FILE MANAGEMENT METHOD, DISTRIBUTED STORAGE SYSTEM, AND MANAGEMENT NODE

TECHNICAL FIELD

[0001] The present invention relates to the field of storage technologies, and specifically, to a file management method, a distributed storage system, and a management node.

BACKGROUND

[0002] A next-generation nonvolatile ~~storage medium~~memory may include a nonvolatile memory that has a capability of being addressed and accessed by byte (byte), such as a phase change memory (Phase Change Memory, PCM), a resistive random access memory (Resistive RAM, ReRAM), or a magnetoresistive~~magnetic~~ random access memory (MagnetoresistiveMagnetic Random Access Memory, MRAM). The next-generation nonvolatile memory~~storage medium~~ can provide an interface similar to that of a dynamic random access memory (Dynamic Random Access Memory, DRAM). In a practical application, the next-generation nonvolatile memory may be managed in a page management manner similar to that of the DRAM. Currently, a storage system based on the next-generation nonvolatile memory~~storage medium~~ is a research focus in the industry.

[0003] Generally, a conventional storage system, a distributed file system, or the like records a location of file data in a storage medium by using metadata, to manage the file data. The metadata uses a global contents-based namespace, and generally uses a data structure such as a B-tree (B-tree). For a data block of a file, the data block of the file is traced in a manner of recording a storage node on which the data block is located or even recording a location on a storage node. This manner is applicable to a block device-based scenario in which a conventional hard disk drive (Hard Disk Drive, HDD), solid state drive (Solid State Drive, SSD), or the like is used. After a storage system uses a new nonvolatile ~~storage~~ memory (non-~~violate~~volatile memory, NVM)~~ medium~~, although an original method can be applied to the new NVM~~ medium~~, because the metadata occupies relatively large storage space, and information such as a storage

location corresponding to each data block of the file needs to be recorded in the metadata, management burden is relatively heavy, efficiency of accessing the storage node by a host by obtaining the metadata is also low, and consequently, file management efficiency is relatively low.

SUMMARY

[0004] Embodiments of the present invention provide a file management method, a distributed storage system, and a management node, and can improve file management efficiency.

[0005] According to a first aspect, a file management method is provided, where the method is applied to a distributed storage system, the distributed storage system includes a management node and multiple storage nodes, and the method includes:

receiving, by the management node, a file creation request sent by a host, where the file creation request is used to request to create a file in the distributed storage system;

allocating, by the management node to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, where local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space;

recording, by the management node, metadata of the file, where the metadata of the file includes information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

sending, by the management node, a file creation success response to the host, where the file creation success response carries the information about the first virtual space.

[0006] With reference to the first aspect, in a first possible implementation manner, an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

[0007] With reference to the first aspect or the first possible implementation manner of the first aspect, in a second possible implementation manner, the address bit quantity of the global virtual address space is twice the address bit quantity of the

[0011] With reference to any one of the first aspect, or the first to the fifth possible implementation manners of the first aspect, in a sixth possible implementation manner, the method further includes:

deleting, by the management node, the metadata of the file, and reclaiming the space that is allocated from the global virtual address space to the file;

determining, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

sending a space reclaiming message to the storage node that is used to store the file, where the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.

[0012] According to a second aspect, a distributed storage system is provided, including a management node and multiple storage nodes, where

local virtual address space of each of the multiple storage nodes is corresponding to a part of global virtual address space of the distributed storage system; and

the management node is configured to: receive a file creation request sent by a host, where the file creation request is used to request to create a file in the distributed storage system; allocate, to the file, first virtual space from the global virtual address space according to the file creation request; record metadata of the file, where the metadata of the file includes information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and send a file creation success response to the host, where the file creation success response carries the information about the first virtual space.

[0013] With reference to the second aspect, in a first possible implementation manner, an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

[0014] With reference to the second aspect or the first possible implementation manner of the second aspect, in a second possible implementation manner, the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, where

is used to store the file; and send a space reclaiming message to the storage node that is used to store the file, where the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file; where

the storage node that is used to store the file is configured to reclaim, according to the space reclaiming message, the local virtual address space and the local physical storage space that are corresponding to the file.

