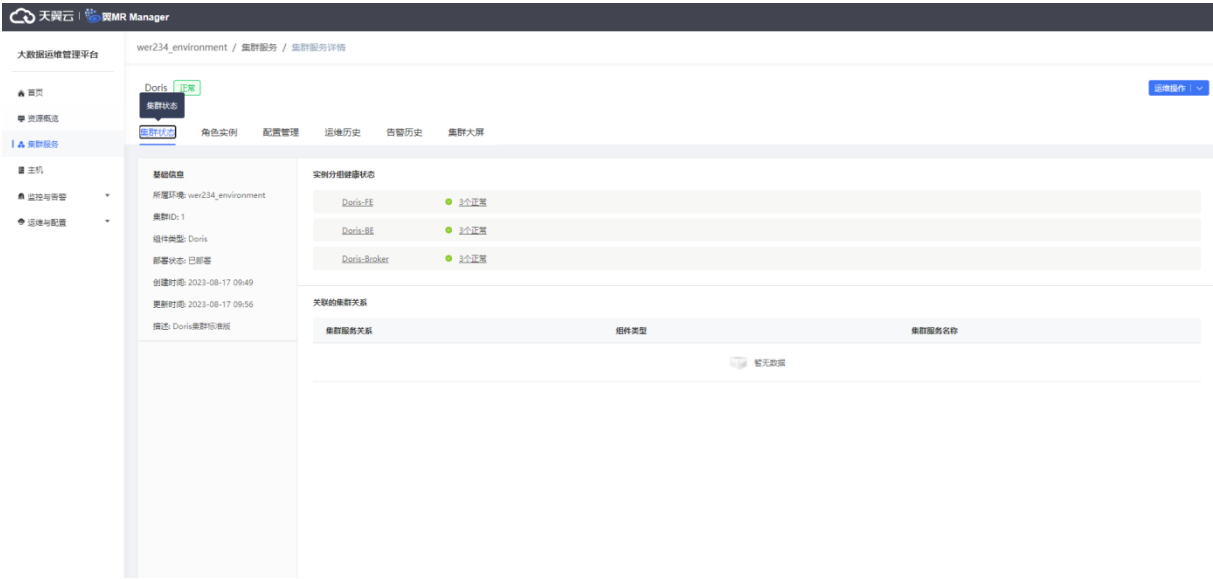
用户在日常使用中，可以在翼 MR Manager 管理所有组件（含角色实例）和主机的状态及指标信息：

- 状态信息，包含运行、健康及角色实例状态统计。
- 指标信息，各组件的主要监控指标项。

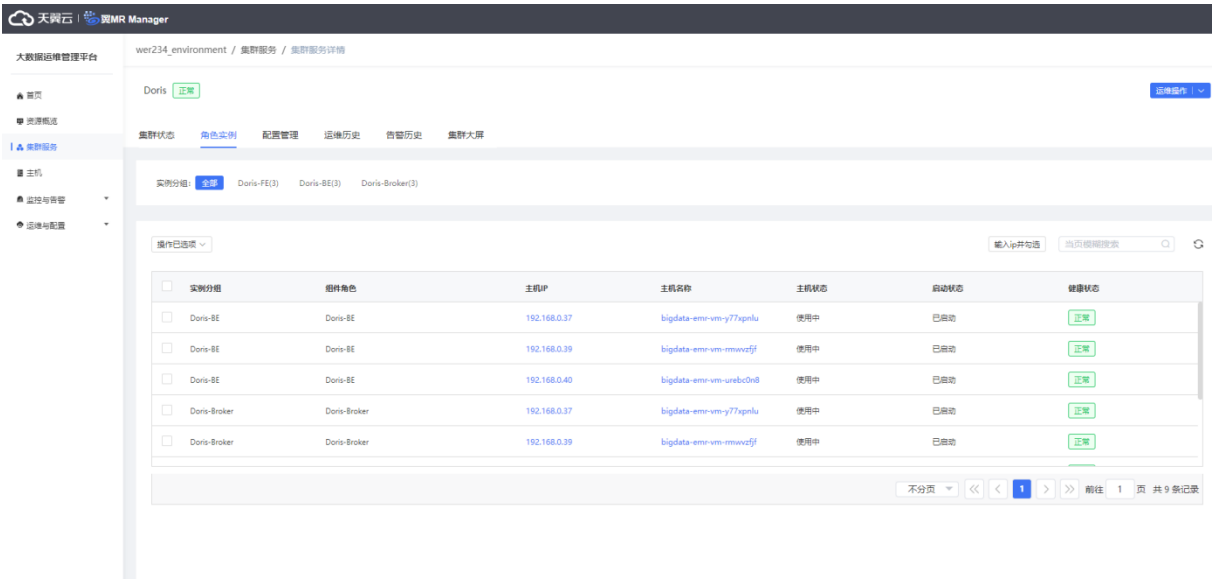
管理组件监控

- (1) 在翼 MR Manager 页面，点击左侧菜单栏“集群服务”。
- (2) 在集群服务列表点击任一集群进入。
- (3) 选择“集群状态”页，展示实例分组健康状态。

状态	描述
告警	集群实例服务曾出现过异常。
故障	集群实例服务产生故障。
正常	集群实例服务已启动。
未监控	集群实例未安装 prometheus，服务未监控。

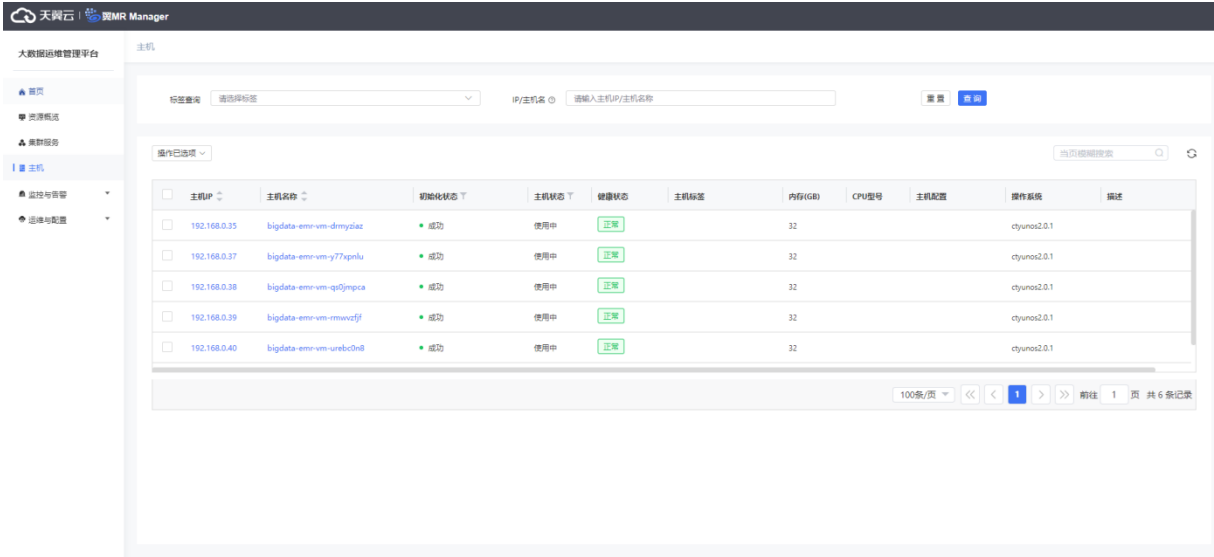


(4) 点击任一实例的健康状态后，跳转到角色实例页面，按实例分组显示每个实例组件角色，主机 IP，主机名称，主机状态，启动状态和健康状态。



管理主机监控

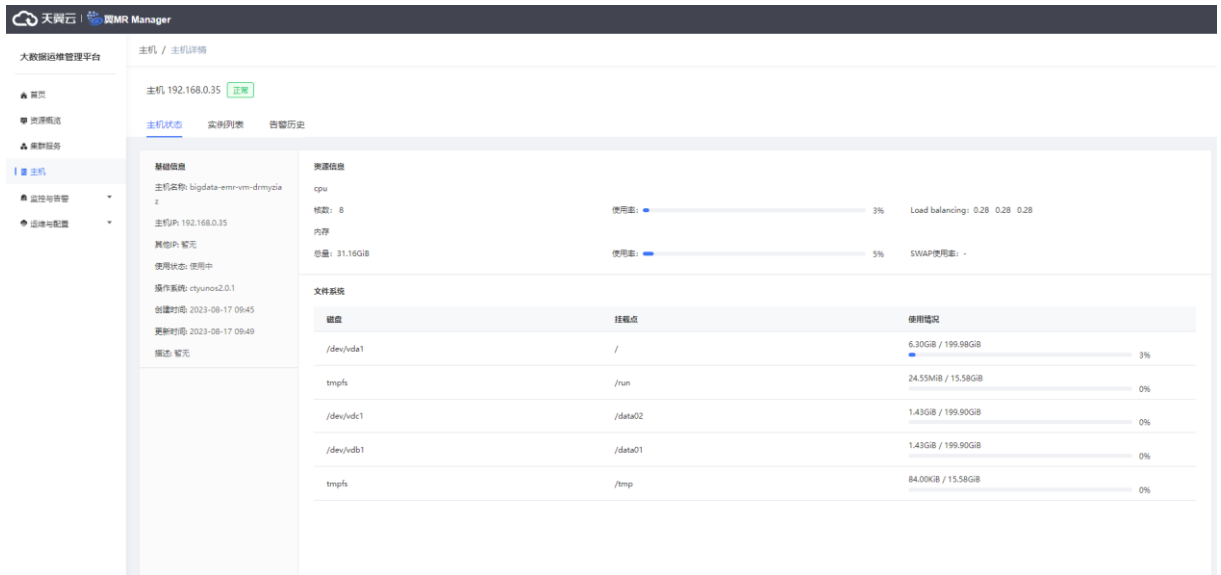
- (1) 在翼 MR Manager 页面，点击左侧菜单栏“主机”。
- (2) 主机页展示主机 IP、主机名称、初始化状态、主机状态、健康状态、主机标签、内存、CPU 型号、主机配置、操作系统、描述等字段。



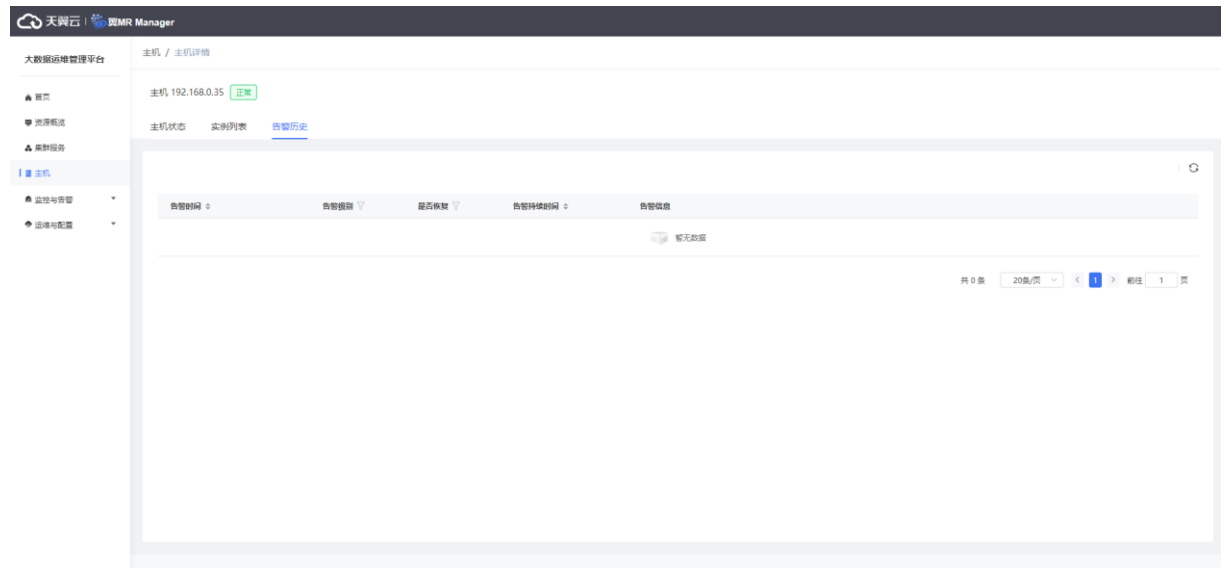
- (3) 健康状态栏，展示主机健康状态。

状态	描述
告警	集群实例服务曾出现过异常。
故障	集群实例服务产生故障。
正常	集群实例服务已启动。
未监控	集群实例未安装 prometheus，服务未监控。

- (4) 点击主机 IP，跳转到主机状态信息栏，展示主机状态、实例列表和告警历史。



(5) 点击告警历史，可追溯该主机 IP 下的历史告警。



3.3. 健康检查

3.3.1. 健康检查

本章节主要介绍健康检查操作和功能，主要是从集群健康检查维度进行介绍。

健康检查

操作场景

该任务指导用户在日常运维中完成集群进行健康检查的工作，以保证集群各项参数、配置以及监控没有异常、能够长时间稳定运行。

操作说明

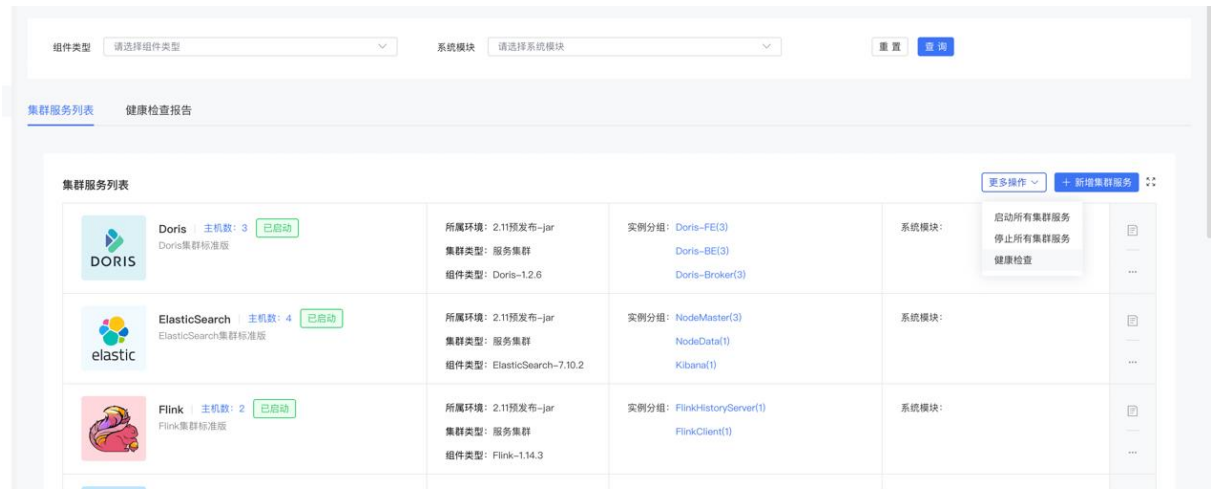
集群健康检查的范围包含 **Manager** 服务级别的各个组件检查：服务级别关注组件是否能够提供正常的服务、告警情况以及各个组件差异化的检查指标。

操作步骤

集群健康检查

(1) 手动执行所有服务的健康检查。

在集群详情页，单击集群服务列表右上角“更多操作 > 启动集群健康检查”。



(2) 周期执行所有服务的健康检查。

在集群服务>健康检查报告>点击列表右上方设置按钮>跳出周期任务设置弹框，支持用户根据天/周/月维度进行间隔配置，支持配合具体时刻，创建成功后，对支持的组件服务会定期发起健康检查。

健康检查间隔



说明

- 在翼 MR Manager 界面，选择“集群设置 > 更多操作 > 健康检查”，可以执行集群健康检查，集群列表页面发起的健康检查针对列表中所有的组件服务。
- 点击健康检查之后页面则跳转至健康检查报告 Tab 页面，展示集群的健康检查列表，列表置顶为最新发起的健康检查，点开一级列表下方展示各个组件的检查详情。
- 点击一级列表操作中的导出报告，则支持导出 CSV 文件，用户可本地查阅健康报告。

查看并导出检查报告

操作场景

为了满足对健康检查结果的进一步具体分析，您可以在翼 MR 中查看以及导出健康检查的结果。

操作说明

平台健康检查的范围包含 Manager 服务级别的健康检查。

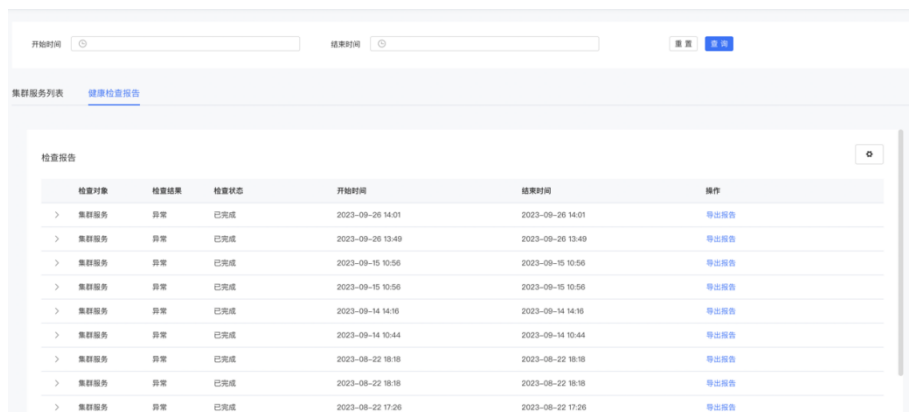
集群健康检查可以包含三方面检查项：各检查对象的服务状态、告警信息、以及每个组件差异化的相关指标。

前提条件

已执行健康检查。

操作步骤

- (1) 在集群详情页，点击健康检查报告“管理操作 > 查看集群健康检查报告”。



检查对象	检查结果	检查状态	开始时间	结束时间	操作
> 集群服务	正常	已完成	2023-09-26 14:01	2023-09-26 14:01	导出报告
> 集群服务	正常	已完成	2023-09-26 13:49	2023-09-26 13:49	导出报告
> 集群服务	正常	已完成	2023-09-15 10:56	2023-09-15 10:56	导出报告
> 集群服务	正常	已完成	2023-09-15 10:56	2023-09-15 10:56	导出报告
> 集群服务	正常	已完成	2023-09-14 14:16	2023-09-14 14:16	导出报告
> 集群服务	正常	已完成	2023-09-14 10:44	2023-09-14 10:44	导出报告
> 集群服务	正常	已完成	2023-08-22 18:18	2023-08-22 18:18	导出报告
> 集群服务	正常	已完成	2023-08-22 18:18	2023-08-22 18:18	导出报告
> 集群服务	正常	已完成	2023-08-22 17:26	2023-08-22 17:26	导出报告

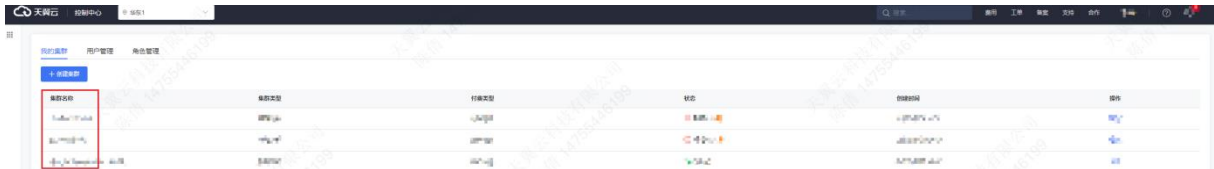
- (2) 在健康检查的报告面板上单击“导出报告”导出健康检查报告，下载后可本地查看检查项的完整信息，集群与主机健康检查报告导出格式均为 CSV。

3.3.2. 查看翼 MR 服务操作日志

本章节主要介绍查看翼 MR 服务操作日志。

操作日志入口

- (1) 在集群列表页面，在“集群名称”列下，单击需要查看日志的集群名称进入集群信息页面。



(2) 进入集群信息页面后，单击如下图所示“操作日志”即可进入操作日志页面。



操作类型

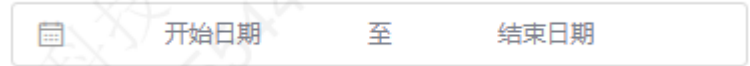
目前翼 MR 服务操作日志，提供了筛选框帮助用户快速定位问题原因。用户在选择筛选条件后，单击“查询”按钮进行查询日志，“重置”按钮清空筛选条件。



(1) 操作状态：操作状态筛选框如下图所示，单击可选择筛选条件，包括“成功”、“失败”和“进行中”三种条件。



(2) 时间范围：时间范围筛选框如下图所示，单击“开始日期”和“结束日期”可筛选条件。



日志字段

日志字段参数说明。

参数	参数说明
----	------

操作名称	记录执行的操作名称，包括：新建集群、配置升级等。
状态	记录操作的状态，包括：成功、失败、进行中。
操作范围	记录操作的范围，包括：集群、节点、节点组等。
操作用户	记录执行操作的用户。
开始时间/结束时间	记录操作开始执行时间和执行结束时间。
备注	对操作失败的备注说明。

3.4. 节点管理

3.4.1. 管理节点操作

本章节主要介绍翼 MapReduce 服务如何管理节点。

操作场景

当节点故障异常时，用户可能需要在翼 MapReduce 停止节点上的所有集群服务，对节点进行维护检查。故障清除后，启动节点上的所有集群服务恢复节点业务。

操作步骤

- (1) 登录翼 MR 控制台，在“我的集群”页面点击正常运行的集群名称，进入集群详情页面。
- (2) 在集群详情页，单击“翼 MR Manager”，点击”前往翼 MR Manager”。
- (3) 在 Manager “集群服务”页面，点击“更多操作”，并点击“停止所有集群服务” 执行相应操作。



3.4.2. 配置升级

本章节主要介绍翼 MapReduce 服务如何升级节点配置。

当 master 或者 core 节点 ECS 实例的规格（vCPU 和内存）无法满足您的业务需求时，您可以使用配置升级功能提升 ECS 实例规格。

背景信息

配置升级前需要前往 Manager“集群服务—更多操作—停止所有集群服务”，手动关闭大数据集群服务。

操作步骤

- (1) 登录翼 MapReduce 管理控制台。
- (2) 从“我的集群”中，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 选择“节点管理”，在需要升级配置的节点组的“操作”列并单击“配置升级”，进入升级节点配置页面。只有运行中的集群才能进行配置升级操作。

基础信息 节点管理 翼MR Manager 用户权限 操作日志 访问链接与端口					
运行中		请输入节点名称/内网IP/外网IP		查询	重置 刷新
节点组名称	节点组类型	付费类型	节点数量	规格	操作
▶ master	MASTER	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级
▶ core	CORE	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级

- (4) 配置升级前需要前往 Manager“集群服务—更多操作—停止所有集群服务”，手动关闭大数据集群服务，否则无法执行配置升级操作。

大数据运维管理平台

首页

资源概览

集群服务

主机

租户与资源

监控与告警

运维与配置

sjh-wx-test-0906-2_environment / 集群服务

组件类型 请选择组件类型

重置 查询

集群服务列表

健康检查报告

集群服务列表

HDFS

HDFS | 主机数: 6 已启动

HDFS集群标准版

集群类型: 服务集群

组件类型: HDFS-3.3.3

实例分

NameNode(2)

DataNode(3)

JournalNode(3)

下载客户端

更多操作

启动所有集群服务

停止所有集群服务

健康检查

- (5) 在配置升级页面，选择“升级后配置”，确认配置升级金额，勾选协议，并单击确定。
- (6) 配置升级完成后，您可以按需前往 Manager 的“配置管理”页面进行配置更改。可以参考“配置修改建议说明”。配置修改完成后，要前往 Manager“集群服务—更多操作—启动所有集群服务”手动拉起大数据集群服务。



说明

- 集群正在升级配置时，不允许继续升级节点配置/远程连接/绑定及解绑弹性 IP。
- 配置升级的操作会被记录在“操作日志”中，可以通过“状态”变化查看配置升级是否执行成功。
- 配置升级成功后，可以在“节点管理”查看集群的节点信息。
- 配置升级前一定要停止集群服务，否则无法执行配置升级操作。
- 配置升级时不能进行 IAM 用户同步操作，若操作 IAM 用户同步会出现同步失败。

3.4.3. 远程连接

本章节主要介绍翼 MapReduce 服务如何进行远程连接。

操作场景

本章节主要介绍如何通过控制台提供的远程登录功能（即 VNC 方式）登录到弹性云主机上，查看实例操作系统的运行状态或问题。

前提条件

ECS 实例已设置登录密码。

操作步骤

- (1) 登录翼 MR 控制台，点击正常运行的集群名称，进入集群详情页面。
- (2) 进入“节点管理”页面，点击“节点组名称”列的下拉按钮，展开对应的节点信息，点击“操作”列的“远程连接”，在新的页面上出现 shell 命令窗口。

< sjh-wx-test-0906-222

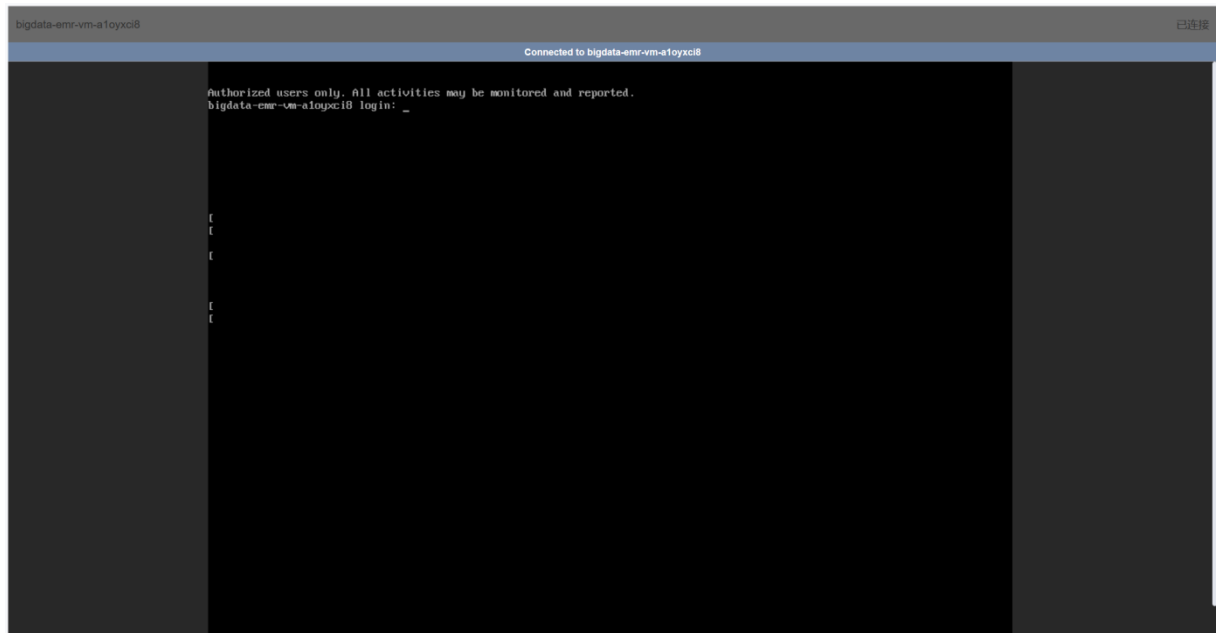
基础信息 **节点管理** 翼MR Manager 用户权限 操作日志 访问链接与端口

请选择节点状态 请输入节点名称/内网IP/外网IP

节点组名称	节点组类型	付费类型	节点数量	规格	操作
<input checked="" type="checkbox"/> master	MASTER	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级

节点名称	节点状态	内网IP	外网IP	角色实例	规格	磁盘信息	操作
bigdata-emr-vm-gnmv	运行中			-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程连接 更多
bigdata-emr-vm-shtpy	运行中			-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程连接 更多
bigdata-emr-vm-jxfjnr	运行中			-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程连接 更多

(3) 您可以根据界面提示，通过命令行实现远程登录机器。



3.4.4. 绑定/解绑弹性 IP

本章节主要介绍翼 MapReduce 服务如何绑定/解绑弹性 IP。

背景信息

创建的翼 MapReduce 集群中所有的节点都会默认被分配内网 IP, 外网 IP 需要用户自己去创建。

绑定弹性 IP 操作步骤

(1) 登录[翼 MapReduce 管理控制台](#)，在我的集群页面，点击集群具体名称，进入集群详情页。

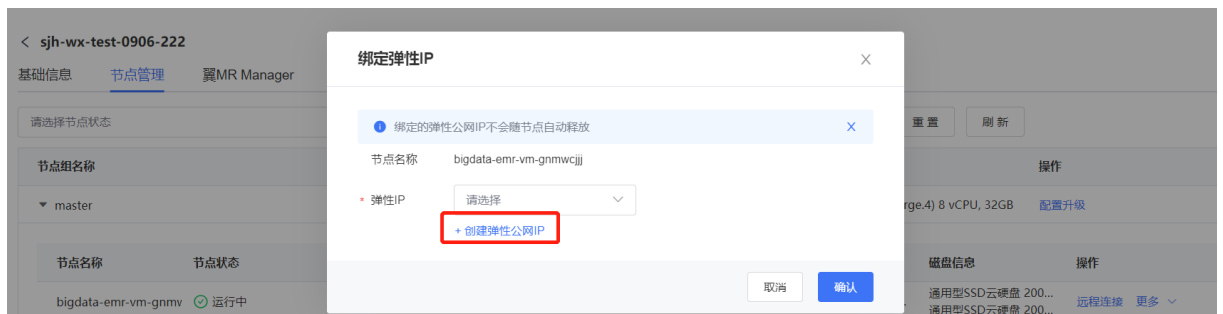
集群名称	集群类型	付费类型	状态
za20231207_01	数据分析	包年包月	自动中
test_doris_zjk	数据分析	包年包月	异常终止
test_es_zjk_no_del	云搜索	包年包月	已终止
emr_new_test_01	云搜索	包年包月	已终止
test-keypair	云搜索	包年包月	已终止
flvtest123	云搜索	包年包月	运行中
no_del_test	数据湖	包年包月	运行中
no_del_test_kafka	数据湖	包年包月	已终止
no_delete-manager	数据湖	包年包月	已终止
not-delete-test-kafka	数据湖	包年包月	已终止

(2) 在“节点管理”页面点击“节点组名称”列的下拉按钮，展开对应的节点信息，点击“操作”列的“更多”。

节点组名称	节点组类型	付费类型	节点数量	规格	操作
master	MASTER	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级

节点名称	节点状态	内网IP	外网IP	角色实例	规格	磁盘信息	操作
bigdata-emr-vm-gnmv	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程连接 更多
bigdata-emr-vm-shpty	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程 绑定弹性IP
bigdata-emr-vm-jxf8jn	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程 解绑弹性IP

(3) 绑定：点击“绑定弹性 IP”，在“绑定弹性 IP”的弹框中，如果您账号下没有可用的弹性公网 IP，需要点击“+创建弹性公网 IP”按钮跳转至新页面进行创建；如果您账号下有可用的弹性公网 IP，可以通过下拉“弹性 IP”的选择框选择 IP 后，点击“确认”进行绑定。需要注意的是：绑定的弹性公网 IP 不会随节点自动释放，需要您前往弹性 IP 控制台进行操作。



解绑弹性 IP 操作步骤

(1) 登录[翼 MapReduce 管理控制台](#)，在我的集群页面，点击集群具体名称，进入集群详情页。

天翼云 控制中心 华东1

我的集群 用户管理 角色管理

+ 创建集群 请选择集群状态 请选择集群类型

集群名称	集群类型	付费类型	状态
za20231207_01	数据分析	包年包月	自动中
test_doris_jzk	数据分析	包年包月	异常终止
test_es_jzk_no_del	云搜索	包年包月	已终止
emr_new_test_01	云搜索	包年包月	已终止
test-keypair	云搜索	包年包月	已终止
flinkdfla123	云搜索	包年包月	运行中
no_del_test	数据湖	包年包月	运行中
no_del_test_kafka	数据湖	包年包月	已终止
no_delete-manager	数据湖	包年包月	已终止
not-delete-test-kafka	数据湖	包年包月	已终止

(2) 在“节点管理”页面点击“节点组名称”列的下拉按钮，展开对应的节点信息，点击“操作”列的“更多”。

< sjh-wx-test-0906-222

基础信息 节点管理 翼MR Manager 用户权限 操作日志 访问链接与端口

请选择节点状态 请输入节点名称/内网IP/外网IP 查询 重置 刷新

节点组名称	节点组类型	付费类型	节点数量	规格	操作
master	MASTER	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级

节点名称	节点状态	内网IP	外网IP	角色实例	规格	磁盘信息	操作
bigdata-emr-vm-gnmv	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程连接 更多
bigdata-emr-vm-shtpy	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程 绑定弹性IP
bigdata-emr-vm-jxfbjr	运行中		-	-	(s7.2xlarge.4) 8 vCPU...	通用型SSD云硬盘 200... 通用型SSD云硬盘 200...	远程 解绑弹性IP

(3) 解绑：点击“解绑弹性 IP”，在“解绑弹性 IP”的弹框中，点击“确认”进行解绑。需要注意的是：解绑的弹性公网 IP 不会随解绑操作自动释放，需要您前往弹性 IP 控制台进行操作。

节点组名称	节点组类型	付费类型	节点数量	规格	操作
master	MASTER	包年包月	3	(s7.2xlarge.4) 8 vCPU, 32GB	配置升级

节点名称	节点状态	内网IP	外网IP	角色实例	规格	磁盘信息	操作
	运行中		-	1个角色实例	(s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB * 1 通用型SSD云硬盘 200GB * 2	远程连接 更多
	运行中		-	1个角色实例	(s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB * 1 通用型SSD云硬盘 200GB * 2	远程 更多
	运行中		-	1个角色实例	(s7.2xlarge.4) 8 vCPU, 32GB	通用型SSD云硬盘 200GB * 1 通用型SSD云硬盘 200GB * 2	远程 绑定弹性IP 解绑弹性IP

core CORE 包年包月 3 (s7.2xlarge.4) 8 vCPU, 32GB 扩容 配置

3.5. 组件管理

3.5.1. 对象管理简介

本章节主要介绍对象管理功能。

翼 MapReduce（简称：翼 MR）集群包含了各类不同的基本对象，不同对象的描述介绍如下表所示：

对象	描述	举例
集群服务	可以完成具体业务的一类功能集合	例如：HDFS 集群服务、YARN 集群服

		务。
服务角色	组成一个完整服务的一类功能实例，一般情况下可使用角色标识	例如：HDFS 由 NameNode、DataNode、HDFSClient、ZKFC、JournalNode 组成，在组件中起到不同的作用，就像扮演着不同的角色，相互通信协调，提供分布式文件存储能力。
实例分组	相同角色的实例组合成实例分组，通常实例分组对应一个角色，名称一般和角色的名称一样	例如：JournalNode 实例分组、DataNode 实例分组。
角色实例	服务角色在主机节点上运行的具体实例	例如：运行在 Host2 上的 KerberosAdmin，运行在 Host3 上的 KerberosServer。
主机	用于运行角色实例的物理主机或虚拟主机	例如：Host1~Host5。
集群	由多台主机组成的可以提供多种服务的逻辑实体	例如：名为 Cluster1 的集群由（Host1~Host5）5 个主机组成，提供了 HDFS 和 YARN 等服务。

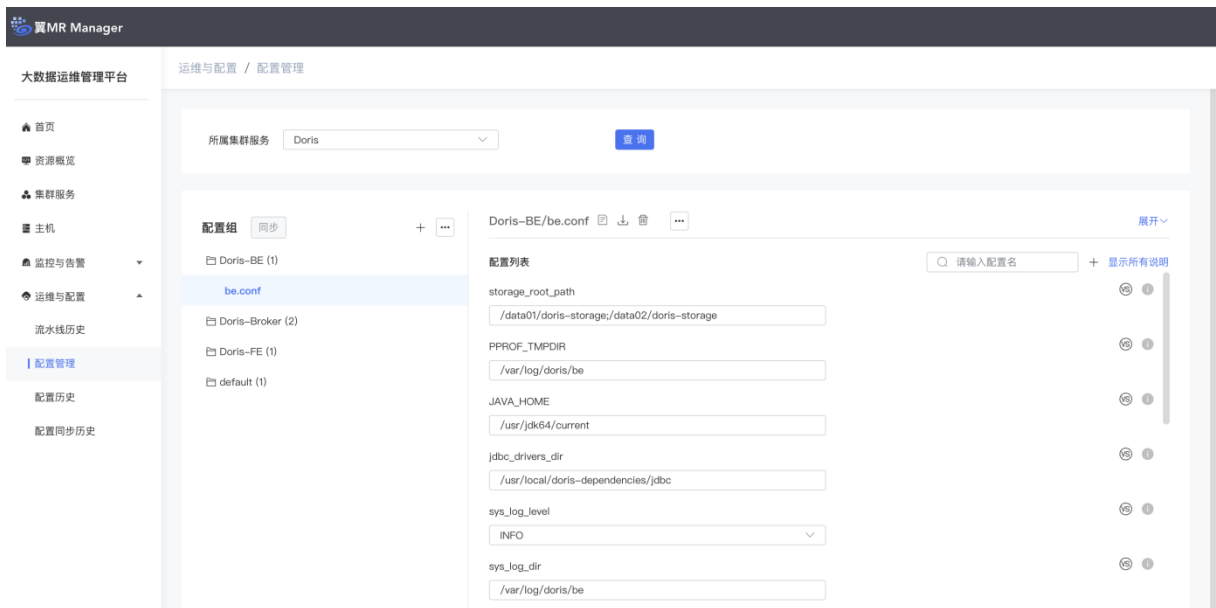
3.5.2. 查看配置

本章节主要介绍如何查看配置。

用户可以在翼 MapReduce（简称：翼 MR）上查看集群服务的配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务，即可查看该集群服务的相关配置。如图所示：



3.5.3. 管理服务操作

本章节主要介绍如何进行管理集群服务操作。

用户可以在翼 MapReduce（简称：翼 MR）：

- 启动集群服务
- 停止集群服务
- 滚动重启集群服务

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。

(6) 点击操作按钮“运维操作”，单击“启动集群服务”、“停止集群服务”、“滚动重启集群服务”执行相应操作。集群服务之间存在依赖关系。对某集群服务执行启动、停止滚动重启操作时，与该集群服务存在依赖关系的服务将受到影响，具体影响如下：启动某集群服务，该集群服务依赖的下层集群服务需先启动，集群服务功能才可生效。

- a. 停止某集群服务，依赖该集群服务的上层集群服务将无法提供功能。
- b. 滚动重启某集群服务，依赖该集群服务且启动的上层集群服务需重启后才可生效。

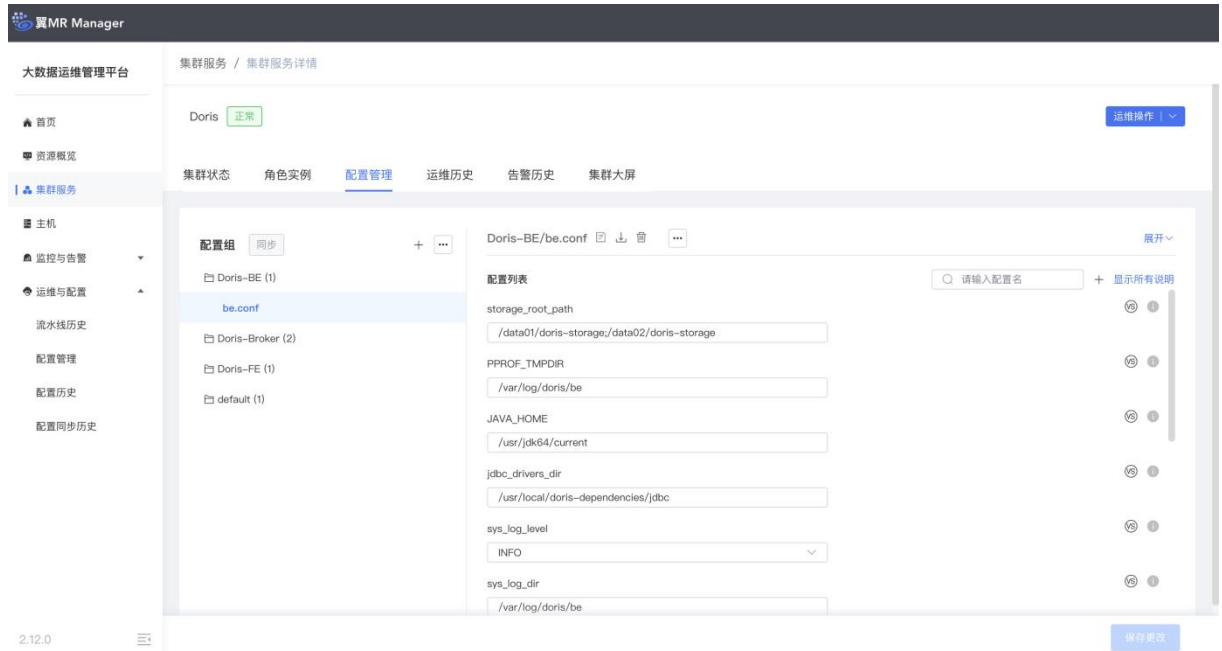
3.5.4. 配置服务参数

本章节主要介绍如何配置集群服务参数。

用户可以根据实际业务场景，在翼 MapReduce（简称：翼 MR）中快速查看和修改集群服务默认的配置。

修改集群服务配置操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。
- (6) 单击“配置管理”tab，选择要修改的配置文件，修改后，点击右下角“保存更改”按钮。



- (1) 单击“同步”按钮，对变更的配置文件进行同步。
- (2) 配置同步完成后，再次进入到该集群服务的详情页面。
- (3) 单击“运维操作”，单击“滚动重启集群”对该集群服务重启后，修改后的配置即可配置生效。

3.5.5. 同步服务配置

本章节主要介绍如何同步集群服务配置。

操作场景

当用户发现部分集群服务的配置存在待同步配置文件时，可以使用配置同步功能对待同步配置进行同步。

对系统的影响

同步集群服务配置后，需要滚动重启集群服务配置才会生效。重启时对应的集群服务不可用。

操作步骤

- (1) 登录翼 MR 管理控制台。

- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击“同步”按钮，对待同步的配置文件进行同步。
- (8) 配置同步完成后，再次进入到该集群服务的详情页面。
- (9) 单击“运维操作”，单击“滚动重启集群”对该集群服务重启后，修改后的配置即可配置生效。

3.5.6. 管理角色实例操作

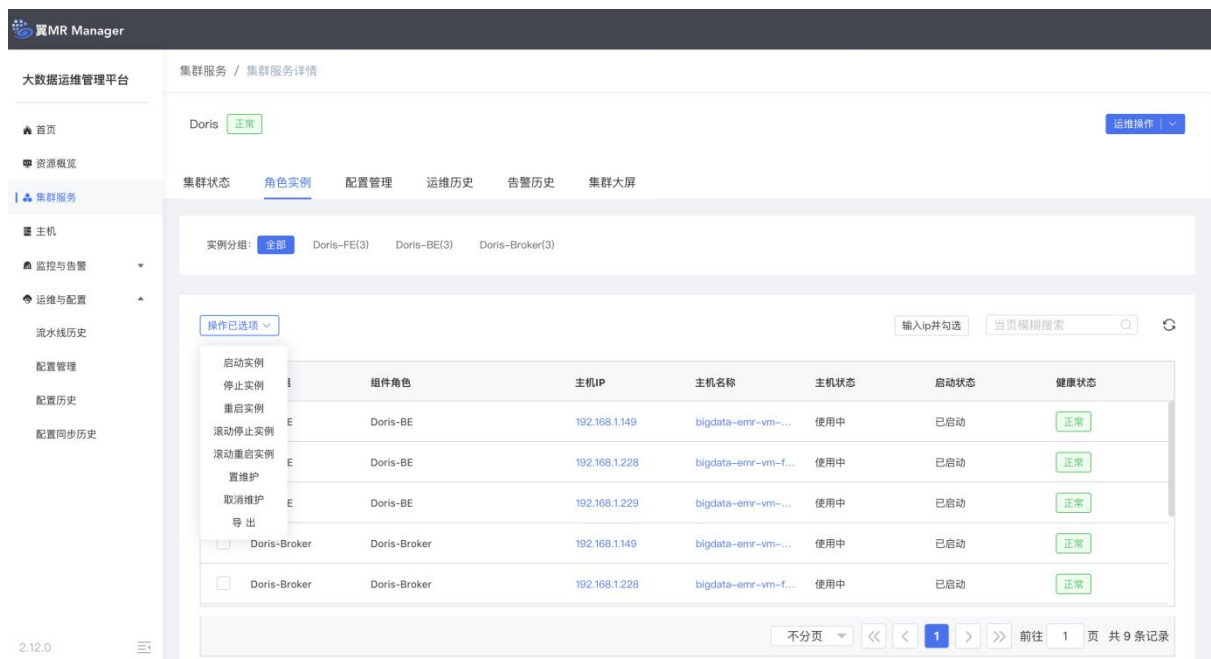
本章节主要介绍管理角色实例操作。

用户可以在翼 MapReduce（简称：翼 MR），对指定角色实例执行：

- 启动实例
- 停止实例
- 重启实例
- 滚动停止实例
- 滚动重启实例
- 置维护
- 取消维护

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项”，单击“启动实例”、“停止实例”、“重启实例”、“滚动停止实例”、“滚动重启实例”、“置维护”或“取消维护”等，执行相应操作。



3.5.7. 启动及停止集群

本章节主要介绍如何启动及停止集群。

集群是包含着服务组件的集合。用户可以启动或停止集群中所有集群服务。

启动及停止集群

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 单击“更多操作”，单击“启动所有集群服务”或“停止所有集群服务”执行相应的操作。

3.5.8. 下载集群的配置数据

本章节主要介绍如何下载配置文件。

操作场景

为了满足实际业务的需求，用户可以下载配置文件，用于快速更新集群服务配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。

- (6) 单击“配置管理”tab。
- (7) 单击要下载的配置文件名称，单击下载按钮即可。

3.5.9. 支持滚动重启

本章节主要介绍如何进行滚动重启。

在修改了大数据组件的配置项后，需要重启对应的服务来使得配置生效，使用普通重启方式会并发重启所有服务或实例，可能引起业务断服。为了确保服务重启过程中，尽量减少或者不影响业务运行，可以通过滚动重启来按批次重启服务或实例（对于有主备状态的实例，会先重启备实例，再重启主实例）。滚动重启方式的重启时间比普通重启时间久。

使用限制

- 请在低业务负载时间段进行滚动重启操作。例如：在滚动重启 Kafka 服务时候，如果 kafka 服务业务吞吐量很高（100M/s 以上的情况下），会出现 Kafka 服务滚动重启失败的情况。例如：在滚动重启 HBase 服务时候，如果原生界面上每个 RegionServer 上每秒的请求数超过 1W，需要增大 handle 数来预防重启过程中负载过大导致的 RegionServer 重启失败。
- 重启前需要观察当前 HBase 的负载请求数（原生界面上每个 rs 的请求数如果超过 1W，需要增大 handle 数来预防到时候负载不过来）。
- 在集群 Core 节点个数小于 6 个的情况下，可能会出现业务短时间受影响的情况。
- 请优先使用滚动重启操作来重启实例或服务，并勾选“仅重启配置过期的实例”。

滚动重启集群服务

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。
- (6) 单击“运维操作”，单击“滚动重启集群服务”。
- (7) 弹出确认滚动重启集群服务操作弹框，单击确定。
- (8) 进入到滚动重启集群服务操作页面，依次点击“下一步”，滚动重启任务完成后，单击“完成”。

滚动重启实例

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到 Manager 操作界面，单击“集群服务”菜单。
- (5) 选择指定集群服务，单击集群服务 ICON 进入到集群服务的详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。

- (8) 单击“操作已选项”，单击“滚动重启实例”。
- (9) 弹出确认滚动重启实例操作弹框，单击确定。
- (10) 进入到滚动重启实例操作页面，依次点击“下一步”，滚动重启任务完成后，单击“完成”。

滚动重启参数说明

滚动重启参数说明如下表所示。

参数名称	描述
单批实例数	按照角色实例维度，每 1 个批次的并发实例数，例如 HDFS-DataNode 默认为 1，输入限制为 int 的最大值，仅支持输入正整数，当角色实例数小于并发数时，以当前已有实例数为准。 例如：并发数设置为 2，HDFS-DataNode 的实例数为 3，则第一批次执行的实例数是 2，第 2 批次执行的实例数则为 1。
批次间隔时	上个批次成功后与下个批次开始的间隔时间，默认 30 秒，取值范围 1-1800，仅支持输入正整数，单位默认为：秒，不可改。 例如：设置时间为 600S 批次间隔则表示，前 1 个批次运行完成后，等待 600S 后开始下一个批次。
批次等待时间	上个批次开始后，开始下个批次的等待时间默认 30 秒，取值范围 1-1800，仅支持输入正整数；单位默认为：秒，不可改。 例如：设置时间为 600S 批次等待则表示，前 1 个批次开始运行，等待 600S 后不管前一个批次是否完成运行，都开始下一个批次的运行。
实例容错阈值	实例执行失败的容错次数为 2，默认为 0，输入限制为 int 的最大值，仅支持输入正整数。 例如：设置为 0 时，即表示任意一个角色实例的操作失败后，滚动操作终止；设置为 2 时，即表示任意 3 个角色实例的操作失败后，滚动操作才会终止。

3.6. 作业管理

3.6.1. 运行 MapReduce 作业

本章节主要介绍在翼 MapReduce 集群的后台提交新的 MapReduce 作业。

用户可将自己开发的程序提交到翼 MR 中，执行程序并获取结果。

前提条件

用户已经将运行作业所需的程序包和数据文件上传至 HDFS 系统中。

通过后台提交作业

例如安装路径为“/usr/local/hadoop3”。具体以实际为准。

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。
- (5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。

- (6) 集群默认开启 Kerberos 认证，执行以下命令认证当前用户。

示例：

```
Klist -kt /etc/security/keytabs/hdfs.keytab 获取 keytab 的  
principalname  
kinit /etc/security/keytabs/hdfs.keytab_x0005_ keytab 的  
principalname
```

- (7) 执行如下命令运行 Example 的 wordcount 作业。

```
hadoop jar /usr/local/hadoop3/share/hadoop/mapreduce/hadoop-  
mapreduce-client-jobclient-3.3.3-tests.jar TestDFSIO -  
Ddfs.replication=1 -write -nrFiles 100 -fileSize 100MB
```

3.6.2. 运行 SparkSubmit 作业

本章节主要介绍在翼 MR 集群的后台如何提交一个新的 Spark 作业。

用户可将自己开发的程序提交到翼 MR 中，执行程序并获取结果。

前提条件

用户已经将运行作业所需的程序包和数据文件上传至 HDFS 系统中。

通过后台提交作业

例如安装路径为“/usr/local/spark3”。具体以实际为准。

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。
- (5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。
- (6) 集群默认开启 Kerberos 认证，执行以下命令认证当前用户。

