

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 17/30 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200610142365.3

[43] 公开日 2008年4月16日

[11] 公开号 CN 101162462A

[22] 申请日 2006.10.11

[21] 申请号 200610142365.3

[71] 申请人 国际商业机器公司

地址 美国纽约

[72] 发明人 任鲲鹏 梁高中 刘展华 杨文

傅毓勤 孙冰江 魏可伟

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 李春晖

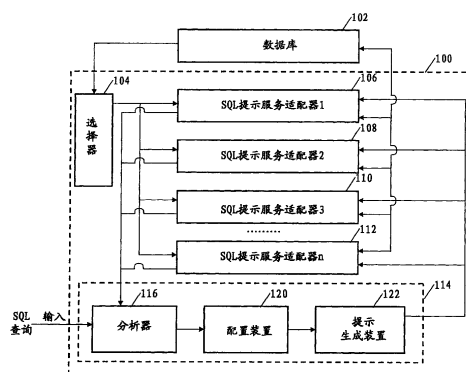
权利要求书2页 说明书11页 附图6页

[54] 发明名称

提示定制工具与方法

[57] 摘要

本申请涉及提示的定制工具与方法。根据本发明，提供至少一个提示服务适配器，用于搜集不同的数据库平台的相关信息，将定制好的提示在真实的数据库平台上进行验证并提供反馈，并用于将定制好的提示部署到数据库平台上，并提供提示定义装置，用于与平台无关地定制提示，并将定制好的提示提供给所述至少一个提示服务适配器。



1. 一种提示定制工具，包括：

至少一个提示服务适配器，用于搜集不同的数据库平台的相关信息，将定制好的提示在真实的数据库平台上进行验证并提供反馈，并用于将定制好的提示部署到数据库平台上；

提示定义装置，用于与平台无关地定制提示，并将定制好的提示提供给所述至少一个提示服务适配器。

2. 如权利要求1所述的提示定制工具，还包括选择器，用于在连接到数据库平台时分析数据库平台的类型，并选择相应的提示服务适配器。

3. 如权利要求1或2所述的提示定制工具，其中，所述提示定义装置包括：

分析器，用于分析输入的SQL查询，获得表连接图和默认表联接顺序图，表联接顺序图包括表节点和表联接方法节点；

配置装置，用于修改所述默认表联接顺序图中各表的联接顺序，获得表联接顺序图，并配置所述默认表联接顺序图或者表联接顺序图中各表节点以及表联接方法节点的属性；以及

提示生成装置，用于依据分析器和配置装置的结果生成提示。

4. 如权利要求1到3之一所述的提示定制工具，还包括实时验证装置，用于根据相应的提示服务适配器所提供的有关数据库平台的相关信息，在所述配置装置的修改和配置过程中对要生成的提示进行实时验证，并提供反馈。

5. 一种提示定制方法，包括下述步骤：

a. 根据数据库平台的类型选择相应的提示服务适配器；

b.分析输入的SQL查询，获得表连接图和默认表联接顺序图，表联接顺序图包括表节点和表联接方法节点；

c.修改所述默认表联接顺序图中各表的联接顺序，获得表联接顺序图，并配置所述默认表联接顺序图或者表联接顺序图中各表节点以及表联接方法节点的属性；

d.依据在前一步骤的结果生成提示；

e.通过提示服务适配器在数据库平台环境中验证所生成的提示，并在必要时将验证结果反馈给步骤c，重复步骤c到e，直到得到合乎需要的提示；

f.通过提示服务适配器将生成的提示部署到数据库平台中。

6.如权利要求5所述的提示定制方法，在步骤a之前还包括连接到数据库平台并分析数据库平台的类型的步骤，并且，在步骤a中依据该步骤的结果选择相应的提示服务适配器。

7.如权利要求5或6所述的提示定制方法，还包括对步骤c的实时验证步骤，用于根据所述提示服务适配器所提供的有关数据库平台的相关信息，对步骤c中的修改和配置操作进行实时验证，并提供反馈，根据反馈重复步骤c。

