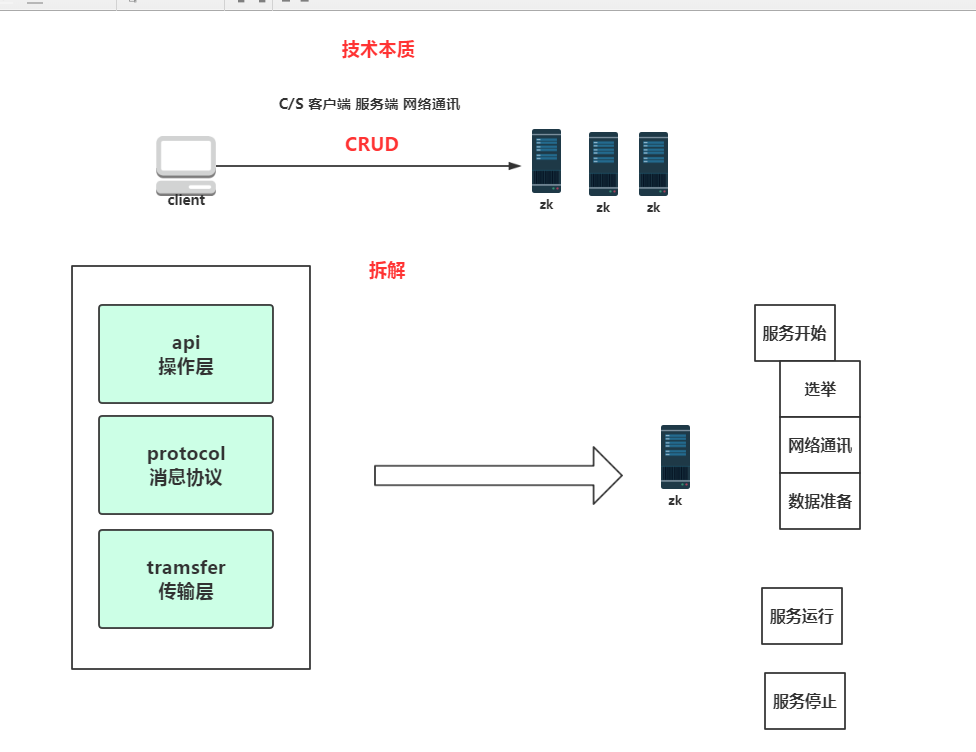
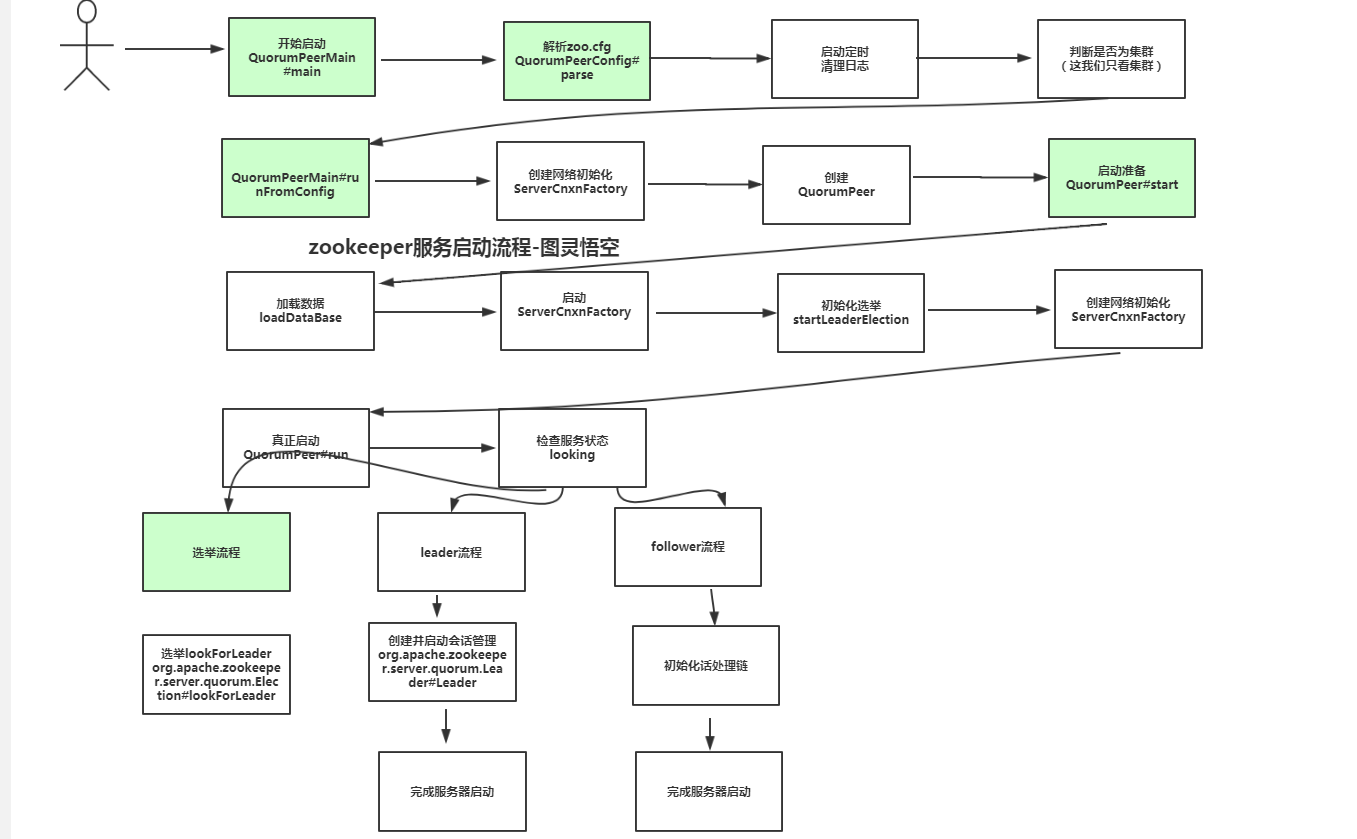
# 源码原理分析：

