



(12)发明专利

(10)授权公告号 CN 105573861 B

(45)授权公告日 2019.01.25

(21)申请号 201510897346.0

G06F 16/182(2019.01)

(22)申请日 2015.12.08

(56)对比文件

(65)同一申请的已公布的文献号
申请公布号 CN 105573861 A

CN 102063438 A,2011.05.18,
US 2008/0126832 A1,2008.05.29,
CN 102567136 A,2012.07.11,

(43)申请公布日 2016.05.11

审查员 万洋

(73)专利权人 曙光信息产业股份有限公司
地址 300384 天津市西青区华苑产业区(环
外)海泰华科大街15号1-3层

(72)发明人 康撼宇 郭照斌 姜国梁 季旻
苗艳超 杨鹏

(74)专利代理机构 北京新知远方知识产权代理
事务所(普通合伙) 11397
代理人 艾凤英

(51)Int.Cl.

G06F 11/14(2006.01)

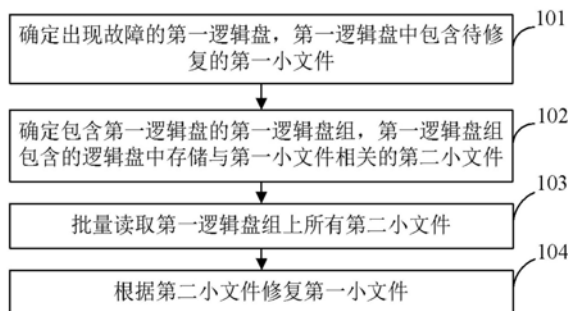
权利要求书2页 说明书9页 附图5页

(54)发明名称

分布式文件系统中小文件的修复方法和装置

(57)摘要

本发明提供了一种分布式文件系统中小文件的修复方法和装置,属于计算机技术领域。所述方法包括:确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;批量读取第一逻辑盘组上所有第二小文件;根据第二小文件修复第一小文件。本发明确定出现故障的逻辑盘,该逻辑盘包含待修复的第一小文件;批量读取包含该逻辑盘的逻辑盘组上所有与第一小文件相关的第二小文件;根据第二小文件修复第一小文件,使逻辑盘组中不同逻辑盘的相关小文件被批量读取、修复,避免逐个修复小文件带来的频繁小块读写操作,提高小文件的修复速度,提升修复性能。



1. 一种分布式文件系统中小文件的修复方法,其特征在于,包括:
 - 确定出现故障的第一逻辑盘,所述第一逻辑盘中包含待修复的第一小文件;
 - 确定包含所述第一逻辑盘的第一逻辑盘组,所述第一逻辑盘组包含的逻辑盘中存储与所述第一小文件相关的第二小文件;
 - 批量读取所述第一逻辑盘组上所有第二小文件;
 - 根据所述第二小文件修复所述第一小文件;
 - 所述确定出现故障的第一逻辑盘之前,还包括:
 - 将磁盘切分为至少一个第二逻辑盘;
 - 将存储第三小文件的第二逻辑盘组成第二逻辑盘组;
 - 所述第二逻辑盘组中存储的第三小文件互为相关。
2. 根据权利要求1所述的方法,其特征在于,所述第二小文件为所述第一小文件和所述第一小文件的备份小文件,或所述第一小文件和与所述第一小文件为同一文件的部分数据形成的小文件,或所述第一小文件、所述第一小文件的备份小文件和与所述第一小文件为同一文件的部分数据形成的小文件。
3. 根据权利要求2所述的方法,其特征在于,所述第二逻辑盘组中存储的第三小文件互为相关,包括:所述第二逻辑盘组中存储的所有第三小文件互为备份,或所述第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或所述第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。
4. 根据权利要求2所述的方法,其特征在于,若组成第二逻辑盘组的第二逻辑盘为多个,则组成所述第二逻辑盘组的第二逻辑盘位于不同节点上。
5. 根据权利要求4所述的方法,其特征在于,节点出现故障,所述确定出现故障的第一逻辑盘,包括:
 - 确定出现故障的节点上的各磁盘;
 - 确定各磁盘切分出的第三逻辑盘;
 - 将所有第三逻辑盘均作为第一逻辑磁盘。
6. 根据权利要求4所述的方法,其特征在于,磁盘出现故障,所述确定出现故障的第一逻辑盘,包括:
 - 确定出现故障的磁盘切分出的第四逻辑盘;
 - 将所有第四逻辑盘均作为第一逻辑磁盘。
7. 根据权利要求1所述的方法,其特征在于,所述根据所述第二小文件修复所述第一小文件之后,还包括:
 - 将修复后的第一小文件写入第五逻辑盘,所述第五逻辑盘与所述第一逻辑盘的盘号相同,所述第五逻辑盘的物理位置与所述第一逻辑盘的物理位置不同;
 - 将所述第一逻辑盘映射至所述第五逻辑盘。
8. 一种分布式文件系统中小文件的修复装置,其特征在于,包括:
 - 第一确定模块,用于确定出现故障的第一逻辑盘,所述第一逻辑盘中包含待修复的第一小文件;
 - 第二确定模块,用于确定包含所述第一确定模块确定的第一逻辑盘的第一逻辑盘组,所述第一逻辑盘组包含的逻辑盘中存储与所述第一小文件相关的第二小文件;

