



- (51) International Patent Classification:
G06F 15/16 (2006.01)
- (21) International Application Number:
PCT/US2016/027322
- (22) International Filing Date:
13 April 2016 (13.04.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
201510213538.5 30 April 2015 (30.04.2015) CN
- (71) Applicant: **ALIBABA GROUP HOLDING LIMITED**
[—/US]; Fourth Floor, One Capital Place, P.O. Box 847,
George Town, Grand Cayman (KY).
- (72) Inventors: **DONG, Yuanyuan**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN). **XU, Li**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West

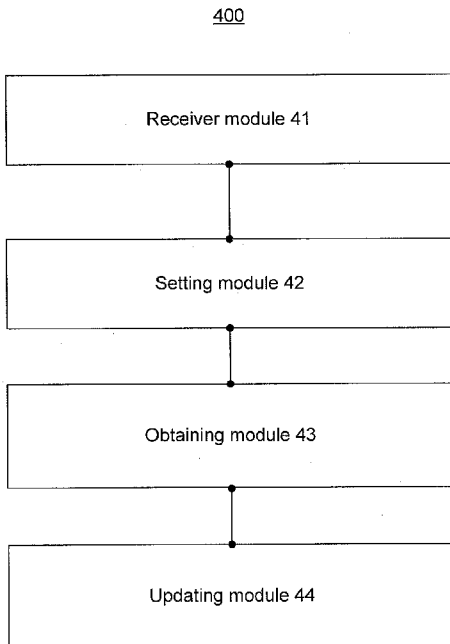
Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN). **YAN, Weibin**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN). **LI, Kai**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN). **YANG, Puyuan**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN). **SUI, Yemin**; Alibaba Group Legal Department, 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou 311121 (CN).

(74) Agent: **CAPRON, Aaron, J.**; Finnegan, Henderson, Farabow, Garrett & Dunner LLP, 901 New York Avenue, NW, Washington, DC 20001-4413 (US).

(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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

[Continued on next page]

(54) Title: SYSTEM, METHOD, AND APPARATUS FOR UPDATING DATA IN A DISTRIBUTED STORAGE SYSTEM



(57) Abstract: A method for updating data in a distributed storage system is disclosed. The method may comprise receiving, by a sender terminal device, an update request; obtaining, by the sender terminal device, update data and control information from the update request; sending, by the sender terminal device, the update data, via a data update request, to a receiver terminal device; and sending, by the sender terminal device, the control information, via a control information update request, to the receiver terminal device.

Fig. 4

WO 2016/176045 A1



HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,

Published:

— with international search report (Art. 21(3))

**SYSTEM, METHOD, AND APPARATUS FOR UPDATING DATA IN A
DISTRIBUTED STORAGE SYSTEM**

CROSS REFERENCE TO RELATED APPLICATION

[001] The present application is based on and claims the benefits of priority to Chinese Application No. 201510213538.5, filed April 30, 2015, the entire contents of which are incorporated herein by reference.

TECHNICAL FIELD

[002] The present application relates to methods, apparatus, and systems for updating data in a distributed storage system.

BACKGROUND

[003] In distributed storage systems, each data block may be stored as multiple replicates and when the data block is updated, the same update needs to be implemented among the replicates as well. The update among replicates can be implemented by a coordinator. In a master-slave replication system, the coordinator can be a master server and update process can be: the coordinator sequences updates data blocks, assigns new data versions, generates update requests, and sends the requests to slave servers. The update request can comprise two parts: data update and control information update. The control information update includes the new data version and metadata of the data block, such as a data block identification, the original version of the data block, the new version of the data block, and metadata information for other slave servers. Each slave server receives and processes the update request and then replies to the coordinator. The coordinator determines, according to a number of normal responses from the slave servers, the

number of update requests sent, and a data consistency protocol, whether the data update is successful. If yes, it will modify the data block to the new version.

[004] In the process of updating the data, the new version is generated based on the current version of the data block. Thus, the version monotonically increases and only one latest version is included in the update requests. If, instead, the slave server receives update requests including multiple new versions, the complexity will increase for maintaining consistency in updating the data. Therefore, in the update process, if the coordinator receives multiple update requests to the same data block, the request has to be queued and only after completing a current update request, it will process the next.

[005] In order to achieve consistency in updating replicates of the data block, update requests of the same data block are sequenced, i.e., in serial execution of the data update requests. The data update process includes: network transmission, logic processing of the request, and write disk IO parts. At present, the disk can be solid state discs (SSD) and cluster write performance is relatively high. But carrying, in an update request, multiple data updates as batch processing can have the data update requests processed at once, reducing delay in the logic processing. The main bottleneck of data updates is in transmission network: the batch processing contains multiple updates and the amount of data transmission is large, causing a significant transmission delay.

SUMMARY

[006] One aspect of the present disclosure is directed to a method for updating data in a distributed storage system. The method may comprise receiving, by a sender terminal device, an update request; obtaining, by the sender terminal

device, update data and control information from the update request; sending, by the sender terminal device, the update data, via a data update request, to a receiver terminal device; and sending, by the sender terminal device, the control information, via a control information update request, to the receiver terminal device.

[007] Another aspect of the present disclosure is directed to a method for updating data in a distributed storage system. The method may comprise receiving, by the receiver terminal device, a data update request from a sender terminal device; setting, by the receiver terminal device, corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table; receiving, by the receiver terminal device, a control information update request corresponding to the data update request from the sender terminal device; obtaining, by the receiver terminal device, the new data corresponding to the data block identification and the new data table identification carried by the control information update request from the data update relation table; and updating, by the receiver terminal device, the data based on the control information update request and the obtained new data.

