



WXG技术能力提升

首页


团队文章

团队文档

团队活动

团队视频

团队文集



如何实现一个分布式定时器-ruoyuliu刘若愚.pdf

ruoyuliu

2021-04-26 20:49

大小(3MB)

收藏(2)

下载

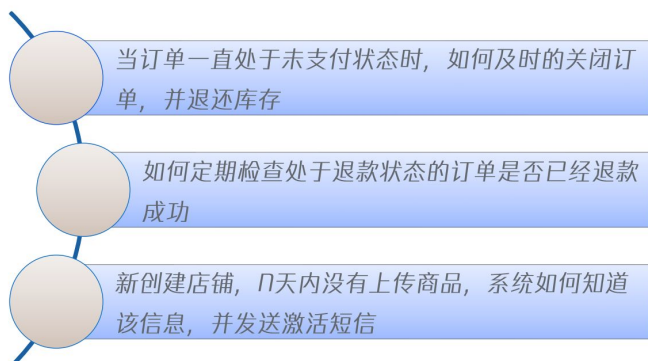
附件来源: [文章]如何实现一个分布式定时器



1 定时器是什么

1 定时器应用场景

你是否有遇到如下需求？



?

2 定时器是什么



本质
deadline越近的任务拥有越高优先级的一个数据结构

基本操作

- Add 新增任务
- Delete 删除任务
- Run 执行到期的任务/到期通知对应业务处理
- Update 更新到期时间 [可选]

Run的工作方式

1.轮询

每隔一个时间片就去查找哪些任务已经到期

2.睡眠/唤醒

- 不停地查找deadline最近的任务，如到期则执行，否则sleep
- 在sleep期间，如果有任务被Add或删除，线程会被唤醒并重新进行1的逻辑

3 设计目标

5点要求

- 支持任务提交〔消息发布〕、任务删除、任务通知〔消息订阅〕等基本功能
- 消息传输可靠性：消息进入延迟队列以后，保证至少被消费一次〔到期通知保证At-least-once，追求Exactly-once〕
- 数据可靠性：数据需要持久化，防止丢失
- 高可用性：至少得支持多实例部署。挂掉一个实例后，还有后备实例继续提供服务，可横向扩展
- 实时性：尽最大努力准时交付信息，允许存在一定的时间误差〔几百ms到秒级左右〕，误差范围可控

exactly-once的相关做法往往需要在接收端进行必要的处理，代价大

Any more? 高级特性：高性能、可伸缩...



6

2 业界实现方案



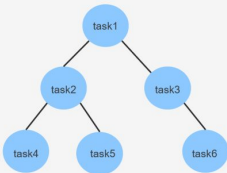
7

1 数据结构

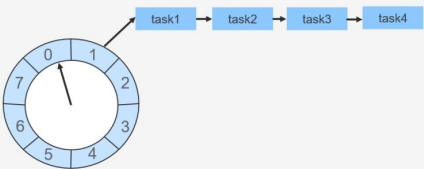
有序链表



堆



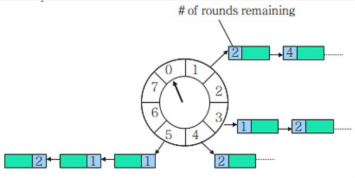
时间轮/多级时间轮



8

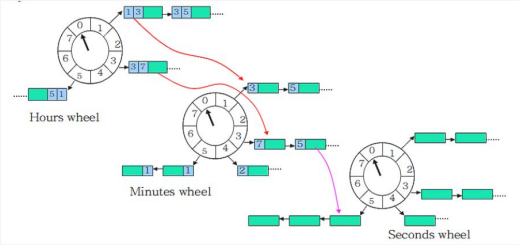
2 时间轮 [TimeWheel]

