SOFAStack

AntStack Plus
技术白皮书

产品版本：AntStack Plus 1.11.0

文档版本：20221021

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.云游

1.1. 什么是云游

云游是蚂蚁集团提供的一站式专有云规划、交付、运维平台，管理着蚂蚁产品从诞生到落地的整个生命周期，由云游 Global 与云游 Local 两部分共同组成。

云游旨在通过自动化、智能化的工具平台，解决专有云交付中三大痛点：

- 多环境部署带来的重复性工作。
- 手工部署引发的低效和易出错。
- 需要同时使用多个系统平台操作繁琐。

云游的工作周期可以分为如下四个阶段：

1. 云产品准备期

   准备产品部署所需的所有配置信息。每一个产品版本包含应用镜像、部署所需的资源（服务器、负载均衡、数据库、文件存储、缓存）、产品依赖信息（被依赖的产品也在云游上）、API 信息（供其他产品调用）、启动和部署参数、健康检查配置、发布日志等。每一个产品版本在部署到实际生产环境之前，都需要云游 Global 上进行模拟部署及功能验证，以确保产品版本的可用性和稳定性。

2. 解决方案规划期

   按照客户需求，综合考虑客户环境的硬件网络情况、所选云底座和过往部署情况，从云游 Global 上选择合适的云产品并确定其版本、拓扑和规格，量身定做一份解决方案。解决方案中将会自动汇总所有产品的部署信息，并由交付工程师从实施角度确认解决方案的可行性。

3. 建站部署期

   为客户项目搭建对应的环境，安装云游。通过 AKE 云原生平台（AKE3）集群管控平台初始化集群并添加节点，准备好相应的容器资源。在云游 Local 上导入制作好的解决方案后，根据实际需要，执行应用部署、扩容和缩容、自动化测试等子任务，最终将环境升级到解决方案指定的终态。交付完成后，发布成功的所有产品及其版本、拓扑、规格信息将形成当前环境的基线信息，用于后续的产品运维。基线信息通过文件回传到云游，为下一次产品交付或升级提供基础。

4. 运维期

   建站部署完成后，您可以在云游 Local 上完成日常的运维操作，如应用上线、下线、重启、扩容、缩容，以及资源查询与信息查看等操作。

1.2. 产品优势

云游基于 Kubernetes 实现云原生应用的调度、发布和运维。在发布业务时，云游提供可监控、可灰度、可回滚的管控能力，保证业务的无损。云游控制台提供差异对比功能，让您可以明确看到业务变更。发布单里会展示每个步骤的执行过程、上下文以及执行结果，方便您定位问题。

除此之外，云游还具有以下优势：

- 自动化

  云游可实现自动化部署，您只需通过解决方案指定终态，云游会自动计算需要执行的步骤，并依次完成。

- 统一元数据

  云游可汇总云产品、专有云环境的全部元数据，并提供一站式展示、查询，避免在多个系统间来回切换、数据不一致或无法配合等问题，简化规划部署流程。

- 多角色高效协作

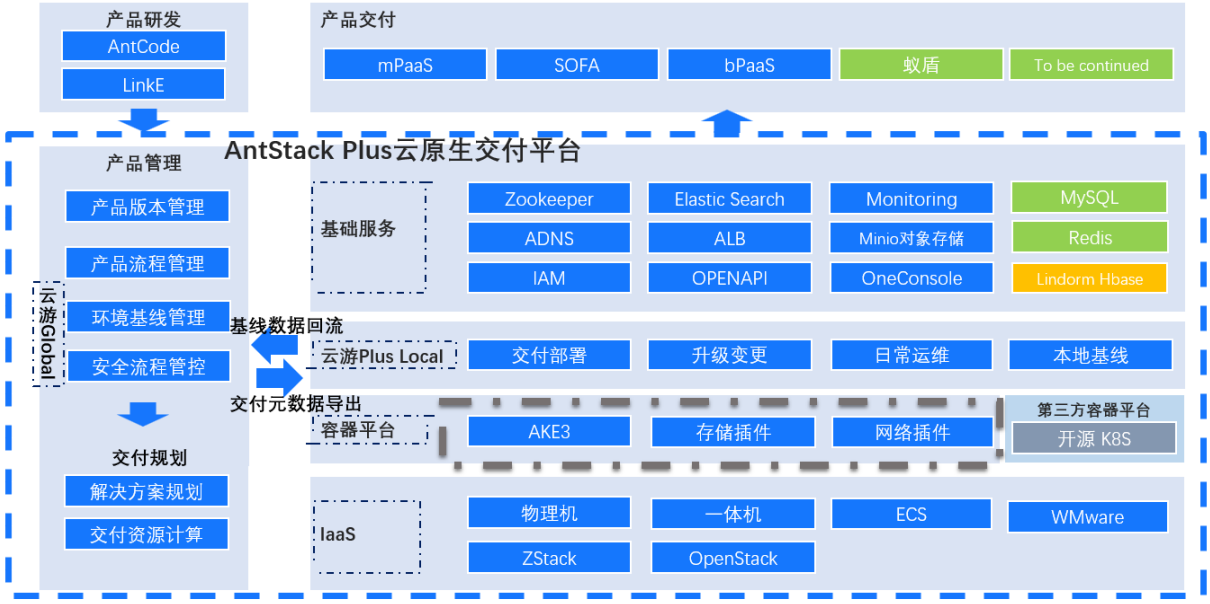
  专有云的部署涉及研发、规划、部署、运维多阶段的流程，参与的角色有产品研发、云架构师、测试工程师、产品集管理员、环境负责人、交付工程师、运维工程师、产品维保工程师等。云游按照业务阶段清晰规划流程，使不同角色既可以专注完成自己的任务，又可以共享透明、全面的信息和数据，降低沟通成本。

1.3. 产品架构

云游采用星型结构，即以云游 Global 为中心，根据您的项目需求搭建多个云游 Local 环境。云游 Global 部署在蚂蚁集团内部，存储所有的产品元数据和环境元数据，以及根据这些元数据制作出的所有解决方案。云游 Global 和各云游 Local 环境之间通过解决方案进行沟通，不同的云游 Local 环境之间相互独立，互不影响。

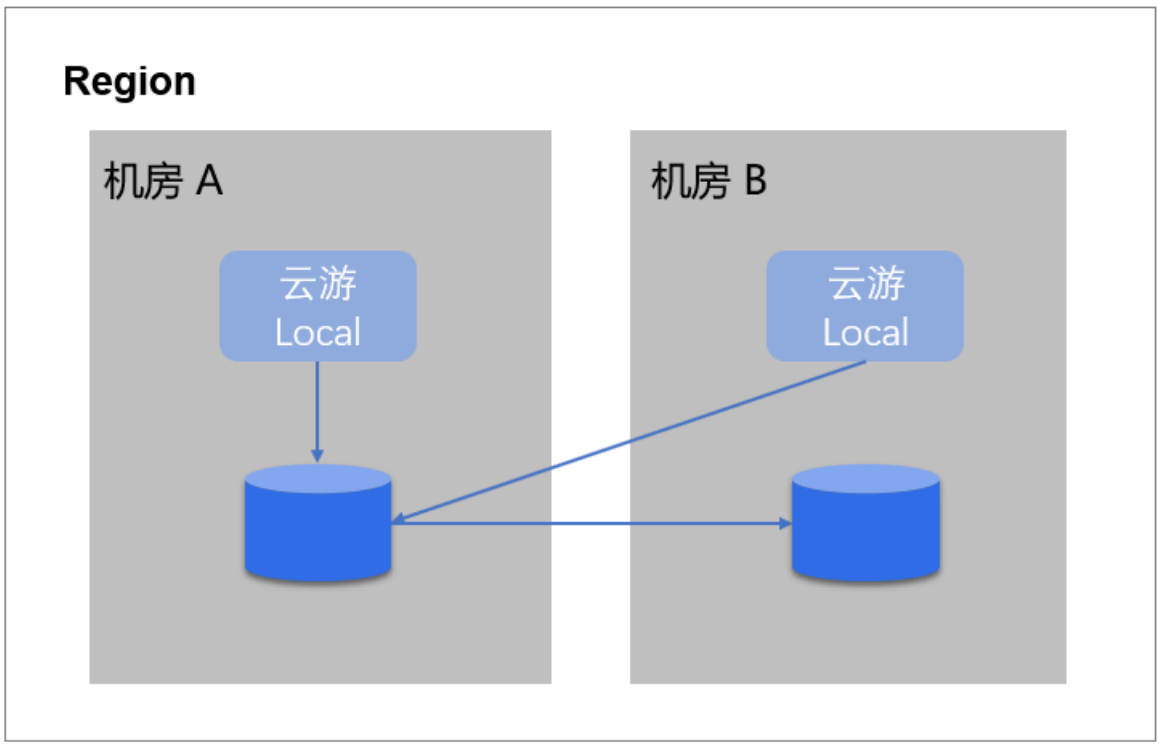
系统架构

云游负责发布和运维上层给客户的各种产品，例如 mPaaS、SOFA 等；下层通过 AKE 云原生平台（AKE3）屏蔽不同底座（如物理机、ECS、OpenStack 等）下的 IaaS 差异。云游内部通过流程编排、资源管理、配置变更等模块支撑对产品的各种运维操作。



