区块链与分享型数据库

主讲： 张　　召

EAST CHINA NORMAL UNIVERSITY

主讲老师简介

- **张召** 长期从事区块链系统和分布式数据库相关的教学和研究。
- 邮件：zhzhang@dase.ecnu.edu.cn
- 电话、微信：18502118359
- 课程钉钉群，作业提交，讲义分发等。

区块链2024春季班级

钉钉扫码加入班级

学习方式

- 本科生：理论+上机课程实践

  掌握已有的区块链系统的架构和关键技术，能够基于已有架构开发区块链应用。

  上课+上机

- 研究生：理论+课后实践

  掌握已有的区块链系统的架构和关键技术，能够对区块链某些关键核心算法进行改进和二次开发。

  上课+上机，上机可选

考核方式及教材

◆平时成绩：

- 课堂表现、出勤
- 上机实践、实验报告

◆期末评价

- 课程大项目

◆教材

- 区块链导论：原理、技术与应用，高等教育出版社，张召等编著

对同学们的期待和要求

- 认真上课
- 勤于思考
- 勇于实践
- 乐于交流

课堂调查

- 听说过区块链，比特币吗？
- 听过Web 3.0吗？
- 为什么选这门课程？
- 对这门课程有什么期待？

课程目标

- 比特币系统并不等于区块链系统，本课程主要专注于支撑比特币系统的核心技术——区块链技术。

- 本课程主要剖析区块链系统所涉及到的基本原理与关键技术，以及其与传统分布式数据库之间的关联。具体授课内容包括典型的区块链系统介绍、密码学基本原理、分布式一致性原理、共识协议，以及基于默克尔树的数据存储与组织等。

能力目标：

- 目标1：本课程通过区块链及主流区块链系统的介绍，使得学生掌握区块链系统的基本原理和框架。
- 目标2：通过本课程的学习掌握区块链系统所涉及的关键核心技术。

- 目标3：基于已有的区块链平台开发区块链应用系统。
- 目标4：理解开源区块链系统的源代码，并在此基础上对开源平台进行必要的功能扩展和性能提升。
- 目标1和目标2是理论能力要求，目标3和目标4是实践能力要求。

思考

- 设想一个去中心化的数字货币系统应该是什么样的？
- 设想一个永不停机的世界计算机系统应该是什么样的？
- 设想一下元宇宙中支撑虚拟资产交易的基础设施是什么？
- 设想一个支撑Web 3.0应用的基础架构是什么样的？

互联网的发展

- Web 1.0 (1991-2004)
  - 门户网站，例如谷歌、雅虎、搜狐、新浪等
- Web 2.0 (2004-现在)
  - 基于博客、社交媒体、在线社区等平台，用户可以自主生成内容，与网站和他人交互协作。
- Web3.0 （2014-）

是什么样的互联网呢?

Web 2.0 存在的问题

- 用户在个人身份、数据和服务等方面丧失了自主权。
- 中心化平台对网络资源的垄断造成价值分配的不均衡。
- 中心式基础设施让用户的数据安全和隐私保护面临更大的威胁。

Web 3.0 是什么

- 1998年由 HTTP 协议的发明者 Tim Berners Lee 提出，指的是一个集成互联网各种应用和系统的通信框架，使得人们能够访问任何网络资源的网络生态系统

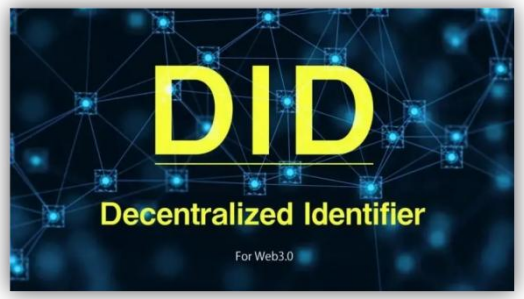
- 2014年以太坊的联合创始人、Plokdot的创建者 Gavin Wood 将Web3.0重新定义为块链技术支撑下的一种“无需信任的交互系统”

- Web 3.0系统应该具备的特点是什么?

Web3.0的设计目标

- 核心目标：数据确权。
- 通过构建去中心化的网络服务，实现用户对其身份、数据、资产等的所有权和支配权。

使用去中心化数字身份



数据管理从以App为中心转变为以用户为中心

以区块链作为事实层和激励层以实现去信任

Web3.0的发展现状

现阶段Web3.0尚处于概念阶段或初级部署阶段，底层基础设施的搭建尤为重要。

宜信区块链研究院Web3.0的核心技术堆栈全景图

Web3.0现存问题

- 应用方面

企业是否有意愿构建去中心化的应用程序？去中心化应用的产品如何迭代更新？用户体验是否良好？加密组件对用户来说是否极具挑战？

- 监管治理方面

基础层的治理将如何跨越不同的政治意识形态和文化？
政府和监管机构如何应对Web3.0？

- 技术方面

区块链能否扩展为数百万或数十亿用户提供服务？

课程所涉及的内容

课程所涉及的主要内容

- 典型区块链系统介绍
  - 比特币、以太坊和超级账本Fabric
- 密码学基础知识
  - 密码学哈希、对称、非对称、数字签名等
- 共识算法
  - 分布式一致性、RAFT、PAXOS、POW、PBFT、DAG

课程所涉及的主要内容

- 数据存储与验证
  - 默克尔树、MPT树
- 以太坊和Fabric的部署与安装
- 以太坊、Fabirc中智能合约的编写

先从比特币系统开始

01 加密货币——比特币

- 建立在计算机科学，密码学和经济学的基础上并采用共识机制实现的一套去中心化的数字货币系统

加密货币 — 传统货币

- 有政府授权的银行、支付平台等机构集中进行管理，包括：
  - 身份管理
    - 在相应银行建立账户，包括个人信息，账户信息
  - 相应服务
    - 提供转账和支付服务
  - 交易记录管理
    - 跟踪账户交易记录并为第三方提供审计服务
  - 信任
    - 由政府进行许可授权

加密货币 — 比特币

- 比特币依托分布式网络，没有集中的管理机构；
- 通过计算机科学、数学、密码学原理确保交易成功进行；
- 通过POW(proof of work)共识机制对比特币的价值进行背书；
- 货币的发行是通过比特币协议规定，总量固定，并且随着时间发行的数量递减。

加密货币 — 比特币

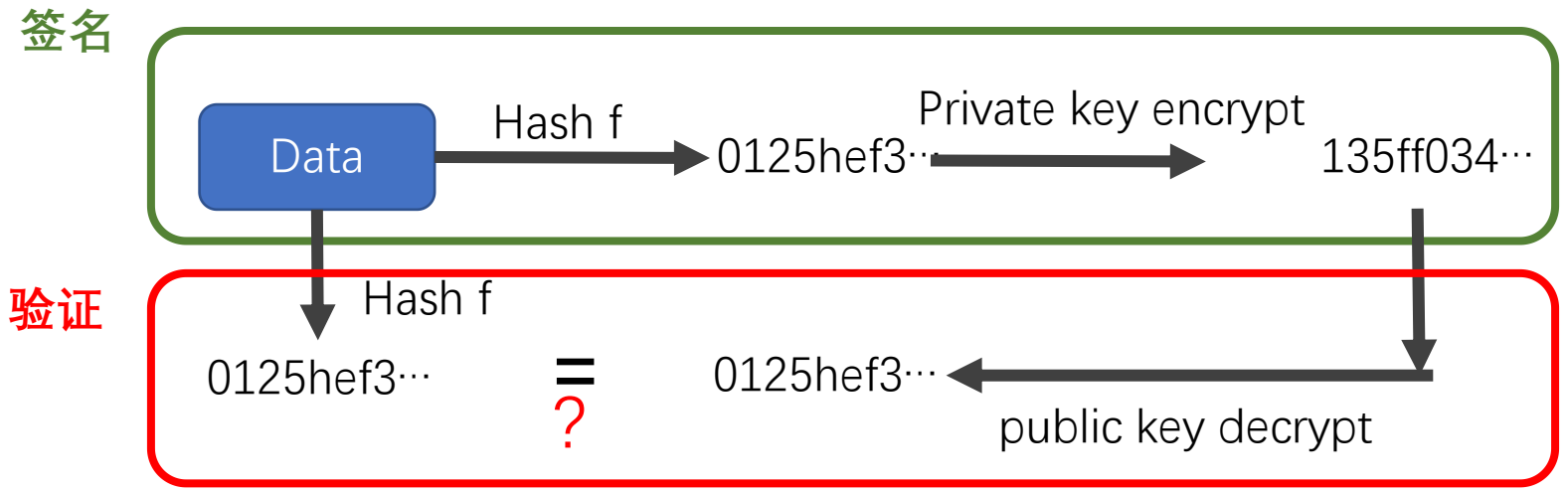
- 身份管理
- 交易服务
- 分布式账本
- 非可信共识

比特币 — 身份管理

- 以地址形式存在的账户，该账户能够拥有、转出和转入比特币；
  - 比特币网络中的地址是一个2^{160}的数字地址，从而确保任意两个账户不会重复。
- 比特币通过公有密钥和私有密钥标识一个唯一的身份。
  - 私有密钥是随机生成的，而公有密钥根据私有密钥产生；
  - 公有密钥是交易方对外公布的地址，用于接收转入；
  - 私有密钥则用于支付。

区块链 — 密码学基础

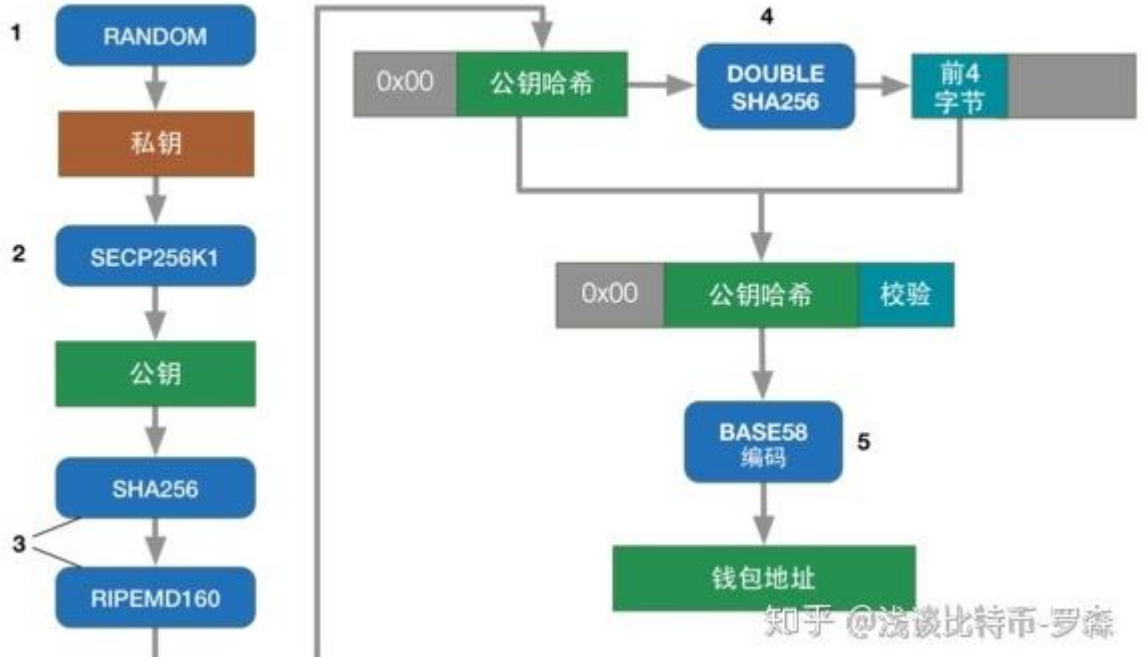
- 哈希+双向加密：
  - Hashes are “signed” with private keys
  - Hashes are verified with public keys



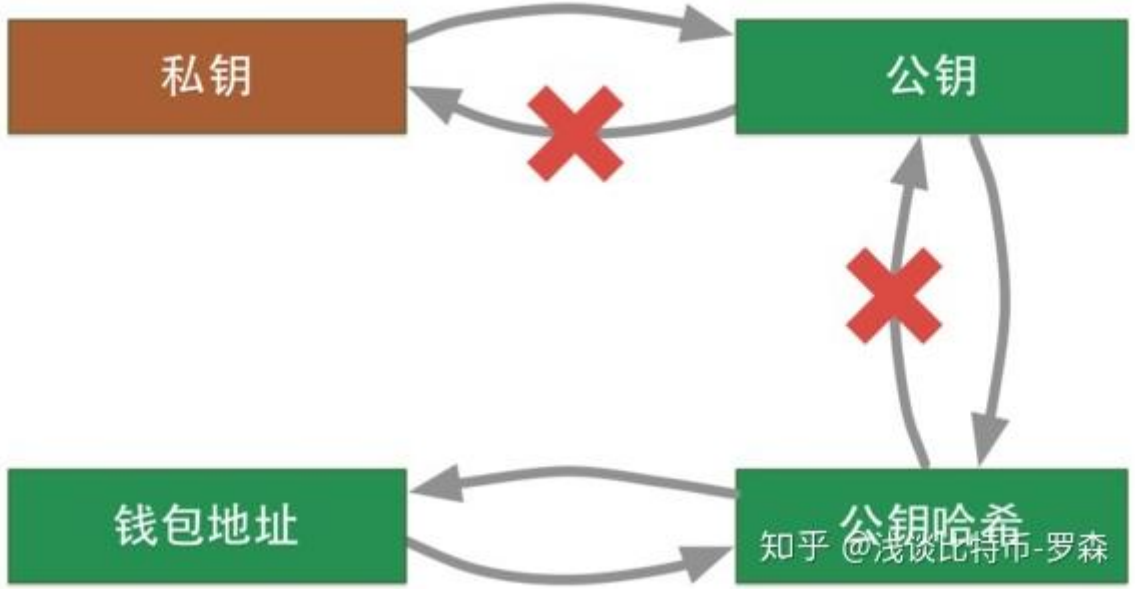
- 随机数发生器生成一个『私钥』
- 『私钥』经过SECP256K1算法处理生成了『公钥』。SECP256K1是一种椭圆曲线算法。
- 同SHA256一样，RIPEMD160也是一种Hash算法
- 将一个字节的地址版本号连接到『公钥哈希』头部（对于比特币网络的pubkey地址，这一字节为“0”），然后对其进行两次SHA256运算，将结果的前4字节作为『公钥哈希』的校验值，连接在其尾部。

