弹性 MapReduce

产品简介

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】

腾讯云

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

产品简介
- 产品概述
- 产品优势
- 产品架构
- 产品功能
- 应用场景
- 约束与限制
- 技术支持范围
- 产品发行版
  - 版本概述
  - 组件版本概览
  - 组件与 API 部署时映射关系

版权所有：腾讯云计算（北京）有限责任公司

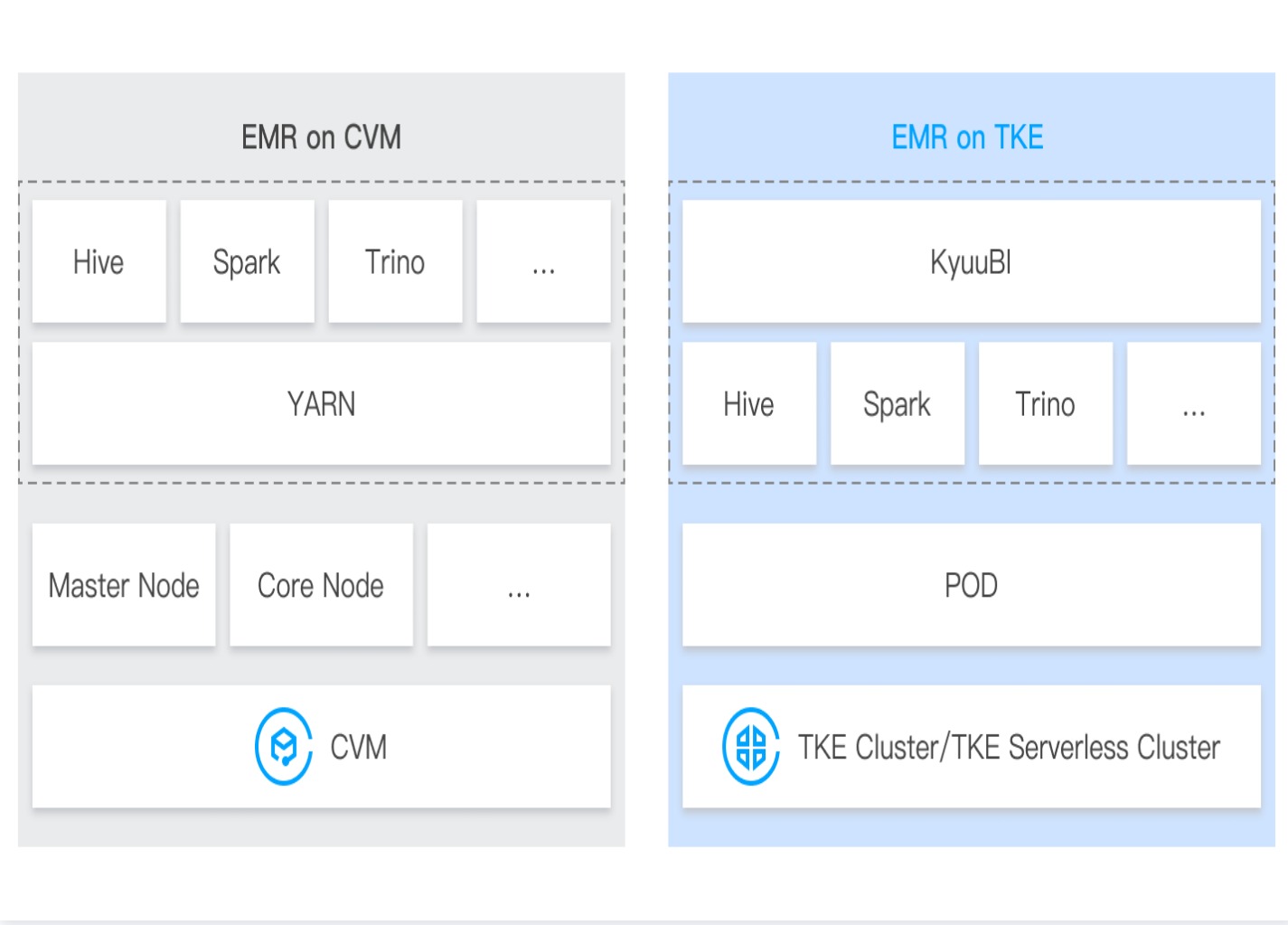
产品简介
产品概述

最近更新时间：2024-09-05 09:48:11

弹性 MapReduce (EMR) 是基于云原生技术和泛 Hadoop 生态开源技术的安全、低成本、高可靠的开源大数据平台。提供易于部署及管理的 Hive、Spark、HBase、Flink、StarRocks、Iceberg、Alluxio 等开源大数据组件，帮助客户高效构建云端企业级数据湖技术架构。EMR 支持基于 CVM 和 TKE 两种资源部署。

支持形态

腾讯云 EMR 提供基于云服务器（CVM）和容器服务（TKE）两种部署运行方式：



形态	描述
EMR on	EMR 负责将开源大数据组件安装部署在 CVM 上，并启动相应的服务。您可以通过 EMR 控制台完成对集群 CVM 及服务的运维操作。

CVM	
EMR on TKE	如您已准备好一个TKE标准集群或 Serverless 集群，EMR 将基于 TKE 的资源安装部署开源大数据组件，实现开源大数据平台的容器化运行，您可减少对于底层资源的运维关注。

相关云产品

服务名称	说明
云服务器（CVM）	EMR on CVM 产品形态下，CVM 实例用于 EMR 集群节点。
云硬盘（CBS）	CVM 节点和 POD 节点根据需要可搭配不同规格的云硬盘进行数据存储。
容器服务（TKE）	EMR on TKE 产品形态下，可选择 TKE 集群中的 POD 作为 EMR 集群节点。
云数据库 MySQL（TencentDB for MySQL）	EMR 集群部署了 Hive、Hue、Ranger 等组件时，集群将同步购买一个云数据库 MySQL用于存储组件元数据。
对象存储（COS）	存算分离场景可以使用 EMR 读写对象存储中的数据。
云HDFS（Cloud HDFS，CHDFS）	存算分离场景可以使用 EMR 读写云 HDFS 的数据。
负载均衡（Cloud Load Balancer，CLB）	EMR on TKE 产品形态下，部分服务可以设置负载均衡提供对外访问。
访问管理（CAM）	通过访问管理对 EMR 产品授予服务角色权限或对协作者/子用户进行操作授权。
腾讯云可观测平台（TCOP）	EMR集群的指标类和事件类监控数据会上报至 TCOP，通过 TCOP 可以拉取集群监控数据并对需要重点观测指标配置告警通知策略。

