



(12) 发明专利申请

(10) 申请公布号 CN 113190236 A

(43) 申请公布日 2021.07.30

(21) 申请号 202110585750.X

(22) 申请日 2021.05.27

(71) 申请人 中国工商银行股份有限公司

地址 100140 北京市西城区复兴门内大街
55号

(72) 发明人 王泽普 张世瑛 赵吉昆 梁晔华

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

代理人 任默闻 王涛

(51) Int. Cl.

G06F 8/41 (2018.01)

G06F 16/242 (2019.01)

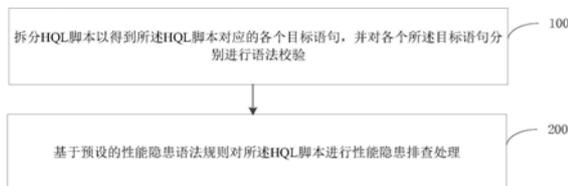
权利要求书2页 说明书15页 附图11页

(54) 发明名称

HQL脚本校验方法及装置

(57) 摘要

本申请实施例提供一种HQL脚本校验方法及装置,可用于大数据技术领域,方法包括:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;以及,基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。本申请能够有效提高HQL脚本校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,进而能够提高HQL脚本的应用可靠性并降低人力成本。



1. 一种HQL脚本校验方法,其特征在于,包括:
拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;
以及,基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。
2. 根据权利要求1所述的HQL脚本校验方法,其特征在于,所述对各个所述目标语句分别进行语法校验,包括:
基于预设的语句加工规则将各个所述目标语句分别转化为执行计划语句;
根据各个所述目标语句各自对应在分布式集群中的执行路径,在该分布式集群中对各个所述执行计划语句分别进行语法校验,以得到所述HQL脚本对应的语法校验结果。
3. 根据权利要求2所述的HQL脚本校验方法,其特征在于,所述在该分布式集群中对各个所述执行计划语句分别进行语法校验,以得到所述HQL脚本对应的语法校验结果,包括:
基于各个所述目标语句各自对应的用户登陆信息登陆所述分布式集群;
将各个所述执行计划语句提交至所述分布式集群进行语法校验,以得到该分布式集群生成的所述HQL脚本对应的语法校验结果。
4. 根据权利要求1所述的HQL脚本校验方法,其特征在于,所述性能隐患排查语法规则包括:全表扫描和全表插入规则;
所述基于预设的性能隐患排查语法规则对所述HQL脚本进行性能隐患排查处理,包括:
基于所述全表扫描和全表插入规则对属于分区表且未做分区字段筛选的HQL脚本进行全表扫描,以得到对应的全表扫描结果;
生成包含有所述全表扫描结果的性能隐患排查结果。
5. 根据权利要求1所述的HQL脚本校验方法,其特征在于,所述性能隐患排查语法规则包括:笛卡尔积查询规则;
所述基于预设的性能隐患排查语法规则对所述HQL脚本进行性能隐患排查处理,包括:
基于所述笛卡尔积查询规则对所述HQL脚本进行笛卡尔积查询,以得到对应的笛卡尔积查询结果;
生成包含有所述笛卡尔积查询结果的性能隐患排查结果。
6. 根据权利要求1所述的HQL脚本校验方法,其特征在于,所述性能隐患排查语法规则包括:关键字检索规则;
所述基于预设的性能隐患排查语法规则对所述HQL脚本进行性能隐患排查处理,包括:
基于所述关键字检索规则对所述HQL脚本进行性能隐患排查关键字检索,以得到对应的性能隐患排查关键字检索结果;
生成包含有所述性能隐患排查关键字检索结果的性能隐患排查结果。
7. 根据权利要求1至6任一项所述的HQL脚本校验方法,其特征在于,还包括:
输出经所述语法校验得到的所述HQL脚本的语法校验结果和经所述性能隐患排查处理得到的所述HQL脚本的性能隐患排查结果。
8. 一种HQL脚本校验装置,其特征在于,包括:
语法校验模块,用于拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;
性能校验模块,用于基于预设的性能隐患排查语法规则对所述HQL脚本进行性能隐患排查

处理。

9. 一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现权利要求1至7任一项所述的HQL脚本校验方法。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行时实现权利要求1至7任一项所述的HQL脚本校验方法。

HQL脚本校验方法及装置

技术领域

[0001] 本申请涉及数据处理技术领域,特别涉及大数据技术领域,具体涉及HQL脚本校验方法及装置。

背景技术

[0002] 企业的大数据服务云平台为各类专业大数据分析应用提供数据接入、存储、计算、安全管理、资源管理等服务,随着平台构建的不断完善,其技术体系日益庞大,承载的业务功能日益丰富,在平台运维体系不断扩大的同时,各大企业将大数据视作战略资源,Hadoop生态系统的应用范围愈发广泛,越来越多的分析人员基于Hive组件的HQL语句进行相关专业领域的灵活查询、分析建模,不仅HQL语句语法的正确性成为相关分析人员需要重点关注的技术问题,各个应用场景的具体模型和逻辑在新技术架构下的性能表现逐渐成为产品研发流程关注的焦点之一。

[0003] 目前,对HQL脚本进行校验的具体方式为:各开发人员需要通过HQL脚本对数据湖或数据仓库中的基础数据或主题数据进行业务逻辑加工等关联操作,由于编写HQL脚本的开发人员编码风格各异,同时对于HQL语法的掌握水平参差不齐,在各项目开发过程中会掺杂各种低效写法导致作业运行效率缓慢,同时消耗大数据分布式集群资源巨大,且无法保证校验的准确性。

发明内容

[0004] 针对现有技术中的问题,本申请提供一种HQL脚本校验方法及装置,能够有效提高HQL脚本校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,进而能够提高HQL脚本的应用可靠性并降低人力成本。

[0005] 为解决上述技术问题,本申请提供以下技术方案:

[0006] 第一方面,本申请提供一种HQL脚本校验方法,包括:

[0007] 拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;

[0008] 以及,基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。

[0009] 进一步地,所述对各个所述目标语句分别进行语法校验,包括:

[0010] 基于预设的语句加工规则将各个所述目标语句分别转化为执行计划语句;

[0011] 根据各个所述目标语句各自对应分布在分布式集群中的执行路径,在该分布式集群中对各个所述执行计划语句分别进行语法校验,以得到所述HQL脚本对应的语法校验结果。

[0012] 进一步地,所述在该分布式集群中对各个所述执行计划语句分别进行语法校验,以得到所述HQL脚本对应的语法校验结果,包括:

[0013] 基于各个所述目标语句各自对应的用户登陆信息登陆所述分布式集群;

[0014] 将各个所述执行计划语句提交至所述分布式集群进行语法校验,以得到该分布式集群生成的所述HQL脚本对应的语法校验结果。

- [0015] 进一步地,所述性能隐患语法规则包括:全表扫描和全表插入规则;
- [0016] 所述基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,包括:
- [0017] 基于所述全表扫描和全表插入规则对属于分区表且未做分区字段筛选的HQL脚本进行全表扫描,以得到对应的全表扫描结果;
- [0018] 生成包含有所述全表扫描结果的性能隐患排查结果。
- [0019] 进一步地,所述性能隐患语法规则包括:笛卡尔积查询规则;
- [0020] 所述基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,包括:
- [0021] 基于所述笛卡尔积查询规则对所述HQL脚本进行笛卡尔积查询,以得到对应的笛卡尔积查询结果;
- [0022] 生成包含有所述笛卡尔积查询结果的性能隐患排查结果。
- [0023] 进一步地,所述性能隐患语法规则包括:关键字检索规则;
- [0024] 所述基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,包括:
- [0025] 基于所述关键字检索规则对所述HQL脚本进行性能隐患关键字检索,以得到对应的性能隐患关键字检索结果;
- [0026] 生成包含有所述性能隐患关键字检索结果的性能隐患排查结果。
- [0027] 进一步地,还包括:
- [0028] 输出经所述语法校验得到的所述HQL脚本的语法校验结果和经所述性能隐患排查处理得到的所述HQL脚本的性能隐患排查结果。
- [0029] 第二方面,本申请提供一种HQL脚本校验装置,包括:
- [0030] 语法校验模块,用于拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;
- [0031] 性能校验模块,用于基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。
- [0032] 第三方面,本申请提供一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现所述的HQL脚本校验方法。
- [0033] 第四方面,本申请提供一种计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现所述的HQL脚本校验方法。
- [0034] 由上述技术方案可知,本申请提供一种HQL脚本校验方法及装置,方法包括:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验;以及,基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理;通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现对HQL脚本的自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现对HQL脚本的自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

附图说明

[0035] 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,显而易见地,下面描述中的附图是本申请的一些实施例,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。

