

腾讯云存储 Redis 产品白皮书

[2016.08.03]

[V2.0]



腾讯云

【版权声明】

©2015-2016 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。

目录

| | |
|------------------------|----|
| 腾讯云存储 Rsdis 产品白皮书..... | 1 |
| 1 前言..... | 5 |
| 2 产品定义..... | 6 |
| 2.1 云存储 Redis 介绍..... | 6 |
| 2.2 云存储 Redis 优势..... | 7 |
| 3 产品功能..... | 9 |
| 3.1 功能介绍..... | 9 |
| 3.2 功能列表..... | 9 |
| 3.3 使用限制..... | 13 |
| 3.3.1 主从版..... | 13 |
| 3.3.2 集群版..... | 13 |
| 3.4 产品示意图..... | 18 |
| 4 系统架构及技术原理..... | 20 |
| 4.1 系统架构..... | 20 |
| 4.1.1 在线存储系统..... | 21 |
| 4.1.2 数据高可用系统..... | 22 |
| 4.1.3 运维支撑系统..... | 23 |

| | |
|-------------------------|----|
| 4.1.4 支撑系统..... | 24 |
| 4.2 存储机主从同步技术原理..... | 25 |
| 4.3 故障恢复技术原理..... | 26 |
| 4.4 数据恢复技术原理..... | 27 |
| 4.5 数据备份技术原理..... | 28 |
| 4.6 在线弹性扩容技术原理..... | 29 |
| 4.7 RDB 文件导入导出技术原理..... | 29 |
| 4.8 多租户隔离技术原理..... | 30 |
| 5 实例规格与性能说明..... | 31 |
| 6 案例..... | 31 |
| 6.1 胡莱游戏..... | 31 |
| 6.2 小红书..... | 32 |
| 6.3 e 袋洗..... | 32 |
| 7 服务等级协议..... | 32 |
| 7.1 数据存储的持久性..... | 32 |
| 7.2 数据可销毁性..... | 33 |
| 7.3 数据可迁移性..... | 33 |
| 7.4 数据私密性..... | 33 |

| | | |
|------|---------------|----|
| 7.5 | 数据知情权..... | 33 |
| 7.6 | 数据可审查性..... | 34 |
| 7.7 | 业务功能..... | 34 |
| 7.8 | 业务可用性..... | 34 |
| 7.9 | 业务资源调配能力..... | 35 |
| 7.10 | 故障恢复能力..... | 35 |

1 前言

互联网的迅猛发展使得传统的关系型数据库在应对海量数据访问处理时面临巨大的挑战，而作为非关系型数据库 NoSQL 代表的 Key-Value 数据库以卓越的性能和极高的扩展性为弥补传统关系型数据库在互联网环境中应用的不足提供了解决方案。

Redis 是目前业界最为流行的高性能 Key-Value 数据库，它通过提供多种 Key-Value 数据类型来适应不同场景下的存储需求，并提供许多高层级的接口使其可以胜任如缓存、队列系统等不同的角色。不过，在 Redis 传统应用中，一般都存在内存利用率不高，数据落地过程中影响性能，扩容建设周期过长等问题。

随着云计算技术的不断成熟和腾讯云基础产品服务的不断发展，针对 Redis 在传统应用的痛点和难点，腾讯云推出了云存储 Redis。其包括了高性能、高可靠的 Redis 服务，整合了自动化管理工具，最大程度减少开发人员在部署、监控、扩容和故障恢复等方面的投入，使开发者可以集中精力进行产品开发和运营。

2 产品定义

2.1 云存储 Redis 介绍

云存储 Redis (Cloud Redis Store ，以下简称 CRS) 是腾讯云基于分布式缓存领域多年技术沉淀和 Redis 类业务运营的需求，打造的一款高可用、高可靠的 Redis 服务平台。CRS 支持集群版和主从版两种型号；通过主从热备自动容灾提供高可用性；通过多个节点存储多份数据拷贝保证高可靠性；集群版通过分布式存储摆脱单机的性能和容量限制，从而提高 Redis 的服务能力。

在使用上，CRS 兼容 Redis 协议，支持字符串、链表、集合、有序集合、哈希表等多种数据类型。

2.2 云存储 Redis 优势

腾讯云 CRS 将 Redis 存储能力作为一种服务提供给用户,使它相对于传统自建 Redis 更容易部署、管理和扩展;同时具有公有云按需申请按量付费的特点,使其成本效益更好,详见下表:

| 维度 | 腾讯云 CRS | 自建 Redis |
|-------|---|---|
| 价格优势 | 无软硬件投入,提供多种选择(集群版,主从版),按需付费 | 硬件:单台存储服务器成本高(如果搭主从,需要购买2台,资源冗余) 软件:需要招聘专业 DBA,人力成本高 |
| 服务可用性 | 服务采用双机热备架构,主机故障后,服务秒级切换到备机,服务切换不影响线上业务 | 需自行处理故障,自建主从 |
| 数据可靠性 | 拥有完善的数据自动备份和无损恢复机制(实时热备,3天内任意时刻数据恢复),让您用得放心 | 自行保障,依赖硬件的故障发生率,依赖技术人员的 Redis 管理水平 |
| 系统安全性 | 防DDoS攻击;及时修复各种安全漏洞 | 自行部署,价格高昂;自行处理 Redis 安全漏洞 |
| 实时 | 多维度监控,故障预警,让您用得 | 需自行开发监控系统,运维人员需 |

| | | |
|---------------|------------------------------|-----------------------------------|
| 监控 | 安心 | 半夜处理故障 |
| 业务 扩容 | 一键式按需扩容，快速部署，早日 上线，让您用得舒心 | 需自行完成硬件采购，机房托管， 应用重新部署等工作，周期较长 |
| 资源 利用 率 | 按需申请，资源利用率 100%，不 浪费您一分钱 | 峰值效用，机器的平均负载不高， 资源利用率低 |

针对自建 Redis 常出现的自建集群困难；运维麻烦；数据全量存储在内存费用昂贵；且作为 Key-Value 数据库使用时，回档不灵活，服务可用性和可靠性不高等难题，腾讯云 CRS 都有自己独有的优势：

1. 支持分布式服务

无缝升级，平滑扩容，不中断服务；分布式环境中支持 mget、mset 等多 Key 操作；支持事务操作

2. 稳定可靠

支持备份多拷贝，数据高可靠；主从热备，自动容灾，服务高可用；缓存持久化过程中不影响对外服务

3. 数据落地存储

支持数据持久化；支持数据操作流水日志；数据自动备份，多种恢复方式

4. 便捷管理

全 web 管理，提升运维效率；立体化监控，提前规避运营风险

3 产品功能

3.1 功能介绍

腾讯云 CRS 本质上是一种支持 Redis 协议的存储服务，不仅具有 Redis 的优点，而且也具有一些特有的功能，主要有以下几点：

1. 数据持久化

支持 Redis 协议，数据可持久化，且落地存储过程中不影响对外服务。适用于缓存和 Key-Value 数据库等多种场景。

2. 数据管理和备份回档

提供多种数据导入方式，轻松初始化 Redis。支持自动备份，CRS 可根据备份文件将数据恢复至 3 日内任意时刻，更支持指定 Key 不停机恢复数据。

3. 安全接入和自动容灾

支持私有网络 VPC，接入更安全。CRS 支持主从热备，主节点发生故障，秒级切换备节点，自动容灾。

4. 自定义监控告警

通过浏览器即可查看 CRS 丰富的性能监控数据，支持针对不同的监控项自定义阈值告警。

3.2 功能列表

| 特性 | 一级子特性 | 二级子特性 | 描述 |
|-------------|-------|-------|-----------------|
| 使用介绍 | 使用向导 | | 腾讯云首页顶部菜单-云产品-云 |

