



(12) 发明专利申请

(10) 申请公布号 CN 117331963 A

(43) 申请公布日 2024. 01. 02

(21) 申请号 202311597223.6

(22) 申请日 2023.11.28

(71) 申请人 恒生电子股份有限公司

地址 310053 浙江省杭州市滨江区江南大道3588号恒生大厦11楼

(72) 发明人 赵其桂 谢鹏 张敏 张君华

(74) 专利代理机构 北京超凡宏宇知识产权代理有限公司 11463

专利代理师 张文娥

(51) Int. Cl.

G06F 16/242 (2019.01)

G06F 16/25 (2019.01)

G06F 16/28 (2019.01)

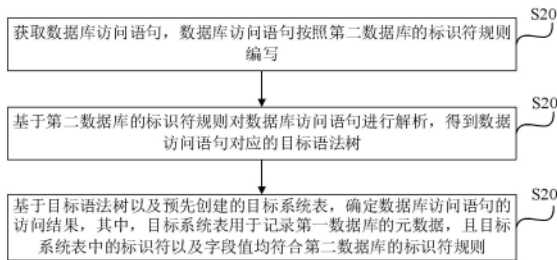
权利要求书2页 说明书13页 附图4页

(54) 发明名称

数据访问处理方法、装置、电子设备及存储介质

(57) 摘要

本申请提供了一种数据访问处理方法、装置、电子设备及存储介质,其中,该方法包括:获取数据库访问语句,数据库访问语句按照第二数据库的标识符规则编写;基于第二数据库的标识符规则对数据库访问语句进行解析,得到数据库访问语句对应的目标语法树;基于目标语法树以及预先创建的目标系统表,确定数据库访问语句的访问结果,其中,目标系统表用于记录第一数据库的元数据,且目标系统表中的标识符以及字段值均符合第二数据库的标识符规则。本申请通过对第一数据库的内核系统中语法解析和系统表的改进,使第一数据库呈现第二数据库的兼容模式,不需要根据数据库的特性重新维护业务代码,从而提高了数据库的开发和维护效率。



1. 一种数据访问处理方法,其特征在于,应用于第一数据库的内核系统,所述方法包括:

获取数据库访问语句,所述数据库访问语句按照第二数据库的标识符规则编写;

基于所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树;

基于所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,其中,所述目标系统表用于记录所述第一数据库的元数据,且所述目标系统表中的标识符以及字段值均符合所述第二数据库的标识符规则。

2. 根据权利要求1所述的方法,其特征在于,所述基于所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树,包括:

基于所述第二数据库的标识符规则对所述数据库访问语句进行词法解析,得到所述数据库访问语句的词法解析结果;

对所述词法解析结果进行语法解析,得到所述目标语法树。

3. 根据权利要求2所述的方法,其特征在于,所述基于所述第二数据库的标识符规则对所述数据库访问语句进行词法解析,得到所述数据库访问语句的词法解析结果,包括:

确定所述数据库访问语句的正则表达式;

根据所述数据库访问语句的正则表达式以及所述第二数据库的标识符规则,得到所述数据库访问语句的词法解析结果。

4. 根据权利要求3所述的方法,其特征在于,所述第二数据库为Oracle数据库;

所述根据所述数据库访问语句的正则表达式以及所述第二数据库的标识符规则,得到所述数据库访问语句的词法解析结果,包括:

若所述数据库访问语句中目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,则将所述目标标识符中除双引号之外的字符作为所述目标标识符的词法解析结果,其中,所述目标标识符为所述数据库访问语句中的任一标识符;

若所述目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,则将所述目标标识符中的字符转换为大写字符,得到所述目标标识符的词法解析结果。

5. 根据权利要求1所述的方法,其特征在于,所述基于所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,包括:

从所述目标系统表中查找所述目标语法树中的各标识符,得到标识符查找结果;

根据所述标识符查找结果,确定所述数据库访问语句的访问结果。

6. 根据权利要求5所述的方法,其特征在于,所述根据所述标识符查找结果,确定所述数据库访问语句的访问结果,包括:

若所述标识符查找结果指示所述目标系统表中不存在所述目标语法树中的至少一个标识符,则确定所述访问结果为访问失败;

否则,根据所述目标语法树,访问所述第一数据库中的至少一个数据表,得到访问结果。

7. 根据权利要求1-6任一项所述的方法,其特征在于,所述获取数据库访问语句之前,还包括:

基于所述第二数据库的标识符规则,进行所述第一数据库的实例化。

8. 根据权利要求7所述的方法,其特征在於,所述基於所述第二数据库的标识符规则,进行所述第一数据库的实例化,包括:

基於所述第二数据库的标识符规则,生成所述目标系统表;

基於所述第二数据库的标识符规则,生成所述第一数据库的初始化脚本;

基於所述目标系统表运行所述初始化脚本。

9. 根据权利要求8所述的方法,其特征在於,所述基於所述第二数据库的标识符规则,生成所述目标系统表,包括:

在所述第二数据库为Oracle数据库时,将初始系统表中的各标识符中的字符转换为大写字符,得到所述目标系统表。

10. 根据权利要求9所述的方法,其特征在於,所述方法还包括:

获取数据库更新语句,所述数据库更新语句用於在所述第一数据库中创建新的数据表或更新所述第一数据库中数据表的表名或字段名;

根据所述数据库更新语句,按照所述第二数据库的标识符规则,向所述目标系统表中写入数据。

11. 一种数据访问处理装置,其特征在於,应用於第一数据库的内核系统,所述装置包括:

获取模块,用於获取数据库访问语句,所述数据库访问语句按照第二数据库的标识符规则编写;

解析模块,用於基於所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树;

确定模块,用於基於所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,其中,所述目标系统表用於记录所述第一数据库的元数据,且所述目标系统表中的标识符以及字段值均符合所述第二数据库的标识符规则。

12. 一种电子设备,其特征在於,包括:处理器和存储器,所述存储器存储有所述处理器可执行的机器可读指令,当电子设备运行时,所述处理器执行所述机器可读指令,以执行如权利要求1至10任一项所述的数据访问处理方法的步骤。

13. 一种计算机可读存储介质,其特征在於,该计算机可读存储介质上存储有计算机程序,该计算机程序被处理器运行时执行如权利要求1至10任一项所述的数据访问处理方法的步骤。

数据访问处理方法、装置、电子设备及存储介质

技术领域

[0001] 本申请涉及数据库技术领域,具体而言,涉及一种数据访问处理方法、装置、电子设备及存储介质。

背景技术

[0002] 目前,大部分业务系统都以Oracle数据库为底座进行开发设计,但随着互联网技术的发展,适配业务的数据库越来越多,需要维护的代码版本也越多,致使数据安全问题日趋严峻。其中,普遍存在的问题是一个数据库的数据导入另一个数据库的过程中,由于两者数据库特性的差异,导致实际存储在数据库的数据存在差异。在同一个业务、同一段代码中查询数据库获取的结果不同,会最终导致整个业务执行逻辑异常。

[0003] 现有技术主要通过创建数据库对象时不使用特殊符号对标识符进行处理,或者使用特殊符号对标识符进行处理两种方式来解决上述的问题。其中,在不使用特殊符号对标识符进行处理的方式中,开发人员需要知道每个查询语句返回的标识符的大小写情况,从而编写符合预期的逻辑判断,得到正确的执行结果。在使用特殊符号对标识符进行处理的方式中,需要在创建数据库对象和引用数据库对象时添加特殊符号,并且在数据查询时添加对应的特殊符号才能在数据库中检索到正确的结果。

