

CLAIMS:

1. A method for data deduplication in a storage system, wherein the storage system comprises a controller and a storage device, and the method comprising:

receiving, by the controller, a first write request, wherein the first write request includes a first logical block address and first data to be written in the storage system;

obtaining, by the controller, a first fingerprint of the first data;

inserting, by the controller, the first fingerprint into a fingerprint table when the fingerprint table does not contain the first fingerprint, and allocating a first storage address from the storage device to the first data; wherein there is no storage address corresponding to the first fingerprint in the fingerprint table;

storing, by the controller, the first data at the first storage address of the storage device;

establishing, by the controller, a mapping relationship between the first logical block address and the first storage address;

receiving, by the controller, a second write request, wherein the second write request includes a second logical block address and second data to be written;

obtaining, by the controller, a second fingerprint of the second data, wherein the second fingerprint of the second data is identical to the first fingerprint;

allocating, by the controller, a second storage address from the storage device to the second data when the controller queries the fingerprint table according to the second fingerprint of the second data and determines that no storage address corresponding to the first fingerprint exists;

storing, by the controller, the second data at the second storage address;

establishing, by the controller, a mapping relationship between the first fingerprint and the second storage address in the fingerprint table; and

establishing, by the controller, a mapping relationship between the second logical block address and the second storage address.

2. The method according to claim 1, wherein allocating, by the controller, a first storage address from the storage device to the first data comprises: allocating, by the controller, the first storage address from a first storage area of the storage device to the first data; and allocating, by the controller, a second storage address from the storage device to the second data comprises: allocating, by the controller, the second storage address from a second storage area of the storage device to the second data, wherein storage performance of the second storage area is higher than storage performance of the first storage area.

3. The method according to claim 1, further comprising:

receiving, by the controller, a third write request, wherein the third write request includes a third logical block address and third data to be written;

obtaining, by the controller, a third fingerprint of the third data, wherein the fingerprint of the third data is identical to the first fingerprint; and

establishing, by the controller, a mapping relationship between the third logical block address and the second storage address when querying the fingerprint table according to the third fingerprint of the third data determines that a storage address corresponding to the first fingerprint is identical to the second storage address.

4. The method according to claim 1, further comprising:

recording, by the controller, a quantity of reference times of the second storage address in the fingerprint table.

5. The method according to claim 1, further comprising:

recording, by the controller in the fingerprint table, time when the first fingerprint is inserted.

6. The method according to claim 5, further comprising:

deleting, by the controller, a fourth fingerprint from the fingerprint table, in response to determining that no storage address corresponding to the fourth fingerprint exists, and duration for which the fourth fingerprint exists in the fingerprint table exceeds preset duration.

7. The method according to claim 1, further comprising:

counting, by the controller, a quantity of times that a received write request includes data to be written whose fingerprint is identical to the first fingerprint, storing, by the controller, the second data at the second storage address, and establishing the mapping relationship between the first fingerprint and the second storage address in the fingerprint table when the quantity of times is greater than a preset quantity N of times, wherein N is an integer not less than 1.

8. A storage system, comprising:

a controller; and

a storage device, wherein the controller is configured to:

receive a first write request having a first logical block address and first data to be written;

obtain a first fingerprint of the first data;

insert the first fingerprint into a fingerprint table when the fingerprint table does not contain the first fingerprint, and allocate a first storage address from the storage device to the first data; wherein there is no storage address corresponding to the first fingerprint in the fingerprint table;

store the first data at the first storage address of the storage device; and
 establish a mapping relationship between the first logical block address and the first storage address;
 wherein the controller is further configured to:
 receive a second write request having a second logical block address and second data to-be-written;
 obtain a second fingerprint of the second data, wherein the second fingerprint of the second data is identical to the first fingerprint;
 allocate a second storage address from the storage device to the second data when querying the fingerprint table according to the second fingerprint of the second data determines that no storage address corresponding to the first fingerprint exists;
 store the second data at the second storage address;
 establish a mapping relationship between the second logical block address and the second storage address; and
 establish a mapping relationship between the first fingerprint and the second storage address in the fingerprint table; and
 wherein the storage device is configured to provide the first storage address and the second storage address.

