# 大数据管理之星 Stars

用户使用手册

# 目录

目表	录		1
1.	引言		3
	1.1 编写	目的	3
	1.2 读者》	对象	3
	1.3 定义、	、首字母缩写和缩略语	3
2.	操作说明.		4
	2.1. 首	页	4
	2. 1. 1	功能描述	4
	2. 1. 2	用户界面	4
	2.2. 集和	群管理	5
	2. 2. 1.	功能描述	5
	2. 2. 2.	接入集群	5
	2. 2. 3.	我的集群	6
	2.3. 申请	请审批	9
	2. 3. 1.	创建工单	9
	2. 3. 2.	待办工单	11
	2. 3. 3.	审批提醒	12
	2.4. 应)	用管理	13
	2. 4. 1.	我的应用	13
	2. 4. 2.	应用任务查看	13
	2. 4. 3.	告警查询	14
	2. 5. 监持	控服务	14
	1. 3. 1	告警查询	15
	1. 3. 2	监控图表	15
	2.6. 运约	维管理	15
	2. 6. 1.	cm 管理	16
	2. 6. 2.	任务查看	16

理16	2. 6. 3.	
	2. 6. 4.	
	7. 文档	2. 7
	2.7.1.	
	2.7.2.	
	8. 系统	2.8
	2. 8. 1.	
18	2. 8. 2.	
19	2. 8. 3.	

# 1. 引言

#### 1.1 编写目的

本手册的编写目的是明确本系统的功能、软件的作用、功能的操作,帮助用户理解及操作本系统。

# 1.2 读者对象

本系统涉及的应用人员、管理员和运维人员

## 1.3 缩略语定义

CDH (Cloudera Distribution Including Apache Hadoop): CDH 是 Cloudera 公司推出的基于 Apache Hadoop 生态的商业发行版,集成了 HDFS、MapReduce、Spark 等核心组件,并通过 Cloudera Manager 实现集群的集中管理,为企业提供稳定、安全且易于运维的大数据处理平台。

CDP (Cloudera Data Platform): CDP 是 Cloudera 推出的新一代企业级数据平台,作为 CDH 的继任者,它以统一架构支持多云和混合云环境,整合了数据湖、数据仓库、流处理及机器学习等功能,同时强化了数据治理、安全性和弹性扩展能力,旨在为企业提供从数据采集到分析的全流程解决方案,满足复杂数据场景下的多样化需求。

MRS (MapReduce Service): MRS 是华为云提供的托管式分布式大数据处理服务,基于开源 Hadoop 生态,集成了 Spark、Hive、Flink 等主流组件,支持批处理、流处理等多种计算模式。它通过托管模式降低了集群运维成本,深度适配华为云生态,可与对象存储、数据仓库等服务无缝协同,帮助用户快速构建云端大数据分析平台。

# 2. 操作说明

# 2.1. 首页

# 2.1.1 功能描述

首页汇集了当前登录账号的全览信息,便于快速、直观的获取重要信息。

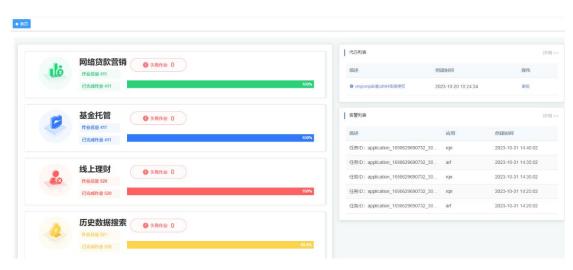
#### 2.1.2 用户界面



系统管理员首页



集群管理员首页



应用管理员首页

## 2.2. 集群管理

#### 2.2.1. 功能描述

用于 CDH、CDP、MRS 集群接入进系统,接入集群的信息展示、集群下属应用的台账信息、集群数据权限控制、集群运行的监控、集群告警信息查询、集群 HDFS 的文件浏览、YARN 任务查看以及每个集群独立的配置中心。

# 2.2.2. 接入集群

目前系统支持 CDH6. 3. 1 和 CDP7. 6. 5 和 MRS3. 3. 0 三个版本的集群。 执行接入操作时,需根据实环境在"API 版本"下拉框里选择对应 API。