示例：

```
Klist -kt /etc/security/keytabs/spark.keytab 获取 spark.keytab  
的 principalname  
kinit /etc/security/keytabs/spark.keytab spark.keytab 的  
principalname
```

- (7) 执行如下命令运行计算圆周率作业。

```
spark-submit --master yarn --deploy-mode client --queue  
default --class org.apache.spark.examples.SparkPi  
/usr/local/spark3/examples/jars/spark-examples_2.12-3.2.2.jar  
100
```

3.6.3. 运行 HiveSQL 作业

本章节主要介绍在翼 MR 集群后台如何提交一个新的 Hive Sql 作业。

用户可将自己开发的程序提交到翼 MR 中，执行程序并获取结果。

Hive SQL 作业用于提交 SQL 语句和 SQL 脚本文件查询和分析数据，包括 SQL 语句和 Script 脚本两种形式，如果 SQL 语句涉及敏感信息，请使用 Script 提交。

前提条件

用户已经将运行作业所需的程序包和数据文件上传至 HDFS 系统中。

通过后台提交作业

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。

(5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。

(6) 集群默认开启 Kerberos 认证，执行以下命令认证当前用户。

示例：

```
Klist -kt /etc/security/keytabs/hive.keytab 获取 hive.keytab 的  
principalname  
kinit /etc/security/keytabs/hive.keytab hive.keytab 的  
principalname
```

(7) 执行 Hive Shell 命令。

```
hive  
>show databases;
```

3.6.4. 运行 SparkSQL 作业

本章节主要介绍翼 MapReduce 如何运行 SparkSql 作业。

用户可将自己开发的程序提交到翼 MR 中，执行程序并获取结果。本章节教您在翼 MR 集群后台如何提交一个新的 SparkSql 作业。SparkSQL 作业用于查询和分析数据，包括 SQL 语句和 Script 脚本两种形式，如果 SQL 语句涉及敏感信息，请使用 Spark Script 提交。

前提条件

用户已经将运行作业所需的程序包和数据文件上传至 HDFS 系统中。

通过后台提交作业

例如安装路径为“/usr/local/spark3”。具体以实际为准。

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。
- (5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。

(6) 集群默认开启 Kerberos 认证，执行以下命令认证当前用户。

示例：

```
Klist -kt /etc/security/keytabs/spark.keytab 获取 spark.keytab  
的 principalname  
kinit /etc/security/keytabs/spark.keytab spark.keytab 的  
principalname
```

(7) 打开 spark-sql 命令行，进入 spark-sql 命令行后可执行 SQL 语句，执行命令如下：

```
cd $SPARK_HOME  
./bin/spark-sql --conf spark.yarn.principal=default
```

若需要执行 SQL 文件，需要上传 SQL 文件（如上传到“/opt/”目录），上传文件后执行命令如下：

```
cd $SPARK_HOME  
./bin/spark-sql --conf spark.yarn.principal=default -  
f /opt/script.sql
```

3.6.5. 运行 Kafka 作业

本章节主要介绍翼 MapReduce 运行 kafka 作业。

用户可将自己开发的程序提交到翼 MR 中，执行程序并获取结果。本章节教您在翼 MR 集群后台如何提交一个新的 kafka 作业。

通过后台提交作业

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。
- (5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。
- (6) 当前集群默认开启 Kerberos 认证，执行以下命令认证当前用户。

示例：

```
Klist -kt /etc/security/keytabs/kafka.keytab 获取 kafka.keytab  
的 principalname  
kinit /etc/security/keytabs/kafka.keytab kafka.keytab 的  
principalname
```

- (7) 创建 kafka 的 topic。

```
/usr/local/kafka/bin/kafka-topics.sh --create --zookeeper
```

- (8) 消费 topic 中的消息。

```
/usr/local/kafka/bin/kafka-console-consumer.sh
--bootstrap-server <KafkaBroker_IP>:9092 --topic <TopicName>
--
consumer.config /usr/local/kafka/config/consumer.properties{
可选: --from-beginning --max-messages 5
}
```

注意

--from-beginning: 只能消费未被消费的历史数据，未消费数据不会出现。

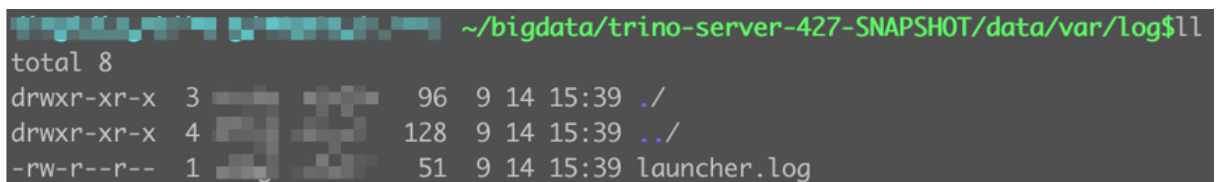
--max-messages: 最多消费多少条数据。

3.6.6. 查看主机和组件日志

本章节介绍如何查看主机和组件日志信息。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 选择“我的集群”，选中一个运行中的集群并单击集群名称，进入集群信息页面。
- (3) 在“节点管理”页选中单击 Master 节点，选择要进入的 Master 节点。
- (4) 单击该节点右侧的“远程连接”。
- (5) 根据界面提示，输入 Master 节点的用户名和密码，用户名、密码分别为 root 和创建集群时设置的密码。
- (6) 进入对应组件的日志目录，查看相关日志。



```
~/bigdata/trino-server-427-SNAPSHOT/data/var/log$ ll
total 8
drwxr-xr-x  3  96  9 14 15:39 ./
drwxr-xr-x  4 128  9 14 15:39 ../
-rw-r--r--  1  51  9 14 15:39 launcher.log
```

3.7. 告警管理

本章节主要介绍如何查看告警列表。

查看告警列表

告警历史显示了集群中的所有告警信息。

告警历史默认按时间顺序排列，时间最近的告警显示在最前端。

告警信息中的各字段说明如下表所示：

参数	参数说明
告警时间	产生告警的时间。
告警级别	灾难：集群服务不可用。 严重：集群服务出现可能影响服务的故障。 错误：集群服务发生异常，需要进一步处理，单集群服务仍提供正常服务。 警告：并不确认告警出现是否影响服务运行，提示可能的风险，需要进一步观测和排查。
是否恢复	告警是否还存在。
告警持续时间	从发生告警到现在的时长。
主机 IP	发生告警的主机 IP 地址。
告警信息	告警信息。
所属集群服务	发生告警的集群服务、实例。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“监控与告警> 告警历史”。
- (5) 查询条件选择指定集群服务、角色实例，或主机 IP，单击“查询”按钮即可查看相关告警信息。

4. 使用翼 MR 客户端

4.1. 安装客户端

本章节主要介绍客户端的安装。

该操作指导安装工程师安装翼 MR 集群所有服务（不包含 Flume）的客户端。Flume 客户端安装请参见[使用 Flume 客户端](#)。

修改集群内组件的服务端配置后，建议重新安装客户端，否则客户端与服务端版本将不一致，可能影响正常使用。

前提条件

- 当安装客户端节点为集群外节点时，该节点必须能够与集群内节点网络互通，否则安装会失败。
- 待安装客户端节点必须启用 NTP 服务，并保持与服务端的时间一致，否则安装会失败。
- 在节点上安装客户端，可以使用 root 用户或任意操作系统用户进行操作，要求该用户对客户端文件存放目录和安装目录具有操作权限，两个目录的权限为 775。本章节以使用操作系统用户 user_client 安装客户端进行举例，安装目录为/opt/hadoopclient。
- 需要允许用户使用密码方式登录 Linux 弹性云服务器（SSH 方式）。

安装客户端

- (1) 执行 NTP 时间同步，使当前安装客户端节点的时间与翼 MR 集群时间同步。
 - a. 执行 `vim /etc/ntp.conf` 命令编辑 NTP 客户端配置文件，并增加翼 MR 集群中 Master 节点的 IP 并注释掉其他 server 的地址。添加如下命令：

```
server master1_ip prefer
server master2_ip
```

- b. 执行 `service ntpd stop` 命令关闭 NTP 服务。
 - c. 执行 `/usr/sbin/ntpdate “主 Master 节点的 IP 地址”` 命令手动同步一次时间。
 - d. 执行 `service ntpd start` 或 `systemctl restart ntpd` 命令启动 NTP 服务。
 - e. 执行 `ntpstat` 命令查看时间同步结果。
- (2) 登录翼 MR 管理控制台。
 - (3) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
 - (4) 单击“前往下载客户端”，完成客户端压缩包和集群服务配置文件压缩包下载。
 - (5) 以 user_client 用户登录将要安装客户端的机器上。
 - (6) 解压客户端软件包。

进入安装包所在目录，例如“/opt/hadoopclient”。

```
cd /opt/hadoopclient
```

执行如下命令解压安装包到本地目录：

```
tar -zxvf CCDP-3.3.3-x86_64-client.tar.gz
```

- (7) 进入客户端软件包目录，根据需要安装的组件进入对应组件目录，执行如下命令安装

客户端到指定目录。

```
./install.sh 客户端安装目录
```

例如执行 `./install.sh /opt/hadoopclient` 命令安装客户端，等客户端安装完成。

说明

- 具体可参考软件包目录下的 README.txt。
- 机器上需要安装好 `krb5-workstation` 相关包。
- 机器上需要安装好 `jdk` 并设置好 `JAVA_HOME` 环境变量。
- `/etc/hosts` 内需要写入 `server` 端服务器的 IP 和主机名。
- 该安装包解压的目录即为客户端安装目录。
- `conf` 目录下需放置各个组件所需的配置文件，各组件所需的配置文件如附录 1 所示。
- `keytab` 目录下需放置各个组件所需的 `keytab` 文件。
- `krb5.conf` 配置文件拷贝需要 `root` 或有 `sudo` 权限的用户。若已完成 `krb` 配置或安装用户没有权限，则需确认 `krb` 已正确配置后，运行 `sh install.sh -s` 跳过 `krb5.conf` 配置，直接安装大数据客户端。
- 运行 `sh install.sh` 安装客户端，若某组件客户端无需安装，将组件目录下的 `install.sh` 改名即可，如 `mv hadoop-3.3.3/install.sh hadoop-3.3.3/install.sh.backup`。
- 运行 `sh uninstall.sh` 卸载客户端。

(8) 安装运行的日志保存在 `/tmp/yunyi_bigdata_client_install.log` 中。

(9) 集群默认启用 Kerberos 认证，执行以下命令认证当前用户。

```
kinit 翼 MR 集群用户
```

例如：

```
kinit admin
```

(10) 直接执行组件的客户端命令。

例如使用 HDFS 客户端命令查看 HDFS 根目录文件，执行 `hdfs dfs -ls /`

(11) 客户端安装完成后，可参考“[各组件客户端使用实践](#)”使用客户端。

4.2. 各组件客户端使用实践

4.2.1. 使用 YARN 客户端

本章节主要介绍 YARN 客户端的使用。

操作场景

该任务指导用户在运维场景或业务场景中使用 YARN 客户端。

前提条件

已安装客户端。

例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 YARN 客户端

- (1) 安装客户端。
- (2) 用户登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

- (4) 集群默认为安全模式，执行以下命令进行用户认证。

```
kinit 组件业务用户
```

- (5) 直接执行 YARN 命令。例如：

```
yarn application -list
```

4.2.2. 使用 Flume 客户端

本章节主要介绍使用 Flume 的操作指导。

操作场景

Flume 支持将采集的日志信息导入到 Kafka。

前提条件

- 已创建启用 Kerberos 认证的流集群。
- 已在日志生成节点安装 Flume 客户端，例如安装目录为“/opt/Flumeclient”。以下操作的客户端目录只是举例，请根据实际安装目录修改。
- 已配置网络，使日志生成节点与流集群互通。

使用 Flume 客户端

- (1) 安装客户端，具体请参考[安装客户端](#)章节。
- (2) 登录安装 Flume 客户端节点，切换到客户端安装目录。
- (3) 执行以下命令，修改 Flume 客户端配置文件“flume-env.sh”：

vi Flume 客户端安装目录/conf/flume-env.sh

在“-XX:+UseCMSCompactAtFullCollection”后面，增加以下内容：

```
undefined -Djava.security.krb5.conf=Flume 客户端安装目录 /flume-1.9.0/conf/kdc.conf -  
Djava.security.auth.login.config=Flume 客户端安装目录 /flume-1.9.0/conf/jaas.conf -  
Dzookeeper.request.timeout=120000
```

例如： "-XX:+UseCMSCompactAtFullCollection -Djava.security.krb5.conf=Flume 客户端安装目录 /Flume 组件版本号/conf/kdc.conf -Djava.security.auth.login.config=Flume 客户端安装目录/Flume 组件版本号/conf/jaas.conf -Dzookeeper.request.timeout=120000"

请根据实际情况，修改“Flume 客户端安装目录”，然后保存并退出。

(4) 假设 Flume 客户端安装路径为“/opt/FlumeClient”，执行以下命令，重启 Flume 客户端。

```
cd /opt/FlumeClient/Flume 组件版本号/bin./flume-manage.sh restart
```

(5) 执行以下命令，修改 Flume 客户端配置文件“properties.properties”。

```
vi Flume 客户端安装目录/Flume 组件版本号/conf/properties.properties
```

将以下内容保存到文件中：


```
#####
client.sources = static_log_source
client.channels = static_log_channel
client.sinks = kafka_sink
#####
#
#LOG_TO_HDFS_ONLINE_1

client.sources.static_log_source.type = spoolDir
client.sources.static_log_source.spoolDir = PATH
client.sources.static_log_source.fileSuffix = .COMPLETED
client.sources.static_log_source.ignorePattern = ^$
client.sources.static_log_source.trackerDir = PATH
client.sources.static_log_source.maxBlobLength = 16384
client.sources.static_log_source.batchSize = 51200
client.sources.static_log_source.inputCharset = UTF-8
client.sources.static_log_source.deserializer = LINE
client.sources.static_log_source.selector.type = replicating
client.sources.static_log_source.fileHeaderKey = file
client.sources.static_log_source.fileHeader = false
client.sources.static_log_source.basenameHeader = true
client.sources.static_log_source.basenameHeaderKey = basename
client.sources.static_log_source.deletePolicy = never

client.channels.static_log_channel.type = file
client.channels.static_log_channel.dataDirs = PATH
client.channels.static_log_channel.checkpointDir = PATH
client.channels.static_log_channel.maxFileSize = 2146435071
client.channels.static_log_channel.capacity = 1000000
client.channels.static_log_channel.transactionCapacity = 612000
client.channels.static_log_channel.minimumRequiredSpace = 524288000

client.sinks.kafka_sink.type = org.apache.flume.sink.kafka.KafkaSink
client.sinks.kafka_sink.kafka.topic = flume_test
client.sinks.kafka_sink.kafka.bootstrap.servers =
XXX.XXX.XXX.XXX:210079092,XXX.XXX.XXX.XXX:21007,XXX.XXX.XXX.XXX:21007
client.sinks.kafka_sink.flumeBatchSize = 1000
client.sinks.kafka_sink.kafka.producer.type = sync
client.sinks.kafka_sink.kafka.security.protocol = SASL_PLAINTEXT
client.sinks.kafka_sink.kafka.kerberos.domain.name = hadoop.XXX.com
client.sinks.kafka_sink.requiredAcks = 0

client.sources.static_log_source.channels = static_log_channel
client.sinks.kafka_sink.channel = static_log_channel
```

请根据实际情况，修改以下参数，然后保存并退出。

- spoolDir
- trackerDir
- dataDirs
- checkpointDir

- e. topic, 如果 kafka 中该 topic 不存在, 默认情况下会自动创建该 topic。
- f. kafka.bootstrap.servers, 默认情况下, 安全集群对应端口 21007, 普通集群对应端口 9092。
- g. kafka.security.protocol, 安全集群请配置为 SASL_PLAINTEXT, 普通集群请配置为 PLAINTEXT。

(6) Flume 客户端将自动加载“properties.properties”的内容。

当“spoolDir”生成新的日志文件, 文件内容将发送到 Kafka 生产者, 并支持 Kafka 消费者消费。

4.2.3. 使用 HBase 客户端

本章节主要介绍使用 HBase 的操作指导。

操作场景

该任务指导用户在运维场景或业务场景中使用 HBase 客户端。

前提条件

- 已安装客户端。例如安装目录为“/opt/hadoopclient”, 以下操作的客户端目录只是举例, 请根据实际安装目录修改。
- 非 root 用户使用 HBase 客户端, 请确保该 HBase 客户端目录的属主为该用户, 否则请参考如下命令修改属主。

```
chown user:group -R 客户端安装目录/HBase
```

使用 HBase 客户端

- (1) 安装客户端, 具体请参考[安装客户端](#)章节。
- (2) 用户登录安装客户端的节点。
- (3) 执行以下命令切换到客户端目录。

```
cd /opt/hadoopclient
```

- (4) 执行以下命令配置环境变量。

```
source bigdata_env
```

(5) 如果当前集群已启用 Kerberos 认证, 执行以下命令认证当前用户, 当前用户需要具有创建 HBase 表的权限, 参考[创建用户](#)章节, 为用户绑定对应角色。如果当前集群未启用 Kerberos 认证, 则无需执行此命令。

```
kinit 组件业务用户
```

例如：

```
kinit hbaseuser
```

(6) 直接执行 HBase 组件的客户端命令。

```
hbase shell
```

4.2.4. 使用 HDFS 客户端

本章节主要介绍使用 HDFS 的操作指导。

操作场景

该任务指导用户在运维场景或业务场景中使用 HDFS 客户端。

前提条件

- 已安装客户端。
- 例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 HDFS 客户端

- (1) 安装客户端。
- (2) 登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

- (4) 执行以下命令配置环境变量。

```
source bigdata_env
```

- (5) 如果集群为安全模式，执行以下命令进行用户认证。普通模式集群无需执行用户认证。

```
kinit 组件业务用户
```

- (6) 直接执行 HDFS Shell 命令。例如：

```
hdfs dfs -ls /
```

4.2.5. 使用 Hive 客户端

本章节主要介绍使用 Hive 的操作指导。

操作场景

该任务指导用户在运维场景或业务场景中使用 Hive 客户端。

前提条件

已安装客户端，例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 Hive 客户端

- (1) 安装客户端，具体请参考[安装客户端](#)章节。
- (2) 用户登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

- (4) 执行以下命令配置环境变量。

```
source bigdata_env
```

- (5) 根据集群认证模式，完成 Hive 客户端登录。
 - a. 安全模式，则执行以下命令，完成用户认证并登录 Hive 客户端。

```
kinit 组件业务用户  
beeline
```

- b. 普通模式，则执行以下命令，登录 Hive 客户端，如果不指定组件业务用户，则会以当前操作系统用户登录。

```
beeline -n 组件业务用户
```

说明

beeline 连接后可以编写并提交 HQL 语句执行相关任务。

4.2.6. 使用 Kafka 客户端

本章节主要介绍使用 Kafka 的操作指导。

操作场景

用户可以在集群客户端完成 Topic 的创建、查询、删除等基本操作。

前提条件

已安装客户端，例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 Kafka 客户端

- (1) 安装客户端，具体请参考[安装客户端](#)章节。
- (2) 登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端目录，例如：

```
cd /opt/hadoopclient/Kafka/kafka/bin
```

- (4) 执行以下命令，配置环境变量。

```
source /opt/hadoopclient/bigdata_env
```

- (5) 如果当前集群已启用 Kerberos 认证，执行以下命令认证当前用户。如果当前集群未启用 Kerberos 认证，则无需执行此命令。

```
kinit Kafka 用户
```

- (6) 创建一个 Topic。

```
sh kafka-topics.sh --create --topic 主题名称--partitions 主题占用的分区数 --replication-factor 主题的备份个数 --zookeeper ZooKeeper 角色实例所在节点 IP 地址:clientPort/kafka
```

- (7) 执行以下命令，查询集群中的 Topic 信息。

```
sh kafka-topics.sh --list --zookeeper ZooKeeper 角色实例所在节点 IP 地址:clientPort/kafka
```

- (8) 删除创建的 Topic。

```
sh kafka-topics.sh --delete --topic 主题名称 --zookeeper ZooKeeper  
角色实例所在节点 IP 地址:clientPort/kafka
```

输入 "y", 回车。

4.2.7. 使用 Trino 客户端

本章节主要介绍使用 Trino 的操作指导。

操作场景

该任务指导用户使用 Trino 客户端。

前提条件

已安装客户端。

例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 Trino 客户端

- (1) 将 trino-cli-406-SNAPSHOT-executable.jar 拷贝到/opt/hadoopclient 下。
- (2) 执行如下命令启动客户端。

```
java -jar trino-cli-406-SNAPSHOT-executable.jar --server  
Trino 服务的 IP:8080
```

说明

具体使用可参考 <https://trino.io/docs>。

- (3) 执行成功后的客户端界面。

```
trino:sf1> use tpch.sf1;  
USE  
trino:sf1>
```

4.2.8. 使用 ZooKeeper 客户端

本章节主要介绍使用 ZooKeeper 的操作指导。

前提条件

已安装客户端。

例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 ZooKeeper 客户端

- (1) 安装客户端并安装了 JDK8。
- (2) 登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

- (4) 集群默认为安全模式，执行以下命令进行用户认证。

```
kinit 组件业务用户
```

- (5) 直接执行命令。例如：

```
zkCli.sh -服务器 IP:端口号
```

4.2.9. 使用 Kyuubi 客户端

本章节主要介绍使用 Kyuubi 的操作指导。

前提条件

已安装客户端并且部署了 ZooKeeper 集群。

例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 Kyuubi 客户端

- (1) 安装客户端。
- (2) 登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

- (4) 集群默认为安全模式，执行以下命令进行用户认证。

```
kinit 组件业务用户
```

(5) 直接执行命令。例如：

```
bin/beeline -u 'jdbc:subprotocol://host:port'  
- subprotocol:kyuubi or hive2  
- host: IP address of the kyuubi server
```

说明

具体可参考：https://kyuubi.readthedocs.io/en/v1.7.1-rc0/client/jdbc/kyuubi_jdbc.html

4.2.10. 使用 Spark 客户端

本章节主要介绍使用 Spark 的操作指导。

前提条件

已安装客户端。

例如安装目录为“/opt/hadoopclient”，以下操作的客户端目录只是举例，请根据实际安装目录修改。

使用 Spark 客户端

- (1) 安装客户端。
- (2) 登录安装客户端的节点。
- (3) 执行以下命令，切换到客户端安装目录。

```
cd /opt/hadoopclient
```

(4) 如果集群为安全模式，执行以下命令进行用户认证。普通模式集群无需执行用户认证。

```
kinit 组件业务用户
```

(5) 直接执行命令。例如：

```
spark-shell --master spark://IP:port
```


5. 访问翼 MR 集群上托管的开源组件 Web 页面

5.1. 开源组件 Web 站点

本章节主要介绍开源组件 Web 站点。

场景介绍

翼 MR 集群默认在集群的 Master 节点和 Core 节点创建并托管了不同组件的 Web 站点，用户可以通过这些 Web 站点查看组件相关信息。

访问开源组件 Web 站点步骤：

- (1) 登录翼 MR 控制台管理页面。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“访问链接与端口”tab 即可。

Web 站点一览

详见下表：开源组件 Web 站点地址

站点类型	WebURL	例子	URL 说明
HDFS	http://nn-ip(nn-hostname):50070	http://127.0.0.1:50070	NameNode Web UI
HDFS	http://dn-ip(dn-hostname):1006	http://127.0.0.1:1006	DataNode Web UI
HDFS	http://jn-ip(jn-hostname):8480	http://127.0.0.1:8480	JournalNode Web UI
HDFS	http://rt-ip(rt-hostname):50071	http://127.0.0.1:50071	Router Web UI
YARN	http://rm-ip(rm-hostname):8088	http://127.0.0.1:8088	ResourceManger Web UI
YARN	http://nm-ip(nm-hostname):8042	http://127.0.0.1:8042	NodeManager Web UI
YARN	http://mrhis-ip(mrhis-hostname):19888	http://127.0.0.1:19888	JobHistory Web UI

HBase	http://hm-ip(hm-hostname):16010	http://127.0.0.1:16010	HMaster Web UI
HBase	http://rs-ip(rs-hostname):16030	http://127.0.0.1:16030	RegionServer Web UI
Spark	http://sparkhis-ip(sparkhis-hostname):18081	http://127.0.0.1:18081	SparkHistory Web UI
Doris	http://fe-ip(fe-hostname):8035	http://127.0.0.1:8035	FE Web UI
Doris	http://be-ip(be-hostname):8045	http://127.0.0.1:8045	BE Web UI
ElasticSearch	http://node-ip(node-hostname):9200	http://127.0.0.1:9200	ESNode Web UI
Ranger	http://rgadm-ip(raadm-hostname):6080	http://127.0.0.1:6080	RangerAdmin Web UI

5.2. 开源组件端口列表

本章节主要介绍开源组件端口列表。

背景

随着纳管的组件越来越多，各节点混部时，尤其在测试验证阶段，不同节点完全混部，会出现一些常见端口冲突，导致一些服务无法启动。

方法

通过 netstat 对应 pid 命令、官方配置文件，当前组件配置文件来确认组件常见固定端口。

结果

产出各组件常用端口，通过 linux 机器初始化时，对这些常用端口进行 net.ipv4.ip_local_reserved_ports 设置，预留这些端口。

常见组件 LIST

组件有 HDFS、YARN、HBase、Hive、Kafka、Spark、ZooKeeper、Kerberos、Trino、Flink、OpenLDAP、Doris、ElasticSearch、Kyuubi、Flume、Ranger 等。

HDFS 常用端口

官方文档参考 1: <https://hadoop.apache.org/docs/r3.3.3/hadoop-project-dist/hadoop-hdfs/hdfs-default.xml>

官方文档参考 2: <https://hadoop.apache.org/docs/r3.3.3/hadoop-project-dist/hadoop-hdfs-rbf/hdfs-rbf-default.xml>

版本: 3.3.3

配置参数	默认端口	当前配置端口	端口说明
dfs.namenode.http-address.x.x	9870	50070	namenode Web UI 绑定端口
dfs.namenode.https-address.x.x	9871	9871	namenode https 绑定端口
dfs.namenode.rpc-address.x.x	8020	54310	namenode 响应 client 请求 RPC 绑定端口
dfs.namenode.servicerpc-address.x.x	无	53310	namenode 响应内部服务请求 RPC 绑定端口, 例如 datanode 等 server 请求
com.sun.management.jmxremote.port	无	8006	namenode jmx 端口
dfs.datanode.address	9866	1004	datanode 数据传输端口
dfs.datanode.http.address	9864	1006	datanode http webUI 端口
dfs.datanode.ipc.address	9867	9867	datanode rpc 响应端口
com.sun.management.jmxremote.port	无	8016	datanode jmx 端口
dfs.datanode.https.address	9865		datanode https 端口
dfs.journalnode.rpc-address	8485	8485	journalnode rpc 端口
dfs.journalnode.http-address	8480	8480	journalnode http 端口

dfs.journalnode.https-address	8481	8481	journalnode https 端口
com.sun.management.jmxremote.port	无	8046	journalnode jmx 端口
dfs.ha.zkfc.port	8019	8019	zkfc rpc 端口
com.sun.management.jmxremote.port	无	8056	zkfc jmx 端口
dfs.federation.router.rpc-address	8888	55310	hdfs router 响应 client 请求 rpc 处理端口
dfs.federation.router.admin-address	8111		hdfs router 响应 admin 请求 rpc 处理端口
dfs.federation.router.http-address	50071		hdfs router http 端口
dfs.federation.router.https-address	50072		hdfs router https 端口
com.sun.management.jmxremote.port	无	8106	hdfs router jxm 端口

YARN 常用端口

官方文档参考：<https://hadoop.apache.org/docs/r3.3.3/hadoop-yarn/hadoop-yarn-common/yarn-default.xml>

版本：3.3.3

配置参数	默认端口	当前配置端口	端口说明
yarn.resourcemanager.address	8032	8032	RM 应用管理端口
yarn.resourcemanager.scheduler.address	8030	8030	RM 调度端口
yarn.resourcemanager.webapp.address	8088	8088	RM http webUI 端口
yarn.resourcemanager.webapp.https.address	8090	8090	RM https webUI 端口

yarn.resourcemanager.resource-tracker.address	8031	8031	RM jmx 端口
yarn.resourcemanager.admin.address	8033	8033	RM admin 端口
yarn.nodemanager.address	无	45454	NM container 管理端口
com.sun.management.jmxremote.port	无	8026	RM jmx 端口
yarn.nodemanager.localizer.address	8040	8040	NM 本地化 IPC 端口
yarn.nodemanager.collector-service.address	8048	8048	NM 控制服务端口
yarn.nodemanager.webapp.address	8042	8042	NM http webUI 端口
yarn.nodemanager.webapp.https.address	8044	8044	NM https webUI 端口
yarn.timeline-service.address	10200	10200	timeline server RPC 端口
yarn.timeline-service.webapp.address	8188	8188	timeline server http webUI 端口
yarn.timeline-service.webapp.https.address	8190	8190	timeline server https webUI 端口
yarn.sharedcache.admin.address	8047	8047	SCM 服务 admin RPC 端口
yarn.sharedcache.webapp.address	8788	8788	SCM 服务 webUI 端口
yarn.sharedcache.uploader.server.address	8046	8046	SCM 服务 节点管理 RPC 端口
yarn.sharedcache.client-server.address	8045	8045	SCM 服务响应 client RPC 端口
yarn.nodemanager.amrmproxy.address	8049	8049	在 yarn federation 下开启 amrmproxy 绑定的服务端口

yarn.router.webapp.address	8089	8089	yarn route http webUI 端口
yarn.router.webapp.https.address	8091	8091	yarn route https webUI 端口
com.sun.management.jmxremote.port	8036	8036	NM jmx 端口
com.sun.management.jmxremote.port	无	8086	mrhistory jmx 端口
mapreduce.jobhistory.address	10020	10020	mrhistory IPC 端口
mapreduce.jobhistory.webapp.address	19888	19888	mrhistory http webUI 端口
mapreduce.jobhistory.webapp.https.address	19890	19890	mrhistory https webUI 端口
mapreduce.jobhistory.admin.address	10033	10033	mrhistory admin RPC 端口
mapreduce.shuffle.port	13562	13562	MR shuffle 端口
spark.shuffle.service.port	7337	7337	spark shuffle 端口

HBase 常用端口

官方文档参考: https://hbase.apache.org/2.4/book.html#hbase_default_configurations

版本: 2.4.12

配置参数	默认端口	当前配置端口	端口说明
hbase.master.port	16000	16000	HMaster 服务绑定端口
hbase.master.info.port	16010	16010	HMaster webUI 绑定端口
com.sun.management.jmxremote.port	无	10101	HMaster jmx 端口
hbase.regionserver.port	16020	16020	regionserver 服务绑定端口

hbase.regionserver.info.port	16030	16030	regionserver webUI 绑定端口
com.sun.management.jmxremote.port	无	10102	regionserver jmx 端口
hbase.zookeeper.peerport	2888	2888	zk 服务端口
hbase.zookeeper.leaderport	3888	3888	zk 选举 leader 端口
hbase.zookeeper.property.clientPort	2181	2181	连接 zk client 端口
hbase.rest.port	8080	8080	hbase rest server 端口
hbase.status.multicast.address.port	16100	16100	hbase 多播端口

Hive 常用端口

官方文档参考：参考机器/usr/local/hive/conf/hive-default.xml.template

版本：3.1.2

配置参数	默认参数	当前配置端口	端口说明
hive.server2.thrift.port	10000	10000	hive.server2.transport.mode 模式为 binary hs2 thrift 端口
hive.server2.webui.port	10002	10002	hs2 webUI 端口
hive.server2.thrift.http.port	10001	10003	hive.server2.transport.mode 模式为 http hs2 thrift 端口
com.sun.management.jmxremote.port	无	9097	hs2 jmx 端口

hive.llap.management.rpc.port	15004	15004	LLAP 管理 rpc 端口
hive.llap.daemon.yarn.shuffle.port	15551	15551	LLAP yarn shuffle 端口
hive.llap.daemon.web.port	15002	15002	LLAP webUI 端口
hive.llap.daemon.output.service.port	15003	15003	LLAP output 服务端口
hive.zookeeper.client.port	2181	2181	连接 zk client 端口
hive.metastore.port	9083	9083	metastore RPC 绑定端口
com.sun.management.jmxremote.port	无	9093	metastore jmx 端口

Kafka 常用端口

官方文档参考: <https://github.com/apache/kafka/tree/2.8.1/config>

版本: 2.8.1

配置参数	默认参数	当前配置端口	端口说明
com.sun.management.jmxremote.port	无	8096	kafka broker jmx 端口
zookeeper.connect	2181	2181	连接 zk client 端口
listeners(SASL_PLAINTEXT)	无	9092	kafka SASL_PLAINTEXT 端口
listeners(PLAINTEXT)	无	9091	kafka PLAINTEXT 端口

Spark 常用端口

官方文档参考: <https://spark.apache.org/docs/3.2.2/configuration.html#spark-properties>

版本: 3.2.2

配置参数	默认参数	当前配置端口	端口说明
spark.yarn.historyServer.address	无	18081	spark history webUI 端口
spark.shuffle.service.port	7337	7337	spark shuffle 端口
spark.ui.port	4040	4040	应用 dashboard UI 端口

ZooKeeper 常用端口

官方文档参考：

https://zookeeper.apache.org/doc/r3.7.1/zookeeperAdmin.html#sc_adminserver_config

版本：3.7.1

配置参数	默认参数	当前配置端口	端口说明
admin.serverPort	8080	8080	zk admin 服务端口
clientPort	无	2181	连接 zk client 端口
com.sun.management.jmxremote.port	无	2182	zk jmx 端口
server.x	无	2888:3888;2181	zk 服务端口、选举 leader 端口，client 端口。

Kerberos 常用端口

配置参数	默认端口	当前配置端口	端口说明
kdc_ports	无	88	kdc 端口
iprop_port	无	744	iprop 端口

kpasswd	无	464	kpasswd 端口
admin_server	无	749	kerberos admin server 端口

Trino 常用端口

配置参数	默认端口	当前配置端口	端口说明
erver.http.port	无	9808	trino http 端口

OpenLDAP 常用端口

官方文档参考: <https://www.openldap.org/doc/admin24/slapdconfig.html>

版本: 2.4.44

配置参数	默认端口	当前配置端口	端口说明
provide	389	389	ldap TLS connections 端口
provide	636	636	ldap legacy SSL connections 端口

Doris 常用端口

官方文档参考:

FE: <https://doris.apache.org/zh-CN/docs/1.2/admin-manual/config/fe-config>

BE: <https://doris.apache.org/zh-CN/docs/1.2/admin-manual/config/be-config>

版本: 1.2.6

配置参数	默认端口	当前配置端口	端口修改	端口说明
be_port	9060	9060		BE 上 thrift server 的端口号, 用于接收来自 FE 的请求。
brpc_port	8060	8060		BE 上的 brpc 的端口, 用于 BE

				之间通讯。
edit_log_port	9010	9010		bdbje 端口
heartbeat_service_port	9050	9050		BE 上心跳服务端口（thrift），用于接收来自 FE 的心跳。
http_port	8030	8030	8035	FE http 端口，当前所有 FE http 端口都必须相同，生产环境被改成了 8035。
https_port	8050			FE https 端口，当前所有 FE https 端口都必须相同。
query_port	9030	9030		Doris FE 通过 mysql 协议查询连接端口
rpc_port	9020	9020		FE Thrift Server 的端口
single_replica_load_brpc_port	9070			单副本数据导入功能中，Master 副本和 Slave 副本之间通信的 RPC 端口。Master 副本 flush 完成之后通过 RPC 通知 Slave 副本同步数据，以及 Slave 副本同步数据完成后通过 RPC 通知 Master 副本。系统为单副本数据导入过程中 Master 副本和 Slave 副本之间通信开辟了独立的 BRPC 线程池，以避免导入并发较大时副本之间的数据同步抢占导入数据分发和查询任务的线程资源。
single_replica_load_download_port	8050			单副本数据导入功能中，Slave 副本通过 HTTP 从 Master 副本下载数据文件的端口。系统为单副本数

				据导入过程中 Slave 副本从 Master 副本下载数据文件开辟了独立的 HTTP 线程池，以避免导入并发较大时 Slave 副本下载数据文件抢占其他 http 任务的线程资源。
webserver_port	8040	8045	8045	BE 上的 http server 的服务端口

ElasticSearch 常用端口

官方文档参考：<https://www.elastic.co/guide/en/elasticsearch/reference/7.10/modules-network.html>

版本：Elasticsearch-7.10.2

配置参数	默认端口	当前配置端口	端口说明
com.sun.management.jmxremote.port	无	9400	es jmx 端口
http.port	无	9200	es http 端口
transport.port	无	9300	es transport 端口

Kyuubi 常用端口

官方文档参考：<https://kyuubi.readthedocs.io/en/v1.6.0-incubating/deployment/settings.html#kyuubi-configurations>

版本：1.6.0

配置参数	默认参数	当前配置端口	端口说明
kyuubi.frontend.bind.port	10009		kyuubi fe 端口
kyuubi.frontend.mysql.bind.port	3309		kyuubi fe mysql 绑定端口

kyuubi.frontend. rest.bind.port	10099		kyuubi fe rest 绑定端口
kyuubi.frontend.t hrift.binary.bind. port	10009		kyuubi thrift fe 绑定端口
kyuubi.metrics.p rometheus.port	10019		kyuubi 暴露给 Prometheus metrics http 端口

Ranger 常用端口

官方文档参考: <https://cwiki.apache.org/confluence/display/RANGER/Ranger+Installation+Guide>

版本: 2.2.0

配置参数	默认参数	当前配置端口	端口说明
ranger.usersync.port	5151		ranger usersync 服务端口
com.sun.management.j mxremote.port	无	60081	ranger usersync jmx 端口
ranger.unixauth.service. port	5151		ranger unixauth 服务端口
ranger.service.shutdown .port	6085		ranger 服务 shutdown 端口
ranger.audit.elasticsearc h.port	9200		ranger audit 配置的 es 服务端口
ranger.service.http.port	6080		ranger-admin http webUI 端口
ranger.service.https.port	6182		ranger-admin https webUI 端口

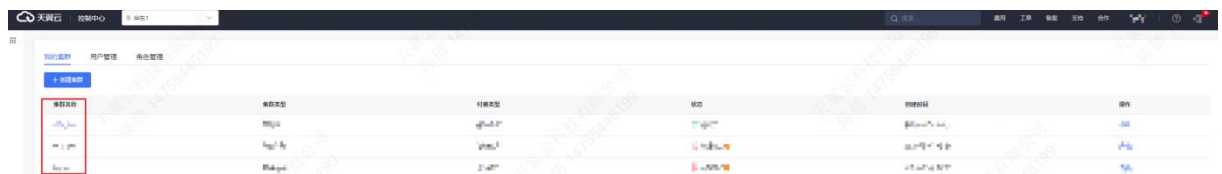
5.3. 通过弹性公网 IP 访问

本章节主要介绍通过弹性公网 IP 访问。

为了方便用户访问开源组件的 Web 站点，翼 MapReduce 集群支持通过为集群绑定弹性公网 IP 的方式，访问集群上托管的开源组件。该方式更加简便易操作，推荐使用该方式访问开源组件的 Web 站点。

绑定弹性公网 IP

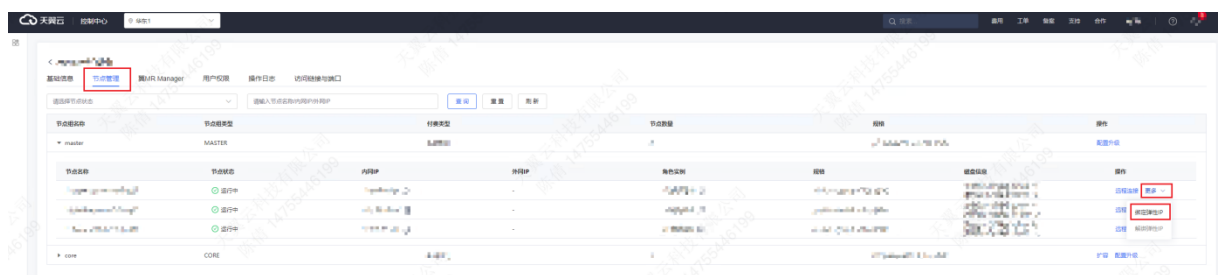
(1) 在集群列表页面，单击“集群名称”列下需要绑定 IP 的集群名称，进入该集群信息页面。



(2) 确定需要开通公网访问组件所在的节点。可以通过单击“访问链接与端口”，在“集群服务名称”列中找到需要开通公网访问的组件，其“原生 UI 地址”中包含所在节点内网 IP 地址。



(3) 单击“节点管理”，点击“节点组名称”列的下拉按钮，展开对应的节点信息，确定需要开通公网访问的节点，单击“操作”列的“更多”。



(4) 单击“绑定弹性 IP”，在“绑定弹性 IP”的弹框中，如果您账号下没有可用的弹性公网 IP，需要点击“+创建弹性公网 IP”按钮跳转至新页面进行创建；如果您账号下有可用的弹性公网 IP，可以通过下拉“弹性 IP”的选择框选择 IP 后，点击“确认”进行绑定。

绑定弹性IP

X

绑定的弹性公网IP不会随节点自动释放

X

节点名称

*

弹性IP

请选择

+ 创建弹性公网IP

取消

确认

说明

绑定的弹性公网 IP 不会随节点自动释放，需要您前往弹性 IP 控制台进行操作。

(5) 绑定弹性 IP 后，根据需要对集群安全组规则进行相应的添加或修改，变更端口的访问权限。之后可以通过“公网 IP：端口”的方式访问该开源组件的 Web 站点。

6. Manager 操作指导

6.1. 从这里开始

6.1.1. Manager 入门指导

本章节主要介绍翼 MR Manager 的操作。

概述

翼 MR Manager 由天翼云自主研发，主要为运维工程师提供日常的大数据组件运维管理操作能力。通过可视化、流程化的方式对系统中的各个系统资源和数据资产进行管理，并支持自动化的运维调度、统一的配置文件管理、统一运维监控，支持组件集群服务管理、多租户管理、资源管理等功能。在提高运维人员工作的效率同时，并为大数据运维工作者，提供专业全面的大数据运维能力，大大降低大数据平台运维门槛。

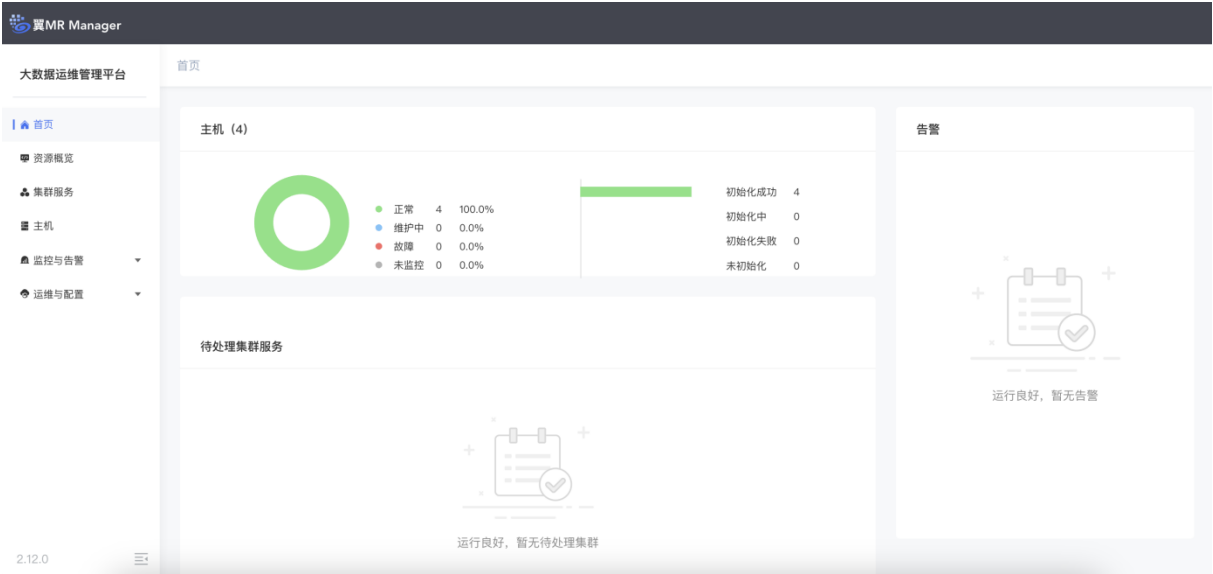
- 提供集群服务管理，为用户快速掌握环境、集群、主机、组件服务等数据信息。
- 提供运维自动化管理，提高大数据运维效率，降低人力成本。

浏览器支持

建议使用 Google Chrome 浏览器。

系统界面简介

翼 MR Manager 提供统一的集群管理平台，帮助您快捷、直观的完成集群的运行维护。
详见下图：Manager 系统界面。



界面左侧为菜单导航区域，右侧为显示区域和操作区域。

菜单导航区域的详细功能如下表所示：

菜单导航	功能描述
首页	展示平台主机健康状态、初始化状态汇总； 故障和告警集群信息； 平台告警信息。
资源概览	展示所有主机的 CPU、内存、网络等信息。
集群服务	展示所有集群，集群列表按组件类型，列表视图列出，在集群列表处一键启动所有集群、一键停止所有集群； 支持单集群的启动、停止、重启、滚动重启等操作； 支持角色实例的启动、停止、重启、滚动停止、滚动重启等操作； 支持查看集群的运维历史、告警历史和监控看板等。

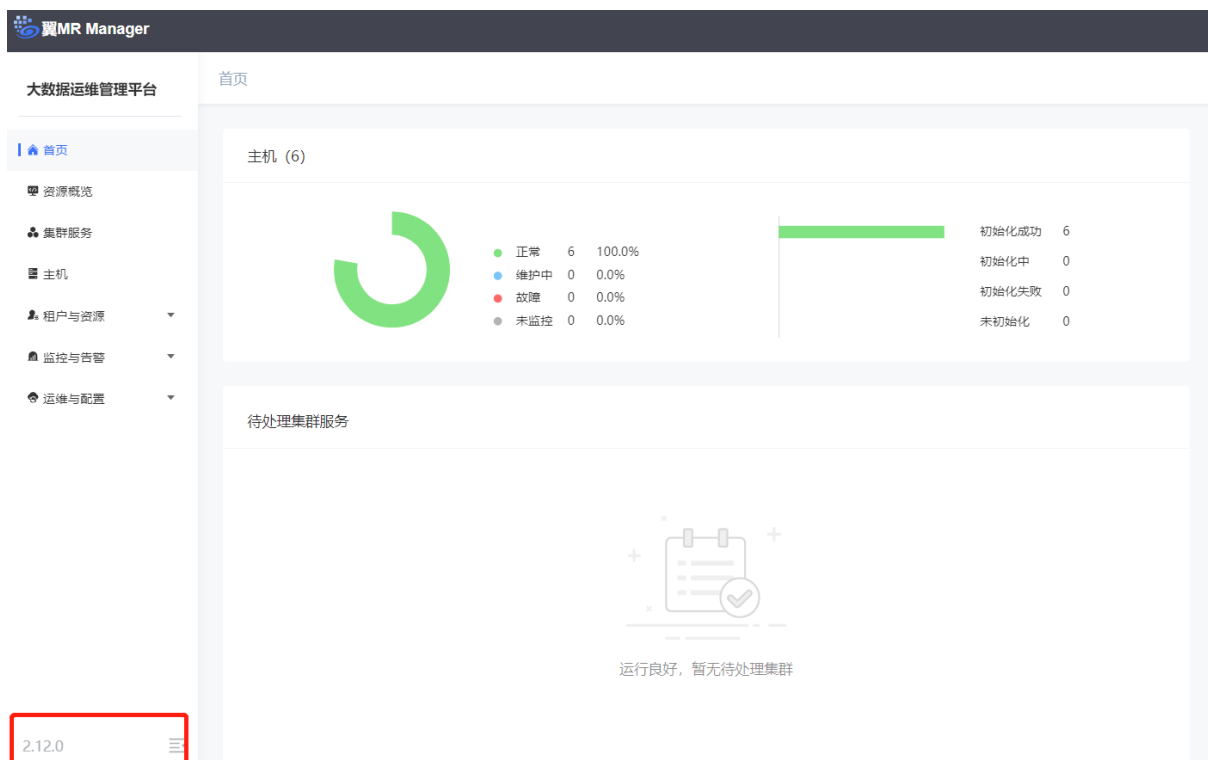
主机	<p>展示当前平台的所有主机；</p> <p>支持对主机进行置维护、取消维护操作；</p> <p>支持查看单个主机基础信息、主机资源使用信息、主机上的文件系统信息；</p> <p>支持查看单个主机上安装的角色实例、告警历史。</p>
监控与告警	<p>支持查询角色实例级、主机级的监控指标；</p> <p>支持指标结果的绘图操作，让用户更直观获取监控项变化；</p> <p>支持按照集群级、角色实例级、主机级查询告警内容。</p>
运维与配置	<p>支持对各集群服务的配置文件进行修改、同步、回滚等操作；</p> <p>支持查看支持查看所有流水线的运行历史记录，以及操作人。</p>

6.1.2. 查询 Manager 版本号

本章节主要介绍查看翼 MR Manager 版本号的步骤。

操作步骤如下：

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab。
- (4) 单击“前往翼 MR Manager”，进入到 Manager 界面，在 Manager 界面左下角即可查看到 Manager 的版本号。



The screenshot displays the Wing MR Manager web interface. The top navigation bar includes the logo and the text "翼MR Manager". Below this, a sidebar on the left lists various management options: "大数据运维管理平台", "首页", "资源概览", "集群服务", "主机", "租户与资源", "监控与告警", and "运维与配置". The main content area shows a dashboard for "主机 (6)" with a circular progress indicator and a table of status metrics:

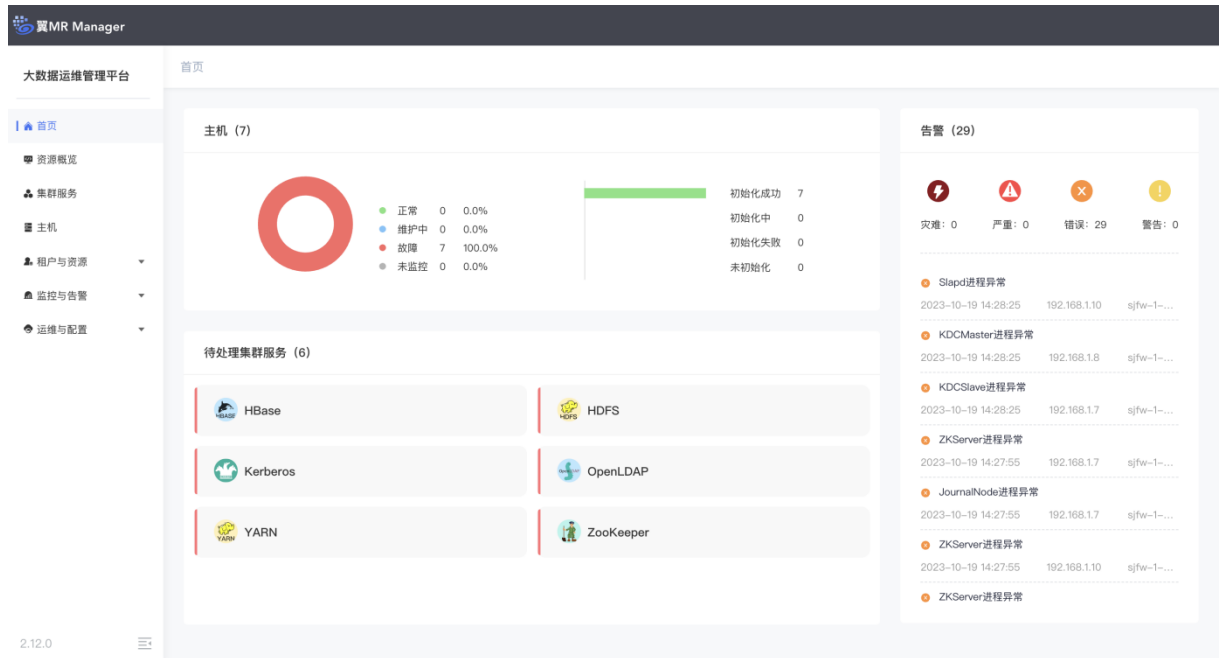
状态	数量	百分比
正常	6	100.0%
维护中	0	0.0%
故障	0	0.0%
未监控	0	0.0%

Below the status table, there is a section for "待处理集群服务" (Pending Cluster Services) which shows a message: "运行良好，暂无待处理集群" (Running well, no pending clusters). In the bottom left corner of the interface, the version number "2.12.0" is displayed within a red rectangular box.

