



(12)发明专利申请

(10)申请公布号 CN 111753141 A

(43)申请公布日 2020.10.09

(21)申请号 201910240947.2

(22)申请日 2019.03.26

(71)申请人 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 田文罡

(74)专利代理机构 广州三环专利商标代理有限公司 44202

代理人 郝传鑫 熊永强

(51) Int. Cl.

G06F 16/901(2019.01)

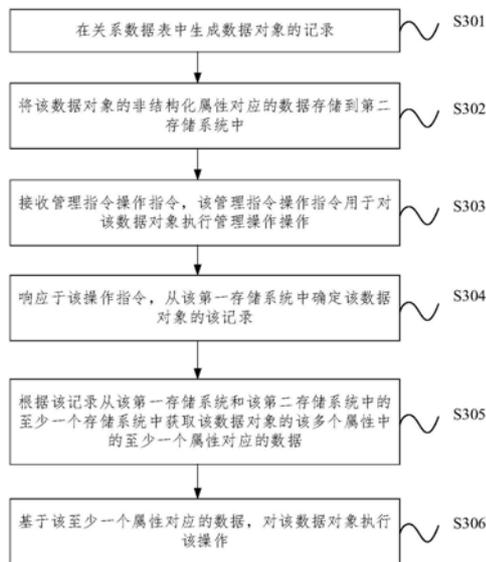
权利要求书7页 说明书20页 附图4页

(54)发明名称

一种数据管理方法及相关设备

(57)摘要

本申请实施例公开了一种数据管理方法及相关设备,用于实现同一个数据对象在跨多个类型的存储系统存储的数据一致性。其中,数据对象的结构化属性对应的数据存储在第一存储系统,比如关系型数据库,数据对象的非结构化属性对应的数据存储在其他类型的存储系统,比如KV系统、文件系统等。通过存储在数据库中的关系数据表来记录数据对象的结构化数据和非结构化数据的关联关系。在对数据对象进行操作时,先访问数据库中的记录,从数据库的记录中获取非结构化属性的键值和路径,然后再通过其他类型的存储系统的接口访问非结构化属性对应的数据。这样,可以通过数据库的事务一致性和规定的访问顺序实现多种存储类型系统之间的数据一致性。



1. 一种数据管理方法,其特征在于,所述方法包括:

在关系数据表中生成数据对象的记录,所述数据对象具有多个属性,所述多个属性包括结构化属性和非结构化属性,所述记录包含所述结构化属性对应的数据,以及所述数据对象的结构化属性和非结构化属性的关联关系,所述关系数据表存储于第一存储系统中;

将所述数据对象的非结构化属性对应的数据存储到第二存储系统中;

接收操作指令,所述操作指令用于对所述数据对象执行操作;

响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录;

根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据;

基于所述至少一个属性对应的数据,对所述数据对象执行所述操作。

2. 根据权利要求1所述的方法,其特征在于,所述在关系数据表中生成数据对象的记录,包括:

接收插入指令或更新指令,所述插入指令用于插入所述数据对象,所述更新指令用于更新所述数据对象;所述插入指令和所述更新指令均包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据;

根据所述对象类型确定所述数据对象对应的关系数据表;

根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;

提交所述插入指令或更新指令对应的事务;

其中,所述插入指令或更新指令对应的事务在所述数据对象的非结构化属性对应的数据存储到第二存储系统中之后提交。

3. 根据权利要求2所述的方法,其特征在于,接收的指令为所述插入指令,所述数据对象的非结构化属性包括键值KV属性,所述第二存储系统为KV存储系统;

所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:

根据第一版本标识和所述KV属性对应的数据中的第一键值生成第二键值;

生成所述数据对象的记录,其中,所述记录中的KV属性对应的数据包括所述第二键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

4. 根据权利要求2所述的方法,其特征在于,接收的指令为所述插入指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;

所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:

根据所述第一版本标识和所述文件属性对应的数据中的第一路径生成第二路径;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第二路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

5. 根据权利要求2所述的方法,其特征在于,接收的指令为所述更新指令,所述数据对象的非结构化属性包括KV属性,所述第二存储系统为KV存储系统;

所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据

对象对应的关系数据表中生成所述数据对象的记录,包括:

根据第二版本标识和所述KV属性对应的数据中的第一键值生成第三键值;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第三键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

6. 根据权利要求2所述的方法,其特征在于,接收的指令为所述更新指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;

所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:

根据所述第二版本标识和所述文件属性对应的数据中的第一路径生成第三路径;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第三路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

7. 根据权利要求3~6任一项所述的方法,其特征在于,所述第二存储系统中存储的非结构化属性对应的数据包括所述非结构化属性的标识和内容;所述关系数据表中存储的所述非结构化属性对应的数据包括所述非结构化属性的标识。

8. 根据权利要求1~7任一项所述的方法,其特征在于,所述操作指令包括查询指令,所述查询指令中包括查询条件;

所述从所述第一存储系统中确定所述数据对象的所述记录,包括:

从所述第一存储系统中选取满足所述查询条件的数据对象的记录;

所述根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据,包括:

根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的所述多个属性对应的数据;

所述基于所述至少一个属性对应的数据,对所述数据对象执行所述操作,包括:

根据获取的所述多个属性对应的数据返回查询结果。

9. 根据权利要求8所述的方法,其特征在于,所述记录的非结构化属性包括KV属性,所述根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据,包括:

根据键值从所述第二存储系统中读取所述键值对应的KV数据,并去除所述键值中的版本标识;所述键值为所述记录中的KV属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;

其中,所述数据对象的KV属性对应的数据包括去除版本标识后的键值和所述KV数据;所述数据对象的结构化属性对应的数据包括所述记录中结构化属性对应的数据。

10. 根据权利要求8所述的方法,其特征在于,所述记录的非结构化属性包括文件属性,所述根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据,包括:

根据路径从所述第二存储系统中读取所述路径对应的文件数据,并去除所述路径中的版本标识;所述路径为所述记录中的文件属性对应的数据,所述版本标识包括第一版本标

识和第二版本标识；

其中，所述数据对象的文件属性对应的数据包括，去除版本标识后的路径和所述文件数据；所述数据对象的结构化属性对应的数据包括，所述记录中结构化属性对应的数据。

11. 根据权利要求1~7任一项所述的方法，其特征在于，所述操作指令包括删除指令，所述删除指令包括所述数据对象的对象类型，以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据；

其中，所述从所述第一存储系统中确定所述数据对象的所述记录，包括：

根据所述对象类型确定所述数据对象对应的关系数据表；

从所述关系数据表中确定所述数据对象的所述记录，所述记录中的结构化属性对应的数据与所述数据对象的结构化属性对应的数据相同；

所述根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的多个属性中的至少一个属性对应的数据，包括：

所述根据所述记录从所述第一存储系统中获取所述数据对象的所述多个属性对应的数据；

所述基于所述至少一个属性对应的数据，对所述数据对象执行所述操作，包括：

从所述第一存储系统中删除所述数据对象的所述多个属性对应的数据；

提交所述删除指令对应的事务。

12. 根据权利要求1~7任一项所述的方法，其特征在于，所述非结构化属性包括KV属性，所述第二存储系统为KV存储系统，所述方法还包括：

当接收到校验指令时或者当检测到满足校验条件时，遍历所述第二存储系统中的键值，所述键值为所述第二存储系统中的KV属性对应的数据；

在遍历所述键值的过程中，若在所述第一存储系统存储的关系数据表中无法查找到与第四键值相同的键值，则在所述第二存储系统中删除所述第四键值和所述第四键值对应的KV数据，所述第四键值为所述第二存储系统中多个键值中的一个。

13. 根据权利要求1~7任一项所述的方法，其特征在于，所述非结构化属性包括文件属性，所述第二存储系统为文件存储系统，所述方法还包括：

当接收到校验指令时或者当检测到满足校验条件时，遍历所述第二存储系统中的路径，所述路径为所述第二存储系统中的文件属性对应的数据；

在遍历所述路径的过程中，若在所述第一存储系统存储的关系数据表中无法查找到与第四路径相同的路径，则在所述第二存储系统中删除所述第四路径和所述第四路径对应的文件数据，所述第四路径为所述第二存储系统中多个路径中的一个。

14. 根据权利要求2~6任一项所述的方法，其特征在于，在根据所述对象类型确定所述数据对象对应的关系数据表之前，所述方法还包括：

接收针对所述数据对象所属对象类型的定义指令，所述定义指令中包含所述对象类型的定义信息，所述定义信息用于定义所述对象类型的关系数据表的结构；

根据所述定义指令在所述第一存储系统中，生成所述对象类型的关系数据表。

15. 根据权利要求2~6任一项所述的方法，其特征在于，所述根据所述对象类型确定所述数据对象对应的关系数据表，包括：

根据所述插入指令或更新指令确定所述数据对象所属的对象类型；

根据所述对象类型确定所述数据对象对应的关系数据表。

16. 一种数据管理设备,其特征在於,所述数据管理设备包括生成单元,存储单元,接收单元,确定单元,获取单元,操作单元:

所述生成单元,用于在关系数据表中生成数据对象的记录,所述数据对象具有多个属性,所述多个属性包括结构化属性和非结构化属性,所述记录包含所述结构化属性对应的数据,以及所述数据对象的结构化属性和非结构化属性的关联关系,所述关系数据表存储于第一存储系统中;

所述存储单元,用于将所述数据对象的非结构化属性对应的数据存储到第二存储系统中;

所述接收单元,用于接收操作指令,所述操作指令用于对所述数据对象执行操作;

所述确定单元,用于响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录;

所述获取单元,用于根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据;

所述操作单元,用于基于所述至少一个属性对应的数据,对所述数据对象执行所述操作。

17. 根据权利要求16所述的数据管理设备,其特征在於,所述生成单元具体用于:

接收插入指令或更新指令,所述插入指令用于插入所述数据对象,所述更新指令用于更新所述数据对象;所述插入指令和所述更新指令均包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据;

根据所述对象类型确定所述数据对象对应的关系数据表;

根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;

提交所述插入指令或更新指令对应的事务;

其中,所述插入指令或更新指令对应的事务在所述数据对象的非结构化属性对应的数据存储到第二存储系统中之后提交。

18. 根据权利要求17所述的数据管理设备,其特征在於,接收的指令为所述插入指令,所述数据对象的非结构化属性包括键值KV属性,所述第二存储系统为KV存储系统;

所述生成单元具体用于:

根据第一版本标识和所述KV属性对应的数据中的第一键值生成第二键值;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第二键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

19. 根据权利要求17所述的数据管理设备,其特征在於,接收的指令为所述插入指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;

所述生成单元具体用于:

根据所述第一版本标识和所述文件属性对应的数据中的第一路径生成第二路径;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第二路径,所述记录中的结构化属性对应的数据包括所述数