部署架构

您可以在双机房部署 2 个云游 Local 的副本，从而保障机房内以及机房级容灾。底层数据库通过阿里云 RDS、OceanBase 或 XDB 方案实现数据主备同步。双机房的云游 Local 都连主库，只有在主库发生故障时，才会将连接切换到备库。



1.4. 功能原理

本文介绍云游支持的功能和实现原理。

支持功能	原理说明
部署	部署是从无到有将产品安装起来的操作，包括为应用创建资源，初始化配置，以及启动应用。这一系列操作都是通过编排引擎实现的自动化，云游通过产品定义的终态描述，将产品一键拉起。
升级	升级与部署类似，区别是升级很少需要创建资源，通常是应用版本的变更，云游通过 AKE 云原生平台提供的能力，会尽量让升级后的应用 Pod 做到节点保持、IP 保持，即常说的“原地升级”。
自动化测试	产品在部署或升级之后，要验证是否与预期的一致，能正常对外提供服务，这个过程也是自动化的。云游通过运行产品提供的自动化测试镜像，根据用例的结果判断是否正确。
扩容和缩容	修改应用的副本数。

支持功能	原理说明
重启	重新启动应用的进程。应用在修改某些配置后，要重启进程才能生效，因此云游提供了重启的功能。
上线和下线	下线即停止应用的进程；上线即开始应用的进程。云游通过替换应用的镜像实现该功能。
回流	将环境基线里的产品或应用的版本、参数和资源信息导出，这些信息都记录在云游的数据库里，用来为下次升级做准备。
卸载	当产品不再需要时，用户可以选择卸载产品。一旦卸载，云游会将产品的资源和对应的数据信息全部删除。

1.5. 附录：基础术语

本文介绍云游相关的基础术语。

标准应用

指提供业务逻辑的应用。

部署规格

指部署拓扑所需要的资源规格配置情况，可以指定资源的数量以及每个资源实例的规格。

部署任务

云游通过计算当前基线（初始态）与解决方案（终态）之间的差异，生成一系列需要执行的子任务。完成所有子任务后，则达到终态，部署完成。

部署拓扑

指一个产品的部署结构，包含哪些应用以及这些应用需要使用的配套基础资源，如服务器、负载均衡、数据库、文件存储。

测试应用

指提供产品自动化测试镜像的应用。

产品版本

指可以独立输出的版本，包含该版本的镜像以及部署配置。

基线

一个环境中，正在运行的、为上层应用服务的所有云产品及其版本、拓扑、规格、参数、资源等信息，称之为这个环境的基线。基线可以类比为一个专有云环境的“水位”，描述一个环境当前真实、稳定的状态。

集群

集群将多台机器的资源整合起来，统一调度、统一管理，是承载部署应用的载体。

节点

云游可以将物理机、虚拟机导入到集群成为节点，节点内的资源会被统一管理和调度。

扩容

为应用增加容器。

容器

是一种轻量级的虚拟化技术，在云游中我们使用了 docker 作为容器引擎。

容器变配

改变容器的规格，如从 2C4G（双核 CPU、4G 内存） 变为 4C8G。

数据回流

指将云游的基线数据用 JSON 文件导出后回收，作为日后升级的基础参考。

缩容

为应用减少容器。

网络

定义了集群中应用容器互相访问的网络配置。

一次性任务应用

指在产品部署过程中，完成一些产品提供方个性化、非标的、云游无法完成的逻辑，例如复杂的数据订正、数据库初始化等。

云产品

对外输出的最小完整单元，能给用户提供一组完整的功能。一个产品中包含多个应用以及这些应用需要用到的基础资源（服务器、数据库、负载均衡、文件存储等）。

2.容器底座

2.1. 什么是容器底座

云原生容器底座是一个基于 Kubernetes 和 Containerd 的应用管理平台，用于部署蚂蚁科技各产品组件。主要由容器引擎 AKE 云原生平台（AKE3）和产品控制台 Captain Plus 构成。

容器引擎 AKE 云原生平台（AKE3）

容器引擎 AKE3（Ant Financial Kubernetes Engine Plus）是将底层物理资源按照计算、网络、存储等进行切分和抽象的容器引擎。AKE3 通过使用 Kubernetes 和 Containerd 技术将整个物理资源进行池化，向上层服务提供按量使用的计算、网络和存储资源。

AKE3 包括以下组件：

- 分布式存储 etcd

  为整个 AKE3 的数据提供存储能力，其本身通过多副本提供数据一致性和高可靠性的能力。

- 中枢控制系统

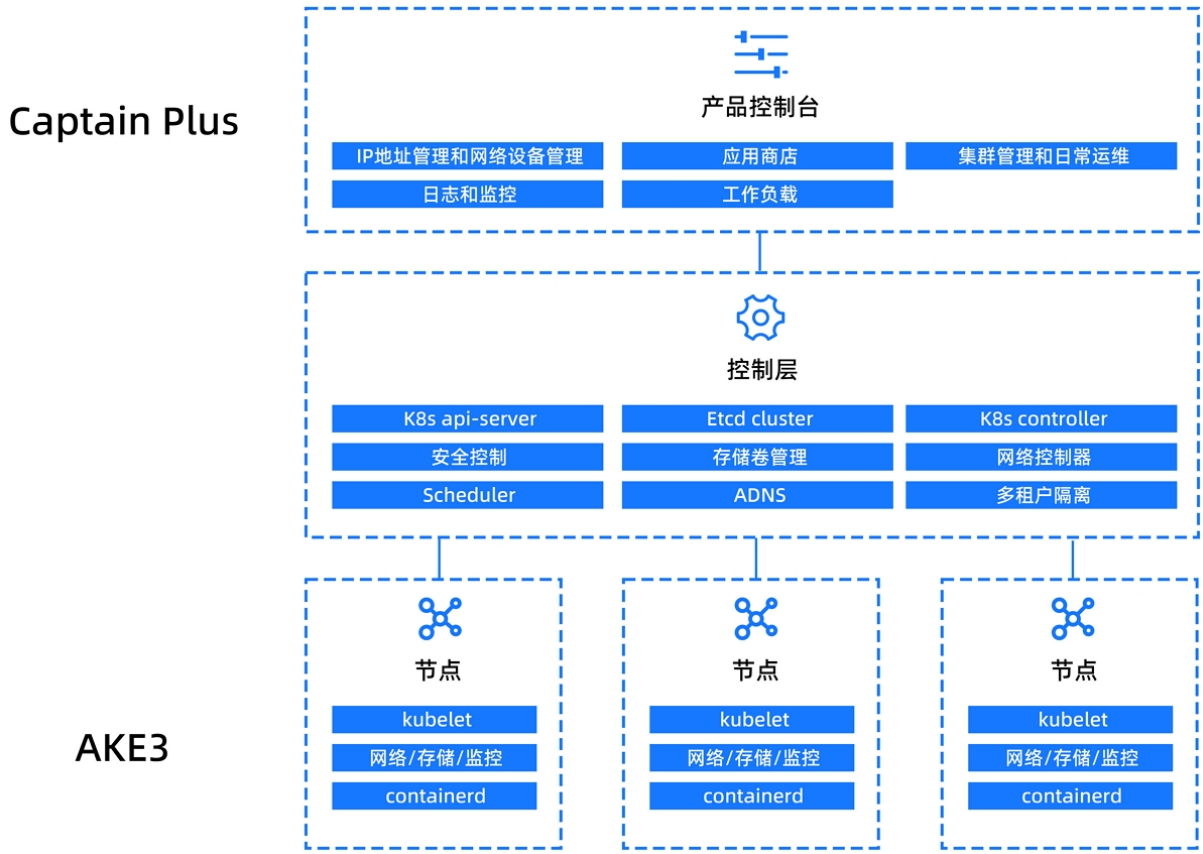
  包括 api-server、controller-manager、network-controller 以及 CNI-service 等。这些系统协同提供如网络管理、资源调度、存储调度、日志和事件等能力。

