



- (51) International Patent Classification:
G06F 17/30 (2006.01) *G06F 17/40* (2006.01)
- (21) International Application Number:
PCT/US2013/073591
- (22) International Filing Date:
6 December 2013 (06.12.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, Texas 77070 (US).
- (72) Inventors: **JOTHIVELAVAN, Sivashanmugam**; Sy. No. 192, Whitefield Road Mahadevapura Post, Bangalore 560048 (IN). **RAJKUMAR, Kannan**; Sy. No. 192, Whitefield Road Mahadevapura Post, Bangalore 560048 (IN). **RAMESH, Kannan K.**; Sy. No. 192, Whitefield Road Mahadevapura Post, Bangalore 560048 (IN). **ANNMARY, Roy**; Sy. No. 192, Whitefield Road Mahadevapura Post, Bangalore 560048 (IN). **JAIPAL, Pendyala**; Sy. No. 192, Whitefield Road Mahadevapura Post, Bangalore 560048 (IN). **COUTINHO, Hamilton De Freitas**; Avenida Ipiranga 6681 Building 99A, 90619-900 Porto Alegre

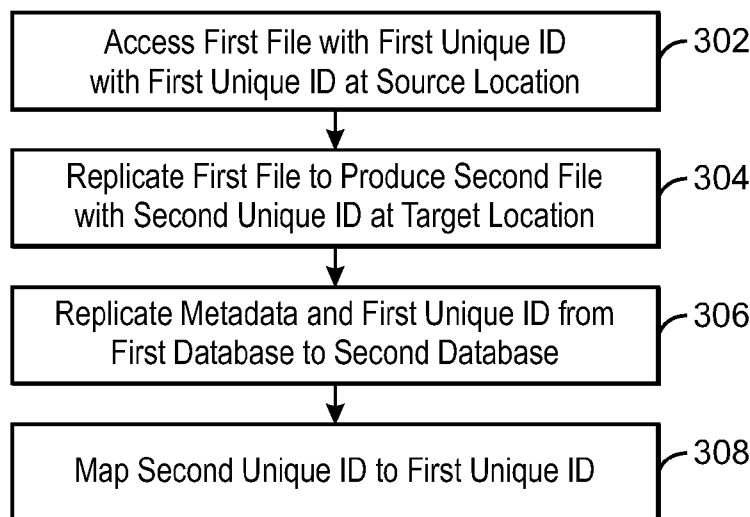
(BR). **MAGALHAES, Guilherme De Campos**; Avenida Ipiranga 6681 Building 91A, 90619-900 Porto Alegre (BR). **CONDOTTA, Marcelo Bandeira**; 165 Dascomb Road, Andover, Massachusetts 01810-5897 (US). **KEETON, Kimberly**; 1501 Page Mill Road, Palo Alto, California 94304-1100 (US). **MORREY III, Charles B.**; 1501 Page Mill Road, Palo Alto, California 94304-1100 (US). **SPITZER, Michael J.**; Columbia Tech Center, 1115 SE 164th Avenue, Vancouver, Washington 98683 (US).

(74) Agents: **ORTEGA, Arthur S.** et al.; Hewlett-Packard Company, Intellectual Property Administration, Mail Stop 35, 3404 E. Harmony Road, Fort Collins, CO 80528 (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, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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.

[Continued on next page]

(54) Title: REPLICATING METADATA ASSOCIATED WITH A FILE



(57) Abstract: The present disclosure is generally related to replicating metadata. A method includes accessing a first file with a first unique identifier at a source location in a storage device, wherein metadata corresponding to the first file is stored in a first database with the first unique identifier. The method includes replicating the first file to produce a second file at a target location, wherein the second file has a second unique identifier. The method includes replicating the metadata and the first unique identifier to a second database. The method includes mapping the second unique identifier to the first unique identifier in the second database.



(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, 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 the identity of the inventor (Rule 4.17(i))*
- *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))*

REPLICATING METADATA ASSOCIATED WITH A FILE

BACKGROUND

[0001] Metadata for a file stored in a file system contains information describing the data contained in the file. The metadata may contain the file's unique identifier, among other attributes associated with the file. If the file is replicated to a different file system, the metadata may be replicated as well.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] Certain examples are described in the following detailed description and in reference to the drawings, in which:

[0003] Fig. 1 is a block diagram of a computing system configured for replicating metadata, in accordance with examples of the present disclosure;

[0004] Fig. 2 is a block diagram illustrating metadata replication from a source file system to a target file system;

[0005] Fig. 3 is a process flow diagram of a method for replicating metadata, in accordance with examples of the present disclosure; and

[0006] Fig. 4 is a block diagram of a tangible, non-transitory computer-readable medium containing instructions configured to direct a processor to replicate metadata, in accordance with examples of the present disclosure.