[0019] According to a third aspect, a management node is provided, including:

a receiving unit, configured to receive a file creation request sent by a host, where the file creation request is used to request to create a file in a distributed storage system;

a management unit, configured to: allocate, to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, where local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space; and record metadata of the file, where the metadata of the file includes information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

a sending unit, configured to send a file creation success response to the host, where the file creation success response carries the information about the first virtual space.

[0020] With reference to the third aspect, in a first possible implementation manner, an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

[0021] With reference to the third aspect or the first possible implementation manner of the third aspect, in a second possible implementation manner, the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, where N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

[0022] With reference to the third aspect, or the first or the second possible

another embodiment of the present invention;

[0034] FIG. 6 is a schematic flowchart of a file management method according to still another embodiment of the present invention;

[0035] FIG. 7 is a schematic flowchart of a file management method according to yet another embodiment of the present invention;

[0036] FIG. 8 is a schematic block diagram of a distributed storage system according to an embodiment of the present invention;

[0037] FIG. 9 is a schematic block diagram of a management node according to an embodiment of the present invention; and

[0038] FIG. 10 is a schematic block diagram of a management node according to another embodiment of the present invention.

DESCRIPTION OF EMBODIMENTS

[0039] The following clearly and completely describes the technical solutions in the embodiments of the present invention with reference to the accompanying drawings in the embodiments of the present invention. Apparently, the described embodiments are some rather than all of the embodiments of the present invention. All other embodiments obtained by a person of ordinary skill in the art based on the embodiments of the present invention without creative efforts shall fall within the protection scope of the present invention.

[0040] It should be understood that technical solutions of the embodiments of the present invention may be applied to various storage systems, and particularly, to a distributed storage system that is based on a nonvolatile memory~~storage medium~~.

[0041] It should also be understood that, in the embodiments of the present invention, a term "virtual address space" may also be expressed as "virtual space", that is, the "virtual address space" and the "virtual space" are referred to each other.

[0042] FIG. 1 is a schematic diagram of an application scenario according to an embodiment of the present invention. In FIG. 1, a host 130, a storage node 120, and a metadata management node 110 are interconnected by using a network. The storage node 120 and the metadata management node 110 constitute a distributed storage system.

[0043] The host 130 can access a file in the distributed storage system. The host 130 may be a computer, or may be other user equipment, and the host 130 may

communicate with the distributed storage system by using the network. The host 130 may include any computation device known in the prior art, such as a server or a desktop computer. An operating system and another application program are installed in the host 130. There may be multiple hosts 130.

[0044] The storage node 120 is used to store file data. The storage node 120 may include at least one memory. For example, the memory may include a new nonvolatile memory~~storage medium~~ NVM that has a capability of being addressed and accessed by byte (byte), such as a phase change memory (Phase Change Memory, PCM), a resistive random access memory (Resistive RAM, ReRAM), or a magnetoresistive~~magnetic~~ random access memory (Magnetoresistive~~Magnetic~~ Random Access Memory, MRAM); or may include another memory. FIG. 1 uses only the NVM as an example in which the storage node 120 manages physical storage space that includes the NVM and the like on the storage node 120.

[0045] The metadata management node 110 is responsible for global virtual address space management, and provides a function such as file metadata querying.

[0046] The metadata management node 110 and the storage node 120 may be combined together, that is, one node may have functions of both a metadata management node and a storage node.

[0047] In the embodiments of the present invention, a management node may be a metadata management node (such as the metadata management node 110 in FIG. 1), or may be a node that has functions of both a metadata management node and a storage node. For ease of description, the following uses the management node as an example for description.

[0048] FIG. 2 shows a schematic flowchart of a file management method 200 according to an embodiment of the present invention. The method 200 is applied to a distributed storage system, and the distributed storage system includes a management node and multiple storage nodes. The method 200 is executed by the management node.

[0049] S210. The management node receives a file creation request sent by a host, where the file creation request is used to request to create a file in the distributed storage system.

[0050] S220. The management node allocates, to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, where in this embodiment of the present invention, local virtual address space of each storage node in the distributed storage system is corresponding

to a part of the global virtual address space.

[0051] In this embodiment of the present invention, a file in the distributed storage system is managed based on the global virtual address space. The global virtual address space is virtual space of the distributed storage system. The local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space.

