软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, [doi: 10.13328/j.cnki.jos.000000]

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

一种可扩展的区块链架构模型

余 跃[1], 李 卓[2], 梁 然[3], 赵 诚 成[1,]

[1](Ultiledger 研发部, 广东 深圳 518000)
[2](道富信息科技（浙江）有限公司,浙江 杭州 310013)
[3](深圳共赢链研发部, 广东 深圳 518000)
通讯作者: 李 卓 , E-mail: lizhuo@zju.edu.cn

摘 要: 区块链是一种融合了分布式数据存储、点对点网络、共识算法、密码学等多项技术的新兴技术，构建在区块链上的应用具有去中心化、不可篡改、可溯源等特性。然而，由于底层传输机制、计算性能及共识算法在架构上的制约，现有的区块链系统都面临着潜在的交易拥堵。目前，比特币及以太坊的每秒交易数 (TPS)都还在个位数或十位数徘徊，这导致其无法承载真正企业级应用的场景。本文将提出一种多链架构模型，其具有一条主链和多条平行子链，实现彻底的资源隔离和服务治理。主链和子链、以及子链之间都可以跨链交互，从而实现业务分片和整体系统交易处理速度的大幅提升。

关键词: 区块链; 多链; 分片; 可扩展性; MCS;
中图法分类号: TP311

A Scalable Blockchain Architecture Model

YU Yue[1], LI Zhuo[2], LIANG Ran[1], ZHAO Chengcheng[1],

[1](Ultiledger, Shenzhen 518000, China)
[2](State Street Technology (Zhejiang) Ltd., ZheJiang Hangzhou 310013, China)
[3](Winwin Chainbill, GuangDong ShenZhen 518000, China)
Corresponding Author: LI Zhuo, E-mail: lizhuo@zju.edu.cn

Abstract: Blockchain is a cuttingedge technology which combines distributed data storage, point-to-point network, consensus algorithms, cryptography, etc. Applications built on top of blockchain are decentralized, non-tamperable, traceable. However, due to architecture constrains from current underlying transmission mechanism, computing performance and consensus algorithm, existing blockchain systems faces the potential transaction congestion. At present, both Bitcoin and Ethereum System throughput (TPS) are still on single digit or double-digit level, which makes it impossible to carry real enterprise scenarios. This paper proposes a multi-chain architecture model with a main chain and multiple parallel sub-chains, to implement complete resource isolation and service governance. Main chain and sub-chain, or among sub-chaims are both inter-operable, which will make business shading possible, and improve system transaction processing throughput significantly.

Key words: blockchain; multi-chain; sharding; scalability; MCS

1 研究背景

中本聪提出比特币的概念已经过去有些时间[1]，其底层技术区块链也得到了长足的发展。随着比特币价格走高，对区块链技术的管账户也与日俱增。区块链并不是全新的单一技术，而是由许多其他成熟技术融合而来，包括分布式数据存储、点对点网络、共识算法、密码学等。构建在区块链上的应用具有去中心化、不可篡改、可溯源等特性。然而，由于当下底层传输机制、计算性能及共识算法在架构上的制约，现有的公链都面临着交易处理性能的瓶颈。目前，比特币及以太坊的每秒交易处理数（TPS）都还在个位数或者两位数的水平，这导致其无法承载真正的企业级应用场景。在实际的经济活动中，所需要的 TPS 往往需要达到数千笔每秒。而传统的电商平台、支付和交易所等，实际的 TPS 甚至经常达到每秒成千上万次。

本文提出了一种多链分片（Multi Chain Sharding）的体系架构，包括单条主链和多条平行子链。各链之间实现彻底的资源隔离，架构本身提供服务治理，从而提升了整体系统架构的可扩展性和灵活性。不同业务可以分布在不同子链，天然地实现了业务分片。本文后续部分结构如下：第二章介绍区块链架构在吞吐量上面临的挑战；第三章介绍 MCS 的体系结构；第四章从单一业务扩展的角度对 MCS 架构进行扩展介绍；最后一张总结 MCS 和其他区块链技术的对比，以及实践情况并展望下一阶段的工作。