| | | | |
|------|--------|--------|--|
| | | | 存储 Redis wiki 里面提供 Redis 使用帮助文档 |
| 实例管理 | 新购实例 | 地域设置 | 可选择实例所在地域 |
| | | 可用区设置 | 可根据地域选择可用区 |
| | | 网络设置 | 可选择基础网络或者 VPC 网络 |
| | | 容量选择 | 主从版：1~60G |
| | | | 集群版：60~300G （需要更大容量可提交工单申请） |
| | 类型 | 提供热备 | |
| | 实例列表 | 显示实例信息 | 显示实例名称、运行状态、所属项目、所属地域、所属可用区、所属网络、内网地址、端口、总容量和已使用容量、创建时间、到期时间 |
| | | 批量操作 | 分配至项目 |
| | | 单实例操作 | 初始化、修改密码、续费、实例改名、管理、升级 |
| | 实例访问 | 内网访问 | |
| 实例监控 | 监控指标选择 | | 8 个监控指标： |

| | | | |
|-------------|--------------|--------|---|
| | | 指标 | <ol style="list-style-type: none"> 1. 内网入流量 2. 内网出流量 3. GET 数 4. SET 数 5. 已使用容量 6. CPU 负载 7. QPS 8. 缓存命中率 |
| | 监控时间窗口 选择 | | 今天、昨天、近 7 天、近 15 天、 近 30 天 |
| 数据管理 | 清空实例 | | 可在线对 Redis 实例进行清空 |
| | 数据在线管理 | | 可通过 Web 页面模拟登陆 Redis 实例进行命令行操作 |
| 数据备份与 恢复 | 备份管理 | | 提供备份数据管理，可按时间段进行查询 |
| | 手动备份 | | 手动执行数据备份任务 |
| | 自动备份 | 自动备份设置 | 可设置自动任备份周期和备份执行时间 |

| | | | |
|------|-------|----------|--|
| | 按时间回档 | | 可回档到最近三天任意时间的数据 |
| | 按流水回档 | 整体回档 | 可通过冷备整体回档数据 |
| | | 单 key 回档 | 可通过冷备按 key 回档数据 |
| | | 备份导出和下载 | 可以导出备份数据进行下载 |
| 任务列表 | 任务列表 | 任务类型 | 支持多种异步任务类型，可按任务类型进行筛选：RDB 导入、备份导出、恢复实例、备份实例、实例回档 |
| | | 处理状态 | 可按任务处理状态筛选：执行中、任务完成、任务失败、任务废弃 |
| | | 任务进度 | 可查看当前任务执行进度 |
| | | 任务信息 | 可查看任务开始时间 |
| | 搜索任务 | 任务搜索 | 可用实例 ID 或实例名模糊搜索该实例下的任务 |
| | | 日期搜索 | 可查询指定时间段内所有任务 |

3.3 使用限制

3.3.1 主从版

主从版完全支持开源 Redis 协议，兼容版本 Redis v2.8.23。

3.3.2 集群版

集群版基本兼容 Redis 协议，但由于分布式架构的特性，有部分命令暂不支持，具体如下：

下表：

支持的命令

| Key(键) | String(字符串) | Hash(哈希表) | List(列表) | Set(集合) | SortedSet(有序集合) |
|----------------|-------------|------------------|----------|------------|-----------------|
| EXPIRE AT | APPEND | HEXISTS | LINDEX | SADD | ZADD |
| PERSIST | BITCOUNT | HGET | LINSERT | SCARD | ZCARD |
| PEXPIRE | BITOP | HGETALL | LLEN | SDIFF | ZCOUNT |
| PEXPIRE EAT | DECR | HINCRBY | LPOP | SDIFFSTORE | ZINCRBY |
| PTTL | DECRBY | HINCRBYF LOAT | LPUSH | SINTER | ZRANGE |

| | | | | | |
|---------|-------------|--------|-----------|-------------|------------------|
| RESTORE | GET | HKEYS | LPUSHX | SINTERSTORE | ZRANGEBYSCORE |
| TTL | GETBIT | HLEN | LRANGE | SISMEMBER | ZRANK |
| TYPE | GETRANGE | HMGET | LREM | SMEMBERS | ZREM |
| DEL | GETSET | HMSET | LSET | SMOVE | ZREMRANGEBYRANK |
| DUMP | INCR | HSET | LTRIM | SPOP | ZREMRANGEBYSCORE |
| EXISTS | INCRBY | HSETNX | RPOP | SRANDMEMBER | ZREVRANGE |
| EXPIRE | INCRBYFLOAT | HVALS | RPOPLPUSH | SREM | ZREVRANGEBYSCORE |
| | MGET | HSCAN | RPUSH | SUNION | ZREVRANK |
| | MSET | HDEL | RPUSHX | SUNIONSTORE | ZSCORE |
| | MSETNX | | | SSCAN | ZUNIONSTORE |
| | PSETEX | | | | ZINTERSTORE |

| | | | | | |
|--|----------|--|--|--|----------------|
| | SET | | | | ZSCAN |
| | SETBIT | | | | ZRANGEBYLEX |
| | SETEX | | | | ZLEXCOUNT |
| | SETNX | | | | ZREMRANGEBYLEX |
| | SETRANGE | | | | |
| | STRLEN | | | | |

其他支持的命令

| Transaction (事务) | Connection (连接) |
|--------------------|-------------------|
| DISCARD | AUTH |
| EXEC | ECHO |
| MULTI | PING |
| UNWATCH | QUIT |
| WATCH | SELECT(仅 Select0) |

即将支持的命令

| List (列表) | Script (脚本) |
|-------------|---------------|
| | |

| | |
|------------|------|
| BLPOP | EVAL |
| BRPOP | |
| BRPOPLPUSH | |

不支持的命令

| Key (键) | Script(脚本) | Server (服务 器) | Pub/Sub (发 布/订阅) | HyperLogLo g |
|---------------|------------------|--------------------|-----------------------|-----------------|
| KEYS | EVALSHA | BGREWRITEA OF | PSUBSCRIBE | PFADD |
| MIGRATE | SCRIPT EXISTS | BGSAVE | PUBLISH | PFCOUNT |
| MOVE | SCRIPT FLUSH | CLIENT GETNAME | PUBSUB | PFMERGE |
| OBJECT | SCRIPT KILL | CLIENT KILL | PUNSUBSCR IBE | |
| RANDOMKE Y | SCRIPT LOAD | CLIENT LIST | SUBSCRIBE | |
| RENAME | | CLIENT SETNAME | UNSUBSCRI BE | |
| RENAMENX | | CONFIG GET | | |

| | | | | |
|------|--|---------------------|--|--|
| SORT | | CONFIG RESETSTAT | | |
| SCAN | | CONFIG REWRITE | | |
| | | CONFIG SET | | |
| | | DBSIZE | | |
| | | DEBUG OBJECT | | |
| | | DEBUG SEGFAULT | | |
| | | FLUSHALL | | |
| | | FLUSHDB | | |
| | | INFO | | |
| | | LASTSAVE | | |
| | | MONITOR | | |
| | | PSYNC | | |
| | | SAVE | | |
| | | SHUTDOWN | | |

| | | | | |
|--|--|---------|--|--|
| | | SLAVEOF | | |
| | | SLOWLOG | | |
| | | SYNC | | |
| | | TIME | | |

其他限制

Key 大小限制：127Byte，单 Key 下 Value 数据大小限制：1,000,000 Byte.