据对象的结构化属性对应的数据。

20. 根据权利要求17所述的数据管理设备,其特征在于,接收的指令为所述更新指令,所述数据对象的非结构化属性包括KV属性,所述第二存储系统为KV存储系统;

所述生成单元具体用于:

根据第二版本标识和所述KV属性对应的数据中的第一键值生成第三键值;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第三键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

21. 根据权利要求17所述的数据管理设备,其特征在于,接收的指令为所述更新指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;

所述生成单元具体用于:

根据所述第二版本标识和所述文件属性对应的数据中的第一路径生成第三路径;

在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第三路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

22. 根据权利要求18~21任一项所述的数据管理设备,其特征在于,所述第二存储系统中存储的非结构化属性对应的数据包括所述非结构化属性的标识和内容;所述关系数据表中存储的非结构化属性对应的数据包括所述非结构化属性的标识。

23. 根据权利要求16~22任一项所述的数据管理设备,其特征在于,所述操作指令包括查询指令,所述查询指令中包括查询条件;

所述确定单元具体用于:

响应于所述操作指令,从所述第一存储系统中选取满足所述查询条件的数据对象的记录;

所述获取单元具体用于:

根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的所述多个属性对应的数据;

所述操作单元具体用于:根据获取的所述多个属性对应的数据返回查询结果。

24. 根据权利要求23所述的数据管理设备,其特征在于,所述记录的非结构化属性包括KV属性,所述第二存储系统为KV存储系统,所述获取单元具体用于:

根据键值从所述第二存储系统中读取所述键值对应的KV数据,并去除所述键值中的版本标识;所述键值为所述记录中的KV属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;

其中,所述数据对象的KV属性对应的数据包括去除版本标识后的键值和所述KV数据;所述数据对象的结构化属性对应的数据包括所述记录中结构化属性对应的数据。

25. 根据权利要求23所述的数据管理设备,其特征在于,所述记录的非结构化属性包括文件属性,所述第二存储系统为文件存储系统,所述获取单元具体用于:

根据路径从所述第二存储系统中读取所述路径对应的文件数据,并去除所述路径中的版本标识;所述路径为所述记录中的文件属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;

其中,所述数据对象的文件属性对应的数据包括去除版本标识后的路径和所述文件数据;所述数据对象的结构化属性对应的数据包括所述记录中结构化属性对应的数据。

26. 根据权利要求16~22任一项所述的数据管理设备,其特征在于,所述操作指令包括删除指令,所述删除指令包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据;

其中,所述确定单元具体用于:

根据所述对象类型确定所述数据对象对应的关系数据表;

从所述关系数据表中确定所述数据对象的所述记录,所述记录中的结构化属性对应的数据与所述数据对象的结构化属性对应的数据相同;

所述获取单元具体用于:

所述根据所述记录从所述第一存储系统中获取所述数据对象的所述多个属性对应的数据;

所述操作单元具体用于:

从所述第一存储系统中删除所述数据对象的所述多个属性对应的数据;

提交所述删除指令对应的事务。

27. 根据权利要求16~22任一项所述的数据管理设备,其特征在于,所述非结构化属性包括KV属性,所述第二存储系统为KV存储系统,所述数据管理设备还包括校验单元:

所述校验单元,用于当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的键值,所述键值为所述第二存储系统中的KV属性对应的数据;

在遍历所述键值的过程中,若在所述第一存储系统存储的关系数据表中无法查找到与第四键值相同的键值,则在所述第二存储系统中删除所述第四键值和所述第四键值对应的KV数据,所述第四键值为所述第二存储系统中多个键值中的一个。

28. 根据权利要求16~22任一项所述的数据管理设备,其特征在于,所述非结构化属性包括文件属性,所述第二存储系统为文件存储系统,所述数据管理设备还包括校验单元:

所述校验单元,用于当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的路径,所述路径为所述第二存储系统中的文件属性对应的数据;

在遍历所述路径的过程中,若在所述第一存储系统存储的关系数据表中无法查找到与第四路径相同的路径,则在所述第二存储系统中删除所述第四路径和所述第四路径对应的文件数据,所述第四路径为所述第二存储系统中多个路径中的一个。

29. 根据权利要求17~22任一项所述的数据管理设备,其特征在于,所述生成单元还用于:

接收针对所述数据对象所属对象类型的定义指令,所述定义指令中包含所述对象类型的定义信息,所述定义信息用于定义所述对象类型的关系数据表的结构;

根据所述定义指令在所述第一存储系统中,生成所述对象类型的关系数据表。

30. 根据权利要求17~22任一项所述的数据管理设备,其特征在于,所述生成单元具体用于:

根据所述插入指令或更新指令确定所述数据对象所属的对象类型;

根据所述对象类型确定所述数据对象对应的关系数据表。

31. 一种数据管理设备,其特征在于,包括处理器和存储器,其中,所述存储器用于存储

程序指令,所述处理器用于根据所述程序指令执行如权利要求1~15任一项所述的方法。

32.一种计算机可读存储介质,其特征在于,所述计算机存储介质存储有程序指令,所述程序指令当被计算机执行时使所述计算机执行如权利要求1~15任一项所述的方法。

## 一种数据管理方法及相关设备

### 技术领域

[0001] 本申请涉及计算机技术领域,尤其涉及一种数据管理方法及相关设备。

### 背景技术

[0002] 在计算机系统中,数据一般分为结构化数据和非结构化数据,其中,结构化数据是由二维表结构来逻辑表达和实现的数据,遵循数据格式与长度规范,例如销售信息,财产信息等。非结构化数据是数据结构不规则或不完整,没有预定义的数据,例如文档、图片,音频和视频等。

[0003] 对数据的存储而言,结构化数据一般存储在关系型数据库中,大的非结构数据一般存储在文件存储系统中,小的非结构数据一般存储在键值(key value,KV)系统中。在实际应用中,一个数据对象可能同时包含结构化数据和非结构化数据。举例而言,数据对象为一张图片的信息,这张图片的名称、大小、拍摄时间和拍摄地的经纬度信息等图片的属性信息为结构化数据,这张图片本身为非结构化数据,基于这张图片生成的缩略图为非结构化数据。那么,图片的名称、大小、拍摄时间和拍摄地的经纬度信息等数据将存储在关系型数据库中,图片本身将存储在文件存储系统中,基于图片生成的缩略图将存储在KV存储系统中。可以看出,由于一个数据对象可能同时包含结构化数据和非结构化数据,因此一个数据对象可能存在跨多个数据系统进行存储的情况。

[0004] 现有技术中,在一个数据对象跨多个数据系统进行存储的情况下,用户一般可以对多个数据系统中的数据分别进行操作,可能会存在同一数据对象在各个数据系统中的不一致的问题。例如,用户从文件存储系统中删除了一张图片的文件,用户仍旧可以通过数据库获取到图片的属性信息,但由于该图片的文件已被删除,该图片无法正常显示。如何让数据在跨多个数据系统存储的情况下保持数据一致性是本领域技术人员亟待解决的问题。

### 发明内容

[0005] 本申请实施例提供一种数据管理方法及相关设备,用于实现数据对象在跨多个数据系统存储的情况下保持数据一致性。

[0006] 第一方面,本申请实施例提供了一种数据管理方法,该方法包括:在关系数据表中生成数据对象的记录,所述数据对象具有多个属性,所述多个属性包括结构化属性和非结构化属性,所述记录指示了所述数据对象的结构化属性和非结构化属性的关联关系,所述关系数据表存储于第一存储系统中;将所述数据对象的非结构化属性对应的数据存储到第二存储系统中;接收操作指令,所述操作指令用于对所述数据对象执行操作;响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录;根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据;基于所述至少一个属性对应的数据,对所述数据对象执行所述操作。由于该数据对象跨多个数据系统存储的多个属性对应的数据均通过该记录获取,可以让该数据对象在跨多个数据系统存储的情况下保持数据一致性。

[0007] 结合第一方面,在一种可能的实现方式中,所述在关系数据表中生成数据对象的记录,包括:接收插入指令或更新指令,所述插入指令用于插入所述数据对象,所述更新指令用于更新所述数据对象;所述插入指令和所述更新指令均包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据;根据所述对象类型确定所述数据对象对应的关系数据表;根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;提交所述插入指令或更新指令对应的事务;其中,所述插入指令或更新指令对应的事务在所述数据对象的非结构化属性对应的数据存储到第二存储系统中之后提交。

[0008] 结合第一方面,在一种可能的实现方式中,接收的指令为所述插入指令,所述数据对象的非结构化属性包括键值KV属性,所述第二存储系统为KV存储系统;所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:根据第一版本标识和所述KV属性对应的数据中的第一键值生成第二键值;根据记录中的结构化属性对应的数据,以及记录中的KV属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第二键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

[0009] 结合第一方面,在一种可能的实现方式中,接收的指令为所述插入指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:根据所述第一版本标识和所述文件属性对应的数据中的第一路径生成第二路径;根据记录中的结构化属性对应的数据,以及记录中的文件属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第二路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

[0010] 结合第一方面,在一种可能的实现方式中,接收的指令为所述更新指令,所述数据对象的非结构化属性包括KV属性,所述第二存储系统为KV存储系统;所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:根据第二版本标识和所述KV属性对应的数据中的第一键值生成第三键值;根据记录中的结构化属性对应的数据,以及记录中的KV属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第三键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

[0011] 结合第一方面,在一种可能的实现方式中,接收的指令为所述更新指令,所述数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统;所述根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录,包括:根据所述第二版本标识和所述文件属性对应的数据中的第一路径生成第三路径;根据记录中的结构化属性对应的数据,以及记录中的文件属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第三路径,所述记录中的结构化属性对应的数据包

括所述数据对象的结构化属性对应的数据。

[0012] 结合第一方面,在一种可能的实现方式中,所述第二存储系统中存储的所述数据对象的非结构化属性对应的数据,包括所述非结构化属性对应的标识和内容;所述第一存储系统中存储的所述记录中非结构化属性对应的数据包括非结构化属性对应的标识。

[0013] 结合第一方面,在一种可能的实现方式中,所述操作指令包括查询指令,所述查询指令中包括查询条件;所述响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录,包括:响应于所述操作指令,从所述第一存储系统中选取满足所述查询条件的数据对象的记录;所述根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据,包括:根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的所述多个属性对应的数据;所述基于所述至少一个属性对应的数据,对所述数据对象执行所述操作,包括:根据所述数据对象的多个属性中每个属性对应的数据和所述记录中多个属性的顺序建立所述数据对象;将所述数据对象作为查询结果返回。

[0014] 结合第一方面,在一种可能的实现方式中,所述记录的非结构化属性包括KV属性,所述第二存储系统为KV存储系统,所述根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据,包括:根据键值从所述第二存储系统中读取所述键值对应的KV数据,并去除所述键值中的版本标识;所述键值为所述记录中的KV属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;其中,所述数据对象的KV属性对应的数据包括去除版本标识后的键值和所述KV数据;所述数据对象的结构化属性对应的数据包括所述记录中结构化属性对应的数据。

