



Databricks 数据洞察 管理集群

文档版本: 20220519



法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。 如果您阅读或使用本文档,您的阅读或使用行为将被视为对本声明全部内容的认可。

- 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档,且仅能用 于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息,您应当严格 遵守保密义务;未经阿里云事先书面同意,您不得向任何第三方披露本手册内容或 提供给任何第三方使用。
- 未经阿里云事先书面许可,任何单位、公司或个人不得擅自摘抄、翻译、复制本文 档内容的部分或全部,不得以任何方式或途径进行传播和宣传。
- 由于产品版本升级、调整或其他原因,本文档内容有可能变更。阿里云保留在没有 任何通知或者提示下对本文档的内容进行修改的权利,并在阿里云授权通道中不时 发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠 道下载、获取最新版的用户文档。
- 4. 本文档仅作为用户使用阿里云产品及服务的参考性指引,阿里云以产品及服务的"现状"、"有缺陷"和"当前功能"的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引,但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的,阿里云不承担任何法律责任。在任何情况下,阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害,包括用户使用或信赖本文档而遭受的利润损失,承担责任(即使阿里云已被告知该等损失的可能性)。
- 5. 阿里云网站上所有内容,包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计,均由阿里云和/或其关联公司依法拥有其知识产权,包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意,任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外,未经阿里云事先书面同意,任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称(包括但不限于单独为或以组合形式包含"阿里云"、"Aliyun"、"万网"等阿里云和/或其关联公司品牌,上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司)。
- 6. 如若发现本文档存在任何错误,请与阿里云取得直接联系。

通用约定

格式	说明	样例
⚠ 危险	该类警示信息将导致系统重大变更甚至故 障,或者导致人身伤害等结果。	⚠ 危险 重置操作将丢失用户配置数据。
▲ 警告	该类警示信息可能会导致系统重大变更甚 至故障,或者导致人身伤害等结果。	警告 重启操作将导致业务中断,恢复业务 时间约十分钟。
〔〕) 注意	用于警示信息、补充说明等,是用户必须 了解的内容。	大意 权重设置为0,该服务器不会再接受新 请求。
? 说明	用于补充说明、最佳实践、窍门等 <i>,</i> 不是 用户必须了解的内容。	⑦ 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击 设置> 网络> 设置网络类型 。
粗体	表示按键、菜单、页面名称等UI元素。	在 结果确认 页面,单击 确定 。
Courier字体	命令或代码。	执行 cd /d C:/window 命令,进入 Windows系统文件夹。
斜体	表示参数、变量。	bae log listinstanceid
[] 或者 [alb]	表示可选项,至多选择一个。	ipconfig [-all -t]
{} 或者 {a b}	表示必选项,至多选择一个。	switch {active stand}

目录

1.集群配置	05
1.1. 创建集群	05
1.2. 查看集群列表信息	07
1.3. 访问Web UI	08
1.4. 设置安全组白名单	12
1.5. 访问外部数据源	14
1.6. 配置弹性伸缩	15
1.6.1. 弹性伸缩概述	15
1.6.2. 开启并配置弹性伸缩	15
1.6.3. 按时间伸缩规则配置	16
1.6.4. 按负载伸缩规则配置	17
1.6.5. 关闭弹性伸缩	20
1.6.6. 查看弹性伸缩记录	21
1.7. 用户管理	21
1.8. 集群扩容	23
1.9. 释放集群	24
2.RAM访问控制	27
2.1. 角色授权	27
2.2. 服务关联角色	28
2.3. 为RAM用户授权	30
3.运维监控	33
3.1. 集群运维与监控	33
3.2. 项目空间运维与监控	38
3.3. Spark作业原生运维与监控	42

1.集群配置

1.1. 创建集群

本节介绍如何使用Databricks数据洞察控制台创建集群。

前提条件

已注册阿里云账号,并完成实名认证。详情请参见阿里云账号注册流程。

操作步骤

- 1. 使用阿里云账号登录Databricks数据洞察控制台。
- 在Dat abricks数据洞察控制台页面,选择所在的地域(Region)。
 创建的集群将会在对应的地域内,一旦创建后不能修改。
- 3. 在左侧导航栏中,单击集群。
- 4. 在集群管理页面,单击创建集群。
- 5. 设置基础信息。

参数	描述
集群名称	集群的名字。长度限制为1~64个字符,仅可使用中 文、字母、数字、连接号(-)和下划线(_)
Knox账号	为了更好的安全性,Web UI访问(如Zeppelin Notebook、Spark UI、Ganglia UI)需要Knox账号和 密码,来保障您的账号安全。若无RAM子账号,请前 往RAM控制台进行创建
Knox密码	两次确认Knox密码,登录Web Ul时候使用,请您牢记
Databricks Runtime版本	Databricks Runtime的版本信息,版本号与 Databricks官方保持一致,包含Scala和Spark的版 本。版本详情请参见 <mark>Databricks Runtime版本说明</mark>
Python版本	默认版本为Python 3
付费类型	Dat abricks数据洞察支持包年包月和按量付费两种付费 类型

参数	描述
可用区	可用区为在同一地域下的不同物理区域,可用区之间 内网互通。 一般选择默认的可用区即可,亦可选择与已购阿里云 产品部署在同一个可用区。
ECS实例	由Master和Worker两种类型的节点组成: Master节点:主要负责集群资源管理和作业调度。 默认节点个数为1。 Worker节点:集群的计算节点,主要负责作业的执行。最小节点数量为3。
元数据选择	推荐选择数据湖元数据。 数据湖元数据:表示元数据存储在数据湖中。如果没有开通数据湖构建服务,需要单击请前往开通。参见数据湖元数据管理 统一meta数据库:表示元数据存储在集群本地环境的MySQL数据库中 独立RDS MySQL:表示使用自建的阿里云RDS作为元数据库,更多信息请参见 共享独立RDS元数据库

6. 设置高级信息。

高级信息包括如下两方面:

○ Spark设置

参数	描述
Spark配置	输入Spark的配置信息。配置的属性值将会更改到 <i>spark-defaults.conf</i> 文件中。支持的配置列表为 <i>spark.apache.org/docs/latest/configuration.html#spark-properties</i>
环境变量	您可以自定义Spark执行的环境变量。配置的属性将 会更新到 <i>spark-env.sh</i> 中。

∘ 服务目录

参数	描述
类型	包括以下两种类型: ■ 默认值 ■ 自定义
OSS路径	该目录用来存放集群服务组件的临时文件等。 该目录会作为产品的根目录来使用。当用户有多个集 群时,不需要为每个集群单独指定服务目录。不同 Region需要有不同的服务目录,产品会为每个集群在 服务目录下创建子目录,即 oss://\${specified-buck et-or-dir}/ddi-\${clusterid}/。

- 7. 阅读并勾选服务条款。
- 8. 单击创建。

集群创建需要时间,当状态更新为空闲时表示创建成功,请您耐心等待。

问题反馈

您在使用阿里云Databricks数据洞察过程中有任何疑问,欢迎用钉钉扫描下面的二维码加入钉钉群进行反馈。

Databricks数据洞察产品... 💷 🥪



该群属于"大数雄+AI"部门群,仅组织内部成员可以加入,如果组织 外部人员收到此分享。需要先申请加入该组织。

1.2. 查看集群列表信息

本文介绍如何查看已创建集群的详情。

前提条件

已创建集群,详情请参见创建集群。

操作步骤

1.

2.

3. 集群管理页面展示您所拥有的所有集群的基本信息,以及各集群支持的操作。

参数	说明
集群ID/名称	集群ID 是产品自动分配的集群唯一标识; 名称 是用户 在集群创建时自定义的集群名称。
集群类型	集群的付费类型。
状态	 集群的状态: 初始化中:集群正在构建,包括两个阶段:一是物理ECS机器的创建;二是集群服务的启动,稍等片刻即可达到运行中的状态。 空闲:集群目前没有作业运行。 运行中:集群处于正常运行状态。 构建失败:创建过程中遇到异常,已经创建的ECS机器会自动回滚,在集群列表页面单击状态右边的问号,可以查看异常明细。 终止中:目前集群处于终止状态。 终止中:目前集群处于终止状态。 终止失败:终止集群时失败。 已终止:集群已终止。已终止的集群无法运行笔记本或作业。 异常:表示集群异常。
创建时间/运行时间	集群创建的时间以及运行的时长。
付费类型	集群的付费类型。
操作	 支持的集群操作: 详情:进入集群的详情页,查看集群创建后的详细 信息。 展示已创建集群的详细信息,包括集群信息、网络 信息、软件信息和主机信息四部分。 Spark UI: Apache Spark history server提供的 Web UI。您可以在此界面查看Spark作业的运行信 息。 Ganglia监控:用来监控集群内节点的运行状况。 Notebook:进入集群对应的DataInsight Notebook页面,Notebook相关操作请参见管理 Notebook。 释放:释放当前集群,详情请参见释放集群。

1.3. 访问Web UI

Databricks数据洞察集群提供了多个Web UI的访问入口,包括Notebook、Spark UI、Yarn UI和Ganglia监控。

增加安全组白名单

为了用户的数据安全,DD提供了安全组白名单机制来访问集群Web UI。默认不添加白名单是不能进行访问的,具体添加入口是**集群基础信息->集群操作->安全组白名单。**

白名单的具体配置方法看文档-安全组白名单。

使用概述

增加安全组白名单之后,用户在集群详情页面单击Web UI的链接,会跳转到Knox账号的验证页面。输入 Knox账号和密码即可登录到相应的Web UI页面。



Web UI登录

首次登录Web UI时,用户可能会在浏览器看到如下告警。用户可以根据浏览器告警的提示,进行操作。通常 情况下,用户会看到下面两种告警。

告警提示一



您的连接不是私密连接

NET::ERR_CERT_AUTHORITY_INVALID

□ 将<u>您访问的部分网页的网址、有限的系统信息以及部分网页内容</u>发送给 Google,以帮助我们提升 Chrome 的安全性。<u>隐私权政策</u>

高级

返回安全连接

单击"高级"按钮,展开隐藏详情后,会出现急需前往链接。单击该链接即可访问Web Ul。



您的连接不是私密连接

NET::ERR_CERT_AUTHORITY_INVALID

□ 将<u>您访问的部分网页的网址、有限的系统信息以及部分网页内容</u>发送给 Google,以帮助我们提升 Chrome 的安全性。<u>隐私权政策</u>







告警提示二

出现"您的连接不是私密连接"的告警提示,且点开高级按钮,没有继续访问链接时,请在当前页面直接键 盘盲输入11个字符: thisisunsafe



您的连接不是私密连接

攻击者可能会试图从 knox.c http://www.characteria.cnhangzhou.databricks.aliyuncs.com 窃取您的信息(例如:密码、通讯内容或信用卡信 息)。<u>了解详情</u>

NET::ERR_CERT_INVALID

□ 将您访问的部分网页的网址、有限的系统信息以及部分网页内容发送给 Google,以帮助我们提升 Chrome 的安全性。<u>隐私权政策</u>

隐藏详情

重新加载

knox.c n-hangzhou.databricks.aliyuncs.com 通常会使用加密技术 来保护您的信息。Google Chrome 此次尝试连接到 knox.c not hangzhou.databricks.aliyuncs.com 时,此网站发回了异常的错误凭据。这可能是因为有 攻击者在试图冒充 knox con-hangzhou.databricks.aliyuncs.com, 或 Wi-Fi 登录屏幕中断了此次连接。请放心,您的信息仍然是安全的,因为 Google Chrome 尚未进行任何数据交换便停止了连接。

您目前无法访问 knox.c-为此网站发送了 Google Chrome 无法处理的杂乱凭据。网络错误和攻击通常是暂时的,因此,此网页稍后可能会恢复正常。

1.4. 设置安全组白名单

背景

安全组白名单功能是通过设置安全组规则,控制外部授权对象(指定IP地址)访问集群knox和以及Spark Thrift Server。

○ 注意

目前产品开放了8443(knox)、10001(Spark Thrift Server)2个端口,需要更多端口请提交工单。

安全组白名单规则

建立数据通信前,安全组匹配安全组规则查询是否放行访问请求,一条安全组规则由规则方向、授权策略、 协议类型、端口范围、授权对象等属性确定,具体详见下表。

属性	说明
规则方向	仅支持网络连接的 入方向
授权策略	仅支持 允许策略
协议类型	仅支持 T CP协议
端口范围	8443(knox)、10001(Spark Thrift Server)
优先级	无
授权对象	IP地址

操作步骤

- 1. 登录阿里云Databricks 数据洞察控制台。
- 2. 在左侧导航栏的菜单下, 单击集群管理。
- 3. 选择想要设置安全组白名单的集群,单击集群ID进入集群详情页面。
- 4. 在集群基础信息页面中单击集群管理按钮下的安全组白名单按钮。

■ (-)阿里云	资源组	▼ 华北2(北京) ▼					费用 工单	备案 服务	简体中文 🖂 🗘 📜 4 💽
Databricks DataInsight	Databri	icks 数据洞察 / 集構	詳管理 / 集群详情						
首页	ا		- 方向)						 Notebook
集群管理		KEMI11+ ()()	(10)						`
Notebook		添加白名单	⑥ 您可以通过白名单功能,	只允许被设置的网络IP可以访问您I	的集群。当前支持 <mark>kr</mark>	nox和Spark Thrift Server	r 2种端口范围限制 查	看详情	
项目空间		授权策略	授权对象	端口范围	描述	*		操作	
元数据	斜	允许	请输入授权IP	8443/8443 (knox)	~	请输入描述		保存 取消	jzhou
库	17	允许	192.168.0.0/16	8443/8443				删除	20年2月24日 10:46:02
帮助	用	允许	192.168.0.0/16	10001/10001				删除	
	ŧ	允许	192.168.0.0/16	10001/10001				删除	
· · · · · · · · · · · · · · · · · · ·	< c	允许	192.168.0.0/16	10001/10001				删除	
	<u>\$</u> .	允许	192.168.0.0/16	10001/10001				删除	
	ŧ	允许	192.168.0.0/16	10001/10001				删除	
	N	允许	192.168.0.0/16	10001/10001				删除	
	•	允许	192.168.0.0/16	10001/10001				删除	80GB SSD云盘*4
	*								80GB SSD云盘*4
								确认 取消	

- 5. 单击添加白名单。
- 6. 输入授权对象、选择端口范围,单击保存。
- 7. 单击确认按钮,完成。

1.5. 访问外部数据源

本文介绍如何在Databricks数据洞察实现访问外部数据源的需求。

背景信息

Databricks数据洞察为了满足您在计算任务里访问您在阿里云上已有的数据,支持通过添加外部数据源的方式,打通集群到您自己VPC下的网络环境。

绑定数据源

绑定数据源的本质是打通不同VPC之间的网络,即将数据源集群所在VPC与目标Databricks数据洞察集群所在 VPC的网络打通。数据源绑定之后,您可以在Notebook或Spark作业里直接访问对应数据源的数据。

⑦ 说明 对于数据源绑定场景,如果多个数据源共用一个VPC下的交换机,打通其中一个数据源意味 着相同交换机下的所有数据源一并打通。因此,只能打通同一Region下的数据源。

