

宝兰德分布式缓存软件 V3.2 白皮书

北京宝兰德软件股份有限公司

Beijing Baolande Software Corporation

版权所有 侵权必究

All rights reserved

目 录

第 1 章 概述	2
1.1 产品介绍	2
1.2 产品架构	3
1.3 应用场景	4
第 2 章 产品特性	5
2.1 多种部署模式	5
2.1.1 主从模式	5
2.1.2 哨兵模式	5
2.1.3 集群模式	6
2.2 高性能代理	6
2.3 OPERATOR 工具	7
2.4 持久化	7
2.5 数据安全	8
2.6 高可用	9
2.7 数据迁移	9
2.8 监控运维	9
2.9 多租户	10
第 3 章 产品优势	11
3.1 高效性能	11
3.2 强大的管理工具	11
3.3 安全可靠	11
3.4 灵活部署	11
3.5 云原生适配	11
3.6 平滑替换 REDIS	11
第 4 章 产品价值	12
4.1.1 提升业务性能	12
4.1.2 大幅降低成本	12
4.1.3 简化业务实现难度	12
第 5 章 联系我们	13

第1章 概述

早期的 Web 应用因为访问人数不多，并发量较小，使用关系型数据库即可满足需求，随着互联网迅猛发展，用户规模和数据规模越来越大，对 Web 网站的性能提出了更高的要求，使用关系型数据的 Web 网站因磁盘 I/O 的读写瓶颈，已经无法满足业务快速发展的需求，对性能提出了更高的要求，主要体现在以下几个方面：

1、低延迟的读写

需要使用更低延迟的读写速度，帮助 Web 网站高速响应用户请求，提升用户满意度。

2、支撑海量数据

当前大数据盛行的时代，需要能承受 PB 级别的数据检索和百万级别的流量。

3、庞大的运营成本考量

IT 部门希望能大幅降低企业的硬件成本、软件成本和运维成本。

为了解决以上问题，缓存技术应运而生，它具有以下特点：

1、高性能

使用更灵活的数据储存格式，数据结构简单，并通过缓存在内存中，实现快速更新和查询使用，可达毫秒级。

2、高扩展性

去掉关系型数据库的关系型特性，数据之间无关系，在架构层面非常容易扩展。

3、高可用

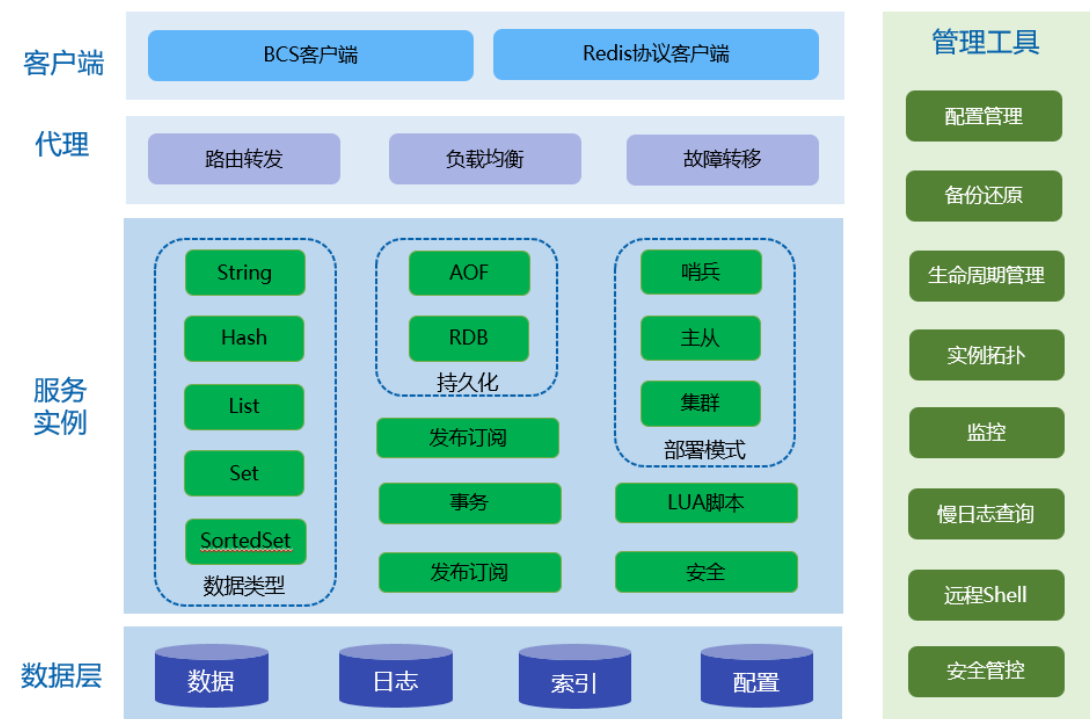
提供主从多节点的部署方式，在主从节点之间同步数据，保证多节点之间数据的一致性，在主节点故障后，由从节点自动接管，对外提供数据查询和储存服务。