[0004] 但是,无论采取上述何种方式,都需要根据数据库的特性重新维护业务代码,导致数据库的开发和维护效率低下。

发明内容

[0005] 本申请的目的在于,针对上述现有技术中的不足,提供一种数据访问处理方法、装置、电子设备及存储介质,以解决现有技术中需要根据数据库的特性重新维护业务代码,导致数据库的开发和维护效率低下的问题。

[0006] 为实现上述目的,本申请实施例采用的技术方案如下:

第一方面,本申请实施例提供了一种数据访问处理方法,所述方法包括:

获取数据库访问语句,所述数据库访问语句按照第二数据库的标识符规则编写;

基于所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树;

基于所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,其中,所述目标系统表用于记录所述第一数据库的元数据,且所述目标系统表中的标识符以及字段值均符合所述第二数据库的标识符规则。

[0007] 作为一种可能的实现方式,所述基于所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树,包括:

基于所述第二数据库的标识符规则对所述数据库访问语句进行词法解析,得到所述数据库访问语句的词法解析结果;

对所述词法解析结果进行语法解析,得到所述目标语法树。

[0008] 作为一种可能的实现方式,所述基于所述第二数据库的标识符规则对所述数据库访问语句进行词法解析,得到所述数据库访问语句的词法解析结果,包括:

确定所述数据库访问语句的正则表达式;

根据所述数据库访问语句的正则表达式以及所述第二数据库的标识符规则,得到所述数据库访问语句的词法解析结果。

[0009] 作为一种可能的实现方式,所述第二数据库为Oracle数据库;

所述根据所述数据库访问语句的正则表达式以及所述第二数据库的标识符规则,得到所述数据库访问语句的词法解析结果,包括:

若所述数据库访问语句中目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,则将所述目标标识符中除双引号之外的字符作为所述目标标识符的词法解析结果,其中,所述目标标识符为所述数据库访问语句中的任一标识符;

若所述目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,则将所述目标标识符中的字符转换为大写字符,得到所述目标标识符的词法解析结果。

[0010] 作为一种可能的实现方式,所述基于所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,包括:

从所述目标系统表中查找所述目标语法树中的各标识符,得到标识符查找结果;

根据所述标识符查找结果,确定所述数据库访问语句的访问结果。

[0011] 作为一种可能的实现方式,所述根据所述标识符查找结果,确定所述数据库访问语句的访问结果,包括:

若所述标识符查找结果指示所述目标系统表中不存在所述目标语法树中的至少一个标识符,则确定所述访问结果为访问失败;

否则,根据所述目标语法树,访问所述第一数据库中的至少一个数据表,得到访问结果。

[0012] 作为一种可能的实现方式,所述获取数据库访问语句之前,还包括:

基于所述第二数据库的标识符规则,进行所述第一数据库的实例化。

[0013] 作为一种可能的实现方式,所述基于所述第二数据库的标识符规则,进行所述第一数据库的实例化,包括:

基于所述第二数据库的标识符规则,生成所述目标系统表;

基于所述第二数据库的标识符规则,生成所述第一数据库的初始化脚本;

基于所述目标系统表运行所述初始化脚本。

[0014] 作为一种可能的实现方式,所述基于所述第二数据库的标识符规则,生成所述目标系统表,包括:

在所述第二数据库为Oracle数据库时,将初始系统表中的各标识符中的字符转换为大写字符,得到所述目标系统表。

[0015] 作为一种可能的实现方式,所述方法还包括:

获取数据库更新语句,所述数据库更新语句用于在所述第一数据库中创建新的数据表或更新所述第一数据库中数据表的表名或字段名;

根据所述数据库更新语句,按照所述第二数据库的标识符规则,向所述目标系统表中写入数据。

[0016] 第二方面,本申请实施例提供了一种数据访问处理装置,应用于第一数据库的内核系统,所述装置包括:

获取模块,用于获取数据库访问语句,所述数据库访问语句按照第二数据库的标识符规则编写;

解析模块,用于基于所述第二数据库的标识符规则对所述数据库访问语句进行解析,得到所述数据库访问语句对应的目标语法树;

确定模块,用于基于所述目标语法树以及预先创建的目标系统表,确定所述数据库访问语句的访问结果,其中,所述目标系统表用于记录所述第一数据库的元数据,且所述目标系统表中的标识符以及字段值均符合所述第二数据库的标识符规则。

[0017] 作为一种可能的实现方式,所述解析模块具体用于:

基于所述第二数据库的标识符规则对所述数据库访问语句进行词法解析,得到所述数据库访问语句的词法解析结果;

对所述词法解析结果进行语法解析,得到所述目标语法树。

[0018] 作为一种可能的实现方式,所述解析模块还用于:

确定所述数据库访问语句的正则表达式;

根据所述数据库访问语句的正则表达式以及所述第二数据库的标识符规则,得到所述数据库访问语句的词法解析结果。

[0019] 作为一种可能的实现方式,所述第二数据库为Oracle数据库,所述解析模块还用于:

若所述数据库访问语句中目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,则将所述目标标识符中除双引号之外的字符作为所述目标标识符的词法解析结果,其中,所述目标标识符为所述数据库访问语句中的任一标识符;

若所述目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,则将所述目标标识符中的字符转换为大写字符,得到所述目标标识符的词法解析结果。

[0020] 作为一种可能的实现方式,所述确定模块具体用于:

从所述目标系统表中查找所述目标语法树中的各标识符,得到标识符查找结果;根据所述标识符查找结果,确定所述数据库访问语句的访问结果。

[0021] 作为一种可能的实现方式,所述确定模块还用于:

若所述标识符查找结果指示所述目标系统表中不存在所述目标语法树中的至少一个标识符,则确定所述访问结果为访问失败;

否则,根据所述目标语法树,访问所述第一数据库中的至少一个数据表,得到访问结果。

[0022] 作为一种可能的实现方式,所述装置还包括:

实例化模块,用于基于所述第二数据库的标识符规则,进行所述第一数据库的实例化。

[0023] 作为一种可能的实现方式,所述实例化模块具体用于:

基于所述第二数据库的标识符规则,生成所述目标系统表;

基于所述第二数据库的标识符规则,生成所述第一数据库的初始化脚本;

基于所述目标系统表运行所述初始化脚本。

[0024] 作为一种可能的实现方式,所述实例化模块还用于:

在所述第二数据库为Oracle数据库时,将初始系统表中的各标识符中的字符转换为大写字符,得到所述目标系统表。

[0025] 作为一种可能的实现方式,所述实例化模块还用于:

获取数据库更新语句,所述数据库更新语句用于在所述第一数据库中创建新的数据表或更新所述第一数据库中数据表的表名或字段名;

根据所述数据库更新语句,按照所述第二数据库的标识符规则,向所述目标系统表中写入数据。

[0026] 第三方面,本申请实施例提供了一种电子设备,包括:处理器、存储介质和总线,所述存储介质存储有所述处理器可执行的机器可读指令,在电子设备运行时,所述处理器与所述存储介质之间通过总线通信,所述处理器执行所述机器可读指令,以执行如上述第一方面任一所述数据访问处理方法的步骤。

[0027] 第四方面,本申请实施例提供了一种计算机可读存储介质,所述计算机可读存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行如上述第一方面任一所述数据访问处理方法的步骤。

