区块链共识

北京交通大学
计算机与信息技术学院
信息安全系

李超 (li.chao@bjtu.edu.cn)
段莉 (duanli@bjtu.edu.cn)

教学安排

第 1 周	区块链技术概述	总体介绍
第 2 周	区块链的密码学原理	
第 3 周	区块链共识	
第 4 周	比特币及其原理	比特币
第 5 周	比特币关键技术	
第 6 周	比特币热点问题	
第 7 周	以太坊基础	以太坊
第 8 周	智能合约基础	
第 9 周	去中心化应用开发	
第 10 周	以太坊实验	
第 11 周	超级账本基础	超级账本
第 12 周	共识机制及智能合约	
第 13 周	区块链扩容技术	热点领域
第 14 周	区块链中心化度量技术	
第 15 周	区块链跨链技术	

六层模型

目录 | CONTENT

分布式共识

什么是分布式共识?

- *分布式的节点就某个全局状态达成一致。*

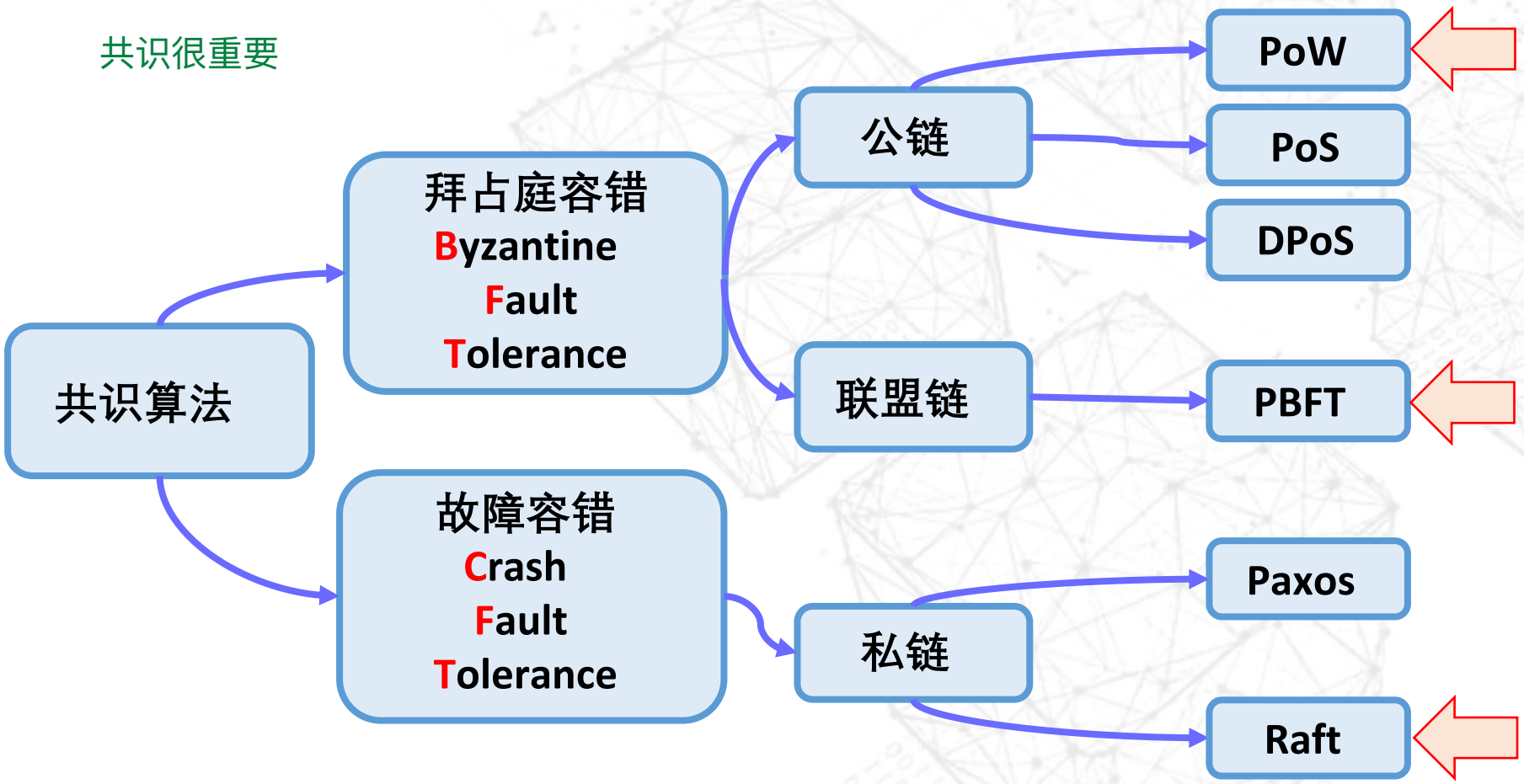
中心化的共识

去中心化的分布式共识

区块链类型与共识协议

分布式共识算法分类

共识很重要



拜占庭将军问题

经典共识协议

Raft

http://thesecretlivesofdata.com/raft/

共识算法	Raft
适用环境	私链
算法通信复杂度	O(n)
最大故障和容错节点	N-f>f
流程对比	• 初始化leader选举（谁快谁当） • 共识过程 • 重选leader机制

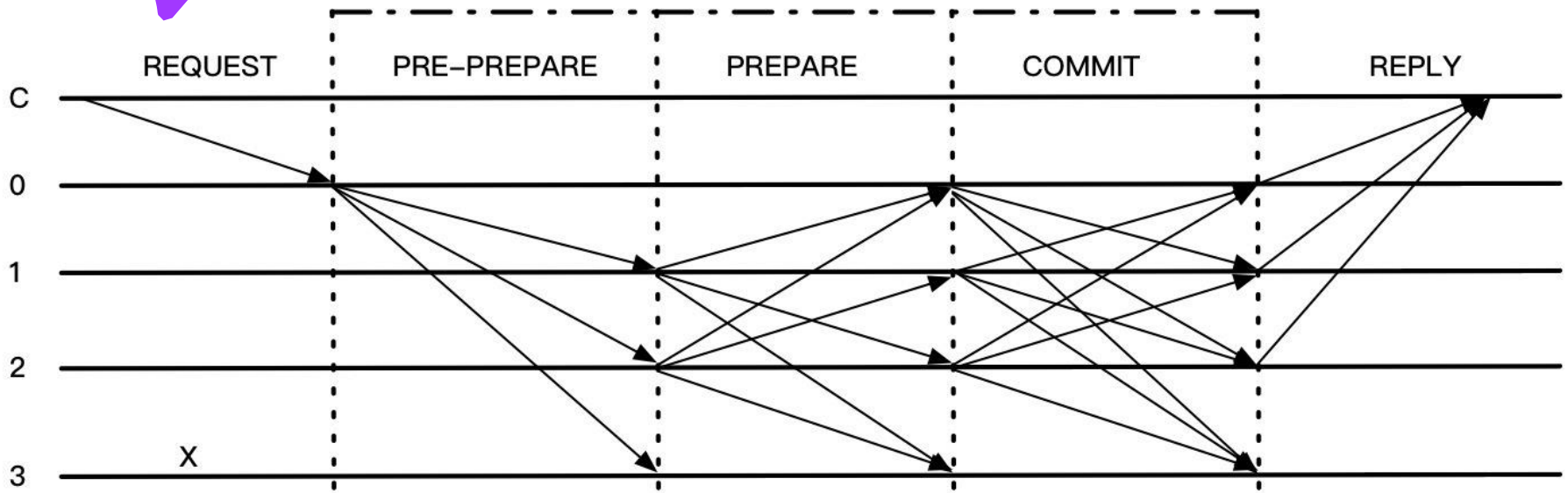
PBFT

- 节点：有一个节点会被当做**主节点**，而其他节点都是**备份节点**。
- 通信：系统内的所有节点都会相互通信。
- 目标：**大家能以少数服从多数的原则达成数据的共识**。如果主节点出现明显的撒谎迹象，其他的节点也可以联合起来更换主节点。