DETAILED DESCRIPTION

[0007] The present disclosure is generally related to replicating metadata. When a file located in a source file system is replicated to a target file system, the metadata associated with the file can be replicated as well. However, custom metadata that a user associates with the file may not be automatically replicated, as the custom metadata may be external to the file, and may reside in a database. One method to replicate the metadata is to manually run a script to export the metadata from the source file system's express query database, and import the metadata to the target file system's database, where it is associated with the path name of the replicated file. However, this method can be prone to errors. For example, a change in the path name of the replicated file can result in invalid association between the replicated file and the metadata.

[0008] Described herein is a method to automatically associate metadata with a replicated file in a target file system following file replication. An original file in a source file system can have its metadata associated with a unique identifier of the file. When the original file is replicated to a target file system, the metadata associated with the unique identifier of the file can be replicated as well. In the target file system, a unique identifier of the replicated file can be mapped to the unique identifier of the original file, such that the metadata is then associated with the unique identifier of the replicated file. In this way, the metadata replication and association can be performed automatically without user intervention. The metadata association can also be unaffected by changes or errors in the path name of the replicated file. Furthermore, the replicated metadata can be stored in a scalable pipelined database. The pipelined database may use a mechanism of lazy ingestion of file system events. The metadata associated with the replicated file may be stored in a query-able authority table in the pipelined database.

[0009] Fig. 1 is a block diagram of a computing system configured for replicating metadata, in accordance with examples of the present disclosure. The computing system 100 may include, for example, a server computer, a mobile phone, laptop computer, desktop computer, or tablet computer, among others. The computing system 100 may include a processor 102 that is adapted to execute stored instructions. The processor 102 can be a single core processor, a multi-core processor, a computing cluster, or any number of other appropriate configurations.

[0010] The processor 102 may be connected through a system bus 104 (e.g., AMBA®, PCI®, PCI Express®, Hyper Transport®, Serial ATA, among others) to an input/output (I/O) device interface 106 adapted to connect the computing system 100 to one or more I/O devices 108. The I/O devices 108 may include, for example, a keyboard and a pointing device, wherein the pointing device may include a touchpad or a touchscreen, among others. The I/O devices 108 may be built-in components of the computing system 100, or may be devices that are externally connected to the computing system 100.

[0011] The processor 102 may also be linked through the system bus 104 to a display device interface 110 adapted to connect the computing system 100 to display devices 112. The display devices 112 may include a display screen that is a built-in component of the computing system 100. The display devices 112 may also include

computer monitors, televisions, or projectors, among others, that are externally connected to the computing system 100.

[0012] The processor 102 may also be linked through the system bus 104 to a memory device 114. In some examples, the memory device 114 can include random access memory (e.g., SRAM, DRAM, eDRAM, EDO RAM, DDR RAM, RRAM®, PRAM, among others), read only memory (e.g., Mask ROM, EPROM, EEPROM, among others), non-volatile memory (PCM, STT_MRAM, ReRAM, Memristor), or any other suitable memory systems.

[0013] The processor 102 may also be linked through the system bus 104 to a storage device 116. The storage device 116 may contain one or more files 118 in a file system. The file 118 may be a document, application, media, or any other virtual item that can be stored. The storage device may also contain metadata 120, which provides information regarding the file 118. Such information may include time of file creation, ownership of the file, and file access permissions. In some examples, the metadata 120 may be custom metadata containing information that a user has manually associated with the file 118. A replication module 122 in the storage device can include instructions to direct the processor 102 to replicate the file 118 from a source location in the storage device 116 to a target location. The target location may be in a second storage device inside the computing system 100, or in an external device coupled to the computing system 100 via wired or wireless means. For example, an external storage device 124 may be linked to the system bus 104 via a communications port 126. The replication module 122 can also replicate the metadata 120 to the target location. The replication module 122 can map the replicated file to the original file 118, such that the replicated file is associated with the metadata.

