SOFAStack

单元化应用服务 LHC

产品简介

产品版本：AntStack Plus 1.11.0

文档版本：20220928

蚂蚁集团
ANT GROUP

法律声明

蚂蚁集团版权所有©2022，并保留一切权利。

商标声明

免责声明

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 `cd /d C:/window` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *Instance_ID*
[] 或者 [a|b]	表示可选项，至多选择一个。	`ipconfig [-all|-t]`
{} 或者 {a|b}	表示必选项，至多选择一个。	`switch {active|stand}`

目录

1.单元化应用服务

1.1. 什么是单元化应用服务

随着云原生技术在业界的持续升温，越来越多的金融客户希望将能把云搭建在拥有云原生能力的平台之上，利用各类云原生的红利迅速实现技术转型，促进业务的敏态发展和持续创新。

单元化应用服务（LDC Hybrid Cloud，简称 LHC）定位于在云原生基础设施之上，在多机房、多地域的Kubernetes多集群场景，提供应用管理、发布运维、流量调拨、配置同步等能力。

作为开发运维人员日常接触的 PaaS 管控层产品，帮助解决应用和逻辑单元管理、按单元的配置变配、网络流量调拨、监控元数据配置等能力。不仅满足金融场景下同城多活和跨地域容灾的业务需求，同时能够让基础设施享受到容器化基础设施、云原生架构的技术红利。

LHC 旨在提供从单 Kubernetes 集群向多活联邦集群演进的能力，提供具备容灾能力的同城双活、 两地三中心及更多机房级多活容灾场景。并可以配合 SOFAStack 各中间件产品、OceanBase 分布式数据库，形成单元化异地多活架构解决方案。

1.2. 产品优势

高可用与容灾

金融行业（尤其是银行）出于监管需求，其生产环境通常是要求在自建 IDC 里，即专有云环境。这对云原生底盘 Kubernetes 的高可用和稳定性提出了极高的要求，比如在双机房情况下如何做到双活，在有异地机房时，如何保证通讯的延迟不导致上层业务的大量超时而影响到最终用户的资金、账户安全。在灾难发生时，如何减少影响面，尽可能减少受影响的用户，并缩短影响的时间。

变更管控与保障

针对金融级场景下大规模分布式系统的特点，提供了丰富的发布策略以满足不同的场景，如：分组发布、Beta 发布、灰度发布等，帮助传统架构平滑过渡，适应金融技术风险保障需求，实现大规模金融级运维场景下的容器服务落地。

- 提供发布单追溯功能，记录和升级容器应用变更和版本管理。
- 支持多应用批量发布、设置应用间发布依赖顺序。
- 提供分组分批次、可灰度、可暂停、可回滚的发布策略和控制能力。
- 容器支持按需、按资源水位、按计划弹性伸缩。

无限弹性扩展

随着金融行业思路的转变，将金融业务内嵌到个人生活服务中，必然会产生诸多高频场景。

采用分布式架构，可以让业务在出现热点后，进行拆分、扩容，以应对流量激增。但如果是由于物理资源不够，或者数据层瓶颈，仅采用应用层的分布式就不够了。如何做到从上至下有一套完整的体系，从接入层到应用层最后到数据层进行规整，可以作为一个整体单元进行快弹、快缩。

