



(51) International Patent Classification:
G06F 13/00 (2006.01)

(21) International Application Number:
PCT/US2011/042219

(22) International Filing Date:
28 June 2011 (28.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/828,213 30 June 2010 (30.06.2010) US

(71) Applicant (for all designated States except US): **EMC CORPORATION** [US/US]; 176 South Street, Hopkinton, MA 01748 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DUTCH, Michael, John** [US/US]; 21010 Bank Mill Road, Saratoga, CA 95070 (US). **CLAUDATOS, Christopher, Hercules** [US/US]; 842 Portswood Circle, San Jose, CA 95120 (US). **RAO, Mandavilli, Navneeth** [IN/US]; 900 Pepper Tree Lane, Apt.# 1715, Santa Clara, CA 95051 (US).

(74) Agent: **HUANG, Victor**; EMC CORPORATION, 2421 Mission College Boulevard, Santa Clara, CA 95054 (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, 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, IS, JP, KE, KG, KM, KN, KP,

[Continued on next page]

(54) Title: DATA ACCESS DURING DATA RECOVERY

(57) Abstract: A method, article of manufacture, and apparatus for accessing data during data recovery. In some embodiments, this includes sending an I/O request from an application to an object, wherein the object is being recovered, establishing an I/O intercept, intercepting the application's I/O request with the I/O intercept, and redirecting the I/O request based on the status of the object's sub-objects.

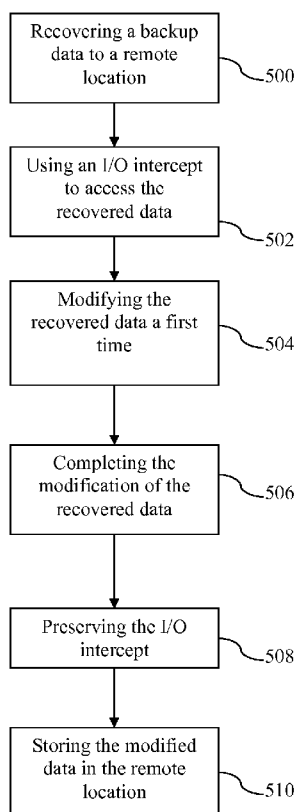


FIG. 5



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, PE, PG, PH, PL, PT, RO, RS, RU, 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.

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, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

[0001] This application is related to co-pending U.S. Patent Application No. 12/828,198 (Attorney Docket No. **EMC-10-169**) for PRIORITIZED BACKUP SEGMENTING and filed concurrently herewith, which is incorporated herein by reference for all purposes; and to co-pending U.S. Patent Application No. 12/828,205 (Attorney Docket No. **EMC-10-186**) for DYNAMIC PRIORITIZED RECOVERY and filed concurrently herewith, which is incorporated herein by reference for all purposes; and to co-pending U.S. Patent Application No. 12/828,218 (Attorney Docket No. **EMC-10-188**) for POST ACCESS DATA PRESERVATION and filed concurrently herewith, which is incorporated herein by reference for all purposes.

FIELD OF THE INVENTION

[0002] The present invention relates generally to data systems, and more particularly, to systems and methods of efficiently protecting and accessing data.

BACKGROUND OF THE INVENTION

[0003] Modern data systems contain vast amounts of data. Such data systems may include, for example, email servers, commercial websites, banking records, government databases, etc.

[0004] Increasingly, there is an expectation that these data systems must always be accessible to the user. For example, if a commercial website is not accessible to the user, (e.g. Amazon.com went down), the business may lose millions of dollars of revenue for every day the website is inaccessible.

[0005] With an ever increasing amount of data, and an ever increasing expectation that the data be accessible, there is tremendous pressure to protect data

case of disasters. However, these conventional data protection schemes are either too slow in recovering the data, or too expensive. Further, conventional data protection schemes require completely recovering the data before allowing access to the data.

[0006] Unfortunately, there may be times when data in a data system needs to be accessible before the recovery operation is complete.

[0007] There is a need, therefore, for an improved method, article of manufacture, and apparatus for protecting and accessing data in data systems.

[0008] The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

[0009] FIG. 1 is a diagram of a data system in accordance with some embodiments.