- 将上一步结果使用BASE58进行编码(比特币定制版本)，就得到了『钱包地址』。

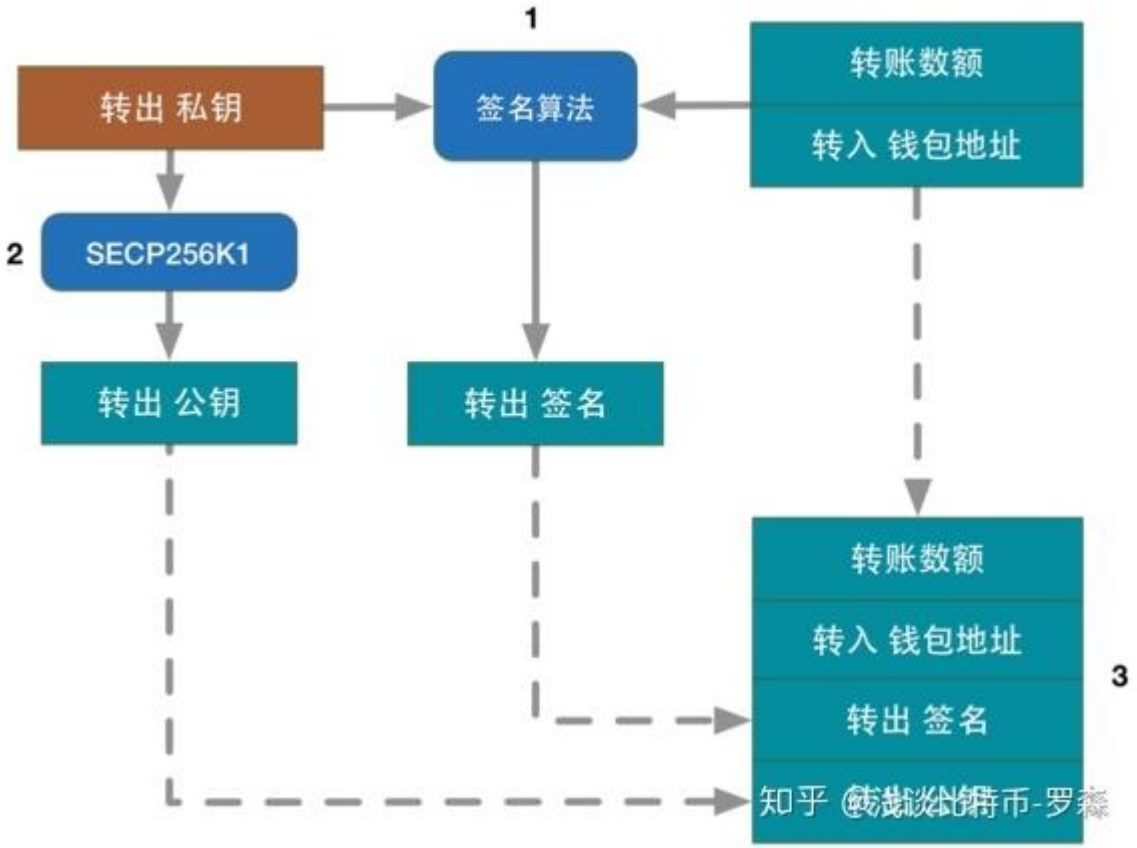
比如, 1A1zP1eP5QGefi2DMPTfTL5SLmv7DivfNa

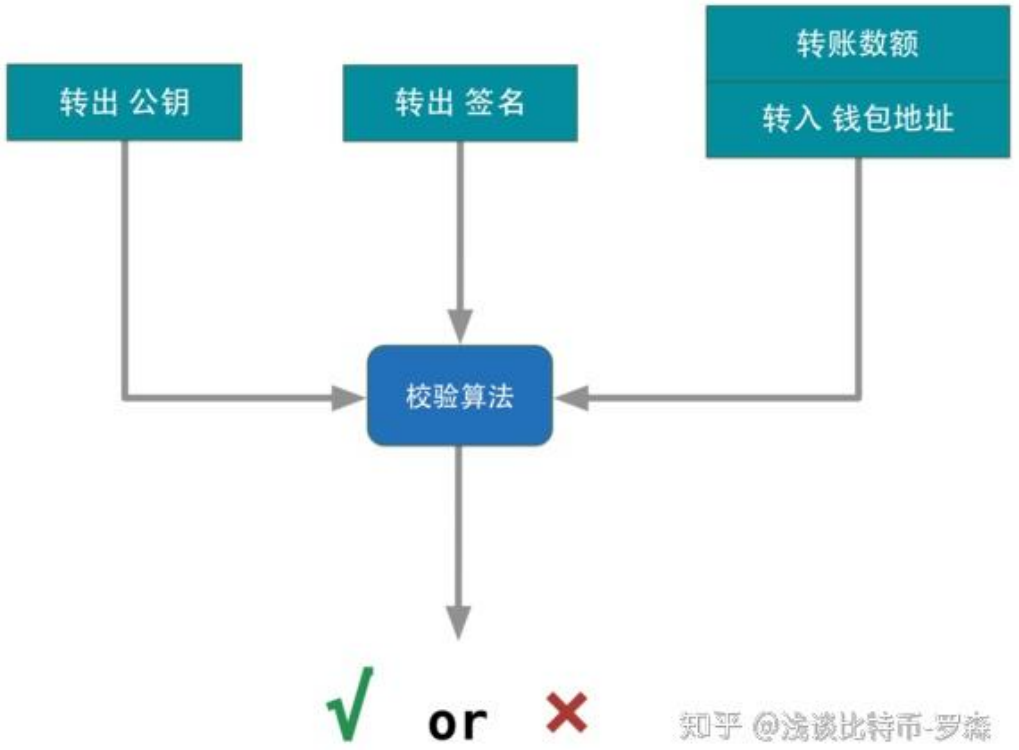


公私钥之间的关系



使用私钥签名、公钥验签


1
转出 私钥
签名算法
转账数额
转入 钱包地址
2
SECP256K1
转出 公钥
转出 签名
转账数额
转入 钱包地址
转出 签名
3
转出 公钥
知乎 @浅谈比特币-罗森

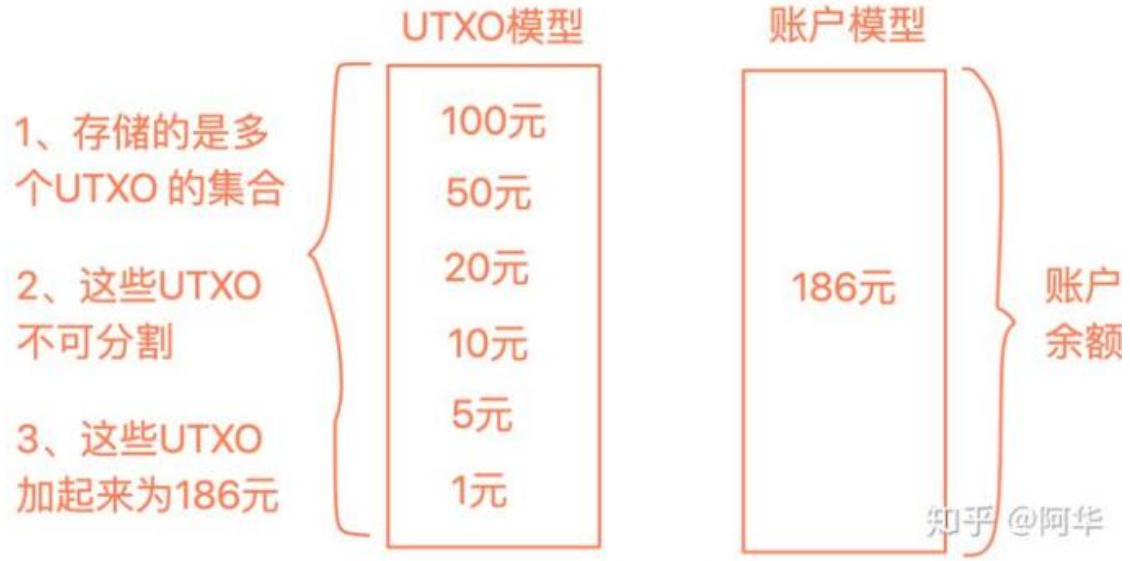

转出 公钥
转出 签名
转账数额
转入 钱包地址
校验算法
√
or
×
知乎 @浅谈比特币-罗森

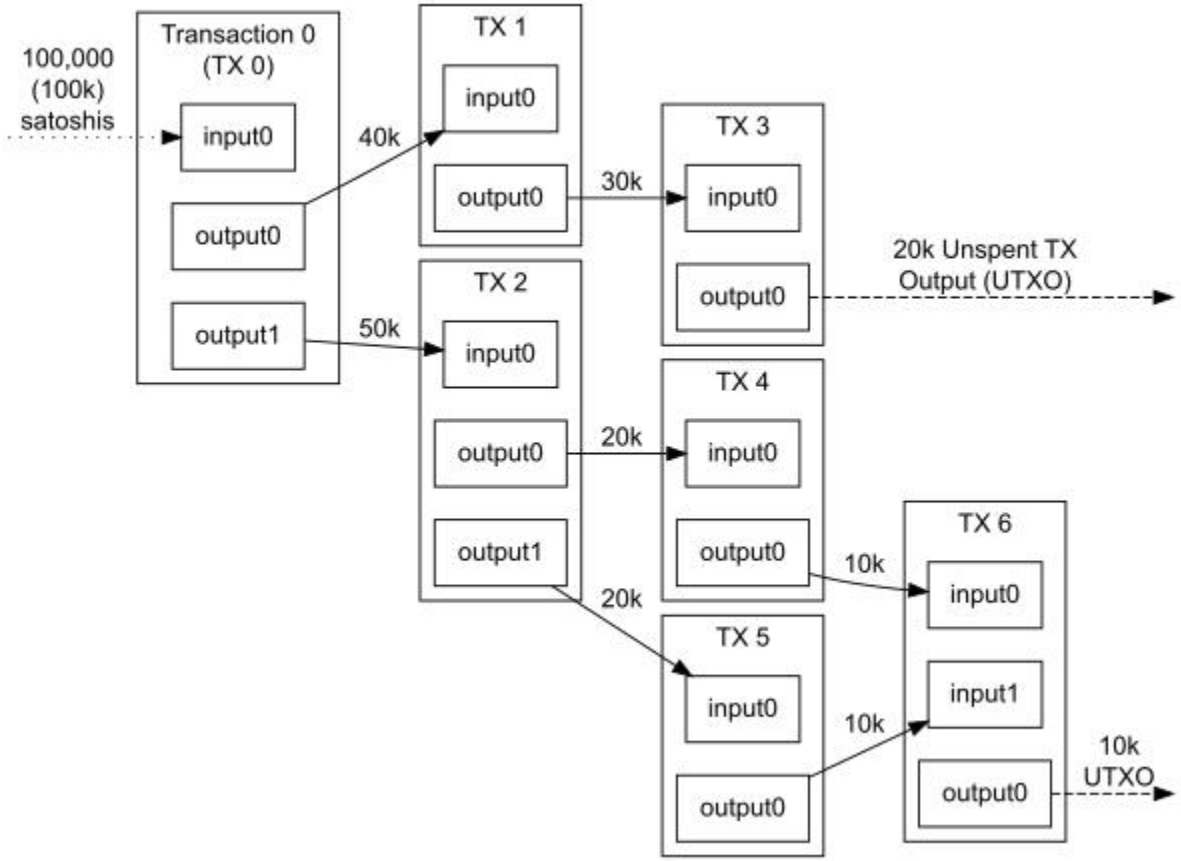
比特币 — 交易服务

- 安全可靠的转账服务，这需要满足下列条件：
  - 1）身份证明；
  - 2）足够的余额；
  - 3）余额不被其他交易使用。
- 比特币网络的每个账户拥有一组UTXOs(Unspent Transaction Outputs)，每个UTXO仅仅会支付一次并且会被整体支付。

UTXO（Unspent Transaction Out）未花费的交易输出

一切交易可追溯，交易与交易之间组成了网状关系，一个交易的输出，成为了下一个交易的输入；下一个交易的输出，又成了下下一个交易的输入





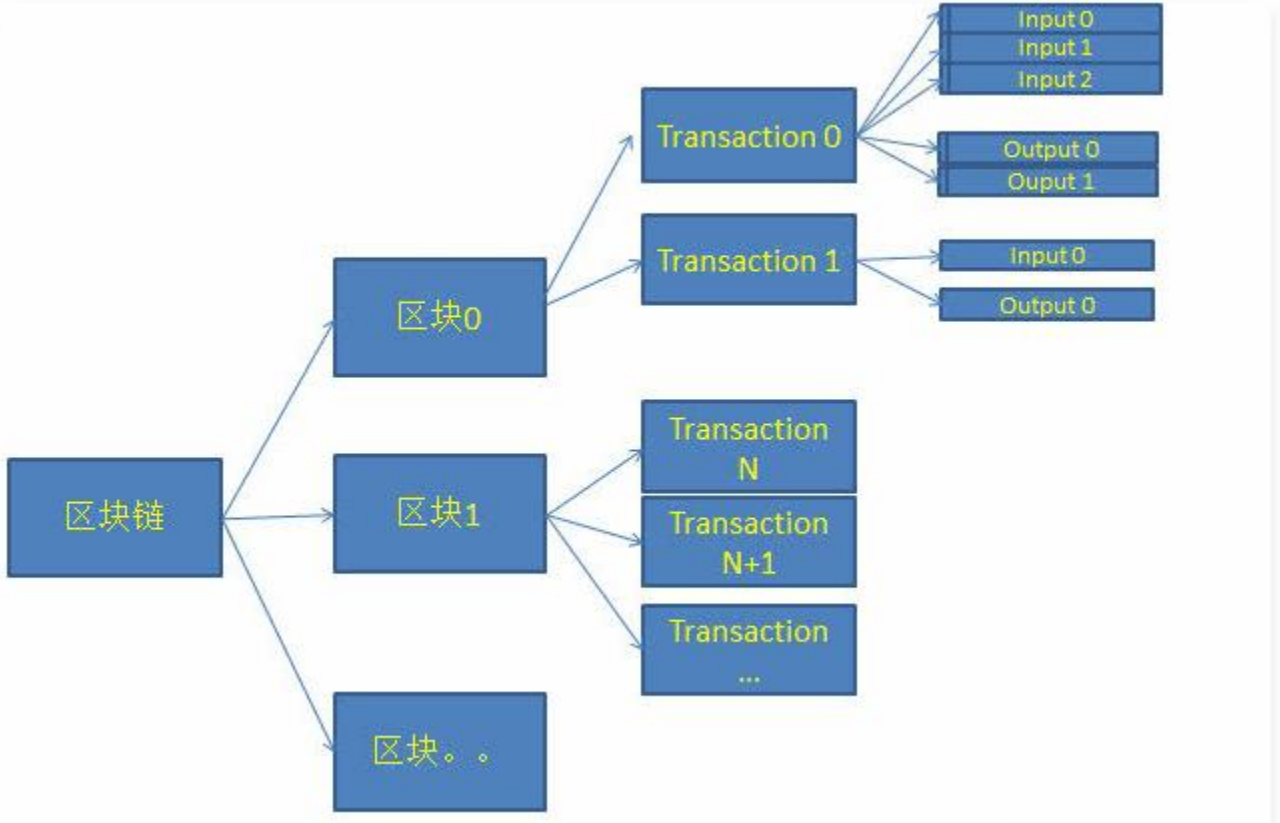
UTXO模型的转账例子

钱包

- 某个人的钱包的余额 = 属于他的UTXO的总和
- 在银行里，会存储每个账号剩余多少钱。
- 在比特币系统中，存的是一笔笔的交易，也就是一笔笔的UTXO，每个账户的余额是通过UTXO计算出来的，而不是直接存储余额。

UTXO总结

- 1个区块链，有多个区块;
- 1个区块，多笔交易（两三千笔）；
- 每笔交易，多笔Input，多笔Output。



UTXO总结

- 任何1笔Transaction，会花费多个UTXO（Input），同时也产生多个新的UTXO（Output），属于多个不同的收款人。

- 1个UTXO，具有如下的表达形式：

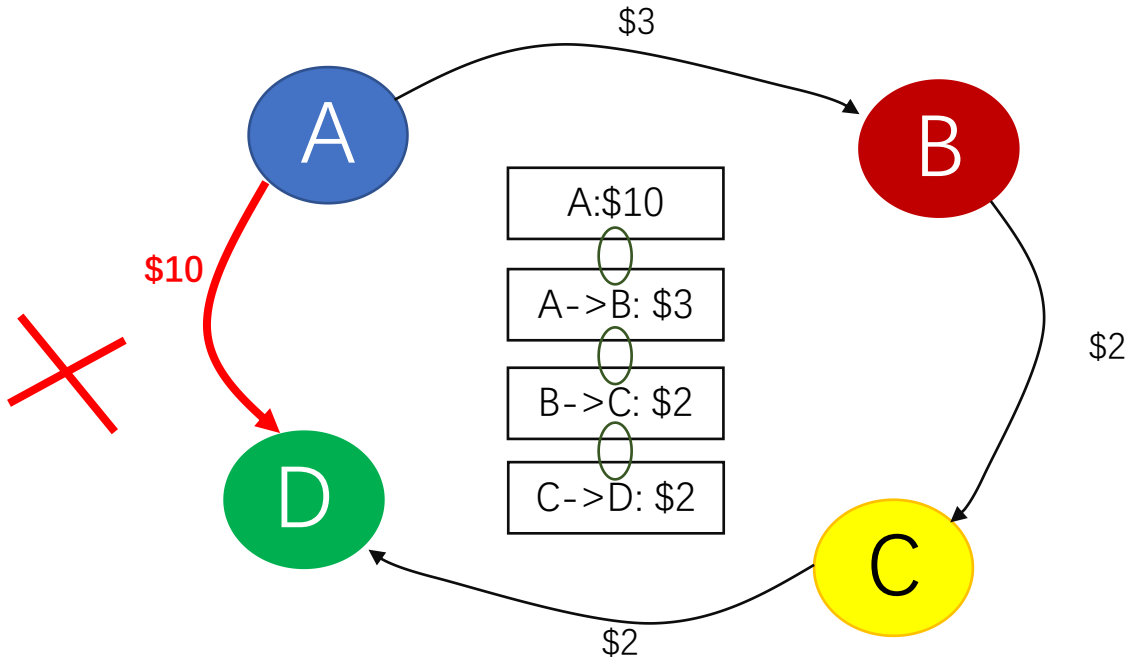
  1个UTXO = 1个Transaction ID + Output Index