6.2. 首页概述

本章节主要介绍翼 MR Manager 的首页特性。

进入到翼 MR Manager 以后，默认进入到首页，如图所示：



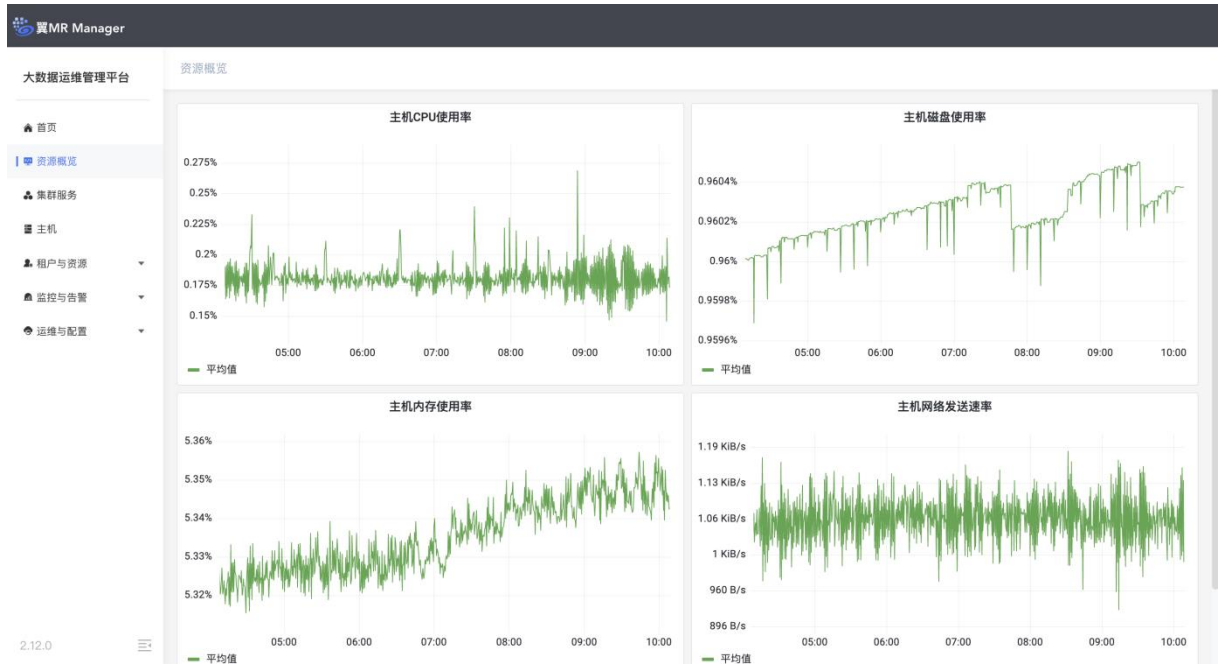
- 主机：展示平台主机健康状态汇总、主机初始化状态汇总。
- 待处理集群：展示故障和告警的集群信息。
- 告警：展示待处理的告警信息，支持查看不同级别的告警信息。

6.3. 资源概览

本章节主要介绍翼 MR Manager 的资源概览特性。

进入到翼 MR Manager 以后，点击菜单“资源概览”，进入资源概览页面。

资源概览展示当前平台所有主机最近 5 个小时的 CPU 使用率、磁盘使用率、内存使用率、网络发送速率、网络接收速率信息。如图所示：

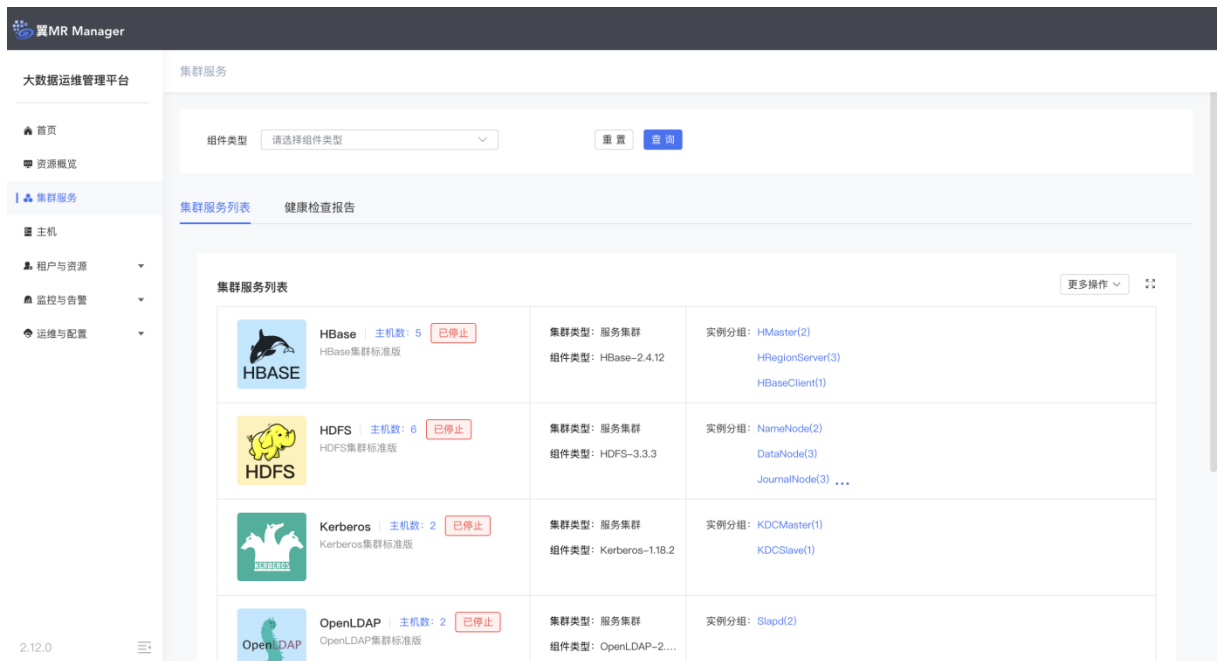


6.4. 集群服务

6.4.1. 集群服务管理概述

本章节主要介绍翼 MR Manager 的集群服务管理特性。

进入到翼 MR Manager 以后，点击菜单“集群服务”，进入集群服务页面，如图所示：



- 集群服务列表上方为查询区域，支持组件类型进行查询。
- 集群服务列表按组件类型，列表视图列出。
- 集群服务列表在右上角，可支持以下操作：启动所有集群服务、停止所有集群服务、健康检查。

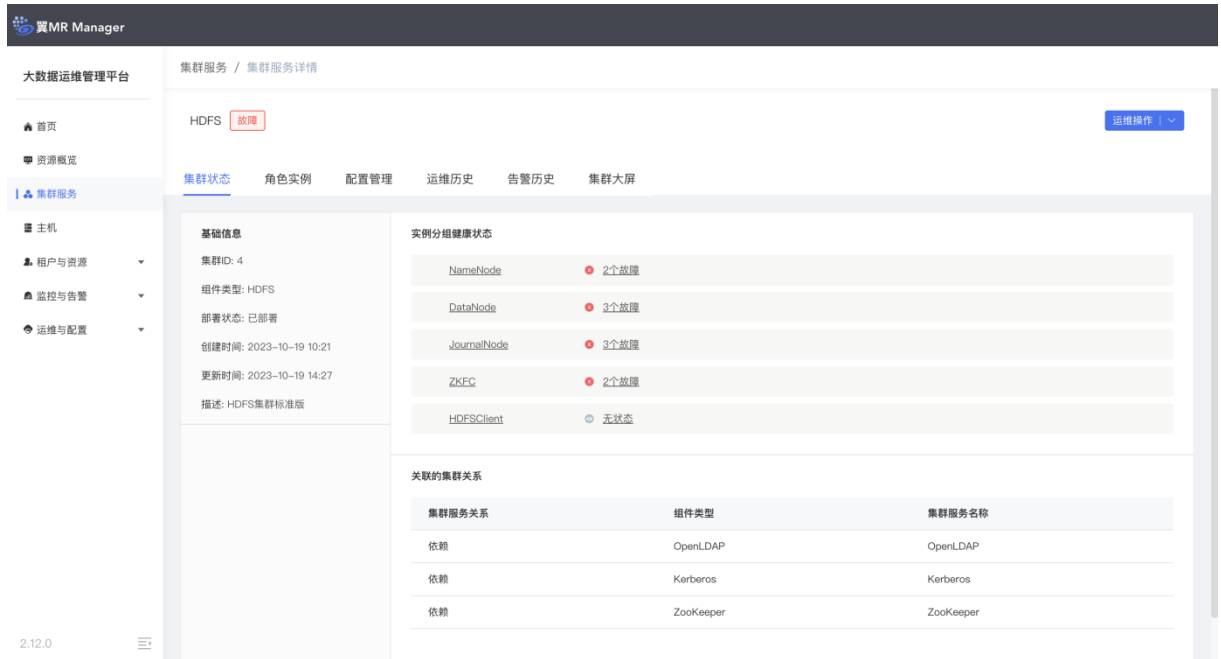
- 点击集群服务 ICON 或集群服务名称可进入到集群服务详情页面。
- 点击主机数可进入到主机菜单页面。
- 点击实例分组可进入到集群详情-角色实例页面。

6.4.2. 集群服务详情页面概述

本章节主要介绍翼 MR Manager 的集群服务详情页面。

进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。

单击指定集群服务的名称进入集群服务详情页面，如图所示：



- 详情页面上面区域为集群服务名称、集群服务的健康状态。
- 详情页面右上角为集群服务的运维操作，包含：启动集群服务、停止集群服务、滚动重启集群服务等。
- 详情页面上面区域则依次为该集群服务的：集群状态、角色实例、配置管理、运维历史、告警历史、集群大屏等信息。

6.4.3. 下载客户端

本章节主要介绍翼 MR 的下载客户端操作。

操作场景

未部署服务端的服务器上需要安装组件客户端时，需要下载客户端安装包和配置文件。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群详情页面。
- (3) 单击“基础信息”页面>软件信息中的“前往下载客户端”按钮。

集群信息

集群名称	emr_dw_1023	高可用	默认开启
创建时间	2023.10.23 17:07	付费类型	包年包月
运行时间	3天 1小时	Kerberos认证	默认开启
到期时间	2023.11.23 17:07		

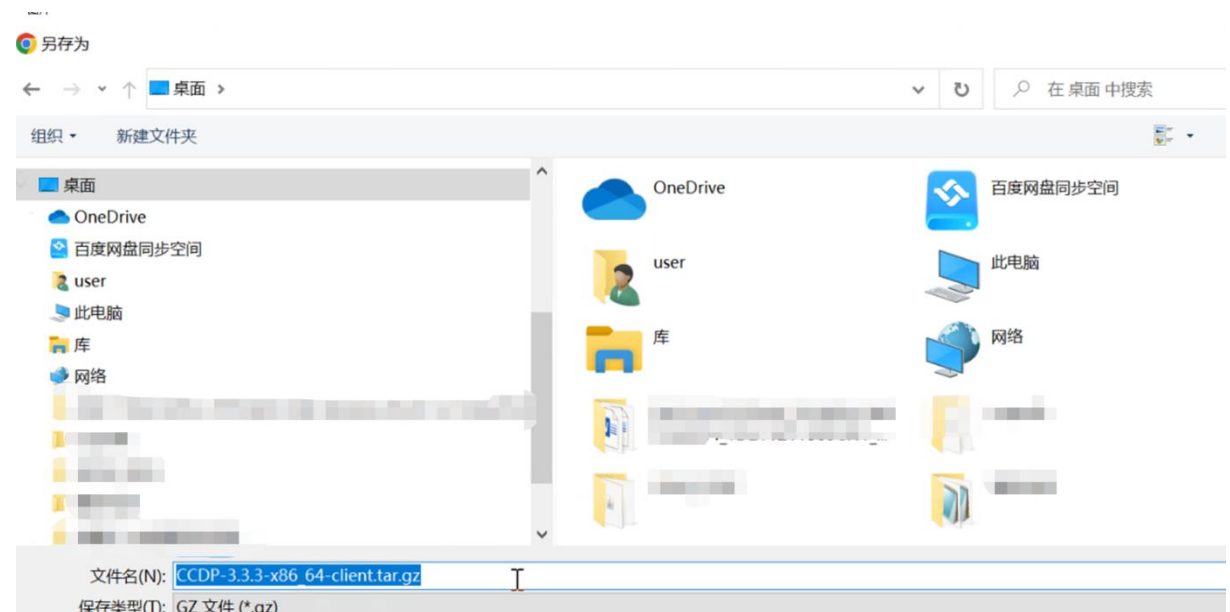
软件信息

产品版本	EMR-2.12.0	业务场景	数据湖	客户端	前往下载客户端
软件信息	<div>OpenLDAP 2.4.50Kerberos 1.16.2ZooKeeper 3.7.1HDFS 3.3.3YARN 3.3.3Hive 3.1.2Spark 3.2.2Ranger 2.2.0</div>				

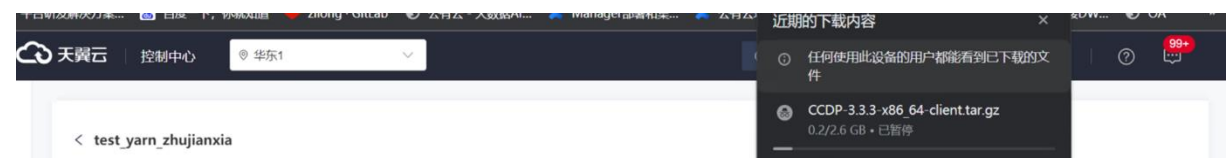
网络硬件

区域集群	华东1 (可用区1)	VPC ID	vpc-k2gplg0	安全组ID	sg-72ksozfn
硬件配置	MASTER: (x7.2large.4) 8 vCPU, 32GB3CORE: (x7.2large.4) 8 vCPU, 32GB3				

(4) 进入到组件的客户端安装包、MD5 值和配置文件的下载页面，选择下载路径，点击保存即可。



(5) 文件容量较大，为 2.6GB，请您耐心等待。



注意

- 下载完整客户端包含：客户端安装包、MD5 值和客户端配置文件。
- 客户端安装包包含组件：Hadoop（HDFS、YARN）、Hive、HBase、Flink、Kafka、Spark、ZooKeeper、Trino、Kyuubi、Kerberos。
- 客户端配置文件：包含当前环境公共空间下所有客户端配置文件。

6.4.4. 启动、停止所有集群服务

本章节主要介绍翼 MR Manager 的启动所有集群服务、停止所有集群服务操作。

启动所有集群服务

操作场景

一键启动所有集群服务。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 单击“更多操作>启动所有集群服务”，弹出确认操作弹框。



- (6) 单击“确定”，开始启动所有集群服务，等待所有集群服务完成启动后，单击“完成”即可。

注意

- 启动所有集群服务时按照集群依赖关系串行执行。
- 在执行过程中，可以取消当前操作，已启动的集群服务无法取消。

停止所有集群服务

操作场景

一键停止所有集群服务。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 单击“更多操作>停止所有集群服务”，弹出确认操作弹框。
- (6) 单击，开始停止所有集群服务，等待所有集群服务完成停止后，单击“完成”即可。



注意

- 停止所有集群服务时按照集群依赖关系串行执行。
- 在执行过程中，可以取消当前操作，已停止的集群服务无法取消。

6.4.5. 启动、停止单集群服务

本章节主要介绍翼 MR Manager 的单个集群服务的启动、停止。

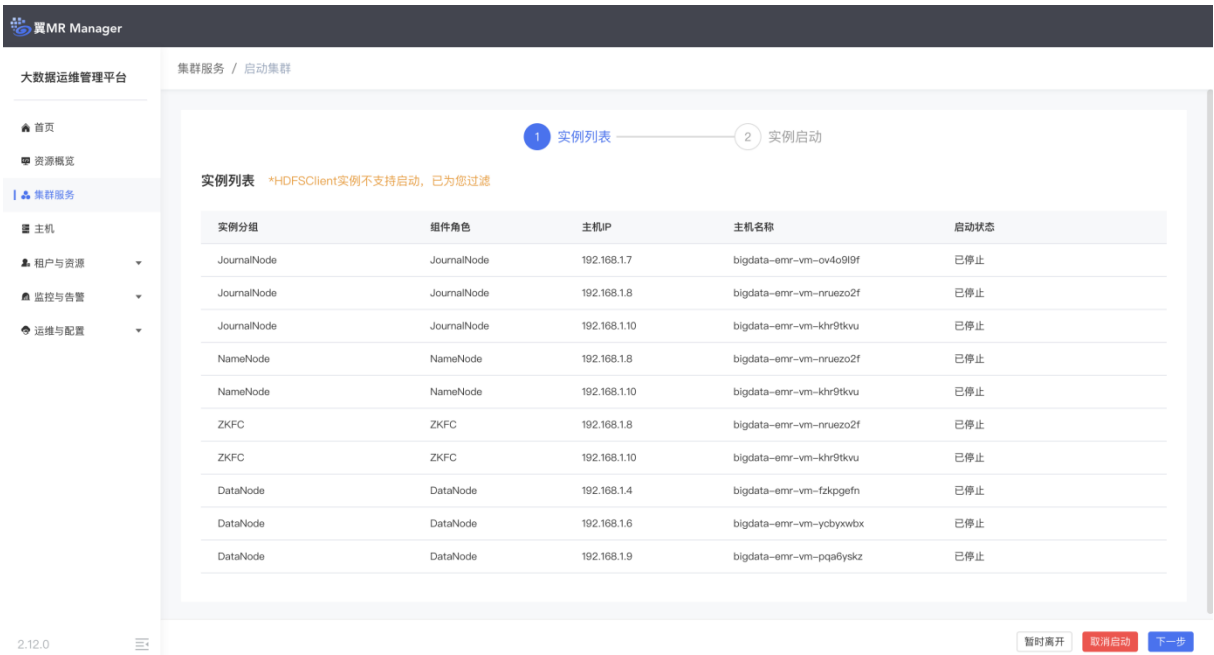
启动集群服务

操作场景

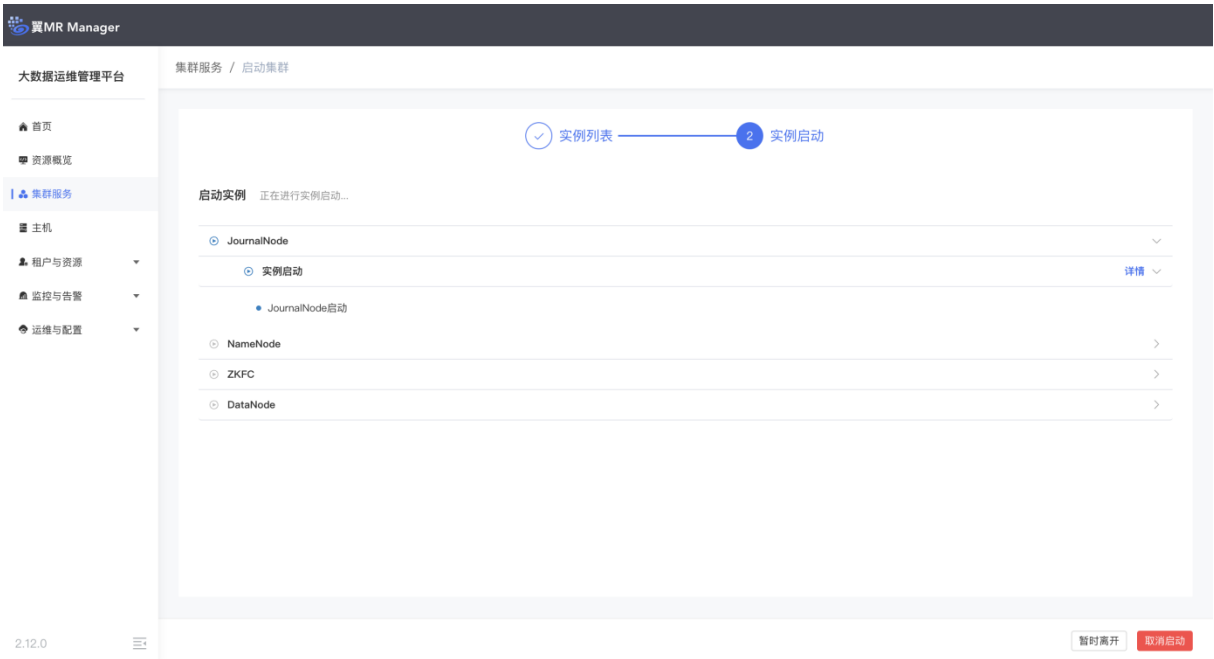
一键启动单个集群服务。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“运维操作>启动集群服务”，弹出确认操作弹框。
- (7) 单击“确定”，进入到启动集群服务页面，展示当前集群服务支持启动的实例列表。如图示：



(8) 单击“下一步”，按照实例分组顺序进行实例启动，等待所有实例完成启动后，单击“完成”即可。如图所示：



停止集群服务

操作场景

一键停止单个集群服务。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“运维操作>停止集群服务”，弹出确认操作弹框。
- (7) 单击“确定”，进入到停止集群服务页面，展示当前集群服务支持停止的实例列表。
- (8) 单击“下一步”，按照实例分组顺序进行实例停止，等待所有实例完成停止后，单击“完成”即可。

6.4.6. 滚动重启单个集群服务

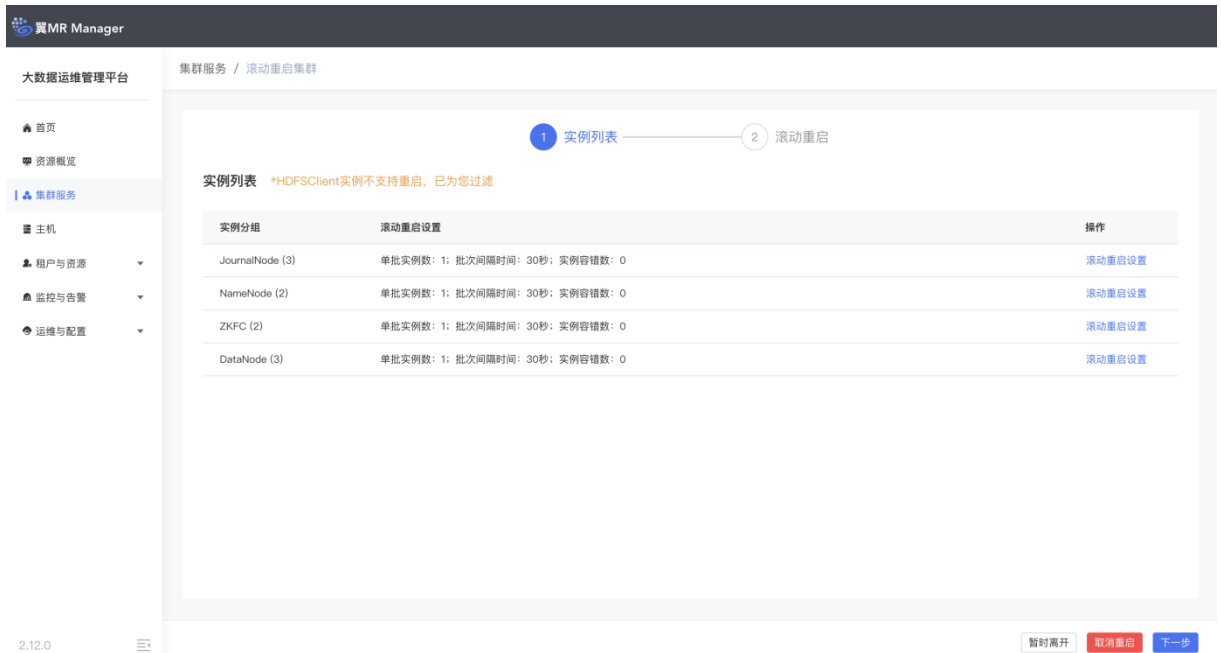
本章节主要介绍翼 MR Manager 的滚动重启集群操作。

操作场景

一键滚动重启单个集群服务。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“运维操作>滚动重启集群服务”，弹出确认操作弹框。
- (7) 单击“确定”，进入到滚动重启集群服务页面，展示当前集群服务支持滚动重启的实例列表。如图所示：



(8) 选择指定实例分组，单击“滚动重启设置”，对默认滚动重启设置进行修改，修改完成后单击“确定”，如图所示：

滚动重启设置

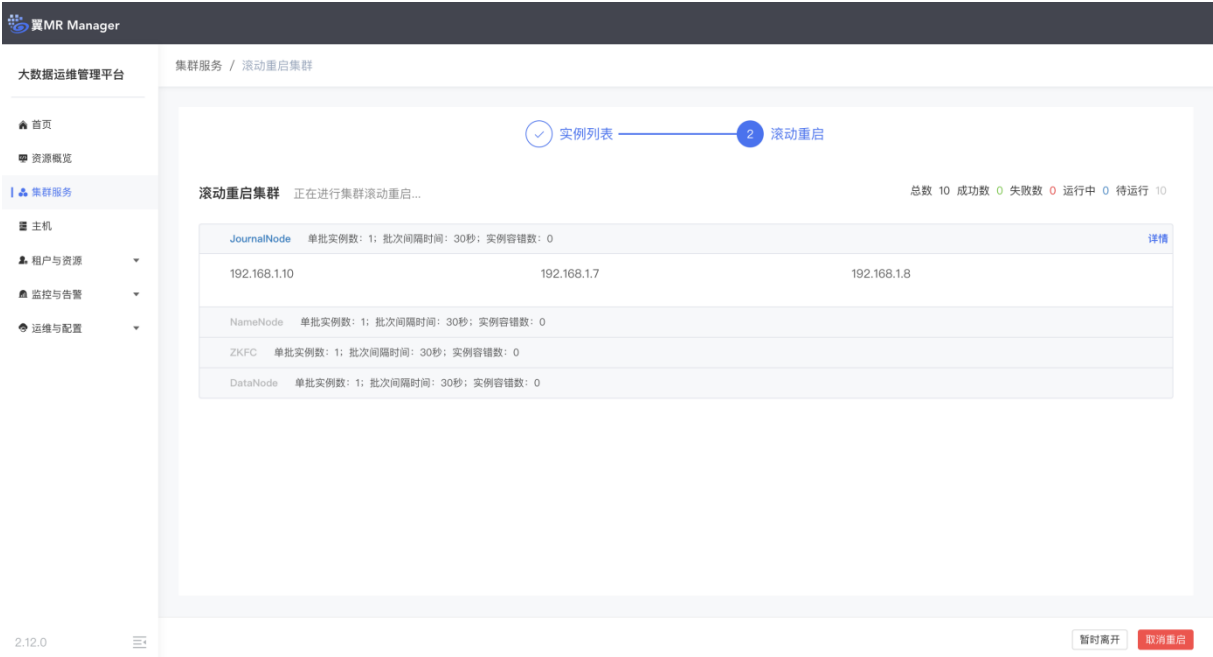
单批实例数

时间设置 ☒ 批次间隔时间 ☐ 批次等待时间

S(秒)

实例容错阈值

(9) 单击“下一步”，按照实例分组顺序进行滚动重启，等待所有实例完成重启后，单击“完成”即可。如图所示：



滚动重启参数说明

滚动重启参数说明如下表所示：

参数名称	描述
单批实例数	<div><div>1. 按照角色实例维度，每 1 个批次的并发实例数，例如 HDFS-DataNode。</div><div>2. 默认为 1，输入限制为 int 的最大值，仅支持输入正整数。</div><div>3. 当角色实例数小于并发数时，以当前已有实例数为准例如：并发数设置为 2，HDFS-DataNode 的实例数为 3；则第一批次执行的实例数是 2，第 2 批次执行的实例数则为 1。</div></div>
批次间隔时间	<div><div>1.上个批次成功后与下个批次开始的间隔时间。</div><div>2. 默认 30 秒，取值范围 1-1800，仅支持输入正整数；单位默认为：秒，不可改 举例：设置时间为 600S 批次间隔则表示，前 1 个批次运行完成后，等待 600S 后开始下一个批次。</div></div>
批次等待时间	<div><div>1. 上个批次开始后，开始下个批次的等待时间。</div><div>2. 默认 30 秒，取值范围 1-1800，仅支持输入正整数；单位默认为：秒，不可改 举例：设置时间为 600S 批次等待则表示，前 1 个排次开始运行，等待 600S 后不管前一个批次是否完成运行，都开始下一个批次的运行。</div></div>

实例容错阈值	<ol style="list-style-type: none"> 实例执行失败的容错次数。 默认为 0，输入限制为 int 的最大值，仅支持输入正整数例如：设置为 0 时，即表示任意一个角色实例的操作失败后，滚动操作终止设置为 2 时，即表示任意 3 个角色实例的操作失败后，滚动操作才会终止。
--------	---

6.4.7. 启动、停止、重启实例

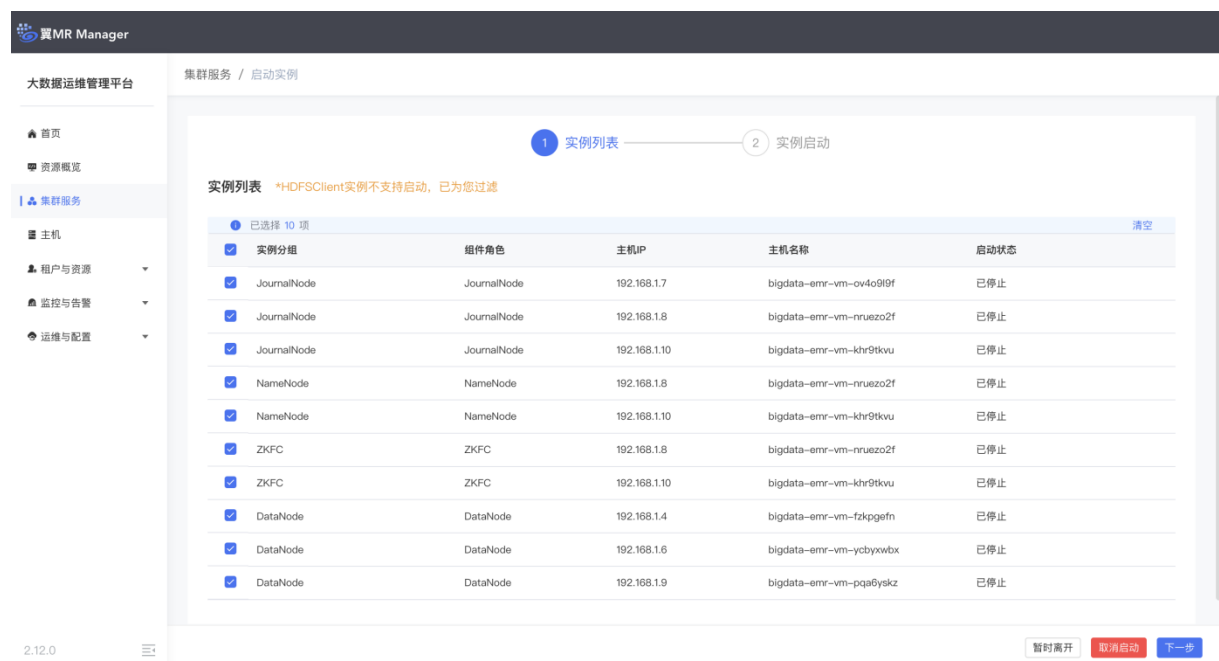
本章节主要介绍翼 MR Manager 的实例的启动、停止、重启操作。

启动实例

操作步骤

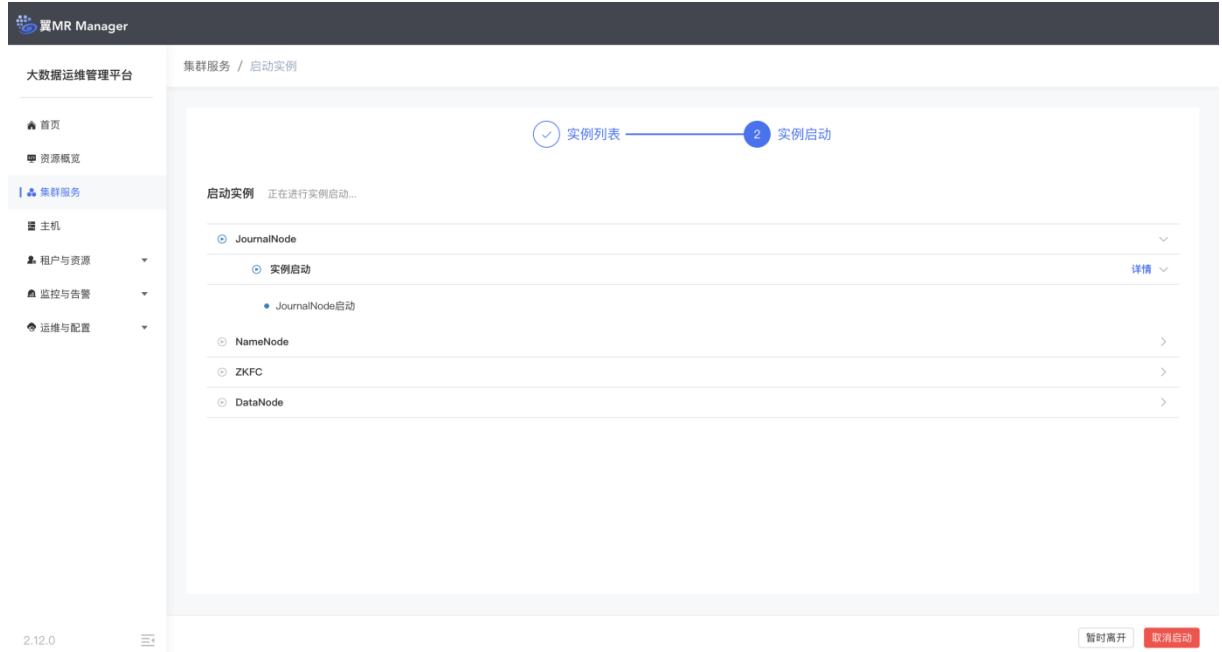
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 启动实例”，弹出确认操作弹框。
- (9) 单击“确定”，进入到启动实例页面，展示当前集群服务支持启动的实例列表。如图所示：

示：



- (10) 单击“下一步”，按照实例分组顺序进行实例启动，等待所有实例完成启动后，单击

“完成”即可。如图所示：



停止实例

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 停止实例”，弹出确认操作弹框。
- (9) 单击“确定”，进入到停止实例页面，展示当前集群服务支持停止的实例列表。
- (10) 单击“下一步”，按照实例分组顺序进行实例停止，等待所有实例完成停止后，单击“完成”即可。

重启实例

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。

- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 重启实例”，弹出确认操作弹框。
- (9) 单击“确定”，进入到重启实例页面，展示当前集群服务支持重启的实例列表。
- (10) 单击“下一步”，按照实例分组顺序进行实例重启，等待所有实例完成重启后，单击“完成”即可。

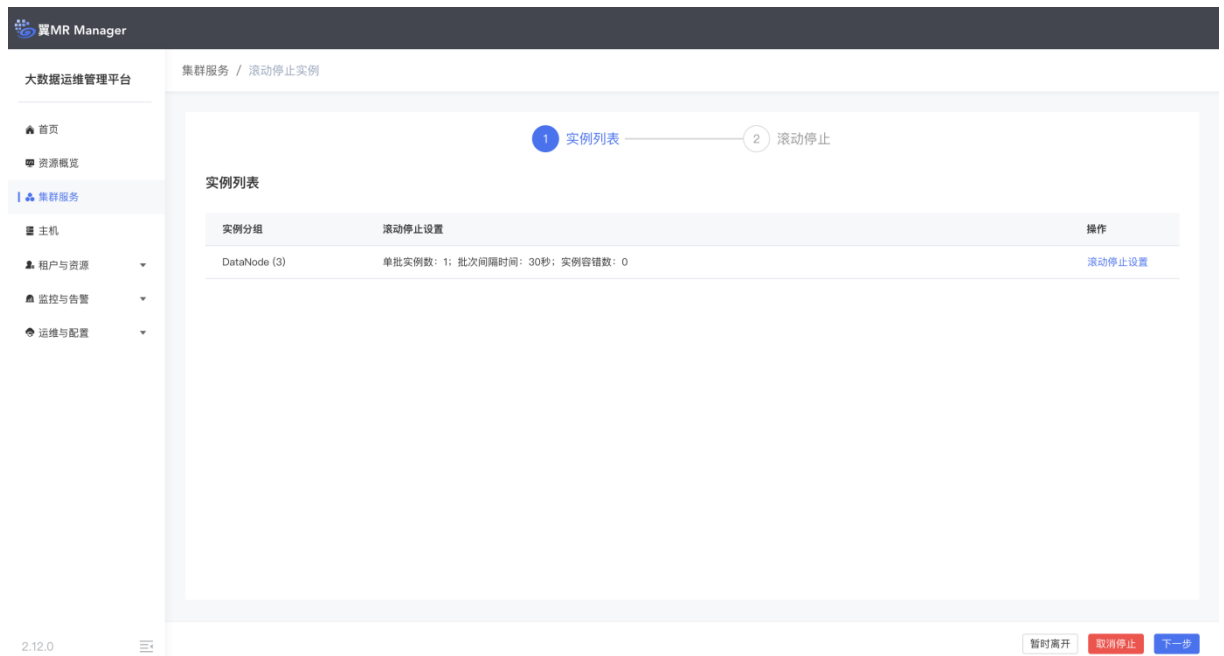
6.4.8. 滚动停止、滚动重启实例

本章节主要介绍翼 MR Manager 的实例的滚动停止、滚动重启操作。

滚动停止实例

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 滚动停止实例”，弹出确认操作弹框。
- (9) 单击“确定”，进入到滚动停止实例页面。如图所示：



- (10) 选择指定实例分组，单击“滚动停止设置”，对默认滚动停止设置进行修改，修改完成后单击“确定”。如图所示：

滚动停止设置

单批实例数 ?

1

^

v

时间设置

☒ 批次间隔时间 ?☐ 批次等待时间 ?

30

^

v

S(秒)

实例容错阈值 ?

0

^

v

取消

确定

(11) 单击“下一步”，按照实例分组顺序进行滚动停止，等待所有实例完成停止后，单击“完成”即可。如图所示：

大数据运维管理平台

集群服务

主机

租户与资源

监控与告警

运维与配置

集群服务 / 滚动停止实例

实例列表

2 滚动停止

滚动停止实例 正在进行实例滚动停止...

总数 3 成功数 0 失败数 0 运行中 0 待运行 3

DataNode

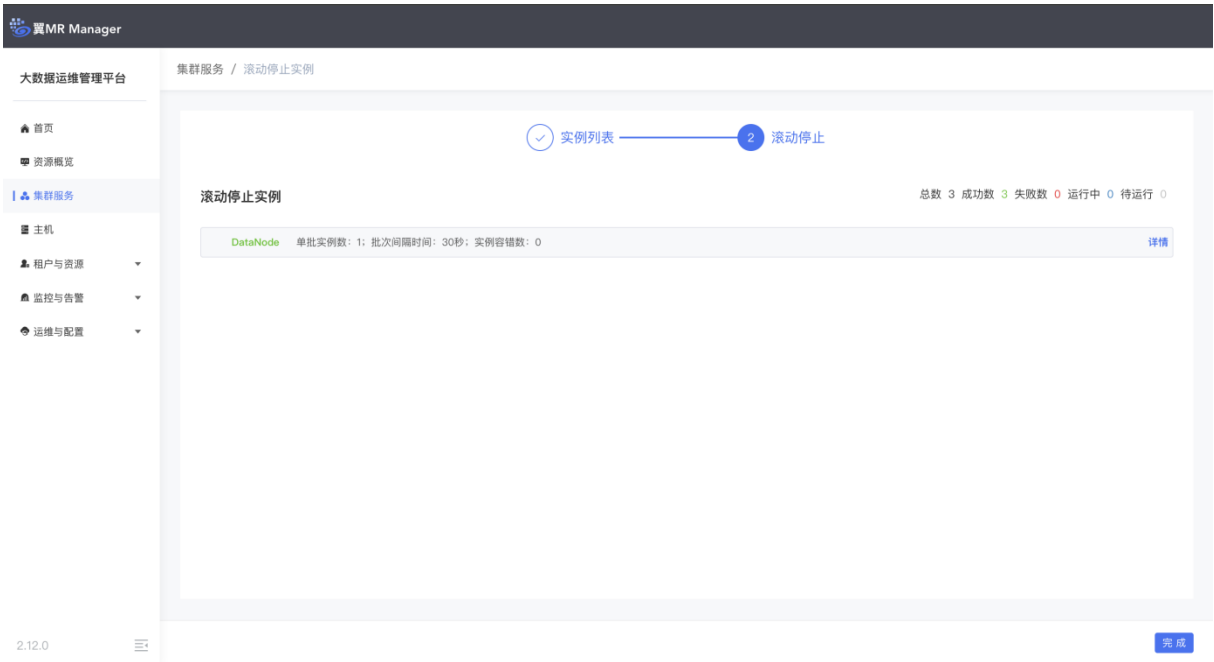
单批实例数: 1; 批次间隔时间: 30秒; 实例容错数: 0

详情

192.168.1.4	192.168.1.6	192.168.1.9
-------------	-------------	-------------

暂时离开

取消停止



滚动重启实例

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 滚动重启实例”，弹出确认操作弹框。
- (9) 单击“确定”，进入到滚动重启实例页面。
- (10) 选择指定实例分组，单击“滚动重启设置”，对默认滚动重启设置进行修改，修改完成后单击“确定”。
- (11) 单击“下一步”，按照实例分组顺序进行滚动重启，等待所有实例完成重启后，单击“完成”即可。

滚动设置参数说明

滚动设置参数说明如下表所示：

参数名称	描述
单批实例数	<ol style="list-style-type: none">1. 按照角色实例维度，每 1 个批次的并发实例数，例如 HDFS-DataNode。2. 默认为 1，输入限制为 int 的最大值，仅支持输入正整数。

	3. 当角色实例数小于并发数时，以当前已有实例数为准例如：并发数设置为 2，HDFS-DataNode 的实例数为 3；则第一批执行的实例数是 2，第 2 批次执行的实例数则为 1。
批次间隔时间	1. 上个批次成功后与下个批次开始的间隔时间。 2. 默认 30 秒，取值范围 1-1800，仅支持输入正整数；单位默认为：秒，不可改举例：设置时间为 600S 批次间隔则表示，前 1 个批次运行完成后，等待 600S 后开始下一个批次。
批次等待时间	1. 上个批次开始后，开始下个批次的等待时间。 2. 默认 30 秒，取值范围 1-1800，仅支持输入正整数；单位默认为：秒，不可改举例：设置时间为 600S 批次等待则表示，前 1 个批次开始运行，等待 600S 后不管前一个批次是否完成运行，都开始下一个批次的运行。
实例容错阈值	1. 实例执行失败的容错次数。 2. 默认为 0，输入限制为 int 的最大值，仅支持输入正整数例如：设置为 0 时，即表示任意一个角色实例的操作失败后，滚动操作终止设置为 2 时，即表示任意 3 个角色实例的操作失败后，滚动操作才会终止。

6.4.9. 实例置维护/取消维护

本章节主要介绍翼 MR Manager 的实例的置维护、取消维护操作。

置维护

操作场景

用户需要对某个实例进行置维护。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“角色实例”tab。
- (7) 勾选需要操作角色实例前的复选框。
- (8) 单击“操作已选项 > 置维护”，弹出置维护操作弹框。如图所示：

置维护 置维护后，将屏蔽实例的告警信息

实例数1

* 备注

请输入本次操作的原因

取消

确定

(9) 输入备注，单击“确定”即可。注意

- 实例置维护后，该实例的健康状态显示为维护中。
- 实例置维护后，则会屏蔽该实例的告警信息。

取消维护

操作场景

用户需要对某个实例取消置维护。

操作步骤

- 登录翼 MR 管理控制台。
- 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- 单击“角色实例”tab。
- 勾选需要操作角色实例前的复选框。
- 单击“操作已选项 > 取消维护”，弹出取消维护操作弹框。如图所示：

取消维护

实例数1

* 备注

请输入本次操作的原因

取消

确定

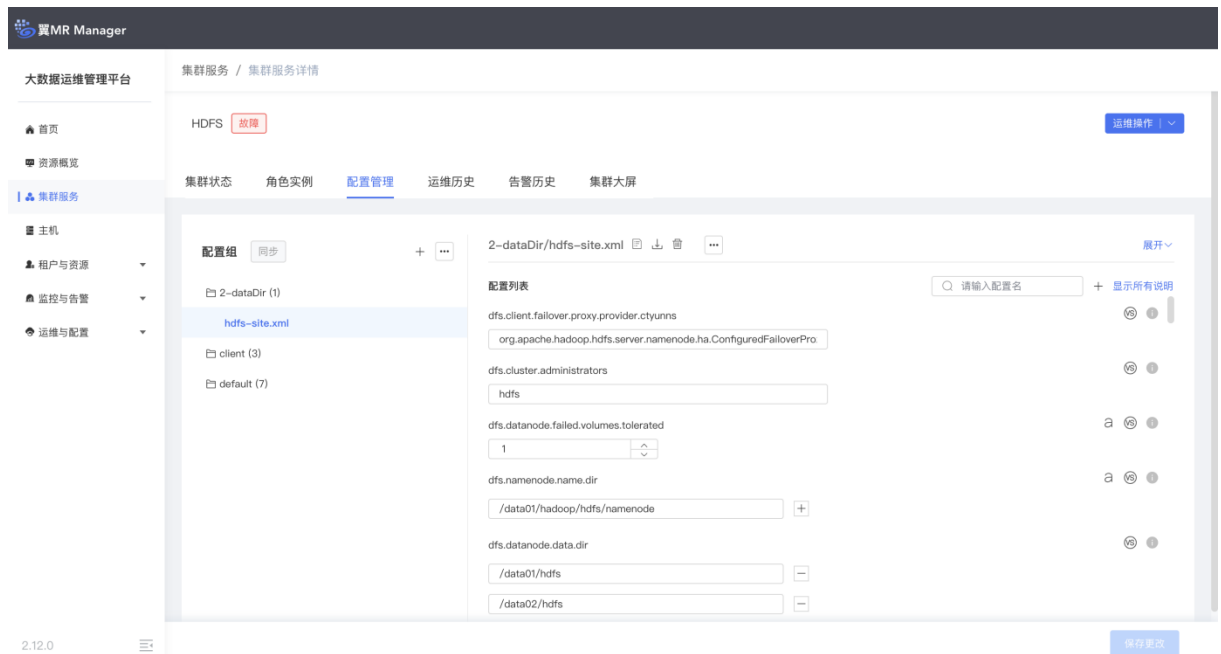
(9) 输入备注，单击“确定”即可。

6.4.10. 查看集群服务配置

本章节主要介绍翼 MR Manager 的集群服务配置的查看操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab，即可查看该集群服务的所有配置。如图所示：



6.4.11. 新增配置组

本章节主要介绍翼 MR Manager 的集群服务新增配置组的操作。

操作场景

用户可以根据不同的分组标准，对集群服务中的主机进行分组，一个配置分组下的主机共享一套配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。

- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击“新增配置组”，进入新增配置组页面。
- (8) 根据需要选择分组方式，选择主机，单击右下角“保存”按钮即可。如图所示：

运维与配置 / 配置管理 / 新增配置组

分组方式

☐ 全部
 ☐ 按组件角色
 ☐ 按实例分组
 ☐ 按主机配置
 ☐ 单个主机
 ☒ 自定义

配置组名称

主机列表 (6)

<input type="checkbox"/>	主机IP	主机名称	组件角色	主机配置
<input type="checkbox"/>	192.168.0.49	bigdata-emr-vm-bbqwvney	NameNode.JournalNode.ZKFC	虚拟核数:8C / 内存:32GB
<input type="checkbox"/>	192.168.0.52	bigdata-emr-vm-0tr2bdvt	NameNode.JournalNode.ZKFC	虚拟核数:8C / 内存:32GB
<input type="checkbox"/>	192.168.0.55	bigdata-emr-vm-mxiofiw	DataNode	虚拟核数:8C / 内存:32GB
<input type="checkbox"/>	192.168.0.51	bigdata-emr-vm-w9utlsdt	DataNode	虚拟核数:8C / 内存:32GB
<input type="checkbox"/>	192.168.0.54	bigdata-emr-vm-pkqc8sij	HDFSClient.JournalNode	虚拟核数:8C / 内存:32GB
<input type="checkbox"/>	192.168.0.53	bigdata-emr-vm-v6fj0ezf	DataNode	虚拟核数:8C / 内存:32GB

注意

- 分组方式为：全部、按组件角色、按实例分组、按主机配置时，配置组名称默认且不可改。
- 分组方式为：单个主机时，需要选择要分组的主机，配置组名称为所选主机的名称且不可改。
- 分组方式为：自定义时，需要选择分组的主机，配置组名称可以自定义但不能与其他配置组名称相同。

