X.Blockchain

Yongseok Kwon

2017-05-23 Rev 1.

Abstract

出现 Bitcoin 并利用它的交易迅速增长，证明块环链技术作为交易账本，可靠而安全。
块环链技术备受瞩目的主要理由如下：与现有方式不同，在确保可靠性的问题上，排除可信任的第三方(Trusted Third Party, TTP)；所有交易数据分散保存于所有网络参与者，所以根本不可能操控交易。
块环链技术的最重要核心概念是“Decentralization（去中心化）”和“Distributed Ledger（分布式账本）”。对于现有方式，所有交易都记录在一个中央服务器，对相关交易的可靠性被中央服务器（可信任的第三方）“保证”。但在块环链上所发生的交易传输给所有网络参与者，进行“验证”并获得“同意”，然后以“块”为单位组合，依次(线性）连接。

块环链记录所有交易数据，随着累计交易数量增加，即，时间的推移，其大小会增加，这意味着将来总有一天会面临所有网络参与者无法存储并管理全体块环链的情况。再说，具有能存储并管理全体块环链的性能的系统（节点）数量逐渐减少，相对形成少数节点群。这样将会带来另一形式的中心化。在相对少数节点群管理全体块环链的情况下，交易可靠性只能依赖于少数节点群。即，块环链的根本概念-“去中心化”会受到严重损失。

本文章是对应用为保护电子文件的块环链技术，提出将块环链的连接结构从现有线性结构转变为多层次结构的 X.Blockchain，以探讨如何解决全体块环链容量问题及其导致的节点中心化问题。

Problems

块环链大小随着时间的推移，与累计交易数量成比例，会持续增加。块环链的基本概念是交易账本分散保存于所有参与块环链的节点，以此来没有可信任的第三方也可确保可靠性。但块环链大小持续增加导致限制节点的参与。即，作为存储并管理巨大块环链的全节点[1]参加，需要具有足够存储空间等一定水平的性能。该性能水平会与块环链大小成比例，会提高。这

[1] 全节点是指存储全体块环链并开采新块的节点。

样导致节点减少，最终带来另一形式的“中心化”[2]。截至 2017 年 5 月，包括 Bitcoin 的交易数据的全体块环链大小已超过 115G[3]，而 Ethereum 的块环链也最近超过 20G。

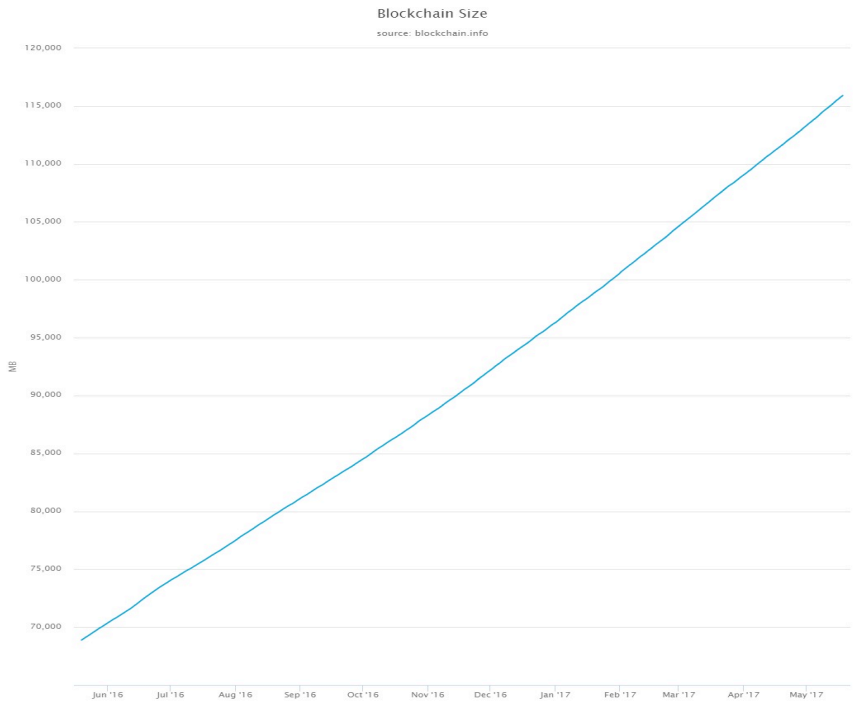

Figure 1. Bitcoin’s full blockchain size

此外，据统计，以 Bitcoin 为准，当前全世界全节点数量约为 5,000 ~ 7,000 个。
而不到一定水平的众多用户(包含移动设备)受到限制作为全节点参与。所以他们无法自己(没有可信任的第三方）判定交易的可靠性，而只能“委托”少数节点群并一味“接受”他们提供的结果。在此，少数节点群起像“可信任的第三方”一样的作用。