提示定制工具与方法

技术领域

本申请涉及数据库管理系统中的SQL（结构化查询语言）提示机制（也称为优化提示），尤其涉及提示的定制，包括提示的验证。

背景技术

几家主要的DBMS（数据库管理系统）供应商比如IBM、微软和Oracle公司都支持SQL（结构化查询语言）提示机制（SQL hint mechanism）（也称为优化提示等）。SQL提示机制为有经验的数据库管理员（DBA）提供了一种建议或者要求SQL优化器基于DBA指定的标准定制访问计划（access plan）的某些部分或者全部的途径。例如，用户可能知道，对于特定的查询，某个索引的选择性更好，如果选择该索引的话效果更佳。基于他的知识，用户可能会想使用SQL提示来指引优化器使用该索引，而不管优化器的决定是什么。但是，为了使用当前的SQL提示，它要求很高的SQL调优技巧或者经验，而且包括创建、验证、部署在内的整个过程是一个非常冗长而费时的过程。

原因在于：

1. SQL提示是专属地生成的。例如，为了生成提示，对于DB2 for z/OS, DBA需要更新访问计划表（PLAN_TABLE）并设置特定的注册信息（register）。对于Oracle或者SQL Server, DBA需要通过SQL追加或注释的方式来使用SQL提示，对于DB2 for LUW, DBA需要在SQL语句的末尾附加一些XML文本。在如今的企业异构数据库环境中，这极大地增加了DBA的负担，因为他们需要掌握每一个平台SQL提示机制的细节。

2. 需要手工指定SQL提示。对于非常复杂的查询，例如包含100个谓词（predicate）和20个表的查询，DBA要使提示正确是很难的而

且冗长的工作。

3. 需要手工验证SQL提示。提示在语法方面正确是不够的。提示可能因为冲突而无效。尽管各种DBMS都提供了对提示的使用的一些反馈，但是它们通常对用户都不甚友好。

显然，需要有一种有效的解决方案来解决至少部分上述问题。

发明内容

因此，本发明的目的之一是提供一种提示定制工具，能够使用户与数据库平台无关地定制提示。本发明的另一个目的是提供一种与平台无关地定制提示的方法。

为实现上述目的，根据本发明的第一方面，提供了一种提示定制工具，包括：至少一个提示服务适配器，用于搜集不同的数据库平台的相关信息，将定制好的提示在真实的数据库平台上进行验证并提供反馈，并用于将定制好的提示部署到数据库平台上；以及提示定义装置，用于与平台无关地定制提示，并将定制好的提示提供给所述至少一个提示服务适配器。

根据本发明的第二方面，提供了一种提示定制方法，包括下述步骤：a.根据数据库平台的类型选择相应的提示服务适配器；b.分析输入的SQL查询，获得表连接图和默认表联接顺序图，表联接顺序图包括表节点和表联接方法节点；c.修改所述默认表联接顺序图中各表的联接顺序，获得表联接顺序图，并配置所述默认表联接顺序图或者表联接顺序图中各表节点以及表联接方法节点的属性；d.依据在前一步骤的结果生成提示；e.通过提示服务适配器在数据库平台环境中验证所生成的提示，并在必要时将验证结果反馈给步骤c，重复步骤c到e，直到得到合乎需要的提示；f.通过提示服务适配器将生成的提示部署到数据库平台中。

附图说明

下面结合附图对本发明的具体实施方式加以说明。附图中：

图1是本发明的提示定制工具的一种优选实施方式的框图；
图2是本发明的提示定制工具的另一种优选实施方式的框图；
图3A是表联接图的一个举例；
图3B是图3A所示表联接图所对应的默认表联接顺序图的一个举例；
图4是定义节点属性的实施方式的举例；
图5是本发明的提示定制方法的一种优选实施方式的流程图；
图6是本发明的提示定制方法的另一种优选实施方式的流程图。

具体实施方式

本发明的主要目的在于解决DBA需要针对不同的数据库平台定制提示的问题。为此，本发明提出了一种新的解决方案，即，提供基于SQL查询的通用的提示定义装置，以及SQL提示服务适配器（以下简称适配器），利用适配器作为数据库平台和提示定义装置之间的中介，在数据库平台中验证提示定义装置所定义的提示，并在数据库平台中部署通过验证的提示。