[0010] FIG. 2 is a flowchart of a method to backup data in accordance with some embodiments.

[0011] FIG. 3 is a flowchart of a method to recover data in accordance with some embodiments.

[0012] FIG. 4 is a flowchart of a method to access data in accordance with some embodiments.

[0013] FIG. 5 is a flowchart of a method to preserve data in accordance with some embodiments.

[0014] A detailed description of one or more embodiments of the invention is provided below along with accompanying figures that illustrate the principles of the invention. While the invention is described in conjunction with such embodiment(s), it should be understood that the invention is not limited to any one embodiment. On the contrary, the scope of the invention is limited only by the claims and the invention encompasses numerous alternatives, modifications, and equivalents. For the purpose of example, numerous specific details are set forth in the following description in order to provide a thorough understanding of the present invention. These details are provided for the purpose of example, and the present invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

[0015] It should be appreciated that the present invention can be implemented in numerous ways, including as a process, an apparatus, a system, a device, a method, or a computer readable medium such as a computer readable storage medium containing computer readable instructions or computer program code, or as a computer program product, comprising a computer usable medium having a computer readable program code embodied therein. In the context of this disclosure, a computer usable medium or computer readable medium may be any medium that can contain or store the program for use by or in connection with the instruction execution system, apparatus or device. For example, the computer readable storage medium or computer usable medium may be, but is not limited to, a random access memory (RAM), read-only memory (ROM), or a persistent store, such as a mass storage device, hard drives,

flash memory), or any magnetic, electromagnetic, infrared, optical, or electrical means system, apparatus or device for storing information. Alternatively or additionally, the computer readable storage medium or computer usable medium may be any combination of these devices or even paper or another suitable medium upon which the program code is printed, as the program code can be electronically captured, via, for instance, optical scanning of the paper or other medium, then compiled, interpreted, or otherwise processed in a suitable manner, if necessary, and then stored in a computer memory. Applications, software programs or computer readable instructions may be referred to as components or modules. Applications may be hardwired or hard coded in hardware or take the form of software executing on a general purpose computer or be hardwired or hard coded in hardware such that when the software is loaded into and/or executed by the computer, the computer becomes an apparatus for practicing the invention. Applications may also be downloaded in whole or in part through the use of a software development kit or toolkit that enables the creation and implementation of the present invention. In this specification, these implementations, or any other form that the invention may take, may be referred to as techniques. In general, the order of the steps of disclosed processes may be altered within the scope of the invention.

[0016] An embodiment of the invention will be described with reference to a data system configured to store files, but it should be understood that the principles of the invention are not limited to data systems. Rather, they are applicable to any system capable of storing and handling various types of objects, in analog, digital, or other form. Although terms such as document, file, object, etc. may be used by way of example, the principles of the invention are not limited to any particular form of

to any object capable of representing information.

[0017] FIG. 1 illustrates a data system in accordance with some embodiments of the present invention. Data System 10 includes Namespace 100, Object 102, Recovery Service 104, Container 106, Intercept 108, Application 110, and Version 112. Though FIG. 1 illustrates Namespace 100, Object 102, Recovery Service 104, Container 106, Intercept 108, Application 110, and Version 112 to be separate components, it should be understood that these components may reside in one or multiple machines. For example, Recovery Service 104 and Container 106 may reside in the same machine, while Intercept 108 and Application 110 reside in a different machine, and Object 102 and Namespace 100 reside in yet another different machine.

[0018] Namespace 100 includes a set of name-address pairs. For example, in some embodiments, Namespace 100 may include a file system directory structure. Namespace 100 is used by Application 110 to locate Object 102.

[0019] Container 106 contains backup data. The backup data may include several files, and several versions of the same file, as depicted by Version 112 in FIG. 1. For example, a first backup may store a preliminary version of a file. A second backup may store a final version of the file. During both backups, the file may have the same file name. Instead of overriding the preliminary version of the file, the second backup may store a second version of the file.

[0020] Version 112 may be segmented into sub-objects. It should be understood that though the term “Version” is used, there need not be multiple versions of a file in a container. For example, a container could have a File X, and only have one version of File X. In this case, it would be File X that is segmented into sub-objects.