2 相关研究

作为区块链早期的应用案例，比特币成功已经向人们展示了区块链技术的优势和潜能。考虑到越来越多不同类型的服务将连接到区块链系统，现有的区块链架构很难支撑日益复杂的业务需求。比特币是目前市值最大的公有链，但是其吞吐量理论最大值不超过 7[1]，而现有金融体系的吞吐量远大于此，例如 VISA 平均值为 2000tps，峰值可达到 56000tps[13]。

以太坊[2]是目前功能性最完备的公有链之一，其最先在区块链系统引入虚拟机（EVM）[2]，并且提供一个图灵完备的开发语言。基于以太坊出现了针对不同行业和场景多种应用，比如 2017 年底流行起来的迷恋猫[12]，这是一个类似宠物养成的游戏，每只猫都是独特的、交易于区块链的不可变更物体。但以太坊当前的 TPS 也只是 7-15 之间，完全无法承载体量稍大的应用需求。比如迷恋猫，上线不久就导致了以太坊网络长时间的拥堵，一度占据了整个以太坊网络 15%的流量。值得一提的是，以太坊规划的下一代 PoS 协议 Casper，设计的目标 TPS 在 100,000 或以上，也采用了分片设计思想。[14]

现有的大多数区块链系统都是单链架构的模式，即基于该系统的所有应用都将数据存储在单一链，遵循同样的数据块结构，网络通信也基于单一链。这种单体的架构具有架构简明、数据复制程度高的特点，但是也存在单体系统的常见缺点，例如可扩展性差、数据冗余度过高、可维护性差等。

3 多链分片架构

如前所述，现有大部分区块链系统都是单链架构，每一个节点都是一个全功能的单体节点。在执行一个具体的业务逻辑的时候，每一个节点都需要重复计算大量相同的任务，因而也造成了大量的资源浪费。此外，若系统出现交易拥堵，系统极其容易出现性能的大幅降级。这也是基于区块链构建世界计算机体系的最大阻碍。另外，由于基于单链模型的所有应用都共享同一个账本，这必然会对系统上的应用产生竞争关系，一个应用的火爆可能影响到其他应用的正常运行，比如前面提到的迷恋猫就影响了以太坊的正常转账交易。

显而易见地，单链的架构体系很难满足日益增长的需求并且应用间的资源隔离性很差。像单体应用拆分为微服务来实现资源隔离和可扩展，单链架构向多链架构转变是一个很自然的思路。本文提出一种多链可分片（Multi Chain Sharding，后文简写为 MCS）的区块链体系架构，试图解决业务多样性和可扩展性要求。

