

引用格式:顾浩,贺寰烨,林果园. 基于联盟链的云存储完整性审计机制研究[J]. 微电子学与计算机,2020,37(1):53-58. [GU H, HE H Y, LIN G Y. Research on cloud storage integrity auditing mechanism based on consortium blockchains[J]. Microelectronics & Computer, 2020, 37(1): 53-58.]

基于联盟链的云存储完整性审计机制研究

顾 浩¹, 贺寰烨¹, 林果园^{1,2}

(1 中国矿业大学 计算机科学与技术学院, 江苏 徐州 221116;

2 南京大学 软件新技术国家重点实验室, 江苏 南京 210093)

摘 要: 针对云存储完整性审计公正性问题, 提出一种基于联盟区块链的云存储完整性审计模型(CSACB, Cloud Storage Integrity Auditing Model Based on Consortium Blockchains). 首先, 该模型以树型结构描述审计联盟(AC, Audit Consortium)构成, 同时利用层级证书链(LCC, Layer Certificate Chain)对联盟成员进行身份标识和权限控制. 其次, 采用完整性审计链与动态操作链的双链形式支持可变云存储审计. 最后, 利用智能合约(SC, Smart Contract)并结合数据块标签索引机制构建公正的动态操作审计模型, 理论分析和实验结果表明该模型在安全性和性能上具备明显优势.

关键词: 云存储; 联盟区块链; 完整性审计; 动态操作

中图分类号: TP309.2

文献标识码: A

文章编号: 1000—7180(2020)01—0053—06

Research on cloud storage integrity auditing mechanism based on consortium blockchains

GU Hao¹, HE Huan-ye¹, LIN Guo-yuan^{1,2}

(1 School of Computer Science and Technology, CUMT, Xuzhou 221116, China;

2 State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

Abstract: Aiming at the issue of fairness in cloud storage integrity audit, a Cloud Storage Integrity Auditing Model Based on Consortium Blockchains (CSACB) is proposed. First of all, the model describes the composition of audit consortium (AC) by a tree structure, and uses the layer certificate chain (LCC) to identify and control the members of the alliance. Second, variable cloud storage auditing is supported by a two-chain form of integrity auditing chain and dynamic operating chain. Finally, the fair dynamic operation audit model is constructed by using the smart contract (SC) and the data block label index mechanism, and the results of theoretical analysis and experimental show that the model has obvious advantages in terms of security and performance.

Key words: cloud storage; consortium blockchains; integrity audit; dynamic operation

1 引言

随着数据时代的来临,云存储便捷和高效等优势愈发明显.但是,它将数据的拥有权与使用权分离,从而导致用户无法确定云端数据的完整性.数据

完整性是指云端数据和迁移前的本地数据保持一致的特性^[1-2].面对云存储完整性问题,审计云端数据既是保障用户数据安全也是对云信誉的认证.目前,国内外研究者已经取得了许多成果.文献[3]分析了云数据删除的证明状况并提出了一种云存储删除验

证协议,但其并未给出完整的审计方案.秦志光等^[4]采用分层索引进行动态操作审计,但其并未考虑证据安全性问题.Chen 等人^[5]描述了基于区块链技术的实时审计,论证了区块链之上的审计服务所具备的安全和公正性,但其并未给出具体实现方法.联盟区块链是一项新兴的面向多组织的区块链技术.多组织共同参与机制与云存储完整性审计有着相当好的契合度,不但可以增加审计的可信性和提高可追溯性,而且可以对于审计中隐私数据进行控制保护^[6].与此同时,将审计逻辑融入智能合约,利用区块链与云端分工并行计算证据以提升审计性能,并在合约中加入身份验证、权限管理、数据校验以增强审计的可信性与公正性.

2 相关知识

乘法循环群 G 可用二元组 (a, p) 描述,其中 a 为 G 的生成元, p 为 G 中的元素个数. G 中任意元素皆为 a 的乘方,即 $G = \{am \mid m \in Z_p\}$.函数 H 是一种可将数据映射到乘法循环群 G 中的哈希函数,即 $H(\{0,1\}^*) \rightarrow G$.