[0036] 图1是本申请实施例中的HQL脚本校验装置分别与客户端设备及大数据服务云等之间的交互示意图。

[0037] 图2是本申请实施例中的HQL脚本校验方法的第一种流程示意图。

[0038] 图3是本申请实施例中的HQL脚本校验方法的第二种流程示意图。

[0039] 图4是本申请实施例中的HQL脚本校验方法的第三种流程示意图。

[0040] 图5是本申请实施例中的HQL脚本校验方法的第四种流程示意图。

[0041] 图6是本申请实施例中的HQL脚本校验方法的第五种流程示意图。

[0042] 图7是本申请实施例中的HQL脚本校验方法的第六种流程示意图。

[0043] 图8是本申请实施例中的HQL脚本校验方法的第七种流程示意图。

[0044] 图9是本申请实施例中的HQL脚本校验装置的结构示意图。

[0045] 图10是本申请应用实例提供的大数据性能容量扫描装置的功能示意图。

[0046] 图11是本申请应用实例提供的HQL语法校验装置的校验接口校验流程示意图。

[0047] 图12是本申请应用实例提供的全表扫描和全表插入规则的应用流程示意图。

[0048] 图13是本申请应用实例提供的笛卡尔积查询规则的应用流程示意图。

[0049] 图14是本申请应用实例提供的关键字检索规则的应用流程示意图。

[0050] 图15是本申请实施例中的电子设备的结构示意图。

具体实施方式

[0051] 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、完整的描述,显然,所描述的实施例是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,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,都属于本申请保护的范围。

[0052] 需要说明的是,本申请公开的HQL脚本校验方法及装置可用于大数据技术领域,也可用于除大数据技术领域之外的任意领域,本申请公开的HQL脚本校验方法及装置的应用领域不做限定。

[0053] 以某商业银行为例,接入大数据服务云的离线批处理业务系统已达数百个。在每期基于大数据服务云的项目研发上线、迭代优化过程中,各开发人员需要通过hive SQL(以下简称HQL)脚本对数据湖或数据仓库中的基础数据或主题数据进行业务逻辑加工等关联操作。由于编写HQL脚本的开发人员编码风格各异,同时对于HQL语法的掌握水平参差不齐,在各项目开发过程中会掺杂各种低效写法导致作业运行效率缓慢,同时消耗大数据分布式集群资源巨大,该类离线批处理脚本存在着较大优化空间。而对于复杂逻辑的HQL脚本实现或基于大表(TB级别)的基础数据关联HQL脚本实现则更需要注重性能隐患,因此,对HQL脚本中存在的技术问题进行检查并反馈整改的重要性不言而喻。

[0054] 基于上述内容,本申请提供了语法校验和性能容量静态扫描方式。一是通过利用

分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。二是制定HQL常见性能隐患语法规则,同时对各接入大数据服务服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本发明对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0055] 针对现有的HQL脚本校验方式存在效率低、准确性差且人工成本高等问题,本申请实施例分别提供一种HQL脚本校验方法、HQL脚本校验装置、电子设备计算机可读存储介质,通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现对HQL脚本的自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现对HQL脚本的自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

[0056] 在本申请的一个或多个实施例中,分布式集群可以指大数据服务云或HBase集群等,其中,HBase集群(Hadoop Database)是指高可靠性、高性能、面向列、可伸缩的分布式存储系统,由主节点Master和从节点Region Server组成,其中,Master也可以专门写作HMaster,Region Server也可以写作HRegionServer或者RegionServer等。

[0057] 基于上述内容,本申请还提供一种用于实现本申请一个或多个实施例中提供的HQL脚本校验方法的HQL脚本校验装置,参见图1,该HQL脚本校验装置可以自行或通过第三方服务器等与各个用户持有的客户端设备、大数据服务云以及接入大数据服务云的各个离线批处理业务系统之间通信连接,HQL脚本校验装置可以为一种服务器,自客户端设备或者各个离线批处理业务系统接收针对大数据服务云的HQL脚本校验请求或者待校验的各个HQL脚本,也可以自客户端设备、第三方数据库或本地获取相关的由用户预先设置各类规则配置文件,例如本申请的一个或多个实施例中提及的语句加工规则、全表扫描和全表插入规则、笛卡尔积查询规则及关键字检索规则等中的至少一个。所述HQL脚本校验装置在得到经所述语法校验得到的所述HQL脚本的语法校验结果和经所述性能隐患排查处理得到的所述HQL脚本的性能隐患排查结果之后,还可以将所述语法校验结果和性能隐患排查结果发送至所述客户端设备进行显示或发送至离线批处理业务系统等,以使用户及时获知所述HQL脚本在大数据服务云的校验结果等。

[0058] 可以理解的是,所述客户端设备可以包括智能手机、平板电子设备、网络机顶盒、便携式计算机、台式电脑、个人数字助理(PDA)、车载设备、智能穿戴设备等。其中,所述智能穿戴设备可以包括智能眼镜、智能手表、智能手环等。

[0059] 上述的客户端设备可以具有通信模块(即通信单元),可以与远程的服务器进行通信连接,实现与所述服务器的数据传输。所述服务器可以包括任务调度中心一侧的服务器,其他的实施场景中也可以包括中间平台的服务器,例如与任务调度中心服务器有通信链接的第三方服务器平台的服务器。所述的服务器可以包括单台计算机设备,也可以包括多个服务器组成的服务器集群,或者分布式装置的服务器结构。

[0060] 上述服务器与所述客户端设备之间可以使用任何合适的网络协议进行通信,包括在本申请提交日尚未开发出的网络协议。所述网络协议例如可以包括TCP/IP协议、UDP/IP协议、HTTP协议、HTTPS协议等。当然,所述网络协议例如还可以包括在上述协议之上使用的RPC协议(Remote Procedure Call Protocol,远程过程调用协议)、REST协议(Representational State Transfer,表述性状态转移协议)等。

[0061] 具体通过下述各个实施例及应用实例分别进行详细说明。

[0062] 为了解决现有的HQL脚本校验方式存在效率低、准确性差且人工成本高等问题,本申请提供一种HQL脚本校验方法的实施例,参见图2,由HQL脚本校验装置执行的所述HQL脚本校验方法具体包含有如下内容:

[0063] 步骤100:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验。

[0064] 在步骤100中,HQL脚本校验装置可以通过利用分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。

[0065] 步骤200:基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。

[0066] 在步骤200中,HQL脚本校验装置可以根据用户预先制定并存储的HQL常见的性能隐患语法规则,同时对各接入大数据服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本发明对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0067] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现对HQL脚本的自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现对HQL脚本的自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

[0068] 为了有效节约分布式集群资源来进行校验,在本申请提供的HQL脚本校验方法的一个实施例,参见图3,所述HQL脚本校验方法中的步骤100具体包含有如下内容:

[0069] 步骤110:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句。

[0070] 步骤120:基于预设的语句加工规则将各个所述目标语句分别转化为执行计划语句。

[0071] 可以理解的是,所述语句加工规则是指在各个所述目标预计之前加入如EXPLAIN等字样,当分布式集群到EXPLAIN字样后,执行该语句时只执行、不下推,可以有效节约集群资源来进行校验。

[0072] 步骤130:根据各个所述目标语句各自对应分布在分布式集群中的执行路径,在该分布式集群中对各个所述执行计划语句分别进行语法校验,以得到所述HQL脚本对应的语法校验结果。

[0073] 具体来说,可以预先采用Python语言开发,通过输入各待检验HQL脚本的路径对相应路径下的HQL脚本先进行语法校验,查看是否有语法错误,否则返回给开发错误信息。

[0074] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过基于预设的语句加工规则将各个所述目标语句分别转化为执行计划语句,能够实现各个目标语句在分布式集群中只执行、不下推,可以有效节约分布式集群资源来进行校验,并能够有效提高应用分布式集群进行语法校验的可靠性及稳定性。

[0075] 为了提高应用分布式集群进行语法校验的安全性,在本申请提供的HQL脚本校验方法的一个实施例,参见图4,所述HQL脚本校验方法中的步骤130具体包含有如下内容:

[0076] 步骤131:根据各个所述目标语句各自对应分布在分布式集群中的执行路径,基于各个所述目标语句各自对应的用户登陆信息登陆所述分布式集群。

[0077] 步骤132:将各个所述执行计划语句提交至所述分布式集群进行语法校验,以得到该分布式集群生成的所述HQL脚本对应的语法校验结果。

[0078] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过基于各个所述目标语句各自对应的用户登陆信息登陆所述分布式集群并提交执行计划语句进行语法校验,能够有效提高应用分布式集群进行语法校验的安全性及可靠性。

[0079] 为了进一步提高对所述HQL脚本进行性能隐患排查处理的准确性,在本申请提供的HQL脚本校验方法的一个实施例,所述性能隐患排查语法规则包括:全表扫描和全表插入规则;参见图5,所述HQL脚本校验方法中的步骤200具体包含有如下内容:

[0080] 步骤211:基于所述全表扫描和全表插入规则对属于分区表且未做分区字段筛选的HQL脚本进行全表扫描,以得到对应的全表扫描结果。

[0081] 步骤212:生成包含有所述全表扫描结果的性能隐患排查结果。

[0082] 具体来说,HQL脚本校验装置可以首先添加规则配置,提供配置分区表查询服务信息,供检查规则使用服务信息查询分区表信息,其中的分区表信息包括分区表名称、分区字段名称;然后使用SQL检查服务框架,解析出SQL语法树,遍历该语法树查找数据表(SQLIdentifier)。由于是静态检查规则,无法判断标识符是否为表或视图,因此对此不作检查全部作为表处理。通过分区表查询服务查询表是否为分区表及分区字段,若非分区表则对当前表不做全表扫描检查,继续处理下一个表。对于SQL查询目标为分区表,若对该表未做分区字段筛选(ON条件及WHERE条件中),则判定语句全表扫描,返回检查结果命中。对于子查询嵌套的SQL,数据库在执行时会做执行计划优化,如果外层查询存在对分区字段的查询,判定无全表扫描,返回未命中。

