

ALIBABA CLOUD

# 阿里云

## 应用高可用服务 产品简介

文档版本：20200828

 阿里云

## 法律声明

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

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

# 通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 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。	 注意 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，不是用户必须了解的内容。	 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 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[ ] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

# 目录

- 1.什么是应用高可用服务AHAS ----- 05
- 2.名词解释 ----- 10
- 3.产品版本说明 ----- 12
- 4.探针版本说明 ----- 15
  - 4.1. 应用高可用探针版本说明 ----- 15
  - 4.2. 应用高可用Java探针版本说明 ----- 16

# 1.什么是应用高可用服务AHAS

应用高可用服务（Application High Availability Service）是一款专注于提高应用高可用能力的SaaS产品，包含架构感知、流量防护、故障演练和功能开关四大独立的功能模块。

- **架构感知：**
  - 自动感知应用的拓扑结构。
  - 以可视化的方式直观呈现应用对基础架构的依赖关系和组件间的依赖关系。
  - 持续记录上述依赖关系。
- **流量防护：**
  - 专业化多样化的限流手段。
  - 实时秒级的监控。
  - 立即生效的规则管理。
- **故障演练：**
  - 提供基于真实线上故障的高可用能力演练服务。
  - 根据您的应用架构智能推荐故障演练场景。
- **功能开关：**
  - 按照统一的方式定义、管理业务运行时功能切换。
  - 支持定义任意类型的配置项，且自动根据类型做自动拆装，代码层无需关注类型信息，简化运行时配置项使用流程。

## 架构感知

- **功能**

服务器、存储、网络是现代云平台的基础设施。随着上云战略的推进，越来越多的大型企业将业务、服务、系统构建在云平台上。开源软件和云服务的多样性，开发语言的异构性，以及企业IT团队的组织能力差异，都提高了标准化的复杂性。

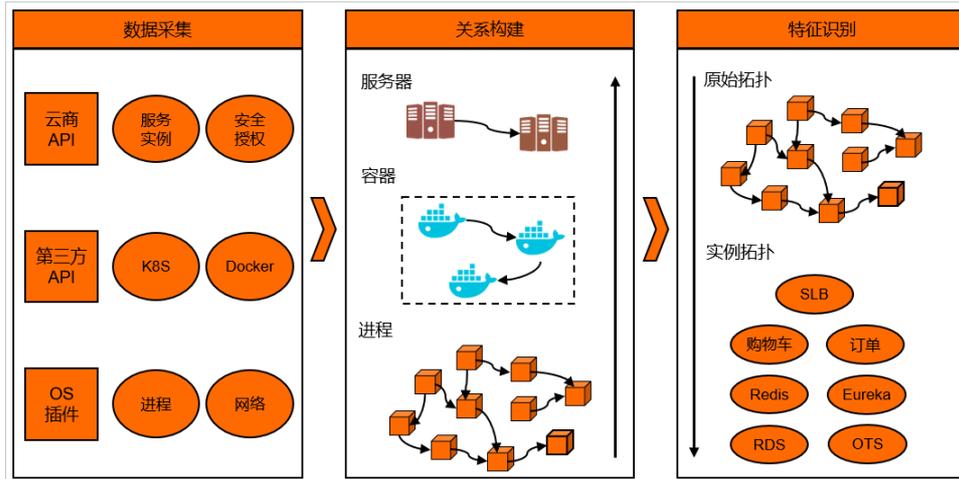
在此背景下，架构感知功能应运而生。它会采集和分析操作系统及第三方标准接口，捕捉进程级的调用关系，并使用特征库算法识别进程所使用的技术组件，最后在服务器、容器和进程这三个维度上以可视化的方式展示应用架构。

