阿里云

企业级分布式应用服务 EDAS 应用托管(K8s)

文档版本: 20220708

(一) 阿里云

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。

- 1. 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。
- 2. 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。
- 3. 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。	
△)注意	用于警示信息、补充说明等,是用户必须 了解的内容。	(大) 注意 权重设置为0,该服务器不会再接受新请求。
② 说明	用于补充说明、最佳实践、窍门等,不是 用户必须了解的内容。	② 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 Instance_ID
[] 或者 [a b]	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.应用托管概述(K8s)	11
2.资源管理	14
2.1. K8s资源管理概述	14
2.2. 管理微服务空间	14
2.2.1. 使用控制台管理微服务空间	14
2.2.2. 使用API管理微服务空间	16
2.3. K8s集群管理	19
2.3.1. 使用控制台管理集群(K8s)	19
2.3.2. 使用API管理K8s集群	23
3.创建和部署应用(K8s)	28
3.1. 创建和部署应用概述(K8s)	28
3.2. Java应用	28
3.2.1. JAR包或WAR包部署	29
3.2.1.1. 在容器服务K8s集群中使用JAR包或WAR包部署应用	29
3.2.1.2. 在Serverless K8s集群中使用JAR包或WAR包部署应用	34
3.2.1.3. 使用API在K8s集群中创建Java应用(JAR包或WAR包)	37
3.2.2. 镜像部署	40
3.2.2.1. 制作应用镜像	40
3.2.2.2. 在K8s集群中使用镜像部署Java微服务应用	46
3.2.2.3. 在Serverless K8s集群中使用镜像方式部署应用	51
3.2.2.4. 使用API创建Java应用(镜像部署)	55
3.2.3. 转化Deployment为EDAS应用	57
3.2.3.1. 使用控制台转化Deployment为EDAS应用	57
3.3. PHP应用	59
3.3.1. 从源码构建PHP应用并部署至容器服务K8s集群	59
3.3.2. 如何在使用源码部署应用时配置SSH公钼	66

3.4. 多语言应用	67
3.4.1. 部署多语言微服务应用	67
3.4.2. 使用API在K8s集群中创建多语言应用	76
3.5. 高级配置	78
3.5.1. 配置调度规则	78
3.5.2. 配置启动命令	82
3.5.3. 配置环境变量	85
3.5.4. 配置持久化存储	91
3.5.5. 配置本地存储	96
3.5.6. 配置应用生命周期的钩子和探针	99
3.5.7. 配置日志收集	105
3.5.8. 配置Tomcat	108
3.5.9. 配置Java启动参数	110
3.5.10. 配置挂载	114
3.5.11. 配置微服务治理	117
3.5.12. 配置同可用区优先路由	127
3.5.13. 实现K8s集群应用的限流降级	129
3.5.14. 配置标签	131
3.5.15. 配置注解	132
4.升级和回滚应用(K8s)	134
4.1. 升级和回滚应用概述(K8s)	134
4.2. 单批发布(K8s)	134
4.2.1. 使用控制台单批发布应用(K8s)	135
4.2.2. 使用API在K8s集群中单批发布应用	136
4.2.3. 使用Cloud Toolkit插件单批发布应用(K8s)	141
4.3. 分批发布(K8s)	147
4.3.1. 使用控制台分批发布应用(K8s)	147
4.3.2. 使用API在K8s集群中分批发布应用	150

4.3.3. 使用toolkit-maven-plugin插件分批发布应用(K8s)	156
4.4. 金丝雀发布(K8s)	164
4.4.1. 使用控制台金丝雀发布应用(K8s)	164
4.4.2. 使用API在K8s集群中金丝雀发布应用	168
4.5. 回滚应用(K8s)	174
4.5.1. 使用控制台回滚应用(K8s)	174
4.5.2. 使用API在K8s集群中回滚应用到目标历史版本	176
4.5.3. 使用API在K8s集群中终止应用变更并回滚	178
5.CI/CD	181
5.1. CI/CD概述(K8s)	181
5.2. 云效	181
5.2.1. 使用云效2020进行持续集成和部署(K8s)	181
5.2.2. 使用云效2020实现其他类型应用的持续集成和部署(K8s)	191
5.3. Jenkins	194
5.3.1. 使用Jenkins创建持续集成	194
5.3.2. 在Jenkins中使用EDAS插件构建应用部署的任务	200
5.3.3. 在Jenkins中使用edas-jenkins-plugin部署应用的参数配置	206
6.应用监控	212
6.1. 应用监控概述	212
6.2. 监控	213
6.2.1. 查看应用总览	213
6.2.2. 应用故障自动诊断	216
6.2.3. 查看Prometheus监控指标	220
6.2.4. 应用实例监控	221
6.2.4.1. 概览	221
6.2.4.2. JVM监控	224
6.2.4.3. 主机监控	226
6.2.4.4. Pod监控	227

6.2.4.5. SQL调用分析	230
6.2.4.6. NoSQL调用分析	231
6.2.4.7. 异常分析	232
6.2.4.8. 错误分析	233
6.2.4.9. 上游应用	234
6.2.4.10. 下游应用	236
6.2.4.11. 调用链查询	238
6.2.4.12. 日志	239
6.2.4.13. 内存快照	240
6.2.5. 服务和接口监控	247
6.2.6. 数据库调用监控	250
6.2.7. NoSQL调用	253
6.2.8. 外部调用	254
6.2.9. MQ监控	255
6.2.10. 应用诊断	256
6.2.10.1. 实时诊断	256
6.2.10.2. 异常分析	258
6.2.10.3. 线程分析	259
6.2.10.4. Arthas诊断(新版)	261
6.2.10.5. 日志分析	267
6.2.11. 应用环境	268
6.2.12. 应用监控设置	269
6.2.12.1. 自定义配置	269
6.2.12.2. 标签管理	273
6.2.12.3. 监控方法自定义	274
6.2.12.4. 自定义参数	275
7.告警管理	278
7.1. 告警管理升级说明	278

	7.2. 名词解释	280
	7.3. 联系人管理	282
	7.3.1. 联系人	282
	7.3.2. 联系人组	284
	7.3.3. 钉钉机器人	285
	7.3.4. 企业微信机器人	286
	7.3.5. 通过Webhook自定义告警通知人	287
	7.3.6. 获取钉钉机器人Webhook地址	292
	7.4. 通知策略	295
	7.5. 升级策略	298
	7.6. 应用监控告警规则(新版)	299
	7.7. 在钉钉群中处理告警	302
	7.8. 查看告警发送历史	310
	7.9. 查看告警事件历史	315
	7.10. 创建应用监控告警并发送告警通知	317
8	.应用运维	323
	8.1. 应用运维概述	323
	8.2. 管理应用生命周期	324
	8.3. 访问应用(发布服务)	325
	8.3.1. 访问应用概述	325
	8.3.2. 应用路由选型	326
	8.3.3. 添加负载均衡SLB	327
	8.3.4. 复用负载均衡SLB	331
	8.3.5. 添加服务Service	332
	8.3.6. 创建应用路由(Nginx Ingress)	335
	8.3.7. 创建应用路由(ALB Ingress)	339
	8.3.8. 使用SLB访问K8s上的EDAS应用	343
	8.4. 应用扩缩容	348

8.4.1. 手动扩缩容	348
8.4.2. 自动弹性扩缩容	349
8.5. 限流降级	354
8.5.1. 限流降级简介	354
8.5.2. 限流降级原则	355
8.5.2.1. 限流降级原则概述	355
8.5.2.2. 服务提供方或消费方流控	356
8.5.2.3. 削峰填谷	358
8.5.2.4. 预热启动(Warm Up)	359
8.5.2.5. 关联限流	360
8.5.2.6. 弱依赖降级	361
8.5.3. 配置规则	362
8.5.3.1. 配置流控规则	362
8.5.3.2. 配置降级规则	365
8.5.3.3. 配置热点规则	367
8.5.4. 实现K8s集群应用的限流降级	369
8.5.5. 查看限流降级操作日志	372
8.6. 变更记录	372
8.6.1. 查看应用变更	372
8.6.2. K8s集群应用部署失败的原因分析	375
8.7. 事件中心	376
8.7.1. 事件中心	376
8.8. 日志管理	380
8.8.1. 查看实时日志	380
8.8.2. 查看日志目录	381
8.8.3. 查看文件日志	384
8.9. 应用管理	384
8.9.1. 导出应用	384

8.9.2.	导入应用	38
8.9.3.	克隆应用	 386

1.应用托管概述(K8s)

EDAS支持以容器的形式托管应用到阿里云Kubernetes集群或混合云集群(其他云域或IDC内自建Kubernetes集群),您无需理解容器服务底层细节,便可以低门槛拥抱容器技术,最大化地利用资源并完成应用在容器里的全生命周期管理,以及监控、运维等服务。本文介绍K8s环境支持的应用类型和应用托管功能。

K8s环境类型

EDAS不仅支持在阿里云上使用Kubernetes管理容器应用,也支持使用在混合云集群(其他云域或IDC内自建 Kubernetes集群)管理容器应用。



EDAS支持的类型

阿里云Kubernetes集群

混合云集群

● 阿里云Kubernetes集群

EDAS无缝对接阿里云容器服务Kubernetes版,支持导入在容器服务Kubernetes版控制台创建的专有版 Kubernetes集群、托管版Kubernetes集群和Serverless Kubernetes集群。

混合云集群(其他云域或IDC内自建Kubernetes集群)
 在日常运维过程中,存在同时在云上和IDC里拥有多个Kubernetes集群,或者同时拥有多个云域上的
 Kubernetes集群的情况。在这种现状下,您可以将这些集群注册到ACK控制台,然后再导入到EDAS。

? 说明

混合云集群、专有版和托管版Kubernetes集群导入到EDAS中对应的是容器服务K8s集群; Serverless Kubernetes集群导入到EDAS中对应的是Serverless K8s集群。

更多阿里云Kubernetes集群和混合云集群的相关信息,请参见集群纳管。

K8s环境支持的应用类型

EDAS支持以容器的形式托管Java应用、PHP应用以及多语言(包含Node.js、Go和Python等多种语言)应用。



EDAS支持的托管应用类型







● Java:支持JAR包、WAR包以及镜像部署,不同的部署方式需选择的应用运行环境也会有所不同。

● PHP: 支持从源码构建PHP应用并托管到EDAS。

● 多语言:支持使用自定义的镜像将多语言应用托管到EDAS。

更多托管应用类型的部署方式以及运行环境的信息,请参见创建和部署应用概述(KBs)。

K8s环境支持的应用托管功能

您可以将应用托管到阿里云Kubernetes集群和混合云集群(其他云域或IDC内自建Kubernetes集群),并使用EDAS为您提供的一系列应用托管能力,包括资源管理、应用部署、CI/CD、应用监控和应用运维等能力。





使用场景	功能
资源管理	EDAS是面向应用的PaaS平台,为应用提供ECS和K8s环境及相关资源。EDAS的K8s环境中主要包含命名空间和集群纳管的资源管理功能。更多信息,请参见K8s资源管理概述。
应用部署	 创建和部署(首次部署):完成应用开发和测试后,您可以在Kubernetes集群中创建并部署应用。更多信息,请参见创建和部署应用概述(K8s)。 升级和回滚:应用在EDAS中创建并部署后,还会不断迭代,需要升级。如果升级的应用版本发现问题,需要将应用回滚到历史版本。更多信息,请参见升级和回滚应用概述(K8s)。
CI/CD	应用在不断迭代过程中,需要持续集成(CI)和持续部署(CD)。EDAS支持通过 <mark>Jenkins和云效</mark> 对您部署的应用进行CI/CD。具体操作,请参见CI/CD概述(K8s)。
应用监控	EDAS集成了应用实时监控服务ARMS,为部署在EDAS中的应用提供关键指标的监控能力和报警能力。更多信息,请参见 <mark>应用监控概述</mark> 。
应用运维	EDAS为应用提供了一系列运维功能,包括负载均衡、应用扩缩容、限流降级、变更记录、事件中心、 日志管理等。更多信息,请参见 <mark>应用运维概述</mark> 。

2.资源管理

2.1. K8s资源管理概述

EDAS不仅支持使用阿里云Kubernet es集群管理容器应用,也支持使用混合云集群(其他云域或IDC内自建集群)管理容器应用。如果您对应用的运行环境有安全隔离的需求,可以创建并使用微服务空间,实现资源和服务的隔离。

微服务空间

微服务空间为应用提供互相隔离的运行环境,如开发、测试和生产环境等,您可以使用微服务空间实现资源 和服务的隔离。

EDAS默认提供一个默认微服务空间,也支持自建微服务空间。

集群纳管

EDAS不仅支持使用阿里云Kubernetes集群管理容器应用,也支持使用混合云集群(其他云域或IDC内自建集群)管理容器应用。您只需要将K8s集群导入到EDAS控制台,无需其他复杂操作,便可以直接托管应用到K8s集群。

EDAS不提供创建K8s集群的跳转入口,您需要到<mark>容器服务管理控制台</mark>创建阿里云集群或者注册混合云集群,容器服务通过将Kubernetes网络和阿里云VPC的深度集成,提供了稳定高性能的容器网络,不同VPC逻辑上完全隔离。

● 阿里云Kubernetes集群

EDAS无缝对接阿里云容器服务Kubernetes版,支持导入在容器服务Kubernetes版控制台创建的专有版 Kubernetes集群、托管版Kubernetes集群和Serverless Kubernetes集群。3种集群形态的对比详情,请参见什么是容器服务Kubernetes版。

- 创建专有版Kubernetes集群的具体操作,请参见创建Kubernetes专有版集群。
- 创建托管版Kubernetes集群的具体操作,请参见创建Kubernetes托管版集群。
- 创建Serverless Kubernetes集群的具体操作,请参见ASK使用快速入门。

专有版Kubernetes集群和托管版Kubernetes集群导入到EDAS中对应的是容器服务K8s集群,Serverless Kubernetes集群导入到EDAS中对应的是Serverless K8s集群。

● 混合云集群

在日常运维过程中,存在同时在云上和IDC里拥有多个Kubernetes集群,同时拥有多个云上的Kubernetes集群的情况。在这种现状下,您可以将这些集群注册到容器服务Kubernetes版控制台,实现统一的安全治理、应用管理和监控、日志等能力,降低运维成本。

注册外部集群到ACK控制台的具体操作,请参见创建注册集群并接入本地数据中心集群。 混合云集群导入到EDAS对应的是容器服务K8s集群。

② 说明 在阿里云Kubernetes集群或混合云集群托管应用原理类似,都是将应用托管到一个Pod或一组Pod。

2.2. 管理微服务空间

2.2.1. 使用控制台管理微服务空间

微服务空间为应用提供互相隔离的运行环境,如开发、测试和生产环境等,您可以使用微服务空间实现服务的隔离。在同一个地域内不允许创建两个同名的微服务空间。

应用场景

例如,您在同一个VPC内有三套环境,分别用于应用开发、测试和生产,可以为这三套环境分别创建一个微服务空间(如Dev、Test和Prod)。不同微服务空间下的应用和服务是相互隔离的,不能进行服务调用和配置推送。

默认微服务空间

在EDAS创建应用时,可以选择**默认**微服务空间,默认微服务空间即为没有微服务空间的状态,表示该应用不使用任何微服务空间来实现服务的隔离。

在应用列表内选择某地域下的默认微服务空间时,即表示筛选出来的应用都没有归属于任何微服务空间。

创建微服务空间

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中选择资源管理 > 微服务空间。
- 3. 在微服务空间页面右上角单击创建微服务空间。
- 4. 在创建微服务空间对话框配置微服务空间参数,然后单击创建。



参数	描述
注册配置中心	 MSE Nacos: 您购买的阿里云MSE Nacos引擎,可与EDAS无缝结合,实现服务注册及配置管理。 EDAS注册配置中心: EDAS提供的免费服务注册配置中心,若应用对性能及稳定性有比较高的要求,建议使用MSE Nacos作为服务注册配置中心。
MSE Nacos实例	MSE构建的Nacos实例。具体操作,请参见 <mark>创建MSE</mark> Nacos实例。
归属地域	当前微服务空间所归属的地域 <i>,</i> 不可更改。
允许远程调试	当您想对应用进行端云互联时,您在该应用所在的微服务空间的编辑页面手动开启 允许远程调试 。端云调试的相关操作,请参见 <mark>端云互联简介</mark> 。
描述	请输入一段文字来描述微服务空间。

编辑微服务空间

- 1. 在微服务空间列表的操作列单击编辑。
- 2. 在**编辑微服务空间**对话框,您可修改**微服务空间**名称及**描述**,也可开启或关闭远程调试功能。在编辑完成后单击**确定**。

删除微服务空间

删除微服务空间需要满足以下前提条件:

- 微服务空间下无任何集群。
- 微服务空间下无任何微服务应用。
- 微服务空间下无任何ECS实例。
 - 1. 在微服务空间列表的操作列中单击删除。
 - 2. 在确认对话框中单击删除。

2.2.2. 使用API管理微服务空间

本文介绍如何使用EDAS提供的Java SDK调用API创建、编辑和删除自定义微服务空间。

前提条件

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定微服务空间创建的地域,假设为cn-hangzhou。

创建或编辑微服务空间

以下代码用于快速创建或编辑自定义微服务空间。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.InsertOrUpdateReqionRequest;
import com.aliyuncs.edas.model.v20170801.InsertOrUpdateRegionResponse;
public class InsertOrUpdateRegion {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun_user_ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //需要创建或编辑的微服务空间所在地域ID。
       String region_id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       InsertOrUpdateRegionRequest request = new InsertOrUpdateRegionRequest();
       //微服务空间ID,格式为"地域ID:微服务空间标识符",例如: cn-hangzhou:doc。
       request.setRegionTag("cn-hangzhou:doc");
       //微服务空间名称,例如: APIdoc。
       request.setRegionName("APIdoc");
       //微服务空间的描述信息。
       request.setDescription("使用API创建的微服务空间");
       //此参数为空或者为0表示创建微服务空间,其他取值表示编辑微服务空间。
       request.setId((long) 0);
       //是否允许远程调试, true表示允许, false表示不允许。
       request.setDebugEnable(true);
       try {
           InsertOrUpdateRegionResponse response = client.getAcsResponse(request);
           System.out.println("RegionName=" + response.getUserDefineRegionEntity().getRegi
onName() + "\nRegionId=" + response.getUserDefineRegionEntity().getRegionId()+ "\nId=" + re
sponse.getUserDefineRegionEntity().getId());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的执行结果如下:

```
RegionName=APIdoc
RegionId=cn-hangzhou:doc
Id=17926
```

② 说明 此处执行结果中的*Id=17926*,是系统中定义微服务空间的唯一标识序号。如果您需要删除微服务空间,则需要通过这个标识进行删除。

删除微服务空间

删除微服务空间需要满足以下前提条件:

- 微服务空间下无任何集群。
- 微服务空间下无任何微服务应用。
- 微服务空间下无任何ECS实例。

如果您在创建或编辑微服务空间时,已记录微服务空间的唯一标识序号(如Id=17926),则可以直接进行删除操作。

如果您没有记录微服务空间的唯一标识,则需要调用ListUserDefineRegion接口,获取微服务空间的唯一标识序号。

以下代码用于快速删除目标微服务空间。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeleteUserDefineRegionRequest;
import com.aliyuncs.edas.model.v20170801.DeleteUserDefineRegionResponse;
public class DeleteUserDefineRegion {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //需要删除的微服务空间所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeleteUserDefineRegionRequest request = new DeleteUserDefineRegionRequest();
       //待删除的微服务空间的唯一标识序号。
       request.setId((long) 17926);
       try {
           DeleteUserDefineRegionResponse response = client.getAcsResponse(request);
           System.out.println("RegionName=" + response.getRegionDefine().getRegionName() +
"\nRegionId=" + response.getRegionDefine().getRegionId()+ "\nMessage=" + response.getMessag
e());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的执行结果如下所示:

```
RegionName=APIdoc
RegionId=cn-hangzhou:doc
Message=success
```

结果验证

您在进行创建、编辑或删除微服务空间操作后,都可以调用ListUserDefineRegion接口查询您的自定义微服务空间列表,根据返回结果中的RegionName、RegionId和Description等参数进行判断您是否操作成功。

调用ListUserDefineRegion接口返回的结果如下:

```
"Message": "success",

"RequestId": "849E908D-E2DE-4798-937B-7210E00FFDD8",

"UserDefineRegionList": {

"UserDefineRegionEntity": [

{

"Description": "使用API创建的微服务空间",

"RegionName": "APIdoc2",

"UserId": "native_*****",

"DebugEnable": true,

"Id": 17925,

"RegionId": "cn-hangzhou:doc2",

"BelongRegion": "cn-hangzhou"

}

}

Code": 200

}
```

2.3. K8s集群管理

2.3.1. 使用控制台管理集群(K8s)

为了让您可以方便地在阿里云上使用Kubernetes管理容器应用,EDAS集成了容器服务Kubernetes版。您可以在容器服务Kubernetes版控制台创建Kubernetes集群,也可以将混合云集群(其他云域或IDC内自建集群)注册到容器服务Kubernetes版控制台,然后将集群导入EDAS用于应用部署,EDAS可提供高可用的容器应用管理能力,提高运维效率,让您专注于容器化的应用的开发与管理。

Kubernetes集群功能简介

现EDAS支持导入在容器服务Kubernetes版控制台创建的专有版Kubernetes集群、托管版Kubernetes集群和 Serverless Kubernetes集群,3种集群形态的对比详情,请参见什么是容器服务Kubernetes版。

专有版Kubernetes集群和托管版Kubernetes集群导入到EDAS中对应的是容器服务K8s集群,Serverless Kubernetes集群导入到EDAS中对应的是Serverless K8s集群。

混合云集群(其他云域或IDC内自建的集群)需要先将这些集群注册到容器服务Kubernetes版控制台,然后将集群导入EDAS。混合云集群导入到EDAS对应的是容器服务K8s集群。

混合云集群注册到容器服务Kubernetes版的具体操作,请参见创建注册集群并接入本地数据中心集群。

在EDAS中管理的K8s集群相对自建K8s集群来说有以下功能优势:

- 在云原生Kubernetes之上,提供应用托管能力,以应用视角一站式完成开源微服务治理和K8s应用轻量化运维:
 - 应用为中心的视角,管理K8s的原生工作负载如Deployment、Pod等,提供多AZ实例打散的高可用部署。

- 提供分批发布、按流量比例、请求参数的金丝雀灰度发布,借助EDAS全维度监控的发布变更单,让您的变更记录可跟踪。
- EDAS对接了主流DevOps系统,助力企业CI/CD落地,降本增效。
- 在开源微服务体系之上,对于使用市面上近五年的Spring Cloud和Dubbo框架自建的微服务应用无需修改任何代码即可迁移至EDAS,支持所有应用框架的微服务治理:
 - 支持应用发布过程中的无损下线、服务压测。
 - 应用运行时的服务鉴权、限流降级、离群实例摘除。
 - 应用运维的服务查询、服务测试。
- 通过产品化的方式,输出阿里巴巴应用安全生产的三板斧理念,实现可观测、可灰度、可回滚,让您的企业立即落地安全生产。
 - 可观测:通过应用总览、新版发布变更记录和发布后自动生成发布报告来实现多维度全流程监控。
 - 可灰度: 支持金丝雀发布, 支持应用按照流量比例或请求内容策略配置实现灰度。
 - 可回滚:支持发布过程一键回滚,并支持已运行应用回退至某历史版本。

Kubernetes集群工作流程

在容器服务K8s集群中部署应用,您需在容器服务Kubernetes版控制台中创建或注册集群,并将该集群导入到EDAS控制台,然后在导入的容器服务Kubernetes版集群中使用部署包或镜像来部署应用。



② 说明 阿里云Kubernetes集群和混合云集群在导入EDAS后,管理集群的操作流程类似,本文以阿里云容器服务K8s集群为例。

前提条件

- 阿里云账号已开通EDAS服务,请参见开通EDAS。
- 该阿里云账号已开通容器服务Kubernetes版并完成角色授权,请参见容器服务默认角色。
- 如果使用RAM用户进行操作,需要保证该RAM用户已经由阿里云账号授予登录等操作权限。

在容器服务Kubernetes版控制台创建集群

- 在EDAS中使用容器服务K8s集群,请在容器服务Kubernetes版控制台创建托管版Kubernetes集群或专有版Kubernetes集群,请参见:
 - o 快速创建Kubernetes托管版集群
 - o 创建Kubernetes专有版集群
- 在EDAS中使用Serverless K8s集群,请在容器服务Kubernetes版控制台创建Serverless Kubernetes集群,请参见创建Serverless Kubernetes集群。

在EDAS控制台导入Kubernetes集群

在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Promet heus监控组件(ack-arms-promet heus)。

1. 登录EDAS控制台。

- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,选择要导入集群的目标微服务空间,然后单击**同步容器服务Kubernetes集** 群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在导入Kubernetes集群对话框选择微服务空间,选择是否打开服务网格开关,然后单击导入。

? 说明

- 如果您未创建微服务空间,在此步骤可不选择微服务空间,使用默认微服务空间。
- 如果您需要部署多语言应用,请打开**服务网格**开关。
- 开启服务网格时,默认会创建两个私网规格SLB实例(slb.s1.small)实现管控,并暴露两个 SLB实例的端口(6443端口和15011端口)。更多信息,请参见<mark>背景信息</mark>。 默认生成的两个小规格SLB实例(slb.s1.small)会产生费用。具体收费标准,请参见包年包月 收费标准或按量计费收费标准。

当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

在EDAS控制台取消导入和清理Kubernetes集群

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群)。
- 3. 在顶部菜单栏选择地域,在页面上方选择集群所在的微服务空间,然后在目标K8s集群的操作列单击取消导入,并在提示对话框单击确定。
 - 取消导入的集群如果**集群状态**为运行中,则集群可以被再次导入至EDAS来创建应用。
 - 取消导入的集群如果**集群状态**为已删除,则可以在集群的操作列单击清理来将该集群从EDAS的集群列表中删除。
 - ② 说明 删除集群的相关操作和常见问题,请参见删除集群。

在EDAS控制台查看Kubernetes集群详情

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群)。
- 3. 单击集群ID进入**集群详情**页面。
 - 在集群详情页面您可查看该集群的详细信息、ECS实例、Deployments和应用列表等信息。
 - 集群信息:该区域展示了集群基本信息,如集群ID、csClusterId、集群名称、微服务空间、集群类型、VPC ID、网络类型、集群状态和描述信息等信息。
 - 在集群信息区域右上角单击查看详情,可查看集群的详细信息。
 - 在集群信息区域右上角单击编辑,可修改集群的描述信息。
 - ECS实例:该区域展示了该集群包含的ECS实例列表和ECS实例概要信息。
 - Deployments:该区域展示您转换和托管到EDAS的第三方平台自建Deployment。
 - **应用列表**:该区域展示了集群内包含的应用,您可以查看应用名称、JDK版本、应用运行环境、实例总数、运行中的实例数和应用负责人等信息。您可以单击应用名称进入应用详情页面。

K8s集群导入操作约束

- 集群安全组配置约束:
 - 需要保证集群所有节点在集群的安全组内或者与集群的安全组能互通。互通的具体操作,请参见<mark>为什么容器之间网络不通</mark>?。
 - 不能删除ACK为集群安全组设置的默认规则。
- 集群节点配置约束:
 - 需要为集群预留可分配的CPU,内存,Pod资源,确保EDAS管控组件能正常运行。
 - 不能删除ACK为节点配置的RAM角色 KubernetesWorkerRole-* 。
- 集群apiServer SLB配置约束:
 - 不要阻拦来自内部访问地址100.104.0.0/16的访问请求。
 - 不要删除SLB上的ACK添加的内置标签。
 - 不要复用SLB的6443端口。
- 集群Helm Chart操作约束:
 - 不能删除EDAS安装的ahas-sentinel-pilot、arms-eventer、arms-pilot和arms-prom组件,包括这些 Helm Chart安装的所有资源。
 - 不能安装开源的oam-runtime、kubevela、keda或flagger。
 - 不能删除或修改任何在edas-oam-system命名空间下的K8s资源。
- 集群ClusterRole操作约束:
 - 不能使用ACK控制台、kubectl或者第三方界面删除或修改edas-default-cluster-role。
- ClusterRoleBinding操作约束:
 - 不能使用ACK控制台、kubectl或者第三方界面删除或修改edas-default-cluster-role-binding、edas-oam-cluster-role-binding、keda-hpa-controller-external-metrics。
- CRD和CR操作约束:

- 不能直接操作以下CRD和CR资源,列表如下:
 - alert providers.f lagger.app
 - applicationconfigurations.core.oam.dev
 - applications.oam-domain.alibabacloud.com
 - applicationscopes.core.oam.dev
 - autoscalings.edas.aliyun.oam.com
 - basecomponents.oam-domain.alibabacloud.com
 - canaries.flagger.app
 - component schematics.core.oam.dev
 - crdreleases.clm.cloudnativeapp.io
 - dynamiclabels.extension.oam.dev
 - imagebuilders.edas.aliyun.oam.com
 - logcollectors.edas.aliyun.oam.com
 - meshtraits.edas.aliyun.oam.com
 - metrictemplates.flagger.app
 - mserulet rait s.edas.aliyun.oam.com
 - packageversions.oam-domain.alibabacloud.com
 - rollout s.edas.aliyun.oam.com
 - scaledobjects.keda.k8s.io
 - scalingrules.oam-domain.alibabacloud.com
 - serviceregistrytraits.edas.aliyun.oam.com
 - servicetraits.edas.aliyun.oam.com
 - sources.clm.cloudnativeapp.io
 - traits.core.oam.dev
 - triggerauthentications.keda.k8s.io
 - workloadtypes.core.oam.dev
- o 不能修改EDAS创建的aliyunlogconfigs.log.alibabacloud.com(含有以下标签: edas-domain: edas-admin) 资源。

相关链接

- 在K8s集群中使用镜像部署Java微服务应用
- 在容器服务K8s集群中使用JAR包或WAR包部署应用
- 在Serverless K8s集群中使用镜像方式部署应用
- 在Serverless K8s集群中使用JAR包或WAR包部署应用
- 从源码构建PHP应用并部署至容器服务K8s集群

2.3.2. 使用API管理K8s集群

本文介绍如何使用EDAS提供的Java SDK调用API导入和取消导入K8s集群。

前提条件

在导入或取消导入K8s集群前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定导入或取消导入K8s集群的地域,假设为cn-hangzhou。
- 在容器服务Kubernetes版控制台创建集群。具体操作,请参见:
 - 在EDAS中使用容器服务K8s集群,请在容器服务Kubernetes版控制台创建托管版Kubernetes集群或专有版Kubernetes集群,请参见:
 - 快速创建Kubernetes托管版集群
 - 创建Kubernetes专有版集群
 - 在EDAS中使用Serverless K8s集群,请在容器服务Kubernetes版控制台创建Serverless Kubernetes集群,请参见创建Serverless Kubernetes集群。

如果已经创建容器服务K8s集群或Serverless K8s集群,请调用GetK8sCluster接口查询容器服务K8s集群或Serverless K8s集群列表,获取目标集群的Clusterld,假设为*da60f685-558a-4e00-b549-15e9143d*****。

● 创建微服务空间。具体操作,请参见创建命名空间。 如果已经存在命名空间,您可以调用ListUserDefineRegion接口查询用户自定义命名空间列表,获取目标命 名空间的*RegionId*。

Kubernetes集群功能简介

现EDAS支持导入在容器服务Kubernetes版控制台创建的专有版Kubernetes集群、托管版Kubernetes集群和 Serverless Kubernetes集群,3种集群形态的对比详情,请参见什么是容器服务Kubernetes版。

专有版Kubernetes集群和托管版Kubernetes集群导入到EDAS中对应的是容器服务K8s集群,Serverless Kubernetes集群导入到EDAS中对应的是Serverless K8s集群。

混合云集群(其他云域或IDC内自建的集群)需要先将这些集群注册到容器服务Kubernetes版控制台,然后将集群导入EDAS。混合云集群导入到EDAS对应的是容器服务K8s集群。

混合云集群注册到容器服务Kubernetes版的具体操作,请参见创建注册集群并接入本地数据中心集群。

在EDAS中管理的K8s集群相对自建K8s集群来说有以下功能优势:

- 在云原生Kubernetes之上,提供应用托管能力,以应用视角一站式完成开源微服务治理和K8s应用轻量化 运维:
 - 应用为中心的视角,管理K8s的原生工作负载如Deployment、Pod等,提供多AZ实例打散的高可用部署。
 - 提供分批发布、按流量比例、请求参数的金丝雀灰度发布,借助EDAS全维度监控的发布变更单,让您的变更记录可跟踪。
 - EDAS对接了主流DevOps系统,助力企业CI/CD落地,降本增效。
- 在开源微服务体系之上,对于使用市面上近五年的Spring Cloud和Dubbo框架自建的微服务应用无需修改任何代码即可迁移至EDAS,支持所有应用框架的微服务治理:
 - 支持应用发布过程中的无损下线、服务压测。
 - 应用运行时的服务鉴权、限流降级、离群实例摘除。
 - 应用运维的服务查询、服务测试。
- 通过产品化的方式,输出阿里巴巴应用安全生产的三板斧理念,实现可观测、可灰度、可回滚,让您的企业立即落地安全生产。
 - o 可观测:通过应用总览、新版发布变更记录和发布后自动生成发布报告来实现多维度全流程监控。
 - 可灰度: 支持金丝雀发布, 支持应用按照流量比例或请求内容策略配置实现灰度。
 - 可回滚:支持发布过程一键回滚,并支持已运行应用回退至某历史版本。

导入集群

以下代码用于导入K8s集群到目标微服务空间。

□ 注意 导入K8s集群包含一些操作约束,请您仔细阅读。更多信息,请参见。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.ImportK8sClusterRequest;
import com.aliyuncs.edas.model.v20170801.ImportK8sClusterResponse;
public class ImportK8sCluster {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //导入K8s集群的所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun user sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       ImportK8sClusterRequest request = new ImportK8sClusterRequest();
       //指定K8s集群的集群ID。
       request.setClusterId("da60f685-558a-4e00-b549-15e9143d****");
       //指定K8s集群导入到的目标微服务空间。
       request.setNamespaceId("cn-hangzhou:doc");
       //是否安装服务网格, true表示安装, false表示不安装。
       request.setEnableAsm(true);
       try {
           ImportK8sClusterResponse response = client.getAcsResponse(request);
           System.out.println("Data=" + response.getData() + "\nMessage=" + response.getMe
ssage() + "\nRequestId=" + response.getRequestId());
        } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的执行结果如下:

```
Data=4d45f6db-2942-4747-a7ba-b0d43273****

Message=success
RequestId=847CF607-B352-4A71-828A-C0F8DECDEDF6
```

取消导入集群

以下代码用于取消导入集群。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeleteClusterRequest;
import com.aliyuncs.edas.model.v20170801.DeleteClusterResponse;
public class DeleteCluster {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //取消导入K8s集群的所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeleteClusterRequest request = new DeleteClusterRequest();
       //需要取消导入集群的集群ID。
       request.setClusterId("da60f685-558a-4e00-b549-15e9143d****");
           DeleteClusterResponse response = client.getAcsResponse(request);
           System.out.println("Data=" + response.getData() + "\nMessage=" + response.getMe
ssage() + "\nRequestId=" + response.getRequestId());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的执行结果如下:

```
Data=true
Message=success
RequestId=68CDDEC0-593E-4C1A-9CDC-96F9FEF0924E
```

结果验证

在导入或者取消导入K8s集群的操作后,您可以调用GetK8sCluster接口获取容器服务K8s集群或Serverless K8s集群列表,根据返回结果中的集群导入状态(ClusterImportStatus)参数值判断您是否操作成功。

调用 Get K8sClust er接口返回的结果如下:

```
"RequestId": "809FCC08-FFE1-417D-8ADF-F5D5BFD968D7",
"Message": "success",
"ClusterPage": {
  "ClusterList": {
    "Cluster": [
     {
        "ClusterImportStatus": 1,
        "NodeNum": 6,
       "ClusterId": "da60f685-558a-4e00-b549-15e9143d****",
      },
        "ClusterImportStatus": 0,
        "NodeNum": 3,
        "ClusterId": "6dca9bc0-b928-4f2c-ae76-396bd108****",
     }
   ]
  "PageSize": 12,
  "CurrentPage": 0,
 "TotalSize": 12
},
"Code": 200
```

请查看上述执行结果中的集群导入状态ClusterImportStatus参数值,通过该值判断导入或取消导入K8s集群是否成功。ClusterImportStatus的取值如下:

● 0: 没有导入

● 1: 导入成功

● 2: 导入失败

● 3: 正在导入

● 4: 集群已删除

3.创建和部署应用(K8s)

3.1. 创建和部署应用概述(K8s)

EDAS支持使用ECS环境和K8s环境托管应用,当您选择K8s环境托管应用时,EDAS支持托管Java应用、PHP应用以及多语言(包含Node.js、Go和Python等多种语言)应用。本文介绍不同应用类型所支持的部署方式。

Java应用

Spring Cloud、Dubbo或HSF应用开发完成后,可通过WAR包、JAR包或镜像的方式部署至K8s集群。

部署方式	介绍	运行环境	监控及治理方案	示例文档
镜像部署	适用于在K8s集群中 使用自定义的镜像 部署应用。	自定义	默认自动挂载Java Agent进行精细化监 控,并提供完整微 服务治理方案(金 丝雀发布、服务鉴 权、限流降级 等)。	在K8s集群中使用镜 像部署Java微服务应 用
JAR包部署	适用于Dubbo和 Spring Boot应用。	Java		在容器服务K8s集群中使用JAR包或WAR 包部署应用
WAR包部署	适用于Dubbo和 Spring应用。	Java、Apache		
WAR包或者FatJar包 部署	适用于部署HSF应 用。	EDAS-Container		

② 说明 如果您想在K8s环境中同时创建多个Java应用,可以调用Insert K8sApplication接口创建应用。更多信息,请参见Insert K8sApplication。

PHP应用

目前支持从源码构建PHP应用并托管到EDAS。在部署应用时,需要选择PHP版本,EDAS提供了*PHP 7.3/Apache/Centos 7*运行环境,该运行环境内已经安装了常用的PHP插件。

部署PHP应用时默认自动挂载PHP Agent进行精细化监控,并通过Istio(Service Mesh)提供完整微服务治理方案(金丝雀发布、服务鉴权、限流降级等)。

示例文档:从源码构建PHP应用并部署至容器服务K8s集群

多语言应用

目前支持使用自定义的镜像将多语言应用托管到EDAS。

部署多语言应用时通过Ist io(Service Mesh)提供完整微服务治理方案(金丝雀发布、服务鉴权、限流降级等)。

示例文档: 在K8s环境中通过镜像部署多语言微服务应用。

② 说明 如果您想同时在K8s环境中创建多个多语言应用,可以调用Insert K8sApplication接口创建应用。更多信息,请参见Insert K8sApplication。

3.2. Java应用

3.2.1. JAR包或WAR包部署

3.2.1.1. 在容器服务K8s集群中使用JAR包或WAR包部署应用

EDAS在云原生K8s之上,全面集成了容器服务Kubernetes版(Alibaba Cloud Container Service for Kubernetes,简称容器服务ACK),支持K8s容器化应用的全生命周期管理。容器服务K8s集群集成了阿里云虚拟机、存储、网络和安全等能力,拥有云端极佳的K8s容器化应用运行环境。本教程介绍如何在容器服务K8s集群中使用官方Demo的JAR包或WAR包来部署应用。

背景信息

在容器服务K8s集群中部署应用,您需在容器服务Kubernetes版控制台中创建容器服务Kubernetes版集群,并将该集群导入到EDAS控制台,然后在导入的容器服务K8s集群中使用部署包或镜像来部署应用。



前提条件

- 您的阿里云账号已同时开通EDAS和容器服务Kubernetes版,请参见:
 - o 开通EDAS
 - o 首次使用容器服务Kubernetes版
- 创建微服务空间
- 在容器服务Kubernetes版完成角色授权,请参见容器服务默认角色。

步骤一: 创建容器服务Kubernetes集群

登录<mark>容器服务Kubernetes版控制台</mark>,创建容器服务Kubernetes集群。具体操作,请参见快速创建Kubernetes托管版集群。

如果创建Serverless Kubernetes集群,专有网络请选择自动创建(选择已有网络,创建完毕后请查看集群资源是否包含VPC和交换机资源),服务发现请选择PrivateZone,以便Serverless Kubernetes集群在导入EDAS后可以使用服务网格。

步骤二:在EDAS控制台中导入容器服务K8s集群

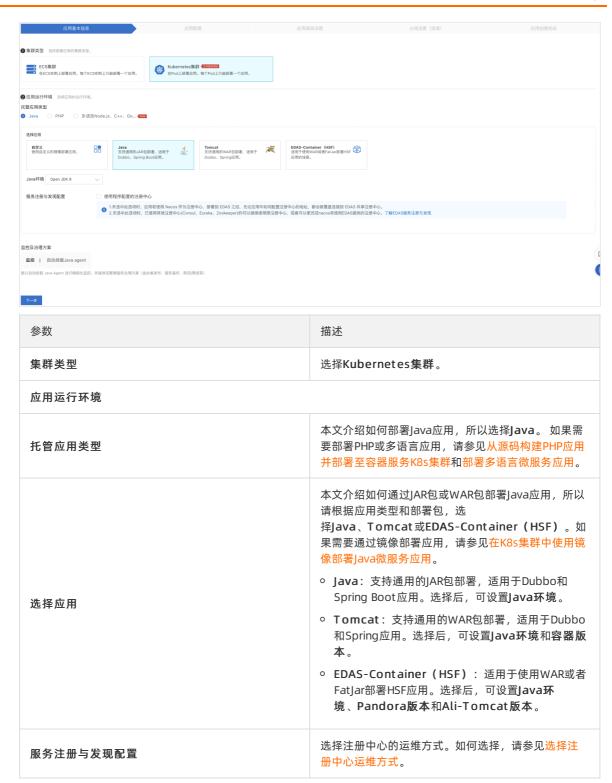
在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Promet heus监控组件(ack-arms-promet heus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,然后单击同步容器服务Kubernetes集群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在**导入Kubernet es集群**对话框,从下拉列表中选择目标微服务空间,根据需要选择是否开启**服务网格**,然后单击**导入**。

当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

步骤三:在容器服务K8s集群中部署应用

- ② 说明 WAR和IAR的部署步骤一致,本文以通过IAR为例介绍如何部署应用。
- 1. 登录EDAS控制台,在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列表**页面左上角单击**创建应用**。
- 2. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。



3. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。

参数	描述	

参数	描述
微服务空间	应用所属空间。选择您创建的微服务空间,如果您未创建微服务空间或不做选择,微服务空间则设置为默认。 如果您没有创建微服务空间,或者需要创建新的微服务空间,您可以直接单击创建微服务空间,创建一个全新的微服务空间。具体操作,请参见使用控制台管理微服务空间。
集群	在右侧下拉列表内选择已导入的容器服务K8s集群。如果您选择的K8s集群未导入到EDAS,请选中该集群在EDAS初次使用,勾选后,将会在应用创建时将该集群导入EDAS,因此会产生一定耗时。并确认是否开启服务网格。 ② 说明 集群导入的微服务空间可以不是应用所属微服务空间。
K8s Namespace	K8s Namespace通过将系统内部的对象分配到不同的Namespace中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。 default:没有其他命名空间的对象的默认命名空间。 kube-system:系统创建的对象的命名空间。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。此处以选择default为例。如果您需要自定义创建K8s Namespace,请单击创建K8s Namespace并设置K8s Namespace名称。名称设置要求只能包含数字、小写字母和短划线(-),且首尾只能是字母或数字,长度为1~63个字符。
应用名称	输入应用名称,必须以字母开头,允许数字、字母、 短划线(-)组合。最多可输入36个字符。
应用描述	输入应用描述 ,最大长度为128个字符。
部署包来源	 ● 自定义程序 需选择文件上传方式: ■ 上传JAR包:选择下载好的JAR包并上传。 ■ JAR包地址:输入Demo包地址。 ○ 官方Demo EDAS提供的Demo类型有Spring Cloud服务端应 用、Spring Cloud客户端应用、Dubbo服务端应用和Dubbo客户端应用,请依据实际需求选择。

参数	描述
版本	请输入版本,您可以自定义版本号,也可以单击右侧的用时间戳作为版本号自动生成版本号。
时区	设置应用的时区信息。
OpenJDK基础镜像	配置镜像 ○ 阿里云镜像服务选择当前账号。 选择镜像所属地域、容器镜像服务、镜像仓库命名空间、镜像仓库名和镜像版本。 ○ 阿里云镜像服务选择为其他阿里云账号。 ■ 如您的镜像存放在公开仓库中,那么您配置完整镜像地址即可。 ■ 如您的镜像存放在私有仓库中,那么您需要使用免密插件拉取容器镜像,请参见使用免密组件拉取容器镜像。 ② 说明 在应用基本信息页面Java环境处,选择自定义OpenJDK时,参数OpenJDK基础镜像才会出现。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

- 4. (可选)配置应用高级设置。
 - 。 配置调度规则
 - 。 配置启动命令
 - o 配置环境变量
 - 。 配置持久化存储
 - 。 配置本地存储
 - 配置应用生命周期的钩子和探针
 - 。 配置日志收集
 - o 配置Tomcat
 - o 配置Java启动参数
 - o 实现K8s集群应用的限流降级
 - 。 配置挂载
- 5. 完成设置后单击**创建应用。** 应用创建可能需要几分钟,创建过程中,可以通过变更记录<mark>查看应用总览</mark>跟踪应用的创建流程。创建完

成后,返回**应用总览**页面查看实例Pod的运行状态若为运行中则说明应用发布成功,单击Pod的运行状态可以查看应用实例的工作负载、容器组(Pod)和启动命令等高级配置信息。

后续步骤

应用创建完毕后,可以通过添加公网SLB实现公网访问,添加私网SLB实现同VPC内所有节点够能通过私网负载均衡访问您的应用。相关操作,请参见添加负载均衡SLB或复用负载均衡SLB。

联系我们

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.2.1.2. 在Serverless K8s集群中使用JAR包或WAR包部署应用

EDAS在云原生Kubernetes之上,全面集成了容器服务Kubernetes版,支持K8s容器化应用的全生命周期管理。Serverless Kubernetes适合快速伸缩的敏捷业务场景。本教程介绍如何在Serverless K8s集群中使用官方Demo的IAR包或WAR包来部署应用。

背景信息

在Serverless K8s集群中部署应用,您需在容器服务Kubernetes版控制台中创建Serverless K8s集群,并将该集群导入到EDAS控制台,然后在导入的Serverless K8s集群中使用部署包或镜像来部署应用。



前提条件

- 您的阿里云账号已同时开通EDAS和容器服务Kubernetes版,请参见:
 - o 开通EDAS
 - o 首次使用容器服务Kubernetes版
- 创建微服务空间
- 在容器服务Kubernetes版完成角色授权,请参见容器服务默认角色。

步骤一: 创建Serverless Kubernetes集群

登录容器服务Kubernetes版控制台,创建Serverless Kubernetes集群。具体操作,请参见创建Serverless Kubernetes集群。

步骤二:在EDAS控制台中导入Serverless Kubernetes集群

在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Prometheus监控组件(ack-arms-prometheus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > Serverless K8s 集群。
- 3. 在顶部菜单栏选择地域,选择要导入集群的目标微服务空间,然后单击同步Serverless K8s集群。
- 4. 在导入的Serverless K8s集群的操作列单击导入。 当Serverless K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示Serverless K8s集群已成功导入到EDAS。

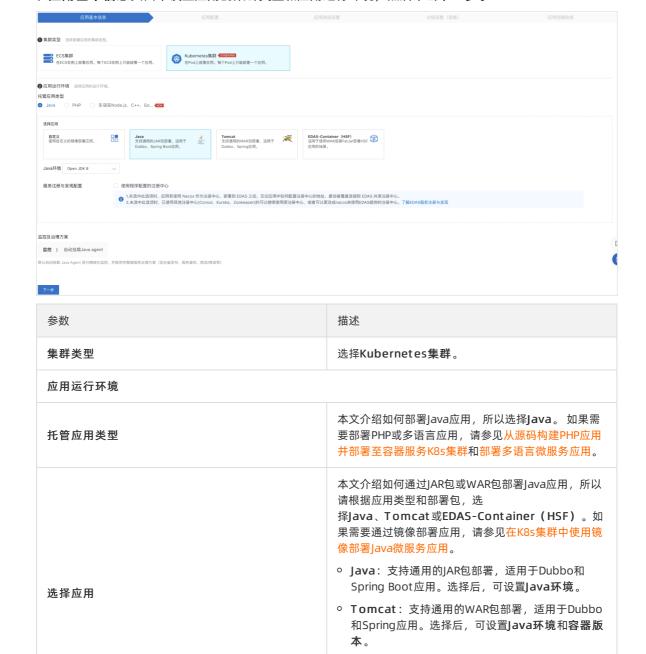
步骤三:在Serverless K8s集群中部署应用

② 说明 WAR和JAR的部署步骤一致,本文以通过JAR为例介绍如何部署应用。

1. 登录EDAS控制台。

服务注册与发现配置

- 2. 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列表** 表页面左上角单击**创建应用**。
- 3. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。

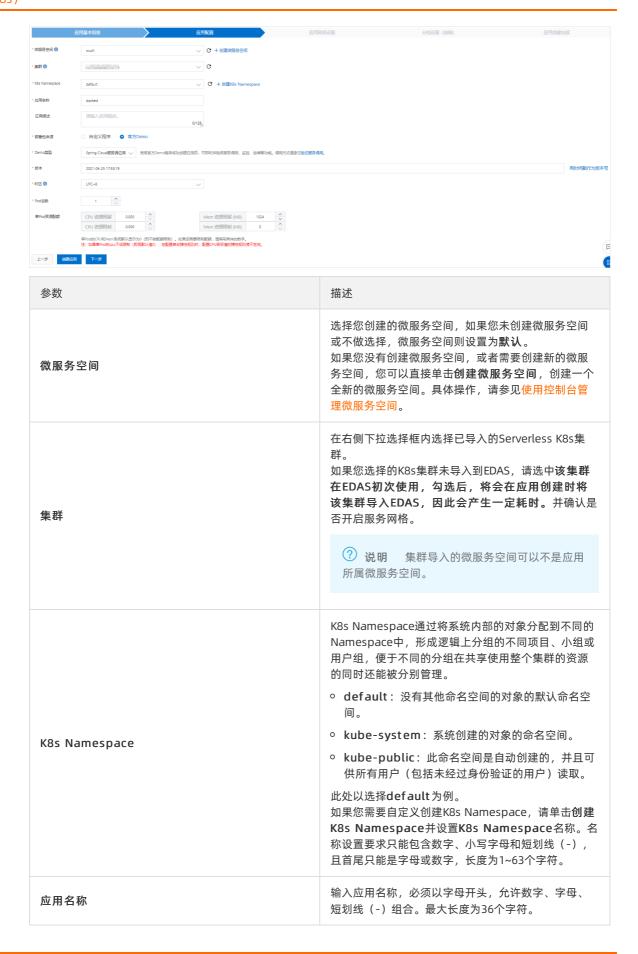


4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。

册中心运维方式。

 EDAS-Container (HSF): 适用于使用WAR或者 FatJar部署HSF应用。选择后,可设置Java环 境、Pandora版本和Ali-Tomcat版本。

选择注册中心的运维方式。如何选择,请参见选择注



参数	描述
应用描述	输入应用描述 , 最大长度为128个字符。
部署包来源	 ○ 自定义程序 需选择文件上传方式: ■ 上传JAR包: 选择下载好的JAR包并上传。 ■ JAR包地址: 输入Demo包地址。 ○ 官方Demo EDAS提供的Demo类型有Spring Cloud服务端应 用、Spring Cloud客户端应用、Dubbo服务端应用和Dubbo客户端应用,请依据实际需求选择。
版本	请输入版本,您可以自定义版本号也可以使用时间戳 作为版本号。
时区	设置应用的时区信息。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

5. (可选)配置应用高级设置。

- 。 配置启动命令
- 。 配置环境变量
- 。 配置持久化存储
- 。 配置应用生命周期的钩子和探针
- o 配置Tomcat
- o 配置lava启动参数
- o 实现K8s集群应用的限流降级
- 6. 完成设置后单击创建应用。

应用创建可能需要几分钟,创建过程中,可以通过变更记录查看应用总览跟踪应用的创建流程。创建完成后,返回应用总览页面查看实例Pod的运行状态若为运行中则说明应用发布成功,单击Pod的运行状态可以查看应用实例的工作负载、容器组(Pod)和启动命令等高级配置信息。

后续步骤

应用创建完毕后,可以通过添加公网SLB实现公网访问,添加私网SLB实现同VPC内所有节点够能通过私网负载均衡访问您的应用。具体操作,请参见为Kubernetes集群中的应用添加负载均衡SLB。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.2.1.3. 使用API在K8s集群中创建Java应用(JAR包或WAR包)

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中创建Java应用(JAR包或WAR包)。

前提条件

调用API创建应用前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定应用部署地域。
- 创建微服务空间。具体操作,请参见<mark>创建命名空间。</mark> 如果已经存在命名空间,您可以调用ListUserDefineRegion接口查询用户自定义命名空间列表,获取目标命 名空间的*RegionId*。
- 导入目标集群(假设集群ID为 da60f685-558a-4e00-b549-15e9143d****) 到命名空间。具体操作,请参见导入集群。
 - 如果已经导入集群到目标命名空间,您可以调用ListCluster接口查询集群列表,获取目标集群的ClusterId。
- 已将应用部署包(JAR包或WAR包)上传至目标地址。本示例以OSS存储路径为例,假设为*https:doc***.o ss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.1-SNAPSHOT.jar*。

创建Java应用(JAR包或WAR包)

以下代码适用于在EDAS K8s环境中快速创建Java应用(JAR包或WAR包)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见Insert K8sApplication。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationResponse;
public class InsertK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun_user_sk = "<yourAccessKeySecret>";
       //需要创建应用的目标地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       InsertK8sApplicationRequest request = new InsertK8sApplicationRequest();
       //自定义设置应用名称。必须以字母开头,允许数字、字母、短划线(-)组合。最多可输入36个字符。
       request.setAppName("edas-SDK0414java");
       request.setClusterId("da60f685-558a-4e00-b549-15e9143d****");
       //应用实例数。
       request.setReplicas(2);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setLimitCpu(0);
       request.setLimitMem(0);
       request.setRequestsCpu(0);
       request.setRequestsMem(0);
       //EDAS微服务空间对应ID,非默认微服务空间须填写。
       request.setLogicalRegionId("cn-hangzhou:doc");
       //部署包类型、地址和版本,支持设置为FatJar和WAR。
       request.setPackageType("FatJar");
       request.setPackageUrl("https:doc***.oss-cn-hangzhou.aliyuncs.com/sc-***-D-0.0.1-SN
APSHOT.jar");
       request.setPackageVersion("20210423:1442");
       //部署的包依赖的JDK版本。可选的参数值为Open JDK 7和Open JDK 8。
       request.setJDK( "Open JDK 8");
       try {
           InsertK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("AppId=" + response.getApplicationInfo().getAppId() + "\nApp
Name=" + response.getApplicationInfo().getAppName() + "\nChangeOrderId=" + response.getAppl
icationInfo().getChangeOrderId());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行代码,返回示例如下:

```
AppId=8ae9cfdb-c20c-4de0-9b21-2b31e7b6****
AppName=edas-SDK0414java
ChangeOrderId=44a1402a-9e59-4494-baa9-f18877e0****
```

结果验证

您在创建应用后,可以调用Get ChangeOrderInf o接口查看应用变更流程详情,获取创建应用的变更状态。 具体API参数详情,请参见Get ChangeOrderInf o。

调用Get Change OrderInfo接口查看应用变更流程详情,返回的执行结果如下:

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断创建应用是否成功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示创建应用失败。您可以根据根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新创建应用。

3.2.2. 镜像部署

3.2.2.1. 制作应用镜像

您可以在本地的开发工具中使用命令将应用程序打包为WAR包或JAR包,基于WAR包或JAR包制作镜像,再将镜像上传到阿里云镜像仓库来部署应用。本文介绍如何制作不同框架应用镜像的Dockfile以及如何将镜像上传到阿里云镜像仓库。

在制作应用镜像前,请查阅<mark>附录:运行时环境变量信息</mark>并按照规约,同时参考以下内容制作EDAS应用镜像。

使用说明

本文涉及的Dockerfile及镜像内容可通过Git hub仓库自行查阅。

使用规范和限制

通过Dockerfile制作自定义镜像时,需要遵循以下规范及限制。

● 服务信息

包含运行时所需连接的EDAS域名、端口等信息。

资源

资源类型	资源名称	描述
ConfigMap	edas-envs	EDAS服务信息

环境变量

环境变量Key	类型	描述
EDAS_ADDRESS_SERVER_DOMAIN	String	配置中心服务域名或IP
EDAS_ADDRESS_SERVER_PORT	String	配置中心服务端口
EDAS_CONFIGSERVER_CLIENT_POR T	String	ConfigServer服务端口

● 应用运行时环境变量 EDAS部署时会提供以下环境变量Env,保证应用运行正常,请勿覆盖配置。 环境变量

环境变量Key	类型	描述
POD_IP	String	POD IP地址
EDAS_APP_ID	String	EDAS应用ID
EDAS_ECC_ID	String	EDAS ECC ID
EDAS_PROJECT_NAME	String	同EDAS_APP_ID,用于调用链解析。
EDAS_JM_CONTAINER_ID	String	同EDAS_ECC_ID,用于调用链解析。
EDAS_CATALINA_OPTS	String	中间件运行时所需CATALINA_OPTS 参数
CATALINA_OPTS	String	同EDAS_CATALINA_OPTS,默认 TOMCAT启动参数。

环境变量Key	类型	描述
CATALINA_HOME	String	所安装的TOMCAT路径
PANDORA_LOCATION	String	所安装的Pandora路径,HSF应用中可见。

创建标准Dockerfile

Dockerfile是一个文本格式的配置文件, 您可以使用Dockerfile快速创建镜像。

您可以根据实际应用框架,通过Dockerfile分别制作HSF、Spring Cloud或Dubbo应用镜像。下面将通过示例分别介绍如何创建这些不同框架应用的Dockerfile。

一个EDAS标准Dockerfile描述了EDAS创建应用运行环境的所有指令,包括下载、安装和启动OpenJDK、Tomcat、WAR包和JAR包等。您还可以通过改变Dockerfile内容,完成如OpenJDK版本替换、Tomcat配置修改、改变运行时环境等操作。关于OpenJDK的使用指导,请参见自定义镜像:使用新版本的OpenJDK。

关于Dockerfile的使用可参考如下文档:

- HSF应用的Dockerfile示例
- Spring Cloud或Dubbo应用的Dockerfile示例

HSF应用的Dockerfile示例

```
# Using the centos7 + openjdk8 + ali tomcat7 + pandora 3.5.9 for the base
FROM apaas/edas-centos-openjdk8-alitomcat7:latest
MAINTAINER Alibaba Cloud EDAS Team<edas-dev@list.alibaba-inc.com>
ENV TZ="Asia/Shanghai"
# Default put your application package into /home/admin/app/
ENV APP HOME /home/admin/app/
#####---> First case: deploy a fat-jar file
ARG APP LOCATION=https://edas-hz.oss-cn-hangzhou.aliyuncs.com/prod/demo/HSF PROVIDER.jar
#####---> Second case: deploy a war file, simply change the file name your wanted.
#ARG APP LOCATION=https://edas-hz.oss-cn-hangzhou.aliyuncs.com/prod/demo/HSF PROVIDER.war
#####---> Third case: deploy a local storage file.
#ARG APP LOCATION=/Users/yanliang.lyl/workspace/java/demo/HSF PROVIDER.war
# Then download to package into /home/admin/app/
ADD ${APP LOCATION} ${APP HOME}/
# Default working dir is set to /home/admin
# WORKDIR /home/admin
# Entry point set to /home/admin/bin/start.sh,
# which including inside basic image (apaas/edas-centos-openjdk8-alitomcat8)
# CMD ["/home/admin/bin/start.sh"]
```

Spring Cloud或Dubbo应用的Dockerfile示例

```
# Using centos7 + openjdk8 + tomcat8 for the base
FROM apaas/edas:latest
# Default put your application package into /home/admin/app/
ENV APP_HOME /home/admin/app/
##### SIMPLE REPLACE YOUR PACKAGE FILE
#####----> First case: deploy a fat-jar file
ARG APP_LOCATION=https://edas-hz.oss-cn-hangzhou.aliyuncs.com/prod/demo/DUBBO_PROVIDER.jar
#####----> Second case: deploy a war file, simply change the file name your wanted.
#ARG APP_LOCATION=https://edas-hz.oss-cn-hangzhou.aliyuncs.com/prod/demo/DUBBO_PROVIDER.war
#####----> Third case: deploy a local storage file.
#ARG APP_LOCATION=/Users/yanliang.lyl/workspace/java/demo/SPRINT_CLOUD_PROVIDER.war
# Then download to package into /home/admin/app/
ADD ${APP_LOCATION} ${APP_HOME}/
```

EDAS基础镜像说明

您可以根据您应用的运行环境,自由选择以下的基础镜像:

镜像名称	OS版本	Java版本	Tomcat版本	Pandora版本	备注
apaas/edas	Cent OS 7	OpenJDK 1.8	8.5.78	N/A	与apaas/edas- centos- openjdk8- tomcat8相同, 最新版本号为: latest
apaas/edas- centos- openjdk8- tomcat8	Cent OS 7	OpenJDK 1.8	8.5.78	N/A	普通 SpringCloud或 Dubbo应用推 荐使用此版本, 最新版本号为: latest
apaas/edas- centos- openjdk8- tomcat7	Cent OS 7	OpenJDK 1.8	7.0.93	N/A	基于Tomcat 7.x的普通 SpringCloud或 Dubbo应用推 荐使用此版本, 最新版本号为: latest
apaas/edas- centos- openjdk8- alitomcat8	Cent OS 7	OpenJDK 1.8	8.5.37	3.5.9	基于Tomcat8.x 的HSF应用推荐 使用此版本,最 新版本号为: latest
apaas/edas- centos- openjdk8- alitomcat7	Cent OS 7	OpenJDK 1.8	7.0.92	3.5.9	基于Tomcat7.x 的HSF应用推荐 使用此版本,最 新版本号为: latest

镜像名称	OS版本	Java版本	Tomcat版本	Pandora版本	备注
apaas/edas- centos- openjdk8- alitomcat8-v4	Cent OS 7	OpenJDK 1.8	8.5.37	4.0.2	基于Tomcat8.x 的HSF 4.0应用 推荐使用此版 本,最新版本号 为:latest
apaas/edas- centos- openjdk8- alitomcat7-v4	Cent OS 7	OpenJDK 1.8	7.0.92	4.0.2	基于Tomcat7.x 的HSF 4.0应用 推荐使用此版 本,最新版本号 为:latest

更多镜像请访问aPaaS Docker Hub官方镜像仓库;同时,如果您需要额外的基础镜像,请<mark>提交工单</mark>给我们;如果您对我们的基础镜像有任何疑问,请参见基础镜像Dockerfile的Git Hub地址。

自定义镜像:使用新版本的OpenIDK

您可以在创建的标准Dockerfile上根据实际需求进行自定义设置。

升级OpenJDK: 在Dockerfile中,可下载和安装新版本的OpenJDK,以下示例为下载安装OpenJDK 9。

```
# 先卸载之前的JDK,然后再下载安装OpenJDK 9
RUN yum erase -y java; yum -y install java-1.9.0-openjdk-devel
```

自定义镜像:升级EDAS Container版本(HSF应用)

对于部署到容器服务K8s集群中的HSF应用,为使用中间件的新特性或修复原有版本的问题,可以升级应用容器版本。升级步骤如下:

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择应用列表。
- 3. 在应用列表页面单击创建应用,在应用基本信息页面应用运行环境区域查看、获取Pandora版本和Ali-Tomcat版本。
- 4. 替换Dockerfile中的版本号信息,如3.5.4。

```
# 设置 EDAS-Container/Pandora 应用容器版本
ENV UPGRADED_CONTAINER_VERSION 3.5.4
RUN /home/admin/bin/upgradecontainer.sh
```

5. 重新制作并发布应用镜像。

自定义镜像:在镜像中调整IVM启动参数

基于基础镜像中的JVM参数,我们将通过JAVA_OPTS环境变量进行传递,下例中说明了如何修改JVM启动参数的方式:

```
FROM apaas/edas:latest
# 设置JVM参数ENV

JAVA_OPTS="\
-Xmx3550m \
-Xms3550m \
-Xmn2g \
-Xss128k"
```

自定义镜像: SpringBoot类型的应用修改应用配置

基于EDAS基础镜像中的部署FatJar类型的应用,如果需要修改Tomcat的启动配置如:Context Path、启动端口、参数等,将借助APP ARGS环境变量进行传递,举例如下:

```
# 基础镜像中的默认配置为:

# — context path为 "/" ,

# — uri编码为ISO-8859-1

# — 应用启动端口为8080,

# — tomcat处理线程池数量为400

# 您也可以通过复写APP_ARGS这个环境变量来修改默认配置,如下:

ENV APP_ARGS="--server.context-path=/\
--server.tomcat.uri-encoding=ISO-8859-1 \
--server.port=8080 \
--server.tomcat.max-threads=400"
```

在本地构建镜像

1. 从本地命令行进入Dockerfile所在的目录,执行docker build命令构建镜像:

```
docker build -t [标签名称,最好取应用名]:[版本号] .
或docker build -t [标签名称,最好取应用名]:[版本号] -f /path/to/custom_dockerfile_name . #
假如您创建好的Dockerfile在其他位置或名称不为Dockerfile时适用。
```

例如:

```
docker build -t hsf-provider:1.0.0 .
```

2. 然后使用 docker images | grep<镜像标签名称> 命令查看本地打包好的镜像。

上传镜像到镜像仓库

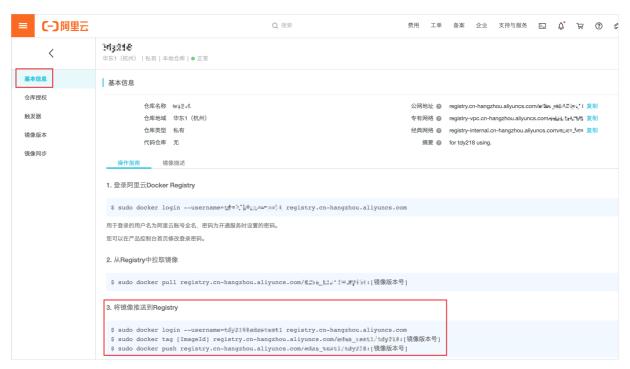
您在本地生成构建好的应用镜像就可以上传到阿里云提供的容器镜像仓库(需要您提前登录<mark>容器镜像仓库控制台</mark>,并根据您EDAS中已经创建或尚未创建的应用所在的地域,在阿里云镜像仓库控制台切换到对应的地域,查看镜像仓库列表,选择一个已经存在或者新创建的镜像仓库来存放您打包好的应用镜像,仓库名称最好是应用名称,便于识别)。

执行指定镜像仓库基本信息页面提供的 "将镜像推送到Registry" 相关命令,将本地镜像上传到该镜像仓库。

```
docker login --username=[当前登录的账号名] registry.[region_id].aliyuncs.com #这一步输入的 密码为阿里云容器镜像服务默认实例 -> 访问凭证页面设置的固定或者临时密码,并非阿里云账号密码。docker tag [本地应用镜像ID] registry.[region_id].aliyuncs.com/[命令空间名称]/[镜像仓库名称]:[镜像版本号] docker push registry.[region_id].aliyuncs.com/[命令空间名称]/[镜像仓库名称]:[镜像版本号]
```

例如:

```
docker login --username=tdy218@gmail.com registry.cn-hangzhou.aliyuncs.com
docker tag 2b64f63601a7 registry.cn-hangzhou.aliyuncs.com/webapps/hsf-provider:1.0.0
docker push registry.cn-hangzhou.aliyuncs.com/webapps/hsf-provider:1.0.0
```



阿里云镜像仓库网络加速: https://cr.console.aliyun.com/#/accelerator

附录:运行时环境变量信息

通过Dockerfile制作自定义镜像时,有一些环境变量是EDAS在运行时会自动注入的环境信息,这些信息如下表:

环境变量Key	描述
POD_IP	POD IP
EDAS_APP_ID	EDAS应用ID
EDAS_ECC_ID	EDAS ECC ID
EDAS_PROJECT_NAME	同EDAS_APP_ID,用于调用链解析
EDAS_JM_CONTAINER_ID	同EDAS_ECC_ID,用于调用链解析
EDAS_CATALINA_OPTS	中间件运行时所需CATALINA_OPTS参数
CATALINA_OPTS	同EDAS_CATALINA_OPTS,默认TOMCAT启动参数
CATALINA_HOME	所安装的TOMCAT路径
PANDORA_LOCATION	所安装的Pandora路径,HSF应用中可见

同时,请勿使用/home/admin/.spas_key/default文件,此文件会在Pod运行时被覆盖。

3.2.2.2. 在K8s集群中使用镜像部署Java微服务应用

EDAS在云原生K8s之上,全面集成了容器服务ACK,支持K8s容器化应用的全生命周期管理。容器服务ACK集成了阿里云虚拟机、存储、网络和安全等能力,拥有云端极佳的K8s容器化应用运行环境。本文介绍了如何在K8s集群中使用镜像部署Java微服务应用。