[008] Another aspect of the present disclosure is directed to an apparatus, implemented as a sender terminal device, for updating data in a distributed storage system. The apparatus may comprise a receiving module configured to receive an update request and to obtain update data and control information from the update request, and a sending module configured to send the update data, via a data update request, to a receiver terminal device and to send the control information, via a control information update request, to the receiver terminal device.

[009] Another aspect of the present disclosure is directed to an apparatus, implemented as a receiver terminal device, for updating data in a distributed storage

system. The apparatus may comprise a receiver module configured to receive a data update request from a sender terminal device and to receive a control information update request corresponding to the data update request from the sender terminal device, a setting module configured to set corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table, an obtaining module configured to obtain the new data corresponding to the data block identification and the new data table identification carried by the control information update request from the data update relation table, and an updating module configured to update the data based on the control information update request and the obtained new data.

[010] Another aspect of the present disclosure is directed to a non-transitory computer-readable medium storing one or more programs, the one or more programs comprising instructions which, when executed by a processor of a computer system, cause the computer system to perform a method. The method may comprise receiving, by the computer system, an update request; obtaining, by the computer system, update data and control information from the update request; sending, by the computer system, the update data, via a data update request, to a receiver terminal device; and sending, by the computer system, the control information, via a control information update request, to the receiver terminal device.

[011] Another aspect of the present disclosure is directed to a non-transitory computer-readable medium storing one or more programs, the one or more programs comprising instructions which, when executed by a processor, cause the computer system to perform a method. The method may comprise receiving, by the receiver terminal device, a data update request from a sender terminal device; setting, by the receiver terminal device, corresponding relations among a data block

identification, a new data table identification, and new data in a data update relation table; receiving, by the receiver terminal device, a control information update request corresponding to the data update request from the sender terminal device; obtaining, by the receiver terminal device, the new data corresponding to the data block identification and the new data table identification carried by the control information update request from the data update relation table; and updating, by the receiver terminal device, data based on the control information update request and the obtained new data.

[012] Additional features and advantages of the present disclosure will be set forth in part in the following detailed description, and in part will be obvious from the description, or may be learned by practice of the present disclosure. The features and advantages of the present disclosure will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims.

[013] It is to be understood that the foregoing general description and the following detailed description are exemplary and explanatory only, and are not restrictive of the invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[014] The accompanying drawings, which constitute a part of this specification, illustrate several embodiments and, together with the description, serve to explain the disclosed principles.

[015] FIG. 1 is a flow diagram illustrating a method for updating data in a distributed storage system, according to an exemplary embodiment.

[016] FIG. 2 is a flow diagram illustrating another method for updating data in a distributed storage system, according to an exemplary embodiment.

[017] FIG. 3 is a block diagram illustrating an apparatus for updating data in a distributed storage system, according to an exemplary embodiment.

[018] FIG. 4 is a block diagram illustrating another apparatus for updating data in a distributed storage system, according to an exemplary embodiment.

DETAILED DESCRIPTION

[019] Reference will now be made in detail to exemplary embodiments, examples of which are illustrated in the accompanying drawings. The following description refers to the accompanying drawings in which the same numbers in different drawings represent the same or similar elements unless otherwise represented. The implementations set forth in the following description of exemplary embodiments consistent with the present invention do not represent all implementations consistent with the invention. Instead, they are merely examples of systems and methods consistent with aspects related to the invention as recited in the appended claims.

[020] FIG. 1 is a flow diagram illustrating a method 100 for updating data in a distributed storage system, according to an exemplary embodiment.

[021] At step 101, a sender terminal device receives an update request and obtains update data and control information from the update request. The update request may include a data block identification. In one embodiment, when the sender terminal device determines that a number of update requests corresponding to the data block exceeds a predetermined threshold, the sender terminal device obtains the update data and the control information.

[022] For example, the update data and the control information may be both from the same update request, i.e., corresponding to the data block identification.

The update request may comprise the data block identification of the data block to be updated. All update requests received by the sender terminal device may be stored in a queue. When the sender terminal device determines that the number of update requests corresponding to the data block exceeds the predetermined threshold, the sender terminal device may obtain the update data and the control information.

[023] At step 102, the sender terminal device sends the update data, via a data update request, to a receiver terminal device. The update data, to be sent to the receiver terminal device, may include the data block identification, new data table identification, and new data. The new data table identification may be configured by the sender terminal device. The new data table identification may be configured to uniquely identify or correspond to the data update request.

[024] The sender terminal device may save the new data table identification, carried by the data update request, to a table of sent items and may, in the future, retrieve the new data table identification from the table of sent items.

[025] After sending the data update request, the sender terminal device may save the new data table identification to the table of sent items, so that it may retrieve the new data table identification from the table of sent items for sending the data update request in the future.

[026] In some embodiments, the update data may comprise multiple pieces of sub-data. The sender terminal device may send the multiple pieces of sub-data to the receiver terminal device respectively via multiple data update requests. Each of the multiple data update requests may carry a new data table identification. The sender terminal device may first receive the multiple pieces of sub-data, each piece of the sub-data comprising the same data block identification and corresponding new

data. The new data table identification and the new data may be uniquely associated with the data update request. After sending the data update requests, the sender terminal device may save the new data table identifications to the table of sent items. For example, the data update request, comprising the data block identification, new data table identification, and the new data sent by the sender terminal device, may be illustrated in the following table. Data update request 1 may carry data block identification 1, new data table identification 1.0, and new data 1. Data update request 2 may carry data block identification 1, new data table identification 1.1, and new data 2. Data update request 3 may carry data block identification 1, new data table identification 1.2, and new data 3. Data update request 4 may carry data block identification 1, new data table identification 1.3, and new data 4. Thus, the same data block identification, but different new data table identifications and different new data are carried by the data update requests. The new data table identifications may be saved in the table of sent items.

