



(12)发明专利

(10)授权公告号 CN 104202387 B

(45)授权公告日 2017. 11. 24

(21)申请号 201410427894.2

(56)对比文件

(22)申请日 2014.08.27

CN 101059807 A, 2007.10.24,

(65)同一申请的已公布的文献号

审查员 李腾飞

申请公布号 CN 104202387 A

(43)申请公布日 2014.12.10

(73)专利权人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 刘洋 张克星

(74)专利代理机构 广州三环专利商标代理有限公司

公司 44202

代理人 郝传鑫 熊永强

(51)Int. Cl.

H04L 29/08(2006.01)

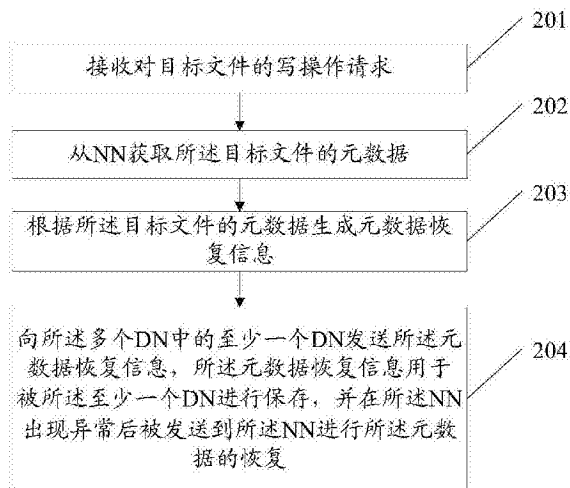
权利要求书4页 说明书17页 附图6页

(54)发明名称

一种元数据恢复方法及相关装置

(57)摘要

本发明实施例公开了一种元数据恢复方法及相关装置,该方法可包括:接收对目标文件的写操作请求,其中,所述目标文件保存在多个数据节点上;从元数据节点获取所述目标文件的元数据,所述目标文件的元数据包含所述多个数据节点的信息;根据所述目标文件的元数据生成元数据恢复信息;向所述多个数据节点中的至少一个数据节点发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个数据节点进行保存,并在所述元数据节点出现异常后被发送到所述元数据节点进行所述元数据的恢复。本发明实施例可以恢复NN异常时的元数据。



1. 一种元数据恢复方法,其特征在于,包括:

接收对目标文件的写操作请求,其中,所述目标文件保存在多个数据节点上;

从元数据节点获取所述目标文件的元数据,所述目标文件的元数据包含所述多个数据节点的信息;

根据所述目标文件的元数据生成元数据恢复信息;

向所述多个数据节点中的至少一个数据节点发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个数据节点进行保存,并在所述元数据节点出现异常后被发送到所述元数据节点进行所述元数据的恢复。

2. 如权利要求1所述的方法,其特征在于,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个数据节点上;

则,所述从元数据节点获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息包括:

从所述元数据节点获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的数据节点的信息;

根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

3. 如权利要求2所述的方法,其特征在于,所述从所述元数据节点获取所述多个数据块的元数据包括:

依次接收所述元数据节点发送的所述每个数据块的元数据;或者,接收所述元数据节点根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

4. 如权利要求2或3所述的方法,其特征在于,所述接收的元数据至少包括第一数据块的元数据和第二数据块的元数据,其中:

所述生成的所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述生成的所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

5. 一种元数据恢复方法,其特征在于,包括:

接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从元数据节点获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息.;

在所述元数据节点出现异常后,向所述元数据节点发送所述元数据恢复信息,元数据恢复信息被所述元数据节点用于进行所述元数据的恢复。

6. 如权利要求5所述的方法,其特征在于,所述目标文件被划分为多个数据块,所述接收客户端发送的目标文件的元数据恢复信息,包括:

接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

7. 如权利要求6所述的方法,其特征在于,所述接收所述客户端发送的所述目标文件的数据块的元数据恢复信息包括:

依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

8. 如权利要求6或7所述的方法,其特征在于,所述接收的元数据恢复信息至少包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

9. 一种元数据恢复方法,其特征在于,包括:

接收客户端发送的对目标文件进行写操作请求,其中,所述目标文件保存在多个数据节点上;

生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个数据节点的信息;

当本端出现异常时,接收所述多个数据节点中的至少一个数据节点发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据所述元数据生成的,由所述客户端发送至所述至少一个数据节点进行保存的元数据恢复信息。

10. 如权利要求9所述的方法,其特征在于,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个数据节点上;

则所述生成所述目标文件的元数据,并向所述客户端发送所述元数据包括:

生成所述多个数据块的元数据,并向所述客户端发送所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的数据节点的信息;

所述接收所述多个数据节点中的至少一个数据节点发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,包括:

接收所述多个数据节点中的至少一个数据节点发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

11. 如权利要求10所述的方法,其特征在于,所述向所述客户端发送所述多个数据块的元数据包括:

依次向所述客户端发送所述多个数据块的元数据;或者,根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

12. 如权利要求10或11所述的方法,其特征在于,所述多个数据块至少包括第一数据块和第二数据块,其中:

所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

13. 一种客户端,其特征在于,包括:接收单元、获取单元、生成单元和发送单元,其中:

所述接收单元,用于接收对目标文件的写操作请求,其中,所述目标文件保存在多个数据节点上;

所述获取单元,用于从元数据节点获取所述目标文件的元数据,所述目标文件的元数据包含所述多个数据节点的信息;

所述生成单元,用于根据所述目标文件的元数据生成元数据恢复信息;

所述发送单元,用于向所述多个数据节点中的至少一个数据节点发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个数据节点进行保存,并在所述元数据节点出现异常后被发送到所述元数据节点进行所述元数据的恢复。

14. 如权利要求13所述的客户端,其特征在于,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个数据节点上;

则,所述获取单元用于从所述元数据节点获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的数据节点的信息;

所述生成单元用于根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

15. 如权利要求14所述的客户端,其特征在于,所述获取单元用于依次接收所述元数据节点发送的所述每个数据块的元数据;或者,所述获取单元用于接收所述元数据节点根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

16. 如权利要求14或15所述的客户端,其特征在于,所述接收的元数据至少包括第一数据块的元数据和第二数据块的元数据,其中:

所述生成的所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述生成的所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

17. 一种数据节点,其特征在于,包括:接收单元和发送单元,其中:

所述接收单元,用于接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从元数据节点获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息.;

所述发送单元,用于在所述元数据节点出现异常后,向所述元数据节点发送所述元数据恢复信息,元数据恢复信息被所述元数据节点用于进行所述元数据的恢复。

18. 如权利要求17所述的数据节点,其特征在于,所述目标文件被划分为多个数据块,所述接收单元用于接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

19. 如权利要求18所述的数据节点,其特征在于,所述接收单元用于依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,所述接收单元用于接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

20. 如权利要求17或18所述的数据节点,其特征在于,所述接收的元数据恢复信息至少包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

21. 一种元数据节点,其特征在于,包括:接收单元、生成单元和恢复单元,其中:

所述接收单元,用于接收客户端发送的对目标文件进行写操作请求,其中,所述目标文

件保存在多个数据节点上；

所述生成单元,用于生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个数据节点的信息；

所述恢复单元,用于当本端出现异常时,接收所述多个数据节点中的至少一个数据节点发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据所述元数据生成的,由所述客户端发送至所述至少一个数据节点进行保存的元数据恢复信息。

22. 如权利要求21所述的元数据节点,其特征在于,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个数据节点上；

所述生成单元用于生成所述多个数据块的元数据,并向所述客户端发送所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的数据节点的信息；

所述恢复单元用于当本端出现异常时,接收所述多个数据节点中的至少一个数据节点发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

23. 如权利要求21所述的元数据节点,其特征在于,所述生成单元用于生成所述多个数据块的元数据,并依次向所述客户端发送所述多个数据块的元数据;或者,所述生成单元用于生成所述多个数据块的元数据,并根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

24. 如权利要求22或23所述的元数据节点,其特征在于,所述多个数据块至少包括第一数据块和第二数据块,其中:

所述第一数据块的元数据恢复信息包括:

所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

所述第二数据块的元数据恢复信息包括:

所述第二数据块在所述目标文件中的位置信息。

一种元数据恢复方法及相关装置

技术领域

[0001] 本发明涉及通信领域,尤其涉及一种元数据恢复方法及相关装置。

背景技术

[0002] Hadoop是目前通信领域中应用十分广泛的分布式系统架构,在Hadoop中包含一个非常重要的存储组件,即Hadoop分布式文件系统(Hadoop Distributed File System,简称HDFS)。其中,HDFS包括元数据节点(NameNode,简称NN)和数据节点(DataNode,简称DN),NN主要用来提供文件的元数据服务,该服务中NN在内存会生成该文件的元数据,元数据可以表示该文件的数据结构,而DN主要用来存储文件的实际数据。在实际应用中当NN出现异常时,这样该NN就会出现元数据丢失的问题,以导致该NN失效。

[0003] 其中,目前主要通过如下技术解决上述问题:

[0004] 在NN内部开启一个定时线程,该线程定期将NN内存中的元数据写到本地磁盘或者远端存储设备形成一个特定文件。当NN出现异常后,NN在重新启动过程中加载该特定文件,以恢复之前的某一时段的元数据,以实现NN恢复到这一时段的内存状态。

[0005] 但上述技术方案中,由于是定期的存储NN内的元数据,例如:每隔十分钟存储一次元数据,这样当NN出现异常的时间点还未到十分钟一次的存储时间点时,那么NN内的出现异常时的元数据不会存储,从而这部分元数据就无法恢复。可见,上述技术方案中无法恢复NN异常时的元数据。

发明内容

[0006] 本发明提供了一种元数据恢复方法及相关装置,可以恢复NN异常时的元数据。

[0007] 第一方面,本发明提供一种元数据恢复方法,包括:

[0008] 接收对目标文件的写操作请求,其中,所述目标文件保存在多个DN上;

[0009] 从NN获取所述目标文件的元数据,所述目标文件的元数据包含所述多个DN的信息;

[0010] 根据所述目标文件的元数据生成元数据恢复信息;

[0011] 向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。

[0012] 在第一方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上;

[0013] 则,所述从NN获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息包括:

[0014] 从所述NN获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息;

[0015] 根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

[0016] 结合第一方面的第一种可能的实现方式,在第二种可能实现方式中,所述从所述NN获取所述多个数据块的元数据包括:

[0017] 依次接收所述NN发送的所述每个数据块的元数据;或者,接收所述NN根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

[0018] 结合第一方面的第一种可能的实现方式或者第二种可能的实现方式,在第三种可能实现方式中,所述接收的元数据至少包括第一数据块的元数据和第二数据块的元数据,其中:

[0019] 所述生成的所述第一数据块的元数据恢复信息包括:

[0020] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0021] 所述生成的所述第二数据块的元数据恢复信息包括:

[0022] 所述第二数据块在所述目标文件中的位置信息。

[0023] 第三方面,本发明提供一种元数据恢复方法,包括:

[0024] 接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从NN获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息.;

[0025] 在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。

[0026] 在第二方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述接收客户端发送的目标文件的元数据恢复信息,包括:

[0027] 接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

[0028] 结合第二方面的第一种可能的实现方式,在第二方面的第二种可能的实现方式中,所述接收所述客户端发送的所述目标文件的数据块的元数据恢复信息包括:

[0029] 依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

[0030] 结合第二方面的第一种可能的实现方式或者第二种可能的实现方式中,在第二方面的第三种可能的实现方式中,所述接收的元数据恢复信息至少包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

[0031] 所述第一数据块的元数据恢复信息包括:

[0032] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0033] 所述第二数据块的元数据恢复信息包括:

[0034] 所述第二数据块在所述目标文件中的位置信息。

[0035] 第三方面,本发明提供一种元数据恢复方法,包括:

[0036] 接收客户端发送的对目标文件进行写操作请求,其中,所述目标文件保存在多个DN上;

[0037] 生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个DN的信息;

[0038] 当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据

所述元数据生成的,由所述客户端发送至所述至少一个DN进行保存的元数据恢复信息。

[0039] 在第三方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上;

[0040] 则所述生成所述目标文件的元数据,并向所述客户端发送所述元数据包括:

[0041] 生成所述多个数据块的元数据,并向所述客户端发送所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息;

[0042] 所述接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,包括:

[0043] 接收所述多个DN中的至少一个DN发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

[0044] 结合第三方面的第一种可能的实现方式,在第三方面的第二种可能的实现方式中,所述向所述客户端发送所述多个数据块的元数据包括:

[0045] 依次向所述客户端发送所述多个数据块的元数据;或者,根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

[0046] 结合第三方面的第一种可能的实现方式或者第二种可能的实现方式,在第三方面的第三种可能的实现方式中,所述多个数据块至少包括第一数据块和第二数据块,其中:

[0047] 所述第一数据块的元数据恢复信息包括:

[0048] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0049] 所述第二数据块的元数据恢复信息包括:

[0050] 所述第二数据块在所述目标文件中的位置信息。

[0051] 第四方面,本发明提供一种客户端,包括:接收单元、获取单元、生成单元和发送单元,其中:

[0052] 所述接收单元,用于接收对目标文件的写操作请求,其中,所述目标文件保存在多个DN上;

[0053] 所述获取单元,用于从NN获取所述目标文件的元数据,所述目标文件的元数据包含所述多个DN的信息;

[0054] 所述生成单元,用于根据所述目标文件的元数据生成元数据恢复信息;

[0055] 所述发送单元,用于向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。

[0056] 在第四方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上;

[0057] 则,所述获取单元用于从所述NN获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息;

[0058] 所述生成单元用于根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

[0059] 结合第四方面的第一种可能的实现方式,在第四方面的第二种可能的实现方式中,所述获取单元用于依次接收所述NN发送的所述每个数据块的元数据;或者,所述获取单元用于接收所述NN根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

[0060] 结合第四方面的第一种可能的实现方式或者第二种可能的实现方式,在第四方面的第三种可能的实现方式中,所述接收的元数据至少包括第一数据块的元数据和第二数据块的元数据,其中:

[0061] 所述生成的所述第一数据块的元数据恢复信息包括:

[0062] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0063] 所述生成的所述第二数据块的元数据恢复信息包括:

[0064] 所述第二数据块在所述目标文件中的位置信息。

[0065] 第五方面,本发明提供一种DN,包括:接收单元和发送单元,其中:

[0066] 所述接收单元,用于接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从NN获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息.;

[0067] 所述发送单元,用于在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。

[0068] 在第五方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述接收单元用于接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

[0069] 结合第五方面的第一种可能的实现方式,在第五方面的第二种可能的实现方式中,所述接收单元用于依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,所述接收单元用于接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

[0070] 结合第五方面的第一种可能的实现方式或者第二种可能的实现方式,在第三种可能的实现方式中,所述接收的元数据恢复信息至少包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

[0071] 所述第一数据块的元数据恢复信息包括:

[0072] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0073] 所述第二数据块的元数据恢复信息包括:

[0074] 所述第二数据块在所述目标文件中的位置信息。

[0075] 第六方面,本发明提供一种NN,包括:接收单元、生成单元和恢复单元,其中:

[0076] 所述接收单元,用于接收客户端发送的对目标文件进行写操作请求,其中,所述目标文件保存在多个DN上;

[0077] 所述生成单元,用于生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个DN的信息;

[0078] 所述恢复单元,用于当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据所述元数据生成的,由所述客户端发送至所述至少一个DN进行保存的元数据恢复信息。

[0079] 在第六方面的第一种可能的实现方式中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上;

[0080] 所述生成单元用于生成所述多个数据块的元数据,并向所述客户端发送所述多个

数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息;

[0081] 所述恢复单元用于当本端出现异常时,接收所述多个DN中的至少一个DN发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

[0082] 结合第六方面的第一种可能的实现方式,在第二种可能的实现方式中,所述生成单元用于生成所述多个数据块的元数据,并依次向所述客户端发送所述多个数据块的元数据;或者,所述生成单元用于生成所述多个数据块的元数据,并根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