[0014] Fig. 2 is a block diagram illustrating metadata replication from a source file system to a target file system. The examples discussed herein can be performed by a computer containing a processor and at least one storage device. A first file 202a stored in a source file system 204 of the storage device can be replicated to produce an identical second file 202b stored in a target file system 206. The target file system 206 may be in a second storage device in the computer itself, an external storage device connected to the computer, or a server coupled to the computer in a network.

[0015] The first file 202a can include a unique identifier and associated with metadata. The metadata can contain at least one key and value pair. The key is the name of a metadata element, while the value pertains to the information contained in the metadata element. In one example, the metadata may be custom metadata describing a color of the first file 202a. The key of the custom metadata may read “color”, while the value of the custom metadata may read “red”. The unique identifier and the metadata can be stored in a first database 208 of the source file system 204. The unique identifier and metadata may be associated with one another and stored in a table of the first database 208. The first file 202a can also include an extended attribute 210, which contains the unique identifier and a timestamp of the metadata. The timestamp can refer to when the metadata was created or last modified.

[0016] The first file 202a can be replicated to produce the identical second file 202b to be stored in the target file system 206. The second file 202b can use a different unique identifier from the first file 202a. The extended attribute 210, which contains the unique identifier of the first file 202a and the timestamp of the metadata, can be replicated to the target file system 206 as well. Furthermore, the table in the first database 208 can also be replicated to a second database 212 in the target file system 206.

[0017] The unique identifier of the second file 202b can be mapped to the unique identifier of the first file 202a in a temporary table in the second database 212. As a result, the unique identifier of the second file 202b becomes associated with the metadata. Thus, the metadata can correspond to both the first file 202a and the second file 202b. The process of associating the second file 202b to the metadata can be done automatically in response to replication of the first file 202a. The second database 212 can be a pipelined database wherein the association between the metadata and the second file 202b can be stored in a query-able table.

[0018] Fig. 3 is a process flow diagram of a method for replicating metadata, in accordance with examples of the present disclosure. The method 300 can be performed by a computing system 100 (as seen in Fig. 1) containing a processor and a storage device.

[0019] At block 302, the processor accesses a first file with a first unique identifier at a source location in a storage device. Metadata corresponding to the first file can be stored in a first database with the first unique identifier, such that the first unique identifier is associated with the metadata. The first file may include an extended

attribute that contains the first unique identifier and a timestamp corresponding to the metadata.

[0020] At block 304, the processor replicates the first file to produce a second file at a target location. The target location may be in a second storage device, either contained in the computing system or coupled externally. The extended attribute of the first file can be replicated to the target location as well. The second file can have a second unique identifier.

[0021] At block 306, the processor replicates the metadata and the first unique identifier to a second database. The second database may be at the target location. The metadata and the first unique identifier may be associated together in a temporary table in the second database.

[0022] At block 308, the processor maps the second unique identifier to the first unique identifier in the second database. As a result, the second unique identifier is associated with the metadata corresponding to the first file.

[0001] Fig. 4 is a block diagram of a tangible, non-transitory computer-readable medium containing instructions configured to direct a processor to replicate metadata, in accordance with examples of the present disclosure. The tangible, non-transitory computer-readable medium 400 can include RAM, a hard disk drive, an array of hard disk drives, an optical drive, an array of optical drives, a non-volatile memory, a universal serial bus (USB) drive, a digital versatile disk (DVD), or a compact disk (CD), among others. The tangible, non-transitory computer-readable media 400 may be accessed by a processor 402 over a computer bus 404. Furthermore, the tangible, non-transitory computer-readable medium 400 may include instructions configured to direct the processor 402 to perform the techniques described herein.

[0023] As shown in Fig. 4, the various components discussed herein can be stored on the non-transitory, computer-readable medium 400. A file access module 406 is configured to access a first file at a source location in a storage device, wherein metadata corresponding to the first file is stored in a first database with the first unique identifier. A file replication module 408 is configured to replicate the first file to produce a second file at a target location, wherein the second file has a second unique identifier. A metadata replication module 410 is configured to replicate the metadata and the first unique identifier to a second database. An

identifier mapping module 412 is configured to map the second unique identifier to the first unique identifier in the second database.

[0024] The block diagram of Fig. 4 is not intended to indicate that the tangible, non-transitory computer-readable medium 400 are to include all of the components shown in Fig. 4. Further, the tangible, non-transitory computer-readable medium 400 may include any number of additional components not shown in Fig. 4, depending on the details of the specific implementation.