PBFT

The resiliency of our algorithm is optimal: $3f+1$ is the minimum number of replicas that allow an asynchronous system to provide the safety and liveness properties when up to f replicas are faulty (see [2] for a proof). This many replicas are needed because it must be possible to proceed after communicating with $n-f$ replicas, since f replicas might be faulty and not responding. However, it is possible that the f replicas that did not respond are not faulty and, therefore, f of those that responded might be faulty. Even so, there must still be enough responses that those from non-faulty replicas outnumber those from faulty ones, i.e., $n-2f > f$. Therefore $n > 3f$.

3阶段协议



<REUEST, *o*, *t*, *c*>σ*c*

* *o*是operation
* *t*是timestamp
* *c*是client
* σ*c*是*c*对消息的签名

节点说明：

* C是client
* 0是primary
* 1、2是没有问题的节点
* 3是有问题的节点
* C给primary发送REUEST

<<PRE-PREPARE, *v*, *n*, *d*>σ*p*, *m*>

* *v*是view
* *n*是序列号
* *d*是*m*的消息摘要
* *m*是捎带给*backup*的消息
* σ*p*是primary对消息的签名

REQUEST消息处理：

* 判断正在处理的消息是否超出限制
* 超出限制就先缓存起来再打包发送
* 生成序列号*n*
* 广播PRE-PREPARE消息给backup节点

<PREPARE, *v*, *n*, *d*, *i*>σ*i*

* *v*是view
* *n*是序列号
* *d*是*m*的消息摘要
* *i*是节点*i*的编号
* σ*i*是节点*i*对消息的签名

PRE-PREPARE消息处理：

* 检查PRE-PREPARE消息签名是否正确，消息的摘要是否正确
* 判断当前节点的view是否是*v*
* 检查是否收到过相同*v*和*n*的消息，摘要却不一样
* 判断*n*是否在*h*和*H*之间
* 如果都通过的话就广播PREPARE消息给其他节点，包括primary节点

<COMMIT, *v*, *n*, *D(m)*, *i*>σ*i*

* *v*是view
* *n*是序列号
* *D(m)*是*m*的消息摘要
* *i*是节点*i*的编号
* σ*i*是节点*i*对消息的签名

PREPARE消息处理：

* 检查PREPARE消息签名是否正确，消息的摘要是否正确
* 判断当前节点的view是否是*v*
* 判断*n*是否在*h*和*H*之间
* 如果都通过的话就把收到的消息写入log
* 如果都通过的话就把收到的消息写入log：PRE-PREPARE和PREPARE消息等
* 如果收到的正确的PREPARE消息大于等于2f，就进入COMMIT阶段，广播COMMIT给其他节点

<REPLY, *v*, *t*, *c*, *i*, *r*>σ*i*

* *v*是view
* *t*是timestamp
* *c*是client
* *i*是节点*i*的编号
* *r*是节点*i*的回复
* σ*i*是节点*i*对消息的签名

COMMIT消息处理：

* 检查COMMIT消息签名是否正确，消息的摘要是否正确
* 判断当前节点的view是否是*v*
* 判断*n*是否在*h*和*H*之间
* 如果都通过的话就把收到的消息写入log
* 如果收到的正确的COMMIT消息大于等于2f+1个(包括自己的)，就回复REPLY消息给client

Raft vs. PBFT

共识算法	Raft	PBFT
适用环境	私链	联盟链
算法通信复杂度	O(n)	O(n2)
最大故障和容错节点	N-f>f	N-2f>f
流程对比	• 初始化leader选举（谁快谁当） • 共识过程 • 重选leader机制	• 初始化leader选举（按编号轮流） • 共识过程 • 重选leader机制

N是网络节点总数，f是网络中异常节点数。

N–f的含义：

①正常的节点数>异常节点数，②正常节点数要多于一半，③异常节点数不多于50%

N–2f的含义：

①异常节点数不多于1/3。

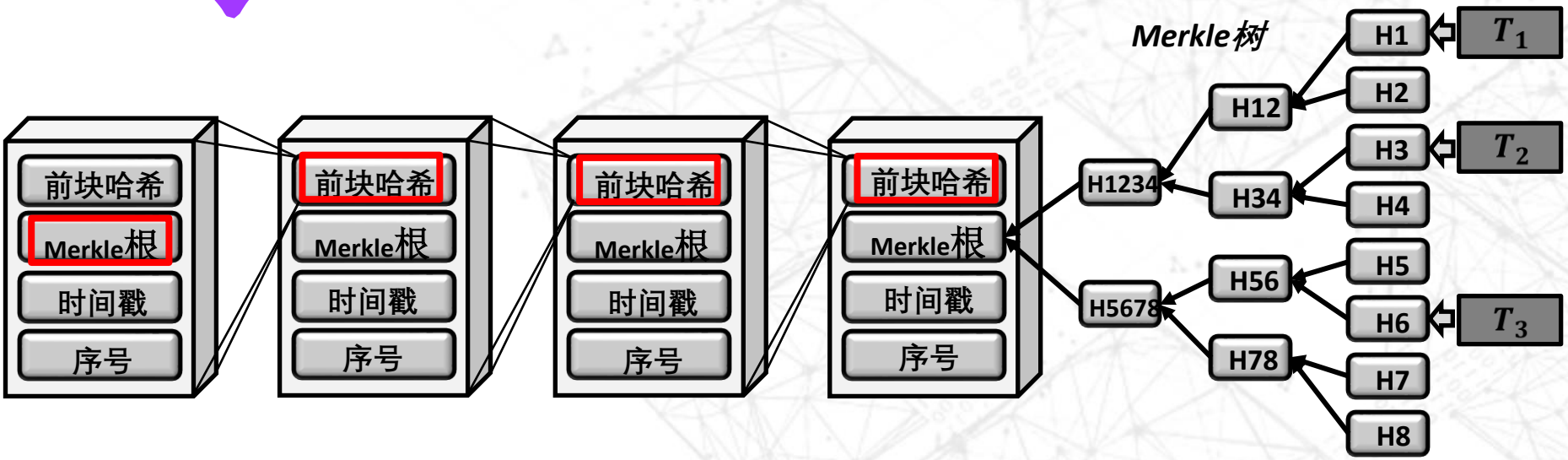
拜占庭将军问题

MAX 

MIN 

30秒讨论时间

PoW (Proof-of-Work)