版权所有：腾讯云计算（北京）有限责任公司

产品优势

最近更新时间：2024-09-10 16:48:11

腾讯云 EMR 提供了易于部署和管理的企业级开源大数据服务，可以快速搭建 Hadoop、Spark、HBase、Trino、StarRocks 等开源大数据服务。相较于自建开源大数据平台有以下产品优势：

丰富可靠的开源组件

- **组件丰富**：提供高性能、高稳定性、按需灵活搭配的 Hive、Spark、Presto、StarRocks、HBase、Flink、Iceberg、Alluxio 等丰富开源大数据组件。
- **持续迭代**：随开源版本升级迭代，适配开源组件，避免开源组件之间的版本兼容性问题。
- **开源增强**：基于开源组件深度优化，提供 Alluxio 透明加速、Iceberg Z-Order 算法等优化技术提升性能。

易于部署运维

- **部署方便**：只需几分钟即可基于 CVM 或 TKE 构建一个开源大数据集群。
- **运维便捷**：支持基于时间和负载的自动化容量管理、可视化集群参数配置、支持资源调度以及联邦等应用级策略配置。
- **监控丰富**：支持从资源到服务运行情况的全面监控，可从运行指标趋势、重点事件监控、日志搜索等功能快速诊断集群基础运行问题。
- **应用分析**：支持 HDFS、YARN、Hive、HBase、Impala 等重点服务应用级分析，提升应用级问题定位效率。

成本节约

- **资源弹性**：按需购买、根据业务特点自动伸缩集群，减少资源闲置成本。
- **集群联邦**：结合统一 Hive 元数据库以及统一对象存储，实现跨集群的同数据集分析架构，集群按需创建或销毁，灵活节省集群成本。
- **存算分离架构**：计算资源和存储资源分开购买，且可根据访问频率选择不同存储方案，降低存储和计算成本；支持温冷数据的对象存储 COS/CHDFS 存储，成本有效降低28% - 50%。
- **在离线混合部署**：支持基于容器服务 TKE 部署，错峰复用算力，降低资源成本。

安全可靠

- **网络安全**：VPC 网络隔离和安全组保证网络安全可信。
- **访问安全**：提供集群级别的 Kerberos 认证，保障集群访问安全；支持基于 Ranger 对本地及 COS 数据细粒度权限管控。
- **防护服务**：腾讯云安全加固服务为 EMR 集群提供一体化的安全服务，涵盖网络防护、入侵检测、漏洞防护等。
- **容灾架构**：Master 节点容灾设计，备节点秒级拉起，保障大数据服务可用性。

版权所有：腾讯云计算（北京）有限责任公司

- **存储可靠**：支持将 Hive 元数据存放于 MetaDB，元数据可靠性达99.9996%；支持分析存放于 COS 的高存储耐久性的 PB 级数据。

版权所有：腾讯云计算（北京）有限责任公司

产品架构

最近更新时间：2024-09-10 16:48:11

弹性 MapReduce 产品逻辑架构如图所示：

EMR 主要由开源组件、腾讯云基础设施和集群管理三部分构成：

- 开源组件
  - 集成 Apache 社区 Hadoop、Hive、Spark、Hbase、Presto、Flink、Alluxio、Iceberg 等几十种丰富、前沿的开源大数据组件（详见 组件版本概览 ）。
  - 基于Iceberg、Alluxio 等开源组件优化，提供 Iceberg Z-Order 算法、Alluxio 透明 URI 等性能及功能增强。
- 腾讯云基础设施

  - 可基于云服务器（CVM）、裸金属云服务器等多种底层计算资源部署，支持容器化部署。
  - 数据可存储在本地盘、云硬盘，或腾讯云对象存储（COS）、云 HDFS 服务上。
  - 私有网络（VPC）、网络 ACL 、安全组为 EMR 提供安全隔离的网络环境。
- 集群管理
  - 快速创建、灵活扩缩、自动伸缩的云端智能部署管理。
  - 服务配置管理、批量节点管理、服务运维可视化操作等丰富便捷的运维工具。
  - 多维度指标监控、事件、巡检、告警、日志搜索等完善的集群监控诊断能力。

版权所有：腾讯云计算（北京）有限责任公司

产品功能

最近更新时间：2024-08-15 14:44:01

弹性 MapReduce 结合云技术和 Hive、Spark、Presto、StarRocks、HBase、Flink、Iceberg 等社区开源技术，为您提供安全、低成本、高可靠、可弹性伸缩的云端 Hadoop 服务。其主要功能体现在以下方面：

集群管理

通过腾讯云弹性 MapReduce 服务可以简单高效的配置和管理您的开源大数据集群，能够无缝衔接云上基础设施服务。

模块	功能点
配置集群	• 支持超过30+开源大数据组件，灵活按需部署 • 支持基于云服务器 CVM 或容器服务 TKE 部署开源大数据组件 • 支持通过引导操作部署第三方应用组件 • 支持提前预设关键软件配置参数 • Hive 元数据库支持关联已有元数据库实现多集群共享元数据库 • 支持设置对象存储或云 HDFS 存放业务数据
管理集群	• 支持按需升级节点硬件配置 • 按需扩缩 Task 节点或 Router 节点 • 支持根据业务运行时间特点和负载特点自动伸缩 Task 节点 • 本地盘换盘后可以便捷修复后恢复使用 • 支持自动探测 Task 节点或 Router 节点异常，并自动更换出现异常的节点 • 支持扩容用作数据盘的云硬盘 • 支持通过集群脚本对多个节点进行统一运维操作

服务管理

服务管理为已部署的开源大数据组件提供了基础运维操作，同时支持部分重点组件的高阶运维和服务策略配置管理。

模块	功能点
服务基础运维	• 服务角色状态以及启停管理 • 配置管理调整服务参数配置 • 支持查看服务原生 WebUI • 支持查看服务客户端配置列表
服务高阶运维	• 支持 HDFS 主备切换及数据均衡可视化操作 • 支持 YARN 主备切换及刷新队列可视化操作

	• 支持 HBase RIT 修复
服务策略配置	• 支持设置 HDFS 联邦管理策略 • 支持设置 YARN 资源调度策略 • 支持 OpenLDAP 用户管理

监控告警

Recd弹性 MapReduce 为集群中节点和服务提供全面的监控告警服务，特别的对于部分核心组件提供应用层分析，以便您更高效的诊断应用问题。

模块	功能点
基础监控	• 支持查看节点和服务运行指标，并可通过 API 获取指标 • 支持节点和服务事件监控 • 支持系统运行事件监控和配置事件监控策略 • 支持服务角色运行日志搜索 • 支持主动巡检集群运行情况
应用分析	• HDFS 支持存储文件分析 • YARN支持作业查询，并提供 MR、Tez 、Spark 类型作业运行情况洞察 • Hive 支持查询管理和数据表分析 • Impala 支持查询管理 • HBase 支持数据表分析 • Kudu 支持数据表分析
告警管理	• 支持默认指标以及时间告警策略 • 支持设置节点和服务运行指标变动告警 • 支持设置节点和服务事件告警