读取模块,用于批量读取所述第二确定模块确定的第一逻辑盘组上所有第二小文件;
修复模块,用于根据所述读取模块读取的第二小文件修复所述第一小文件;
所述装置,还包括:

切分模块,用于将磁盘切分为至少一个第二逻辑盘;

组合模块,用于将存储第三小文件的第二逻辑盘组成第二逻辑盘组;

所述第二逻辑盘组中存储的第三小文件互为相关。

9. 根据权利要求8所述的装置,其特征在于,所述第二小文件为所述第一小文件和所述第一小文件的备份小文件,或所述第一小文件和与所述第一小文件为同一文件的部分数据形成的小文件,或所述第一小文件、所述第一小文件的备份小文件和与所述第一小文件为同一文件的部分数据形成的小文件。

10. 根据权利要求9所述的装置,其特征在于,所述第二逻辑盘组中存储的第三小文件互为相关,包括:所述第二逻辑盘组中存储的所有第三小文件互为备份,或所述第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或所述第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。

11. 根据权利要求9所述的装置,其特征在于,若组成第二逻辑盘组的第二逻辑盘为多个,则组成所述第二逻辑盘组的第二逻辑盘位于不同节点上。

12. 根据权利要求11所述的装置,其特征在于,节点出现故障,所述第一确定模块,包括:

第一确定单元,用于确定出现故障的节点上的各磁盘;

第二确定单元,用于确定所述第一确定单元确定的各磁盘切分出的第三逻辑盘;

第三确定单元,用于将所述第二确定单元确定的所有第三逻辑盘均作为第一逻辑磁盘。

13. 根据权利要求11所述的装置,其特征在于,磁盘出现故障,所述第一确定模块,包括:

第四确定单元,用于确定出现故障的磁盘切分出的第四逻辑盘;

第五确定单元,用于将所述第四确定单元确定的所有第四逻辑盘均作为第一逻辑磁盘。

14. 根据权利要求8所述的装置,其特征在于,所述装置,还包括:

写入模块,用于将修复后的第一小文件写入第五逻辑盘,所述第五逻辑盘与所述第一逻辑盘的盘号相同,所述第五逻辑盘的物理位置与所述第一逻辑盘的物理位置不同;

映射模块,用于将所述第一逻辑盘映射至所述第五逻辑盘。

分布式文件系统中小文件的修复方法和装置

技术领域

[0001] 本发明涉及计算机技术领域,尤其涉及分布式文件系统中小文件的修复方法和装置。

背景技术

[0002] 对于分布式文件系统,每块磁盘上可以存储大量文件。对于小文件而言,每块磁盘上存储的数据量更加可观。当磁盘出现故障时,会造成大量小文件出现部分数据不可用的情况。分布式文件系统会对存在故障的小文件进行修复。

[0003] 分布式文件系统对磁盘故障的容忍能力有限,为了避免修复完成前,由于系统中产生新的故障造成某些小文件的完全损坏,需要在尽可能短的时间内,对存在故障的小文件进行修复。

[0004] 目前,分布式文件系统中小文件的修复方法为:对每个小文件,进行读取、修复。当修复的小文件个数较多时,目前的方法会造成磁盘上大量小块随机读取,由于分布式文件系统采用的存储介质在进行小块随机读取时性能较低,会导致修复性能偏低。

发明内容

[0005] 为解决上述问题,本发明实施例提出了一种分布式文件系统中小文件的修复方法和装置。

[0006] 一方面,本发明实施例提供了一种分布式文件系统中小文件的修复方法,包括:

[0007] 确定出现故障的第一逻辑盘,所述第一逻辑盘中包含待修复的第一小文件;

[0008] 确定包含所述第一逻辑盘的第一逻辑盘组,所述第一逻辑盘组包含的逻辑盘中存储与所述第一小文件相关的第二小文件;

[0009] 批量读取所述第一逻辑盘组上所有第二小文件;

[0010] 根据所述第二小文件修复所述第一小文件。

[0011] 可选地,所述第二小文件为所述第一小文件和所述第一小文件的备份小文件,或所述第一小文件和与所述第一小文件为同一文件的部分数据形成的小文件,或所述第一小文件、所述第一小文件的备份小文件和与所述第一小文件为同一文件的部分数据形成的小文件。

[0012] 可选地,所述确定出现故障的第一逻辑盘之前,还包括:

[0013] 将磁盘切分为至少一个第二逻辑盘;

[0014] 将存储第三小文件的第二逻辑盘组成第二逻辑盘组;

[0015] 所述第二逻辑盘组中存储的第三小文件互为相关。

[0016] 可选地,所述第二逻辑盘组中存储的第三小文件互为相关,包括:所述第二逻辑盘组中存储的所有第三小文件互为备份,或所述第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或所述第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。

