目录

[集群 1](#_Toc40966529)

[一、 节点 1](#_Toc40966530)

[(一) 启动节点 2](#_Toc40966531)

[(二) 集群数据结构 2](#_Toc40966532)

[(三) CLUSTER MEET的实现 4](#_Toc40966533)

[二、 槽指派 4](#_Toc40966534)

[(一) CLUSTER ADDSLOTS命令的实现 5](#_Toc40966535)

[三、 在集群中执行命令 6](#_Toc40966536)

[(一) 计算键的属于哪个槽 6](#_Toc40966537)

[(二) MOVED错误 6](#_Toc40966538)

[(三) 被隐藏的MOVED错误 7](#_Toc40966539)

[(四) 节点数据库的实现 7](#_Toc40966540)

[四、 重新分片 8](#_Toc40966541)

[五、 ASK错误 9](#_Toc40966542)

[(一) ASK错误和MOVE错误的区别 10](#_Toc40966543)

[六、 复制与故障转移 10](#_Toc40966544)

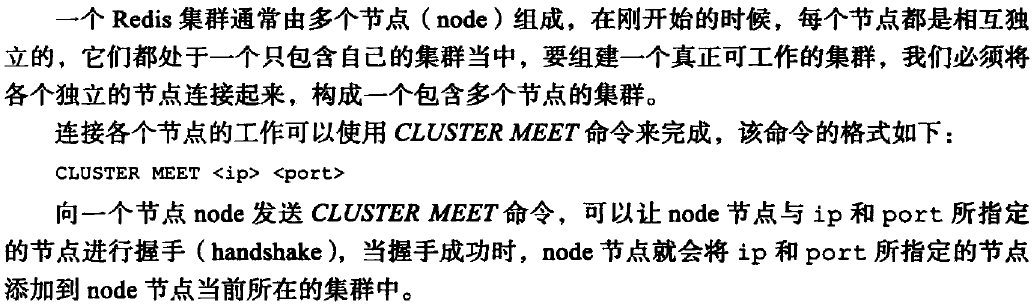
[(一) 设置从节点（CLUSTER REPLACE <node\_id>） 10](#_Toc40966545)

[(二) 故障检测 11](#_Toc40966546)