6.4.12. 查看配置组详情

本章节主要介绍翼 MR Manager 的集群服务查看配置组详情的操作。

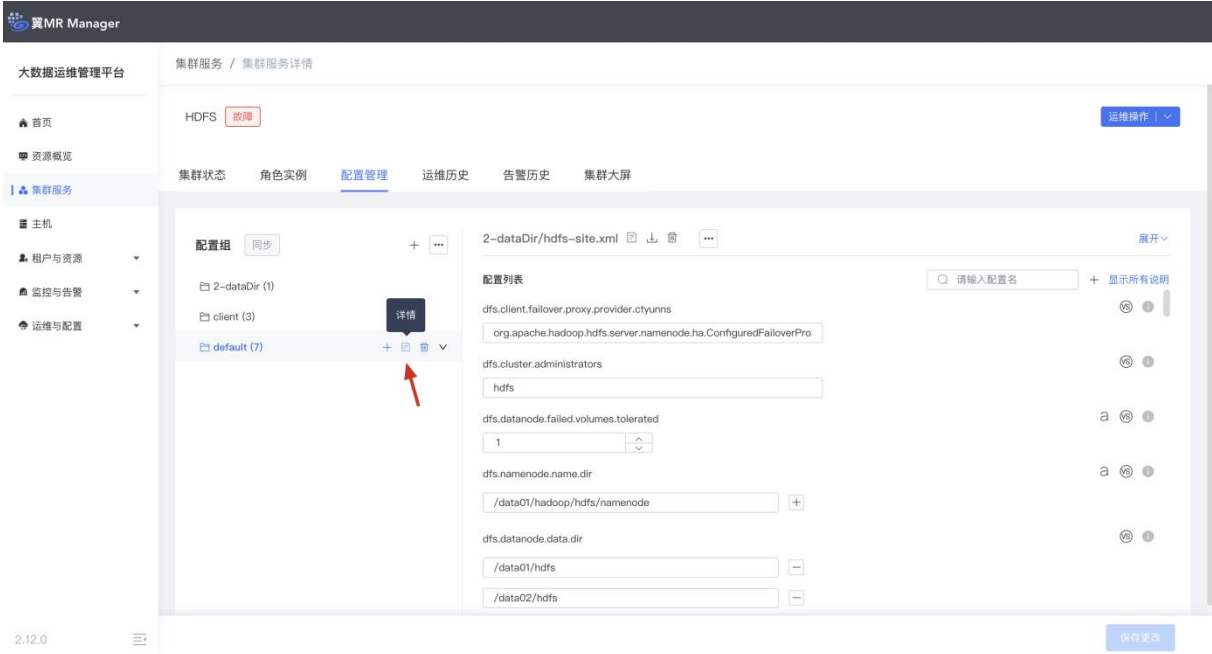
操作场景

用户可以查看不同配置组的详细信息，了解该配置组所包含的主机信息。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。

- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要查看的配置组，鼠标移入到配置组名称上，单击右侧的“详情”图标，进入到配置组详情页面。如图所示：



运维与配置 / 配置管理 / 配置组详情

分组方式 全部主机

配置组名称 default

主机列表 (6)

主机名	IP	组件角色	主机配置
bigdata-emr-vm-bbqwvney	192.168.0.49	NameNode JournalNode ZKFC	虚拟核数:8C/内存:32GB
bigdata-emr-vm-0tr2bdvt	192.168.0.52	NameNode JournalNode ZKFC	虚拟核数:8C/内存:32GB
bigdata-emr-vm-mxiojfw	192.168.0.55	DataNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-w9utlsdt	192.168.0.51	DataNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-pkqc8sij	192.168.0.54	HDFSClient JournalNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-v6fj0ezf	192.168.0.53	DataNode	虚拟核数:8C/内存:32GB

关闭

6.4.13. 删除配置组

本章节主要介绍翼 MR Manager 的集群服务删除配置组的操作。

操作场景

用户可以删除不需要的配置组。

前置条件

配置组下的配置文件数量必须为 0 才能删除。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要删除的配置组，鼠标移入到配置组名称上，单击右侧的“删除”图标，出现删除确认弹框。



- (8) 单击“确定”即可删除该配置组。

6.4.14. 新增配置

本章节主要介绍翼 MR Manager 的集群服务新增配置的操作。

在翼 MR Manager 中支持 4 种方式的新增配置。

方式 1：新建

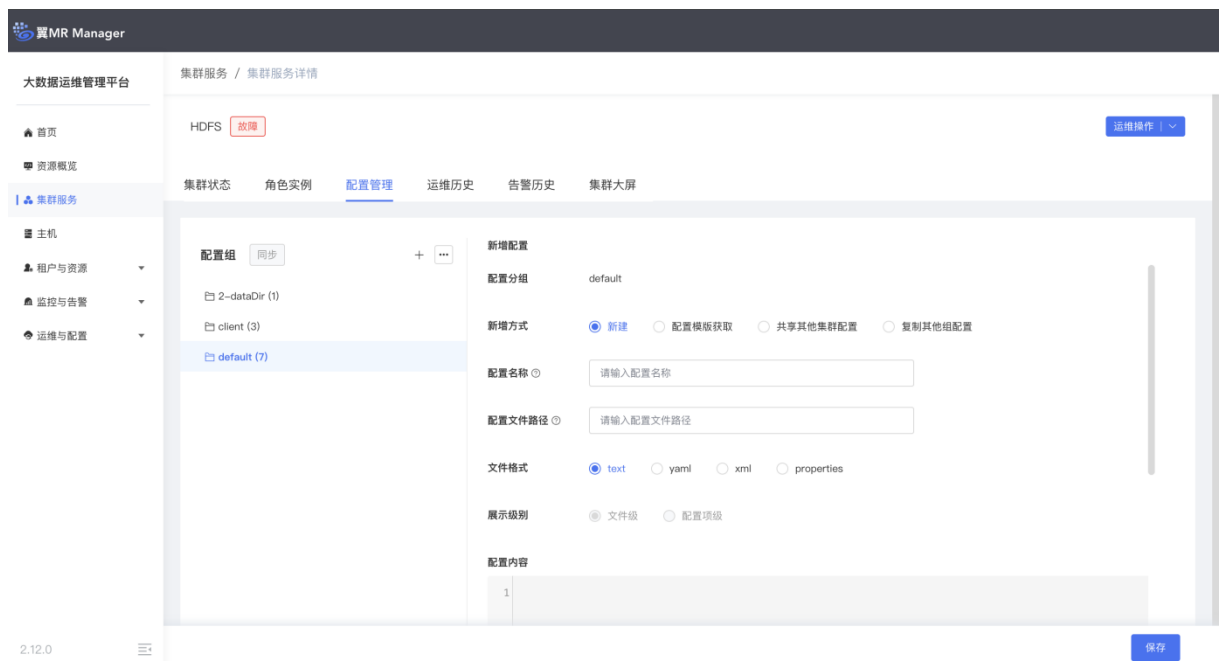
操作场景

无需共享其他集群的配置，且配置模板不符合使用要求时，用户为集群服务新增一条全新的配

置文件。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要新增配置的配置组，鼠标移入到配置组名称上，单击右侧的“新增配置”图标，页面右侧为新增配置。如图所示：



- (8) 选择新增方式为“新建”，输入配置名称、配置文件路径、选择文件格式，输入配置内容或者添加配置项。
- (9) 单击“保存”完成配置新增。

注意事项

- (1) 配置名称：仅支持输入英文、数组、“_.”，且不能以数字开头。
- (2) 文件路径：仅支持输入英文、数组、“_.”；且配置路径必须在白名单中。
- (3) 文件格式为 text 时，仅需要输入配置内容即可。如图所示：

文件格式

☒ text

☐ yaml

☐ xml

☐ properties

展示级别

☒ 文件级

☐ 配置项级

配置内容

1

(4) 文件格式非 text，为 yaml、xml 或者 properties 时，需要：选择展示级别、选择配置规则组、配置项。如图所示：

文件格式

☐ text

☒ yaml

☐ xml

☐ properties

展示级别

☐ 文件级

☒ 配置项级

配置规则组

请选择配置规则组

▼

配置项

配置项名称	配置值	配置描述	操作
<input type="text"/>	<input type="text"/>	<input type="text"/>	保存 取消
<div>+ 添加</div>			

(5) 展示级别配置项级，则配置的展示为一条条的配置项。如图所示：

default/HDFS_vars.yaml 文件 下载 删除 ...

收起^

配置文件路径: /etc/HDFS_vars.yaml

文件格式: YAML

展示级别: 配置项级

配置规则组:

配置列表

Q 请输入配置名

+ 显示所有说明

dfs_clusterId

VS 删除 信息

ctyunns

active_namenode

VS 删除

bigdata-emr-vm-oynkzmji

kdc_master

VS 删除

192.168.0.87

ldap_uri

VS 删除

ldap://bigdata-emr-vm-n7ct3uqr:389,ldap://bigdata-emr-vm-ja7si

(6) 文件级，则配置的展示为文本样。如图所示：

default/hadoop-env.sh 文件 下载 删除 ...

收起^

配置文件路径: /etc/hadoop/conf/hadoop-env.sh

文件格式: TEXT

展示级别: 文件级

配置内容

```
1 export JAVA_HOME="/usr/jdk64/current"
2 export HADOOP_HOME="/usr/local/hadoop3"
3 export HADOOP_HEAPSIZE="4096"
4 NN_OPTS="-Xmx12g -Xms12g"
5 DN_OPTS="-Xmx4g -Xms4g"
6 export JOURNALNODE_OPTS="-Xmx4g -Xms4g "
7 export HADOOP_OPTS="-Djava.net.preferIPv4Stack=true ${HADOOP_OPTS}"
8 USER="$(whoami)"
9 export HADOOP_LOG_DIR=/var/log/hadoop/$USER
10 export
   HADOOP_CLASSPATH=$HADOOP_CLASSPATH:$HADOOP_HOME/share/hadoop/mapreduce/*:$HADOOP_HOME/share/hadoop/mapredu
   ce/lib/*:$HADOOP_HOME/share/hadoop/common/*:$HADOOP_HOME/share/hadoop/common/lib/*:$HADOOP_HOME/share/hado
   op/yarn/*:$HADOOP_HOME/share/hadoop/yarn/lib/*:$HADOOP_HOME/share/hadoop/hdfs/*:$HADOOP_HOME/share/hadoop/
   hdfs/lib/*:/usr/lib/hadoop/lib/*
11 export JMX_OPTS="-Dcom.sun.management.jmxremote.authenticate=false -
```

(7) 配置规则组：即为系统中所维护的配置规则，选择配置规则组后，下方的配置项则展示该配置规则组中所包含的配置项；可编辑配置规则组带出的配置项的配置值和描述；继续添加新的配置项。

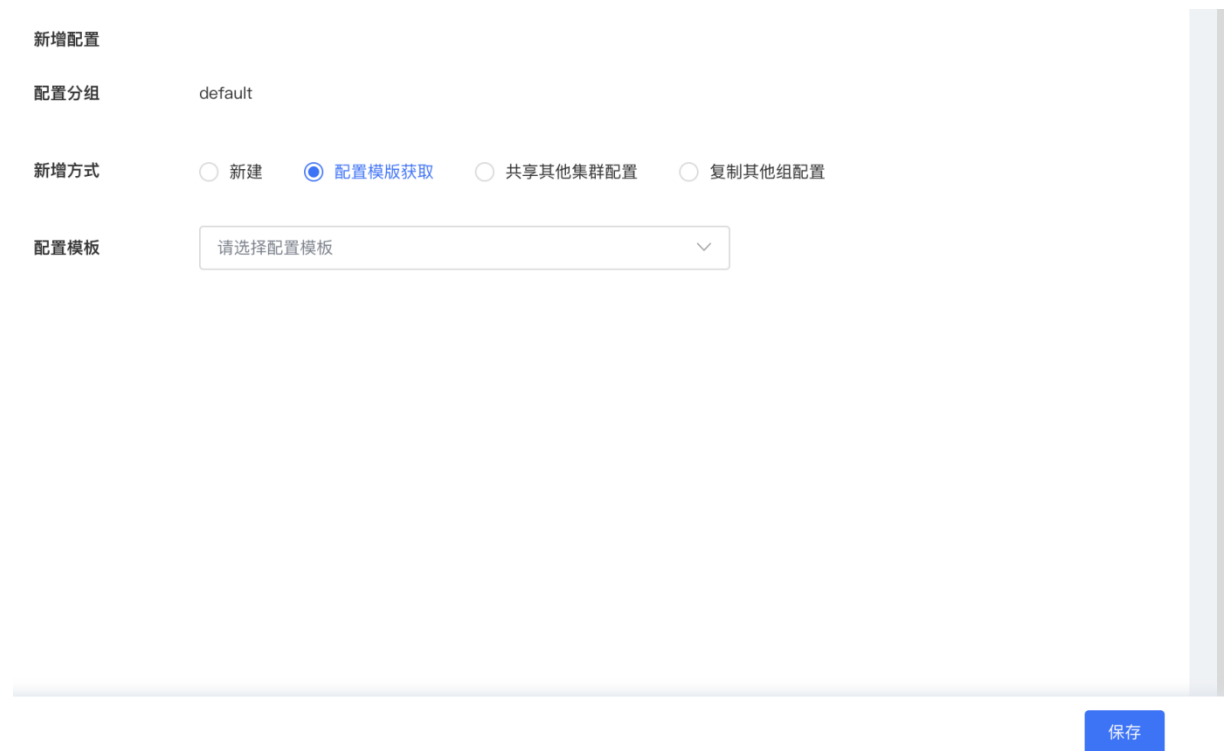
方式 2：配置模板获取

操作场景

直接使用配置模板为集群服务新增配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要新增配置的的配置组，鼠标移入到配置组名称上，单击右侧的“新增配置”图标，页面右侧为新增配置。
- (8) 选择新增方式为“配置模板获取”。如图所示：



- (9) 选择配置名称。
- (10) 单击“保存”完成配置新增。

方式 3：共享其他集群配置

操作场景

当前集群服务与其他集群的配置存在依赖管理，可直接共享其他集群服务的配置。

操作步骤

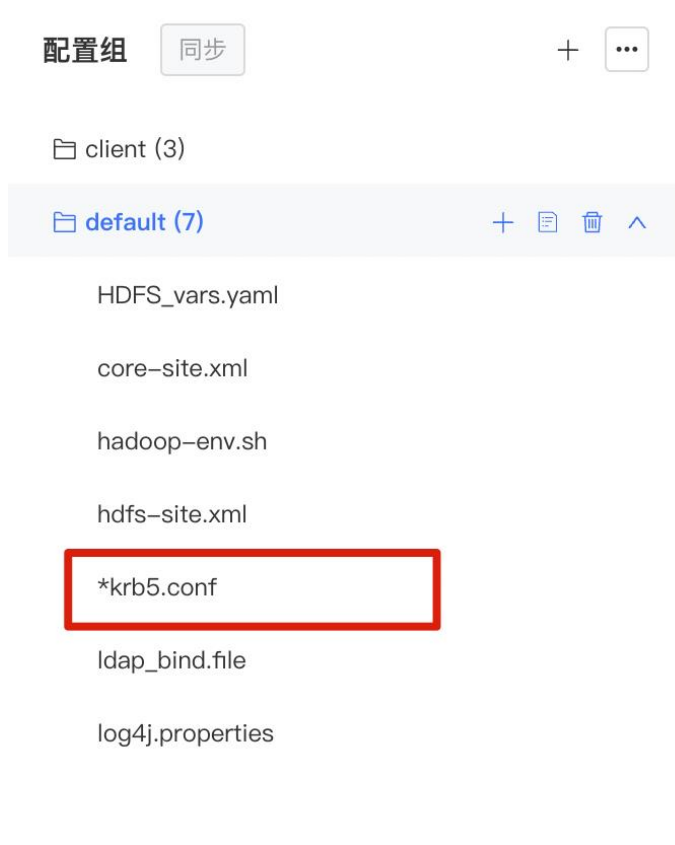
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要新增配置的配置组，鼠标移入到配置组名称上，单击右侧的“新增配置”图标，页面右侧为新增配置。
- (8) 选择新增方式为“共享其他集群配置”。如图所示：



- (9) 选择要共享的集群和配置。
- (10) 单击“保存”完成配置新增。

注意事项

- (1) 与其他集群共享的配置，在配置文件名称前会有“*”进行标识。如图所示：



(2) 与其他集群共享的配置，在任一个集群服务下进行配置修改后，共享该配置的集群服务的配置也会同步进行修改。

方式 4：复制其他组配置

操作场景

当前配置在其他组已存在，无需再新建，可以直接从其他组复制到当前组。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 选择要新增配置的配置组，鼠标移入到配置组名称上，单击右侧的“新增配置”图标，页面右侧为新增配置。
- (8) 选择新增方式为“复制其他组配置”。如图所示：

新增配置

配置分组

default

新增方式

☐ 新建

☐ 配置模版获取

☐ 共享其他集群配置

☒ 复制其他组配置

其他组配置

请选择

▼

保存

- 选择要复制的配置组和配置。
- 单击“保存”完成配置新增。

注意事项

对复制的配置文件进行修改，或者是修改被复制的配置，不会互相影响。

6.4.15. 查看配置详情

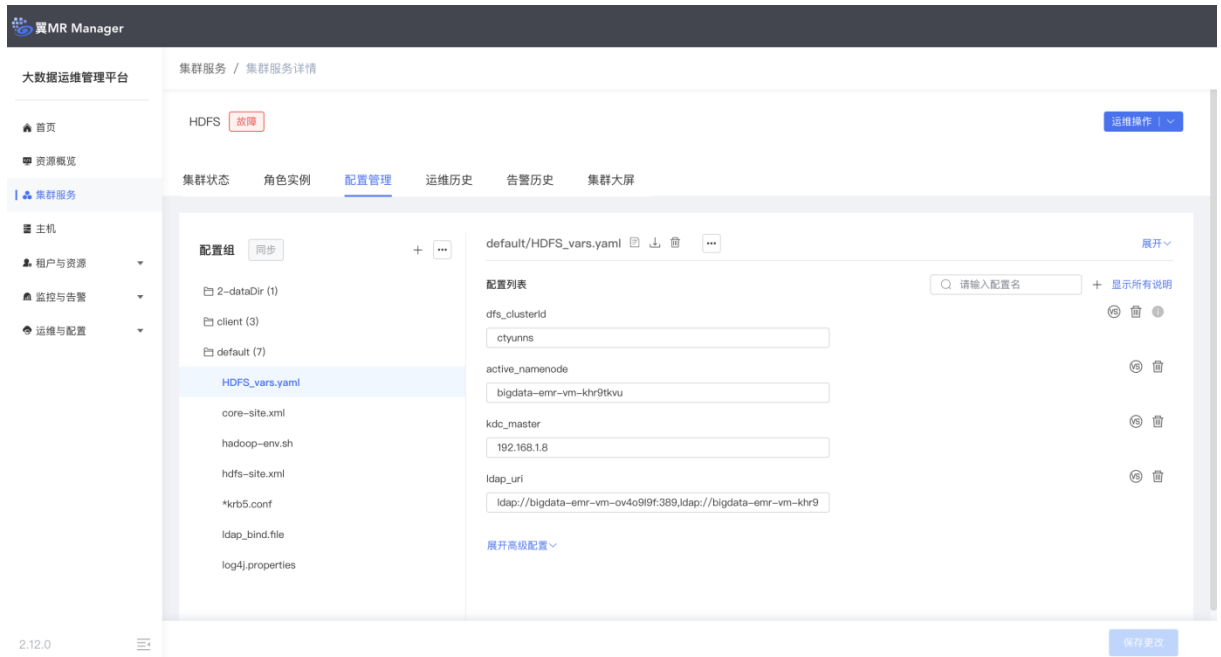
本章节主要介绍翼 MR Manager 的集群服务查看配置详情的操作。

操作场景

用户可以查看某个配置的详细信息。

操作步骤

- 登录翼 MR 管理控制台。
- 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- 单击“配置管理”tab。
- 单击配置组，单击要查看详情的配置名称，页面右侧展示该配置的详细信息。如图所示：



6.4.16. 编辑配置

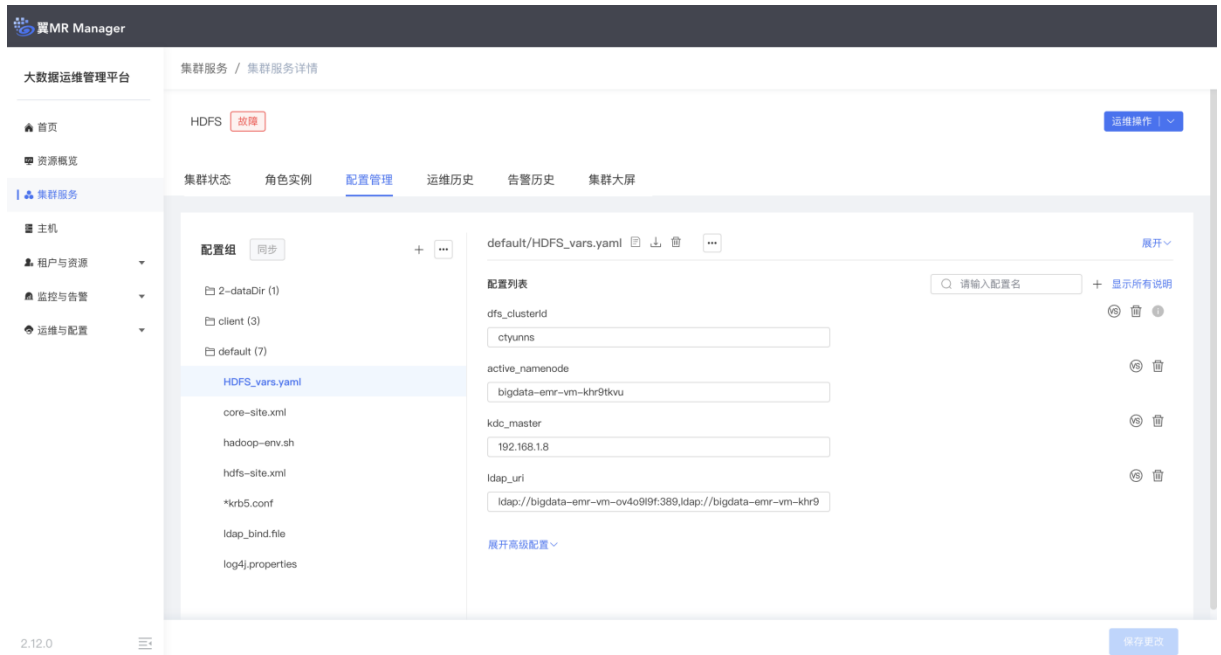
本章节主要介绍翼 MR Manager 的集群服务编辑配置的操作。

操作场景

用户可以对配置文件进行修改。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击要修改的配置名称，页面右侧展示该配置的详细信息。如图所示：

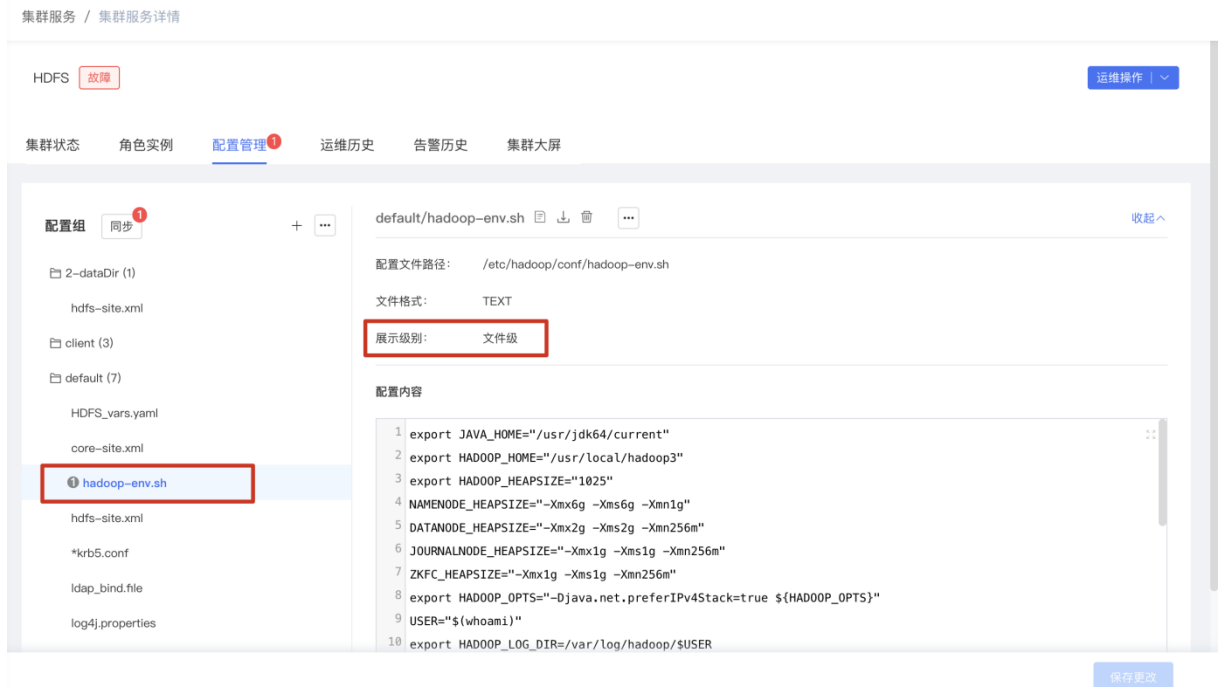


(8) 直接对配置文件修改即可，修改完成后点击“保存更改”。

注意事项

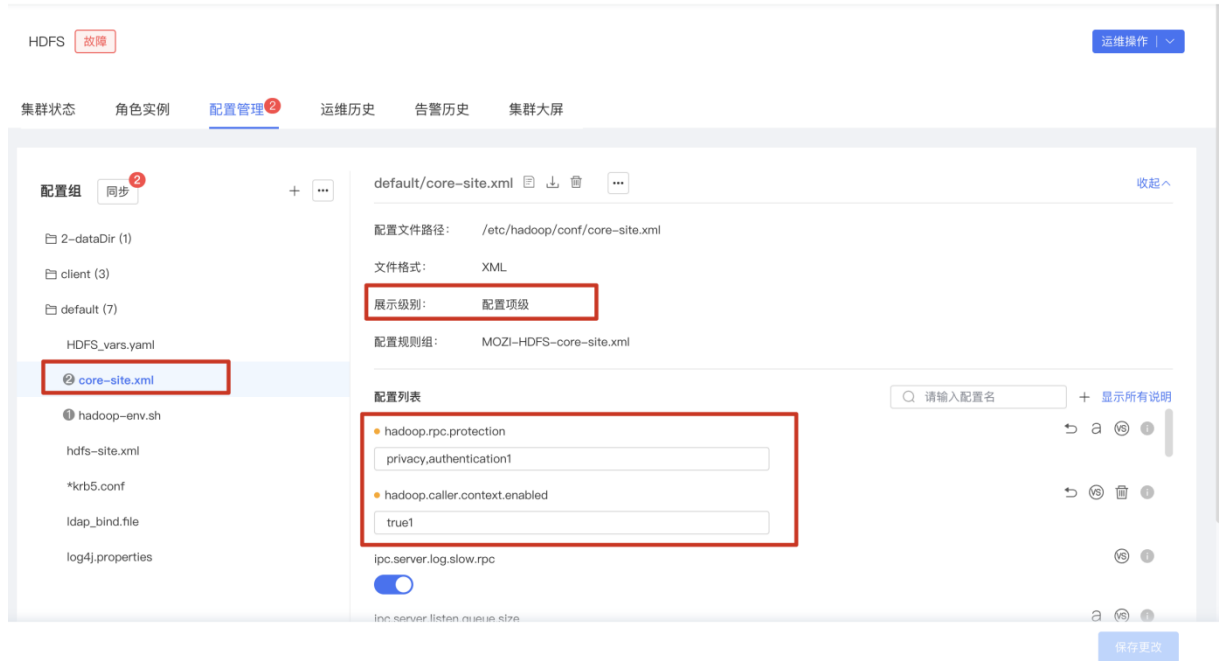
(1) 展示级别为：文件级。

- 直接在文本框中进行编辑或修改。
- 修改后且当前配置文件未同步时，配置名称前展示数字 1。如图所示：



(2) 展示级别为：配置项级。

- 需要逐条修改配置项。
- 修改后且当前配置文件未同步时，配置名称前展示所修改的配置项数，所修改的配置项名称前展示修改标识。如图所示：



6.4.17. 配置项对比

本章节主要介绍翼 MR Manager 的集群服务配置项对比的操作。

操作场景

用户可以查看同一集群下相同配置项名称在不同配置组中的值。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 选择要对比的配置项，单击该配置项名称右侧的“配置项对比按钮”，出现配置项对比弹框。如图所示：

HDFS 故障

运维操作

集群状态 角色实例 配置管理 运维历史 告警历史 集群大屏

配置组 同步

2-dataDir (1)

client (3)

core-site.xml

hadoop-env.sh

hdfs-site.xml

default (7)

client/core-site.xml

配置列表

hadoop.rpc.protection

privacy,authentication

io.compression.codec.lzo.class

com.hadoop.compression.lzo.LzoCodec

io.compression.codecs

org.apache.hadoop.io.compress.GzipCodec,org.apache.hadoop.io.c...

ipc.client.connection.maxidletime

30000

hadoop.security.authentication

kerberos

hadoop.security.authorization

true

展开

请输入配置名

显示所有说明

配置对比

保存更改

配置项对比

×

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

修改相同值

default

30000

5/500

client

20000

5/500

关闭

保存更改

注意事项

- (1) 配置项对比弹框中，当多个配置项的值不相同时，进行逐条展示。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

修改相同值

default

30000

5/500

client

20000

5/500

关闭

保存更改

(2) 配置项对比弹框中，当多个配置项的值相同时，进行合并展示。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

修改单个值

default

client

30000

关闭

保存更改

(3) 为可直接对配置项值进行修改，并支持修改相同值，也支持修改单个配置项的值。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

default

client

30000

修改单个值

关闭

保存更改

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

default

client

30000

5/500

30000

5/500

修改相同值

关闭

保存更改

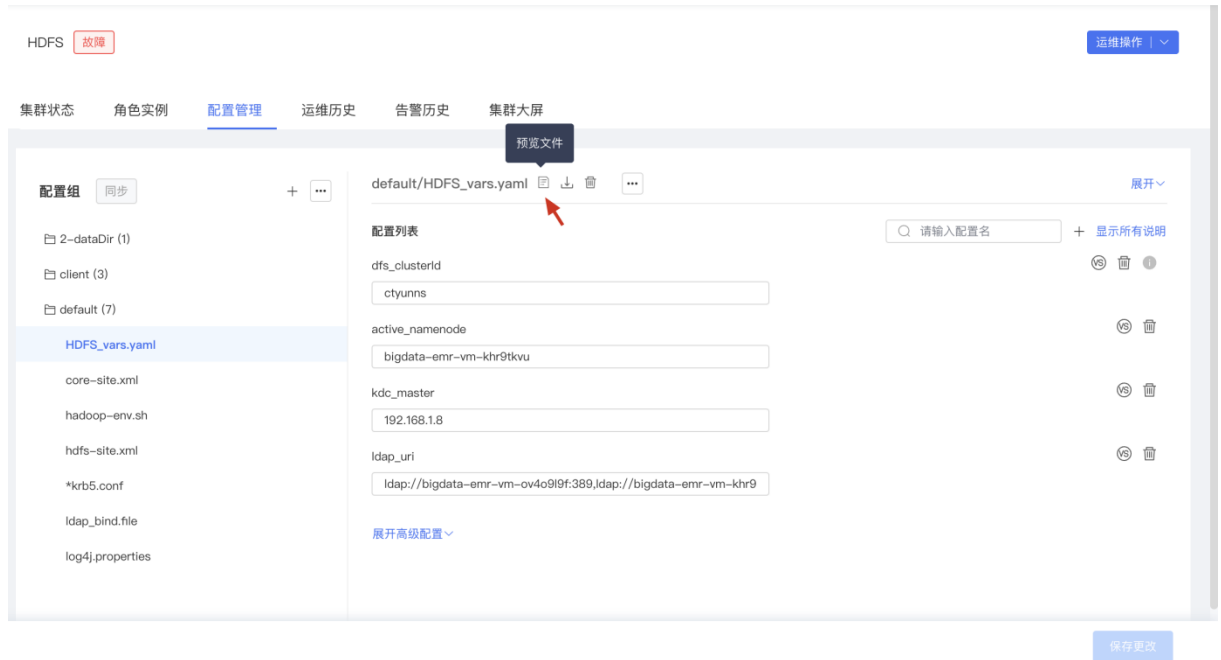
6.4.18. 配置文件预览

本章节主要介绍翼 MR Manager 的集群服务配置文件预览的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 单击“预览文件”，出现配置文件预览弹框。如图所示：

集群服务 / 集群服务详情



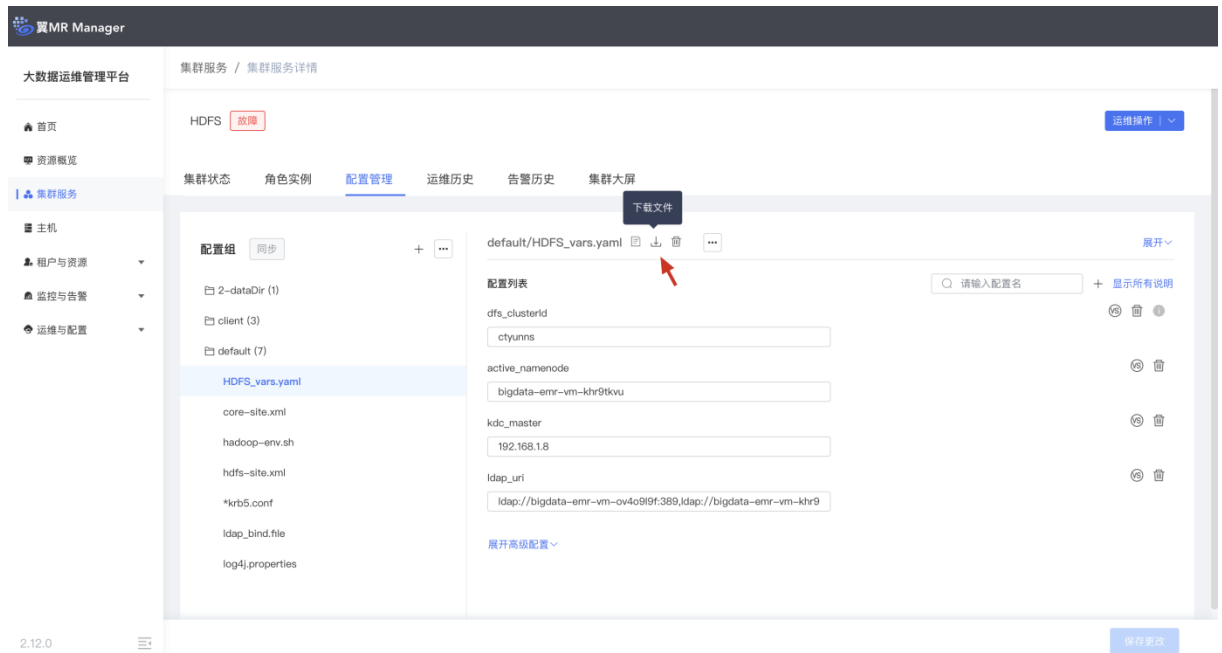


6.4.19. 下载配置文件

本章节主要介绍翼 MR Manager 的集群服务配置文件下载的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 单击“下载文件”，等待下载完成即可。如图所示：



6.4.20. 配置文件历史对比

本章节主要介绍翼 MR Manager 的集群服务配置文件历史对比的操作。

操作场景

支持用户查看当前配置文件与不同历史版本的差异。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 单击“更多 > 历史对比”，出现历史对比弹框。如图所示：

翼MR Manager

大数据运维管理平台

集群服务 / 集群服务详情

HDFS 故障 运维操作

集群状态 角色实例 **配置管理** 运维历史 告警历史 集群大屏

配置组 同步 + ...

- 2-dataDir (1)
- client (3)
- default (7)
 - HDFS_vars.yaml
 - core-site.xml**
 - hadoop-env.sh
 - hdfs-site.xml
 - *krb5.conf
 - ldap_bind.file
 - log4j.properties

default/core-site.xml 历史对比 查看历史版本 请输入配置名 显示所有说明

配置列表

- hadoop.rpc.protection: privacy.authentication
- hadoop.caller.context.enabled: true
- ipc.server.log.slow.rpc: ☒
- ipc.server.listen.queue.size: 65535
- topology.script.file.name: /etc/hadoop/conf/RackAware.py
- io.compression.codec.lzo.class: com.hadoop.compression.lzo.LzoCodec

2.12.0 保存更改

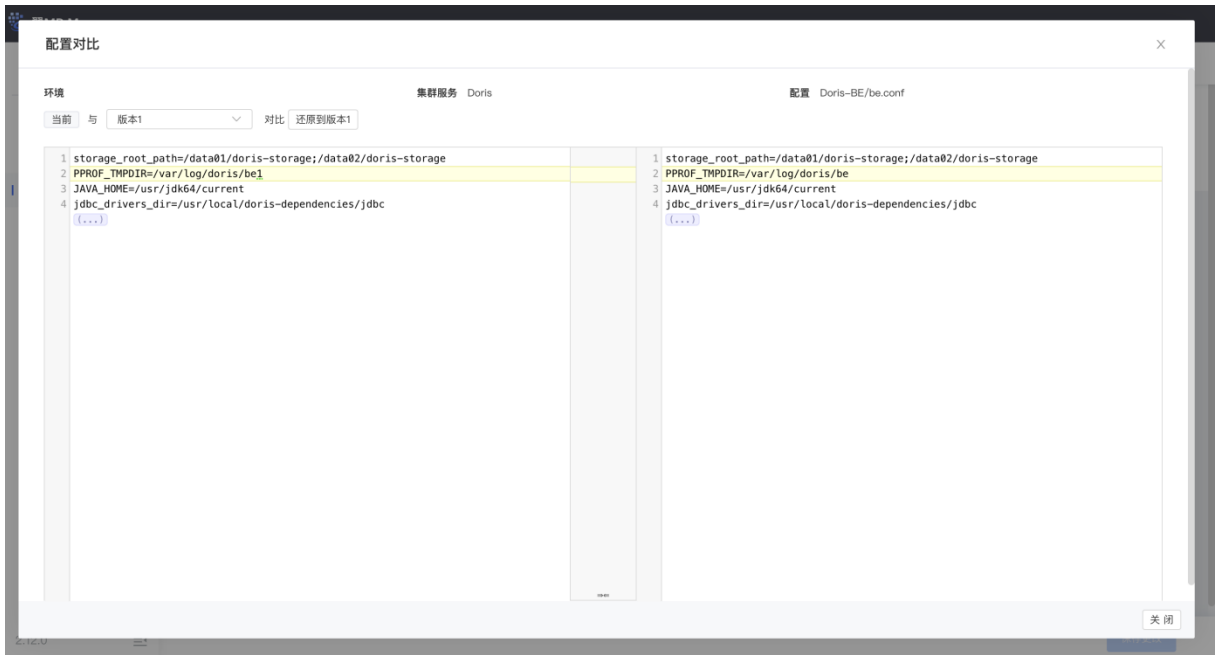
配置对比

环境 集群服务 HDFS 配置 default/core-site.xml

当前 与 版本1 对比 还原到版本1

12	(...)	12	(...)
13	<property>	13	<property>
14	<name>ipc.server.listen.queue.size</name>	14	<name>ipc.server.listen.queue.size</name>
15	<value>32769</value>	15	<value>32768</value>
16	<description>接受客户端连接的服务器的侦听队列的长度</description>	16	<description>接受客户端连接的服务器的侦听队列的长度</description>
	</property>		</property>
	(...)		(...)

关闭



6.4.21. 查看配置历史版本

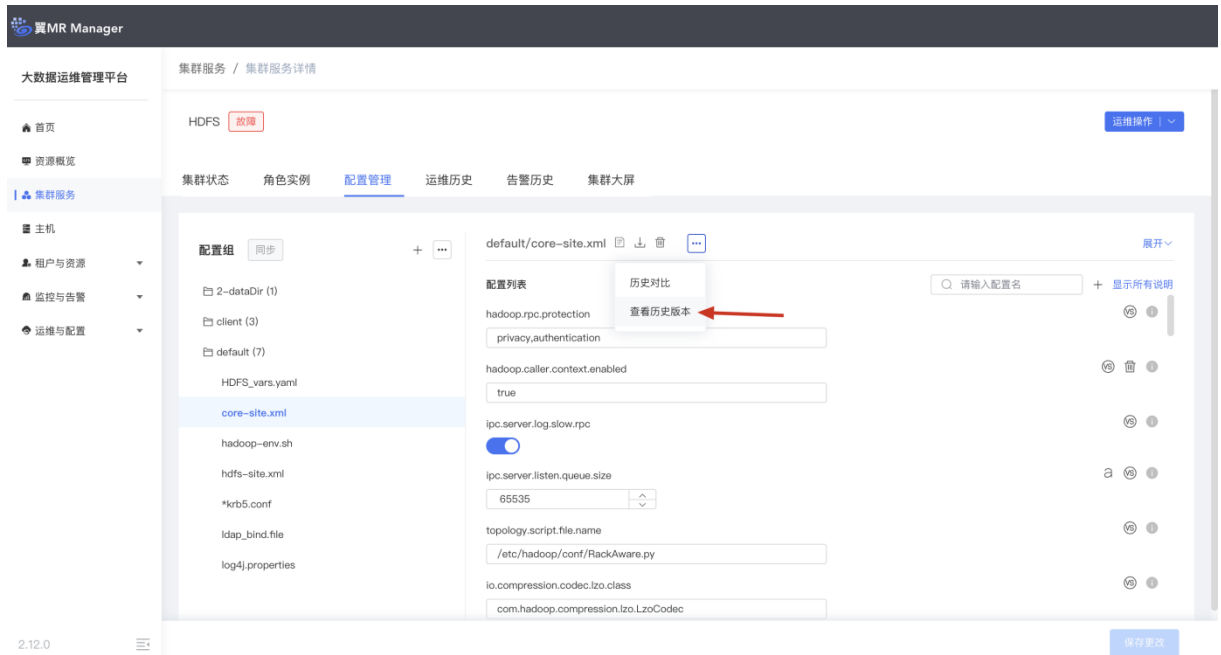
本章节主要介绍翼 MR Manager 的集群服务查看配置历史版本的操作。

操作场景

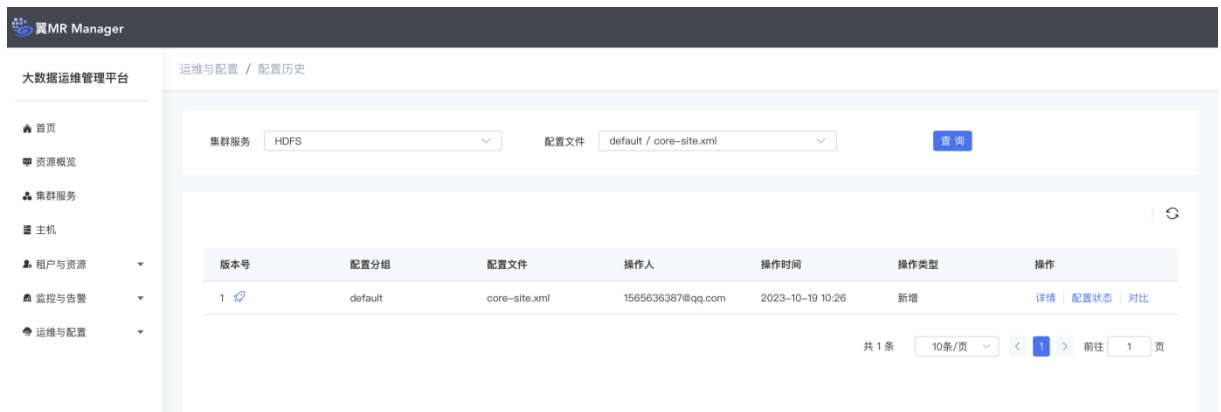
支持用户查看指定配置文件的历史版本，不同版本的配置内容以及差异。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 单击“更多 > 查看历史版本”。如图所示：



(9) 进入到配置历史页面，配置历史页面展示当前配置的所有历史版本。如图所示：

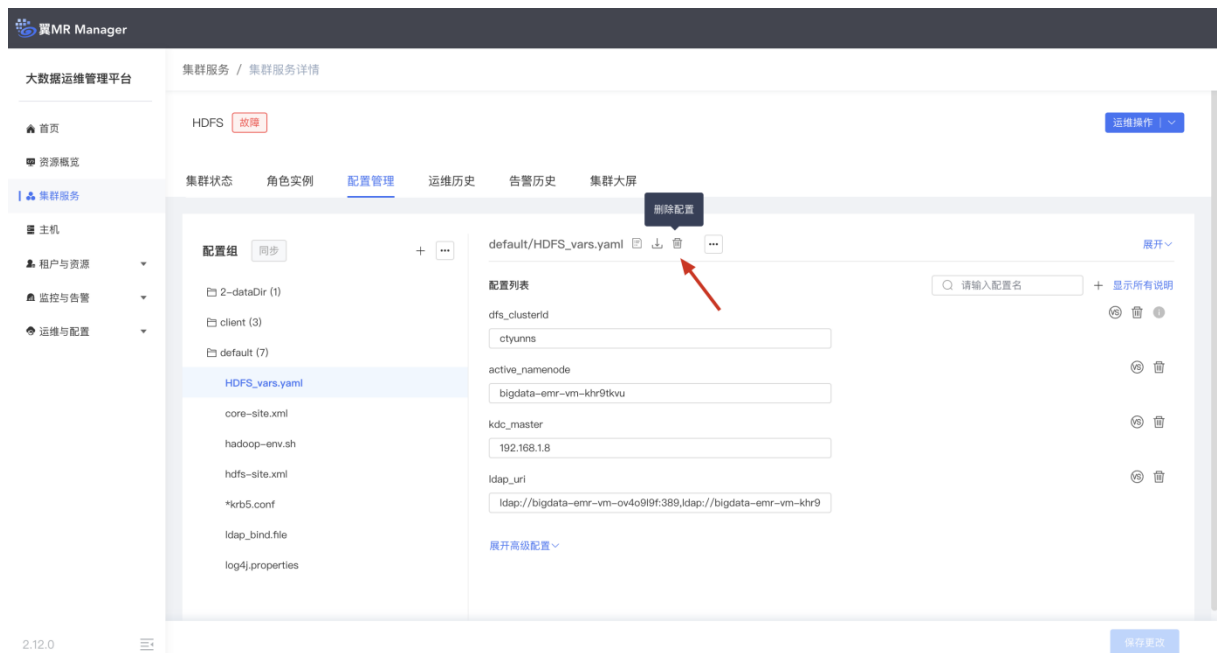


6.4.22. 删除配置

本章节主要介绍翼 MR Manager 的集群服务删除配置的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 出现删除确认弹框，单击“确定”即可。如图所示：