发生“全节点的中心化”源于如上述所说，存储全体块环链以及生成(开采)块时被要求很高的 computing power。在此，为什么要求存储全体块环链，是因为块环链是线性结构，只取所需块没有意义。非连接块群无法确保可靠性，也没有任何价值。

这就是在加密货币“交易”时无法避免的问题。对记录多少货币从哪个账户转移到哪个账户，无法予以限制或分类，不能以货币的“量”或用于货币转移的“账户”为准分类。所有交易有一样

[2] 为解决该问题，Bitcoin 提出 SPV。除 Transaction 数据外，仅通过块头信息验证交易，以 Resource 最小化，在为电子文件的块环链应用方面，其是必需的。对加密货币，Transaction 的构成因素为交易数据，而对电子文件应用，Transaction 的构成因素就是电子文件数据。在此，文件大小很大，根本比不上加密货币的交易数据，所以块环链不包含文件数据本身。
除非在本文件中提及，“块环链大小”就意味着“块头大小”。

[3] 是包含指描述交易的所有 Transaction 数据和块头信息的全体块环链大小。

意义，所有具有同样意义的交易以任何标准也不能分类。鉴于此，脱离线性结构，根本不可能。

相反，“电子文件”与加密货币不同，多样文件相关记录都以相关“文件”为中心进行。所以文件的创建、修改、传输、阅览及报废等都针对相关“文件”进行。

这意味着文件相关记录可以其“文件”为准分类，这样块环链的链也可分成多条链，而不是一个线性结构。

N-Dimensional Blockchain – X.Blockchain

X.Blockchain 是反映前面所说的电子文件特性，对所有文件相关记录(Transaction)，按“文件”或相当于它的某个“标准”来分类，而不是利用一个线性结构来连接。而且，按其标准，形成多数链条，以提出多层次的块环链。
例如，以“文件”为准，“最初生成”文件被记录在具有与现有块环链一样线性结构的块环链(main-chain)中。但对已记录在 main-chain 的文件发生修改等附加记录(Transaction)，那么其不是记录在 main-chain 中，而记录在将 main-chain 的相关块当作 genesis block[4]的另一个块环链——sub-chain 中。

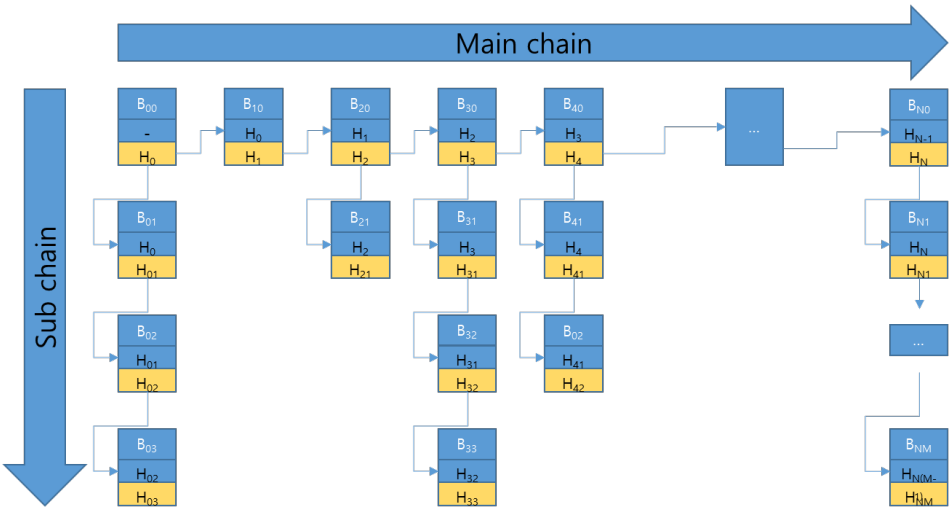

Figure 2. X.Blockchain

上图表示二维结构的 X.Blockchain。如对前面所述的“文件”，以 main-chain 为准时，构成 main-chain 的每个块(B_{00} ~ B_{n0})包含所有新文件的生成记录，同时也可称为每个 sub-chain 的 genesis block。例如，对在 B_{20} 有其生成记录的电子文件 E_{20} 首次发生修改，其会被记录

