



(51) International Patent Classification:

H04L 9/00 (2006.01) G06F 21/62 (2013.01)

G06F 16/23 (2019.01) H04L 9/06 (2006.01)

G06F 16/27 (2019.01) H04L 9/08 (2006.01)

(21) International Application Number:

PCT/CA2020/050006

(22) International Filing Date:

03 January 2020 (03.01.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/790,527 10 January 2019 (10.01.2019) US

(71) Applicant: BITNOBI, INC. [CA/CA]; MaRS Center,

West Tower, 661 University Ave., Suite 465, Toronto, Ontario M5G 1M1 (CA).

(72) Inventors: LITOIU, Marin; 266 St. Clements Ave.,

Toronto, Ontario M4R 1H5 (CA). SHTERN, Mark; 6010 Bathurst St., Apt. 605, Toronto, Ontario M2R 2H4 (CA).

(74) Agent: CHUMAK, Yuri; c/o Dickinson Wright LLP, 199

Bay Street, Suite 2200, Toronto, ON M5L 1G4 (CA).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH,

(54) Title: DISTRIBUTED GOVERNANCE FOR SHARING OF BIG DATA

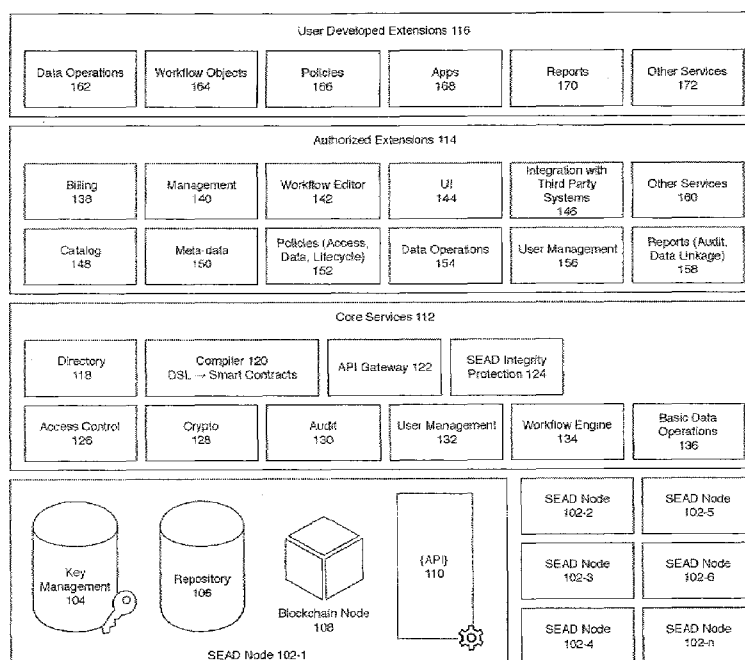


FIG. 1

(57) **Abstract:** Methods and systems for secure, encrypted and distributed ownership and usage of big data are provided. According to one example, a server maintains a local key management data store, a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy. A data operation from a user electronic device is received. The server verifies that the user electronic device has access against the local key management data store, runs the data operation and records metadata about the data operation, and writes data blocks to the data blockchain copy, the audit blockchain copy, and the metadata blockchain copy. The server broadcasts the updated blockchain copies to the peer-to-peer network for replication.



GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

- *with international search report (Art. 21(3))*
- *with amended claims and statement (Art. 19(1))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

DISTRIBUTED GOVERNANCE FOR SHARING OF BIG DATA

Field of the Invention

[0001] The field of the invention is cryptography and, in particular, a platform for secure, encrypted and distributed ownership, storage and usage of big data.

Background

[0002] The following description includes information that may be useful in understanding the present invention. It is not an admission that any of the information provided herein is prior art or relevant to the presently claimed invention, or that any publication specifically or implicitly referenced is prior art.

[0003] Many modern software applications rely heavily on big data, either as a core asset of their basic use cases or as supporting infrastructure used to improve quality of service. At the same time, privacy and security requirements constrain how the data can be stored and used in an accountable manner. Governing big data becomes even more challenging when it is shared among entities that desire shared responsibility for data ownership, storage and usage.

[0004] All publications herein are incorporated by reference to the same extent as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. Where a definition or use of a term in an incorporated reference is inconsistent or contrary to the definition of that term provided herein, the definition of that term provided herein applies and the definition of that term in the reference does not apply.

[0005] In some embodiments, the numbers expressing quantities of features used to describe and claim certain embodiments of the invention are to be understood as being modified in some instances by the term “about.” Accordingly, in some embodiments, the numerical parameters set forth in the written description and attached claims are approximations that can vary depending upon the desired properties sought to be obtained by a particular embodiment. In some embodiments, the numerical parameters should be construed in light of the number of reported significant digits and by applying ordinary rounding techniques. Notwithstanding that the numerical ranges and parameters setting forth the broad scope of some embodiments of the invention are approximations, the numerical values set forth in the specific examples are reported

as precisely as practicable. The numerical values presented in some embodiments of the invention may contain certain errors necessarily resulting from the standard deviation found in their respective testing measurements.

[0006] As used in the description herein and throughout the claims that follow, the meaning of “a,” “an,” and “the” includes plural reference unless the context clearly dictates otherwise. Also, as used in the description herein, the meaning of “in” includes “in” and “on” unless the context clearly dictates otherwise.

[0007] The recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g. “such as”) provided with respect to certain embodiments herein is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the invention.

[0008] Groupings of alternative elements or embodiments of the invention disclosed herein are not to be construed as limitations. Each group member can be referred to and claimed individually or in any combination with other members of the group or other elements found herein. One or more members of a group can be included in, or deleted from, a group for reasons of convenience and/or patentability. When any such inclusion or deletion occurs, the specification is herein deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

[0009] Thus, there is still a need for a system that allows for secure, encrypted and distributed ownership and usage of big data and to mitigate some of the obstacles to data sharing.

Brief Description of The Drawings

[0010] Various objects, features, aspects and advantages of the inventive subject matter will become more apparent from the following detailed description of preferred embodiments, along with the accompanying drawing figures in which like numerals represent like components.

[0011] FIG. 1 is a block diagram of a system architecture for secure, encrypted and distributed sharing of data in accordance with an example of the present specification;

[0012] FIG. 2A is a block diagram of a server for running data operations according to the system architecture of FIG. 1 in accordance with an example;

[0013] FIG. 2B is a schematic diagram of a system for secure, encrypted and distributed sharing of data in accordance with an example of the present specification;

[0014] FIG. 3 is a sequence diagram of a data operation running on the system of FIG. 2B in accordance with an example;

[0015] FIG. 4 is a sequence diagram of an audit operation running on the system of FIG. 2B in accordance with an example;

[0016] FIG. 5 is a sequence diagram of a data storage operation running on the system of FIG. 2B in accordance with an example;

[0017] FIG. 6 is a sequence diagram of a metadata operation running on the system of FIG. 2B in accordance with an example;

[0018] FIG. 7 is a sequence diagram of a blockchain analytics operation running on the system of FIG. 2B in accordance with an example;

[0019] FIG. 8 is a sequence diagram of an integrity protection operation running on the system of FIG. 2B in accordance with an example; and

[0020] FIG. 9 is a flowchart of a data operation, according to an example of the present specification.

Detailed Description

[0021] Throughout the following discussion, numerous references will be made regarding servers, services, interfaces, engines, modules, clients, peers, portals, platforms, or other systems formed from computing devices. It should be appreciated that the use of such terms is deemed to represent one or more computing devices having at least one processor (e.g., ASIC, FPGA, DSP, x86, ARM, ColdFire, GPU, multi-core processors, etc.) configured to execute software instructions stored on a computer readable tangible, non-transitory medium (e.g., hard drive, solid state drive, RAM, flash, ROM, etc.). For example, a server can include one or more computers operating as a web server, database server, or other type of computer server in a manner to fulfill described roles, responsibilities, or functions. One should further appreciate the disclosed algorithms, processes, methods, or other types of instruction sets can be embodied as a computer program product comprising a non-transitory, tangible computer readable media storing the instructions that cause a processor to execute the disclosed steps. The various servers, systems, databases, or interfaces can exchange data using standardized protocols or algorithms, possibly based on HTTP, HTTPS, AES, public-private key exchanges, web service APIs, known financial query protocols, or other electronic information exchanging methods. Data exchanges can be conducted over a packet-switched network, the Internet, LAN, WAN, VPN, or other type of packet switched network.