1. 在Databricks数据洞察控制台,进入集群详情页面。

- 2. 点击详情页面数据源标签, 在添加数据源弹窗选择要添加的数据源类型。
- 3. 在所选类型的数据源列表里勾选希望绑定的数据源类型,如果是EMR集群或ECS实例可以直接选择对应 类型;更通用的方式请选择通用网络,自助选择要打通的网络和安全组。
- 4. 建议补充数据源描述信息,便于辨识已绑定数据源实例。
- 5. 点击下一步,确认安全组和交换机信息。

② 说明 对于Aliyun EMR HDFS和Aliyun EMR Kaf ka类型数据源,目前支持各自添加一个集群。Aliyun ECS类型可以多选,如果是自建集群(如Kaf ka或HDFS),只需要选择集群中的一个实例即可。

数据源访问说明

对于Aliyun EMR HDFS集群,数据源打通之后您可以通过以下方式访问集群数据。

对于HA集群,默认使用emr-cluster作为hostname。

sc.textFile("hdfs://emr-cluster/tmp/user0/airline_statistic_usa.csv").count()

对于非HA集群,请直接使用EMR HDFS集群namenode的ip访问。

sc.textFile("hdfs://192.168.xxx.xxx:9000/tmp/user0/airline_statistic_usa.csv").count()

② 说明 对于Aliyun EMR Kafka集群,支持通过ip或者hostename访问。

解绑数据源

解绑数据源本质是将数据源所在VPC与目标Databricks数据洞察集群VPC网络隔离。如果多个数据源共用一 个交换机, 解绑操作会使得当前Databricks数据洞察集群无法继续访问该交换机下所有数据源集群。

- 1. 在Databricks数据洞察控制台,进入集群详情页面。
- 2. 点击详情页面数据源标签。

3. 在已绑定数据源列表里选择要解绑的交换机,点击解绑即可。

1.6. 配置弹性伸缩

本章节介绍Databricks数据洞察弹性伸缩功能,您可以根据业务需求和策略设置伸缩策略。弹性伸缩开启并 配置完成后,当业务需求增长时Databricks数据洞察会自动为您增加Task节点以保证计算能力,当业务需求 下降时Databricks数据洞察会自动减少Task节点以节约成本。

1.6.1. 弹性伸缩概述

本文介绍Databricks数据洞察的弹性伸缩功能,您可以根据业务需求和策略设置伸缩策略。弹性伸缩开启并 配置完成后,当业务需求增长时Databricks数据洞察会自动为您增加Task节点以保证计算能力,当业务需求 下降时Databricks数据洞察会自动减少Task节点以节约成本。

应用场景

在以下场景中,开启Dat abricks数据洞察的弹性伸缩功能,可帮助您节省成本,提高执行效率。

- 临时需要按照时间段添加Task节点,补充计算能力。
- 为确保重要作业按时完成,需要按照某些集群指标扩充Task节点。

功能介绍

Dat abricks数据洞察支持按时间伸缩和按负载伸缩两种伸缩策略,但使用时两者只能二选一。如果切换伸缩策略,原伸缩规则会保留,但处于失效状态,不会被触发执行;当前已扩容的节点也会保留,除非缩容规则触发,否则不会被缩容。

Databricks数据洞察弹性伸缩目前只支持按量付费一种实例。

1.6.2. 开启并配置弹性伸缩

当您的业务量需求不断波动时,建议您开启弹性伸缩功能并配置相应的伸缩规则,以使Databricks数据洞察可按业务量波动来增加和减少Task节点。

操作步骤

- 1. 登录阿里云Databricks控制台。
- 2. 在顶部菜单栏处,根据实际情况选择地域(Region)和资源组。
- 3. 单击上方的集群管理页签。
- 4. 在集群管理页面,单击相应集群所在行的详情。
- 5. 在详情页上方, 单击弹性伸缩。
- 6. 单击新建弹性伸缩机器组-->填写机器组名称-->点击创建机器组。
- 7. 在弹性伸缩配置页面, 配置伸缩策略。

关键配置项说明如下:

配置项	说明
规定时间伸缩	按时间触发计算资源的扩容和缩容。根据业务需求,如果 能够明确需要按照一定的时间周期来伸缩计算资源,建议 您选择此项。详细配置说明请参见 <mark>按时间伸缩</mark> 。
规定负载伸缩	按集群负载触发计算资源的扩容和缩容。根据业务需求, 如果需要对计算资源进行伸缩的时间段较多,但需要根据 YARN的特定指标伸缩计算资源,建议您选择此项。详细 配置说明请参见 <mark>按负载伸缩</mark> 。

8. 完成伸缩策略配置后,单击下方的保存。

9. 单击开启弹性伸缩。

弹性伸缩配置完成后,后续触发时,您可在弹性伸缩记录中查看弹性伸缩的历史执行记录以及每次执行的详 细结果,详情请参见<u>查看弹性伸缩记录</u>。

1.6.3. 按时间伸缩规则配置

如果集群计算量在一定的周期内存在明显的波峰和波谷,则您可以设置在每天、每周或每月的固定时间段扩展一定量的Task节点来补充计算能力,这样在保证作业完成的同时,也可以节省成本。

配置伸缩实例数

- 最大节点数:弹性伸缩的Task节点上限。一旦达到上限,即使满足弹性伸缩的规则,也不会继续进行弹性 伸缩的动作。目前可设置的弹性伸缩最大上限为500。
- 最小节点数:弹性伸缩的Task节点下限。如果弹性伸缩规则中设置的增加或减少Task节点数小于此处的最小节点数,那么在首次执行时,集群会以最小节点数为准进行伸缩。

例如,设置弹性扩容规则为每天零点动态添加1个节点,但最小节点数为3。那么系统在第一天的零点时会添加3个节点,以满足最小节点数的要求。

配置伸缩规则

在Dat abricks数据洞察中开启弹性伸缩时,如果选择按时间配置伸缩规则,则根据以下说明配置相关参数即可

伸缩规则分为扩容规则和缩容规则。集群关闭弹性伸缩功能后,所有规则会被清空,再次开启弹性伸缩功能时,需要重新配置伸缩规则。

添	加弹性伸缩规则 - 接	时间扩	容		×
实例					
ecs.	* 规则名称:				
_		规则不	可以重得		
配置:		() 重复	执行(只执行一次	
大小:		每天	\sim		
		每 1		天执行一次	
配置:		弹性伸缩	的时间规	则间隔: 10 minutes	
大小:	*执行时间:	2021	-03-02 1	:17 💼	
释实例1(* 规则有效期:	2021	-03-02 1	:17 🛗	
	* 重试过期时间(秒):	0	重	战过期时间范围是 0-21600秒	
	* 扩容(台):	1	增	uTask节点数范围是 1-10 台	
	* 冷却时间(秒):	0	冷	即时间范围是 0-86400秒	
					确定 取消

- 规则名称:在同一个集群中,伸缩规则名称(包括扩容规则和缩容规则)不允许重复。
- 规则执行周期:
 - 只执行一次:集群在指定的时间点执行一次弹性伸缩动作。
 - 重复执行:您可以选择每天、每周或每月的某一特定时间点执行一次弹性伸缩动作。
- 重试过期时间:弹性伸缩在到达指定时间时可能由于各种原因不能执行,通过设置重试过期时间,系统会 在该时间范围内每隔30秒尝试执行一次,直到在满足条件时执行伸缩。设置范围为0~21600秒。

假设在指定时间段需要进行弹性伸缩动作A,如果有其他弹性伸缩动作B正在执行或正处在冷却期,则动作 A无法执行。在您设置的重试过期时间内,每隔30秒会重试一次,尝试执行A,一旦条件满足,集群会立 刻执行弹性伸缩。

- 扩容或减少Task节点数:规则被触发时,集群每次增加或减少Task节点数量。
- 冷却时间:每次弹性伸缩动作执行完成,到可以再次进行弹性伸缩的时间间隔。在冷却时间内,不会发生 弹性伸缩动作。

配置伸缩规格

弹性伸缩配置可以指定伸缩的节点的硬件规格。您只能在开启弹性伸缩功能时配置,保存后不能更改。如果 特殊情况确实需要修改,可以关闭弹性伸缩功能后,再次开启。

- 选择vCPU和内存规格时,系统会根据您的选择自动匹配出满足条件的实例,并显示在备选实例列表中。您 需要选中备选的实例,以便集群按照已选的实例规格进行伸缩。
- 为避免由于ECS库存不足造成的弹性伸缩失败,您最多可以选择3种ECS实例。
- 无论是选择高效云盘还是SSD云盘,数据盘最小设置为40GB。

1.6.4. 按负载伸缩规则配置

在使用Dat abricks数据洞察集群时,如果您无法准确的预估大数据计算的波峰和波谷,则可以使用按负载伸 缩配置的策略。在Dat abricks数据洞察中开启弹性伸缩时,如果选择按负载配置伸缩规则,则根据以下说明 配置相关参数即可。详细配置步骤请参见开启并配置弹性伸缩。

配置伸缩实例数

- 最大实例数:弹性伸缩的Task节点上限。一旦达到上限,即使满足弹性伸缩的规则,也不会继续进行弹性 伸缩的动作。目前可设置的弹性伸缩最大上限为500。
- 最小实例数: 弹性伸缩的Task节点下限。
 - 如果弹性伸缩规则中设置的增加后的Task节点总数小于最小节点数,那么在首次执行时,集群会以最小 节点数为准,伸缩至最小节点数。例如,当前Task节点数为0,设置弹性扩容规则为每天零点动态添加1 个节点,但最小节点数为3。那么系统在第一天的零点时会添加至3个节点,此后每天零点动态添加1个 节点,以满足最小节点数的要求。
 - 如果弹性伸缩规则中设置减少后的Task节点总数小于最小节点数,那么集群不会进行弹性伸缩操作。

配置负载伸缩规则

伸缩规则分为**扩容规则和缩容规则**。集群关闭弹性伸缩功能后,所有规则会被清空,再次开启弹性伸缩功 能时,需要重新配置伸缩规则。切换伸缩策略时(例如从按负载伸缩切换到按时间伸缩),原策略下的伸缩 规则处于失效状态,不会被触发,但已经扩容的节点会继续保留,不会被释放。

	act.1	· * + + + + + + + + + + + + + + + + + +	* 抑则乞称:
	观则不可以重名		ייניישרגאנאי
ng virtual cores	ARN.PendingVCores 🗸	标: Y	* 集群负载指标:
	~ 分钟	期: 1	* 统计周期:
	3均値 ~ >= ~]则: 平	* 统计规则:
	~	容: 1	* 重复几次后扩容:
	□ 增加Task节点数范	台): 1	* 扩容(台):
	□冷却时间范围是(秒): 0	* 冷却时间(秒):

定 取消

- 规则名称:在同一个集群中,伸缩规则名称(包括扩容规则和缩容规则)不允许重复。
- 集群负载指标: 在YARN的负载指标中获取,具体可以参见Hadoop官方文档。Databricks数据洞察弹性伸 缩指标与YARN负载指标的对应关系如下:

E-MapReduce弹性伸缩指标	YARN Metrics	说明
YARN.AvailableVCores	availableVirtualCores	The number of available virtual cores

E-MapReduce弹性伸缩指标	YARN Metrics	说明
YARN.PendingVCores	pendingVirtualCores	The number of pending virtual cores , Databricks数据洞察补充采 集
YARN.AllocatedVCores	allocatedVirtualCores	The number of allocated virtual cores
YARN.ReservedVCores	reservedVirtualCores	The number of reserved virtual cores
YARN.AvailableMemory	availableMB	The amount of memory available in MB
YARN.PendingMemory	pendingMB	The amount of memory pending in MB,Databricks数据洞察补充采 集
YARN.AllocatedMemory	allocatedMB	The amount of memory allocated in MB
YARN.ReservedMemory	reservedMB	The amount of memory reserved in MB
YARN.AppsRunning	appsRunning	The number of applications running
YARN.AppsPending	appsPending	The number of applications pending
YARN.AppsKilled	appsKilled	The number of applications killed
YARN.AppsFailed	appsFailed	The number of applications failed
YARN.AppsCompleted	appsCompleted	The number of applications completed
YARN.AppsSubmitted	appsSubmitted	The number of applications submitted

E-MapReduce弹性伸缩指标	YARN Metrics	说明
YARN.AllocatedContainers	containersAllocated	The number of containers allocated
YARN.PendingContainers	containersPending	The number of containers pending
YARN.ReservedContainers	containersReserved	The number of containers reserved

- 统计周期与统计规则: 您选定的集群负载指标在一个统计周期内, 按照选定的聚合维度(平均值、最大值、最小值), 达到触发阈值为一次触发。
- 重复几次后扩容或缩容: 负载指标聚合后达到阈值触发的次数, 达到该次数后触发集群弹性伸缩的动作。
- 扩容或缩容(台):规则被触发时,集群每次执行增加或减少的Task节点数量。
- 冷却时间(秒):每次弹性伸缩动作执行完成,到可以再次进行弹性伸缩的时间间隔。在冷却时间内,即 使满足弹性伸缩条件也不会发生弹性伸缩动作。即忽略本次在冷却时间内触发的弹性伸缩动作,直到下一 次满足伸缩条件且不在冷却时间内再执行。

配置伸缩规格

弹性伸缩配置可以指定伸缩的节点的硬件规格。您只能在开启弹性伸缩功能时配置,保存后不能更改。如果 特殊情况确实需要修改,可以关闭弹性伸缩功能后,再次开启。

- 选择vCPU和内存规格时,系统会根据您的选择自动匹配出满足条件的实例,并显示在备选实例列表中。您 需要选中备选的实例,以便集群按照已选的实例规格进行伸缩。
- 为避免由于ECS库存不足造成的弹性伸缩失败,您最多可以选择3种ECS实例。
- 无论是选择高效云盘还是SSD云盘,数据盘最小设置为40GB。

1.6.5. 关闭弹性伸缩

开启弹性伸缩功能后,后续如果您的业务量需求趋于稳定,则您可关闭弹性伸缩功能。背景信息

背景信息

关闭弹性伸缩后,当前已经通过弹性伸缩功能扩充的Task节点将会全部释放,但HDFS存储的位于Core节点的数据不会受影响。如果关闭弹性伸缩后,集群计算资源不足,则建议您通过扩容来增加集群计算资源,详 情请参见集群扩容。

操作步骤

- 1. 登录阿里云Databricks控制台。
- 2. 在顶部菜单栏处,选择地域(Region)。
- 3. 单击左侧集群页签。
- 4. 在集群管理页面,单击相应集群所在行的详情。

- 5. 在详情页面上部页签栏, 单击弹性伸缩。
- 6. 在弹性伸缩页面,单击关闭弹性伸缩划窗。
- 7. 在弹出的弹性伸缩关闭提醒对话框中,单击确定。

1.6.6. 查看弹性伸缩记录

本节介绍弹性伸缩执行完成后,如何查看弹性伸缩活动的执行记录。

操作步骤

- 1. 登录阿里云Databricks控制台。
- 2. 在顶部菜单栏处,根据实际情况选择地域(Region)和资源组。
- 3. 单击上方的集群管理页签。
- 4. 在集群管理页面,单击相应集群所在行的详情。
- 5. 在详情页面上部页签栏, 单击弹性伸缩。
- 6. 右上角单击弹性伸缩记录。
 - i. 正在执行:弹性伸缩活动正在执行。
 - ii. 成功:根据伸缩规则,所有弹性伸缩中的所有节点成功加入或移出集群。
 - iii. 失败: 根据伸缩规则, 没有一个节点被加入或移出集群。
 - iv. 拒绝: 当运行伸缩规则后的实例数大于最大实例数或者小于最小实例数时, 就会拒绝该规则运行。