(9) 单击“删除配置”，操作“确认”即可。

6.4.23. 配置同步

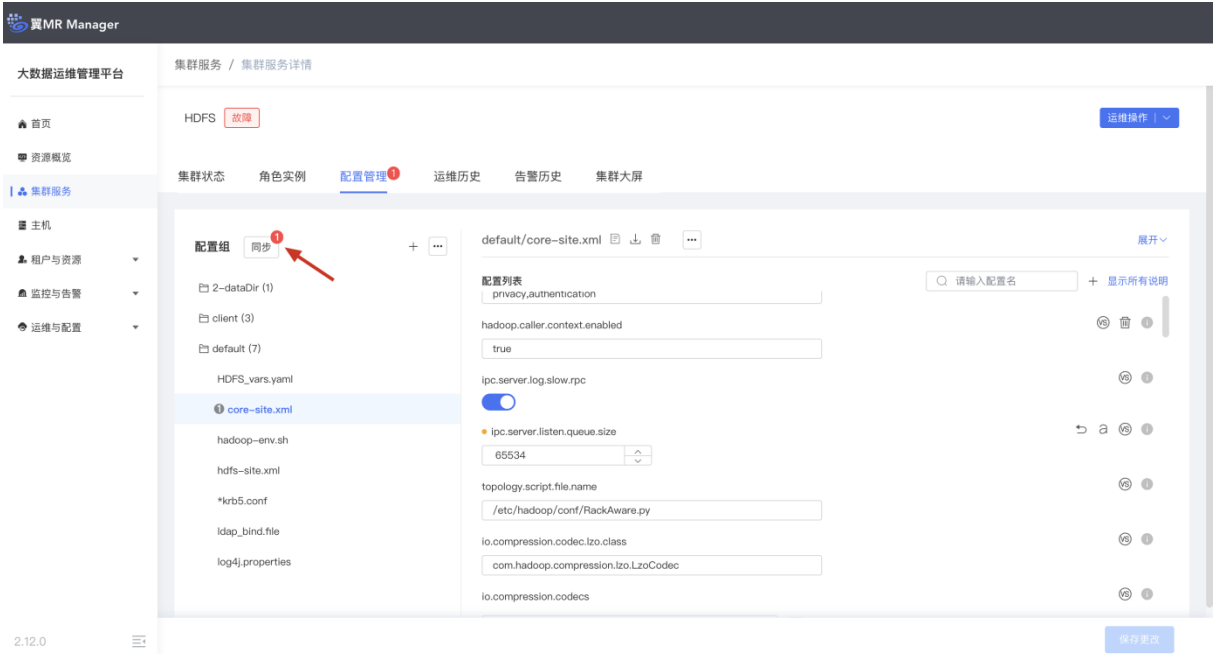
本章节主要介绍翼 MR Manager 的集群服务配置同步的操作。

操作场景

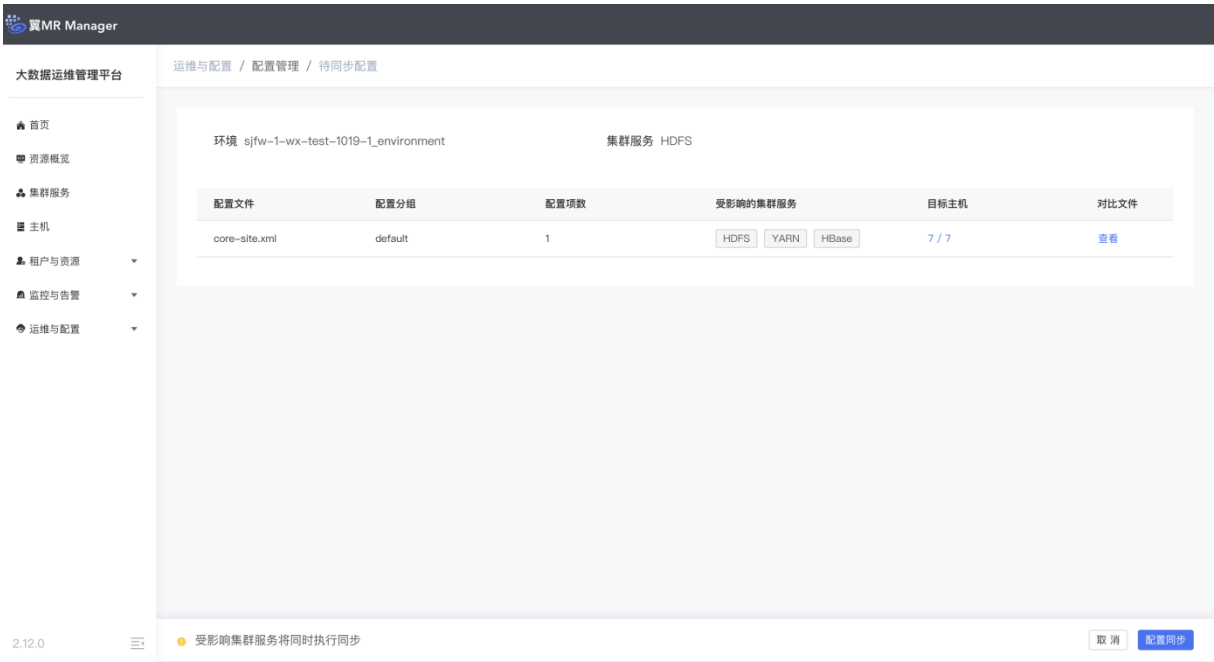
支持将新的配置文件或修改后的配置同步到相应主机上。

操作步骤

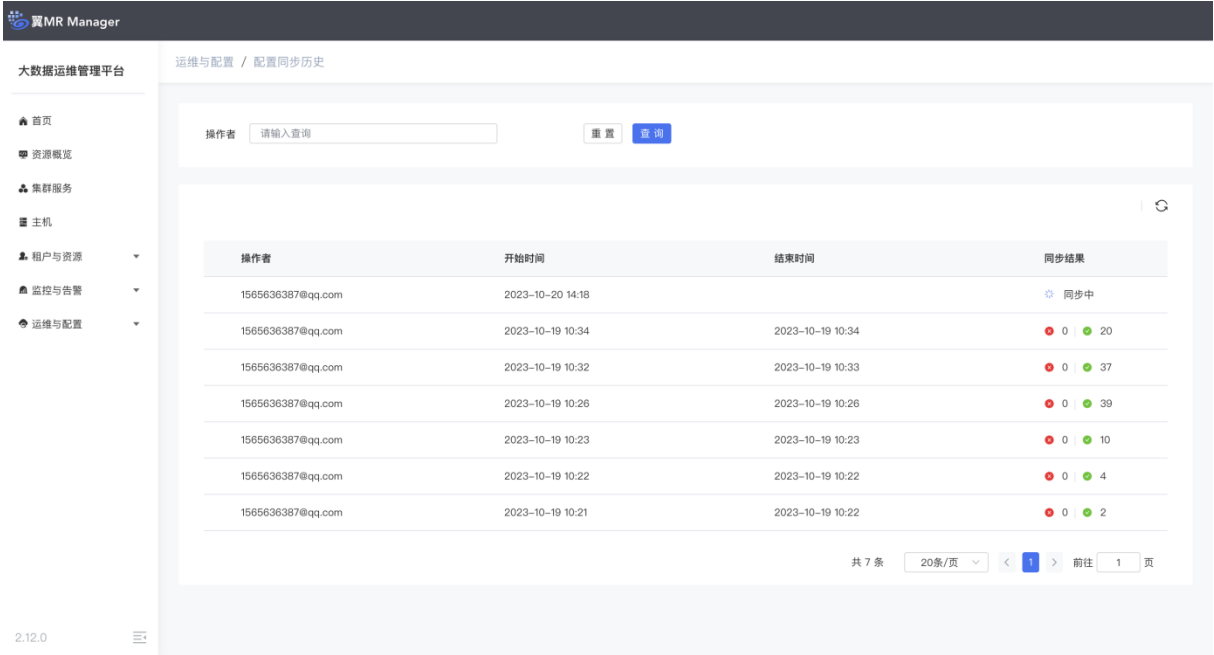
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击“同步”按钮。如图所示：



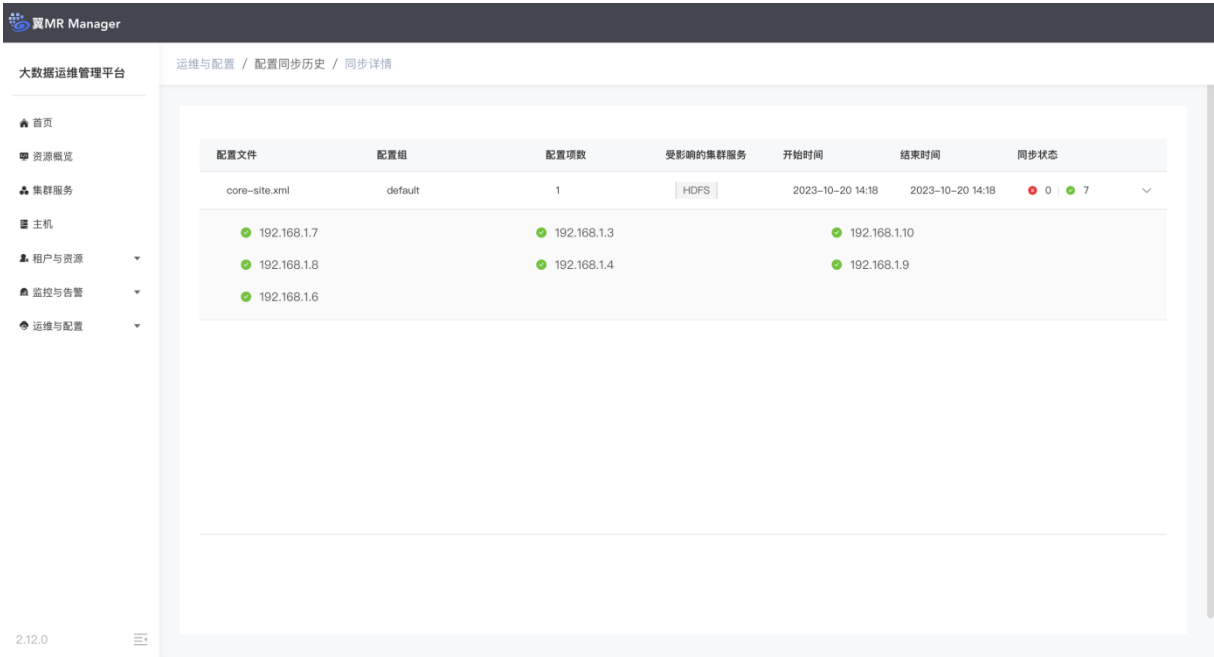
(8) 进入待同步配置页面。如图所示：



(9) 单击“配置同步”按钮，跳转到配置同步菜单页面。如图所示：



(10) 单击“同步中”，进入到配置同步详情页面，查看配置同步进度，等待同步完成。如图所示：



6.4.24. 查看配置状态

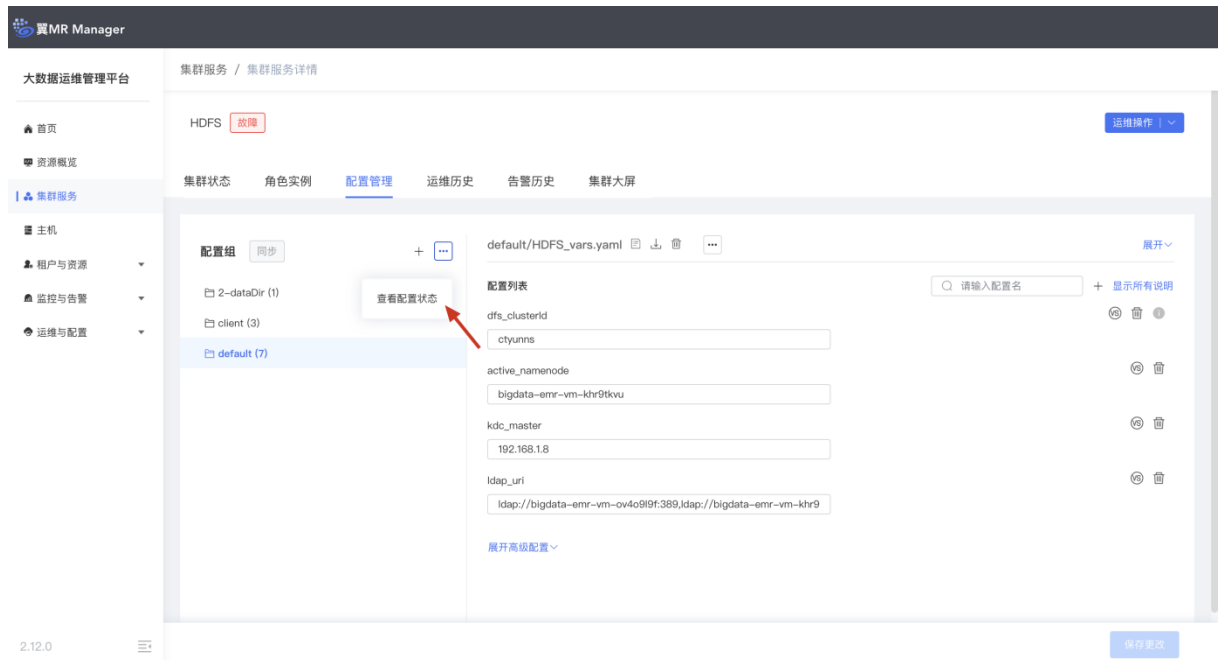
本章节主要介绍翼 MR Manager 的集群服务查看配置状态的操作。

操作场景

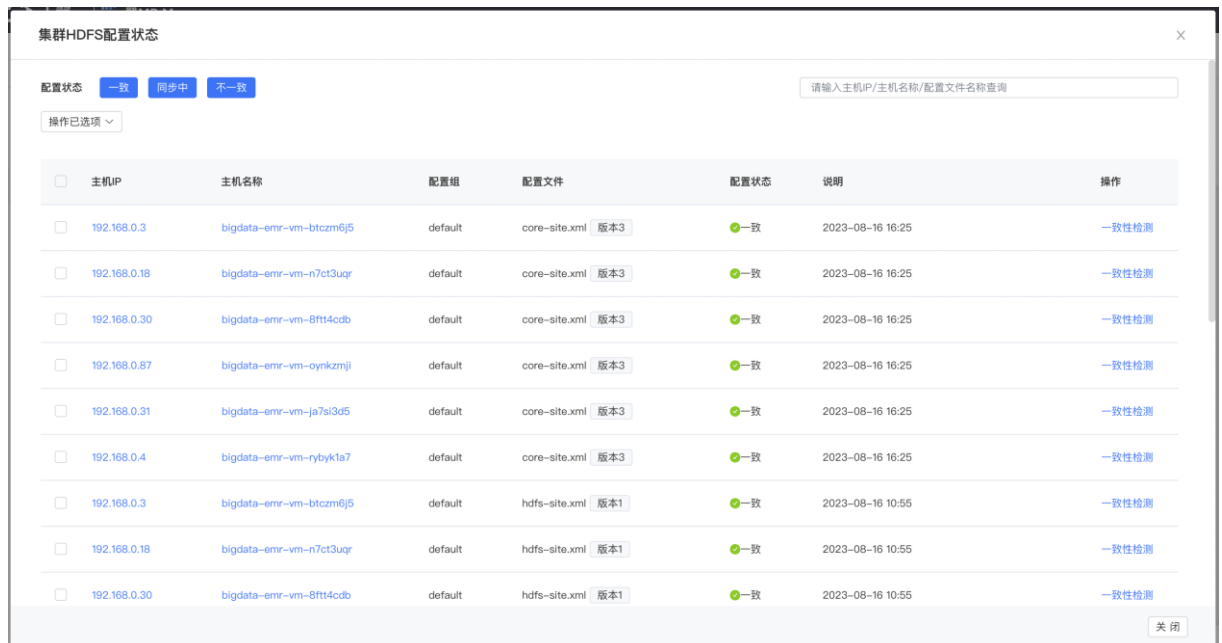
支持查看指定集群服务所有配置文件的同步状态。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击“更多 > 查看配置状态”按钮。如图所示：



- (8) 出现配置同步状态弹框，展示所有配置文件的同步状态。如图所示：



6.4.25. 重新同步

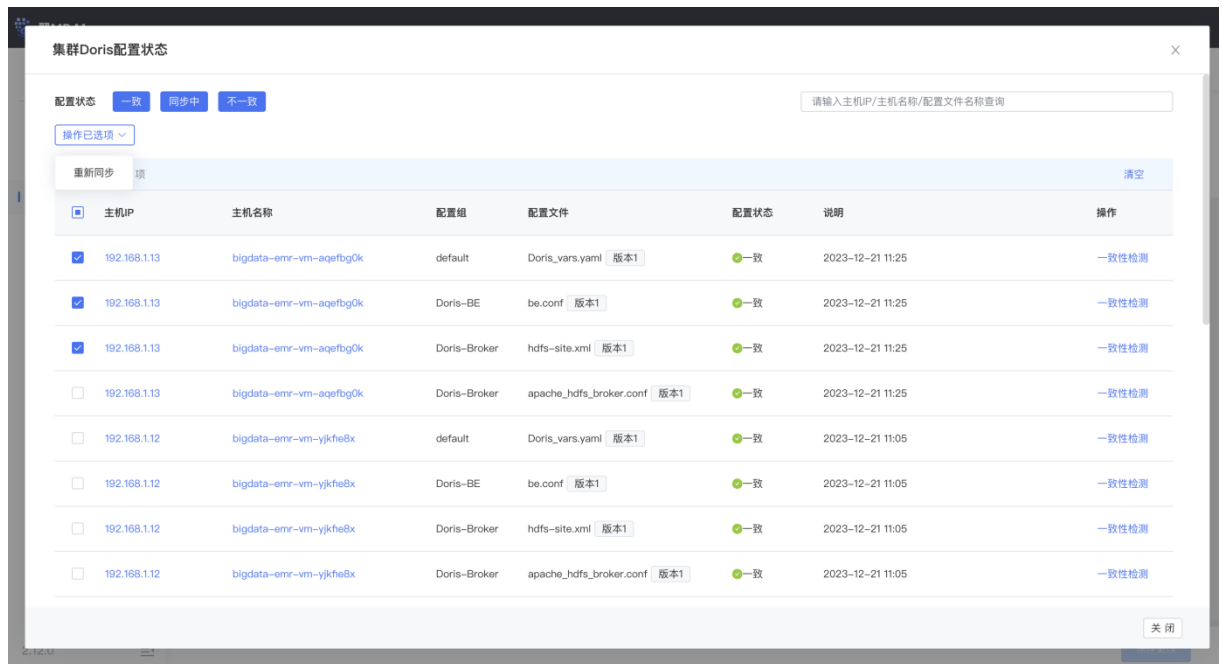
本章节主要介绍翼 MR Manager 的集群服务重新同步的操作。

操作场景

支持对同步失败的配置或配置状态不一致的进行重新同步。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击“更多 > 查看配置状态”按钮，出现配置同步状态弹框。
- (8) 勾选需要操作主机前的复选框。
- (9) 单击“操作已选项 > 重新同步”，即可支持配置同步操作。如图所示：



6.4.26. 配置回滚

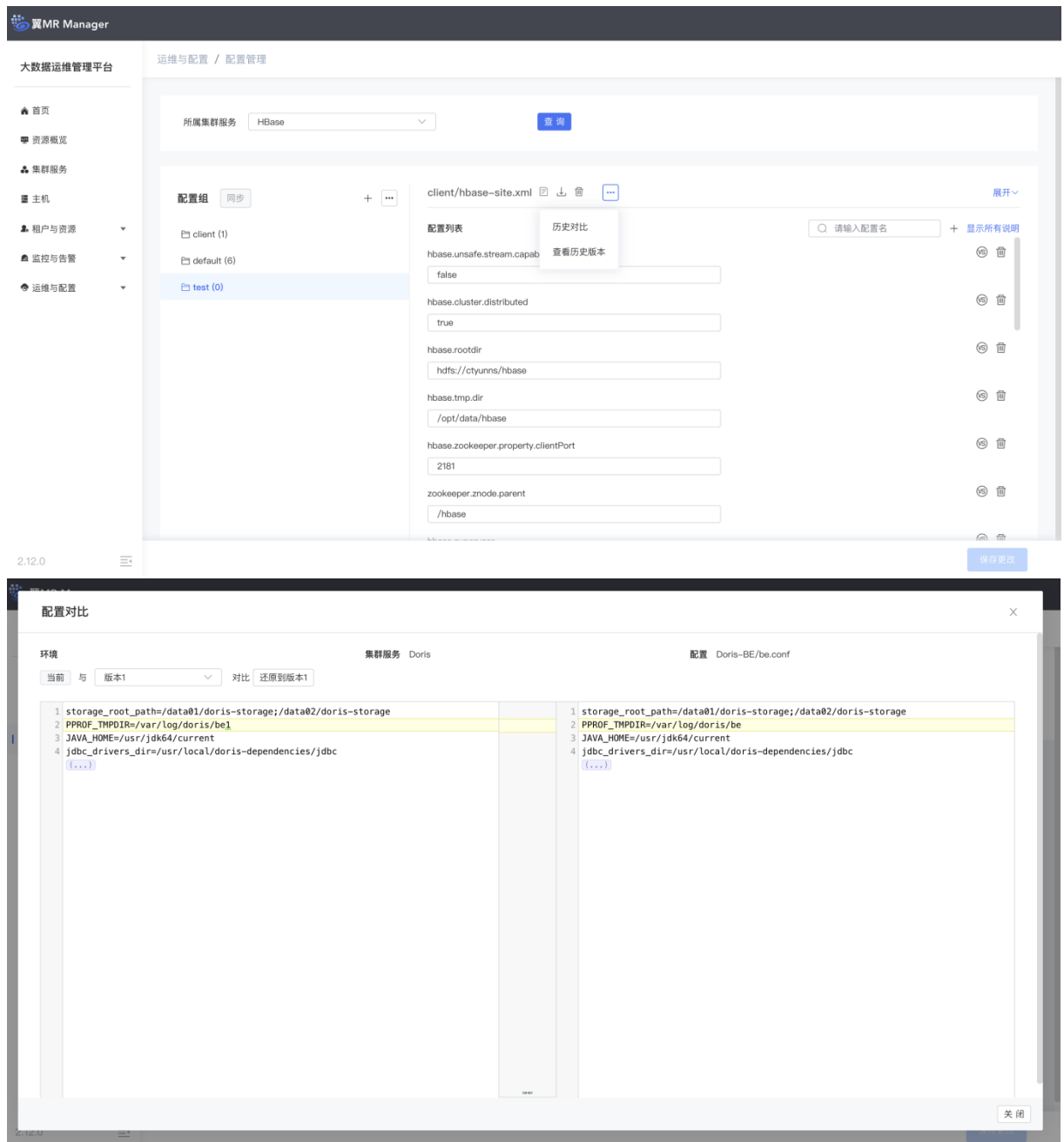
本章节主要介绍翼 MR Manager 的集群服务配置回滚的操作。

操作场景

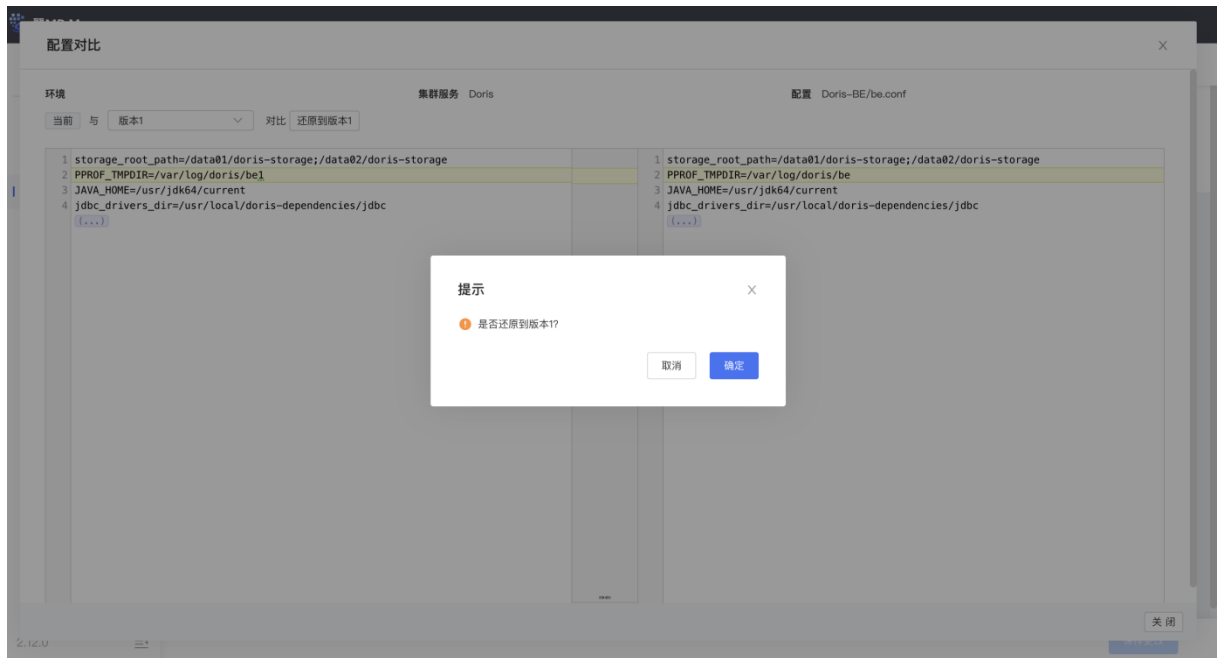
支持回退指定配置文件为某个历史版本。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 单击“更多 > 历史对比”，出现历史对比弹框。如图所示：



- (9) 单击“还原到版本 X”按钮，出现确认操作提示弹框。如图所示：



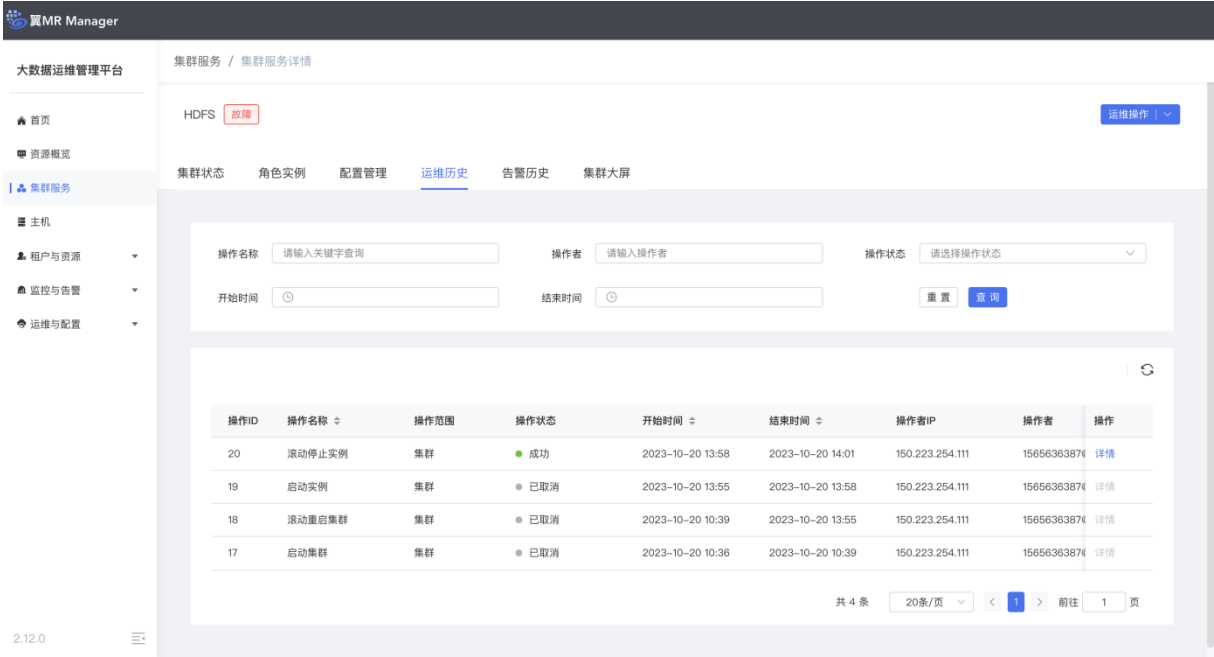
- (10) 单击“确定”。
- (11) 单击“关闭”按钮，关闭配置对比弹框。
- (12) 查看配置文件，则配置文件内容与所还原的版本内容一致。
- (13) 单击“同步”，对配置文件进行同步操作即可。

6.4.27. 查看运维操作历史

本章节主要介绍翼 MR Manager 的集群服务运维操作历史的查看操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“运维历史”tab，即可查看该集群服务的所有运维操作历史。如图所示：

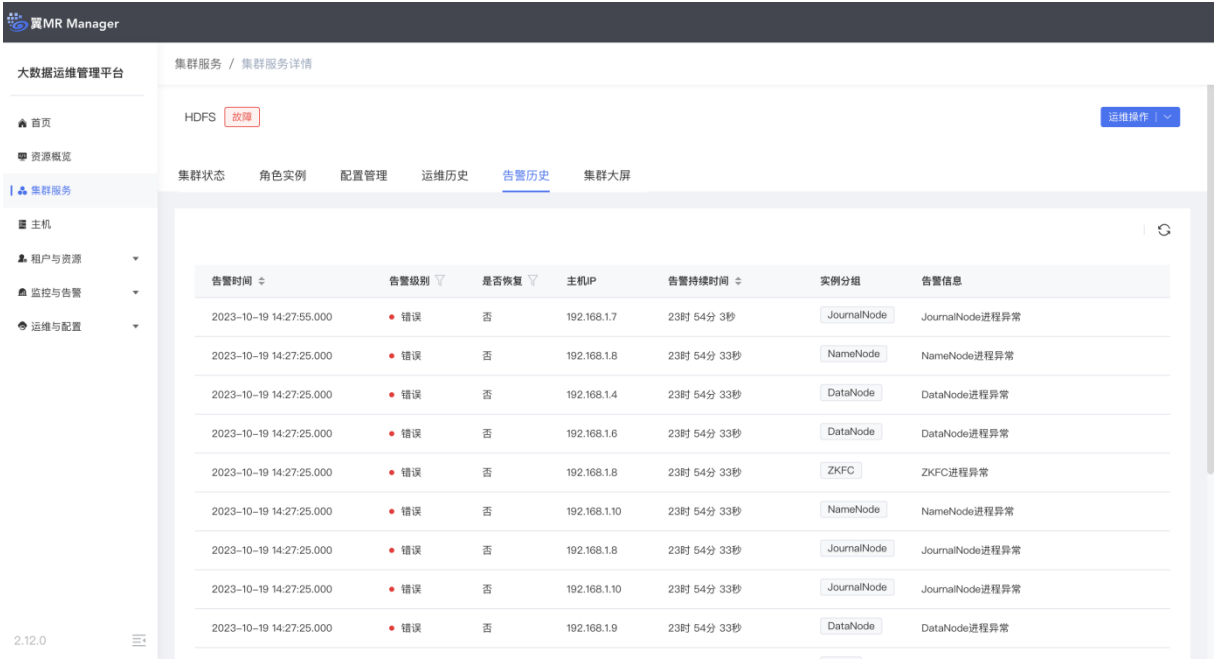


6.4.28. 查看告警历史

本章节主要介绍翼 MR Manager 的集群服务告警历史的查看操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“告警历史”tab，即可查看该集群服务的所有告警信息。如图所示：

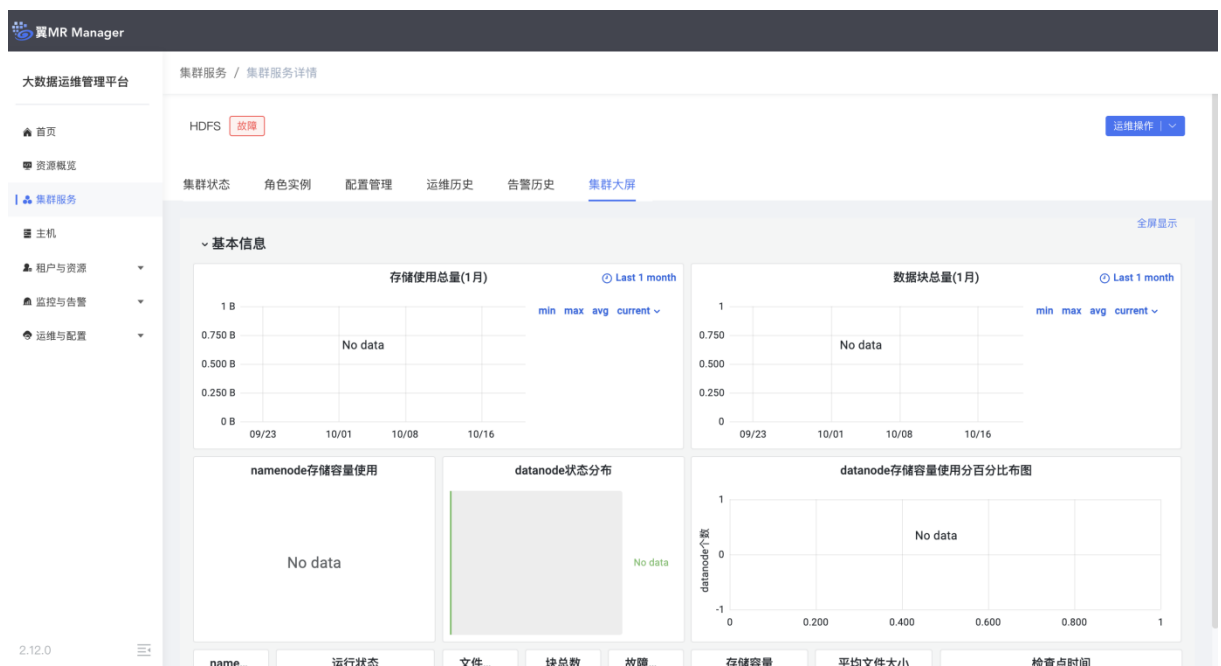


6.4.29. 查看集群服务大屏

本章节主要介绍翼 MR Manager 的集群服务大屏的查看操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“集群服务”，进入集群服务列表页面。
- (5) 选择指定集群服务，单击集群服务名称进入集群服务详情页面。
- (6) 单击“集群大屏”tab，即可查看该集群服务的监控大屏。如图所示：

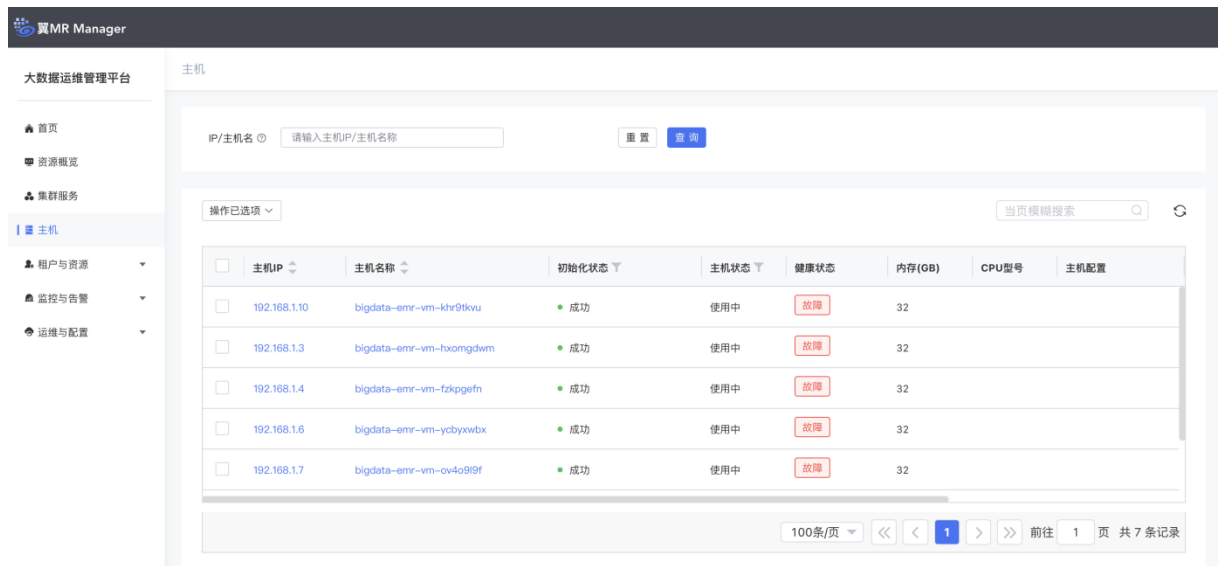


6.5. 主机

6.5.1. 主机管理页面概述

本章节主要介绍翼 MR Manager 的主机管理特性。

进入到翼 MR Manager 以后，点击菜单“主机”，进入主机页面。如图所示：



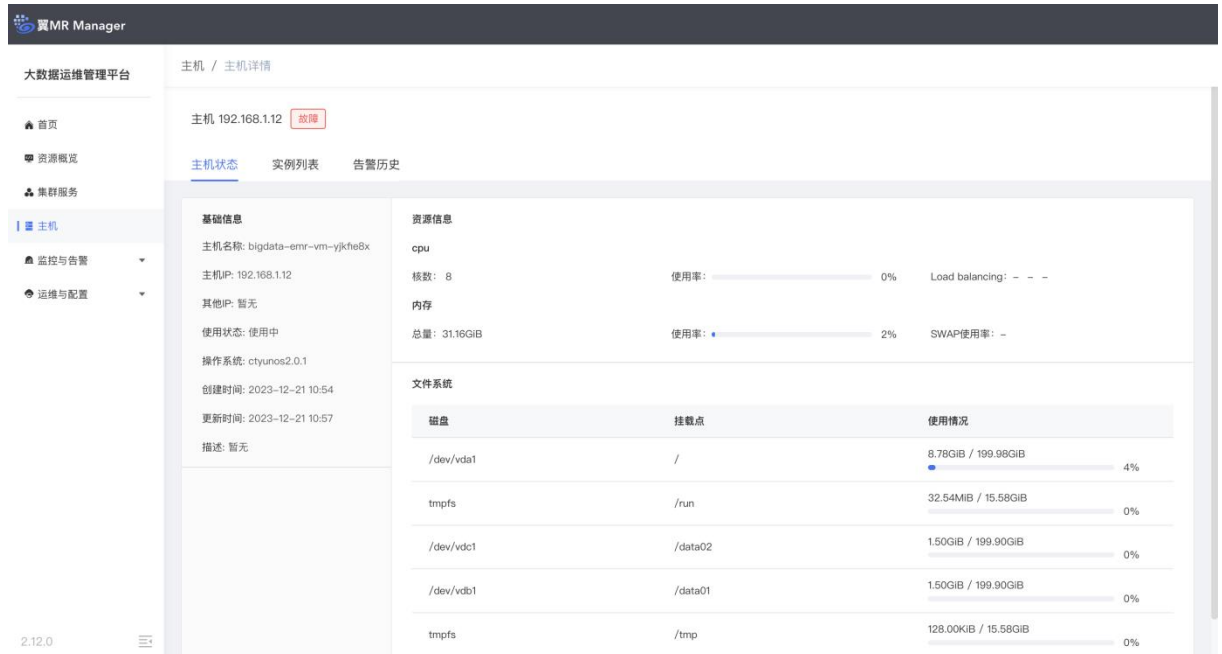
- 主机列表上方为查询区域，支持主机 IP/主机名进行查询。
- 主机列表以数据列表视图列出。
- 主机列表在左上角为操作区，可支持操作：置维护、取消维护、导出。
- 单击主机 IP 或主机名称可进入到主机详情页面。

6.5.2. 主机详情页面概述

本章节主要介绍翼 MR Manager 的主机详情页面。

进入到翼 MR Manager 以后，点击菜单“主机”，进入主机列表页面。

单击指定主机 IP 或主机名称进入到主机详情页面。如图所示：



详情页面下方依次为该主机的：主机状态、实例列表、告警历史信息。

6.5.3. 主机置维护、取消维护

本章节主要介绍翼 MR Manager 的主机置维护、取消维护的操作。

置维护

操作场景

用户需要对某个或一批主机进行置维护。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“主机”，进入主机列表页面。
- (5) 勾选需要操作主机前的复选框。
- (6) 单击“操作已选项 > 置维护”，弹出置维护操作弹框。如图所示：



置维护 ⓘ 置维护后，将屏蔽主机和关联实例的告警信息 X

主机数 1

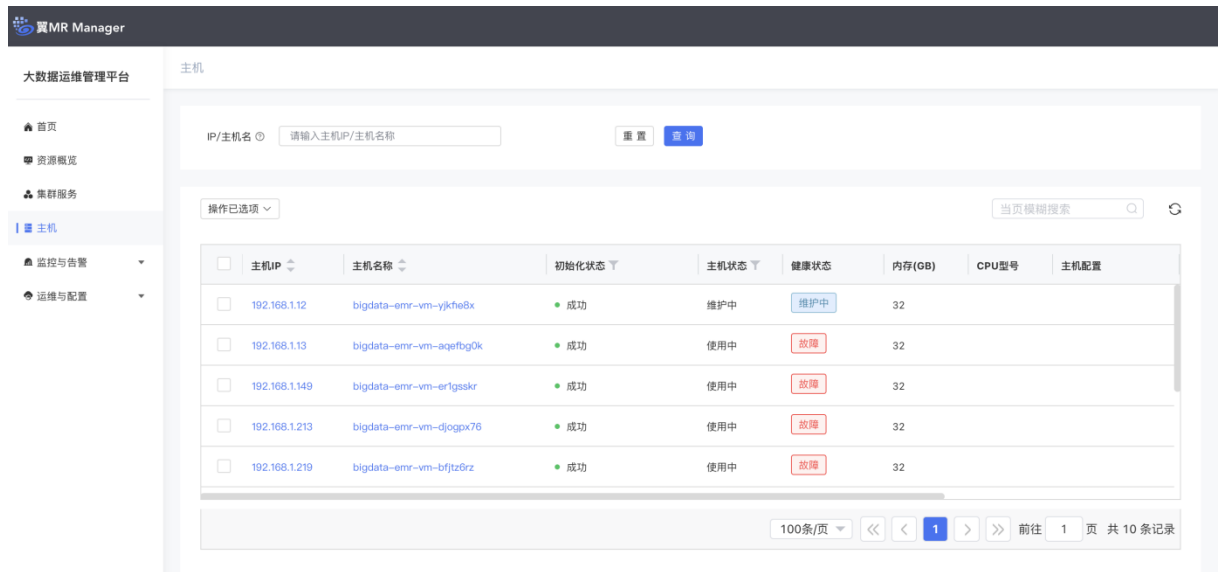
* 备注 请输入本次操作的原因

取消 确定

- (7) 输入备注，单击“确定”即可。

注意事项

- (1) 主机置维护后，该主机的健康状态显示为：维护中，如图所示：



(2) 主机置维护后，将屏蔽主机和关联实例的告警信息。

取消维护

操作场景

用户需要对某个或一批主机进行取消维护。

操作步骤

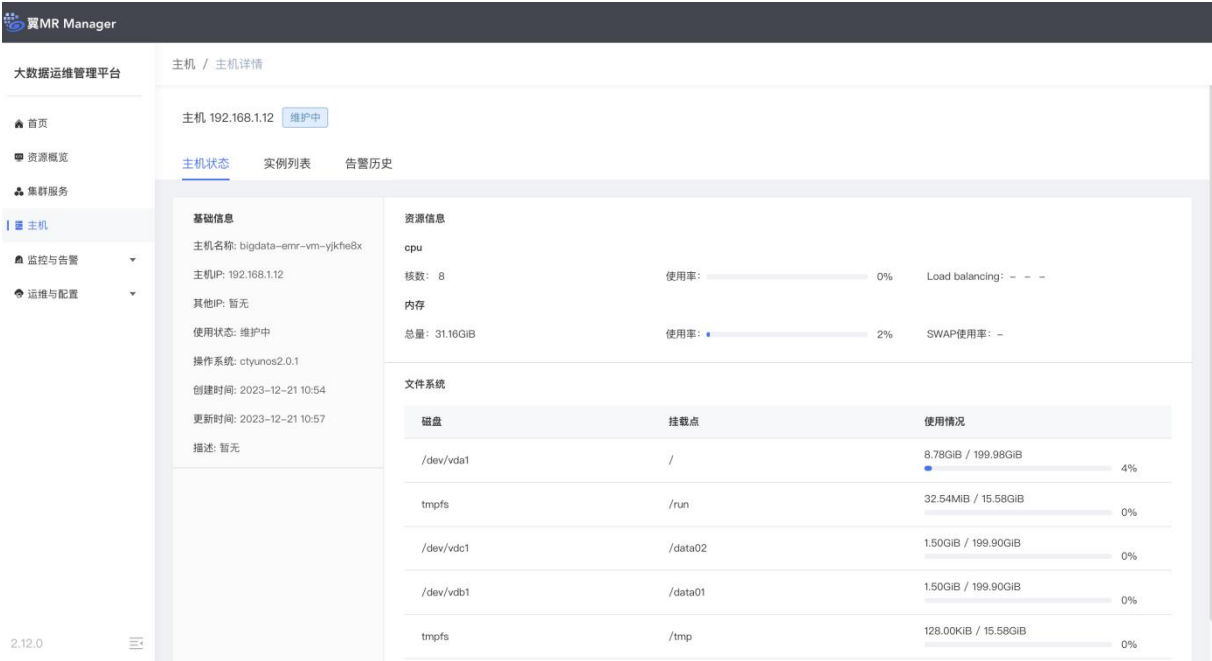
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“主机”，进入主机列表页面。
- (5) 勾选需要操作主机前的复选框。
- (6) 单击“操作已选项 > 取消维护”，弹出取消维护操作弹框。
- (7) 输入备注，单击“确定”即可。

6.5.4. 查看主机状态

本章节主要介绍翼 MR Manager 查看主机状态的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“主机”，进入主机列表页面。
- (5) 单击指定主机的主机 IP 或主机名称进入到主机详情页面，默认进入到主机状态页面。如图所示：

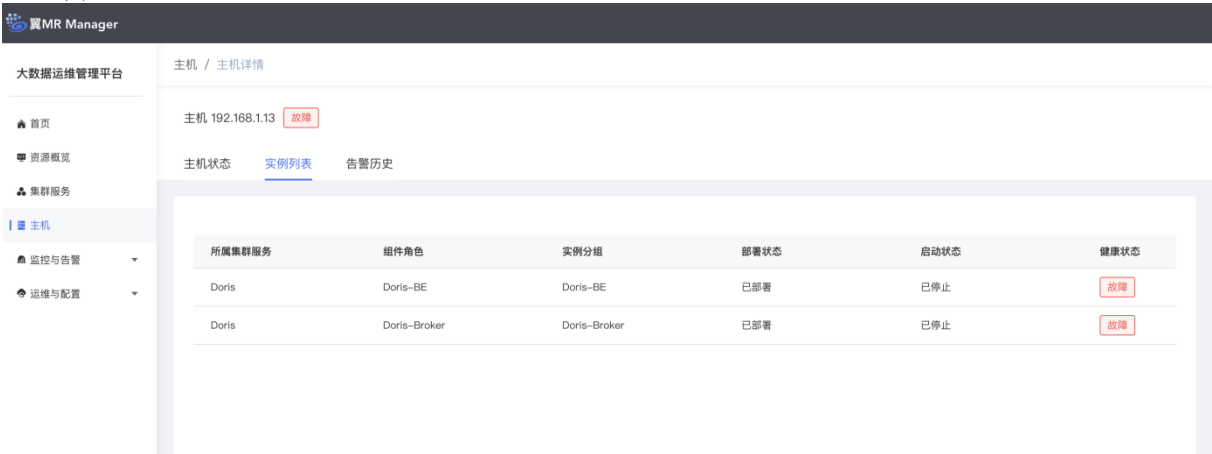


6.5.5. 查看主机实例列表

本章节主要介绍翼 MR Manager 的查看主机实例列表的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“主机”，进入主机列表页面。
- (5) 单击指定主机的主机 IP 或主机名称进入到主机详情页面。
- (6) 单击“实例列表”tab，即可查看该主机上安装的实例服务。如图所示：

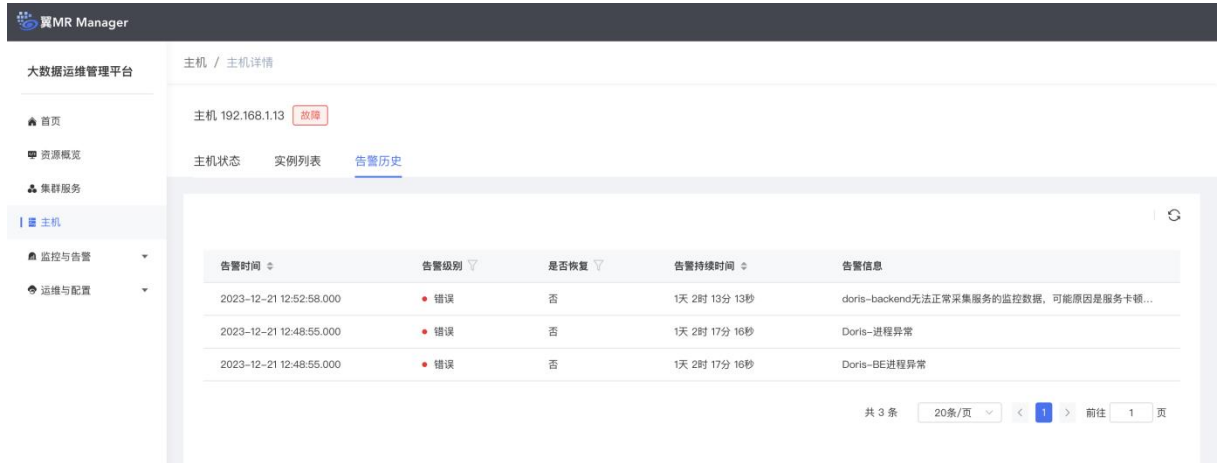


6.5.6. 查看主机告警历史

本章节主要介绍翼 MR Manager 的查看主机告警历史的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“主机”，进入主机列表页面。
- (5) 单击指定主机的主机 IP 或主机名称进入到主机详情页面。
- (6) 单击“告警历史”tab，即可查看该主机的所有告警信息。如图所示：



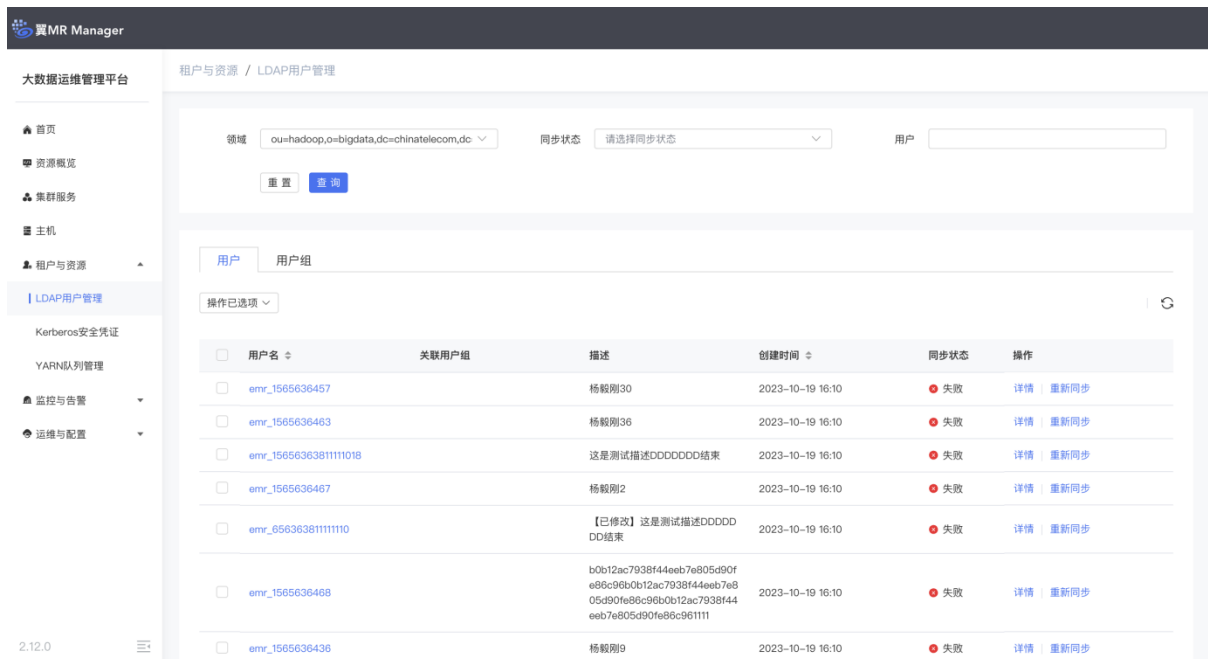
6.6. 租户与资源

本章节主要介绍翼 MR Manager 的租户与资源特性。

翼 MR Manager 提供的租户与资源主要包含：LDAP 用户管理、Kerberos 安全凭证、YARN 队列管理。

LDAP 用户管理

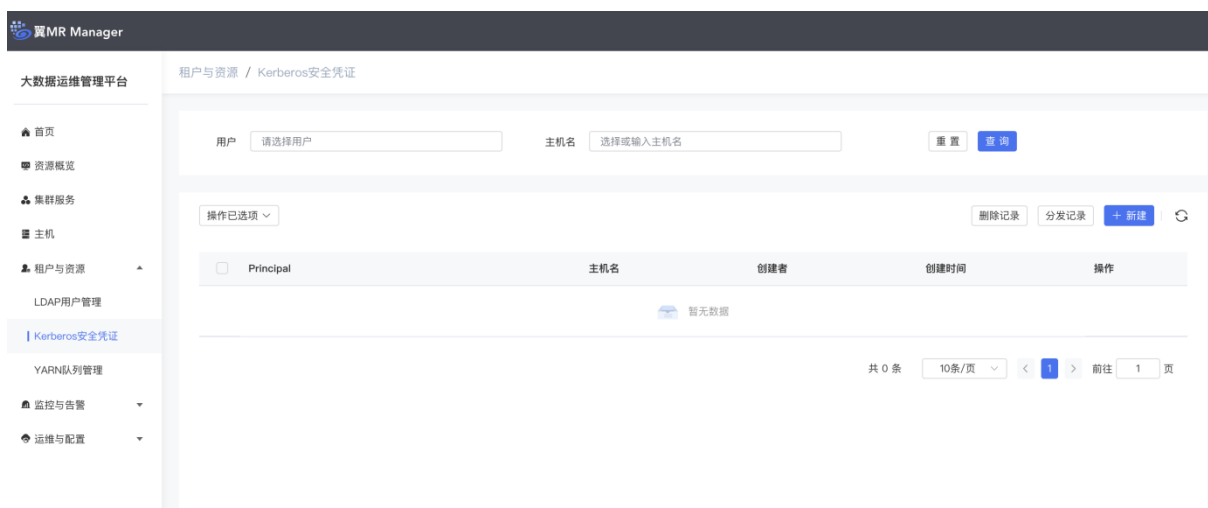
进入到翼 MR Manager 以后，点击菜单“租户与资源 > LDAP 用户管理”，进入 LDAP 用户管理页面。如图所示：



