



## (12) 发明专利

(10) 授权公告号 CN 107622103 B

(45) 授权公告日 2021.12.28

(21) 申请号 201710812945.7

G06F 16/2455 (2019.01)

(22) 申请日 2012.04.30

G06F 16/245 (2019.01)

(65) 同一申请的已公布的文献号

G06F 16/901 (2019.01)

申请公布号 CN 107622103 A

G06F 16/2453 (2019.01)

(43) 申请公布日 2018.01.23

(30) 优先权数据

(56) 对比文件

13/098,823 2011.05.02 US

CN 102007476 A, 2011.04.06

(62) 分案原申请数据

US 2003120682 A1, 2003.06.26

201280021472.4 2012.04.30

CN 1790324 A, 2006.06.21

(73) 专利权人 起元技术有限责任公司

CN 1756190 A, 2006.04.05

地址 美国马萨诸塞州

CN 101868792 A, 2010.10.20

(72) 发明人 I·谢克特 G·J·艾琳

CN 101971165 A, 2011.02.09

(74) 专利代理机构 北京林达刘知识产权代理事

Graefe et al.. "Encapsulation of parallelism and architecture-independence in extensible database query execution".

务所(普通合伙) 11277

《IEEE Transactions on Software

代理人 刘新宇

Engineering》.1993,第19卷(第8期),全文.

审查员 陈丽媛

(51) Int. Cl.

G06F 16/242 (2019.01)

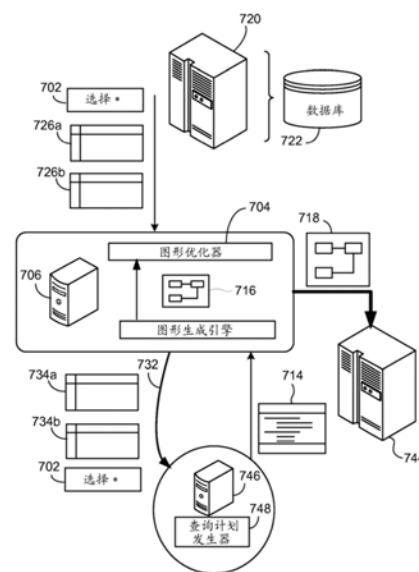
权利要求书11页 说明书14页 附图12页

(54) 发明名称

管理数据查询

(57) 摘要

本申请涉及管理数据查询。一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:接收查询;基于所述查询来识别第一数据源;基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作。



1. 一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,生成所述操作集合包括:

基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及至少一个链路,其代表所述第一节点的操作的输出数据被提供作为向所述数据流图形的第二节点的操作的输入数据。

2. 根据权利要求1所述的方法,其中,根据能够应用于关系数据库的查询语言来表示所述查询。

3. 根据权利要求1所述的方法,其中,所述第一数据源和所述第二数据源将数据存储在有形的非暂时性计算机可读介质中。

4. 根据权利要求1所述的方法,其中,所述数据流图形还包括:第三节点,其与运行系统为了从所述第一数据源接收数据而能够使用的信息相关联;以及第四节点,其与运行系统为了从所述第二数据源接收数据而能够使用的信息相关联。

5. 根据权利要求4所述的方法,其中,所述第三节点与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述第一数据源的系统。

6. 根据权利要求5所述的方法,其中,还包括以下步骤:判断为管理所述第一数据源的系统能够执行基于所述查询计划所描述的所述第二步骤而选择的所述至少一个操作。

7. 根据权利要求5所述的方法,其中,还包括以下步骤:判断为管理所述第二数据源的

系统不能执行基于所述查询计划所描述的所述第二步骤而选择的所述至少一个操作。

8. 根据权利要求1所述的方法, 其中, 还包括以下步骤: 实例化所述数据流图形的节点和第二数据流图形的节点之间的链路, 其中所述第二数据流图形包括代表除基于所述查询计划所描述的任何步骤而选择的操作以外的要运行的操作的至少一个节点。

9. 根据权利要求1所述的方法, 其中, 所述查询包含向可运行功能的引用, 并且所述数据流图形包括代表所述可运行功能的节点。

10. 根据权利要求1所述的方法, 其中, 所述第一数据源包括关系数据库表格。

11. 根据权利要求10所述的方法, 其中, 所述第二数据源包括除关系数据库表格以外的数据的一部分。

12. 根据权利要求10所述的方法, 其中, 所述第二数据源包括平面文件。

13. 根据权利要求1所述的方法, 其中, 所述查询包括SQL。

14. 根据权利要求1所述的方法, 其中, 还包括以下步骤: 接收所述查询的结果作为所述程序代码的输出。

15. 根据权利要求1所述的方法, 其中, 所述查询计划器包括查询优化器。

16. 根据权利要求1所述的方法, 其中, 所述查询计划包括所述查询中的参数的数据类型。

17. 根据权利要求1所述的方法, 其中, 所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联, 并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

18. 根据权利要求1所述的方法, 其中, 基于所述查询来生成向查询计划器的请求包括: 产生所述第一数据源的结构代表, 其中所述结构代表包括反映所述第一数据源的格式化的数据组织, 并且缺少所述第一数据源的数据; 以及

将向所述结构代表的引用包括在向所述查询计划器的请求中。

19. 根据权利要求1所述的方法, 其中, 所述操作集合是由所述运行系统生成的。

20. 根据权利要求1所述的方法, 其中, 所述操作集合是由除所述运行系统以外的系统生成的。

21. 根据权利要求1所述的方法, 其中, 所述操作集合包括基于所述第二数据源的一些数据处理能力或所述第一数据源的一些数据处理能力而选择的至少一个数据处理操作。

22. 根据权利要求1所述的方法, 其中, 数据处理能力包括被配置为执行至少一个数据处理指令的可运行功能性。

23. 根据权利要求1所述的方法, 其中, 所述组合操作包括结合操作。

24. 根据权利要求1所述的方法, 其中, 所述数据概括操作包括聚合操作。

25. 一种计算机实现的方法, 用于运行查询, 所述方法包括以下步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源, 其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

26. 一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,基于所述查询来生成向查询计划器的请求包括:

产生所述第一数据源的结构代表,其中所述结构代表包括反映所述第一数据源的格式化的数据组织,并且缺少所述第一数据源的数据;以及

将向所述结构代表的引用包括在向所述查询计划器的请求中。

27. 一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,数据处理能力包括被配置为执行至少一个数据处理指令的可运行功能性。

28. 一种计算机系统,其包括一个或多个处理器并且用于运行查询,所述系统被配置为运行包括以下的操作:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,生成所述操作集合包括:

基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

至少一个链路,其代表所述第一节点的操作的输出数据被提供作为向所述数据流图形的第二节点的操作的输入数据。

29. 根据权利要求28所述的计算机系统,其中,所述第一数据源包括关系数据库表格。

30. 根据权利要求29所述的计算机系统,其中,所述第二数据源包括除关系数据库表格

以外的数据的一部分。

31. 根据权利要求29所述的计算机系统,其中,所述第二数据源包括平面文件。

32. 根据权利要求28所述的计算机系统,其中,所述操作还包括:接收所述查询的结果作为所述程序代码的输出。

33. 根据权利要求28所述的计算机系统,其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

34. 一种计算机系统,其包括一个或多个处理器并且用于运行查询,所述系统被配置为运行包括以下的操作:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

35. 一种非暂时性的计算机可读介质,其包括计算机可运行指令,所述计算机可运行指令用于执行查询,所述计算机可运行指令使计算机执行包括以下的步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少

一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,生成所述操作集合包括:

基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

至少一个链路,其代表所述第一节点的操作的输出数据被提供作为向所述数据流图形的第二节点的操作的输入数据。

36. 根据权利要求35所述的计算机可读介质,其中,所述第一数据源包括关系数据库表格。

37. 根据权利要求36所述的计算机可读介质,其中,所述第二数据源包括除关系数据库表格以外的数据的一部分。

38. 根据权利要求36所述的计算机可读介质,其中,所述第二数据源包括平面文件。

39. 根据权利要求35所述的计算机可读介质,其中,所述步骤还包括:接收所述查询的结果作为所述程序代码的输出。

40. 根据权利要求35所述的计算机可读介质,其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

41. 一种非暂时性的计算机可读介质,其包括计算机可运行指令,所述计算机可运行指令用于执行查询,所述计算机可运行指令使计算机执行包括以下的步骤:

接收查询;

基于所述查询来识别第一数据源;

基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从所述第一数据源和所述第二数据源接收数据;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联。

42. 一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:

接收与根据能够应用于关系数据库的查询语言所表示的查询相对应的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;

生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;

从第一数据源和第二数据源接收数据,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;以及

在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果,

其中,生成所述操作集合包括:基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

至少一个链路,其代表所述第一节点的操作的输出数据被提供作为向所述数据流图形的第二节点的操作的输入数据;

或者,

其中,所述第一数据源与用于对所述第一数据源所存储的数据执行操作的可运行功能性相关联,并且所述第二数据源不与用于对所述第二数据源所存储的数据执行操作的可运行功能性相关联;

或者,

其中,数据处理能力包括被配置为执行至少一个数据处理指令的可运行功能性。

43. 一种计算机实现的方法,用于对来自数据源的数据进行操作,所述数据存储在该数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述操作包括在除管理关系数据库的系统以外的运行系统上运行所述查询,所述运行包括:

接收查询;

基于所述查询来识别数据源,其中所述数据源与管理所述数据源的系统相关联;



识别除管理关系数据库的系统以外的运行系统；

基于所述查询来生成向查询计划器的请求；

将所述请求提供至所述查询计划器；

接收所述查询计划器基于所述请求所生成的查询计划，其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述；

基于所述查询计划来生成实例化数据流图形的数据结构，其中所述数据流图形包括：

第一节点，其代表要运行的至少一个操作，

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联，

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联，以及

所述操作是基于所述查询计划所描述的第一步骤而选择的，以及

第二节点，其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联，其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统；

从所述数据源接收数据；以及

在所识别的运行系统上运行基于所述数据流图形的程序代码。

44. 根据权利要求43所述的方法，其中，所述数据流图形还包括至少一个链路，其中所述至少一个链路代表所述第一节点的操作的输出数据被提供作为向所述数据流图形的第三节点的操作的输入数据。

45. 根据权利要求43所述的方法，其中，所述运行还包括：判断为管理所述数据源的系统能够执行基于所述查询计划所描述的所述第二步骤而选择的所述至少一个操作。

46. 根据权利要求43所述的方法，其中，生成实例化数据流图形的数据结构包括：将所述查询计划的一部分翻译为适合发送至管理所述数据源的系统的查询。

47. 根据权利要求43所述的方法，其中，所述第二节点与运行系统为了从所述数据源接收数据而能够使用的信息相关联。

48. 根据权利要求43所述的方法，其中，所述数据流图形还包括第三节点，所述第三节点代表除基于所述查询计划所描述的任何步骤而选择的操作以外的要运行的操作。

49. 根据权利要求43所述的方法，其中，运行基于所述数据流图形的程序代码包括：实例化所述数据流图形的节点和第二数据流图形的节点之间的链路，其中所述第二数据流图形包括代表除基于所述查询计划所描述的任何步骤而选择的操作以外的要运行的操作的至少一个节点。

50. 根据权利要求43所述的方法，其中，所述数据源包括关系数据库表格。

51. 根据权利要求43所述的方法，其中，所述数据源包括除关系数据库表格以外的数据的一部分。

52. 根据权利要求43所述的方法，其中，所述数据源包括平面文件。

53. 根据权利要求43所述的方法，其中，所述查询包括SQL查询。

54. 根据权利要求43所述的方法，其中，所述运行还包括以下步骤：接收所述查询的结果作为基于所述数据流图形的程序代码的输出。

55. 根据权利要求43所述的方法,其中,所述查询计划器包括查询优化器。

56. 根据权利要求43所述的方法,其中,所述查询计划包括所述查询中的参数的数据类型。

57. 根据权利要求43所述的方法,其中,所述查询包含向可运行功能的引用,并且所述数据流图形包括代表所述可运行功能的节点。

58. 根据权利要求43所述的方法,其中,基于所述查询来生成向查询计划器的请求包括:

