区块链（Blockchain）——核心技术概览

运用之妙夺造化，存乎一心胜天工。

有人可能会遇到这样的问题：

- 跨境商贸合作中签订的合同，怎么确保对方能严格遵守和及时执行？
- 酒店宣称刚打捞上来的三文鱼，怎么追踪捕捞和运输过程中的时间和卫生？
- 现代数字世界里，怎么证明你是谁？怎么证明某个资产属于你？
- 经典囚徒困境中的两个人，怎样才能达成利益的最大化？
- 宇宙不同文明之间的“黑暗森林”猜疑链，有没有可能被彻底打破？

这些看似很难解决的问题，在区块链的世界里已经有了初步的答案。本文将带领大家探索区块链的核心技术，包括其定义与原理、关键的问题等，还将探讨区块链技术的演化，并对未来发展的趋势进行展望。最后，对一些常见的认识误区进行了澄清。

定义与原理

1、定义

公认的最早关于区块链的描述性文献是中本聪所撰写的文章《Bitcoin：A Peer-to Peer Electronic Cach System》，但该文献重点在于讨论比特币系统，实际上并没有明确提出区块链的定义和概念，在其中指出，区块链是用于记录比特币交易账目历史的数据结构。

另外，Wikipedia 上给出的定义中，将区块链类比为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、不可篡改的数据记录。

区块链技术最早的应用出现在比特币项目中。作为比特币背后的分布式记账平台，在无集中式管理的情况下，比特币网络稳定运行了八年时间，支持了海量的交易记录，并且从未出现严重的漏洞，这些都与巧妙的区块链结构分不开的。

区块链技术自身仍然在飞速发展中，目前相关规范和标准还在进一步成熟中。

2、基本原理

区块链的基本原理理解起来并不复杂。首先，区块链包括三个基本概念：

- 交易（transaction）：一次对账本的操作，导致账本状态的一次改变，如添加一条转账记录；
- 区块（block）：记录一段时间内发生的所有交易和状态结果，是对当前账本状态的一次共识；
- 链（chain）：由区块按照发生顺序串联而成，是整个账本状态变化的日志记录。

如果把区块链作为一个状态机，则每次交易就是试图改变一次状态，而每次共识生成的区块，就是参与者对于区块中交易导致状态改变的结果进行确认。

在实现上，首先假设存在一个分布式的数据记录账本，这个账本只允许添加、不允许删除。账本底层的基本结构是一个线性的链表，这也是其名字“区块链”的来源。链表由一个个“区块”串联组成（如图 2-1 所示），后继区块记录前导区块的哈希值（pre hash）。新的数据要加入，必须放到一个新的区块中。而这个块（以及块里的交易）是否合法，可以通过计算哈希值的方式快速检验出来。任意维护节点都可以提议一个新的合法区块，然而必须经过一定的共识机制来对最终选择的区块达成一致。

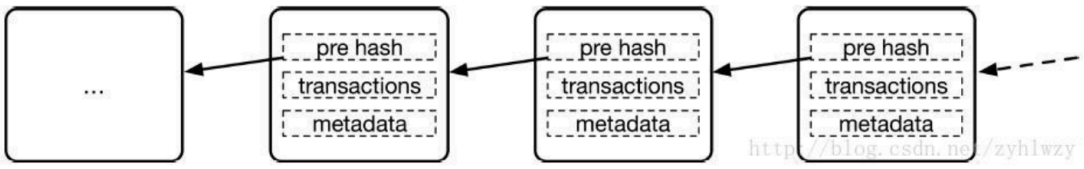

图 2-1 区块链结构示例

3、以比特币为例理解区块链工作过程

以比特币网络为例，可以具体看其中如何使用了区块链技术。

首先，比特币客户端发起一项交易，广播到比特币网络中并等待确认。网络中的节点会将一些收到的等待确认的交易记录打包在一起(此外还要包括前一个区块头部的哈希值等信息)，组成一个候选区块。然后，试图找到一个 nonce 串（随机串）放到区块里，使得候选区块的哈希结果满足一定条件（比如小于某个值）。这个 nonce 串的查找需要一定的时间去进行计算尝试。