- 旧的UTXO不断消亡，新的UTXO不断产生。所有的UTXO，组成了UTXO Set的数据库，存在于每个节点

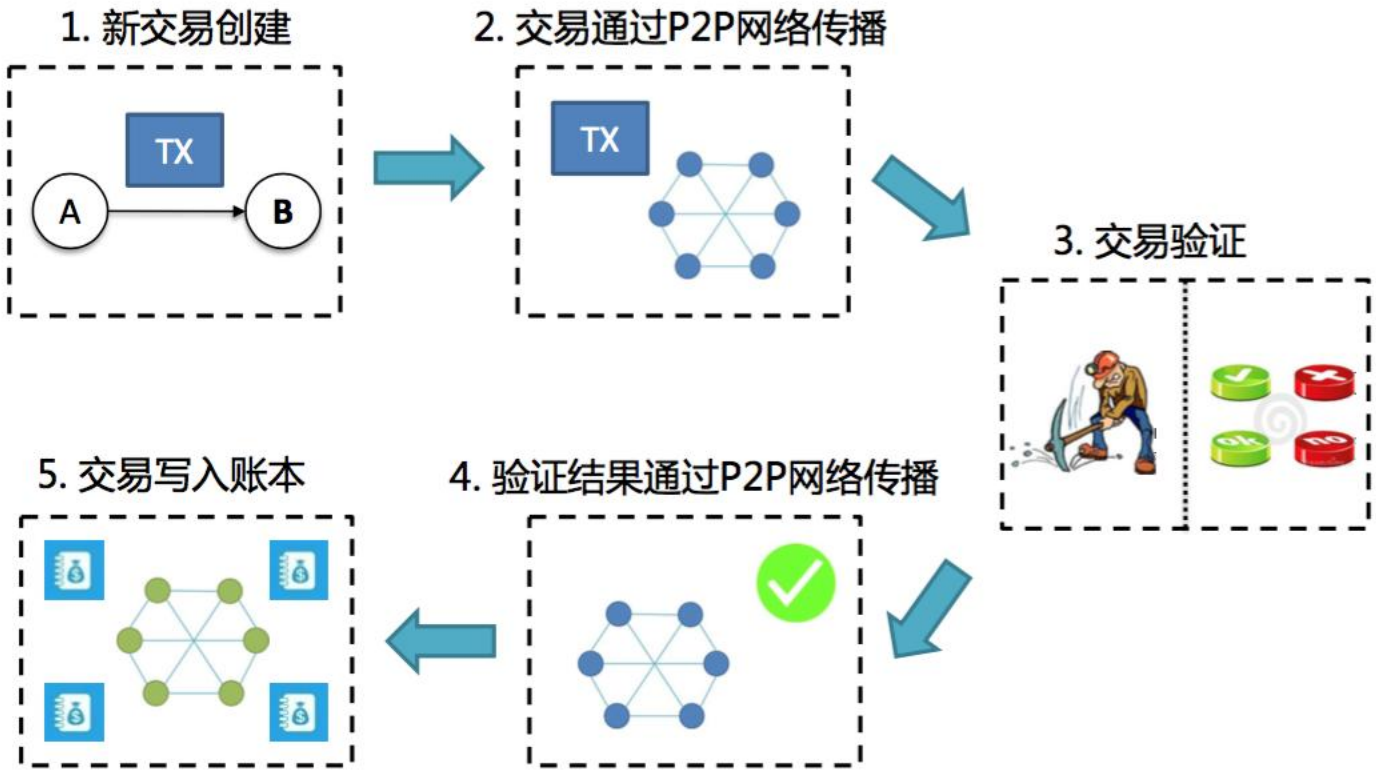
- 任何1笔UTXO，有且仅可能被1个交易花费1次

比特币 — 转账示例

- 数据层：成块的数据由密码学技术链在一起



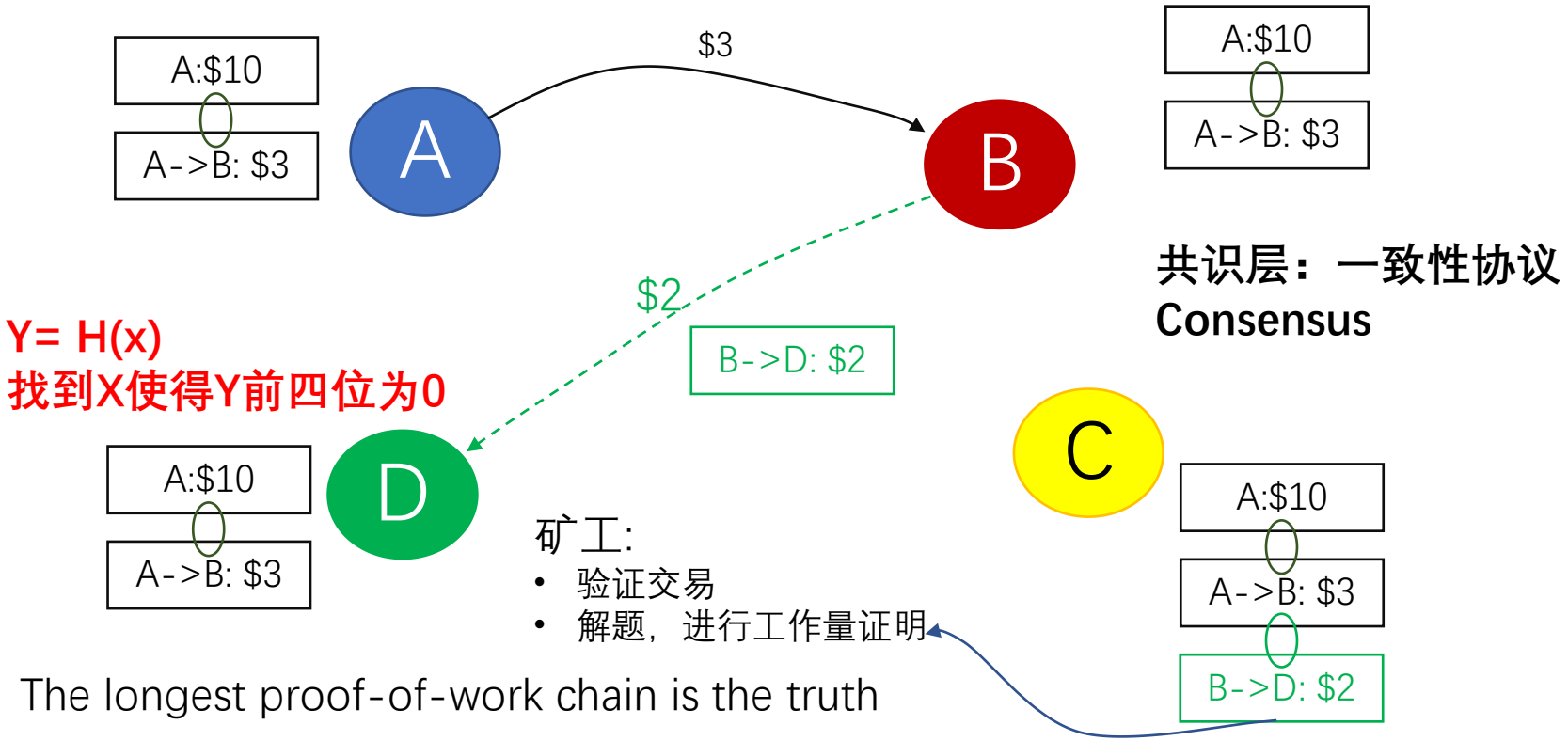
比特币 — 交易流程


1. 新交易创建
TX
A
B
2. 交易通过P2P网络传播
TX
3. 交易验证
4. 验证结果通过P2P网络传播
5. 交易写入账本

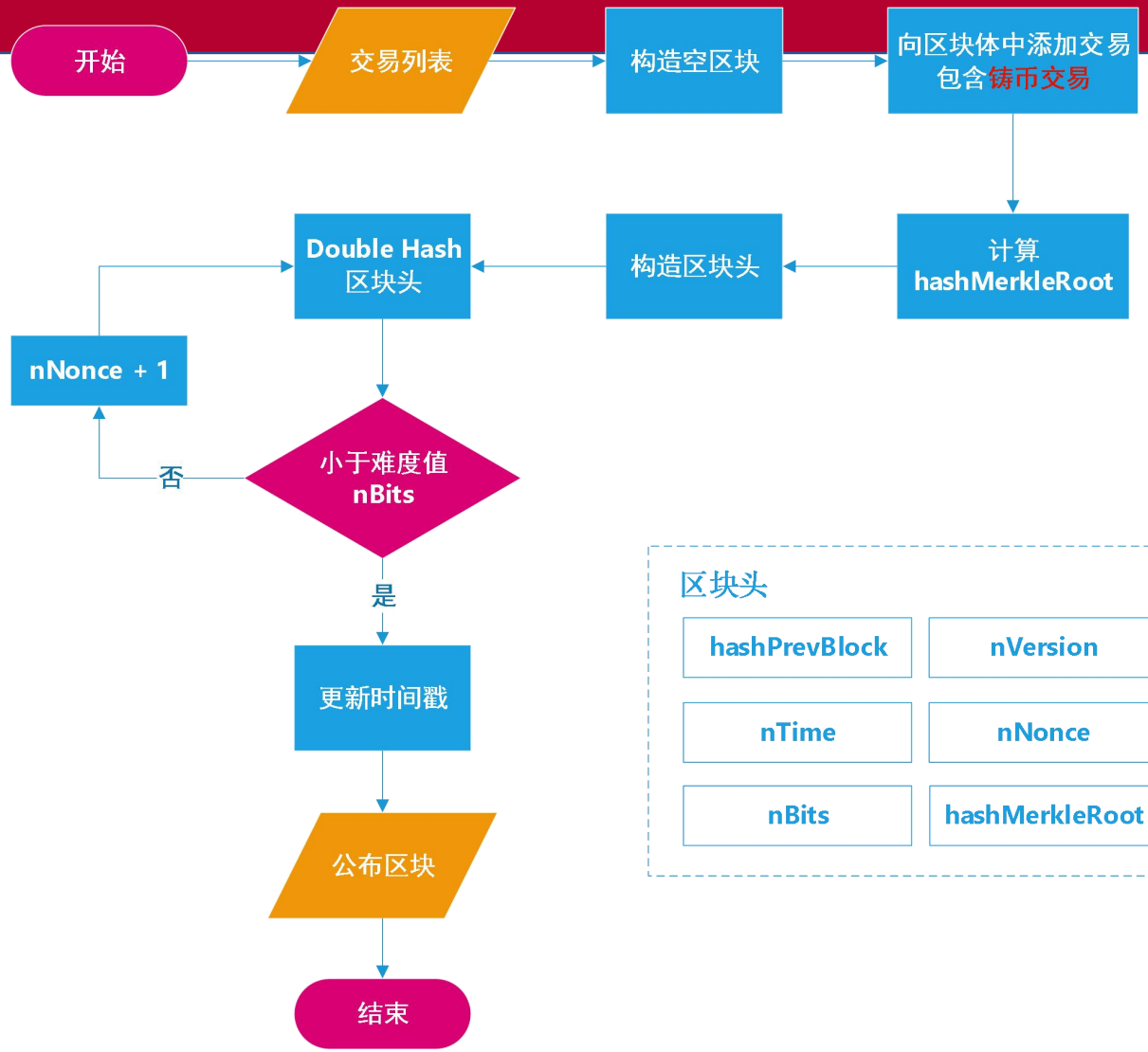
比特币 — 工作量证明（POW）

- Proof of Work - POW
  - 区块头包含一个随机数，使得区块的随机散列值出现了所需的0个数；
  - 节点通过反复尝试来找到这个随机数，这样就构建了一个工作量证明机制。
- **工作量证明机制的本质是一CPU一票，“大多数”的决定表达为最长的链，因为最长的链包含了最大的工作量。**

比特币 — POW


A:$10
A->B: $3
A
$3
B
A:$10
A->B: $3
共识层：一致性协议
Consensus
$2
B->D: $2
Y= H(x)
找到X使得Y前四位为0
A:$10
A->B: $3
D
C
A:$10
A->B: $3
B->D: $2
矿工:
• 验证交易
• 解题，进行工作量证明
The longest proof-of-work chain is the truth

比特币POW过程



区块头

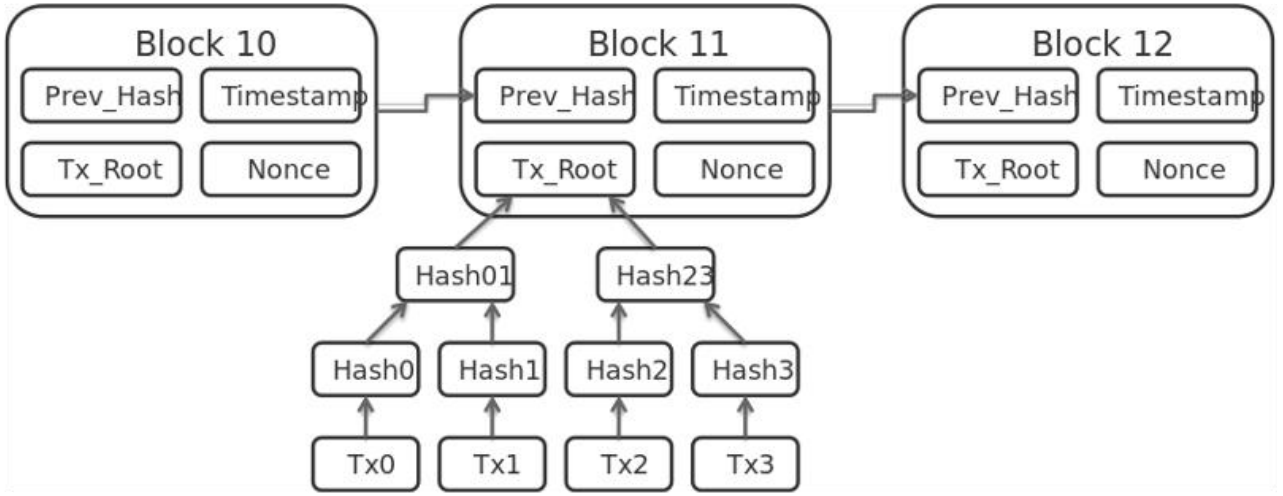
hashPrevBlock	nVersion
nTime	nNonce
nBits	hashMerkleRoot

```
//
// Search
//
unsigned int nStart = GetTime();
uint256 hashTarget = CBigNum().SetCompact(pblock->nBits).getuint256();
uint256 hash;
loop
{
    BlockSHA256(&tmp.block, nBlocks0, &tmp.hash1);
    BlockSHA256(&tmp.hash1, nBlocks1, &hash);


    if (hash <= hashTarget)
    {
        pblock->nNonce = tmp.block.nNonce;
        assert(hash == pblock->GetHash());

            //// debug print
            printf("BitcoinMiner:\n");
            printf("proof-of-work found  \n  hash: %s  \ntarget: %s\n",
            hash.GetHex().c_str(), hashTarget.GetHex().c_str());
            pblock->print();
```

```
// Update nTime every few seconds
if ((++tmp.block.nNonce & 0x3ffff) == 0)
{
    CheckForShutdown(3);
    if (tmp.block.nNonce == 0)
        break;
    if (pindexPrev != pindexBest)
        break;
    if (nTransactionsUpdated != nTransactionsUpdatedLast && GetTime() - nSta
        break;
    if (!fGenerateBitcoins)
        break;
    tmp.block.nTime = pblock->nTime = max(pindexPrev->GetMedianTimePast()+1,
}
```