版权所有：腾讯云计算（北京）有限责任公司

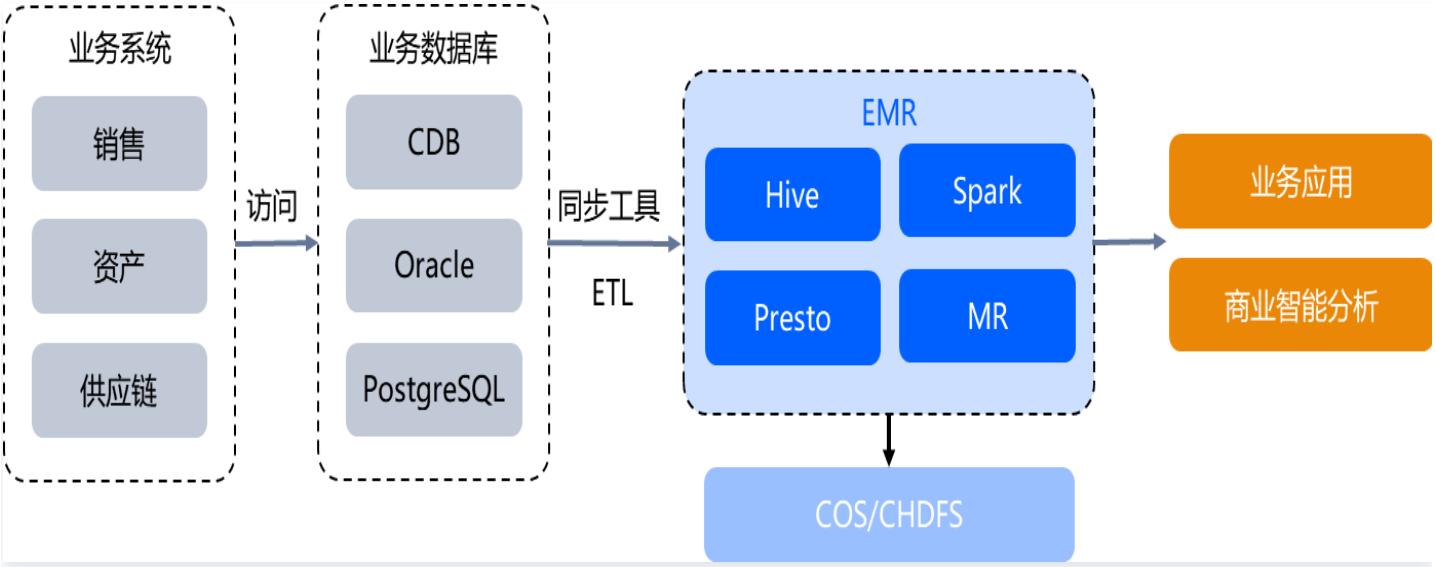
应用场景

最近更新时间：2024-09-10 16:48:11

腾讯云 EMR 支持开源组件丰富，应用场景广泛，本文为您介绍 EMR 的主要应用场景。

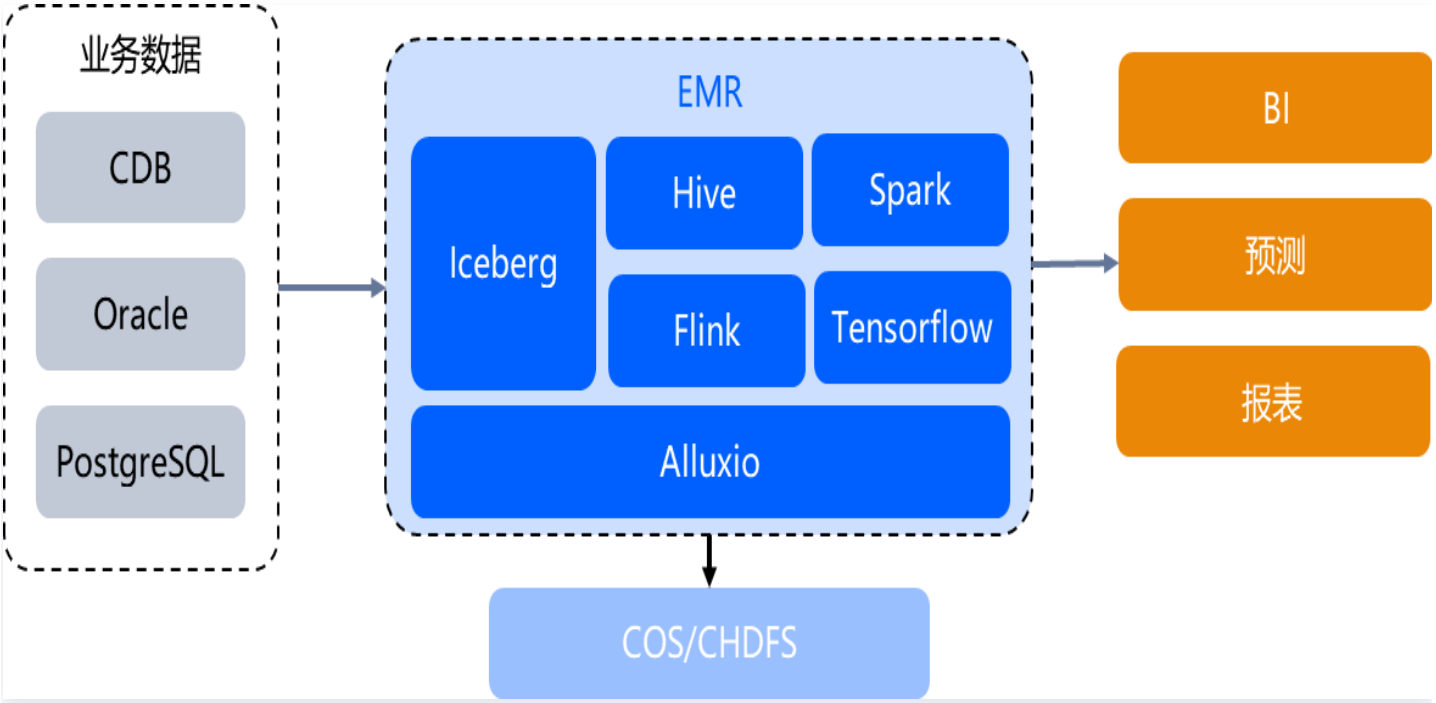
企业级数据仓库构建

对销售、资产、供应链等业务数据进行汇总分析，需要结合不同数据源，提取不同源的数据、然后借助 EMR 强大的 PB 级数据分析能力，以及原生支持腾讯云COS、CHDFS 存储，提供高性能存算分离数仓方案，对海量数据进行计算，发现数据中隐藏的商业价值，进行业务决策。



