软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2018,29(1):150−159 [doi: 10.13328/j.cnki.jos.005434]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

区块链与可信数据管理:问题与方法*

钱卫宁, 邵奇峰, 朱燕超, 金澈清, 周傲英

(华东师范大学 数据科学与工程学院,上海 200062)

通讯作者: 钱卫宁, E-mail: wnqian@dase.ecnu.edu.cn

摘 要: 作为支撑比特币实现无中心高可信的账本管理的技术,区块链在金融领域得到了广泛关注.区块链实现了不完全可信环境中的可信数据管理,具有去中心化、防篡改、不可抵赖、强一致和完整性等特性,但同时也存在高延迟和低吞吐率的性能问题.在互联网技术发展、新型应用层出不穷的大背景下,借鉴区块链在数字加密货币应用中的成功经验,探索可信数据管理的理论、技术,并设计、实现系统,是学术界所面临的重要问题.从可信数据管理角度,介绍了区块链相关的技术和研究进展,包括分布式共识、智能合约、数据溯源等,并分析了应用对可信数据管理所提出的需求和研究挑战.

关键词: 区块链;可信数据管理;智能合约;数据溯源;分布式共识

中图法分类号: TP311

中文引用格式: 钱卫宁,邵奇峰,朱燕超,金澈清,周傲英.区块链与可信数据管理:问题与方法.软件学报,2018,29(1):150−159. http://www.jos.org.cn/1000-9825/5434.htm

英文引用格式: Qian WN, Shao QF, Zhu YC, Jin CQ, Zhou AY. Research problems and methods in blockchain and trusted data management. Ruan Jian Xue Bao/Journal of Software, 2018,29(1):150−159 (in Chinese). http://www.jos.org.cn/1000-9825/5434.htm

Research Problems and Methods in Blockchain and Trusted Data Management

QIAN Wei-Ning, SHAO Qi-Feng, ZHU Yan-Chao, JIN Che-Qing, ZHOU Ao-Ying

(School of Data Science and Eningneering, East China Normal University, Shanghai 200062, China)

Abstract: As a supporting technology of Bitcoin for decentralized ledger management, blockchain has gain much attention in financial domain. Blockchain achieves trusted data management in not fully trusted computation environments. It has the advantage of decentralization, immutability, strong consistency and integrity, however, also suffers from poor performance with high latency and low throughput. With ever growing Internet technology and applications, the success of blockchain technology in cryptocurrency may shed light on the research of new trusted data management theories, technologies and systems. This paper introduces the blockchain related technologies, including distributed consensus, smart contract and data provenance, from the perspective of trusted data management. The requirements and research challenges of trusted data management are also analyzed.

Key words: blockchain; trusted data management; smart contract; data provenance; distributed consensus

区块链(blockchain 或 block chain)是指通过数据加密、数据链式钩稽、多副本存储和分布式共识等机制,实现去中心化的分布式数据管理技术.它最早是由中本聪提出,并在比特币(bitcoin)中加以实现和应用[1].随着比特币应用的快速发展,区块链技术所具有的防篡改、不可抵赖、强一致和完整性等特性,特别是它的对等网络(peer-to-peer network)去中心化本质,得到了工业界和学术界的广泛关注.在加密货币[1]、分布式账本[2]、单据管理[3]、首次代币发售(ICO)和众筹[4]、慈善[5]等领域,区块链技术得到了广泛的探索和应用.

* 基金项目: 国家自然科学基金(61432006, 61672232, 61332006); 国家高技术研究发展计划(863)(2015AA015307)

Foundation item: National Natural Science Foundation of China (61432006, 61672232, 61332006); National High Technology Research and Development Program of China (863) (2015AA015307)

收稿时间: 2017-09-17; 修改时间: 2017-10-16; 采用时间: 2017-11-21

另一方面,最早的区块链技术被设计用于比特币这一特殊的虚拟货币应用.它与应用紧密结合,所能提供的数据管理功能简单,同时基于工作量证明(proof-of-work,简称 PoW)的共识机制的计算量耗费巨大,导致极低的系统吞吐率和很长的系统延迟.如何提供丰富的数据管理和数据处理功能,提高系统性能,成为区块链研究、开发和应用所关心的热点.以以太坊(Ethereum)[6]和 Hyperledger[7]等为代表的开源项目则提供了相对完善的区块链的开发与应用基础,推动了区块链普及、应用的快速增长,以及新问题的发现与研究.

从数据管理角度看,区块链的本质是一个构建在对等网络上、提供了可信数据管理功能的数据库系统.一个可信数据库管理系统从 3 个层面确保系统的可信性,即存储的可信性、处理的可信性以及外部访问的可信性,如图 1 所示.

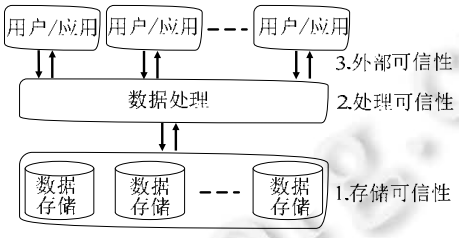


Fig.1 The architecture of trusted data management systems

图 1 可信数据管理系统示意图

存储可信性是指数据处理结果一旦被确认,不会丢失或被篡改.它要求系统提供传统数据库管理系统[8]和事务处理[9]中所要求的事务持久性(durability),但同时也要求系统在存储、通信故障,甚至在蓄意攻击时,仍能确保数据存储的正确性.