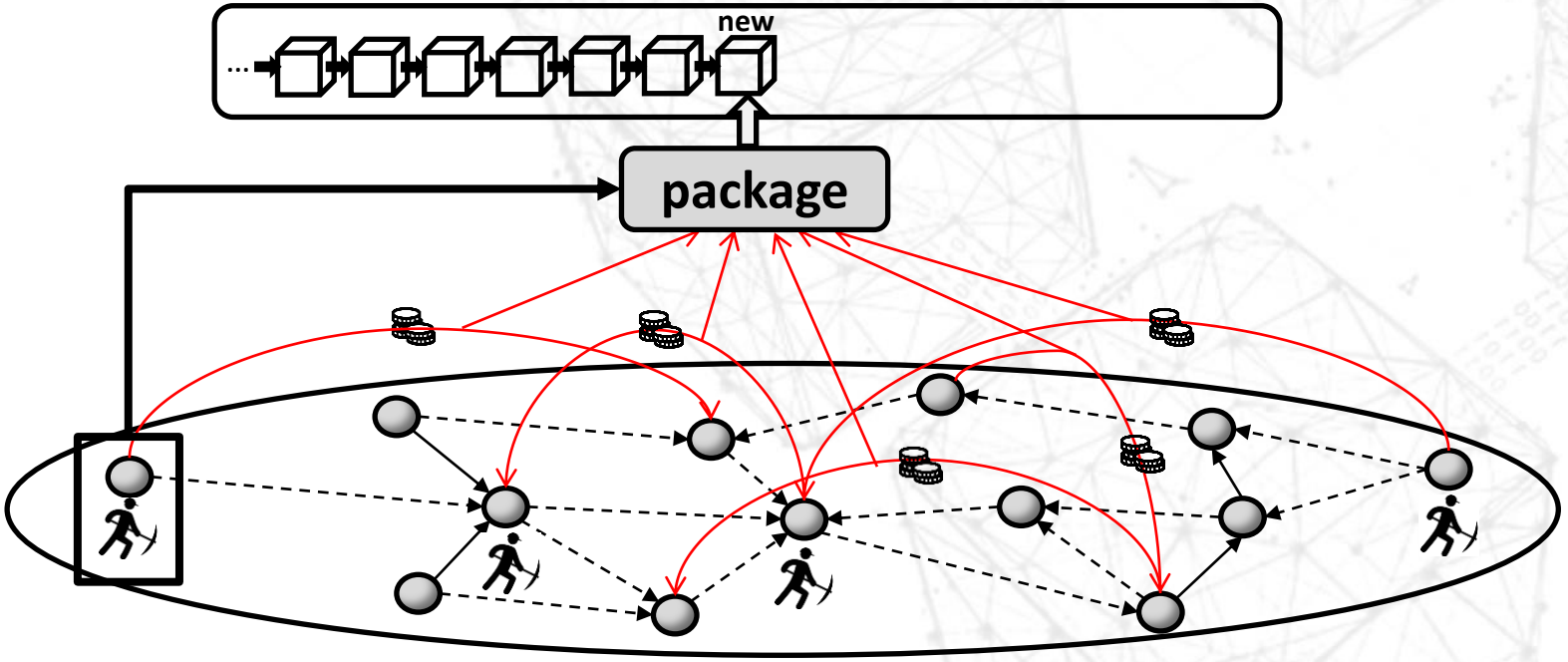
Merkle树： 嵌套的哈希结构，任何一个交易(T)数据的改动，都会导致Merkle根变动。
前块哈希： 前一个区块(block)中任何数据的改动，都会导致前块哈希值的变动。