- 支持 LDAP 中的用户和用户组。
- 支持为用户关联用户组、用户组关联用户以及取消关联关系。
- 支持以用户为维度的 Principal 管理以及 Keytab 的分发。

Kerberos 安全凭证

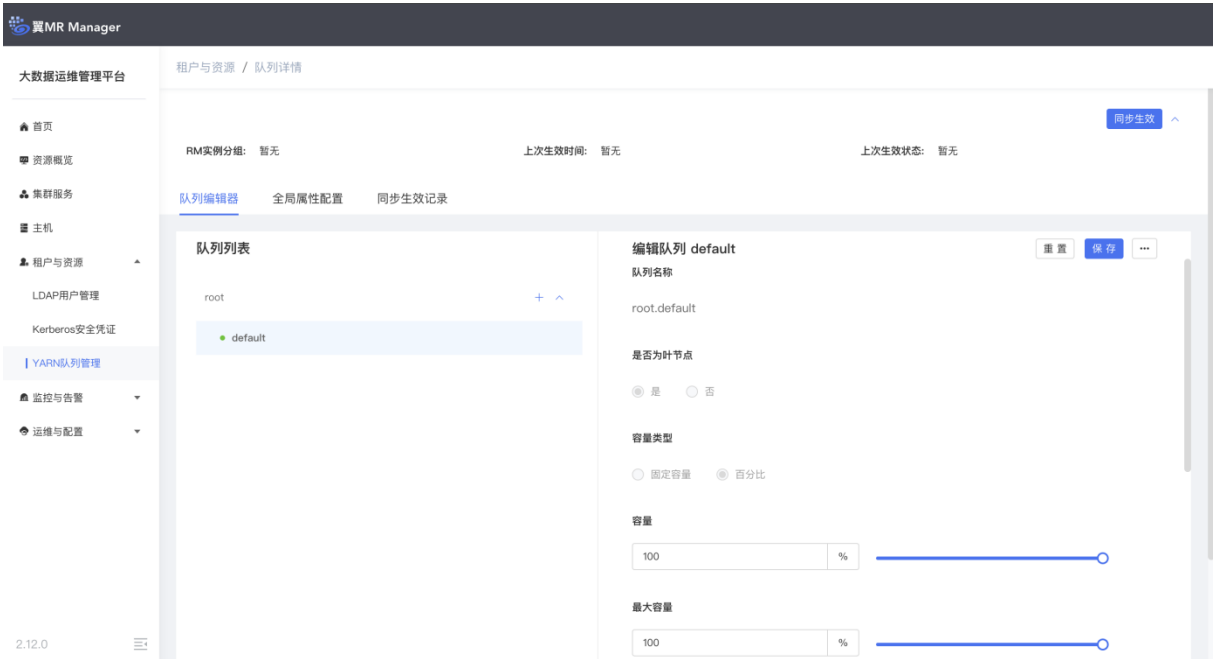
进入到翼 MR Manager 以后，点击菜单“租户与资源 > Kerberos 安全凭证”，进入 Kerberos 安全凭证页面。如图所示：



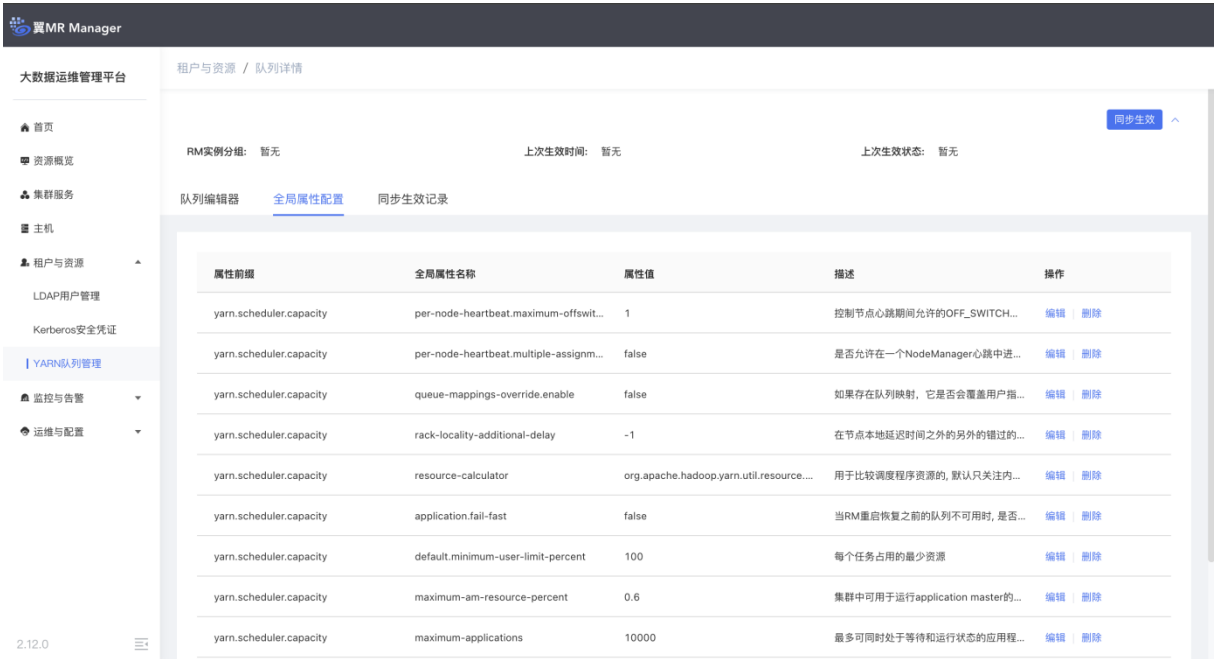
- 支持新建 Principal、删除 Principal。
- 支持 Keytab 分发与下载，并支持查看 Keytab 的分发记录。

YARN 队列管理

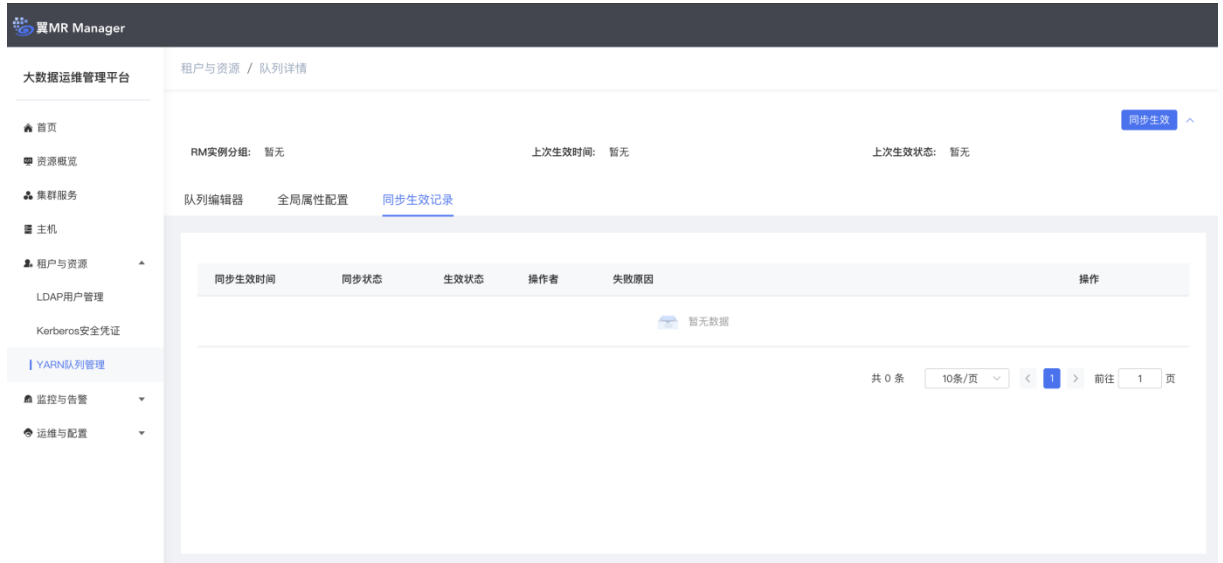
进入到翼 MR Manager 以后，点击菜单“租户与资源 > YARN 队列管理”，进入 YARN 队列管理页面。如图所示：



- 支持 YARN 队列新建、编辑与删除。
- 支持 YARN 队列启动与停止。
- 支持 YARN 队列容量设置和 YARN 队列授权。
- 支持 YARN 队列 capacity-scheduler.xml 的全局属性配置。



- 支持 YARN 队列的同步生效并支持查看同步生效记录。

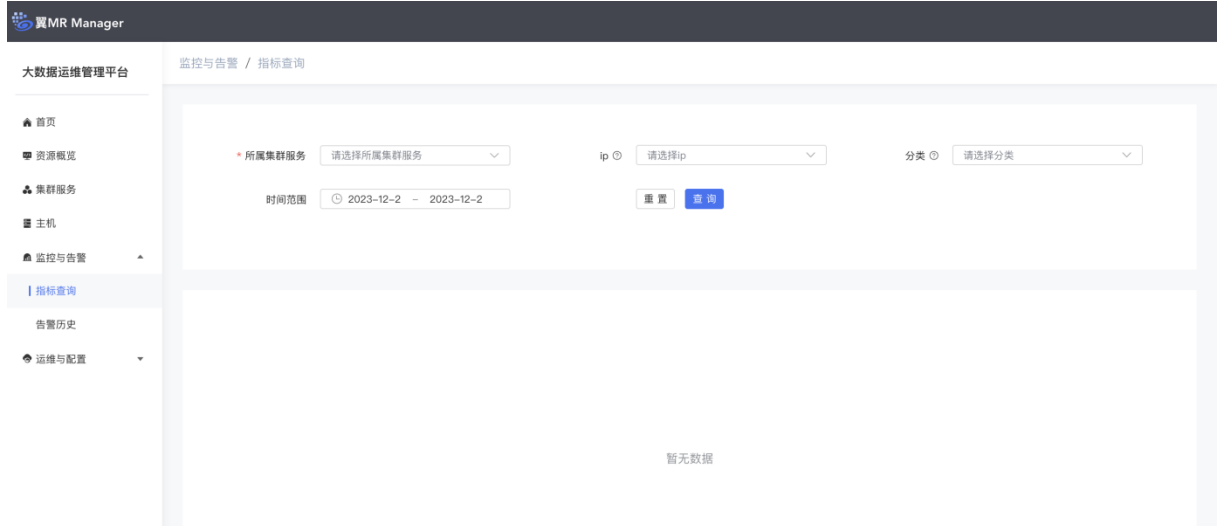


6.7. 监控与告警

6.7.1. 指标查询-指标查询概述

本章节主要介绍翼 MR Manager 的指标查询特性。

进入到翼 MR Manager 以后，点击菜单“监控与告警 > 指标查询”，进入指标查询页面。如图所示：



大数据运维管理平台

首页

资源概览

集群服务

主机

监控与告警

指标查询

告警历史

运维与配置

2.12.0

监控与告警 / 指标查询

* 所属集群服务 Doris / Doris-FE

ip 请选择ip

分类 请选择分类

时间范围 2023-12-2 - 2023-12-2

重置 查询

常用指标

Doris-FE

网络

cpu

磁盘

内存

edac

文件描述符

系统

文件系统

进程

socket

软中断

时钟

大数据运维管理平台

首页

资源概览

集群服务

主机

监控与告警

指标查询

告警历史

运维与配置

2.12.0

监控与告警 / 指标查询

* 所属集群服务 Doris / Doris-FE

ip 请选择ip

分类 请选择分类

时间范围 2023-12-2 - 2023-12-2

重置 查询

常用指标

os_network_receive_bps 网卡接收比特率

网卡接收传输速度，单位为比特率，即Bits Per Second。由一段时间间隔内，网卡接收数据量计算得出

绘图 ...

os_network_receive_drop_pps 网卡接收丢包率

网卡接收丢包率，单位为pps，即Packets Pr Second。由一段时间间隔内，网卡接收丢包数计算得出

绘图 ...

os_network_receive_err_pps 网卡接收错误包率

网卡接收错误包率，单位为pps，即Packets Pr Second。由一段时间间隔内，网卡接收错误包数计算得出

绘图 ...

os_network_transmit_bps 网卡发送比特率

网卡发送传输速度，单位为比特率，即Bits Per Second。由一段时间间隔内，网卡发送数据量计算得出

绘图 ...

os_network_transmit_drop_pps 网卡发送丢包率

网卡发送丢包率，单位为pps，即Packets Pr Second。由一段时间间隔内，网卡发送丢包数计算得出

绘图 ...

os_network_transmit_err_pps 网卡发送错误包率

网卡发送错误包率，单位为pps，即Packets Pr Second。由一段时间间隔内，网卡发送错误包数计算得出

绘图 ...

os_cpu_busy_usage_percent cpu使用率

绘图 ...

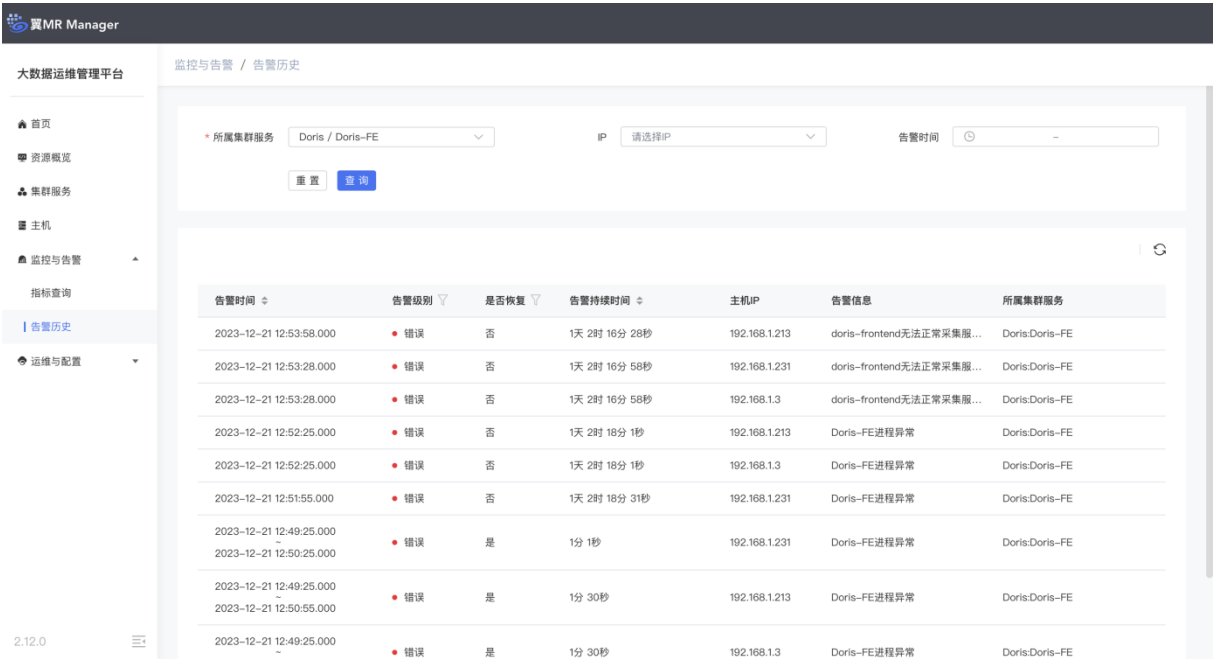


- 页面上方为查询区域，各种查询条件进行组合查询。
- 支持查询角色实例级、主机级的监控指标。
- 支持指标结果的绘图操作，让用户更直观获取监控项变化。

6.7.2. 告警历史-告警历史概述

本章节主要介绍翼 MR Manager 的告警历史特性。

进入到翼 MR Manager 以后，点击菜单“监控与告警 > 告警历史”，进入告警历史页面。如图所示：



- 页面上方为查询区域，各种查询条件进行组合查询。

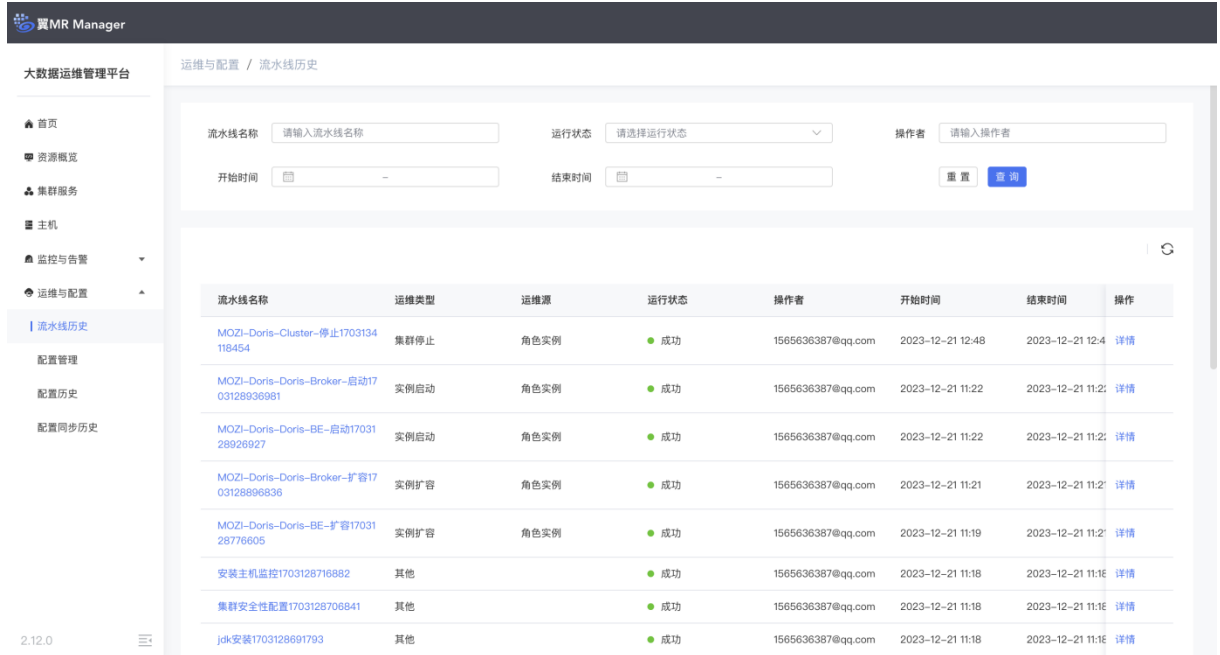
- 支持查询集群服务级、角色实例级、主机级的告警信息。

6.8. 运维与配置

6.8.1. 流水线历史-流水线历史概述

本章节主要介绍翼 MR Manager 的流水线历史特性。

进入到翼 MR Manager 以后，点击菜单“运维与配置 > 流水线历史”，进入流水线历史页面。如下图所示：



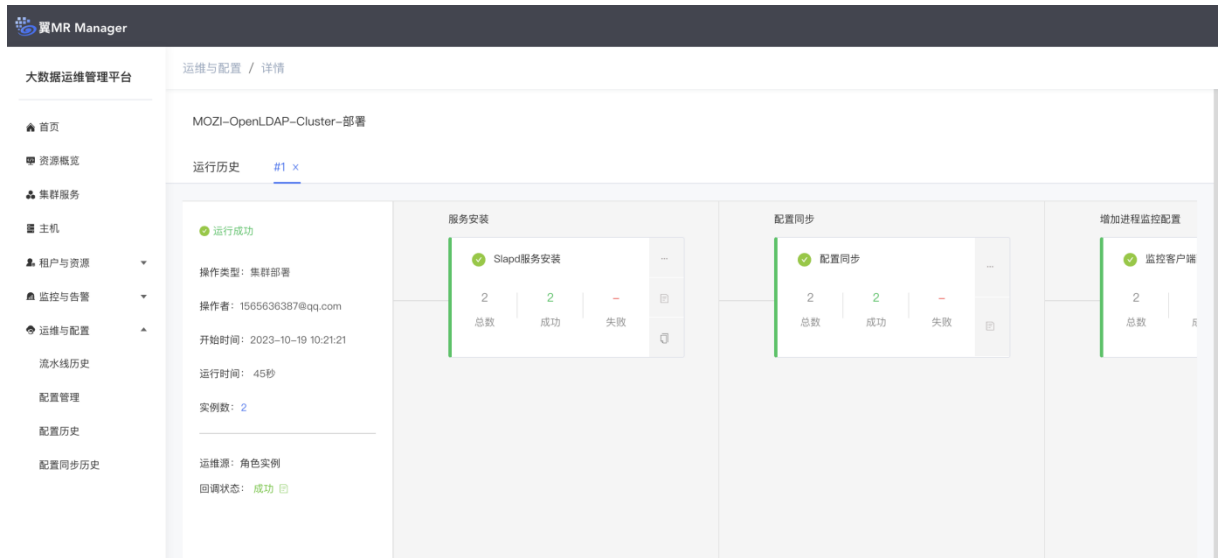
- 页面上方为查询区域，各种查询条件进行组合查询。
- 展示展示所有流水线的运行历史记录，以及操作人。
- 单击流水线名称或详情按钮，可查看运行详情。

6.8.2. 流水线历史-查看流水线运行详情

本章节主要介绍翼 MR Manager 的查看流水线运行详情的操作。

操作步骤

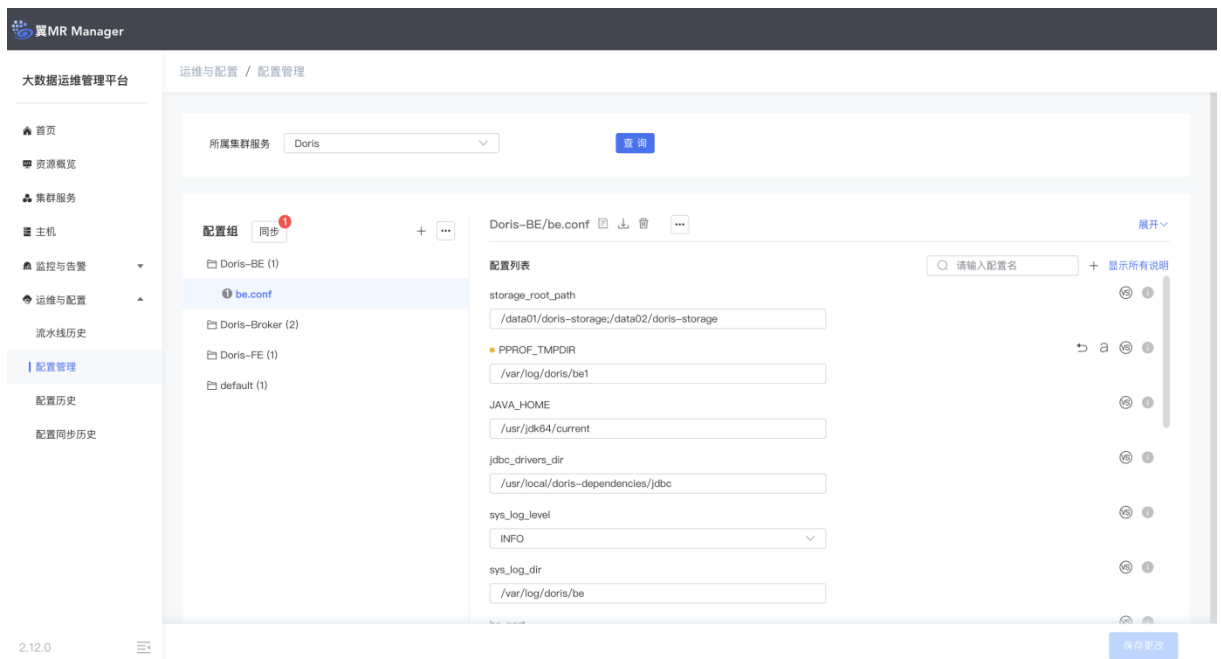
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 流水线运行历史”。
- (5) 单击要查看运行详情的流水线名称或详情按钮，进入到该运行详情页面。如图所示：



6.8.3. 配置管理-配置管理概述

本章节主要介绍翼 MR Manager 的配置管理特性。

进入到翼 MR Manager 以后，点击菜单“运维与配置 > 配置管理”，进入配置管理页面。如图所示：



- 支持查看不同集群服务的配置。
- 支持为不同集群服务新增配置组、查看配置组信息、删除配置组。
- 支持新增配置文件、查看配置文件详情、预览和下载配置文件，以及删除配置文件。
- 支持按主机维度的配置同步、查看配置状态等。

6.8.4. 配置管理-新增配置组

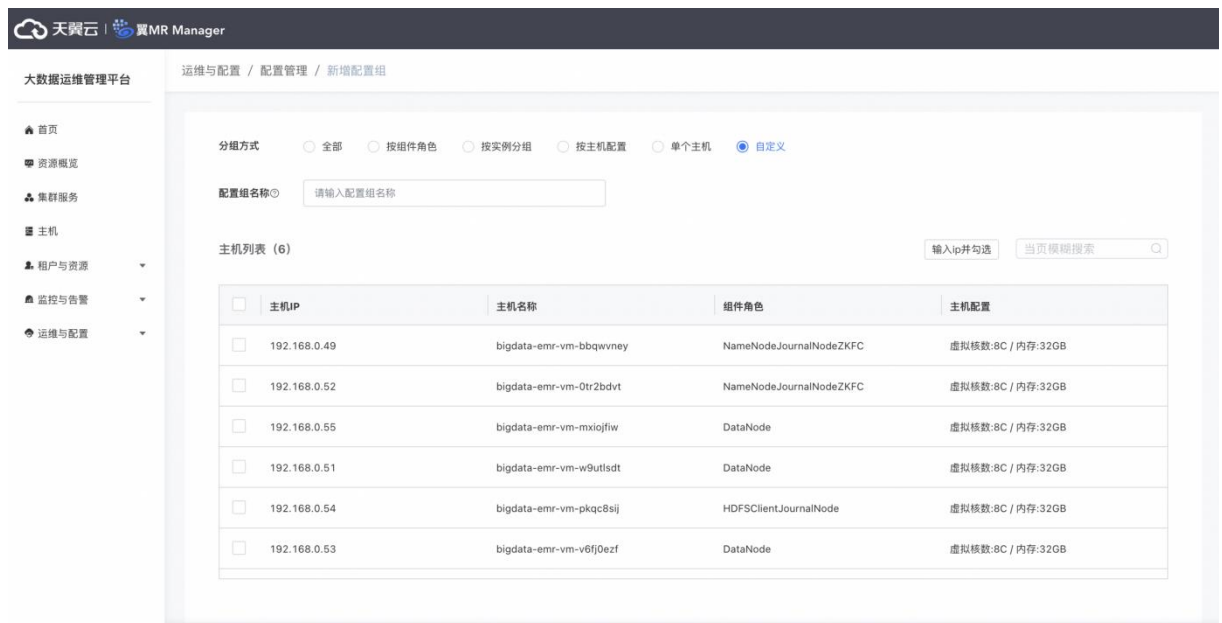
本章节主要介绍翼 MR Manager 的新增配置组的操作。

操作场景

用户可以根据不同的分组标准，对集群服务中的主机进行分组，一个配置分组下的主机共享一套配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“配置管理”tab。
- (7) 单击“新增配置组”，进入新增配置组页面。
- (8) 根据需要选择分组方式，选择主机，单击右下角“保存”按钮即可。如图所示：



注意

- 分组方式为：全部、按组件角色、按实例分组、按主机配置时，配置组名称默认且不可改。
- 分组方式为：单个主机时，需要选择要分组的主机，配置组名称为所选主机的名称且不可改。
- 分组方式为：自定义时，需要选择分组的主机，配置组名称可以自定义但不能与其他配置组名称相同。

6.8.5. 配置管理-查看配置组详情

本章节主要介绍翼 MR Manager 的查看配置组详情的操作。

操作场景

用户可以查看不同配置组的详细信息，了解该配置组所包含的主机信息。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 选择要查看的配置组，鼠标移入到配置组名称上，单击右侧的“详情”图标，进入到配置组详情页面。如图所示：

运维与配置 / 配置管理 / 配置组详情

分组方式	全部主机	配置组名称	default
主机列表	(6)		
主机名	IP	组件角色	主机配置
bigdata-emr-vm-bbqwvney	192.168.0.49	NameNode JournalNode ZKFC	虚拟核数:8C/内存:32GB
bigdata-emr-vm-0tr2bdvt	192.168.0.52	NameNode JournalNode ZKFC	虚拟核数:8C/内存:32GB
bigdata-emr-vm-mxiojfw	192.168.0.55	DataNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-w9utisdt	192.168.0.51	DataNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-pkqc8sij	192.168.0.54	HDFSClient JournalNode	虚拟核数:8C/内存:32GB
bigdata-emr-vm-v6fj0ezf	192.168.0.53	DataNode	虚拟核数:8C/内存:32GB

关闭

6.8.6. 配置管理-删除配置组

本章节主要介绍翼 MR Manager 的删除配置组的操作。

操作场景

用户可以删除不需要的配置组。

前置条件

配置组下的配置文件数量必须为 0 才能删除。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 选择要删除的配置组，鼠标移入到配置组名称上，单击右侧的“删除”图标，出现删除确认弹框。
- (7) 单击“确定”即可删除该配置组。

6.8.7. 配置管理-新增配置

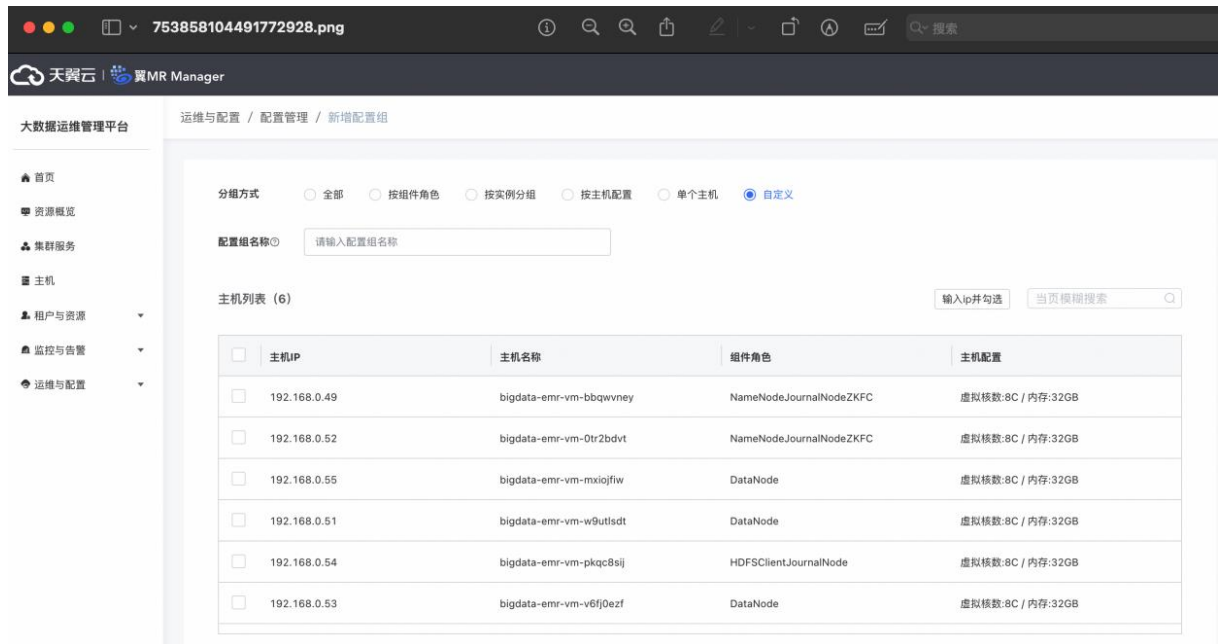
本章节主要介绍翼 MR Manager 的新增配置组的操作。

操作场景

用户可以根据不同的分组标准，对集群服务中的主机进行分组，一个配置分组下的主机共享一套配置。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“新增”，进入新增配置组页面。
- (7) 根据需要进行选择分组方式，选择主机，单击右下角“保存”按钮即可。如图所示：



注意

- 分组方式为：全部、按组件角色、按实例分组、按主机配置时，配置组名称默认且不可改。
- 分组方式为：单个主机时，需要选择要分组的主机，配置组名称为所选主机的名称且不可改。
- 分组方式为：自定义时，需要选择分组的主机，配置组名称可以自定义但不能与其他配置组名称相同。

6.8.8. 配置管理-查看配置详情

本章节主要介绍翼 MR Manager 的查看配置组详情的操作。

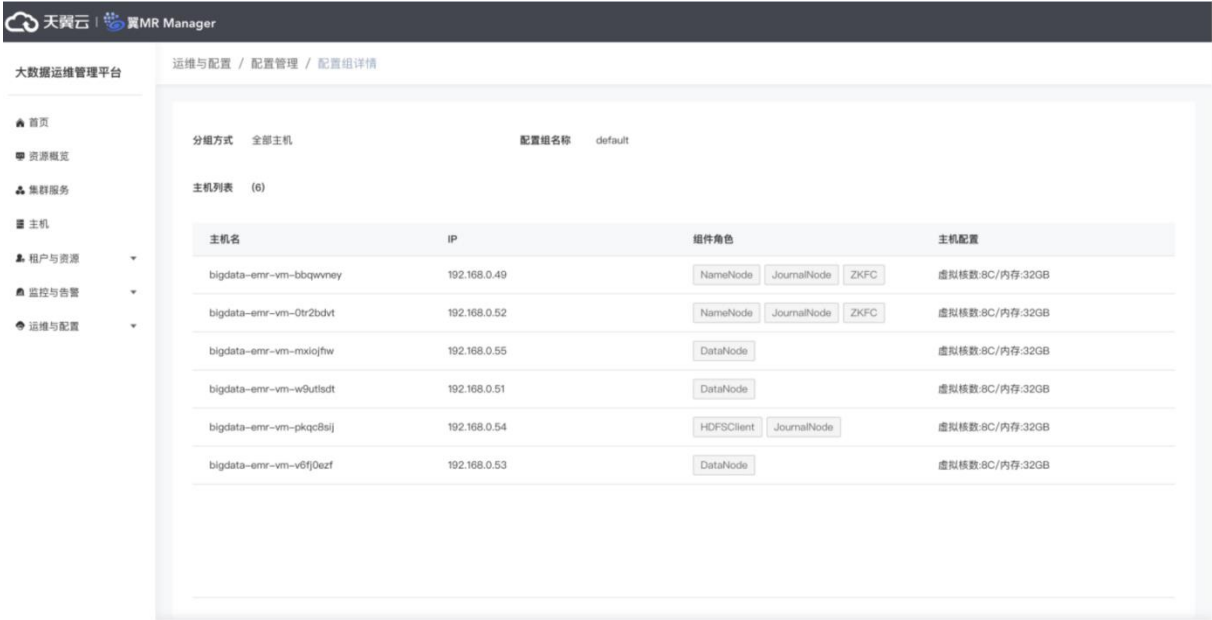
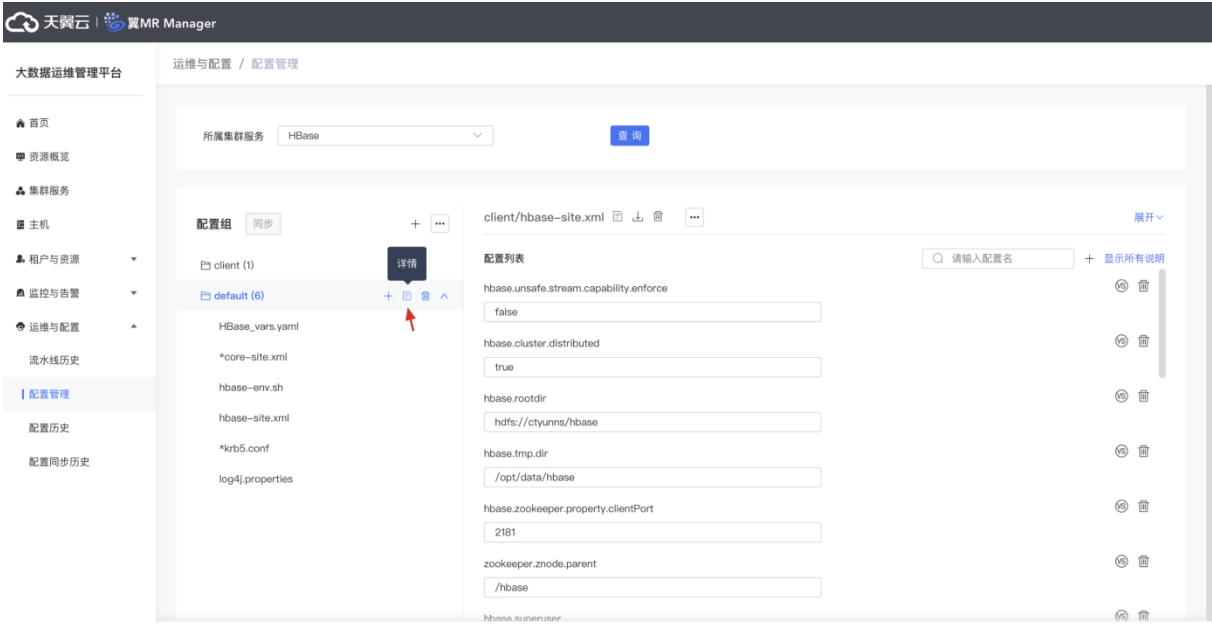
操作场景

用户可以查看不同配置组的详细信息，了解该配置组所包含的主机信息。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。

(6) 选择要查看的配置组，鼠标移入到配置组名称上，单击右侧的“详情”图标，进入到配置组详情页面。如图所示：



6.8.9. 配置管理-编辑配置

本章节主要介绍翼 MR Manager 的集群服务编辑配置的操作。

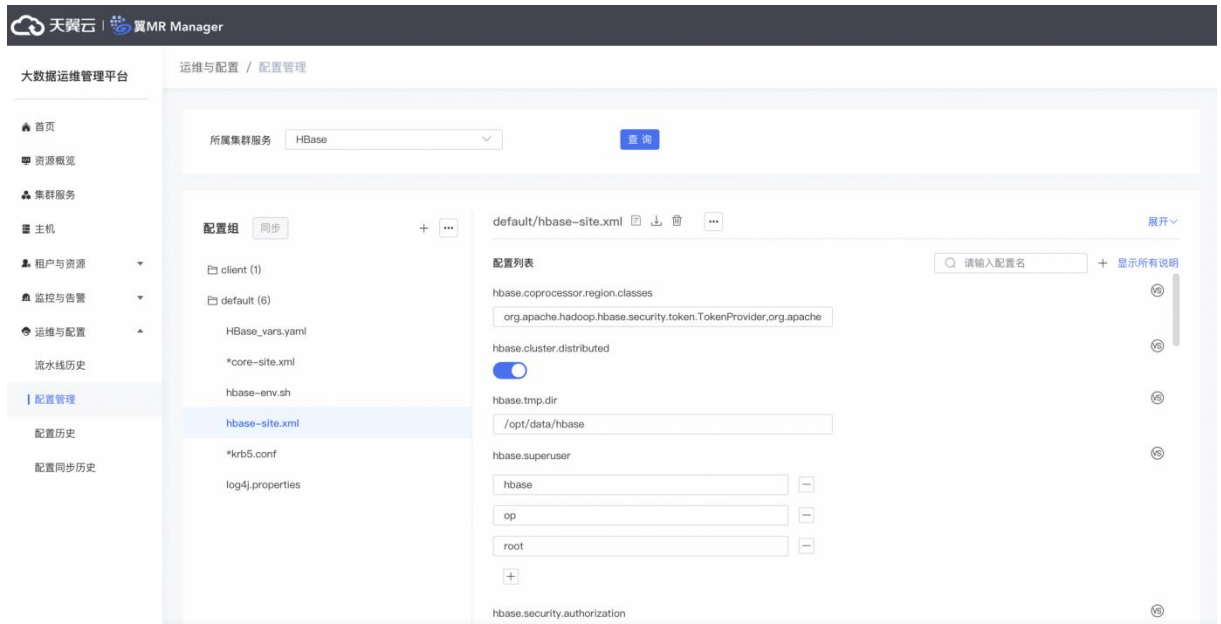
操作场景

用户可以对配置文件进行修改。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。

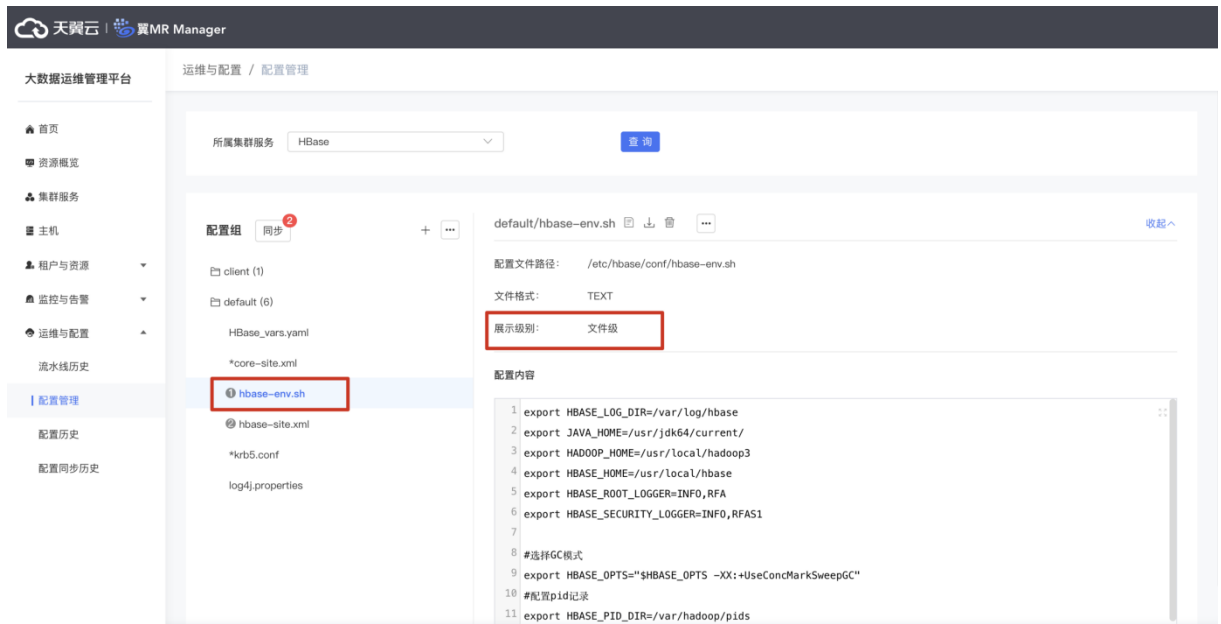
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击要修改的配置名称，页面右侧展示该配置的详细信息。如图所示：



- (7) 直接对配置文件修改即可，修改完成后点击“保存更改”。

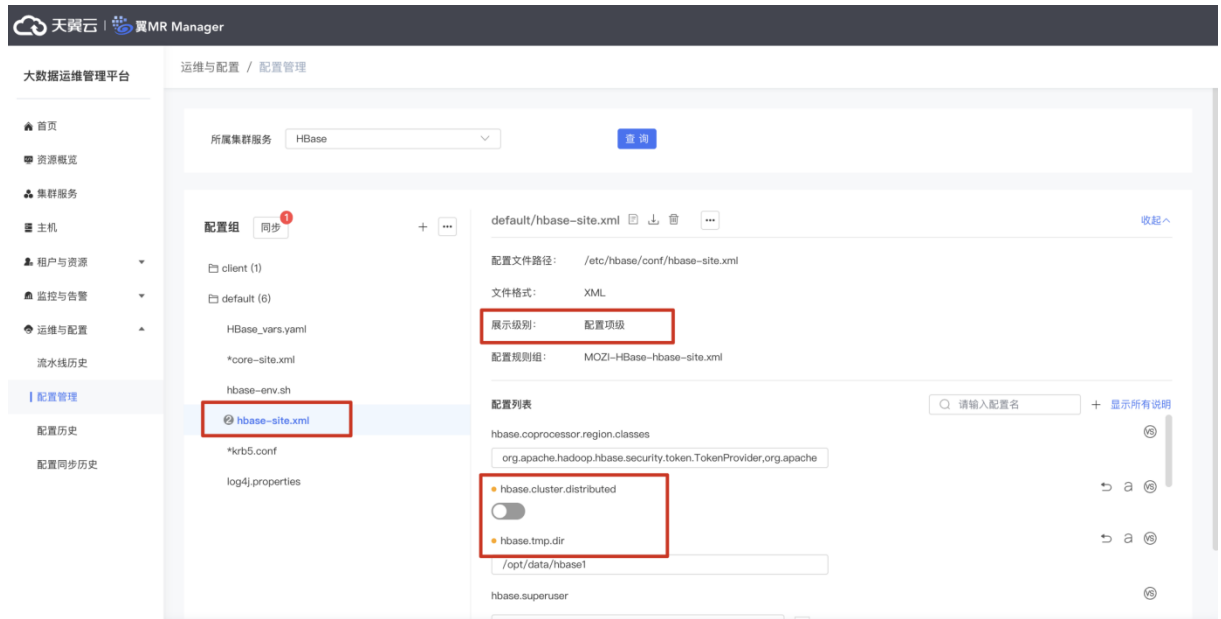
注意事项

- (1) 展示级别为：文件级。
 - 直接在文本框中进行编辑或修改。
 - 修改后且当前配置文件未同步时，配置名称前展示数字 1。如图所示：



- (2) 展示级别为：配置项级。

- 需要逐条修改配置项。
- 修改后且当前配置文件未同步时，配置名称前展示所修改的配置项数，所修改的配置项名称前展示修改标识。如图所示：



6.8.10. 配置管理-配置项对比

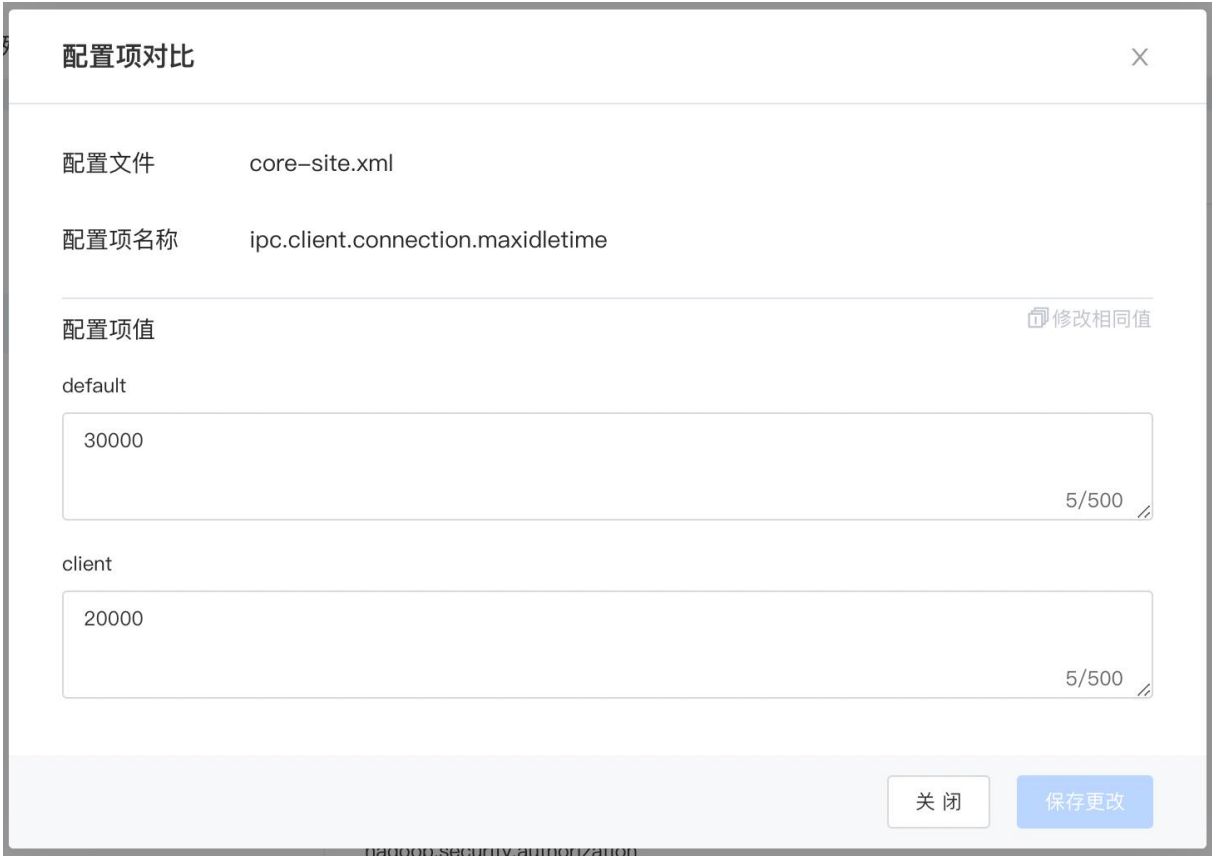
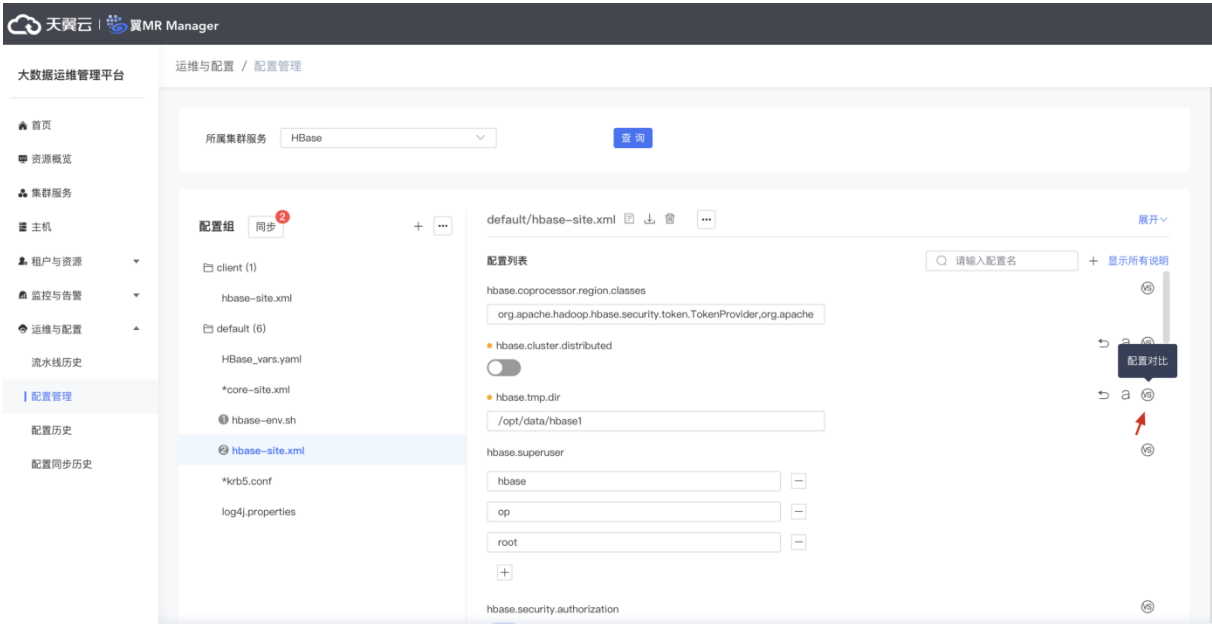
本章节主要介绍翼 MR Manager 的集群服务配置项对比的操作。