[0025] While the present techniques may be susceptible to various modifications and alternative forms, the examples discussed above have been shown only by way of example. It is to be understood that the technique is not intended to be limited to the particular examples disclosed herein. Indeed, the present techniques include all alternatives, modifications, and equivalents falling within the true spirit and scope of the appended claims.

CLAIMS

What is claimed is:

1. A method, comprising:
accessing a first file with a first unique identifier at a source location in a storage device, wherein metadata corresponding to the first file is stored in a first database with the first unique identifier;
replicating the first file to produce a second file at a target location, the second file having a second unique identifier;
replicating the metadata and the first unique identifier to a second database;
and
mapping the second unique identifier to the first unique identifier in the second database.
2. The method of claim 1, wherein the first file comprises an extended attribute that contains the first unique identifier.
3. The method of claim 2, comprising replicating the extended attribute of the first file to the target location.
4. The method of claim 1, wherein the second database is a pipelined database.
5. The method of claim 1, wherein the metadata comprises a key and a value.
6. A system, comprising:
a replication module to provide instructions that replicate a file with metadata from a source location to a target location;
a processor to execute the instructions provided by the replication module, wherein the instructions direct the processor to:
access a first file with a first unique identifier at the source location in a storage device, wherein metadata corresponding to the first file is stored in a first database with the first unique identifier;
replicate the first file to produce a second file at the target location, the second file having a second unique identifier;

replicate the metadata and the first unique identifier to a second database; and
map the second unique identifier to the first unique identifier in the second database.

7. The system of claim 6, the first file comprising an extended attribute that contains the first unique identifier.
8. The system of claim 7, the processor to replicate the extended attribute of the first file to the target location.
9. The system of claim 7, wherein the second database is a pipelined database.
10. The system of claim 6, the metadata comprising a key and a value.
11. A tangible, non-transitory, computer-readable medium, comprising instructions configured to direct a processor to:
 - access a first file with a first unique identifier at a source location in a storage device, wherein metadata corresponding to the first file is stored in a first database with the first unique identifier;
 - replicate the first file to produce a second file at a target location, the second file having a second unique identifier;
 - replicate the metadata and the first unique identifier to a second database;
 - and
 - map the second unique identifier to the first unique identifier in the second database.
12. The tangible, non-transitory, computer-readable medium of claim 11, the first file comprising an extended attribute that contains the first unique identifier.
13. The tangible, non-transitory, computer-readable medium of claim 12, comprising instructions configured to direct a processor to replicate the extended attribute of the first file to the target location.

14. The tangible, non-transitory, computer-readable medium of claim 12, wherein the second database is a pipelined database.