一旦节点算出来满足条件的 nonce 串，这个区块在格式上就被认为是“合法”了，就可以尝试在网络中将它广播出去。其他节点收到候选区块，进行验证，发现确实符合约定条件了，就承认这个区块是一个合法的新区块，并添加到自己维护的区块链上。当大部分节点都将区块添加到自己维护的区块链结构上时，该区块被网络接受，区块中所包括的交易也就得到确认。

当然，在实现上还会有很多额外的细节。这里面比较关键的步骤有两个：

- 一个是完成对一批交易的共识（创建区块结构）；
- 一个是新的区块添加到区块链结构上，被大家认可，确保未来无法被篡改。

比特币的这种基于算力寻找 nonce 串的共识机制称为工作量证明（Proof of Work，PoW）。目前，要让哈希结果满足一定条件，并无已知的快速启发式算法，只能进行尝试性的暴力计算。尝试的次数越多（工作量越大），算出来的概率越大。

通过调节对哈希结果的限制，比特币网络控制平均约 10 分钟产生一个合法区块。算出区块的节点将得到区块中所有交易的管理费和协议固定发放的奖励费（目前是 12.5 比特币，每四年减半），这个计算新区块的过程俗称为挖矿。

读者可能会关心，比特币网络是任何人都可以加入的，如果网络中存在恶意节点单，能否进行恶意操作来对区块链中的记录进行篡改，从而破坏整个比特币网络系统。比如最简单的，故意不承认收到的别人产生的合法候选区块，或者干脆拒绝来自其他节点的交易等。

实际上，比特币网络中存在大量（据估计数千个）的维护节点，而且大部分节点都是正常工作的，默认都只承认所看到的最长的链结构。只要网络中不存在超过一半的节点提前勾结一起采取恶意行动，则最长的链将很大概率上成为最终合法的链。而且随着时间增加，这个概率会越来越大。例如，经过 6 个区块生成后，即便有一半的节点联合起来想颠覆被确认的结果，其概率也仅为（1/2）6≈1.6%，即低于 1/60 的可能性。

当然，如果整个网络中大多数的节点都联合起来作恶，可以导致整个系统无法正常工作。要做到这一点，往往意味着付出很大的代价，跟通过作恶得到的收益相比，得不偿失。

技术的演化与分类

区块链技术自比特币网络设计中被大家发掘关注，从最初服务数字货币系统，到今天在分布式账本场景下发挥着越来越大的技术潜力。

1、区块链的演化

比特币区块链已经支持了简单的脚本计算，但仅限于数字货币相关的处理。除了支持数字货币外，还可以将区块链上执行的处理过程进一步泛化，即提供智能合约（smart contract）。智能合约可以提供除了货币交易功能外更灵活的合约功能，执行更为复杂的操作。

这样，扩展之后的区块链已经超越了单纯数据记录的功能，实际上带有一点“智能计算”的意味；更进一步，还可以为区块链加入权限管理和高级编程语言支持等，实现更强大的、支持更多商用场景的分布式账本。

从计算特点上，可以看到现有区块链技术的三种典型演化场景，如表 2-1 所示。

场景	功能	智能合约	一致性	权限	类型	性能	编程语言	代表
工信的数字货币	记账功能	不带有或较弱	PoW	无	公有链	较低	简单脚本	比特币网络
工信的交易处理	智能合约	图灵完备	PoW、PoS	无	公有链	受限	特定语言	以太坊网络
带权限的分布式账本处理	商业处理	多种语言，图灵完备	包括CFT、BFT在内的多种机制，可插拔	支持	联盟链	可拓展	高级语言	超级账本

表 2-1　区块链技术的三种典型演化场景

2、区块链与分布式记账

现代复式记账系统（double entry bookkeeping）由意大利数学家卢卡·帕西奥利于 1494 年在《Summa de arithmetica，geometrica，proportioni et proportionalità》一书中最早制定。复式记账法对每一笔账目同时记录来源和去向，首次将对账验证功能引入记账过程，提升了记账过程的可靠性。

从这个角度来看，区块链是首个自带对账功能的数字记账技术实现。

更广泛地看，区块链属于一种去中心化的记录技术。参与到系统上的节点，可能不属于同一组织，彼此无需信任；区块链数据由所有节点共同维护，每个维护节点都能复制获得一份完整或部分记录的拷贝。