[4] 块环链的第一个块

在将 B_{20} 当作 genesis block 的 sub-chain 的 B_{21} 中，而不是 main-chain 中下一个块——B_{30}。

Figure 3. 线性块环链

在线性块环链结构上，即使是同一个文件，对相关文件的修改等附加记录，也需要块环链的附加块。
上图表示以线性构成将文件 D_0 ~ D_3 追加到块环链中的例子。其中，文件 D_n 表示生成，D_{n-m} 表示文件 D_n 所发生的修改及传输等附加记录。文件 D_0 包括生成在内共发生 6 次记录(~D_{0-5})，而文件 D_2 共发生 3 次记录(~D_{2-2})。
对如上图所示的线性块环链，所有客户为自己验证某个文件的可靠性，应确保并存储全体块。即，客户即使只需要文件 D_2，也要确保实际上不需要的对 D_0、D_1 及 D_3 的块和包括 D_0 附加记录的块——D_{0-1}~D_{0-5} 等共 11 个块。

以 X.Blockchain 构成同一情况，如下：

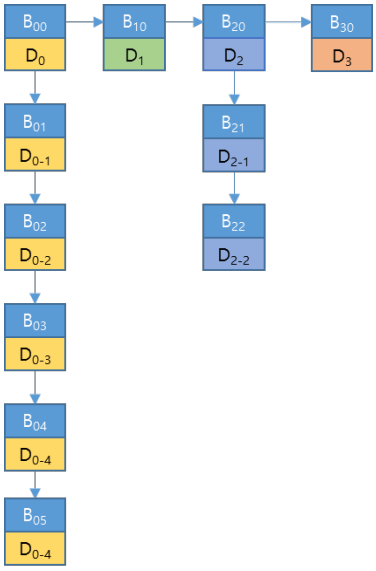

Figure 4. X.Blockchain

如上图所示，B_{00} ~ B_{30} 构成 main-chain，还构成将 B_{00} 和 B_{20} 当作 genesis block 的 sub-chain。对于如此多层次结构的块环链，不用所有客户都拥有全体块环链。如前面例子所示，当客户只需文件 D_2 时，不需要全体块环链，而确保将 D_2 当作 genesis block 的 sub-chain 和其上级 main-chain，就可自己判定文件 D_2 的可靠性。即，拥有由 main-chain 的块 B_{00} ~ B_{30} 和 sub-chain 的块 B_{21} ~ B_{22} 构成的共 6 个块信息，即可。

在此，适用于 Main-chain 的“分类标准”不一定是“文件”，要么是按服务的“部门”要么是有关联性的文件组合。

此外，按情况，可构成 sub-chain 为其他 sub-chain 的 main-chain。对于土地档案，将大范围“地区”作为 main-chain，将小单位市、郡及区作为第一 sub-chain，并将土地区别单位作为第二 sub-chain，这样可构成如下图所示的三维结构，而不是如上图所示的二维结构。