处理可信性一方面是指数据处理的正确性,另一方面是指处理过程和结果可审计与可溯源.前者要求事务的并发控制,而后者则要求系统不仅保存数据的最终状态,还要保存数据处理的过程.数据处理的正确性是对传统数据管理系统的基本要求.但是,传统的数据库管理系统是集中式的,保持事务的 ACID 属性已有成熟并相对高效的技术[8,9].对等网络环境中的数据管理,大都专注于查询处理的性能[10,11].虽然已有大量关于分布式系统中的共识(consensus)机制研究[12,13],但在数据管理系统中,由于性能问题,共识机制和跨节点的协调通常只被用于选举主控节点,而较少被直接应用于事务处理或被尽量避免[14].因此,在区块链这样的去中心化对等网络环境中,如何在确保系统“正确”的同时,实现高效事务处理,就成为一个突出的问题.

处理过程和结果的可审计及可追溯也是重要的研究问题.在传统的数据库管理系统中,数据库中存储、维护的是当前的数据状态,处理过程和数据的历史信息通常存储在数据库日志中,仅被用于故障恢复[8,9],并不直接提供查询服务;在系统无故障正常运行的情况下,也不参与查询的处理.在节点不可信的对等网络环境中,一些查询和事务在处理时需要验证数据的历史状态,以确保当前状态的正确性.因此,传统的数据管理技术无法被直接应用于这一场景.

数据溯源(data provenance)是数据管理中的一项重要技术,在科学数据管理和数据仓库中有着广泛的应用[15].然而,很多数据溯源技术仅针对集中式数据库或节点可信的分布式环境,在区块链的应用场景下无法直接应用.

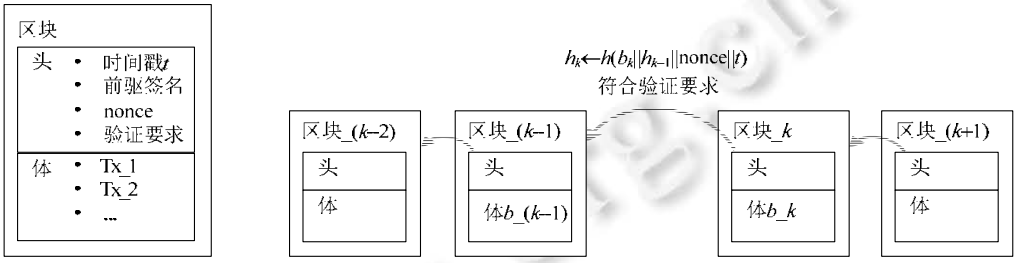
外部访问可信性是指对用户访问的认证.在实现机制上,它依赖于分布式身份认证等技术,也与具体的应用场景和业务紧密相关.本文的综述不涉及外部访问可信性.

与已有的从数字货币[16]、安全[17]、协议[18]、系统架构[19]、私有链[20]和研究挑战[21]角度所进行的区块链技术综述不同,本文从可信数据管理的角度梳理区块链与相关数据管理技术的关联,介绍在不完全可信的对等网络环境中的数据管理问题和相关技术,并分析它们在新型应用场景中的适用性.由于外部可信性一方面与应用的具体模式紧密关联,另一方面又可以部分地依赖于分布式认证[22]技术解决,因此,本文聚焦于存储可信性和处理可信性技术.

本文第 1 节简单介绍区块链的基本数据结构和概念.第 2 节从分布式共识的角度介绍存储可信性保障技术.第 3 节介绍处理可信性,包括智能合约及其问题、数据溯源技术以及可认证查询处理.第 4 节简要介绍主要的区块链系统和应用.最后,第 5 节对可信数据管理技术所面临的研究挑战进行分析.

1 区块链基础

区块链的基本数据结构包括两部分,即区块内结构与区块间链式结构,分别如图 2(a)和图 2(b)所示.一个区块包含头信息和体信息.头信息是区块的元数据,用于验证区块,并与其前驱和后继区块建立关联.通常,头信息包含自身时间戳、前驱区块的签名值、一个特殊值(称为 nonce)、验证要求(如难度目标).体信息则是交易的序列.



(a) 区块内结构 (b) 区块间链式结构

Fig.2 Logic data structures of blockchains

图 2 区块链逻辑结构

只有当一个区块的签名结果满足验证要求时,一个区块才能通过验证.例如,在比特币中,区块散列(签名)后的结果值必须小于某个特定值(该值由难度目标决定,随着时间的变化,难度逐渐增加)[1].当一个区块需要与其前驱建立关联时,其体信息、前驱区块签名值、自身时间戳、验证要求等信息都已经确定,唯一能调整以获得不同自身签名值的变量就是 nonce 值.只有当获得了合法的 nonce 值后,区块才能通过验证,与前驱区块进行链接.

区块内的交易序列常通过特殊的数据结构,如 Merkle-tree[23],进行组织.Merkle-tree 是一种树型数据结构,最初提出时为二叉树,但可被拓展为多叉树,其叶子节点为数据项或数据项的散列值,每一个内部节点的值为其所有子节点的散列值,从而根节点的值可被视为整棵树的签名.利用这一性质,Merkle-tree 可被方便地用来实现数据集相等测试、定位修改以及零知识证明.因此,Merkle-tree 在区块链中被用于检测区块副本是否相同.

区块链的逻辑结构确保区块间的关系可验证.在系统中,一个区块存储于多个节点,以应对由于节点或网络故障所引起的区块副本丢失问题.