- **AHAS支持的数据源**

	服务器	容器		进程		
	服务器	第三方容器服务	云容器服务	客户业务服务	第三方服务	云 SaaS
特征数据	Cloud OpenAPI	Docker OpenAPI	Docker OpenAPI	探针	探针	Cloud OpenAPI
	探针	K8s OpenAPI	K8s OpenAPI			
			Cloud OpenAPI			
依赖数据	计算	计算	Cloud OpenAPI	探针	探针	Cloud OpenAPI
			探针			
运行时数据	Cloud OpenAPI	Docker OpenAPI	Docker OpenAPI	探针	探针	Cloud OpenAPI
	探针	K8s OpenAPI	K8s OpenAPI			
	云商 API	开源 API	AHAS 能力	无		

- **工作流程**

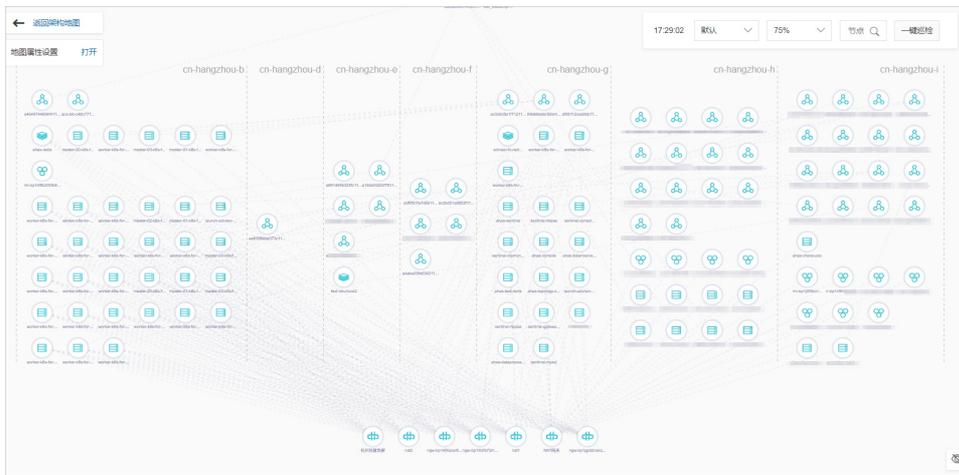
架构感知工作流程



AHAS架构感知的工作流程包括四个步骤:

- i. 数据采集
  - ii. 关系构建
  - iii. 特征识别
  - iv. 架构可视化
- 界面展现

架构可视化



## 流量防护

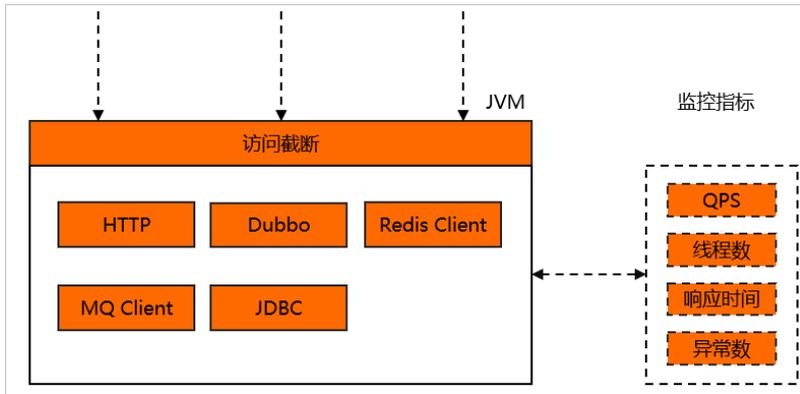
- 功能

在当今日益复杂的应用环境下，应用设计应该遵循“面向失败”的设计原则，对上下游的依赖零信任。借助流量控制、熔断降级模块，您的应用有能力对流量采取限流控制，并对下游依赖进行降级处理。