- [0017] 可选地,若组成第二逻辑盘组的第二逻辑盘为多个,则组成所述第二逻辑盘组的第二逻辑盘位于不同节点上。
- [0018] 可选地,节点出现故障,所述确定出现故障的第一逻辑盘,包括:
- [0019] 确定出现故障的节点上的各磁盘;
- [0020] 确定各磁盘切分出的第三逻辑盘;
- [0021] 将所有第三逻辑盘均作为第一逻辑磁盘。
- [0022] 可选地,磁盘出现故障,所述确定出现故障的第一逻辑盘,包括:
- [0023] 确定出现故障的磁盘切分出的第四逻辑盘;
- [0024] 将所有第四逻辑盘均作为第一逻辑磁盘。
- [0025] 可选地,所述根据所述第二小文件修复所述第一小文件之后,还包括:
- [0026] 将修复后的第一小文件写入第五逻辑盘,所述第五逻辑盘与所述第一逻辑盘的盘号相同,所述第五逻辑盘的物理位置与所述第一逻辑盘的物理位置不同;
- [0027] 将所述第一逻辑盘映射至所述第五逻辑盘。
- [0028] 另一方面,本发明实施例提供了一种分布式文件系统中小文件的修复装置,包括:
- [0029] 第一确定模块,用于确定出现故障的第一逻辑盘,所述第一逻辑盘中包含待修复的第一小文件;
- [0030] 第二确定模块,用于确定包含所述第一确定模块确定的第一逻辑盘的第一逻辑盘组,所述第一逻辑盘组包含的逻辑盘中存储与所述第一小文件相关的第二小文件;
- [0031] 读取模块,用于批量读取所述第二确定模块确定的第一逻辑盘组上所有第二小文件;
- [0032] 修复模块,用于根据所述读取模块读取的第二小文件修复所述第一小文件。
- [0033] 可选地,所述第二小文件为所述第一小文件和所述第一小文件的备份小文件,或所述第一小文件和与所述第一小文件为同一文件的部分数据形成的小文件,或所述第一小文件、所述第一小文件的备份小文件和与所述第一小文件为同一文件的部分数据形成的小文件。
- [0034] 可选地,所述装置,还包括:
- [0035] 切分模块,用于将磁盘切分为至少一个第二逻辑盘;
- [0036] 组合模块,用于将存储第三小文件的第二逻辑盘组成第二逻辑盘组;
- [0037] 所述第二逻辑盘组中存储的第三小文件互为相关。
- [0038] 可选地,所述第二逻辑盘组中存储的第三小文件互为相关,包括:所述第二逻辑盘组中存储的所有第三小文件互为备份,或所述第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或所述第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。
- [0039] 可选地,若组成第二逻辑盘组的第二逻辑盘为多个,则组成所述第二逻辑盘组的第二逻辑盘位于不同节点上。
- [0040] 可选地,节点出现故障,所述第一确定模块,包括:
- [0041] 第一确定单元,用于确定出现故障的节点上的各磁盘;
- [0042] 第二确定单元,用于确定所述第一确定单元确定的各磁盘切分出的第三逻辑盘;
- [0043] 第三确定单元,用于将所述第二确定单元确定的所有第三逻辑盘均作为第一逻辑

磁盘。

[0044] 可选地,磁盘出现故障,所述第一确定模块,包括:

[0045] 第四确定单元,用于确定出现故障的磁盘切分出的第四逻辑盘;

[0046] 第五确定单元,用于将所述第四确定单元确定的所有第四逻辑盘均作为第一逻辑磁盘。

[0047] 可选地,所述装置,还包括:

[0048] 写入模块,用于将修复后的第一小文件写入第五逻辑盘,所述第五逻辑盘与所述第一逻辑盘的盘号相同,所述第五逻辑盘的物理位置与所述第一逻辑盘的物理位置不同;

[0049] 映射模块,用于将所述第一逻辑盘映射至所述第五逻辑盘。

[0050] 有益效果如下:

[0051] 确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;批量读取第一逻辑盘组上所有第二小文件;根据第二小文件修复第一小文件,使第一逻辑盘组中不同逻辑盘的第二小文件被批量读取、根据第二小文件修复第一小文件,避免逐个读取修复小文件带来的频繁输入操作,从而提高小文件的修复速度,提升修复性能。

附图说明

[0052] 下面将参照附图描述本发明的具体实施例,其中:

[0053] 图1示出了本发明一种实施例中提供的一种分布式文件系统中小文件的修复方法的流程示意图;

[0054] 图2示出了本发明另一种实施例中提供的另一种分布式文件系统中小文件的修复方法的流程示意图;

[0055] 图3示出了本发明另一种实施例中提供的第一种分布式文件系统中小文件的修复装置的结构示意图;

[0056] 图4示出了本发明另一种实施例中提供的第二种分布式文件系统中小文件的修复装置的结构示意图;

[0057] 图5示出了本发明另一种实施例中提供的一种第一确定模块的结构示意图;

[0058] 图6示出了本发明另一种实施例中提供的另一种第一确定模块的结构示意图;

[0059] 图7示出了本发明另一种实施例中提供的第三种分布式文件系统中小文件的修复装置的结构示意图。

具体实施方式

[0060] 为了使本发明的技术方案及优点更加清楚明白,以下结合附图对本发明的示例性实施例进行进一步详细的说明,显然,所描述的实施例仅是本发明的一部分实施例,而不是所有实施例的穷举。并且在不冲突的情况下,本说明中的实施例及实施例中的特征可以互相结合。