CRS 事务说明

- 1) 若事务部分成功，则失败部分不能回滚
- 2) 事务进行时，所涉及的 Key 均被锁定，此时读写 Key 会失败，返回冲突提示：“conflict with another client”
- 3) 已支持的命令，均能在事务中使用

3.4 产品示意图

1. 实例列表

| 实例名称 | 实例状态 | 所属项目 | 所属地域 | 可用区 | 网络 | 内网地址 | 已使用/总容量 | 到期时间 | 创建时间 | 管理 |
|-----------------|-------|-------|--------|--------|----------------|------------------|---------|----------------------|----------------------|----|
| kevinshu_tian | 实例运行中 | 默认项目 | 上海金融专区 | 上海金融一区 | 请用一个vpc测试-L... | 10.0.0.2-6379 | 0MB/1G | 2016-02-11 10:10:... | 2015-11-13 10:10:... | 升级 |
| 改名 | 实例运行中 | 默认项目 | 华南区-广州 | 广州二区 | 基础网络 | 10.66.89.4-6379 | 0MB/4G | 2016-02-07 16:39:... | 2015-11-09 16:39:... | 升级 |
| 91d3b5de-daa... | 实例运行中 | 测试项目2 | 华南区-广州 | 广州二区 | 基础网络 | 10.66.88.17-6379 | 0MB/1G | 2015-12-22 09:56:... | 2015-09-23 09:56:... | 升级 |
| d320cf2c-9a4... | 实例运行中 | 测试项目2 | 华南区-广州 | 广州二区 | 基础网络 | 10.66.88.14-6379 | 0MB/1G | 2015-12-21 14:26:... | 2015-09-22 14:26:... | 升级 |
| 腾讯云交易使... | 实例运行中 | 默认项目 | 华南区-广州 | 广州二区 | 基础网络 | 10.66.87.92-6379 | 0MB/1G | 2016-09-19 00:00:... | 2015-09-11 11:16:... | 升级 |

2. 实例详情

云存储Redis < 返回 | fc1721cd-1d65-45d4-b29d-dacc6df0c5f1

实例列表 任务管理

实例详情 系统监控 备份与恢复

基本信息

实例ID: fc1721cd-1d65-45d4-b29d-dacc6df0c5f1
 运行状态: 实例运行中
 所属网络: 基础网络
 所属地域: 华南区-广州
 所属项目: 默认项目
 内网地址: 10.66.127.163
 端口: 6379

配置信息

容量: 2G, 已用0MB
 到期时间: 2016-01-12 18:31:45
 创建时间: 2015-10-14 18:31:45

3.系统监控

< 返回 | 侧耳

实例详情 系统监控 备份与恢复

今天 昨天 最近7天 最近15天 最近30天

访问

- 内网入流量
- 内网出流量
- Get数
- Set数

运行状态

- 已使用容量
- CPU负载
- QPS
- 缓存命中率

内网入流量

内网出流量

4.备份与恢复

云存储Redis < 返回 | fc1721cd-1d65-45d4-b29d-dacc6df0c5f1

实例列表 任务管理

实例详情 系统监控 备份与恢复

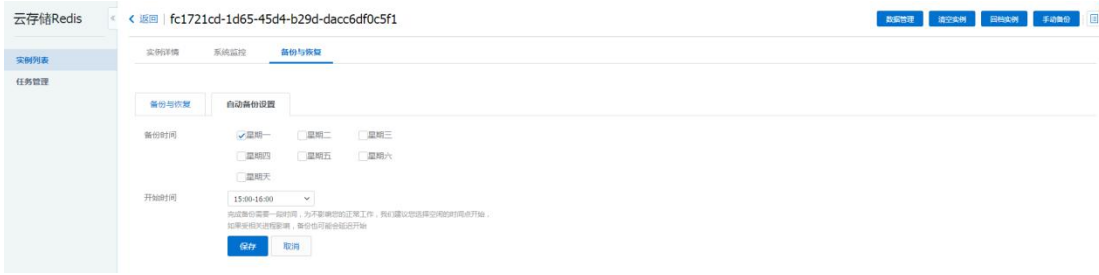
备份与恢复 自动备份设置

2015-11-12 至 2015-11-19

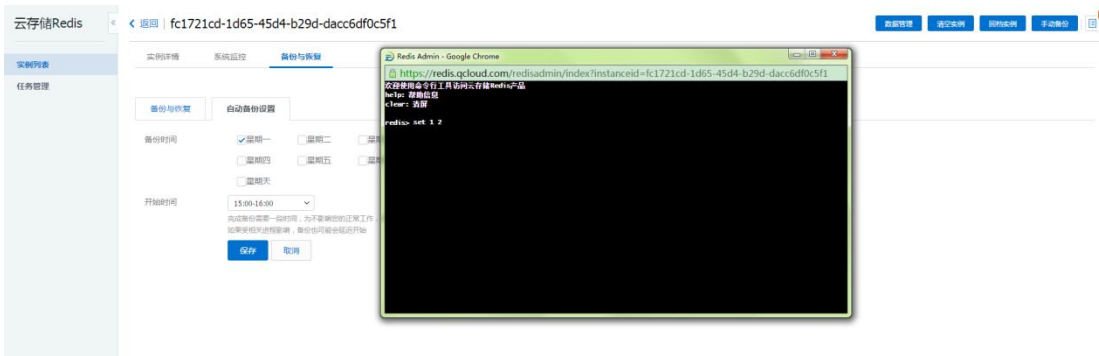
| 备份时间 | 备份方式 | 备注 | 操作 |
|---------------------|------|----------|--------------|
| 2015-11-19 10:24:06 | 手动备份 | 4_coding | 恢复实例 恢复指定Key |
| 2015-11-17 17:01:09 | 手动备份 | test_gbk | 恢复实例 恢复指定Key |
| 2015-11-16 20:35:15 | 手动备份 | 1234567a | 恢复实例 恢复指定Key |
| 2015-11-16 16:55:13 | 手动备份 | test | 恢复实例 恢复指定Key |
| 2015-11-16 15:16:19 | 自动备份 | -- | 恢复实例 恢复指定Key |
| 2015-11-12 17:50:04 | 自动备份 | -- | 恢复实例 恢复指定Key |