[0022] One should appreciate that the systems and methods of the inventive subject matter provide various technical effects, including providing secure, encrypted and distributed ownership and usage of big data.

[0023] The following discussion provides many example embodiments of the inventive subject matter. Although each embodiment represents a single combination of inventive elements, the inventive subject matter is considered to include all possible combinations of the disclosed elements. Thus if one embodiment comprises elements A, B, and C, and a second embodiment comprises elements B and D, then the inventive subject matter is also considered to include other remaining combinations of A, B, C, or D, even if not explicitly disclosed.

[0024] As used herein, and unless the context dictates otherwise, the term "coupled to" is intended to include both direct coupling (in which two elements that are coupled to each other contact each other) and indirect coupling (in which at least one additional element is located

between the two elements). Therefore, the terms "coupled to" and "coupled with" are used synonymously.

[0025] As used herein, ledger, or data repository, refers to a distributed record of transactions. A ledger may be implemented in a blockchain. A blockchain is a distributed network of peer-to-peer electronic devices that process and record transactions as part of a chain of blocks (blocks referring to electronic records or transactions). Once a block is completed, the block is added to the blockchain and the ledger is thereby updated. In many instances, a blockchain may be a ledger of transactions or data in chronological order or may be presented in any other order that is suitable for use by the blockchain network. In some configurations, transactions recorded in the blockchain may include a destination address and a currency or other fields, such that the blockchain records how much currency is attributable to a specific address. In many instances, additional information is captured, such as a time-stamp, source address, and the like.

[0026] As used herein, a digital asset is any electronic object such as an access control rule, or a data record, etc. Digital assets can be stored in a ledger. A smart contract is a data operation performed on a ledger. In this specification, virtual data, or derived data, or metadata, refers to data produced as a result of a smart contract (also referred to as an electronic contract) or a data operation. A metadata blockchain contains virtual data. A dataset, or data source, refers to a list of records or virtual records. Datasets can be defined by one or more schemas (e.g., a primary schema and optional schemas that may specify one or more data fields). Datasets and schemas can be stored in a data blockchain. An external data repository refers to data storage outside of the ledger. An external application refers to an application for consuming data from a ledger. An audit log contains records of modifications of data or modifications of data governance. An access control blockchain can be used to protect access to an audit log. Typically, data governance refers to the overall management of the availability, usability, integrity, and security of the data employed in an organization.

[0027] The specification is directed to a platform for secure, encrypted and distributed ownership, storage and usage of data, based on blockchain technology. The platform is general but in one example the platform offers significant advantages in an ecosystem of big data stakeholders that define data usage, storage and access based on electronic contracts.

[0028] The skilled reader will appreciate that many modern software applications rely heavily on big data, either as a core asset of their basic use cases or as supporting infrastructure to improve their quality of service. At the same time, privacy and security requirements including regulations mandated by different countries, such as HIPAA and GDPR, constrain how data can be stored and used in an accountable manner. Data governance becomes even more challenging when multiple stakeholders belong to a business ecosystem in which data is a core commodity that is exchanged and shared. In these ecosystems, to increase trust and transparency, it is beneficial to have transaction records jointly owned and stored.

[0029] According to examples of the present specification, blockchain technology is suitable as a transport layer to implement techniques of distributed transactions. A blockchain implements a distributed ledger, that is a distributed database, jointly operated by the parties or participants of the ecosystem or network. Transactions on the ledger are committed by electronic consensus, speeding up the technical operations of participants.

[0030] The present specification discloses a secure and auditable distributed data access and processing platform (“SEAD”). The platform uses blockchain as a data and transaction storage mechanism as well as a mechanism to define non-repudiable electronic contracts on data operations. The platform includes a data workflow definition and execution engine. In this example, a user electronic device specifies the data sources, the sequence of operations on data, as well as the access policies on data and operations. The SEAD system enables participants to: a) share data in a secure and auditable manner on a shared infrastructure; b) store and exchange data in a secure way; c) store and share metadata; d) audit the data access and the operations performed by different users; e) perform analytics on the recorded data and transaction in the blockchain; and f) integrate with public blockchain systems, among other functions.

[0031] Disclosed examples according to the present specification offer significant advantages. For example, the SEAD system gives and takes away control of data and operations and it provides full lineage of the data. The SEAD system can also provide lineage of the operations. The activities of a specific user can be replayed. The participants in the ecosystem can decide if the data is to be distributed or if the operations are to run on a specific participant’s infrastructure. In one example, the SEAD system is appropriate for Internet of Things (“IoT”) systems where there are many participants and small but many distributed data sets.

[0032] The term “big data” is generally used to describe collections of data of a relatively large size and complexity, such that the data becomes difficult to analyze and process within a reasonable time, given computational capacity (e.g., available database management tools and processing power). Thus, the term “big data” can refer to data collections measured in gigabytes, terabytes, petabytes, exabytes, or larger, depending on the processing entity’s ability to handle the data. As used herein, and unless the context dictates otherwise, the term “big data” is intended to refer to collections of data stored in one or more storage locations and can include collections of data of any size. Thus, unless the context dictates otherwise, the use of the term “big data” herein is not intended to limit the applicability of the inventive subject matter to a particular data size range, data size minimum, data size maximum, or particular amount of data complexity.

[0033] FIG. 1 illustrates the architecture of the SEAD system 100. The SEAD system 100 incorporates a distributed ledger (shown as blockchain nodes 108) and shares data among many participants by using a distributed data repository 106, across many SEAD nodes 102. FIG. 1 illustrates multiple SEAD nodes given by reference numbers 102-1, 102-3, 102-4, 102-5, 102-6, ... 102-n. According to one example, each SEAD node 102 resides under a different participant’s administrative domain. Each SEAD node 102 includes a blockchain node 108 and a key management data store 104 as well as a set of APIs 110. FIG. 1 shows the core services 112 of the SEAD system 100: a directory 118 provides an entry point to core basic operations that operates on digital assets stored in the data repository 106; an access control service 126 ensures that each participant accesses only specific digital assets from the data repository 106 and performs only the set of the operations that it has been given permissions to. With the SEAD system 100, selected data segments or digital assets can be encrypted, and copies can be distributed among selected participants using a cryptographic service 128; however, according to one example, the cryptographic keys can be managed by the SEAD system 100 to ensure that the digital assets in the data repository 106 are protected even if distributed to other SEAD nodes 102. The layer of core services 112 offers a basic data operations service 136 that provides atomic operations on the digital assets stored in the data repository 106 of the SEAD nodes 102. Higher level operations and actions, such as data transformations (e.g., for anonymization or privacy reasons) can be expressed in a domain specific language (“DSL”) and then compiled into blockchain specific electronic contracts using a compiler 120.

[0034] Still with reference to FIG. 1, according to one example of the present specification, the services of the SEAD system 100 can be extended by the user, shown as authorized extensions 114 and user developed extensions 116. Authorized extensions 114 include billing 138, that is, a crypto-currency service that can be configured to monetize access to data and operations, a user management module 140 that can be configured to integrate with enterprise identity management services. User developed extensions 116 include other data policies 166 and workflow object editors 164. Generally speaking, the user developed extensions 116 layer also allows for integration with external applications as well as external data repositories.

[0035] With continued reference to FIG. 1, the user management service 132 supports different roles for users including: a data creator, the owner of new imported data or the owner of a data job created from existing data; and an operation creator, the creator of a data operation. A data operation is any data processing operation such as a query, insert, join, and the like. A data operation has several properties: deterministic (generates the same result from the same input – true/false); open source (source code available and binaries can be regenerated on demand – true/false); operation creator (list of participants or actors that approved the data operation; price (optional). All data operations may have a clearly defined input and output. In one example, a data operation can be implemented as a model and can be executed as an electronic contract (blockchain smart contract) or an external application. Every execution of data operation can be recorded on the blockchain node 108 or ledger. Higher level data operations include

- a. Cryptographic transformation – action for encryption/decryption of digital assets
- b. Filters – action for removing digital assets
- c. Transformation – action for modifying digital assets (e.g. anonymization)
- d. Export – action for streaming digital assets to external application or system
- e. Query
- f. Data import – acting for streaming digital assets for data protection
- g. Save into data repository 106
- h. Save into blockchain node 108 or ledger

[0036] According to examples of the present specification, data operations can be run as electronic contracts that execute a sequence of data operations after receiving approval or permission from an access control blockchain (maintained as part of blockchain node 108). According to this example, after a user is granted permission from the access control blockchain, this information can be stored in the blockchain node 108. While the user's permission is not expired and the access control blockchain is not modified, the user can run operations without seeking subsequent permission. The operation can run following a lazy execution principle; the operation or transformation is executed if data should be exported and some data actions are non-deterministic, query or execution was forced by user. If all data operations are deterministic, transformed data will not be recorded into system except when a subset of the digital assets is marked for deletion; or a request for storing digital assets is explicit.

