



Serverless 应用引擎 应用管理

文档版本: 20210903



法律声明

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

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

通用约定

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

目录

1.应用管理概述				
2.应用访问				
2.1. 管理SLB	09			
2.1.1. SLB使用说明	09			
2.1.2. SAE SLB配置实践	09			
2.1.3. 为应用绑定SLB	10			
2.2. 配置网关路由	14			
2.2.1. 使用约束	14			
2.2.2. 为应用配置网关路由	15			
3.应用设置	21			
3.1. 变更实例规格	21			
3.2. 切换安全组和vSwitch	22			
3.3. 复制应用	24			
3.4. 管理标签	25			
3.4.1. 为应用添加标签	25			
3.5. 一键启停应用	27			
3.6. 管理定时启停规则	30			
4.应用运维	37			
4.1. 管理应用生命周期	37			
4.2. 重启应用实例	39			
4.3. 删除应用实例	40			
4.4. 使用Webshell诊断应用	42			
4.5. 应用变更记录	46			
4.5.1. 查看应用事件	46			
4.5.2. 查看变更记录	47			
4.6. 远程调试	49			

4.6.1. Java远程调试	49
4.6.2. 实例远程访问及上传下载文件	54
5.服务治理	59
5.1. 限流降级(Java)	59
5.1.1. 规则管理	59
5.1.1.1. 配置流控规则	59
5.1.1.2. 配置隔离规则	64
5.1.1.3. 配置降级规则	68
5.1.1.4. 配置系统规则	71
5.1.1.5. 配置热点规则	73
5.1.2. 设置限流降级	75
5.1.3. 查看Java应用概览信息	77
5.1.4. 查看节点监控指标	78
5.1.5. 查看操作日志	81
5.2. 微服务灰度流量	81
5.2.1. 灰度发布微服务	82
5.3. 优雅下线	86
5.3.1. 设置微服务无损下线	86
6.升级和回滚应用	91
6.1. 管理灰度规则(Java)	91
6.2. 灰度发布应用	93
7.分布式配置管理	99
7.1. 配置管理概述	99
7.2. 创建或导入配置	99
7.3. 查看推送状态	104
7.4. 查询推送轨迹	104
7.5. 查询和回滚版本	105
7.6. 创建和使用加密配置	106

8.手动和自动扩缩	109
8.1. 手动扩缩	109
8.2. 配置弹性伸缩策略	111

1.应用管理概述

应用托管到SAE以后,通过SAE控制台对应用本身及其相关的资源进行管理,如SLB绑定、应用生命周期管理、应用事件查看、弹性配置等。

绑定SLB 通过添加SLB实现对的应用访问。绑定SLB时您可以复用您原有 的←B 也可以使用通过SAE帮您代购新的SLB,进行TCP或者 HTTPS监听。
应用生命周期管理 通过SAE对已部署的应用进行启停、更新、删除和扩缩容等应用 生命周期管理操作。 ┃ ┃ ┃
管理命名空间 命名空间以应用的服务调用与分布式配置推送为视角,为应用运 行提供逻辑隔离环境。
查看应用事件信息 SAE支持查看K8s原生的应用事件,帮助您了解应用运行时的状态,方便快速聚焦问题。
查看变更记录 记录SAE应用在生命周期内所有的操作记录,便于管理应用及定 (分词) (分词) (分词) (分词) (分词) (分词) (分词) (分词)
変更 实例规格 SAE支持按需修改应用实例规格,确保应用正常运行。 ●●●●
配置弹性伸缩 通过自动弹性伸缩实现应用实例数高效利用,降低应用资源成 本。

- 健尼停应用 - 理言行应用环境,即开即用,大幅度降低资源成本。
配置网关路由 加里你的业务违式重要公告经甘他服务或老应用,现么你可以为
如未忘的亚另谓不需安力及结实他服另或省应用,加互忘可以为 应用配置网关路由,实现请求路由分发。
配置管理
SAE集成了应用配置管理(ACM),方便您对应用配置进行集中
通知报警
SAE集成了ARMS的通知报警功能,方便您实时掌握应用动态。
使用Webshell诊断应用
SAE提供了Webshell切能,帮助您完成基本的运维需求。
标签管理
随着创建的应用逐渐增多,您可以利用标签将资源进行分组管理 和归类便于搜索和资源聚合。

2.应用访问

2.1. 管理SLB

2.1.1. SLB使用说明

在SAE中部署应用后,可以通过添加公网SLB(Server Load Balancer,负载均衡)实现公网访问,也可以添加私网SLB实现同VPC内所有应用间互相访问。在绑定SLB前您可以了解SLB的原理、限制等内容,让您更好的使用SLB功能。

SLB原理

SLB为阿里云产品,用于流量负载均衡,具体详情请参见负载均衡产品文档。

使用限制

在SAE中应用访问是通过SLB实例来实现,SLB实例中的一条监听将您的业务暴露,以便被访问,但在使用 SLB时还存在一定的约束条件。

SLB使用约束条件如下所示:

- SLB侧的监听配置请勿同时配置在SAE侧。
 SAE为您提供的实例为非ECS容器,具有易变、无状态的特性。因此,SLB转发监听流量到SAE实例的配置
 需要经常变更。您的应用每次部署,以及SAE每次重启,都会触发一次重新配置。从而,请勿与SLB产品
 同时配置一条监听,在SAE产品侧的默认配置以及您的配置将会覆盖SLB产品侧的配置。
- 通过SAE购买的SLB实例在配置移除后,该SLB实例会被释放,无法购买与之相同的实例,配置移除时请谨慎。

如果其他产品也有类似购买以及回收的行为,复用SLB时请慎用,不建议复用,建议通过SAE购买新的SLB 实例或者在SLB侧购买新的SLB实例。

- 共享型SLB实例无法在SAE中使用。
- 其他
 其他SLB限制请参见使用限制。

2.1.2. SAE SLB配置实践

在SAE中使用SLB时有一些操作是有一定风险的,例如跨过SAE、通过SLB控制台或者OpenAPI方式直接配置 SLB实例,并进行自定义配置,例如更改监听名称、更改后端服务器组名称和添加额外服务器组等。本文介 绍SLB在SAE中的配置边界,即哪些可以自定义配置、哪些禁止自定义配置。

SLB配置概述

SAE对SLB的管控为托管式,帮助您配置SLB实例监听,并在部署、重启、扩缩容等场景协助您维护该配置的 正确性;而在SLB产品侧对其进行配置动作,SAE将视为干扰配置,将放弃后续配置或者覆盖掉SLB原来的配 置,甚至会出现无定义的行为。因此,在使用SAE时,强烈建议将SLB的监听配置功能全权由SAE完成。

对于SAE不支持的场景,您可以在SLB侧配置,但是存在一定的界限,详情请参见<mark>SAE代购的SLB实例和SAE</mark> 托管的SLB实例。

SAE代购的SLB实例

SAE代购的SLB实例信息中包含sae.do.not.delete前缀。

当您删除应用或为应用解除SLB绑定时,该实例会被立即释放,且不可恢复。因此,该类SLB实例不建议您对 其进行管控。如果有特殊临时需求,那么请参见SAE托管的SLB实例。

同理,其他云产品亦不可管控此类SLB。

SAE托管的SLB实例

SAE托管的SLB实例,即由用户购买并交由SAE代为配置的实例,通常单个应用只会管控一个监听。对于您所 创建监听,您可以对其进行任何配置,属于合法行为(如创建新的监听、将流量导向SAE托管的监听或SAE 创建的虚拟后端服务器组)。但是对于SAE创建的监听以及监听关联的虚拟后端服务器组,如果您需要进行 修改,请检查操作的合法性,合法性如下表所示。

类型	修改项	合法性	
实例	实例名称	Υ	
	实例标签	禁止以下操作: ● 更改SAE创建的标签 ● 删除SAE创建的标签	
	EIP绑定	Ν	
	升降配	Υ	
	带宽	Υ	
	监听名称	Ν	
	带宽	Υ	
	调度算法	Υ	
	虚拟服务器组ID	Ν	
	访问控制	Υ	
监听	添加请求x-forward-for字段	Υ	
	gzip压缩	Υ	
	会话保持	Υ	
	健康检查	Υ	
	超时时间	Υ	
	证书配置	Ν	
虎拟服务器组	名称	Ν	
应 1以加入了 码 扫	后端服务器(权重、实例、端口等)	Ν	

2.1.3. 为应用绑定SLB

在SAE中部署应用后,您可以通过添加负载均衡SLB(Server Load Balancer)实现公网访问,也可以添加私网SLB实现同VPC内所有应用间的互相访问。本文介绍如何为应用绑定SLB。

前提条件

为应用绑定SLB前,请先了解以下文档:

- SLB使用限制
- SLB使用说明
- SAE SLB配置实践

不同场景下绑定SLB的前提条件如下所示:

- 场景一: 绑定已有SLB
 - i. 上传已有SSL证书
 - ii. 在SAE控制台部署应用
 - ↓ 注意 在SAE中复用的SLB实例需要满足以下条件:
 - 必须为非性能共享型SLB实例。
 - 必须为非容器服务独占的SLB实例。
 - 必须为通过SLB控制台购买的SLB实例。SAE不复用其他产品代购或者独占的SLB实例,以 防出现监听配置冲突。
 - 必须与部署在SAE上的应用所在的实例处于同一个VPC内。
- 场景二: 绑定新建SLB
 - i. 上传已有SSL证书
 - ii. 创建SLB实例
 - iii. 在SAE控制台部署应用

场景一: 绑定已有SLB

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面默认显示的基本信息页签,找到应用访问设置区域绑定SLB。
 - 添加私网SLB: 单击添加私网SLB访问。
 - 添加公网SLB: 单击添加公网SLB访问。

本文以添加公网SLB访问为例。

✓ 息用の同設置			
●使用SAE SLE功能之前,请先邀署SLE成置的来 典型SLE场展展性实践 自2020年4月1日,凡是在SAEPA员中创建编辑出目后,再去SLEPA员中将改造所编口、协议、SSL证书、SLE tag、虚拟服务器组合等SLE成置约束中限制股件的资源,终改后的或置部会被SAE保存的或置信息强制原盖,请一定注意避免。			
私网动间地站 漫加公网518访问 公网动间地址 漫加公网518访问			

a. 单击添加公网SLB访问。

添加公网SLB访问			×
 1. 设置公照负载均衡,能通过公网访问您的应用。您可以选择已有的51.8服务,也可以创建新的51.8。 2. 选择新创建后,系统全为您的应用目动购买一个公网51.8服务,按使用量计费,详双51.8件量件情,购买的51.8倍意可以在负载均衡控制台查看。 			
请选择SLB: 8.	5d1) >> SLB复用说明 ?		
TCP协议 HTTP协议	化 HTTPS协议		
网络协议	SLB第一	容凝病口	操作
TCP	80	8080	Î
+添加下一条监听			
			稳定 取消

b. 在添加公网SLB访问对话框的请选择SLB的下拉列表中选择已有的SLB。

- c. 配置SLB监听端口。
 - TCP协议:
 - SLB端口:提供公网访问应用的SLB端口,可设置范围为1~65535。
 - 容器端口: 进程监听端口, 由程序定义, 例如: Web服务默认使用8080端口。
 - HTTP协议:
 - HTTP端口:提供公网访问应用的SLB端口,可设置范围为1~65535。
 - 容器端口: 进程监听端口, 由程序定义, 例如: Web服务默认使用8080端口。
 - HTTPS协议:
 - HTTPS端口:提供公网访问应用的SLB端口,可设置范围为1~65535。
 - SSL证书: SSL协议证书, 在下拉列表中选择已上传的SSL证书。
 - 容器端口: 进程监听端口。由程序定义, 例如: Web服务默认使用8080端口。

d. 单击确定。

4. 结果验证。

复制配置的SLB的IP地址及端口,例如 192.168.0.184:80 ,在浏览器中输入地址并回车,即可分别进入各自的应用首页。

如果访问地址区域未出现IP地址和端口信息,则表示绑定SLB失败,请查看变更记录并修复失败问题。更 多信息,请参见<mark>查看变更记录</mark>。

场景二: 绑定新建SLB

如果您需要SAE为您全新代购SLB并将其绑定,具体操作,请参见<mark>绑定已有SLB</mark>,在步骤3选择SLB时,在**请选** 择SLB下拉列表中选择**新建**。

选择**新建**之后,SAE自动进行SLB配额检查和账户余额检查,检查通过后为应用自动购买全新的SLB实例,并 在下方显示具体SLB信息。

添加公网SLB访问						\times
 1. 设置公网负载均衡,能通过公网访问您 2. 选择新创建后,系统会为您的应用目录 	的应用。您可以选择已有的SLB服务,也可以创建新的SLB。 加购买一个公网SLB服务,按使用量计费,详见 <mark>SLB计费详情。</mark> 购买的	ISLB信息可以在负载均衡	时控制台查看。			
请选择SLB: 新建 🗸 SLB复用说明?						
检查项目	状态	说明				
SLB配额检查	成功●					
账户余额检查	成功 ❷					
产品类别	产品配置	数量	付费方式	购买周期	资费	
负载均衡SLB - 公网	地域: 华南1 公网带宽:按使用流量计费	1	按量付费	无	查看价格	
TCP协议 HTTP协议 HTTPS协议						
网络协议	SLB溯口	容器端口			操作	
ТСР	80	8080			Ī	
+添加下一条监听						
					确定	取消

相关操作

修改SLB访问设置

1. 在应用基本信息页面默认显示的基本信息页签,找到应用访问设置区域,并根据网络需求单击编辑私 网SLB访问或编辑公网SLB访问。

>> 应用访问设置			
●使用SAE S16功能之前,请先查看S16配置的果 典型S16晒量最佳实践 自2020年4月1日,凡是在SAE产品中创建,编唱S18后,再去S18产品中修改追听演口、协议、SS1E书、S18 tag、或拟服务器组名等S18配置的東中探制操作的资源,修改后的配置都会被SAE保存的配置信息强制重要,请一定注意通免。			
私网访问地址: 80 TCP	公网访问地址: 80 TCP		
编辑私网SLB访问 删除私网SLB访问	编辑公网SLB访问 删除公网SLB访问		

2. 在弹出的编辑私网SLB访问或编辑公网SLB访问对话框,修改所需信息并单击确定。

删除SLB访问设置

1. 在应用基本信息页面默认显示的基本信息页签,找到应用访问设置区域,单击删除私网SLB访问或删 除公网SLB访问。

∨ 应用访问设置							
● 使用Sat SLB功能之前,请先查看SLB配置的東 典型SLB后展最佳支援 自2020年4月1日,凡是在Sat产局中创造/编辑SLB后,再去SLB产局中修改监听编口、协议、SSL证书、SLB tag、虚拟服务器组名等SLB配置约束中探制操作的资源,修改后的配置都会被Sati保存的配置信息强制要量,请一定注意追急。							
私网は河岸地址: 80 TCP (瞬時45(网SLE)2)河 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	公网访问地址: 80 TCP 编辑公网SLE访问 删除公网SLE访问						

2. 在弹出的删除私网SLB访问或者删除公网SLB访问对话框,单击确定。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



2.2. 配置网关路由

2.2.1. 使用约束

SAE利用七层负载均衡能力,将请求流量按照转发规则分发到指定应用实例上。在使用SAE网关路由功能前,您需要了解负载均衡SLB相关约束。

约束

SAE的负载均衡实质是7层的负载均衡服务,通过为您配置SLB实例的监听来实现。

从SLB产品侧的角度来讲,您在SLB配置行为和SAE侧的SLB配置行为不作区分的,但是SLB侧对SAE侧的SLB配置没有保护,允许在SLB侧篡改或破坏SAE所维护的SLB配置,致使业务不符预期。为了确保您在SAE侧的业务不受影响,请仔细阅读以下约束信息:

● 在SAE侧配置的一条网关路由,在SLB侧将会映射为一条监听配置,SAE的网关路由功能不独占您的SLB实例。

如果您在SAE上创建了SLB监听,那么在该SLB上进行监听新建、配置、维护等操作属于合法行为。

- SAE网关路由功能还支持为您创建需要关联的虚拟服务器组,您对该虚拟服务器有完全的操作权限。
- 对于SAE创建的监听或虚拟服务器组,不建议您进行任何的修改或引用,否则对您的业务有一定的影响。 如果您需要进行修改,请检查操作的合法性,合法性如下表所示。

类型	修改项	合法性
	实例名称	Υ
实例	实例标签	禁止以下操作: • 更改SAE创建的标签。 • 删除SAE创建的标签。
	EIP绑定	Ν
	升降配	Υ
	带宽	Υ
	监听名称	Ν
	带宽	Υ
	调度算法	Υ
	虚拟服务器组ID	Ν
	访问控制	Υ
11 6 n.c.	添加请求x-forward-for字段	Υ
监听	gzip压缩	Υ
	会话保持	Υ
	健康检查	Υ

类型	修改项	合法性
	超时时间	Υ
	证书配置	Ν
	名称	Ν
虚拟服务器组	后端服务器(权重、实例、端口 等)	Ν
路由规则	所有配置修改	Ν

2.2.2. 为应用配置网关路由

应用托管到SAE后,如果您的业务请求需要分发给其他服务或者应用,那么您可以为应用配置网关路由,实现请求路由分发。本文介绍如何为您的应用配置网关路由。

前提条件

配置网关路由前,您需要完成以下操作:

- 应用托管至SAE。
- 创建SLB实例。
- 了解使用约束。
- 确保待配置网关路由的应用和后端接收请求的应用在同一命名空间内。

背景信息

网关路由功能适用场景如下:

- 单应用或多应用使用相同域名但存在不同路径流量转发。
- 单应用或多应用有不同域名流量转发,即不同域名解析的访问IP地址是同一个。

网关路由配置完成后,您可以通过域名:访问端口/Path访问相应的后端服务或者应用。

创建网关路由

- 1. 登录SAE控制台。
- 2. 进入创建网关路由页面。
 - 从应用基本信息页面进入。
 - a. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 单击具体应用名称。
 - b. 在基本信息页面的网关路由设置区域,单击添加转发策略,在网关路由页面,单击新建。
 - 从命名空间详情页面进入。
 - a. 在左侧导航栏单击命名空间,在命名空间页面上方选择地域,单击具体命名空间名称。
 - b. 在命名空间详情页面的左侧导航栏,单击网关路由,在网关路由页面,单击新建。
- 3. 在新建路由规则面板,设置网关入口的基本信息,单击下一步。

新建路由规则	IJ		×
	配置网关入口	配置转发策略	
* 名称:	test		
* 网关类型:	公网 私网 通过公网网关转发的流量按照实际转发流量计费,通过私网网关转发的流量不 转发;	际计费,但只能在当前VPC内部	
* SLB:	auto_named_slb()		
* 协议类型:	http https		
* 访问端口:	8080		
下一步	取消		

参数说明如下:

参数	说明
名称	设置路由规则的名称。
网关类型	选择待转发请求的网络类型。 • 公网 :通过公网网关转发的请求SAE按照实际转发 流量计费。 • 私网 :通过私网网关转发的请求不计费,仅在当前 VPC内部转发。

参数	说明
	选择您创建的SLB实例。
SLB	⑦ 说明 SLB实例数量上限与SLB控制台上保 持一致。关于SLB实例的使用约束,请参见SLB使 用说明。
协议类型	选择请求转发协议,支持HTTPS和HTTP两种协议。
SSL证书	如果您使用HTTPS协议,需要选择其相应的SSL证书。 如果您账户下没有SSL证书,请登录 <mark>SLB控制台</mark> 创建。
	设置接收请求,并向后端服务或者应用进行请求转发 的监听端口,端口范围为1~65535,例如80。
访问端口	 注意 • 在同一个负载均衡实例内,监听端口不可 重复。 • 多条网关路由规则禁止使用同一个SLB的 同一个端口。

4. 在新建路由规则面板,配置转发策略,单击确定。

新建路由规则				×
配置网关入 ③ 域名规范:由英文字母、数字、	□ 中划线(-)和点\:组	1成,城名不区分英文大小写,不支持	配置转发策略 泛域名定义。 标准域名如:	
foo.example.com。 path规范: 长度限制为1-80个字 成。	符。path以'/'开乡	头,后面由字母,数字,'-'、'/'、 ''、	"%'、'?'、'#'、'&'、这些字符	拍90个或多个组
自定义转发策略				
域名 端口	Path 💡	后端应用	容器端	口 操 作
例如: www. : 8080	/ + 添加	清选择	▶ 清補	
+ 添加域名				
默认转发策略(推荐设置) 🥐				
SLB	SLB端口	后端应用		容器端口
auto_named_slb(118.31	8080	请选择	\checkmark	请输入】
上一步 确定 取消				

参数说明如下:

参数	说明				
	■ 域名: 输入要转发的请求域名。				
	 ⑦ 说明 由英文字母、数字、短划线(-)和半角句号(.)组成,域名不区分英文大小写,仅支持*.example.com类型的泛域名定义。标准域名示例:foo.example.com。 Path与域名不能为空。 				
自定义转发策略	 Path: 输入请求转发路径。例如 foo.example.co m/product/sae ,请求转发路径 为/product/sae。 				
	 ⑦ 说明 Path规范:长度限制为1~80个字符。 Path以正斜线(/)开头,后面由字母、数字和 -/.%? #& 等字符组成。 Path与域名不能为空。 				
	 后端应用:接收转发请求的应用,该应用须与转发请求的应用在同一个命名空间内。 容器端口:设置访问您后端应用的容器端口,例如8080。 				
	 后端应用:选择需要转发的后端应用。 容器端口:设置访问您后端应用的容器端口,例如8080。 				
默认转发策略	◆ 注意 没有匹配自定义策略转发的请 求, SAE会指定默认的转发策略。如果您不设置默 认转发策略, 会导致访问报错。				

5. 验证结果。

• 在应用事件页面查看是否配置成功。具体操作,请参见查看应用事件。

く基本信息		应用事 命名空间:	件 cn-shenzhen:	qualitation	来源类型:	网关路由(Ingre	ss) 🗸		来源名:	8-wr807%bendaninge	34a-80-10di
变更记录	3	事件原因:	请输入事件原因		事件等级:	全部	~		搜索	重置	
网关路由		1总 🕕	十12 项记录,当前()	2保留7天内的应用事件。							
事件信息			事件等级	事件原因		来源类型	来源名			首次发生时间	最近发生时间
		+	Normal	ConfigForwardRulesSucceed		Ingress	k-acit/hiteplaney	- 2014 82 1848-4		2020-01-15 11:39:57	2020-01-15 11:39:59
		+	Normal	ConfigBackendServerGroupSucceed		Ingress	k-sch?hbsplanrg	2114-02-0284		2020-01-15 11:39:56	2020-01-15 11:39:59
		+	Normal	ConfigBackendServerGroupSucceed		Ingress	k-scit/7sbsplanrg	p.2014-80-1048-4		2020-01-15 11:39:55	2020-01-15 11:39:59
		+	Normal	ConfigLoadBalancerListenerSucceed		Ingress	k-scit/7kbsplanrp	p 25ra 80 15464	·	2020-01-15 11:39:55	2020-01-15 11:39:59
		+	Normal	ConfigDefaultBackendServerGroupSuc	cceed	Ingress	b-sc077cbeplanrq	p.25ra-80-35a64		2020-01-15 11:39:53	2020-01-15 11:39:58
		-	Normal	ConfigIngres		Ingress	k-scit/Teleplanes	- 70% RT 1048-4		2020-01-15 11:39:58	2020-01-15 11:39:58
		C	onfig ingress sl-	-cn-shenzhen-	/lb-wz	9t77e3wq4aovngv	2bta-80-bbdb48 succe	eded			
		+	Normal	CleanBackendServerGroupsSucceed		Ingress	Ib-w2977k3wplacurg	p 25ra 80 10404		2020-01-15 11:39:58	2020-01-15 11:39:58
		+	Normal	ConfigForwardRulesSucceed		Ingress	Ib-wc977e3wq4aovrq	p 251a-80-3548-4		2020-01-15 11:39:57	2020-01-15 11:39:57
		+	Normal	ConfigBackendServerGroupSucceed		Ingress	Ib-ec977e3eplacerg	p-251a-80-1048-4		2020-01-15 11:39:56	2020-01-15 11:39:56
		+	Normal	ConfigBackendServerGroupSucceed		Ingress	h-scit7hbsplannp	p 244a - 80 - 58 - 68		2020-01-15 11:39:55	2020-01-15 11:39:55