- 工作流程

流量防护模块支持主流的Java框架，包括HTTP和Dubbo。该模块可以实时监控框架的QPS（Queries per Second，每秒查询数）、线程数、响应时间、异常数等指标，并有选择地截断对这些框架的访问，从而保护应用的可用性。此外，利用AHAS提供的SDK，您还可以采取更细粒度的代码级限流降级防护措施。

### 流量防护工作流程

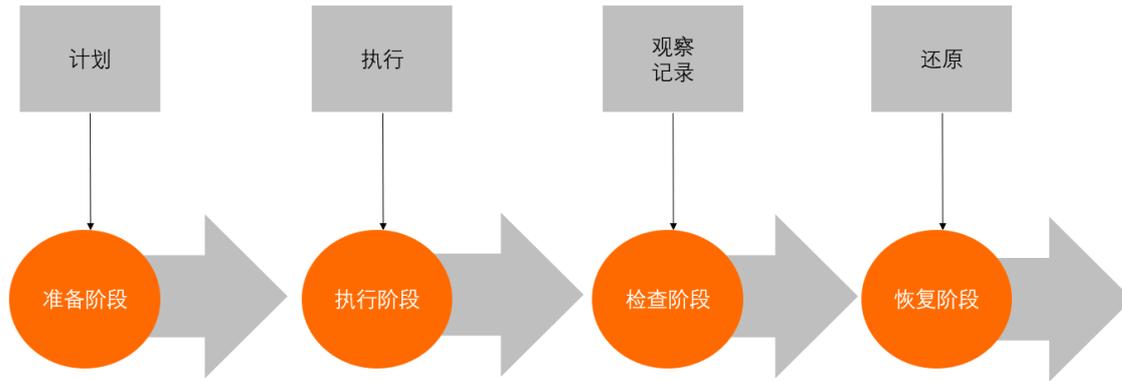


## 故障演练

故障演练是一款遵循混沌工程实验原理并融合了阿里巴巴内部实践的产品，提供丰富故障场景实现，能够帮助分布式系统提升容错性和可恢复性。

### ● 流程

故障演练建立了一套标准的演练流程，包含准备阶段、执行阶段、检查阶段和恢复阶段。通过四阶段的流程，覆盖用户从计划到还原的完整演练过程，并通过可视化的方式清晰的呈现给用户。



### ● 适用场景

故障演练可适用于以下典型场景：

- 衡量微服务的容错能力。
- 验证容器编排配置是否合理。
- 测试PaaS层是否健壮。
- 验证监控告警的时效性。
- 定位与解决问题的应急能力。

更多信息，请参见[故障演练概述](#)。

## 功能开关

功能开关是一个轻量级的框架，既可用于集中管理各应用的业务开关，也可以管理应用的运行时基础配置。

### ● 功能描述

- 业务开关管理：按照统一的方式定义、管理业务运行时功能切换。
- 运行时配置管理：支持定义任意类型的配置项，且自动根据类型做自动拆装，代码层无需关注类型信息，简化运行时配置项使用流程。

### ● 使用场景

- 业务开关管理。一个业务经常由多个系统、多个功能模块组成，为保证某些业务的动态性，后端程序通常会用开关来控制程序的逻辑，以达到在系统运行时切换运行逻辑的目的。
- 运行时配置管理。无需写死的URL、接口名、阈值和读取文件用的编码、黑白名单等，直接使用功能开关即可方便快速创建运行时能覆盖的动态配置。

## AHAS使用场景

- 微服务高可用

如今，Spring Cloud、Kubernetes的微服务架构对服务的高可用性提出了前所未有的挑战。借助AHAS，您可以在零代码改动的前提下，快速使用SaaS化的高可用服务，包括架构可视化、架构变化追踪、故障演练和流量防护的保护。