双线性映射可用四元组 (p, G_1, G_2, e) 描述,设 G_1 和 G_2 是两个素数阶乘法循环群,阶数为大质数 p ,映射关系 $e:G_1 \times G_1 \rightarrow G_2$,满足以下性质^[7]:

- (1) 双线性:对于任意 $a, b \in Z_p$ 和 $R, S \in G_1$,有 $e(R^a, S^b) = e(R, S)^{ab}$.
- (2) 非退化性:存在 $R, S \in G_1$,使得 $e(R, S) \neq 1_{G_2}$ (1_{G_2} 代表 G_2 的单位元).
- (3) 可计算性:存在有效的算法对任意的 $R, S \in G_1$,计算 $e(R, S)$ 的值.

3 CSACB

3.1 模型架构

CSACB 模型由云用户、云交互、区块生成与审计背书四方构成.云用户节点负责接收服务请求;云交互节点群负责联盟链与云存储端交互;区块生成节点群负责将审计打包成区块发至 AC 网络中同步;审计背书节点群负责执行云存储端证据验证和动态操作认证.

图 1 中,云用户通过云用户节点向审计背书节点群发起背书请求,背书的本质是对审计证据和动态操作进行征信.然后,各审计背书节点依据请求类型执行相应 SC,SC 负责验证证据或认证操作,再向云用户节点反馈结果.云用户节点依据征信结果,即背书的一致性,判定该次审计的有效性,进而决定是

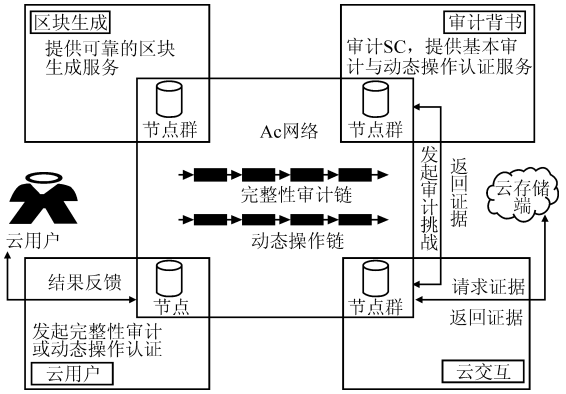


图 1 CSACB 模型架构

否向区块生成群发起打包请求以在相应区块链中永久存储.其中完整性审计链负责存储完整性审计信息,而动态操作链存储动态操作记录和认证结果.

3.2 联盟 LCC

CSACB 模型中,为了各组织在 AC 中有序地完成自身职能和防止节点跨越组织职能从而滥用权限,引入联盟 LCC 机制.联盟 LCC 以树型结构描述身份,然后以证书链验证身份.联盟 LCC 共包含两种映射.(1)节点身份到组织映射,即节点组织身份.(2)组织节点到 AC 映射,即节点联盟身份.

图 2 中,身份以根 R 衍生, R 被所有组织共同认可,其包含 3 部分, $[RC, Rpk, Rsk]$,即自签名证书 RC ,公钥 Rpk 以及私钥 Rsk . RC 不仅是根证书,也是身份标识, Rpk 用于验证 RC 使用 Rsk 签发的子证书有效性. RC 签发组织管理员(OM)证书 OMC 和联盟管理员(AM)证书 AMC . OM 负责组织身份生成,而 AM 负责联盟身份创建.

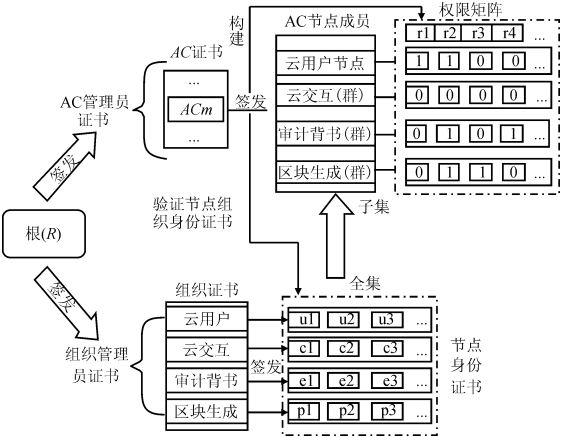


图 2 联盟 LCC

$OM = [OMC, Opk, Osk]$,其中 OM 证书 OMC 是用 Rsk 加密 RC 和随机序列 r 生成,即 $Rsk(RC, r) \rightarrow OMC$.而 Opk 和 Osk 分别是其公钥和