产生所述数据源的结构代表,其中所述结构代表包括反映所述数据源的格式化的数据组织,并且缺少所述数据源的数据;以及

将向所述结构代表的引用包括在向所述查询计划器的请求中。

59. 根据权利要求43所述的方法,其中,所述数据结构是由所识别的运行系统生成的。

60. 根据权利要求43所述的方法,其中,所述数据结构是由除所识别的运行系统以外的系统生成的。

61. 一种计算机系统,包括一个或多个处理器,所述一个或多个处理器用于对来自数据源的数据进行操作,所述数据存储在所述数据源的有形的非暂时性的计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述系统被配置为运行包括以下的操作:

接收查询;

基于所述查询来识别数据源,其中所述数据源与管理所述数据源的系统相关联;

识别除管理关系数据库的系统以外的运行系统;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;

基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;

从所述数据源接收数据;以及

在所识别的运行系统上运行基于所述数据流图形的程序代码。

62. 根据权利要求61所述的计算机系统,其中,所述操作还包括:判断为管理所述数据源的系统能够执行基于所述查询计划所描述的所述第二步骤而选择的所述至少一个操作。

63. 根据权利要求61所述的计算机系统,其中,生成实例化数据流图形的数据结构包

括:将所述查询计划的一部分翻译为适合发送至管理所述数据源的系统的查询。

64. 根据权利要求61所述的计算机系统,其中,所述第二节点与运行系统为了从所述数据源接收数据而能够使用的信息相关联。

65. 根据权利要求61所述的计算机系统,其中,所述数据源包括关系数据库表格。

66. 根据权利要求61所述的计算机系统,其中,所述数据源包括除关系数据库表格以外的数据的一部分。

67. 根据权利要求61所述的计算机系统,其中,所述操作还包括:接收所述查询的结果作为基于所述数据流图形的程序代码的输出。

68. 一种非暂时性计算机可读介质,其包括计算机可运行指令,所述计算机可运行指令用于对来自数据源的数据进行操作,所述数据存储在所述数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述操作包括在除管理关系数据库的系统以外的运行系统上运行所述查询,所述计算机可运行指令使计算机执行包括以下的步骤:

接收查询;

基于所述查询来识别数据源,其中所述数据源与管理所述数据源的系统相关联;

识别除管理关系数据库的系统以外的运行系统;

基于所述查询来生成向查询计划器的请求;

将所述请求提供至所述查询计划器;

接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;

基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;

从所述数据源接收数据;以及

在所识别的运行系统上运行基于所述数据流图形的程序代码。

69. 根据权利要求68所述的计算机可读介质,其中,所述步骤还包括:判断为管理所述数据源的系统能够执行基于所述查询计划所描述的所述第二步骤而选择的所述至少一个操作。

70. 根据权利要求68所述的计算机可读介质,其中,生成实例化数据流图形的数据结构包括:将所述查询计划的一部分翻译为适合发送至管理所述数据源的系统的查询。

71. 根据权利要求68所述的计算机可读介质,其中,所述第二节点与运行系统为了从所述数据源接收数据而能够使用的信息相关联。

72. 根据权利要求68所述的计算机可读介质,其中,所述数据源包括关系数据库表格。

73. 根据权利要求68所述的计算机可读介质,其中,所述数据源包括除关系数据库表格以外的数据的一部分。

74. 根据权利要求68所述的计算机可读介质,其中,所述步骤还包括:接收所述查询的结果作为基于所述数据流图形的程序代码的输出。

75. 一种计算机系统,其包括一个或多个处理器,所述一个或多个处理器用于对来自数据源的数据进行操作,所述数据存储在该数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述系统包括:

用于接收查询的部件;

用于基于所述查询来识别数据源的部件,其中所述数据源与管理所述数据源的系统相关联;

用于识别除管理关系数据库的系统以外的运行系统的部件;

用于基于所述查询来生成向查询计划器的请求的部件;

用于将所述请求提供至所述查询计划器的部件;

用于接收所述查询计划器基于所述请求所生成的查询计划的部件,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;

用于基于所述查询计划来生成实例化数据流图形的数据结构的部件,其中所述数据流图形包括:

第一节点,其代表要运行的至少一个操作,

所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,

所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及

所述操作是基于所述查询计划所描述的第一步骤而选择的,以及

第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;

用于从所述数据源接收数据的部件;以及

用于在所识别的运行系统上运行基于所述数据流图形的程序代码的部件。

## 管理数据查询

[0001] (本申请是申请日为2012年4月30日、申请号为201280021472.4、发明名称为“管理数据查询”的申请的分案申请)

### 技术领域

[0002] 该说明书涉及管理数据查询。

### 背景技术

[0003] 数据可存储在数据库中并按照诸如数据库表格的各种形式来安排。数据库表格可包括具有公共主题或目的的数据的集合。数据库表格的安排可通过数据库方案来定义,并且多个数据库表格可具有相似或相同安排。此外,当调整、附加或删除数据时,数据库的内容及其关联数据库表格可随着时间而改变。可使用各种技术来将数据传递到数据库里面和外面,并操纵数据库中的数据。

### 发明内容

[0004] 在一个方面,一般来说,用于生成代表数据库查询的数据流图形(graph)的方法包括从计划发生器接收查询计划,该查询计划代表用于对于代表数据源的至少一个输入运行数据库查询的操作,从该查询计划产生数据流图形,其中该数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路,并基于代表数据源的至少一个输入的至少一个特性来改变该数据流图形的一个或多个组件。

[0005] 各方面可包括以下特征中的一个或多个。改变该数据流图形的一个或多个组件包括去除该数据流图形的至少一个组件。该图形的组件对应于该查询计划所代表的操作。所述至少一个操作能够由与所述至少一个输入所代表的数据源关联的可运行功能性执行。所述至少一个输入的特性包括与该输入所代表的数据源关联的可运行功能性。该可运行功能性包括分类功能性。改变该数据流图形的一个或多个组件的步骤包括将代表可应用于第一输入的第一操作的第一组件和代表可应用于第二输入的第二操作的第二组件合并为单一组件,该单一组件代表等效于该第一操作和该第二操作的单一操作,该单一操作可应用于该第一输入和该第二输入。改变该数据流图形的一个或多个组件的步骤包括将代表第一数据源的第一输入组件和代表第二数据源的第二输入组件合并为单一输入组件,该单一输入组件包括访问该第一数据源和该第二数据源的数据的操作。改变该数据流图形的一个或多个组件的步骤包括去除代表可应用于第一输入的第一操作的第一组件,并向该第一输入应用第二组件,该第二组件代表对应于与该第一操作相同的功能性的第二操作,并且该第二操作可应用于第二输入。该方法还包括标识和与代表所述至少一个数据源的组件关联的数据库关联的功能性,并基于该标识,配置该组件以向该数据库提供数据库查询。所述代表数据源的至少一个输入包括数据文件、数据库表格、第二数据流图形的输出、和网络套接字中的至少一个。将该数据流图形的输出分配到数据文件、数据库表格、第二数据流图形、和网

络套接字中的至少一个。该数据库查询包括SQL查询。该数据流图形包括被配置为从该计划发生器接收输出的组件。

[0006] 在另一方面,一般,一种计算机可读介质存储有用于生成代表数据库查询的数据流图形的计算机程序,并且该计算机程序包括指令,用于促使计算机:从计划发生器接收查询计划,该查询计划代表用于对于代表数据源的至少一个输入运行数据库查询的操作;从该查询计划产生数据流图形,其中该数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路;和基于代表数据源的至少一个输入的至少一个特性,来改变该数据流图形的一个或多个组件。

[0007] 在另一方面,一般,一种用于生成代表数据库查询的数据流图形的系统,该系统包括:用于从计划发生器接收查询计划的部件,该查询计划代表用于对于代表数据源的至少一个输入运行数据库查询的操作;用于从该查询计划产生数据流图形的部件,其中该数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路;和用于基于代表数据源的至少一个输入的至少一个特性、来改变该数据流图形的一个或多个组件的部件。

[0008] 在另一方面,一般,一种用于生成代表数据库查询的数据流图形的系统,该系统包括:计算机系统,被配置为从计划发生器接收查询计划,该查询计划代表用于对于代表数据源的至少一个输入运行数据库查询的操作;从该查询计划产生数据流图形,其中该数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路;和基于代表数据源的至少一个输入的至少一个特性,来改变该数据流图形的一个或多个组件。

[0009] 在另一方面,一般,一种用于生成代表数据库查询的数据流图形的系统,该系统包括:接口,与提供查询计划的计划发生器通信,该查询计划代表用于对于代表数据源的至少一个输入运行数据库查询的操作;和处理器,被配置为从该计划发生器提供的查询计划产生数据流图形,其中该数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路,和基于代表数据源的至少一个输入的至少一个特性,来改变该数据流图形的一个或多个组件。

[0010] 在另一方面,一般,一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:接收查询;基于所述查询来识别第一数据源;基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;从所述第一数据源和所述第二数据源接收数据;以及在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果。

[0011] 在另一方面,一般,一种系统,其包括一个或多个处理器并且用于运行查询,所述系统被配置为运行包括以下的操作:接收查询;基于所述查询来识别第一数据源;基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;从所述第一数据源和所述第二数据源接收数据;以及在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果。

[0012] 在另一方面,一般,一种非暂时性的计算机可读介质,其包括计算机可运行指令,所述计算机可运行指令用于执行查询,所述计算机可运行指令使计算机执行包括以下的步骤:接收查询;基于所述查询来识别第一数据源;基于所述查询来识别第二数据源,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;从所述第一数据源和所述第二数据源接收数据;以及在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果。

[0013] 在另一方面,一般,一种计算机实现的方法,用于运行查询,所述方法包括以下步骤:接收与根据能够应用于关系数据库的查询语言所表示的查询相对应的查询计划,其中所述查询计划包括由管理关系数据库的系统执行的一个或多个步骤的描述;生成包括用于运行的至少一个数据处理操作的操作集合,所述至少一个数据处理操作包括数据选择操作、分类操作、数据概括操作和组合操作中的至少一个操作,其中所述至少一个数据处理操作是基于所述查询计划所描述的特定步骤而选择的,并且所述特定步骤与数据选择操作、分类操作、数据概括操作和组合操作中的所述至少一个操作相对应;从第一数据源和第二数据源接收数据,其中所述第二数据源同与所述第一数据源的至少一些数据处理能力不同的至少一些数据处理能力相关联;以及在除管理关系数据库的系统以外的运行系统上运行基于所述操作集合的程序代码,使得运行所述程序代码产生所述查询的结果。

[0014] 在另一方面,一般,一种计算机实现的方法,用于对来自数据源的数据进行操作,所述数据存储在该数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述操作包括在除管理关系数据库的系统以外的运行系统上运行所述查询,所述运行包括:接收查询;基于所述查询来识别数据源,其

中所述数据源与管理所述数据源的系统相关联;识别除管理关系数据库的系统以外的运行系统;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:第一节点,其代表要运行的至少一个操作,所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及所述操作是基于所述查询计划所描述的第一步骤而选择的,以及第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;从所述数据源接收数据;以及在所识别的运行系统上运行基于所述数据流图形的程序代码。

[0015] 在另一方面,一般,一种系统,包括一个或多个处理器,所述一个或多个处理器用于对来自数据源的数据进行操作,所述数据存储在该数据源的有形的非暂时性的计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述系统被配置为运行包括以下的操作:接收查询;基于所述查询来识别数据源,其中所述数据源与管理所述数据源的系统相关联;识别除管理关系数据库的系统以外的运行系统;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:第一节点,其代表要运行的至少一个操作,所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及所述操作是基于所述查询计划所描述的第一步骤而选择的,以及第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;从所述数据源接收数据;以及在所识别的运行系统上运行基于所述数据流图形的程序代码。

[0016] 在另一方面,一般,一种非暂时性计算机可读介质,其包括计算机可运行指令,所述计算机可运行指令用于对来自数据源的数据进行操作,所述数据存储在该数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述操作包括在除管理关系数据库的系统以外的运行系统上运行所述查询,所述计算机可运行指令使计算机执行包括以下的步骤:接收查询;基于所述查询来识别数据源,其中所述数据源与管理所述数据源的系统相关联;识别除管理关系数据库的系统以外的运行系统;基于所述查询来生成向查询计划器的请求;将所述请求提供至所述查询计划器;接收所述查询计划器基于所述请求所生成的查询计划,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;基于所述查询计划来生成实例化数据流图形的数据结构,其中所述数据流图形包括:第一节点,其代表要运行的至少一个操作,所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相



