消息队列 CKafka

产品简介

产品文档

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2019 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

【商标声明】

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。

【服务声明】

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。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模式的承诺或保证。

文档目录

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：2019-08-22 10:46:55

什么是消息队列 CKafka

消息队列 CKafka（Cloud Kafka）是基于开源 Apache Kafka 消息队列引擎，提供高吞吐性能、高可扩展性的消息队列服务。消息队列 CKafka 完美兼容 Apache kafka 0.9、0.10版本接口，在性能、扩展性、业务安全保障、运维等方面具有超强优势，让您在享受低成本、超强功能的同时，免除繁琐运维工作。

产品功能

- **收发解耦**
  有效解耦生产者、消费者之间的关系。在确保同样的接口约束的前提下，允许独立扩展或修改生产者/消费者间的处理过程。

- **可扩展性**
  由于消息的处理过程被解耦，只需要水平扩展处理过程，即可有效增加消息的入队效率和处理效率，十分灵活。

- **削峰填谷**
  消息队列 CKafka 能够抵挡突增的访问压力，不会因为突发的超负荷的请求而完全崩溃，有效提升系统健壮性。

- **可恢复性**
  当系统的部分组件出现故障时，整个系统不会因此受到影响，增加了系统的容错能力。即使某一处理消息的进程故障，队列中的消息依然可以在系统恢复后被处理。

版权所有：腾讯云计算（北京）有限责任公司

- **顺序读写**
  消息队列 CKafka 能够保证一个 Partition 内消息的有序性。和大部分的消息队列一致，消息队列 CKafka 可以保证数据按照顺序进行处理，极大提升磁盘效率。

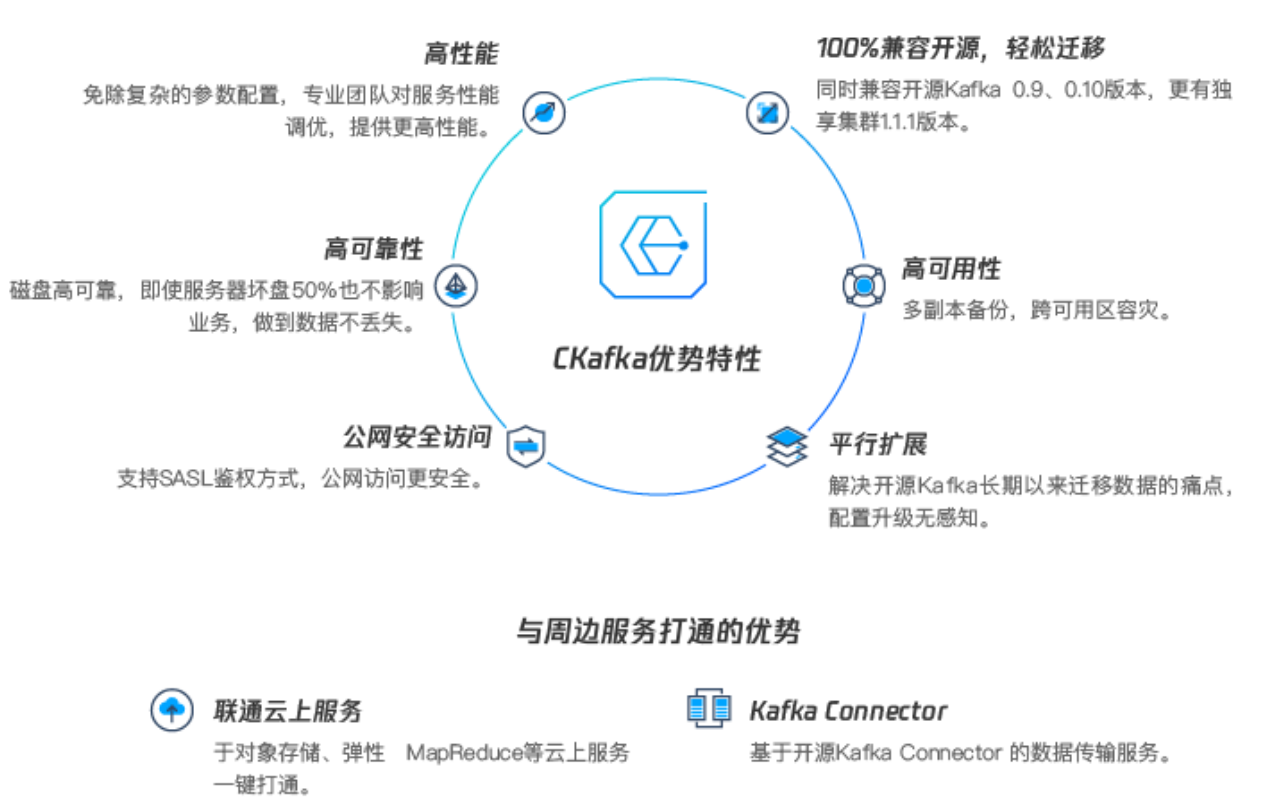
- **异步通信**
  在业务无需立即处理消息的场景下，消息队列 CKafka 提供了消息的异步处理机制，访问量高时仅将消息放入队列中，在访问量降低后再对消息进行处理，缓解系统压力。