私钥, OM 利用 Osk 来生成每个组织组织证书,即云用户组织证书 UC 、云交互组织证书 IC 、审计背书组织证书 EC 和区块生成组织证书 PC ,同样的,它们拥有各自的密钥对。每个组织利用自身证书和私钥为节点签发证书。节点 u_x 到云用户组织的映射链为 $R \rightarrow OM \rightarrow Urg \rightarrow u_x$, u_x 证书 $u C_x$ 由层级证书链 LCC 构成。

$$u C_x = Usk(UC, Osk(OMC, Rsk(RC, r_1), r_2), r_3) \quad (1)$$

公式(1)中, Usk 为云用户组织私钥, $u C_x$ 由各级父证书层级运算生成,若要验证 u_x 是否为云用户组织成员,需对其 LCC 进行校验,校验中有一步未通过,则校验失败,过程如下。

(1) 使用云用户组织公钥 Usk 解密 $u C_x$, 提取首字段, 校验其有效性(该字段是否为真实 UC , 下同)。

(2) 使用 Opk 解密(1)中除 UC 以外字段, 校验首字段 OMC 有效性。

(3) 使用 Rpk 解密(2)中除 OMC 以外字段, 校验首字段 RC 有效性。

与 OM 一样, $AM = [AMC, Apk, Ask]$, 其中 AMC 是 AC 管理员身份证书, Ask 用于创建 AC_m 时签发其证书 ACC_m 。 AC_m 同样拥有私钥 A_msk 与公钥 A_mpk , 与组织不同的是, AC_m 不签发联盟中组织证书, 而是利用节点的组织身份证书生成节点的 AC 身份证书 $A_m u C_x$ 。

$$A_m u C_x = A_msk(ACC_m, u C_x) \quad (2)$$

公式(2)中, 二元组 $(ACC_m, u C_x)$ 可唯一标识联盟节点。验证 u_x 为 AC_m 云用户节点的过程如下。

(1) 使用 A_msk 解密 $A_m u C_x$ 获取二元组 $(ACC_m, u C_x)$ 。

(2) 使用 Apk 解密 ACC_m , 验证 AMC 有效性, 然后使用 Rpk 验证 RC 有效性。

(3) 校验 $u C_x$ 的 LCC , 判断其是否为云用户组织合法节点。

AC 中每种组织节点拥有各自的权限, 当 AC_m 签发联盟成员身份证书的同时会构造权限矩阵 CM_m , 该矩阵由 AC_m 使用自身证书和私钥进行签名。权限矩阵从上到下共 4 行, 分别对应联盟中云用户节点、云交互节点、审计背书节点和区块生成节点的权限, 而其每列则代表一种联盟区块链操作权限。

$$CM_m[i, j] = \begin{cases} 1, & \text{拥有权限} \\ 0, & \text{没有权限} \end{cases} \quad (3)$$

公式(3)中, $CM_m[i, j]$ 可取值 1 或 0, 当其取值

为 1 时, 表示 AC_m 中第 i 种组织节点拥有第 j 种权限, 反之则无该权限。

3.3 联盟双链与云存储审计

3.3.1 数据迁移验证

为保证审计公正, AC 提供完整性审计和动态操作认证之前, 云用户本地数据需经过 AC 验证后, 才可迁移至云端。云用户将文件 F_1 等分成 n 块 $(f_1, f_2, f_3, \dots, f_n)$, 取 t 为小于 p 的大正整数, 数据标签私钥 $S_{key} = t$, 公钥 $P_{key} = g^t$ 。数据块标签由文件标识符 FI (这里为 F_1)、数据块 f_i 、 f_i 摘要 $d(f_i)$ 、数据块版本号 v_i (初值为 1)、伪随机数 k_i 组成, 其中 $k_i \in R(S)$, R 为伪随机函数, S 是随机种子。数据块 i 的标签计算如下。

$$\text{Tag}f_i = \{H(F_1, v_i, i)g^{k_i f_i d(f_i)}\}^{S_{key}} \quad (4)$$

公式(4)中, 由于 $H(FI, v_i, i) \in G$, 同时 $g^{k_i f_i d(f_i)} \in G$, 因此 $\text{Tag}f_i \in G$ 。云用户在本地计算标签与数据块匹配证据 $\text{Tag}P = \sum_{i=1}^n k_i f_i d(f_i)$ 。数据块迁移验证合约参数为 $[F_1, \text{Tag}P, d(F_1), \text{Tag}f_1, \dots, \text{Tag}f_n]$, 其中 $d(F_1)$ 为整个文件摘要, 具体过程如下。

