Bitcoin基础：共识机制

吴嘉婧 副教授

中山大学

计算机学院

Outline & Keywords of this Class

Part 1: What is Consensus

Part 2: Why Bitcoin needs consensus

& What can consensus do for Bitcoin

Outline & Keywords of this Class

Part 1: What is Consensus

Part 2: Why Bitcoin needs consensus

& What can consensus do for Bitcoin

什么是共识?

Definition: **Consensus** is

- A process of agreement between distrusting nodes on the final state of data.
- 不信任节点之间就数据的最终状态达成一致的过程。

To achieve consensus, different algorithms are used

An agreement

- between two nodes is easy;
- It becomes pretty a challenge to achieve consensus among multiple nodes

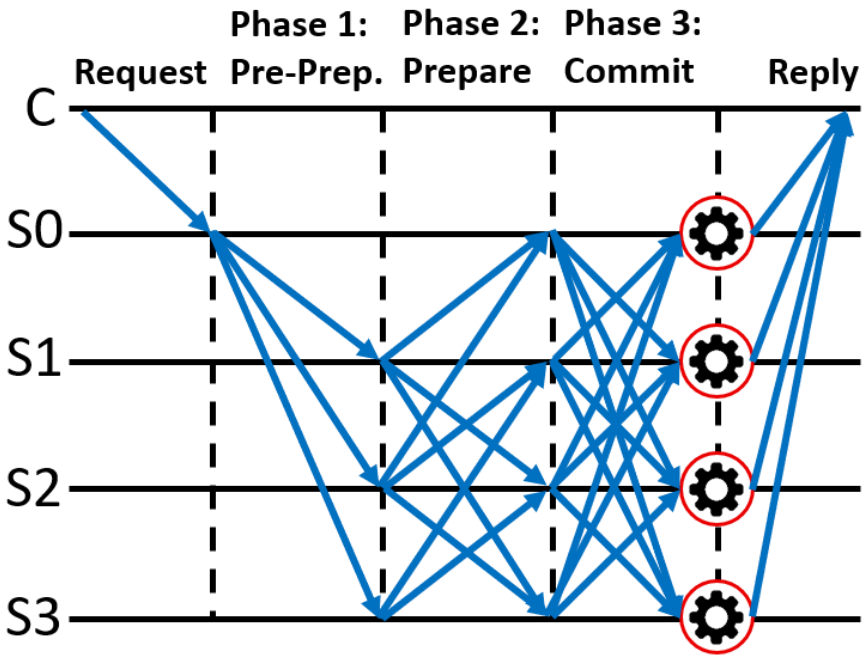
Distributed consensus

- To attain a common state/value among multiple nodes despite the failure of some nodes

什么是 共识机制?

共识机制指的是

- 在一个分布式系统中（例如区块链）的大多数或全部节点就某个状态或取值达成一致的一系列步骤



什么是 区块链的共识?

It is the backbone of a blockchain

It provides decentralization of control through an optional process known as mining（挖矿）

Not all consensus mechanisms are suitable for all types of blockchains

- Public permissionless blockchains: PoW（工作量证明）
- Private blockchains: *Proof of Authority*（权威证明共识）
- To choose an appropriate consensus algorithm for a particular blockchain system

Requirement of 共识机制

- **Agreement 达成共识**
  - honest nodes decide on the same value可信节点共同决定的值
- **Termination 可终止性**
  - honest nodes terminate execution of the consensus process

  不会陷入死循环
- **Validity 合法性**
  - agreed value == initial value proposed by at least one honest node

  每轮共识的结果值，都是由最少一个可信节点提议的
- **Fault tolerant 容错性**
  - should be able to run in the presence of faulty or malicious nodes (Byzantine nodes)在出现错误和恶意节点时也可以运行
- **Integrity 完整性**
  - no node can make the decision more than once in a single consensus cycle没有一个节点可以在单个共识周期中多次做出决策

举例：共识机制

相信在大家读书的时候，每个班上但都有班花，是大家心目中的女神

举例：共识机制

所谓的班花 指班级中被大家公认或选举的最漂亮的女孩子，也可以理解为班级里最受大家认可、优秀的女生。

Agreement 达成共识

确认班花是谁，这就是区块链伟大的共识机制，能在众多学生节点中达到一种较为平衡的状态，保障班级能在班花的带领之下正常开展一系列活动。

Requirement of 共识机制

- **Agreement 达成共识**
  - honest nodes decide on the same value可信节点共同决定的值
- **Termination 可终止性**
  - honest nodes terminate execution of the consensus process

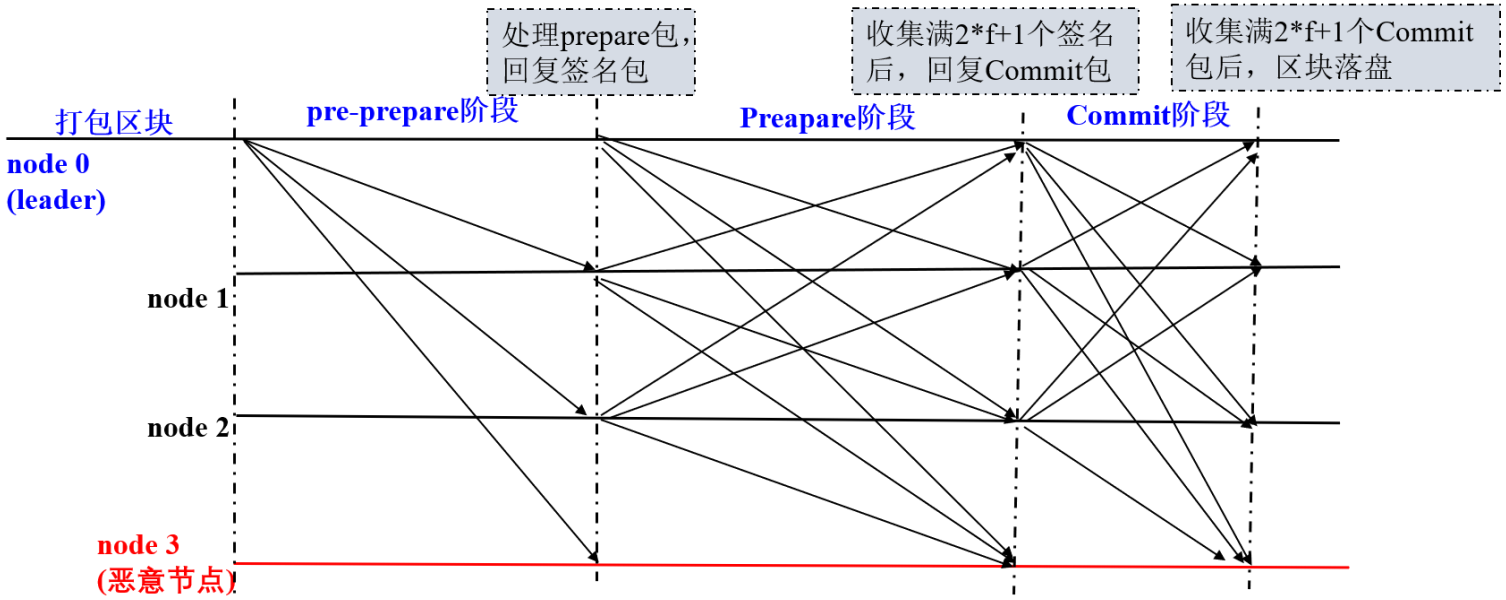
  不会陷入死循环
- **Validity 合法性**
  - agreed value == initial value proposed by at least one honest node

  每轮共识的结果值，都是由最少一个可信节点提议的
- **Fault tolerant 容错性**
  - should be able to run in the presence of faulty or malicious nodes (Byzantine nodes)在出现错误和恶意节点时也可以运行
- **Integrity 完整性**
  - no node can make the decision more than once in a single consensus cycle没有一个节点可以在单个共识周期中多次做出决策

Purpose of 共识机制

All consensus mechanisms are designed to

- deal with faults in a distributed system
- allow distributed systems to reach a final state of agreement

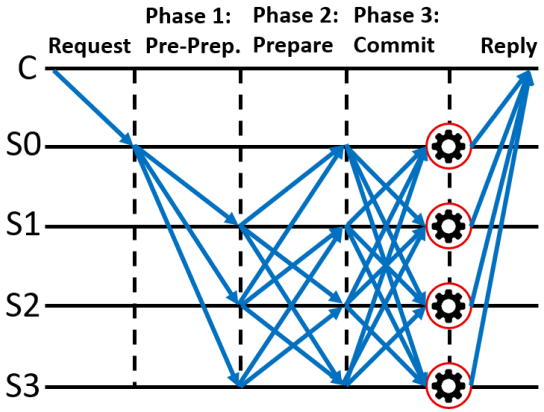


Types of 共识机制

两大类

- **拜占庭容错类** (BFT: Byzantine Fault Tolerance)
- **领导人选举类**（LE：Leader Election-based）

BFT-based



Leader Election-based
e.g.,
PROOF OF WORK

(Supplementary)

Permissioned / Permissionless Blockchains

- Permissionless
  - 几乎人人都可以参与，每个参与者都是匿名
  - 参与者之间不存在任何信任
  - 采用"挖矿"作为激励机制
- Permissioned
  - 在一组已知的、已识别的、且经常经过审查的参与者中进行区块链的操作
  - 具有一定的信任
  - 不需要挖矿机制进行

共识机制类型-1： BFT-based

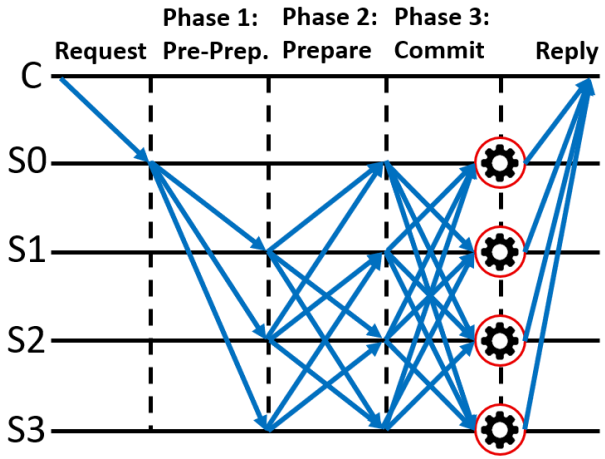
没有计算密集型操作，例如部分哈希反转

它依赖于一个简单的节点方案，这些节点是发布者签名的消息

最后，当收到一定数量的消息时，就会达成协议。

优缺点?

- 在节点数量有限的时候表现良好
- 但是可扩展性差



共识机制类型-2： Leader Election-based

它要求节点参与领导人选举的竞争

只有获胜的节点才能提出最终值

- e.g., 比特币的工作量证明

优缺点?

- 可扩展性好
- 计算慢

Outline & Keywords of this Class

Part 1: What is Consensus

Part 2: Why Bitcoin needs consensus

& What can consensus do for Bitcoin

Outline: Section II

比特币共识机制

- 身份确认
- UTXO交易模型
- 交易信息记录
- 工作量证明

比特币共识机制 关键词

- **比特币身份确认　身份确认：**公私钥体系
- **UTXO交易模型　交易服务**
- **链式交易信息记录　记录管理**
- **工作量证明(POW)　信任规则**

为何需要共识?

从交易的角度，不共识会有什么样的场景?

- 一个提议 (e.g., a Tx, or a block) 广播出去，有人收到，有人没收到

- 假如有人发送假消息，怎么办?
  - 假如有人尝试：一个 coin 花两次

共识机制：区块链的核心引擎

定义：

一种多方协作机制，用于协调多参与方达成共同接受的唯一结果，且保证此过程难以被欺骗，且持续稳定运行

共识模型里体现的博弈论

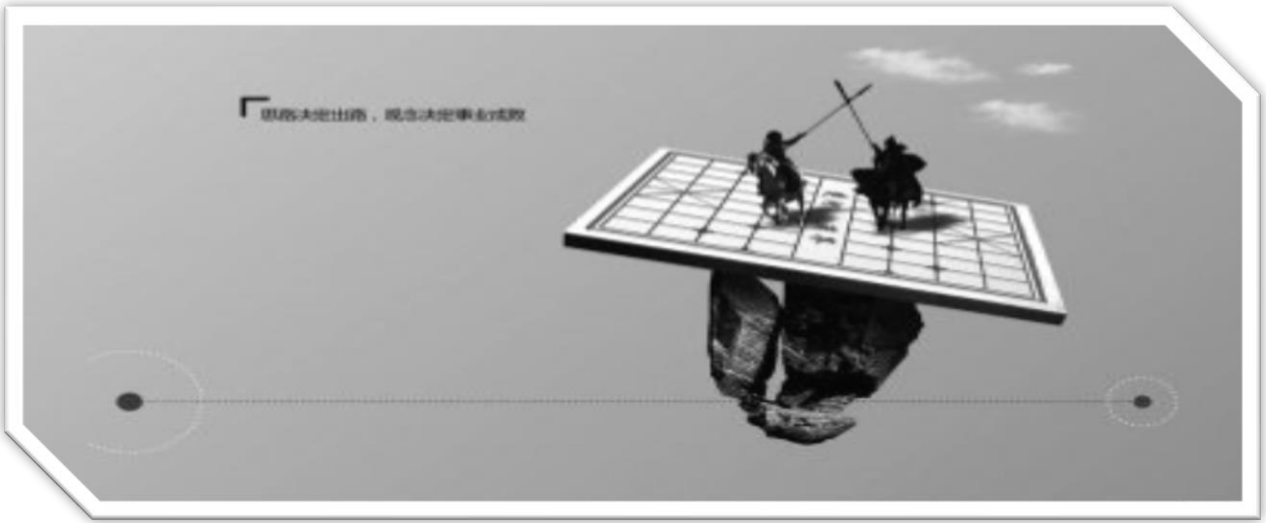
拥有记账权的人更倾向在维护整个体系过程中获利（纳什均衡+帕累托最优）

使用网络的人需要付出一定的成本（手续费、计算费）以免滥用（避免公地悲剧）

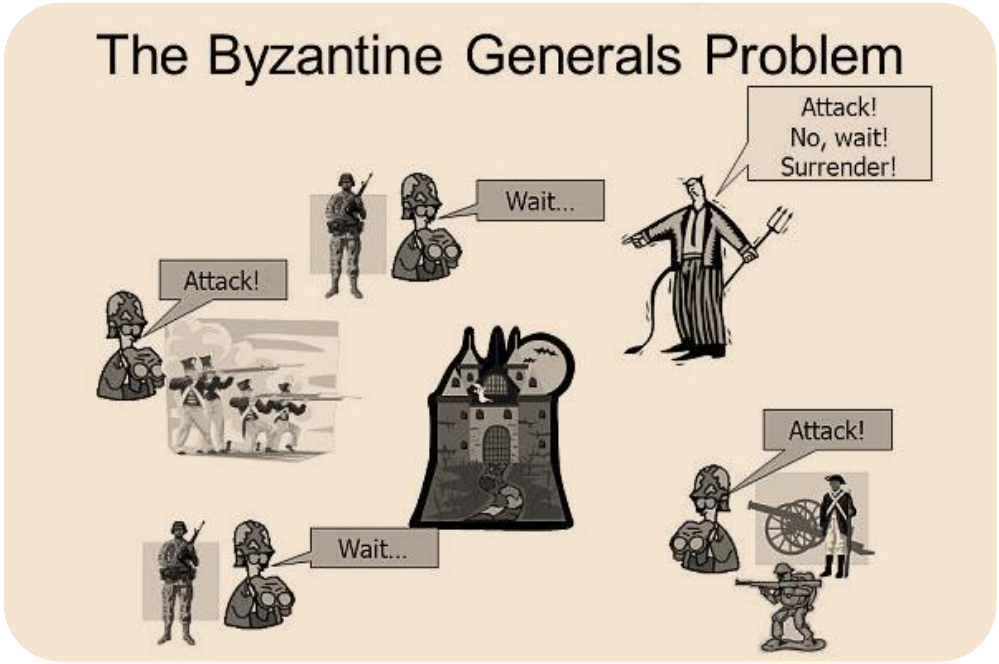
少数人作恶的成功几率很低，参考“赌徒破产问题”（Gambler‘s Ruin problem）

只有极端势力才有可能不顾一切的颠覆这个体系

整个局势不存在“确定性”，一直在动态的多方博弈

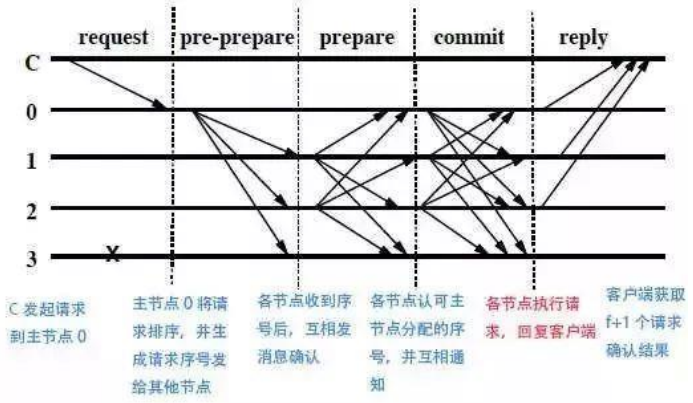


拜占庭将军问题和解决方案 (Byzantine Fault Tolerance)



预设条件

- 至少4个以上参与者
- 每轮次有一个发令者
- 少于1/3的参与者作恶或失效
- 极大概率可达的网络（区块链网络）
- 可控的网络规模（少于100参与者）



BFT拜占庭将军问题：Lamport，Shostak 和 Pease 于1982年的一篇学术论文中引入，Miguel Castro 和 Babara Liskov 在1999年提出 PBFT，放松了约束来解决拜占庭问题。Liskov于2008年获得了图灵奖

拜占庭将军问题相关的图灵奖得主

Lamport分布式计算理论奠定了这门学科的基础。他在1978年发表的论文《分布式系统内的时间、时钟事件顺序（Time, Clocks, and the Ordering of Events in a Distributed System）》成为计算机科学史上被引用最多的文献。他为“并发系统的规范与验证”研究贡献了核心原理。

2008年，美国计算机协会(ACM) 宣布Barbara为当年年度图灵奖获得者，以表彰其在程序设计语言与系统设计，特别是在数据抽象、容错和分布式计算领域的实践和理论基础方面的贡献。