不同的数据库平台需要不同的适配器。在本发明的一种实施方式中，可以针对不同的数据库平台提供不同的适配器，由DBA根据所使用的数据库平台来选择相应的适配器。

为了进一步便利DBA，可以提供一个选择器，其探测数据库平台的相关信息，据以确定数据库平台的类型和版本，从而选用合适的适配器。

图1图示了该优选的实施方案的提示定制工具100。如图所示，有多个适配器106、108、110、112，每一个适配器对应于一种数据库平台。选择器104根据从数据库102获得数据库的类型和版本，从而选择合适的适配器。DBA使用提示定义装置114基于输入的SQL定制提示，并通过适配器来在实际的数据库平台中验证，并将通过验证的提示部署到数据库平台中。

作为更为优选的实施方式，图1所示的提示定制工具100的提示定义装置114进一步包括分析器116、配置装置120和提示生成装置122。

分析器分析输入的SQL查询，结合从适配器得到的数据库平台的系统信息，得到表联接图和默认表联接顺序图。表联接图表示一个SQL查询所涉及的表以及表之间的相互关系。依据SQL分析数据，可以提取出SQL执行过程中所涉及到的所有表以及表之间的相关关系。依据从数据库得到的数据库平台的系统信息，可以获知它们更进一步的相关信息，例如表中定义的列、索引（index）以及统计信息等。表联接顺序图表示执行该SQL查询时，访问各表的顺序。表联接顺序可以是合乎表联接图的逻辑的任意顺序。但是，可以使表联接顺序是合乎表联接图的逻辑的最优顺序。此外，还可以基于查询的历史和过去已定义的提示来对该顺序加以约束，给出默认表联接顺序图。

所述表联接图和默认表联接顺序图可以在图形用户界面上予以直观形象的显示。如图3A所示，每一个方框节点代表一个表。框之间的连线表示所连的两个框相关，对应于查询中的谓词（predicate）。该图用来展示查询中的表是如何联接的。联接图不仅指出查询中的表是如何联接的，还包含原存储方案的各种属性。在初始状态，可以提供联接图的自动布局，以改善用户的视觉体验。用户还可以用拖放操作对其重新进行手工布局。在图上，由可视方案提示提供了两个切换按钮（trigger），允许用户显示本地谓词和联接谓词。

图3B所示为基于图3A的表联接图的默认表联接顺序图。图中，各表在图形用户界面上的排列顺序就表示了表联接顺序。联接每两个表的节点表示联接方法。用户可以从图3A的联接图推知表以什么顺序联接是最有效率的。默认的联接顺序是在数据库优化器不使用提示情况下确定的表间联接关系。通过调整默认联接顺序，用户能够容易地生成自己提示所需的联接顺序。

配置装置120可以对图3B所示的表联接顺序图进行修改。例如，可以对表进行拖放，移动到图3B所示表联接顺序图中的适当位置，从而完成对表联接顺序的修改。当用户定义其联接顺序图时，可以高亮标记已经使用了的节点，自动检查首表等，以协助用户的定义。

一般来说，除了表的联接顺序之外，提示还可以包括以下内容：

1. 如何访问表。例如，是优先选择索引扫描还是表扫描（r-scan），使用什么索引，是否使用列表预取（list prefetch）等；
2. 多表访问时的表联接方法；
3. 并行性，比如并行模式和并行度。

这些提示内容可以视为表节点和联接方法节点的属性。按照一种优选的实施方式，配置装置120被设计为当选择某个表节点或者联接方法节点进行配置时，弹出如图4所示的对话框，在其中可以填写或者选择有关的属性。具体来说，例如，对于DB2 for z/OS平台可以有下表所述的属性：