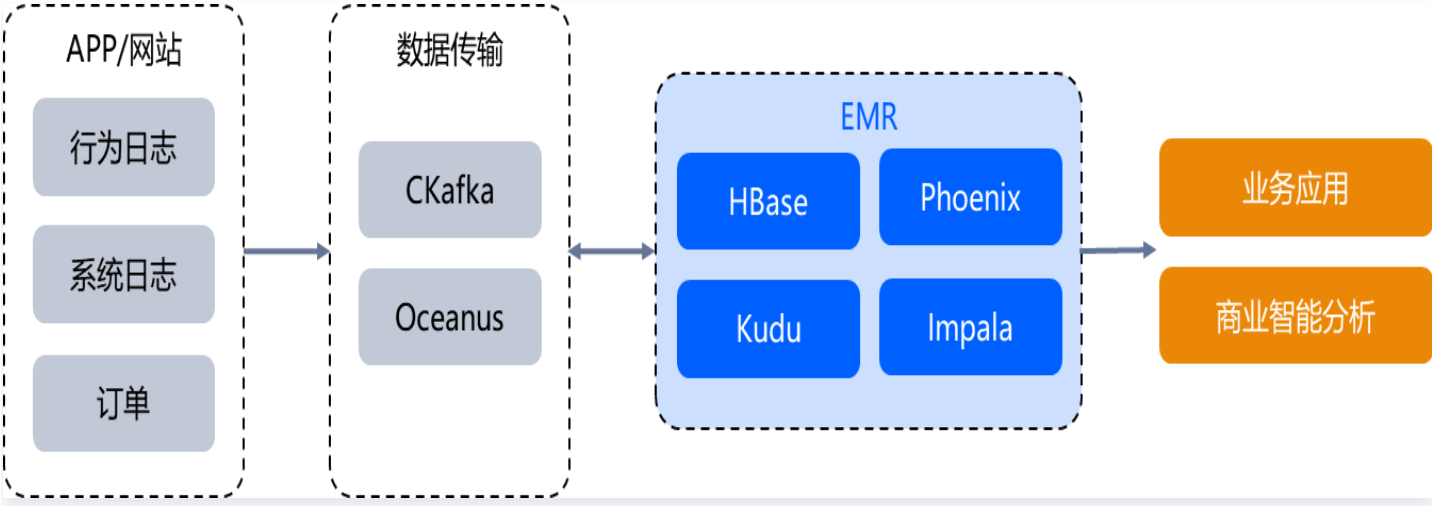
企业数据湖构建

在企业不断积累业务发展的全量数据领域，需要将各种类型的数据存储并适配多种场景的数据分析任务，借助 EMR 提供的数据湖格式及缓存加速能力构建数据湖，可以充分利用各种资源并适配如离线计算、流式计算、交互式分析、机器学习等场景，赋予客户更高的数据敏捷度、更低的数据分析成本。



高并发在线数据查询

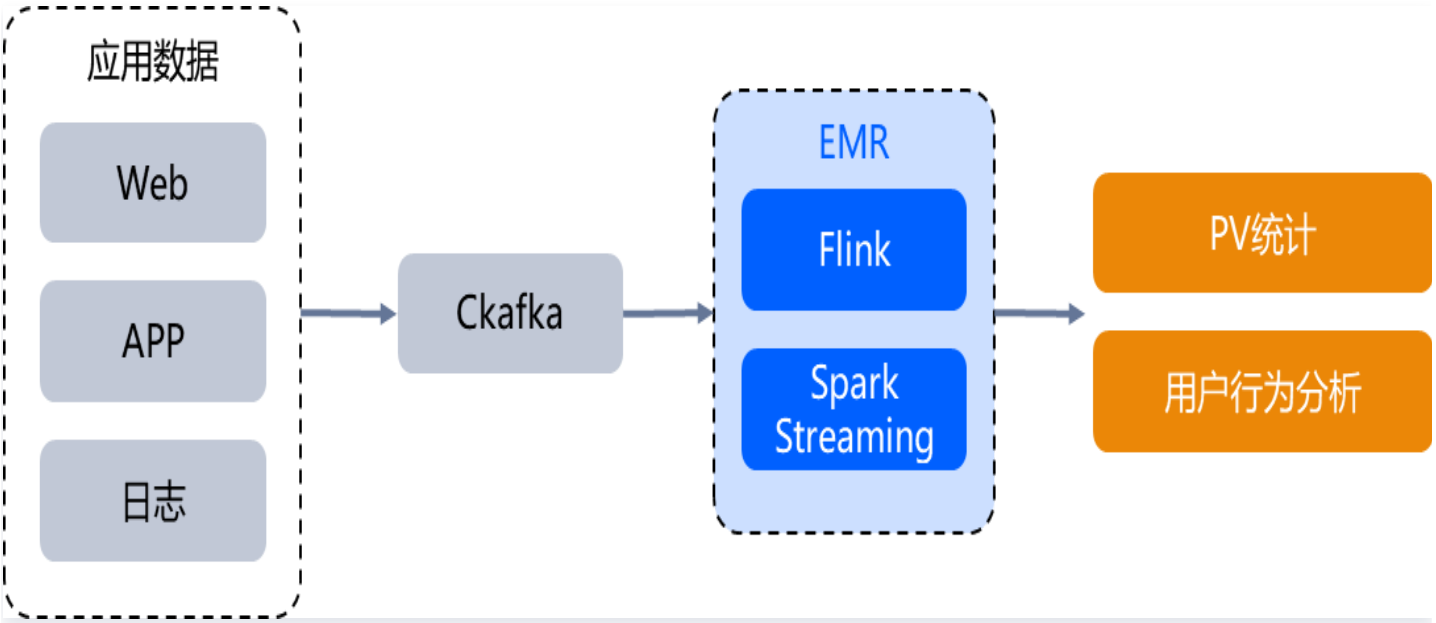
对用户行为、系统日志、订单等结构化或半结构化数据进行高效分析，需要收集在线网站、App、系统的用户行为及系统日志等各种业务数据，通过 EMR 丰富的计算组件及分钟级集群构建与平行扩展能力，支撑在线业务实时查询，提高业务响应效率。