○ 登录SLB控制台,查看监听是否已经配置。

负载均衡 SLB	负素	战均衡 SLB / 实例管理 / Fine ingress-interne	at/47.113	转发策略				C)添加域名和路行	至转发 🗙
概览	~	- Loga-ingress-inter	net/	长度限制为2-80′ * 域名与URL请至	个字符,只能使用字母、数 少填写一项。	字、"="、"/"、"."、"%"	,"?"、"#"、"&"这些字符;	URL不能只为"/",但必须	i以"/"开头。	
实例管理		实例详情 监听 虚拟服务器组	默认服	添加转发策略						
回收站		添加监听		域名	URL		虚拟服务器组	备注		操作
证书管理		名称	前端协	请输入域名	1		请选择 🗸 🗸	请输入备注		删除
访问控制				添加域名	添加規則					
日志管理		ingress/80/lb-		添加转发策略						
操作日志		wz9t77e3wq4aovngv2bta- 8/b067721d-3748-11ea-92f4-	HTTP:	转发策略列表						
访问日志				城名	URL	虚拟服务器组	备注	1	果作	
健康检查日志					1	ingress/80/b067	ingress/eda	is.si	◎辑 删除 ❷	API
SLB 实验室 ^				edas.site	/consumer/	ingress/80/b067	ingress/eda	as.si	扁辑 制除	
闲置实例										
配额管理				确定 取消						

○ 通过域名:访问端口/Path访问相应的后端服务或者应用,查看是否生效。

网关路由管理

创建路由规则后,您可以在网关路由页面进行转发策略查看、网关路由的编辑、删除以及相关事件查看。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



3.应用设置

3.1. 变更实例规格

应用部署到SAE后,您可以按需修改应用运行所需的实例规格,例如CPU和内存。

前提条件

为了保证您的业务连续性,在变更实例规格前,请先配置最小存活实例数,配置步骤请参见回退历史版本。

背景信息

应用运行一定时间后,如果应用实例运行所需的规格过小导致应用无法正常工作,例如内存过小致使内存溢出(OOM),那么需要在应用配置不变的情况下,修改应用实例规格,确保应用正常运行。

↓ 注意 变更实例规格后, SAE自动重新部署该应用, 其中单实例应用在短暂几秒内不可用, 请在业务量较小时段进行操作, 控制业务风险; 多实例应用无影响。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在基本信息 > 基本信息页签, 在应用信息区域的实例规格所在单元格中, 单击变更规格。

基本信息 实例部署信息								
◇ 计量数据 🔗 了解计量方式								
本月CPU资源使用量 🛛	本月Memory资源使用量 💿	实时CPU资源使用量 🔮	实时Memory资源使用量 📀					
18443 Coremin	36886 Gi8-min	1 Coremin	2 GiB-min					
✓ 应用信息								
应用名: spring-boot-consumer	应用名: spring-boot-consumer							
应用id:		VPC: doc-test-	VPC: doc-test-					
VSwitch doc-test 参VSwitch都書 运行实例数 当前1个实例/目标1个实例								
实例规想: 1Core, 2GB. 系统直接盘空间20GB 安更原格 程序地名: spring-boot-consumer-1.0-SNAPSHOT.jar.1.0								
应用标题中间: 2020年5月18日 125324 应用标签: 编辑标签								
应用描述:								

4. 在变更实例规格对话框,依据需求选择实例规格,并单击确定。

变更	巨实例规格		×
提示	: 变更规格时系统会自动执行一次应用部署	, 单实例的应用可能会短暂几秒不可用, i	青自行控制业务风险。多实例的应用不会有任何影响。
	规格族	CPU	Memory
С	〕 通用型8	0.5 Core	1 GiB
С	〕 通用型9	0.5 Core	2 GiB
С	〕 通用型2	1 Core	2 GiB
۲	通用型10	1 Core	4 GiB
С	〕 通用型4	2 Core	4 GiB
С	〕 通用型11	2 Core	8 GiB
С	〕 通用型6	4 Core	8 GiB
С	〕 通用型12	4 Core	16 GiB
			确定

- 5. 结果验证。
 - 方法一:在应用详情页面的基本信息页签查看,如果实例规格已经变更为所更改内容,则表示规格 变更成功。
 - 方法二: 在应用的**变更记录**页面查看, 如果变更记录显示成功, 则表示规格变更成功。

变更记录									
所有变更类型	~	模糊搜索	✓ 根据描述信息进行模糊搜索	Q					
创建时间		结束时间	变更类型		描述信息	变更状态	交更人	来源	摄作
2020-07-13 19:27:42		2020-07-13 19:27:43	调整实例规格		目标应用配置: CPU 1.0核 内存 2048MB	● 执行成功		console	立石
2020-05-18 10:58:41		2020-05-18 10:59:32	绑定SLB		N/A	● 执行成功	主账号	console	宣春
2020-05-18 10:58:24		2020-05-18 10:58:24	解绑SLB		N/A	● 执行成功	主账号	console	查看

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



3.2. 切换安全组和vSwitch

应用部署到SAE后,您可以按需修改应用所在的安全组和vSwitch。

背景信息

安全组是应用实例的虚拟防火墙,用于设置单个或多个实例的网络访问控制,每个应用实例必须属于一个安 全组。您可以通过切换应用的安全组修改应用间的网络访问控制。

交换机(vSwitch)是组成专有网络的基础网络模块,用来连接不同的云资源。成功创建交换机后,您可以 在交换机中创建云资源、绑定路由表或者绑定网络ACL。更多信息,请参见使用交换机。

切换安全组

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在基本信息页面的基本信息页签, 在应用信息区域的安全组所在单元格中, 单击切换安全组。

基本信息 实例部署信息				
∨ 计量数据 🔗 了解计量方式				
本月CPU资源使用量 👔	本月Memory资源使用量	实时CPU资源使用量 🕜	实时Memory资源使用量 📀	
65,230 _{kemin}	130,460 GiBmin	2 _{棱min}	4 GiB-min	
✓ 应用信息				
应用名: test		命名空间: cn-hangzhou		
应用id: aff6d4f0-2f20-4fff-		VPC: doc-test		
vSwitch: doc-test (可用ip数为222) 多vSwitch部	8	安全组: edas.serverless.donot.delete.vpc. 切脑安全组		
实例规格: 1Core, 2GiB, 系统盘磁盘空间20GiB 变更规格		运行实得数 当前2个实例/目标2个实例		
应用创建时间: 2021年6月18日 14:18:45		揭像地址: registry-vp.c.n-hangzhou.aliyuncs.com/		
应用标签: 编辑标签		应用描述∠ 编辑		

4. 在弹出的切换安全组对话框中选择需要切换的安全组,单击确定。

切换安全组	×
1 切除安全组。需要用户手动执行启动、重命和邮客应用时新实例生效	
安全钼: edss.serverless.donot.delete.vpc	
	論定 取消
 注意 切换安全组后,您需要手动重启或者部署实例才可以生效,详细操作请 命周期。 	参见管理应用生

5. 结果验证。

在应用详情页面的基本信息页签查看,如果安全组已经变更为所更改内容,则表示安全组切换成功。

切换vSwitch

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在基本信息页面的基本信息页签, 在应用信息区域的vSwitch所在单元格中, 单击多vSwitch部署。
- 4. 在弹出的多vSwitch部署对话框中选择需要的vSwitch,单击确定。

? 说明 vSwitch支持多选。

5. 结果验证。

在**应用详情**页面的基本信息页签查看,如果vSwitch已经变更为所更改内容,则表示vSwitch部署成功。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



3.3. 复制应用

如果您希望将当前应用的配置完全复制到其他命名空间,您可以使用SAE的应用复制功能。

操作步骤

< ↓ 注意

- SAE不支持跨地域的应用复制。
- 一次只可以复制一个应用。
- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域。
- 3. 在应用列表页面,单击目标应用操作列的复制。

全部	默认(cn-shenzhen)	test				
elite di	四月 应用名称	✓ 満輸入应用名称	Q. 标签			\$ C
	应用名称	标签	命名空间	弹性策略启用状态	当前实例数/目标实例数	操作
	mysae	ø	cn-shenzhen:test	未设置	1/2	管理 复制
	test	ø	cn-shenzhen	已启用 (监控指标单性: 实例数2)		管理 复制
	test1	ø	cn-shenzhen:test	未设置	0/2	管理 复制
		批量启动应用 批量设置标签 (已选中	2个应用) 🕜		每页显示: 10	▶ 共3条 〈 上一页 】 下一页 〉

单击**复制**后,页面会自动跳转至**创建应用**界面,您可以按需修改配置信息。如果您不需要修改配置信息,您可以按照界面提示完成当前应用的创建。

1 一应用基本信息	1			esses and
您当前账户了	下cn-shenzhen的应用总实例数限制为300个,目前您已使用0个	,如需增加额度请提工单申请。		
* 应用名称:	mysae-copy			
* 专有网络配置:	● 自定义配置 ○ 自动配置			
* 命名空间 🕐:	test	~		
	命名空间和VPC是——映射关系,如果需要修改VPC请前往命	ia空间详情页。	士 特名注\	
• vswitch:	可用区	• G (ST425464)	
* 安全组:	10.00	$^{\sim}$ G		
应用实例数 ?:	2 个 (最多创建50个)			
* VCPU:	0.5Core 1Core 2Core 4Core 8Core 16Core	32Core		
* 内存:	1GiB 2GiB			
	建议0.5core 1GiB的实例规格仅在开发测试环境使用			
应用描述:	介绍应用的基本情况			
配置费用: ¥ 0.00 参考价格, 具体扣费请	040509/min 省:¥0.0005787/min ①优惠 拟账单为准。了解计费详情	洋情		下一步:应用部署配置

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



3.4. 管理标签

3.4.1. 为应用添加标签

为了方便您管理托管在SAE上的应用,SAE提供了标签功能。本文介绍如何为应用添加标签。

前提条件

在为应用添加标签前,确保您已完成以下工作:

- 开通SAE服务
- 创建命名空间
- 创建VPC
- 部署应用到SAE

背景信息

随着您所创建的应用逐渐增多,利用标签将资源进行分组管理和归类便于搜索、聚合资源。

- 为不同的环境(如生产环境和测试环境)、操作系统(如Windows Server和Linux)或者客户端平台(如 iOS和Android)绑定不同的标签。
- 在团队或者项目管理中,您可以添加以群组、项目或部门为维度的标签(如Cost Center: aliyun),便于项

目管理。

使用说明

- 标签都由一对键值对(Key-Value)组成。
- 资源的任一标签的标签键(Key)必须唯一。
- 不同地域中的标签信息不互通。
- 解绑标签时,如果解绑之后该标签不再绑定任何资源,则该标签会自动被删除。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 选择您需要添加标签的具体应用。

⑦ 说明 当应用较多时,您可以在应用列表页面的搜索栏选择应用名称、应用ID、SLB ID或实例IP,并输入相应的参数,单击 图标,搜索目标应用。

- 3. 在具体应用的标签列,将鼠标悬停至 🖏 图标,然后单击标签气泡内的编辑标签。
- 4. 在编辑标签面板,单击新增,并设置标签的标签键和标签值,然后单击确定。

← 编辑标签			×
 推荐使用标签编辑功能,您可以给相同业务属性的 	了解更多		
您正在为应用批量设置标签			
* 标签键		标签值	
key	:	value	Ē
新增			
注:每个应用最多可绑定20个标签。单次操作绑定/解绑	示签	的数量分别不能超过 20个。	
当给多个应用编辑标签时,编辑框内仅显示新增加的标签	,不	显示历史已有的标签。	
确定取消			

5. (可选)单击新增,可为应用添加多个标签。

? 说明

- 。 每个应用最多可绑定20个标签。
- 。 单次绑定或者解绑标签的数量均不超过20个。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



3.5. 一键启停应用

本文介绍如何将处于同一命名空间的应用一键批量启停,提高您的工作效率,释放闲置资源,节约成本。

前提条件

待批量启停的应用须处于同一命名空间内。

背景信息

随着您在SAE控制台上托管的应用逐渐增多,部分应用占用了资源却处于闲置状态,造成了资源的浪费。例 如应用开发完成后对其进行测试联调,完成测试联调后未将应用实例释放。您可以使用一键批量启停功能将 闲置应用停止以释放资源。当需要再次使用时可以一键批量启动应用,继续执行相关业务。

批量停止应用

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域。
- 3. 在应用列表页面上方选择需要停止的应用所在的命名空间。

Serverless应用引擎SAE / 应用列表							
应用列表							
全部 默认(cn-hangzhou)	namespace:						
创建应用 应用名称	✓ 満輸入应用名称	Q 标签				© ©	
应用名称	标签	命名空间	弹性策略启用状态	当前实例数/目标实例数	操作		
wayne-consumer	ø	cn-hangzhou	未设置	3/3	管理		
wayne-testapp	ø	cn-hangzhou:	未启用	3/3	管理		

- 4. 在应用列表中选中要批量停止的应用。
- 5. 单击批量停止应用,在弹出的批量停止应用面板,选择停止策略,单击确定。
 - **立即停止**:选择**立即停止**后, SAE将立即停止所选的所有应用。适用于临时性的单次运维场景。

← 批量停止应用	×
已选中的应用: 共2个	
wayne-consumer wayne-testapp	
停止策略	
● 立即停止 ? ○ 定时停止 ?	
1 批量停止应用后,SAE将物理删除这些应用下的所有实例,并停用所有已启用的弹性策略。业务将会中断,资源计费也会停止,但SAE会保存应用的基本配置信息和负载均衡设备信息,以便下次启动应用时可秒级完成。请注意,停止后再启动应用时,实例IP会变。确定批量停止应用?	
确定取消	

定时停止:选择定时停止后,将跳转到定时启停规则页面。定时启停适用于计划性单次或重复运维的场景,例如开发测试环境在固定时段的定时启停。更多信息,请参见管理定时启停规则。

<□ 注意

- 停止应用后,SAE将删除选中应用下所有实例、中断业务和停止计费,仅保留应用的原基本配置信息和负载均衡设备信息,以便再次启动应用。
- 再次启动应用后,应用的实例IP将会发生改变。
- 如果应用启用前配置了弹性伸缩,停止应用后该应用上所设的弹性伸缩规则不可用。再次 启动后,该规则不会被自动启用,需要手动启用。
- 6. 您可以参考以下方法验证应用是否批量停止成功。
 - 方法一:在应用列表页面查看应用的当前实例数/目标实例数,如果应用的当前实例数为0,表示
 应用批量停止成功。
 - 方法二:在命名空间的变更记录中查看变更详情,如果变更详情显示执行成功,表示应用批量停止成功。

変更记录 / 変更详情				终止变更
变更流程ID: 51933301-		发布分批数: 1	分批间处理方式: N/A	
执行状态: 💿 执行成功		变更类型: 批量停止应用	发布时间: 2020-11-19 20:36:46	
描述信息:		变更对象: ···test.spring-boot-pro		
	应用执行状态 全部 V			
✔ 批量傳止应用	应用名称	我们行	大态	操作
	lubu-test	• 执	将无成功	查看详情
	spring-cloud-netflix-eureka-provider	• 扨	行成功	查看详情
	spring-boot-provider	● 失	数 (等待重试)	报错信息 查看发布单 重试
	spring-cloud-netflix-eureka-consumer	● 失	敗 (等待重试)	报错信息 查看发布单 重试
	my-sae	• 执	行成功	查看详情
	spring-boot-consumer	• 执	行成功	查看详情 📖
	く 1 下一页 >			Bô

- ⑦ 说明 如果您批量操作存在失败记录,您可以执行以下操作:
 - 单击**报错信息**查看详细的错误日志。
 - 单击查看发布单查看应用部署的过程信息。
 - 单击**重试**重新停止应用,但重试次数不超3次。

批量启动应用

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域。
- 3. 在应用列表页面上方选择需要停止的应用所在的命名空间。

Serverless应用引擎SAE / 应用列表						
应用列表						
全部 默认(cn-hangzhou)	namespace					
创建应用 应用名称	➤ 请输入应用名称	Q. 标签				\$ C
应用名称	标签	命名空间	弹性策略启用状态	当前实例数/目标实例数	操作	
wayne-consumer	ø	cn-hangzhou	未设置	3/3	管理	
wayne-testapp	ø	cn-hangzhou:	未启用	3/3	管理	

- 4. 在应用列表页面选中需要批量启动的应用。
- 5. 单击批量启动应用,在弹出的批量启动应用面板,单击确定。
 - **立即启动**:选择**立即启动**后, SAE将立即启动所选的所有应用。适用于临时性的单次运维场景。
 - 定时启动:选择定时启动后,将跳转到定时启停规则页面。定时启停适用于计划性单次或重复运维的场景,例如开发测试环境在固定时段的定时启停。更多信息,请参见管理定时启停规则。

↓ 注意 启动应用后,应用的实例ⅠP地址将会发生改变。

- 6. 您可以参考以下方法验证应用是否批量启动成功。
 - 方法一:在应用列表页面查看应用的当前实例数/目标实例数,如果应用的当前实例数与目标实例数相同,表示应用批量启动成功。
 - 方法二:在命名空间的变更记录中查看变更详情,如果变更详情显示执行成功,表示应用批量启动成功。

変更记录 / 変更详情				终止	交更
変更流程ID: 51933301-		发布分批数: 1	分批间处理方式: N/A		
执行状态: 💿 执行成功		变更关型: 批量停止应用	发布时间: 2020-11-19 20:36:46		
描述信息:		变更对象:			
	应用执行状态				
▶ 批畫停止应用	应用名称	执行状态	5	操作	
	lubu-test	• 执行	成功	查看详情	
	spring-cloud-netflix-eureka-provider	• 执行	成功	查看详情	
	spring-boot-provider	• 失敗	(等待重试)	报错信息 查看发布单 重试	
	spring-cloud-netflix-eureka-consumer	• 失敗	(等待重试)	报错信息 查看发布单 重试	
	my-sae	• 执行	成功	查看详情	
	spring-boot-consumer	• 执行	成功	查看详情	
	く 1 下一页 >				BX

⑦ 说明 如果您批量操作存在失败记录,您可以执行以下操作:

- 单击**报错信息**查看详细的错误日志。
- 单击查看发布单查看应用部署的过程信息。
- 单击重试重新停止应用,但重试次数不超3次。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



3.6. 管理定时启停规则

定时启停功能可实现按命名空间定时批量启停应用,例如定时启停开发环境或测试环境的全部应用,帮助您 提高硬件资产利用率、降低资源闲置成本。

准备工作

命名空间应用定时启停功能依赖运维编排服务OOS,请先创建OOSServiceRole RAM角色,并为该角色赋予管理SAE的完整权限AliyunSAEFullAccess。

⑦ 说明 如果您已经创建OOSServiceRole角色,则只需为其添加权限AliyunSAEFullAccess。

- 1. 登录访问控制RAM控制台。
- 2. 在左侧导航栏单击RAM角色管理,并在RAM角色管理页面单击创建RAM角色。
- 3. 在创建RAM角色面板的选择类型页面上选择阿里云服务,并单击下一步。

创建 RAM 角色	×
1 选择类型 2 配置角色 3 创建完成	
当前可信实体类型	
阿里云账号 受信云账号下的子用户可以通过扮演该RAM角色来访问您的云资源,受信云账号可以是当前云账号,也可以是其他云账号	3
● 阿里云服务 受信云服务可以通过扮演RAM角色来访问您的云资源	
 ○ 身份提供商 身份提供商功能,通过设置SSO可以实现从企业本地账号系统登录阿里云控制台,帮您解决企业的统一 用户登录认证要求 	
	22
下一步 关闭	

4. 在**配置角色**页面的角色类型区域选择普通服务角色,输入角色名称*OOSServiceRole*,在选择受信服务区域选择运维编排服务,并单击完成。

创建 RAM 角色	×
✓ 选择类型 2 配置角色 3 创建完成	
选择可信实体类型 阿里云服务	
角色类型	
● 普通服务角色 ○ 服务关联角色	
* 角色名称	
OOSServiceRole	
备注	
* 选择受信服务	
运维编排服务	E?
上一步 完成 关闭	

- 5. 在创建完成页面上, 单击为角色授权。
- 6. 在**添加权限**面板的**选择权限**区域,选择**系统策略**,在输入框中输入关键字*sae*,单击搜索结果中的AliyunSAEFullAccess,然后单击确定。

添加权限				×
* 被授权主体				
OOSServiceRole@role.	.onaliyunservice.com X			
* 选择权限				
 系统策略 自定义策略 	新建权限策略		已选择 (1)	清除
sae		G	AliyunSAEFullAccess X	
权限策略名称	备注			
AliyunSAEReadOnlyAccess	只读访问Serverless应用引擎 (SAE) 的权限			
AliyunSAEFullAccess	管理Serverless应用引擎 (SAE) 的权限			
				E?
				ÖŎ
确定取消				

创建定时启停规则

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击命名空间,在顶部菜单栏选择目标地域。
- 3. 在命名空间页面上,执行以下操作之一以进入命名空间详情页面。
 - 单击目标命名空间的名称。
 - 在目标命名空间的操作列中单击详情。
- 4. 在左侧导航栏单击定时启停规则, 在定时启停规则页面单击创建定时启停规则。
- 5. 在创建定时启停规则面板上,配置规则信息,并单击确定。

← 创建定时启停规则				×
* 规则名称	请输入规则名称 必须以小写字母开头,以数字或小写字母 过63个字符。	结尾,仅可包含小写字母、	中文字符、数字及短划线(-)。不超	
* 触发动作	批量启动应用	\sim		
* 执行频率	◉ 仅执行一次 ○ 定期执行			
	Jun 8, 2020 16:47:30	Ē		
执行前通知配置	请输入钉钉机器人Webhook地址	如何获	取钉钉机器人Webhook地址	
	SAE將在每次开始执行定时任务前向此处打 信息。	指定的钉钉机器人发送通知	1, 方便您第一时间获取系统运维动作	
* 触发的应用列表				
批量停止应用后, SAE将物 均衡设备信息, 以便下次启	理删除这些应用下的所有实例,并停用所有已 动应用时可秒级完成。请注意,停止后再启;	3启用的弹性策略。业务将 动应用时,实例IP会变。	会中断,资源计费也会停止,但SAE会保	存应用的基本配置信息和负载
应用名称	标签	命名空间	弹性策略启用状态	当前实例数/目标实例数
没有数据				
			每页显示: 10 2 共0条	〈 上─页 】 下─页 〉
确定 取消				

参数说明如下。

参数	说明
规则名称	定时启停规则的名称。必须以小写字母开头,以数字 或小写字母结尾,仅可包含小写字母、中文字符、数 字及短划线(-),不超过63个字符。
触发动作	 批量启动应用:批量启动本命名空间下选中的应用。 批量停止应用:批量停止本命名空间下选中的应用。

参数	说明
执行频率	 仅执行一次:在指定日期和时间执行一次后不再执行。 定期执行:按照指定的日期和时间循环执行。 日粒度:指定每个月需要执行的具体日期和时间,每月循环。 周粒度:指定每周需要执行的具体日期和时间,每周循环。 周粒度:指定每周需要执行的具体日期和时间,每周循环。 高级:使用Crontab表达式定义执行时间,各字段的取值范围如下所示: 分(分钟):0分~59分。 时(小时):0时~23时,0表示午夜0时。 日(每个月的具体日期):1日~31日。 月:不可修改,1月~12月。 周(星期):1~7,1表示星期日,2表示星期一,依次类推。
执行前通知配置	配置一个钉钉机器人Webhook地址,SAE将在每次开 始执行定时任务前向此处指定的钉钉机器人发送通 知,方便您第一时间获取系统运维动作信息。
触发的应用列表	选中需要执行启动或停止操作的应用。

创建好的规则将显示在**定时启停规则**页面上,且默认为启用状态。

使用Crontab表达式

借助Crontab表达式,可以定义更灵活的定时启停规则触发时间和循环规律。

使用Crontab表达式时需要注意以下事项:

- 只能按日或周指定启停触发时间,如果指定了其中一个,另一个则必须填写通配符?,表示不加限制。
 例如,如果指定每个月第1日、第2日,则必须在周输入框中填写?。
- 在Cront ab表达式内必须使用英文半角字符,例如英文通配符星号(*)、问号(?)。
- 枚举值之间用英文逗号(,)分隔。
- 值的区间用短划线(-)表示,例如:日输入框中填写 1-5,表示第1日~第5日。
- 填写周时, 1表示周日, 7表示周六, 即每周从周日到周六的数字为1~7。

含义	填写方法	Crontab表达式
每周一至周五22:30	30 分 22 时 ? 日 * 月 1-5 周	30 22 * 1-5
每个月第5日、第15日和第25日的0 时	0分0时5,15,25日*月? 周	0 0 5,15,25 * ?

