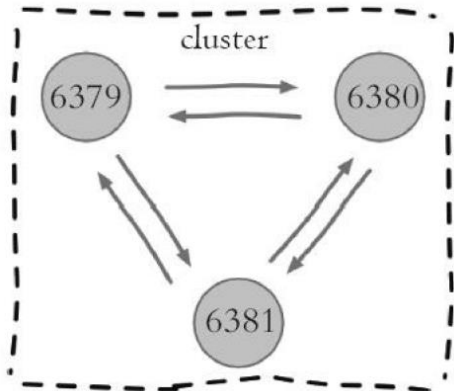


牛客大学高薪加成系列课

集群中节点的通信

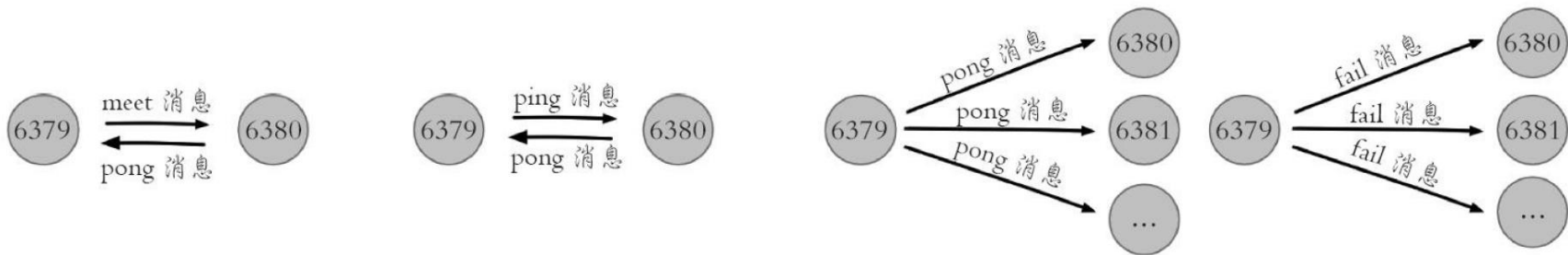


- 在分布式存储中需要提供维护节点元数据信息的机制，所谓元数据是指：节点负责哪些数据，是否出现故障等状态信息。常见的元数据维护方式分为：集中式和P2P方式。
- Redis集群采用P2P的Gossip（流言）协议，Gossip协议的工作原理就是节点彼此不断通信交换信息，一段时间后所有的节点都会知道集群完整的信息，这种方式类似流言传播。
- 通信过程说明：
 1. 集群中的每个节点都会单独开辟一个TCP通道，用于节点之间彼此通信，通信端口号在基础端口号上加10000；
 2. 每个节点再固定周期内通过特定规则选择几个节点发送ping消息；
 3. 接收ping消息的节点用pong消息作为响应。



Gossip协议的主要职责就是信息交换，而信息交换的载体就是节点彼此发送的Gossip消息，Gossip消息分为：meet消息、ping消息、pong消息、fail消息等。

- meet消息：用于通知新节点加入，消息发送者通知接受者加入到当前集群。meet消息通信正常完成后，接收节点会加入到集群中并进行周期性的ping、pong消息交换。
- ping消息：集群内交换最频繁的消息，集群内每个节点每秒向多个其他节点发送ping消息，用于检测节点是否在线和交换彼此状态信息。ping消息封装了自身节点和一部分其他节点的状态数据。
- pong消息：当接收到meet、ping消息时，作为响应消息回复给发送方确认消息正常通信。pong消息内封装了自身状态数据，节点也可以向集群内广播自身的pong消息来通知整个集群对自身状态进行更新。
- fail消息：当节点判定集群内另一个节点下线时，会向集群内广播一个fail消息，其他节点接收到fail消息之后把对应节点更新为下线状态。



虽然Gossip协议的信息交换机制具有天然的分布式特性，但它是有成本的。因为Redis集群内部需要频繁地进行节点信息交换，而ping/pong消息会携带当前节点和部分其他节点的状态数据，势必会加重带宽和计算的负担。所以，Redis集群的Gossip协议需要兼顾信息交换的实时性和成本的开销。

- 集群里的每个节点默认每隔一秒钟就会从已知节点列表中随机选出五个节点，然后对这五个节点中最长时间没有发送过PING消息的节点发送PING消息，以此来检测被选中的节点是否在线。
- 如果节点A最后一次收到节点B发送的PONG消息的时间，距离当前时间已经超过了节点A的超时选项设置时长的一半（`cluster-node-timeout/2`），那么节点A也会向节点B发送PING消息，这可以防止节点A因为长时间没有随机选中节点B作为PING消息的发送对象而导致对节点B的信息更新滞后。
- 每个消息主要的数据占用包含：`slots`槽数组（2KB）和整个集群1/10的状态数据（10个节点状态数据约1KB）。



牛客大学

- 专业求职辅导 -

THANKS



关注【牛客大学】公众号
回复“牛客大学”获取更多求职资料