[0028] 根据本申请实施例的数据访问处理方法、装置、电子设备及存储介质,在第一数据库的内核系统中,获取按照第二数据库的标识符规则编写的数据库访问语句,基于第二数据库的标识符规则对数据库访问语句进行解析,生成与数据库访问语句对应的目标语法树,并预先在第一数据库的内核系统中创建记录第一数据库的元数据的目标系统表,且目标系统表中的标识符以及字段值均符合第二数据库的标识符规则。基于此,对第一数据库的内核系统进行改进,具体的,通过对第一数据库的内核系统中语法解析以及系统表的改进,使第一数据库呈现第二数据库的兼容模式,进而使得第一数据库兼容第二数据库,能够解决第二数据库中标识符的大小写问题,所以无需再根据数据库的特性重新维护业务代码,从而提高了数据库的开发和维护效率。

附图说明

[0029] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,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相关的附图。

[0030] 图1示出了本申请实施例提供的一种数据访问处理方法的应用场景示意图;

图2示出了本申请实施例提供的一种数据访问处理方法的流程图;

图3示出了本申请实施例提供的一种第一数据库实例化方法的流程图;

图4示出了本申请实施例提供的一种元数据写入方法的流程示意图;

图5示出了本申请实施例提供的一种目标语法树构建方法的流程示意图;

图6示出了本申请实施例提供的一种词法解析方法的流程示意图;

图7示出了本申请实施例提供的另一种词法解析方法的流程示意图;

图8示出了本申请实施例提供的一种访问结果确定方法的流程示意图;

图9示出了本申请实施例提供的一种数据访问处理装置的结构示意图;

图10示出了本申请实施例提供的一种电子设备的结构示意图。

具体实施方式

[0031] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,应当理解,本申请中附图仅起到说明和描述的目的,并不用于限定本申请的保护范围。另外,应当理解,示意性的附图并未按实物比例绘制。本申请中使用的流程图示出了根据本申请的一些实施例实现的操作。应该理解,流程图的操作可以不按顺序实现,没有逻辑的上下文关系的步骤可以反转顺序或者同时实施。此外,本领域技术人员在本申请内容的指引下,可以向流程图添加一个或多个其他操作,也可以从流程图中移除一个或多个操作。

[0032] 另外,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以以各种不同的配置来布置和设计。因此,以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请的实施例,本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0033] 需要说明的是,本申请实施例中将会用到术语“包括”,用于指出其后所声明的特征的存在,但并不排除增加其它的特征。

[0034] 现有技术中虽然提供了两种方式来解决数据库兼容的问题,但是无论采取上述何种方式,都要根据适配的数据库特性重新维护一份业务代码,且随着业务域适配的数据库越来越多,维护的代码版本也越多,不仅数据库开发前期工程量大,且后期维护成本又高,从而导致数据库的开发和维护效率低下。

[0035] 本申请基于上述问题,提出一种数据访问处理方法,通过改进第一数据库的内核系统,使得第一数据库的内核系统能够兼容第二数据库(Oracle数据库)的功能,从而无需对业务代码进行任何改动即可正常地访问第一数据库中的数据。

[0036] 为便于解释本申请的技术方案,本申请以下实施例中以第一数据库为LightDB数据库,第二数据库为Oracle数据库为例进行说明。

[0037] 图1示出了本申请实施例提供的一种数据访问处理方法的应用场景示意图。参照图1所示,本申请的数据库访问方法可以应用于LightDB数据库的内核系统,其中,该LightDB数据库的内核系统由原始LightDB数据库的原始内核系统改进而来,能够兼容Oracle数据库的功能,因此工作在Oracle兼容模式下,该LightDB数据库相应可以服务于任意的以Oracle数据库为底座进行开发设计的业务系统。这些业务系统可以按照Oracle的标识符规则编写数据库访问语句并访问该LightDB数据库。

[0038] 继续参照图1所示,LightDB数据库的内核系统还可以工作在LightDB兼容模式下。在这种情况下,无需对LightDB数据库的内核系统进行改进。工作在LightDB模式下的LightDB数据库可以服务于按照LightDB规则进行开发设计的业务系统。这些业务系统可以按照LightDB的标识符规则编写数据库访问语句并访问该LightDB数据库。

[0039] 由于不同数据库对标识符的处理逻辑不同,所以LightDB数据库不兼容以Oracle数据库的标识符规则编写的处理逻辑,所以在针对以Oracle数据库为底座开发设计的同一个业务系统,同一段查询代码,在访问LightDB数据库时会出现访问异常情况。

[0040] 示例性的,LightDB数据库的内核系统即数据库内核,实质上是一个数据库引擎,是完成数据库管理系统(database managementsystem,DBMS)对数据操作的核心功能集合。一般的,数据库内核包括解析器、计划器、优化器、执行器和存储引擎等五部分。解析器是对数据库用户输入的结构化查询语言即SQL语句进行词法、语法的解析,判断用户的查询意图;计划器是根据SQL语句的解析结果生成相应的执行计划;优化器是对初步生成的执行计划进行优化,以便高效地完成用户请求的操作;执行器按照优化后的执行计划根据计划中的每一个操作步骤进行具体执行,得到用户需要的结果;存储引擎负责上述过程中的读写和索引等功能,从而为其他功能部件提供输入、输出能力。

[0041] 基于此,本申请对LightDB数据库的内核系统进行改进,改进后的LightDB数据库的内核系统中维护两个版本的系统表,小写版本对应LightDB兼容模式,大写版本对应Oracle兼容模式,并且对LightDB数据库的内核系统的语法解析进行改进,从而在获取利用Oracle数据库的标识符规则编写的数据库访问语句后,利用LightDB数据库的数据库内核的解析器对数据库访问语句进行语法、词法的解析,并通过计划器、优化器、执行器完成用户访问请求,得到访问结果。由此,使得LightDB数据库能够呈现并兼容Oracle模式,从而提高了数据库的开发和维护效率。

[0042] 下面结合上述图1示出的内容,对本申请实施例提供的数据库访问处理方法进行详细说明。

[0043] 参照图2所示,为本申请实施例提供的一种数据库访问处理方法的流程示意图,应用于第一数据库的内核系统,该方法具体包括如下步骤:

S201、获取数据库访问语句,数据库访问语句按照第二数据库的标识符规则编写。

[0044] 可选的,数据库访问语句是按照Oracle数据库的标识符规则编写的SQL语句,数据库访问语句例如可以是查询语句、更新语句、删除语句或新增语句等,以实现业务数据的增删改查等。在Oracle数据库中,标识符通常指表名、列名、变量名等命名对象的名称,例如在Oracle创建表时,需要给每个表命名一个唯一的名称,该名称便是一个标识符。

[0045] 可选的,标识符规则包括标识符命名规则、标识符解析规则等。具体的,标识符命名规则相对比较灵活,可以包含字母、数字和下划线,但是命名必须以字母开头,且长度不能超过64个字符。此外,在Oracle数据库中标识符是区分大小写的,也就是说,不同大小写的标识符被视为不同的对象。标识符解析规则为在当前schema中查找该标识符指代的对象,若找到则使用该对象,若在schema中未找到则在public schema中继续查找,找到则使用该对象,若在public schema中还未找到该标识符指代的对象,则说明该对象不存在。