多地域应用服务统一发布

多地域机房架构下，同一个应用服务的配置中支持多地域差异化配置，保证一次发布可以在多个机房中生效，从而有效降低多机房应用服务发布层面的运维成本。

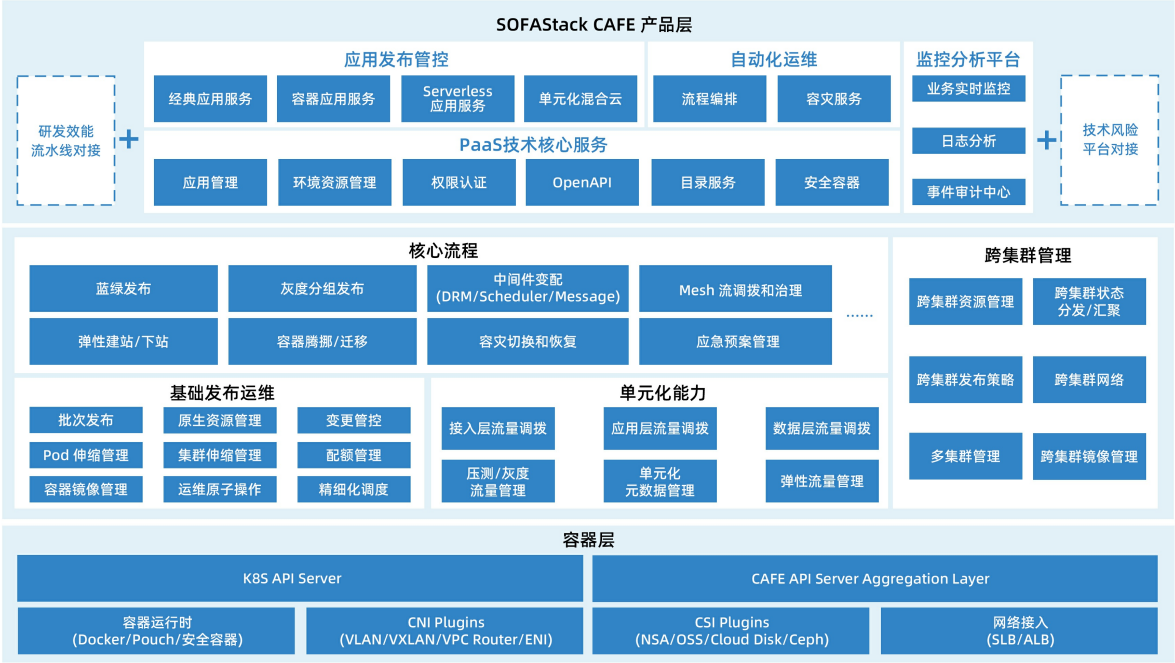
1.3. 产品架构

单元化应用服务，基于 Kubernetes 基础设施提供跨集群资源管理、发布运维及单元化管控能力。其纳管的每个集群，均为完整、标准的 Kubernetes 集群及相关扩展。

在集群之上，通过联邦管控平面，协调各集群资源、应用和配置，以提供应用变更管控、分组发布、镜像管理、流量调拨、元数据管理、集群资源管理等功能。用户可以通过控制台、命令行或 SDK 以标准方式对单集群及联邦管控层进行交互。

作为分布式架构的平台管控层，单元化应用服务还承担与各相关系统的对接能力。

- 研发效能产品，可提供持续集成与交付功能，生成 LHC 产品的发布单。
- 与各 PaaS 基础产品对接，提供权限、资源、应用元数据、工作空间、部署单元等领域模型的管理能力。
- 与监控分析平台对接，向用户提供完整集群资源和业务应用的可观测性。



1.4. 功能特性

跨集群应用管理

- **应用服务（FedAppService/AppInstanceGroup）**
  - 创建应用服务模板，根据部署单元（Cell）做相应的属性同步或变配。
  - 应用配置版本管理。
- **发布单管理**
  - 有效管控单个或一组应用的发布状态和策略，检查发布变更。
  - 在发布准备时，可以配置应用依赖信息，自动实现串行和并行发布策略。
  - 在发布过程中，确保前置、发布、后置任务事件完整记录，方便追踪发布历程。
- **应用服务级别监控、日志和事件配置和展现**
  - 配置跨集群的监控，并且可以聚合、显示。
  - 配置跨集群的日志收集，并且可以聚合、搜索、显示。

- 展示应用服务级别事件。

跨集群资源管理

- Service：是 Kubernetes 自带的核心资源，用于抽象向多个 Pods 访问的接口，类似于负载均衡的概念。目前支持 Cluster IP、Loadbalancer、Node Port、Headless、ExternalDNS 类型。支持通过 YAML 和 UI 方式进行 Service 的创建和修改。
- ConfigMap：存储应用所需配置信息的资源类型，用于保存配置数据的键值对。ConfigMap 可通过 YAML 和 UI 方式进行创建和修改。
- Secret：一种用于存储工作负载所需要认证信息、密钥的敏感信息等的资源类型。Secret 可通过 YAML和 UI 方式进行创建和修改。
- 工作负载（Kubernetes NativeWorkloads）：工作负载包括 Kubernetes 原生的 Deployment、StatefulSet、DaemonSet、Job、Pod，支持工作负载配置、监控、更新、删除等生命周期管理。