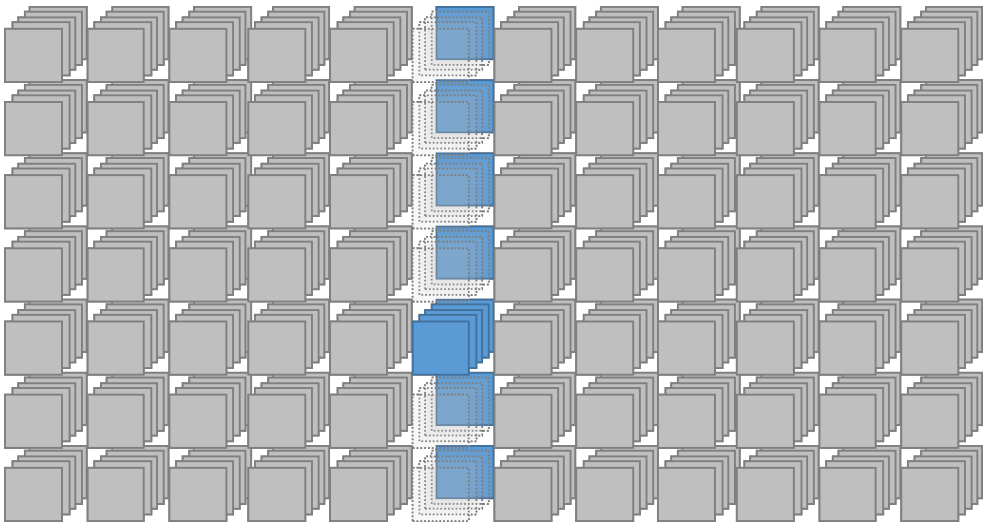

Figure 5. 3 维结构的 X.Blockchain

在 X.Blockchain 上负责开采新块的全节点仍然需要拥有对所有块的信息。然而不是开采，而是判定文件的可靠性时，客户(用户设备)不需要拥有全体块。用户拥有包括需要验证的文件的 sub-chain 和其上级 main-chain，就足以没有“可信任的第三方”而自己判定相关文件的可靠性。在此，全节点不是确定可靠性的可信任第三方，而起执行“开采”的作用。

Parallel Transaction Processing

在线性块环链结构上，形成一个块时所发生的交易(Transaction)都处于等待状态，直到开始开采下一个块。尤其是一个块大小[5]不能无限扩大，所以也有可能要等到再开采下一个块。分配给块环链网络的 transaction 从包括在块中到确定所等待时间使“每秒处理 Transaction 数量” -TPS[6]减少，如此，处理 transaction 会延迟。

但在 X.Blockchain 上各 sub-chain 同时进行开采块。例如，在 main-chain 上开采块 B_{n0} 和在 sub-chain 上开采块 B_{2m} 可同时进行。因为 B_{n0} 和 B_{2m} 分别属于不同的块环链，互相没有关联性。

[5] Bitcoin 的块大小限为 1MB。限制块大小，导致限制块能包含的交易数量，从而造成每单位时间处理的交易数量(Transaction Per Seconds - TPS)减少。当前，就扩大 Bitcoin 的块大小进行讨论，但不受开采者的支持。

[6] 当前 Bitcoin 的 TPS 约为 7 TPS。

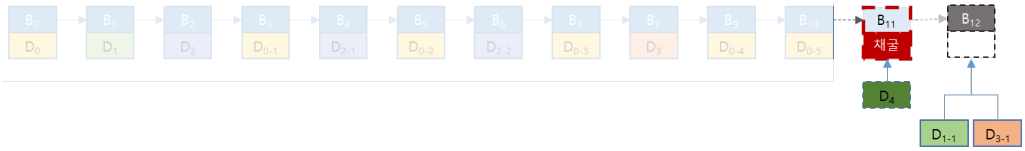

Figure 6. Transaction Processing - 线性结构的块环链

上图假设在前面所提出的示例中新生成 D_4，接着 D_1 和 D_3 发生修改记录。在线性块环链上，D_1 和 D_3 的修改记录，在对新生成 D_4 进行开采期间一直处于等待状态。对 D_4 的块 B_{11} 开采完成，再开始开采块 B_{12} 时，D_1 和 D_3 的修改内容包括在 B_{12} 中并开采完后，最终连接到块环链上，对块环链的记录都完成。

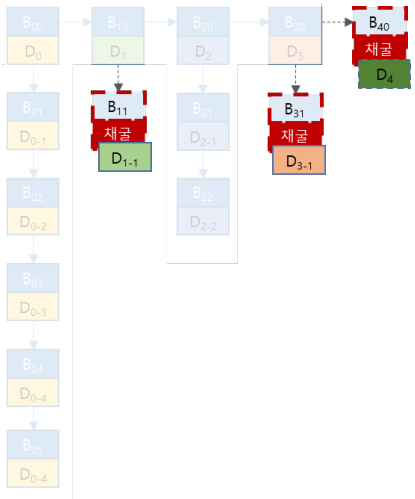

Figure 7. Transaction Processing – X.Blockchain

但在多层次结构的 X.Blockchain 上，如图所示，对文件 D_1 及 D_3 修改事项的块 B_{11} 及 B_{31} 不用等到对文件 D_4 生成的块 B_{40} 开采完成后再开始开采。如图所示，对所有新块 B_{40}、B_{11} 及 B_{31} 的开采可以独立而同时进行。