In some embodiments, segmenting may be done using a predefined interval (e.g. create segments at every 10 kB). In some embodiments, segmenting may be done according to access pattern. For example, suppose Version 112 is a backup of a virtual machine. If the first 30kB of Version 112 is necessary to boot the virtual machine, it may be preferable to segment the first 30kB. The rest of Version 112 may be segmented into 30kB segments, or may be segmented according to some other criteria (e.g. 40kB segments, segmented according to other known access patterns, etc.). The remaining segments need not be identical in size to each other, or to the initial 30kB segment. Segmenting may also be based upon metadata, such as file history maintained by an application, or upon policy or user command.

[0022] In some embodiments, segments may be assigned priorities. Priorities may be assigned through a variety of methods. In some embodiments, access patterns may be used as criteria in assigning priority. In the example above, it may be preferable to assign the highest priority to the first 30kB of the backup of the virtual machine. Since any user of the virtual machine would first have to access the first 30kB to boot the virtual machine, it would be preferable to recover the first 30kB as quickly as possible. By assigning the highest priority to the first 30kB, a recovery operation would be able to see the highest priority assigned to the first 30kB, and recover it as soon as possible. Other access patterns may be used to determine priority. For example, suppose Version 112 is a backup of a Microsoft Word application, and that the last 5kB of the backup contains the “most recently accessed files” list. A user may decide that they want to know which files were most recently accessed and assign that 5kB segment a high priority. Similarly, a policy could be set such that the most recently accessed files are most likely the first to be accessed upon

priority. Low priorities may be set using similar logic. For example, if certain sub-objects are not likely to be accessed during a recovery operation, those sub-objects may be assigned a low priority.

[0023] Segments (e.g. sub-objects) or versions (e.g. objects) may be “pinned.” Pinning an object or sub-object restricts the pinned object or sub-object to a “Remote I/O” setting during a recovery operation as described herein. Pinning may be preferable when the object being recovered is highly sensitive, and data integrity is of the utmost importance (outweighs performance importance). For example, if an object being recovered is a large database, data integrity may be compromised if recovery is done in parallel sub-objects. In this case, it may be preferable to quickly restore the large database to a remote location (but “local” to container, allowing rapid data transfer between the two), and allow remote access from the client. In this way, data integrity is protected, but the client may notice a decline in performance due to bandwidth and resource issues associated with accessing a remote location.

[0024] In cases where backup data includes a transactional database, it may be preferable to also backup a transaction log along with the data objects. For example, when a transactional database fails, the state of the objects being accessed at the time of the failure may be unknown. In some cases, the last write request may have been done, but there was no confirmation. In other cases, the last write request may not have been done at all. With a transaction log, it can be determined whether a write request had been properly executed, allowing a recovery operation to determine which version of a file to recover.

[0025] FIG. 2 illustrates a method to backup data in accordance with some embodiments. In step 200, an object is backed up to a repository. In step 202, the

segmenting of objects may be based on a variety of criteria. For example, segmenting may be based on access pattern, metadata, policy or command, among others, or any combination of criteria. In step 204, a priority is assigned to a sub-object. As discussed above, priority may also be assigned based on a variety of criteria. For example, priority may be based on access pattern, metadata, policy, or user command, among others, or any combination of criteria. It should be noted that not every sub-object needs a priority assigned. In some embodiments, it may be preferable to “pin” a sub-object to Remote I/O during a recovery operation. In some embodiments, the prioritization itself may be stored as metadata in the backup repository, or as a header in each of the segments.

[0026] Application 110 includes one or more computing processes executing on one or more machines, where each computing process may perform input/output (I/O) from one or more devices. For example, in some embodiments, Application 110 may include a Microsoft Exchange Server client.

[0027] Recovery Service 104 manages the recovery of Object 102. In some embodiments, Recovery Service 104 may track changes to Object 102 and Version 112, such as during Remote I/O.

[0028] FIG. 3 illustrates a method to recover data in accordance with some embodiments. In step 300, an object is selected to be recovered. In step 302, the sub-objects of the object are recovered based on the priorities assigned to the sub-objects. In step 304, the sub-objects are reprioritized based on an application’s I/O during recovery. For example, an application may want to access a sub-object with a low priority. If the sub-object has not been recovered yet, the sub-object’s priority may be

object, and allow the application to access the sub-object.