1.1 产品介绍

BES CacheServer（简称 BCS）是一款宝兰德自研的分布式高性能 KV 存储数据库，完全兼容 Redis 协议标准。支持基于内存和文件的持久化存储，保证数据的安全可靠；使用标准 C 编写，将所有数据加载内存中，速度非常快；支持包括字符串（strings），散列（hashes），列表（lists），集合（sets），有序集合（sorted sets）在内的多种数据结构；支持单机、主从、集群等多种部署架构；支持数据分片，数据分布在多个实例中；

客户端支持 Python、Java、Lua、PHP、Ruby、Perl 等多种语言。主要解决高并发、大数据量场景下的数据访问性能问题。

1.2 产品架构



• 客户端

BCS 提供客户端与缓存实例进行通讯，读写缓存实例中的数据，用户也可以使用符合 Redis 协议的客户端。

• 缓存实例

BCS 实例支持多种数据类型，包括 String、Hash、List、Set、SortedSet，支持包括 RDB、AOF 等多种持久化策略，支持单实例、主从、哨兵和集群等多种部署架构，用户可以按需选择。支持对实例通信通道进行加密，加密协议包括 TLS、国密等，支持对实例密码进行加密储存等。

• 管理中心

提供可视化的管理中心，批量管理所有实例的配置、生命周期，并提供详细的监控性能视图，帮助用户轻松管理大规模的实例。

• 数据层

数据持久化到磁盘、数据库中。

1.3 应用场景

1、缓存热点数据

所有大中型网站都在使用的技术，不仅能够提高网站的访问速度，还能大大降低数据的压力。

2、计数器

提供自增命令实现计数器功能，内存操作，性能非常好。

3、Session 共享

在分布式应用系统中，Session 统一储存在内存数据库中，方便快速读写和集中管理，防止应用单点故障导致的 Session 丢失。

4、分布式锁

利用 `setnx` 功能编写分布式锁，如果返回 1 说明获取锁成功，否则失败。

5、排行榜

提供有序集合数据实现各种复杂的排行榜应用。

6、社交网络

使用哈希、集合等数据结构能很方便的实现社交网站的点赞，粉丝，共同爱好，推送，下拉刷新等基本功能

7、消息队列

提供发布/订阅及阻塞队列功能，实现一个简单的消息队列。

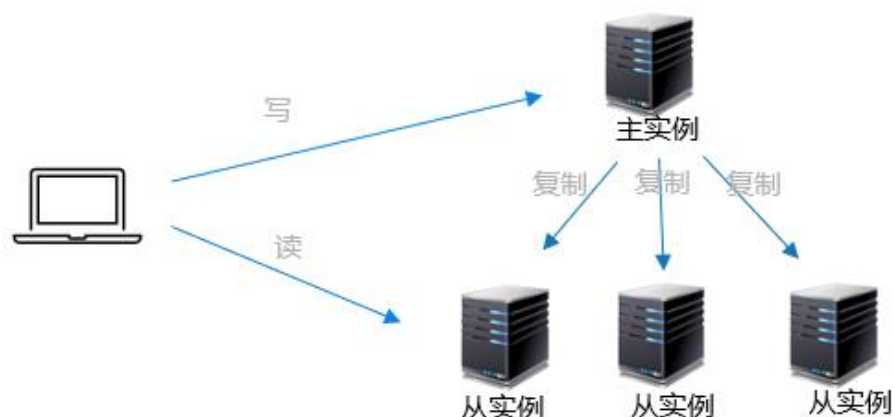
第2章 产品特性

2.1 多种部署模式

2.1.1 主从模式

每个主实例下有一个以上的从实例，主从模式能够实现读写分离，主实例对外提供写操作，从实例对外提供读操作，降低主实例的并发访问量。从实例在主实例宕机或者网络断开后，需要手动将从实例升级为主实例，继续对外提供服务。

