

## FoxPRO 2.5 程序设计讲座 (五)

罗 辉 (湖南省双峰工行)

何丹鸿 (湖南省娄底人行)

### 第七讲索引与 Rushmore 技术

为快速存储和检索数据库记录,或按某种顺序对记录加以处理,所有的数据库系统都提供了数据库索引功能。Foxpro 2.5 在索引技术方面独树一帜,它不仅大大丰富了索引类型,还提供了独有的 Rushmore 技术,大大优化了索引性能。

#### 一、索引类型和特点

##### 1. 索引类型

Foxpro2.5 支持多种索引类型。

(1) 按一个索引文件中索引的个数来分

① 单入口索引 IDX。为保持与老版本的 Fox 产品的兼容性, Foxpro2.5 仍支持 IDX 索引文件。每个索引放在一个独立的文件中,文件扩展名为 .IDX。若你希望按索引查询,那么在每次添加或编辑数据库文件时,必须打开所有索引文件。这样索引文件很多,如果有哪个索引文件忘了打开,其索引将不被修改,此后当你仍使用按老顺序排列的索引文件时,新添加的记录将无法检索到。

② 独立复合索引 CDX。与 IDX 索引不同, CDX 索引文件中可包含多个索引。每个索引可以按单入口索引的方式表达索引,指定由一个或多个字段组成的索引表达式进行索引。其中每个索引的称谓标志(TAG)的索引标识符,最长可达十个字符并可包括字母、数字或下划线,但必须以字母或下划线开头。索引文件名也可任意设定,可与数据库文件名不同,扩展名为 .CDX。

③ 结构复合索引 CDX。它是一种特殊的复合索引。无论你何时打开与它相应的数据库文件,它都自动

打开。结构复合索引文件名与数据库文件的文件名相同,扩展名为 .CDX。

(2) 按索引文件的存储方式来分

① 非压缩索引。索引信息未经压缩存盘。在单入口索引文件中一般使用这种存储方式。为保证 Foxpro 的应用程序的索引文件与 FoxBASE+ 的兼容性,那么应用程序中的索引文件必须使用这种非压缩的单入口索引方式。

② 压缩索引。复合索引文件都自动使用压缩索引方式。它建立的索引文件所占的磁盘空间远小于非压缩的 IDX 索引,约只为它的 1/6。这种索引既减少了空间占用,又提高了速度,改进了性能。单入口索引也可显式地使用压缩索引方式,但这种单入口索引将不能保持与 FoxBASE+ 的兼容性。

##### 2. 复合索引的特点

较之单入口索引,复合索引具有如下特点:

(1) 复合索引自动采用压缩存储方式,可节省大量磁盘空间;同时进行索引时只需做少量的磁盘存取,其执行速度相当快。

(2) 复合索引文件可同时容纳许多种索引,索引体积大大减小;体积小,则更适合放到内存中,以腾出内存空间打开更多的文件。

(3) 储存在复合索引文件中的索引数目即标志数目仅受内存空间和磁盘空间的限制,因而实际能打开的索引数目无限制。

(4) 结构复合索引文件随相关的数据库的打开而自动打开,不必花费许多时间去打开众多索引文件,也不必担心会忘记打开索引文件,索引能自动的及时更新。

(5) 结构复合索引有时会降低系统速度。如果应用

中存在大量索引,数据修改时,修改索引的时间将十分的长,大大增加更新索引所耗时间。因此,这时不使用结构复合索引可能会更好。

## 二、索引的操作

### 1. 建立索引

Foxpro2.5 有多种方式建立索引。交互式建立方法有:利用 Table Structure 对话框建立索引,它特别适应于建立结构复合索引;利用 INDEX 对话框建立索引,它适用建立单入口索引文件或独立的复合索引文件。批处理方式建立索引,是在程序或命令窗口中使用 INDEX ON 命令建立索引。

Foxpro2.5 对单入口索引和复合索引分别提供 INDEX ON 的命令格式。

对单入口索引,其命令格式是:

```
INDEX ON <索引表达式> TO <文件名>
[COMPACT]
[ASCENDING DESCENDING] [UNIQUE]
[FOR <表达式>][ADDITIVE]
```

使用 COMPACT 子句,可以建立压缩的单入口索引。

对复合索引,命令格式是:

```
INDEX ON <索引表达式> TAG <标志名>
OF <文件名>
[ASCENDING DESCENDING] [UNIQUE]
[FOR <表达式>][ADDITIVE]
```

使用 OF <文件名> 子句可建立独立的复合索引文件,它将不随数据库的打开而自动打开。复合索引自动以紧缩结构存储,而不需使用 COMPACT 子句。

你可以用 ASCENDING / DESCENDING 标明索引以升(降)序的顺序访问,也可用 UNIQUE 进行唯一性索引,或用 FOR 建立一个筛选条件进行有选择性的索引。

索引表达式可以是单个字段名,也可以用各种运算符、函数组合多个字段构成表达式。必须注意的是:

(1)索引表达式所用字段类型必须相同,不同的话,必须运用有关函数将它们转换成相同的类型。

(2)对两个字符型字段进行索引,可直接用“+”将两个字段串连起来。但如要对两个数值型字段建立不是两数之和的索引,就必须将每一个数值字段变成字符型进行索引。

(3)对日期型字段进行索引,必须注意日期的不同表达形式对索引结果的影响。例如一个日期,用美国表达方式为月/日/年,将不能得到正确的日期顺序索引,可用 DTOS()函数将日期数转换成[年月日]的字符串格式,即可对日期顺序进行准确的索引。

### 2. 使用索引

(1)指定主索引。操作复合索引文件,有时必须指定主索引。主索引决定记录的显示或访问顺序。用 SET ORDER TO[<单入口索引文件> <标志名>]说明该索引或标志符为主索引。若未说明索引或标识符,便取消主索引,因此记录按自然顺序排列。还可将 SET ORDER TO 与说明打开记录顺序的数字或表达式连用。

(2)重新索引。当你不是使用结构复合索引而是使用其它索引类型,在需要时,你必须用 REINDEX 命令重新索引,以便使对数据库的修改及时反应到索引中去。

(3)索引查找。利用 FIND 或 SEEK 可迅速定位到要查找的记录,其中查找表达式的类型必须与建立索引时的索引表达式类型匹配,否则会出错。你可利用 FOUND() 函数确证记录是否找到,也可使用 SEEK(表达式, [, 别名])函数来定位,它将 SEEK 命令和 FOUND()函数的功能组合到一起,若找到匹配则返回真,否则为假。

(4)查找后记录指针的定位。按索引查找后,如果找到记录,其指针即指向该记录。如果没找到匹配的记录,FOUND()函数返回假,EOF()返回真,同时记录指针将移到文件尾。但如用了 SET NEAR ON 命令同时没找到记录,则 FOUND()返回假,而 EOF() 则不一定为假,记录将移向关键字值紧跟搜索值之后或相类似的记录上,而不一定是在文件尾。(5)索引删除。单入口索引的删除,直接删除该索引文件即可。对复合索引的删除,必须用 DELETE TAG <标志列表> 选择索引标志进行部分删除。如果复合索引中的所有标志都被删除,相应

的索引文件即自行删除了。

(6)单入口索引与复合索引的转换。将一些单入口索引文件并到一个复合索引文件或从某复合索引文件产生单入口索引文件,有时也很有必要。这可通过:

```
COPY INDEXES <单入口索引文件列表>|ALL
[TO <CDX 文件>]
```

```
和 COPY TAG <标志> [OF <CDX 文件>] TO
<单入口索引文件>
```

两条命令分别完成。

### 三、Rushmore 技术

Rushmore 是一种存取一组记录较之单一被索引记录还要有效的数据存取技术。Foxpro 2.5 采用这一技术,实现了对系统记录的有效快速存取。Rushmore 技术的操作速度与欲处理的记录数成正比,记录数越多,使用 Rushmore 技术的优势越明显,从而在微机机上处理容纳数以百万计的记录的数据库的速度可与大型机相媲美。

Rushmore 技术能自动利用任何已打开的索引文件中包含的信息,包括传统的单入口索引,以加速数据查询速度。它采用标准 B- 树索引,而不需要任何其它文件或索引。Rushmore 技术也不特意依赖于压缩索引格式。

Rushmore 技术已经运用在某些命令的算法中,只要适当地运用这些命令,系统自动采用 Rushmore 技术来优化命令的执行速度。

能从 Rushmore 技术中受益的带 FOR 子句的命令有: AVERAGE、BROWSE、CALCULATE、CHANGE、COUNT、COPY TO、COPY TO ARRAY、DELETE、DISPLAY、EDIT、EXPORT、LABEL、LIST、LOCATE、RECALL、REPLACE、REPORT、SCAN、SORT、SUM、TOTAL 等命令。

使用 Rushmore 技术的命令其 FOR 子句中表达式必须是可优化或部分可优化的,同时如果命令还带有命令范围子句,则只能选 ALL 或 NEXT n,其它范围选定将使 Rushmore 无效。

最简单的应用 Rushmore 技术的形式是:在单一数据库中已被索引的索引字段域上利用 FOR 子句对一组

记录进行操作,Rushmore 技术将使操作速度成倍提高;如果 SET FILTER 是有效时,且过滤条件是建立在被索引的索引字段域上时,Rushmore 技术也将使某些命令速度大幅度加快。

Rushmore 不能使用一个过滤的或单一的索引(用 FOR 或 UNIQUE 等建立的索引),同时不需建立数据库主索引文件。你可以用 SET ORDER TO 0 确保按数据库物理顺序存取数据。