[0061] 在对分布式文件系统中小文件进行修复时,对每个小文件,进行读取、修复。当修复的小文件个数较多时,目前的方法会造成磁盘上大量小块随机读取,由于分布式文件系统采用的存储介质在进行小块随机读取时性能较低,会导致修复性能偏低。本发明提出了

一种分布式文件系统中小文件的修复方法,该方法应用于一种终端,该终端上运行有分布式文件系统中小文件的修复装置,该分布式文件系统中小文件的修复装置如下述图3至图7任一图所示的实施例所述的分布式文件系统中小文件的修复装置。该分布式文件系统中小文件的修复装置可以确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;批量读取第一逻辑盘组上所有第二小文件;根据第二小文件修复第一小文件,使第一逻辑盘组中不同逻辑盘的第二小文件被批量读取、根据第二小文件修复第一小文件,避免逐个读取修复小文件带来的频繁输入操作,从而提高小文件的修复速度,提升修复性能。

[0062] 结合上述实施环境,参见图1所示的实施例,本实施例提供了一种双副本并行数据库的查询方法,本实施例提供的方法流程具体如下:

[0063] 101:确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;

[0064] 102:确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;

[0065] 可选地,第二小文件为第一小文件和第一小文件的备份小文件,或第一小文件和与第一小文件为同一文件的部分数据形成的小文件,或第一小文件、第一小文件的备份小文件和与第一小文件为同一文件的部分数据形成的小文件。

[0066] 103:批量读取第一逻辑盘组上所有第二小文件;

[0067] 104:根据第二小文件修复第一小文件。

[0068] 可选地,确定出现故障的第一逻辑盘之前,还包括:

[0069] 将磁盘切分为至少一个第二逻辑盘;

[0070] 将存储第三小文件的第二逻辑盘组成第二逻辑盘组;

[0071] 第二逻辑盘组中存储的第三小文件互为相关。

[0072] 可选地,第二逻辑盘组中存储的第三小文件互为相关,包括:第二逻辑盘组中存储的所有第三小文件互为备份,或第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。

[0073] 可选地,若组成第二逻辑盘组的第二逻辑盘为多个,则组成第二逻辑盘组的第二逻辑盘位于不同节点上。

[0074] 可选地,节点出现故障,确定出现故障的第一逻辑盘,包括:

[0075] 确定出现故障的节点上的各磁盘;

[0076] 确定各磁盘切分出的第三逻辑盘;

[0077] 将所有第三逻辑盘均作为第一逻辑磁盘。

[0078] 可选地,磁盘出现故障,确定出现故障的第一逻辑盘,包括:

[0079] 确定出现故障的磁盘切分出的第四逻辑盘;

[0080] 将所有第四逻辑盘均作为第一逻辑磁盘。

[0081] 可选地,根据第二小文件修复第一小文件之后,还包括:

[0082] 将修复后的第一小文件写入第五逻辑盘,第五逻辑盘与第一逻辑盘的盘号相同,第五逻辑盘的物理位置与第一逻辑盘的物理位置不同;

[0083] 将第一逻辑盘映射至第五逻辑盘。

[0084] 有益效果：

[0085] 确定出现故障的第一逻辑盘，第一逻辑盘中包含待修复的第一小文件；确定包含第一逻辑盘的第一逻辑盘组，第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件；批量读取第一逻辑盘组上所有第二小文件；根据第二小文件修复第一小文件，使第一逻辑盘组中不同逻辑盘的第二小文件被批量读取、根据第二小文件修复第一小文件，避免逐个读取修复小文件带来的频繁输入操作，从而提高小文件的修复速度，提升修复性能。

[0086] 结合上述实施环境，参见图2所示的实施例，本实施例提供了一种分布式文件系统中小文件的修复方法，本实施例提供的方法流程具体如下：

[0087] 201：将磁盘切分为至少一个第二逻辑盘；

[0088] 本实施例不对将磁盘切分为至少一个第二逻辑盘的具体切分方法进行限定，包括但不限于，根据磁盘空间与小文件之间的关系将磁盘切分为至少一个第二逻辑盘。

[0089] 本实施例也不对第二逻辑盘的具体数量进行限定。

[0090] 202：将存储第三小文件的第二逻辑盘组成第二逻辑盘组；

[0091] 其中，第二逻辑盘组中存储的第三小文件互为相关。

[0092] 分布式文件系统进行文件存储时，会将一个文件分布存储成多个小文件，也会将小文件做备份。因此，对于第二逻辑盘组中存储的第三小文件互为相关的具体关系，本实施例不进行限定，包括但不限于：第二逻辑盘组中存储的所有第三小文件互为备份，或第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件，或第二逻辑盘组中存储的第三小文件中一部分互为备份，另一部分为同一文件的部分数据形成的小文件。

[0093] 优选的，若组成第二逻辑盘组的第二逻辑盘为多个，则组成第二逻辑盘组的第二逻辑盘位于不同节点上。

[0094] 一个逻辑盘可以存储多个小文件，但是存储的多个小文件均位于同一逻辑盘组。因此。经过步骤201和步骤202，逻辑盘组、逻辑盘和小文件形成如下关系：一个逻辑盘组包含一个或多个逻辑盘，一个逻辑盘存储一个或多个小文件。一个文件的所有小文件位于一个逻辑盘组，一个逻辑盘属于一个逻辑盘组。