[0046] 其中,schema是数据库的组织结构,schema对象可以是表、列、数据类型、视图、存储过程、关系、主键、外键等。创建表等其他schema对象时不指定任何模式,而在默认情况下,会将创建的表等对象自动放入名为“public”的模式中,即形成上述的public schema。

[0047] S202、基于第二数据库的标识符规则对数据库访问语句进行解析,得到数据库访问语句对应的目标语法树。

[0048] 可选的,由于第一数据库的内核系统能够兼容Oracle数据库的功能,因此,可以基于Oracle数据库的标识符解析规则对按照Oracle数据库的标识符命名规则编写的数据库访问语句进行解析。具体的,在Oracle数据库中,若数据库访问语句中的标识符使用双引号括起来,则在语法、词法解析时不改变字符串的大小写,按照原文进行存储。若数据库访问

语句中的标识符不使用双引号括起来,则默认按照大写进行存储。

[0049] 示例性的,基于词法分析器如flex工具,利用正则表达式描述词素的模式。在利用flex工具进行解析的过程中,在识别运算符、特殊符号、关键字等词法单元后,需要确定各词法单元出现的行数、开始位置和结束位置等。开始的位置为上一个识别的词法单元结束的位置,结束的位置为开始位置加上词素长度,最后返回token值完成一个词素的识别。而对于注释、空格、换行符等特殊词法单元,只需要进行识别,而不需要输出到解析结果中。如此,在对数据库访问语句进行解析后即得到数据库访问语句对应的目标语法树。目标语法树每个节点都有一个语法符号,根节点是开始符,每个节点的子节点从左到右连在一起符合语法规则。

[0050] S203、基于目标语法树以及预先创建的目标系统表,确定数据库访问语句的访问结果,其中,目标系统表用于记录第一数据库的元数据,且目标系统表中的标识符以及字段值均符合第二数据库的标识符规则。

[0051] 可选的,目标系统表是预先在LightDB数据库的内核系统中创建的符合Oracle数据库模式的系统表。为了实现数据库系统的控制,必须提供数据字典的功能,数据字典不仅存储各种对象的描述信息,而且存储系统管理所需的各种对象的细节信息。而在LightDB数据库系统中,系统表扮演着数据字典的角色,即存放结构元数据的地方,在LightDB数据库系统中表现为存放有系统信息的普通表或者视图,用户可以删除并重建表、增加列、插入和更新数值。

[0052] 需要说明的是,LightDB数据库和Oracle数据库中都有一套系统表,LightDB兼容模式对应一套小写的系统表,Oracle兼容模式对应一套大写的系统表,而目标系统表则是预先在LightDB数据库的内核系统中创建的记录LightDB数据库的元数据的系统表,并且目标系统表中的标识符和字段值均符合Oracle数据库的标识符规则。

[0053] 本实施例中,在第一数据库的内核系统中,获取按照第二数据库的标识符规则编写的数据库访问语句,基于第二数据库的标识符规则对数据库访问语句进行解析,生成与数据库访问语句对应的目标语法树,并预先在第一数据库的内核系统中创建记录第二数据库的元数据的目标系统表,且目标系统表中的标识符以及字段值均符合第二数据库的标识符规则。通过对第一数据库的内核系统中语法解析方式以及系统表的改进,使第一数据库呈现第二数据库模式,进而使得第一数据库兼容第二数据库,能够解决第二数据库的标识符的大小写问题,所以无需再根据数据库的特性重新维护业务代码,从而提高了数据库的开发和维护效率。

[0054] 作为一种可能的实现方式,在上述步骤S201获取数据库访问语句之前,该方法还包括:基于第二数据库的标识符规则,进行第一数据库的实例化。

[0055] 可选的,按照Oracle数据库的标识符规则对LightDB数据库进行数据库实例化,数据库实例化是指将数据库的逻辑结构转化为实际存储结构的过程。示例性的,数据库实例化的过程包括创建表、定义字段、设置约束、建立索引等操作,最终将数据库的逻辑结构映射到物理存储空间上。

[0056] 示例性的,如图3所示,基于第二数据库的标识符规则,进行第一数据库的实例化,具体包括如下步骤:

S301、基于第二数据库的标识符规则,生成目标系统表。

[0057] 示例性的,在第二数据库为Oracle数据库时,将初始系统表中的各标识符中的字符转换为大写字符,得到目标系统表。

[0058] 其中,初始系统表可以是LightDB数据库的默认系统表,初始系统表中的标识符均遵循LightDB的标识符规则。

[0059] 在进行第一数据库的实例化之前,需要先确定实例化Oracle兼容模式还是LightDB兼容模式,如果为LightDB兼容模式,不需要生成大写的系统表,也不需要词法解析和语义分析进行处理。只有在实例化Oracle兼容模式时,由于Oracle数据库中系统表的标识符为大写字符,所以在LightDB数据库中创建目标系统表时,需要将初始系统表中的各标识符中的字符转换为大写字符。需要注意的是,初次生成的目标系统表实质上是空系统表,只记录了一些表的表名等标识符,但并未记录标识符对应的字段值以及元数据。当工作在Oracle兼容模式下的第一数据库中创建数据表时,会相应向目标系统中写入相应的字段值和元数据。

[0060] 示例性的,可利用perl工具生成目标系统表。具体的,将初始系统表中的各标识符中的字符转换为大写字符包括将.h头文件和.dat数据文件中系统表定义的结构体中的字段全部改写为大写,在使用.h头文件定义的结构体时,将所引用的字段也转换为大写。此外,perl工具包含模板文件和执行文件,在生成目标系统表的过程中,同样需要将模板文件所引用的系统表里的字段转换为大写,在执行文件执行后,生成的与.h头文件对应的.d.h文件中宏定义的字母也变成大写,以用来创建索引。

[0061] 需要注意的是,在字符串的比较过程中不可避免应用硬编码。在LightDB数据库的内核系统中,涉及到对字符串的比较,数据库内核中比较使用函数strcmp,比较的字符串常量同样需要转换为大写。

[0062] S302、基于第二数据库的标识符规则,生成第一数据库的初始化脚本。

[0063] 示例性的,数据库初始化脚本也即数据库所需初始化视图以及全文检索函数。按照Oracle数据库的标识符命名规则、标识符解析规则等标识符规则,对LightDB数据库进行一些初始化配置,以确保数据库的正常运行。具体的,生成LightDB数据库的初始化脚本的过程包括创建数据库、创建表、插入数据、查询数据、更新数据和删除数据等。通过执行相应的命令行,即可完成对LightDB数据库的初始化。需要说明的是,初始化脚本中用到的在where条件中使用到的系统表字段的字符串也需要全部转换为大写。

[0064] S303、基于目标系统表运行初始化脚本。

[0065] 示例性的,在LightDB数据库的内核系统中生成目标系统表以及生成LightDB数据库的初始化脚本后,按照目标系统表运行初始化脚本,以使新创建的数据库具备数据库的基本能力,能够正常运转。

[0066] 基于此,按照Oracle数据库的标识符规则进行LightDB数据库的实例化,以在LightDB数据库的内核系统中生成兼容Oracle模式的目标系统表,以及生成兼容Oracle模式的初始化脚本,从而确保LightDB数据库的初始化脚本能够正常运行。

[0067] 作为一种可能的实现方式,由于目标系统表中目前只记录了兼容Oracle模式的标识符,尚未记录LightDB数据库的元数据,所以在生成目标系统表后,还需向目标系统表中写入元数据。如图4所示,该方法还包括:

S401、获取数据库更新语句,数据库更新语句用于在第一数据库中创建新的数据

表或更新第一数据库中数据表的表名或字段名。

[0068] 示例性的,数据库更新语句同样按照Oracle数据库的标识符规则编写的,而利用数据库更新语句在LightDB数据库中创建新的数据表,或者更新LightDB数据库中数据表的表名或字段名,其目的在于向目标系统表中写入元数据,为后续的数据查询奠定基础。

[0069] S402、根据数据库更新语句,按照第二数据库的标识符规则,向目标系统表中写入数据。

[0070] 示例性的,目标系统表是由列和行组成的二维结构,利用数据库更新语句将元数据添加到目标系统表的过程中,需要获知元数据对应的标识符在目标系统表中的行列值,从而将各标识符对应的元数据准确存入目标系统表的对应位置中。

[0071] 需要注意的是,在将元数据存入目标系统表之前首先需要打开目标系统表,而目标系统表无法直接打开,需要通过对象标识符(Object Identifier,OID)打开目标系统表,对象标识符OID是与对象相关联的用来无歧义地标识对象的全局唯一的值,能够保证元数据在存入目标系统表的过程中对元数据的存储位置进行准确定位和管理。

[0072] 基于此,在生成目标系统表的同时,根据数据库更新语句,并按照第二数据库的标识符规则向目标系统表中存入记录LightDB数据库的元数据,以便确定数据库访问语句的访问结果。

[0073] 作为一种可能的实现方式,如图5所示,上述步骤S202基于第二数据库的标识符规则对数据库访问语句进行解析,得到数据库访问语句对应的目标语法树,具体包括如下步骤:

S501、基于第二数据库的标识符规则对数据库访问语句进行词法解析,得到数据库访问语句的词法解析结果。

[0074] 可选的,词法解析工具flex读取数据库访问语句,此时将输入到词法解析工具flex中的数据库访问语句作为LEX源文件,内含正则表达式以及相应模式的处理代码,词法解析工具flex通过对源文件的扫描自动生成相应的词法解析函数,例如int yylex(),并将其输出到文件名规定为lexyy.c的文件中,该文件即为LEX的输出文件或输出的词法分析器,进而利用词法分析器按照Oracle数据库的标识符解析规则对数据库访问语句进行词法解析,得到数据库访问语句的词法解析结果。

[0075] 示例性的,如图6所示,上述步骤S501具体包括如下步骤:

S601、确定数据库访问语句的正则表达式。

[0076] 示例性的,正则表达式又称规则表达式,是对字符串和特殊字符操作的一种逻辑公式,通常被用来检索、替换符合某个模式或规则的文本。目前,许多程序设计语言都支持利用正则表达式进行字符串操作,例如在词法解析工具flex中就内建了一个功能强大的正则表达式引擎,利用事先定义好的一些特殊字符、以及这些特殊字符的组合,组成一个规则字符串用来表达对字符串的一种过滤逻辑。因此,可利用词法解析工具flex为含双引号标识符和不含双引号标识符分别创建对应的正则规则。

[0077] S602、根据数据库访问语句的正则表达式以及第二数据库的标识符规则,得到数据库访问语句的词法解析结果。

[0078] 示例性的,若数据库访问语句中目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,则将目标标识符中除双引号之外的字符作为目标标识符的词法解析结

果,其中,目标标识符为数据库访问语句中的任一标识符;

若目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,则将目标标识符中的字符转换为大写字符,得到目标标识符的词法解析结果。

[0079] 示例性的,如图7所示,首先判断数据库访问语句中目标标识符是否使用双引号,并利用flex工具为双引号的正则表达式匹配创建规则,并进行语法解析,在执行器处理后即得到词法解析结果。具体的,若数据库访问语句中目标标识符使用了双引号,则在数据库访问语句进行词法解析时,目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,即利用正则表达式去匹配xdstart规则,以将双引号内的字符串按照原文保存到结构体yy1val的str成员中,也即将目标标识符中除双引号之外的字符作为目标标识符的词法解析结果。若数据库访问语句中目标标识符没有使用双引号,则在数据库访问语句进行词法解析时,目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,即利用正则表达式去匹配identifier规则,修改LightDB数据库的内核系统中词法解析identifier正则表达式逻辑处理部分,将目标标识符中的字符转换为大写字符,例如将yytext中yy1leng长度的字符串转换成大写。

[0080] 需要说明的是,在判断数据库访问语句中目标标识符是否使用双引号的过程中,以及在利用目标标识符的正则表达式与不含双引号标识符或含双引号标识符对应的正则规则匹配的过程中,都需要进行字符串比较,而涉及到字符串比较的地方都需要用到硬编码处理。

[0081] S502、对词法解析结果进行语法解析,得到目标语法树。

[0082] 示例性的,语法解析一个非常重要的功能就是构建一个树形数据结构,无论是解释器执行当前语句还是编译器将语句转换为低级语言,解析树都是一个承上启下的重要结构。所以根据词法解析工具flex生成的词法解析结果和语法规则构建语法解析树,将符合文法规则的词法解析结果,词法解析结果实质上也是字符串,即将字符串转换成一个层次结构,即得到目标语法树。

[0083] 具体地,生成目标语法树的过程可以包括词法分析、语法分析和语义分析三个步骤。其中,词法分析是指将词法分析结果的源代码分割成一系列单词,即词法单元,如变量名、数字、运算符等。语法分析是指根据词法分析器生成的词法单元序列和语法规则构建语法解析树。语义分析是指对语法解析树进行上下文分析,检查是否存在语义错误,例如检查变量类型是否匹配、函数参数是否正确等。

[0084] 需要注意的是,在目标语法树中,树的节点表示语法规则的符号或终端符号,而树的边则表示规则应用的过程。对于一个给定的词法解析结果也即字符串,其对应的语法树是唯一的。

[0085] 基于此,对词法解析结果进行语法解析,生成目标语法树,相对于语法列表或语法图,语法树能够准确地表示词法解析结果中的语法结构,并且更便于程序分析、优化和代码生成。此外,语法树提供了更丰富的语法信息,利用语法树也更容易进行语法分析和错误检测。

[0086] 作为一种可能的实现方式,如图8所示,上述步骤S203基于目标语法树以及预先创建的目标系统表,确定数据库访问语句的访问结果,具体包括如下步骤:

S801、从目标系统表中查找目标语法树中的各标识符,得到标识符查找结果。

[0087] 示例性的,对词法解析结果进行语法解析生成目标语法树后,目标语法树中包括了多个标识符,例如指示数据库访问语句所要访问的数据表的表名等。然后根据目标语法树中的各标识符从目标系统表中查找与之对应的标识符,例如,目前目标语法树中存在的标识符有ABC、ABE、DEF,目标系统表中存储有标识符ABC、ABD、ABE、BCD、DEF及各标识符对应的元数据,则得到标识符查找结果即为目标系统表中与目标语法树中标识符相对应的标识符ABC、ABE、DEF。也就是说,标识符查找结果表征了目标系统表中是否存在与目标语法树中的标识符相同的标识符。

[0088] S802、根据标识符查找结果,确定数据库访问语句的访问结果。

[0089] 示例性的,若标识符查找结果指示目标系统表中不存在目标语法树中的至少一个标识符,则确定访问结果为访问失败;否则,根据目标语法树,访问第一数据库中的至少一个数据表,得到访问结果。