实时流式数据计算

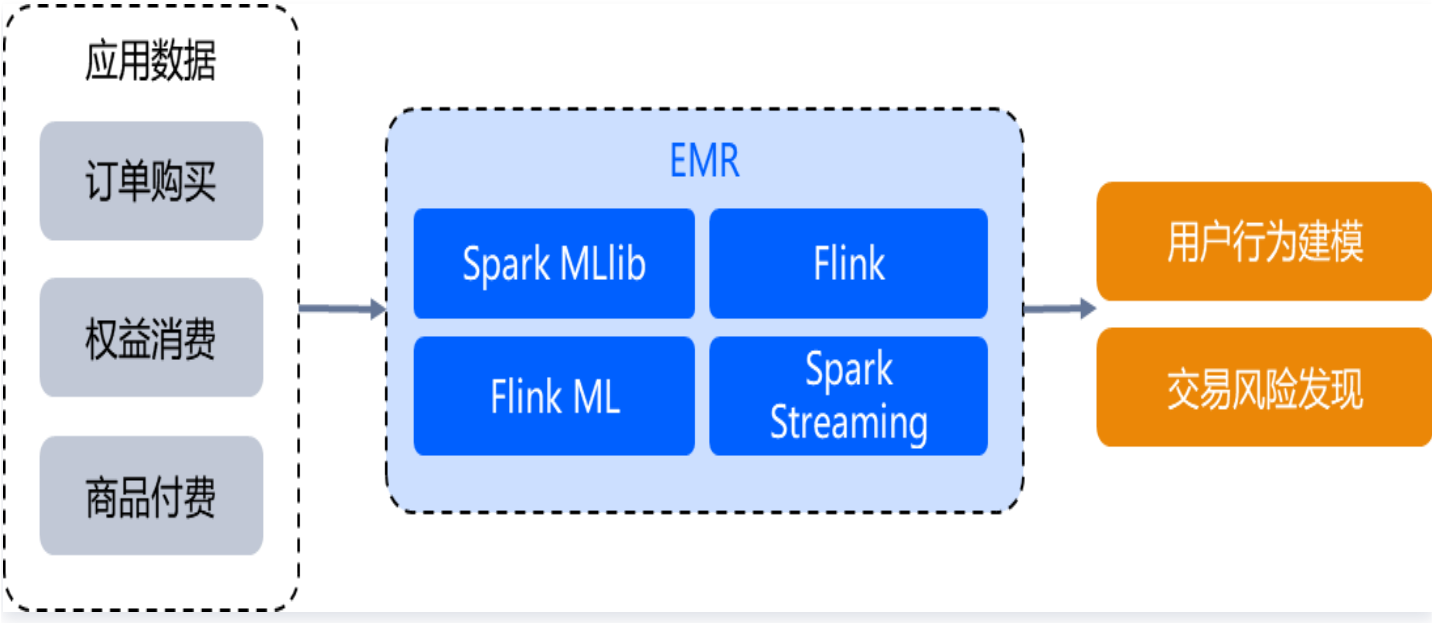
在企业构建实时计算的业务领域，借助于 EMR 云端流计算服务，能够分钟级构建实时分析，对用户行为数据进行实时汇聚分析，帮助改进用户体验。通过同时搭建批、流处理系统，实现批流一体，降低资源投入，提升数据处理速

度，及时分析业务运营效果，快速调整业务策略，支撑主流业务更好发展。



数据挖掘及分析

在实时风控、实时推荐等需要快速计算以及数据挖掘能力的业务场景，EMR 可以提供高效的云端流计算服务以及数据挖掘组件能力支撑，帮助业务监控储户异常交易，及时发现金融漏洞，确保资金安全，以及构建用户画像模型，帮助企业准确及时地了解用户群体。



约束与限制

最近更新时间：2024-04-01 14:37:51

在使用腾讯云 EMR 服务之前，请您详细阅读并了解以下使用限制：

- 新建集群为保证集群网络安全，集群将放置在同一 VPC 中，请勿随意变更已有集群或节点的 VPC，避免造成集群网络不互通。
- 在创建集群时，EMR 可以帮您创建一个新的安全组，您也可以手动选择已有 EMR 安全组。请仔细确认手动选择的安全组具备 EMR 必要的出入站规则，并在集群创建后不要随意删除或更改正在使用的安全组，避免造成集群通信异常，影响服务。
- 请根据业务需要提前规划节点的存储空间，并及时扩充存储节点，避免因存储空间不足造成数据及节点运行风险。目前 EMR 集群 Core、Task、Router 三类节点支持挂载多云盘，单个节点最多云盘累计不超过20块，黑石2.0机型和本地盘机型（IO 系列和 D 系列）集群暂不支持多云盘。
- 使用 EMR 服务时，请您尽可能避免在 CVM 控制台进行操作，如关机、重启、切换私有网络、调整安全组规则等，以免造成集群异常；OS 重装、实例销毁/变配、续费、变更计费类型等操作也会受限，您可以在 EMR 控制台进行集群维护的必要操作。
- 公网 IP 会在一定程度上增加 Master 节点被网络攻击的可能性，请您管理和监控相关风险。弹性公网 IP（含辅助网卡上的 IP）会在集群销毁后继续保留，闲置 IP 会产生费用，如不需保留，请到对应资源管理器页面进行释放。
- 在创建集群时，EMR 提供了满足通用场景的组件初始化参数，在使用组件服务前，建议您检查 HDFS/HBase 等组件参数以确保匹配您的业务场景。如需相关组件初始化指南，可以联系技术支持人员获取。
- 请您妥善保管 EMR 集群的主机登录密码。在您配置节点间访问免密登录后，腾讯云安全可能会检测到漏洞风险并对您进行提示。
- 集群出现异常状态时，会仍然持续计费，建议及时联系我方技术支持人员处理。如需登录集群排查的故障，技术支持人员会在征得您同意后请您提供账号密码。

在您使用及维护 EMR 集群时，一些非预期的操作可能会导致集群不可用或不稳定，您在控制台执行部分操作前会有相应的风险提示，本文也为您列举了一些禁止及高危操作：

禁止操作

操作	操作风险
在 CVM 中操作关机不收费	集群不可用，EMR 不支持关机不收费，费用仍会持续结算，如成本考虑可进行集群缩容或销毁
在 CVM 中修改 EMR 节点内网 IP	节点通信异常、集群不可用
在集群运行中修改 CVM 节点的安全组	节点通信异常、组件服务不可用
删除节点上已有进程/应用程序/文件	集群/组件服务不可用

删除或者修改/etc 目录下的 hosts 文件	集群关联不到节点上的服务，导致服务异常
删除或者修改 HDFS 元数据文件 edit log	导致 HDFS 集群不可用
手动修改 Hive 元数据库的数据	Hive 数据解析错误，服务异常
删除 ZooKeeper 相关数据目录	相关依赖组件无法运行