[0083] 结合第六方面的第一种可能的实现方式或者第二种可能的实现方式,在第三种可能的实现方式中,所述多个数据块至少包括第一数据块和第二数据块,其中:

[0084] 所述第一数据块的元数据恢复信息包括:

[0085] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0086] 所述第二数据块的元数据恢复信息包括:

[0087] 所述第二数据块在所述目标文件中的位置信息。

[0088] 上述技术方案中,接收对目标文件的写操作请求,再从NN获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息,再向所述多个DN中的至少一个DN发送所述元数据恢复信息。这样该元数据恢复信息就可以用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。从而可以恢复NN出现异常时的元数据。

附图说明

[0089] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0090] 图1是本发明实施例提供的一种可选的元数据的示意图;

[0091] 图2是本发明实施例提供的一种元数据恢复方法的流程示意图;

[0092] 图3是本发明实施例提供的另一种元数据恢复方法的流程示意图;

[0093] 图4是本发明实施例提供的一种可选的数据信息的示意图;

[0094] 图5是本发明实施例提供的另一种元数据恢复方法的流程示意图;

[0095] 图6是本发明实施例提供的另一种元数据恢复方法的流程示意图;

[0096] 图7是本发明实施例提供的另一种元数据恢复方法的流程示意图;

[0097] 图8是本发明实施例提供的一种客户端的结构示意图;

[0098] 图9是本发明实施例提供的一种DN的结构示意图;

[0099] 图10是本发明实施例提供的一种NN的结构示意图;

[0100] 图11是本发明实施例提供的另一种元数据恢复设备的结构示意图;

[0101] 图12是本发明实施例提供的一种DN设备的结构示意图;

[0102] 图13是本发明实施例提供的一种NN设备的结构示意图。

具体实施方式

[0103] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0104] 本发明实施例中,元数据可以是用于表示文件数据结构的数据,其中,文件的元数据具体可以表示该文件所在的HDFS系统的版本信息和该文件的属性信息,例如:文件的元数据包括的HDFS系统的版本信息可以包括如下一项或者多项:

[0105] HDFS系统的版本信息(例如:imgVersion)、名称空间标识(例如:namespaceID)、HDFS系统的文件个数(例如:numFiles)和HDFS系统的时间戳(例如:genStamp);

[0106] 文件的元数据包括的该文件的属性信息可以包括如下一项或者多项:

[0107] 文件的路径(例如:Path)、文件的副本数(例如:replicas)、文件的访问时间(例如:atime)、文件的每个数据块大小(例如:blockSize)、文件的数据块的块数(例如:numBlock)、文件占据的空间的限制信息(例如:nsQuota和dsQuota)和用户权限信息(例如:username、group和perm)。

[0108] 另外,当文件包括多个数据块时,文件的元数据还可以包括各数据块的元数据,其中,数据块的元数据包括该数据块所属的文件的属性信息和该数据块的属性信息;其中,数据块的属性信息可以包括如下一项或者多项:

[0109] 数据块标识(例如:BlockId)、数据块存储的数据量(例如:numBytes)和数据块的时间戳(例如:genStamp)。另外,文件的属性信息上面已经介绍,此处不重复说明。

[0110] 例如:文件的元数据可以如图1所示,文件的元数据可以包括图1中的三行数据,其中,第一行表示HDFS系统的版本信息,第二行表示文件的属性信息,第三行表示数据块的元数据。即文件的元数据可以包括HDFS系统的版本信息、文件的属性信息和数据块的元数据。

[0111] 另外,本发明实施例应用的场景可以是单个设备,即该设备上包括HDFS的客户端、至少一个NN和至少一个DN;或者本发明实施例应用的场景可以是设备集群,即HDFS的客户端、至少一个NN和至少一个DN部署在多个设备上。其中,上述设备可以是网络侧设备,例如:基站、服务器、控制器等设备,上述设备还可以是用户侧设备,例如:计算机、平板电脑、手机、电子阅读器、遥控器、笔记本电脑、车载设备、网络电视、可穿戴设备等具有网络功能的智能设备。

[0112] 请参阅图2,图2是本发明实施例提供的一种元数据恢复方法的流程示意图,如图2所示,包括以下步骤:

[0113] 201、接收对目标文件的写操作请求,其中,所述目标文件保存在多个DN上。

[0114] 上述目标文件保存在多个DN上可以是响应上述写操作请求,将上述目标文件保存在多个DN上,其中,该保存过程的时间不作限定,例如:可以是在步骤204之后,或者在步骤204之前。

[0115] 202、从NN获取所述目标文件的元数据,所述目标文件的元数据包含所述多个DN的信息。

[0116] 可选的,当接收到上述写操作请求时,本端就可以向NN申请用于保存目标文件的

多个DN,从而NN就可以为该目标文件分配NN,以及生成该目标文件的元数据,并将该目标文件的元数据返回到上述本端。另外,上述多个DN的信息可以是各DN保存上述目标文件的路径信息,例如:上述介绍的文件的属性信息中的文件的路径。

[0117] 203、根据所述目标文件的元数据生成元数据恢复信息。

[0118] 可选的,上述元数据恢复信息可以包括该目标文件的元数据,例如:包括目标文件的属性信息。其中,该属性信息可以包括该目标文件的如下一项或者多项:

[0119] HDFS系统的版本信息、文件名、文件占据的空间的限制信息、用户权限信息、文件的每个数据块大小、文件的数据块的块数。

[0120] 其中,上述文件名还可以携带有文件的路径信息。

[0121] 204、向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。

[0122] 可选的,上述方法可以应用于HDFS的客户端(Client),即该客户端可以实现上述方法。

[0123] 上述技术方案中,接收对目标文件的写操作请求,再从NN获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息,再向所述多个DN中的至少一个DN发送所述元数据恢复信息。这样该元数据恢复信息就可以用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。从而可以恢复NN出异常时的元数据。

[0124] 请参阅图3,图3是本发明实施例提供的另一种元数据恢复方法的流程示意图,如图3所示,包括以下步骤:

[0125] 301、接收对目标文件的写操作请求,其中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上。

[0126] 可选的,对目标文件进行数据写操作可以是以数据块为单位进行写操作,即一个目标文件可以被划分为多个数据块中,其中,每个数据块的大小可以是相同。例如:每个数据块可以是64兆(MB)或者128MB等。另外,上述数据块的大小可以是数据块的存储空间大小,但在实际应用中一个文件的各数据块存储的该文件的数据量可以是不同的,例如:一个文件包含150MB数据,而每个数据块的大小为64MB,那么该文件的第一个数据块就可以存储该文件的64MB数据,而第二个数据块就可以存储该文件的64MB数据,而第三个数据块就可以存储该文件的22MB数据。

[0127] 302、从所述NN获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息。

[0128] 可选的,当接收到上述写操作请求时,本端就可以向NN申请用于保存目标文件的多个数据块,以及保存数据块的多个DN。这样从而NN就可以为该目标文件分配数据块,以及数据块所在的DN,以及生成每个数据块的元数据,并将每个数据块的元数据返回到上述本端。另外,上述数据块的信息可以是数据块的属性信息,例如:数据块标识和数据块的时间戳。上述多个DN的信息可以是各DN保存数据块的路径信息,例如:上述介绍的文件的属性信息中的文件的路径。

[0129] 303、根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

[0130] 可选的,上述数据块的元数据恢复信息可以包括该数据块的元数据,例如:包括目标文件的属性信息。

[0131] 304、向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述目标文件的元数据的恢复。

[0132] 由于上述元数据恢复信息被一个或者多个DN保存,这样可以防止单个DN出现异常后,导致该元数据恢复信息丢失。

[0133] 可选的,上述从所述NN获取所述多个数据块的元数据可以包括:

[0134] 依次接收所述NN发送的所述每个数据块的元数据。

[0135] 例如:本端可以是每次向NN申请一个数据块,这样NN就会为目标文件分配一个数据块,以及生成该数据块的元数据,并将该数据块的元数据返回到上述本端。当上述数据块被写完时,而目标文件还没有被写完,本端再向NN申请另一个数据块,NN就会为目标文件分配另一个数据块,以及生成该数据块的元数据,并将该数据块的元数据返回到上述本端。当上述数据块被写完时,而目标文件还没有被写完,继续向NN申请数据块,直到目标文件被写完。从而步骤303可以是本端每获取一个数据块的元数据,再生成该数据块的元数据恢复信息,步骤304再将该数据块的元数据恢复信息发送至上述至少一个DN。