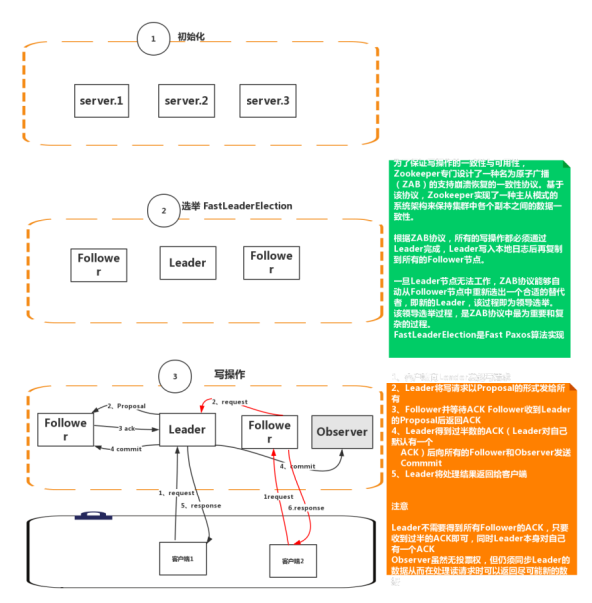
# 本质：



## 服务端：

org.apache.zookeeper.server.quorum.QuorumPeerMain

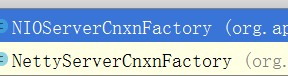




单机、集群

Start方法 线程、Thread Runable

Znode 临时节点特性：



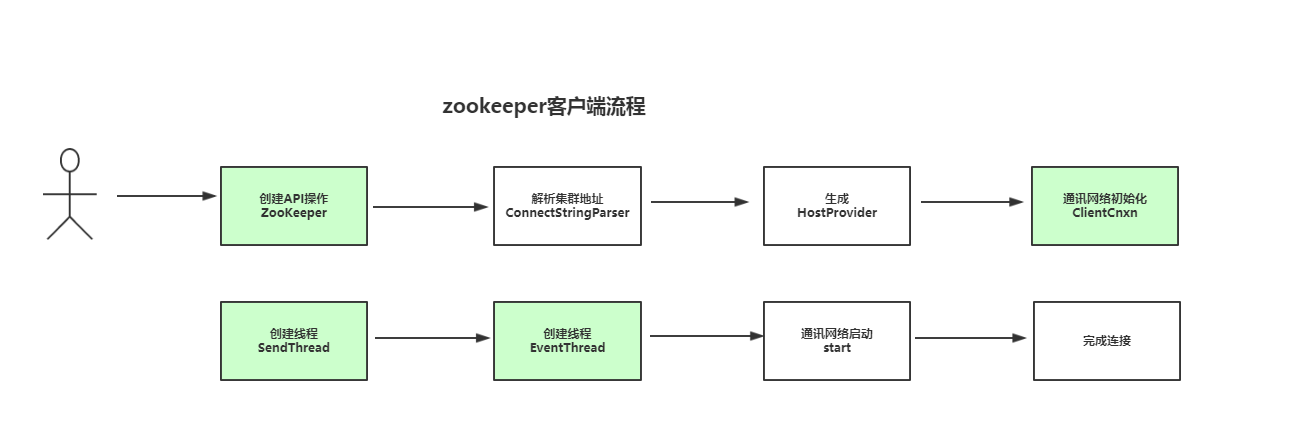
start Thread>run方法

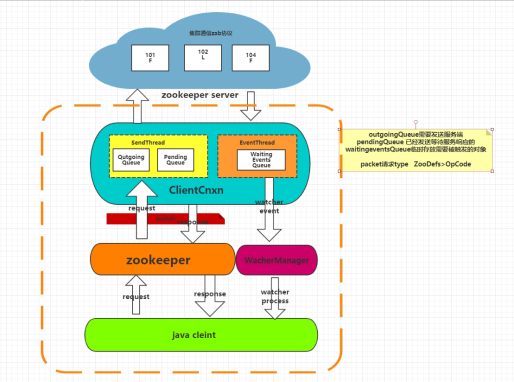
org.apache.zookeeper.server.quorum.QuorumPeer#startLeaderElection 选举