问题1： 当图中最左侧区块中某一交易数据发生改动，图中最右侧区块中会发生什么变化？

30秒讨论时间

PoW

共识1：区块链算力竞赛的优胜者有权生成区块。

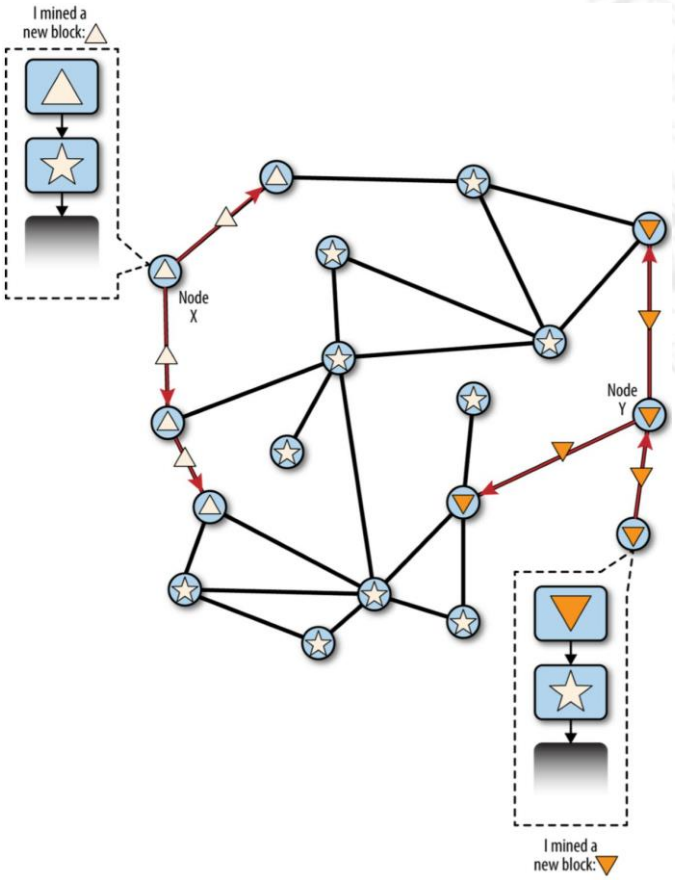


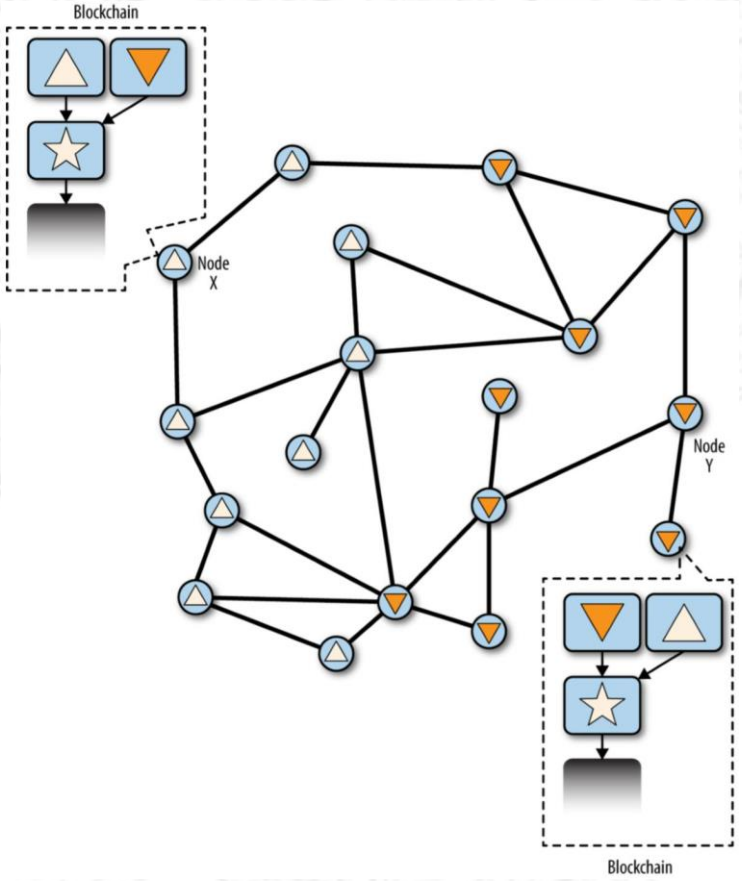
问题：同时有两个优胜者怎么办？

信息的传输有时延，先听到谁的，就认为谁是优胜者。

PoW

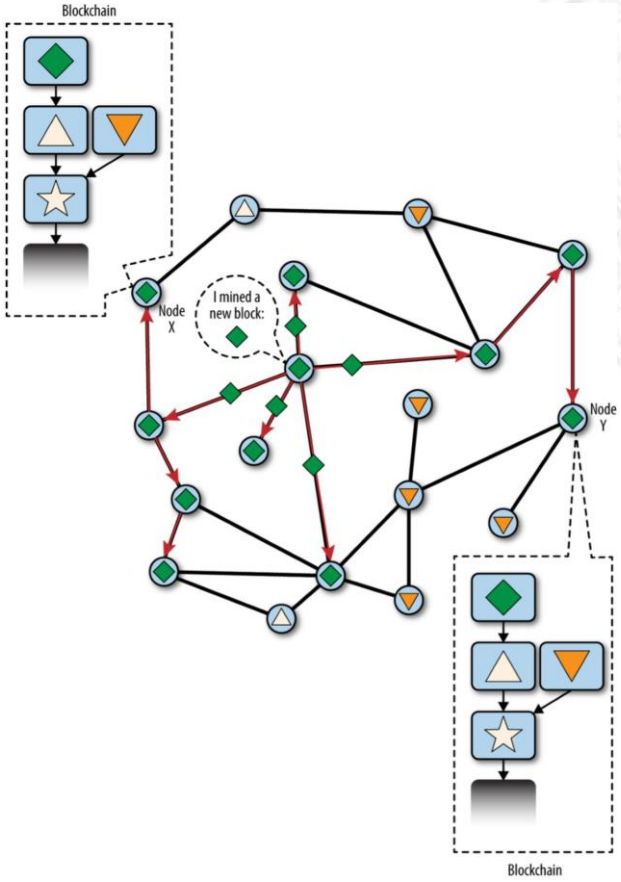
共识2：多个优胜者时，以先接收到为暂时优胜者。

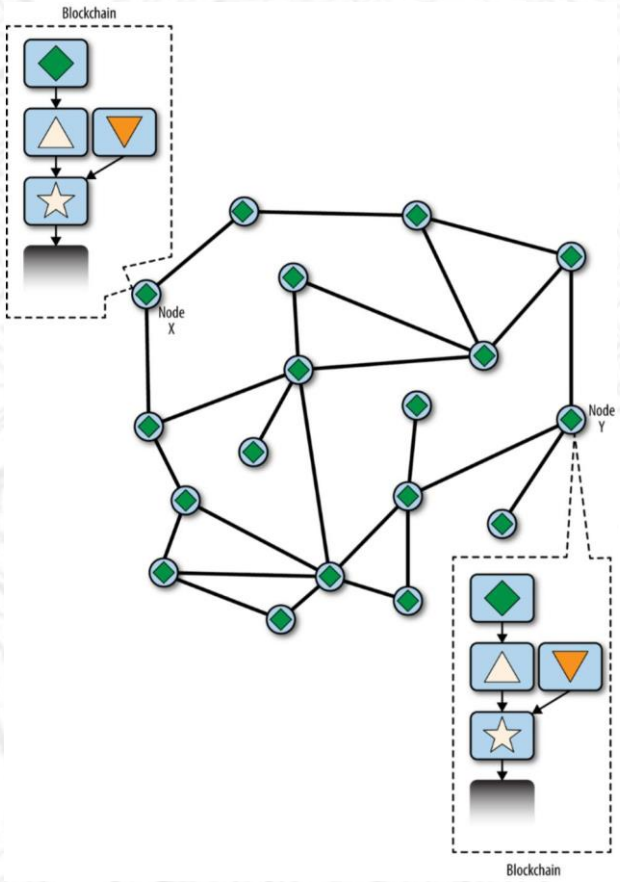




PoW

共识3：以最长区块链为准。





PoW

问题：当连续两轮算力竞赛同时出现两个优胜者时，怎么办呢？

答案：等第三轮。

PoW

对比要看，很重要。

共识算法	Raft	PBFT	PoW
适用环境	私链	联盟链	公链
算法通信复杂度	O(n)	O(n2)	O(n)
最大故障和容错节点	N-f>f	N-2f>f	51%
流程对比	• 初始化leader选举（谁快谁当） • 共识过程 • 重选leader机制	• 初始化leader选举（按编号轮流） • 共识过程 • 重选leader机制	• 解题 • 如果解开，广播答案，此轮胜出；如未解开，进入下轮。 • 以最长链为合法链。

PoS (Proof of Stake)