比特币 — 分布式账本

- 比特币网络设计了区块链（Blockchain）结构，提供可靠、无法被恶意篡改的数字货币账本功能；
- 交易各参与方可以对该账本进行自由访问，并且任何个体无法对所记录的交易数据进行恶意篡改。



为什么需要Merkle树这样的结构

- 如何验证或确保一个数字货币的交易已经在对应区块链的一个区块中？

  成为一个全节点？

  成为一个轻节点？

SPV验证过程

- SPV钱包节点无需下载区块链完整数据，而只需下载区块链的每块不包含交易的头部数据；
- 在验证某一个交易真实性的时候，SPV钱包节点只需要把该交易哈希值向网络中连接的全节点（Full Node：同步了全部区块链数据的节点）发起询问网络里面的全节点只需要回复最小量必要数据给SPV钱包，即可验证交易真实性；
- 如果SPV钱包不信任提供交易验证数据的全节点，还可以同时发起多个全节点的询问，来确保交易验证的最大可靠性。

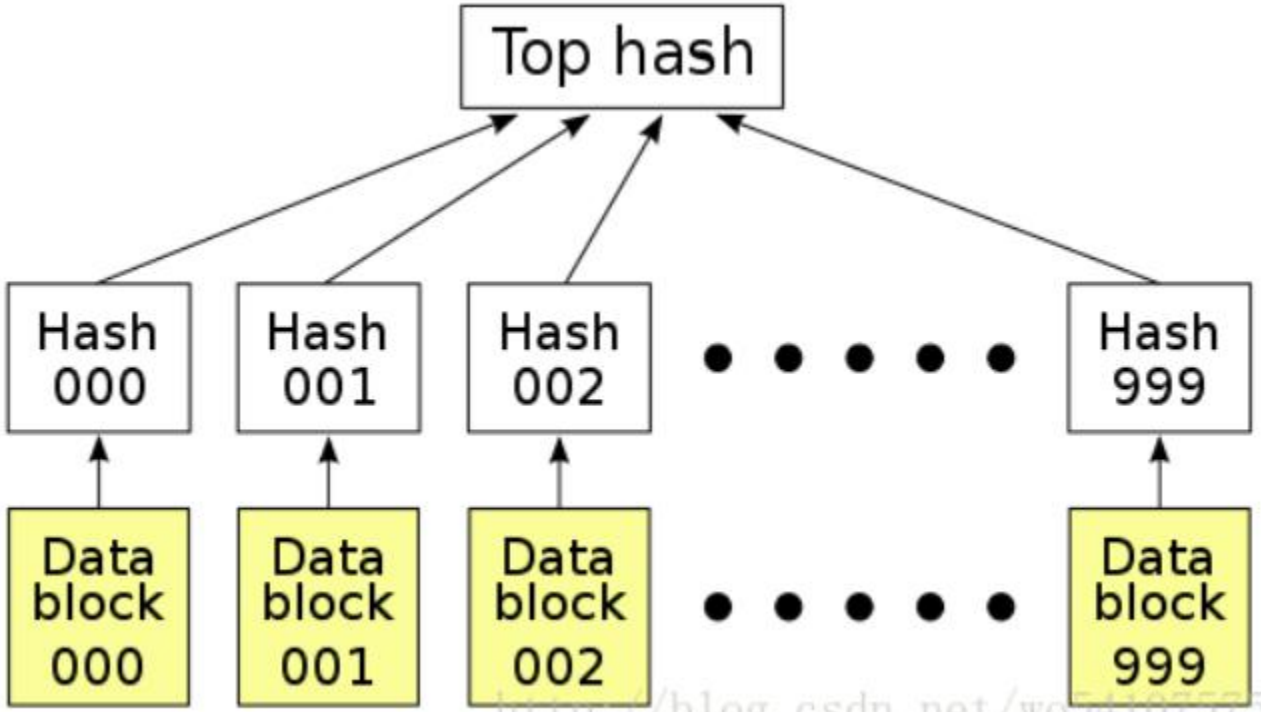
SPV节点的验证过程

- 0. 从网络上获取并保存最长链的所有block header至本地；
- 1. 计算该交易的hash值tx_hash；
- 2. 定位到包含该tx_hash所在的区块，验证block header是否包含在已知的最长链中；
- 3. 从区块中获取构建merkle tree所需的hash值；
- 4. 根据这些hash值计算merkle_root_hash；
- 5. 若计算结果与block header中的merkle_root_hash相等，则交易真实存在；
- 6. 根据该block header所处的位置，确定该交易已经得到多少个确认。

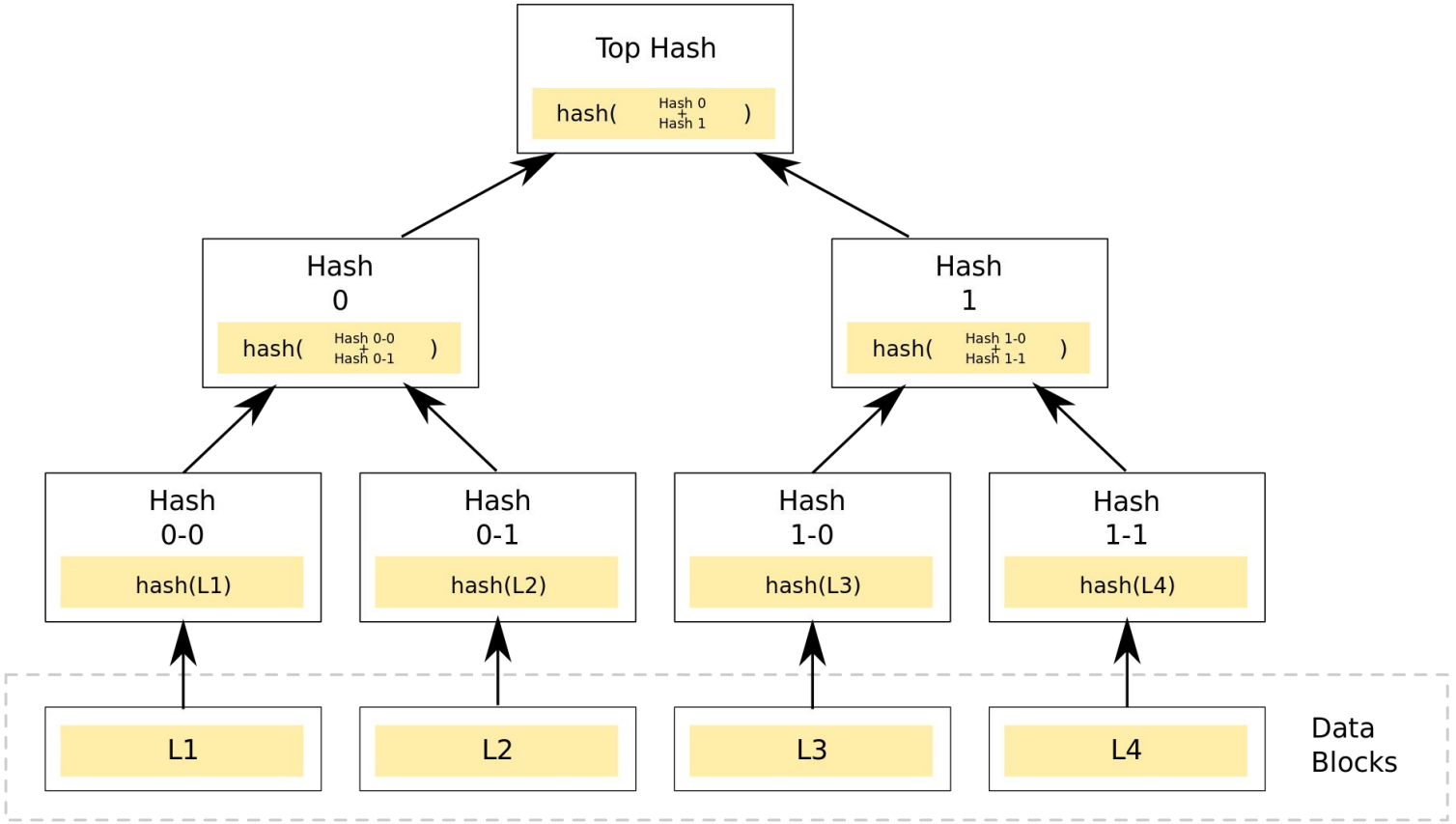
在P2P网络下载中的应用

- **Hash List**
- **Merkle Tree**
- 以上两种数据结构在确认数据块是否有损坏时有什么差别？

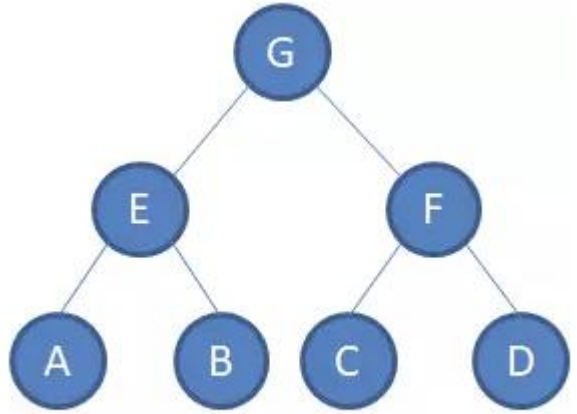
使用Hash List可以吗?



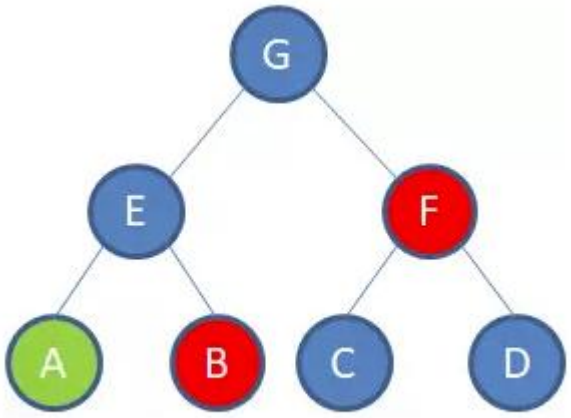
区块链 - Merkle Tree



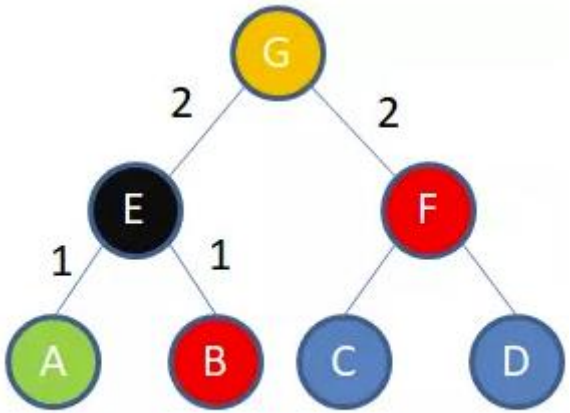
比特币中的默克尔验证



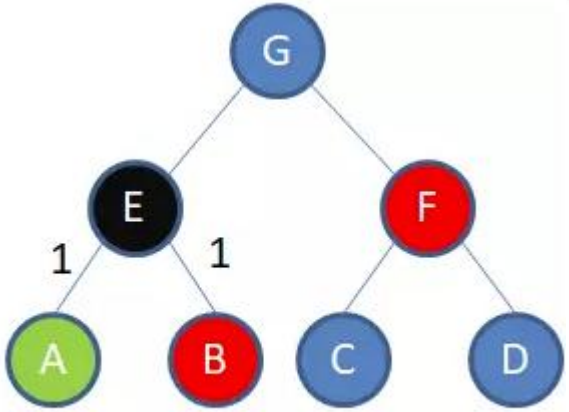
默克尔树对应的树形图



红色节点对应的就是节点A的默克尔树分支



验证第二步是得到节点G

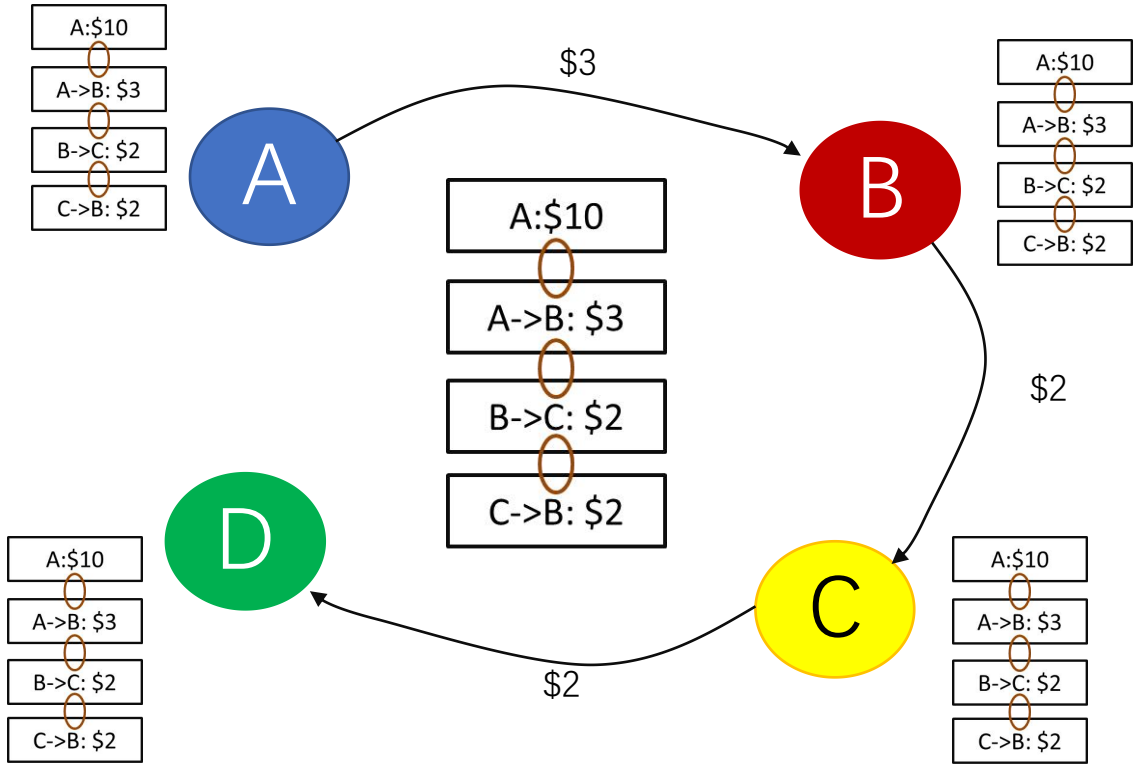


验证第一步是得到节点E

交易索引0对应的默克尔树分支

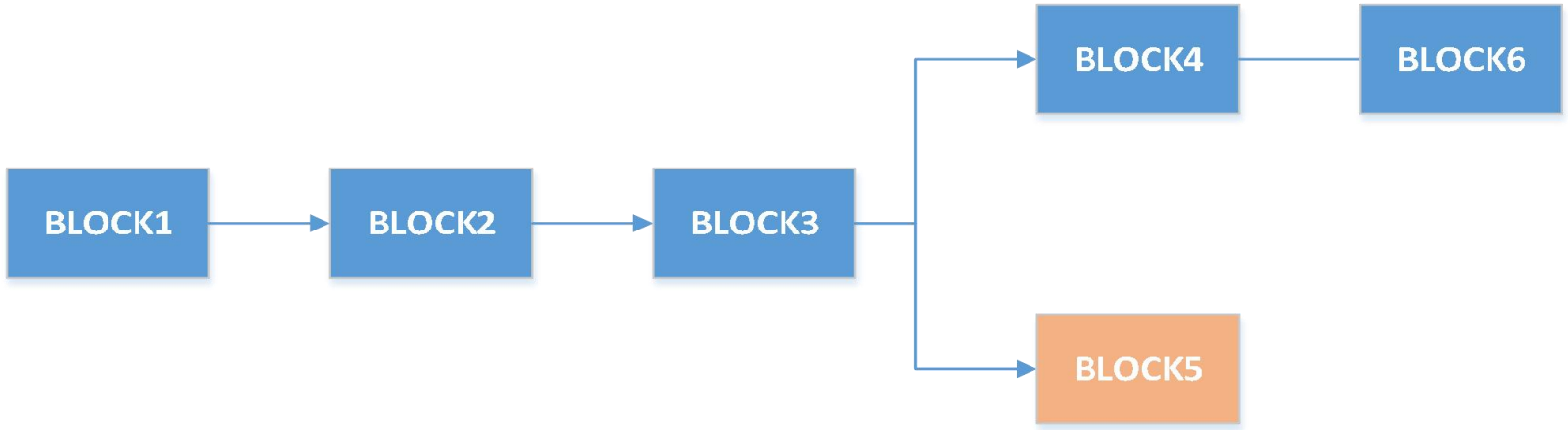
比特币 — 转账账本存储

网络层：P2P
数据层：全复制



比特币 — 分叉 (Fork)