背景信息

在容器服务K8s集群中部署应用,您需在容器服务Kubernetes版控制台中创建容器服务Kubernetes版集群,并将该集群导入到EDAS控制台,然后在导入的容器服务K8s集群中使用部署包或镜像来部署应用。



前提条件

- 您的阿里云账号已同时开通EDAS和容器服务Kubernetes版,请参见:
 - o 开通EDAS
 - o 首次使用容器服务Kubernetes版
- 创建微服务空间
- 在容器服务Kubernetes版完成角色授权,请参见容器服务默认角色。

步骤一: 创建容器服务Kubernetes集群

登录<mark>容器服务Kubernetes版控制台</mark>,创建容器服务Kubernetes集群。具体操作,请参见快速创建Kubernetes托管版集群。

如果创建Serverless Kubernetes集群,**专有网络**请选择**自动创**建(选择已有网络,创建完毕后请查看集群资源是否包含VPC和交换机资源),**服务发现**请选择**PrivateZone**,以便Serverless Kubernetes集群在导入EDAS后可以使用服务网格。

```
{
 1
          "Version": "1",
 2
          "Statement": [
 3
               {
4
                    "Effect": "Allow",
5
                    "Action": [
                        "edas:*"
 7
                    "Resource": [
9
                        "*"
10
11
                    1
12
13
14
```

步骤二:在EDAS控制台中导入容器服务K8s集群

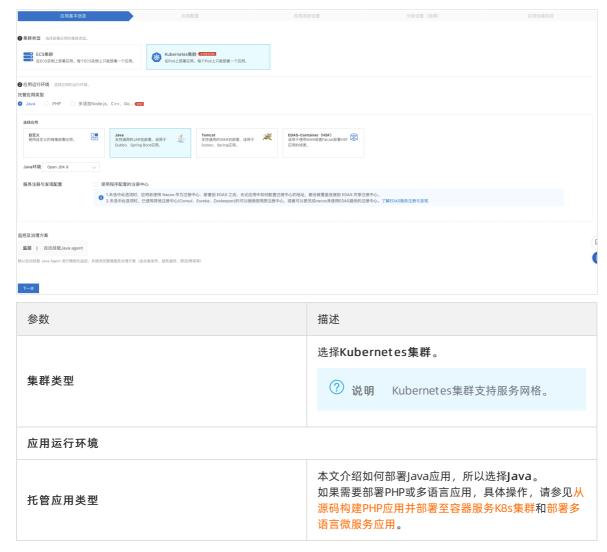
在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Promet heus监控组件(ack-arms-promet heus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,然后单击同步容器服务Kubernetes集群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在**导入Kubernet es集群**对话框,从下拉列表中选择目标微服务空间,根据需要选择是否开启**服务网格**,然后单击**导入**。

当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

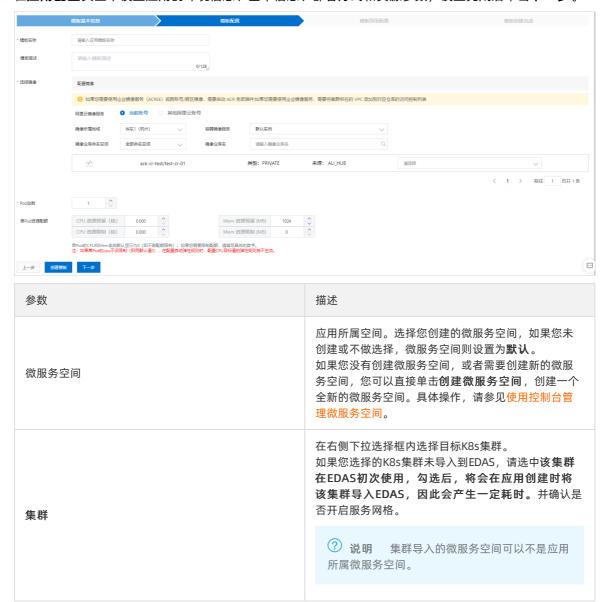
步骤三:在K8s集群中部署应用

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列表** 表页面左上角单击**创建应用**。
- 3. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。



参数	描述
选择应用	应用的运行环境。不同 托管应用类型 ,不同部署包类型,需要选择不同的Java应用环境。本文中选择 自定义 ,即使用镜像部署Java应用。应用运行环境包含在镜像中,无需选择。
服务注册与发现配置	选择注册中心的运维方式。如何选择,请参见选择注册中心运维方式。
监控及治理方案	K8s集群默认自动挂载Java Agent进行精细化监控,并提供完整微服务治理方案(金丝雀发布、服务鉴权、限流降级等)。

4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。



参数	描述
K8s Namespace	K8s Namespace通过将系统内部的对象分配到不同的Namespace中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。 default:没有其他命名空间的对象的默认命名空间。 kube-system:系统创建的对象的命名空间。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。此处以选择default为例。如果您需要自定义创建K8s Namespace,请单击创建K8s Namespace并设置K8s Namespace名称。名称设置要求只能包含数字、小写字母和短划线(-),且首尾只能是字母或数字,长度为1~63个字符。
应用名称	输入应用名称,必须以字母开头,允许数字、字母、 短划线(-)组合。最多支持输入36个字符。
(可选) 应用描述	输入应用描述 , 最大长度为128个字符。
镜像类型	 配置镜像 ■ 阿里云镜像服务选择当前账号。 选择镜像所属地域、容器镜像服务、镜像仓库 命名空间、镜像仓库名和镜像版本。 ■ 阿里云镜像服务选择为其他阿里云账号。 ■ 如您的镜像存放在公开仓库中,那么您配置完整镜像地址即可。 ■ 如您的镜像存放在私有仓库中,那么您需要使用免密插件拉取容器镜像,请参见使用免密组件拉取容器镜像。 ○ Demo镜像 选择EDAS提供的Demo镜像和镜像版本。 ② 说明 专有云暂不支持Demo镜像。 ② 说明 如果您以RAM用户且使用企业版容器镜像仓库里的镜像创建应用时,需要阿里云账号为该RAM用户授权。具体信息,请参见配置仓库的RAM访问控制。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

- 5. (可选)配置应用高级设置。
 - 。 配置调度规则
 - 。 配置启动命令
 - 。 配置环境变量
 - 。 配置持久化存储
 - 。 配置本地存储
 - 。 配置应用生命周期的钩子和探针
 - 。 配置日志收集
 - o 配置Tomcat
 - o 配置Java启动参数
 - o 实现K8s集群应用的限流降级
 - o 配置挂载
- 6. 完成设置后单击创建应用。

应用创建可能需要几分钟,创建过程中,可以通过变更记录查看应用总览跟踪应用的创建流程。创建完成后,返回应用总览页面查看实例Pod的运行状态若为运行中则说明应用发布成功,单击Pod的运行状态可以查看应用实例的工作负载、容器组(Pod)和启动命令等高级配置信息。

后续步骤

应用创建完毕后,可以通过添加公网SLB实现公网访问,添加私网SLB实现同VPC内所有节点够能通过私网负载均衡访问您的应用。相关操作,请参见添加负载均衡SLB或复用负载均衡SLB。

联系我们

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.2.2.3. 在Serverless K8s集群中使用镜像方式部署应用

EDAS在云原生Kubernetes之上,全面集成了容器服务Kubernetes版,支持K8s容器化应用的全生命周期管理。Serverless K8s适合快速伸缩的敏捷业务场景以及单个或多个任务处理的场景。本教程介绍如何在Serverless K8s集群中使用镜像来部署应用。

前提条件

- 您的阿里云账号已同时开通EDAS和容器服务Kubernetes版,请参见:
 - o 开通EDAS
 - o 首次使用容器服务Kubernetes版
- 创建微服务空间
- 在容器服务Kubernetes版完成角色授权,请参见容器服务默认角色。

步骤一: 创建Serverless Kubernetes集群

登录容器服务Kubernetes版控制台,创建Serverless Kubernetes集群。具体操作,请参见创建Serverless Kubernetes集群。

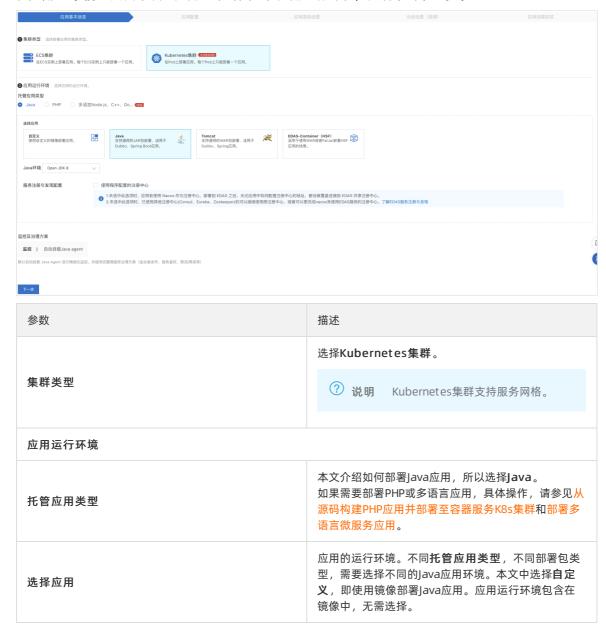
步骤二:在EDAS控制台中导入Serverless Kubernetes集群

在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Prometheus监控组件(ack-arms-prometheus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > Serverless K8s 集群。
- 3. 在顶部菜单栏选择地域,选择要导入集群的目标微服务空间,然后单击同步Serverless K8s集群。
- 4. 在导入的Serverless K8s集群的操作列单击导入。 当Serverless K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示Serverless K8s集群已成功导入到EDAS。

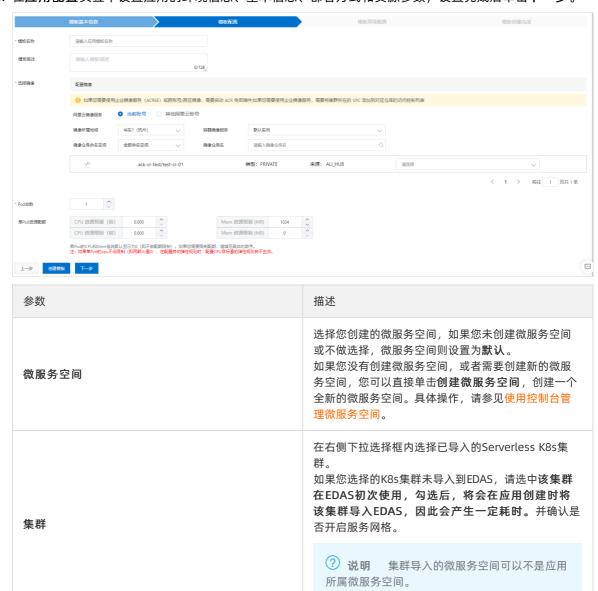
步骤三:在Serverless K8s集群中创建应用

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。



参数	描述
服务注册与发现配置	选择注册中心的运维方式。如何选择,请参见选择注册中心运维方式。
监控及治理方案	K8s集群默认自动挂载Java Agent进行精细化监控,并提供完整微服务治理方案(金丝雀发布、服务鉴权、限流降级等)。

4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。



参数	描述
	K8s Namespace通过将系统内部的对象分配到不同的 Namespace中,形成逻辑上分组的不同项目、小组或 用户组,便于不同的分组在共享使用整个集群的资源 的同时还能被分别管理。
	o default:没有其他命名空间的对象的默认命名空间。
K8s Namespace	。 kube-system:系统创建的对象的命名空间。
	。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。
	此处以选择default为例。 如果您需要自定义创建K8s Namespace,请单击创建 K8s Namespace并设置K8s Namespace名称。名 称设置要求只能包含数字、小写字母和短划线(-), 且首尾只能是字母或数字,长度为1~63个字符。
应用名称	输入应用名称,必须以字母开头,允许数字、字母、 短划线(-)组合。最多可输入36个字符。
应用描述	输入应用描述 ,最大长度为128个字符。
镜像类型	 配置镜像 ■ 阿里云镜像服务:选择为当前账号。选择镜像所属地域、容器镜像服务、镜像仓库命名空间、镜像仓库名和镜像版本。 ■ 阿里云镜像服务:选择为其他阿里云账号。 ■ 如您的镜像存放在公开仓库中,那么您配置完整镜像地址即可。 ■ 如您的镜像存放在私有仓库中,那么您需要使用免密插件拉取容器镜像。 O Demo镜像选择EDAS提供的Demo镜像和镜像版本。 ② 说明 如果您以RAM用户且使用企业版容器镜像仓库里的镜像创建应用时,需要阿里云账号为该RAM用户授权。具体信息,请参见配置仓库的RAM访问控制。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

5. (可选)配置应用高级设置。

- 。 配置启动命令
- 配置环境变量

- 。 配置持久化存储
- 配置应用生命周期的钩子和探针
- o 配置Tomcat
- o 配置lava启动参数
- o 实现K8s集群应用的限流降级
- 6. 完成设置后单击创建应用。

应用创建可能需要几分钟,创建过程中,可以通过变更记录查看应用总览跟踪应用的创建流程。创建完成后,返回应用总览页面查看实例Pod的运行状态若为运行中则说明应用发布成功,单击Pod的运行状态可以查看应用实例的工作负载、容器组(Pod)和启动命令等高级配置信息。

后续步骤

应用创建完毕后,可以通过添加公网SLB实现公网访问,添加私网SLB实现同VPC内所有节点够能通过私网负载均衡访问您的应用。相关操作,请参见添加负载均衡SLB或复用负载均衡SLB。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.2.2.4. 使用API创建Java应用(镜像部署)

本文介绍使用EDAS提供的Java SDK调用API快速创建Java应用(镜像部署)的方法。

前提条件

调用API创建应用前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定应用部署地域。
- 创建微服务空间。具体操作,请参见创建命名空间。 如果已经存在命名空间,您可以调用ListUserDefineRegion接口查询用户自定义命名空间列表,获取目标命名空间的*RegionId*。
- 导入目标集群(假设集群ID为 da60f685-558a-4e00-b549-15e9143d****) 到命名空间。具体操作,请参见导入集群。

如果已经导入集群到目标命名空间,您可以调用ListCluster接口查询集群列表,获取目标集群的ClusterId。

● 已将应用镜像保存在目标镜像仓库中。假设镜像仓库为 *image-demo-project*,镜像地址为 *registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provider:1.0*。

创建lava应用(镜像部署)

以下代码适用于在EDAS K8s环境快速创建应用(镜像部署)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见InsertK8sApplication。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationResponse;
public class InsertK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //要执行API调用的应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       InsertK8sApplicationRequest request = new InsertK8sApplicationRequest();
       //应用名称。
       request.setAppName("edas-SDK0311image");
       //集群ID。
       request.setClusterId("da60f685-558a-4e00-b549-15e9143d****");
       //应用实例数。
       request.setReplicas(2);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setLimitCpu(0);
       request.setLimitMem(0);
       request.setRequestsCpu(0);
       request.setRequestsMem(0);
       //EDAS命名空间对应ID,非默认命名空间须填写。
       request.setLogicalRegionId("cn-hangzhou:doc");
       //镜像仓库和镜像地址。
       request.setImageUrl("registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provi
der:1.0");
       request.setRepoId("image-demo-project");
           InsertK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("AppId=" + response.getApplicationInfo().getAppId() + "\nApp
Name=" + response.getApplicationInfo().getAppName());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行代码,返回示例如下:

```
AppId=d8130adc-2298-49ea-9c64-f19165fa****
AppName=edas-SDK0311image
```

结果验证

您在创建应用后,可以调用Get ChangeOrderInf o接口查看应用变更流程详情,获取创建应用的变更状态。 具体API参数详情,请参见Get ChangeOrderInf o。

调用Get ChangeOrderInf o接口查看应用变更流程详情,返回的执行结果如下:

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断创建应用是否成功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示创建应用失败。您可以根据根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新创建应用。

3.2.3. 转化Deployment为EDAS应用

3.2.3.1. 使用控制台转化Deployment为EDAS应用

在阿里云Kubernet es集群或混合云集群创建的Deployment 应用,可以不用在集群侧运维应用,而选择在EDAS托管该Deployment 应用,借助EDAS的全生命周期管理、监控、运维等能力,大幅降低运维成本。本文介绍如何将K8s集群的Deployment 应用转换为EDAS应用。

前提条件

- 快速创建Kubernetes托管版集群
- 创建无状态工作负载Deployment

背景信息

使用EDAS托管Deployment应用,您需要在容器服务Kubernetes版控制台中创建或注册集群,并将该集群导入到EDAS控制台,然后在导入的集群中一键转换Deployment应用。



② 说明 阿里云Kubernet es集群和混合云集群在导入EDAS后,管理集群的操作流程类似,本文以阿里云容器服务K8s集群为例。

在EDAS控制台中导入容器服务Kubernetes集群

在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Promet heus监控组件(ack-arms-promet heus)。

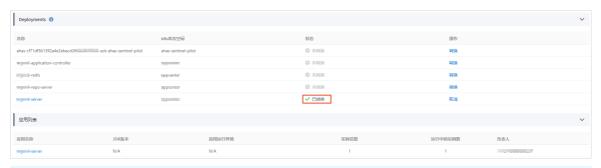
- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,然后单击同步容器服务Kubernetes集群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在**导入Kubernet es集群**对话框,从下拉列表中选择目标微服务空间,根据需要选择是否开启**服务网格**,然后单击**导入**。

当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

在EDAS控制台完成Deployment转换

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在容器服务K8s集群页面的顶部菜单栏选择地域,在页面中选择目标微服务空间。
- 4. 在容器服务K8s集群页面,选择目标集群,单击集群D。
- 5. 在集群详情页面的Deployments区域,单击操作列下的转换。
- 6. 在转换应用对话框,单击确定。
 - ⑦ 说明 转换Deployment应用会导致应用重启,不要在业务高峰期操作。

转换完成后,该Deployment应用状态显示为**已转换**,且在**应用列表**区域出现转换的Deployment应用。



② 说明 当您不再需要EDAS托管你的Deployment应用时,可以取消转换,取消转换 Deployment应用等同于删除应用。

相关步骤

Deployment 应用转换完毕后,和通过JAR包、WAR包和镜像部署的应用完全相同,同样可以使用升级和回滚应用、持续集成和持续部署、应用监控和应用运维等功能。

相关文档,请参见:

- 升级和回滚应用概述(K8s)
- CI/CD概述 (K8s)
- 应用监控概述
- 应用运维概述

3.3. PHP应用

3.3.1. 从源码构建PHP应用并部署至容器服务K8s集群

EDAS现支持从源码构建PHP应用并部署至EDAS的容器服务K8s集群或Serverless K8s集群,并提供了应用扩缩、监控报警、负载均衡和限流降级等能力。部署PHP应用至容器服务K8s集群或Serverless K8s集群的操作步骤类似,本文以部署一个从源码构建的PHP应用至容器服务K8s集群为例来说明。

前提条件

- 您的阿里云账号已同时开通EDAS和容器服务Kubernetes版,请参见:
 - o 开通EDAS
 - o 首次使用容器服务Kubernetes版
- 创建微服务空间
- 在容器服务Kubernetes版完成角色授权,请参见容器服务默认角色。

步骤一: 创建容器服务Kubernetes集群

登录<mark>容器服务Kubernetes版控制台</mark>,创建容器服务Kubernetes集群。具体操作,请参见<mark>快速创建Kubernetes托管</mark>版集群。

如果创建Serverless Kubernetes集群,**专有网络**请选择**自动创建**(选择已有网络,创建完毕后请查看集群资源是否包含VPC和交换机资源),**服务发现**请选择**PrivateZone**,以便Serverless Kubernetes集群在导入EDAS后可以使用服务网格。

```
{
 1
           "Version": "1",
 2
           "Statement": [
 3
               {
 4
                    "Effect": "Allow",
 5
                    "Action": [
 6
                         "edas:*"
 7
                    "Resource":
 9
                         "*"
10
11
12
13
14
```

步骤二:在EDAS控制台中导入容器服务K8s集群

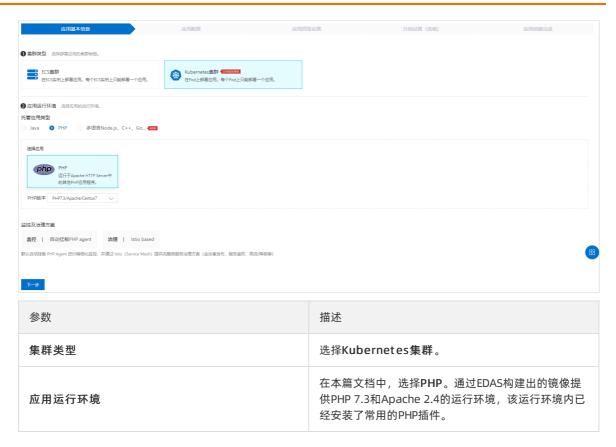
在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Prometheus监控组件(ack-arms-prometheus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,然后单击同步容器服务Kubernetes集群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在**导入Kubernet es集群**对话框,从下拉列表中选择目标微服务空间,根据需要选择是否开启**服务网格**,然后单击**导入**。

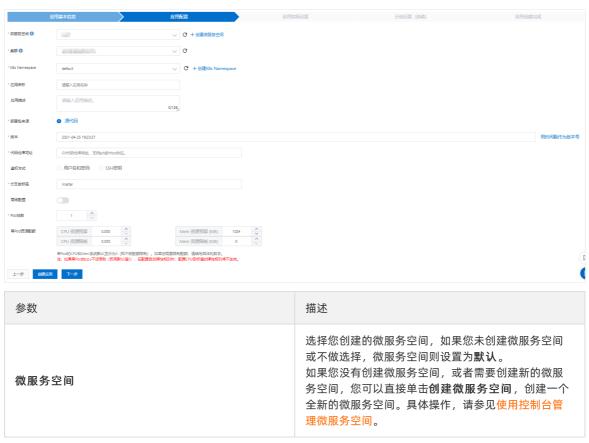
当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

步骤三:在容器服务K8s版集群中部署PHP应用

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列表** 表页面左上角单击**创建应用**。
- 3. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。



4. 在**应用配置**页签中设置应用的环境信息、基本信息、源码及镜像构建相关信息,设置完成后单击下一步。



参数	描述
集群	在右侧下拉列表内选择已导入的容器服务K8s集群。如果您选择的K8s集群未导入到EDAS,请选中该集群在EDAS初次使用,勾选后,将会在应用创建时将该集群导入EDAS,因此会产生一定耗时。并确认是否开启服务网格。 ② 说明 集群导入的微服务空间可以不是应用所属微服务空间。
K8s Namespace	K8s Namespace通过将系统内部的对象分配到不同的 Namespace中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。 odefault:没有其他命名空间的对象的默认命名空间。 kube-system:系统创建的对象的命名空间。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。 此处以选择default为例。 如果您需要自定义创建K8s Namespace,请单击创建 K8s Namespace并设置K8s Namespace名称。名称设置要求只能包含数字、小写字母和短划线(-),且首尾只能是字母或数字,长度为1~63个字符。
应用名称	输入应用名称,必须以字母开头,允许数字、字母、 短划线(-)组合。最大长度为36个字符。
应用描述	输入应用描述 , 最大长度为128个字符。
部署包来源	默认选择 源代码 。
版本	可选择时间戳作为版本,或自定义版本号。
代码仓库地址	输入您的PHP应用的代码仓库的公网访问地址,现支持 HTTPS和Git两种协议。
鉴权方式	如果代码仓库是私有,需要配置鉴权方式:

参数	描述
分支或标签	输入代码库的分支或标签。
高级配置(可选)	打开高级配置,完成下述参数的配置: • 克隆子模块: 子模块允许您将一个Git仓库作为另一个Git仓库的子目录。它能让您将另一个仓库克隆到自己的项目中,同时还保持提交的独立。 • 工作目录: 工作目录是保存您目前正在处理档案的目录,Git相关的操作都会在这个目录下完成。 • 产出物路径: 指定Git的产出文件的存放目录。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

- 5. (可选)配置应用高级设置,设置完成后单击创建应用。
 - i. 设置环境变量。

在创建应用过程中,将所填环境变量注入到即将生成的容器中,这样可以避免重复添加常用环境变量。

- 如果您需要设置JVM堆内存、JVM属性参数、Java Agent等参数,可以在设置环境变量这一步添加相关参数。
 - 变量名: CATALINA_OPTS。
 - 变量值: [您需要添加的参数] \$(EDAS_CATALINA_OPTS)。



- 如您使用MySQL镜像,可以参考以下环境变量:
 - MYSQL_ROOT_PASSWORD(必选项):用于设置MySQL的root密码。
 - MYSQL_USER和MYSQL_PASSWORD(可选项):用于添加除root之外的账号并设置密码。
 - MYSQL_DATABASE(可选项):用于设置生成容器时需要新建的数据库。
- 如使用其它类型的镜像,请根据实际需求进行配置。

ii. 设置持久化存储。

在阿里云容器服务Kubernetes集群中,原生的Volume对象所对应的物理储存为非持久化的,它的生命周期与Kubernetes Pod一致,是一个暂态的储存对象。使用持久化的储存服务(文件储存NAS),可以永久保存实例数据,应用在升级或迁移后不丢失。

② 说明 设置持久化存储前,请确保您的EDAS账号已开通文件存储NAS。使用NAS请保证当前账户拥有足够的余额或计费方式为后付费。NAS服务的计费规则,请参见计费项。

参数描述

NAS文件存储

② 说明 使用NAS文件存储服务时,请确保您的集群安装了flexvolume组件,对于Serverless集群请确保NAS挂载目录存在。NAS存储按实际使用量付费,详细收费标准,请参见NAS计费说明。

存储类型	默认为 NAS文件存储 ,不可配置。
存储服务类型	目前仅支持SS D性能型 ,不可配置。
选择NAS	■ 新购一个NAS:选择NAS挂载目录和本地挂载目录。单一地域内NAS文件系统上限为10个,若已超过10个会创建失败。如有特殊需求,请提交工单。 ■ 使用已有NAS:选择一个已有的NAS。您最多可以创建两个挂载点。不符合要求的NAS将不会显示在下拉列表中。
挂载目录	分别设置NAS挂载目录和容器挂载目录。

PVC挂载

② 说明 通过修改YAML方式配置的PVC挂载,在部署应用时将不会显示已存在的PVC挂载,仅部署时会保留这些配置。推荐您在使用EDAS部署应用时配置PVC挂载。

选择PVC	在下拉列表中选择PVC。如果没有可选的PVC,请参见创建持久化存储卷声明。
挂载配置	分别设置 挂载目录 和 挂载模式 。

iii. 设置应用生命周期管理。

参数	描述
PostStart配置	一种容器钩子。该钩子在容器被创建后立刻触发,通知容器它已经被创建。该钩子不需要向其所对应的hook handler传入任何参数。如果该钩子对应的hook handler执行失败,则该容器会终止运行,并根据该容器的重启策略决定是否要重启该容器。更多信息,请参见Container Lifecycle Hooks。
PreStop配置	一种容器钩子。该钩子在容器被删除前触发,其所对应的hook handler必须在删除该容器的请求发送给Docker daemon之前完成。在该钩子对应的hook handler完成后不论执行的结果如何,Docker daemon会发送一个SIGT ERN信号量给Docker daemon来删除该容器。更多信息,请参见Container Lifecycle Hooks
Liveness配置	一种探测容器状态的探针,探测应用是否处于健康 状态。如果不健康,则删除重建容器。更多信息, 请参见Pod Lifecycle。
Readiness配置	一种探测容器状态的探针,探测应用是否启动完成并且处于正常服务状态。如果不正常,则更新容器的状态。更多信息,请参见Pod Lifecycle。

6. 完成设置后单击创建应用,然后在应用创建完成页签单击确定创建应用。 应用创建可能需要几分钟,创建过程中,可以在变更记录页面查看应用的创建流程。创建完成后,返 回应用总览页面查看实例Pod的运行状态若为运行中则说明应用发布成功,单击Pod的运行状态可以查 看应用实例的工作负载、容器组(Pod)和启动命令等高级配置信息。

常见问题

怎么查看镜像构建出错日志?

在进行应用部署的时候,如果镜像构建失败,可以在变更记录页面查看日志。具体操作,请参见<mark>查看应用变</mark>更。



后续步骤

应用创建完毕后,可以通过添加公网SLB实现公网访问,添加私网SLB实现同VPC内所有节点够能通过私网负载均衡访问您的应用。相关操作,请参见添加负载均衡SLB或复用负载均衡SLB。

联系我们

如果您在部署EDAS多语言微服务应用过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23307994 加入钉钉群进行反馈。



3.3.2. 如何在使用源码部署应用时配置SSH公钥

在EDAS里通过源码部署应用的时候,如果采用Git协议并且鉴权方式选择为SSH密钥时,您需在代码仓库配置EDAS提供的SSH公钥,本文以在阿里云Code上配置公钥为例来说明。

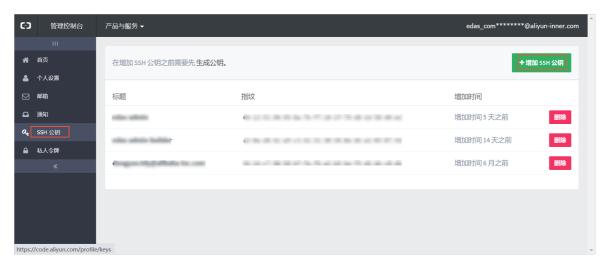
步骤一:复制EDAS提供的公钥

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在**应用基本信息**页面中设置应用的**集群类型为Kubernet** es**集群**,并在**应用运行环境**区域选择PHP,然后单击下一步。
- 4. 在应用配置配置向导页面中选择鉴权方式为SSH密钥,然后单击SSH公钥右侧的一键复制。



步骤二: 在代码仓库配置公钥

- 1. 登录阿里云Code控制台。
- 2. 在左侧导航栏单击设置。
- 3. 在左侧导航栏单击SSH公钥,然后在页面右上角单击增加SSH公钥。



4. 在公钥文本框中粘贴步骤中拷贝的公钥,设置标题,然后单击增加公钥。



配置完成后可在SSH公钥页面查看增加的公钥。

相关操作

如需删除创建的公钥,可在SSH公钥页面单击公钥最右侧的删除,然后根据页面提示完成删除。

3.4. 多语言应用

3.4.1. 部署多语言微服务应用

随着Python、Node.js等语言的快速发展,多语言微服务应用越来越多。EDAS能够通过服务网格支持部署多语言微服务应用,且提供应用托管和服务发现、链路追踪、负载均衡等服务治理能力。

背景信息

应用从最初的单体架构演变到目前的微服务架构,在带来便利的同时也大大增加了服务部署、运维的复杂度。而微服务本身可以是任意语言开发的,在部署多语言服务后,如何对多语言的微服务提供通用的链路追踪、服务发现、负载均衡等能力,一种做法是提供多语言的SDK,一种是服务网格。SDK对应用有侵入性,而服务网格在无侵入性的同时也能提供服务发现、负载均衡、链路追踪等能力,EDAS对多语言的支持正是采用的后者——服务网格。

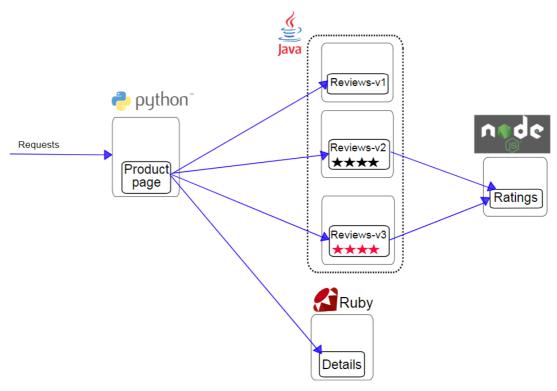
服务网格是致力于解决服务间通讯的基础设施层。它负责在现代云原生应用程序的复杂服务拓扑中可靠地传递请求,通常是通过一组轻量级网络代理,与应用程序部署在一起来实现,而无需感知应用程序本身。

示例场景说明

示例应用BookInfo模仿在线书店的一个分类,显示一本书的信息。页面上会显示一本书的描述,书籍的细节(ISBN、页数等),以及关于这本书的一些评论。

BookInfo是一个异构应用,几个微服务应用是由不同的语言编写的。这些服务构成了一个有代表性的服务网格的例子:由多个服务、多个语言构成,并且Reviews服务具有多个版本。

多语言服务架构



BookInfo应用包含四个单独的服务:

- Productpage: 为Python服务,会调用Details和Reviews两个服务,用来生成页面。同时,Productpage 还包含登录和登出功能。
- Details: 为Ruby服务,包含了书籍的信息。
- Reviews: 为Java服务,包含了书籍相关的评论,还会调用Ratings服务。Reviews包含3个版本:
 - ∘ v1版本不会调用Ratings服务。
 - v2版本会调用Ratings服务,并使用1到5个黑色星形图标来显示评分信息。
 - v3版本会调用Ratings服务,并使用1到5个红色星形图标来显示评分信息。
- Ratings: 为Node.js服务,包含了由书籍评价组成的评级信息。

前提条件

将示例应用制作成镜像,并上传到阿里云镜像仓库。上传镜像的具体操作,请参见构建仓库与镜像。

示例应用下载地址: BookInfo Sample。

步骤一: 创建容器服务Kubernetes集群

登录<mark>容器服务Kubernetes版控制台</mark>,创建容器服务Kubernetes集群。具体操作,请参见<mark>快速创建Kubernetes托管</mark>版集群。

如果创建Serverless Kubernetes集群,**专有网络**请选择**自动创建**(选择已有网络,创建完毕后请查看集群资源是否包含VPC和交换机资源),**服务发现**请选择**PrivateZone**,以便Serverless Kubernetes集群在导入EDAS后可以使用服务网格。

```
{
 1
           "Version": "1",
           "Statement": [
 3
 4
               {
                    "Effect": "Allow",
 5
                    "Action": [
 6
 7
                         "edas:*"
                    "Resource":
 9
                         "*"
10
11
12
13
14
```

步骤二:在EDAS控制台导入容器服务K8s集群

在EDAS控制台导入容器服务K8s集群时,会默认安装应用防护(限流降级)组件(ack-ahas-sentinel-pilot)、ARMS监控组件(ack-arms-pilot)以及Prometheus监控组件(ack-arms-prometheus)。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择资源管理 > 容器服务K8s集群。
- 3. 在顶部菜单栏选择地域,选择要导入集群的目标微服务空间,然后单击**同步容器服务Kubernetes集** 群。
- 4. 在导入的容器服务K8s集群的操作列单击导入。
- 5. 在导入Kubernetes集群对话框选择微服务空间,并打开服务网格开关,然后单击导入。

? 说明

- 如果您未创建微服务空间,在此步骤可不选择微服务空间,使用默认微服务空间。
- 如果您的集群已经导入且未开启服务网格,可在集群列表的**服务网格**列单击开关来开启。
- 开启服务网格时,默认会创建两个私网规格SLB实例(slb.s1.small)实现管控,并暴露两个 SLB实例的端口(6443端口和15011端口)。更多信息,请参见<mark>背景信息</mark>。 默认生成的两个小规格SLB实例(slb.s1.small)会产生费用。具体收费标准,请参见SLB收费 标准。

当容器服务K8s集群状态显示为**运行中**,并且**导入状态**显示为**导入成功**时,表示容器服务K8s集群已成功导入到EDAS。

(可选)

步骤三:添加链路追踪功能

EDAS控制台部署多语言应用采用的是基于Istio的服务网格链路追踪的监控方案。Istio代理虽能够自动发送Span信息,但应用程序仍然需要携带HTTP表头信息,以实现在发送Span信息时将Span信息正确地关联到单个跟踪中。

应用程序需要携带以下HTTP表头,并将其从传入请求传播到任意传出请求。

- x-request-id
- x-b3-traceid
- x-b3-spanid
- x-b3-parentspanid
- x-b3-sampled
- x-b3-flags
- x-ot-span-context

此处仅介绍示例多语言应用中部分服务的应用程序中携带的HTTP表头信息。

以Python语言实现的Product page服务的应用程序中携带的HTTP表头如下:

```
def getForwardHeaders(request):
   headers = {}
    \# x-b3-*** headers can be populated using the opentracing span
   span = get_current_span()
   carrier = {}
    tracer.inject(
       span context=span.context,
       format=Format.HTTP HEADERS,
       carrier=carrier)
   headers.update(carrier)
   incoming_headers = ['x-request-id']
    for ihdr in incoming_headers:
       val = request.headers.get(ihdr)
       if val is not None:
           headers[ihdr] = val
    return headers
```

以Java语言实现的Reviews服务的应用程序中携带的HTTP表头如下:

步骤五:在容器服务K8s版集群中部署多语言应用

您需要将场景示例中的几个服务以应用的形式分别部署到EDAS中。下面介绍单个服务的部署过程。

- ② 说明 多语言应用目前仅支持使用镜像部署。
- 1. 登录EDAS控制台。
- 3. 在应用基本信息页面中设置应用的集群类型和应用运行环境,然后单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	在托管应用类型区域选择多语言Node.js、C++、Go。

4. 在应用配置页签中设置应用的环境信息、基本信息及镜像相关配置,设置完成后单击下一步。

参数	描述
微服务空间	选择您导入的K8s集群所在的微服务空间,如果您未创建微服务空间或不做选择,将默认选择默认微服务空间。如果您没有创建微服务空间,或者需要创建新的微服务空间,您可以直接单击 创建微服务空间 ,创建一个全新的微服务空间。具体操作,请参见使用控制台管理微服务空间。

参数	描述
集群	在右侧下拉列表内选择已导入的容器服务K8s集群。如果您选择的K8s集群未导入到EDAS,请选中该集群在EDAS初次使用,勾选后,将会在应用创建时将该集群导入EDAS,因此会产生一定耗时。并确认是否开启服务网格。 ② 说明 集群导入的微服务空间可以不是应用所属微服务空间。
K8s Namespace	K8s Namespace通过将系统内部的对象分配到不同的Namespace中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。 odefault:没有其他命名空间的对象的默认命名空间。 kube-system:系统创建的对象的命名空间。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。此处以选择default为例。如果您需要自定义创建K8s Namespace,请单击创建K8s Namespace并设置K8s Namespace名称。名称设置要求只能包含数字、小写字母和短划线(-),且首尾只能是字母或数字,长度为1~63个字符。
应用名称	输入应用名称,必须以字母开头,允许数字、字母、 短划线(-)组合。最大长度为36个字符。
应用描述	输入应用描述 , 最大长度为128个字符。
版本	可在右侧单击 生成版本号 来生成版本,也可按照页面 提示自定义版本号。
镜像类型	选择 配置镜像 ,然后选择镜像所属地域和目标镜像, 并设置镜像版本。
Pod总数	设置该应用要部署的Pod个数。
单Pod资源配额	设置单个Pod的CPU和内存,如果您需要限额,请填写 具体的数字,使用默认值0则代表不限额。

5. 配置应用高级设置,完成配置后单击创建应用。

i. 配置服务网格。

参数	描述
协议	在下拉列表选择协支持的协议。
服务名	请输入应用提供的服务名,要和应用代码中的服务 名一致,以保证服务能成功注册和被调用。
服务端口	请输入应用提供的服务端口,要和应用代码中的服 务端口一致,以保证服务能成功注册和被调用。

- ii. (可选)配置以下高级设置。
 - 配置调度规则
 - 配置启动命令
 - 配置环境变量
 - 配置持久化存储
 - 配置本地存储
 - 配置应用生命周期的钩子和探针
 - 配置日志收集
 - 配置Java启动参数
- 6. 完成设置后单击**创建应用**,然后在**应用创建完成**页签单击**确定创建应用**。 应用创建可能需要几分钟,创建过程中,可以在页面上方单击**查看详情**,跳转到应用的**变更记录**页面 查看部署进度及相关日志。创建完成后,返回**应用总览**页面查看实例Pod的运行状态若为运行中则说明 应用发布成功,单击Pod的运行状态可以查看应用实例的工作负载、容器组(Pod)和启动命令等高 级配置信息。
- 7. 重复执行上述步骤, 部署示例应用的其他服务。

结果验证

服务部署完成后,访问主服务,页面上会显示一本书的描述,书籍的细节(ISBN、页数等),以及关于这本书的一些评论。同时还能进行登录、登出操作。

- 1. 在应用总览页面的访问方式配置区域单击负载均衡(公网)右侧的图标。+
- 2. 在负载均衡 (公网) 对话框中设置SLB和监听参数, 然后单击确认。
 - i. 在**选择SLB**右侧的下拉列表中选择*新建SLB*。 如果您有SLB实例,可以在列表中选择已有的SLB实例。
 - ii. 选择协议,然后在协议右侧单击**添加新的监听**,然后将SLB端口和容器端口分别设置 为*80*和*18082*。
 - ② 说明 如果您选择了HTTPS协议,则还需要选择SSL证书。

添加公网SLB大概需要30秒。添加完成后,**访问方式配置区域的负载均衡(公网)**右侧会显示公网SLB的地址,格式为 SLB实例IP:端口 。

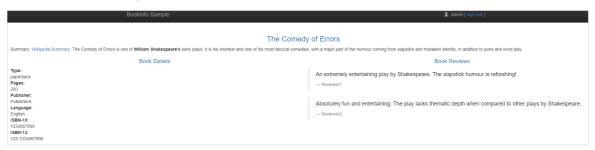
3. 复制公网SLB地址,在浏览器中粘贴该公网地址进入到示例应用(在线书店)首页。



- 4. 在客户端应用首页单击Normal user或者Test user。
- 5. 在页面顶部单击Sign in,输入User Name和Password,然后单击Sign in登录示例应用。 本示例的User Name和Password均为*admin*。



在该示例多语言应用里,您可看到书籍的详细信息和相关评论。



应用监控

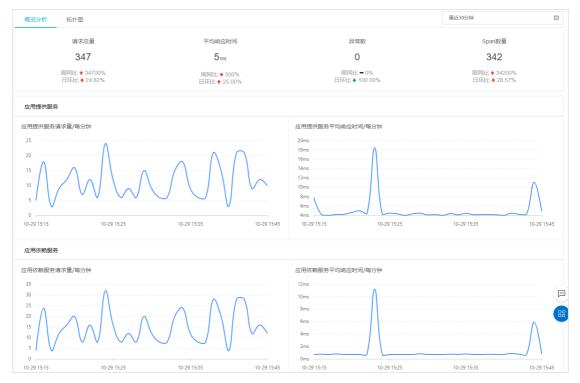
示例多语言应用部署完成后,您可以查看应用的健康状况关键指标,包括请求总量、平均响应时间等总体指标,应用所提供的服务、所依赖的服务的相关指标;以及CPU使用量、内存使用量等系统信息。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列**表页面单击具体的应用名称。
- 3. 查看应用的健康状况指标。
 - i. 在左侧导航栏单击**应用总览**。

75

ii. 在**应用总览**页面右侧区域单击概览分析页签。

您可在**概览分析**页签查看该应用的请求总量、平均响应时间等总体指标和应用所提供的服务、所依赖的服务的相关指标。更多信息,请参见<mark>概览分析</mark>。



iii. 在**应用总览**页面右侧区域单击拓扑图页签。

您可在**拓扑图**页签查看上下级组件及调用关系,以及请求数、响应时间和错误率等指标。更多信息,请参见<mark>应用拓扑</mark>。



- 4. 查看应用的系统使用量情况。
 - i. 在左侧导航栏选择监控 > Prometheus。

ii. 在Promet heus页面左上角选择目标namespace和Pod,然后在右上角选择目标时间段。您可在该页面查看该应用的Pod IP地址、Pod状态、Pod容器,以及Pod CPU使用率、Pod内存例

您可在该页面查看该应用的Pod IP地址、Pod状态、Pod容器,以及Pod CPU使用率、Pod内存使用率等系统信息。更多信息,请参见大盘列表。



EDAS多语言应用交流群

如果您在部署EDAS多语言微服务应用过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23307994 加入钉钉群进行反馈。

3.4.2. 使用API在K8s集群中创建多语言应用

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中创建多语言应用。

前提条件

调用API创建多语言应用前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定应用部署地域,假设为cn-hangzhou。
- 创建微服务空间。具体操作,请参见创建命名空间。
 如果已经存在命名空间,您可以调用ListUserDefineRegion接口查询用户自定义命名空间列表,获取目标命名空间的RegionId。
- 导入目标集群(假设集群ID为 da60f685-558a-4e00-b549-15e9143d****) 到命名空间。具体操作,请参见导入集群。
 - 如果已经导入集群到目标命名空间,您可以调用ListCluster接口查询集群列表,获取目标集群的ClusterId。
- 已将多语言应用镜像保存在目标镜像仓库中。假设镜像仓库为*image-demo-project*,其中一个应用的镜像地址为*regist ry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/product page-python:v1*。

创建多语言应用

以下代码适用于在EDAS K8s环境快速创建多语言应用。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见InsertK8sApplication。

② 说明 本示例以创建一个名为 *Product page*的应用为例,如需创建多语言中的其他应用,请修改应用名称和镜像地址并重新运行程序。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.InsertK8sApplicationResponse;
public class InsertK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun_user_ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //需要创建应用的目标地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       InsertK8sApplicationRequest request = new InsertK8sApplicationRequest();
       //自定义设置应用名称。必须以字母开头,允许数字、字母、短划线(-)组合。最多可输入36个字符。
       //此处以Productpage为例,应用名称在微服务空间内必须唯一。
       request.setAppName("Productpage");
       //集群ID。
       request.setClusterId("da60f685-558a-4e00-b549-15e9143d****");
       //应用实例数。
       request.setReplicas(2);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setLimitCpu(0);
       request.setLimitMem(0);
       request.setRequestsCpu(0);
       request.setRequestsMem(0);
       //EDAS微服务空间对应ID,非默认命名空间须填写。
       request.setLogicalRegionId("cn-hangzhou:doc");
       //镜像仓库和镜像地址。
       request.setImageUrl("registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/produ
ctpage-python:v1");
       request.setRepoId("image-demo-project");
           InsertK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("AppId=" + response.getApplicationInfo().getAppId() + "\nApp
Name=" + response.getApplicationInfo().getAppName());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
```

运行代码,返回示例如下:

> 文档版本: 20220708 77

```
AppId=3a2a3a70-fcb2-491c-a8cf-fd68097a****
AppName=Productpage
```

结果验证

您在创建应用后,可以调用Get ChangeOrderInf o接口查看应用变更流程详情,获取创建应用的变更状态。 具体API参数详情,请参见Get ChangeOrderInf o。

调用Get Change OrderInf o接口查看应用变更流程详情,返回的执行结果如下:

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断创建应用是否成功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示创建应用失败。您可以根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新创建应用。

3.5. 高级配置

3.5.1. 配置调度规则

在容器服务K8s集群中创建或部署应用时,您可以根据实际需求修改应用实例的调度规则,将Pod实例尽可能多可用区和多节点部署,来增强应用的可用性。

在创建应用时配调度规则

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node、C++、Go:适用于任何语言的使用自定义镜像来部署的微服务应用。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开调度规则,并根据您的业务需求设置调度模式和规则。
 - 选择内置配置:

可用区调度:

- 默认调度:实例默认按照集群资源情况调度,所有实例有可能部署于单可用区,可用性一般。
- 尽量多可用区部署:实例会利用反亲和性,尽可能分散部署到多可用区,可用性高。

节点调度:

- 默认调度:实例默认按照节点资源负载情况调度。
- 尽量多节点部署:同一应用的实例,将尽量打散到不同节点上部署,以实现应用高可用部署。
- 选择自定义配置:

如果内置的按节点与可用区的配置调度规则无法满足您对于复杂场景的需求时,您可以使用EDAS提供的自定义配置功能,根据实际业务场景来设计调度规则。

- 添加节点亲和性调度规则
- 添加应用亲和性调度规则
- 添加应用反亲和性调度规则
- 添加调度容忍规则

添加节点亲和性调度规则

> 文档版本: 20220708 79

参数名	描述
选择器	您可以通过配置 节点标签名、操作符号和节点标签 值来指定节点的调度规则。
权重	权重值允许设置0-100,默认权重值为1。

添加应用亲和性调度规则

参数名	描述
K8s 命名空间	应用关联的K8s命名空间。
拓扑域	K8s 集群中为方便管理划分的一种"位置"关系,如节点、可用区、区域等。
选择器	您可以通过配置 应用标签名、操作符号 和 应用标签 值 来指定调度约束条件。
权重	权重值允许设置0-100,默认权重值为1。

添加应用反亲和性调度规则

参数名	描述
K8s 命名空间	应用关联的K8s命名空间。
拓扑域	K8s 集群中为方便管理划分的一种"位置"关系,如节点、可用区、区域等。
选择器	您可以通过配置 应用标签名、操作符号和应用标签 值 来指定调度约束条件。
权重	权重值允许设置0-100,默认权重值为1。

添加调度容忍规则

参数名	描述
标签名	调度容忍规则作用的对象。
操作符号	操作符定义了其对调度规则的作用。
标签值	标签对应的值。调度容忍是应用由 标签名、标签 值和操作符以及效果这四部分去匹配合适的污点。
效果	匹配到节点对应污点后,应用运行的效果。
容忍时间(秒)	匹配到节点对应污点后,应用在节点上可以继续运行的时间。若不指定该字段则表示应用可以一直运行在该节点。当操作符号取值为NoExecute时,您才可以配置容忍时间。

- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配调度规则

您可在创建应用时配置调度规则,也可以在部署应用时配置或更新调度规则。

- 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开**调度规则**,根据您的需求完成配置后单击**确定**。
 - 选择内置配置:

可用区调度:

- 默认调度:实例默认按照集群资源情况调度,所有实例有可能部署于单可用区,可用性一般。
- 尽量多可用区部署:实例会利用反亲和性,尽可能分散部署到多可用区,可用性高。

节点调度:

- 默认调度:实例默认按照节点资源负载情况调度。
- 尽量多节点部署: 同一应用的实例,将尽量打散到不同节点上部署,以实现应用高可用部署。
- 选择自定义配置:

如果内置的按节点与可用区的配置调度规则无法满足您对于复杂场景的需求时,您可以使用EDAS提供的自定义配置功能,根据实际业务场景来设计调度规则。

- 添加节点亲和性调度规则
- 添加应用亲和性调度规则
- 添加应用反亲和性调度规则
- 添加调度容忍规则

添加节点亲和性调度规则

参数名	描述
选择器	您可以通过配置 节点标签名、操作符号和节点标签 值 来指定节点的调度规则。
权重	权重值允许设置0-100,默认权重值为1。

添加应用亲和性调度规则

参数名	描述
K8s 命名空间	应用关联的K8s命名空间。

参数名	描述
拓扑域	K8s 集群中为方便管理划分的一种"位置"关系,如节点、可用区、区域等。
选择器	您可以通过配置 应用标签名、操作符号 和 应用标签 值 来指定调度约束条件。
权重	权重值允许设置0-100,默认权重值为1。

添加应用反亲和性调度规则

参数名	描述
K8s 命名空间	应用关联的K8s命名空间。
拓扑域	K8s 集群中为方便管理划分的一种"位置"关系,如节点、可用区、区域等。
选择器	您可以通过配置 应用标签名、操作符号和应用标签 值 来指定调度约束条件。
权重	权重值允许设置0-100,默认权重值为1。

添加调度容忍规则

参数名	描述
标签名	调度容忍规则作用的对象。
操作符号	操作符定义了其对调度规则的作用。
标签值	标签对应的值。调度容忍是应用由 标签名、标签 值和操作符以及效果这四部分去匹配合适的污点。
效果	匹配到节点对应污点后,应用运行的效果。
容忍时间(秒)	匹配到节点对应污点后,应用在节点上可以继续运行的时间。若不指定该字段则表示应用可以一直运行在该节点。当操作符号取值为NoExecute时,您才可以配置容忍时间。

□ 注意 单击确定后,该应用将会被重启,请在业务较少的时间段进行。

3.5.2. 配置启动命令

在EDAS控制台使用镜像、JAR包或WAR包部署应用至K8s集群时,如果您精通Docker和JVM,并且希望在启动前进行特殊配置或不采用EDAS中预设的启动参数,您可设置容器启动和运行时需要的命令。

背景信息

制作镜像时,已通过Dockerfile文件中的ENTRYPOINT或CMD配置了容器的启动配置,启动时所配置的内容会优先被执行。

□ 注意

- 如果您不了解原Dockerfile镜像的CMD或ENTRYPOINT内容,不建议您自定义启动命令和启动参数,错误的自定义命令将导致应用创建失败。
- 由于Docker运行时仅支持一条ENTRYPOINT命令,所以在EDAS控制台中设置的启动命令将会覆盖 在制作容器Docker镜像中设置的ENTRYPOINT和CMD命令。

例如在Dockerfile中配置 ENTRYPOINT: [nginx, '-g', 'daemon off;'] 命令,该命令将会在容器启动时被第一个执行。

```
FROM ubuntu
ENTRYPOINT [nginx, '-g', 'daemon off;']
```

启动命令配置可以在创建、部署应用过程中设置,也可以在应用部署完成后设置。本文介绍在不同场景下设置启动命令的步骤。

- 在创建应用时配置启动命令的操作步骤,请参见在创建应用时配置启动命令。
- 如果您在应用部署完成后配置启动命令,那么该应用将会被重启,请在业务较少的时间段进行。具体操作,请参见在更新应用时配置启动命令。

在创建应用时配置启动命令

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes 集群 。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开启动命令,并根据您的业务需求设置启动命令。



以 ENTRYPOINT: [nginx, '-g', 'daemon off;'] 为例。

参数	描述
启动命令	输入启动命令,例如: nginx 。
启动参数	每个参数占据一行。例如:在文本框中输入 -g , 然后单击 添加 ,在新的参数行中增加输入参数 daemo n off 。

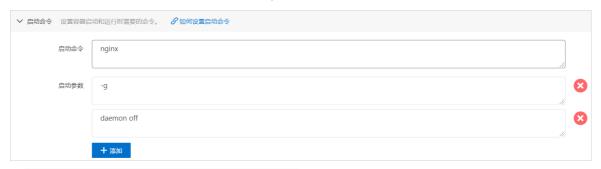
- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置启动命令

您可在创建应用时配置启动命令,也可以在部署应用时配置或更新启动命令。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开启动命令,根据您的需求完成配置后单击确定。



以 ENTRYPOINT: [nginx, '-g', 'daemon off;'] 为例。



☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

杳看执行结果

方法一·

在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。

● 方法二:

在**应用总览**页面单击**运行状态**右侧的Pod运行状态链接,在容**器组(Pod)**区域查看Pod的**状态**,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

3.5.3. 配置环境变量

应用在系统中运行需要配置特定的环境变量,例如Java应用程序在配置JAVA_HOME及其PATH后,其相关命令才可以被执行。本文介绍了如何在EDAS控制台上为应用配置环境变量。

背景信息

目前EDAS控制台支持三种环境变量类型:

- 自定义: 创建应用时自定义环境变量。
- 配置项:提前配置好的一类不需要加密的配置信息,应用需要和配置项在相同地域、相同集群、相同K8s 命名空间,否则无法使用配置项。

> 文档版本: 20220708 85

创建配置项的具体操作,请参见创建配置项。

● 保密字典:提前配置好的一类需要加密的密码和证书等信息,应用需要和保密字典在相同地域、相同集群、相同K8s命名空间,否则无法使用保密字典。 创建保密字典的具体操作,请参见创建保密字典。

在创建应用时设置环境变量

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node.js、C++、Go:适用于在K8s集群中使用自定义的镜像部署应用。此处未介绍多语言应用的配置流程,如需了解更多信息,请参见在K8s环境中通过镜像部署多语言微服务应用。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开环境变量,并根据您的业务需求设置环境变量。

在创建应用过程中,将所填环境变量注入到即将生成的容器中,这样可以避免重复添加常用环境变量。目前EDAS支持自定义、配置项和保密字典三种配置环境变量的方式,同时还支持单个添加和批量添加的方式,请根据实际需求选择配置方式:

○ 单个添加环境变量。

■ 自定义配置环境变量。



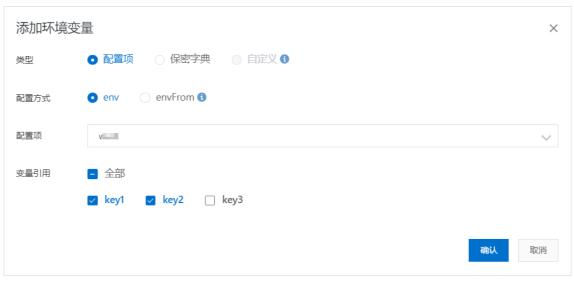
■ 使用配置项配置环境变量。 在**变量引用**参数下方选择目标配置项。一个配置项可能存在多个键值信息,请选择正确的键值信息。 息。



■ 使用保密字典配置环境变量。 在**变量引用**参数下方选择目标保密字典。一个保密字典可能存在多个键值信息,请选择正确的键值信息。



○ 批量添加环境变量。



参数	描述
类型	根据需要选择 配置项 或 保密字典,自定义 类型无法 批量添加。
	支持env和envFrom两种配置方式。 ■ env: 选择目标配置项后,还需要选择配置项中的目标变量。 ■ envFrom:选择目标配置项后,不能选择变量,默认选中全部变量。
配置方式	② 说明 ■ 批量添加环境变量时,最大支持添加 100个环境变量。 ■ 通过编辑YAML添加envFrom方式导入 的环境变量数据,也会生效,但是不会 在此显示。
配置项	从下拉列表中选择目标配置项。
变量引用	选择env配置方式时需要选择变量。

- ② 说明 批量添加配置项的变量后,还需要根据实际需要设置环境变量的名称。
- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创 建应用**。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置环境变量

如果您在创建应用时没有配置环境变量,可以在部署应用时进行配置。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开**环境变量**,根据您的需求完成配置后单击**确定**。 在创建应用过程中,将所填环境变量注入到即将生成的容器中,这样可以避免重复添加常用环境变量。 目前EDAS支持**自定义、配置项和保密字典**三种配置环境变量的方式,同时还支持单个添加和批量添加的方式,请根据实际需求选择配置方式:
 - 单个添加环境变量。
 - 自定义配置环境变量。



■ 使用配置项配置环境变量。

在**变量引用**参数下方选择目标配置项。一个配置项可能存在多个键值信息,请选择正确的键值信息。



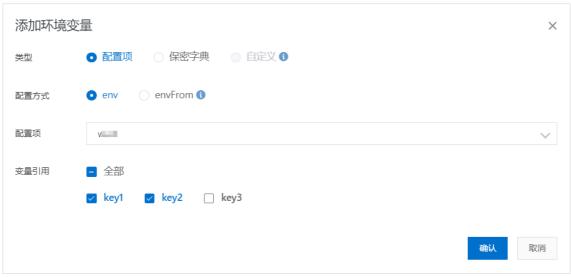
> 文档版本: 20220708 89

8s)

■ 使用保密字典配置环境变量。 在**变量引用**参数下方选择目标保密字典。一个保密字典可能存在多个键值信息,请选择正确的键值信息。



○ 批量添加环境变量。



参数	描述
类型	根据需要选择 配置项或保密字典,自定义 类型无法 批量添加。
	支持env和envFrom两种配置方式。 env: 选择目标配置项后,还需要选择配置项中的目标变量。 envFrom: 选择目标配置项后,不能选择变量,默认选中全部变量。
配置方式	② 说明 ■ 批量添加环境变量时,最大支持添加 100个环境变量。 ■ 通过编辑YAML添加envFrom方式导入 的环境变量数据,也会生效,但是不会 在此显示。
配置项	从下拉列表中选择目标配置项。
变量引用	选择env配置方式时需要选择变量。

- ② 说明 批量添加配置项的变量后,还需要根据实际需要设置环境变量的名称。

查看执行结果

● 方法一:

在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。

方法二:

在**应用总览**页面单击**运行状态**右侧的Pod运行状态链接,在**容器组(Pod)**区域查看Pod的**状态**,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.5.4. 配置持久化存储

在通常情况下,存储在容器中的数据是非持久化的,在容器销毁以后数据也将丢失。在当代高速发展的互联网环境下,数据丢失意味着线上生产环境的灾难性事件。EDAS的K8s集群支持NAS文件存储和PVC挂载功能,实现了应用实例数据持久化存储和实例间多读共享数据。

前提条件

- 开通NAS服务,使用NAS请保证当前账户拥有足够的余额或账号的计费方式为预付费。
- 创建文件系统并管理挂载点。
- 创建持久化存储卷声明。

在创建应用时配置持久化存储

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境, 然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开持久化部署,设置持久化存储数据和日志等信息。



 参数 描述

NAS文件存储

② 说明 使用NAS文件存储服务时,请确保您的集群安装了flexvolume组件,对于Serverless集群请确保 NAS挂载目录存在。NAS存储按实际使用量付费,详细收费标准,请参见NAS计费说明。

存储类型	默认为 NAS文件存储 ,不可配置。
存储服务类型	目前仅支持SSD 性能型 ,不可配置。
选择NAS	 新购一个NAS:选择NAS挂载目录和本地挂载目录。单一地域内NAS文件系统上限为10个,若已超过10个会创建失败。如有特殊需求,请提交工单。 使用已有NAS:选择一个已有的NAS。您最多可以创建两个挂载点。不符合要求的NAS将不会显示在下拉列表中。
挂载目录	分别设置NAS挂载目录和容器挂载目录。

PVC挂载

② 说明 通过修改YAML方式配置的PVC挂载,在部署应用时将不会显示已存在的PVC挂载,仅部署时会保留这些配置。推荐您在使用EDAS部署应用时配置PVC挂载。

选择PVC	在下拉列表中选择PVC。如果没有可选的PVC,请参见创建持久化存储卷声明。
挂载配置	分别设置 挂载目录和挂载模式 。

- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

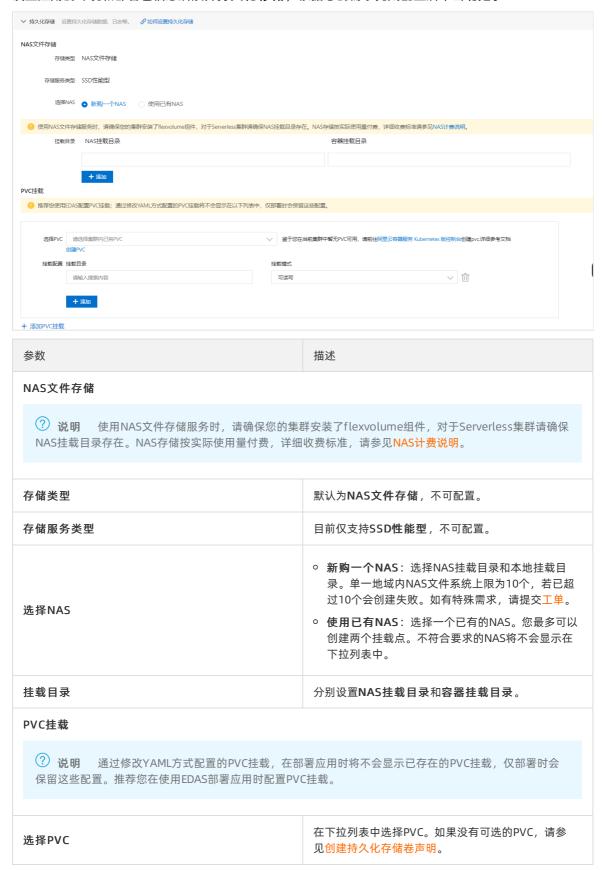
在更新应用时配置持久化存储

您除了在创建应用时可以配置NAS文件存储和PVC挂载,也可以在部署应用时配置或更新NAS文件存储和PVC挂载。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。

> 文档版本: 20220708 93

- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开持久化存储,根据您的需求完成配置后单击确定。



参数	描述
挂载配置	分别设置 挂载目录和挂载模式 。

☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

结果验证

● (推荐)方法1:在应用实例中执行以下命令来查询应用中是否存在NAS挂载信息。

```
cat /proc/mount | grep nfs
```

当回显以下信息时,表示NAS挂载成功。

sh-4.2# cat /proc/mounts | grep nfs rw,relatime,vers=3,rsize =1048576,wsize=1048576,namlen=255,hard,nolock,noresvport,proto=tcp,timeo=600,ret rans=2,sec=sys,mountaddr=10.0.0.80,mountvers=3,mountport=4002,mountproto=tcp,loc al_lock=all,addr=10.0.0.80 0 0 sh-4.2#

方法2:使用Webshell进行文件读写验证NAS挂载是否成功。 示例:

准备2个应用实例A和B,分别挂载NAS。对A执行写入命令 echo "hello" > tmp.txt , 对B执行读取命令 cat tmp.txt 。如果B中能够读取到在A中写入的 hello ,表示NAS挂载成功。

- (推荐)在NAS控制台的NAS挂载点的已**挂载客户端**页面中,查看是否存在该应用的实例IP地址。
 - ☐ 注意 已挂载客户端中仅显示近一分钟内正在使用NAS的客户端IP地址。
 - i. 登录NAS控制台。
 - ii. 进入您创建好的NAS详情页面,并单击已挂载的客户端。



iii. 在已挂载的客户端查看是否存在该应用实例的IP地址,然后在EDAS控制台中查看该应用的Pod IP地址。

如果IP地址相同,则表示NAS挂载成功。

● 在EDAS控制台查看:在变更流程详情页查看变更状态,如果变更状态显示为**执行成功**,表示应用部署成功,说明NAS挂载成功。

Linux/Unix/macOS系统

Windows系统

取消持久化存储

挂载NAS后,如果您不再使用持久化存储,可以通过重新部署应用来取消挂载NAS和PVC。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击应用列表,在顶部菜单栏选择地域并在页面上方选择命名空间,在集群类型下拉

列表中选择容器服务/Serverless K8s集群,然后在页面单击具体的应用名称。

- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开持久化存储, 取消挂载NAS后单击确定。

☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

在EDAS控制台取消挂载NAS后,您在NAS中所存储的数据仍然存在,不会被删除。

常见问题

- 如何查看NAS文件系统中的内容,有没有页面可以查看? 如果需要查看NAS文件系统中的内容,需要将NAS文件系统挂载到具体的ECS或容器上查看。具体操作, 请参见Windows系统挂载NFS文件系统或推荐的挂载方式。
- 文件日志指定的目标路径,可以与NAS指定的存储路径相同吗? 不可以。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.5.5. 配置本地存储

您可以通过配置本地存储将宿主机的文件系统上的文件或目录挂载到您的Pod中。

背景信息

在EDAS的容器服务K8s集群中配置本地存储主要通过hostPath卷来实现,hostPath卷使得容器可以使用宿主机的高速文件系统来存储文件。

在创建应用时配置本地存储

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列** 表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。

参数	描述
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node.js、C++、Go:适用于在K8s集群中使用自定义的镜像部署应用。此处未介绍多语言应用的配置流程,如需了解更多信息,请参见在K8s环境中通过镜像部署多语言微服务应用。

- 4. 在**应用配置**页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击**下一步**。
- 5. 在**应用高级设置**页签展开**本地存储**,并根据您的业务需求配置本地存储。

本地存储目前支持宿主机本地文件和emptyDir两种类型。

○ 宿主机本地文件

文件类型	Value	说明
默认	空字符串	不执行类型检查,直接挂载。
(新建)文件目录	DirectoryOrCreate	文件目录,如果不存在,则新建。
文件目录	Directory	文件目录,不存在则容器启动失 败。
(新建)文件	FileOrCreate	文件,如果不存在,则新建。
文件	File	文件,不存在则容器启动失败。
Socket	Socket	标准Unix Socket文件,不存在则容器启动失败。
CharDevice	CharDevice	字符设备文件,不存在则容器启 动失败。
BlockDevice	BlockDevice	块储存设备文件,不存在则容器 启动失败。

emptyDir

emptyDir类型的存储卷主要用于临时空间或者不同容器之间的数据共享。

参数	说明
挂载模式	包含 <i>可读写</i> 和 <i>只读</i> 。
挂载路径 (mountPath)	emptyDir的挂载路径。
subPathExpr	使用 subPathExpr 字段基于环境变量来构造 subPath目录名。

- 6. 完成设置后单击创建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置本地存储

您可在创建应用时配置本地存储,也可以在部署应用时配置或更新本地存储。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开**本地存储**,根据您的需求完成配置后单击**确定**。 本地存储目前支持**宿主机本地文件**和**emptyDir**两种类型。

○ 宿主机本地文件

文件类型	Value	说明
默认	空字符串	不执行类型检查,直接挂载。
(新建) 文件目录	DirectoryOrCreate	文件目录,如果不存在,则新建。
文件目录	Directory	文件目录,不存在则容器启动失 败。
(新建) 文件	FileOrCreate	文件,如果不存在,则新建。
文件	File	文件,不存在则容器启动失败。
Socket	Socket	标准Unix Socket文件,不存在则 容器启动失败。

文件类型	Value	说明
CharDevice	CharDevice	字符设备文件,不存在则容器启动失败。
BlockDevice	BlockDevice	块储存设备文件,不存在则容器 启动失败。

o emptyDir

emptyDir类型的存储卷主要用于临时空间或者不同容器之间的数据共享。

参数	说明
挂载模式	包含 <i>可读写</i> 和 <i>只读</i> 。
挂载路径 (mountPath)	emptyDir的挂载路径。
subPathExpr	使用 subPathExpr 字段基于环境变量来构造 subPath目录名。

杳看执行结果

● 方法一:

在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。

● 方法二:

在**应用总览**页面单击运行状态右侧的Pod运行状态链接,在容器组(Pod)区域查看Pod的状态,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

3.5.6. 配置应用生命周期的钩子和探针

在EDAS控制台可以配置PostStart和PreStop的生命周期钩子,使您在应用启动后和停止前都可以添加代码来自定义一些操作动作。同时在EDAS控制台可以配置Liveness和Readiness探针,来探测容器的存活性和可读性。本文将展示如何配置应用的生命周期钩子,以及如何配置容器的存活探针和可读性探针。

背景信息

通常应用运行前后和停止前后都会有一些操作,例如运行前部署资源、停止前优雅下线应用或通知其他服务或应用。EDAS依托于K8s完美集成了配置生命周期钩子的功能,支持为容器的生命周期配置PostStart和PreStop钩子。

当您在K8s集群中部署应用的时候,需要读取Pod的存活状态并确定Pod是否准备好对外提供服务。EDAS依托于K8s完美集成了配置Pod探针的功能,支持配置Liveness探针来确定何时重启容器,以及配置Readiness探针来确定容器是否已经就绪可以接受流量。

在创建应用时配置应用生命周期的钩子和探针

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node.js、C++、Go:适用于在K8s集群中使用自定义的镜像部署应用。此处未介绍多语言应用的配置流程,如需了解更多信息,请参见在K8s环境中通过镜像部署多语言微服务应用。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开应用生命周期管理,设置以下参数。

参数	描述
PostStart配置	一种容器钩子。该钩子在容器被创建后立刻触发,通知容器它已经被创建。该钩子不需要向其所对应的Hook Handler传入任何参数。如果该钩子对应的Hook Handler执行失败,则该容器会终止运行,并根据该容器的重启策略决定是否要重启该容器。更多信息,请参见Container Lifecycle Hooks。
PreStop配置	一种容器钩子。该钩子在容器被删除前触发,其所对应的Hook Handler必须在删除该容器的请求发送给Docker Daemon之前完成。在该钩子对应的Hook Handler完成后不论执行的结果如何,Docker Daemon会发送一个SIGTERN信号给Docker Daemon来删除该容器。更多信息,请参见Container Lifecycle Hooks
Liveness配置	一种探测容器状态的探针,探测应用是否处于健康状态。如果不健康,则删除并重建容器。更多信息,请参见Pod Lifecycle。
Readiness配置	一种探测容器状态的探针,探测应用是否启动完成并且处于正常服务状态。如果不正常,则更新容器的状态。更多信息,请参见Pod Lifecycle。

- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置应用生命周期钩子和探针

您除了在创建应用时可以配置应用生命周期钩子和探针,也可以在部署应用时配置或更新钩子和探针。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开应用生命周期管理,根据您的需求完成配置后单击确定。

参数	描述
PostStart配置	一种容器钩子。该钩子在容器被创建后立刻触发,通知容器它已经被创建。该钩子不需要向其所对应的Hook Handler传入任何参数。如果该钩子对应的Hook Handler执行失败,则该容器会终止运行,并根据该容器的重启策略决定是否要重启该容器。更多信息,请参见Container Lifecycle Hooks。
PreStop配置	一种容器钩子。该钩子在容器被删除前触发,其所对应的Hook Handler必须在删除该容器的请求发送给Docker Daemon之前完成。在该钩子对应的Hook Handler完成后不论执行的结果如何,Docker Daemon会发送一个SIGTERN信号给Docker Daemon来删除该容器。更多信息,请参见Container Lifecycle Hooks
Liveness配置	一种探测容器状态的探针,探测应用是否处于健康状态。如果不健康,则删除并重建容器。更多信息,请参见Pod Lifecycle。
Readiness配置	一种探测容器状态的探针,探测应用是否启动完成并且处于正常服务状态。如果不正常,则更新容器的状态。更多信息,请参见Pod Lifecycle。

☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

配置应用生命周期管理

PostStart和PreStop配置

PostStart和PreStop的配置方式相同,均支持以下两种模式:

- Exec模式:在容器中执行指定的命令,如果命令退出时返回码为0则判定为执行成功。
- HttpGet模式:对容器上的指定端点执行HTTP GET请求,如果响应的状态码大于等于200且小于400,则

判定为执行成功。

下面以PostStart的配置为例,讲解如何在EDAS中配置应用的PostStart和PreStop容器生命周期钩子。

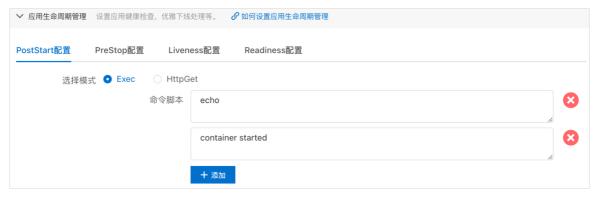
● Exec模式

以容器在创建成功后执行以下命令为例:

```
echo "container started"
```

在EDAS中有以下两种配置方式:

方式一:将命令拆分为不同的参数。这种方式适合命令比较短,参数比较少的情况。



方式二:启动一个Shell,在Shell中执行命令。这种方式适合命令比较长,参数比较多的情况。

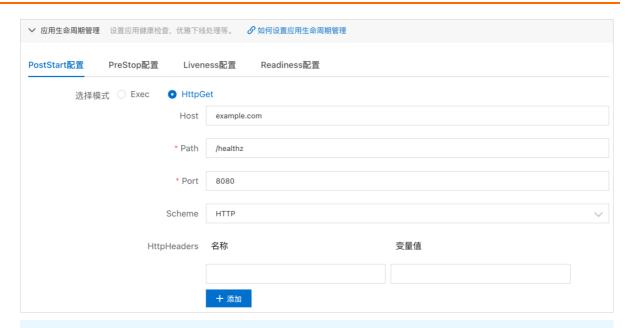


● HttpGet模式

HttpGet模式支持以下配置参数:

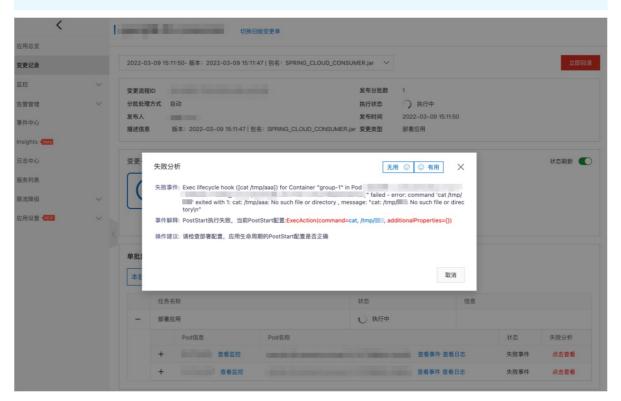
- Host: HTTP请求主机地址,不设置时默认为Pod的IP地址。
- Path: HTTP请求路径,默认值是"/"。
- Port: HTTP请求端口号。
- Scheme:协议类型,支持HTTP和HTTPS协议,默认是HTTP。
- HttpHeaders: 自定义HTTP请求头。

以PostStart的HttpGet钩子为例,在容器启动后会向 http://example.com:8080/healthz 这个地址发送GET请求。



? 说明

如果PostStart或PreStop钩子执行失败,容器会被自动重启,在应用的变更记录中可以查询到 PostStart或PreStop执行失败的信息。



Liveness和Readiness配置

EDAS支持以下两种K8s容器探针:

- LivenessProbe: 用于检测容器是否存活。如果检测失败,K8s会自动杀死容器,然后根据重启策略重启容器。
- ReadinessProbe: 用于检测容器是否就绪可以接收流量。如果检测成功,K8s会将Pod添加到Pod所属的 Service列表中,从而使Pod开始接收流量。

> 文档版本: 20220708 103

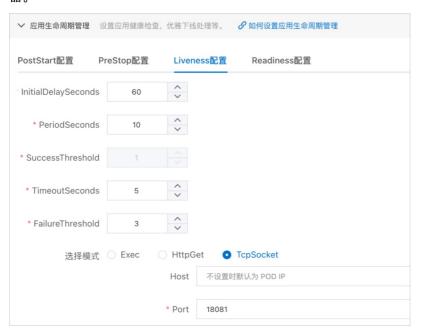
Liveness和Readiness探针支持以下三种模式:

- Exec: 在容器中执行指定的命令,如果命令退出时返回码为0则认为诊断成功。
- HttpGet:对容器上的特定端点执行HTTP GET请求,如果响应的状态码大于等于200且小于400,则认为探测成功的。
- TcpSocket: 检查容器的IP地址上的指定的TCP端口是否打开, 如果打开则认为探测成功。

Liveness和Readiness所有模式支持以下配置参数:

- *InitialDelaySeconds*:容器启动后到探针第一次执行的延迟时间,以秒为单位,最小值为1秒,最大值为1000秒。
- PeriodSeconds: 探测周期,以秒为单位,最小值为1秒,最大值为1000秒。
- *SuccessThreshold*:探针探测失败后,只有连续成功次数达到该阈值才被认为再次成功。最小值为1,最大值为1000,在Liveness探针中,该值固定为1。
- Timeout Seconds: 探针超时时间,以秒为单位,最小值为1秒,最大值为1000秒。
- FailureThreshold: 探针探测成功后,只有连续失败次数达到该阈值才被认为探测失败。最小值为1,最大值为1000。

下图以Liveness配置为例,使用TcpSocket模式,探测的地址为空,因此默认目标地址是Pod的IP地址,探测端口是18081。在容器启动后60秒才开始进行首次探测,探测周期为10秒,只要有1次探测成功即认为Pod存活。每次探测的超时时间是5秒,出现连续3次探测失败,K8s就会终结容器,然后根据重启策略重启容器。



结果验证

当配置了应用生命周期钩子后,可通过以下示例步骤验证配置结果。

例如在启动应用前,配置了 Hello from the postStart handler 的写入脚本。

 应用部署后,使用Webshell查看文件中是否存在 Hello from the postStart handler , 如果存在,那么表示配置成功。



EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.5.7. 配置日志收集

在容器服务K8s集群中创建或部署应用时,您可开通日志服务SLS,将业务文件日志、容器标准输出stdout、stderr日志输出到SLS,方便您无限制行数地查看日志和自聚合分析日志。

前提条件

- 开通日志服务SLS
- 确保应用中每个实例预留了0.25核CPU和250MB内存的可用资源。

在创建应用时配置日志收集

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境, 然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。

> 文档版本: 20220708 105

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在**应用高级设置**页签展开**日志收集设置**,开启**开通日志收集到SLS日志服务功能**并根据您的业务需求设置日志收集信息。



□ 注意

- 。 日志源的存放目录须包含日志的文件名,如/tmp0/cjsc.log。
- 请勿在日志源的存放目录中存放其他重要文件, 谨防目录内的文件被覆盖。
- 如果需要配置多条收集规则,请单击**添加日志源**。
- 6. 完成设置后单击创建应用。

7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置日志收集

如果您在创建应用时没有配置日志收集,可以在部署应用时进行配置。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开**日志收集设置**,开启**开通日志收集到SLS日志服务功能**并根据您的业务需求设置日志收集信息。然后单击**确定**。



□ 注意

- 单击确定后,该应用将会被重启,请在业务较少的时间段进行。
- 日志源的存放目录须包含日志的文件名,如/tmp0/cjsc.log。
- 请勿在日志源的存放目录中存放其他重要文件, 谨防目录内的文件被覆盖。

结果验证

应用部署完成后,EDAS依据所配的日志收集规则收集日志并存放到指定的文件内。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在左侧导航栏单击**日志中心**,然后单击**文件日志(SLS采集)**页签。 在**文件日志(SLS采集)**页签可以查看收集的日志信息,如果存在日志数据,则表示日志收集配置成功。
 - 您可以在操作列单击跳转,在日志服务控制台查询并分析日志。更多信息,请参见查询与分析。

○ 您可以单击管理日志源,在管理日志源对话框修改LogStore名称、删除日志源或添加日志源。

常见问题

- EDAS日志功能是否支持滚动日志? EDAS支持使用通配符方式滚动日志。 指定收集日志文件名称,区分大小写,您可以使用通配符,例如 *.log 。 Linux下的文件名通配符包括 * 、 ? 和 [...] 。
- 查看文件日志时,在日志服务控制台无数据展示,如何处理? 通常日志服务默认查询最近15分钟内的查询结果,并且在配置SLS收集前产生的文件内容不会被采集到SLS 中。如果日志服务无数据显示,建议使用Webshell查看您的应用是否有文件日志信息。
 - 如果有,请创建工单。
 - 如果没有,请检查您的应用。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.5.8. 配置Tomcat

在容器服务K8s集群或Serverless K8s集群中以WAR包或JAR包方式创建或部署应用时,您可以配置应用端口、应用访问路径以及最大线程数等参数信息。

在创建应用时配置Tomcat

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	选择Java应用类型,通过JAR包或WAR包部署。 o 自定义:支持镜像部署,选择这种方式表示使用程序本身的Tomcat配置,无需配置Tomcat参数。故如果您选择自定义的运行环境,在高级设置中将没有Tomcat配置选项。 o Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 o EDAS-Container(HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在应用高级设置页签展开Tomcat配置,选中自定义配置,然后单击自定义配置右侧的编辑,您可

在编辑Tomcat配置对话框配置以下参数,配置完成后单击确定。

参数	描述
应用端口	端口范围为1024~65535,小于1024的端口需要Root 权限才能操作。因为容器配置的是Admin权限,所以 请填写大于1024的端口。如果不配置,则默认为 8080。
Tomcat Context	选择应用的访问路径。 程序包名字: 无需填写自定义路径,应用的访问路径是JAR包或WAR包名称。 根目录: 无需填写自定义路径,应用的访问路径是/。 自定义: 需要在下面的自定义路径中填写自定义的路径。
最大线程数	配置连接池的连接数大小,对应参数为maxThreads,默认大小是400。 警告 此项配置对应用性能有很大影响,请由专业人士配置。
Tomcat编码	Tomcat的编码格式,包括UTF-8、ISO-8859-1、GBK 和GB2312。如果不设置则默认为ISO-8859-1。
使用高级设置	当上述参数类型和具体参数无法满足您的需求时,可以选中 使用高级设 置,直接对Tomcat的Server.xml文件进行编辑。 ① 注意 只有使用WAR包部署时支持高级配置。

- 6. 完成设置后单击创建应用。
- 7. 在应用创建完成页签确认应用基本信息、应用配置和应用高级设置等信息,确认完毕后单击确定创建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置Tomcat

您可在创建应用时配置Tomcat,也可以在部署应用时配置或更新Tomcat。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。

8s)

5. 设置应用的环境和部署包信息后展开Tomcat配置,选中自定义配置,然后单击自定义配置右侧的编辑,您可在编辑Tomcat配置对话框配置以下参数,配置完成后单击确定。

参数	描述
应用端口	端口范围为1024~65535,小于1024的端口需要Root 权限才能操作。因为容器配置的是Admin权限,所以 请填写大于1024的端口。如果不配置,则默认为 8080。
Tomcat Context	选择应用的访问路径。 程序包名字: 无需填写自定义路径,应用的访问路径是JAR包或WAR包名称。 根目录: 无需填写自定义路径,应用的访问路径是/。 自定义: 需要在下面的自定义路径中填写自定义的路径。
最大线程数	配置连接池的连接数大小,对应参数为maxThreads,默认大小是400。 警告 此项配置对应用性能有很大影响,请由专业人士配置。
Tomcat编码	Tomcat的编码格式,包括UTF-8、ISO-8859-1、GBK 和GB2312。如果不设置则默认为ISO-8859-1。
使用高级设置	当上述参数类型和具体参数无法满足您的需求时,可以选中 使用高级设 置,直接对Tomcat的Server.xml文件进行编辑。 ① 注意 只有使用WAR包部署时支持高级配置。

查看执行结果

● 方法一:

在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。

● 方法二:

在**应用总览**页面单击运行状态右侧的Pod运行状态链接,在容器组(Pod)区域查看Pod的状态,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

3.5.9. 配置Java启动参数

在容器服务K8s集群或Serverless K8s集群中创建或部署应用时,您可以通过Java启动参数对JVM进行内存、GC(垃圾回收)策略以及服务注册与发现等配置。正确配置Java启动参数有助于降低GC(垃圾回收)开销,从而缩短服务器响应时间并提高吞吐量。

在创建应用时配置Java启动参数

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

集群类型 选择Kubernetes集制	群。
署应用。 Java: 支持通用的 Spring Boot应用。 应用运行环境 Tomcat: 支持通 和Spring应用。在 本。 EDAS-Container	K8s集群中使用自定义的镜像部 JAR包部署,适用于Dubbo和在选择后可更改Java环境。 用的WAR包部署,适用于Dubbo选择后可更改Java环境和容器版 (HSF):适用于使用WAR或者

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在**应用高级设置**页签展开**Java启动参数配置**,单击**Java启动参数配置**右侧的**编辑**,并根据您的业务需求完成以下配置。

参数	描述	

参数	描述
内存配置	您可完成以下参数的配置: • 初始堆内存:初始化的堆大小,单位为MB,0表示不限制使用大小。 • 最大堆内存:最大堆内存,单位为MB,可用范围为0~0.85*应用容器的Limit内存。 • 初始年轻代:初始化年轻代大小,单位为MB。 • 最大年轻代:最大的年轻代的大小,单位为MB,填写max_uintx意味着无上限的使用内存。 • Eden/Survivor比例:Eden/Survivor内存大小的比例。 • 老年代/新生代比例:老年代/新生代内存大小的比例。 • 初始持久代:永久代的初始化大小,单位为MB。 • 最大持久代:永久代的最大大小,单位为MB。 • 最大直接内存:允许使用NIO类型的直接内存的最大值大小,单位为MB。
应用	HSF应用可以配置以下应用信息: 绑定端口:指定HSF的启动服务绑定端口,默认值为12200。 序列化方式:指定HSF的序列化方式,默认值为hessian。 线程池最小:指定HSF的服务端最小线程池大小,默认值为50。 线程池最大:指定HSF的服务端最大线程池大小,默认值为720。
GC策略	您可配置以下垃圾回收策略: 年轻代GC策略:用于设置年轻代垃圾回收器。 老年代GC策略:用于设置老年代垃圾回收器,必须
协程特性	当您想使用协程特性时,可以开启 协程特性 开关。 ② 说明 只有在选择Dragonwell作为JDK时才会出现 协程特性 配置项。
工具	设置GC日志文件的日志目录、文件个数和文件大小,以及内存溢出时的Dump路径。

参数	描述
服务注册和发现配置	当您想使用自建Nacos注册中心时,可以勾选该配置项。当未选中此选项时,将自动替换解析为EDAS共享版注册中心。
自定义	当上述配置参数不能满足您的配置需求时,您可以配置自定义参数,多个参数用空格分隔。

- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置Java启动参数

您可在创建应用时配置lava启动参数,也可以在部署应用时配置或更新lava启动参数。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择容器**服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开Java启动参数配置,单击Java启动参数配置右侧的编辑,根据您的需求完成配置后单击确定。

参数	描述
内存配置	您可完成以下参数的配置: 初始堆内存:初始化的堆大小,单位为MB,0表示不限制使用大小。 最大堆内存:最大堆内存,单位为MB,可用范围为0~0.85*应用容器的Limit内存。 初始年轻代:初始化年轻代大小,单位为MB。 最大年轻代:最大的年轻代的大小,单位为MB,填写max_uintx意味着无上限的使用内存。 Eden/Survivor比例:Eden/Survivor内存大小的比例。 老年代/新生代比例:老年代/新生代内存大小的比例。 初始持久代:永久代的初始化大小,单位为MB。 最大持久代:永久代的最大大小,单位为MB。 最大直接内存:允许使用NIO类型的直接内存的最大值大小,单位为MB。 线程栈大小:线程栈大小,单位为KB。

参数	描述
应用	HSF应用可以配置以下应用信息: 绑定端口:指定HSF的启动服务绑定端口,默认值为12200。 序列化方式:指定HSF的序列化方式,默认值为hessian。 线程池最小:指定HSF的服务端最小线程池大小,默认值为50。 线程池最大:指定HSF的服务端最大线程池大小,默认值为720。
GC策略	您可配置以下垃圾回收策略: 年轻代GC策略:用于设置年轻代垃圾回收器。 老年代GC策略:用于设置老年代垃圾回收器,必须 先配置年轻代垃圾回收器。 并发GC线程数:配置并发GC将使用的线程数。 并行GC线程数:配置并行GC将使用的线程数。
协程特性	当您想使用协程特性时,可以开启 协程特性 开关。 ② 说明 只有在选择Dragonwell作为JDK时才会出现 协程特性 配置项。
工具	设置GC日志文件的日志目录、文件个数和文件大小,以及内存溢出时的Dump路径。
服务注册和发现配置	当您想使用自建Nacos注册中心时,可以勾选该配置项。当未选中此选项时,将自动替换解析为EDAS共享版注册中心。
自定义	当上述配置参数不能满足您的配置需求时,您可以配置自定义参数,多个参数用空格分隔。

☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

查看执行结果

● 方法一:

在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。

● 方法二:

在**应用总览**页面单击**运行状态**右侧的Pod运行状态链接,在**容器组(Pod)**区域查看Pod的**状态**,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

3.5.10. 配置挂载

您可以将已经创建好的配置项或保密字典以文件的形式挂载到容器的指定目录,从而向容器中注入配置信息。本文介绍如何在EDAS控制台为应用配置挂载。

背景信息

目前EDAS控制台支持两种挂载的配置类型:

- 配置项:提前配置好的一类不需要加密的配置信息,应用需要和配置项在相同地域、相同集群、相同K8s 命名空间,否则无法使用配置项。 创建配置项的具体操作,请参见创建配置项。
- 保密字典:提前配置好的一类需要加密的密码和证书等信息,应用需要和保密字典在相同地域、相同集群、相同K8s命名空间,否则无法使用保密字典。 创建保密字典的具体操作,请参见创建保密字典。

在创建应用时配置挂载

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列**表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择Kubernetes集群。
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node.js、C++、Go:适用于在K8s集群中使用自定义的镜像部署应用。此处未介绍多语言应用的配置流程,如需了解更多信息,请参见在K8s环境中通过镜像部署多语言微服务应用。

- 4. 在应用配置页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击下一步。
- 5. 在**应用高级设置**页签展开**配置管理**,单击**添加**,并根据您的业务需求配置挂载。 在创建或部署应用过程中,将配置项或保密字典以文件挂载的形式,向容器中注入配置信息。

参数	描述
配置类型	可选择 配置项 或 保密字典 。
挂载类型	可选择 挂载到目录 或 挂载到文件 。
挂载源	需要挂载到应用容器文件系统中的配置项或保密字典 的名称。
挂载主目录	设置容器的挂载主目录。路径必须以正斜线(/)开头,且字符长度大于两位。
文件挂载方式	参数 挂载类型 选择挂载到文件时呈现此参数。 • 覆盖源文件 : 隐藏原目录下的文件,仅保留本次挂载文件。 • 保留原有文件 : 保留原目录下的文件,添加本次挂载文件,注意,相同文件名将保留本次挂载文件,隐藏原目录重名文件。
挂载配置	参数 挂载类型 选择挂载到文件时呈现此参数。 • 要挂载的Key :需要挂载到应用容器文件系统中的配置项或保密字典指定的Key。 • 挂载文件子路径 :相对挂载主目录的子路径。

- 6. 完成设置后单击创建应用。
- 7. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创 建应用**。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置挂载

如果您在创建应用时没有将配置项或保密字典以文件的形式挂载到容器的指定目录,可以在部署应用时进行配置。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开**配置管理**,根据您的需求完成配置后单击**确定**。 在创建或部署应用过程中,将配置项或保密字典以文件挂载的形式,向容器中注入配置信息。

参数	描述
配置类型	可选择 配置项或保密字典 。
挂载类型	可选择 挂载到目录 或 挂载到文件 。

参数	描述
挂载源	需要挂载到应用容器文件系统中的配置项或保密字典的名称。
挂载主目录	设置容器的挂载主目录。路径必须以正斜线(/)开头,且字符长度大于两位。
文件挂载方式	参数 挂载类型 选择挂载到文件时呈现此参数。 • 覆盖源文件 :隐藏原目录下的文件,仅保留本次挂载文件。 • 保留原有文件 :保留原目录下的文件,添加本次挂载文件,注意,相同文件名将保留本次挂载文件,隐藏原目录重名文件。
挂载配置	参数 挂载类型 选择挂载到文件时呈现此参数。 • 要挂载的Key :需要挂载到应用容器文件系统中的配置项或保密字典指定的Key。 • 挂载文件子路径 :相对挂载主目录的子路径。

☆ 注意 単击确定后,该应用将会被重启,请在业务较少的时间段进行。

查看执行结果

- 方法一:
 - 在应用的变更记录页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。
- 方法二:

在**应用总览**页面单击运行状态右侧的Pod运行状态链接,在容器组(Pod)区域查看Pod的状态,如果显示为代表运行中的绿色圆圈,则说明部署成功,配置已生效。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

3.5.11. 配置微服务治理

在EDAS中部署的SpringCloud和Dubbo应用能够无侵入对接MSE微服务治理能力。当前EDAS支持您在创建和部署应用时为应用配置服务注册与发现、推空保护和无损上线微服务治理能力。

背景信息

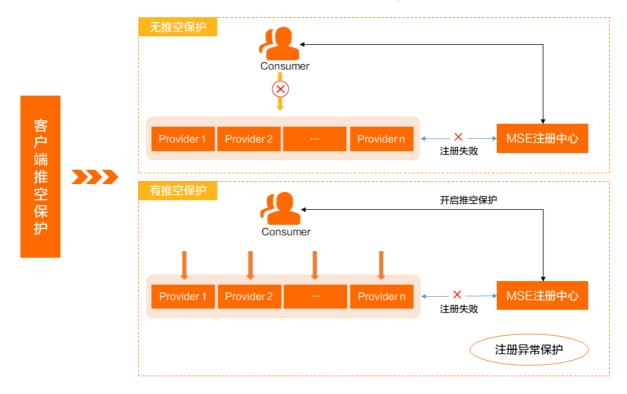
服务注册与发现

Java微服务架构中常见的注册中心包含Eureka、ZooKeeper和Nacos等,用于实现服务的注册与发现,能够 屏蔽、解耦服务之间的相互依赖,以便对微服务进行动态管理。

推空保护

推空保护功能用于处理客户端在请求注册中心订阅服务端地址列表时,在服务端注册异常的场景下,注册中心返回了空列表,此时客户端忽略该空返回的变更,从缓存中获取上一次正常的服务端地址进行服务访问。能够在注册中心在进行变更(变配、升降级)或遇到突发情况(例如,可用区断网断电)或其他不可预知情况下的列表订阅异常时,增加客户端可靠性。

服务消费者(Consumer)会从注册中心上订阅服务提供者(Provider)的实例列表。当注册中心在进行变更(变配、升降级)或遇到突发情况(例如,可用区断网断电)时,都有可能导致订阅异常,影响服务消费者(Consumer)的可用性。为了应对在不可预知的情况下订阅列表异常,可以在Consumer侧配置推空保护。



无损上线

对于任何一个线上应用来说,发布、扩容、缩容和重启等操作不可避免,EDAS集成了MSE提供的一套无损上线方案,针对应用启动的多个阶段提供了相应的保护能力,具体功能包含服务预热、服务延迟注册以及无损滚动发布等。

- 服务预热:通过设置预热时间和预热曲线,降低新上线的服务在短时间内无法处理突增流量导致进程异常的风险。
- 延迟注册: 延迟注册能够在应用启动时进行延时等待,在等待延迟时间后,再将服务注册到注册中心。能够为应用提供异步请求、计算资源的准备时间,降低在应用发布时由于资源未初始化成功的请求异常。对于初始化过程繁琐的应用,由于注册通常与应用初始化过程同步进行,从而出现应用还未完全初始化就已经被注册到注册中心供外部消费者调用,此时直接调用可能会导致请求报错。通过设置延迟注册,可让应用在充分初始化后再注册到注册中心对外提供服务。
- 无损滚动发布: Health接口为微服务体系外的客户端提供应用就绪状态检查的方法,基于该接口能够同步 微服务框架下与非微服务框架下的服务状态检查。支持服务注册状态检查和服务预热状态检查。 K8s的滚动发布是将一次完整的发布过程分成多个批次,每次发布一个批次,成功后,再发布下一个批 次,最终完成所有批次的发布。EDAS能够为K8s应用打通微服务生命周期与K8s服务生命周期,通过提供 确保在通过就绪检查前完成服务注册以及确保在就绪检查前完成服务预热等无损上线能力,帮助您在应用的滚动发布过程中,始终保证有可用的服务实例(副本)在运行,从而实现应用的无损滚动发布。
 - 通过就绪检查前完成服务注册:微服务生命周期关联K8s就绪检查功能开启后,可通过Agent无侵入为应用提供一个检测其是否完成注册的端口,当应用注册完成返回200帮助K8s判定应用已就绪,未完成注册返回500帮助K8s判定应用未就绪。

○ 通过就绪检查前完成服务预热:除了可以单独使用服务预热,还可以将其关联K8s就绪检查机制。当应用服务预热完成后,K8s才将应用置为就绪状态。

使用限制

仅Spring Cloud和Dubbo应用支持无侵入接入推空保护和无损上线微服务治理能力,对应支持的版本如下:

限制项	推空保护	无损上线
Spring Cloud	支持Spring Cloud Edgware及以上版本 支持Spring Cloud Edgware及	
	支持Dubbo 2.5.3~2.7.8版本	支持Dubbo 2.7.3~2.7.8版本
Dubbo	? 说明 如需使用Dubbo 3.0及以上版本,请提交工单。	? 说明 如需使用Dubbo 2.6或 Dubbo 3.0及以上版本,请提交 <mark>工单</mark> 。

- 支持Nacos、Eureka和Zookeeper注册中心类型。
- 网关应用一般不通过注册中心调用,而是通过直接对外暴露API的方式调用,因此MSE当前所支持的小流量 预热功能对该类应用不生效。另一方面,网关应用本身很少变更而且没有太多业务逻辑,因此也不太需要 服务预热功能。
- 金丝雀发布第一批灰度批次的流量不支持进行服务预热,后续批次才能进行服务预热。

在创建应用时配置微服务治理

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,然后在**应用列** 表页面左上角单击**创建应用**。
- 3. 在应用基本信息页签选择集群类型和应用运行环境,然后在页面最下方单击下一步。

参数	描述
集群类型	选择 Kubernetes集群 。

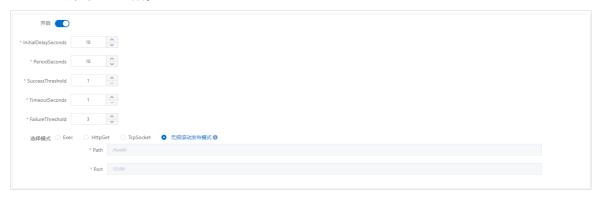
参数	描述
应用运行环境	根据选择的托管应用类型选择应用运行环境。 Java: 自定义:适用于在K8s集群中使用自定义的镜像部署应用。 Java:支持通用的JAR包部署,适用于Dubbo和Spring Boot应用。在选择后可更改Java环境。 Tomcat:支持通用的WAR包部署,适用于Dubbo和Spring应用。在选择后可更改Java环境和容器版本。 EDAS-Container (HSF):适用于使用WAR或者FatJar部署HSF应用。在选择后可更改Java环境、Pandora版本和Ali-Tomcat版本。 PHP:适用于运行在Apache HTTP Server中的PHP应用程序。 多语言Node.js、C++、Go:适用于在K8s集群中使用自定义的镜像部署应用。此处未介绍多语言应用的配置流程,如需了解更多信息,请参见在K8s环境中通过镜像部署多语言微服务应用。 ② 说明 选择EDAS-Container (HSF)部署应用时,微服务治理仅支持配置服务注册与发现。 选择PHP或多语言Node.js、C++、Go部署应用时,不支持配置微服务治理。

- 4. 在**应用配置**页签中设置应用的环境信息、基本信息、部署方式和资源参数,设置完成后单击**下一步**。
- 5. 在**应用高级设置**页签展开**微服务治理**,根据您的业务需求配置服务注册于发现、推空保护和无损上线。

参数	描述
服务注册与发现 配置	注册中心运维方式。您可以选择使用EDAS共享注册中心或使用应用程序配置的注册中心。 具体操作,请参见 <mark>选择注册中心运维方式</mark> 。
推空保护	是否开启推空保护。 开启推空保护: Consumer订阅到空列表,推空保护生效,丢弃变更,保障业务服务可用。不开启推空保护: Consumer订阅到空列表,业务中断报错。

是否开启无损上线。开启时,需配置以下参数: 预热时长(秒) :应用实例下一次启动的预热时间,默认预热时长为120秒。服务预热	参数	描述
时长设置范围为0~86400秒(即24小时)。 • 预热曲线:基于已配置的预热时长,被预热的应用流量权重会根据配置的预热曲线呈指数型增长。在指定预热时长内,预热曲线值越大被预热应用刚启动时分配的流量权重越小,以满足需要较长时间进行预热的复杂应用的预热需求。默认为2(适合于一般预热场景),表示在预热周期内服务提供者的流量接收曲线形状呈2次曲线形状。预热曲线设置范围为0~20。相同预热时间,预热曲线值越大,表示预热开始将接收的流量越小,临近预热结束时接收的流量增幅越大。 • 延迟注册时间(秒):设置范围为0~86400秒(即24小时)。 • 无损滚动发布:配置应用无损滚动发布。 ■ 通过就绪检查前完成服务注册:为应用无侵入提供55199端口用于检查微服务是否已经完成注册,如果已注册完成,端口返回200,否则返回500。 ■ 通过就络检查前完成服务预热:为应用无侵入提供55199端口用于检查微服务是否已经完成预热,如果已预热完成,端口返回200,否则返回500。	无损上线	 预热时长(秒):应用实例下一次启动的预热时间,默认预热时长为120秒。服务预热时长设置范围为0~86400秒(即24小时)。 预热曲线:基于已配置的预热时长,被预热的应用流量权重会根据配置的预热曲线呈指数型增长。在指定预热时长内,预热曲线值越大被预热应用刚启动时分配的流量权重越小,以满足需要较长时间进行预热的复杂应用的预热需求。默认为2(适合于一般预热场景),表示在预热周期内服务提供者的流量接收曲线形状呈2次曲线形状。预热曲线设置范围为0~20。相同预热时间,预热曲线值越大,表示预热开始将接收的流量越小,临近预热结束时接收的流量增幅越大。 延迟注册时间(秒):设置范围为0~86400秒(即24小时)。 无损滚动发布:配置应用无损滚动发布。 通过就绪检查前完成服务注册:为应用无侵入提供55199端口用于检查微服务是否已经完成注册,如果已注册完成,端口返回200,否则返回500。 通过就络检查前完成服务预热:为应用无侵入提供55199端口用于检查微服务是否已经完成预热,如果已预热完成,端口返回200,否则返回500。 ② 说明 若同时开启了通过就绪检查前完成服务注册、通过就结检查前完成

6. 在**应用高级设置**页签展开**应用生命周期管理**,选择**Readiness**配置选项,单击开启开关,使Readiness探针配置生效。



应用未配置就绪探针时,启用无损上线的无损滚动发布配置后,会在就绪探针模式中新增无损滚动发布模式,选择该模式会为您的应用默认配置路径为/health,端口为55199的就绪状态检查。您可以直接使用该无损滚动发布模式以打通微服务与K8s的生命周期,也可以结合业务需求使用其他模式结合/health接口配置更严谨的业务就绪检查方法。

- 7. 完成设置后单击创建应用。
- 8. 在**应用创建完成**页签确认**应用基本信息、应用配置和应用高级设置**等信息,确认完毕后单击**确定创** 建应用。

在应用**应用总览**页面顶部提示信息中单击**查看详情**,可在**变更记录**页面查看变更流程。应用部署需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。

在更新应用时配置微服务治理

如果您在创建应用时没有配置应用的服务注册与发现、推空保护或无损上线,可以在部署应用时进行配置。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择容器服务/Serverless K8s集群,然后在页面单击具体的应用名称。
- 3. 在应用总览或基本信息页面的右上角选择部署 > 部署。
- 4. 在选择部署模式页面选择具体的部署方式,然后在所选方式区域的右上角单击开始部署。
- 5. 设置应用的环境和部署包信息后展开配置管理,根据您的需求完成配置后单击确定。
 - i. 展开微服务治理,根据您的业务需求配置服务注册与发现、推空保护和无损上线。

参数	描述
服务注册与发现配置	注册中心运维方式。您可以选择使用EDAS共享注册中心或使用应用程序配置的注册中心。具体操作,请参见选择注册中心运维方式。
推空保护	是否开启推空保护。 一 开启推空保护: Consumer订阅到空列表,推空保护生效,丢弃变更,保障业务服务可用。 一 不开启推空保护: Consumer订阅到空列表,业务中断报错。
无损上线	是否开启无损上线。开启时,需配置以下参数: 预热时长(秒) :应用实例下一次启动的预热时间,默认预热时长为120秒。服务预热时长设置范围为0~86400秒(即24小时)。 预热曲线 :基于已配置的预热时长,被预热的应用流量权重会根据配置的预热曲线呈指数型增长。在指定预热时长内,预热曲线值越大被预热应用刚启动时分配的流量权重越小,以满足需要较长时间进行预热的复杂应用的预热需求。默认为2(适合于一般预热场景),表示在预热周期内服务提供者的流量接收曲线形状呈2次曲线形状。预热曲线设置范围为0~20。相同预热时间,预热曲线值越大,表示预热开始将接收的流量越小,临近预热结束时接收的流量增幅越大。 延迟注册时间(秒) :设置范围为0~86400秒(即24小时)。 无损滚动发布:配置应用无损滚动发布。 通过就绪检查前完成服务注册:为应用无侵入提供55199端口用于检查微服务是否已经完成注册,如果已注册完成,端口返回200,否则返回500。 通过就绪检查前完成服务预热:为应用无侵入提供55199端口用于检查微服务是否已经完成预热,如果已预热完成,端口返回200,否则返回500。

- ii. 展开限流降级配置,开启AHAS接入开关。

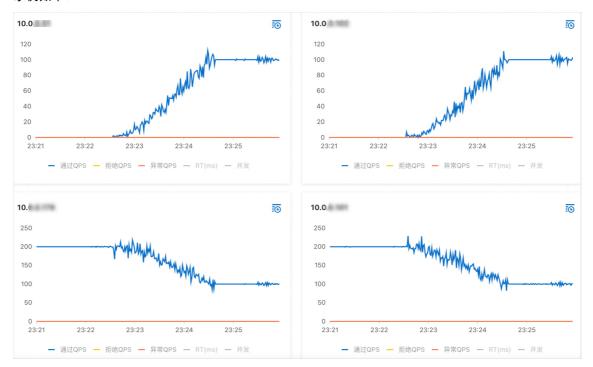
结果验证

在应用的**变更记录**页面中查看应用变更详情,如果显示执行成功,则表示部署成功,即配置已生效。 无损上线在应用部署后,可以查看预热效果。

1. 登录AHAS控制台。

- 2. 在AHAS控制台左上角,选择应用接入的地域。
- 3. 在控制台左侧导航栏中选择流量防护 > 应用防护。
- 4. 在应用防护页面单击目标应用卡片。
- 5. 在左侧导航栏中选择机器监控。
- 6. 在机器监控页面,查看应用的预热效果。
 - o 服务预热

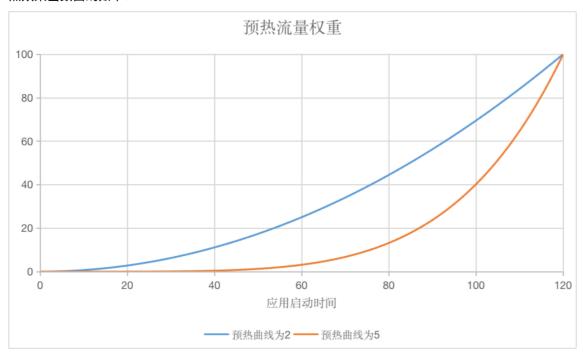
应用部署选择分批发布,并配置应用在部署时开启服务预热,同时应用接入AHAS,应用的预热效果示例如下:



在稳定的业务流量下,可以看到新发布的应用所占流量比率呈预热曲线缓慢增长。

8s)

无损上线功能通过配置应用的预热时长和预热曲线来控制新发布的应用流量百分比,用户可以根据业务需求按需配置预热时长,并根据程序复杂度控制预热曲线来调整应用刚启动时分配的流量权重,预热效果函数曲线如下:

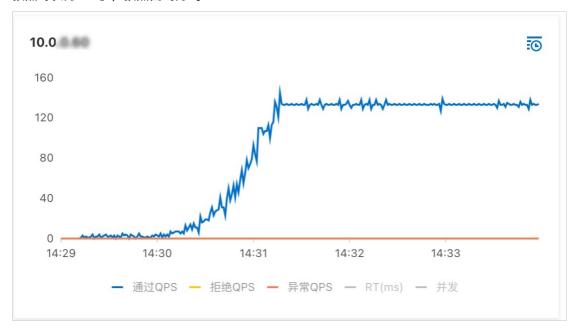


应用部署时,配置不同的预热曲线效果展示示例如下:

■ 预热时长为120秒, 预热曲线为2。



■ 预热时长为120秒,预热曲线为5。



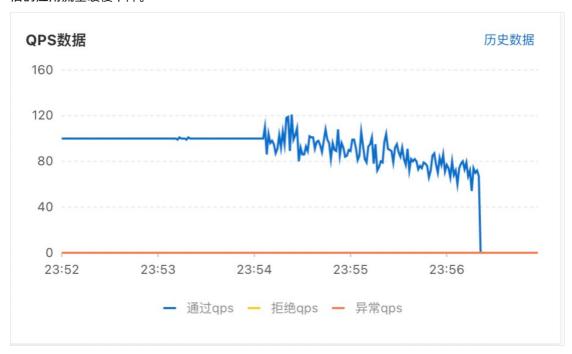
○ 无损滚动发布

开启无损滚动发布后能将微服务的生命周期与K8s的生命周期打通,在分批发布时选择配置服务预热并开启无损滚动发布,可以在发布过程中看到以下效果:

■ 新发布的应用接收的流量呈预热曲线缓慢增长。



■ 旧的应用流量缓慢下降。

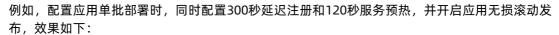


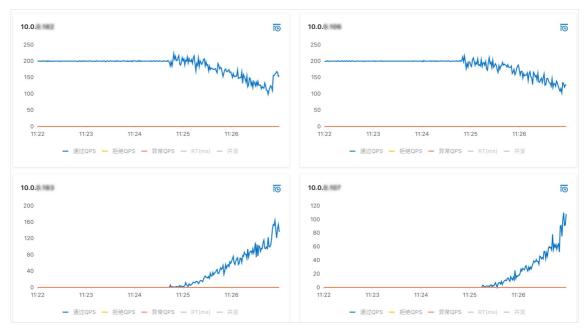
在应用发布的过程中可在应用事件中看到Unhealthy告警事件,由于应用的就绪探针未检查通过,此时新旧版本实例共存,旧实例在预热过程中接收额外流量。在预热结束后,就绪探针请求返回200,旧实例才被删除以实现服务预热并打通与K8s的生命周期。

> [13][2022-04-14T15:45:00Z]Pulled	→ 执行成功
> [14][2022-04-14T15:45:00Z]Created	→ 执行成功
> [15][2022-04-14T15:45:01Z]Started	→ 执行成功
> [16][2022-04-14T15:45:50Z]Unhealthy	! 告警事件
> [17][2022-04-14T15:46:40Z]Unhealthy	! 告警事件

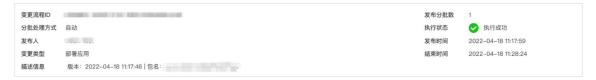
o 延迟注册

配置延迟注册后,能够让应用在启动后等待指定的延迟时间后再注册到注册中心以提供服务,能够给应用充分的启动时间,以避免应用未完全加载完成导致的流量损失。





在应用的变更记录页面中查看应用变更记录。



可以看到应用在11:17:59进行变更,在11:24:44前新建的Pod都没有接收任何流量,且在11:24:44时间点后开始进行流量预热。观察应用变更事件能够找到Unhealt hy告警事件。考虑到应用启动过程包含Pod调度、镜像拉取、程序启动等事件,从应用的发布时间到变更流程结束时间大致满足应用启动进行延迟注册和服务预热的时间。

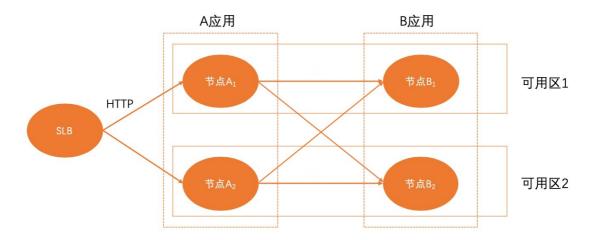
3.5.12. 配置同可用区优先路由

可用区优先是指在应用调用服务时,优先调用同可用区的服务提供者。Spring Cloud、Dubbo和HSF协议应用都支持配置可用区优先调用。本文介绍如何为K8s应用开启可用区优先调用。

背景信息

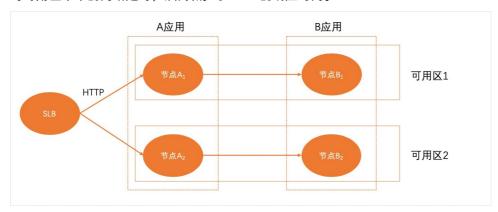
8s)

当应用部署在多个可用区的时候,应用之间互相调用会出现跨可用区的情况。调用链路如下图所示:



图中可用区1的A应用如果调用可用区2的B应用,即节点A1上的应用调用节点B2上的应用,跨可用区调用出现网络延时增加,并导致HTTP响应时间增加。

当应用服务开启同可用区优先后,消费者应用会优先调用同可用区的提供者,调用链路如下图所示:避免了跨可用区带来的网络延时,从而减少了HTTP的响应时间。



如果当前可用区没有服务提供者时,服务调用者会调用到其他可用区的服务提供者。

开启同可用区优先

要使用同可用区优先功能,您可以在服务提供者侧开启同可用区优先。

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用总览页面的服务治理区域,单击同可用区优先右侧的开启。
- 4. 在应用总览页面的服务治理区域,单击同可用区优先右侧的≥图标并设置阈值。

 ② 说明 只有在当前可用区实例数量占实例总数比例大于所设置的阈值时,同可用区优先调用策略才会生效。阈值填写范围为0~100。

3.5.13. 实现K8s集群应用的限流降级

部署在容器服务K8s集群或Serverless K8s集群中的应用支持通过重新部署来接入AHAS的试用版,并可在EDAS内嵌的监控页面中实时监控流量,全面可视化地保障您的应用可用性。本教程以在Dubbo应用中配置限流降级规则为例,介绍了如何实现微服务的限流降级。

背景信息

EDAS中默认开通AHAS试用版,仅包含入门级防护,您只能为同一个应用类设置2条相同类型的防护规则,如需设置更多规则,请前往AHAS控制台开通专业版。

如果您想废弃或修改已创建的规则,您可以在应用总览页面的左侧导航栏选择**限流降级 > 规则管理**,选择要删除或修改的规则页签,然后在规则列表中选择规则来进行管理。

接入AHAS

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览页面右上角选择部署 > 部署。
- 4. 在选择部署模式页面的其中一种部署模式区域的右上角单击开始部署。
- 5. 在单批发布页面,修改部署参数,并展开限流降级配置。
- 6. 打开接入AHAS开关,然后单击确定。
 - □ 注意 如果该应用已经通过SDK方式接入了AHAS,打开此开关将会有冲突风险。

打开接入AHAS后,该应用将自动接入AHAS。在变更流程执行完毕后,您可前往**监控 > 服务/接口监控**页面配置限流降级规则。

配置限流降级规则

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在左侧导航栏选择监控 > 服务/接口监控。
- 4. 在服务/接口监控页面单击限流降级页签。
 - 如果您已开通过AHAS,或在部署的应用中已通过SDK方式接入了AHAS,那么您可以在此页面直接观测到接口的请求流量并配置限流降级规则。
 - 如果您未开通过AHAS,将可以在此页面提示中单击**立即开通**来跳转到应用部署页面,通过重新部署来开启AHAS试用版。具体操作,请参见接入AHAS。
- 5. 在**限流降级**页签选择请求链路并观测该请求链路的QPS和RT,然后根据观测结果来配置限流降级规则。
 - 您可在请求链路的操作列单击限流,在弹出的限流规则对话框中配置限流规则。具体操作,请参

见配置流控规则。

○ 您可在请求链路的**操作**列单击**降级**,在弹出的**降级规则**对话框中配置降级规则。具体操作,请参见配置熔断规则。

支持组件列表

AHAS为了简化应用的接入流程,对主流框架进行了适配。本文将列出AHAS支持的第三方组件和框架列表。

组件支持版本	从EDAS直接接入	配置SDK接入
Dubbo 2.5.x+	2.5.x和2.6.x版本已支持	全部支持
Web Servlet	已支持	支持SDK 2.x或以上版本接入
Spring Boot 1.3.x+	已支持	参见说明
Spring MVC 4.x+	已支持	参见说明
Spring Cloud Gateway 2.x	已支持	参见说明
Zuul 1.3.x	已支持	参见 <mark>说明</mark>
GRPC-Java 1.13+	已支持	需另引入sentinel-grpc-adapter依 赖
Jetty 8.x+	已支持	通过Servlet支持
Tomcat 7.x+	已支持	通过Servlet支持
WebLogic 10.3	已支持	通过Servlet支持
HttpClient 3.x+	不支持	不支持
HttpClient 4.x+	不支持	不支持
JDK HTTP 1.7.x+	不支持	不支持
OKHttp 2.x+	不支持	不支持
MyBatis 3.x+	不支持	支持SDK 1.4.1或以上版本接入
MySQL JDBC 5.0.x+	已支持	不支持
Oracle JDBC 12.x	已支持	不支持
PostgreSql JDBC 9.4+	不支持	不支持
SQLServer JDBC 6.4+	不支持	不支持
Redis Client (Jedis)	已支持	不支持
MemCached 2.8+	已支持	不支持
MongoDB 3.7+	不支持	不支持

组件支持版本	从EDAS直接接入	配置SDK接入
RocketMQ(callback模式)4.x	已支持	需手动埋点
RabbitMQ 3.7+	已支持	需手动埋点
Feign 10.x+	已支持	需引入spring-cloud-starter- alibaba-sentinel

? 说明

- 支持的AHAS Agent版本和SDK版本,请参见版本说明。
- Spring MVC、Spring Boot和Spring Cloud Web应用需要引入 ahas-sentinel-client 并进行相 应配置即可使用(参考Web Servlet支持的版本);若需要通过*Spring properties*配置启动参数,则需要单独引入 spring-boot-starter-ahas-sentinel-client 依赖。
- Spring Cloud Gateway网关需要引入 spring-cloud-gateway-starter-ahas-sentinel 依赖;
 Zuul 1.x网关需要引入 spring-cloud-zuul-starter-ahas-sentinel 依赖, 无需引入其它依赖。
- Dubbo 2.7.x以上版本暂不支持在EDAS控制台部署接入AHAS,您可通过在应用中添加防护依赖的方式来接入AHAS。具体操作,请参见接入Dubbo应用。

3.5.14. 配置标签

本文主要介绍在创建或部署应用时为应用的Pod配置标签。在集群中有大量的应用时,使用标签能够方便地对配置了特定标签的Pod进行分组管理。

背景信息

在微服务架构中,K8s集群中可能会部署大量的应用,且不同的应用可能存在多个版本,此时集群中会出现大量的Pod。基于K8s的Label在EDAS控制台对应用配置Label,能够为该应用的所有Pod配置标签,以键值对的形式对一组Pod进行标识。结合标签选择器,能够在大量的Pod中筛选出具有特定标签标识的一组Pod。

操作步骤

此处以部署应用的操作为例,讲解如何进行标签配置。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览页面右上角选择部署 > 部署。
- 4. 在选择部署模式页面的其中一种部署模式区域的右上角单击开始部署。

- 5. 在单批发布页面,修改部署参数,并展开标签 (Label)配置。
- 6. 进行自定义标签键值对配置,根据您的需求完成配置后单击**确定**。 等待应用部署成功后,通过kubectl命令根据指定标签键值查询Pod。

kubectl get pod -l demo-app-name=sc-server-demo

结果示例如下:

```
NAME READY STATUS RESTARTS AGE sc-server-demo-group-1-1-5795d5487c-*** 2/2 Running 0 2m58s
```

基于Label实现应用亲和性

在部署K8s应用时,我们通常需要考虑不同组件间的依赖关系,比如将某些Pod调度到同一拓扑域下,或不能将某些Pod调度到同一拓扑域。此时为应用配置Label可以辅助配置自定义的应用亲和性和反亲和性。下面将介绍如何通过Label实现应用亲和性。

前提条件

已经按照操作步骤部署了标签为demo-app-name=sc-server-demo的应用。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览页面右上角选择部署 > 部署。
- 4. 在选择部署模式页面的其中一种部署模式区域的右上角单击开始部署。

5. 在单批发布页面,修改应用的环境和部署包信息后,展开调度规则。

选择模式配置为自定义配置,在应用亲和性页签添加规则,参数配置如下:

- o 拓扑域选择 kubernetes.io/hostname
- o 应用标签名配置为 demo-app-name
- 操作符号选择 In
- o 应用标签值配置为 sc-server-demo

等待应用部署成功后,通过kubectl命令检查两个应用Pod所在的节点。

```
kubectl get pod -l demo-app-name=sc-client-demo -o wide
kubectl get pod -l demo-app-name=sc-server-demo -o wide
```

结果示例如下:

NAME	READY ST	TATUS	RESTARTS	AGE	IP
NODE	NOMINATED NODE READING	ESS GATES			
sc-cl	ient-demo-group-1-1-5c5cbb9889-*** 2/2 Rur	nning 0		24s	10.85.***
.*4	cn-hangzhou.192.168.AAA.BBB <none></none>	<none></none>			
NAME	READY ST	TATUS	RESTARTS	AGE	IP
NODE	NOMINATED NODE READING	ESS GATES			
sc-se	erver-demo-group-1-1-5795d5487c-*** 2/2 Rur	nning 0		35m	10.85.***
.*3	cn-hangzhou.192.168.AAA.BBB <none></none>	<none></none>			

从结果中可以看到两个应用Pod所在节点为同一节点。

3.5.15. 配置注解

本文主要介绍在创建或部署应用时为应用添加非标识的元数据信息。通过使用注解,可以将元数据附加到应用Pod上。注解值能够支持结构化及非结构化数据,能够支持包含标签不允许的字符。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览页面右上角选择部署 > 部署。
- 4. 在选择部署模式页面的其中一种部署模式区域的右上角单击开始部署。
- 5. 在单批发布页面,修改部署参数,并展开注解(Annotation)配置。
- 6. 进行自定义注解键值对配置,根据您的需求完成配置后单击确定。

注解键值对参数配置示例如下:

- o 名称配置为 demo-annotation
- o 值配置为 demo-annotation-value

等待应用部署成功后,通过kubectl命令查询应用的注解配置。

kubectl get pod

可以看到应用Pod的metadata.annotations存在demo-annotation: demo-annotation-value键值对。

4.升级和回滚应用(K8s)

4.1. 升级和回滚应用概述(K8s)

应用在EDAS中创建并部署后,还会不断迭代,需要升级。如果升级的应用版本发现问题,需要将应用回滚到历史版本。本文介绍EDAS为您提供的应用升级策略、方式,以及在需要时如何回滚应用。

升级应用

由于新版本应用的状态不同,升级的场景也不同,您需要使用的升级策略、方式也不同。

升级策略	适用场景	升级方式	相关文档
单批发布(直接升级)	应用规模较小或服务架构 比较简单,在开发、测试 完成后,可以直接升级, 即直接部署新版本。	控制台API插件或工具	 使用控制台单批发布应用(K8s) DeployK8sApplication 使用Cloud Toolkit插件单批发布应用(K8s)
分批发布	应用规模较大或服务架构 比较复杂,可以按批次进 行应用部署。每次仅对应 用的一部分实例进行升 级,直至全量升级完成。 分批发布过程中如果出现 故障,可以及时终止并回 滚,待问题修复后重新发 布。	控制台API插件或工具	 使用控制台分批发布应用(K8s) DeployK8sApplication 使用toolkit-maven-plugin插件分批发布应用(K8s)
金丝雀发布	应用的新版本需要先进行 小规模验证,验证后,再 将应用全量升级到新版 本。	控制台	使用控制台金丝雀发布应 用(K8s)

回滚应用

在应用升级过程中,出现部署包不可用、健康检查失败等异常报错会导致应用升级失败,当前应用变更会自动终止并回滚。

如果您发现新版本应用有问题或异常,可以手动回滚应用,有以下两种场景:

- 升级中回滚:可以终止升级变更流程并回滚应用。
- 升级后回滚,可以通过部署历史版本回滚应用。

EDAS提供控制台可视化操作和调用API接口两种方式实现回滚应用。更多信息,请参见:

- 使用控制台回滚应用(K8s)
- 终止升级变更流程并回滚: Abort And Rollback Change Order
- 回滚历史版本: RollbackApplication

4.2. 单批发布(K8s)

4.2.1. 使用控制台单批发布应用(K8s)

应用在EDAS中创建并部署后,还会不断迭代,需要升级。当您的应用规模较小或服务架构比较简单,在开发、测试完成后,可以直接升级,即直接部署新版本。本文介绍如何使用EDAS控制台单批升级应用。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中选择应用管理 > 应用列表。
- 3. 在**应用列表**页面的**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后单击具体应用名称。
- 4. 在应用总览页面右上角选择部署 > 部署。
- 5. 在选择部署模式页面单批发布区域单击开始部署。
- 6. 在单批发布页面上传应用新版本的部署包。

参数	描述		
应用运行环境	默认标准Java应用运行环境。		
Java环境	选择Open JDK 8 [最新版本: 1.8.0_191]、Open JDK 7、JDK 8、JDK 7或Dragonwell 8。		
当前环境	显示当前应用的运行环境。仅是JAR包和WAR包部署的应用,才会显示当前应用的运行环境。EDAS会自动将您应用的Java环境或应用运行环境升级到最新版本。		
	可选择 上传JAR包或JAR包地址 两种方式进行上传。		
文件上传方式	② 说明 文件上传方式需要和上一次部署的方式一致,包含JAR包、WAR包或镜像,本文以JAR包为例。		
上传JAR包	文件上传方式选择上传JAR包时,单击选择文件,上 传JAR包。		
	文件上传方式选择JAR包地址时,输入JAR包地址。		
JAR包地址	② 说明 对于带认证签名的OSS文件访问 URL, EDAS在部署时会缓存该文件用于之后的回滚、扩容等操作。		
版本	输入JAR版本,也可单击右侧的用时间戳作为版本 号,自动生成相应的时间戳。		
时区	参考指定地域对应的UTC时区。		
服务注册与发现配置	选择注册中心的运维方式。如何选择 <i>,</i> 请参见 <mark>选择注</mark> 册中心运维方式。		

- 7. (可选)根据您的实际需求,选择是否需要配置**调度规则、启动命令、环境变量、持久化存储、本地存储、应用生命周期管理**和日志收集设置等高级配置。更多信息,请参见在K8s集群中使用镜像部署Java微服务应用和(可选)配置应用高级设置。
- 8. 设置完成后,单击确定。

结果验证

在应用中查看应用变更记录,查看单批发布的状态,待执行状态显示为执行成功则说明单批发布成功。

4.2.2. 使用API在K8s集群中单批发布应用

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中单批发布应用(JAR包、WAR包和镜像)。

前提条件

单批发布应用前,您必须提前完成以下准备工作:

- 确定应用部署地域,假设为cn-hangzhou。
- 在EDAS K8s环境中创建应用。目前支持多种创建应用的方式,请根据实际需求选择:
 - o 使用API在K8s集群中创建lava应用(JAR包或WAR包)
 - o 使用API创建Java应用(镜像部署)
 - o 使用API在K8s集群中创建多语言应用

如果已经创建应用,您可以调用ListApplication接口查询应用列表,获取目标应用的APPId,假设为6bbc57 a2-a017-4bec-b521-49a15bd3****。

- 已将应用升级部署包(JAR包、WAR包和镜像)上传至目标地址。
 - JAR包或WAR包:本示例以OSS存储路径为例,假设为https://doc***.oss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.2-SNAPSHOT.jar。
 - 镜像: 假设镜像仓库为*image-demo-project*, 镜像地址为*registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provider:2.0*。

使用IAR包或WAR包单批发布应用

以下代码适用于在EDAS K8s集群中单批发布应用(JAR包)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

② 说明 如果您需要使用WAR包部署应用,请根据需要在示例代码中增加以下两个参数:

//部署包依赖的Tomcat版本。适用于通过WAR包部署的Spring Cloud和Dubbo应用。

request.setWebContainer("apache-tomcat-7.0.91");

//部署包依赖的EDAS Container版本。适用于通过WAR包部署的HSF应用。

request.setEdasContainerVersion("3.5.9");

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //应用实例数。
       request.setReplicas(2);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //JAR包或者WAR包地址、版本。
       request.setPackageUrl("https:doc***.oss-cn-hangzhou.aliyuncs.com/sc-***-D-0.0.2-SN
APSHOT.jar");
       request.setPackageVersion("20210227");
       //部署包依赖的JDK版本。可选的参数值为Open JDK 7和Open JDK 8。
       request.setJDK("Open JDK 8");
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
       //CPU能使用的最大值,单位:核数。0表示不限制。
       request.setMcpuLimit(0);
       //变更记录描述。
       request.setChangeOrderDesc("单批发布(JAR包)");
           DeployK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("ChangeOrderId=" + response.getChangeOrderId() + "\nMessage=
" + response.getMessage());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的结果如下:

ChangeOrderId=1c18b409-07ac-4a37-990b-66bd87d7****
Message=success

使用镜像单批发布应用

以下代码适用于在EDAS K8s集群中单批发布应用(镜像)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun_user_ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //要执行API调用的应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       request.setAppId("c44227c7-fc5e-46ca-80ac-8f342d45****");
       //应用实例数。
       request.setReplicas(2);
       //镜像地址。
       request.setImage("registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provider
:2.0");
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
       //CPU能使用的最大值,单位:核数。0表示不限制。
       request.setMcpuLimit(0);
       //变更记录描述。
       request.setChangeOrderDesc("单批发布(镜像)");
           DeployK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("ChangeOrderId=" + response.getChangeOrderId() + "\nMessage=
" + response.getMessage());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
```

运行程序,返回的结果如下:

```
ChangeOrderId=1c18b409-07ac-4a37-990b-66bd87d7****
Message=success
```

结果验证

您在单批发布应用后,可以调用Get Change OrderInf o接口查看应用变更流程详情,获取单批发布应用的变更状态。具体API参数详情,请参见Get Change OrderInfo。

调用Get Change OrderInf o接口查看应用变更流程详情,返回的执行结果如下:

● 使用IAR包或WAR包单批发布应用的变更单详情。

● 使用镜像单批发布应用的变更单详情。

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断单批发布应用的变更是否成功。changeOrderInfo.Status的取值如下:

● 0: 准备

● 1: 执行中

• 2: 执行成功

● 3: 执行失败

• 6: 终止

- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败
 - ② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示单批发布应用失败。您可以根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新单批发布应用。

4.2.3. 使用Cloud Toolkit插件单批发布应用(K8s)

您在本地完成应用的开发、调试和测试后,可以在Intellij IDEA中安装Cloud Toolkit插件实现EDAS应用的快速部署。本文档将介绍如何在Intellij IDEA中安装Cloud Toolkit,并快速部署应用到容器服务K8s集群。

前提条件

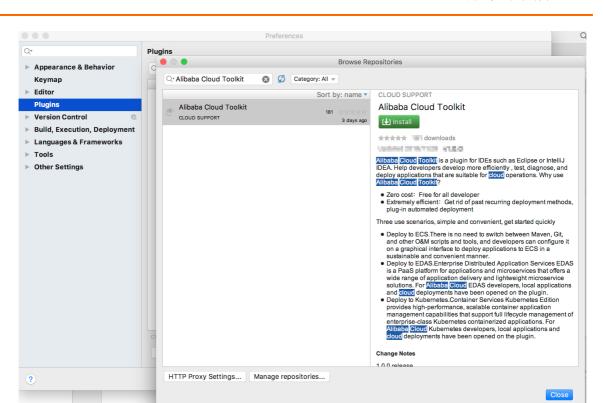
- 下载并安装JDK 1.8或更高版本。
- 下载并安装IntelliJ IDEA (2018.3或更高版本)。
 - ② 说明 因Jet Brains插件市场官方服务器在海外,如遇访问缓慢无法下载安装的,请加入文末交流群,向Cloud Toolkit 技术支持人员获取离线包安装。
- 在容器服务K8s集群中创建了应用并完成了首次部署,相关操作请参见在K8s集群中使用镜像部署Java微服务应用。

步骤一:安装Cloud Toolkit

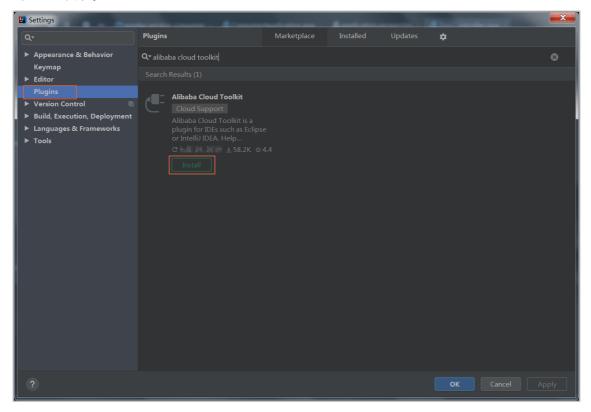
- 1. 启动IntelliJ IDEA。
- 2. 在Intellij IDEA中安装最新版本的插件。

如您之前已安装过旧版安装包,请升级至最新版本。

○ mac系统: 进入Preference配置页面,在左侧导航栏选择Plugins,在右侧搜索框输入Alibaba Cloud Toolkit,然后单击Install安装。



○ Windows系统:在左侧导航栏选择Plugins,在右侧搜索框输入Alibaba Cloud Toolkit,并单击Install安装。

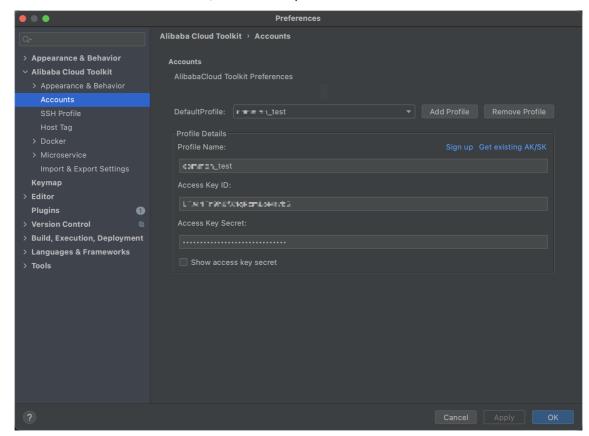


3. 在Intellij IDEA中插件安装成功后,重启Intellij IDEA,您可以在工具栏看到Alibaba Cloud Toolkit的图标(【一】。

步骤二:配置Cloud Toolkit账号

 在安装完Alibaba Cloud Toolkit后,您需使用AccessKey ID和AccessKey Secret来配置Cloud Toolkit的账号。

- 1. 启动IntelliJ IDEA。
- 2. 在Intellij IDEA菜单栏选择Tools > Alibaba Cloud > Preference...。
- 3. 在左侧导航栏选择Alibaba Cloud Toolkit > Accounts。
- 4. 在Accounts界面中设置 AccessKey ID和AccessKey Secret, 然后单击OK。
 - 如果您使用RAM用户的AccessKey ID和AccessKey Secret,请确认该RAM用户至少拥有部署应用的权限。
 - 如果您是EDAS专有云企业版用户,还需配置Endpoint才能正常使用Cloud Toolkit功能。



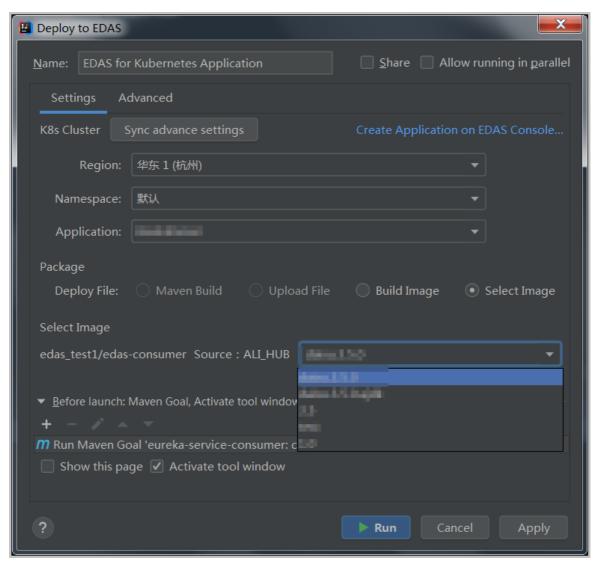
- 如果您已经注册过阿里云账号,在Accounts界面中单击Get existing AK/SK,进入阿里云登录页面。用已有账号登录后,跳转至安全信息管理页面,获取AccessKey ID和AccessKey Secret。
- 如果您还没有阿里云账号,在Accounts界面中单击Sign up,进入阿里云账号注册页面,注册账号。注册完成后按照上述方式获取AccessKey ID和AccessKey Secret。