需要注意的是,对于区块(k–1),可能存在多个区块 $k1,k2,\ldots,kp$,都能通过验证,成为(k–1)的后继.如何让参与区块链的所有节点对区块链结构达成一致,其本质是分布式共识(consensus)[12]问题.通常,区块链仅承认链最长的那条链.下一节将对区块链中的分布式共识机制进行介绍.

区块链系统的另一个重要方面是其提供服务的接口.在比特币应用中,区块链仅提供转账,即事务的执行与查询.而随着应用需求以及以太坊和 HyperLedger 等系统的发展,新的区块链平台提供了称为“智能合约(smart contract)”的用户代码执行机制.从可信性角度看,智能合约不仅可被执行,且其执行历史将被记录,执行过程和结果可审计、可追溯.第 3.1 节将介绍区块链中的智能合约处理机制,而对于数据溯源这一特殊问题则在第 3.2 节中加以介绍.

2 存储可信性

2.1 工作量证明机制

如前所述,存储可信性解决区块的容错一致问题,其本质是分布式共识问题.比特币中的区块链采用了被称为工作量证明(proof of work,简称 PoW)的机制来解决这一问题.PoW 基于如下技术和假设:根据 b_k,t 和 h_{k-1} 计算使 h_k 满足验证要求的 nonce 需要耗费算力,每次计算 nonce 所需的算力在一定时间段内相当.这一计算过程被称为“挖矿”.因此,如果需要篡改或伪造记录,则需要构造一条比当前被公认的区块链(主链)更长的链,因此需要的算力需要超过整个区块链中的其他(正在进行正常挖矿运算的)算力.或者,更准确地说,在考虑网络延迟时,攻击者的算力接近 50%就会破坏比特币区块链的正确性[24].而当考虑“自私挖矿(selfish mining)”——也就是当自身“挖矿”所获得的链比别人的链长时,不发布自己的链,在自己的链上继续挖;当自身的链和别人已发布的链相比等长或者更短时,立即发布自己的链,并在别人已发布的链上继续“挖矿”,那么,攻击者接近 1/4 算力即会危及比特币的正确性[25].

PoW 共识机制的另一个问题是其性能问题.如 Vukolic[26]和 Tseng[13]对 PoW 和传统的拜占庭容错问题进行了详细的对比分析所述,由于比特币区块链为“公有链”,即其参与读取、交易以及共识机制的用户是开放的,其用户规模是动态的,参与者是匿名的.这直接导致了 PoW 机制的低吐率和高延迟.但从另一个角度看,PoW 机制实现了系统的高可扩展性,支持从数千到数十万个参与者,这一网络规模远远大于绝大多数金融机构信息系统的规模.

2.2 实用拜占庭容错机制

并非所有区块链应用的需求和对环境的假设都与比特币相同.例如,在私有链(private blockchain 或 permissioned blockchain)或联盟链(consortium blockchain)中,节点(参与者)就不再是匿名的,节点规模远小于公有链,且可信程度也远比在公有链中要高.实用拜占庭容错机制(practical Byzantine fault tolerance,简称 PBFT)可被用于该场景[27,28].与 PoW 不同,采用 PBFT 时,区块仅有被选举出的唯一主控节点生成.PBFT 由请求、预准备、准备、提交这 4 个阶段构成.预准备由主控节点发起,准备阶段各节点分别验证主控节点发起的共识请求的正确性,并将验证结果返回给主控节点,并由主控节点汇总后在提交阶段确定是否提交.与 PoW 相比,PBFT 适用于节点数少于 20 个的场景,可拜占庭容错少于 1/3 的节点的攻击,即有少于 1/3 的节点存在漏发、错发或选择性错发消息情况,主要开销在于网络消息传输带宽,吞吐率可达数千,并将延迟降到毫秒级.此外,PBFT 可确保系统的最终一致性.由于具有这些特性,PBFT 被应用于 HyperLedger Fabric.

2.3 Paxos和BVP

PoW 和 PBFT 考虑的都是拜占庭容错问题.在私有链的场景下,若假设节点或参与者不进行攻击,则可进一步放宽假设.Paxos 是重要的非拜占庭场景下的共识机制[29,30],可被用于私有链场景.与 PBFT 相比,Paxos 的吞吐率可进一步提升到超过 4 万 tps[31].

Paxos 的改进版本也能处理拜占庭容错场景,被称为拜占庭 Paxos[32].Abraham 和 Malkhi 提出了 BVP[33],用以利用 TPM(trusted platform module)加密处理器[34]提供高性能的拜占庭容错.

2.4 其他面向区块链的共识机制

PoW、PBFT 和 Paxos 分别是 3 个典型的可用于区块链的共识机制.除此以外,不同的区块链项目也采用它们的改进版本或其他机制.PPCoin 采用权益证明(proof of stake,简称 PoS),面向公有链,避免了 PoW 导致的算力消耗和能源消耗[35].PoS 通过奖励机制鼓励参与节点成为验证者节点,区块的产生由随机选取的验证者节点或验证者节点集合验证获批.PoS 避免了 PoW 导致的大量算力和电力消耗.Ripple 为另一个公有链平台,采用其自身的 RPCA 机制实现共识[36].RPCA 首先将共识问题归结到系统中的一组“受信任”节点,然后采用类似于 PBFT 的投票选取主控节点方式,实现共识.

此外,还有 Proof-of-Luck[37]、Raft[38]等共识机制被应用于区块链系统或应用.