关联,所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及所述操作是基于所述查询计划所描述的第一步骤而选择的,以及第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;从所述数据源接收数据;以及在所识别的运行系统上运行基于所述数据流图形的程序代码。

[0017] 在另一方面,一般,一种系统,其包括一个或多个处理器,所述一个或多个处理器用于对来自数据源的数据进行操作,所述数据存储在所述数据源的有形的非暂时性计算机可读介质中,所述操作基于根据能够应用于关系数据库的查询语言所表示的查询,所述系统包括:用于接收查询的部件;用于基于所述查询来识别数据源的部件,其中所述数据源与管理所述数据源的系统相关联;用于识别除管理关系数据库的系统以外的运行系统的部件;用于基于所述查询来生成向查询计划器的请求的部件;用于将所述请求提供至所述查询计划器的部件;用于接收所述查询计划器基于所述请求所生成的查询计划的部件,其中所述查询计划包括管理关系数据库的系统要执行的一个或多个步骤的描述;用于基于所述查询计划来生成实例化数据流图形的数据结构部件,其中所述数据流图形包括:第一节点,其代表要运行的至少一个操作,所述第一节点与运行系统为了调用可运行程序代码以执行所述操作而能够使用的信息相关联,所述第一节点与运行系统为了使数据能够用于该程序代码而能够使用的信息相关联,以及所述操作是基于所述查询计划所描述的第一步骤而选择的,以及第二节点,其与运行系统为了调用如下的可运行程序代码而能够使用的信息相关联,其中该可运行程序代码使得将用以执行基于所述查询计划所描述的第二步骤而选择的至少一个操作的请求发送至管理所述数据源的系统;用于从所述数据源接收数据的部件;以及用于在所识别的运行系统上运行基于所述数据流图形的程序代码的部件。

## 附图说明

- [0018] 图1是数据流图形的框图。
- [0019] 图2图示了数据库查询和数据流图形。
- [0020] 图3是数据库系统和关联组件的概括。
- [0021] 图4代表数据流图形的运行。
- [0022] 图5是数据库表格和数据库表格的结构代表。
- [0023] 图6图示了其每一个包括可运行功能的数据库查询和数据流图形。
- [0024] 图7是数据库系统和关联组件的概括。
- [0025] 图8图示了查询计划和数据流图形。
- [0026] 图9图示了数据库查询用户界面。
- [0027] 图10—11图示了查询计划和数据流图形。
- [0028] 图12是用于数据库系统的操作的流程图。

## 具体实施方式

- [0029] 1数据库、查询和图形
- [0030] 数据库管理系统处置一个或多个数据库中存储的数据。为了存储这样的数据,数

数据库储存器能采用可作为组织为诸如行和列的数据结构的数据组的、诸如数据库表格的一个或多个形式。在一个图中，每一行代表数据的记录，而每一列代表每一行中的区域(field)。

[0031] 数据库中包括的信息可使用数据库查询来访问和处理(例如,修改)。数据库查询是描述数据库内容的子集和对该子集中的数据采取的动作的指令的集合。例如,一些数据库系统执行按照诸如结构化查询语言(SQL)的专用数据库查询语言书写的数据库查询。在这些数据库系统中,SQL查询是用于操纵数据库的内容的主要手段。

[0032] 在一些实现中,数据库查询和其他计算在图形代表中与数据库管理系统关联。例如,数据可被代表为通过一组操作(称为数据流)。在一个安排中,数据流可通过直接图形提供,其中计算的组件与图形的顶点关联,并且组件之间的数据流对应于图形的链路(弧、边缘)。图形是模块实体,并且可与其他模块图形连接或组合。每一图形可由一个或多个图形组成,并且特定图形可以是较大图形中的组件。图形开发环境(GDE)提供用于指定可运行图形并定义图形组件的参数的用户界面。实现这样的基于图形的计算的系统在美国专利第5966072号EXECUTING COMPUTATIONS EXPRESSED AS GRAPHS中描述,通过引用将其全部合并在这里。

[0033] 参考图1,数据流图形101(如示范用户界面中代表的)的示例包括输入组件102,用于提供数据流图形101的可运行组件104a-104j所要处理的一组数据。在一些示例中,输入组件102是能包括与数据库系统关联的数据记录或者与交易处理系统关联的交易的数据库源。数据记录可按照诸如数据库表格的各种形式来存储。每一可运行组件104a-104j与整个数据流图形101所定义的计算的一部分关联。工作元素(例如,来自数据组或数据库表格的单独数据记录)进入组件的一个或多个输入端口,并且输出工作元素(在一些情况下,是输入工作元素,或输入工作元素的已处理版本)典型地离开该组件的一个或多个输出端口。在数据流图形101中,来自组件104e、104g和104i的输出工作元素被存储在输出数据组件102a-102c中。

[0034] 输入组件102和输出数据组件102a-102c中的每一个例如能提供对于数据文件、数据库表格、第二数据流图形的输出、网络套接字、或另一数据源中的任一个的访问。

[0035] 使用一些基于图形的数据库系统来处理数据库查询。例如可将数据库查询应用到一个或多个数据库表格,以提取例如用于在数据流中处理的数据库表格内容的所标识的子集。在一些实现中,基于图形的数据库系统接受和运行数据流图形形式的数据库查询。其他数据库系统可使用其他类型数据库查询。

[0036] 在一些情况下,将一个或多个数据库表格从使用SQL的另一类数据库系统移动到基于图形的数据库系统。所述另一数据库系统可具有与基于图形的数据库系统不兼容的已书写的许多SQL查询。在一些实现中,可将SQL查询变换为与基于图形的数据库系统兼容的数据库查询。

[0037] 图2示出了按照SQL书写的数据库查询200的示例。该示例数据库查询200意欲在基于图形的数据库管理系统所管理的数据库表格202、204上操作。基于图形的数据库管理系统可识别数据库表格202、204中的数据,因为数据库表格具有基于图形的数据库管理系统所固有(native)的格式。数据库表格202、204可包括从其他数据库表格(例如源自识别SQL查询并不使用数据流图形的数据库系统的数据库表格)获得的数据。

[0038] 然而,基于图形的数据库管理系统可以不具有用于处理SQL查询的内建功能性,这样可产生基于图形的数据库查询以模拟SQL数据库查询200。基于图形的数据库查询由基于图形的数据库系统识别。例如,数据库查询200可(使用这里描述的技术)从SQL查询变换206为数据流图形208。数据流图形208通过接受数据库表格202、204作为输入而对它们进行操作,并提供数据库查询200的运行结果作为输出。

#### [0039] 2查询计划

[0040] 一些数据库系统基于查询计划(有时也称为解释计划)进行诸如SQL查询的数据库查询。查询计划是如果运行数据库查询可执行的数据库操作的描述。该查询计划可描述数据库操作的一个可能安排,即使其他操作安排或不同操作集合将实现相同结果。

[0041] 为了提供这样的查询计划,数据库系统可包括查询计划发生器(有时也称为查询计划器)。例如,当正运行数据库查询时,查询计划发生器能产生查询计划,或者在进行关于运行查询的任何决定之前,查询计划发生器能生成查询计划。

[0042] 在一些安排中,数据库操作可按照各种顺序运行,同时仍提供等效输出。这样,查询计划发生器可具有确定最佳查询计划的功能性。例如,最佳查询计划可以是描述以下数据库操作的安排的查询计划,所述数据库操作用于在最小数量时间中运行数据库查询、或使用诸如数据储存空间的最小数量数据库资源、或按照别的方式在该数据库系统已标识的约束下实现该数据库查询。查询计划发生器的用于确定最佳查询计划的功能性可包括对许多可能查询计划进行打分或分级的功能性,并且还可以包括将可能查询计划重排为最佳或有效配置的功能性。

[0043] 单一数据库查询可被运行多次,并且每一运行可具有唯一最佳查询计划。例如,数据库表格中的数据可在数据库查询的两次运行之间改变。在该示例中,在数据库查询的第一次运行中生成的查询计划中描述的操作在数据库查询的第二次运行期间可能需要比第一次运行期间或多或少的运行时间。在该情况下,不同的操作安排可以更好地适应数据库查询的第二次运行,例如,相同操作的不同安排、或不同操作的安排。考虑到数据库表格的瞬时状态,可生成对于数据库查询的第二次运行优化的查询计划用于该运行。

#### [0044] 3系统概括

[0045] 在产生模拟诸如SQL查询的另一类数据库查询的基于图形的数据库查询时,可使用查询计划发生器。图3示出了用于准备数据库查询302以在数据库管理计算机系统304上运行的数据库查询管理系统300。示出的数据库管理计算机系统304包括图形运行引擎306,其处置实现为数据流图形的数据库操作。数据库查询管理系统300还包括图形生成计算机系统308,其具有图形生成引擎310,能由数据流图形要执行的操作的描述构建数据流图形312。例如,这些操作的描述可以是查询计划314。图形生成计算机系统308被配置有基于代表数据源的输入的一个或多个特性来改变数据流图形的组件的能力,如下面在描述“联合查询”的第6部分中更详细描述的那样。

[0046] 数据库查询管理系统300还包括查询计划计算机系统316,其运行查询计划发生器318。查询计划发生器318可以是数据库查询产生查询计划的任何查询计划发生器,并且不需要被设计有与数据流图形或图形生成有关的任何功能性。此外,数据库查询管理系统300还包括数据库计算机系统320,其在数据储存器(例如,硬盘驱动器、光盘等)中具有数据库322,并包括一个或多个数据库表格324a、324b和324c。

[0047] 尽管示出了单独计算机系统用于数据库管理计算机系统304、图形生成计算机系统308、查询计划计算机系统316、和数据库计算机系统320,但是这些计算机系统中的一个或多个可以是相同计算机系统、或相同计算机系统的组件。这些计算机系统全部具有用于运行它们的相应可运行组件的至少一个处理器和至少一个数据储存系统。计算机系统能使用诸如局域网 (LAN)、广域网 (WAN) 的计算机网络、诸如因特网的网络、或另一种类计算机网络彼此连接。

[0048] 为了表明从一个或多个数据库查询302产生数据流图形312,在数据流图形的运行之前,由图形生成计算机系统308接收和处理数据库查询和数据库表格326。图形生成计算机系统308从数据库计算机系统320接收数据库表格326。

[0049] 数据库表格326可采用几个形式中的任一个。例如,数据库表格326可以是关系数据库表格、部分数据库表格、平面文件、或另一类型数据文件或一组数据文件。在一些示例中,数据库表格326可按照关于数据库表格的信息的形式接收,例如关于数据库表格的元数据、或数据库表格的描述。

[0050] 在一些实现中,数据库表格326可通过与数据库计算机系统320关联或者按照别的方式可由图形生成计算机系统308访问的数据登记簿来标识。数据登记簿可以是查找文件目录的形式,例如其可包括与数据库表格326关联的数据文件位置、以及与数据库表格关联的主关键字和索引信息。数据登记簿还可以提供关于不同类型数据库表格的数据格式的信息。此外,数据登记簿还可以提供关于数据流图形312能如何访问数据库表格的信息。

[0051] 图形生成计算机系统308还接收要向数据库表格326应用的数据库查询302。例如,数据库查询可以是SQL查询。数据库查询302可从任意数目的可能源接收。例如,数据库查询302可从其中用户330已键入数据库查询的用户界面328接收。在一些示例中,数据库查询302从数据储存系统接收,或者该数据库查询从诸如因特网的计算机网络接收,或者该数据库查询基于另一先前接收的数据库查询来生成。