[0029] Object 102 includes a data object that is being recovered by Recovery Service 104 while Application 110 is using the data object. Object 102 is restored using Version 112, which is part of Container 106. In some embodiments, Object 102 may be segmented into sub-objects. Segmenting Object 102 into sub-objects allows for parallel recovery of the sub-objects.

[0030] As illustrated by Object 102 in FIG. 1, Application 110 may access data objects through multiple methods. One method is Local I/O. This is when the data (e.g. sub-object) has been fully recovered to the local site. During local I/O, Intercept 108 allows Application 110 to locally access a sub-object in Object 102. In some embodiments, during Local I/O, Intercept 108 may act as an intermediary between Application 110 and the sub-object. For example, Application 110 may send a request to what Application 110 thinks is the sub-object. Intercept 108 intercepts the request, retrieves the requested data from the sub-object, and sends the requested data to Application 110. In some embodiments, during Local I/O, Intercept 108 may act as a gatekeeper, and allow Application 110 to communicate directly with the sub-object. For example, Application 110 may send a request to what Application 110 thinks is the sub-object. Intercept 108 intercepts the request, determines whether the request may proceed or not, and if so, allows the request to go through. Application 110 then receives the requested data directly from the sub-object. In some embodiments, after Intercept 108 has determined that Application 110 may directly communicate with the sub-object, subsequent requests by Application 110 to retrieve data from the sub-object need not undergo the determination phase as described herein (e.g. Intercept 108 will no longer need to determine if Application 110 may access the sub-object.)

accessing data that is on the same machine as Application 110. Rather, “Local” may include machines that are geographically distant, but may be connected by a comparatively higher bandwidth connection. “Local” may also include machines that are normally accessed by Application 110. For example, a Microsoft Exchange client resident in California may typically connect to data that is stored in Oregon. A backup copy of the data may be stored in Washington. During recovery, the backup copy is restored to Oregon. When a sub-object is restored to Oregon, the California client may “locally” access the Oregon sub-object via “Local I/O.”

[0031] Another method of Application 110 accessing data is Remote I/O. This occurs when a sub-object has not been restored to the local site, but Application 110 needs access to the sub-object. This may also occur when the object or sub-objects being restored are “pinned” as described herein. In some embodiments, Intercept 108 will intercept Application 110’s request, and redirect to Recovery Service 104. Recovery Service 104 will perform Application 110’s request on a copy of the sub-object maintained by Recovery Service 104. Data objects which are accessed via Remote I/O are spun-up (e.g. booted, mounted, recovered, etc.) by Recovery Service 104 on a system with fewer resource constraints than at the primary system. For example, Recovery Service 104 may mount a virtual machine on a physical machine that is “local” to the container, resulting in a high bandwidth connection between the two, and allowing for a much faster restoration of the virtual machine. A remote connection is established between Application 110 and the spun-up data objects, allowing I/O operations. This may be useful when the sub-object has not been restored to the local site, but when immediate access to the sub-object is required or preferred. For example, it may take a significant amount of time to repair or replace

access data. Rather than wait for the replacement components to arrive, Application 110 may use Remote I/O to access data that is spun up by Recovery Service 104. Further, Remote I/O may be preferable when impact of concurrent recovery on application services using available network bandwidth is unacceptable, or when the risk of data integrity is too high, such as in the case when an object being restored is a large database.

[0032] Another method of Application 110 accessing data is Recover on Access. This includes Application 110 sending an I/O request to the sub-object, and Intercept 108 delaying the I/O request to the sub-object until the sub-object has been recovered. This may be preferable when the performance impact due to the delayed I/O request is tolerable.

[0033] In some embodiments, Intercept 108 is code that is transparent to Application 110. In other words, Application 110 thinks it is accessing Object 102 locally when in fact Intercept 108 is rerouting Application 110's instructions to Recovery Service 104.