跟传统的记账技术相比，基于区块链的分布式账本应该包括如下特点：

- 维护一条不断增长的链，只可能添加记录，而发生过的记录都不可篡改；
- 去中心化，或者说多中心化，无需集中控制而能达成共识，实现上尽量采用分布式；
- 通过密码学的机制来确保交易无法被抵赖和破坏，并尽量保护用户信息和记录的隐私性。

3、分类

根据参与者的不同，可以分为公开（public）链、联盟（consortium）链和私有（private）链。

- 公有链，顾名思义，任何人都可以参与使用和维护，如比特币区块链，信息是完全公开的；

如果进一步引入许可机制，可以实现私有链和联盟链两种类型：

- 私有链，由集中管理者进行管理限制，只有内部少数人可以使用，信息不公开；
- 联盟链则介于两者之间，由若干组织一起合作维护一条区块链，该区块链的使用必须是带有权限的限制访问，相关信息会得到保护，如供应链机构或银行联盟。

目前来看，公有链更容易吸引市场和媒体的眼球，但更多的商业价值会在联盟链和私有链上落地。

根据使用目的和场景的不同，又可以分为以数字货币为目的的货币链，以记录产权为目的的产权链，以众筹为目的的众筹链等，也有不局限特定应用场景的通用链。

现有大部分区块链实现都至少包括了网络层、共识层、智能合约和应用层等结构，联盟链实现往往还会引入一定的权限管理机制。

关键问题和挑战

从技术角度讲，区块链所涉及的领域比较繁杂，包括分布式系统、存储、密码学、心理学、经济学、博弈论、控制论、网络协议等，这也就意味着大量工程实践上的技术挑战。

下面列出了目前业内关注较多的一些技术话题。

1、抗抵赖与隐私保护

- 怎么防止交易记录被篡改?
- 怎么证明交易双方的身份?
- 怎么保护交易双方的隐私?

密码学的发展为解决这些问题提供了不少手段。传统方案包括 Hash 算法、加解密算法、数字证书和签名（盲签名、环签名）等。

随着区块链技术的应用，新出现的需求将刺激密码学的进一步发展，包括更高效的随机数产生、更高强度的加密、更快速的加解密处理等。同时，量子计算等新技术的出现，也会带来更多的挑战，例如，RSA 算法等目前商用的加密算法，在未来可能无法提供足够的安全性。

能否满足这些新的需求，将依赖于数学科学的进一步发展和新一代计算技术的突破。

2、分布式共识

这是个经典的技术难题，学术界和业界都已有大量的研究成果（包括 Paxos、拜占庭系列算法等）。

问题的核心在于如何解决某个变更在分布式网络中得到一致的执行结果，是被参与多方都承认的，同时这个信息是被确定的，不可推翻的。

该问题在公开匿名场景下和带权限管理的场景下需求差异较大，从而导致了基于概率的算法和确定性算法两类思想。

最初，比特币区块链考虑的是公开匿名场景下的最坏保证。通过引入了“工作量证明”策略来规避少数人的恶意行为，并通过概率模型保证最后参与方共识到最长链。算法在核心思想上是基于经济利益的博弈，让恶意破坏的参与者损失经济利益，从而保证大部分人的合作。同时，确认必须经过多个区块的生成之后达成，从概率上进行保证。这类算法的主要问题在于效率的低下。类似算法还有以权益为抵押的 PoS、DPoS 和 Casper 等。

后来更多的区块链技术（如超级账本）在带权限管理的场景下，开始考虑支持更多的确定性的共识机制，包括经典的拜占庭算法等，可以解决快速确认的问题。

共识问题在很长一段时间内都将是极具学术价值的研究热点，核心的指标将包括容错的节点比例、决策收敛速度、出错后的恢复、动态特性等。PoW 等基于概率的系列算法理论上允许少于一半的不合作节点，PBFT 等确定性算法理论上则允许不超过 1/3 的不合作节点。

3、交易性能

虽然一般来说，区块链不适用于高频交易的场景，但由于金融系统的需求，业界目前十分关心如何提高区块链系统交易的吞吐量，同时降低交易的确认延迟。

目前，公开的比特币区块链只能支持平均每秒约 7 笔的吞吐量，一般认为对于大额交易来说，安全的交易确认时间为一个小时左右。以太坊区块链的吞吐量略高一些，但交易性能也被认为是较大的瓶颈。