[0095] 203：确定出现故障的第一逻辑盘；

[0096] 其中，第一逻辑盘中包含待修复的第一小文件。

[0097] 根据出现故障的位置不同，第一逻辑盘确定方法也不同，具体的，

[0098] ●若节点出现故障，则确定出现故障的第一逻辑盘的方法为：

[0099] 1) 确定出现故障的节点上的各磁盘；

[0100] 2) 确定各磁盘切分出的第三逻辑盘；

[0101] 3) 将所有第三逻辑盘均作为第一逻辑磁盘。

[0102] ●若磁盘出现故障，则确定出现故障的第一逻辑盘的方法为：

[0103] 1) 确定出现故障的磁盘切分出的第四逻辑盘；

[0104] 2) 将所有第四逻辑盘均作为第一逻辑磁盘。

[0105] 需要说明的是，步骤203中确定的第一逻辑磁盘可能为多个。对于多个第一逻辑磁盘情况，可以1) 选择一个第一逻辑磁盘执行下述步骤204至步骤208，完成对第一逻辑磁盘中的小文件的修复，2) 重复执行1)，直至所有第一逻辑磁盘中的小文件均被修复。

[0106] 204:确定包含第一逻辑盘的第一逻辑盘组;

[0107] 其中,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件。第一逻辑盘组包含至少一个逻辑盘,若第一逻辑盘组包含多个逻辑盘,则包含的逻辑盘位于不同节点上。第二小文件中包括第一小文件,即部分第二小文件与第一小文件相同。

[0108] 本步骤中的第一逻辑盘组为步骤202中组成的第二逻辑盘组中的一个。

[0109] 另外,由于第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件,且第一逻辑盘组包含逻辑盘中包括第一逻辑磁盘,因此,第二小文件中包括第一小文件。第一小文件与第二小文件之间的关系为:第二小文件为第一小文件和第一小文件的备份小文件,或第一小文件和与第一小文件为同一文件的部分数据形成的小文件,或第一小文件、第一小文件的备份小文件和与第一小文件为同一文件的部分数据形成的小文件。

[0110] 205:批量读取第一逻辑盘组上所有第二小文件;

[0111] 具体的,批量读取逻辑盘组上所有小文件至内存;

[0112] 为了避免逐个小文件处理带来的频繁输入操作,本步骤将逻辑盘组上所有小文件形成大块,进而批量读取逻辑盘组上所有小文件至内存,以提高小文件的修复速度,提升修复性能。

[0113] 另外,磁盘容量相对于内存较大,内存无法存储磁盘中所有文件,本实施例提供的方法中,逻辑盘组包括的各个逻辑盘均是所属磁盘的一部分,逻辑盘组上所有小文件对应的数据量小于磁盘容量,因此,步骤201的合理切分,可以保证内存能够存储逻辑盘组上所有小文件。

[0114] 此外,步骤202中,将同一文件的所有小文件存储于一个逻辑盘组,本步骤批量读取逻辑盘组上所有小文件,可以保证与待修复第一小文件相关的所有小文件全部读取至内存,保证后续修复时使用。

[0115] 206:根据第二小文件修复第一小文件;

[0116] 具体的,在内存中根据第二小文件修复第一小文件。

[0117] 本实施例不对在内存中修复读取到的小文件的具体方式进行限定,例如:在内存中根据第二小文件批量修复第一小文件。

[0118] 具体的修复方法可以按照现有修复方法进行,本实施例不再描述。

[0119] 207:将修复后的第一小文件写入第五逻辑盘;

[0120] 其中,第五逻辑盘与第一逻辑盘的盘号相同,第五逻辑盘的物理位置与第一逻辑盘的物理位置不同。

[0121] 本实施例不对将修复后的小文件写入第五逻辑盘的具体方式进行限定,例如:将修复后的第一小文件批量写入第五逻辑盘。

[0122] 具体的写入方法可以按照现有写入方法进行,本实施例不再描述。

[0123] 208:将第一逻辑盘映射至第五逻辑盘。

[0124] 具体的映射方法可以按照现有映射方法进行,本实施例不再描述。

[0125] 步骤204至步骤208循环执行,即对于每一个第一逻辑磁盘,均执行一次步骤204至步骤208,直至所有第一逻辑磁盘中的小文件均被修复。

[0126] 另外,本实施例中逻辑盘或逻辑盘组前的第一、第二、第三、第四、第五等仅为序号,用于区分不同阶段的逻辑盘或逻辑盘组,无其他区别。第一逻辑盘与第五逻辑盘可以相

同也可以不同,第一逻辑盘组与第二逻辑盘组可以相同也可以不同,同样,第一、第二、第三、第四、第五逻辑盘之间均可以相同也可以不同,本实施例不作任何限定。

[0127] 有益效果:

[0128] 确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;批量读取第一逻辑盘组上所有第二小文件;根据第二小文件修复第一小文件,使第一逻辑盘组中不同逻辑盘的第二小文件被批量读取、根据第二小文件修复第一小文件,避免逐个读取修复小文件带来的频繁输入操作,从而提高小文件的修复速度,提升修复性能。