Data update request	Data block identification	New data table identification	New data
Data update request 1	Data block identification 1	1.0	New data 1
Data update request 2	Data block identification 1	1.1	New data 2
Data update request 3	Data block identification 1	1.2	New data 3
Data update request 4	Data block identification 1	1.3	New data 4

[027] At step 103, the sender terminal device sends the control information, via a control information update request, to the receiver terminal device. The control information, to be sent to the receiver terminal device, may include a data block identification, a new data table identification, and new version information. The new data table identification can be obtained from the table of sent items. This method applies similarly to multiple new data table identifications. For example, if the new data table identifications save in the table of sent items are 1.0, 1.1, 1.2, and 1.3, the new data table identifications carried by the sender terminal device via the control information update request are 1.0, 1.1, 1.2, and 1.3, as shown in the following table.

Control information update request	Data block identification	New data table identification
Control information update request	1	1.0, 1.1, 1.2, 1.3

[028] In some embodiments, if the sender terminal device receives a message corresponding to the data block identification, indicating that the data block has finished updating, and if some of the data update requests have not been sent, the sender terminal device may continue sending the remaining data update requests to the receiver terminal device. Also, the sender terminal device may combine the remaining data update requests and remaining control information update requests into correspondingly restored update requests and may send the restored update requests to the receiver terminal device.

[029] Consistent with exemplary embodiments of this disclosure, the sender terminal device can receive the update request and obtain update data and control information from the update request. The sender terminal device may send the update data, via a data update request, to the receiver terminal device. The sender terminal device may send the control information, via a control information update request, to the receiver terminal device. The update data and the control information may be sent separately. The update data of the update request can be sent to the receiver terminal device while data is updated in the receiver terminal device. This method can significantly lower wait time in transferring data during updates, therefore lowering update delay time and enhancing system performances.

[030] FIG. 2 is a flow diagram illustrating another method 200 for updating data in a distributed storage system, according to an exemplary embodiment.

[031] At step 201, the receiver terminal device receives the data update request from the sender terminal device.

[032] At step 202, the receiver terminal device configures/sets corresponding relations among the data block identification, the new data table identification, and the new data in a data update relation table.

[033] For example, the received data update request comprises the data block identification, the new data table identification, and the new data. The new data table identification may be configured by the sender terminal device to uniquely identify the data update request, while the sender terminal device is sending the data update request. Exemplary relations among the data block identification, the new data table identification, and the new data are illustrated in the following data update relation table. The receiver terminal device may configure/set, in the data update relation table: corresponding relations among data block identification 1, new data

table identification 1.0, and new data 1 carried by data update request 1; corresponding relations among data block identification 1, new data table identification 1.1, and new data 2 carried by data update request 2; corresponding relations among data block identification 1, new data table identification 1.2, and new data 3 carried by data update request 3; and corresponding relations among data block identification 1, new data table identification 1.3, and new data 4 carried by data update request 4.

Data block identification	New data table identification	New data
1	1.0	New data 1
1	1.1	New data 2
1	1.2	New data 3
1	1.3	New data 4

[034] In some embodiments, the receiver terminal device may delete the corresponding relation(s), in the data update relation table, or cleanup invalid update data. Deleting the corresponding relation(s) may be achieved through at least one of the following methods.

[035] Method 1: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, the receiver terminal device sets a timer for the corresponding relations and deletes the corresponding relations from the data update relation table when the timer expires.

[036] Method 2: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, the receiver terminal device sets a timer for the corresponding relations. When a memory space used by the receiver terminal device for the data update relation table exceeds a predetermined threshold, the receiver terminal device deletes a corresponding relation that consumes most of the memory space or a corresponding relation established earliest in the data update relation table.

[037] Method 3: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, the receiver terminal device sets a timer for the corresponding relations. When the timer expires. the receiver terminal device deletes the corresponding relations from the data update relation table. When a memory space used by the receiver terminal device for the data update relation table exceeds a predetermined threshold, the receiver terminal device deletes a corresponding relation that consumes most of the memory space or a corresponding relation established earliest in the data update relation table.

[038] At step 203, the receiver terminal device receives, from the sender terminal device, the control information update request corresponding to the data update request, and obtain, from the data update relation table, the new data corresponding to the data block identification and the new data table identification carried by the control information update request.

[039] The receiver terminal device may receive the control information carried by the control information update request. The control information update request corresponding to the data update request may correspond to the control information, along with the update data of the data update request, carried by the

data update request. The data block identification/new data table identification carried by the control information update request may be the data block identification/new data table identification carried by the data update request. After sending the data update request, the sender terminal device may save the new data table identification to the table of sent items. When sending the control information update request, the sender terminal device may retrieve the new data table identification from the table of sent items. For example, the control information update request may carry new data table identifications 1.0, 1.1, 1.2, and 1.3, and the data block identification 1. The new data received by the receiver terminal device, based on the new data table identification and the data block, is new data 1, new data 2, new data 2, and new data 4.

[040] In some embodiments, a control information update request may correspond to multiple data update requests. A control information update request may comprise multiple new data table identifications. Thus, multiple pieces of new data may be obtained based on data block identification(s) and new data table identification(s) of a control information update request.

[041] At step 204, the receiver terminal device updates data based on the control information update request and the obtained new data.

[042] The receiver terminal device may restore the control information update request and the obtained new data to a restored update request and may update the data based on the restored update request. In some embodiments, the receiver terminal device responds to the sender terminal device after finishing the data update, e.g., by sending a message/signal of completion.