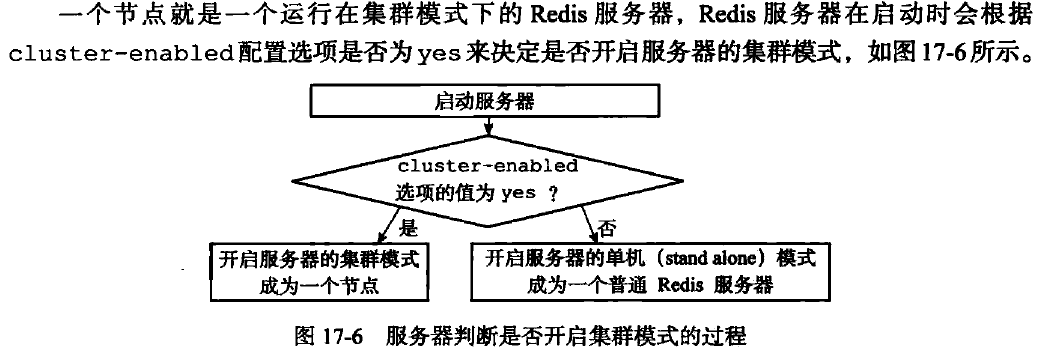
集群

Redis集群式redis提供的分布式数据库方案，集群通过分片（sharding）来进行数据库共享，并提供复制和故障转移功能。

# 节点



## 启动节点



## 集群数据结构

每一个节点都会使用一个clusterNode结构来记录自己的状态，并为集群中的所有其他节点都创建一个相应的clusterNode结构，以此来记录其他节点的状态：

Struct clusterNode{

mstime\_t ctime; //创建节点的时间

char name[REDIS\_CLUSTER\_NAMELEN]; //节点的名字，有40个十六进制字符组成

int flags; //节点标识，记录节点角色以及当前所处的状态

uint64\_t configEpoch; //系欸但当前的配置纪元，用于实现故障转移

char ip[REDUS\_IP\_STR\_LEN]; //节点的IP地址

int port; //节点的端口号

clusterLink \*link; //保存连接节点所需的有关信息

//…

}

clusterNode结构的link属性式一个clusterLink结构，该结构保存了连接节点所需的有关信息，比如套接字描述符，输入输出缓冲区：