3 处理可信性

比特币区块链仅支持“挖矿”和转账,功能上仅适用于数字货币,具有很大的局限性.在传统的数据库管理系统中,用户通过提交事务来处理数据.事务常由过程型语言与SQL语句组合共同构成,事务执行的过程或结果通过日志进行记录.以太坊首先采用智能合约实现区块链中的数据处理[6],而区块链的逻辑结构本身就和日志具有相似之处.

3.1 智能合约

智能合约是指通过信息技术手段实现的可自动执行的任务合约[39],其概念出现远早于区块链技术.智能合约包含执行条件和执行逻辑.当条件满足时,执行逻辑会被自动执行.从数据管理角度看,智能合约与数据管理系统中的触发器和存储过程[40]具有相似性.另一方面,与传统数据库管理系统中的事务不同,不仅智能合约所做的处理结果需要在区块链中保存,智能合约本身也需要被保存在区块链中,并在系统的各个节点间同步,以确保不同节点和用户所看到的智能合约的一致性.

比特币区块链仅提供非常简单的脚本语言,用以实现智能合约;Ripple 不提供智能合约;以太坊提供图灵完备的智能合约脚本语言;而 HyperLedger Fabric 则提供 Go 和 Java 撰写智能合约的功能.

智能合约扩展了区块链处理数据的能力,但同时也对其使用者以及系统的安全性提出了更高的要求.以太坊智能合约曾发生 TheDAO 攻击[41].Maurice Herlihy 对区块链中由并发控制等因素导致的智能合约问题进行了系统的梳理[42].

3.2 数据溯源

类似于数据库日志,区块链维护了区块链上所有操作和处理的记录.但区块链所提供的数据查询及分析处理功能较为简单.作为一种可信数据管理系统,对区块链上的数据进行溯源,是一个重要的问题.虽然理论上,在像比特币这样的区块链平台上,每一笔交易都能够回溯到“挖矿”所获得的原始比特币,但是如何在引入更为复杂的智能合约以后,在区块链平台所管理的数据随着应用增多、规模扩大以后越来越多时,高效处理数据溯源查询,是区块链技术发展及在更多应用中推广使用所面临的研究题目.

数据溯源(data provenance)是指对于数据处理流程的管理,解决回答数据为什么是该状态(why)、数据从哪儿来(where)以及如何获得(how)的问题[15,43−45].数据溯源的研究在科学数据管理、数据仓库、数据资产管理(data curation)的背景下进行.

数据溯源方法可分成两大类,即基于批注(annotation-based)的方法[46−49]和非批注(non-annotation-based)的方法[50,51].对于非批注的方法,在处理数据的过程中,不需要对源数据和目标数据(处理的结果)附加额外的信息.但是,此时需要了解存储、维护数据进行了何种处理.当处理是可逆的时候,通过目标数据,就能反推得到源数据.需要注意的是,虽然如 SPJ(select-project-join)这样的查询,数据处理是可逆的,但是很多数据库常用查询是不可逆的.例如,很多聚集函数是不可逆的.非标注的数据溯源可用于数据变换、数据集成过程的调试[52].当源数据与目标数据之间的数据模式改变时,这类方法尤为有用.

基于批注的方法将每个数据项变换为$\langle s,d,i\rangle$三元组标签,其中,s 表示数据项源,d 表示目标数据(当前数据),而 i 则表示中间数据结果.通过在数据处理过程中进行标签传播,实现数据的勾连,以支持数据溯源.基于批注的数据溯源系统包括 DBNotes[53]和 Mondrian[54].

数据溯源的查询表达具有严格的代数学基础[55,56],可在关系数据库上实现[57].CuratedDB[58]和 Trio[59,60]是两个重要的数据溯源原型系统.

数据溯源的理论和技术与数据的结构化模式之间关联紧密.虽然在当前的区块链应用中,数据未必一定是结构化的,但是,随着应用的发展,区块链数据管理中数据模式的管理将成为一个重要的问题,也将是现有数据溯源方法能否被成功应用的关键问题.

3.3 可认证数据查询与处理

在查询区块链中的数据时,确保每个区块中数据的正确性是确保最终结果可信的前提,该问题类似于外包数据库中的可认证查询处理[61,62].与可认证查询处理类似,区块链也常用基于 Merkle-tree 的结构来维护一个区块内的事务之间的关系.

随着区块链上查询需求的增长,在链式日志结构上的索引技术也正在成为重要的研究问题[63].

除了查询处理,近年来在如云计算平台这样的硬件和操作系统不可信的非可信计算平台上,提供可信的数据处理,也成为了研究热点.例如,Haven 系统原型通过在飞地(enclave)中利用 Intel SGX 芯片和 Linux LibOS 实现屏蔽模块(shield module),提供包括线程、虚存、调度与文件系统的抽象,提供了无需进行修改的应用程序与 Windows 操作系统间的相互调用[64].

VC3 是由微软研究院研发的另一个原型系统,用以在 Hadoop 平台上运行 MapReduce 程序,确保数据与处理是可信的[65].与 Haven 不同,VC3 并不在 SGX 中加载操作系统库,而只在其中加载 Map/Reduce 程序与处理的数据.程序与数据仅在处理器上运行时是解密的.