管理定时启停规则

如需管理定时启停规则,例如查看规则详情、编辑规则等,请在定时启停规则页面上按照以下步骤执行:

- 如需查看规则详情,单击规则名称。
- 如需查看最近一次定时规则触发后的执行过程信息和执行结果,单击最近一次触发时间列中的变更单详 情,并在变更详情页面查看详细信息。
- 如需启用或停用规则,单击操作列中的启用或停用。
- 如需编辑规则,单击操作列中的编辑,并按照创建定时启停规则的说明编辑规则内容。
- 如需删除规则, 单击操作列中的删除。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



相关文档

• 什么是运维编排服务
4.应用运维

4.1. 管理应用生命周期

在应用托管至SAE后,您可以对应用执行更新、扩缩容、启停、删除和监控启停等应用生命周期管理操作。

访问应用概览页

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击概览页, 在概览页页面查看全部地域下应用的各项指标。



- 3. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 单击具体应用名称。
- 4. 在基本信息页面,管理目标应用的生命周期。

基本信息		部署应用 回退历史版	本 手动扩缩 自动扩缩 停止应用 更多 >>				
 1.应用在VPC内默认是不能访问公网的,如果需要访问公网请按 2.如果你的应用需要访问RDS数据库,请参考:如何设置RDS白名 	以下提示操作: 应用如何访问公网,常见访问公网场 <mark>景最佳</mark> 示例 名单						
基本信息 实例部署信息							
∨ 计量数据 🔗 了解计量方式							
本月CPU资源使用量	本月Memory资源使用量()	实时CPU资源使用量 💡	实时Memory资源使用量 💡				
1,656 _{i表min}	3,312 GiBmin	2 _{核min}	4 GiB-min				
∨ 应用信息							
应用名: ••••		命名空间: cn-shanghai					
应用id: 1c26c0dc-055a-4627-8388		VPC: vpc-					
vSwitch: vswitch (可用ip数为248) 多vSwitch	印書	安全组 alikafka切换安全组					
实例规格: 1Core. 2GiB. 系统盘磁盘空间20GiB 变更规格		运行实例数:当前2个实例/目标2个实例					
应用创建时间: 2021年5月7日 13:23:53		辑像地址: registry-vpc.cn-shanghai.aliyuncs.com/					
应用标签: 编辑标签		应用描述∠ 编辑					

更新应用

在基本信息页面右上角,单击部署应用,并选择需要的发布方式。具体操作,请参见:

- 灰度发布应用
- 分批发布应用

回退历史版本

在基本信息页面右上角,单击回退历史版本,在回退历史版本页面选择版本,设置发布策略。具体操作, 请参见回退历史版本。

手动扩缩应用

在基本信息页面右上角,单击手动扩缩。具体操作,请参见手动扩缩。

? 说明

- 扩容时,如果扩容的实例数与原应用的实例数总和,大于应用的资源使用限制,则扩容失败。如果需要继续增加应用实例,请提交工单。
- 如果您的扩容缩为非紧急需求,建议使用弹性伸缩。

自动扩缩应用

在基本信息页面的实例部署信息页签,单击+添加弹性策略。具体操作,请参见配置弹性伸缩策略。

启停应用

单应用的启停步骤如下所示。关于多应用的启停步骤,请参见一键启停应用。

停止应用:在基本信息页面右上角单击停止应用,在弹出的停止应用对话框中单击确定,停止应用运行。

当应用处于运行状态时,可以根据业务需求随时停止应用。停止应用时,该应用的最小存活实例数设置为 0,实例数设置为0,其所使用的计算资源将停止计费或者计量。

↓ 注意 应用停止后,如果您的应用运行时还依赖其他产品或者服务,例如绑定SLB和使用VPC, 该部分产品或者服务仍处于计费状态。

 启动应用:在基本信息页面右上角单击启动应用,在弹出的启动应用对话框中单击确定,启动处于停止 状态的应用。

当应用处于停止状态时,您可以根据需求启动该应用。应用启动后,启动的实例数恢复为停止前所配置 值,同时重新开始计量或计费。

重启应用

对于正在运行的应用实例,在基本信息页面右上角选择更多 > 重启应用,在弹出的重启应用对话框中设置最小存活实例数,并单击确定。

⑦ 说明 在应用程序正常、业务异常运行时,请勿手动重启整个应用,您需要先使用健康检查功能对应用实例状态进行诊断。如果检查结果表明实例状态不健康会自动重启实例。具体操作,请参见设置健康检查。

删除应用

在基本信息页面右上角,选择更多 > 删除应用,在弹出的删除应用对话框中单击确定,删除已经部署的应用。

↓ 注意 如果您的应用运行时还依赖其他产品或者服务,例如绑定SLB和配置VPC,那么删除前请确认 所依赖的产品或者服务是否需要继续使用,若不需要则停止该服务,停止计费。

启停应用监控

在基本信息页面右上角,启停应用监控。具体操作,请参见启停应用监控。

- 开启应用监控:单击更多 > 开启应用监控,在弹出的开启应用监控对话框中单击确定,开启对应用的 监控。
- 停止应用监控:单击更多 > 停止应用监控,在弹出的停止应用监控对话框中单击确定,关闭已开启的应用监控。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



4.2. 重启应用实例

SAE支持应用实例的单个重启或者批量重启,您可以按需选择需要重启的实例。本文介绍在SAE控制台上重 启应用实例的操作步骤。

操作步骤

○ 注意

- 应用实例重启, 只重启实例中业务容器, 不会重建实例。
- 应用实例重启会造成业务中断,请在业务较小的时间段进行。
- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面,单击实例部署信息页签。
- 4. 在**实例部署信息**页签的实例列表中,勾选需要重启的实例,单击**批量重启实例**。

基本信息			部署应用	回退历史版本	手动扩缩 自动推	「缩 停止应用	更多 ~
 1. 应用 2. 如果 	在VPC内默认是不能访问公网的,如果需要访问公网请按以下提示操作:应用如何访问 你的应用需要访问RDs数据库,请参考:如何设 <mark>置</mark> RD <mark>s白名单</mark>	9公网,常见访问公网场景载佳示例					
基本信息	实例部署信息						
> 弹性伸缩	8						
∨ 默认分约	l 状态:运行2个实例						
显示实例II	e信息						
	实例名称	VSwitch	版本	运行状态	操作		
	7snqt	可用区B wuch-sh_switch	1606372602290	Running	实时日志 │ V	Vebshell 事件 :	
	mqlvx	可用区B wuch-sh_switch	1606372602290	Running	实时日志 V	Vebshell 事件 :	
批量重启实	9 当前包息计选中了2个实例					< 1	

单击**批量重启实例**后,页面会自动跳转至**变更详情**页面。如果**执行状态**显示**执行成功**,表示实例重启 成功。

交更记录 / 交更详情							
变更流程ID: 7e2bf978-		发布分批数: 1		分担	如处理方式: 自动		
执行状态: 执行成功	变更类型: 批量重启实例		发行	时间: 2020-12-01 14:37:38			
描述信息: N/A		变更对象: N/A					
第1批变更							
父 重启实例					批量	重试执行失败的实例	刷新
	实例ID		运行状态		配置规格		
	> mysae-	0.000.000.000.000	成功		CPU: 0.5 核 内存: 1024MB		
	> mysae-		成功		CPU: 0.5 核 内存: 1024MB		
					共有2条,每页显示:20条	1 > 前往	1 页

5. (可选)如果您选择的部分实例重启失败,您可以在**发布详情**页面单击**批量重试执行失败的实例**,重 新重启实例。

変更记录 / 変更详情							
变更流程ID: 7e2bf978-	10.00	发布分批数: 1		分批间处理方式: 自动			
执行状态: 🚺 执行中	变更类型: 批量重启实例		发布时间: 2020-12-01 14:37:38				
描述信息: N/A 变更对象: N/A							
第1批变更							
つ 重启实例	实例状态: 全部			能量電話执行失敗的实例	新		
	实例ID		运行状态	配置规格			
	> mysae-	100 000 Late 150	€ 执行中	CPU: 0.5 核 内存: 1024MB			
	> mysae-	100000000000000000000000000000000000000	€ 执行中	CPU: 0.5 核 内存: 1024MB			
				共有2条, 每页显示: 20条 〈 <mark>1</mark> 〉 前往 1	页		

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



4.3. 删除应用实例

当您的应用实例出现异常或无法处理业务时,您可以删除相关实例,系统会自动重建新的实例。本文介绍在 SAE控制台上删除应用实例的操作步骤。

操作步骤

○ 注意 应用实例被删除后, SAE会自动为您创建一个新的实例, 但实例IP会发生变化, 请谨慎操作。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面,单击实例部署信息页签。
- 4. 在实例部署信息页签的默认分组区域中,单击实例列表中目标实例操作列的 图标,在下拉框中单

击删除。

∨ 默认:	◇ 款\/分组 状态: 進行2个実例										
显示实例P值息											
	实例名称	VSwitch	镜像	运行状态	运行时间	操作					
	ks44s	可用区A sae-demo-switch	/sae-demo-image/provider:1.0	Running	1月	实时日志 \	Webshell	事件 ∫	1		
	nb5cg	可用区A sae-demo-switch	/sae-demo-image/provider:1.0	Running	3月	实时日志 \	Webshell	事件	重启		
									删除		

5. 在删除实例对话框中单击确认。

单击确认后,页面会自动跳转至变更详情页面。如果执行状态显示执行成功,表示实例删除成功。

交更记录 / 交更详情							
变更流程ID: fca772bb-bc6e-4977-a082-	变更流程ID: fca772bb-bc6e-4977-a082-				分批间处理方式: 自动		
执行状态: 执行成功		变更类型: 删除实例			发布时间: 2021-01-05 21:53:11		
描述信息: N/A		变更对象: N/A					
第1批变更							
❷ 删除实例	实例状态: 全部					批量重试执行失败的实例	刷新
	实例ID		运行状态		配置规格		
	> mysae-1daa7236-3844-4	f36-b39a-	成功		CPU: 0.5 核 内存: 1024MB		
					共有1条,每页显示: 20条	〈 1 〉 前	1 页

6. (可选)如果您选择的部分实例删除失败,您可以在**变更详情**页面单击**批量重试执行失败的实例**,重 新删除实例。

変更记录 / 変更详情								
变更流程ID: fca772bb-bc6e-4977-a082-	发布分批数: 1			分批间处理方式: 自动				
执行状态: 执行成功		变更类型: 删除实例			发布时间: 2021-01-05 21:53:11			
描述信息: N/A		变更对象: N/A						
第1批变更								
❷ 删除实例	实例状态: 全部					批量重试执行失败的家	実例 刷新	
	实例ID		运行状态		配置规格			
	> mysae-1daa7236-3844-41	36-b39a-	成功		CPU: 0.5 核 内存: 1024MB			
					共有1条,每页显示: 20条	< 1 →	前往 1 页	

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



4.4. 使用Webshell诊断应用

本文介绍SAE的Webshell功能,以及SAE应用的网络环境(包括容器内的环境)等运维背景知识,并在此基础上介绍如何利用Webshell完成基本的运维需求。

Webshell简介

您可以通过阿里云控制台直接获取ECS的Shell来完成运维需求。如果ECS内开启了SSH服务,且ECS存在弹性 公网IP,那么在本地可以通过SSH服务获取ECS的Shell完成运维需求。

在SAE的场景中,容器是一个暂态的、供应用运行的环境,通常不需要进行运维。为了方便进行线上问题定 位和排查,SAE在控制台提供了简易版Webshell,供您查看并调试自己的容器。

SAE应用容器的基础镜像是面向应用运行时的且暂态的,因此您的镜像不需要在镜像中启动SSH服务,仅需要带有可执行的 /bin/bash 即可,同时建议您带上所需的运维工具,方便排查。

↓ 注意 Webshell不支持Windows镜像。

查看Webshell

使用Webshell完成基本运维的操作步骤如下:

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面, 单击实例部署信息页签。
- 4. 在实例部署信息页签的默认分组区域,单击目标实例操作列的Webshell。

基本信息			部署应用	回退历史版本	手动扩缩	自动扩缩	停止应	囲 更多 >
 1. 应用在VPC内默认是不能防 2. 如果你的应用需要访问RDS 	问公网的,如果需要访问公网请按以下提示操作:应 数据库,请参考:如何设置RDS白名单	用如何访问公网,常见访问公网场景最佳示例						
基本信息 实例部署信息								
> 弹性伸缩								
✓ 默认分组 状态: 运行24	下实例							
显示实例IP信息								
实例各称 v	Switch	镜像	运行状态 小	运行时间	操作			
	可用区A vswitch	/sae-demo-image/	● Running 未配置健康检查	13小时	实时日志	Webshell 事作	+ 重启 {	删除
	可用区A vswitch	/sae-demo-image,	● Running 未配置健康检查	13小时	实时日志	Webshell 事例	* 重启 {	删除
							<	(1下页)

5. (可选)打开Webshell后,您可以单击窗口右上角的全屏图标,将窗口全屏显示。

connected)	С×
sn-4.2#	

6. 按需在Webshell窗口执行命令, 查看并调试您的容器。

获取实例内部元数据

您通过SAE成功部署应用后,即可获取实例的基本元数据,包括安全组ID、VPC ID和vSwitch ID。

说明				
安全组ID元数据所在的路径。				
VPC ID元数据所在的路径。				
vSwitch ID元数据所在的路径。				
⑦ 说明 如果有多个vSwitch,将以半角逗号(,)分隔开多个 vSwitch ID。				

您可以登录SAE控制台,在Webshell中执行以下命令,验证安全组ID、VPC ID和vSwitch ID的元数据信息。

1. 执行以下命令,进入 podinfo 文件内。

cd /etc/podinfo/

⑦ 说明 如果执行该命令时显示 can't cd to /etc/podinfo/ 报错,您可以选择重启实例或重新部 署应用解决该问题。

2. 执行以下命令, 查看实例元数据信息的所有文件。

ls

预期输出。

sgld vpcld vswitchld

- 3. 按需执行以下命令,验证实例内部相应的元数据信息。
 - 获取安全组Ⅳ。

cat sgld

预期输出。

sg-2ze4ddm2rqgenvcx***sh-4.2#

○ 获取VPC ID。

cat vpcld

预期输出。

vpc-2zeqku2oevokwm66b****sh-4.2#

○ 获取vSwitch ID。

cat vswitchId

预期输出。

vsw-2zewmy4ynhljytzhu****sh-4.2#

应用的网络环境

SAE应用置于自建的VPC网络,并提供了命名空间功能,可以将中间件层面的服务调用进行逻辑隔离。命名 空间与VPC内的vSwitch为绑定关系,一个命名空间仅能对应一个vSwitch,一个vSwitch可以对应多个命名 空间,即VPC内的IP地址为局域网地址,不同VPC内应用间无法访问。命名空间主要用于中间件逻辑隔离,不 同命名空间内的应用在中间件层面是隔离的,如服务发现和配置下发等。

⑦ 说明 关于VPC的原理和产品介绍,请参见VPC基础架构。

鉴于VPC的产品特性和当前的SAE的产品特性,容器无法直接触达VPC外的服务(OSS、镜像服务等阿里云产 品除外)。在没有额外配置的情况下,您的容器单独运行在网络环境,因此您无法直接接触SAE应用容器。

容器无法触达公网的代码示例如下。



容器内如果需要访问公网服务,须购买NAT,并在VPC内配置vSwitch的SNAT规则。具体操作,请参见应用 如何访问公网。

SNAT规则可以让VPC内地址访问公网地址,从而能够调用公网显露的服务,获取公网资源。

构建镜像的方法

基于阿里云容器镜像服务,SAE集成了构建镜像和管理镜像的功能。用于构建的基础镜像为 centos:7,并 为您配置了语言与编码方式、时区和Open JDK运行环境等运行环境。 容器存在的目的是为了让应用运行起来,SAE不可能以占用所有用户运行资源为代价,集成过多的工具。因此,若您对容器内工具有更多需求,请自行构建镜像或者按需从OSS获取。更多信息,请参见制作应用容器 Docker镜像。

诊断应用

通常线上容器运维不必要。如果您需要进入容器并进行运维,存在一定的业务风险:

- 单点应用业务:可能导致容器内存耗尽,从而导致分钟级别的业务中断。
- 多点部署业务:可能导致业务秒级中断。

SAE应用的诊断有常规检查和上传搜集的日志两种方式。

常规检查

常规检查方法众多。以Java应用为例,有进程检查、线程以及JVM的健康状态检查。

○ 执行命令 ps -ef | grep java 检查应用的Java进程是否存在。

⑦ 说明 容器内通常使用主进程启动应用,如果应用被停止,则容器也会退出,SAE自动将退出的容器重新启动,防止业务中断。

如果进程不存在,请执行 dmesg | grep - i kill 命令检查OOM日志。如果日志存在,表示应用的进程被停止,需要检查工作目录下 hs_err_pid{PID}.log 日志文件,并定位具体原因。

Java类型应用在线分析还可以使用阿里巴巴开源软件Arthas,建议在测试镜像中集成Arthas工具进行常规诊断。Arthas能够实时查看Java类加载情况,方便观察方法出参、入参和环境变量等。
 您可以执行以下命令下载并运行Arthas:

wget https://alibaba.github.io/arthas/arthas-boot.jar java -jar arthas-boot.jar

② 说明 接入Arhas前,请先连接公网。

• 日志上传

由于容器内工具匮乏,推荐将容器内搜集到的日志上传到云端,并下载到本地进行分析。 目前SAE没有容器内日志下载功能,由于OSS服务连通阿里云下所有网络环境,推荐使用阿里云OSS服务 进行日志上传下载。

通过OSS服务上传并下载日志的操作方法如下:

i. 在容器内部安装OSS命令行工具。具体操作,请参见安装OSS命令行工具。 本文中以64位Centos系统的root用户为例,在没有打通公网的情况下可以选择在本地下载,然后将 这个文件上传到OSS,然后获取OSS的VPC内地址下载文件。

wget http://gosspublic.alicdn.com/ossutil/1.5.0/ossutil64 chmod 755 ossutil64

ii. 配置所需的OSS命令行工具,并附上当前地域VPC内的Endpoint,填写用于接收上传文件的账号的AccessKey,查看已经创建的Bucket,检查您的OSS服务是否可用。

⑦ 说明 请确保账号(不必是当前账号,任意开通阿里云OSS服务的账号均可)已开通OSS服务。

a. 按照提示配置您的AccessKey、Endpoint信息,无需填写STS Token。

./ossutil64 config

b. 检查账号是否可用,如果报错则配置错误,如果没有Bucket,则建议前往OSS控制台创建,命令 行工具也支持创建。

./ossutil64 ls

c. 创建一个模拟的日志文件, 用于上传。

echo "Hello" > edas-app.log ./ossutil64 cp edas-app.log {bucket-address,例如: oss://test-bucket,可以从上述命令"./ossutil6 4 ls"中查看}

iii. 从OSS控制台或其他工具中找到您的日志文件,下载到本地,并使用您熟悉的工具进行分析。

4.5. 应用变更记录

4.5.1. 查看应用事件

SAE支持查看K8s原生应用的事件,帮助您了解应用运行时的状态,方便快速聚焦问题。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面的左侧导航栏,单击应用事件,查看应用事件。

← 应用列表	(sprin	ng-boot-cons	umer)				
基本信息	应用事件	14					
交更记录	0.8	息计 25185 项记录,当前仅保	留7天内的应用事件。				
应用事件 日志管理 ✓	全部将	来源英型 >	全部事件等级	∨ 来源名	✓ 清緒入来源名,比如应用名// Q		C
基础监控		事件等级	事件原因	来源类型	来源名	首次发生时间	最近发生时间
应用监控 ~	+	Normal	Killing	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:02:30	2021-01-27 14:03:30
服务列表	-	Warning	Unhealthy	Pod	spring-boot-consumer-7e579e20-5a12-41e3-(2020-12-30 17:02:10	2021-01-27 14:03:10
限流降级(仅支持Java) 🏏 远程调试(仅支持Java)		Liveness probe failed: 自排查建议: 健康检测(live	ness or readiness)失败,请确	认健康检查配置是否正确 点击	2#		
通知报警 🗸	+	Normal	Started	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:01:51	2021-01-27 14:02:53
	+	Normal	Created	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:01:49	2021-01-27 14:02:51
	+	Normal	Pulled	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:01:49	2021-01-27 14:02:51
	+	Normal	Pulling	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:01:49	2021-01-27 14:02:50
	+	Normal	Killing	Pod	spring-boot-consumer-7e579e20-5a12-41e3-	2020-12-30 17:02:30	2021-01-27 14:02:20

- (可选)您可以设置事件搜索过滤条件,按需查看应用事件。
 搜索条件如下:
 - 来源类型:包括应用实例(Pod)、SLB(Service)、应用(Deployment)、自动弹性 (HorizontalPodAutoscaler)。
 - 事件等级: 包含Warning和Normal, Warning级别事件需要关注。
 - 来源名:选择事件的来源名,例如应用名称、应用实例名称。
 - 其他: 输入来源名, 例如应用名称, 单击 Q 图标, 即可按关键字搜索应用事件。

后续操作

查看应用事件后,您可以依据具体的事件信息以及您的业务,进行相应的应用操作。

• 管理应用生命周期

- 查看变更记录
- 变更实例规格

常见问题

部署在SAE上的应用处于运行状态时,出现的问题一般为Pod问题。关于Pod问题,请参见Pod问题。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



4.5.2. 查看变更记录

当您在SAE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。

操作指南

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用**基本信息**页面左侧导航栏,单击**变更记录**,查看该应用的所有变更操作以及该操作的变更状态。

