人才稀缺的区块链，程序员转型入门必看这四项技能

2018-02-27 卿苏德 CSDN

点击上方“CSDN”，选择“置顶公众号”

关键时刻，第一时间送达！

作者 | 卿苏德

区块链（BlockChain），是区块（Block）和链（Chain）的直译，其数据结构如图 1 所示，即每个区块保存规定时间段内的数据记录，并通过密码学的方式，构建一条安全可信的链条，形成一个不可篡改、全员共有的分布式账本。

比特币的区块分为区块头和区块体两部分。区块头的大小为 80 字节，包括 4 字节的版本号、32 字节（256 位）的上一区块哈希值、32 字节的 Merkle 根节点、4 字节的时间戳、4 字节的难度值和 4 字节的随机数。区块体包含 10 分钟内选定的交易记录，第一笔交易（coinbase 交易）是用于奖励矿工比特币的特殊交易，由矿工自己添加进区块。

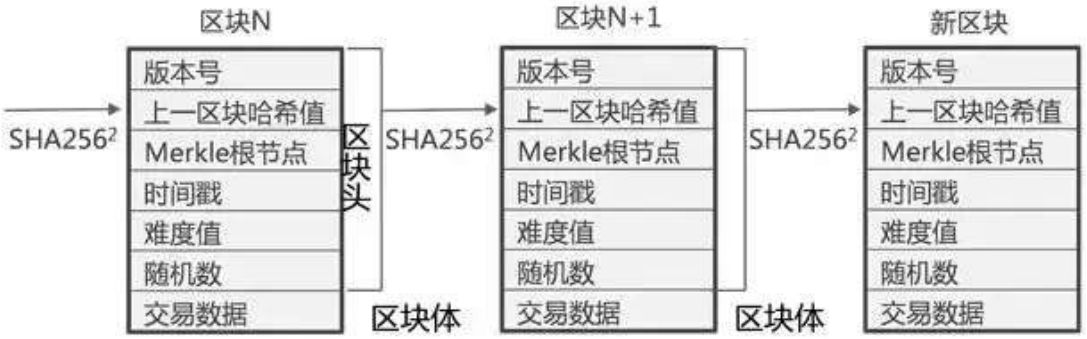

图 1　区块链的数据结构示意图

基本概念

区块链是很多现有技术交叉融合在一起的集成创新。因此，要了解区块链，首先要了解区块链到底集成了哪些技术。

P2P 网络

如图 2 所示，P2P（Peer-to-Peer）网络是一种端到端的网络。P2P 网络分为结构化（例如基于 Chord 的 P2P 网络）和非结构化的 P2P 网络（例如 Gnutella）。比特币的区块链采用的是非结构化 P2P 网络，整个网络没有中心化的硬件或管理机构，任一节点既是服务端，也是客户端。任何节点只要安装相应的客户端软件，就能接入 P2P 网络（例如 BT 软件），参与区块链的记录和验证，不超过 1/3 节点的损坏、退出甚至被植入恶意代码，都不会影响整个系统的运作。