[0090] 示例性的,继续以上述目标语法树中存在的标识符有ABC、ABE、DEF,目标系统表中存储有标识符ABC、ABD、ABE、BCD、DEF及各标识符对应的元数据为例,目标系统表中存在目标语法数中的标识符ABC、ABE、DEF,即表示访问结果为访问成功,获取LightDB数据库中标识符ABC、ABE、DEF对应的数据表的元数据即得到访问结果。反之,若目标语法树中存在的标识符有BCD、DEF,而目标系统表中存储的标识符为ABC、ABD、ABE,则目标系统表中不存在目标语法树中的标识符,也就说明LightDB数据库中不存在与标识符BCD、DEF对应的数据表,即访问失败。

[0091] 需要说明的是,在判断目标系统表中是否存在目标语法树中的至少一个标识符的过程中,同样需要进行字符串比较,而涉及到字符串比较的地方同样需要用到硬编码处理。

[0092] 基于此,根据本申请实施例的一种数据访问处理方法,在进行LightDB数据库的初始化实例时,生成符合Oracle兼容模式大写的系统表。在对数据库访问语句进行词法解析时,首先确定数据库访问语句中目标标识符是否使用了双引号,并将不加引号的目标标识符转换成大写。由此,LightDB数据库的数据库内核通过兼容模式以及硬编码处理将涉及到字符串比较的关键字全部转换为大写,并且比较时严格进行标识符大小写比较。

[0093] 此外,本申请实施例对LightDB数据库和Oracle数据库的数据库性能并未造成影响,反而基于LightDB数据库的内核系统,实现对Oracle标识符大小写的兼容,使得基于Oracle数据库为底座开发的业务系统能够适配LightDB数据库,且并不影响LightDB数据库的正常运行。此外,在基于LightDB数据库的内核系统兼容Oracle标识符大小写的过程中,仅仅涉及到修改标识符及对应字符串的大小写,开发人员不需要重新维护业务代码,因此大大缩短了开发周期,且能够实现快速上线,从而提高了数据库的开发和维护效率。

[0094] 基于同一发明构思,本申请实施例中还提供了与数据访问处理方法对应的数据访问处理装置,由于本申请实施例中的装置解决问题的原理与本申请实施例上述数据访问处理方法相似,因此装置的实施可以参见方法的实施,重复之处不再赘述。

[0095] 参照图9所示,为本申请实施例提供的一种数据访问处理装置的结构示意图,所述数据访问处理装置900包括:获取模块901、解析模块902、确定模块903,其中:

获取模块901,用于获取数据库访问语句,数据库访问语句按照第二数据库的标识符规则编写;

解析模块902,用于基于第二数据库的标识符规则对数据库访问语句进行解析,得

到数据库访问语句对应的目标语法树；

确定模块903,用于基于目标语法树以及预先创建的目标系统表,确定数据库访问语句的访问结果,其中,目标系统表用于记录第一数据库的元数据,且目标系统表中的标识符以及字段值均符合第二数据库的标识符规则。

[0096] 由此,根据本申请实施例的数据访问处理装置,在LightDB数据库的内核系统中,获取按照Oracle数据库的标识符规则编写的数据库访问语句,基于Oracle数据库的标识符规则对数据库访问语句进行解析,生成与数据库访问语句对应的目标语法树,并预先在LightDB数据库的内核系统中创建记录LightDB数据库的元数据的目标系统表,且目标系统表中的标识符以及字段值均符合Oracle数据库的标识符规则。基于此,对LightDB数据库的内核系统进行改进,具体的,通过对LightDB数据库的内核系统中语法解析以及系统表的改进,使LightDB数据库呈现Oracle模式,进而使得LightDB数据库兼容Oracle数据库,能够解决Oracle标识符的大小写问题,所以无需再根据数据库的特性重新维护业务代码,从而提高了数据库的开发和维护效率。

[0097] 一种可能的实施方式中,上述解析模块902具体用于:

基于第二数据库的标识符规则对数据库访问语句进行词法解析,得到数据库访问语句的词法解析结果;

对词法解析结果进行语法解析,得到目标语法树。

[0098] 一种可能的实施方式中,上述解析模块902还用于:

确定数据库访问语句的正则表达式;

根据数据库访问语句的正则表达式以及第二数据库的标识符规则,得到数据库访问语句的词法解析结果。

[0099] 一种可能的实施方式中,上述第二数据库为Oracle数据库,上述解析模块902还用于:

若数据库访问语句中目标标识符的正则表达式与含双引号标识符对应的正则规则匹配,则将目标标识符中除双引号之外的字符作为目标标识符的词法解析结果,其中,目标标识符为数据库访问语句中的任一标识符;

若目标标识符的正则表达式与不含双引号标识符对应的正则规则匹配,则将目标标识符中的字符转换为大写字母,得到目标标识符的词法解析结果。

[0100] 一种可能的实施方式中,上述确定模块903具体用于:

从目标系统表中查找目标语法树中的各标识符,得到标识符查找结果;

根据标识符查找结果,确定数据库访问语句的访问结果。

[0101] 一种可能的实施方式中,上述确定模块903还用于:

若标识符查找结果指示目标系统表中不存在目标语法树中的至少一个标识符,则确定访问结果为访问失败;

否则,根据目标语法树,访问第一数据库中的至少一个数据表,得到访问结果。

[0102] 一种可能的实施方式中,上述数据访问处理装置900还包括:

实例化模块,用于基于第二数据库的标识符规则,进行第一数据库的实例化。

[0103] 一种可能的实施方式中,上述实例化模块具体用于:

基于第二数据库的标识符规则,生成目标系统表;

基于第二数据库的标识符规则,生成第一数据库的初始化脚本;

基于目标系统表运行初始化脚本。

[0104] 一种可能的实施方式中,上述实例化模块还用于:

在第二数据库为Oracle数据库时,将初始系统表中的各标识符中的字符转换为大写字符,得到目标系统表。

[0105] 一种可能的实施方式中,上述实例化模块还用于:

获取数据库更新语句,数据库更新语句用于在第一数据库中创建新的数据表或更新第一数据库中数据表的表名或字段名;

根据数据库更新语句,按照第二数据库的标识符规则,向目标系统表中写入数据。

[0106] 关于装置中的各模块的处理流程、以及各模块之间的交互流程的描述可以参照上述方法实施例中的相关说明,这里不再详述。

[0107] 本申请实施例还提供了一种电子设备1000,如图10所示,为本申请实施例提供的电子设备1000结构示意图,包括:处理器1001、存储器1002,可选的,还可以包括总线1003。所述存储器1002存储有所述处理器1001可执行的机器可读指令,当电子设备1000运行时,所述处理器1001与所述存储器1002之间通过总线1003通信,所述机器可读指令被所述处理器1001执行时执行如上任一项所述数据访问处理方法的步骤。

[0108] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机程序,该计算机程序被处理器运行时执行如上任一项所述数据访问处理方法的步骤。

[0109] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统 and 装置的具体工作过程,可以参考方法实施例中的对应过程,本申请中不再赘述。在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,所述模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,又例如,多个模块或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口,装置或模块的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0110] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0111] 以上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,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。