共6条 1/1页

5.自动备份设置



6.数据管理



7.任务管理



4 系统架构及技术原理

4.1 系统架构

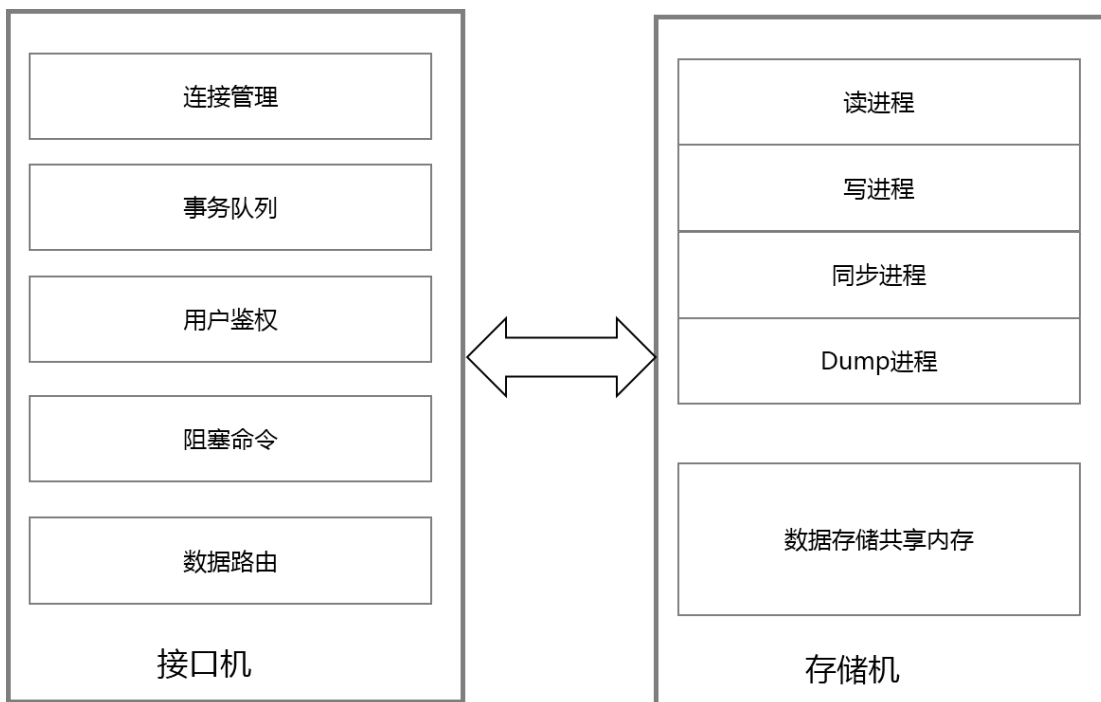
CRS 是利用分布式计算实现 Redis 相关协议的存储系统，将数据的存储和用户的接入独立开，同一个用户的数据分布在多台机器上，从而突破单机内存容量的限制。整个系统包括如下几部分：

- 在线存储系统：接入集群、存储集群和导入导出服务；
- 数据高可靠系统：主备同步模块、流水系统和冷备中心；

- 运维监控系统：日志中心和多维监控系统；
- 支持系统：任务中心、配置中心和路由系统；

4.1.1 在线存储系统

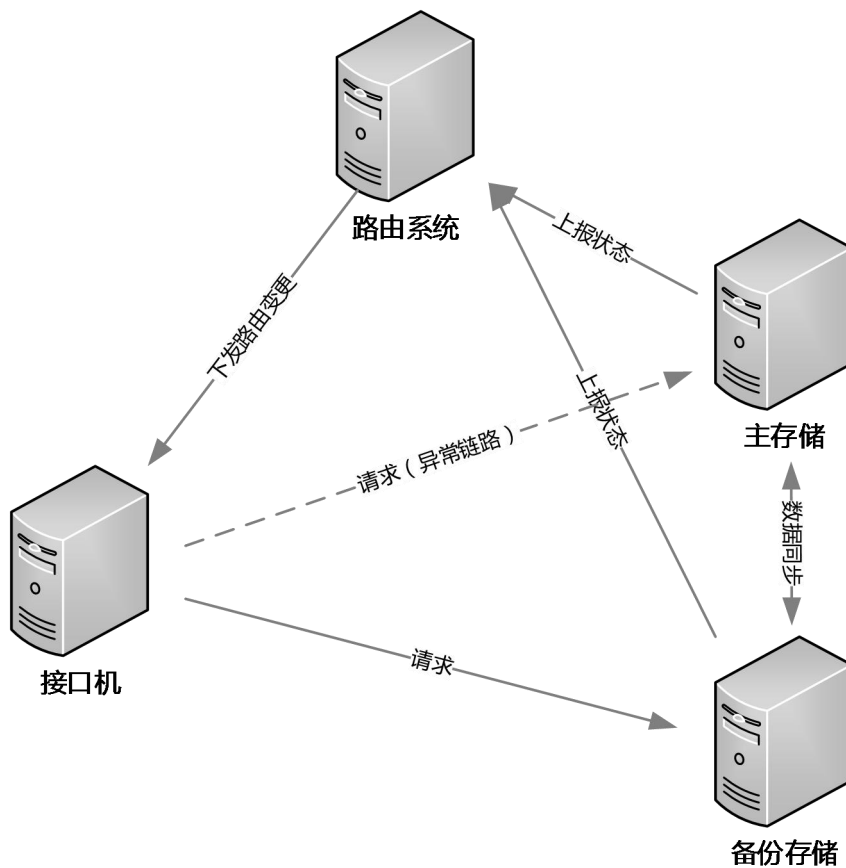
CRS 在线服务部分拆分为：接入集群和存储集群，接入集群负责管理客户应用发起的连接，接收用户命令，将命令转发到存储机；在主存储机发生故障时，将请求迁移到备机。同时，一些特殊的 Redis 命令也在接口机上完成计算工作。用户的数据按照一致性 hash 算法，散落到存储集群的多台机器上，保证数据无限平行扩展、以最小的迁移代价完成系统扩容，并降低单机死机对单个用户数据丢失的影响。存储机上，使用共享内存来存储数据，将数据读、写、同步和 dump 的功能以独立的进程来实现，避免了进程异常退出对数据的影响，同时解耦了对业务的操作代码，各种功能可以按需部署。