(1) 计算 $\text{left} = e(\prod_{i=1}^n \text{Tag}f_i, g)$, $\text{right} = e(g \text{Tag}P \prod_{i=1}^n H(F_1, v_i, i), P_{key})$ 。如果 $\text{left} = \text{right}$, 则请求云交互节点将 $[F_1, d(F_1), n]$ 发至云端, 否则验证失败。

(2) 云端请求云用户迁移数据。云用户向云端迁移数据。云端收到所有数据块后, 计算文件摘要, 使用 n 和 $d(F_1)$ 验证数据块个数与文件摘要。验证一致后, 向云交互节点响应迁移成功, 否则响应迁移失败。

(3) 审计背书节点获得云交互节点响应结果, 若然后将其则返回至云用户节点。

云用户节点验证征信结果。若结果一致且为迁移成功, 该验证参数及结果通过区块生成节点群生成区块并广播, AC 节点将该区块入本地动态操作链, 否则不生产区块。

3.3.2 完整性审计链与完整性审计

完整性审计 SC 实现对云数据完整性审计, 审计结果被永久保存在完整性审计链。云用户节点向 AC 发起完整性审计, 参数包含文件名 F_1 、随机种子 S_2, S_3 等。审计背书节点首先验证: (1) 发起者 LCC 有效性; (2) F_1 存在性。验证通过后, 然后执行完整

性审计 SC,图 3 中,该 SC 流程分为挑战生成、证据计算、证据验证三部分。

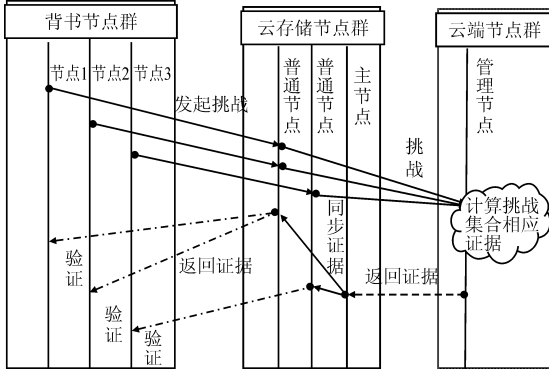


图 3 完整性审计合约流程

(1) 挑战生成. 各审计背书节点使用 $R(S_2)$ 生成同样的挑战集合 C , 其大小为 l . 紧接着选取云交互节点向云管理节点发起挑战, 审计背书节点待其返回挑战证据并在期间计算标签随机数 $t_i = R(S_3)_i (i = 1, \dots, l)$ 以及挑战标签证据 $\text{tag} = \prod_{i=1}^{\text{size}(C_g)} \text{tag}_{f_{C_g[i]}}$. 云交互主节点与云端数据交互以减轻云端响应压力。

(2) 证据计算. 云端首先验证 C 一致性, 验证通过后, 管理节点依据 C 收集数据块信息, 进而计算挑战证据. 证据公式如下。

$$E = \sum_{i=1}^l R(S)_{C_g[i]} R(S_3)_i f_{C[i]} d(f_{C[i]}) \quad (5)$$

公式(5)中, $R(S)_{C_g[i]}$ 为挑战集合中第 i 项所对应的数据块标签系数, $R(S_3)_i$ 为挑战集合中第 i 项系数, $f_{C[i]}$ 为挑战集合中第 i 项所对应的数据块, $d(f_{C[i]})$ 为其摘要. 云管理节点加密 E (审计背书节点共享密钥), 发至云交互主节点并在组织中同步. 普通云交互节点收到证据后, 响应审计背书节点请求。