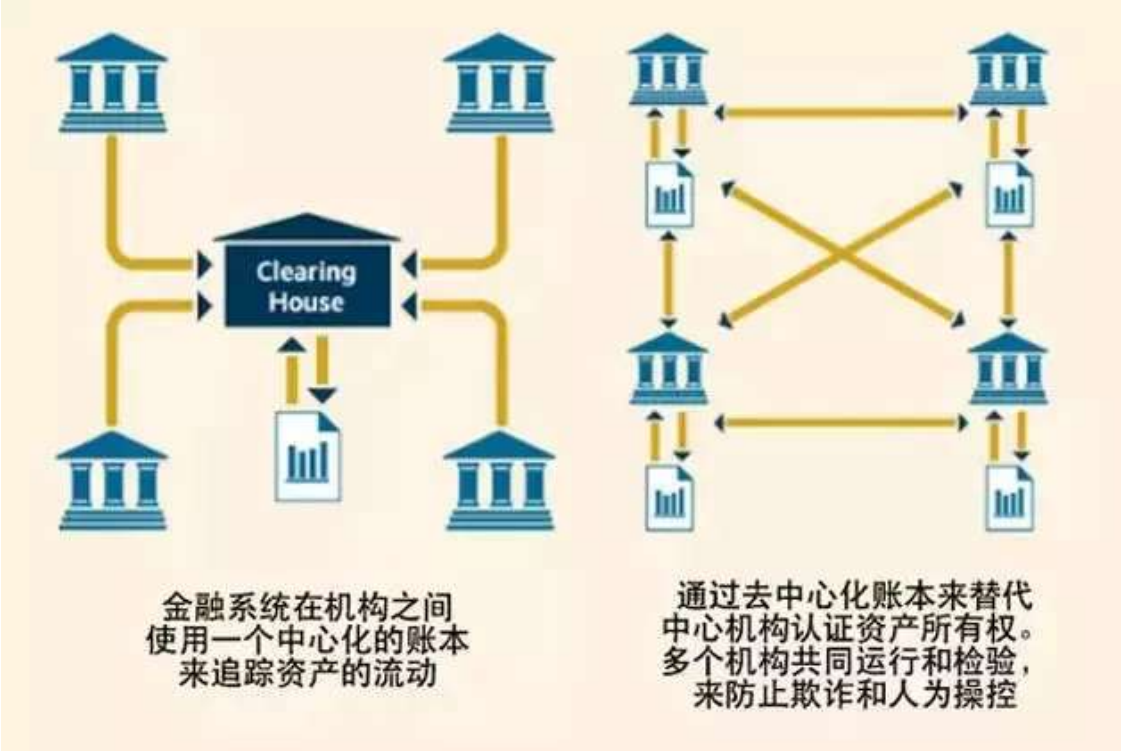

图 2　传统中心化系统和 P2P 网络的拓扑对比图

加密算法和数字签名

加密技术分为对称、非对称和哈希（Hash）加密。对称加密是指用同样的密钥来进行加密和解密，非对称加密是指用一个密钥对来进行加密和解密，哈希加密主要是通过对数据进行哈希运算，用固定的哈希结果值验证信息是否被篡改。

- **非对称加密**

在非对称加密技术中，对外公开、分发出去的密钥叫做公钥，不能公开、自己留存的密钥叫做私钥。公钥加密的，对应的私钥才能解密。反之亦然。如图 3 所示。

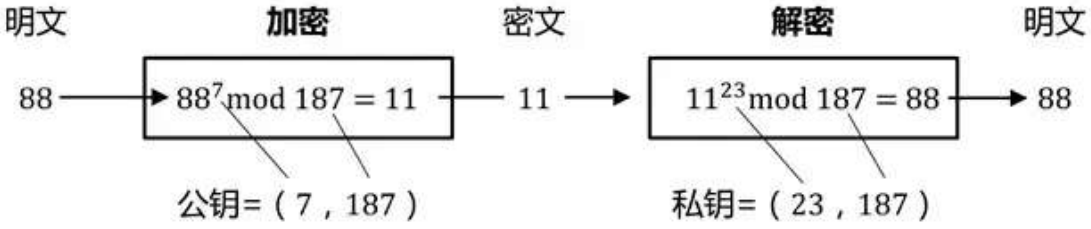

图 3　非对称加密 RSA 算法的简化示例图

非对称加密算法有 RSA、DSA 和 ECC 等种类，区块链使用的是基于椭圆曲线加密技术的数字签名（ECDSA），具体实现是

secp256k1。ECDSA 相当于是 DSA 和非对称加密 ECC 的结合。相比 RSA 算法，ECDSA 具有计算量小、存储空间小、带宽要求低等特点。

- **数字签名**

基于数字签名的通信机制工作原理，如图 4 所示，发送报文时，发送方用一个哈希函数从报文文本中生成文件摘要，然后用自己的私钥对摘要进行加密，加密后的摘要将作为报文的数字签名和报文一起发送给接收方。接收方首先用与发送方一样的哈希函数从接收到的原始报文中计算出报文摘要，接着再用发送方的公钥来对报文附加的数字签名进行解密，如果得到的明文相同，那么接收方就能确认传输的文件并未受到篡改，是安全可信的。