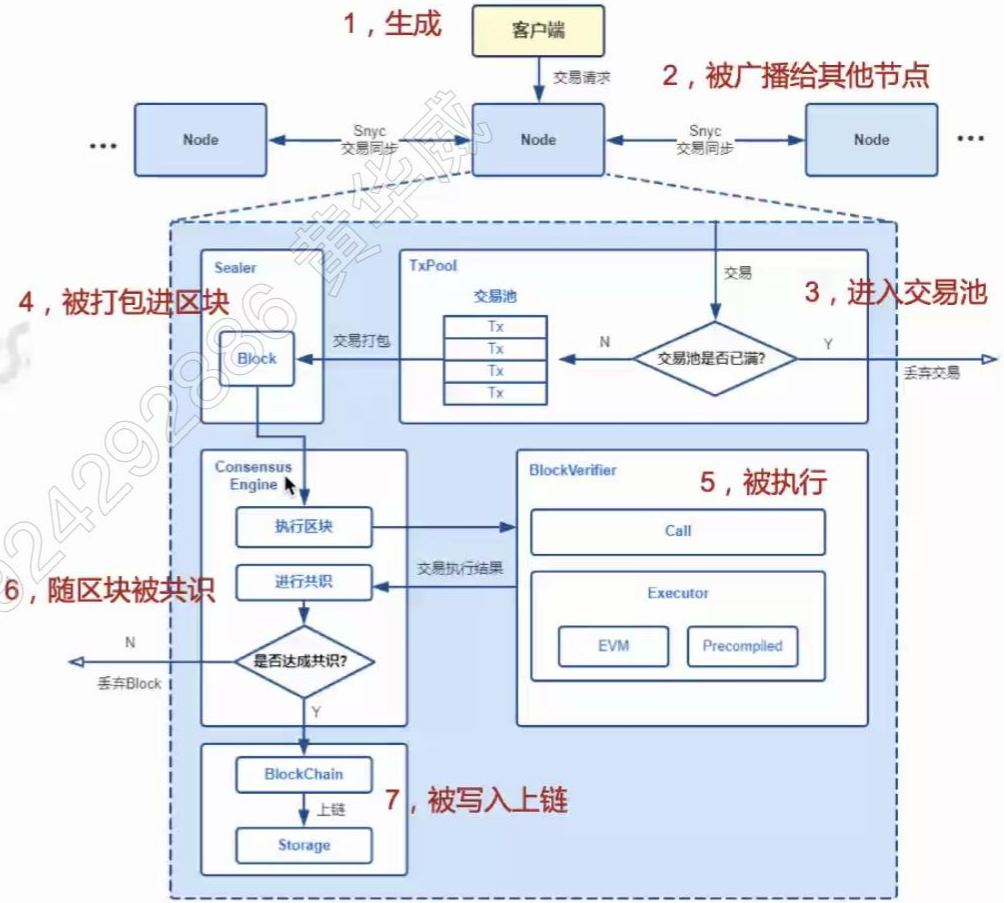
比特币共识机制

- ~~比特币身份确认~~ 身份确认
- UTXO交易模型 交易服务
- 链式交易信息记录 记录管理
- 工作量证明(POW) 信任规则

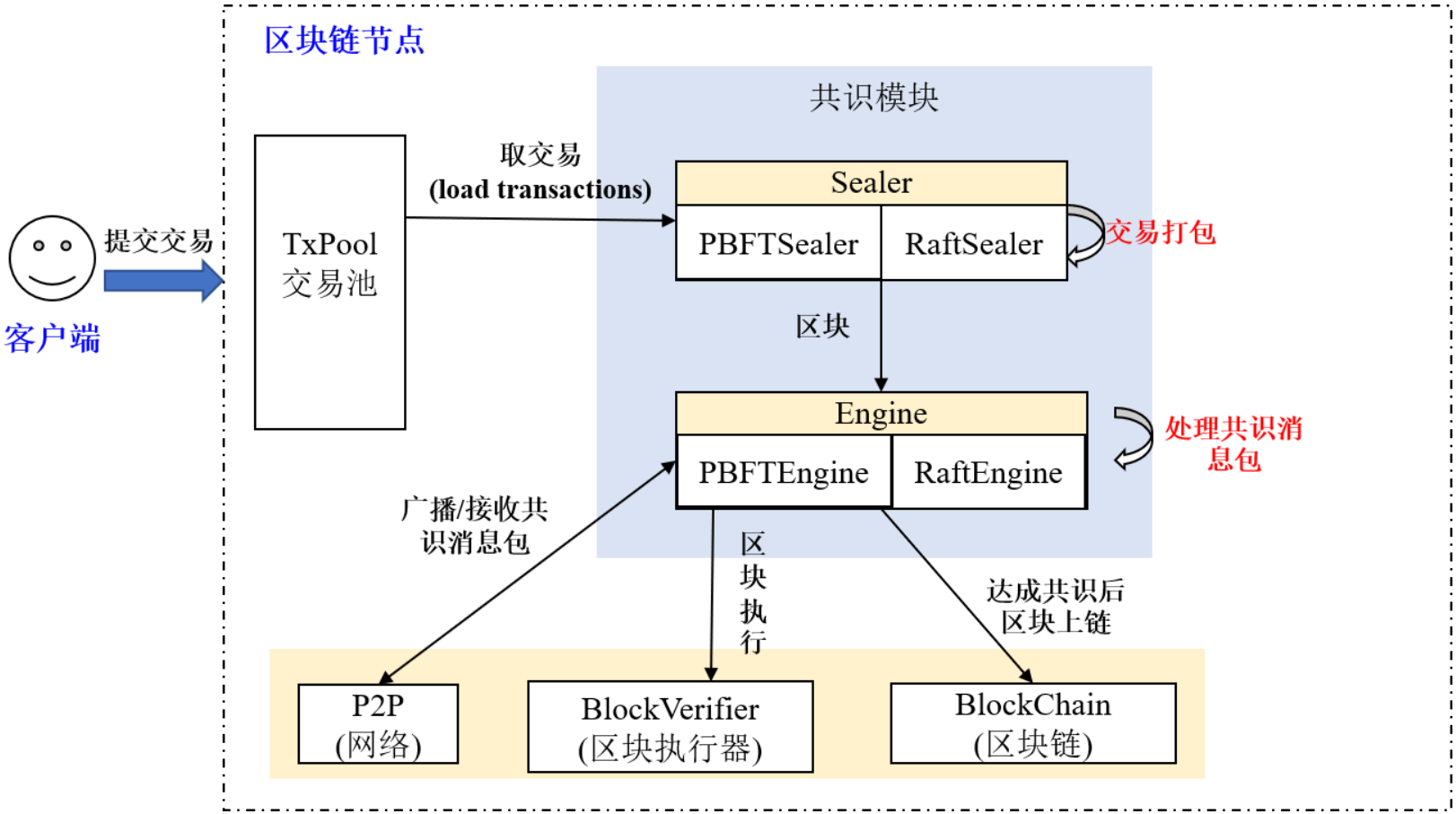
补充知识：交易生命周期

交易生命周期

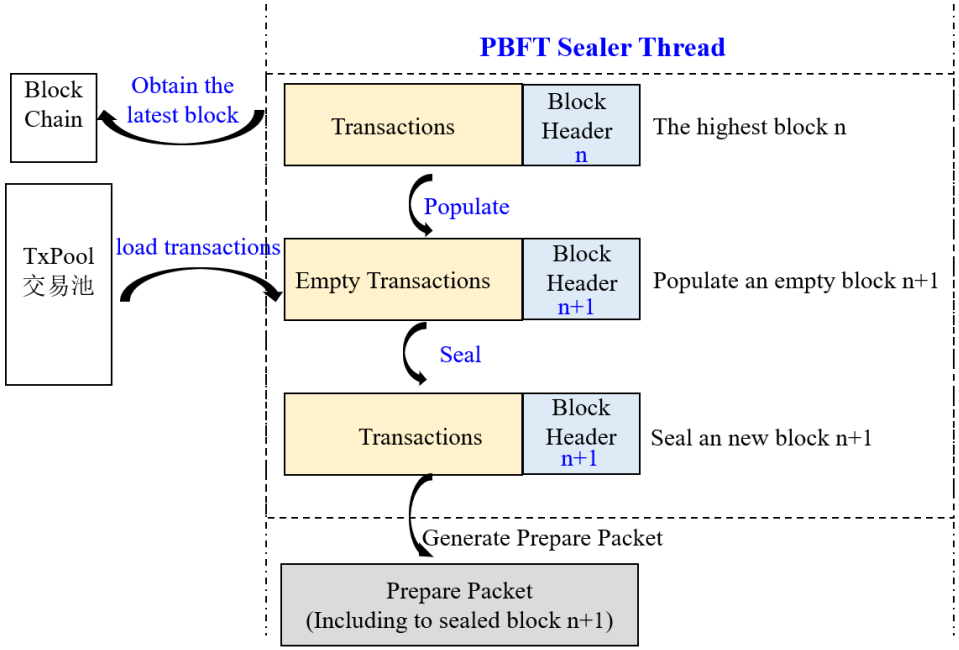
- 网络模块：广播交易和区块
- 交易池模块：缓存交易
- 打包共识模块：交易打包，节点间共识
- 交易执行模块：执行交易，获得状态变更
- 存储模块：落盘存储交易、区块等数据



补充知识：交易打包共识框架



补充知识：交易打包（以PBFT为例）



1. **产生新的空块**: 通过区块链(BlockChain) 获取当前最高块，并基于最高块产生新空块

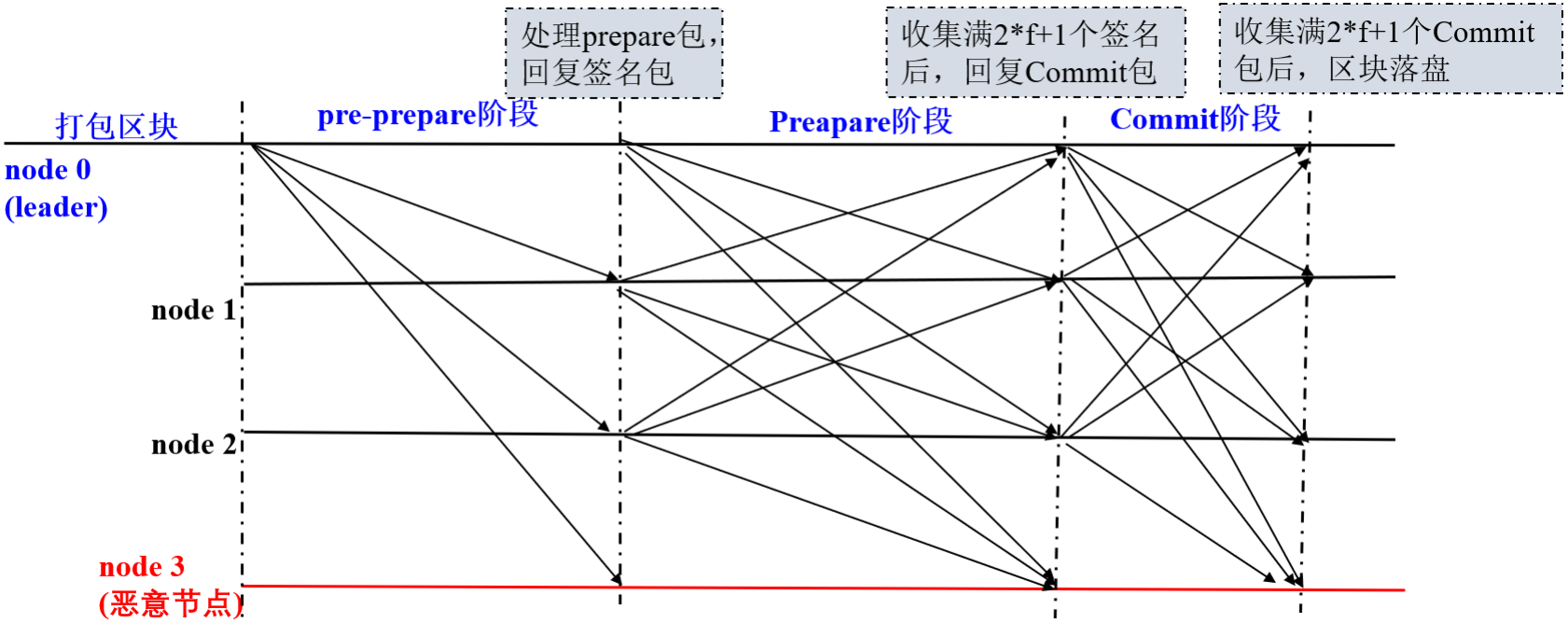
2. **从交易池打包交易**: 产生新空块后，从交易池中获取交易，并将获取的交易插入到产生的新区块中；

3. **组装新区块**: Sealer线程打包到交易后，将新区块的打包者(Sealer字段)置为自己索引，并根据打包的交易计算出所有交易的transactionRoot；

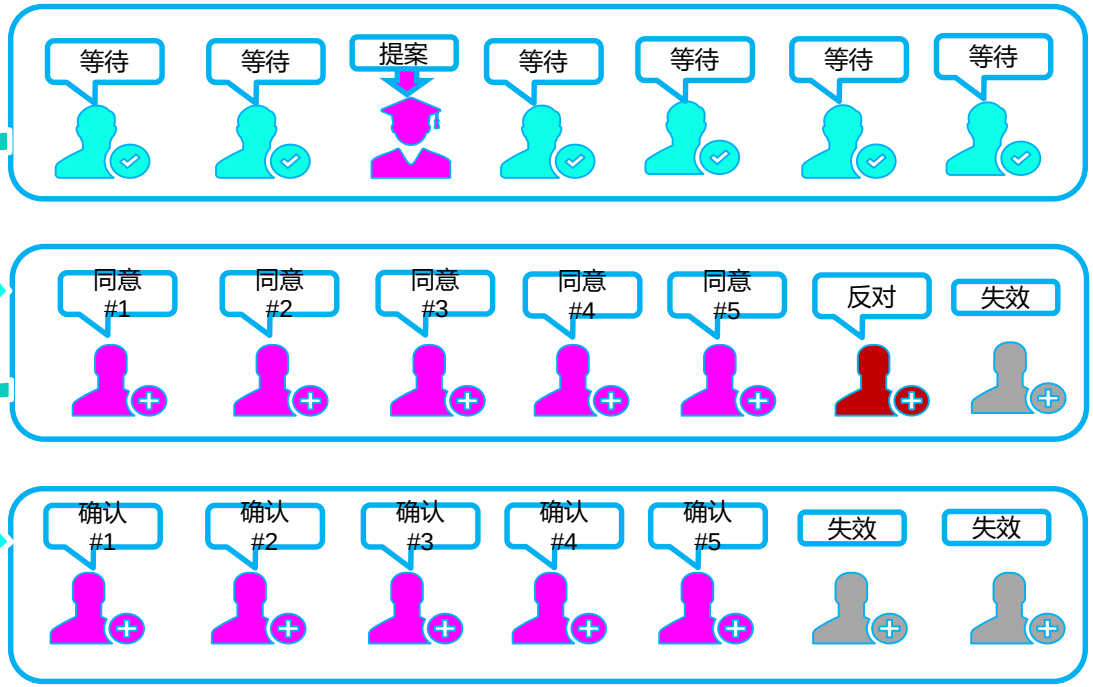
4. **产生Prepare包**: 将组装的新区块编码到Prepare包内，通过PBFTEngine线程广播给组内所有共识节点，其他共识节点收到Prepare包后，开始进行三阶段共识

补充知识： PBFT共识算法

- 广播模型，三次广播，容错1/3节点故障



补充知识：一次PBFT协商过程: 民主集中制



(提议：本小组周二上午开会)

确认记账者列表,每一轮次选出新的提案人,提案人排序打包，广播提案

(投票：周二上午开会,同意/反对/不表态)

所有记账者针对提案进行检验（检查交易，运行合约等），都通过的话发出同意投票，超过2/3进入下一轮

(确认：大家设定日程，并反馈参加确认消息)

所有记账者表示可以收妥提案，如果超过2/3人表示收妥，则提交存储，进入下一轮

- 实际的处理过程**非常复杂**，需要考虑：

公平高效的选出记账者列表 | 议长轮换和存活检测 | 超时进行轮次切换 | 共识时间 | 网络波动 | 广播流量 | 交易计算量 | 区块同步校验 | 过多节点不共识 | 网络规模太大 | 极端情况下的崩溃恢复

TRANSACTIONS
交易服务

交易有效性

- 如何确认一笔交易的有效性?
  - 所有权确认（签名）
  - 具有可动用的资金
  - 其他交易不会用到同一笔资金

所有权-确认

Alice 要给 Bob 5个币，他可以用别人的5个币吗?

- 每一笔交易包含所有者的公钥（确定归属）
- 如果要花该交易收入的币，你必须提供相应的私钥（签名）
- 因此，如果你能拿到别人的私钥 ……

UTXO交易模型

- ~~所有权确认~~
- **具有可动用的资金**
- **其他交易不会用到同一笔资金**

- 由于缺乏中心化的机构管理，与传统银行中使用账户结余不同，比特币使用了 **Unspent Transaction Output (UTXO)** 来确保
  - 同一笔资金只出现在一笔交易中

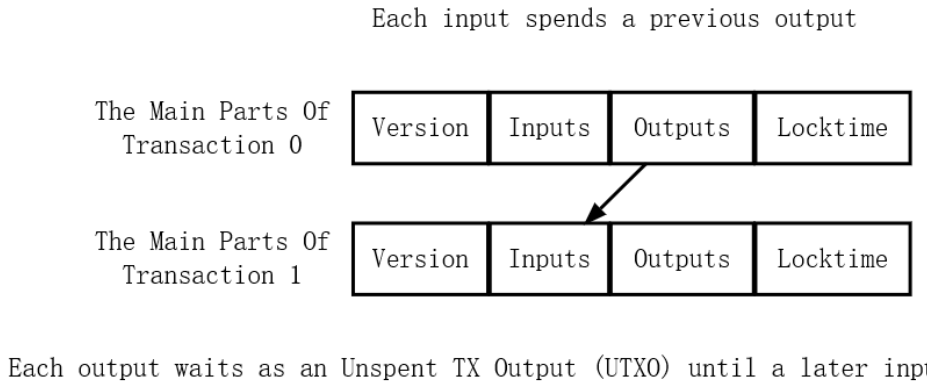
- **Unspent Transaction Output (UTXO)：未花费的交易输出**

UTXO交易模型

- **传统**：每笔收到的资金存储在保险箱里，每次交易多少就从里面取多少

VS

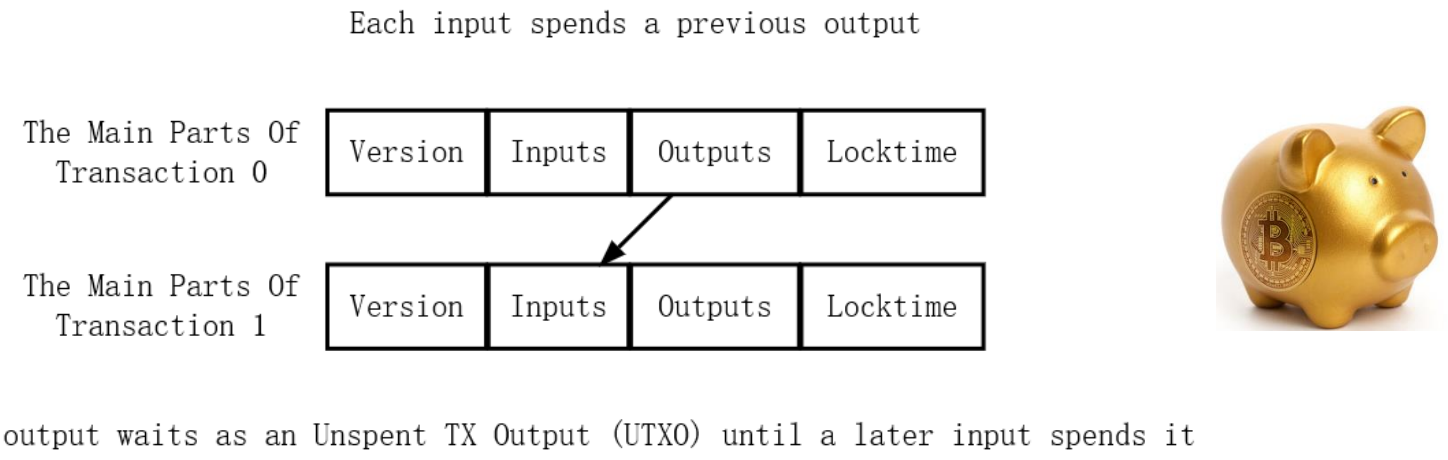
- **UTXO**：每一笔收到的资金都存在一个储蓄罐里，每次交易都需要打碎一个或多个储蓄罐



❖**每笔新的交易都需要打破旧的 “UTXO”以生成新的“UTXO”**

注：比特币交易是可分的，最小的单位叫做 satoshi，等于0.0000001BTC

UTXO交易模型



- **UTXO交易模型基本形式：**

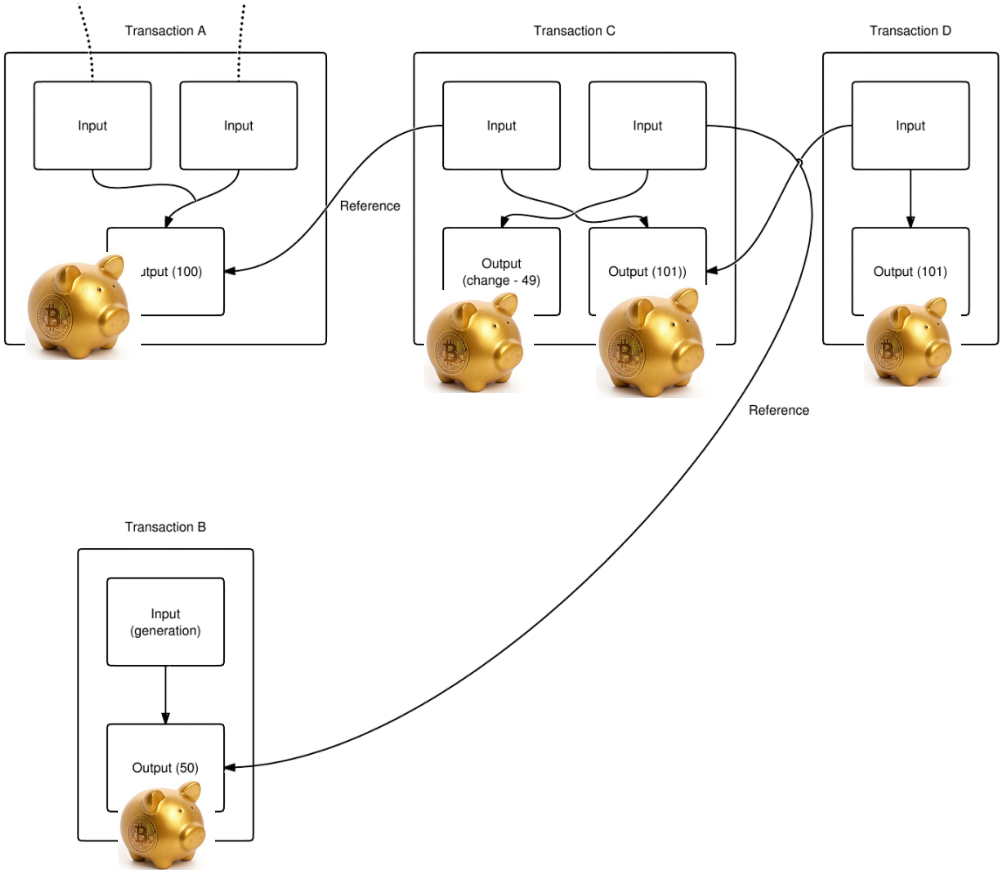
  版本号+Input: 未花费交易UTXO + Output:支付地址及数量+锁定时间