- 执行节点组件

  包括 Containerd、kubelet、CSI-plugin、CNI-plugin 等组件，运行在每台提供负载能力的物理机或者虚拟机上。

产品控制台 Captain Plus

Captain Plus 是 AKE3 的产品化入口，提供了用户鉴权、集群管理、网络管理、工作负载管理、配置管理、应用商店、镜像中心等能力。整体组件和相互关系如下图所示：



2.2. 产品优势

云原生容器底座是一个基于 Kubernetes 和 Containerd 的应用管理平台，使用时具备如下优势。

网络控制增强

- 提供网络分配调度，支持 VLAN、Overlay（IPIP 和 Vxlan）等网络，提供灵活的网段管理接入能力。
- 实现 Pod 原地升级，不改变 IP 的情况下完成变配、启停操作，满足经典运维机制需求。

资源调度增强

- 支持虚拟机部署。
- 支持大规模集群调度、支持物理服务器、虚拟机形态部署，通过多重互斥标签，实现虚拟机场景下的高可用。

多集群资源隔离

在 K8s 原生基础上增加了多集群能力，一套管理平台可以管理可以管理多个互相独立的 K8s 集群，并满足租户之间的隔离需求，从而降低管理成本。

异构平台支持能力

- 支持产品的优雅升级和发布，包括分组能力、健康检查等。
- 支持物理机、各种虚拟化平台等多种部署平台，满足客户的多样化需求。

运维监控能力

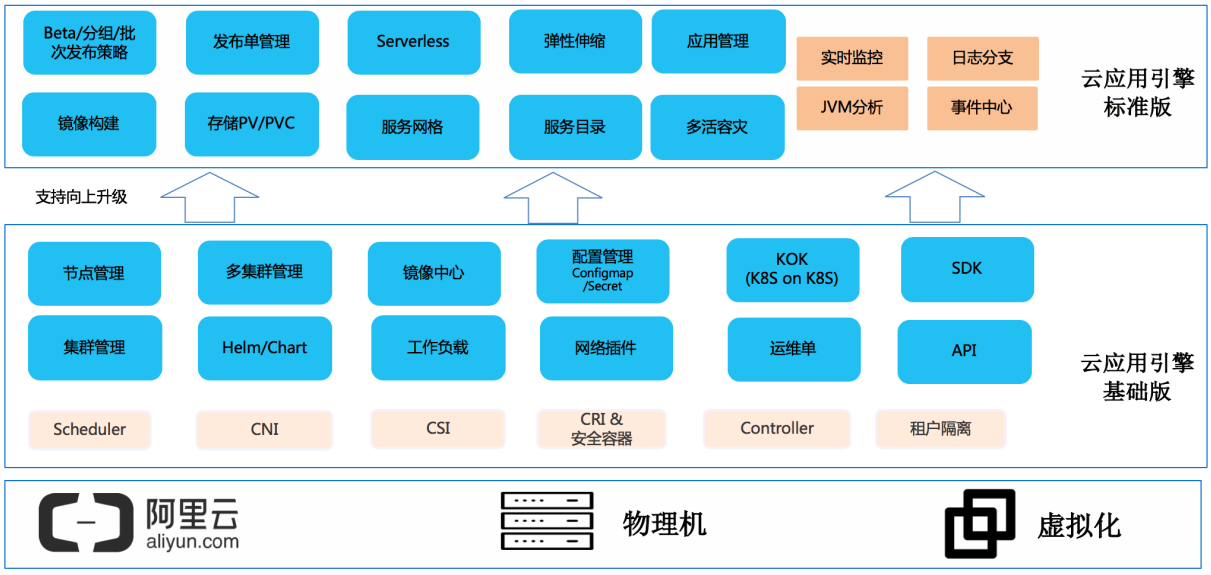
- 结合核心态监控，支持集群、主机、平台组件、应用、容器等多层次、多维度监控告警。
- 监控数据持久化存储，支持多种预警机制，通过预置和开放 API 满足企业多维度监控。
- 提供容器性能指标的可视化监控，可以实时展示平台组件、应用服务等运行状态，提供监控大盘。

2.3. 产品架构

系统架构

云应用引擎基础版中的 AKE 云原生平台（AKE3）通过对底层资源的适配，可以完成对物理机、阿里云、OpenStack、VMware、ZStack 等底层 IaaS 资源适配，向上层服务提供基于容器的计算、网络和存储能力，从而解决各种 IaaS 平台异构问题，满足统一管理的需求。

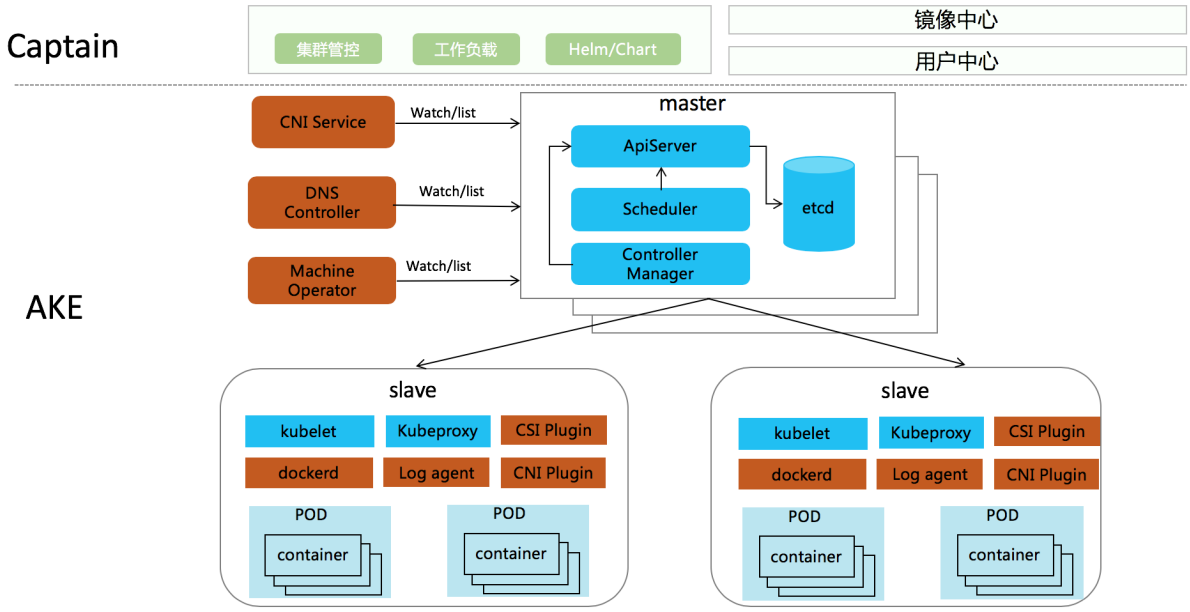
云应用引擎基础版对用户提供了集群管理、应用商店、镜像中心、工作负载等能力，同时也可以平滑升级到云应用引擎标准版。