区块链系统跟传统分布式系统不同，其处理性能很难通过单纯增加节点数来进行横向扩展。实际上，传统区块链系统的性能，在很大程度上取决于单个节点的处理能力。高性能、安全、稳定性、硬件辅助加解密能力，都将是考查节点性能的核心要素。

这种场景下，为了提高处理性能，一方面可以提升单个节点的性能（如采用高配置的硬件），同时设计优化的策略和算法；另外一方面试图将大量高频的交易放到链外来，只用区块链记录最终交易信息，如比特币社区提出的“闪电网络”等设计。类似地，侧链（side chain）、影子链（shadow chain）等思路在当前阶段也有一定的借鉴意义。类似设计可以将交易性能提升 1～2 个数量级。

此外，在联盟链的场景下，参与多方存在一定的信任前提和利益约束，可以采取更优化的设计，换来性能的提升。以超级账本 Fabric 项目为例，在普通虚拟机配置下，单客户端交易吞吐量可达几百次每秒（transactions per second，tps）；在有一定工程优化或硬件加速情况下可以达到每秒数千次的吞吐量。

客观地说，目前开源区块链系统已经可以满足不少应用场景的性能需求，但离大规模交易系统在峰值每秒数万笔的吞吐性能还有较大差距。

4、扩展性

常见的分布式系统可以通过增加节点来横向扩展整个系统的处理能力。对于区块链网络系统来说，根据共识机制的不同，这个问题往往并非那么简单。

例如，对于比特币和以太坊区块链而言，网络中每个参与维护的核心节点都要保持一份完整的存储，并且进行智能合约的处理。此时，整个网络的总存储和计算能力取决于单个节点的能力。甚至当网络中节点数过多时，可能会因为一致性的达成过程延迟降低整个网络的性能。尤其在公有网络中，由于存在大量低性能处理节点，导致这个问题将更加明显。

要解决这个问题，根本上是放松对每个节点都必须参与完整处理的限制（当然，网络中节点要能合作完成完整的处理），这个思路已经在超级账本中得到应用；同时尽量减少核心层的处理工作。

在联盟链模式下，还可以专门采用高性能的节点作为核心节点，相对较弱的节点仅作为代理访问节点。

5、安全防护

区块链目前最热门的应用场景是金融相关的服务，安全自然是讨论最多、挑战最大的话题。区块链在设计上大量采用了现代成熟的密码学算法。但这是否就能确保其绝对安全呢？

世界上并没有绝对安全的系统。

系统是由人设计的，系统也是由人来运营的，只要有人参与的系统，就难免出现漏洞。如下几个方面是很难避免的。

首先是立法。对区块链系统如何进行监管？攻击区块链系统是否属于犯罪？攻击银行系统是要承担后果的。但是目前还没有任何法律保护区块链（特别是公有链）以及基于它的实现。

其次是软件实现的潜在漏洞。考虑到使用了几十年的 OpenSSL 还带着那么低级的漏洞（heart bleeding），而且是源代码完全开放的情况下，让人不禁对运行中的大量线上系统持谨慎

态度。而对于金融系统来说，无论客户端还是平台侧，即便是很小的漏洞都可能造成难以估计的损失。

另外，公有区块链所有交易记录都是公开可见的，这意味着所有的交易即便被匿名化和加密处理，但总会在未来某天被破解。安全界一般认为，只要物理上可接触就不是彻底的安全。实际上，已有文献证明，比特币区块链的交易记录很大可能是能追踪到真实用户的。

作为一套完全分布式的系统，公有的区块链缺乏有效的调整机制。一旦运行起来，出现问题也难以修正。即使是让它变得更高效、更完善的修改，只要有部分既得利益者联合起来反对，就无法得到实施。比特币社区已经出现过多次类似的争论。

最后，运行在区块链上的智能合约应用可能是五花八门的，可能存在潜在的漏洞，必须有办法进行安全管控，在注册和运行前需要有合理的机制进行探测，以规避恶意代码的破坏。

2016 年 6 月 17 日发生的“DAO 系统漏洞被利用”事件，直接导致价值 6000 万美元的数字货币被利用者获取。尽管对于这件事情的反思还在进行中，但事实再次证明，目前基于区块链技术进行生产应用时，务必要细心谨慎地进行设计和验证。必要时，甚至要引入“形式化验证”和人工审核机制。