typedef struct clusterLink{

mstime\_t ctim e; //连接创建的时间

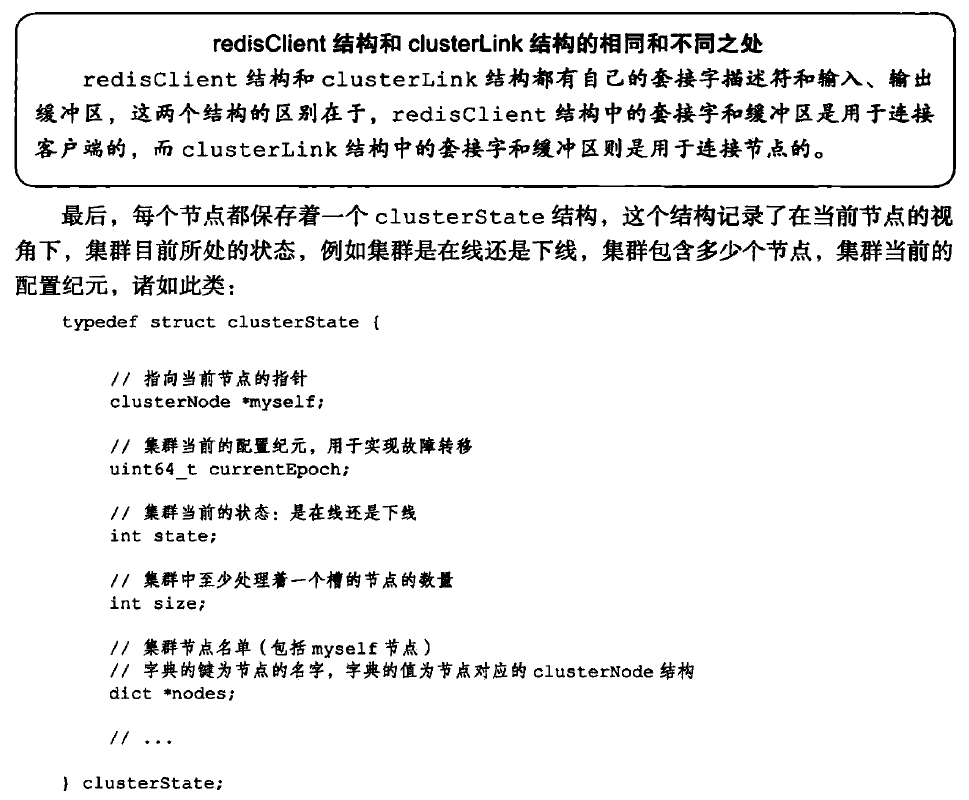
int fd; //TCP套接字描述符

sds sndbuf; //输出缓冲区

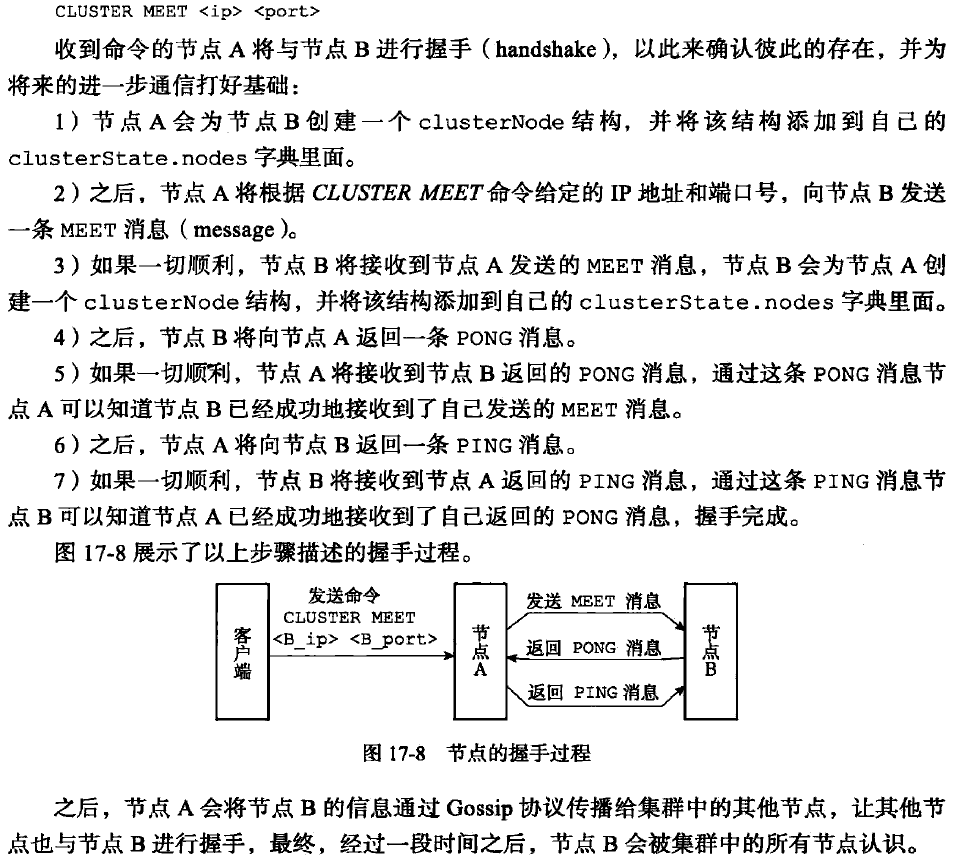
sds rcvbuf; //输入缓冲区

struct clusterNode \*node; //与这个连接相关联的节点，如果没有的化就为NULL

}clusterLink；



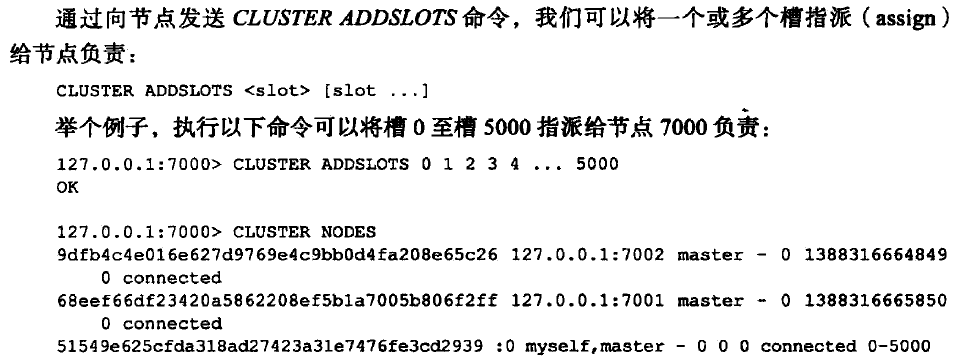
## CLUSTER MEET的实现

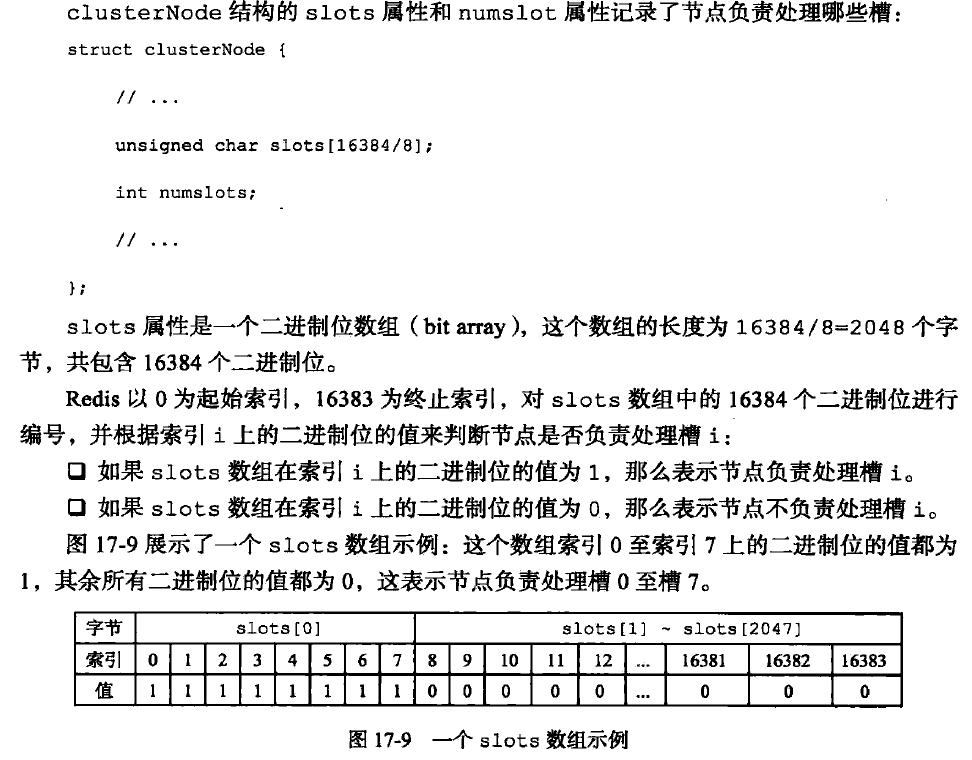


# 槽指派

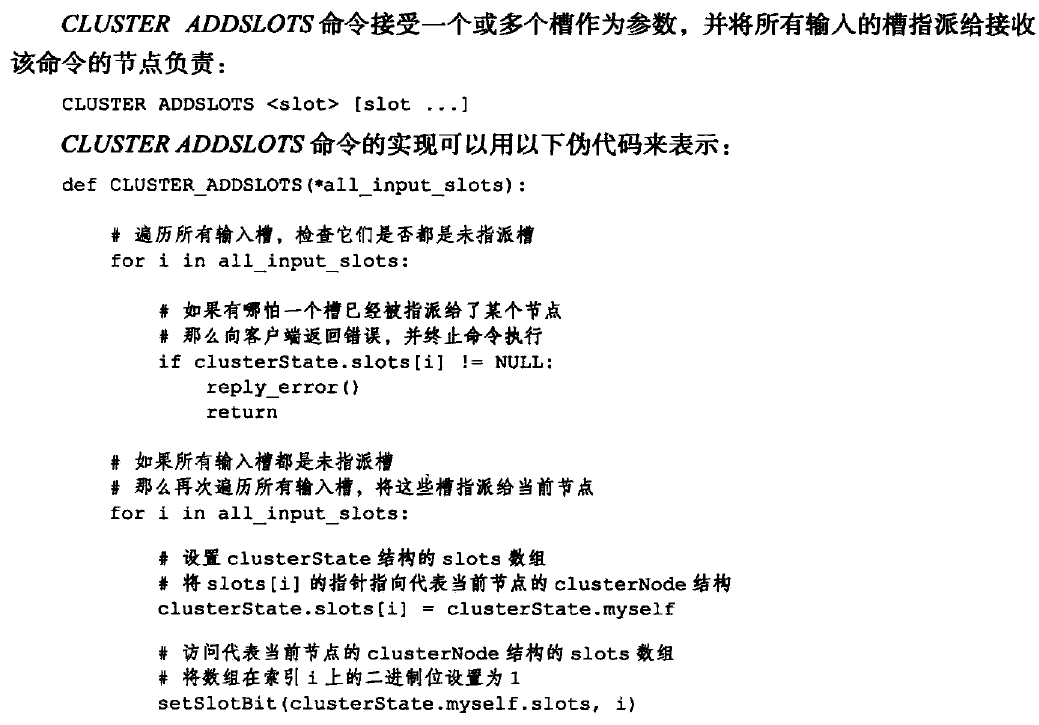
Redis集群通过分片的方式来保存数据库中的键值对：集群的整个数据库被分为16384个槽(slot)，数据库中的每个键都数据这16384个槽的其中一个，集群中的每个节点可以处理0个或作堆16384个槽。