15. The tangible, non-transitory, computer-readable medium of claim 11, the metadata comprising a key and a value.

1/4

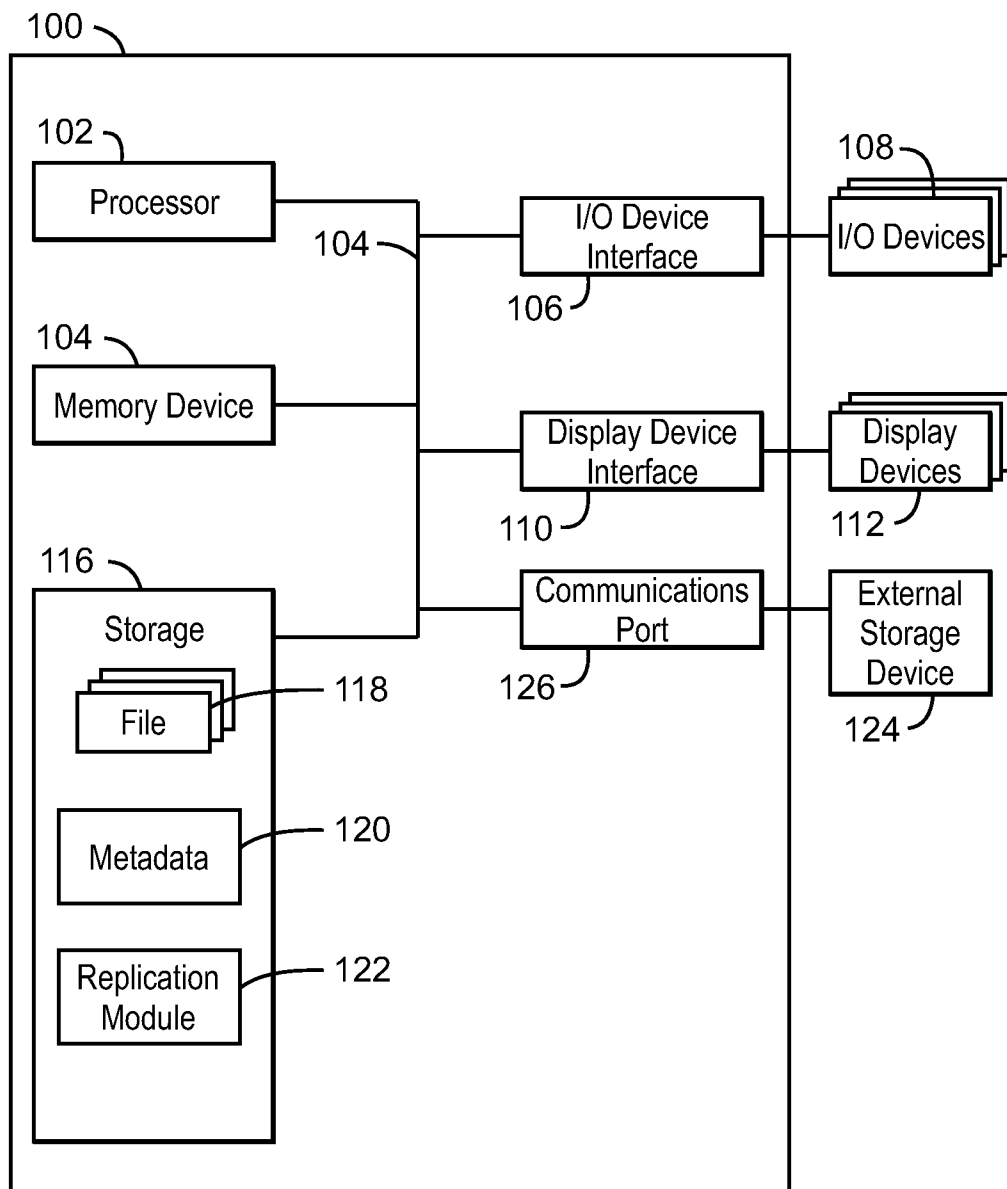
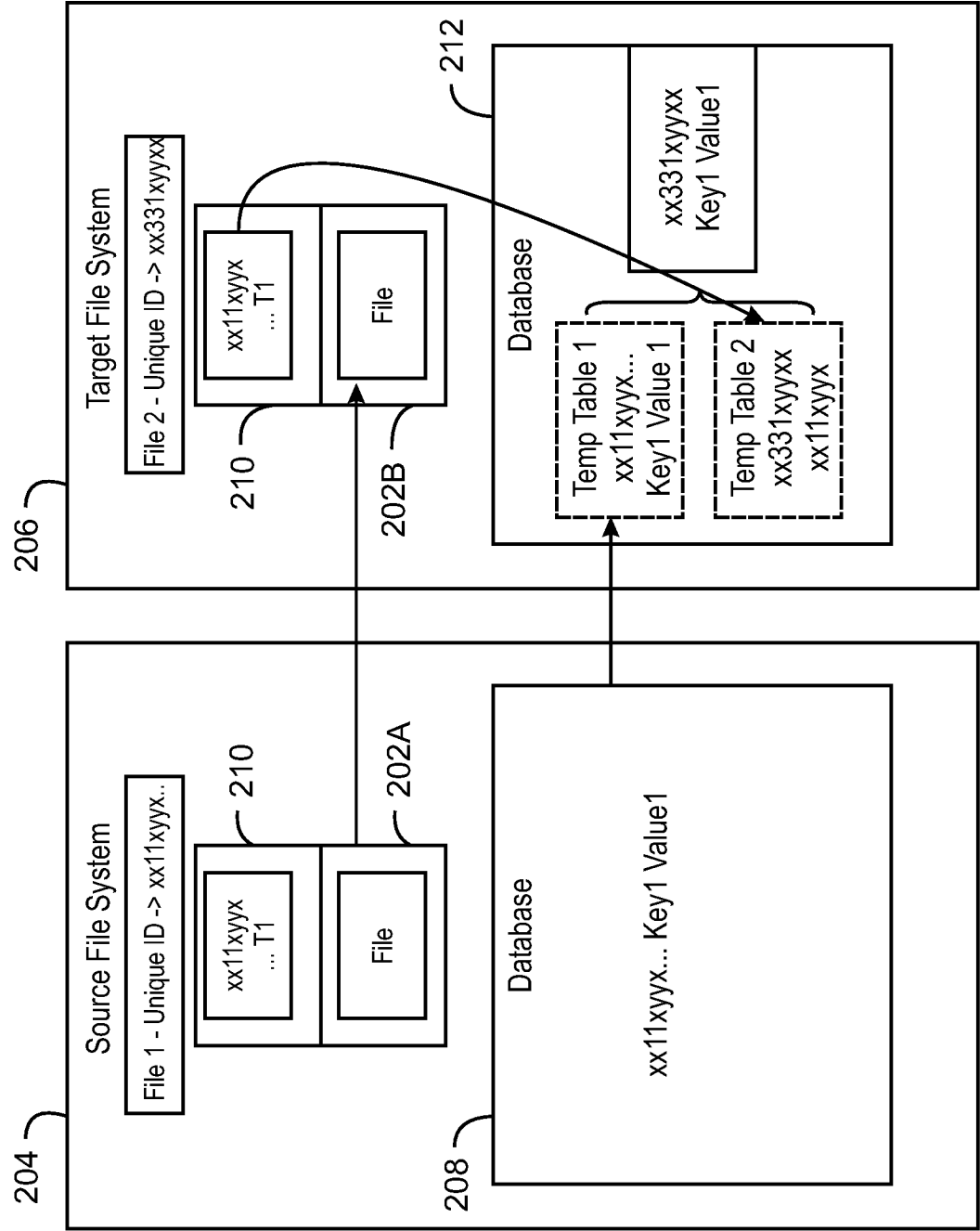