4.1.2 数据高可用系统

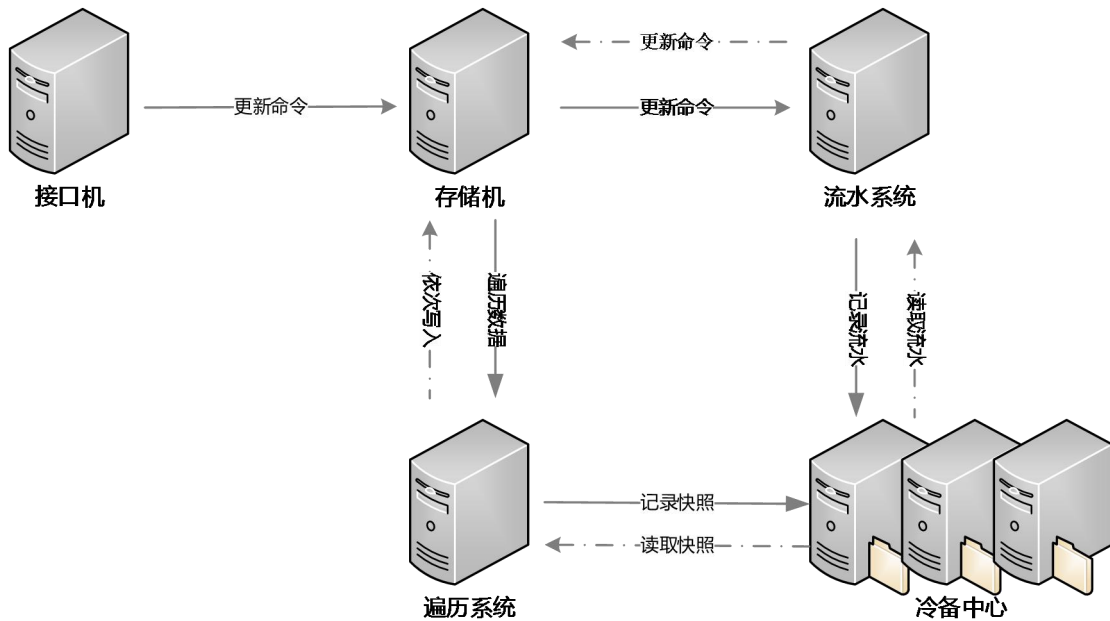
数据高可用系统包括：存储机的准实时同步，记录所有数据变更的流水系统以及某个时间点数据快照的冷备中心。

存储机准实时同步通过主备方式实现，数据更新命令，作用在主机上以后，会在短时间内，以异步的方式同步到备用服务器，使得备份服务器上的数据和主存储机上的数据差异非常小。在路由系统的配合下，及时发现主存储机异常，并将接口机对主存储机的请求，转发到对应的备份存储机上，以接近零数据丢失的代价，完成故障切换。



存储机数据保存在内存中，内存是易失性存储，重启后数据会丢失，即使有主备同步，也不能保证数据 100%的可靠性，为了应对主备同时死机的风险，系统提供了数据备份到磁盘文件的能力。备份包括两个维度：全量快照和增量流水，全量快照是某个时间点内存数

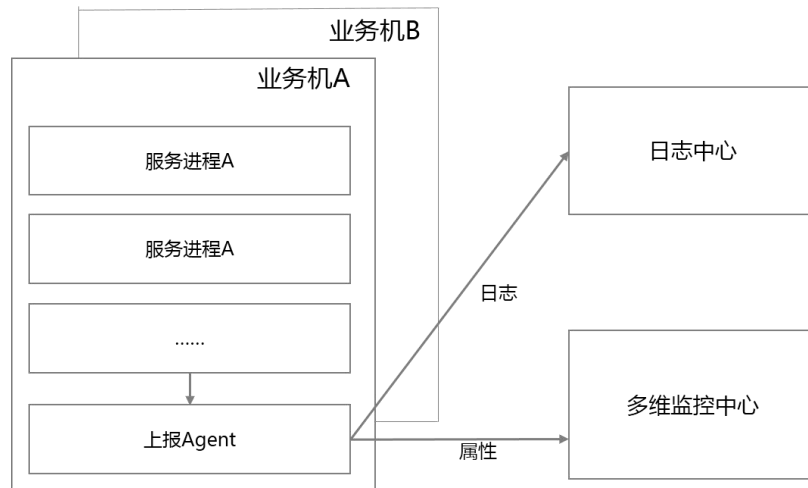
据的全量拷贝 增量流水记录对数据的每一个变更操作和对应的时间戳。利用流水和快照，系统能将数据恢复到三天内，任意时间点的状态；也可以恢复到指定快照创建时间点。备份的文件，存放在分布式文件系统中，分布式文件系统提供了 3 份拷贝，保证备份文件的可靠性。



4.1.3 运维支撑系统

运维支撑系统负责系统运行状态的监控、错误定位和为用户提供监控数据，主要由集中的日志系统和多维度属性统计系统。在分布式架构下，服务进程运行在多台机器上，日志也分布在多台机器上，定位问题时，需要登录多台机器查看日志，非常耗时，而且无法完成跨机器日志的联合分析，需要集中的日志中心，统一查询和存放日志文件。多维属性统计系统记录服务器某一个事件发生的次数，这些数据用来分析系统故障，比如：某一个错误发生的次数，同时记录用户命令执行情况，提供监控数据，比如：用户某一个命令执行的次数，每个命令的耗时。

系统结构上，都使用了 agent 上报，集中存储的方式，如下图所示：



4.1.4 支撑系统

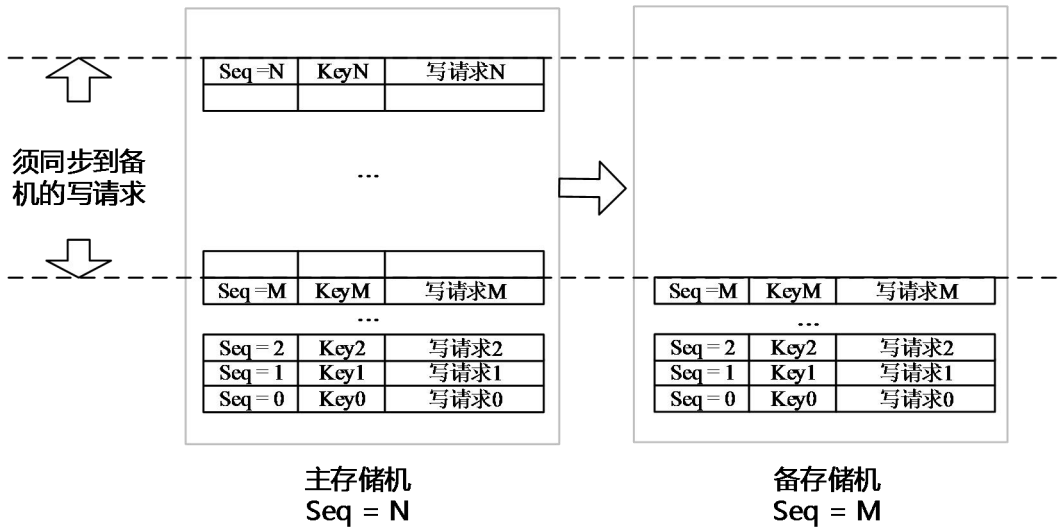
支撑系统是系统和业务运行必须的一些系统，包括：配置中心，路由系统和任务中心。

配置中心保存所有机器、实例的配置信息，并为每个模块提供接口，为各个业务提供透明的配置更新服务，某个配置项修改后，配置中心主动下发配置到接入的机器，供业务使用。

路由系统提供逻辑存储服务器到物理存储服务器的转换，接口机通过配置的存储机列表和一致性 hash 算法，计算出当前 key 需要访问的逻辑服务器，路由服务将逻辑服务器地址转换到物理存储机地址。路由系统另外一个作用是，在主存储机发生故障时，通过修改路由，将请求转换到备份存储机，实现故障恢复。