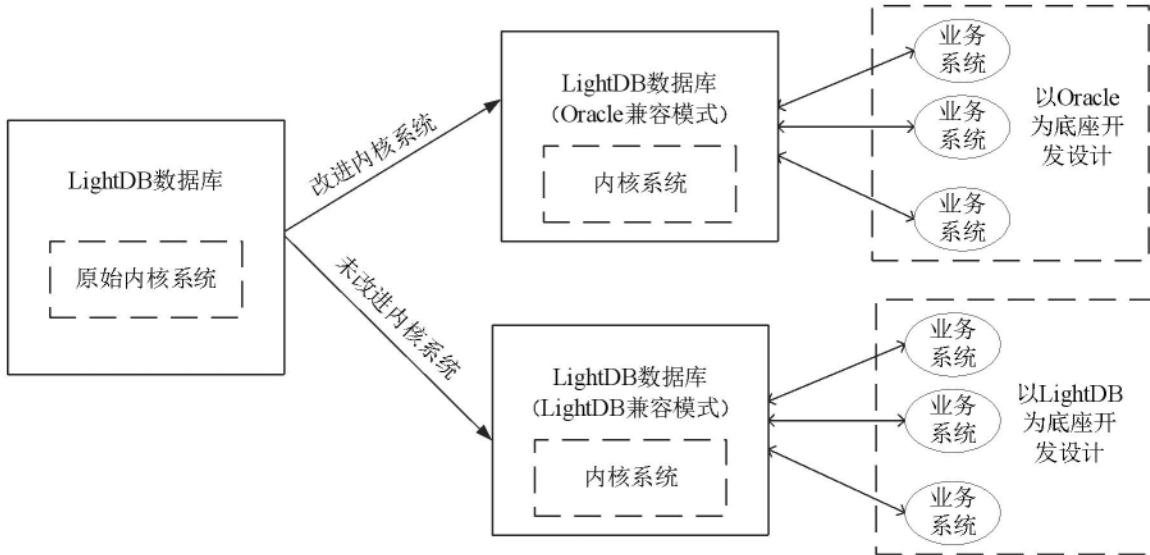


图1

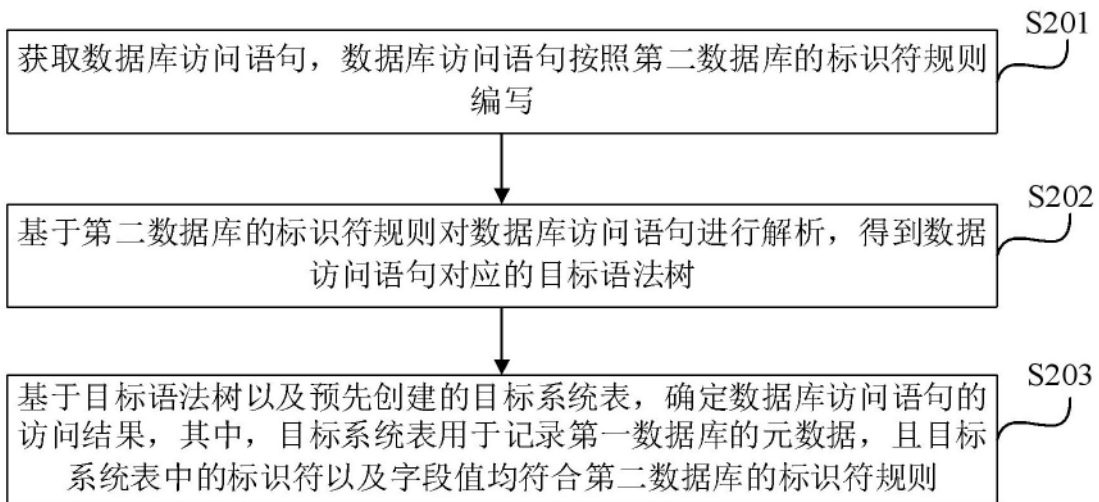


图2

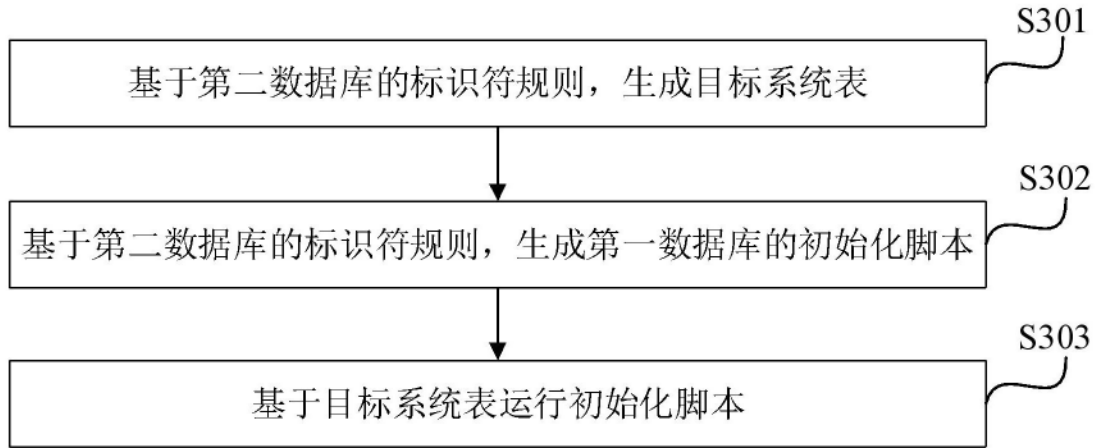


图3

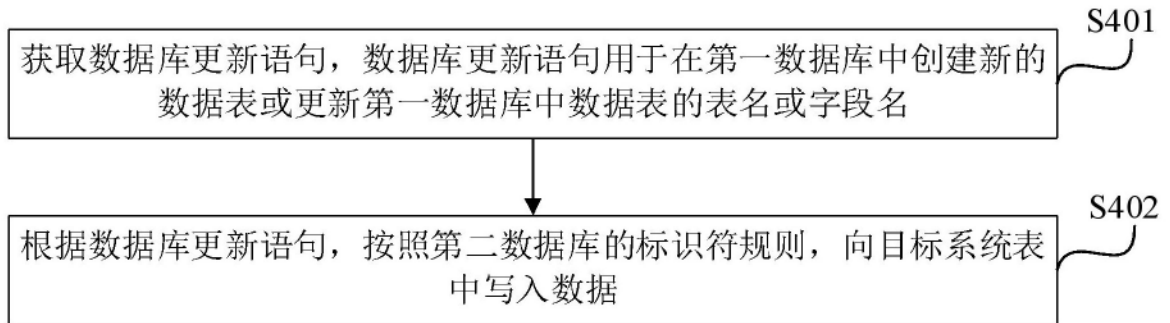


图4

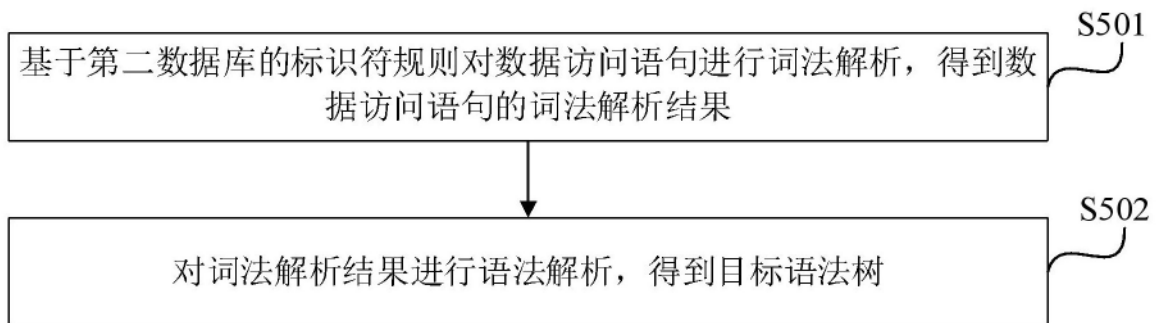


图5