步骤三: 部署应用到容器服务K8s集群

1. 在Intellij IDEA菜单栏选择Tools > Alibaba Cloud > Deploy to EDAS > EDAS for Kubernetes Application。

您也可以在Intellij IDEA菜单栏选择Run > Edit Configurations,在Run/Debug configurations对话框左侧选择Templates > EDAS on Alibaba Cloud > EDAS for Kubernetes Application。

2. 在Deploy to EDAS对话框配置应用部署参数。



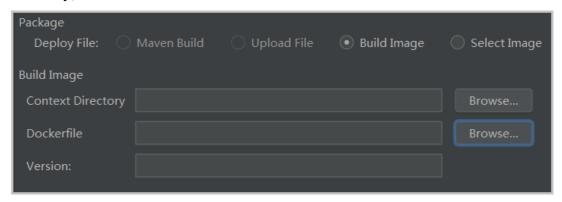
i. 在配置页面中根据您的实际需求选择应用的Region、Namespace和Application。

■ Region:应用所在地域。

■ Namespace: 应用所在微服务空间。

■ Application: 应用名称。

- ii. 设置部署方式,可选的部署方式由您在控制台第一次部署的方式决定。
 - **Maven Build**:选择Maven Build方式来构建应用时,系统会默认添加一个Maven任务来构建部署包。
 - ② 说明 如果您想部署多模块工程,请参见部署多模块工程。
 - **Upload File**:选择Upload File方式来构建应用时,选择上传您的WAR包或者JAR包,然后进行部署。
 - Build Image: 选择Build Image方式来构建应用时,需要依次输入工作目录Content Directory,文件路径Dockerfile和镜像标签Version。



- Select Image: 选择Build Image方式来构建应用时,可在Select Image的下拉列表选择需要部署的镜像。
- iii. 单击Advanced进入高级部署参数设置页签,并配置Application Environment、Startup Command、Environments、Persistence、Storage 和Appcalition Management等参数。
- 3. 单击**Apply**,再单击**Run**。
 Intellij IDEA的Console区域会打印部署日志。您可以根据日志信息检查部署结果。

部署多模块工程

实际工作中碰到的大部分Maven工程都是多模块的,各个项目模块可以独立开发,其中某些模块又可能会使用到其他的一些模块的功能,这样的项目工程就是多模块工程。

如果您的工程项目为Maven多模块工程并且想部署工程中的某子模块,那么需要保证EDAS Deployment Configurations页面中的Before launch中的Maven构建任务中最后一个任务为该子模块的构建任务。管理Maven构建任务的具体操作,请参见Maven构建任务的管理。

例如:一个CarShop工程存在以下示例的子模块。

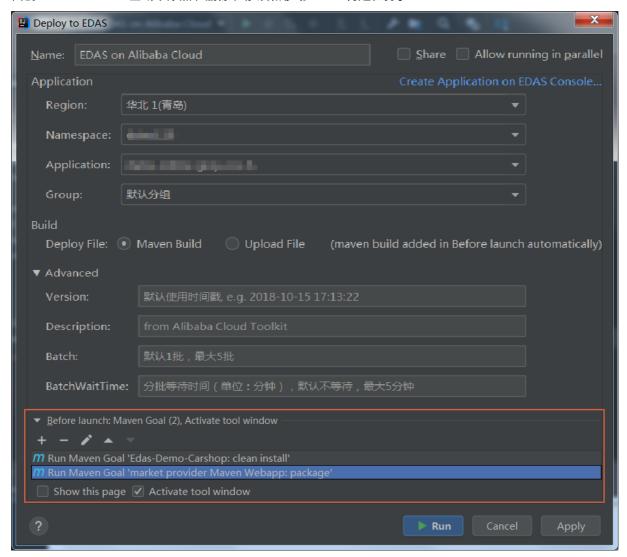


其中itemcenter和detail为子模块,且都依赖于itemcenter-api模块,现在想部署itemcenter模块,应该怎么做?只需要在配置页面中的**Before launch**中增加如下两个Maven任务:

- 增加一个在父工程carshop中执行 mvn clean install 的Maven任务。
- 增加一个在子模块itemcenter中执行 mvn clean package 的Maven任务。

Maven构建任务的管理

在Intellij IDEA中安装的Cloud Toolkit内可以部署Maven的构建任务。您也可以在 **Deploy to EDAS**的配置页面的**Before launch**区域来添加、删除、修改和移动Maven构建任务。



在添加Maven构建任务编辑框中,您可以单击右侧的文件夹按钮选择当前工程的所有可用模块,并在Command line中编辑构建命令。



常见问题

● 应用列表获取不到应用 通常出现这种情况为使用RAM用户来部署应用,且RAM用户没有同步到EDAS系统或者没有进行正确授 权,从而导致在应用列表下拉框中看不到应用。您可以通过RAM授权或EDAS子账号授权来确保RAM用户 已经同步到EDAS并且得到授权。

o RAM授权

该授权方式可使RAM用户访问EDAS的所有资源。

- a. 在RAM控制台左侧导航栏中选择人员管理 > 用户。
- b. 在用户页面上找到需要授权的RAM用户,单击操作列中的添加权限。
- c. 在**添加权限**面板的**选择权限**区域中,搜索**Aliyun**EDASFullAccess权限,单击权限策略将其添加至右侧的已**选择**列表中,然后单击**确定**。
- d. 在添加权限的授权结果页面上,查看授权信息摘要,并单击完成。
- e. 使用阿里云账号登录EDAS控制台,在左侧导航栏选择**系统管理 > 子账号**,单击**子账号**页面右上角的同步子账号。

○ EDAS子账号授权

该授权方式可使子账号细粒度授权访问EDAS的资源。

- a. 使用阿里云主账号登录EDAS控制台。
- b. 在左侧导航栏选择系统管理 > 角色, 单击页面右上角的创建角色。
- c. 输入一个角色名称,在可**选权限**区域框中,选择**应用列表 > 基本信息 > 部署应用**,单击**添加**将 部署应用添加到已**选权限**,然后单击**确**定。
- d. 在左侧导航栏选择系统管理 > 子账号, 单击子账号页面右上角的同步子账号。
- e. 选择需进行授权的子账号,在**EDAS旧鉴权**列单击**管理角色**,在左侧搜索并选择目标角色,将该角色添加到右侧已选角色列表中,然后单击**确定**。
- f. 选择需进行授权的子账号,在**EDAS旧鉴权**列单击**授权应用**,在左侧搜索并选择目标应用,将该应用添加到右侧已选应用列表中,然后单击**确定**。

● 配置Endpoint

② 说明 如果使用EDAS专有云企业版,还需要按以下步骤在Cloud Toolkit中配置Endpoint。 Endpoint请联系EDAS技术支持获取。

- i. 在Preference (Filtered) 对话框的左侧导航栏中选择Appearance & BehaviorEndpoint。
- ii. 在Endpoint界面中设置Endpoint,配置完成后,单击Apply and Close。

问题反馈

如果您在使用工具部署应用的过程中有任何疑问,欢迎您扫描下面的二维码或搜索群号34556175,加入钉钉群进行反馈。

4.3. 分批发布(K8s)

4.3.1. 使用控制台分批发布应用(K8s)

应用规模较大或服务架构比较复杂,可以按批次进行部署。每次仅对应用的一部分实例进行升级,直至全量升级完成。本文介绍如何使用EDAS控制台分批发布应用。

分批发布简介

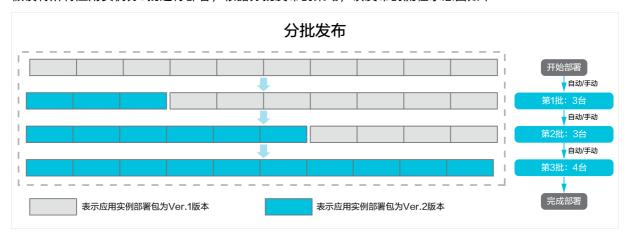
分批发布是按照一定的批次,每次只对应用的一部分实例进行升级的发布过程。分批发布过程中如果出现故障,您可以终止变更过程并进行回滚,待问题修复后重新发布。

在K8s集群中,当一个应用进行分批发布时,是将该应用内的应用实例数均分到每一个批次进行部署,如果不能均分则批次靠前数量较小,批次靠后数量较大。

场景示意

现有一个应用包含10个应用实例,每个应用实例的部署版本为Ver.1版本,现需将每个应用实例升级为Ver.2版本。

假设将所有应用实例分3批进行部署,根据分批发布的策略,该发布的流程示意图如下:



使用须知

在K8s集群中,应用使用分批发布时会新建Deployment进行部署。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击应用列表,在顶部菜单栏选择地域并在页面上方选择微服务空间。
- 3. 在**应用列表**页面的**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后单击具体应用名称。
- 4. 在应用总览页面右上角选择部署 > 部署。
- 5. 在选择部署模式页面分批发布区域单击开始部署。
- 6. 在分批发布页面上传应用新版本的部署包。

标准Java应用运行环境。
《Open JDK 8 [最新版本:1.8.0_191]、Open 7、JDK 8、JDK 7或Dragonwell 8。
S当前应用的运行环境。仅是JAR包和WAR包部署的 B,才会显示当前应用的运行环境。EDAS会自动将 I用的Java环境或应用运行环境升级到最新版本。
接择上传JAR包或JAR包地址两种方式进行上传。 说明 文件上传方式需要和上一次部署的方式一致,包含JAR包、WAR包或镜像,本文以JAR 2为例。

参数	描述
上传JAR包	文件上传方式选择上传JAR包时,单击选择文件,上 传JAR包。
JAR包地址	文件上传方式选择JAR包地址时,输入JAR包地址。
	② 说明 对于带认证签名的OSS文件访问 URL, EDAS在部署时会缓存该文件用于之后的回滚、扩容等操作。
版本	输入JAR版本,也可单击右侧的 用时间戳作为版本 号,自动生成相应的时间戳。
时区	参考指定地域对应的UTC时区。
服务注册与发现配置	选择注册中心的运维方式。如何选择,请参见选择注册中心运维方式。

7. 设置发布策略。



发布策略参数说明。

参数	描述
发布批次	应用实例按照设定的批次完成分批发布。
分批间处理方式	包含自动和手动两种方式。 • 自动:根据间隔自动分批发布。间隔即剩余批次间的发布时间间隔,单位为min。 • 手动:手动触发下一个批次的发布。 ② 说明 仅发布批次大于1时,才会显示分批间处理方式。
批次內部署间隔	如果应用实例数大于1,应用实例间的部署时间间隔,单位为秒。

8. (可选)根据您的实际需求,选择是否需要配置**调度规则、启动命令、环境变量、持久化存储、本地存储、应用生命周期管理**和日志收集设置等高级配置。更多信息,请参见在K8s集群中使用镜像部署Java微服务应用和(可选)配置应用高级设置。

9. 设置完成后,单击确定。

结果验证

在应用中查看应用变更记录,查看分批发布的状态,待所有批次都执行成功则说明分批发布成功。

在应用详情页查看实例部署信息,当实例的版本已变为Ver.2并且所有实例运行状态均为Running时,则说明发布成功。

回滚应用

在分批发布期间,只要至少有一个应用实例未升级为新版本,就认为发布处于**执行中**状态。在应用升级时,如果您注意到第一批升级的应用实例停止响应,您可以进入实例变更详情页面单击**回滚**,将已发布的实例回滚到以前的服务包和配置。

在分批发布时,可能会遇到异常。如何排查并解决异常,请参见变更流程问题排查指南。

- 部署包不可用、健康检查失败等异常报错会导致应用升级失败,当前应用变更会自动终止并回滚。
- 升级时单批次最大超时时间为30分钟,如果是超时导致变更流程暂停,您需进入**变更详情**页面手动终止 发布流程并回滚。

4.3.2. 使用API在K8s集群中分批发布应用

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中分批发布应用(JAR包、WAR包和镜像)。

前提条件

调用API分批发布应用前,您必须提前完成以下准备工作:

- 在EDAS k8s环境中创建应用。目前支持多种创建应用的方式,请根据实际需求选择:
 - 使用API在K8s集群中创建Java应用(JAR包或WAR包)
 - o 使用API创建Java应用(镜像部署)
 - o 使用API在K8s集群中创建多语言应用

如果已经创建应用,您可以调用**ListApplication**接口查询应用列表,获取目标应用的APPId,假设为*6bbc57 a2-a017-4bec-b521-49a15bd3*****。

- 已将应用升级部署包(JAR包、WAR包和镜像)上传至目标地址。
 - JAR包或WAR包:本示例以OSS存储路径为例,假设为https://doc***.oss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.2-SNAPSHOT.iar。
 - 镜像: 假设镜像仓库为 *image-demo-project*, 镜像地址为 *registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provider:2.0*。

背景信息

使用API分批发布应用时,推荐选用自动分批发布方式。如果您选用手动分批发布方式,则还需要调用ContinuePipeline接口手动确认执行下一批发布。具体操作,请参见手动确认下一批发布。

此处提供两种示例分批发布策略,请根据你的实际场景选择并修改策略。

● Pod实例分2批发布+手动分批。

```
{"type":"BatchUpdate","batchUpdate":{"batch":2,"releaseType":"manual"}}
```

• Pod实例分2批发布+自动分批+分批间隔2分钟。

```
{"type": "BatchUpdate", "batchUpdate": { "batch": 2, "releaseType": "auto", "batchWaitTime": 2} }
```

使用IAR包或WAR包分批发布应用

以下代码适用于在EDAS K8s集群中分批发布应用(JAR包)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

② 说明 如果您需要使用WAR包部署应用,请根据需要在示例代码中增加以下两个参数:

```
//部署包依赖的Tomcat版本。适用于通过WAR包部署的Spring Cloud和Dubbo应用。
request.setWebContainer("apache-tomcat-7.0.91");
//部署包依赖的EDAS Container版本。适用于通过WAR包部署的HSF应用。
request.setEdasContainerVersion("3.5.9");
```

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       //应用ID。
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //JAR包或者WAR包地址、版本。
       request.setPackageUrl("https:doc***.oss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.2-SN
APSHOT.jar");
       request.setPackageVersion("2021-04-15 14:01:12");
       //部署包依赖的JDK版本。可选的参数值为Open JDK 7和Open JDK 8。
       request.setJDK("Open JDK 8");
       //自定义分批发布策略。此处示例值代表"Pod实例分2批发布+自动分批+分批间隔2分钟"。
       request.setUpdateStrategy("{\"type\":\"BatchUpdate\",\"batchUpdate\":{\"batch\":2,\
"releaseType\":\"auto\",\"batchWaitTime\":2}}");
       //应用实例数。
       request.setReplicas(3);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
       //CPU能使用的最大值,单位:核数。0表示不限制。
       request.setMcpuLimit(0);
       //变更记录描述。
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=e16bdf30-00c9-4b12-87b5-0ad35f3f****
Message=success
```

使用镜像分批发布应用

以下代码适用于在EDAS K8s集群中分批发布应用(镜像)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

② 说明 本示例采用的是手动确认分批发布方式,您可以根据需要修改为自动分批发布方式。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun_user_ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //镜像地址。
       request.setImage("registry-vpc.cn-hangzhou.aliyuncs.com/edas-demo-project/provider:
2.0");
       //自定义分批发布策略。此处示例值代表"Pod实例分2批发布+手动分批",您可以根据需要修改为自动分批
策略。
       request.setUpdateStrategy("{"type":"BatchUpdate","batchUpdate":{"batch":2,"releaseT
ype":"manual"}}");
       //应用实例数。
       request.setReplicas(3);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
       //CPU能使用的最大值,单位:核数。0表示不限制。
       request.setMcpuLimit(0);
       //变更记录描述。
       request.setChangeOrderDesc("分批发布(镜像)");
       trv {
           DeployK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("ChangeOrderId=" + response.getChangeOrderId() + "\nMessage=
" + response.getMessage());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
   }
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=7ed96c93-cbd5-4614-a304-861ede6c****
Message=success
```

结果验证

您在分批发布应用后,可以调用Get ChangeOrderInf o接口查看变更流程详情,获取分批发布应用的变更状态。具体API参数详情,请参见Get ChangeOrderInfo。

调用Get Change OrderInf o接口查看变更流程详情,返回的执行结果如下:

● 使用JAR包或WAR包分批发布应用的变更流程详情。

● 使用镜像分批发布应用的变更流程详情。

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断分批发布应用的变更是否成功。changeOrderInfo.Status的取值如下:

● 0: 准备

- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败
 - ② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示分批发布应用失败。您可以根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新分批发布应用。

手动确认下一批发布

如果您使用的是手动确认分批发布方式,则需要调用ContinuePipeline接口进行手动确认下一批发布。

1. 调用Get ChangeOrderInfo接口,获取分批发布应用变更的相关批次的流程ID(PipelineId)。 调用Get ChangeOrderInfo接口,返回结果如下:

② 说明 本示例使用的是Pod实例分2批发布+手动确认的分批发布场景,在调用Get ChangeOrderInf o接口返回结果中会出现2个流程ID(PipelineId),第二次出现的流程ID(PipelineId)代表需要手动确认发布的批次。按此规律,如果是Pod实例分3批发布+手动确认的分批发布场景,则会出现3个流程ID(PipelineId),第二次和第三次出现的流程ID(PipelineId)代表需要手动确认发布的批次。

2. 调用ContinuePipeline接口,手动确认下一批发布。

返回结果如下:

```
8s)
```

```
{
  "Message": "success",
  "RequestId": "B8ECF6F7-672D-40E9-91DD-1C33F06D4FD8",
  "Code": 200
}
```

⑦ 说明 本示例使用的是Pod实例分2批发布+手动确认的分批发布场景,只需要手动确认一批。如果您是需要手动确认多批的场景,请选择正确批次的流程ID (Pipelineld) 并重复执行此步骤。重复执行此步骤时,请等待**批次内部署间隔**时长,默认10秒。

4.3.3. 使用toolkit-maven-plugin插件分批发布应用 (K8s)

您可以使用toolkit-maven-plugin插件分批发布已经部署在EDAS的Spring Cloud、Dubbo和HSF应用。

自动化部署

通过toolkit-maven-plugin插件自动化部署应用的流程为:添加插件依赖,配置插件,构建部署。

1. 添加插件依赖

在pom.xml文件中增加如下所示的插件依赖。

? 说明 toolkit-maven-plugin建议使用最近版本。

2. 配置插件

配置插件主要包含账号配置,打包配置及部署配置。如果需要更多自定义配置项,请参见<mark>打包参数</mark>和部 署参数。

i. 账号配置

在打包工程的根目录下创建文件格式为 YAML 的账号配置文件,命名

为 toolkit_profile.yaml 并填入如下信息:

```
regionId: #应用所在区域,如北京为"cn-beijing",上海为"cn-shanghai",杭州为"cn-han gzhou"。
jarPath: #部署包路径,配置该参数可忽略Maven打包流程从而直接使用配置路径内的包部署,
您可配置绝对路径和相对路径。
accessKeyId: #访问阿里云资源的AccessKey ID,建议使用子账号或RAM用户的AccessKey ID,以
降低安全风险。
accessKeySecret: #访问阿里云资源的AccessKey Secret,建议使用子账号或RAM用户的AccessKey Secret,以降低安全风险。
```

ii. 打包配置

在打包工程的根目录下创建文件格式为 YAML 的打包配置文件。如果打包工程为Maven的子模块,则需要在子模块的目录下创建该文件,并命名为 toolkit package.yaml ,填入如下信息:

```
apiVersion: V1
kind: AppPackage
spec:
packageType: #应用部署包类型,支持War、FatJar、Image、url; 您只有在该处配置了URL,那么packageUrl才能生效。
packageUrl才能生效。
packageUrl: #如果应用部署包类型为War或FatJar,可填入此字段,不填则使用当前Maven构建的包进行部署。
imageUrl: #如果部署包类型为Image,可填入此字段; Image类型也可以在本地构建Docker镜像进行部署。
```

iii. 部署配置

在打包工程的根目录下创建文件格式为 YAML 的部署文件,命名为 toolkit_deploy.yaml ,并填入如下信息:

```
apiVersion: V1
kind: AppDeployment
 type: kubernetes
 target:
             #部署应用的ID。如果配置了appId,则无需配置namespaceId和appName。
  namespaceId: #应用所属的微服务空间ID。如果不清楚appId,可使用namespaceId和appName (
应用名称)进行部署。
            #应用名称。如果不清楚appId,可使用namespaceId和appName进行部署。
   appName:
 updateStrategy:
  type: GrayBatchUpdate
                      #灰度分批发布。
                      #灰度设置。
   grayUpdate:
                       #整数值,含义为灰度实例数量。
    gray: x
                     #分批设置。
  batchUpdate:
                      #整数值,含义为批次数量。
                    #发布方式:auto表示自动发布,manual表示手动发布。
    releaseType: xxx
                      #整数值,含义为批次间隔时间,单位分钟。
    batchWaitTime: x
```

3. 构建部署

进入 pom.xml 所在的目录(如果部署Maven子模块,则进入子模块 pom.xml 所在的目录),执行如下命令:

```
mvn clean package toolkit:deploy -Dtoolkit_profile=toolkit_profile.yaml -Dtoolkit_packa
ge=toolkit_package.yaml -Dtoolkit_deploy=toolkit_deploy.yaml
```

命令参数含义为:

- o toolkit:deploy: 在打包完成后进行应用部署。
- o -Dtoolkit_profile : 指定账号配置文件。如果账号文件跟 pom.xml 在同一个目录下,且名字为 .toolkit_profile.yaml (注意:文件名最前面有个小数点),可不填此参数,插件会自动获取。
- o -Dtoolkit_package : 指定打包文件。如果打包文件跟 pom.xml 在同一个目录下,且名字为 .t oolkit_package.yaml (注意:文件名最前面有个小数点),可不填此参数,插件会自动获取。
- o -Dtoolkit deploy : 指定部署文件。如果部署文件跟 pom.xml 在同一个目录下,且名字为 .to

olkit deploy.yaml (注意:文件名最前面有个小数点),可不填此参数,插件会自动获取取。

- o -Ddeploy version : 指定部署的版本号,优先级高于部署配置文件中的version配置。
 - ? 说明 toolkit-maven-plugin插件 1.0.6 及以后版本支持配置该参数。

执行该打包命令后,系统显示如下结果,当回显信息中显示"BUILD SUCCESS"表示部署成功。

更多配置项

1. 打包参数

打包文件支持的参数如下所示。

```
apiVersion: V1
kind: AppPackage
packageType: #应用部署包类型,支持War、FatJar、Image、url;您只有在该处配置了URL,那么packa
geUrl才能生效。
imageUrl:
          #镜像地址,Image包类型应用可填入。
packageUrl: #包地址, War、FatJar类型应用可填入。
build:
  docker:
    dockerfile: #Docker镜像构建文件。如您希望在本地构建镜像部署,需填入此字段。
    imageRepoAddress: #阿里云镜像仓库地址。如您希望在本地构建镜像部署,需填入此字段。
    imageTag: #镜像Tag。如您希望在本地构建镜像部署,需填入此字段。
                  #阿里云镜像仓库用户名。如您希望在本地构建镜像部署,需填入此字段。
    imageRepoUser:
    imageRepoPassword: #阿里云镜像仓库密码。如您希望在本地构建镜像部署,需填入此字段。
                   #目标存储桶名称。如您希望使用自定义的OSS仓库存储部署包,需填入此字
        bucket:
段。
                    #OSS自定义路径。如您希望使用自定义的OSS仓库存储部署包,需填入此字
        key:
段。
        accessKeyId: #OSS账号。如您希望使用自定义的OSS仓库存储包,需填入此字段。
        accessKeySecret: #OSS密码。如您希望使用自定义的OSS仓库存储包,可填入此字段。
```

2. 部署参数

单击此处展开部署文件支持的参数。

```
apiVersion: V1
kind: AppDeployment
spec:
 type: kubernetes
 target:
   appName:
            #应用名称。
   namespaceId: #应用所在微服务空间。
            #应用ID。插件会使用此应用进行部署,如未填入则使用namespaceId和appName来查找应
  appId:
用进行部署。
           #部署版本号,默认使用日时分秒格式。
           #部署的包依赖的JDK版本,JDK支持版本为Open JDK 7和Open JDK 8。镜像不支持。
   webContainer: #部署的包依赖的Tomcat版本, WebContainer支持apache-tomcat-7.0.91。镜像不
支持。
   batchWaitTime: #分批等待时间。
   command: #镜像启动命令。该命令必须为容器内存在的可执行的对象。例如: sleep。设置该命令
将导致镜像原本的启动命令失效。
```

```
commandArgs: #現隊后切而令梦致。上述后切而令所需梦致。
           #容器环境变量参数。
 envs:
   - name: envtmp0
    value: '0'
   - name: envtmp1
    value: '1'
 liveness: #容器健康检查,健康检查失败的容器将停止运行并恢复。
   exec: #exec、tcpSocket和httpGet三者选其一配置
    command:
      - sleep
      - 1s
   tcpSocket: #exec、tcpSocket和httpGet三者选其一配置
    host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
    port: "18081" #端口(注意,此处为字符串类型)
   httpGet: #exec、tcpSocket和httpGet三者选其一配置
    host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
    port: "18081" #端口 (注意,此处为字符串类型)
    path: "/health"
    scheme: "HTTP" #可选值为HTTP、HTTPS
    httpHeaders:
     - name: "color"
       value: "blue"
   initialDelaySeconds: 5
   timeoutSeconds: 11
   periodSeconds: 5
   successThreshold: 1 #固定为1,不可修改
   failureThreshold: 3
 readiness: #应用启动状态检查,多次健康检查失败的容器将停止运行并重启。不通过健康检查的容器将
不会有SLB流量进入。
   exec: #exec、tcpSocket和httpGet三者选其一配置
    command:
      - sleep
      - 1s
   tcpSocket: #exec、tcpSocket和httpGet三者选其一配置
    host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
    port: "18081" #端口 (注意,此处为字符串类型)
   httpGet: #exec、tcpSocket和httpGet三者选其一配置
    host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
    port: "18081" #端口 (注意,此处为字符串类型)
    path: "/health"
    scheme: "HTTP" #可选值为HTTP、HTTPS
    httpHeaders:
     - name: "color"
       value: "blue"
   initialDelaySeconds: 5
   timeoutSeconds: 11
   periodSeconds: 5
   successThreshold: 2
   failureThreshold: 3
 preStop: #生命周期挂钩,在容器删除前触发执行
   exec:
    command:
      - /bin/bash
```

```
- ls /tmp
 httpGet:
   host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
   port: "18081" #端口(注意,此处为字符串类型)
   path: "/health"
   scheme: "HTTP" #可选值为HTTP、HTTPS
   httpHeaders:
    - name: "color"
     value: "blue"
postStart: #生命周期挂钩,在容器创建后立即执行
   command:
    - /bin/bash
    - -c
    - ls /tmp
 httpGet:
   host: "192.168.1.109" #此配置项可选,如未设置则默认为Pod IP
   port: "18081" #端口(注意,此处为字符串类型)
   path: "/health"
   scheme: "HTTP" #可选值为HTTP、HTTPS
   httpHeaders:
    - name: "color"
     value: "blue"
configMountDescs: #配置管理
 - type: "ConfigMap" #可选值为 ConfigMap、Secret
   name: "configtest"
   mountPath: "/home/admin" #如此值为目录,则需配置mountItems来挂载到文件
   mountItems: #挂载到文件
    - key: "test-name"
      path: "test"
   useSubPath: true #当挂载到文件时,此值为true表示保留源文件,为false表示覆盖源文件
javaStartUpConfig: #Java启动参数配置,根据实际情况设置相应值
 initialHeapSize: #初始堆内存
  original: 1000
   startup: "-Xms1000m"
 maxHeapSize: #最大堆内存
  original: 1000
  startup: "-Xmx1000m"
 newSize: #初始新生代
   original: 200
   startup: "-XX:NewSize=200m"
 maxNewSize: #最大新生代
   original: 200
   startup: "-XX:MaxNewSize=200m"
 survivorRatio: #Eden/Survivor比例
  original: 2
   startup: "-XX:SurvivorRatio=2"
 newRatio: #老年代/新生代比例
  original: 8
   startup: "-XX:NewRatio=8"
 permSize: #持久代内存
   original: 512
   startup: "-XX:PermSize=512m"
maxPermSize: #最大持久代内存
```

```
original: 512
     startup: "-XX:MaxPermSize=200m"
   maxDirectMemorySize: #直接内存
     original: 100
     startup: "-XX:MaxDirectMemorySize=100m"
   threadStackSize: #栈大小
     original: 500
     startup: "-XX:ThreadStackSize=500"
   hsfserverPort: #HSF端口
     original: 12200
     startup: "-Dhsf.server.port=12200"
   hsfserverMinPoolSize: #HSF最小线程池
     original: 50
     startup: "-Dhsf.server.min.poolsize=50"
   hsfserverMaxPoolSize: #HSF最大线程池
     original: 720
     startup: "-Dhsf.server.max.poolsize=720"
   youngGarbageCollector: #年轻代GC策略
     original: "UseSerialGC" #可选值为 UseSerialGC、UseG1GC、UseParNewGC、UseParallelGC
     startup: "-XX:+UseSerialGC" # 当值为UseG1GC、UseParNewGC 或 UseParallelGC 时,此值
为空字符串
   oldGarbageCollector:
                         #老年代GC策略
     original: "UseConcMarkSweepGC" #可选值为 UseConcMarkSweepGC、UseSerialGC、UseG1GC
, UseConcMarkSweepGC, UseParNewGC, UseParallelOldGC, UseParallelGC
     startup: "-XX:+UseConcMarkSweepGC" #格式为-XX:+<GC策略>
   concGCThreads:
                  #并发GC线程数
     original: 5
     startup: "-XX:ConcGCThreads=5"
   parallelGCThreads: #并行GC线程数
     original: 5
     startup: "-XX:ParallelGCThreads=5"
   glHeapRegionSize: #Gl Region文件大小
     original: 50
     startup: "-XX:G1HeapRegionSize=50m"
   gclogFilePath: #gc日志目录
     original: "/tmp/"
     startup: "-Xloggc:/tmp/"
   useGCLogFileRotation: #GC日志滚动
     original: true
     startup: "-XX:+UseGCLogFileRotation"
   numberOfGCLogFiles: #GC日志数量
     original: 5
     startup: "-XX:NumberOfGCLogFiles=5"
   gclogFileSize: #GC日志大小
     original: 100
     startup: "-XX:GCLogFileSize=100m"
   heapDumpOnOutOfMemoryError: #打开OOM Dump
     original: true
     startup: "-XX:+HeapDumpOnOutOfMemoryError"
   heapDumpPath: #OOM Dump文件路径
     original: "/tmp/dumpfile"
     startup: "-XX:HeapDumpPath=/tmp/dumpfile"
   customParams: #自定义参数
     original: "-Dtest=true"
```

```
startup: "-Dtest=true"
deployAcrossZones: "true" #尽量多可用区部署 (注意,此值为字符串类型)
deployAcrossNodes: "true" #尽量多节点部署(注意,此值为字符串类型)
customTolerations: #调度容忍
  - key: aa
   operator: Exists
   effect: NoSchedule
  - key: bb
   operator: Equal
   value: "111"
    effect: "NoExecute"
    tolerationSeconds: 111
customAffinity: #自定义调度规则
  nodeAffinity: #节点亲和性规则
    requiredDuringSchedulingIgnoredDuringExecution:
     nodeSelectorTerms:
     - matchExpressions:
        - key: "beta.kubernetes.io/arch"
         operator: "Gt"
         values:
          - "11"
    \verb|preferredDuringSchedulingIgnoredDuringExecution:|\\
    - weight: 1
     preference:
       matchExpressions:
        - key: "beta.kubernetes.io/arch"
         operator: "Gt"
         values:
          - "11"
  podAffinity: #应用 (Pod) 亲和性规则
    required \verb|DuringSchedulingIgnoredDuringExecution|:
    - namespaces:
      - "default"
     topologyKey: "failure-domain.beta.kubernetes.io/zone"
     labelSelector:
       matchExpressions:
        - key: "beta.kubernetes.io/arch"
         operator: "In"
         values:
          - "11"
    \verb|preferredDuringSchedulingIgnoredDuringExecution:|\\
    - weight: 1
     podAffinityTerm:
       namespaces:
        - "default"
       topologyKey: "failure-domain.beta.kubernetes.io/region"
       labelSelector:
         matchExpressions:
         - key: "beta.kubernetes.io/arch"
           operator: "Exists"
           values: []
                   #应用(Pod)反亲和性规则
  podAntiAffinity:
    required {\tt DuringSchedulingIgnoredDuringExecution:}
    - namespaces:
```

```
- "default"
  topologyKey: "kubernetes.io/hostname"
 labelSelector:
   matchExpressions:
    - key: "beta.kubernetes.io/arch"
     operator: "In"
      values:
      - "11"
preferredDuringSchedulingIgnoredDuringExecution:
- weight: 2
 podAffinityTerm:
   namespaces:
    - "default"
   topologyKey: "kubernetes.io/hostname"
   labelSelector:
      matchExpressions:
      - key: "beta.kubernetes.io/arch"
       operator: "In"
       values:
        - "11"
```

典型场景示例

典型部署场景及相关配置示例。

- 场景一:本地构建War(或FatJar)包进行部署 假设您在北京环境有WAR(或FatJar)类型的EDAS应用,期望本地构建WAR(或FatJar)进行部署。打包 配置和部署配置如下所示:
 - 。 打包文件:

```
apiVersion: V1
kind: AppPackage
spec:
   packageType: War
```

○ 部署文件:

```
apiVersion: V1
kind: AppDeployment
spec:
    type: kubernetes
    target:
    appId: #应用ID。插件会使用此应用进行部署,如未填入则使用namespaceId和appName来查找
应用进行部署。
    namespaceId: #【可选】微服务空间,如不清楚appId,可使用此微服务空间及应用名称进行部署。
    appName: #【可选】应用名称,如不清楚appId,可使用此微服务空间及应用名称进行部署。
```

● 场景二:使用已有镜像地址部署镜像类型应用 假设您在北京环境有一个镜像类型应用,期望使用已有的镜像(registry.cnbeijing.aliyuncs.com/test/gateway:latest)部署应用。打包配置和部署配置如下所示:

。 打包文件:

```
apiVersion: V1
kind: AppPackage
spec:
   packageType: Image
   imageUrl: registry.cn-beijing.aliyuncs.com/test/gateway:latest
```

○ 部署文件:

```
apiVersion: V1
kind: AppDeployment
spec:
    type: kubernetes
    target:
    appId: #应用ID。插件会使用此应用进行部署,如未填入则使用namespaceId和appName来查找
应用进行部署。
    namespaceId: #【可选】微服务空间,如不清楚appId,可使用此微服务空间及应用名称进行部署。
    appName: #【可选】应用名称,如不清楚appId,可使用此微服务空间及应用名称进行部署。
```

● 场景三:本地构建镜像上传至仓库并部署应用 假设您在北京环境有镜像类型应用,期望在本地编译并构建为镜像,并上传到阿里云镜像仓库进行部署, 打包配置和部署配置如下所示:

。 打包文件:

```
apiVersion: V1
kind: AppPackage
spec:
   packageType: Image
   build:
   docker:
        dockerfile: Dockerfile #指定Dockerfile。
        imageRepoAddress: #镜像仓库地址。
        imageTag: #镜像Tag。
        imageRepoUser: #镜像仓库用户名。
        imageRepoPassword: #镜像仓库密码。
```

○ 部署文件:

```
apiVersion: V1
kind: AppDeployment
spec:
    type: kubernetes
    target:
    appId: #应用ID。插件会使用此应用进行部署,如未填入则使用namespaceId和appName来查找
应用进行部署。
    namespaceId: #【可选】微服务空间,如不清楚appId,可使用此微服务空间及应用名称进行部署。
    appName: #【可选】应用名称,如不清楚appId,可使用此微服务空间及应用名称进行部署。
```

4.4. 金丝雀发布(K8s)

4.4.1. 使用控制台金丝雀发布应用(K8s)

对于部署在K8s集群中的Spring Cloud或Dubbo微服务应用,为了确保升级操作的安全性,可以使用金丝雀发布(即灰度发布)进行小规模验证,验证通过后再全量升级。

使用限制

- HSF应用:不支持金丝雀发布。
- Dubbo应用: 没有使用限制。
- Spring Cloud应用:如果应用依赖Deployment.Metadata.Name或Deployment.Metadata.Uid的功能与配置,请勿使用金丝雀发布,否则发布后该原生功能将出现异常。

视频教程

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击应用列表,在顶部菜单栏选择地域并在页面上方选择微服务空间。
- 3. 在左侧导航栏单击应用列表,在顶部菜单栏选择地域,并在页面上方选择微服务空间。
- 4. 在**应用列表**页面的**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 5. 在应用总览页面右上角选择部署 > 部署。
- 6. 在选择部署模式页面的金丝雀发布(灰度)区域右上角单击开始部署。
- 7. 在金丝雀发布(灰度)页面设置新版本应用的部署参数、发布策略和灰度规则,然后单击确定。

i. 设置部署参数。

部署参数说明

参数	描述
配置镜像(仅适用于镜像部署的应用)	镜像不可更改, 只能更新镜像版本。
应用运行环境 (适用于JAR包和WAR包部署的应用)	需要和上一次部署保持一致。 ■ JAR包部署:应用运行环境为标准Java应用运行环境,不可更改。 ■ WAR包部署:应用运行环境为Apache Tomcat,类型不可更改。如果需要,可以更换版本。
Java环境(适用于JAR包和WAR包部署的应用)	根据需求在列表中选择。
当前环境	显示当前应用的运行环境。仅是JAR包和WAR包部署的应用,才会显示当前应用的运行环境。EDAS会自动将您应用的Java环境或应用运行环境升级到最新版本。
文件上传方式(适用于JAR包和WAR包部署的应用)	部署包类型(WAR包或JAR包)需要和上一次部署保持一致,不可更改。上传方式可以根据实际需求选择上传部署包(上传JAR包或上传WAR包)或输入部署包地址两种方式(JAR包地址或WAR包地址)。
版本(适用于JAR包和WAR包部署的应用)	部署包版本,可以用时间戳作为版本号。
时区(适用于JAR包和WAR包部署的应用)	根据实际需求在列表中选择。
服务注册与发现配置	选择注册中心的运维方式。如何选择,请参见 <mark>选择</mark> 注册中心运维方式。

ii. 在**发布策略**区域配置发布策略参数。

发布策略参数说明

参数	说明			
首批灰度数量	首批发布的应用实例数量。右侧会显示应用当前实例数,为了保证应用稳定性,灰度实例数不能超过应用实例总数的50%。			
	② 说明 灰度分组发布后,必须手动开始剩余分批发布。			
剩余批次	首批发布后,剩余的应用实例按照此处设定的批次完成分批发布。			
分批间处理 方式	支持以下处理方式: 自动:根据间隔自动分配分批发布。间隔即剩余批次间的发布时间间隔,单位为分钟。 手动:手动触发下一个批次的发布。 ① 说明 仅当剩余批次大于1时,才会显示分批间处理方式。			
批次内部署间隔	每一批次内,如果应用实例数大于1,则应用实例会按此时间间隔部署,单位为秒。			

右侧的发布策略配置信息会根据配置显示金丝雀发布流程。

iii. 设置灰度规则。

EDAS支持的灰度规则包括按内容灰度和按比例灰度。

灰度规则参数说明

页签	参数	说明
按内容灰度条件模式条件列表	协议类型	■ Spring Cloud: 需要设置path。 ■ Dubbo: 需要选择服务和方法。
	条件模式	选择同时满足下列条件或满足下列任一条件。
	条件列表	■ Spring Cloud:可根据Cookie、Header或Parameter类型设置具体的参数。 ■ Dubbo:根据应用实际的参数、参数值获取表达式设置。
按比例灰度	流量比例	流量会按配置的比例被转发到当前的灰度分组。

② 说明 单击创建流量规则,即可创建多个入口流量规则,多个规则可以同时生效。

iv. (可选)配置应用高级设置。

金丝雀发布启动后,EDAS将先在指定的灰度分组中部署新版本应用。**变更记录**页面将显示部署进度和状态。

- ② 说明 如需监控灰度流量是否符合预期,请参见监控灰度流量。
- 8. 灰度流量验证完成后,在**变更记录**页面右侧单击**开始下一批**,完成后续分批发布。 如果在验证过程中发现新版本应用有问题,可以在**变更记录**页面右上角单击**回滚**,并在弹出的**确认**对 话框单击**确认**。

结果验证

金丝雀发布完成后,在应用总览页面顶部查看部署包是否为新部署的应用版本。

4.4.2. 使用API在K8s集群中金丝雀发布应用

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中金丝雀发布应用(JAR包、WAR包和镜像)。

前提条件

金丝雀发布应用前,您必须提前完成以下准备工作:

- 确定应用部署地域,假设为cn-hangzhou。
- 在EDAS k8s环境中创建应用。目前支持多种创建应用的方式,请根据实际需求选择:
 - o 使用API在K8s集群中创建Java应用(JAR包或WAR包)
 - 使用API创建Java应用(镜像部署)
 - o 使用API在K8s集群中创建多语言应用

如果已经创建应用,您可以调用ListApplication接口查询应用列表,获取目标应用的APPId,假设为6bbc57 a2-a017-4bec-b521-49a15bd3****。

- 已将应用升级部署包(JAR包、WAR包和镜像)上传至目标地址。
 - JAR包或WAR包:本示例以OSS存储路径为例,假设为https:doc***.oss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.2-SNAPSHOT.jar。
 - 镜像: 假设镜像仓库为*image-demo-project*, 镜像地址为*registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/provider:2.0*。

背景信息

使用API金丝雀发布应用,在首批灰度发布后,您必须手动确认继续剩余批次的发布。在确认剩余批次发布前,您可以进行小规模验证,验证后,再继续剩余批次的发布,即将应用全量升级到新版本。

使用API金丝雀发布应用,设置剩余批次(除首批灰度发布外)的发布方式时推荐选用自动分批发布。如果您选用手动分批发布方式,则还需要调用ContinuePipeline接口手动确认执行下一批发布。

此处提供两种金丝雀发布策略示例,请根据您的实际场景选择并修改策略。

● 首批灰度发布2个Pod实例+剩余Pod实例分2批发布+自动分批+分批间隔1分钟。

```
{"type":"GrayBatchUpdate", "batchUpdate":{"batch":2, "releaseType":"auto", "batchWaitTime":1
}, "grayUpdate":{"gray":2}}
```

● 首批灰度发布2个Pod实例+剩余Pod实例分2批发布+手动分批。

```
{"type":"GrayBatchUpdate", "batchUpdate":{"batch":2, "releaseType":"manual"}, "grayUpdate":{
   "gray":2}}
```

使用JAR包或WAR包金丝雀发布应用

1. 运行以下示例代码,使用JAR包或WAR包金丝雀发布应用。

以下代码适用于在EDAS K8s集群中金丝雀发布应用(JAR包)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

② 说明 如果您需要使用WAR包金丝雀应用,请根据需要在示例代码中增加以下两个参数:

```
//部署包依赖的Tomcat版本。适用于通过WAR包部署的Spring Cloud和Dubbo应用。
request.setWebContainer("apache-tomcat-7.0.91");
//部署包依赖的EDAS Container版本。适用于通过WAR包部署的HSF应用。
request.setEdasContainerVersion("3.5.9");
```

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun use
r ak, aliyun user sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       //应用ID。
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //JAR包或者WAR包地址、版本。
       request.setPackageUrl("https:doc***.oss-cn-hangzhou.aliyuncs.com/sc-****-D-0.0.
2-SNAPSHOT.jar");
       request.setPackageVersion("2021-04-15 16:41:52");
       //部署包依赖的JDK版本。可选的参数值为Open JDK 7和Open JDK 8。
       request.setJDK("Open JDK 8");
       //自定义分批发布策略。此处示例值代表"首批灰度发布2个Pod实例+剩余Pod实例分2批发布+自动分批
+分批间隔1分钟"。
       request.setUpdateStrategy("{\"type\":\"GrayBatchUpdate\",\"batchUpdate\":{\"bat
ch\":2,\"releaseType\":\"auto\",\"batchWaitTime\":1},\"grayUpdate\":{\"gray\":2}}");
       //灰度发布流量控制策略。示例策略表示: 50%的流量会被转发到灰度分组。
       request.setTrafficControlStrategy("{\"http\":{\\"rules\\":[{\\"conditionType\\":\\"p
ercent\",\"percent\":50}]}}");
       //应用实例数。
       request.setReplicas(4);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=5886e6f9-05b1-42f0-a3d4-5d90558e****
Message=success
```

2. 调用Get ChangeOrderInfo接口,获取金丝雀发布应用变更的相关批次的流程ID(PipelineId)。 调用Get ChangeOrderInfo接口,返回结果如下:

② 说明 示例代码中使用的是首批灰度发布2个Pod实例+剩余Pod实例分2批发布+自动分批+分批间隔1分钟的金丝雀发布策略,在首批灰度发布2个Pod实例流程运行完以后,需要手动确认继续剩余批次的发布。在调用Get ChangeOrderInf o接口返回结果中出现3个流程ID(PipelineId),第二次出现的流程ID(PipelineId)代表需要手动确认继续剩余批次的发布。剩余批次的策略是剩余Pod实例分2批发布+自动分批+分批间隔1分钟,则无需手动确认,系统自动完成剩余批次的发布。

3. 调用ContinuePipeline接口,手动确认继续剩余批次的发布。

返回结果如下:

```
{
  "Message": "success",
  "RequestId": "050392EE-959D-4144-AA02-3A1E7CE2F736",
  "Code": 200
}
```

使用镜像金丝雀发布应用

1. 运行程序,使用镜像金丝雀发布应用。

以下代码适用于在EDAS K8s集群中金丝雀发布应用(镜像)。代码中未设置调度规则、启动命令、环境变量等高级参数,如需了解更多API参数信息,请参见DeployK8sApplication。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun use
r ak, aliyun user sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       //应用ID。
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //镜像地址。
       request.setImage("registry-vpc.cn-hangzhou.aliyuncs.com/image-demo-project/prov
ider:2.0");
       //自定义金丝雀发布策略。此处示例值代表"首批灰度发布2个Pod实例+剩余Pod实例分2批发布+手动分
批"。
       request.setUpdateStrategy("{\"type\":\"GrayBatchUpdate\",\"batchUpdate\":{\"bat
ch\":2,\"releaseType\":\"manual\"},\"grayUpdate\":{\"gray\":2}}");
       //灰度发布流量控制策略。示例策略表示: 50%的流量会被转发到灰度分组。
       request.setTrafficControlStrategy("{\"http\":{\"rules\":[{\"conditionType\":\"p
ercent\", \"percent\":50}]}}");
       //应用实例数。
       request.setReplicas(4);
       //应用运行过程中,应用实例的CPU限额、内存限额等,0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=a7431f9e-7e05-4590-aef8-17088f0****
Message=success
```

2. 调用GetChangeOrderInfo接口,获取金丝雀发布应用变更的相关批次的流程ID(PipelineId)。 调用GetChangeOrderInfo接口,返回结果如下:

② 说明 示例代码中使用的是首批灰度发布2台+剩余分2批发布+手动分批的金丝雀发布策略,在首批灰度发布2个Pod实例流程运行完以后,需要手动确认继续剩余批次的发布。在调用Get ChangeOrderInf o接口返回结果中出现3个流程ID(PipelineId),第二次出现的流程ID(PipelineId)代表需要手动确认继续剩余批次的发布,第三次出现的流程ID(PipelineId)代表剩余批次中第一批完成后需要手动确认继续下一批次的发布。

3. 调用ContinuePipeline接口,手动确认继续剩余批次发布。 返回结果如下:

```
"Message": "success",
"RequestId": "B8ECF6F7-672D-40E9-91DD-1C33F06D4FD8",
"Code": 200
```

4. 调用ContinuePipeline接口,手动确认下一批发布。

返回结果如下:

```
"Message": "success",
 "RequestId": "7BB4F043-7C28-4A0E-B6A5-D4023EB24388",
 "Code": 200
}
```

② 说明 示例代码中使用的是首批灰度发布2个Pod实例+剩余Pod实例分2批发布+手动分批的金 丝雀发布策略,运行到剩余批次发布过程,只需要手动确认一批。如果您是剩余批次需要手动确认 多批的场景,请选择正确批次的流程ID (PipelineId)并重复执行此步骤。

重复执行此步骤时,请等待批次内部署间隔时长,默认10秒。

结果验证

您在金丝雀发布应用后,可以调用Get Change OrderInf o接口查看变更流程详情,获取金丝雀发布应用的变 更状态。具体API参数详情,请参见Get Change OrderInfo。

调用Get Change OrderInfo接口查看变更流程详情,返回的执行结果如下:

```
"Message": "success",
"RequestId": "10DA427D-AF74-4ECE-BB0A-EECC0EBF7548",
"Code": 200,
"changeOrderInfo": {
   "Status": 2,
   "Desc": "金丝雀发布 (***)",
   "PipelineInfoList": {
       "PipelineInfo": [
           {
              //此处未罗列全部执行结果,仅供参考。
```

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断金丝雀发布应用的变更是否成 功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

4.5. 回滚应用(K8s)

4.5.1. 使用控制台回滚应用(K8s)

应用在升级过程中或升级完成后,发现新版本应用有问题或异常,可以回滚应用,待问题解决后再升级。本文介绍如何使用EDAS控制台回滚应用。

背景信息

在应用升级过程中,出现部署包不可用、健康检查失败等异常报错会导致应用升级失败,当前应用变更会自动终止并回滚。

如果您发现新版本应用有问题或异常,可以手动回滚应用,有以下两种场景:

- 升级中回滚:可以终止升级变更流程并回滚应用。
- 升级后回滚,可以通过部署历史版本回滚应用。

升级中回滚

本示例以在手动模式的分批发布变更过程中回滚应用为例。当您需要分批发布应用时,可优先考虑使用手动模式。手动模式的分批发布更加可控,当前面批次发布完成之后,可以验证新版本应用的正确性,以决定下一批次是否继续发布。

- 经验证,新版本应用符合预期,则手动确认继续下一批次的发布。
- 经验证,新版本应用存在异常问题,则可以回滚应用。待修复问题后,再重新发布应用。
 - 1. 分批发布升级应用。

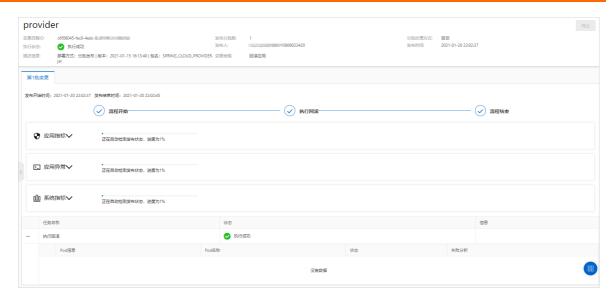
具体操作,请参见使用控制台分批发布应用(K8s)。

- 2. 在左侧导航栏,单击变更记录。
- 3. 在变更记录选择目标变更记录,单击操作列下的查看。
- 4. 在右上角单击回滚。



5. 在确认对话框,单击确认。

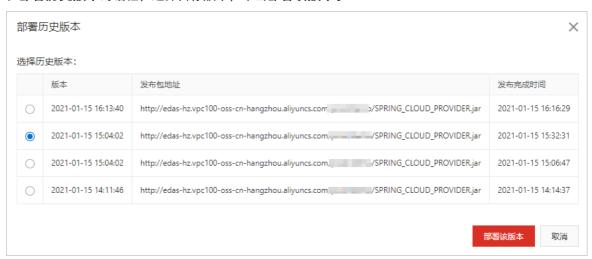
确认手动回滚应用后,自动终止应用升级流程并回滚到之前的版本。**执行状态**显示为**执行成功**,即回滚应用成功。



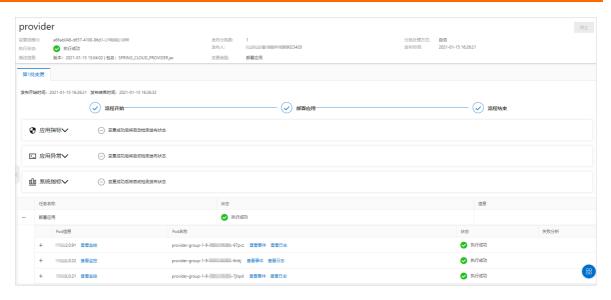
升级后回滚

本示例以完成分批发布升级应用后,通过部署历史版本回滚应用为例。

- 1. 在应用总览页面,选择部署 > 部署历史版本。
- 2. 在部署历史版本对话框,选择目标版本,单击部署该版本。



- 3. 在左侧导航栏,单击变更记录。
- 4. 在**变更记录**选择目标变更记录,单击操作列下的查看。 确认部署历史版本后,系统开始回滚到之前的版本。**执行状态**显示为**执行成功**,即部署历史版本应用 成功。



4.5.2. 使用API在K8s集群中回滚应用到目标历史版本

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中回滚应用到目标历史版本。

前提条件

回滚应用到目标历史版本前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确定应用部署地域,假设为cn-hangzhou。
- 您的应用已经升级版本。目前支持多种升级应用的方式,请根据实际需求选择:
 - o 使用API在K8s集群中单批发布应用
 - o 使用API在K8s集群中分批发布应用
 - o 使用API在K8s集群中金丝雀发布应用
- 调用List Hist ory Deploy Version接口查询应用历史版本列表,获取目标历史版本的Package Version.ld,假设为 *e4aeb2972cde64d5ff0618966102*****。

背景信息

您可以调用DeployK8sApplication接口,部署应用的历史版本,即可实现回滚应用到历史版本。本文仅提供简单示例代码实现部署应用的历史版本,如您需要了解更多部署应用的信息,请参见DeployK8sApplication。

回滚到目标历史版本

以下代码适用于在EDAS K8s集群中回滚应用到目标历史版本。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationRequest;
import com.aliyuncs.edas.model.v20170801.DeployK8sApplicationResponse;
public class DeployK8sApplication {
   public static void main(String[] args) {
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun_user_ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun_user_sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       DeployK8sApplicationRequest request = new DeployK8sApplicationRequest();
       //应用ID。
       request.setAppId("6bbc57a2-a017-4bec-b521-49a15bd3****");
       //通过ListHistoryDeployVersion接口查询到的目标历史版本的PackageVersion.Id。
       request.setPackageVersionId("e4aeb2972cde64d5ff0618966102****");
       //应用实例数。
       request.setReplicas(4);
       //应用运行过程中,应用实例的CPU限额、内存限额等, 0表示不限制。
       request.setCpuLimit(0);
       request.setMemoryLimit(0);
       request.setCpuRequest(0);
       request.setMemoryRequest(0);
       //CPU最小资源需求,单位:核数。0表示不限制。
       request.setMcpuRequest(0);
       //CPU能使用的最大值,单位:核数。0表示不限制。
       request.setMcpuLimit(0);
       //变更记录描述。
       request.setChangeOrderDesc("回滚到目标历史版本");
       trv {
           DeployK8sApplicationResponse response = client.getAcsResponse(request);
           System.out.println("ChangeOrderId=" + response.getChangeOrderId() + "\nMessage=
" + response.getMessage());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
    }
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=1b0bec74-7ec3-43f7-b9c8-f8e746d2****
Message=success
```

结果验证

您在回滚应用到目标历史版本后,可以调用Get ChangeOrderInfo接口查看变更流程详情,获取回滚应用的变更状态。具体API参数详情,请参见Get ChangeOrderInfo。

调用Get Change OrderInf o接口查看变更流程详情,返回的执行结果如下:

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断回滚应用到历史版本的变更是否成功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示应用回滚到目标版本失败。您可以根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新回滚应用到目标版本。

4.5.3. 使用API在K8s集群中终止应用变更并回滚

本文介绍如何使用EDAS提供的Java SDK调用API在K8s集群中终止应用升级变更流程并回滚。

前提条件

在终止应用升级变更流程并回滚前,您必须提前完成以下准备工作:

- 安装EDAS的Java SDK。更多信息,请参见Java SDK接入指南。
- 确保目标应用处于升级变更流程中,请根据实际需求选择应用升级方式:
 - o 使用API在K8s集群中单批发布应用
 - o 使用API在K8s集群中分批发布应用

- o 使用API在K8s集群中金丝雀发布应用
- (可选)调用ListRecent ChangeOrder接口,查询应用变更记录,获取应用正在升级变更的变更单流程ID(ChangeOrderId)。假设为47e2b863-adc8-4a0e-ac23-8361b95a****。
 如果您在升级应用时已经记录变更单流程ID(ChangeOrderId),则无需调用该接口。

示例代码

以下代码适用于在EDAS K8s集群中快速终止应用升级变更流程并回滚。

```
import com.aliyuncs.DefaultAcsClient;
import com.aliyuncs.exceptions.ClientException;
import com.aliyuncs.exceptions.ServerException;
import com.aliyuncs.profile.DefaultProfile;
import com.aliyuncs.edas.model.v20170801.AbortAndRollbackChangeOrderRequest;
import com.aliyuncs.edas.model.v20170801.AbortAndRollbackChangeOrderResponse;
public class AbortAndRollbackChangeOrder {
    public static void main(String[] args)
       //阿里云账号或RAM用户的AccessKey ID。
       String aliyun user ak = "<yourAccessKeyId>";
       //阿里云账号或RAM用户的AccessKey Secret。
       String aliyun user sk = "<yourAccessKeySecret>";
       //应用所在地域ID。
       String region id = "cn-hangzhou";
       DefaultProfile defaultProfile = DefaultProfile.getProfile(region id, aliyun user ak
, aliyun user sk);
       DefaultAcsClient client = new DefaultAcsClient(defaultProfile);
       //创建API请求,并设置参数。
       AbortAndRollbackChangeOrderRequest request = new AbortAndRollbackChangeOrderRequest
();
       //应用的升级变更单流程ID。
       request.setChangeOrderId("47e2b863-adc8-4a0e-ac23-8361b95a****");
       try {
           AbortAndRollbackChangeOrderResponse response = client.getAcsResponse(request);
           System.out.println("ChangeOrderId=" + response.getData().getChangeOrderId() + "
\nMessage=" + response.getMessage() + "\nTraceId=" + response.getTraceId());
       } catch (ServerException e) {
           e.printStackTrace();
       } catch (ClientException e) {
           e.printStackTrace();
```

运行程序,返回的执行结果如下:

```
ChangeOrderId=aef94ee3-3142-4abc-9d7c-57249686****
Message=success
TraceId=210e815316184954351417778d****
```

结果验证

您在终止应用升级变更流程并回滚后,可以调用Get ChangeOrderInf o接口查看变更流程详情,获取回滚应用的变更状态。具体API参数详情,请参见Get ChangeOrderInfo。

调用Get ChangeOrderInf o接口查看变更流程详情,返回的执行结果如下:

请查看上述执行结果中的changeOrderInfo.Status参数值,通过该值判断回滚应用的变更是否成功。changeOrderInfo.Status的取值如下:

- 0: 准备
- 1: 执行中
- 2: 执行成功
- 3: 执行失败
- 6: 终止
- 8: 手动分批发布模式下,等待手工确认执行下一批。
- 9: 自动分批发布模式下,等待下一批执行中。
- 10: 系统异常执行失败

② 说明 如果返回结果中的changeOrderInfo.Status值为 3, 即表示终止应用变更并回滚失败。您可以根据返回结果中的TaskMessage参数的值分析错误原因,解决问题后重新升级应用。

5.CI/CD

5.1. CI/CD概述(K8s)

应用在不断迭代过程中,需要持续集成(CI)和持续部署(CD)。EDAS支持通过Jenkins和云效对您部署的应用进行CI/CD。

使用Jenkins进行持续集成和部署

Jenkins是一个开源工具,帮助您持续、自动的构建和测试软件项目、监控外部任务的运行。

您可以使用Git Lab托管您的代码,然后使用Jenkins实现EDAS应用的持续集成。更多信息,请参见使用Jenkins创建持续集成。

在Jenkins环境中集成了edas-jenkins-plugin插件,您可以使用该插件在Jenkins内快速完成应用持续部署到EDAS。更多信息,请参见在Jenkins中使用EDAS插件构建应用部署的任务。

使用云效进行持续集成和部署

阿里云云效是云原生时代新DevOps平台。我们可以使用阿里云云效中的云效流水线Flow以完成应用的持续 集成和部署,云效流水线Flow提供了灵活易用的持续集成、持续验证和持续发布功能。

您可以将代码上传到云效流水线Flow支持的代码库,然后使用云效流水线Flow完成应用的持续集成和部署。更多信息,请参见使用云效2020进行持续集成和部署(K8s)。

5.2. 云效

5.2.1. 使用云效2020进行持续集成和部署(K8s)

应用在不断迭代过程中,需要持续集成(CI)和持续部署(CD),通过自动化的构建集成、测试和循环部署,实现快速交付高质量产品并降低人力成本。本文介绍使用云效2020完成应用的持续集成和部署。

前提条件

- 在EDAS控制台已经部署应用,具体操作步骤,请参见在容器服务K8s集群中使用JAR包或WAR包部署应用。
- 开通云效2020
- 将业务代码上传至代码库。
 - 云效支持的代码库,请参见代码源。
 - 业务代码中须包含应用的Dockerfile文件,具体制作步骤,请参见构建仓库与镜像。

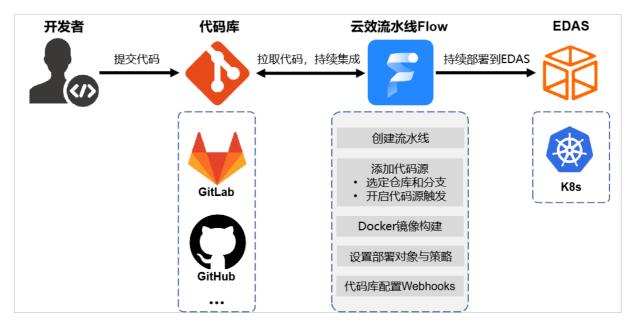
背景信息

使用云效2020实现应用的持续集成和部署,根据您期望的部署环境不同,有以下两种场景:

- 部署到EDAS ECS环境: 只能使用JAR包或者WAR包构建、部署。示例文档,请参见使用云效2020部署Java应用至ECS集群。
- 部署到EDAS K8s环境:只能使用镜像构建、部署。

② 说明 在EDAS K8s环境部署应用时,根据不同应用类型,包含JAR包、WAR包和镜像3种部署方式;在云效2020侧持续集成和部署应用到EDAS K8s环境时,只能以镜像方式构建和部署。

使用云效2020实现应用的持续集成和部署,您只需要将开发好的新版本应用代码提交到代码库,云效流水线 Flow会监听代码事件,当满足触发事件时会触发流水线运行,部署新版本应用到EDAS K8s环境。



代码开发与提交一般由开发人员执行,本文不做介绍,本文仅介绍在云效流水线Flow侧如何创建流水线,并配置代码源、镜像仓库、部署对象和发布策略等操作流程。

(可选)

步骤一:在云效创建企业

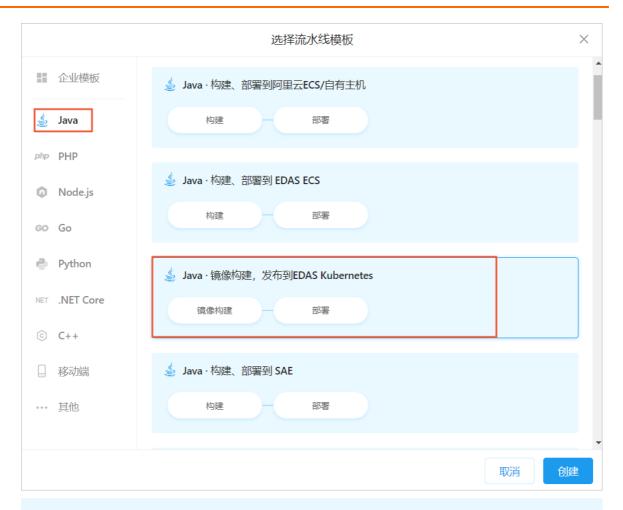
如果您是第一次使用云效Flow,则需要在云效上创建您的企业。

- 1. 登录云效Flow。
- 2. 设置企业、组织或团队名称并选择使用规模,单击完成创建。



步骤二:在云效创建流水线

- 1. 登录云效Flow。
- 2. 在左侧导航栏,单击我的流水线,然后单击右上角的新建流水线。
- 3. 在选择流水线模板对话框,在左侧单击Java,在右侧选择Java·镜像构建,发布到EDAS Kubernetes,然后单击创建。



☐ 注意 在EDAS控制台创建应用时使用的部署方式(JAR、WAR和镜像),必须与在云效流水线设置保持一致。例如:在EDAS上使用镜像方式部署应用,那么在云效设置流水线时,构建环节必须是镜像相关配置。

流水线创建完成后,自动跳转到流水线配置页面。

步骤三:配置流水线

- 1. 添加代码源。
 - i. 在流水线的源区域,单击添加代码源。
 - ii. 在添加流水线源面板,配置流水线源相关信息,然后单击添加。



云效中提供的流水线源有3种:代码源、Jenkins和Flow流水线。本文以代码源作为示例。

参数 描述

参数	描述
选择代码源	选择您业务代码所在的代码源,本示例以 阿里云 Code为例。
	② 说明 不同代码源需要配置的参数不同, 请根据实际界面进行配置。
代码仓库	选择您的代码仓库。
默认分支	选择代码分支。
开启代码源触发	开启代码源触发功能,将Webhook添加到代码仓库,云效流水线Flow会监听代码事件,满足条件时触发流水线运行。 阿里云Code包含以下两种触发事件: 《代码提交: 从本地Push分支到远程分支(或在代码管理平台的Web端进行文件修改),云效流水线Flow会将Push的目标分支与触发分支进行匹配,如果匹配成功,则会触发流水线,并使用Push的目标分支作为流水线运行时使用的分支。 TAG创建: 从本地Push Tag到服务端Tag(或在代码管理平台的Web端进行tag创建),云效流水线Flow会将Push的目标Tag名称与触发分支进行匹配,如果匹配成功,则会触发流水线,并使用目标Tag作为流水线运行时使用的代码版本。如果使用的是其他代码仓库,可能会涉及到其他触发事件。更多触发事件的信息,请参见触发事件与触发分支。

2. Docker镜像构建。

i. 在流水线的构建区域,单击Docker镜像构建,然后展开任务步骤镜像构建并推送阿里云。



ii. 配置任务步骤参数,配置完成后在页面右上角单击**保存并运行**。

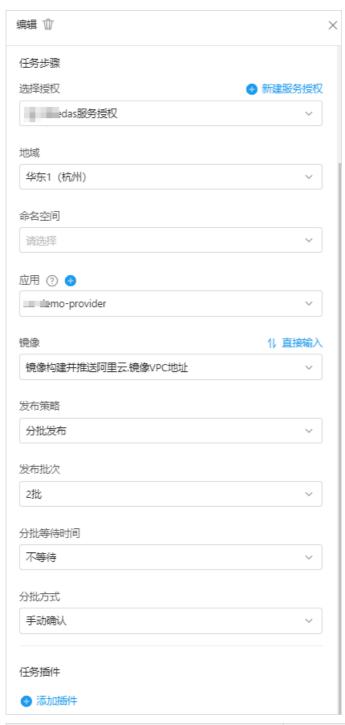


参数	描述
步骤名称	输入您自定义的步骤名称,不修改则显示为默认名称。

参数	描述
选择授权	选择任务的服务授权,使云效能使用容器镜像服务。
	⑦ 说明 如果您从未授权过,请先单击新 建服务授权,在云资源授权页面单击同意授 权。
地域	在下拉列表中选择地域。
仓库	在下拉列表中选择仓库。
标签	标签为Docker镜像Tag。支持固定参数例如1.0,或者动态参数例如\${DATETIME}。
Dockerfile路径	Dockerfile路径为Dockerfile文件相对于代码库根目录所在路径,如 <i>META/config/Dockerfile</i> 或 <i>Dockerfile</i> 。
ContextPath	ContextPath为docker build命令执行上下文路径。 填写相对于代码根目录的路径,如 <i>target</i> ,如果不 填则为Dockerfile文件所在目录。
不使用缓存	如果选中,Docker Build将使用no-cache=true参数进行镜像构建。
构建参数	构建参数为运行时会以 -build-arg 的形式传递 到 build 命令中的一组参数。单击添加参数可 以新增并配置构建参数。

3. 部署到EDAS K8s环境。

- i. 在部署区域,单击部署到EDAS K8s。
- ii. 在编辑面板配置任务信息,然后单击右上角保存并运行。



参数名	说明
任务名称	自定义的任务名称,不修改则显示为默认名称。
构建集群	可为任务选择不同的构建集群。

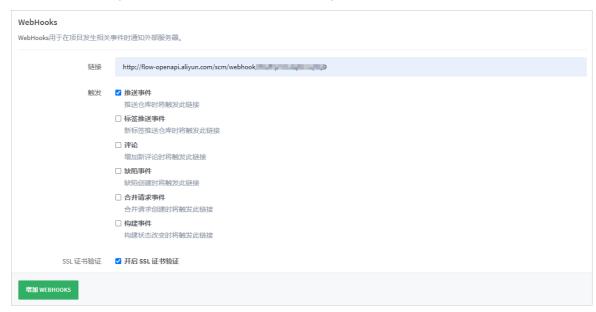
参数名	说明
选择授权	选择任务的服务授权,使云效能在EDAS上部署应用。
	② 说明 如果您从未授权过,请先单击新 建服务授权,根据跳转完成阿里云RAM授权后 再进行相应配置。
地域	选择EDAS侧所创建应用所在的地域。
应用	选择EDAS侧所创建的应用。
镜像	选择Docker镜像构建所创建的产出物名称。
发布策略	支持 分批发布和灰度发布 ,本示例以分批发布为例。
发布批次	选择 单批次发布 ,表示不分批发布;选择批次,如 2批,表示按照指定的批次来分批发布。 本示例以 2批 为例。
分批等待时间	相邻发布批次之间的等待时间。
分批方式	可选择 手动确认 或 自动确认 。例如,如果您需在完成第一批发布时先观察发布结果再决定后续操作,则可选择 手动确认 。
(可选) 任务插件	您可以根据需要配置任务插件来发送流水线通知。 「钉钉机器人通知插件 webhook地址:配置钉钉机器人的webhook地址,钉钉机器人的创建步骤,请参见钉钉机器人配置。 运行时机:您可以选择任务插件发送通知的时机。 「邮件通知 「邮件地址:配置需要发送通知的邮件地址。」运行时机:您可以选择任务插件发送通知的时机。
	■ webhook地址:配置需要发送通知的 webhook地址,地址必须公网可访问。 ■ 运行时机:您可以选择任务插件发送通知的时机。

步骤四: 代码库配置Webhook

在代码库添加云效流水线Flow的代码源触发功能生成的Webhook地址,云效流水线Flow会监听代码事件, 当满足触发事件时会触发流水线运行。

本文使用阿里云Code作为代码库,在代码库配置Webhook的步骤仅作参考,其他代码库请以实际为准。

- 1. 登录阿里云Code, 单击目标项目。
- 2. 在左侧导航栏单击设置。
- 3. 在左侧导航栏单击WebHooks。
- 4. 在WebHooks页面,设置Webhook地址并选中推送事件,然后单击增加WEBHOOKS。



结果验证

- 1. 开发者上传应用新版本代码到代码库。
- 2. 云效流水线Flow监听代码事件, 当满足触发事件时会触发流水线运行。

在本示例流水线中配置了分批发布。故在构建任务运行成功后,您可在部署任务的区域执行以下操作:

- 终止变更: 可终止本次流水线的运行。
- 恢复变更: 可继续下一批的发布。
- **当前发布单详情地址**:可跳转到EDAS控制台查看变更详情。
- 日志: 可查看流水线的运行日志。



3. 第一批发布完成后,您可以先验证新版本应用。验证无误后,单击**恢复变更**,继续下一批次的发布。 您也可以单击**当前发布单详情地址**,在EDAS控制台的**变更记录**页面单击**开始下一批**,继续下一批次 的发布。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

5.2.2. 使用云效2020实现其他类型应用的持续集成和部署(K8s)

EDAS K8s环境支持部署Golang、Node.js、Python和PHP等编程语言开发的应用,包含JAR包、WAR包和镜像3种部署方式;在云效2020侧持续集成和部署应用到EDAS K8s环境时,只能以镜像方式构建和部署。本文介绍如何使用云效2020实现此类应用的持续集成和部署。

背景信息

使用云效2020实现Golang、Node.js、Python、PHP等应用的持续集成和部署,与实现Java应用的持续集成和部署的操作流程类似。本文仅描述关键步骤和配置,如果您了解云效2020,则可以依据本文内容完成此类应用的持续集成和部署。具体操作,请参见使用云效2020进行持续集成和部署(K8s)。

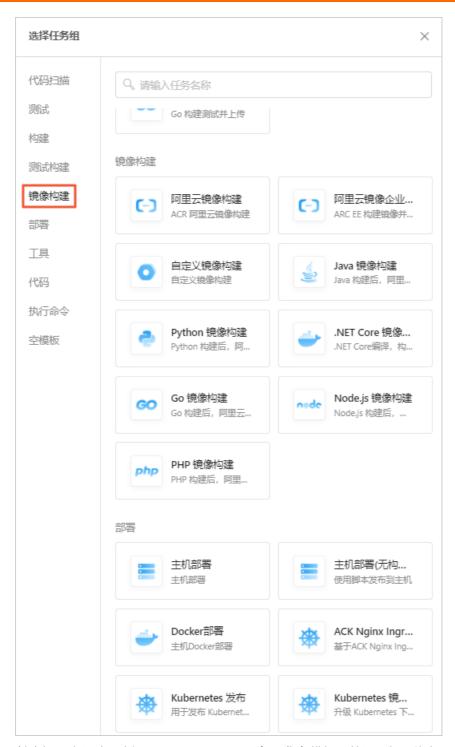
关键步骤和配置

- 在EDAS控制台创建并部署应用。
- 在新建流水线时选择相应的编程语言和流水线模板。



□ 注意 在云效上仅可通过镜像方式实现Golang、Node.js、Python、PHP等应用的持续继承和部署到EDAS K8s环境。

● 在基础流水线中删除原有的**构建**和**部署**阶段。创建构建阶段时,选择相应编程语言的镜像模板。



● 创建部署阶段时,选择EDAS Kubernetes应用发布模板,并设置部署信息。



5.3. Jenkins

5.3.1. 使用Jenkins创建持续集成

使用Jenkins可以构建EDAS应用的持续集成方案。该方案涉及下面的计算机语言或开发工具,阅读本文需要对下述的语言或工具有一定的理解。

前提条件

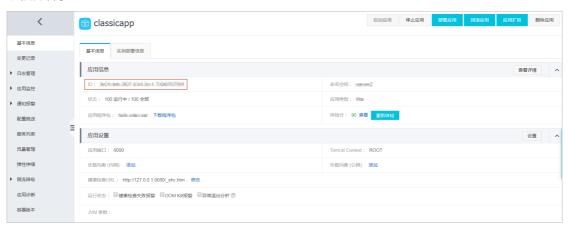
在开始持续集成之前,需要完成下述的准备工作。

1. 获取阿里云的AccessKey ID和AccessKey Secret。

- i. 使用已经开通了EDAS服务的主账号登录阿里云官网。
- ii. 进入Access Key控制台,创建AccessKey ID和AccessKey Secret。



- 2. 在使用Jenkins自动部署应用之前,需要先在EDAS控制台中创建并部署应用。
 - i. 登录EDAS控制台。
 - ii. 参考<mark>应用创建和部署概述(ECS)和创建和部署应用概述(K8s)</mark>,部署应用。 如果已经创建了应用,请忽略此步。
 - iii. 在左侧导航栏中单击**应用列表**。找到您在上一步中创建的应用并单击进入详情页面,获取应用ID的字段内容。



- 3. 使用Git Lab托管您的代码。您可以自行搭建Git lab或者使用阿里云Code。本文使用通过自行搭建的Git Lab做演示,关于Git lab的更多信息,请参见Git Lab。
- 4. 了解并使用Jenkins。关于Jenkins的更多详细信息,请参见Jenkins官网。

背景信息

② 说明 ECS集群和容器服务Kubernetes集群都支持使用Jenkins创建持续集成。

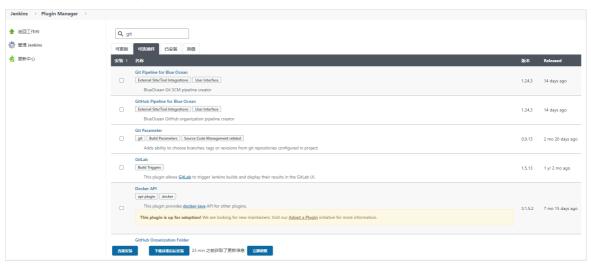
工具	说明
Maven	Maven是一个项目管理和构建的自动化工具。
Jenkins	Jenkins 是一个可扩展的持续集成引擎。
GitLab	GitLab 是一个利用Ruby on Rails开发的开源应用程序,实现一个自托管的Git项目仓库,可通过Web界面进行访问公开的或者私人项目。 它拥有与Git Hub类似的功能,能够浏览源代码,管理缺陷和注释。

配置项目

修改项目配置,添加toolkit-maven-plugin及部署信息,具体操作,请参见使用toolkit-maven-plugin插件升级应用(ECS集群)。您在修改完项目配置后,建议在本地使用Maven构建验证配置是否正确。

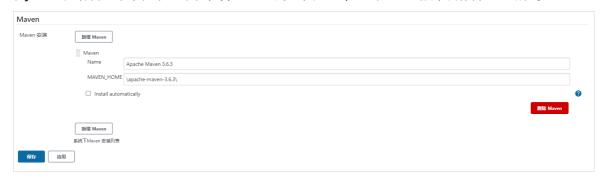
安装和配置Jenkins

- 1. 进入Jenkins官网下载安装Jenkins。
- 2. 在Jenkins控制台的菜单栏中选择系统管理 > 插件管理, 安装Git和GitLab插件。
 - 。 安装GIT Client Plugin和GIT Plugin插件可以帮助Jenkins拉取Git仓库中的代码。
 - 安装Git lab Hook Plugin插件可以帮助Jenkins在收到Git lab发来的Hook后触发一次构建。



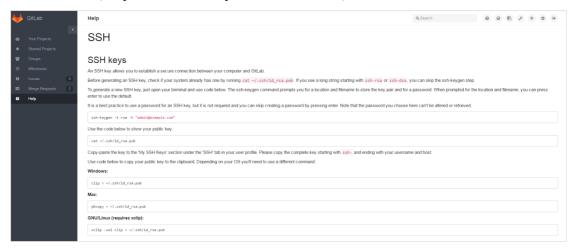
3. 安装Maven构建工具,请参见Maven官网。

在Jenkins控制台的菜单栏中选择系统管理 > 全局工具配置,选择Maven版本名称并配置路径。

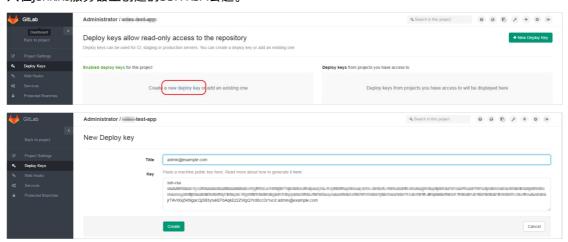


4. 在Jenkins服务器上生成SSH RSA密钥对,并将公匙导入Git Lab,实现Jenkins拉取Git Lab代码时自动认证。

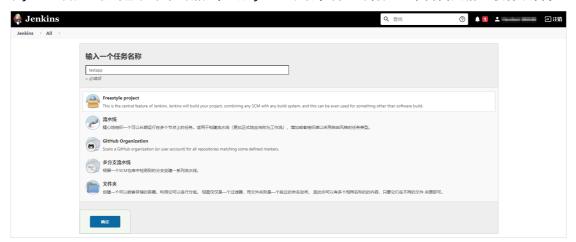
i. 参考Git Lab文档,在Jenkins服务器运行Jenkins软件的用户下,生成SSH RSA密钥对。



ii. 进入GitLab首页,在菜单栏选择Settings > Deploy Keys,并单击new deploy key添加key,导入在Jenkins服务器上创建的SSH RSA公匙。



- 5. 创建Jenkins任务。
 - i. 在Jenkins首页左侧导航栏中单击新建,创建Jenkins任务,并选择构建一个自由风格的软件项目。



- ii. 在**源码管理**页面中选择Git , 并设置相关参数。
 - Repository URL: 您的项目的Git协议地址。
 - Credentials:安全凭证,选择无即可(前提是运行Jenkins软件的用户的SSH RSA公匙已添加到该Git项目所在的GitLab中,否则这里会报错)。



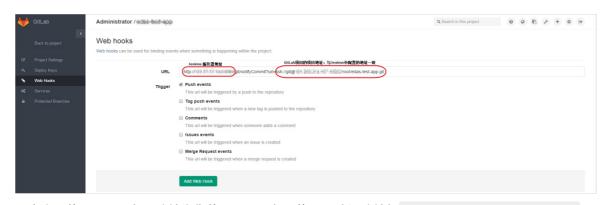
- iii. 单击构建触发器页签, 勾选轮询SCM。
- iv. 单击**构建环境**页签,勾选Add timestamps to the Console Output(为控制台输出的信息添加时间戳)。
- v. 单击构建页签, 然后单击增加构建步骤。
- vi. 在**调用顶层Maven目标**区域设置**Maven版本**和目标。如果您想部署多模块工程,请参见<mark>创建多模块工程的Jenkins任务。</mark>
 - Maven Version: 单击该选项后面的下拉框,选择在全局工具配置里配置的Maven版本名称。
 - Goals:填入clean package toolkit:deploy(如有其它参数,请根据实际情况填入)



配置Gitlab的Web Hook,实现自动构建

- 1. 右键单击GitLab工程,然后选择Setting > Web Hooks。
- 2. 在**Web Hooks**页面的在**URL文本框**中输入 http://jenkins**服务器地址:**jenkins**服务器监听端口**/git/not ifyCommit?url=本项目的git协议地址 。

例如: http://123.57.xx.xxx:8080/git/notifyCommit?url=git@code.aliyun.com:tdy218/hello-edas.git



图中表示的Jenkins服务器地址为您的Jenkins服务器的Web访问地址如 http://123.57.xx.xxx:8080 。

3. 配置完成后,单击Test Hook,进行测试。



配置正确后,提交变更到GitLab

如果上述步骤配置正确,这次提交会触发一次Git Lab Hook。Jenkins在接受到这个Hook后会构建您的Maven项目,并在构建结束时调用EDAS POP API脚本触发部署。

提交部署成功输出的日志信息(Build Number > 控制台输出)。

如果部署失败,可以登录EDAS控制台,在左侧导航栏中单击**应用列表**,在应用列表页面单击具体应用名称,进入应用详情页面。在左侧导航栏单击**变更记录**来查看此次部署任务的执行过程。

创建多模块工程的Jenkins任务

创建多模块工程的Jenkins任务和安装和配置Jenkins第5步基本相同,只需要调整下**调用顶层Maven目标**。如果工程为多模块工程,想在Jenkins中部署子模块的话,那么需要在父模块中调用 mvn clean install 命令,然后在子模块中调用 mvn clean package toolkit:deploy 命令。以Demo工程为例,工程结构如下:

```
sh-3.2# tree -L 1 carshop
carshop
— detail
— itemcenter
— itemcenter-api
— pom.xml
```

其中,detail、itemcenter、itemcenter-api为子模块,现在想部署itemcenter模块的话,那么需要在父工程中设置一个clean install构建目标,然后在itemcenter模块中设置 clean package toolkit:deploy 构建目标。



5.3.2. 在Jenkins中使用EDAS插件构建应用部署的任务

在Jenkins环境中集成了edas-jenkins-plugin插件,您可以使用该插件在Jenkins内快速完成应用持续部署到EDAS。本文介绍了在Jenkins控制台的白屏化部署过程,使用流水线的部署步骤以及使用Job-dsl的任务模板来快速生成部署任务的简介。

前提条件

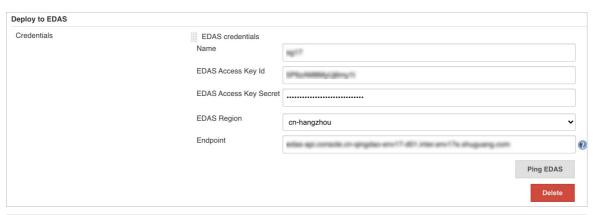
已安装Job-dsl插件。

安装edas-jenkins-plugin

- 1. 下载edas-jenkins-plugin安装包。
- 2. 登录Jenkins首页,单击Manage Jenkins,在Manage Jenkins页面单击Manage Plugins。
- 3. 在Plugin Manager页面选择Advanced页签,在Upload Plugin区域单击**选择文件**,选择下载的插件,然后单击Upload。
- 4. 在Installing Plugins/Upgrades页面等待插件安装完成。

配置账号信息

- 1. 登录Jenkins首页,单击Manage Jenkins,在Manage Jenkins页面单击Configure System。
- 2. 在Deploy to EDAS区域内配置账号信息,单击Add,然后单击Advanced输入地域信息。



参数	描述
EDAS credentials name	为AccessKey设置名称。
EDAS Access Key ID	输入账号的AccessKey信息,获取方法,请参见 <mark>获取</mark> AccessKey。
EDAS Access Key Secret	输入账号的AccessKey密钥,获取方法,请参见 <mark>获取</mark> AccessKey。
EDAS Region	请输入您想部署的应用的地域ID,例如cn-hangzhou。
Endpoint(专有云)	在专有云中使用Jenkins前,需在该配置项中输入EDAS的EndPoint,Endpoint请联系EDAS技术支持获取。
	⑦ 说明 在公共云中使用Jenkins时,无需配置 该参数。

- 3. 配置完成后在Deploy to EDAS区域单击Ping EDAS,当页面显示success即说明账号信息配置完成。
 - ? 说明 您可以单击Add设置多个EDAS账号。

在Jenkins中创建Freestyle project任务来部署应用

- 1. 登录Jenkins首页, 单击New Item。
- 2. 在Enter an item name区域输入一个项目名称,选择Freestyle project然后单击OK。
- 3. 在Post-build Actions区域单击Add post-build actions, 然后在下拉列表中单击Deploy to EDAS。
- 4. 在Deploy to EDAS区域单击Add,在下拉列表中可以选择集群来部署。

EDAS ECS Application

Create EDAS ECS Application

EDAS K8s Application

Create EDAS K8s Application

参数	描述
EDAS ECS Application	已经部署在ECS集群中的应用在Jenkins插件中升级部署,详细配置参数,请参见在ECS集群中部署应用。
Create EDAS ECS Application	使用Jenkins插件在ECS集群中创建应用并部署,详细配置参数,请参见 <mark>在ECS集群中创建并部署应用</mark> 。
EDAS K8s Application	已经部署在容器服务K8s集群中的应用在Jenkins插件中升级部署,详细配置参数,请参见在容器服务K8s集群中部署应用。
Create EDAS K8s Application	使用Jenkins插件在容器服务K8s集群中创建应用并部署,详细配置参数,请参见在容器服务K8s集群中创建并部署应用。

- 5. 设置完成后在页面左下角单击Save或Apply。
- 6. 在工程任务页面的左侧导航栏单击Build Now来运行任务。



您将会在Build History区域查看构建记录,单击构建记录可以进入构建记录的详情页面查看构建详情。

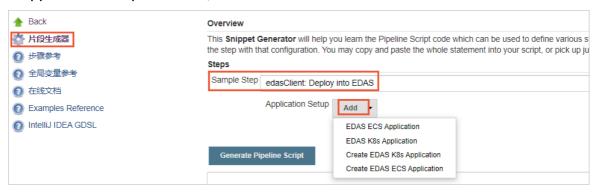


使用流水线插件部署应用到EDAS

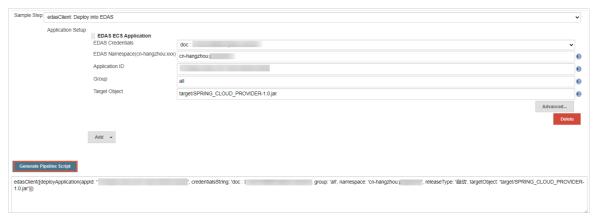
edas-jenkins-plugin已经支持流水线预发的片段生成器,在Jenkins中构建流水线部署应用到EDAS。

- 1. 登录Jenkins首页, 单击New Item。
- 2. 在Enter an item name区域输入项目名称,选择流水线然后单击OK。
- 3. 在Pipeline区域单击Pipeline Syntax。

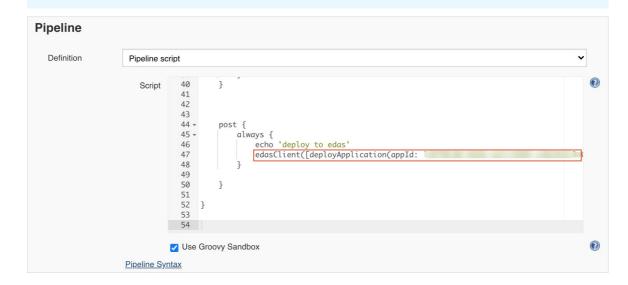
- ② 说明 如果您已创建了流水线,可在项目列表单击流水线名称进入流水线管理页面,然后在左侧导航栏单击流水线语法进入片段生成器页面。
- 4. 在**片段生成器**页面的Sample Step下拉框选择edasClient: Deploy into EDAS,然后单击Application Setup右侧的Add,然后在下拉列表中选择集群和应用部署类型。



- 5. 配置部署参数,具体配置信息,请参见在Jenkins中使用edas-jenkins-plugin部署应用的参数配置。
- 6. 完成参数配置后单击Generate Pipeline Script , 拷贝生成的流水线脚本。



- 7. 在左侧导航栏单击Back返回流水线管理页面,然后在左侧导航栏单击Configure,然后在Pipeline区域按照下图示例输入脚本内容。
 - ② 说明 下图橙色框内的内容需替换为上面步骤生成的脚本。

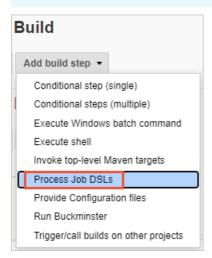


- 8. 设置完成后在页面左下角单击Save或Apply。
- 9. 在工程任务页面的左侧导航栏单击Build Now来运行任务。 您将会在Build History区域查看构建记录,单击构建记录可以进入构建记录的详情页面查看构建详 情。

使用Job-dsl插件来构建部署任务

在Jenkins中您可以使用Job-dsl插件来构建一个任务模板,使您可以快速根据模板来构建部署任务。

- 1. 登录Jenkins首页, 单击New Item。
- 2. 在Enter an item name区域输入项目名称,选择Freestyle project然后单击OK。
- 3. 在Build区域单击Add build step, 然后在下拉列表中选择Process Job DSLs。
 - ② 说明 如果您已创建了流水线,可在项目列表单击流水线名称进入流水线管理页面,然后在左侧导航栏单击流水线语法进入片段生成器页面。



4. 在Process Job DSLs区域单击Use the provided DSL script,然后在脚本输入框内输入您的配置脚本。

```
job('edas') { //目标任务的名称
  publishers { //EDAS插件的类型
   edasClient { //EDAS插件调用名称
       deployApplication('', '', '', '') //调用的EDAS方法
   }
}
```

支持的方法示例如下:

```
//在ECS集群中部署应用

void deployApplication(
    String namespace,
    String credentialsString,
    String group,
    String appId,
    String targetObject);

void deployApplication(
    String namespace,
    String credentialsString,
    String group.
```

```
------ 9------
       String appId,
       String targetObject,
       String versionLabelFormat,
       String versionDescriptionFormat,
       String batch,
       String batchWaitTime,
       String releaseType);
//在ECS集群中创建并部署应用
void insertApplication(
       String namespace,
       String credentialsString,
       String applicationName,
       String packageType,
       String clusterId,
       String targetObject,
       String ecuInfo);
void insertApplication(
       String namespace,
       String credentialsString,
       String applicationName,
       String packageType,
       String clusterId,
       String targetObject,
       String ecuInfo,
       String versionLabelFormat,
       String versionDescriptionFormat,
       String buildPackId,
       String healthCheckUrl);
//在容器服务K8s集群中部署应用
void deployK8sApplication(
       String namespace,
       String credentialsString,
       String appId,
       String targetObject,
       Boolean image);
void deployK8sApplication(
       String namespace,
       String credentialsString,
       String appId,
       String targetObject,
       Boolean image,
       String edasContainerVersion,
       String webContainer,
       String jdk,
       String versionLabelFormat,
       String envs,
       String startupCommand,
       String args,
       String cpuLimit,
       String memoryLimit,
       String cpuRequest,
       String memoryRequest,
       String replicas,
       String postStart,
```

```
String preStop,
        String readiness,
        String liveness,
        String updateStrategy);
//在容器服务K8s集群中创建并部署应用
void insertK8sApplication(
       String namespace,
       String credentialsString,
       String applicationName,
       String packageType,
       String clusterId,
       String targetObject,
       String k8sNamespace,
       String jdk);
void insertK8sApplication(
       String namespace,
       String credentialsString,
       String applicationName,
       String packageType,
       String clusterId,
       String targetObject,
       String k8sNamespace,
       String jdk,
       String descFormat,
       String edasContainerVersion,
       String webContainer,
       String versionLabelFormat,
       String envs,
       String startupCommand,
       String args,
       String cpuLimit,
       String memoryLimit,
       String cpuRequest,
       String memoryRequest,
       String replicas,
       String postStart,
       String preStop,
       String readiness,
        String liveness);
```

- 5. 设置完成后在页面左下角单击Save或Apply。
- 6. 在工程任务页面的左侧导航栏单击Build Now来运行任务。 您将会在Build History区域查看构建记录,单击构建记录可以进入构建记录的详情页面查看构建详 情。

更多信息

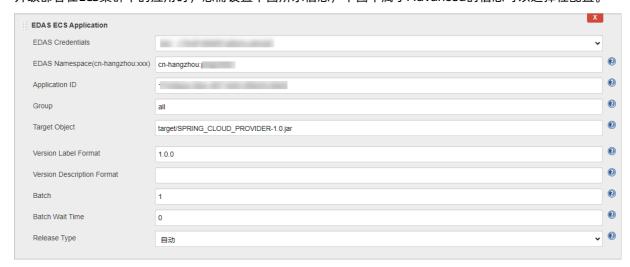
在Jenkins中使用edas-jenkins-plugin部署应用的参数配置

5.3.3. 在Jenkins中使用edas-jenkins-plugin部署应用的参数配置

在Jenkins中使用edas-jenkins-plugin部署应用到EDAS时,您既可以在ECS集群中创建或部署应用,也可以在容器服务K8s集群中创建或部署应用,本文介绍了4种选项的参数配置说明。

在ECS集群中部署应用

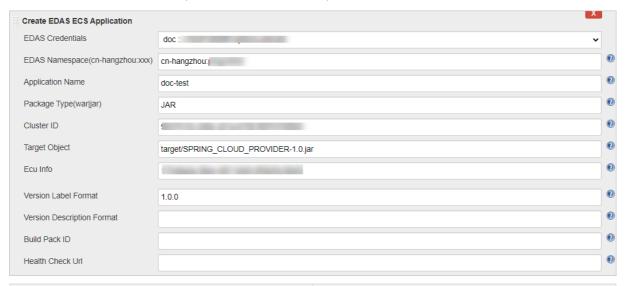
升级部署在ECS集群中的应用时,您需设置下图所示信息,下图中属于Advanced的信息可以选择性配置。



参数	描述
EDAS Credentials	选择AccessKey的名称。
EDAS Namespace	输入部署应用所在的微服务空间,格式为 地域ID:微服务空间ID ,例如cn-hangzhou:doc。
Application ID	输入部署应用的应用ID。
Group	输入应用分组,all代表所有分组。
Target Object	要部署的目标文件的地址链接,您再次部署的包类型需跟应用第一次部署的包类型保持一致。
Version Label Format	版本号。
Version Description Format	格式化的版本描述信息。
Batch	部署分批次数。
Batch Wait Time	批次间隔时间。
Release Type	分批发布类型,您可以选择手动或者自动部署下一批。

在ECS集群中创建并部署应用

在ECS集群中创建应用并部署时,您需设置下图所示信息,下图中属于Advanced的信息可以选择性配置。



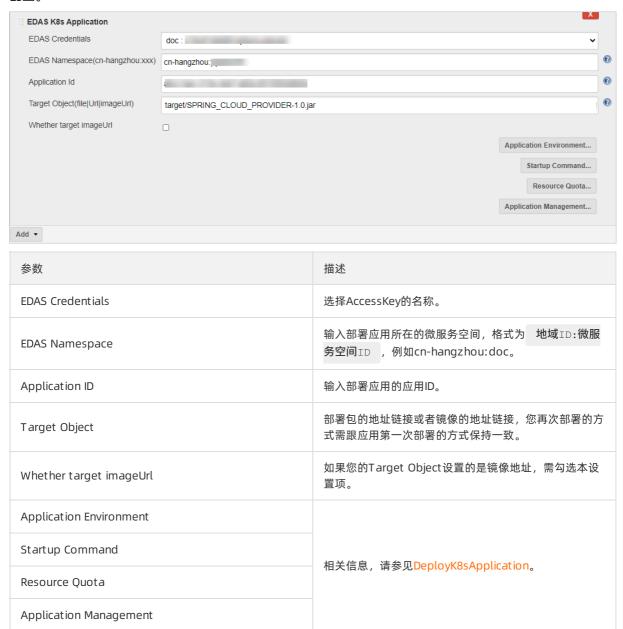
参数	描述
EDAS Credentials	选择AccessKey的名称。
EDAS Namespace	输入部署应用所在的微服务空间,格式为 地域ID:微服务空间ID ,例如cn-hangzhou:doc。
Application Name	输入部署应用的应用名称。
Package Type	输入部署包类型:WAR或JAR。
Cluster ID	输入应用所在的集群ID。
Target Object	要部署的目标文件的地址链接,您再次部署的包类型需跟应用第一次部署的包类型保持一致。
Ecu Info	输入需要使用的ECS实例的ECU ID,相关信息,请参见ListScaleOutEcu。
Version Label Format	版本号。
Version Description Format	格式化的版本描述信息。
Build Pack ID	容器版本ID。
Health Check Url	健康检查URL。

在容器服务K8s集群中部署应用

? 说明

通过edas-jenkins-plugin更新Spring Boot应用至K8s集群时,您可以在EDAS控制台左侧导航栏的**应用列**表栏,选择K8s集群中的应用进入**应用总览**详情页面,单击**部署 > 分批发布 > 开始部署 > 复制部署**参数生成 UpdateStrategy Json 文件。

升级部署在容器服务K8s集群中的应用时,您需设置下图所示信息。下图中的Application Environment、Startup Command、Resource Quota和Application Management信息可以选择性配置。



在容器服务K8s集群中创建并部署应用

Create EDAS K8s Application	X	
EDAS Credentials	profit programme and the second	
EDAS Namespace(cn-hangzhou:xxxx)	cn-shenzhen:	?
Cluster ID		•
K8s Namespace	default	•
Application Name	\${BUILD_TAG}	?
Application Description	Prop. In as found	
Target Object(filelUrllimageUrl)	target/SPRING_CLOUD_PROVIDER-1.0.jar	?
Package Type(FatJarlWARIImage)	FatJar	
EDAS Container Version		•
Web Container		?
JDK Version	Open JDK 8	?
Version Label Format		•
Environments Json		•
	Startup Command	
	Resource Quota	
	Application Management	
۵. ساد د ما د	1#\L	

参数	描述
EDAS Credentials	选择AccessKey的名称。
EDAS Namespace	输入部署应用所在的微服务空间,格式为 地域ID:微服务空间ID ,例如cn-hangzhou:doc。
Cluster ID	输入应用所在的集群ID。
K8s Namespace	K8s Namespace通过将系统内部的对象分配到不同的 Namespace中,形成逻辑上分组的不同项目、小组或用 户组,便于不同的分组在共享使用整个集群的资源的同时 还能被分别管理。此处设置为default即可。
Application Name	输入部署应用的应用名称。
Application Description	输入应用描述。
Target Object	请输入Fatjar包、WAR包或镜像的地址链接。

参数	描述	
Package Type	输入部署包的类型: FatJar、WAR或Image。	
EDAS Container Version	部署包依赖的EDAS Container版本。适用于通过WAR包部署的HSF应用,镜像不支持此参数。	
Web Container	部署包依赖的Tomcat版本。适用于通过WAR包部署的 Spring Cloud和Dubbo应用,镜像不支持此参数。	
JDK Version	输入JDK版本,例如 OPEN JDK 8。	
Version Label Format	版本号。	
Environments Json	部署环境变量,格式必须符合JSON对象数组。	
Startup Command		
Resource Quota	相关信息,请参见DeployK8sApplication。	
Application Management		

6.应用监控

6.1. 应用监控概述

EDAS集成了应用实时监控服务ARMS,为部署在EDAS中的应用提供关键指标的监控能力和报警能力。

应用监控

EDAS集成了应用实时监控服务ARMS,为部署在EDAS中的应用提供关键指标的监控能力,帮助您快速定位出错接口和慢接口、重现调用参数,从而大幅提升线上问题诊断的效率。

监控粒度	功能描述	相关文档
应用总览	应用详情页,您可以查看K8s集群信息,包括地域、微服务空间和运行状态以及应用诊断报告等;也可以查看应用的健康状况关键指标,包括总请求量、平均响应时间等总体指标,应用所提供的服务、所依赖的服务的相关指标,以及CPU使用量、内存使用量等系统信息。	查看应用总览
Prometheus监控	Prometheus监控提供的预置监控仪表板包括Pod基本信息、CPU信息、内存信息和网络信息,您可以通过这些仪表板查看丰富的Prometheus监控指标,并按需更改仪表板数据的时间区间、刷新频率等属性。	查看Prometheus监控指标
实例监控	EDAS提供了多种应用监控指标: IVM监控功能用于监控重要的JVM指标,包括堆内存指标、非堆内存指标、非堆内存指标、内存映射缓冲区指标、GC(垃圾收集)累计详情和JVM线程数等。 JVM监控可以直观展示指定时间段内的多项内存指标,然而虽然情况,但无法显示具体信息,因此不能帮助您排查问题的原,通过详细的日志查看内存占用的详细信息。 Internal control co	 JVM监控 内存快照 主机监控
服务与接口监控	用于监控应用下的接口调用详情,包括SQL分析、NoSQL调用分析、异常分析、错误分析、链路上下游和接口快照。	服务和接口监控

监控粒度	功能描述	相关文档
应用诊断-实时诊断	适用于在短时间内密切监控应用性能 和定位问题原因的场景。	实时诊断
应用诊断-异常分析	应用异常分析功能,包括异常数量统计、每类异常次数统计以及异常发生的端口等。	异常分析
应用诊断-线程分析	提供线程粒度的CPU耗时和每类线程数量的统计,并且每5分钟记录一次线程的方法栈并聚合,可真实还原代码执行过程,帮助您快速定位线程问题。	线程分析

报警

您可以针对特定监控对象制定报警规则。当规则被触发时,系统会以您指定的报警方式向报警联系人分组发送报警信息,以通知联系人采取必要的问题解决措施。

相关文档:

- 创建联系人
- 创建联系人组
- 应用监控告警规则(新版)
- 在钉钉群中处理告警
- 创建应用监控告警并发送告警通知

6.2. 监控

6.2.1. 查看应用总览

在应用总览页面上,您可以查看K8s集群信息,包括地域、微服务空间和运行状态以及应用诊断报告等;也可以查看应用的健康状况关键指标,包括总请求量、平均响应时间等总体指标,应用所提供的服务、所依赖的服务的相关指标,以及CPU使用量、内存使用量等系统信息。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。

查看应用详情

在**应用总览**页面查看K8s集群概览信息,包括集群名称、应用诊断报告、地域、集群类型、运行状态、创建时间、规格和负载均衡等。



- 应用诊断报告,请参见查看诊断报告。
- 应用概览分析,请参见概览分析。
- 应用拓扑,请参见应用拓扑。
- 更改EDAS规格,请参见手动扩缩容。
- 配置负载均衡,请参见添加负载均衡SLB。
- 配置服务Service, 请参见添加服务Service。
- YAML配置,请参见YAML配置。
- 应用部署和回滚,请参见管理应用生命周期。

查看诊断报告

② 说明 EDAS标准版不能查看诊断报告,EDAS专业版和铂金版支持查看诊断报告。

概览分析

概览分析页签上展示以下关键指标:

● 选定时间内的总请求量、平均响应时间、错误数、FullGC次数、慢SQL次数、异常次数和慢调用次数,以

 及这些指标和上周的同比、上一天的同比升降幅度。

总请求量	平均响应时间	错误数	Full GC	慢SQL	异常	慢调用
18.6K	2.04ms	12375次	0 次	0次	23.9K ↑	0 ↑
周同比 1 856300% 日同比 6 1776.7%	周同比 ◆ 203.8% 日同比 ◆ 203.8%	周同比 ↑ 1237500% 日同比 ↑ 1237500%	周同比 = 0% 日同比 = 0%	周同比 = 0% 日同比 = 0%	周同比 ◆ 2389800% 日同比 ◆ 2389800%	周同比 - 0% 日同比 - 0%

● 应用提供服务: 应用提供服务的请求量和平均响应时间的时序曲线。在EDAS企业版和铂金版中,支持单击时序曲线中出现的突增点进行深度分析。



● 应用依赖服务:应用依赖服务的请求量、平均响应和应用实例数的时序曲线,以及HTTP-状态码统计。



● 系统信息: CPU、MEM和负载的时序曲线。



● 慢调用:慢调用的时序曲线和调用详情。



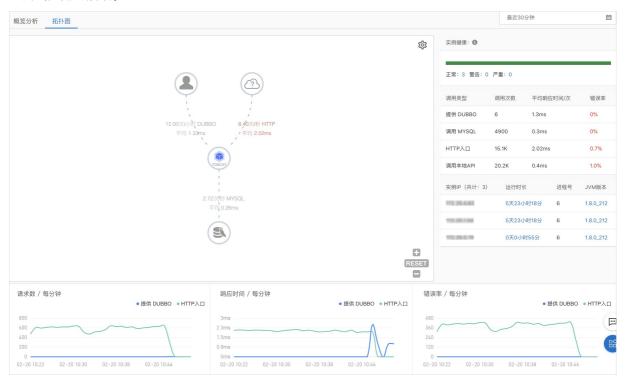
● 统计分析:接口慢调用分析和异常类型分析。



应用拓扑

② 说明 EDAS标准版不支持拓扑图, EDAS专业版和铂金版支持拓扑图。

在**拓扑图**页签上,您可以通过拓扑图更加直观地看到应用的上下游组件以及与它们的调用关系,从而更快速 地找出应用的瓶颈。



6.2.2. 应用故障自动诊断

在EDAS中进入应用总览页面时,会按照您自定义的时间段对该应用的运行情况做一个自动诊断。如果发现问题,则会在应用总览页面上方闪现一个红色盾牌图标♥,单击该图标即可弹出故障诊断报告,故障报告上的故障定界和根因分析可以帮助您更好地发现和解决问题。

常见的故障自动诊断场景

RT突增

- 下游业务导致的本应用的RT突增,您可以联系下游业务的负责人进行排查。
- 应用变更导致的RT突增,您可以查看此次变更的具体变更进行排查。
- 应用的某个服务导致RT突增,可以排查以下情况:
 - 服务是否在此时有发生异常。
 - 调用该服务的下游服务是否有出现RT较高的情况。
 - 该服务调用的某服务的RT是否较高。
- 因为单机的问题导致RT突增:
 - 线程池满,此时故障诊断报告中会给出线程数的时序图。
 - 单机FullGC。
 - 单机磁盘读写异常。
 - 单机OOM。

错误请求或大请求的占比过高

- 某应用服务的错误请求次数出现突增,导致错误请求占比过高。
- 某段时间出现的大请求和大响应占比过高,导致序列化和反序列化耗时高。

宿主机负载过高

宿主机的负载过高导致容器提供服务的能力降低。

网络问题

机房发生网络故障导致应用运行异常。

查看故障自动诊断报告

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面上单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用总览页面的概览分析页签右上角选择您的自定义诊断时间。

□ **注意** 当诊断到自定义时间内应用运行异常,将会在页面上方的应用名称右侧出现红色盾牌图标。当应用没有被诊断出故障时,不代表应用不存在潜在的问题。

- i. 在**应用总览**页面上方的应用名称右侧单击**♀**。
- ii. 在弹出的诊断报告中查看故障现象和原因分析。

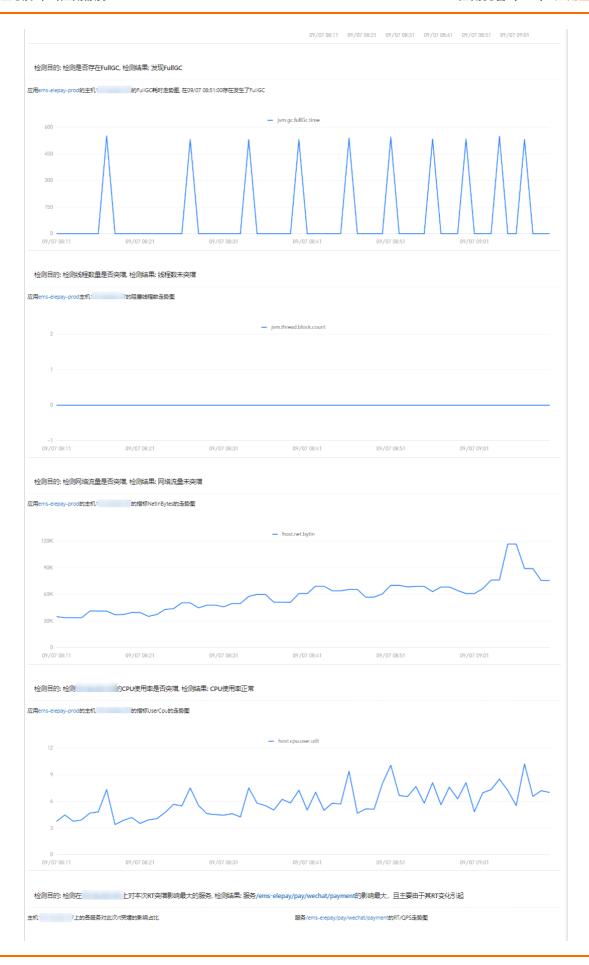
诊断报告组成

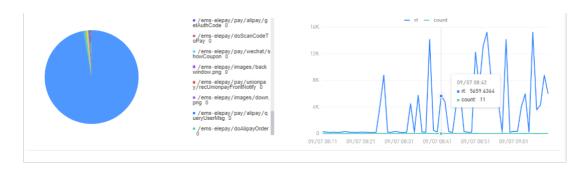
诊断报告主要由4个部分组成:诊断详情、故障定界、根因分析和数据支持。

- 诊断详情:这一部分主要包含了诊断应用、诊断时间和故障现象等3部分内容。
- 故障定界:这一部分包含了诊断模型推测出的导致应用故障的浅层原因,一般包含以下3种情况:
 - 应用的某个实例故障导致的整体故障。
 - 应用的某个接口或者服务故障导致的整体故障。
 - 。 应用的下游应用故障导致本应用故障。
- 根因分析:这一部分包含了诊断模型推测出来的深层原因,深层原因比较多,根据实际的情况有所不同。
- 数据支持:这一部分包含了得到推测结论的数据支持,不同故障原因的诊断报告中包含的分析数据将不一样。

故障诊断报告示例图-RT诊断







6.2.3. 查看Prometheus监控指标

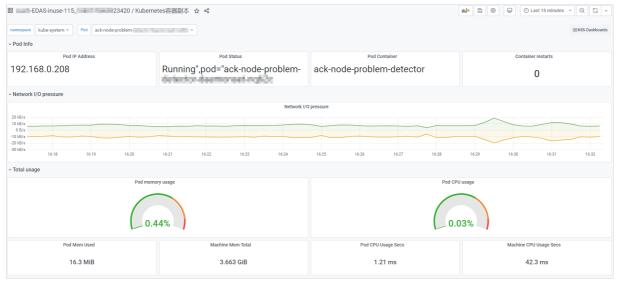
EDAS支持Promet heus监控功能,您可以在监控仪表板查看Pod基本信息、CPU信息、内存信息和网络信息,还可以更改仪表板数据的时间区间、刷新频率等属性。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择**监控 > Promet heus**。 系统自动跳转到Graf ana的Kubernet es容器副本页面。

查看监控指标

在Kubernetes容器副本页面,您可以切换Pod,查看不同Pod的监控指标。



该仪表板展示的监控指标主要包括:

- Pod基本信息: Pod IP地址、Pod状态、Pod容器、容器重启次数。
- 总体使用量信息:例如Pod CPU使用率、Pod内存使用率。
- CPU信息:例如Pod CPU使用率、全部进程CPU使用率。

● 内存信息:例如Pod内存使用率、全部进程内存使用率。

● 网络信息:例如网络I/O压力、Pod网络I/O、所有进程网络I/O。

常用操作

单击面板名称后,可进行以下操作:



- 全屏查看当前面板: 单击View, 或按快捷键V。再次按快捷键V或Esc即可退出全屏模式。
- 编辑当前面板: 单击Edit / 或按快捷键E。在Edit Panel页面修改当前面板的信息 / 然后单击右上角的Apply。
- 分享当前面板:单击**Share**,或依次按下P和S打开分享对话框,获得当前面板的分享链接、嵌入链接或快 照链接。
- 探索当前面板: 单击Explore,或按快捷键X。在Explore页面检查指标、排查故障或探索数据。
- 将当前面板的数据导出为CSV文件:选择Inspect > Data,然后在Data页签中单击Download CSV导出数据。
- 查看当前面板的查询指标:选择Inspect > Query,然后在Query页签中查看请求和响应。
- 获得当前面板的JSON代码:选择Inspect > Panel JSON,然后在JSON页签中拷贝JSON代码。
- 拷贝并粘贴当前面板:选择More > Duplicate,或依次按下P和D即可拷贝当前面板并自动粘贴至当前仪表盘中。
- 拷贝当前面板:选择More > Copy,可以拷贝当前面板。

6.2.4. 应用实例监控

6.2.4.1. 概览

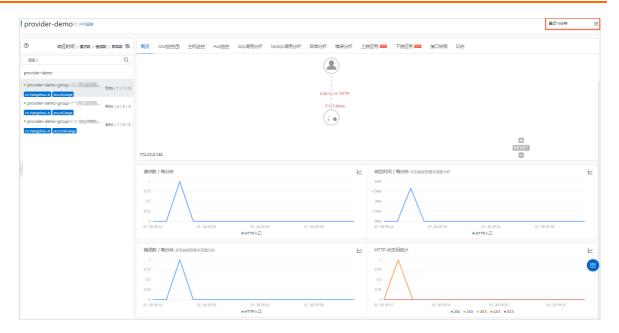
本文说明如何查看应用概览,从而了解应用拓扑、请求数、响应时间、错误数、HTTP状态码等信息。

背景信息

EDAS标准版不支持拓扑图、专业版和铂金版支持拓扑图。

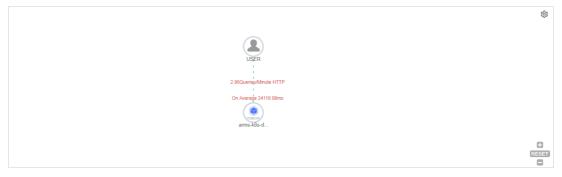
功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - o 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您需要查看的节点,并在页面右侧单击概览页签。



应用拓扑

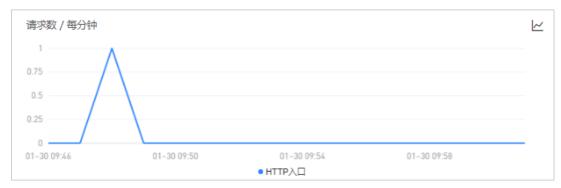
应用拓扑区域显示该应用在指定时间段的内部服务的调用关系拓扑图。



- 1. (可选)在应用拓扑区域,您可以执行以下操作:
 - 单击段图标,设置应用拓扑图的显示设置。
 - ? 说明 设置会被浏览器存储,下次依然生效。
 - 单击加号图标或向上滑动鼠标滚轮,放大应用拓扑图。
 - 单击减号图标或向下滑动鼠标滚轮,缩小应用拓扑图。
 - 单击RESET图标,将应用拓扑图恢复至默认大小。

请求数

请求数区域显示该应用在指定时间段的请求数时序曲线。



- 1. (可选)在请求数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

响应时间

响应时间区域显示该应用在指定时间段的响应时间时序曲线。



- 1. (可选)在响应时间区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击曲线突增点,进行深度分析。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

错误数

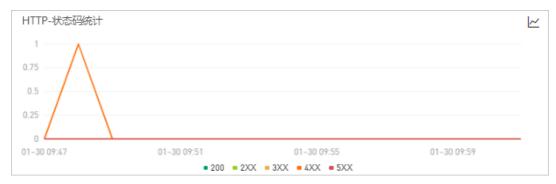
错误数区域显示该应用在指定时间段的错误数时序曲线。



- 1. (可选)在错误数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击曲线突增点,进行深度分析。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

HTTP状态码

HTTP状态码区域显示该应用在指定时间段的HTTP状态码时序曲线。



- 1. (可选)在HTTP-状态码统计区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击图例,隐藏或显示数据。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

6.2.4.2. JVM监控

JVM监控功能用于监控重要的JVM指标,包括堆内存指标、非堆内存指标、直接缓冲区指标、内存映射缓冲区指标、GC(垃圾收集)累计详情和JVM线程数等。本文介绍JVM监控功能和查看JVM监控指标的操作步骤。

功能介绍

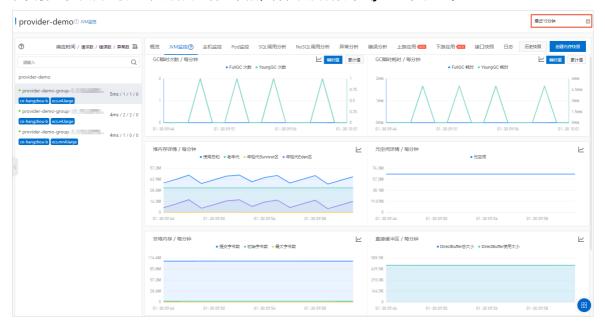
JVM监控功能可监控以下指标:

- GC (垃圾收集) 瞬时和累计详情
 - FullGC次数
 - YoungGC次数
 - FullGC耗时
 - YoungGC耗时
- 堆内存详情
 - 堆内存使用总和
 - 。 堆内存老年代字节数
 - 堆内存年轻代Survivor区字节数
 - 堆内存年轻代Eden区字节数
- 非堆内存
 - 非堆内存提交字节数

- 非堆内存初始字节数
- 非堆内存最大字节数
- 元空间 元空间字节数
- 直接缓冲区
 - DirectBuffer总大小(字节)
 - Direct Buffer使用大小(字节)
- JVM线程数
 - 。 线程总数量
 - 。 死锁线程数量
 - 新建线程数量
 - 阻塞线程数量
 - 。 可运行线程数量
 - 终结线程数量
 - 。 限时等待线程数量
 - 。 等待中线程数量

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在**实例监控**页面左侧选择您需要查看的节点,并在页面右侧单击JVM监控页签。



查看JVM监控指标

您可在**JVM监控**页签查看GC瞬时次数、GC瞬时耗时、堆内存详情、非堆内存详情、元空间详情、直接缓冲区和IVM线程数的时序曲线。

- 单击**GC瞬时次数/每分钟**区域右上角的**瞬时值**或**累计值**按钮,可以切换GC瞬时次数的瞬时值和累计值的时序曲线,默认为瞬时值。
- 单击**GC瞬时耗时/每分钟**区域右上角的**瞬时值**和**累计值**按钮,可以切换GC瞬时耗时的瞬时值和累计值的时序曲线,默认为瞬时值。
- 单击各监控区域的指标名称,可以打开或关闭该指标在图标中的可见性。
 - ② **说明** 每个图表必须至少有一个指标设为可见,这意味着当图表中只有一个指标时,您无法关闭该指标的可见性。

6.2.4.3. 主机监控

主机监控功能用于监控CPU、内存、Disk(磁盘)、Load(负载)、网络流量和网络数据包的各项指标。本文介绍主机监控功能和查看主机监控指标的操作步骤。

功能介绍

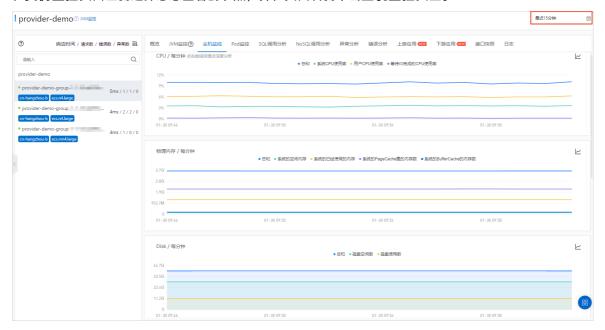
主机监控功能可监控以下指标:

- CPU
 - CPU使用率总和
 - 系统CPU使用率
 - 用户CPU使用率
 - 。 等待IO完成的CPU使用率
- 物理内存
 - 系统总内存
 - 。 系统空闲内存
 - 。 系统已使用内存
 - 。 系统PageCache中的内存
 - o 系统BufferCache中的内存
- Disk (磁盘)
 - 系统磁盘总字节数
 - 系统磁盘空闲字节数
 - 系统磁盘使用字节数
- Load (负载)系统负载数
- 网络流量
 - 。 网络接收的字节数
 - 。 网络发送的字节数
- 网络数据包
 - 。 每分钟网络接收的报文数
 - 每分钟网络发送的报文数

- 。 每分钟网络接收的错误数
- 每分钟网络丢弃的报文数

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击主机监控页签。



查看主机监控指标

您可以在**主机监控**页签查看CPU、物理内存、Disk(磁盘)、Load(负载)、网络流量和网络数据包的时序曲线。

单击各区域的指标名称(例如系统CPU使用率),可打开或关闭该指标在图表中的可见性。

② 说明 每个图表必须至少有一个指标设为可见,这意味着当图表中只有一个指标时,您无法关闭该指标的可见性。

6.2.4.4. Pod监控

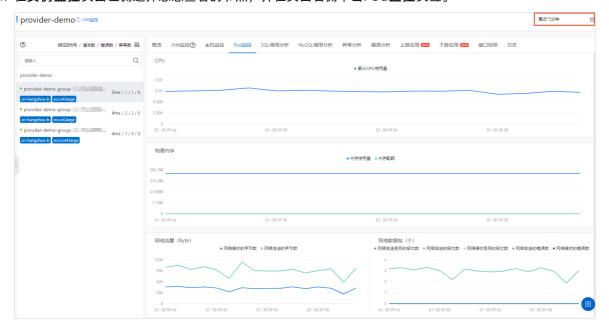
本文说明如何查看Pod监控,从而了解应用的Pod情况,包括CPU、物理内存、网络流量、网络数据包等信息。

应用限制

EDAS标准版不支持Pod监控, 专业版和铂金版支持Pod监控。

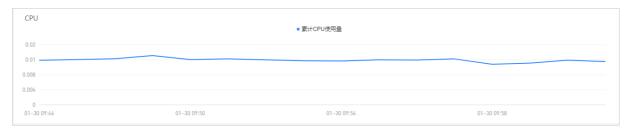
功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击Pod监控页签。



CPU

CPU区域显示该应用在指定时间段的Pod的CPU情况(累计CPU使用量)。



您可以将光标移到统计图上,查看具体时刻的累计CPU使用量。

物理内存

物理内存区域显示该应用在指定时间段的Pod的物理内存情况,包括以下指标:

- 内存使用量
- 内存配额

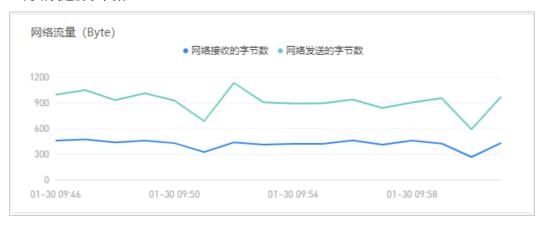


- 1. (可选)在物理内存区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 单击图例,隐藏或显示数据。

网络流量

网络流量区域显示该应用在指定时间段的Pod的网络流量情况,包括以下指标:

- 网络接收的字节数
- 网络发送的字节数



- 1. (可选)在网络流量区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 单击图例,隐藏或显示数据。

网络数据包

网络数据包区域显示该应用在指定时间段的Pod的网络数据包情况,包括以下指标:

- 网络发送丢弃的报文数
- 网络发送的报文数
- 网络接收丢弃的报文数
- 网络发送的错误数
- 网络接收的错误数



- 1. (可选)在网络数据包区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 单击图例, 隐藏或显示数据。

6.2.4.5. SQL调用分析

本文说明如何查看SQL调用分析,从而了解应用的SQL调用情况。

应用限制

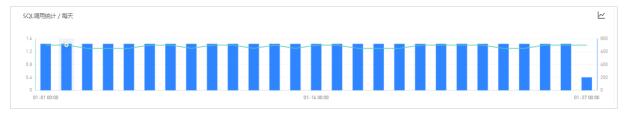
EDAS标准版不支持SQL调用分析,专业版和铂金版支持SQL调用分析。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在**实例监控**页面左侧选择您想查看的节点,并在页面右侧单击SQL调用分析页签。

SOL调用统计

调用统计区域显示该应用在指定时间段的SQL调用时序曲线。



- 1. (可选)在SOL调用统计页签下,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

SOL语句列表

SQL语句列表显示该应用在指定时间段的所有SQL语句的列表。



- 1. (可选)在SQL语句列表,您可以执行以下操作:
 - 在SQL语句的操作列,单击调用统计,查看该SQL语句的SQL调用时序曲线。
 - 在SQL语句的操作列,单击接口快照,查看该SQL语句调用的接口的快照。 接口快照说明,请参见接口快照。

6.2.4.6. NoSQL调用分析

本文说明如何查看NoSQL调用分析,从而了解应用的NoSQL调用情况。

应用限制

EDAS标准版不支持NoSQL调用分析,专业版和铂金版支持NoSQL调用分析。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击NoSQL调用分析页签。

NoSQL调用统计

调用统计区域显示该应用在指定时间段的NoSQL调用时序曲线。



- 1. (可选)在NoSQL调用统计页签下,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

操作命令列表

操作命令列表显示该应用在指定时间段的所有NoSQL调用的操作命令的列表。



- 1. (可选)在操作命令列表,您可以执行以下操作:
 - 在操作命令右侧操作列,单击调用统计,查看该操作命令的NoSQL调用统计。
 - 在操作命令右侧操作列,单击接口快照,查看该操作命令的接口的快照。 接口快照说明,请参见接口快照。

6.2.4.7. 异常分析

本文说明如何查看异常分析,从而了解应用的异常情况。

应用限制

EDAS标准版不支持异常分析,专业版和铂金版支持异常分析。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击异常分析页签。

异常统计

异常统计区域显示该应用在指定时间段的异常的堆积柱状图和异常列表。



- 1. (可选)在异常统计区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - o 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

异常列表

异常列表显示该应用在指定时间段的所有异常的列表。

1. (可选)在异常列表,您可以执行以下操作:



- 在异常的操作列,单击调用统计查看该异常的堆积柱状图。
- 在异常的操作列,单击接口快照查看该异常的接口快照。 接口快照说明,请参见调用链查询。
- 在异常的操作列,单击详情查看该异常的详细信息。

6.2.4.8. 错误分析

本文说明如何查看错误分析,从而了解应用的错误情况。

应用限制

EDAS标准版不支持错误分析,专业版和铂金版支持错误分析。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击错误分析页签。



错误数

错误数区域显示该应用在指定时间段的错误时序曲线。



- 1. (可选)在错误数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

HTTP状态码

HTTP-状态码统计区域显示该应用在指定时间段的HTTP状态码时序曲线。



- 1. (可选)在HTTP-状态码统计区域,您可以执行以下操作:
 - 将光标移到统计图上,查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击图例,隐藏或显示数据。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

错误列表

错误列表显示该应用在指定时间段的所有错误的列表。



在错误右侧的TraceId列,单击TraceId名称,可以查看该错误的调用链。

6.2.4.9. 上游应用

对于某个应用而言,上游应用是指向该应用发送数据的应用。本文说明如何查看上游应用情况,包括响应时间、请求数、错误数等信息。

应用限制

EDAS标准版不支持查看上游应用情况,专业版和铂金版支持查看上游应用情况。

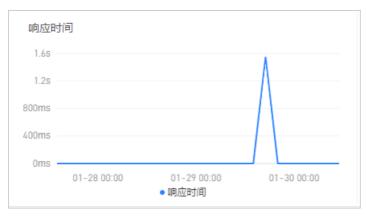
功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击上游应用页签。



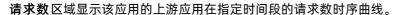
响应时间

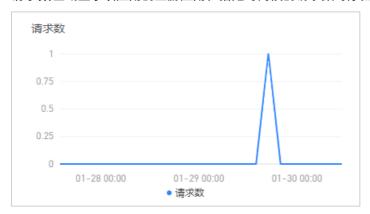
响应时间区域显示该应用的上游应用在指定时间段的响应时间时序曲线。



- 1. 在响应时间区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

请求数

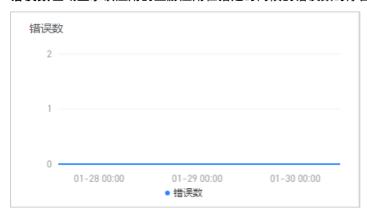




- 1. 在请求数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

错误数

错误数区域显示该应用的上游应用在指定时间段的错误数时序曲线。



- 1. 在错误数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

6.2.4.10. 下游应用

对于某个应用而言,下游应用是指从该应用接收数据的应用。本文说明如何查看应用的下游应用情况,包括响应时间、请求数、错误数等。

应用限制

EDAS标准版不支持查看下游应用情况,专业版和铂金版支持查看下游应用情况。

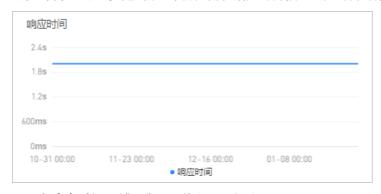
功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。

- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击下游应用页签。

响应时间

响应时间区域显示该应用的下游应用在指定时间段的响应时间时序曲线。



- 1. 在响应时间区域, 您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

请求数

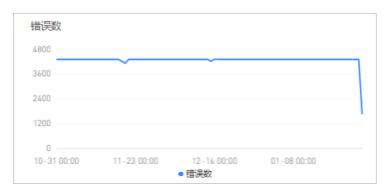
请求数区域显示该应用的下游应用在指定时间段的请求数时序曲线。



- 1. 在请求数区域,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

错误数

错误数区域显示该应用的下游应用在指定时间段的错误数时序曲线。



- 1. 在错误数区域,您可以执行以下操作:
 - 将光标移到统计图上,查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。

6.2.4.11. 调用链查询

本文说明如何查看应用的所有接口的被调用情况,包括产生时间、耗时、状态等信息。

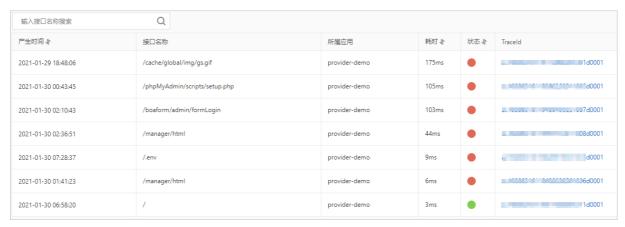
功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击调用链查询页签。



调用链查询

调用链查询页签下显示该应用在指定时间段内被调用的所有接口的列表。



- 1. (可选)在调用链查询页签下,您可以执行以下操作:
 - 在搜索框,输入接口名称,然后单击\图标,查看指定接口的快照。
 - 在接口的TraceId列,单击TraceId名称,查看该接口的调用链路。

6.2.4.12. 日志

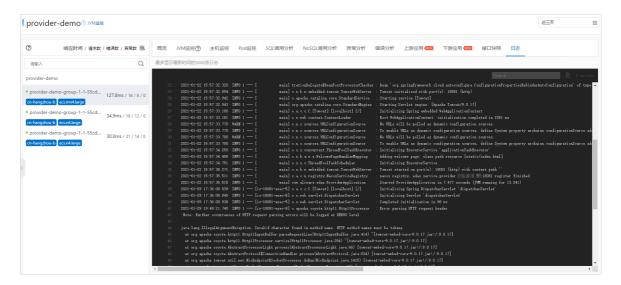
本文说明如何查看应用的Pod日志,从而了解Pod情况。

应用限制

EDAS标准版不支持查看应用的Pod日志,专业版和铂金版支持查看应用的Pod日志。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在实例监控页面左侧选择您想查看的节点,并在页面右侧单击日志页签。



日志

日志页签下显示该应用的Pod实例的最新时间的日志。

? 说明 最多显示最新时间的5000条日志。