Effectiveness

假设将居民登记手写本做成电子文件，并适用块环链技术管理时，下面说明 X.Blockchain 和现有块环链有什么差异。
假设每一个人有一个居民登记手写本，其再构成一个块，每年按迁移人数发生居民登记手写本更新，其也记录于一个块。

[单位：千名，%，千件，来源：统计厅 「国内人口迁移统计」

		2016
总迁移	迁移人数	7,378
	迁移率(%)	14.4
	迁移申报件数	4,636
	迁移人性比(女性=100)	103.9
按地区 纯迁移人数	首都圈	-1
	中部圈	41
	湖南圈	-16
	岭南圈	-40

据国家统计门户(http://kosis.kr)发表，截至 2015 年底，韩国总人口为 51,525,338 名。每一个人有一份居民登记手写本，而在每次迁移时，要更新居民登记手写本。据上表数据，在 2016 年共有 7,378,000 次[7]更新居民登记手写本。

就此，如以线性结构构成块环链，构成相当于总人口数的块，还追加相当于每年迁移人数的块。如从 2016 年开始适用，以 2016 年年底为准，块环链的块数量如下：

51,525,338 (初期块数量) + 7,378,000 (2016 年修改块数量) = 58,903,338 个

而且假设每年有 7,000,000 名迁移，每年会加 7 百万个块。在此情况下，每个块的大小为 80 byte[8]，累计 10 年记录的全体块环链大小如下：

块环链大小= (51,525,338+7,000,000*10)*80/1024^3= 9.1G

[7] 据国家统计门户发表，2016 年准确的迁移人数为 7,378,383 名。
[8] Bitcoin 的块头大小为 81byte。

即，对线性结构的块环链，累计 10 年的块环链大小为 9.1G，而每年会线性增加存储修改数据的 0.52G。

如将同一条件适用于 X.Blockchain，全体块数量和大小一样，但每年增加的修改块数量不是以线性连接到 main-chain，而被构成为 sub-chain。即，累计 10 年的 70,000,000 个修改块会分散在由 51,525,338 个块构成的 main-chain 的 sub-chain。简单计算修改块分散在 main-chain 的 sub-chain 的程度，main-chain 的每一个块拥有一个 sub-chain，而每一个 sub-chain 拥有 1.35 个[9]块。鉴于此，每一个人的块环链大小如下：

sub-chain 平均大小 = (7,000,000 * 10 / 51,525,338) * 80 = 108.68B
main-chain 大小 = 51,525,338 * 80 / 1024^3 = 3.83G

与线性结构的块环链不同， X.Blockchain 可筛选管理所需数据。如有特殊原因，可对 100 万名特定人口管理居民登记手写本等。在此情况下，验证 100 万名人口的 10 年来居民登记手写本变更履历，所需的总存储大小如下：

3.83 + 108.68 * 1,000,000 / 1024^3 = 3.93

对于块环链的大小，逐年会增加相当于 100 万名的每年平均变更块的大小。
本文章没有考虑由于出生而 main-chain 增加以及由于死亡、结婚和离婚而 sub-chain 增加[10]。因此，实际上可能需要更大的全体块大小。此外，main-chain 的分类基准不一定是每一个人，而且一个块不一定只包含一个文件(Transaction)。所以，前面求得数值不具有实际性意义，只是为将线性结构的块环链和多层次结构的 X.Blockchain 做比较，具有相对值的意义。

前面例子是只适用于一种文件——居民登记手写本，但实际上，可同时适用于多种公共文件。追加的文件越多，与线性结构的块环链相比，多层次结构的块环链的效率越高。如假设在上面例子中追加一种其他公共机构的证明文件，而且变更履历等 Transaction 以类似于居民登记手写本类似比率发生，那么，对于线性结构的块环链，每次增加一种文件，需要约一倍的块。
但对于 X.Blockchain，main-chain 容量没有变化，而增加文件导致追加块，都只在 sub-chain 上进行，总之，多层次结构的 X.Blockchain 相对效率会提高。

[9] 是每一个人的 10 年来平均迁移次数。
[10] sub-chain 越增加，多层次结构的 X.Blockchain 效率越高。

Conclusion

如上述所说，全体块环链大小没有改变。多层次结构的 X.Blockchain 的主要优点是按特定基准，可筛选并管理数据(块）。用户可自己筛选并存储所要范围的块环链，所以在相关范围之内，没有可信任的第三方也可自己解决文件的可靠性问题。