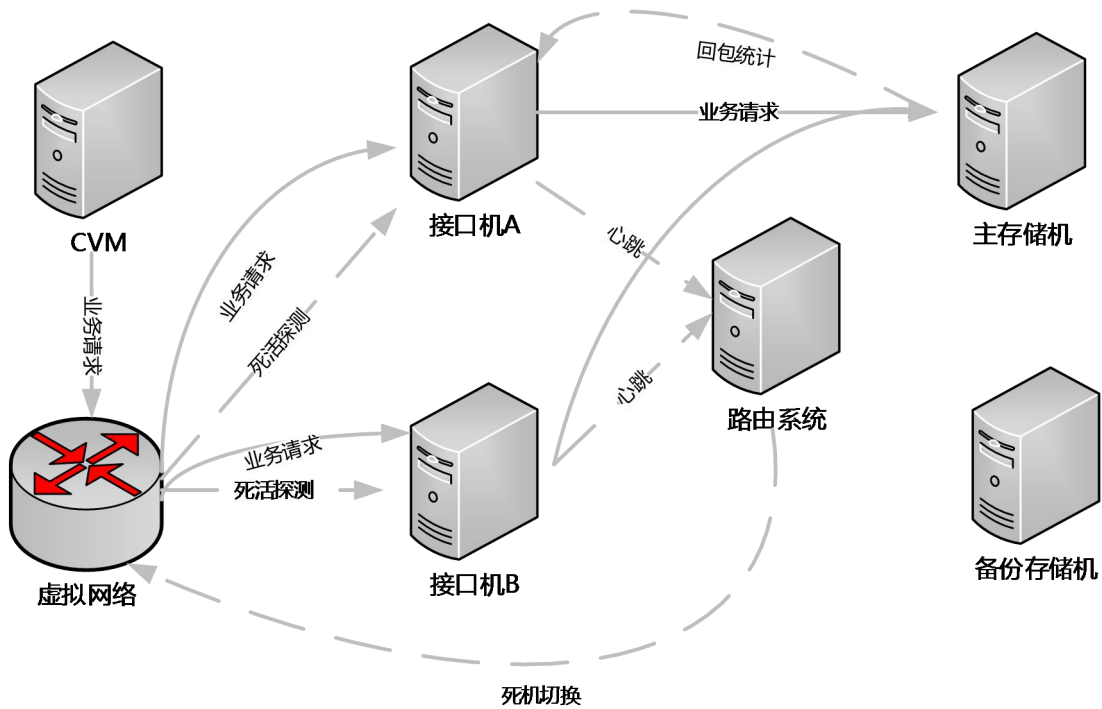
任务中心是一个异步的流程引擎，用户通过官网或 api 提交的任务，有些非常耗时，让用户一直等待执行完后再返回，这样的体验非常不友好。首先将任务记录在流程引擎中，然后流程引擎定时的异步执行这些任务，最后将结果返回给用户。比如：创建备份、恢复实例的任务。

4.2 存储机主从同步技术原理



存储机主从同步是保证数据安全性最重要的策略，系统使用序列号的方式同步，初始化时，主备存储机的序列号都初始化为 0，每一次数据更新时，序列号加 1，同时把 <seq, key, 写请求> 信息记录下来，也就是 binlog，当主存储机和备存储机的 seq 不相同步时，表明有写请求需要同步到备存储机，此时主存储机会依次把这些写请求从 binlog 中读出来，发送给备存储机进行同步。

4.3 故障恢复技术原理



用户在

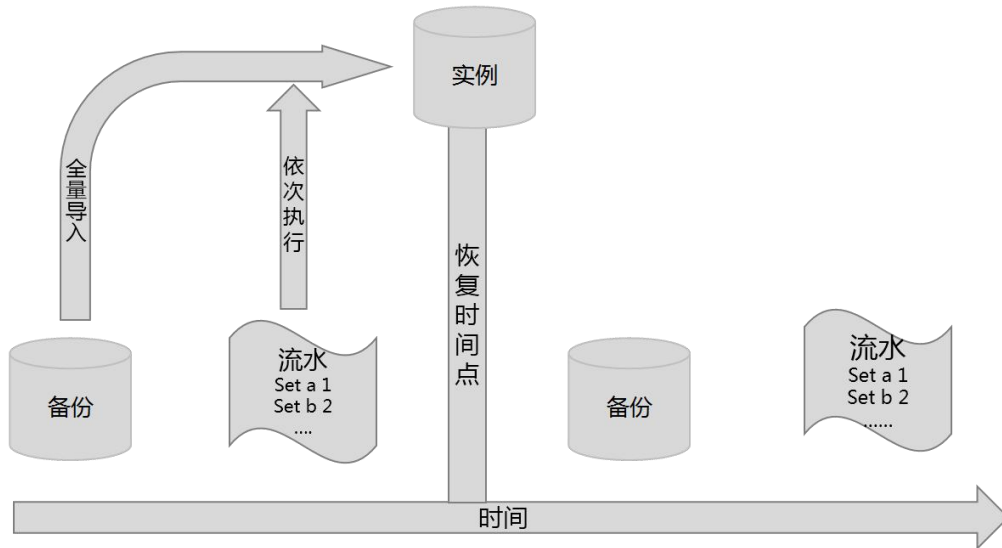
线服务涉及到接口和存储机，故障恢复系统对应包括：接口机和存储机故障恢复。故障恢复通常包括两个步骤：故障发现和故障恢复。

接口机故障发现包括两部分：CVM(云主机)虚拟机网络模块死活探测和接口机死活探测，CVM 虚拟机网络模块是一个路由模块，将用户使用的 ip 转换成 CRS 接口机服务部署的物理机，通过定时向物理机服务端口发送探测死活的 tcp 数据包，及时发现出现异常服务。接口机的死活探测通过接口机上的 agent 主动上报心跳实现。由于一个实例有多台接口机提供服务，所以，发现接口机故障时，只需要将对应的接口从路由列表中剔除即可。对于已经建立好的长连接，接口机服务故障，会向客户端发送连接单向断开的请求，客户端需要处理这种情况，否则，会在下一次发送请求时，收到连接重置的错误。