[0015] 结合第一方面,在一种可能的实现方式中,所述记录的非结构化属性包括文件属性,所述第二存储系统为文件存储系统,所述根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据,包括:根据路径从所述第二存储系统中读取所述路径对应的文件数据,并去除所述路径中的版本标识;所述路径为所述记录中的文件属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;其中,所述数据对象的文件属性对应的数据包括去除版本标识后的路径和所述文件数据;所述数据对象的结构化属性对应的数据包括所述记录中结构化属性对应的数据。

[0016] 结合第一方面,在一种可能的实现方式中,所述操作指令还包括删除指令,所述删除指令包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据;其中,所述响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录,包括:根据所述对象类型确定所述数据对象对应的关系数据表;从所述关系数据表中确定所述数据对象的所述记录,所述记录中的结构化属性对应的数据与所述数据对象的结构化属性对应的数据相同;所述根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的多个属性中的至少一个属性对应的数据,包括:所述根据所述记录从所述第一存储系统中获取所述数据对象的所述多个属性对应的数据;所述基于所述至少一个属性对应的数据,对所述数据对象执行所述操作,包括:从所述第一存储系统中删除所述数据对象的所述多个属性对应的数据;提交所述删除指令对应的事务。

[0017] 结合第一方面,在一种可能的实现方式中,所述非结构化属性包括KV属性,所述第

二存储系统为KV存储系统,所述方法还包括:当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的键值,所述键值为所述第二存储系统中的KV属性对应的数据;在遍历所述键值的过程中,若在所述第一存储系统存储的关系数据表中无法查找到与第四键值相同的键值,则在所述第二存储系统中删除所述第四键值和所述第四键值对应的KV数据,所述第四键值为所述第二存储系统中多个键值中的一个。

[0018] 结合第一方面,在一种可能的实现方式中,所述非结构化属性包括文件属性,所述第二存储系统为文件存储系统,所述方法还包括:当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的路径,所述路径为所述第二存储系统中的文件属性对应的数据;在遍历所述路径的过程中,若在所述第一存储系统存储的关系数据表中无法查找到与第四路径相同的路径,则在所述第二存储系统中删除所述第四路径和所述第四路径对应的文件数据,所述第四路径为所述第二存储系统中多个路径中的一个。

[0019] 结合第一方面,在一种可能的实现方式中,在根据所述对象类型确定所述数据对象对应的关系数据表之前,所述方法还包括:接收针对所述数据对象所属对象类型的定义指令,所述定义指令中包含所述对象类型的定义信息,所述定义信息用于定义所述对象类型的关系数据表的结构;根据所述定义指令在所述第一存储系统中,生成所述对象类型的关系数据表。

[0020] 结合第一方面,在一种可能的实现方式中,所述根据所述对象类型确定所述数据对象对应的关系数据表,包括:根据所述插入指令或更新指令确定所述数据对象所属的对象类型;根据所述对象类型确定所述数据对象对应的关系数据表。

[0021] 第二方面,本申请实施例提供了一种数据管理设备,该数据管理设备包括用于执行第一方面或第一方面的各种可能的实现方式所描述的方法所对应的单元。

[0022] 上述数据管理设备可以是电子设备,也可以是电子设备中的用于实现数据管理的装置(例如,操作系统、数据库管理系统),也可以是服务器,比如数据库服务器,应用服务器等。

[0023] 上述数据管理设备置包括的单元可以是硬件电路,也可是软件,也可以是硬件电路结合软件实现。

[0024] 第三方面,本申请实施例提供了另一种数据管理设备,包括处理器和存储器,该处理器和存储器相互连接,其中,该存储器用于存储程序指令,该处理器用于调用该存储器中的程序指令来执行上述第一方面或者第一方面的任一可能的实现方式所描述的方法。

[0025] 第四方面,本申请实施例提供了一种计算机可读存储介质,该计算机存储介质存储有程序指令,该程序指令当被处理器运行时,该处理器执行上述第一方面或者第一方面的任一可能的实现方式所描述的方法。

[0026] 第五方面,本申请实施例提供了一种计算机程序,该计算机程序在处理器上运行时,该处理器执行上述第一方面或者第一方面的任一可能的实现方式所描述的方法。

[0027] 在本申请实施例中,数据管理设备可以在关系数据表中生成数据对象的记录,该记录指示了数据对象的结构化数据和非结构化数据的关联关系,该关系数据表存储在所述第一存储系统中,该数据对象的非结构化数据存储在第二存储系统中。在需要对数据进行操作时,可以从该第一存储系统中获取该数据对象的记录,根据该记录从该第一存储系统和/或该第二存储系统中获取该数据对象的多个属性中至少一个属性对应的数据;再基于该至少

一个属性对应的数据,对该数据对象执行该操作。由于该数据对象跨多个数据系统存储的多个属性对应的数据均通过关系数据表中的记录获取,可以让该数据对象在跨多个数据系统存储的情况下保持数据一致性。

### 附图说明

[0028] 为了更清楚地说明本申请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。

[0029] 图1A是本申请实施例提供了一种数据管理设备的示意图;

[0030] 图1B是本申请实施例提供的又一种数据管理设备的示意图;

[0031] 图1C是本申请实施例提供的又一种数据管理设备的示意图;

[0032] 图2是本申请实施例提供了一种数据管理系统的架构示意图;

[0033] 图3是本申请实施例提供了一种数据管理方法的流程图;

[0034] 图4是本申请实施例提供的又一种数据管理设备的示意图;

[0035] 图5是本申请实施例提供的又一种数据管理设备的示意图。

### 具体实施方式

[0036] 下面对本申请实施例中的技术方案进行更详细的描述。

[0037] 本申请实施例提供的的数据管理方法可应用于数据管理设备,该数据管理设备包括第一存储系统和第二存储系统。第一存储系统和第二存储系统为不同类型的存储系统。其中,第一存储系统可以用于存储数据对象的结构化属性及对应的结构化数据,第二存储系统用于存储数据对象的非结构化属性及对应的非结构化数据。结构化属性是用于描述或定义结构化数据的特征的属性,非结构化属性是用于描述或定义非结构化数据的特征的属性。结构化数据也称作行数据,是由二维表结构来逻辑表达和实现的数据,主要通过关系型数据库进行存储和管理;非结构化数据是数据结构不规则或不完整,没有预定义的数据模型,不方使用数据库二维逻辑表来表现的数据,比如文档、文本、图片、报表、图像、音频/视频信息等。在一个实施例中,该第一存储系统可以为数据库,比如关系型数据库。可选的,数据对象的非结构化属性可以为多个,则该第二存储系统可以为多个,举例而言,该数据对象的非结构化属性可以包括键值(key-value,KV)属性和文件(file)属性,则该数据管理设备可以包括KV存储系统(比如KV数据库)和文件存储系统(简称“文件系统”)。具体的,该数据管理设备可以包括手机、平板电脑、个人数字助理(personal digital assistant,PDA)、移动互联网设备(mobile internet device,MID)等终端设备,也可以包括数据库服务器、应用服务器等具有数据存储和处理功能的设备,本发明实施例不作限定。

[0038] 该数据管理设备可以通过自身运行的应用程序接收用户输入的针对数据对象的操作指令,并对数据对象执行该操作指令。举例而言,该应用程序可以为存储图像或者视频的相册,可以接收用户输入的对图像或者视频的操作指令;或者该应用程序可以为创建文本的文本软件,可以接收用户输入的对文本的操作指令;或者该应用程序可以为即时通讯软件,可以接收用户输入的对该软件中的办公文档、文本、图片、图像和音频/视频等数据的操作指令。

[0039] 参见图1A,是本申请实施例提供了一种数据管理设备的示意图。该数据管理设备

包括应用模块,操作模块,接口模块,存储系统模块。以下将对这些模块进行进一步的介绍。

[0040] 其中,该应用模块可以包含一个或多个应用程序,这些应用程序可以接收用户输入的针对数据对象的操作指令。举例而言,该应用程序可以包括相册、邮箱、文档处理软件等等。

[0041] 操作模块,是一个数据对象管理组件,它提供接口给应用模块,应用模块通过操作模块可以实现对数据对象进行定义、插入、修改、删除、查询、校验等操作。具体地,操作模块可以根据从应用模块接收到的指令,执行该指令指示的操作。接收到的指令不同,该操作模块执行的操作不同,以下针对几种不同的指令来进行举例说明。

[0042] 若该指令为定义指令,该操作模块可以根据该定义指令确定数据对象该对象类型的定义信息,再通过该接口模块将该定义信息存储到存储系统模块中。其中,该数据对象该对象类型的定义信息包括该数据对象的结构化属性和非结构化属性。

[0043] 若该指令为插入指令,该操作模块可以根据该插入指令确定需要存储的该数据对象的信息,再通过该接口模块将需要插入的该数据对象的信息存储到存储系统模块中。其中,操作模块会将不同类型的数据保存到不同的存储系统,具体的,结构化数据保存到数据库中,文件数据保存到文件存储系统中,键值数据保存到KV存储系统中。

[0044] 若该指令为更新指令,该操作模块可以根据该更新指令确定需要更新的该数据对象的信息,再通过该接口模块将需要更新的该数据对象的信息存储到存储系统模块中。其中,操作模块会将不同类型的数据保存到不同的存储系统,具体的,结构化数据保存到数据库中,文件数据保存到文件存储系统中,键值数据保存到KV存储系统中。

[0045] 若该指令为查询指令,该操作模块可以根据该查询指令确定查询条件,再根据该查询条件通过接口模块从存储系统模块中,选取满足查询条件的数据对象反馈给应用模块。

[0046] 若该指令为删除指令,该操作模块可以根据该删除指令确定需要删除的数据对象,再通过接口模块从存储系统模块中,删除该需要删除的数据对象。

[0047] 若该指令为校验指令,该操作模块可以根据该校验指令通过接口模块对存储系统模块中存储的数据进行校验,以清除无效的数据。

[0048] 接口模块,提供访问存储系统模块的接口,操作模块可以通过该接口模块访问存储系统模块中的数据。在一个实施例中,如图1B所示,接口模块包含第一存储系统接口子模块和第二存储系统接口子模块;存储系统模块包含第一存储系统子模块和第二存储系统子模块。操作模块可以通过第一存储系统接口子模块访问第一存储系统中的数据,该操作模块可以通过第二存储系统接口子模块访问第二存储系统中的数据。