[0129] 基于同一发明构思,参见图3所示的实施例,本实施例提供了一种分布式文件系统中小文件的修复装置,由于这些装置解决问题的原理与一种分布式文件系统中小文件的修复方法相似,因此这些装置的实施可以参见方法的实施,重复之处不再赘述。

[0130] 参见图3,该装置,包括:

[0131] 第一确定模块301,用于确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;

[0132] 第二确定模块302,用于确定包含第一确定模块301确定的第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;

[0133] 读取模块303,用于批量读取第二确定模块302确定的第一逻辑盘组上所有第二小文件;

[0134] 修复模块304,用于根据读取模块303读取的第二小文件修复第一小文件。

[0135] 其中,第二小文件为第一小文件和第一小文件的备份小文件,或第一小文件和与第一小文件为同一文件的部分数据形成的小文件,或第一小文件、第一小文件的备份小文件和与第一小文件为同一文件的部分数据形成的小文件。

[0136] 参见图4,该装置,还包括:

[0137] 切分模块305,用于将磁盘切分为至少一个第二逻辑盘;

[0138] 组合模块306,用于将存储第三小文件的第二逻辑盘组成第二逻辑盘组;

[0139] 其中,第二逻辑盘组中存储的第三小文件互为相关。

[0140] 其中,第二逻辑盘组中存储的第三小文件互为相关,包括:第二逻辑盘组中存储的所有第三小文件互为备份,或第二逻辑盘组中存储的所有第三小文件为同一文件的部分数据形成的小文件,或第二逻辑盘组中存储的第三小文件中一部分互为备份,另一部分为同一文件的部分数据形成的小文件。

[0141] 其中,若组成第二逻辑盘组的第二逻辑盘为多个,则组成第二逻辑盘组的第二逻辑盘位于不同节点上。

[0142] 参见图5,节点出现故障,第一确定模块301,包括:

[0143] 第一确定单元3011,用于确定出现故障的节点上的各磁盘;

[0144] 第二确定单元3012,用于确定第一确定单元3011确定的各磁盘切分出的第三逻辑盘;

[0145] 第三确定单元3013,用于将第二确定单元3012确定的所有第三逻辑盘均作为第一逻辑磁盘。

[0146] 参见图6,磁盘出现故障,第一确定模块301,包括:

[0147] 第四确定单元3014,用于确定出现故障的磁盘切分出的第四逻辑盘;

[0148] 第五确定单元3015,用于将第四确定单元3014确定的所有第四逻辑盘均作为第一逻辑磁盘。

[0149] 参见图7,该装置,还包括:

[0150] 写入模块307,用于将修复后的第一小文件写入第五逻辑盘,第五逻辑盘与第一逻辑盘的盘号相同,第五逻辑盘的物理位置与第一逻辑盘的物理位置不同;

[0151] 映射模块308,用于将第一逻辑盘映射至第五逻辑盘。

[0152] 有益效果如下:

[0153] 确定出现故障的第一逻辑盘,第一逻辑盘中包含待修复的第一小文件;确定包含第一逻辑盘的第一逻辑盘组,第一逻辑盘组包含的逻辑盘中存储与第一小文件相关的第二小文件;批量读取第一逻辑盘组上所有第二小文件;根据第二小文件修复第一小文件,使第一逻辑盘组中不同逻辑盘的第二小文件被批量读取、根据第二小文件修复第一小文件,避免逐个读取修复小文件带来的频繁输入操作,从而提高小文件的修复速度,提升修复性能。

[0154] 上述实施例中,均可以采用现有的功能元器件模块来实施。例如,处理模块可以采用现有的数据处理元器件,至少,现有定位技术中采用的定位服务器上便具备实现该功能元器件;至于接收模块,则是任意一个具备信号传输功能的设备都具备的元器件;同时,处理模块进行的A、n参数计算、强度调整等采用的都是现有的技术手段,本领域技术人员经过相应的设计开发即可实现。

[0155] 为了描述的方便,以上所述装置的各部分以功能分为各种模块或单元分别描述。当然,在实施本发明时可以把各模块或单元的功能在同一个或多个软件或硬件中实现。

[0156] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0157] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0158] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0159] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0160] 尽管已描述了本发明的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例作出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