[0052] 在一些实现中,图形生成计算机系统308向查询计划计算机系统316提供(箭头332所代表的)从关于数据库表格326的信息产生的数据库表格326的版本。例如,向查询计划计算机系统316提供的数据库表格326的版本可以是尺寸比该数据库表格更小的数据库表格的结构代表334,并由此需要更少计算资源来处理。数据库表格326的结构代表334可包括关于数据库表格的信息,但是缺少数据库表格的一些或全部数据。例如,数据库表格326的结构代表334可包括反映数据库表格的格式的格式,诸如数据库表格的列、行或区域。数据库表格326的结构代表334还可包括关于数据的信息,诸如数据库表格中的元素的数据储存尺寸、或数据库表格中的元素的数据类型。

[0053] 图形生成计算机系统向查询计划计算机系统316提供(箭头332所代表的)数据查询302和数据库表格326的结构代表334。查询计划计算机系统316运行查询计划发生器318,该查询计划发生器318产生为了在数据库表格326上运行数据库查询而优化的查询计划。数据库表格326的结构代表334供应与数据库表格326自己将供应的信息相同的由查询计划发生器318使用的信息,例如数据尺寸、数据类型、以及关于数据库表格中包括的数据的其他信息。在一些实现中,数据库表格326的结构代表334包括用来优化数据查找和检索的数据元素的索引。查询计划发生器318可使用该索引来计算对来自被编索引的数据库表格326的数据元素进行标识和检索的速度。

[0054] 图形生成计算机系统308从查询计划计算机系统316接收查询计划314。查询计划314描述能用来在数据库表格326上运行数据库查询302的数据库操作的安排。例如,查询计划314中的操作能对应于数据流图形312的节点。查询计划314还可以包括关于查询计划中的操作所使用的数据类型的信息。例如,查询计划中的操作可具有参数,并且这些参数的数据类型可在该查询计划中描述。

[0055] 一旦产生,就将查询计划314提供到图形生成计算机系统308,用于由图形生成引擎310的数据流图形312产生。图形生成引擎310输出与该查询计划314对应的数据流图形312。在一些实现中,数据流图形312具有代表查询计划中描述的操作的节点、以及代表操作之间的数据流动的节点链路。因为数据流图形312可以对于准备数据库查询用于运行的每一实例生成,所以图形生成引擎310能足够快地生成数据流图形,以对运行数据库查询的实时请求作出应答。在一些实现中,图形生成引擎310可在小于一秒中从查询计划生成数据流图形。

[0056] 图形生成计算机系统308将图形生成引擎316所生成的数据流图形312提供(如箭头336所代表的)到数据库管理计算机系统304。在一些实现中,图形生成计算机系统308还准备数据库表格326用于由数据库管理计算机系统使用,并提供准备的数据库表格338。例如,图形生成计算机系统308可将数据库表格326从图形生成计算机系统308所使用的格式变换为数据库管理计算机系统304所使用的格式。

[0057] 一旦提供到数据库管理计算机系统304,数据流图形312就准备运行。如图4中所示,数据库管理计算机系统304可运行数据流图形312的操作并使用数据库表格326以便产生数据库查询的结果402。数据库管理计算机系统304将数据库表格326提供到数据流图形312的一个或多个节点404a、404b、404c,并使用图形运行引擎306来运行该数据流图形。图形运行引擎306执行数据流图形312的节点404a、404b、404c所代表的操作,这些操作对应于用于运行下层数据库查询的数据库操作。此外,节点之间的链路408a、408b、408c代表当处理数据库表格时在数据库操作之间的数据流动。数据流图形312输出数据库查询的结果402。

[0058] 4结构代表

[0059] 图5示出了数据库表格500和数据库表格的结构代表502(不包括来自数据库表格的数据)的示例。数据库表格500具有列504a-504i,其可包括相似或不同类型数据。数据库表格500还具有行508a-508e,其每一个包括与每一列对应的区域。行的每一区域包括对应列的数据类型的数据元素510a、510b、510c(例如,字符串、整数、浮点数等)。此外,每一行508a-508e具有固有数据储存尺寸512a-512e。例如,数据储存尺寸512a-512e可以是当该行驻留在诸如存储器的数据储存器中时、由该行的数据元素使用的储存空间的数量。

[0060] (图3中示出的图形生成计算机系统308所产生的)数据库表格500的结构代表502具有和数据库表格相同的列514a-514i,包括和原始列504a-504i相同的数据类型。数据库表格500的结构代表502还具有和数据库表格的行508a-508e对应的行518a-518e。然而,行518a-518e不包括来自数据库表格500的数据元素510a、510b、510c。每一行518a-518e与数据储存尺寸值520a-520e关联。尽管数据库表格500中的行的数据储存尺寸512a-512e可根据各个数据元素510a、510b、510c的数据储存尺寸来计算,但是每一行518a-518e的数据储存尺寸值520a-520e自己可以是在数据库表格的结构代表502中沿着每一行存储的数据元

素。数据库表格500的结构代表502的总数据储存尺寸可以占到数据库表格的总数据储存尺寸的小百分比,因为数据库表格的结构代表的区域522a、522b、522c缺少来自数据库表格的区域的数据元素510a、510b、510c。在一些实现中,数据库表格500的结构代表502可保持来自数据库表格的一些数据元素,例如,数据库表格的结构代表可保持与关键字值524、526对应的数据元素、或者在数据库表格的结构、组织或格式中使用的其他数据元素。在一些实现中,数据库表格500的结构代表502可包括提供在数据检索中使用的信息的索引或另一数据结构。在一些实现中,数据库表格500的结构代表502可包括关于数据库表格中包括的数据的其他统计值。

#### [0061] 5函数调用

[0062] 如图6中所示,在一些实现中,定制图形函数602能嵌入在数据库查询604中。例如,定制图形函数602可代表在数据流图形的运行期间数据流图形606能进行的操作。定制图形函数602可以不具有在数据库查询604的上下文中的任何功能性,并被放置在要传递给生成的数据流图形606的数据库查询中。例如,定制图形函数602可被添加到数据库查询604以准备它用于变换到数据流图形,只是替换或增加现有数据库查询的部分。此外,查询计划发生器可以不具有关于定制图形函数602的功能性的信息。定制图形函数602可以是不具有数据库查询604的语言中的等效函数的函数。在一些实现中,查询计划发生器可知道定制图形函数602的输入数据类型或输出数据类型。当生成608数据流图形606时,定制图形函数602可充当数据流图形的节点610。

[0063] 在一些实现中,定制图形函数602可以是用来访问特定或定制格式的数据的函数,或者定制图形函数可以是用来运行规则表达式或图案匹配表达式的函数,或者定制图形函数可以是例如数据库管理计算机系统的用户的用户所实现的函数。

#### [0064] 6联合查询

[0065] 在一些实现中,查询计划发生器(例如,图3中示出的查询计划发生器318)能对这样的数据库查询操作,该数据库查询被构造为对从多个数据源和/或多类数据源获取的输入数据进行操作。例如,该数据库查询可被构造为对多个数据库表格操作,并且数据库表格可属于不同类型并具有不同内部功能性的数据库。而且,该数据库查询可被构造为对其他数据源中可访问的数据进行操作,诸如多个数据文件、多个网络套接字、其他数据流图形的多个输出、或这些类型数据源或使用数据库查询能访问的其他类型数据源的任何组合。当图形生成引擎从查询计划生成数据流图形时,可取决于数据源的数目和相应类型,基于相应数据源的特性,来优化得到的数据流图形,用于访问代表数据源的每一输入。

[0066] 图7示出了包括图形生成引擎908和图形优化器704的图形生成计算机系统706。图形生成计算机系统706接收一个或多个数据库表格726a、726b。例如,图形生成计算机系统706可从一个或多个数据库计算机系统(例如,在数据储存系统(例如,硬盘驱动器、光盘等)中具有数据库722的数据库计算机系统720)接收数据库表格726a、726b。图形生成计算机系统706还接收要向数据库表格726a、726b应用的数据库查询702。例如,数据库查询702可从用户界面接收,或者数据库查询702可从数据储存系统接收,或者数据库查询可从诸如因特网的计算机网络接收,或者数据库查询可基于另一先前接收的数据库查询而生成。

[0067] 图形生成计算机系统706向查询计划计算机系统746提供(如箭头932所代表的那样)数据库表格726a、726b的版本734a、734b。在一些实现中,数据库表格726a、726b的版本

734a、734b是数据库表格726a、726b自己。在一些实现中，数据库表格726a、726b的版本734a、734b是数据库表格726a、726b的结构代表。例如，数据库表格726a、726b的版本734a、734b可以是按照与图5中示出的数据库表格500的结构代表502相同的方式格式化的结构代表。

[0068] 图形生成计算机系统706向查询计划计算机系统746提供(如箭头932所代表的那样)数据库查询702和数据库表格726a、726b的版本734a、734b。查询计划计算机系统746运行查询计划发生器748,该查询计划发生器748产生用于在数据库表格726a、726b上运行数据库查询702的查询计划714。在一些实现中,如果数据库表格726a、726b的版本734a、734b是数据库表格726a、726b的结构代表,则所述结构代表供应与数据库表格726a、726b自己将供应的信息相同的由查询计划发生器748使用的信息。

[0069] 图形生成计算机系统706从查询计划计算机系统746接收查询计划714。与图3中示出的查询计划314相同,该查询计划714描述能用来在数据库表格726a、726b上运行数据库查询702的数据库操作的安排。图形生成计算机系统706使用该查询计划714作为输入来运行图形生成引擎908。图形生成引擎908基于该查询计划714的内容生成数据流图形716。在一些实现中,数据流图形716具有代表该查询计划714中描述的操作的节点,并且节点链路代表这些操作之间的数据的流动。

[0070] 数据流图形716可由图形优化器704变换为优化的数据流图形718。在一些实现中,数据流图形718可由计算机系统(例如,数据库管理计算机系统744)运行,以进行与该查询计划714定义的操作对应的操作。数据流图形716进行的一些操作可以是多余的,并能去除或与其他操作合并。例如,数据流图形716可包括能合并为单一组件的一组组件,所述单一组件执行与所述一组组件的操作相同的操作。图形优化器704分析数据流图形716并执行优化,以改变数据流图形,例如,去除多余组件、合并组件、以及按照别的方式减少数据流图形716中的组件的数目以生成优化的数据流图形718。在生成优化的数据流图形718之后,可将优化的数据流图形718提供到数据库管理计算机系统744用于运行。

[0071] 在一些实现中,数据流图形716和/或优化的数据流图形718包括这样的组件,该组件被配置为向数据流图形716的其他组件提供查询计划发生器748的输出(例如,查询计划714)。例如,数据流图形716或优化的数据流图形718可在运行(“生效(on the fly)”)期间接收数据库查询702,并且这样数据流图形716或优化的数据流图形718的组件也在运行期间能向查询计划发生器748请求与数据库查询702对应的查询计划714。查询计划发生器748的提供输出的组件然后能向数据流图形718的其他组件提供该查询计划714。

[0072] 在一些实现中,图形优化器704可在生成优化的数据流图形718的时候增加数据流图形716中的组件的数目。例如,如果用一组更有效的组件替换复杂组件,则在运行时,数据流图形716可操作更快或更有效。

[0073] 在一些实现中,图形生成计算机系统706包括查询计划优化器,用于对查询计划714操作,并在变换为数据流图形之前对查询计划714进行优化。

[0074] 图8示出了从查询计划806生成的数据流图形802和优化的数据流图形804的示例。查询计划806代表以下SQL查询:

[0075] 选择count(\*)作为num\_transactions,

[0076] 求和(transaction\_amt)作为总数,

[0077] account\_id

[0078] 从federated\_transactions

[0079] 根据account\_id分组

[0080] 根据account\_id排序

[0081] 该SQL查询从称为federated\_transactions的数据库表格中选择记录,并对它们进行分组和排序。查询计划806描述可执行以运行查询的三个操作。查询计划806的操作是数据选择操作810、分类操作812和称为组聚合操作814的数据概括操作。这些操作可被翻译为具有大致对应于这些操作的组件的数据流图形802。数据流图形802具有数据选择组件820、分类组件822、称为汇总(rollup)组件824的数据概括组件、以及输出组件826。

