1. 复制
2. 概述

Redis中，用户可以通过执行SLAVE OF 命令或者设置slaveof选项，让一个服务器去复制另一个服务器，我们称呼被复制的服务器为主服务器，另一个服务器为从服务器。进行复制中的主从服务器双方的数据库将保存着相同的数据，这种现象称为数据库状态一致。Redis2.8版本更新了复制功能，这里我们会介绍旧版和新版的复制功能。

1. 旧版复制功能
2. 概述

旧版的复制功能分为同步和命令传播两个操作。

1. 同步

将从服务器的数据库状态更新至主服务器当前所处的数据库状态。

1. 命令传播

在主服务器的数据库状态被修改，导致主从服务器的数据库状态不一致时，让主从服务器的数据库重新回到一致状态。

1. 同步

同步操作通过向主服务器发送SYNC命令完成，当主服务器收到该命令后，执行BGSAVE命令，在后台生成一个RDB文件，并使用一个缓冲区记录从现在开始执行的所有写命令。当BGSAVE命令执行完毕时，主服务器将RDB文件发送给从服务器，从服务器接受这个RDB文件，更新自己的数据库状态到主服务器执行BGSAVE时的状态，然后主服务器会将缓冲区中的所有写命令发送给从服务器，从服务器执行这些命令，更新数据库状态。

1. 命令传播

同步之后主从服务器的一致状态并不是一成不变的，每当主服务器都执行客户端发送的写命令后，主从服务器状态就会不一致了。因此主服务器会对从服务器执行命令传播操作，将自己执行的写命令不断发送给从服务器执行，这样两者的状态又会恢复一致了。

1. 旧版复制缺陷

旧版的复制分为两种情况：初次复制和断线后重复制。

1. 初次复制

从服务器以前没有复制过任何主服务器，或者这次的主服务器与上次不同。

1. 断线后重复制

处于命令传播阶段的主从服务器因为网络原因中断了复制，从服务器通过自动重连接连上了主服务器，并继续复制。这种情况的效率很低，因为重连之哦户，从服务器会向主服务器再次发送SYNC命令，主服务器就会将包含它所有数据的RDB文件发送给从服务器，而并不一定要把之前的所有数据都发送给它，因为从服务器之前已经有了部分的数据了，因此这种操作很低效。

1. 新版复制功能
2. 概述

新版的Redis采用PSYNC命令来代替SYNC命令执行同步操作，有完整重同步和部分重同步两种模式。完整重同步用于初次复制，部分重同步用于断线后重复制。

1. 实现

部分重同步由三个部分组成：主从服务器的复制偏移量，主服务器的复制积压缓冲区，服务器的运行ID。

1. 复制偏移量

主服务器每次向从服务器中传播N个字节的数据时，就将自己的复制偏移量加上N，同理，从服务器接受后也会加上N，这个属性用于对比主从服务器的一致性。

1. 复制积压缓冲区

主服务器维护的一个固定长度的FIFO队列，默认大小1MB，当主服务器进行命令传播时，不仅将写命令发送给所有的从服务器，还会入队到复制积压缓冲区中，保存着一部分最近传播的写命令，当从服务器重连后，从服务器通过PSYNC命令将自己的复制偏移量offset发送给主服务器，如果offset之后的数据仍然存在于复制缓冲区中，则进行部分重同步；否则进行完整重同步操作。

这个区域的大小是可以调整的，可以根据断线重连的情况来尽量调大一点。

1. 服务器运行ID

每个Redis服务器都有自己的运行ID，当从服务器对主服务器进行初次复制时，主服务器会将自己的运行ID发送给从服务器，从服务器会保存起来。之后从服务器重连时，会将该运行ID发送给当前连接的主服务器，如果一致则尝试进行部分重同步，否则完整重同步。

1. PSYNC命令

该命令的调用方法有两种，一种是开启新的复制，向主服务器发送PSYNC ? -1命令，主动请求完整重同步；另一种是从服务器已经复制过某个服务器了，在一次新的复制时就会发送PSYNC <runId> <offset>命令，runID是保存的上一次的主服务器ID，offset是复制偏移量。

如果主服务器返回 +FULLRESYNC 以及该服务器的运行ID以及复制偏移量，则进行完整重同步；

如果返回+CONTINUE，则表示执行部分重同步；

如果返回-ERR，则表示主服务器版本低于2.8，识别不了PSYNC，从服务器会发送SYNC命令，进行完整重同步。

1. 复制的实现

通过向从服务器发送SLAVEOF命令，从服务器可以复制主服务器。

1. 设置主服务器的地址和端口

客户端向从服务器发送SLAVEOF命令后，SLAVEOF命令将主服务器的IP和端口设置到服务器状态的masterhost属性和masterport属性中。

1. 建立套接字连接

从服务器根据命令设置的IP和端口，创建连向主服务器的套接字连接，如果成功连接，从服务器将为这个套接字关联一个专用于处理复制操作的文件事件处理器。

而主服务器接受这个套接字连接后，将为该套接字创建相应的客户端状态，并将从服务器当做一个客户端看待。

1. 发送PING命令

从服务器成为主服务器的客户端之后，第一件事就是向主服务器发送一个PING命令，该命令两个作用，检查套接字读写状态是否正常，检查主服务器能否正常处理命令请求。如果从服务器收到PONG回复表示连接正常，否则断开重连。

1. 身份验证

收到PONG回复后，决定是否进行身份验证，根据从服务器设置的masterauth选项，向主服务器发送一条AUTH命令。

1. 发送端口信息

从服务器向主服务器发送自己的监听端口号，主服务器收到命令后，会记录端口号到从服务器对应的客户端状态的监听属性中。

1. 同步

从服务器向主服务器发送PSYNC命令，执行同步操作，在这之后，主从服务器双方都会变为对方的客户端。

1. 命令传播

主服务器一直将自己执行的写命令发送给从服务器即可。

1. 心跳检测

命令传播阶段，从服务器默认一秒一次向主服务器发送命令 REPLCONF ACK <offset>，offset是从服务器当前的复制偏移量，检测连接状态，检测命令丢失情况。

1. 检测连接

通过向主服务器发送INFO replication命令，可以列出从服务器的信息，lag属性表示延迟，应该在0-1秒之间，否则就是连接出错。

1. 检测命令丢失

如果因为网络故障，主从之间的命令传播丢失了，那么当从服务器向主服务器发送REPLCONF ACK时，如果携带的offset和主服务器的不一致，说明数据丢失了，这样主服务器就会从复制积压缓冲区中找到丢失的数据重新发送。

REPLCONF ACK命令和复制积压缓冲区都是Redis2.8新增的，为了数据一致性请使用2.8及以上版本。