背景知识CAP 一致性协议

WkRequestHeader{protocolVersion=45, lastZxidSeen=0, timeOut=0, sessionId=21474836480000, passwd=[]}

## 客户端：

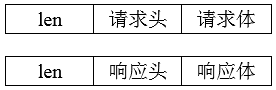


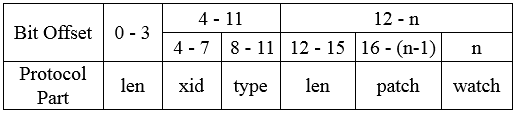


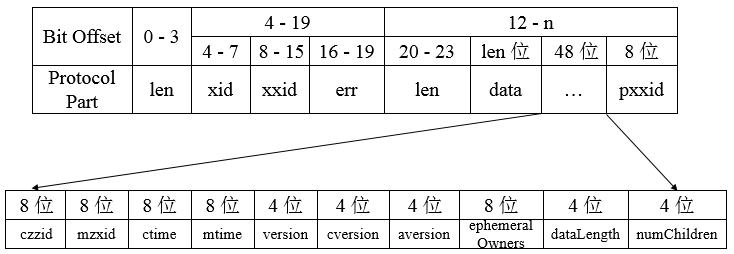
## 序列化：

OutputArchive和InputArchive分别是Jute底层的序列化器和反序列化器。

数据传输结构：







# ZAB协议Zk选举流程：

选举分5个步骤：

1. 初始化leader选举，投票

给自己投票 myid zxid 0

1. 每个服务器接受投票

(1 0) (2,0) (3,0) >myid 谁最大谁就是leader 33错了 32

1. 处理投票

别的服务器 pk

1. 统计投票

**过半票数n=n/2+1;**

1. 服务器状态变更

Leader，其他票数低 f

http://thesecretlivesofdata.com/raft/#election

### 作业：zk服务启动顺序颠倒

服务器具有四种状态，分别是LOOKING、FOLLOWING、LEADING、OBSERVING。

　　LOOKING：寻找Leader状态。当服务器处于该状态时，它会认为当前集群中没有Leader，因此需要进入Leader选举状态。

　　FOLLOWING：跟随者状态。表明当前服务器角色是Follower。

　　LEADING：领导者状态。表明当前服务器角色是Leader。

　　OBSERVING：观察者状态。表明当前服务器角色是Observer。

# 运维：

**四字命令：**

| **ZooKeeper四字命令** | **功能描述** |
| --- | --- |
| conf | 3.3.0版本引入的。打印出服务相关配置的详细信息。 |
| cons | 3.3.0版本引入的。列出所有连接到这台服务器的客户端全部连接/会话详细信息。包括"接受/发送"的包数量、会话id、操作延迟、最后的操作执行等等信息。 |
| crst | 3.3.0版本引入的。重置所有连接的连接和会话统计信息。 |
| dump | 列出那些比较重要的会话和临时节点。这个命令只能在leader节点上有用。 |
| envi | 打印出服务环境的详细信息。 |
| reqs | 列出未经处理的请求 |
| ruok | 测试服务是否处于正确状态。如果确实如此，那么服务返回"imok"，否则不做任何相应。 |
| stat | 输出关于性能和连接的客户端的列表。 |
| srst | 重置服务器的统计。 |
| srvr | 3.3.0版本引入的。列出连接服务器的详细信息 |
| wchs | 3.3.0版本引入的。列出服务器watch的详细信息。 |
| wchc | 3.3.0版本引入的。通过session列出服务器watch的详细信息，它的输出是一个与watch相关的会话的列表。 |
| wchp | 3.3.0版本引入的。通过路径列出服务器watch的详细信息。它输出一个与session相关的路径。 |
| mntr | 3.4.0版本引入的。输出可用于检测集群健康状态的变量列表 |

使用：echo zk命令|nc ip port

如：echo mntr |nc 192.168.0.31 2181

**注意：**

Linux中nc命令是一个功能强大的网络工具,全称是netcat

在线安装：yum install -y nc

**容灾：**N=N/2+1

**过半票数n=n/2+1; 选举 ledader 集群**