[043] Consistent with the exemplary embodiments of this disclosure, the receiver terminal device may receive the data update request from the sender

terminal device. The receiver terminal device may configure/set, in the data update relation table, corresponding relations among the data block identification, the new data table identification, and the new data, carried by the data update request. The receiver terminal device may receive, from the sender terminal device, the control information update request corresponding to the data update request. The receiver terminal device may obtain, from the data update relation table, the new data corresponding to the data block identification and the new data table identification carried by the control information update request. The receiver terminal device may update the data based on the obtained new data and the control information update request. By sending the update data of the update request while the receiver terminal device updates the data, this method can significantly lower wait time in transferring data during updates, therefore lowering update delay time and enhancing system performances.

[044] FIG. 3 is a block diagram illustrating an apparatus 300 for updating data in a distributed storage system, according to an exemplary embodiment. Apparatus 300 may comprise a number of components, some of which are optional. Apparatus 300 may be the sender terminal device described above. Apparatus 300 may include a receiving module 31, sending module 32, and a saving module 33.

[045] Receiving module 31 may be configured to receive an update request and to obtain update data and control information from the update request. The update request may include a data block identification. Receiving module 31 may be configured to obtain the update data and the control information, when the sender terminal device determines that a number of update requests corresponding to the data block exceeds a predetermined threshold.

[046] Sending module 32 may be configured to send the update data, via a data update request, to a receiver terminal device. Sending module 32 may also be configured to send the control information, via a control information update request, to the receiver terminal device.

[047] Saving module 33 may be configured to save the new data table identification to the table of sent items, and to retrieve the new data table identification from the table of sent items.

[048] Consistent with exemplary embodiments of this disclosure, the sender terminal device can receive the update request and obtain update data and control information from the update request. The sender terminal device may send the update data, via a data update request, to the receiver terminal device. The sender terminal device may send the control information, via a control information update request, to the receiver terminal device. The update data and the control information may be sent together or separately. By sending the update data of the update request to the receiver terminal device while data is updated in the receiver terminal device, this method can significantly lower wait time in transferring data during updates, therefore lowering update delay time and enhancing system performances.

[049] FIG. 4 is a block diagram illustrating another apparatus 400 for updating data in a distributed storage system, according to an exemplary embodiment. Apparatus 400 may comprise a number of components, some of which are optional. Apparatus 400 may be the receiver terminal device described above. Apparatus 400 may include a receiver module 41, setting module 42, obtaining module 43, and updating module 44.

[050] Receiver module 41 may be configured to receive the data update request from the sender terminal device.

[051] Setting module 42 may be configured to configure/set corresponding relations among the data block identification, the new data table identification, and the new data in a data update relation table.

[052] In some embodiments, setting module 42 may also be configured to delete the corresponding relation(s), in the data update relation table, or cleanup invalid update data. Deleting the corresponding relation(s) may be achieved through at least one of the following methods.

[053] Method 1: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, setting module 42 sets a timer for the corresponding relations and deletes the corresponding relations from the data update relation table when the timer expires.

[054] Method 2: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, setting module 42 sets a timer for the corresponding relations. When a memory space used by the receiver terminal device for the data update relation table exceeds a predetermined threshold, setting module 42 deletes a corresponding relation that consumes most of the memory space or a corresponding relation established earliest in the data update relation table.

[055] Method 3: when configuring/setting the corresponding relations among the data block identification, the new data table identification, and the new data, setting module 42 sets a timer for the corresponding relations. When the timer expires, setting module 42 deletes the corresponding relations from the data update relation table. When a memory space used by the receiver terminal device for the data update relation table exceeds a predetermined threshold, setting module 42

deletes a corresponding relation that consumes most of the memory space or a corresponding relation established earliest in the data update relation table.

[056] Receiving module 41 may also be configured to receive, from the sender terminal device, the control information update request corresponding to the data update request.

[057] Obtaining module 43 may be configured to obtain, from the data update relation table, the new data corresponding to the data block identification and the new data table identification carried by the control information update request.

[058] Updating module 44 may be configured to update data based on the control information update request and the obtained new data.

[059] Updating module 44 may also be configured to restore the control information update request and obtained new data to a restored update request and to update the data based on the restored update request.

[060] Consistent with the exemplary embodiments of this disclosure, the receiver terminal device may receive the data update request from the sender terminal device. The receiver terminal device may configure/set, in the data update relation table, corresponding relations among the data block identification, the new data table identification, and the new data, carried by the data update request. The receiver terminal device may receive, from the sender terminal device, the control information update request corresponding to the data update request. The receiver terminal device may obtain, from the data update relation table, the new data corresponding to the data block identification and the new data table identification carried by the control information update request. The receiver terminal device may update the data based on the obtained new data and the control information update request. By sending the update data of the update request while the receiver

terminal device updates the data, this method can significantly lower wait time in transferring data during updates, therefore lowering update delay time and enhancing system performances.

[061] A person skilled in the art can further understand that various exemplary logic blocks, modules, circuits, and algorithm steps described with reference to the disclosure herein may be implemented as electronic hardware, computer software, or a combination of electronic hardware and computer software. For example, the modules/apparatus and corresponding steps may be implemented by a processor executing software instructions stored in the computer readable medium. The apparatus/terminal device may be a mobile phone, a personal computer, a server, or an internet device that implements the methods/steps described above.

