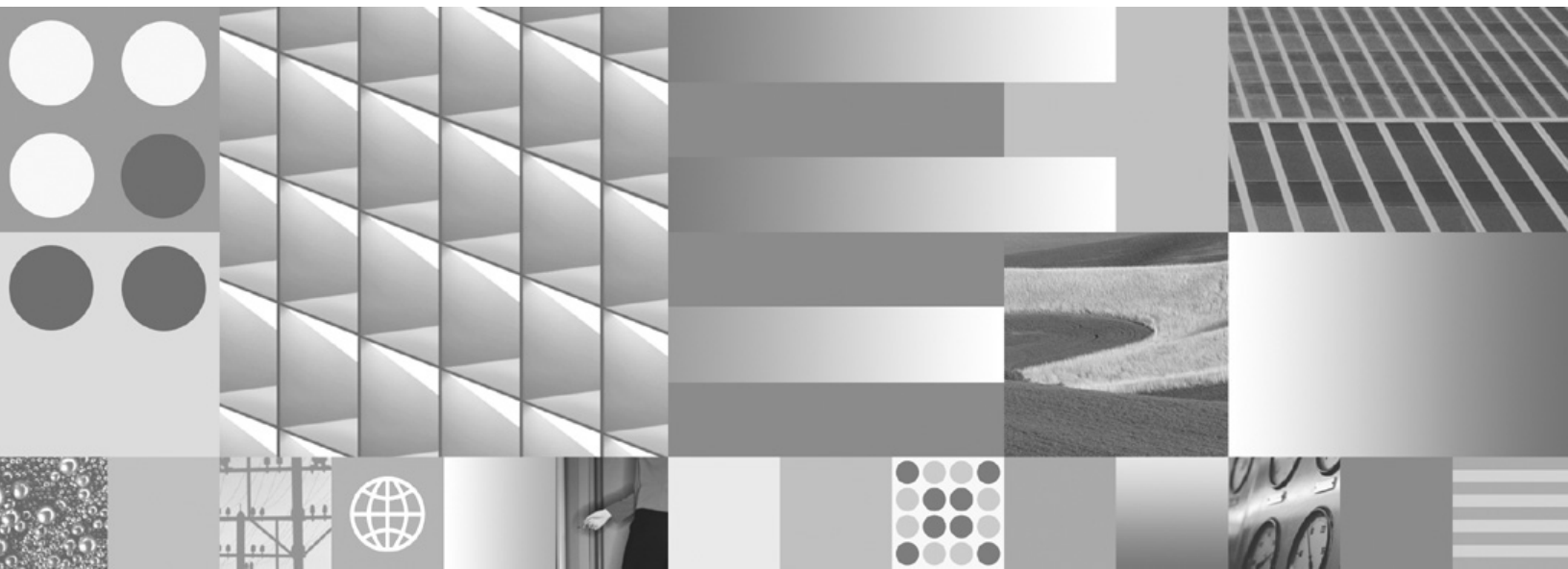


## 分区和集群指南





## 分区和集群指南

**注意**

使用此信息及其支持的产品前，请先阅读第 381 页的附录 E，『声明』下的常规信息。

**修订版声明**

此文档包含 IBM 的所有权信息。它在许可协议中提供，且受版权法的保护。本出版物中包含的信息不包括对任何产品的保证，且提供的任何语句都不需要如此解释。

您可在线或通过当地的 IBM 代表处订购 IBM 出版物。

- 要在线订购出版物，请转至 IBM 出版物中心，网址为：[www.ibm.com/shop/publications/order](http://www.ibm.com/shop/publications/order)
- 要查找当地的 IBM 代表处，请转至 IBM 全球联系人目录，网址为：[www.ibm.com/planetwide](http://www.ibm.com/planetwide)

要从美国或加拿大的 DB2 市场和销售部订购 DB2 出版物，请致电 1-800-IBM-4YOU (426-4968)。

当您向 IBM 发送信息时，即同意授予 IBM 独一无二的权力以它认为适当且不会对您造成任何影响的方式使用或分发该信息。

© Copyright International Business Machines Corporation 1993, 2007. All rights reserved.

# 目录

|                   |     |
|-------------------|-----|
| 关于本书 . . . . .    | vii |
| 本书的使用对象 . . . . . | vii |
| 本书的结构 . . . . .   | vii |
| 突出显示约定 . . . . .  | x   |

## 第 1 部分 规划和设计注意事项 . . . . . 1

### 第 1 章 分区数据库和表 . . . . . 3

|                                      |    |
|--------------------------------------|----|
| 设置分区数据库环境 . . . . .                  | 3  |
| 跨多个数据库分区的数据库分区 . . . . .             | 4  |
| 分区数据库认证注意事项 . . . . .                | 5  |
| 数据库分区组 . . . . .                     | 5  |
| 数据库分区组设计 . . . . .                   | 6  |
| 分发映射 . . . . .                       | 7  |
| 分布键 . . . . .                        | 8  |
| 表并置 . . . . .                        | 9  |
| 分区兼容性 . . . . .                      | 9  |
| 分区表 . . . . .                        | 10 |
| 表分区 . . . . .                        | 10 |
| 数据分区和范围 . . . . .                    | 12 |
| 数据组织方案 . . . . .                     | 12 |
| DB2 和 Informix 数据库中的数据组织方案 . . . . . | 17 |
| 表分区键 . . . . .                       | 21 |
| 分区表的装入注意事项 . . . . .                 | 23 |
| 复制具体化查询表 . . . . .                   | 25 |
| 数据库分区组中的表空间 . . . . .                | 26 |
| 表分区和多维集群表 . . . . .                  | 26 |

### 第 2 章 范围集群表 . . . . . 31

|                           |    |
|---------------------------|----|
| 范围集群表结构的优点 . . . . .      | 31 |
| 范围集群表的不兼容性 . . . . .      | 32 |
| 范围集群表和超出范围的记录键值 . . . . . | 32 |
| 范围集群表锁定 . . . . .         | 33 |

### 第 3 章 多维集群 (MDC) 表 . . . . . 35

|                               |    |
|-------------------------------|----|
| 多维集群表 . . . . .               | 35 |
| 常规表与 MDC 表的比较 . . . . .       | 35 |
| 选择 MDC 表维 . . . . .           | 37 |
| 创建 MDC 表时的注意事项 . . . . .      | 43 |
| MDC 表的装入注意事项 . . . . .        | 48 |
| MDC 表的记录注意事项 . . . . .        | 48 |
| MDC 表的块索引注意事项 . . . . .       | 48 |
| MDC 表的块索引 . . . . .           | 49 |
| 方案: 多维集群 (MDC) 表 . . . . .    | 51 |
| MDC 表的块索引和查询性能 . . . . .      | 54 |
| 在 INSERT 操作期间自动维护集群 . . . . . | 57 |
| MDC 表的块映射 . . . . .           | 59 |
| 从 MDC 表中删除 . . . . .          | 60 |
| 更新 MDC 表 . . . . .            | 61 |
| 表分区和多维集群表 . . . . .           | 61 |

## 第 4 章 并行数据库系统 . . . . . 67

|                       |    |
|-----------------------|----|
| 并行性 . . . . .         | 67 |
| 分区数据库环境 . . . . .     | 70 |
| 数据库分区和处理器环境 . . . . . | 72 |

## 第 2 部分 安装注意事项 . . . . . 79

### 第 5 章 安装先决条件 . . . . . 81

|  |    |
|--|----|
| 安装 DB2 服务器 (Windows) . . . . .                   | 81 |
| 为分区 DB2 服务器准备环境 (Windows) . . . . .              | 82 |
| 快速通信管理器 (Windows) . . . . .                      | 84 |
| 安装 DB2 服务器产品的概述 (Linux 和 UNIX) . . . . .         | 84 |
| DB2 安装方法 . . . . .                               | 85 |
| 使用“DB2 安装”向导来安装 DB2 服务器 (Linux 和 UNIX) . . . . . | 87 |

### 第 6 章 安装之前 . . . . . 91

|   |    |
|---|----|
| 其他分区数据库环境预安装任务 (Linux 和 UNIX) . . . . .       | 91 |
| 更新用于分区 DB2 安装的环境设置 (AIX) . . . . .            | 91 |
| 建立工作集合以将命令分发至 ESE 工作站 (AIX) . . . . .         | 92 |
| 验证 NFS 是否在运行 (Linux 和 UNIX) . . . . .         | 93 |
| 验证参与的计算机上的可用端口范围 (Linux 和 UNIX) . . . . .     | 94 |
| 为分区 DB2 服务器创建文件系统 (Linux) . . . . .           | 95 |
| 创建分区数据库系统 (AIX) 的 DB2 本地文件系统 . . . . .        | 96 |
| 在分区数据库环境中为 DB2 服务器安装创建必需的用户 (Linux) . . . . . | 98 |
| 在分区数据库环境中为 DB2 服务器安装创建必需的用户 (AIX) . . . . .   | 99 |

### 第 7 章 安装 DB2 服务器产品 . . . . . 101

|   |     |
|---|-----|
| 设置分区数据库环境 . . . . .                               | 101 |
| 使用响应文件在参与的计算机上安装数据库分区服务器 (Windows) . . . . .      | 103 |
| 使用响应文件在参与的计算机上安装数据库分区服务器 (Linux 和 UNIX) . . . . . | 104 |

### 第 8 章 安装之后 . . . . . 105

|                                       |     |
|---------------------------------------|-----|
| 验证安装 . . . . .                        | 105 |
| 验证分区数据库环境安装 (Windows) . . . . .       | 105 |
| 验证分区数据库服务器安装 (Linux 和 UNIX) . . . . . | 105 |

## 第 3 部分 实施和维护 . . . . . 107

### 第 9 章 创建数据库之前 . . . . . 109

|                         |     |
|-------------------------|-----|
| 设置分区数据库环境 . . . . .     | 109 |
| 创建节点配置文件 . . . . .      | 110 |
| DB2 节点配置文件的格式 . . . . . | 111 |

|  |     |
|--|-----|
| 指定分区数据库环境中的机器列表 . . . . .                | 117 |
| 除去分区数据库环境中机器列表中的重复条目 . . . . .           | 117 |
| 更新节点配置文件 (Linux 和 UNIX) . . . . .        | 117 |
| 设置多逻辑节点 . . . . .                        | 119 |
| 配置多逻辑分区 . . . . .                        | 119 |
| 启用分区内查询并行性 . . . . .                     | 120 |
| 对查询启用分区内并行性 . . . . .                    | 121 |
| 数据服务器容量的管理 . . . . .                     | 122 |
| 快速通信管理器 . . . . .                        | 123 |
| 快速通信管理器 (Windows) . . . . .              | 123 |
| 快速通信管理器 (Linux 和 UNIX) . . . . .         | 123 |
| 使用 FCM 通信启用数据库分区之间的通信 . . . . .          | 123 |
| 启用数据库分区服务器之间的通信 (Linux 和 UNIX) . . . . . | 124 |

## 第 10 章 创建和管理分区数据库环境 127

|   |     |
|---|-----|
| 初始数据库分区组 . . . . .                          | 127 |
| 创建数据库分区组 . . . . .                          | 127 |
| 数据库分区组中的表空间 . . . . .                       | 128 |
| 管理数据库分区 . . . . .                           | 128 |
| 在分区数据库环境中添加数据库分区 . . . . .                  | 129 |
| 将数据库分区添加到正在运行的数据库系统 . . . . .               | 130 |
| 将数据库分区添加至已停止的数据库系统 (Windows) . . . . .      | 131 |
| 将数据库分区添加到已停止的数据库系统 (UNIX) . . . . .         | 131 |
| 添加数据库分区时的错误恢复 . . . . .                     | 133 |
| 删除数据库分区 . . . . .                           | 134 |
| 列实例中的数据库分区服务器 . . . . .                     | 134 |
| 将数据库分区服务器添加至实例 (Windows) . . . . .          | 135 |
| 使用“添加分区”向导将数据库分区添加至实例 . . . . .             | 136 |
| 更改数据库分区 (Windows) . . . . .                 | 136 |
| 向数据库分区上的 SMS 表空间添加容器 . . . . .              | 137 |
| 从实例中删除数据库分区 (Windows) . . . . .             | 138 |
| 使用“删除分区”启动板从实例中删除数据库分区 . . . . .            | 138 |
| 方案: 对数据库中的数据进行分区 . . . . .                  | 139 |
| 在分区数据库环境中发出命令 . . . . .                     | 141 |
| rah 和 db2_all 命令概述 . . . . .                | 141 |
| 指定 rah 和 db2_all 命令 . . . . .               | 142 |
| 以并行方式运行命令 (Linux 和 UNIX) . . . . .          | 143 |
| 扩展 rah 命令以使用树逻辑 (AIX 和 Solaris) . . . . .   | 143 |
| rah 和 db2_all 命令 . . . . .                  | 143 |
| rah 命令前缀序列 . . . . .                        | 144 |
| 控制 rah 命令 . . . . .                         | 146 |
| 指定与 rah 一起运行的 . 文件 (Linux 和 UNIX) . . . . . | 147 |
| 确定 rah 的问题 (Linux 和 UNIX) . . . . .         | 148 |
| 监视 rah 进程 (Linux 和 UNIX) . . . . .          | 149 |
| 在 Windows 上为 rah 设置缺省环境概要文件 . . . . .       | 150 |

## 第 11 章 创建表和其他相关表对象 . . . 151

|                         |     |
|-------------------------|-----|
| 分区数据库环境中的表 . . . . .    | 151 |
| 分区表中的大对象行为 . . . . .    | 152 |
| 创建分区表 . . . . .         | 153 |
| 定义分区表的范围 . . . . .      | 153 |
| 将现有表和视图迁移到分区表 . . . . . | 156 |

|                               |     |
|-------------------------------|-----|
| 已分区的具体化查询表 (MQT) 行为 . . . . . | 158 |
| 创建范围集群表 . . . . .             | 161 |
| 用于范围集群表的算法 . . . . .          | 161 |
| 范围集群表索引 . . . . .             | 162 |
| 与常规表的差别 . . . . .             | 162 |
| 关于使用范围集群表的准则 . . . . .        | 163 |
| SQL 编译器如何使用范围集群表 . . . . .    | 163 |
| 方案: 范围集群表 . . . . .           | 163 |
| 创建 MDC 表时的注意事项 . . . . .      | 165 |

## 第 12 章 改变数据库 . . . . . 171

|                            |     |
|----------------------------|-----|
| 改变实例 . . . . .             | 171 |
| 更改多个数据库分区中的数据库配置 . . . . . | 171 |
| 改变数据库 . . . . .            | 171 |
| 改变数据库分区组 . . . . .         | 171 |
| 从“控制中心”管理数据库分区 . . . . .   | 171 |

## 第 13 章 改变表和其他相关表对象 . . . 173

|                          |     |
|--------------------------|-----|
| 改变分区表 . . . . .          | 173 |
| 有关改变分区表的准则和限制 . . . . .  | 174 |
| 连接数据分区 . . . . .         | 174 |
| 将数据分区连接至分区表的准则 . . . . . | 176 |
| 拆离数据分区 . . . . .         | 178 |
| 已拆离数据分区的属性 . . . . .     | 180 |
| 对分区表添加数据分区 . . . . .     | 181 |
| 删除数据分区 . . . . .         | 182 |
| 方案: 旋转分区表中的数据 . . . . .  | 183 |
| 方案: 转入和转出分区表数据 . . . . . | 184 |

## 第 14 章 装入 . . . . . 189

|                      |     |
|----------------------|-----|
| 并行性和装入 . . . . .     | 189 |
| 的多维集群注意事项 . . . . .  | 190 |
| 分区表的装入注意事项 . . . . . | 190 |

## 第 15 章 在分区数据库环境中装入数据 193

|   |     |
|---|-----|
| 装入概述 - 分区数据库环境 . . . . .                  | 193 |
| 在分区数据库环境中装入数据 - 提示与技巧 . . . . .           | 194 |
| 在分区数据库环境中装入数据 . . . . .                   | 195 |
| 使用 LOAD QUERY 命令在分区数据库环境中监视装入操作 . . . . . | 200 |
| 在分区数据库环境中继续、重新启动或终止装入操作 . . . . .         | 201 |
| 为分区数据库环境装入配置选项 . . . . .                  | 203 |
| 分区数据库环境中的装入会话 - CLP 示例 . . . . .          | 207 |
| 迁移和版本兼容性 . . . . .                        | 209 |

## 第 16 章 迁移分区数据库环境 . . . . . 211

|                   |     |
|-------------------|-----|
| 迁移分区数据库 . . . . . | 211 |
|-------------------|-----|

## 第 17 章 使用快照和事件监视器 . . . . . 213

|                              |     |
|------------------------------|-----|
| 使用快照监视器数据来监视分区表的重组 . . . . . | 213 |
| 分区数据库系统上的全局快照 . . . . .      | 220 |
| 为分区数据库创建事件监视器 . . . . .      | 221 |

## 第 18 章 开发好的备份和恢复策略 . . . 225

|                             |     |
|-----------------------------|-----|
| 崩溃恢复 . . . . .              | 225 |
| 从分区数据库环境中的事务故障中恢复 . . . . . | 226 |

|                   |     |
|-------------------|-----|
| 从数据库分区服务器的故障恢复    | 228 |
| 重建分区数据库           | 228 |
| 使用 db2adutl 来恢复数据 | 229 |
| 使分区数据库环境中的时钟同步    | 235 |

## 第 19 章 故障诊断 . . . . . 237

|               |     |
|---------------|-----|
| DB2 故障诊断      | 237 |
| 诊断分区数据库环境     | 237 |
| 在分区数据库环境中发出命令 | 237 |

## 第 4 部分 性能问题 . . . . . 239

### 第 20 章 数据库设计中的性能问题 . . . 241

|            |     |
|------------|-----|
| 性能增强功能     | 241 |
| 表分区和多维集群表  | 241 |
| 分区表的优化策略   | 244 |
| MDC 表的优化策略 | 249 |

### 第 21 章 索引 . . . . . 253

|               |     |
|---------------|-----|
| 分区表中的索引       | 253 |
| 了解分区表上的索引行为   | 253 |
| 了解分区表上的集群索引行为 | 256 |

### 第 22 章 设计顾问程序 . . . . . 259

|                        |     |
|------------------------|-----|
| 使用设计顾问程序从单一分区迁移到多分区数据库 | 259 |
|------------------------|-----|

### 第 23 章 管理并行性 . . . . . 261

|                         |     |
|-------------------------|-----|
| 表的锁定方式和 MDC 表的 RID 索引扫描 | 261 |
| MDC 表的块索引扫描的锁定          | 264 |
| 对分区表的锁定行为               | 266 |

### 第 24 章 代理程序管理 . . . . . 269

|             |     |
|-------------|-----|
| 分区数据库中的代理程序 | 269 |
|-------------|-----|

### 第 25 章 优化访问方案 . . . . . 271

|                                     |     |
|-------------------------------------|-----|
| 索引访问和集群比率                           | 271 |
| MDC 表的表和索引管理                        | 271 |
| 集群                                  | 273 |
| 分区内并行性的优化策略                         | 273 |
| db2expln 和 dynexpln 输出的示例           | 275 |
| 示例一: 无并行性                           | 275 |
| 示例二: 具有分区内并行性的单一分区方案                | 276 |
| 示例三: 具有分区间并行性的多分区方案                 | 278 |
| 示例四: 具有分区间并行性和分区内并行性的多分区方案          | 280 |
| 连接                                  | 282 |
| 数据库分区组对查询优化的影响                      | 283 |
| 分区数据库中的连接策略                         | 283 |
| 分区数据库环境中的连接方法                       | 285 |
| 分区数据库环境中重复的具体化查询表                   | 290 |
| 第 4 课 在分区数据库环境中改进访问方案               | 292 |
| 使用访问方案图                             | 292 |
| 在分区数据库环境中运行没有索引和统计信息的查询             | 292 |
| 在分区数据库环境中使用 runstats 来收集表和索引的当前统计信息 | 295 |

|                         |     |
|-------------------------|-----|
| 对分区数据库环境的查询中用于连接表的列创建索引 | 299 |
| 对分区数据库环境中的表列创建其他索引      | 303 |

## 第 26 章 数据再分发 . . . . . 307

|                               |     |
|-------------------------------|-----|
| 对 Redistribute 实用程序的限制        | 308 |
| 确定是否要再分发数据                    | 309 |
| 在数据库分区之间再分发数据                 | 309 |
| 在数据库分区组中再分发数据                 | 310 |
| 数据再分发的日志空间要求                  | 311 |
| 再分发事件日志记录和错误恢复                | 311 |
| STEPWISE_REDISTRIBUTE_DBPG 过程 | 314 |

## 第 27 章 配置自调整内存 . . . . . 317

|                  |     |
|------------------|-----|
| 分区数据库环境中的自调整内存   | 317 |
| 在分区数据库环境中使用自调整内存 | 318 |

## 第 28 章 DB2 配置参数和变量 . . . . . 321

|             |     |
|-------------|-----|
| 配置跨多个分区的数据库 | 321 |
| 分区数据库环境变量   | 322 |
| 分区数据库环境配置参数 | 323 |
| 通信          | 323 |
| 并行处理        | 327 |

## 第 5 部分 管理 API、命令和 SQL 语句 . . . . . 329

### 第 29 章 管理 API . . . . . 331

|                               |     |
|-------------------------------|-----|
| sqlleadn - 添加数据库分区服务器到分区数据库环境 | 331 |
| sqlcgran - 在数据库分区服务器上创建数据库    | 332 |
| sqlqpan - 删除数据库分区服务器上的数据库     | 334 |
| sqlqdrpn - 检查是否可以删除数据库分区服务器   | 335 |
| sqlqgrpn - 为行获取数据库分区服务器号      | 336 |

### 第 30 章 命令 . . . . . 339

|                                       |     |
|---------------------------------------|-----|
| REDISTRIBUTE DATABASE PARTITION GROUP | 339 |
| db2nchg - 更改数据库分区服务器配置                | 345 |
| db2nprt - 将数据库分区服务器添加至实例              | 346 |
| db2ndrop - 从实例中删除数据库分区服务器             | 348 |

### 第 31 章 SQL 语言元素 . . . . . 349

|                        |     |
|------------------------|-----|
| 数据类型                   | 349 |
| 与数据库分区兼容的数据类型          | 349 |
| 专用寄存器                  | 350 |
| CURRENT DBPARTITIONNUM | 350 |

### 第 32 章 SQL 函数 . . . . . 351

|                  |     |
|------------------|-----|
| DATAPARTITIONNUM | 351 |
| DBPARTITIONNUM   | 351 |

### 第 33 章 SQL 语句 . . . . . 353

|                                 |     |
|---------------------------------|-----|
| ALTER DATABASE PARTITION GROUP  | 353 |
| CREATE DATABASE PARTITION GROUP | 356 |

|   |            |
|---|------------|
| <b>第 34 章 受支持的 SQL 管理例程和视图</b>                        | <b>359</b> |
| ADMIN_CMD 存储过程和关联的管理 SQL 例程                           | 359        |
| 使用 ADMIN_CMD 过程的 GET STMM TUNING DBPARTITIONNUM 命令    | 359        |
| 使用 ADMIN_CMD 过程的 UPDATE STMM TUNING DBPARTITIONNUM 命令 | 360        |
| 配置管理 SQL 例程和视图  | 360        |
| DB_PARTITIONS   | 360        |
| 按步骤再分发管理 SQL 例程                                       | 361        |
| STEPWISE_REDISTRIBUTE_DBPG 过程 - 再分发部分数据库分区组           | 361        |
| <b>第 6 部分 附录</b>                                      | <b>363</b> |
| <b>附录 A. 作为非 root 用户安装</b>                            | <b>365</b> |
| 作为非 root 用户来安装 DB2 产品                                 | 365        |
| <b>附录 B. 使用备份</b>                                     | <b>367</b> |
| 使用备份  | 367        |
| <b>附录 C. 分区数据库环境目录视图</b>                              | <b>369</b> |

|                                |     |
|--------------------------------|-----|
| SYSCAT.BUFFERPOOLDBPARTITIONS  | 369 |
| SYSCAT.DATAPARTITIONEXPRESSION | 369 |
| SYSCAT.DATAPARTITIONS          | 369 |
| SYSCAT.DBPARTITIONGROUPDEF     | 370 |
| SYSCAT.DBPARTITIONGROUPS       | 371 |
| SYSCAT.PARTITIONMAPS           | 371 |

|                               |            |
|-------------------------------|------------|
| <b>附录 D. DB2 技术信息概述</b>       | <b>373</b> |
| 硬拷贝或 PDF 格式的 DB2 技术库          | 373        |
| 订购印刷版的 DB2 书籍                 | 375        |
| 从命令行处理器显示 SQL 状态帮助            | 376        |
| 访问不同版本的 DB2 信息中心              | 376        |
| 在 DB2 信息中心中以您的首选语言显示主题:       | 377        |
| 更新安装在您的计算机或内部网服务器上的 DB2 信息中心: | 377        |
| DB2 教程                        | 379        |
| DB2 故障诊断信息                    | 379        |
| 条款和条件                         | 379        |
| <b>附录 E. 声明</b>               | <b>381</b> |
| <b>索引</b>                     | <b>385</b> |



---

## 关于本书

欢迎使用《分区和集群指南》！

DB2® 关系数据库管理系统的功能在很大程度上受分区和集群功能影响，这两种功能允许管理员和系统操作员有效地增强数据库性能并将许多数据库对象分布在硬件资源中。更快的数据检索速度、将对象分布在不断增多的硬件资源中的能力以及利用并行性和存储器容量，这些最终会极大地提高效率。本书包含 DB2 库中的一个有组织的主题集，它形成一个丰富的信息源，只重点讨论规划、设计、实施、使用以及数据库分区、表分区、表集群、表范围集群、多维集群表和并行性的维护。

---

## 本书的使用对象

本书主要面向需要设计、实施和/或维护由本地和远程客户机访问的分区和/或集群数据库的数据库管理员、系统管理员、安全管理员和系统操作员。应用程序开发者和其他用户也可通过本书获得丰富信息源，并从中了解 DB2 关系数据库管理系统的分区、集群和并行性功能的管理和操作方式。对于那些关注我们此处讨论的任何或所有主要功能的将来实施的人来说，本书是难得的参考资料。