- **安全性：**

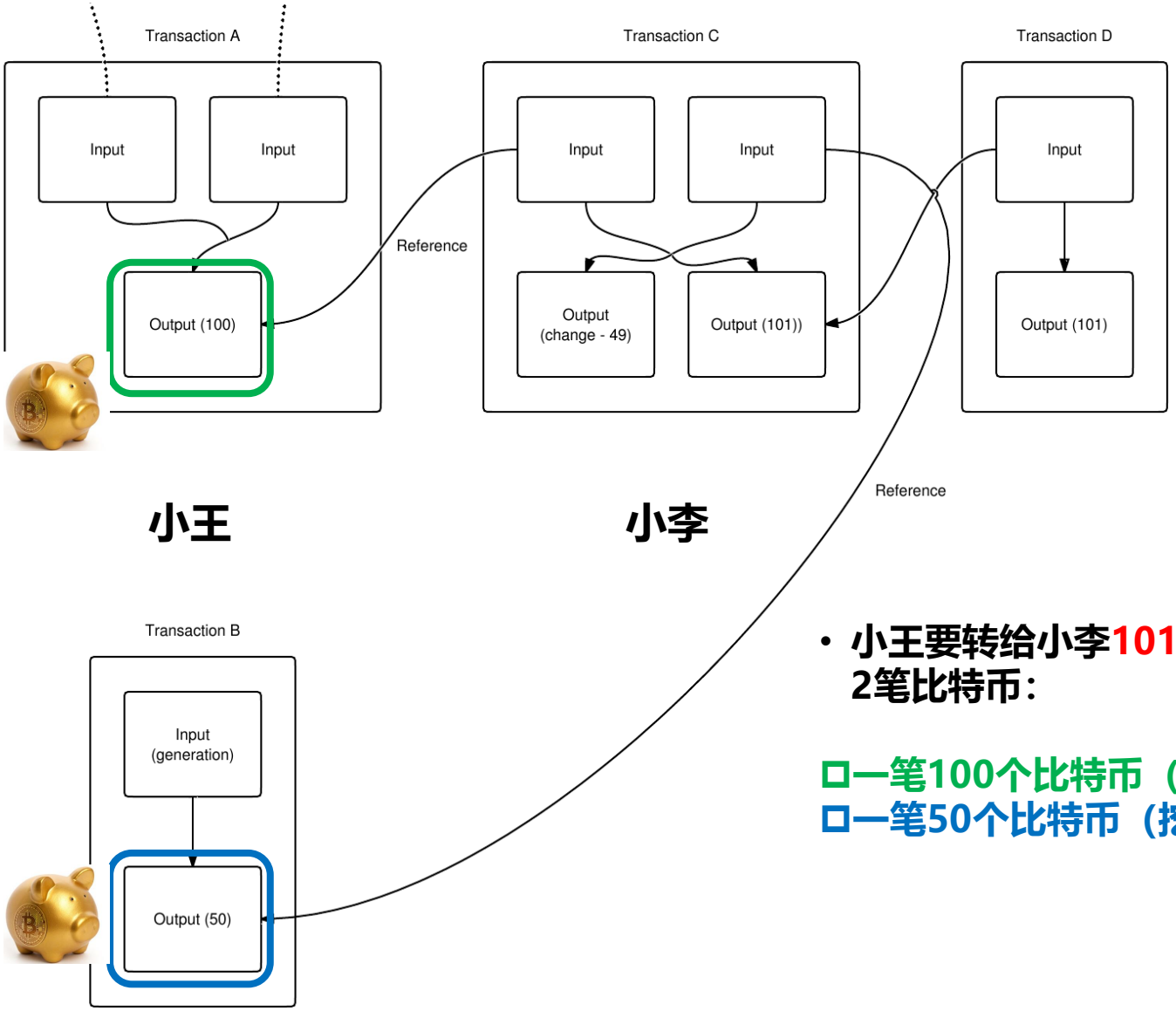
  结合所有权确认，每个交易都包含着比特币拥有者的签名，所有交易记录可溯源，记录着每个比特币从“出生”开始的所有“主人”

UTXO交易模型

UTXO：从比特币自身角度上，每次交易都是在某块“**比特币地皮**”的“**所有权证**”上改改名字而已



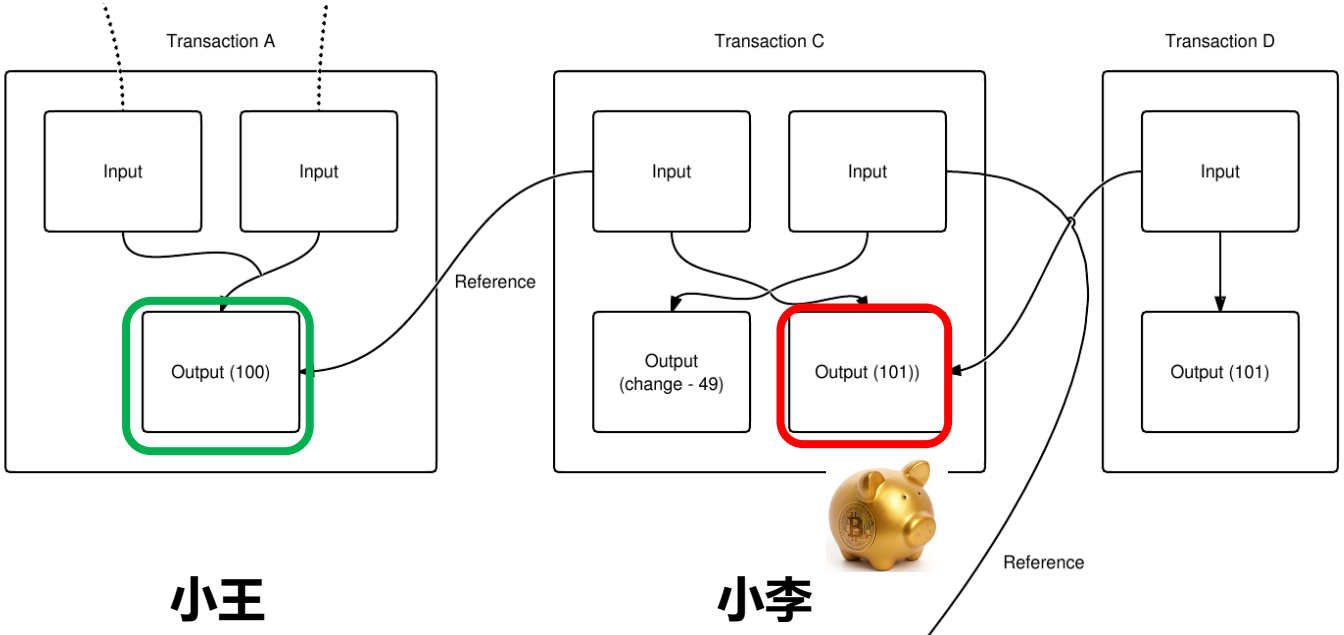
UTXO交易实例



- 小王要转给小李101个比特币，手头有2笔比特币：

  口一笔100个比特币（老王给的）
  口一笔50个比特币（挖矿）

UTXO交易实例

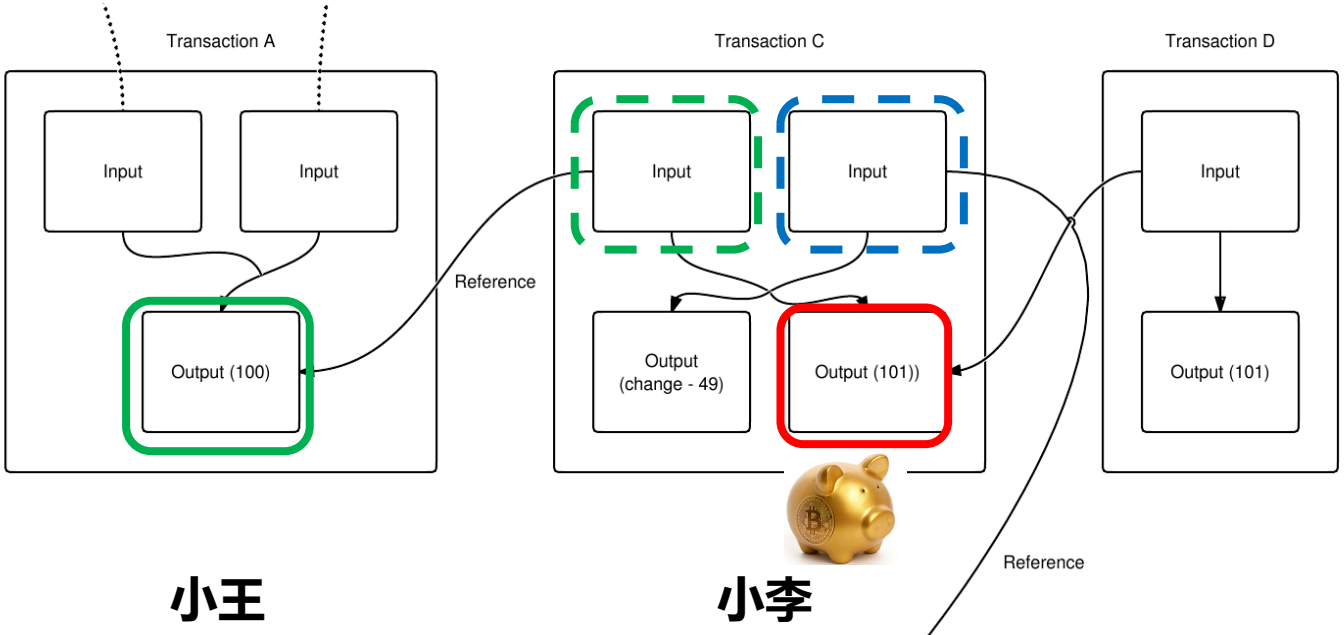


- 小王要转给小李101个比特币，手头有2笔比特币：

口一笔100个比特币（老王给的）
口一笔50个比特币（挖矿）

- 101=

UTXO交易实例

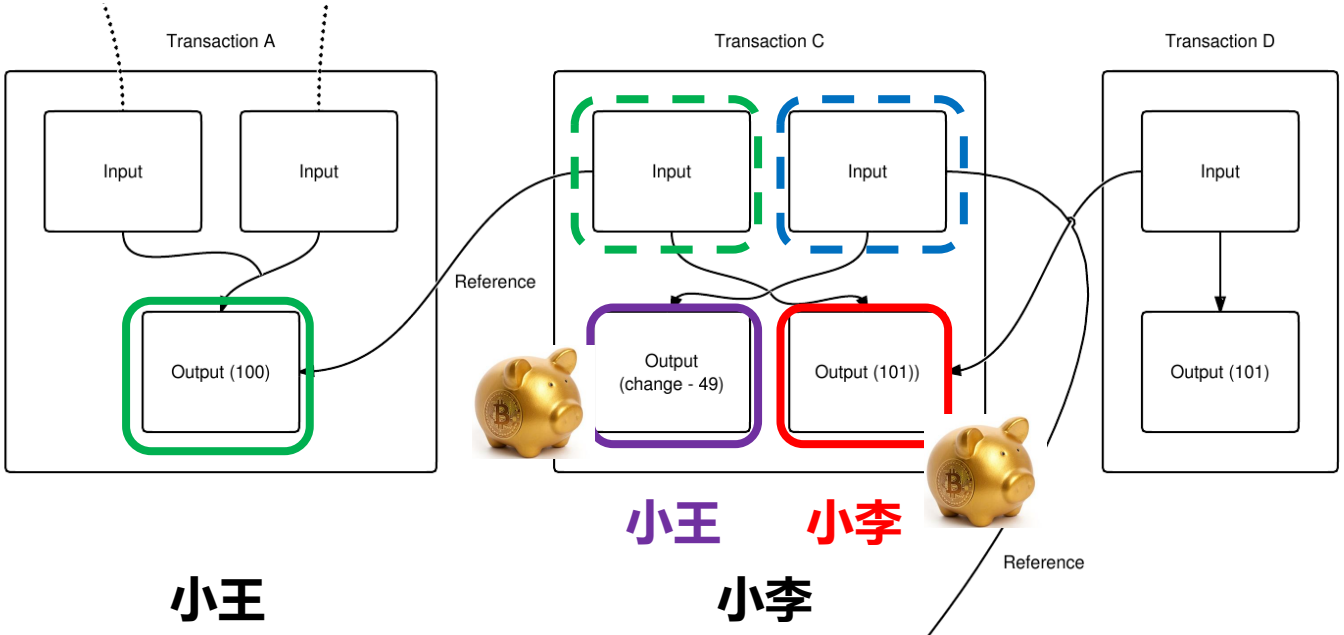


- 小王要转给小李101个比特币，手头有2笔比特币：

口一笔100个比特币（老王给的）
口一笔50个比特币（挖矿）

- 101=支付（100+50）

UTXO交易实例

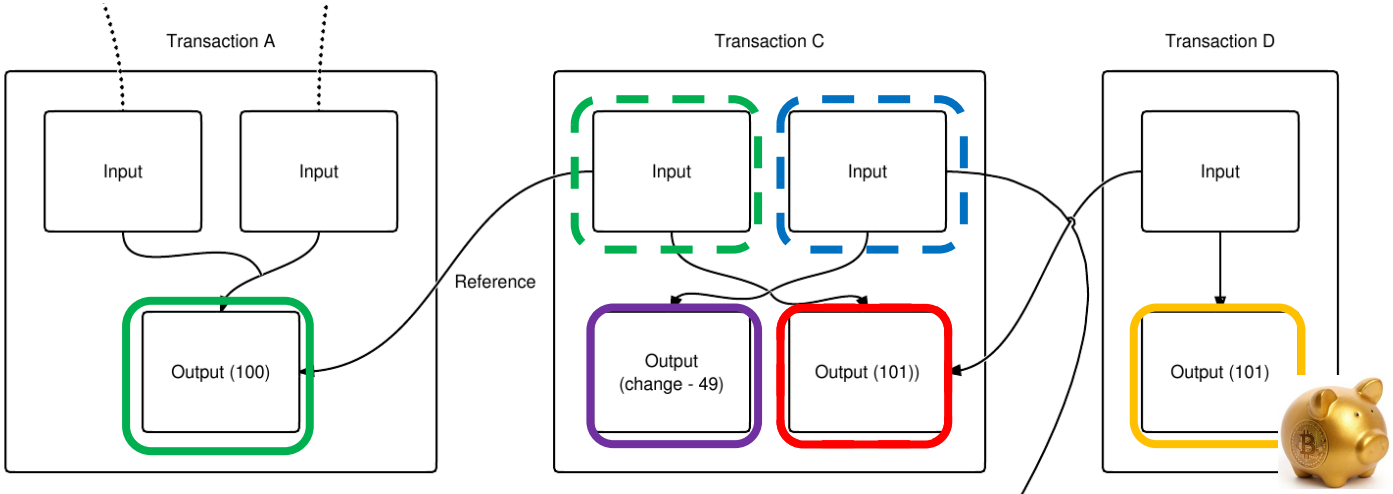


- 小王要转给小李101个比特币，手头有2笔比特币：

口一笔100个比特币（老王给的）
口一笔50个比特币（挖矿）

- 101=支付（100+50）- 找零（49）

UTXO交易实例



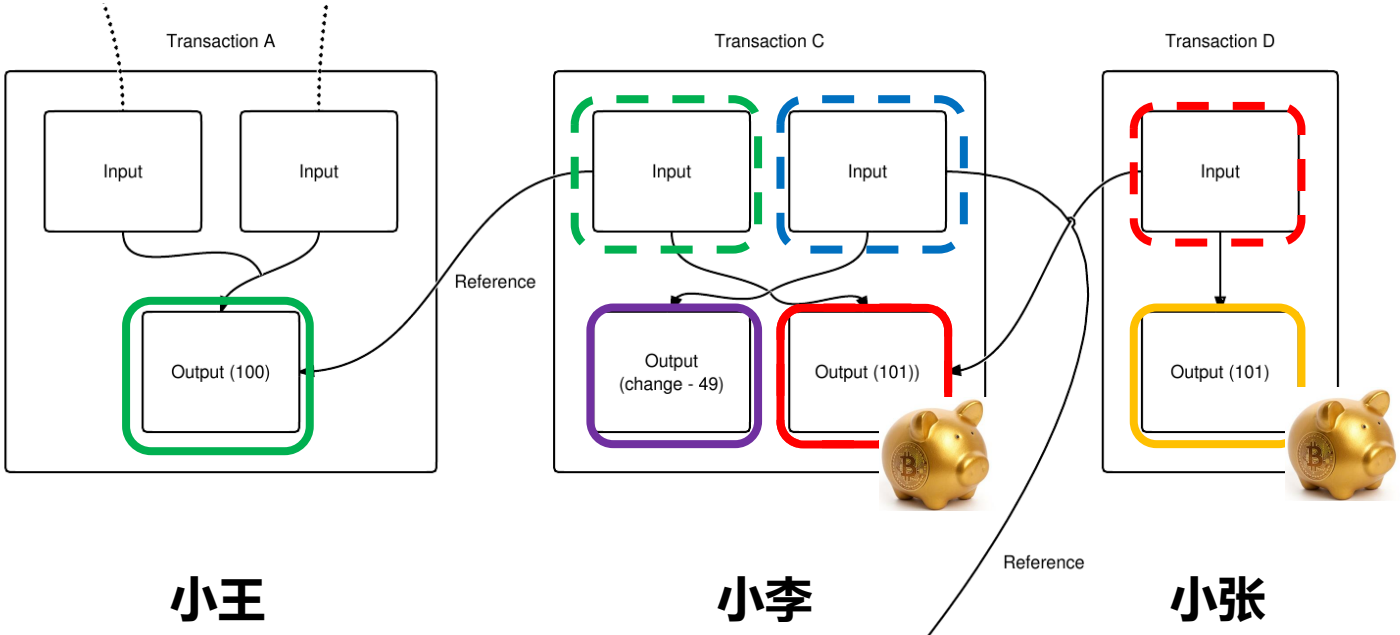
- 小王要转给小李101个比特币，手头有2笔比特币：

口一笔100个比特币（老王给的）
口一笔50个比特币（挖矿）

- 101=支付（100+50）- 找零（49）

- 小李就可以继续支付给小张101个比特币

UTXO交易实例



- 小王要转给小李101个比特币，手头有2笔比特币：

  口一笔100个比特币（老王给的）
  口一笔50个比特币（挖矿）

- 101=支付（100+50）- 找零（49）

- 小李就可以继续支付给小张101个比特币

UTXO模型信息

Transaction View information about a bitcoin transaction

75eb73a617e1aaef1a187d9e19bdaee7c38476e1399e8cab8f95e9e5e83bd4b7

1KoFB5SQq3kPZ8z6KFt97aTEQW1fjs75SS 148mARsoudWiUdJi9sDoUa9tRtfogMGNdz	→	1NEFQJ6rQNnRaUtQwaxu8dV14McHMzFfC5 13MbEETZszRA2tF5s6MGxVtK4P5vw7X8W8	0.065773 BTC 0.00010771 BTC