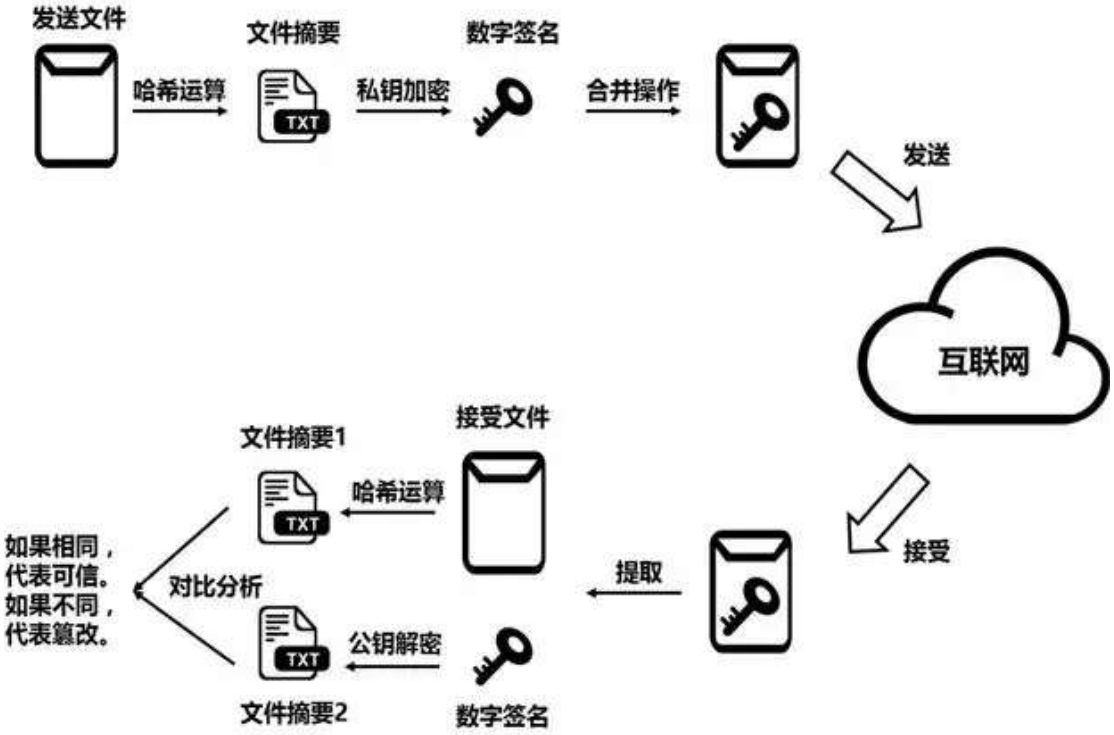

图 4 数字签名的流程示意图

- **哈希加密**

安全哈希算法（Secure Hash Algorithm，SHA）是由美国国家安全局研发，由美国国家标准与技术研究院（NIST）发布的一系列密码哈希函数，包括 SHA-0、SHA-1、SHA-2 和 SHA-3 等系列。比特币的区块链使用的是 SHA-256 哈希加密算法，于 2001 年发布，属于 SHA-2 分支。由于 SHA256 伪随机性的特点，只要是相

同的数据输入，一定会得到相同的结果，如果输入数据稍有变化，将得到一个千差万别的结果，如图 5 所示。SHA256 还是一个单向不可逆的算法，即根据一个输入数算 SHA256 的结果很容易，但根据 SHA256 的结果反算输入数几乎是不可能。除此之外，比特币还使用 ripemd160 算法来生成比特币钱包的地址。

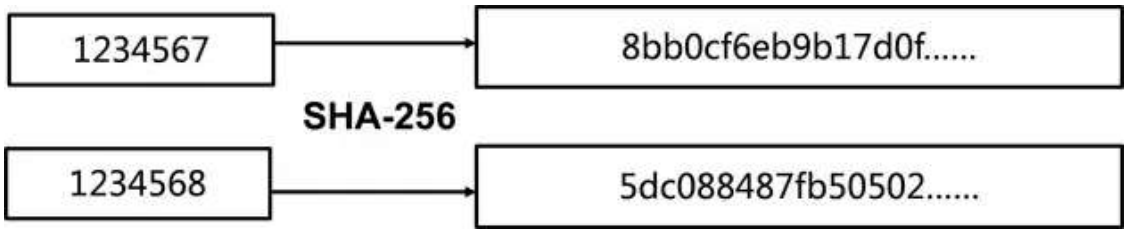

图 5　哈希加密的示意图

梅克尔树

梅克尔（Merkle）树是区块链的基本组成部分。如果没有梅克尔树，区块链也是可以运转，但是要在区块头里包含所有交易记录，扩展性方面存在很大挑战。如图 6 所示，区块链中的每个区块，由区块头和区块体构成，区块头中含有一个 Merkle 根节点的字段，通过对区块体中所有交易记录，以二叉树的形式迭代地两两拼接、进行哈希操作，可以得到一个最终的哈希值，我们称之为 Merkle 根哈希。Merkle 根哈希相当于是对区块中所有交易记录进行了一个快照，区块中交易记录的任意改动都可以通过比较 Merkle 根哈希而很容易地察觉。Merkle 根哈希主要用于简单支付验证（SPV），在验证某个交易是否在区块中时，也能极大地减少网络传输成本。