---

## 本书的结构

DB2 库中的此主题集提供一个丰富的信息源，它只重点讨论 DB2 分区、集群和并行性功能。为了使您能够方便有效地使用本书，它分为六个主要部分，前五个部分表示管理员、系统操作员和应用程序开发者所关心的主要管理主题。本书主要部分中的主题可映射至表示 DB2 库中另一本书的内容的主题，从而使您能够方便地交叉引用更多一般信息，因为它涉及大量其他 DB2 功能和对象。例如，在阅读第 4 部分第 20 章中关于多维集群表的优化策略如何改善性能的主题后，您可能希望通过阅读该特定示例主题映射至的《调整数据库性能》一书，来检查常规表上可配置的其他一般性能增强功能。在下面的表 1 中，您将发现本书中映射至其他书籍的主要部分，通过参阅这些书籍，您可以了解关于类似主题包含的其他 DB2 对象和功能的更多信息。

表 1. 本书各部分至 DB2 库中的其他书籍的映射

| 《分区和集群指南》中的各部分   | 映射至 DB2 库中的书籍                        |
|------------------|--------------------------------------|
| 第 1 部分 规划和设计注意事项 | 《数据服务器、数据库和数据库对象指南》<br>《数据库安全性指南》    |
| 第 2 部分 安装注意事项    | 《数据服务器、数据库和数据库对象指南》<br>《DB2 服务器快速入门》 |

表 1. 本书各部分至 DB2 库中的其他书籍的映射 (续)

| 《分区和集群指南》中的各部分           | 映射至 DB2 库中的书籍   |
|--------------------------|---|
| 第 3 部分 实施和维护             | 《数据移动实用程序指南和参考》<br>《数据恢复和高可用性指南和参考》<br>《数据服务器、数据库和数据库对象指南》<br>《迁移指南》<br>《系统监视器指南和参考》<br>《故障诊断指南》<br>《Visual Explain 教程》<br>《XQuery 参考》  |
| 第 4 部分 性能问题              | 《数据服务器、数据库和数据库对象指南》<br>《调整数据库性能》<br>《Visual Explain 教程》   |
| 第 5 部分 管理 API、命令和 SQL 语句 | <i>Administrative API Reference</i><br><i>Administrative Routines and Views</i><br><i>Command Reference</i><br><i>Developing ADO.NET and OLE DB Applications</i><br><i>Developing Embedded SQL Applications</i><br><i>Developing Java Applications</i><br><i>Developing Perl and PHP Applications</i><br><i>Developing User-defined Routines (SQL and External)</i><br><i>Getting Started with Database Application Development</i><br><i>SQL Reference</i> |
| 第 6 部分 附录                | 《数据恢复和高可用性指南和参考》<br>《DB2 服务器快速入门》<br><i>SQL Reference</i>   |

本书的各章节中讨论了下列主要主题领域:

### 第 1 部分 规划和设计注意事项

下列所有章节都包含与规划和设计数据库/表相关的概念性信息，将对这些数据库/表进行分区或集群，或者将它们用于并行数据库系统中。

- 第 1 章『分区数据库和表』引入了关于对数据库和表进行分区的功能和好处的相关概念。

- 第 2 章『范围集群表』提供了关于使用范围集群表的功能和优点的一般概念性信息。
- 第 3 章『多维集群 (MDC) 表』描述了使用多维集群作为对表中数据进行集群的极佳方法。
- 第 4 章『并行数据库系统』描述了如何利用并行性极大地提高性能。

## 第 2 部分 安装注意事项

下列章节提供了关于在准备数据库分区时需要执行的预安装和安装任务的信息。

- 第 5 章『安装先决条件』描述了与准备分区数据库环境中涉及的 DB2 服务器相关的先决条件和限制。
- 第 6 章『安装之前』讨论了 UNIX<sup>®</sup> 和 Linux<sup>®</sup> 系统中的其他预安装任务和注意事项。
- 第 7 章『安装 DB2 服务器产品』描述了如何安装数据库分区服务器和设置分区数据库环境。
- 第 8 章『安装之后』描述了如何验证 Windows<sup>®</sup>、UNIX 和 Linux 系统上的安装。

## 第 3 部分 实施和维护

完成规划、设计和安装步骤后，下列章节讨论了如何实施和维护先前准备工作生成的功能和/或对象。

- 第 9 章『创建数据库之前』描述了在创建数据库之前应考虑的事项（如启用并行性），以及在创建分区数据库环境、创建和配置节点/分区及建立数据库分区之间的通信之前应考虑的事项。
- 第 10 章『创建和管理分区数据库环境』描述了如何创建和管理数据库分区及分区组。
- 第 11 章『创建表和其他相关表对象』提供了关于如何创建和设置分区表、范围集群表及 MDC 表的信息。
- 第 12 章『改变数据库』描述了如何改变实例和/或数据库。
- 第 13 章『改变表和其他相关表对象』提供了关于如何修改分区表的信息。
- 第 14 章『装入』讨论了在并行性、多维集群和分区表情况下的装入注意事项。
- 第 15 章『在分区数据库环境中装入数据』描述了如何在分区数据库环境中插入、继续、重新启动或终止数据装入操作。
- 第 16 章『迁移分区数据库环境』简要概述了如何迁移分区数据库和对更详细信息的引用。
- 第 17 章『使用快照和事件监视器』除了描述如何使用 CREATE EVENT MONITOR 语句外，还提供了关于如何使用快照监视器结果来监视表重组过程或访问分区数据库系统的全局状态的相关信息。
- 第 18 章『开发好的备份和恢复策略』描述了分区数据库环境中的崩溃恢复相关概念，这有助于在出现故障前开发备份和恢复策略。
- 第 19 章『故障诊断』简要概述了关于如何在实例中的所有计算机上或所有数据库分区服务器上发出有助于进行故障诊断的命令（如 db2trc）的故障诊断和有用信息。

## 第 4 部分 性能问题

下列章节包含允许您提高分区和/或集群环境的性能的相关信息。

- 第 20 章『数据库设计中的性能问题』描述了表分区和多维集群的性能增强功能，包括表分区和多维集群的优化策略。
- 第 21 章『索引』提供有助于了解分区表索引的概念性信息。
- 第 22 章『设计顾问程序』描述了如何使用设计顾问程序来获取关于从单一分区迁移至多分区数据库的信息，以及关于分布数据和创建新索引、具体化查询表及多维集群表的建议。
- 第 23 章『管理并行性』提供了关于锁定方式的信息。
- 第 24 章『代理程序管理』描述了如何优化用于处理应用程序请求的数据库代理程序。
- 第 25 章『优化访问方案』描述了如何改善访问方案以及优化器如何使用通过各种扫描获得的信息来优化数据访问策略，并且包括关于连接策略的信息，所有这些都是为了提高使用并行性的分区数据库环境、集群表和/或系统的性能。
- 第 26 章『数据再分发』帮助您确定是否应执行数据再分发，如果应执行，如何在数据库分区之间再分发数据。
- 第 27 章『配置自调整内存』讨论了分区数据库环境中自调整内存功能的使用，并提供了配置建议。
- 第 28 章『DB2 配置参数和变量』提供了关于如何在多个数据库分区间设置数据库配置参数和环境变量的信息，并列示了与分区数据库环境和并行性功能相关的参数和变量。

## 第 5 部分 管理 API、命令和 SQL 语句

下列章节将与分区数据库环境相关的管理 API、命令和 SQL 元素的信息全部合并在一起。

- 第 29 章『管理 API』提供了仅与分区数据库环境相关的 API 的信息。
- 第 30 章『命令』提供了仅与分区数据库环境相关的命令的信息。
- 第 31 章『SQL 语言元素』提供了与数据库分区兼容的数据类型和专用寄存器。
- 第 32 章『SQL 函数』描述了仅与分区数据库环境相关的 SQL 函数。
- 第 33 章『SQL 语句』描述了仅与分区数据库环境相关的 SQL 语句。
- 第 34 章『受支持的管理 SQL 例程和视图』描述了仅与分区数据库环境相关的 SQL 例程和视图。

## 第 6 部分 附录

- 附录 A『作为非 root 用户安装』描述了作为非 root 用户将 DB2 产品安装在 UNIX 和 Linux 系统上的过程。
- 附录 B『使用备份』描述了如何使用 BACKUP DATABASE 命令。
- 附录 C『分区数据库环境目录视图』列示了分区数据库环境的特殊目录视图。

---

## 突出显示约定

本书中使用了下列突出显示约定。

---

|           |                        |
|-----------|------------------------|
| <b>粗体</b> | 指示命令、关键字和其他由系统预定义名称的项。 |
|-----------|------------------------|

---

---

斜体

指示下列其中一种情况:

- 必须由用户提供的名称或值（变量）
- 一般强调
- 引入新术语
- 引用其他信息源

---

等宽字体

指示下列其中一种情况:

- 文件和目录
  - 指示您应在命令提示符处或窗口中输入的信息
  - 特定数据值的示例
  - 类似于系统可能显示的文本的文本示例
  - 系统消息示例
  - 程序代码示例
-



---

## 第 1 部分 规划和设计注意事项





---

# 第 1 章 分区数据库和表

---

## 设置分区数据库环境

必须在创建数据库之前决定创建多分区数据库。在作出数据库设计决定时，必须确定是否应利用数据库分区可以提供的性能提高。

在分区数据库环境中，仍然使用 `CREATE DATABASE` 命令或 `sqlcrea()` 函数来创建数据库。无论使用哪种方法，都可以通过 `db2nodes.cfg` 文件中列示的任何分区来发出请求。

在创建多分区数据库之前，必须选择将作为数据库的目录分区的数据库分区。然后，可以直接从该数据库分区创建数据库，也可以从连接至该数据库分区的远程客户机创建数据库。您要连接并对其执行 `CREATE DATABASE` 命令的数据库分区成为该特定数据库的目录分区。

目录分区是用于存储所有系统目录表的数据库分区。对系统表的所有访问都必须通过此数据库分区进行。所有联合数据库对象（例如，包装器、服务器和昵称）都存储在此数据库分区上的系统目录表中。

若可能的话，应该在独立的实例中创建每个数据库。若不可能做到此点（即，必须在每个实例中创建多个数据库），应该将目录分区分布至可用的数据库分区中。这样可以减少在单个数据库分区中对目录信息的争用。

**注：**应该定期备份目录分区，同时，因为其他数据会增加备份所需的时间，所以要避免将用户数据置于该节点上（任何可能的時候）。

当创建数据库时，它会在 `db2nodes.cfg` 文件中定义的所有数据库分区之间自动创建。

创建系统中的第一个数据库时，就会形成一个系统数据库目录。并追加有关您创建的任何其他数据库的信息。在 UNIX 上工作时，系统数据库目录是 `sqlbdir`，位于主目录下的 `sqllib` 目录中或安装 DB2 数据库的目录下面。在 UNIX 上工作时，由于组成分区数据库环境的所有数据库分区只有一个系统数据库目录，所以此目录必须位于共享文件系统（例如，UNIX 平台上的 NFS）上。在 Windows 上工作时，系统数据库目录位于实例目录中。

位于 `sqlbdir` 目录中的还有系统意向文件。它称为 `sqlbins`，用于确保数据库分区保持同步。该文件也必须位于共享文件系统中，因为所有数据库分区中只有一个目录。该文件由组成数据库的所有数据库分区共享。

必须修改配置参数，才能利用数据库分区。使用 `GET DATABASE CONFIGURATION` 和 `GET DATABASE MANAGER CONFIGURATION` 命令以了解特定数据库或数据库管理器配置文件中的个别条目的值。要修改特定数据库或数据库管理器配置文件中的个别条目，可分别使用 `UPDATE DATABASE CONFIGURATION` 和 `UPDATE DATABASE MANAGER CONFIGURATION` 命令。

影响分区数据库环境的数据库管理器配置参数包括 `conn_elapse`、`fcm_num_buffers`、`fcm_num_channels`、`max_connretries`、`max_coordagents`、`max_time_diff`、`num_poolagents` 和 `stop_start_time`。

## 跨多个数据库分区的数据库分区

在将数据分布在分区数据库的多个数据库分区（节点）上时，数据库管理器提供了很大的灵活性。用户可以通过声明分布键来选择分布数据的方式，并且可以通过选择应存储数据的数据库分区组和表空间来确定数据可以分布在哪些数据库分区上以及分布在多少个数据库分区上。

此外，可更新的分发映射指定分布键值至数据库分区的映射。这使得可以灵活地将大型表的工作负载并行分布在分区数据库上，并同时允许将较小的表存储在一个或少数数据库分区中（如果应用程序设计者选择这样做的话）。每个本地数据库分区都可以有它所存储数据的本地索引，以便提供高性能的本地数据访问。

在分区数据库中，分布键用于将表数据分布到一组数据库分区中。也会分配索引数据及其相应的表，并将它们本地存储在每个数据库分区中。

在使用数据库分区存储数据之前，必须对数据库管理器定义这些数据库分区。在名为 `db2nodes.cfg` 的文件中定义数据库分区。

在 `CREATE TABLE` 语句或 `ALTER TABLE` 语句中为分区数据库分区组中的表空间中的表指定分布键。如果没有为表指定分布键，那么缺省情况下将根据主键的第一列创建分布键。如果未定义主键，那么缺省分布键是该表中定义的第一个非长型或 `LOB` 数据类型的列。分区数据库中的表必须至少有一列既非长型也非 `LOB` 数据类型。只有在显式指定后，单一分区数据库分区组中的表空间中的表才具有分布键。

按如下所示将行放置在数据库分区中：

1. 对所有分布键列应用散列算法（数据库分区函数），这将生成分发映射索引值。
2. 分发映射中该索引值处的数据库分区号标识将存储行的数据库分区。

数据库管理器支持部分分区，这意味着可以将表分布到系统中的数据库分区子集上（即，数据库分区组）。不必将表分布在系统中的所有数据库分区上。

在访问用于连接或子查询的数据时，数据库管理器能够识别这些数据是否位于同一数据库分区组中的同一数据库分区上。这称为表并置。在并置的表中具有相同分布键值的那些行位于同一数据库分区上。数据库管理器可以选择在存储该数据的数据库分区中执行连接或子查询处理。这样做的优点是可以显著改善性能。

并置的表必须：

- 位于将不再分发的同一数据库分区组中。（再分发期间，数据库分区组中的表可能正使用不同分发映射 - 它们不是并置的。）
- 具有与列数相同的分布键数。
- 具有与数据库分区兼容的相应分布键列。
- 位于在同一数据库分区中定义的单一分区数据库分区组中。

## 分区数据库认证注意事项

在一个分区数据库中，必须为数据库的每个分区定义同一组用户和组。如果这些定义不相同，那么用户也许是被授权在不同的分区上做不同的事情。建议所有分区保持一致。

## 数据库分区组

数据库分区组是定义为属于某个数据库的一个或多个数据库分区的集合。想要为数据库创建表时，首先创建用来存储表空间的数据库分区组，然后创建用来存储表的表空间。

可以在数据库中定义一个或多个数据库分区组成的命名子集。您定义的每个子集称为数据库分区组。包含多个数据库分区的每个子集称为多分区数据库分区组。多分区数据库分区组只能使用属于相同实例的数据库分区定义。数据库分区组可以只包括一个数据库分区，也可以包括数据库中的所有数据库分区。

图 1 给出了一个含五个数据库分区的数据库示例，在这个示例中：

- 数据库分区组横跨除一个数据库分区外的所有其他分区（数据库分区组 1）。
- 数据库分区组包含一个数据库分区（数据库分区组 2）。
- 数据库分区组包含两个数据库分区（数据库分区组 3）。
- 数据库分区组 2 中的数据库分区与数据库分区组 1 共享并重叠。
- 数据库分区组 3 中有一个数据库分区与数据库分区组 1 共享并重叠。

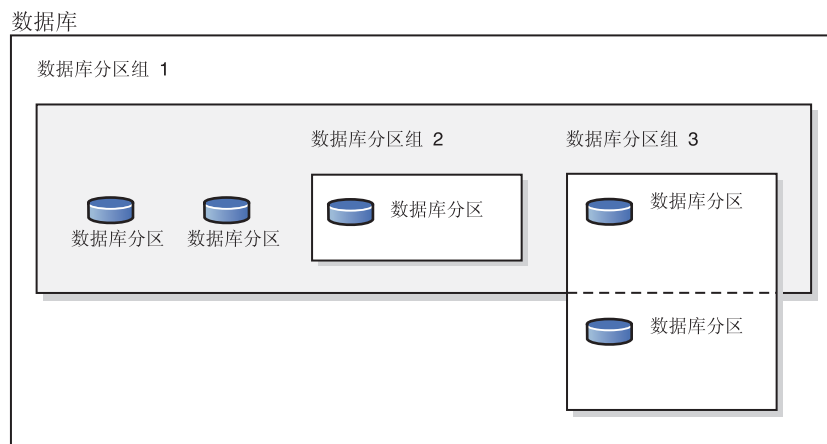


图 1. 数据库中的数据库分区组

使用 `CREATE DATABASE PARTITION GROUP` 语句创建新的数据库分区组。可以使用 `ALTER DATABASE PARTITION GROUP` 语句对其进行修改。数据分布在数据库分区组中的所有数据库分区上，并且可以从数据库分区组添加或删除一个或多个数据库分区。如果正在使用多分区数据库分区组，那么必须查看几个数据库分区组设计注意事项。

作为数据库系统配置一部分的每个数据库分区，必须已在名为 `db2nodes.cfg` 的数据库分区配置文件中定义。数据库分区组可以只包括一个数据库分区，也可以包括为该数据库系统定义的所有数据库分区。

创建或修改数据库分区组时，分发映射与它是相关联的。数据库管理器将分发映射和分布键及散列法联合使用来确定数据库分区组中的哪个数据库分区将存储给定的数据行。

在非分区数据库中，不需要任何分布键或分发映射。数据库分区是数据库的一部分，与用户数据、索引、配置文件和事务日志相辅相成。创建数据库时创建的缺省数据库分区组将由数据库管理器使用。IBM`CATGROUP` 是包含系统目录的表空间的缺省数据库分区组。IBM`TEMPGROUP` 是系统临时表空间的缺省数据库分区组。IBM`DEFAULTGROUP` 是包含用户定义的表的表空间的缺省数据库分区组（可选择在其中放置用户定义的表）。已声明临时表的用户临时表空间可在 IBM`DEFAULTGROUP` 或用户创建的任何数据库分区组中创建，但不能在 IBM`TEMPGROUP` 中创建。

处理数据库分区组时，您可以：

- 创建数据库分区组。
- 更改与数据库分区组相关联的注释。
- 将数据库分区添加至数据库分区组。
- 从数据库分区组中删除数据库分区。
- 在数据库分区组中再发表数据。

## 数据库分区组设计

如果使用单分区数据库，那么不考虑数据库分区组设计。“DB2 设计顾问程序”是一个可以用来建议数据库分区组的工具。可以从控制中心来访问“DB2 设计顾问程序”，也可以从命令行处理器中使用 `db2adviz` 来访问“DB2 设计顾问程序”。

若使用多分区数据库分区组，应考虑下列设计要点：

- 在多分区数据库分区组中，如果唯一索引是该分布键的超集，那么只能创建唯一索引。
- 根据该数据库中的数据库分区的数目，可能有一个或多个单分区数据库分区组，且存在一个或多个多分区数据库分区组。
- 必须给每个数据库分区指定唯一的编号。在一个或多个数据库分区组中，可能会发现相同的数据库分区。
- 要确保快速恢复包含系统目录表的数据库分区，应避免将用户表放在同一个数据库分区上。这可通过将用户表放在不包括 IBM`CATGROUP` 数据库分区组中的那些数据库分区的数据分区组中来实现。

除非要与一个更大的表并置，否则应该将小表放在单分区数据库分区组中。并置是将不同表中包含相关数据的行放置在同一个数据库分区中。并置的表允许 DB2 数据库 Linux 版、UNIX 版和 Windows 版利用更有效的连接策略。并置的表可以位于单分区数据库分区组中。如果表位于多分区数据库分区组中，且在分布键中具有相同数目的列，对应列的数据类型是兼容的，那么将这些表视为并置。在并置的表中具有相同分布键值的行被放置在同一个数据库分区上。这些表可以位于相同数据库分区组中的单独表空间中，且仍被视为是并置的。

应该避免将中等大小的表扩展到太多数据库分区上。例如，100 MB 的表在 16 个分区的数据分区组中可能比在 32 个分区的数据分区组中执行得更好。

可以使用数据库分区组将联机事务处理（OLTP）表与决策支持（DSS）表分开，以确保不会对 OLTP 事务的性能产生负面影响。

## 分发映射

在分区数据库环境中，数据库管理器必须知道到哪里去查找所需的数据。数据库管理器使用一个称为分发映射的映射来查找数据。

分发映射是一个内部生成的数组，对于多分区数据库分区组，它包含 4096 个条目，对于单分区数据库分区组，它只包含一个条目。对于单分区数据库分区组，分发映射只有一个条目，该条目包含存储数据库表的所有行的数据库分区的编号。对于多分区数据库分区组，数据库分区组的编号指定方式使得能够一个接一个地使用每个数据库分区，以确保整个映射分发均匀。正如使用网格将城市地图划分为区一样，数据库管理器使用分布键来确定存储数据的位置（数据库分区）。

例如，假定您将一个数据库创建在四个数据库分区（编号为 0 到 3）上。此数据库的 IBMDEFAULTGROUP 数据库分区组的分发映射将是：

0 1 2 3 0 1 2 ...

如果已使用数据库分区 1 和 2 在该数据库中创建了一个数据库分区组，那么该数据库分区组的分发映射将是：

1 2 1 2 1 2 1 ...

如果要装入到数据库的一个表的分布键是一个可能在范围 1 到 500000 之间的整数，那么会将分布键散列至 0 到 4095 之间的一个编号。将该编号用作分发映射中的索引，以选择用于该行的数据库分区。

图 2 显示如何将具有分布键值 (c1, c2, c3) 的行映射至编号 2，然后引用数据库分区 n5。

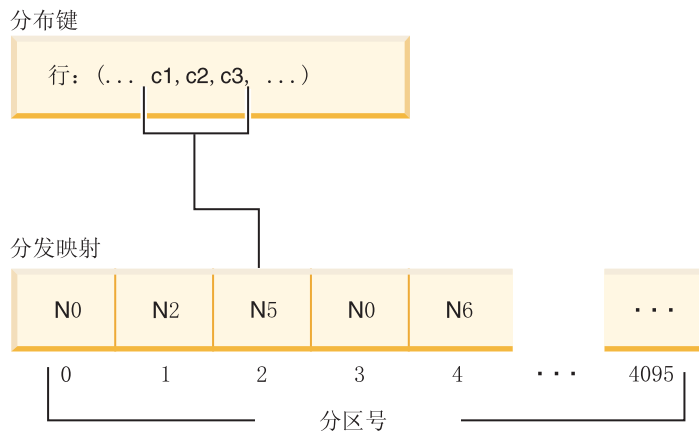


图 2. 使用分发映射的数据分发

分发映射可以灵活地控制将数据存储在多分区数据库中的哪个位置。若需要更改数据库中各数据库分区上的数据分发，可以使用数据再分发实用程序。此实用程序允许重新平衡或调整数据分发的偏差。

可以使用 `sqlugtpi` API 来获取您可以查看的分发映射的副本。

## 分布键

分布键是一列（或一组列），用于确定存储特定数据行的数据库分区。分布键是使用 `CREATE TABLE` 语句在表上定义的。如果没有为分布在数据库分区组中的多个数据库分区中的表空间中的表定义分布键，在缺省情况下将会根据主键的第一列创建分布键。

如果未指定主键，那么缺省分布键是在该表中定义的第一个非长字段列。（长型包括所有长型数据类型和所有大对象（LOB）数据类型）。如果正在与单一分区数据库分区组相关的表空间中创建表，且您希望有分布键，那么必须显式定义该分布键。缺省情况下，不会创建该分布键。

如果没有列满足缺省分布键的需求，那么会创建不带分布键的表。仅允许没有分布键的表存在于单一分区数据库分区组中。以后可以使用 `ALTER TABLE` 语句来添加或删除分布键。仅可对其表空间与单一分区数据库分区组相关的表改变分布键。

选择好的分布键很重要。应考虑下列各项：

- 如何访问表
- 查询工作负载的性质
- 数据库系统所使用的连接策略

如果并置不是主要的注意事项，那么好的表分布键就是将数据均匀分布在数据库分区组中的所有数据库分区上的分布键。与数据库分区组相关联的表空间中每个表的分布键确定这些表是否是并置的。下列情况中，表被认为是并置的：

- 表被放置在同一数据库分区组中的表空间内
- 每个表中的分布键具有相同数量的列
- 对应列的数据类型是分区兼容的

这些特征确保并置的表中具有相同分布键值的那些行位于同一个数据库分区上。

不适当的分布键会导致数据分布不均匀。不应选择数据分布不均匀的列和含有少数单值的列作为分布键。单值的数目必须足够大，才能确保将行均匀分布在数据库分区组中的所有数据库分区上。应用分发算法的成本与分布键的大小是成正比的。分布键不能超过 16 列，而且列越少，性能越好。不应将不需要的列包括在分布键中。

当定义分布键时，应该考虑以下几点：

- 不支持创建只包含长型数据类型（`LONG`、`VARCHAR`、`LONG VARCHAR`、`BLOB`、`CLOB` 或 `DBCLOB`）的多分区表。
- 不能改变分布键定义。
- 分布键应该包括最频繁连接的列。
- 分布键应该由经常参与 `GROUP BY` 子句的列组成。
- 任何唯一键或主键必须包含所有分布键列。
- 在联机事务处理（OLTP）环境中，分布键中的所有列都应该使用带常量或主变量的等于（=）谓词来参与该事务。例如，假定有一个在事务中经常使用的职员号 `emp_no`，如：

```
UPDATE emp_table SET ... WHERE  
emp_no = host-variable
```

在此情况下，EMP\_NO 列对于 EMP\_TABLE 而言是一个不错的单列分布键。

数据库分区是确定表中每一行的位置的方法。该方法的原理如下：

1. 将散列算法应用于分布键的值，并生成一个介于零（0）与 4095 之间的编号。
2. 创建数据库分区组时将创建分发映射。每个编号依次以循环方式重复，以填写该分发映射。
3. 将该编号用作分发映射中的索引。分发映射中该位置处的编号是存储该行的数据库分区的编号。