[0034] FIG. 4 illustrates a method to access data during data recovery in accordance with some embodiments. In step 400, an I/O request is sent from an application to an object, wherein the object is being recovered. In step 402, an I/O intercept is established. In step 404, the application's I/O request is intercepted by the I/O intercept. In step 406, the I/O request is redirected based on the status of the object's sub-objects. In some embodiments, once an application is able to access a sub-object via Local I/O, the I/O intercept no longer needs to intercept the application's I/O requests because the sub-object has been recovered. In such cases,

the application's I/O requests to proceed uninhibited.

[0035] Recovery may be prioritized. As discussed above, segments of backup data (e.g. objects or versions) may have assigned priority. When a recovery operation detects these priorities, it may recover segments based on priority. For example, suppose the first 30kB of an object is a segment containing the boot up information for a virtual machine, and the first 30kB has been assigned the highest priority. When a recovery operation scans the object, it will identify the first 30kB as having the highest priority, and begin restoring the first 30kB before lower priority segments.

[0036] Priority of restoration may be dynamic during a recovery operation. For example, suppose the last 100kB of a backup image, the backup image being a version, contains the two most recently accessed files, File A and File B. A post-back up operation may have identified these files to be a high priority since they were the two most recently accessed files, and a policy dictated that recently accessed files are to be assigned a high priority. However, during a recovery process, a user may attempt to access the middle 50kB, which contains a rarely accessed File C. Since File C is rarely accessed, a policy may have dictated that it be assigned a low priority, and a recovery operation placed a low priority on restoring File C. When the recovery operation detects that an application is trying to access File C, or when an interceptor tells recovery operation that an application is trying to access File C, the recovery operation may change the priority of File C to high. Depending on the performance tolerance of File C, or other restrictions placed by policy or user, access to File C may be Remote I/O, or Recover on Access. For example, if the user considers File C to be a very important file, where data integrity is of the utmost importance, the user may force the recovery operation to allow access to File C via Remote I/O. In another

tolerate the performance decline (e.g. waiting for File C to be restored before accessing), then the recovery operation may use Recovery on Access as described herein.

[0037] Further, priorities may be altered during an object close. For example, suppose an object, object Z is being restored. Object Z has three sub-objects, sub-objects A, B, and C. A has been restored, and is accessed by an application via Local I/O. B is being restored, but has not yet been restored. C has been “pinned,” and is restricted to remote I/O. During the recovery operation, an application accesses only sub-object A. Before recovery of B or C, the application is finished with object Z, and closes object Z. At this point, it may be preferable to re-assign B and C to a low priority since the application may be done with object Z and need not access it for the remainder of the recovery operation.

[0038] In some embodiments, it may be preferable to retain a copy of an altered sub-object as a result of Remote I/O. Using the above example, suppose the application had accessed sub-object A via Local I/O, and sub-object C via Remote I/O. Before recovery has been completed, the application is finished with object Z. In this case, it may be beneficial for the recovery operation to preserve the changes made to sub-object C (e.g. preserve the remote copy of sub-object C), and to maintain the I/O intercept. Maintaining the I/O intercept would be helpful in case the application again wants to access sub-object C during the recovery operation. Preserving the remote copy of C would be also be helpful in case the application gain wants to access sub-object C during the recovery operation. Once the application has closed object Z, the recovery operation may begin restoring the remote copy of C to

to re-assignment in priority or other settings determined by policy or user.

[0039] FIG. 5 illustrates a method to preserve data changes made during data recovery in accordance with some embodiments. In step 500, a backup is recovered to a remote location. In step 502, an I/O intercept is used to access the recovered data. In step 504, the recovered data is modified a first time. In step 506, the modification of the recovered data is complete (e.g. an application has closed the object). In step 508, the I/O intercept is preserved. In step 510, the modified data is stored in the remote location. In some embodiments, the application may again access the modified data after closing the object (e.g. access the backup data a second time). In this case, it is helpful to maintain the I/O intercept to facilitate Remote I/O. In some embodiments, after the modified data has been recovered to the local location (e.g. the remote location data has been fully restored to the local location), the modified data stored at the remote location may be deleted, may be added to the backup data, or may replace the backup data. The modified data may also be deleted from the remote location once a copy has been restored to the local location.