表1 表节点和联接方法节点的属性举例

1. 定义表的访问方式（表节点的属性）		
1.1 确定所要定义的表。为了确定提示中所要定义的表，下面几个属性用于唯一地标识该表。	所需设置的属性	说明
	CREATOR	表的创建者标识，如果TNAME不能唯一地标识所定义的表，则需要指定该属性。
	TNAME	表名，同时也可以是视图名或表表达式名。
	TABNO	表编号，如果指定CREATOR、TNAME、CORRELATION_NAME以及QBLOCKNO还不能够唯一地标识所定义的表，则需要指定该属性。
	CORRELATION_NAME	在SQL中表的别名
	QBLOCKNO	在存储访问计划中表所在的查询块
1.2 定义表的访问类型，定义数据库引擎访问该表的方法，例如该属性设置为I时表示将采用全索引扫描的方法访问表；设置为R时	ACCESSTYPE	其有效值必须为I, I1, N, M, R, RW, T, 或V。

表示采用全表扫描的方法访问		
1.3 定义访问时所使用的索引,但在1.2中定义的表访问方法为索引扫描时,用户可以定义下列两个属性指定索引扫描时所使用的索引(Index)。	ACCESSCREAT OR	索引的创建者标识
	ACCESSNAME	索引名
1.4 定义预取方式	PREFETCH	有效值必须为S,L或空; 该值定义DB2所使用预取方式: sequential prefetch (S): 顺序预取 list prefetch (L): 列表预取 no prefetch (空)
1.5 定义数据扫描时是否限制在某个或某几个分区表空间中。	PAGE_RANGE	有效值必须为Y,N或空
1.6 定义表访问时是否尝试使用数据行直接访问	PRIMARY_ACC	有效值为D或空
	ESSTYPE	D: 尝试使用数据行直接访问 空: 不试图尝试使用数据行直接访问
2. 在多表访问中,表的联接 (联接方法节点的属性)		
2.1 指定表联接方法	METHOD	有效值必须为0, 1, 2, 3, or 4. 0: 第一个访问的表 (leading table) 1: 嵌套循环联接 (Nested loop join) 2: 合并扫描连接 (Merge scan join) 3: ORDER BY, GROUP BY, SELECT DISTINCT, UNION,所需要的排序 4: 混合联接 (Hybrid join)
2.2 多表的联接顺序	PLANNO	对于一个查询块用连续的数字表示表的联接顺序。
2.3 关于联接时的排序	SORTN_JOIN	有效值必须为Y,N。联接时内表 (Inner table) 是否需要排序
	SORTC_JOIN	有效值必须为Y,N。联接时合成表 (Composite

		table) 是否需要排序
3. 并行性 (联接方法节点的属性)		
3.1 指定并行模式	PARALLELISM_MODE	指定使用的并行模式, 如果可以并行访问, 它的有效值为I, C, X或空.
3.2 指定表访问并行度	ACCESS_DEGREE	并行任务数
3.2 指定表联接并行度	JOIN_DEGREE	联接内表和合成表时的并行任务数

上表所示的“2.2 多表的联接顺序”是通过直接拖放表节点来确定的。但是, 在节点属性的对话框中, 也可以显示出该属性。并且也可以通过修改其值来改变联接顺序。

需要注意的是, 上述属性本身只是作为举例。在实现本发明的过程中, 可以尽量囊括所有数据库平台的提示可能涉及到的属性, 并可以在显示属性对话框时, 根据DBA所指定的或者选择器所确定的数据库平台类型或者版本等来定制该属性对话框, 屏蔽掉不必要的属性选项。

为了协助用户定义提示, 在上述图形用户界面中还可以提供了下述功能:

1. 提供已有的访问计划信息, 帮助用户决定提示;
2. 显示可变的提示, 隐藏不可变的提示;
3. 显示本地谓词, 在表联接图上加入谓词, 以帮助用户决定提示;

在完成对各节点的配置之后, 提示生成装置122即可生成提示, 即, 将定制信息最终转换成系统内部的提示定义形式(它是提示定义的抽象表示)。因为内部提示定义与数据库的平台和提示类型无关, 因此它可以用来为任何数据库平台和任何提示类型自动地生成提示。同时, 还可以使用户避免了前面提到手工操作中的错误, 例如输入、语法错误以及语义冲突等。