## 表并置

您可能会发现，作为对特定查询的响应，两个或多个表频繁地提供数据。在此情况下，您会希望这样的表中的相关数据的位置尽可能地靠近。在数据库被物理地划分为两个或多个数据库分区的环境中，必须有一种方法可将划分的表的相关碎片尽可能地靠近。完成此过程的功能称为表并置。

当表存储在同一个数据库分区组中且它们的分布键兼容时，这些表就是并置的。将两个表置于同一个数据库分区组中以确保存在公共的分发映射。这些表可能位于不同的表空间，但是这些表空间必须与相同数据库分区组相关联。每个分布键中对列的数据类型必须是分区兼容的。

当访问用于连接或子查询的多个表时，DB2 数据库 Linux 版、UNIX 版和 Windows 版能够识别要连接的数据是否位于相同数据库分区上。当发生这种情况时，DB2 可以在存储该数据的数据库分区上执行连接或子查询，而不必在数据库分区之间移动数据。此功能的优点是可以显著改善性能。

## 分区兼容性

对分布键的对应列的基本数据类型进行比较，并可将它们声明为分区兼容。分区兼容的数据类型具有如下属性：具有相同值但不同类型的两个变量会按相同的分区算法映射至同一个编号。

分区兼容性具有下列特征：

- 基本数据类型与另一个相同的基本数据类型兼容。
- 内部格式用于 DATE、TIME 和 TIMESTAMP 数据类型。它们彼此都不兼容，且都不与 CHAR 兼容。
- 分区兼容性不受带有 NOT NULL 或 FOR BIT DATA 定义的列的影响。
- 对兼容数据类型的 NULL 值的处理是完全相同的；对不兼容数据类型的 NULL 值的处理可能不相同。
- 用户定义的类型的基本数据类型用于分析分区兼容性。
- 对分布键中值相同的小数的处理是完全相同的，即使它们的小数位和精度不同也是如此。
- 字符串中（CHAR、VARCHAR GRAPHIC 或 VARGRAPHIC）的尾部空格会被散列算法忽略。
- BIGINT、SMALLINT 和 INTEGER 是兼容的数据类型。
- 不同长度的 CHAR 和 VARCHAR 是兼容的数据类型。

- GRAPHIC 和 VARGRAPHIC 是兼容的数据类型。
- 分区兼容性不适用于 LONG VARCHAR、LONG VARGRAPHIC、CLOB、DBCLOB 和 BLOB 数据类型，因为不支持它们作为分布键。

---

## 分区表

表分区功能可用于 DB2 企业服务器版 Linux 版、UNIX 版和 Windows 版版本 9.1 和更高版本。分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。

数据分区或范围是一个表的一部分，它包含该表的一部分行，并且与其他行集分开存储。根据 CREATE TABLE 语句的 PARTITION BY 子句中指定的内容，给定表的数据被划分到多个数据分区或范围中。这些数据分区或范围可位于不同表空间和/或同一表空间中。如果一个表是使用 PARTITION BY 子句创建的，那么该表是分区表。

指定的所有表空间在下列方面必须相同：页大小、扩展数据块大小、存储机制（DMS 和 SMS）和类型（常规或大型），并且所有表空间必须位于同一数据库分区组中。

分区表能够简化表数据的转入和转出，并且，与普通的表相比，分区表包含的数据可以多得多。分区表最多可以有 32767 个数据分区。可以对分区表添加数据分区、将数据分区与分区表相连以及断开数据分区与分区表的连接，并且，可以将一个表的多个数据分区范围存储在一个表空间中。

**限制：**不支持在分区表中使用分区的分层表或临时表、范围集群表和分区视图。也不支持在分区表中使用 XML 列类型。

## 表分区

表分区功能是一种数据组织方案，即，表数据根据一个或多个表列中的值分布到多个存储对象（称为数据分区或范围）中。每个数据分区都是单独存储的。这些存储对象可以在不同的表空间中，也可以在相同表空间中。

存储对象的行为与单个表的行为非常类似，它通过使用 ALTER TABLE ...ATTACH 语句将现有表合并到分区表中，可以很容易实现快速转入。同样，使用 ALTER TABLE ...DETACH 语句很容易实现转出。查询处理同样可以利用分离的数据来避免扫描不相关数据，从而使许多数据仓库样式查询具有更好地查询性能。

按照 CREATE TABLE 语句的 PARTITION BY 子句中指定那样将表数据分区。此定义中使用的列被称为表分区键列。

这种组织方案可以单独地使用，也可与其他组织方案结合使用。通过组合使用 CREATE TABLE 语句的 DISTRIBUTE BY 和 PARTITION BY 子句，可以将数据分布在跨多个表空间的数据库分区之间。该组织方案包括：

- DISTRIBUTE BY HASH
- PARTITION BY RANGE
- ORGANIZE BY DIMENSIONS

表分区功能可用于 DB2 企业服务器版 Linux 版、UNIX 版和 Windows 版版本 9.1 和更高版本。



## 表分区的优点

如果下列任何情况适用于您和您的组织，请考虑表分区的许多优点：

- 更容易转入和转出表数据使数据仓库受益
- 数据仓库包括大型表
- 您正在考虑从先前发行版或某个竞争数据库产品迁移至版本 9.1 数据库
- 需要更有效地使用分层存储管理（HSM）解决方案

表分区简化了表数据转入和转出以及管理工作，并且提高了索引位置的灵活性和查询处理效率。

### 有效转入和转出

表分区功能提高了表数据的转入和转出效率。这可通过使用 ALTER TABLE 语句的 ATTACH PARTITION 和 DETACH PARTITION 子句来实现。通过转入分区表数据，可以方便地将新范围作为附加数据分区合并到分区表中。通过转出分区表数据，可以方便地从分区表中分离出某些范围的数据，以进行后续清除或归档处理。

### 更容易管理大型表

由于您可以在各个数据分区上执行管理任务，因此表级别管理更灵活。这些任务包括：拆离和重新连接数据分区、备份和复原各个数据分区以及重组各个索引。通过将花费较长时间的维护操作分解成一系列较小的操作，可以缩短这种维护操作的执行时间。例如，将数据分区放置在单独的表空间中后，备份操作可以逐个处理数据分区。因此，可以一次备份分区表的一个数据分区。

### 灵活的索引位置

现在，可以将索引放置在不同的表空间中，从而允许对索引位置进行更精细地控制。这种新设计具有以下一些好处：

- 提高了删除索引和联机索引创建的性能
- 能够针对每个表索引之间的任何表空间特征使用不同的值（例如，为了确保更好的空间利用率，对每个索引使用不同的页大小可能更合适）。
- 减少 IO 争用并提供对表索引数据更有效的并发访问。
- 删除各个索引时，空间将立即可供系统使用，而无需进行索引重组。
- 如果您选择执行索引重组，可以重组单个索引。

DMS 和 SMS 表空间都支持在不同于表的另一个位置使用索引。

### 提高了商业智能样式查询的性能

增强了查询处理功能，能够根据查询谓词自动消除某些数据分区。此功能称为数据分区消除，它为许多决策支持查询带来好处。

以下示例创建了一个表 *customer*，其中包含 `l_shipdate >= '01/01/2006'` 和 `l_shipdate <= '03/31/2006'` 的行存储在表空间 *ts1* 中，包含 `l_shipdate >= '04/01/2006'` 和 `l_shipdate <= '06/30/2006'` 的行存储在表空间 *ts2* 中，等等。

```
CREATE TABLE customer (l_shipdate DATE, l_name CHAR(30))
IN ts1, ts2, ts3, ts4, ts5
PARTITION BY RANGE(l_shipdate) (STARTING FROM ('01/01/2006')
ENDING AT ('12/31/2006') EVERY (3 MONTHS))
```

## 数据分区和范围

分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。可以自动生成对每个数据分区指定的范围，也可以在创建表时手动生成这些范围。

数据分区在整个 DB2 库中以多种方式被引用。下表列示了最常见的引用方式：

- **DATAPARTITIONNAME** 是创建给定的表时对数据分区指定的永久名称。此列值存储在 **SYSCAT.DATAPARTITIONS** 目录视图中。连接或拆离操作不保留此名称。
- **DATAPARTITIONID** 是创建给定的表时对数据分区指定的永久标识。该标识用于唯一地标识给定表中的特定数据分区。连接或拆离操作不保留此标识。此值由系统生成，它会出现在各个实用程序的输出中。
- **SEQNO** 表示特定数据分区范围相对于表中其他数据分区范围的顺序，其中，拆离数据分区排列在所有可视数据分区和连接数据分区后面。

## 数据组织方案

在引入表分区之后，DB2 数据库提供了一种三级数据组织方案。**CREATE TABLE** 语句中有三个子句包含指示应如何组织数据的算法。

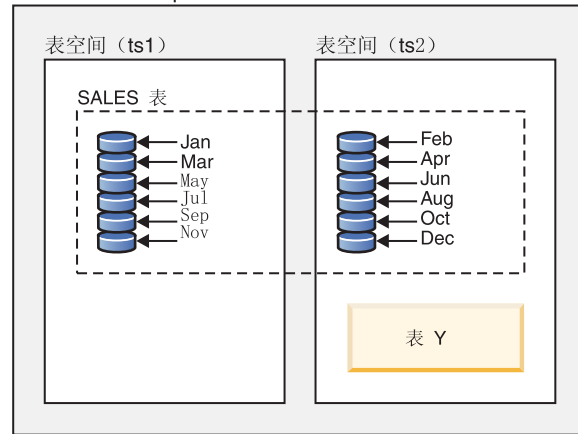
下列三个子句演示了可以任意组合在一起使用的数据组织级别：

- **DISTRIBUTE BY** 用于将数据均匀地分布在数据库分区上（以启用查询内并行性并平衡每个数据库分区上的负载）（数据库分区）
- **PARTITION BY** 用于对同一个数据分区中具有类似单维值的行进行分组（表分区）
- **ORGANIZE BY** 用于对同一表扩展数据块中在多个维上具有类似值的行进行分组（多维集群）

此语法允许子句之间保持一致并允许进一步的数据组织算法。每个子句可以单独地使用，也可与其他子句结合使用。通过组合使用 **CREATE TABLE** 语句的 **DISTRIBUTE BY** 和 **PARTITION BY** 子句，可以在跨多个表空间的数据库分区之间分布数据。此方法允许类似于 Informix<sup>®</sup> Dynamic Server 和 Informix Extended Parallel Server 混合功能的行为。

在一个表中，可以组合每个数据组织方案中使用的子句来创建更先进的分区方案。例如，DB2 数据库分区功能部件（DPF）不仅兼容，而且是对表分区的补充。

数据库分区 (dbpart1)



图注



图 3. 演示表分区组织方案，其中表示月销售数据的表划分到多个数据分区中。该表还跨越两个表空间 (ts1 和 ts2)。

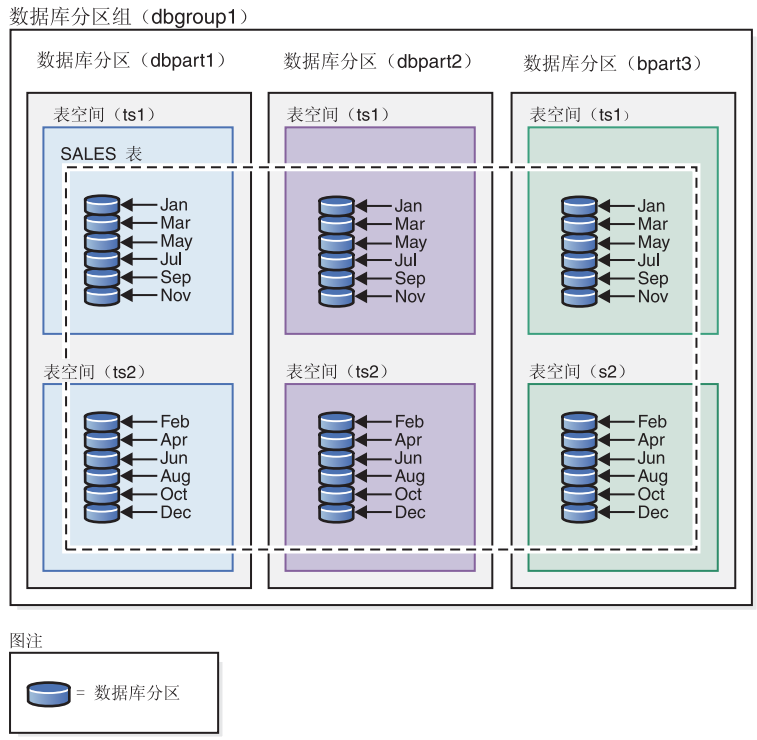


图 4. 演示数据库分区与表分区相互补充的组织方案。表示月销售数据的表划分到多个数据分区中并且跨两个表空间 (ts1 和 ts2)，而这两个表空间又分布在数据库分区组 (dbgroup1) 的多个数据库分区 (dbpart1、dbpart2 和 dbpart3) 中。

多维集群 (MDC) 与表分区之间的显著区别是多维和单维。MDC 适合于立方体 (即, 具有多维的表), 而当有一个维是数据库设计的中心时 (例如, DATE 列), 表分区就能很好地起作用。当同时符合这两个条件时, MDC 和表分区互补。这在第 15 页的图 5 中进行了演示。

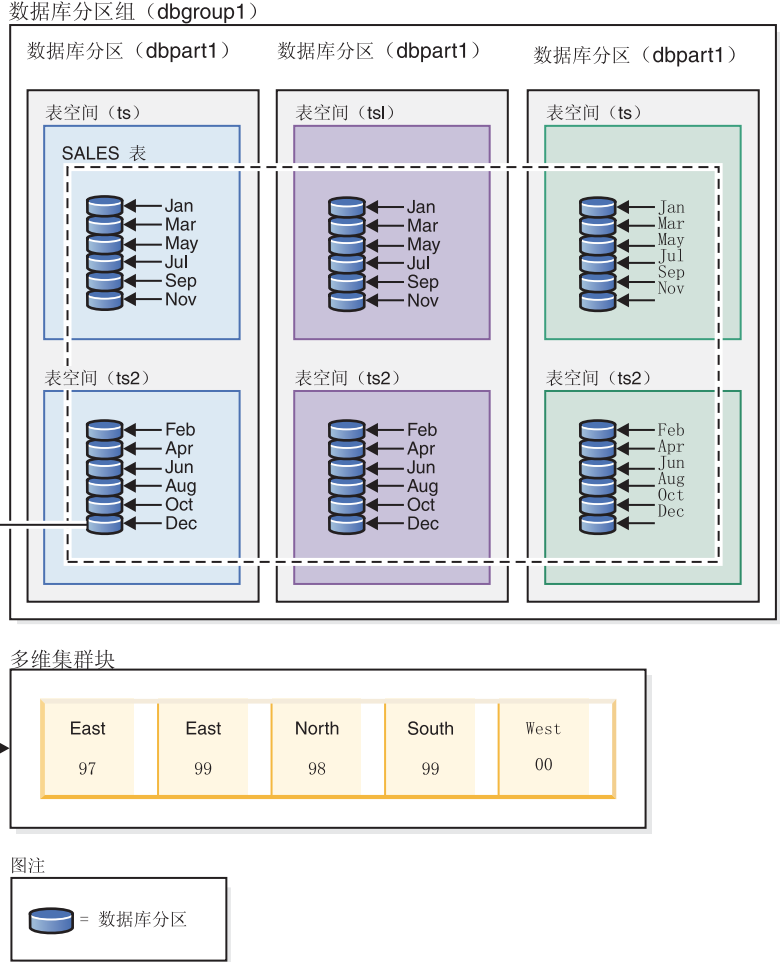


图 5. 数据库分区、表分区和多维组织方案的表示法，其中 SALES 表中的数据不仅分布在多个数据库分区中、划分到表空间 ts1 和 ts2 上，而且还对 date 和 region 维上具有类似值的行进行分组。

还有一种数据组织方案，它不能与上面列示的任何方案一起使用。此方案是 ORGANIZE BY KEY SEQUENCE。它用于将每个记录插入到在创建表（范围集群的表）时为该记录保留的行中。

### 数据组织术语

#### 数据库分区 (Database partitioning)

一种数据组织方案，即，表数据根据该表中的一个或多个分布键列中的散列值以及使用的数据库分区的分发映射分布到多个数据库分区中。给定表的数据根据 CREATE TABLE 语句的 DISTRIBUTE BY HASH 子句中指定的内容进行分布。

#### 数据库分区 (Database partition)

数据库分区服务器上的一部分数据库，它由自己的用户数据、索引、配置文件和事务日志组成。数据库分区可以是逻辑或物理的。

### 表分区 (Table partitioning)

一种数据组织方案，即，表数据根据该表中一个或多个分区列中的值分布到多个数据分区中。根据 CREATE TABLE 语句的 PARTITION BY 子句中指定的内容，给定表的数据被划分到多个存储对象中。这些存储对象可以在不同的表空间中。

### 数据分区 (Data partition)

表的一部分行，这些行不与其他部分的行存储在一起，并且按照 CREATE TABLE 语句的 PARTITION BY RANGE 子句中指定的内容分组。

### 多维集群 (Multidimensional clustering, MDC)

一个表，其数据按 ORGANIZE BY DIMENSIONS 子句中指定的一个或多个维或者集群键以物理方式组织成块。

## 每种数据组织方案的优势

了解每种数据组织方案的优势可以帮助您在规划、设计或重新评估数据库系统需求时确定最佳方法。表 2 提供了普通客户需求的高级视图，并显示了各种数据组织方案可以帮助您满足这些需求。

表 2. 将表分区与数据库分区功能部件配合使用

| 问题            | 建议的方案   | 说明                              |
|---------------|---------|---------------------------------|
| 数据转出          | 表分区     | 使用拆离来转出大量数据，并且只出现最少中断。          |
| 并行查询执行 (查询性能) | 数据库分区功能 | 提供查询并行性以改善查询性能                  |
| 数据分区消除 (查询性能) | 表分区     | 提供数据分区消除以改善查询性能                 |
| 获得最佳查询性能      | 两者      | 一起使用时可以获得最佳查询性能: 查询并行性和数据分区消除互补 |
| 管理员的工作负载较重    | 数据库分区功能 | 为每个数据库分区执行许多任务                  |

表 3. 将表分区与 MDC 表配合使用

| 问题         | 建议的方案 | 说明                                       |
|------------|-------|--|
| 转出期间的数据可用性 | 表分区   | 使用 DETACH PARTITION 子句来转出大量数据，并且只出现最少中断。 |
| 查询性能       | 两者    | MDC 最适合用来查询多个维。表分区通过数据分区消除提高性能。          |
| 最少重组       | MDC   | MDC 维护集群，从而减少进行重组的必要性。                   |

注: 对于 UNION ALL 视图，目前建议使用表分区。

## DB2 和 Informix 数据库中的数据组织方案

表分区功能是一种数据组织方案，即，表数据根据一个或多个表列中的值分布到多个存储对象（称为数据分区或范围）中。每个数据分区都是单独存储的。这些存储对象可以在不同的表空间中，也可以在相同表空间中。

按照 CREATE TABLE 语句的 PARTITION BY 子句中指定那样将表数据分区。此定义中使用的列被称为表分区键列。DB2 表分区功能与 Informix Dynamic Server 和 Informix Extended Parallel Server 提供的分段组织方法相对应。

### Informix 方法

Informix 支持若干种数据组织方案，在 Informix 产品中，这些方案称为分段。其中一种较常使用的分段类型是 FRAGMENT BY EXPRESSION。这种类型的分段的工作方式与 CASE 语句非常相似，其中有一个与表的每个分段相关联的表达式。检查这些表达式以便确定行的放置位置。

### Informix 与 DB2 数据库系统的比较

DB2 数据库提供了丰富的补充功能，这些功能与 Informix 数据组织方案直接对应，这使客户能够相对容易地从 Informix 语法转换为 DB2 语法。DB2 数据库管理器将生成列与 CREATE TABLE 语句的 PARTITION BY RANGE 子句配合使用，以处理复杂的 Informix 方案。表 4 对 Informix 和 DB2 数据库产品中使用的数据组织方案作了比较。

表 4. 所有 Informix 与 DB2 数据组织方案的映射

| 数据组织方案   | Informix 语法             | DB2 版本 9.1 语法                           |
|--|-------------------------|---|
| • Informix: 基于表达式<br>• DB2: 表分区                | FRAGMENT BY EXPRESSION  | PARTITION BY RANGE                      |
| • Informix: 循环法<br>• DB2: 缺省                   | FRAGMENT BY ROUND ROBIN | 没有语法: DB2 数据库管理器自动在容器之间分布数据             |
| • Informix: 范围分布<br>• DB2: 表分区                 | FRAGMENT BY RANGE       | PARTITION BY RANGE                      |
| • Informix: 系统定义的散列<br>• DB2: 数据库分区            | FRAGMENT BY HASH        | DISTRIBUTE BY HASH                      |
| • Informix: HYBRID<br>• DB2: 在进行表分区的情况下进行数据库分区 | FRAGMENT BY HYBRID      | DISTRIBUTE BY HASH 和 PARTITION BY RANGE |
| • Informix: 不适用<br>• DB2: 多维集群                 | 不适用                     | ORGANIZE BY DIMENSION                   |

### 示例

下列示例详细说明了在 DB2 中如何实现与任何使用表达式的 Informix 分段方案等同的结果。

示例 1: 下面这个基本的 Create Table 语句显示了 Informix 分段以及等同的 DB2 数据库系统表分区语法:

Informix 语法:

```
CREATE TABLE demo(a INT) FRAGMENT BY EXPRESSION
a = 1 IN db1,
a = 2 IN db2,
a = 3 IN db3;
```

DB2 语法:

```
CREATE TABLE demo(a INT) PARTITION BY RANGE(a)
(STARTING(1) IN db1,
STARTING(2) IN db2,
STARTING(3) ENDING(3) IN db3);
```

Informix XPS 支持称为 hybrid 的两层分段方案，在此方案中，使用一个表达式来在协作服务器之间分布数据，并使用第二个表达式来在协作服务器内分布数据。这使所有协作服务器都能够参与查询（即，在所有协作服务器上都有数据），并使查询能够利用数据分区消除功能的优势。

通过结合使用 CREATE TABLE 语句的 DISTRIBUTE BY 和 PARTITION BY 子句，DB2 数据库系统实现了与 Informix hybrid 等同的组织方案。

示例 2: 以下示例显示了组合子句的语法:

Informix 语法

```
CREATE TABLE demo(a INT, b INT) FRAGMENT BY HYBRID HASH(a)
EXPRESSION b = 1 IN dbs11,
b = 2 IN dbs12;
```

DB2 语法

```
CREATE TABLE demo(a INT, b INT) IN dbs11, dbs12
DISTRIBUTE BY HASH(a),
PARTITION BY RANGE(b) (STARTING 1 ENDING 2 EVERY 1);
```

此外，可以使用多维集群来进一步组织数据:

```
CREATE TABLE demo(a INT, b INT, c INT) IN dbs11, dbs12
DISTRIBUTE BY HASH(a),
PARTITION BY RANGE(b) (STARTING 1 ENDING 2 EVERY 1)
ORGANIZE BY DIMENSIONS(c);
```

这样，列 **a** 值相同的所有行都在同一个数据库分区中。所有列 **b** 值相同的行都在同一个表空间中。对于具有给定的 **a** 和 **b** 值的行，会将再具有相同 **c** 值的所有行集中到一个磁盘上。此方法非常适合于 OLAP 类型的下寻操作，这是因为，仅需扫描单个数据库分区上单个表空间中的一个或数个扩展数据块就可以满足此类查询。

## 应用表分区以解决常见的应用程序问题

下列各节讨论如何应用各种 DB2 表分区功能以解决常见的应用程序问题。每一节都特别侧重于采取最佳措施来将各种 Informix 分段方案映射到等同的 DB2 表分区方案。

### 创建简单数据分区范围时的注意事项

其中一种最常见的表分区应用是根据日期键对大型事实表进行分区。如果需要创建大小统一的日期范围，请考虑使用自动生成的 CREATE TABLE 语法格式。



## 示例

示例 1: 以下示例显示自动生成的语法格式:

```
CREATE TABLE orders
(
  l_orderkey DECIMAL(10,0) NOT NULL,
  l_partkey INTEGER,
  l_suppkey INTEGER,
  l_linenummer INTEGER,
  l_quantity DECIMAL(12,2),
  l_extendedprice DECIMAL(12,2),
  l_discount DECIMAL(12,2),
  l_tax DECIMAL(12,2),
  l_returnflag CHAR(1),
  l_linestatus CHAR(1),
  l_shipdate DATE,
  l_commitdate DATE,
  l_receiptdate DATE,
  l_shipinstruct CHAR(25),
  l_shipmode CHAR(10),
  l_comment VARCHAR(44)
  PARTITION BY RANGE(l_shipdate)
  (STARTING '1/1/1992' ENDING '12/31/1993' EVERY 1 MONTH);
```

这将创建 24 个范围, 即对 1992-1993 的每个月创建一个范围。尝试插入 l\_shipdate 超出该范围的行将导致错误。

示例 2: 将上一示例与以下 Informix 语法作比较:

```
create table orders
(
  l_orderkey decimal(10,0) not null,
  l_partkey integer,
  l_suppkey integer,
  l_linenummer integer,
  l_quantity decimal(12,2),
  l_extendedprice decimal(12,2),
  l_discount decimal(12,2),
  l_tax decimal(12,2),
  l_returnflag char(1),
  l_linestatus char(1),
  l_shipdate date,
  l_commitdate date,
  l_receiptdate date,
  l_shipinstruct char(25),
  l_shipmode char(10),
  l_comment varchar(44)
) fragment by expression
l_shipdate < '1992-02-01' in ldbs1,
l_shipdate >= '1992-02-01' and l_shipdate < '1992-03-01' in ldbs2,
l_shipdate >= '1992-03-01' and l_shipdate < '1992-04-01' in ldbs3,
l_shipdate >= '1992-04-01' and l_shipdate < '1992-05-01' in ldbs4,
l_shipdate >= '1992-05-01' and l_shipdate < '1992-06-01' in ldbs5,
l_shipdate >= '1992-06-01' and l_shipdate < '1992-07-01' in ldbs6,
l_shipdate >= '1992-07-01' and l_shipdate < '1992-08-01' in ldbs7,
l_shipdate >= '1992-08-01' and l_shipdate < '1992-09-01' in ldbs8,
l_shipdate >= '1992-09-01' and l_shipdate < '1992-10-01' in ldbs9,
l_shipdate >= '1992-10-01' and l_shipdate < '1992-11-01' in ldbs10,
l_shipdate >= '1992-11-01' and l_shipdate < '1992-12-01' in ldbs11,
l_shipdate >= '1992-12-01' and l_shipdate < '1993-01-01' in ldbs12,
l_shipdate >= '1993-01-01' and l_shipdate < '1993-02-01' in ldbs13,
l_shipdate >= '1993-02-01' and l_shipdate < '1993-03-01' in ldbs14,
l_shipdate >= '1993-03-01' and l_shipdate < '1993-04-01' in ldbs15,
l_shipdate >= '1993-04-01' and l_shipdate < '1993-05-01' in ldbs16,
l_shipdate >= '1993-05-01' and l_shipdate < '1993-06-01' in ldbs17,
```

```

l_shipdate >= '1993-06-01' and l_shipdate < '1993-07-01' in ldfs18,
l_shipdate >= '1993-07-01' and l_shipdate < '1993-08-01' in ldfs19,
l_shipdate >= '1993-08-01' and l_shipdate < '1993-09-01' in ldfs20,
l_shipdate >= '1993-09-01' and l_shipdate < '1993-10-01' in ldfs21,
l_shipdate >= '1993-10-01' and l_shipdate < '1993-11-01' in ldfs22,
l_shipdate >= '1993-11-01' and l_shipdate < '1993-12-01' in ldfs23,
l_shipdate >= '1993-12-01' and l_shipdate < '1994-01-01' in ldfs24,
l_shipdate >= '1994-01-01' in ldfs25;