* API版本	6.3.2	×	*集群类型	CDH
* cm账号	admin		* cm密码	****
* cm地址	192.168.61.80:7180	0		

接入集群					
* API版本	3.3.0		*集群类型 MRS		
*cm账号	wsd-lwz		* cm 密码		
* cm批址	172.16.67.127:28443	•			
		<b>童</b> 询 重置			
注意:	cm 地址不含 ht	tp://			
各项信	息填写完毕后,	点击【查询】会弹出	集群选	择窗口。	
	<b>∢ cm地址</b> : 1	92.168.61.80		✓ cm账号:	admin
	集群名称		集群url		集群版本
急: 请确保	cdh80 接入的集群已经在AD域中	http://cdh03cm01.wisdom. 中初始化了hdfs_manager和hive_ma 下一步			为该集群的超级用户
勾选本	i接入的集群已经在AD域中 次要接入的集郡	中初始化了hdfs_manager和hive_ma 下一步 羊,点击【下一步】	nager,新接 取消	入的集群将以这两个用户作	为该集群的超级用户
勾选本	i接入的集群已经在AD域中 次要接入的集郡	・ 中初始化了hdfs_manager和hive_ma 下一步	nager,新接 取消	入的集群将以这两个用户作	为该集群的超级用户
勾选本 <b>注意:</b>	i接入的集群已经在AD域中 次要接入的集郡	<del>で初始化了hdfs_manager和hive_maでします。でしままである。できまままでは、「一ちまままままままままままままままままままままままままままままままままままま</del>	nager,新接 取消	入的集群将以这两个用户作	为该集群的超级用户
勾选本 <b>注意:</b> ■	為的集群已经在AD域中 次要接入的集制 请确保 AD 域中	<del>で初始化了hdfs_manager和hive_maでします。でしままである。できまままでは、「一ちまままままままままままままままままままままままままままままままままままま</del>	nager, 新接 取消 ager 和	入的集群将以这两个用户作 hive_manager 账 集群名称: cdh80	为该集群的超级用户
勾选本 <mark>注意:</mark> ■	為的集群已经在AD域中 次要接入的集制 请确保 AD 域中	で初始化了hdfs_manager和hive_ma 下一歩 詳,点击【下一歩】 已经配置好 hdfs_man	nager, 新接 取消 ager 和	入的集群将以这两个用户作 hive_manager 账 集群名称: cdh80	为该集群的超级用户
勾选本 <mark>注意:</mark> ■	接入的集群已经在AD域中 次要接入的集制 请确保 AD 域中 CM地址: 192.168.61.80 集群url: http://cdh03cn	で初始化了hdfs_manager和hive_ma 下一歩 詳,点击【下一歩】 已经配置好 hdfs_man	nager, 新接 取消 ager 和	入的集群将以这两个用户作 hive_manager 账 集群名称: cdh80	为该集群的超级用户
勾选本 <mark>注意:</mark> ■	接入的集群已经在AD域中 次要接入的集制 请确保 AD 域中 CM地址: 192.168.61.80 集群url: http://cdh03cn	下一步 并,点击【下一步】 已经配置好 hdfs_man	nager, 新接 取消 ager 和	入的集群将以这两个用户作 hive_manager 账 集群名称: cdh80	为该集群的超级用户

【点击上传】选择集群使用的 krb5 文件,点击【确认接入】

# 2.2.3. 我的集群

"我的集群"中作为本系统已接入的所有集群的入口。每个集群卡片中的"主机"、

"组件"、"告警"的图标均可点击跳转的对应标签页。集群的名称,可点击 <sup>∠</sup> 修改。点击 卡片中的集群名称可进入相应集群的详情页。

对于新集群接入后,访问集群详情时会自动跳转到配置中心,必须完成配置并保存才能访问集群其他信息。



#### 1) 组件列表

展示集群组件信息,检测健康即组件的可用性检测



#### 2) 集群权限

被添加的用户,在"我的集群"中才能查看到相应集群



#### 3) 监控图表

展示集群的主要监控信息。包含 hdfs 空间使用率、集群 CPU 使用、磁盘 IO、内存使用情况。各图显示顺序可配置,各图显示与否可控制。



#### 4) 告警查询

展示当前集群的告警信息,告警信息的设置位于集群配置中心。支持告警类型、告警创建时间的搜索。



#### 2.3. 申请审批

本功能模块进行工单审批流的配置、申请以及审批。

在正式创建工单之前,应先完善"申请审批"-"审批提醒"中每个工单的提醒对象。

#### 2.3.1. 创建工单

#### 1) 数据平衡

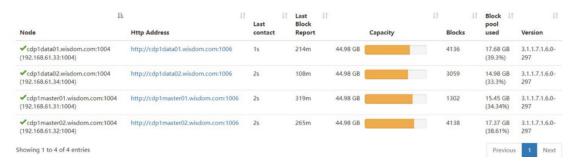
此工单用来解决 HDFS 由于数据节点异常或者其他原因,导致的各数据节点数据目录空间使用率差异过大。