[062] The flowcharts and block diagrams in the accompanying drawings show system architectures, functions, and operations of possible implementations of the system and method according to multiple embodiments of the present invention. In this regard, each block in the flowchart or block diagram may represent one module, one program segment, or a part of code, where the module, the program segment, or a part of code includes one or more executable instructions used for implementing specified logic functions. It should also be noted that, in some alternative implementations, functions marked in the blocks may also occur in a sequence different from the sequence marked in the drawing. For example, two consecutive blocks actually can be executed in parallel substantially, and sometimes, they can also be executed in reverse order, which depends on the functions involved. Each block in the block diagram and/or flowchart, and a combination of blocks in the block diagram and/or flowchart, may be implemented by

a dedicated hardware-based system for executing corresponding functions or operations, or may be implemented by a combination of dedicated hardware and computer instructions.

[063] As will be understood by those skilled in the art, embodiments of the present disclosure may be embodied as a method, a system or a computer program product. Accordingly, embodiments of the present disclosure may take the form of an entirely hardware embodiment, an entirely software embodiment, or an embodiment combining software and hardware. Furthermore, embodiments of the present disclosure may take the form of a computer program product embodied in one or more computer available storage media (including but not limited to a magnetic disk memory, a CD-ROM, an optical memory and so on) containing computer available program codes.

[064] Embodiments of the present disclosure are described with reference to flow diagrams and/or block diagrams of methods, devices (systems) and computer program products according to embodiments of the present disclosure. It will be understood that each flow and/or block of the flow diagrams and/or block diagrams, and combinations of flows and/or blocks in the flow diagrams and/or block diagrams, can be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general-purpose computer, a special-purpose computer, an embedded processor, or other programmable data processing devices to produce a machine, such that the instructions, which are executed via the processor of the computer or other programmable data processing devices, create a means for implementing the functions specified in one or more flows in the flow diagrams and/or one or more blocks in the block diagrams.

[065] These computer program instructions may also be stored in a computer readable memory that can direct a computer or other programmable data processing devices to function in a particular manner, such that the instructions stored in the computer readable memory produce a manufactured product including an instruction means which implements the functions specified in one or more flows in the flow diagrams and/or one or more blocks in the block diagrams.

[066] These computer program instructions may also be loaded onto a computer or other programmable data processing devices to cause a series of operational steps to be performed on the computer or other programmable devices to produce processing implemented by the computer, such that the instructions, which are executed on the computer or other programmable devices, provide steps for implementing the functions specified in one or more flows in the flow diagrams and/or one or more blocks in the block diagrams. In a typical configuration, a computer device includes one or more Central Processing Units (CPUs), an input/output interface, a network interface and a memory. The memory may include forms of a volatile memory, a random access memory (RAM) and/or non-volatile memory and the like, such as a read-only memory (ROM) or a flash RAM in a computer readable medium. The memory is an example of the computer readable medium.

[067] The computer readable storage medium refers to any type of physical memory on which information or data readable by a processor may be stored. Thus, a computer readable storage medium may store instructions for execution by one or more processors, including instructions for causing the processor(s) to perform steps or stages consistent with the embodiments described herein. The computer readable medium includes non-volatile and volatile media, removable and non-

removable media, wherein information storage can be implemented with any method or technology. Information may be modules of computer readable instructions, data structures and programs, or other data. Examples of a computer storage medium include, but are not limited to, a phase-change random access memory (PRAM), a static random access memory (SRAM), a dynamic random access memory (DRAM), other types of random access memories (RAMs), a read-only memory (ROM), an electrically erasable programmable read-only memory (EEPROM), a flash memory or other memory technologies, a compact disc read-only memory (CD-ROM), a digital versatile disc (DVD) or other optical storage, a cassette tape, tape or disk storage, other magnetic storage devices, or any other non-transmission media which may be used to store information capable of being accessed by a computer device. The computer readable medium is non-transitory, and does not include transitory media, such as modulated data signals and carrier waves.

[068] The specification has described methods, apparatus, and systems for updating data in a distributed storage system. The illustrated steps are set out to explain the exemplary embodiments shown, and it should be anticipated that ongoing technological development will change the manner in which particular functions are performed. Thus, these examples are presented herein for purposes of illustration and not limitation. For example, steps or processes disclosed herein are not limited to being performed in the order described but may be performed in any order, and some steps may be omitted consistent with disclosed embodiments. Further, the boundaries of the functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternative boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed. Alternatives (including equivalents, extensions, variations, deviations,

etc., of those described herein) will be apparent to persons skilled in the relevant art(s) based on the teachings contained herein. Such alternatives fall within the scope and spirit of the disclosed embodiments.

[069] While examples and features of disclosed principles are described herein, modifications, adaptations, and other implementations are possible without departing from the spirit and scope of the disclosed embodiments. Also, the words “comprising,” “having,” “containing,” “including,” and other similar forms, are intended to be equivalent in meaning and be open ended in that an item or items following any one of these words is not meant to be an exhaustive listing of such item or items, or meant to be limited to only the listed item or items. It must also be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural references unless the context clearly dictates otherwise.

[070] It will be appreciated that the present invention is not limited to the exact construction that has been described above and illustrated in the accompanying drawings, and that various modifications and changes can be made without departing from the scope thereof. It is intended that the scope of the invention should only be limited by the appended claims.

WHAT IS CLAIMED IS:

1. A method for updating data in a distributed storage system, comprising:
 - receiving, by a sender terminal device, an update request;
 - obtaining, by the sender terminal device, update data and control information from the update request;
 - sending, by the sender terminal device, the update data, via a data update request, to a receiver terminal device; and
 - sending, by the sender terminal device, the control information, via a control information update request, to the receiver terminal device.

2. The method of claim 1, wherein the receiver terminal device is configured to:
 - set corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table;
 - obtain the new data corresponding to the data block identification and the new data table identification carried by the control information update request; and
 - update the data based on the control information update request and the obtained new data.