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：2019-08-20 12:14:07

本文主要介绍消息队列 CKafka 相比于自建开源 Apache Kafka 所具备的优势。



自身优势特性

100%兼容开源，轻松迁移

消息队列 CKafka 不仅兼容开源 Kafka 0.9和0.10版本，更有独享集群1.1.1版本。
消息队列 CKafka 业务系统基于现有的开源 Apache Kafka 生态的代码，无需任何改造，即可迁移上云，享受到腾讯云提供的高性能消息队列 CKafka 服务。迁移方法请参考 迁移数据到 CKafka。

高性能

腾讯云消息队列专业团队对服务性能进一步调优，免除复杂的参数配置，提供更高性能。

高可用性

依托腾讯技术工程多年监控平台的技术积累，对集群全方位多角度监控，更有专业运维团队7 × 24小时处理告警保障消息队列 CKafka 服务的高可用性。

版权所有：腾讯云计算（北京）有限责任公司

更有多可用区容灾方案可选，零感知服务迁移。

高可靠性

磁盘高可靠，即使服务器坏盘50%也不影响业务。
默认2副本，支持3副本，副本越多可靠性越高。

平行扩展

解决开源 Kafka 长期以来迁移数据的痛点，配置升级无感知。

公网安全访问

支持 SASL 鉴权方式，公网访问更安全。

数据安全

消息队列 CKafka 提供鉴权与授权机制、主子账号等功能，提供企业级的安全防护。
腾讯云私有网络（VPC）：支持腾讯云 VPC 访问，网络环境安全。
主子账号：全面支持腾讯云 CAM 主子账号、协作者等功能，实现主子账号之间以及企业间跨账号的授权服务。

与周边服务打通的优势

联通云上服务

消息队列 CKafka 支持与对象存储、弹性 MapReduce（EMR）等云上服务一键打通。

Kafka Connector

支持基于开源 Kafka Connector 的数据传递服务，两个 Kafka 集群间可互相传递数据。

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：2019-09-05 18:32:57

消息队列 CKafka 广泛应用于大数据领域，如网页追踪行为分析、日志聚合、监控、流式数据处理、在线和离线分析等。
您可以通过以下方式让数据集成变得简单：

- 将消息队列 CKafka 中的消息导入到腾讯云平台的 COS、流计算等数据仓库。
- 通过 SCF 触发器的方式连接云上其他产品。

网页追踪

消息队列 CKafka 通过实时处理网站活动（PV、搜索、用户其他活动等），并根据类型发布到 Topic 中，这些信息流可以被用于实时监控或离线统计分析等。
由于每个用户的 page view 中会生成许多活动信息，因此网站活动跟踪需要很高的吞吐量，消息队列 CKafka 可以完美满足高吞吐、离线处理等要求。