- 同一时间段内全网不止一个节点能计算出随机数，即会有多个节点在网络中广播它们各自打包好的临时区块（都是合法的）。
- 某一节点若收到多个针对同一前续区块的后续临时区块，则该节点会在本地区块链上建立分支，多个临时区块对应多个分支。



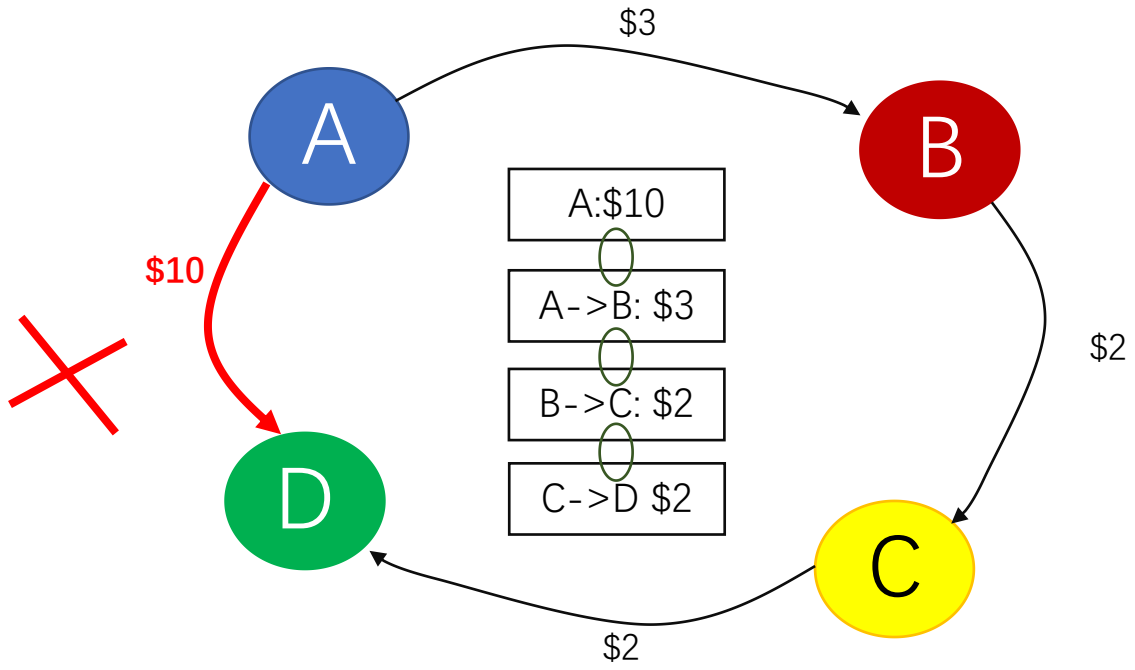
比特币 — 非可信共识

- 比特币网络中交易方是匿名的，不可信的。
- 存在：Double Spend attack 和 Sybil attach。

比特币 — Double Spend attack

- 双花，即二重支付，指攻击者几乎同时将同一笔钱用作不同交易
  - 节点在把新收到的交易单加入区块之前，会顺着交易的发起方的公钥向前遍历检查，即检查当前交易所用的币是否确实属于当前交易发起方；
  - 多份交易单可以任意序进行广播，但它们最终被加入区块时必定呈现一定的顺序；
  - 区块之间以Hash值作为时间戳，则任意一笔交易资金来源都可以被确定的回溯。

比特币 – 转账示例中的双花



02

支撑比特币的关键技术——区块链

- 基于比特币网络提炼的区块链技术，逐步受到关注并让大家看到了更高效、更安全的未来商业网络的可能。

区块链

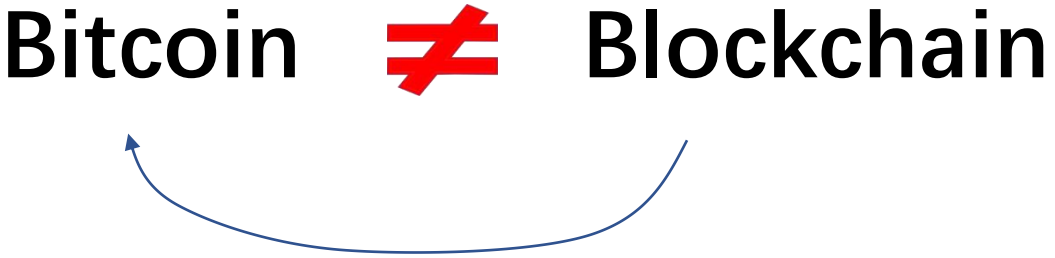
- 中本聪在2008年，于《比特币白皮书》中提出“区块链”概念，并在2009年创立了比特币网络，开发出第一个区块，即“创世区块”。
- 区块链是一个分布式账本，一种通过去中心化、去信任的方式由参与各方维护一个可靠、不可篡改交易记录的技术方案。
- 从数据库角度可将区块链比作一种分布式数据库技术，通过维护数据块的链式结构，可以维持持续增长的、不可篡改的数据库记录。

区块链

- 区块链技术被明确写入国务院的“十三五规划、十四五计划”中。
- 2019年10月24日中共中央政治局就区块链技术发展现状和趋势进行第十八次集体学习，明确要把区块链作为核心技术自主创新重要突破口。
- 2018年美国众议院多次召开区块链听证会,上升为美国国家战略。
- 2017年以来BAT公司相继宣布全面拥抱区块链技术。
- 应用范围
  - 银行票据流转
  - 大宗商品仓单流转
  - 沃尔玛，京东，天猫溯源应用
  - Arcade City -- 去中心化打车应用
  - Steemit -- 去中心化“简书”
  - NFT，元宇宙
  - Web 3.0

区块链 — 认识的误区

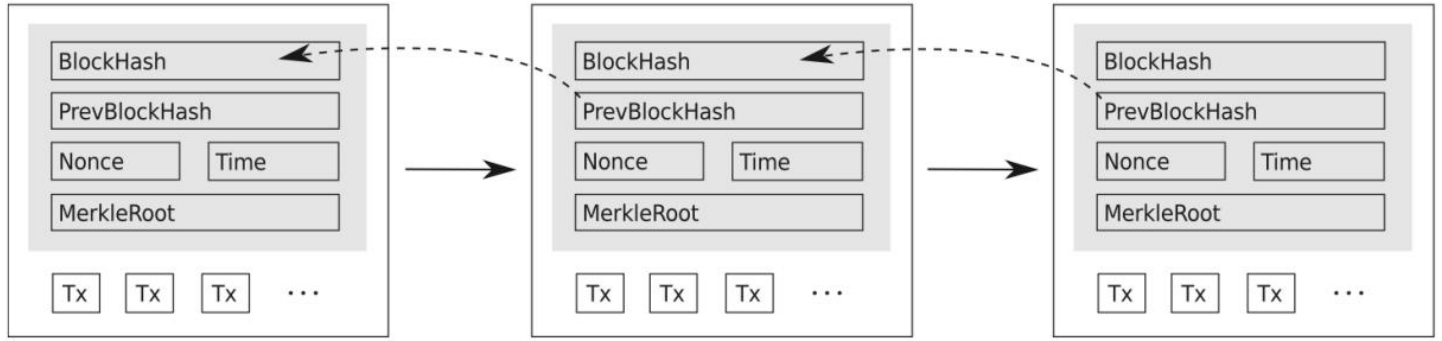
- 比特币不等于区块链
- 区块链不等于数据库
- 区块链不是万能的



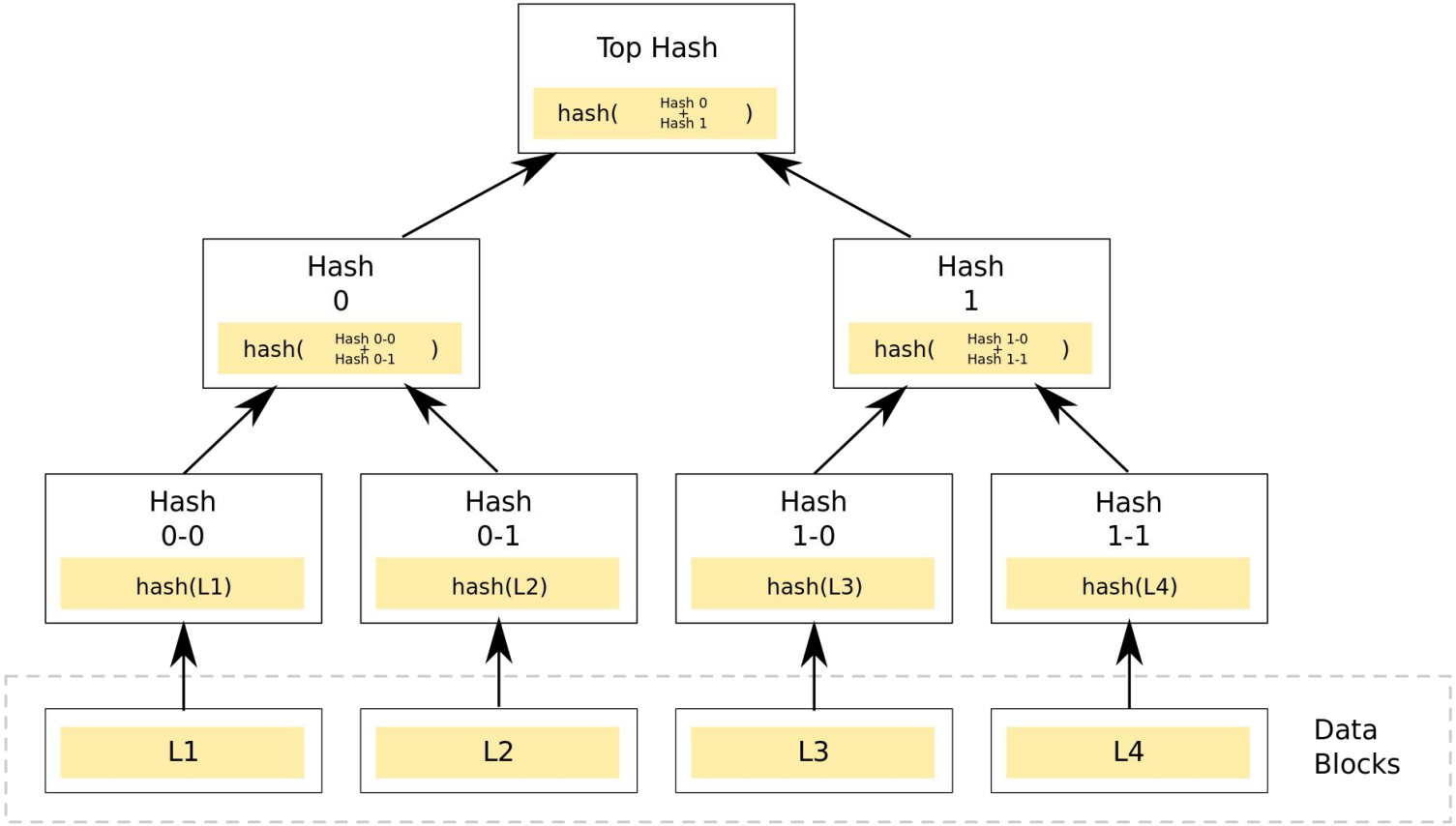
区块链是支持比特币转账的底层技术

区块链 — 逻辑结构

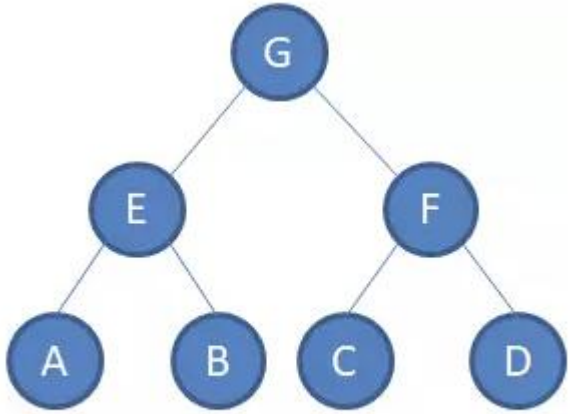
- 区块（Blocks）
- 以哈希值的形式编码成Merkle树存储批量有效的交易
- 每个区块存储前导区块的哈希值，从而构成从最新区块到创世区块的链状结构
- 每次共识达成的区块，就是参与各方对状态变更结果的一次确认
- 不可篡改，可追溯



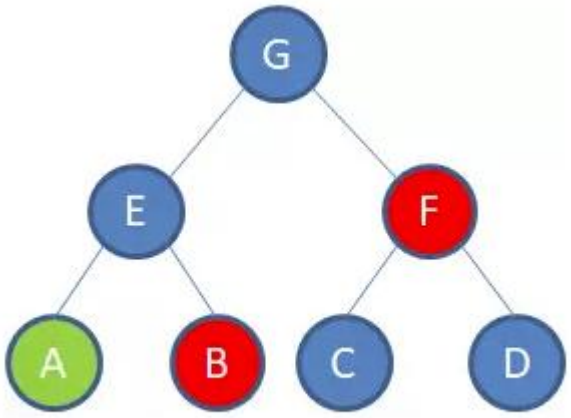
区块链 — Merkle Tree



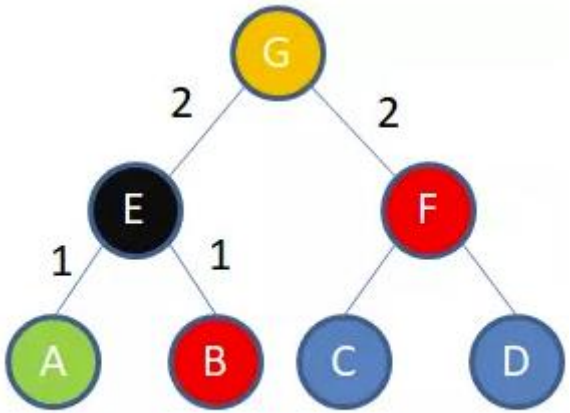
比特币中的默克尔验证



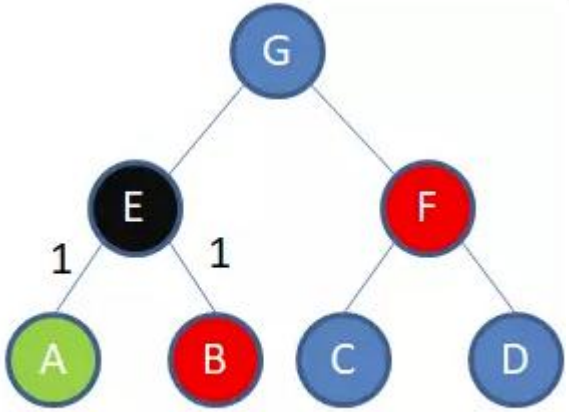
默克尔树对应的树形图



红色节点对应的就是节点A的默克尔树分支



验证第二步是得到节点G



验证第一步是得到节点E

交易索引0对应的默克尔树分支

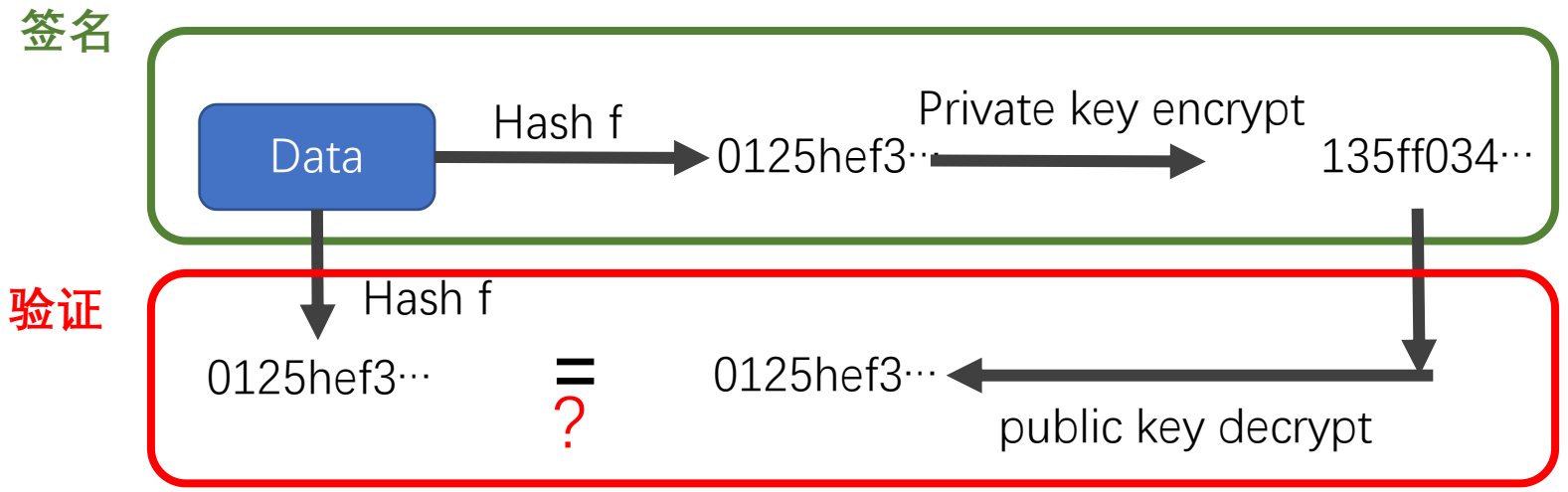
区块链 — Hash算法

- 将任何一串数据输入到SHA256将得到一个256位的Hash值（散列值）
- 特点：相同的数据输入将得到相同的结果，且结果无法事先预知
  - 正向计算（由数据计算其对应的Hash值）十分容易
  - 逆向计算（俗称“破解”，即由Hash值计算出其对应的数据）极其困难，在当前科技条件下被视作不可能