⑦ 说明 在执行完应用变更操作后,在应用基本信息页面上方出现变更提示应用有变更流程正在执行,处于执行中状态,您可以单击查看详情阅览该应用的变更信息及实时状态。

变更记录								
所有变更类型	✓ 所有変更状态	~ 模糊搜索	> 根据描述信息进行模糊搜索	Ķ Q				G
创建时间	结束时间	变更类型	描述信息	变更状态	变更人	来源	操作	
2020-12-30 16:34:54	2020-12-30 16:35:48	部署应用	版本: 1609317203289 包 名: spring-boot-consumer -1.0-SNAPSHOT.jar	● 执行成功	14,01007000040	console	효용	
2020-12-30 16:26:45	2020-12-30 16:27:38	部署应用	版本: 1609316802473 包 名: spring-boot-consumer -1.0-SNAPSHOT.jar	● 执行成功		console	查看	
2020-12-24 15:01:45	2020-12-24 15:02:35	绑定SLB		● 执行成功	10.0	console	查看	
2020-12-24 15:00:01	2020-12-24 15:01:17	部署应用	版本: 1608793197574 包 名: spring-boot-consumer -1.0-SNAPSHOT.jar	● 执行成功		console	查看	
2020-11-19 21:23:53	2020-11-19 21:24:36	启动应用		● 执行成功	主账号	console	查看	
2020-11-19 21:05:19	2020-11-19 21:06:06	停止应用		● 执行成功	主账号	console	查看	

4. 在操作列单击查看,查阅变更详情及操作的详细信息。

変更记录 / 変更详情					
变更流程ID: 3f36cd3f-809e-		发布分批数: 1	分批间处理方式: 自动		
执行状态: 执行成功		· 変更类型: 部署应用	发布时间: 2020-12-30 16:34:54		
描述信息: 版本: 1609317203289 包名: spring-boot	t-consumer-	変更对象: 默认分组			
第1批变更					
❷ 部署应用			刷新		
	构建镜像	⊘ 成功	>		
	初始化环境	📀 成功	>		
	创建或更新部署配置	🕑 ಹೆಸು	>		
	执行应用部署	⊘ 成功	>		

该页面包含变更概要信息和变更流程执行信息。

- 变更概要信息:包括变更流程ⅠD、执行状态、变更类型等信息。
- · 变更流程执行信息:包含整个变更流程,以及变更阶段中具体的变更任务。
- 5. 在变更流程执行信息区域的右侧展开具体任务,可以查看该变更过程中本环节的操作日志。

如果流程中出现失败任务,系统自动展开日志,请排查任务失败原因。更多信息,请参见常见问题。

常见问题

如果应用的变更记录中出现失败记录,您可以在变更详情中查询变更流程中哪些步骤执行失败,以及失败的 原因。

排查变更失败问题

1. 在应用基本信息页面左侧的导航栏,单击变更记录。

変更记录							
所有变更类型	✓ 所有変更状态	∨ 模糊搜索	> 根据描述信息进行模糊搜	ж Q			C
创建时间	结束时间	变更类型	描述信息	变更状态	変更人	来源	操作
2020-12-31 15:00:07	2020-12-31 15:00:08	部署应用	版本: 1590044655025 镜 像名称:	● 执行失败	0.000	console	查若
						毎页显示: 10 🗸	- 共1条 〈 上一页 1 下一页 〉

如果变更失败,则在**变更状态**列会显示执行失败。

- 2. 在执行失败的变更记录所在行,单击查看。
- 3. 在**变更详情**页签,在变更区域单击失败的阶段,例如**部署应用**,在右侧实例中查看失败步骤的具体日志,定位问题。

更流程ID: adf434b5-8ebd-4ae2		发布分批数: 1	分批间处理方式: 自动	
行状态: 执行失败 🔒 (变更流程错误排查)	皆南)	变更类型: 部署应用	发布时间: 2020-12-31 15:00:07	
述信息: 版本: 1590044655025 镜像名称: 👘	and the second sec	变更对象: 默认分组		
◎ 部署应用	初始化环境 🔒	⊗ 失敗		
	auto config failed, the			

应用变更典型异常及处理

端口健康检查异常导致变更失败

変更记录 / 変更详情							
变更筛程D: 分批间处理方式: 自动							
执行状态: 🖲 执行失败 🕄		发布人:	发布时间: 2019-03-05 17:54:39				
描述信息:版本:V1.0.1 包名:	in product 2-040/04/04	变更类型: 部署应用	变更对象:默认分组				
第1批变更 1							
✓ SLB下线	实例状态 : 全部	◇ 查看全部任务:					
⊗ 执行部署 2	IP	运行状态	配置规格				
● SLB上线	~	● 失败	CPU:1核 内存:2048MB				
3 第口健康检查 ① 含 実改 3 [B1Z] 2019-03-05 17.56 57.508 Task timeout when running 1 exec failed [agent] 2019-03-05 17.54 56 CST [0322]. Taskid: 1 - 32 - healthCheckWihPort appPort.8000, result. Failes [B1Z] 2019-03-05 17.54 56 CST [0322]. Taskid: 1 - 48 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0322]. Taskid: 1 - 48 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0402]. Taskid: 1 - 32 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0400]. Taskid: 1 - 32 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0400]. Taskid: 1 - 32 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0400]. Taskid: 1 - 32 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2019-03-05 17.54 56 CST [0400]. Taskid: 1 - 48 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2011-03-05 17.54 56 CST [0400]. Taskid: 1 - 42 - healthCheckWihPort backendPort 0, result. Failes 1 [agent] 2011-03-05 17.54 56 CST [0400]. Taskid: 1 - 43 - healthCheckWihPort 1000 LankendPort 0, result. Failes 1							

• 可能原因

- 应用启动时应该被正常占用的65000端口不可访问。
- 客户的Tomcat容器没有正常启动。
- 处理方法

查看应用日志是否有相应的异常日志。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



4.6. 远程调试

4.6.1. Java远程调试

如果您想要远程诊断部署在SAE上的应用,可以使用SAE远程调试功能。SAE目前仅支持Java应用的远程调试。

配置Java应用远程调试

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用详情页面的右上角,单击部署应用。
- 4. 在部署应用页面下方,展开启动命令设置区域,设置启动命令。设置完成后单击确认。

系统默认启动命令	\$JAVA_HOME/bin/java \$Options -jar \$CATALINA_OPTS "\$package_path" \$args
	启动命令格式说明:Java [-Options] -jar jarfile[args]
options设置	-agentlib:jdwp=transport=dt_socket,address=9000,server=y,suspend=n
arcs设置	
2	

参数说明如下。

参数	说明
系统默认启动命令	SAE默认的启动命令。
	请配置以下命令:
options设置	- agentlib:jdwp=transport=dt_socket,address= 9000,server=y,suspend=n
	transport:远程调试间的数据传输方式。address:远程调试的端口。
args设置	配置标准输出和错误输出的重定向命令,例如 1>>/t mp/std.log>&1 。

↓ 注意 单击确认后,该应用将会被重启,请在业务较少的时间段部署。

- 5. 设置启动命令后,在基本信息页面的左侧导航栏,单击远程调试(仅支持Java)。
- 6. 在远程调试 (java) 页面, 打开开启远程调试开关, 单击远程调试设置。

远程调试 (java)	
● 如果您想要远程诊断部署在SAE上的应用,您可以使用远程调试功能。该功能目前仅支持Java应用,长时间开启远程调试可能存在应用数据泄露等风险,因此远程诊断后请及时关闭该功能,详情请参见如何使用远程调试 (java)。	
开启远程调试	

7. 在远程调试设置面板中,设置远程调试参数,单击确定。

∨ *远程调试ip和端口设置						
请选择公网SLB	新建SLB		\sim			
检查项目	状态	说明				
SLB配额检查	成功♥					
账户余额检查	成功♥					
产品类别	产品配置	数量	付费方式	购买周期	资费	
负载均衡SLB - 公网	地域:华南1 公网带宽 : 按使用流量计费	1	按量付费	无	查看价格	
调试端口						
∨ *远程调试指定实例						
_DEFAULT_GROUP	~					
实例名称	实例ip		镜像		运行状态	
○ test- 88e9-8	1. The second se		registry-vpc.cr	n-shenzhen.aliyu	Running	
O test 88e9-8	4-10- A-10-1		registry-vpc.cr	n-shenzhen.aliyu	Running	
			每页显示	10 ~	,共2条 く 上一页	1 下一页 >
确定 取消						

远程调试参数说明如下。

参数	说明
请选择公网SI B	选择SLB完成本地与远程实例的连接。如果您之前没有 创建过SLB,您可以选择 新建SLB 来创建SLB进行连 接。关于SLB的详细说明,请参见 <mark>什么是传统型负载均</mark> 衡CLB。
	↓ 注意 使用SLB时会产生额外费用,SLB的 计费方式,请参见按量计费。
调试端口	输入 <mark>步骤4</mark> 中设置应用启动命令时配置的address调试 端口。
	选择需要远程调试的实例。
远程调试指定实例	⑦ 说明 只有运行状态为Running的实例才 可以设置远程调试。

远程实例设置完成后,页面自动进入变更详情页面,查看实例是否连接成功。

变更记录 / 变更详情							
变更流程ID: 14aa1c9a-b7aa-4a12-94	aa-a1baa69f7e03	发布分批数: 1	分批调处理方式: 自动				
执行状态: 执行成功		变更类型: 绑定SLB	发布时间: 2020-07-28 14:05:54				
描述信息: N/A		变更对象: N/A					
第1批变更	第1批变更						
C HERE'S R							
	配置SLB信息 🥑 成功		>				

当执行状态为执行成功时,说明实例已经成功连接。

8. 在左侧导航栏单击远程调试(仅支持Java),获取远程调试IP及端口信息,对应用进行远程诊断。

远程调试 (java)						
① 如果您想要远	程诊断部署在SAE上的应用,	您可以使用远程调试功能。	该功能目前仅支持Java应用,	长时间开启远程调试可能存在应用数据泄露等风险	,因此远程诊断后请及时关闭该功能,	详情请参见如何使用远程调试(java)。
开启远程调试						
远程调试IP和端口						
远程调试指定实例	test-					
修改远程调试设置						

↓ 注意 长时间开启远程调试可能存在数据泄露等风险,因此远程诊断完后请及时关闭开启远程调试开关,在弹出的确认解绑远程Debug对话框中单击确认,关闭远程调试功能。

修改远程调试设置

SAE一次仅支持一个应用的远程调试,如果您需要远程调试其他应用,请修改远程调试设置。

↓ 注意 修改远程调试设置前,请先确保需要远程调试的应用已设置启动命令。关于如何设置启动命令,请参见配置Java应用远程调试。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 单击具体应用名称。
- 3. 在基本信息页面的左侧导航栏,单击远程调试(仅支持Java)。
- 在远程调试(java)页面,单击修改远程调试设置。在远程调试设置面板中修改设置信息,单击确 定。

. 修改	了远程调试IP和端口设置后,原有设置的	调试IP和端口会立即不生效。		
◇ *远程	调试ip和端口设置			
调试端口	8080			
◇ *远程	调试指定实例 test-	The second dense to		
_DEFAI	ULT_GROUP	~		
	实例名称	实例ip	镜像	运行状态
۲	test 88e9-8	40.000	registry-vpc.cn-shenzhen.aliyu	Running
0	test- 88e9-8	10.00	registry-vpc.cn-shenzhen.aliyu	Running
			每页显示: 10 🗸 共2务	• 〈 上一页 】 下一页 〉
确定	取消			

修改远程调试的参数说明如下所示。

参数	说明
请选择公网SLB	选择SLB完成本地与远程实例的连接。如果您之前没有 创建过SLB,您可以选择 新建SLB 来创建SLB进行连 接。关于SLB的详细说明,请参见 <mark>什么是传统型负载均</mark> 衡CLB。
	↓ 注意 使用SLB时会产生额外费用,SLB的 计费方式,请参见按量计费。
调试端口	输入 <mark>步骤4</mark> 中设置应用启动命令时配置的address调试 端口。
	选择需要远程调试的实例。
远程调试指定实例	⑦ 说明 只有运行状态为Running的实例才 可以设置远程调试。

远程实例设置完成后,页面自动进入变更详情页面,查看实例是否连接成功。

变更记录 / 变更详情			终止变更
变更流程ID: 14aa1c9a-b7aa-4a12-94	aa-a1baa69f7e03	发布分批数: 1	分批间处理方式: 自动
执行状态: 执行成功		变更类型: 绑定SLB	发布时间: 2020-07-28 14:05:54
描述信息: N/A		变更对象: N/A	
第1批变更			
Ø 绑定SLB			Reter
	配置SLB信息 🛛 🕑 成功		>

当执行状态为执行成功时,说明实例已经成功连接。

5. 在左侧导航栏单击远程调试(仅支持Java),获取远程调试IP及端口信息,对应用进行远程诊断。

远程调试 (java)	
① 如果您想要远程诊断部署在	SAE上的应用,您可以使用远程调试功能。该功能目前仅支持Java应用,长时间开启远程调试可能存在应用数据泄露等风险,因此远程诊断后请及时关闭该功能,详情请参见如何使用远程调试(java)。
开启远程调试	
远程调试IP和端口	
远程调试指定实例	test
修改远程调试设置	

↓ 注意 长时间开启远程调试可能存在数据泄露等风险,因此远程诊断完后请及时关闭开启远程调试开关,在弹出的确认解绑远程Debug对话框中单击确认,关闭远程调试功能。

验证远程调试结果

您可以通过应用的变更记录查看远程调试是否成功开启或关闭。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击**应用列表**,在**应用列表**页面上方选择**地域**,单击具体应用名称。
- 3. 在基本信息页面的左侧导航栏,单击变更记录。

变更记录								
所有变更类型	✓ 所有変更状态	> 模糊投索	> 根据描述信息进行模糊搜索 Q					с
创建时间	结束时间	交更类型	描述信息	变更状态	变更人	来源	操作	
2020-07-28 14:08:45	2020-07-28 14:08:45	解鉀SLB		● 执行成功	1.1000	console	重要	
2020-07-28 14:05:54	2020-07-28 14:06:45	绑理SLB		● 执行成功		console	重要	
2020-07-28 11:48:18	2020-07-28 11:48:19	解師SLB		● 执行成功	11000	console	重要	
2020-07-28 11:35:26	2020-07-28 11:36:08	绑理SLB	-	• 执行中止	1.10000.000000	console	查看	

变更记录说明如下所示:

- 绑定SLB成功说明远程调试已成功连接。
- 解绑SLB成功说明远程调试已成功关闭。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



4.6.2. 实例远程访问及上传下载文件

OpenSSH是SSH(Secure Shell)协议的免费开源实现,OpenSSH提供了服务端后台程序和客户端工具,用 来加密远程控件和文件传输过程的中的数据。您可以通过安装OpenSSH,借助远程调试的能力实现SSH远程 登录以及SFTP、SCP上传下载文件的功能,进而方便您临时调试SAE的某个应用实例。

前提条件

如果您的应用配置了安全组,请确保22端口(SSH端口)允许开放。

步骤一:开启远程调试

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用基本信息页面,单击实例部署信息。在实例部署信息页签的默认分组区域,单击目标实例操 作列的Webshell。
- 4. 在命令框中执行passwd命令,为当前Root角色设置访问密码,作为SSH, SFTP远程登录密码。



⑦ 说明 您也可以新建用户并为其设置密码。

5. 在命令框中执行以下命令安装OpenSSH。

(?) 说明 对于以Centos及Apline操作系统为基础镜像的应用,SAE提供bash脚本便于您快速安装 openssh。如果您需要自定义OpenSSH相关配置或者因自身镜像原因导致执行脚本异常,请自行安 装OpenSSH。

各地域的安装命令如下所示。

地域	安装命令
华东1(杭州)	weet http://edas-hz.oss-cn-hangzhou-internal.a livuncs.com/sae_openssh.sh && chmod 777 sae_ openssh.sh && ./sae_openssh.sh
华东2(上海)	weet http://edas-sh.oss-cn-shanghai-internal.al ivuncs.com/sae_openssh.sh && chmod 777 sae_o penssh.sh && ./sae_openssh.sh

地域	安装命令
华北2(北京)	weet http://edas-bi.oss-cn-beiiing-internal.alivu ncs.com/sae_openssh.sh && chmod 777 sae_ope nssh.sh && ./sae_openssh.sh
华南1(深圳)	weet http://edas-sz.oss-cn-shenzhen-internal.al ivuncs.com/sae_openssh.sh && chmod 777 sae_o penssh.sh && ./sae_openssh.sh
华北3(张家口)	wget http://edas-zik.oss-cn-zhangijakou-intern al.alivuncs.com/sae_openssh.sh && chmod 777 s ae_openssh.sh && ./sae_openssh.sh

返回结果如下,说明安装成功。

3e26-nlwlq (connected)		ĽС>
Running transaction		
Installing : fipscheck-1.4.1-6.el7	.x86_64	1/5
Installing : fipscheck-lib-1.4.1-6	5.el7.x86 64	2/5
Installing : openssh-7.4p1-21.el7.	x86 64	3/5
Installing : tcp wrappers-libs-7.6	-77.el7.x86 64	4/5
Installing : openssh-server-7.4p1-	-21.el7.x86 64	5/5
Verifying : fipscheck-lib-1.4.1-6	5.el7.x86 64	1/5
Verifying : tcp wrappers-libs-7.6	5-77.el7.x86 64	2/5
Verifying : fipscheck-1.4.1-6.el7	- .x86 64	3/5
Verifying : openssh-7.4p1-21.el7.	x86 64	4/5
Verifying : openssh-server-7.4p1-	-21.el7.x86_64	5/5
Installed: openssh-server.x86_64 0:7.4p1-21.e	217	
Dependency Installed:		
fipscheck.x86_64 0:1.4.1-6.el7	fipscheck-lib.x86_64 0:1.4.1-6.el7	
openssh.x86_64 0:7.4p1-21.el7	tcp_wrappers-libs.x86_64 0:7.6-77.el7	
Name lata l		
download openssn-server success		
run sshd start		
run sshd success		
sh-4.2#		

- 6. 开启实例远程调试。
 - i. 在左侧导航栏单击**远程调试(仅支持Java)**,在远**程调试(java)**页面,打开**开启远程调试**开 关。
 - ii. 单击远程调试设置,在远程调试设置面板,设置调试信息,单击确定。

← 远程调试设置					×
✓ *远程调试ip和端口设置					
请选择公网SLB	新建SLB	``	/		
检查项目	状态	说明			
SLB配额检查	成功 🗢				
账户余额检查	成功♥				
产品类别	产品配置	数量	付费方式	购买周期	资费
负载均衡SLB - 公网	地域:华东1 公网带宽:按使用流量计费	1	按量付费	无	查看价
调试端口	22				
∨ *远程调试指定实例	4833-8f3a-78d3be433e	26-nlwlq			
_DEFAULT_GROUP	\sim				
实例名称	实例ip		镜像		运行
•1dd3-483 7	3-8f3a-				● F
O 7	.0.73 (0.73				• F
				每页显示: 10 🔻	/ 共2条
(÷
确定 取消					

参数说明如下:

- 请选择公网SLB: 您可以选择新建SLB实例或选择已有的SLB实例。
- 调试端口:设置为22。
- 远程调试指定实例:选择需要调试的目标实例。

开启远程调试后,您可以在远程调试 (java)页面,获取远程调试IP和端口。

远程调试 (java)	
① 如果您想要远程诊断部署在	SAE上的应用,您可以使用远程调试功能,该功能目前仅支持Java应用,长时间开启远程调试可能存在应用数据泄露等风险,因此远程诊断后请及时关闭该功能,详情请参见如何使用远程调试(java)。
开启远程调试	
远程调试IP和端口	123: 22
远程调试指定实例	4833-8f3a-78d3be433e26-nlwlq
修改远程调试设置	

步骤二:执行远程调试

远程调试开启成功后,您可以通过获取的远程调试IP和端口,在可以访问公网,且具有SSH Client 的主机中完成实例的远程调试。

• 远程调试:执行ssh root@ip命令,登录目标实例,完成远程调试。

[root@	-40ea-96dc-e6e04bfd340f-pkdrk ~]# ls	5
anaconda-ks.cfg	nacos	
[root@	-40ea-96dc-e6e04bfd340f-pkdrk ~]# pw	vd
/root		
[root@	-40ea-96dc-e6e04bfd340f-pkdrk ~]#	

• 远程上传下载文件:执行sftp root@ip命令,登录目标实例,远程上传下载文件。

sftp> ls				
anaconda-ks.cfg nacos				
sftp> put a.txt				
Uploading a.txt to /root/a.txt				
a.txt	100%	3	0.6KB/s	00:00
sftp> ls				
a.txt anaconda-ks.cfg nacos				
sftp> get anaconda-ks.cfg				
Fetching /root/anaconda-ks.cfg to anaconda-ks.cfg				
/root/anaconda-ks.cfg	100% 34	15	389.5KB/s	00:00
sftp>				

○ 注意

- 如果您在步骤4中新建了用户,您需要将命令中的root改成您新建的用户名。
- 本方案仅作为临时调试。当您的应用重新进行部署时,或者执行了任何会导致实例重启的操作或 者发生的现象,您都需要重新完成实施步骤的配置:配置登录密码,安装OpenSSH,修改远程调 试的目标实例。
- 当您修改远程调试的目标实例后再次登录时,需先执行ssh-kevgen-Rip命令,删除原有证书,否则会出现WARNING: REMOTE HOST IDENTIFICATION HAS CHANGED Host key verification failed.
 错误。

5.服务治理

5.1. 限流降级 (Java)

5.1.1. 规则管理

5.1.1.1. 配置流控规则

SAE可以通过AHAS监控应用或服务流量的QPS指标,当指标达到流控规则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,从而保障应用的高可用性。本文介绍配置流控规则的操作步骤,以及三种常用 场景的流控配置规则。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见<mark>设置限流降级</mark>。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 规则管理。
- 4. 在规则管理页面,单击流控规则,在流控规则页签,单击新增流控规则。
- 5. 在新增流控规则对话框,设置参数,单击新建。

参数说明如下:

○ 单机模式:关闭集群流控的状态。

default		
● 当前接口	○ 关联接口	○ 链路入口
用于接口调用流热 前接口来自于来源	2。该接口被来源应用调 原应用的请求进行流控	用次数超过阈值时, 会对当
100		
● 快速失败	○ 預热启动	○ 排队等待
常规流控方式。当 息,如文本/静态	≦前接□超过设置阈值的 页面等。	流量,直接返回默认流控信
() 该规则打	斤, 创建后即生效	
	default ● 当前後口 用于输口调用顶距 前線口湯自手未退 100 ● 快速失敗 第回加起方式、当 思、如文本術表: ● 検測別打去	default ・ ・ ・

待流控的资源名称。

描述

接口名称

参数	描述
来源应用	 该规则针对的来源应用,默认来源应用设为 default,代表不区分来源应用。 若是Dubbo服务,请填写对应调用方的Dubbo applicationName,注意调用方也需要接入 Sentinel。 若是Web服务,请参见扩展接口。 若是通过注解、自定义埋点,请参见常用类及其方法,ContextUtil只在调用链入口生效,即首个埋点生效。 当前账号下已经接入了流量防护的其他应用,填写 default 即表示不特殊区分。 ↓注意 来源数目不宜过多,200个以内内存使用量可控,过多则会影响性能。
	の一定用B 「 「 「 」 」 」 」 」 」 」 」 」 」 」 」 」
统计维度	 选择资源调用关系进行流控。 当前接口:直接控制来自来源应用中调用来源的 访问流量,如果来源应用为 default 则不区分调 用来源。通常应用于流量匀速通过的场景,更多信息,请参见常用场景1。 关联接口:控制当前资源的关联资源的流量。通 常应用于资源争抢时,留足资源给优先级高接口的 场景,更多信息,请参见常用场景2。 链路入口:控制该资源所在的调用链路的入口流 量。选择链路入口后需要继续配置入口资源,即 该调用链路入口的上下文名称。通常应用于预热启 动避免大流量冲击的场景,更多信息,请参见常用 场景3。
单机QPS 阈值	触发对流控接口的统计维度对象的QPS阈值。