(3) 证据验证. 审计背书节点收到挑战证据后, 解密获得 E . 证据验证方式如下。

$$e(\text{tag}, g) \rightarrow \begin{cases} = e(\prod_{i=1}^l H(F_1, v_{C[i]}, C[i])^{t_i} g^E, P_{\text{key}}) \\ \neq e(\prod_{i=1}^l H(F_1, v_{C[i]}, C[i])^{t_i} g^E, P_{\text{key}}) \end{cases} \rightarrow \begin{cases} \text{True} \\ \text{False} \end{cases}$$

如果返回 True, 则表示通过挑战. 各审计背书节点将结果返回给云用户节点, 验证征信结果. 如果一致, 则发送该审计至区块生成节点群以生成区块, 否则丢弃该次审计。

3.3.3 动态操作链与动态操作认证

动态操作 SC 完成对动态操作云数据认证并将认证结果存储于动态操作链. 数据块动态操作包括增加、更新、删除, 详情如下。

(1) 增加数据块是指云用户在原数据块间任意位置插入新数据块. 增加数据块认证参数包含操作类型 add、文件标识 F_1 、插入位置 q 、新数据块摘要 h 、标签随机系数 $k_{\frac{2q-1}{2}}$ 、新数据标签 $\text{tag}_{\frac{2q-1}{2}}$ 、伴随标签 $\text{ctag}_{\frac{2q-1}{2}}$ 、标签验证证据 $\text{tag}p = k_{\frac{2q-1}{2}}(f_{\frac{2q-1}{2}} + cf_{\frac{2q-1}{2}})$. 为确保新标签与新数据块匹配验证中不暴露数据块信息, 采用伴随数据块 $cf_{\frac{2q-1}{2}}$ 与新数据块聚合验证来隐藏新数据块. 伴随数据块是云用户对新数据块随机修改生成, 其标签 $\text{ctag}_{\frac{2q-1}{2}}$ 与 $\text{tag}_{\frac{2q-1}{2}}$ 使用同一系数. 审计背书节点收到请求后, 执行增加操作 SC. SC 认证过程中任意一项不通过, 则直接认证失败。

(1) SC 首先在标签索引中查询 F_1 数据块数, 再结合 q 判断请求插入位置是否越界。

(2) 数据块标签匹配验证, 即 $e(\text{tag}_{\frac{2q-1}{2}} v \text{tag}_{\frac{2q-1}{2}}, g)$ 如果与 $e(g^{\text{Tag}P}(H(F_1, 1, \frac{2q-1}{2}))^2, P_{\text{key}})$ 相等, 则选取云交互节点将新数据块信息 $[F_1, h, q, k_{\frac{2q-1}{2}}]$ 发送至云端。

(3) 云管理节点验证新数据块信息的一致性, 若验证一致则请求云用户发送新数据块 $f_{\frac{2q-1}{2}}$. 否则, 返回云交互节点验证失败。

(4) 云端收到新数据块后, 判断其摘要是否与 h 一致, 若一致则将数据块添加到指定位置, 返回添加结果。

(5) 审计背书节点获得响应, 若结果为成功, 则发起完整性审计挑战, 挑战集合为 $[q-1, q, q+1]$. 审计背书节点验证云端返回的挑战证据后, 将验证结果发送至云用户节点, 进而记录在 AC 中。

图 4 中, 新增加数据块后, 序号 q 变更为新增加数据块信息, 其后数据块序号依次递增, 实现了数据块在文件中位置的更新同时便于挑战数据块按序号抽取. 新数据块标签存储在 N_{new1} 区块中, 即包含新增数据块认证的区块, 而偏移则是新标签的存储位置。

(2) 更新数据块是指云用户以新数据块替换指定文件中指定数据块. 其 SC 与增加数据块 SC 逻辑相似, 不同的是待插入位置 q 变为待更新序号 u , 云端收到新数据块 u 后将旧数据块替换. 图 4 中, 数据块 u 原尾索引指向新的标签存储信息. 其中, 版本号变更为 2, 表示该数据块更新 1 次, 尾部索引为 null,

表示该版本为数据块最新版本. 如果再次对 u 进行更新, u 的索引将继续延伸, 其间所有版本可作为更

新操作记录.