高危操作

操作	操作风险	建议
在 CVM 中对 EMR 集群的节点进行关机、重启	重启、关机导致服务不可用	确认操作必要性，并详细阅读 CVM 相关操作限制
在 CVM 控制台对 EMR 节点挂载磁盘	EMR 无法识别和初始化，导致磁盘不可用	建议通过 Core 节点扩缩实现，或在技术人员指导下进行
在 CVM 控制台对 EMR 节点卸载磁盘	会导致数据丢失或集群不可用	建议通过 Core 节点扩缩实现，或在技术人员指导下进行
直接在 CVM 上修改组件配置文件的参数	在 EMR 控制台下发配置并重启服务后，会导致在 CVM 修改的参数被覆盖	在 EMR 控制台上修改参数配置，特殊情况请在技术支持人员指导下进行
删除或者修改/etc 目录下的 resolv.conf 文件	集群关联不到节点上的服务，导致服务异常	确认操作必要性，并在技术指导下进行
修改 EMR 节点 Hostname	集群关联不到节点上的服务，导致服务异常	确认操作必要性，并在技术指导下进行
修改 MetaDB 密码	EMR 依赖 MetaDB 中配置的密码，修改后导致 Hive/Ranger 等服务不可用	在 EMR 控制台同步修改配置，并在技术人员指导下进行
修改 MetaDB 浮动 IP	EMR 依赖 MetaDB 中配置的 IP，修改后导致 Hive/Ranger 等服务不可用	在 EMR 控制台同步修改配置，并在技术人员指导下进行
修改 MetaDB 安全组	导致 MetaDB 与集群通信受阻，Hive/Ranger 等服务不可用	在技术人员指导下进行

版权所有：腾讯云计算（北京）有限责任公司

技术支持范围

最近更新时间：2024-09-10 16:48:11

腾讯云弹性 MapReduce EMR on CVM 产品形态下集群资源归属于用户，EMR 提供基于该资源的半托管云服务能力，用户对集群拥有完全的管理操作权限，集群日常运维管理由用户负责。为了更好的支持使用，现将 EMR 产品技术支持服务规范进行说明。

支持的服务

- 集群购买创建及销毁流程
  支持客户成功完成软件配置、区域与硬件配置、基础配置等购买创建全部流程，支持销毁。
- 集群扩容和缩容流程
  支持客户选择不同的节点类型，完成扩容及缩容全部流程。
- 集群配置变更流程
  支持创建成功后客户完成选择变更机型配置，单独选择云数据盘扩容变更配置流程。
- 集群服务功能
  支持在可选组件范围内新增组件，支持启停服务功能及相关管理功能。
- 集群告警监控功能
  支持客户在控制台查看集群节点运行状态，支持设置监控事件规则和巡检时间，查看告警历史，搜索日志功能。
- 集群自动伸缩功能
  支持开启或关闭自动伸缩，开启后客户可选择自定义伸缩或托管伸缩。

协助支持的服务

- 协助排查 EMR 产品开源组件的缺陷或需求，将根据产品规划迭代排期解决。包括但不限于与开源社区积极沟通、提供社区及业界验证明确可行的解决方案等，但鉴于开源组件性质，腾讯云无法承诺超越社区进度的解决方案，开源组件详情请参见 组件版本概览 。
- 协助排查 EMR 产品依赖的其他腾讯云基础产品的缺陷或需求，流转对应产品团队负责解决，依赖的其他产品如下：
  - 云服务器（CVM）、裸金属云服务器等底层计算产品。
  - 本地盘、云硬盘，或腾讯云对象存储（COS）等存储产品。
  - 用于存储元数据的云数据库 MySQL。
  - 私有网络（VPC）、网络 ACL 、安全组等网络环境。

不支持的服务

- EMR 提供服务配置管理、批量节点管理、服务运维可视化操作等丰富便捷的运维工具，并保证工具功能的可靠性和可用性，不提供具体集群和组件管理的运维操作。
- 在集群组件无明显异常或明确产品缺陷的情况下，不负责单个作业的问题排查。

版权所有：腾讯云计算（北京）有限责任公司

- 不支持非产品标准能力的服务，例如 Core 节点缩容、磁盘容量清理等。
- 不支持客户业务应用开发相关问题处理。
- 不支持用户自行安装的第三方组件问题处理。
- 不支持用户进行非产品预期操作导致集群不稳定或不可用问题处理，详情请参见 约束与限制 。

支持方式

如果您在使用弹性 MapReduce（EMR）过程中需要相关技术支持，可 提交工单 联系客服。

产品发行版
版本概述

最近更新时间：2025-04-03 10:35:02

产品发行版本号格式

1. EMR On CVM 采用 EMR-Va.b.c 格式的版本号，详细说明如下：
   1.1 在不同集群中代表的版本含义如下：
      - Hadoop 集群类型中 a 等于2为支持 Hadoop 为2.x 版本，a 等于3表示为支持 Hadoop3.x 版本。
      - Kafka 集群中 a 代表当前版本支持 Kafka 的版本，a 等于1为支持 Kafka1.x，a 等于2为支持 Kafka2.x。
      - StarRocks 集群中 a 代表当前版本支持 StarRocks 的版本，a 等于1为支持 StarRocks2.x，a 等于2为支持 StarRocks3.x。
   1.2 b 代表版本中新增组件或支持组件版本升级。
   1.3 c 代表功能优化。
2. EMR on TKE 采用 EMR-TKE-Va.b.c格式的版本号，详细说明如下：
   2.1 a 代表的版本整体有较大的变动。
   2.2 b 代表版本中新增组件或支持组件版本升级。
   2.3 c 代表功能优化。

注意

- 每一个版本上捆绑的组件和组件的版本都是固定的，不支持组件的多个不同版本的选择，也不支持用户自行更改组件的版本。例如在 EMR-V2.7.0 版本中内置的是 Hadoop 2.8.5、Spark 3.2.1 等。
- 一旦选择了 EMR 某个版本创建集群，该集群使用的 EMR 版本和组件版本不会自动升级，例如选 EMR-V2.7.0 版本，那么 Hadoop 就一直保持在2.8.5版本，Spark 就一直保持在3.2.1 版本。后续如果版本升级到了 EMR-V2.8.0 版本，Hadoop 到了更高版本，Spark 到了3.3.0版本，也不会影响已创建的集群。只有新的集群才会使用新的镜像。
- 当您通过数据迁移的方式升级集群版本时（例如，从 EMR-V2.6.0 版本升级到 EMR-V2.7.0 版本），为防止一些升级不兼容、环境变化等问题的出现，请务必测试需要迁移的任务，以确保在新的软件环境中可以正常运行。
- EMR 发行版本中 jdk 基于腾讯 kona（基于 OpenJDK8），基于云场景的支撑及特性，我们在 kona 进行了开发及优化，kona 详情可参考 Kona JDK 。

