

CLAIMS

What is claimed is:

1. A data backup method, wherein the method is applied to a data backup system, the data backup system comprises a primary cluster, a secondary cluster, and a control device, and the method comprises:

controlling, by the control device based on a first data backup policy, the primary cluster or the secondary cluster to back up, to the secondary cluster, a plurality of data sets related to a first service that are in the primary cluster and that are at a first moment, wherein the first data backup policy comprises information about the plurality of data sets related to the first service and the first moment.

2. The method according to claim 1, wherein the controlling, by the control device based on a first data backup policy, the primary cluster or the secondary cluster to back up, to the secondary cluster, a plurality of data sets related to a first service that are in the primary cluster and that are at a first moment comprises:

sending, by the control device, a first instruction to the primary cluster, to instruct the primary cluster to send, to the secondary cluster, data corresponding to snapshots of the plurality of data sets related to the first service that are at the first moment; or

sending, by the control device, a second instruction to the secondary cluster, to instruct the secondary cluster to replicate, from the primary cluster, data corresponding to snapshots of the plurality of data sets related to the first service that are at the first moment and that are in the primary cluster.

3. The method according to claim 2, wherein before the sending, by the control device, a first instruction to the primary cluster, or the sending, by the control device, a second instruction to the secondary cluster, the method further comprises:

sending, by the control device, a third instruction to the primary cluster, wherein the third instruction comprises the information about the plurality of data sets related to the first service and the first moment, and the third instruction instructs the primary cluster to obtain the snapshots of the plurality of data sets related to the first service that are at the first moment.

4. The method according to any one of claims 1 to 3, wherein the method further comprises:

sending, by the control device, a fourth instruction to the primary cluster, wherein the fourth instruction instructs the primary cluster to synchronize user data to the secondary cluster; or

obtaining, by the control device, user data stored in the primary cluster and the secondary cluster, and adjusting, by the control device based on the user data stored in the primary cluster, the user data stored in the secondary cluster.

5. The method according to any one of claims 1 to 4, wherein the method further comprises: configuring, by the control device, the first data backup policy for the first service based on the information that is about the plurality of data sets related to the first service and that is entered by a user and the first moment.

5 6. The method according to any one of claims 1 to 5, wherein the method further comprises: configuring, by the control device, a second data backup policy for a second service, wherein the second data backup policy comprises information about a plurality of data sets related to the second service and a second moment; and

controlling, by the control device based on the second data backup policy, the primary cluster
10 or the secondary cluster to back up, to the secondary cluster, the plurality of data sets related to the second service that are in the primary cluster and that are at the second moment.

7. The method according to any one of claims 1 to 6, wherein the plurality of data sets related to the first service comprise a data set processed or stored by a first component in the primary cluster and a data set processed or stored by a second component in the primary cluster.

15 8. The method according to any one of claims 1 to 7, wherein the control device comprises a primary client and a secondary client, the primary client is configured to detect first status information of the primary cluster, the secondary client is configured to detect second status information of the secondary cluster, and the method further comprises:

obtaining, by the control device, the first status information obtained through detection of the
20 primary client and the second status information obtained through detection of the secondary client;
and

when the first status information indicates that the primary cluster is a secondary identity or
the cluster fails, and the second status information indicates that the secondary cluster is a primary
identity, determining, by the control device, that the secondary client is a client accessed by an
25 application.

9. The method according to claim 8, wherein the method further comprises:
prompting, by the control device, the user with information indicating that the primary cluster
is faulty; and

adjusting, by the control device, an identity of the secondary cluster from the secondary
30 identity to the primary identity in response to an identity adjustment operation of the user for the
secondary cluster.

10. The method according to any one of claims 1 to 9, wherein the control device is deployed
in an isolated manner from the primary cluster.

11. The method according to any one of claims 1 to 10, wherein a same clock source is set in
35 the control device, the primary cluster, and the secondary cluster.

12. The method according to any one of claims 1 to 11, wherein the primary cluster and/or the secondary cluster comprise/comprises a cluster constructed based on a hadoop architecture.

13. A data backup method, wherein the method is applied to a data backup system, the data backup system comprises a primary cluster, a secondary cluster, and a control device, and the method comprises:

obtaining, by the primary cluster, an instruction delivered by the control device, wherein the instruction comprises information about a plurality of data sets related to a first service and a first moment; and

backing up, by the primary cluster to the secondary cluster based on the instruction, the plurality of data sets related to the first service that are in the primary cluster and that are at the first moment.

14. The method according to claim 13, wherein the backing up, by the primary cluster to the secondary cluster based on the instruction, the plurality of data sets related to the first service that are in the primary cluster and that are at the first moment specifically comprises:

obtaining, by the primary cluster based on the information about the plurality of data sets related to the first service and the first moment, snapshots of the plurality of data sets related to the first service that are in the primary cluster and that are at the first moment; and

sending, by the primary cluster, data corresponding to the snapshots to the secondary cluster based on the snapshots.