1.7. 用户管理

本文介绍如何通过Dat abricks 数据洞察的用户管理功能,管理集群DDI用户。DDI用户信息存储在集群自带的 OpenLDAP中,主要用于在DDI集群内的身份认证。DDI用户可以用于访问链接与端口,查看组件Web UI时的 用户身份认证,也可以在开启组件LDAP认证之后进行身份认证。

准备条件

- 1. 已创建集群,详情请参见集群创建。
- 2. 已创建RAM用户,详情请参见创建RAM用户。

? 说明

因为在Dat abricks 数据洞察(简称DDI)用户管理中添加的用户,只能是与RAM用户同名的DDI用户,所以需要先创建RAM用户。

用户分类

用户管理中的DDI用户以列表的形式来进行展示和操作。登录控制台账号不同,所拥有的用户管理权限也不一样,主要分2种。

- 主账号:主账号管理员可以查看集群中所有用户列表,并对所有用户执行重置密码、删除、修改备注等操作,也可以添加用户。
- RAM子账号: 默认RAM账号(仅授予系统策略AliyunDDIDevelopAccess),只能在用户列表中查看与自己 同名的DDI用户,并只能进行重置密码、修改备注的操作,不能添加和删除用户。

↓ 注意

如果当前用户使用的是RAM账号,想要和主账号一样的查看所有用户列表和编辑权限,您可以使用以下 方式

- 1. 主账号登录RAM控制台为RAM账号授予AliyunDDIUserAdminAccess、 AliyunDDIDevelopAccess和AliyunRAMReadOnlyAccess的权限。
- 主账号登录RAM控制台为RAM账号授予AliyunDDIFullAccess和AliyunRAMReadOnlyAccess权限 (不推荐)。

添加用户

- 1. 主账号或被授予相应权限的子账号登录阿里云Databricks数据洞察控制台。
- 2. 在左侧导航栏的菜单下, 单击集群管理。
- 3. 选择集群,单击集群ID。
- 4. 在右侧菜单栏中选择用户管理。
- 5. 在用户管理页面,单击添加用户。

Databricks 数据洞察	Databricks 数据洞察 / 集群	管理 / 集群 11/4) / userManagement			
首页	← Im-test-2	0210126	● Notebook 🗗			
集群	集群基础信息 Spar	rk配置 Spark ULP	Yarn UI M Gandlia 监控 M 库 操作F	志 数据源 用户管	7	
Notebook						
项目空间	添加用户	添加用户			×	帮助
元数据	用户名称					
库	dwm	* 用户名:	请选择	~		
帮助		* 密码:	请输入密码			
			长度限制为8-30个字符,只允许包含字母、数字、-、_			
		* 确认密码:	请输入密码			
			长度限制为8-30个字符,只允许包含字母、数字、-、_			
				商定 取消		
				1007-50 1007		
						8

- 6. 选择RAM用户,并设置密码和确认密码。
- 7. 单击**确认**。

操作用户

- 1. 重置密码: 您可以修改已添加用户的密码。
- 2. 删除: 您可以删除已添加的用户。

常见问题

用户管理中的用户只在当前集群中生效,所以每个集群的DDI用户不互通。例如,在cluster-1上添加DDI用户A之后,并不会共享给cluster-2。如果需要在cluster-2上使用DDI用户A,则需要在cluster-2上重新添加用户A。

1.8. 集群扩容

当Dat abricks 数据洞察集群计算资源或存储资源不足时,您可以对集群进行水平扩展。目前支持Worker 实例的扩容。

扩容操作步骤

- 1. 登录阿里云Databricks 数据洞察控制台。
- 2. 在左侧导航栏的菜单下, 单击集群管理。
- 3. 选择想要扩容的集群,单击**集群ID**进入集群详情页面。
- 4. 在集群基础信息页面中单击资源管理按钮下的扩容按钮。

按量付费集群扩容

- 1. 在弹出的扩容对话框中,设置相应节点的扩容参数。
- 2. 单击确认按钮,完成扩容。

Databricks 数据洞察	Databricks 数据洞察 / 集群管理 / 集群	
首页	۲œ X	
集群		-
Notebook	Worker实例组	
项目空间	Worker实例组 机器组名称: 核心零例组	
元数据		
库	配置: ecs.go.xiarge 4検 166 ESSU云蓝 80GB*4袂	znou / cn-nangzno :47:04
帮助	付费类型:按量付费	
	当前Worker数量: 3台	
	塘加数量: 2 台	
	现价 ¥ 0.01/小时 省 ¥ 5.638/小时	
	确定	80GB ESSD云盘* 1
		80GB ESSD云盘* 4
	关闭	

包年包月集群扩容

- 1. 在弹出的扩容对话框中,设置相应节点的扩容参数。
- 2. 单击确认按钮, 弹出订单弹窗。

Databricks 数据洞察	Databricks 数据洞察 / 集群管理 / 集都 / 详销
首页	← fjl-dbr7数据源测试-202101231
集群	集群基础信息 Spark配置 Spark UI 0 Yarn UI 0 Ganglia 监控 0 库 操作日志 数据源 用户管理
Notebook	
项目空间	
元数据	
库	集創名称: 街-db/7数館 订単创建中 付置类型: 包午包月 301 311 312 312 313 313 314
帮助	39期时间: 2021-02-27
	您的订单已创建成功,请立即支付。
	Databricks Runtimel%
	取消订单 前往支付 完成 主机信息
	Master 类例组 包午包月
	◆主机設量: 1 ◆ CPU: 4模 ◆ 内存: 16GB ◆ Instance: ecs.g6.xlarge ◆ 数据盘配置: 80GB ESSD云盘* 1
	Worker 实例编 包年包月
	◆主机数量: 3

- 3. 单击前往支付按钮,进入订单页面,完成支付。
- 4. 回到控制台单击完成按钮完成扩容。

查看扩容状态

你可以在基础信息页面中主机实例中查看扩容状态。

Databricks 数据洞察	Databricks 数据洞察 / 集群管理 / 集群	. / 详情					
首页	← fjl-dbr7数据源测i	武-202101231	Notebook 🗗				
集群	集群基础信息 Spark配置 Spa	rk UL F? Yarn UL F? Ganolia I		据源 田户管理			
Notebook				1111 BUT 11			
项目空间	■ 资源交配 ∨ 四 集群操作 ∨	· 國费用管理 > C					
元数据	集群信息						
库	集群名称: fil-dbr7数据源测试-202101231 ∠ 付费举型: 句도句月	集群 ID: 当前状态: ⑤ 空闲	集群类型:标准型 运行时间:9小时21分21	矽	地域 / 可用区 ID: cn-hangz 创建时间: 2021-01-26 12:5	thou / cn-hangzho	
帮助	到期时间: 2021-02-27 00:00:00	自动续费: 未开通	Knox账号: 查看		EXECUTE: 101		
	服务目录: ""	and the second s					
	软件信息						
	Databricks Runtime版本: DBR 7.3, Spark 3.0	1, Scala 2.12					
	主机信息						
	Master 实例组 包年包月						
	◆ 主机数量: 1 ◆ C	CPU: 4核 ◆ P	∮存: 16GB ◆	Instance: ecs.g6.xlarge	♦ 数据盘配置: 8	80GB ESSD云盘* 1	
	Worker 实例组包年包月 ()扩容中 扩容	§实例数: 2个 🥝					
	◆ 主机数量: 3 ◆ C	CPU: 4核 ◆ P	9存: 16GB ◆	Instance: ecs.g6.xlarge	♦ 数据盘配置: 8	80GB ESSD云盘* 4	
							the local sectors in the local

1.9. 释放集群

当集群不再使用时,您可以随时进行释放,以节约成本。

? 说明

待释放集群的状态必须是创建中、运行中或空闲中,其他状态不支持释放。

按量付费集群释放

- 1. 使用阿里云账号登录Databricks数据洞察控制台。
- 2. 在Databricks数据洞察控制台页面,选择所在的地域(Region)。
- 3. 在左侧导航栏中,单击集群。
- 4. 设置释放。
 - 在集群管理页面,单击待释放集群所在行的释放。

	C-FDC01742873A522E test-20210126	◎ 标准型	Ѕ 空闲	2021-01-26 17:47:04 5天16小时41分35秒	按量付费	详情 Notebook 🗗 Spa	ark UI 🗗 Yarn UI 🗗	释放			
0	单击待释放集郡	¥的集群ID	,在集群	基础信息页面,	单击集群	≨操作> 释放 。					
	Databricks 数据洞察 / 集群管理 / 集群 (C-) / 详信 ← -test-20210126 C- ○ ○ ○ Notebook @ 集群系統信息 Spark UI @ Yam UI @ Ganolia 监控 @ 廣作日志 教振源 用户管理										
	 	群操作 へ CG 放	费用管理 🗸	C							

5. 在弹出的集群管理-释放对话框中,单击释放。

包年包月集群释放

1. 使用阿里云账号登录Databricks数据洞察控制台。

2. 选择费用单击用户中心。

9 G pre-data	ioneks.console.aliyun.com/#/en-nangznou/cluster									м н л	Update :
(-) 阿里云 #病1	(积州) 👻				Q 提示文档、控制台、API、解决方面	1000 20	备案 企业	支持 官同	۵.	¥ Ø	₿ <mark>预发2.0</mark>
ricks 数据洞察	🟮 "Notebook"、"Spark UI"、"Ganglia監控"、"Yarn UI"制	§Web Ui在首次登录会提示"这不是一个私密连接"告	暨,您可以单击"帮助文档"查看详细解决方法 17			充值					
	Databricks 数据调察 / 集群管理					订单					
	集群管理					发展					
nak.						续费管理					
10	□ 回建集群 请输入集群名称 请法	S持集群状态 V Q.搜索 Ø				用户中心 🗲					G
ag	集群ID/名称	集群类型 状态	创建时间/运行时间	付费类型	操作 🕑						
3.选择	订单管理单击退订	丁 管理。									
4. 产品	选择。										
产品名称:	data 🔨	实例ID:		创建时间: 起始日期		⊞ 搜索	I.K.				
	Databricks数据洞察(包年	实例名称/ID		地域	Я	F始/结束时间		操作	F		
	智能数据助理										

5. 按集群ID退订资源(也可以退续费)。

管理集群·集群配置

产品名称:	Databricks数据洞察	实例ID: 创建时间: 起始日期	結束日期 首 搜索		
	产品名称	案例名称/ID	地域	开始/结束时间	操作
		C-F761AB343C25FDAD	华东1(杭州)	2020-10-12 11:48:21 2020-11-13 00:00:00	8 A
		C-B80B5EA3B2AB6C81	华东1(杭州)	2020-10-15 11:47:01 2020-11-16 00:00:00	且订资源
		C-88FD4D80603D5149	华北5(呼和浩特)	2020-10-27 21:11:45 2020-11-28 00:00:00	退订资源
		C-9181957FABF52F7E	华东1(杭州)	2020-10-27 10:02:38 2020-11-28 00:00:00	退订资源
		C-6748D93457E87B91	华北5(呼和浩特)	2020-10-27 18:54:26 2020-11-28 00:00:00	退订资源
	-	C-B189B21351504080	华东1(杭州)	2020-10-29 17:19:59 2021-01-30 00:00:00	退订资源 退结费
		C-D2BF5B88E8995211	华东1(杭州)	2020-09-27 14:22:54 2021-03-28 00:00:00	退订资源
	-	C-4D2359C1BC1FF6A4	华东1(杭州)	2020-10-20 20:24:05 2020-11-21 00:00:00	退订资源
		C-C8DB3C298DABB3DB	华东1(杭州)	2020-10-26 15:57:35 2020-11-27 00:00:00	退订资源
	-	C-AA97DBFE8D229325	华东1(杭州)	2020-11-02 14:47:38 2020-12-03 00:00:00	退 下资源

2.RAM访问控制

2.1. 角色授权

首次使用Databricks数据洞察服务创建集群时,需要使用主账号为Databricks数据洞察服务授权名为 AliyunDDlEcsDefaultRole的系统默认角色。同时需要您创建一个系统目录存储Bucket。

背景信息

关于角色详细信息,具体可以参见RAM角色概览。

 通过授予AliyunDDIEcsDefault Role角色,您创建的Databricks数据洞察集群可以以免AK的方式访问阿里云OSS资源,详细信息请参见基于MetaService免AccessKey访问阿里云资源。注意首次使用 Databricks数据洞察服务时,必须用主账号完成默认角色授权和Bucket创建,否则子账号和主账号不能使用Databricks数据洞察。

角色授权流程

1. 首次使用Databricks数据洞察服务创建集群时, 会弹窗提示授权。

Databricks 数据洞察 / 集	详管理 / 创建集群		
← 创建集群	1		
2。基础设置			当前配置
* 集群名称:	长度限制为1-64个字符,只允许包含中文、字母、数字、-、_		基础配置
* Knox账号: 🕜			Databricks Runtime版本: DBR 8.2, Spark 3.1.1, Scala 2.12 Python版本: 3
* Knox密码:	角色授权		付费类型: 包年包月 付费时长: 1月 可用区: 可用区
	产品的使用需要 主账号 授权以下DDI角色		Master 实例
Databricks Runtime版z	AliyunDDIEcsDefaultRole		突例类型:ecs.g6.xlarge 高可用:关闭 系统盘配置:ESSD云盘
Python版本:		单击前往RAM进行授权 查看角色说明	系统盘大小: 120GiB*1块 数据盘配置: ESSD云盘
付费类型:	包年包月按量付费		数据蓝大小: 80GIB * 1 块 数量: 1

2. 点击**单击前往RAM进行授权**。单击**同意授权**,将默认角色AliyunDDIEcsDefaultRole授予给Databricks数据洞察服务(spark)。

管理集群·RAM访问控制

三 C-) 阿里云 1 命 工作台	Q 搜索	费用	工单	ICP 备案	企业	支持	App	>_	¢.	₽ 0	简体	0
云资源访问授权 如需传动角色权限,请能往 RAM 控制台角色管理中设置,需要注意的是,错误的配置可能导致	云产品无法获取到必要的权限。											
spark 请求获取访问您云资源的权限。 下方是系统创建的可供 spark 使用的角色, 授权后, spark 拥有对您云资源相应的访问权限。 AliyunDDIEcsDefaultRole 展开详情												
同意接权 取消												

3. 完成以上授权后,您需要刷新Databricks数据洞察控制台,然后即可进行相关操作。如果您想查 看AliyunDDIEcsDefaultRole相关的详细策略信息,您可登录RAM的控制台查看。

AliyunDDIEcsDefaultRole权限内容

默认角色AliyunDDIEcsDefaultRole包含系统权限策略为AliyunDDIEcsDefaultRolePolicy,OSS相关权限内容如下。

```
"Action": [
   "oss:GetObject",
   "oss:ListObjects",
   "oss:PutObject",
   "oss:DeleteObject",
   "oss:ListBuckets",
   "oss:AbortMultipartUpload",
   "oss:ListMultipartUploads"
]
```

2.2. 服务关联角色

本文介绍Databricks数据洞察服务关联角色AliyunServiceRoleForDDI以及如何删除该角色。

背景信息

Dat abricks数据洞察服务关联角色AliyunServiceRoleForDDI是Dat abricks数据洞察在某些情况下,为了完成自身的某个功能,需要获取其他云服务的访问权限而提供的RAM角色。更多关于服务关联角色的信息请参见服务关联角色。

AliyunServiceRoleForDDI应用场景

Dat abricks数据洞察集群创建及数据源绑定功能需要访问<mark>云服务器ECS、专有网络VPC</mark>等云服务的资源时,需 要通过服务关联角色AliyunServiceRoleForDDI获取访问权限。

AliyunServiceRoleForDDI权限说明

AliyunServiceRoleForDDI具备以下云服务的访问权限:

"Action": ["vpc:DescribeVSwitches", "ecs:CreateNetworkInterface", "ecs:DeleteNetworkInterface", "ecs:DescribeNetworkInterfaces", "ecs:CreateNetworkInterfacePermission", "ecs:DeleteNetworkInterfacePermission", "ecs:CreateSecurityGroup", "ecs:AuthorizeSecurityGroup", "ecs:AuthorizeSecurityGroupEgress"]

删除AliyunServiceRoleForDDI

如果您需要删除AliyunServiceRoleForDDI服务关联角色,需要先释放依赖这个服务关联角色的Databricks数据洞察集群。

具体操作步骤如下:

- 1. 登录RAM控制台,在左侧导航栏中单击RAM角色管理。
- 2. 在RAM角色管理页面的搜索框中,输入AliyunServiceRoleForDDI,自动搜索到名称为 AliyunServiceRoleForDDI的RAM角色。
- 3. 在右侧操作列,单击删除。
- 4. 在删除RAM角色对话框,单击确定。
 - i. 如果当前账号下存在关联的Databricks数据洞察集群,则需先释放集群后才能删除 AliyunServiceRoleForDDI,否则提示删除失败。
 - ii. 如果当前账号下已释放所有Databricks数据洞察集群,则可直接删除AliyunServiceRoleForDDI。

常见问题

为什么我的RAM用户无法自动创建Databricks数据洞察服务关联角色 AliyunServiceRoleForDDI?