0.06588071 BTC

Summary	
Size	436 (bytes)
Weight	1744
Received Time	2015-03-24 00:08:11
Included In Blocks	348915 (2015-03-24 00:12:26 + 4 minutes)
Confirmations	176377 Confirmations
Visualize	View Tree Chart

Inputs and Outputs	
Total Input	0.06598071 BTC
Total Output	0.06588071 BTC
Fees	0.0001 BTC
Fee per byte	22.936 sat/B
Fee per weight unit	5.734 sat/WU
Estimated BTC Transacted	0.065773 BTC
Scripts	Show scripts & coinbase

来源：blockchain.info

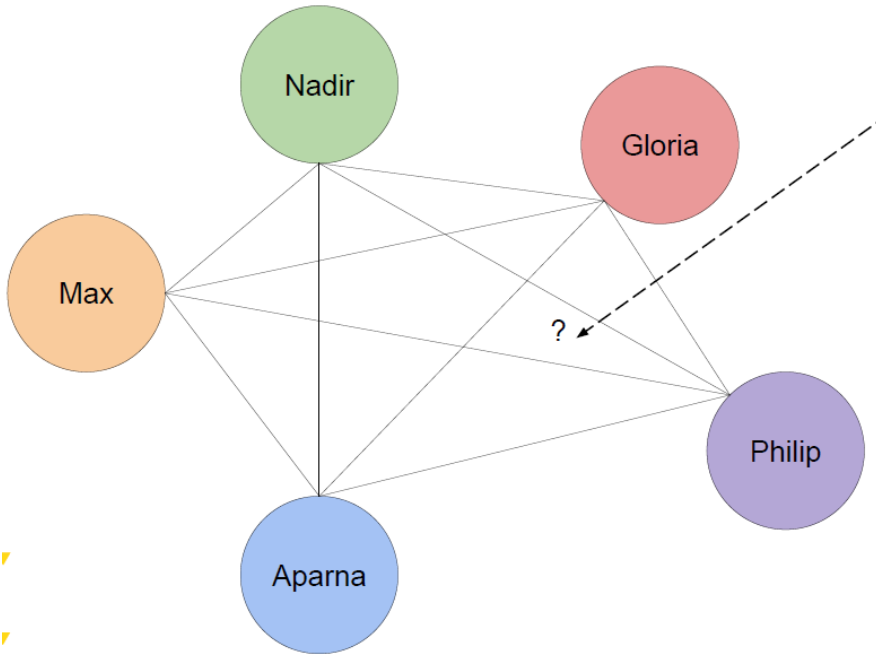
比特币共识机制

- ~~比特币身份确认~~ 身份确认
- ~~UTXO交易模型~~ 交易服务
- 链式交易信息记录 记录管理
- 工作量证明(POW) 信任规则

RECORD
记录管理

记录：分布式数据库

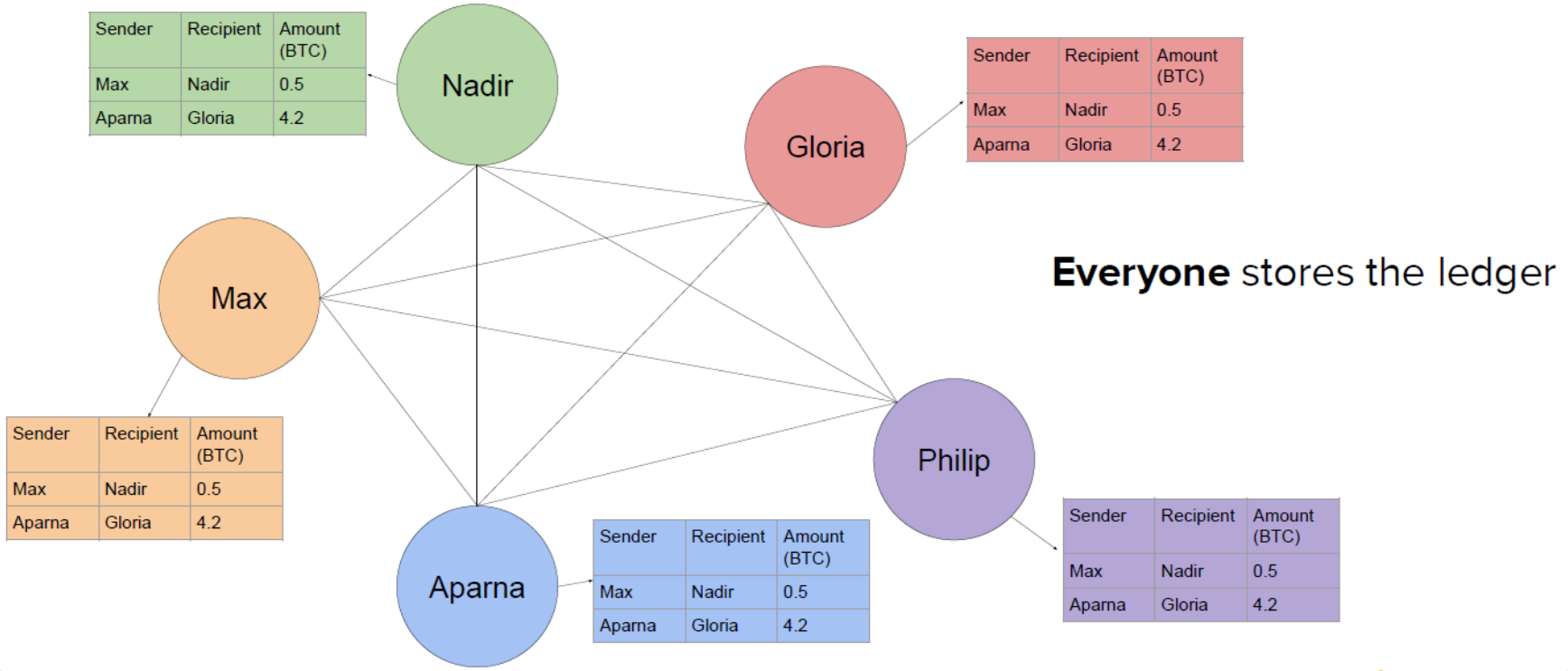
- 在身份确认，交易有效后，如何存储交易记录?
- 如何保障交易账本的可溯性?



Sender	Recipient	Amount (BTC)
Max	Nadir	0.5
Aparna	Gloria	4.2

记录：人人即银行

- 每个比特币网络上的节点都存储一个完整的账本



记录格式：区块链

- 每一段时间生成一个区块，区块里记录着这个时间段内的交易记录，区块与区块间按时间安全紧密地串联在一起，就成了区块链

Sender	Recipient	Amount (BTC)
Max	Nadir	0.5
Aparna	Gloria	4.2
Philip	Gloria	23
Max	Philip	3.2
Nadir	Aparna	0.3
Gloria	Philip	17

Batching

Sender	Recipient	Amount (BTC)
Max	Nadir	0.5
Aparna	Gloria	4.2

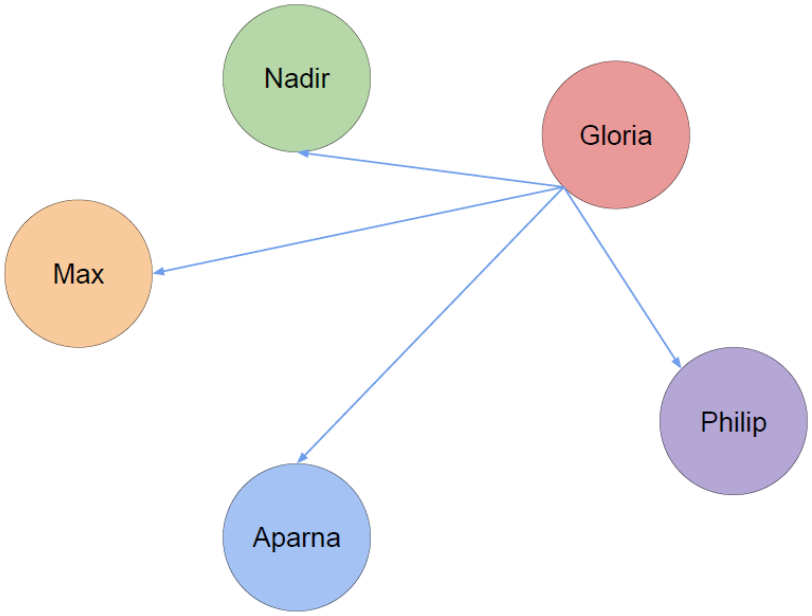
Sender	Recipient	Amount (BTC)
Philip	Gloria	23
Max	Philip	3.2

Sender	Recipient	Amount (BTC)
Nadir	Aparna	0.3
Gloria	Philip	17

AGREEMENT（CONSENSUS）
共识机制

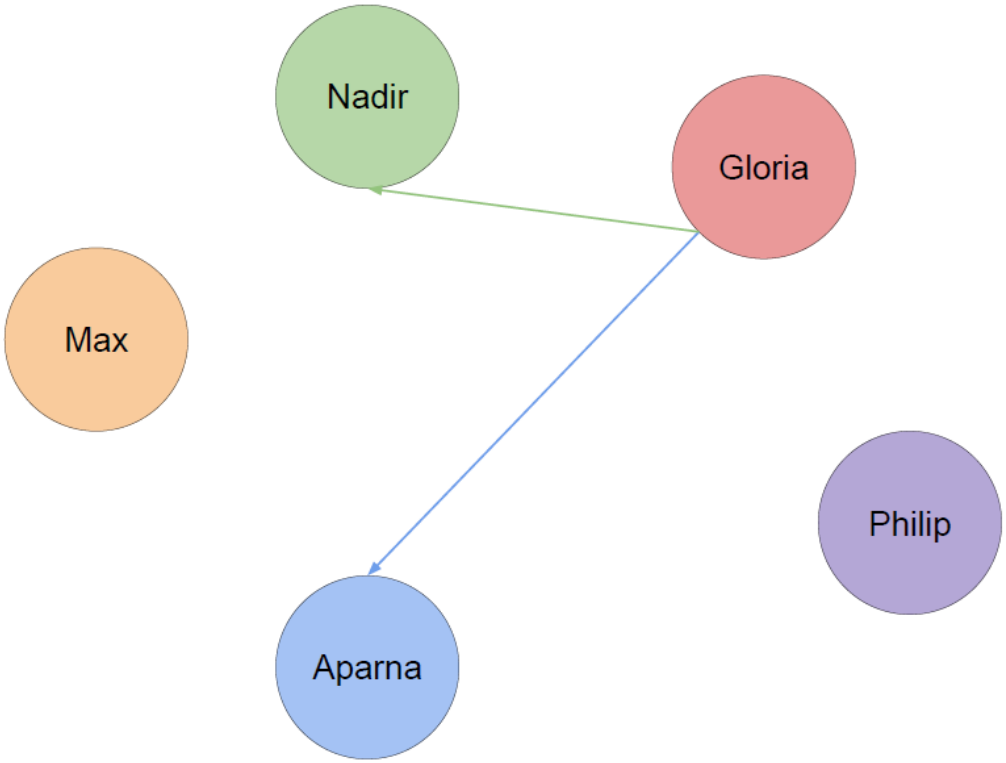
共识机制

- 在一般的公有区块链上，如何达成共识?
  - 如果每个区块链上的用户只要接收到有效记录后，都记录到自己账本，会发生什么



共识机制中常见的攻击

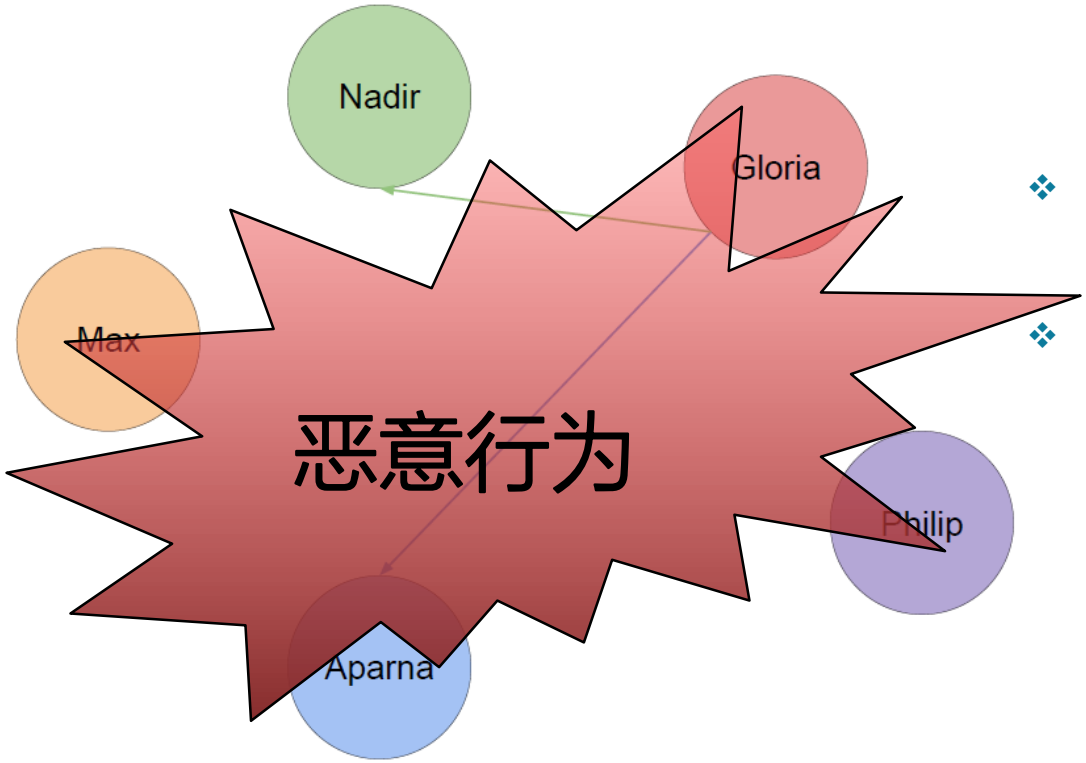
- **Double Spend Attack 双花攻击**



- Gloria答应给Aparna 10个比特币，同时答应给Nadir 10个比特币.....但是她总共只有10个比特币
- Gloria的行为就是**双花攻击**
- **将导致每个节点记录的账本信息不一致**

共识机制常见的攻击

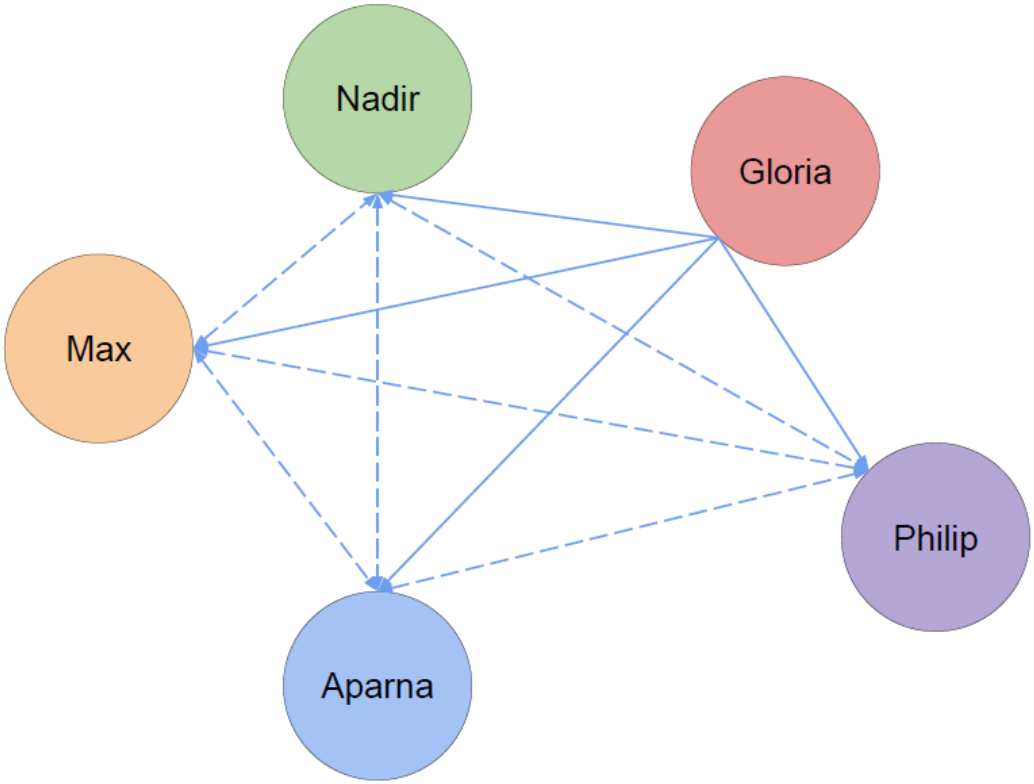
- **Double Spend Attack 双花攻击**



- Gloria答应给Aparna 10个比特币，同时答应给Nadir 10个比特币.....但是她总共只有10个比特币
- Gloria的行为就是双花攻击
- **如何保证这种 独立的 记录过程不存在这种恶意行为?**

共识机制常见的攻击

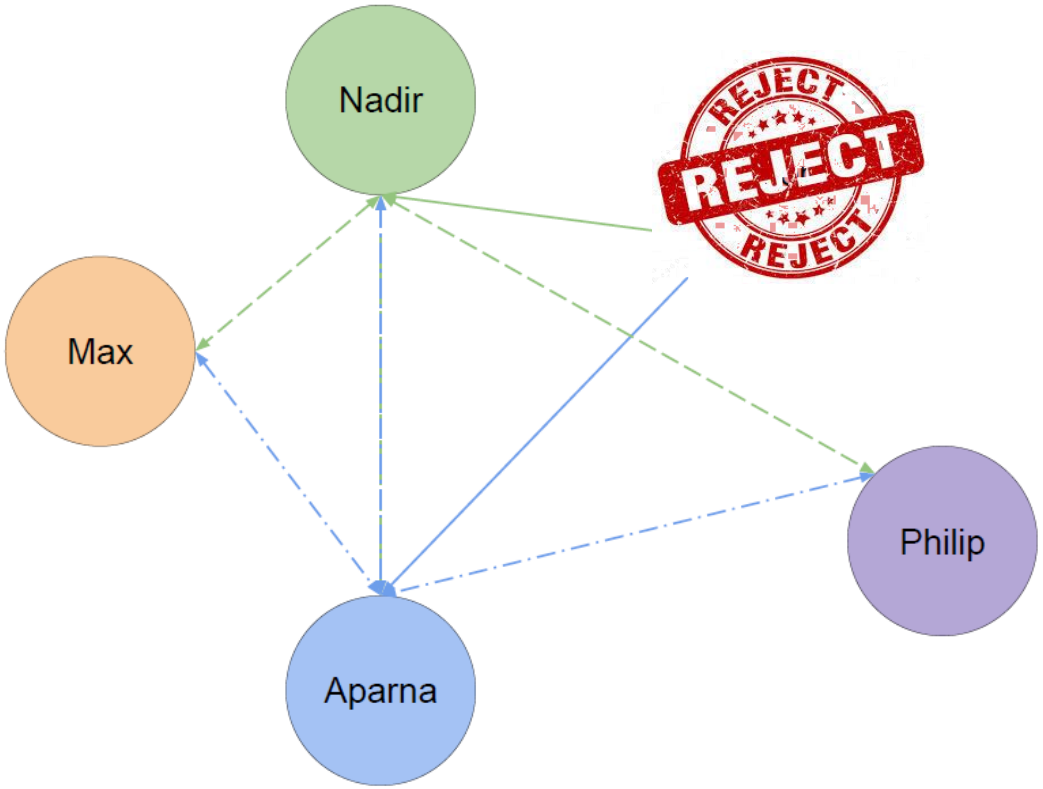
❖ 同等验证



- ❖ **取代每个用户孤立的决定**
  - 提交者向其他用户提交一条交易信息
  - 其他用户进行投票
  - 当获得一定数目投票后，大家同意将交易信息进行保存
- ❖ **保证所有节点存储相同的交易账单**