版权所有：腾讯云计算（北京）有限责任公司

日志聚合

消息队列 CKafka 的低延迟处理特性，易于支持多个数据源和分布式的数据处理（消费）。相比于中心化的日志聚合系统，消息队列 CKafka 可以在提供同样性能的条件下，实现更强的持久化保证以及更低的端到端延迟。
消息队列 CKafka 的特性决定它非常适合作为"日志收集中心"；多台主机/应用可以将操作日志"批量""异步"地发送到消息队列 CKafka 集群，而无需保存在本地或者 DB 中；消息队列 CKafka 可以批量提交消息/压缩消息，对于生产者而言，几乎感觉不到性能的开支。此时消费者可以使用 Hadoop 等其他系统化的存储和分析系统对拉取日志进行统计分析。

大数据场景

在一些大数据相关的业务场景中，需要对大量并发数据进行处理和汇总，此时对集群的处理性能和扩展性都有很高的要求。消息队列 CKafka 在实现上的数据分发机制，磁盘存储空间的分配、消息格式的处理、服务器选择以及数据压缩等方面，也决定其适合处理海量的实时消息，并能汇总分布式应用的数据，方便系统运维。
在具体的大数据场景中，消息队列 CKafka 能够很好地支持离线数据、流式数据的处理，并能够方便地进行数据聚合、分析等操作。

云函数触发器

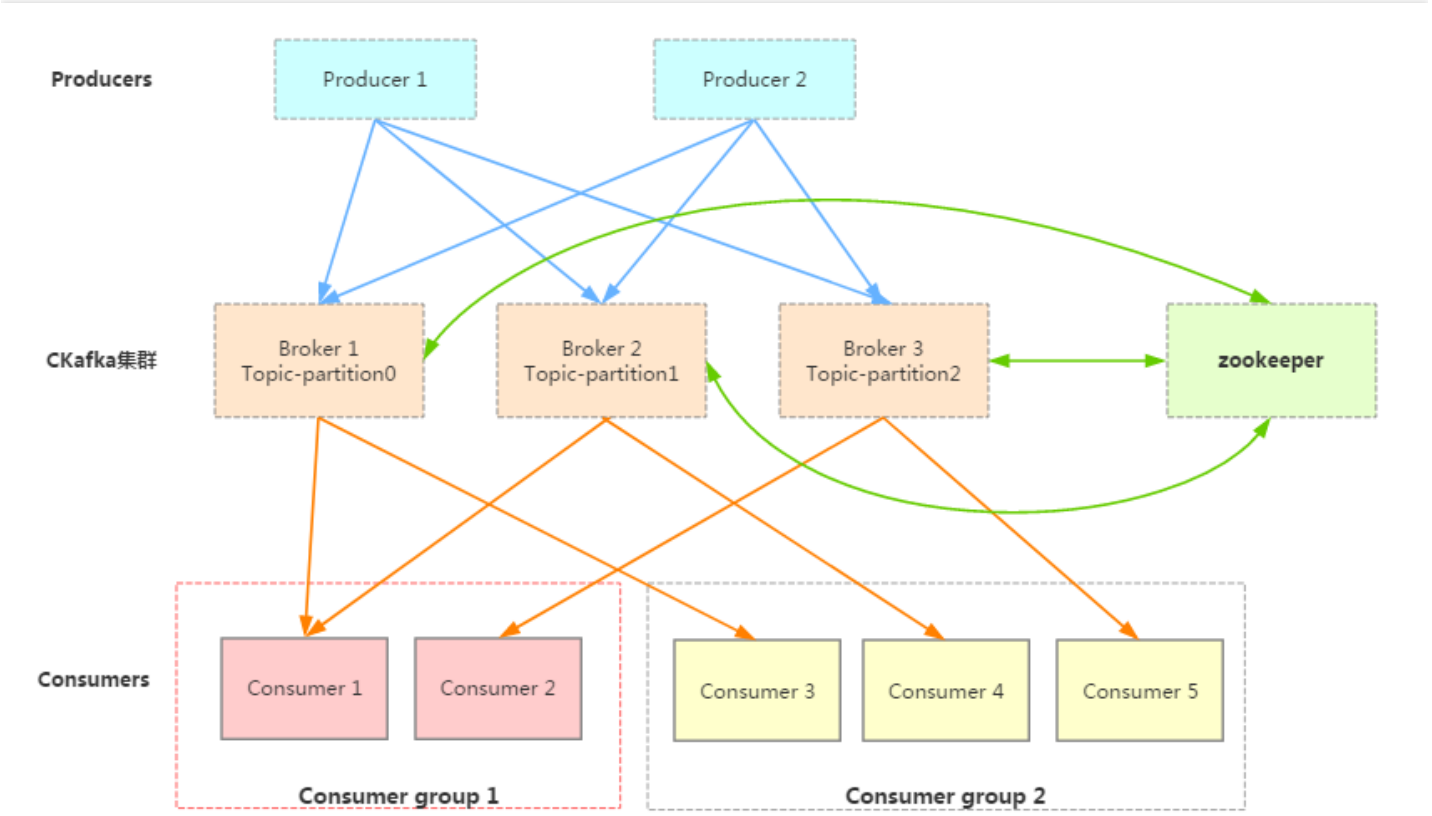
消息队列 CKafka 可以作为云函数触发器，在消息队列中接收到消息时将触发云函数的运行，并会将消息作为事件内容传递给云函数。例如，Ckafka 触发云函数时，云函数可以对消息进行结构变换、内容过滤等处理或者将消息投递到 Elasticsearch Service（ES）中。