技术架构

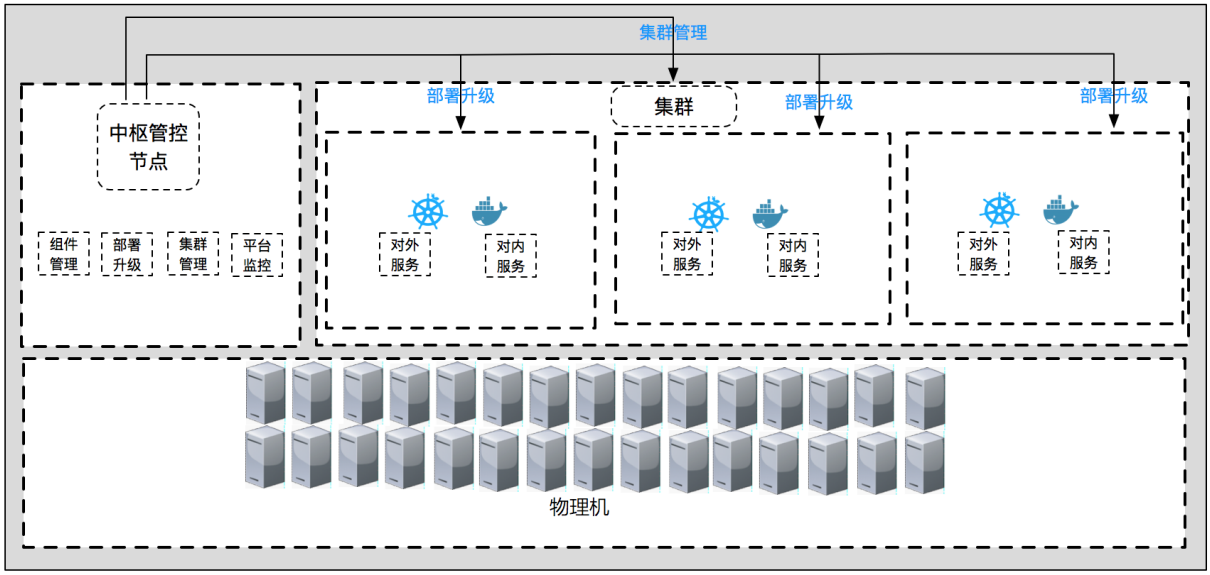
云应用引擎基础版中的 AKE3 构建在开源的 Kubernetes 和 Docker 技术之上。通过蚂蚁集团自研的组件，如网络、调度、存储、日志等，对开源产品进行了适配和增强，更适合开箱即用和追求稳定的用户场景。

Captain 通过打通用户中心和镜像中心，提供了可视化操作界面，方便对 AKE3 集群进行管理升级，也可以进行工作负载的发布运维，同时集成的应用商店可以一键部署已定义和自定义组件。

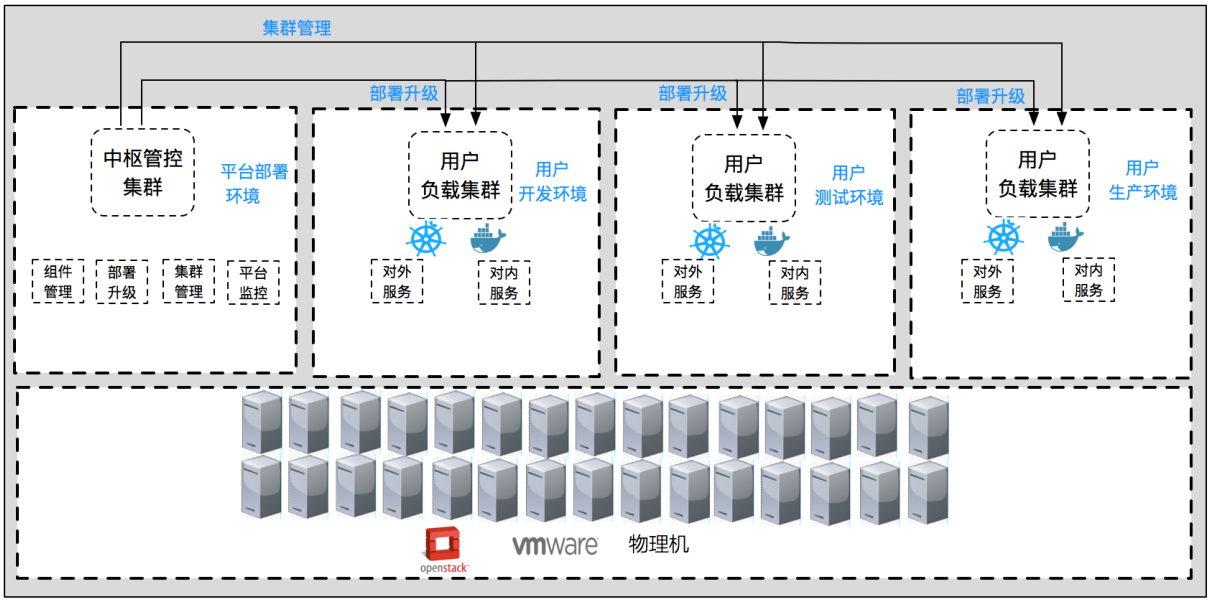


部署架构

云应用引擎基础版支持多种部署架构。在 AntStack Plus 产品输出时，IaaS 层一般为物理机，因此 AKE3 提供的是一个集群，通过为各节点服务器打标，完成产品部署。由云游完成所有 IaaS、PaaS 产品的部署和升级。



AKE3 还支持部署在 OpenStack、VMWare 等环境下，通过使用 kubenetes-on-kubernetes 的技术从物理上将负载集群和中枢管控集群分开，并通过 Captain 可以对多个集群进行统一管理。



2.4. 功能原理

本文介绍容器底座支持的功能和实现原理。

集群管理

提供集群创建、节点管理、容器管理、网络管理等能力。满足一键建站和日常运维所需，支持基于物理机、OpenStack、Aliyun、VMWare类型的 IaaS 底座。

网络管理

提供配置集群网络拓扑、创建网络、查看网络状态、管理网络等能力。支持二层 VLAN 网络、三层 Flannel 网络、Aliyun ENI、Aliyun VRouter 类型的网络方案。

工作负载管理

提供创建工作负载、管理工作负载、删除工作负载、查询工作负载及其容器组等能力。支持管理 Deployment、StatefuleSet、DaemonSet、Job、Pod 类型的工作负载。

配置管理

提供集群配置管理、密钥管理等能力，可以轻松进行配置管理。

权限管理

权限管理基于角色的权限控制系统 RBAC（Role-Based Access Control），可以对用户、角色、权限进行灵活的分配。在 RBAC 系统中，有用户、角色、对象、操作、许可权五个基本数据元素。

权限是资源的使用资格，不用的权限下能够调用、使用的资源数量和类型不同。操作的完成依赖于资源，因此操作的实施以获取相应的权限为前提。当权限被分配给不同的角色时，每个角色拥有一组针对资源的权限。当一个角色被指定给一个用户时，权限就会被传递到拥有该角色的用户。每个用户可以拥有一或多个角色，一个用户想要完成一个操作，需要得到具备相应权限的角色后才能实施。AKE 云原生平台（AKE3）支持多级角色定义，满足企业内资源隔离和细粒度权限隔离的需求。

以下是 Captain Plus 支持的对象和操作示例：

镜像-删除	镜像-push	镜像-pull	镜像-查询
集群-查询	集群-创建	集群-修改	集群-删除
主机-查看	主机-创建	主机-修改	主机-删除
容器-查看	--	--	--

AKE3 平台提供角色模版功能，管理员可以便捷的将预定义模版中的权限添加到角色中，用户可在此基础上进一步修改。另外，用户也可以自定义角色模板。

用户根据实际运维场景，可以使用角色模板快速设定各类角色的权限，包括平台超级管理员、集群管理员、集群运维工程师、集群扩容工程师等。各种角色的权限可以覆盖对工作空间、应用、服务实例、镜像仓库、数据库资源、主机、集群等各类资源的增、删、改、查操作。

以设定某个空间开发者的角色为例，管理员可以先把工作空间的增、删、改、查权限定义成一个角色模板，然后调出该模板、修改工作空间的名称，即可快速设置出特定工作空间的开发人员角色。