[0052] An address bit quantity of the local virtual address space is a bit quantity of an operating system (Operating System, OS). For example, for a 64-bit OS, the address bit quantity of the local virtual address space is 64 bits, and a size of the local virtual address space is 2^64 bytes (^ represents the power). An address bit quantity of the global virtual address space may be 128 bits, and a size of the global virtual address space is 2^128 bytes. In this case, the local virtual address space of each storage node may be corresponding to a part of the global virtual address space.

[0053] In a case, a correspondence (which may also be referred to as a mapping relationship) between the global virtual address space and the local virtual address space may be established in advance. For example, the correspondence between the global virtual address space and the local virtual address space is configured in advance on the management node. After the management node allocates the space to the file, the allocated global virtual address space may be corresponding to the local virtual address space of a storage node according to the correspondence established in advance. In another case, a correspondence (which may also be referred to as a mapping relationship) between the global virtual address space and the local virtual address space may be established when the space is being allocated to the file. For example, when allocating the space to the file, the management node may establish the correspondence between the allocated global virtual address space and the local virtual address space, and save the correspondence between the allocated global virtual address space and the local virtual address space. The allocated global virtual address space may be corresponding to the local virtual address space of the storage node according to the correspondence. In this embodiment of the present invention, how to establish the correspondence between the global virtual address space and the local virtual address space is not limited.

[0054] When receiving the file creation request sent by the host, the management node allocates, to the file, a continuous segment of space (represented as the first virtual space) from the global virtual address space of the distributed storage system.

global virtual address space, without a need to process complex storage location information of a data block of the file, so that file management efficiency can be improved.

[0061] In this embodiment of the present invention, optionally, the local virtual address space of the storage node may be ~~an~~ a segment of unused ~~interval~~ space in virtual address space of the OS. For example, a size of virtual address space of the 64-bit OS is 2 raised to the power of 64. The virtual address space of the 64-bit OS includes reserved virtual address space in addition to user mode process virtual address space and kernel virtual address space. The local virtual address space may be the reserved virtual address space of the OS. Alternatively, the virtual address space of the operating system may be re-divided according to a need, and a segment of virtual address space is obtained by means of division and used as the local virtual address space for data storage by the storage node.

[0062] In this embodiment of the present invention, the global virtual address space is used to manage together all files in the distributed storage system. The global virtual address space is not less than local virtual address space of all storage nodes in the distributed storage system.

[0063] In an embodiment of the present invention, optionally, an address bit quantity of the global virtual address space is greater than the address bit quantity of the local virtual address space of the storage node in the distributed storage system. When the address bit quantity of the local virtual address space is the bit quantity of the OS, the address bit quantity of the global virtual address space is greater than the bit quantity of the OS. In this case, some address bits of the global virtual address space may be used to point to the storage node, and the other address bits may be used to point to the local virtual address space of the storage node. Specifically, if the address bit quantity of the global virtual address space is M, and the address bit quantity of the local virtual address space of the storage node is N, M–N most significant bits of an address of the global virtual address space may be used to point to the storage node. For example, the M–N most significant bits may be corresponding to an identifier and/or routing information of the storage node. N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node. For example, the N least significant bits may be corresponding to an address of the local virtual address space of the storage node.

[0064] In an embodiment of the present invention, optionally, the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system. That is, if the address bit quantity of the local virtual address space of the storage node is N, the address bit quantity of the global virtual address space may be 2N. For example, when the address of the local virtual address space of the storage node has 64 bits, and the address of the global virtual address space has 128 bits; in this case, a size of the local virtual address space of the storage node is 2^64 bytes, and a size of the global virtual address space is 2^128 bytes.

[0065] Optionally, N most significant bits of the address of the global virtual address space are used to point to the storage node, and the N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node. Specifically, the N least significant bits of the address of the global virtual address space are corresponding to the address of the local virtual address space of the storage node and the N most significant bits of the address of the global virtual address space are corresponding to the identifier and/or the routing information of the storage node. For example, when the address of the local virtual address space of the storage node has 64 bits, and the address of the global virtual address space has 128 bits, 64 least significant bits of the address of the global virtual address space are corresponding to the address of the local virtual address space of the storage node, and 64 most significant bits of the address of the global virtual address space are corresponding to the identifier and/or the routing information of the storage node.