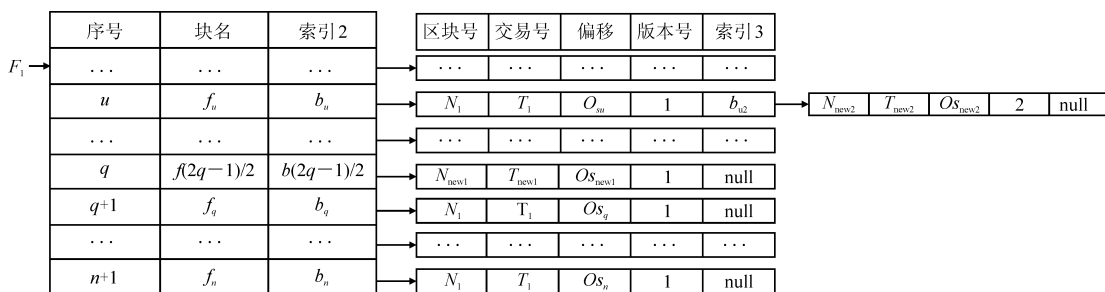


图 4 动态数据块标签索引表

(3) 删除数据块是指云用户删除指定文件中指定数据块. 数据块删除 SC 验证待删除数据块 m 有效性, 待云端响应删除后, 向其发起挑战, 挑战块集合为 $[m-1, m]$. 只有云端以删除前序号为 $m-1$ 和 $m+1$ 的数据块进行证据计算才能通过验证. 数据块 m 被删除后, 标签索引表中, 原序号 $m+1$ 的数据块变更为 m , 其下所有数据块序号增 1.

4 安全性与性能分析

4.1 安全性分析

完整性审计证据验证方法证明如下.

证明

$$\begin{aligned}
e(\text{tag}, g) &= e(\prod_{i=1}^l \text{tag} f_{C[i]} g) \\
&= e(\prod_{i=1}^l \{H(F_1, u_{C[i]}, Cg[i]) g^{k_{C[i]} f_{C[i]}^{d \leftarrow f_{C[i]}}} \}_{i}^{S_{key} t}, g) = \\
&e(g^{\sum_{i=1}^l k_{C[i]} f_{C[i]}^{d \leftarrow f_{C[i]}}} \prod_{i=1}^l H(F_1, v_{C[i]}, C[i])^{t_i}, g^{S_{key}}) \\
&= e(\prod_{i=1}^l H(F_1, v_{C[i]}, C[i])^{t_i} g^E, P_{key}).
\end{aligned}$$

证毕

下面

下面将从标签匹配、数据泄露、操作伪造 3 个方面分析本模型的安全优势:

(1) 标签匹配. CSACB模型中,本地数据块标签需经合约验证才可存入区块,如果恶意云用户使用不匹配标签来诬陷云端,将无法通过合约审查.另外,数据块标签存储于不可篡改的动态操作链中且对其访问权限控制,云端只需提供挑战指定数据块证据,无需再提供数据块标签.云端若想通过挑战,便不可再用其它数据块伪造挑战数据块.

(2) 数据泄露. 云端证据由随机系数、数据块以及数据块摘要构成, 大量挑战次数下, 恶意节点可能

通过线性方程获得每项数据块. 采用数据块摘要与数据块自身乘积聚合, 即 $f_i d(f_i)$, 由于联盟中没有存储其中任意一个因子, 因此无法解出数据块内容.

(3)操作伪造. 对于增加和更新操作,动态审计SC可确保云端正确执行操作,但对于删除操作,虽可确定云端将数据块从文件组织中删除,但不能确定其是否还在云端. 此行为目前是^[10]不可审计的,而利用动态操作链中操作记录永久存储且不可篡改的特性,可为数据块动态操作提供可追溯的证据.

4.2 性能分析

CSACB 模型无需第三方审计机构,相较于其他审计方案,本模型特征如表 1 所示.