9. The storage system according to claim 8, wherein the controller is configured to allocate the first storage address from a first storage area of the storage device to the first data and allocate the second storage address from a second storage area of the storage device to the second data, wherein storage performance of the second storage area is higher than storage performance of the first storage area.

10. The storage system according to claim 8, wherein the controller is further configured to:
 receive a third write request having a third logical block address and third data to-be-written;
 obtain a third fingerprint of the third data, wherein the third fingerprint of the third data is identical to the first fingerprint;
 establish a mapping relationship between the third logical block address and the second storage address when querying the fingerprint table according to the third fingerprint of the third data determines that a storage address corresponding to the first fingerprint is identical to the second storage address.

11. The storage system according to claim 8, wherein the controller is further configured to record a quantity of reference times of the second storage address in the fingerprint table.

12. The storage system according to claim 8, wherein the controller is further configured to record, in the fingerprint table, time when the first fingerprint is inserted.

13. The storage system according to claim 12, wherein the controller is further configured to delete a fourth fingerprint, in response to determining that no storage address corresponding to the fourth fingerprint exists, and duration for which the fourth fingerprint exists in the fingerprint table exceeds preset duration.

14. The storage system according to claim 8, wherein the controller is further configured to:
count a quantity of times that a received write request includes data to be written whose fingerprint value is identical to the first fingerprint, store the second data at the second storage address, and establish the mapping relationship between the first fingerprint and the second storage address in the fingerprint table when the quantity of times is greater than a preset quantity N of times,, wherein N is an integer not less than 1.

15. A controller applied to a storage system, wherein the storage system comprises the controller and a storage device, the controller comprising:

an interface; and

a processor communicatively coupled to the processor, wherein the interface is configured to receive a first write request having a first logical block address and first data to be written;

wherein the processor is configured to:

obtain a first fingerprint of the first data;

insert the first fingerprint into a fingerprint table when the fingerprint table does not contain the first fingerprint, and allocate a first storage address from the storage device to the first data; wherein there is no storage address corresponding to the first fingerprint in the fingerprint table;

store the first data at the first storage address; and

establish a mapping relationship between the first logical block address and the first storage address;

wherein the interface is further configured to receive a second write request having a second logical block address and second data to be written; and

wherein the processor is further configured to:

obtain a second fingerprint of the second data, wherein the second fingerprint of the second data is identical to the first fingerprint;

allocate a second storage address from the storage device to the second data when querying the fingerprint table according to the second fingerprint of the second data determines that no storage address corresponding to the first fingerprint exists in the fingerprint table;

store the second data at the second storage address;

establish a mapping relationship between the second logical block address and the second storage address; and

establish a mapping relationship between the first fingerprint and the second storage address in the fingerprint table.

16. The controller according to claim 15, wherein the processor is configured to allocate the first storage address from a first storage area of the storage device to the first data and allocate the second storage address from a second storage area of the storage device to the second data, wherein storage performance of the second storage area is higher than storage performance of the first storage area.

17. The controller according to claim 15, wherein the interface is further configured to receive a third write request having a third logical block address and third data to be written; and

the processor is further configured to:

obtain a third fingerprint of the third data, wherein the third fingerprint of the third data is identical to the first fingerprint; and

establish a mapping relationship between the third logical block address and the second storage address when querying the fingerprint table according to the third fingerprint of the third data determines that a storage address corresponding to the first fingerprint is identical to the second storage address.

18. The controller according to a claim 15, wherein the processor is further configured to record a quantity of reference times of the second storage address in the fingerprint table.

19. The controller according to claim 15, wherein the processor is further configured to record, in the fingerprint table, time when the first fingerprint is inserted.

20. The controller according to claim 19, wherein the processor is further configured to delete a fourth fingerprint from the fingerprint table, in response to determining that no storage address corresponding to the fourth fingerprint exists, and duration for which the fourth fingerprint exists in the fingerprint table exceeds preset duration.

21. The controller according to claim 15, wherein the processor is further configured to: count a quantity of times that a received write request includes data to be written whose fingerprint value is identical to the first fingerprint, store the second data at the second storage address, and establish the mapping relationship between the first fingerprint and the second storage address in the fingerprint table when the quantity of times is greater than a preset quantity N of times,, wherein N is an integer not less than 1.