[0136] 可选的,上述从所述NN获取所述多个数据块的元数据可以包括:

[0137] 接收所述NN根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

[0138] 例如:本端可以是一次性向NN申请多个数据块,这样NN就会为目标文件分配多个数据块,以及生成这些数据块的元数据,并将这些数据块的元数据返回到上述本端。从而步骤303可以是一次性生成多个数据块的元数据恢复信息,步骤304再将这多个数据块的元数据恢复信息发送至上述至少一个DN。

[0139] 可选的,上述接收的元数据至少包括第一数据块的元数据和第二数据块的元数据,即目标文件至少被划分为第一数据块和第二数据块。

[0140] 其中,上述生成的所述第一数据块的元数据恢复信息可以包括:

[0141] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息。

[0142] 其中,目标文件的属性信息可以包括:

[0143] HDFS系统的版本信息、文件名、文件的副本数、文件的每个数据块大小、文件占据的空间的限制信息和用户权限信息。

[0144] 上述第一数据块在所述目标文件中的位置信息可以是指,第一数据块的目标文件的数据块序号,例如:第一数据块可以是目标文件的序号为1的数据块,其中,序号为1的数据块可以是目标文件中时间戳最早的一个数据块;上述第一数据块在所述目标文件中的位置信息可以包括:

[0145] 数据块序号和文件结束数据块信息。

[0146] 上述生成的所述第二数据块的元数据恢复信息可以包括:

[0147] 所述第二数据块在所述目标文件中的位置信息。

[0148] 上述第二数据块在所述目标文件中的位置信息可以包括:

[0149] 数据块序号和文件结束数据块信息。

[0150] 另外,第二数据块的元数据恢复信息还可以包括HDFS系统的版本信息和文件名,

这样就可以清楚地知道第二数据块的元数据恢复信息是用于恢复哪个文件的数据块,以及该文件的HDFS系统的版本信息。

[0151] 例如:如图4a所示,第一数据块的元数据恢复信息可以包括:

[0152] HDFS系统的版本信息、文件名、文件的副本数、文件的每个数据块大小、文件占据的空间的限制信息、用户权限信息、数据块序号和文件结束数据块信息。

[0153] 如图4b所示,第一数据块的元数据恢复信息可以包括:

[0154] HDFS系统的版本信息、文件名、数据块序号和文件结束数据块信息。

[0155] 其中,文件结束数据块信息可以用当前数据块是否为结束Block的信息表示,例如:上述第一数据块的元数据恢复信息中的文件结束数据块信息可以用假(false)表示;当上述目标文件只包括两个数据块时,那么,上述第一数据块的元数据恢复信息中的文件结束数据块信息可以用真(true)表示;当上述目标文件包括三个或者三个以上的数据块时,上述第一数据块的元数据恢复信息和上述第二数据块的元数据恢复信息中的文件结束数据块信息可以用假(false)表示;即该实施方式中,只有目标文件的最后一个数据块对应的文件结束数据块信息才用真(true)表示,其余的都用假(false)表示,从而可以根据文件结束数据块信息可以确定目标文件包括的数据块的个数。

[0156] 另外,由于一个文件内各数据块的元数据都可以包括文件的属性信息,而文件的属性信息是相同的,例如:文件占据的空间的限制信息、用户权限信息、数据块的大小信息和副本数等信息是相同的,这样就只需要在其中一个数据块的元数据恢复信息中写入这些信息即可,从而可以节省其他DN的存储空间。

[0157] 另外,上述仅举例了所述第一数据块和第二数据块的元数据恢复信息的多种可选的实施方式,本实施例中对此不作限定,例如:第一数据块的元数据恢复信息也可以包括:

[0158] HDFS系统的版本信息、文件名、文件的副本数、文件的访问时间、文件的数据块的块数、文件占据的空间的限制信息、用户权限信息、数据块标识、数据块的大小和数据块时间戳。

[0159] 可选的,本方法中每获取NN发送的数据块的元数据之前,都可以是向NN发送数据块的申请请求,NN接收到该请求后,就可以为目标文件分配数据块,以及存储该数据块的DN。另外,在步骤301之前,所述方法还可以包括以下步骤:

[0160] 创建目标文件,并生成该目标文件的概要信息。例如:生成该文件的文件名等信息。

[0161] 可选的,上述方法可以应用于HDFS的客户端(Client),即该客户端可以实现上述方法。

[0162] 上述技术方案中,接收对目标文件的写操作请求,从NN获取多个数据块的元数据,再根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息,从而可以向所述多个DN中的至少一个DN发送所述元数据恢复信息。这样元数据恢复信息可以用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。从而当NN出现异常后,可以及时恢复NN在出现异常时生成的元数据。

[0163] 请参阅图5,图5是本发明实施例提供的另一种元数据恢复方法的流程示意图,如图5所示,包括以下步骤:

[0164] 501、接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信

息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从NN获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息。

[0165] 其中,上述目标文件所述目标文件保存在多个DN上,且上述元数据包括所述多个DN的信息。

[0166] 可选的,步骤501可以是将元数据恢复信息存储在本地,或者将元数据恢复信息存储第三方面存储设备上。

[0167] 502、在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。

[0168] 可选的,本方法可以应用于HDFS的DN。

[0169] 上述技术方案中,接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息;在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。从而可以实现恢复NN出异常时的元数据。

[0170] 请参阅图6,图6是本发明实施例提供的另一种元数据恢复方法的流程示意图,如图6所示,包括以下步骤:

[0171] 601、接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,所述客户端根据该数据块的元数据生成的元数据恢复信息。

[0172] 其中,上述目标文件被划分为多个数据块,且上述每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息。

[0173] 可选的,步骤601接收到元数据恢复信息可以是一个或者多个数据块的元数据恢复信息。例如:目标文件被划分为多个数据块,其中,执行本方法的DN可以是保存其中一个或者多个数据块的元数据恢复信息。

[0174] 可选的,步骤601可以是将上述元数据恢复信息保存在一个特定文件内,例如:保存在命名为:blk_blockId_genstamp.blockmeta的文件内,其中,该文件的名称可以包括数据块的标识,且该文件名称还可以包括数据块的时间戳,例如:DN在接收到数据块的元数据恢复信息的同时,还可以接收到客户端发送的数据的标识和数据块的时间戳。另外,该文件内还可以将客户端在上述数据块中写入的该目标文件的数据保存至另一个特定的文件内,例如:保存在命名为:blk_blockId的文件内。当客户端还写入了上述数据块的循环冗余校验码(Cyclic Redundancy Check,CRC),本方法还可以将该CRC保存用于保存上述元数据恢复信息的特定文件内,即将上述元数据恢复信息和CRC存储在同一个特定文件内,当然,本方法还可以将CRC存储在一个专门用于存储CRC的特定文件内。另外,上述两个特定文件可以是接收到上述数据信息之前设置好的,例如:接收到客户端发送的写请求时,生成上述两个特定文件。由于步骤601将上述数据信息以文件的形式进行存储,这样可以实现数据仓库工(例如:Hive)功能,或者Spark功能,其中,Spark是一个基于内存计算的开源的集群计算系统,目的是让数据分析更加快速。

[0175] 另外,上述元数据恢复信息还可以是预定义的数据格式进行保存,例如:以HBase的数据格式进行保存,其中,HBase是一个分布式的、面向列的开源数据库。

[0176] 可选的,在获取到上述元数据恢复信息时,还可以在内存中保存上述元数据恢复信息,这样可以更快速地发送该元数据恢复信息。可以直接将内存的元数据恢复信息上传

到NN。另外,当实现本方法的DN重新启动后,可以扫描本地存储的元数据恢复信息,再将该元数据恢复信息上传给NN。

[0177] 602、在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述目标文件的元数据的恢复。

[0178] 例如:DN在NN重新启动时,会接收到该NN发送的通知消息,该通知消息用于请求发送上述数据信息,这样就可以响应消息通知消息,以返回上述元数据恢复信息。