日志和监控

容器底座支持如下日志和监控功能：

- 集群、主机、平台组件、应用、容器等多层次、多维度监控告警。
- 监控数据可持久化存储。
- 支持多种预警机制。
- 通过预置和开放 API 满足企业多维度监控。
- 提供智能基线功能，可以智能配置监控数据阈值和区间。
- 支持邮件、短信和即时通信软件的通知方式，并可在告警发生后可按照预案自动进行弹性扩容，故障自愈等。
- 提供容器性能指标的可视化监控，可以实时展示平台组件、应用服务等运行状态，提供监控大盘。
- 提供完善的日志采集、存储和查询能力，也可以通过 HDFS shipper 等转存到自建存储中。
- 提供各种关键事件的管理能力。

2.5. 附录：基础术语

AKE 云原生平台（AKE3）

AKE3 （Ant Financial Kubernetes Engine）是在开源 Kubernets 和 Containerd 基础上蚂蚁自研的容器云平台，通过蚂蚁自研组件进行了能力增强。对外提供统一的计算、存储、网络资源抽象，屏蔽物理机和虚拟机等各种环境的差异。

Captain Plus

云应用引擎基础版的控制台，通过可视化界面满足日常运维需求。

Overlay 网络

采用软件定义的一层虚拟化网络，提供了更高定制化的网络，更加灵活可配，可显著降低对二层网络地址的损耗。缺点是网络性能有一定的损失。

VLAN 网络

VLAN 网络为每个容器提供一个或多个独立 IP，用于二三层可达的网路访问，比较适合传统运维模式。VLAN 的网段配置，包含 subnet、gateway 和 VLAN ID，如果有可用 IP 限制需配置 IP 范围。对于物理 VLAN 网络场景，如果交付的多个 VLAN 网段配置在不同交换机上，还必须提供交换机、网络、物理机列表的映射关系。

保密字典（Secret）

Kubernetes 的原生概念，用于存储用户的加密内容。

部署（Deployment）

Kubernetes 的原生概念，表示无状态的应用集合。

调度器

AKE3 在社区 CPU manager 之上增强的调度，可提供共享、独占以及共享绑核的能力。

工作负载（Workload）

工作负载指应用程序运行态的载体及其上层聚合，通常包括部署（Deployments）、有状态副本集（StatefulSet）、守护进程集（DaemonSet）、任务（Job）、容器组（Pod）。

集群（Cluster）

集群将多台机器的资源整合起来，统一调度，统一管理，是承载部署应用的载体。

节点（Node）

节点指用于部署和管理容器的服务器（可以是虚拟机实例或者物理服务器），该服务器会被注册到集群中，为集群提供相应的计算，网络，存储资源。

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。

命名空间（Namespace）

命名空间为 Kubernetes 集群提供虚拟的隔离作用。Kubernetes 集群初始有 3 个命名空间，分别是默认命名空间 default、系统命名空间 kube-system 和 kube-public ，除此以外，管理员可以创建新的命名空间以满足需求。

任务（Job）

Kubernetes 的原生概念，表示一次性执行的任务。

容器组（Pod）

Kubernetes 的原生概念，表示一个应用容器。

守护进程集（Daemonset）

Kubernetes 的原生概念，表示运行在所有节点上的守护应用。

网络（Network）

定义了集群中应用容器互相访问的网络配置。

网络控制器

AKE3 中用来实现网络能力的控制，可以提供网络分配调度，和网络插件一起配合可以满足网络隔离需求。

网络、存储、日志、监控插件

运行于节点之上的，提供相应网络、存储、日志、监控等功能的插件。

业务集群

业务集群与元集群同等地位，且创建方式一样。通过 cluster 控制器管理生命周期，业务集群的控制容器均独立运行在业务集群中节点。

有状态副本集（Statefulset）

Kubernetes 的原生概念，表示有状态的应用集。

元集群（Meta cluster）

在一个网络环境内的第一个集群，通过 Static Pod 部署控制组件。

云游（Yunyou）

云游是一个 Web 应用，主要提供的功能包括产品发布和管理、权限管理等。

3.AntStack DNS

3.1. 什么是 AntStack DNS

AntStack DNS（ADNS）为 AntStack Plus 平台上的内部应用提供稳定高可用的 DNS 域名注册和解析服务。ADNS 还与 AKE 云原生平台（AKE3）集成，为 AKE3 集群提供服务发现的能力。此外，ADNS 封装了若干集群治理和运维的能力，提高了 DNS 的整体稳定性和可运维能力。

3.2. 产品优势

AntStack DNS 提供了一套较完备的 DNS 管控和高可用的解决方案，具备以下优势：

- 高可用能力保障

  支持一键主备 DNS master 切换、DNS 数据一致性巡检和自动弥补。自带虚拟路由冗余协议 VRRP（Virtual Router Redundancy Protocol）高可用能力。

- 简化了交付和运维成本

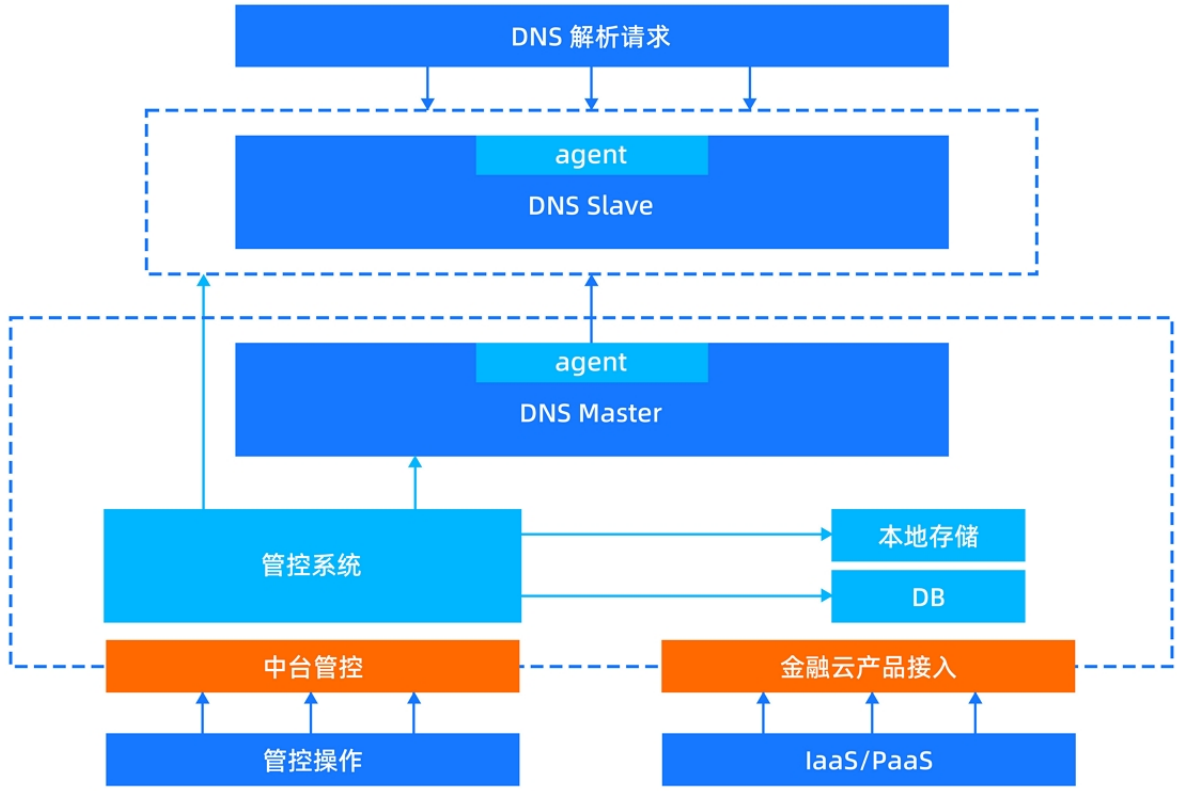
  提供 DNS zone 和记录等资源的变配，方便运维和管理。

3.3. 产品架构

本文介绍 AntStack DNS 的产品架构。

逻辑架构

管控系统管理 DNS 集群（包括 DNS Master 和 DNS Slave）对外提供 DNS 变配和管理服务。DNS 的变配会写入 DNS Master，并异步落库。同时 DNS Slave 会根据区域传输协议从 DNS Master 同步数据，对外提供解析服务。根据不同底座，DNS Master 和 Slave 通过 VRRP 或四层负载均衡提供高可用接入能力。逻辑架构如下：