存储机故障发现使用回包统计策略：接口机通过统计单位时间内网络的成功率，网络成功率低于某个阈值时，任务服务已经不可用，需要发起故障恢复，通过回包统计，可以

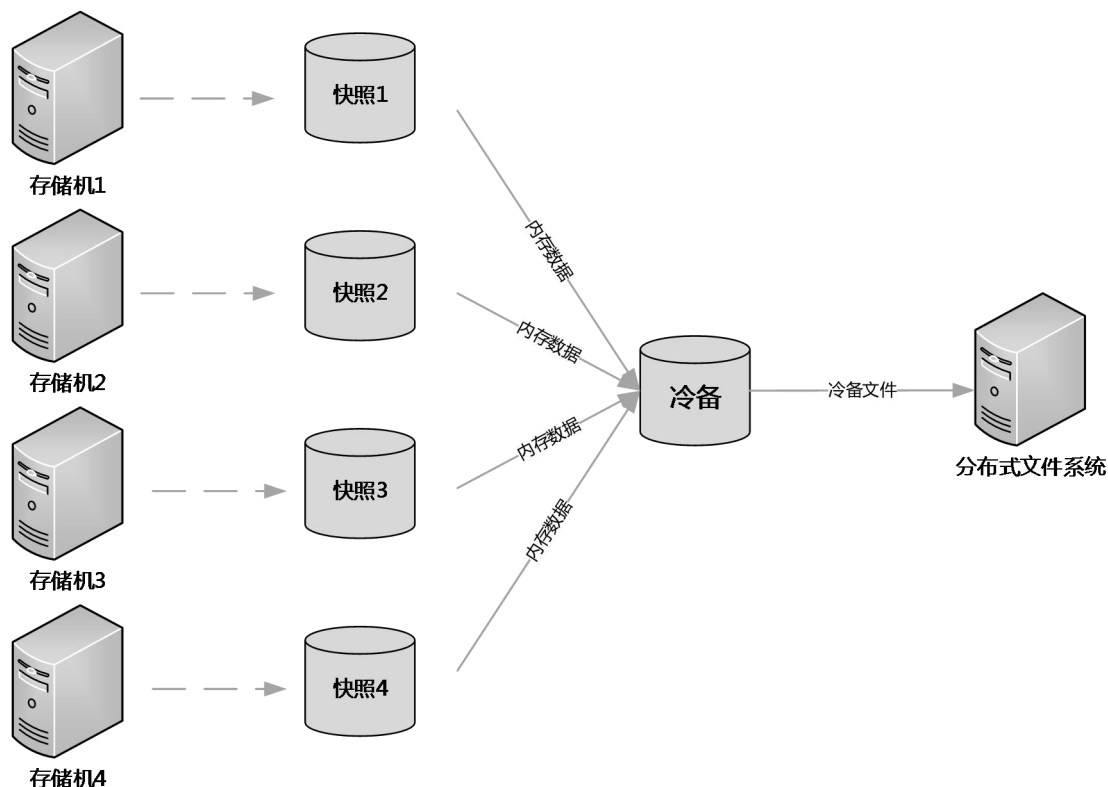
较快的发现故障；发现故障后，需要通知存储系统将备份存储机升级为主存储机，然后将请求转发到新的主存储机。

4.4 数据恢复技术原理



流水系统记录用户对数据的每个更新操作和对应的时间点，在恢复时，先找一个距离恢复时间点最近的全量备份，将全量备份恢复到 Redis 内存中，然后从全量备份的时间点开始，按照时间顺序依次重新执行记录的更新命令。由于更新命令非常多，目前只保留最近三天的流水，可实现三日内的任意时间点数据回档。

4.5 数据备份技术原理



冷备系统获取和保存某个时间点的内存镜像保存到磁盘文件，在需要时，将备份的磁盘文件恢复回内存中。开发商可以设置自动备份或手动触发备份，在需要的时候，指定特定的备份文件，将数据恢复到实例中。

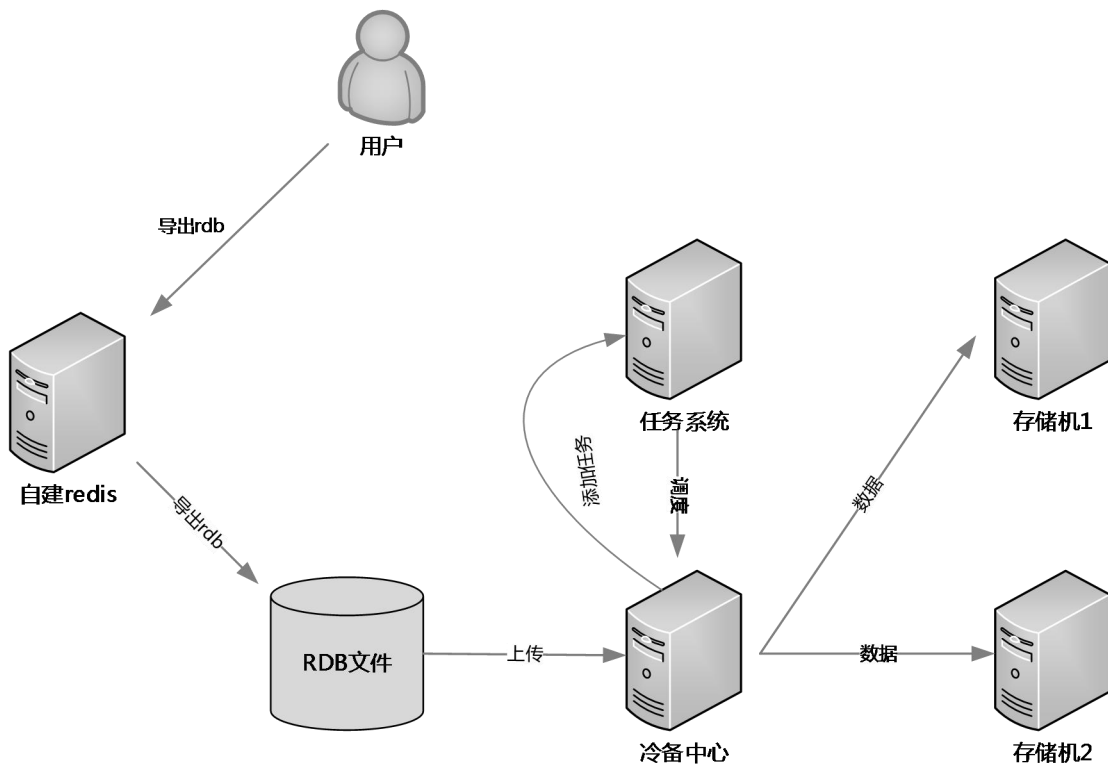
冷备系统由遍历系统、数据恢复模块和分布式文件系统组成，由于数据分布在多台机器上，一次导出所以机器的内存数据是不可能的，因此，需要在开始导出备份的数据点上，在所有存储机器上做一个快照，然后再将依次读取各个机器内存快照数据，从而保证导出的数据是某个时间点的拷贝，确保数据的一致性。分布式文件系统自身的数据冗余可以保证备份文件的安全，容量可以平行扩展，满足备份文件逐步增加的需求。

4.6 在线弹性扩容技术原理

对应主从版，受到单机可用内存的影响，如果实例扩容大小低于可用内存大小，可以直接修改实例相关配置，完成扩容。对于扩容增加容量大小大于可用内存的实例，需要先将实例迁移到内存容量足够的机器上，再调整实例相关配置。涉及到数据的迁移，这个过程相对时间较长。

与主从版受单机内存限制不同，集群版每个实例的容量大小并不受可用内存的影响，因此，在集群版中，在线扩容只需要修改实例的相关配置，即可达到扩容的目的。

4.7 RDB 文件导入导出技术原理

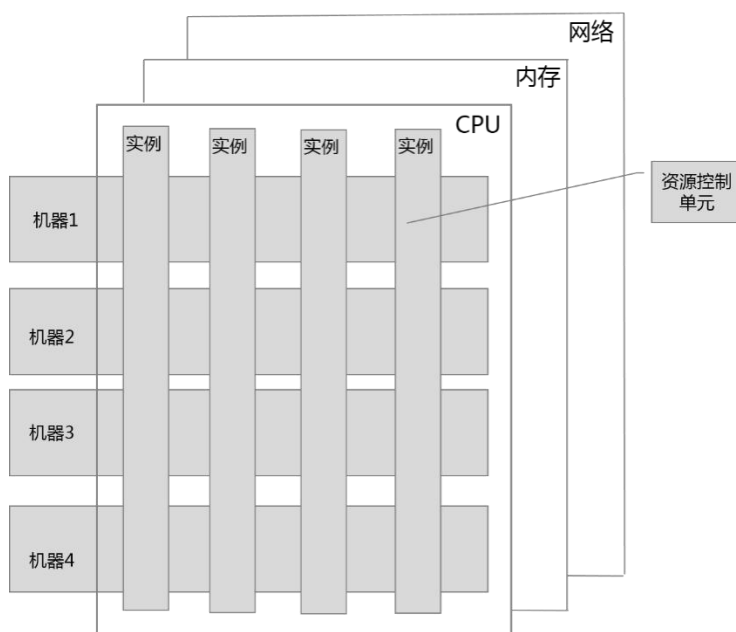


集群版 CRS 目前暂时不支持 Slaveof 的方式导入导出，需要借助 RDB 为载体离线导入导出。导入的步骤如下：

1. 开发商通过 BGSAVE 等方式，将 Redis 内存持久化到磁盘 RDB 文件
2. 在 CVM 上，通过 rit 工具，将导出的 RDB 文件，上传到冷备中心，并在任务中心提交导入任务；
3. 任务中心异步执行导入任务，将 RDB 文件内容导入到存储机