```
2021-01-22 15-67-32 300 1BF0 1 — [ axis] tratically equiple and traces of the control of the con
```

1. 在日志页签下的搜索框,您可以输入关键字,单击置图标,筛选日志。

6.2.4.13. 内存快照

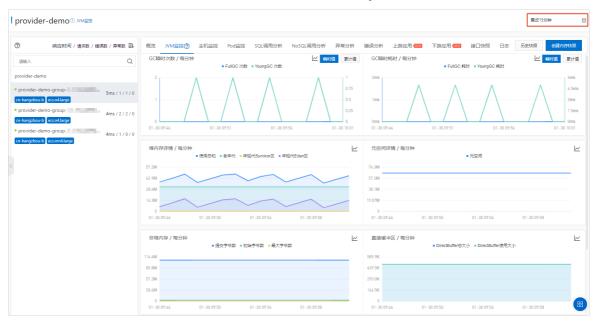
JVM监控可以直观展示指定时间段内的多项内存指标,然而虽然图表能体现出内存使用量过大的情况,但无法显示具体信息,因此不能帮助您排查问题的原因。此时您可以创建内存快照,通过详细的日志查看内存占用的详细信息。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部

菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。

- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 实例监控。
- 4. 在**实例监控**页面左侧选择您需要查看的节点,并在页面右侧单击JVM监控页签。



创建内存快照

- 1. 在JVM监控页签右上角,单击创建内存快照。
 - ② 说明 如果单击创建内存快照时,上一个快照任务仍在运行,则系统会弹出错误消息。请您耐心等待上一个快照任务运行完毕。目前仅支持为Linux系统新建内存快照。
- 2. 在**创建内存快照**对话框中选择一个IP, 并单击保存。



⑦ 说明 如果在应用详情页面左侧已选择目标实例,则IP字段会默认选中该实例的IP地址。

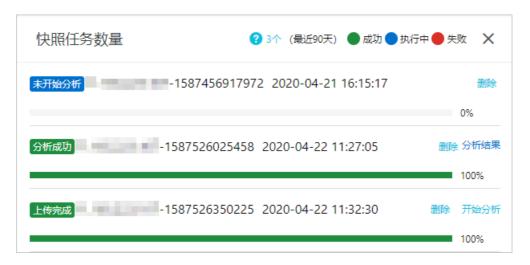
查看内存快照详情

1. 在IVM监控页签右上角,单击历史快照。

在**快照任务数量**面板展示了任务执行状态:绿色表示快照任务执行成功,蓝色表示快照任务执行中, 红色表示快照任务执行失败。

快照任务名称的信息依次为:

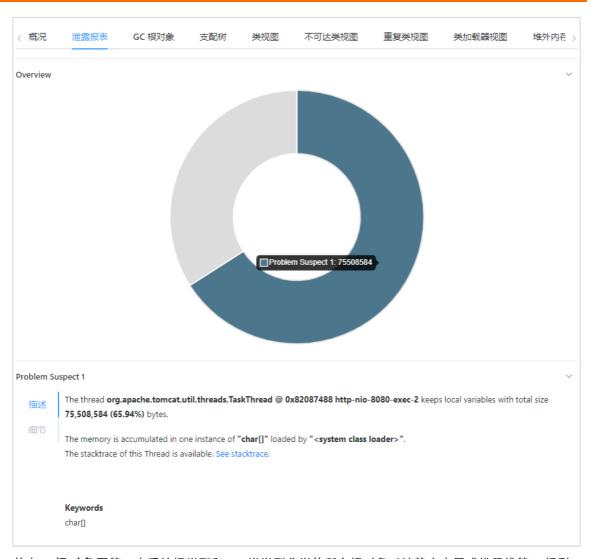
- 。 内存分析状态
- 由IP地址和时间戳组成的快照任务ID
- 。 快照创建时间



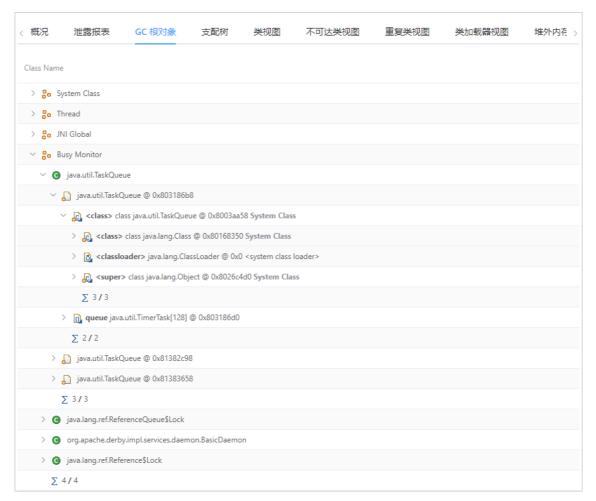
- 2. 单击开始分析,弹出提示对话框,单击确认。
- 3. 单击**分析结果**,打开内存分析页面查看内存分析详情,帮助您查找内存泄漏信息,用于减少内存消耗。
 - 单击**概况**页签,查看堆使用大小、类数量、对象数量、类加载器数量和根对象数量等统计信息,以及以环形图显示的内存占用情况。



○ 单击**泄露报表**页签,可以查看消耗过多内存的可疑对象。单击页面下方的Problem Suspect 查看可 疑对象的具体实例、内存占用的情况以及类加载信息。



○ 单击GC根对象页签,查看按根类型和Java类类型分类的所有根对象(被静态变量或线程栈等GC根引用的对象)。



○ 单击**支配树**页签,查看堆中对象之间的支配关系,可以识别出占用了大量内存的对象和它们之间的依赖关系。



○ 单击类视图页签,查看各个类型的堆使用情况和对应实例的数量。



○ 单击不可达类视图页签,查看堆中未被引用的对象的大小及类型情况。



单击重复类视图页签,查看被多个类加载器加载的类型情况。



○ 单击**类加载器视图**页签,查看应用当前使用到的所有类加载,以及每个类加载器加载类的情况。例如加载了哪些类,这些类有多少实例等。



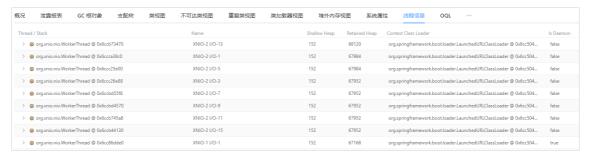
○ 单击**堆外内存视图**页签,查看应用目前使用的所有java.nio.DirectByteBuffer对象以及对应的堆外内存信息,用于排查由堆外内存导致物理内存消耗过多的问题。



○ 单击**系统属性**页签,查看系统参数和环境变量。



○ 单击**线程信息**页签,查看线程的相关信息,例如线程名、堆占用情况、调用栈信息、局部变量等。通过该视图您可以分析线程过多、死锁、调用栈过深等问题。



○ 单击OQL页签,查询堆中的各种信息,例如所有长度大于2000的字符串。



6.2.5. 服务和接口监控

服务和接口监控功能用于监控应用下的接口调用详情。

支持的框架

本功能模块可自动发现和监控以下Web框架和RPC框架中提供的接口:

- Tomcat 7+
- Jetty 8+
- Resin 3.0+
- Undertow 1.3+
- WebLogic 11.0+
- SpringBoot 1.3.0+
- HSF 2.0+
- Dubbo 2.5+

功能入口

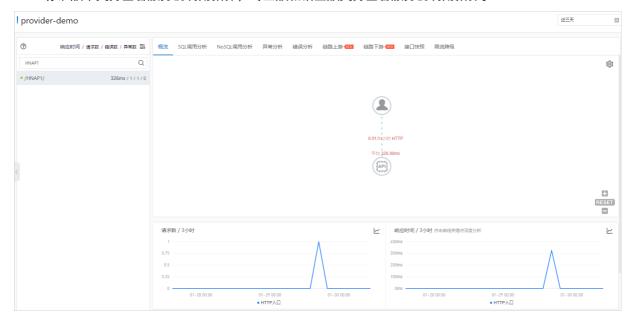
1. 登录EDAS控制台。

- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击应用列表,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群类型下 拉列表中选择容器服务/Serverless K8s集群,然后单击目标应用名称。
- 3. 在应用详情页面的左侧导航栏选择监控 > 服务/接口监控。

查看服务概览信息

概览页签列出了应用监控探针自动发现的所有接口。您可以按照响应时间、请求数、错误数或异常数对该列 表排序。选中一个服务,即可在**概览**页签上查看该服务的详细调用拓扑,以及请求数、响应时间、错误数的 时序曲线。

EDAS标准版不支持查看服务的调用拓扑,专业版和铂金版支持查看服务的调用拓扑。



查看SOL和NoSOL调用分析

EDAS标准版不支持SQL调用分析和NoSQL调用分析,专业版和铂金版支持SQL调用分析和NoSQL调用分析。

在SQL调用分析页签和NoSQL调用分析展示了左侧选中服务的代码段内所发起的SQL和NoSQL请求列表。借助此页签,您可以找出是哪一个SQL或NoSQL造成某个服务过慢。您还可以单击某个SQL或NoSQL中的调用链查询来查看一个SQL或NoSQL执行逻辑所处的完整代码链路。

查看异常分析

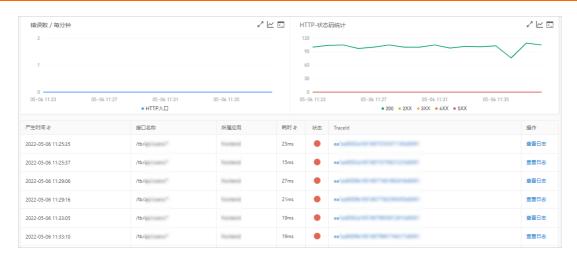
EDAS标准版不支持查看异常分析,专业版和铂金版支持查看异常分析。

在**异常分析**页签展示了左侧选中服务的代码段内所抛出的Java异常。您还可以单击某个异常中的**调用链查** 询来查看一个异常堆栈所处的完整代码链路。

查看错误分析

EDAS标准版不支持查看错误分析,专业版和铂金版支持查看错误分析。

在错误分析页签展示了应用的错误和HTTP状态码统计数据。您还可以单击需要查看的Traceld,即可在新页面查看调用链路相关信息。



查看链路上游和链路下游的接口调用情况

EDAS标准版不支持查看链路上游和链路下游的接口调用情况,专业版和铂金版支持查看链路上游和链路下游的接口调用情况。

链路上游和链路下游页签分别列出了应用上游(调用应用的一方)和应用下游(被应用调用的一方)的接口及其调用性能指标,包括响应时间、请求数和错误数。

在链路上游和链路下游页签上,可按需执行以下操作:

- 在页签顶部单击**全部折叠/展开**,即可折叠或展开下方的所有接口。
- 在页签顶部的搜索框内输入应用名称或接口(Span)名称的关键字,并单击搜索图标,即可筛选出符合 条件的接口。
- 单击接口信息所在的折叠面板,或者单击行末的上箭头或下箭头,即可展开或折叠该接口的性能指标信息。

查看服务接口的调用情况

调用链查询页签展示了该服务接口中的参数详情,以及根据该参数详情生成的总调用链次数和调用链响应时间两个图表。



● 在总调用链次数图表中,超过500毫秒被定义为慢调用。慢调用受应用设置中阈值影响,但不包含因为异步产生的调用本地API的情况。

② **说明** 在总调用链次数图表中,正常调用次数指的是总调用链次数。慢调用次数包含在正常调用次数里。

- 在调用链响应时间图表中,其展示的数据是基于接口调用数据的二次统计,并受采样率影响。
- 在参数详情表格中,单击Traceld可以查看调用链路和业务轨迹。单击**查看日志**可以查看该接口的调用日志。

查看限流降级

EDAS标准版不支持查看限流降级,专业版和铂金版支持查看限流降级。

部署在容器服务K8s集群或Serverless K8s集群中的应用支持通过重新部署来接入AHAS的试用版,并可在EDAS内嵌的监控页面中实时监控流量规则,全面可视化地保障您的应用可用性。

接入AHAS和配置限流降级规则的操作步骤,请参见实现K8s集群应用的限流降级。

6.2.6. 数据库调用监控

本文说明如何查看数据库调用,从而了解应用的数据库调用情况。

功能入口

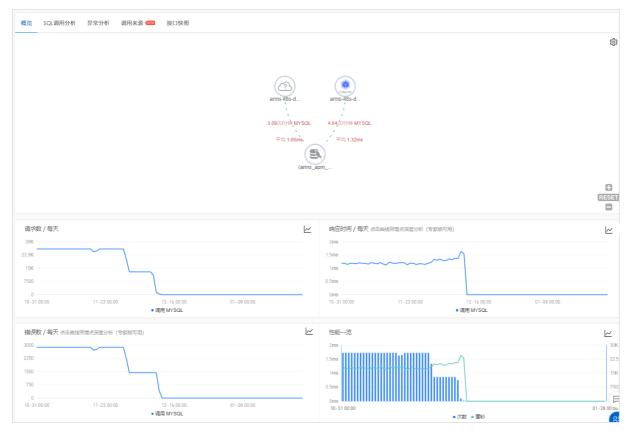
- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情页面的左边导航栏选择监控 > 数据库调用。



概览

EDAS标准版不支持查看调用关系拓扑图,专业版和铂金版支持查看调用关系拓扑图。

概览页签下显示数据库的调用关系拓扑、请求数时序曲线、响应时间时序曲线、错误数时序曲线等信息。



- 1. (可选)在概览页签下,您可以执行以下操作:
 - 单击微图标,设置应用拓扑图的显示设置。
 - ? 说明 设置会被浏览器存储,下次依然生效。
 - 单击加号图标或向上滑动鼠标滚轮,放大应用拓扑图。
 - 单击减号图标或向下滑动鼠标滚轮,缩小应用拓扑图。
 - 单击RESET图标,将应用拓扑图恢复至默认大小。
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

SQL调用分析

SQL调用分析页签下显示数据库的SQL调用数柱状图、响应时间时序曲线、SQL语句列表等信息。



- 1. (可选)在SQL调用分析页签下,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。
 - 在SQL语句的操作列,单击调用统计,查看该SQL语句的SQL调用统计。
 - 在SQL语句的操作列,单击接口快照,查看该SQL语句调用的接口的快照。

异常分析

EDAS标准版不支持查看数据库的异常情况,专业版和铂金版支持查看数据库的异常情况。

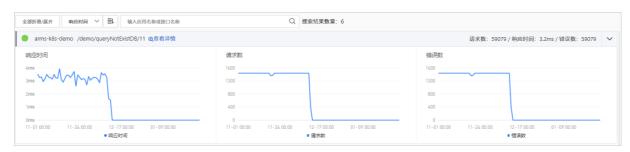
异常分析页签下显示数据库的异常情况。



- 1. (可选)在异常分析页签下,您可以执行以下操作:
 - 将光标移到统计图上, 查看统计情况。
 - 使用光标选中一段时间,查看指定时间段的统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。
 - 在异常的**操作**列,单击**调用统计**,查看该异常的统计情况。
 - 在异常的操作列,单击**详情**,查看该异常的详情。

调用来源

调用来源页签下显示数据库的调用来源情况。



- 1. (可选)在调用来源页签下,您可以执行以下操作:
 - 在搜索框,输入应用或接口名称,单击○图标,查看应用或接口的情况。
 - 在调用来源应用的接口右侧,单击查看详情,查看该接口的快照。
 - 将光标移到统计图上, 查看统计情况。
 - 单击|~图标,查看该指标在某个时间段的统计情况或对比不同日期同一时间段的统计情况。

数据库的调用链情况

调用链查询页签下显示数据库的所有接口的调用情况。



- 1. (可选)在调用链查询页签下,您可以执行以下操作:
 - 在搜索框,输入接口名称,然后单击 ②图标,查看接口的调用情况。
 - 在接口的TraceId列,单击TraceId名称,查看该接口的调用链路。

6.2.7. NoSQL调用

本文说明如何查看NoSQL调用,从而了解应用的NoSQL调用情况,包括概况、NoSQL调用、异常、接口等信息。

应用限制

- EDAS标准版不支持NoSQL调用,专业版和铂金版支持NoSQL调用。
- 该功能需要升级探针版本至2.7.1.1 及以上。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单

击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。

- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情页面的左边导航栏选择监控 > NoSOL调用。

功能介绍

- 概览页签下显示NoSQL数据库的调用关系拓扑、请求数时序曲线、响应时间时序曲线、错误数时序曲线等信息。
- NoSQL调用分析页签下显示NoSQL数据库的NoSQL调用数柱状图、响应时间时序曲线、操作命令列表等信息。
- 在**异常分析**页签展示了左侧选中服务的代码段内所抛出的Java异常。您还可以单击某个异常中的调用链来 查看一个异常堆栈所处的完整代码链路。
- 调用链查询页签下显示该应用在指定时间段内被调用的所有接口的列表。

6.2.8. 外部调用

当您需要定位应用外部调用缓慢或出错的问题时,可以使用外部调用功能实现问题定位。

应用限制

EDAS标准版不支持查看外部调用详情,专业版和铂金版支持查看外部调用详情。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情页面的左边导航栏选择监控 > 外部调用。

外部调用页面的左侧列表展示了应用的所有外部调用。您可以按照响应时间、请求数、错误数或异常数对该列表排序。

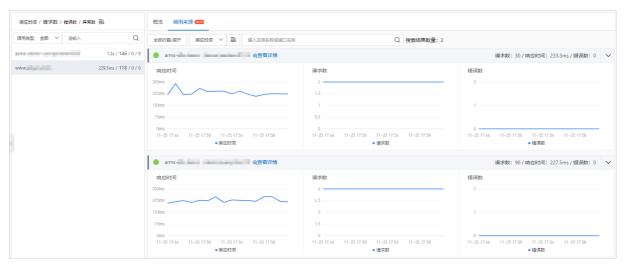
概览

在左侧列表选中其中一个外部调用,可在概览页签上查看该外部调用的请求数、响应时间、错误数以及HTTP-状态码的时序曲线。



调用来源

在左侧列表选中其中一个外部调用,可在**调用来源**页签上查看该外部调用所有接口的响应时间、请求数和错误数的时序曲线。



在调用来源页签上,可按需执行以下操作:

- 在页签顶部单击**全部折叠/展开**,即可折叠或展开下方的所有接口。
- 在页签顶部的搜索框内输入应用名称或接口(Span)名称的关键字,并单击搜索图标,即可筛选出符合 条件的接口。
- 单击接口信息所在的折叠面板,或者单击行末的上箭头或下箭头,即可展开或折叠该接口的性能指标信息。

6.2.9. MQ监控

MQ监控可展示消息队列Rocket MQ版的Topic发布和订阅消息的情况。

应用限制

EDAS标准版不支持MQ监控,专业版和铂金版支持MQ监控。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情页面的左边导航栏选择监控 > MQ监控。



功能介绍

- 在概览页签中以拓扑图展示应用与MQ数据源之间的消息发布和订阅关系。
- 在发布端统计页签展示消息发布的统计数据,包括请求数、响应时间和错误数。
- 在**订阅端统计**页签展示消息订阅的统计数据,包括消息请求数、响应时间和错误数。
- 在**调用链查询**页签提供关于消息发布和订阅的调用链,您可以通过TraceId链接查看完整调用链以及诊断问题原因。

6.2.10. 应用诊断

6.2.10.1. 实时诊断

实时诊断功能适用于在短时间内密切监控应用性能和定位问题原因的场景。本文介绍实时诊断功能的使用方法。

背景信息

当您需要密切监控一小段时间内的应用性能时,例如发布应用或者对应用进行压测时,可以使用实时诊断功能。开启实时诊断后,会持续监控应用5分钟,并在这5分钟内全量上报调用链数据。接下来,您就能以出现性能问题的调用链路为起点,通过方法栈瀑布图和线程剖析等功能定位问题原因。

应用限制

EDAS标准版不支持应用的实时诊断功能,专业版和铂金版支持应用的实时诊断功能。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击应用列表,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群类型下 拉列表中选择容器服务/Serverless K8s集群,然后单击目标应用名称。
- 3. 在应用详情的左侧导航栏选择监控 > 应用诊断-实时诊断。

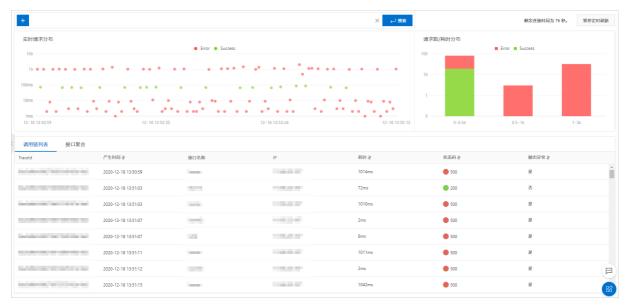
开启和终止实时诊断

首次进入**实时诊断**页面时,默认自动开启实时诊断。其他情况下,如需开启实时诊断,请单击页面右上角的**开启实时诊断**。

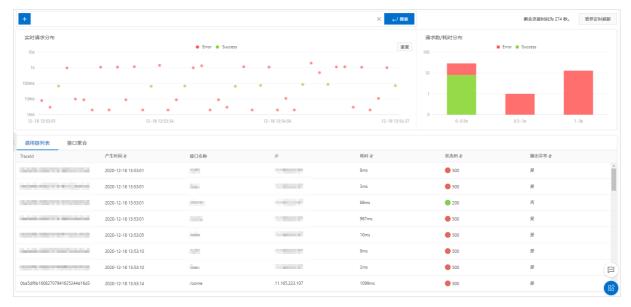
实时诊断将于自动开启5分钟后自动终止。如需提前终止实时诊断,请单击页面右上角的暂停定时刷新。

查看实时监控数据

在**实时请求分布**和**请求数/耗时分布**区域,您可以查看截至当前时间点捕捉到的最后1000次请求统计数据。



在**实时请求分布**区域的图表中,用鼠标框选一段时间区间,即可将所选时间区间设为数据可视时间区间,即图表中仅显示该时间区间内的数据。此后,单击图表右上角的重置即可恢复为默认视图。



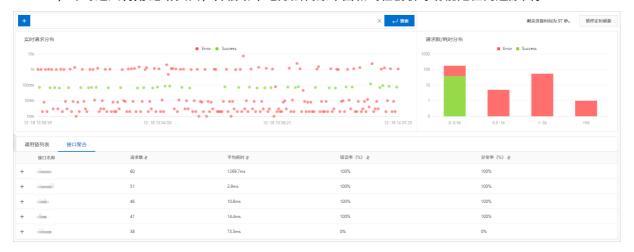
筛选监控数据

您可以按照接口名称和IP筛选页面上显示的请求监控数据。

- 1. 在实时请求分布区域上方单击+图标。
- 2. 在下拉框中选择一个API或IP, 并单击**搜索**。 仅选中接口的请求监控数据会显示在页面上。

查看调用链信息

在调用链列表和接口聚合页签上,您可以查看相应时间区间内捕捉到的全部调用链信息。单击一个 Traceld,即可进入调用链路页面,并借助本地方法栈瀑布图和线程剖析等功能定位问题原因。



6.2.10.2. 异常分析

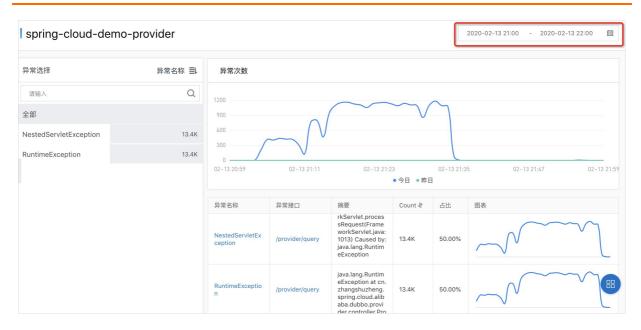
EDAS提供了应用异常分析功能,包括异常数量统计、每类异常次数统计以及异常发生的端口等。本文介绍 异常分析功能的使用方法。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 应用诊断-异常分析。

异常分析

异常分析页面的左侧列表展示了应用的全部异常。您可以在右上角选择查看的时间段,在异常列表中选中某一异常线程,然后根据右侧的**异常次数**时序曲线图分析异常变化,例如分析不同时间点的异常次数是否过多。



6.2.10.3. 线程分析

线程分析功能提供线程粒度的CPU耗时和每类线程数量的统计,并且每5分钟记录一次线程的方法栈并聚合,可真实还原代码执行过程,帮助您快速定位线程问题。当发现集群的CPU使用率过高,或者出现大量慢方法时,可以通过线程分析功能找到消耗CPU最多的线程或方法。

应用限制

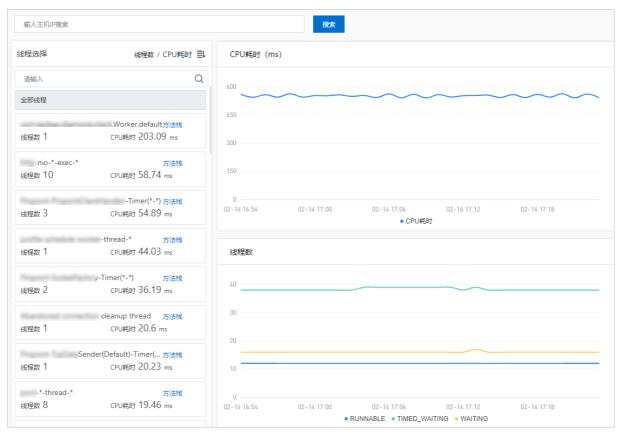
EDAS标准版不支持线程分析功能,专业版和铂金版支持线程分析功能。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 应用诊断-线程分析。

进行线程分析

线程分析页面的左侧列表展示了应用的全部线程,您可以根据**CPU耗时**统计快速发现异常线程。选中某一异常线程后,再根据右侧的**CPU耗时**和**线程数**曲线图分析CPU耗时与线程数变化,例如分析每分钟的线程总数是否过多。



您还可以单击异常线程的**方法栈**,查看指定时间内的真实运行方法栈,例如查看处于BLOCKED状态的线程对应的方法,从而优化指定代码段,以便降低CPU使用率。



6.2.10.4. Arthas诊断(新版)

Arthas是诊断Java领域线上问题的利器,利用字节码增强技术,可以在不重启JVM进程的情况下,查看程序的运行情况。

前提条件

- 新开通EDAS的账号或15天内未使用Arthas诊断功能的账号仅支持使用新版Arthas诊断功能。
- 已接入应用监控。具体操作,请参见应用监控概述。 EDAS新版Arthas诊断仅对以下地域开放。

地域名称	所在城市	Region ID
华东1	杭州	cn-hangzhou
华东2	上海	cn-shanghai
华北2	北京	cn-beijing

地域名称	所在城市	Region ID
华北3	张家口	cn-zhangjiakou
华南1	深圳	cn-shenzhen

- 应用的编程语言需要为Java。
- 3658端口未被占用。

背景信息

EDAS提供的Arthas诊断功能主要用于补齐EDAS在实时诊断方面的能力。EDAS的Arthas诊断功能包括以下几种类型:

- JVM概览: 查看当前JVM进程实时的内存使用情况、系统信息、系统变量和环境变量。
- 线程分析:查看当前JVM进程的线程耗时情况以及指定线程的实时方法栈。
- 方法执行分析: 抓取任意方法(非JDK方法)满足指定条件的一次执行记录,记录该方法的参数、异常、返回值以及方法内部各个方法执行耗时。
- 对象查看器: 查看任意类的某个实例实时的属性取值情况。
- 实时看板:常见组件的实时看板,例如,Druid连接池的实时看板可以看到连接池的配置、使用情况以及 SQL执行耗时情况。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择监控 > 应用诊断-Art has诊断。
- 4. 在Arthas诊断页面,在顶部下拉列表选择待诊断的应用实例。
 - 若该实例的Agent版本未升级到2.7.1.3或以上,页面会提示您需要先升级Agent。
 - 若该实例的Agent版本已升级到2.7.1.3或以上,页面会显示该实例的Arthas诊断信息。

JVM概览

JVM概览支持查看应用的JVM相关信息,包括JVM内存、操作系统信息、变量信息等,帮助您了解JVM的总体情况。

Arthas诊断页面默认显示JVM概览页签,您可以在JVM概览页签查看以下信息:

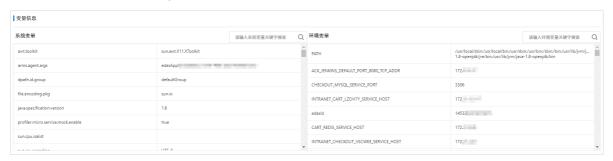
● JVM内存: JVM内存的相关信息,包括堆内存使用情况、非堆内存使用情况、GC情况等。



● 操作系统信息:操作系统的相关信息,包括平均负载情况,操作系统名称、操作系统版本、Java版本等。



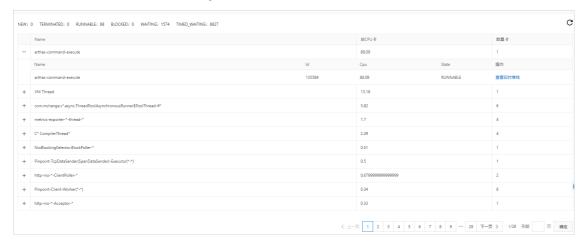
● **变量信息**:变量的相关信息,包括系统变量和环境变量。



线程耗时分析

线程耗时分析支持显示该应用的所有线程和查看线程的堆栈信息,帮助您快速定位耗时较高的线程。

1. 在Arthas诊断页面,单击**线程耗时分析**页签。 **线程耗时分析**页签会实时获取当前JVM进程的线程耗时情况,并将相似线程聚合。



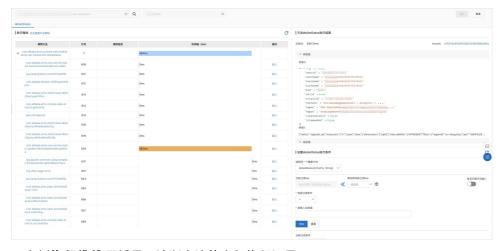
- 2. 单击线程左侧的+图标展开线程明细,可以查看线程的ID、CPU使用率和状态。
- 3. 如需查看目标线程的堆栈信息,您可以在目标线程右侧的操作列,单击查看实时堆栈。



方法执行分析

方法执行分析支持抓取方法的某一次执行的耗时、入参、返回值等信息和钻入,帮助您快速定位导致慢调用的根本原因,以及问题线下无法复现或日志缺失等场景。

- 1. 在Arthas诊断页面,单击方法执行分析页签。
- 2. 在方法执行分析页签的搜索框中输入类名的关键词,然后单击 Q图标。
- 3. 在搜索到的类中选择需要诊断的类,然后在右侧方法选择框选取该类的某个方法,单击**确定**。 页面将会显示EDAS随机抓取的该方法的某一次执行的信息。



- 左侧执行堆栈区域显示诊断方法的内部执行记录。
 - 如需钻入某个内部方法,在其右侧**操作**列,单击**钻入**。

■ 如需查看方法源码,单击**执行堆栈**区域顶部的**查看方法源码**。 如下图所示,每一次内部方法的执行耗时都会以注释的方式显示在源代码中。

- 右侧方法执行结果区域显示方法执行的参数值、返回值、异常、成员变量以及此次方法执行的 TraceID。
- o 右侧**设置执行条件**区域执行以下步骤,可以设置方法执行条件来抓取满足条件的方法执行记录。
 - a. 选择当前方法中的一个重载方法。
 - b. 在请选择初始过滤key下拉框选择初始过滤Key的类型,单击《图标。

初始过滤Key的类型:

■ params[n]: 方法的第n个参数。

■ returnObj: 方法的返回值。

■ 方法执行耗时:方法执行的耗时。

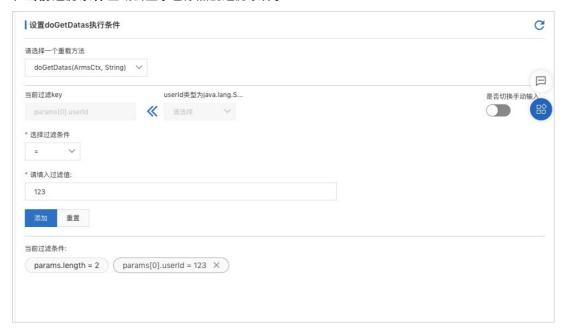
■ 是否抛出异常:方法执行时是否抛出异常。

② **说明** 如果选取的初始过滤Key为嵌套类型,则还需要继续选择该嵌套类型的内部字段,直到选择字段为基础类型。

- c. 选择过滤条件。
- d. 输入过滤值。

e. 单击添加。

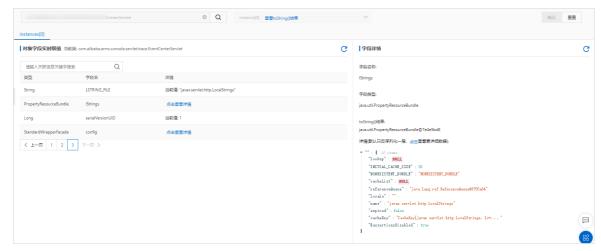
在当前过滤条件区域会显示已添加的过滤条件。



对象查看器

对象查看器用于查看一些单例对象当前的状态,用于排查应用状态异常问题,例如应用配置、黑白名单、成员变量等。

- 1. 在Arthas诊断页面,单击**对象查看器**页签。
- 2. 在对象查看器页签的搜索框中输入类名的关键词,然后单击搜索图标。
- 3. 在搜索到的类中选择需要诊断的类,然后在右侧实例选择框选择该类的某个实例,单击**确定**。 页面则会显示该实例中当前字段的实时取值。



- 对于简单类型的字段,在左侧对象字段实时取值区域的详情列会是字段的取值。
- 对于复制类型的字段,在左侧**对象字段实时取值**区域的**详情**列单击**点击查看详情**,在右侧字**段详** 情区域查看字段取值详情。

字段详情区域仅支持将复杂类型字段反序列化一层进行展示,如果需要查看更具体的数据,单击点击查看字段更详细的数据。

实时看板

实时看板用于查看系统中用到的关键组件的实时状态,例如查看数据库连接池的使用情况、HTTP连接池的使用情况等,有利于排查资源类型的问题。

- 1. 在Arthas诊断页面,单击**实时看板**页签。
- 2. 在**实时看板**页签的下拉列表中选择一个看板,然后在右侧实例选择框选择该看板的一个实例,单击**添加到看板**。

页面显示该看板的实时情况。如下图显示为一个Druid连接池的实时状态信息,包括基础配置、连接池状态、执行耗时分布等。



6.2.10.5. 日志分析

当应用出现业务异常问题时,您可以分析业务日志,精准定位业务异常。

应用限制

EDAS标准版不支持日志分析,专业版和铂金版支持日志分析。

步骤一:关联业务日志

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择应用设置 > 应用监控设置。
- 4. 在应用监控设置页面,单击自定义配置页签。
- 5. 在**业务日志关联设置区**域,设置是否在应用的业务日志中关联调用链的Traceld信息。

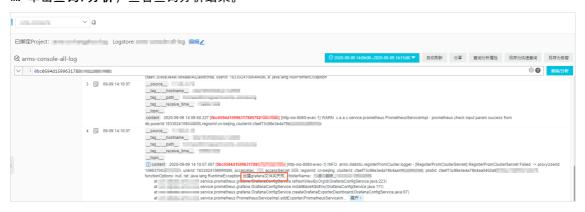


○ 开启此开关后,会在业务日志中自动生成调用链的Traceld。

- 如果您想要实现精准定位,还需要绑定Project和Logstore来实现。
- 6. 在自定义配置页签底部,单击保存。

步骤二:查询并分析日志

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择资源管理 > 容器服务K8s集群(或资源管理 > Serverless K8s集群),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单击集群ID,然后在集群详情页面的应用列表区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情页面的左边导航栏选择监控 > 应用诊断-日志分析。
- 4. 在右侧页面, 执行以下操作:
 - i. 在搜索框中输入查询分析语句。查询分析语句由查询语句和分析语句构成,格式为查询语句|分析语句,查询分析语句语法请参见查询语法、SQL分析语法。
 - ii. 设置查询分析的时间范围。您可以设置相对时间、整点时间和自定义时间。
 - ? 说明 查询结果有1分钟以内的误差。
 - iii. 单击查询/分析, 查看查询分析结果。



6.2.11. 应用环境

本文说明如何查看应用的环境信息,包括IP地址、运行时长、进程号、JVM版本等信息。

应用限制

EDAS标准版不支持查看应用运行环境,专业版和铂金版支持查看应用运行环境。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容器服务K8s集群或Serverless K8s集群页面单

击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。

- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用详情的左边导航栏中选择**监控 > 应用环境**。 **应用环境**页面显示应用的环境信息。



参数说明

参数	说明
实例IP	应用的实例的IP地址。
运行时长	应用从启动到当前的运行时长。
进程号	应用所对应的进程的唯一标识。
JVM版本	应用的Java虚拟机版本。
ARMS探针版本	应用所安装的ARMS探针的版本。
JVM参数	应用的Java虚拟机参数配置。
JVM参数检测	根据应用所在的主机的物理内存推荐的Java虚拟机参数配置。
启动时间	应用启动的时间。
主机名	应用所在的主机的名称。

6.2.12. 应用监控设置

6.2.12.1. 自定义配置

应用监控的一些常用设置,例如调用链采样率、Agent开关、慢SQL阈值等,可直接在**自定义配置**页签上配置。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:

- 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择应用设置 > 应用监控设置。
- 4. 在应用监控设置页面,单击自定义配置页签。

配置调用链采样设置

在**调用链采样设置**区域,可以打开或关闭调用链采样开关,并设置采样率。**采样率设置**字段输入百分比的数字部分即可,例如输入*10*代表采样10%。



设置完成后,单击页签底部的保存。

配置Agent(探针)开关和日志级别

在Agent开关配置区域,可以打开或关闭探针总开关以及各插件开关,并配置日志级别。

○ 注意 探针总开关和日志级别的修改即时生效,无需重启应用。如果关闭探针总开关,则系统将无法监控您的应用,请谨慎操作。要使对各插件开关的修改生效,必须手动重启应用。



设置完成后,单击页签底部的保存。

配置阈值设置

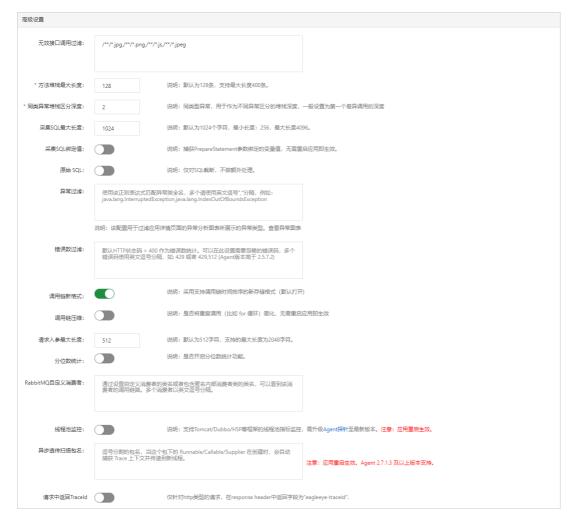
在阈值设置区域,可以设置慢SQL查询阈值、接口响应时间阈值和限流阈值。

阈值设置		
* 慢SQL查询阈值:	500	说明:单位为毫秒,默认500ms。当SQL查询的耗时大于该阈值的时候,该查询会被标记为侵SQL。
*接口响应时间阈值:	500	说明:单位为毫秒,默认500ms。当接口响应时间大于该阈值的时候,该接口会被标记为慢调用。
* 限流阈值:	100	说明:Agent端每秒最大可处理请求数,默认100条。大于该阈值的调用链,不被收集。

设置完成后,单击页签底部的保存。

配置高级设置

在高级设置区域,可以设置需过滤的接口、方法堆栈最大长度等。



- 无效接口调用过滤:输入不需要查看调用情况的接口,从而将其从接口调用页面隐去。
- 方法堆栈最大长度: 默认为128条, 最大值为400条。
- **同类异常堆栈区分深度**: 同类型异常,用于作为不同异常区分的堆栈深度,一般设置为第一个差异调用的深度。
- 采集SQL最大长度: 默认为1024个字符, 最小值为256, 最大值为4096。
- 采集SQL绑定值:捕获PrepareStatement参数绑定的变量值,无需重启应用即可生效。
- 原始SQL: 仅对SQL截断,不做额外处理。
- 异常过滤: 此处输入的异常不会显示在应用详情和异常分析页面的图表中。
- **错误数过滤**:默认情况下,大于400的状态码会计入错误数,您可以自定义大于400但不计入的HTTP状态码。
- 週用链新格式:采用支持调用链时间排序的新存储格式(默认打开)。
- 调用链压缩: 是否将重复调用(比如for循环)简化, 无需重启应用即生效。
- 请求入参最大长度: 默认为512字符, 支持的最大长度为2048字符。
- 分位数统计:是否开启分位数统计功能。
- 应用紧急事件报警: 支持针线程死锁、OOM等紧急报警。Agent版本须为2.5.8+。
- Rabbit MQ自定义消费者:通过设置自定义消费者的类名或者包含匿名内部消费者的类名,可以看到该消费者的调用链路。多个消费者以英文逗号(,)分隔。
- **线程池监控**: 支持Tomcat/Dubbo/HSF等框架的线程池指标监控,需升级Agent探针至最新版本。该设置需重启应用后生效。

设置完成后,单击页签底部的保存。

配置线程设置

在线程设置区域,可以打开或关闭线程诊断方法栈开关、线程剖析总控开关,并设置慢调用监听触发阈值。



② 说明 服务调用耗时超过慢调用监听触发阈值(默认值为1000毫秒)时才会启动监听,并一直持续 到该次调用结束或超过15秒。建议将此阈值设为调用耗时的第99百分位数。假设有100次调用,则按耗 时从小到大排序,排在第99位的耗时就是第99百分位数。

设置完成后,单击页签底部的保存。

配置内存快照设置

在**内存快照设置**区域,可以启用或停用内存快照功能。打开此开关后,出现内存泄漏时将自动转储内存(一天至多一次)。



设置完成后,单击页签底部的保存。

设置关联业务日志与TraceId

在业务日志关联设置区域,可以设置是否在应用的业务日志中关联调用链的Traceld信息。



? 说明

- 开启此开关后,会在业务日志中自动生成调用链的Traceld。
- 如果您还需要绑定Project和Logstore,实现精准定位业务异常问题。

设置完成后,单击页签底部的保存。

配置URL收敛规则

 在URL收敛设置区域,可以打开或关闭收敛功能的开关,并设置收敛阈值、收敛规则和排查规则。URL收敛是指将具有相似性的一系列URL作为一个单独的个体展示,例如将前半部分都为/service/demo/的一系列URL集中展示。收敛阈值是指要进行URL收敛的最低数量条件,例如当阈值为100时,则符合规则正则表达式的URL达到100时才会对它们进行收敛。

URL收敛设置		
收敛URL:		
收敛阈值:	100	说明: 大于此阈值即进行收敛。
收敛规则正则:		
排除规则正则:	示例: /service/(.*?)/demo。	注: 多个规则之间以英文逗号分隔。
31FP赤苏拉州JIE州J:		
		> 4 4 4 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	示例: /service/\d+/demo。	注: 多个规则之间以英文逗号分隔。

设置完成后,单击页签底部的保存。

设置业务监控

在业务监控设置区域,可以打开或关闭业务监控开关,并设置HTTP编码。



设置完成后,单击页签底部的保存。

6.2.12.2. 标签管理

您可以在应用设置页面上设置和管理应用的自定义标签。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择应用设置 > 应用监控设置。
- 4. 在应用监控设置页面,单击标签管理页签。

设置标签

- 1. 在应用标签管理区域框,单击管理应用标签。
- 2. 在弹出的**管理应用标签**对话框,单击❶图标,输入您自定义的标签名称并单击**确认**。
- 3. 在弹出的成功对话框,单击确认。
- 4. 设置完成后,单击标签管理页签底部的保存。

修改和删除标签

- 1. 在应用标签管理区域框,单击管理应用标签。
- 2. 在弹出的管理应用标签对话框,将鼠标悬浮到您想编辑或删除的标签上。
 - 如需编辑标签,则单击右侧的 图标,并在文本框内输入新的标签。
 - 如需删除标签,则单击左侧的×图标。

□ **注意** 如果在**管理应用标签**对话框中删除现有标签,则所有已经启用该标签的应用将会失去该标签。

- 3. 在管理应用标签对话框,单击确认。
- 4. 在弹出的成功对话框,单击确认。
- 5. 设置完成后,单击标签管理页签底部的保存。

6.2.12.3. 监控方法自定义

除了被探针自动发现的所有方法和接口外,如果您需要监控应用中的其他方法和接口,则可以设置自定义监控方法。

功能入口

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择应用设置 > 应用监控设置。
- 4. 在应用监控设置页面,单击监控方法自定义页签。

自定义监控方法

- 1. 在监控方法自定义页签,单击添加方法。
- 2. 在添加自定义方法对话框,设置监控方法相关参数,然后单击确认。



参数	描述
方法全名	自定义设置监控的方法的名称,在当前应用下方法名 称需要唯一。
	开启此功能,即可监控此方法,并且可在本地方法栈 中展示此方法。默认为开启状态。
开启此功能	⑦ 说明 支持动态开启或关闭此功能,并且无需重启应用。
设为业务调用入口	设置为业务调用入口后,则可通过调用链对其进行业 务查询,并且可在 接口调用 模块中展示对应接口,更 多信息,请参见 <mark>服务和接口监控</mark> 。默认为关闭状态。

自定义监控方法添加成功后,自动显示在方法列表中。

6.2.12.4. 自定义参数

查询调用链功能所支持的内置查询字段不具备业务意义,无法满足有业务场景的查询需求。您可以为应用配置自定义参数提取规则,应用监控会获取相应业务参数并追加到调用链中,此后您可以根据自定义参数来查询调用链。

添加自定义参数

1. 登录EDAS控制台。

- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏选择应用设置 > 应用监控设置。
- 4. 在应用监控设置页面,单击自定义参数页签。
- 5. 在自定义参数页签,单击添加自定义参数。
- 6. 在添加自定义参数对话框,设置自定义参数,然后单击确认。



参数	描述
参数提取规则	针对HTTP接口,支持从Parameter、Cookie、Method和Header中提取所需的业务字段,且支持设置多条参数提取规则。如需从业务请求接口 curl "http://{domain}/api/buy?brand=SIEMENS" -H "account: 123456"中提取brand和account数据,则添加以下参数提取规则: Parameter: brand Header: accoun
高级设置	
适用接口	默认情况下,参数提取规则针对所有接口,表达式为/**。如需针对特定的接口,请按实际需求设置。
调用链全采	开启后对应接口将忽略应用调用链采样率限制并进行 全量采集。默认关闭。

7.告警管理

7.1. 告警管理升级说明

EDAS告警管理全面升级,本次升级包括页面与功能的新增和变更。本文介绍新版告警管理基于旧版所做的一系列变更,帮助您了解新版与旧版的区别,快速上手新版告警管理。

② 说明 升级后,原有报警规则均可以正常运行,告警也可以正常发送。

新增

通知策略

通知策略功能可以对产生的告警事件进行分派、处理和发送。当规则被触发时,系统会以您指定的通知方式向处理人发送告警信息,以提醒处理人采取必要的解决措施。

升级策略

对于一段时间内未解决的告警,您可以选择升级通知来提醒联系人及时解决。在通知策略中添加升级策略 后,系统会以您指定的通知方式向处理人发送告警信息,以提醒处理人采取必要的解决措施。

钉群通知

新版告警管理支持在钉钉群里通过卡片方式发送告警。您还可以在钉钉群里直接进行告警屏蔽、告警认领等操作。

变更

- 报警规则页面
 - 移除原有报警规则页面。在应用监控中新增告警管理界面,您可在新增页面创建、查看、编辑和管理特定 类型的告警规则。
- 通知方式
 - 原有告警规则联系人、通知模板、恢复通知、静默时间和生效时间等功能替换为通知策略功能,现已无法编辑。若需编辑,您需在编辑页面一键迁移为通知策略后通过通知策略进行变更。
- 下线报警模板管理功能。
- 默认告警规则创建
 - 新导入应用监控的应用不再自动创建默认告警规则。

告警规则 创建应用监控告警规则 Q 请输入报警名称 报警名称 类型 所属应用 告警规则 更新时间 状态 操作 -应用紧急 默认应用紧急报警 默认应用紧急报警, 当应用出现: 内存泄漏异 ●运行中 ● | 报警历史 • srs 事件报警 常、死锁异常、进程退出情况时,发送紧急报警 15:25:38 ■ -应用紧急事件 报警 2021-07-27 默认应用紧急报警, 当应用出现: 内存泄漏异 ● | 报警历史 常、死锁异常、进程退出情况时,发送紧急报警 ■ -应用紧急事件 2021-07-27 默认应用紧急报警, 当应用出现: 内存泄漏异 默认应用紧急报警 ●运行中 ⊕ | 报警历史 报警 常、死锁异常、进程退出情况时,发送紧急报警 em 最近1分钟节点机用户使用cpu_百分比最大值大于 2021-07-23 • em 应用监控报警 ●运行中 21:22:31 报警历史 于2000 | 最近5分钟数据库调用错误次数平均值大于等于3000 | 最近5分钟数据库调用错误次数平均值大于等于1 最近5分钟数据库调用响应时间 ms平均值大干等 2021-07-211 编辑 | 启动 | 删除 | - 应用监控默认报警-数 默认应用监控报警 ●已停止 5:54:06 最近5分钟调用响应时间_ms平均值大于等于2000 |最近5分钟调用错误次数平均值大于等于1 2021-07-211 编辑 | 启动 | 删除 | ●已停止 -应用监控默认报警-异 默认应用监控报警 最近5分钟节点机用户使用cpu_百分比平均值大于 2021-07-211 编辑 | 启动 | 删除 | -应用监控默认报警-主 机监控报警 默认应用监控报警 ●已停止 等于90 | 最近5分钟节点机空闲磁盘 byte平均值 报警历史 小于等于1048576

升级后,您可以在告警管理 > 告警规则中创建、查看、编辑和管理应用监控告警规则。

升级前,旧版和新版告警管理功能均支持在告警规则中选择已创建的联系人和联系人组进行通知,并且可以设置告警的通知模板、静默时间和生效时间等。当告警触发时,会按照告警规则的相关设置发送告警。

升级后,新版告警管理新增通知策略功能,用于告警通知。在创建通知策略时,您可以指定匹配告警规则的分派规则、对告警事件的处理和分组方式、通知人、通知方式、通知模板和静默时间或升级策略等。在创建告警规则时,您只能新建通知策略或者选择已有的通知策略,无法像升级之前直接指定通知相关的内容。

建议事先按需创建特定的通知策略,然后在创建告警规则时选择已有的通知策略来对告警事件进行处理和发送。或者您还可以选择在创建告警规则时均不指定通知策略,然后在创建通知策略时通过创建模糊匹配的分派规则来对满足分派条件的告警规则进行处理和发送。

□ 注意

- 已创建的旧版告警规则中,未使用通知策略的告警规则对应的告警事件和告警发送记录在新版无法查看。您只能在告**警管理 > 告警发送历史**中,单击**返回旧版**进行查看。
- 在编辑通过旧版创建的告警规则时,通知方式相关的内容均已经无法编辑。若需编辑,则需单击编辑页面上的一**键迁移**,将当前告警规则的通知方式自动转化为对应的通知策略,然后编辑该通知策略并保存。但需注意:
 - 请详细核对自动创建的通知策略各项是否满足您的告警需求。通知策略部分新增功能均为默认值,您可以参考文档,按需进行调整。
 - 使用通知策略后,告警通知的格式和内容会发生改变。请注意将钉群设置的关键词包含 在新的告警内容中,否则钉群将无法接收告警。
 - 告警规则和通知策略是松耦合,一个告警规则产生的告警事件可以被多个通知策略匹配 到,一个通知策略也可以匹配多个告警规则产生的告警事件。

新版告警管理移除了报警模板功能。新版中告警规则支持同时选择多个应用进行监控,建议使用此方式代替原有的报警模板功能。

升级前,在创建或导入应用监控应用和前端监控站点时,EDAS会自动根据模板分别创建默认的告警规则。 升级后,在创建或导入应用监控应用和前端监控站点时,将不再自动创建默认告警规则。但已创建的默认

升级后,在创建或导入应用监控应用和前端监控站点时,将不再自动创建默认告警规则,但已创建的默认告警规则仍可正常运行。

默认告警规则创建变更

7.2. 名词解释

本文列出了EDAS告警管理的主要术语。

B|C|G|H||L|R|S|T|Y

В

补零、补一、补Null

报警数据修订策略可选择为补零、补一或补Null(默认)。此功能一般用于无数据、复合指标和环比同比等异常的数据修复。

● 补零: 将被判断的数值修复为0。

• 补一: 将被判断的数值修复为1。

● 补Null: 不会触发报警。

应用场景:

● 异常情况一:无数据

用户A想利用报警功能监控页面访问量。创建报警时,选择前端监控报警,设置报警规则为N=5时页面访问量的总和小于等于10则报警。若该页面一直没有被访问,则没有数据上报,不会发送报警。为解决此类问题,可将报警数据修订策略勾选为补零,将没有收到数据视为收到零条数据,符合报警规则,即可发送报警。

● 异常情况二:复合指标异常

用户B想利用报警功能监控商品的实时单价。创建报警时,选择自定义监控报警,设置变量a的数据集为当前总价,变量b的数据集为当前商品总数,报警规则为N=3时(当前总价)/(当前商品总数)的最小值小于等于10则报警。若当前商品总数为0时,复合指标(当前总价)/(当前商品总数)的值不存在,则不会发送报警。为解决此类问题,可将报警数据修订策略勾选为补零,将复合指标(当前总价)/(当前商品总数)的值视为0,符合报警规则,即可发送报警。

• 异常情况三: 指标环比、同比异常

用户C想利用报警功能监控节点机用户使用CPU百分比。创建报警时,选择应用监控,设置报警规则为N=3时节点机用户使用CPU百分比的平均值环比下降100%则报警。若最近N分钟用户的CPU故障无法工作,即α无法获取,导致环比结果不存在,则不会发送报警。为解决此类问题,可将报警数据修订策略勾选为补一,将环比结果视为下降100%,符合报警规则,即可发送报警。

C

重复升级规则

如果升级策略判断过所有升级规则后,仍未解决告警,则将重新开始升级判断。默认为重复1次,您也可以将其调整为9以内的其他次数。

[回到顶部]

G

告警

指需要通知联系人去解决的告警。只有通过通知策略触发的需要解决的事件,才会创建告警。

告警管理

指整个告警管理功能。

告警卡片

钉钉群的机器人通过卡片的形式向钉钉群发送告警。联系人可以通过卡片查看、处理、解决告警。您可

281

以在控制台的钉群中配置钉钉群的机器人。

告警自动解决

当告警的事件都已恢复时,告警会自动设置为已解决。

故障

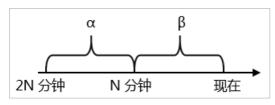
业务的规则会引起监控数据异常,监控工具检测到异常后会生成事件。

[回到顶部]

Н

环比上升 / 下降%

若β为最近N分钟的数据(可选择为平均值、总和、最大值和最小值), α为前2N分钟到前N分钟的数据, 环比为β与α做比较。



L

联系人

处理告警的运维人员,联系人可以通过钉钉查看、处理和解决告警。

[回到顶部]

R

认领告警

如果联系人想将自己设置为尚未解决的告警的处理人,则可以认领该告警。

[回到顶部]

S

升级超时

升级超时后,告警管理会将未确认或未解决的告警通知升级策略的联系人。默认升级超时设置为10分钟,您也可以将其设置在90分钟内的任一时间。

升级策略

在告警升级超时,告警管理会依次通知升级规则定义的不同批次的联系人。升级策略可以添加到特定的 通知策略中。

事件分组

事件管理将多个事件汇总到单一告警中,用来减少需要联系人处理的告警。同时汇总关键信息,降低通知疲劳。

事件静默

不重要的事件可以将其静默,被静默的事件没有对应的告警。

事件自动恢复

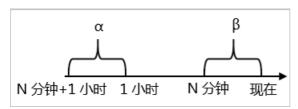
您可以设置在指定的时间后自动恢复事件。此时间过后,事件将自动被认定为已解决。事件自动恢复时间默认为5分钟。

[回到顶部]

Т

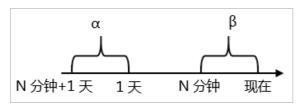
同比上升 / 下降% (与上小时)

若β为最近N分钟的数据(可选择为平均值、总和、最大值和最小值), α 为上小时最近N分钟的数据,与上小时同比为 β 与 α 做比较。



同比上升 / 下降% (与昨日)

若β为最近N分钟的数据(可选择为平均值、总和、最大值和最小值), α 为昨日同一时刻最近N分钟的数据,与昨日同比为 β 与 α 做比较。



通知

触发告警后,告警管理会立即发送通知给联系人。通知消息中包含触发的告警对象和告警标题。通知将通过短信、邮件、微信、钉群、WebHook等方式发送。

通知策略

通知策略确定了如何将事件汇总到单一告警中,以及如何将分配的告警通知发送给联系人。您可以通过设置电话,短信,邮件或钉群等方式,将需要解决或需要联系人知晓的告警1分钟内通知到他。

[回到顶部]

Υ

用户

指阿里云用户和RAM用户,用户可以通过EDAS控制台实现以下操作:

- 创建、编辑、修改通知策略。
- 创建、编辑、修改升级策略。
- 查看、处理、解决告警。

[回到顶部]

7.3. 联系人管理

7.3.1. 联系人

通知策略的分派规则被触发时可以向您指定的联系人发送通知。联系人支持通过电话、短信、邮件和钉钉等 方式接收告警通知。

创建联系人

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 联系人。
- 4. 在联系人页签上, 单击右上角的新建联系人。
- 5. 在**新建联系人**对话框中输入联系人姓名,根据实际需求输入联系人手机号码、邮箱或钉钉机器人地址, 设置是否接收系统通知,然后单击**确认**。

□ 注意

- 手机号码、邮箱和钉钉机器人至少填写一项。每个手机号码或邮箱只能用于一个联系人。
- 最多可添加100个联系人。
- 获取钉钉机器人地址的方法,请参见获取钉钉机器人Webhook地址。
- 仅验证过的手机号码可以在通知策略中使用电话的通知方式,验证手机号的操作,请参见<mark>验证手机号。</mark>

验证手机号

仅验证过的手机号码可以在通知策略中使用电话的通知方式。

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 联系人。
- 4. 在联系人页签上,选择为一个或多个联系人验证手机号。
 - 如需为单个联系人验证手机号: 单击未验证手机号右侧的未验证。
 - 如需为多个联系人批量验证手机号:选择需要验证手机号的联系人,然后单击**批量验证**。 系统将会给各联系人发送验证手机号短信。
- 5. 使用浏览器打开短信中的链接。

【阿里云】【阿里云-告警运维中 心】请验证手机号,点击此链接前去 验证,

https://cloud-ops.aliyun.com/phone
verify?secretKey=

30分钟内有效 本消息由阿里云用户:

在ARMS报警中

配置。

6. 在验证页面确认手机号信息,然后根据信息内容进行验证。

管理联系人

创建联系人后,您可以在**联系人**页签查询、编辑或删除联系人:

- 如需搜索联系人,请在**联系人**页签上,从搜索下拉框中选择**姓名、手机号或Email**,然后在搜索框中输入 联系人姓名、手机号码或邮箱的全部或部分字符,并单击图标。**Q**
- 如需编辑联系人,请单击联系人右侧操作列中的编辑,在编辑联系人对话框中编辑信息,然后单击确 **认**。
- 如需删除单个联系人,请单击联系人右侧操作列中的删除,然后在提示对话框中单击确认。
- 如需删除多个联系人,请选择目标联系人,单击**批量删除**,然后在提示对话框中单击**确认**。

7.3.2. 联系人组

创建通知策略时,您可以将联系人组指定为通知对象,当通知策略的分派规则被触发时,EDAS告警管理会向该联系人组中的联系人通过电话、短信、邮件和钉钉等方式发送告警通知。

前提条件

已创建联系人,具体操作,请参见创建联系人。

创建联系人组

1. 登录EDAS控制台。

- 2. 在**应用列表**页面集群类型的下拉列表中,选择EDAS自带K8s集群或容器服务/Serverless K8s集群,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 联系人。
- 4. 在联系人页面单击联系人组页签,并单击右上角的新建联系组。
- 5. 在新建联系组对话框中输入组名,选择报警联系人,并单击确认。

管理联系人组

创建联系人组后,您可以在**联系人组**页签查询、编辑或删除联系人:

- 如需搜索联系组,请在联系人组页签的搜索框中输入联系人组名称的全部或部分字符,并单击图标。Q
- 如需编辑联系组,请单击联系人组右侧的∠图标,并在编辑联系组对话框中编辑相关信息,然后单击确认。
 - ② 说明 仅支持编辑当前用户创建的联系人组。
- 如需查看联系组中的联系人信息,请单击联系人组左侧的 ✓ 图标展开联系组。
 - ② 说明 您可以在展开模式下移除联系组中当前用户创建的联系人。如需移除,请单击目标联系 人操作列中的移除。
- 如需删除联系组,请单击联系人组右侧的 ※ 图标,然后在弹出的提示对话框中单击确认。

□ 注意

- 删除联系组之前,请确保目标联系组没有添加至通知策略中,否则可能导致告警通知无法发 详。
- 仅支持删除当前用户创建的联系人组。

7.3.3. 钉钉机器人

在EDAS告警管理中创建钉钉机器人后,您可以在通知策略中指定对应的钉钉群用于接收告警。当通知策略的分派规则被触发时,系统会自动向您指定的钉钉群发送告警通知。钉钉群收到通知后,您可以在钉钉中对告警进行管理。

前提条件

已在钉钉客户端创建用于接收告警通知的钉钉群。

在EDAS控制台中创建告警通知钉群

- 1. 在钉钉群中创建自定义机器人并获取机器人Webhook地址,具体操作,请参见获取钉钉机器人Webhook地址。
- 2. 登录EDAS控制台。
- 3. 在**应用列表**页面顶部菜单栏选择地域,在页面中选择微服务空间,在**集群类型**下拉列表中选择**ECS集** 群,然后单击目标应用名称。
- 4. 在左侧导航栏中选择告警管理 > 联系人。
- 5. 在联系人页面单击IM机器人页签。

- 6. 在IM机器人页签上,单击右上角的新建机器人。
- 7. 在新建机器人对话框中设置以下参数,然后单击确认。

(二) 注意 在通知策略中需要选择**通知方式**为钉钉才能在钉钉群中接收告警,具体操作,请参见通知策略。

参数	说明
名称	自定义钉钉机器人的名称。
类型	选择 类型 为 钉钉 。
机器人地址	输入钉钉机器人的Webhook地址。
机器人是否发送每日统计	选中后,需要输入每日统计信息发送的时间点,使用英文逗号(,)分隔多个发送时间点,时间点格式为 HH:SS 。EDAS告警管理将在设置的时间点发送今日产生告警的总数、解决数和待解决数。

在钉钉群中管理告警

在钉钉群中收到告警通知后,您可以在钉钉群里查看并管理告警。更多信息,请参见在告警通知群中处理告警。

相关操作

创建机器人后,您可以在IM机器人页签查询、编辑或删除机器人:

- 如需搜索指定的机器人,在搜索框中输入机器人名称的关键字,然后单击Q图标。
- 如需编辑机器人,单击机器人右侧操作列中的编辑,在编辑机器人对话框中修改信息,然后单击确认。
- 如需删除机器人,单击机器人右侧操作列中的删除,并在提示对话框中单击确认。

7.3.4. 企业微信机器人

在EDAS告警管理中创建企业微信机器人后,您可以在通知策略中指定对应的企业微信群用于接收告警。当通知策略的分派规则被触发时,系统会自动向您指定的企业微信群发送告警通知。企业微信群收到通知后,您可以在企业微信群中对告警进行管理。

前提条件

已在企业微信中创建用于接收告警通知的企业微信群。

步骤一: 获取机器人Webhook地址

- 1. 打开并登录企业微信。
- 2. 单击企业微信群右上角的 ... 图标, 然后单击添加群机器人。
- 3. 在添加机器人页面单击新创建一个机器人。
- 4. 在创建机器人页面自定义机器人名称,然后单击添加机器人。
- 5. 单击复制地址,保存企业微信机器人的Webhook地址。



步骤二:在EDAS控制台中创建机器人

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面顶部菜单栏选择地域,在页面中选择微服务空间,在**集群类型**下拉列表中选择**ECS集 群**,然后单击目标应用名称。
- 3. 在左侧导航栏中选择告警管理 > 联系人。
- 4. 在联系人页面单击IM机器人页签。
- 5. 在IM机器人页签上,单击右上角的新建机器人。
- 6. 在新建机器人对话框中设置以下参数,然后单击确认。

参数	说明
名称	自定义企业微信机器人的名称。
类型	选择 类型 为 企业微信 。
机器人地址	输入企业微信机器人的Webhook地址。
机器人是否发送每日统计	选中后,需要输入每日统计信息发送的时间点,使用英文逗号(,)分隔多个发送时间点,时间点格式为 HH:SS 。EDAS告警管理将在设置的时间点发送今日产生告警的总数、解决数和待解决数。

在企业微信群中管理告警

在企业微信群中收到告警通知后,您可以在企业微信群里查看并管理告警。更多信息,请参见在告警通知群中处理告警。

相关操作

创建机器人后,您可以在IM机器人页签查询、编辑或删除机器人:

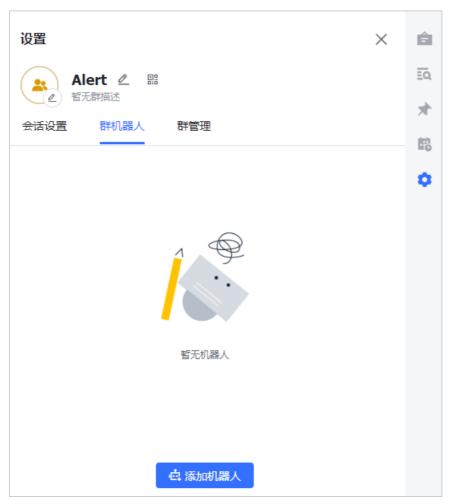
- 如需搜索指定的机器人,在搜索框中输入机器人名称的关键字,然后单击Q图标。
- 如需编辑机器人,单击机器人右侧操作列中的编辑,在编辑机器人对话框中修改信息,然后单击确认。
- 如需删除机器人,单击机器人右侧操作列中的删除,并在提示对话框中单击**确认**。

7.3.5. 通过Webhook自定义告警通知人

创建通知策略时,您可以将告警通知发送到自定义的Webhook地址中。ARMS告警管理支持对飞书、微信、钉钉等群组发送Webhook告警,本文以飞书为例,介绍如何创建Webhook告警。

步骤一: 获取Webhook地址

- 1. 打开并登录飞书。
- 2. 单击+图标, 然后单击创建群组, 新建一个用于发送告警的群组。
- 3. 单击群组设置图标,然后单击群机器人页签。
- 4. 在群机器人页签单击添加机器人。



5. 在添加机器人面板选择Custom Bot。



6. 在配置页设置显示名称和描述,然后单击添加。



7. 在添加情况区域单击复制链接,然后单击完成。



步骤二: 创建Webhook联系人

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或容**器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 联系人。
- 4. 在联系人页签上,单击右上角的新建webhook。
- 5. 在创建Webhook对话框中输入配置信息。

基本参数描述如下所示。

参数	说明
Webhook名称	必填,自定义Webhook名称。
Post和Get	必填,设置请求方法。URL不可超过100个字符。 此例中选择Post,并将 <mark>步骤一:获取Webhook地址</mark> 中保存的Webhook地址粘贴至右侧文本 框。
Header和 Param	非必填,设置请求头,不可超过200个字符。 单击 +添加 ,可以添加其他Header信息或 Param信息。默认请求头为Content-Type: text/plain; charset=UTF-8,Header和Param个 数总数不能超过6个。 此例中设置以下两个Header: • Arms-Content-Type: json • Content-Type: application/json

```
参数
             说明
             告警触发时发送的通知模板,非必填,在Post方法下出现,可使用$content占位符输出通知
             内容,不可超过500个字符。通知模板如下:
              "告警名称":"{{ .commonLabels.alertname }}{{if
              .commonLabels.clustername }}",
              "集群名称":"{{ .commonLabels.clustername }} {{ end }}{{if eq "app"
              .commonLabels._aliyun_arms_involvedObject_kind }}",
              "应用名称":"{{ .commonLabels. aliyun arms involvedObject name }} {{
              end }}",
              "通知策略":"{{ .dispatchRuleName }}",
              "告警内容":"{{ for .alerts }} {{ .annotations.message }} {{ end }}"
通知模板
             此处以飞书为例可以设置如下文本格式:
               "msg_type": "text",
               "content": {
                 "text": "报警名称: {{ .commonLabels.alertname }}\n报警时间: {{
              .startTime }}\n报警内容: {{ for .alerts }} {{ .annotations.message }}
              {{if .generatorURL }} 详情链接: {{.generatorURL}} {{ end }} \n {{ end
              } } "
              }
```