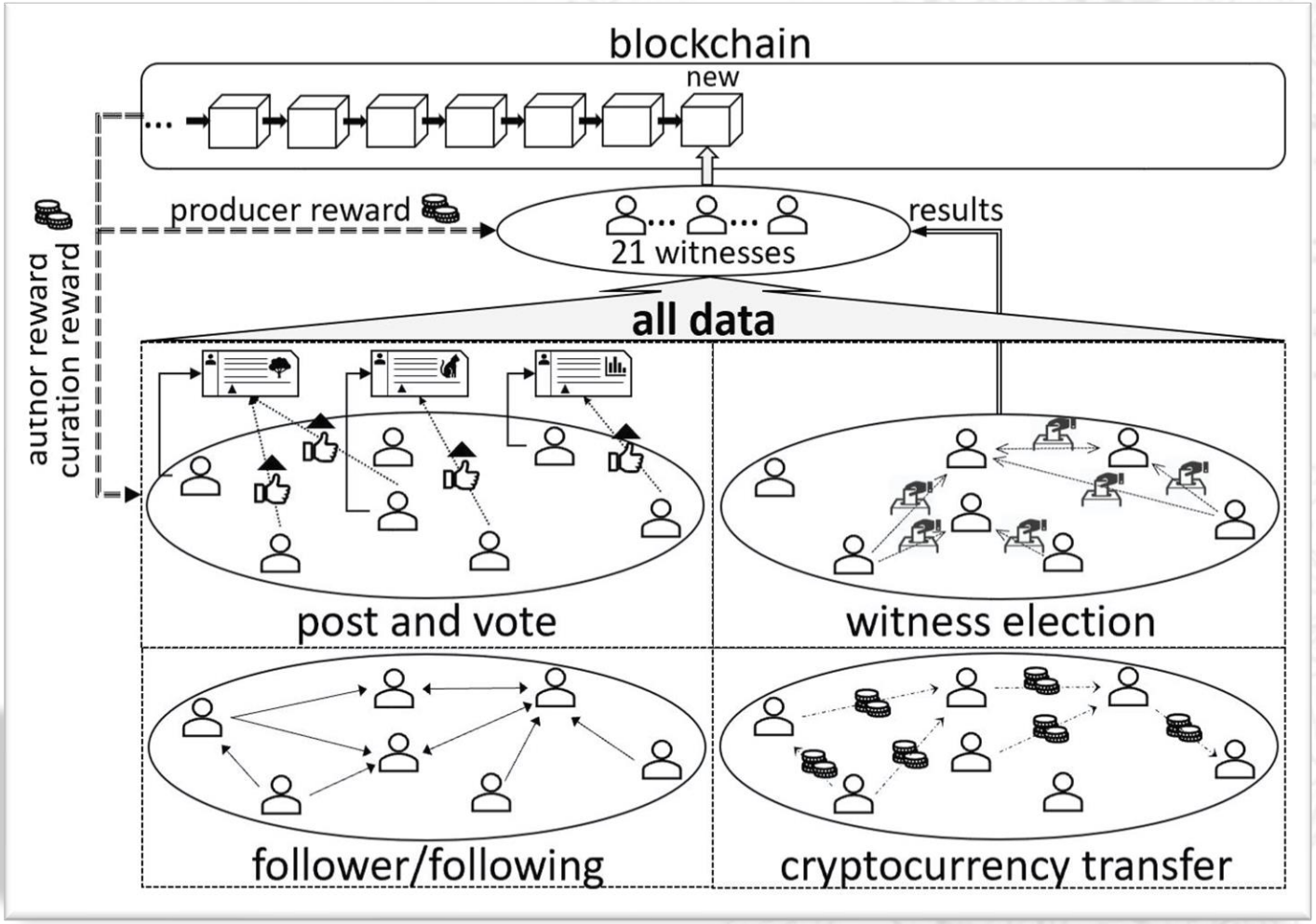
POS和POW的区别：

- 在 PoS 中，一个账户的余额越多，就越容易发现下一个区块。

越容易拿到记账权，越有钱的人越有动机去维护区块链的安全性。

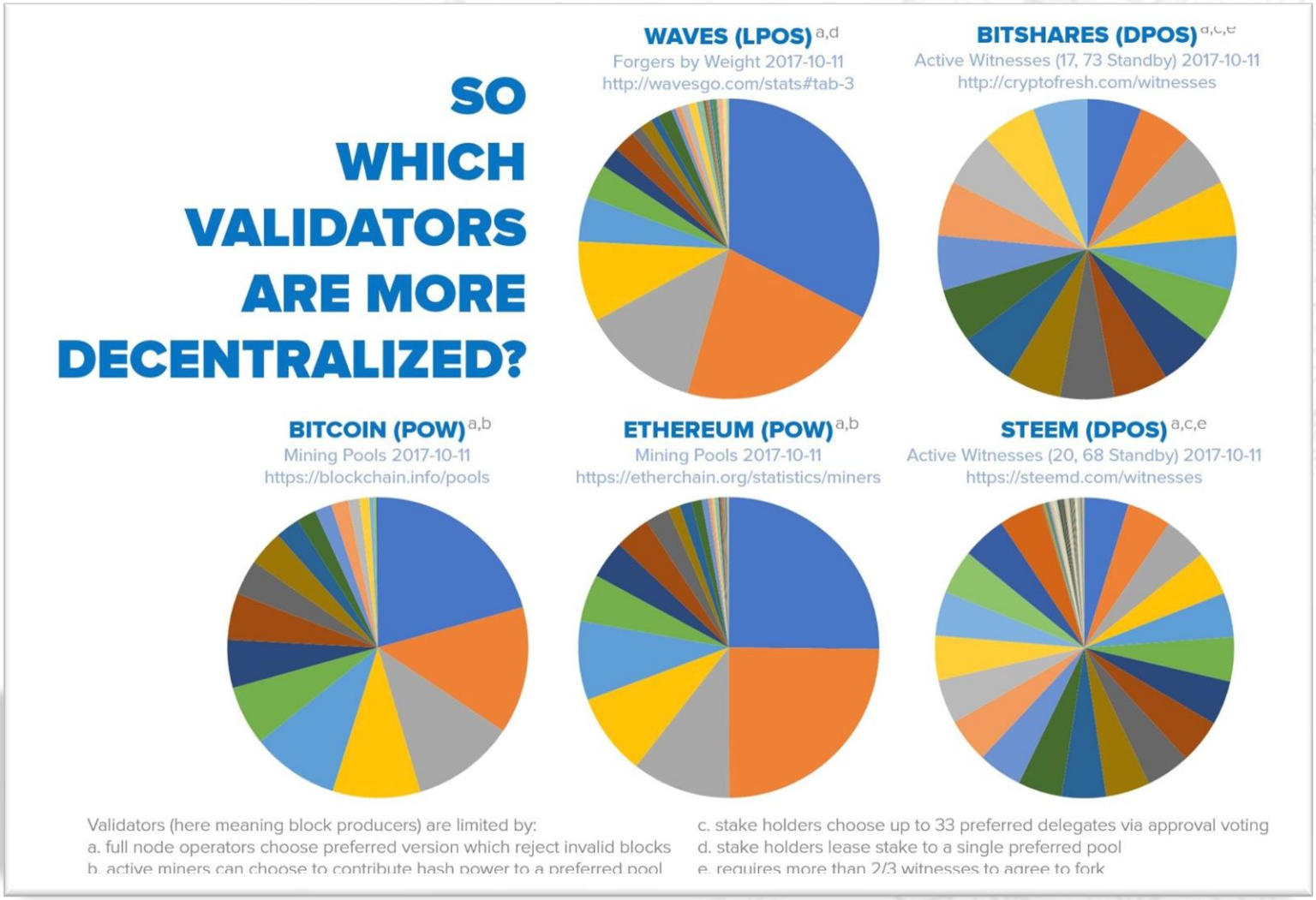
- 优点：
  - 降低资源浪费
  - 区块生产者与区块链发展的经济利益一致
- 攻击
  - 长程攻击 （Long Range Attack）
  - 无利害攻击（Nothing at Stake）

DPoS


blockchain
new
producer reward
21 witnesses
results
autnor reward
curation reward
all data
post and vote
witness election
follower/following
cryptocurrency transfer