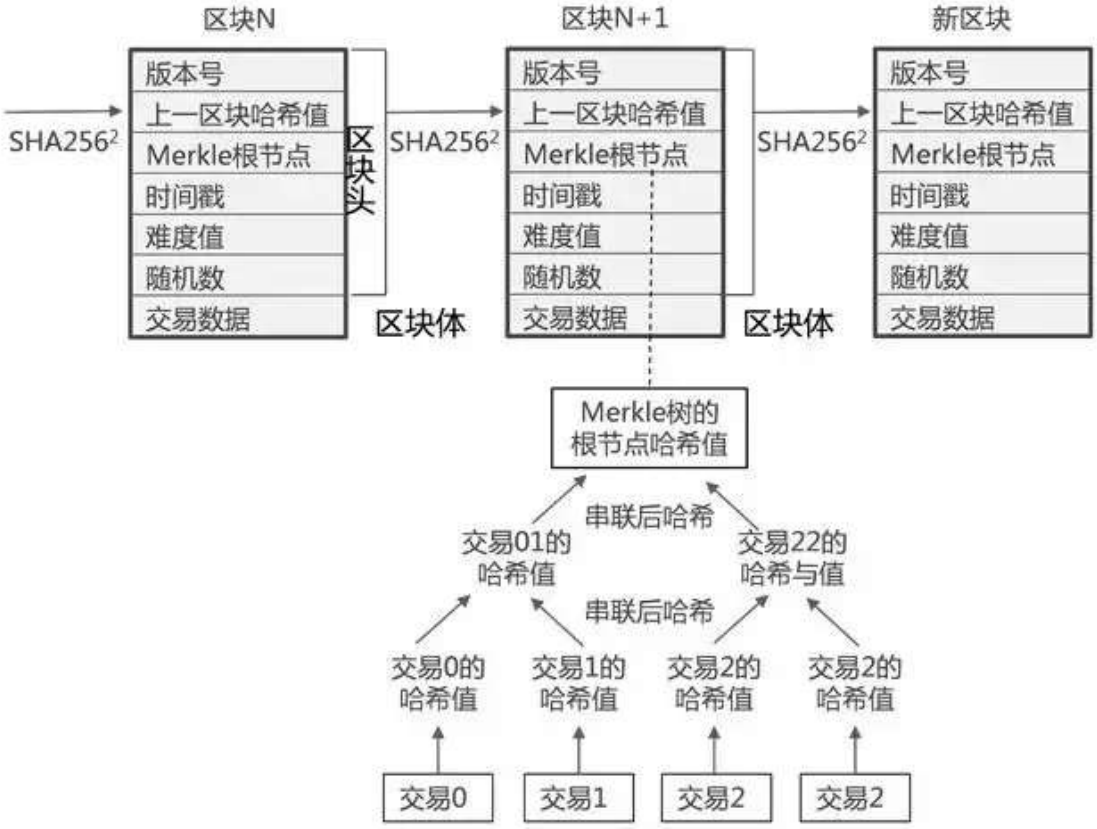

图 6　Merkle 树示意图

工作量证明机制

工作量证明机制，简单地说，就是一种共识机制，用来确认你是否做过一定量工作的证明。比特币的区块链主要是依托计算数学难题来衡量工作量。每个区块，当选定一定数量的交易记录之后，填充版本号、时间戳、难度值，生成相应的 Merkle 根哈希。很容易看到，这些数值在选定交易记录以后，都是确定的，唯一能够改变的就只有随机数（Nonce）这个值。如图 7 所示，系统根据难度值，要求计算整个区块头的两次 SHA256 算法，得到的哈希结果要小于一个阈值。根据前面描述的 SHA256 算法的伪随机性，只有通过不断地尝试和枚举，才能找到相应的随机数，证明自己的工作量。