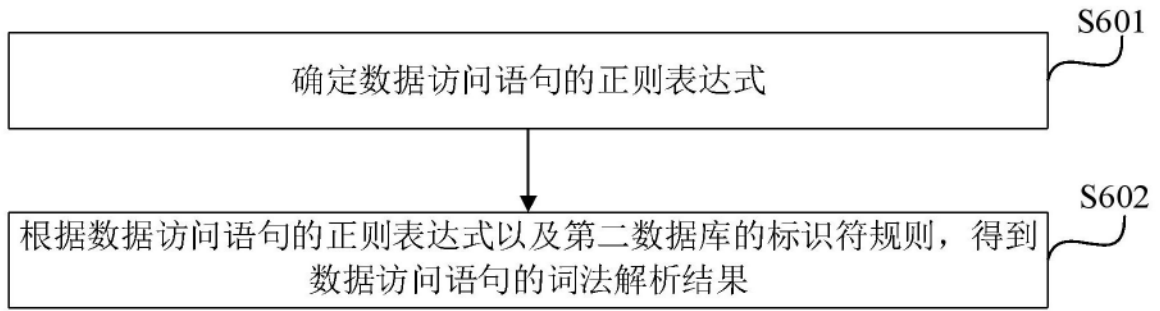


图6

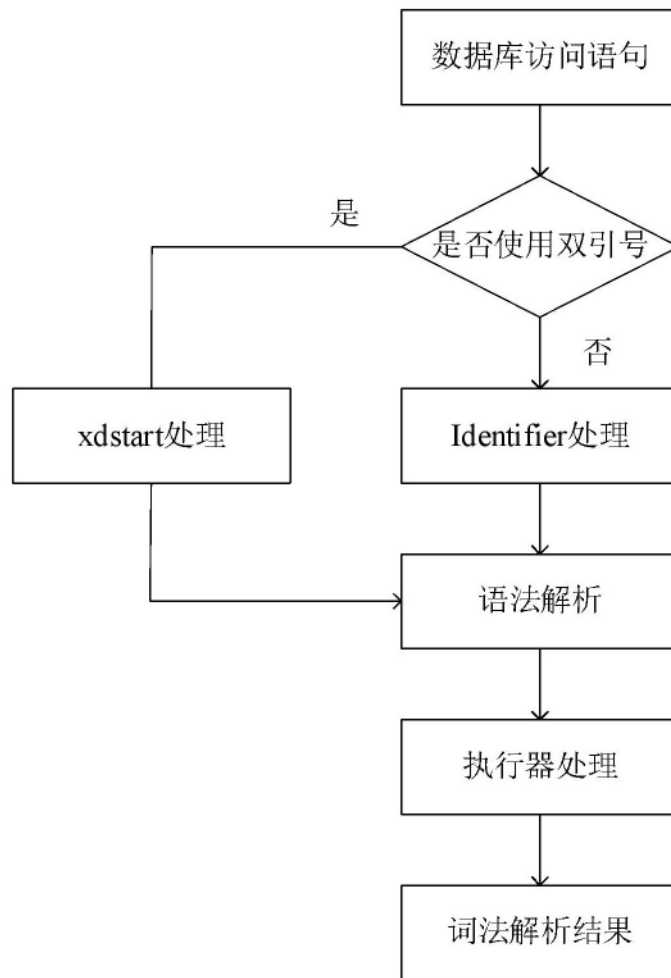


图7

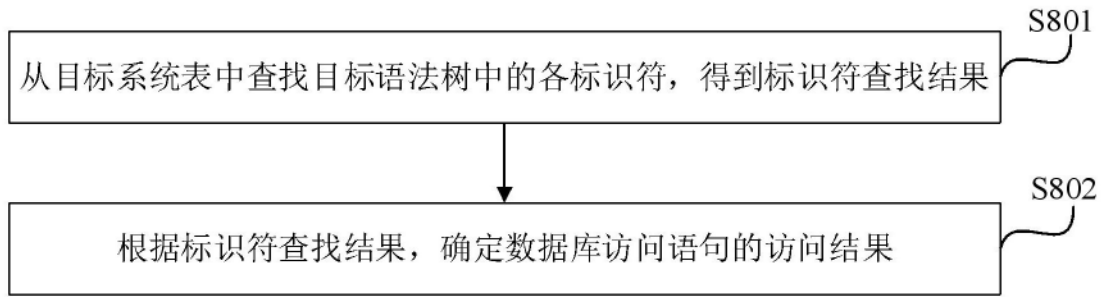


图8

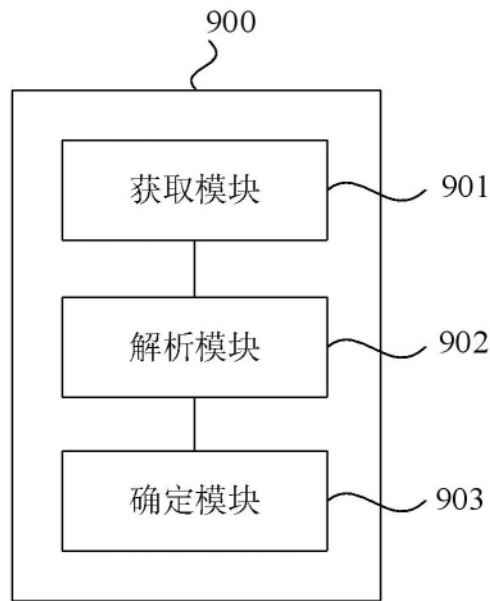


图9

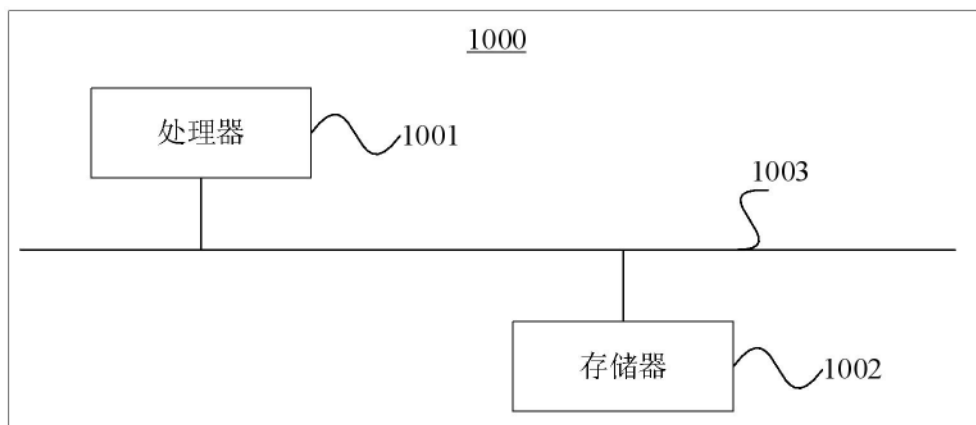


图10