[0040] For the sake of clarity, the processes and methods herein have been illustrated with a specific flow, but it should be understood that other sequences may be possible and that some may be performed in parallel, without departing from the spirit of the invention. Additionally, steps may be subdivided or combined. As disclosed herein, software written in accordance with the present invention may be stored in some form of computer-readable medium, such as memory or CD-ROM, or transmitted over a network, and executed by a processor.

[0041] All references cited herein are intended to be incorporated by reference. Although the present invention has been described above in terms of specific

no doubt become apparent to those skilled in the art and may be practiced within the scope and equivalents of the appended claims. More than one computer may be used, such as by using multiple computers in a parallel or load-sharing arrangement or distributing tasks across multiple computers such that, as a whole, they perform the functions of the components identified herein; i.e. they take the place of a single computer. Various functions described above may be performed by a single process or groups of processes, on a single computer or distributed over several computers. Processes may invoke other processes to handle certain tasks. A single storage device may be used, or several may be used to take the place of a single storage device. The present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein. It is therefore intended that the disclosure and following claims be interpreted as covering all such alterations and modifications as fall within the true spirit and scope of the invention.

What is claimed is:

Claims

1. A method for accessing data during data recovery, comprising:
sending an I/O request from an application to an object, wherein the object is being recovered;
establishing an I/O intercept;
intercepting the application's I/O request with the I/O intercept; and
redirecting the I/O request based on the status of the object's sub-objects.
2. The method as recited in claim 1, wherein redirecting the I/O request includes redirecting the I/O request to a local copy of the object's sub-objects.
3. The method as recited in claim 1, wherein redirecting the I/O request includes redirecting the I/O request to a remote copy of the object's sub-objects.
4. The method as recited in claim 1, wherein redirecting the I/O request includes holding the I/O request until a sub-object is recovered.
5. The method as recited in claim 3, wherein redirecting the I/O request to a remote copy is based on a pinned status of the object's sub-objects.
6. The method as recited in claim 2, further comprising abolishing the I/O intercept with regard to the local copy after the application has successfully performed I/O on the local copy.

and a processor configured to select send an I/O request from an application to an object, wherein the object is being recovered, establish an I/O intercept, intercept the application's I/O request with the I/O intercept, and redirect the I/O request based on the status of the object's sub-objects.

8. The system as recited in claim 1, wherein redirect the I/O request includes redirecting the I/O request to a local copy of the object's sub-objects.

9. The system as recited in claim 1, wherein redirect the I/O request includes redirecting the I/O request to a remote copy of the object's sub-objects.

10. The system as recited in claim 1, wherein redirecting the I/O request includes holding the I/O request until a sub-object is recovered.

11. The system as recited in claim 9, wherein redirect the I/O request to a remote copy is based on a pinned status of the object's sub-objects.

12. The system as recited in claim 8, the processor further configured to abolish the I/O intercept with regard to the local copy after the application has successfully performed I/O on the local copy.

13. A computer program product for accessing data during data recovery, comprising a computer usable medium having machine readable code embodied therein for:

recovered;

establishing an I/O intercept;

intercepting the application's I/O request with the I/O intercept; and

redirecting the I/O request based on the status of the object's sub-objects.

14. The computer program product as recited in claim 13, wherein redirecting the I/O request includes redirecting the I/O request to a local copy of the object's sub-objects.

15. The computer program product as recited in claim 13, wherein redirecting the I/O request includes redirecting the I/O request to a remote copy of the object's sub-objects.

16. The computer program product as recited in claim 13, wherein redirecting the I/O request includes holding the I/O request until a sub-object is recovered.

17. The computer program product as recited in claim 15, wherein redirecting the I/O request to a remote copy is based on a pinned status of the object's sub-objects.

18. The computer program product as recited in claim 14, further comprising machine readable code for abolishing the I/O intercept with regard to the local copy after the application has successfully performed I/O on the local copy.