[0082] 数据选择组件820代表数据的源。例如,数据选择组件820可以是数据库表格、第二数据流图形的输出、网络套接字、或另一数据源。

[0083] 输出组件826代表数据目的地。例如,输出组件826可以是数据文件、数据库表格、第二数据流图形的输入、网络套接字、或输出数据的另一目的地中的任一个。

[0084] 数据选择组件820可代表具有可运行功能性的数据源。例如,数据选择组件820可代表数据库中的以下数据库表格,该数据库表格可运行数据库查询并基于诸如上面示出的SQL查询的运行的数据库查询来返回数据。数据流图形802可得到(have access to)与存储数据库表格的数据库关联的功能性。例如,数据流图形802可具有能向数据库发送执行与可运行功能性关联的操作的请求的组件。

[0085] 在一些实现中,SQL查询对属于该数据库的数据操作。例如,数据选择组件820所代表的federated\_transactions数据库表格可属于能执行与分类组件822和汇总组件824所执行的功能相同的功能的数据库。优化的数据流图形804能从利用该内建功能性的查询计划806生成。优化的数据流图形804仅具有两个组件:包括汇总功能性的数据选择组件830、和输出组件832。

[0086] 优化的数据流图形804可比从查询计划806直接生成的数据流图形802具有更少组件,并且优化的数据流图形804也可更有效地操作。例如,在示出的示例中,数据流图形802从federated\_transactions数据库表格选择796096个记录,一旦完成了汇总操作,则记录的数目减少为3754个记录。然而,优化的数据流图形804不需要采用选择和处理796096个记录的全部数目的步骤,因为其获取已分类并概括的3754个记录。

[0087] 在一些实现中,可通过将查询计划806的一部分翻译回适于具有优化的数据流图形804所利用的功能性的数据源的查询,来执行该优化。图9示出了与优化的数据流图形804的数据选择组件830关联的用户界面940的示例。用户界面940允许用SQL查询942来配置数据选择组件830。当运行与数据选择组件830关联的数据流图形804时,SQL查询942被传递到由数据选择组件830所代表的、并由与数据选择组件830关联的数据库功能性所运行的数据的源。例如,如果数据选择组件830与数据库计算机系统720(图7)所存储的数据库722关联,则SQL查询942可被传递到数据库计算机系统720用于运行。按照该方式,数据选择组件830能使用SQL查询942来访问与数据库722关联的功能性。

[0088] 代表具有多个数据源作为输入的查询计划的数据流图形可基于多个输入的特性来优化。图10示出了经受到优化的数据流图形1030的转换的查询计划1010。查询计划1010代表联合查询,并具有两个数据库表格1012、1014作为输入。查询计划能变换1016为数据流



图形1020。例如，图7中示出的图形生成引擎908能执行该变换。

[0089] 数据流图形1020包括代表数据库表格之一1012的数据选择组件1022和代表另一数据库表格1014的数据组件1026。数据选择组件1022将数据输出到汇总组件1024，而数据组件1026将数据输出到第二汇总组件1028。数据选择组件1022具有内部功能性，能用来执行与汇总组件1024在数据流图形1020中执行的操作相同的操作。例如，数据选择组件1022可代表具有可运行功能性的数据库系统，并能对其提供的数据库执行其自己的操作。例如，数据选择组件1022可代表在诸如图7中示出的数据库722的数据库上可访问的数据库表格。

[0090] 数据组件1026不具有能用来执行与汇总组件1028在数据流图形1020中执行的操作相同的操作的功能性。例如，数据组件1026可以是诸如不包括可运行功能性的数据文件的实体。

[0091] 数据流图形1020可(通过诸如图7中示出的图形优化器704的图形优化器)被变换1018为优化的数据流图形1030，其中已基于数据选择组件1022和数据组件1026的内部功能性而去除了一些组件。因为数据选择组件1022包括汇总功能性，所以数据选择组件1022及其关联汇总组件1024能被合并为组合数据选择组件1032，该组合数据选择组件1032并入该汇总组件1024按照别的方式执行的汇总操作。组合数据选择组件1032可被提供有数据库查询，指示该组合数据选择组件1032在将数据提供给优化的数据流图形1030之前执行汇总操作。例如，该组合数据选择组件1032能被提供诸如图8中示出的SQL查询942的数据库查询。与数据组件1026关联的汇总组件1028保留在优化的数据流图形1030中，因为数据组件1026不具有该汇总组件1028按照别的方式复制的内部汇总功能性。

[0092] 图11示出了经受到优化的数据流图形1130的转换的查询计划1110。查询计划1110代表联合查询，并具有两个数据库表格1112、1114作为输入。查询计划能变换1116为数据流图形1120。例如，图7中示出的图形生成引擎908能执行该变换。

[0093] 数据流图形1120包括与查询计划1110的操作对应的组件。该数据流图形1120包括代表查询计划1110的数据库表格1112、1114的数据源组件1122、1124，还包括结合组件1126，其对数据源组件1122、1124所提供的数据库操作以将数据库表格1112、1114的数据组合为单一数据源。

[0094] 数据库表格1112、1114可属于相同数据库，例如图7中示出的数据库722。数据库722可具有能提供两个数据库表格1112、1114的数据作为单一数据源(例如，作为结合操作的结果的单一数据库表格)的功能性。图形优化器(例如图7中示出的图形优化器704)能标识与数据库表格1112、1114关联的数据库722的功能性，并确定是否能通过减少数据流图形1120中的组件的数目来简化数据流图形1120。图形优化器704能确定该数据库722能对于数据库表格1112、1114执行结合操作，并生成包括合并数据源组件1132的优化的数据流图形1030，该合并数据源组件1132代表当数据库722对于这两个数据库表格1112、1114执行结合操作时该数据库722所提供的数据库。优化的数据流图形1030不包括结合组件1126，因为由数据库722取代执行该结合组件1126所代表的功能性。

[0095] 图12是示出了(图7中示出的)图形生成计算机系统706的示范操作的流程图1200。在步骤1202，图形生成计算机系统接收查询计划，诸如由查询计划计算机系统的查询计划发生器产生的查询计划。该查询计划代表用于运行对于代表数据源的至少一个输入的数据库查询的操作。在步骤1204，图形生成计算机系统使用图形生成引擎，来基于查询计划中包

括的操作从查询计划产生数据流图形。得到的数据流图形包括代表该查询计划所代表的至少一个操作的至少一个节点,并包括代表与该查询计划关联的至少一个数据流的至少一个链路。在步骤1206,图形生成计算机系统基于代表数据源的输入的特性来改变数据流图形的组件。可改变这些组件以优化数据流图形,例如减少数据流图形中的组件的数目。

[0096] 上面描述的数据库查询管理方案可使用用于在计算机上运行的软件来实现。例如,该软件形成一个或多个计算机程序中的过程,所述程序在一个或多个已编程或可编程计算机系统(其可以是诸如分布、客户机/服务器、或网格的各种架构)上运行,所述计算机系统中的一个包括至少一个处理器、至少一个数据储存系统(包括易失性和非易失性存储和/或储存元素)、至少一个输入装置或端口、以及至少一个输出装置或端口。该软件可形成较大程序的一个或多个模块,诸如,提供与计算图形的设计和配置相关的其他服务。该图形的节点和元素可被实现为计算机可读介质中存储的数据结构、或符合数据仓库中存储的数据模型的其他组织数据。

[0097] 该软件可被提供在一般或特定目的可编程计算机可读取的诸如CD-ROM的储存介质上,或者通过网络的通信介质向运行其的计算机传递(在广播信号中编码)。所有功能可在特定目的计算机上执行,或使用诸如协处理器的特定目的硬件来执行。该软件可按照分布方式实现,其中该软件所规定的计算的不同部分由不同计算机执行。每一这样的计算机程序优选被存储在或下载到一般或特定目的可编程计算机可读取的储存介质或装置上(例如,固态存储器或介质、或者磁或光介质),用于当所述储存介质或装置由计算机系统读取以执行其中描述的过程时,配置和操作计算机。本发明系统还可以被看作实现为配置有计算机程序的计算机可读储存介质,其中这样配置的储存介质促使计算机系统按照特定和预定义方式操作以执行其中描述的功能。

[0098] 已描述了本发明的多个实施例。但是,将理解的是,可进行各种修改,而不脱离本发明的精神和范围。例如,上面描述的一些步骤可以是顺序独立的,并由此能按照与描述的顺序不同的顺序执行。

[0099] 应理解的是,前述描述意欲是示意性的,而不限由所附权利要求的范围限定的本发明的范围。例如,上面描述的多个功能步骤可按照不同顺序执行,而基本上不影响总体处理。其他实施例在以下权利要求的范围中。