[0066] It should be understood that the correspondence between the global virtual address space and the local virtual address space may be another correspondence. For example, multiple storage nodes form one group, the N most significant bits of the address of the global virtual address space are corresponding to a group number, and the N least significant bits are corresponding to a location in the group. When space is being allocated to the file, the group number is determined first, and then a continuous virtual address space segment of a storage node in the group is determined, so that virtual pages of the file are distributed in strips in the group. For example, if eight nodes form one group, the first page of the file is stored on the first node in the group, the eighth page falls on the eighth node, the ninth page falls on the first node, the sixteenth page falls on the eighth node, and so on.

[0067] FIG. 3 is a schematic diagram of an example of global virtual address space. As shown in FIG. 3, local virtual address space of a storage node may be mapped to local physical storage space of the storage node. In an embodiment shown in FIG. 3, an address of the local virtual address space of the storage node has 64 bits, and a size of the local virtual address space of the storage node is 2^64 bytes. An address of the global virtual address space has 128 bits, and a size of the global virtual address space is 2^128 bytes. The local virtual address space of the storage node may be corresponding to a part of the global virtual address space. That is, the global virtual address space may be mapped to local virtual address space of each storage node.

[0068] In this embodiment of the present invention, a file in a distributed storage system may be managed in the foregoing global virtual address space whose size is 2^128 bytes. Each file occupies a continuous segment of the global virtual address space. The global virtual address space segment allocated to the file is mapped to the local virtual address space of each storage node, and is mapped to local physical storage space of the storage node by using the local virtual address space. For example, a continuous segment of space is allocated from the global virtual address space to a file 1, where the continuous segment of space allocated to the file 1 is corresponding to a segment of local virtual address space of a first storage node, and is mapped to local physical storage space of the first storage node.

[0069] In an embodiment of the present invention, optionally, the global virtual address space includes multiple areas, and different areas in the multiple areas have different space allocation granularities. Specifically, the global virtual address space may be divided into multiple areas that have different space allocation granularities. For example, as shown in FIG. 4, the global virtual address space is divided into a large file area, a medium file area, a small file area, and a spare area. A space allocation granularity of the small file area is the smallest, a space allocation granularity of the medium file area is the second smallest, and a space allocation granularity of the large file area is the largest. For example, the space allocation granularity of the small file area may be 1 MB, and if space is allocated from the small file area to a file, 1-MB space is allocated to the file. The space allocation granularity of the medium file area may be 100 MB, and if space is allocated from the medium file area to a file, 100-MB space may be allocated to the file. The space allocation granularity of the large file area may be 1 GB, and if space is allocated

from the large file area to a file, 1-GB space is allocated to the file. During creation of a file, first, space may be allocated from the small file area to the file according to the space allocation granularity of the small file area. Then, when a size of the file exceeds the previously allocated space, space is reallocated, to the file according to the space allocation granularity of the file area, from the file area that is one level higher than the small file area.

[0070] Optionally, as shown in FIG. 5, in an embodiment of the present invention, S220 of the method 200 may be specifically as follows:

[0071] S221. The management node allocates, to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area.

[0072] The first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated. For example, the first area may be the small file area in FIG. 4. During creation of the file, first, space is allocated from the small file area to the file according to the space allocation granularity of the small file area.

[0073] S210, S230, and S240 in FIG. 5 are the same as those in FIG. ~~1~~2, and details are not repeatedly described herein.

[0074] Optionally, as shown in FIG. 6, in an embodiment of the present invention, the method 200 may further include the following steps:

[0075] S251. The management node receives a file space expansion request sent by the host, where the file space expansion request is used to request to expand space of the file.

[0076] S252. Allocate, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, where the space allocation granularity of the second area is greater than the space allocation granularity of the first area.

[0077] S253. Update the metadata of the file, where an updated metadata of the file includes information about the second virtual space allocated to the file.

[0078] S254. Send a file space expansion response to the host, where the file space expansion response carries the information about the second virtual space.

[0079] Specifically, when the space of the file needs to be expanded, the host sends the file space expansion request to the management node. The management node allocates, to the file, space (represented as the second virtual space) from the

file.

[0085] According to the file management method in this embodiment of the present invention, in global virtual address space of a distributed storage system, space is allocated to a file and the space is reclaimed; therefore, management complexity is low, and file management efficiency can be improved.