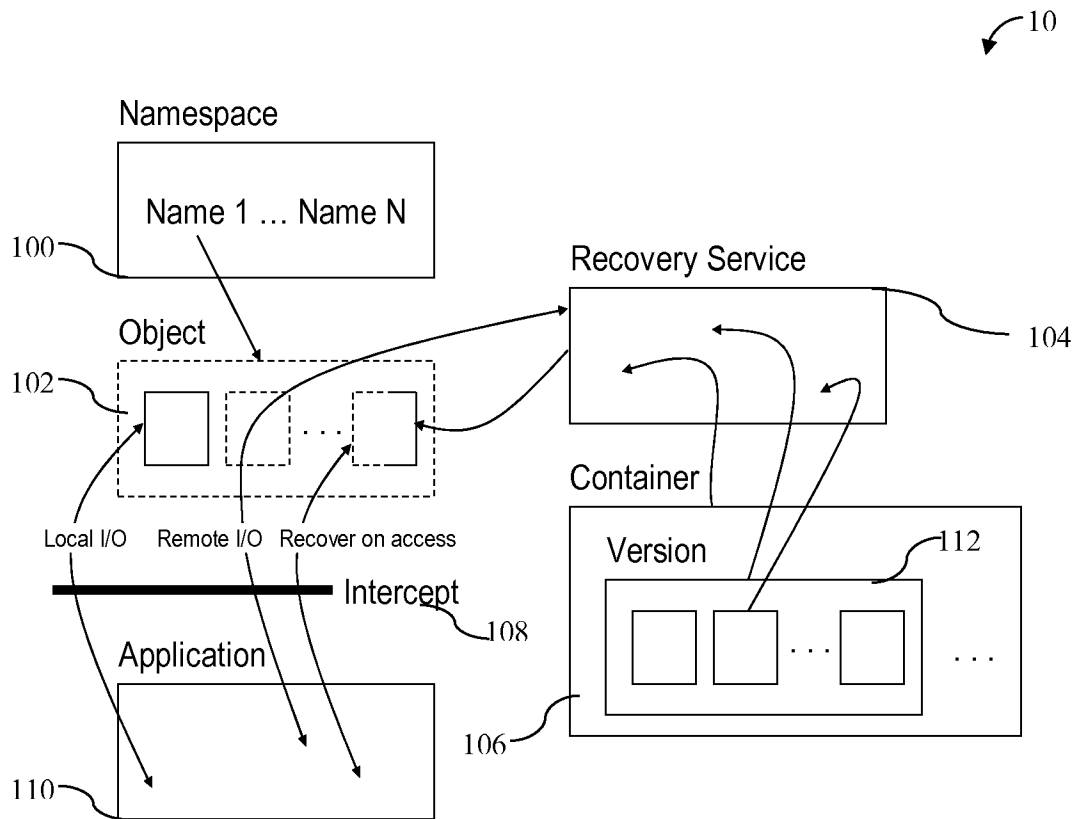
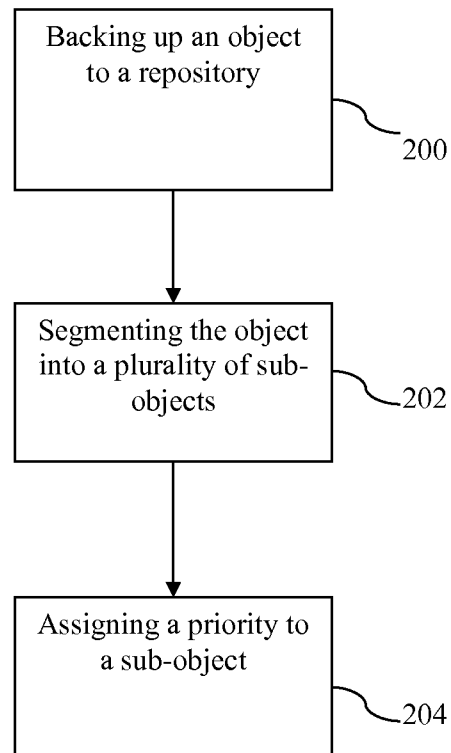
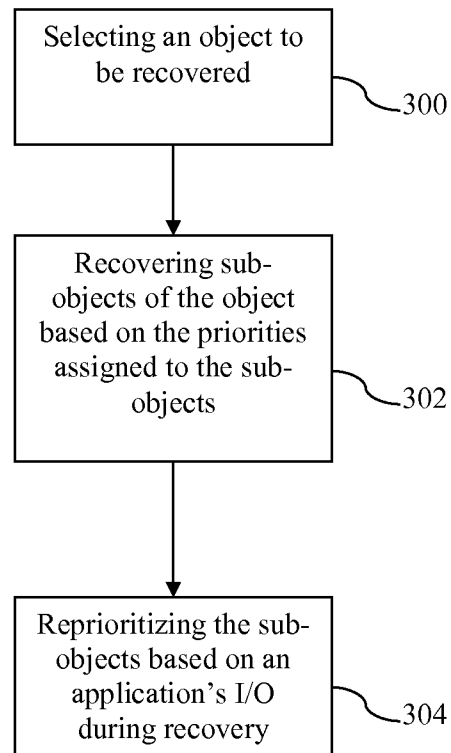
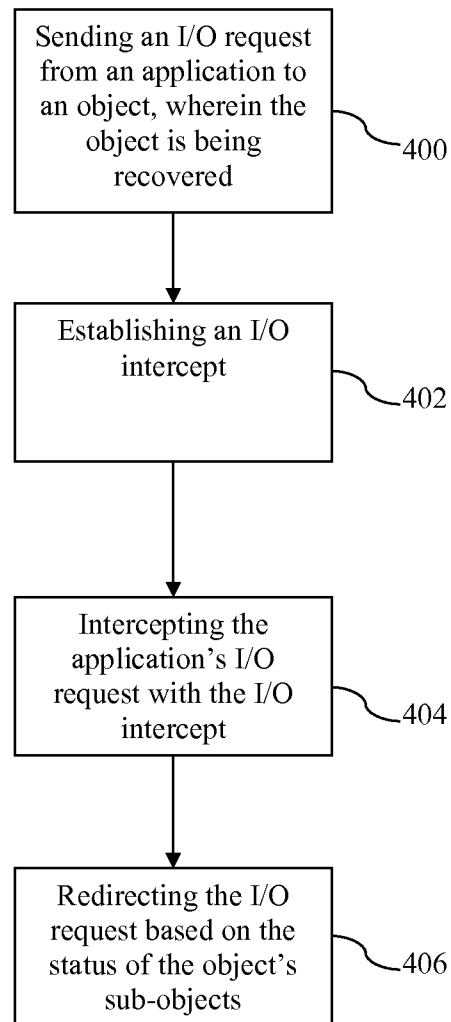