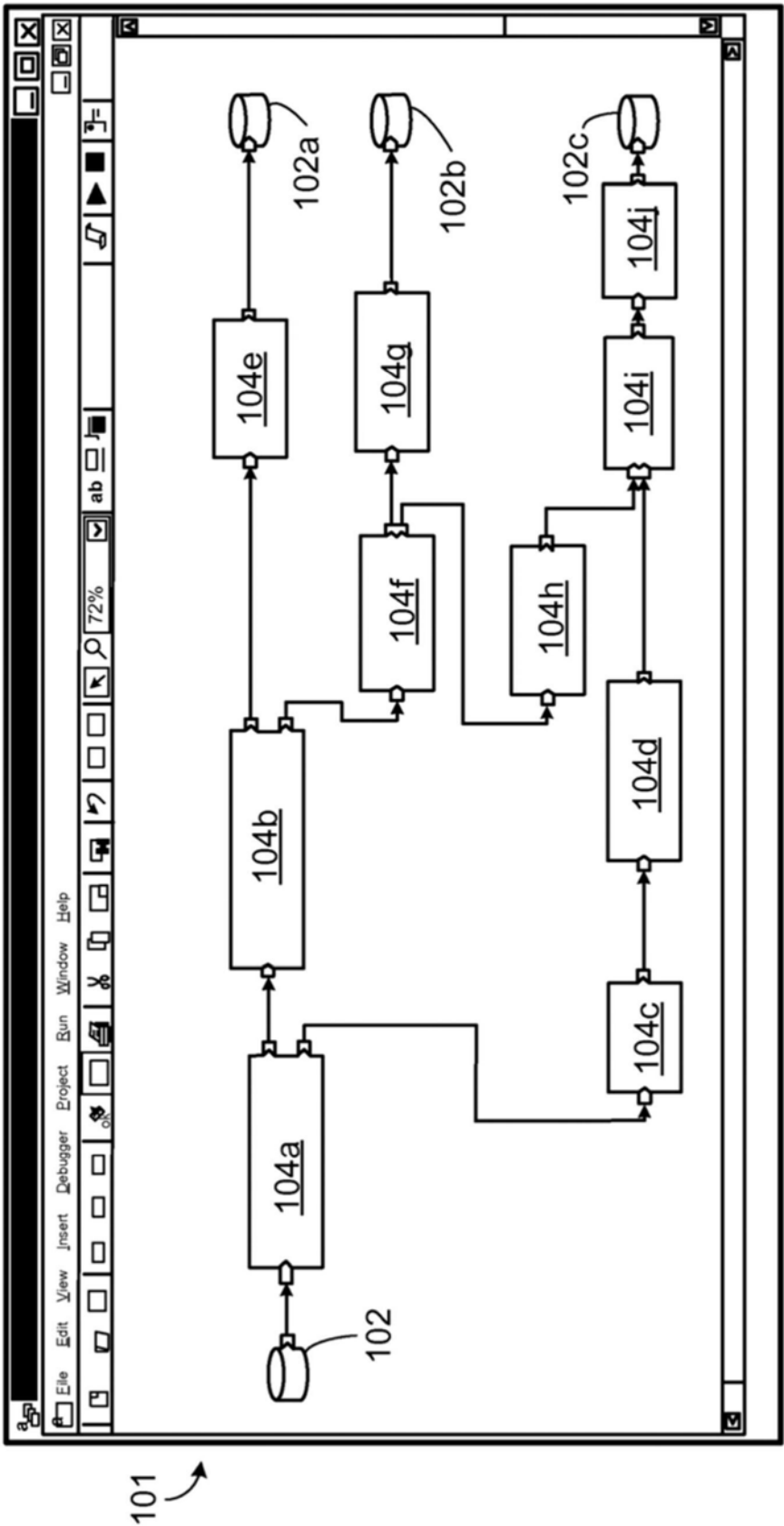


图1

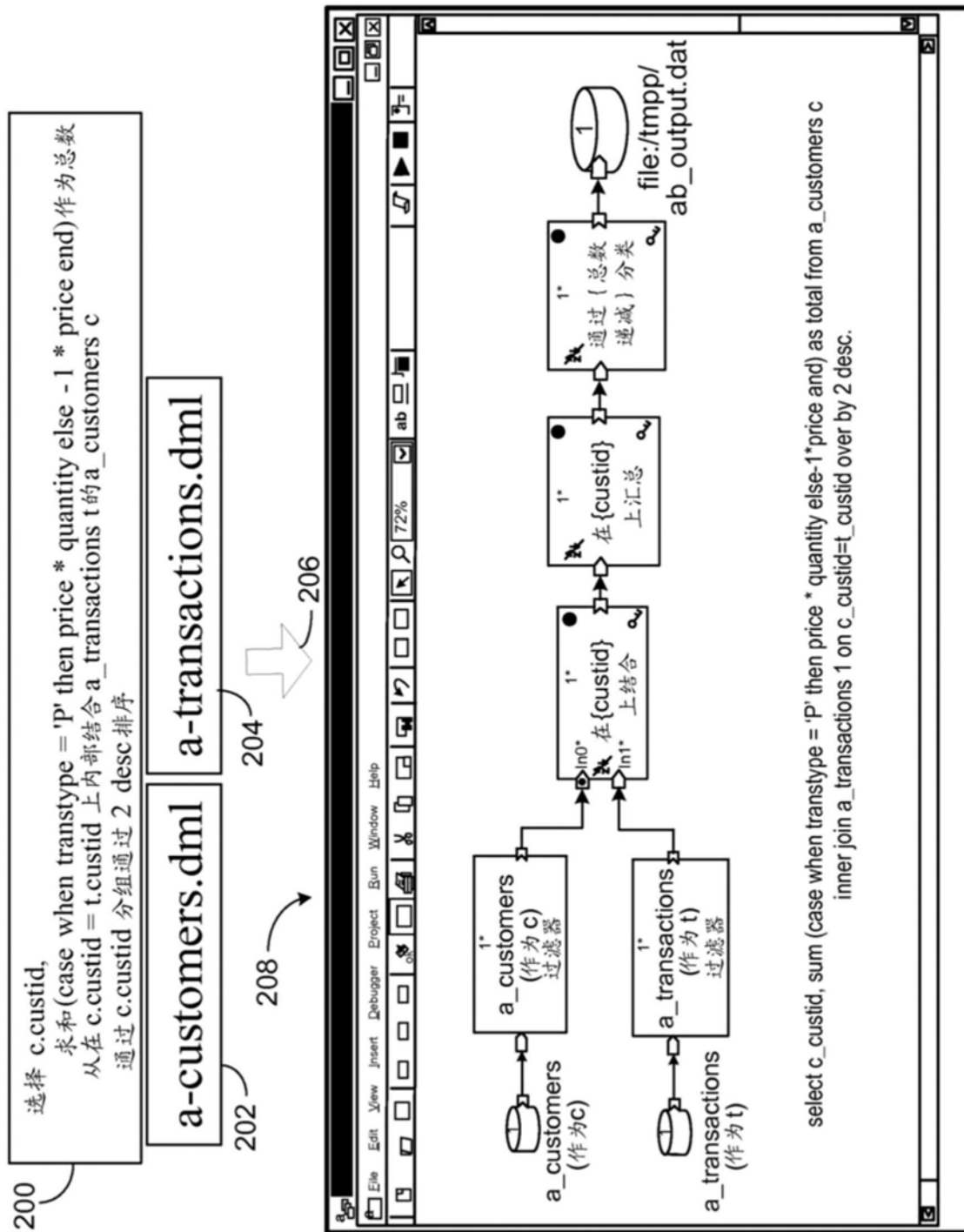


图2

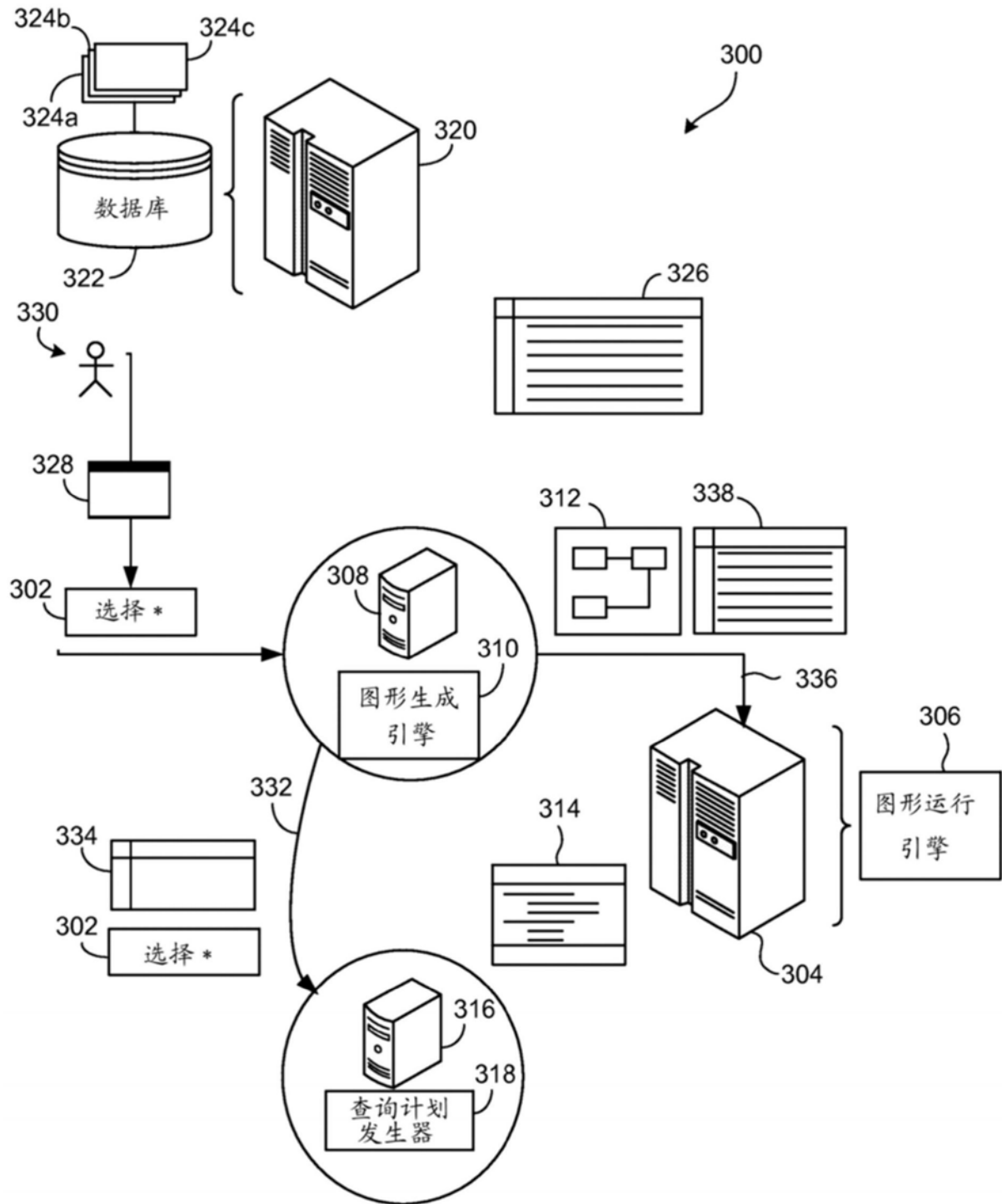


图3

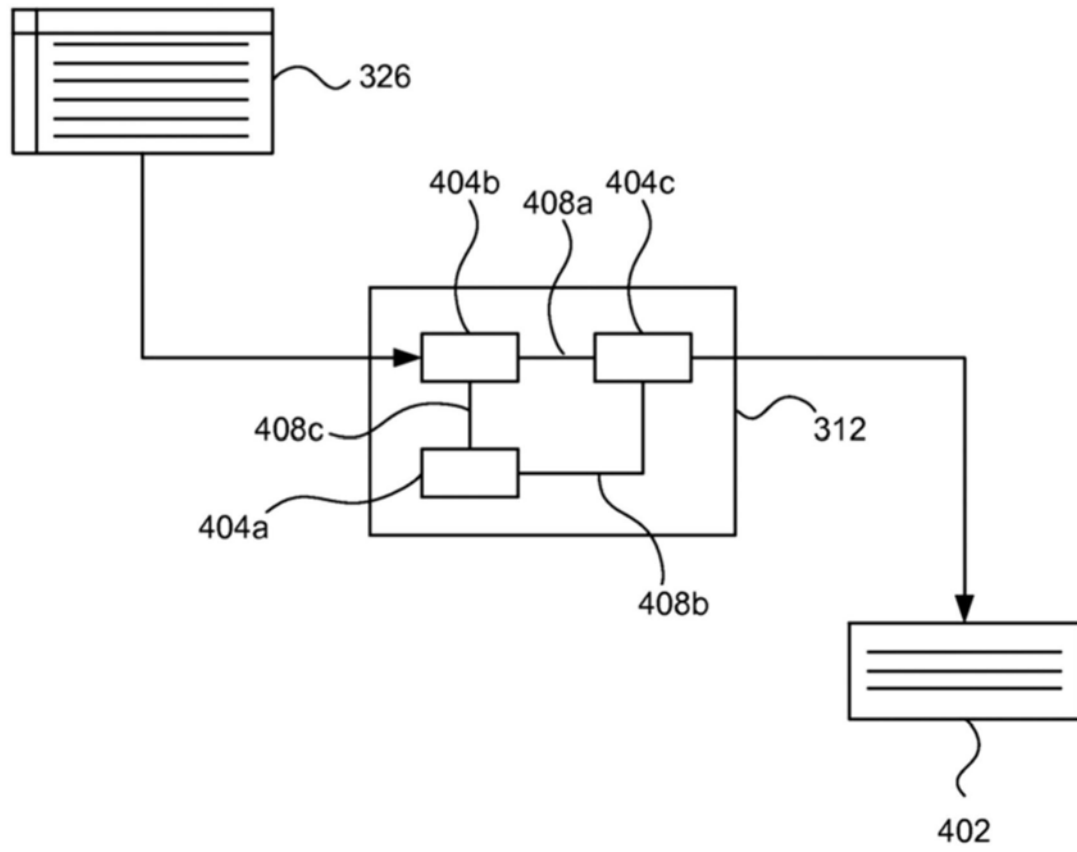


图4