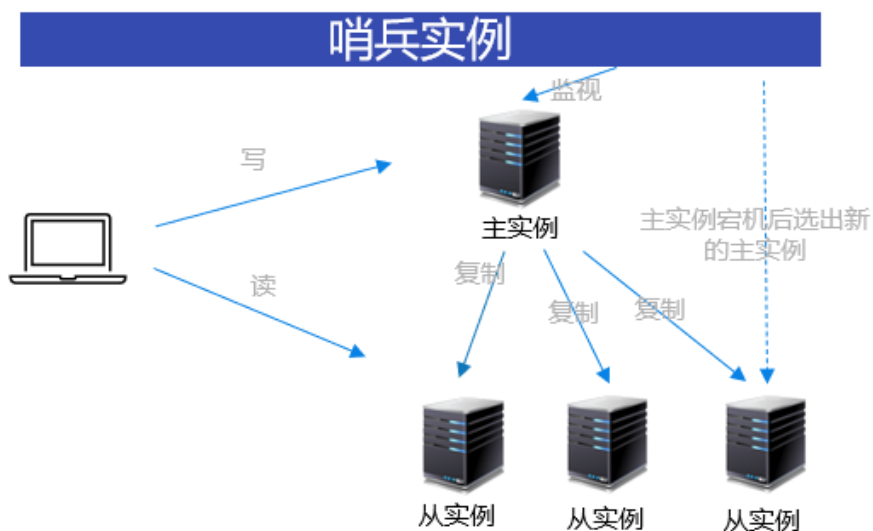
- 主实例：对外提供读写服务，并向从实例复制数据。
- 从实例：对外提供读服务，接收来自主实例的数据。如果配置了哨兵实例，在主实例宕机后会有一个从实例升级为主实例。



2.1.2 哨兵模式

和主从模式相比，多了哨兵实例（可选），哨兵实例在监控到主实例宕机或者网络断开后，自动将其中一个从实例升级为主实例，继续对外提供服务，保障服务的高可用性。

- 主实例：对外提供读写服务，并向从实例复制数据。
- 从实例：对外提供读服务，接收来自主实例的数据。如果配置了哨兵实例，在主实例宕机后会有一个从实例升级为主实例。
- 哨兵实例：不断地检查主实例和从实例是否运作正常。当主实例不能运行正常时，进行自动故障迁移，升级从实例为主实例。