例:如下两种用于查找的程序结构,效果都很好。尽管第二种没有指定主关键字,而且使用 LOCATE 而不是 SEEK 查找记录,但由于 Rushmore 的作用,效果有时会更好。

```
SET ORDER TO <主关键字>
IF SEEK <查找值>
SCAN WHILE <主关键字 = 查找值>
    && 处理一条记录
ENDSCAN
```

```
ENDIF
```

<查找程序结构一>

```
LOCATE FOR <关键字 = 查找值>
DO WHILE FOUND( )
    &&处理一条记录
CONTINUE
```

```
ENDDO
```

<查找程序结构二>

通过设置 DELEDE 为 OFF,可以改善 Rushmore 的性能,特别是在查询结果返回大量的记录时。

优化时不必对数据库进行排序。如果用户建立索引文件,那么,这种排序是自动进行的。如果用户既想对一个大型数据库最大限度地发挥 Rushmore 技术的优势,又要求其中记录按特殊要求排序,那么应先用 SET ORDER TO 命令关掉索引控制,然后执行排序命令 SORT。

### 四、优化表达式

Rushmore 技术能否产生作用,取决于在 FOR 子句中的表达式是否可优化。通过一个基本的或复合的 FOR 子句优化表达式,可大大提高该命令的数据处理速

度。一个基本可优化表达式可形成一个完整的表达式，或只作为复合表达式的一部分出现。

**1.基本可优化表达式**

一个基本可优化表达式可以有两种形式：

<索引表达式> <关系运算符> <常量表达式>  
或 <常量表达式> <关系运算符> <索引表达式>

在一个基本可优化表达式中，<索引表达式>必须与建立索引文件时的索引关键词表达式精确匹配，而且索引关键词表达式中不能包含别名；<关系运算符>必须是“<、>、=、#、! =、<=、>=、<>”各运算符之一，但不包括符号\$、==；<常量表达式>可以是任何表达式，包括内存变量或与当前数据库无关的其它数据库字段变量。

表 1 复合表达式的优化规则

基本表达式	逻辑运算符	基本表达式	复合表达式
优化	AND	优化	完全优化
优化	OR	优化	完全优化
优化	AND	不优化	部全优化
优化	OR	不优化	不优化
不优化	AND	不优化	不优化
不优化	OR	不优化	不优化
-	NOT	优化	完全优化
-	NOT	不优化	不优化

例：若用户已用表达式 BOOKNAME、UPPER (BOOKNAME)、BOOKNO、PUBLICDATE 等建立了索引文件，则下面的表达式是基本可优化表达式，用于支持 Rushmore 技术的命令的 FOR 子句中，将大大提高数据处理的速度。

```
BOOKNAME='Foxpro2.5 程序设计'
UPPER(BOOKNAME)='FOXPRO2.5 程序设计'
BOOKNO=10001
PUBLICDATE>{07/30/94}
```

**2.复合的优化表达式**

通过 AND、OR 或 NOT 等逻辑运算符可将多个基本表达式组合成一个复合的 FOR 子句表达式。组成复合表达式的基本表达式如果有一部分是可优化的，Rushmore 将尽可能地优化整个复合表达式。

如果各基本表达式都是可优化的，则它们组成的复合表达式也是完全可优化的。若其中有一个或多个基本表达式不能优化，则复合表达式可能是部分可优化的也可能不可优化。其遵循一定的优化规则，见表 1。

**五、Rushmore 技术的局限**

尽管我们能从 Rushmore 技术中获得许多好处，但 Rushmore 技术并不能适用于所有的场合。

在下列情况下，Rushmore 技术不可使用或会失效：

1.一旦使用 Rushmore 的命令执行，Rushmore 就建立一个临时记录集，这一记录集满足 FOR 准则。你可对这组记录进行处理。但此时如果你却又要对这组记录的索引关键字进行修改，将导致不正确的结果。可以在命令的后面加上 NOOPTIMIZE 子句，或者输入 SET OPTIMIZE OFF 使 Rushmore 失效，直至执行了 SET OPTIMIZE ON 才使之恢复有效。

2.在少数情况下，Rushmore 技术可能会失效。譬如，可获得 Rushmore 技术支持的命令如果其命令范围选定不是 ALL 或 NEXT n，则 Rushmore 技术将被屏蔽；又如，尽管其命令的 FOR 子句带可优化的表达式，但如果命令中还包含有 WHILE 子句，则该命令得不到 Rushmore 技术的优化。

3.在 Foxpro2.5 的标准版本下，如果所有打开的数据库中记录总数超过 50 万个，Foxpro2.5 将暂停 Rushmore 技术的作用。这时，你可改用扩展版 Foxpro2.5，它在记录数超过 50 万个或不足 50 万个的情况，都可以发挥 Rushmore 优化作用，加快命令的执行速度。但扩展版要求计算机带扩展存储器 and 至少 386 以上的处理器。

4.在内存较少的情况下，Rushmore 技术不能优化数据命令。这时尽管不会出现数据丢失，而且程序将继续运行，但 Rushmore 技术在查询时已无法发挥作用。