6、数据库和存储系统

区块链网络中的大量信息需要写到文件和数据库中进行存储。

观察区块链的应用，大量的读写操作、Hash 计算和验证操作，跟传统数据库的行为十分不同。当年，人们观察到互联网应用大量非事务性的查询操作，而设计了非关系型（NoSQL）数据库。那么，针对区块链应用的这些特点，是否可以设计出一些特殊的针对性的数据库呢？

LevelDB、RocksDB 等键值数据库，具备很高的随机写和顺序读、写性能，以及相对较差的随机读的性能，被广泛应用到了区块链信息存储中。但目前来看，面向区块链的数据库技术仍然是需要突破的技术难点之一，特别是如何支持更丰富语义的操作。

大胆预测，未来将可能出现更具针对性的“块数据库”（BlockDB），专门服务类似区块链这样的新型数据业务，其中每条记录将包括一个完整的区块信息，并天然地跟历史信息进行关联，一旦写入确认则无法修改。所有操作的最小单位将是一个块。为了实现这种结构，需要原生支持高效的签名和加解密处理。

7、集成和运营

即便大量企业系统准备迁移到区块链平台上，但在相当长的一段时间内，基于区块链的新业务系统必将与已有的中心化系统集成共存。

两种系统如何共存，如何分工，彼此的业务交易如何进行合理传递？出现故障如何排查和隔离？已有数据如何在不同系统之间进行迁移和灾备？区块链系统自身又该如何进行运营（如网络的设计选择、状态监控、灾备等）？

这些都是迫切要解决的实际问题。若解决不好，将是区块链技术落地的不小阻碍。

趋势与展望

关于区块链技术发展趋势的探讨和争论，自其诞生之日起就从未停息。或许，读者从计算技术的演变历史中能得到一些启发。计算技术的发展历史如图 2-3 所示。

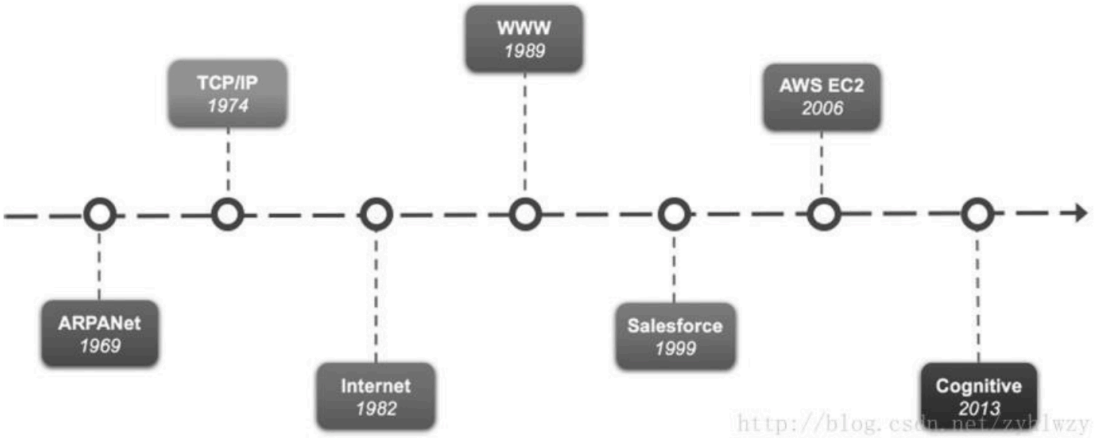

图 2-3　计算的历史

以云计算为代表的现代计算技术，其发展历史上有若干重要的时间点和事件：

- 1969——ARPANet（Advanced Research Projects Agency Network）：现代互联网的前身，由美国高级研究计划署（Advanced Research Project Agency）提出，其使用 NCP 协议，核心缺陷之一是无法做到和个别计算机网络交流；
- 1973——TCP/IP：Vinton.Cerf（文特·瑟夫）与 Bob Karn（鲍勃·卡恩）共同开发出 TCP 模型，解决了 NCP 的缺陷；
- 1982——Internet：TCP/IP 正式成为规范，并被大规模应用，现代互联网诞生；
- 1989——WWW：早期互联网的应用主要包括 telnet、ftp、email 等，Tim Berners-Lee（蒂姆·伯纳斯-李）设计的 WWW 协议成为互联网的杀手级应用，引爆了现代互联网，从那开始，互联网业务快速扩张；
- 1999——Salesforce：互联网出现后，一度只能进行通信应用，但 Salesforce 开始以云的理念提供基于互联网的企业级服务；
- 2006——AWS EC2：奠定了云计算的业界标杆，直到今天，竞争者们仍然在试图追赶 AWS 的脚步；