共识机制常见的攻击

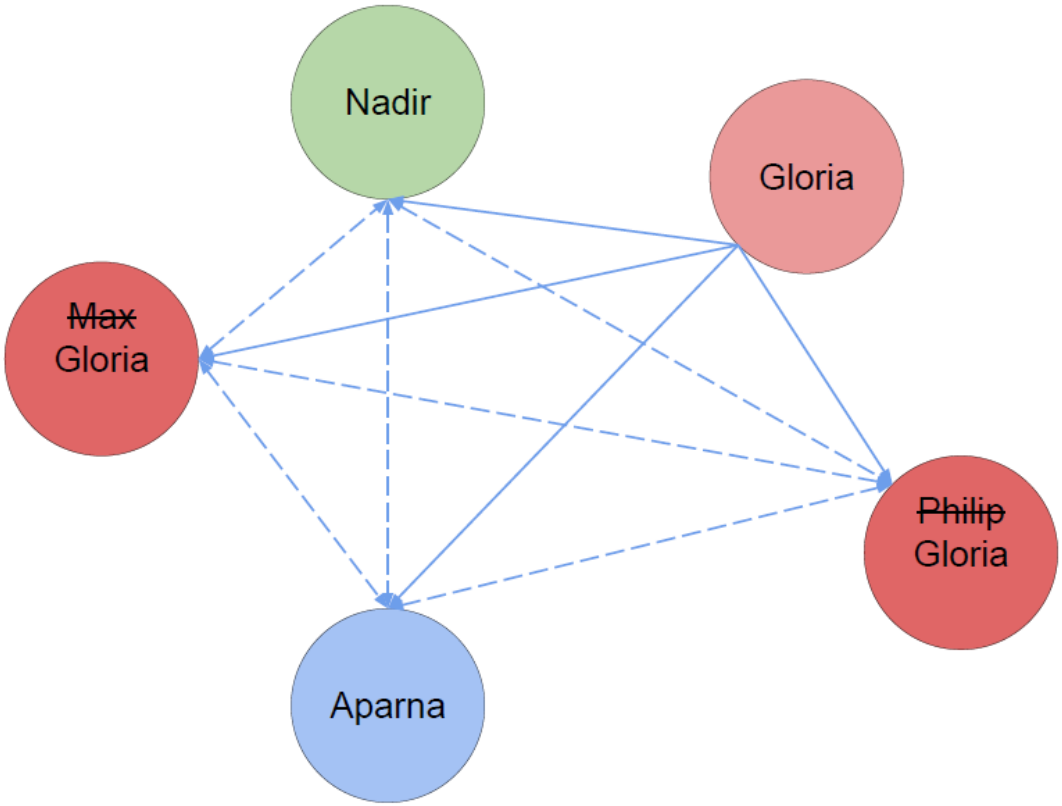
❖ 拒绝双花



- 现在，当Gloria准备进行双花时，会被其他用户拒绝接受该笔交易。
- 节点在观测到多笔交易使用同一笔比特币时，会对Gloria的提案投出反对票。

共识机制常见的攻击

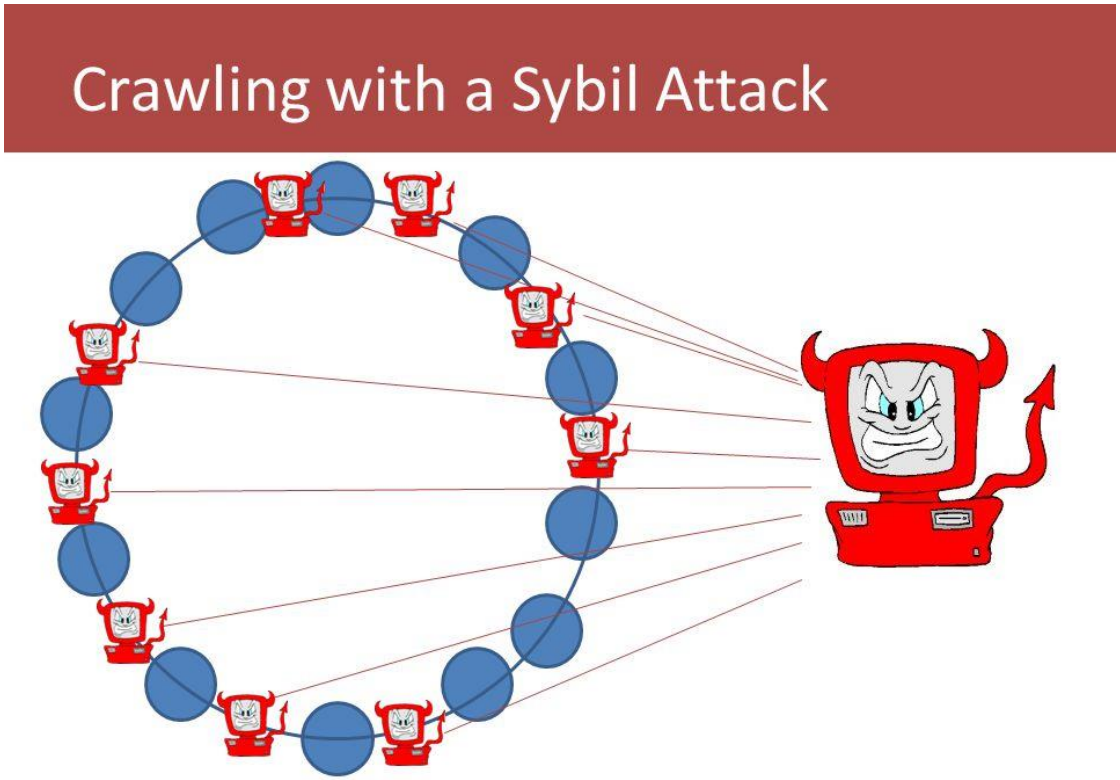
Sybil Attack 多重身份攻击/女巫攻击



- 比特币作为无中心登记的匿名服务
  - 创建多重身份代价极低
  - 多重身份意味着多重的投票权利

共识机制常见的攻击

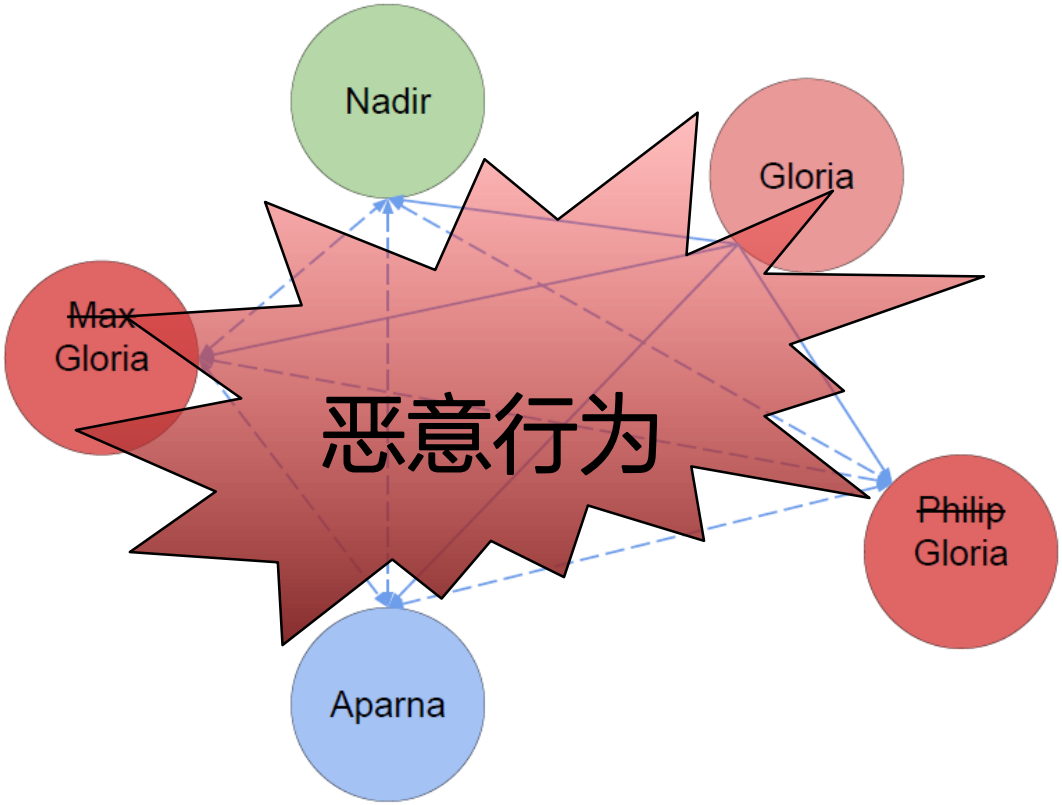
- **Sybil Attack 多重身份攻击/女巫攻击**



- 比特币作为无中心登记的匿名服务
- 创建多重身份代价极低
- 多重身份意味着多重的投票权利

共识机制常见的攻击

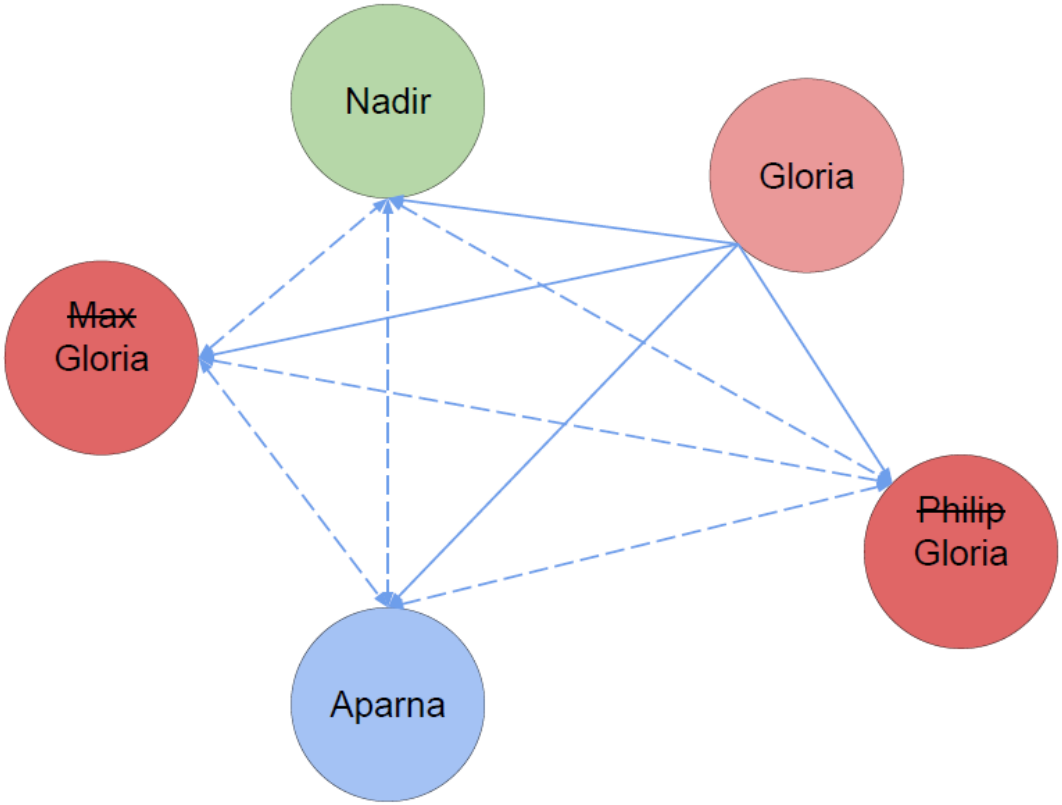
❖ Sybil Attack 多重身份攻击



- ❖ 比特币作为无中心登记的匿名服务
  - 创建多重身份代价极低
  - 多重身份意味着多重的投票权利
- ❖ **Gloria可以实行多重身份攻击，从而允许她的双花行为**

共识机制常见的攻击

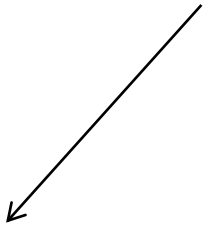
❖ Sybil Attack 多重身份攻击

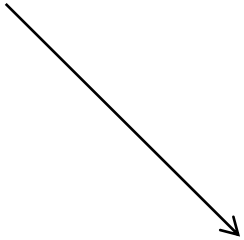


- ❖ 原因：**投票几乎没有成本**！
- ❖ 取代用身份投票的机制，我们采用**资源成本**进行投票
- ❖ **提高作恶的代价！**

工作量证明

Proof-of-Work





证明

资源消耗

POW应用

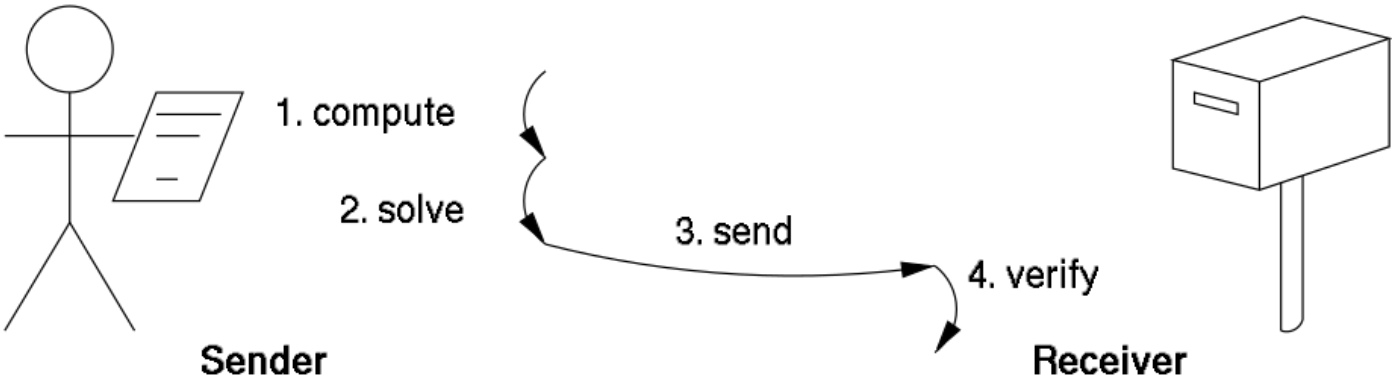
工作量证明（POW），此一概念最早由 Cynthia Dwork 和 Moni Naor 于 1992 年的学术论文提出，而工作量证明一词则是在 1999 年由 Markus Jakobsson 与 Ari Juels 所发表

POW 最早被用于阻止拒绝服务攻击（DDOS）、反垃圾邮件等一些服务滥用的经济对策

POW应用

第一个POW 应用是1996年 Adam Back 开发的 "Hashcash" 应用，它采用工作量证明共识机制来过滤垃圾邮件，微软也将其应用在 Hotmail, Exchange, Outlook 等电邮服务上。

具体做法是要求所有收到的邮件都使用强 PoW 附件（邮票），比如Receiver向所有想发送电子邮件的 Sender 都分发一个"标准质询"。



此系统使得垃圾邮件发送者在大量发送邮件时在经济成本上不可行。

POW in Life

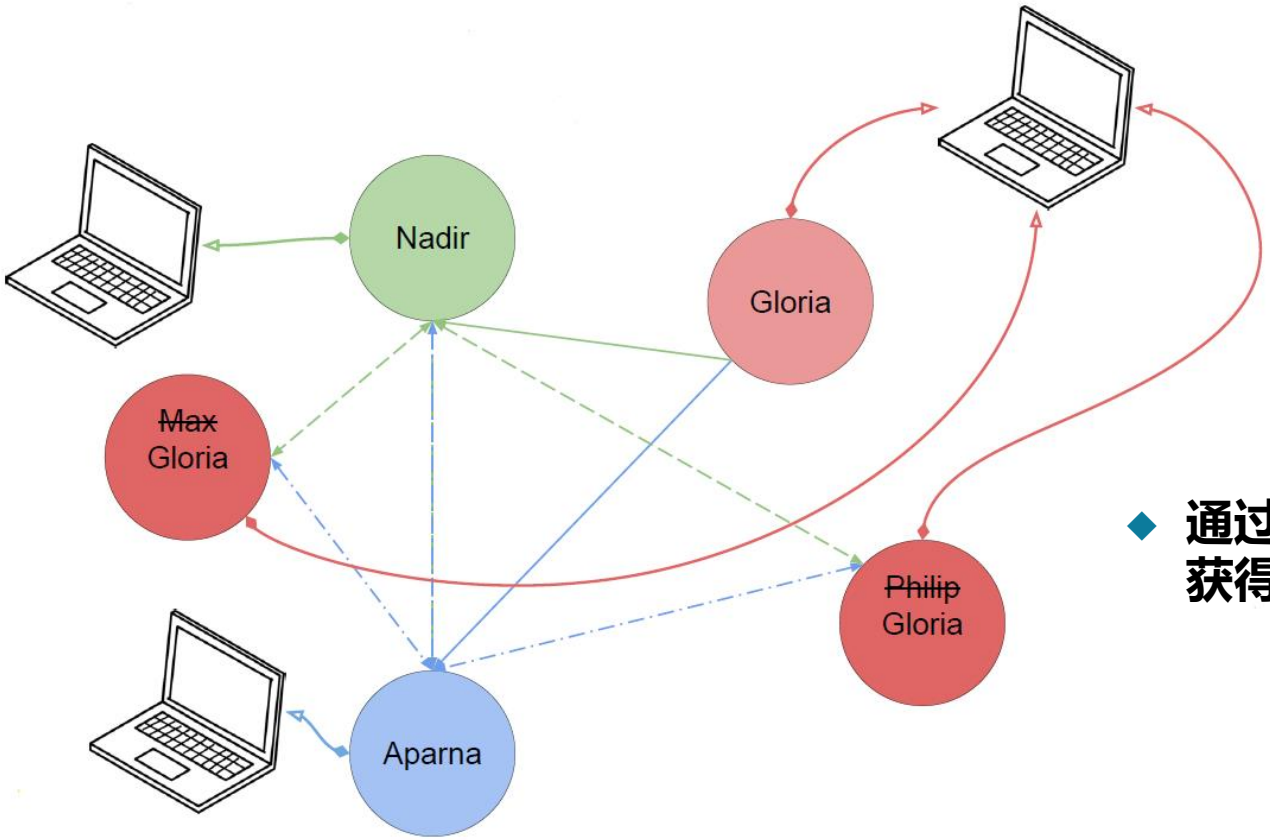
- 你想到一家公司去工作，这家公司会让你先实习一段时间，公司会考量你的实习质量来决定是否录用你，这段实习的时间就是你的工作量证明。

- 某男女谈恋爱，某女为了考验其的忠心，便让某男去给她买套房，某男便用光老爸老妈所有的钱买了套房，某女可以轻易的检验某男是否购房，这个行为也可以看做是工作量证明。

✓ **不容易完成**
✓ **容易验证**

工作量证明

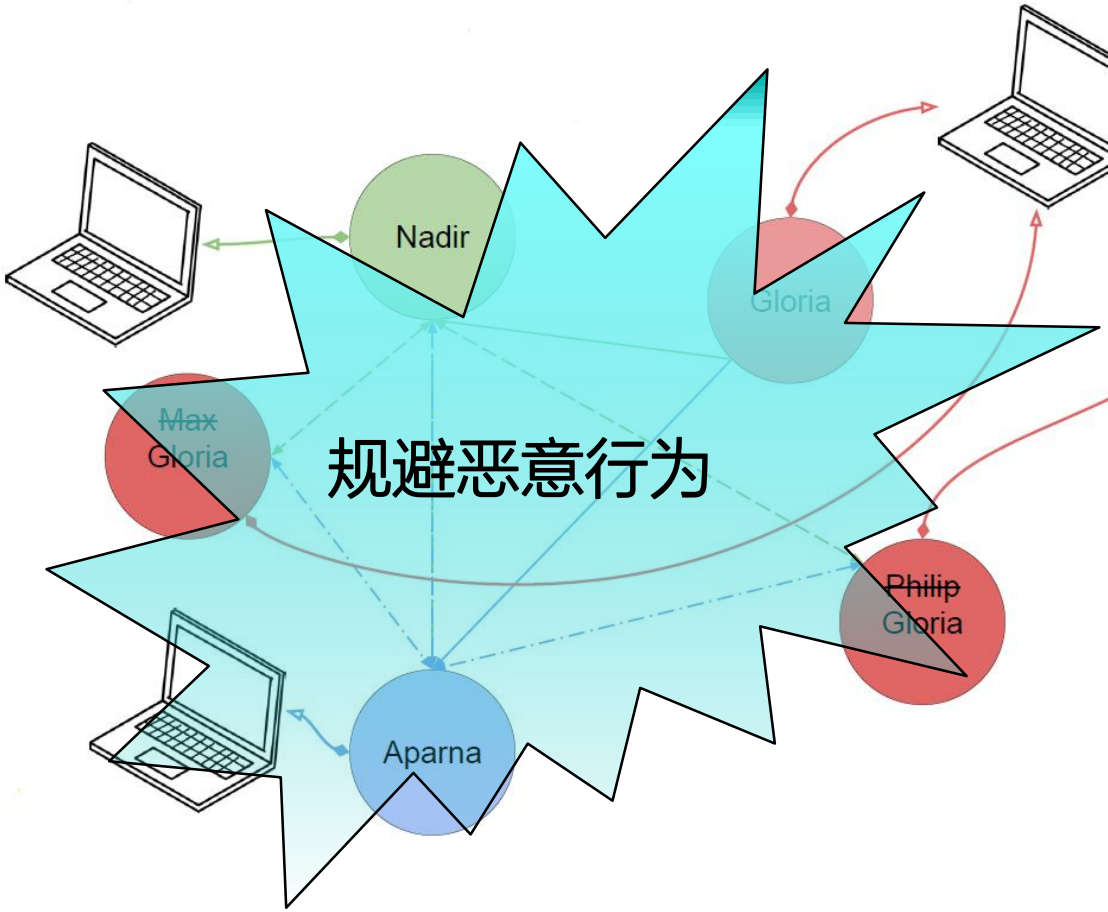
- Proof-of-Work 工作量证明



- **通过消耗资源解决一个问题获得投票资格（即记账权）**

工作量证明

- Proof-of-Work 工作量证明



- **记账权必须通过花费计算资源来获得**，比如说通过蛮力解决一个问题；**通过记账奖励鼓励投入资源**
- 即使Gloria有多个身份，也只对应到单个计算资源，从而保证记账的公正性
- 更准确来说像买彩票。投入的越多（花费越多），中奖几率越大（成功记账）