当数据库中的16384个槽都有节点在处理时，集群处于上线状态（OK）；相反，如果数据库中有任何一个槽没有得到处理，那么进去处于下线转台（fail）。

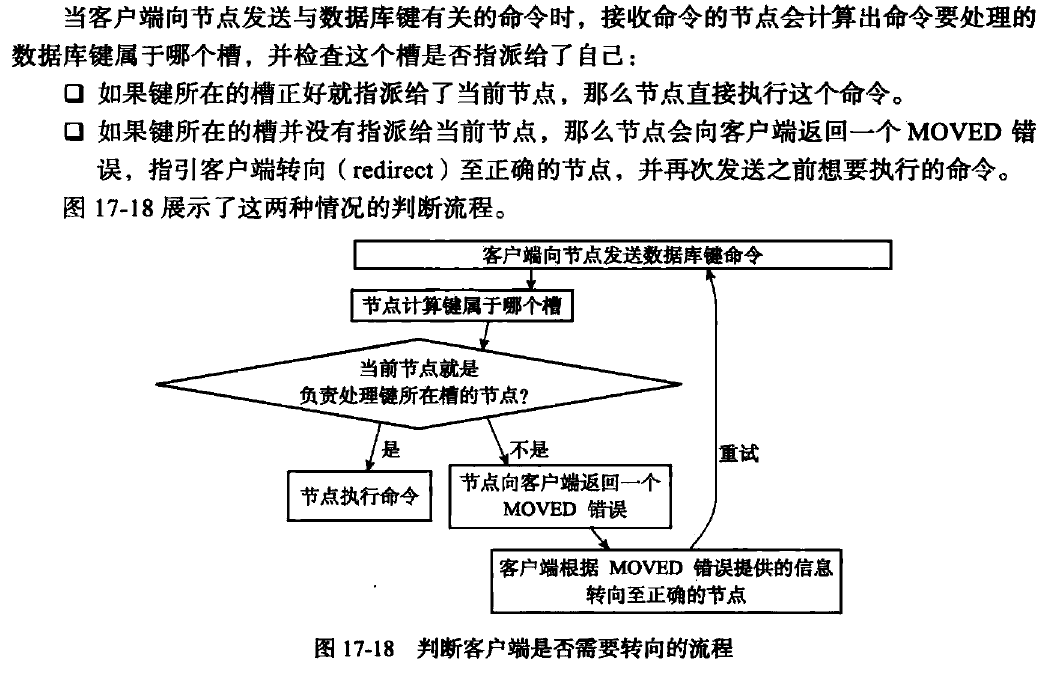




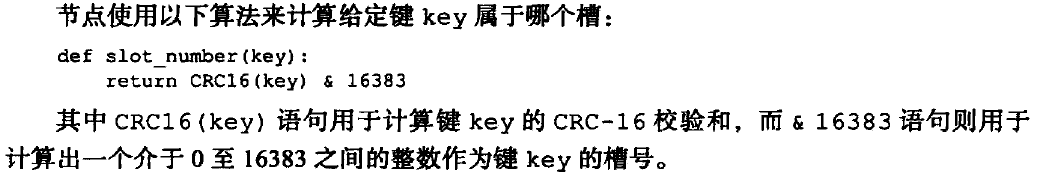
## CLUSTER ADDSLOTS命令的实现



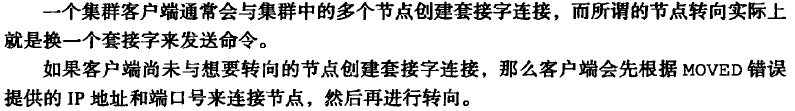
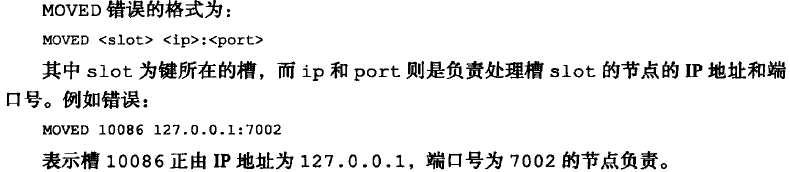
# 在集群中执行命令