AHAS支持Spring Cloud组件的快速发现、故障演练能力，以及SpringBoot应用的一键流量防护。

AHAS支持阿里云容器服务Kubernetes应用的快速接入。

- 传统应用高可用

由于缺乏高可用能力，对于传统的Java单体应用和分布式应用而言，如何保障稳定性和连续性向来是个难题。

AHAS提供了应用高可用保障所必需的架构实时展现与追踪、架构高可用性故障演练，以及Java应用零代码改动接入流量防护的能力。即便是已上线的应用，也无需升级改造。此外，对于引入的开源或第三方组件，AHAS可以进行自动化智能评估，并提供优化使用方面的指导。

## 相关文档

- [第三方组件和云服务支持列表](#)
- [开源Sentinel与商业化AHAS流量防护对比](#)

## 2. 名词解释

本文列出了应用高可用服务 AHAS 的主要术语。

### 故障演练相关术语

#### 标签

即用户对演练自定义的内容。一个演练可以有多个标签，便于用户对演练进行分类和管理。

#### 场景

即活动的分类，例如磁盘满载和磁盘IO高都属于磁盘这个场景。

#### 草稿

即未配置完成的演练或修改后未保存的演练。

#### 活动

即演练流程中的一个可运行的节点，一个演练由多个活动组成。

#### 演练

即一个完整的、可执行的流程配置。

#### 演练任务

即一次演练的执行记录，每执行一次演练就产生一个任务。

#### 演练报告

即一个演练任务的结果，一个演练任务对应一份报告。

#### 演练参数

即一个活动运行时所需要的参数，参数决定了活动执行的结果。

#### 演练对象

即实施演练的目标主机或设备。目前以活动为最小颗粒度，不同活动可以对不同的主机或设备执行。

## 流量防护相关术语

### 规则

即围绕应用资源的实时状态，设定规则（Rule）控制资源调用。规则类型包括：流控规则、降级规则、隔离规则。

### 流量防护

AHAS 流量防护（限流降级）提供专业多样化的流量控制、熔断降级手段，并提供强大的监控信息，以及实时的动态规则推送能力，为您的应用和服务保驾护航。

### 流量控制

即监控应用流量的 QPS 或并发线程数等指标，控制达到阈值的流量，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。

### 熔断降级

即限制不稳定资源的调用，让请求快速失败，避免影响到其它的资源而导致级联错误。

### 系统保护

即从整体维度对应用入口流量进行控制，结合应用的 Load、总体平均 RT、入口 QPS 和线程数等维度的监控指标，让系统的入口流量和系统的负载达到一个平衡，让系统尽可能跑在最大吞吐量的同时保证系统整体的稳定性。

② 说明 Sentinel 基于 TCP BBR 拥塞控制算法实现流量塑形，对入口流量进行系统保护。

# 3. 产品版本说明

本文介绍应用高可用服务AHAS（Application High Availability Service）产品的发布历史。

## AHAS版本说明

变更时间	变更分类	变更内容	产品文档
2020年8月	新功能	集群流控可以精确地控制整个集群的调用总量，结合单机限流兜底，更好地发挥流量防护的效果。	请参见 <a href="#">集群流控</a> 。
2020年8月	新功能	通过Nginx Sentinel模块可以快速接入至AHAS，当有请求流量时，可以在网关防护中查看Nginx网关请求的实时QPS与RT数据。	请参见 <a href="#">接入Nginx</a> 。
2020年8月	体验优化	<ul style="list-style-type: none"> <li>流量防护全新UI，应用概览数据统计更直观，体验更佳。</li> <li>防护监控数据增加分位值、TOP统计及日粒度统计，便于您更好的查看接口。</li> </ul>	请参见 <a href="#">接入应用概述</a> 。
2020年4月	体验优化	AHAS概览页全新升级，高可用体系直观呈现	无
2020年3月	新功能	<ul style="list-style-type: none"> <li>功能预案功能。</li> <li>AHAS流量防护接入方式新增GO语言。</li> </ul>	请参见 <a href="#">接入Go应用</a> 。
2019年12月	新功能	AHAS计费模型优化，增加入门级防护模式。	请参见 <a href="#">应用基础设置</a> 。
2019年6月	新功能	<ul style="list-style-type: none"> <li>架构感知支持云服务Redis的实例识别。</li> <li>故障演练支持Java故障注入。</li> </ul>	无
	新功能	<ul style="list-style-type: none"> <li>展示资源使用情况。</li> <li>架构感知主机层支持按照可用区划分。</li> </ul>	无
2019年5月	新功能	架构感知： <ul style="list-style-type: none"> <li>主机层增加集群过滤条件。</li> <li>集成云监控的指标报警信息（ECS、RDS、SLB）。</li> <li>增加对CPU使用率和Load监控。</li> </ul>	无
	新功能	流控降级：增加网关流控功能，支持API分组。	请参见 <a href="#">网关防护概述</a> 。
	新功能	架构感知： <ul style="list-style-type: none"> <li>主机层增加集群过滤条件</li> <li>集成云监控的指标报警信息（ECS、RDS、SLB）</li> <li>RDS、SLB实例节点展示具体实例描述</li> </ul>	无
	新地域	应用高可用服务AHAS上海Region开通上线。	无

