# raft-extended

## 中文翻译

<https://www.jianshu.com/p/2a2ba021f721>

<http://blog.luoyuanhang.com/2017/02/02/raft-paper-in-zh-CN/>

<https://www.cnblogs.com/YaoDD/p/6172011.html>

<https://blog.csdn.net/hfty290/article/details/44281205>

# 工作流程

## 三种状态

### 领导者（Leader）

### 追随者（Follower）

### 候选人（Candidate）

## 三种RPC

### Request Vote RPC

### Append Entries RPC

#### 心跳（不带有任何日志条目）

#### 日志复制（带有一条或多条）

### Install Snapshot RPC

## 三种时间

Broadcast Time << Election Timeout << MTBF

Broadcast Time应该比Election Timeout小一个数量级，为的是使领导人能够持续发送心跳信息来阻止追随者们开始选举。

### 广播时间（Broadcast Time）

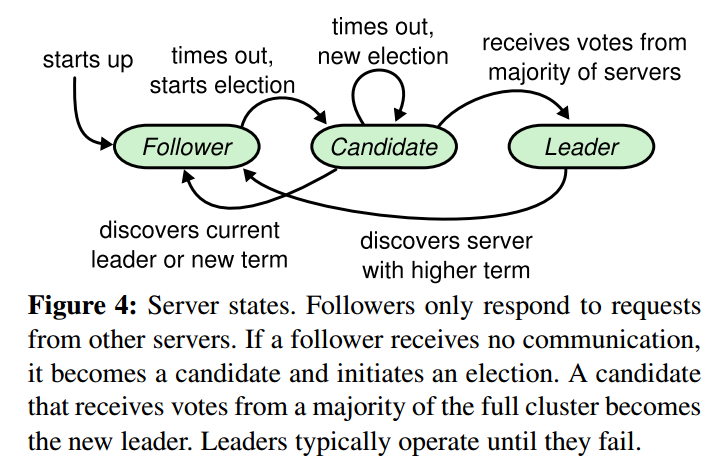
一台服务器并行的向集群中的其他服务器发送 RPC 并且收到它们的响应的平均时间。

### 选举超时（election timeout）

### 服务器平均故障时间（MTBF）

单个服务器发生故障的间隔时间的平均数。

## 状态图



## 状态转移逻辑

### 追随者

#### 服务器启动时，自动初始化为追随者

#### 追随者在选举超时之内没有收到心跳消息

一个追随者会自增它的当前任期并且转换状态为候选人。然后，它会给自己投票并且给集群中的其他服务器发送 Request Vote RPC。

### 候选人

#### 它赢得了选举

一个候选人如果在一个任期内收到了来自集群中大多数服务器的投票就会赢得选举

#### 另一台服务器赢得了选举

当一个候选人等待别人的选票时，它有可能会收到来自其他服务器发来的声明其为领导人的 Append Entries RPC。

如果这个领导人的任期（包含在它的 RPC 中）比当前候选人的当前任期要大，则候选人认为该领导人合法，并且转换自己的状态为追随者。

如果在这个 RPC 中的任期小于候选人的当前任期，则候选人会拒绝此次 RPC， 继续保持候选人状态。

#### 一段时间后没有任何一台服务器赢得了选举

Raft 使用随机的选举超时时间来确保第三种情形很少发生，并且能够快速解决。

### 领导者

# 数据维护

## 每个服务器需要维护的数据

### 需要持久化的数据

### 不需要持久化的数据

# 领导选举（Leader election）

# 日志复制

# 安全性

# 动态改变集群成员