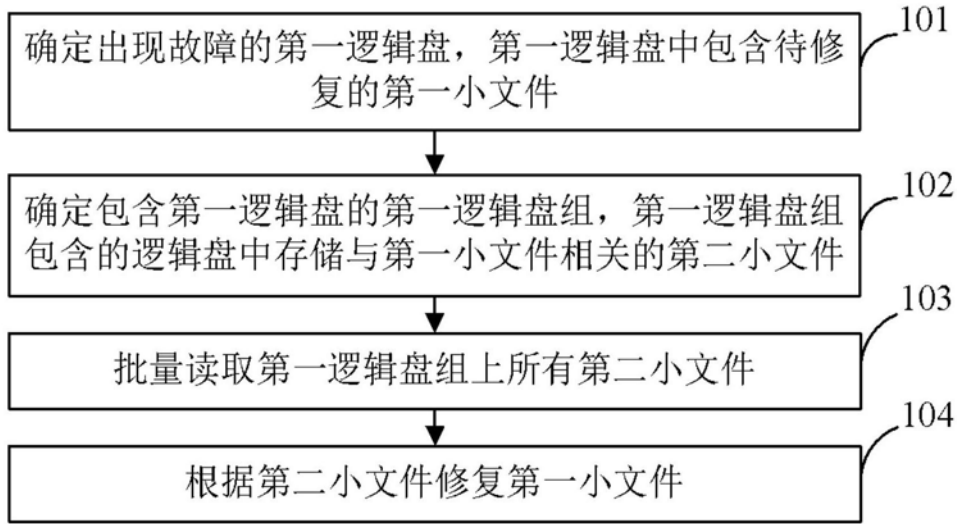


图1

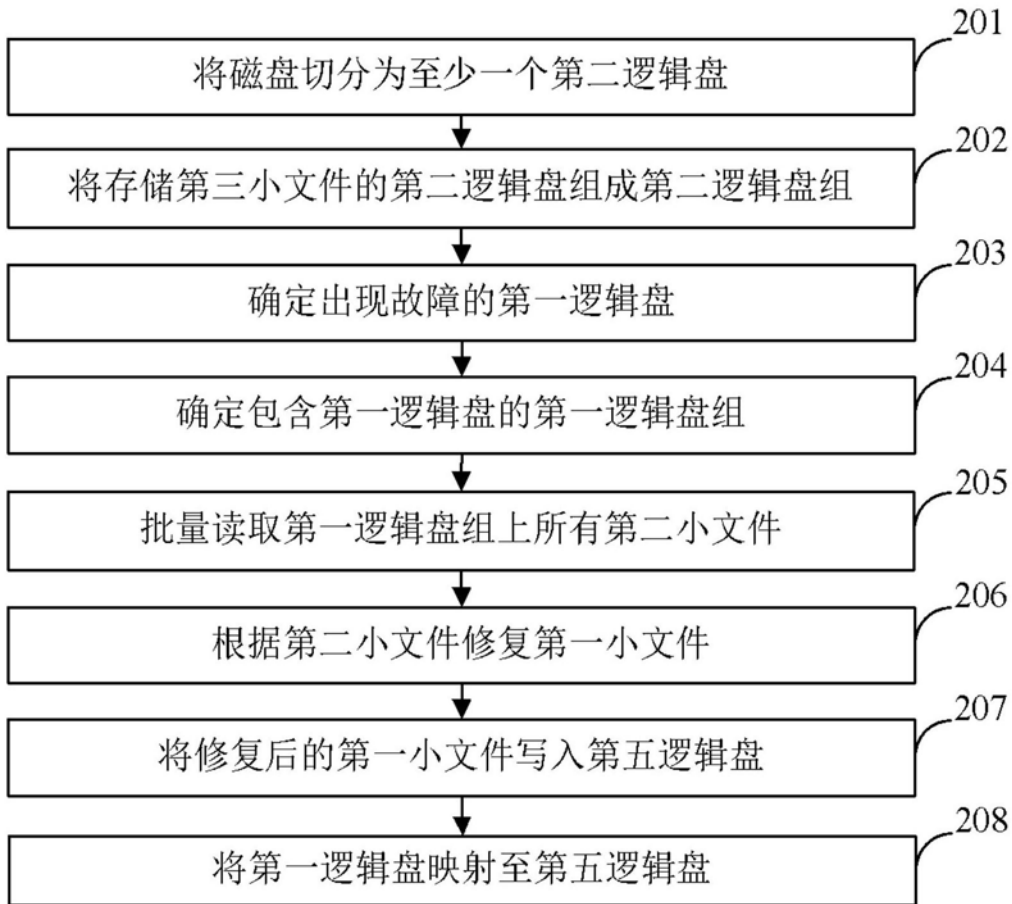


图2

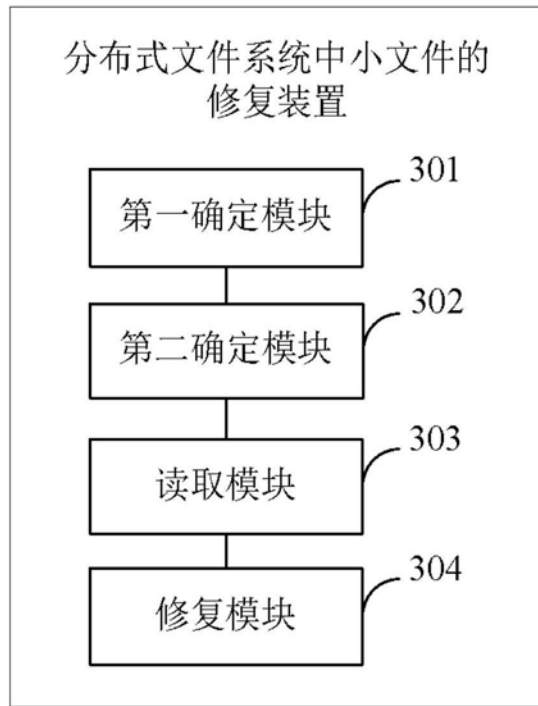


图3

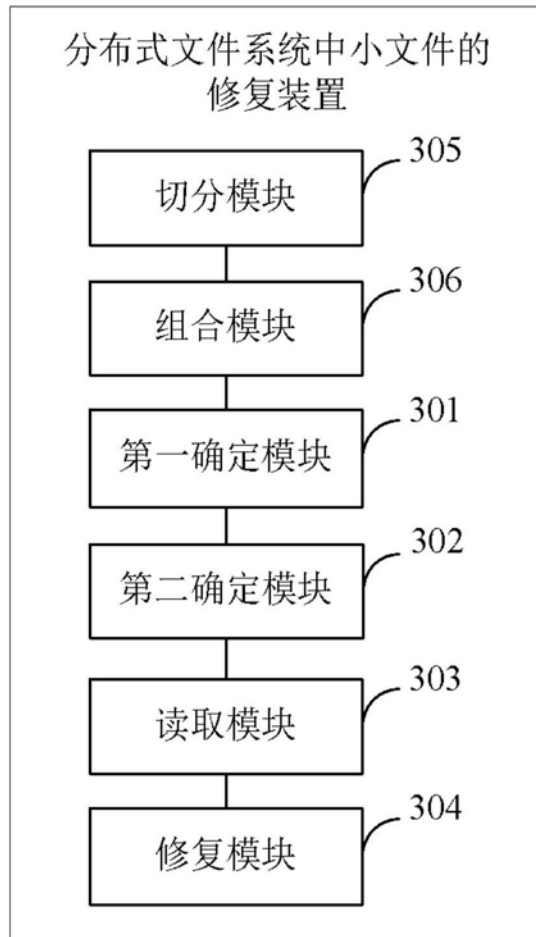