您需要拥有指定的权限,才能自动创建或删除AliyunServiceRoleForDDI。因此,在RAM用户无法自动创建 AliyunServiceRoleForDDI时,您需为其添加以下权限策略。参见创建自定义权限策略

管理集群·RAM访问控制

```
{
    "Statement": [
        {
            "Action": [
                "ram:CreateServiceLinkedRole"
            ],
            "Resource": "acs:ram:*:主账号ID:role/*",
            "Effect": "Allow",
            "Condition": {
                "StringEquals": {
                    "ram:ServiceName": [
                        "ddi.aliyuncs.com"
                    1
                }
            }
        }
   1.
    "Version": "1"
}
```

说明请将 主账号ID 替换为您实际的阿里云账号(主账号)ID。

相关文档

• 服务关联角色

2.3. 为RAM用户授权

为确保RAM用户能正常使用Dat abricks 数据洞察控制台的功能,您需要使用云账号登录访问控制 RAM(Resource Access Management),授予RAM用户相应的权限。

背景信息

访问控制RAM是阿里云提供的资源访问控制服务,更多详情请参见什么是访问控制。以下举例访问控制RAM 的典型场景:

- 用户:如果您购买了多台Databricks数据洞察集群实例,您的组织里有多个用户(如运维、开发或数据分析)需要使用这些实例,您可以创建一个策略允许部分用户使用这些实例。避免了将同一个AccessKey泄露给多人的风险。
- 用户组:您可以创建多个用户组,并授予不同权限策略,授权过程与授权用户过程相同,可以起到批量管理的效果。

权限策略

权限策略分为**系统策略**和自定义策略。

- 系统策略: 阿里云提供多种具有不同管理目的的默认权限策略。Databricks 数据洞察经常使用的系统策略:
 - AliyunDDIFullAccess:管理Databricks数据洞察的权限,主要包括对Databricks数据洞察的所有资源的 所有操作权限。
 - AliyunDDIDevelopAccess: Databricks 数据洞察开发者权限,与AliyunDDIFullAccess策略相比,不授予 集群的创建和释放等操作权限。

 自定义策略:需要您精准地设计权限策略,适用于熟悉阿里云各种云服务API以及具有精细化控制需求的 用户。详细方法可以参见创建RAM自定义策略。

○ 注意

系统策略默认仅为RAM用户提供查看OSS Bucket和Object列表权限, RAM用户无法编辑OSS Bucket和 Object。如需更多OSS权限策略请参见OSS数据权限隔离

授权建议

首次开通购买,如果安全合规条件允许,我们推荐使用阿里云主账号或具有AdministratorAccess权限的 RAM子用户/角色。

如果需要更精细的权限管理,您需要使用阿里云主账号为您的RAM用户账号做如下授权。

- 授权系统策略 AliyunDDIFullAccess
- 如需购买包年包月集群,需要使用自定义策略授权支付订单的权限点 bss:PayOrder
- 因Databricks数据洞察产品依赖RAM跨服务授权,首次开通产品服务需要RAM的管理权限进行授权,建议 通过主账号授权AliyunRAMFullAccess策略或者添加如下自定义策略:

```
{
    "Statement": [
        {
            "Action": [
                "ram:CreateRole",
                "ram:AttachPolicyToRole"
            ],
            "Resource": [
                "acs:ram:*:*:role/AliyunDDIEcsDefaultRole",
                "acs:ram:*:system:policy/AliyunDDIEcsDefaultRolePolicy"
            ],
            "Effect": "Allow"
        },
        {
            "Action": "ram:CreateServiceLinkedRole",
            "Resource": "acs:ram:*:*:role/*",
            "Effect": "Allow",
            "Condition": {
                "StringEquals": {
                    "ram:ServiceName": [
                        "ddi.aliyuncs.com"
                    1
                }
            }
        }
   ],
    "Version": "1"
}
```

○ 注意

以上三个鉴权策略, AliyunDDIFullAccess影响后续Dabricks数据洞察产品使用; 另外两个策略只在您 首次开通和购买Dabricks数据洞察集群时需要授予。

授权RAM用户

执行以下步骤在访问控制RAM控制台授权RAM用户Dat abricks 数据洞察相关权限。

- 1. 使用云账号登录RAM控制台。
- 2. 单击左侧导航栏的人员管理 > 用户。
- 3. 单击待授权RAM用户所在行的添加权限。
- 4. 单击需要授予RAM用户的权限策略, 单击确定。具体权限策略请参见权限策略。
- 5. 单击完成。完成授权后,权限立即生效。

3.运维监控

3.1. 集群运维与监控

本文主要介绍Databricks数据洞察在集群维度的监控与告警能力。

背景信息

目前Databricks数据洞察主要采用**阿里云监控**服务,完整对接DBR指标体系,实现多维度、多实例、多时段的指标图表监控,用户可以根据业务情况,灵活自定义告警规则,摆脱底层的束缚,根据平台产品层面的能力,高效感知问题,运维业务。

用户可以登录Databricks数据洞察监控,选择对应的**集群属地与集群名称**,即可开启集群运维应用,示例如 下图:

📃 (-) 阿里云 🗆 🌣	工作台		Q 报索	费用 工单 ICP 备案 企业 支持 App 🖾 🛕	H 🖉 🕸
云监控(回到旧版)	云监控 / 云产品监控				
假览	← Databricks数据洞察 > ##2 (#原) >				
应用分组				创建报警规则	报警规则列表
主机监控	基本指标 Hadoop-queue指标				
云产品监控	资源类型: role role:emr-worker-6				
Kubernetes容器监控	1,087 2,037 6,037 12,087 17 2 T 14 T 20009658328 09:42	2022年5月12日 00:42 前			毎行展示3 >
Dashboard		TOTTO SAL DAT			
网络分析与监控 へ	网络流入速率(bits/s)	网络流出通率(bits/s)	🏚 🦯 cpu_id	idle(old)(%)	ė 🗸
站点监控(定时拨测)	平均值 ~ (周期:1分钟)	平均值 ~ (周期:1分钟)	平均值	<u>a</u> ~ (周期:1分钟)	
一次性拔测工具	4.418M	41.784K 39.063K	100.0	000	
事件监控 ~	3.815M	29.297K	80.0	000	
日志监控	2.861M		60.0	200	
自定义监控	1.907M	19.5316	40.0	000	
报警服务 ~	976.563K	9.766K	20.0	000	
资源消耗	0.000 08:56:40 09:13:20 09:30:00 09:41:00	0.000 08:43:00 08:56:40 09:13:20	09:30:00 09:41:00 0	000 08:43:00 08:56:40 09:13:20 09:30:00	0 09:41:00
企业云监控 へ	● 网络流入速率	● 网络流出速率		● cpu空闲率	
监控大盘					

具体内容可参考云监控文档。

指标监控

目前默认的指标可以分为两个维度,即机器节点基本指标和Hadoop-queue相关指标。

后者可以总结为YARN中队列的相关指标内容,下文均称为队列指标,将分别给出详细定义。

机器节点基本指标

这部分指标主要围绕CPU、内存、磁盘、网络等监控项,满足主机的基本运维需求。云监控支持的监控项, 请参见监控项说明。

队列监控指标

该部分指标通过YARN当中的Resource。Manager统计获取,通过JMX的方式进行暴露,平台选取了较为重要的指标进行监控,直接描述了集群的资源与作业情况。

具体指标项(均为分钟级)见下表:

监控项名称	监控项含义	单位	Dimension s	Statistics	说明
-------	-------	----	----------------	------------	----

管理集群·运维监控

监控项名称	监控项含义	单位	Dimension s	Statistics	说明
ActiveUsers	Active的用户数量	count	用户维度、集 群维度	Maximum 、 Minimum 、 Average	
AggregateCo ntainersAlloc ated	总共分配的container 个数	count	用户维度、集 群维度	Average	
AggregateCo ntainersRele ased	总共释放的container 个数	count	用户维度、集 群维度	Average	
AllocatedCo ntainers	每分钟内分配的 container个数	count/分钟	用户维度、集 群维度	Average	描述该集群资 源现状下的调 度能力状况
AppsComple ted	已完成的作业数	count	用户维度、集 群维度	Average	
AppsFailed	失败的作业数	count	用户维度、集 群维度	Average	
AppsKilled	被杀死的作业数	count	用户维度、集 群维度	Average	
AppsPending	等待的作业数	count	用户维度、集 群维度	Average	描述资源不足 的作业数
AppsRunning	运行中的作业数	count	用户维度、集 群维度	Average	
AppsSubmitt ed	提交的作业数	count	用户维度、集 群维度	Maximum 、 Minimum 、 Average	
AvailableMB	当前队列可用的内存大 小	Mbyte	用户维度、集 群维度	Maximum 、 Minimum 、 Average	描述当前集群 的可用内存

Dat abricks 数据洞察

监控项名称	监控项含义	单位	Dimension s	Statistics	说明
AvailableVCo res	当前队列可用的vCores 个数	count	用户维度、集 群维度	Maximum 、 Minimum 、 Average	描述当前集群 的可用vCore
PendingCont ainers	等待的container个数	count	用户维度、集 群维度	Maximum 、 Minimum 、 Average	
ReservedCon tainers	保留的container个数	count	用户维度、集 群维度	Average	描述预分配的 container数

自定义指标配置

为了支持用户更加具有自由度的指标监控,平台提供了自定义指标,步骤如下:

1. 首先用户需要自定义创建大盘,创建成功后点击自定大盘名称,如下图:

☰ (−)阿里云	ດ I∩	台				Q 搜索	费用 工单 ICP 备案	企业 支持 🔂 🖸	🔮 ## 🕥 🤉 🖌 🕻
云监控(回到旧版)		Dashboard							
概范		Dashboard							
应用分组		自定义大盘	网络监控大盘	云产品监控大盘	秒级监控				
主机监控		(200) 十倍々約	0						Stort &
云产品监控		2) 占击已	创建的大母名称		1) 创建白完义大母				3970 X 12
Kubernetes容器监控		test 🗲							
Dashboard		ID: one cost of the data on			4				
网络分析与监控	^	创建时间:2022年5月12日 12:15:08			T				
站点监控(定时拨测)		修改时间:2022年5月12日 12:15:08							
一次性拨测工具									
事件监控	~								
日志监控									
自定义监控									
报警服务	~								
资源消耗									

2. 其次添加业务所需的自定义指标项,点击右上的添加图表,可从不同图表形式描述自定义指标,如下图:

M/11/2142					
▲ 折线图	面积图	▲ TopN表格	热力图	U 供图	
云产品监控	自定义监控				
品名称	Databricks数据洞察				\sim
图表名称	队列中的pending内存量				
			Y轴显示	范围: 0	
488.281K					
390.625K 292.969K 195.313K 97.656K 0.000 15:24:00 15	5:28:20 15:36:40 ● roo版列中per	15:45:00 15:53 ding的内存量-平均值- ddi-tc	3:20 16:01:40 3:40 19:40 3:20 19:4721D	16:10:00 }-emr-header-1	16:22:00
390.625K 292.969K 195.313K 97.656K 0.000 15:24:00 15 监控项	5:28:20 5:28:20 「root版列中pending的内存量 () 应用分组 ④ 实例	15:45:00 15:53 ding的内存量-平均值- ddi-te	3:20 16:01:40 est(C-6ACA5EEC1194F21D	16:10:00)-emr-header-1	16:22:00

上图给出了"队列中pending的内存量"指标的基本添加方式,用户可以根据自己的需求进行配置。 这部分的更详细的操作可以参考管理自定义大盘。

告警配置

当用户需要对指标及时感知,可以配置具体的告警规则。若资源指标的监控达到报警条件,云监控会自动按 配置的告警信息进行反馈,高效辅助用户排查问题,进行运维。

报警规则配置

用户可以通过在Databricks的云产品监控页面,选择创建新的报警规则。

本文给出一个配置示例,当用户希望感知集群作业资源状态时,可以针对队列指标进行告警规则配置,其中 队列指标中AppsPending就可以描述上述需求,步骤如下:

1. 登录Dat abricks数据洞察监控,选择对应的集群属地与集群名称,点击右上角的创建报警规则。

Dat abricks 数据洞察

云监控 / 云产品监控		
← Databricks 致 据 洞祭 × ^{¥北2(北京)} ×		
		第一步创建报警规则报警规则列表
基本指标 Hadoop-queue指标		
资源类型: role:emr-worker-6		
1小时 3小时 6小时 12小时 1天 3天 7天 14天 2022年5月12日 08:42	- 2022年5月12日 09:42	毎行展示3 ~
网络流入速率(bits/s) 🏚 🖉	网络流出速率(bits/s) 🛕 🖉	cpu_idle(old)(%)
平均值 💙 (問期:1 分钟)	平均值 >> (周期:1分钟)	平均值 🗸 (周期:1 分钟)
4.418M	41.784K	100.000
3.815M	39.053K	80.000
2.861M	29.297K	60.000
1.907M	19.531K	
		40.000
976.563K	9.766K	40.000
976.563K	9.766K	40.000 20.000 0.000
976.563K 0.000 08.43.00 08.56.40 09.13.20 09.30.00 09.41.00	9.766K 0.000 08:43:00 08:56:40 09:13:20 09:30:00 08:41:00	40.000 20.000 0.000 08.43.00 08.56.40 08:13.20 09:30:00 09:41:00