[0083] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过基于所述全表扫描和全表插入规则对属于分区表且未做分区字段筛选的HQL脚本进行全表扫描,能够有效提高对所述HQL脚本进行性能隐患排查处理的准确性、全面性及可靠性。

[0084] 为了进一步提高对所述HQL脚本进行性能隐患排查处理的准确性,在本申请提供的HQL脚本校验方法的一个实施例,所述性能隐患语法规则包括:笛卡尔积查询规则;参见图6,所述HQL脚本校验方法中的步骤200还具体包含有如下内容:

[0085] 步骤221:基于所述笛卡尔积查询规则对所述HQL脚本进行笛卡尔积查询,以得到对应的笛卡尔积查询结果。

[0086] 步骤222:生成包含有所述笛卡尔积查询结果的性能隐患排查结果。

[0087] 具体来说,HQL脚本校验装置可以添加规则配置,提供配置小数据量列表(简称小表),供检查规则读取(白名单功能);然后使用SQL检查服务框架,解析出SQL语法树,由根节点向叶子节点遍历语法树中SQLJoin节点。对于JOIN运算两侧操作子均为标识符SQLIdentifier的情况,检查两个表(标识符)是否为参数表中配置的小表。对于JOIN运算两侧操作子存在非标识符的情况,分别检查左右两个操作子,判断其查询结果大小情况。对于UNION操作,基于预设的运算规则划分大结果集或小结果集;针对其他INTERSECT等非扩大运算,其结果均认为是小结果集,另外,大表笛卡尔积判断还需要预设的特殊情形。

[0088] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过基于所述笛卡尔积查询对所述HQL脚本进行笛卡尔积查询,以得到对应的笛卡尔积查询结果,能够有效提高对所述HQL脚本进行性能隐患排查处理的准确性、有效性及可靠性。

[0089] 为了进一步提高对所述HQL脚本进行性能隐患排查处理的准确性,在本申请提供的HQL脚本校验方法的一个实施例,所述性能隐患语法规则包括:关键字检索规则;参见图7,所述HQL脚本校验方法中的步骤200还具体包含有如下内容:

[0090] 步骤231:基于所述关键字检索规则对所述HQL脚本进行性能隐患关键字检索,以得到对应的性能隐患关键字检索结果。

[0091] 步骤232:生成包含有所述性能隐患关键字检索结果的性能隐患排查结果。

[0092] 具体来说,HQL脚本校验装置可以在数据库中添加关键字规则配置,供检查规则读取,然后使用SQL检查服务框架,解析出SQL语法树,由根节点向叶子节点遍历语法树中SQLJoin节点;如果出现select*、regexp、regexp_extract、regexp_replace、rlike、orde及count(distinct)by等,均可以判断成有性能隐患。

[0093] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过基于所述关键字检索规则对所述HQL脚本进行性能隐患关键字检索,以得到对应的性能隐患关键字检索结果,能够有效提高对所述HQL脚本进行性能隐患排查处理的准确性、有效性及可靠性。

[0094] 为了提高用户获知HQL脚本校验结果的效率及便捷性,在本申请提供的HQL脚本校验方法的一个实施例,参见图8,所述HQL脚本校验方法中还具体包含有如下内容:

[0095] 步骤300:输出经所述语法校验得到的所述HQL脚本的语法校验结果和经所述性能隐患排查处理得到的所述HQL脚本的性能隐患排查结果。

[0096] 具体来说,HQL脚本校验装置可以由Python语言开发,通过输入各待检验HQL脚本的路径对相应路径下的HQL脚本先进行语法校验,查看是否有语法错误,否则则返回给开发错误信息,然后进行自动化静态扫描,结合既定规则的对比判断出脚本是否存在性能隐患

语法。若存在相应的写法,则将脚本名称、语法校验结果、HQL语句段落、性能隐患语法三部分内容输出至最终的汇总文件中。

[0097] 从上述描述可知,本申请实施例提供的HQL脚本校验方法,通过输出经所述语法校验得到的所述HQL脚本的语法校验结果和经所述性能隐患排查处理得到的所述HQL脚本的性能隐患排查结果,能够有效提高用户获知HQL脚本校验结果的效率及便捷性,以有效提高用户体验。

[0098] 从软件层面来说,为了解决现有的HQL脚本校验方式存在效率低、准确性差且人工成本高等问题,本申请提供一种用于执行所述HQL脚本校验方法中全部或部分内容的HQL脚本校验装置的实施例,参见图9,所述HQL脚本校验装置具体包含有如下内容:

[0099] 语法校验模块10,用于拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验。

[0100] 在语法校验模块10中,HQL脚本校验装置可以通过利用分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。

[0101] 性能校验模块20,用于基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。

[0102] 在性能校验模块20中,HQL脚本校验装置可以根据用户预先制定并存储的HQL常见的性能隐患语法规则,同时对各接入大数据服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本发明对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0103] 本申请提供的HQL脚本校验装置的实施例具体可以用于执行上述实施例中的HQL脚本校验方法的实施例的处理流程,其功能在此不再赘述,可以参照上述方法实施例的详细描述。

[0104] 从上述描述可知,本申请实施例提供的HQL脚本校验装置,通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

[0105] 为了进一步说明书本方案,本申请还提供一种应用大数据性能容量扫描装置实现的具有语法校验和性能容量静态扫描功能的HQL脚本校验方法的具体应用实例,涉及大数据平台离线批处理技术领域,本申请应用实例通过两个技术方案实现了语法校验和性能容

量静态扫描方法。一是通过利用分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。二是制定HQL常见性能隐患语法规则,同时对各接入大数据服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本申请应用实例对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0106] 1. 整体流程:

[0107] 大数据性能容量扫描装置由Python语言开发,通过输入各待检验HQL脚本的路径对相应路径下的HQL脚本先进行语法校验,查看是否有语法错误,否则则返回给开发错误信息,然后进行自动化静态扫描,结合既定规则的对比判断出脚本是否存在性能隐患语法。若存在相应的写法,则将脚本名称、语法校验结果、HQL语句段落、性能隐患语法三部分输出至最终的汇总文件中。参见图10,大数据性能容量扫描装置的功能包含有:接入大数据服务云的业务逻辑加工hql脚本、规则制定、脚本拆分、脚本解析、规则判定及输出脚本扫描结果。

[0108] 2. 规则简述:

[0109] 在大数据性能容量扫描装置中,最重要的一部分为高危性能容量语法判定规则的设定,在综合考虑实际脚本内容和性能提升之后,该装置制定了以下3大类规则:

[0110] (1) 全表扫描和全表插入规则;

[0111] (2) 笛卡尔积查询;

[0112] (3) 关键字检索(使用了低效的关键语句);

[0113] 1) SELECT*或SELECT t.*;

[0114] 2) 使用order by进行排序;

[0115] 3) 使用COUNT (DISTINCT) 操作;

[0116] 4) union all部分个数大于2;

[0117] 5) 使用collect_list,collect_set函数。

[0118] 3. 技术实现:

[0119] (1) 将基于执行计划的HQL语法校验装置

[0120] 参见图11,HQL语法校验装置的校验接口校验流程主要通过如下四个步骤实现:

[0121] 1) 应用侧发送待校验的语句至URL;

[0122] 2) 将URL获取到的语句进行加工成执行计划语句;

[0123] 3) 根据语句的标识去数据库获取用户登陆信息;

[0124] 4) 使用该用户信息登陆集群提交语句;

[0125] 5) 获取到集群返回的语句校验信息;

[0126] 其中执行计划语句相对于普通SQL语句前有EXPLAIN字样,集群读到explain字样后,执行该语句时只执行不下推,可以有效节约集群资源来进行校验。

[0127] 装置的URL为所在服务器的IP,用户连接URL提交语句后,装置的TOMCAT组件会捕捉语句,并将其记录在数据库中,待后续校验的进行。

[0128] (2) 性能静态校验

[0129] 利用Python代码将HQL脚本以分号为间隔拆分为多个HQL语句,依次检查每一句HQL语句是否存在性能问题:通过调用Python中SQLparse等库对HQL脚本进行拆分,根据数据类型提取出相关信息(SQLparse会将HQL语句拆分为token,每一个token都对应一个数据类型);对于脚本中是否含有规则中关键词的判断,采用逐行扫描脚本中的每一行语句,判断这行语句中是否存在规则中涉及到的关键词,并将这个判断功能封装为一个接口,后续通过调用接口的方式来实现对所有HQL脚本的分析。

[0130] 2-1) 参见图12,全表扫描和全表插入规则:

[0131] ①添加规则配置,提供配置分区表查询服务信息,供检查规则使用服务信息查询分区表信息。

[0132] 分区表信息包括分区表名称、分区字段名称。

[0133] ②使用SQL检查服务框架,解析出SQL语法树,遍历该语法树查找数据表(SQLIdentifier)。

[0134] 由于是静态检查规则,无法判断标识符是否为表或视图,因此对此不作检查全部作为表处理。

[0135] ③通过分区表查询服务查询表是否为分区表及分区字段,若非分区表则对当前表不做全表扫描检查,继续处理下一个表。

[0136] ④对于SQL查询目标为分区表,若对该表未做分区字段筛选(ON条件及WHERE条件中),则判定语句全表扫描,返回检查结果命中。

[0137] ⑤对于子查询嵌套的SQL,数据库在执行时会做执行计划优化,如果外层查询存在对分区字段的查询,判定无全表扫描,返回未命中。

[0138] 2-2) 参见图13,笛卡尔积查询:

[0139] ①添加规则配置,提供配置小数据量列表(简称小表),供检查规则读取(白名单功能)。

[0140] ②使用SQL检查服务框架,解析出SQL语法树,由根节点向叶子节点遍历语法树中SQLJoin节点。

[0141] ③对于JOIN运算两侧操作子均为标识符SQLIdentifier的情况,检查两个表(标识符)是否为参数表中配置的小表。

[0142] 若两个表均不是小表,则认为可能是两个大表做笛卡尔积运算,继续判断两个表是否各自都有ON条件做筛选。

[0143] 若上述两表各自都没有ON条件做筛选,则判定语句中存在两个大表做笛卡尔积运算,直接抛出规则命中异常,返回检查结果为命中。

[0144] 其余情况认为无规则命中。

[0145] 由于是静态检查规则,无法判断标识符是否为表或视图,因此对此不作检查全部作为表处理。

[0146] ④对于JOIN运算两侧操作子存在非标识符的情况,分别检查左右两个操作子,判断其查询结果大小情况。

[0147] 判断逻辑使用步骤②开始的逻辑进行递归判断,操作子中存在大表笛卡尔积运算则抛出命中异常。

[0148] 左右操作子中不存在JOIN运算时,如果是小表查询,则判断为小数据集,如果是大表查询且无条件筛选,则判断为大数据集。

[0149] 左右操作子为以下三种情况时,判定语句中存在大表笛卡尔积运算,抛出规则命中异常,返回检查结果为命中。

[0150] 大表JOIN大结果集=命中;

[0151] 大表JOIN大表=命中;

[0152] 大结果集JOIN大结果集=命中。

[0153] ⑤对于UNION操作,运算规则如下:

[0154] (大结果集/大表) UNION (大结果集/大表/小结果集/小表) = 大结果集。

[0155] (小结果集/小表) UNION (小结果集/小表) = 小结果集。

[0156] ⑥其他INTERSECT等非扩大运算,其结果均认为是小结果集。

[0157] ⑦大表笛卡尔积判断需要考虑但不限于以下情况:

[0158] JOIN的ON条件中出现 $1=1$, $t1.a=t1.a$,作为笛卡尔积处理。

[0159] `SELECT t1.*,t2.*FROM t1 JOIN t2 ON t1.date>='2010/01/01'AND t2.date>='2010/01/01'`,两表均有筛选,不作为笛卡尔积处理。

[0160] 2-3) 参见图14,关键字查询:

[0161] ①在数据库中添加关键字规则配置,供检查规则读取。

[0162] ②使用SQL检查服务框架,解析出SQL语法树,由根节点向叶子节点遍历语法树中SQLJoin节点。

[0163] ③如果出现`select*,regexp,regexp_extract,regexp_replace,rlike,order,count(distinct)by`等,均可以判断成有性能隐患。

[0164] 基于上述技术方案,本申请应用实例提供的基于大数据性能容量扫描装置实现HQL脚本校验方法,具备如下有益效果:

[0165] (1) 基于大数据服务云平台的离线批处理作业运行效能提升

[0166] 该装置通过对HQL脚本中的低效、高风险、错误语法进行扫描遍历,针对符合既定规则的语法做出提醒并反馈开发人员整改,整改后的批量作业性能得以调优,相关基于大数据平台的离线分析挖掘场景的应用作业运行时效得到大幅提升。

[0167] (2) 大数据分布式集群计算资源集约

[0168] 该装置通过对HQL语句中低效语法的优化规避可以极大地减少单位时间内作业对大数据分布式集群的计算资源消耗,节省无需占用的资源,使整个系统有条不紊地进行运转,提高大数据平台的整体工作效率。

[0169] (3) 脚本性能分析自动化

[0170] 通过大数据性能容量扫描装置可以根据规则筛选出不符合性能要求的脚本,并标注出不符合的具体规则种类,实现了快速精准定位问题,在后续进行脚本性能优化的时候确定优化方向。这一过程完全由程序进行自动化操作,无需人为干预,从而使脚本校验的规

则和流程更加规范化,不会出现因人为主观因素导致的对脚本的判断失误,同时大大减少了人力分析脚本所需时间,提升了分析脚本的效率。

[0171] (4) 灵活性高,支持单个、批量分析,规则库可不断丰富完善

[0172] 大数据性能容量扫描装置支持大批量脚本分析,也支持单个脚本检验,因此它不仅可以用来发现脚本中存在的性能问题,也可用于后续判断修改后脚本的性能问题是否得到解决:用大数据性能容量扫描装置分析修改后的脚本,判断之前的性能问题是否的到改善并检验是否存在新出现的性能问题,可以根据需求灵活调整输入脚本的数量,在操作上具有很强的灵活性。同时它适用于所有HQL脚本的分析,且装置中的规则可以不断进行修改、完善和添加,对后续整个装置的更新迭代设置了十分灵活的机制,能够通过对代码的微小调整实现功能的更新,是一个十分友好的可持续调整优化的装置。也可通过修改内部规则实现其他的功能,例如对其他类型语法脚本的分析,适用性较强。

[0173] 从硬件层面来说,为了解决现有的HQL脚本校验方式存在效率低、准确性差且人工成本高等问题,本申请提供一种用于实现所述HQL脚本校验方法中的全部或部分内容的电子设备的实施例,所述电子设备具体包含有如下内容:

[0174] 图15为本申请实施例的电子设备9600的系统构成的示意框图。如图15所示,该电子设备9600可以包括中央处理器9100和存储器9140;存储器9140耦合到中央处理器9100。值得注意的是,该图15是示例性的;还可以使用其他类型的结构,来补充或代替该结构,以实现电信功能或其他功能。

[0175] 在一实施例中,HQL脚本校验功能可以被集成到中央处理器中。其中,中央处理器可以被配置为进行如下控制:

[0176] 步骤100:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验。

[0177] 在步骤100中,HQL脚本校验装置可以通过利用分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。

[0178] 步骤200:基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。

[0179] 在步骤200中,HQL脚本校验装置可以根据用户预先制定并存储的HQL常见的性能隐患语法规则,同时对各接入大数据服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本发明对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0180] 从上述描述可知,本申请实施例提供的电子设备,通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现对HQL脚本的自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够

有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现对HQL脚本的自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

[0181] 在另一个实施方式中,HQL脚本校验装置可以与中央处理器9100分开配置,例如可以将HQL脚本校验装置配置为与中央处理器9100连接的芯片,通过中央处理器的控制来实现HQL脚本校验功能。

[0182] 如图15所示,该电子设备9600还可以包括:通信模块9110、输入单元9120、音频处理器9130、显示器9160、电源9170。值得注意的是,电子设备9600也并不是必须要包括图15中所示的所有部件;此外,电子设备9600还可以包括图15中没有示出的部件,可以参考现有技术。

[0183] 如图15所示,中央处理器9100有时也称为控制器或操作控件,可以包括微处理器或其他处理器装置和/或逻辑装置,该中央处理器9100接收输入并控制电子设备9600的各个部件的操作。

[0184] 其中,存储器9140,例如可以是缓存器、闪存、硬驱、可移动介质、易失性存储器、非易失性存储器或其它合适装置中的一种或更多种。可储存上述与失败有关的信息,此外还可存储执行有关信息的程序。并且中央处理器9100可执行该存储器9140存储的该程序,以实现信息存储或处理等。

[0185] 输入单元9120向中央处理器9100提供输入。该输入单元9120例如为按键或触摸输入装置。电源9170用于向电子设备9600提供电力。显示器9160用于进行图像和文字等显示对象的显示。该显示器例如可为LCD显示器,但并不限于此。

[0186] 该存储器9140可以是固态存储器,例如,只读存储器(ROM)、随机存取存储器(RAM)、SIM卡等。还可以是这样的存储器,其即使在断电时也保存信息,可被选择性地擦除且设有更多数据,该存储器的示例有时被称为EPROM等。存储器9140还可以是某种其它类型的装置。存储器9140包括缓冲存储器9141(有时被称为缓冲器)。存储器9140可以包括应用/功能存储部9142,该应用/功能存储部9142用于存储应用程序和功能程序或用于通过中央处理器9100执行电子设备9600的操作的流程。