Haven 和 VC3 都可以运行未经修改的代码,并处理数据.针对机器学习问题,如决策树、SVM、神经元网络、矩阵分解、*K*-means 聚类等算法,Ohrimenko 等人提出了运行于 SGX 的保护隐私的机器学习方法,他们通过使用实现了 oblivious 原语的 libO 库,重写机器学习算法,达到了可验证的安全机器学习的目的[66].

类似地,Sinha 等人提出,通过将代码分为包含程序逻辑的用户程序和包含内存管理和加密通信原语的运行时库两部分,当运行时库满足信息发布约束(information release confinement,简称 IRC)时,系统的数据处理的安全性是可验证的[67].

4 区块链数据管理系统与应用

除比特币、以太坊和 HyperLedger 以外,近年来还出现了一大批区块链相关的系统.

BigChainDB 试图同时实现传统数据库管理系统的高性能和区块链系统的可伸缩性[68].它采用两层架构.底层依赖于 MongoDB,实现事务和故障恢复,高层利用基于权威的区块链协议应对攻击情况下的容错.由于没有采用拜占庭容错机制,BigChainDB 具有较高的性能.BigChianDB 的目标应用为数字资产管理.

Bitcoin-NG 的设计目标为“下一代”比特币,以具备更好的可伸缩性[69].与其他同类系统相似,它也采用两层协议,一层选主.选举得到的主控节点负责事务的串行化执行.由于避免了事务执行中的 PoW,Bitcoin-NG 可实现其更好的可伸缩性的目标.

Blockstack 为普林斯顿大学研发的基于区块链的命名和存储服务[70].Blockstack 将用于控制的元数据管理与数据存储分开管理,并用 skip-list 管理区块,以避免大规模地扫描区块链.实验结果表明,它能极大地减少计算资源的消耗.

在应用方面,CrowdBC 为基于以太坊的去中心化的众包平台[71].众包中的任务分发、回收等操作都采用智能合约实现.通过采用区块链技术,与集中式众包平台相比,CrowdBC 可以更好地保护用户的隐私.ProvChain 则通过采用区块链技术,在云计算服务中提供数据溯源和数据验证服务[72].

与以上系统和应用不同,Weaver 是一个图数据库,提供了对于区块链结构的高效查询[73].

Quorum(https://www.jpmorgan.com/global/Quorum)是 Morgan 基于以太坊开发的面向企业的区块链平台,与以太坊相比,Quorum 实现了拜占庭容错的共识算法,其事务吞吐率可达到近千事务/秒.而 Corda(https://www.corda.net)则是 R3 公司主导研发的开源分布式账本平台.Corda 面向金融应用,与其他区块链平台不同,它并不在所有节点上维护所有数据的副本,并且,在分布式共识机制的基础上提供了便利的业务逻辑编写机制.

5 小结与展望

通过分析可见,PoW 共识机制、智能合约等区块链技术是面向金融应用,特别是数字加密货币而设计的.它们在确保对等网络中数据和处理的可信性方面,具有很好的特性,可伸缩性尤其突出.但同时,现有的区块链技

术在如下 3 个方面仍然存在着缺陷.首先,现有区块链系统和平台的服务接口通常是过程性的,需要用户撰写复杂的智能合约,与数据库系统声明性的数据操纵相比,容易导致错误和漏洞产生.第二,现有区块链系统大都不支持复杂模式数据管理,不能提供通用的数据建模和模式管理功能,导致系统及平台与应用耦合度高、应用开发难度大.第三,由于部分区块链平台,特别是公有链系统,为了保障系统可伸缩性,采用了 PoW 共识机制,在性能上,特别是延迟和吞吐率方面,无法与传统的可信数据管理系统相比,无法满足大多数关键任务应用的需要.

当考虑更复杂场景下的可信数据管理问题时,现有的区块链技术和系统无法被直接应用.另一方面,信息技术发展是国家政策和新兴商业模式落实应用的前提:可信数据管理对于构建社会信用体系,从机制上提供信用保障,至关重要.

例如,在共享经济、大宗商品交易、数字资产增值利用开发、安全监督、政府治理等应用中,业务可能涉及双方或多方,业务间关联模式各不相同,数据的结构化程度不同,事务的复杂程度和并发数不同,数据处理的及时性响应要求也不相同,借鉴区块链在数字加密货币应用中的成功经验,特别是其在系统可伸缩性、完全去中心化、灵活的智能合约撰写、验证以及可信执行上的特点,在对现有区块链和相关技术进行梳理和分析的基础上,探索大规模分布式环境下的可信数据管理基础理论,设计针对特定应用的可信数据管理系统,或称其为“分享型数据库(sharing database)”系统,提出安全、高效的可信数据管理系统实现方法,是研究的重要问题.

当前所直接面临的研究问题包括:在特定场景下的高性能分布式共识机制、区块链上的结构化数据管理方法、区块链上的分布式数据索引构造和维护方法、链式结构或日志结构上的高效查询处理和优化技术等.

References:

[1] Satoshi N. Bitcoin: A peer-to-peer electronic cash system. 2008. https://bitco.in/pdf/bitcoin.pdf

[2] Walport M. Distributed ledger technology: Beyond block chain. Technical Report, gs-16-1, UK Government Office for Science, 2016.

[3] Cuende LI. Systems and methods for using a block chain to certify the existence, integrity, and/or ownership of a file or communication. U.S. Patent 9,679,276, 2017.