2.创建具体的报警规则,可以看到如下左图中的配置项,我们先关注最核心的**增加规则**配置,点击后产生如 右图的具体指标告警规则配置,选取AppsPending作为监控指标,配置不同程度的指标阈值,这部分的具体 值需要参考业务规模和具体的容忍度,图中值仅参考。

创建报警规则	×	添加规则描述
/* &		
Databricks数据洞察		规则名称
资源范围		大量作业资源告警
全即资源 应用分组 奖例		
关联资源		指标类型
C-B9F518B2704BF6C2 ×	\sim	单指标 多指标 动态阈值
规则描述		
规则名称 规则描述 资源描述		监控指标
没有数据		root队列中pending的作业数 V
第二步核心		请洗择维廖
个亮级设置		
报警职系人组		阈值及报警级别
请选择		连续 3 个周期(1周期=1分钟) V 电话+短信+邮件+钉钉机器人
URL頻繁回调和損量触发		紧急 Critical 平均值 ✓ >= ✓ 10 count
10. m (a) 40		
		连续 3 个周期(1周期=1分钟) ☆ 短信+邮件+钉钉机器人
		Warn 平均值 V >= V 5 count
消息服务 MNS — topic		養通 连续 3 个周期(1周期=1分钟) ✓ 邮件+钉钉机器人
无数据报警处理方法		Info 平均值 V >= V 2 count
18定 取3州		

 其他告警配置,可根据业务要求配置,详细参数描述请参见创建阈值报警规则。配置报警联系人详情请 参见创建报警联系人或报警联系组。创建成功后,点击右上角的报警规则列表,可以查询到已创建的 报警规则,如下图:

☰ (-) 阿里云 🚓 :	I#6	Q 搜索	费用 工单	ICP 备案 企业 支持	🔮 🕸 ି ୍ କ୍ ତି ସ 🗄
云监控(回到旧版)	投辦规则				
概览	报警规则列表				
应用分组					
主机监控	自該批監批則	×	u		6
云产品监控		联资源	报警规则	报警联系人组	操作
Kubernetes容器监控	大量作业资源pending Databucke数规则数 の	via-emr-hearler_1 re	root队列中pending的作业数(Critical) 连续3个周期平均值>=10 count 就通知 (Marn) 连续3 个周期平均值>=5 count	test	详情 报警历史
Dashboard	uuid_ae0e1a21ff0a7482e0		就通知(Info) 连续 3 个周期平均值>=1	1001	1922、「删除数 1
网络分析与监控 ^			count without		
站点监控(定时拨测)	日 删除				
一次性拔测工具					
事件监控 >					
日志监控	2				
自定义监控					

4.当告警异常发生时,首先在云监控的报警列表当中可以发现异常的告警规则项,如下图:

报警规则列表

013 1 214	强整规则	状态 V Databricks数据	 	/ 規則名称 /	Q			G
	状态	规则名称 小	产品名称	关联资源	报警规则	报警联系人组	操作	
	\$ 报警	大量作业资源pending uuid_bfce1e39f86db082d	Databricka数据洞察	role=emr-header-1,re	root队列中pending的作业数(Critical) 连续3个周期平均值>=10 count 就通知 (Warn) 连续3个周期平均值>=3 count 就通知(Info) 连续3个周期平均值>=1	test	详情 报警历史 修改 删除 【	

符合具体的报警规则时,会进行相应的渠道反馈,如下图:



3.2. 项目空间运维与监控

本文主要介绍Databricks数据洞察中项目空间具有的运维与监控能力。

背景信息

项目空间的运维能力主要辅助作业开发和工作流生产,前者提供作业开发过程中的日志监控、版本监控和操 作审计监控;后者兼容前者的功能,更提供针对调度场景下的监控与告警。

在项目维度,设置了运维中心,协助用户监控作业状态,运维管理。其中包含作业总览、作业运行记录、工 作流记录、流式作业监控和审计日志监控等功能,全方位提升用户运维效率。

☴ (-) 阿里云 🎄 エ	2作台 🔮 华北2(北京) 🗸			Qł	夏索	费用 工单 ICP 备案 企业	支持 🗗 🖸 🛕	🛉 🖗 🕐 🦉
Databricks 数据洞察	く返回 TEST_API ✓ 作业开发	项目管理 运续中心						● 帮助
首页	概览 作业运行记录 工作	流记录 流式作业 审计日志						
集群	请选择作业类型 🖌 请输入	作业实例ID 请输入作业ID	起始日期	- 结束日期	曲 请选择状态 丶	/ 报索		
Notebook	作业实例ID	作业ID	作业名称	作业类型	开始时间	结束时间	状态	操作
项目空间	FJI-3D1C11A78E228F2E	FJ-CB2718DD8DA5B5C2	Spark-perf-test-largeScale	X SPARK	2022-05-12 23:45:49	2022-05-12 23:51:17	⊗ FAILED	详情 停止
元数据			•					
库	FJI-C0FE783A24CD4160	FJ-E67F97C14BCEC23D	Spark-perf-test-smallScale	√ 3 SPARK	2022-05-12 23:31:16	2022-05-12 23:32:55	⊗ FAILED	详情 停止
帮助	FJI-C6E1A53F8E2B3B96	FJ-E67F97C14BCEC23D	Spark-perf-test-smallScale	✓I SPARK	2022-05-12 23:23:52	2022-05-12 23:29:45	⊗ FAILED	详情 停止
	FJI-1F06FA3C84CD8278	FJ-E67F97C14BCEC23D	Spark-perf-test-smallScale	₽ SPARK	2022-05-12 19:38:37	2022-05-12 19:45:18	KILLED	详情 停止
	FJI-A73007C7C078B5EF	FJ-E67F97C14BCEC23D	Spark-perf-test-smallScale	✓T SPARK	2022-05-12 15:51:34	2022-05-12 15:59:13	⊗ FAILED	详情 停止
	FJI-D8578C448A54C6C7	FJ-E67F97C14BCEC23D	Spark-perf-test-smallScale	J SPARK	2022-05-12 15:47:43	2022-05-12 15:51:35		详情 停止

作业运维

这里我们模拟一个错误作业作为示例,给出项目空间在对该作业运维的操作。

日志监控

当我们在项目空间中运行测试作业时,若运行过程中作业报错或作业失败,在下方的调试框内会给出相应的 执行日志以及错误判断信息,在运行记录当中可以查看过往的提交信息,以供用户排查作业问题,如下图:

18 - V 1		
日志 运行记录 审计日志 版本控制 + 插入OSS路	至 ∂去OSS控制台上传	d ~ ~
2022-05-13 13:03:07.507 [main] BRNOK c.a.e.f.a.j.l.impl.CommonShellJobLauncherImpl - [FJI-59405999AA814F12_0] Failed to execute command, exit code=1 2022-05-13 13:03:07.508 [main] INFO c.a.e.f.a.j.l.impl.CommonShellJobLauncherImpl - [COMMAND][FJI-59405999AA814F12_0] Finished command line, exit code=1. Fri May 13 13:03:07 CST 2022 [JobLauncherRunner] INFO Closing job launcher		详细日志内容
2022-05-13 15/03/07.510 [main] INFO Cratemiritor.agent.jobs.launcher.soub.auncher.soub.cr.[pt] = [pt]=5805959AAB4F12_0] Elosing 2022-05-13 13/03/07.510 [main] INFO crate.fr.aj.l.impl.CommonShellJobLauncherEngl = [pt]=5805959AAB4F12_0] Stopping command executor Fri May 13 13/03/07 CSF 2022 [LocalJobLauncherAM] URKOR Failed init or run, diagnostics info:		
. ## DIAGNOSTICS ##] ← 初步判断信息		
221: diagnostics: User class threw exception: org.apache.spark.SparkException: Job aborted.		
293: Caused by: org.apache.spark.SparkException: Job aborted due to stage failure: Task 10 in stage 7.0 failed 4 times, most recent failure: Lost task 10.3 i r.a cluster_10256 organize 11.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1	n stage 7.0 (TID 34)	(emr-worke
294: at scala.svs.packages.error(package.scala:30)		
328: at org.apache.spark.scheduler.DASScheduler.failJobAndIndependentStages(DAGScheduler.scala:2783)		
335: at org.apache.spark.scheduler.DAGScheduler.\$anonfun\$handleTaskSetFailed\$1(DAGScheduler.scala:1260)		
336: at org.apache.spark.scheduler.DAGScheduler.\$anonfun\$handleTaskSetFailed\$1\$adapted(DAGScheduler.scala:1260)		
338: at org.apache.spark.scheduler.DAGScheduler.handleTaskSetFailed(DAGScheduler.scala:1260)		
343: Caused by: java.lang.RuntimeException: Could not find dsdgen at /home/hadoop/tools/dsdgen or //home/hadoop/tools/dsdgen. Run install		
344: at scala.sys.package\$.error(package.scala:30)		
381: final status: FAILED		
385: 22/05/13 13:03:06 ERROR [main] Client: Application diagnostics message: User class threw exception: org.apache.spark.SparkException: Job aborted.		
45): Caused by: org.apache.spark.sparkException: Job aborted due to stage failure: Task 10 in stage /.0 failed 4 times, most recent failure: Lost task 10.3 i	h stage 7.0 (TID 34)	(emr-worke
r-s.Gluster-suck30 executor 1): java.lang.kuntimetxception: Could not find dsdgen at /nome/hadcop/tools/dsdgen or //nome/hadcop/tools/dsdgen. kun install		
400; at organological to (pachage solation)		
422: at Dig apacies spark, solieuter i physicieuter i atriobalitatiegenicello carges (physicieuter i scalar / 03) 400: at organizache gnark schedular Dischedular sanothusbandlataskestasiladi (fDischedular scalar / 120)		
500: a cross anache spark schedular. DAGSchedular sumon universität status a construction a cons		
502: at org.apache.spark.scheduler.DAGScheduler.handleTaskSetFailed(DAGScheduler.scala:1260)		
507: Caused by: java.lang.RuntimeException: Could not find dsdgen at /home/hadoop/tools/dsdgen or //home/hadoop/tools/dsdgen. Run install		
508: at scala.sys.package\$.error(package.scala:30)		
541: Exception in thread "main" org.apache.spark.SparkException: Application application_1651802247269_0029 finished with failed status		88
### DIAGNOSTICS ###.		

日志监控图

日志	运行记录	宙计日志	版本控制					+插入OSS路径	⊘ 去OSS控制台上传 ┏	~ ~
		×=	시그 대학 / 대							C刷
运行实	例ID	冱	行头例	116求	开始时间	结束时间	状态	操作		
FJI-59	405999AA814	1F12			2022-05-13 13:01:20	2022-05-13 13:03:09	FAILED	详情	停止作业实例	

运行记录列表图

点击运行记录下的**详情**查看,如下图可以看到完整的作业基本信息和提交日志,这部分的信息内容更加全面。

✓ 返回列表 作业实例信息 提交日志 YARN容器列表 审计日志	5 (the second se
基本信息	
ID: FJI-068D92F606C81C04	名称: Logitach-poc
作业类型: SPARK_STREAMING	作业ID: FJ-EE5E283F1A81F96D
状态: FAILED	执行时长: 2分49秒
开始时间: 2022-05-13 14:53:56	结束封间: 2022-05-13 14:58:45
执行集群: C-6ACA5EEC1194F21D (ddi-test)	客户谳: emr-header-1.cluster-302350
外部AppID: LocalApplication_1651802239760_000029	外部容器: container_1651802239760_000029_01_000001
外部状态: FAIL	
执行信息	
执行参数	
cyctime:2022-05-13 14:53:56	
诊断信息	
396: 22/65/13 14:55:42 [main] Client: Application diagnostics message: User application exited with status 1 391: Exception in thread "main" org.agente.spark.sparkException: Application application_165180247469_0035 finish	d vith failed status
### DiAconstics ###. (7) WWY FAILED: CHECK STATUS: leunch status is FAILED	

对于长时间段的日志查询,可以在提交日志中自定义查询具体时间,方便排查具体时间段暴露的问题,如下 图:

★返回 TEST_API → 作业开发 項目管理 返進中心
概览 作业运行记录 工作流记录 流式作业 审计日志
✓ 返回列表 作业实例信息 提交日志 YARN容器列表 审计日志
加载器新 更早 选择时间段 🕜 使用时间窗口模式加载日志,如日志比较分散,请尝试使用"指定时间"的方式。
Dload Upload Total Spent Left Speed
100 87 0 46 100 41 8542 7613::: 9200 {"state":"SUCCESS","mmessage":null,"data":"{}"}
Fri May 13 14:56:43 CST 2022 [LocalJobLauncherAM] INFO Emr flow launcher is quit.
2022-05-13 14:56:43.797 [Shutdown-FJI-06BD92F606CB1C04_0] INFO c.a.emr.flow.agent.jobs.launcher.JobLauncherBase - [FJI-06BD92F606CB1C04_0] Call shutdown hook.
2022-05-13 14:56:43.797 [Shutdown-FJI-06BD92F606CB1C04_0] INFO c.a.emr.flow.agent.jobs.launcher.JobLauncherBase - [FJI-06BD92F606CB1C04_0] Closing
2022-05-13 14:56:43.797 [Shutdown-FJI-05BD92F606CB1C04_0] INFO c.a.emr.flow.agent.jobs.launcher.JobLauncherBase - [FJI-06BD92F606CB1C04_0] This launcher is closed already, skip.
######END_0F_L0G######

操作审计与版本控制

这部分功能可以辅助用户进行运维决策,当作业需要回滚,可以通过创建快照和版本控制结合使用,查询之前的作业版本,有效恢复生产业务,步骤如下图:

? 说明

创建快照操作应首先保存当前作业的内容,确保为最新更新内容。

₿ Spa	rk-perf-test-2 × 🛿 Spark-perf-test-largeScale × 🖉 Spark-perf-test-smallSc	ale ×			
SPA	RK FJ-E67F97C14BCEC23D 作业内容: @	•		保存 ② 运行 作业设	置 👌 上锁 创建快照
1	class com.databricks.spark.sql.perf.tpcds.GenTPCDSData				- THE
2	deploy-mode cluster			弗 一	I PARTIES.
3	name generate_dataset—test				20.
4	queue default				
5	master yarn		+		
6	conf spark.yarn.submit.waitAppCompletion=true				
7	conf spark.driver.cores=2				
8	conf spark.driver.memory=4G				
10					
10	conf spark.executor.memory=86				-
12					Ť
13	conf spark.default.narallelism=200				
14	conf spark.shuffle.service.enabled=true				+
15	conf spark.sgl.autoBroadcastJoinThreshold=-1				
16	oss://databricks-tpc-ds-test/tpcds-jar/spark-sql-perf-assembly	.jar			-
17	-m yarn				
18	-y 1				2
19	实际运行(仅供参考)				^
日志	运行记录 审计日志 版本控制			+ 插入OSS路径	♂去OSS控制台上传 ₫ ヘ ∨
提交用户	请输入提交用户 修订版本号 请输入修订版本号 搜索			第二步,回滚捞	操作
当前版	本 修订版本号	提交用户	创建时间	备注信息	操作
~	e1edf50bb8e54dca9910fa91808c4eea	1499208184838857	2022-05-13 14:28:09	test4	回滚 查看详情
	e987b195e92d40e18420838ffc8c47ce	1499208184838857	2022-05-13 14:27:44	test3	回滾 查看详情 💬
	ddf821816193411ebdae4a6cb131cf34	1499208184838857	2022-05-13 14:25:13	test2。	回滚 查看详情 🔛
	093194a700d1431e9d4f65e35a047046	1499208184838857	2022-05-13 14:20:38	test,	回滚