变更时间	变更分类	变更内容	产品文档
2019年4月	新功能	架构感知： <ul style="list-style-type: none"> <li>支持RDS云服务的实例识别。</li> <li>容器组中Service Ingress支持流量配比展示。</li> <li>应用名称识别与流控名称同步。</li> <li>支持展示SLB实例监控信息，可识别SLB实例信息。</li> </ul>	请参见 <a href="#">架构感知概述</a> 。
	新功能	<ul style="list-style-type: none"> <li>支持展示Ingress入口。</li> <li>支持展示非Running状态的Pod。</li> <li>架构感知记录回溯时间修改为7天。</li> </ul>	无
	新功能	故障演练： <ul style="list-style-type: none"> <li>支持系统指标图的检测。</li> <li>支持磁盘、网卡的下拉选择。</li> </ul>	请参见 <a href="#">创建演练</a> 。
2019年3月	新功能	流控降级： <ul style="list-style-type: none"> <li>新增子账号授权功能。</li> <li>增加规则变更的操作日志。</li> </ul>	请参见 <a href="#">流量防护RAM授权</a> 和 <a href="#">应用概览</a> 。
	体验优化	故障演练：演练探针卸载时增加提醒。 流控降级： <ul style="list-style-type: none"> <li>规则弹窗中默认折叠来源应用。</li> <li>水位详情增加文案说明。</li> <li>支持TPS跌0的展示。</li> <li>支持监控图的缩放。</li> <li>增加RT指标坐标轴</li> </ul>	无
2019年1月	新功能	故障演练服务全新升级，提供更流畅的演练体验。	请参见 <a href="#">故障演练概述</a> 。
	新功能	架构感知支持采集并展示容器服务Kubernetes事件。	请参见 <a href="#">事件列表</a> 。
2018年12月	新功能	支持开源Sentinel SDK接入，允许用户独立使用AHAS的流控降级功能。	请参见 <a href="#">接入Dubbo应用</a> 。
	新功能	AHAS Agent最新版本支持探针的自动拉起与升级。	请参见 <a href="#">管理探针</a> 。
	体验优化	AHAS Agent性能优化，降低资源消耗。	无
	新功能	架构感知支持表格化展示所有节点，并支持导出表格数据。	请参见 <a href="#">列表信息</a> 。
	新地域	应用高可用服务AHAS公网Region开通上线，支持不限地域、不限厂商的用户接入。	请参见 <a href="#">应用高可用服务接入</a> 系列文档。