- 2013——Cognitive：以 IBM Watson 为代表的认知计算开始进入商业领域，计算开始变得智能，进入“后云计算时代”。

从这个历史中能看出哪些端倪呢？

首先，技术领域也存在着周期律。这个周期目前看是7～8年左右。或许正如人有“七年之痒”，技术也存在着七年这道坎，到了这道坎，要么自身突破迈过去，要么就被新的技术所取代。如果从比特币网络上线（2009 年 1 月）算起，到今年正是在坎上。因此，现在正是相关技术进行突破的好时机。

其次，最早出现的未必是先驱。创新固然很好，但过早播撒的种子，若没有合适的土壤，往往也难长大。技术创新与科研创新很不同的一点是，技术创新必须立足于需求，过早过晚都会错失良机。科研创新则要越早越好，比如 20 世纪出现的物理学巨匠们，超前的研究成果奠定了后续一百多年的科技革命的基础。

最后，事物的发展往往是延续的、长期的。新生事物大都不是凭空蹦出来的，往往是解决了前辈未能解决的问题，或是出现了之前未曾出现过的场景。而且很多时候，新生事物的出现需要长期的孵化；坚持还是放弃，故事不断重复。笔者认为，只要是朝着提高生产力的正确方向努力，迟早会有出现在舞台上的一天。

目前，区块链在数字货币领域（以比特币为代表）的应用已经相对成熟，而在智能合约和分布式账本方向尚处于初步实践阶段。但毫无疑问的是，区块链技术在已经落地的领域，确实带来了生产力提升。因此可以相信，随着相关技术的进一步发展，区块链技术必然会在更多的领域中大放异彩，特别是金融科技相关领域。

认识上的误区

目前，由于区块链自身仍是一种相对年轻的技术，不少人对区块链的认识还存在一些误区。下面是需要注意的一些问题：

首先，区块链不等于比特币。虽说区块链的基本思想诞生于比特币的设计中，但发展到今日，比特币和区块链已经俨然成为了两个不太相关的技术。前者更侧重从数字货币角度发掘比特币的实验性意义；后者则从技术层面探讨和研究可能带来的商业系统价值，试图在更多的场景下释放智能合约和分布式账本带来的科技潜力。

其次，区块链不等于数据库。虽然区块链也可以用来存储数据，但它要解决的核心问题是多方的互信问题。单纯从存储数据角度，它的效率可能不高，也不推荐把大量的原始数据放到区块链系统上。当然，现在已有的区块链系统中，数据库相关的技术十分关键，直接决定了区块链系统的吞吐性能。

最后，区块链并非一门万能的颠覆性技术。作为融合多项已有技术而出现的新事物，区块链跟现有技术的关系是一脉相承的。它在解决多方合作和可信处理上向前多走了一步，但并不意味着它是万能的，更不会彻底颠覆已有的商业模式。很长一段时间里，区块链所适用的场景仍需不断摸索，并且跟已有系统也必然是长期合作共存的关系。

小结

本章剖析了区块链的相关核心技术，包括其定义、工作原理、技术分类、关键问题和认识上的误区等。通过本章的学习，读者可以对区块链的相关核心技术形成整体上的认识，并对区块链在整个信息科技产业中的位置和发展趋势形成更清晰的认知。

除了数字货币应用外，现在业界越来越看重区块链技术可能带来的面向商业应用场景的计算能力。开源社区发起的开放的“以太坊”和“超级账本”等项目，让用户可以使用它们来快速设计复杂的分布式账本应用。

有理由相信，随着更多商业应用场景的出现，区块链技术将在未来金融和信息技术等领域占据越来越重要的地位。

摘自：区块链原理、设计与应用

原文链接：https://blog.csdn.net/zyhlwzy/article/details/78637404