3. The method of claim 2, before sending the control information, via the control information update request, to the receiver terminal device, further comprising:
 - saving, by the sender terminal device, the new data table identification, carried by the data update request, to a table of sent items; and
 - retrieving, by the sender terminal device, the new data table identification from the table of sent items.

4. The method of claim 2, wherein sending the update data, via the data update request, to the receiver terminal device comprises:

sending, by the sender terminal device, multiple pieces of sub-data to the receiver terminal device respectively via multiple data update requests, the update data including the multiple pieces of sub-data and the multiple data update requests carrying corresponding new data table identifications.

5. The method of claim 1, wherein obtaining the update data and the control information from the update request comprises obtaining, by the sender terminal device, the update data and the control information, when the sender terminal device determines that a number of update requests corresponding to the data block exceeds a predetermined threshold.

6. A method for updating data in a distributed storage system, comprising:

receiving, by the receiver terminal device, a data update request from a sender terminal device;

setting, by the receiver terminal device, corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table;

receiving, by the receiver terminal device, a control information update request corresponding to the data update request from the sender terminal device;

obtaining, by the receiver terminal device, the new data corresponding to the data block identification and the new data table identification carried by the control information update request from the data update relation table; and

updating, by the receiver terminal device, the data based on the control information update request and the obtained new data.

7. The method of claim 6, wherein updating the data based on the control information update request and the obtained new data comprises:

restoring, by the receiver terminal device, the control information update request and the obtained new data to a restored update request; and

updating, by the receiver terminal device, the data based on the restored update request.

8. The method of claim 6, after setting the corresponding relations among the data block identification, the new data table identification, and the new data in the data update relation table, further comprising at least one of:

setting, by the receiver terminal device, a timer for the corresponding relations, and deleting, by the receiver terminal device, the corresponding relations from the data update relation table when the timer expires; or

setting, by the receiver terminal device, a timer for the corresponding relations and deleting, by the receiver terminal device, a corresponding relation that consumes most of the memory space or a corresponding relation established earliest in the data update relation table, when a memory space used by the receiver terminal device for the data update relation table exceeds a predetermined threshold.

9. An apparatus, implemented as a sender terminal device, for updating data in a distributed storage system, comprising:

a receiving module configured to receive an update request and to obtain update data and control information from the update request; and

a sending module configured to send the update data, via a data update request, to a receiver terminal device and to send the control information, via a control information update request, to the receiver terminal device.

10. The apparatus of claim 9, wherein the receiver terminal device is configured to:

set corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table;

obtain the new data corresponding to the data block identification and the new data table identification carried by the control information update request; and

update the data based on the control information update request and the obtained new data.

11. The apparatus of claim 10, further comprising:

a saving module configured to save the new data table identification, carried by

the data update request, to a table of sent items and to retrieve the new data table identification from the table of sent items.

12. The apparatus of claim 9, wherein:

the update data includes the multiple pieces of sub-data;

the sending module is configured to send multiple pieces of sub-data to the receiver terminal device respectively via multiple data update requests; and

the multiple data update requests carry corresponding new data table identifications.

13. The apparatus of claim 9, wherein:

the update request includes a data block identification; and

the obtaining module is further configured to obtain the update data and the control information, when the sender terminal device determines that a number of update requests corresponding to the data block exceeds a predetermined threshold.

14. A non-transitory computer-readable medium storing one or more programs, the one or more programs comprising instructions which, when executed by a processor of a computer system, cause the computer system to perform a method, comprising:

receiving, by the computer system, an update request;

obtaining, by the computer system, update data and control information from the update request;

sending, by the computer system, the update data, via a data update request, to a receiver terminal device; and

sending, by the computer system, the control information, via a control information update request, to the receiver terminal device.

15. The non-transitory computer-readable medium of claim 14, wherein the receiver terminal device is configured to:

set corresponding relations among a data block identification, a new data table identification, and new data in a data update relation table;

obtain the new data corresponding to the data block identification and the new data table identification carried by the control information update request; and

update the data based on the control information update request and the obtained new data.

16. The non-transitory computer-readable medium of claim 15, before sending the control information, via the control information update request, to the receiver terminal device, further comprising:

saving, by the computer system, the new data table identification, carried by the data update request, to a table of sent items; and

retrieving, by the computer system, the new data table identification from the table of sent items.

17. The non-transitory computer-readable medium of claim 15, wherein sending the update data, via the data update request, to the receiver terminal device comprises:

sending, by the computer system, multiple pieces of sub-data to the receiver terminal device respectively via multiple data update requests, the update data including the multiple pieces of sub-data and the multiple data update requests carrying corresponding new data table identifications.

18. The non-transitory computer-readable medium of claim 14, wherein obtaining the update data and the control information from the update request comprises obtaining, by the computer system, the update data and the control information, when the computer system determines that a number of update requests corresponding to the data block exceeds a predetermined threshold.