[0037] Still with reference to FIG. 1, the data repository 106 is considered system protected data storage. For information stored in the data repository 106, a metadata and integrity checksum can be stored on the blockchain nodes 108 (the ledger). Indexes, system information can be stored in the data repository 106 and the data is not replicated by default between different SEAD nodes 102. The replication can be configured without violating data governance. In one example, all digital assets in the data repository 106 can encrypted by default.

[0038] The key management store 104 is a local system deployed for each SEAD node 102 or participant. In one example, the key management store 104 is not a replicated dataset among the SEAD nodes 102. According to one example, the key management store 104 follows these definitions and procedures:

- a. Key guard – an electronic contract that protects encryption keys.
- b. Key guard family – a collection of key guards that protects the same set of keys. A key guard family assigns a key guard for the temporary giving of access to an encryption key.
- c. Key release process – when an access control rule provides a user with permission to access data, the key guard family is notified and requested to release a specific encryption key to a specific cryptographic transformation for data decryption. After the key guard notifies the key management store 104 and requests

temporary key release, the key management store 104 provides a URL to the key, and then the key guard can encrypt the URL with a public key of the corresponding cryptographic transformation.

- d. Key creation process – When encryption is required on a new dataset, a request is made of the key guard family for a new encryption key. The assigned key guard creates a new public/private key pair if it is a new dataset or new schema, otherwise it offers an existing public key. The private key is added to the key management store 104.
- e. Encryption keys are stored in the key management store 104. Every key is protected by a key guard family.

[0039] With continued reference to FIG. 1, the access control service 126 manages the access control rules and access control blockchain contained within blockchain node 108. As used herein, an access control rule is a condition on “link” data operations to data or virtual data (virtual data is the data that is generated as a result of data operations). An access control rule records all accesses to data into the blockchain node 108 (ledger). When an access control rule is attached to digital assets or data, the rule can automatically follow all data transformation(s) as long as the rule is not explicitly removed. An access control rule can be implemented as an electronic contract (smart contract). A data owner is allowed to add/remove access control rules. According to one example, all access control rules and access control rule operations can be permanently recorded on the blockchain node 108 (ledger). As used herein, an access control blockchain is a chain of access control rules that protects access to data. The access control blockchain can be contracted automatically for virtual (derived) data which is contracted using data actions.

[0040] In this specification, methods and systems for secure, encrypted and distributed ownership and usage of big data are provided.

[0041] FIG. 2A provides a block diagram of a server 204 for running data operations according to the system architecture of FIG. 1. The server 204 can include one or more computing devices programmed to perform the data operations and blockchain replication functions. Thus, the server 204 can include at least one processor 206, at least one non-transitory computer-readable

storage medium shown as memory 208 (e.g., RAM, ROM, flash drive, solid-state memory, hard drives, optical media, etc.) storing computer readable instructions that cause the processors to execute functions and processes of the inventive subject matter, and communication interfaces, such as network interface devices, that enable the server 204 to perform data exchanges with other servers 204 and to create a network of peer-to-peer servers 204. The computer-readable instructions (shown as OS 210 and Programs 212) that the server 204 uses to carry out its functions can be instructions allowing the server 204 to access, retrieve, and process data operations to authorized parties, access control functions, blockchain replication, etc. The server 204 can include input/output interfaces 216 (e.g., keyboard, mouse, touchscreen, display 218, sound output devices, microphones, sensors, etc.) that allow an administrator or other authorized user to enter information into and receive output from the server 204. Examples of suitable computing devices for use as a server 204 can include server computers, desktop computers, laptop computers, tablets, smartphones, smartwatches, wearables, IoT devices, etc.

[0042] FIG. 2B is a schematic overview of a networked ecosystem 200 for secure, encrypted and distributed sharing of data. The ecosystem 200 includes one or more nodes 202 (which can be the servers 204 of FIG. 2A) that maintain SEAD nodes 102 (shown as SEAD nodes 102-1, 102-2 and 102-3 in FIG. 1). Each of the components of the ecosystem 200 (shown as the nodes 202-1, 202-2, 202-3, 202-4, 202-5, 202-6, 202-7, 202-8 in FIG. 2) can be communicatively coupled with each other via one or more data exchange networks (e.g., Internet, cellular, Ethernet, LAN, WAN, VPN, wired, wireless, short-range, long-range, etc.).

[0043] A user interface of the server 204 can access big data within the SEAD nodes 102. The user interface can comprise one or more computing devices that enables a user or data analyst to access data from SEAD nodes 102 by creating and submitting query jobs.

[0044] Flowcharts illustrating examples of data operations are shown in FIG. 3 to FIG. 9. These operations or methods can be carried out by applications or software executed by, for example, the processor of the server 204. The method can contain additional or fewer processes than shown and/or described, and can be performed in a different order. Computer-readable code executable by at least one of the processors to perform the method can be stored in a computer-readable storage medium, such as a non-transitory computer-readable medium.

Secure Basic Data Operations

[0045] With reference to FIG. 3, the electronic device of a user 302 accesses and modifies a workflow object stored in the ledger or blockchain node 108. As used herein, a workflow is a program that defines how data is processed. The sequence of operations involved in the modification of a workflow process is shown in FIG. 3.

Audit Operations

[0046] In FIG. 4, the SEAD system 100 can be configured to store in an audit blockchain, all data operations, the input and outputs of each operation, and the workflows, together with user information. At the same time, it logs all changes in the system such as who changed/created an operation. Only authorized users will be able to view data in the audit blockchain. FIG. 4 shows a sequence of operations involved in this scenario up to the Basic Data Operation service. From there on, the calls go through the Access Control, Crypto and Blockchain services (similar to FIG. 3).

Data Storage Operations

[0047] With reference to FIG. 5, the commodity used by the workflow editors is data, users can configure SEAD to store data records on data blockchain. When a data sharing platform receives new or generates derived data, it can store it on the data blockchain. To protect data, all information can be stored with encryption. FIG. 5 shows the sequence diagram of this use case.

Metadata Operations

[0048] As shown in FIG. 6, the SEAD system 100 can be configured to store data sources (raw or derived) in a metadata blockchain.

Analytics Operations

[0049] As shown in FIG. 7, the SEAD system 100 can be configured to consume data from the audit, data and metadata blockchains and then to run analytic jobs on top of data extracted from the blockchain node 108. To increase performance, data can be offloaded into local storage of

the server 204 and then processed locally using a hosting framework for workflow editing and execution. The results can be stored back on the blockchain node 108 or cryptographically sealed and the cryptographic signature will be stored on the blockchain node 108.

Integrity Operations

[0050] One problem with having a small ecosystem of blockchain nodes is that if one participant controls the majority of the blockchain nodes in a blockchain platform, that participant can modify the blocks. To avoid that, the SEAD system 100 can be modified whereby a number of blocks or transactions must be signed by an external system. The signature can be stored externally and available to all participants (FIG. 8).

[0051] With reference to FIG. 9, a method 900 starts at 905 and, at 910, the server 204 maintains a SEAD node 102. At 915, the server 204 receives a data operation. At 920, the server 204 verifies access from the key management store 104 of the SEAD node 102. At 925, 930, 935, and 940, the access blockchain, data blockchain, audit blockchain, and metadata blockchains are updated, respectively. At 945, a cryptographic transformation is performed. At 950, the data operation is completed, and the method returns to monitor for data operations at 915.

[0052] A system of one or more computers can be configured to perform particular operations or actions by virtue of having software, firmware, hardware, or a combination of them installed on the system that in operation causes or cause the system to perform the actions. One or more computer programs can be configured to perform particular operations or actions by virtue of including instructions that, when executed by data processing apparatus, cause the apparatus to perform the actions.