[0187] 存储器9140还可以包括数据存储器9143,该数据存储器9143用于存储数据,例如联系人、数字数据、图片、声音和/或任何其他由电子设备使用的数据。存储器9140的驱动程序存储部9144可以包括电子设备的用于通信功能和/或用于执行电子设备的其他功能(如消息传送应用、通讯录应用等)的各种驱动程序。

[0188] 通信模块9110即为经由天线9111发送和接收信号的发送机/接收机9110。通信模块(发送机/接收机)9110耦合到中央处理器9100,以提供输入信号和接收输出信号,这可以和常规移动通信终端的情况相同。

[0189] 基于不同的通信技术,在同一电子设备中,可以设置有多个通信模块9110,如蜂窝网络模块、蓝牙模块和/或无线局域网模块等。通信模块(发送机/接收机)9110还经由音频处理器9130耦合到扬声器9131和麦克风9132,以经由扬声器9131提供音频输出,并接收来自麦克风9132的音频输入,从而实现通常的电信功能。音频处理器9130可以包括任何合适的缓冲器、解码器、放大器等。另外,音频处理器9130还耦合到中央处理器9100,从而使得可

以通过麦克风9132能够在本机上录音,且使得可以通过扬声器9131来播放本机上存储的声音。

[0190] 本申请的实施例还提供能够实现上述实施例中的HQL脚本校验方法中全部步骤的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,该计算机程序被处理器执行时实现上述实施例中的执行主体为服务器或客户端的HQL脚本校验方法的全部步骤,例如,所述处理器执行所述计算机程序时实现下述步骤:

[0191] 步骤100:拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验。

[0192] 在步骤100中,HQL脚本校验装置可以通过利用分布式技术的执行计划特性,对普通的查询语句进行再编写成执行计划语句(语句前加EXPLAIN),集群识别执行后可以详细的返回出执行语句对应的MR代码(MR过程为HQL语句在集群侧分解任务的过程),装置通过遍历输出内容中的Error、Failed、关键字找出所校验语句是否存在语法问题以及所存在的具体问题。相关数据分析人员能够在第一时间对自己的HQL语句进行语法校验而无需在将语句提交到Yarn上才能发现语法问题,为HQL语句的编写、提交集群执行动作增加了一种事前校验机制,使用HQL语句进行的相关灵活查询、分析建模工作效能得以提升。

[0193] 步骤200:基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理。

[0194] 在步骤200中,HQL脚本校验装置可以根据用户预先制定并存储的HQL常见的性能隐患语法规则,同时对各接入大数据服务云平台的业务逻辑加工HQL脚本进行拆分,而后基于既定的规则对HQL脚本进行高危语法提取,最终实现了对HQL脚本的性能容量隐患排查。同时本发明对上述方法进行自动化的技术实现,通过该自动化HQL脚本静态扫描装置克服了人工排查性能隐患的人力成本,实现了在短时间内即可全量解析HQL脚本性能隐患问题的功能。

[0195] 从上述描述可知,本申请实施例提供的计算机可读存储介质,通过拆分HQL脚本以得到所述HQL脚本对应的各个目标语句,并对各个所述目标语句分别进行语法校验,能够实现HQL脚本的自动语法校验,且能够有效提高对HQL脚本进行语法校验的准确性及有效性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度;通过基于预设的性能隐患语法规则对所述HQL脚本进行性能隐患排查处理,能够实现HQL脚本的自动性能隐患排查,进而能够有效提高HQL脚本的应用可靠性,并能够有效提高HQL脚本校验的效率、自动化程度及智能化程度,并有效降低人工排查HQL脚本性能隐患的人力成本,提高用户体验。

[0196] 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、装置、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。本发明是参照根据本发明实施例的方法、设备(装置)、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。

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。

[0197] 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。

[0198] 本发明中应用了具体实施例对本发明的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本发明的方法及其核心思想;同时,对于本领域的一般技术人员,依据本发明的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本发明的限制。