```
参数
             说明
             告警恢复时发送的通知模板,非必填,在Post方法下出现,可使用$content占位符输出通知
             内容,不可超过500个字符。恢复模板如下:
              "告警名称":"{{ .commonLabels.alertname }}{{if
               .commonLabels.clustername }}",
              "集群名称":"{{ .commonLabels.clustername }} {{ end }}{{if eq "app"
               .commonLabels._aliyun_arms_involvedObject_kind }}",
              "应用名称":"{{ .commonLabels. aliyun arms involvedObject name }} {{
              end }}",
              "通知策略":"{{ .dispatchRuleName }}",
              "恢复时间":"{{ .endTime }}",
              "告警内容":"{{ for .alerts }} {{ .annotations.message }} {{ end }}"
             此处以飞书为例可以设置如下文本格式:
恢复模板
                "msg type": "text",
                "content": {
                  "text": "报警名称: {{ .commonLabels.alertname }}\n恢复时间: {{
               .endTime }}\n报警内容: {{ for .alerts }} {{ .annotations.message }}
              {{if .generatorURL }} 详情链接: {{.generatorURL}} {{ end }} \n {{ end
               }
```

- 6. (可选)单击测试,验证配置是否成功。
- 7. 单击创建。

步骤三:设置通知策略

在通知策略中需要选择通知方式为WebHook才能在自定义的Webhook中接收告警。

- 1. 新建或编辑通知策略,选择**通知人**为Webhook联系人,在**通知方式**区域选中**WebHook**。具体操作,请参见通知策略。
- ② 说明 Webhook告警的超时时间为5秒,如果发出请求后5秒内没有返回,即没有收到告警信息,则表示发送失败。

7.3.6. 获取钉钉机器人Webhook地址

设置钉群或在联系人中设置钉钉机器人时需要先在钉钉群中获取自定义机器人Webhook地址。本文介绍如何获取钉钉机器人Webhook地址。

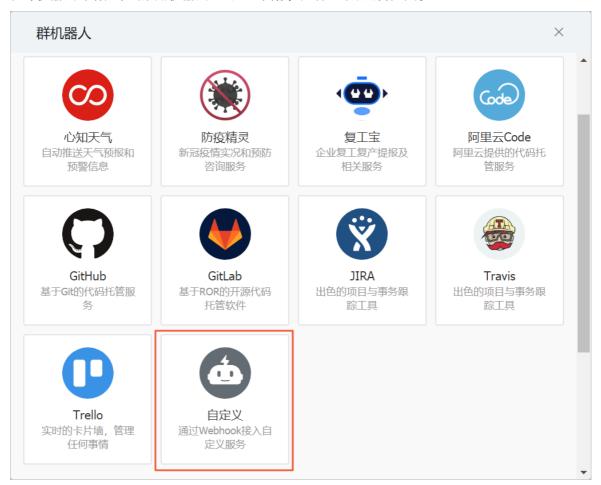
前提条件

已创建用于接收告警通知的钉钉群。

操作步骤

请按照以下步骤在钉钉群中添加自定义钉钉机器人并获取Webhook地址。

- 1. 在PC版钉钉上打开您想要添加报警机器人的钉钉群,并单击右上角的群设置图标。
- 2. 在群设置面板中单击智能群助手。
- 3. 在智能群助手面板单击添加机器人。
- 4. 在群机器人对话框单击添加机器人区域的+图标,然后选择添加自定义。



- 5. 在机器人详情对话框单击添加。
- 6. 在添加机器人对话框中执行以下操作。



- i. 设置机器人头像和名字。
- ii. 安全设置选中自定义关键词,设置关键词为报警和分派。
- iii. 选中我已阅读并同意《自定义机器人服务及免责条款》。
- iv. 单击完成。
- ② 说明 更多关于钉钉机器人的操作,请参见自定义机器人接入。
- 7. 在添加机器人对话框中复制生成的机器人Webhook地址,然后单击完成。



7.4. 通知策略

通过设置通知策略,您可以制定针对告警事件的分派规则。当规则被触发时,系统会以您指定的通知方式向处理人发送告警信息,以提醒处理人采取必要的问题解决措施。

前提条件

创建联系人,具体操作,请参见创建联系人。

新建通知策略

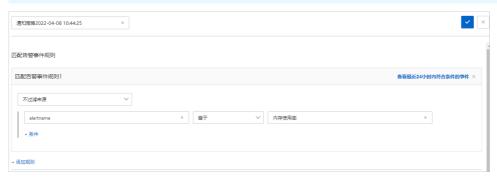
- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 通知策略。
- 4. 在通知策略页面右上角单击创建通知策略。
- 5. 在右侧区域顶部文本框中输入通知策略名称。
- 6. 在分派规则区域的进行以下操作。
 - i. 选择告警过滤来源或不过滤来源。
 - 指定过滤来源:通知策略会针对指定来源的告警事件进行分派规则过滤并发送通知。
 - 不过滤来源:通知策略会针对所有告警事件进行分派规则过滤并发送通知。

- ii. 设置分派条件表达式,您可以自定义标签或选择已有的标签。例如:alert name等于PodRest art。 已有的标签包括以下三种:
 - 告警规则表达式指标中携带的标签。
 - Promet heus监控通过告警规则创建的标签。
 - EDAS系统自带的默认标签,默认标签说明如下。

标签	说明
alertname	告警名称,格式为:告警名称_集群名称。
_aliyun_arms_alert_level	告警等级。
_aliyun_arms_alert_type	告警类型。
_aliyun_arms_alert_rule_id	告警规则对应的ID。
_aliyun_arms_region_id	地域ID。
_aliyun_arms_userid	用户ID。
_aliyun_arms_involvedObject_type	关联对象子类型,如 ManagedKubernetes,ServerlessKubernetes。
_aliyun_arms_involvedObject_kind	关联对象分类,如app,cluster。
_aliyun_arms_involvedObject_id	关联对象ID。
_aliyun_arms_involvedObject_name	关联对象名称。

? 说明

- 如果需同时满足多个分派条件才告警,则单击**+条件**编辑第二条分派条件。
- 如果需满足任意一个分派条件就告警,则单击**+添加规则**编辑第二条分派条件。

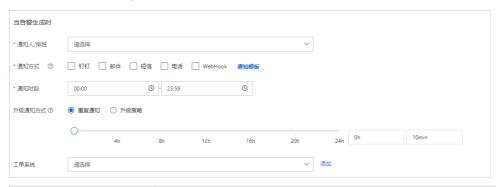


7. 在事件处理区域,设置以下参数。



参数	说明	
	生成告警:将监测到的告警发送给联系人。不告警(丢弃事件):监测到告警后不发送给联系人。	
处理方式	② 说明 如果同时设置了两个相同的通知策略且处理方式分别设置为 <i>生成告警和不告警(丢弃事件)</i> ,则 <i>不告警(丢弃事件)</i> 的优先级高于 <i>生成告警</i> ,即不会给联系人发送告警信息。	
高级配置		
事件分组	不分组:所有告警会以一条信息发送给处理人。指定相同字段内容的事件分到一个组:设置分组字段,相同字段的告警内容会分别通过独立信息发送给处理人。	
告警自动恢复	当告警下面全部事件都恢复时,告警状态是否自动恢复为已解决。当告警恢复时, 系统将会发送通知给处理人。	

8. 在当告警生成时区域,设置以下参数。



参数	说明
通知人	通知人支持设置联系人、联系人组和钉群。联系人创建方法,请参见 <mark>创建联系人</mark> 。

参数	说明	
通知方式	通知方式支持钉钉、邮件、WebHook、短信和电话,可以同时选择多种方式。 ② 说明 未验证手机号的联系人无法使用电话通知方式。验证手机号的操作,请参见验证手机号。 单击通知模板,可以在通知模板对话框中设置邮件、短信和电话的通知信息格式。	
通知时段	告警会在设置的通知时段内重复发送告警通知。	
升级通知方式	对于长期未解决的告警可以选择重复通知或升级通知来提醒联系人及时解决。 重复通知:选择重复通知后,需要设置告警发送频率。所有告警会在指定通知时段内以设置的告警发送频率循环发送告警信息直至告警恢复。升级策略:选择升级策略后,需要选择已有的升级策略或添加一个新的升级策略。具体操作,请参见升级策略。 ② 说明 单击详情,可以查看当前选中的升级策略的详细信息。	
工单系统	选择告警需要推送到的工单系统。工单系统的集成,请参见 <mark>通过Jira账号信息集成Jira工单系统</mark> 。	

9. 设置完成后,单击右上角的 🗸 图标。

管理通知策略

通知策略新建完成后会在通知策略列表中显示,您可以在通知策略列表中执行以下操作:

- 单击通知策略名称右侧的更多图标,选择编辑、停用、启用、复制或删除可以管理该通知策略。
- 单击目标通知策略,在右侧页面右上角选择相应的图标,可以编辑、刷新、复制和删除该通知策略。

7.5. 升级策略

对于长期未解决的告警,可以选择升级通知来提醒联系人及时解决。在通知策略中添加升级策略后,系统会以您指定的通知方式向处理人发送告警信息,以提醒处理人采取必要的问题解决措施。

前提条件

创建联系人,具体操作,请参见创建联系人。

新建升级策略

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 通知策略。
- 4. 在通知策略列表区域选择已创建的策略,单击。图标,并单击编辑按钮。

- 5. 在当告警生成时区域的升级通知方式选择升级策略,并单击添加。
- 6. 在弹出的对话框顶部文本框中输入升级策略名称。
- 7. 在升级规则区域,设置通知条件,即当告警在一段时间内未被认领或未被解决时发送告警通知。例如: 当告警10分钟未认领时发送升级通知。
 - ② 说明 单击+添加规则,可以新增一条升级规则。

参数	说明	
通知人	通知人支持设置联系人、联系人组和钉群。创建方法,请参见 <mark>创建联系人</mark> 。	
	通知方式支持钉钉、邮件、WebHook、短信和电话,可以同时选择多种方式。	
通知方式	⑦ 说明 未验证手机号的联系人无法使用电话通知方式。验证手机号的操作,请参见验证手机号。	
通知时段	满足通知条件的告警会在设置的通知时段内重复发送告警通知。	
重复发送告警通知的次数。当告警不满足通知条件后,告警通知将不再发		
重复次数	? 说明 至少发送一次告警,即设置为0时也会发送一次告警。	

8. 设置完成后,单击右上角的 🗸 图标。

7.6. 应用监控告警规则(新版)

通过创建应用监控告警规则,您可以制定针对特定应用监控的告警规则。当告警规则被触发时,系统会以您 指定的通知方式向告警联系人或钉群发送告警信息,以提醒您采取必要的解决措施。

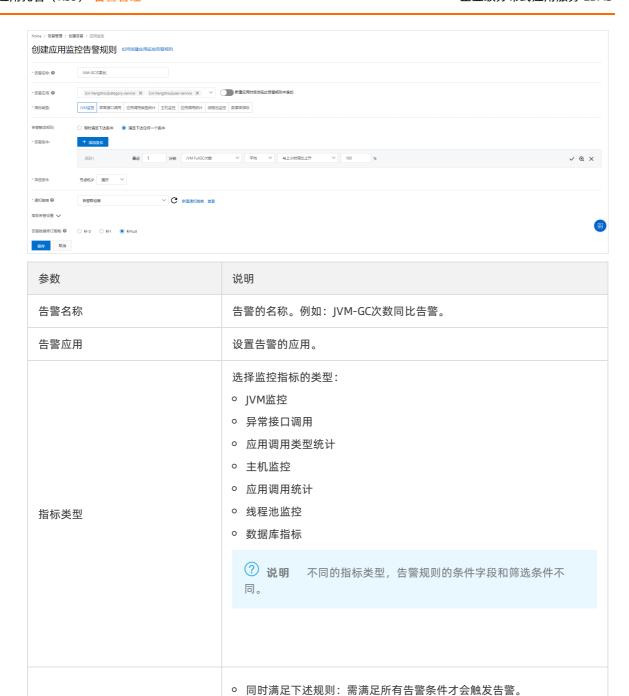
前提条件

已接入应用监控,请参见应用监控接入概述。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 告警规则。
- 4. 在告警规则页面的右上角单击创建应用监控告警规则。
- 5. 在创建应用监控告警规则页面输入所有必填信息,完成后单击保存。

告警触发规则



300 > 文档版本: 20220708

。 满足下述一条规则: 满足任意一条告警条件就会触发告警。

参数	说明
告警条件	单击+添加条件,设置告警规则表达式。例如:最近5分钟JVM FullGC次数平均与上小时同比上升100%。然后单击右侧✔图标。 ② 说明 ○ 单击规则右侧 图标,可以预览应用在当前规则下的指标走势图。 ○ 单击规则右侧 X图标,可以删除该告警规则表达式。 ○ 若需设置多条告警规则,单击+添加条件,即可编辑第二条告警规则。
筛选条件	告警指标的维度: 无: 告警内容中透出这个维度所有数值的和。 =: 告警中只透出当前设置维度的内容。 遍历: 会在告警内容中透出实际触发告警的维度内容。
通知策略	 不指定通知规则:告警被触发时不会发送告警,仅当通知策略的分派规则被触发时才会发送告警。 指定通知规则发送告警:告警被触发时,ARMS通过指定通知策略的通知方式发送告警信息。您可以选择已有的通知策略,也可以新建一个通知策略。更多信息,请参见通知策略。 说明 单击查看,可以查看选中的通知策略详情。
高级告警设置	
告警数据修订策略	用于无数据、复合指标和环比同比等异常数据的修复。当告警指标没有达到设置的条件时,告警数据补0、补1或不补充数据。

管理告警

创建的应用监控告警规则在**告警规则**页面上,您可以对告警规则执行启动、停止、编辑、删除、查看告警详情等操作。

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或容**器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏中选择告警管理 > 告警规则。

- 4. (可选)在告警规则页面的搜索框中输入告警名称,并单击搜索图标。
 - ? 说明 您可以输入告警名称的一部分内容进行模糊搜索。
- 5. 在搜索结果列表的操作列中,按需对目标告警规则采取以下操作:



- 如需编辑告警规则,请单击**编辑**,在**编辑告警**页面中编辑告警规则,并单击**保存**。
- 如需删除告警规则,请单击**删除**,并在提示对话框中单击**确认**。
- 如需启动已停止的告警规则,请单击**启动**,并在**提示**对话框中单击**确认**。
- 如需停止已启动的告警规则,请单击**停止**,并在**提示**对话框中单击**确认**。
- 如需查看告警事件历史和告警发送历史,请单击告警历史,在告警事件历史和告警发送历史页面上 查看相关记录。

相关文档

- 查看告警发送历史
- 查看告警事件历史

7.7. 在钉钉群中处理告警

在钉钉群中收到告警通知后,您可以查看、认领和关闭相关告警。本文介绍如何在钉钉群中查看并处理告警。

查看告警卡片

在钉钉群中收到告警通知后,您可以通过以下操作查看告警卡片。

1. 在钉钉群中单击告警卡片中的查看详情。



- 2. 在钉钉群中首次查看告警信息时会弹出绑定手机号面板,输入手机号,并单击发送验证码。
 - ② 说明 如果输入的手机号和已有联系人的手机号相同,那么可以将联系人和钉钉绑定。绑定后系统发送告警到钉钉群时会@该联系人。具体操作,请参见企业微信机器人。



3. 收到验证码后输入验证码,然后单击确定,进入告警详情面板。

告警详情面板

在告警详情面板可以查看当前告警的详细信息。

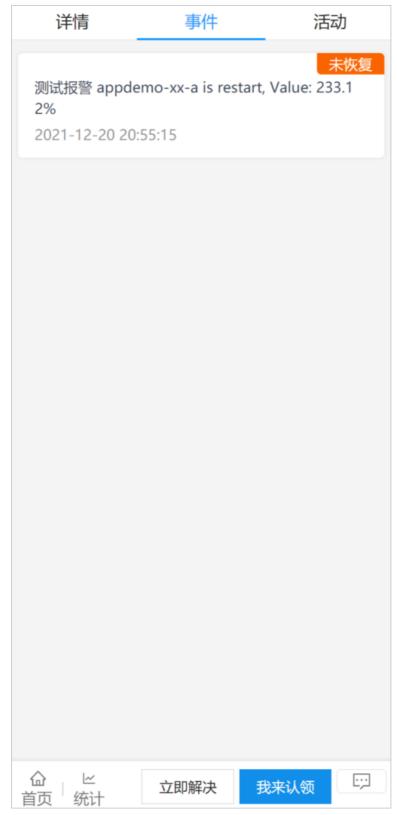
● 在详情页签,可以查看告警名称、告警等级、告警状态、告警内容、告警对象、处理人和通知人。



在详情页签,您可以执行以下操作:

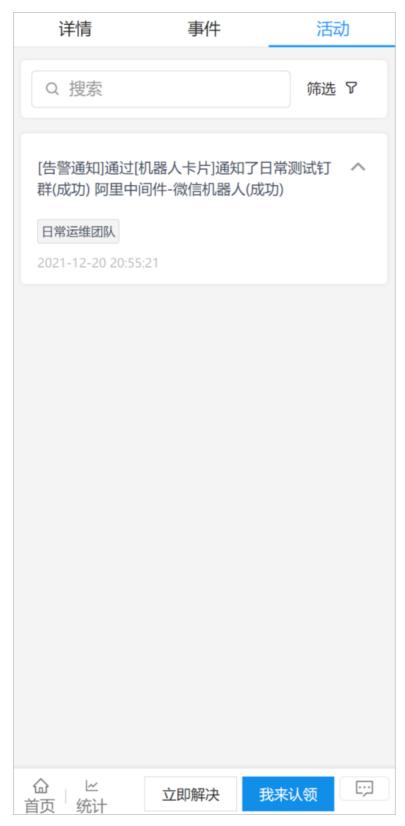
对于待认领状态的告警,单击右上角的**设置**,可以指定钉钉群里的联系人为告警处理人或修改告警等级。

● 在**事件**页签,可以查看告警关联的告警事件及状态。



在**事件**页签,您可以执行以下操作: 单击目标事件,查看事件的基本信息、监控数据和扩展字段。

• 在活动页签,可以查看告警的通知和处理记录。



在活动页签,您可以执行以下操作:

单击**筛选**,选择筛选类型为日志内容或日志对象,然后在文本框中输入关键字后单击回车,可以筛选活动记录。

在告警详情面板,您可以执行以下操作:

- 对于待认领状态的告警,单击**我来认领**,可以将自己设置为告警处理人。
- 对于待认领或处理中状态的告警,单击**立即解决**,可以将告警状态设置为已解决。
- 单击告警详情面板左下角的统计,可以进入统计面板。更多信息,请参见统计面板。
- 单击告警详情面板左下角的**首页**,可以进入首页面板。更多信息,请参见<mark>首页面板</mark>。

统计面板

统计面板显示了钉钉群里所有告警的统计信息。包括今日告警数量、近7天告警趋势图、近7天告警平均解决时间趋势图和今日告警处理人统计表。



在统计面板, 您可以执行以下操作:

- 在近7天告警趋势图和近7天告警平均解决时间趋势图上单击目标时间点,可以查看目标时间点具体的告警数和告警处理时间。
- 单击**返回首页**,可以进入首页面板。更多信息,请参见<mark>首页面板</mark>。

首页面板

可以在首页面板查看钉钉群里所有的告警信息。



在首页面板,您可以执行以下操作:

• 在面板顶部文本框中输入告警全部或部分名称,可以搜索告警。

- 在未解决、我的和全部页签,可以分别查看未解决的告警、我认领的告警和全部告警。
- 单击目标告警,可以进入目标告警的详情面板,查看目标告警的详细信息。
- 单击**告警统计**,可以进入统计面板。更多信息,请参见统计面板。

7.8. 查看告警发送历史

通过告警发送历史页面,您可以筛选并查看根据通知策略分派的告警内容,并管理告警。

前提条件

EDAS告警管理功能仅对2021年04月30日0点之后开通EDAS的阿里云账户支持。

功能入口

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面集**群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏选择告警管理 > 告警发送历史。

告警

告警页面显示了告警的等级、名称、处理人、创建时间和状态。



在告警页面,您可以执行以下操作:

● 设置筛选字段,然后单击**搜索**,查看对应的告警发送历史。

? 说明 打开更多开关,可以设置更多的过滤筛选字段。

字段	说明
告警名称	创建的告警规则的名称。
告警状态	告警目前的处理状态,共有以下3种状态: 待认领 处理中 已解决

字段	说明
告警等级	告警的严重级别。 告警级别有以下几种: 默认 P1 P2 P3 P4 P5
处理人	告警的处理人。
更多	
通知策略	告警对应的通知策略。
创建时间	告警产生的时间段。

- 单击告警名称,可以查看目标告警的详细信息。更多信息,请参见<mark>告警详情</mark>。
- 对于未解决的告警,可以认领、解决、指定告警处理人或修改告警等级。具体操作,请参见<u>处理告警</u>。

告警详情

告警详情页面显示了告警等级、发送信息、状态,以及告警基本信息、关联事件和活动记录。



在告警详情页面,您可以执行以下操作:

对于未解决的告警,可以认领、解决、指定告警处理人或修改告警等级。具体操作,请参见处理告警。

在详情、事件和活动页签可以分别查看以下信息:

- **详情**页签显示了告警创建时间、告警对象、处理人和通知人。
- **事件**页签显示了告警关联事件、事件创建时间和事件状态。单击事件名称,可以查看目标事件的详细信息。更多信息,请参见事件详情。

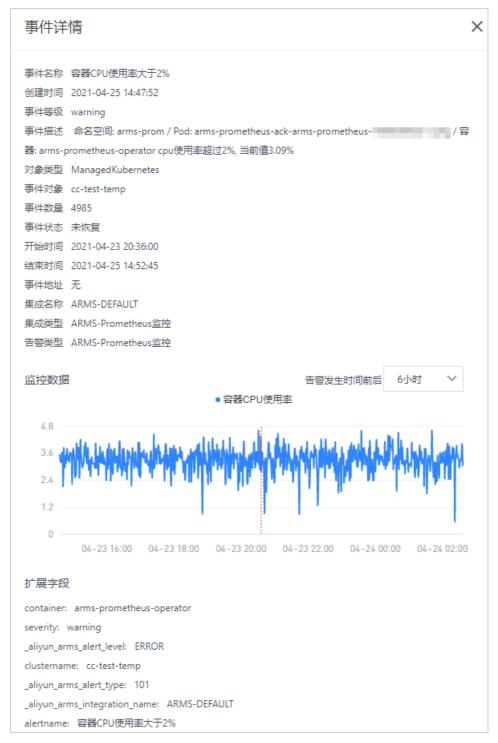


• 活动页签显示了告警的活动记录。



事件详情

事件详情面板显示了事件的基本信息、监控数据和扩展字段。



在事件详情面板的监控数据区域,您可以执行以下操作:

- 在**监控数据**区域右上角,设置数据显示的时间段为告警发生时间前后的6小时、12小时或1天。
- 使用光标选中一段时间,可以查看指定时间段的监控数据。单击重置,可以将曲线图恢复。

处理告警

在控制台中,对于未解决的告警,可以认领、解决、指定告警处理人或修改告警等级。

1. 绑定手机号。

i. 在钉钉群中单击告警卡片中的查看详情。



ii. 在钉钉群中首次查看告警信息时会弹出**绑定手机号**面板,输入手机号,并单击**发送验证码**。



- iii. 收到验证码后输入验证码,然后单击确定。
- 2. 登录EDAS控制台。
- 3. 在左侧导航栏选择告警管理 > 告警发送历史。
- 4. 在告警页面右上角的选择操作人下拉框选择操作人。

? 说明

- 。 操作人的名称为钉钉昵称。
- 由于一个阿里云账号可能存在多人同时使用的情况,所以需要通过选择操作人来区分实际每次操作告警的人员。
- 5. 在目标告警右侧,或者单击目标告警,在告警详情页面:
 - 单击**认领**,可以将自己设置为当前告警的处理人。
 - 单击解决,可以关闭目标告警。
 - 单击更多图标,然后单击指定告警处理人,可以指定钉钉群里的联系人为告警处理人。

7.9. 查看告警事件历史

通过事件列表页面,您可以筛选并查看所有告警事件。

前提条件

EDAS告警管理功能仅对2021年04月30日0点之后开通EDAS的阿里云账户支持。

功能入口

- 1. 登录EDAS控制台。
- 2. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 3. 在左侧导航栏选择告警管理 > 告警事件历史。

事件列表

事件列表页面显示了未恢复告警和已恢复告警的事件名称、通知策略、创建时间、事件数量、事件状态、事件对象和对象类型。



在事件列表页面, 您可以执行以下操作:

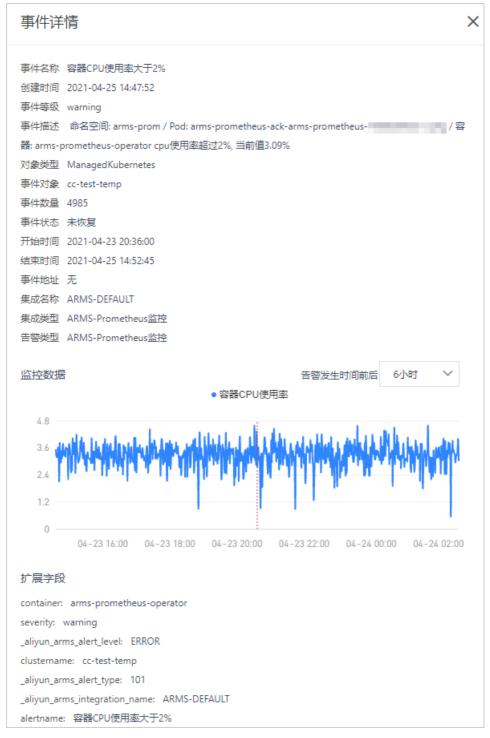
- 设置筛选字段,然后单击**搜索**,可以查看对应的告警事件。
 - ② 说明 打开更多开关,可以设置更多的过滤筛选字段。

字段	说明	
事件名称	创建的告警规则的名称。	
事件状态	告警事件的状态,共有以下3种状态: 未恢复:告警事件持续被触发。静默:在设置的自动恢复告警事件的时间内,和第一条告警事件名称和等级相同的告警事件的状态为静默。已恢复:在设置的时间内,告警事件不再触发。	
更多		
事件对象	监控任务名称或集群名称。	
对象类型	前端站点、应用或Prometheus告警。	

- 单击事件所在行,查看目标事件的详细信息。更多信息,请参见事件详情。
- 单击页面右上角的**发送测试事件**,在弹出的对话框中设置集成名称和事件内容,可以发送一次告警事件测试到指定集成中。

事件详情

事件详情面板显示了事件的基本信息、监控数据和扩展字段。



在事件详情面板的监控数据区域,您可以执行以下操作:

- 在监控数据区域右上角,设置数据显示的时间段为告警发生时间前后的6小时、12小时或1天。
- 使用光标选中一段时间,可以查看指定时间段的监控数据。单击重置,可以将曲线图恢复。

7.10. 创建应用监控告警并发送告警通知

本文通过示例说明如何创建应用监控告警并处理告警通知。

前提条件

已接入应用监控,具体操作,请参见接入应用监控。

背景信息

EDAS应用监控告警支持对采集到的各种应用指标配置告警规则。您可以根据需求和业务场景配置合理的告警规则,即可在应用不符合预期时收到通知,以便及时做出响应。本文以下述场景为例说明如何使用EDAS应用监控告警规则。

示例场景:核心业务请求延时过高或成功率下降。

示例应用名称: user-service。user-service应用是部署在阿里云容器服务Kubernetes版集群上面的一个Java应用,且已经成功接入了EDAS应用监控。

user-service应用是公司的核心服务,如果其接口可用性下降将导致严重的后果,因此,当此应用提供的接口调用延时增加或者成功率下降时,运维人员需及时收到通知并且迅速响应。

在此背景下,运维人员需要关注的指标有:

- 应用提供服务的平均响应时间
- 应用提供服务的请求成功率

步骤一:验证应用是否成功接入应用监控

执行以下步骤验证user-service应用是否成功接入应用监控。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择应用列表,在展开的列表中查找并单击user-service应用名称。
 - i. 在左侧导航栏单击**应用总览**。
 - ii. 进入**应用总览**页面,查看**应用总览**页面是否有数据。

如果页面有数据,说明应用已经成功接入应用监控;如果页面没有任何数据,说明应用没有成功接入,请将应用重新接入应用监控。

② 说明 示例应用是容器服务Kubernetes版集群的Java应用,如果您的应用是其它类型,接入应用监控成功之后的页面显示可能与上图不同,但是只要页面有指标数据就说明已经接入成功。

步骤二: 讲入创建应用监控告警规则页面

- 1. 登录EDAS控制台。
- 2. 在顶部菜单栏,选择地域。
- 3. 在**应用列表**页面**集群类型**的下拉列表中,选择**EDAS自带K8s集群**或**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 4. 在左侧导航栏中选择告警管理 > 告警规则。
- 5. 在告警规则页面右上角,单击创建应用监控告警规则。
- 6. 在**创建应用监控告警规则**页面,设置告警应用为user-service、指标类型为应用调用类型统计。

方法一

方法二

步骤三: 创建告警规则

本示例中需要创建以下两条告警规则,并在告警触发时通过短信,邮件和钉钉群通知联系人:

● user-service应用提供请求延时过高的告警规则:

最近3分钟应用提供服务调用响应时间大于等于500 ms。

● user-service应用提供请求成功率下降的告警规则:

最近3分钟应用提供服务调用错误率小于等于90%。

- ② 说明 一般认为接口在短时间内有所波动为正常现象,持续性的异常才需要关注并且响应。所以在添加告警规则时,可以设置当最近3分钟的平均延时都比较高或者请求失败率低于90%的时候才需要告警。您在创建应用监控告警规则时,可以根据需要指定不同的时间和指标计算方法。
- 1. 在**创建应用监控告警规则**页面输入告警名称,添加告警条件。 此处先创建user-service应用提供请求延时过高的告警条件。
- 2. 设置通知策略。
 - 如果您已经创建过通知策略,则在通知策略区域选择已有的通知策略。
 - 如果您之前没有创建过通知策略,则执行步骤3新建通知策略。新建完通知策略后,在**创建应用监控** 告警规则页面刷新并选择通知策略。
- 3. (可选)单击通知策略右侧的新建通知策略。

在右侧弹出的新建通知策略面板中,选择通知人,并设置其他参数,然后单击确定。

- 如果您已经创建过联系人,则在通知人区域选择已有的联系人。
- 如果您之前没有创建过联系人,则执行步骤4新建联系人。新建完联系人后,在**新建通知策略**面板刷新并选择通知人。

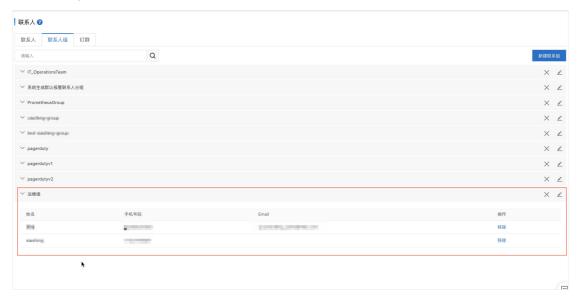


4. (可选)单击通知人右侧的新建联系人。

在联系人页面创建联系人、联系人组和钉群。

- i. 在联系人页签,单击右上角的新建联系人。
- ii. 在新建联系人对话框中输入联系人参数,然后单击确认。
- iii. 在联系人组页签,单击右上角的新建联系组。

iv. 在**新建联系组**对话框设置组名,并将上面创建的联系人添加至联系人组,然后单击**确认**。 创建完成后,在**联系人组**页签中可以查看已创建的联系人组。



- v. 在钉群页签,单击右上角的新建钉群。
- vi. 在**新建钉群**对话框设置钉群参数,然后单击**确认**。 创建完成后,在**钉群**页签中可以查看已创建的钉群。



- 5. 告警规则创建完成后,单击保存。
- 6. 重复以上步骤创建user-service应用提供请求成功率下降的告警规则。 创建完成后,在控制台的**告警管理 > 告警规则**页面可以查看创建的两条告警规则。



查看告警通知

当满足告警条件时,收到的告警通知如下所示:

● 短信:

【阿里云】告警名称:

user-service提供请求延时过高

通知策略: 运维组-测试

告警时间: <u>2021-07-18 16:44:00</u> 告警内容: 应用调用类型统计

应用提供服务调用响应时间_ms 最近3分钟求平均 >= 500.0, 当前

值10497.1443

本消息由阿里云用户:

中配置。

• 邮件:



● 钉钉群:



在钉钉群中收到告警通知后,您可以直接在钉钉群中处理告警。具体操作,请参见在钉钉群中处理告警。

处理告警

收到通知后,您可以在告警管理中查看告警的发送历史和详细告警时间信息。

1. 登录EDAS控制台。

2. 在左侧导航栏,选择**告警管理 > 告警发送历史**。 在**告警**页面可以查看告警发送历史详细告警时间信息。更多信息,请参见<mark>查看告警发送历史</mark>。



8.应用运维

8.1. 应用运维概述

EDAS为应用提供了一系列运维功能,包括生命周期管理、应用访问方式管理、应用扩缩容、限流降级、变更记录、事件中心、日志管理等。

应用运维能力

使用场景	功能
生命周期管理	将应用托管至EDAS K8s集群后,可以在EDAS控制台上管理该应用的生命周期,包括升级、回滚、扩缩容、启动、停止、删除等操作,也可以通过编辑YAML对该应用进行更精细的生命周期管理。更多信息,请参见管理应用生命周期。
应用访问方式管理	应用托管至EDAS K8s集群后,可能希望暴露服务端口,有以下两种场景: • 对外部暴露服务端口。可以为应用绑定公网或私网SLB,并提供负载均衡能力。更多信息,请参见添加负载均衡SLB或复用负载均衡SLB。 • 对集群内部暴露服务端口。可以为应用添加服务Service,并提供负载均衡能力。更多信息,请参见添加服务Service。
应用扩缩容	当流量变动较大时,您可以进行手动或自动扩缩容应用,增加或减少实例数量,提高资源利用率,降低应用资源成本。更多信息,请参见 <mark>手动扩缩容或自动弹性扩缩容</mark> 。
限流降级	EDAS支持使用应用高可用服务AHAS(Application High Availability Service)实现Spring Cloud、Dubbo和HSF 应用的限流降级,支持实时查看限流降级详情和动态变更规则,从而全面保障您的应用可用性。更多信息,请参见限流降级简介。
变更记录	在EDAS上进行应用部署、启动、扩容、缩容等生命周期操作后,可以跳转到应用详情页查看当前变更状态,也可以通过变更记录页面查看该应用的历史变更记录。更多信息,请参见查看应用变更。
事件中心	事件中心目前接入了EDAS的应用变更事件、ARMS的报警事件、0-1事件(如死锁、OOM和应用启动等)、MSE的 微服务管控事件和K8s集群事件。当您的应用使用了相关的产品,对应的事件会自动接入事件中心进行统一的分析展示,方便您查看与分析。更多信息,请参见事件中心。

使用场景	功能
日志管理	EDAS提供以下三种日志: ● 部署在EDAS K8s集群的应用的实时日志,当应用出现异常情况的时候,可以通过查看实时日志来排查容器(Pod)相关问题。更多信息,请参见查看实时日志。 ● EDAS提供了日志目录收藏、日志查看、日志搜索的功能。您可以在线查看实例和应用级别的日志来排查应用异常问题。更多信息,请参见查看日志目录。 ● EDAS对接了日志服务SLS。如果您在创建或部署应用时开启了日志服务功能,则可以查看该应用的文件日志。更多信息,请参见查看文件日志。

8.2. 管理应用生命周期

将容器服务K8s集群中的应用托管至EDAS后,可以在EDAS控制台上管理该应用的生命周期,包括部署新版本的应用、部署历史版本的应用、扩缩容等操作,此外也可以通过编辑YAML对该应用进行更精细的生命周期管理。

访问应用总览页面

- 1. 登录EDAS控制台。
- 2. 登录EDAS控制台,在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群类型下拉列表中选择容器服务/Serverless K8s集群,然后单击目标应用名称。
- 3. 在应用总览页面管理应用的生命周期。



部署新版本的应用

在应用总览页面右上角选择部署 > 部署,并选择需要的发布方式。请参见:

- 使用控制台分批发布应用(K8s)
- 使用控制台金丝雀发布应用(K8s)
 - ② 说明 开始部署后,可在变更记录页面查看变更记录信息。

部署历史版本的应用

在应用总览页面右上角选择部署 > 部署历史版本,在部署历史版本对话框选择要部署的历史版本,并单击部署该版本。

在部署过程中回滚应用

当应用正在部署时,在**应用总览**页面顶部单击**查看详情**,然后在**变更记录**页面右上角单击回滚,并在**确** 认对话框中单击**确认**。

启停应用

当应用处于运行状态时,可以根据业务需求在**应用总览**页面右上角单击**停止**来停止应用。停止应用是通过将Pod数缩容至0个来实现的。应用内至少包含1个运行中的Pod时,才能进行应用停止操作。

☐ 注意 应用停止时,如果其运行依赖其他产品,例如SLB和VPC,所依赖的产品或者服务依然处于计费状态。

当应用处于停止状态时,可以根据业务需求来启动应用。在**应用总览**页面右上角单击启动,然后在启动应用对话框中设置应用Pod数,然后单击确认。

重启应用

对于正在运行的应用实例,在**应用总览**页面右上角单击**重启应用**,将会通过滚动发布,先启动新Pod,再停止旧Pod来重启应用。

扩缩容

在应用总览页面右上角选择扩缩 > 手动扩缩或者扩缩 > 弹性扩缩。具体操作,请参见手动扩缩容或自动弹性扩缩容。

查看应用配置详情

在**应用总览**页面右上角单击**配置详情**,在面板中查看应用配置详情,例如工作负载(Deployment)、容器组(Pod)和启动命令等高级配置信息。

删除应用

在应用总览页面右上角单击删除,并在确认对话框中单击确认。

编辑YAML

在**应用总览**页面的基本信息区域的YAML右侧单击编辑,然后在编辑YAML对话框中编辑应用的YAML文件,从而对应用进行更精细的生命周期管理。编辑完毕后单击保存。

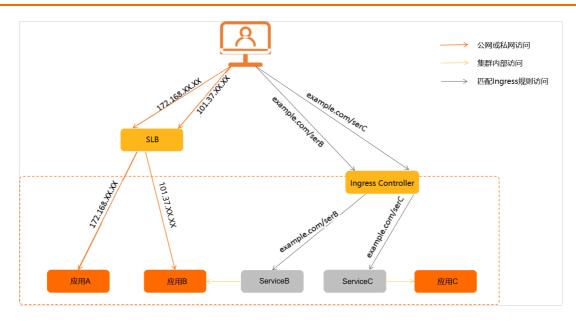
8.3. 访问应用(发布服务)

8.3.1. 访问应用概述

应用托管到EDAS K8s环境后,无法直接被其他应用所访问。本文介绍发布应用的几种方式。

服务发布方式

服务访问示例



EDAS K8s环境目前支持以下几种服务发布方式:

服务发布方式	描述	文档链接
公网访问	使用Service的LoadBalancer模式, 绑定一个公网负载均衡SLB,公网IP 可以直接访问到服务后端的Pod。	● 添加负载均衡SLB
VPC内网访问	使用Service的LoadBalancer模式, 绑定一个VPC内的私网负载均衡 SLB,私网IP可以直接访问到服务后 端的Pod。	● 复用负载均衡SLB
集群内部访问	使用Service的ClusterIP模式,添加服务Service会自动生成一个固定IP地址(即ClusterIP)。 在集群内部访问ClusterIP,此类访问将会转发给服务后端Pod,并为这些Pod提供负载均衡。	添加服务Service
匹配Ingress转发规则访问	支持为服务Service添加一条或多条应用路由(K8s Ingress)。 当外部HTTP请求访问后端服务 时,Ingress Controller解析Ingress 路由规则,根据规则将HTTP流量路 由转发到对应的Service。	创建应用路由(Nginx Ingress)

8.3.2. 应用路由选型

EDAS支持为K8s应用配置Nginx Ingress和ALB Ingress类型应用路由作为应用的外部访问入口。本文介绍 Nginx Ingress和ALB Ingress推荐使用场景和具体差异。

推荐使用场景

Nginx Ingress

Nginx Ingress是基于开源Nginx构建的K8s Ingress网关,能够使应用快捷接入Ingress,且Nginx使用广泛,是K8s集群默认的接入Ingress网关方案,提供了基础的安全、路由及可观测等能力,但Nginx Ingress组件需要您手工运维。如果您的业务流量低、对安全、扩展性和稳定性要求低,且可以接受手工运维网关,或您对网关定制有强烈需求,可以选择Nginx Ingress。

ALB Ingress

ALB Ingress基于阿里云应用型负载均衡ALB(Application Load Balancer),具备处理复杂业务路由和证书自动发现的能力,支持自动弹性、全托管免运维,单个ALB实例支持100万QPS,提供强大的Ingress流量管理。如果您对网关有高弹性和大规模七层流量处理的需求,可以选择ALB Ingress。

对比差异

Nginx Ingress和ALB Ingress的具体差异,请参见Nginx Ingress与ALB Ingress对比。

8.3.3. 添加负载均衡SLB

在EDAS中的容器服务Kubernetes集群中创建应用后,为应用添加公网或私网负载均衡SLB来实现应用的公网或私网访问。本文以容器服务Kubernetes集群中的应用为例介绍如何为应用添加公网SLB和私网SLB。

背景信息

如果需要为应用关联负载均衡SLB,必须在EDAS控制台完成关联SLB实例的监听配置管理。

☐ 注意 禁止从负载均衡控制台直接修改SLB实例的监听配置(含监听+证书),可能导致SLB控制台修改的配置回退,引发应用入口相关的故障。

为容器服务K8s集群或自建K8s集群中的应用绑定SLB有以下区别:

- 自建K8s集群:应用绑定SLB都需新购SLB实例,每个应用独享SLB实例。
- 容器服务K8s集群:应用绑定SLB可以新购SLB实例来独享SLB实例,也可以复用SLB实例。

约束限制

- 不支持使用容器服务(ACK)控制台、kubectl或者第三方界面删除或修改EDAS创建的Service资源(EDAS 资源含有以下label: edas-domain:edas-admin)。
- 不支持使用SLB控制台修改或删除EDAS代购的SLB实例。
- 不支持使用SLB控制台修改或删除EDAS托管的SLB实例的监听配置。
- 不支持使用命令行工具kubectl或在容器服务控制台更新Service的外部流量策略配置,否则可能导致修改被回退,引发应用访问相关的故障。

绑定新建SLB(公网)

- 1. 登录EDAS控制台,在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 2. 在应用总览页面的访问方式配置区域,单击负载均衡(公网)右侧的+图标。
 - ⑦ 说明 如果您已经配置过负载均衡,则在此处会显示负载均衡的IP和端口信息,您可单击
- 3. 在负载均衡(公网)对话框,设置负载均衡参数,完成后单击确认。

参数	描述	
选择SLB	您可以在下拉列表中选择 新建SLB 或已有SLB。	
调度算法	您可以选择轮询(RR)或加权轮询(WRR)。	
外部流量策略	当您选择新建SLB时,外部流量策略区域下方可以查看SLB配额检查和账户余额检查的状态和说明以及负载均衡SLB-公网的产品配置、数量、付费方式、购买周期和资费。您可以单击查看价格了解SLB价格。您可以选择Local和Cluster两种模式: Local:流量只发给本机的Pod。配置为Local模式时,只有集群中存在Service对应Pod的节点会被添加到SLB后端,且流量不会转发到集群中其他节点的Pod中。结合SLB的负载均衡能力,能够获得较好的性能且该模式下能够保留四层源IP地址。 ② 说明 当配置为Local模式时,推荐配置SLB的调度算法为加权轮询(WRR)。在容器集群中,CCM组件会根据节点上部署的Pod的数量来调整节点在SLB后端的权重,使用WRR算法能够获得较好的负载均衡能力。 Cluster:流量可以转发到集群中其他节点上的Pod。 ② 说明 配置为Cluster模式时,集群的所有节点将被挂载到SLB的后端服务器组中,即使所访问的节点上没有启动对应的Pod,也能通过SNAT转发访问到Pod服务。但是该模式无法保留四层源IP地址。	
该SLB已有的监听信息	当您选择已有SLB时,该SLB已有的监听信息参数才会出现。您可以查看网络协议、端口号和容器端口(Target port)。	
ТСР НТТР 协议	 SLB端口(应用名):公网负载均衡前端端口,通过该端口访问应用,可设置范围为1~65535。 容器端口(TargetPort):进程监听的端口。一般由程序定义,例如:Web应用默认使用8080端口。 	

参数	描述
HTTPS协议	HTTPS端口(应用名):公网负载均衡前端端口,通过该端口访问应用,可设置范围为1~65535。
	。 SSL证书 : SSL协议证书,在下拉菜单中选择已上传的SSL证书。
	。 容器端口(TargetPort): 进程监听的端口。一般由程序定义,例如:Web应用默认使用 <i>8080</i> 端口。
	② 说明 该监听配置支持设置多个监听,您可以单击添加新的监听进行配置。

- 4. 如果您的应用访问需求量已经超出了当前设置的SLB规格,则重复步骤~步骤,添加多个SLB。
 - ② 说明 SLB的规格指标主要有最大连接数、每秒新建连接数和每秒查询数,如果应用访问量超过了SLB的规格指标,您可以通过添加多个SLB来分担流量压力。

结果验证

在浏览器地址栏中输入 <负载均衡IP>:<端口号> ,例如 115.xx.XX.XX:80 ,再按回车键即可进入各自的应用首页。

如果负载均衡右侧未出现IP和端口信息,则表示绑定负载均衡失败,请进入**变更记录**查看变更详情,根据变更记录排查并修复失败原因。

绑定新建SLB(私网)

- 1. 登录EDAS控制台,在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群类型下拉列表中选择容器服务/Serverless K8s集群,然后单击目标应用名称。
- 2. 在应用总览页面的访问方式配置区域,单击负载均衡(私网)右侧的+图标。
 - ⑦ 说明 如果您已经配置过负载均衡,则在此处会显示负载均衡的IP和端口信息,您可单击
- 3. 在负载均衡(私网)对话框,设置负载均衡参数,完成后单击确认。

参数	描述
选择SLB	您可以在下拉列表中选择 新建SLB 或已有SLB。
调度算法	您可以选择 轮询(RR)或加权轮询(WRR) 。

参数	描述
外部流量策略	当您选择新建SLB时,外部流量策略区域下方可以查看SLB配额检查的状态和说明以及负载均衡SLB-私网的产品配置和数量。 您可以选择Local和Cluster两种模式: Local:流量只发给本机的Pod。配置为Local模式时,只有集群中存在Service对应Pod的节点会被添加到SLB后端,且流量不会转发到集群中其他节点的Pod中。结合SLB的负载均衡能力,能够获得较好的性能且该模式下能够保留四层源IP地址。 ② 说明 当配置为Local模式时,推荐配置SLB的调度算法为加权轮询(WRR)。在容器集群中,CCM组件会根据节点上部署的Pod的数量来调整节点在SLB后端的权重,使用WRR算法能够获得较好的负载均衡能力。 Cluster:流量可以转发到集群中其他节点上的Pod。 ③ 说明 配置为Cluster模式时,集群的所有节点将被挂载到SLB的后端服务器组中,即使所访问的节点上没有启动对应的Pod,也能通过SNAT转发访问到Pod服务。但是该模式无法保留四层源IP地址。
该SLB已有的监听信息	当您选择已有SLB时,该参数才会出现。您可以查看网络协议、端口号和容器端口(Target port)。
ТСР НТТР 协议	 SLB端口(应用名): 私网负载均衡前端端口,通过该端口访问应用,可设置范围为1~65535。 容器端口(TargetPort): 进程监听的端口。一般由程序定义,例如: Web应用默认使用8080端口。
HTTPS协议	 HTTPS端口(应用名): 私网负载均衡前端端口,通过该端口访问应用,可设置范围为1~65535。 SSL证书: SSL协议证书,在下拉菜单中选择已上传的SSL证书。 容器端口(TargetPort):进程监听的端口。一般由程序定义,例如:Web应用默认使用8080端口。 ① 说明 该监听配置支持设置多个监听,您可以单击添加新的监听进行配置。

- 4. 如果您的应用访问需求量已经超出了当前设置的SLB规格,则重复步骤~步骤,添加多个SLB。
 - ② 说明 SLB的规格指标主要有最大连接数、每秒新建连接数和每秒查询数,如果应用访问量超过了SLB的规格指标,您可以通过添加多个SLB来分担流量压力。

结果验证

在浏览器地址栏中输入 <负载均衡IP>:<端口号> / 例如 115.xx.XX.XX:80 / 再按回车键即可进入各自的 应用首页。

如果负载均衡右侧未出现IP和端口信息,则表示绑定负载均衡失败,请进入**变更记录**查看变更详情,根据变更记录排查并修复失败原因。

问题反馈

如果在使用容器服务Kubernetes集群过程中有任何疑问,欢迎您加入钉钉群进行反馈。

8.3.4. 复用负载均衡SLB

容器服务Kubernetes集群支持复用SLB,您在EDAS的容器服务Kubernetes集群中部署的多个应用可以添加同一个SLB,以便节省资源,同时还能提供应用的固定IP地址,方便运维。

前提条件

- 升级容器服务Kubernetes集群的Cloud Controller Manager (CCM) 组件。升级步骤,请参见管理组件。
 容器服务Kubernetes集群包含多个模板,需要将标准托管集群 (Managed Kubernetes) 和标准专有集群 (Dedicated Kubernetes) 模板的CCM组件升级到以下指定版本:
 - 标准托管集群(Managed Kubernetes)的CCM需为v1.9.3.112-g93c7140-aliyun及以上版本。
 - 标准专有集群 (Dedicated Kubernetes) 的CCM需为v1.9.3.106-q3f39653-aliyun及以上版本。
 - □ 注意 如果未升级容器服务Kubernetes集群的CCM组件,将导致复用SLB失败、已有监听被删除、应用流量跌零。
- 创建实例。

□ 注意

- 您在EDAS应用总览页添加负载均衡时EDAS为您代购的SLB (新建SLB)不能被复用。
- 创建的SLB和待添加SLB的应用须在同一个VPC内。
- 了解SLB使用限制,请参见使用限制。

背景信息

如果需要为应用关联负载均衡SLB,必须在EDAS控制台完成关联SLB实例的监听配置管理。

□ 注意 禁止从负载均衡控制台直接修改SLB实例的监听配置(含监听+证书),可能导致SLB控制台修改的配置回退,引发应用入口相关的故障。

为容器服务K8s集群或自建K8s集群中的应用绑定SLB有以下区别:

- 自建K8s集群:应用绑定SLB都需新购SLB实例,每个应用独享SLB实例。
- 容器服务K8s集群:应用绑定SLB可以新购SLB实例来独享SLB实例,也可以复用SLB实例。

复用SLB

复用SLB即为应用绑定已有的SLB。绑定复用的SLB与绑定新建SLB操作相同,仅在选择SLB时不同,具体操作,请参见添加负载均衡SLB。本节介绍复用SLB时的注意事项。

- 需要选择在SLB控制台创建的SLB,不能使用EDAS为您代购的SLB。
- 配置监听时,需要选择未被占用的容器端口。
- 监听覆盖:在编辑EDAS绑定的SLB时,能够设置监听覆盖。 在复用SLB时,需要为应用配置监听。在编辑负载均衡时,可以选择打开或关闭**监听覆**盖开关。



- 开启监听覆盖后,在EDAS修改的监听配置会覆盖SLB控制台的配置。
- 关闭监听覆盖时,在EDAS修改的监听配置不会生效,请您前往SLB控制台修改监听配置。
- 在EDAS中配置的监听会显示该监听对应的虚拟服务器。

② 说明 SLB的规格指标主要有最大连接数、每秒新建连接数和每秒查询数,如果应用访问量超过了SLB的规格指标,您可以通过添加多个SLB来分担流量压力。

8.3.5. 添加服务Service

本文主要介绍如何为应用添加服务Service实现集群内部服务间的快速访问。

背景信息

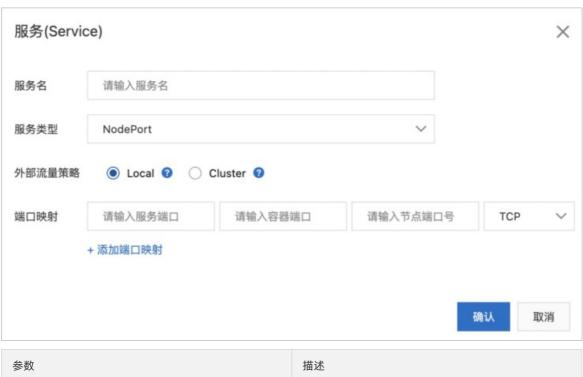
在EDAS中的容器服务Kubernetes集群中创建的应用往往都是由多个运行相同镜像的一组Pod组成,Pod虽然拥有独立的IP,但Pod会快速地创建和删除,所以通过访问一个个可能被删除的Pod来访问应用是不现实的。

通过服务 (Service) 能够解耦前端和后端的关联,从而实现松耦合的微服务设计。

在EDAS中部署的K8s应用,除了通过公网或者私网负载均衡SLB实现服务对外访问外,您还可为服务添加NodePort类型的Service,通过节点IP及节点端口提供服务外部访问。此外,在EDAS中的同一个K8s集群中创建多个服务后,如果集群内部服务间需要互相访问,无法通过公网或者私网负载均衡SLB实现访问,此时您可以为服务添加ClusterIP类型Service,以实现集群内部服务间的访问。

操作步骤

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用总览页面的访问方式配置区域,单击服务 (Service)右侧的+图标。
 - ② 说明 如果您已经配置过Service,则在此处会显示服务名称和IP信息,您可单击 ≥ 图标修改 Service的端口和协议类型,或可单击 面图标删除Service。
- 4. 在服务 (Service) 对话框,完成参数配置后单击确认。



多数 描述 自定义设置服务名称,服务名称必须唯一,不能和已有服务名称重复。 支持小写字母、数字和短划线(-),且必须以字母开 头,字母或数字结尾,长度范围为2~32个字符。

参数	描述
服务类型	 虚拟集群IP: 即ClusterIP, 指通过集群的内部IP暴露服务。选择该值,服务只能够在集群内部可以访问,这也是默认的Service类型。 节点端口: 即NodePort,通过每个Node上的IP和静态端口(NodePort)暴露服务。NodePort服务会路由到ClusterIP服务,这个ClusterIP服务会自动创建。通过请求<nodeip>:<nodeport>,可以从集群的外部访问一个NodePort服务。</nodeport></nodeip>
外部流量策略	Local:外部访问流量只转发给本机Pod。Cluster:外部访问流量可以转发给集群中其它节点的Pod。
服务端口	服务前端端口,通过该端口访问应用,可设置范围为 1~65535。
容器端口	进程监听的端口。一般由程序定义,可设置范围为 1~65535。
节点端口	节点对外暴露该服务的端口。一般由程序定义,可设置范围为30000~32767。 ② 说明 当前仅支持服务类型为NodePort时设置节点端口。
协议	服务协议,包含TCP和UDP,默认值为TCP。

? 说明

为应用添加服务Service时, 您需要关注以下两点:

- 系统支持在同一个Service里,添加多个端口映射。但是如果您需要将不同的端口映射,设置为不同的服务名称,则可以重复执行步骤3~步骤4,设置多个Service。
- 在同一个服务Service,添加多个端口映射时,服务端口和协议两两之间不得完全一样。如8 0/8080/T CP和 80/8081/T CP,这两个端口映射无法保存,提示端口映射存在重复项。

结果验证

在当前应用中添加服务Service后,您可以登录K8s集群中的任一Pod,在Pod内根据Service固定IP地址访问应用。

本示例验证的是访问Web服务,如果是其他类型服务,验证方法不尽相同。

- 1. 登录容器服务管理控制台。
- 2. 在左侧导航栏,单击集群。
- 3. 在集群列表页面,选择目标集群,在操作列下单击更多 > 通过CloudShell管理集群。 在页面下方自动打开集群所关联的CloudShell。打开集群关联的CloudShell时,系统会自动加载集群的 kubeconfig文件。您可以通过kubectl直接管理您的集群。

4. 执行以下命令,查询集群中所有的Pod。

kubectl get pods

查询出的Pod列表如下:

NAME	READY	STATUS	RESTARTS
AGE			
store-pre-***-group-1-19-***7569b-f7***	1/1	Running	0
28h			
store-prod-***duct-group-1-1-***7f894-zh***	1/1	Running	0
28h			

5. 执行以下命令,登录任一Pod。

```
kubectl exec -it store-prod-***duct-group-1-1-***7f894-zh*** /bin/sh
```

6. 执行以下命令,根据Service固定IP访问应用。