参数	描述
流控效果	 选择流控方式来处理被拦截的流量。 快速失败:达到阈值时,立即拦截请求。按照应用系统设置中的适配模块配置信息,进行内容返回。 预热启动:需设置具体的预热时间。详情请参见Warm Up(冷启动)。 如果系统在此之前长期处于空闲的状态,当流量突然增大的时候,该方式会让处理请求的速率缓慢增加,经过设置的预热时间以后,到达系统处理请求速率的设定值。默认会从设置的QPS阈值的1/3开始慢慢往上增加至设置的QPS值,多余请求会按照快速失败处理。 排队等待:请求匀速通过,允许排队等待,通常用于请求调用削峰填谷等场景。需设置具体的超时时间,达到超时时间后请求会快速失败。详情请参见削峰填谷。
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。开关修改之后会立即生效。

• 集群流控模式:开启集群流控的状态。

* 接口名称	/doAnotherThing
是否集群流控 🚺	
*接口集群总 QPS 🚯	1000
阔值模式 🚯	 集群阈值 单机阈值
	即以集醉作为整体来查看词值。
*集群阈值	100
* 统计窗口时长	1
	Ð ∨
通信请求模式 🚺	 每次请求 批量请求
失败退化策略 🚯	 退化到单机限流 直接通过
	当出现连接失败、通信失败或 Token Server 不可用时,退化到根据单 机阈值来进行流控。

参数	描述
是否集群流控	开启集群流控,对集群内此资源的调用总量进行限 制。
接口集群总QPS	该接口预估的集群最大QPS,表示最大流量,用于为 Token Server自动分配提供参考,当流量超出该值的 请求会退化到单机模式。

参数	描述
阈值模式	 可选择设置集群阈值或单机QPS阈值。 集群阈值:设置的阈值等同于整个集群的总阈值。 单机QPS阈值:设置的阈值等同于单机能够承受的限额,Token Server会根据连接数来计算总的阈值。
失败退化策略	 当出现连接失败、通信失败或Token Server不可用等情况时,流控规则是退化到单机限流的模式或是直接通过忽略失败情况。 退化到单机限流:当出现失败的情况时,退化到根据的单机阈值来进行流控。需要设置退化单机阈值,代表单机的兜底阈值。 直接通过:当出现失败的情况时,忽略失败情形,直接通过。

常用场景1: 削峰填谷, 使流量匀速通过

请求流量具有波峰波谷的特点,流控的原理是将前面的峰值流量延迟(排队时长)到后面再处理,既能最大 化满足所有请求,又能保证用户体验。详情请参见<u>削峰填谷</u>。

在新建流控规则对话框中配置以下规则信息:

- 统计维度选择当前接口。
- 流控效果选择排队等待。
- 配置匀速模式下请求单机QPS阈值为5。
- 等待时长为5s。

系统则每200 ms处理一条请求,多余的处理任务将排队;同时设置了等待时长为5s,则预计排队时长超过5s 的处理任务将快速失败,直接返回默认流控信息,如文本、静态页面等。

*接口名称	/doAnotherThing
是否集群流控 🚺	
* 来源应用	default
统计维度 🚺	 ● 当前接口 ○ 关联接口 ○ 链路入口
	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
*单机QPS阈值	5
流控效果 🕦	
	适用于流星勾速器场景。脱发流控后,多余请求会按照则穿萼符,这到萼 得时长后失败
超时时间	5000 ms
是否开启	该规则打开, 创建后即生效

常用场景2: 当资源争抢时, 需留足资源给优先级高的接口

read_db 和 write_db 这两个资源分别代表数据库读写。为保证提交的数据不丢失, write_db 接口优先 级更高。当写库操作过于频繁时,读数据的请求会被限流。详情请参见关联限流。

在新建流控规则对话框中配置以下规则信息:

- 统计维度选择关联接口。
- 流控效果选择快速失败。
- 并发数阈值为10。

当 write_db 资源的QPS超过10之后, read_db 会被限流以保证留足资源给 write_db , 避免 write_db 数 据丢失。

* 接口名称	read_db
是否集群流控 🚺	
* 来源应用	default
統计维度 🚯	○ 当前接口 ● 关联接口 ● 链路入口
	用于资源争给情况。当关联接口被来源应用调用QPS超过阈值时,会对当前接口来目于来源应用的请求进行流控
* 关联接口名	write_db
* 并发数阈值	10
流控效果 🚺	 使速失败 预热启动 排队等待
	常规添控方式。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。
是否开启	该规则打开, 创建后即生效

常用场景3:预热启动避免大流量冲击

流控的原理是在流量入口处控制流量,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,以便系统可以预热。最适合突发流量的场景。详情请参见Warm Up(冷启动)。

在新建流控规则对话框中配置以下规则信息:

- 统计维度选择链路入口。
- 流控效果选择预热启动。
- 单机QPS阈值为60。
- 预热时间为2s。

预热流控方式下,默认会从设置的QPS阈值的1/3开始慢慢往上增加至QPS设置值。本示例中,当入口的QPS 超过20(即60÷3)时,会在预热的2s内缓慢增长至60。

* 接口名称	funtion
是否集群流控 🚺	
* 来源应用	default
统计维度 🌔	○ 当前接口 ○ 关联接口 ● 链路入口
	用于应用內callstack调用流控情况。当callstack入口被未愿应用调用 QPS数适过阈值时,会对当前接口来自于未源应用的源实进行流控
* callstack入口	test
*单机QPS阈值	60
流控效果 🚺	○ 快速失敗 ● 預热启动 ○ 排队等待
	适用于避免束增流量瞬间中击系统的场景。请求流量会在预热时间内 缓步增加至阈值量级,多余请求会直接把绝掉。
预热时间	2 5
是否开启	该规则打开, 创建后即生效

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.1.2. 配置隔离规则

隔离规则是通过控制接口或依赖的并发线程数,来保证系统的稳定性。该规则适用于应用内部或下游依赖出现不稳定的场景,例如慢SQL、下游应用响应时间变长等。本文介绍配置隔离规则的操作步骤。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见设置限流降级。

背景信息

当强依赖的方法或接口出现不稳定的时候,可以通过配置并发线程数来限制不稳定的强依赖并发数,起到隔 离异常的效果。若运行该请求的响应时间变长,会导致线程的并发数变大。当并发数超过阈值以后,AHAS 将拒绝多余的请求,直到堆积的任务完成,并发线程数变少。达到将异常隔离,减小不稳定性的效果。

如何设定并发线程数阈值,可参见以下内容:

- 并发线程数 = 期望QPS*响应时间+冗余量。
- 例如预期的SQL执行时间为20毫秒,预期该请求每秒有20个,并发最大时候是6个,建议并发线程数按照以下逻辑设置: Max (20/1000*20, 6) = 6,再加上冗余量2,则建议并发数阈值设置为8。
- 设置好后,当这个SQL发生死锁或者有性能问题,SQL运行特别慢成为慢SQL时,即使请求不断的进来,也 仅仅会占用8个线程,不会因为持续进来的请求(请求也无法在短时间内退出),从而耗光进程的活跃线 程。
- 当这个SQL恢复正常后,并发数会迅速减少。当并发数减少至低于预设的阈值时,系统就不会拒绝请求,应用的处理能力也快速的恢复。通过这样的方式,起到了根据响应时间自动调节的效果,隔离了不稳定的应用。

隔离规则配置通常用于强依赖隔离场景,详情请参见强依赖隔离。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 规则管理。
- 4. 在规则管理页面,单击隔离规则,在隔离规则页签,单击新增隔离规则。
- 5. 在新增流控规则对话框,设置参数,单击新建。

新增隔离规则 🕦	×
1 根据实时占用并发	线程数来控制接口流量,超过阈值时通过配置的方式进行流控。查看详情
* 接口名称	请输入接口名称
是否开启	该规则打开, 创建后即生效
* 来源应用	default
统计维度 🚹	● 当前接口 ○ 关联接口 ○ 链路入口
	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
* 并发数阈值	阈值
	▲ 隐藏高级选项
	新建取消

参数说明如下:

参数	描述
接口名称	待隔离的资源名称。

参数	描述
来源应用	 该规则针对的来源应用,默认来源应用设为default,表示不区分来源应用。 若是Dubbo服务,请填写对应调用方的DubboapplicationName,注意调用方也需要接入Sentinel。 若是Web服务,请参见扩展接口。 若是通过注解、自定义埋点,请参见常用类及其方法,ContextUtil只在调用链入口生效,即首个埋点生效。 当前账号下已经接入了流量防护的其他应用,填写default即表示不特殊区分。 〕注意 来源数目不宜过多,200个以内内存使用量可控,过多则会影响性能。
	default为全部来源应用
统计维度	 选择资源调用关系进行隔离流控。 当前接口:直接控制来自来源应用中调用来源的访问流量,如果来源应用为default则不区分调用来源,通常应用于保障自身资源充足的场景,请参见常用场景1。 关联接口:控制当前资源的关联资源的流量。通常应用于资源争抢时,留足资源给优先级高接口的场景,请参见常用场景2。 链路入口:控制该资源所在的调用链路的入口流量。选择链路入口后需要继续配置callstack入口,即该调用链路入口的上下文名称。通常应用于接口有多入口资源的场景,请参见常用场景3。
并发数阈值	资源的并发线程数(即该资源正在执行的线程数)阈 值。

常用场景1:保障自身资源充足

当运行该请求的响应时间变长,会导致线程的并发数变大。当并发数超过阈值以后,AHAS将拒绝多余的请求,直到堆积的任务完成,并发线程数变少。达到将异常隔离,减小不稳定性的效果。例如某个SQL执行时间为20毫秒,预期该请求每秒有20个。

在新建隔离规则对话框中配置以下规则信息:

- 填写接口名称和来源应用。
- 统计维度选择当前接口。
- 并发数阈值为10。

* 接口名称	handleServiceK
* 来源应用	default
统计维度 🚺	● 当前接口 关联接口 链路入口
	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 口来目于来源应用的请求进行流控
* 并发数阈值	10
是否开启	该规则打开,创建后即生效

设置完成后,当这个SQL发生死锁或者存在性能问题时,该SQL运行变慢,成为慢SQL,此时即使请求不断进来,也仅仅会占用10个线程,不会因为持续进来的请求(请求也无法在短时间内退出),从而耗光进程的活跃线程。当这个SQL恢复正常后,并发数会迅速减少。当并发数减少至低于预设的阈值时,系统就不会拒绝请求,应用的处理能力也快速的恢复。通过这样的方式,起到了根据响应时间自动调节的效果,隔离了不稳定的应用。

常用场景2:有一定相关联性的接口

当关联接口被来源应用调用OPS超过阈值时,会对当前接口来源应用的请求进行限流,有一定的相关性的方法来配置规则。例如 read_db 和 write_db 这两个资源分别代表数据库读写, write_db 接口优先级更高。

为保证读写资源争抢时, write_db 的接口可以留足资源, 可在新建隔离规则对话框中配置以下规则信息:

- 接口名称为 write_db 。
- 统计维度选择关联接口。
- 关联接口名为 read_db 。
- 并发数阈值为10。

* 接口名称	write_db
* 来源应用	default
统计维度 👔	 ○ 当前接口 ● 关联接口 ○ 链路入口
	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经
* 关联接口名	read_db
* 并发数阈值	10
是否开启	(该规则打开, 创建后即生效

这样在 read db 接口被调用QPS超过10次后, 会对 write_db 接口来自于来源应用的请求进行隔离限流, 保证 write_db 的足够资源。

常用场景3:针对入口链路来配置隔离规则

从入口处将资源进行分别隔离,以保障更高优先级入口。当callstack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隔离流控。

在新建隔离规则对话框中配置以下规则信息:

- 填写接口名称和来源应用。
- 统计维度选择链路入口。
- **并发数阈值**设置为10。

*接口名称	user_test	
* 来源应用	default	
统计维度 🚺	○当前接口 ○关联接口	● 链路入口
	用于应用内callstack调用流控情况。当ca QPS数超过阈值时,会对当前接口来自于	allstack 入口被来源应用调用 F来源应用的请求进行流控
* callstack入口	db	
* 并发数阈值	10	
是否开启	该规则打开, 创建后即生效	

当callstack入口的接口被调用超过10次,当前接口 user_test 会对来自于来源应用的请求进行隔离流控。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.1.3. 配置降级规则

SAE可以通过AHAS监控应用内部或者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。在指定的时间内,系统不调用该不稳定的资源,避免应用受到影响,从而保障应用高可用性。当指定时间过后,再重新恢复对该资源的调用。本文介绍配置降级规则的操作步骤。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见设置限流降级。

背景信息

除了流量控制以外,对调用链路中不稳定的方法或者下游依赖进行熔断也是重要措施之一。由于调用关系的 复杂性,如果调用链路中的某一环节出现了错误,会导致这个请求失败,甚至会放大不稳定性,导致整个链 路无法正常服务。熔断功能会在调用链路中某个方法出现不稳定时(例如某方法出现Timeout或异常比例升 高),对这个方法的调用进行限制,让请求快速失败,避免此错误影响整个链路。关于熔断的相关信息,请 参见熔断结构图。

熔断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。

操作步骤

1. 登录SAE控制台。

- 2. 在左侧导航栏单击**应用列表**,在**应用列表**页面上方选择**地域**,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 规则管理。
- 4. 在规则管理页面,单击降级规则,在降级规则页签,单击新增降级规则。
- 5. 在**新增降级规则**对话框,设置参数,单击**新建**。

新增降级规则 🕦 💦 🗙 🗙			
① 一般用于对弱依赖接口的降级调用,以便保证所在应用整体的可用性。查看详情			
* 接口名称	请输入接口名称		
* 统计窗口时长	20 秒 🔨	/	
阈值类型	● 慢调用比例 (%) ○ 异常比例 (%)		
* 慢调用RT	请求的响应时间超过该值则统计为慢调用 m	IS	
* 降级阈值 🚹	80 9	16	
* 熔断时长	- 10 + 秒		
	即为接口降级的时间。在该时间段内,该接口的请求都会快速失败。		
是否开启	该规则打开,创建后即生效		
* 最小请求数目 🕦	10		
熔断恢复策略 🚯	● 单次探测恢复 ○ 渐进式恢复		
	∧ 隐藏高级选项		
	新建取消		

参数说明如下:

参数	描述
接口名称	适用该规则的应用资源。
统计窗口时长	统计的时间窗口长度,取值范围为1秒~120分钟。
最小请求数目	触发熔断的最小请求数目,若当前统计窗口内的请求 数小于此值,即使达到熔断条件规则也不会触发。

参数	描述
	选择以慢调用比例或异常比例作为阈值。
阈值类型	3 选择改设调用比例作为阈值,需要设值允许的设调用RT(即最大的响应时间),请求的响应时间大于该值则统计为慢调用。 在降级阈值中设置触发熔断的慢调用比例。规则开启后,在单位统计时长内请求数目大于设置的最小请求数目,并且慢调用的比例大于阈值,则接下来的熔断时长内请求会自动被熔断。经过熔断时长后熔断器会进入探测恢复状态,若接下来的一个请求响应时间小于设置的慢调用RT则结束熔断,若大于设置的慢调用RT则会再次被熔断。
	 选择以异常比例作为阈值,需要在降级阈值中设置 触发熔断的异常比例。 规则开启后,在单位统计时长内业务异常数目大于 设置的最小请求数目,并且异常的比例大于阈值, 则接下来的熔断时长内请求会自动被熔断。
熔断时长	即熔断触发后持续的时间。资源进入熔断状态后,在 配置的熔断时长内,请求都会快速失败。
	熔断器进入恢复阶段(半开启状态)的恢复策略。
	 单次探测恢复:经过熔断时长后,熔断器会对接下 来的一个请求进行探测,若该请求符合预期(不为 慢调用或没有异常),则结束熔断;否则重新回到 熔断阶段。
	 渐进式恢复:需要设置恢复阶段数和每步最小通 过数目。
	 经过熔断时长后,熔断器按照设定的恢复阶段数 进行渐进式恢复,若该阶段内请求达到一定量即 每步最小通过数目,则触发检查。检查的请求若 都未超过阈值,则逐步提高允许通过的请求比 例,直到请求完全恢复;若某一步的指标超出阈 值,则重新回到熔断阶段。
浴断恢复束略	 请求比例T=100/恢复阶段数N,则第一阶段请求 比例为T,第二阶段为2T直到100%。
	例如恢复阶段数为3,每步最小通过数目为5,则 三个阶段分别按照33%、67%和100%的比例放 入请求,当每阶段的请求数目大于等于5时进行 检查,若请求的指标未超阈值则进入下一恢复阶 段,直至完全恢复。
	⑦ 说明 渐进式恢复策略功能需要Java SDK 版本≥1.6.2。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.1.4. 配置系统规则

系统规则从整体维度对应用入口流量进行控制,根据应用负载、CPU使用率、总体平均RT、入口QPS和并发 线程数等几个维度的监控指标,结合自适应的流控策略,让系统的入口流量和系统的负载达到一个平衡,保 证系统在最大吞吐量的状态下稳定运行。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见<mark>设置限流降级</mark>。

背景信息

系统支持开启自适应流控或手动设置系统规则:

- 自适应流控:当开启自适应流控,系统会根据CPU使用率动态调节应用的入口流量,在尽可能保证吞吐量的同时保证高负载下系统稳定。
- 系统规则:当关闭自适应流控,需要您手动设置系统规则。系统规则从整体维度对应用入口流量进行控制,结合应用的负载、CPU使用率、总体平均RT、入口QPS和并发线程数等几个维度的监控指标,让系统的入口流量和系统的负载达到一个平衡,保证系统在最大吞吐量状态下稳定运行。

系统保护规则是应用整体维度的,而不是资源维度的,并且仅对入口流量生效。入口流量指的是进入应用的 流量。例如Web服务或Dubbo服务端接收的请求。通过自定义埋点接入和通过注解接入的应用,入口流量 为 EntryType.IN 的逻辑被调用时产生的流量。配置系统规则的对象必须是入口流量,且默认接口为Dubbo 或HSF才生效。如默认埋点的Dubbo、Servelet等,或者通过修改 EntryType.IN 来让系统规则对该方法生 效, EntryType.IN 的详情请参见通过自定义埋点接入和接入Spring应用。

系统规则支持以下的模式:

- Load (仅对Linux、Unix-like机器生效): 当系统Load 1超过阈值且系统当前的并发线程数超过系统容量 时才会触发系统保护。
- CPU使用率: 当系统CPU使用率超过阈值(0.0~1.0)即触发系统保护。
- RT:当单台机器上所有入口流量的平均RT达到阈值即触发系统保护。
- 线程数:当单台机器上所有入口流量的并发线程数达到阈值即触发系统保护。
- 入口QPS:当单台机器上所有入口流量的QPS达到阈值即触发系统保护。

对于一个应用来说,每种相同的系统保护规则最多只能存在一条,即一个应用最多配置五条系统保护规则。 配置系统规则的原理请参见<mark>系统防护</mark>。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。

- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 规则管理。
- 4. 在规则管理页面,单击系统规则,在系统规则页签,单击新增系统保护规则。
- 5. 在新增系统保护规则对话框,设置参数,单击新建。

新增系统保	护规则 🕦	×
 根据单 看详情 	节点的系统负载和请求性能,控制入口流量以便保证最大吞吐量下的系统稳定性。查	
统计维度	● CPU 使用率 Load	
	用于根据CPU使用率衡量系统负载稳定性的场景。通过控制入口总流量,使其不 超过CPU利用率阈值。	
* CPU 阈值	CPU使用率,例如0.6	
是否开启	该规则打开,创建后即生效	
	新建取消	Ľ.

参数说明如下:

参数	描述	使用场景说明
CPU使用率	当系统CPU使用率超过阈值即触发 系统保护,阈值设置范围为 0.0~1.0(代表0%~100%)。	适用于设置基础资源水位的场景, 比如需要保证一定的冗余水位。但 系统水位不宜过高,需要留部分水 位。
Load	当系统的Load1超过阈值,且系统 当前的并发线程数超过系统容量时 才会触发系统保护。系统容量由系 统的maxQps * minRt计算得出。	适用于设置基础资源水位的场景, 比如需要保证一定的冗余水位。但 系统水位不宜过高,需要留部分水 位。
线程数	当单台机器上所有入口流量的并发 线程数达到阈值即触发系统保护。	适用于设置基础资源水位的场景, 比如需要保证一定的冗余水位。但 系统水位不宜过高,需要留部分水 位。
入口平均RT	当单台机器上所有入口流量的平均 RT达到阈值即触发系统保护,单位 是毫秒。	适用于衡量入口请求的场景。
入口总QPS	当单台机器上所有入口流量的QPS 达到阈值即触发系统保护。	适用于衡量入口请求的场景。

问题反馈
如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.1.5. 配置热点规则

为应用配置热点规则后, SAE将分析资源调用过程中调用次数高的参数, 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, 保障系统的稳定性。本文介绍配置热点规则的操作步骤。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见<mark>设置限流降级</mark>。

背景信息

热点即经常被访问的数据。例如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数据,并对其访问进行限制。

- 针对一段时间内最频繁购买的商品ID进行限制,防止击穿缓存而导致大量请求到数据库的情形。
- 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。

AHAS利用LRU(Least Recently Used)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。

使用说明

热点规则的使用说明和常见场景请参见配置热点规则。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 规则管理。
- 4. 在规则管理页面,单击热点规则,在热点规则页签,单击新增热点限流规则。
- 5. 在新增热点限流规则对话框,设置参数,单击新建。

新增热点限流规则		×
根据周期内数据	的访问次数或占用资源数,来限制多余请求,以保护系统稳定性。 <mark>查看详情</mark>	
* 接口名称	请输入需要进行热点限流的接口名,需与埋点名保持一致	
* 参数位置索引 🚯	想要进行热点限流统计的参数位置,从0开始	
统计维度	● 通过请求数 ○ 井发数	
	根据统计周期内调用次数来进行限制	
* 统计周期时间 👔	1	秒
* 单机阈值 🐧	请输入周期内对应参数被调用的统计阈值	
流控效果 🚹	● 快速失败 ○ 排队等待	
	多余请求立即失败,可以设置一个缓冲请求数量,多余缓冲区的请求数 立即失败。	则
* 缓冲请求数	0	
是否开启	该规则打开,创建后即生效	
	新建取	肖

参数说明如下:

参数	描述
接口名称	适用该规则的资源名称,与埋点传入的资源名保持一 致。
参数位置索引	埋点传入参数的索引位置。对应 SphU.entrv(xxx.arg s) 中的参数索引位置。例如 SphU.entrv(resourceN ame.Entry Type.IN,1,paramA.paramB) 埋点中, paramA 的参数索引是0, paramB 的参数索引是 1。
统计维度	可选择通过请求数或并发线程数。 • 通过请求数:限制一段时间内的调用次数。 • 并发数:限制该资源调用的并发数。

参数	描述
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10s,QPS阈值为5代表限制10s内每个热点参数访问 不超过5次。
单机阈值	是作用于每个热点参数的阈值。
流控效果	 当统计维度为通过请求数时,可以选择流控效果来处理被拦截的流量。 快速失败:达到阈值时,立即拦截请求。该模式下可以额外设置一个缓冲请求数,即针对突发请求额外允许的请求数目。 排队等待:请求匀速通过,允许排队等待,通常用于消息队列削峰填谷等场景。需设置具体的超时时间,排队时会计算预计的排队时长,若超过最大超时时间则请求会直接被拒绝。例如,单机阈值配置为5,则代表请求每200 ms才能通过一个,多出的请求将排队等待通过。超时时间配置1000 ms,则当前排队请求超过5个(>1000 ms)时,新到来的请求将会直接被拒绝。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.2. 设置限流降级