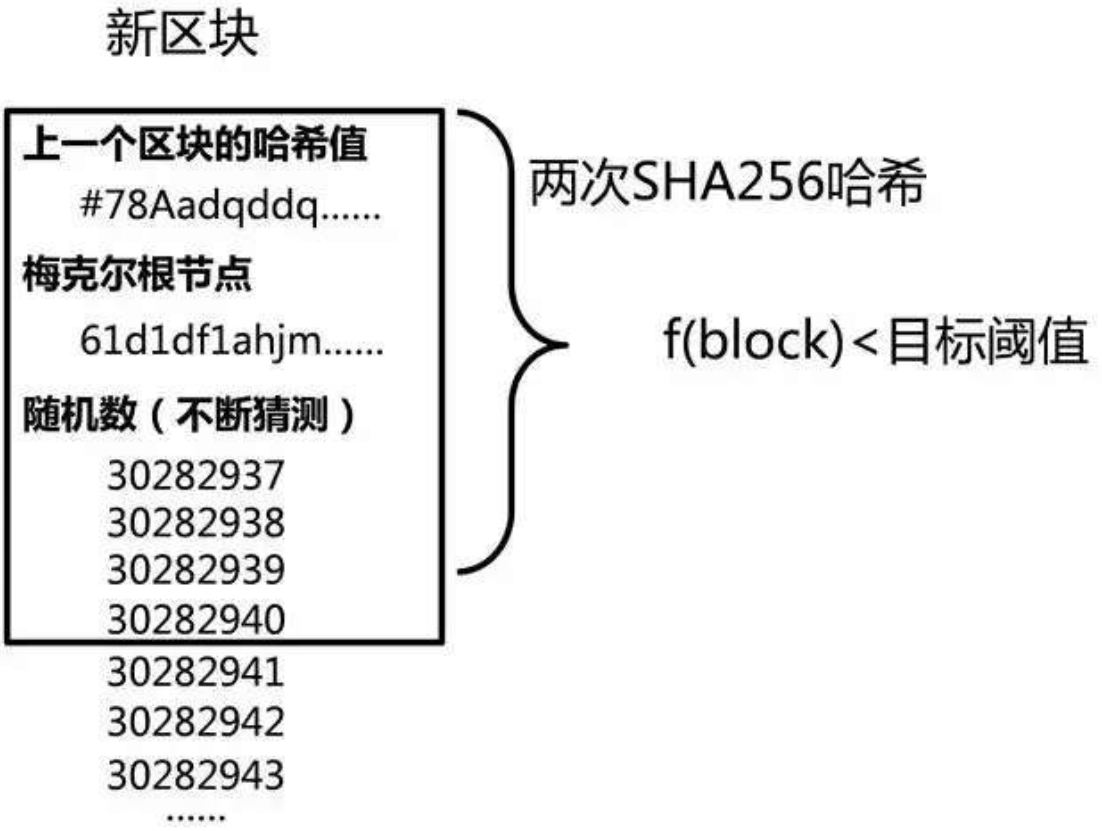

图 7　工作量证明机制示意图

除了工作量证明机制（PoW）这类共识机制之外，还有股权证明机制（PoS）、授权股权证明机制（DPoS）、拜占庭容错机制（BFT）、实用拜占庭容错机制（PBFT）这些在不可信环境下的共识机制以及要求在可信环境下的共识机制，例如 PaxOS 和 Raft。表 1 是做了简单的对比。

表 1　共识机制的简单对比表

	zookeeper	etcd	bitcoin	eris	hyperledger
共识机制	Paxos	Raft	PoW	BFT	PBFT
一致性	强一致性	强一致性	弱一致性	弱一致性	弱一致性
网络组织	主从	主从	对等	对等	对等
数据库	适配	自身	levelDB	levelDB	rocksDB
允许失败的节点数	<1/2	<1/2	<1/3	<1/3	<1/3
恶意节点	不允许	不允许	允许	允许	允许
虚拟机	无	无	无	有	有
需要代币	无	无	有	有	无

运行机制

接入网络和验证

节点通过安装相应的软件（例如比特币核心），接入区块链。节点启动以后，主要是在 P2P 网络上发现邻居节点、链接邻居节点、传递 P2P 消息和下载区块链验证。节点可以选择下载全量的区块链进行验证，或者是只下载区块头，通过 Merkle 树节点来进行简单支付验证（SPV）。

钱包软件可以分为移动钱包、桌面钱包、互联网钱包和纸钱包，都支持保存用户的私钥，钱包也可以根据私钥是否是种子产生的，而分为决定性钱包和非决定性钱包，关键区别在于私钥的备份和易恢复性。

区块链的存储和接受

比特币的区块链使用 Berkeley DB（文件数据库）作为钱包数据库，使用 LevelDB（键值数据库）存储区块的索引和 UTXO（Unspent Transaction Output，未开销的比特币交易输出）。节点在启动的时候，将整个区块链的索引从 LevelDB 加载入内存。当收到一个新区块时，节点对新区块中的所有交易进行检测，验证交易格式、交易大小、交易签名、UTXO 是否匹配、交易签名、脚本合规等方面。

如果验证成功，检查上一区块头与链头区块哈希值是否一致，如果是一致，则更新 UTXO 数据库和回滚交易数据库，如果不是，则将该区块放在孤儿区块池中 。当节点发现网络中存在另一条更长的区块链时，就需要断开现有的区块并对区块链进行重组。如果验证不成功，会抛弃该区块，继续等待新区块的到来（矿工会继续计算新区块的数学难题）。

- **区块链的工作量证明计算机制**

“矿工”角色的节点一直收集网络中广播的交易记录，并致力于计算新区块的数学难题，即工作量证明。如果其他节点发来的新区块验证成功，节点除了更新 UTXO 数据库和回滚交易数据库，节点会立即开始下一个新区块的计算。新区块的构建优先选取交易内存池

