EDU

CSDN学院　IT实战派

比特币和区块链的相关技术细节

分布式系统共识

讲师：康烁

本节目标

1. 了解分布式系统的共识算法
2. 了解拜占庭将军容错

分布式系统

1. 分布式系统案例：售票系统

2. 分布式系统的共识算法

在一个分布式系统中，如何保证集群中所有节点中的数据完全相同并且能够对某个提案（Proposal）达成一致是分布式系统正常工作的核心问题，而共识算法就是用来保证分布式系统一致性的方法

3. 区块链和共识算法的关系

数字货币 -> 双花问题 -> 顺序账本（区块链） -> 分布式系统的共识算法

分布式系统的同步和异步

- 同步系统：消息不丢失且秒到
- 异步系统：消息有延迟而且可能丢失

分布式系统的一致性

- 强一致性：任何时刻保持一致
- 弱一致性：某一时刻保持一致

不能达成一致性的两种情况

我们假设通信是可靠的。那么我们把造成不能达成一致性的故障情况分为两种：

1. 节点只是故障状态，不存在恶意节点，那我们称为“非拜占庭错误”

2. 存在恶意节点的分布式网络，我们称为“拜占庭错误”。我们区块链面临的的一致性问题为“拜占庭将军问题”

非拜占庭错误的两种解决方案

PAXOS：

核心思想：Paxos解决这一问题利用的是选举，少数服从多数的思想，只要2N+1个节点中，有N个以上同意了某个决定，则认为系统达到了一致，客户端不必与所有服务器通信，选择与大部分通信即可；无需服务器都全部处于工作状态，只有保证半数以上存活着，整个过程也能持续下去，容错性相当好