PoW vs. DPoS

不考



共识安全

双花问题

同一笔钱，支付2次

双花：重复支付。

中心化系统中很好解决。

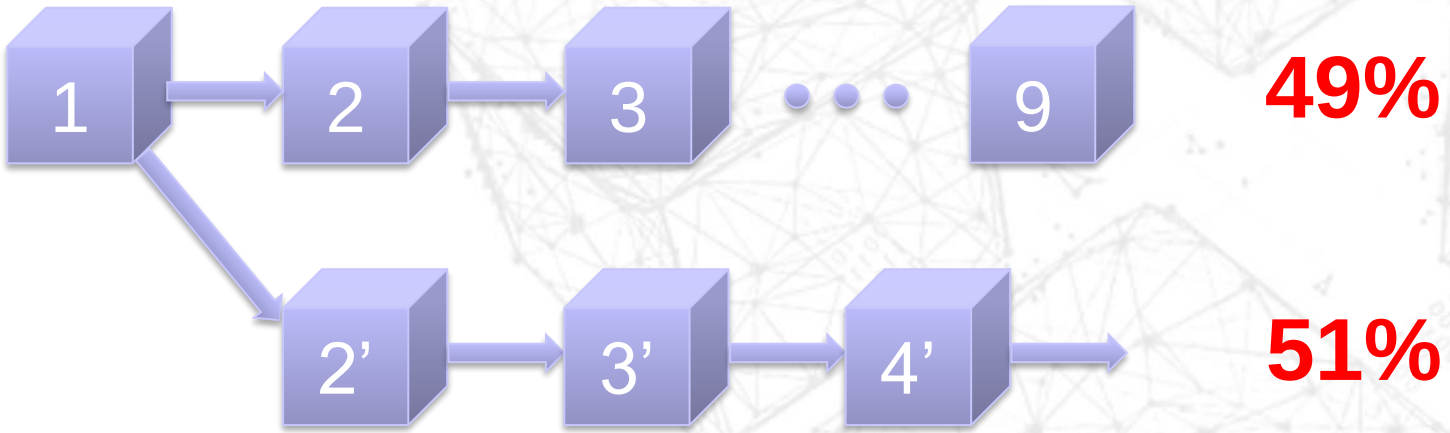
去中心化系统中较难解决。

51%攻击

比特币的共识机制依赖于这样一个前提，那就是绝大多数的矿工，出于自己利益大化的 考虑，都会 通过诚实地挖矿来维持整个比特币系统。

然而，当一个或者一群拥有了整个系统中大量算力的矿工出现 之后，他们就可 以通过攻击比特币的共识机制来达到破坏比特币网络的安全性和可靠性的目的。

51%攻击



博弈论与风险模型

- 许多去中心化的协议都假设至少 51% 的网络协议参与者都是诚实可靠的。
- 比特币假设超过 51% 的参与者都会按照符合自己经济利益最大化的规则行事。
- 激励的相容性，即意味着每一个协议参与者自身的最优决策也就是整个系统的最优决策。