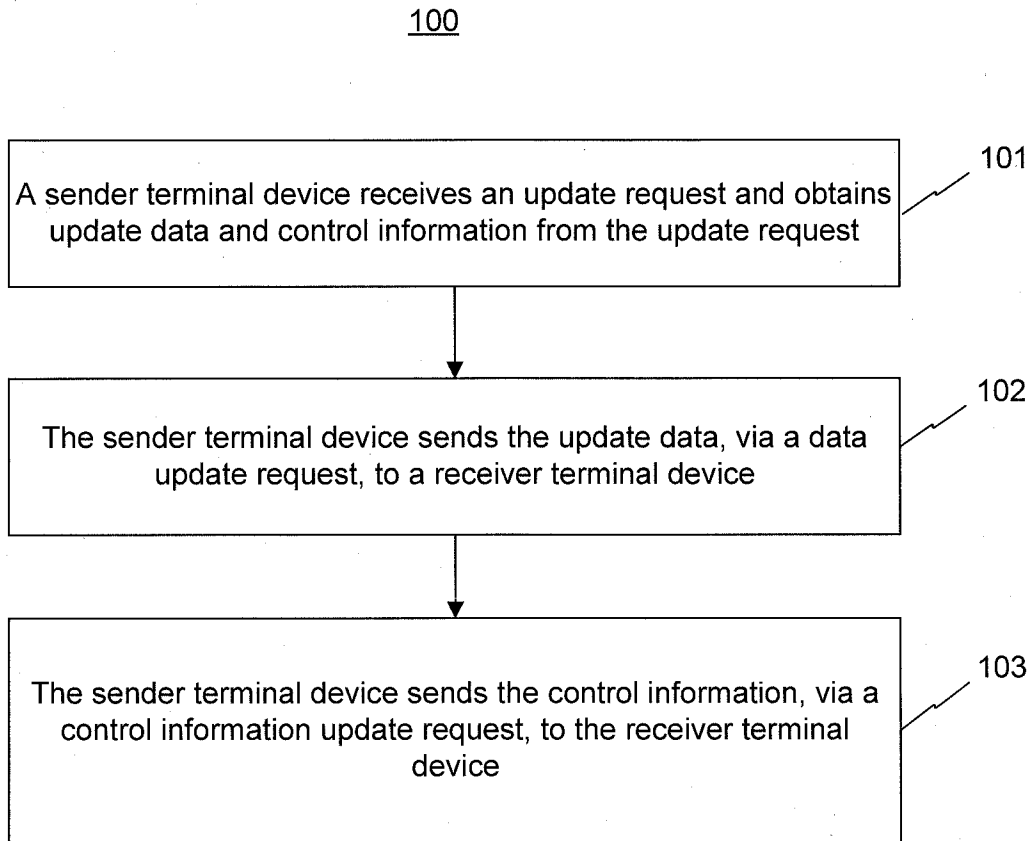
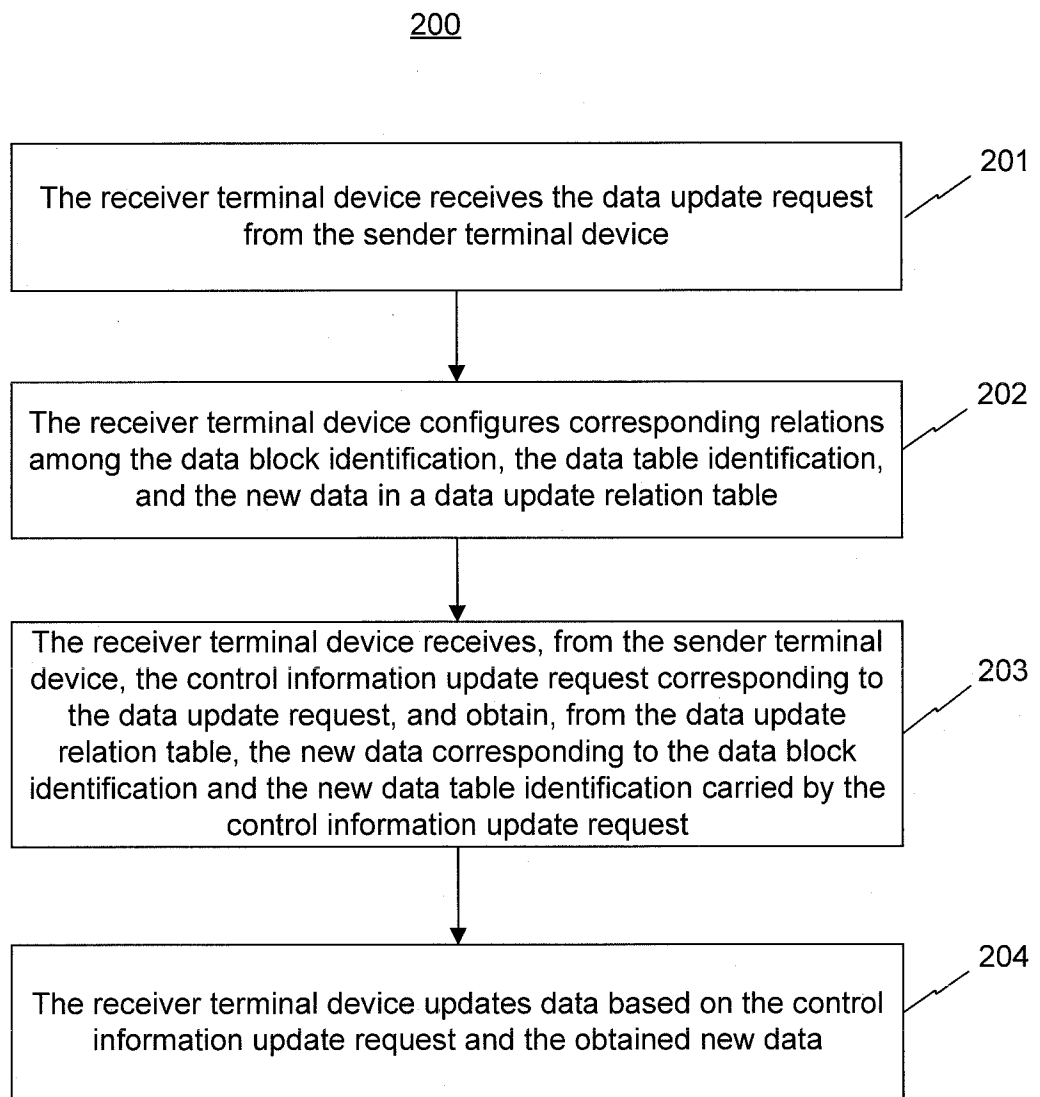