中优先级高的交易记录。优先级的计算方式为：

如果自己的工作量证明计算成功，节点会第一时间将这个区块广播至整个网络中，其他节点收到该新区块，如上所述，会进行相应的验证和存储。

整个区块链的运转机制如图 8 所示。

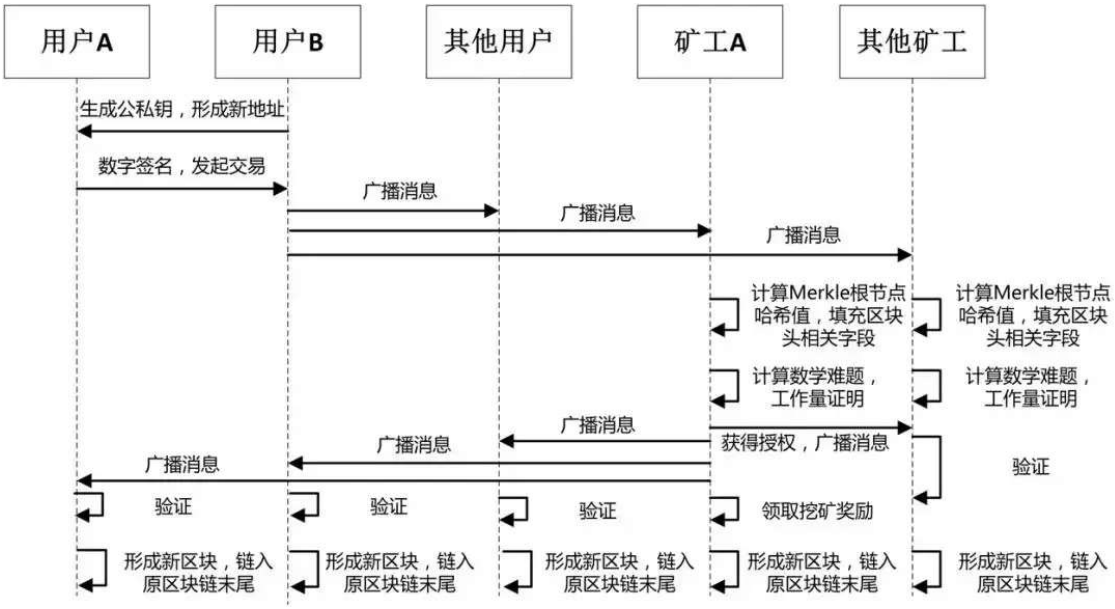

图 8　区块链运转机制示意图

其他相关

脚本语言

区块链采用的脚本语言并不是图灵完备的语言，不支持循环，只能进行堆栈式操作。这种脚本语言的好处是，不允许矿工提交一个死循环的脚本，更注重的是安全方面的考量，但其扩展能力有限。从以太坊为首的区块链编程平台支持图灵完备的编程语言，引领区块链跨入 2.0 时代。由于支持循环等复杂操作，以太坊用 Gas（燃料）机制来防止死循环的出现，确保系统的安全。

消息队列

比特币区块链采用 Zero MQ（ZMQ）作为消息分发和消息队列管理工具。与很多人熟悉的 RabbitMQ 相比，ZMQ 不像传统意义的消息服务器，更像一个底层的网络通信库，在多个线程、内核和主机盒之间弹性伸缩，在 Socket API 之上将网络通信、进程通信和

线程通信抽象为统一的 API 接口。

挖矿设备和算法演进

挖矿设备从支持复杂指令（CISC）、适合串行计算的 CPU 矿机时代，经由基于众核体系、适合并行简单计算的 GPU 挖矿和低功耗却价格昂贵的 FPGA 挖矿，逐渐向集约高速的 ASIC 矿机和规模效应的矿池演进。

基于工作量证明机制的算法，容易导致矿工算力集中的问题。有人将这种“中心化”的责任归咎于 SHA256 算法。此时，基于 SCRYPT 算法的莱特币（Litecoin）进入了人们视线，其占用内存多、计算时间长、并行计算困难的特点，限制了矿工的“军备竞赛”。莱特币的成功催生了更多算法的交叉融合，衍生出串联算法（夸克币）、并联算法（HeavyCoin）和多用途算法（在工作量证明的同时，寻找大素数的素数币，PrimeCoin）。

开源项目和工具

区块链的开源项目

BitCoin

BitCoin 是最早、也是现网运行区块链最成功的一个开源项目，核心技术框架采用 C++开发，共识算法采用 PoW，每秒交易量（TPS）为不多于 7 笔，开源许可协议为 MIT。

- 官方编程语言：C++
- 开源许可协议：MIT
- 开源项目地址：https://github.com/bitcoin/bitcoin

Ethereum

以太坊（Ethereum）是一个支持图灵完备脚本运行的区块链开发平台，基于智能合约，降低用户搭建 DApp 应用的门槛。目前以太坊正式运行的版本是 1.0，采用的是 POW 共识算法，公网 TPS 是 25 笔，未来将采用类 POS 的 Casper 算法，区块链的确认速度将得到大幅提升。在规划的 2.0 版本中，TPS 有望可以达到