DNS Master

多个 Master 中，只有一个 active，其余的 Master 为 standby。在物理机模式，VIP 会绑定在 active 的 Master 上。active Master 定时做以下动作：

- 将和 DB 中的 zone 相关的数据同步至本地。
- 将 active Master 的 DNS 缓存同步至 DB。
- 将 DB 中的数据与 standby Master 的 DNS 配置进行同步。
- 将 active Master 的 DNS 缓存与 Slave 的 DNS 通过区域传输协议同步数据。

DNS Slave

Slave 提供 RPC 接口，由 DNS Master 调用接口同步 DNS 数据。

Master 和 Slave 主要功能是同步数据以及生成 zone 配置文件，DNS解析服务由 bind9 提供。

- 通过 DNS API 添加 RR 记录时，会通过 bind9 的服务端口动态更新 RR，再异步写入 DB。
- 通过 DNS API 添加 zone 时，则是先同步写 DB，再调用各个 DNS 的 RPC 接口添加 zone。

在数据同步的过程中，当 DB 中的 zone 与本地的 zone 不一致时（例如 DB 中有 zone A，本地没有 zone A），以 DB 中的 zone 作为数据基准。当 DB 中的某个 zone 的 RR 记录与本地的不一致，则以本地的 RR 记录作为数据基准。

内部调用链路

Master 通过 HTTP 接口提供 DNS 数据的增删改查、主备切换以及节点加入与移除等功能。

- DNS 的增删改查：Master 会调用 Slave 的 RPC 接口动态更新 DNS 数据。
- 主备切换：更新 DB 中的数据，同步 standby Master 的数据。
- 节点加入与移除：更新 DB 中的数据，调用所有 Master 的 sync-ha-state 接口。

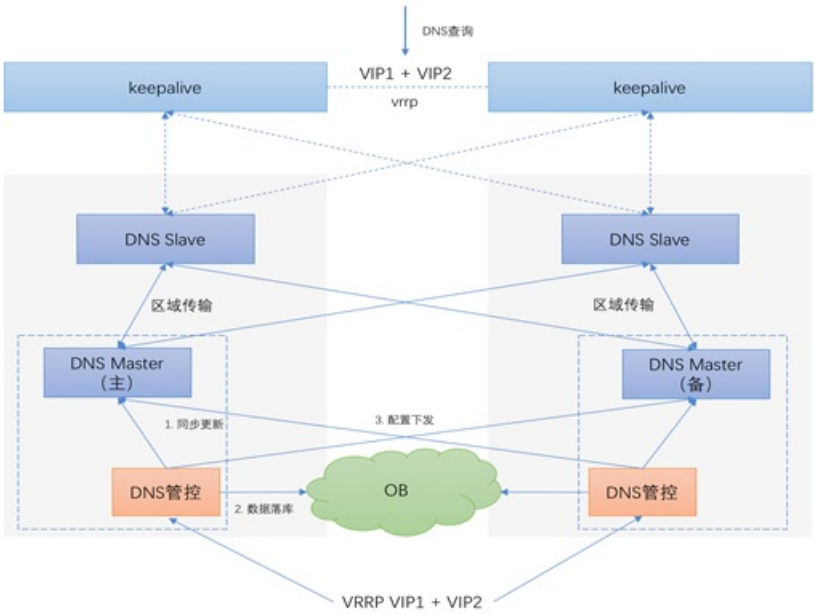
产品依赖

ADNS 与其他产品没有依赖，也没有部署依赖。

部署架构

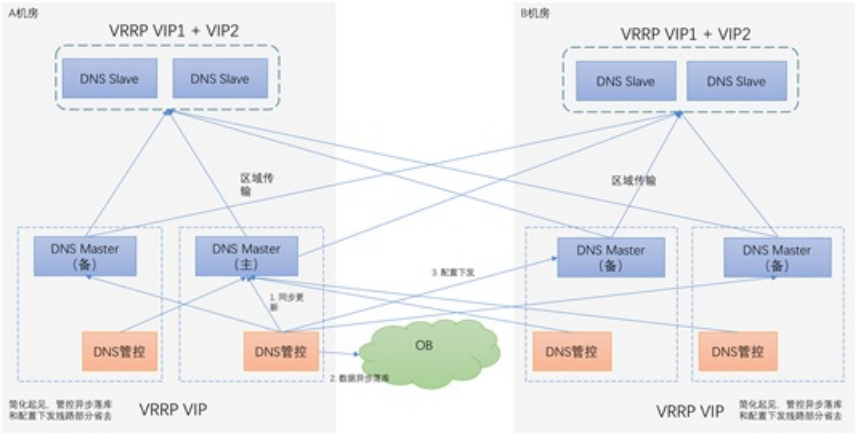
- 单机房部署

  单机房部署架构如下：



- 双机房部署

  双机房部署架构如下：



3.4. 功能原理

本文介绍 AntStack DNS（ADNS） 支持的功能和实现原理。

支持功能	原理说明

支持功能	原理说明
域名注册	ADNS 在保持向前兼容 Knocker 和 OCT 的同时，还提供了更多的 API，方便用户对域名进行增、删、改、查。域名的注册生效在秒级。
zone 管理	为 DNS 集群添加或删除 zone。支持 forward zone 和 hosted zone。
一键切换主备	提供一键切换 DNS master 主备的能力，减少了繁杂的 DNS 手工运维工作，提高了应急容灾的能力。
集群管理	集群内 DNS 节点的管理（增删）。
数据、配置巡检	定时巡检 master、database、slave 上的数据和配置，根据巡检结果同步数据和配置，保证整个集群的数据、配置一致性。
高可用能力	ADNS 自带基于 VRRP 的高可用解决方案，满足输出场景下无四层负载均衡时对 DNS 服务高可用能力的要求。

3.5. 附录：基础术语

本文介绍 AntStack DNS（ADNS） 的基础术语。

DNS

域名解析系统DNS（Domain Name System）提供主机名字和 IP 地址间的转换，是一个具有树状层次结构的联机分布式数据库。

域名

域名是一个易于记忆的名称，用于对某一互联网资源进行分配。例如 `alipay.com` 就属于一个域名。

A记录、AAAA记录

代表主机名称和 IP 地址的对应关系，即将名称转换为 IP 地址。

- A 记录是对一个 IPv4 地址指定一个域名。
- AAAA 记录是对一个 IPv6 地址指定一个域名。

zone

DNS zone 也称 DNS 区域，是为了便于根据实际情况来分散 DNS 名称管理工作的负荷，将 DNS 名称空间划分为 zone 来进行管理。例如 `www.alipay.com` 这个域名就在 `alipay.com` 这个 zone 内。