[4] Jacynycz V, Calvo A, Hassan S, Sánchez-Ruiz AA. Betfunding: A distributed bounty-based crowdfunding platform over ethereum. In: Proc. of the 13th Int’l Conf. on Distributed Computing and Artificial Intelligence. Springer Int’l Publishing, 2016. 403−411. [doi: 10.1007/978-3-319-40162-1_44]

[5] Elizabeth W. How blockchain can bring financial services to the poor. MIT Technology Review. 2017. https://www.technologyreview.com/s/604144/how-blockchain-can-lift-up-the-worlds-poor/

[6] Gavin W. Ethereum: A secure decentralisedgeneralised transaction ledger. Ethereum Project Yellow Paper, 151, 2014.

[7] Christian C. Architecture of the hyperledger blockchain fabric. In: Proc. of the Workshop on Distributed Cryptocurrencies and Consensus Ledgers. 2016. https://www.zurich.ibm.com/dccl/papers/cachin_dccl.pdf

[8] Garcia-Molina H, Ullman JD, Widom J. Database Systems—The Complete Book. 2nd ed., Upper Saddle River: Pearson Education, 2009.

[9] Gray J, Reuter A. Transaction Processing: Concepts and Techniques. San Francisco: Morgan Kaufmann Publishers, 1993.

[10] Qian WN. Data management in peer-to-peer systems [Ph.D. Thesis]. Shanghai: Fudan University, 2004 (in Chinese with English abstract). [doi: 10.7666/d.y650588]

[11] Vu QH, Lupu M, O BC. Peer-to-Peer Computing: Principles and Applications. Springer-Verlag, 2010. [doi: 10.1007/978-3-642-03514-2]

[12] Lynch NA. Distributed Algorithms. San Francisco: Morgan Kaufmann Publishers, 1996.

[13] Tseng L. Recent results on fault-tolerant consensus in message-passing networks. In: SIROCCO. 2016. 92−108. [doi: 10.1007/978-3-319-48314-6_7]

[14] Bailis P, Fekete A, Franklin MJ, Ghodsi A, Hellerstein JM, Stoica I. Coordination avoidance in database systems. PVLDB, 2014,8(3):185−196. [doi: 10.14778/2735508.2735509]

[15] Cheney J, Chiticariu L, Tan WC. Provenance in databases: Why, how, and where. Foundations and Trends in Databases, 2009,1(4): 379−474. [doi: 10.1561/1900000006]

[16] Tschorsch F, Scheuermann B. Bitcoin and beyond: A technical survey on decentralized digital currencies. IEEE Communications Surveys and Tutorials, 2016,18(3):2084−2123. [doi: 10.1109/COMST.2016.2535718]

[17] Lin IC, Liao TC. A survey of blockchain security issues and challenges. Int'l Journal of Network Security, 2017,19(5):653−659. [doi: 10.6633/IJNS.201709.19(5).01]

[18] Pass R, Seeman L, Shelat A. Analysis of the blockchain protocol in asynchronous networks. EUROCRYPT, 2017,(2):643−673. [doi: 10.1007/978-3-319-56614-6_22]

[19] Xu XW, Weber I, Staples M, Zhu LM, Bosch J, Bass L, Pautasso C, Rimba P. A taxonomy of blockchain-based systems for architecture design. In: Proc. of the ICSA. 2017. 243−252. [doi: 10.1109/ICSA.2017.33]

[20] Dinh TTA, Liu R, Zhang MH, Chen G, Ooi BC, Wang J. Untangling blockchain: A data processing viewof blockchain systems. Technical Report, NUS., 2017.

[21] Hull R. Blockchain: Distributed event-based processing in a data-centric world: Extended abstract. In: Proc. of the DEBS. 2017. 2−4. [doi: 10.1145/3093742.3097982]

[22] Sirbu MA, Chuang JCI. Distributed authentication in kerberos using public key cryptography. In: Proc. of the NDSS. 1997. 134−143. [doi: 10.1109/NDSS.1997.579231]

[23] Merkle RC. A digital signature based on a conventional encryption function. In: Proc. of the CRYPTO. 1987. 369−378. [doi: 10.1007/3-540-48184-2_32]

[24] Garay JA, Kiayias A, Leonardos N. The bitcoin backbone protocol: Analysis and applications. EUROCRYPT, 2015,(2):281−310. [doi: 10.1007/978-3-662-46803-6_10]

[25] Eyal I, Sirer EG. Majority is not enough: Bitcoin mining is vulnerable. In: Proc. of the Financial Cryptography. 2014. 436−454. [doi: 10.1007/978-3-662-45472-5_28]

[26] Vukolic M. The quest for scalable blockchain fabric: Proof-of-Work vs. BFT replication. In: Proc. of the iNetSeC. 2015. 112−125. [doi: 10.1007/978-3-319-39028-4_9]

[27] Castro M, Liskov B. Proactive recovery in a Byzantine-fault-tolerant system. In: Jones MB, Kaashoek MF, eds. Proc. of the 4th Symp. on Operating System Design and Implementation. San Diego: USENIX Association, 2000. 273−288.