FIG. 1

**FIG. 2**

**FIG. 3**

**FIG. 4**

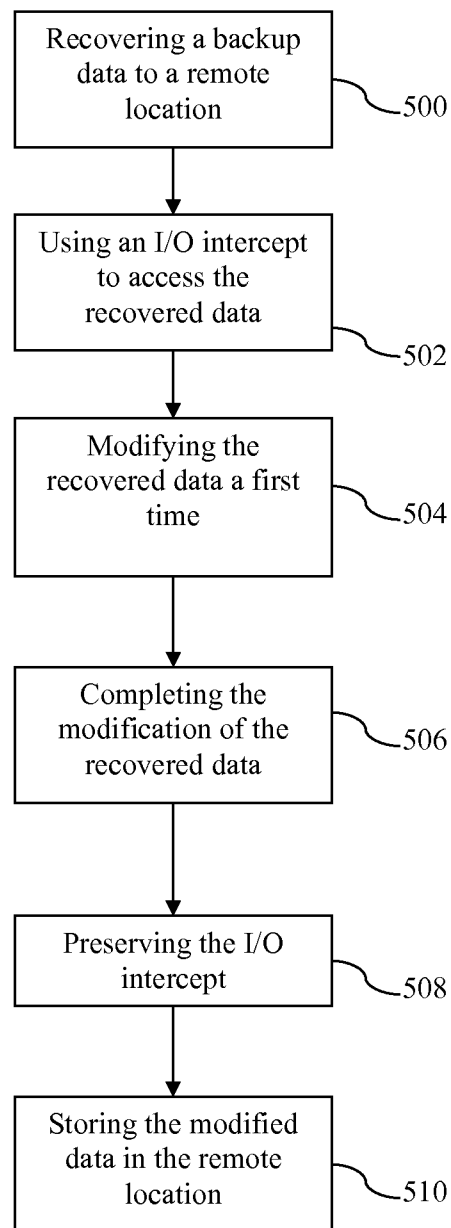


FIG. 5