3.1 多链架构

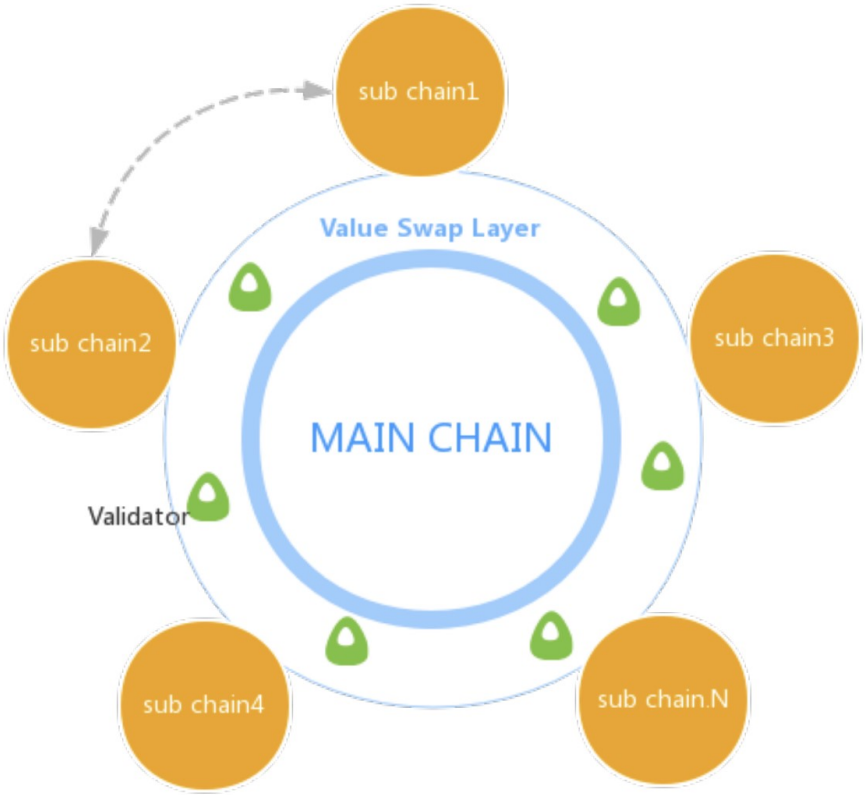

图 1　多链架构模型

MCS 架构的核心包括一个高可用的主链，一组平行子链以及一个连接主链与子链的价值交换层。图 1 是 MCS 架构的模型示意图。

3.1.1　主链

主链是系统的核心，在设计上以最小数据量、最小计算量及最小网络带宽需求为目标，主要提供基础协议、账本、智能合约、信任体系和价值体系服务，起到了整个架构的基础设施作用。它所拥有的最高验证权来源于全部参与节点和联邦拜占庭共识机制。主链一方面存储来自子链的资产交易摘要，另一方面也存储来自子链与子链间的交易和交互记录，一旦被打包到主链区块链上，将形成最安全的且不可逆的共识。主链的另一重要功能是为不同子链上产生的交易提供集中清算。由于每个子链都与主链相连接，在子链上形成共识的交易通过主链共识验证后被打包到主链区块链上，因此所有子链都可以共享总账，实现不同子链间的清算。

3.1.2　子链

子链是系统的参与者，子链节点通常承载大量具体业务的计算，比如产品溯源、汇兑交易、供应链管理、物联网等等。由于 MCS 特有的跨链见证和交互能力，不同子链可以使用不同的共识算法和账本技术。这样各个子链可以根据自己的实际业务特点选择合适的区块链技术，可以是基于联盟链的 Hyperledger[5,8]，也可以是基于私有链的 Ripple[3]、EOS[4]等。

子链包含如下关键要素：初始块、智能合约地址、网关及子链地址协议，它们保证了子链内部交易的可靠性、安全性和有效性。每条子链的起始块是由主链签署的特别块，并为子链定义一个特别账号，称为智能合

约代理网关，以代理子链与外部的一切交易。子链的信任可以交由价值交换层验证，形成双重验证，见证过程中受主链及其制订的规则制约、达到最终验证权归于基础区块链的结果。

3.1.3 价值交换层

价值交换层是主链和子链集合的连接器，它负责子链和主链信息交互，以及子链和子链间的跨链交易。价值交换层充当着代理作用，在不显著增加主链负担的情况下，可以提供对子链交易状态的实时更新已经查询能力，可有效防止子链双花攻击。

价值交换层的处理流程如下：

收到来自子链的包→验证包的合法性→验证交易合法性→更新价值交换层状态→向子链提交收据

价值交换层中终始保存着在子链交易的地址或账户的准确状态，如果需要，还可进一步向存储节点核实子链上的详细交易记录。

3.2 业务分片

多链机制的设计使得原本集中在单链上的工作被分不到单个主链和一系列的子链上去，并且每条子链按具体的使用场景都可以拥有自己不同的特性，比如可有不同的共识算法、不同出块机制和出块时间，这大大提高了系统的灵活性。此机制将不同组织、机构、企业的业务按需求分散到不同子链，不需额外处理，形成了天然的分片，并且保证了各个组织、机构、企业的数据隔离性。