"Hello, world!0" => 1312af178c253f84028d480a6adc1e25e81caa44c749ec81976192e2ec934c64
"Hello, world!1" => e9afc424b79e4f6ab42d99c81156d3a17228d6e1eef4139be78e948a9332a7d8
"Hello, world!2" => ae37343a357a8297591625e7134cbea22f5928be8ca2a32aa475cf05fd4266b7
...
"Hello, world!4248" => 6e110d98b388e77e9c6f042ac6b497cec46660deef75a55ebc7cfdf65cc0b965
"Hello, world!4249" => c004190b822f1669cac8dc37e761cb73652e7832fb814565702245cf26ebb9e6
"Hello, world!4250" => 0000c3af42fc31103f1fdc0151fa747ff87349a4714df7cc52ea464e12dcd4e9

区块链 — 密码学基础

- 哈希+双向加密：
  - Hashes are “signed” with private keys
  - Hashes are verified with public keys



区块链 — 账号模型

- **UTXO （Unspent Transaction Output) 模型**
  - 比特币
- **Account-based模型**
  - 以太坊
  - Fabric

区块链 - 去中心化存储

- P2P网络，节点地位相同，不存在中心化的设备和管理机构
  - 每个节点都允许获得一份完整的数据库拷贝，任一节点失效，其余节点仍能正常工作；
  - 通过大规模数据复制和计算信任来确保数据质量。

区块链 — 开放式访问

- 交易参与方
  - 任何人或设备都可以参与到区块链网络
- 网络中的节点
  - 网络中的一个节点，无需得到官方机构许可

区块链 — 主要特征

- 匿名：交易参与方与网络中的相关节点都是匿名的；
- 不可信：开放的环境需要通过数字签名技术进行验证，按照系统既定的规则运行；
- 去中心化：基于P2P，不存在中心化的设备和管理机构
  - 任何人都可以参与到区块链网络，每一台设备都能作为一个节点，每个节点都允许获得一份完整的数据库拷贝；
  - 节点间基于一套共识机制，通过竞争计算共同维护整个区块链；
  - 任一节点失效，其余节点仍能正常工作。

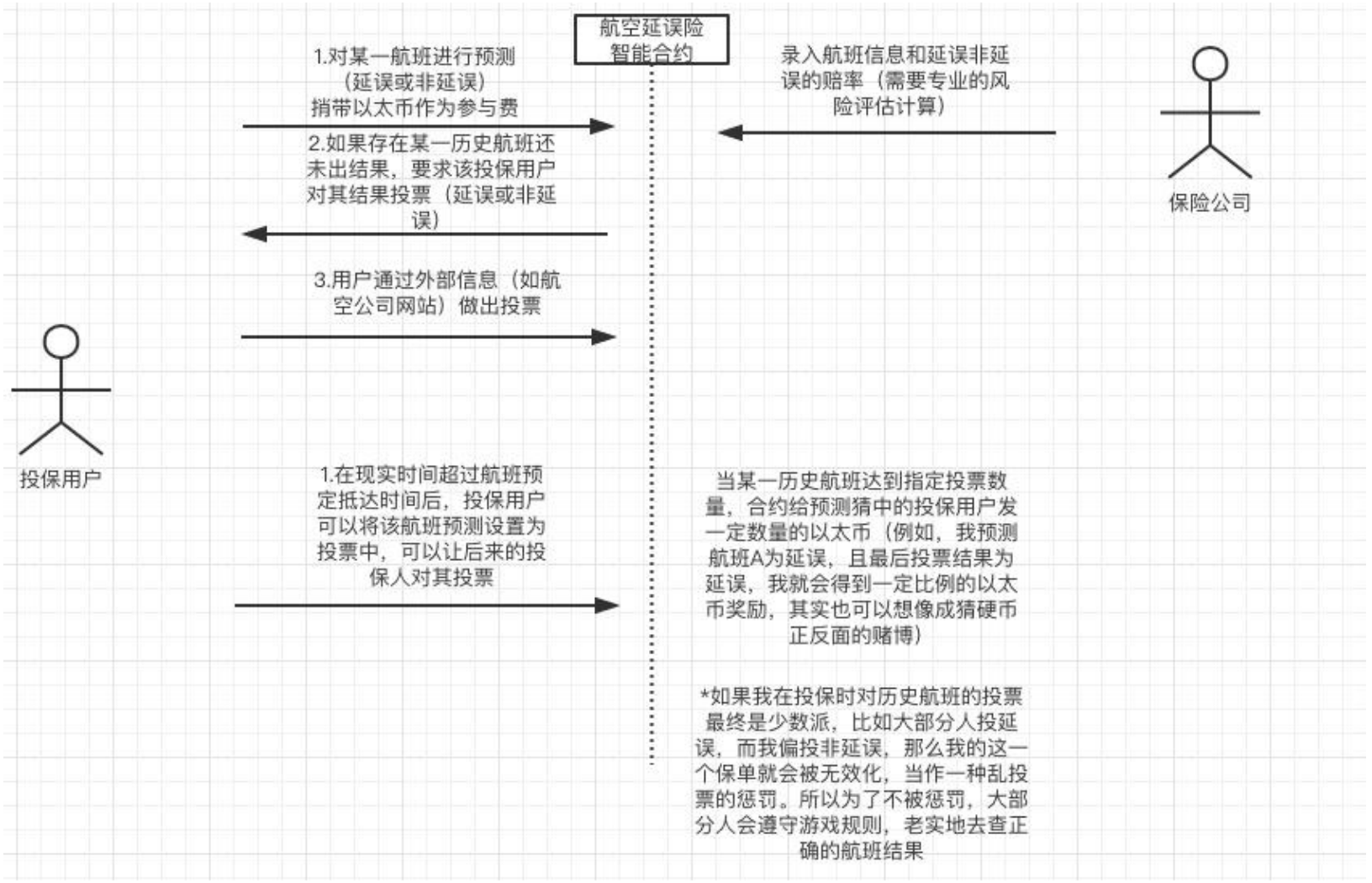
区块链 — 主要特征

- 不可篡改
  - 批量交易一旦达成共识，就会作为一个区块添加到区块链中，会被永久的存储起来
  - 单个甚至多个节点对交易数据的修改是无效的，除非能控制整个网络中超过51%的节点同时修改，这几乎不可能发生。
- 可追溯
  - 区块链中的每一笔交易都通过密码学方法与相邻两个区块串联，因此可以追溯到任何一笔交易的前世今生。
- 共识
  - 采用协商一致的规范和协议（如POW）来保证数据质量，从而共同维护整个区块链。

区块链 — 适用场景

- 多方参与
- 互不信任
- 自治，设计规则鼓励遵守规则
- 链下数据可信的保证，ORACLE预言机

区块链 — 某航空公司延误险理赔

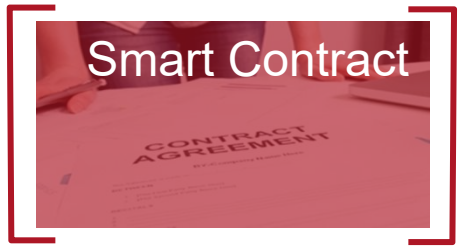


区块链 — 主要分类

- **公有链，Public blockchains**
  - 公有链无官方组织及管理机构，无中心服务器，参与的节点按照系统规则自由接入网络、不受控制，节点间基于共识机制开展工作。
- **私有链，Private blockchains**
  - 建立在某个企业内部，系统的运作规则根据企业要求进行设定，修改甚至是读取权限仅限于少数节点，同时仍保留着区块链的真实性和部分去中心化的特性。
- **联盟链**
  - 联盟链由若干机构联合发起，介于公有链和私有链之间，兼具部分去中心化的特性。联盟链管理员可以根据业务要求限制某些用户的读取权限，并限制某些节点参与共识计算。

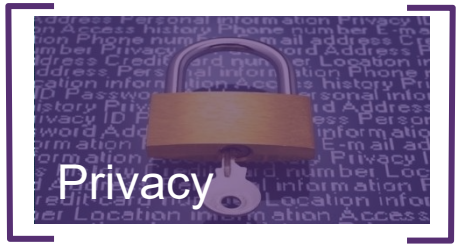
区块链 — 业务

Append-only distributed **system of record** shared across business network



Business terms embedded in transaction database & executed with transactions

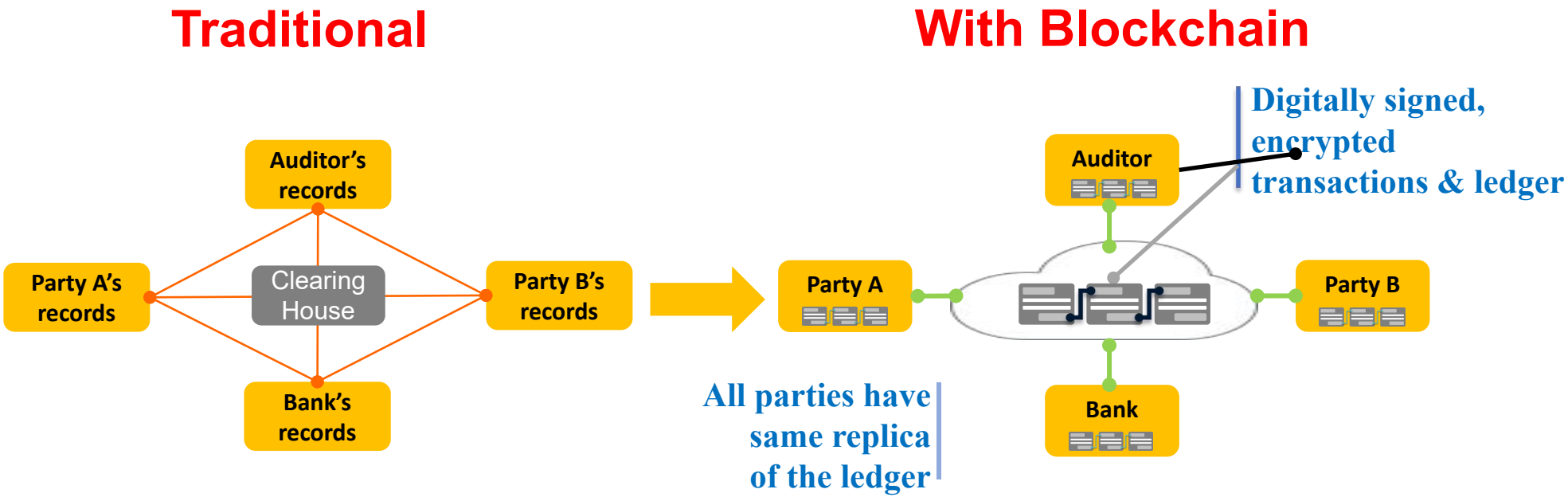
Ensuring appropriate visibility; transactions are secure, authenticated & verifiable



All parties agree to network verified transaction

… Broader participation, lower cost, increased efficiency

区块链 –商业模式改变



… Inefficient, expensive, vulnerable

… Consensus, provenance, immutability, finality

区块链 2.0

区块链 2.0

- 具有智能合约（smart contract）功能的公共区块链平台

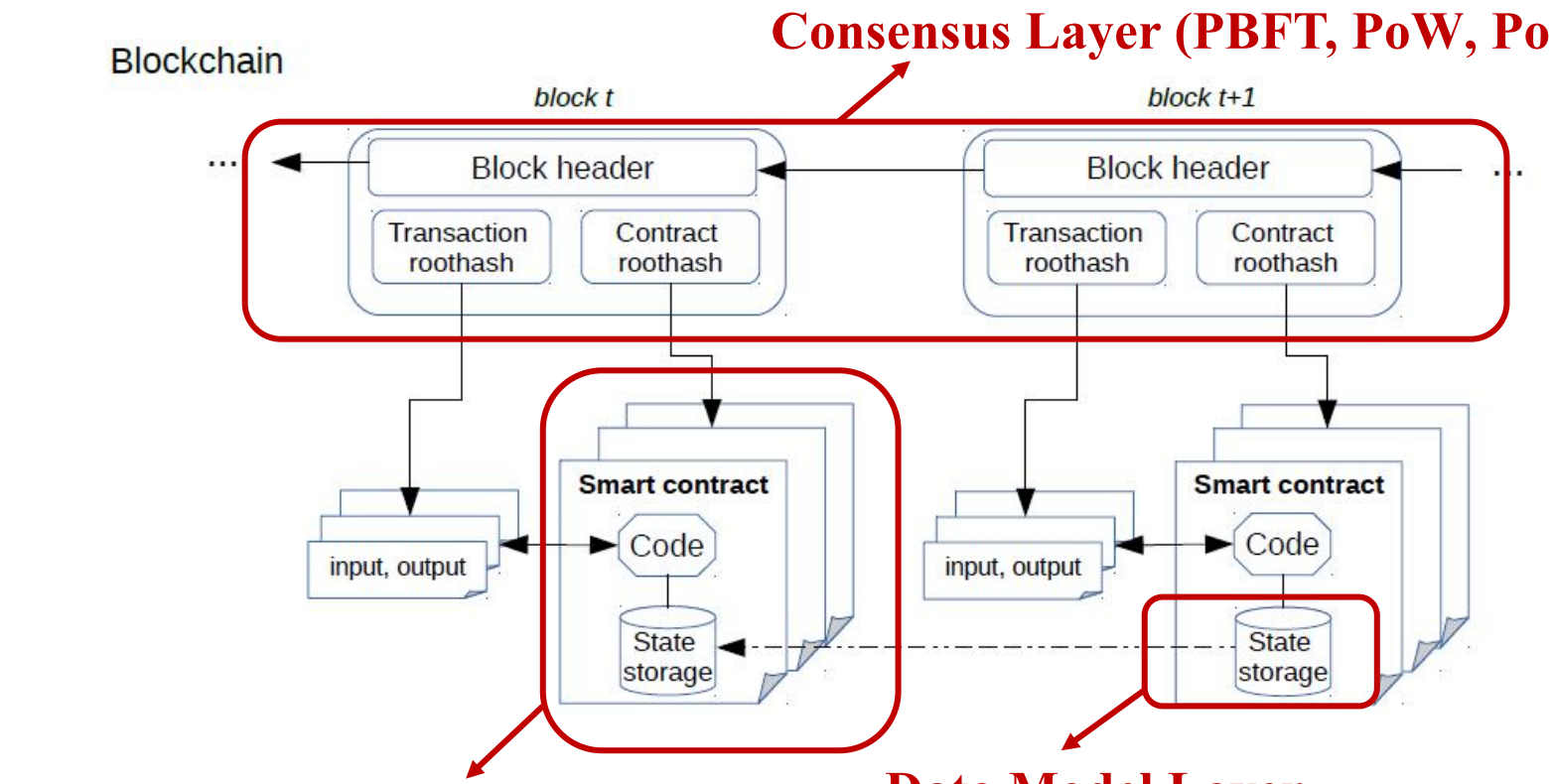
Transaction(State + Code)

Property	Value
A456Sget498038DG	SD203948A
Code	Function….