> 说明：
> 云函数的可用性详见《云函数服务等级协议》。

技术原理

最近更新时间：2019-07-11 14:25:00

消息队列 CKafka 的架构图如下所示：



- 生产者 Producer 可能是网页活动产生的消息、服务日志等信息。生产者通过 push 模式将消息发布到 Cloud Kafka 的 Broker 集群。
- 集群通过 Zookeeper 管理集群配置，进行 leader 选举，故障容错等。
- 消费者 Consumer 被划分为若干个 Consumer Group。消费者通过 pull 模式从 Broker 中消费消息。

高吞吐

消息队列 CKafka 中存在大量的网络数据持久化到磁盘和磁盘文件通过网络发送的过程。这一过程的性能直接影响 Kafka 的整体吞吐量，主要通过以下几点实现：

- **高效使用磁盘**：磁盘中顺序读写数据，提高磁盘利用率。
  - 写 message：消息写到 page cache，由异步线程刷盘。
  - 读 message：消息直接从 page cache转入 socket 发送出去。

  - 当从 page cache 没有找到相应数据时，此时会产生磁盘 IO，从磁盘加载消息到 page cache，然后直接从 socket 发出去。
- **Broker 的零拷贝（Zero Copy）机制**：使用 sendfile 系统调用，将数据直接从页缓存发送到网络上。
- **减少网络开销**
  - 数据压缩降低网络负载。
  - 批处理机制：Producer 批量向 Broker 写数据、Consumer 批量从 Broker 拉数据。

数据持久化

消息队列 CKafka 的数据持久化主要通过如下原理实现：

- **Topic 中 Partition 存储分布**
  在消息队列 CKafka 文件存储中，同一 Topic 有多个不同 Partition，每个 Partition 在物理上对应一个文件夹，用户存储该 Partition 中的消息和索引文件。例如，创建两个 Topic，Topic1 中存在5个 Partition，Topic2 中存在10个 Partition，则整个集群上会相应生成5 + 10 = 15个文件夹。

- **Partition 中文件存储方式**
  Partition 物理上由多个 segment 组成，每个 segment 大小相等，顺序读写，快速删除过期 segment, 提高磁盘利用率。

水平扩展（Scale Out）

- 一个 Topic 可包含多个 Partition，分布在一个或多个 Broker 上。
- 一个消费者可订阅其中一个或者多个 Partition。
- Producer 负责将消息均衡分配到对应的 Partition。
- Partition 内消息是有序的。

Consumer Group

- 消息队列 CKafka 不删除已消费的消息。
- 任何 Consumer 必须属于一个 Group。
- 同一 Consumer Group 中的多个 Consumer 不同时消费同一个 Partition。
- 不同 Group 同时消费同一条消息，多元化（队列模式、发布订阅模式）。