SAE支持使用AHAS实现Java应用的限流降级,全面保障您的应用的可用性。本文介绍如何在SAE设置限流降级。

前提条件

- 您已成功部署Java应用,详情请参见在SAE控制台使用WAR包部署Java Web应用。
- 您已开通AHAS,详情请参见开通AHAS。

⑦ 说明 使用AHAS时会产生单独费用, AHAS的计费说明请参见价格及欠费说明。

背景信息

限流: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。开启限流降级功能并设置限流策略后,如果流量在一定时间内超过了预设阈值,则会触发限流策略,进行限流处理。

 降级:降级通常用于对下游出现超时的非核心应用进行低优先级调用,确保上游核心应用不被影响的场景。开启限流降级功能并设置降级策略后,如果下游应用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。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用详情页面的右上角, 单击部署应用。

等,详情请参见以下文档。

4. 在部署应用页面,展开限流降级设置区域,打开启用限流降级功能开关,单击确认。

✓ 限流降级设置 设置应用是書	后自用限海岸级功能,全面提升业务稳定性。 🔗 如何设置限规律级	
雇用限选择级功能 开启后,SAE自动内到	■ ■AHAS現流降级能力,调至应用详细页的限流降级常单配置规则,计量方式请参考AHAS定价,请注意:如果应用已经通过SDK方式接入AHAS,开启此开关有一定冲突风险。	E B
		職认取消
↓ 注意	单击 确认 后,该应用将会被重启,请在业务较少的时间段进行。	

开启限流降级功能后,您可以在**限流降级(仅支持J**ava)菜单查看应用概览、机器监控、配置规则

功能	参考文档
应用概览	查看Java应用概览信息
机器监控	查看节点监控指标
配置流控规则	配置流控规则
配置隔离规则	配置隔离规则
配置降级规则	配置降级规则
配置系统规则	配置系统规则
配置热点规则	配置热点规则
查看操作日志	查看操作日志

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.3. 查看Java应用概览信息

开启限流降级后, SAE将通过AHAS监控各应用、接口、机器的实时数据, 从而评估系统的整体表现, 并为流 控降级规则提供重要依据。本文介绍查看Java应用限流降级概览信息的操作步骤。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见<mark>设置限流降级</mark>。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 应用概览。
- 4. 在应用概览页面,查看应用信息。
 - 限流指标详情:统计了近5分钟应用接口的通过QPS、流控QPS、异常QPS、响应时间、系统负载情况 以及防护事件。

⑦ 说明 应用概览页面涉及到的QPS、响应时间、并发线程均为应用入口接口的统计,不包括 应用内部方法调用的统计。

时间周期:5分钟 节点总数:2 应用概范中涉及到的QPS/F	RT/并发线程均为应用入口接口的统计,不包括应用内部方法调用的	961+.			
QPS 统计 →	响应时间 →	系統负載 →	防护事件 →		
1446.0k 146144.9k 0 👞	61 昨日网站 0.00%~		[簇点眼流]	2020-10-16 17:37:37	详情
通过QPS 流控QPS 异常QPS 环比	周期环比 0.00%~		[簇点限流]	2020-10-16 15:48:23	详情
50.0w 1	320 1	100 1	[簇点限流]	2020-10-16 15:48:23	详情
15.0w With the second strategy and the second	240 -	75	[簇点限流]	2020-10-16 15:48:23	详情
30.0w -	160 -	50 -	[簇曲限流]	2020-10-16 15:48:23	详情
15 Day .	80	25	[旗曲限流]	2020-10-16 15:48:23	详情
			[旗曲限流]	2020-10-16 15:48:23	详情
17:32 17:33 17:34 17:35 17:36	17:32 17:33 17:34 17:35 17:36	17:32 17:33 17:34 17:35 17:36			

- 单击QPS统计、响应时间、系统负载的→图标,可以查看各个指标更详细的历史数据。
- 单击防护事件的 → 图标,可以进入事件中心页面,查看防护事件的详细信息,包括级别、类型、 起始时间等。单击目标事件操作列的查看详情,可以查看该接口的历史监控数据。
- QPS热力图:包括通过QPS热力图、流控QPS热力图、异常QPS热力图等。

通过QPS	流控QPS	异常QPS	平均RT	CPU	Load			
TOP节点						五衢全部	图表	表格
平均值(avg): 3	11.5 标准整(std): 0.71						

- 鼠标悬浮在热力图的图标中,可以查看各指标TOP节点的设备ID等相关信息;单击热力图中的图标,可查看各指标的历史数据。
- 单击右侧的图表或表格,可以选择以不同的形式展现TOP节点的信息。
- 单击右侧的查看全部,进入机器监控页面,查看各节点的QPS、CPU、Load等详细信息。
- TOP接口列表会动态刷新,按通过QPS由大到小排列接口。

TOP接口						查看全部
接口名称	通过QPS	限流QPS	异常QPS	平均RT(ms)	操作	
quick-service	1097	2914	0	0	流控 隔离 降级	
/doSomething	608	0	0	19	流控 隔离 降级	
handleServiceE	50	0	0	9	流控 隔离 降级	
handleServiceD	50	0	0	16.48	流控 隔离 降级	

- 单击目标资源操作列的流控、隔离或降级,可为该资源配置相应规则,详情请参见以下文档:
 - 配置流控规则
 - 配置降级规则
 - 配置隔离规则
- 单击接口名称进入接口详情页面,可以查看该接口的详细信息。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



5.1.4. 查看节点监控指标

SAE控制台记录了所有节点的通过QPS、限流QPS、异常QPS、RT、并发等指标。一个节点对应一个JVM进程,当多个JVM接入应用后,即展示为多个节点。您可以通过节点指标查看应用中所有资源的数据分布及对比情况,并管理接口的各项规则。本文介绍查看节点监控指标的操作步骤。

前提条件

请确保您的Java应用已开启限流降级功能。具体步骤,请参见<mark>设置限流降级</mark>。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java) > 机器监控。
- 4. 您可以在机器监控页面,执行以下操作:

ahas-token-client-demo 机器监	空 ◎	1 回放时间 2021-01-08 16:49:31 箇
调输入节点名称	节点概览 JVM 监控	节点对比
节点名称 通过QPS / 限流QPS / 异常QPS / RT ↓	应用描标 CPU LOAD 物理内存 Disk 网络流量 网络数据包	кл КУ
全部节点	展示描标 通过QPS 扫绘QPS 展带QPS RT(ms) 并发	
0/ 0/ 0/ 0		
2	800	
	600 400 200 0 0 0 0 0 0 0 0 0 0 0 0	
	3	共有1条 〈 上一页 1 下一页 〉

• (图标①)在页面右上角选择回放时间,查看该时间前5分钟内的历史数据。

? 说明 最多保留7天的历史数据。

- (图标②)在**节点名称**区域,罗列了全部节点和对应的通过QPS、限流QPS、异常QPS、RT等信息。 单击节点名称可以查看对应的各数据时序图。
- (图标③)在时序图区域,可以执行以下操作:
 - 在**节点概览**页签,单击QPS、CPU、LOAD等页签,可以分别查看全部节点相关指标的时序图,还 可以选择要展示或隐藏的指标。
 - 单击 <u>同</u> 图标, 可以查看该接口指标的历史数据。

? 说明 最多保留7天的历史数据。

- 单击节点名称后, 会在右侧**节点概览**页展示该节点对应的各数据时序图。
 - 单击JVM监控页签, 查看该接口的GC次数、GC耗时等时序图。
 - 单击接口详情页签,筛选查看不同接口的数据。
 - 单击callstack信息页签,查看所有接口的信息,并可以设置该接口的限流规则、查看历史数据。
 - **平铺展示**:不区分调用链路关系,平铺展示接口的运行情况。
 - 树状展示:根据接口的调用链路关系,展示树状结构。
 - 单击目标接口操作列中的流控、隔离或降级,可以快速管理限流规则。更多信息,请参见配置流控规则、配置隔离规则和配置熔断规则。
 - 单击目标接口操作列中的更多 > 查看监控, 可查看该接口指标的历史数据和数据对比情况。
 - 单击**异常统计**页签,查看该应用近5分钟的异常情况。
 - 单击**异常统计数据**柱形图中的柱形,则在右侧**异常TOP**列表中,展示该异常的具体接口、类型、次数。
 - 单击异常统计数据右上角的历史数据,可以查看任意时间段发生的异常次数,最多可以查看 7天内的历史数据。
 - 您在**异常统计数据**区域查看到异常数据后,可在应用本地日志中查看异常记录。

⑦ 说明 控制台查看以秒为单位的异常统计情况,异常详细内容您可以在本地异常记录 文件中查看。

- 异常记录的文件地址:与metrics.log同个目录。
- 异常记录的文件名称:
 - {appName}-ahas_exceptions.log.{日期,精确到天}
 - {appName}-ahas_exceptions.log.{日期,精确到天}.idx
- 异常记录格式: {时间戳}|{时间,精确到毫秒}|{异常接口}|{异常详情}

样例如下所示:

1619749703000|2021-04-30 10:28:23.062|/exception1|com.alibaba.csp.sentinel.demo.exceptio n.TestException: test exception_1 15

1619749703000|2021-04-30 10:28:23.065|/exception|com.alibaba.csp.sentinel.demo.exception .AnotherException: test another exception 71

1619749705000|2021-04-30 10:28:25.065|/exception1|com.alibaba.csp.sentinel.demo.exceptio n.TestException: test exception_1 38

- 单击目标资源操作列的流控、隔离或降级,可为该资源配置相应规则,详情请参见以下文档:
 - 配置流控规则
 - 配置降级规则
 - 配置隔离规则

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.1.5. 查看操作日志

SAE记录了所有应用及资源相关的操作详情,包括操作时间、操作内容、操作人ID等信息。您可以通过这些信息追踪应用和资源的变更。本文介绍如何查看操作详情。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏中选择限流降级(仅支持Java)>操作日志。
- 4. 在操作日志页面,查看操作详情。

foo 操作日志					
请输入资源名称进行简选	Q				
操作	类别	资源名	摄作内容	摄作时间	摄作人ID
删除	流控规则	test	制除规则	2020-09-22 20:00:57	
新増	流控规则	test	状态开启,描标项缓程数调置:1.0,来源应用:default,流控 模式,直接,流控方式,快速失败	2020-09-22 13:58:12	1010/014040
添加例外项	热点规则	XXXX	英型:Int;例外项参数值:123.调值:12.0	2019-12-16 18:19:13	0.0.000000
新增	热点规则	X000X	状态:开启,指标项:QPS,单机调值:22.0,参数跟引:1,统计圈口 时长:1,流控效果快速失败,burst count:0	2019-12-16 18:18:03	12.02
新增	流控规则	test	状态:开启,指标项QPS,测道:12.0,来源应用:default,流控構 式,直接,流控方式快速失败	2019-11-27 21:18:55	10000
					共有臻 〈 上一页 1 下一页 〉

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



5.2. 微服务灰度流量

5.2.1. 灰度发布微服务

灰度发布是服务治理领域中最重要的策略之一。当需要上线新的功能时,您可以使用灰度发布功能在小范围 内进行新版本测试,待验证通过后再进行全量升级。本文以Spring Cloud微服务应用为例,介绍在SAE灰度 发布微服务的操作步骤。

前提条件

您已在SAE创建并部署Spring Cloud或Dubbo微服务应用,其中包含1个Consumer应用和1个Provider应用。 具体操作,请参见以下文档:

- 使用Spring Cloud开发微服务应用并部署至SAE
- 使用Spring Boot开发Dubbo应用并部署至SAE

Consumer应用和Provider应用的示例工程如下:

- Spring Cloud
- Dubbo

背景信息

Consumer应用调用Provider应用时, Provider应用有版本V1和版本V2。通过创建灰度规则,指定Header的 Version参数为V1的流量都灰度到Provider应用的V1版本。灰度发布微服务流程,如下图所示:



步骤一: 创建灰度规则

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,单击新建灰度规则。
- 右新建灰度规则面板,执行以下操作,然后单击确定。
 SAE支持的灰度规则是按内容灰度,您所需配置的参数会因应用框架而异。

i. 按需配置以下参数。

参数	说明
规则名称	灰度规则的名称。支持以小写字母开头,以数字或小写字母结尾,仅可包含小写字 母、中文字符、数字及短划线(-)。不超过64个字符。
规则描述	灰度规则的描述信息。不超过64个字符。
框架类型	配置灰度规则应用的框架类型,支持以下框架: Spring Cloud:需要设置Path。 Dubbo:需要选择服务方法。
条件模式	选择同时满足下列条件或满足下列任一条件。

ii. 单击添加新的规则条件,按需配置灰度规则条件。

参数	说明
条件列表	 Spring Cloud:根据Cookie、Header或Parameter参数类型,设置相应的参数、条件以及值。 Dubbo:根据应用实际情况,设置参数、参数值获取表达式、条件以及值。

iii. (可选)单击新建流量规则创建多个入口流量规则,多个规则可以同时生效。Spring Cloud应用灰度规则示例

则名称	provider-	test-rule				
	支持以小写 超过64个字符	字母开头,以数字或/ 穿。	小写字母结尾,仅可	「包含小写字母、	中文字符、数字及	短划线(-)。不
则描述						
						0/64
度类型:	按内容灰度	0				
灰度规则1						
* 框架类型	Sprir	ng Cloud 🛛 🔿 D	ubbo			
Path						
Path HTTP相对路	径, 例如/a/b,注	意严格匹配, 留空代	表任何路径			切换为选择路径
Path HTTP相对路	经,例如/a/b,注i	意严格匹配, 留空代	表任何路径			切换为选择路径
Path HTTP相对路行 *条件模式	圣, 例如/a/b,注	意严格匹配,留空代	表任何路径			切换为选择路径
Path HTTP相对路4 * 条件模式 同时满足下 	译,例如/a/b注) 列条件 〇 ;	意严格匹配,留空代 满足下列任一条件	表任何路径			切换为选择路径
Path HTTP相对路台 * 条件模式 • 同时满足下 * 条件列表	径,例如/a/b,注) 列条件 〇 ;	意严格匹配, 留空代 满足 下列任一条件	表任何路径			切换为选择路径
Path HTTP相对路 * 条件模式 • 同时满足下: * 条件列表 参数类型	径,例如/a/b.注 列条件 ○ 氵	愈严格匹配,留空代 满足下列任一条件 参数	表任何路径	ŧ	值	切换为选择路径
Path HTTPHE对路: * 条件模式 ● 同时满足下: * 条件列表 参数类型 Header	径,例如/a/b.注 列条件 ○ : ✓	意严格匹配,留空代 満足下列任一条件 参数 headergrey	表任何路径	ŧ = ~	值 xx	切換为选择路径

在本示例中,参数类型选择Header。

创建成功后,您将在**灰度规则**页面查看到刚创建的灰度规则。同时,新建灰度规则按钮已置灰。

步骤二:查看规则详情

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,找到您需查看的灰度规则,单击规则名称。
- 5. 在灰度规则详情面板,查看规则内容。

← 灰度规则详	桂 月			
观则名称: 灰度类型:	provider-test-rule 按内容灰度 🕜	规则描述:		
规则1				
框架类型 Path	springcloud	条件模式	同时满足下列	川条件
条件列表	参数类型	参数	条件	值
	header	headergrey	==	XX

步骤三: 配置灰度发布策略

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在应用详情页面的右上角, 单击部署应用。
- 4.
- 5. 在发布策略设置区域内配置灰度发布参数。

灰度发布参数说明如下。

- 6. 单击确认完成发布设置。
- 7. 您可以选择以下任一方式验证发布结果。
 - 方法一:在应用的**变更记录**页面中查看应用变更详情,查看发布状态,如果所有批次都执行成功,则
 说明应用更新成功。
 - 方法二:在应用基本信息页面的实例部署信息页签查看实例的运行状态。如果执行状态显示为Running,且实例的版本已变更,表示应用部署成功。

ン 加度公组 - 「状本・法行1个定例	
✓ 显示实例IP信息 注意:实例ip非固定ip,会在部署应用前后、底层主机宕机迁移时发生变化。请不要强依赖实例ip信息。	
实例名称 vSwitch 实例ip 版本 运行状态 \h	
□f4vk6 可用区H	康检查
批量重启实例 当前您总计选中了0个实例	
> 默认分组 状态: 运行1个实例	
✓ 显示实例IP信息 注意:实例ip非固定ip,会在部署应用前后、底层主机宕机迁移时发生变化。请不要强依赖实例ip信息。	
□ 实例名称 vSwitch 实例ip 版本 运行状态↓ ト	

步骤四:通过Consumer访问验证

通过请求Consumer应用访问Provider应用,验证流量灰度发布到了指定实例。具体操作,请参见使用 Webshell诊断应用。

sh-4.2# curl --location --request GET 'http://127.0. /echo-rest/rest-res t1?pa=10' \--header 'headergrey: xx' \ > --header 'Cookie: Cookie_2=value; cookie2=xx' provider: .2.79: Tue Aug 10 09:47:04 CST 2021:79588a50-f6f4-4146-b521-68a49 5bc :(rest-rest1,10,xx,Cookie_2=value; cookie2=xx,xx)sh-4.2# curl --location t1?pa=10' \--header 'headergrey: xx' --header 'Cookie: Cookie_2=value; cookie2=x x' provider: .2.79: Tue Aug 10 09:47:07 CST 2021:79588a50-f6f4-4146-b521-68a49 5bc :(rest-rest1,10,xx,Cookie_2=value; cookie2=xx,xx)sh-4.2# curl --location t1?pa=10' \--header 'headergrey: xx' --header 'Cookie: Cookie_2=value; cookie2=x x' provider: .2.79: Tue Aug 10 09:47:08 CST 2021:79588a50-f6f4-4146-b521-68a49 5bc :(rest-rest1,10,xx,Cookie_2=value; cookie2=xx,xx)sh-4.2# curl --location t1?pa=10' \--header 'headergrey: xx' --header 'Cookie: Cookie_2=value; cookie2=x x' provider: .2.79: Tue Aug 10 09:47:08 CST 2021:79588a50-f6f4-4146-b521-68a49 5bc :(rest-rest1,10,xx,Cookie_2=value; cookie2=xx,xx)sh-4.2# curl --location t1?pa=10' \--header 'headergrey: xx' --header 'Cookie: Cookie_2=value; cookie2=x x' provider: .2.79: Tue Aug 10 09:47:10 CST 2021:79588a50-f6f4-4146-b521-68a49 5bc :(rest-rest1,10,xx,Cookie_2=value; cookie2=xx,xx)sh-4.2#

5.3. 优雅下线

5.3.1. 设置微服务无损下线

如果您期望托管至SAE的微服务应用处理完请求后再停止应用,您可以使用SAE无损下线功能。SAE支持 Spring Cloud和Dubbo应用无损下线。本文介绍如何在SAE控制台上配置无损下线。

应用部署时设置微服务无损下线

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击创建应用。
- 3. 在应用基本信息页签设置应用相关信息,并单击下一步: 应用部署配置。
- 4. 在应用部署配置页签,选择技术栈语言和应用部署方式,设置部署参数。
- 5. 在应用部署配置页面,展开微服务无损下线配置区域,单击微服务无损下线配置向导。



6. 在微服务无损下线配置向导面板,设置相关参数,单击预览配置。

微服务无	损下线配置向导		
	 ————————————————————————————————————	2	
	无损下线参数 ^融 署	预览配置	
配置完无损下线	8054 发的参数,这些参数将会在应用创建/部署的环境变量处	4、生命周期管理-优雅下线超时设置处自动渲染。 配置参数说明可参考文档	
* 是否开启优新	性下线服务治理能力(profiler.micro.service.http.serv	ver.enable)	
true	\checkmark		
SAE优雅下线J	的能占用端囗(profiler.micro.service.http.server.por	rt)	
5000	ms		
* 是否开启自动	加等待功能(micro.service.shutdown.auto.wait)		
true	\sim		
		预览配置取消	

参数说明如下。

变量名称	参数说明
是否开启优雅下线服务治理能力	默认为true。取值说明如下: o true:开启。 o false:关闭。
SAE优雅下线功能占用端口	默认端口为54199。如果此默认端口和您应用的端口 产生冲突,请配置新的端口。
收到Spring的ContextClosedEvent事件后,进程动态 等待关闭的时间	默认取值为5000,即进程收到kill信号后,会等待 5000 ms再关闭。 配置该参数后,需要设置是否开启自动等待功能参数 为false。
是否开启自动等待功能	默认为true。取值说明如下: • true:开启。进程可能会随着流量停止而提前关闭。 • false:关闭。进程直接等待收到Spring的 ContextClosedEvent事件后,进程动态等待关闭的时间所配置的值的对应时长后再关闭。

7. 在**应用生命周期管理设置**区域,确认相关配置,单击确认。

✔ 应用生命周期管理设置	2 生命周期脚本定义,	管理应用容器在运行	前和关闭前的—些	动作,如环境准备、(优雅下线等 🔗	如何设置应用	生命
周期 启动后处	理 (PostStart设置)	停止前处理	(PreStop设置)	优雅下线超时	讨论置 (Termi	< >	
应用容器被终 命令行 脚 本	止前的任务,用于优雅;	关闭应用程序、通知	其他系统等				
/bin/bash							
-c curl http://	/127.0.0.1:54188/offlir	ش ش					
+ 添加		Ш					
					上一步	确认	取消

8. 单击下一步:确认规格。

9. 在确认规格页签,查看您所创建应用的详细信息以及配置费用情况,并单击确认创建。

应用部署完成后设置微服务无损下线

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 单击具体应用名称。
- 3. 在应用详情页面的右上角, 单击部署应用。
- 4. 在部署应用页面,展开微服务无损下线配置区域,单击微服务无损下线配置向导。

∨ 微服务无损下线配置	设置Spring Cloud/Dubbo应	Z用优雅下线,在应用处理完请求后再下线实	例。 🔗 如何设置微服务无损下线
启用微服务 , 微服务无损	后损下线功能: (仅 5.线配置向导	राजेSpring Cloud/Dubbo JAVA应用生效)	

5. 在微服务无损下线配置向导面板,设置相关参数,单击预览配置。

微服务无损下线配置	向导		
	 — 		2
	无损下线参数	Ţ	领览配置
配置完无损下线的参数, 这些参	₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩₩	生命周期管理-优雅下线超时设置处自动渲	染。 配置参数说明可参考文档
* 是否开启优雅下线服务治理能	力(profiler.micro.service.http.serve	r.enable)	
true 🗸			
SAE优雅下线功能占用端口(pi	rofiler.micro.service.http.server.port)		
收到Spring的ContextClosedEv	vent事件后,进程动态等待关闭的时间	(micro.service.shutdown.wait.time)	
5000 ms			
* 是否开启自动等待功能(mico	ro.service.shutdown.auto.wait)		
true 🗸			
			预览配置取消

参数说明如下。

变量名称	参数说明
是否开启优雅下线服务治理能力	默认为true。取值说明如下: o true:开启。 o false:关闭。
SAE优雅下线功能占用端口	默认端口为54199。如果此默认端口和您应用的端口 产生冲突,请配置新的端口。
收到Spring的ContextClosedEvent事件后,进程动态 等待关闭的时间	默认取值为5000,即进程收到kill信号后,会等待 5000 ms再关闭。 配置该参数后,需要设置是否开启自动等待功能参数 为false。
是否开启自动等待功能	默认为true。取值说明如下: • true:开启。进程可能会随着流量停止而提前关闭。 • false:关闭。进程直接等待收到Spring的 ContextClosedEvent事件后,进程动态等待关闭的 时间所配置的值的对应时长后再关闭。