单级时间轮



如果任务的时间
跨度很大，数量
也多

多级时间轮



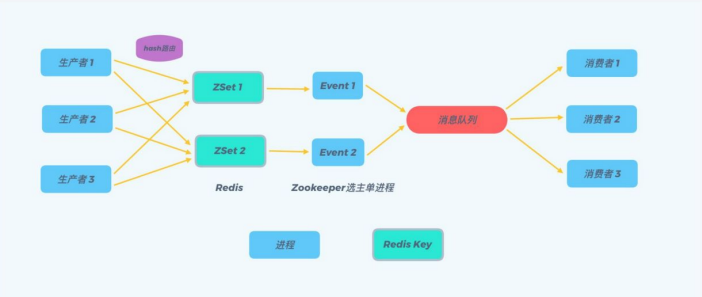
将Wheel按时间粒度分级，类似水表的思想

核心：利用**进制**的做法可
以使用较少的空间表示更
大的时间跨度

3 业界工程实践

1. 基于Redis ZSet

一个例子

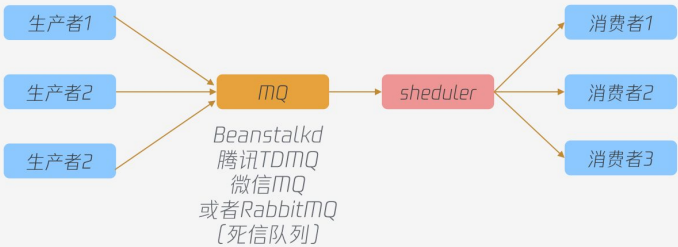


1. 消息hash到Redis key上
2. 每个key对应一个Event进程
3. Event进程负责分发消
息，消息队列负责投递
4. Event采用zookeeper选
主单进程部署

优点：理解起来较为直观，可以快速落地
可以依赖Redis自身的持久化来实现持久化
使用Redis集群来支持高并发和高可用

3 业界工程实践

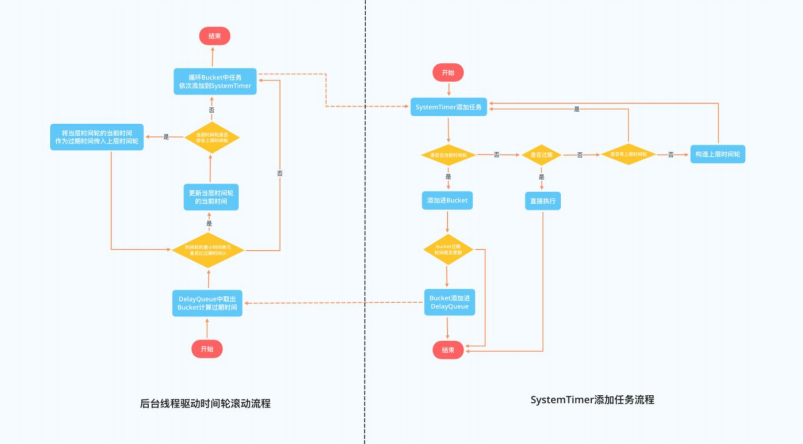
2. 采用某些自带延时选项的队列



优点：可以很好的利用一些MQ的特性，比如消息可靠发送、消息可靠投递
通过MQ集群的特性，可以很好的解决单点故障问题，不会因为单个节点挂
掉导致延迟队列不可用或者消息丢失