[0053] One general aspect includes a method including the steps of: maintaining a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein each blockchain copy comprises data blocks that are linked and secured using cryptography, receiving a data operation from a user electronic device, verifying that the user electronic device has access against the local key data management store, responsive to the verifying, approving the data operation, after the data operation has been approved, running the data operation and recording metadata about the data operation; updating the plurality of blockchain copies comprising writing first data

blocks with the results of the data operation to the data blockchain copy, writing second data blocks with details of the user electronic device to the audit blockchain copy, and writing third data blocks with the recorded metadata to the metadata blockchain copy, and time-stamping and broadcasting the updated plurality of blockchain copies to the peer-to-peer network for replication.

[0054] Implementations may include one or more of the following features: the verifying includes requesting a specific encryption key for decrypting data associated with the data operation, receiving the specific encryption key as a key URL, generating an encrypted key URL by encrypting the key URL for providing temporary access to the specific encryption key, and providing the encrypted key URL to the user electronic device. The data operations can include action for removing data, for modifying data, for exporting data to an external application, for querying data, for importing data from an external application, for saving into a local data repository, and for saving into a blockchain node. The data operations can include modifying a workflow object stored in the data blockchain copy. Details of the user electronic device can be written to the audit blockchain copy and includes writing the data operations, an input and an output of the data operation, one or more workflows associated with the data operation, and user information. The data operation can include derived data from a data repository. The method can include performing blockchain replication activities, and before replicating a given block or a given transaction, verifying that the given block or the given transaction has been signed by a trusted external system. The method can include a public data blockchain copy and the method can include reading a checksum from the data blockchain copy, writing fourth data blocks with the checksum to the public data blockchain copy, and receiving a confirmation of the writing of the fourth data blocks to verify the integrity of the data blockchain copy. The data operation can include a data analytics operation that consumes data from the audit blockchain copy, the data blockchain copy and the metadata blockchain copy and further includes running a data analytics job using the consumed data.

[0055] One general aspect includes at least one non-transitory computer-readable storage medium storing instructions that, when executed by at least one processor, cause the at least one processor to maintain a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein each blockchain copy comprises data blocks that are linked and secured using

PAT2186US00

cryptography, receive a data operation from a user electronic device, verify that the user electronic device has access against the local key management data store, responsive to the verifying, approving the data operation, after the data operation has been approved, run the data operation and recording metadata about the data operation, update the plurality of blockchain copies comprising writing first data blocks with the results of the data operation to the data blockchain copy, writing second data blocks with details of the user electronic device to the audit blockchain copy, and writing third data blocks with the recorded metadata to the metadata blockchain copy, and time-stamp and broadcast the updated plurality of blockchain copies to a peer-to-peer network for replication.

[0056] It should be apparent to those skilled in the art that many more modifications besides those already described are possible without departing from the inventive concepts herein. The inventive subject matter, therefore, is not to be restricted except in the spirit of the appended claims. Moreover, in interpreting both the specification and the claims, all terms should be interpreted in the broadest possible manner consistent with the context. In particular, the terms “comprises” and “comprising” should be interpreted as referring to elements, components, or steps in a non-exclusive manner, indicating that the referenced elements, components, or steps may be present, or utilized, or combined with other elements, components, or steps that are not expressly referenced. Where the specification claims refers to at least one of something selected from the group consisting of A, B, C and N, the text should be interpreted as requiring only one element from the group, not A plus N, or B plus N, etc.

CLAIMS

What is claimed is:

1. A method comprising the steps of:

at a server comprising a processor, a memory, and a network interface device connected to a peer-to-peer network,

maintaining a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein the blockchain copies comprise data blocks that are linked and secured using cryptography;

receiving a data operation from a user electronic device;

verifying that the user electronic device has access against the local key management data store;

responsive to the verifying, approving the data operation;

after the data operation has been approved, running the data operation and recording metadata about the data operation to generate results of the data operation;

updating the plurality of blockchain copies comprising writing first data blocks with the results of the data operation to the data blockchain copy, writing second data blocks with details of the user electronic device to the audit blockchain copy, and writing third data blocks with the recorded metadata to the metadata blockchain copy; and

time-stamping and broadcasting the updated plurality of blockchain copies to the peer-to-peer network for replication.

2. The method of claim 1 wherein the verifying further comprises:

requesting a specific encryption key for decrypting data associated with the data operation;

receiving the specific encryption key as a key URL;

generating an encrypted key URL by encrypting the key URL for providing temporary access to the specific encryption key; and

providing the encrypted key URL to the user electronic device.

3. The method of claim 1 wherein the data operations comprise action for removing data, for modifying data, for exporting data to an external application, for querying data, for importing data from the external application, for saving into a local data repository, and for saving into a blockchain node.

4. The method of claim 1 wherein the data operation comprises modifying a workflow object stored in the data blockchain copy.

5. The method of claim 1 wherein writing second data blocks with details of the user electronic device to the audit blockchain copy comprises writing the data operation, an input and an output of the data operation, one or more workflows associated with the data operation, and user information.

6. The method of claim 1 wherein data operations comprises deriving data from a data repository.

7. The method of claim 1 further comprising:

in communication with a network of peer-to-peer servers, performing blockchain replication activities; and

before replicating one of a given block and a given transaction, verifying that the one of a given block and the given transaction has been signed by a trusted external system.

8. The method of claim 1 wherein the plurality of blockchain copies further comprises a public data blockchain copy and the updating the plurality of blockchain copies further comprises:

reading a checksum from the data blockchain copy;

writing fourth data blocks with the checksum to the public data blockchain copy; and

receiving a confirmation of the writing of the fourth data blocks to verify the integrity of the data blockchain copy.

9. The method of claim 1 wherein the data operation comprises a data analytics operation that consumes data from the audit blockchain copy, the data blockchain copy and the metadata blockchain copy and further comprising running a data analytics job using the consumed data.

10. At least one non-transitory computer-readable storage medium storing instructions that, when executed by at least one processor, cause the at least one processor to:

maintain a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein the blockchain copies comprise data blocks that are linked and secured using cryptography;

receive a data operation from a user electronic device;

verify that the user electronic device has access against the local key management data store;

responsive to the verifying, approve the data operation;

after the data operation has been approved, run the data operation and record metadata about the data operation to generate results of the data operation;

update the plurality of blockchain copies comprising write first data blocks with the results of the data operation to the data blockchain copy, write second data blocks with details of the

user electronic device to the audit blockchain copy, and write third data blocks with the recorded metadata to the metadata blockchain copy; and

time-stamp and broadcast the updated plurality of blockchain copies to a peer-to-peer network for replication.

AMENDED CLAIMS
received by the International Bureau on 25 May 2020 (25.05.2020)

CLAIMS

What is claimed is:

1. A method comprising the steps of:

at a server comprising a processor, a memory, and a network interface device connected to a peer-to-peer network,

maintaining a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein the blockchain copies comprise data blocks that are linked and secured using cryptography;

receiving a data operation comprising action for anonymizing digital assets stored in the data blockchain from a user electronic device;

verifying that the user electronic device has access against the local key management data store and verifying that the data operation follows access control rules;

responsive to the verifying, approving the data operation;

after the data operation has been approved, running the data operation and recording metadata about the data operation to generate results of the data operation;

updating the plurality of blockchain copies comprising writing first data blocks with the results of the data operation to the data blockchain copy, writing second data blocks with details of the user electronic device to the audit blockchain copy, and writing third data blocks with the recorded metadata to the metadata blockchain copy; and

time-stamping and broadcasting the updated plurality of blockchain copies to the peer-to-peer network for replication.

2. The method of claim 1 wherein the verifying further comprises:

requesting a specific encryption key for decrypting data associated with the data operation;

receiving the specific encryption key as a key URL;

generating an encrypted key URL by encrypting the key URL for providing temporary access to the specific encryption key; and

providing the encrypted key URL to the user electronic device.

3. The method of claim 1 wherein the data operations comprise action for removing data, for modifying data, for exporting data to an external application, for querying data, for importing data from the external application, for saving into a local data repository, and for saving into a blockchain node.

4. The method of claim 1 wherein the data operation comprises modifying a workflow object stored in the data blockchain copy.

5. The method of claim 1 wherein writing second data blocks with details of the user electronic device to the audit blockchain copy comprises writing the data operation, an input and an output of the data operation, one or more workflows associated with the data operation, and user information.