6. 在**应用生命周期管理设置**区域,确认相关配置,单击**确认**。

✔ 应用生	命周期管理设置 生命	周期脚本定义,管理	里应用容器在运行	前和关闭前的一	-些动作,如环境	准备、优雅下线	₩ 🔗 ¢	叩何设置应用	生命
周期	启动后处理 (Pos	stStart设置)	停止前处理	(PreStop设置) 优雅下	〔 线超时设置 (Termi <	$\langle \rangle$	
	应用容器被终止前的任	务,用于优雅关闭	应用程序、通知其	其他系统等					
	命令行脚本								
	/bin/bash								
	-c	Ū							
	curl http://127.0.0.1	:54188/offlir							
	+ 添加								
						Ŀ	步	确认	取消

7. 配置完成后单击确认。

↓ 注意 单击确认后,该应用将会被重启,请在业务较少的时间段进行。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



6.升级和回滚应用

6.1. 管理灰度规则(Java)

对于部署在SAE的Spring Cloud或Dubbo微服务应用,为了确保升级操作的安全性,您可以通过启用灰度发 布(即金丝雀发布)的灰度规则进行小规模验证,验证通过后再全量升级。本文介绍如何在SAE控制台上创 建、查看、编辑和删除灰度规则。

前提条件

您已在SAE创建并部署Spring Cloud或Dubbo微服务应用。具体操作,请参见以下文档:

- 使用Spring Cloud开发微服务应用并部署至SAE
- 使用Spring Boot开发Dubbo应用并部署至SAE

背景信息

如果您的应用有多个应用实例且需要升级,那么您可以按需通过灰度或分批发布方式升级应用实例版本。其中,针对Spring Cloud和Dubbo微服务的灰度发布需要启动灰度规则才能实现。

更多信息,请参见灰度发布应用。

创建灰度规则

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,单击新建灰度规则。
- 5. 在新建灰度规则面板,执行以下操作,然后单击确定。

SAE支持的灰度规则是按内容灰度,您所需配置的参数会因应用框架而异。

i. 按需配置以下参数。

参数	说明
规则名称	灰度规则的名称。支持以小写字母开头,以数字或小写字母结尾,仅可包含小写字母、中文字符、数字及短划线(-)。不超过64个字符。
规则描述	灰度规则的描述信息。不超过64个字符。
框架类型	配置灰度规则应用的框架类型,支持以下框架: Spring Cloud:需要设置Path。 Dubbo:需要选择服务方法。
条件模式	选择同时满足下列条件或满足下列任一条件。

ii. 单击添加新的规则条件,按需配置灰度规则条件。

参数	说明
条件列表	 Spring Cloud:根据Cookie、Header或Parameter参数类型,设置相应的参数、条件以及值。 Dubbo:根据应用实际情况,设置参数、参数值获取表达式、条件以及值。

iii. (可选)单击新建流量规则创建多个入口流量规则,多个规则可以同时生效。

Spring Cloud应用灰度规则示例

E 12 11 - C 1 47/1	rule-test							
273 H13.	支持以小写学 过64个字符。	字母开头,以数3	字或小写字母结屑	1、仅可包含小5	言字母、 中文	(字符、数字及)	豆划线(-)。不超	
则描述	test							
							4/64	
度类型:	按内容灰度	0						
灰度规则1								
* 框架类型	Sprin	ng Cloud	Dubbo					
Path								
/path							1774/2 21 2/5477 00 (7	
* 冬산道士							切换为远绊路径	
 条件模式 同时满足下 条件列表 	列条件 〇 ;	满足下列任一条 _{条数}	(‡	冬(仕		信	切换为选择确任	福作
 条件模式 同时满足下 条件列表 参数类型 	列条件 〇;	满足下列任一条 参数	件	条件		值	切换为逻样确全	操作
 条件模式 同时满足下 条件列表 参数类型 Cookie 	列条件 〇 ;	满足下列任一条 参数	件 5/64	条件 =	~	值 test		操作
 条件模式 同时满足下 条件列表 参数类型 Cookie +添加新的规则 	列条件 〇 ; ~ [—] 川条件	满足下列任一条 参数	件 5/64	条件 =	~	值 test		操作
 条件模式 同时满足下 条件列表 参数类型 Cookie +添加新的规则 + 新建流量规则 	列条件 〇 ; ~ 川条件	满足下列任一条 参数 *****	/‡ 5/64	条件 =	~	值 test		操作
 条件模式 同时满足下 条件列表 参数类型 Cookie + 添加新的规则 + 新建流量规则 	列条件 ○; ~ U条件	满足下列任一条 参数 	/# 5/64	条件 =	~	值 test		操作

创建成功后,您将在**灰度规则**页面查看到刚创建的灰度规则。同时,新建灰度规则按钮已置灰。

查看规则详情

1. 登录SAE控制台。

- 2. 在左侧导航栏单击**应用列表**,在**应用列表**页面上方选择**地域**,单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,找到您需查看的灰度规则,单击规则名称。
- 5. 在灰度规则详情面板,查看规则内容。

← 灰度规则详	情			×
规则名称: 灰度类型:	rule-test 按内容灰度 🗿	规则描述:	test	
规则1				
框架类型	springcloud	条件模式	同时满足	2下列条件
Path	/path			
条件列表	参数类型	参数	条件	值
	param	.name	==	test

编辑灰度规则

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表, 在应用列表页面上方选择地域, 单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,找到您需编辑的灰度规则,在其操作列,单击编辑。
- 5. 在编辑灰度规则面板,按需更新规则内容。
- 6. 单击确定。

删除灰度规则

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在左侧导航栏,选择微服务治理 > 灰度规则。
- 4. 在灰度规则页面,找到您需删除的灰度规则,在其操作列,单击删除。
- 5. 在弹出的提示对话框,单击确认。

6.2. 灰度发布应用

如果您的应用有多个应用实例且需要升级,那么您可以按需通过灰度或分批发布方式升级应用实例版本。本 文介绍灰度发布方式的背景信息、场景示例和操作步骤;此外,还涉及应用回滚的说明。

背景信息

- 灰度发布:又名金丝雀发布,在原有部署版本可用的情况下,同时部署新版本应用作为金丝雀,测试新版本的性能,在保证整体系统稳定的情况下,帮助您尽早发现问题和修复问题。
 应用进行灰度发布时,为了保证应用稳定性,灰度发布的应用实例数不能超过应用实例总数的50%,剩余的应用实例按照指定的批次分批发布。
- 分批发布: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。分批发布过程中如果出现故障,则可以及时终止并回退,待问题修复后重新发布。

应用进行分批发布时,应用内的实例数会平均分配到每个批次进行部署。如果无法平均分配,则靠前的批次分配到的实例数少,靠后的批次分配到的实例数多。更多信息,请参见分批发布应用。

场景示例

某应用包含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,灰度发布流程如下图所示。



操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击**应用列表**,在**应用列表**页面上方选择**地域**,单击具体应用名称。
- 3. 在应用详情页面的右上角,单击部署应用。
- 4. 配置部署参数。

⑦ 说明 部署方式由应用首次部署方式决定,请根据所需的部署方式设置参数。

• WAR包部署:重新上传WAR包或者输入新部署WAR包的地址,并完成相关环境和参数设置。

配置JAR包。			
应用名称	sae-demo-consumer		
* Java环境:	openjdk-8u191-jdk-alpine3.9	\sim	
* 文件上传方式:	上传WAR包	\sim	
*WAR包地址:	spring-boot-		
	下親JAR包样例		
* 版本:	1.0 3/16 使用时间颠为版本号		
* 时区设置:	UTC+8	\sim	
✓ 高级选项			

○ JAR包部署:重新上传JAR包或者输入新部署JAR包的地址,并完成相关环境和参数设置。

配置JAR包 *		
应用名称	sae-demo-consumer	
* Java环境:	openjdk-8u191.jdk-alpine3.9	\sim
* 文件上传方式:	上(#JAR8)	\sim
* JAR包地址:	spring-boot	×
	下载JAR包祥例	
* 版本:	1.0 3/16 使用时间能为版本号	
* 时区设置:	UTC+8	\sim
∨ 高级选项		

○ 镜像: 在部署应用区域单击修改镜像, 在修改镜像面板重新选择镜像仓库或镜像版本。

← 修改镜像					×
配置镜像 *					
我的阿里云镜像	Demo镜像	公有镜像	其它阿里云账号私有镜	意像	
镜像服务个人版	🦳 镜像服务企业)	扳			
如果您应用所在的reg	gion和阿里云镜像仓库	不在同一个regi	ion,需要跨region pull镜像	,请在ACR产品进行配置,	查看详情。
镜像仓库命名空间:	sae_test1 🗸	镜像仓库名:	请输入镜像仓库名:	Q	
	sae_test1,		类型:PRIVATE	来源:ALI_HUB	请选择
	sae_test1/	sister -	类型:PRIVATE	来源:ALI_HUB	请选择
确定 取消					

5. 在发布策略设置区域内配置灰度发布。

权仲策阻	次度及中 、	×			发布策略配置信息	
					The second second second second	
友度数量	1	(当前应用实例总数:	: 2)		→ 开始部署	
o de la compañía de	· 为了保证应用稳定性、友度实例数2	不能招讨应用实例总数的1	1/2			
	75 5 UKEEELIMAKEEELI MABEKUSAK		/-			第1批: 11
灰度后剩余批次	1				○ 部署结束	
	灰度发布后,剩余的应用实例将按照	曾指定的批次发布完成				
最小存活实例数	1 4	☑ 使用系统推荐(έ.			
	建议最小友活动例数设置2-1 尽可	计能保证业务不由断 加度	电沿弯为0 应用在升	场过程中令中断业务		
灰度规则配置详情						
秋沢川	and and and	A7 (4.+##P				
Path	/path	39614-136255		6/E 1793361+		
条件列表	参数类型	参数	条件	值		
	param	.name		test		

灰度发布参数说明如下。

配置	是否必选	说明
发布策略	是	选择 灰度发布 。
灰度数量	是	设置首先需要进行灰度发布的应用 实例数量。
灰度后剩余批次	是	灰度发布后 <i>,</i> 剩余的应用实例按照 设定的批次完成发布。
最小存活实例数	是	每次滚动升级最小存活的实例数。 选择 使用系统推荐值 后SAE将根据 您的需求为应用设置最佳的最小存 活实例数。
启用微服务灰度规则	否	您为Spring Cloud或Dubbo应用创 建的灰度规则。具体操作,请参 见 <mark>管理灰度规则(Java)</mark> 。

6. (可选)高级配置。

高级设置	是否必选	参考文档
		设置PHP应用监控
设置PHP应用监控	否	⑦ 说明 仅技术栈语 言为PHP时可以设置。

高级设置	是否必选	参考文档
设置PHP应用配置文件	否	设置PHP应用配置文件 ⑦ 说明 仅技术栈语 言为PHP时可以设置。
启动命令	否	设置启动命令
环境变量	否	设置环境变量
Hosts绑定	否	设置Hosts绑定
应用健康检查	否	设置健康检查
应用生命周期管理	否	设置应用生命周期管理
日志收集	否	设置日志收集
持久化存储	否	设置NAS存储
配置管理	否	注入配置信息

7. 单击确认完成发布设置。

- 8. 您可以选择以下任一方式验证发布结果。
 - 方法一:在应用的变更记录页面中查看应用变更详情,查看发布状态,如果所有批次都执行成功,则
 说明应用更新成功。
 - 方法二:在应用基本信息页面的实例部署信息页签查看实例的运行状态。如果执行状态显示为Running,且实例的版本已变为Ver.2,表示应用部署成功。

应用回滚

采用灰度发布或者分批发布方式升级应用实例时,如果应用实例中存在未完成升级的实例,则当前应用升级 状态处于**进行中**。

在实时跟踪应用升级时,如果首批应用实例升级突发异常停止响应,为了保证业务不受影响,请在**变更详** 情页面中单击**立即回滚**,将已升级的实例回退至升级前版本,并将配置还原为升级前原有配置。

在应用变更过程中,如果出现部署包不可用、健康检查失败等变更流程异常情况,将导致应用升级失败,SAE将停止当前应用并进行回退。

在SAE上进行应用升级耗时约为30分钟,超出该时间后SAE将上报超时异常并暂停变更流程,请在**变更详** 情页面手动终止发布流程并回滚应用。

更多信息

- 您在SAE部署完应用后,可以对应用进行更新、扩缩容、启停、删除应用等生命周期管理操作。具体操作,请参见管理应用生命周期。
- 您在SAE部署完应用后,可以对应用进行自动弹性伸缩、SLB绑定和批量启停等提升应用性能的操作。具体操作,请参见以下文档:

o 绑定SLB

o 配置弹性伸缩策略

- 一键启停应用
- 配置管理
- 变更实例规格
- 您在SAE部署完应用后,还可以对应用进行日志管理、监控管理、应用事件查看和变更记录查看等聚焦应 用运行状态的操作。具体操作,请参见以下文档:
 - 日志管理
 - o 监控管理
 - 应用事件查看
 - 变更记录查看
 - 使用Webshell诊断应用

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



7.分布式配置管理

7.1. 配置管理概述

SAE集成了ACM(Application Configuration Management,应用配置管理)中的配置管理功能,您可以在SAE中使用ACM对应用配置进行集中管理。



您可以将应用开发过程中产生的大量的参数和变量等信息,提取到配置文件中统一管理和运维。代码运行时,可以从该配置文件中获取相应的变量和参数,以便快速更改变量和参数,降低代码维护成本。本文介绍如何在SAE中创建、导入或者导出配置文件。

背景信息

本文创建的配置文件适用于应用开发及运维场景,如果您需要创建容器运行环境相关的配置项,请参见<mark>管理</mark> 和使用配置项(K8s)。

创建配置

- 1. 登录SAE控制台。
- 2. 在左侧导航栏选择配置管理 > 配置列表。
- 3. 在顶部菜单栏,选择地域。
- 4. 在配置列表页面,选择命名空间并单击创建配置。
- 5. 在创建配置面板, 输入配置信息并单击创建。

← 创建配置	×
所属地域	
华东2(上海)(cn-shanghai)	
命名空间	
默认(31696bd1-1623-4fb1-b3fb-5ed314bd)	
Data ID * 🔞	
支持大小写字母、数字、下划线(_)、短横线(-)、点号(.)和半角冒号(:),长度不超过236个字符。	0/236
Group * 🔞	
支持大小写字母、数字、下划线(_)、短横线(-)、点号(.)和半角冒号(:),长度不超过128个字符。	0/128
数据加密 🖉	
配置格式 📀	
● TEXT ○ JSON ○ XML ○ YAML ○ HTML ○ Properties	
1	×.
記号描述の	
	0/128
◇ 更多配置	
应用 ❷	
支持大小写字母、数字、下划线(_)、短横线(-)、点号(.)和半角冒号(:),长度不超过128个字符。	0/128
标签	
请输入标签	0/64
创建取消	

参数说明如下:

⑦ 说明 配置的数据模型为<Namespace+Data ID+Group,Content>。Data ID和Group无需申请,须在同一命名空间内保持唯一。不同命名空间之间的配置相互隔离。

参数	说明
Data ID	配置ID。采用类似 package.class (如 com.taoba o.tc.refund.log.level)的命名规则保证全局唯一 性。建议根据配置的业务含义定义 <i>class</i> 部分。仅允许 使用英文小写字符和以下4种特殊字符:英文句号 (.)、冒号(:)、短划线(-)和下划线(_)。不超 过236个字符。
Group	配置分组(命名空间)。建议填写 产品名:模块 名 (如 ACM:Test)保证唯一性。后续可以根据 Group进行鉴权。仅允许使用英文字符和以下4种特殊 字符:英文句号(.)、冒号(:)、短划线(-)和下 划线(_)。不超过128个字符。
	配置数据是否加密。如果您的配置中包含敏感数据, 建议您使用加密存储功能,降低配置泄漏风险。
数据加密	↓ 注意 使用前必须开通密钥管理服务,并授 权ACM使用密钥管理服务进行加解密。因为ACM 数据加密功能依赖密钥管理服务,为其配置加 密。加密配置的Data ID均以cipher-开头,具体操 作,请参见创建和使用加密配置。
配置格式	选择配置格式。SAE会根据您选择的格式进行数据校 验。默认选择TEXT,支 持TEXT、JSON、XML、YAMLHTML和Properties 格式。
配置内容	输入配置的内容。建议不超过10 KB,最大不超过100 KB。
配置描述	配置描述信息。便于理解配置含义,不超过128个字 符。
更多配置	
应用	配置归属应用名。仅允许使用英文小写字符和以下4种 特殊字符:英文句号(.)、冒号(:)、短划线(-) 和下划线(_)。不超过128个字符。
标签	配置标签。方便您根据自己的维度管理配置,最多支 持5个标签,每个标签不超过64个字符。

导入或者导出配置

SAE支持导入或者导出跨命名空间的配置文件,方便配置复用。

1. 登录SAE控制台。

- 2.
- 3. 在左侧导航栏选择配置管理 > 配置列表。
- 4. 在左侧导航栏选择配置管理 > 配置列表, 在配置列表页面上方选择目标配置所在的命名空间。
- 5. 在目标配置所在的命名空间下,选择需要导出的配置,单击导出,并在导出配置对话框中单击导出。

导出配置		×
源命名空间:		
默认(31696bd1-1623-4fb1-b3fb-5ed314b))		
配置数:		
1个条目已选中		
	导出	取消

- 6. 解压已导出的配置文件。
- 7. 在**配置列表**页面上方选择需要导入配置的命名空间,单击**导入配置**,并在**导入配置**对话框中单击上传 文件。

导入配置		×
目标命名空间: 默认(31696bd1-1623-4fb1-b3fb-5ed314b)		
相同配置:		
终止导入		\sim
配置文件		
	导入	取消

8. 在打开的对话框中选择解压的配置文件,单击打开,然后在导入配置对话框中单击导入。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



7.3. 查看推送状态

修改配置后,您需要查看修改后的配置信息是否已推送到监听该配置的机器上。此查询仅对使用了监听配置 接口监听配置的客户端有效。本文介绍如何查看配置信息的推送状态。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏选择配置管理 > 监听查询。
- 3. 在顶部菜单栏,选择地域。
- 4. 在监听查询页面,设置相关参数。
 - i. 选择命名空间。
 - ii. 在查询条件输入区域,在下拉列表选择查询维度:配置或查询维度: IP,根据维度输入待查询配置的参数,并单击Q 图标。

您可以在**监听查询**页面,查看配置项的推送状态。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。



7.4. 查询推送轨迹

修改完配置后,您可以根据推送轨迹查看配置是否推送完成。如果配置修改完发现某台机器不生效,那么可 以根据推送轨迹排查是否推到。本文介绍如何查询推送轨迹。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏选择配置管理 > 推送轨迹。
- 3. 在顶部菜单栏,选择地域。
- 4. 在推送轨迹页面,设置相关参数。
 - i. 选择命名空间。
 - ii. 在查询条件输入区域,在下拉列表选择查询维度:配置或查询维度:IP,根据维度输入待查询配置的参数,并单击Q图标。
 - iii. (可选)在右侧日历面板选择具体时间。

⑦ 说明 如果不选择具体时间,系统默认查询当日。

您可以在**推送轨迹**页面,查看配置项的推送时间与推送轨迹。

问题反馈

如果您在使用SAE的过程中有任何疑问和建议, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉 钉群与我们交流。



7.5. 查询和回滚版本

使用配置管理功能完成配置修改后,如果出现配置更改错误,须查看历史记录、立即回滚。本文介绍如何查 询和回滚配置文件版本。

操作步骤

- 1. 登录SAE控制台。
- 2. 在左侧导航栏选择配置管理 > 历史版本。
- 3. 在顶部菜单栏,选择地域。
- 4. 在历史版本页面,设置相关参数。
 - i. 选择命名空间。
 - ii. 在**请选择Group**和Data ID下拉列表选择相应的参数,在**请输入Data ID**文本框输入完整的Data

ID, 并单击 Q 图标。

- 5. 查询历史版本详情或者回滚版本。
 - 查看: 单击操作列中的查看, 查阅该配置的历史版本详情。
 - 回滚: 单击操作列中的回滚, 并在回滚版本页面单击确定。

历史版本									
默认 test demo-test	el Ja								
() 系统只保留30天的变更记录。									
Group: DEFAULT_GROUP	✓ Data ID	✓ cipher-test	٩						G
Data ID	Group		更新时间		所属应用	操作类型	操作		
cipher-test	DEFAULT_GROUP		2021年8月3日 16:31:6	06	N/A	更新	查看	回滾	
cipher-test	DEFAULT_GROUP		2021年8月3日 16:30:	17	N/A	更新	查看		
cipher-test	DEFAULT_GROUP		2021年8月3日 16:28:	22	N/A	插入	查看		

? 说明 配置管理功能目前仅保存30天的变更记录。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



7.6. 创建和使用加密配置

为了确保敏感配置(数据源、Token、用户名和密码等)的安全性,降低用户配置的泄露风险,SAE提供了 创建加密配置的功能。

前提条件

- 开通ACM服务
- 开通密钥管理服务

背景信息

↓ 注意 使用加解密服务必须使用阿里云账号或RAM用户的AccessKey,不支持使用ACM专用 AccessKey。

SAE提供的配置加密方法有KMS加密和KMS AES-128加密。

- KMS加密:配置内容禁止超过6 KB。
- KMS AES-128加密:配置内容允许超过6 KB,建议不超过10 KB,最大不超过100 KB。配置内容的明文数 据不会传输到KMS系统,安全性更高。

创建加密配置

按照以下步骤在控制台创建加密配置。

- 1. 登录SAE控制台。
- 2. 在顶部菜单栏,选择地域。
- 3. 在配置列表页面,选择命名空间并单击创建配置。

4. 在创建配置页面,设置相关参数。

关于配置参数的更多信息,请参见创建或导入配置。

5. 打开数据加密开关,选择KMS加密或KMS AES-128加密方式,填写完成后单击创建。

数据加密 🕜	
● KMS 加密	○ KMS AES-128 加密

参数	描述
KMS加密	直接调用KMS服务对配置进行加解密,加解密数据的大 小不超过4 KB,最大不超过6 KB,对特殊符号如 and(&)会解密错误,不推荐使用。
KMS AES-128加密	使用KMS的信封加解密方法,配置内容可以超过6 KB, 最大不超过100 KB。配置内容的明文数据不会传输到 KMS系统,安全性更高,推荐使用。

↓ 注意

- 首次使用此功能时,须开通密钥管理服务,并授权ACM使用您的密钥管理服务进行加解密,因为ACM数据加密功能依赖密钥管理服务,并为配置加密。
- 为方便您管理配置,控制台上均显示明文,实际均为加密存储。

使用加密配置

按照以下步骤在控制台获取加密所需的AccessKey ID和AccessKey Secret等初始化参数。

- 1. 登录SAE控制台。
- 2. 在顶部菜单栏,选择地域。
- 3. 在配置列表页面,选择命名空间并查看相关参数。
 - i. 单击查看获取方法,即可获取AccessKey ID和AccessKey Secret等初始化参数。

Endpoint: addr-sh-internal.edas.aliyun.com	命名空间: 31696bd1-1623-4fb1-b3fb-5ed314bd
AccessKey/SecretKey: 推荐用于测试环境, 查看获取方法	通过ECS示例RAM角色访问ACM:推荐用于生产环境,查看获取方法

⑦ 说明 使用阿里云账号的AccessKey ID和AccessKey Secret可直接获取配置数据,建议您填写RAM用户的AccessKey ID和AccessKey Secret。如果使用RAM用户的AccessKey ID和AccessKey Secret,则必须提前为其授予AliyunACMFullAccess和AliyunKMSCryptoAccess权限。更多信息,请参见RAM用户授权。

ii. 查看示例代码。

a. 选择目标配置,单击配置的Data ID。

b. 在配置详情页面,单击示例代码页签,选择目标语言,即可获取示例代码。

⑦ 说明 Java、Python的SDK已经整合KMS-SDK,您可以添加解密过滤器来自动解密。其他语言解密,请参见Decrypt。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。