3 业界工程实践

3.基于Timing-Wheel时间轮算法



一个例子：
Kafka中的时
间轮

优点：能够支持
大量任务的高效
触发

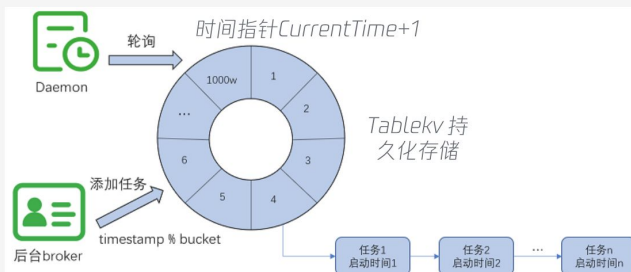
3 实现思想归类



3 本人实践

1 存储设计

存储数据结构



tablekv

微信kv又称paxosstore, 天然支持按uin分表, 是一种支持单表亿行、千万以上分表, 低延迟低成本, 强一致高可用的分布式存储系统

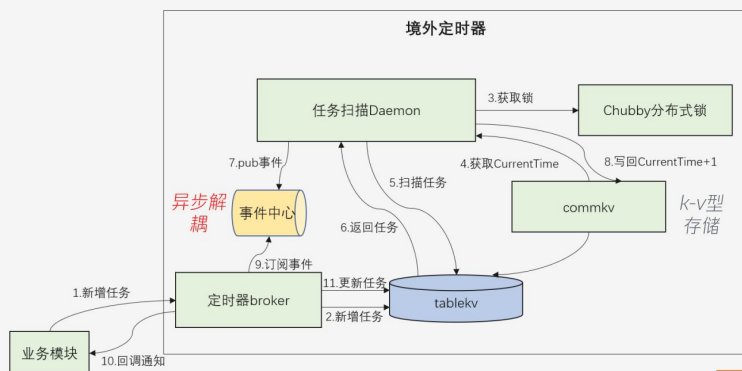
分表方式: 按任务执行时间表
调度算法: 时间轮
精度: 秒级

? 为什么不使用多轮时间轮?
为什么要mod分表?



15

2 架构设计



任务扫描Daemon: 充当Scheduler的角色。扫描所有到期任务, 投递到事件中心, 让它通知broker, 由broker的Notifier通知业务方

定时器broker: 集业务接入、Notifier两者功能于一身

事件中心

微信可靠事件中心微信内部的高可用、高可靠的消息平台, 支持事务和非事务消息



16

4 重点&难点问题探究



17

1 时间轮空转问题

1000万分表，大部分Bucket为空，时间轮的指针推进存在低效问题如何解决？



Kafka引入了DelayQueue来解决这个问题，避免线程空转



优化：维护一个有序队列，保存不为空的bucket



2 业务隔离

目的：业务隔离的目的在于使各个接入业务不受其他业务的影响



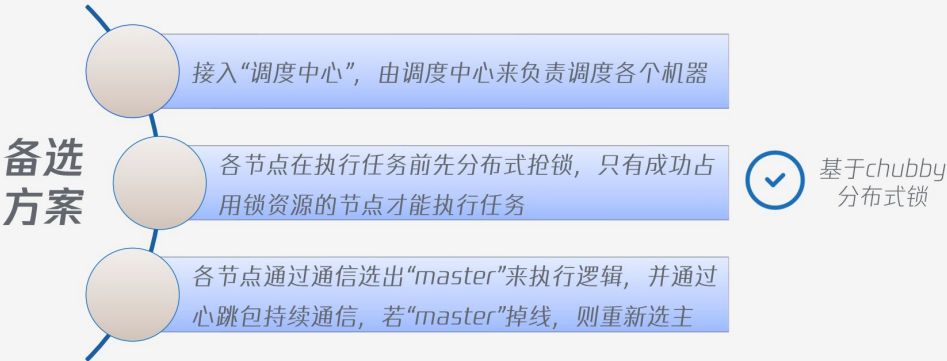
措施：对不同业务执行不同限频策略



3 分布式单实例容灾

目的：Daemon具有容灾能力，若有Daemon实例异常挂起或退出，其他机器的实例进程可以继续执行任务。但同时同一时刻只允许一个实例运行

这里这允许一个实例运行的原因是避免了多进程对一个时间桶进行操作的并发问题



我顶 (0) 收藏 (2)

Copyright©1998-2021 Tencent Inc. All Rights Reserved

腾讯公司研发管理部 版权所有

[广告申请](#) [反馈问题](#)

[56/73/341 ms]