[0086] It should be understood that, specific examples in this embodiment of the present invention are merely intended to help a person skilled in the art better understand this embodiment of the present invention, but are not intended to limit the scope of this embodiment of the present invention.

[0087] It should also be understood that sequence numbers of the foregoing processes do not mean execution sequences in various embodiments of the present invention. The execution sequences of the processes should be determined according to functions and internal logic of the processes, and should not be construed as any limitation on the implementation processes of the embodiments of the present invention.

[0088] The foregoing describes in detail the file management method according to the embodiment of the present invention. The following describes in detail a distributed storage system and a management node according to an embodiment of the present invention.

[0089] FIG. 8 shows a schematic block diagram of a distributed storage system 800 according to an embodiment of the present invention. As shown in FIG. 8, the distributed storage system 800 includes: a management node 810 and multiple storage nodes 820.

[0090] Local virtual address space of each of the multiple storage nodes 820 is corresponding to a part of global virtual address space of the distributed storage system 800.

[0091] The management node 810 is configured to: receive a file creation request sent by a host, where the file creation request is used to request to create a file in the distributed storage system 800; allocate, to the file, first virtual space from the global virtual address space according to the file creation request; record metadata of the file, where the metadata of the file includes information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node 820 that is used to store the file; and send a file creation success response to the host, where the file creation success response carries

granularity of the second area, where the space allocation granularity of the second area is greater than the space allocation granularity of the first area; update the metadata of the file, where an updated metadata of the file includes information about the second virtual space allocated to the file; and send a file space expansion response to the host, where the file space expansion response carries the information about the second virtual space.

[0099] In this embodiment of the present invention, optionally, the management node 810 is further configured to: delete the metadata of the file, reclaim the space that is allocated from the global virtual address space to the file; determine, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node 820 that is used to store the file; and send a space reclaiming message to the storage node 820 that is used to store the file, where the space reclaiming message is used to instruct the storage node 820 that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.

[0100] The storage node 820 that is used to store the file is configured to reclaim, according to the space reclaiming message, the local virtual address space and the local physical storage space that are corresponding to the file.

[0101] The management node 810 and the storage node 820 in the distributed storage system 800 according to this embodiment of the present invention may be corresponding to a management node and a storage node in the foregoing method embodiment, and the management node 810 can execute corresponding procedures of the foregoing methods. For brevity, details are not described herein.

[0102] FIG. 9 shows a schematic block diagram of a management node 900 according to an embodiment of the present invention. As shown in FIG. 9, the management node 900 includes:

a receiving unit 910, configured to receive a file creation request sent by a host, where the file creation request is used to request to create a file in a distributed storage system;

a management unit 920, configured to: allocate, to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, where local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space; and record metadata of the file, where the metadata of the file includes

information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

a sending unit 930, configured to send a file creation success response to the host, where the file creation success response carries the information about the first virtual space.

[0103] According to the management node in this embodiment of the present invention, space is allocated to a file from global virtual address space of a distributed storage system, and the file may be managed based on the allocated global virtual address space, without a need to process complex storage location information of a data block of the file, so that file management efficiency can be improved.

[0104] In this embodiment of the present invention, optionally, an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

[0105] In this embodiment of the present invention, optionally, the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, where N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

[0106] In this embodiment of the present invention, optionally, the information about the first virtual space includes a start address and a size, of the file, in the global virtual address space.

[0107] In this embodiment of the present invention, optionally, the global virtual address space includes multiple areas, and different areas in the multiple areas have different space allocation granularities.

[0108] The management unit 920 is specifically configured to allocate, to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, where the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

[0109] In this embodiment of the present invention, optionally, the receiving unit 910 is further configured to receive a file space expansion request sent by the host,

CLAIMS

What is claimed is:

1. A file management method, wherein the method is applied to a distributed storage system, the distributed storage system comprises a management node and multiple storage nodes, and the method comprises:

receiving, by the management node, a file creation request sent by a host, wherein the file creation request is used to create a file in the distributed storage system;

allocating, by the management node to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, wherein local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space;

recording, by the management node, metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

sending, by the management node, a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

2. The method according to claim 1, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

3. The method according to claim 1 or 2, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

4. The method according to any one of claims 1 to 3, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

5. The method according to any one of claims 1 to 4, wherein the global virtual address space comprises multiple areas, and different areas in the multiple areas have different space allocation granularities; and

the allocating, by the management node to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request comprises:

allocating, by the management node to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

6. The method according to claim 5, wherein the method further comprises:

receiving, by the management node, a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand space of the file;

allocating, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area;

updating the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

sending a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

7. The method according to any one of claims 1 to 6, wherein the method further comprises:

deleting, by the management node, the metadata of the file, and reclaiming the space that is allocated from the global virtual address space to the file;

determining, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

sending a space reclaiming message to the storage node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.