[0049] 在一个实施例中,第一存储系统为数据库系统,比如关系型数据库,则第一存储系统接口子模块为数据库系统接口;第二存储系统为用于存储非关系型数据的存储系统,比如KV存储系统和/或文件存储系统,相应地,第二存储系统接口子模块包括KV系统接口子模块和/或文件存储系统接口子模块。数据管理设备将同一数据对象的结构化属性对应的数据存储在第一存储系统中,将该数据对象的非结构化数据对应的数据存储在第二存储系统中。进一步地,数据管理设备在第一存储系统中生成关系数据表,以建立该数据对象的结构化属性和非结构化属性的关联关系。该关系数据表包含该数据对象的记录,该记录包括了该数据对象的结构化属性的名称、非结构化属性的名称、结构化属性对应的数据内容(值)、

以及非结构化属性对应的数据的标识(非结构化属性对应的数据内容本身存储在第二存储系统)。当操作模块从应用模块接收到针对该数据对象的操作指令后,先从存储在第一存储系统的关系数据表中确定该数据对象的记录,然后可以根据该记录从关系数据表中获取该数据对象的结构化属性对应的数据,以及可以基于该记录确定非结构化属性对应的数据的标识,进而根据该标识从第二存储系统中获取该数据对象的非结构化属性对应的数据(数据内容)。最后,数据管理设备基于获取的该数据对象的结构化属性对应的数据,和/或非结构化属性对应的数据,对该数据对象执行相应的操作。数据管理设备生成数据对象的关系数据表,存储数据以及操作数据的详细过程可以参见图3相关的实施例。

[0050] 在一个实施例中,第一存储系统子模块可以是支持多版本并发控制(multi-version concurrency control,MVCC)的数据库,比如轻型数据库(SQLite)。其中,MVCC可以为数据库中的每条记录维护多个快照副本,通过起始时间戳(begin timestamp)和结束时间戳(end timestamp)维护副本的可见性。第二存储系统子模块是支持持久化的存储系统,比如KV系统,闪存文件系统(flash friendly file system,F2FS),第四代扩展文件系统(fourth extended filesystem,EXT4)等。其中,持久化(persistence)的含义是将数据在持久状态和瞬时状态间转换的机制。通俗的讲,就是瞬时数据(比如内存中的数据)持久化为持久数据,该持久化数据能够长久地被保存。数据管理设备在对数据对象进行访问时,先访问该数据对象在数据库中的记录,根据记录内容再操作存储在第二存储系统中的数据,从而能够借助数据库的并发控制实现第二存储系统(如文件系统,KV系统)的并发控制访问。对于数据对象的插入和修改操作,数据管理设备必须在文件系统,KV系统操作完成后,才能提交数据库的事务;对于删除操作,数据管理设备必须先操作数据库中的数据,提交事务后,才能再操作文件系统,KV系统中的数据。

[0051] 图1C示出了数据管理设备的一种更具体的实现方式。根据图1C,操作模块可以包括数据定义子模块,数据插入、更新、删除、查询子模块,数据校验子模块。其中,该数据定义子模块用于根据定义指令确定数据对象的定义信息。该数据插入、更新、删除、查询子模块,数据校验子模块可以根据插入指令、更新指令、删除指令、查询指令对实际存储的数据对象执行插入、更新、删除、查询等操作。该数据校验子模块可以对不同存储系统中的数据进行校验,以清除无效的数据,保证多个存储系统中的数据一致性。需要说明的是,还可以按照需要对操作模块进行其他的划分方式,例如,将该数据插入、更新、删除、查询子模块划分为数据插入子模块、数据更新子模块、数据删除子模块、数据查询子模块,本申请实施例不作具体的限制。

[0052] 可选的,该操作模块还可以包括第一存储系统操作子模块和第二存储系统操作子模块。其中,该第一存储系统操作子模块用于对第一存储系统中的数据执行操作,第二存储系统操作子模块用于对第二存储系统中的数据执行操作。

[0053] 可选的,由于该存储系统模块可以包括多个存储系统,该操作模块还可以包括系统适配子模块,该系统适配子模块可以对数据对象进行处理,使得该数据对象可以适配多个存储系统,或者使得操作反馈的数据对象适应不同应用程序的应用环境。例如,可以包括第一存储系统适配子模块和第二存储系统适配子模块。举例而言,该第一存储系统适配子模块可以包括数据库适配子模块,它可以对接不同的数据库,能够方便进行数据库切换,对数据库操作接口封装,对上层业务提供类似数据库的接口,包括打开数据库(open),执行数

数据库的操作创建(create),插入(insert),更新(update),删除(delete),查询(query)操作,以及执行开始事务(begin),提交事务(commit)等事务操作。在又一种可能的实现方案中,该第二存储系统适配子模块可以包括KV存储系统适配子模块,可以对接不同的KV存储系统,能够方便进行KV存储切换,对KV操作接口封装,对上层业务提供类似KV的接口,包括输入(put),输出(get),删除(delete)等操作。在又一种可能的实现方案中,该第二存储系统适配子模块还可以包括文件存储系统适配子模块,可以对接不同的文件存储系统,能够方便文件存储系统切换,对文件存储系统接口封装,对上层业务提供类似文件存储系统的接口,包括打开文件(open),读文件(read),写文件(write),关闭文件(close)等操作。参见图1C,是本申请实施例提供的又一种数据管理设备的示意图。

[0054] 在又一种可能的实现方式中,本申请实施例的数据管理方法还可以应用于一种数据管理系统,参见图2,是本申请实施例提供的一种数据管理系统的架构示意图,该系统包括客户端和数据管理设备。以下对这两种设备进行进一步的介绍。

[0055] 客户端,是为客户提供本地服务的设备。除了一些只在本地运行的应用程序之外,一般客户端的运行需要与服务器互相配合。较常用的客户端包括了如万维网使用的网页浏览器,收寄电子邮件时的电子邮件客户端,存储图像或者视频的相册客户端,创建文本的文本客户端以及即时通讯的客户端软件等。在本申请实施例中,该客户端可以接收用户输入的针对数据对象的操作指令,该操作指令可以包括插入指令,更新指令,定义指令,查询指令,删除指令,校验指令,等等。举例而言,该客户端可以为存储图像或者视频的相册客户端,可以接收用户输入的对图像或者视频的操作指令;该客户端可以为创建文本的文本客户端,可以接收用户输入的对文本的操作指令;该客户端可以为即时通讯的客户端软件,可以接收用户输入的对该软件中的办公文档、文本、图片、图像和音频/视频等数据的操作指令。

[0056] 数据管理设备,是为客户端提供数据存储和处理服务的设备,可以实现对数据的管理,例如,该管理可以包括定义,存储,更新,删除,校验,等等。该客户端和该数据管理设备为独立的两台设备,该客户端与该服务器通过网络或者数据线进行通信。该数据管理设备可以接收来自客户端的操作指令,再对数据对象执行该操作指令。该数据管理设备的结构可以参照上述图1A~1C所描述的结构,仅将上述图1A~1C中示意的“应用模块”替换为“接收模块”,该接口模块用于从客户端接收针对数据对象的操作指令。另外,上述图1A~1C中示意的模块中除该应用模块之外的剩余模块功能均可参照上述的描述,此处不再赘述。

[0057] 参见图3,是本申请实施例提供的一种数据管理方法的流程图。下面描述的数据管理设备可以是图1A~1C及图2任一所示的数据管理设备;该方法包括但不限于如下步骤。

[0058] S301、在关系数据表中生成数据对象的记录。

[0059] 其中,该数据对象具有多个属性,该多个属性包括结构化属性和非结构化属性。结构化属性是用于描述或定义结构化数据的特征的属性,非结构化属性是用于描述或定义非结构化数据的特征的属性。结构化数据是由二维表结构来逻辑表达和实现的数据,主要通过关系型数据库进行存储和管理;非结构化数据是数据结构不规则或不完整,没有预定义的数据模型,不便于数据库二维逻辑表来表现的数据,比如文档、文本、图片、报表、图像、音频/视频信息等。在关系数据表中生成的记录包括所述数据对象的结构化属性和该结构化属性对应的数据,以及结构化属性和非结构化属性的关联关系,所述关系数据表存储于

第一存储系统中。

[0060] S302、将该数据对象的非结构化属性对应的数据存储到第二存储系统中。

[0061] 关系数据表中一个数据对象的记录可以包含结构化属性字段和非结构化属性字段,结构化属性字段的值为结构化属性对应的数据,非结构化字段的值为非结构化属性对应的数据的标识,如键值、路径等。进一步地,非结构化属性对应的非结构化数据存储在第一存储系统中。也就是说,通过第一存储系统中的关系数据表,可以将数据对象的结构化属性和非结构化属性关联起来,关系数据表中的记录包含了结构化属性和非结构化属性,以及结构化属性和非结构化属性各自对应的数据。需要说明的是,在关系数据表的记录中,结构化属性对应的数据为数据本身,即数据内容或数据值,而非结构化属性对应的数据,并非原始数据内容,而是数据的一个标识,真正的数据内容存储在第二存储系统中。下面的实施例会进行详细说明。

[0062] S303、接收操作指令,该操作指令用于对该数据对象执行操作。

[0063] 具体的,该操作指令可以是使用数据库定义语言(data definition language, DDL),数据操纵语言(data manipulation language, DML)等描述的查询语句(query),校验语句(check),删除语句(delete)或者是函数调用语句等等。该操作指令指示了操作所涉及的该数据对象所属对象类型。可选的,该操作指令还可以包括对该数据对象执行操作所需的数据,比如该数据对象的结构化属性对应的数据和非结构化属性对应的数据。

[0064] S304、响应于该操作指令,从该第一存储系统中确定该数据对象的记录。

[0065] S305、根据该记录从该第一存储系统和该第二存储系统中的至少一个存储系统中获取该数据对象的该多个属性中的至少一个属性对应的数据。

[0066] S306、基于该至少一个属性对应的数据,对该数据对象执行该操作。

[0067] 在一个实施例中,在该数据管理设备生成该数据对象的记录之前,可以基于数据定义指令来定义该数据对象,即定义该数据对象的各个属性的名称和类型,比如,定义的该数据对象的一个或多个结构化属性,以及一个或多个非结构化属性。在一种具体的实现方式中,该数据对象的定义过程包括:该数据管理设备接收针对所述数据对象所属对象类型的定义指令,所述定义指令中包含所述对象类型的定义信息,所述定义信息用于定义所述对象类型的关系数据表的结构;根据所述定义指令在所述第一存储系统中,生成所述对象类型的关系数据表以关联该数据对象的结构化属性和非结构化属性。

[0068] 举例而言,该定义信息可以是“图片”这种对象类型的定义信息,该数据对象的非结构化属性可以包括文件和键值。该“图片”类型的定义信息可以为“picture (name STRING, size INT, path FILE, latitude DOUBLE, longitude DOUBLE, time\_taken STRING, thumbnail KV)”,根据该定义信息生成的所述对象类型“图片”的关系数据表可以参照表一所示。

[0069] 表一

[0070]

属性名称	name	size	path	latitude	longitude	time_taken	thumbnail
属性类型	STRING	INT	FILE	DOUBLE	DOUBLE	STRING	KV

