第11讲｜深入区块链技术（三）：共识算法与分布式一致性算法

2018-04-18 陈浩

第11讲｜深入区块链技术（三）：共识算法与分布式一致性算法

朗读人：黄洲君 11'44" | 5.38M

共识机制的概念，我们在前面的文章"浅说区块链共识机制"中已经讲解了一部分，但是，共识算法其实是一个非常大的话题，一篇文章肯定 没有办法做到面面俱全。

那么今天的内容，我会将重点放在梳理技术的脉络上，详细分析的部分会少一点。如果你对共识算法有兴趣的话，可以自行查找相关内容，也可以和其他的资料进行相互补充的阅读。

从相亲大会说起：分布式系统的模型

由于区块链就是一种分布式系统，所以这篇文章我就从这一概念开始讲起。 为了让您更容易理解分布式系统，我们先来构建一个模型。

在"浅说区块链共识机制"那篇文章中，我举了一个村庄举办相亲大会的例子，我们来回顾一下。

1. 大村子因为人口增长变成 11 个小村落分散在地图各地；
2. 村落之间的通信只能依靠信鸽；
3. 一只信鸽可能无法完全覆盖所有村落，需要有中继村落代为传输消息。

相亲大会的举办权会为村子带来巨大收益，为了产生合理的举办者，人们约定了几条规则：

1. 大会举办权从 A 和 B 两个村子中产生，他们每一届都是候选村；
2. 投票时所有村落仅能投 A 或 B；
3. 用投票的方式产生举办者，少数服从多数。

所有村子会为了举办权都会使出浑身解数，比如延迟发送投票结果、篡改别人的投票结果、假装没有接收到通知等等。

其实这是一个典型的分布式系统，可以看成是我们简化版的区块链网络环境，那么这个分布式系统会遇到什么样的问题呢？

分布式系统面临的问题

分布式系统面临了几个问题：一致性问题，可终止性问题，合法性问题。

可终止性可以理解为系统必须在有限的时间内给出一致性结果，合法性是指提案必须是系统内的节点提出，当然其中面对的最重要也是最基础的问题，就是我们常说的一致性问题。

一致性是指在某个分布式系统中，任意节点的提案能够在约定的协议下被其他所有节点所认可。

需要提醒你区分的一点是：这里的"认可"表示所有节点对外呈现的信息一致，而不是对信息的内容认可。一致性也分严格一致性、 最终一致性，这些我们在后文会说到。

我们回到上面的例子，我们提到了 所有的村子只能投 A 或 B。其实这个投票的动作可以理解为提案。

在"投票过程被大家所认可"这个语境下，"被大家所认可"表示某个村落投票的结果已经被记录。用于 最后统计结果，而不是认可投给 A 或者投给 B，这也是我在上述强调你要注意区分的一点。

那我们这里所说的一致性到底体现在那里呢？

主要体现在下面两种类型的问题上。

1. **非人为恶意的意外投票过程。**非人为恶意篡改可归类为信鸽半路挂掉、信鸽迷路、信鸽送错目的地、信鸽送信途中下雨导致 信件内容模糊、接收信件的人不在家、天气变化信鸽延迟送达等等，这些对应到分布式系统面临的问题就是：消息丢包、网络拥堵、消息延迟、消息内容校验失败、节点宕机等。
2. **人为恶意篡改投票过程。**人为恶意篡改包括"精神分裂式投票"，中继篡改上一个村落的投票信息，对应到分布式系统面临的问题就是：消息被伪造、系统安全攻击等等。发生的人为恶意篡改的过程就可以称之为系统发生了拜占庭错误（Byzantine Fault），如果系统可以容忍拜占庭错误而不至于崩溃，也就是在发生系统被恶意篡改的情况下仍然可以达成一致，我们将这样系统称作为做拜占庭容错系统。

问题 1 我们已经有较成熟的方案了。分布式系统本质上是一种并行异步操作，如果通过中心化的手段将系统中的"并行不确定"操作变更为"同步串行"操作就能解决上述的问题。

比如让第三方机构介入托管所有人的投票：或者构造一个不可伪造令牌，大家轮流将投票统一写到令牌上。

这些也是现代分布式系统经常使用的方法，但是这些方法有个缺陷，如果在分布式系统中被过多地使用以后，系统便会越来越像单点系统。

我们设计分布式系统的初衷就是为了克服单点系统的可用性不足、扩展性不好、单点性能上限等缺陷，这种退化的方案可能不是我们想要的。

而问题 2 要求设计拜占庭容错系统，这个在 IT 行业并不常见，因为多数 IT 系统是中心化的，所以如果我们想要解决问题 2，这就引出了我们今天要介绍的共识算法与分布式一致性算法。

有关分布式系统的定理

我们在介绍具体的分布式一致性算法之前，先介绍两个定理，做一下铺垫。

- **第一个是 FLP 不可能性**，简单来说是：即使网络通信完全可靠，只要产生了拜占庭错误，就不存在一个确定性的共识算法能够为异步分布式系统提供一致性。换句话来说就是，不存在一个通用的共识算法 可以解决所有的拜占庭错误。
- **第二个是 CAP 定理**，CAP 定理是分布式系统领域最重要的定理之一，这个我们在"理解区块链的常见误区"一文中稍微讲到过，也就是在设计分布式系统的过程中，"一致性""可用性""分区容忍性"三者中，我们只能选择两个作为主要强化的点，另外一个必然会被弱化。

我们由 CAP 定理可以推导出 严格一致性和 最终一致性。严格一致性是指在约定的时间内，通常是非常短、高精度的时间内，系统达到一致性的状态，这种系统很难实现，即使实现也很难有高的性能。