组件版本概览

最近更新时间：2025-05-16 15:38:41

EMR 已停止购买版本说明

部分历史 EMR 发行版因开源组件版本较低，无法体验社区新特性，现已下线。下线的 EMR 发行版不再支持新集群创建，已购的集群仍可正常使用扩缩容等服务。
EMR on CVM 已下线版本：
Hadoop 集群类型： EMR-V1.3.1、EMR-V2.0.1、EMR-V2.1.0、EMR-V2.2.0、EMR-V2.3.0、EMR-V2.4.0、EMR-V2.5.0、EMR-V2.5.1、EMR-V3.0.0、EMR-V3.2.0、EMR-V3.2.1、EMR-TIANQIONG-V1.0.0。
StarRocks 集群类型：StarRocks-V1.0.0、StarRocks-V1.1.0、StarRocks-V1.2.0、StarRocks-V1.3.0、StarRocks-V2.1.0。
推荐您使用各集群类型最新稳定发行版进行集群创建，以获得更多的特性及更好的稳定性。

EMR on CVM 发行版本更新记录

EMR on CVM 支持 Hadoop 、Kafka 、StarRocks 等集群类型，Hadoop 集群类型当前提供标准版本和 beta 版本两种选择。

Hadoop 集群标准版本 Hadoop2.x支持组件版本

组件名称	EMR-V 2.8.0	EMR-V 2.7.0	EMR-V 2.6.0
发布时间	2024.03	2022.07	2021.07
hdfs（必选组件）	2.8.5	2.8.5	2.8.5
yarn（必选组件）	2.8.5	2.8.5	2.8.5
zookeeper（必选组件）	3.6.3	3.6.3	3.6.1
openldap（必选组件）	2.4.44	2.4.44	2.4.44
knox（必选组件）	1.6.1	1.6.1	1.2.0
tez	0.10.1	0.10.1	0.9.2
hive	2.3.9	2.3.9	2.3.7
spark	3.3.4	3.2.1	3.0.2

Livy	–	0.8.0	0.8.0
kyuubi	1.7.3	1.4.1	1.4.1
kylin	4.0.3	4.0.1	2.5.2
presto	–	–	–
trino(prestosql)	435	385	332
kudu	1.16.0	1.15.0	1.12.0
impala	3.4.1	3.4.0	3.4.0
storm	–	1.2.3	1.2.3
flink	1.16.1	1.14.3	1.12.1
hbase	2.4.5	2.4.5	1.4.9
phoenix（集成在 hbase 中）	5.1.2	5.1.2	4.14.3
alluxio	2.9.3	2.8.0	2.5.0
iceberg	1.4.3	0.13.0	0.11.0
hudi	0.14.1	0.11.0	0.7.0
Hue	4.10.0	4.10.0	4.6.0
oozie	5.2.1	5.2.1	5.1.0
zeppelin	–	0.10.1	0.9.1
superset	2.1.1	1.4.1	0.35.2
tensorFlowSpark	–	1.4.4	1.4.4
jupyter	7.0.6	–	–
sqoop	1.4.7	1.4.7	1.4.7
flume	1.9.0	1.9.0	1.9.0
ranger	2.1.0	2.1.0	1.2.0
kerberos（仅创建时支持选择）	1.15.0	1.15.0	1.15.0

版权所有：腾讯云计算（北京）有限责任公司

ganglia	-	3.7.2	3.7.2
goosefs	1.4.2	1.2.0	-
keepalived（Trino HA场景自动创建，白名单功能）	2.1.5	2.1.5	-
uniffle	0.10.0	0.10.0	-

Hadoop 集群标准版本 Hadoop3.x支持组件版本

组件名称	EMR-V3.6.0	EMR-V3.5.0	EMR-V3.4.0	EMR-V3.3.0
发布时间	2023.08	2022.10	2022.04	2021.09
hdfs（必选）	3.2.2	3.2.2	3.2.2	3.2.2
yarn（必选）	3.2.2	3.2.2	3.2.2	3.2.2
zookeeper（必选组件）	3.6.3	3.6.3	3.6.3	3.6.1
openldap（必选组件）	2.4.44	2.4.44	2.4.44	2.4.44
knox（必选组件）	1.6.1	1.6.1	1.6.1	1.2.0
tez	0.10.2	0.10.2	0.10.1	0.10.1
hive	3.1.3	3.1.3	3.1.2	3.1.2
spark	3.3.2	3.2.2	3.2.1	3.0.2
livy	0.8.0	0.8.0	0.8.0	0.8.0
kyuubi	1.7.0	1.6.0	1.4.1	1.1.0
kylin	4.0.3	4.0.1	4.0.1	4.0.1
presto	-	-	-	-
trino（prestosql）	414	389	372（改名trino）	350

impala	4.1.1	4.1.0	4.0.0	3.4.0
kudu	1.16.0	1.16.0	1.15.0	1.15.0
hbase	2.4.5	2.4.5	2.4.5	2.3.5
phoenix（集成在hbase 中）	5.1.2	5.1.2	5.1.2	5.1.2
flink	1.16.1	1.14.5	1.14.3	1.12.1
hue	4.10.0	4.10.0	4.10.0	4.10.0
oozie	5.2.1	5.2.1	5.1.0	5.1.0
zeppelin	0.10.1	0.10.1	0.10.1	0.9.1
superset	2.0.1	1.5.1	1.4.1	1.4.1
alluxio	2.8.0	2.8.0	2.8.0	2.5.0
iceberg	1.1.0	0.13.1	0.13.1	0.11.0
hudi	0.13.0	0.12.0	0.11.0	0.8.0
flume	1.11.0	1.10.0	1.9.0	1.9.0
sqoop	1.4.7	1.4.7	1.4.7	1.4.7
ranger	2.3.0	2.3.0	2.1.0	2.1.0
kerberos（仅创建时支持选择）	1.15.1	1.15.1	1.15.1	1.15.1
ganglia	-	3.7.2	3.7.2	3.7.2
deltalake	2.2.0	2.0.0	-	-
goosefs	1.4.2	1.4.0	1.2.0	-
keepalived（Trino HA场景自动创建，白名单功能）	2.1.5	2.1.5	2.1.5	-
uniffle	0.10.0	0.10.0	-	-