多副本

多副本设计可增强系统可用性、可靠性。
Replica 均匀分布到整个集群，Replica 的算法如下：

1. 将所有 Broker（假设共 n 个 Broker）和待分配的 Partition 排序。
2. 将第 i 个 Partition 分配到第（i mod n）个 Broker 上。
3. 将第 i 个 Partition 的第 j 个 Replica 分配到第（(i + j) mode n）个 Broker 上。

Leader Election 选举机制

消息队列 CKafka 在 ZooKeeper 中动态维护了一个 ISR（in-sync replicas），ISR 里的所有 Replica 都跟上了 Leader。只有 ISR 里的成员才有被选为 Leader 的可能。

- ISR 中 f + 1个 Replica，一个 Partition 能在保证不丢失已 commit 的消息的前提下容忍 f 个 Replica 的失败。
- 共有 2f + 1个 Replica（包含 Leader 和 Follower），commit 之前必须保证有 f + 1个 Replica 复制完消息，为了保证正确选出新的 Leader，fail 的 Replica 不能超过 f 个。

版权所有：腾讯云计算（北京）有限责任公司

产品对比

最近更新时间：2019-05-27 13:18:53

消息队列 CKafka 与其他消息服务产品的性能对比详情如下：

特性	CKafka	Apache Kafka	RabbitMQ	RocketMQ	CMQ
优点	吞吐量非常大； 扩展性非常灵活； 运维成本极低	吞吐量大	可靠性高	可靠性高	可靠性非常高； 金融等强一致性场景
缺点	极端情况可能丢失消息	可能丢失消息； 扩展性不够灵活； 依赖组件多，运维量大； 安全防护功能有限，隔离和兼容性差	性能较差 扩展不灵活	HA 切换需要手动支持，不能自动化	为保证强一致性，吞吐量一般
开发语言	Scala	Scala	Erlang	Java	C++
可扩展性	非常灵活、易于扩展，发送消息只需指明 VIP 地址，Broker 的变化对于收发消息都透明	不够灵活，发送消息需指明 Broker 地址，接收消息需 ZooKeeper 协调调度	不够灵活，发送消息需要指明 Broker 地址	较灵活，发送方、接收方和 Name Server 连接	灵活、平滑、水平扩展，逻辑上单个 Queue可跨多个集群提供服务
吞吐量	非常大	较大	一般	一般	一般
常规性能	百万级QPS	百万级QPS	十万级QPS	十万级QPS	十万级QPS
2C 4GB 压测	读写22万QPS	读写20万QPS	读写10万QPS	读写10万 QPS	读写12万QPS
同步算法	ISR（Replica）	ISR（Replica）	GM	同步双写	Raft
可用性	可用性很高，主从自动切换，腾讯云消息服务承诺可用性 99.95%	可用性高，主从自动切换，但由于异步刷盘和复制，切换后可能会丢消息	主备自动切换，用 mirror queue 支持m/s，master 提供服务，slave 仅备份	不支持主从自动切换，master 不可用时 slave 只读不写	可用性很高，Broker 中存在2节点即可提供高可用服务

版权所有：腾讯云计算（北京）有限责任公司

特性	CKafka	Apache Kafka	RabbitMQ	RocketMQ	CMQ
消费方式	拉取方式	拉取方式	拉取和推送方式	拉取和推送方式	拉取和推送方式
消息可靠性	可靠性较高； 可通过三副本方式提升可靠性，集群容灾性能好，故障情况极少发生	可靠性低； Broker 只有异步刷盘机制并主备只有异步复制，可能会导致丢失部分消息	可靠性高； 发送消息时，指定消息为持久化就会写入到磁盘	可靠性高； Broker 同步双写，主备都写成功才返回成功	可靠性极高； 保证消息不丢失同步刷盘，数据持久性99.999999%
数据校验	CRC	CRC	无	CRC	Checksum
消息回溯	支持	支持	不支持	不支持	支持
安全防护	支持	不支持	不支持	不支持	支持
监控告警	支持	不支持	不支持	不支持	支持
服务支持	支持	不支持	不支持	不支持	支持

说明：

“2C 4GB压测”表示在2核 CPU 4GB内存服务器上压测的结果。