6. The method of claim 1 wherein data operations comprises deriving data from a data repository.

7. The method of claim 1 further comprising:

in communication with a network of peer-to-peer servers, performing blockchain replication activities; and

before replicating one of a given block and a given transaction, verifying that the one of a given block and the given transaction has been signed by a trusted external system.

8. The method of claim 1 wherein the plurality of blockchain copies further comprises a public data blockchain copy and the updating the plurality of blockchain copies further comprises:

reading a checksum from the data blockchain copy;

writing fourth data blocks with the checksum to the public data blockchain copy; and

receiving a confirmation of the writing of the fourth data blocks to verify the integrity of the data blockchain copy.

9. The method of claim 1 wherein the data operation comprises a data analytics operation that consumes data from the audit blockchain copy, the data blockchain copy and the metadata blockchain copy and further comprising running a data analytics job using the consumed data.

10. At least one non-transitory computer-readable storage medium storing instructions that, when executed by at least one processor, cause the at least one processor to:

maintain a local key management data store and a plurality of blockchain copies comprising a data blockchain copy, an audit blockchain copy, and a metadata blockchain copy, wherein the blockchain copies comprise data blocks that are linked and secured using cryptography;

receive a data operation comprising action for anonymizing digital assets stored in the data blockchain copy from a user electronic device;

verify that the user electronic device has access against the local key management data store and verify that the data operation follows access control rules;

responsive to the verifying, approve the data operation;

after the data operation has been approved, run the data operation and record metadata about the data operation to generate results of the data operation;

update the plurality of blockchain copies comprising write first data blocks with the results of the data operation to the data blockchain copy, write second data blocks with details of the user electronic device to the audit blockchain copy, and write third data blocks with the recorded metadata to the metadata blockchain copy; and

time-stamp and broadcast the updated plurality of blockchain copies to a peer-to-peer network for replication.

STATEMENT ACCOMPANYING AMENDMENT OF CLAIMS UNDER ARTICLE 19
PCT/CA2020/050006 filed January 3, 2020

Claims 1-10 are replaced by amended claims 1-10. Claims 1 and 10 are amended.

Claims 1 and 10 are amended to clarify the claimed invention. In particular, claim 1 is amended to clarify that the verifying step also verifies the data operation and the data operation comprises action for anonymizing data. The basis for amendment can be found at least in paragraphs [0033], [0035] and [0039] of the Applicant's original specification.

The amendment distinguishes the reference cited in the International Search Report of March 23, 2020, that is, US 2017/0163733 to Grefen et al. ("Grefen"). Grefen discloses recording data from a number of devices in a distributed network system in a manner adaptable for auditing the device output. However, Grefen does not disclose or suggest the technique of action for anonymizing data or verifying the data operation as recited in claims 1 and 10 as amended. As far as Applicant understands, the solution of Grefen requires sensor data to be stored in "two locations" (para. [0095], i.e., content and stripped content), rather than permitting data operations including action for anonymizing data to be executed (and verified against access rules prior to execution) in respect of digital assets stored in a data blockchain.