算力集中的现象

矿机分类

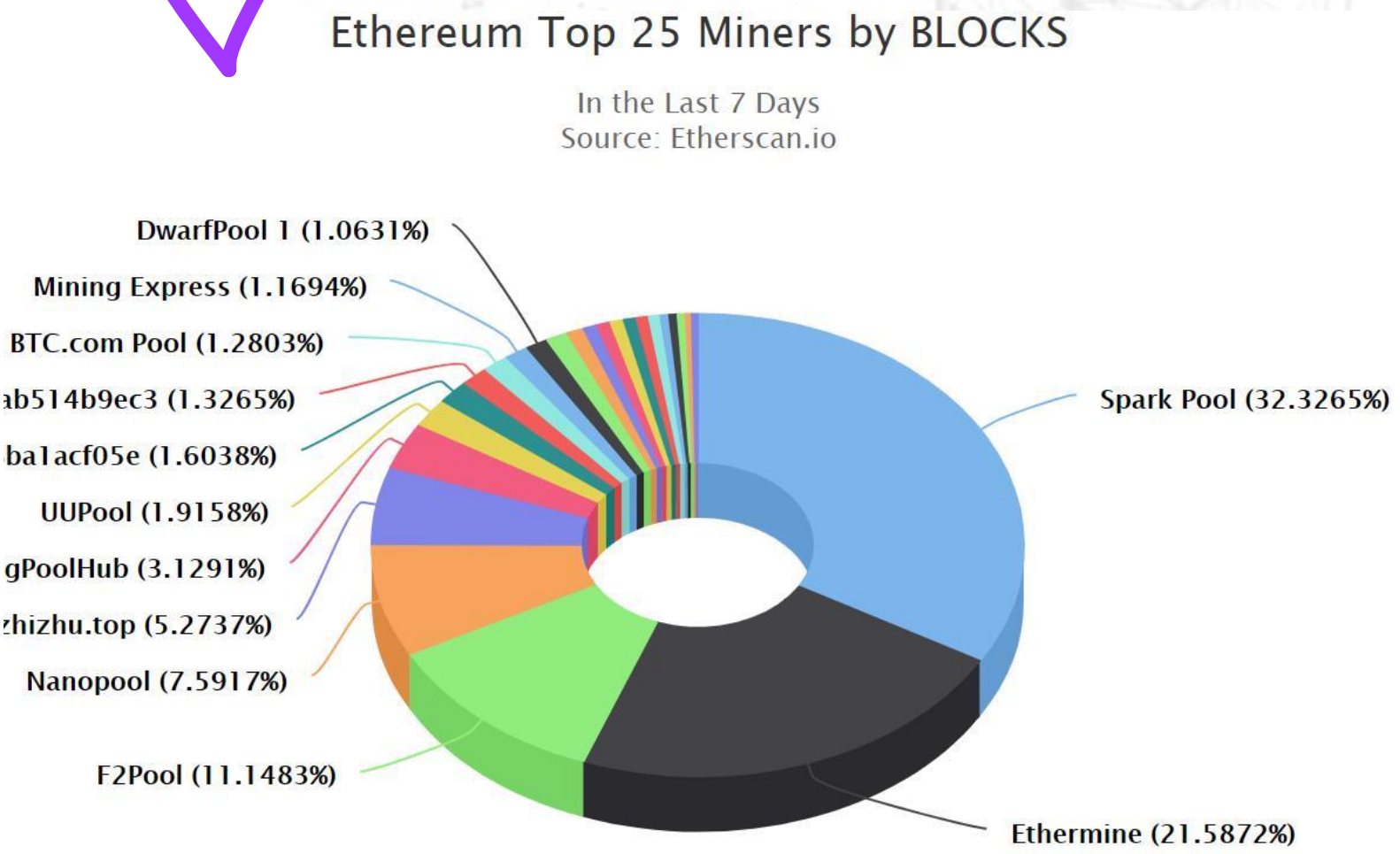
CPU

GPU

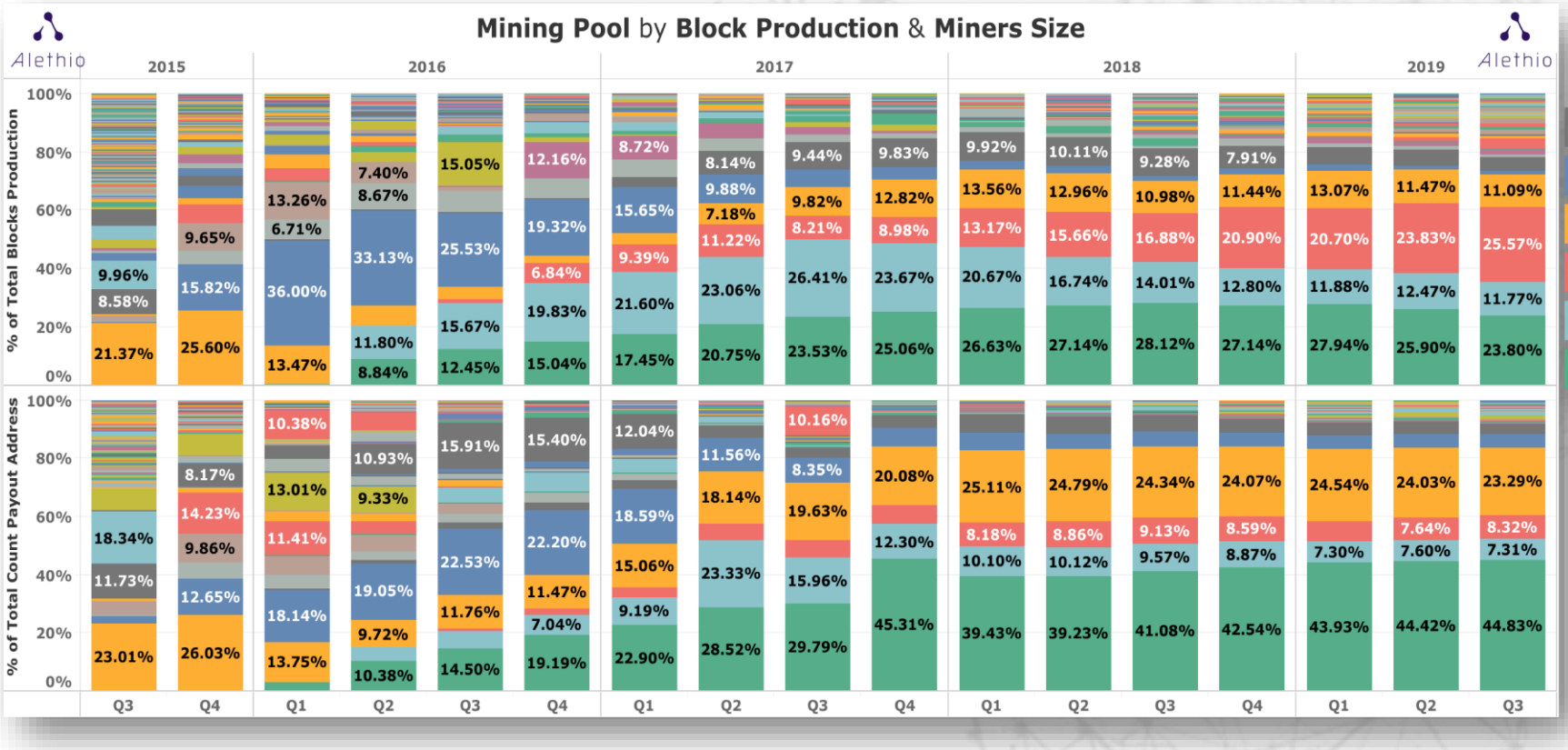
ASIC

大矿池

低算力的矿工抱团，不再单打独斗，而是集中算力，抽成。



大矿池

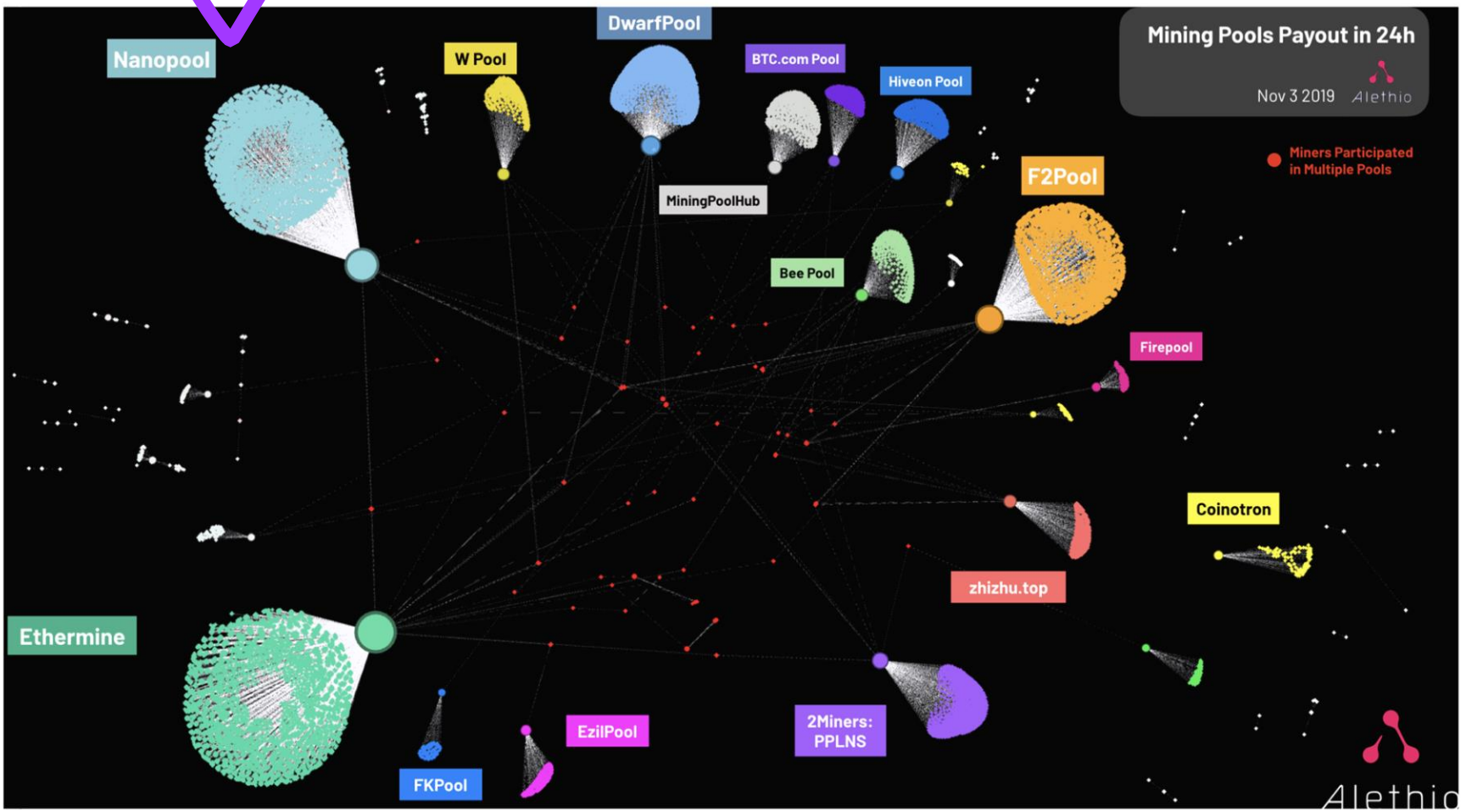


Ethereum： Mining Pool Block Production and Mining Payout as percentages of totals | 2015 – 2019