Hadoop 集群 Beta 版本支持组件版本

组件名称	EMR-V3.6.1
发布时间	2024.04
hdfs（必选）	3.2.2
yarn（必选）	3.2.2
zookeeper（必选组件）	3.6.3
openldap（必选组件）	2.4.46
knox（必选组件）	1.6.1
tez	0.10.2
hive	3.1.3
spark	3.2.2
kyuubi	1.7.0
kylin	4.0.3
trino（prestosql）	414
flink	1.16.1
oozie	5.2.1
superset	2.0.1
alluxio	2.8.0
iceberg	1.1.0
hudi	0.13.0
flume	1.11.0
sqoop	1.4.7
ranger	2.3.0
kerberos（仅创建时支持选择）	1.15.1
deltalake	2.2.0

版权所有：腾讯云计算（北京）有限责任公司

Kafka 集群已支持组件产品版本

组件名称	KAFKA- V2.0.0	KAFKA-V 1.0.0
发布时间	2023.03	2021.05
Kafka（必选组件）	2.4.1	1.1.1
kafkamanager（必选组件）	2.0.0.2	2.0.0.2
knox（必选组件）	1.2.0	1.2.0
zookeeper（必选组件）	3.6.3	3.6.1

StarRocks 集群已支持组件产品版本

组件名称	STARROCKS-V2.2.0	STARROCKS-V2.1.1	STARROCKS-V2.0.0	STARROCKS-V1.4.0
发布时间	2025.01	2024.08	2023.09	2023.04
starrocks（必选组件）	3.3.7	3.2.9	3.1.2	2.5.3
knox（必选组件）	1.2.0	1.2.0	1.2.0	1.2.0
openldap（必选组件）	2.4.44	2.4.44	-	-
ranger	2.3.0	2.3.0	-	-

RSS 集群已支持组件产品版本

组件名称	RSS-V1.0.0
发布时间	2025.01
uniffle（必选组件）	0.10.0
knox（必选组件）	1.6.1

版权所有：腾讯云计算（北京）有限责任公司

EMR on TKE 发行版本更新记录

EMR on TKE 支持 数据湖和机器学习集群类型。

数据湖集群已支持组件产品版本

组件名称	EMR-TKE V1.1.0	EMR-TKE V1.0.1	EMR-TKE-DLC V1.0.0	EMR-TKE V1.0.0
发布时间	2024.6	2023.11	2023.11	2023.5
spark	3.3.2	3.2.2	-	3.2.2
virtualspack	-	-	3.2.2	-
kyuubi	1.7.0	1.7.0	1.7.1	1.6.0
zookeeper	3.6.3	3.6.3	3.6.3	3.6.3
knox	1.6.1	1.6.1	-	1.6.1
hiveserver2	3.1.3	3.1.3	-	3.1.3
metastore	3.1.3	3.1.3	-	3.1.3
trino	414	389	-	389
ranger	2.3.0	2.3.0	-	2.3.0
hue	4.10.0	4.10.0	-	4.10.0
uniffle	0.7.1	0.7.1	-	0.6.0
openldap	2.4.44	2.4.44	-	-
presto	-	-	0.242	-
cosranger	6.0.2	-	-	-

机器学习集群已支持组件产品版本

组件名称	EMR-TKE-AI- V1.1.0
发布时间	2024.12
JupyterLab	4.2.5

eg	3.2.2
MLflow	2.16.2

组件与 API 部署时映射关系

最近更新时间：2024-05-31 08:59:31

组件名与 API 对应数字映射关系或进程名与 API 对应数字映射关系，主要用于**扩容集群节点**（ScaleOutCluster）API 接口的 ServiceNodeInfo.N 和 SoftDeployInfo.N字段参数 long 型映射填写。

组件名	组件名映射 ID	进程名	进程名映射 ID
KYUUBI	22	KyuubiServer	87
YARN	2	ResourceManager	6
		NodeManager	7
		JobHistoryServer	14
		TimeLineServer	102
ZOOKEEPER	0	QuorumPeerMain	0
ZEPPELIN	15	Zeppelin	27
TEZ	13	Tez	25
		Tomcat	103
TENSORFLOWONSPARK	31	–	–
SUPERSET	28	Superset	56
STORM	11	Nimbus	22
		Supervisor	23
SQOOP	7	Sqoop	13
SPARK	6	Spark	12
		SparkJobHistoryServer	17
RANGER	16	Ranger	28
		RangerUsersync	68

		Solr	95
PRESTOSQL	33	PrestoSqlWorker	66
		PrestoSqlCoordinator	67
PRESTO	5	PrestoWorker	11
		PrestoCoordinator	10
OOZIE	9	Oozie	16
LIVY	27	LivyServer	55
KYLIN	14	Kylin	26
KUDU	32	KuduMaster	64
		KuduServer	65
KRB5	21	Krb5Kdc	43
KNOX	24	Gateway	50
OPENLDAP	39	Slapd	51
IMPALA	23	ImpalaServer	47
		ImpalaServerCoordinator	45
		ImpalaServerExecutor	46
		ImpalaStateStore	48
		ImpalaCataLog	49
HUE	8	Hue	15
HUDI	26		54
HIVE	4	HiveServer2	31
		HiveMetaStore	30
		HiveWebHcat	32

版权所有：腾讯云计算（北京）有限责任公司

HIVESERVER2	47	HiveServer2	31
		HiveWebHcat	32
METASTORE	48	HiveMetaStore	30
HDFS	1	NameNode	1
		Router	88
		DataNode	2
		zkfc	8
		JournalNode	3
HBASE	3	HMaster	4
		HRegionServer	5
		HbaseThrift	37
GANGLIA	10	Httpd	36
		Gmond	21
		Gmetad	19
FLUME	18	Flume	38
FLINK	12	Flink	24
FILEBEAT	25	Filebeat	53
DRUID	29	router	62
		coordinator	57
		overlord	58
		historical	59
		middleManager	60
		broker	61
COSRANGER	34	CosRangerServer	71
CLICKHOUSE	30	ClickHouseServer	63

版权所有：腾讯云计算（北京）有限责任公司

ALLUXIO	20	AlluxioMaster	41
		AlluxioWorker	42
		AlluxioJobMaster	69
		AlluxioJobWorker	70
GOOSEFS	38	GooseFSMaster	82
		GooseFSWorker	83
		GooseFSJobMaster	84
		GooseFSJobWorker	85
		GooseFSProxy	101
DORIS	35	DorisFeFollower	72
		DorisBe	73
		DorisBroker	74
		DorisFeObserver	75
KAFKA	36	Kafka	76
KAFKAMANAGER	37	kafkamanager	77
ICEBERG	40	Iceberg	44
DELTA	46	Delta	44
STARROCKS	41	StarRocksFeFollowe r	78
TRINO	42	StarRocksBe	79
		StarRocksBroker	80
		StarRocksFeObserv er	81
		StarRocksCn	104
TRINO	42	TrinoWorker	89
		TrinoCoordinator	90

版权所有：腾讯云计算（北京）有限责任公司

KEEPALIVED	54	keepalived	107