导出 RDB 文件过程和导入过程相反

4.8 多租户隔离技术原理



由于使用分布式架构，多个开发商的实例可以在同一个集群中共存，在多租户环境下，需要避免各租户之间的相互影响。集群配额就是避免多租户相互影响的一种策略，将集群的资源划分为 CPU、内存、网络、磁盘等多个维度，对不同规格的实例，指定可用的资源度量。同时，这些资源分布在集群的多台机器上，在数据分布均匀的前提下，单个实例在单台机器的资源配额是集群配/集群机器数。因此，集群配合控制的执行是在每台机器上，配额是集群配额值均分到每台机器的值。

5 实例规格与性能说明

| 版本 | 容量 | 类型 | 性能(QPS) |
|-----|----------------------------|------|------------|
| 主从版 | 1~60G | 主从热备 | 4000 次/G/秒 |
| 集群版 | 60~300G (需要更大容量可提交工单申请) | 主从热备 | 4000 次/G/秒 |

注：随着容量增长，QPS 会有衰减，最大可达 15WQPS

6 案例

6.1 胡莱游戏

腾讯云 CRS——游戏业务爆发式增长的可靠保证。

CRS 应用：

- 分布式服务摆脱单机容量限制，支持全区全服架构
- 按 Key 备份、回档支持对游戏按人物等特定场景不停服回档
- 根据游戏生命周期需求，一键扩容/缩容，快速开服

6.2 小红书

高性能、高可靠腾讯云 CRS，提供完美用户体验，提升同行竞争力。

CRS 应用：

- 集群版提供分布式服务，性能通过业务的高并发验证
- 主从热备的极高稳定性提供稳定接入保证
- 完善的监控告警机制，保证数据的高可用性

6.3 e 袋洗

腾讯云 CRS——海量业务访问的可靠保障

CRS 应用：

- 主从版完美兼容 Redis 协议，百万级 Key 交并集运算性能验证
- 一主多备、异地接入，自动容灾，运维无忧
- 数据存储安全，读写效率高，回档机制健全

7 服务等级协议

7.1 数据存储的持久性

1.1 数据存储的持久性是指在合同期内数据保存不丢的概率，即每月完好 key-value 数量 / (每月完好 key-value 数量 + 每月丢失 key-value 数量)。以自然月为统计周期，不满一个月按一个月计。

1.2 云存储 Redis 的持久性为 99.99%。意味每月用户存储 10000 个 key，即每月只有 1 个丢失的可能性。

7.2 数据可销毁性

2.1 数据可销毁性是指用户要求删除云存储 Redis 表数据时，数据将从内存以及磁盘中删除，并且原磁盘使用空间将会被覆盖，从而实现彻底删除，无法复原。

2.2 下线的存储服务器报废时，磁盘通过消磁销毁里面的数据。

7.3 数据可迁移性

云存储 Redis 提供表数据导入和导出功能。导入和导出的数据文件按照云存储 Redis 的备份文件格式组织。文件组织格式可以提供给用户。

7.4 数据私密性

云存储 Redis 通过鉴权中心防火墙机制，限制只能被同一云账号的服务器访问，从而隔离其他账号的非法访问。在无用户授权的情况下，腾讯云不能查看用户存储的数据，不能查看用户的操作流水。

7.5 数据知情权

用户对存储在云存储 Redis 的数据有以下知情权：

- (1) 数据存储的数据中心位置（可以通过企业 QQ 咨询）；
- (2) 数据备份数量以及备份数据存储的数据中心位置（可以通过企业 QQ 咨询）；
- (3) 用户有权选择本身数据的数据中心位置，无权选择备份数据的数据中心位置；
- (4) 存储数据所在地与数据中心相关法律（可以通过企业 QQ 咨询）；

(5) 用户所有数据不会提供给任意第三方，除政府监管部门监管审计需要。

(6) 云存储 Redis 不会分析用户存储的数据。

7.6 数据可审查性

腾讯云在依据现有法律法规体系下，出于配合政府监管部门的监管或安全取证调查等原因的需要，在符合流程和手续完备的情况下，可以提供云存储 Redis 相关信息，包括关键组件的运行日志、运维人员的操作记录、用户操作记录等信息。

7.7 业务功能

云存储 Redis 提供购买、数据增删改查、数据导入、数据导出、扩容、自助清理、查看监控信息等功能。所有功能均已提供详细的功能介绍和使用说明文档。每项会影响用户数据结果的功能变更均通过企业 QQ、电话或者邮件方式通知用户。

7.8 业务可用性

1、业务可用性是指在合同期内云存储 Redis 可用的概率，即 $(\text{总服务时间}-\text{无法提供正常服务时间}\times\text{受影响的访问比例})/\text{总服务时间}\times 100\%$ 。以自然月为统计周期，不满一个月按一个月计。无法提供正常服务的时间定义为从业务无法访问云存储 Redis 算起，到云存储 Redis 服务恢复至正常水平结束的时间如果超过 2 分钟则计入无法提供正常服务的时间。

2、受影响的访问比例定义为 $(\text{故障前 5 分钟的访问数量}-\text{故障时成功访问的数量})/\text{故障前 5 分钟的访问数量}\times 100\%$

3、业务更换接入机，或者接入机死机但是业务自身没有使用云存储 Redis 提供的多台接入机做容灾而导致不可访问，或者其他业务自身的原因而导致的不可访问，不计入无法提供正常服务时间。

4、云存储 Redis 提供 99.95% 的可用性。即业务每月的数据可用时间应为 $30 \times 24 \times 60 \times 99.95\%$ 分钟 = 43178.4 分钟，即存在 $43200 - 43178.4 = 21.6$ 分钟数据不可以访问的可能性。

7.9 业务资源调配能力

腾讯云承诺用户，申请计算资源扩容时，现有资源 50% 以下容量，并且扩容实例资源小于 10 个，1 小时完成；申请资源小于 30 个，24 小时内完成；多于 30 个，请联系企业 QQ 咨询完成时间。每次单用户最大可扩展到 200G 容量，最小 1GB 的容量，若有更大容量需求，请联系企业 QQ 咨询扩容。

7.10 故障恢复能力

腾讯云提供专业团队 7x24 小时帮助维护。