Fig. 1

**Fig. 2**

300

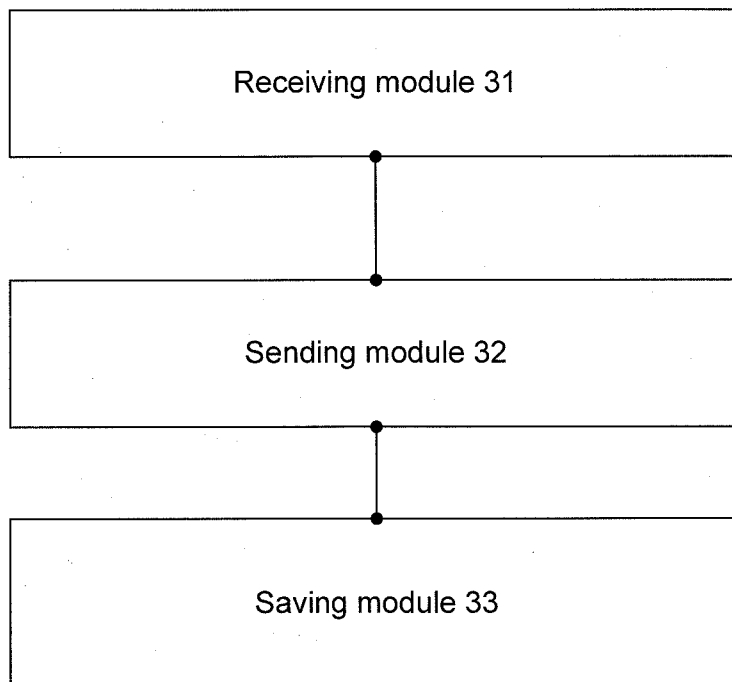


Fig. 3

400

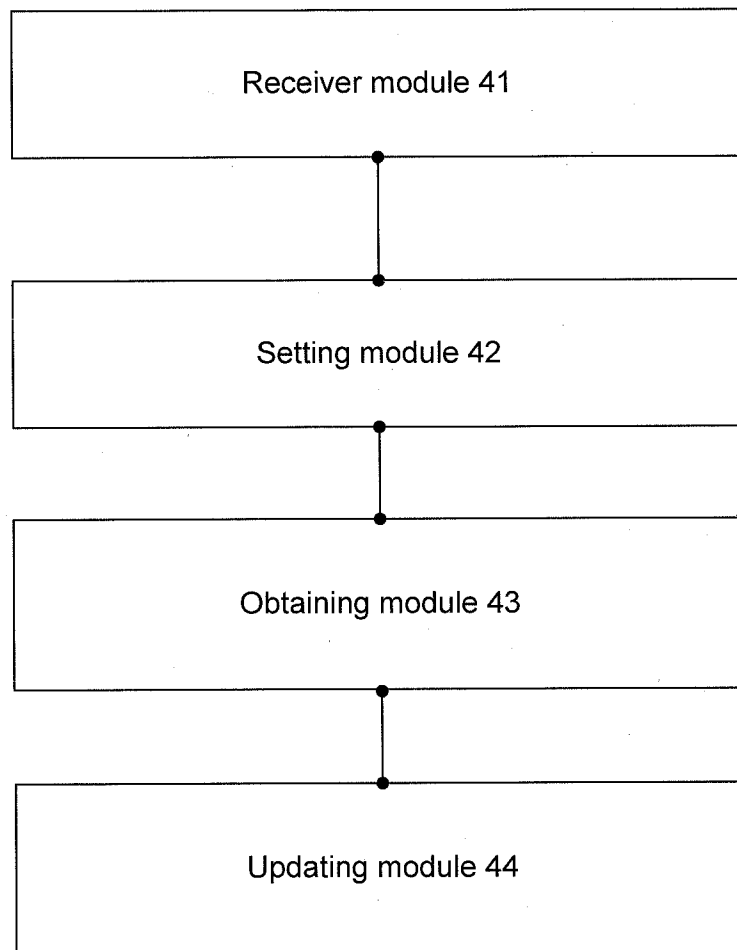


Fig. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 16/27322

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06F 15/16 (2016.01)

CPC - H04L 29/08072; H04L 29/06; H04L 29/08135

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(8): G06F 15/16 (2016.01)

CPC: H04L 29/08072; H04L 29/06; H04L 29/08135

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

IPC(8): G06F 15/16 (2016.01); CPC: H04L 29/08072; H04L 29/06; H04L 29/08135; H04L 29/08144; H04L 29/0809; USPC: 709/201; 709/217; 709/213 (keyword limited; terms below)

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

PatBase; Google(Web); Search terms used: updating data distributed storage sender terminal update request control information receiver expiration timer delete restore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011/0289058 A1 (Anzai et al.) 24 November 2011 (24.11.2011), entire document especially abstract; Fig. 1, 20, 24; para [0044], [0128]-[0139], [0171], [0172]	1-18
Y	US 2006/0149740 A1 (Tsutazawa et al.) 06 July 2006 (06.07.2006), entire document especially para [0002], [0024], [0070], [0116], [0253], [0299]	1-18
Y	US 2014/0207913 A1 (Maddah-Ali et al.) 24 July 2014 (24.07.2014), abstract; para [0082], [0085], [0095]	7
Y	US 2003/0233442 A1 (Futatsugi) 18 December 2003 (18.12.2003), abstract; para [0120]	8

 Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" 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)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" 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

"X" 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

"Y" 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

08 June 2016

Date of mailing of the international search report

18 JUL 2016

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774