操作场景

用户可以查看同一集群下相同配置项名称在不同配置组中的值。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“配置管理”tab。
- (7) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (8) 选择要对比的配置项，单击该配置项名称右侧的“配置项对比按钮”，出现配置项对比弹框。如图所示：



注意事项

- (1) 配置项对比弹框中，当多个配置项的值不相同时，进行逐条展示。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

修改相同值

default

30000

5/500

client

20000

5/500

关闭

保存更改

(2) 配置项对比弹框中，当多个配置项的值相同时，进行合并展示。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

修改单个值

default

client

30000

关闭

保存更改

(3) 为可直接对配置项值进行修改，并支持修改相同值，也支持修改单个配置项的值。如图所示：

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

default

client

30000

修改单个值

关闭

保存更改

配置项对比

配置文件

core-site.xml

配置项名称

ipc.client.connection.maxidletime

配置项值

default

client

30000

5/500

30000

5/500

修改相同值

关闭

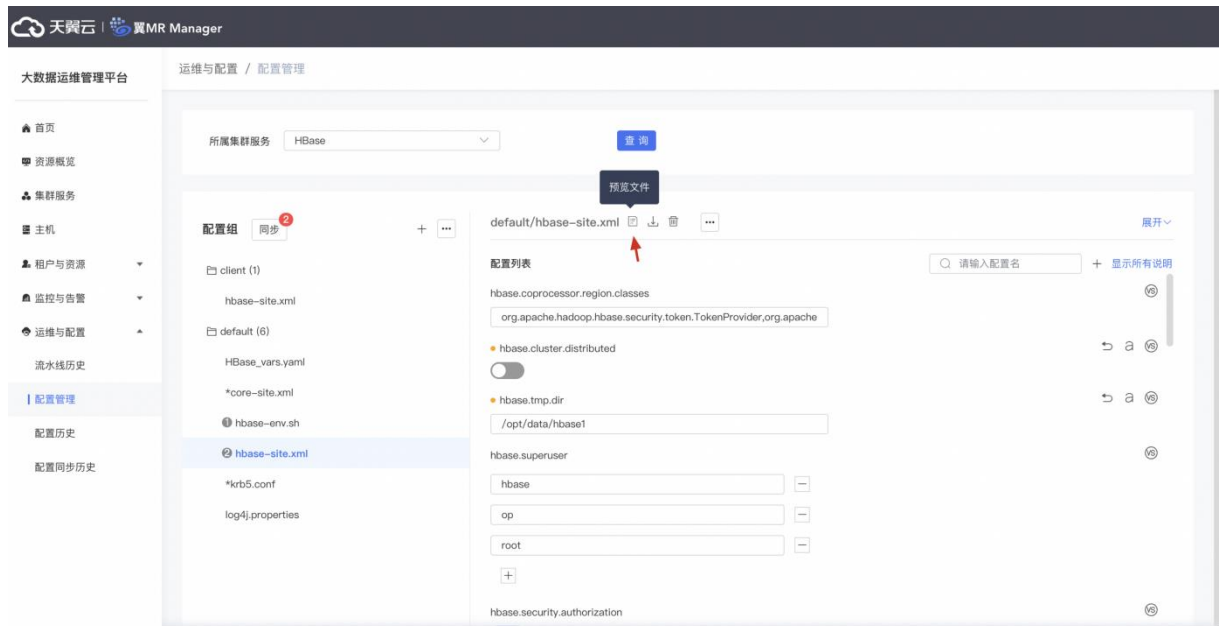
保存更改

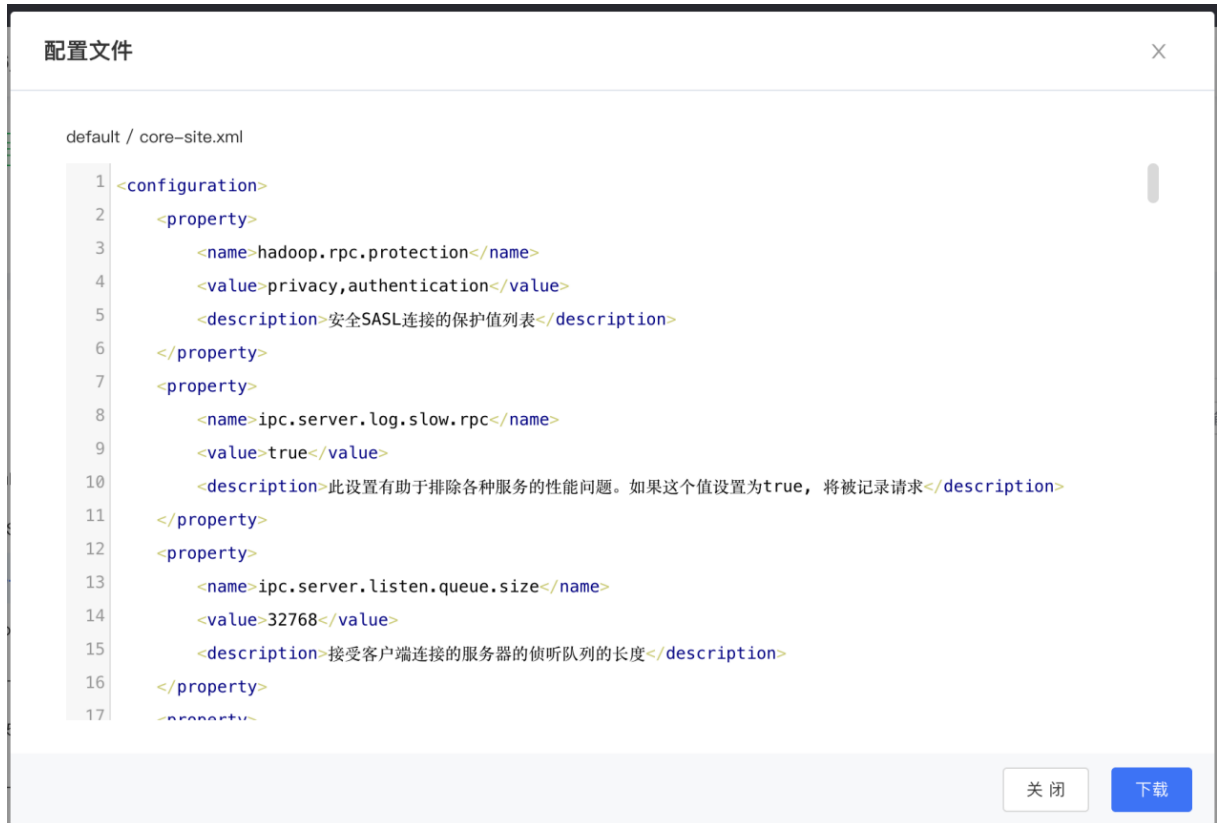
6.8.11. 配置管理-配置文件预览

本章节主要介绍翼 MR Manager 的集群服务配置文件预览的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“预览文件”，出现配置文件预览弹框。如图所示：



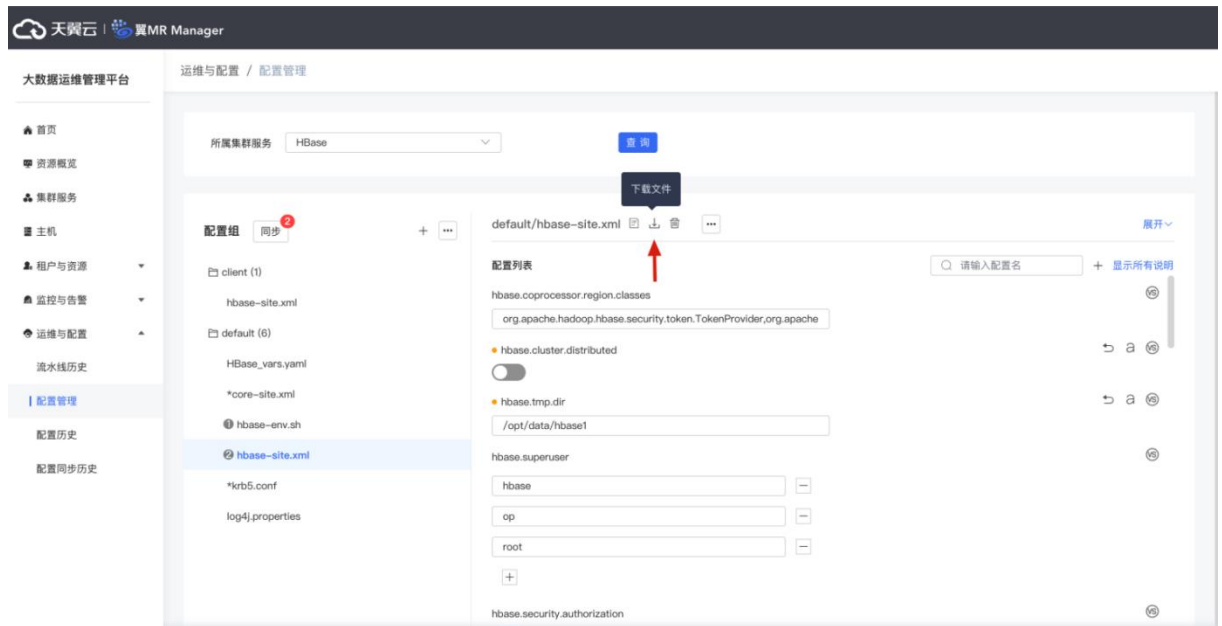


6.8.12. 配置管理-下载配置文件

本章节主要介绍翼 MR Manager 的集群服务配置文件下载的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“下载文件”，等待下载完成即可。如图所示：



6.8.13. 配置管理-配置文件历史对比

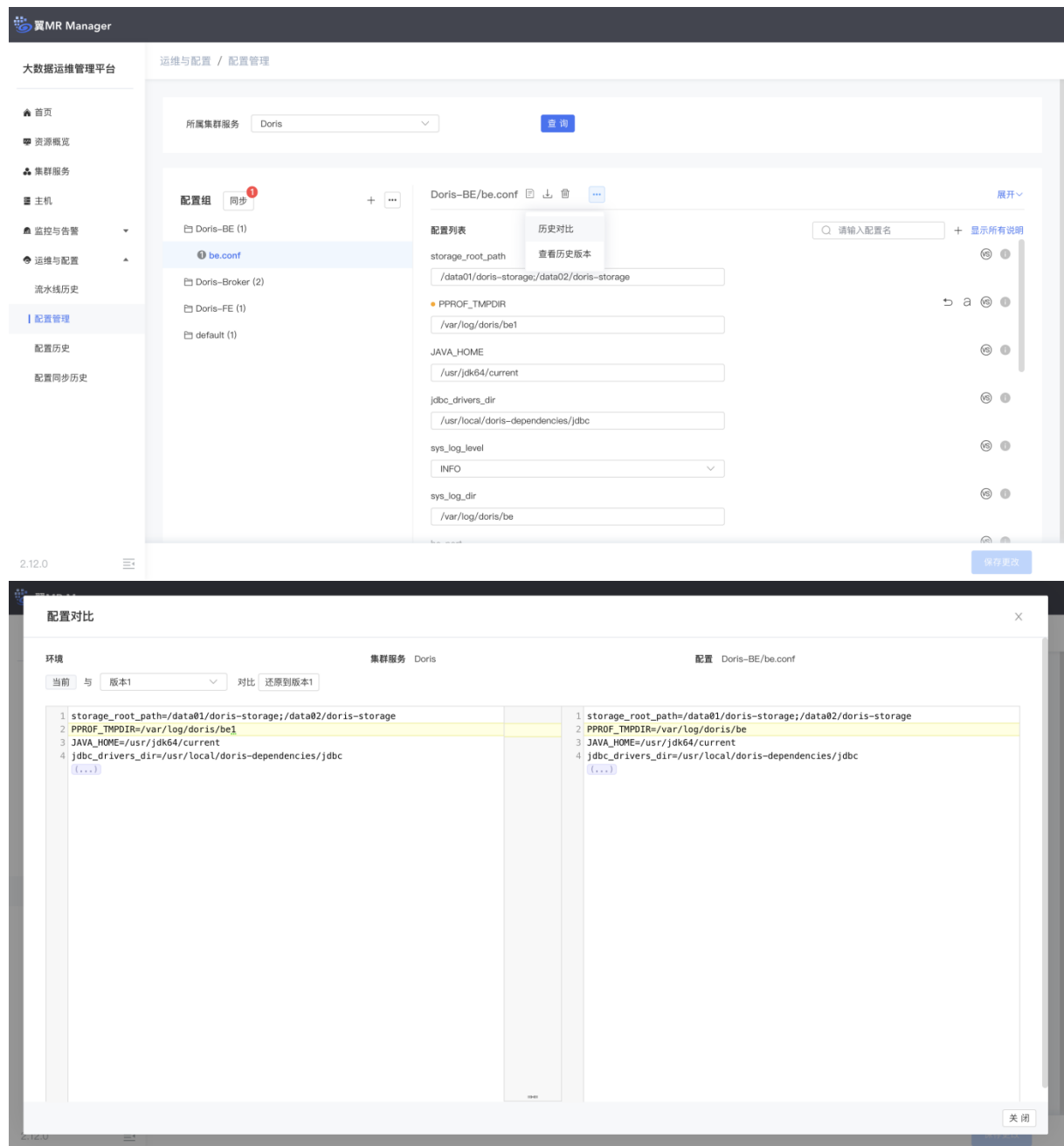
本章节主要介绍翼 MR Manager 的集群服务配置文件历史对比的操作。

操作场景

支持用户查看当前配置文件与不同历史版本的差异。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“更多 > 历史对比”，出现历史对比弹框。如图所示：



6.8.14. 配置管理-查看配置历史版本

本章节主要介绍翼 MR Manager 的集群服务查看配置历史版本的操作。

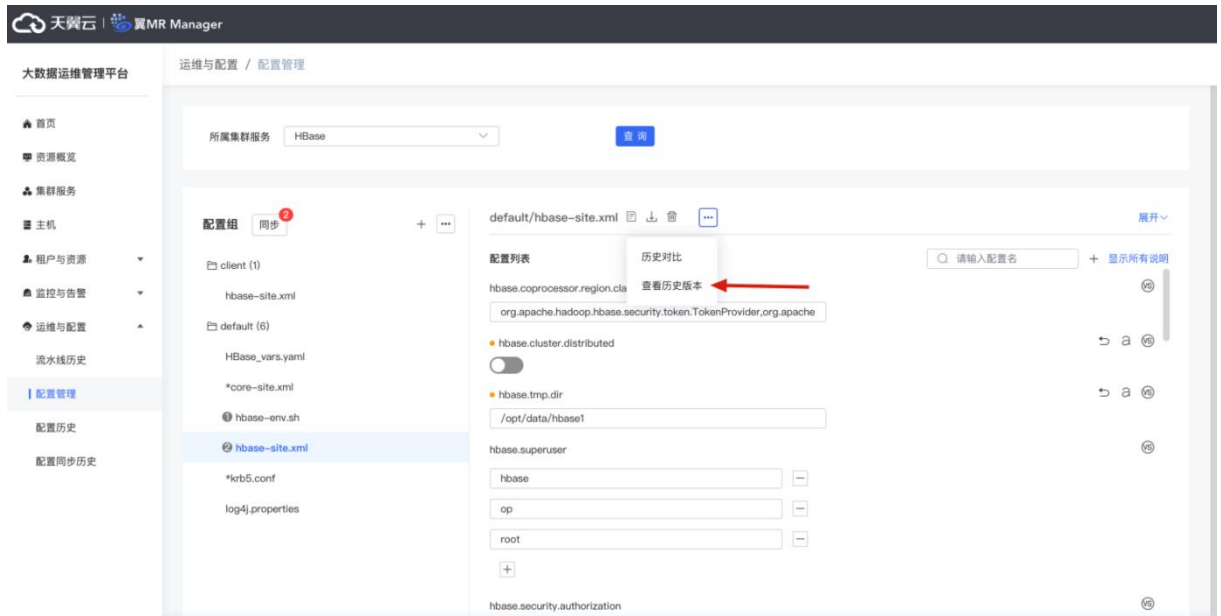
操作场景

支持用户查看指定配置文件的历史版本，不同版本的配置内容以及差异。

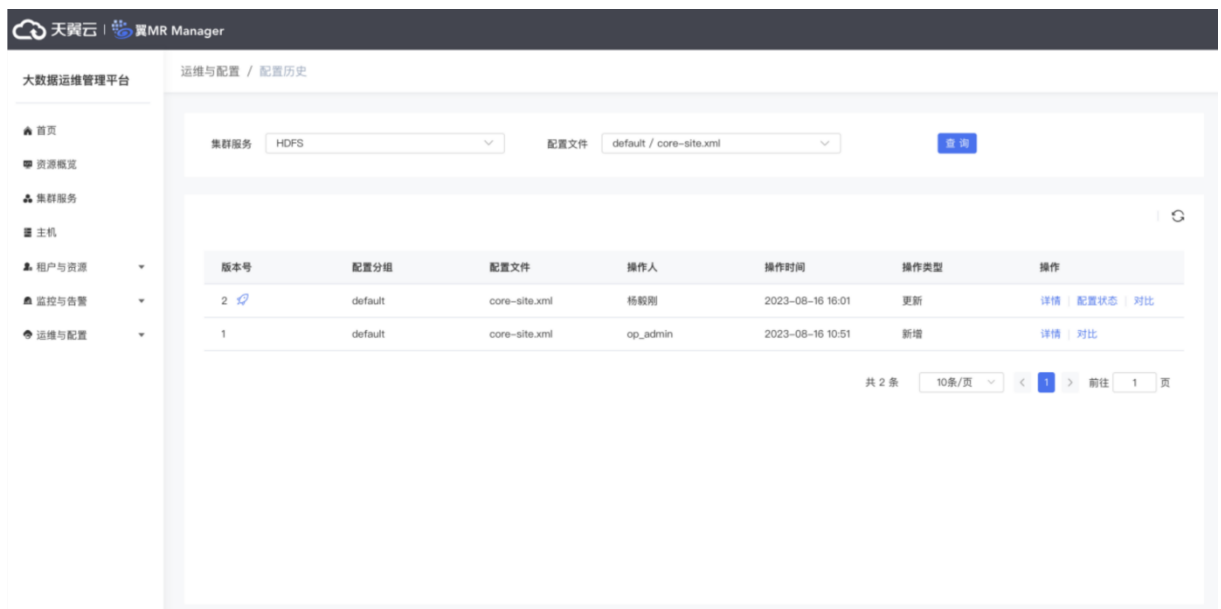
操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。

- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“更多 > 查看历史版本”。如图所示：



- (8) 进入到配置历史页面，配置历史页面展示当前配置的所有历史版本。如图所示：

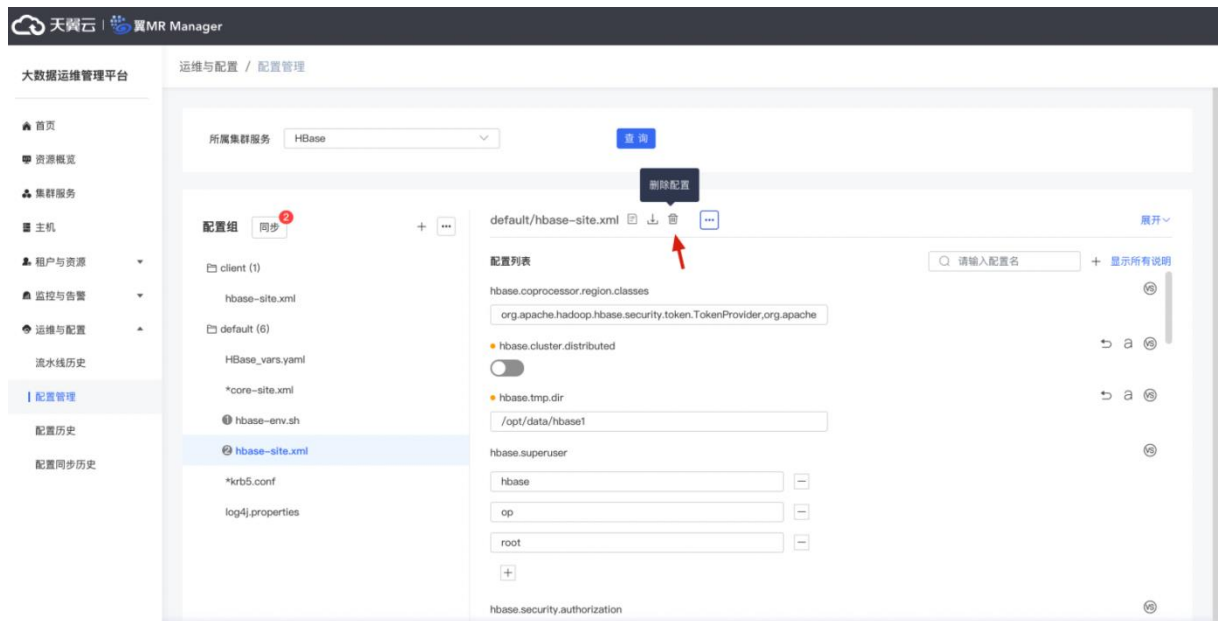


6.8.15. 配置管理-删除配置

本章节主要介绍翼 MR Manager 的集群服务删除配置的操作。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“删除配置”。如图所示：



- (8) 出现删除确认弹框，单击“确定”即可。如图所示：



6.8.16. 配置管理-配置同步

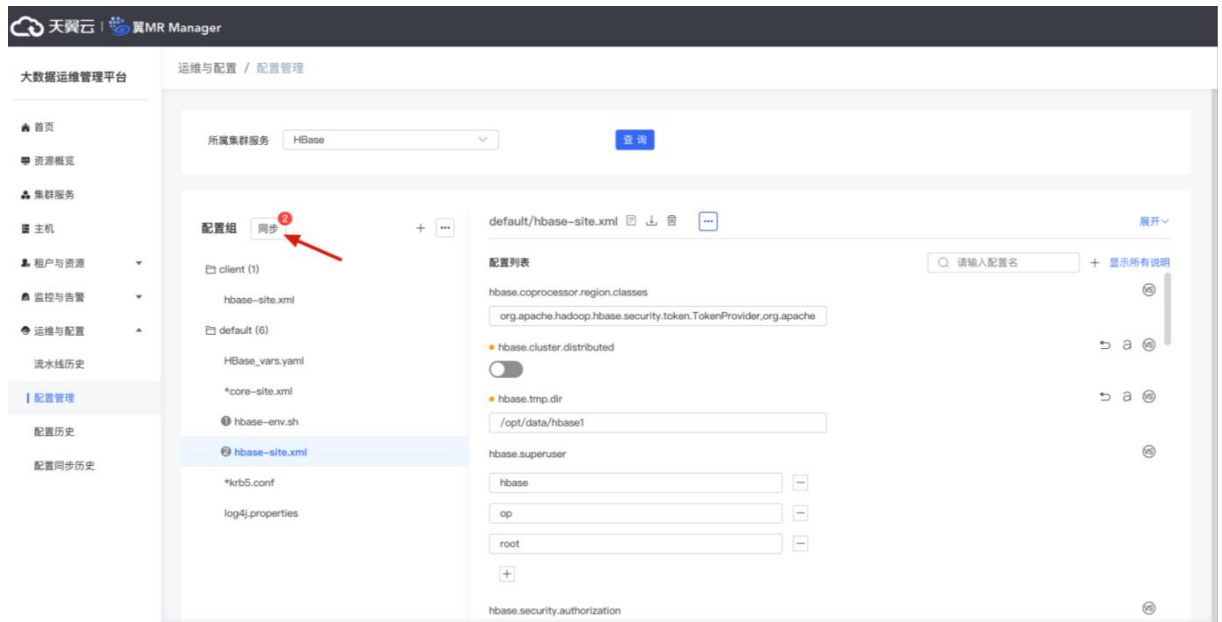
本章节主要介绍翼 MR Manager 的集群服务配置同步的操作。

操作场景

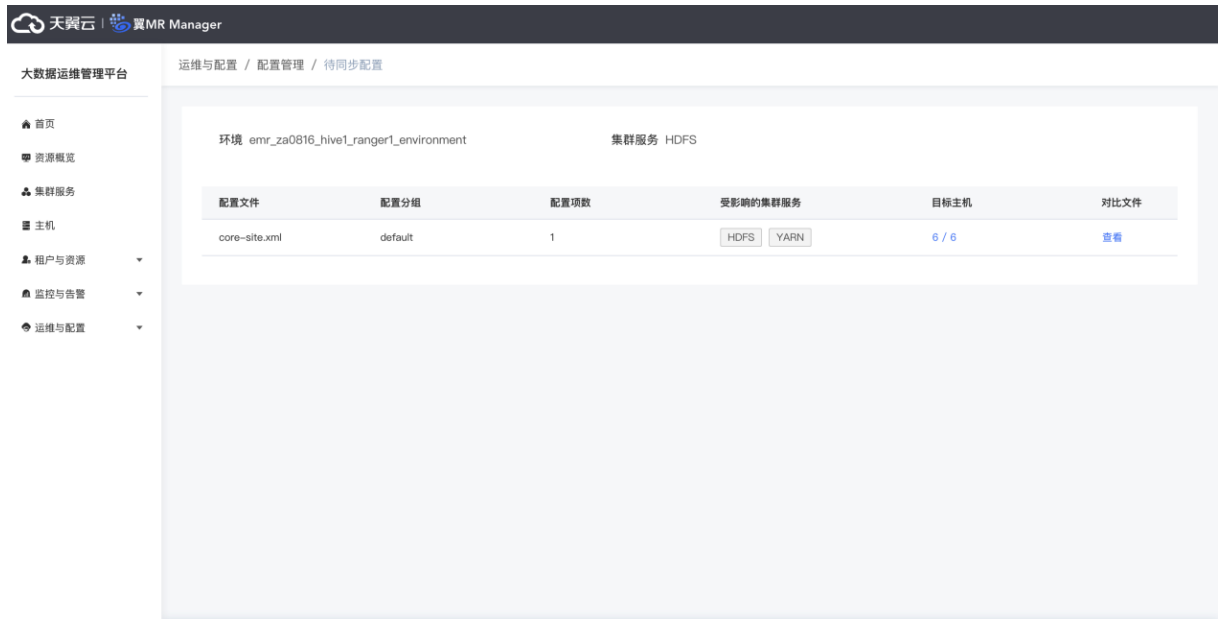
支持将新的配置文件或修改后的配置同步到相应主机上。

操作步骤

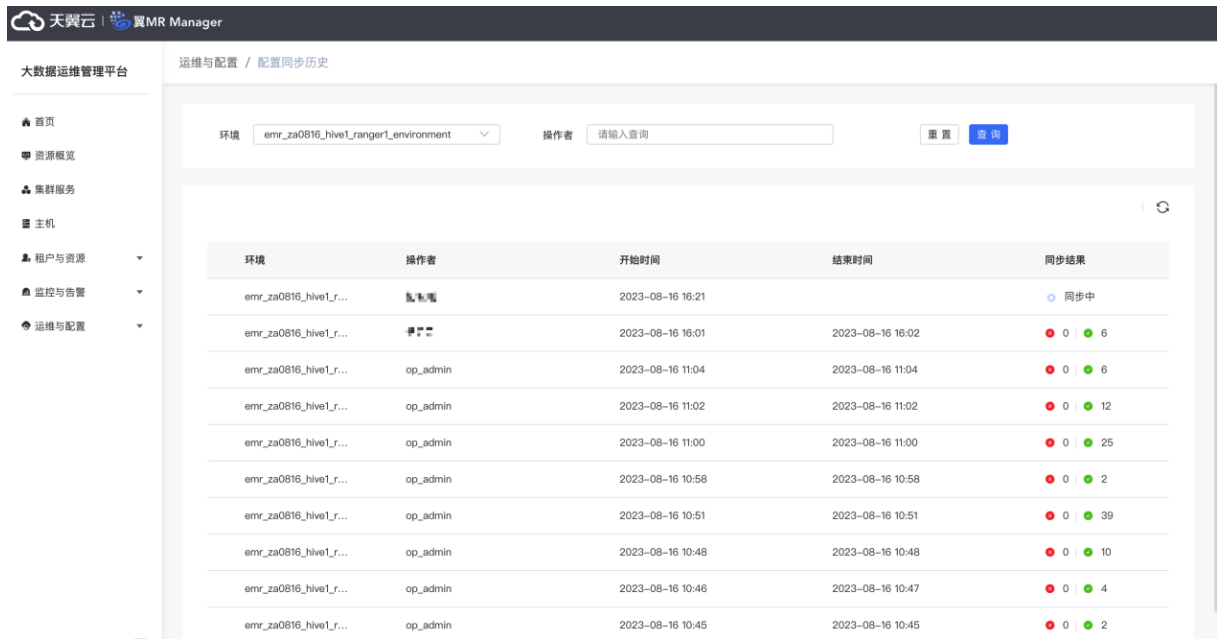
- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“同步”按钮。如图所示：



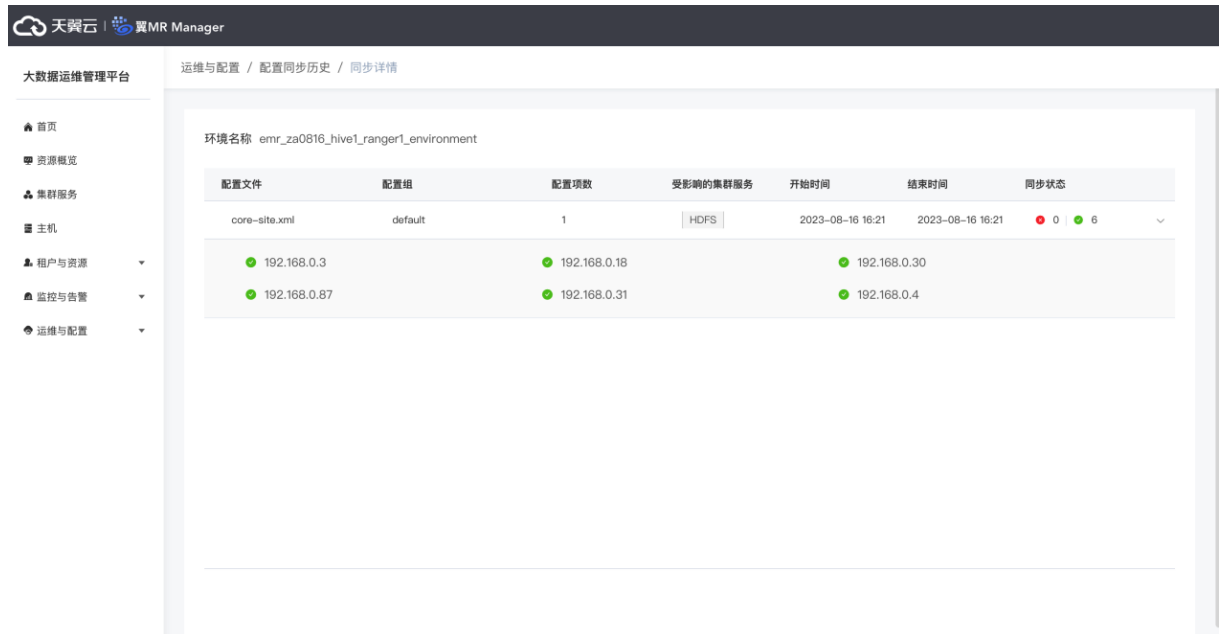
- (7) 进入待同步配置页面。如图所示：



(8) 单击“配置同步”按钮，跳转到配置同步菜单页面。如图所示：



(9) 单击“同步中”，进入到配置同步详情页面，查看配置同步进度，等待同步完成。如图所示：



6.8.17. 配置管理-查看配置状态

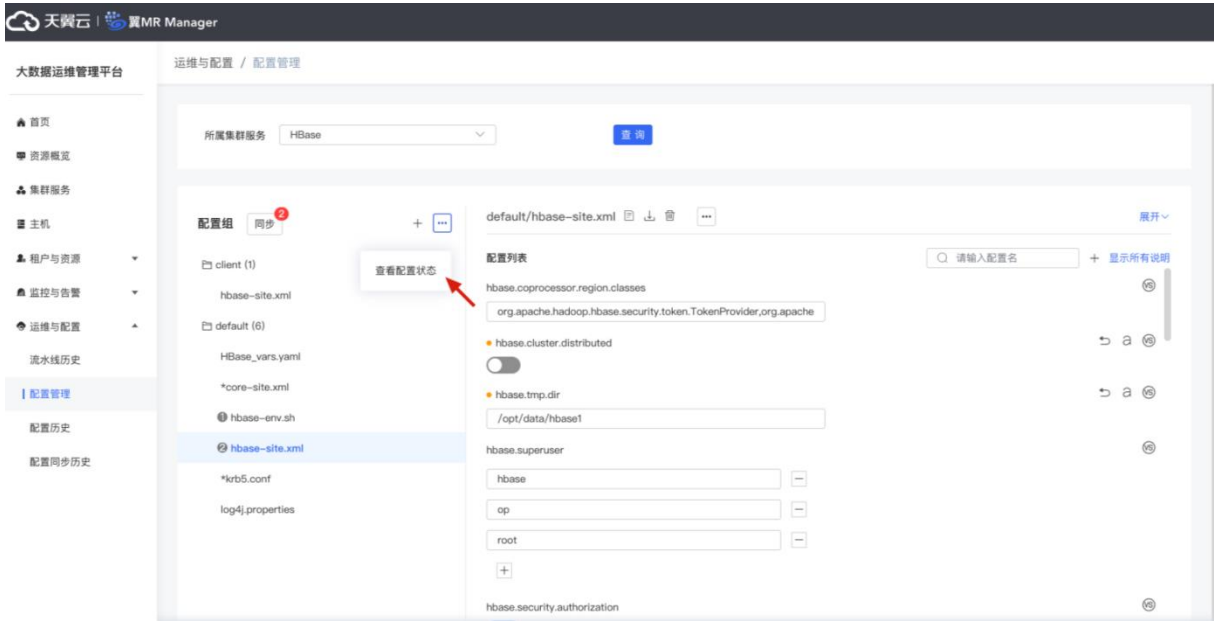
本章节主要介绍翼 MR Manager 的集群服务查看配置状态的操作。

操作场景

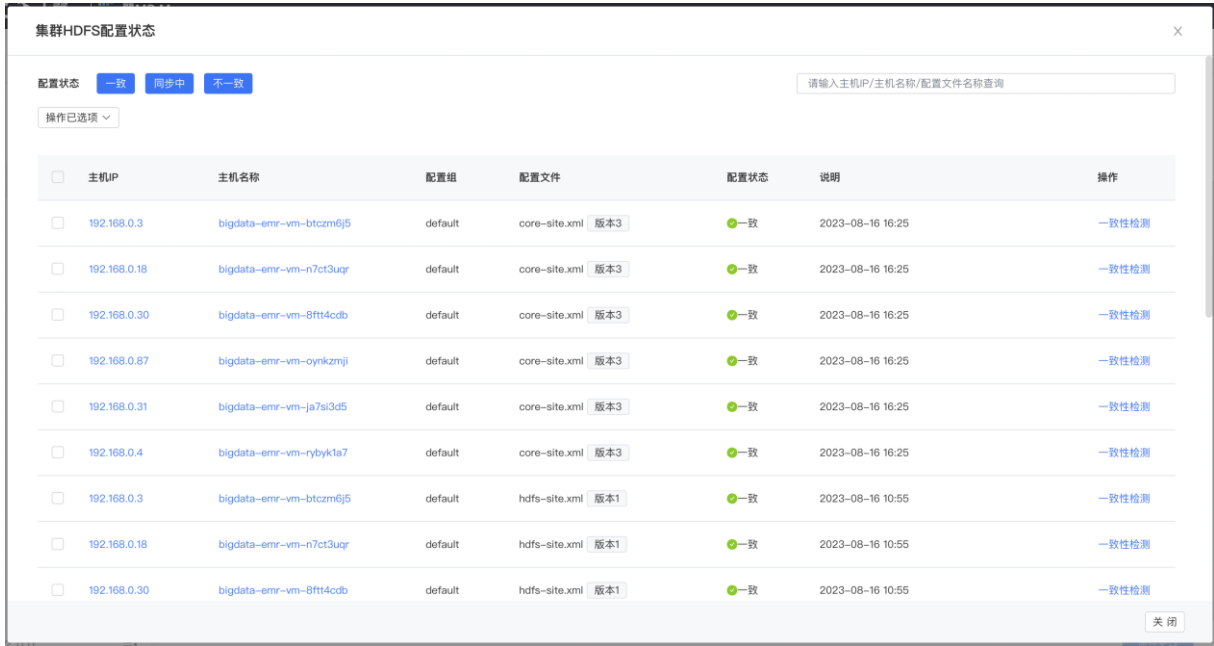
支持查看指定集群服务所有配置文件的同步状态。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“更多 > 查看配置状态”按钮。如图所示：



(7) 出现配置同步状态弹框，展示所有配置文件的同步状态。如图所示：



6.8.18. 配置管理-重新同步

本章节主要介绍翼 MR Manager 的集群服务重新同步的操作。

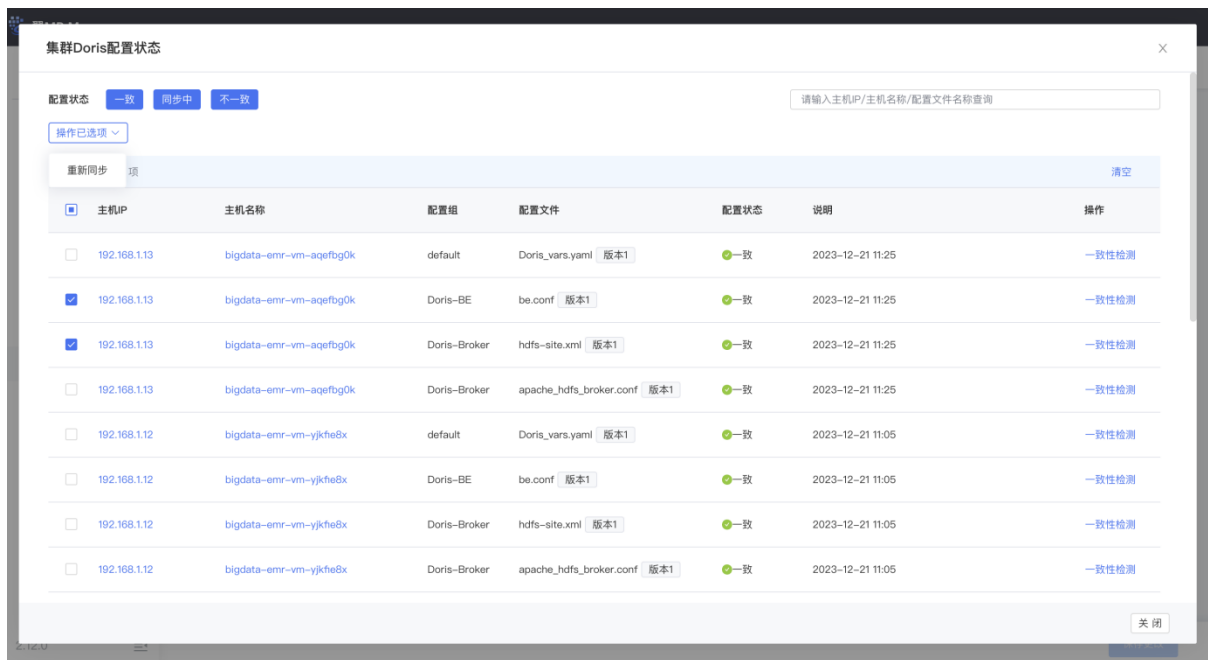
操作场景

支持对同步失败的配置或配置状态不一致的进行重新同步。

操作步骤

(1) 登录翼 MR 管理控制台。

- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击“更多 > 查看配置状态”按钮，出现配置同步状态弹框。
- (7) 勾选需要操作主机前的复选框。
- (8) 单击“操作已选项 > 重新同步”，即可支持配置同步操作。如图所示：



6.8.19. 配置管理-配置回滚

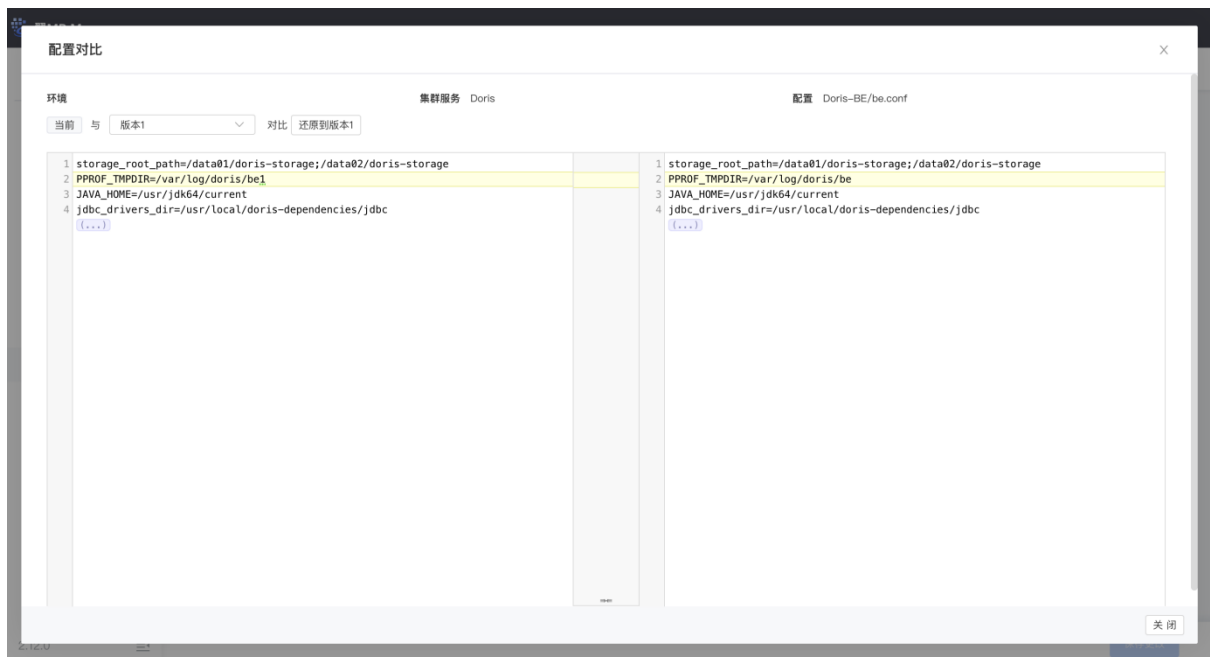
本章节主要介绍翼 MR Manager 的集群服务配置回滚的操作。

操作场景

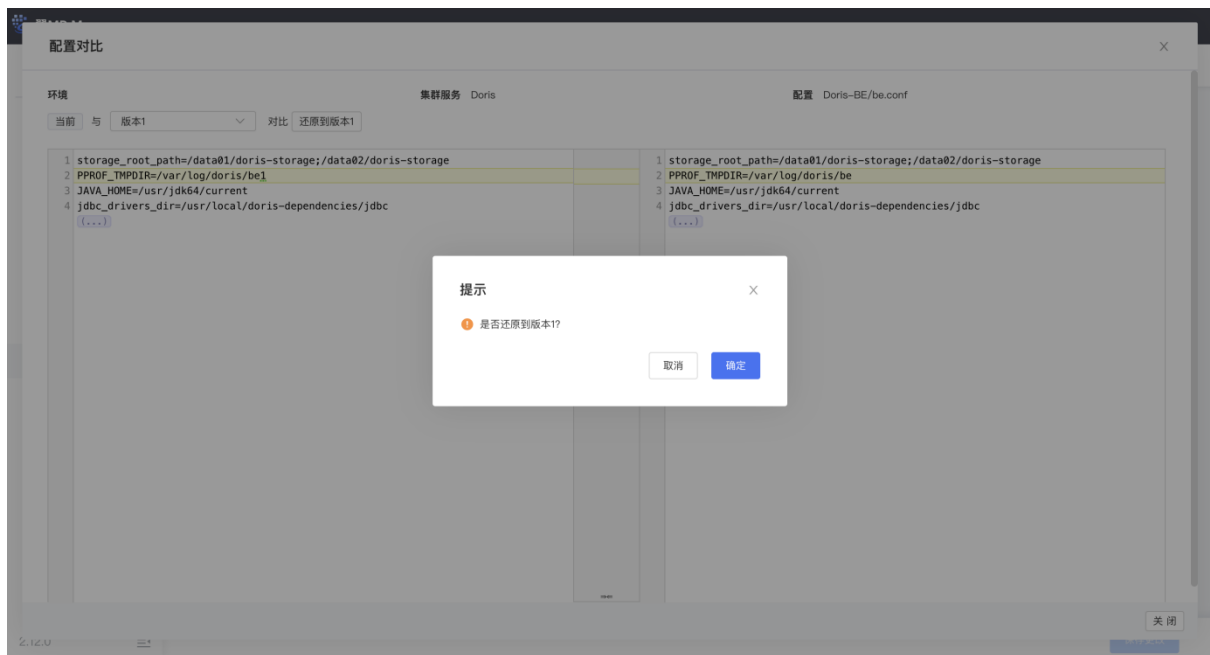
支持回退指定配置文件为某个历史版本。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置管理”。
- (5) 查询条件选择指定集群服务。
- (6) 单击配置组，单击配置文件名称，页面右侧展示该配置的详细信息。
- (7) 单击“更多 > 历史对比”，出现历史对比弹框。如图所示：



(8) 单击“还原到版本 X”按钮，出现确认操作提示弹框。如图所示：



(9) 单击“确定”。

(10) 单击“关闭”按钮，关闭配置对比弹框。

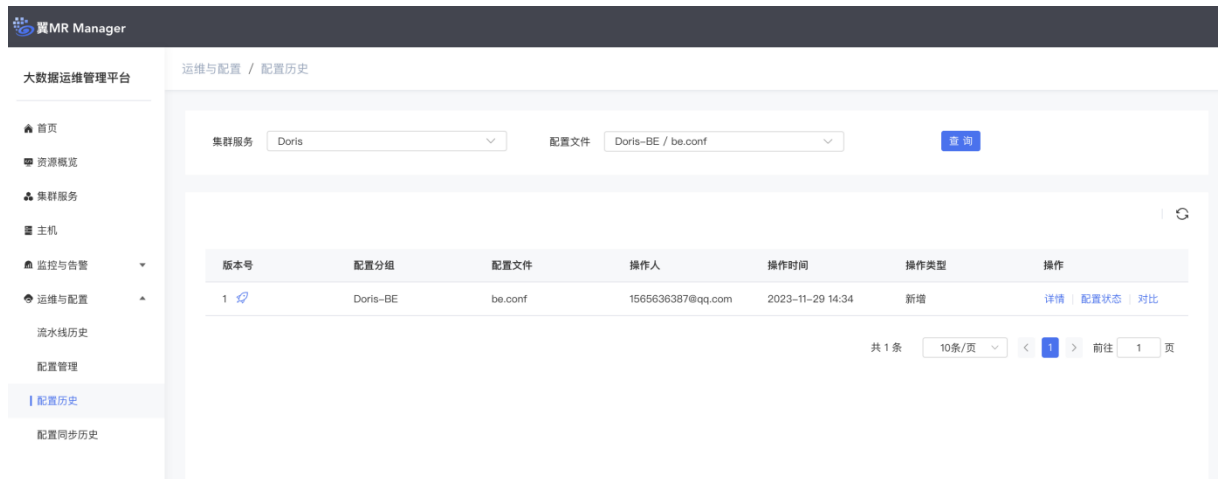
(11) 查看配置文件，则配置文件内容与所还原的版本内容一致。

(12) 单击“同步”，对配置文件进行同步操作即可。

6.8.20. 记登历史-配置历史概述

本章节主要介绍翼 MR Manager 的配置历史特性。

进入到翼 MR Manager 以后，点击菜单“运维与配置 > 配置历史”，进入配置历史页面。如图所示：



- 页面上方为查询区域。
- 支持查看不同配置文件的所有历史版本信息。
- 支持查看配置文件不同版本的配置内容、并支持不同版本之前的内容对比。
- 支持查看配置文件当前版本的同步详情。

6.8.21. 配置历史-查询配置历史

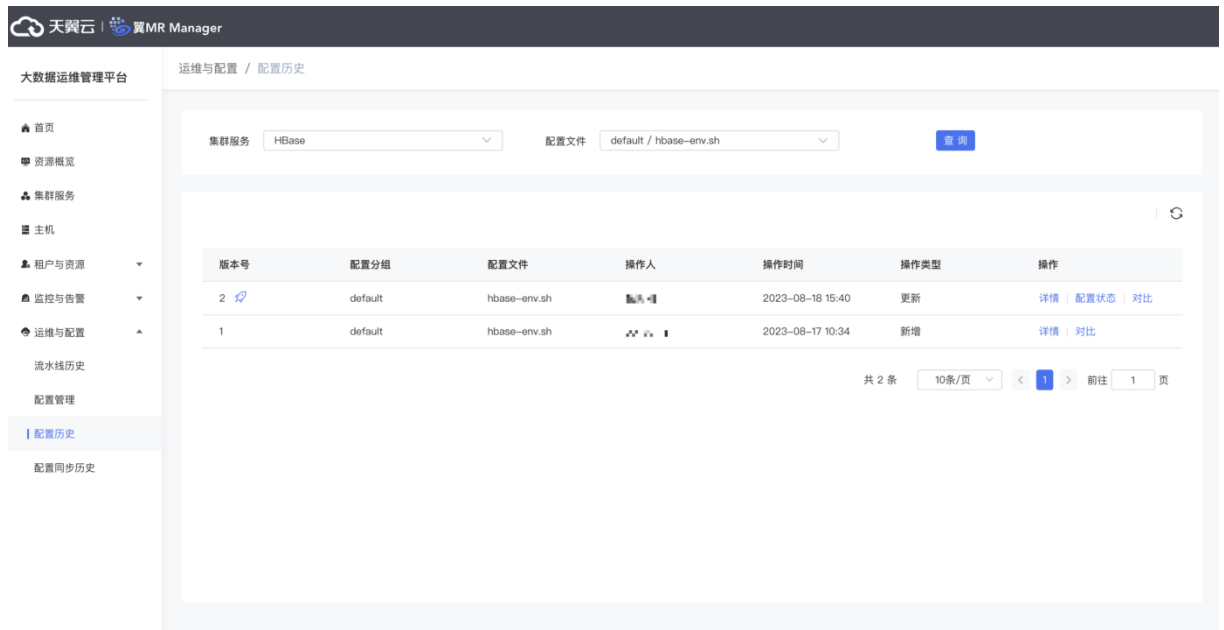
本章节主要介绍翼 MR Manager 的查询配置历史的操作。

操作场景

用户可以查询某个配置文件的历史版本。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置历史”。
- (5) 查询条件选择指定集群服务、选择指定配置组、配置文件，下方展示该配置文件的所有历史版本。如图所示：