[28] Castro M, Liskov B. Practical Byzantine fault tolerance and proactive recovery. ACM Trans. on Computer Systems, 2002,20(4): 398−461. [doi: 10.1145/571637.571640]

[29] Lamport L. The part-time parliament. ACM Trans. on Computer Systems, 1998,16(2):133−169. [doi: 10.1145/279227.279229]

[30] Lamport L. Fast paxos. Distributed Computing, 2006,19(2):79−103. [doi: 10.1007/s00446-006-0005-x]

[31] Guo JW, Chu JJ, Cai P, Zhou MQ, Zhou AY. Low-Overhead Paxos replication. Data Science and Engineering, 2017,2(2):169−177. [doi: 10.1007/s41019-017-0039-z]

[32] Lamport L. Byzantizing Paxos by refinement. In: Proc. of the DISC. 2011. 211−224. https://doi.org/10.1007/978-3-642-24100-0_22

[33] Abraham I, Malkhi D. BVP: Byzantine vertical Paxos. In: Proc. of the Distributed Cryptocurrencies and Consensus Ledgers (DCCL). 2016. https://www.zurich.ibm.com/dccl/papers/abraham_dccl.pdf

[34] Morris T. Trusted platform module. In: Proc. of the Encyclopedia of Cryptography and Security. Springer US, 2011. 1332−1335. [doi: 10.1007/978-1-4419-5906-5_796]

[35] King S, Nadal S. PPCoin: Peer-to-Peer crypto-currency with proof-of-stake. 2012. http://peerco.in/assets/paper/peercoin-paper.pdf

[36] Schwartz D, Youngs N, Britto A. The Ripple protocol consensus algorithm. Technical Report, Ripple Labs Inc., 2014. https://ripple.com/files/ripple_consensus_whitepaper.pdf

[37] Milutinovic M, He W, Wu H, Kanwal M. Proof of luck: An efficient blockchain consensus protocol. IACR Cryptology ePrint Archive, 2017. 249. [doi: 10.1145/3007788.3007790]

[38] Ongaro D, Ousterhout JK. In search of an understandable consensus algorithm. In: Gibson G, Zeldovich N, eds. Proc. of the USENIX Annual Technical Conf. Philadelphia: USENIX Association, 2014. 305−319.

[39] Nick S. Smart contracts: Building blocks for digital markets. 1996. http://www.alamut.com/subj/economics/nick_szabo/smart Contracts.html

[40] Stonebraker M, Brown P, Moore D. Object-Relational DBMSs: The Next Great Wave. 2nd ed., San Francisco: Morgan Kaufmann Publishers, 1998.

[41] Buterin V. Hard Fork completed. 2016. https://blog.ethereum.org/2016/07/20/hard-fork-completed/2016

[42] Herlihy M. Blockchains and the future of distributed computing. In: Proc. of the PODC. 2017. 155. [doi: 10.1145/3087801.3087873]

[43] Buneman P, Khanna S, Tan WC. Why and where: A characterization of data provenance. In: Proc. of the ICDT. 2001. 316−330. [doi: 10.1007/3-540-44503-X_20]

[44] Dong XL, Kementsietsidis A, Tan WC. A time machine for information: Looking back to look forward. SIGMOD Record, 2016, 45(2):23−32. [doi: 10.1145/3003665.3003671]

[45] Chen A, Wu Y, Haeberlen A, Loo BT, Zhou WC. Data provenance at Internet scale: Architecture, experiences, and the road ahead. In: Kossmann D, Balazinska M, Stonebraker M, eds. Proc. of the 8th Biennial Conf. on Innovative Data Systems Research. 2017.

[46] Wang YR, Madnick SE. A polygen model for heterogeneous database systems: The source tagging perspective. In: McLeod D, Sacks-Davis R, Schek HJ, ed. Proc. of the 16th Int'l Conf. on Very Large Data Bases. Queensland: Morgan Kaufmann Publishers, 1990. 519−538.

[47] Buneman P, Khanna S, Tan WC. On propagation of deletions and annotations through views. In: Proc. of the PODS. 2002. 150−158. [doi: 10.1145/543613.543633]

[48] Cong G, Fan WF, Geerts F. Annotation propagation revisited for key preserving views. In: Proc. of the CIKM. 2006. 632−641. [doi: 10.1145/1183614.1183705]

[49] Buneman P, Cheney J, Vansummeren S. On the expressiveness of implicit provenance in query and update languages. In: Proc. of the ICDT. 2007. 209−223. [doi: 10.1007/11965893_15]

[50] Woodruff A, Stonebraker M. Supporting fine-grained data lineage in a database visualization environment. In: Proc. of the ICDE. 1997. 91−102. [doi: 10.1109/ICDE.1997.581742]

[51] Cui YW, Widom J, Wiener JL. Tracing the lineage of view data in a warehousing environment. ACM Trans. on Database Systems, 2000,25(2):179−227. [doi: 10.1145/357775.357777]

[52] Chiticariu L, Tan WC. Debugging schema mappings with routes. In: Dayal U, Whang KY, Lomet DB, *et al*., eds. Proc. of the 32nd Int'l Conf. on Very Large Data Bases. Seoul: ACM Press, 2006. 79–90.

[53] Bhagwat D, Chiticariu L, Tan WC, Vijayvargiya G. An annotation management system for relational databases. In: Nascimento MA, Özsu MT, Kossmann D, *et al*., eds. Proc. of the 30th Int'l Conf. on Very Large Data Bases. Toronto: Morgan Kaufmann Publishers, 2004. 900−911.