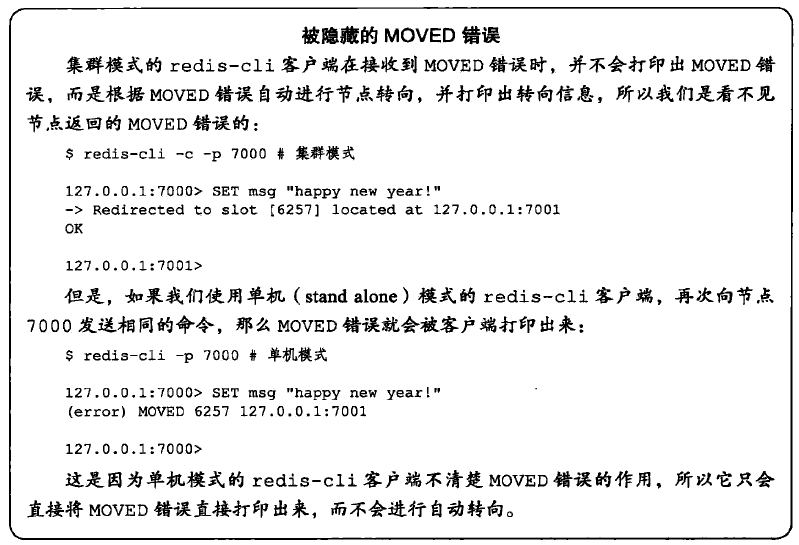
## 计算键的属于哪个槽



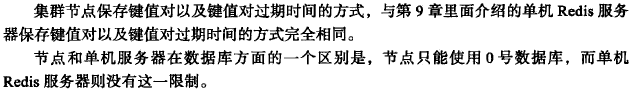
## MOVED错误

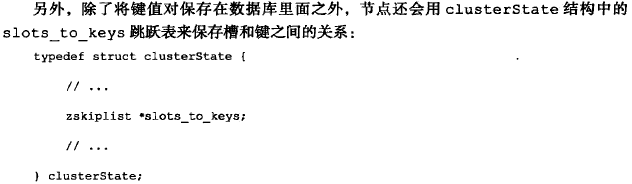


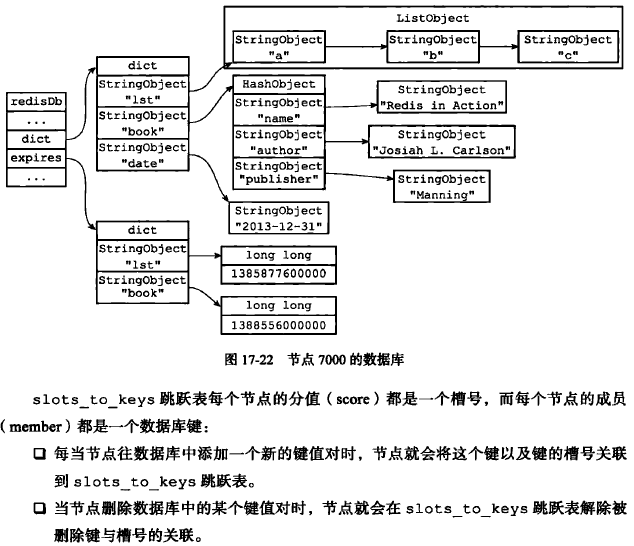
## 被隐藏的MOVED错误



## 节点数据库的实现

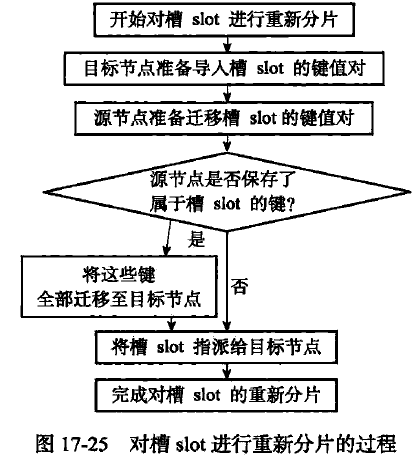


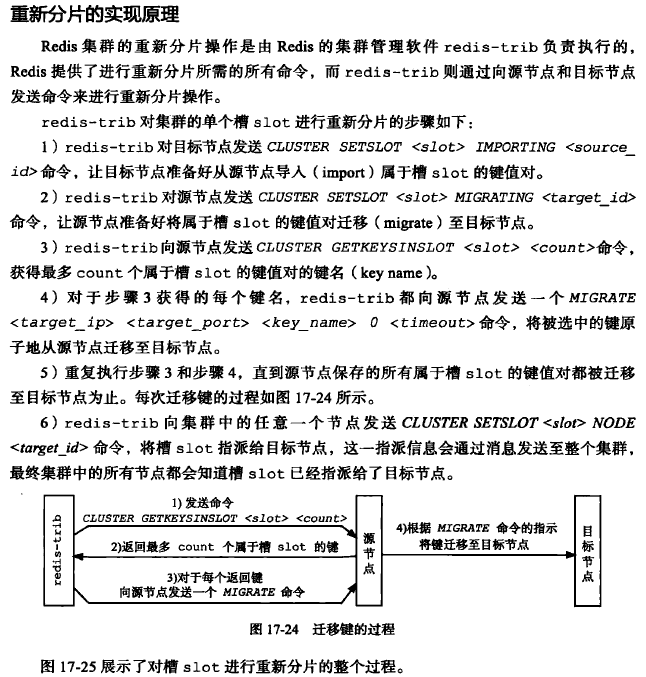