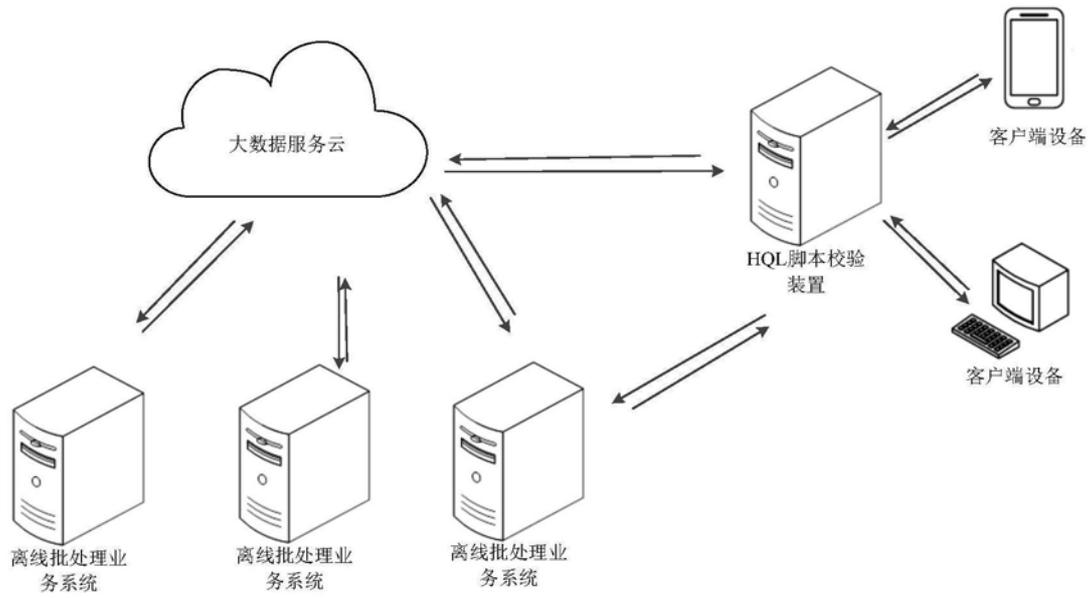


图1

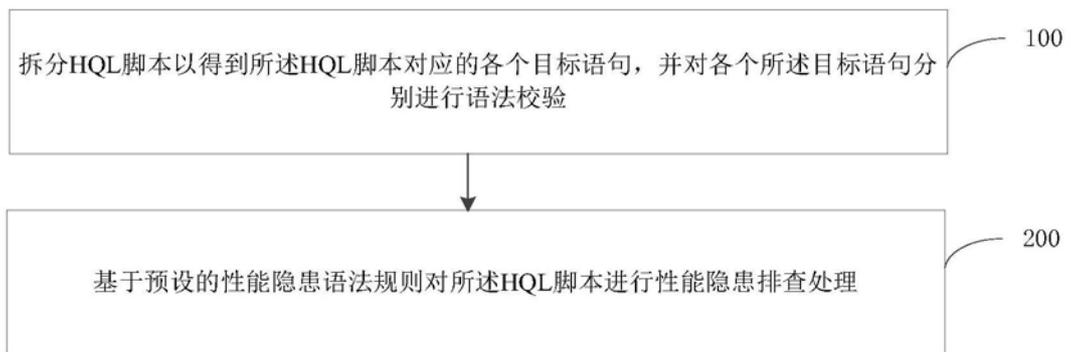


图2

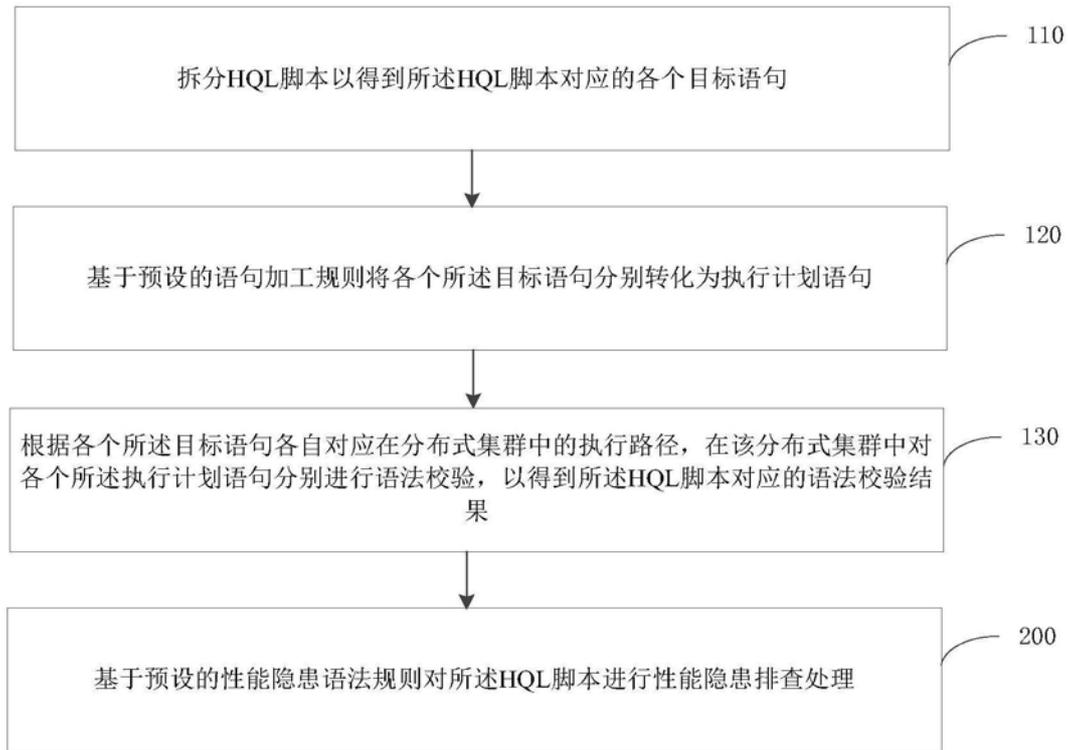


图3

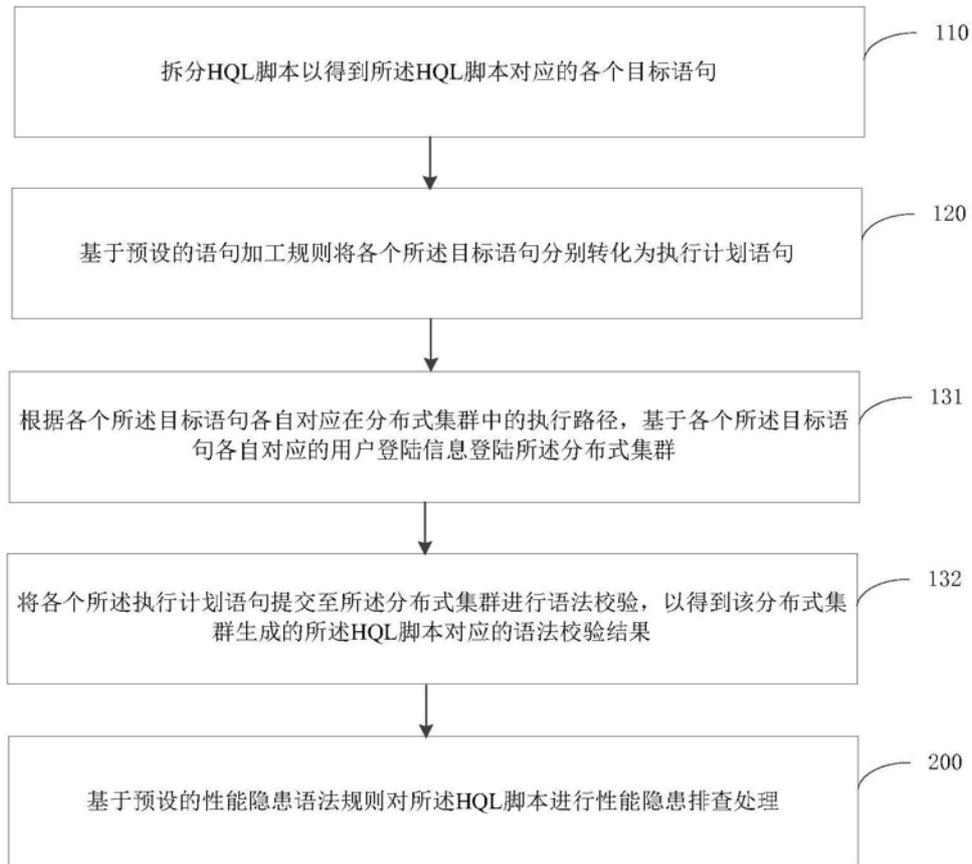


图4

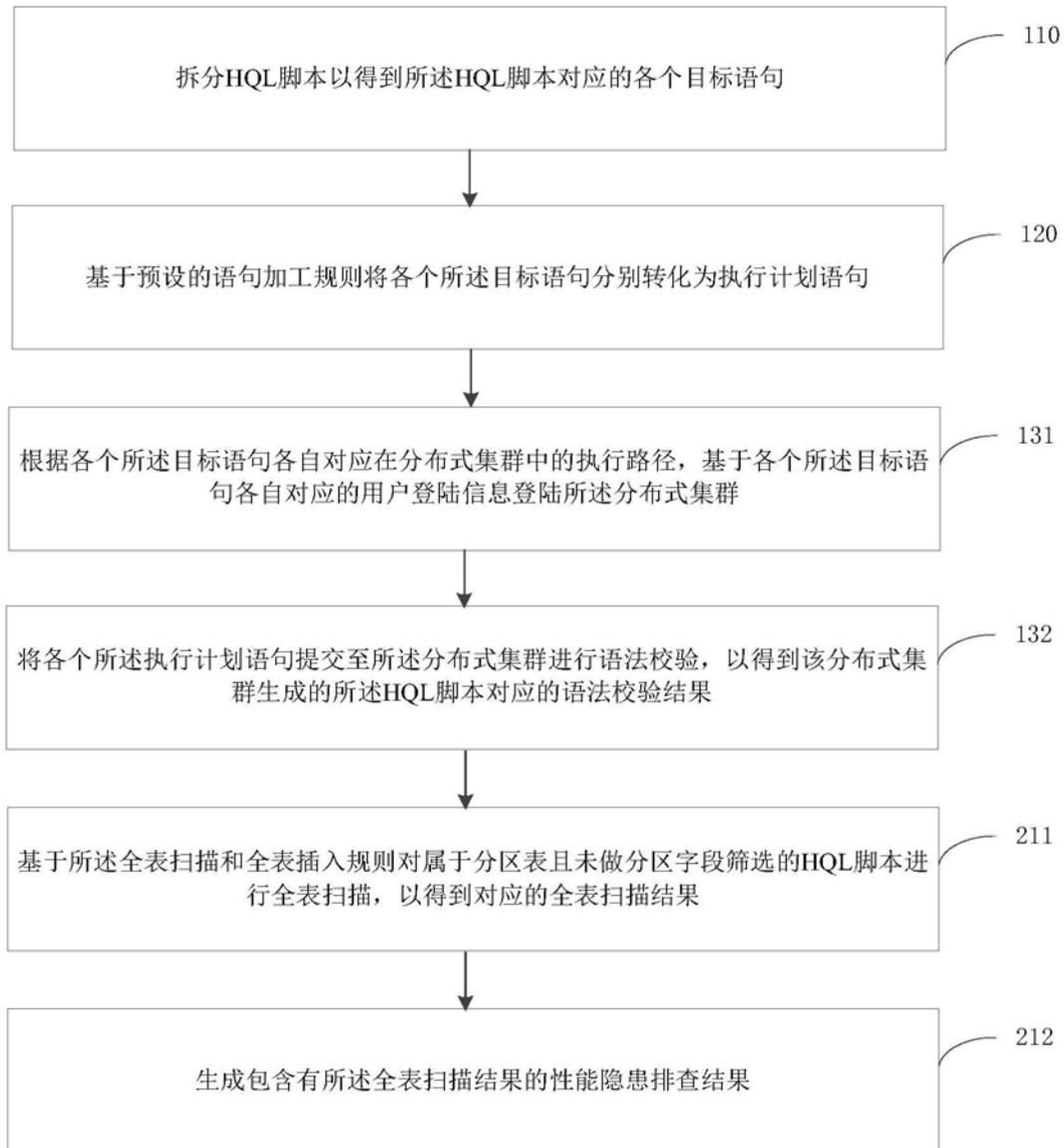


图5