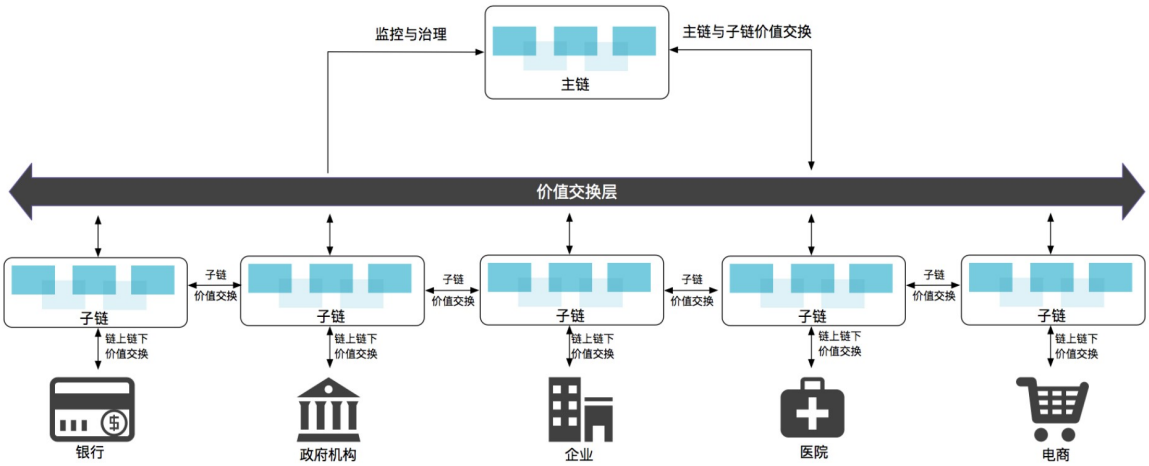

图 2 基于子链的业务分片

如图 2 所示，在多链体系中，不同于大多数单链架构，不同组织、机构或者企业的相关业务可以运行在不同子链。在此种分片模式下，子链可以灵活的选择适合自己业务的具体技术，包括不同的共识算法或者不同的账本技术，例如银行的子链可以采用 PoW[9]，供应链金融的子链可以采用 PBFT[6]，电商子链采用 tendermit[7]或者 POS[10]，甚至在某些完全授信环境下子链可以选择非拜占庭容错的 RAFT[11]算法。

就性能而言，基于该架构的各个应用不再同时将交易提交到主链，通过分而治之的方式，提交交易信息到各自子链，再由子链向主链发起见证。根据不同的业务需求，每个子链的 TPS 需求可以不同，但整个系统的 TPS 却是经过子链简化过后的需要见证的 TPS 的总和。

4 多链分片架构扩展

子链作为一个完整的区块链系统实现，也会拥有自己的存储、清算、交易体系。一方面，独立的账本体系可以依托子链内部节点快速达成共识，实现子链业务的高效进行；另一方面有助于子链内部生态的构建，实现子链系统价值最大化。这也间接推动整个 MCS 架构下系统真正走向高内聚低耦合。

但在某些复杂的业务环境下，一个特定的组织或企业的业务体量可能特别巨大，单一子链承载其整体业务会导致巨量交易拥堵子链本身。针对这种情况，MCS 也支持多级分片，将单一组织业务进一步拆分到不同部门，再分布到不同子链。本节将着重介绍 MCS 为了满足这种需要，提供的多级分片、子链交易见证和跨链交易的。