申请时,确定数据失衡集群并,选择好数据平衡可使用的最大带宽和数据平衡停止的阈值,工单审批通过后集群将按照申请参数执行平衡操作。



执行数据平衡前



执行数据平衡后

#### 2) 目录扩容

此工单用于应用的库目录、账户目录配额不足时,申请增加配额。扩容分为临时库容和

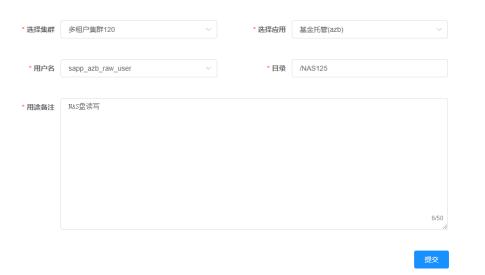
固定扩容。

临时扩容: 扩容的配额具有时效性, 到期后临时配额会被系统收回。

* 选择集群	多租户集群120	* 应用	基金托管(azb)	*选择目录	/user/hive/warehouse/sapp_azb.dl ∨
当前固定配额		当前临时配额		使用情况	17.58GB/3TB+0B
扩容选项	● 临时扩容 □ 固定扩容	* 到期时间	© 2023-12-01 00:00:00	* 申请配额	1 TB v
					提交
固知	E扩容:申请扩容的	的配额是永久	有效的		
*选择集群	多租户集群120	* 应用	基金托管(azb)	*选择目录	/user/hive/warehouse/sapp_azb.dl ∨
当前固定配额		当前临时配额		使用情况	17.58GB/3TB+0B
扩容选项	○ 临时扩容	*申请配额	1 TB ~		

#### 3) 目录授权

此工单用于申请对 HDFS 目录的权限



#### 4) 库表授权

此工单用于跨应用访问或者访问公共应用库没有权限时,申请权限。包含"读权限"、 "读写权限"、"all 权限"。

* 目标集群	多租户集群120 ~	* 目标库	sapp_azb ~	* 待赋权应用	C端消费贷款(skt)	
* 待赋权用户	sapp_skt_raw_user ∨	* 申请权限	读权限			
* 目标表	maker					
					5/51:	2
* 申请用途	申请临时访问azb. maker					
					15/5	0
					#B->-	
					提交	