8.手动和自动扩缩

8.1. 手动扩缩

应用扩缩是通过改变应用的实例数来增加或减少应用的计算容量。您可以在应用的实例负载过高时以手动方 式添加新应用实例,在应用闲置时减少应用实例,能够高效利用应用资源、降低成本。本文介绍如何在SAE 通过手动方式实现应用实例数的增加或减少。

前提条件

- 部署应用
- (可选) 配置弹性伸缩策略

背景信息

配置弹性伸缩策略包括手动扩缩和自动扩缩两种方式:

- 当应用扩缩容为紧急需求时,例如突发性流量高峰,您可以选择手动扩缩方式。
- 当应用扩缩容为为非紧急需求时,例如周期性的流量高峰,您可以选择自动扩缩方式。具体操作,请参见配置弹性伸缩策略。

操作步骤

⑦ 说明 扩容时,如果扩容的实例数与原应用的实例数总和,大于应用的资源使用限制,则扩容失败。如果需要继续增加应用实例,请提交工单。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在基本信息页面右上角单击手动扩缩。
- 4. 在应用扩缩对话框内设置相关参数。
 - 未部署或未开启弹性伸缩策略时:
 滑动指针设置应用目标实例数,并单击确定。

应用扩缩							×
* 应用目标实例数:	0 5 (应用当前实)	10 15 20 列数3个)	25 30	35 40	45 50	- 5 + 1	
						确定	取消

部署并开启弹性伸缩策略时:

a. 滑动指针设置应用目标实例数。

- b. 选择手动扩缩后恢复自动弹性方式,并单击确定。参数说明如下:
 - 系统自动恢复:定时弹性将在当前周期内的下一个时间段窗口生效。如果不满足条件,将递延到下一个循环周期生效。监控指标弹性策略和混合弹性策略在手动扩缩后立即自动生效,您需要检查弹性最小最大值的设置,避免对业务SLA造成影响。

应用扩缩		×
* 应用目标实例数:	■ 5 10 15 20 25 30 35 40 45 50 - 5 + 个 (应用当前实例数2个)	
手动扩缩后恢复自动弹性方式:	● 系统自动恢复 ○ 人工启用恢复	
送择系统自动恢复弹性后 监控指标弹性/混合弹性在	定时增性特在当前周期内的下一个时间段窗口立即生效,如不满足条件,将递延到下一个循环周期生效。 手动扩缩后立即目动生效,请一定注意检查到世最小最大值的设置,避免造成业务SLA影响。查看详情	
	確定 取消	ij

■ **人工启用恢复**:手动恢复自动弹性。

应用扩缩	>	<
* 应用目标实例数:	0 5 10 15 20 25 30 35 40 45 50 - 10 + ↑ (应用当前实例数2个)	
手动扩缩后恢复自动弹性方式:	○ 系统自动恢复 ● 人工启用恢复	
	确定 取得	
 ⑦ 说明 ■ 手动扩始的 	1日左京例数2恢复收按长行递供的具大店	

- 手动扩缩的目标实例数≤恢复监控指标弹性的最大值。
- 您必须先部署并开启弹性伸缩策略,才能够选择**手动扩缩后恢复自动弹性方式**。
- 5. 结果验证。

进入**变更详情**页面,查看变更记录。

- 方法一: 在基本信息页面上方单击查看详情, 在变更详情页面查看变更的具体执行状态。
- 方法二:在左侧导航栏单击变更记录,在变更记录页面查看变更状态。如果变更状态显示执行成功,则表示成功扩容或缩容应用。单击操作列的查看,可以在变更详情页面查看变更的具体执行状态。

问题反馈

如果您在使用SAE过程中有任何疑问, 欢迎您扫描下面的二维码或搜索钉钉群号23198618, 加入钉钉群与我 们交流。



8.2. 配置弹性伸缩策略

在分布式应用管理中,弹性伸缩是很重要的运维能力,它能够根据实例状态自动增加或减少实例数量,即扩容或缩容,从而提高资源利用率、降低资源成本。本文介绍配置SAE定时弹性策略、监控指标弹性策略和混合弹性策略的操作及验证步骤。

前提条件

部署应用

背景信息

配置弹性伸缩策略包括手动扩缩和自动扩缩两种方式:

- 当应用扩缩容为紧急需求时,例如突发性的流量高峰,您可以选择手动扩缩方式。具体操作,请参见手动 扩缩。
- 当应用扩缩容为为非紧急需求时,例如周期性的流量高峰,您可以选择自动扩缩方式,即配置弹性伸缩策略。SAE弹性伸缩策略包括定时弹性伸缩策略、监控指标弹性伸缩策略和混合弹性伸缩策略。关于配置弹性伸缩策略的最佳实践,请参见SAE弹性伸缩最佳实践。

适用场景

- 配置定时弹性伸缩策略:适用于资源使用率有周期性规律的应用场景,多用于证券、医疗、政府和教育等行业。
- 配置监控指标弹性伸缩策略:适用于突发流量和典型周期性流量的应用场景,多用于互联网、游戏和社交 平台等行业。
- 配置混合弹性伸缩策略:适用于兼备资源使用率有周期性规律和有突发流量、典型周期性流量的应用场景,多用于互联网、教育和餐饮等行业。

使用须知

- 您最多创建5条定时策略、1条监控指标策略或1条混合弹性策略,三种策略不能同时使用。
- 弹性策略启用时,请勿手动进行应用生命周期管理操作,例如应用扩缩、部署应用、更改规格、重启应用 和停止应用。如果您需要对应用进行该类操作,先停用弹性策略后,再手动执行操作。
- 如果当前应用处于扩容、缩容、部署(单批、分批或灰度)、更改规格、重启和停止等过程中,该应用暂时无法添加或启动弹性策略。
- 如果单个应用需要弹出超过50个实例,需提交工单申请白名单。

配置定时弹性伸缩策略

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击**应用列表**, 在**应用列表**页面上方选择**地域**, 单击具体应用名称。

在基本信息页面单击实例部署信息,在实例部署信息页签展开弹性伸缩折叠区域,单击+添加弹性策略。

基本信息	实例部署信息
∨ 弾性伸缩	
+ 添加弹性策	A 弹性策略使用须知 ?

4. 在添加弹性策略面板,设置策略参数。

添加弹性	策略				
弹性策	1 () () () () () () () () () () () () ()				2 预览定时策略
策略类型	定时策略 监控指标策略				
* 策略名称	口请输入策略名称				
*选择时间	请输入以小写字母开头,小写 短期 长期	字母、数字、中划线组成的1到	32字符的字符串。		
	周期 每天 >				
	从 08:00 ④	之后保持实例数	10	^ 🛅	
	从 20:00 ①	之后保持实例数	3	↑ Î	
	+添加				
	1. 定时策略的触发时间不能 2. 两个邻近触发时间间隔要	小于系统的当前时间。 大于5分钟。			
. 新建	的弹性策略默认不启用,需要在	弹性策略列表中手动启用。			
取消	下一步: 预党定时策略				

参数说明如下:

参数	说明
策略类型	选择定时策略。

参数	说明
策略名称	自定义的策略名称,必须以小写字母开头,仅可包含 小写字母、数字及短划线(-),不超过32个字符。
选择时间	 短期:适用于需要指定执行定时弹性伸缩策略的起止日期时的场景。 长期:适用于不需要指定执行定时弹性伸缩策略的结束日期时的场景。
周期	执行定时弹性伸缩策略的周期。 • 每天:每天指定时间执行定时策略。 • 每周:每周指定天数的指定时间执行定时策略。 • 每月:每月指定日期的指定时间执行定时策略。
单天内的触发时间	设置弹性伸缩策略触发的时间,以及该时间段内需要 保持的应用实例数。示例如下: 如果周期设为每天,第一条定时策略触发开始时间为 08:00,目标实例数为10个;第二条定时策略开始时间 为20:00,目标实例数为3个。则在08:00至20:00之 间,SAE依据所设的规则,将该应用的实例数保持为10 个,在20:00至次日08:00之间,应用实例数保持为3 个。

< ♪ 注意

- 单条定时策略内设置的触发时间,不能小于系统当前时间;如果小于当前时间,该策略将在 下个执行周期内生效。
- 单条定时策略内的两个邻近时间间隔须大于5分钟。如果您设置了多条定时策略, SAE会检查重复日期内的多个触发时间点的时间间隔。
- 如果设置了多条定时策略,并且两个触发点时间重复,那么遵循以新盖旧的原则,以最后添加的触发点策略为准。设置定时策略时请您避免时间冲突的情况。
- 5. 单击下一步:预览定时策略确认策略信息,单击确认。
- 6. 策略设置完成后,在**实例部署信息**页签,单击目标策略操作列的启用。
- 在启用弹性策略对话框中单击确认, 启用该策略。
 当该策略的状态切换至已启用时, 说明策略启用成功。

定时策略列表					
策略名称	状态	触发时间和目标实例数	创建时间	最后一次停用策略时间	操作
1070	 已启用 	预览定时策略	2021-06-09 11:12:59	2021-06-17 10:40:57	启用 停用 编辑 :

配置监控指标弹性伸缩策略

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 在基本信息页面单击实例部署信息,在实例部署信息页签展开弹性伸缩折叠区域,单击+添加弹性策略。

基本信息	实例部署信息			
∨ 弾性伸縮				
+ 添加弹性策略	B 弹性策略使用须	页知 ?		

4. 在**添加弹性策略**对话框中单击监控指标策略页签,在监控指标策略页签设置策略参数,单击确认。

11111111111111111111111111111111111111	策略				
策略类型	定时策略 监控	指标策略 混合弹性策略	Beta 3		
* 策略名称	请输入策略名称	ζ.			
触发条件	必须以小写字母开的 CPU使用率。	头,仅可包含小写字母、教	数字及短划线(-),不超过3 5 & RT 近30分钟内无数据,5	2个字符。 6法设置触发条件。详情	
	CPU使用率目标值		%		
	或 内存使用率目标	值	%		
	SAE会自动伸缩实例	谢数, 无限接近您设置的监	a 容指标目标值。		
	当应用监控指标的	实际值小于目标值时, SAI	E会自动缩容实例。反之, SA	E会自动扩容实例。 <mark>详情</mark>	
实例数设置	● 应用最小实例数	č 2 个		● 应用最大实例数:	50 个
实例数设置	 应用最小实例数 1 	13	25	● 应用最大实例数: 37	50 1
实例数设置	 应用最小实例数 1 	2 ↑ 13	25	 ● 应用最大实例数: 37 	50 个 50
实例数设置	 应用最小实例数 1 的弹性策略默认不启标 	2 个 13 +	25 十 中手动启用。	● 应用最大实例数: 37 ↓	50 个 50
文例数设置 ① 新建	 应用最小实例数 1 的弹性策略默认不启用 	2 个 13 用,需要在弹性策略列表中	25 十 中手动启用。	● 应用最大实例数: 37 ↓	50 个 50
实例数设置	 应用最小实例数 1 的弹性策略默认不启的 	2 个 13 用,需要在弹性策略列表。	25 十 中手动启用。	● 应用最大实例数: 37	50 个 50
实例数设置	 应用最小实例数 1 的增性策略默认不启病 	2 个 13 月,需要在弹性策略列表中	25 中手动启用。	● 应用最大实例数: 37	50 ↑ 50
○ 新建	 应用最小实例数 1 的弹性策略默认不启的 	2 个 13 日,需要在弹性策略列表中	25 十 中手动启用。	● 应用最大实例数: 37 ■	50 个 50
文例数设置	● 应用最小实例数 1 ● 的弹性策略默认不启序	2 个 13 月,需要在弹性策略列表。	25 十 中手动启用。	● 应用最大实例数: 37 ↓	50 1
文例数设置 ③ 新建	● 应用最小实例数 1 ● 的弹性策略默认不启	2 ↑ 13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	25 十 中手动启用。	● 应用最大实例数: 37 ■	50

参数说明如下:

参数	说明
策略类型	选择 监控指标策略 。
策略名称	自定义的策略名称,必须以小写字母开头,仅可包含 小写字母、数字及短划线(-),不超过32个字符。

参数	说明
	设置以下监控指标中的一个或多个。 • CPU使用率 • 内存使用率
触发条件	 说明 当应用监控页面最近30分钟内有应用 QPS或响应时间(RT)数据时才可以设置 该项指标。 当设置多个监控指标时,只要任何一个指 标大于或者等于目标值,则应用进行扩 容,扩容后实例数不高于应用最大实例 数;如果所有指标均小于目标值,则应用 进行缩容,缩容后实例数不低于应用最 小实例数。
实例数设置	 您可以直接设置输入实例数或者滑动下方的指针设置 实例数。 应用最小实例数:触发弹性伸缩条件后,应用缩 容,其实例数可达到的目标值。 应用最大实例数:触发弹性伸缩条件后,应用扩 容,其实例数可达到的目标值。 ① 说明 手动扩缩与添加弹性策略能够设置 的最大实例数同步。

5. 策略设置完成后,在实例部署信息页签,单击目标策略操作列的启用。

在启用弹性策略对话框中单击确认,启用该策略。
 当该策略的状态切换至已启用时,说明策略启用成功。

监控指标策略列表									G
策略名称	状态	指标目标值	指标当前值	触发时间 扩缩触发条件 🕜	应用最大实例数	应用最小实例数	创建时间	最后—次停用策略时间	操作
1000	● 已启用	CPU: 50% MEMORY: 59%			10个	2个	2021-06-09 11:1 1:49	2021-06-16 14:12:15	启用 停用 編輯 :

配置混合弹性伸缩策略

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 在基本信息页面单击实例部署信息,在实例部署信息页签展开弹性伸缩折叠区域,单击+添加弹性策略。

基本信息	实例部署信息
∨ 弾性伸缩	
+ 添加弹性策略	弹性策略使用须知 ?

4. 在添加弹性策略对话框中单击混合弹性策略页签,在混合弹性策略页签设置策略参数,单击下一步:预览定时策略。

添加弹	性策略										
策略类型	定时策略 监控指标	Beta 策略 混合弹性策略									
* 策略名称	请输入策略名称										
	必须以小写字母开头, (又可包含小写字母、数字及短划线(-),不超	过32个字符。								
触发条件	CPU使用率、内存	使用率文 该应用 QPS & RT 近30分钟内无数据,	,无法设置触发条件。详情								
	CPU使用率目标值 %										
	或内存使用率目标值 %										
	SAE会自动伸缩实例数,无限接近您设置的监控指标目标值。										
	当应用监控指标的实际	直小于目标值时, SAE会自动缩容实例。反之,	SAE会自动扩容实例。详情								
* 应用最大	最小实例数设置:										
默认i	设置										
点 ●	如用最小实例数: 2	\uparrow	 应用最大实例数; 	50 个							
结种	时间段沿署										
选择	时间										
0 \$	短期 💿 长期										
周期	期 每天 🗸										
单	天内的触发时间										
	特殊时间段	应用最小实例数 应用	用最大实例数	操作							
	没有数据										
	+添加										
1. ī 2. 1	1. 两条策略的时间段不能重合。 2. 最多添加20条策略。										
! 新	建的弹性策略默认不启用, ;	需要在弹性策略列表中手动启用。									
		-									
取消	下一步:预览定时策略										

参数说明如下:

参数	说明
策略类型	选择 混合弹性策略 。
策略名称	自定义的策略名称,必须以小写字母开头,仅可包含 小写字母、数字及短划线(-),不超过32个字符。
触发条件	 设置以下监控指标中的一个或多个。 CPU使用率 内存使用率 了 说明 当应用监控页面最近30分钟内有应用 QPS或响应时间(RT)数据时才可以设置 该项指标。 当设置多个监控指标时,只要任何一个指 标大于或者等于目标值,则应用进行扩 容,扩容后实例数不高于应用最大实例 数;如果所有指标均小于目标值,则应用 进行缩容,缩容后实例数不低于应用最 小实例数。
应用最大最小实例数设置	 默认设置应用实例数。 默认设置 应用最小实例数:触发弹性伸缩条件后,应用缩容,其实例数可达到的目标值。 应用最大实例数:触发弹性伸缩条件后,应用扩容,其实例数可达到的目标值。 最小存活实例数:每次部署最小存活的实例数。 单击设置,在设置最小存活实例数对话框中设置最小存活实例数对话框中设置最小存活实例数量。 特殊时间设置 短期:适用于需要指定执行定时弹性伸缩策略的起止日期时的场景。 长期:适用于不需要指定执行定时弹性伸缩策略的结束日期时的场景。 周期:在周期列表中,选择每天、每周或每月,并选择相应参数。 单天内的触发时间:单击+添加,选择特殊时间段,并输入应用最小实例数和应用最大实例数。 ① 说明 手动扩缩与添加弹性策略能够设置的最大实例数同步。每次滚动部署最小存活的实例数建议大于等于1,保证业务不中断。如果设置为0,应用在升级过程中将会中断业务。

参数	说明
特殊时间段设置	 短期:适用于需要指定执行定时弹性伸缩策略的起止日期时的场景。 长期:适用于不需要指定执行定时弹性伸缩策略的结束日期时的场景。
周期	 每天:每天指定时间执行定时策略。 每周:每周指定天数的指定时间执行定时策略。 每月:每月指定日期的指定时间执行定时策略。
单天内的触发时间	设置弹性伸缩策略触发的时间,以及该时间段内需要 保持的应用实例数。示例如下: 如果周期设为每天,第一条定时策略触发开始时间为 08:00,目标实例数为10个;第二条定时策略开始时间 为20:00,目标实例数为3个。则在08:00至20:00之 间,SAE依据所设的规则,将该应用的实例数保持为10 个,在20:00至次日08:00之间,应用实例数保持为3 个。

? 说明

- 。 混合弹性策略设置的多个特殊时间段之间不可重叠。
- 单条混合弹性策略内设置的触发时间,不能小于系统当前时间;如果小于当前时间,该策略 将在下个执行周期内生效。
- 5. 单击下一步:预览定时策略确认策略信息,单击确认。
- 6. 策略设置完成后,在**实例部署信息**页签,单击目标策略操作列的启用。
- 在启用弹性策略对话框中单击确认,启用该策略。
 当该策略的状态切换至已启用时,说明策略启用成功。

混合弹性策略列表								C
策略名称	状态	指标目标值	指标当前值	触发时间 扩缩触发条件 🕜	当前应用最大实例数	当前应用最小实例数	6 操作	
dische	● 已启用	CPU: 12% MEMORY: 12% QPS: 21/砂 RT: 21ms	CPU: 1% MEMORY: 17% QPS: 5/秒 RT: 994.15ms	触发时间:2021-06-16 17:26:39之后 扩容: 满足扩容: 満足cpu > 13% 或 mem ory > 13% 或 qps > 22/砂 或 rt > 21.001 ms 缩容: 満足缩容: 満足cpu < 10% 且 mem ory < 10% 且 qps < 18/砂 且 rt < 18.375 ms	50个	2个	2 2 2	停用 编辑

验证弹性伸缩策略

启用弹性策略后, SAE会自动监控应用状态, 当满足触发条件时, SAE依据弹性策略进行应用实例扩容或缩容。您可以通过以下步骤验证弹性伸缩策略是否生效。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 在基本信息页面的应用信息区域查看应用运行实例数。 如果运行实例数与弹性策略中设置的实例数一致,说明弹性策略已生效。

✓ 应用信息							
应用名:	命名空间: cn-shanghai						
应用id: 1c26c0dc-055a-4627-8388.	VPC: vpc-						
vSwitch:vswitch	安全组: alikafka 切换安全组						
实例规格: 1Core, 2GiB, 系統盘磁盘空间20GiB 变更规格	运行实例数:当前2个实例/目标2个实例						
应用创建时间:2021年5月7日 13:23:53	晓像地址 registry-vpc.cn-shanghai.aliyuncs.com/						
应用标签: 编辑标签	应用描述∠ 编辑						

- 4. (可选)您可以通过以下方式查看弹性伸缩原因。
 - 查看应用事件: 在左侧导航栏单击应用事件,在左侧事件原因中选择自动弹性(Horizont alPodAutoscaler), 查看弹性伸缩原因。
 - 查看自动弹性事件:
 - a. 在基本信息页面单击实例部署信息,在实例部署信息页签展开弹性伸缩折叠区域。

b. 单击目标策略操作列的 : 图标并在快捷菜单中单击事件, 查看自动弹性原因。

应用事	#						
0 4	3计 4 项记录,当前仅保留7天内	的应用事件。					
自認	弹性 (HorizontalPod) ~	全部事件等级 >	来源名 >	-	٩		C
	事件等级	事件原因	来源类型	来源名		首次发生时间	最近发生时间
+	Warning	FailedComputeMetricsReplicas	HorizontalPodAutoscaler	and the local		2020-09-27 18:05:57	2020-09-27 18:05:57
+	Warning	FailedGetResourceMetric	HorizontalPodAutoscaler	10111000-1111-000		2020-09-27 18:05:57	2020-09-27 18:05:57
+	Normal	SuccessfulRescale	HorizontalPodAutoscaler	100.0034-001.004		2020-09-27 17:38:15	2020-09-27 17:38:15
+	Normal	SuccessfulRescale	HorizontalPodAutoscaler			2020-09-27 17:35:15	2020-09-27 17:35:15
						毎页显示: 10 ン 共	4条 く 上一页 1 下一页 >

查看和管理已配置的弹性策略

为应用配置弹性伸缩策略后,您可以在实例部署信息页签上查看已配置策略的详细信息。

- 1. 登录SAE控制台。
- 2. 在左侧导航栏单击应用列表,在应用列表页面上方选择地域,单击具体应用名称。
- 3. 在基本信息页面单击实例部署信息页签,展开弹性伸缩折叠区域,在指定策略的操作列中,您可以执 行以下操作。

基本信息 实	列部署信息											
∨ 弾性伸縮												
+ 添加弹性策略	弹性策略使用须	20										
定时策略列表												
策略名称	状态		触发时间和目标实例数		创建时间		最后一次停用策略时间]	操作			
10.00	• 28)	ŧ	预觉定时策略		2021-06-09 11:12:59		2021-06-17 10:40:57		启用 停用 编辑	:		
监控指标策略列表											C	
策略名称	状态	指标目标值	指标当前值	触发时间 扩缩触发条	件 😧	应用最大实例数	应用最小实例数	创建时间	最后一次停用策略时间	操作		
	 已停用 	CPU: 50% MEMORY: 59%				10个	2个	2021-06-09 11:1 1:49	2021-06-17 10:43:27	启用 停	用「编辑」	:
混合弹性策略列表											G	
策略名称	状态	指标目标值	指标当前值	触发时间 扩缩触发条件	0	当前应用最大实例数	当前应用最小实例	數 创建时间	最后一次停用策略	操作		
44470	 ● 已停用 	CPU: 12% MEMORY: 22% QPS: 12/秒 RT: 33ms				-		2021-06-10 6:52	5 14:4	启用丨《	明 编辑	

○ 启用: 单击启用来启动策略。

- **停止**: 单击**停止**来停止策略。
- 编辑: 单击编辑来编辑策略。
- 删除: 单击 ⋮ 图标并在快捷菜单中单击删除来删除策略
- 查看事件: 单击 ⋮ 图标并在快捷菜单中单击事件来查看自动弹性事件。

问题反馈

如果您在使用SAE过程中有任何疑问,欢迎您扫描下面的二维码或搜索钉钉群号23198618,加入钉钉群与我 们交流。