8. A distributed storage system, comprising a management node and multiple storage nodes, wherein

local virtual address space of each of the multiple storage nodes is corresponding to a part of global virtual address space of the distributed storage system; and

the management node is configured to:

receive a file creation request sent by a host, wherein the file creation request is used to create a file in the distributed storage system;

allocate, to the file, first virtual space from the global virtual address space according to the file creation request;

record metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

send a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

9. The distributed storage system according to claim 8, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

10. The distributed storage system according to claim 8 or 9, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

11. The distributed storage system according to any one of claims 8 to 10, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

12. The distributed storage system according to any one of claims 8 to 11, wherein the global virtual address space comprises multiple areas, and different areas in the multiple areas have different space allocation granularities; and

the management node is specifically configured to allocate, to the file, the first virtual space from a first area in the global virtual address space of the distributed

storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

13. The distributed storage system according to claim 12, wherein the management node is further configured to:

receive a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand space of the file;

allocate, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area;

update the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

send a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

14. The distributed storage system according to any one of claims 8 to 13, wherein the management node is further configured to:

delete the metadata of the file, and reclaim the space that is allocated from the global virtual address space to the file;

determine, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

send a space reclaiming message to the storage node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file; wherein

the storage node that is used to store the file is configured to reclaim, according to the space reclaiming message, the local virtual address space and the local physical storage space that are corresponding to the file.

15. A management node, comprising:

a receiving unit, configured to receive a file creation request sent by a host, wherein the file creation request is used to create a file in a distributed storage system;

a management unit, configured to: allocate, to the file, first virtual space from global virtual address space of the distributed storage system according to the file

creation request, wherein local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space; and record metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

a sending unit, configured to send a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

16. The management node according to claim 15, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

17. The management node according to claim 15 or 16, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

18. The management node according to any one of claims 15 to 17, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

19. The management node according to any one of claims 15 to 18, wherein the global virtual address space comprises multiple areas, and different areas in the multiple areas have different space allocation granularities; and

the management unit is specifically configured to allocate, to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

20. The management node according to claim 19, wherein

the receiving unit is further configured to receive a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand

space of the file;

the management unit is further configured to: allocate, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area; and update the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

the sending unit is further configured to send a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

21. The management node according to any one of claims 15 to 20, wherein the management unit is further configured to: delete the metadata of the file, and reclaim the space that is allocated from the global virtual address space to the file; and determine, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

the sending unit is further configured to send a space reclaiming message to the storage node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.

CLAIMS

What is claimed is:

1. A file management method, wherein the method is applied to a distributed storage system, the distributed storage system comprises a management node and multiple storage nodes, and the method comprises:

receiving, by the management node, a file creation request sent by a host, wherein the file creation request is used to request to create a file in the distributed storage system;

allocating, by the management node to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, wherein local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space;

recording, by the management node, metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

sending, by the management node, a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

2. The method according to claim 1, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

3. The method according to claim 1 or 2, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

4. The method according to any one of claims 1 to 3, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

5. The method according to any one of claims 1 to 4, wherein the global virtual address space comprises multiple areas, and different areas in the multiple areas have different space allocation granularities; and

the allocating, by the management node to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request comprises:

allocating, by the management node to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

6. The method according to claim 5, wherein the method further comprises:

receiving, by the management node, a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand space of the file;

allocating, by the management node and to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area;

updating, by the management node, the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

sending, by the management node, a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

7. The method according to any one of claims 1 to 6, wherein the method further comprises:

deleting, by the management node, the metadata of the file, and reclaiming the space that is allocated from the global virtual address space to the file;

determining, by the management node and according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

sending, by the management node, a space reclaiming message to the storage

node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.