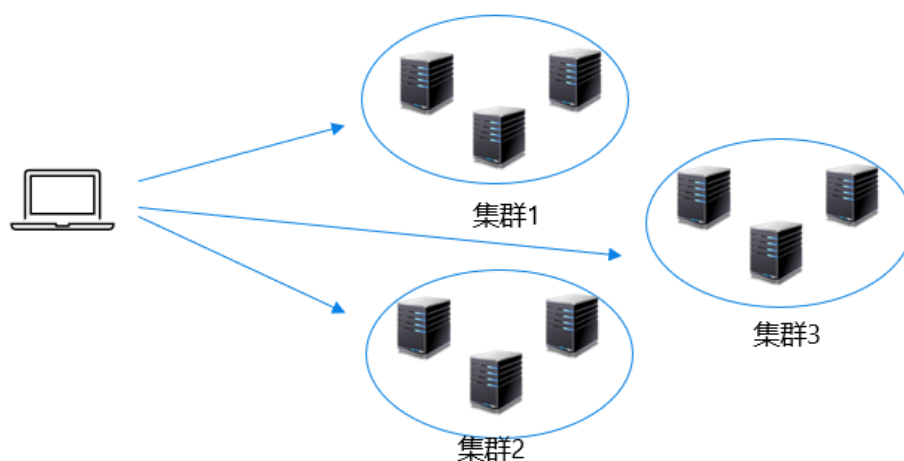


2.1.3 集群模式

集群采用分区模式，将数据由之前的全部分配到一个节点改为自动分配多个节点上，当一部分实例宕机或不可达，集群依然能继续工作。包含多个主实例，多个从实例。

主实例：对外提供读写服务，并向从实例复制数据。某个宕机主实例的从实例发起升级选举时，进行投票。

从实例：对外提供读服务，接收来自主实例的数据。在主实例宕机后，从实例会发起一次选举，获得多数选票的从实例将会升级为主实例。



2.2 高性能代理

提供代理服务器，提供路由转发、负载均衡与故障转移等能力，帮助用户设计更高效

的业务架构

- 1、读写分离：控制写请求转发给主节点，读请求按照比例转发给从节点。主从实例支持自定义设置读权重，代理根据负载均衡策略，将请求转发到不同的数据实例中。
- 2、负载均衡和路由转发：代理和后端的数据实例建立连接，负责请求负载均衡和路由转发操作。操作单个 Key 命令时，会根据 Key 所属的槽位，将请求发送给所属的数据分片。操作多个 Key 命令时，如果这些 Key 储存在不同的数据实例上，代理会拆分成多个命令分别发送给对应的分片。
- 3、强大的权限控制机制：支持自定义不同用户的权限，权限分为读、写、管理三种权限，每类权限可以指定生效的 Key 前缀，方便用户隔离不同业务系统的数据操作权限。
- 4、多数据中心：支持定义数据实例的数据中心概念，在 BCS 跨数据中心部署时，优先将读请求分发给本数据中心的数据实例，避免跨数据中心访问带来的性能延迟。
- 5、丰富的统计信息，包括 CPU、内存、请求和响应等。

2.3 Operator 工具

随着云原生技术的蓬勃发展，大量业务上云，作为有状态的缓存服务，在上云后无法平滑的做容器化部署，宝兰德分布式缓存软件提供 Operator 工具，实现在云原生环境上的自动化安装部署、动态扩缩容和备份还原等能力，从而极大地简化了应用程序的部署和管理，大幅提升应用程序的可靠性和安全性，并减少故障率和安全漏洞，为用户创造更高效、更安全、更稳定的运维环境。

- 1、支持主从、哨兵和集群架构的快速部署，以 StatefulSet 部署有状态的 BCS 实例，并提供挂载持久化数据的方案，在实例漂移后恢复对应的数据。
- 2、横向扩缩容，自动调整分片实例数量。
- 3、纵向扩缩容，自动调整实例的内存容量。
- 4、提供备份还原机制，定期对实例数据进行备份，并在发生故障后，通过备份数据还原。

2.4 持久化

提供了两种持久化方式：RDB 和 AOF，其中 RDB 持久化可以在指定的时间间隔内生成数据集的时间点快照（point-in-time snapshot）。AOF 持久化记录服务器执行的所有写操作命令，并在服务器启动时，通过重新执行这些命令来还原数据集，新命令会被追加到文件的末尾，还可以在后台对 AOF 文件进行重写（rewrite），使得 AOF 文件的体积不会超出保存数据集状态所需的实际大小。

可以同时使用 AOF 持久化和 RDB 持久化。在这种情况下，当 BES CacheServer 重启时，它会优先使用 AOF 文件来还原数据集，因为 AOF 文件保存的数据集通常比 RDB 文件所保存的数据集更完整。

	RDB	AOF
优点	非常紧凑的文件，保存了某个时间点的数据集，非常适用于数据集的备份	每次发生数据变更会被立即记录到磁盘，性能较差但数据完整性比较好
	方便传送到另一个远端数据中心，非常适用于灾难恢复	
	在保存时父进程只需要 fork 出一个子进程，由子进程来做工作，性能较好	
	比 AOF 方式速度更快	
缺点	意外停止工作后，会丢失部分数据	文件体积大于 RDB 文件
	当数据集比较大的时候，会短暂影响响应客户端请求	速度慢于 RDB

2.5 数据安全

- 通过多种手段保障数据的安全性，包括：
- 权限控制
- 完善的用户权限定义和授权，有效隔离数据，限制用户操作，提升系统数据安全性。
- SSL 加密访问
- 支持启用 SSL 加密，在传输层对网络连接进行加密，在提升通信数据安全性的同时，保证数据的完整性。加密算法支持 SM2、SM3 和 SM4 等。
- 密码加密
- 对配置中的密码信息进行加密保存，避免敏感信息泄露导致的安全隐患。
- 审计日志
- 完整的记录用户的操作，方便事后追溯，审计违规行为，有效保护系统的安全性。