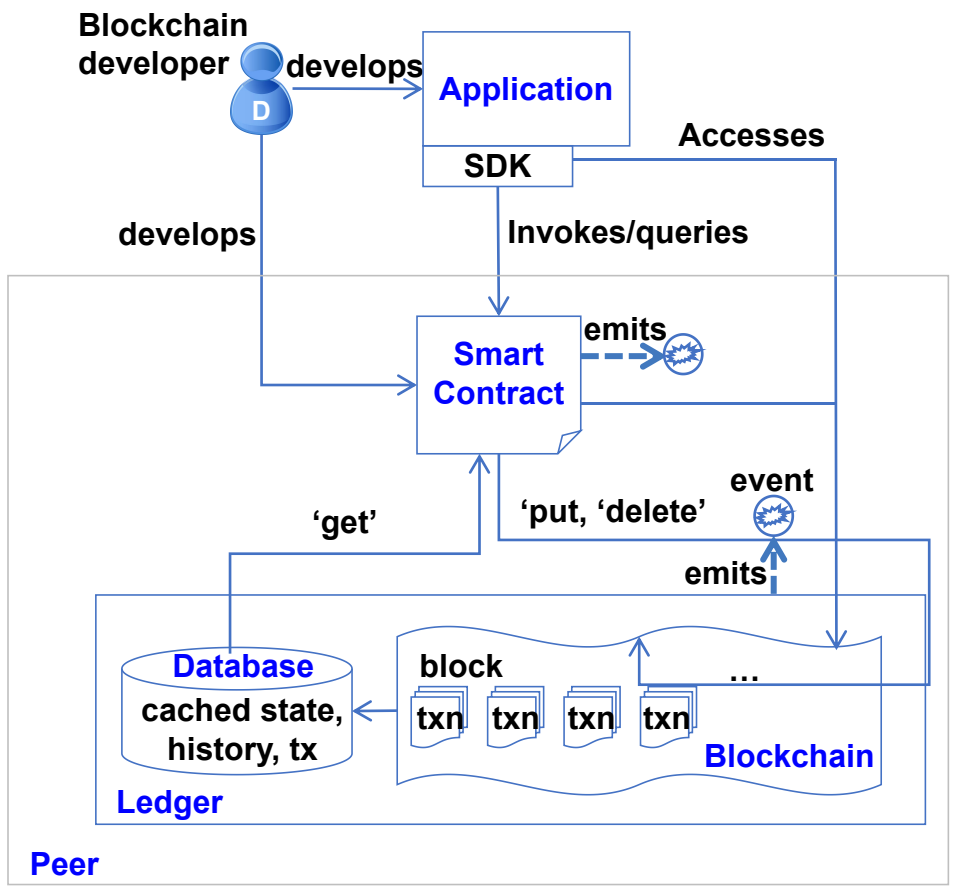
区块链 2.0 — 智能合约

- 智能合约（Smart Contract）
  - 智能合约程序不只是一个可以自动执行的计算机程序，其自己就是一个系统参与者。
  - 它可以对接收到的信息进行回应，可以接收和储存价值，还可以向外发送信息和价值。

区块链 2.0 — 软件架构


Consensus Layer (PBFT, PoW, PoS, etc.)
Blockchain
block t
block t+1
Block header
Transaction roothash
Contract roothash
Block header
Transaction roothash
Contract roothash
Smart contract
Code
input, output
State storage
Smart contract
Code
input, output
State storage
Smart Contract Execution Engine (Virtual Machine, Docker, etc.)
Data Model Layer (LevelDB, RocksDB, etc.)

区块链 2.0 — 业务流程



- **Developers create application and smart contracts (chaincodes)**
  - Chaincodes are deployed on the network and control the state of the ledger
  - Application handles user interface and submits transactions to the network which call chaincodes
- **Network emits events on block of transactions allowing applications to integrate with other systems**

从分布式数据库的角度解读区块链

去中心化的分布式数据

- 全复制——数据分布
- 拜占庭问题下的数据一致——POW/PBFT
- P2P网络——gossip协议

如何保证不可篡改

- 哈希链的数据结构
  - 数字摘要
  - 链式存储
- 数据的全复制分布

挑战：拜占庭将军问题

- 拜占庭问题讨论的是允许存在少数节点作恶（消息可能被伪造）场景下的一致性达成问题。
- 是Leslie Lamport 在1982年提出用来解释一致性问题的一个虚构模型
  - 拜占庭是古东罗马帝国的首都，由于地域宽广，守卫边境的多个将军(系统中的多个节点)需要通过信使来传递消息，以达成某些一致的决定。
  - 但由于将军中可能存在叛徒(系统中节点出错)，这些叛徒将努力向不同的将军发送不同的消息，以试图干扰一致性的达成。
- POW是个选主的过程

共识算法

- 基于工作量证明51%
  - 算力浪费、不确定性分叉
- 基于协议PBFT最多有1/3的非诚实节点
  - 网络开销太大，存在扩展性不好
- 基于概率的协议 **Algorand**

03 应用情况

- 区块链目前的商业应用情况

应用情况

- 《中国区块链技术和应用发展白皮书（2016）》
  - 金融服务
  - 供应链管理
  - 智能制造
  - 社会公益
  - 文化娱乐
  - 教育就业

应用情况 — 金融服务应用进展

- 金融服务是区块链最早的应用领域之一，也是区块链应用数量最多、普及程度最高的领域之一
  - 防金融欺诈、资产托管交易、金融审计、跨境支付、对账与清结算、供应链金融以及保险理赔等；
  - 典型案例：基于区块链的机构间对账平台、差异账检查系统，以及通过区块链技术改造的跨境直联清算业务系统等。

应用情况 — 供应链管理应用进展

- 供应链核心企业、商业银行、电商平台等相关力量不断加强区块链在供应链管理领域的应用探索。
  - 防伪溯源：京东、蚂蚁金服、众安科技等科技企业；
  - 供应链金融：央行数字货币研究所、央行深圳市中心支行推动“粤港澳大湾区贸易金融区块链平台”，万向区块链、平安壹账通、京东、腾讯等众多企业；
  - 防伪溯源和物流等领域：更注重与物联网、人工智能等技术的融合发展。

应用情况 — 区块链与智慧城市

- 区块链在建设智慧城市中的应用涵盖智慧园区、智慧物联网、智慧资产、智慧交通、能源电力、电子政务、法律应用等广阔领域。
  - 采用分布式点对点的网络结构，可以使设备之间保持共识，实现点对点传输数据，减少甚至无需与中心服务器的数据库进行验证，避免对中心化设施的依赖；
  - 利用区块链数据防篡改、可追溯的特性，可以帮助打通政府各部门数据孤岛，为公众提供更加可信和有价值的服务。

应用情况 — 区块链与公共服务

- **部分地方政府大力推进“区块链+政务”服务，以满足公共服务在信息共享、权限控制和隐私保护等方面的高要求。**
  - 鉴证确权：将公民财产、数字版权相关的所有权证明存储在区块链账本中，大大优化权益登记和转让流程，减少产权交易过程中的欺诈行为；
  - 身份验证：将身份证、护照、驾照、出生证明等存储在区块链账本中，实现无需任何物理签名即可在线处理繁琐的流程，并能实时控制文件的使用权限；
  - 信息共享：区块链技术用于机构内部以及机构之间的信息共享和实时同步，能有效解决各行政部门间协同工作中流程繁琐、信息孤立等问题。

04 技术发展

- 区块链技术的发展情况

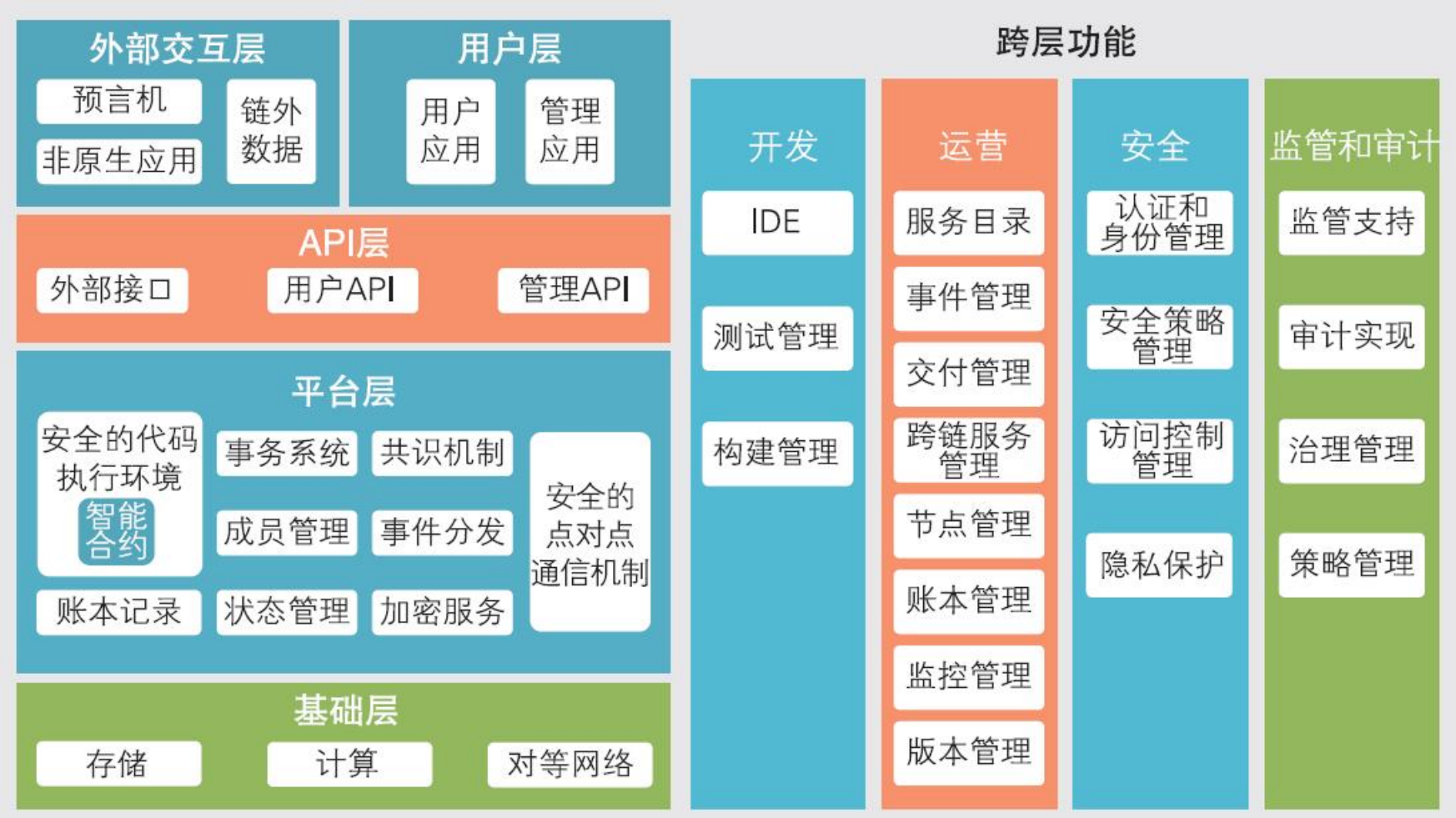
技术发展

- 区块链系统架构
- 安全技术
- 隐私保护技术
- 跨链技术
- 分片技术
- 数据存储
- 共识机制
- 智能合约

技术发展 — 区块链系统架构

- 区块链1.0 – 数字代币为典型特征
- 区块链2.0 – 以智能合约为典型特征
- 区块链3.0 – 未形成共识

技术发展 — 区块链系统架构



技术发展 — 安全技术

- 主要包括：数字摘要算法、数字签名和加密算法。
- 经典算法：
  - SHA256、SM3等主流的数字摘要算法
  - RSA、ECDSA、SM2等常用的数字签名算法
  - AES、SM4等对称加密算法
  - RSA、SM2等非对称加密算法。

技术发展 — 安全技术

- 挑战
  - 量子计算的技术演进对现有的密码学安全机制的巨大影响；
  - 根据Shor算法，经典非对称算法（基于大数分解、离散对数等算法，如RSA、ECDSA和SM2等）可以被稳定、可用的量子计算机攻破；
  - 密码学家正积极探索能够抵抗量子计算机攻击的密码机制，如基于格的密码机制、基于纠错码的密码机制、多变量密码机制等。

技术发展 — 隐私保护技术

- 目标：“身份的隐私性”和“数据的机密性”
  - “身份的隐私性”主要是对区块链参与者身份的保护；
  - “数据的机密性”主要是对记录内容、合约逻辑等数据的保护。

技术发展 — 隐私保护技术

- 相关技术：环签名、同态加密、零知识证明和安全多方计算等。
  - 环签名允许一个成员代表一个群组进行签名而不泄漏签名者信息，可以实现签名者完全匿名。
  - 同态加密除了具有一般的加密操作之外，还能实现直接对密文的计算操作，通常分为加法同态、乘法同态、全同态等类型。
  - 零知识证明是指一方（证明者）向另一方（验证者）证明某个事实的论断，同时不透露该事实的其他信息的方法。
  - 安全多方计算能够在保证输入数据隐私的前提下，为缺乏信任的参与方提供协同计算功能。

技术发展 — 跨链技术

- 跨链泛指两个或多个不同区块链上资产和状态通过特定的可信机制互相转移、传递和交换的技术。
- 跨链分为同构链的跨链和异构链的跨链。目前主流的跨链技术有：公证人机制、侧链/中继、哈希锁定、分布式私钥控制等。
- 未来跨链技术的重点发展方向包括：加快交易速度，减轻主链负担，发展多链并行处理计算，支持海量交易，提升安全性和加强隐私保护等。

- 比特币的闪电网络？

技术发展 — 分片技术

- 分片技术本身是一种传统数据库技术，此前主要用于将大型数据库分成更小、更快、更容易管理的数据碎片。在区块链中，可将区块链网络分成很多更小的部分，即进行“分片”处理。
- 分片技术主要有网络分片、交易分片和状态分片三类
  - 网络分片是利用随机函数随机抽取节点形成分片，支持更海量的共识节点。
  - 交易分片分为同账本分片和跨账本分片，主要思想是确保双花交易在相同的分片中或在跨分片通信后得到验证。
  - 状态分片关键是将整个存储区分开，让不同的碎片存储不同的部分，每个节点只负责托管自身的分片数据，而不是存储完整的区块链状态。

技术发展 — 数据存储

- 区块的存储由链式结构发展为有向无环图（DAG），DAG 区块链在并行性、可扩展性上有较大改善。
- 链外数据的存储，除了传统集中的数据中心存储、云存储以外，产生了新的互联网点对点文件系统
  - 如融合Git、自证明文件系统（SFS）、BitTorrent 和DHT 等技术的星际文件系统（IPFS）

技术发展 — 共识机制

- 主要集中在：提高系统吞吐量及降低网络带宽
- **CFT**
  - Paxos, Raft
- **BFT**
  - POW系列：POW、POS、DPOS
  - PBFT、DBFT
  - DAG

技术发展 — 智能合约

- 主要方面：形式化验证框架与通用型合约编程语言
- 安全性是智能合约的关键性问题
- 智能合约编程语言逐渐从脚本型语言向通用型语言演变
- 智能合约的执行逐渐从显式调用执行向由链上触发器（如预言机机制）自动触发执行的方向发展