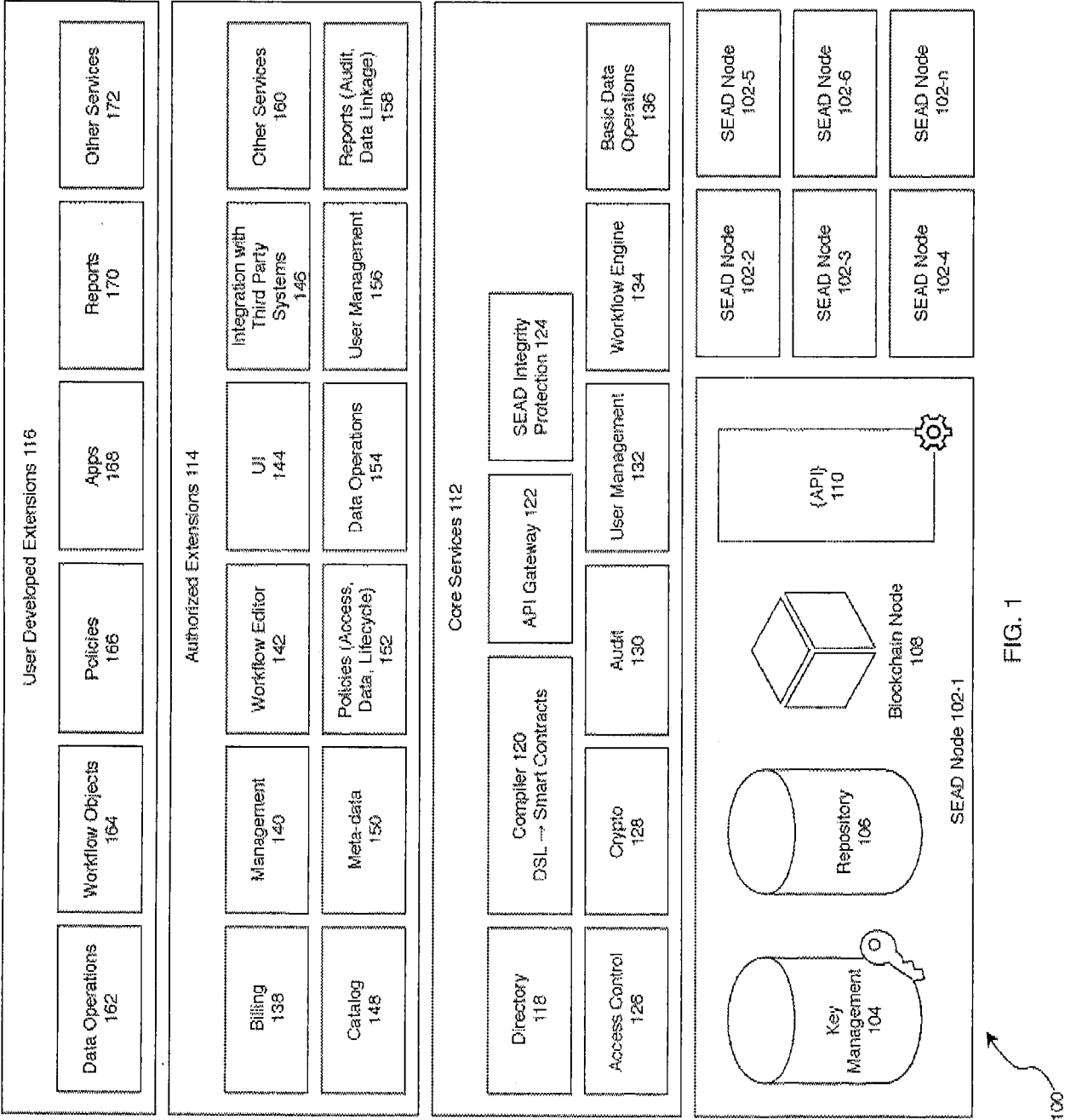


FIG. 1

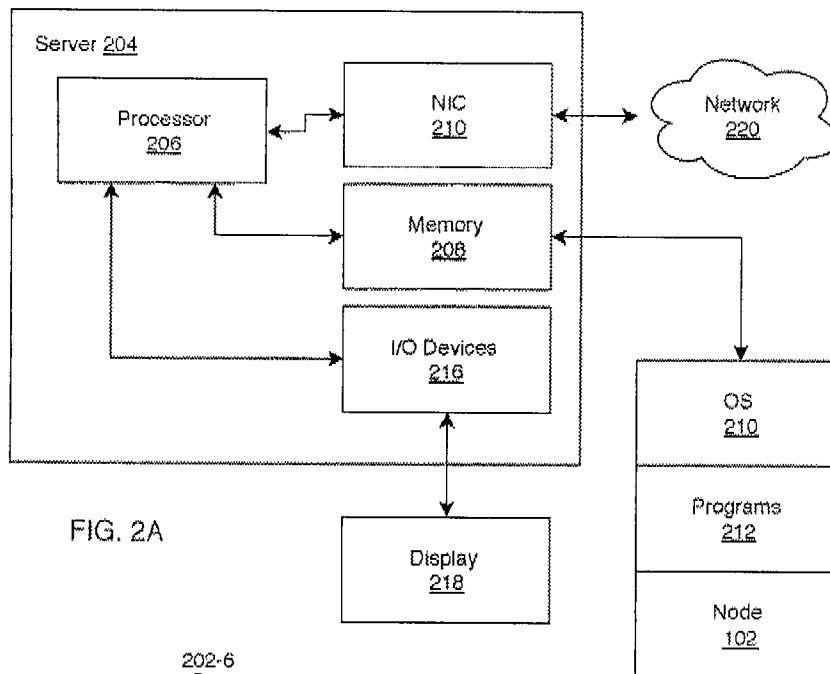


FIG. 2A

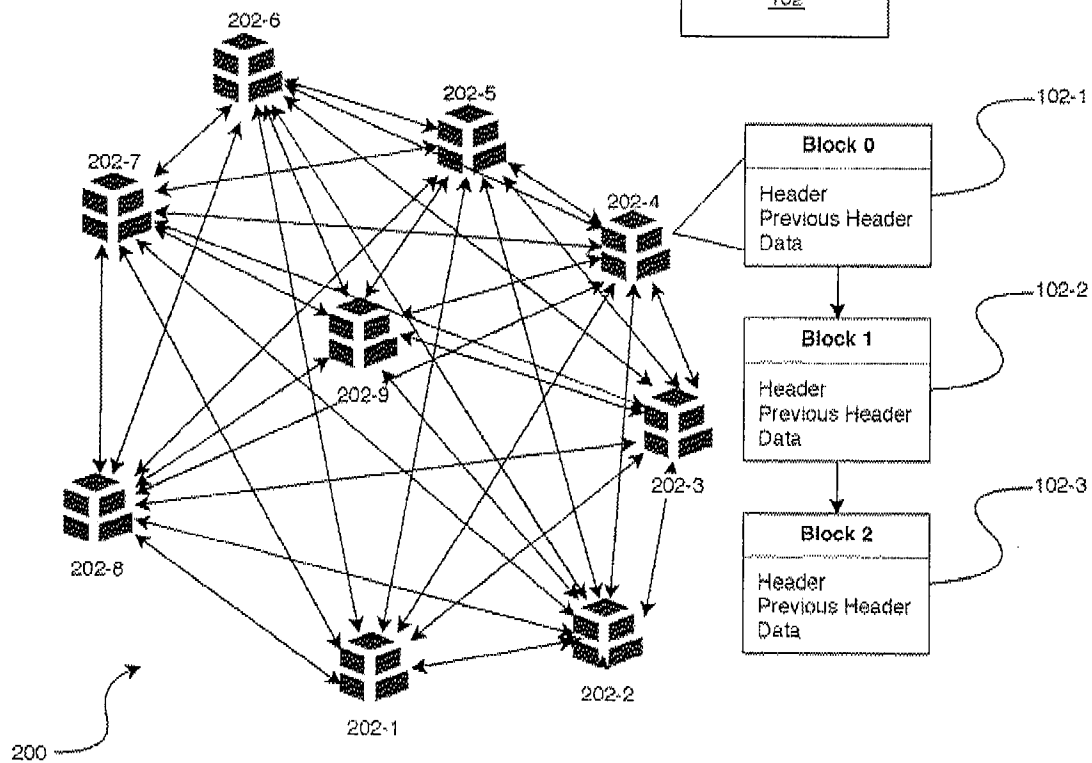


FIG. 2B

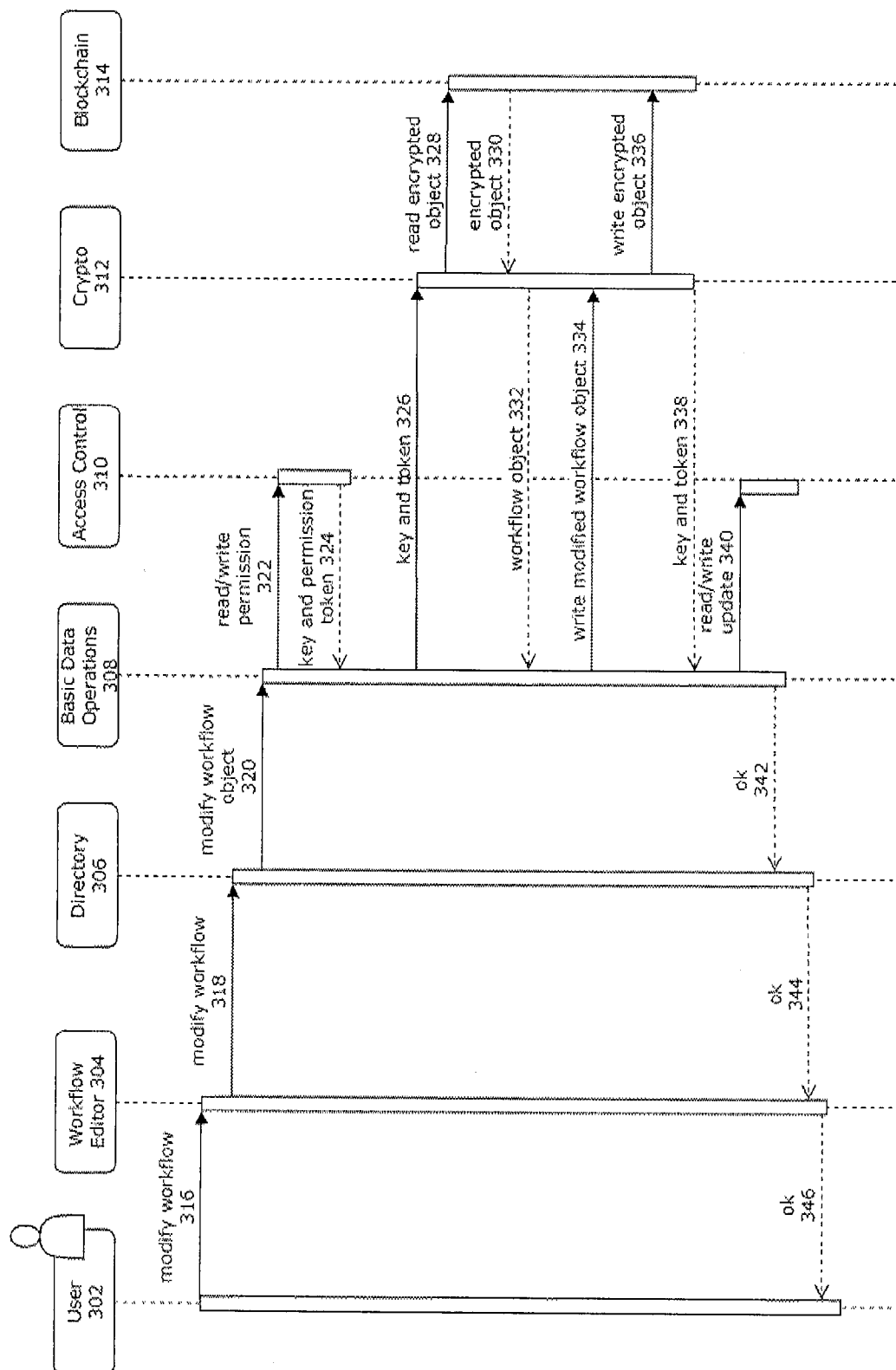


FIG. 3

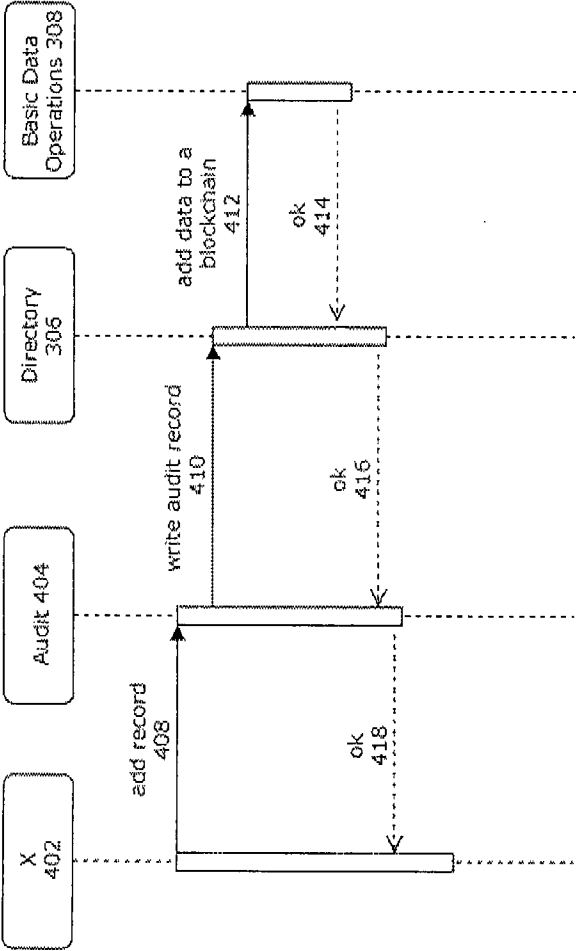


FIG. 4

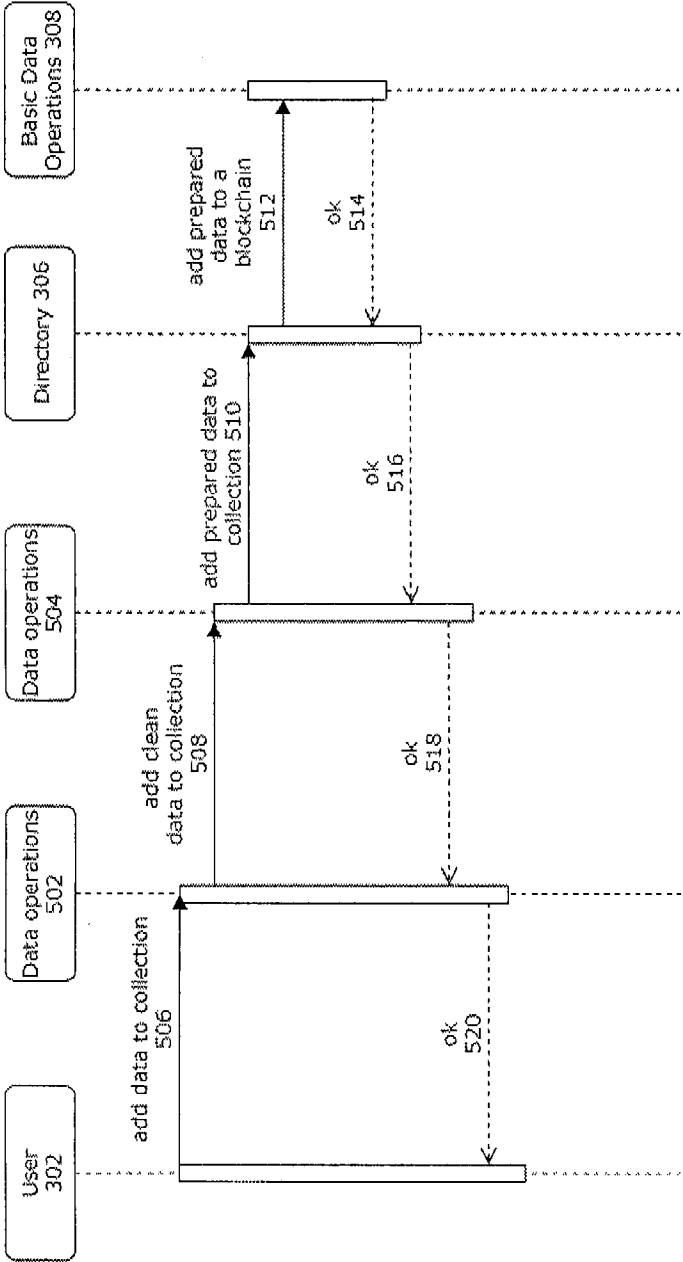


FIG. 5

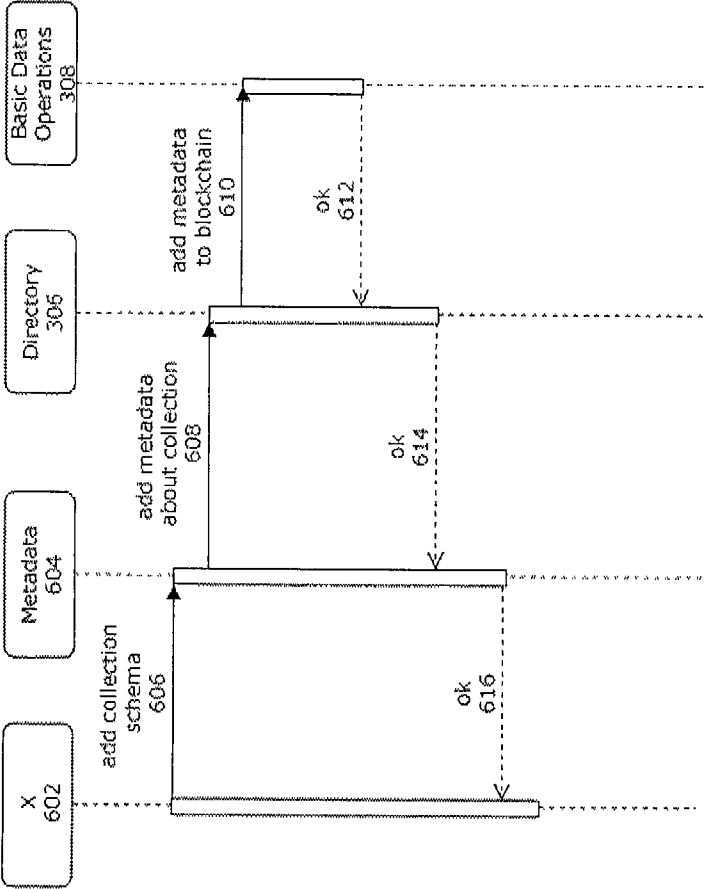


FIG. 6

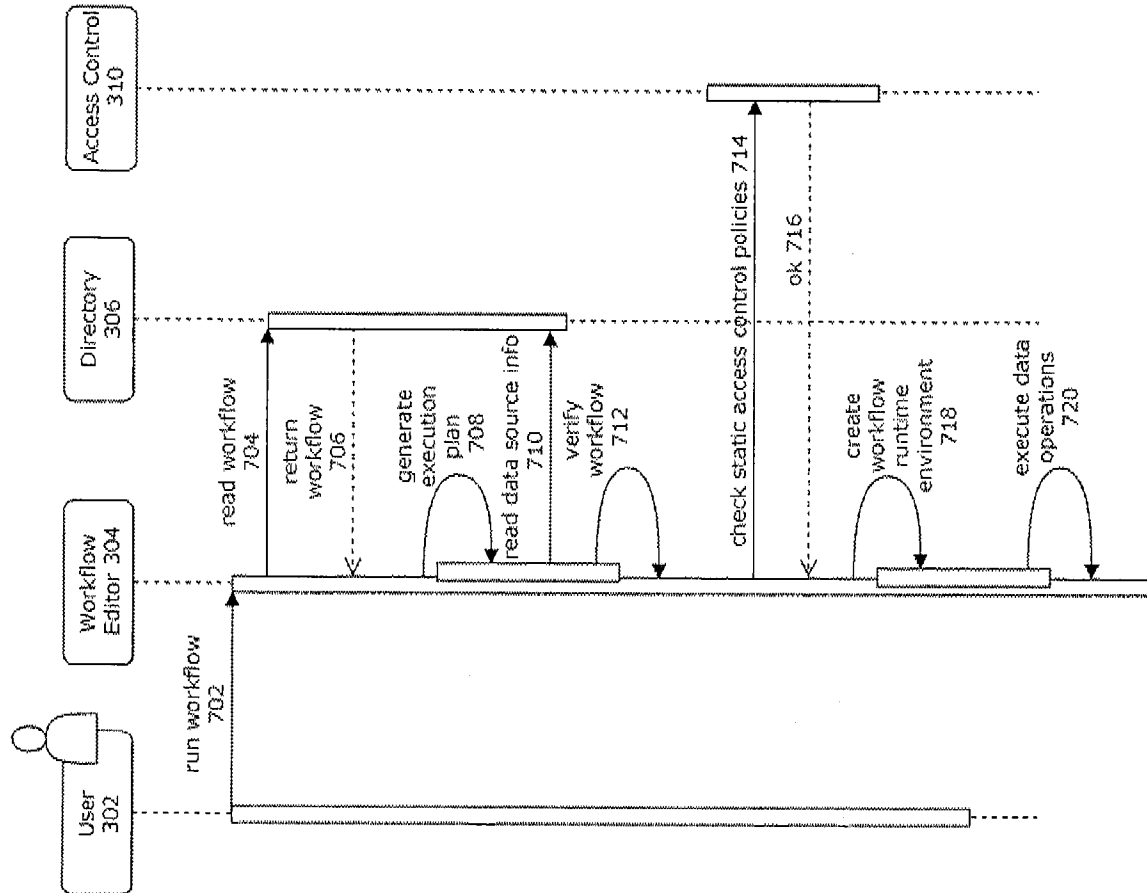


FIG. 7

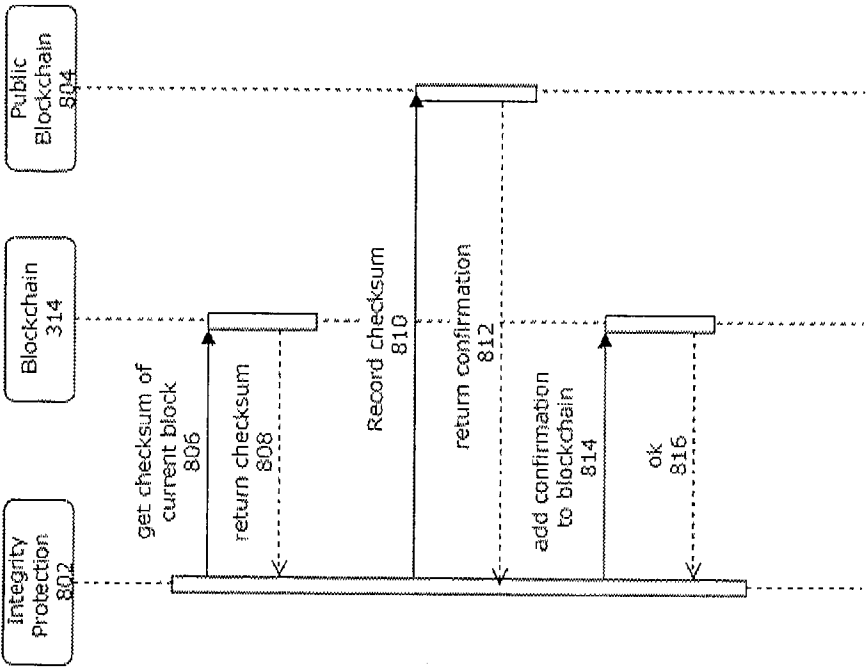


FIG. 8

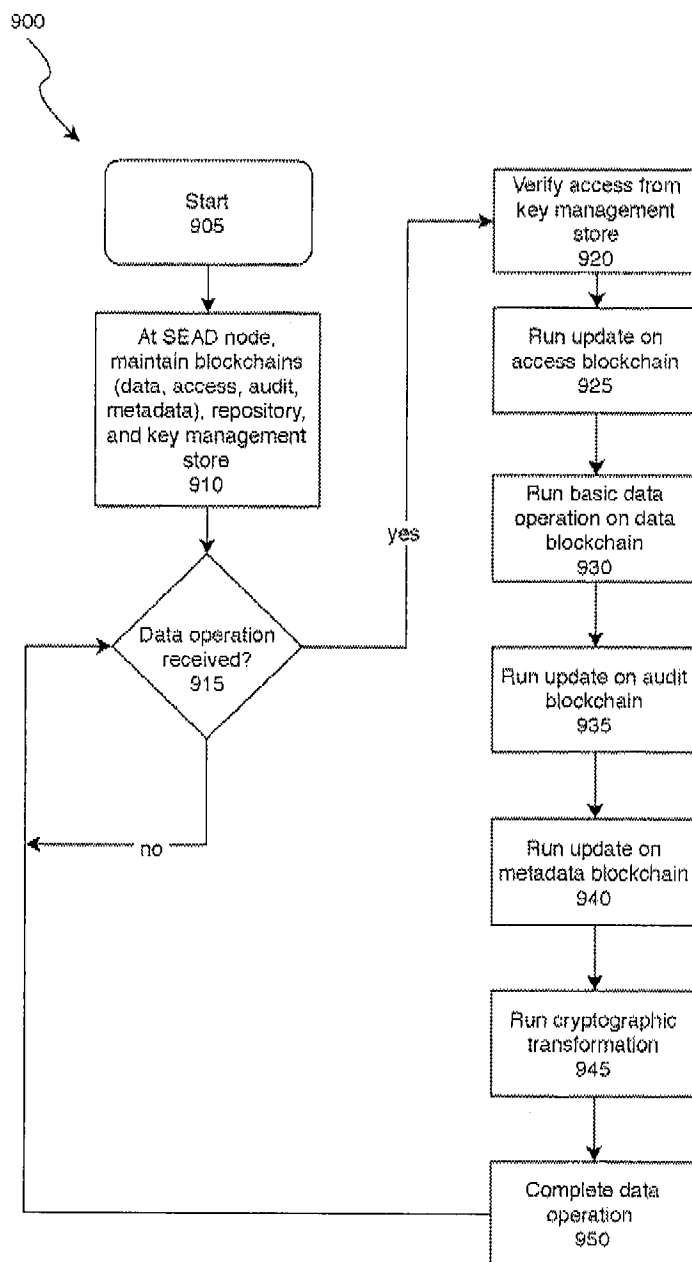


FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA2020/050006

A. CLASSIFICATION OF SUBJECT MATTER

IPC: **H04L 9/00** (2006.01), **G06F 16/23** (2019.01), **G06F 16/27** (2019.01), **G06F 21/62** (2013.01),
H04L 9/06 (2006.01), **H04L 9/08** (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: **H04L 9/00** (2006.01), **G06F 16/23** (2019.01), **G06F 16/27** (2019.01), **G06F 21/62** (2013.01), **H04L 9/06** (2006.01), **H04L 9/08** (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)

Questel Orbit

Search terms: blockchain, copy, audit, checksum, encryption, URL, key, workflow object

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	D1 US 2017/0163733 A1, (GREFEN, S. et al.), 8 June 2017 (08-06-2017) *See entire document.*	1, 3, 5-7, 9-10 2, 4, 8