[0071] 其中,该“图片”类型的定义信息中的结构化属性为“命名(name)”“大小(size)”“纬度(latitude)”“经度(longitude)”“拍摄时间(time\_taken)”,该非结构化属性为“路径

(path)”和“缩略图(thumbnail)”,其中,“path”为文件属性,而“thumbnail”为KV属性。包含该定义信息的定义指令可以为“create (picture (name STRING, size INT, path FILE, latitude DOUBLE, longitude DOUBLE, time\_taken STRING, thumbnail KV))”。其中,“图片 (picture)”为数据对象所属对象类型。

[0072] 步骤S301中,该数据管理设备在关系数据表中生成数据对象的记录可能存在两种情况。第一种情况是基于接收到的插入指令生成该数据对象的记录,第二种情况是基于接收到的更新指令生成该数据对象的记录。其中,插入指令用于插入该数据对象,更新指令用于更新该数据对象。针对该数据对象的插入指令和更新指令均指示了所述数据对象的对象类型,以及所述数据对象的属性对应的数据。以下将对这两种情况进行具体的介绍。

[0073] 针对第一种情况,该数据管理设备接收到用于插入所述数据对象的插入指令后,在关系数据表中生成数据对象的记录的过程包括:根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;以及,在所述数据对象的非结构化属性对应的数据存储在第二存储系统中之后提交所述插入指令或更新指令对应的事务。

[0074] 其中,该数据管理设备可以根据所述插入指令确定所述数据对象所属的对象类型;根据所述对象类型确定所述数据对象对应的关系数据表。

[0075] 举例而言,该针对数据对象的插入指令为“insert (picture (“Snoopy”, 2M, (data/Snoopy.jpg, Snoopy.jpg), 39.92, 116.46, “2018-10-12”, (t\_snoopy\_key, t\_snoopy.jpg)))”。该数据管理设备可以根据该插入指令确定该数据对象所属的对象类型为“图片”,再根据“图片”这种对象类型确定该数据对象对应的关系数据表为,第一存储系统中的“图片”这种对象类型的关系数据表。该关系数据表可以参照上述表一。

[0076] 具体的,该数据管理设备可以从该插入指令中获取该数据对象的结构化属性对应的数据和非结构化属性对应的数据。举例而言,该插入指令为“insert (picture (“Snoopy”, 2M, (data/Snoopy.jpg, Snoopy.jpg), 39.92, 116.46, “2018-10-12”, (t\_snoopy\_key, t\_snoopy.jpg)))”。该数据管理设备可以从更新指令中获取该数据对象的结构化属性“命名 (name)”,“大小 (size)”,“纬度 (latitude)”,“经度 (longitude)”,和“拍摄时间 (time\_taken)”各自对应的数据“Snoopy”, 2M, 39.92, 116.46, “2018-10-12”,并在关系数据表中将这些数据填充至对应的结构化属性字段中。同时,该数据管理设备可以获取非结构化属性“路径 (path)”和“缩略图 (thumbnail)”各自对应的数据 (data/Snoopy.jpg, Snoopy.jpg), (t\_snoopy\_key, t\_snoopy.jpg),并在关系数据表中将这些数据填充至对应的非结构化属性字段中。可以看出,数据对象的非结构化属性对应的数据包括数据的标识和内容。例如,非结构化属性“路径 (path)”对应的数据 (data/Snoopy.jpg, Snoopy.jpg) 包括非结构化数据的标识,即路径: data/Snoopy.jpg, 以及数据内容,即Snoopy.jpg文件;非结构化属性“缩略图 (thumbnail)”对应的数据 (t\_snoopy\_key, t\_snoopy.jpg) 包括缩略图标识 t\_snoopy\_key 和缩略图的内容 t\_snoopy.jpg。

[0077] 以下介绍在插入数据对象的过程中,该数据管理设备在所述数据对象对应的关系数据表中生成所述数据对象的记录的详细过程。

[0078] 在一个实施例中,该数据对象的非结构化属性包括KV属性,所述第二存储系统为KV存储系统。该数据管理设备根据所述结构化属性对应的数据以及所述非结构化属性对应

的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录的方法包括:根据第一版本标识和该KV属性对应的数据中的第一键值生成第二键值;根据记录中的结构化属性对应的数据,以及记录中的KV属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第二键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

[0079] 其中,该版本标识用于指示该记录的版本,在数据对象的更新过程中,该记录的版本标识可以区分该记录为更新前的数据对象的记录,还是更新后的数据对象的记录。可选的,该第一版本标识代表的含义是该数据对象更新前已在数据库中存储的记录的版本标识。该第二版本标识代表的含义是该数据对象更新后在数据库中存储的记录的版本标识。在一种可能的情况下,该数据管理设备中仅存在两个版本标识,例如version1和version2。若该第一版本标识为version1,则该数据管理设备确定第二版本标识为version2;若该第一版本标识为version2,则该数据管理设备确定第二版本标识为version1。在又一种可能的情况下,该数据管理设备中可以存在多个版本标识,例如version1,version2和version3等等。若该第一版本标识为version1,则该数据管理设备确定第二版本标识为version2或者除version1的其他版本标识;若该第一版本标识为version2,则该数据管理设备确定第二版本标识为version3或者除version2的其他版本标识。

[0080] 可选的,根据第一版本标识和该KV属性对应的数据中的第一键值生成第二键值的方式可以包括:在该KV属性对应的数据中的第一键值中添加该第一版本标识以生成该第二键值。需要说明的是,还可以存在其他根据第一版本标识和该KV属性对应的数据中的第一键值生成第二键值的方式,此处不做限制。

[0081] 在一种实施例中,该数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统。该数据管理设备根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录的方法包括:根据所述第一版本标识和所述文件属性对应的数据中的第一路径生成第二路径;根据记录中的结构化属性对应的数据,以及记录中的文件属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第二路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。

[0082] 可选的,根据该第一版本标识和该文件属性对应的数据中的第一路径生成第二路径的方式可以包括:在该文件属性对应的数据中的第一路径中添加该第一版本标识以生成该第二路径。其中,若在执行该插入指令之前,该第二路径不存在,则该数据管理设备在该数据库中创建该第二路径。需要说明的是,还可以存在其他根据该第一版本标识和该文件属性对应的数据中的第一路径生成第二路径的方式,此处不做限制。

[0083] 以下以数据对象picture (“Snoopy”, 2M, (data/Snoopy.jpg, Snoopy.jpg), 39.92, 116.46, “2018-10-12”, (t\_snoopy\_key, t\_snoopy.jpg)), 在关系数据表中的记录的生成为例,介绍在关系数据表中生成数据对象的记录生成过程,其中,该数据对象包括的非结构化属性为文件属性 (data/Snoopy.jpg, Snoopy.jpg), KV属性 (t\_snoopy\_key, t\_snoopy.jpg)。针对该数据对象的该插入指令为“insert (picture (“Snoopy”, 2M, (data/Snoopy.jpg, Snoopy.jpg), 39.92, 116.46, “2018-10-12”, (t\_snoopy\_key, t\_snoopy.jpg)))”, 该数据对

象包含7个属性,该第一版本标识为version1。其中,该数据对象的第1个属性为结构化属性,则该记录中的该结构化属性对应的数据为该结构化属性对应的数据“Snoopy”。此外,该数据对象的7个属性中除该第1个结构化属性的其他结构化属性也可以参照该方式,此处不再赘述。该数据对象的第3个属性为文件属性,则根据第一版本标识version1和该文件属性对应的数据“(data/Snoopy.jpg,Snoopy.jpg)”中的第一路径“data/Snoopy.jpg”生成第二路径“data/version1/Snoopy.jpg”,则该记录中的文件属性对应的数据为该第二路径“data/version1/Snoopy.jpg”。该数据对象的第7个属性为KV属性,则根据第一版本标识version1和该的KV属性对应的数据“(t\_snoopy\_key,t\_snoopy.jpg)”中的第一键值“t\_snoopy\_key”生成第二键值“t\_snoopy\_key\_version1”,则该记录中的KV属性对应的数据为该第二键值“t\_snoopy\_key\_version1”。之后,根据记录中的结构化属性对应的数据,记录中的KV属性对应的数据,以及记录中的文件属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录。该记录可以参照表二所示。

[0084] 表二

[0085]

属性名称	name	size	path	latitude	longitude	time_taken	thumbnail
数据	Snoopy	2M	data/version1/Snoopy.jpg	39.92	116.46	2018-10-12	t_snoopy_key_version1

[0086] 数据管理设备会将该数据对象的KV属性对应的数据“(t\_snoopy\_key\_version1,t\_snoopy.jpg)”存储在KV存储系统中,将该数据对象的文件属性对应的数据“(data/version1/Snoopy.jpg,Snoopy.jpg)”存储在文件存储系统中。KV存储系统存储的KV属性对应的数据可以参照表三所示:

[0087] 表三

[0088]

属性名称	键值	缩略图
数据	t_snoopy_key_version1	t_snoopy.jpg

[0089] 文件存储系统存储的文件属性对应的数据可以参照表四所示:

[0090] 表四

[0091]

属性名称	路径	图片
数据	data/version1/Snoopy.jpg	Snoopy.jpg

[0092] 在该数据管理设备将该数据对象的非结构化属性对应的数据存储到第二存储系统之后,该数据管理设备将提交该插入指令对应的数据库事务。需要说明的是,该数据管理设备提交数据库事务的操作需要在第一存储系统和第二存储系统均保存了相应的数据之后执行,这种操作方式可以确保该数据对象的数据在各个存储系统中均成功存储。

[0093] 步骤S301中,针对第二种情况,该数据管理设备接收到用于更新所述数据对象的更新指令后,生成该数据对象的关系数据表的过程包括:在所述数据对象对应的关系数据表中生成所述数据对象的记录;提交所述插入指令或更新指令对应的事务;其中,所述插入指令或更新指令对应的事务在所述数据对象的非结构化属性对应的数据存储到第二存储系统之后提交。

[0094] 其中,该数据管理设备可以根据所述更新指令确定所述数据对象所属的对象类型;根据所述对象类型确定所述数据对象对应的关系数据表。

[0095] 举例而言,该针对数据对象的更新指令为“update (picture (“Snoopy”,2M, (data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,”2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg)))”。该数据管理设备可以根据该更新指令确定该数据对象所属的对象类型为“图片”,再根据“图片”这种对象类型确定该数据对象对应的关系数据表为,第一存储系统中的“图片”这种对象类型的关系数据表。该关系数据表可以参照上述表一所示。

[0096] 以下介绍在更新过程中,该数据管理设备根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录的方法。