TTL

DNS 记录的生存时间，表示被外部 DNS 服务器缓存的时间总和。以秒为单位。

4.身份访问 IAM

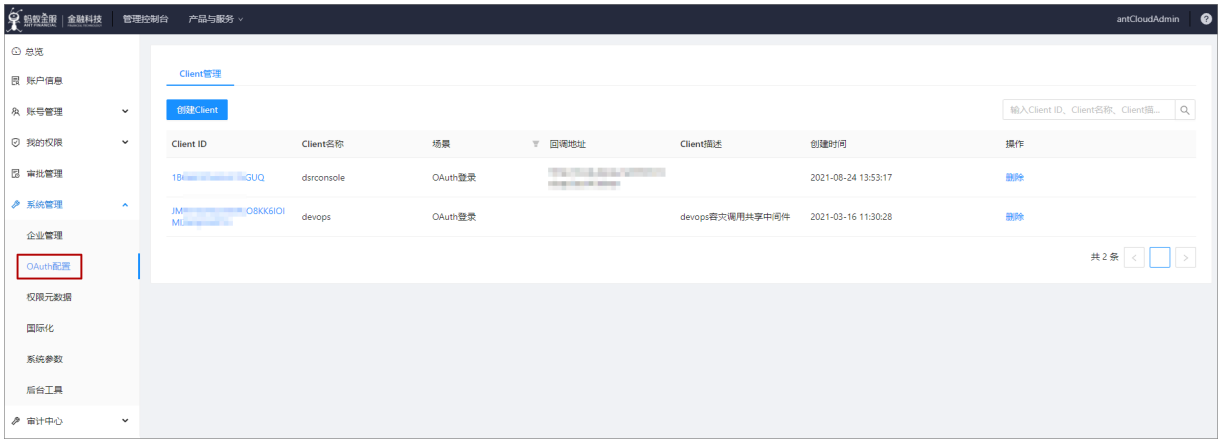
4.1. 什么是 IAM

IAM（Identity and Access Management）是基于蚂蚁集团多年发展，结合业内其他厂商的方案，推出的一套通用、灵活的身份管理、认证及访问控制解决方案。

IAM 包括应用集成、身份管理、身份认证、权限控制、透明审计、身份审批等功能：

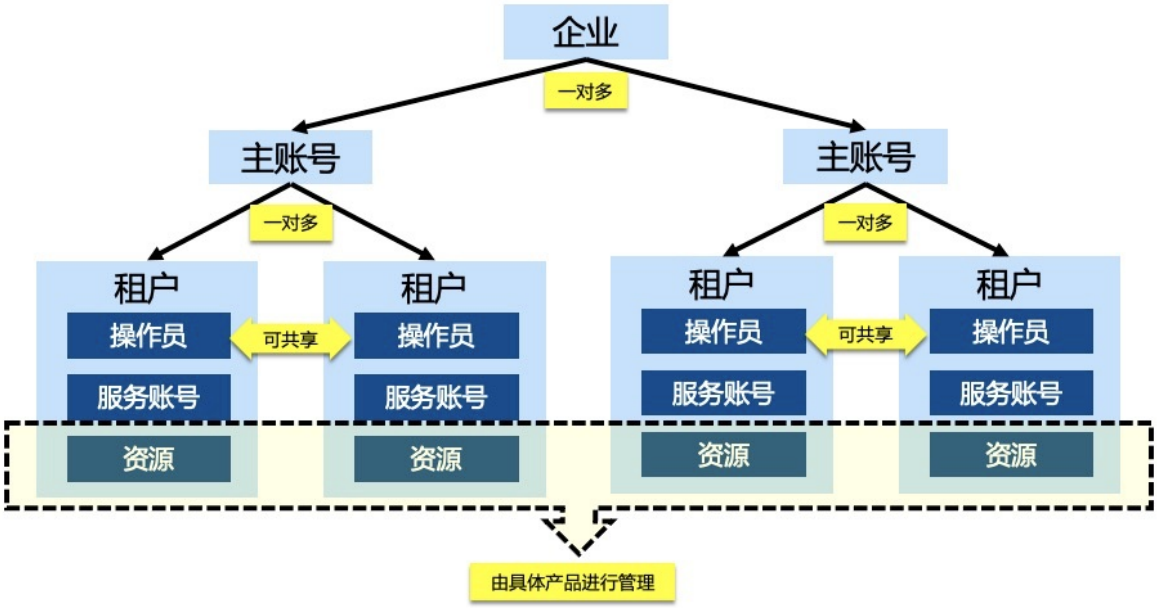
应用集成

为帮助产品更便捷地接入，IAM 提供基于标准 OAuth2.0 的接入方案，同时提供了产品化的界面对 OAuth 客户端进行统一管理。



身份管理

IAM 提供蚂蚁产品统一的身份管理能力，可以实现用户身份企业隔离，且在企业下多租户共享。当企业下多个 BU 需要独享使用该产品，但使用者的身份有重叠穿插时，产品的各种功能或资源还可进行租户隔离。



相关概念介绍：

- **企业**：IAM 最上层概念，可以理解为企业基本信息。

- **主账号**：主账号可用于登录平台，具有 root 权限，可作为独立的结算个体。一个企业下可以有多个主账号。
- **租户**：IAM 支持多租户模式的核心概念，一个企业下可对应多个租户，每个租户有对应的一个主账号。
- **操作员**：IAM 用户类型的一种。由自然人持有，可用于登录平台。企业下可创建多个操作员，操作员可被添加到企业下的不同租户中。
- **服务账号**：IAM 用户类型的一种，由应用或系统持有。服务账号拥有 AccessKey，可用于调用 OpenAPI。
- **资源**：金融云产品所管理的业务资源，例如发布部署平台的应用服务、中间件平台的消息 Topic 等。通常由具体产品自行管理，并且通过 IAM 提供的租户概念实现租户隔离。

身份认证

- 自有身份认证

  IAM 支持多种自有身份类型，不同地身份类型都有对应的认证方式。

  - 基于帐密的产品控制台访问认证。
  - 基于 AccessKey 的 OpenAPI 访问认证。
  - 基于 STS Token 的虚拟身份认证。
  - MFA（Multi-Factor Authentication）双因子校验能力。 支持短信校验码二次验证和多种客户端，包括谷歌验证器（Google Authenticator）在内的令牌二次验证能力，企业可灵活管控，支持管理员强制企业下全部用户开启和个人自定义是否开启。

- 密码管理策略

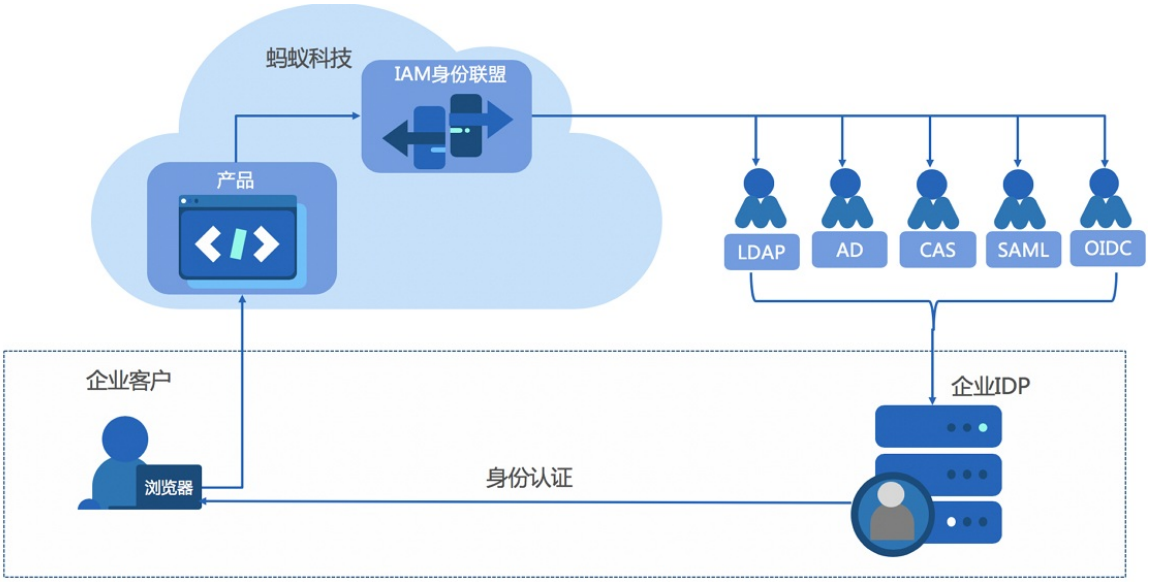
  企业的管理员可以根据需求设置该企业下用户的密码策略，可灵活应对包括 PCI 等在内的认证需求，包括但不限于以下策略：

  - 密码长度控制。
  - 密码字符组成控制。
  - 密码有效期控制。
  - 密码重试次数控制。
  - 过期重置密码，与历史密码校验控制。

- 身份联盟

  IAM 旨在建立与客户的身份联盟体系，不仅可以作为 IDP，提供基于 OAuth 的三方认证能力，便于客户系统对接。同时也可以以 SP 身份，对接客户标准的认证系统，目前支持的认证协议有 CAS、SAM、OIDC、LDAP、AD。