Raft：

核心思想：Raft相比paxos的优点是容易理解，容易实现。它强化了leader的地位，把整个协议可以清楚的分割成两个部分，并利用日志的连续性做了一些简化

Leader在时。由Leader向Follower同步日志

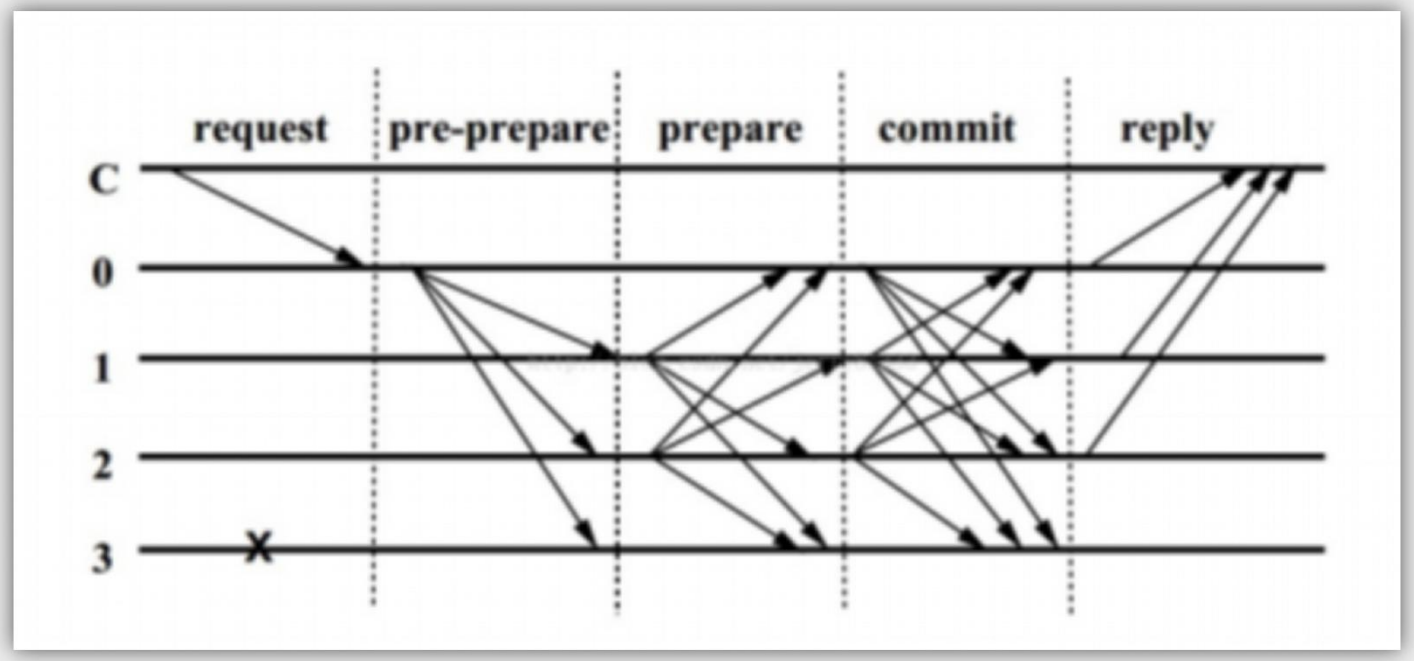
Leader挂掉了，选一个新Leader，Leader选举算法

拜占庭错误的解决方案

1. 对于可以容忍拜占庭错误的算法：PBFT，中本聪共识（POW），POS和DPOS四种算法
2. PBFT：更加实用的拜占庭容错方法。早期的BFT的缺陷：1、假定是同步场景；2、性能太慢（超过100个节点则不可用）
3. PBFT算法的核心理论是n>=3f+1

   n是系统中的总节点数，f是允许出现故障的节点数。如果这个系统允许出现f个故障，这个系统必须包括n个节点，才能解决故障。
4. http://pmg.lcs.mit.edu/papers/osdi99.pdf

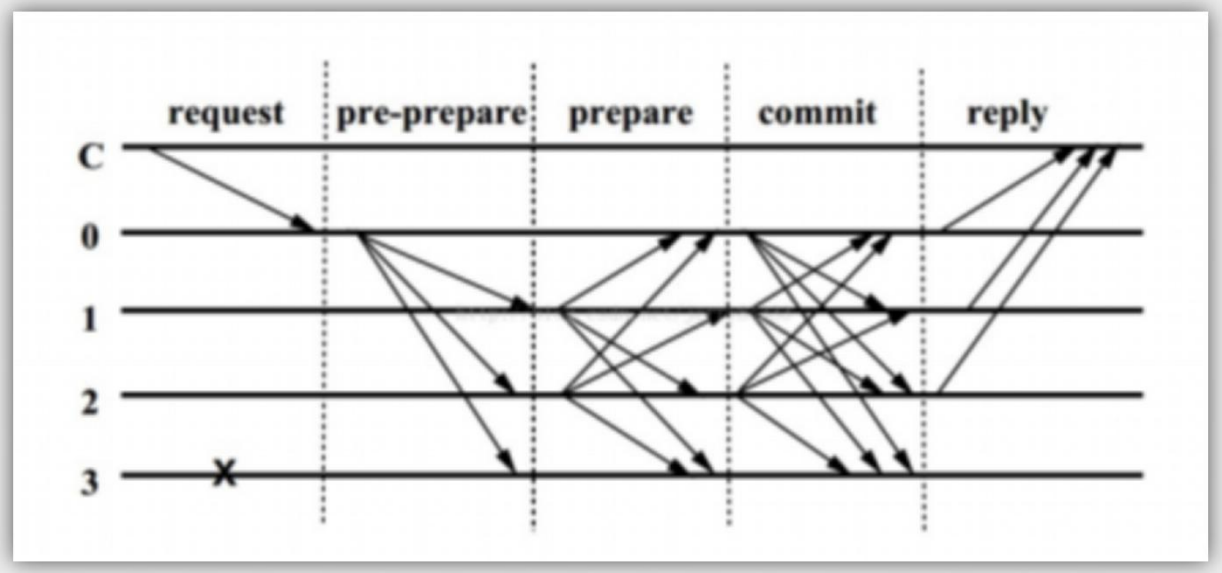
PBFT算法在区块链中的应用



步骤：

1. Request: 从全网节点选举出一个主节点（Leader），新区块由主节点0负责生成
2. Pre-Prepare：每个节点把客户端发来的交易向全网广播，主节点0将从网络收集到的交易，并把搜集到的多个交易在新区块中排序后存入列表，并将该列表向全网广播，扩散至123

PBFT算法在区块链中的应用（续）



步骤（续）：

3. Prepare：每个节点接收到交易列表后，根据排序模拟执行这些交易。所有交易执行完后，基于交易结果计算新区块的哈希摘要，并向全网广播，1->023，2->013，3因为宕机无法广播
4. Commit：如果一个节点收到的2f（f为可容忍的拜占庭节点数）个其它节点发来的摘要都和自己相等，就向全网广播一条commit消息
5. Reply：如果一个节点收到2f+1条commit消息，即可提交新区块及其交易到本地的区块链和状态数据库

总结

1. 分布式系统的共识算法
2. 什么是拜占庭将军容错

布置作业

- 必做内容：
- 理解拜占庭将军问题

EDU

CSDN学院　IT实战派