```

注意, Informix 语法提供了上下不封顶的范围以捕获预期范围外的日期。通过添加使用 MINVALUE 和 MAXVALUE 的范围, 可以将 DB2 语法修改为与 Informix 语法匹配。

示例 3: 以下示例将示例 1 修改为与 Informix 语法匹配:

```

CREATE TABLE orders
(
  l_orderkey DECIMAL(10,0) NOT NULL,
  l_partkey INTEGER,
  l_suppkey INTEGER,
  l_linenummer INTEGER,
  l_quantity DECIMAL(12,2),
  l_extendedprice DECIMAL(12,2),
  l_discount DECIMAL(12,2),
  l_tax DECIMAL(12,2),
  l_returnflag CHAR(1),
  l_linestatus CHAR(1),
  l_shipdate DATE,
  l_commitdate DATE,
  l_receiptdate DATE,
  l_shipinstruct CHAR(25),
  l_shipmode CHAR(10),
  l_comment VARCHAR(44)
) PARTITION BY RANGE(l_shipdate)
(STARTING MINVALUE,
 STARTING '1/1/1992' ENDING '12/31/1993' EVERY 1 MONTH,
 ENDING MAXVALUE);

```

这种技术允许将任何日期插入到表中。

## 使用生成列按表达式进行分区

虽然 DB2 数据库并不直接支持按表达式进行分区, 但支持按生成列进行分区, 因此有可能获得相同的结果。

在决定是否使用此方法前, 请考虑下列用法准则:

- 生成列是真实的列, 它占用物理磁盘空间。使用生成列的表会略微变大。
- 对于在分区表进行分区时所基于的列, 不能改变其生成列表达式。尝试执行此操作将产生消息 SQL0190。如果按下一节描述的方式将新数据分区添加到使用生成列的表中, 通常要求改变定义生成列的表达式。目前, 不支持改变定义生成列的表达式。
- 当表使用生成列时, 对于何时可以应用数据分区消除是有限制的。

## 示例

示例 1: 以下示例使用 Informix 语法, 在这种情况下适合使用生成列。在本示例中, 进行分区时所基于的列存放了加拿大的省和地域。由于省列表不大可能更改, 因此生成列表达式也不大可能更改。

```

CREATE TABLE customer (
  cust_id INT,
  cust_prov CHAR(2))

```

```

FRAGMENT BY EXPRESSION
cust_prov = "AB" IN dbspace_ab
cust_prov = "BC" IN dbspace_bc
cust_prov = "MB" IN dbspace_mb
...
cust_prov = "YT" IN dbspace_yt
REMAINDER IN dbspace_remainder;

```

示例 2: 在此示例中, 使用生成列对 DB2 表进行分区:

```

CREATE TABLE customer (
  cust_id      INT,
  cust_prov   CHAR(2),
  cust_prov_gen GENERATED ALWAYS AS (CASE
    WHEN cust_prov = 'AB' THEN 1
    WHEN cust_prov = 'BC' THEN 2
    WHEN cust_prov = 'MB' THEN 3
    ...
    WHEN cust_prov = 'YT' THEN 13
    ELSE 14 END))
IN tbspace_ab, tbspace_bc, tbspace_mb, .... tbspace_remainder
PARTITION BY RANGE (cust_prov_gen)
(STARTING 1 ENDING 14 EVERY 1);

```

这里, CASE 语句中的表达式与 FRAGMENT BY EXPRESSION 子句中的相应表达式匹配。CASE 语句将每个原始表达式映射到一个数字, 该数字存储在生成列(在本示例中是 cust\_prov\_gen)中。此列是存储在磁盘上的真实列, 因此, 表占用的空间量会比 DB2 通过表达式直接支持的分区所必需的空间量略多。本示例使用短语法格式。因此, 必须在 CREATE TABLE 语句的 IN 子句中列示用来放置数据分区的表空间。如果使用长语法格式, 那么每个数据分区都需要不同的 IN 子句。

注: 这种技术可以应用于任何 FRAGMENT BY EXPRESSION 子句。

## 表分区键

表分区键是一个或多个表列的有序集合。表分区键列中的值用来确定每个表行所属的数据分区。

要对表定义表分区键, 请使用指定了 PARTITION BY 子句的 CREATE TABLE 语句。

选择有效的表分区键列对于充分利用表分区功能的优点来说十分关键。下列准则可以帮助您为分区表选择最有效的表分区键列:

- 将范围定义成与数据转入大小相匹配。最常见的情况是根据日期或时间列对数据进行分区。
- 将范围详细程度定义为与数据转出相匹配。最常见的情况是使用月份或季度。
- 根据有益于消除分区的列进行分区。

## 支持的数据类型

表 5 显示了支持用作表分区键列的数据类型 (包括同义词):

表 5. 支持的数据类型

| 数据类型列 1  | 数据类型列 2 |
|----------|---------|
| SMALLINT | INTEGER |
| INT      | BIGINT  |
| FLOAT    | REAL    |

表 5. 支持的数据类型 (续)

| 数据类型列 1                        | 数据类型列 2                   |
|--------------------------------|---------------------------|
| DOUBLE                         | DECIMAL                   |
| DEC                            | NUMERIC                   |
| NUM                            | CHARACTER                 |
| CHAR                           | VARCHAR                   |
| DATE                           | TIME                      |
| GRAPHIC                        | VARGRAPHIC                |
| CHARACTER VARYING              | TIMESTAMP                 |
| CHAR VARYING                   | CHARACTER FOR BIT DATA    |
| CHAR FOR BIT DATA              | VARCHAR FOR BIT DATA      |
| CHARACTER VARYING FOR BIT DATA | CHAR VARYING FOR BIT DATA |
| 用户定义的类型 (单值)                   |                           |

## 不支持的数据类型

分区表可以包含下列数据类型，但不支持将它们用作表分区键列:

- 用户定义的类型 (结构化)
- LONG VARCHAR
- LONG VARCHAR FOR BIT DATA
- BLOB
- BINARY LARGE OBJECT
- CLOB
- CHARACTER LARGE OBJECT
- DBCLOB
- LONG VARGRAPHIC
- REF
- C 变长字符串
- Pascal 变长字符串

在分区表中，不支持 XML 数据类型。

如果您选择使用 CREATE TABLE 语句的 EVERY 子句来自动生成数据分区，那么只能将一列用作表分区键。如果您选择通过在 CREATE TABLE 语句的 PARTITION BY 子句中指定每个范围来手动生成数据分区，那么可以将多个列用作表分区键，如以下示例所示:

```
CREATE TABLE sales (year INT, month INT)
PARTITION BY RANGE(year, month)
(STARTING FROM (2001, 1) ENDING (2001,3) IN tbsp1,
ENDING (2001,6) IN tbsp2, ENDING (2001,9)
IN tbsp3, ENDING (2001,12) IN tbsp4,
ENDING (2002,3) IN tbsp5, ENDING (2002,6)
IN tbsp6, ENDING (2002,9) IN tbsp7,
ENDING (2002,12) IN tbsp8)
```

这将生成 8 个数据分区，即 2001 年和 2002 年的每个季度有一个数据分区。

注:

1. 当将多个列用作表分区键时, 将把这些列视为组合键(类似于索引中的组合键), 其中, 后面的列依赖于前面的列。指定的每个起始值或结束值(所有列一起)不能超出 512 个字符。此限制与 SYSCAT.DATAPARTITIONS 目录视图中的 LOWVALUE 和 HIGHVALUE 列大小对应。如果指定超出 512 个字符的起始值或结束值, 就会导致错误 SQL0636N, 原因码为 9。
2. 表分区是多列的, 而不是多维的。在表分区中, 使用的所有列都包含在单个维中。

## 生成列

可以将生成列用作表分区键。此示例创建包含 12 个数据分区的表, 即每个月一个数据分区。对于任何年份, 一月份的所有行都将被放到第一个数据分区中, 二月份的行将被放到第二个数据分区中, 依此类推。

示例 1

```
CREATE TABLE monthly_sales (sales_date date,
sales_month int GENERATED ALWAYS AS (month(sales_date)))
PARTITION BY RANGE (sales_month)
(STARTING FROM 1 ENDING AT 12 EVERY 1);
```

注:

1. 对于表分区键中使用的生成列, 不能改变或删除其表达式。不允许对表分区键中使用的列添加生成列表表达式。对于表分区键中使用的列, 如果尝试添加、删除或改变该列的生成列表表达式, 就会导致错误 (SQL0270N, 原因码为 52)。
2. 如果生成列不是单调的, 或者优化器无法检测出该列是否是单调的, 就不会对范围谓词使用数据分区消除功能。如果存在非单调表达式, 那么只能对等价或 IN 谓词执行数据分区消除功能。有关单调性的详细讨论和示例, 请参阅第 43 页的『创建 MDC 表时的注意事项』。

## 分区表的装入注意事项

对目标表进行分区时, 将支持所有现有装入功能, 但存在以下常规限制:

- 当分区代理程序数大于 1 时, 不支持一致点。
- 不支持将数据装入到数据分区子集中的同时保持其余数据分区完全联机。
- 装入操作使用的异常表不能分区。
- 当 load 实用程序以插入方式或重新启动运行并且装入目标表具有任何已拆离的从属时, 那么不能重建唯一索引。
- 与装入 MDC 表相同, 装入分区表时将不会保留输入数据记录的精确排序。只有在单元或数据分区中才保留排序。
- 在每个数据库分区上利用多个格式化程序的装入操作仅保留输入记录的大致排序。在每个数据库分区上运行单个格式化程序, 将输入记录按单元或表分区键进行分组。要在每个数据库分区上运行单个格式化程序, 应显式请求 CPU\_PARALLELISM 为 1。

### 一般装入行为

load 实用程序将数据记录插入到正确的数据分区中。在装入之前, 不需要使用外部实用程序(如分割程序)来对输入数据进行分区。

load 实用程序不访问任何拆离的或相连的数据分区。数据仅插入到可视数据分区中。可视数据分区不会拆离，也不会相连。此外，装入替换操作不会截断拆离或相连的数据分区。因为 load 实用程序获取针对目录系统表的锁定，所以它将等待任何未落实的 ALTER TABLE 事务。这些事务将获取针对目录表中的相关行的互斥锁定，并且互斥锁定必须先终止，装入操作才能继续。这意味着装入操作运行期间，可能没有未落实的 ALTER TABLE ...ATTACH、DETACH 或 ADD PARTITION 事务。将拒绝目标为拆离或相连的数据分区的所有输入源记录，但可从异常表中检索它们（如果指定了异常表）。会有一条参考消息写入消息文件，以指示某些目标表数据分区处于相连或拆离状态。针对对应于目标表的相关目录表行的锁定使得用户无法通过在 load 实用程序运行时发出 ALTER TABLE ...ATTACH、DETACH 或 ADD PARTITION 操作来更改目标表的分区。

### 处理无效行

当 load 实用程序遇到的记录不属于任何可视数据分区时，将拒绝该记录并且 load 实用程序继续进行处理。因为范围限制违例而拒绝的记录个数不会显式显示出来，但会包括在拒绝的记录的总行数中。因为范围违例而拒绝记录不会增加行警告的数目。会有一条消息（SQL0327N）写入 load 实用程序消息文件，指示发现范围违例，不会对每一个记录来记录消息。除了目标表中的所有列之外，异常表还包括用于描述特定行发生的类型违例的列。包含无效数据的行（包括未分区的数据）将写至转储文件。

因为异常表插入成本很高，所以可以控制插入到异常表中的约束违例。例如，load 实用程序的缺省行为是将本来有效但因为范围限制或唯一约束违例而拒绝的行插入到异常表中。通过对 FOR EXCEPTION 子句分别指定 NORANGEEXC 或 NOUNIQUEEXC 可以关闭此行为。如果指定不应将这些约束违例插入到异常表中，或者未指定异常表，那么有关违反范围限制或唯一约束的行的信息将会丢失。

### 历史记录文件

如果目标表已分区，那么相应的历史记录文件条目不会包括目标表跨越的表空间列表。另一操作详细程度标识（“R”而不是“T”）指示对分区表运行了装入操作。

### 终止装入操作

终止装入替换操作将完全截断所有可视数据分区，而终止装入插入操作会将所有可视数据分区截断至装入前的长度。如果联机装入操作在装入复制阶段失败，那么在终止该操作期间，索引会变得无效。在终止接触索引的脱机装入操作时，索引也会变得无效（因为重建索引方式或者增量维护期间插入了键而使得索引处于不一致状态，从而导致索引无效）。将数据装入到多个目标中不会影响装入恢复操作，但将无法从装入阶段期间获取的一致点重新启动装入操作。在此情况下，如果对目标表进行分区，那么将忽略 SAVECOUNT 装入选项。此行为与将数据装入到 MDC 目标表中的行为一致。

### 生成列

如果生成列在任何分区、维或分布键中，那么会忽略 generatedoverride 文件类型修饰符并且 load 实用程序会生成值，就像指定了 generatedignore 文件类型修饰符一样。在此情况下，装入错误的生成列值可能导致将记录放置在错误的物理位置上，例如，错误的分区、MDC 块或数据库分区。例如，一旦记录放在错误的分区上，设置完整性就必须将其移至另一物理位置，这不能在联机设置完整性操作期间完成。

### 数据可访问性

当前联机装入算法扩展至分区表。在 LOAD 命令上指定的联机装入（ALLOW READ ACCESS）允许阅读器并行访问整个表，包括装入和非装入数据分区。

### 数据分区状态

在某些情况下，成功装入后可视数据分区可能切换至“设置完整性暂挂”状态和/或“只读访问”表状态。如果该表存在装入操作不能保留的约束，那么数据分区可能会置于这些状态。这种约束可能包括检查约束和拆离的具体化查询表。失败的装入操作会导致所有可视数据分区处于“装入暂挂”表状态。

### 错误隔离

不支持在数据分区级别进行错误隔离。隔离错误意味着在运行时未出现错误的数据库分区上继续装入，而在运行时出现错误的数据库分区上停止装入。可在不同数据库分区间隔离错误，但 load 实用程序不能在一个可视数据分区子集上落实事务，而回滚其余可视数据分区。

### 其他注意事项

- 如果有任何索引标记为无效，那么不支持递增。如果索引需要重建或拆离的从属项需要使用 SET INTEGRITY 语句进行验证，那么认为索引无效。
- 同时支持装入到分区表中，这些表使用按范围分区、按散列分布或按维算法组织的任何组合进行分区。
- 对于包括受装入影响的对象和表空间标识列表的日志记录，这些日志记录的大小（LOAD START 和 COMMIT (PENDING LIST)）可能增长得很快，并且因此而降低可供其他应用程序使用的活动日志空间量。
- 当表同时进行了分区和分布时，分区数据库装入可能不会影响所有数据库分区。只有输出数据库分区上的对象才会更改。
- 在装入操作期间，分区表的内存消耗会随表数的增加而增加。注意，总增加量不是线性的，因为仅总内存需求的一小部分与数据分区数成正比。

## 复制具体化查询表

具体化查询表是由查询定义的表，也用于确定表中的数据。具体化查询表可用来改进查询的性能。如果 DB2 数据库 Linux 版、UNIX 版和 Windows 版确定查询的一部分可使用具体化查询表来解决，那么数据库管理器可以重写该查询以使用具体化查询表。

在分区数据库环境中，可以复制具体化查询表并使用它们来改善查询性能。复制具体化查询表基于这样一个表：可能已经在单一分区数据库分区组中创建该表，但是您想在另一个数据库分区组中的所有数据库分区之间复制该表。要创建复制具体化查询表，调用带 REPLICATED 关键字的 CREATE TABLE 语句。

通过使用复制具体化查询表，可将未典型并置的表并置。对于大事实表和小维表的连接，复制具体化查询表特别有用。要将所需的额外存储器以及必须更新每个副本所带来的影响降至最小，要复制的表应较小，且更新不频繁。

**注：**还应考虑复制那些不常更新的更大的表：复制的单个成本可通过并置获得的性能效益来抵消。

通过在定义复制表所用的 subselect 子句中指定适当的谓词，可以复制选择的列和/或选择的行。

## 数据库分区组中的表空间

通过将表空间放在多分区数据库分区组中，就将该表空间内的所有表划分或分区到该数据库分区组的每个数据库分区中。

由此该表空间被创建到了一个数据库分区组中。一旦位于某个数据库分区组中，该表空间就必须保留在该处；而不能更改至另一数据库分区组。CREATE TABLESPACE 语句用于将表空间与数据库分区组关联。

## 表分区和多维集群表

一个表可以同时是多维集群表和分区表。在同时是多维集群表和分区表的表中，可以同时使用表分区的范围分区规范和 MDC 键中使用列。与只单独使用多维集群或分区功能相比，这更有助于获取较详细的数据分区和块消除。在许多应用中将 MDC 键列指定为不同于对表进行分区的列很有用。应该注意的是，表分区是多列，而 MDC 是多维。

### 主流 DB2 V9.1 数据仓库的特征

下列建议主要针对对于 DB2 V9.1 来说是新产品的典型主流仓库。假定下列特征：

- 数据库在多台机器或多个 AIX 逻辑分区上运行。
- 使用数据库分区功能（DPF）（表是使用 DISTRIBUTE BY HASH 子句创建的）。
- 有 4 到 50 个数据分区。
- 考虑其 MDC 和表分区的表是主要事实表。
- 表包含 1 亿到 1000 亿行。
- 在各种时间范围装入新数据：每夜、每周和每月。
- 每日接受量为 1 万到 1 亿条记录。
- 数据量变化：最多的一个月是最少的月的 5 倍。同样，最大维数（生产线，区域）具有 5 倍大小范围。
- 获取 1 到 5 年的详细数据。
- 每月或每个季度转出到期数据。
- 表使用大范围的查询类型。但是，相对于 OLTP 工作负载来说，该工作负载通常是具有下列特征的分析查询：
  - 较大的结果集，最多有 2 百万行
  - 大多数或全部查询都命中视图，而不是基本表
- SQL 子句按范围（BETWEEN 子句）、列表中的项等选择数据。

### 主流 DB2 V9.1 数据仓库事实表的特征

一个典型仓库事实表可能采用以下设计：

- 在 Month 列中创建数据分区。
- 为转出的每个时间段（例如，1 个月和 3 个月）定义数据分区。
- 在 Day 和 1 到 4 个其他维的基础上创建 MDC 维。典型的维有：生产线和区域。
- 所有数据分区和 MDC 集群都分布在所有数据库分区中。

MDC 和表分区具有一些相同的好处。下表列示组织中的潜在需求并根据先前确定的特征确定建议的组织方案。



表 6. 将表分区与 MDC 表配合使用

| 问题  | 建议的方案    | 建议  |
|---|----------|---|
| 转出期间的数据可用性  | 表分区      | 使用 DETACH PARTITION 子句来转出大量数据，并且只出现最少中断。  |
| 查询性能  | 表分区和 MDC | MDC 最适合用来查询多个维。表分区通过数据分区消除提高性能。   |
| 最少重组  | MDC      | MDC 维护集群，从而减少进行重组的必要性。  |
| 在传统脱机窗口中转出一个<br>月或更长时间的数据                               | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。  |
| 在短时间脱机窗口（小于 1 分钟）<br>期间转出一个月或更长时间的数据                    | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。  |
| 转出一个月或更长时间的数据，<br>并同时在不损失任何服务的情况下使表<br>对于提交查询的企业用户完全可用。 | MDC      | MDC 只能解决一部分此需求。由于表处于脱机状态的时间段太短，表分区并不适合。   |
| 每天装入数据（联机或脱机）   | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。   |
| “连续”装入数据（联机）  | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。   |
| “传统 BI”查询的查询执行性能  | 表分区和 MDC | MDC 特别适合用来查询立方体/多个维。表分区通过分区消除提高性能。  |
| 通过消除进行重组的必要性或减少执行任务所产生的不良影响，<br>使重组所带来的不良影响降到最低。        | MDC      | MDC 维护集群，从而减少进行重组的必要性。如果使用 MDC，那么数据分区不提供递增好处。但是，如果不使用 MDC，那么范围分区通过在分区级别维护一些粗粒度集群，有助于减少重组的必要性。 |

示例 1:

考虑一个具有键列 YearAndMonth 和 Province 的表。一种合理的规划此表方法是按日期进行分区，每 2 个月添加一个数据分区。此外，还可以按 Province 进行组织，以任何两个月日期范围内的特定省份的所有行集群在一起，如第 28 页的图 6 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (Province);
```

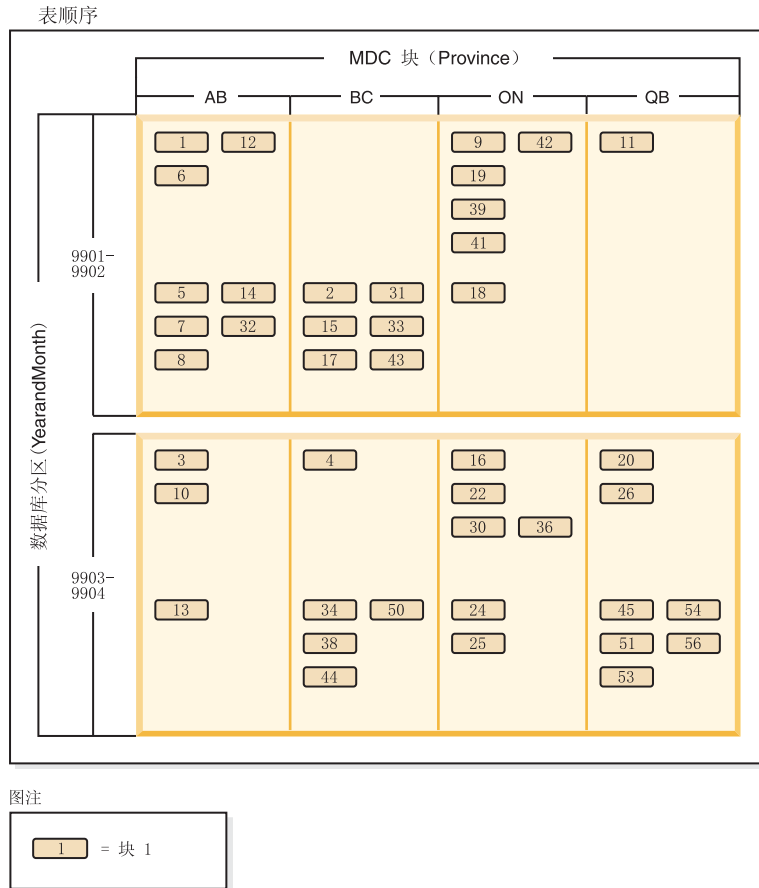


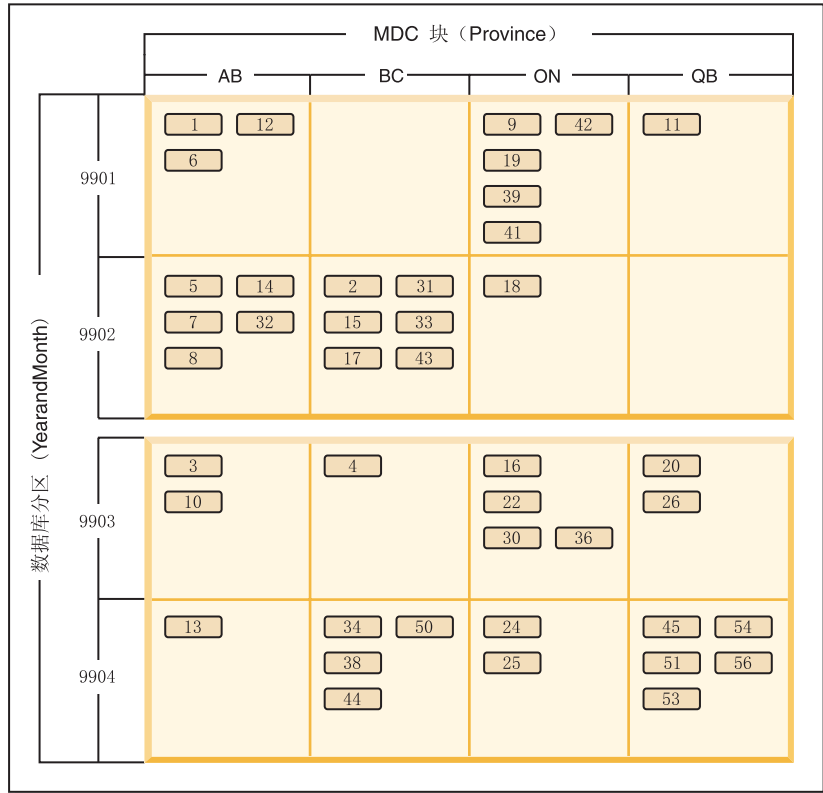
图 6. 按 YearAndMonth 分区并按 Province 组织的表

示例 2:

通过将 YearAndMonth 添加至 ORGANIZE BY 子句，可以获得较高的详细程度，如第 29 页的图 7 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (YearAndMonth, Province);
```