4.1 多级分片

多级分片实施类似第一级分片，可拆分为以下几个步骤：

1) 初始化子链，该子链主要用于鉴权、验证；
2) 识别业务边界；
3) 按业务边界映射不同业务到各个孙链；
4) 约定孙链和子链、孙链和孙链的交互协议；

以上是两级分片的步骤，同理在业务需要的情况下，可以推广到任意多级分片。多级分片的叶子链是承担具体业务的执行者，叶子链的父链作为交易的见证鉴权，父链的父链也如此，直至主链收集到所有子链的信息。

4.2 子链交易见证

多级分片面临的第一个挑战就是如何保持全网共识。通常情况下，子链都承载着大量具体业务的处理。但是子链的独立账本系统必须被主链账本记录认可才能真正达成全网共识，主链系统的账本记录就是来源于子链共识的记录。换句话说，子链的账本系统脱离主链账本的话，就无法真正获得多链系统的益处。举个例子：在子链 A 上共识的交易如果没被打包至主链，它仅仅是子链内部的共识，在全网并不一定会接受。如果被打包至主链并通过网络节点的验证之后，才形成真正的全网共识，并拥有最高的共识级别。

子链交易的见证流程如下：

1) 子链打包其业务交易到一个子链区块中；
2) 子链通过区块信息构建一个关于该区块的默克尔树（Merkle Tree）；
3) 子链广播默克尔树的根哈希（Root Hash）、时间戳、子链元数据到价值交换层；
4) 价值交换层的验证节点验证签名和其他信息是否合法；
5) 主链接受验证合法的子链消息，该消息作为一个交易被打包到主链区块中；

4.3 跨链交易

很多时候，资产有从一个组织或者机构流通到另外一个组织或者机构的需求，此时一个可以保证交易原子性和安全性的可信机制显得尤为重要。MCS 架构能够通过跨链交易，天然支持这样的价值流动需求。

例如，存在着两个子链 Ca 和 Cb， Ca 和 Cb 间不能够建立直接的交易通道，但是 Ca 和 Cb 均可以同主链 M 建立交易通道，Ca 想支付 Cb 总共 X 个通证（Token）T，该交易的过程如下所示：

1) 子链 C_b 通过其私钥生成一个随机数 N，子链 C_b 再通计算哈希函数 hash(N)得到一个哈希值 H；
2) 子链 C_b 将哈希 H 发送给子链 Ca；
3) 子链 Ca 通过交易 T_1，发送 X 个通证 T 到主链 M，并且附带一个附加协议，M 可以认领该交易但

必须满足一下两个条件；

a) M 必须提供一个 N’满足：

hash(N’) = H

b) M 必须在 P_1 秒内提供满足条件的 N’

4) M 创建一笔交易 T_2，发送 X 个通证 T 到 C_b，并且附带一个附加协议， C_b 可以认领该交易但必须满足一下两个条件：

a) C_b 必须提供一个 N”满足：

hash(N”) = H

b) C_b 必须 P_2 秒内提供满足条件的 N’

5) C_b 拥有原始的随机数 N，C_b 完成在 P_1 秒内完成交易 T_2，并且将完全收到来自 M 的 X 个通证 T。M 此时获取到随机数 N，M 使用 N 完成交易 T_1。

5 总结与展望

本文提出了一种基于异构多链分片的区块链架构模型 MCS，包括主链、平行子链以及起到中介连接作用的价值交换层，通过异构子链天然支持业务分片，并且可以将子链扩展到多级分片模型，结合子链到主链的交易见证，以及子链之间的跨链交易支持，做到灵活适配不同业务场景的需要，并且拥有更好的系统可扩展性和处理性能。表 1 总结了 MCS 与常见区块链系统架构在不同方面的异同。

	吞吐量（TPS）	共识算法	可扩展性	资源隔离性	复杂性
比特币	7	POW	低	低	低
以太坊	7-15 [14]	POW->POW/POS 混合->POS	中	低	中
Hyperledger Fabric	50-2250(不同配置以及节点数量) [8]	PBFT	中	中	中
MCS	可随子链数量增长而增长	主链任意 BFT 算法，不同子链灵活选择不同算法	高	高	高