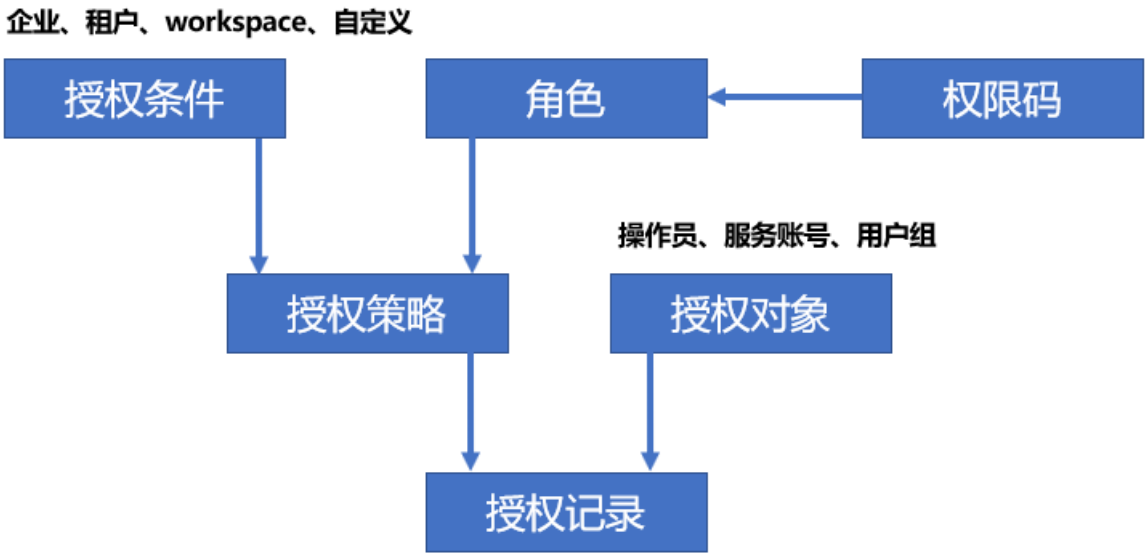
如果客户不是标准的协议，又迫切希望我们复用其身份体系，IAM 提供了基于 SPI 形式的认证方式，客户主要按照我们的标准实现相关接口，便可实现与客户的身份打通。



权限控制

IAM 提供基于角色和条件的访问控制。角色是权限码的集合，便于将一批权限授予某个对象，权限码对应于页面的一个功能或一个对外暴露的 API，角色和权限码做垂直权限控制。

条件根据产品需求自由定制，做水平权限控制，例如租户，应用等，任何一个产品提供给用户的资源或概念都可以作为条件。同时提供组的管理能力，多个账号可加入到同一个组，对该组进行授权，则组内的成员便都将具有对应权限。



- 权限码：操作或权限，对应一个原子的权限点。例如页面的一个功能或一个对外暴露的 API。
- 角色：权限码的集合，方便您将一批权限授予某个授权实体。角色和权限码做垂直权限控制。
- 授权条件：权限生效的水平条件。IAM 提供了企业、租户的通用条件隔离能力，进行基于租户维度的水平权限校验。当产品有需求做更细粒度的水平权限控制，可在 IAM 定义条件元数据，例如工作空间 worksapce。
- 授权策略：授权策略是垂直权限和水平权限的模型载体。

- 授权对象：IAM 支持的所有身份类型都是授权对象，例如操作员账号、服务账号、用户组等。
- 授权记录：建立授权策略与授权对象的绑定关系。

透明审计

为了帮助外部银行站点满足监管合规的要求，IAM 提供包含以下基本能力的审计中心产品能力：

- **事件上报**：支持包含 WebAPI、OpenAPI 多种来源的事件上报能力。
- **归档存储**：自持 OSS 或者自定义服务器的归档存储能力。
- **消息订阅**：支持短信、邮件、钉钉多种渠道的消息订阅通知能力。
- **多维查询**：支持 AccessKey、操作人、资源、来源 IP 等多种维度的查询能力。

申请审批

针对的不同业务场景，IAM 支持以下多种事件类型的申请审批能力：

- 账号申请
- 权限申请
- 服务账号申请
- 用户组申请
- 加入租户申请

4.2. 产品优势

身份访问 IAM（Identity and Access Management）是基于蚂蚁集团的多年发展，结合业内其他厂商（AWS、阿里云、GoogleCloud 等）的方案，推出的一套通用、灵活的身份管理、认证及访问控制解决方案。

若您的产品可快速对接 IAM 的标准接口，便可以快速具备身份、权限和认证管理能力。同时，针对诸多机构已有的身份认证体系，IAM 提供基于标准协议的对接方式，例如 Oauth、LDAP 及 AD 等。

4.3. 功能原理

本文介绍身份访问 IAM 的功支持的功能和原理说明。

身份管理（Identity Manager）

- 多企业多租户的用户身份管理体系

  用户身份企业维度隔离，且在企业下多租户共享，适用于云上、云下的各种使用场景。对接 IAM 的一套产品同时服务于多个企业时，每个企业维护各自的用户身份信息，当企业下多个业务单元需要独享使用该产品，但使用者的身份有重叠穿插时，产品的各种功能或资源可通过租户隔离，同一个企业下的用户身份租户间共享。

- 多种用户身份类型
  - 操作员：拥有登录 ID 和登录密码的用户，用于访问 IAM 或对接产品的控制台。由自然人持有。
  - 服务账号：拥有一对 AccessKey，用于访问 IAM 或对接产品提供的标准 API 接口，由 App 或 System 持有。
  - 虚拟身份：虚拟身份可授信与其他企业租户或租户产品扮演，拥有一对临时 AccessKey。扮演改虚拟身份的用户可通过 API 访问虚拟身份所在企业的资源。

- 灵活多变的密码管理策略

  企业的管理员可以根据需求设置该企业下用户的密码策略，可灵活应对包括 PCI 等在内的认证需求，包括但不限于以下策略：

  - 密码长度控制。
  - 密码字符组成控制。
  - 密码有效期控制。
  - 密码重试次数控制。
  - 过期重置密码与历史密码校验控制。

身份认证（Identity Authenticator）

- 基于帐密或 AccessKey 的身份认证

  提供不同类型的认证方式，满足控制台访问和 API 调用的不同访问方式的身份认证。

- SSO（Single Sign On）

  一个系统登录，可访问相互信任的其他产品系统，实现单点登录。

- 标准 Oauth 三方认证

  基于标准的 Oauth 协议开放的三方认证能力，便于第三方应用集成对接时，共用一个身份体系。

- MFA（Multi-Factor Authentication）

  支持短信校验码二次验证和多种客户端（包括谷歌验证器 Google Authenticator）的令牌二次验证能力，具有灵活的管控能力。支持管理员强制企业下全部用户开启和个人自定义是否开启。

访问控制（Access Manager）

- 页面访问控制（WEB Access Control）

  提供标准 Restful 前端接口的访问控制能力。针对个别技术栈（例如 sofa-mvc）提供 SDK 快速接口，集成认证和权限控制能力。

- API 访问控制（API Access Control）

  基于 AccessKey 进行身份认证，对于开放的 API 接口，提供访问控制能力。如使用了金融云网关，还提供集成 SDK，可以方便的将一个后台服务以标准化的网关协议对外暴露，且具有访问控制。

- CLI 访问控制

  命令行操作的权限控制，支持基于 k8s 的权限控制方案。

- R&CBAC（Role&Condition Based Access Control）

  基于角色和条件的访问控制。

  - 角色是权限码的集合，便于将一批权限授予某个对象；权限码对应于页面的一个功能或一个对外暴露的 API，角色和权限码进行垂直权限控制。
  - 条件可根据产品需求自由定制，实现水平权限控制。例如租户、应用等任何一个产品提供给用户的资源或概念都可以作为条件。同时提供组的管理能力，可将多个身份可加入到同一个组，并对该组进行授权，则组内的成员便都将具有对应权限。