8. A distributed storage system, comprising a management node and multiple storage nodes, wherein

local virtual address space of each of the multiple storage nodes is corresponding to a part of global virtual address space of the distributed storage system; and

the management node is configured to:

receive a file creation request sent by a host, wherein the file creation request is used to request to create a file in the distributed storage system;

allocate, to the file, first virtual space from the global virtual address space according to the file creation request;

record metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

send a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

9. The distributed storage system according to claim 8, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

10. The distributed storage system according to claim 8 or 9, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

11. The distributed storage system according to any one of claims 8 to 10, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

12. The distributed storage system according to any one of claims 8 to 11, wherein the global virtual address space comprises multiple areas, and different areas

in the multiple areas have different space allocation granularities; and

the management node is specifically configured to allocate, to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is the smallest in areas that can be currently allocated.

13. The distributed storage system according to claim 12, wherein the management node is further configured to:

receive a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand space of the file;

allocate, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area;

update the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

send a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

14. The distributed storage system according to any one of claims 8 to 13, wherein the management node is further configured to:

delete the metadata of the file, and reclaim the space that is allocated from the global virtual address space to the file;

determine, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

send a space reclaiming message to the storage node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file; wherein

the storage node that is used to store the file is configured to reclaim, according to the space reclaiming message, the local virtual address space and the local physical storage space that are corresponding to the file.

15. A management node, comprising:

a receiving unit, configured to receive a file creation request sent by a host,

wherein the file creation request is used to request to create a file in a distributed storage system;

a management unit, configured to: allocate, to the file, first virtual space from global virtual address space of the distributed storage system according to the file creation request, wherein local virtual address space of each storage node in the distributed storage system is corresponding to a part of the global virtual address space; and record metadata of the file, wherein the metadata of the file comprises information about the first virtual space allocated to the file, and the information about the first virtual space is used to point to local virtual address space of a storage node that is used to store the file; and

a sending unit, configured to send a file creation success response to the host, wherein the file creation success response carries the information about the first virtual space.

16. The management node according to claim 15, wherein an address bit quantity of the global virtual address space is greater than an address bit quantity of local virtual address space of a storage node in the distributed storage system.

17. The management node according to claim 15 or 16, wherein the address bit quantity of the global virtual address space is twice the address bit quantity of the local virtual address space of the storage node in the distributed storage system, wherein N most significant bits of an address of the global virtual address space are used to point to the storage node, N least significant bits of the address of the global virtual address space are used to point to the local virtual address space of the storage node, and N is the address bit quantity of the local virtual address space of the storage node.

18. The management node according to any one of claims 15 to 17, wherein the information about the first virtual space comprises a start address and a size, of the file, in the global virtual address space.

19. The management node according to any one of claims 15 to 18, wherein the global virtual address space comprises multiple areas, and different areas in the multiple areas have different space allocation granularities; and

the management unit is specifically configured to allocate, to the file, the first virtual space from a first area in the global virtual address space of the distributed storage system according to the file creation request and a space allocation granularity of the first area, wherein the first area is an area whose space allocation granularity is

the smallest in areas that can be currently allocated.

20. The management node according to claim 19, wherein

the receiving unit is further configured to receive a file space expansion request sent by the host, wherein the file space expansion request is used to request to expand space of the file;

the management unit is further configured to: allocate, to the file, second virtual space from a second area in the global virtual address space according to the file space expansion request and a space allocation granularity of the second area, wherein the space allocation granularity of the second area is greater than the space allocation granularity of the first area; and update the metadata of the file, wherein an updated metadata of the file comprises information about the second virtual space allocated to the file; and

the sending unit is further configured to send a file space expansion response to the host, wherein the file space expansion response carries the information about the second virtual space.

21. The management node according to any one of claims 15 to 20, wherein the management unit is further configured to: delete the metadata of the file, and reclaim the space that is allocated from the global virtual address space to the file; and determine, according to a preset correspondence between the global virtual address space and the local virtual address space, the storage node that is used to store the file; and

the sending unit is further configured to send a space reclaiming message to the storage node that is used to store the file, wherein the space reclaiming message is used to instruct the storage node that is used to store the file to reclaim the local virtual address space and local physical storage space that are corresponding to the file.