15. The method according to claim 13 or 14, wherein the method further comprises:
synchronizing, by the primary cluster, user data to the secondary cluster.

16. The method according to any one of claims 13 to 15, wherein the primary cluster and/or the secondary cluster comprise/comprises a cluster constructed based on a hadoop architecture.

17. A control device, wherein the control device is located in a data backup system, the data backup system further comprises a primary cluster and a secondary cluster, and the control device comprises:

a control module, configured to control, based on a first data backup policy, the primary cluster or the secondary cluster to back up, to the secondary cluster, a plurality of data sets related to a first service that are in the primary cluster and that are at a first moment, wherein the first data backup policy comprises information about the plurality of data sets related to the first service and the first moment.

18. The control device according to claim 17, wherein the control module is specifically configured to:

send a first instruction to the primary cluster, to instruct the primary cluster to send, to the secondary cluster, data corresponding to snapshots of the plurality of data sets related to the first

service that are at the first moment; or

send a second instruction to the secondary cluster, to instruct the secondary cluster to replicate, from the primary cluster, data corresponding to snapshots of the plurality of data sets related to the first service that are at the first moment and that are in the primary cluster.

5 19. The control device according to claim 18, wherein the control device further comprises:

a communication module, configured to: before the sending, by control device, a first instruction to the primary cluster or the sending, by the control device, a second instruction to the secondary cluster, send a third instruction to the primary cluster, wherein the third instruction comprises the information about the plurality of data sets related to the first service and the first
10 moment, and the third instruction instructs the primary cluster to obtain the snapshots of the plurality of data sets related to the first service that are at the first moment.

20. The control device according to any one of claims 17 to 19, wherein the control device further comprises:

a communication module, configured to send, a fourth instruction to the primary cluster,
15 wherein the fourth instruction instructs the primary cluster to synchronize user data to the secondary cluster; or

the control module, further configured to: obtain, user data stored in the primary cluster and the secondary cluster, and adjust, based on the user data stored in the primary cluster, the user data stored in the secondary cluster.

20 21. The control device according to claims 17 to 20, wherein the control device further comprises a configuration module, configured to configure the first data backup policy for the first service based on the information that is about the plurality of data sets related to the first service and that is entered by a user and the first moment.

22. The control device according to any one of claims 17 to 21, wherein the control device
25 further comprises a configuration module, configured to configure a second data backup policy for a second service, wherein the second data backup policy comprises information about a plurality of data sets related to the second service and a second moment; and

the control module is further configured to control, based on the second data backup policy, the primary cluster or the secondary cluster to back up, to the secondary cluster, the plurality of
30 data sets related to the second service that are in the primary cluster and that are at the second moment.

23. The control device according to claims 17 to 22, wherein the plurality of data sets related to the first service comprise a data set processed or stored by a first component in the primary cluster and a data set processed or stored by a second component in the primary cluster.

35 24. The control device according to any one of claims 17 to 23, wherein the control device

comprises a primary client and a secondary client, the primary client is configured to detect first status information of the primary cluster, the secondary client is configured to detect second status information of the secondary cluster, and the control device further comprises:

a communication module, configured to obtain the first status information obtained through detection of the primary client and the second status information obtained through detection of the secondary client; and

a determining module, configured to: when the first status information indicates that the primary cluster is a secondary identity or the cluster fails, and the second status information indicates that the secondary cluster is a primary identity, determine that the secondary client is a client accessed by an application.

25. The control device according to claim 24, wherein the control device further comprises a prompting module and an adjustment module, wherein

the prompting module is configured to prompt the user with information indicating that the primary cluster is faulty; and

the adjustment module is configured to adjust an identity of the secondary cluster from the secondary identity to the primary identity in response to an identity adjustment operation of the user for the secondary cluster.

26. The control device according to any one of claims 17 to 25, wherein the control device is deployed in an isolated manner from the primary cluster.

27. The control device according to any one of claims 17 to 26, wherein a same clock source is set in the control device, the primary cluster, and the secondary cluster.

28. The control device according to any one of claims 17 to 27, wherein the primary cluster and/or the secondary cluster comprise/comprises a cluster constructed based on a hadoop architecture.

29. A data backup system, wherein the data backup system comprises a control device, a primary cluster, and a secondary cluster, wherein

the control device is configured to perform the method according to any one of the method claims 1 to 12;

the primary cluster is configured to perform the method according to any one of the method claims 13 to 16; and

the secondary cluster is configured to obtain and store a data set backed up from the primary cluster.

30. A control device, wherein the computing device comprises a processor and a memory, wherein

the processor is configured to execute instructions stored in the memory, to enable the

computing device to perform the method according to any one of claims 1 to 12.

31. A computer-readable storage medium, comprising instructions, wherein when the instructions run on a computing device, the computing device is enabled to perform the method according to any one of claims 1 to 12.

5 32. A computer program product comprising instructions, wherein when the computer program product runs on a computing device, the computing device is enabled to perform the method according to any one of claims 1 to 12.