Source: https://consensys.net/research/measuring-blockchain-decentralization/

51%攻击的4个诱因

大矿池



Ethereum： Miner relationship with mining pools, visualized by payouts | 11.03.19

算力市场

- 矿机所有者可以把矿机租赁给更有经验的矿工，增加了算力市场所有参与者的利润，创造了一个双赢的局面。

- 然而，算力市场的负面效果在于为攻击者实施攻击提供了一个可以快速租借大量算力的平台。

- 算力市场对使用专用挖矿设备的密码学货币远没有那么大影响。

硬件熊市

- 硬件熊市不论是对使用多链硬件还是使用专用硬件的密码学货币都会产生影响。
- 现在有专用于以太坊和 **Zcash** 挖矿的 **ASIC** 了，这使得租用 **GPU** 以攻击更低价值密码学货币的成本大大降低。
- 面对硬件熊市，比特币也不能幸免于难。据估计，多达三分之一的比特币算力被已经破产的矿场甩卖。
- 熊市对制造商造成的影响尤为严重。据估计，比特大陆、芯动科技、台积电，甚至三星都因硬件价格突然下跌遭受重大损失。

区块奖励的影响

每隔4年，区块奖励除以2，长期以往下去会向0接近，当后期挖的钱少了，挖的人少了，那么就很容易集中矿力来进行51%攻击。

- 由于硬件的获取和操作成本非常高昂，密码学货币对“双花”攻击的抵御很大程度上取决于其区块奖励。
- 密码学货币抵御攻击的能力与其挖矿硬件数量成正比，因此，如果区块奖励低到没有大量硬件用于该货币的挖矿，这种货币受到的保护将变得微乎其微。

成本评估

- 谈到密码学货币的安全性，我们必须考虑发起一次 51% 攻击所需的美元成本。如果一种密码学货币的所有矿机总价值是一百万美元，那么显而易见，任何超过一百万美元的交易都极易遭受 51% 攻击，因为交易对手只需要花一百万美元来购买或制造矿机，就能发动“双花”攻击。

- 评估一种密码学货币的矿机总价值难乎其难，评估制造一套足够发动 51% 攻击的硬件设施的成本亦是如此，但根据一般经验，该成本应该等同于 6 到 24 个月的区块奖励。通常矿机市场的开放竞争会确保其价值处于这个范围。

密码学货币空头

- 问其他人借来币，在当前的价格卖出。
- 坐等价格下跌到5元。
- 用下跌后的低价买币，还币。

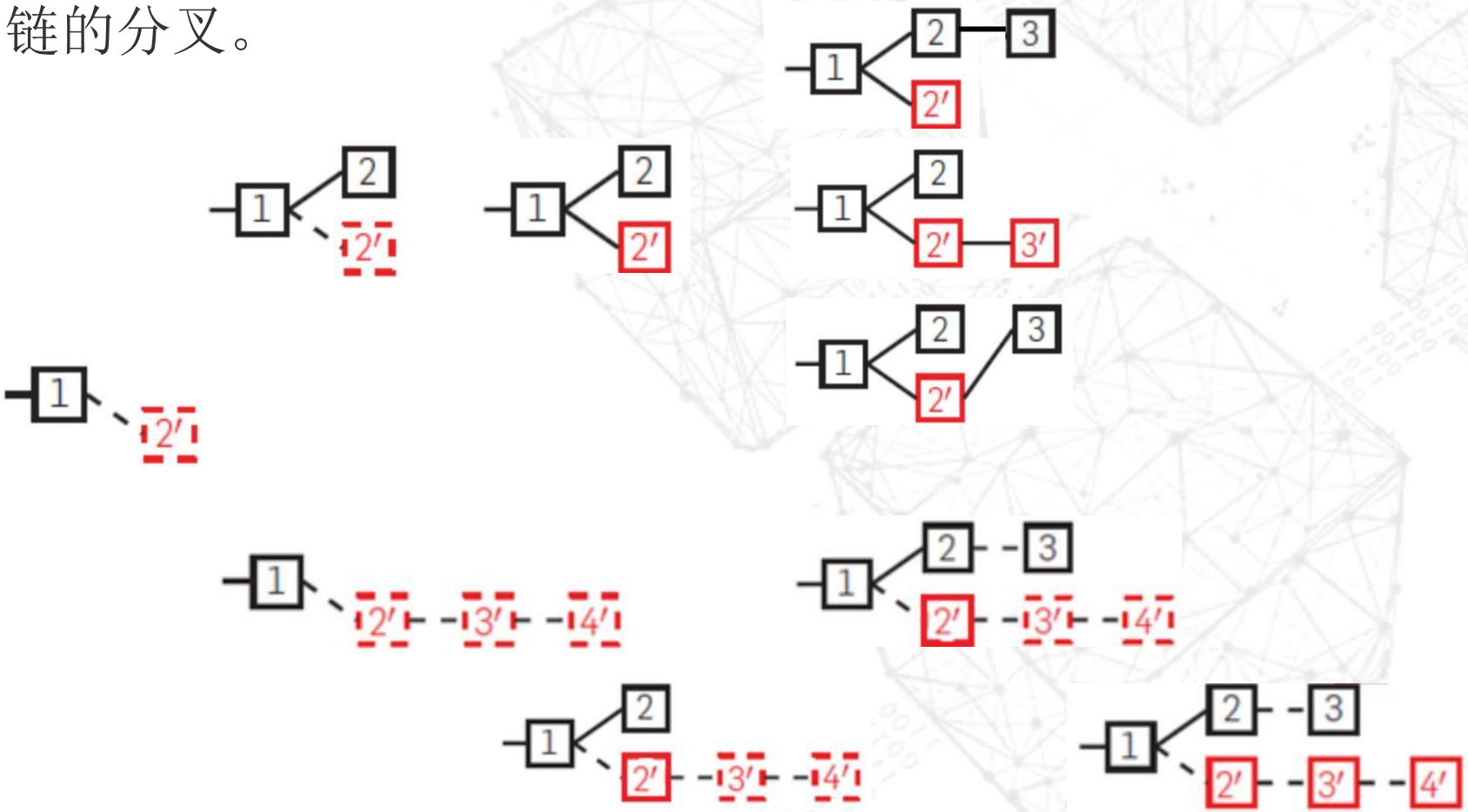
现有防御方法

增加确认时长

地址黑名单

自私挖矿

“自私挖矿”攻击的核心思想是“自私挖矿”矿池故意延迟公布其计算得到的新块，并构造一条自己控制的私有分支，造成链的分叉。



总结 | CONCLUSION

1 分布式共识

2 区块链类型与共识协议

3 经典共识协议

3 共识安全