工单审批通过后,即具备目标表的相应权限。

# 2.3.2. 待办工单

在登录一些被设置了审批权限的账号时,所有待审批的工单都集中在此。



点击工单的"明细"可以了解工单的详细信息



点击"审批",进入审批弹窗,点击同意时可不输入意见,但若不同意则必须输入审批 意见。

P		● 拒绝时审批意见必填	
	审批		×
比状	请输	入审批意见	
批		同意	i绝

## 2.3.3. 审批提醒

在"审批提醒"中点击提醒方案的"配置",进入明细标签页。

注: 方案编号、实体不支持自定义。



新增一条提醒方案明细。

单据状态,可自定义的状态值

状态名称: 为我的工单中显示的"状态"代表着工单审批执行到了哪一步。

流程顺序: 用来在多条明细中,调整明细调目的顺序,按 0,1,2,3,4,5···的顺序由上向下排序。

方案类型:运维



确定后点击"提醒对象",即可设置此条明细触发时将通知哪些角色和用户。



提醒与角色集: 多个角色集之间执行"与"运算,如选择了"集群管理员"、"系统管理员",则只有同时具备以上两个角色的用户才会收到工单处理通知。

提醒或角色集: 多个角色集之间执行"或"运算,如选择了"集群管理员"、"系统管理员",则只要具备其中一个角色的用户就会收到工单处理通知。

提醒用户集:被选中的用户才会收到工单处理通知。

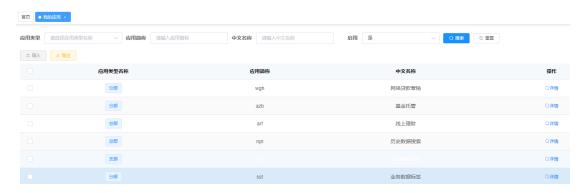
排除用户集:被选中的用户不会收到工单处理通知。

注: 最终通知对象是以上每个提醒集的结果取并集。

## 2.4. 应用管理

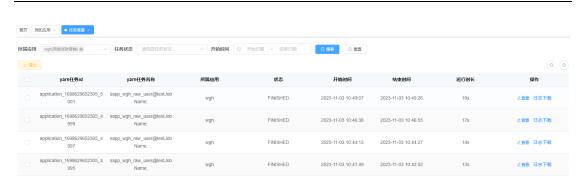
#### 2.4.1. 我的应用

"我的应用"用于展示当前账号有权限的相应应用信息。



### 2.4.2. 应用任务查看

仅可查看当前账号管理的应用下辖的任务,可根据应用、任务状态、任务的运行开始时间进行检索。



"操作"栏的"查看"功能,用于查看任务详细信息,当任务是 FINISHED 状态时,内存分配和 CPU 分配显示为-1,表示系统不再更新任务信息。

当任务在运行时,还可通过"查杀"功能直接杀掉相应任务。



#### 2.4.3. 告警查询

用于应用管理员,查询有数据权限的应用作业告警

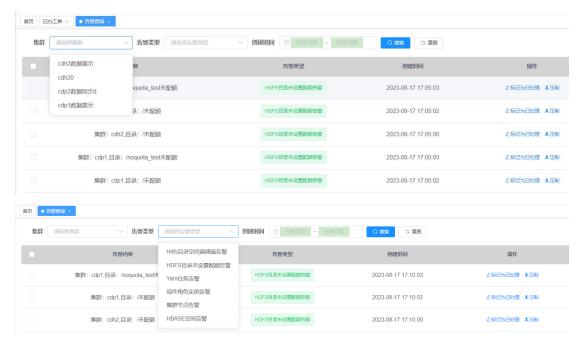


# 2.5. 监控服务

监控服务主要向运维人员开放,所展示的数据类型与集群详情中类似。

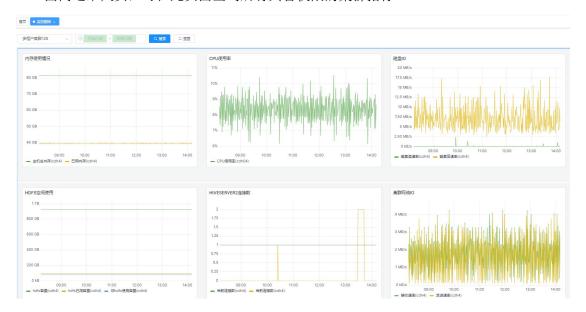
## 1.3.1 告警查询

面向运维人员,汇集了系统下所有集群的告警信息,可通过集群和告警类型进行检索。



## 1.3.2 监控图表

面向运维人员, 可在此页面查询所有具备权限的集群指标



## 2.6. 运维管理

主要向运维人员开放,集中列出运维管理中关注的数据。

# 2.6.1. cm 管理

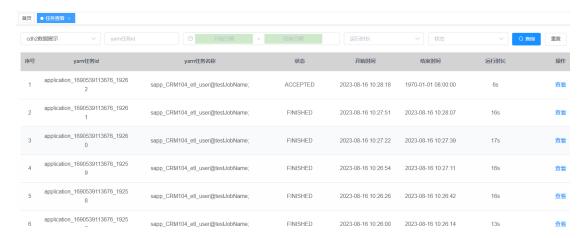
展示出系统下已接入集群的 cm 信息,实时更新 cm 的状态。

注: 删除 CM 意味着取消对应集群的接入。



### 2.6.2. 任务查看

有别于"应用管理"中的任务查看,此处汇集系统下辖的所有集群的任务信息。 同样具备任务的搜索和查杀。



# 2.6.3. keytab 管理

keytab 快速生成小工具,用于一些免交互环境的认证使用。



## 2.6.4. 故障记录

用于常见运维故障处理经验分享,可上传文档分享处理过程中的图片和详细过程



#### 2.7. 文档管理

用于共享文档,针对仅面向应用或者运维人员的文档,可在文档管理中分开发布,并通过账号角色隔离。

## 2.7.1. 应用文档

支持 Excel 表格、PPT 文档、Word 文档、TXT 文档、PDF 文档。

允许 5M 以内文件上传、下载、预览。



## 2.7.2. 运维文档

支持 Excel 表格、PPT 文档、Word 文档、TXT 文档、PDF 文档。

允许 5M 以内文件上传、下载、预览。



# 2.8. 系统管理

# 2.8.1. 审计管理

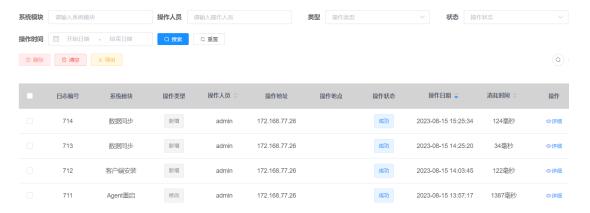
记录系统内所有账号对集群的操作。



# 2.8.2. 日志管理

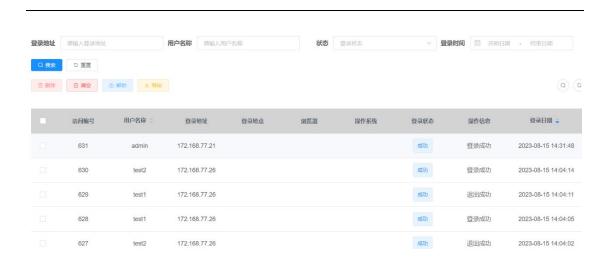
#### 1) 操作日志

记录了用户的操作详情



#### 2) 登录日志

记录用户的登录信息



# 2.8.3. 在线用户

#### 列出当前系统所有的在线用户。