适配器将生成的提示在实际的数据库平台环境中予以验证，最终给用户验证报告。从该报告用户可以知道：

1. 数据库优化器使用的是提示的哪些部分；
2. 提示的哪些部分没有使用，以及为什么不使用。对于这些提示，还可以给出一些建议。
3. 访问路径有何差别。

最后，适配器将通过验证的提示部署到数据库平台中。所谓部署的过程也就是将在提示定义装置114中生成的提示按照各种数据库平台所固有的方式加入到数据库平台中。

图2图示了本发明的提示定制工具的另一种优选实施方式。该优选实施方式基本上与图1所示是一样的，但是在提示定义装置204（对应于图1中的提示定义装置114）中增加了实时验证装置202。实时验证装置202提供客户端运行时验证，用于验证用户通过配置装置120正在定义的提示中的语义错误和冲突。当发生这样的错误时，可以高亮显示这样的错误并给出相应的建议。

在进一步的优选实施方式中，为了减少适配器在实际的数据库平台环境中进行验证所耗用的时间，可以结合适配器所收集并提供的数据库平台的有关信息，对正在定义的提示进行实时验证。这种实时验证虽不能完全替代在实际的数据库平台环境中的验证，但能够实现其绝大多数功能，可以大大缩短验证时间。

在上面的优选实施方式中，如果数据库的有关信息已经为本发明的提示定制工具所知（例如，提示定制工具已经获取过有关信息，且在那之后数据库并未发生有关变化），则在提示定制工具后来的操作中无需连接到数据库。反之，则需要连接到数据库以获取有关信息。

上面描述了本发明的提示定制工具。对于提示定制工具的各个部件的描述，还可以参照下文对提示定制方法的相应描述。

下面描述与上面的提示定制工具相应的提示定制方法。

图5图示了本发明的提示定制方法的一种优选实施方式。首先，

确定所要加载的SQL提示服务适配器（步骤502），以便于做进一步的处理，例如加载编目信息（catalog），收集数据库相关数据生成表联接图以及默认联接顺序图等。SQL提示服务适配器是本发明提供的一种机制：针对不同的数据库平台，应用不同的适配器，这样可以屏蔽掉平台不同而带来的差异。关于适配器，前文结合图1也有描述。

同样，如同在前文所述的提示定制工具中一样，如果数据库的有关信息是已知的，则在执行本发明的方法时无需连接到数据库。反之，则需要首先连接到数据库，检查数据平台及数据库版本信息（图6的步骤602），据此确定适配器。

接着，在步骤504，分析输入的SQL，生成表联接图以及默认联接顺序图，在这一过程中，以下的信息将被收集：

- 数据库编目信息（catalog）。
- 数据库引擎选择的当前访问路径信息，并将它们进行重新组织为一个统一的数据模型。这个数据模型可以用来生成该SQL的表联接图以及默认的表联接顺序图。在本发明中，表联接图以及默认的表联接顺序图中的节点携带有可定制的属性集。
- 加载与平台相关的SQL提示的配置信息，以及从规则库中加载基于知识的平台相关的SQL提示的验证规则，这些规则将被用来做平台相关的SQL提示的验证。

所需的数据信息被正确收集后，即可构建出与平台无关的表联接图和默认表联接顺序图。

接着，在步骤506，用户就可以基于上述表联接图和默认表联接顺序图，进行SQL提示的定制。例如，在表本身层面，用户可以定义表的访问方式以及索引的使用情况等。在表间关系层面，用户可以定义表间的联接顺序以及联接方法等。具体定制方式，可参见前文对配置装置120的描述。

当用户完成提示的定制后，即生成SQL提示（步骤508），并通过适配器将它们发送到数据库引擎端进行验证（步骤510），获得反馈信息，例如给出SQL提示验证报告。提示验证报告中，可以帮助用户

获知以下事项：

1. 根据用户的定义标准和选定的提示类型生成的SQL提示；
2. 不使用提示情况下的访问计划和使用提示的方案之间的区别。

突出显示它们；

3. 提示中的哪些部分已生效；
4. 提示中的哪些部分是无效的或者未使用，为什么不使用，并给出相应的建议。

根据反馈信息，如果用户对当前结果不满意，可以进一步修改表联接顺序图，配置节点属性，来重新定义提示，再对其重新验证，直到得到满意的结果。

如果用户定义的SQL提示就是用户所需要的，则可以通过适配器将定制的提示部署到数据库平台中（步骤512），以便其他应用可以使用。因为提示的定制过程与具体的数据库平台及提示类型无关，因此它可以部署到具有任何提示类型的任何数据库。用户只需要定义一次。

