

文章

[姚鑫](#) · 四月 27, 2021 阅读大约需分钟

第章 解释SQL查询计划

第章 解释SQL查询计划

本章介绍由ShowPlan生成InterSystems SQL查询访问计划中使用的语言和术语。

存储在映射中的表

SQL表存储为一组映射。

每个表都有一个包含表中所有数据的主映射；

表还可以有其他映射，如索引映射和位图。

每个映射可以被描绘成一个维全图，其中一些字段的数据在一个或多个标中，其余字段存储在节点值中。标控制要访问的数据。

- 对于主映射，RowID或IDKEY字段通常用作映射标。
- 对于索引映射，通常将其ID字段用作前导标，将RowID/IDKEY字段用作附加较低级别的标。
- 对于位图，可以将位图层视为附加的RowID标级别。但是，位图只能用于为正整数的RowID。

发展计划

编译SQL查询会生成组指令来访问和返回查询指定的数据。

这些指令表示为.int例程中的ObjectScript代码。

指令及其执行顺序受到SQL编译器中有关查询中涉及的表的结构和内容的数据的影响。

编译器尝试使用表大小和可用索引等信息，以使指令集尽可能高效。

查询访问计划(ShowPlan)是对结果指令集的可读翻译。

查询的作者可以使用这个查询访问计划来查看将如何访问数据。

虽然SQL编译器试图最有效地利用查询指定的数据，但有时查询的作者对存储的数据的某些方面的了解要比编译器清楚得。

在这种情况下，作者可以利用查询计划来开始查询，为查询编译器提供更多信息或更指导。

阅读计划

“ShowPlan”的结果是一系列关于访问和显示查询中指定的数据的处理的语句。

下面提供了关于如何解释ShowPlan语句的信息。

访问映射

一个查询计划可以访问多个表。

当访问一个表时，计划可以访问单个映射(索引或主映射)、两个映射(索引映射后面跟着主映射)，或者，对于索引计划，可以访问多个映射。

在通过映射访问数据时，计划指示使用的标。
它还指示实际的标值是体：一个给定值、一组给定值、一个值范围，或该标在表中显示的所有值。
选择哪一个取决于查询中指定条件。
显然，访问单个或几个标值要比访问该标级别上的所有值快得多。

条件和表达式

当查询运行时，将测试查询指定为条件。
除了前面提到的某些限制标的条件外，ShowPlan输出没有显式地指示条件的测试。
尽管测试条件总是最好的。
测试条件的最佳地点可以从计划细节中推断出来。

类似地，ShowPlan不详细描述表达式和子表达式的计算。
除了简单之外，主要原因是在大多数数据库环境中，表和索引访问结构处理的更重要方面；
检索表数据的速度占总查询速度的主要地位，因为磁盘访问速度仍然比CPU处理慢几个数量级。

循环

当访问一个表中的数据时，经常迭代地检查每个行。
这样的访问是通过一个循环来指示的。
每一次传递要执行的指令称为循环体
它们可以通过缩进直观地显示出来。
涉及多个表的数据库访问通常有循环中的循环。
在这种情况下，每个循环级别都通过与前一个级别相比的进一步缩进表示。

临时文件

定义

查询计划还可能指示构建和使用中间临时文件(TEMP-FILE)。这是本地数组中的“临时”区域。它用于临时结果以用于各种目的，如排序。就像映射一样，临时文件有一个或几个标，可能还有节点数据。

使用

一些临时文件包含处理单个表的数据。在这种情况下，可以将构建临时文件视为对该表中的数据进行预处理。在读取这样的临时文件之后，可以访问源表的主映射，也可以不访问源表的主映射。在其他情况下，临时文件可能包含处理多个表的结果。在其他情况下，临时文件用于存储分组的聚合值、检查DISTINCT等。

模块

临时文件的构建，以及其他处理，可以委托给一个称为模块的独立工作单元。
每个模块都被命名。
当列出单个模块时，该计划将指明调用每个模块的位置。
当模块执行结束时，处理将在模块调用之后的下一条语句中继续进行。

发送给处理的查询

对于通过ODBC或JDBC网关连接链连接的外部表，该计划显示发送到远程SQL gateway connection的查询文本，以从远程表检索所请求的数据。

对于并行查询处理和分片，该计划显示发送到并行处理或在分片上处理哪种查询。
还将显示用于每个查询的计划。

子查询、连接和联合

给定查询中的一些子查询(和视图)也可以单独处理。它们的计划在单独子查询部分中指定。在计划中没有指明子查询部分被调用的精确位置。这是因为它们经常作为条件或表达式处理的一部分被调用。

对于指定OUTER JOIN的查询,如果没有找到匹配的行,该计划可能指示可能生成null行,以满足外部连接语义的要求。

对于UNION,该计划可能指示将来自不同UNION子查询的结果行组合到一个单独模块中,在该模块中可以对这些结果行进行进一步处理。

计划分析

在分析给定查询的计划时,应用程序开发人员有时可能会觉得不同的计划会更有效率。应用程序开发人员有各种方法来影响计划。

首先,计划将在包含实际应用程序数据的环境中正确运行调优表的影响。在类源定义中手动定义一些Tune Table通常计算的值——例如表EXTENTSIZE、字段SELECTIVITY和映射BlockCount——也可以用于实现所需计划。

此外,分析计划可能表明对类定义的某些更改可能导致更有效的计划,例如:

添加一个索引

在某些情况(尽管不总是),使用一个临时文件进行预处理可能意味着向原始表添加一个与临时文件具有相同或类似结构的索引将消除构建临时文件的需要。从查询计划中删除这个处理步骤显然可以使查询运行得更快,但这必须与更新表时维护索引所需工作量进行平衡。

添加字段到索引数据

当计划显示正在使用的索引,然后是对主映射的访问时,这意味着将查询中使用的主映射字段添加到索引节点数据可能会为该查询生成更快的计划。同样,这必须与额外的更新时间以及添加到处理使用该索引的其他查询的额外时间进行平衡,因为索引会更大,因此需要更多的读取时间。

添加连接索引

当计划显示以特定顺序连接两个表时(例如,先检索t1,然后使用连接条件t1.a=t2.b连接到t2),可能相反的表顺序会产生一个更快的计划。例如,如果t2有额外的条件,可以显著限制符合条件的行数。在这种情况下,在t1上添加一个t1索引。a将允许这样一个连接顺序。

[#SQL #Caché #InterSystems IRIS #InterSystems IRIS for Health](#)

源 URL: <https://cn.community.intersystems.com/post/%E7%AC%AC%E4%B8%83%E7%AB%A0-%E8%A7%A3%E9%87%8A%E6%9F%A5%E8%AF%A2%E8%AE%A1%E5%88%92>