多集群管理

- 集群创建与运维
- 节点扩容与展现
- 命名空间管理
- 提供对工作空间与工作空间组的管理

镜像中心

单元化应用服务所包含的镜像中心，提供和用户 Gitlab 代码库关联的镜像构建和版本管理能力，满足从代码到镜像的一键构建需求。同 IAM（身份认证管理系统）对接满足安全保障。

- 镜像构建和版本管理
- 镜像技术栈创建

单元化配置

- 逻辑单元配置
- 部署单元配置

接入层管理

- 接入网关与流量管理
- 应用层流量配置

应用商店

支持通过 HelmChart 应用商店一键编排复杂容器应用。

1.5. 应用场景

LHC 服务于云原生模式下，通过一套应用 PaaS 平台，提供统一的应用、资源管理，以及发布运维视图，实现多集群管理、跨集群应用运维发布、资源管理、流量管理。

具体来说，可细分为以下场景：

- **同城双活（active-active）**

  在同一个地域 Region，建立两个或更多可用区下的多个 Kubernetes 集群。

- **两地三中心**
  - 在同城双活的基础上，增加一个异地机房，做数据和应用备份。根据网络延时和带宽情况，可选择异地热备、温备和冷备三种方案。
  - 在异地存在延迟的情况下、业务可接受的范围内，可以选择异地热备。正常情况下访问同城双活测，在容灾期间切至异地灾备机房，提供业务访问。
- **异地多活（Multi-region active-active）**

  数据层做分片（Sharding），不同的 AZ 可以划分为更多的逻辑单元（Logic Data Center），处理不同的数据分片。 尽量保证数据访问的链路从接入层到应用层再到数据层不会出现跨可用区的调用。这种架构下，可以做到任意数量地域的多活。
- **异构基础设施下的混合云**

  通过 Kubernetes 屏蔽掉底层 IaaS 的差异性，可充分利用公有云上的资源，将业务同时在专有云和公有云上进行部署，并进行统一运维管控。在该场景下，可以帮助金融客户达到以下目的：
  - 减少开发、测试资源的投入：专有云部署生产应用，公有云按需部署开发测试应用。
  - 线下快速容灾需求：应国家监管需求，需要在线下部署一套环境，以应对公有云上的突发情况（例如天弘余额宝）。
  - 弹性扩容：结合异地多活架构，使业务能够按需进行机房级的无限水平扩展。

1.6. 使用限制

专有云场景单元化应用服务（LHC）的功能限制如下。

应用发布

- 同一个应用服务，不能同时进行发布和运维操作。
- 以下操作都会在发布应用服务时删除并重建 Pod，而非原地升级：
  - 调整应用服务中 Pod 的 CPU 和 memory 的值。
  - 改变 Pod 中容器的数量。
- 在应用服务版本中更改所关联的负载均衡（LB）或修改统一接入模式下的转发规则都会导致业务中断。
- 应用服务不支持同时配置多个访问配置。
- 负载均衡（LB）在不同的应用服务间不能复用。

集群

- 由于网络模式限制，当前暂不提供 cluster IPservice，集群内通信只能走 Pod IP 或 LBService。
- 存储：飞天版本依赖 ACK 支持存储类型。物理机版本依赖 YODA 本地存储。

网络

统一接入集群支持节点扩缩容。

镜像

- 镜像构建：限制只能同时进行 3 个构建任务。
- 镜像中心：用户只能往所在租户的 namespace 下 push 镜像，或者从公共的 namespace 拉取镜像。镜像中心只能在 region 内部访问。