在更为优选的实施方式中，如图6所示，可以在修改表联接顺序图、配置节点属性的同时对所定义的提示进行实时验证。如果发现任何的错误和警告信息，系统将会实时地指出发生错误或存在警告提示，并给出相应的建议。总的来说，提示的实时验证可以划分为如下两个类别。

一是一般意义上的验证。这部分的验证基于公共的验证规则集，适用于各种不同的数据库平台，与具体的平台无关。它包括：

- 数据库模式相关的验证。数据库模式的定义对不同数据库平台都是相同的。拿定义于某一数据库模式上的索引为例，对于某个SQL中的表，如果没有索引存在于这个表上，那么一个在这个表上使用索引扫描的SQL提示，就没法应用。

- SQL语句本身的相关验证。SQL对于不同的数据库来说，它都是使用统一的标准。例如，对于一个索引，如果没有任何的列在SQL语句中出现，而且这个SQL中没有试图去计算这个表的记录总数，如果这时SQL提示中要求使用这个索引，这肯定是低效的。

● 在已定义的SQL提示中进行冲突检测。例如，新加入的SQL提示要求针对某个表使用索引扫描，而在以前定义的提示中已经定义了这个表要使用表扫描的方式进行访问，这就产生了冲突。又如在表联接顺序中对环的自动检测。

二是与平台相关的验证。通过基于规则引擎的方法对那些与平台相关的SQL提示进行验证，这些基于知识的规则是由前面提到的SQL提示服务适配器针对不同平台类型进行加载和构建的。例如，对于DB2 z/OS平台，混合联接（Hybrid join）要求内表的访问必须通过基于索引的列表预取（list prefetch）方式进行。

上面描述了本发明的提示定制工具和提示定制方法的优选实施方式。由以上说明可见，本发明具有下述优点：

1. 与特定数据库平台无关。用户不需掌握每一个平台提示机制的细节，极大地减少了DBA的负担；
2. 允许用户可视地定义、验证和部署优化提示，并更容易利用提示；提供运行时验证和建议，令用户消除暗含的语义错误和冲突；
3. 基于图形界面，用户能够定义复杂得多的优化提示。