[0179] 可选的,所述接收所述客户端发送的所述目标文件的数据块的元数据恢复信息包括:

[0180] 依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

[0181] 可选的,所述接收的元数据恢复信息至少可以包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

[0182] 所述第一数据块的元数据恢复信息可以包括:

[0183] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0184] 所述第二数据块的元数据恢复信息可以包括:

[0185] 所述第二数据块在所述目标文件中的位置信息。

[0186] 另外,本实施例中,当所述生成的元数据出现异常后,可以是向NN发送目标文件的所有数据块的元数据恢复信息。例如:当DN接收到NN异常重启的通知消息后,响应该通知消息扫描本地或者内存中所有数据块的元数据恢复信息,并向NN上传所有数据块的元数据恢复信息。

[0187] 可选的,上述方法可以应用于HDFS的DN,即该DN可以实现上述方法。

[0188] 上述技术方案中,接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,并保存所述元数据恢复信息,在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。从而可以及时恢复NN在出现异常时生成的元数据。

[0189] 请参阅图7,图7是本发明实施例提供的另一种元数据恢复方法的流程示意图,如图7所示,包括以下步骤:

[0190] 701、接收客户端发送的对目标文件进行写操作请求,其中,所述目标文件保存在多个DN上。

[0191] 702、生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个DN的信息。

[0192] 可选的,步骤702可以是接收到上述写操作请求后,为上述目标文件分配所述多个DN,再生成上述目标文件的元数据。

[0193] 703、当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据所述元数据生成的,由所述客户端发送至所述至少一个DN进行保存的元数据恢复信息。

[0194] 这样当本端,即NN出现异常时,可以接收到上述至少一个DN发送的元数据恢复信息,从而可以恢复出上述目标文件的元数据。

[0195] 可选的,上述目标文件可以被划分为多个数据块,所述多个数据块被保存在所述多个DN上;

[0196] 则,上述生成所述目标文件的元数据,并向所述客户端发送所述元数据可以包括:

[0197] 生成所述多个数据块的元数据,并向所述客户端发送所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息。

[0198] 上述接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,可以包括:

[0199] 接收所述多个DN中的至少一个DN发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

[0200] 这样可以实现根据多个数据块的元数据恢复信息恢复目标文件的元数据。

[0201] 可选的,向所述客户端发送所述多个数据块的元数据包括:

[0202] 依次向所述客户端发送所述多个数据块的元数据;或者,根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

[0203] 可选的,上述多个数据块至少可以包括第一数据块和第二数据块,其中:

[0204] 所述第一数据块的元数据恢复信息可以包括:

[0205] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0206] 所述第二数据块的元数据恢复信息可以包括:

[0207] 所述第二数据块在所述目标文件中的位置信息。

[0208] 需要说明的是,本实施例中针对元数据和元数据恢复信息的定义可以参考图2或者图3所示的实施例,此处不作重复说明。

[0209] 可选的,上述方法可以应用于HDFS内的NN。

[0210] 上述技术方案中,接收客户端发送的对目标文件进行写操作请求,生成所述目标文件的元数据,并向所述客户端发送所述元数据;这样当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据。从而可以恢复NN出异常时的元数据。

[0211] 下面为本发明装置实施例,本发明装置实施例用于执行本发明方法实施例一至五实现的方法,为了便于说明,仅示出了与本发明实施例相关的部分,具体技术细节未揭示的,请参照本发明实施例一、实施例二、实施例三、实施例四和实施例五。

[0212] 请参阅图8,图8是本发明实施例提供的一种客户端的结构示意图,如图8所示,包括:接收单元81、获取单元82、生成单元83和发送单元84,其中:

[0213] 所述接收单元81,用于接收对目标文件的写操作请求,其中,所述目标文件保存在多个DN上。

[0214] 所述获取单元82,用于从NN获取所述目标文件的元数据,所述目标文件的元数据包含所述多个DN的信息。

[0215] 所述生成单元83,用于根据所述目标文件的元数据生成元数据恢复信息。

[0216] 所述发送单元84,用于向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。

[0217] 在另一个实施例中,所述目标文件被划分为多个数据块,所述多个数据块被保存

在所述多个DN上；

[0218] 则,获取单元82可以用于从所述NN获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息；

[0219] 生成单元83可以用于根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

[0220] 该实施方式中,可以实现生成每个数据块的元数据恢复信息,以使NN根据多个数据块的元数据恢复信息恢复目标文件的元数据。

[0221] 可选的,所述获取单元82可以用于依次接收所述NN发送的所述每个数据块的元数据;或者,所述获取单元82可以用于接收所述NN根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

[0222] 可选的,上述接收的元数据至少可以包括第一数据块的元数据和第二数据块的元数据,其中:

[0223] 所述生成的所述第一数据块的元数据恢复信息可以包括:

[0224] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0225] 所述生成的所述第二数据块的元数据恢复信息可以包括:

[0226] 所述第二数据块在所述目标文件中的位置信息。

[0227] 可选的,上述客户端可以是HDFS的客户端(Client)。

[0228] 上述技术方案中,接收对目标文件的写操作请求,再从NN获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息,再向所述多个DN中的至少一个DN发送所述元数据恢复信息。这样该元数据恢复信息就可以用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。从而可以恢复NN出异常时的元数据。

[0229] 请参阅图9,图9是本发明实施例提供的一种DN的结构示意图,如图9所示,包括:接收单元91和发送单元92,其中:

[0230] 接收单元91,用于接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从NN获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息。。

[0231] 发送单元92,用于在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。

[0232] 在另一个实施例中,所述目标文件被划分为多个数据块,所述接收单元91可以用于接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

[0233] 可选的,接收单元91可以用于依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,所述接收单元91可以用于接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

[0234] 可选的,上述接收的元数据恢复信息至少可以包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

[0235] 所述第一数据块的元数据恢复信息可以包括:

[0236] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

- [0237] 所述第二数据块的元数据恢复信息可以包括：
- [0238] 所述第二数据块在所述目标文件中的位置信息。
- [0239] 可选的，本节点可以是HDFS的DN。
- [0240] 上述技术方案中，接收客户端发送的目标文件的元数据恢复信息，并保存所述元数据恢复信息；在所述NN出现异常后，向所述NN发送所述元数据恢复信息，元数据恢复信息被所述NN用于进行所述元数据的恢复。从而可以实现恢复NN出异常时的元数据。
- [0241] 请参阅图10，图10是本发明实施例提供的一种NN的结构示意图，如图10所示，包括：接收单元101、生成单元102和恢复单元103，其中：
- [0242] 所述接收单元101，用于接收客户端发送的对目标文件进行写操作请求，其中，所述目标文件保存在多个DN上。
- [0243] 所述生成单元102，用于生成所述目标文件的元数据，并向所述客户端发送所述元数据，其中，所述元数据包含所述多个DN的信息。
- [0244] 所述恢复单元103，用于当本端出现异常时，接收所述多个DN中的至少一个DN发送的元数据恢复信息，并根据所述元数据恢复信息恢复所述元数据，其中，所述元数据恢复信息是所述客户端根据所述元数据生成的，由所述客户端发送至所述至少一个DN进行保存的元数据恢复信息。
- [0245] 在另一个实施例中，所述目标文件被划分为多个数据块，所述多个数据块被保存在所述多个DN上；
- [0246] 所述生成单元102可以用于生成所述多个数据块的元数据，并向所述客户端发送所述多个数据块的元数据，每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息；
- [0247] 所述恢复单元103可以用于当本端出现异常时，接收所述多个DN中的至少一个DN发送的所述多个数据块的元数据恢复信息，并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。
- [0248] 可选的，生成单元102可以用于生成所述多个数据块的元数据，并依次向所述客户端发送所述多个数据块的元数据；或者，所述生成单元102可以用于生成所述多个数据块的元数据，并根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。
- [0249] 可选的，上述多个数据块至少可以包括第一数据块和第二数据块，其中：
- [0250] 所述第一数据块的元数据恢复信息可以包括：
- [0251] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息；
- [0252] 所述第二数据块的元数据恢复信息可以包括：
- [0253] 所述第二数据块在所述目标文件中的位置信息。
- [0254] 可选的，上述节点可以是HDFS内的NN。
- [0255] 上述技术方案中，接收客户端发送的对目标文件进行写操作请求，生成所述目标文件的元数据，并向所述客户端发送所述元数据；这样当本端出现异常时，接收所述多个DN中的至少一个DN发送的元数据恢复信息，并根据所述元数据恢复信息恢复所述元数据。从而可以恢复NN出异常时的元数据。
- [0256] 请参阅图11，图11是本发明实施例提供的另一种元数据恢复设备的结构示意图，