# 重新分片

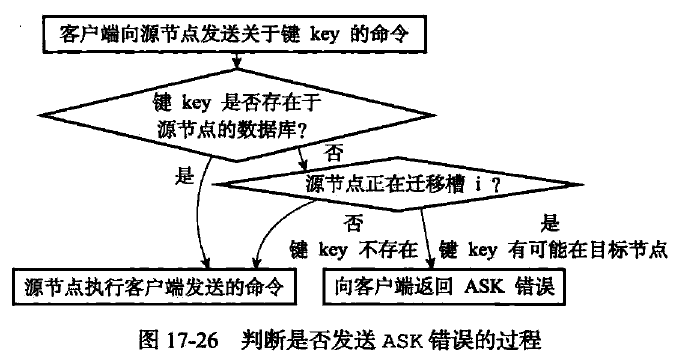
重新分片的处理流程



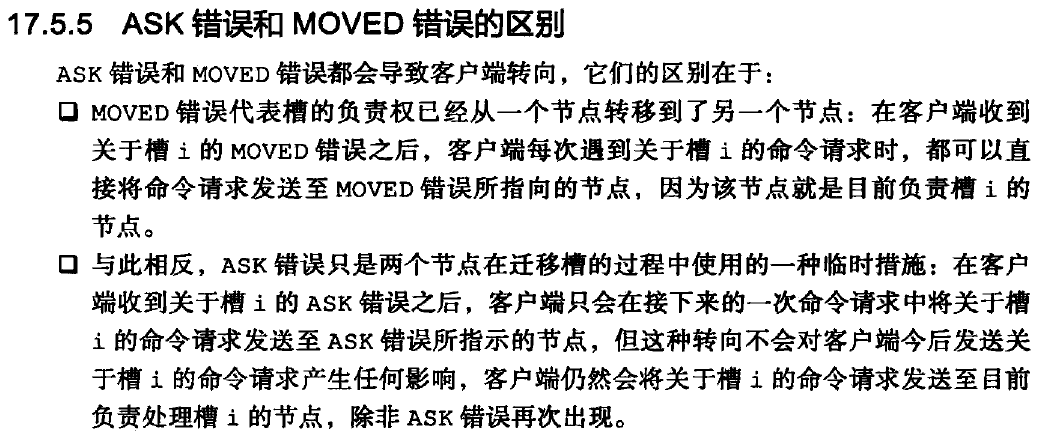


# ASK错误

ASK错误出现的情况：



## ASK错误和MOVE错误的区别

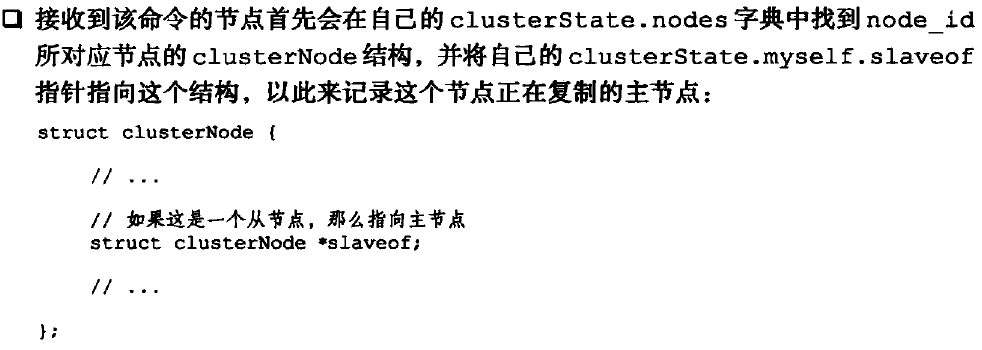


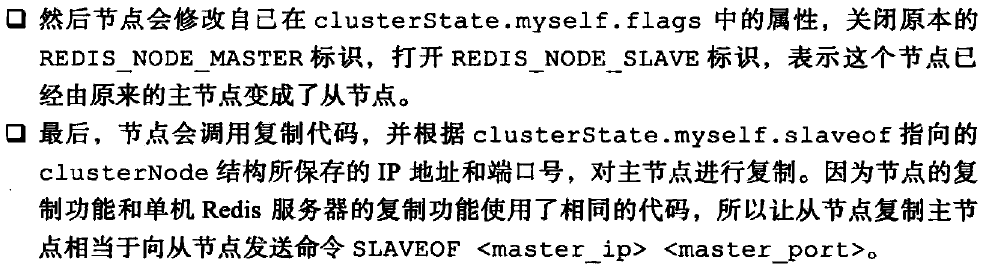
# 复制与故障转移

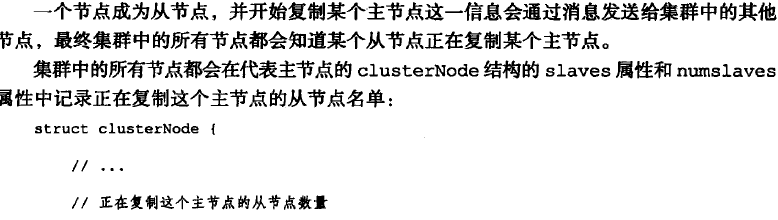
Redis集群中的节点分为主节点和从节点，其中主节点处理槽，从节点用于复制某个节点，并在被复制的主节点下线时，代替下线主节点继续处理命令请求。