变更时间	变更分类	变更内容	产品文档
	新功能	优化流控降级： <ul style="list-style-type: none"> <li>流控降级前端监控改版；</li> <li>展现请求调用链路细节和调用关系；</li> <li>新增机器列表，提供机器IP地址、进程ID等具体信息。</li> </ul>	请参见 <a href="#">应用防护概述</a> 。
	新功能	新增系统规则，从应用级别进行入口流量管控。	请参见 <a href="#">配置系统规则</a> 。
2018年11月	新功能	架构感知中的进程拓扑图支持结构化展示和多维度筛选。  进程拓扑图展示为四层：第一层是云服务或入网、出网进程，第二层是应用服务进程，第三层是存储服务进程，第四层是其它进程。	请参见 <a href="#">架构感知概述</a> 。
	新功能	架构感知支持查看节点详情的监控数据。	请参见 <a href="#">架构感知概述</a> 。
	新地域	应用高可用服务AHAS深圳Region开通上线。	无
	新功能	架构感知支持零门槛体验Demo实例。	无
	新功能	架构感知支持TCP6协议。	无
2018年9月	产品发布 上线	<ul style="list-style-type: none"> <li>架构感知：自动感知应用的拓扑结构，以可视化的方式直观呈现应用对基础架构的依赖关系和组件间的依赖关系，并持续记录上述依赖关系，让您对系统架构一目了然。</li> <li>应用高可用能力测评（现已更名为“故障演练”）：根据架构感知模块捕捉到的架构数据主动制造故障，检验应用系统及其各组件在故障下的可用性表现，从而验证应用系统的高可用能力，提前暴露故障隐患，帮助您针对性地应对风险。</li> <li>应用高可用防护（现已更名为“流控降级”）：提供零代码改动、可一键启用的流控和降级两大可用性防护服务，为您的应用保障高可用性。</li> <li>规则管理：提供统一的流控和降级规则管理模块，方便您集中管理规则。</li> </ul>	请参见 <a href="#">什么是应用高可用服务AHAS</a> 。

## 4. 探针版本说明

### 4.1. 应用高可用探针版本说明

应用高可用探针（即 AHAS 探针），使用架构感知、故障演练功能前需安装此探针。本文介绍 AHAS 探针版本发布历史。

版本号	变更时间	变更内容
1.5.0	2019 年 7 月 10 日	<ul style="list-style-type: none"> <li>• 新增对 Kubernetes Virtual Node 的支持。</li> <li>• 添加了 TLS 认证通信方式。</li> <li>• 添加数据压缩传输的能力。</li> <li>• 优化 Kubernetes 数据上报方式：一个集群 meta 数据只上报一份。</li> <li>• 精简数据上报内容，降低对带宽占用。数据总量降低 1/5。</li> </ul>
1.4.7	2019 年 6 月 25 日	<ul style="list-style-type: none"> <li>• 新增 Java 应用演练场景，和开源功能对齐。</li> <li>• Kubernetes 部署添加 affinity 亲和性配置，防止 Agent 部署在 VK 上。</li> </ul>
1.4.6	2019 年 5 月 13 日	<ul style="list-style-type: none"> <li>• 新增磁盘 IO 数据采集。</li> <li>• 修改 ChaosBlade 安装逻辑。</li> <li>• 新增 ChaosBlade 升级功能。</li> <li>• 修复 /proc 访问限制。</li> </ul>
1.4.5	2019 年 5 月 5 日	<ul style="list-style-type: none"> <li>• 添加非 running pod 上报。</li> <li>• 添加 ingresses 资源上报。</li> <li>• 修改 eventer 版本至 1.0.1，添加推送开关。</li> <li>• ChaosBlade 支持网卡和磁盘查询。</li> </ul>
1.4.4	2019 年 4 月 12 日	<ul style="list-style-type: none"> <li>• 新增 events 组件，抓取 Kubernetes 事件流。</li> <li>• 修复数据推送。</li> <li>• 修改数据推送策略。</li> <li>• 更新 ChaosBlade 版本。</li> </ul>
1.4.3	2019 年 3 月 29 日	<ul style="list-style-type: none"> <li>• 采集数据增量上报。</li> <li>• 建立连接时上报 CPU、内存信息。</li> <li>• 优化 AHAS Agent 安装失败错误提示。</li> <li>• 优化 Kubernetes 下 AHAS Agent Volumes。</li> <li>• 新增 Service EXTERNAL-IP 数据上报。</li> <li>• 修改容器 sar 信息采集周期。</li> <li>• AHAS Agent 接入容器服务应用目录。</li> </ul>
1.4.2	2019 年 2 月 26 日	修复部分容器进程、网络数据未抓取问题。