所以人们从工程的角度提出了最终一致性，最终一致性不要求严格的短时间内达到一致，为了其他两个指标，我们 相当于让一致性在时间上做了妥协。区块链满足了最终一致性，而且处理过程时间比较长。

可用性其实是传统技术 后端架构上非常重要的指标，从单点到主备模式、从主备模式到异地多活， 再到现在的 Paxos 和 Raft 协议。

我们从软件架构上也经历了基于 ESB 的模块化 SOA 模式，到无状态的微服务架构。 从工程的角度来看，根据业务需求达到 4 个 9、6 个 9 就足够了，只是 肯定比不了区块链近乎 100% 的可用性。

分区容忍性在企业内部极少出现，尤其是中心化的服务性应用，所以很少考虑， 然而区块链的 P2P 网络环境十分复杂，所以必须要保证很高的分区容忍性。

通过以上我们可以发现比特币，以太坊等公链是偏重高可用性、分区容忍性（AP），满足最终一致性（C）且 TPS 较低的分布式系统。

所以如果有人号称 他们的区块链能够达到媲美中心化系统上万的 TPS，先别着急投资，你问问他们 技术是不是知道 CAP 定理，再问问他们的去中心化程度如何。

这点我们也可以从 EOS 等高性能的区块链身上佐证。EOS 全球只有 21 个记账节点，而以太坊全球有上万个节点可以随时参与 记账，所以越想去中心化，你的 TPS 就不可能高，这也就是为什么 EOS 的 TPS 高，而以太坊的 TPS 低。

接下来我来介绍一下经典的分布式一致性算法和区块链的共识算法。经典的分布式一致性算法在多数的论文中一般被叫做共识算法，在这里，我为了与区块链的共识算法做出区别，所以 在命名上改成了分布式一致性算法， 但是它们要解决的问题是 一样的。

共识算法与分布式一致性算法

1. 经典的分布式一致性算法

经典分布式一致性算法有 Raft 协议，Raft 协议是一种强 Leader 的一致性算法，它的吞吐量基本就是 Leader 的吞吐量。它无法抵御节点恶意篡改数据的攻击。

稍微复杂一点的就是 Paxos 协议，Paxos 能提供不同场合不同种类的一致性算法，所以 Paxos 有很多变种，经典 Paxos 是 Leaderless 的，有变种是强 Leader 型的，叫做 Fast Paxos， 有关 Paxos 的文献非常丰富，这里就不赘述了。

以上两种都是不提供拜占庭容错的系统，下面介绍一种具有拜占庭容错的一致性算法。

PBFT 全称实用性拜占庭容错系统（Practical Byzantine Fault Tolerance, PBFT），PBFT 是一种状态机， 要求所有节点共同维护一个状态，所有节点采取的行动一致，PBFT 非常适合联盟链等对性能具有较高要求的场合，超级版本项目中的 Fabric 框架默认采用的就是 PBFT 的修改版本。

2. 区块链共识算法

区块链的共识算法，我在某些场合直接称作基于经济学的博弈算法 ，以区别于经典分布式一致性算法思路，它的整体思路就是让攻击者的攻击成本远远大于收益。

区块链中的共识算法目前具有工业级熟度的是 PoW，另外两种比较成熟的是 PoS 和 DPoS，其次还有一些变种和单一币种使用的共识算法，例如 Ripple 共识、PoC 共识（概念性证明）、PoE 共识（存在性证明）。

在使用 PoW 共识算法的情况下，容错阈值是 50%，而 PBFT 及其变种的容错阈值是 33% 左右，这里的百分比是指作弊节点占全网节点的 比例。

PoX 类的算法其实都延续了 PoW 的设计理念，相比较经典分布式一致性算法，PoX 类算法通过 概率选择记账者降低了潜在的提案者，另外是延长了达成最终一致性的时间。

第一条降低了系统通信复杂度，每次记账系统的确定性其实是概率确定的， 又由于被选中需要付出成本，此处才提高了记账成本阈值，结合区块链的基础代币设计，是一个非常天才的想法。

有关 PoW、PoS、DPoS 的内容，我在后面会有专文介绍。

总结

今天我们从相亲大会开始说起，介绍了分布式系统所面临的问题，之后，我们又引出了区块链所要解决的拜占庭容错问题，并重点介绍了分布式系统的基本定理。最后我还介绍了经典共识算法和区块链算法。

区块链并没有造离分布式系统这个理论框架，希望今天的内容能够帮助你分辨出真实的区块链项目。最后给你留一个思考题，区块链行业有哪些公司使用了如 PBFT、Paxos 这样的经典共识算法呢？ 你可以给我留言，我们一起讨论。

感谢你的收听，我们下期再见。

精选留言

阿良 5

跟我个人总结，联盟链基本上使用PBFT及其变种，而公有链大多采用PoX算法，对吗？

2018-04-21

作者回复

嗯，我们理解一致。

2018-04-21

teletime 1

去中心化需要群体意志，群体意志需要共识，抽签是最好方式，谁用机器模拟抽签模拟的好，谁的方法好。POW思想正统。

2018-04-21

作者回复

区块链随机数有的，Cardano项目有公开一种算法。

2018-04-21

吹牛老爹 0

最近的迅雷的公链采用的是Pbft

2018-05-09

D 0

还有人问投资的问题，如果我是主讲人我都不会发表任何观点。

2018-04-26

沃野旷阳 0

EOS 会不会形成泛中心化，即牺牲一致性来得到效率？ 请教老师对此看法。 EOS 值得投资吗？

2018-04-21

作者回复

本专栏不构成任何投资建议哦。

2018-04-26