“工作量”争夺记账权

- **整个争夺记账的过程就是挖矿的过程，也就是比特币发行的过程**

- **“挖矿”争夺记账权奖励**
  - 记账有利润：**比特币奖励** + **交易手续费**
  - 很多人争夺记账权
  - 通过付出计算量解决一个难题，谁先解决谁获得记账权
  - 坏人作恶的成本变高

POW 挖矿

- **挖矿：是参与维护比特币网络的节点，通过协助生成新区块来获取一定量新增的比特币的过程**

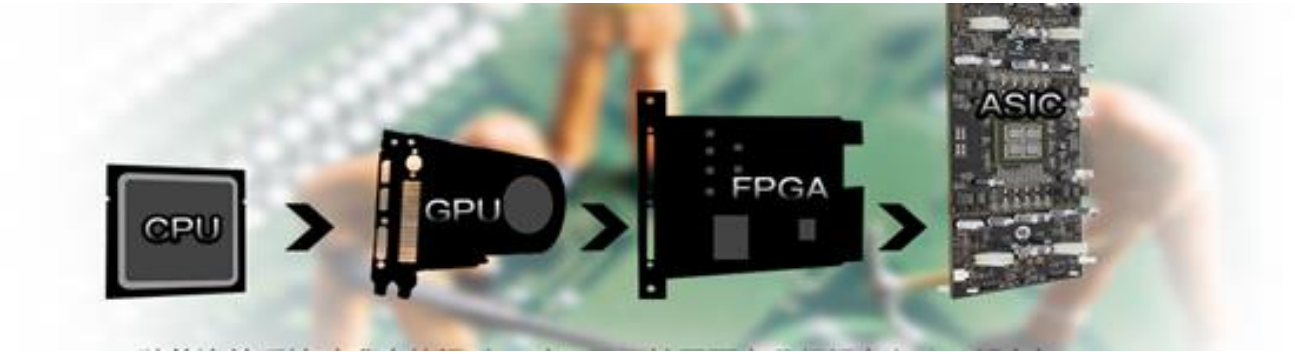
- 每 **10 分钟**左右生成一个**不超过 1 MB** 大小的区块（记录了这 10 分钟内发生的验证过的交易内容），串联到区块链尾部，每个区块的成功提交者可以得到系统 **6.25 个比特币**的奖励（该奖励作为区块内的第一个交易，一定区块数后才能使用），以及用户附加到交易上的支付服务费用

- **注**：每个区块的奖励最初是 **50 个**比特币，每隔 **21 万个区块自动减半**，即 4 年时间，最终在**2140年**比特币总量稳定在 **2100 万个**。因此，比特币是一种通缩的货币

POW 算力

由于 Hash 难题在目前计算模型下需要大量的计算，这就保证在一段时间内，系统中只能出现少数合法提案。反过来，能够提出合法提案，也证明提案者确实已经付出了一定的工作量。这也保障了，如果有人尝试恶意破坏，需要付出大量的经济成本

❖普通的 CPU（2009 年）、到后来的 GPU（2010 年） 和 FPGA（2011 年末）、到后来的ASIC 矿机（2013 年年初，目前单片算力已达每秒数百亿次 Hash 计算）、再到现在众多矿机联合组成矿池（知名矿池包括 F2Pool、BitFury、BTCC 等）



❖截止1/5/2018, 全网的算力已超过**每秒2.6*10^18 次** Hash 计算，超过**世界500强超级计算机算力总和**的**100倍!**

总结：无中心网络需要何种工作量证明?

- 难题设计必须满足如下条件：
  - **不容易完成**（表明需要工作量）
  - **容易验证**（其他节点可以快速确认确实付出了工作量）
  - **工作过程公平**（任何节点没有完成工作的捷径）
  - **具有随机性**（能力越强，只能保证率先完成概率越大）

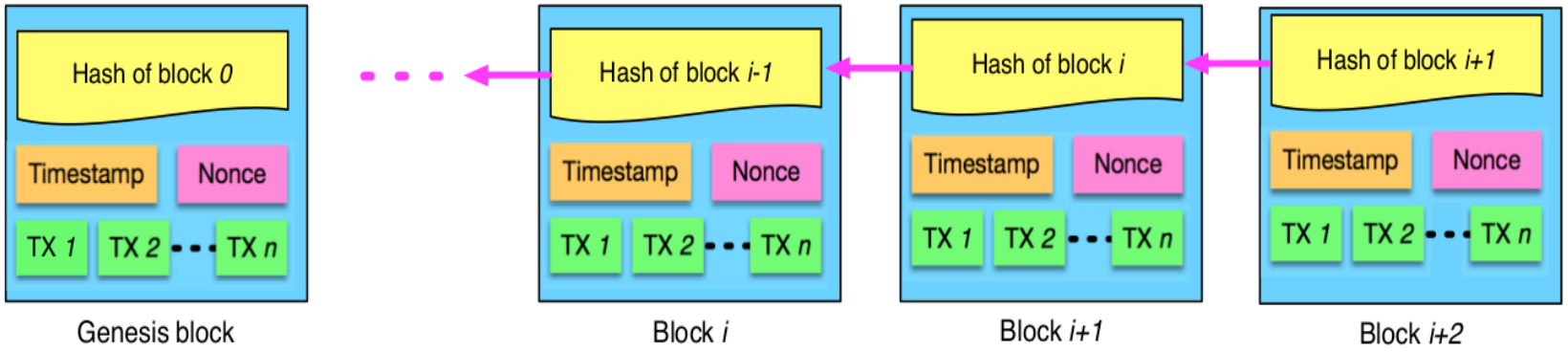
之后的部分用来延伸学习、复习

POW for Blockchain
区块链中的工作量证明

Blockchain

❖ 区块链定义

按照时间顺序将数据区块以顺序相连的方式组合成的一种**链式数据结构**

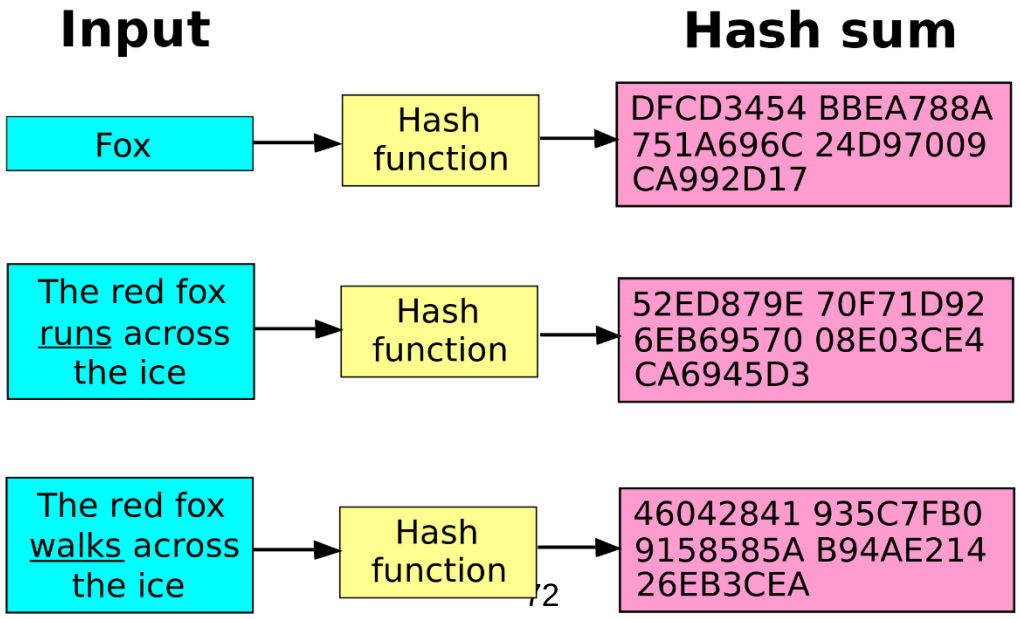


- 区块链是一个**分布式的无法篡改的账本数据库**
- 分布式：每个节点均保存独立的一套账本
- 无法篡改：每一个区块都指向前一个区块 **“牵一发动全身”**
- 无法篡改性由**Hash（哈希）算法**的**单向性**保证

Hash function

- **Hash函数**
  - 信息技术领域**非常基础也重要的技术**
  - 能将**任意长度的信息**（明文）转换成**固定长度**的**无序字符串**（Hash值）

- 例如计算一段话“**hello blockchain world, this is chuan@sysu**”的MD5 hash 值为 **c36c8db7751921054b374a21a22015a2**



Hash function

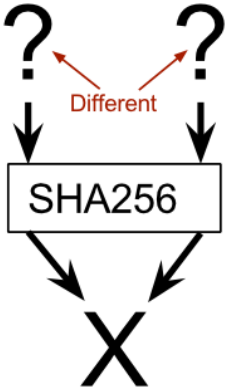
❖Hash算法具有以下特点：

- 正向快速：给定明文和 hash 算法，在有限时间和有限资源内能计算出 hash 值
- 输入敏感：原始输入信息修改一点信息，产生的 hash 值看起来应该都有很大不同
- 逆向困难：给定特定hash 值，在有限时间内很难（基本不可能）逆推出明文
- 冲突避免：很难找到两段内容不同的明文，使得它们的 hash 值一致（发生冲突）

❖区块链中一般采用比MD5安全性更好的**SHA2系列Hash算法**。

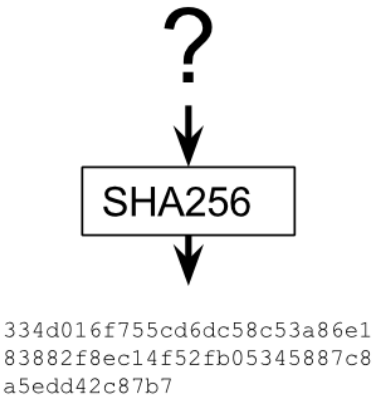
- SHA2系列生成的hash值更长，SHA256，SHA512.....数值代表hash长度

$$hash = (\sum_{n} a[n]s^n) \mod 2^{31}$$

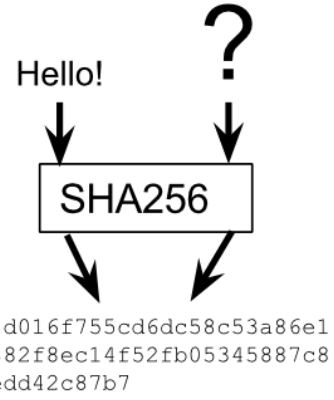


Fast calculation

Collision resistance



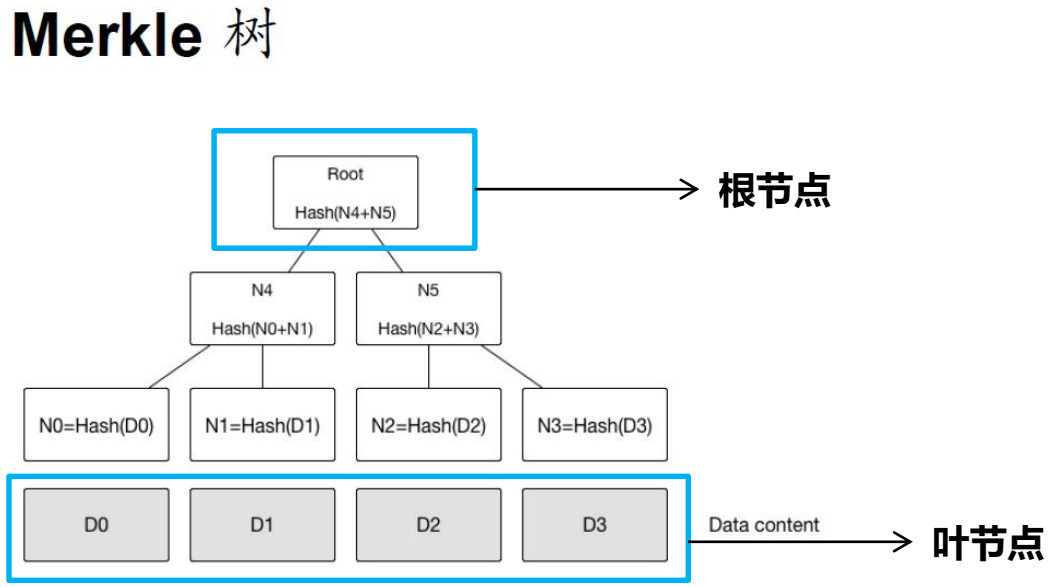
Preimage resistance



Second-preimage resistance

Merkle Tree

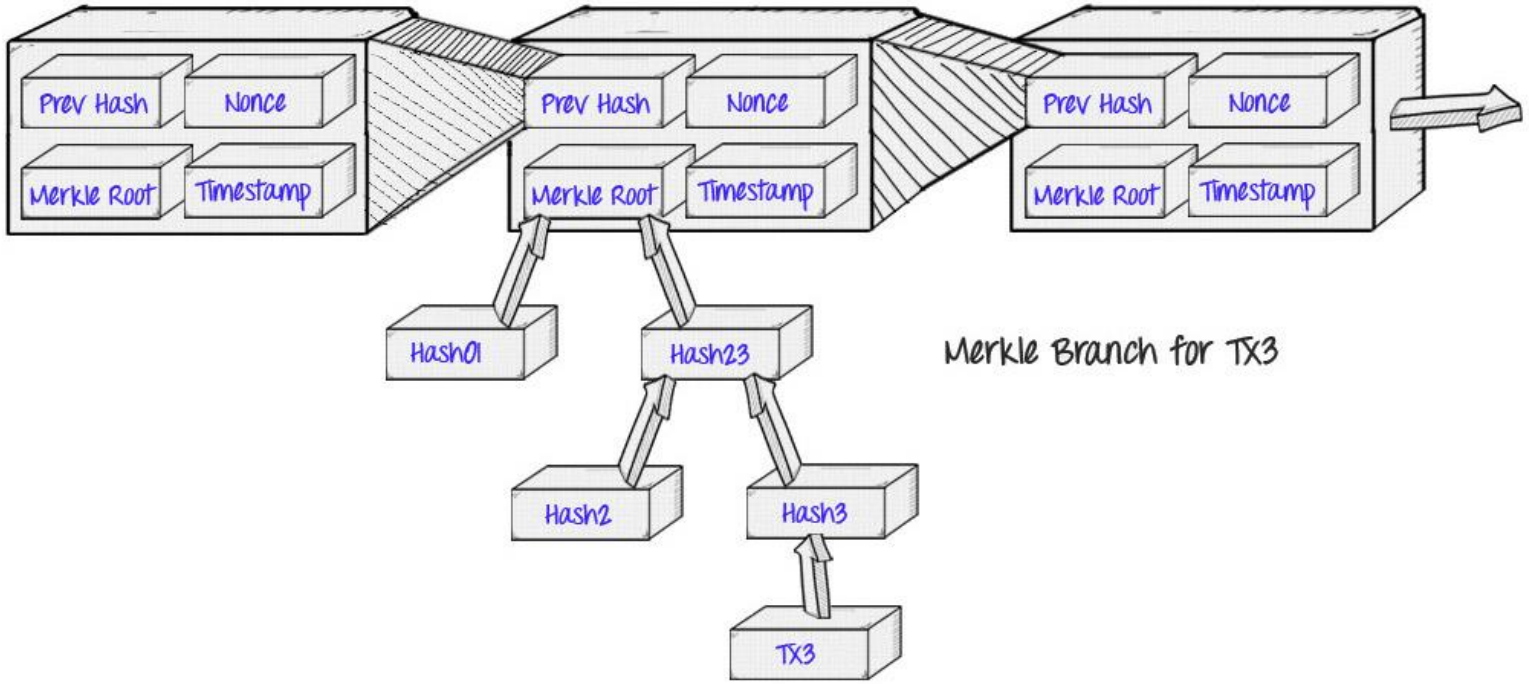
- **Merkle树，又名哈希树**
- 是一种二叉树，由一个根节点、一组中间节点和一组叶节点组成。最下面的叶节点包含存储数据或其哈希值，每个中间节点是它的两个子节点内容的哈希值，根节点也是由它的两个子节点内容的哈希值组成



- 底层数据的任何变动，都会传递到其父节点，一直到根节点

Merkle Tree

- Merkle Tree大多用来进行比对以及验证处理，在处理比对或验证的应用场景中时，特别是在分布式环境下进行比对或验证时，Merkle Tree会大大减少数据的传输量以及计算的复杂度。



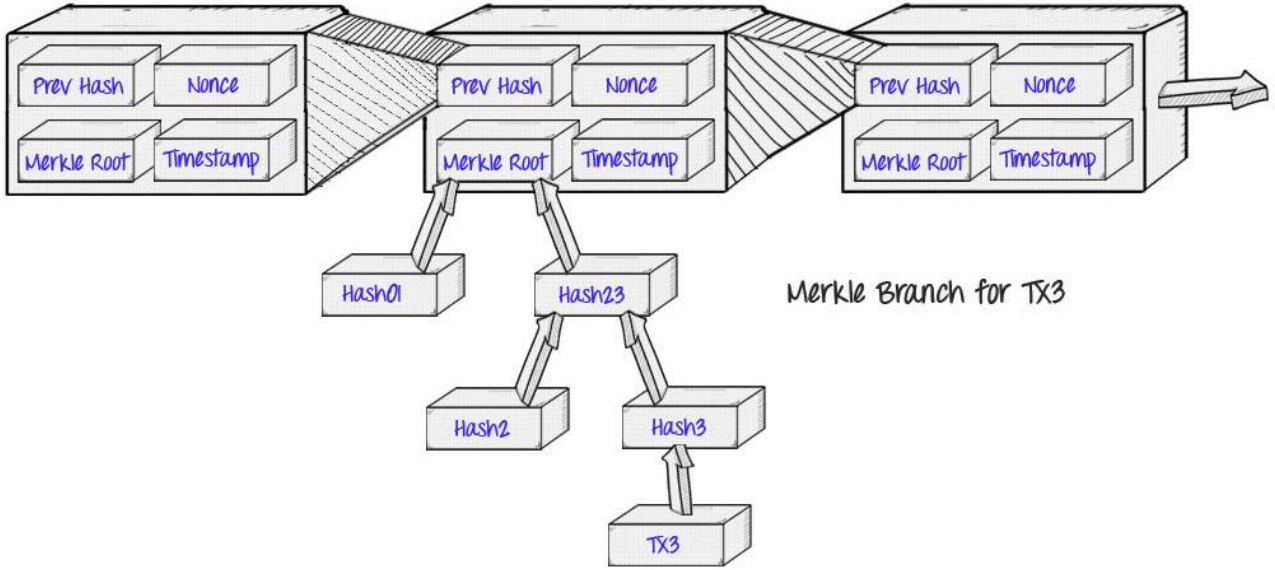
比特币共识 —— 挖矿

Mining is a process by which new blocks are added to the blockchain

挖矿原理

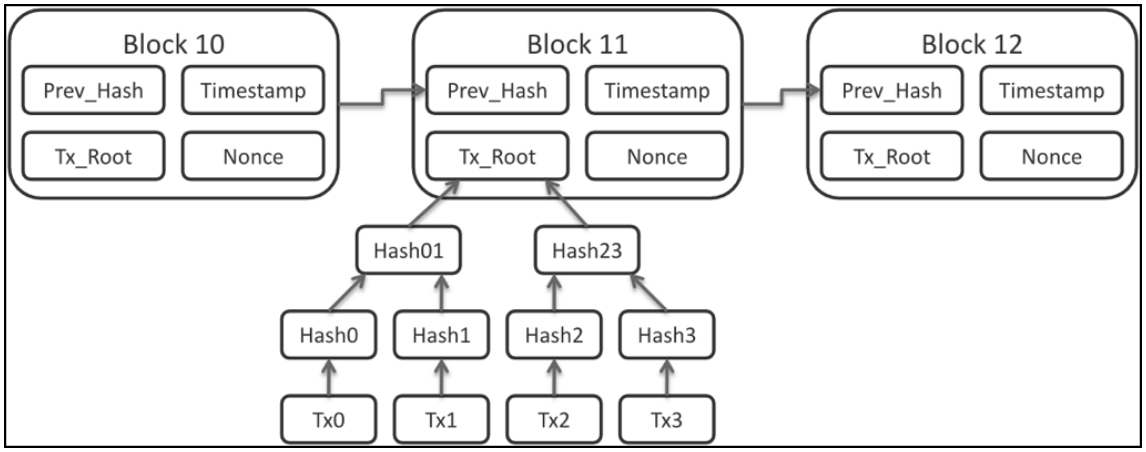
具体挖矿过程：参与者综合**上一个区块的 Hash 值**，上一个区块生成之后的**新的验证过的交易内容的Markle Root值**，加上**猜测的一个随机数Nonce**，再加上**时间**，一起打包到一个候选新区块，让新区块的 Hash 值小于比特币网络中给定的一个数。这是一道面向全体矿工的“计算题”，这个数越小，计算出来就越难。（基于Hash函数特性）