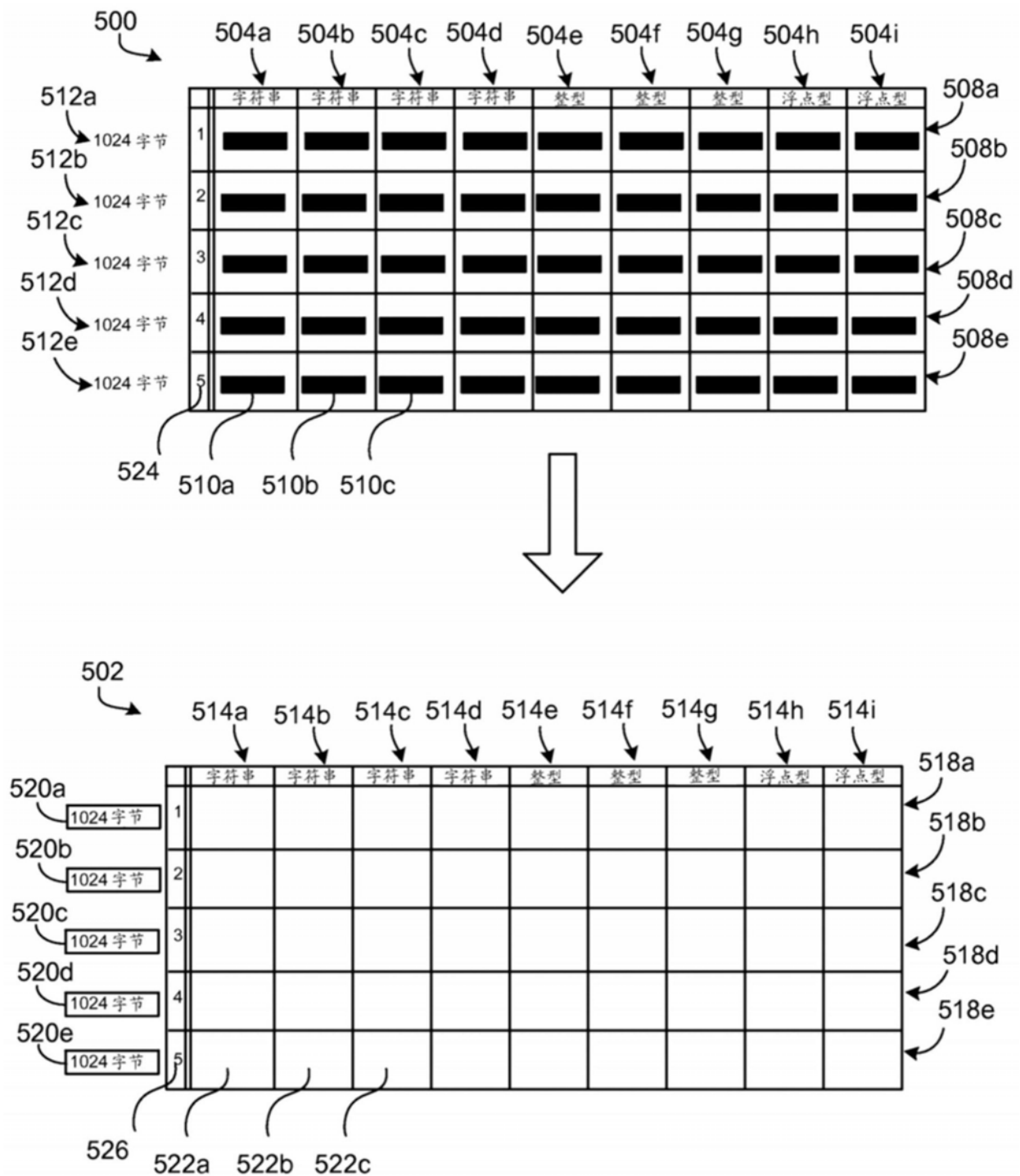


图5

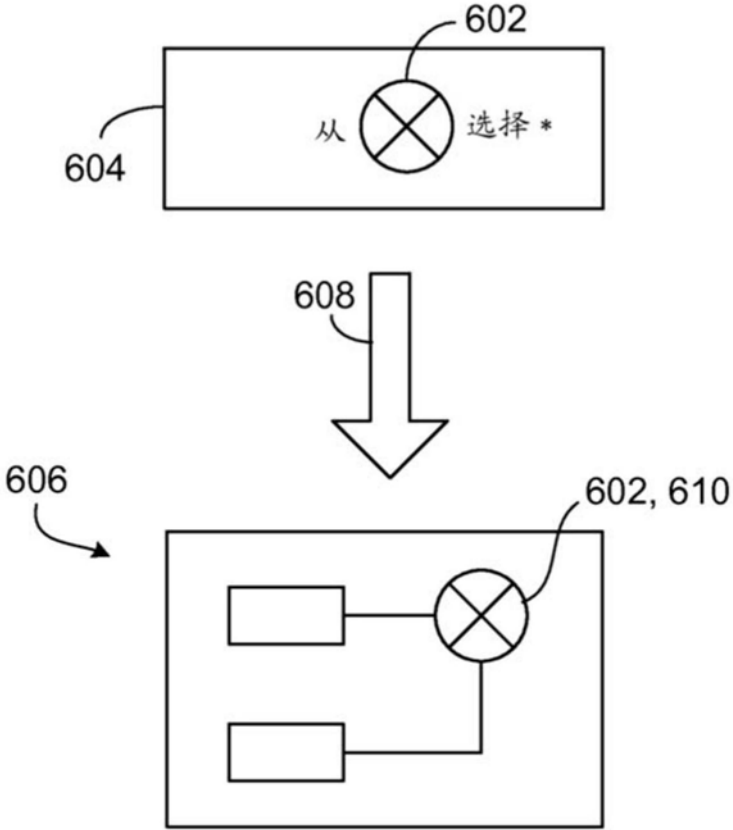


图6



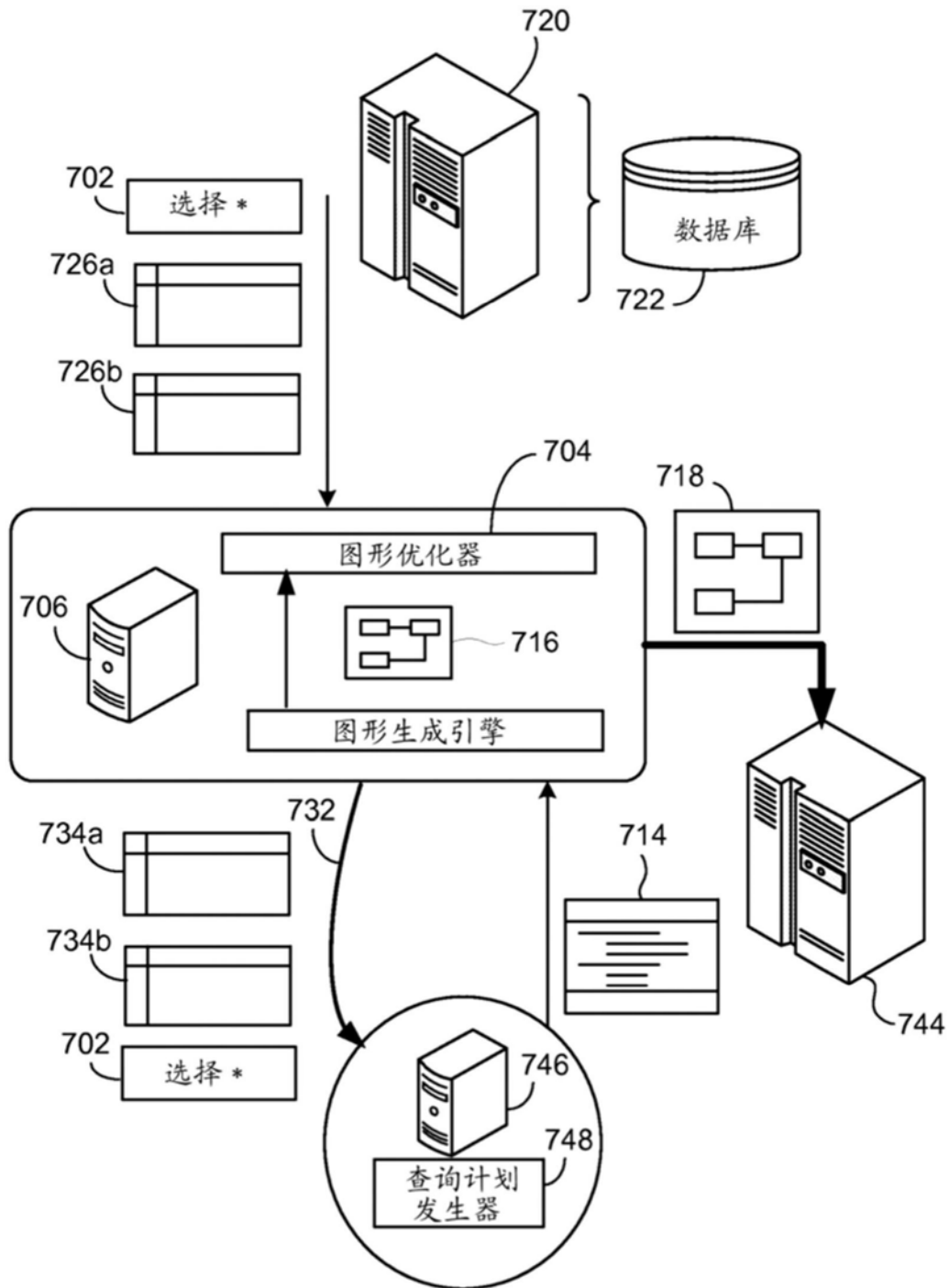


图7

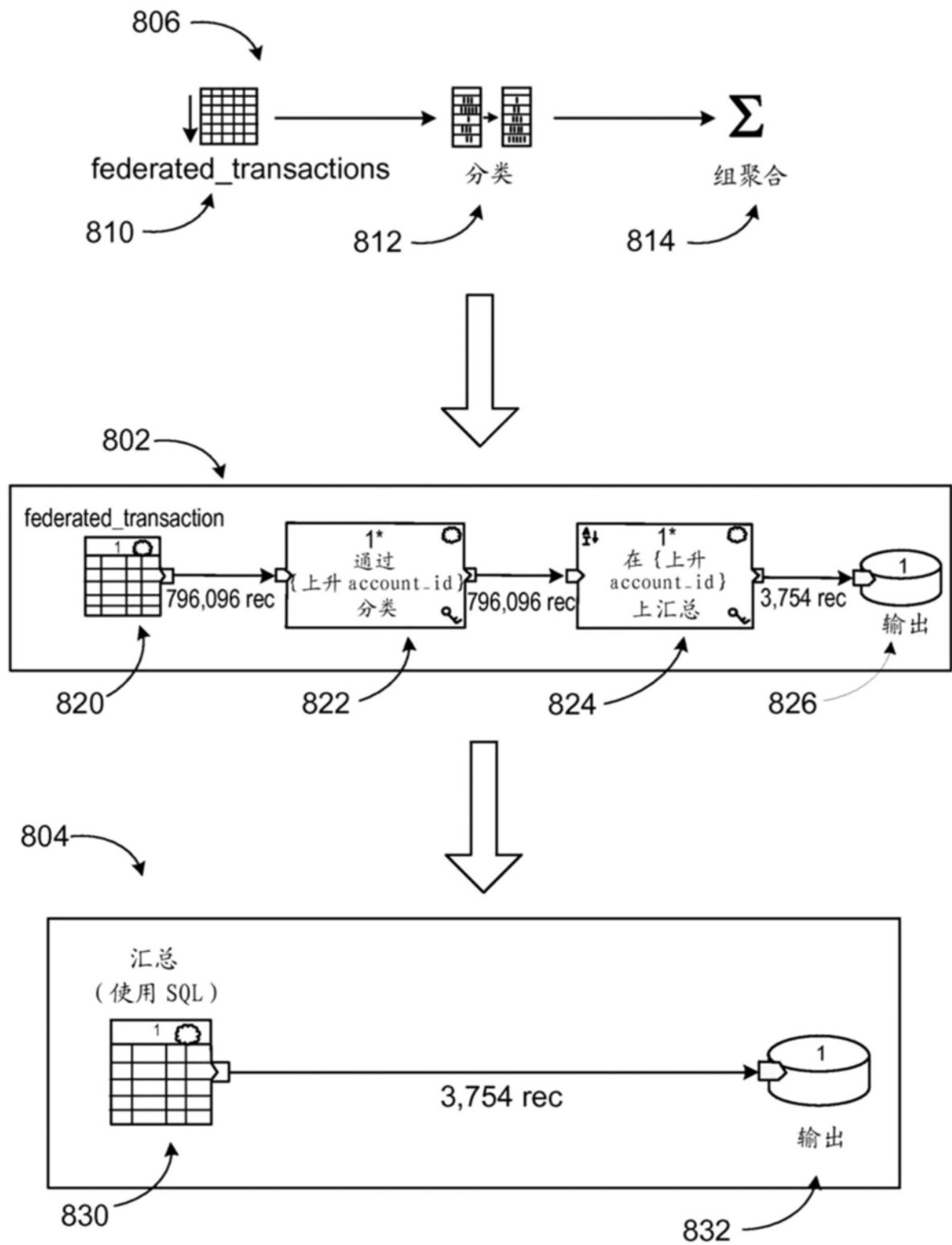


图8