对于需要复盘操作记录时,可以通过审计日志功能,查询具体操作记录,如下图:

B	志	运行记录	审计日志	版本控制						
时	间			实体		类型	操作	操作人	状态	详情
20)22-()5-13 14:20:3	8	FJ-E67F97C1	4BCEC23D	JOB	commit	1499208184838857	SUCCEEDED	# PARAMS: " test。 " # RESULT: "09319
20)22-()5-13 13:01:20	D	FJ-E67F97C1	4BCEC23D	JOB	submit	1499208184838857	SUCCEEDED	# PARAMS: {"jobId":"FJ-E67F97C14BCE
20)22-0)5-13 13:01:19	9	FJ-E67F97C1	4BCEC23D	JOB	modify	1499208184838857	SUCCEEDED	# PARAMS: {"id":"FJ-E67F97C14BCEC2

工作流运维

工作流的运维主要围绕作业信息展示和调度告警展开,其中包含作业流程监控,作业图形化展示、变更审计日志等功能。

实例信息展示

当需要监控工作流生产时,可以通过运维中心的工作流实例来监控,可以看到工作流的运行阶段与作业的运行状态,当出现问题作业时,可以通过点击**详情**,查询的日志报错信息,后续和上文中**日志监控**的内容类似。操作界面如下图:

<	返回 TEST_API V 作业开发 I	项目管理 运维中心									0 帮
	概览 作业运行记录 工作流记	录 流式作业 审计日志									
	工作流实例信息 图形化展示	审计日志									
	ID: FI-6C8E0A66938FDC4E				쇧	称: test-spark-perf					
	工作流ID: F-EF960A4028C6CF95				B.	行集群:C-6ACA5EEC119	94F21D				
	状态: 🔮 FAILED				ŧ	行时长: 1分39秒					
	开始时间: 2022-05-13 15:16:53				经	東时间:2022-05-13 15:	18:32				
	依赖工作流:										
	所属项目	所属工作流	工作流实例ID 🚺			业务时间	调度时间 🚺)	条件状态	D	
					没有数据	ŝ					
								刷新 暂停工作流	恢复工作流	停止工作流	重跑工作流突例
	作业实例ID	作业名称 1	执行集群	作业类型	作业提交节点主	机	开始时间 🖡	作业完成时间 🖡	执行时长	执行状态 🔽	操作
	FNI-90CD19854137B5AB	Spark-perf-test-smallScale	C-6ACA5EEC1194F21D	SPARK	emr-header-1.	luster-302350	2022-05-13 15:16:53	2022-05-13 15:18:31	1分38秒	FAILED	详情

调度告警

调度作业时,需要配置具体的调度属性和告警设置,具体操作如下图:

		工作流调度配置	×
	0.00. RE IN L'AR 20 XH 🕀 🖬 🖉 🖓 🖉 HR 🧑		
50% 新聞日間	9 WR	其木尾性 细度尾性 牛蒡沿署	
0 Spark-perf-test-smallScale + 2 test-spark-perf +			
 	UTH FUSH CITA AD COMM		
Tiden	5997	④ 暂无钉钉告警群,请添加:系统维护	×
拉制学机 2010年	点击工作流配置	执行失败: 🗾 通知到用户告警组: test 🗸	
END	4° Saado-perf-test-small.	通知到钉钉群告警组: 请选择 🗸	
6270	-		
Spark	END	节点失败: 🗾 通知到用户告警组: test 🗸	
Zrepolin	•	□ 通知到钉钉群告警组: 请选择 ∨	
	0		
		执行成功: 🗾 通知到用户告警组: test 🗸	
		□ 通知到钉钉群告警组: 请选择 ∨	
		启动超时: ✔ 通知到用户告警组: test ✓	
		通知到钉钉群告警组: 请选择 >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	
		节点执行超时: ✔ 通知到用户告警组; test >	
		□ 通知到钉钉群告警组: 请选择 ∨	
	90		

对于设置具体的告警人,请参考项目空间告警人设置。当符合告警条件时,系统会根据设置的反馈方式,通 知到具体报警人,高效处理运维,反馈示例如下图:

【Databricks 数据洞察】FlowNode 通知

阿	J里云 战 ≑				
! 收件	人包含非	公司联系人			
(-) 阿里云			品服务 备案专	区 管理控制台 用户中心 帮助中心 联系客	8
尊敬的阿里	云用户:				
您在flow: te	st-spark-perf,	flow实例: F		3942C, job: Spark-perf-test-	
smallScale,	node实例: FN	I-36B52450E929	3AA2运行约	告束,状态为: 失败	
结束时间:2	022-05-13 12	:07:12			
概要: Flow:	node failed: F	NI-1	93AA2	1.1.0 000	
141前: https:/	/databricks.co	ance/Flag and	#/cn-beijing	//workflow/FP- /detail	
15515	A Daily	ances 11-0000	14201	ucun	
阿里云计算有限	公司				
此为系统邮件请勿	BÆ				
	03440				
				Copyright © PIEz; 2009-2021 All Right Reserved	3

3.3. Spark作业原生运维与监控

本文主要介绍基于原生Spark Ul和YARN Ul提供的运维与监控能力。

背景信息

Databricks数据洞察提供了原生的监控运维UI,例如Spark U和YARN UI,用户可根据自己的需求进行查询。 本文主要介绍原生UI透露出来的日志监控和指标查询,配合平台侧的运维与监控能力,更好的服务于用户, 运维Spark作业。集群Web UI的查询方式,请参考:集群Web UI。

日志监控

Spark UI

通过Spark History Server可以下载Spark的日志,也可以直接查询具体Executors的日志,如下图:

Soork	211 ENADEUOT	History Server
Spark.	3.1.1-SNAPSHOT	,

SPOR. 3.1.1-5	SNAPSHOT	tory Server												
Event log direc	tory: jfs://C us	spark-history												
Last updated: 2	022-05-15 11:18:05													
Client local time	zone: Asia/Shangh	ai												
Show 20 ¢	entries												Search:	ge
Version	👌 App ID		👌 App Name	•			Started	👌 Comp	leted	Duration	🍦 Spark l	Jser 💧 Last Upd	ated	👌 Event Log 👌
3.1.1-SNAPSHO	T application	Liboloore o Leso	com.datab	ricks.spark.s	ql.perf.tpcds.GenTPC	DSData\$	2022-05-10 14:25:0	03 2022-	05-10 15:47:04	1.4 h	hadoop	2022-05-	10 15:47:04	Download
3.1.1-SNAPSHO	T application		com.datab	ricks.spark.s	ql.perf.tpcds.GenTPC	DSData\$	2022-05-10 12:17:1	8 2022-	05-10 12:46:10	29 min	hadoop	2022-05-	10 12:46:10	Download
3.1.1-SNAPSHO	T application		com.datab	ricks.spark.s	ql.perf.tpcds.GenTPC	DSData\$	2022-05-09 09:30	48 2022-	05-09 23:17:39	13.8 h	hadoop	2022-05-	09 23:17:40	Download
3.1.1-SNAPSHO	T application		com.datab	ricks.spark.s	ql.perf.tpcds.GenTPC	DSData\$	2022-05-08 17:24:	13 2022-	05-08 18:45:04	1.3 h	hadoop	2022-05-	08 18:45:04	Download
3.1.1-SNAPSHO	T application		com.datab	ricks.spark.s	ql.perf.tpcds.GenTPC	DSData\$	2022-05-08 16:19:3	37 2022-	05-08 16:49:25	30 min	hadoop	2022-05-	08 16:49:25	Download
Showing 1 to 5 c	of 5 entries (filtered	from 6,645 total entries)											[Previous 1 Next
Spork 3.1.1-	snapshot Job	os Stages Storage	Environment	Executors	s SQL							com.databr	icks.spark.sq	I.perf.tp application U
Executors														
Show Additional	Metrics													
Summary														
	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	• Total Tasks	👌 Task Time (G	C Time) 💧	Input 🖕	Shuffle Read	Shuffle Write	$_{\varphi}$ Excluded $_{\varphi}$
Active(30)	0	0.0 B / 310.5 GiB	0.0 B	203	0	0	4084	4084	42.8 h (4.9 h)		27 MiB	459.5 GiB	459.5 GiB	0
Dead(0)	0	0.0 B / 0.0 B	0.0 B	0	0	0	▶ 0	0	0.0 ms (0.0 m	s)	0.0 B	0.0 B	0.0 B	0
Total(30)	0	0.0 B / 310.5 GiB	0.0 B	203	0	0	4084	4084	42.8 h (4.9 h)		27 MiB	459.5 GiB	459.5 GiB	0

Executors															
Show 20 \$	entries												Sear	rch:	
Executor ID	Address	Status 🖕	RDD Blocks 👌	Storage Memory 🖕	Disk Used 👌	Cores	Active Tasks 🝦	Failed Tasks 🖕	Complete Tasks 🖕	Total Tasks 🖕	Task Time (GC Time) 🖕	Input 👌	Shuffle Read 🝦	Shuffle Write	Logs
driver	emr-worker-6.clus	Active	0	0.0 B / 6.2 GiB	0.0 B	0	0	0	0	0	0.0 ms (0.0 ms)	0.0 B	0.0 B	0.0 в 日志链接	stdout stderr
1	emr-worker-5.clus	Active	0	0.0 B / 10.5 GiB	0.0 B	7	0	0	136	136	1.5 h (11 min)	1.3 KiB	16.1 GiB	15.8 GiB	stdout stderr
2	emr-worker-1.clus	Active	0	0.0 B / 10.5 GiB	0.0 B	7	0	0	143	143	1.3 h (6.3 min)	3.4 MiB	16.2 GiB	14.1 GiB	stdout stderr

对于更完整内容的解析,可以参考Spark官方文档,Spark UI解析。

YARN UI

通过平台侧查询作业Application ID,通过YARN UI查询,点击具体的作业链接,可以查询到具体的作业信息。点击Logs链接,可以看到具体的container日志,如下图:

	Appi	ication application_1651907952651_2364	+
lication			
		Application C	workiew
	User:	hakee	
	Application Type:	generals_stratest	
	Application Tags:	1499201184838857,4j=c52462458773682,xpars,4j=c524682459776882_0/j=c2278858343685c2	
	Application Priority:	0 Higher Integer value indicates higher priority)	
	YawApplicationState:	FINSHED	
	FinalStatus Reported by AM	MEDAAAA SALCOSFIDED	
	Started	Tee May 10 12:17:11 +0600 2022	
	Elepsed	28mins, 59ses	
	Tracking UPL:	History	
	Disposition	BALLED .	
	Unmanaged Application:	false	
	Application Node Lobel expression:	-Not set-	
	AM container Node Label expression:	ADPAULT, PARITION:	
		Application	Motrica
		Total Resource Preempted: <pre><memory.0, pre="" vcares.0x<=""></memory.0,></pre>	
		Total Number of Non-AM Containers Preempted: 0	
		I deal relation of two containers recentpled: 0 Researce Prevented from Cannet Alternet: I conserved, VCanis Ca	
		Number of Non-AM Containers Preempted from Current Attempt: 0	
		Aggregate Resource Allocation: 1152553538 MB-seconds, 51953 vcore-seconds	
		Aggregate Preempted Resource Allocation: 0 MB-seconds, 0 vcore-seconds	
- entri	00	Search:	
10	angt ID * Started 0 I	Node © Logs © Nodes blacklisted by the opp © Nodes blacklisted by the system	0
met 1651	07952551 2364 000001 Tue May 10 12-17-11 http:	//em. Logs 0 0	
	+0800 2022 work	E Container 日志時	
	9.69	19	
11010	1.90.2.7996	The Device 1. Nor	Last
		Frit Parka (1) for	Last
The		The Protect 3 for	Last
1		Tre Jone 1 De	Looped i
Ication	Log Tipes staff Log Tipes staff	Pro Point 1 for	Laar
Fration at	Log Type tofer Log Type tofer Log Log/15/06/19 Time Tue May 10 10-06:14 + 6600 302	r dana tak	Logped I
Reation at	Co Tips: Islan Co Tips: Islan Co Lips: Islan Co Lips: Statistic Co Lips: Statistic Figure 1 (Statistic) Figure 1 (Statistic) Figure 1 (Statistic)	n er annan i ber Jan Maa	Last
ication at	Log Type: Haler Log Lingth State Time: The May 10 15-4614 + 0000 3020 Lingth State Time: The May 10 15-4614 + 0000 3020 Lingth State May 10 15-4614 + 00000 3020 Lingth State May 10 15-4614 + 0000 3020 Lingt	The formula 1 for The Mag The Mag Th	Logged is
ication d	Configuration	Tre Annue 1 Mar M M Ma M M Ma M M Ma M M M M M M M M M	Loop Looped N
Reation 41	Log Specific Log Specific Log Specific Log Specific Log Logics (2000) Print-Internet J. promotion Print-Internet J. Promotion Print-Internet J. Promotion Print-Internet J. Promotion Print-Internet J. Promotion Print-Internet J. Promotion Print-Internet J. Print-Print Print-Internet J. Print-Print Print-Internet J. Print-Print Print-Internet J. Print-Print Print-Internet J. Print-Print Print-Internet J. Print-Print Print-Internet J. Print Print-Internet J. Print Print Print-Internet J. Print Pr	r M Ma M Ma M Ma M Ma M Ma M M Ma M	Lagged i
Section at	Log type after Log ty	The Annual 1 Mar M MM W MM	Logged i
ication at	Log Jone The Mark Control (1996) 14 - 6000 520 Log Jones There Mark Control (1996) 14 - 6000 520 Mark Control (1996) 14 - 6000 520 M	The Annual 1 Mar Per Mag Per	Lagged 1
Reation at	La grand a series a s	2 M Ma M Ma M Ma M Manual Mathematical Statistics (MM (1, 1, 100), 1) (1,1,1,100), 400 Manual Andre Jacobs (ma, 1m) Appendix (MM (1)) (MM (1)) M Manual Mathematical Statistics (MM (1)) (MM	Logged i Logged i
Reation at	The second secon	The Annual 1 Mar A Ma	Logged is
Reation at	La gran construction of the second se	Fig. Annual 1 Mar. Fig. Manual 1 Mar. Fig. Mar. </td <td>Last</td>	Last
Beation at	Compared and the second and the	2 MM MM MM MM MM MM MM MM MM MM	Logged 1
Beation at	Constraints Constrain	The Annual Links	Last
Beation at		A MO	Lagged i
Reation at		2 3 3 4	Lagged 1
Scation Scation		The Annual 1 Mar Prof Ma Prof Ma Pro	Long of the second seco
Reation at		A Main	Loss . Logged)
Reation at		1 1 1 1 1 1 1 1 1 1 1 1 1 1	Loged i
Reation as		The Annual Labor The Annual L	Logged i
Reation as		Mon M	Logged &
anton tor		2 2 3 4 5 5 5 5 5 5 5 5 5 5	Logged 1
Reation at s		Port Man	Logged 1
in the first of th		A A	
Leation at		2 MANNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNN MANNNNN MANNNNN MANNNNNNNNNN	
Region of the local			ыла Саррей м. Болд, Ф.(м. 11)11627 1982 - Сал, 6.47 4 4 616-01 1982 - Сал, 6.47 4 4 616-01
al to tot		A A	
Region of the local sector		2 Monte a la la constante a la con	
Reation		<pre>Process Table Process Tab</pre>	
Ration s		A second se	Lager 1 Lager 1 m. lag. 0) et 11101