6.8.22. 配置历史-查看配置详情

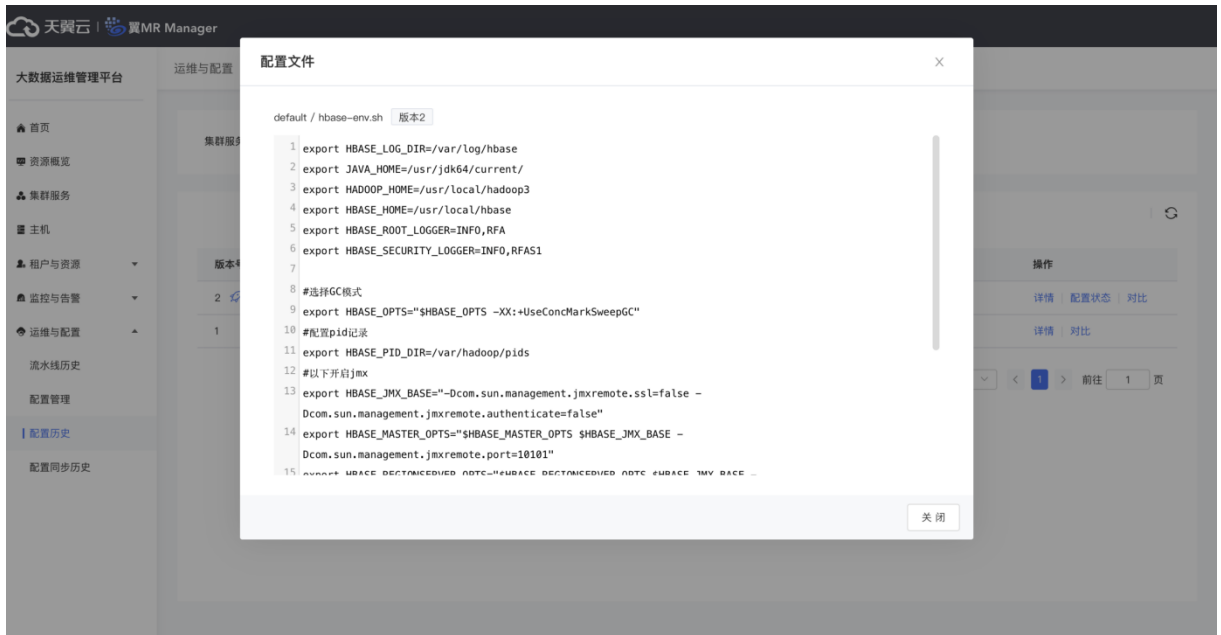
本章节主要介绍翼 MR Manager 的查看配置详情的操作。

操作场景

用户在执行配置同步后，用户需要查看不同版本配置文件的详细信息。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置历史”。
- (5) 查询条件选择指定集群服务、选择指定配置组、配置文件。
- (6) 选择要查看的版本，单击“详情”按钮，出现该版本配置文件的详情弹框。如图所示：



6.8.23. 配置历史-查看配置状态

本章节主要介绍翼 MR Manager 的查看配置状态的操作。

操作场景

支持查看最新版本的配置状态。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置历史”。
- (5) 查询条件选择指定集群服务、选择指定配置组、配置文件。
- (6) 单击“配置状态”按钮，出现配置状态弹框。如图所示：

集群HDFS配置状态						
<div>配置状态 一致 同步中 不一致</div> <div>请输入主机IP/主机名称/配置文件名称查询</div> <div>操作已选项</div>						
<input type="checkbox"/>	主机IP	主机名称	配置组	配置文件	配置状态	说明
<input type="checkbox"/>	192.168.1.7	bigdata-emr-vm-ov4o9lf	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.10	bigdata-emr-vm-khr9kvu	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.8	bigdata-emr-vm-nruezo2f	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.4	bigdata-emr-vm-fzkgpfn	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.9	bigdata-emr-vm-pqa6yskz	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.6	bigdata-emr-vm-ycbyxwbx	default	core-site.xml 版本2	一致	2023-10-20 14:20
<input type="checkbox"/>	192.168.1.7	bigdata-emr-vm-ov4o9lf	default	hadoop-env.sh 版本1	一致	2023-10-19 10:30
<input type="checkbox"/>	192.168.1.10	bigdata-emr-vm-khr9kvu	default	hadoop-env.sh 版本1	一致	2023-10-19 10:30
<input type="checkbox"/>	192.168.1.8	bigdata-emr-vm-nruezo2f	default	hadoop-env.sh 版本1	一致	2023-10-19 10:30
关闭						

6.8.24. 配置历史-查看配置对比

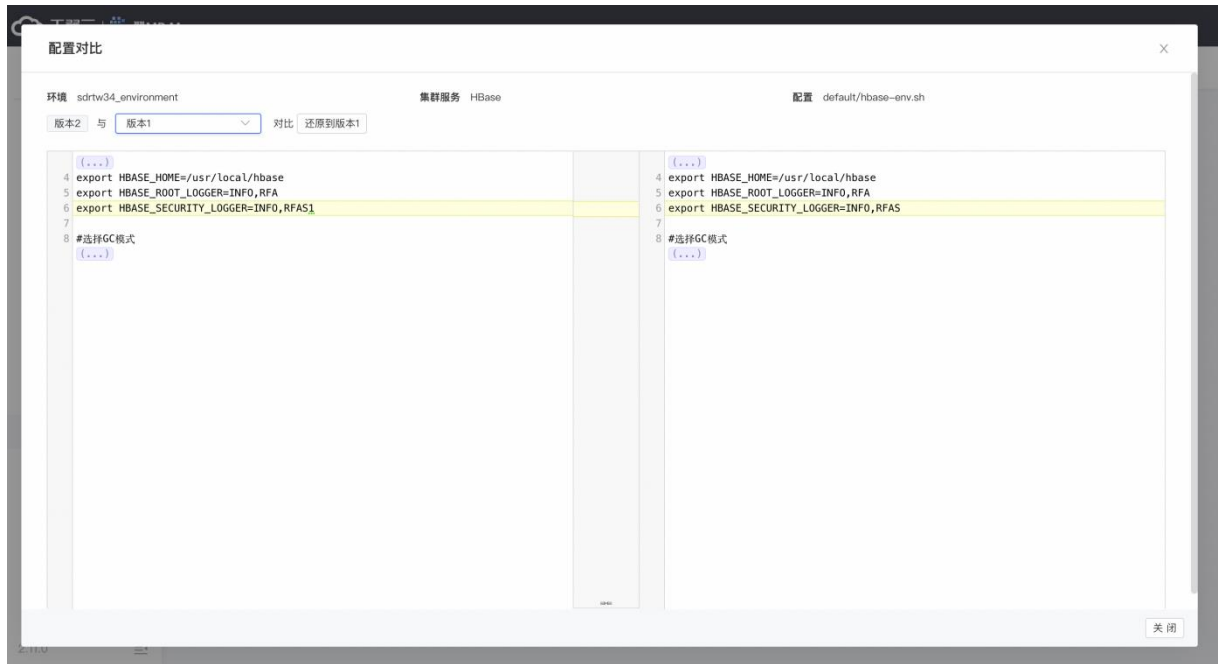
本章节主要介绍翼 MR Manager 的查看配置对比的操作。

操作场景

支持查看配置文件不同版本之前的差异。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置历史”。
- (5) 查询条件选择指定集群服务、选择指定配置组、配置文件。
- (6) 单击“对比”按钮，出现配置对比弹框。如图所示：



6.8.25. 配置历史-重新同步

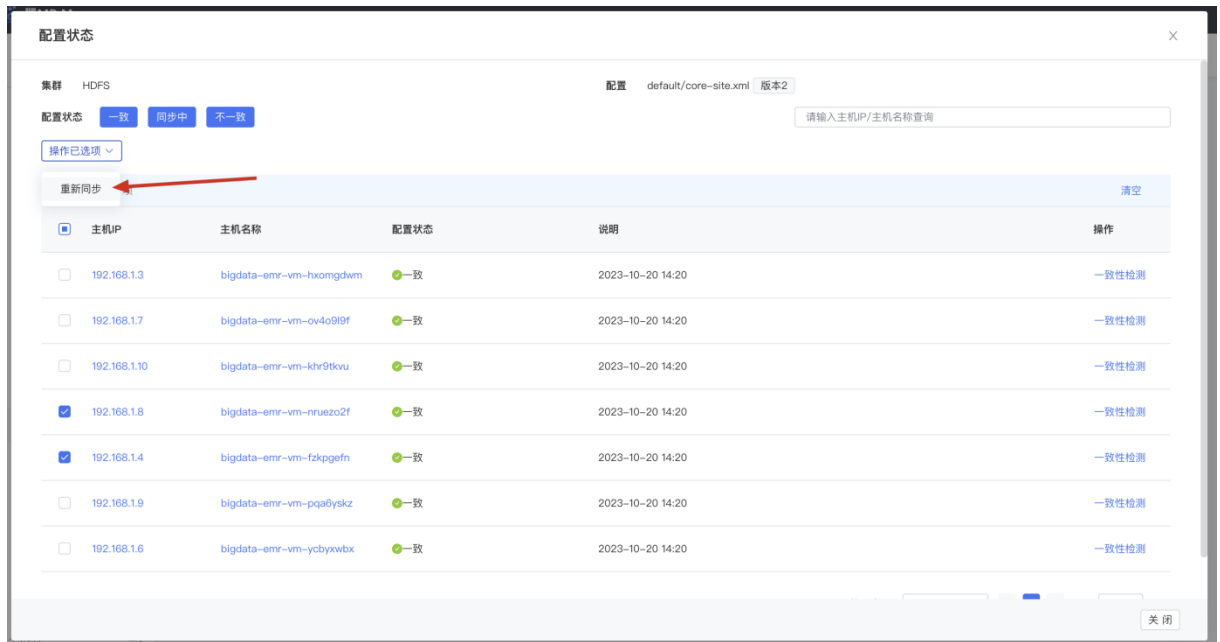
本章节主要介绍翼 MR Manager 的重新同步的操作。

操作场景

支持查看最新版本的配置状态。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置历史”。
- (5) 查询条件选择指定集群服务、选择指定配置组、配置文件。
- (6) 单击“配置状态”按钮，出现配置状态弹框。
- (7) 勾选需要操作主机前的复选框。
- (8) 单击“操作已选项 > 重新同步”，即可支持配置同步操作。如图所示：

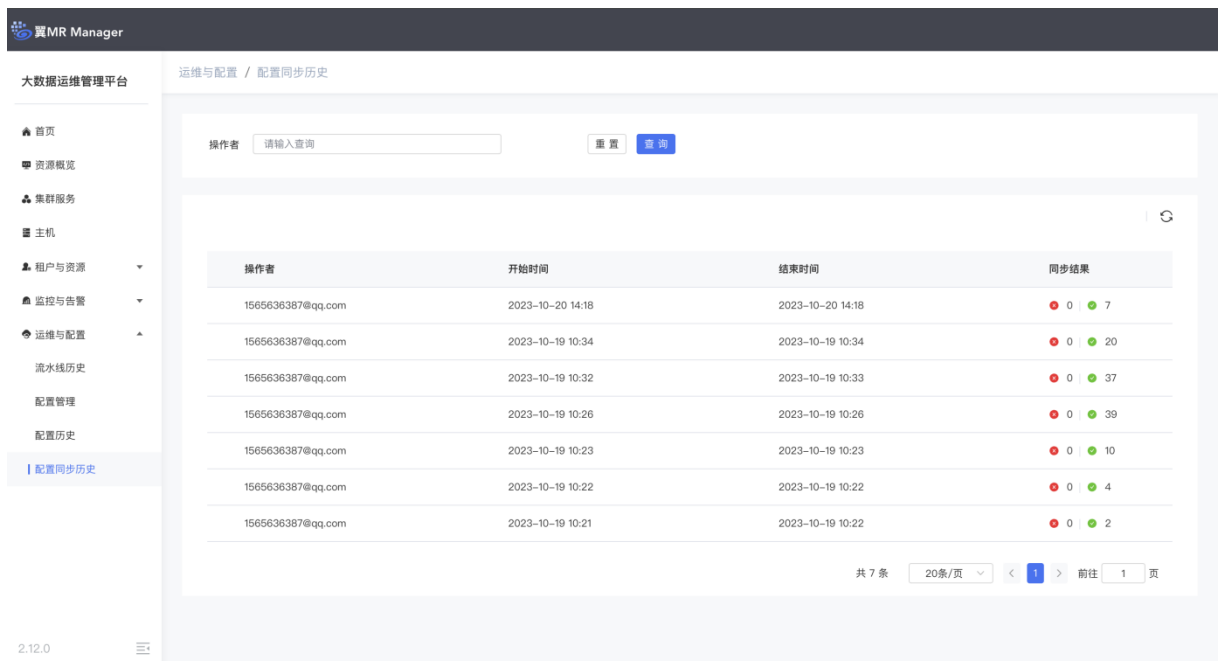


6.8.26. 配置同步历史-配置同步历史概述

本章节主要介绍翼 MR Manager 的配置同步历史特性。

进入到翼 MR Manager 以后,点击菜单“运维与配置 > 配置同步历史”,进入配置同步历史页面。

如图所示:



- 页面上方为查询区域。
- 展示所有配置同步操作的历史、配置同步操作人，以及配置同步详情。

6.8.27. 配置同步历史-查看配置同步状态

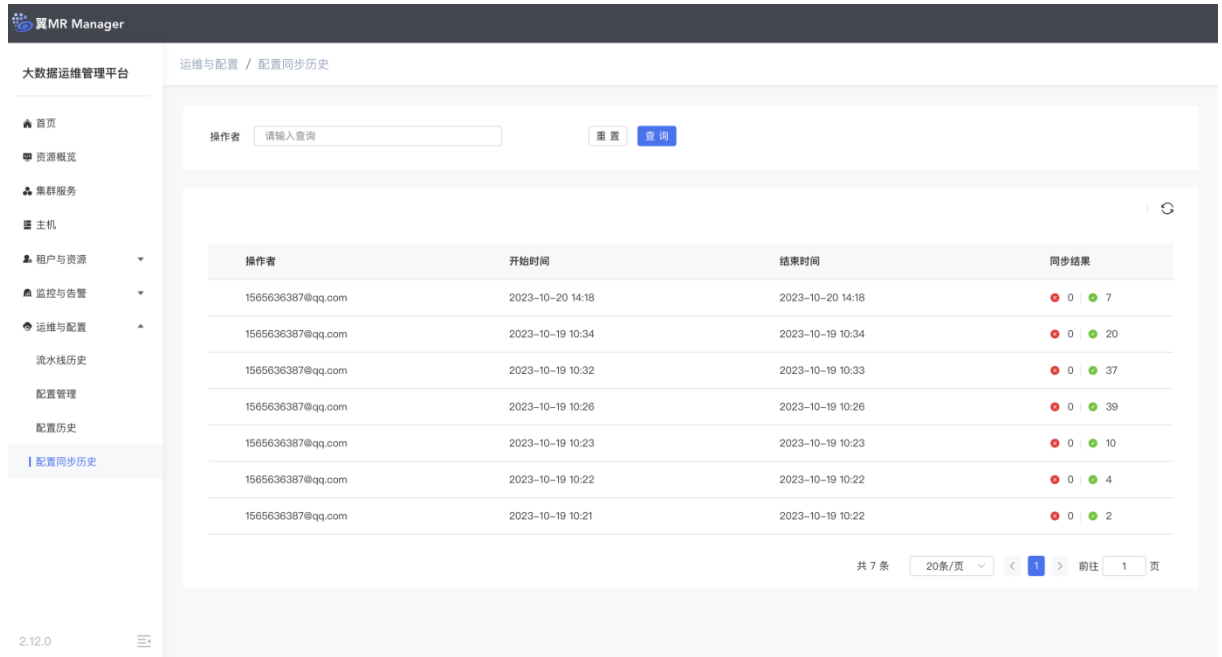
本章节主要介绍翼 MR Manager 的查看配置同步状态的操作。

操作场景

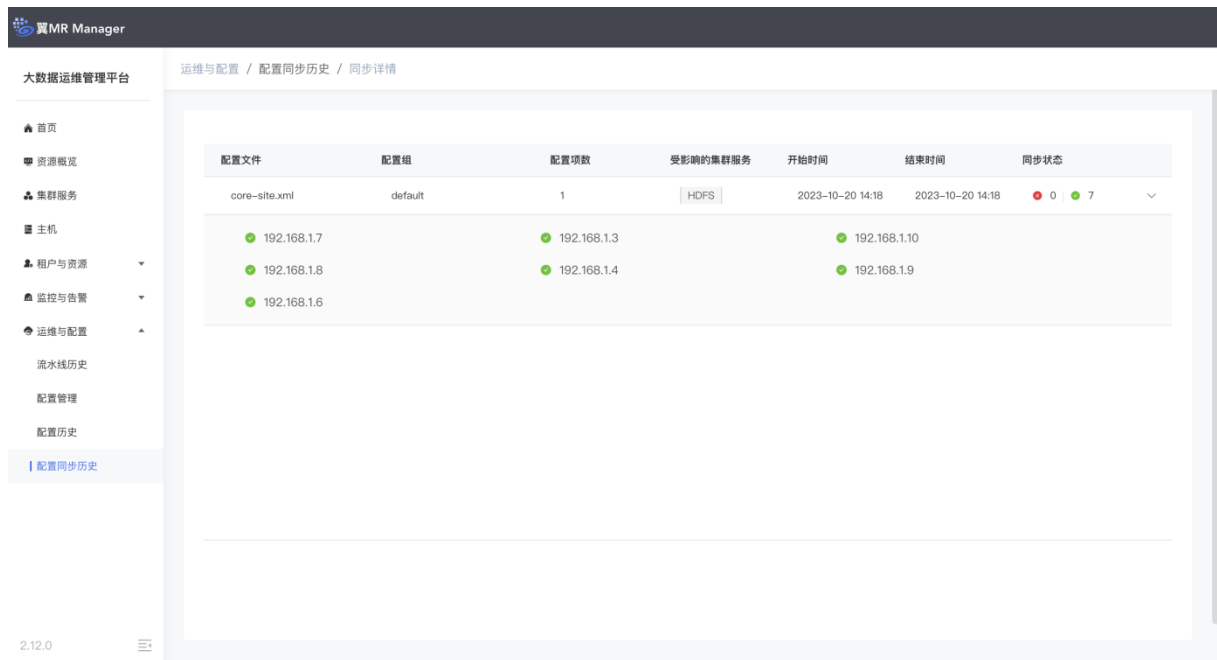
用户在配置管理进行配置同步后，需要查看本次同步操作的同步状态或同步结果。

操作步骤

- (1) 登录翼 MR 管理控制台。
- (2) 单击“我的集群”，单击指定的集群名称，进入集群信息页面。
- (3) 单击“翼 MR Manager”tab，单击“前往翼 MR Manager”。
- (4) 进入到翼 MR Manager 以后，单击菜单“运维与配置 > 配置同步历史”。如图所示：



- (5) 选择要查看的操作记录，单击“同步中”或同步结果，进入配置同步详情页面。如图所示：



7. 安全认证原理和认证机制

本章节主要介绍安全认证原理和认证机制。

功能

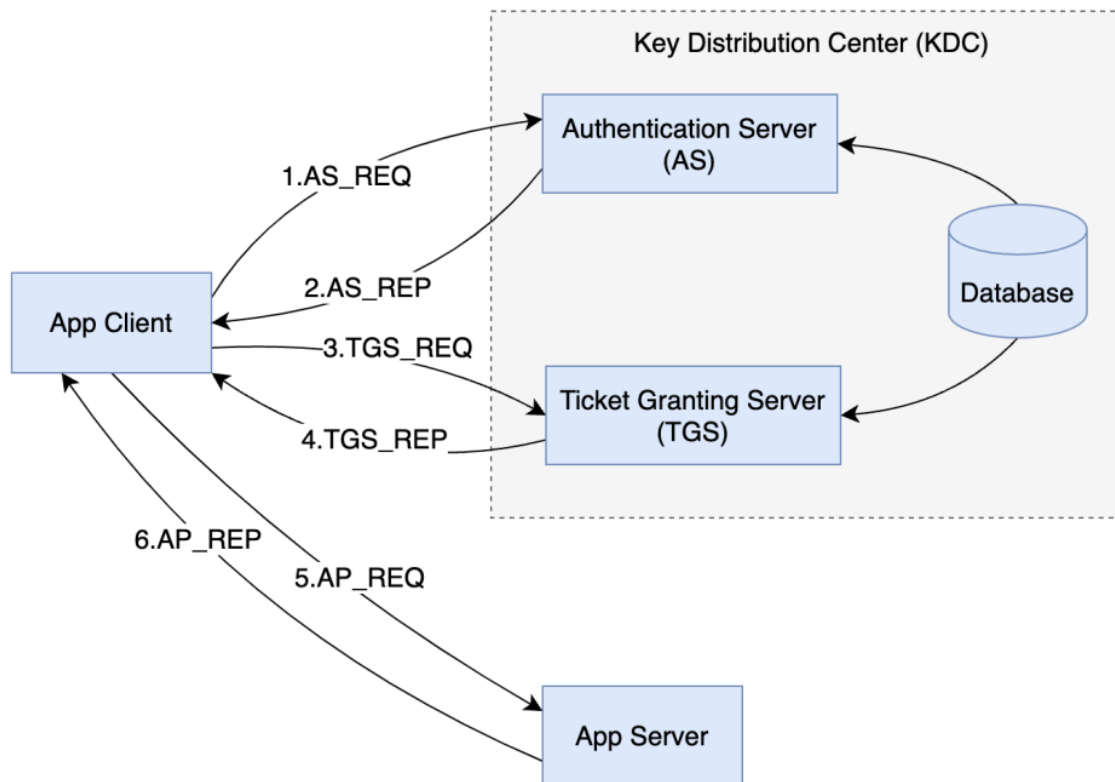
开启了 Kerberos 认证的安全模式集群，进行应用开发时需要进行安全认证。

Kerberos 这一名词来源于希腊神话“三个头的狗——地狱之门守护者”，后来沿用作为安全认证的概念，使用 Kerberos 的系统在设计上采用“客户端/服务器”结构与 AES 等加密技术，并且能够进行相互认证（即客户端和服务端均可对对方进行身份认证）。可以用于防止窃听、防止 replay 攻击、保护数据完整性等场合，是一种应用对称密钥体制进行密钥管理的系统。

原理

Kerberos 的原理架构和各模块的说明如下图所示：

原理架构



模块说明

- **App Client**: 应用客户端，通常是需要提交任务（或者作业）的应用程序。
- **App Server**: 应用服务端，通常是应用客户端需要访问的应用程序。
- **Key Distribution Center (KDC)**: 提供安全认证的服务。**Database**: 存储 Principal 数据
 - **Authentication Server (AS)**: 认证服务器，认证客户端身份，发放客户访问 TGS 的票据授权票据 (TGT)。
 - **Ticket Granting Server (TGS)**: 票据授予服务器，发放应用客户端访问应用服务端所需的服务票据 (ST)。

步骤原理说明

应用客户端 (**App Client**) 可以是集群内某个服务，也可以是客户二次开发的一个应用程序，应用程序可以向应用服务提交任务或者作业。

(1) **AS_REQ**: **App Client** 在提交任务或者作业前，需要向 **AS** 申请 **TGT**，用于建立和 **TGS** 的安全会话。

(2) **AS_REP**: **AS** 在收到 **TGT** 请求后，会解析其中的参数来生成对应的 **TGT**，使用 **App Client** 指定的用户名的密钥进行加密响应消息。

(3) **TGS_REQ**: **App Client** 收到 **TGT** 响应消息后，解析获取 **TGT**，此时，再由 **App Client** (通常是 **rpc** 底层) 向 **TGS** 获取应用服务端所需的 **ST**。

(4) **TGS_REP**: **TGS** 在收到 **ST** 请求后，校验其中的 **TGT** 合法后，生成对应的 **App**

Server 的 ST，再使用 App Server 密钥将响应消息进行加密处理。

(5) AP_REQ: App Client 收到 ST 响应消息后，将 ST 打包到发给 App Server 的消息里面传输给对应的 App Server。

(6) AP_REP: App Server 端收到请求后，使用 App Server 对应的密钥解析其中的 ST，并校验成功后，本次请求合法通过。

说明

(1) Kerberos 认证时需要配置 Kerberos 认证所需要的文件参数，主要包含 keytab 路径，Kerberos 认证的 principal，Kerberos 认证所需要的客户端配置 krb5.conf 文件。

(2) 方法 login()为调用的 UserGroupInformation 的方法执行 Kerberos 认证，生成 TGT 票据。

(3) 方法 doSth()调用 hadoop 的接口访问文件系统，此时底层 RPC 会自动携带 TGT 去 Kerberos 认证，生成 ST 票据。

基本概念

- 票据授权票据 (Ticket-Granting Ticket, TGT)：由 AS 生成，提供给应用程序与 TGS 建立认证安全会话，该票据的默认有效期为 24 小时，24 小时后该票据自动过期。
- 服务票据 (Server Ticket, ST)：由 TGS 生成，提供给应用程序与应用服务建立安全会话，该票据一次性有效。
- 用户身份 (Principal)：用于标识用户或者服务，格式一般为<用户名或服务名>/<主机名>，例如 hdfs/host1, user/host2。

样例

以 HDFS 为例，前提：HDFS 集群已经开启 Kerberos 认证。

通过 UserGroupInformation

```
package bigdata.hdfs.examples;

import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileStatus; import
org.apache.hadoop.fs.FileSystem;

import org.apache.hadoop.fs.Path;

import org.apache.hadoop.security.UserGroupInformation;
public class KerberosTest {
private FileSystem fs;
private Configuration conf;
/**
 * initialize Configuration
 */
private void initConf() {
conf = new Configuration();
//add configuration files
//PATH_TO_HDFS_SITE_XML 是 hdfs-site.xml 的路径
//PATH_TO_CORE_SITE_XML 是 core-site.xml 的路径
conf.addResource(new Path(PATH_TO_HDFS_SITE_XML));
conf.addResource(new Path(PATH_TO_CORE_SITE_XML));
}

/**
 * login Kerberos to get TGT, if the cluster is in security mode
 * @throws IOException if login is failed
 */

private void login() throws IOException {
// not security mode, just return
if (!
"kerberos".equalsIgnoreCase(conf.get("hadoop.security.authenticat
```

```
private void login() throws IOException {
    // not security mode, just return
    if (!
        "kerberos".equalsIgnoreCase(conf.get("hadoop.security.authenticat
        ion"))) {
        return;
    }

    //security mode
    //PATH_TO_KRB5_CONF 是 krb5.conf 的路径

    System.setProperty("java.security.krb5.conf", PATH_TO_KRB5_CONF);

    UserGroupInformation.setConfiguration(conf);

    //PATH_TO_KEYTAB 是 keytab 的路径
    //PRINCIPAL_NAME 是 principal 名称

    UserGroupInformation.loginUserFromKeytab (PRINCIPAL_NAME,
        PATH_TO_KEYTAB);
}

/**
 * initialize FileSystem, and get ST from Kerberos
 * @throws IOException
 */

private void initFileSystem() throws IOException {
    fs = FileSystem.get(conf);
}

/**
 * An example to access the HDFS
 * @throws IOException
 */

private void doSth() throws IOException {
```



```
/**
 * An example to access the HDFS
 * @throws IOException
 */

private void doSth() throws IOException {
    Path path = new Path("/tmp");
    FileStatus fStatus = fs.getFileStatus(path);
    System.out.println("Status of " + path + " is " + fStatus);
    //other thing
}

public static void main(String[] args) throws Exception {
    KerberosTest test = new KerberosTest();
    test.initConf();
    test.login();
    test.initFileSystem();
    test.doSth();
}
}
```

通过 kinit 命令，执行 hadoop 命令

```
kinit -kt /etc/security/keytabs/hdfs.keytab hdfs/host1
hadoop fs -ls /tmp
```

8. 高危操作一览表

本章节主要介绍翼 MapReduce 服务中具有高风险性的操作说明，如不当操作，可能会触发翼 MR 服务的不可用。

禁用操作

下表中描述了在集群操作与维护阶段，进行日常操作时应注意的禁用操作。

类别	操作风险
严禁删除 ZooKeeper 相关数据目录	HDFS/YARN/HBase/Hive 等很多组件都依赖于 ZooKeeper，在 ZooKeeper 中保存元数据信息。删除 ZooKeeper 中相关数据目录将会影响相关组件的正常运行。
严禁手动修改 Hive 元数据库的数据（Hivemeta 数据库）	修改 Hive 元数据可能会导致 Hive 数据解析错误，Hive 无法正常提供服务。
禁止对 Hive 的元数据表手动进行 insert 和 update 操作	修改 Hive 元数据可能会导致 Hive 数据解析错误，Hive 无法正常提供服务。
严禁修改 Hive 私有文件目录的权限	修改该目录权限可能会导致 Hive 服务不可用。
严禁修改 Kafka 配置文件中 broker.id	修改 Kafka 配置文件中 broker.id 将会导致该节点数据失效。
严禁修改节点主机名	主机名修改后会导致该主机上相关实例和上层组件无法正常提供服务，且无法修复。
禁止重装节点 OS	该操作会导致翼 MR 集群进入异常状态，影响翼 MR 集群使用。
禁止使用私有镜像	该操作会导致翼 MR 集群进入异常状态，影响翼 MR 集群使用。

高危操作

以下各表分别描述了各组件在操作与维护阶段，进行日常操作时应注意的高危操作。

集群高危操作

操作名称	操作风险	风险等级	规避措施	重大操作 观察项目
------	------	------	------	--------------

绑定弹性公网 IP	该操作会将集群服务的相关主节点如:HDFS NameNode 所在的 master 节点暴露在公网，会增大来自互联网的网络攻击风险可能性。	★★★★★	请确认绑定的弹性公网 IP 为可信任的公网访问 IP, 确认是否针对开放的相关端口进行设置安全组规则, 只允许可信的 IP 可以访问该端口, 入方向规则不推荐设置允许 0.0.0.0 可以访问。	无
开放集群 22 端口安全组规则	该操作会增大用户利用 22 进行漏洞攻击的风险。	★★★★★	针对开放的 22 端口进行设置安全组规则, 只允许可信的 IP 可以访问该端口, 入方向规则不推荐设置允许 0.0.0.0 可以访问。	无
删除集群或删除集群数据	该操作会导致数据丢失。	★★★★★	删除前请务必再次确认该操作的必要性, 同时要保证数据已完成备份。	无
缩容集群	该操作会导致数据丢失。	★★★★★	缩容前请务必再次确认该操作的必要性, 同时要保证数据已完成备份。	无
卸载磁盘或格式化数据盘	该操作会导致数据丢失。	★★★★★	操作前请务必再次确认该操作的必要性, 同时要保证数据已完成备份。	无

翼 MR Manager 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改日志级别	如果修改为 DEBUG，会导致 Manager 运行速度明显降低。	★★	修改前确认操作的必要性，并及时修改回默认设定。	无
重启下层服务时，如果勾选同时重启上层服务	该操作会导致上层服务业务中断，影响集群的管理维护和业务。	★★★★★	操作前确认操作的必要性，操作时确保同一时间无其它管理维护操作。	观察是否有未恢复的告警产生，观察集群的管理维护是否正常，业务是否正常。
修改 LDAP 端口	修改该参数时，会重启 LdapServer 和 Kerberos 服务和其关联的所有服务，会影响业务运行。	★★★★★	操作前确认操作的必要性，操作时确保同一时间无其它管理维护操作。	无
重启服务	重启过程中会中断服务，如果勾选同时重启上层服务会导致依赖该服务的上层服务中断。	★★★★★	操作前确认重启的必要性。	观察是否有未恢复的告警产生，观察集群的管理维护是否正常，业务是否正常。
修改节点 SSH 默认端口	修改默认端口（22）将导致创建集群、添加服务/实例、添加主机、重装主机等功能无法正常使用，	★★★★★	执行相关操作前将 SSH 端口改回默认值。	无

	并且会导致集群健康检查结果中节点互信等检查项不准确。			
--	----------------------------	--	--	--

Flink 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改日志级别	如果修改为 DEBUG, 会影响任务运行性能。	★★	修改前确认操作的必要性, 并及时修改回默认设定。	无。
修改文件权限	该操作可能导致任务运行失败。	★★★	修改前确认操作的必要性。	观察相关业务操作是否正常。

Flume 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改 Flume 实例的启动参数 GC_OPTS	导致服务启动异常。	★★	修改相关配置项时请严格按照提示描述, 确保修改后的值有效。	观察服务能否正常启动。

HBase 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改加密的相关配置项: hbase.regionserver.wal.ency ption 、 hbase.crypto.keyprovider.par ameters.uri 、	导致服务启动异常。	★★★★★	修改相关配置项时请严格按照提示描述, 加密相关配置项是有关联的, 确保修改后的值有效。	观察服务能否正常启动。

hbase.crypto.keyprovider.parameters.encryptedtext				
已使用加密的情况下关闭或者切换加密算法，关闭主要指修改 hbase.regionserver.wal.encyption 为 false，切换主要指 AES 和 SMS4 的切换	导致服务启动失败，数据丢失。	★★★★★	加密 HFile 和 WAL 内容的时候，如果已经使用一种加密算法加密并且已经建表，请不要随意关闭或者切换加密算法。未建加密表（ENCRYPTION=>AES/SMS4）的情况下可以切换，否则禁止操作。	无。
修改 HBase 实例的启动参数 GC_OPTS、HBASE_HEAPSIZE	导致服务启动异常。	★★	修改相关配置项时请严格按照提示描述，确保修改后的值有效，且 GC_OPTS 与 HBASE_HEAPSIZE 参数值无冲突。	观察服务能否正常启动。
使用 OfflineMetaRepair 工具	导致服务启动异常。	★★★★★	必须在 HBase 下线的情况下才可以使用该命令，而且不能在数据迁移的场景中使用该命令。	观察 HBase 服务是否可以正常启动。

HDFS 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改 HDFS 的 NameNode 的数据存储目录 dfs.namenode.name.dir、DataNode 的数据配置目录 dfs.datanode.data.dir	导致服务启动异常。	★★★★★	修改相关配置项时请严格按照提示描述，确保修改后的值有效。	观察服务能否正常启动。

执行 <code>hadoop distcp</code> 命令时，使用 <code>-delete</code> 参数	Distcp 拷贝时，源集群没有而目的集群存在的文件，会在目的集群删除。	★★	在使用 Distcp 的时候，确保是否保留目的集群多余的文件，谨慎使用 <code>-delete</code> 参数。	Distcp 数据拷贝后，查看目的的数据是否按照参数配置保留或删除。
修改 HDFS 实例的启动参数 <code>GC_OPTS</code> 、 <code>HADOOP_HEAPSIZE</code> 和 <code>GC_PROFILE</code>	导致服务启动异常	★★	修改相关配置项时请严格按照提示描述，确保修改后的值有效，且 <code>GC_OPTS</code> 与 <code>HADOOP_HEAPSIZE</code> 参数值无冲突。	观察服务能否正常启动。
修改 HDFS 的副本数目 <code>dfs.replication</code> ，将默认值由 3 改为 1	1. 存储可靠性下降，磁盘故障时，会发生数据丢失。 2. NameNode 重启失败，HDFS 服务不可用。	★★★★	修改相关配置项时，请仔细查看参数说明。保证数据存储的副本数不低于 2。	观察默认的副本值是否不为 1，HDFS 服务是否可以正常提供服务。
修改 Hadoop 中各模块的 RPC 通道的加密方式 <code>hadoop.rpc.protection</code>	导致服务故障及业务异常。	★★★★★	修改相关配置项时请严格按照提示描述，确保修改后的值有效。	观察 HDFS 及其他依赖 HDFS 的服务能否正常启动，并提供服务。

Hive 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改 Hive 实例的启	修改该参数可能会导	★★	修改相关配置项	观察服务能否正常启

动参数 GC_OPTS	致 Hive 实例无法启动。		时请严格按照提示描述，确保修改后的值有效。	动。
删除 MetaStore 所有实例	Hive 元数据丢失，Hive 无法提供服务。	★★★	除非确定丢弃 Hive 所有表信息，否则不要执行该操作。	观察服务能否正常启动。
使用 HDFS 文件系统接口或者 HBase 接口删除或修改 Hive 表对应的文件	该操作会导致 Hive 业务数据丢失或被篡改。	★★	除非确定丢弃这些数据，或者确保该修改操作符合业务需求，否则不要执行该操作。	观察 Hive 数据是否完整。
使用 HDFS 文件系统接口或者 HBase 接口修改 Hive 表对应的文件或目录访问权限	该操作可能会导致相关业务场景不可用。	★★★	请勿执行该操作。	观察相关业务操作是否正常。
导出表数据覆盖写入本地目录，例如将 t1 表中数据导出，覆盖到“/opt/dir”路径下： insert overwrite local directory '/opt/dir' select * from t1	该操作会删除目标目录，如果设置错误，会导致软件或者操作系统无法启动。	★★★★★	确认需要写入的路径下不要包含任何文件；或者不要使用 overwrite 关键字。	观察目标路径是否有文件丢失。
将不同的数据库、表或分区文件指定至相同路径，例如默认仓库 路 径 “/user/hive/warehouse	执行创建操作后数据可能会紊乱，如果删除其中一个数据库、表或分区，会导致其他对象数据丢失。	★★★★★	请勿执行该操作。	观察目标路径是否有文件丢失。

”				
---	--	--	--	--

Kafka 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
删除 Topic	该操作将会删除已有的主题和数据。	★★★	采用 Kerberos 认证，保证合法用户具有操作权限，并确保主题名称正确。	观察主题是否正常处理。
删除数据目录	该操作将会导致业务信息丢失。	★★★	请勿手动删除数据目录。	观察数据目录是否正常。
修改数据目录下内容（创建文件、文件夹）	该操作将会导致该节点上的 Broker 实例故障。	★★★	请勿手动在数据目录下创建或修改文件及文件夹。	观察数据目录是否正常。
修改磁盘自适应功能“disk.adapter.enable”参数	该操作会在磁盘使用空间达到阈值时调整 Topic 数据保存周期，超出保存周期的历史数据可能被清除。	★★★	若个别 Topic 不能做保存周期调整，将该 Topic 配置在“disk.adapter.topic.blacklist”参数中。	在 KafkaTopic 监控页面观察数据的存储周期。
修改数据目录“log.dirs”配置	该配置不正确将会导致进程故障。	★★★	确保所修改或者添加的数据目录为空目录，且权限正确。	观察数据目录是否正常。
减容 Kafka 集群	该操作将会导致部分 Topic 数据副本数量减少，可能会导致 Topic 无法访问。	★★	请先做好数据副本转移工作，然后再进行减容操作。	观察分区所在备份节点是否都存活，确保数据安全。

单独启停基础组件	该操作将会影响服务的一些基础功能导致业务失败。	★★★	请勿单独启停 ZooKeeper/Kerberos/LDAP 等基础组件启停基础组件请勾选关联服务。	观察服务状态是否正常。
重启/停止服务	该操作将会导致业务中断。	★★	确保在必要时重启/停止服务。	观察服务是否运行正常。
修改配置参数	该操作将需要重启服务使得配置生效。	★★	确保在必要时修改配置。	观察服务是否运行正常。
删除/修改元数据	修改或者删除 ZooKeeper 上 Kafka 的元数据可能导致 Topic 或者 Kafka 服务不可用。	★★★★	请勿删除或者修改 Kafka 在 ZooKeeper 上保存的元数据信息。	观察 Topic 或者 Kafka 服务是否可用。
修改元数据备份文件	修改 Kafka 元数据备份文件, 并被使用进行 Kafka 元数据恢复成功后, 可能导致 Topic 或者 Kafka 服务不可用。	★★★★	请勿修改 Kafka 元数据备份文件。	观察 Topic 或者 Kafka 服务是否可用。

KrbServer 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改 KrbServer 的参数	修改该参数后, 若没有及	★★★★★	修改该参数后, 请	无。

KADMIN_PORT	时重启 KrbServer 服务和其关联的所有服务,会导致集群内部 KrbClient 的配置参数异常,影响业务运行。		重启 KrbServer 服务和其关联的所有服务。	
修改 KrbServer 的参数 kdc_ports	修改该参数后,若没有及时重启 KrbServer 服务和其关联的所有服务,会导致集群内部 KrbClient 的配置参数异常,影响业务运行。	★★★★★	修改该参数后, 请重启 KrbServer 服务和其关联的所有服务。	无。
修改 KrbServer 的参数 KPASSWD_PORT	修改该参数后,若没有及时重启 KrbServer 服务和其关联的所有服务,会导致集群内部 KrbClient 的配置参数异常,影响业务运行。	★★★★★	修改该参数后, 请重启 KrbServer 服务和其关联的所有服务。	无。
修改 Manager 系统域名	若没有及时重启 KrbServer 服务和其关联的所有服务,会导致集群内部 KrbClient 的配置参数异常,影响业务运行。	★★★★★	修改该参数后, 请重启 KrbServer 服务和其关联的所有服务。	无。
配置跨集群互信	该操作会重启 KrbServer 服务和其关联的所有服务,影响集群的管理维护和业务。	★★★★★	更换前确认操作的必要性, 更换时确保同一时间无其它管理维护操作。	观察是否有未恢复的告警产生,观察集群的管理维护是否正常,业务是否正常。

LdapServer 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
修改 LdapServer 的参数 LDAP_SERVER_PORT	修改该参数后，若没有及时重启 LdapServer 服务和其关联的所有服务，会导致集群内部 LdapClient 的配置参数异常，影响业务运行。	★★★★★	修改该参数后，请重启 LdapServer 服务和其关联的所有服务。	无。
恢复 LdapServer 数据	该操作会重启 Manager 和整个集群，影响集群的管理维护和业务。	★★★★★	修改前确认操作的必要性，修改时确保同一时间无其它管理维护操作。	观察是否有未恢复的告警产生，观察集群的管理维护是否正常，业务是否正常。
更换 LdapServer 所在节点	该操作会导致部署在该节点上的服务中断，且当该节点为管理节点时，更换节点会导致重启 OMS 各进程，影响集群的管理维护。	★★★★	更换前确认操作的必要性，更换时确保同一时间无其它管理维护操作。	观察是否有未恢复的告警产生，观察集群的管理维护是否正常，业务是否正常。
修改 LdapServer 密码	修改密码需要重启 LdapServer 和 Kerberos 服务，影响集群的管理维护和业务。	★★★★★	修改前确认操作的必要性，修改时确保同一时间无其它管理维护操作。	无。

节点重启导致 LdapServer 数据损坏	如果未停止 LdapServer 服务，直接重启 LdapServer 所在节点，可能导致 LdapServer 数据损坏。	★★★★★	使用 LdapServer 备份数据进行恢复。	无。
------------------------	--	-------	-------------------------	----

YARN 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
删除或者修改数据目录 yarn.nodemanager. local-dirs 和 yarn.nodemanager. log-dirs	该操作将会导致业务信息丢失。	★★★★	请勿手动删除数据目录。	观察数据目录是否正常。

ZooKeeper 高危操作

操作名称	操作风险	风险等级	规避措施	重大操作观察项目
删除或者修改 ZooKeeper 的数据目录	该操作将会导致业务信息丢失。	★★★★	修改 ZooKeeper 目录时候，严格按照扩容指导操作。	观察服务能否正常启动，关联组件能否正常启动。
修改 ZooKeeper 实例的启动参数 GC_OPTS	导致服务启动异常。	★★	修改相关配置项时请严格按照提示描述，确保修改后的值有效。	观察服务能否正常启动。
设置 ZooKeeper 中 znode 的 ACL 信息	修改 ZooKeeper 中 znode 的权限，可能会导致其他用户无权限访问该 znode，导致系统功能异常。	★★★★★	修改相关配置项时请严格按照“ZooKeeper 权限设置指南”章节操作，确保修改 ACL 信息，不会影响其他组件正常使用 ZooKeeper。	观察项目观察其他依赖 ZooKeeper 的组件能否正常启动，并提供服务。

9. 附录

本章节主要介绍翼 MapReduce 服务的配置修改建议说明。

操作场景

当节点组内 ECS 实例的规格（vCPU 和内存）无法满足您的业务需求时，您可以使用配置升级功能提升 ECS 实例规格。升级配置后需要手动修改 HDFS、YARN 和 Spark 等服务的配置信息。本文为您介绍如何在翼 MR Manager 的“配置管理”页面修改配置项。

前提条件

已创建集群。

操作步骤

在翼 MR Manager 中，单击“运维与配置”。

- (1) 单击“配置管理”。
- (2) 选择“所选集群服务”，点击查询，即可在当前页面修改配置信息。

配置修改建议说明

- (1) Doris: Doris 所在节点配置升级后，Doris FE 建议使用节点一半内存。
- (2) ElasticSearch: 一个 ES 节点，内存建议不超过 64G。ES 所在节点配置升级后，ES 会自动根据节点情况设置内存值，一般无须用户手动修改，但需要重启集群。如果想手动设置，可以修改 jvm.options 文件配置-Xms30g、-Xmx30g 参数，并重启集群。
- (3) HBase: HBase 所在节点配置升级后，修改建议如下：hbase-site.xml:hbase.regionserver.handler.count
 - a. 说明：一般跟 CPU 核数相同 hbase-env.sh:export HBASE_MASTER_OPTS : "-Xmsg -Xmxg"
 - b. 说明：master 不消耗很多内存，一般默认不添加或者分配 2~8G 左右。export HBASE_REGIONSERVER_OPTS: "-Xmsg -Xmxg "
- (4) 说明：regionserver 需要较多内存，一般配置内存配额的一半或更多。HDFS: HDFS 所在节点配置升级后，可以根据 hadoop-env.sh 参数配置进行，通过调整服务的内存大小调整服务的性能，如-Xmx20g -Xms20g -Xmn4g，然后重启服务。Namenode 建议文件、目录、数据块之和 1 亿，配置 50G。
- (5) Hive: Hive 所在节点配置升级后，可以通过 hive-env.sh 统一参数配置来进行，也可以在作业提交时使用额外参数指定来进行。通过调整服务的内存大小调整服务的性能，如-Xmx20g -Xms20g -Xmn4g，然后重启服务。内存大小可以根据机器的总内存而定，建议初始值为总内存大小的 10%，后续根据性能需求调整。
- (6) Kafka: Kafka 所在节点配置升级后，建议配置如下：kafka-env.sh 设置 jvm 配置参数：调整 jvm 堆大小，通过调整参数：export KAFKA_HEAP_OPTS="-Xmx20G -Xms20G -Xmn4g"设置堆大小。
 - a. server.properties 文件建议修改的配置项：num.io.threads: 修改写磁盘的线程数，建议配置

为 CPU 核数的 50%;

- i. num.replica.fetchers: 修改副本拉取线程数, 建议配置为 CPU 核数 50% 的 1/3;
- ii. num.network.threads: 修改数据传输线程数, 建议配置为 CPU 核数的 50% 的 2/3;
- iii. replica.fetch.max.bytes: 副本拉取数据量的大小。内存增加, 可以适当加大该值;
- iv. socket.send.buffer.bytes: 调整 socket 发送的数据量。内存增加, 可以适当加大该值;
- v. socket.receive.buffer.bytes: 调整 socket 接受的数据量。内存增加, 可以适当加大该值;
- vi. socket.request.max.bytes: socket 请求的数据量。内存增加, 可以适当加大该值。

(7) Kerberos: 建议保持默认值, 无需修改配置。

(8) Kibana: kibana 是一个基于 NodeJS 的单页 web 应用, 一般情况下, 对内存 CPU 占用很少, 无须修改内存、CPU 等配置。

(9) Kyubi: Kyubi 一般情况下, 对内存 CPU 占用很少, 无须修改内存、CPU 等配置。

(10) OpenLDAP: 建议保持默认值, 无需修改配置。

(11) Ranger: Ranger 所在节点配置升级后, 修改建议如下: ranger-admin 通过 {installdir}/ews/ranger-admin-services.sh 中变量 ranger_admin_max_heap_size 的值修改 JMX, JAVA_OPTS 修改 Xmx、Xmn 等 JVM 参数, 一般设置 1-8g, 1K policy 建议设置为 1G, 1W policy 建议设置为 8G。

- a. ranger-usersync 通过/{installdir}/ranger-usersync-services.sh 中变量 ranger_usersync_max_heap_size 的值修改 JMX, JAVA_OPTS 修改 Xmx Xmn 等 JVM 参数, 一般设置 1-8g, 1K policy 建议设置为 1G, 1W policy 建议设置为 8G。

(12) Spark: Spark 所在节点配置升级后, 修改建议如下: spark.history.kerberos.principal 和 spark.history.kerberos.keytab 为 spark 读写 eventLog 的租户, 用户如有特殊需求自行更改。

- a. spark.yarn.historyServer.address=: 说明了 history server 的地址, 用户如有特殊需求自行更改。
- b. spark.dynamicAllocation.enabled 和 spark.dynamicAllocation.maxExecutors 分别控制动态和动态开启下能使用的最大资源, 用户如有特殊需求自行更改。
- c. spark.executor.cores 和 spark.executor.memory 确保 spark.executor 的每一个 core 分配到 2~4g 内存, 标准是 4g, 具体视情况而定, 设置 core 的 memory 设置过小 executor 容易 oom。

(13) Trino: Trino 的服务包括 coordinator 和 worker。Trino 所在节点配置升级后, 可以根据 jvm.config 参数配置进行, 通过调整服务的内存大小调整服务的性能, 如-Xmx128g -Xms128g, 然后重启服务。

(14) YARN: YARN 所在节点配置升级后, 可以根据 yarn-env.sh 参数配置进行, 通过调整服务的内存大小来调整服务的性能, 如-Xmx20g -Xms20g -Xmn4g, 然后重启服务。NM 用于集群中作业的内存和 CPU, 需要修改 NM 节点的 yarn-site.xml 中的 yarn.nodemanager.resource.memory-mb 的值, 该值用于所有作业的最多可用内存; 以及 yarn.nodemanager.resource.cpu-vcores 的值, 该值用于所有作业的最多可用虚拟 CPU 核数。

(15) ZooKeeper: ZooKeeper 所在节点配置升级后, 可通过配置 java.env 文件, 在其中添加: export ZK_SERVER_HEAP=2048, (这里设置的单位默认是 MB)。