[0097] 在一种实施例中,该数据对象的非结构化属性包括KV属性,所述第二存储系统为KV存储系统。该数据管理设备根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录的方法包括:根据第二版本标识和所述KV属性对应的数据中的第一键值生成第三键值;根据记录中的结构化属性对应的数据,以及记录中的KV属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的KV属性对应的数据包括所述第三键值,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。其中,该第二版本的含义可以参照上述介绍的内容,另外,根据第二版本标识和该KV属性对应的数据中的第一键值生成第三键值的方式,可以参照上述介绍的,根据第一版本标识和该KV属性对应的数据中的第一键值生成第二键值的方式,此处不再赘述。

[0098] 在一种实施例中,该数据对象的非结构化属性包括文件属性,所述第二存储系统为文件存储系统。该数据管理设备根据所述结构化属性对应的数据以及所述非结构化属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录的方法包括:根据所述第二版本标识和所述文件属性对应的数据中的第一路径生成第三路径;根据记录中的结构化属性对应的数据,以及记录中的文件属性对应的数据,在所述数据对象对应的关系数据表中生成所述数据对象的记录;其中,所述记录中的文件属性对应的数据包括所述第三路径,所述记录中的结构化属性对应的数据包括所述数据对象的结构化属性对应的数据。其中,根据该第二版本标识和该文件属性对应的数据中的第一路径生成第三路径的方式,可以参照上述介绍的,根据该第一版本标识和该文件属性对应的数据中的第一路径生成第二路径的方式,此处不再赘述。

[0099] 以下以数据对象picture (“Snoopy”,2M, (data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,”2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg)),在关系数据表中的记录的生成为例,介绍在关系数据表生成数据对象的记录的生成过程,其中,该数据对象包括的非结构化属性为文件属性 (data/Snoopy.jpg,Snoopy.jpg),KV属性 (t\_snoopy\_key,t\_snoopy.jpg)。该更新指令为“update (picture (“Snoopy”,2M, (data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,”2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg)))”,该数据对象包含7个属性,该第二版本标识为version2。其中,该数据对象的第1个属性为结构化属性,则该记录中的该结构化属性对应的数据包括该数据对象的结构化属性对应的数据“Snoopy”。此外,该数据对象中的7个属性中除该第1个结构化属性的其他结构化属性也可以参照该方式,此处不再赘

述。该数据对象的第3个属性为文件属性，则根据第二版本标识version2和该的文件属性对应的数据“(data/Snoopy.jpg, Snoopy.jpg)”中的第一路径“data/Snoopy.jpg”生成第三路径“data/version2/Snoopy.jpg”，则该记录中的文件属性对应的数据为该第三路径“data/version2/Snoopy.jpg”。该数据对象的第7个属性为KV属性，则根据第二版本标识version2和该的KV属性对应的数据“(t\_snoopy\_key, t\_snoopy.jpg)”中的第一键值“t\_snoopy\_key”生成第三键值“t\_snoopy\_key\_version2”，该记录中的KV属性对应的数据为该第三键值“t\_snoopy\_key\_version2”。之后，根据记录中的结构化属性对应的数据，记录中的KV属性对应的数据，以及记录中的文件属性对应的数据，在所述数据对象对应的关系数据表中生成所述数据对象的记录。该记录可以参照表五所示。

[0100] 表五

[0101]

属性名称	name	size	path	latitude	longitude	time_taken	thumbnail
数据	Snoopy	2M	data/version2/Snoopy.jpg	39.92	116.46	2018-10-12	t_snoopy_key_version2

[0102] 由于在更新数据对象之前，关系数据表中已经存储有该数据对象的更早版本的记录，所以数据管理设备在接收到更新指令，在关系数据库中生成该数据对象的更新版本的记录后，该第一存储系统中存储的关系数据表同时包含了新旧版本的记录，如表六所示：

[0103] 表六

[0104]

属性名称	name	size	path	latitude	longitude	time_taken	thumbnail
数据	Snoopy	2M	data/version1/Snoopy.jpg	39.92	116.46	2018-10-12	t_snoopy_key_version1
数据	Snoopy	2M	data/version2/Snoopy.jpg	39.92	116.46	2018-10-12	t_snoopy_key_version2

[0105] 相应地，KV存储系统存储的KV属性对应的数据可以参照表七所示：

[0106] 表七

[0107]

属性名称	键值	缩略图
数据	t_snoopy_key_version1	t_snoopy.jpg
数据	t_snoopy_key_version2	t_snoopy.jpg

[0108] 文件存储系统存储的文件属性对应的数据可以参照表八所示：

[0109] 表八

[0110]

属性名称	路径	图片
数据	data/version1/Snoopy.jpg	Snoopy.jpg
数据	data/version2/Snoopy.jpg	Snoopy.jpg

[0111] 在该数据管理设备将该数据对象的非结构化属性对应的数据存储到第二存储系统之后，该数据管理设备将提交该更新指令对应的数据库事务。需要说明的是，该数据管

理设备提交数据库事务的操作需要在第一存储系统和第二存储系统均保存了相应的数据之后执行,这种操作方式可以确保该数据对象在各个存储系统中均成功存储。

[0112] 在一实施例中,步骤S303中接收到的操作指令为查询指令,该查询指令中包括查询条件。步骤S304中,从所述第一存储系统中确定所述数据对象的所述记录的过程包括:从所述第一存储系统中选取满足所述查询条件的数据对象的记录。

[0113] 举例来说,该查询指令为“query (picture (“time\_taken $\geq$ 2018-10-12”))”,该查询指令的含义是查询拍摄时间在2018年10月12日或者在该拍摄时间之后的图片。则该数据管理设备遍历该第一存储系统中图片对应的关系数据表,从中选取拍摄时间大于或等于2018-10-12的记录。例如,该获取到的记录可以为picture (“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1)和picture (“Stitch”,1.5M,data/version2/Stitch.jpg,38.23,129.78,“2018-10-17”,t\_Stitch\_key\_version2)。

[0114] 在一实施例中,步骤S303中接收到的操作指令为删除指令,该删除指令包括所述数据对象的对象类型,以及所述数据对象的结构化属性对应的数据和非结构化属性对应的数据。步骤S304中,从所述第一存储系统中确定所述数据对象的所述记录的方法包括:根据所述对象类型确定所述数据对象对应的关系数据表;从所述关系数据表中确定所述数据对象的所述记录,所述记录中的结构化属性对应的数据与所述数据对象的结构化属性对应的数据相同。

[0115] 举例而言,该数据库中存储的记录包括:第一关系数据表中的记录picture (“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1),picture (“Stitch”,1.5M,data/version2/Stitch.jpg,38.23,129.78,“2018-10-17”,t\_Stitch\_key\_version2)和第二关系数据表中的记录video (“Show”,300M,data/version1/Show.avi,47.56,119.73,“2018-10-23”,t\_Show\_key\_version1)。具体的,该删除指令为“delete (picture (“Snoopy”,2M,(data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,“2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg)))”。该删除指令的含义为删除数据对象“picture (“Snoopy”,2M,(data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,“2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg))”。在接收到该删除指令之后,该数据管理设备根据所述对象类型picture确定所述数据对象对应的关系数据表为第一关系数据表。之后,该数据管理设备从所述第一关系数据表中确定所述数据对象的所述记录,所述记录中的结构化属性对应的数据与所述数据对象的结构化属性对应的数据相同。则确定出的该记录为picture (“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1)。

[0116] 以下介绍步骤S305中,该数据管理设备根据所述记录,从所述第一存储系统和所述第二存储系统中获取所述数据对象的所述多个属性对应的数据的详细过程。

[0117] 在一种可能的实现方式中,记录的非结构化属性包括KV属性,所述第二存储系统为KV存储系统。该数据管理设备根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据的方法包括:根据键值从所述第二存储系统中读取所述键值对应的KV数据,并去除所述键值中的版本标识;所述键值为所述记录中的KV属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;其中,所述数据对象的

KV属性对应的数据为,去除版本标识后的键值和所述KV数据;所述数据对象的结构化属性对应的数据为,所述记录中结构化属性对应的数据。

[0118] 在又一种可能的实现方式中,该记录的非结构化属性包括文件属性,所述第二存储系统为文件存储系统。该数据管理设备根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据,包括:根据路径从所述第二存储系统中读取所述路径对应的文件数据,并去除所述路径中的版本标识;所述路径为所述记录中的文件属性对应的数据,所述版本标识包括第一版本标识和第二版本标识;其中,所述数据对象的文件属性对应的数据为,去除版本标识后的路径和所述文件数据;所述数据对象的结构化属性对应的数据为,所述记录中结构化属性对应的数据。

[0119] 以下将以记录picture(“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1)为例,介绍该数据管理设备根据所述记录从所述第一存储系统和所述第二存储系统中获取所述数据对象的多个属性中对应的数据的过程。

[0120] 其中,该记录中的第1个属性为结构化属性,则该数据对象的结构化属性对应的数据为该第1个属性对应的数据“Snoopy”,此外,该记录中的7个属性中除该第1个结构化属性的其他结构化属性也可以参照该方式,此处不再赘述。该记录中的第3个属性为路径,则根据该路径“data/version1/Snoopy.jpg”从该文件存储系统中读取该路径对应的文件数据“Snoopy.jpg”,并去除该路径中的版本标识version1,该数据对象的文件属性对应的数据为去除该版本标识后的路径“data/Snoopy.jpg”和该文件数据“Snoopy.jpg”,即“(data/Snoopy.jpg,Snoopy.jpg)”。该记录中的第7个属性为键值,则根据该键值“t\_snoopy\_key\_version1”从该KV存储系统中读取该键值对应的KV数据“t\_snoopy”,并去除该键值中的版本标识version1,该数据对象的KV属性对应的数据为去除版本标识后的键值“t\_snoopy\_key”和该KV数据“t\_snoopy”,即“(t\_snoopy\_key,t\_snoopy.jpg)”。

[0121] 在上述列举的操作指令为删除指令的实施例中,根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据,包括:所述根据所述记录从所述第一存储系统中获取所述数据对象的所述多个属性对应的数据。举例而言,若该记录为picture(“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1),则该数据管理设备获取该数据对象的该多个数据对应的数据:“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1。

[0122] 在一个实施例中,步骤S306中,若操作指令为查询指令,则基于所述至少一个属性对应的数据,对所述数据对象执行所述操作的过程包括:根据所述至少一个属性生成查询结果,并将查询结果返回给请求发起方,比如应用程序。

[0123] 以上述示例为例,该数据管理设备根据该数据对象中的7个属性中每个属性对应的数据和记录中7个属性的顺序建立该数据对象。该数据对象为“picture(“Snoopy”,2M,(data/Snoopy.jpg,Snoopy.jpg),39.92,116.46,“2018-10-12”,(t\_snoopy\_key,t\_snoopy.jpg))”。以相似的方法,该数据管理设备还可以建立数据对象picture(“Stitch”,1.5M,(data/Stitch.jpg,Stitch.jpg),38.23,129.78,“2018-10-17”,(t\_Stitch\_key\_version2,t\_Stitch.jpg))。之后,该数据管理设备将这两个数据对象作为查询结果。