Y	D2 US 2015/0278820 A1, (MEADOWS, M.), 1 October 2015 (01-10-2015) *See at least paragraphs 0073-0074*	2
Y	D3 WO 2018/170321 A1, (WODETZKI, J. et al.), 20 September 2018 (20-09-2018) *See at least paragraphs 0016 and 0079.*	4
Y	D4 CA 3051065 A1, (QUI, H.), 23 August 2018 (23-08-2018) *See at least Fig. 1, 2, and 4.*	8

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* "A" "D" "E" "L" "O" "P"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance document cited by the applicant in the international application earlier application or patent but published on or after the international filing date document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed	"I" "X" "Y" "&"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art document member of the same patent family
---	--	--------------------------	--

Date of the actual completion of the international search
23 March 2020 (23-03-2020)

Date of mailing of the international search report
23 March 2020 (23-03-2020)

Name and mailing address of the ISA/CA
Canadian Intellectual Property Office
Place du Portage I, C114 - 1st Floor, Box PCT
50 Victoria Street
Gatineau, Quebec K1A 0C9
Facsimile No.: 819-953-2476

Authorized officer

Orkunalp Erkan (873) 354-9788

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CA2020/050006

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
US2017163733A1	08 June 2017 (08-06-2017)	WO2018102749A1	07 June 2018 (07-06-2018)
US2015278820A1	01 October 2015 (01-10-2015)	CA2980707A1 SG10201808013UA SG11201707861UA WO2015148725A2 WO2015148725A3	01 October 2015 (01-10-2015) 30 October 2018 (30-10-2018) 30 October 2017 (30-10-2017) 01 October 2015 (01-10-2015) 24 November 2016 (24-11-2016)
WO2018170321A1	20 September 2018 (20-09-2018)	US2018268506A1	20 September 2018 (20-09-2018)
CA3051065A1	23 August 2018 (23-08-2018)	AU2018221097A1 AU2019101603A4 CN107040582A EP3554051A1 EP3554051A4 KR20190098765A MX2019009291A SG11201906738SA TW201832099A US2019332608A1 WO2018149345A1	01 August 2019 (01-08-2019) 23 January 2020 (23-01-2020) 11 August 2017 (11-08-2017) 16 October 2019 (16-10-2019) 01 January 2020 (01-01-2020) 22 August 2019 (22-08-2019) 30 October 2019 (30-10-2019) 27 September 2019 (27-09-2019) 01 September 2018 (01-09-2018) 31 October 2019 (31-10-2019) 23 August 2018 (23-08-2018)