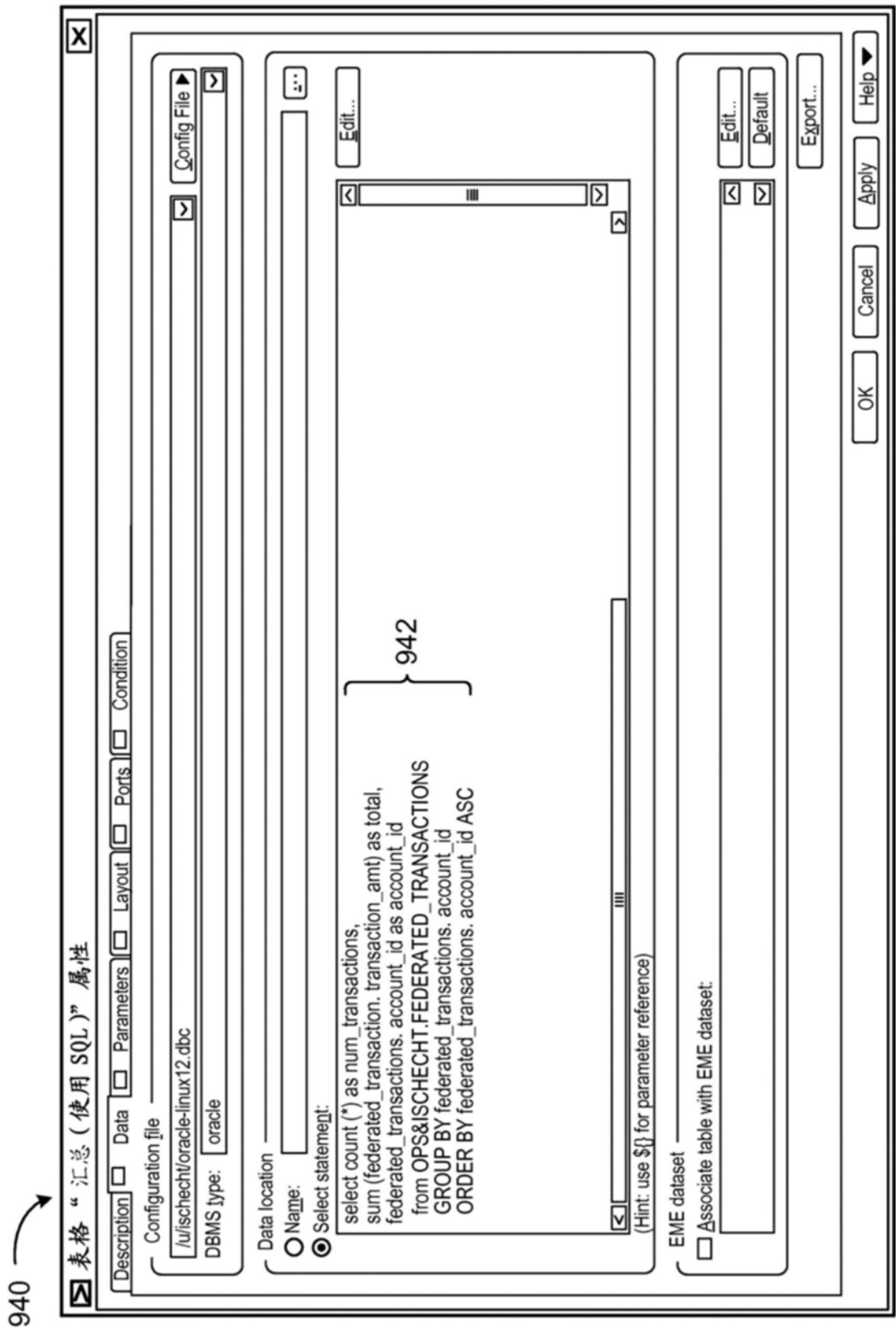


图9

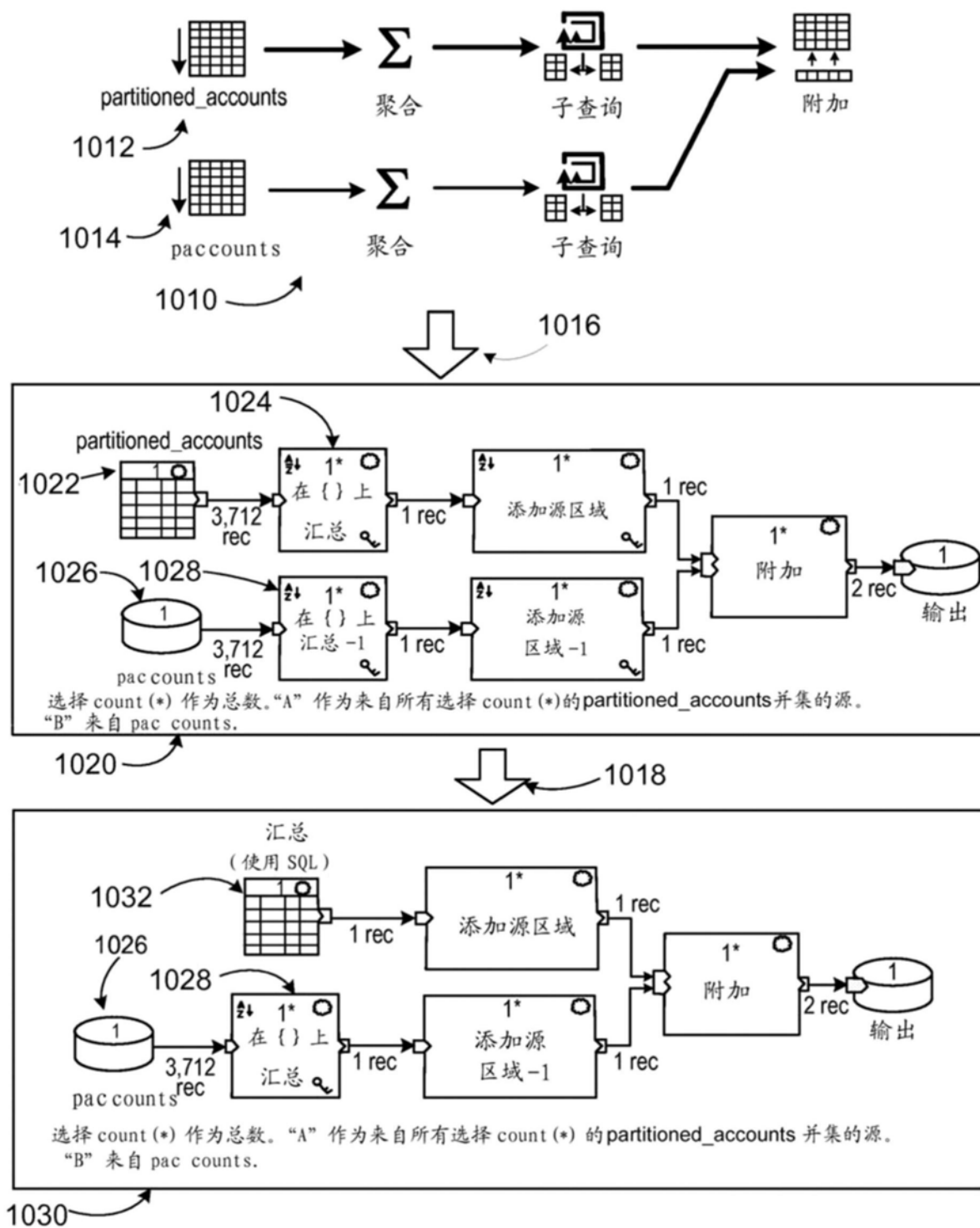


图10

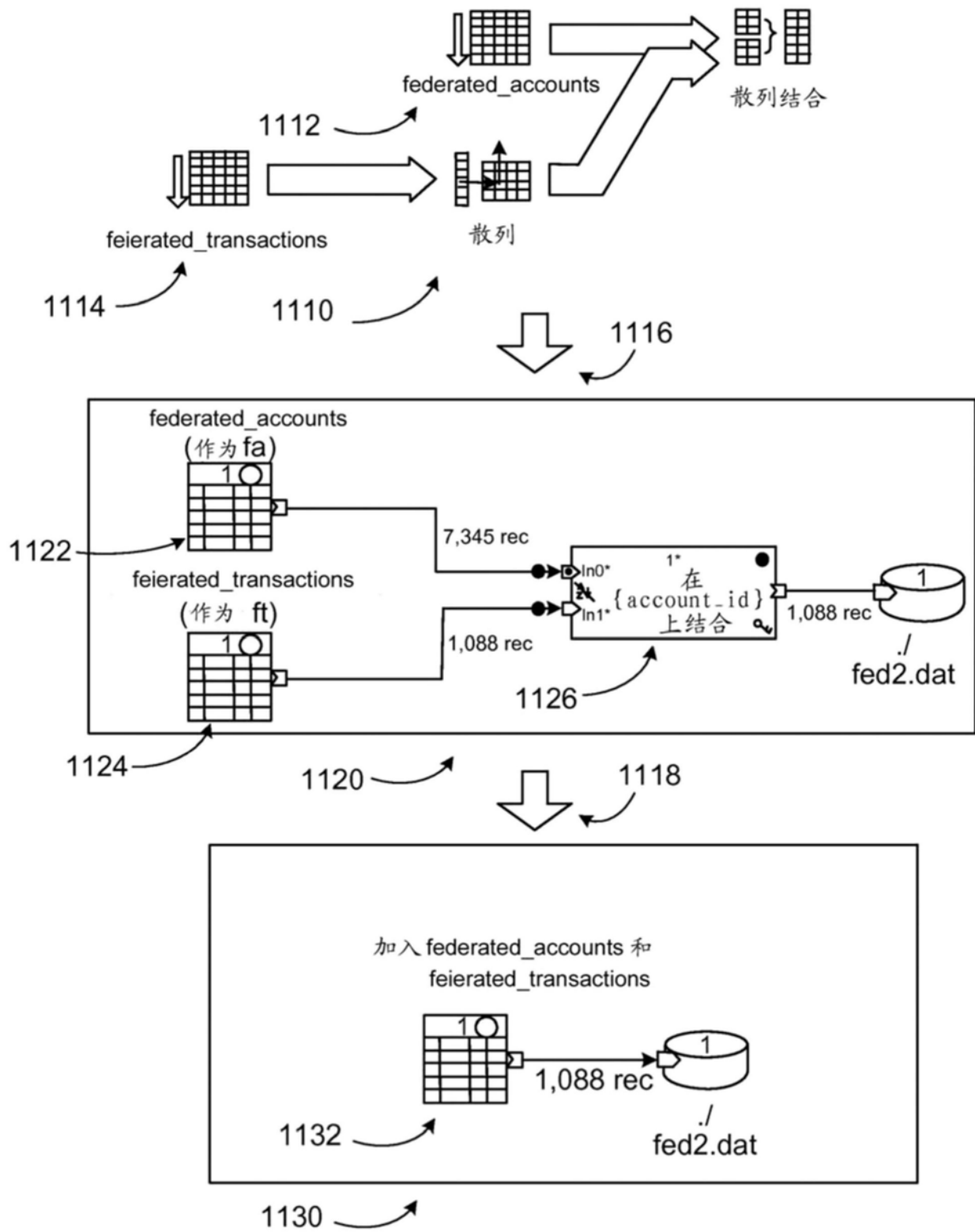


图11

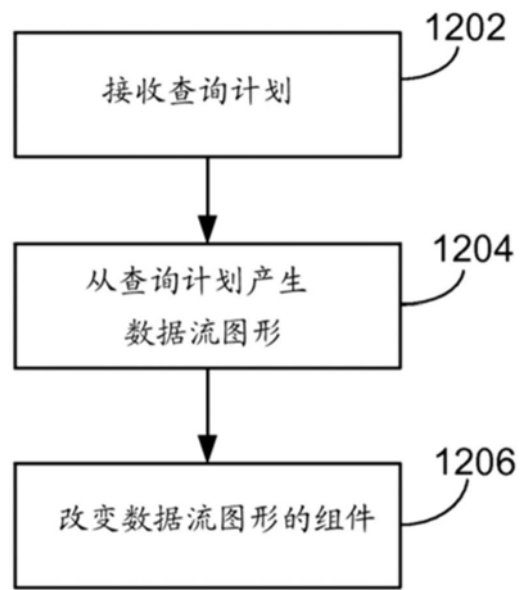


图12