表顺序



图注

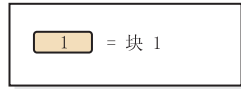


图 7. 按 YearAndMonth 分区并按 Province 和 YearAndMonth 组织的表

如果分区导致每个范围内只有单个值，那么通过在 MDC 键中包括表分区列不能获得任何好处。

### 注意事项

- 与基本表相比，MDC 表和分区表都需要一些存储器。这些存储器需求是附加的，但相对于所带来的好处来说，它们是合理的。
- 如果选择不在分区数据库环境中组合表分区和 MDC 功能，那么在您可以非常自信地预计数据分发情况时（通常也就是此处讨论的系统类型情况），使用表分区是最好的选择。否则，应考虑 MDC。



---

## 第 2 章 范围集群表

范围集群表 (RCT) 是一种表布局方案, 在此方案中, 表中的每一条记录都预先确定了记录标识 (RID), 该标识是用来在表中查找记录的内部标识。

对于用来保存数据的每个表, 应考虑最可能满足您需要的表类型。例如: 如果具有松散集群 (而不是单调增大) 的数据记录, 那么应考虑使用常规表和索引。如果有一些数据记录将在键中具有重复 (不是唯一的) 值, 那么不应使用范围集群表。如果不能在磁盘上为您想要的范围集群表预分配固定的存储量, 那么不应使用这种类型的表。这些因素可帮助您确定是否具有可用作范围集群表的数据。

---

### 范围集群表结构的优点

使用范围集群表有一些优点。

- 直接访问

通过范围集群表的键至 RID 映射功能来进行访问。

- 易于维护

不需要在每次 INSERT、UPDATE 或 DELETE 时都要更新辅助结构 (例如, B+ 树)。

- 日志记录量较少

与大小接近的常规表和相关联的 B+ 树索引比较起来, 范围集群表需要进行的日志记录量较少。

- 需要较少的缓冲池内存

不需要额外的内存用来存储辅助结构。

- 将 B+ 树形表的属性排序

记录的排序与通过 B+ 树形表获得的排序相同, 而不需要额外级别或者 B+ 树下一个键锁定方案。使用了 RCT 之后, 与常规 B+ 树形索引比较起来, 代码路径长度减少了。但是, 要获得这种优点, 必须使用 DISALLOW OVERFLOW 来创建范围集群表, 并且数据必须是密集的而不是稀疏的。

- 少一个索引

将每个键映射至磁盘上的一个位置, 这意味着可以使用少一个索引 (与必需的索引数比较起来) 来创建表。对于范围集群表, 访问表中数据的应用程序需求可能使得不需要另一个独立的索引。您仍然可以选择创建常规索引, 尤其是应用程序要求这种操作的话。

---

## 范围集群表的不兼容性

除了这些注意事项之外，还存在一些不兼容性，它们要么会限制可以使用范围集群表的位置，要么会限制不使用这些表的其他实用程序。

范围集群表的局限性包括：

- 不支持分区表上的范围集群表。

如果尝试创建具有范围集群的分区表，就会返回错误消息 `SQL0270 rc=87`。

- 声明的全局临时表（DGTT）不受支持。

不允许这些临时表使用范围集群属性。

- 自动摘要表（AST）不受支持。

不允许这些表使用范围集群属性。

- `load` 实用程序不受支持。

必须通过导入操作或并行插入应用程序来一次只插入一行。

- `REORG TABLE` 实用程序不受支持。

将不需要重组定义为 `DISALLOW OVERFLOW` 的范围集群表。定义为 `ALLOW OVERFLOW` 的这些范围集群表仍然不允许重组此溢出区域中的数据。

- 仅在一台逻辑机器上的范围集群表。

在具有“数据库分区功能部件”（DPF）的“企业服务器版”上，范围集群表不能存在于包含多个数据库分区的数据库中。

- 设计顾问程序将不会建议范围集群表。
- 根据定义，范围集群的表是已集群的。

这意味着下列集群方案与范围集群表不兼容：

- 多维集群（MDC）表
- 集群索引

- 值和缺省压缩不受支持。
- 不支持对范围集群表进行逆向扫描。
- 不支持 `IMPORT` 命令使用 `REPLACE` 选项。
- 不支持 `ALTER TABLE ... ACTIVATE NOT LOGGED INITIALLY` 语句使用 `WITH EMPTY TABLE` 选项。不支持 `ACTIVATE NOT LOGGED INITIALLY` 语句。

---

## 范围集群表和超出范围的记录键值

使用 `CREATE TABLE` 语句和 `ALLOW OVERFLOW` 选项来控制范围集群表（RCT）的行为，使它允许溢出记录。这样，可以确保立即分配表所需要的在定义范围之内的所有页。

一旦创建了记录，具有属于定义范围内的键的任何记录都将以相同方式工作，无论创建表时是否允许溢出选项。当一条具有键的记录位于定义范围之外时，情况将有所不同。在此情况下，当表允许溢出记录时，将把记录放置在动态分配的溢出区域中。当从定义范围之外添加更多记录时，将把这些记录放置在正在增加的溢出区域中。对涉

及到此溢出区域的表执行的操作将需要更长的处理时间，这是因为必须将溢出区域作为操作的一部分来访问。溢出区域越大，访问溢出区域所花的时间就会越长。在延长了使用溢出区域的时间之后，可考虑通过将表中的数据导出至使用新的扩展范围定义的新范围集群表来减小溢出区域大小。

可能有些时候您不想将记录放置在范围集群表中，以使记录键值属于允许的或定义的范围之外。要使这种类型的 RCT 存在，必须在 CREATE TABLE 语句中使用 DISALLOW OVERFLOW 选项。一旦创建了这种类型的 RCT，如果记录键值属于允许的或定义的范围之外，那么您必须接受错误消息。

---

## 范围集群表锁定

在正常的处理过程中，将锁定记录以确保在给定的任何时候都只有一个应用程序或用户有权访问记录或一组记录。对于范围集群表，不使用键锁定和下一个键锁定，而使用“离散锁定”。

此方法将锁定受到或者可能会受到由应用程序或用户请求的操作影响的所有记录。获得的锁定数取决于隔离级别。

范围集群表中当前是空的但是已经预分配的那些合格行就会被锁定。这样就避免了需要进行下一个键锁定。于是，密集的范围集群表需要更少锁定。





---

## 第 3 章 多维集群 (MDC) 表

---

### 多维集群表

多维集群 (MDC) 提供了一种极佳的方法来将多个表中的数据集群在多个维中, 它具有灵活、连续并且自动完成的特点。MDC 可显著提高查询性能。

此外, MDC 还可以极大地减少在插入、更新和删除操作期间执行数据维护 (例如, 重组) 和索引维护操作的开销。MDC 主要用于数据仓储和大型数据库环境, 但也可用于联机事务处理 (OLTP) 环境中。

---

### 常规表与 MDC 表的比较

常规表的索引是基于记录来建立的。索引的任何集群都仅限于单个维。在版本 8 之前, 数据库管理器仅支持通过集群索引对数据进行单维集群。通过使用集群索引, 在表中插入和更新记录时, 数据库管理器会尝试按索引的键顺序来维护各页上数据的物理顺序。

集群索引极大地提高了具有包含集群索引的一个或多个键的谓词的范围查询的性能。使用好的集群索引可以提高性能, 这是因为只需要访问表的一部分, 并且可以执行更有效地预取。

但是, 使用集群索引的数据集群有一些缺点。首先, 因为随着时间的推移将逐渐填满数据页上的空间, 因此, 将不能保证集群。插入操作将尝试把记录添加至具有相同或相似集群键值的那些页附近的页, 但是, 如果在理想位置没有空间, 那么将把它插入到表中的其他地方。因此, 可能需要定期进行表重组, 以便对表进行重新集群并设置具有附加可用空间的页以容纳将来的集群插入请求。

其次, 只能将一个索引指定为“集群”索引, 其他所有索引都将是非集群索引, 这是因为只能在一维中以物理方式对数据进行集群。这种局限性与集群索引是基于记录的这样一个事实有关, 因为所有索引都是在版本 8.1 之前建立的。

最后, 因为对于表中的每个记录, 基于记录的索引都包含一个指针, 所以它们可能非常大。

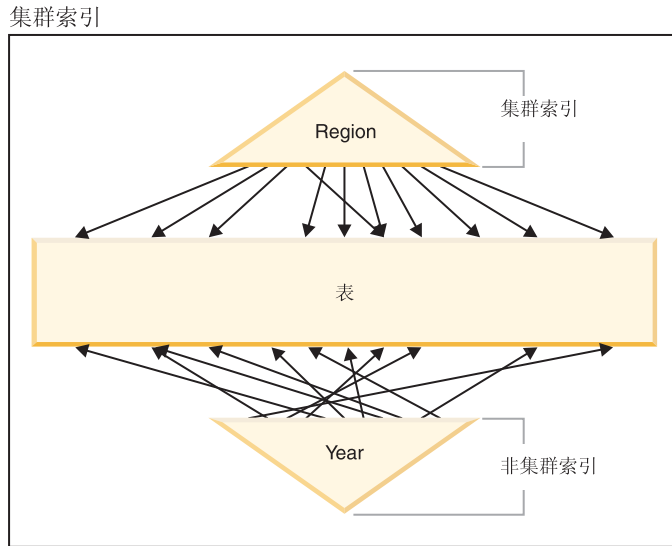


图 8. 具有集群索引的常规表

图 8 中的表上定义了两个基于记录的索引:

- “Region”上的集群索引
- “Year”上的另一个索引

“Region”索引是一个集群索引，它意味着当在索引中扫描键时，通常将在表中的同一页或邻近页上找到相应的记录。相反，“Year”索引是一个非集群索引，它意味着在该索引中扫描键时，很可能将在表的随机页上找到相应的记录。对集群索引进行扫描将获得更好的 I/O 性能，并且在该索引的数据集群程度越高时，顺序预取就会带来更多好处。

MDC 引入了基于块的索引。“块索引”指向记录块或记录组，而不是指向单个记录。通过从物理上根据集群值将 MDC 表中的数据组织成块，然后使用块索引来访问这些块，MDC 不仅能够解决集群索引的所有缺点，还能显著地改善性能。

首先，MDC 使表能够同时在多个键或维上进行物理集群。通过 MDC，多个维或集群键也具备了单维集群的优点。在对表的一个或多个指定维具有集群的情况下，查询性能会提高。这些查询不仅是只访问包含具有正确维值的记录的那些页，还会将这些满足要求的页组成块或扩展数据块。

其次，尽管具有集群索引的表经过一段时间之后可能会变成不是集群的，但是大多数情况下 MDC 表还是能够对所有维自动并持续地维护和保证它的集群。因此，无需频繁重组 MDC 表就可以复原数据的物理顺序。虽然始终将维护块中的记录顺序，但是在插入时（某些情况下甚至在初始装入时）不会维护块的物理排序（即，在块索引扫描时从一个块扫描到另一个块）。

最后，在 MDC 中，集群索引是基于块的。这些索引比常规的基于记录的索引要小很多，因此，占用的磁盘空间更少，并且扫描时速度会更快。

---

## 选择 MDC 表维

一旦您决定使用多维集群表，您选择的维将不仅取决于将使用表并且受益于块级别集群的查询类型，更重要的是取决于实际数据的数量和分布情况。使用相关的概念链接可以看到设计 MDC 表的各个方面以及有关选择适当的维和块大小的一些指导。

### 将受益于 MDC 的查询

选择表的集群维时首先应注意确定哪些查询将受益于块级别的集群。通常，当根据组成将对数据完成的工作的查询来选择维时有一些候选维。这些候选维的等级是很重要的。在等同查询或范围谓词查询中涉及到的列（特别是具有较小基数的列）将最受益于集群维，因此应将它们看作集群维的候选维。您还要考虑为涉及与维表进行星型连接的 MDC 事实表中的外键创建维。记住，对多个维进行自动持续的集群以及扩展数据块或块级别的集群会提高性能。

有许多种查询都可以利用多维集群。以下是这样一些查询的示例。在这些示例的某些示例中，假定存在一个 MDC 表 t1，具有 c1、c2 和 c3 维。在其他示例中，假定存在一个 MDC 表 mdctable，具有 color 和 nation 维。

示例 1:

```
SELECT .... FROM t1 WHERE c3 < 5000
```

此查询涉及到单个维的范围谓词，因此可以在内部重写以使用 c3 的维块索引来访问该表。将扫描索引以查找键值小于 5000 的块标识（BID），并且对块的结果集应用最小关系扫描以检索实际记录。

示例 2:

```
SELECT .... FROM t1 WHERE c2 IN (1,2037)
```

此查询涉及到单个维的 IN 谓词，并且可以触发基于块索引的扫描。可以内部重写此查询以使用 c2 的维块索引来访问该表。将扫描索引以查找具有键值 1 和 2037 的 BID，并且对块的结果集应用最小关系扫描以检索实际记录。

示例 3:

```
SELECT * FROM MDCTABLE WHERE COLOR='BLUE' AND NATION='USA'
```

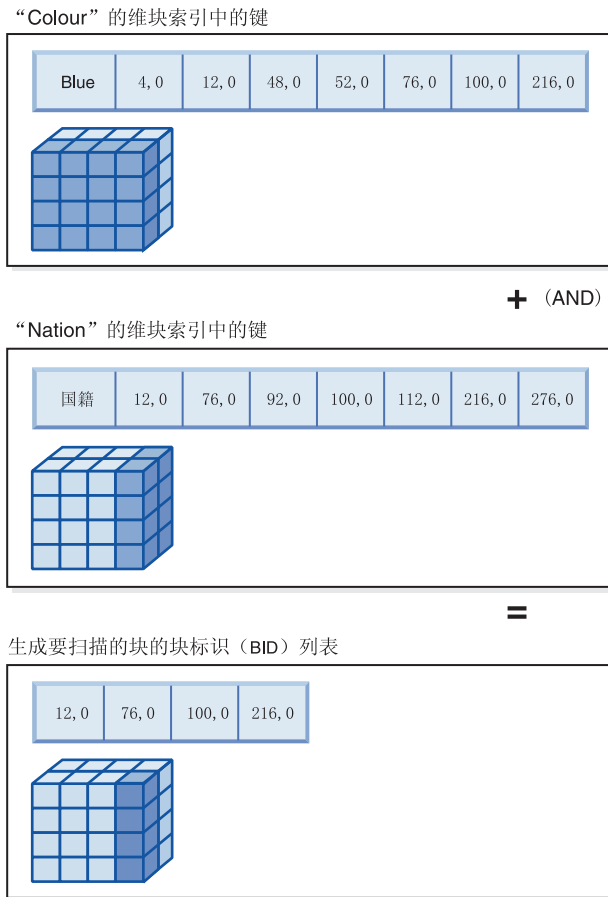


图 9. 对两个块索引使用逻辑 AND 操作的查询请求。

要执行此查询请求，完成下列步骤（显示在图 9 中）：

- 完成维块索引查找：对 Blue 片执行一次，对 USA 片执行一次。
- 执行块逻辑 AND 操作以确定两个片的交集。即，逻辑 AND 操作只确定在这两个片中都可以找到的那些块。
- 对表中获得的块执行最小关系扫描。

示例 4:

```
SELECT ... FROM t1
  WHERE c2 > 100 AND c1 = '16/03/1999' AND c3 > 1000 AND c3 < 5000
```

此查询涉及到 c2 和 c3 的范围谓词和 c1 的等价谓词，并且还要执行逻辑 AND 操作。可以内部重写此查询以访问每个维块索引上的表：

- 扫描 c2 块索引以查找具有大于 100 的键值的 BID
- 扫描 c3 块索引以查找键值在 1000 到 5000 之间的 BID
- 扫描 c1 块索引以查找键值为“16/03/1999”的 BID。

然后，对从每个块扫描获得的 BID 执行逻辑 AND 操作以查找它们的交集，并对块的结果集应用最小关系扫描以查找实际记录。

示例 5:

```
SELECT * FROM MDCTABLE WHERE COLOR='BLUE' OR NATION='USA'
```

要执行此查询请求，完成下列步骤:

- 完成维块索引查找: 对每个片都执行一次。
- 执行逻辑 OR 操作以查找两个片的并集。
- 对表中获得的块执行最小关系扫描。

示例 6:

```
SELECT .... FROM t1 WHERE c1 < 5000 OR c2 IN (1,2,3)
```

此查询涉及到 c1 维的范围谓词和 c2 维的 IN 谓词，以及逻辑 OR 操作。可以内部重写此查询以访问维块索引 c1 和 c2 上的表。扫描 c1 维块索引以查找小于 5000 的值，并对 c2 维块索引执行另一个扫描以查找值 1、2 和 3。对从每个块索引扫描获得的 RID 执行逻辑 OR 操作，然后对块的结果集应用最小关系扫描以查找实际记录。

示例 7:

```
SELECT .... FROM t1 WHERE c1 = 15 AND c4 < 12
```

此查询涉及到 c1 维的等价谓词和不是维的一列的另一个范围谓词以及逻辑 AND 操作。可以内部重写该查询以访问 c1 的维块索引，以获取 c1 值为 15 的表的片的块列表。如果 c4 具有 RID 索引，那么可以进行索引扫描以检索 c4 小于 12 的记录的 RID，然后可以对获得的块列表和此记录列表执行逻辑 AND 操作。此交集将除去 c1 为 15 的块中不存在的 RID，而仅从表中检索合格块中存在的列示的 RID。

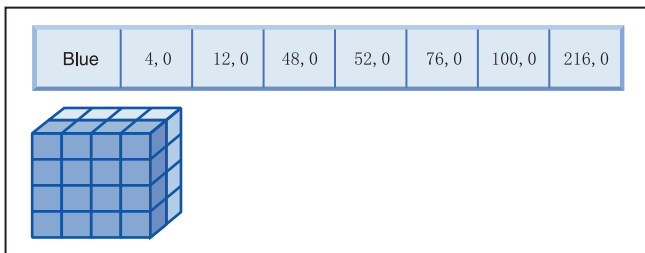
如果 c4 没有 RID 索引，那么可以扫描块索引以查找合格块的列表，并且在每个块的最小关系扫描期间，可以对发现的每个记录应用谓词 c4 < 12。

示例 8:

假定这样一个方案，存在 color、year 和 nation 维以及部件号的行标识 (RID) 索引，那么可以进行以下查询。

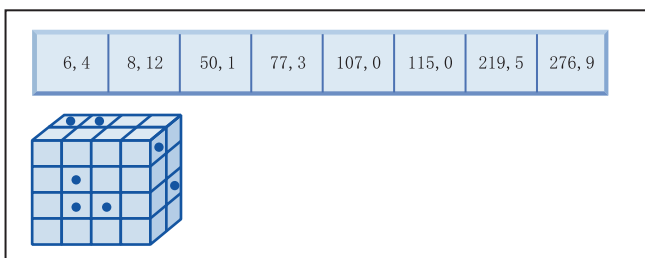
```
SELECT * FROM MDCTABLE WHERE COLOR='BLUE' AND PARTNO < 1000
```

“Colour” 的维块索引中的键



+ (AND)

Partno 的 RID 索引中的行标识 (RID)



=

生成用访存的行标识

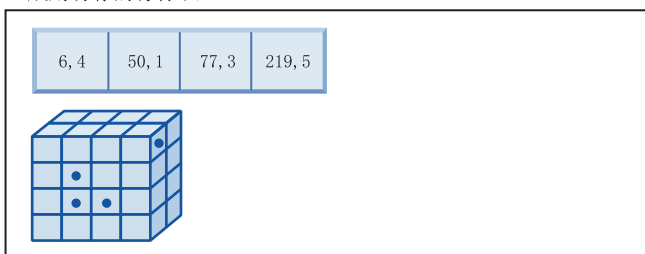


图 10. 对块索引和行标识 (RID) 索引使用逻辑 AND 操作的查询请求

要执行此查询请求，完成下列步骤（显示在图 10 中）：

- 完成维块索引查找和 RID 索引查找。
- 对块和 RID 使用逻辑 AND 操作来确定片与满足谓词条件的那些行的交集。
- 获得的结果仅为还属于满足条件的块的那些 RID。

示例 9:

```
SELECT * FROM MDCTABLE WHERE COLOR='BLUE' OR PARTNO < 1000
```

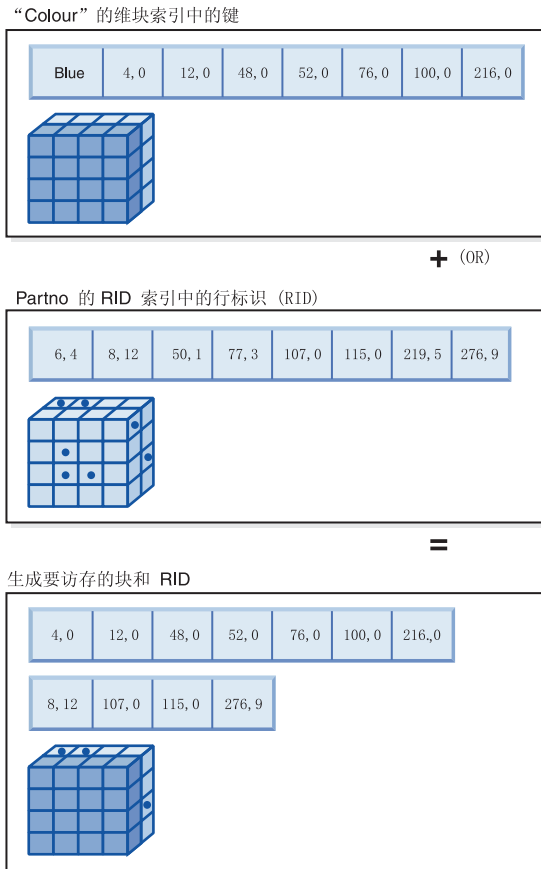


图 11. 使用逻辑 OR 操作的块索引和行标识的工作方式

要执行此查询请求，完成下列步骤（显示在图 11 中）：

- 完成维块索引查找和 RID 索引查找。
- 对块和 RID 使用逻辑 OR 操作来确定片与满足谓词条件的那些行的并集。
- 获得的结果是满足条件的块中的所有行以及满足谓词条件但在满足条件的块外部的其他 RID。对每个块执行最小关系扫描以检索它们的记录并逐个检索这些块之外的其他记录。

示例 10:

```
SELECT ... FROM t1 WHERE c1 < 5 OR c4 = 100
```

此查询涉及到 c1 维的范围谓词和非维列 c4 的等价谓词，以及逻辑 OR 操作。如果 c4 列上具有 RID 索引，那么可以内部重写此查询以使用 c1 上的维块索引和 c4 上的 RID 索引来执行 OR 操作。如果 c4 上没有索引，那么可以选择表扫描，因为必须检查所有记录。逻辑 OR 操作将对 c1 使用块索引扫描以找到小于 4 的值，以及对 c4 使用 RID 索引扫描以找到值 100。对每个满足条件的块执行最小关系扫描，因为这些块中的所有记录都满足条件，还要检索这些块外部的记录的任何其他 RID。

示例 11:

```

SELECT .... FROM t1,d1,d2,d3
WHERE t1.c1 = d1.c1 and d1.region = 'NY'
      AND t2.c2 = d2.c3 and d2.year='1994'
      AND t3.c3 = d3.c3 and d3.product='basketball'

```

此查询涉及星型连接。在此示例中，t1 是事实表并且它具有外键 c1、c2 和 c3（与主键 d1、d2 和 d3 相对应）以及维表。维表不一定是 MDC 表。Region、year 和 product 是可以使用常规或块索引建立索引的各自维表的列（如果维表是 MDC 表的话）。当根据 c1、c2 和 c3 值访问事实表时，可以对这些列的维块索引进行块索引扫描，然后使用获得的 BID 来执行逻辑 AND 操作。当存在块列表时，可以对每个块进行最小关系扫描以获取记录。

## 单元的密度

选择适当的维和扩展数据块大小对于 MDC 设计来说**非常重要**。这些因素将确定表的期望单元密度。它们之所以重要是因为对每个现有单元都分配扩展数据块，无论该单元中的记录数是多少。正确的选择将利用基于块的索引和多维集群，从而获得更好的性能。目标是获得填充密度很高的块，以便从多维集群中获得最多益处并且最佳地利用空间。

因此，设计多维表时的非常重要的注意事项就是表中的期望单元密度（基于现有的和预期的数据）。可以根据查询性能选择一组维，它导致表中的潜在单元数很大（基于每个维的可能值的数目）。表中的可能单元的数目等于每个维的基数的笛卡尔乘积。例如，如果根据 Day、Region 和 Product 维来集群表且涵盖五年的数据，那么表中可能具有  $1821 \text{ days} * 12 \text{ regions} * 5 \text{ products} = 109260$  个不同的单元。任何仅包含几个记录的单元仍需要分配给它的整块的页，以存储该单元的记录。如果块大小较大，那么此表结果可能会比它实际需要的大得多。

要获得最佳单元密度，应考虑以下几种设计因素：

- 维数不同。
- 一个或多个维的详细程度不同。
- 表空间的块（扩展数据块）大小和页大小不同。

执行下列步骤来获得最佳设计：

### 1. 标识候选维。

确定哪些查询将采用块级别集群。检查具有下列某些特征或所有特征的列可能具有的工作负载：

- 任何 IN 列表谓词的范围和等效性
- 数据的转入或转出
- Group-by 和 order-by 子句
- Join 子句（特别是在星型模式环境中）。

### 2. 估计单元数目。

标识在根据一组候选维组织的表中可能存在多少个潜在的单元。确定数据中出现的维值的唯一组合的数目。如果表存在，那么只需要选择每个列中将为表的维的单值的数目来确定当前数据的准确数目。或者，如果只有表的统计信息，那么可以通过乘以候选维的列基数来确定一个大概数目。



**注：**如果表处于分区数据库环境中，并且分布键与考虑的任何维都无关，那么必须通过用所有数据除以数据库分区数来确定每个单元的平均数据量。

### 3. 估计空间占用率或密度。

按平均情况来看，认为每个单元都有一个部分填充的块，该块中只存储了很少的行。当每个单元的行数变得越来越少时，部分填充的块就会更多。另外还要注意，按平均情况来看（假定存在很少甚至没有数据偏差），可以通过用表中的记录数除以单元数来计算每个单元的记录数。但是，如果表处于分区数据库环境中，那么需要考虑每个数据库分区上的每个单元中有多少条记录，因为是按照数据库分区来为数据分配块的。当估计分区数据库环境中的空间占用率和密度时，需要考虑每个数据库分区（而不是整个表）上每个单元的平均记录数。有关更多信息，请参阅『多维集群（MDC）表的创建、布置和使用』一节。