版本号	变更时间	变更内容
1.4.1	2019年2月19日	<ul style="list-style-type: none"><li>• 日志打印修改。</li><li>• 修复启动关闭超时。</li><li>• 修复 helm chart 镜像地址。</li><li>• 新增卸载退出码。</li></ul>
1.4.0	2019年1月30日	<ul style="list-style-type: none"><li>• 新增对 Kubernetes 集群的支持。</li><li>• 新增对 Kubernetes 集群自动升级的支持。</li><li>• 添加 kubectl 安装 Agent。</li><li>• 添加 helm chart 安装 Agent。</li><li>• 修复 stime 为 0 时，进程数据抓取问题。</li><li>• 优化网络抓取逻辑，当服务器存在 5 万连接时，CPU 使用率在 10% 以下。</li><li>• AHAS Java Agent 支持集群限流。</li></ul>
1.3.3	2019年1月25日	<ul style="list-style-type: none"><li>• 修复兜底方案。</li><li>• 设备 connect 新增版本字段。</li></ul>
1.3.2	2019年1月21日	Sentinel SDK 上线。
1.3.1	2018年12月27日	支持探针自动升级。

## 4.2. 应用高可用Java探针版本说明

应用高可用Java探针，这是针对JVM的Java探针，如果需要使用流量防护功能，可安装此探针。本文记录了Java Agent的版本发布说明。

Java Agent探针版本

## Java Agent

版本号	版本说明
1.7.8	<ul style="list-style-type: none"> <li>修复Spring Cloud Gateway、Zuul等网关插件不生效的问题。</li> <li>应用appType支持通过环境变量配置。</li> </ul>
1.7.7	修复JDK 11下agent初始化失败的问题。
1.7.6	熔断降级支持渐进式恢复策略。
1.7.5	<ul style="list-style-type: none"> <li>熔断降级特性改进：引入半开启探测模式，支持任意统计时长，RT模式改进为慢调用比例模式。</li> <li>集群流控通信与异常处理优化。</li> <li>支持张家口region。</li> </ul>
1.7.4	<ul style="list-style-type: none"> <li>支持集群流控特性。</li> <li>支持SAE方式接入。</li> </ul>
1.7.3	支持MySQL CJ Driver的SQL埋点。
1.7.2	<ul style="list-style-type: none"> <li>支持启动参数和环境变量控制所有插件模块的开启和关闭。</li> <li>改进SQL含换行符情况下的处理，添加SQL埋点长度限制。</li> <li>去除RT统计限制。</li> <li>修复部分非daemon线程导致main函数执行结束后无法退出的问题。</li> </ul>
1.7.1	<ul style="list-style-type: none"> <li>同步基线至开源1.7.0版本，优化内存占用。</li> <li>增加通用配置功能，支持簇点数目限制、context数目限制、来源数目限制、最大统计RT等基础配置项，支持动态配置。</li> <li>增加热点规则支持。</li> <li>默认关闭MQ相关插件以提升启动速度。</li> </ul>
1.7.0	<p>新增Memcached (spymemcached)、gRPC、Jedis、RocketMQ (callback模式)、RabbitMQ插件支持。</p> <p> 说明 MQ插件可能会导致应用启动速度变慢。</p>
1.6.0	新增MySQL (5.x driver) 或Oracle SQL埋点支持。