表 1 模型特征

名称	匹配验证	访问控制	可信证明	历史追溯
刘忆宁等 ^[3]	否	否	支持	支持
秦志光等 ^[4]	否	否	否	否
CSACB	支持	支持	支持	支持

除表 1 所示的理论特征外,将搭建联盟区块链以测试 CSACB 模型在审计消耗性能.实验环境采用 OpenStack 云平台中虚拟节点搭建,配置如表 2 所示.乘法循环群与双线性映射函数 e 由椭圆曲线 MNT D159 构造,待审计文件包含 1 000 个数据块.实验采用逐渐增加挑战块数量的方式测试模型性能,每种结果取 20 次测量平均值.

表 2 实验环境配置

类型	CPU	内存	服务组件
联盟节点	至强 X5	8 GB	Fabric
云存储节点	至强 X5	4 GB	Cinder

图 5 中,本模型时间消耗包含 AC 和云端时间消耗两部分,AC 时间消耗主要包括智能合约运行时间、

交易认证时间以及区块生成同步时间,而云端时间消耗主要是证据生成和传输时间. AC 与云端时间消耗之和为审计串行总和,在任何挑战块下,总和时间都短于秦等审计方案,这是由于标签证据计算在 AC 中运用标签索引快速查询计算. 除此之外,数据块标签证据计算完毕直接在 AC 合约中使用,无需从云端传输至 AC. 需要注意的是,串行总和并非该模型真正的时间消耗,因为审计合约中数据块标签证据计算与云端证据计算采用并行化设计,这种并行化节约了大量的时间消耗,随着挑战数据块的增加,并行化所减少的时间消耗越多,当挑战块为全部数据块时,相较于秦等方案,时间缩短近 30%.

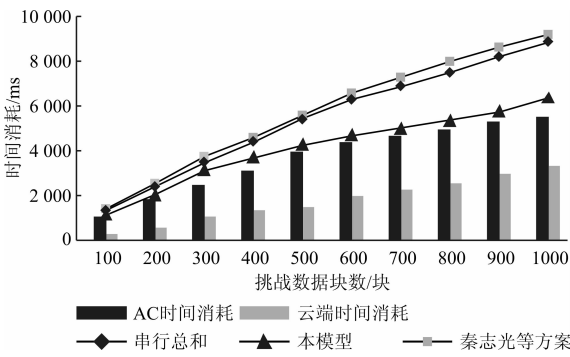


图 5 模型时间消耗

5 结束语

面对云存储审计的公正性与可靠性问题,提出一种基于联盟区块链的云存储完整性审计模型. 该模型首先在区块链联盟的各组织中实现身份和权限管理,有效地保障了审计数据的安全和合约的有序执行. 其次,联盟链智能合约和共识机制确保了审计和认证的公正性,并采用防篡改完整性审计链与动态操作链可信地存储了审计以及动态操作认证结果. 理论和实验证明其安全性与性能均得到提升.

参考文献:

[1] BABAK RAVANDI, IOANNIS PAPAPANAGIOTOU. A self-anized resource provisioning for cloud block storage[J]. Future Generation Computer Systems, 2018(89):765-776.

[2] YU J, WANG H. Strong key-exposure resilient auditing for secure cloud storage[J]. IEEE Trans. Inf. Forensics Security, 2017,12(8):1931-1940.

[3] 刘忆宁,周元健,蓝如师,等. 基于区块链的云数据删除验证协议[J]. 计算机研究与发展, 2018,55(10):2199-2207.

[4] 秦志光,王士雨,赵洋,等. 云存储服务的动态数据完整性审计方案[J]. 计算机研究与发展,2015,52(10):2192-2199.

[5] CHEN X, JI C H. Real-time audit research based on block chain technology[J]. Chinese Certified Public Accountant, 2017,1(4): 67-71.

[6] CHENG J, CHI C, CHEN Y. Blockchain and smart contract for digital certificate[C]//2018 IEEE International Conference on Applied System Invention (ICASI). Chiba, 2018:1046-1051.

[7] KWANGSU LEE, DONG HOON LEE. Two-Input functional encryption for inner products from bilinear maps[C]//IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2018, E101. A(6):915-928.

作者简介:

顾浩 男,(1993-),硕士研究生. 研究方向为云计算安全. Email:18052226151@163.com.

贺寰烨 男,(1995-),硕士研究生. 研究方向为信息安全.

林果园 男,(1975-),博士,副教授. 研究方向为云计算、信息安全.