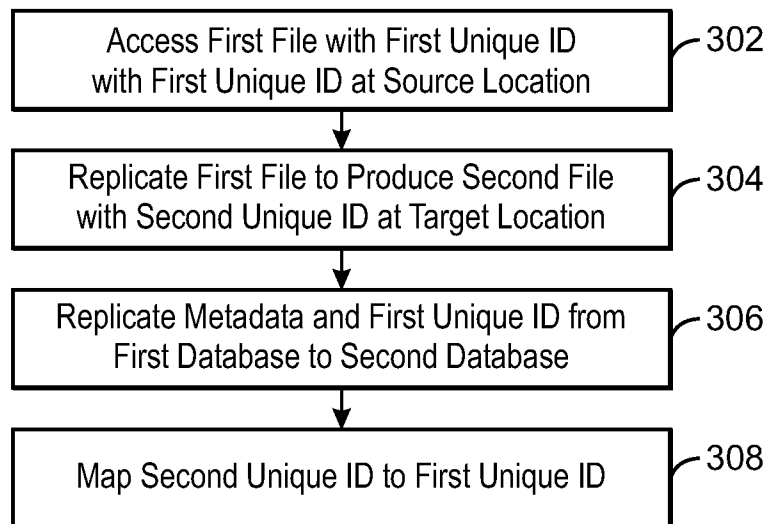
FIG. 1



200

FIG. 2

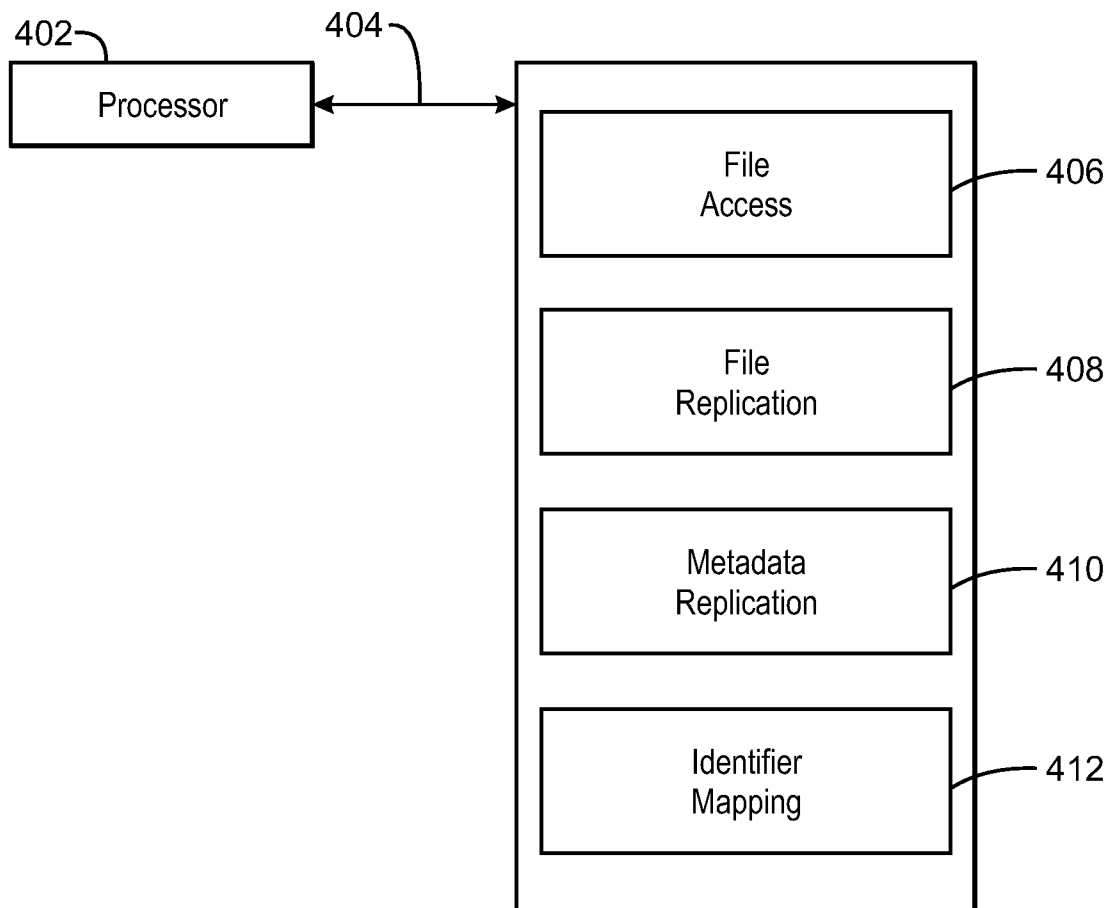
3/4



300
FIG. 3

+

4/4



400
FIG. 4

+

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/073591**A. CLASSIFICATION OF SUBJECT MATTER****G06F 17/30(2006.01)i, G06F 17/40(2006.01)i**

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)

G06F 17/30; G06F 12/00; G06F 17/40

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords:file replication, unique identifier, metadata, and similar terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2007-0022122 A1 (BAHAR CAMERON et al.) 25 January 2007 See paragraphs [0040]-[0042], [0076], [0137]-[0138]; and figures 4, 5B, 16A.	1-15
Y	US 2012-0271793 A1 (PARAG GOKHALE) 25 October 2012 See paragraphs [0028], [0055]; claims 16, 19; and figure 1.	1-15
A	US 2013-0054524 A1 (MATTHEW J. ANGLIN et al.) 28 February 2013 See paragraphs [0037]-[0042]; and figure 6.	1-15
A	US 2011-0078110 A1 (MANOJ PAUL JOSEPH) 31 March 2011 See paragraphs [0028]-[0029]; and figures 1-2.	1-15
A	US 2013-0304777 A1 (BRANDON BILINSKI et al.) 14 November 2013 See paragraphs [0038]-[0042]; and figure 2.	1-15



Further documents are listed in the continuation of Box C.



See patent family annex.

* 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

04 September 2014 (04.09.2014)

Date of mailing of the international search report

04 September 2014 (04.09.2014)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

BYUN, Sung Cheal

Telephone No. +82-42-481-8262



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2013/073591

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007-0022122 A1	25/01/2007	US 2007-0022087 A1 US 2007-0022121 A1 US 2007-022129 A1 US 7571168 B2 US 7574443 B2 US 7599941 B2 WO 2007-014296 A2 WO 2007-014296 A3	25/01/2007 25/01/2007 25/01/2007 04/08/2009 11/08/2009 06/10/2009 01/02/2007 06/12/2007
US 2012-0271793 A1	25/10/2012	US 2009-319585 A1 US 8219524 B2	24/12/2009 10/07/2012
US 2013-0054524 A1	28/02/2013	CN 103765817 A DE 112012002762 T5 GB 201405324 D0 GB 2509023 A US 2013-054523 A1 WO 2013-029173 A1	30/04/2014 10/04/2014 07/05/2014 18/06/2014 28/02/2013 07/03/2013
US 2011-0078110 A1	31/03/2011	None	
US 2013-0304777 A1	14/11/2013	WO 2013-170079 A1	14/11/2013