[0124] 在一个实施例中,步骤S306中,若操作指令为删除指令,则基于所述至少一个属性对应的数据,对所述数据对象执行所述操作的方法为:从所述第一存储系统中删除所述至少一个属性对应的数据;提交所述删除指令对应的事务。以上述实例为例,则该数据管理设备在该第一存储系统的第一关系数据表中删除该数据对象的该多个数据对应的数据:“Snoopy”,2M,data/version1/Snoopy.jpg,39.92,116.46,“2018-10-12”,t\_snoopy\_key\_version1。删除之后,该数据管理设备提交该删除指令对应的事务。

[0125] 在一种可能的实现方式中,若接收到针对数据对象的更新指令,在该数据管理设备提交数据库事务之前,该数据库中的一个数据对象存在两个对应的记录。基于数据库的多版本并发控制(multiversion currency control,MVCC)机制,该数据库只保留更新后的记录。

[0126] 以下将对这种MVCC机制进行进一步介绍。MVCC机制可以为数据库中的每条记录维护多个快照副本,通过起始时间戳(begin timestamp)和结束时间戳(end timestamp)维护副本的可见性。其中,起始时间戳用于指示一条记录何时被创建,结束时间戳用于指示一条记录何时过期(或者被删除)。需要说明的是,时间戳并不存储一条记录创建或者过期发生的实际时间,它存储的是这条记录发生时的系统版本号。该系统版本号会随着事务的创建而不断增长,每个事务在事务开始时记录自身的系统版本号。

[0127] 在一个实施例中,在执行插入指令的过程中,数据对象对应的第一记录的起始时间戳为当前存储事务的系统版本号,该第一记录的结束时间戳未定义。在执行更新指令的过程中,更新的数据对象对应的第二记录的起始时间戳为当前更新事务的系统版本号,该第二记录的结束时间戳未定义;其中,该更新事务的系统版本号大于该存储事务的系统版本号。另外,该第一记录的结束时间戳将定义为该更新事务的系统版本号。在数据管理设备执行了更新指令,提交数据库事务时,将删除该第一记录,即,在该数据管理设备提交数据库事务之后,该数据库中只保留了更新的记录。此外,若在执行更新事务的同时,另一事务对该数据对象执行了读访问,则该另一事务读取到的该数据对象的记录为第一记录,这种方式可以使得数据库的更新和读取互不阻塞。

[0128] 举例而言,在数据管理设备执行更新指令的过程中,数据库中存储的记录可以参照上述表六的内容。其中,该数据库中一个数据对象存在了两个对应的记录,为包含了第一版本标识的记录和包含了第二版本标识的记录。基于数据库的多版本并发控制机制,在该数据管理设备提交数据库事务之后,该数据库中只保留包含了第二版本标识的记录。此时,该数据库中存储的记录可以参照上述表三的内容。若在执行更新指令的过程中,另一事务对该数据对象进行了读访问,则该另一事务读取到的该数据对象的记录为包含了第一版本标识的记录。

[0129] 在执行了更新指令,并且该数据库提交数据库事务之后,该数据库中存储该数据对象对应的一个记录,以上述例举的更新过程为例,最终数据库中存储的记录可以参照表三所示的内容。但对于KV存储系统和文件存储系统而言,会存在该数据对象的两种非结构化属性的数据。对于KV存储系统而言,该数据对象对应的非结构化属性的数据有“t\_snoopy\_key\_version1,t\_snoopy.jpg”和“t\_snoopy\_key\_version2,t\_snoopy.jpg”;对于文件存储系统而言,该数据对象对应的非结构化属性的数据有“data/version1/Snoopy.jpg,Snoopy.jpg”和“data/version2/Snoopy.jpg,Snoopy.jpg”。其中,“t\_snoopy\_

key\_version1,t\_snoopy.jpg”和“data/version1/Snoopy.jpg,Snoopy.jpg”是无效的数据,这种无效的数据可以通过校验操作进行清理,以下将对该数据管理设备的执行校验操作的方法进行介绍。

[0130] 在一实施例中,该非结构化属性包括KV属性,所述第二存储系统为KV存储系统。该执行校验操作的方法包括:当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的键值,所述键值为所述第二存储系统中的KV属性对应的数据;在遍历所述键值的过程中,若在所述第一存储系统存储的记录中无法查找到与第四键值相同的键值,则在所述第二存储系统中删除所述第四键值和所述第四键值对应的KV数据,所述第四键值为所述第二存储系统中多个键值中的一个。

[0131] 以上述例举的更新后的数据对象为例,在KV存储系统中存在第四键值“t\_snoopy\_key\_version1”,但在数据库存储的记录中无法查找到与该第四键值相同的键值,则在该KV存储系统中删除该第四键值“t\_snoopy\_key\_version1”和该第四键值对应的KV数据“t\_snoopy.jpg”。

[0132] 在一实施例中,该非结构化属性包括文件属性,该第二存储系统为文件存储系统。该执行校验操作的方法包括:当接收到校验指令时或者当检测到满足校验条件时,遍历所述第二存储系统中的路径,所述路径为所述第二存储系统中的文件属性对应的数据;在遍历所述路径的过程中,若在所述第一存储系统存储的关系数据表中无法查找到与第四路径相同的路径,则在所述第二存储系统中删除所述第四路径和所述第四路径对应的文件数据,所述第四路径为所述第二存储系统中多个路径中的一个。

[0133] 以上述例举的更新后的数据对象为例,在文件存储系统中存在第四路径“data/version1/Snoopy.jpg”,但在数据库存储的记录中无法查找到与该第四路径相同的路径,则在该文件存储系统中删除该第四路径“data/version1/Snoopy.jpg”和该第四路径对应的文件数据“Snoopy.jpg”。

[0134] 其中,该校验条件可以为当前时刻处于预设的校验周期,或者该数据管理设备中已存储的数据量大于预设值,等等。举例而言,该校验指令可以为“check(picture)”,代表校验“图片”这种对象类型的数据对象。当数据管理设备接收到该校验指令时,该数据管理设备将遍历第二存储系统中的非结构属性对应的数据。通过这种方式,对比数据库中已存储的记录与第二存储系统中的非结构属性对应的数据,可以清除第二存储系统中无效的数据,能够让数据对象在跨多个数据系统存储的情况下保持数据一致性。

[0135] 需要说明的是,该数据管理设备从该第二存储系统中删除与该第一记录对应的非结构化属性所对应的数据的方法,可以参照该数据管理设备的执行校验指令的方法。由于第一记录已被删除,通过执行校验指令,可以清除第二存储系统中与该第一记录对应的非结构化数据。

[0136] 以上描述了本申请的方法实施例,下面对实现上述方法的装置实施例进行介绍。

[0137] 参见图4,是本申请实施例提供的一种数据管理设备,该设备包括数据管理设备包括生成单元401,存储单元402,接收单元403,确定单元404,获取单元405和操作单元406。下面对生成单元401,存储单元402,接收单元403,确定单元404,获取单元405和操作单元406进行介绍。

[0138] 所述生成单元401,用于在关系数据表中生成数据对象的记录,所述数据对象具有

多个属性,所述多个属性包括结构化属性和非结构化属性,所述记录指示了所述数据对象的结构化属性和非结构化属性的关联关系,所述关系数据表存储于第一存储系统中。该生成单元401所执行的操作可以参照上述图3的步骤301中的相关描述。

[0139] 所述存储单元402,用于将所述数据对象的非结构化属性对应的数据存储到第二存储系统中。该存储单元402所执行的操作可以参照上述图3的步骤302中的相关描述。

[0140] 所述接收单元403,用于接收操作指令,所述操作指令用于对所述数据对象执行操作。在一个实施例中,该接收单元403可以为数据传输接口、通信接口或接收器等可被配置用于接收信息的电路或组件,该接收单元403执行的操作可以参照上述图3的步骤303中的相关描述。

[0141] 所述确定单元404,用于响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录。该确定单元404所执行的操作可以参照上述图3的步骤304中的相关描述。

[0142] 所述获取单元405,用于根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据。该获取单元405所执行的操作可以参照上述图3的步骤305中的相关描述。

[0143] 所述操作单元406,用于基于所述至少一个属性对应的数据,对所述数据对象执行所述操作。该获取单元406所执行的操作可以参照上述图3的步骤306中的相关描述。

[0144] 另外,图4中的各个操作的具体实现细节还可以对应参照图3所示的方法实施例的相应描述。上述各个单元可以以硬件,软件或者软硬件结合的方式来实现。在一个实施例中,生成单元401,存储单元402,确定单元404,获取单元405和操作单元406可以为软件实现的功能模块,这些功能模块的功能由存储在存储器中的程序或代码实现,数据管理设备通过至少一个处理器执行这些程序或代码,可以实现各个功能模块的功能。由于该数据对象跨多个数据系统存储的多个属性对应的数据均通过关系数据表中的记录获取,该数据管理设备可以让该数据对象在跨多个数据系统存储的情况下保持数据一致性。

[0145] 参见图5,是本申请实施例提供的又一种数据管理设备。该数据管理设备包括处理器501、存储器502和通信接口503,该处理器501、存储器502和通信接口503通过总线504相互连接。

[0146] 存储器502包括但不限于是随机存储记忆体(random access memory, RAM)、只读存储器(read-only memory, ROM)、可擦除可编程只读存储器(erasable programmable read only memory, EPROM)、或便携式只读存储器(compact disc read-only memory, CD-ROM),该存储器502用于相关指令及数据。

[0147] 通信接口503可以为数据传输接口、通信接口或接收器等可被配置用于接收信息的电路或组件。

[0148] 处理器501可以是一个或多个中央处理器(central processing unit, CPU),在处理器501是一个CPU的情况下,该CPU可以是单核CPU,也可以是多核CPU。

[0149] 该数据管理设备中的处理器501通过读取并执行存储器502中存储的程序代码,执行以下操作:

[0150] 在关系数据表中生成数据对象的记录。其中,所述数据对象具有多个属性,所述多个属性包括结构化属性和非结构化属性,所述记录指示了所述数据对象的结构化属性和非

结构化属性的关联关系,所述关系数据表存储于第一存储系统中。

[0151] 将所述数据对象的非结构化属性对应的数据存储到第二存储系统中。

[0152] 接收操作指令,所述操作指令用于对所述数据对象执行操作。

[0153] 响应于所述操作指令,从所述第一存储系统中确定所述数据对象的所述记录。

[0154] 根据所述记录从所述第一存储系统和所述第二存储系统中的至少一个存储系统中获取所述数据对象的所述多个属性中的至少一个属性对应的数据。

[0155] 基于所述至少一个属性对应的数据,对所述数据对象执行所述操作。

[0156] 图5中的处理器501所执行的各个操作的具体细节还可以对应参照图3所示的方法实施例的相应描述。由于该数据对象跨多个数据系统存储的多个属性对应的数据均通过关系数据表中的记录进行获取,可以让该数据对象在跨多个数据系统存储的情况下保持数据一致性。