指标查询

YARN队列资源指标

Dat abricks数据洞察中YARN组件采用的是Capacity Scheduler,当用户访问Scheduler页面时,可以看到整个集群队列的资源状态,对于单一作业的资源指标状态也可以查询到,如下图:

u: Ca	spacity Used Use	ed (over	capacity)	Max C	apacity	Us	iers Reque	sting Reso	urces									
ae: root																hand when		
uouo: defa	uit) (1.1% used		
																'dı	efault' Qu	Jeue S
						Que	e State:	RUNNIN	3									
					<i>c</i>	Used C	apacity	0.1%										
				0	onfigur	ingured C	apacity:	100.0%										
					Absolu	te Used C	apacity:	0.1%										
				Absolu	rte Cor	figured C	apacity:	100.0%										
			AI	solute C	onfigu	ed Max C	apacity:	100.0%										
						Used Re	sources:	<memor< td=""><td>y:896, w</td><td>Cores:1></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></memor<>	y:896, w	Cores:1>								
			Configur	ed Max A	ation h	ion Mast	er Limit:	25.0		9 uConor:								
			Us	ed Annlic	ation h	laster Re	sources:	smemor	v:896 v	Cores:1>								
			Max Applic	ation Mas	ter Re	sources F	Per User:	<memor< td=""><td>y:18124</td><td>8, vCores:</td><td>></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></memor<>	y:18124	8, vCores:	>							
				Num Se	chedul	able Appl	ications:	1										
			Nu	m Non-Se	chedul	able Appl	ications:	0										
						Num Cor	ications:	10000										
				Ma	x Appl	cations P	Per User:	10000										
			Configu	red Minir	num U	er Limit	Percent:	100%										
				Confi	gured I	Jser Limi	t Factor:	1.0										
					Access	Ible Node	e Labels:	Electron and	inolor" -									
						Pres	g Policy:	disabled	ingroue	y								
				I	ntra-o	Jeue Pres	emption:	disabled										
				Default	Node	Label Exp	pression:	<defau< td=""><td>T_PART</td><td>ITION></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></defau<>	T_PART	ITION>								
				Defa	oult Ap	plication	Priority:	0										
e Users 1	nto																	
er Name	Max Resource		weight Used	Resource			чах АМ К	source		Used	AM Resou	rce		medula	spie App	s Non-	Scheduk	1010 A
																_		
) 	- -						A11 A	nnlic	ation	16							Lapped	() and
ha	daaja						AII A	pplic	atior	าร							Lapped	t in and V
ha	Custer Metrics						AII A	pplica	atior	าร							Lapped	tin and Q
ha	Cueter Metrics Ages Submitted Ages Pe	nding	Apps Running	Apps Co	npietad	Contain	All A	pplica	atior	1S Memory To	ai Mer	nory Reserv	ed VC	ares Use	d VČ	lores Tatal	Lapped	t in as: c
	Cluster Metrics Apps Bulmitted 6665 0	eding	Apps Running 1	Apps Co 9984	npieted	Contain 1	All A	pplic:	atior	Memory To 708 08	ai Mer 08	nory Reserv	ed VC	ares Use	d V(iones Tatal	Lopped VCorer 0	t in and c
line	Cluster Metrics Acces Baltrated Gese Nodes Metrics	eding	Apps Running	Apps Co 9984	npietad	Contain 1	All A	pplic:	atior	Nemary To 708 08	al Mer 0.8	nory Reserv	ed VC	pres Use	d V(384	iones Tatal	Lagged VConer 0	t in and Co s Reserv
	Cluster Metrics 6965 O Cluster Nodes Metrics 4000 Floodes Metrics	nding Decor	Apps Running 1	Apps Co 8984	npieted	Contain 1	All A	pplic:	atior y Used	Nemary To 708 08	al Mor 0.8 Uniteditiva	nory Reserv	ed Vi	ares Use	d Vč 384 Nodos	cores Tatol	Logged VCorer 0	t in as: C s Reserv
	Cluster Metrics Acces Saurmad 566 0 Cluster Nodes Metrics Active Nodes 8 Chodeler Metrics	nding Decor	Apps Running 1 renissioning Nodes	Apps Co 8984	npleted	Contain 1 Decommission	All A	pplic:	v Used Lost Nod	Memory To 708 08 les Q	al Mo 0.8 Uniteditiv/	nory Reserv Nodes	ed VC 1 2	ares Use	d V(384 Nodos	tones Tatal	Lagged VConer 0 Stutistion 7	t in aus c s Reserv
	Cuter Metrics Arcs Rainthol Unter Metrics Cuter Moder Metrics Arche Indone Be Schen Metrics Schen Metrics	nding Decor	Apps Running 1 Scheduldrug Nodes	Apps Co 8984	mpleted	Contair 1 Decommission	All A	pplica energy o	atior v Used Lost Nod	Nerrary To 705 08	al Mor 0.8 Unheathy1	nory Reserv Notes	ed VC 1	ares Use	d V(384 Nodos	anter developed	Logged VConer 0 Shutdown 7	t in auc C s Reserv Nodes
	Cluster Metrics Ages Rainman Ages Pe View Nodes Metrics Custor Prodes Schoolar Metrics Dockater Tomo Dockater Tomo	Decor	Apps Running 1 somissioning Nodes Scheduling Resc rers)	Apps Co 6994	npieted	Dontair 1 Decommission Man memory S2, 1	All A ners Running neel Nadas marr. Alasata Garasts	pplica ene Merro o	y Used Lost Nod	Memory To 708 08 0 Machin, pt 223322, vCon	al Mer 0.8 (Intentity) m Alocation m 525	nory Reserv Notes	ed VC	pros Use lobooted M	d VC 384 Nodos aximum Cli	a ster Applicat	Logged VConer 0 Shutdown 7 Son Priority	t in and G
r aktela Korsa X. Jakobal Milittico Milittico Schilor Schilor Schilor Schilor Schilor Schilor Schilor Schilor Schilor Schilor Milittico	Caster Metrics Area Rainmeil Apea Po Hess 0 Caster Nodes Metrics Reine Nodes Metrics School Winnics School Winnics School Winnics School Winnics	nding Decor	Apps Running 1 Scheduling Resc Thry	Apps Co 5994	nploted	Contain 1 Decommission Man memory 32, y	All A ners Running med Niedes imum Alexant Corrects	pplic:	ution y Used Lost Nod	Memory To 708 08 0 Machine #1205322, #Com	al Mor 0.8 Uniteotity) m Aboation s:S2>	nory Reserv Nodes	ed VC	ores Use tobooted M	d VC 384 Nodes accinum Ch	ores Tatal Q ster Apploat Dear	Lopped VConer 0 Shutdown 7 Ion Priority toth pe	t in and Q
Abels Korns X X X X X X X X X X X X X	Custor Metrics Area Names Area No 1995 - Custor Nodes Metrics Custor Nodes Metrics Scholar Metrics Scholar Metrics Dischar Teor Caperity Scholar Teor Zene 2 - jennes	nding Decor	Apps Running 1 minissioning Nodes Scheduling Reso rev)	Apps Co 9994	nploted	Contain 1 Decommission Man memory 32, 1	All A nors Familing need Nicoles former Allocate Garactiv	pplic: Memo eee Me 0 ton	ution y Used Lost Nod	Memory To 705 08 0 Massim, r123832, #Con	al Mer 0.8 University 1 m Alexandre 1225	Notes	ed VC	ores Use Notes of M	d VC 384 Nodos soinum Cà	a a Ster Apploat	Logged VCorrer 0 Shutdown P Ion Priority this ge	s Reserv Nodes
Abdis Abdis Marriero Abdis Marriero Abdis	Custer Merrice Messi Annotes Active Nodes Metricos Active Nodes Metricos Active Nodes Metricos Counter Volta-Anno Brens - uneres Total Active Notes Brens - uneres Brens - uneres	nding Decor Datavo	Apps Running I Scheduling Nodes Scheduling Reserved Name 1	Apps Co 9994 I wroe Type 2 Application Type 2	Carees 0	Contain 1 Decommission Memory 32, v Application Plication	All A ners Running neel Neeles formen Aleces Corrects	pplica Memo ese Me 0 Ion Reist/Time	v Used Lost Nod orientary State 0	Memory To 768 08 0 Mexim, 123532, «Con FinalCitatus	al Mer 0.8 University/ m. Aboration s:S2> Parming Containery	Nodes Allocated CPU	ed VC 5 0 Allocated Memory	ores Use Motoched M	d V(384 Nodes sximum Ca Sk of Chusher	orres Tassi 0 star Apploat Progress ⁰	Logged VCorrer 0 Shutdown 7 Shutdown 7 Shutd	t in ant C s Reserv
Abda Abda Social Sector Sector Sector Sector Adv Adv Adv Adv Adv Adv Adv Adv Adv Adv	Custor Merica 2 Custor Merica Custor Model Custor Model	Decor Decor	Apps Running 1 Schebuling Nodes rer] Name 2 Name 2	Apps Co 9994 arce Type Application Type 2 Application	Gueve of the first of the first	Contain 1 Decommission Man memory 32, v Application Pharty 0	All A nees Running neel Nados fores:15 StartTime 2 Ta Mer	pplic: Memo ese MB D Ion FinishTime T To Mero	atior y Used Lost Nod ottertor State 0	Memory To 708 08 0 Maskey 120552, 4Con FinalStatus 2 2007FT095	al Mec 0.8 Unhealthy) m Alocation e:22> Parming Containers Containers	Allocated CPU VCres	ed VC	ares Use totootori Mi Guese 0 0	d VC 384 Nedos aximum Câ Ouster 0 0.0	a a ster Accion Progress 0	Logged VCorer 0 Studieven P ton Presity this ge Tracking U = 2	I In an: C Is Reserv Notes
abala dora tora tora tora tora tora tora tora t	Caster Noder Metrics Ages Sammer Ayea Pe Usater Noder Metrics Disability of the Sammer Schoduler Works Disability Table Disability Table Disability Table Disability of the Sammer Metrics Disability of the Sammer Disability	Decor Decor User 1 hedoop	Apps Running ministenting Nodes Bohesking Rese reg Name 2 generate, dataset	Apps Co 9984 arce Type Apploation Type 2 SPARK	Guess default	Contain 1 Decommission Merrory 32, y Application Plicating 2 0	All A ters Purning neel Nades from Allocat Corrects StartTime Tue May 30 125/511	pplic: Memo eee we PrisstTime 5 Tue May 10 12-46-10	v Used Lost Nod orternory State * FNISHED	Namery To 778 08 0 Nations Fractional Succession	al Moo 0.8 Unhealthy) m Alocation es22> Parming Containers Containers Containers	Alocated CPU VOres I N/A	ed VC 3 D Allocated Memory MB 0 N/A	ares Use hotosted M Si of Gana S	d VC 384 Nodos soinum Ca Si of Custor o 0.0	a aner Apploat Progress *	Logged VCorer 0 Shutdown 7 Ion Proofby this ge Tracking UI 7 History	I III III: C II III III: C III IIII III: C IIII IIII IIII IIII IIIIIIII IIIIIIII
abala dora secondaria secondaria secondaria secondaria der	Cuerto Merica Aces Names Apel Ne Horison Aces Ne Scalar Marcia Carlos Ma	Decor Decor User ¹ hedoop	Apps Running I Scheduling Nodec Scheduling Rese Ingenerate, dataset	Apps Co 5984 arror Type Application Type 2 55%RK	Queon o default	Contain 1 Merimanoy X2, v Application Planty 2 0	All A nors Running nod Nodos former Allocat Corrects 2 StartTime 2 Tas May 10 122/211	pplic: Memore exerve 0 FinishTime 2 Tue Ney 10 T2-MEM	v Used Lost Nod orremon State 0 FNISHED	Menany To 706 08 0 Meniny FINASSE	al Mer 0.8 Unicettry 1 m Alexandry 1 Parenting Containers 0 N/A	Allocated CFU VCores I N/A	ed VC 3 9 0 Allocated Memory NMS 0 N/A	tores Use Noteshed Market 0 0.0	d VC 384 Nedes aximum Ch Chuster Cluster C	ater Applicat Progress ?	Logged VConer 0 Shutdown 7 Ion Priority this ge Tracking UI 0 History	I III III C C III III III C C III III II
ateria ateria form form form form form form form form	Counter Marcine Aces In Marcine Boundary Counter Marcine Counter Marcine Marcine Counter Marcine Marcine Schwaler Ma	Decor Decor User * hedoop	Apps Running 1 ministranting Nodes Boheskulning Nodes Reng Name 1 paraminis, dataset	Apps Do 9984 since Type Application Type 2 SPWRK	Queue o construction of the second of the se	Contain 1 Decommission Man memory 22, 1 Application Philarby [©] 0	All A neers Running med Niedes corrects Corrects Tue May 10127211 -02600 2002	D	v Used Lost Not crientry Date 0 FNISHED	Memory To 708 08 0 Machine 1205032, «Com Finalizes.a. 5 SUCCEEDED	al Mee 0.8 Unhealthy) en Alocation er25 Parning Containers 0 N/A	Alocated CPU VCores I N/A	nd VC	ones Use toboosted M Galees 0.0	d VC 384 Nodes Self Cautor 0.0	0 aster Applicet Progress 0	Logged VCorer 0 Shutdown 7 Shutdown 7 Shutdown 7 Shutdown 7 History	Reserv
abala doran Koran X. Savena METTED METTED METTED METTED METTED METTED METTED METTED METTED METTED METTED METTED	Court Mints Area Barristo Met Mantal Met Mantal Met Mints Met Mints Barristo Ba	Decor Decor Liser [©] hadoop	Apps Running I Innisistanting Nodes Scheduling Rese Instrum In	Apps Co 9984 1 urce Type 2 SPARK SPARK	Queue C default	Contain 1 Seconventises Main Plantay 2 0 0	All A ness Pauming neel Neeles corrects StartTime of 122521 10122521 20022 Nan May 9	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	vy Used Lost Nod State 0 FINISHED	Memory To 708 08 0 Mashin, r100532, «Con Finalization guocettobo	al Moo 0.8 University/ ne Alexandre Containens N/A N/A	Alocated CPU VCores I N/A	nd VC 1 D Allocated Memory MB 0 N/A	Intes Use Intervented M 0.0 0.0	d VC 384 Nodes Skot Duster 0 0.0	a a ster Apploat Progress *	Lopped VCorer O Shutdown 7 Shutdown 7 Shutdown 7 Shutdown 7 Lintory History	I in an: C S Perserv Notes 0 0
abela dozna zerito zerito zerito de de de	Cuer Write Courte Write Courte Write Courte Wook Marco Ante Index State Index Bookan D State Index Index Index Index Index Index Index Index Index Index Ind	Inding Decor (MEXAD Liser ® hedoop	Apps Running 1 30-bed-afrag Nodes wrvj Nama 1 paramata, dataaet	Арря Со 9984 игов Таро Аррасеван Тура 2 5РУИК 5РУИК	Queue o default	Contain 1 Neorministo Man Peranty 22, 1 Pisarty 2 0	All A ters Burning need Nodes former Allocat Corrects Too Nay 10 127211 - 40600 2022 Non May 9 0020041	Pplic: Memo 806 MB 0 805 MB 0 10 10 10 10 10 10 10 10 10 10 10 10 1	ation y Used Lost Ned State © FNESHED	Memory To 768 08 0 Mexim, 129532, «Con Finalizes, 5 SUCCEEDED	a Mer 0.8 Unhealthy I in Aboration 525 Parming Containers 0 N/A N/A	Alocated CPU VCores I N/A	ed VC	ores Use bloconed M Garage 0 0.0	d W 384 Nodos 56 ef Cluster 0 0.0	a ater Applicat Progress *	Lopped V/Doner 0 Stutitionn 7 Stutitionn 7 Stutitionn 7 Statistic UI 2 History	I in an: C Is Reserv Notes 0 0
Alexandra Alexan	Control Contro	Decor Decor User ® hadoop	Apps Running I Indestanting Nodes International Parameters International International International International International International International International International International	Apple Do 9994 arrow Type Type 2 SPMRK SPMRK	Queue c default	Contain 1 Decommission Man Pelantay 0 0 0	All A nees Running med Nodes merer Alexads StarrTime 2 Tue May 50 12:15:11 -0600 2022 Man May 9 00:30:41 -0600 2022	Pplic: Memo 866 MB 0 10 10 10 10 10 10 10 10 10	atior y Used Cost Nod State © PNS9-ED PNS9-ED	Atemory To 708 08 0 Markin, FRANSSA, «Con FRANSSA, «Con SUCCEEDED SUCCEEDED	al Mer 0 8 Uneothy 1 m Alocation s S25 Parning Containers 0 N(A	Allocated CFU N/A	ed VK 1 D Allocated Memory N(A	Interes Use Million One One One One One One One One One On	d VC 394 Nodos Si ef Coutor 0 0.0	a a ster Apploat Progress 0	Loged VCore 0 Stutizen / Tradig U 2 Elstor Listor	B Reserv Noties Black Noties 0 0
r r N. MANDA N. MANDA AND	Control Montrol Control Montrol Contro	Decor Decor Loser 0 hadoop hadoop	Apps Running I Schebuhrig Nodes I Name I parente, dataet generate, dataet	Apps Co 6984	appieted Queues default default	Dontain 1 Decommission Memory 32, v Priority 9 0 0 0	All A ners Running nert Nodes cores:1s StartTime 2 StartTime 2 StartTime 2 StartTime 2 StartTime 2 StartSin 012:5:51 1 :02002 StartSin 012:5:51 1 :02002 StartSin 012:5:51 1 :02002 StartSin 012:5:51 1 :02002 StartSin 012:5:51 StartSin 012:5:51 1 :02002 StartSin 012:5:51 StartSin Sta	Denormality (1997)	ation vy Used Lost Nod State 0 FN45HED FN45HED	Memory To 768 08 0 Machine 1233532, 4Con Readings 9, 4Con 19, 4Con	al Mer 0.8 Unhealthy? In Alexandry Containers 0 N/A N/A	Allocated CFU VCores II N/A N/A	ed VC	Interes Use Interested M Galaxie 0.0 0.0	d VC 364 Nodes Nodes Sk of Coulor 0 0.0 0.0	anter Acceloret	Logod VCore 0 Shuttown fr Ion Priority U 2 Litator History	I In an: C S Reserv Modes Black Nodes 0 0 0
Abria Abria Marina Mari	Course Service	nding Decor User 0 hedoop hedoop	Apps Running I ministening Nodes Scheduling Nese Ing Name 2 generate, dataset generate, dataset	Aces Do 9994 arce Tupo Acedeation Type 2 SPARK SPARK SPARK	Queos Constant Constant Constant Constant Constant	Contain 1 Natur Manufactor Starty 7 0 0	All A wes Running med Nodes corrects 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Dentification (Control of Control	ation y Used Lost Nod Same 0 Same 0 Findaneo Findaneo	Atenary To TP8 08 0 Maritm P 129332, «Con FruidPata Succession Succession Succession	at Mee 0.8 Univestity 1 in Abocation e 22> Parning Consilience 0 N/CA N/CA	Allocated CFU VCores I N/A N/A	ed V/C 3	ares Use betworked bio conserve conserv	d W 394 Nodes Self Caster Cast	a a ster Acchort Progress P	Loged VCOrer 0 Skutstern 7 Trasting U 2 Efficient History	I in an: C In Reserv Notes 0 0 0
r Azərla Xara Xara Xara Xara Xara Xara Xara X	Construction Co	Decor Decor Liser ® hedoop hedoop	Apps Running I Interstanting Notice Sobestung Nese (v) Name I paramin_dataset paramin_dataset	Apps Do 9994 8 wrote Tupe Type 2 SPVARK SPVARK	Queue o default default	Dontain 1 Decommission Memory SC, 1 Priority 2 0 0 0	All A hers Running med Rindes former Allocate StartTime	Pplic: 806 MB 0 806 MB 0 0 10 12-460 -0800 2022 Nan May 8 2022 San Asso 10 12-460 -0800 2022 San Asso 10 12-460 -0800 -0	V Used Lost Nod Date 0 FINISHED	Merrary To 778 08 0 Madra FitadStata SUCCEEDED SUCCEEDED SUCCEEDED	al Motorform 0.8 Unitedativy / Running Constants N/A N/A	Allocated CRU VCores I NVA	ed V/V 1 D D Allocated Memory MB 0 NVA	Interes User Interesting No. 0.0 0.0	d VC 384 Nodos admum C3 0.0 0.0 0.0	a a ster Applot	Loged VCDree 0 Stutteren 7 Tracking Litetor History History	Back Nodes
Antonia Antonia Martino Martino Antonia Antoni	Control Marcin Marcine Control Marcine Marcine Control Marcine Marci	nding Decor Decor Nedoop Nedoop	Apps Running I ministenting Nodes Scheduling Resc Int Name 2 generate, dataset generate, dataset	Арея Со 9994 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Queue default default	Dontari 1 Decomission Million Priority 2 0 0 0	All A hers Running mee Plades meer Rudes meer Rudes start Tag Start Start Start Tag Start Start Start Start Start Start Start Start Start St	Pplic: 000000000000000000000000000000000000	ation y Used Lost Nod State 0 FINESHED FINESHED	15 Menory To 708 08 0 Menime 150532, «Con Finalization 5 5 5 5 5 5 5 5 5 5 5 5 5	al Meedow 0.8 Ckriednitry JU Running Constanting NVA	Allocated CRU Vicens II N/A	nd V/C 3 D Allocated Memory NUA NUA	Notes Use Notes the State Contest 0.0 0.0 0.0	d W 384 Notes aximum CS Ounter 0 0.0 0.0 0.0	a a Progress 0	Lopped VCOree 0 Stutidisen 7 Tracking 12 Etistore Etistore Etistore	Reserv Reserv Notes 0 0
r abidis total	Contro Marine Neurosciences Contro Marine Neurosciences Securitor Marine Securitor Securitor Securitor Marine Securitor Marin	Decor Decor	Appo Running I Introducting Nodes Scheduling News Internet, Scheduling News generating, Scheduling generating, Scheduling generating, Scheduling	Appe Do 9984 8 Application Type 2 SINAR SINAR SINAR	Queue o construit default default default	Contain 1 Tecomenisation Memory 32, v Pinanty 2 0 0 0 0	All A were Running med Packs front Allocate StartTime 2 StartTime	Pplic: Merca ele Me	ation y Used Lost Nod State 0 FNE9-ED FNE9-ED FNE9-ED	Merrary To 778 08 0 Markin, 123352, 4Con Finalizesa, 5UCCEEDED SUCCEEDED SUCCEEDED	al More 0.8 (khoatiny)	Allocated Allocated CFU VCores II NIA NIA NIA	ed V/C	Inter Use Intervented Million 0.0 0.0 0.0 0.0	d VC 384 Nados Salmum Ca Danter 0.0 0.0 0.0 0.0	a autor Apploat Progress 0	Lagged VCorrect O Studiewn // Kin Presity Litelory Litelory Litelory	Racel 0 0
r abits bits bits different construction different diffe	Control Marine Control Marine	Decor Decor Decor Decor hadoop hadoop	Appa Running Innessarating Nodes Soleda Ang Neser Nama I penemini, dataset penemini, dataset penemini, dataset	Аран Со 9894 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Queue Contraction of the second of the secon	Contain 1 Montematica Management	All A sets Running med Nodes sets Running med Nodes Sets Set New Sets Set New Set Set Set New Sets Set New Sets Set	Pplic: Merco 666 MB 0 0 10 10 10 10 10 10 10 10	ation y Used Lost Ned Same 0 FINESHED FINESHED FINESHED	15 Memory To 758 08 0 Mexim 120302 500000000 500000000 500000000 500000000	More the second	Allocated CPU VCores I N/A N/A N/A	ed VC 3	ares Use blocked M 0.0 0.0 0.0	d VC 384 Nodes aximum C3 0.0 0.0 0.0 0.0	a a star teoloor Progress 0	Logod VCOrec 0 Studiown / Tracking Efficient Efficient Efficient Efficient Efficient	Reserv Receive 0 0 0
r abida Stot	Control Marine Control	Decor Decor Liser ® hadsop hadsop	Agos Running I Innisistaning Nodes Schesking Hene ry Name 1 generate, dataset generate, dataset generate, dataset	Apps Co 9984 8 Application Type 2 SINARK SINARK SINARK	Queue o construction	Dontain 1 Decommission Minimum 232, v Argenitarion Price V 0 0 0	All A errs Running real Ploades errors Running Corrects StartTime 2 Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start Start StartStart	Pplic: Merce 066.46 0 107 108 108 108 108 108 108 108 108	ation y Used Lost Ned State 0 FNSHED FNSHED FNSHED	Nerray To 38 North Control of Con	III More 0.8 Constraints of the	nory Reserv Allocated CRU VCores II N/A N/A N/A	ed VV	000 Use 00000000 000 000 000	d VC 394 Nostos aximum Ci Duntor 0.0 0.0 0.0	0 anter Apphoe Progress 0	Logged UCores 0 Studiewn II Ion Provity Uniter Uniter Ethiony Ethiony Ethiony	Black Notes 0 0 0
A AND	Course Marine Marine Course Automatic Marine Course Automatic Marine Course Automatic Scolare News Scolare Ne	Decor Decor Liser ® hadoop hadoop	Acce Running Innessaring Nodes Boheakay Nodes Innessaring Nodes Innessaring Nodes generate, dataset generate, dataset generate, dataset	Арра 00 9994 1 1 9994 4рралия Тура 2 50 млк 50 млк 50 млк 50 млк 50 млк 50 млк 50 млк 50 млк 50 млк 50 млк	Queue default default default default	Contain 1 Decommission Membrany St., vi Application 0 0 0 0 0	All A hers Running meet Rundes envers Allendes description 2002 2002 2002 2002 2002 2002 2002 20	Pplic: Merco 866 MB 0 0 10 10 10 10 10 10 10 10	atior y Used Lost Ned Same © Same © Finis-ED Finis-ED Finis-ED Finis-ED	Meenay To 00 789.08 Meenay To 00	I Moodeling I Centrality I Rendering Constanting Constanting Constanting Constanting NUA	Allocated CRU VORes I N/A N/A N/A N/A	ed VC 3	2% of Gamer 0.0 0.0 0.0 0.0	d W 384 384 Nodes Solution Co Outer 0 0.0 0.0 0.0	a anter Appleat	Logod VCOren D D D D D D D D D D D D D D D D D D D	Black Notes
r akhila karana ka karana karana karana karana karana karana karana karana kara	Control Marine Marine State Control Control Marine Control Marine Control Marine Schoold	Decor Decor Liser 1 hadoop hadoop hadoop	Ages Furning I Interestanting Nodes Scheduling Nese Prof Name I paramite, dataset paramite, dataset paramite, dataset	Ages CD 9993 8 8 4cpitostim Acpitostim Soviet Soviet Soviet Soviet Soviet Soviet Soviet Soviet	Queue default default default default	Contained Minimum Advanced Minimum Advanced Applications 0 0 0 0	All A hers Running and Roudes former Allowed Correct- StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime (Correct- StartTime) StartTime) StartTime) StartTime (Correct- StartTime) Star	Menne eles Me 0 men ResetTime 2002 Men Men 2002 Men Men 2002 Men 2002 Men 2002 Men 2002 Men Men 2002 Men Men 2002 Men Men 2002 Men Men 2002 Men Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 2002 Men 200 Men 2002 Men 2002 Men 200 Men Men 200 M	atior y Leef Lost Nod strenor PNSPED PNSPED PNSPED PNSPED PNSPED PNSPED PNSPED	Nemov To 776 08 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	at Mode of Sector Secto	Alcosted CPU N/A N/A N/A	nd VX p p Allocated Memory 0 NA NA NA	1000 Use 1000 000 0.0 0.0 0.0 0.0 0.0	4 W 394 394 Nedes aximum C3 0.0 0.0 0.0 0.0	a atter Apphoe	Logond ViCores O Stuttoren 7 Iour Proutly Usi 2 Litatore Litatore Litatore Litatore Litatore	Base Contraction C