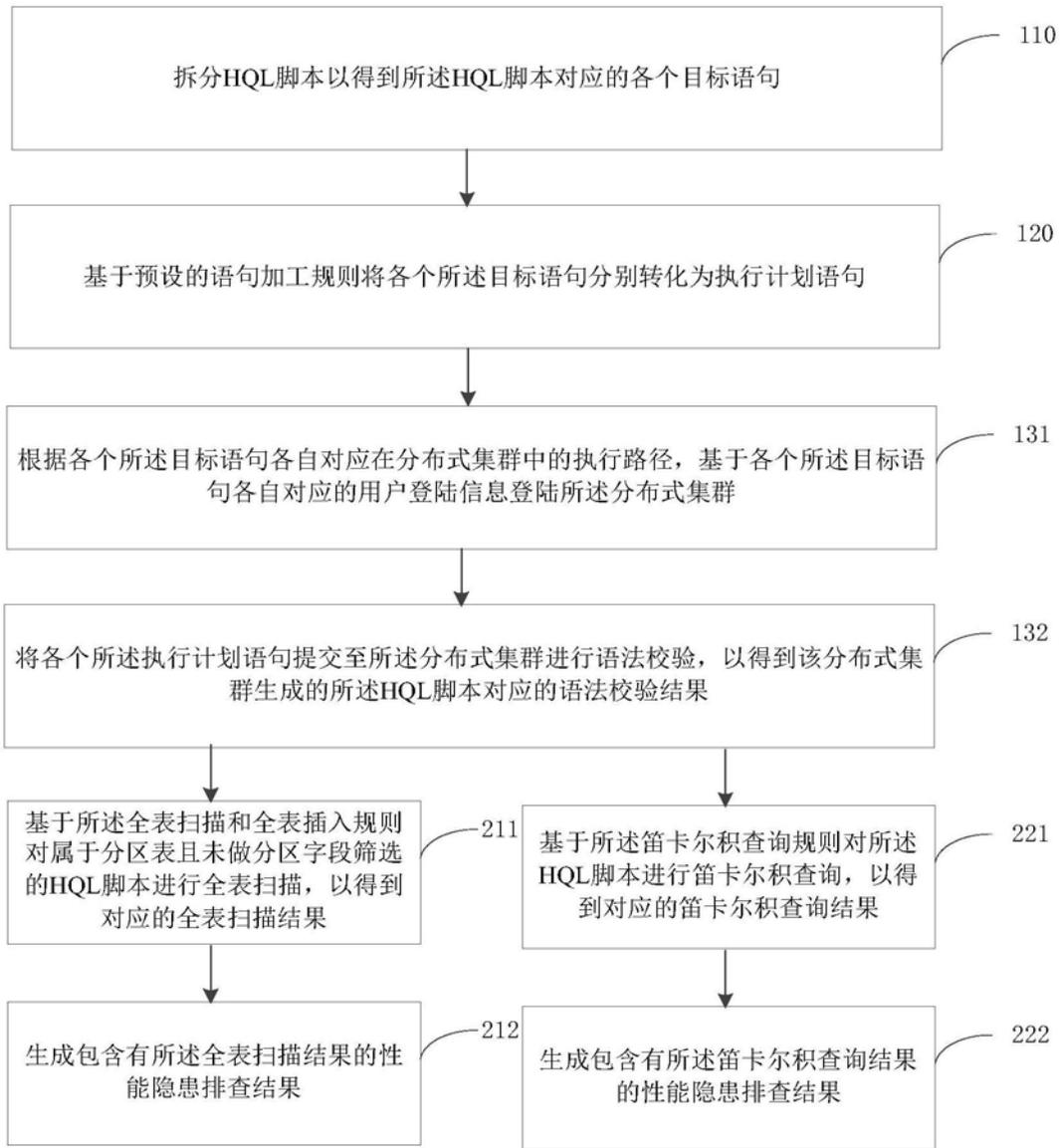


图6

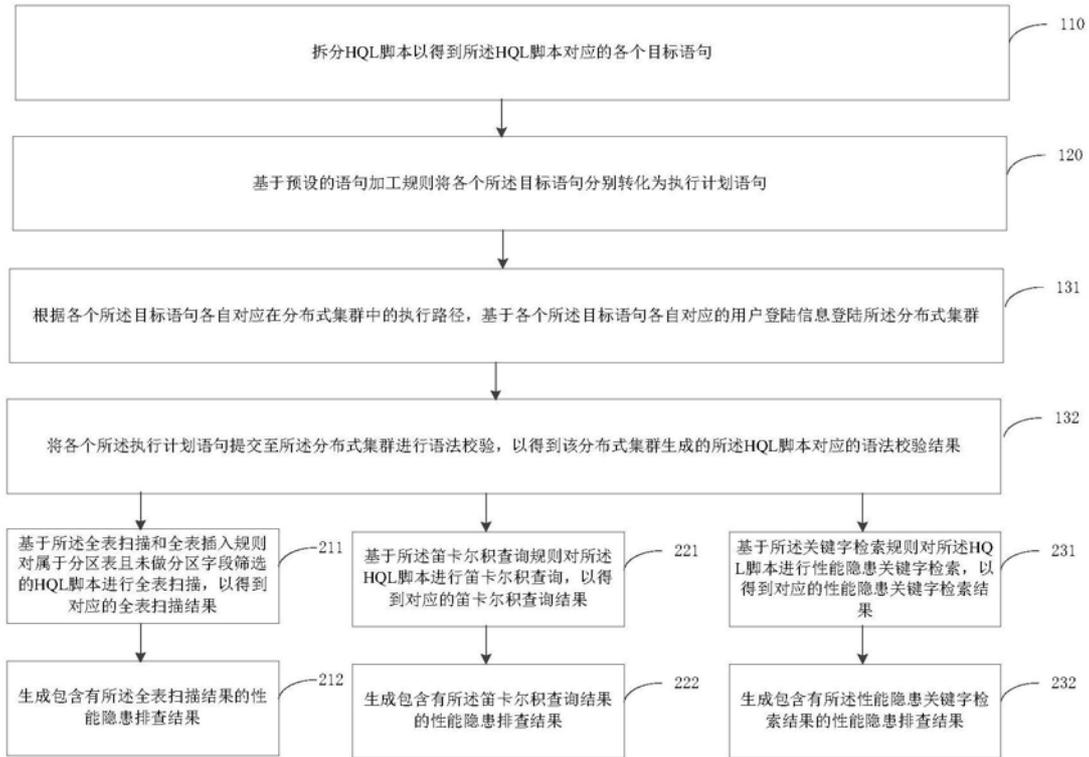


图7

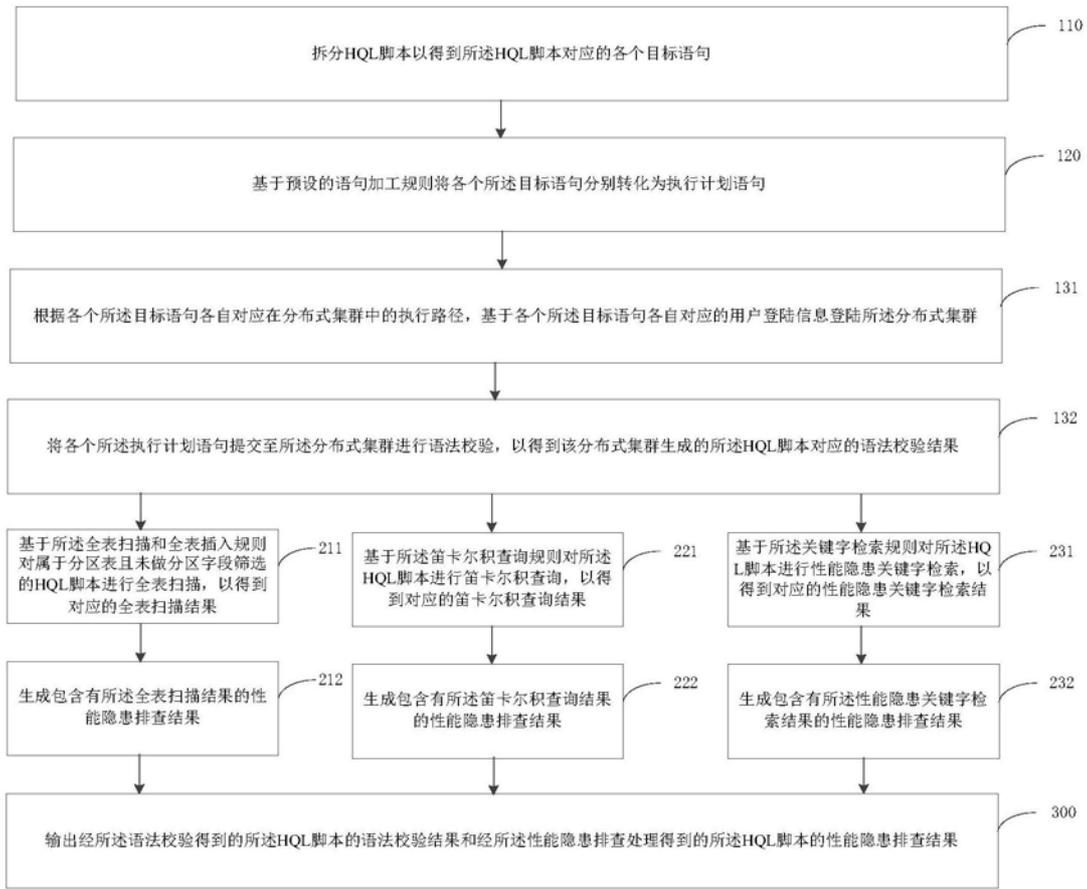


图8

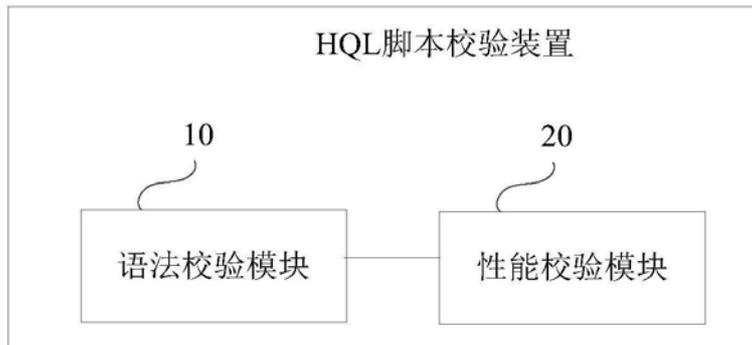


图9

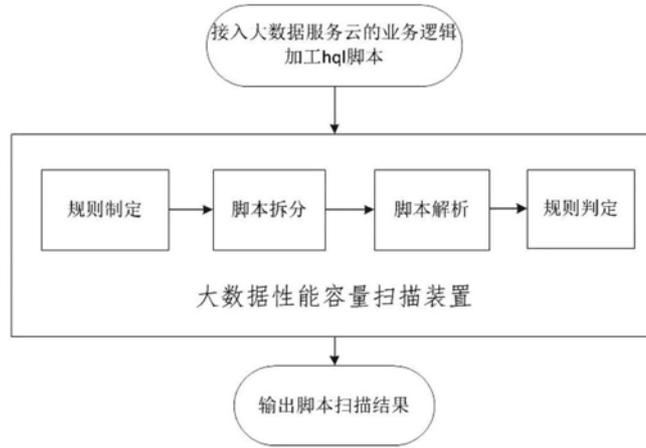


图10