2000TPS。

- 官方编程语言：Go
- 开源许可协议：GPLv3
- 开源项目地址：https://github.com/ethereum

Hyperledger Fabric

Hyperledger Fabric 是 IBM 开源的区块链项目，开发环境可以适配多种环境（virtualbox 虚拟机、自建网络和 IBM 的 BlueMix），支持 Docker，共识算法插件化，注重角色的权限控制和企业级的安全机制。主要开发语言是 Go 语言，支持 JavaScript、Java 和 Python 等语言，交易频率 TPS 最高能够达到 100K。其子项目 Iroha 助力区块链移动应用程序的开发，值得关注和进一步跟踪。

- 官方编程语言：Go
- 开源许可协议：Apache 2.0
- 开源项目地址：https://github.com/hyperledger/fabric

OpenChain

OpenChain 是区块链技术公司 Coinprism 的开源工具，目标是大型企业和金融机构，基于一种独特的分布式账本技术，帮助用户部署自己定制的区块链，减少用户的交易成本和结算时间。

- 官方编程语言：C#
- 开源许可协议：Apache 2.0
- 开源项目地址：https://github.com/openchain

BitShares

比特股（BitShares）提供的 BitUSD 等锚定资产，是虚拟币历史上的一个最重要变革之一，消除了虚拟货币估值波动大的问题。比特股创新地提出了 DPoS 共识算法，核心技术框架采用 C++语言

开发，既适用于公有链，也适合于联盟链。在比特股 2.0 中，交易频率 TPS 最高能够达到 100K。

- 官方编程语言：C++
- 开源许可协议：MIT
- 开源项目地址：http://github.com/bitshares

Ripple

瑞波（Ripple）是世界上第一个开放的支付网络，也是目前最成功的区块链技术公司。其核心产品 Ripple 协议本质上是一个实时结算系统，通过引入新的共识机制 RPCA，只要特殊节点投票，就能在很短时间内完成交易的验证和确认。

- 官方编程语言：C++
- 开源许可协议：ISC
- 开源项目地址：https://github.com/ripple/rippled

Tendermint

美国公司 Tendermint 推出的 Tendermint 是第一个实施分片技术的公共区块链。Tendermint 主核心负责管理所有区块链分区，支持比特币分区和以太坊分区，具有很大的灵活性，共识引擎通过 Tendermint 套接字协议（TMSP）与应用程序进行连接，不依赖于某一特定的编程语言，所以开发人员可以使用任意一种编程语言来编写智能合约。

- 官方编程语言：Go
- 开源许可协议：Apache 2.0
- 开源项目地址：https://github.com/tendermint/tendermint

Corda

Corda 是 R3CEV 于 2016 年 12 月初开源的区块链平台，采用一

种类区块链的分布式账本，基于产业标准工具，通过创新智能合约和数据处理，为金融服务设计一种新型分布式的分类帐平台。

- 官方编程语言：Go
- 开源许可协议：Apache 2.0
- 开源项目地址: https://www.corda.net/

具体对比图如表 2 所示。

表 2　开源项目的对比表

开源项目	共识算法	语言	智能合约	TPS（理论值）	开源许可
BitCoin	Paxos	C++	否	7	MIT
Ethereum	PoW/Casper	Go	是	<2000	GPLv3
Fabric	PBFT为主	Go	是	<100K	Apache 2.0
OpenChain	非区块链	C#	是	即时确认	Apache 2.0
BitShares	DPoS	C++	否	>500	MIT
Ripple	RPCA	C++	否	<1000	ISC
Tendermint	BFT	Go	是	>10000	Apache2.0
Corda	多种共识	Java	是	不详	Apache2.0

以太坊的集成开发环境（IDE）

Remix 是以太坊官方 IDE，它允许开发者在以太坊区块链创建和部署合约及去中心化应用。它包含一个 Solidity 源代码排错器，Solidity 是以太坊开发的智能合约语言，可以将智能合约代码编译成以太坊虚拟机（EVM）可识别的字节码。此外，如要和以太坊节点交互，主要用到的 Web3.js API；与节点进行底层交互，需要用到 JSON RPC API。以太坊主流项目的对比如表 3 所示。

表 3　以太坊主流项目的对比表

项目名称	开发语言	客户端文件	平台界面
go-ethereum	Go	Geth	命令行
pyethapp	Python	pyethapp	命令行
Mist	JavaScript	Mist	图形界面
solidity	C++	solc	命令行
browser-solidity	C++	浏览器	浏览器界面