表 1 常见区块链系统架构比较

关于区块链系统的性能和效率的讨论由来已久，多链的架构设计在近年来得到了越来越多人的关注和实现[1]，也出现了包括联盟链、比特币的闪电网络[15]和以太坊的雷电网络[16]等。Hyperledger Fabric 也支持多节点和多链的设计，但是 Fabric 从本质上说仍然共享同一套账本，在特定的场景下这是一种必须的设计需要，但是应用扩展到覆盖了更多更大的业务场景的时候，单一账本体系的限制就比较明显。而 MCS 基于异构子链的隔离设计，各个子链可以实现完全不同的账本体系和共识机制。当子链出块之后，以主链约定的方式，通过价值交换层完成子链摘要数据到主链的同步，实现更高层面的信任。

和以闪电网络为代表的非主链交易处理的架构比较，MCS 在跨链交易部分也使用了哈希时间锁的机制来保证效率，但是 MCS 整体来说是一个更完整的系统架构模型，能够实现各种资产的链上交换，而闪电网络仅仅是对现有系统的一个补充。

基于 MCS 架构模型，已经有多个不同行业的实际案例的实施[17]，跨越了供应链金融、知识产权保护、跨境结算、医疗大数据、公益事业、分布式能源，初步检验了 MCS 在适配不同业务需求、交易处理效率和跨链协作上的能力。作为下一步，我们需要进一步检验和加强 MCS 的隐私保护能力、开展压力测试探知当

1 这里将数据流动不局限于主链的设计都归于多链架构.

前架构的最优子链层级、探索最佳工程实践，总结提炼框架来提高部署实施的速度。

参考文献:

[1] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. bitcoin.org, 2008.

[2] V. Buterin. Ethereum: Now Going Public. ethereum.org, 2014

[3] D. Schwartz, N. Youngs, and A. Britto. The Ripple Protocol Consensus Algorithm. 2014.

[4] EOS.IO Techinical White Paper v2. March, 2018. Available: https://github.com/EOSIO/Documentation/blob/master/TechnicalWhitePaper.md

[5] Hyperledger- blockchain technologies for business, accessed on Feb. 1 2017, [Online]. Available: https://www.hyperledger.org/.

[6] Castro, M. and Liskov, B. Practical byzantine fault tolerance and proactive recovery. ACM Transactions on Computer Systems (TOCS), 20(4):398–461, 2002.

[7] Kwon, J. Tendermint: Consensus without mining. 2014.

[8] Androulaki, E., Barger, A., Bortnikov, V., Cachin, C., Christidis, K., De Caro, A., Enyeart, D., Ferris., et al. Hyperledger fabric: a distributed operating system for permissioned blockchains. In Proceedings of the Thirteenth EuroSys Conference, page 30. ACM. 2018.

[9] J, Markusm, J, Ari. Proofs of Work and Bread Pudding Protocols. Communications and Multimedia Security. Kluwer Academic Publishers: 258–272, 1999.

[10] Proof-of-stake, Wikipedia, retrieved 2018-08-27

[11] ONGARO, D., AND OUSTERHOUT, J. In search of an understandable consensus algorithm (extended version). http://ramcloud.stanford.edu/raft.pdf.

[12] Cryptokitties, 2017, https://www.cryptokitties.co/

[13] Visa, 2015, https://usa.visa.com/dam/VCOM/download/corporate/media/visa-fact-sheet-Jun2015.pdf

[14] Vitalik Buterin, Etherum Scalability, 2015, https://github.com/vbuterin/scalability_paper/blob/master/scalability.pdf

[15] Dryja & Poon, 2016, https://lightning.network/lightning-network-paper.pdf

[16] Raiden Network, 2018, https://raiden.network/

[17] Ultiledger Whitepaper, 2018, https://www.ultiledger.io/files/ult_tech_whitepaper_v1.0.pdf