SHA(**Markle Root**+**上一个区块Hash值**+**时间**+**Nonce**)<某个数：**挖矿成功！**



挖矿原理

- 具体挖矿过程：参与者综合**上一个区块的 Hash 值**，上一个区块生成之后的**新的验证过的交易内容的Markle Root值**，再加上**猜测的一个随机数Nonce**，再加上**时间**，一起打包到一个候选新区块，让新区块的 Hash值小于比特币网络中给定的一个数。这是一道面向全体矿工的“计算题”，这个数越小，计算出来就越难。（基于Hash函数特性）
- SHA(**Markle Root**+**上一个区块Hash值**+**时间**+**Nonce**)<某个数：**挖矿成功**
- “**0000000000000f15673f1354**” “**0**”的数目体现了挖矿难度



- 系统每隔2016 个区块，会根据上一周期的挖矿时间来调整挖矿难度（通过调整限制数“**0**”数目），来调节生成区块的时间稳定在 10 分钟左右。

比特币的货币发行方式—挖矿（mining）

挖矿即哈希

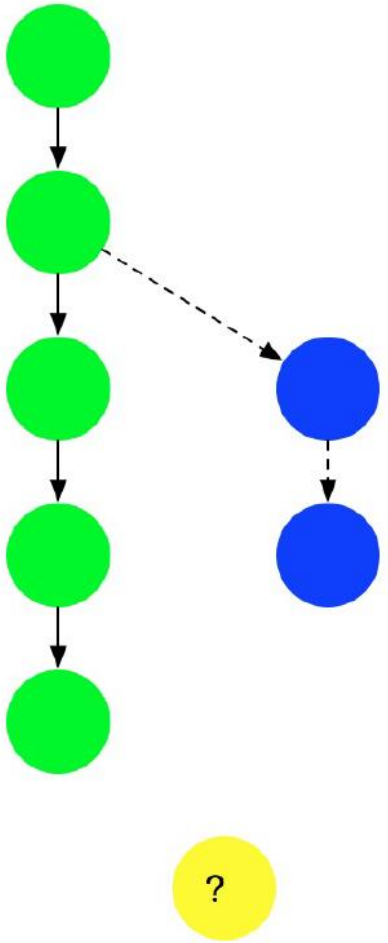
随机公平性：输入相同，则输出相同，输入稍有变化，输出变化巨大，且事先无法预知（输出为均匀分布）

单向不可逆性：给定输入，容易计算输出；给定输出，无法破解输入

POW 最长链机制

- 上述过程还存在什么问题?
- 加入全网同时有两个合法提案会在网络中进行广播，收到的用户进行验证后，会基于用户认为的最长链基础上继续难题的计算。因此，系统中可能出现链的分叉（Fork）
- **解决方案：比特币网络最长链机制**

- 假定超市只有一个出口，付款时需要排成一队，可能有人不守规矩要插队。超市管理员会检查队伍，认为最长的一条队伍是合法的，并让不合法的分叉队伍重新排队。新到来的人只要足够理智，就会自觉选择最长的队伍进行排队



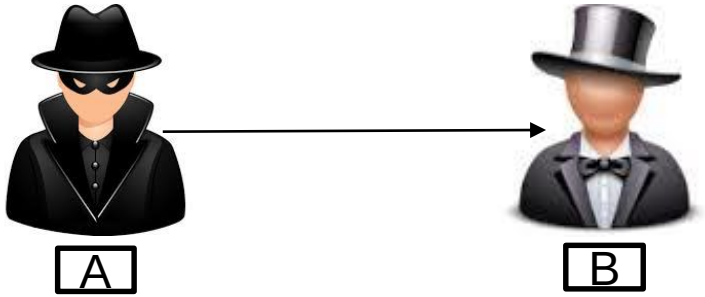
最长链机制-51%攻击

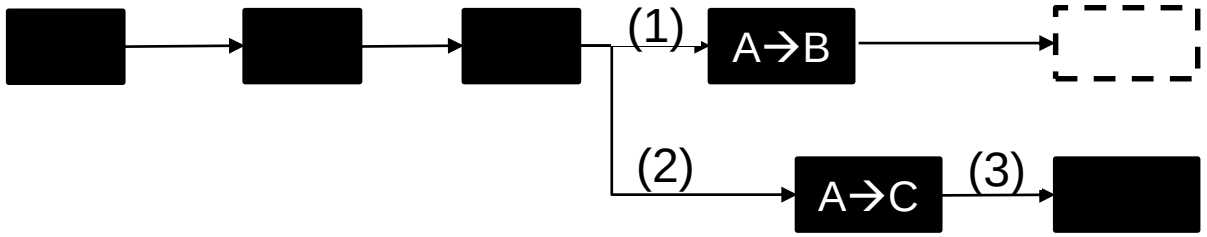
回顾双花支付：一笔钱花两次

(1) A给B支付比特币，交易在一个块中确认

(2) A重新构造一笔交易A→C，并打包进区块公布（分叉）

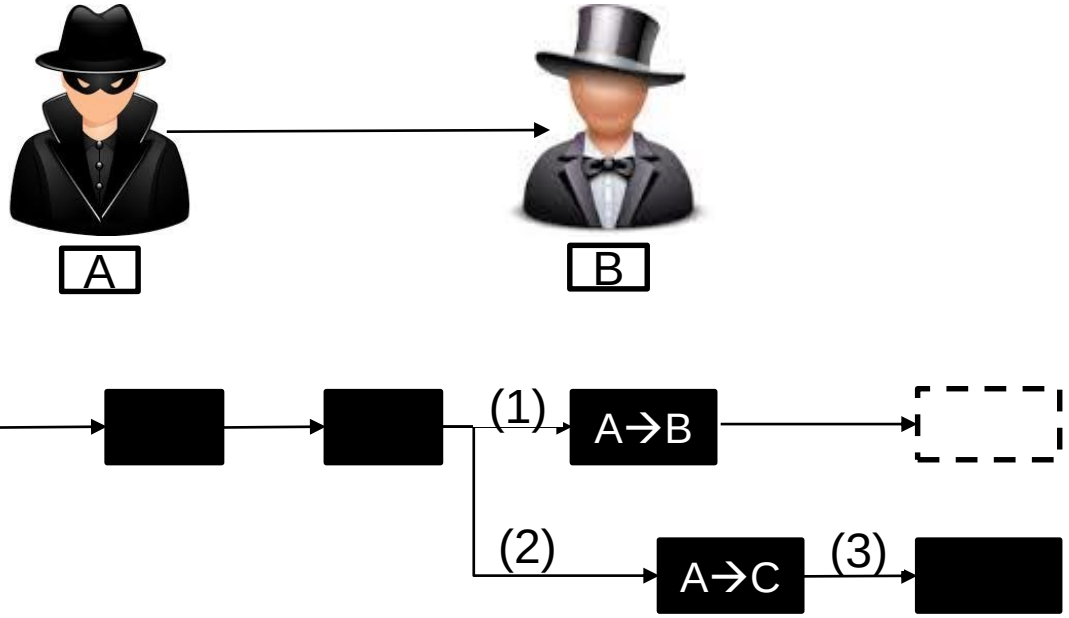
(3) 包含双重支付的块率先找到下一个块，全网认可A→C,交易A→B无效





双重支付-51%攻击

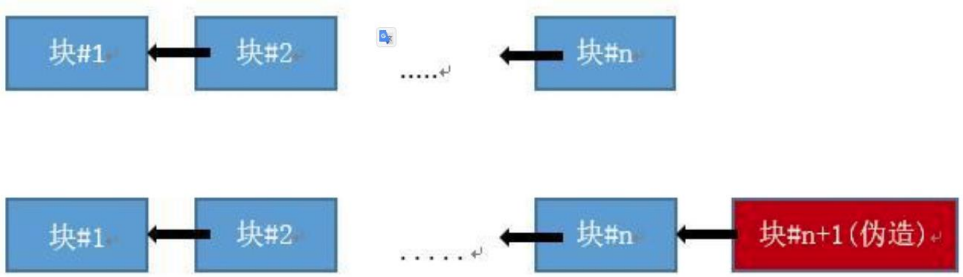
- 双花能够成功，需要攻击者拥有超过一半以上的算力
- 比特币网络的安全，要求一半以上的节点（算力）是可信的



现实中的51%攻击

- **在上面的讨论中，比特币区块的构建和算力的多少紧密相关，因此控制了算力就控制了区块链的生成**

- **设想这样一个现实场景：**
  - Alice 和 Bob之间使用比特币完成了一杯咖啡的交易，Bob在收到Alice的转账通知 (交易提交)，就给Alice提供了咖啡。
  - Alice不想支付这笔钱，在开始之前他把区块里的这笔交易改成Alice转给自己的一笔交易了(更改很容易，只要把接收地址和签名改掉即可)。
  - Alice开始尝试用这个伪区块进行计算(计算正确后这个快会被加入主块中)，因为拥有51%的算力，Alice比别的节点更容易优先计算成功，导致一个伪造的区块加入了主链。



- **一般个体达到 1/3的计算力，比特币网络就存在被破坏的风险了，也就会出现双花问题**

防范 51%攻击

如何防范51%攻击?

- 除了尽量避免计算力放到同一个组织手里，没太好的办法，这是目前 PoW 机制自身造成的

- 绝大多数的矿工，都会通过诚实挖矿来维持整个比特币系统
- 如果他们集体伪造交易，用户对比特币失去了信心，没人在去使用比特币。那么矿工伪造了交易盗取比特币就失去了意义

- 如果真有这样的51%攻击，建议是收款方等到全网的 6 个区块确认之后再交付商品。按照10分钟一个区块的速度，也只需一个小时就可以保证你的钱是否基本肯定收到。

  （6个区块后再对全网进行修改难度很高）

即便如此，POW仍是目前数学上可证的最安全的机制。

总结：比特币共识机制

- **Identity 身份确认：**

  通过共享公钥传输比特币，通过私钥声明比特币所有权

- **Transaction 交易服务：**

  在UTXO模型下，账户余额可由所有未花费交易输出总和计算得到，一笔UTXO只能被花费一次，避免双花

- **Record 记录管理:**

  比特币网络每个参与者均持有一份区块链账本，区块间基于hash安全相连

- **Consensus 共识机制：**

  通过工作量证明执行提案并确认交易信息，避免双花攻击行为发生，最终达成全网一致性

Outline: Section II

比特币共识机制

身份确认

UTXO交易模型

交易信息记录

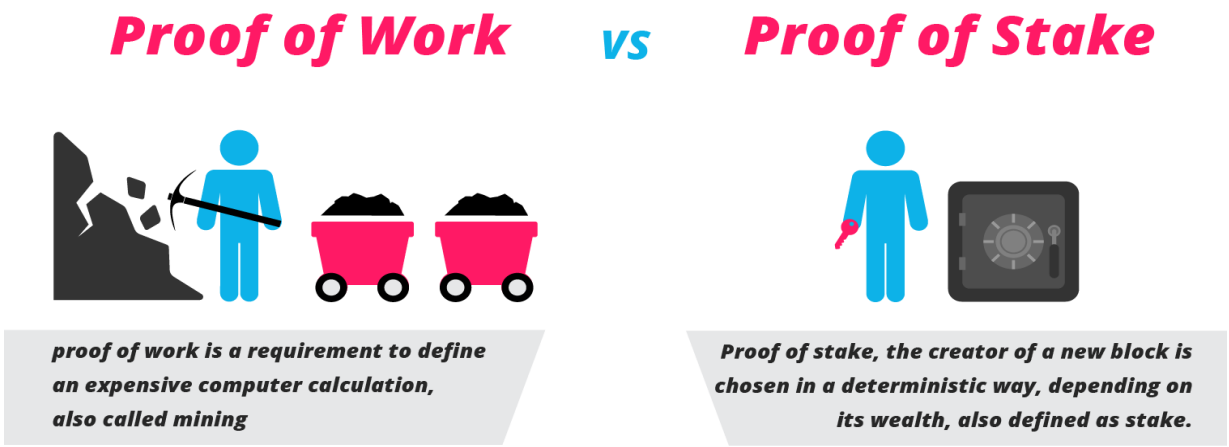
工作量证明

(Optional) 共识机制应用与局限

其他PoX机制

Proof-of-Stake 权益证明

- 最初有Sunny King 在2012年在论文PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake 中提出，这种机制通过计算你持有币数占总币数的百分比，包括你占有币数的时间来决定你获得本次记账权利的概率。**持有越多，获得记账权力概率越大**



- 代表币种：如Qtum币，以太坊也部分采用POS机制
- 优点：相比POW缩短了达成共识时间；节省能源
- 缺点：容易分叉；易中心化（马太效应）

其他PoX机制

Delegated Proof-of-Stake 股份权益证明

- 由BitShares社区首先提出了DPoS机制，与PoS基本相同，主要**区别**在于节点选举若干代理人，由代理人验证和记账。

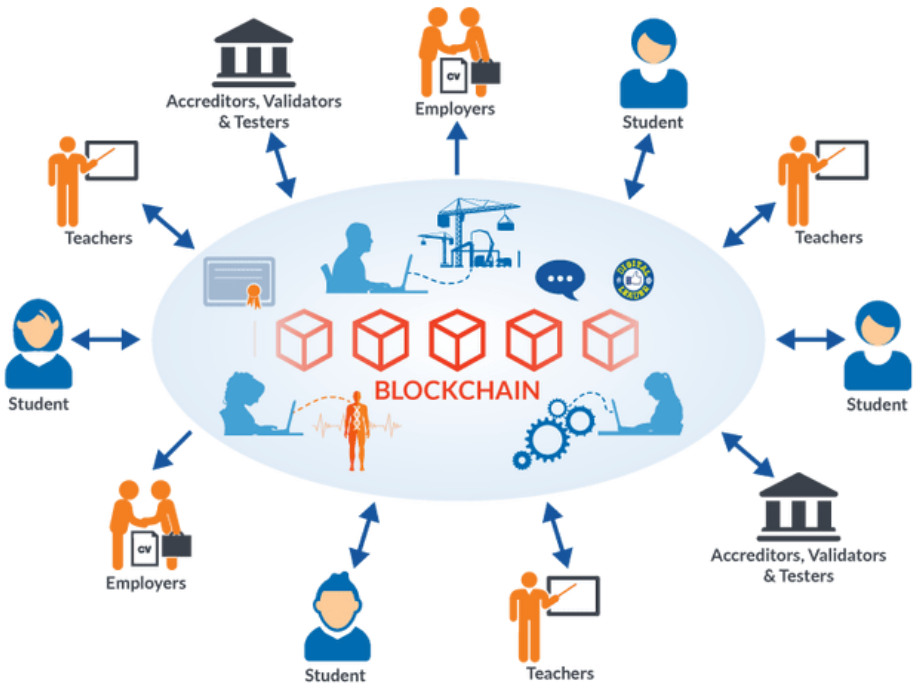
- DPoS机制类似于股份制公司
  - 普通股民进不了董事会，要投票选举代表（受托人）来代替他们做决策。
  - 那些握着加密货币的用户可以通过投票的方式随时更换这些代表，如果他们提供的算力不稳定、计算机宕机，或者试图利用手中的权利作恶，那么他们将会立刻被用户们踢出整个系统。

- 代表币种：如比特股(BTS)
- 优点：相比POS进一步缩短了达成共识时间；能耗更小
- 缺点：投票的积极性并不高；对坏节点处理不及时

区块链共识机制应用

区块链 → 人工智能

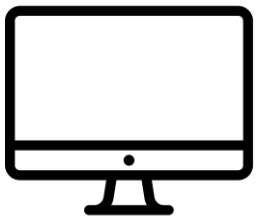
- **数据/模型/应用的共享交易平台**

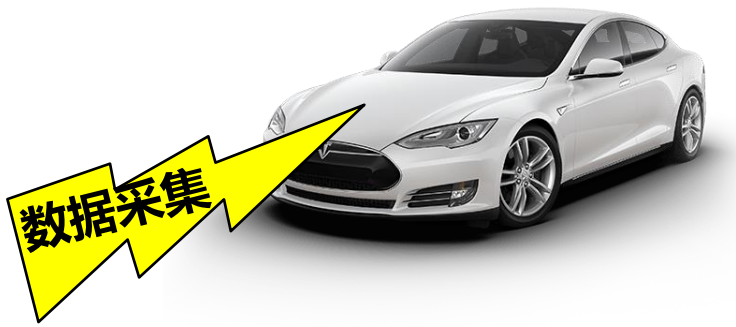


区块链最本质的功能是不可篡改的分布式账本，这一功能恰恰为建立具有公信力的数据-模型-应用共享交易平台提供了技术基础。如果数据-模型-应用都在链上使用，那么其使用情况就能够以可信方式记录，从而能够准确结算，保障各方利益。