总结

区块链凭着数据公开透明、信息安全可靠、来源可证可溯等诸多特性，降低信任构建成本，提高网络协作效率，加速价值的全球流

动，促进下一代信息基础设施持续演进。

本文着重从区块链 1.0 的基本概念、运行机制、相关技术和开源项目及工具四个方面进行了简单的介绍，希望能够帮助读者厘清概念、梳理思路和开拓视野，期待读者针对各行各业痛点，广泛应用区块链技术，有效改善商业规则，营造“区块链+”的新浪潮。

> 注：本文源自卿苏德写于 2017 年的一篇技术文章。
>
> 作者简介：卿苏德，博士，就职于中国信息通信研究院（原工业和信息化部电信研究院）产业与规划所。目前担任 CSDN 区块链知识库特邀编辑。

—————— 推荐阅读 ——————

点击图片即可阅读


谁在控制着 iCloud 中国区账号的密钥?


68存活
游戏外挂衍生黑色产业链，代理月入 10 万元


4 年创 40 亿美元业绩神话，比特币挖矿究竟有多赚钱?

内容转载自公众号

区块链大本营 了解更多

精选留言

该文章作者已设置需关注才可以留言

何永灿CSDN 置顶 7

本文源自卿苏德2017年的技术文章。
作者职位现为中国信息通信研究院高级工程师，可信区块链联盟办公室主任，国际电信联盟ITU-T FG DLT分布式账本焦点组测试评估牵头人。马化腾《数字经济》、周宏仁主任《中国信息化形势分析与预测（2017）》区块链章节撰稿人；中国信息通信研究院《全球区块链应用十大趋势》、《区块链在物联网中的应用》等报告的核心编制成员。

昨天

作者回复 3

更新一下作者信息，让我们都记住他

昨天

Zoro 15

可以的可以的，很强势！

昨天

雨 14

如果区块链工作量证明那里，进行一些有意义的运算，比如说计算人类基因图谱，病毒基因分析，大数据分析等。那么虚拟币可以正名，不会被认为是毫无实际价值的毒瘤了。

昨天

糯米团子 10

无非是海量备份和循环加密而已，每个分支要存全量数据，这是以牺牲存储空间和速度来实现安全的笨办法。如果数据基数大一些的话存储空间更是海量，现在各大互联网公司都是试水项目，保证不存在技术空档而已。别拿比特币说事，如果支付宝用比特币，每个人手机客户端要存几亿交易记录，每一次交易都要同步所有客户端，你认为可能么。

昨天

sun 9

突然觉得区块链不是去中心化，而是多中心化。

昨天

Aka_Sam 8

干货

昨天

skyer chan 6

目前小米公司6万块钱每月招聘区块链工程师

昨天

杨志鹏 6

想请教一下：

1.区块链技术主要是为了防中心的吗？（比如中心篡改账单，中心被黑阔入侵后篡改了账单）

2.区块链技术的应用蛮多的，但似乎都看不到实际的社会效益，貌似

使用区块链技术能做的事传统技术也能做（比如金融行业的应用），当然除了挖矿。但是挖矿看起来又没有社会效益。
3.这篇文章都是在结合挖矿讲述区块链内部技术实现。但我对于实际上能产生社会效益的应用不是很了解，比如千年特区号称要应用区块链技术，不知这方面可有可以学习的途径？
非常感谢！
昨天

AbelCui 5
确实不错
昨天

潇潇暮雨 5
“如果没有梅克尔树，区块链也是可以运转，但是要在区块头里包含所有交易记录，扩展性方面存在很大挑战。”
没有 Merkle Tree 的区块链结构就像一条链表，在这种情况下，校验支付者是否有足够可用的coin时，是一个区块一个区块的向上遍历的，不会把所有交易存在一个区块中，这样单个区块的数据大小就太庞大了吧
昨天

Real 4
有个手书错误　非对称加密里应该是“一对密钥”，作者写成了“一个密钥”。
昨天

作者回复 2
自罚一个鸡腿
昨天

DBL0123 3
干货满满
昨天

ben大神点C 3
绝对干货，看了很多篇文章，这是第一篇我想转发朋友圈的。很清晰。
昨天

陈发 2

我之前提交过兼职翻译的简历，静候佳音。
昨天

Dio 1
兄嵇，每个成员都是某种意义上的中心的话，那这就不是多中心化啦
昨天

wch❀ 1
非常喜欢
昨天

未嗳缒隧 1
终于对区块链有点了解了
昨天

欧阳安
干货满满的
6小时前

徐立秋
写的真好，不会有几处错误希望能改过来，还有做好能补充一下智能合约的介绍
昨天

以上留言由公众号筛选后显示

了解留言功能详情