05 典型区块链系统

- 代表性的区块链

典型区块链系统

- 公有链平台：比特币、以太坊
- 联盟链平台：**Hyperledger Fabric，Corda，Quorum**
- 国内平台：
  - 分布式应用账本（DAppLedger）开源社区重点孵化的开源项目有BCOS和Annchain等
  - BCOS（微众银行），Annchain（众安科技）和蚂蚁链（蚂蚁集团）属于联盟链

谢谢！

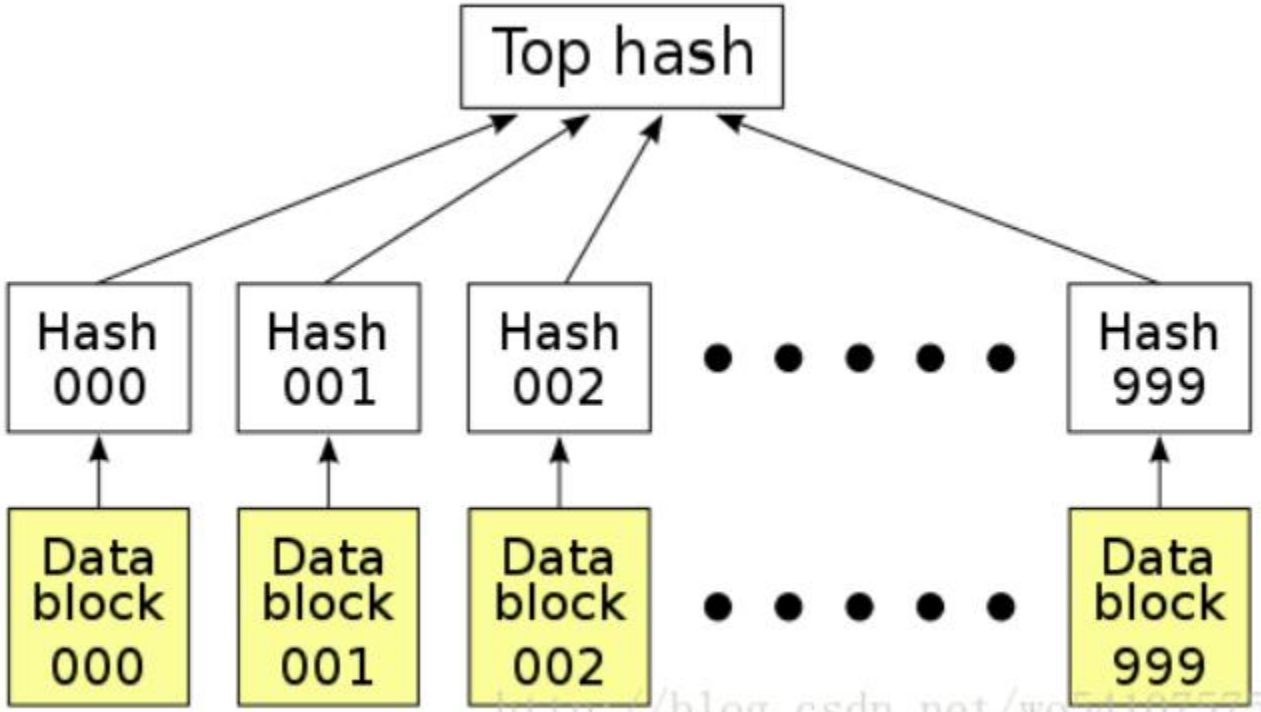
双花

- 如果先发起一笔交易，在它被确认前，进行矛盾的第二笔交易，那么在记账时，这些交易会被拒绝。
- 如果是A分支被认可（B也一样），相应交易确认，拿到商品之后，立刻自己变身矿工，争取到连续两次记账权，然后在B分支上连加两个**block**。

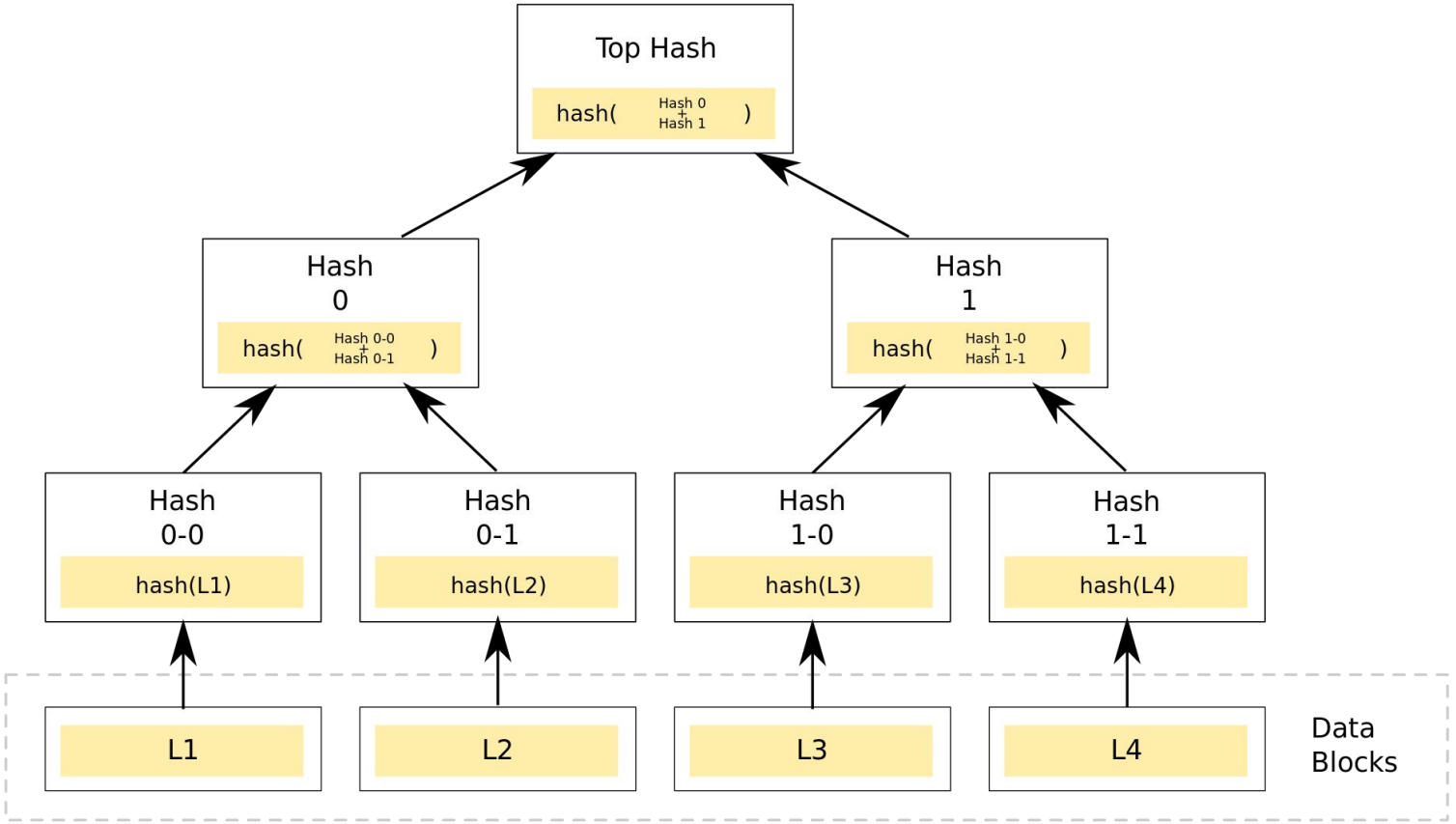
女巫攻击（Sybil Attack）

- 节点具有多个身份标识，通过控制系统的大部分节点来消弱冗余备份的作用
- 恶意节点伪装成多重身份
  - 比如冒充多个身份投票
- 数据冗余备份
- 比特币系统中如何避免这一点？

Hash List



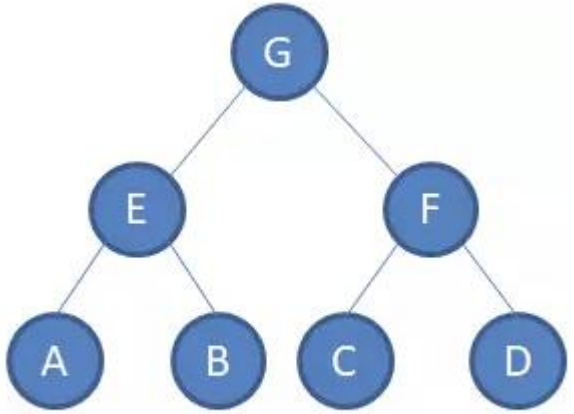
区块链 - Merkle Tree



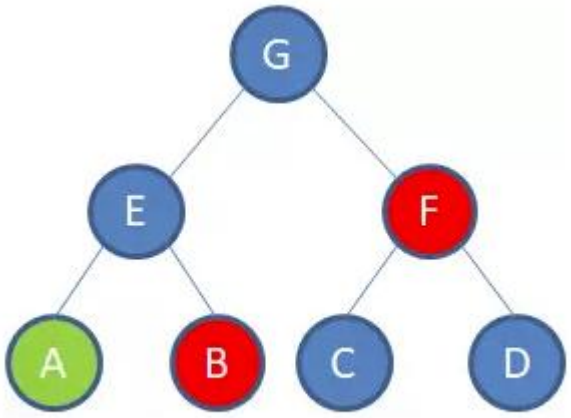
在P2P网络下载中的应用

- **Hash List**
- **Merkle Tree**
- 以上两种数据结构在确认数据块是否有损坏时有什么差别？

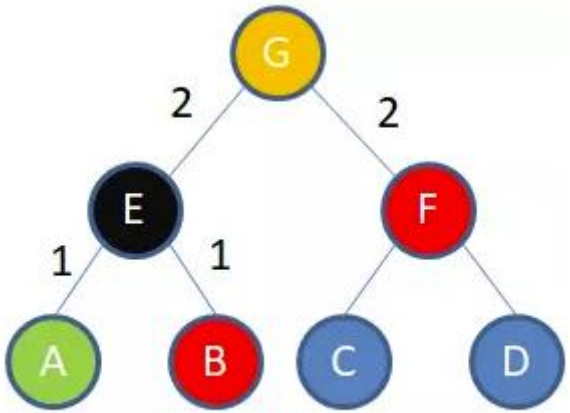
比特币中的默克尔验证



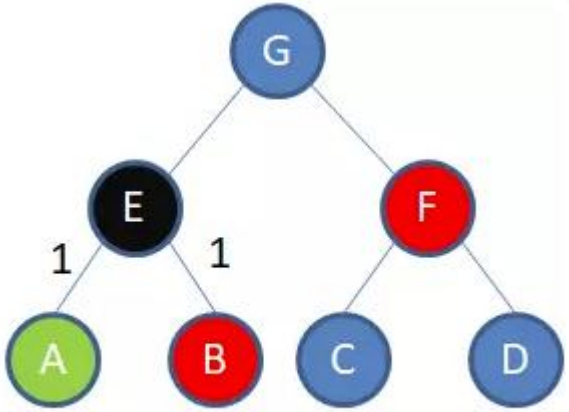
默克尔树对应的树形图



红色节点对应的就是节点A的默克尔树分支



验证第二步是得到节点G



验证第一步是得到节点E

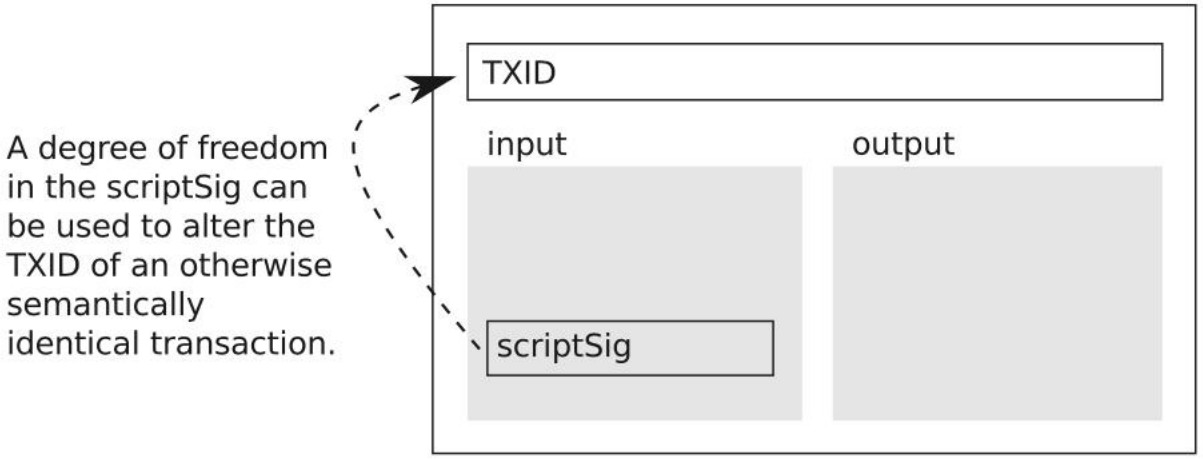
交易索引0对应的默克尔树分支

SPV节点的验证过程

- 0. 从网络上获取并保存最长链的所有block header至本地；
- 1. 计算该交易的hash值tx_hash；
- 2. 定位到包含该tx_hash所在的区块，验证block header是否包含在已知的最长链中；
- 3. 从区块中获取构建merkle tree所需的hash值；
- 4. 根据这些hash值计算merkle_root_hash；
- 5. 若计算结果与block header中的merkle_root_hash相等，则交易真实存在；
- 6. 根据该block header所处的位置，确定该交易已经得到多少个确认。

Transaction Malleability 可锻性攻击

- **A bug in the Bitcoin protocol**
  - Change the TXID without invalidating the transaction
  - TXID: a hash of the transaction data, including the redeem scripts
  - Signature in a script does not cover the redeem script



Transaction Malleability 可锻性攻击

- **ECDSA(elliptic curve digital signature algorithm， 椭圆曲线数字签名算法）生成两个大整数r和s并组合起来作为签名。**
- **而r和BN-s也同样可以作为签名来验证交易(BN＝0xFFFFFFFF FFFFFFFF FFFFFFFE BAAEDCE6 AF48A03B BFD25E8C D0364141)。**

Transaction Malleability 可锻性攻击

- **Result: the signature remains valid, but the TXID changes**
- **Transaction malleability attack**
  - the attacker withdraws coins from an exchange
  - as soon as the attacker receives the respective withdrawing transaction issued by the exchange, he rebroadcasts the altered version of this transaction with a different TXID

延展性攻击，或者锻造性攻击

- 延展性攻击者侦听比特币P2P网络中的交易，利用交易签名算法的特征修改原交易中的input签名，生成拥有一样input和output的新交易，然后广播到网络中形成双花。
- 受影响更大的是交易所，银行之类的离线钱包网站

主要问题是交易所使用TXID进行验证

- 首先需要有足够多的比特币矿机接入网络，以增加伪造的请求被优先处理的可能性。
- 攻击者在第三方交易平台提交一个提款请求并获得一个交易ID。
- 根据交易信息伪造一个签名同时生成一个完全不同的交易ID，并将伪造的请求发出。
- 若伪造的交易被优先处理，则原始交易失败。
- 我们可再次提交提现请求，第三方交易平台确认之前的交易失败后会再次发送提现交易，至此攻击成功。

Pooled Mining

- **Solo mining**
  - A significant reward
  - Only very seldom
- **Group into mining pools**
  - Search parts of the nonce space for a valid nonce
  - Profit is shared

Pooled Mining

- **security**
  - Approach the critical threshold of 51% of the network's hash rate
  - Coalition
  - Solutions
    - Decentralize the mining pool(p2pool)
    - Non-outsourceable proof of work

Pooled Mining

- **P2Pool**
  - A decentralized mining pool
  - Shared chain
    - 30s a new block
  - Distribute reward by shares
  - Additional complexity and significant resource consumption

私自挖矿

- **Selfish mining**
  - Miners keep their discovery private and establish a private chain
  - If the public chain approaches the length of the private chain, the rogue miner broadcasts his chain to catch up
  - key idea: let honest miners waste their power by mining on the public chain

谢谢！