对于这部分指标的解析,可以参考YARN官方文档,参数含义。

jmx的json全类别指标

Hadoop提供了JMX监控接口,可以获取到基于JSON类型的指标数据,其中包含JVM指标、RPC指标、YARN相关指标等等,具体指标含义可参考官方文档,指标详解。

访问方式可以将YARN UI中的链接地址进行拼接修改,具体格式为原域名+/gateway/clustertopo/yarn/jmx

举例如下:

App展示网址: https://knox-c-xxxxx-cn-beijing.databricks.aliyuncs.com:8443/gateway/cluster-topo/yarn/cluster/apps

JMX指标网址: https://knox-c-xxxxx-cn-beijing.databricks.aliyuncs.com:8443/gateway/clustertopo/yarn/jmx

用户可以参考自己的网址进行访问,示例结果如下图:

}, {
 "name": "Hadoop:service=ResourceManager, name=QueueMetrics,q0=root",
 "modelerType": "QueueMetrics,q0=root",
 "tag.Queue": "root",
 "tunning_O": 2,
 "tunning_Id40": 0,
 "AumesourceLiniLVCOres": 0,
 "tag.Queue": 0,
 "Appertunited": 809,
 "AppexQuanting": 2,
 "AppexQuanting": 0,
 "AppexQuanting": 0,
 "AppexQuanting": 0,
 "AppexQuanting": 0,
 "AppexQuanting": 0,
 "AppexQuanting": 0,
 "AppexQuantingTexPresented": 0,
 "AppexQuentingTexPresented": 0,
 "AppexQuantingTexPresented":