图4

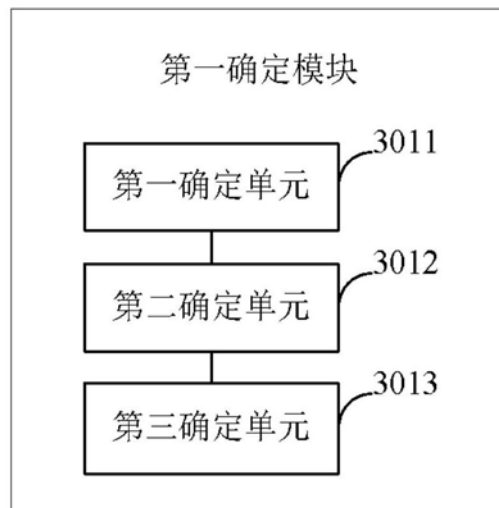


图5

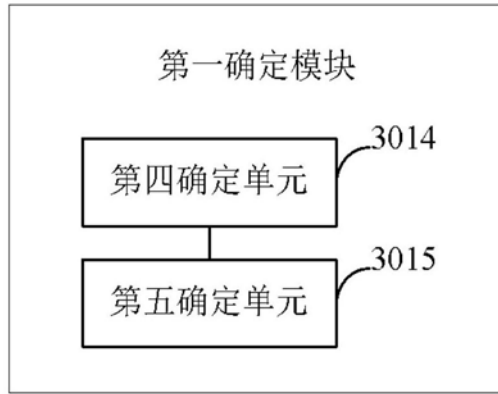


图6

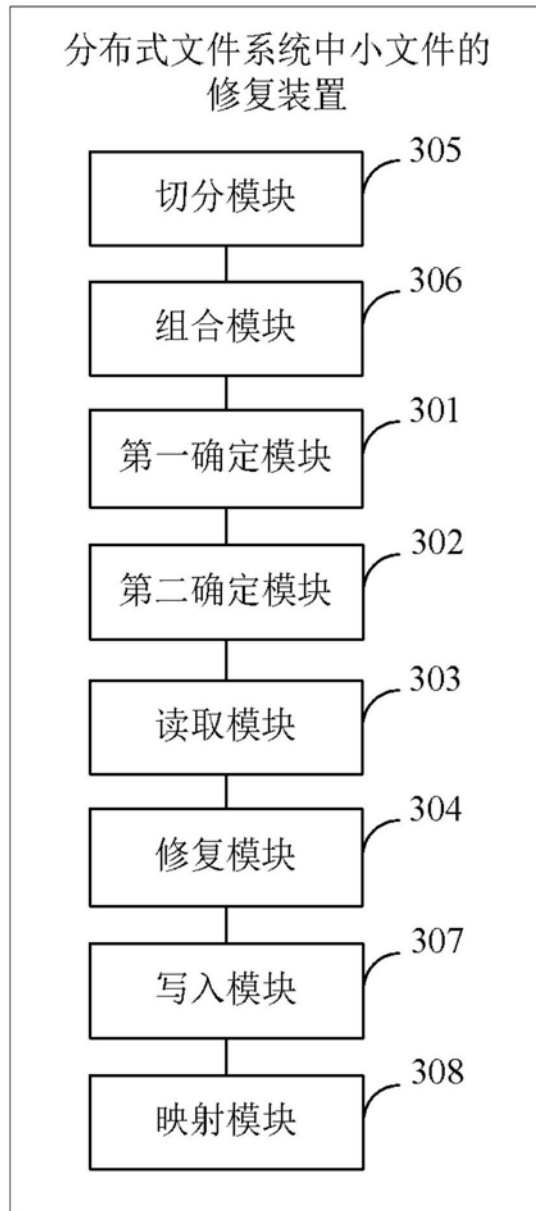


图7