包括至少一个处理器1102(例如CPU),至少一个网络接口1105或者其他通信接口,存储器1106,和至少一个通信总线1103,用于实现这些装置之间的连接通信。处理器1102用于执行存储器1106中存储的可执行模块,例如计算机程序。存储器1106可能包含高速随机存取存储器(RAM:Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过至少一个网络接口1105(可以是有线或者无线)实现该系统网关与至少一个其他网元之间的通信连接,可以使用互联网,广域网,本地网,城域网等。

[0257] 在一些实施方式中,存储器1106存储了程序11061,程序11061可以被处理器1102执行,这个程序包括:

[0258] 接收对目标文件的写操作请求,其中,所述目标文件保存在多个DN上;

[0259] 从NN获取所述目标文件的元数据,所述目标文件的元数据包含所述多个DN的信息;

[0260] 根据所述目标文件的元数据生成元数据恢复信息;

[0261] 向所述多个DN中的至少一个DN发送所述元数据恢复信息,所述元数据恢复信息用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复。

[0262] 在另一个实施例中,所述目标文件可以被划分为多个数据块,所述多个数据块可以被保存在所述多个DN上;处理器1102执行的从NN获取所述目标文件的元数据,根据所述目标文件的元数据生成元数据恢复信息的程序,可以包括:

[0263] 从所述NN获取所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息;

[0264] 根据所述每个数据块的元数据,生成所述每个数据块的元数据恢复信息。

[0265] 可选的,处理器1102执行的从所述NN获取所述多个数据块的元数据的程序,可以包括:

[0266] 依次接收所述NN发送的所述每个数据块的元数据;或者,接收所述NN根据所述目标文件的写操作请求一次性发送的所有数据块的元数据。

[0267] 可选的,所述接收的元数据至少可以包括第一数据块的元数据和第二数据块的元数据,其中:

[0268] 所述生成的所述第一数据块的元数据恢复信息可以包括:

[0269] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0270] 所述生成的所述第二数据块的元数据恢复信息可以包括:

[0271] 所述第二数据块在所述目标文件中的位置信息。

[0272] 可选的,上述设备可以是HDFS的客户端(Client)所属的设备。

[0273] 上述技术方案中,接收对目标文件的写操作请求;从NN获取所述目标文件的元数据;根据所述目标文件的元数据生成元数据恢复信息;向所述多个DN中的至少一个DN发送所述元数据恢复信息。这样元数据恢复信息可以用于被所述至少一个DN进行保存,并在所述NN出现异常后被发送到所述NN进行所述元数据的恢复,从而可以恢复NN出异常时的元数据。

[0274] 请参阅图12,图12是本发明实施例提供的一种DN设备的结构示意图,包括至少一

个处理器1202(例如CPU),至少一个网络接口1205或者其他通信接口,存储器1206,和至少一个通信总线1203,用于实现这些装置之间的连接通信。处理器1202用于执行存储器1206中存储的可执行模块,例如计算机程序。存储器1206可能包含高速随机存取存储器(RAM: Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过至少一个网络接口1205(可以是有线或者无线)实现该系统网关与至少一个其他网元之间的通信连接,可以使用互联网,广域网,本地网,城域网等。

[0275] 在一些实施方式中,存储器1206存储了程序12061,程序12061可以被处理器1202执行,这个程序包括:

[0276] 接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息,其中,所述元数据恢复信息是所述客户端在接收对目标文件的写操作请求时,从NN获取的所述目标文件的元数据,并根据所述元数据生成的元数据恢复信息.;

[0277] 在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。

[0278] 在另一个实施例中,所述目标文件被划分为多个数据块,处理器1202执行的接收客户端发送的目标文件的元数据恢复信息的程序,可以包括:

[0279] 接收所述客户端发送的所述目标文件的数据块的元数据恢复信息,所述数据块的元数据恢复信息为所述客户端根据该数据块的元数据生成的元数据恢复信息。

[0280] 可选的,处理器1202执行的接收所述客户端发送的所述目标文件的数据块的元数据恢复信息的程序,可以包括:

[0281] 依次接收所述客户端发送的所述目标文件的多个数据块的元数据恢复信息;或者,接收所述客户端一次性发送的所述目标文件的多个数据块的元数据恢复信息。

[0282] 可选的,所述接收的元数据恢复信息至少可以包括第一数据块的元数据恢复信息和第二数据块的元数据恢复信息,其中:

[0283] 所述第一数据块的元数据恢复信息可以包括:

[0284] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0285] 所述第二数据块的元数据恢复信息可以包括:

[0286] 所述第二数据块在所述目标文件中的位置信息。

[0287] 可选的,本节点设备可以是HDFS的DN所属的设备。

[0288] 上述技术方案中,接收客户端发送的目标文件的元数据恢复信息,并保存所述元数据恢复信息;在所述NN出现异常后,向所述NN发送所述元数据恢复信息,元数据恢复信息被所述NN用于进行所述元数据的恢复。从而可以实现恢复NN出异常时的元数据。

[0289] 请参阅图13,图13是本发明实施例提供的一种NN设备的结构示意图,包括至少一个处理器1302(例如CPU),至少一个网络接口1305或者其他通信接口,存储器1306,和至少一个通信总线1303,用于实现这些装置之间的连接通信。处理器1302用于执行存储器1306中存储的可执行模块,例如计算机程序。存储器1306可能包含高速随机存取存储器(RAM: Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。通过至少一个网络接口1305(可以是有线或者无线)实现该系统网关与至少一个其他网元之间的通信连接,可以使用互联网,广域网,本地网,城域网等。

[0290] 在一些实施方式中,存储器1306存储了程序13061,程序13061可以被处理器1302

执行,这个程序包括:

[0291] 接收客户端发送的对目标文件进行写操作请求,其中,所述目标文件保存在多个DN上;

[0292] 生成所述目标文件的元数据,并向所述客户端发送所述元数据,其中,所述元数据包含所述多个DN的信息;

[0293] 当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据,其中,所述元数据恢复信息是所述客户端根据所述元数据生成的,由所述客户端发送至所述至少一个DN进行保存的元数据恢复信息。

[0294] 在另一个实施例中,所述目标文件被划分为多个数据块,所述多个数据块被保存在所述多个DN上;处理器1302执行的生成所述目标文件的元数据,并向所述客户端发送所述元数据,可以包括:

[0295] 生成所述多个数据块的元数据,并向所述客户端发送所述多个数据块的元数据,每个数据块的元数据包含每个数据块的信息和所述每个数据块所在的DN的信息。

[0296] 处理器1302执行的接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据的程序,可以包括:

[0297] 接收所述多个DN中的至少一个DN发送的所述多个数据块的元数据恢复信息,并根据所述多个数据块的元数据恢复信息恢复所述目标文件的元数据。

[0298] 可选的,处理器1302执行的向所述客户端发送所述多个数据块的元数据的程序,可以包括:

[0299] 依次向所述客户端发送所述多个数据块的元数据;或者,根据所述目标文件的写操作请求一次性向所述客户端发送所有数据块的元数据。

[0300] 可选的,上述多个数据块至少可以包括第一数据块和第二数据块,其中:

[0301] 所述第一数据块的元数据恢复信息可以包括:

[0302] 所述目标文件的属性信息和所述第一数据块在所述目标文件中的位置信息;

[0303] 所述第二数据块的元数据恢复信息可以包括:

[0304] 所述第二数据块在所述目标文件中的位置信息。

[0305] 可选的,上述节点设备可以是HDFS内的NN所属的设备。

[0306] 上述技术方案中,接收客户端发送的对目标文件进行写操作请求;生成所述目标文件的元数据,并向所述客户端发送所述元数据;当本端出现异常时,接收所述多个DN中的至少一个DN发送的元数据恢复信息,并根据所述元数据恢复信息恢复所述元数据。从而可以恢复NN出异常时的元数据。

[0307] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)或随机存取存储器(Random Access Memory,简称RAM)等。

[0308] 以上所揭露的仅为本发明较佳实施例而已,当然不能以此来限定本发明之权利范围,因此依本发明权利要求所作的等同变化,仍属本发明所涵盖的范围。

| | | | | | | | | | | |
|------|--------|------|-------|------|---------|--------|-------|-------|-----|--|
| 版本信息 | 名称空间标识 | 文件个数 | 时间戳 | | | | | | | |
| 路径 | 副本数 | 访问时间 | 数据块大小 | 数据块数 | 空间的限制信息 | 用户权限信息 | | | | |
| 路径 | 副本数 | 访问时间 | 数据块大小 | 数据块数 | 空间的限制信息 | 用户权限信息 | 数据块标识 | 存储数据量 | 时间戳 | |

图1

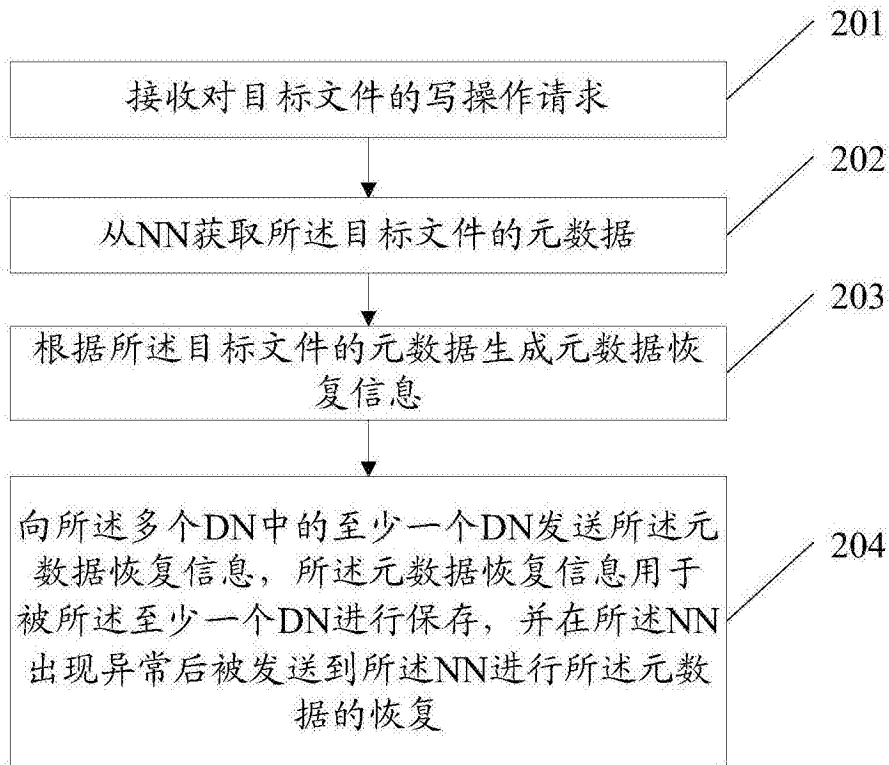


图2

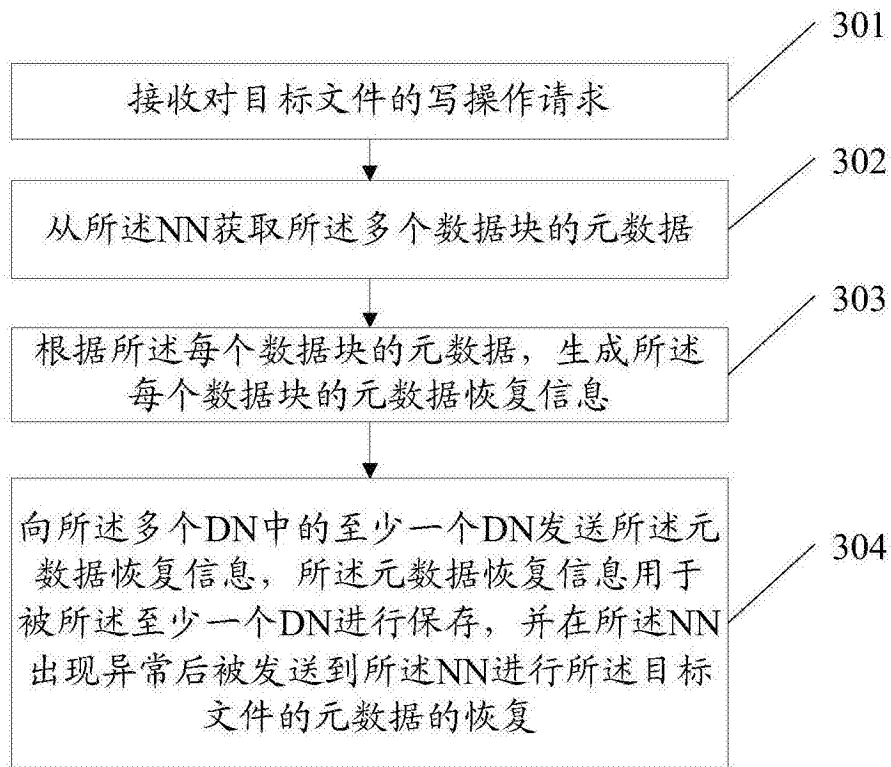


图3

| | | | | | | | |
|-------|-----|---------|--------|-------|-----|-------|---------|
| 系统版本号 | 文件名 | 空间的限制信息 | 用户权限信息 | 数据块大小 | 副本数 | 数据块序号 | 文件结束数据块 |
|-------|-----|---------|--------|-------|-----|-------|---------|

4a

| | | | |
|-------|-----|-------|---------|
| 系统版本号 | 文件名 | 数据块序号 | 文件结束数据块 |
|-------|-----|-------|---------|