本领域普通技术人员知道，对于本说明书中所描述的各种部件和步骤，都有诸多替代方案可供使用。因此，本发明的保护范围不限于说明书中所述，而应包括其所有等效方案。

图1

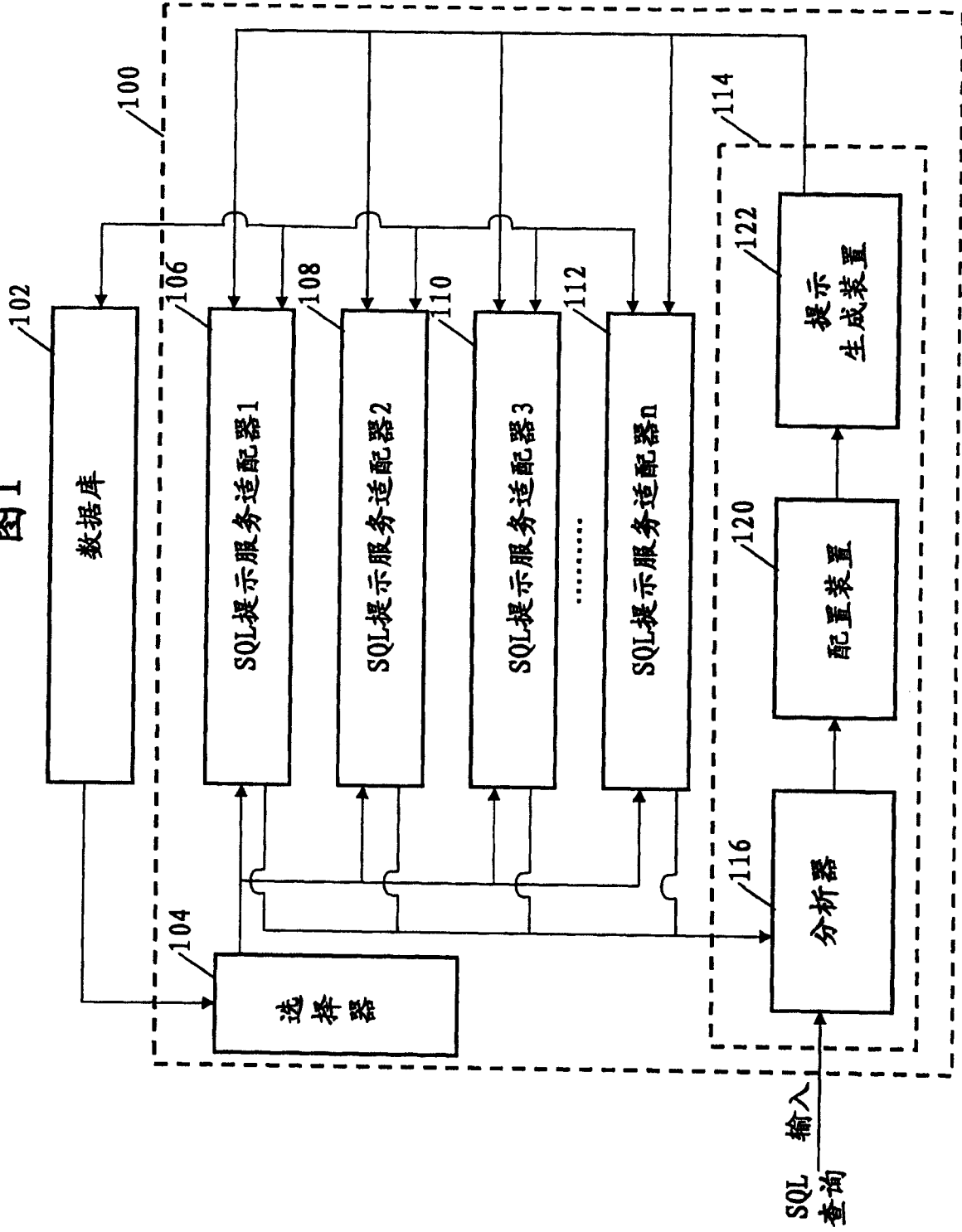


图 3A