可以采用以下几种方法来提高密度：

- 减小块大小，以便使部分填充的块占用较少的空间。

通过适当减小扩展数据块大小来减小每个块的大小。具有部分填充的块的每个单元或者只包含具有很少记录的一个块的每个单元都占用较少的空间。但是，采用一种折衷办法就是，对于具有很多记录的那些单元，需要更多的块来包含它们。这将增加块索引中的这些单元的块标识（**BID**）的数目，使这些索引更大并且潜在地导致更多的索引插入和删除操作，因为会更快地清空和填充块。与数目较少的集群数据的较大组合相比较，它还会导致表中的集群数据的更多更小的组合，以装入更多填充单元值。

- 通过减少维数或者通过增大具有生成列的单元的详细程度来减少单元数。

可以将一个或多个维上滚为较低の詳細程度，以给予它较低的基数。例如，可以仍然根据 **Region** 和 **Product** 对将前一个示例中的数据库进行集群，但将 **Day** 维替换为 **YearAndMonth** 维。这将为 **YearAndMonth**、**Region** 和 **Product** 维给予基数 60（12 个月乘以 5 年）、12 和 5，可能的单元数为 3600。于是，每个单元将具有更大范围的值，并且降低了只包含很少记录的可能性。

还应该考虑对涉及到的列通常使用的谓词，例如，使用得较多的是 **Month of Date**、**Quarter** 或 **Day**。这将影响更改维的詳細程度的期望度。例如，如果大多数谓词是针对特定日期的，并且您根据“月份”对表进行了集群，那么 **DB2** 数据库 **Linux** 版、**UNIX** 版和 **Windows** 版可以使用 **YearAndMonth** 的块索引来快速缩小包含想要的日期的月份的范围并且仅访问这些相关联的块。但是，在扫描块时，必须应用 **Day** 谓词来确定哪些日期合格。但是，如果根据 **Day** 进行集群（并且 **Day** 具有很高的基数），那么可以使用 **Day** 的块索引来确定扫描哪些块，并且只需要对合格的每个单元的第一条记录重新应用 **Day** 谓词。在此情况下，在使用用户定义的函数上滚 **Region** 列（它包含 12 个不同的值）至 **Regions West**、**North**、**South** 和 **East** 时最好考虑上滚其他维之一以增大单元的密度。

---

## 创建 MDC 表时的注意事项

当创建 MDC 表时应该考虑许多因素。下列各节讨论了您当前所处的数据库环境（例如，是否具有分区数据库）和为 MDC 表选择的维对于您创建、布置和使用 MDC 表将产生的影响。还讨论了“**DB2 设计顾问程序**”以及可以如何使用它来提供有关这样一些问题的建议。

## 将数据从现有表移至 MDC 表

要提高数据仓库或大型数据库环境中的查询性能和减少数据维护操作的开销，可以将数据从常规表移至多维集群（MDC）表。要将数据从现有表移至 MDC 表：导出数据，删除原始表（可选），创建多维集群（MDC）表（使用带有 `ORGANIZE BY DIMENSIONS` 子句的 `CREATE TABLE` 语句），并将您自己的数据装入 MDC 表。

可以使用称为 `SYSPROC.ALTOBJ` 的 `ALTER TABLE` 过程来将现有表中的数据转换为 MDC 表中的数据。可以从“DB2 设计顾问程序”中调用此过程。在这两个表之间转换数据所需要的时间可能很长，这取决于表的大小以及需要转换的数据量。

当改变表时，`ALTOBJ` 过程将执行下列操作：

- 删除表的所有从属对象
- 重命名表
- 使用新定义来创建表
- 重新创建表的所有从属对象
- 将表中的现有数据变换为新表中需要的数据。也就是说，从旧表中选择数据，然后将该数据装入新表中，可以使用列函数来将旧的数据类型变换为新的数据类型。

## SMS 表空间中的 MDC 表

如果打算将 MDC 表存储在 SMS 表空间中，那么需要使用多页文件分配。（对于在版本 8.2 和更新版本中新创建的数据库，多页文件分配是缺省值。）这是因为 MDC 表总是按整个扩展数据块来扩展的，并且这些扩展数据块中的所有页在物理上保持连续很重要。因此，禁用多页文件分配不利于利用空间；而且，启用多页文件分配将显著提高每个扩展数据块中的页在物理上连续的可能性。

## DB2 设计顾问程序上的 MDC 顾问程序功能部件

“DB2 设计顾问程序”（`db2advis`）（以前称为“索引顾问程序”）具有 MDC 功能部件。此功能部件建议用于 MDC 表中的集群维（包括基本列的粗糙度）以便提高工作负载性能。粗糙度这个术语表示用来减小集群维的基数（单值的数目）的一个数学表达式。粗糙度的常见示例是日期，日期的粗糙度可为日期、日期所在的星期、日期所在的月份或一年中的季度。

使用“DB2 设计顾问程序”的 MDC 功能部件要求数据库中至少存在几个扩展数据块的数据。“DB2 设计顾问程序”使用数据来对数据密度和基数建立模型。

如果数据库的表中没有数据，那么“DB2 设计顾问程序”不会建议使用 MDC，即使该数据库包含空表，但有一组虚假的统计信息来表示它是一个已填充的数据库。

建议还标识了用来定义维的粗糙度的潜在生成列。建议中不包括可能的块大小。在为 MDC 表提供建议时，使用表空间的扩展数据块大小。假定将在现有表所在的表空间中创建建议的 MDC 表，因此该表将具有相同的扩展数据块大小。对 MDC 维的建议将根据表空间的扩展数据块大小变化而变化，这是因为扩展数据块大小将影响可以填充到块或单元中的记录数。这将直接影响单元的密度。

只考虑单列维而不考虑组合列维，尽管可以为表建议单个维或多个维。MDC 功能部件将建议大多数受支持的数据类型使用的粗糙度，目标是减小所采用的 MDC 解决方案中

的单元的基数。异常的数据类型包括: CHAR、VARCHAR、GRAPHIC 和 VARGRAPH 数据类型。所有受支持的数据类型都将被强制类型转换为 INTEGER 并通过生成的表达式来设置粗糙度。

“DB2 设计顾问程序”的 MDC 功能部件的目标是选择可提高性能的 MDC 解决方案。另一个目标是将数据库的存储器扩充限制在适当的级别。使用统计方法来确定每个表的最大存储器扩充。

顾问程序中的分析操作既会利用块索引访问的优点, 也会受到 MDC 对表的维执行插入、更新和删除操作的影响。对表执行这些操作时可能会导致在各个单元之间移动记录。分析操作还会模拟对特定 MDC 维上的数据进行组织时产生的任何表扩充而对性能造成的潜在影响。

通过对 db2advis 实用程序使用 -m <advise type> 标志来启用 MDC 功能部件。使用“C”建议类型来指示多维集群表。建议类型为: “I”表示索引、“M”表示具体化查询表、“C”表示 MDC, 而“P”表示分区数据库环境。建议类型可以相互组合使用。

注: “DB2 设计顾问程序”将不会处理小于 12 个扩展数据块大小的表。

当提出建议时, 顾问程序将同时分析 MQT 和常规基本表。

MDC 功能部件的输出包括:

- 每个表的生成列表表达式 (用于对 MDC 解决方案中出现的维设置粗糙度)。
- 为每个表建议的 ORGANIZE BY 子句。

对标准输出和作为说明工具的一部分的 ADVISE 表报告了建议。

## MDC 表和分区数据库环境

多维集群可以与分区数据库环境联合使用。实际上, MDC 可以作为分区数据库环境的补充。分区数据库环境用来将一个表中的数据分配到多个物理节点或逻辑节点上, 以便达到下列目的:

- 利用多台机器来并行增加处理请求。
- 增大表的物理大小 (超过单个数据库分区的限制)。
- 提高数据库的可伸缩性。

分发表的原因与该表是 MDC 表还是常规表无关。例如, 选择用来组成分布键的列的规则是相同的。MDC 表的分布键可以包括任何列, 无论这些列是否组成表的维的一部分。

如果分布键与表的某个维完全相同, 那么每个数据库分区都将包含该表的不同部分。例如, 如果作为示例的 MDC 表按颜色分发在两个数据库分区上, 那么将使用 Color 列来划分数据。因此, 可能在一个数据库分区上找到 Red 和 Blue 片, 而在另一个数据库分区上找到 Yellow 片。如果分布键与表中的维不完全相同, 那么每个数据库分区都将具有每个片的数据的子集。当选择维和估计单元占用率时 (请参阅『单元密度』一节), 注意, 平均来说, 每个单元的数据总量是通过用所有数据除以数据库分区数来确定的。

## 具有多个维的 MDC 表

如果知道将在查询中大量使用特定谓词, 那么可以使用 ORGANIZE BY DIMENSIONS 子句根据涉及到的列对表进行集群。

示例 1:

```
CREATE TABLE T1 (c1 DATE, c2 INT, c3 INT, c4 DOUBLE)
  ORGANIZE BY DIMENSIONS (c1, c3, c4)
```

示例 1 中的表根据形成逻辑立方体（即，具有三个维）的三个本地列中的值集群。现在可以在查询处理期间根据一个或多个维对表进行逻辑分片，以便涉及的关系运算符仅处理相应的片或单元中的块。注意，块的大小（页数）将是表的扩展数据块大小。

## 具有基于多列的维的 MDC 表

每个维可以由一列或多列组成。作为一个示例，可以创建一个根据包含两列的一个维来集群的表。

示例 2:

```
CREATE TABLE T1 (c1 DATE, c2 INT, c3 INT, c4 DOUBLE)
  ORGANIZE BY DIMENSIONS (c1, (c3, c4))
```

在示例 2 中，表将根据两个维 c1 和 (c3, c4) 来进行集群。这样，在查询处理期间，表可以根据 c1 维或组合 (c3, c4) 维逻辑分片。该表将与示例 1 中的表具有相同数目的块，但是少一个维块索引。在示例 1 中，将有三个维块索引，列 c1、c3 和 c4 各一个。在示例 2 中，将有两个维块索引，一个针对列 c1，而另一个针对 c3 和 c4。这两个方法的主要不同之处在于，在示例 1 中，仅涉及 c4 的查询可以使用 c4 的维块索引来快速直接地访问相关数据块。在示例 2 中，c4 是维块索引中的辅助键部分，因此仅涉及 c4 的查询涉及更多的处理。但是，在示例 2 中，将少维护和存储一个块索引。

“DB2 设计顾问程序”将不对包含多列的维提供建议。

## 将列表式作为维的 MDC 表

列表式也可用于集群维。根据列表式集群的功能对于将维上滚至更低的详细程度非常有用，例如，将地址上滚为地理位置或区域，或者将日期上滚为星期、月份或年。要以此方式实现维的上滚，可以使用生成列。此类型的列定义将允许使用可表示维的表达式创建列。在示例 3 中，该语句创建根据一个基本列和两个列表式进行集群的表。

示例 3:

```
CREATE TABLE T1(c1 DATE, c2 INT, c3 INT, c4 DOUBLE,
  c5 DOUBLE GENERATED ALWAYS AS (c3 + c4),
  c6 INT GENERATED ALWAYS AS (MONTH(C1)))
  ORGANIZE BY DIMENSIONS (c2, c5, c6)
```

在示例 3 中，c5 列是基于 c3 和 c4 列的表达式，而 c6 列会将 c1 列上滚至更低的详细程度。此语句将根据 c2、c5 和 c6 列中的值集群该表。

## 对生成列维的范围查询

对生成列维的范围查询需要单调列函数。表达式必须是单调的才能为生成列的维派生范围谓词。如果对生成列创建维，那么对基本列的查询将能够利用生成列的块索引来提高性能（有一种情况例外）。要使基本列（例如，日期）的范围查询对维块索引使用范围扫描，用来在 CREATE TABLE 语句中生成列的表达式必须是单调的。尽管列表式

式可以包括任何有效表达式（包括用户定义的函数（UDF）），但是如果表达式不是单调的，那么当等价谓词或 IN 谓词都在基本列上时，它们才能够使用块索引来满足查询。

作为一个示例，假定使用生成列 month 的维来创建 MDC 表，其中  $month = \text{INTEGER}(\text{date})/100$ 。对于该维（month）的查询，可以执行块索引扫描。对于基本列（date）的查询，也可以执行块索引扫描来缩小要扫描的块的范围，然后只将日期的谓词应用于这些块中的行。

编译器将生成要在块索引扫描中使用的其他谓词。例如，对于以下查询：

```
SELECT * FROM MDCTABLE WHERE DATE > "1999/03/03" AND DATE < "2000/01/15"
```

编译器将生成以下谓词：“month >= 199903”和“month < 200001”，它们可以用作维块索引扫描的谓词。当对获得的块进行扫描时，会将原始谓词应用于这些块中的行。

非单调表达式将只允许对该维应用等价谓词。非单调函数的一个较好的示例是 MONTH()，如示例 3 中的 c6 列的定义所示。如果 c1 列是日期、时间戳或日期或时间戳的有效字符串表示法，那么函数将返回范围是 1 到 12 的整数值。尽管函数的输出是确定的，但是实际上它生成的输出与阶跃函数（即，循环模式）相似：

```
MONTH(date('99/01/05')) = 1
MONTH(date('99/02/08')) = 2
MONTH(date('99/03/24')) = 3
MONTH(date('99/04/30')) = 4
...
MONTH(date('99/12/09')) = 12
MONTH(date('00/01/18')) = 1
MONTH(date('00/02/24')) = 2
...
```

尽管此示例中的日期是连续增加的，但是 MONTH(date) 不会增加。更具体地说，每当 date1 大于 date2，并不能保证 MONTH(date1) 大于或等于 MONTH(date2)。这是单调性所要求的。此非单调性是允许的，但是它限制了维，基本列的范围谓词不能生成维的范围谓词。但是，表达式的范围谓词是可以的，例如，where month(c1) between 4 and 6。这可以采用常规方式使用维的索引，起始键为 4 而停止键为 6。

要使此函数单调，必须将年包括为月份的高位部分。版本 9.2 提供对 INTEGER 内置函数的扩展以帮助根据日期定义单调表达式。INTEGER(date) 返回日期的整数表示法，可以分开查找年和月份的整数表示法。例如，INTEGER(date('2000/05/24')) 返回 20000524，因此  $\text{INTEGER}(\text{date}('2000/05/24'))/100 = 200005$ 。函数 INTEGER(date)/100 是单调的。

相似的，内置函数 DECIMAL 和 BIGINT 也具有扩展，所以可以派生单调函数。DECIMAL(timestamp) 返回时间戳的十进制表示法，可以在单调表达式中使用它来派生月份、天、小时或分钟等等的增加的值。BIGINT(date) 返回日期的大整数表示法，类似于 INTEGER(date)。

只要可能，数据库管理器将在为表创建生成列或者根据维子句中的表达式创建维时确定表达式的单调性。特定函数被识别为保留单调性，例如，DATENUM()、DAYS() 和 YEAR()。并且，列和常量的各种算术表达式，例如，除法、乘法或加法是保留单调性的。当 DB2 确定表达式不保留单调性时，或者如果它不能确定这一点，该维将仅支持对其基本列使用等价谓词。

## MDC 表的装入注意事项

如果定期将数据装入到数据仓库中，那么使用 MDC 表会很有帮助。在 MDC 表中，装入会先复用表中先前清空的块，然后才扩展该表并添加新块以便装入余下的数据。

在删除了一组数据之后（例如，一个月的数据），可以使用 Load 实用程序来装入下一个月的数据，并且它可以复用在（已落实）删除之后已经清空的块。在版本 9.5 中，MDC 转出功能是随延迟清除一起添加的。在落实转出（它也是一项删除操作）之后，各个块并未被释放，尚不能复用这些块。将调用后台进程以维护基于记录标识（RID）的索引。完成维护之后，就会释放各个块，从而可以复用这些块。

当将数据装入 MDC 表中时，可以对输入数据排序，也可以不对它进行排序。如果不进行排序，那么考虑执行下列操作：

- 增大 *util\_heap* 配置参数。

增大实用程序堆大小将影响数据库中的所有装入操作（以及备份和复原操作）。

- 增大使用 LOAD 命令的 DATA BUFFER 子句给定的值。

增大此值将影响单个装入请求。实用程序堆大小必须足够大，以满足多个并发装入请求的可能性。

- 确保用于缓冲池的页大小与临时表空间的最大页大小相同。

装入从块边界开始，所以最好将它用于属于新单元的数据或者用于表的初始填充。

## MDC 表的记录注意事项

如果 RID 索引先前索引或另外索引的列现在是维并且因此由块索引来索引，那么索引维护和日志记录将显著地减少。

仅当删除了整个块中的最后一条记录时，数据库管理器才需要从块索引中除去 BID 并记录此索引操作。同样，仅当将记录插入到新块中（如果是逻辑单元的第一条记录或者是插入到当前已满的块的逻辑单元中）时，数据库管理器才需要将 BID 插入到块索引中并记录该操作。因为块可以具有 2 到 256 页记录，所以此块索引维护和日志记录相对较小。仍将记录表和 RID 索引的插入和删除操作。对于转出删除，不会记录被删除的记录。而是会通过重新格式化页的某些部分来使包含这些记录的页看上去是空的。将会记录对重新格式化的部分所做的更改，但不会记录这些记录本身。

## MDC 表的块索引注意事项

当定义 MDC 表的维时，将创建维块索引。此外，当定义了多个维时，还可能会创建组合块索引。但是，如果为 MDC 表只定义了一个维，那么 DB2 将只创建一个块索引，它将同时充当维块索引和组合块索引。

相似的，如果创建的 MDC 表具有针对列 A 和针对（列 A 和列 B）的维，DB2 将对列 A 创建维块索引，并对列 A 和列 B 创建维块索引。因为组合块索引是表中的所有维的块索引，所以列 A 和列 B 的维块索引也将充当组合块索引。

组合块索引也用于查询处理，以访问在表中具有特定维值的数据。注意，组合块索引中的键部分的顺序可能会影响它对于查询处理的使用或适用性。它的键部分的顺序由创建 MDC 表时使用的整个 ORGANIZE BY DIMENSIONS 子句中的列的顺序确定。例如，如果表是使用下列语句创建的：

```
CREATE TABLE t1 (c1 int, c2 int, c3 int, c4 int)
  ORGANIZE BY DIMENSIONS (c1, c4, (c3,c1), c2)
```

那么将对 (c4,c3,c1,c2) 列创建组合块索引。注意，尽管 c1 在维子句中指定了两次，但是它仅在作为组合块索引的键部分时使用了一次，并且是以第一次发现它的顺序使用的。组合块索引中的键部分的顺序对于插入处理没有影响，但对于查询处理可能是有影响的。因此，如果更希望组合块索引具有列顺序 (c1,c2,c3,c4)，那么应使用以下语句来创建表：

```
CREATE TABLE t1 (c1 int, c2 int, c3 int, c4 int)
  ORGANIZE BY DIMENSIONS (c1, c2, (c3,c1), c4)
```

## MDC 表的块索引

本主题说明如何使用块索引组织 MDC 表中的记录。

图 12 中显示的 MDC 表是以物理方式组织的，这样就可以将具有相同“Region”和“Year”值的记录组成为单独的块或扩展数据块。扩展数据块是磁盘上的一组连续页，所以将这些记录的组集群在物理上是连续的数据页上。每个表页都只属于一个块，所有块的大小都相同（即，它们的页数相同）。一个块的大小等于表空间的扩展数据块大小，以便使块边界与扩展数据块边界对齐。在这种情况下，将创建两个块索引，一个针对“Region”维，另一个针对“Year”维。这些块索引包含仅指向表中的块的指针。当对“Region”块索引进行扫描以找到“Region”等于“East”的所有记录时将找到两个合格的块。将在这两个块中找到“Region”等于“East”的所有记录（并且也只能找到这些记录），而且将这些记录集群在这两组连续页或扩展数据块上。同时，完全独立地扫描“Year”索引（1999 与 2000 之间的记录）将找到三个合格的块。对这三个块中的每个块进行数据扫描时都将返回在 1999 与 2000 之间的所有记录，并且也只能返回这些记录，并且您将发现这些记录集群在每个块的连续页上。

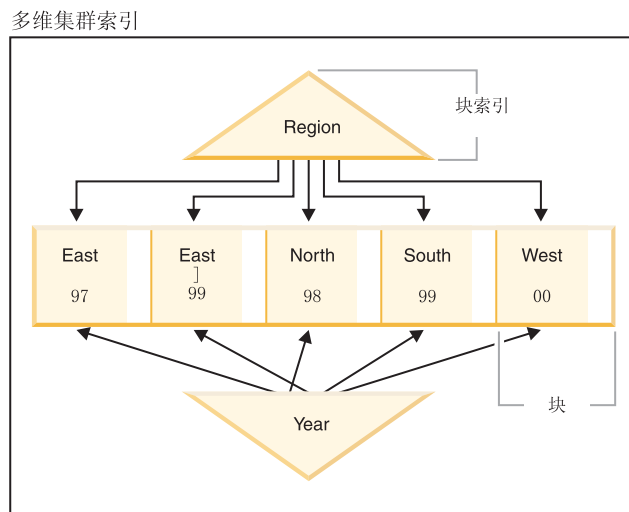


图 12. 多维集群表

除了对集群所作的这些改进之外，MDC 表还具有下列优点：

- 由于与基于记录的索引相比，由于块索引的大小是很小的，所以探测和扫描块索引要快得多
- 块索引和相应的数据组织允许更精细的“数据库分区忽略”或者有选择性地访问
- 利用块索引的查询因为减小了索引大小、优化了对块的预取并且可以保证相应数据的集群而受益
- 对于某些查询，可以减少锁定和谓词求值
- 块索引用于日志记录和维护方面的开销很少，因为仅当将第一条记录添加至块或从块中除去最后一条记录时才需要更新它们
- 转入的数据可以复用先前转出的数据留下的连续空间。

**注：**即使只是使用单个维定义的 MDC 表也可以因这些 MDC 属性而受益，并且可以是具有集群索引的常规表的可行备用。应该根据许多因素来作出此决定，这些因素包括：组成工作负载的查询以及表中数据的性质和分布。请参阅第 37 页的『选择 MDC 表维』和设计顾问程序。

创建表时，可以将一个或多个键指定为集群数据时所采用的维。每个 MDC 维都可以包括类似于常规索引键的一列或多列。将为每个指定的维自动创建维块索引，并且优化器将使用它根据每个维快速并有效地访问数据。还将自动创建组合块索引，它包含所有维中的所有列，并且将用来在插入和更新活动期间维护数据的集群。仅当单个维尚未包含所有的维键列时，才创建组合块索引。优化器还可以选择组合块索引来有效地访问满足一部分列维或所有列维中的值的数据。

**注：**此索引在处理查询期间的用途取决于它的键部分的顺序。键部分顺序是由解析器在对在 CREATE TABLE 语句的 ORGANIZE BY 子句中指定的维进行解析时遇到的列顺序来确定的。有关更多信息，请参阅第 48 页的『MDC 表的块索引注意事项』。

块索引与常规索引在结构上是相同，只不过它们指向的是块而不是记录。块索引比常规索引小的倍数为每页平均记录数乘以块大小所得到的乘积。一个块中的页数等于表空间的扩展数据块大小，它的范围是 2 到 256 页。页大小可以为 4 KB、8 KB、16 KB 或 32 KB。



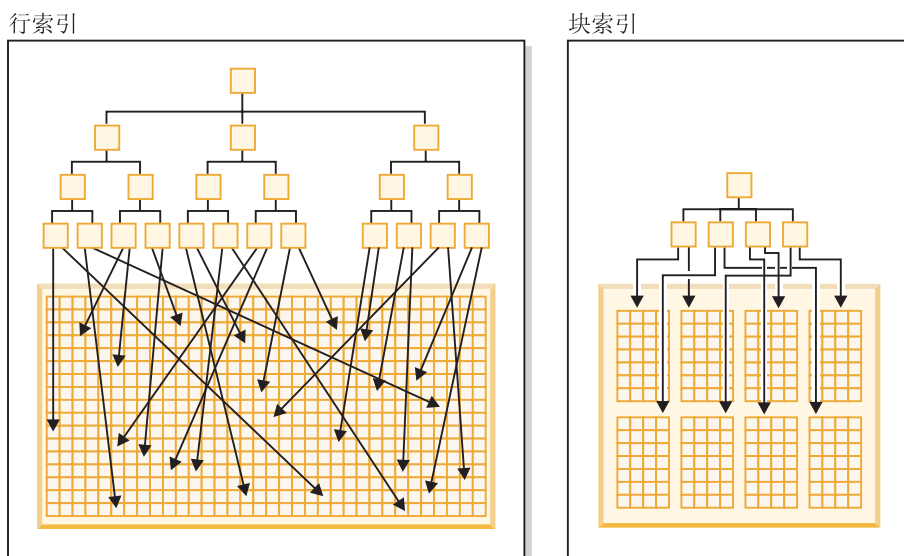


图 13. 行索引与块索引的区别

如图 13 中所示，与每一行具有单个条目相比，在块索引中，每个块都有单个索引条目。因此，块索引显著减少了磁盘的使用，并且极大地提高了数据访问速度。

在 MDC 表中，维值的每个唯一组合组成逻辑单元，它在物理上是由一个或多个页的块组成的。逻辑单元将仅使足够多的块与它相关联，以便存储具有该逻辑单元的维值的记录。如果表中没有具有特定逻辑单元的维值的记录，那么将不会为该逻辑单元分配块。包含具有特定维键值的数据的块的集合称为片。

## 方案：多维集群（MDC）表

通过一个方案说明如何使用 MDC 表：我们假设一个名为“Sales”的 MDC 表，它记录有关国内零售商的销售数据。该表根据“YearAndMonth”和“Region”维进行集群。表中的记录是以块为单位存储的，块包含磁盘上足够的连续页以填充扩展数据块。

在第 52 页的图 14 中，块由矩形表示，并且根据表中分配的扩展数据块的逻辑顺序进行编号。图表中的网格表示这些块的逻辑数据库分区，并且每个正方形表示一个逻辑单元。网格中的列和行表示特定维的片。例如，“Region”列中包含“South-central”值的所有记录都可以在由网格中的“South-central”列定义的片中包含的块中找到。实际上，此片中的每个块也只包含“Region”字段中具有“South-central”的记录。这样，当且仅当一个块包含在“Region”字段中具有“South-central”的记录时，该块才包含在此片或网格的列中。