[54] Geerts F, Kementsietsidis A, Milano D. MONDRIAN: Annotating and querying databases through colors and blocks. In: Proc. of the ICDE. 2006. 82. [doi: 10.1109/ICDE.2006.102]

[55] Green TJ, Karvounarakis G, Tannen V. Provenance semirings. In: Proc. of the PODS. 2007. 31−40. [doi: 10.1145/1265530.1265535]

[56] Green TJ, Tannen V. The semiring framework for database provenance. In: Proc. of the PODS. 2017. 93−99. [doi: 10.1145/3034786.3056125]

[57] Srivastava D, Velegrakis Y. Using queries to associate metadata with data. In: Proc. of the ICDE. 2007. 1451−1453. [doi: 10.1109/ICDE.2007.369033]

[58] Buneman P, Chapman A, Cheney J. Provenance management in curated databases. In: Proc. of the SIGMOD Conf. 2006. 539−550. [doi: 10.1145/1142473.1142534]

[59] Widom J. Trio: A system for integrated management of data, accuracy, and lineage. In: Stonebraker M, Weikum G, DeWitt D, eds. Proc. of the 2nd Biennial Conf. on Innovative Data Systems Research. 2005. 262−276.

[60] Benjelloun O, Sarma AD, Halevy AY, Widom J. ULDBs: Databases with uncertainty and lineage. In: Dayal U, Whang KY, Lomet DB, *et al*., eds. Proc. of the 32nd Int'l Conf. on Very Large Data Bases. Seoul: ACM Press, 2006. 953−964.

[61] Li FF, Yi K, Hadjieleftheriou M, Kollios G. Proof-Infused streams: Enabling authentication of sliding window queries on streams. In: Koch C, Gehrke J, Garofalakis MN, *et al*., eds. Proc. of the 33rd Int'l Conf. on Very Large Data Bases. Vienna: ACM Press, 2007. 147−158.

[62] Li FF, Hadjieleftheriou M, Kollios G, Reyzin L. Authenticated index structures for aggregation queries. ACM Trans. on Information and System Security, 2010,13(4):32:1−32:35. [doi: 10.1145/1880022.1880026]

[63] Zhu YC, Zhang Z, Cai P, Qian WN, Zhou AY. An efficient bulk loading approach of secondary index in distributed log-structured data stores. In: Proc. of the Database Systems for Advanced Applications. 2017. 87−102. [doi: 10.1007/978-3-319-55753-3_6]

[64] Baumann A, Peinado M, Hunt GC. Shielding applications from an untrusted cloud with haven. ACM Trans. Computer Systems, 2015,33(3):8:1−8:26. [doi: 10.1145/2799647]

[65] Schuster F, Costa M, Fournet C, Gkantsidis C, Peinado M, Mainar-Ruiz G, Russinovich M. VC3: Trustworthy data analytics in the cloud using SGX. In: Proc. of the IEEE Symp. on Security and Privacy. 2015. 38−54. [doi: 10.1109/SP.2015.10]

[66] Ohrimenko O, Schuster F, Fournet C, Mehta A, Nowozin S, Vaswani K, Costa M. Oblivious multi-party machine learning on trusted processors. In: Holz T, Savage S, eds. Proc. of the USENIX Security Symp. Austin: USENIX Association, 2016. 619−636.

[67] Sinha R, Costa M, Lal A, Lopes NP, Rajamani SK, Seshia SA, Vaswani K. A design and verification methodology for secure isolated regions. In: Proc. of the PLDI. 2016. 665−681. [doi: 10.1145/2980983.2908113]

[68] Bigchain DB, Gmb H. A BigchainDB Primer. Berlin, 2017. https://www.bigchaindb.com/whitepaper/bigchaindb-primer.pdf

[69] Eyal I, Gencer AE, Sirer EG, van Renesse R. Bitcoin-NG: A scalable blockchain protocol. In: Argyraki KJ, Isaacs R, eds. Proc. of the 13th USENIX Symp. on Networked Systems Design and Implementation. Santa Clara: USENIX Association, 2016. 45−59.

[70] Ali M, Nelson JC, Shea R, Freedman MJ. Blockstack: A global naming and storage system secured by blockchains. In: Gulati A, Weatherspoon H, eds. Proc. of the 2016 USENIX Annual Technical Conf. Denver: USENIX Association, 2016. 181−194.

[71] Li M, Weng J, Yang AJ, Lu W. CrowdBC: A blockchain-based decentralized framework for crowdsourcing. IACR Cryptology ePrint Archive, 2017. 444.

[72] Liang XP, Shetty S, Tosh DK, Kamhoua CA, Kwiat KA, Njilla L. ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In: Proc. of the CCGrid. 2017. 468−477. [doi: 10.1109/CCGRID.2017.8]

[73] Dubey A, Hill GD, Escriva R, Sirer EG. Weaver: A high-performance, transactional graph database based on refinable timestamps. PVLDB, 2016,9(11):852−863. [doi: 10.14778/2983200.2983202]

附中文参考文献:

[10] 钱卫宁.对等计算系统中的数据管理[博士学位论文].上海:复旦大学,2004. [doi: 10.7666/d.y650588]

钱卫宁(1976－),男,浙江上虞人,博士,教授,博士生导师,CCF 专业会员,主要研究领域为面向互联网级应用的数据管理系统,可扩展事务处理,大数据管理系统基准评测,海量数据分析处理及其应用.

邵奇峰(1976－),男,副教授,主要研究领域为大数据,区块链.

朱燕超(1992－),男,博士生,主要研究领域为分布式数据库,区块链.

金澈清(1977－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为基于位置的服务,数据质量,不确定数据管理,区块链.

周傲英(1965－),男,博士,教授,博士生导师,CCF 会士,主要研究领域为 Web 数据管理,数据密集型计算,内存集群计算,分布事务处理,大数据基准测试和性能优化.