[0157] 在本申请的另一实施例中提供一种计算机程序产品,当该计算机程序产品在计算机上运行时,图3所示实施例的方法得以实现。

[0158] 在本申请的另一实施例中提供一种计算机可读存储介质,该计算机可读存储介质存储有计算机程序,该计算机程序被计算机执行时实现图3所示实施例的方法。

[0159] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以权利要求的保护范围为准。

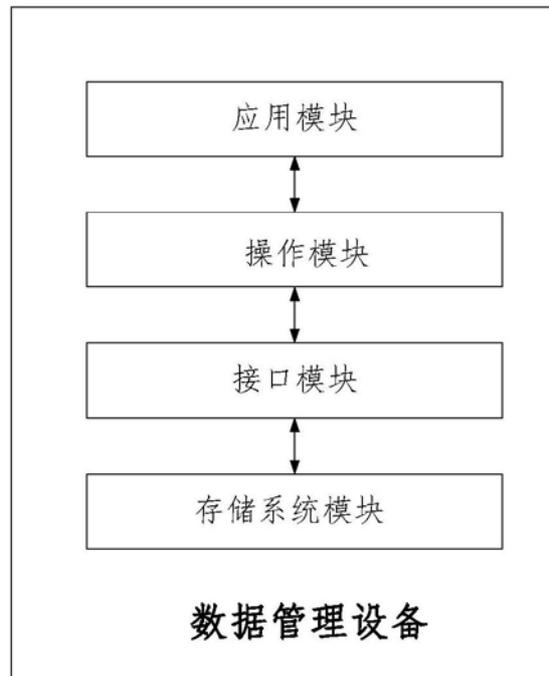


图1A

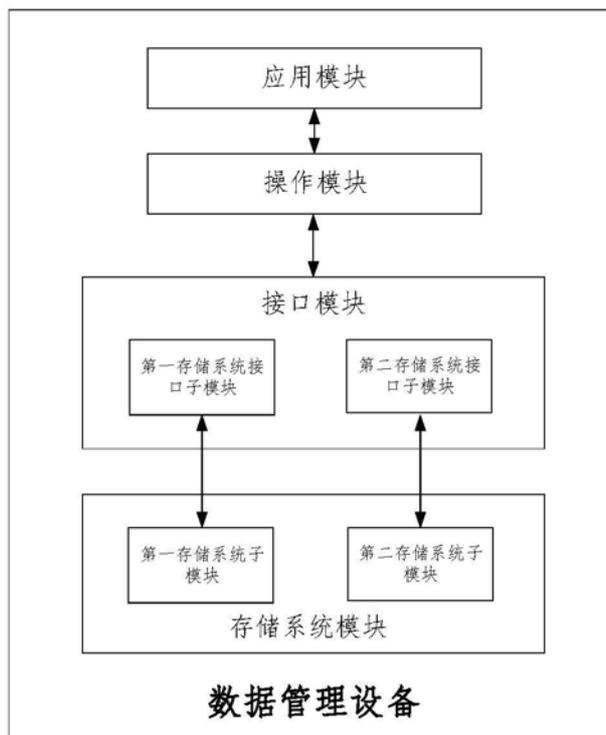


图1B

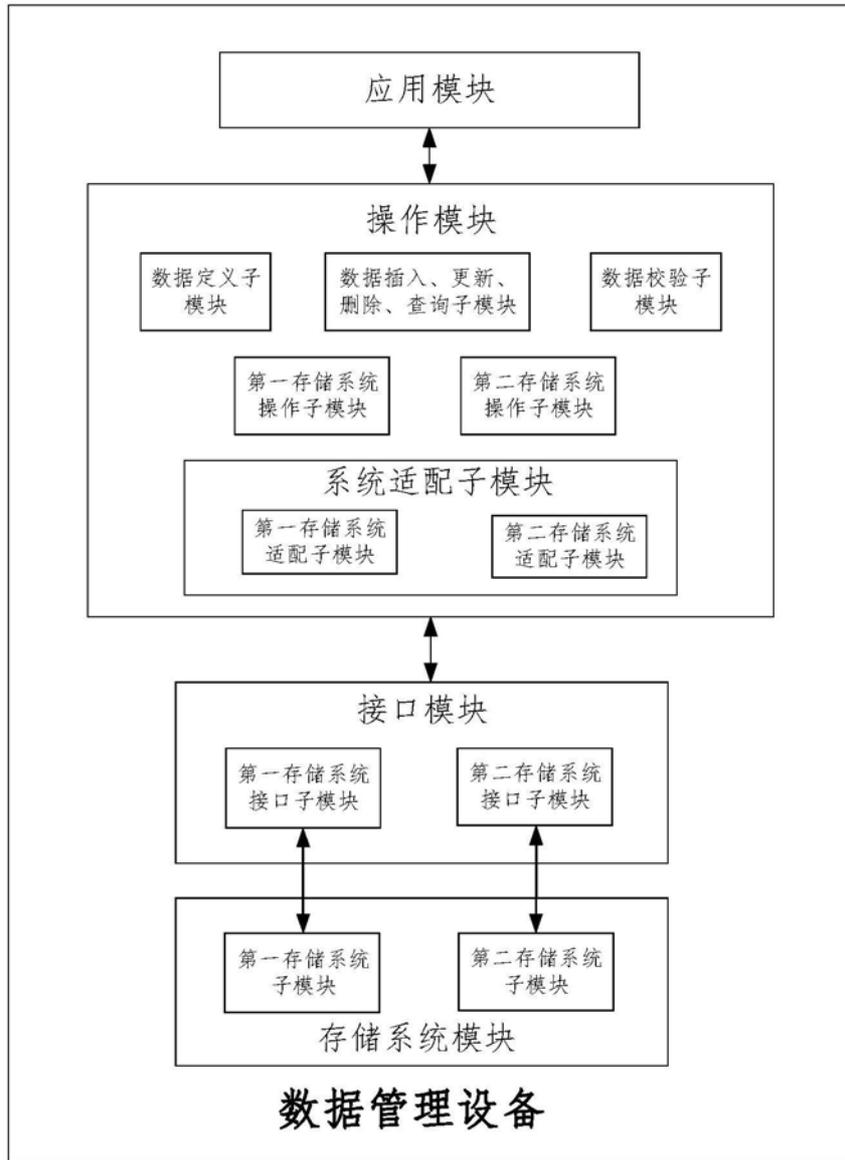


图1C

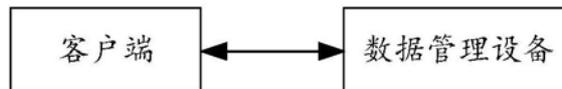


图2

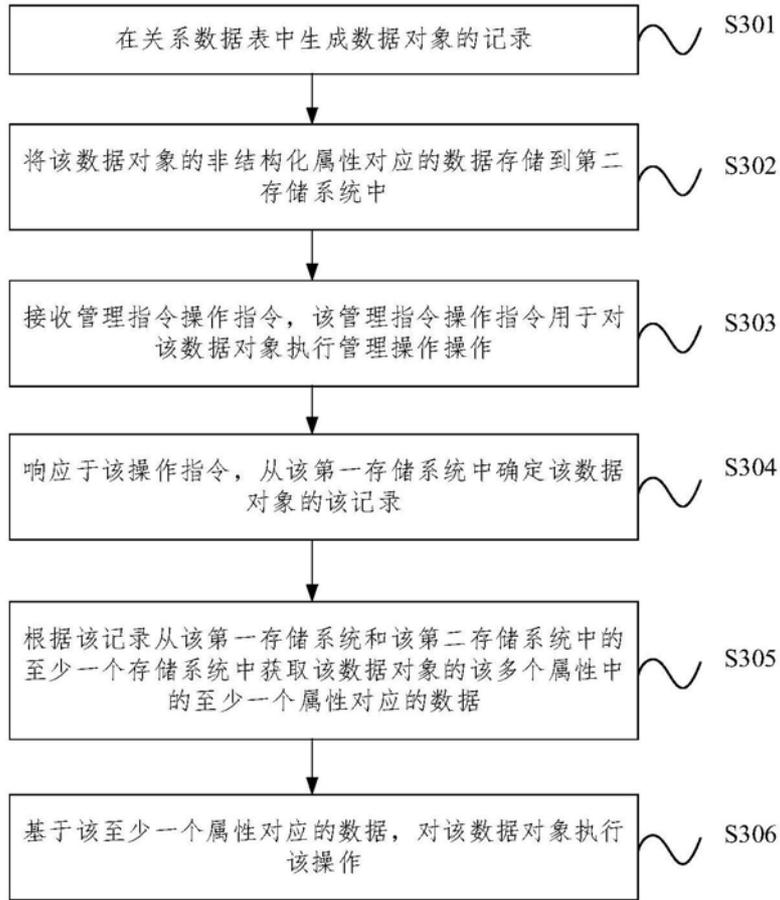


图3

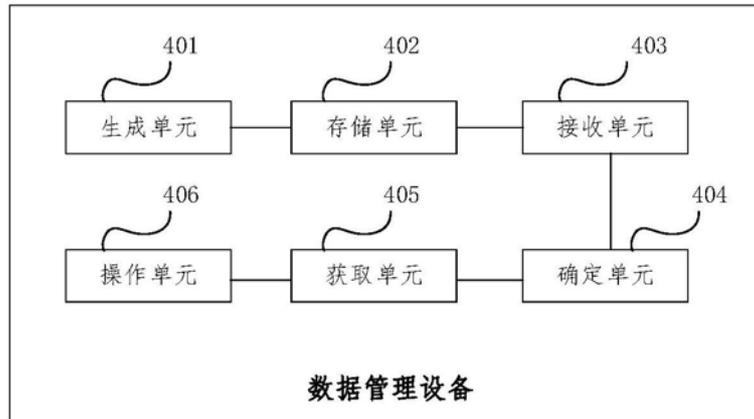


图4

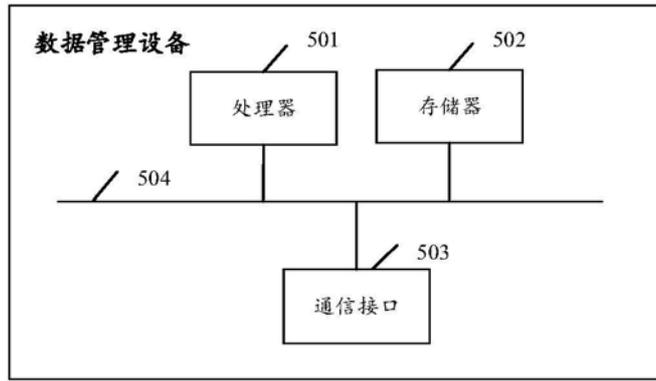


图5