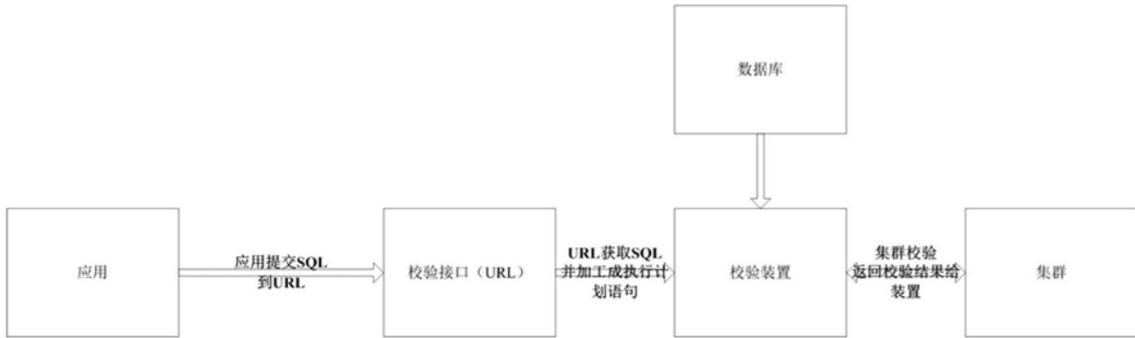


图11

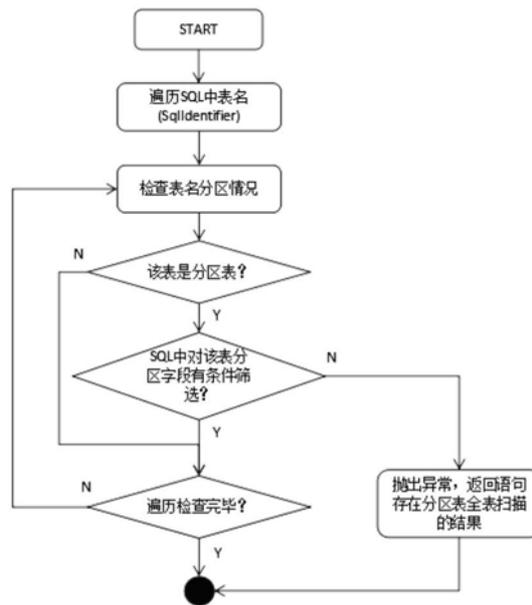


图12

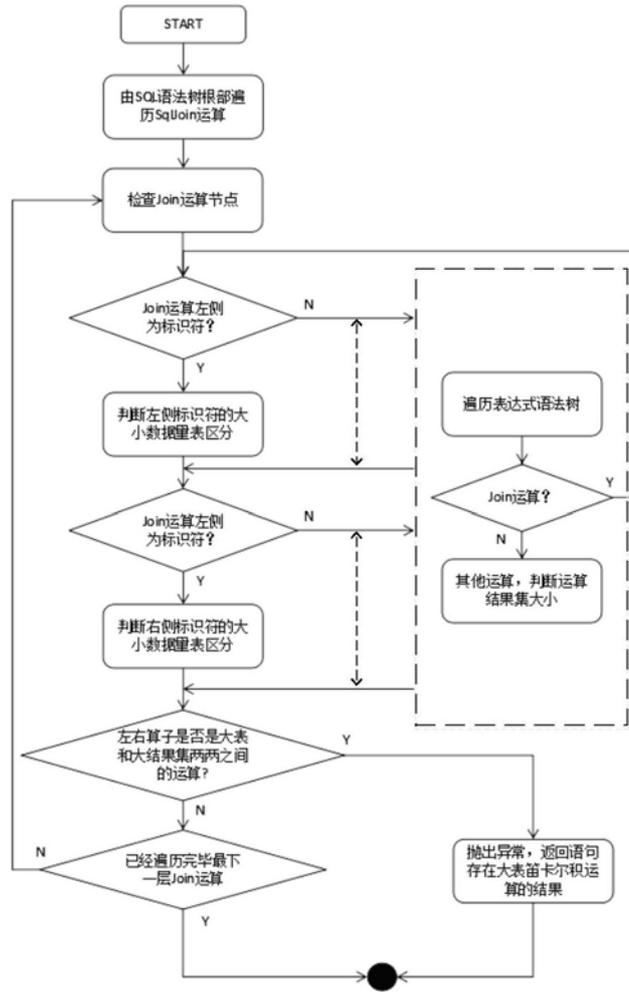


图13

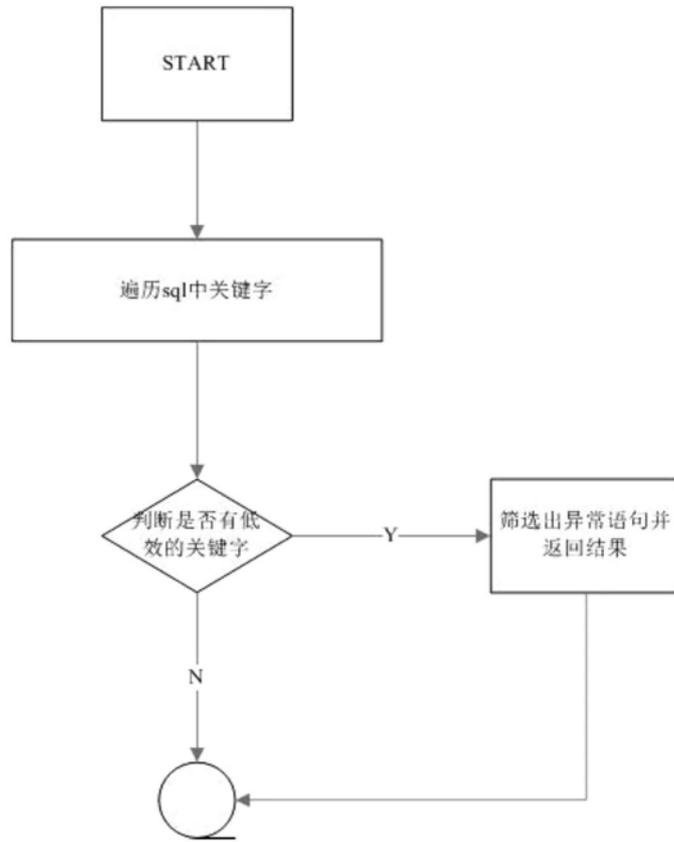


图14

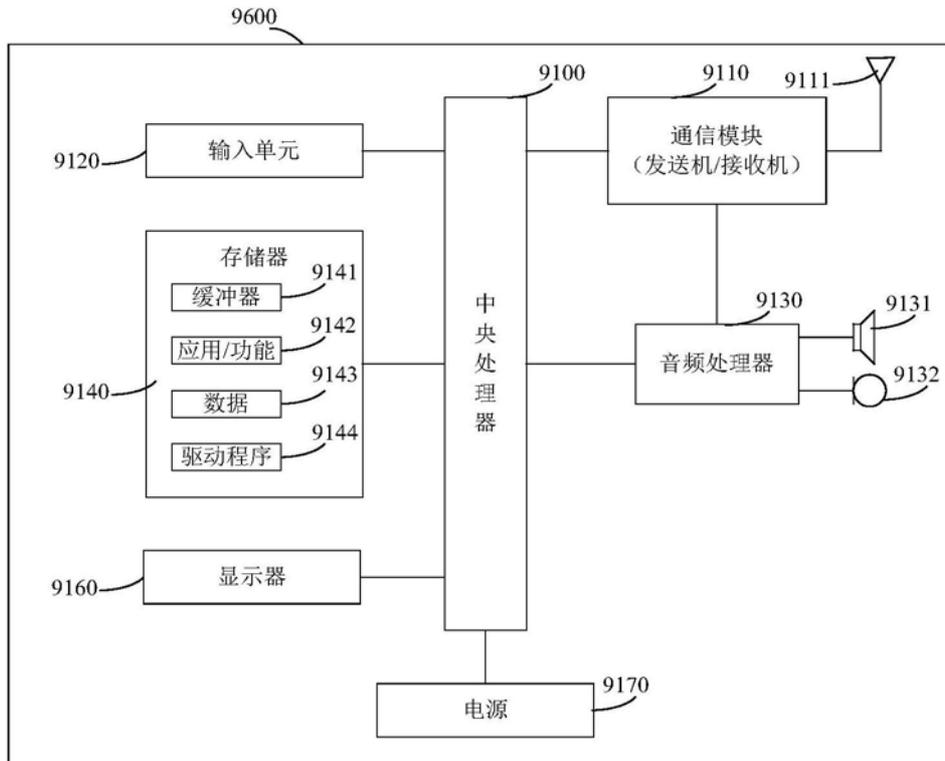


图15