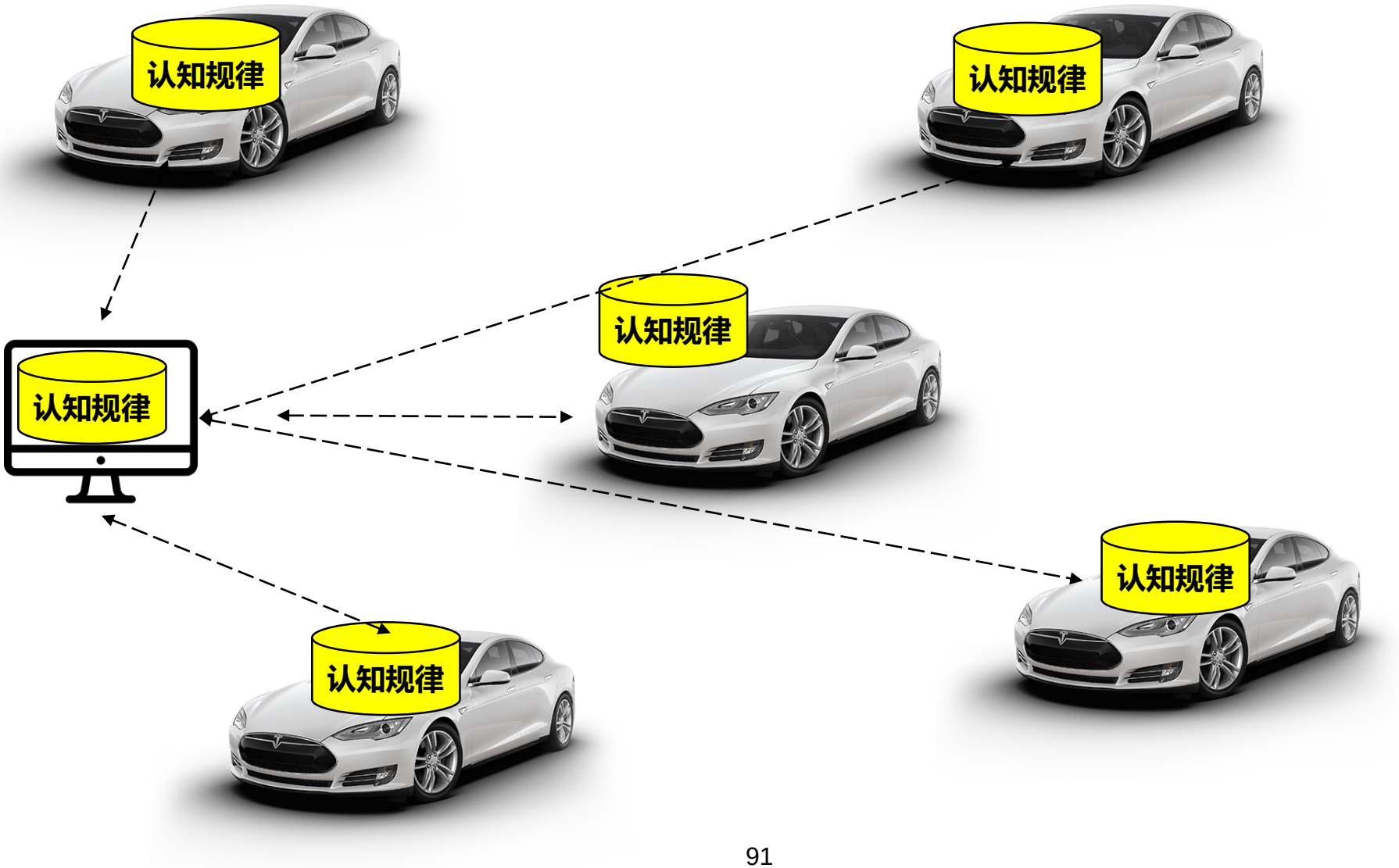
- SingularityNET (侧重数据应用DApp)
- Neuromation (侧重针对AI模型训练的合成数据)
- AI Blockchain (侧重多应用集成)
- BurstIQ (侧重医疗健康数据)
- Medical Token Currency （侧重医疗数据和模型）
- OpenMined project (本地训练模型的数据市场)
- Synapse.ai (数据和模型市场)
- Dopamine.ai (B2B的AI变现平台)
- Neuroseed (AI解决方案市场)

场景分析

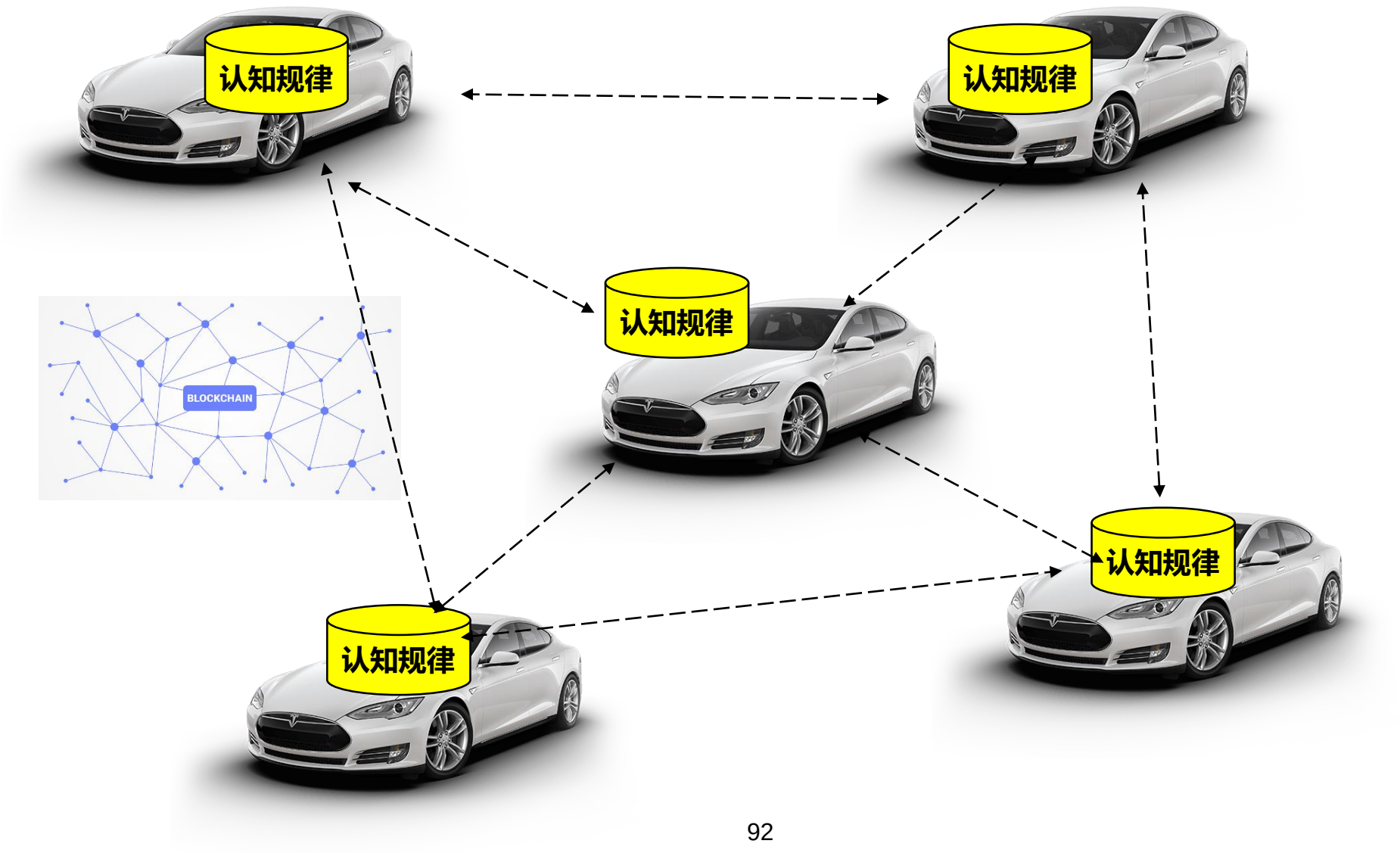




案例分析



案例分析



区块链共识机制应用

区块链 → 人工智能

- **基于区块链的分布式算力**

目前的挖矿计算不具备普世价值，但如果我们能够通过改造挖矿机制（包括算法和使用模式），区块链有可能催生世界最大的、去中心化的计算平台。当前，全球存在大量闲散算力，组织起来需要有精细权衡计算能力和网络带宽的调度方法，也需要可信的、细粒度结算机制保证利益分配，而区块链为后者提供了有效解决路径。

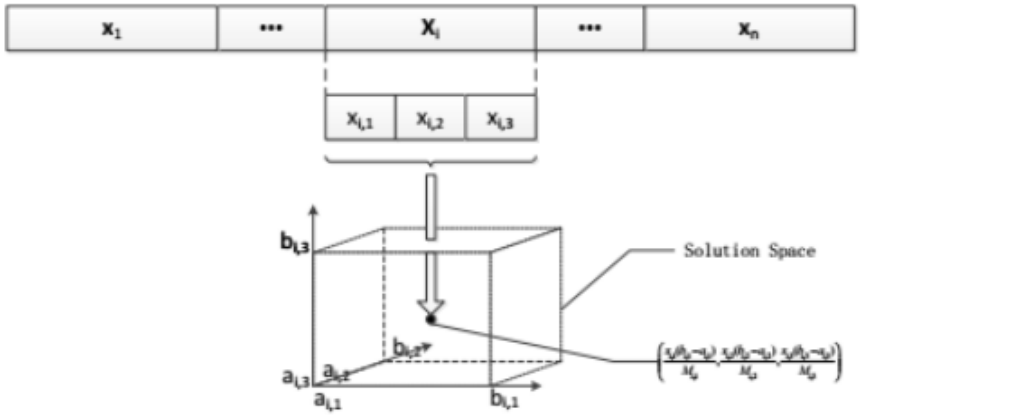
用户可以通过Golem该网络该买卖算力，这意味着用户可以在其他人的计算机上完成需要算力的工作，或者将自己空闲的算力出售给需要的人。

区块链共识机制应用

区块链 → 人工智能

- 基于区块链的分布式算力 Proof of Optimal (PoO)

算力浪费 → 求解实际优化问题



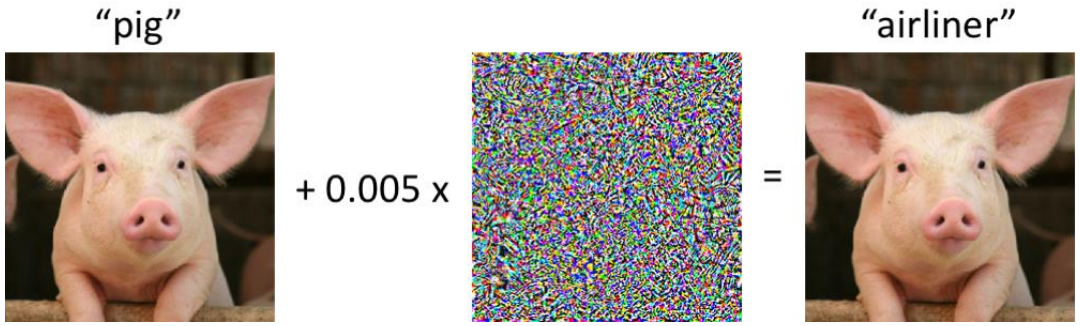
PoO共识机制中所解决的科学问题为**机器学习中的优化问题**，问题由用户提交到问题池（Optimal Pool）中，等待求解，求解结果会记录在区块链上

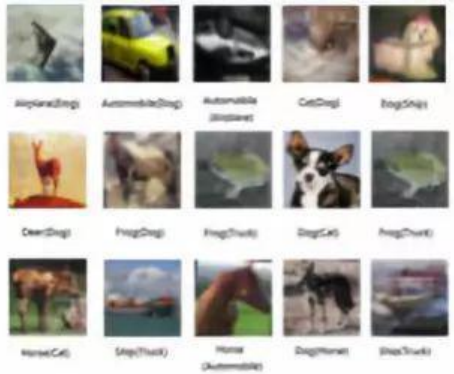
区块链共识机制应用

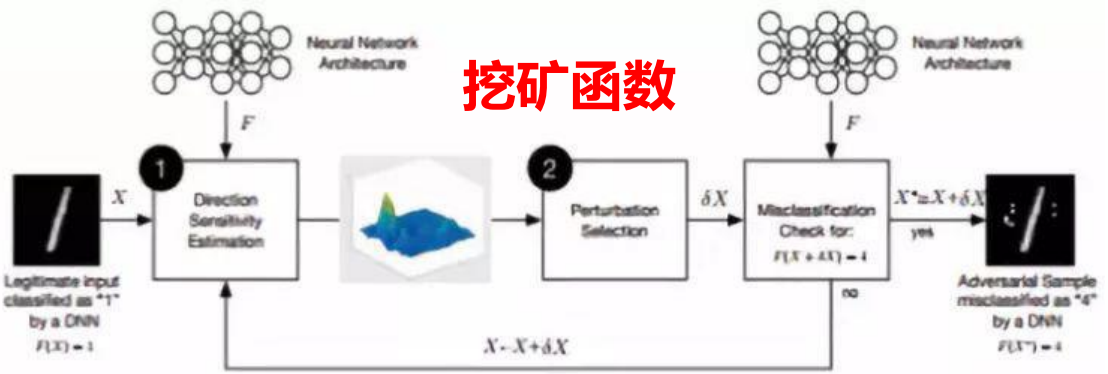
算力为AI服务：MatrixAI->对抗攻击

- 通过挖矿函数设计寻找对抗攻击实例
- 利用区块链算力提升模型鲁棒性

对抗攻击：





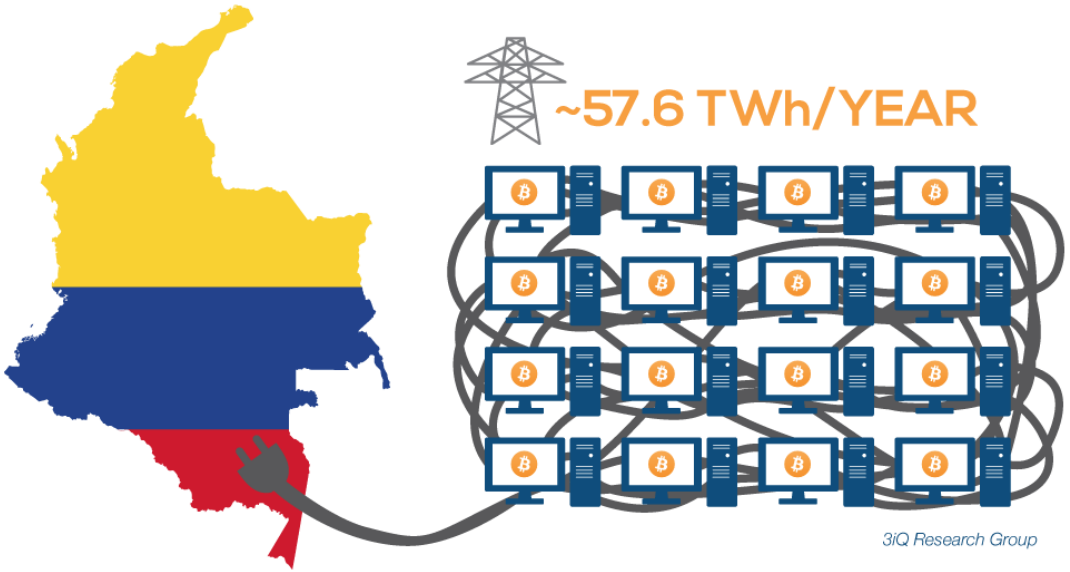


现有共识机制的局限

区块链面临的挑战

- **挖矿能耗**
- **可拓展性**
- **安全性**
- **隐私保护**

如果把全部挖矿的计算能力折算为浮点运算，粗略估算的总体计算能力达到1023FLOPS，已经达到谷歌计算能力的1百万倍，或者全球500强超级计算机总体计算能力的10万倍。如此庞大的计算能力当然以电力作为基础，其总用电量已经超过世界上160多个国家。



比特币网络一年能源消耗量相当于哥伦比亚一年用电量

区块链面临的挑战

- 挖矿能耗
- **可拓展性**
- 安全性
- 隐私保护

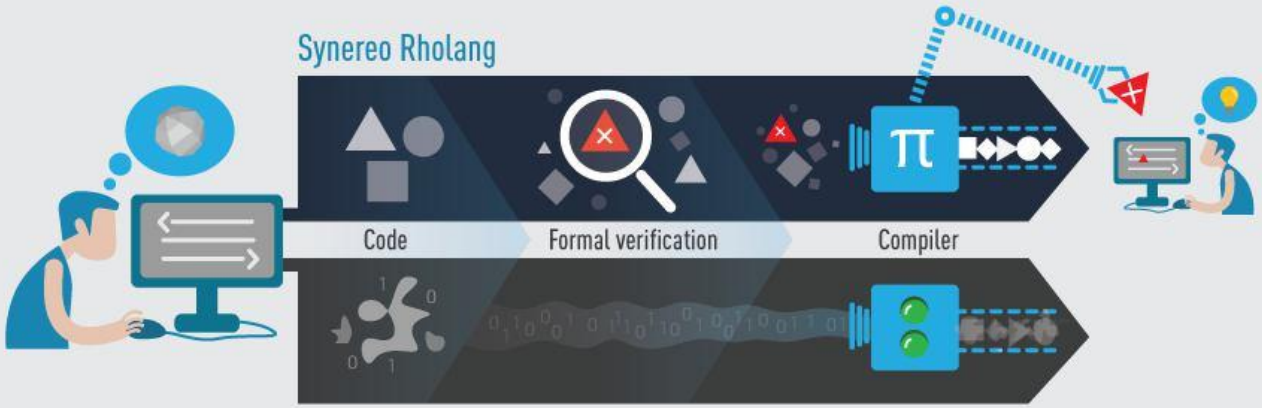
区块链的吞吐率以每秒完成交易数目Transactions Per Second（TPS）表征。比特币的吞吐率为3.3~7TPS，以太坊略高，但也只有30TPS左右。对比而言，使用中心化方式验证交易的VISA信用卡的持续吞吐率能够达到1700TPS以上（VISA官网宣称峰值可达65000TPS）。

Platform	# Transactions per second (TPS)
Visa	2000
Paypal	115
Bitcoin	7

区块链面临的挑战

- 挖矿能耗
- 可拓展性
- **安全性**
- 隐私保护

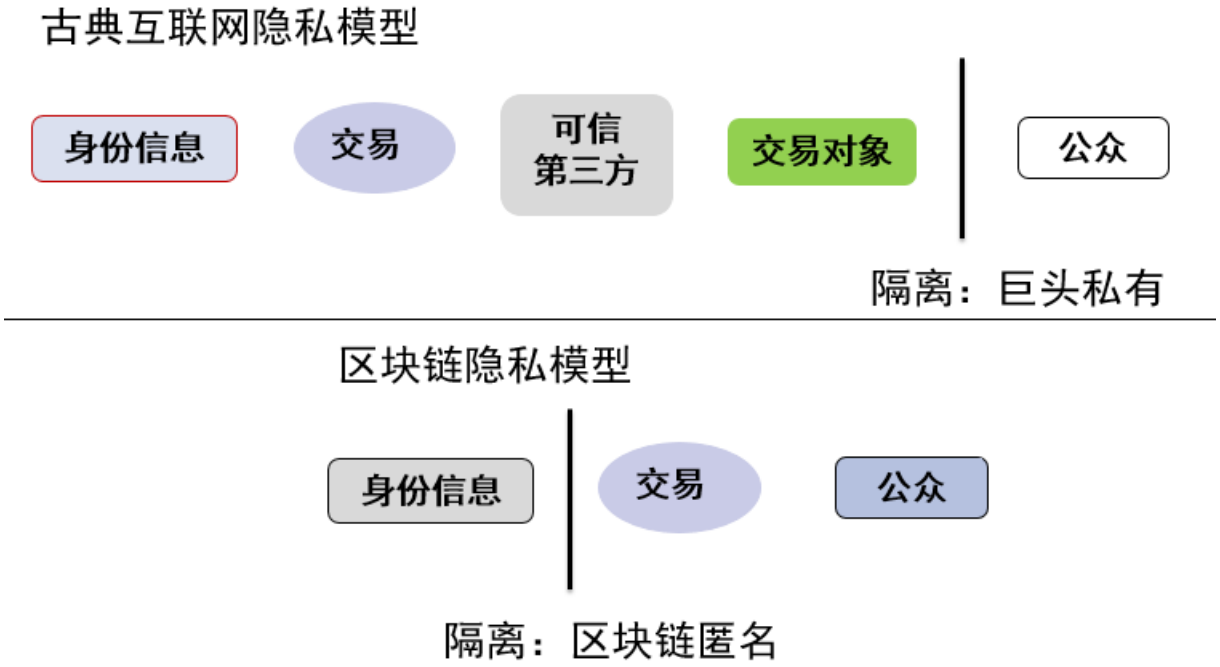
区块链采用了去中心化的共识机制，本身的安全性是比较高的。然而，区块链由网络实现，因此其网络协议的各个层次均有可能受到攻击。更为严重的安全隐患来自于智能合约。由于智能合约是具有图灵完备性的程序，因此其行为更加复杂，而且代码在分布式网络环境上运行时，潜在风险会大大提升。典型案例是以太坊上的众筹项目DAO，它在2017年受到重入攻击，被盗走当时价值6千万美元的以太币。



区块链面临的挑战

- 挖矿能耗
- 可拓展性
- 安全性
- **隐私保护**

在大数据时代，保护数据隐私的重要性不言而喻。目前区块链公链上的数据大体来说是完全开放的。因此，随着区块链应用的不断拓展以及其数据库应用比重的提升，如何在区块链上引入完备的隐私保护机制已经成为亟待解决的问题。



加入我们！

InplusLab

- 8位教授/副教授导师，研究方向涉及 区块链、人工智能、数据挖掘、分布式学习等等方向
- 请随时投简历到邮箱：inpluslab@yeah.net

InplusLab

实验室公众号