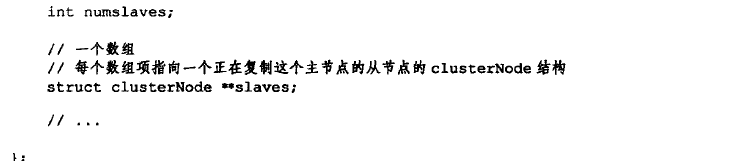
## 设置从节点（CLUSTER REPLACE <node\_id>）

CLUSTER REPLACE <node\_id>命令可以让收到命令的节点成为node\_id所指定节点的从节点。

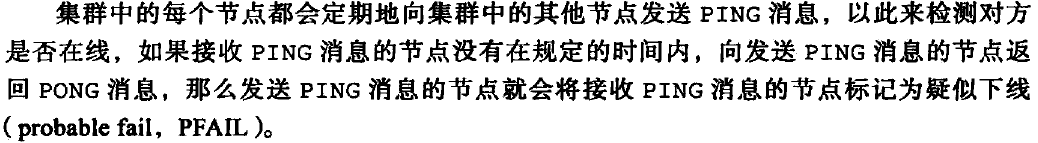


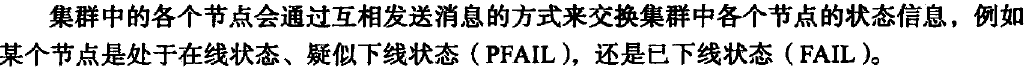


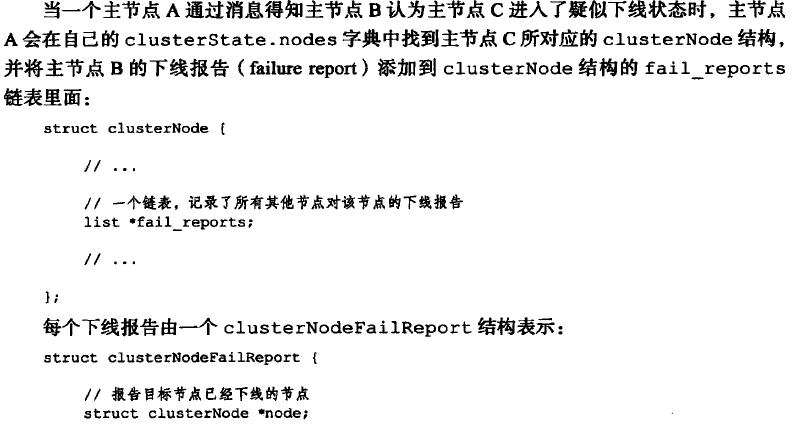


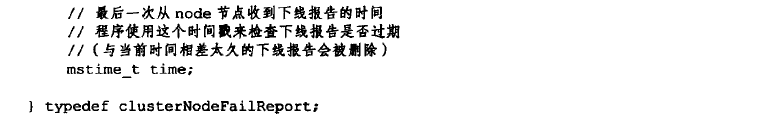


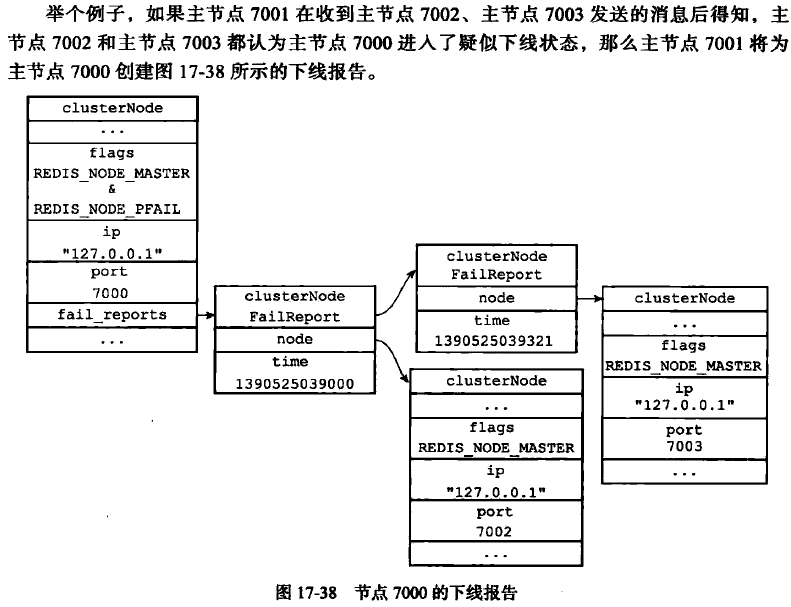
## 故障检测

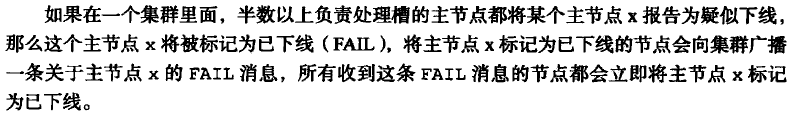