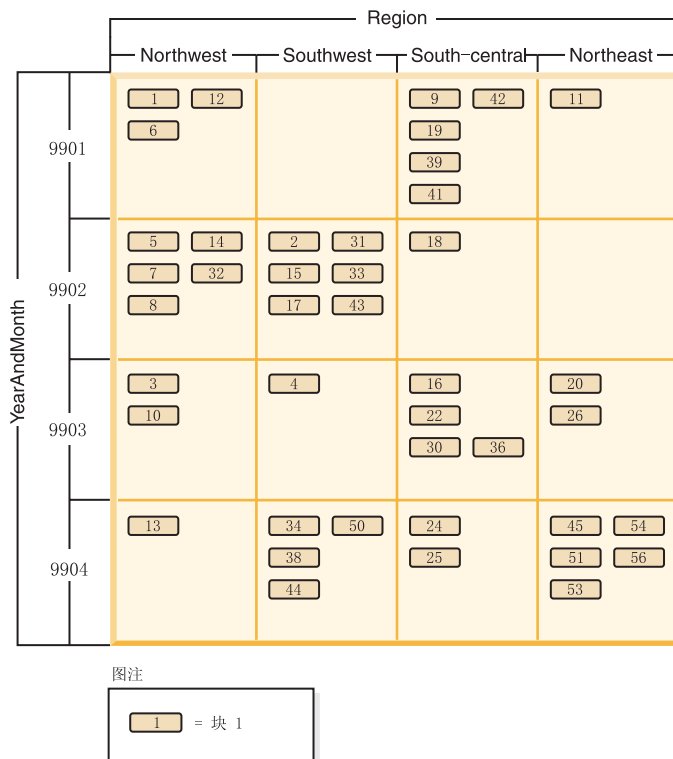


图 14. 具有“Region”和“YearAndMonth”维的称为 Sales 的多维表

为了便于确定哪些块组成片，或者哪些块包含具有特定维键值的所有记录，在创建表时将自动为每个维创建维块索引。

在第 53 页的图 15 中，一个维块索引是根据“YearAndMonth”维创建的，另一个维块索引是根据“Region”维创建的。每个维块索引的结构方式与传统 RID 索引的结构方式相同，但在叶级，这些键指向块标识（BID）而不是记录标识（RID）。RID 通过物理页号和槽号（找到记录的页上的槽）来标识表中记录的位置。BID 通过该扩展数据块的第一页的物理页号和虚槽（0）来表示块。因为块中的所有页在物理上是从该页开始连续的，并且我们知道该块的大小，所以可以使用此 BID 来找到该块中的所有记录。

片，或包含带有在维中具有特定键值的记录的页的一组块，将由该键值的 BID 列表在相关维块索引中表示。

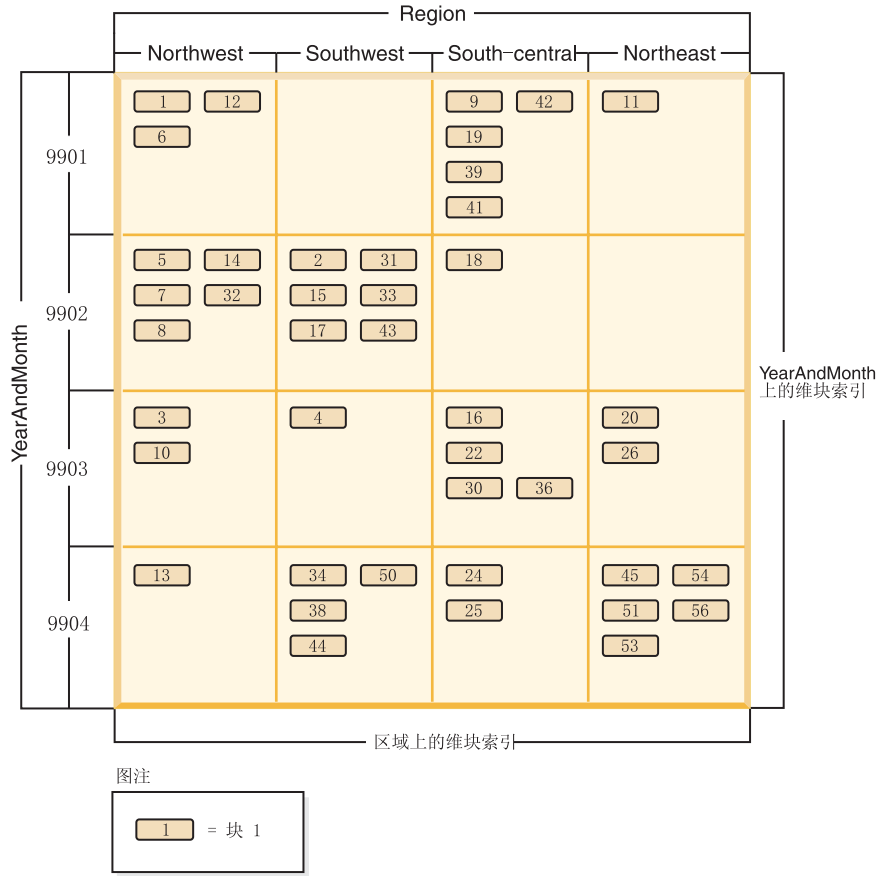


图 15. 显示维块索引且具有“Region”和“YearAndMonth”维的 Sales 表

图 16 显示“Region”的维块索引中的键如何出现。键由键值（即“South-central”）和 BID 列表组成。每个 BID 包含一个块位置。在图 16 中，列示的块号与在 Sales 表的网格中发现的“South-central”片相同（请参阅第 52 页的图 14）。

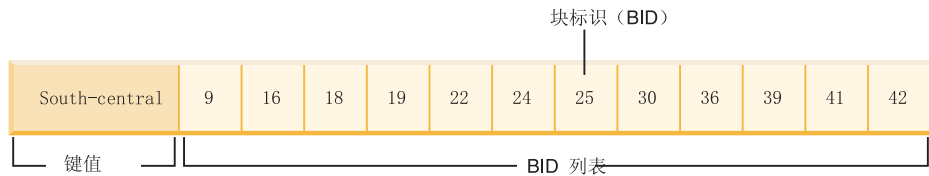


图 16. “Region”的维块索引中的键

类似的，要找到包含“YearAndMonth”维为“9902”的所有记录的块的列表，应在“YearAndMonth”维块索引中查找此值，如第 54 页的图 17 中所示。

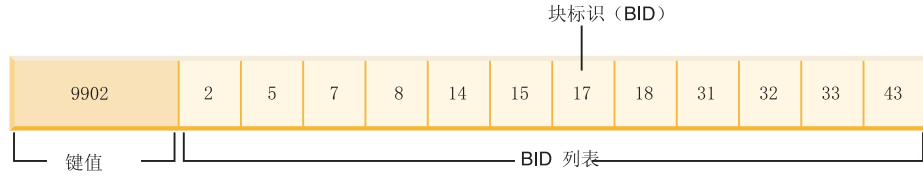


图 17. “YearAndMonth”的维块索引中的键

## MDC 表的块索引和查询性能

扫描 MDC 表的任何块索引将提供集群数据访问，这是因为每个块标识 (BID) 都与保证包含具有指定维值的数据的表中的一组连续页相对应。此外，还可以通过维或片的块索引来单独访问维或片，而不会损害任何其他维或片的集群因子。这就提供了多维集群的多维性。

利用块索引访问的查询在很多方面提高了性能。第一，块索引比常规索引小很多，块索引扫描的效率很高。第二，当使用块索引时，预取数据页不依赖于顺序检测。DB2 在索引中向前查找，使用大块 I/O 将块的数据页预取到内存中并确保在表中访问数据页时扫描不会引起 I/O。第三，表中的数据集群在连续页上，这样就优化了 I/O 并将结果集定位到表的选择的部分。第四，如果在基于块的缓冲池的块大小等于扩展数据块大小的情况下使用该缓冲池，那么将把 MDC 块从磁盘上的连续页预取到内存中，这样就进一步增强了集群对性能所起的作用。最后，每个块中的记录都是通过对它的数据页使用最小关系扫描来检索的，这种扫描数据的方法通常比通过基于 RID 的检索更快。

查询可以使用块索引来缩小表中具有特定维值或某一范围内的值的部分。这就提供了“数据库分区忽略”（即，块忽略）的精确格式。这样可使表具有更好的并行性，因为其他查询、装入、插入、更新和删除可访问表中的其他块，而不需要与此查询的数据集交互。

如果 Sales 表根据三个维进行集群，那么还可以使用个别维块索引来查找包含满足对表的所有维的子集的查询的记录的一组块。如果表具有“YearAndMonth”、“Region”和“Product”维，那么可以认为它是逻辑立方体，如第 55 页的图 18 中所示。

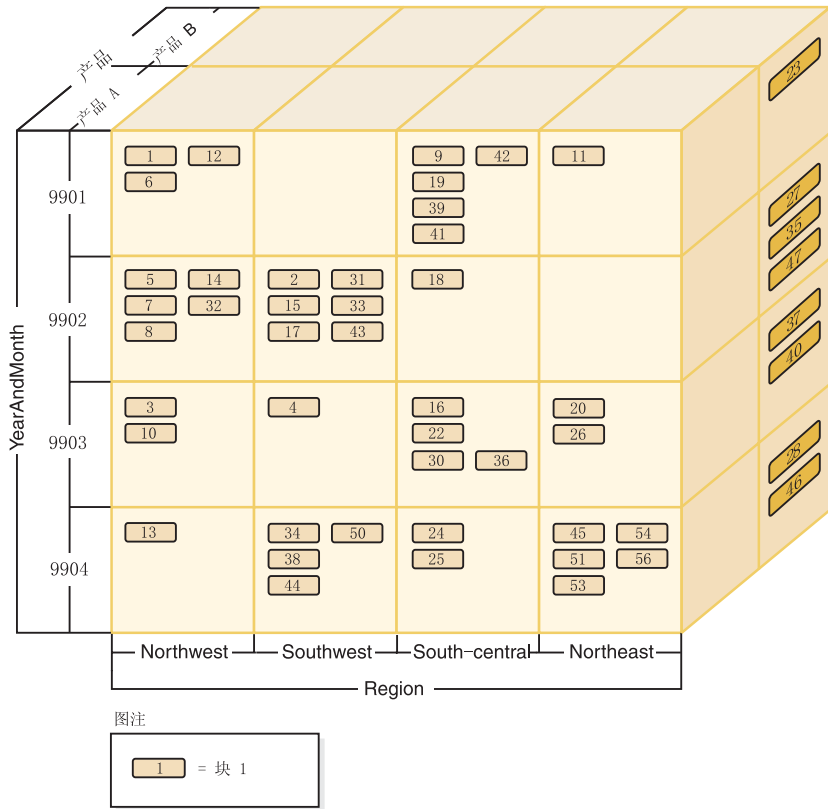


图 18. 具有“Region”、“YearAndMonth”和“Product”维的多维表

将对图 18 中显示的 MDC 表创建四个块索引，为“YearAndMonth”、“Region”和“Product”维每个都创建一个块索引；而另一个块索引将所有这些维列作为它的键。要检索“Product”等于“ProductA”并且“Region”等于“Northeast”的所有记录，数据库管理器将首先从“Product”维块索引中搜索 ProductA 键。（请参阅图 19。）然后，数据库管理器通过在“Region”维块索引中查找“Northeast”键来确定包含“Region”等于“Northeast”的所有记录的块。（请参阅图 20。）

|      |   |   |   |     |    |     |    |    |    |    |    |    |     |    |
|------|---|---|---|-----|----|-----|----|----|----|----|----|----|-----|----|
| 产品 A | 1 | 2 | 3 | ... | 11 | ... | 20 | 22 | 24 | 25 | 26 | 30 | ... | 56 |
|------|---|---|---|-----|----|-----|----|----|----|----|----|----|-----|----|

图 19. “Product”的维块索引中的键

|           |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|-----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Northeast | 11 | 20 | 23 | 26 | 27 | 28 | 35 | 37 | 40 | 45 | 46 | 47 | 51 | 53 | 54 | 56 |
|-----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

图 20. “Region”的维块索引中的键

可以通过使用逻辑 AND 和逻辑 OR 运算符来组合块索引扫描，并且获得的要扫描的块列表还提供集群数据访问。

以上面的示例为例，要查找包含具有这两个维值的所有记录的一组块，就必须找到这两个片的交集。这是通过对两个块索引键中的 BID 列表使用逻辑 AND 操作来完成的。常用的 BID 值有 11、20、26、45、54、51、53 和 56。

以下示例说明如何对块索引使用逻辑 OR 操作来满足具有涉及到两个维的谓词的查询。图 21 假定一个 MDC 表具有两个维：“Colour”和“Nation”。目标是检索 MDC 表中满足以下条件的所有记录：“Colour”为“blue”或者“Nation”名称为“USA”。

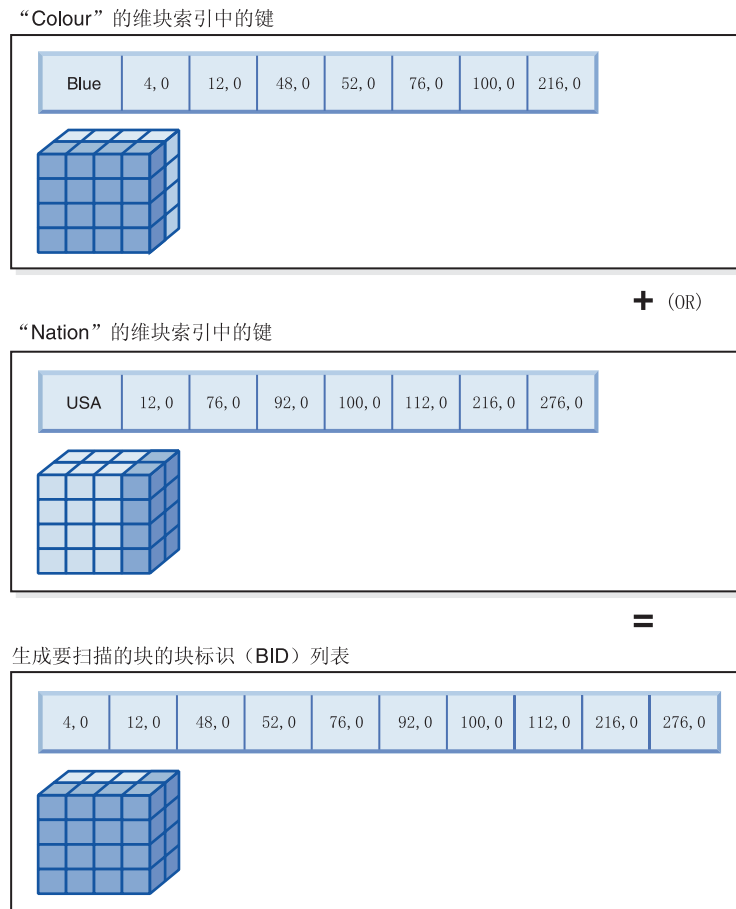


图 21. 可以如何对块索引使用逻辑 OR 操作

本图说明了如何组合两个单独的块索引扫描的结果以确定满足谓词限制的值的范围。（数目指示记录标识 (RID) 和槽字段。）

根据 SELECT 语句中的谓词，完成了两个单独的维块索引；一个是针对 blue 片，另一个是针对 USA 片。在内存中完成了逻辑 OR 操作，以便找到两个片的并集并确定在这两个片中找到的组合块集合（包括除去重复的块）。

一旦数据库管理器具有要扫描的块列表，数据库管理器就可以对每个块执行最小关系扫描。可以完成块的预取，并且对于每个块，此操作将只涉及到一个 I/O，因为每个块在磁盘上作为扩展数据块存储并且可以作为一个单元读取到缓冲池中。如果需要对数

据应用谓词，那么只需要对块中的一个记录应用谓词，因为将保证块中的所有记录都具有相同的维键值。如果存在其他谓词，那么数据库管理器只需要对块中的其余记录检查这些谓词。

MDC 表还支持基于 RID 的常规索引。可以通过使用逻辑 AND 操作或逻辑 OR 操作来将 RID 和块索引与其他索引组合在一起。块索引为优化器提供了可供选择的其他访问方案，但是您仍然可以使用传统访问方案（RID 扫描、连接、表扫描和其他方案）。在用于特定查询的所有其他可能的访问方案中，优化器将采用块索引方案，并且将选择成本最低的方案。

“DB2 设计顾问程序”可以帮助对 MDC 表建议基于 RID 的索引，或者对表建议 MDC 维。

---

## 在 INSERT 操作期间自动维护集群

通过使用组合块索引确保了自动维护 MDC 表中的数据集群。它用来在 INSERT 操作期间根据表维动态管理和维护数据的物理集群。

对于包含记录的表的每个逻辑单元，此组合块索引中只有一个键。因此，在 INSERT 期间使用此块索引来快速高效地确定表中是否存在逻辑单元，并且仅当表中存在逻辑单元时，才能准确地确定哪些块包含具有该单元的一组特定维值的记录。

当进行插入时：

- 将检查组合块索引，以找到与要插入的记录的维值相对应的逻辑单元。
- 如果在索引中找到了该逻辑单元的键，那么它的块标识（BID）列表将提供表中具有该逻辑单元的维值的各个块的完整列表。（请参阅第 58 页的图 22。）这就限制了表的用来搜索要插入记录的空间的扩展数据块数目。
- 如果在索引中找不到该逻辑单元的键，或者如果包含这些值的扩展数据块已满，那么为该逻辑单元指定新块。如果可能的话，在使用页的另一个新扩展数据块（新块）来扩展表之前，先复用该表中的空块。

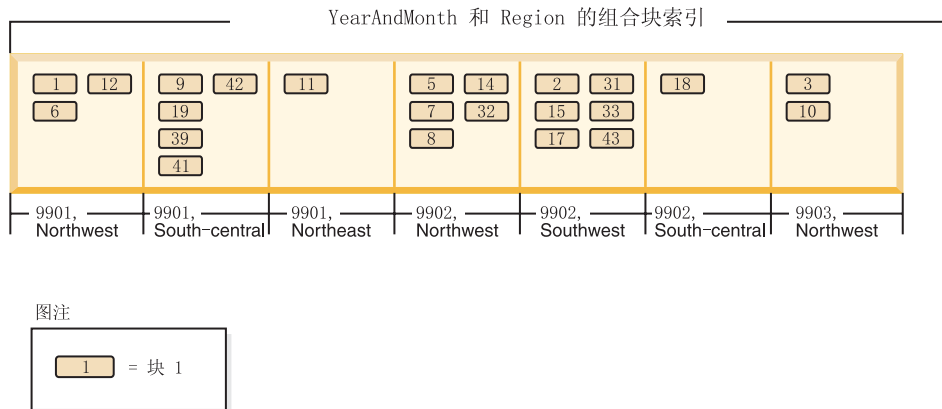


图 22. “YearAndMonth”和“Region”的组合块索引

已经保证了在包含并且仅包含具有特定维值的所有记录的一组块中将找到具有这些值的数据记录。这些块是由磁盘上的连续页组成的。因此，将按顺序访问这些记录，并且将提供集群。通过确保将记录从具有该记录的维值的单元中插入到块中来随时间推移自动维护此集群。当逻辑单元中的现有块已满时，将复用空块，或者将分配新块并将它添加至该逻辑单元的一组块。当一个块中没有数据记录时，将从块索引中除去块标识（BID）。这将使该块不再与任何逻辑单元值相关联，以便将来另一个逻辑单元可以复用该块。因此，将根据需要从表中动态添加和除去单元以及与它们相关联的块索引条目，以便只容纳表中存在的数据。组合块索引用来管理这种情况，因为它将逻辑单元值映射至包含具有这些值的记录的那些块。

因为集群是按这种方式自动维护的，所以从不需要重组 MDC 表来重新集群数据。但是，仍然可以使用重组来回收空间。例如，如果单元具有许多稀疏块（数据只能满足很少的块），或者如果表具有许多指针溢出对，那么重组表时就会将属于每个逻辑单元的记录压缩到需要的最小块数中，并且将除去指针溢出对。

以下示例说明了可以如何将组合块索引用于查询处理。如果想要查找图 22 的表中“Region”为“Northwest”且“YearAndMonth”为“9903”的所有记录，那么数据库管理器将在组合块索引中查找键值 9903, Northwest，如图 23 中所示。键由键值（即“9903, Northwest”）和 BID 列表组成。可以看到列示的 BID 仅有 3 和 10，并且实际上 Sales 表中只有两个块包含具有这两个特定值的记录。

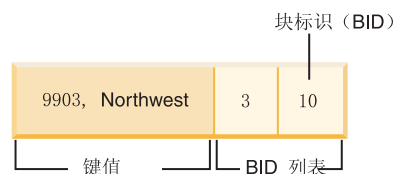


图 23. “YearAndMonth”和“Region”的组合块索引中的键



为了说明在插入期间如何使用组合块索引，我们以插入具有维值 9903 和 Northwest 的另一条记录为例。数据库管理器将在组合块索引中查找此键值，并查找第 3 个和第 10 个块的 BID。这些块包含具有这些维键值的仅有全部记录。如果具有可用空间，那么数据库管理器将把新记录插入到其中一个块中。如果在这些块中的所有页上都没有空间，那么数据库管理器将为表分配新块或使用表中先前清空的块。注意，在此示例中，表当前没有使用第 48 个块。数据库管理器将记录插入到该块中，并通过将该块的 BID 添加至组合块索引和每个维块索引来使此块与当前逻辑单元相关联。请参阅图 24 以获取有关添加第 48 个块之后的维块索引的键的说明。

|                 |   |    |    |    |    |    |    |    |    |    |    |    |
|-----------------|---|----|----|----|----|----|----|----|----|----|----|----|
| 9903            | 3 | 4  | 10 | 16 | 20 | 22 | 26 | 30 | 36 | 48 |    |    |
| Northwest       | 1 | 3  | 5  | 6  | 7  | 8  | 10 | 12 | 13 | 14 | 32 | 48 |
| 9903, Northwest | 3 | 10 | 48 |    |    |    |    |    |    |    |    |    |

图 24. 添加第 48 个块之后维块索引中的键

## MDC 表的块映射

当块被清空时，它就会通过从块索引中除去它的 BID 来不再与它的当前逻辑单元值相关联。然后，该块可被另一个逻辑单元复用。这就不需要使用新块来扩展表。

当需要新块时，需要快速找到先前被清空的块，而不必搜索表来找到它们。

块映射是用来便于找到 MDC 表中的空块的结构。块映射是作为单独的对象存储的：

- 在 SMS 中，作为单独的 .BKM 文件
- 在 DMS 中，作为对象表中的新对象描述符。

块映射是包含表的每个块的条目的一个数组。每个条目组成块的一组状态位。

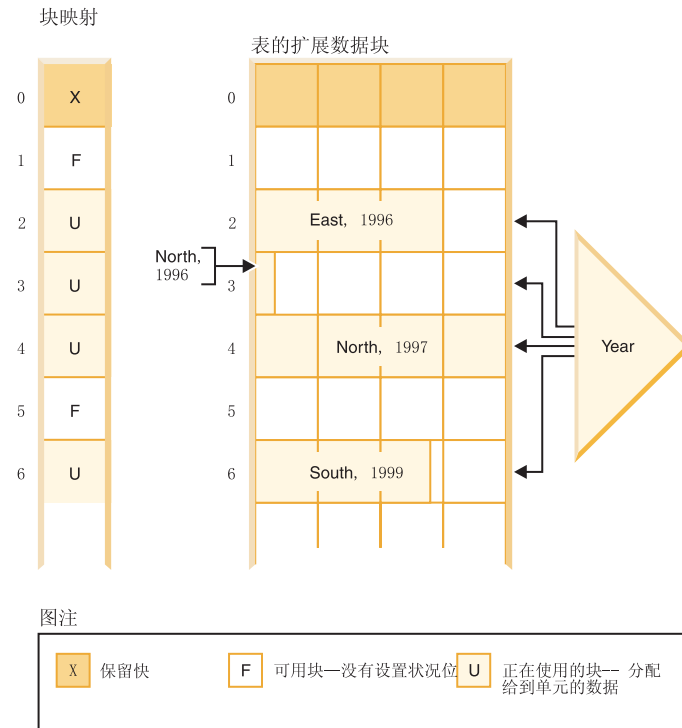


图 25. 块映射的工作方式

在图 25 的左边显示具有表中每个块的不同条目的块映射数组。右边显示了正在如何使用表的每个扩展数据块：某些扩展数据块是可用的，而大多数扩展数据块正处于使用状态，并且仅在块映射中标记为正在使用的块中找到了记录。为了简单起见，在图中只显示两个维块索引的其中一个。

**注：**

1. 块索引中存在仅指向块映射中标记为 **IN USE** 的那些块的指针。
2. 第一个块是保留块。此块包含表的系统记录。

通过扫描块映射以找到可用块（即，未设置任何位的那些块）来很容易找到可用块以便在单元中使用。

表扫描还使用块映射来仅访问当前包含数据的扩展数据块。根本不需要在表扫描中包括任何未在使用的扩展数据块。为了便于说明，此示例中的表扫描（图 25）将从表中的第三个扩展数据块（扩展数据块 2）开始，跳过第一个保留的扩展数据块和紧接着的空扩展数据块，扫描表中的第 2、3 和 4 个块，跳过下一个扩展数据块（不使用该扩展数据块的任何数据页），然后从那里继续扫描。

## 从 MDC 表中删除

当在 MDC 表中删除记录时，如果它不是块中的最后一条记录，那么数据库管理器只删除该记录并从在该表上定义的任何基于记录的索引中除去它的 RID。

但是，当删除操作除去块中的最后一条记录时，数据库管理器将通过更改它的 IN\_USE 状态位并从所有块索引中除去该块的 BID 来释放该块。同样，如果还存在基于记录的索引，那么也会除去它们的 RID。

注：因此，仅当块被彻底清空时，才需要对整个块除去一次块索引条目，而不是对基于记录的索引中已删除的每一行都除去一次。

---

## 更新 MDC 表

在 MDC 表中，更新非维值是在适当位置完成的，正如对常规表更新非维值一样。如果更新影响了变长列并且记录不再适合页，那么将查找另一个具有足够空间的页。

在同一个块中开始搜索此新页。如果该块中没有空间，那么使用插入新记录的算法来查找逻辑单元中具有足够空间的页。除非在单元中找不到空间并且需要将新块添加至该单元，否则不需要更新块索引。