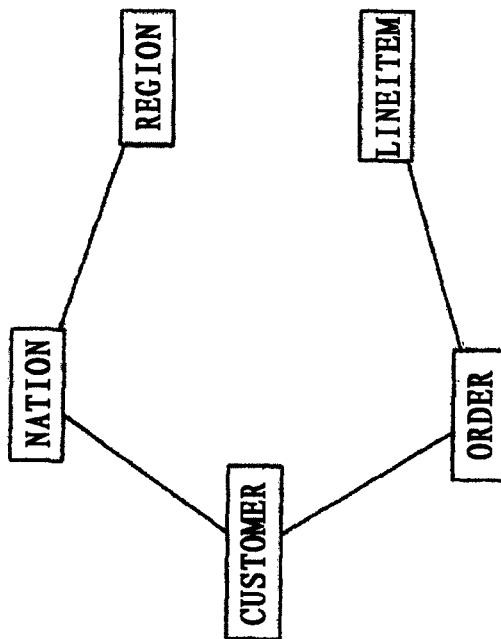


图 3B

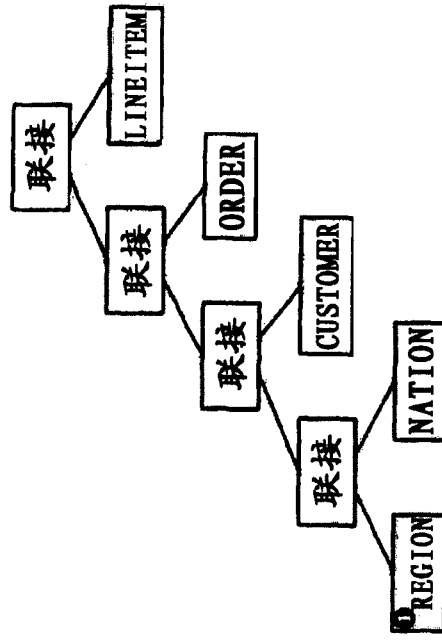


图5

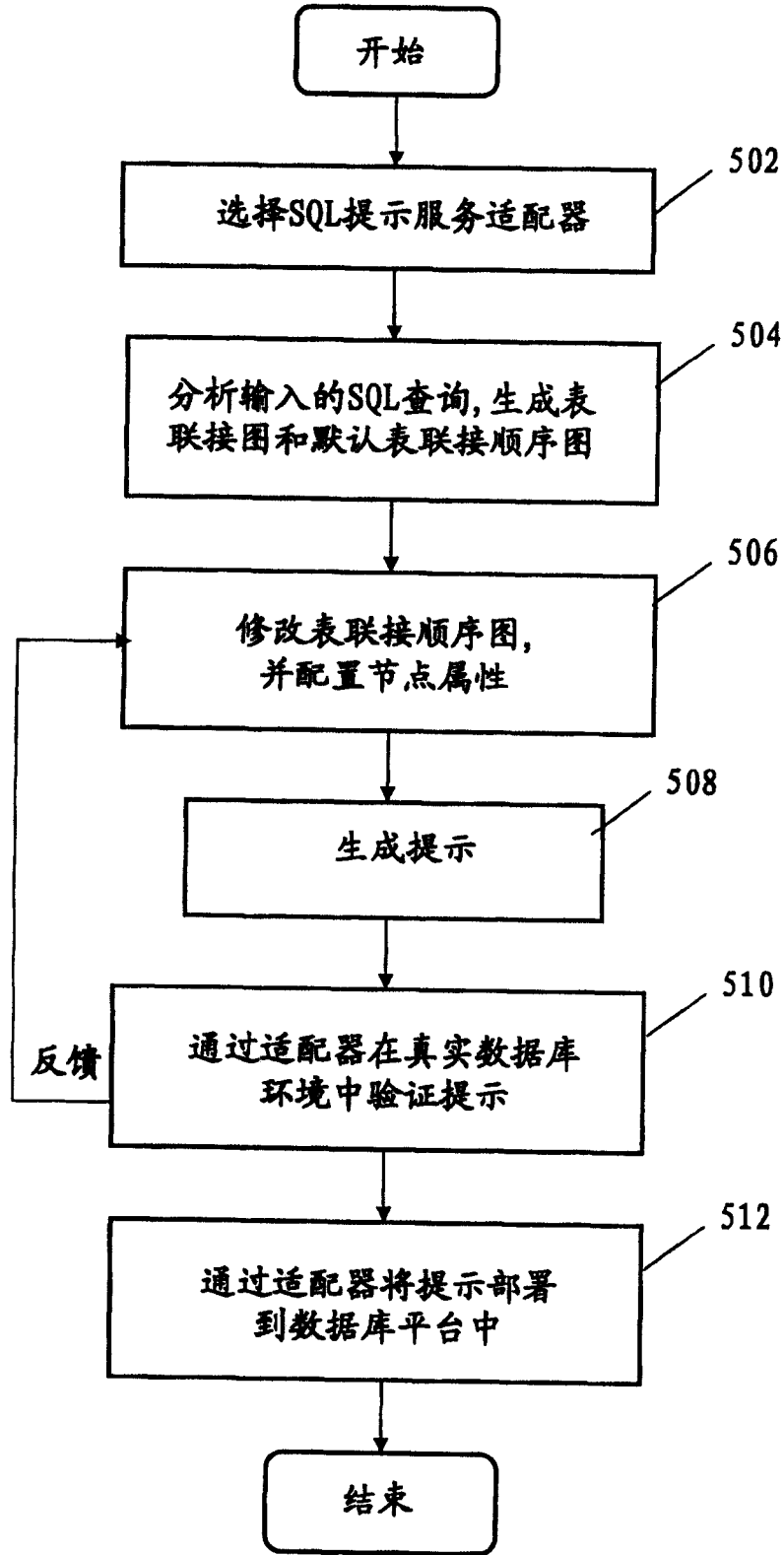


图6