```
wget 10.XX.XX.XX:8081 //IP地址和服务端口需要和您的Service信息保持一致。
```

返回如下类似日志:

② **说明** 本示例验证的是访问Web服务,如果是其他类型服务,回显日志不相同,请以实际情况为准。

8.3.6. 创建应用路由(Nginx Ingress)

EDAS的Ingress是用于将外部HTTP/HTTPS流量路由到内部Service的路由规则集合。您可以通过配置不同的应用路由(K8s Ingress),从而根据不同的路由转发规则访问集群内不同的Service。本文介绍在EDAS控制台如何配置应用路由和转发规则。

前提条件

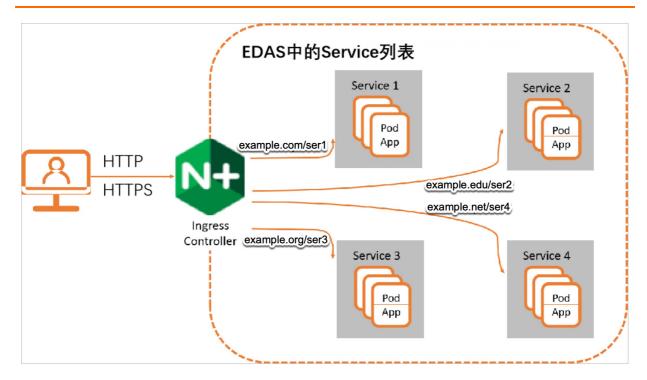
添加服务Service

背景信息

创建K8s集群时默认安装了Ingress Controller组件。K8s集群中的Ingress Controller用于解析Ingress的路由转发规则。Ingress Controller收到请求,匹配Ingress转发规则并转发到后端Service,Service转发到Pod,最终由Pod处理请求。

Service、Ingress与Ingress Controller有着以下关系:

- Service是后端真实服务的抽象,一个Service可以代表多个相同的后端服务。
- Ingress是用于将外部HTTP/HTTPS流量路由到内部Service的路由规则集合。
- Ingress Controller是一个反向代理程序,负责解析Ingress路由规则,根据规则将HTTP/HTTPS流量路由转发到对应的Service。



应用路由配置约束(适用于K8s集群)

不能修改EDAS创建的Ingress资源(含有以下标签: edas-domain: edas-admin 、 edas-domain)。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击应用路由。
- 3. 在应用路由(K8s Ingress)页面顶部菜单栏选择地域,在页面中选择目标微服务空间。
- 4. 在应用路由 (K8s Ingress) 页面中单击创建Nginx Ingress。
- 5. 在创建Nginx Ingress面板中,选择K8s集群、K8s命名空间并设置应用路由名称,单击添加转发规则并设置规则相关参数,然后单击确定。



参数	描述
K8s命名空间	K8s命名空间通过将系统内部的对象分配到不同的命名空间中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。 odefault:没有其他命名空间的对象的默认命名空间。 kube-system:系统创建的对象的命名空间。 kube-public:此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。此处以选择default为例。
重定向至HTTPS	打开 重定向至HTTPS 开关,表示用HTTP协议访问时会重定向到HTTPS协议。
应用路由名称	自定义设置路由名称,在同一个命名空间下应用路由名称必须唯一。 支持小写字母、数字和短划线(-),且必须以字母开头,字母或数字结尾。

转发规则

- ② 说明 在添加应用路由时,您需要了解以下几点:
 - 。 在同一个应用路由下,可以创建多个路由规则:
 - 相同域名和路径的组合,只能对应一个Service。
 - 同一个Service,可以对应多个域名和路径组合。
 - 。 在不同应用路由下,可以创建完全相同的路由规则。

域名	指定访问域名。
路径	指定访问路径,必须以正斜线(/)开头。
应用	从下拉列表中选择当前K8s集群中的目标应用。
服务名称	从下拉列表中选择当前应用的目标服务Service。
服务端口	服务Service的服务端口。
(可选)开启TLS	选中开启TLS,即代表允许外部HTTPS请求路由到内部Service的路由规则集合。如果允许外部HTTPS请求,还需要设置TLS证书。在Secret列表中选择TLS证书类型的保密字典,如果您还没有将HTTPS证书保存到保密字典,请单击新建密钥。创建保密字典的具体操作,请参见创建保密字典。

创建完成后,新建的应用路由出现在应用路由列表。您可以对其进行查看详情、变更路由规则信息和删除应用路由,以及查看应用路由的监控详情。

结果验证

您可在浏览器中访问路由规则中的域名/路径,查看能否访问到Service对应的后端服务。

8.3.7. 创建应用路由(ALB Ingress)

EDAS支持ALB Ingress应用路由,ALB Ingress基于阿里云应用型负载均衡ALB(Application Load Balancer)提供更有效的Ingress流量管理方式,同时兼容Nginx Ingress,具备处理复杂业务路由和证书自动发现的能力。

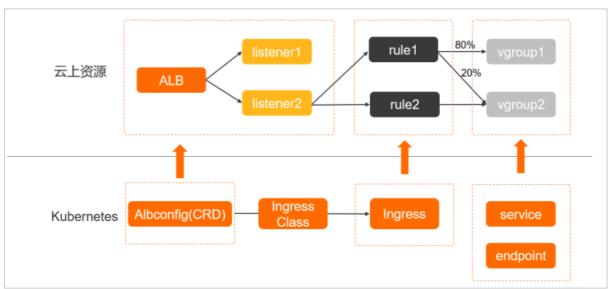
前提条件

- 添加服务Service
- 已安装ALB Ingress Controller组件。
 - 若需要创建集群,请在组件配置配置向导页面,在Ingress字段右侧选择ALB Ingress。
 - 若已创建集群,且未安装ALB Ingress Controller组件。关于如何安装ALB Ingress Controller组件,请参见管理组件。

背景信息

ALB Ingress Controller通过API Server获取Ingress资源的变化,动态地生成AlbConfig,然后依次创建ALB实例、监听、路由转发规则以及后端服务器组。Kubernet es中Service、Ingress与AlbConfig有着以下关系:

- Service是后端真实服务的抽象,一个Service可以代表多个相同的后端服务。
- Ingress是反向代理规则,用来规定HTTP或HTTPS请求应该被转发到哪个Service上。例如:根据请求中不同的Host和URL路径,让请求转发到不同的Service上。
- AlbConfig是在ALB Ingress Controller提供的CRD资源,使用AlbConfig CRD来配置ALB实例和监听。一个AlbConfig对应一个ALB实例。



约束限制

- 为Ingress服务的ALB由EDAS完全托管,您不能在ALB控制台进行配置,否则可能造成Ingress服务的异常。
- 不能修改EDAS创建的Ingress资源(含有以下标签 : edas-domain: edas-admin 、 edas-domain), 不能变更或删除Ingress关联的IngressClass、AlbConfig资源(以 edas-{alb id} 格式命名的资源)。
- 若集群的网络插件类型为Flannel,则仅支持后端Service的类型为NodePort或LoadBalancer。
- 如果需要配置HTTPS类型转发,则需要改Ingress实例的所有规则都开启TLS。

- ALB Ingress支持证书自动发现,路由规则开启TLS后无需进行Secret资源配置,但需要在SSL证书服务中购买对应域名的证书,且需要域名匹配的证书唯一。证书支持泛域名证书和单域名证书。
- 使用应用路由需要依赖已有标准版ALB实例,配置应用路由ALB Ingress后,绑定的ALB实例无法再被其他 集群复用。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击应用路由。
- 3. 在应用路由(K8s Ingress)页面顶部菜单栏选择地域,在页面中选择目标微服务空间。
- 4. 在应用路由 (K8s Ingress) 页面中单击创建ALB Ingress。
- 5. 在创建应用路由(K8s Ingress)面板中,选择K8s集群、K8s命名空间、ALB实例并设置应用路由名称,单击添加转发规则并设置规则相关参数,然后单击确定。



参数	描述
K8s集群	从下拉列表中选择目标K8s集群。
	K8s命名空间通过将系统内部的对象分配到不同的命名空间中,形成逻辑上分组的不同项目、小组或用户组,便于不同的分组在共享使用整个集群的资源的同时还能被分别管理。
K8s命名空间	o default: 没有其他命名空间的对象的默认命名空间。
	 kube-system: 系统创建的对象的命名空间。 kube-public: 此命名空间是自动创建的,并且可供所有用户(包括未经过身份验证的用户)读取。 此处以选择default为例。
重定向至HTTPS	打开 重定向至HTTPS 开关,表示用HTTP协议访问时 会重定向到HTTPS协议。
应用路由名称	自定义设置路由名称,在同一个命名空间下应用路由名称必须唯一。 支持小写字母、数字和短划线(-),且必须以字母开头,字母或数字结尾。
ALB实例	选择路由绑定的ALB实例,列表中仅返回与选择的K8s 集群在同一专有网络下的ALB实例。 若您还没有ALB实例,请参见 <mark>创建应用型负载均衡</mark> 。
	说明仅支持标准版ALB实例。若选择ALB实例绑定到该K8s集群后,无法再被其他集群使用。可以使用一个ALB实例绑定一个集群中的多个应用路由。

转发规则



? 说明 在添加应用路由时,需注意:

- 。 在同一个应用路由下,可以创建多个路由规则:
 - 相同域名和路径的组合,只能对应一个Service。
 - 同一个Service,可以对应多个域名和路径组合。
- 。 在不同应用路由下,可以创建完全相同的路由规则。

域名	指定访问域名。
路径	指定访问路径,必须以正斜线(/)开头。

参数	描述
应用	从下拉列表中选择当前K8s集群中的目标应用。
服务名称	从下拉列表中选择当前应用的目标服务Service。
服务端口	服务Service的服务端口。
(可选)开启TLS	选中开启TLS,即代表允许外部HTTPS请求路由到内部Service的路由规则集合。如果允许外部HTTPS请求,还需要设置TLS证书。在EDAS控制台左侧导航栏配置管理 > Kubernetes配置 > 保密字典列表中选择TLS证书类型的保密字典。如果您还没有将HTTPS证书保存到保密字典,请单击创建保密字典。创建保密字典的具体操作,请参见创建保密字典。

创建完成后,新建的应用路由出现在应用路由列表。您可以对其进行查看详情、变更路由规则信息和删除应用路由等操作。

8.3.8. 使用SLB访问K8s上的EDAS应用

EDAS支持以容器的方式托管应用,EDAS代为构建镜像并部署到Kubernetes集群,有效简化流程降低使用门槛。

背景信息

EDAS支持以容器的方式托管应用,EDAS代为构建镜像并部署到Kubernetes集群,有效简化流程降低使用门槛。阿里云上K8s集群中部署的CCM(Cloud Controller Manager)组件已支持将集群内一个应用的多个容器副本无缝对接到SLB后端,并支持应用进行弹性扩缩容时,自动将应用副本添加到SLB监听后端或从后端移除。

使用SLB的两种方式:

- 在EDAS中部署的应用,可以在应用总览页选择配置新建SLB,使用由EDAS代购的SLB。通过这种方式使用 SLB不支持在SLB控制台直接修改该SLB的配置。
- 在EDAS中复用已有SLB,在应用总览页选择已经在SLB控制台创建的SLB,配置EDAS应用端口监听,在暴露EDAS应用同时不影响SLB已有监听。详细操作,请参考复用负载均衡SLB。

约束限制

- 禁止从负载均衡控制台直接修改EDAS代购的SLB实例的监听配置(含监听和证书),可能导致SLB控制台修改的配置回退,引发应用入口相关的故障。
- 若复用SLB来暴露K8s上EDAS应用,则创建的SLB要和K8s集群在同一个VPC内。
- 新建SLB或复用SLB的使用限制及配额,请参考使用限制。
- 使用SLB暴露出EDAS应用后,在K8s集群内部节点存在无法通过SLB地址访问EDAS应用问题。详细信息,请参考Kubernetes集群中访问LoadBalancer暴露出去的SLB地址不通。

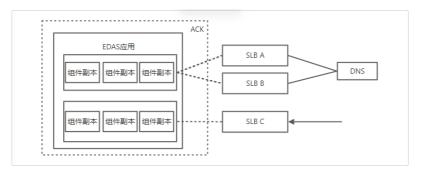
前提条件

您已成功在EDAS控制台导入Kubernetes集群并部署了一个EDAS应用。

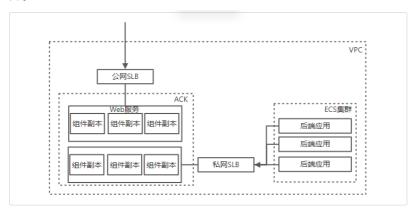
- 使用JAR包或WAR包部署应用。具体操作,请参见在容器服务K8s集群中使用JAR包或WAR包部署应用。
- 使用镜像部署应用。具体操作,请参见在K8s集群中使用镜像部署Java微服务应用。
- 将K8s中Deployment转化为EDAS应用。具体操作,请参见使用控制台转化Deployment为EDAS应用。

使用SLB访问应用

使用SLB可以为一个应用的多个后端副本提供一个稳定的对外访问地址,并进行流量负载均衡。可以使用不同SLB为一个应用的不同程序端口提供对外访问,且不同的SLB可以配置不同的实例规格,以支持不同服务的不同业务流量需求。也可以使用多个SLB绑定同一组后端副本,以扩展单个SLB的规格限制,分担流量压力。



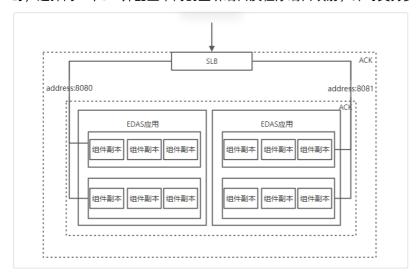
K8s上EDAS应用支持使用公网SLB和私网SLB组合方式提供应用访问。能够使用公网SLB提供Web应用公网访问,使用私网SLB提供后端其他应用间内网访问。您可以根据业务实际需求配置各自的应用扩缩及负载均衡。



- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在**应用总览**页面的**访问方式配置区域**,单击**负载均衡(公网)或负载均衡(私网)**右侧的+图标。 详细信息,请参见添加负载均衡SLB。
 - ② 说明 如果您已经配置过负载均衡,则在此处会显示负载均衡的IP和端口信息,您可单击。
- 4. 在负载均衡对话框,设置负载均衡参数。
- 5. 单击确定,等待应用变更状态正常。
- 6. (可选)如果您的应用访问需求量已经超出了当前设置的SLB规格,则重复<mark>步骤3</mark>和步骤4,添加多个 SLB。

复用SLB

 为了降低成本,减少SLB的实例数,EDAS允许为多个应用共享一个SLB来支持对外服务访问,以支持多个应用通过配置不同的监听端口使用一个SLB对外访问。在EDAS中,找到需要使用SLB的应用,在进行访问配置时,选择同一个SLB并配置不同的监听端口及程序端口映射,即可支持多个应用使用同一个SLB。



? 说明

- 复用SLB需要您持有一个或多个处于运行中状态的SLB实例。具体操作,请参见创建和管理CLB实例。
- 如果需要使用公网SLB绑定应用,则需SLB网络类型为公网;如果需要使用私网SLB绑定应用,则需要SLB网络类型为私网。
- SLB实例和K8s集群位于同一地域。
- 如果SLB实例和K8s集群实例的网络类型均为专有网络,则必须位于同一专有网络。

1. 登录EDAS控制台。

- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在**应用总览**页面的**访问方式配置**区域,单击**负载均衡(公网)**或**负载均衡(私网)**右侧的+图标。 详细信息,请参见复用负载均衡SLB。
 - ⑦ 说明 如果您已经配置过负载均衡,则在此处会显示负载均衡的IP和端口信息,您可单击。
- 4. 在负载均衡对话框,设置负载均衡参数。
- 5. 单击确定,等待应用变更状态正常。
- 6. (可选)如果您的应用访问需求量已经超出了当前设置的SLB规格,则重复步骤3和步骤4,添加多个SLB。

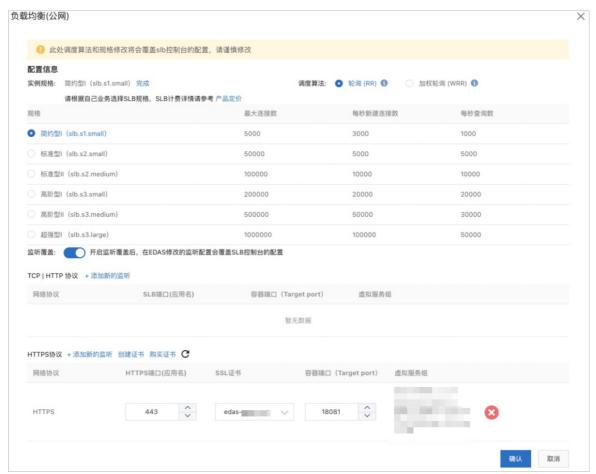
更新SLB配置

当您持有一个绑定了SLB的EDAS应用,需要修改SLB配置时,可参考本小节。应用绑定SLB的具体操作,请参见添加负载均衡SLB。

注意事项

- 在EDAS中支持更改SLB的实例规格、调度算法、监听端口映射,对于EDAS新建SLB和监听的默认配置,请参见新建SLB和监听配置的默认值。
- 对于EDAS代购的SLB禁止通过SLB控制台进行配置更改,以避免多入口更改配置导致的数据不一致问题, 防止变更的配置回退;非EDAS代购的SLB支持在SLB控制台进行配置更改。
 - 1. 登录EDAS控制台。
 - 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
 - 3. 在应用总览页面的访问方式配置区域,单击负载均衡(公网)或负载均衡(私网)右侧的 图标。
- 4. 在负载均衡对话框,设置负载均衡参数。

更改实例规格为简约型,调整调度算法为轮询,并修改已有的HTTPS协议类型443端口监听的容器端口为18081。

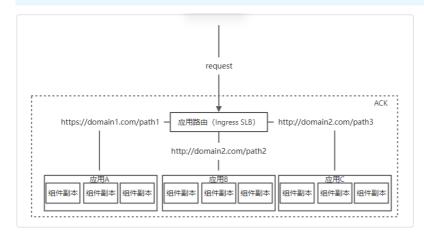


5. 单击确定,等待应用变更状态正常。

K8s中EDAS应用的转发规则配置

在K8s上部署的EDAS应用,能够使用应用路由实现根据域名和URL进行流量分发,以实现七层负载均衡路由转发功能。应用路由中使用的SLB无需用户额外购买及配置,详细信息,请参见创建应用路由(Nginx Ingress)。

② 说明 通过EDAS平台代购的SLB不支持通过SLB控制台进行路由规则配置更改,仅非EDAS代购的 SLB支持在SLB控制台进行路由规则配置。

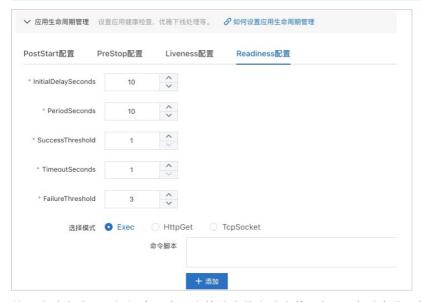


K8s中EDAS应用的健康检查

在K8s上部署的EDAS应用时,可以通过配置应用生命周期的钩子和探针来替代SLB中监听的健康检查配置。

? 说明

通过EDAS平台代购的SLB不支持通过SLB控制台进行健康检查配置更改,非EDAS代购的SLB支持在SLB控制台进行健康检查配置。



使用这种方式可以根据应用中副本的健康状态动态的更新SLB中对应监听的后端,此外还支持更多的健康检查方式。详细信息,请参见配置应用生命周期的钩子和探针。

新建SLB和监听配置的默认值

配置	默认值及相关描述
负载均衡带宽	5120 Mbps
负载均衡的IP版本	IPv4
负载均衡删除保护	打开
负载均衡配置修改 保护	打开
会话保持时间	0秒
会话保持开关	关闭
健康检查开关	TCP类型监听默认开启HTTPS类型监听默认关闭。
健康检查协议	ТСР
健康检查健康阈值	健康检查连续成功N次后,将后端服务器的健康检查状态由fail判定为success,默认值为3。
健康检查不健康阈值	健康检查连续失败多少次后,将后端服务器的健康检查状态由success判定为fail,默认值为3。
健康检查间隔时间	2秒
健康检查响应超时时间	接收来自运行状况检查的响应需要等待的时间,适用于TCP模式。如果后端ECS在指定的时间 内没有正确响应,则判定为健康检查失败,默认值为5秒。
健康检查正常状态码	HTTP_2xx
访问控制功能开关	关闭
TCP监听连接超时时 间	配置TCP连接超时,连接空闲时间超过该时长后,负载均衡会主动断开该连接,默认值为900秒。
HTTPS监听连接空 闲超时时间	15秒
HTTPS监听HTTP2 特性	打开

8.4. 应用扩缩容

8.4.1. 手动扩缩容

本文介绍通过手动扩缩功能,实现应用实例数增加和减少,在保证服务质量的同时,高效的利用应用资源、降低应用资源成本。

背景信息

如果应用配置了Service, SLB, 在发生弹性伸缩时, EDAS会自动将扩容的应用实例添加到后端服务器列表, 并将缩容的实例从后端服务器列表中移除。

如果应用配置了持久化存储,在发生弹性伸缩时,EDAS会自动将扩容的应用实例与配置的持久化存储进行 关联。

操作步骤

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在应用总览页面右上角,选择扩缩 > 手动扩缩。
- 4. 在**应用扩缩容**对话框中设置要扩容或缩容的应用总Pod数,单击确认。

如果将Pod数量设置为0,系统将物理删除该应用下所有的Pod,仅保留应用的基本创建信息。

开始扩缩容后,在**应用总览**页面出现**应用正在变更...**字样,您可以单击**查看详情**,查看应用变更的详细过程。应用变更需要几分钟,请您耐心等待。变更完成后在**应用总览**页面的基本信息区域查看实例的运行状态,如果显示为Pod运行中,说明应用部署成功。



8.4.2. 自动弹性扩缩容

本文介绍如何在EDAS中完成自动弹性伸缩策略配置。分布式应用管理中,弹性伸缩是很重要的一个运维能力。弹性伸缩能够感知应用内各个实例的状态,并根据实例状态动态实现应用扩容、缩容,在保证服务质量的同时,提升应用的可用率。

背景信息

互联网、游戏类等应用在促销活动期间容易出现突发性流量洪流,SLA和资源成本不易平衡,极易造成系统响应延迟、系统瘫痪等问题。EDAS继承阿里巴巴应对双11的流量洪流技术,提供秒级自动弹性功能,保证SLA的同时也节省实例保有成本。多适用于互联网、游戏以及社交平台等行业。

部署在EDAS的容器服务K8s集群和Serverless K8s集群中的应用,可以配置弹性策略。

在弹性策略内,可以配置多个触发器。

- 当K8s集群版本低于V1.15.0时,任意一条触发器运行失败,将不会进行扩缩容。
- 当K8s集群版本为V1.15.0或以上版本时,任意一条触发器运行成功,即可成功扩缩容。当多条触发器运行成功时,应用目标实例数=max(触发器1目标实例数,触发2目标实例数,…,触发器N目标实例数)。

您可前往事件中心订阅应用扩缩容相关事件,当应用发生扩缩容、扩缩容失败(自动扩缩达到设置的最大或最小实例数),您都会接收到消息通知。订阅弹性伸缩相关的K8s事件,请参见事件中心。

(→) 注意

- 弹性策略启用后,无法对应用执行**停止、启动**和**手动扩缩**操作。如果必须执行,请先停止弹性 策略,再执行。
 - 如果应用配置了Service、SLB,在发生弹性伸缩时,EDAS会自动将扩容的应用实例添加 到后端服务器列表,然后将缩容的实例从后端服务器列表中移除。
 - 如果应用配置了持久化存储,在发生弹性伸缩时,EDAS会自动将扩容的应用实例与配置的持久化存储进行关联。
- 执行应用变更(如部署应用、应用扩缩、变更规格等)时,无法添加弹性策略。
- Kubernetes集群不具备节点自动扩容能力时,应用弹性扩容可能无法工作,您可以前往集群页面配置节点弹性伸缩。具体操作,请参见节点自动伸缩。

操作步骤

您可以配置多个触发器,EDAS将根据触发器自动为您扩容和缩容应用实例。

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,然后在**应用列表** 表页面单击具体的应用名称。
- 3. 在应用详情页面的左侧导航栏选择**应用设置 > 弹性伸缩**,然后单击**添加弹性策略**。 您也可以在**应用总览**页面右上角,选择**扩缩 > 弹性扩缩**,然后单击**添加弹性策略**。您也可以通过页面 下方的**快速添加指标策略和快速添加定时策略**区域来添加弹性策略。
- 4. 在策略名称的文本框中输入自定义的弹性策略名称。
 - ② 说明 策略名称必须以字母开头,允许数字、字母、短划线(-)或下划线(_)组成。最大长度30个字符。
- 5. 在**触发条件**区域单击**新增触发器**,然后在右侧弹出的**新增触发器**面板配置触发器规则,配置完成后单击**确**定。

弹性策略触发器参数说明

参数	描述
触发器名称	必须以字母开头,允许数字、字母、短划线(-)或下划线(_)组成。最大长度30个字符。
	选择系统提供的触发器类型,当前支持 定时触发 器和 应用指标触发器 。
	。 定时触发器 适用于负载强周期型应用,例如白天(08:00~ 20:00)负载高、晚上(20:00~08:00)负载低或 工作日(周一~周五)负载高、周末(周六~周 日)负载低的周期性应用。

a. **周期**:您可在下拉列表中选择**每天**、每描述 图 5 5 7 参数 周或每月。 ② 说明 选择每周后根据需求可选择 具体日期来配置触发规则,如果一周内不 同日期的单天内触发规则不同, 可通过创 建多条定时触发器来实现。 b. **单天内的触发时间**:通过单击添加来创建每 天的触发时间和实例数。 注意 ■ 如果错过设定的触发时刻,定时 策略需要等到下一周期才会被触 ■ 两个邻近触发时间间隔要大于10 分钟。 ■ 两个邻近触发时刻设定的实例数 变化需要大于10%, 即 Pod2 / Pod1 > 1.1 或者 Pod2/ Po d1 < 0.9 , 其中Pod1是前一 个时刻设定的实例数, Pod2是后 一个时刻设定的实例数。 c. 之后保持实例数: 跟据实际情况设置弹性策 略执行后保持的实例数,可设置范围为0~50。 d. 触发模式: 触发周期选择每天时呈现此参数, 可在下拉列表中选择边沿触发模式或水平触 发模式。 触发器类型 ■ 水平触发:弹性策略在两个设定时间段内, 一直保持开始时刻的设定值。 ■ 边沿触发:弹性策略只在设定时刻时生效, 超过设定时刻则忽略。 ○ 应用指标触发器 适用于应用指标与容量线性相关型应用,例如通过 压测确定应用单实例可以处理200 QPS, 当QPS达到 1200的时候,将扩容至6个实例的应用。 a. 指标名: 支持选择应用提供服务平均每秒请 求量、应用提供服务平均响应时间(毫 秒)、应用平均CPU使用率和应用平均内存 使用率四种指标。 b. 按均值计算指标数据: 触发器的类型为应用 指标触发器且指标名称选择应用提供服务平 均每秒请求量、应用提供服务平均响应时间 (毫秒)时,呈现此参数。当前系统默认按实 时数据计算指标数据,勾选按均值计算指标 数据后可按照均值计算,可有助于减少数据波 动。按照均值计算的**指标时间窗口**,即用于均 值计算的数据时间范围,参数范围为

> 文档版本: 20220708 351

300~3600秒。

参数	描述 c. 目标值:设置期望的监控指标目标值,系统会描述 帮您自动扩缩实例数,无限接近您设置的目标
	值。
	② 说明 当应用监控指标的实际值小于目标值时,系统会自动帮您缩容实例。 反之,系统会自动帮您扩容实例。
	根据需要选择触发器的运行方式:
	 仅观测: 观测触发器指标, 当指标达到期望值时发送事件, 不触发扩缩容动作。
方式	观测并扩缩: 观测触发器指标, 当指标达到期望值时发送事件, 触发扩缩容动作。
	② 说明 当选择应用平均CPU使用率和应用 平均内存使用率作为触发器指标时,目标值的 0.9~1.1倍系统会认为是合理的波动,当实际值 超过这个范围,才会触发弹性。

- 6. 配置最大应用实例个数和最小应用实例个数。
 - **最大应用实例个数**:可设置的范围为2~1000。触发弹性伸缩条件后,应用扩容,其实例数可达到最大值。
 - **最小应用实例个数**:可设置的范围为0~1000。触发弹性伸缩条件后,应用缩容,其实例数可达到最小值。

? 说明

- 应用目标实例数=max (触发器1目标实例数, 触发2目标实例数, ..., 触发器N目标实例数)
- 当目标实例数大于当前实例数时,EDAS则对应用进行扩容,其应用实例数不超过所设的最大应用实例个数;反之,则进行缩容,其应用实例数不低于所设的最小应用实例个数。
- 最小应用实例个数不得等于最大应用实例个数。
- 建议当前实例数在您设置的**最小应用实例个数**和**最大应用实例个数**范围内,如果不在范围内,保存弹性策略时会提示定时触发器保持实例数已超出最大最小实例数范围,但仍能保存。
- 7. (可选)单击自定义弹性策略开关,启用弹性伸缩策略并配置扩容和缩容行为。
 - **冷却时间(s)**:根据需要设置冷却时间。单位为秒,范围为0~3600,**扩容行为**冷却时间默认为0 秒,缩容**行为**冷却时间默认为300秒。
 - 扩容策略或缩容策略:在**扩容行为**区域单击**添加扩容策略**,或在**缩容行为**区域单击**添加缩容策略**, 在弹出的对话框中,配置扩容或缩容策略,然后单击**确定**。

- **策略类型**:根据需要选择副本个数或百分比作为策略类型。支持Pods和Percent类型,Pods表示副本个数类型,Percent表示百分比类型。
- 步长/百分比:根据选择的策略类型设置该参数值。例如,策略类型选择Pods,则此处值为Pod数量。
- **时间周期(s)**:策略执行周期。例如,**策略类型**选择**Percent**,**步长/百分比**设置为10,时间周期为60秒,则允许在一分钟内最多扩容或缩容当前副本个数的百分之十。

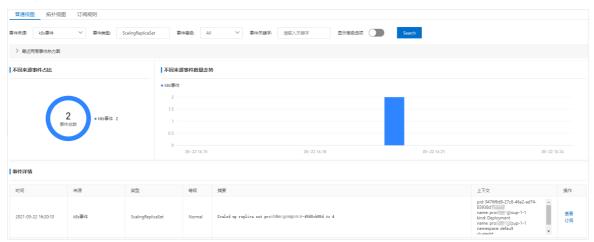
单击同步至缩容策略或同步至扩容策略,可以将配置的扩容或缩容策略同步至缩容或扩容策略。

- 选择策略: 支持Min、Max和Disable三种策略。默认为Max。
- 8. 完成以上配置成后单击保存。
- 9. 在弹性策略页面右上角单击**启用。** 触发弹性策略后,如果EDAS依据所设策略对应用实例进行扩容或者缩容,表示自动弹性伸缩策略设置 成功。

结果验证

启用弹性策略后,EDAS将自动依据弹性策略进行应用实例扩缩容,您可以通过以下步骤来查看扩缩容的详细事件记录。

- 1. 在弹性伸缩页面的右上角单击事件。
- 2. 在事件中心页面右上角,设置查询事件的时间范围。
- 3. 在事件中心页面,单击普通视图页签。
- 4. 在普通视图页签,在事件来源的下拉列表中选择K8s事件,在事件类型的下拉列表中选择ScalingReplicaSet或DryRunSuccessfulReScale,然后单击Search。
 - ScalingReplicaSet:以观测并扩缩方式触发生成的事件记录,可查看到扩缩容的详细记录。
 - DryRunSuccessfulReScale: 以仅观测方式触发生成的事件记录,未触发扩缩容动作。



更多信息

弹性策略启用后,您可以对弹性策略进行删除、停用、启用和编辑,还可以进行一系列应用管理操作。

② 说明 当弹性策略启用后,无法对应用执行**停用、启动**和**手动扩缩**操作。如果必须执行,请先停止弹性策略,再执行。

EDAS容器服务K8s和Serverless K8s交流群

如果您在EDAS中使用容器服务K8s集群和Serverless K8s集群过程中有任何疑问或建议,请提交工单,或使用钉钉搜索钉钉群号 23197114 加入钉钉群进行反馈。

8.5. 限流降级

8.5.1. 限流降级简介

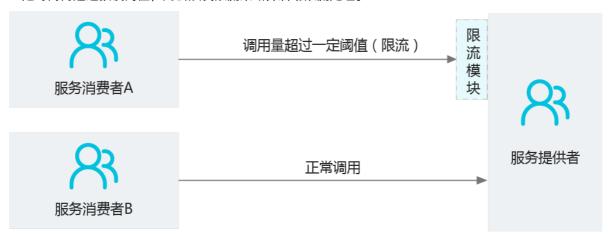
EDAS已支持使用应用高可用服务AHAS(Application High Availability Service)实现Spring Cloud应用、Dubbo应用和HSF应用的限流降级,并支持实时查看限流降级详情和动态变更规则,从而全面保障您的应用可用性。

什么是限流降级?

限流

可以理解为一个控制流量阈值或调节比例的功能。例如,在前端网站面对大流量访问时,可以对流量进行控制,防止大流量对后端核心系统造成破坏并导致服务不可用的情况。总体而言,限流是指通过调节流量阈值来控制通过系统的最大流量值,保证系统安全可靠运行。

在服务提供者端配置限流模块代码,并在EDAS中配置限流策略后,即可使服务提供者具备限流功能。当服务消费者调用服务提供者的服务时,所有访问请求都会通过限流模块进行计算,若服务消费者调用量在一定时间内超过预设阈值,则会触发限流策略并启动限流处理。



● 降级

在EDAS中,降级通常用于对下游出现超时的非核心服务提供者进行低优先级调用,确保上游核心应用(服务消费者)不被影响。

通过在服务消费者端配置降级模块代码,并在EDAS中配置降级策略,使服务消费者具备降级功能。此时服务消费者去调用服务提供者时,若服务提供者服务响应时间超过了预设阈值,则会触发降级策略进行降级处理。



使用须知

在使用EDAS限流降级功能前,您需要关注以下内容:

● 请确保您已开通AHAS。具体操作,请参见开通AHAS。



- 新部署的应用将默认使用AHAS组件进行限流降级。
- 部署的已有应用如果未使用过限流降级,当您使用限流降级时,将默认使用AHAS进行限流降级。
- 部署的已有应用如果已使用过限流降级,将继续使用原有的实现方式。

功能概述

限流降级功能具体包括:

- **流控规则**: 监控应用流量的QPS或线程数等指标,当达到您指定的阈值时立即拦截流量,以避免被瞬时的流量高峰冲垮,从而保障应用的可用性。更多信息,请参见配置流控规则。
- **降级规则**:监控应用下游依赖应用的响应时间或异常比例,当达到您指定的阈值时立即降低下游依赖应用的优先级,避免应用受到影响,从而保障应用的可用性。更多信息,请参见配置降级规则。
- 热点规则:为应用配置热点规则后,EDAS将分析统计热点参数(资源调用过程中的调用次数较高的参数),并根据配置的热点规则对包含热点参数的资源调用进行限流,保护系统稳定性。更多信息,请参见配置热点规则。
- 操作日志:记录当前云账号及当前云账号创建的RAM用户对应用进行的操作,包括具体资源名、操作的内容、操作时间和操作人ID等,方便您追踪应用和资源的变更。更多信息,请参见查看限流降级操作日志。
- **权限管理**:如果RAM用户需要配置和推送AHAS中应用的流控降级规则,您需要对RAM用户单独授予读写权限,从而严格管控权限,降低由于错误推送配置引发线上问题的可能性。

8.5.2. 限流降级原则

8.5.2.1. 限流降级原则概述

限流降级功能是基于Sentinel实现的,本文介绍限流降级原则。

Sentinel原则

- Sentinel是围绕着资源来工作的。
- 编码时,只需要关心如何定义资源,即需要关注保护哪些方法或代码块,而不需要关注如何保护这个资源。

• 通过添加规则来保护资源,规则添加后立即生效。

规则配置原则

- 按照应用处理能力进行流控:
 - 。 服务提供方或消费方流控
 - o 削峰填谷
 - 预热启动 (Warm Up)
 - o 关联限流
- 弱依赖降级

8.5.2.2. 服务提供方或消费方流控

限流降级可以根据服务提供方的能力和服务消费方的分配能力进行流量控制。其中服务提供方(Service Provider)是指对外提供请求的服务或应用;服务消费方(Service Consumer)是指调用该服务的下游应用。

根据服务提供方限流

为了保护服务提供方不被激增的流量拖垮影响稳定性,您可以为其配置QPS模式的流控规则。当每秒的请求量超过设定的阈值时,AHAS将拒绝多余的请求。提供方限流可以分为服务接口限流和服务方法限流。

- 服务接口限流:适用于整个服务接口的QPS不超过一定数值的情况。例如:为对应服务接口资源配置QPS 阈值。
- 服务方法限流:适用于服务的某个方法的QPS不超过一定数值的情况。例如:为对应服务方法资源配置 QPS阈值。

示例:

若应用A为服务提供方,其Web接口 /queryData 对应的方法是 queryData(java.lang.String) 。假设应用A最多每秒只能承受10次调用。若调用次数超过,则应用A宕机。



针对此类场景,可以对应用A按服务方法粒度设置流控规则,限制 queryData(java.lang.String) 方法每秒最多只能被调用10次,具体操作,请参见配置流控规则。若超过阈值,消费方将会收到一个 BlockException 异常,并且快速返回。

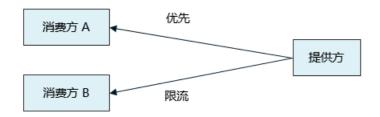
根据服务消费方限流

根据消费方限流是指根据调用方的需求来分配服务提供方的处理能力。

② **说明** 若限流规则未设置调用方(default),则该限流规则对所有调用方生效。若限流规则设置了调用方,则限流规则仅对指定调用方生效。

示例:

有两个服务消费者A和B都向服务提供方发起调用请求,我们希望优先服务消费方A,而只对来自消费方B的请求进行限流。通过对消费方B设置限流规则(limitApp)来实现这个目的。



对于默认框架例如Dubbo, AHAS会自动解析Dubbo消费方的 application name 作为调用方名称 (origin), 在进行资源保护的时候都会带上调用方名称。而对于非默认框架,只需要在Sentinel原有的代码中加入 ContextUtil.enter(resourceName, origin) 和 ContextUtil.exit() 即可。示例代码如下:

```
Entry entry = null;
//这里代表消费者A的调用。
ContextUtil.enter(queryData, "A");
try {
    // 资源名可使用任意有业务语义的字符串。
entry = SphU.entry("queryData");
    // 被保护的业务逻辑。
} catch (BlockException el) {
    // 资源访问阻止,被限流或被降级。
    // 进行相应的处理操作。
} finally {
    if (entry != null) {
        entry.exit();
    }
}
//调用结束。
ContextUtil.exit();
```

? 说明

ContextUtil.enter(xxx) 方法仅在初次调用链路入口时才生效。

更多设置

限流规则中的 limitApp (针对来源)支持以下三种选项,分别对应不同的场景:

- default : 适用于不区分消费者的场景。来自任何调用者的请求都将进行限流统计。如果这个资源名的调用总和超过了这条规则定义的阈值,则触发限流。
- {some_origin_name} : 适用于对特定的消费者限流的场景。只有来自这个调用者的请求才会进行流量控制。
 - 例如,资源A配置了一条针对消费者caller1的规则,那么当且仅当来自消费者caller1对资源A的请求才会触发流量控制。
- other :表示针对除 {some_origin_name} 以外的其余调用方的流量进行流量控制。这个场景适用于资源对大部分消费者都有一个通用的阈值,对特定消费者有不一样的阈值的场景。 例如,资源A可以对大部分消费者可以每秒提供10个请求,但是对于消费者caller1是个例外,对caller1,每秒可以提供200个请求。需配置两条规则,说明如下:
 - o 为消费者caller1配置一条 limitApp 为caller1的限流规则,这条规则每秒的最大请求量设置为200;

○ 同时,配置一条 limitApp 为other的规则,这条规则每秒的最大请求量设置为10,那么任意来自非 caller1对该resource的调用,都不能超过10。

⑦ 说明 同一个资源名可以配置多条规则,规则的生效顺序为:{some_origin_name}>other>default。

相关文档

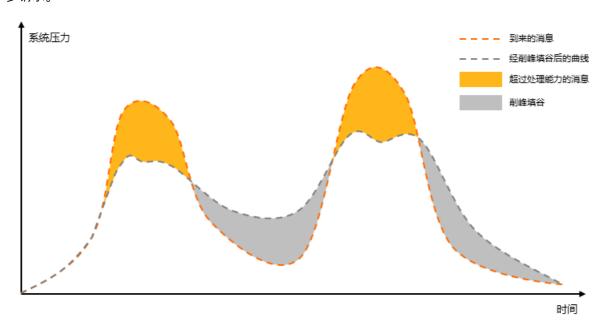
• 配置流控规则

8.5.2.3. 削峰填谷

当消费端请求骤增时,可以为其配置排队等待的流控规则,以稳定的速度逐步处理这些请求,达到削峰填谷的效果,从而避免流量骤增造成系统负载过高。

背景信息

在实际应用中,收到的请求是没有规律的。例如,某应用每秒可处理10个请求,某一秒突然来了30个请求,而接下来两秒都没有请求到达。在这种情况下,如果直接拒绝20个请求,应用在接下来的两秒就会空闲。所以,需要把骤增的请求平均分布到一段时间内,让系统负载保持在请求处理水位之内,同时尽可能地处理更多请求。

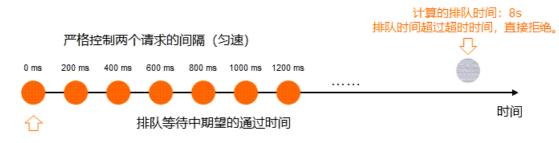


上图中,黄色区域代表超出消息处理能力的部分。将黄色部分的消息平均分布到之后的空闲时间去处理,既可以保证系统负载处在一个稳定的水位,又可以尽可能地处理更多消息。通过配置流控规则,可以达到消息 匀速处理的效果。

功能原理

限流降级的排队等待功能,可以把骤增的大量请求匀速分配,以固定的间隔时间让请求通过,达到削峰填谷的效果,从而避免流量骤增造成系统负载过高的情况。堆积的请求将会被排队处理,当请求的预计排队时间 超过最大超时时长时就会被拒绝。

例如,配置匀速模式下请求QPS为5,则每200 ms处理一条请求,多余的处理任务将排队。另外,由于设置的超时时间为5s,则预计排队时长超过5s的处理任务将被直接拒绝。具体操作步骤,请参见新建流控规则。



许多请求同时到达

8.5.2.4. 预热启动 (Warm Up)

对于长期处于低水位状态的系统,可以使用预热启动功能来避免流量骤增导致的水位瞬间升高和系统不可用的情况。

预热启动原理

当系统长期处于低水位时,若流量突然增加,可能会令系统水位瞬间升高,进而导致系统崩溃。通过配置预 热启动规则,可以让通过的流量缓慢增加,在一定时间内逐渐增加到阈值上限,给冷系统一个预热的时间, 避免冷系统崩溃。

预热启动功能借鉴了Guava的算法,通过随时调整斜率,在指定时间内将流量缓慢调整到特定阈值。

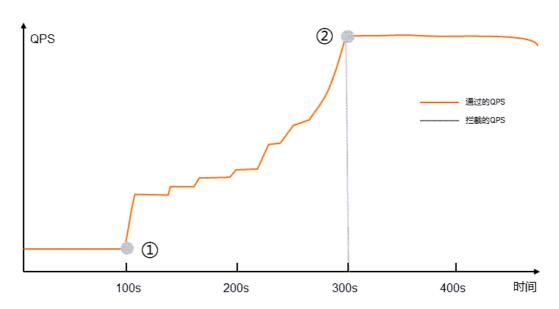
预热启动示例

例如,假设按照以下关键参数设置流控规则:

统计维度:当前接口流控效果:预热启动预热时间:200秒

⑦ 说明 配置流控规则的具体步骤,请参见配置流控规则。

启用该规则后,可以看到流量的增长趋势如下图所示:



上方示意图的含义如下:

- 系统已处于闲置状态一段时间,此时收到大量请求(图示中①处)。
- 预热启动期之后,系统开始按照规则定义的数量来处理请求(图示中②处)。

相关文档

• 配置流控规则

8.5.2.5. 关联限流

使用关联限流策略,可以避免具有关联关系的资源之间过度的争抢,造成的资源不可用问题。 当两个资源之间具有资源争抢或者依赖关系的时候,这两个资源便具有了关联。例如对数据库同一个字段的读操作和写操作存在争抢,读的速度过高会影响写得速度,写的速度过高会影响读的速度。如果放任读写操作争抢资源,则争抢本身带来的开销会降低整体的吞吐量。可使用关联限流来避免具有关联关系的资源之间过度的争抢。

示例

read_db 和 write_db 这两个资源分别代表数据库读写。给 read_db 设置以下规则来达到写优先的目的。具体操作,请参见配置流控规则。

当写库操作过于频繁时,读数据的请求会被限流。 read_db 会在 write_db 资源的QPS超过10之后,调用被拒绝。



相关文档

• 配置流控规则

8.5.2.6. 弱依赖降级

如果依赖的第三方应用出错不会影响整体流程,则将该依赖称为弱依赖。为弱依赖配置降级原则可保证系统稳定性。

背景信息

在实际业务中,应用通常会调用依赖方(远程服务、数据库、第三方API等)来完成服务。例如,支付时需要远程调用银联提供的API。然而依赖方的稳定性是不可控的,若依赖方出现不稳定的情况,则请求和调用依赖方的方法响应时间变长,线程产生堆积,最终可能耗尽自身的线程数,导致应用本身不可用。

在复杂链路中,若某一环不稳定,就可能最终导致整个链路都不可用。针对这种情况,可以使用应用高可用 服务AHAS的应用流控降级功能为依赖方配置降级原则来保证系统稳定性。

功能原理

若应用依赖于多个下游服务(弱依赖),当下游服务调用过慢,则会严重影响当前服务的调用。为调用端配置基于平均响应时间或错误率的降级规则后,当调用链路中某个下游服务调用的平均响应时间或错误率超过阈值,AHAS就会对此调用进行降级操作,拒绝多余的调用,保护应用不被调用端短板影响。

配置降级规则具体操作,请参见配置降级规则。同时可以配合Fallback功能使用,在被降级的时候提供相应的处理逻辑。

相关文档

• 配置降级规则

8.5.3. 配置规则

8.5.3.1. 配置流控规则

配置流控规则的原理是监控应用或服务流量的QPS指标,当指标达到设定的阈值时立即拦截流量,避免应用 被瞬时的流量高峰冲垮,从而保障应用高可用性。本文介绍如何配置和管理流控规则。

背景信息

流量控制在网络传输中是一个常用的概念,常用于调整网络包的发送数据。系统需处理的请求是随机不可控的,而系统的处理能力是有限的,因此就需要根据系统的处理能力对流量进行控制。流控规则的配置通常有以下场景:

- 服务提供方或消费方流控
- 削峰填谷
- 预热启动 (Warm Up)
- 关联限流

新建流控规则

- 1. 登录EDAS控制台。
- 2. 执行以下操作之一以进入应用详情页面:
 - 在左侧导航栏选择**资源管理 > 集群**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群页面的**容器服务K8s集群**页签上单击集群ID或名称,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**应用列表**页面单击具体容器服务K8s集群应用的应用名称。
- 3. 在左侧导航栏中选择限流降级 > 规则管理。
- 4. 单击流控规则页签,然后在页面右上角单击新增流控规则。
- 5. 在新增流控规则对话框,设置规则参数,然后单击新建。



参数	描述
接口名称	待流控的资源名称。
是否集群流控	是否开启集群流控,即对集群内此资源的调用总量进行限制。
是否开启	打开开关表示启用该规则,关闭开关表示禁用该规则。
来源应用	该规则针对的来源应用,默认来源应用为 default ,代表不区分来源应用。

参数	描述
统计维度	统计QPS的维度,包含以下三种维度: 。 当前接口:用于接口调用流控。该接口被来源应用调用次数超过阈值时,会对当前接口来自于来源应用的请求进行流控。 。 关联接口:用于资源争抢情况。当关联接口被来源应用调用QPS超过阈值时,会对当前接口来自于来源应用的请求进行流控。例如, read_db 和 write_db 这两个资源分别代表读数据库和写数据库。可以为 read_db 设置限流规则,将关联接口名设置为 write_db 。当写数据库操作过于频繁时,读数据的请求会被限流。 。 链路入口:用于应用内Call Stack调用流控情况。当Call Stack入口被来源应用调用QPS数超过阈值时,会对当前接口来自于来源应用的请求进行流控。
单机QPS阈值	应用或服务流量的QPS指标。
流控效果	流量控制的方式,包含以下三种方式: 中速失败:常规流控方式,当前接口超过设置阈值的流量,直接返回默认流控信息,如文本、静态页面等。 添热启动:适用于避免突增流量瞬间冲击系统的场景。请求流量会在预热时间内缓步增加至阈值量级,多余请求会直接拒绝。更多信息,请参见预热启动(Warm Up)。 排队等待:适用于流量匀速器场景。触发流控后,多余请求会按照顺序等待,达到超时时间后失败。更多信息,请参见削峰填谷。

配置成功后,新的流控规则出现在规则列表中。

管理规则

在规则管理页面的规则列表中,您可以编辑、删除、启用、禁用、批量启用或批量禁用规则。

- 如需编辑规则,单击目标规则右侧操**作**列中的**编辑**,并按需修改规则内容。
- 如需删除规则,单击目标规则右侧操作列中的**删除**,并在对话框中单击**确定**。
- 如需启用或禁用规则,打开或关闭目标规则右侧**状态**列中的开关,并在对话框中单击**确定**。
- 如需批量启用或禁用规则,选中要操作的一个或多个规则,单击页面底部的**批量启用**或**批量禁用**,并在对话框中单击**确定**。

相关文档

- 服务提供方或消费方流控
- 削峰填谷
- 预热启动 (Warm Up)
- 关联限流

8.5.3.2. 配置降级规则

EDAS的降级功能可以监控应用下游依赖的响应时间或异常比例,当达到指定的阈值时立即降低下游依赖的优先级,避免应用受到影响,从而保障应用高可用性。

背景信息

除了流量控制以外,对调用链路中不稳定的资源进行熔断降级也是重要措施之一。由于调用关系的复杂性,如果调用链路中的某个资源出现了不稳定,最终会导致请求发生堆积。熔断降级功能会在调用链路中某个资源出现不稳定时(例如某资源出现超时或异常比例升高),对这个资源的调用进行限制,让请求快速失败,避免影响到其它的资源而导致级联错误。

降级规则配置通常用于弱依赖降级场景,更多信息,请参见弱依赖降级。

新建降级规则

- 1. 登录EDAS控制台。
- 2. 执行以下操作之一以进入应用详情页面:
 - 在左侧导航栏选择**资源管理 > 集群**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群页面的**容器服务K8s集群**页签上单击集群ID或名称,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**应用列表**页面单击具体容器服务K8s集群应用的应用名称。
- 3. 在左侧导航栏中选择限流降级 > 规则管理。
- 4. 单击降级规则页签, 然后在页面右上角单击新增降级规则。
- 5. 在新增降级规则对话框,设置规则参数,然后单击新建。



参数	描述
接口名称	待降级的资源名称。
统计窗口时长	统计请求数的窗口时长。
阈值类型	 慢调用比例:响应时间超过慢调用RT的请求数所占比例达到降级阈值时,资源进入降级状态。 异常比例:出现异常的请求数所占比例达到降级阈值时,资源进入降级状态。 ① 说明 降级阈值输入框中必须输入0~1的数字,表示比例。
慢调用RT	请求的响应时间超过该值则统计为慢调用,单位为ms。
降级阈值	触发熔断的慢调用比例阈值。

参数	描述
熔断时长	接口降级的时间。在该时间段内,该接口的请求都会快速失败。
是否开启	打开表示启用该规则,关闭表示禁用该规则。
最小请求数目	触发熔断的最小请求数目,若当前统计窗口请求数小 于此值,那么即使达到熔断条件规则也不会触发。
熔断恢复策略	 单次探测恢复:经过熔断时长后,AHAS会放过一个请求进行探测,若该请求符合预期(不为慢调用/没有异常)则结束熔断,否则重新回到熔断阶段。 渐进式恢复:经过熔断时长后,按照设置的恢复阶段数和每步最小通过数目进行渐进式恢复,若该阶段内指标未超过阈值则逐步提高允许通过的请求比例,直到请求完全恢复。(SDK 1.6.2及以上版本支持)

配置成功后,新的降级规则将出现在规则列表中。

管理规则

在规则管理页面的规则列表中,您可以编辑、删除、启用、禁用、批量启用或批量禁用规则。

- 如需编辑规则,单击目标规则右侧操作列中的编辑,并按需修改规则内容。
- 如需删除规则,单击目标规则右侧操作列中的删除,并在对话框中单击确定。
- 如需启用或禁用规则,打开或关闭目标规则右侧状态列中的开关,并在对话框中单击确定。
- 如需批量启用或禁用规则,选中要操作的一个或多个规则,单击页面底部的**批量启用或批量禁用**,并在 对话框中单击**确定**。

相关文档

• 弱依赖降级

8.5.3.3. 配置热点规则

为应用配置热点规则后,EDAS将分析统计热点参数(资源调用过程中的调用次数较高的参数),并根据配置的热点规则对包含热点参数的资源调用进行限流,保护系统稳定性。本文介绍如何为应用配置热点规则。

背景信息

热点即经常被访问的数据。在以下场景中需要统计某个热点数据中访问频次最高的Top K数据,并对其访问进行限制。例如:

- 针对一段时间内最频繁购买的商品ID进行限制,防止击穿缓存压垮数据库。
- 针对一段时间内频繁访问的用户ID进行限制, 防止恶意刷单。

EDAS利用LRU策略统计最近最常访问的热点参数,结合令牌桶算法来进行参数级别的流控。

新建热点规则

- 1. 登录EDAS控制台。
- 2. 执行以下操作之一以进入应用详情页面:
 - 在左侧导航栏选择**资源管理 > 集群**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集**

群页面的容器服务K8s集群页签上单击集群ID或名称,然后在集群详情页面的应用列表区域单击具体应用名称。

- 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**应用列表**页面单击具体容器服务K8s集群应用的应用名称。
- 3. 在左侧导航栏中选择限流降级 > 规则管理。
- 4. 单击热点规则页签,然后在页面右上角单击新增热点限流规则。
- 5. 在新增热点限流规则对话框中,配置规则信息。



参数	描述
接口名称	要进行热点限流的接口名称,需与埋点名保持一致。
参数位置索引	热点参数的索引,对应 SphU.entry(xxx, args) 中的参数索引位置。例如 SphU.entry(res ourceName, EntryType.IN, 1, paramA, param B) 埋点中, paramA的参数索引是0, paramB的参数索引是1。

参数	描述
统计维度	热点参数的统计维度,包含以下两种维度: • 通过请求数:根据统计周期内调用次数来进行限制。 • 并发数:根据统计周期内该参数占用最大并发资源的数量进行限制。
统计周期时间	统计窗口时间长度(单位为秒)。例如统计窗口时长为10秒,单机阈值为5,代表限制10秒内每个热点参数访问不超过5次。
单机阈值	每个热点参数的阈值。
流控效果	选择流控效果来处理被拦截的流量,包含以下两种处理方式: 快速失败: 达到阈值时立即拦截请求。该模式下需要设置缓冲请求数,即针对突发请求额外允许的请求数目。 排队等待: 阈值内的请求会匀速在周期内通过,多余的请求会排队等待,等待时长超过超时时间的会立即失败。 ① 说明 此选项仅适用于按并发数统计的情况。
是否开启	打开则表示启用该规则,关闭则表示禁用该规则。

配置成功后,新的热点限流规则出现在规则列表中。

管理规则

在**规则管理**页面的规则列表中,您可以编辑、删除、启用、禁用、批量启用或批量禁用规则。此外,还可以 针对指定参数值单独设置限流阈值。

- 如需编辑规则,单击目标规则右侧操作列中的编辑,并按需修改规则内容。
- 如需删除规则,单击目标规则右侧操**作**列中的**删除**,并在对话框中单击**确定**。
- 如需启用或禁用规则,打开或关闭目标规则右侧状态列中的开关,并在对话框中单击确定。
- 如需批量启用或禁用规则,选中要操作的一个或多个规则,单击页面底部的**批量启用**或**批量禁用**,并在对话框中单击**确定**。
- 如需为指定参数值单独设置限流阈值,单击目标规则右侧操作列中的添加例外项,并在热点规则例外 项对话框中添加例外项。

? 说明 仅可为基本类型和字符串类型的参数添加例外规则。

8.5.4. 实现K8s集群应用的限流降级

部署在容器服务K8s集群或Serverless K8s集群中的应用支持通过重新部署来接入AHAS的试用版,并可在 EDAS内嵌的监控页面中实时监控流量,全面可视化地保障您的应用可用性。本教程以在Dubbo应用中配置 限流降级规则为例,介绍了如何实现微服务的限流降级。

背景信息

EDAS中默认开通AHAS试用版,仅包含入门级防护,您只能为同一个应用类设置2条相同类型的防护规则,如需设置更多规则,请前往AHAS控制台开通专业版。

如果您想废弃或修改已创建的规则,您可以在应用总览页面的左侧导航栏选择**限流降级 > 规则管理**,选择要删除或修改的规则页签,然后在规则列表中选择规则来进行管理。

接入AHAS

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择**容器服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在应用总览页面右上角选择部署 > 部署。
- 4. 在选择部署模式页面的其中一种部署模式区域的右上角单击开始部署。
- 5. 在单批发布页面,修改部署参数,并展开限流降级配置。
- 6. 打开接入AHAS开关,然后单击确定。
 - □ 注意 如果该应用已经通过SDK方式接入了AHAS,打开此开关将会有冲突风险。

打开接入AHAS后,该应用将自动接入AHAS。在变更流程执行完毕后,您可前往监控 > 服务/接口监控页面配置限流降级规则。

配置限流降级规则

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏中单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择命名空间,在**集群类型**下拉列表中选择容器**服务/Serverless K8s集群**,然后在页面单击具体的应用名称。
- 3. 在左侧导航栏选择监控 > 服务/接口监控。
- 4. 在服务/接口监控页面单击限流降级页签。
 - 如果您已开通过AHAS,或在部署的应用中已通过SDK方式接入了AHAS,那么您可以在此页面直接观测到接口的请求流量并配置限流降级规则。
 - 如果您未开通过AHAS,将可以在此页面提示中单击**立即开通**来跳转到应用部署页面,通过重新部署 来开启AHAS试用版。具体操作,请参见接入AHAS。
- 5. 在限流降级页签选择请求链路并观测该请求链路的QPS和RT,然后根据观测结果来配置限流降级规则。
 - 您可在请求链路的**操作**列单击**限流**,在弹出的**限流规则**对话框中配置限流规则。具体操作,请参见配置流控规则。
 - 您可在请求链路的**操作**列单击**降级**,在弹出的**降级规则**对话框中配置降级规则。具体操作,请参见配置熔断规则。

支持组件列表

AHAS为了简化应用的接入流程,对主流框架进行了适配。本文将列出AHAS支持的第三方组件和框架列表。

组件支持版本	从EDAS直接接入	配置SDK接入
Dubbo 2.5.x+	2.5.x和2.6.x版本已支持	全部支持
Web Servlet	已支持	支持SDK 2.x或以上版本接入
Spring Boot 1.3.x+	已支持	参见 <mark>说明</mark>
Spring MVC 4.x+	已支持	参见 <mark>说明</mark>
Spring Cloud Gateway 2.x	已支持	参见说明
Zuul 1.3.x	已支持	参见 <mark>说明</mark>
GRPC-Java 1.13+	已支持	需另引入sentinel-grpc-adapter依赖
Jetty 8.x+	已支持	通过Servlet支持
Tomcat 7.x+	已支持	通过Servlet支持
WebLogic 10.3	已支持	通过Servlet支持
HttpClient 3.x+	不支持	不支持
HttpClient 4.x+	不支持	不支持
JDK HTTP 1.7.x+	不支持	不支持
OKHttp 2.x+	不支持	不支持
MyBatis 3.x+	不支持	支持SDK 1.4.1或以上版本接入
MySQL JDBC 5.0.x+	已支持	不支持
Oracle JDBC 12.x	已支持	不支持
PostgreSql JDBC 9.4+	不支持	不支持
SQLServer JDBC 6.4+	不支持	不支持
Redis Client (Jedis)	已支持	不支持
MemCached 2.8+	已支持	不支持
MongoDB 3.7+	不支持	不支持
RocketMQ(callback模式)4.x	已支持	需手动埋点
RabbitMQ 3.7+	已支持	需手动埋点
Feign 10.x+	已支持	需引入spring-cloud-starter- alibaba-sentinel

? 说明

- 支持的AHAS Agent版本和SDK版本,请参见版本说明。
- Spring MVC、Spring Boot和Spring Cloud Web应用需要引入 ahas-sentinel-client 并进行相 应配置即可使用(参考Web Servlet支持的版本);若需要通过*Spring properties*配置启动参数,则需要单独引入 spring-boot-starter-ahas-sentinel-client 依赖。
- Spring Cloud Gateway网关需要引入 spring-cloud-gateway-starter-ahas-sentinel 依赖; Zuul 1.x网关需要引入 spring-cloud-zuul-starter-ahas-sentinel 依赖, 无需引入其它依赖。
- Dubbo 2.7.x以上版本暂不支持在EDAS控制台部署接入AHAS,您可通过在应用中添加防护依赖的方式来接入AHAS。具体操作,请参见接入Dubbo应用。

8.5.5. 查看限流降级操作日志

EDAS提供了查看限流降级操作日志功能,您可以查看所有的限流降级操作信息,如操作时间、类别、操纵内容和操作等。

操作步骤

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在**容器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧的导航栏中选择限流降级 > 操作日志。

在**操作日志**页面向您展示了限流降级操作的相关信息,如操作类型、资源名、操作内容、操作时间和操作人ID。

8.6. 变更记录

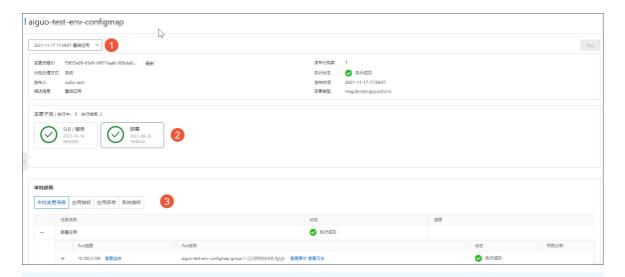
8.6.1. 查看应用变更

当您在EDAS上进行应用部署、启动、扩容、缩容等生命周期操作后,您可以跳转到应用详情页查看当前变更状态,也可以通过变更记录页面查看该应用的历史变更记录。

查看应用变更详情

下面以一次部署应用为例介绍如何查看应用变更。

- 1. 在执行完应用变更操作后,返回**应用总览**页面。 在**应用总览**上方,会提示**应用正在变更...查看详情**。
- 2. 单击提示信息中的查看详情,查看该应用的变更过程信息。



? 说明

- 当流程状态为执行中、执行成功或失败时,如果此时状态不符合您的预期,可以通过单击右 上角的**立即回滚**按钮将应用实例的版本回退到上一个版本。
- 如果需要终止回滚,可以单击右上角的**终止**按钮。终止回滚后,可以部署一个期望的应用版本来进行修复。部署应用的描述,请参见创建和部署应用概述(K8s)。
- 发布单变更信息(图示中①): 默认展示最新一次变更的概要信息,您也可通过展开下拉列表来选择希望展示的历史变更流程,单击击近三个月变更记录,可查看更多变更记录。展示的信息包括变更流程ID、发布批数、分批处理方式、执行状态、发布人等信息。详细信息,请参见流程变更信息说明。
- 变更子项(图示中②):展示此次变更所涉及的资源种类。蓝色表示资源正在变更,绿色表示资源变更成功,红色表示资源变更失败。详细信息,请参见变更子项资源类型及其描述。 发布单正在执行中,页面会定时请求后端状态并刷新页面。如果页面刷新影响您问题排查或事件观测,可关闭页面中部右侧的**状态刷新**开关。需要注意的是,当关闭状态刷新后,页面会停止刷新,但后端仍会在变更中。
- 变更资源的详细信息(图示中③):展示本批变更信息、资源信息、应用指标、应用异常与系统指标。在变更过程中可通过该模块来获取部署状态信息,方便您定位可能出现的问题。相关描述,请参见变更资源的详细信息。

流程变更信息说明

名称	描述
变更流程ID	此ID是该发布单的唯一标识。当出现预期外的未知异常时,您可提交此信息给EDAS支持人员以协助问题排查。
发布分批数	当您应用内的实例数大于1时,可采取分批部署(或金丝雀部署)来逐批变更实例。此字段为具体采取的分批数量。
分批处理方式	当采取分批部署(或金丝雀部署)时,批次之间可采取自动或手动的方 式来进行下一批实例发布。

名称	描述
执行状态	变更流程状态分为执行中、执行成功、失败和终止。 ■ 执行中: 变更开始的状态为执行中。 ■ 成功: 变更执行过程无异常。 ■ 失败: 执行过程中遇到异常,如镜像无法拉取、机器资源不足等,导致变更失败。 ■ 终止: 长时间阻塞时,您可通过执行回滚操作来停止变更,回滚中可以通过单击 <i>终止</i> 按钮来停止回滚。
发布人	此次变更的操作人。
发布时间	此次变更的开始时间。
描述信息	此次变更的版本描述信息。
变更类型	此次变更的操作类型。

变更子项资源类型及其描述

资源类型	描述
部署 (Rollout)	用于控制底层K8s工作负载(如无状态负载Deployment等)。
SLB/服务 (ServiceTrait)	用于配置K8s服务或SLB负载均衡。
镜像构建(ImageBuilder)	用于WAR/JAR类型应用的镜像构建。
日志采集(LogCollector)	用于配置日志收集。
动态标(DynamicLabel)	用于全链路灰度配置。
弹性伸缩(AutoScaling)	用于动态弹性扩缩。

变更资源的详细信息

标签名	描述
本批变更信息	展示本次变更的底层Pod信息,可通过展示的Pod列表进一步查看其监控、事件或日志,并且确认当前的执行状态。如果Pod变更失败,那么可点击失败分析来一键获取失败信息。异常情况排查,请参见 <mark>变更流程问题排查指南</mark> 。
资源信息	展示本次变更资源的事件和详情。根据资源的关系,左侧将其组织为类似于目录的树状结构,上一层资源负责创建并控制下一层资源。最上方的一级目录为控制整个应用变更的资源。您可点击目录上的任一资源节点来查看具体事件或资源详情。
应用指标	展示应用的整体指标,包括响应时间、请求数、错误数等指标。
应用异常	展示应用的异常信息。

标签名	描述
系统指标	展示应用部署前后的CPU和内存信息。

查看应用变更记录

- 1. 登录EDAS控制台。
- 2. 在应用总览页面左侧的导航栏中单击变更记录。

在**变更记录**页中展示应用所有变更信息,如创建时间、时间结束、变更类型、变更状态、变更人等。 如果您还需要进一步了解变更记录的详情,可以单击操作列的查看。

8.6.2. K8s集群应用部署失败的原因分析

当您在EDAS的K8s集群中创建、部署或扩容应用后,可以跳转到应用详情页面查看变更状态,同时还可以在变更记录中查看和分析应用部署失败的常见原因。

查看部署失败的执行结果

- 1. 在执行完应用部署操作后,返回应用总览页面。 在应用总览上方,会提示应用有变更流程正在执行,处于执行中状态。
- 2. 在左侧导航栏单击变更记录,选择变更流程并在操作列单击查看。
- 3. 在任务名称区域可以查看应用变更过程以及执行结果。

如果提示任务执行失败,您可在失败原因列单击点击查看,在弹出的对话框中查看部署失败的原因。



常见部署失败原因

导致应用部署失败的常见原因总结如下:

● 资源不足导致部署失败:



应用容器启动失败导致部署失败:



PostStart启动报错导致部署失败:



• Readiness探针探活报错导致部署失败:



相关文档

● 查看应用变更

8.7. 事件中心

8.7.1. 事件中心

事件中心将云产品所生成的事件数据进行统一管理、存储、分析和展示,已接入EDAS的变更事件、ARMS的报警事件、0-1事件(如死锁、OOM和应用启动等)、MSE的微服务管控事件和K8s集群事件。当您的应用使用了相关的产品,对应的事件会自动接入事件中心进行统一的分析展示,方便您查看与分析。

事件模型

事件中心的一个事件主要由以下参数来定义:

参数	是否必须	描述
source	是	事件来源
type	是	事件类型
level	是	事件等级
time	是	事件发生时间
data	是	事件体(一般为JSON格式)
PID	否	Pod ID
IP	否	IP地址
ClusterId	否	集群ID
PodName	否	Pod名称

进入事件中心

- 1. 登录EDAS控制台。
- 2. 登录EDAS控制台,在左侧导航栏单击**应用列**表,在顶部菜单栏选择地域并在页面上方选择微服务空间,在集群类型下拉列表中选择容器服务/Serverless K8s集群,然后单击目标应用名称。
- 3. 在左侧导航栏单击事件中心。

事件中心模块介绍

事件中心主要包含典型事件和三个页签(分别是普通视图,拓扑视图和订阅规则)。

- 典型事件:展示系统预置的典型事件的数量。更多信息,请参见典型事件。
- 普通视图将与当前应用关联的所有事件进行简单的多维度分析与展示。更多信息,请参见普通视图。
- 拓扑视图将当前应用关联的事件和与当前应用的资源拓扑图进行结合展示。更多信息。请参见拓扑视图。
- 订阅规则以列表形式展示了当前您创建的订阅规则。更多信息,请参见订阅规则。

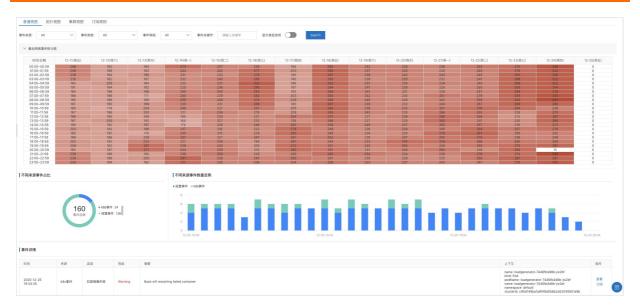
典型事件

典型事件:展示系统预置的典型事件类型在最近30分钟(可在右上角调整时间范围)内发生的次数。在事件下方单击**订阅**,可以编辑该事件的订阅规则。订阅规则的操作,请参见<mark>订阅规则</mark>。



普通视图

普通视图可以按照您指定的检索条件进行搜索,搜索的结果会以四个视图进行展示,分别是:最近两周事件热力图、不同来源事件占比、不同来源事件数量走势和事件详情。



- **最近两周事件热力图**展示近两周内满足过滤条件的事件发生次数按小时统计的热力分布。颜色越深,说明该小时内发生的事件数量越多。
- ◆ 不同来源事件占比展示不同来源事件的数量占比。
- ▼ 不同来源事件数量走势展示不同来源事件在选定时段内的走势。
- 事件详情展示当前所有事件的详情列表。

普通视图查看说明:

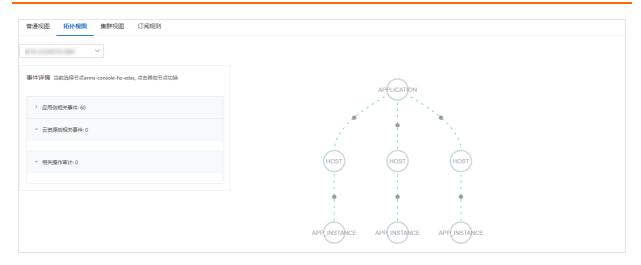
- 单击最近两周事件热力图中的热力方块,可查看该小时内的事件详情。
- 单击不同来源事件数量走势中的立柱,可在事件详情区域查看对应时段内的所有事件列表。
- 单击事件详情右侧的查看,可以查看事件的具体内容。
- 单击事件详情右侧的订阅,可以订阅指定事件。更多信息,请参见<mark>订阅规则</mark>。

拓扑视图

拓扑视图首先会绘制出该应用的资源拓扑,包含该应用使用的ECS,该应用部署的实例Pod,该应用使用的所有RDS和Redis等中间件资源,以及该应用挂载的SLB和NAT等。之后会将获取到的关联事件、操作审计、云监控事件关联到对应的拓扑节点上,单击相应节点,会在左上角的事件详情区域展示与该节点关联的事件:

- 应用侧相关事件: 即事件中心的事件。
- 云资源侧相关事件:即云监控的相关事件。
- 相关操作审计: 即来自于操作审计的审计记录。

该视图可以帮助您在应用出现故障时,快速排查关联的资源各自发生了什么问题。例如在大型企业中,由于某个员工的误操作,重启了生产环境的RDS,导致线上业务故障,利用该视图,可以快速的发现应用访问的RDS出现了重启操作。



订阅规则

订阅规则页签展示您当前所有的订阅规则,您可以在该页面启用、禁用或者修改规则。

□ 注意 您只能修改您自己在控制台创建的规则,无法修改系统内部自动创建的规则。

订阅规则是一条表示您订阅满足指定条件的事件,并将该事件发送到指定Webhook的规则。创建订阅规则有两种方式:

- 方式一: 在订阅规则页签右上角单击新建订阅规则。
- 方式二:在普通视图页签的事件列表区域单击事件的订阅。
 - 1. 单击订阅规则页签, 在页面右上角单击新建订阅规则。
 - 2. 在新建订阅规则面板的填写基本信息页面输入规则名称和规则描述,然后单击下一步。
 - 3. 在选择事件模式页面设置事件规则参数,然后单击下一步。

参数	描述
事件来源	在下拉列表中选择事件来源。
事件类型	在下拉列表中选择事件类型。
事件等级	在下拉列表中选择事件等级。
事件关键字	在文本框中输入事件关键字。
显示高级过滤选项	默认关闭,开启后可设置 集群ID、接口/服务名、主机IP和POD名 的过滤选项。
自定义过滤条件	自定义过滤条件一般用来指定事件体JSON的某个字段需要满足的条件,根节点是data,以 的形式下钻事件体JSON的某个字段。请输入自定义的过滤条件,最多可以设置6条过滤条件。
选择有效字段	以data.x.y的形式选择有效字段,选择完成后,请给选择的字段输入一个别名,该别名可用于在填写Webhook信息的Post请求体中以占位符的形式出现。最多可以设置6条有效字段。

参数	描述
消息通知模板	消息通知模板会作为当指定消息发生时通知给您的内容(如果通知对象为钉钉机器人Webhook,请注意在消息模板中包含创建钉钉机器人Webhook时的关键字)。

4. 在选择联系人页面选择联系人, 然后单击提交。

如果选择联系人的列表中没有联系人信息,在右侧单击新建联系人,创建联系人后再在列表中选择。

- 1. 在普通视图页签的事件详情区域,单击事件详情右侧的订阅。
 - 以该种方式创建的订阅规则,会根据您选择的事件来自动选择事件来源、事件类型和事件等级三个过滤 条件。
- 2. 单击事件JSON文件中某个字段 *Value*的形式来选择自定义过滤条件和有效字段。每次单击会自动生成一个过滤条件和选择字段,同时您还可以手动修改或者删除字段来调整规则,完成事件模式设置后单击下一步。
- 3. 在**选择联系人**页面选择联系人,然后单击**提交**。 如果**选择联系人**的列表中没有联系人信息,在右侧单击**新建联系人**,创建联系人后再在列表中选择。

新建订阅规则

在普通视图中订阅事件

常见问题

对于部署在EDAS的容器服务K8s集群或Serverless K8s集群中的应用,运行时出现的问题一般是Pod问题。解决方法,请参见常见Pod问题。

8.8. 日志管理

8.8.1. 查看实时日志

您可以查看在K8s集群中部署的应用的实时日志。

操作步骤

当应用出现异常情况的时候,可以通过查看实时日志来排查容器 (Pod) 相关问题。

- 1. 登录EDAS控制台。
- 2. 按需执行以下任一操作来进入应用的详情页面:
 - 在左侧导航栏选择**资源管理 > 容器服务K8s集群**(或**资源管理 > Serverless K8s集群**),在顶部菜单栏选择地域并在页面上方选择微服务空间,在容**器服务K8s集群或Serverless K8s集群**页面单击集群ID,然后在**集群详情**页面的**应用列表**区域单击具体应用名称。
 - 在左侧导航栏单击**应用列表**,在顶部菜单栏选择地域并在页面上方选择微服务空间,在**集群类型**下 拉列表中选择**容器服务/Serverless K8s集群**,然后单击目标应用名称。
- 3. 在左侧导航栏中单击日志中心。
- 4. 在日志中心页面单击实时日志(标准输出)页签。
- 5. 在**实时日志(标准输出)**页签的**Pod名称**下拉框内选择要查看日志的Pod,查看实时日志。如果有需要,可以选择开启**自动更新**开关,每10秒更新一次日志。

8.8.2. 查看日志目录

当您的应用出现异常情况的时候,可以通过查看实例和应用级别的日志来排查问题。EDAS提供了日志目录收藏、日志查看、日志搜索的功能。

收藏日志目录

日志目录页面包含EDAS相关的默认日志目录,您可以收藏(添加)应用日志目录。

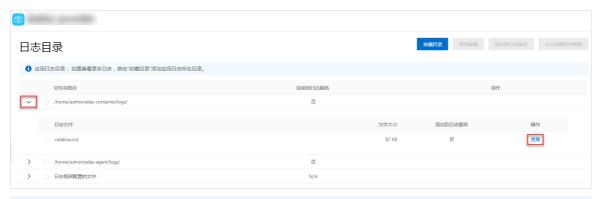
在收藏日志目录时还可以将该目录添加到日志服务,以便在日志搜索页面查看和搜索该目录下的应用日志。

- ? 说明 收藏目录和取消收藏仅对日志目录可用。
- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击应用列表,在顶部菜单栏选择地域并在页面上方选择微服务空间。
- 3. 在**应用列表**页面**集群类型**的下拉列表中,选择**容器服务/Serverless K8s集群**,然后在**应用列表**页面单击目标应用的应用名称。
- 4. 在左侧导航栏中单击日志中心, 然后在日志中心页面单击日志目录(在线查看)页签。
- 5. 在日志目录(在线查看)页签右上角单击添加在线查看。
- 6. 在**添加在线查看**对话框中输入**应用日志目录**,选择是否将该目录添加到日志服务,然后单击**添加**。 输入应用日志目录时,请遵循以下要求:
 - 完整目录中必须包含log或者logs。
 - 目录最后必须以斜杠/结尾,表示添加的是一个文件夹。

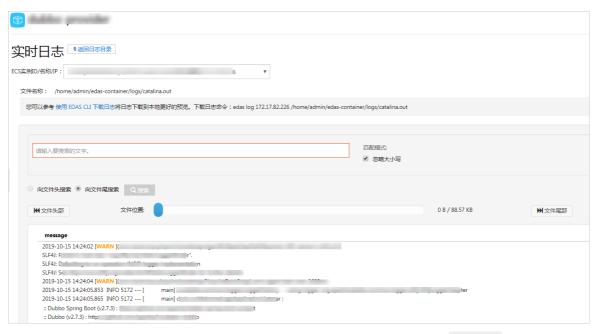
杳看实例日志

在收藏了应用日志后,可以查看实例和应用级别的日志。

1. 在**日志目录(在线查看)**页面单击日志目录左侧的展开按钮,然后在展开的日志文件列表的**操作**列单击查看。



- ② 说明 日志框架配置的文件下的日志文件除了可以查看日志详细信息,还可以修改该日志文件的级别。日志级别由低到高依次为:TRACE、DEBUG、INFO、WARN和ERROR。修改后,系统会将所设置的级别及更高级别的日志显示在页面下方。
- 2. 在**实时日志**页面顶部单击**ECS实例ID/名称/IP**右侧的下拉箭头,选择实例,查看该实例中的日志详细信息。



在页面右下角单击开启实时追加,可以一直加载文件的最新追加内容(类似于 tailf 命令的效果)。

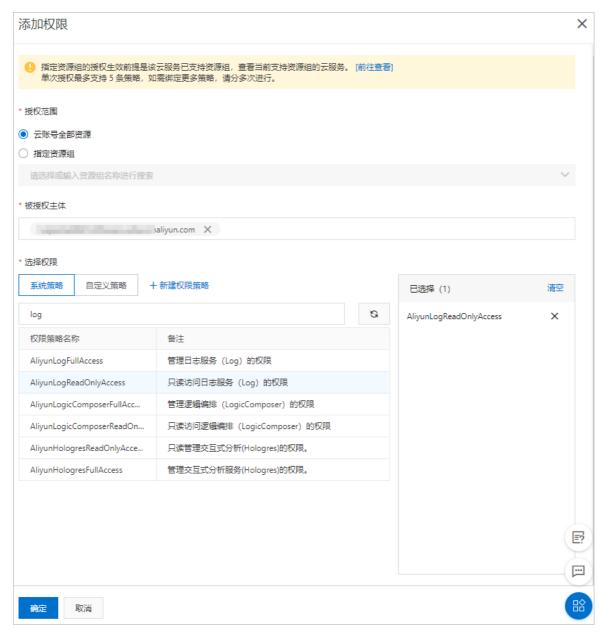
为RAM用户授予日志服务权限

如果您是RAM用户,需要使用日志服务(包括将日志目录或文件添加到日志服务、查看应用日志和分布式搜索),还需要由阿里云账号在访问控制RAM中为RAM用户授权。操作步骤如下:

- 1. 使用云账号登录RAM访问控制控制台。
- 2. 在左侧导航栏中选择**人员管理 > 用户**,然后在**用户**页面的用户列表中找到您要授权的RAM用户,在**操作**列单击**添加权限**。



- 3. 在**添加权限**页面单击**系统策略**页签,在下方的文本框输入*log*,单击**AliyunLogReadOnlyAccess**将该权限添加到右侧列表中,单击**确定**。
 - ② 说明 被授权主体会默认加载,无需设置。如果您需要同时给多个RAM用户授权,在被授权主体下面的文本框通过关键字搜索并添加。



4. 在授权结果页面查看RAM用户和被授权的权限,确认无误后,单击完成。

取消收藏日志目录

取消收藏日志目录即移除指定的应用日志目录。移除后,该日志目录将不再显示在**日志目录**页面上,也不能再查看该目录下的实例日志,但并不会删除实际的日志目录及文件。

- ? 说明 默认目录也可以取消收藏。
- 1. 在日志目录页面选中某个日志目录,然后单击取消收藏。
- 2. 在**取消收藏日志目录**对话框中,确认要取消收藏的日志目录,根据需要选择**并从日志服务中移除**,单击**确**定。

在取消收藏日志目录时,您可以选择并从日志服务中移除。

- 如果未从日志服务中移除,则可以查看该目录下的原有应用日志。
- 如果从日志服务中删除,则不能再查看该目录下的应用日志。

8.8.3. 查看文件日志

EDAS对接了日志服务SLS。如果在EDAS的容器服务Kubernetes集群创建或部署应用时开启了日志服务功能,则可以查看该应用的文件日志。本文介绍如何为应用开启日志服务SLS及查看应用的文件日志。

前提条件

- 开通日志服务。
- 确保应用中每个实例预留了0.25核CPU和25 MB内存的可用资源。
- 已为应用开启日志服务SLS。 在容器服务Kubernetes集群中,日志服务需要在应用创建或部署时开启。具体操作,请参见在K8s集群中使用镜像部署|ava微服务应用或在容器服务K8s集群中使用|AR包或WAR包部署应用。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏单击应用列表。
- 3. 在顶部菜单栏选择地域,并在页面上方选择微服务空间,然后在**应用列表**页面单击开启了日志服务的应用名称。
- 4. 在应用总览页面左侧的导航栏单击日志中心,然后单击文件日志(SLS采集)页签。
- 5. 在文件日志(SLS采集)页签下的文件日志列表中的操作列,单击查看文件日志。
- 6. 在日志服务控制台查询及分析日志。具体操作,请参见查询与分析。

8.9. 应用管理

8.9.1. 导出应用

EDAS支持将应用的配置和程序包导出到本地,帮助您减少运维操作成本。

前提条件

已完成K8s应用部署。具体操作,请参见:

- 在K8s环境中通过镜像部署微服务应用
- 在K8s环境中通过镜像部署多语言微服务应用
 - ② 说明 当前仅支持导出K8s应用。导出应用的数据详情描述,请参见应用信息。

使用场景

- 从公有云EDAS环境迁移应用到其他类型的EDAS环境,例如迁移应用到金融云EDAS环境。
- 公有云EDAS环境中跨账号迁移应用。例如从当前账号导出应用,然后导入到其他账号。

批量导出应用

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择应用管理 > 应用列表。
- 3. 在应用列表页面集群类型的下拉列表中,选择容器服务/Serverless K8s集群。
- 4. 勾选需要导出的应用,在应用列表页面下方,单击批量导出应用。

- ? 说明 当前单次导出的应用数不能超过10个。
- 5. 在导出应用的对话框,单击导出。

您可在**应用列表**页面右上方单击**导入/导出记录**查看应用导出进度和详情信息。成功导出后,即可下载应用包和程序包。

应用信息

当前导入、导出应用数据时,包含的应用信息如下所示:

- 应用名、应用描述信息、应用所处集群信息、应用实例数
- 应用程序包、应用镜像
- CPU、内存请求与限制值
- 启动命令、启动参数、环境变量配置
- 就绪探针、存活探针、启动后脚本、停止前脚本
- 主机目录挂载、emptyDir挂载
- 微服务空间信息、K8s命名空间信息
- 应用运行时: JDK、TomCat、edas-container信息
- IVM启动参数
- 调度亲和性配置
 - ② 说明 不在以上列表中的应用信息暂不支持导入、导出、克隆操作。

8.9.2. 导入应用

您可将本地的应用上传到EDAS进行管理。

前提条件

导出应用

使用场景

- 从其他公有云EDAS环境迁移应用到当前的EDAS环境,例如迁移应用到金融云EDAS环境。
- 公有云EDAS环境中跨账号迁移应用。例如从其他账号导出应用,然后导入到当前账号。

批量导入应用

- 1. 登录EDAS控制台。
- 2. 在左侧导航栏选择应用管理 > 应用列表。
- 3. 在应用列表页面上方单击导入应用。
- 4. 在导入应用页面选择应用文件并上传程序包(可选)。

? 说明

- 在同一个EDAS环境中迁移应用,例如公有云应用迁移到其他公有云、金融云应用迁移到其他金融云,此时您无需上传程序包。
- 。 仅当跨云环境迁移应用时,例如公有云迁移到金融云,您需要执行上传程序包操作。

- 5. 单击确定。
- 在导入确认页面确认需要导入的应用,单击导入。
 您也可以使用快速导入功能,只需填写应用的微服务空间、集群等资源参数,即可完成导入。

8.9.3. 克隆应用

您可以克隆应用实现应用的复制和迁移。

前提条件

已部署应用。

操作步骤

- 1. 登录EDAS控制台。
- 2. 在应用列表页面勾选需要克隆的应用。
 - ? 说明 当前应用的克隆数量不能超过10个。
- 3. 在应用列表页面下方单击批量克隆应用。
- 4. 在克隆确认对话框确认完毕应用信息后,单击确定。
- 5. 查看应用的微服务空间、集群等资源参数,配置**副本**数目,单击**确定**。 您可在**应用列表**页面上方单击**导入/导出记录**查看应用克隆进度和详情信息。