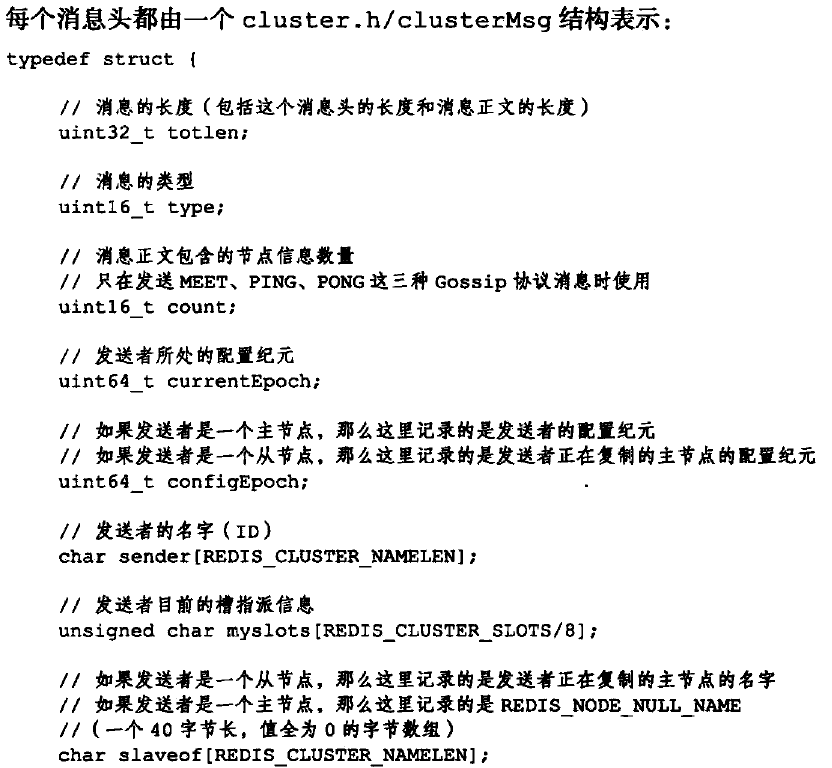
# 消息

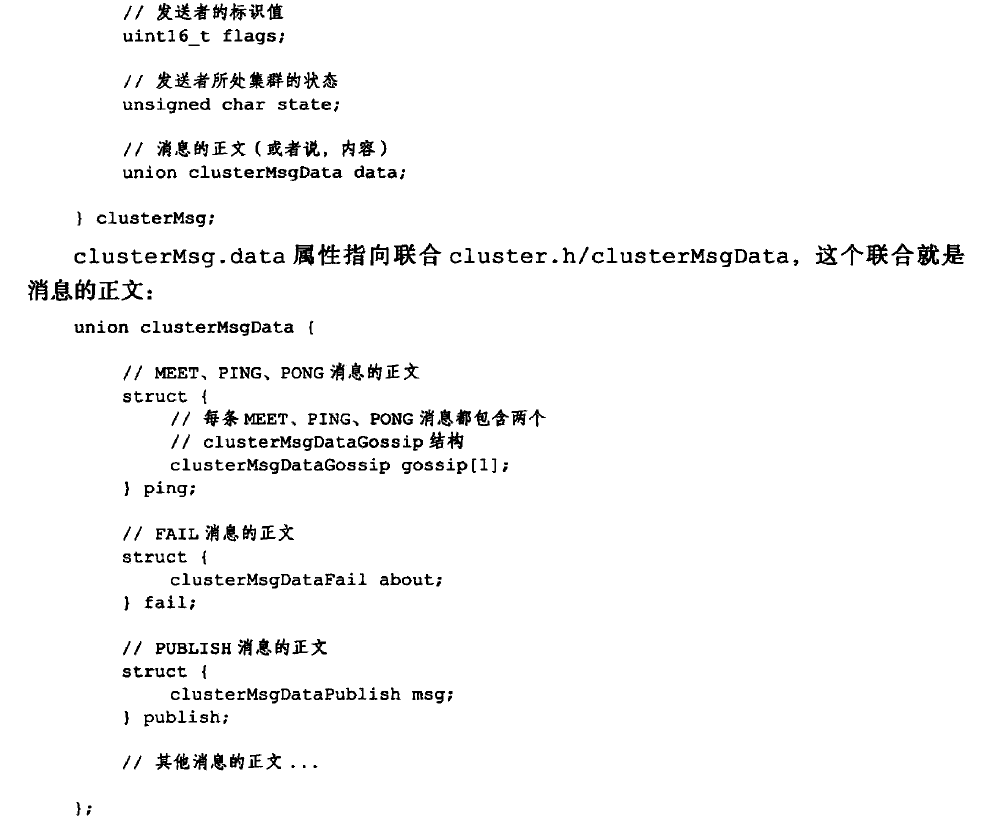
节点发送的消息有以下五种：

1. MEET
2. PING
3. PONG
4. FAIL
5. PUBLISH

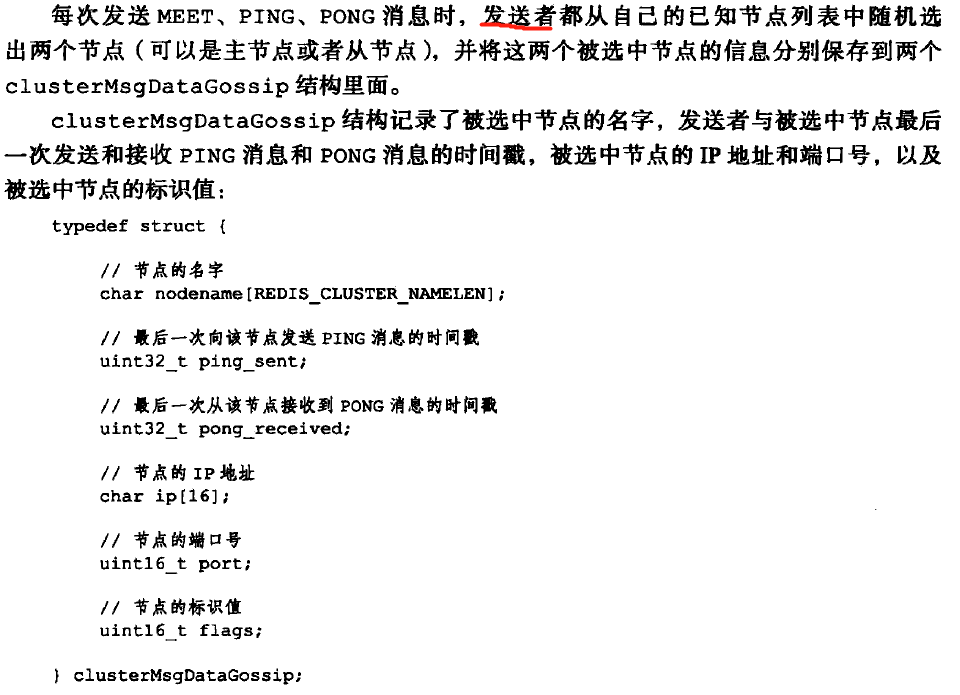
一条消息由消息头和消息正文组成。

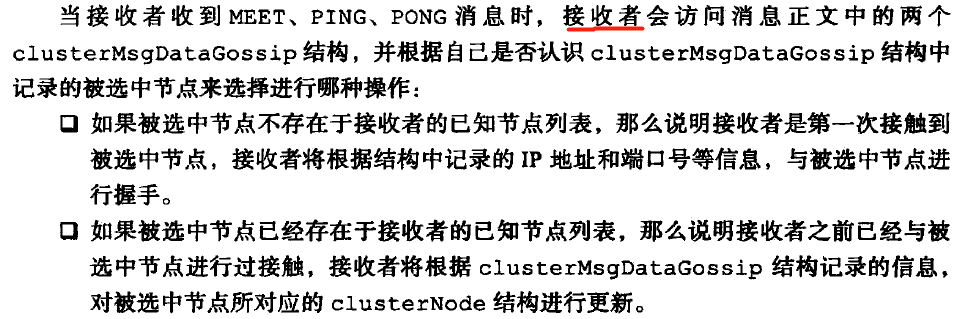
## 消息头



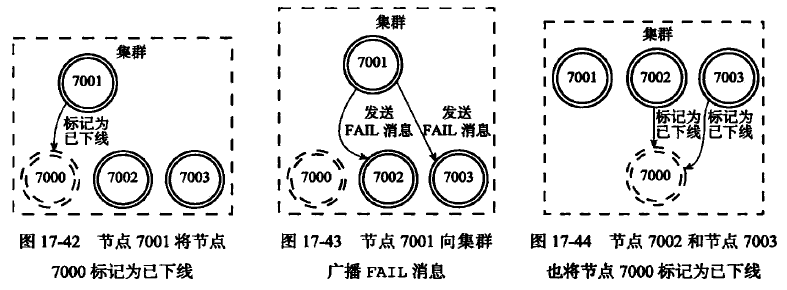


## MEET,PING,PONG消息的实现





## FAIL消息的实现



## PUBLISH消息的实现