4b

图4

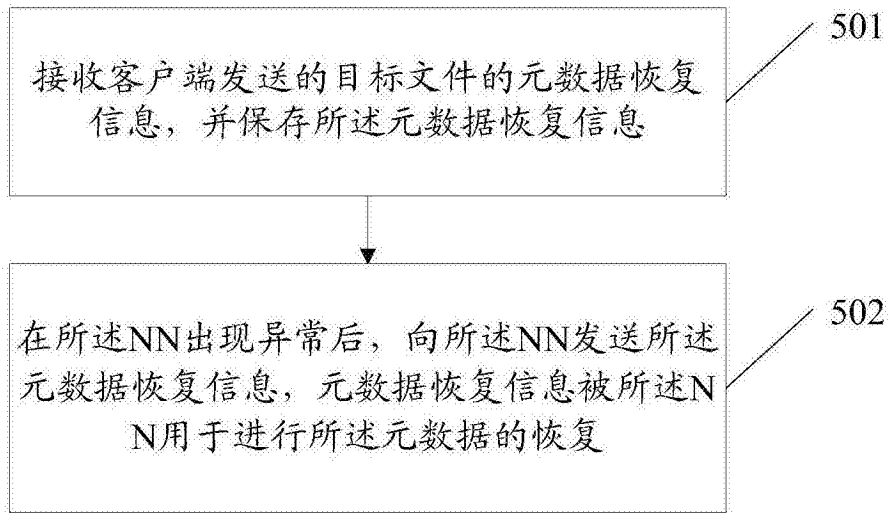


图5

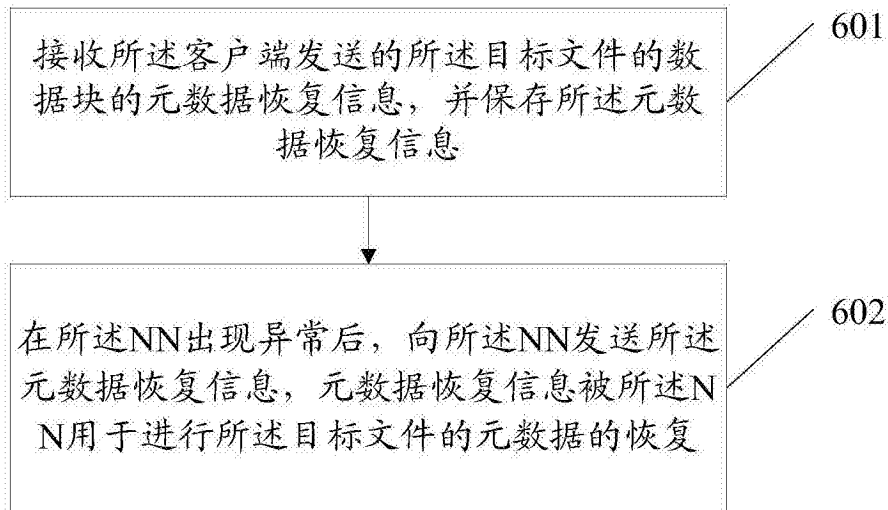


图6

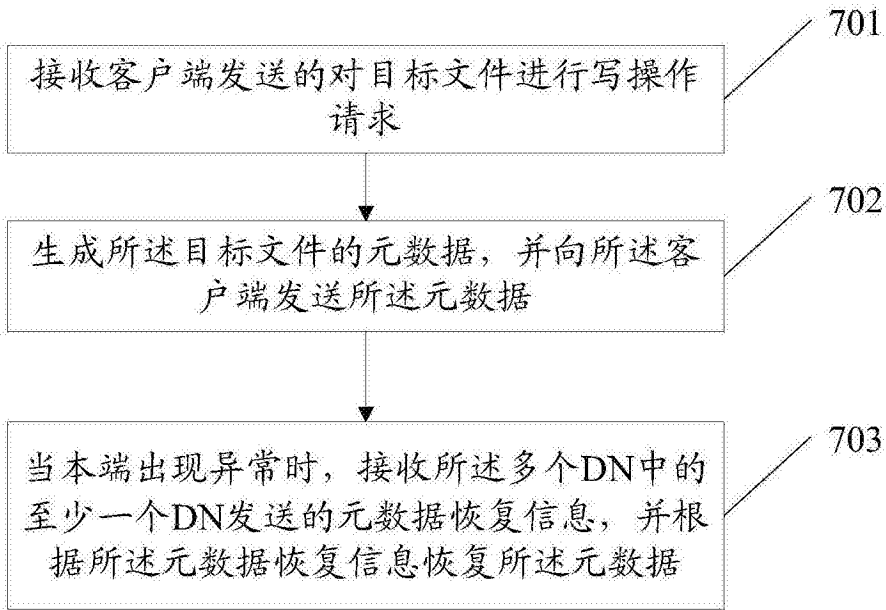


图7

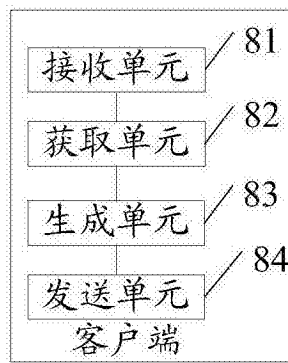


图8

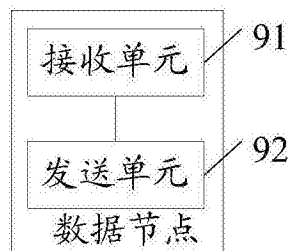


图9

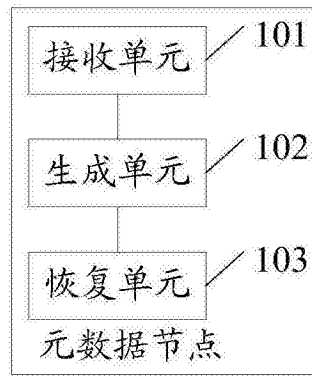


图10

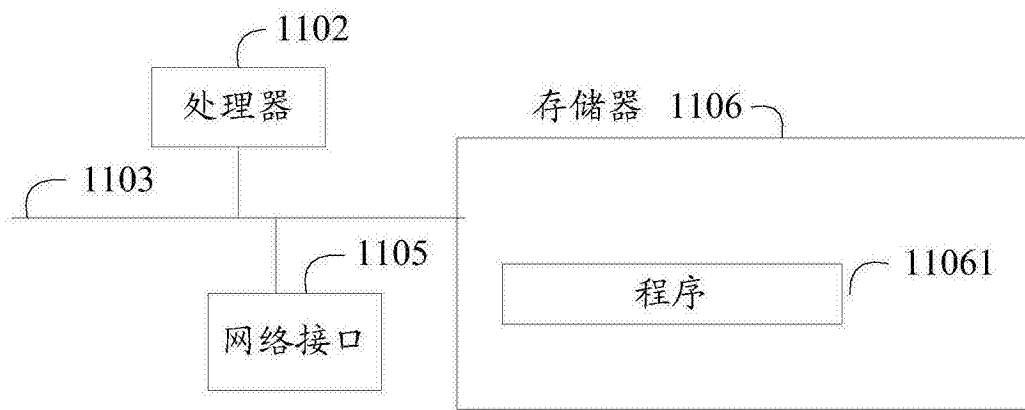


图11

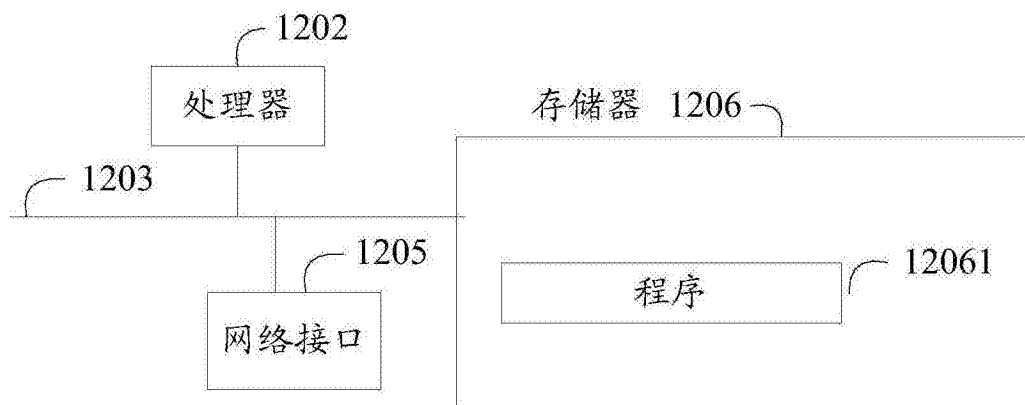


图12

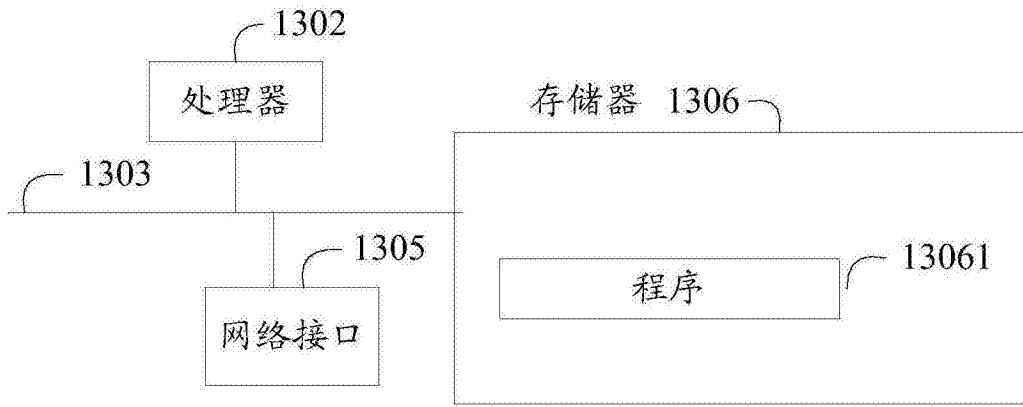


图13