1.7. 基础术语

中文	英文	说明
混合云	Hybrid Cloud	混合云狭义上指“公有云 + 私有部署”的混合形态，通过平台能力抽象统一、自动化部署、配置管理等方面的技术和产品，淡化开发和运维人员对底层基础设施的关注，使应用和数据能够在混合数据中心环境中进行部署和运维。
单元化工作空间	LDC workspace	提供单元化能力，可用于同城双活及异地容灾场景。您可以通过单元化工作空间组对用户资源进行隔离，不同工作空间组下的集群彼此隔离。
工作空间组	WorkspaceGroup	工作空间（Workspace）在多地域的扩展，在多地域内对资源进行分组隔离管理。多个地域的网络可以通过专线高速通道实现互通。
逻辑单元	Zone	一个单元被称为一个 Zone，有 3 种不同类型：RZone、GZone、CZone。单元的特点如下： • 同一个应用在每个单元中拥有独立使用的资源。 • 同一个应用的业务在不同单元中按水平方向拆分。 • 不同单元处理的业务分片不重叠。
单元化架构	LDC architecture	应用层按照数据层相同的拆片维度，将整个请求链路收敛在一组服务器中，从应用层到数据层就可以组成一个封闭的单元。 数据库只需要承载本单元的应用节点的请求，大大节省了连接数。“单元”可以作为一个相对独立整体来挪动，甚至可以将部分单元部署至异地。
部署单元	Cell	部署单元（Cell），是指一个能完成所有业务操作的自包含集合，在这个集合中包含了所有业务所需的所有服务，以及分配给这个单元的数据。 单元化架构就是将单元作为部署的基本单位，在全站所有机房中部署数个单元，每个机房里的单元数目不定，任意一个单元都部署了系统所需的所有应用，数据则是全量数据按照某种维度划分后的一部分。
应用服务	Application service	该概念和 经典应用服务 中的应用服务概念一致。 但由于容器有其特殊性，LHC 中的应用服务会包含一些额外的元数据信息，比如容器规格配置、镜像、调度策略、日志配置等。

中文	英文	说明
镜像	Image	镜像是应用包，将配置和相关软件等打在一起的二进制包，并且符合 Docker Image 规范。镜像可以来自任何可被 LHC 网络访问到的镜像中心，对于私有镜像中心，需要在 LHC 中配置相应的访问信息。
构建	Build	构建用于描述从应用源代码到制作出镜像过程的配置信息，包括源代码地址、分支信息、源镜像访问信息、目标镜像信息、Dockerfile 位置信息等。
集群	Cluster	LHC 中集群用于描述您所创建的一个工作负载集群，由多个节点组成。
节点	Node	节点表示一台装了 Docker 和 Kubelet，用以运行应用负载的物理机或者虚拟机。
容器组	Pod	Kubernetes 中最小的部署及管理单元。一个 Pod 由多个相关的并且共享磁盘的容器组成。
命名空间	Namespace	命名空间和 Kubernetes 中相应的概念保持一致，用于表示一个逻辑隔离的空间，会将 Pod、Service、ReplicaSet 等元素隔离，但通常来说，网络不隔离。
原地升级	Inplace upgrade	原地升级是指应用服务中 Pod 的更新方式。发布后 Pod 的 IP 通常和发布前无法保持一致，所在的节点也可能发生变化。该更新方式在镜像替换时不会导致 Pod 删除。
标签	Label	Kubernetes 的原生概念，用于给相应的资源打上标签，做聚合或者匹配。
污点	Taint	Kubernetes 的原生概念，用于给节点做污点标记，通常用于 Pods 的调度策略。 与之相对应的概念为：容忍（tolerance），若 Pods 上有相对应的 tolerance 标记，则可以容忍节点上的污点，并调度到该节点。
保密字典	Secret	Kubernetes 的原生概念，用于存储用户的加密内容。
应用容器	Container	应用程序所运行的隔离工作空间，通常是 Docker 容器或者 Pouch 等兼容 CRI 接口的具有隔离能力的沙箱工作空间。

中文	英文	说明
工作负载	Workload	应用程序运行态的载体及其上层聚合。通常包括：Pod、Deployment、StatefulSet、DeamonSet、Job 等。
配置项	Configmap	Kubernetes 的原生概念，用于存储用户的配置信息。
存储类型	Storage Class	Kubernetes 的原生概念，通常由系统管理员定义，用于指定所支持的存储类别，不同的类别会有不同的存储 SLA、备份策略等差异性。
存储卷	Persistent Volume	Kubernetes 的原生概念，表示一个由系统管理员创建好的存储资源。
存储卷声明	Persistent Volume Claim	Kubernetes 的原生概念，一个存储卷声明绑定一个存储卷。