2.6 高可用

通过主从模式提供高可用能力，避免实例单点故障导致的服务中断，节点故障后，自动发现并由健康节点对外提供服务。

2.7 数据迁移

提供跨产品异构系统的数据迁移能力，帮助客户快速完成开源 Redis 替换、系统升级等。

- 1、提供一键式迁移外部 Redis 和 Redis 持久化文件的能力，方便用户快速完成 BCS 替换过程。
- 2、提供一键式迁移不同版本、不同架构的 BCS 实例数据的能力，包括主从变更为集群，集群 2 个分片变更为 4 个分片等。

2.8 监控运维

提供可视化管理工具，支持管理上千实例的生命周期和监控指标的采集，包括：

- 1、提供一键式的创建主从、哨兵和集群实例的配置向导，方便用户简单快速地创建出大规模的实例
- 2、提供完整的生命周期的管理，包括启动、停止、卸载、槽位转移、失败转移、加入集群、离开集群等多种命令的执行，方便用户批量管理大规模实例，简化运维难度。
- 3、提供完善的监控采集和展示，包括资源使用，如：CPU 使用率、内存使用率，命令执行情况，如：各类命令执行数、平均响应时间等，Key 统计情况，如：Key 总数、过期 Key 数、命令 Key 情况等，网络流量使用情况，如输入字节数、输出字节数等，客户端使用情况，如：活跃、堵塞的客户连接等，并提供标准的 API 接口对外开放。
- 4、可视化的配置管理，包括监听地址、监听端口、日志、持久化、主从复制、慢日志、哨兵监控、集群等配置项。
- 5、提供远程 WebShell 执行窗口，免登录远程执行运维命令
- 6、展示实例运行状态、运行角色和关系的实例拓扑图，实例关系包括主实例和从实例的关系、哨兵实例监控的主从实例关系等。

2.9 多租户

提供基于多 DB 的租户数据隔离机制，不同业务系统之间共用一套缓存系统，降低系统使用的资源成本。

多个业务系统混合存储在一起，没有做数据隔离时，不同系统的数据存在冲突的可能性，影响读写数据操作的性能，通过写到不同的 DB 上，避免了相互之间的干扰，大幅降低了资源的使用。

第3章 产品优势

3.1 高效性能

支持集群架构下弹性扩展数据库系统的存储空间及吞吐性能，突破海量数据高 QPS 性能瓶颈，轻松应对每秒数十万次的读写需求。

3.2 强大的管理工具

提供强大的可视化的管理工具，管理大规模实例的生命周期、配置、监控等，帮助用户提高运维效率，简化运维难度。

3.3 安全可靠

通过对通讯通道加密、密码加密储存、细粒度的访问权限控制和审计日志等多种安全保障机制，保护缓存数据被安全传输和储存，提供全面的安全保障。

3.4 灵活部署

提供主从、哨兵和集群多种部署模式，并支持高可用、数据分片等特性，满足客户不同应用场景下的要求。

3.5 云原生适配

提供容器化镜像和配套的 Operator 工具，搭配 Kubernetes 高性能、自动化运行在云原生环境上，帮忙客户实现业务系统的快速上云。

3.6 平滑替换 Redis

产品在规划设计时，充分考虑了对 Redis 协议的兼容适配性，确保原使用 Redis 作为缓存服务的应用，在不修改代码的情况下能够无感平滑迁移到本产品中。

第4章 产品价值

4.1.1 提升业务性能

BES CacheServer 使用缓存技术，帮助业务大幅提升业务性能，用户能够对外提供更好的用户体验，从而获得更大的商业成功。

4.1.2 大幅降低成本

BES CacheServer 高性能、高可用性、易维护等特别帮助用户降低采购设备和运营维护的人力成本，减少企业开支。

4.1.3 简化业务实现难度

BES CacheServer 能很方便的帮助用户快速实现复杂的业务场景需求，比如排行榜、Session 共享、分布式锁等，极大的简化了复杂业务的实现难度。

第5章 联系我们

北京宝兰德软件股份有限公司将为您提供专业、全面的技术支持服务。

了解关于宝兰德软件产品的更多内容，请通过以下方式联系我们：

网址：www.bessystem.com

Support Email: support@bessystem.com

Support Tel: 400 650 1976