更新维值被视作删除当前记录之后插入已更改的记录，因为该记录将更改它所属的逻辑单元。如果删除当前记录导致块被清空，那么需要更新块索引。类似地，如果插入新记录要求将它插入到新块中，那么需要更新块索引。

仅当将第一条记录插入到块中或者从块中删除最后一条记录时才需要更新块索引。因此，与块索引进行维护和日志记录相关联的索引开销小于与常规索引相关联的索引开销。对于在其他情况下将为常规索引的每个块索引，极大地减少了维护和日志记录开销。

MDC 表被视作任何现有表；即，可以对这些表定义触发器、引用完整性、视图和具体化查询表。

---

## 表分区和多维集群表

一个表可以同时是多维集群表和分区表。在同时是多维集群表和分区表的表中，可以同时使用表分区的范围分区规范和 MDC 键中使用列。与只单独使用多维集群或分区功能相比，这更有助于获取较详细的数据分区和块消除。在许多应用中，将 MDC 键列指定为不同于对表进行分区的列很有用。应该注意的是，表分区是多列，而 MDC 是多维。

### 主流 DB2 V9.1 数据仓库的特征

下列建议主要针对对于 DB2 V9.1 来说是新产品的典型主流仓库。假定下列特征：

- 数据库在多台机器或多个 AIX 逻辑分区上运行。
- 使用数据库分区功能（DPF）（表是使用 DISTRIBUTE BY HASH 子句创建的）。
- 有 4 到 50 个数据分区。
- 考虑其 MDC 和表分区的表是主要事实表。
- 表包含 1 亿到 1000 亿行。
- 在各种时间范围装入新数据：每夜、每周和每月。
- 每日接受量为 1 万到 1 亿条记录。
- 数据量变化：最多的一个月是最少的月的 5 倍。同样，最大维数（生产线，区域）具有 5 倍大小范围。
- 获取 1 到 5 年的详细数据。

- 每月或每个季度转出到期数据。
- 表使用大范围的查询类型。但是，相对于 OLTP 工作负载来说，该工作负载通常是具有下列特征的分析查询：
  - 较大的结果集，最多有 2 百万行
  - 大多数或全部查询都命中视图，而不是基本表
- SQL 子句按范围（BETWEEN 子句）、列表中的项等选择数据。

## 主流 DB2 V9.1 数据仓库事实表的特征

一个典型仓库事实表可能采用以下设计：

- 在 Month 列中创建数据分区。
- 为转出的每个时间段（例如，1 个月和 3 个月）定义数据分区。
- 在 Day 和 1 到 4 个其他维的基础上创建 MDC 维。典型的维有：生产线和区域。
- 所有数据分区和 MDC 集群都分布在所有数据库分区中。

MDC 和表分区具有一些相同的好处。下表列示组织中的潜在需求并根据先前确定的特征确定建议的组织方案。

表 7. 将表分区与 MDC 表配合使用

| 问题  | 建议的方案    | 建议                                       |
|---|----------|--|
| 转出期间的数据可用性  | 表分区      | 使用 DETACH PARTITION 子句来转出大量数据，并且只出现最少中断。 |
| 查询性能  | 表分区和 MDC | MDC 最适合用来查询多个维。表分区通过数据分区消除提高性能。          |
| 最少重组  | MDC      | MDC 维护集群，从而减少进行重组的必要性。                   |
| 在传统脱机窗口中转出一个<br>月或更长时间的数据                                   | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。   |
| 在短时间脱机窗口（小于 1 分<br>钟）期间转出一个月或更长时间<br>的数据                    | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。   |
| 转出一个月或更长时间的数据，<br>并同时在不损失任何服务的<br>情况下使表对于提交查询的<br>企业用户完全可用。 | MDC      | MDC 只能解决一部分此需求。由于表处于脱机状态的时间段太短，表分区并不适合。  |
| 每天装入数据（联机或脱机）   | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。                |
| “连续”装入数据（联机）  | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。                |
| “传统 BI”查询的查询执行性能  | 表分区和 MDC | MDC 特别适合用来查询立方体/多个维。表分区通过分区消除提高性能。       |

表 7. 将表分区与 MDC 表配合使用 (续)

| 问题   | 建议的方案 | 建议  |
|--|-------|---|
| 通过消除进行重组的必要性或减少执行任务所产生的不良影响，使重组所带来的不良影响降到最低。 | MDC   | MDC 维护集群，从而减少进行重组的必要性。如果使用 MDC，那么数据分区不提供递增好处。但是，如果不使用 MDC，那么范围分区通过在分区级别维护一些粗粒度集群，有助于减少重组的必要性。 |

示例 1:

考虑一个具有键列 YearAndMonth 和 Province 的表。一种合理的规划此表方法是按日期进行分区，每 2 个月添加一个数据分区。此外，还可以按 Province 进行组织，以便任何两个月日期范围内的特定省份的所有行集群在一起，如第 28 页的图 6 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (Province);
```

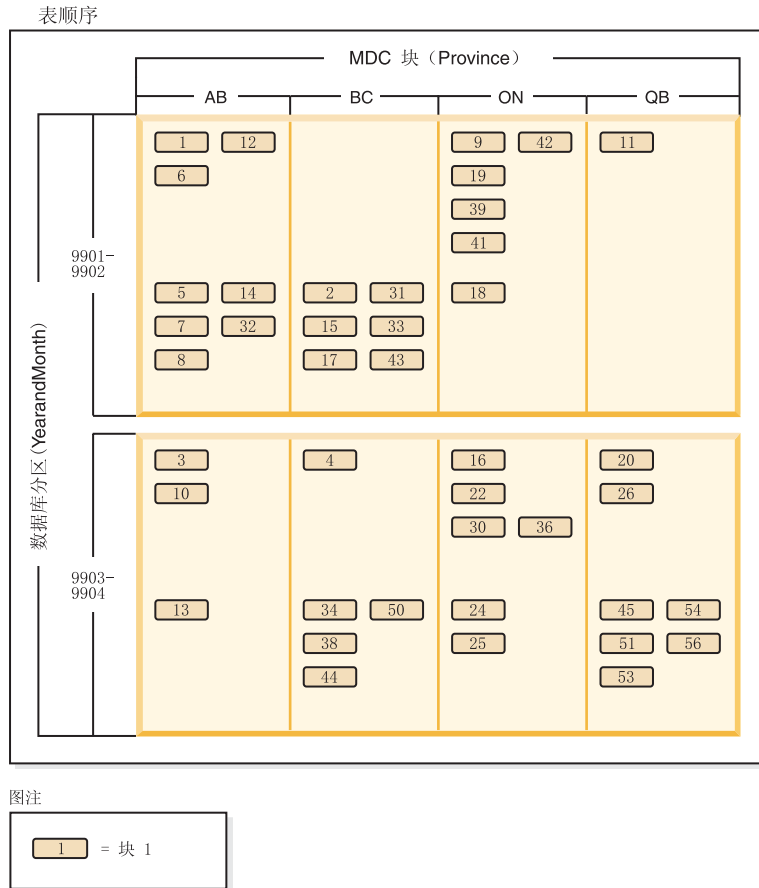


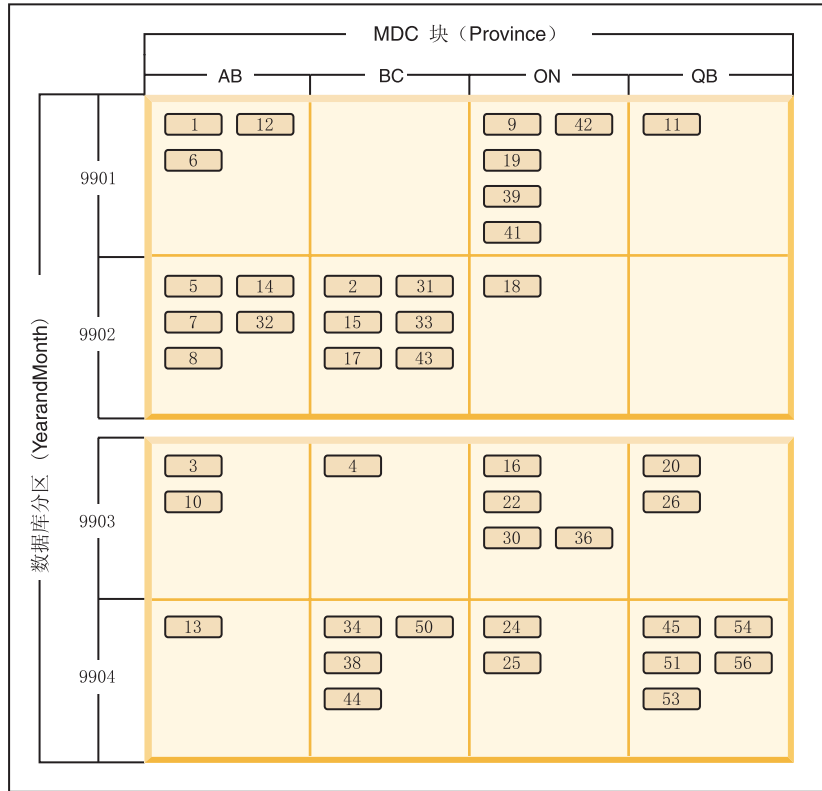
图 26. 按 YearAndMonth 分区并按 Province 组织的表

示例 2:

通过将 YearAndMonth 添加至 ORGANIZE BY 子句，可以获得较高的详细程度，如第 29 页的图 7 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (YearAndMonth, Province);
```

表顺序



图注

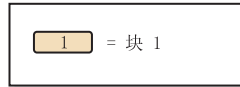


图 27. 按 YearAndMonth 分区并按 Province 和 YearAndMonth 组织的表

如果分区导致每个范围内只有单个值，那么通过在 MDC 键中包括表分区列不能获得任何好处。

### 注意事项

- 与基本表相比，MDC 表和分区表都需要一些存储器。这些存储器需求是附加的，但相对于所带来的好处来说，它们是合理的。
- 如果选择不在分区数据库环境中组合表分区和 MDC 功能，那么在您可以非常自信地预计数据分发情况时（通常也就是此处讨论的系统类型情况），使用表分区是最好的选择。否则，应考虑 MDC。





---

## 第 4 章 并行数据库系统

---

### 并行性

任务（例如，数据库查询）的各个组件可并行运行以大幅提高性能。任务的性质、数据库配置和硬件环境都确定 DB2 数据库产品将如何以并行方式执行任务。

这些注意事项是相关的，在决定数据库的物理和逻辑设计方案时，应当一起考虑。DB2 数据库系统支持下列类型的并行性：

- I/O
- 查询
- 实用程序

#### 输入/输出并行性

当一个表空间有多个容器时，数据库管理器可以利用并行 I/O。并行 I/O 指的是同时处理两个或多个 I/O 设备的写入或读取；这能够显著地改进吞吐量。

#### 查询并行性

有两种类型的查询并行性：查询间并行性和查询内并行性。

**查询间并行性**是指数据库同时接受多个应用程序查询的能力。每个查询以独立于其他查询的方式运行，但数据库管理器同时运行所有查询。DB2 数据库产品始终支持这种类型的并行性。

**查询内并行性**是指使用分区内并行性和/或分区间并行性来同时处理单个查询的各部分。

#### 分区内并行性

**分区内并行性**是指将一个查询分为多个部分的能力。某些 DB2 实用程序也执行此类型的并行性。

分区内并行性将通常认为的单个数据库操作（例如，创建索引、装入数据库或 SQL 查询）细分成多个部分，其中的大部分或全部操作可以在单个数据库分区内以并行方式运行。

第 68 页的图 28 显示了一个查询，它被分为可并行运行的四个部分，返回结果的速度比按串行方式运行该查询的速度快得多。这几部分互为副本。要利用分区内并行性，必须适当地配置数据库。您可以选择并行度，或由系统为您选择。并行度表示一个查询中并行运行的部分数。

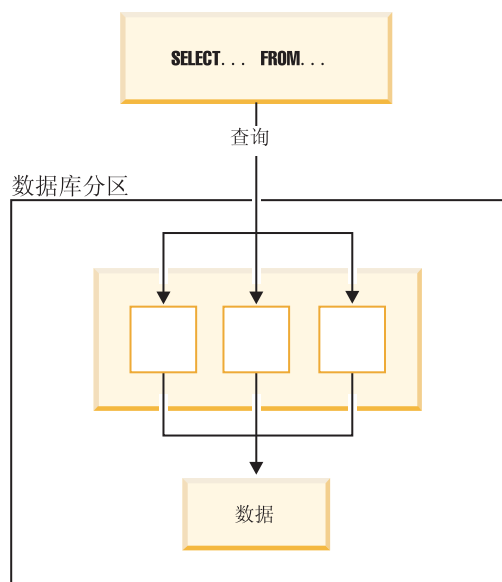


图 28. 分区内并行性

## 分区间并行性

分区间并行性是指将一个查询分为多个部分并将这几部分置于一个分区数据库的多个分区（位于一台或多台机器上）上的能力。该查询以并行方式运行。某些 DB2 实用程序也执行此类型的并行性。

分区间并行性将通常认为的单个数据库操作（例如，创建索引、装入数据库或 SQL 查询）细分成多个部分，其中的大部分或全部操作可以在一台或多台机器上的一个分区数据库的多个分区中以并行方式运行。

第 69 页的图 29 显示了一个查询，它被分为可并行运行的四个部分，返回结果的速度比在单个数据库分区上按串行方式运行该查询的速度快得多。

并行度在很大程度上取决于您创建的数据库分区数和您定义数据库分区组的方式。

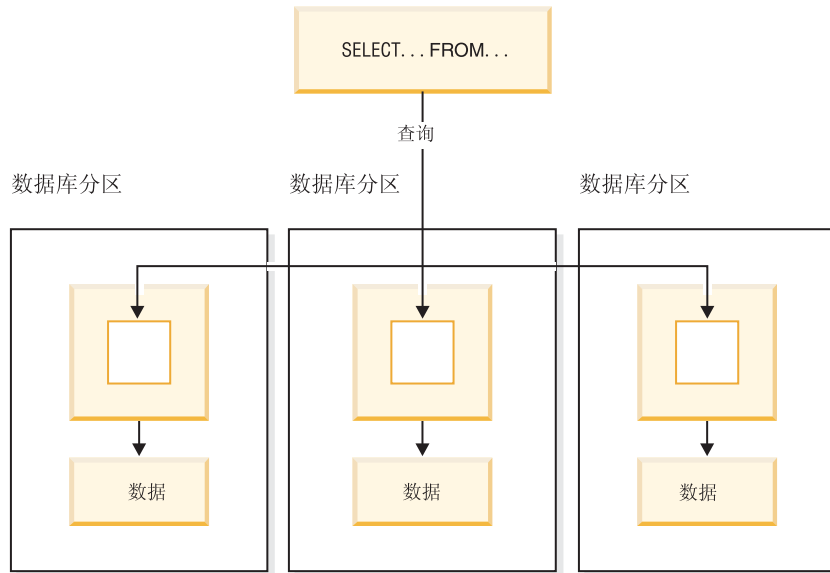


图 29. 分区间并行性

### 同时使用分区内和分区间并行性

可以同时使用分区内并行性和分区间并行性。此组合提供了两种并行性，这也使处理查询的速度显著加快。

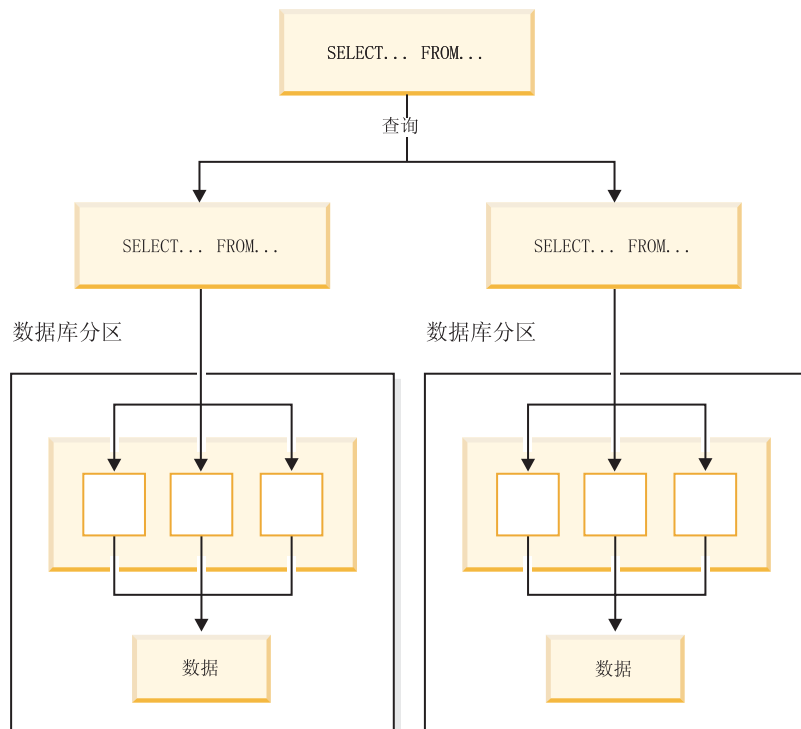


图 30. 同时使用分区间并行性和分区内并行性

## 实用程序并行性

DB2 实用程序可以利用分区内并行性。它们还可以利用分区间并行性；前提是存在多个数据库分区，而实用程序在每个数据库分区内并行执行。

Load 实用程序可以利用分区内并行性和 I/O 并行性。装入数据是一个大量使用 CPU 的任务。Load 实用程序利用多个处理器来执行如解析和格式化数据这类任务。它也可使用并行 I/O 服务器来以并行方式将数据写入容器中。

在分区数据库环境中，LOAD 命令通过在该表所在的每个数据库分区上进行并行调用来利用分区内、分区间和 I/O 并行性。

在创建索引期间，可并行执行数据的扫描和后续排序。在创建索引时，DB2 系统既利用 I/O 并行性又利用分区内并行性。这有助于在发出 CREATE INDEX 语句时、重新启动期间（若一个索引标记为无效）及数据重组期间加快索引创建的速度。

备份和复原数据任务是繁重地涉及 I/O 的任务。在执行备份和恢复操作时，DB2 系统既利用 I/O 并行性又利用分区内并行性。备份操作通过以并行方式读取多个表空间容器并以并行方式异步写入多备份介质，来利用 I/O 并行性。

## 分区数据库环境

“数据库分区功能”（DPF）将数据库管理器扩展至并行多分区环境。

- 数据库分区是数据库的一部分，它由其自己的数据、索引、配置文件和事务日志组成。数据库分区有时称为节点或数据库节点。分区数据库环境是支持将数据分布到各数据库分区上的数据库安装。
- 单一分区数据库是只有一个数据库分区的数据库。该数据库中的所有数据都存储在这个单一数据库分区中。对于这种情况，数据库分区组（如果存在的话）未提供任何其他功能。
- 多分区数据库是有两个或更多个数据库分区的数据库。表可以存储在一个或多个数据库分区中。当表存储在包含多个数据库分区的数据库分区组中时，它的某些行存储在一个数据库分区中，而其他行存储在其他数据库分区中。

通常，在每台物理机器上都存在单个数据库分区，而在每个数据库分区中，数据库管理器使用每个系统上的处理器来管理该数据库的全部数据中属于该分区的那一部分。

由于数据分布在各数据库分区中，所以可以使用多台物理机器上多个处理器的能力来满足对信息的请求。数据检索和更新请求被自动分解成子请求，并在适用的数据库分区中并行执行。数据库分布在各数据库分区中的这个事实对于发出 SQL 语句的用户是透明的。

用户交互通过一个数据库分区发生，该数据库分区称为该用户的协调程序分区。协调程序分区与应用程序在同一个数据库分区中运行，对于远程应用程序来说，协调程序分区在该应用程序所连接至的数据库分区中运行。任何数据库分区都可用作协调程序分区。

数据库管理器允许将数据存储于数据库的多个数据库分区中。这意味着该数据以物理方式存储在多个数据库分区中，但是对其进行访问时，仍可以将数据视为位于同一位置。访问多分区数据库中的数据的应用程序和用户不需要知道该数据的物理位置。

该数据在物理上是分离的，但却将它作为一个逻辑整体来使用和管理。用户可以通过声明分布键来选择如何分布数据。用户还可以通过选择表空间和应存储数据的相关数据库分区组，确定数据可以分布在哪些数据库分区上以及分布在多少个数据库分区上。可以使用 DB2 设计顾问程序来完成有关分布和复制的建议。另外，将可更新的分布图与散列算法一起使用，来指定分布键值到数据库分区的映射，以确定每行数据的位置和检索。因此，可以将大型表的工作负载分布在多分区数据库上，并将较小的表存储在一个或多个数据库分区中。每个数据库分区都有它所存储数据的本地索引，这可以提高本地数据的访问性能。

**注：**并不是必须将所有表都分布到数据库的所有数据库分区中。数据库管理器支持部分分区，这意味着可以将表及其表空间分布到系统中的数据库分区的子集上。

当您想要将表放在每个数据库分区中时，可考虑的替代方法是使用具体化查询表，然后复制那些表。您可以创建包含所需信息的具体化查询表，然后将它复制到每个数据库分区。

非 root 用户安装的 DB2 数据库产品不支持数据库分区。因此，添加节点操作无法运行。不应手动更新 db2nodes.cfg 文件。手动更新将导致显示 SQL6031N 错误。

## 数据库分区和处理器环境

容量是指能访问数据库的用户数和应用程序数。这很大程度上取决于内存、代理程序数、锁定数、I/O 和存储器管理。可伸缩性是指一个数据库随着其增长而继续显示出相同的操作特征和响应时间的能力。

本节对下列硬件环境提供了一个概述：

- 单处理器上的单个数据库分区（单处理器）
- 具有多处理器的单个数据库分区（SMP）
- 多个数据库分区配置
  - 具有单处理器的数据库分区（MPP）
  - 具有多处理器的数据库分区（SMP 集群）
  - 逻辑数据库分区

下面讨论每种环境的容量和可伸缩性。

### 单处理器上的单个数据库分区

此环境由内存和磁盘组成，但仅包含一个 CPU（请参阅图 31）。可使用许多不同的名称来称呼它，包括独立数据库、客户机/服务器数据库、串行数据库、单处理器系统和单节点或非并行环境。

此环境中的数据库可满足一个部门或小办公室的需要，其中的数据和系统资源（包括一个单处理器或 CPU）由单数据库管理器来管理。

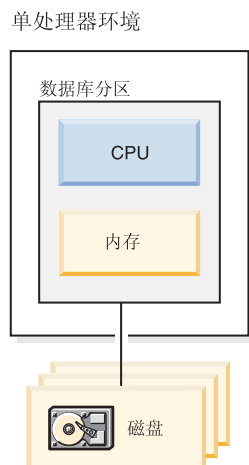


图 31. 单处理器上的单个数据库分区

### 容量和可伸缩性

在此环境中可以添加更多的磁盘。如果让每个磁盘拥有一个或多个 I/O 服务器，那么多个 I/O 操作可同时执行。

单处理器系统受处理器可处理的磁盘空间容量的限制。无论可以添加的其他组件（如内存或磁盘）如何，随着工作负载的增加，单个 CPU 可能无法更快地处理用户请求。如果已达到最大容量或最大可伸缩性，那么可考虑移到一个有多处理器的单个数据库分区系统中。

## 具有多处理器的单个数据库分区

此环境通常由同一台机器中几个能力相等的处理器组成（请参阅图 32），称为对称多处理器（SMP）系统。像磁盘空间和内存这类资源是共享的。

利用可用的多个处理器，可以更快地完成不同的数据库操作。DB2 数据库系统还可将单个查询的工作分布在可用的处理器中，以提高处理速度。其他数据库操作，例如，装入数据、备份和复原表空间以及对现有数据创建索引，都可以利用多个处理器。

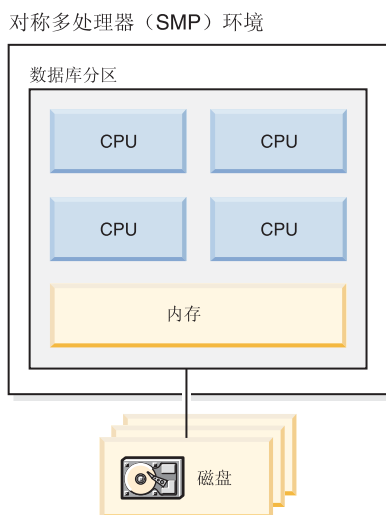


图 32. 单分区数据库对称多处理器环境

### 容量和可伸缩性

在此环境中可以添加更多的处理器。然而，由于不同的处理器可能会试图访问同一数据，所以随着您公司业务量的增加，此环境也可能遇到限制。利用共享内存和共享磁盘，可有效地共享所有数据库数据。

可以通过增加磁盘数来增加与处理器相关的数据库分区的 I/O 容量。可以建立 I/O 服务器以专门处理 I/O 请求。如果让每个磁盘拥有一个或多个 I/O 服务器，那么多个 I/O 操作可同时执行。

如果已达到最大容量或最大可伸缩性，那么可考虑移到有多个数据库分区的系统中。

### 多个数据库分区配置

可以将一个数据库分布在多个数据库分区中，每个数据库分区在它自己的机器上。可将有多个数据库分区的多台机器编组在一起。本节描述下列数据库分区配置：

- 具有单处理器的系统上的数据库分区
- 具有多处理器的系统上的数据库分区
- 逻辑数据库分区

## 具有单处理器的数据库分区

在此环境中，有许多数据库分区。每个数据库分区都位于它自己的机器上，而且它有自己的处理器、内存和磁盘（图 33）。所有机器通过通信工具连接在一起。可使用许多不同的名称来称呼此环境，包括：集群、单处理器集群、大规模并行处理（MPP）环境和“不共享”配置。最后一个名称准确地反映了此环境中的资源安排。与 SMP 环境不同，MPP 环境没有共享的内存或磁盘。MPP 环境消除了由共享内存和磁盘带来的限制。

一个分区数据库环境允许一个数据库保持逻辑上的整体性，但它实际上分布于多个数据库分区中。数据是分布的这一事实对大多数用户是透明的。可以在数据库管理器之间划分工作；每个数据库分区中的每个数据库管理器都只为该数据库中它自己的那部分工作。

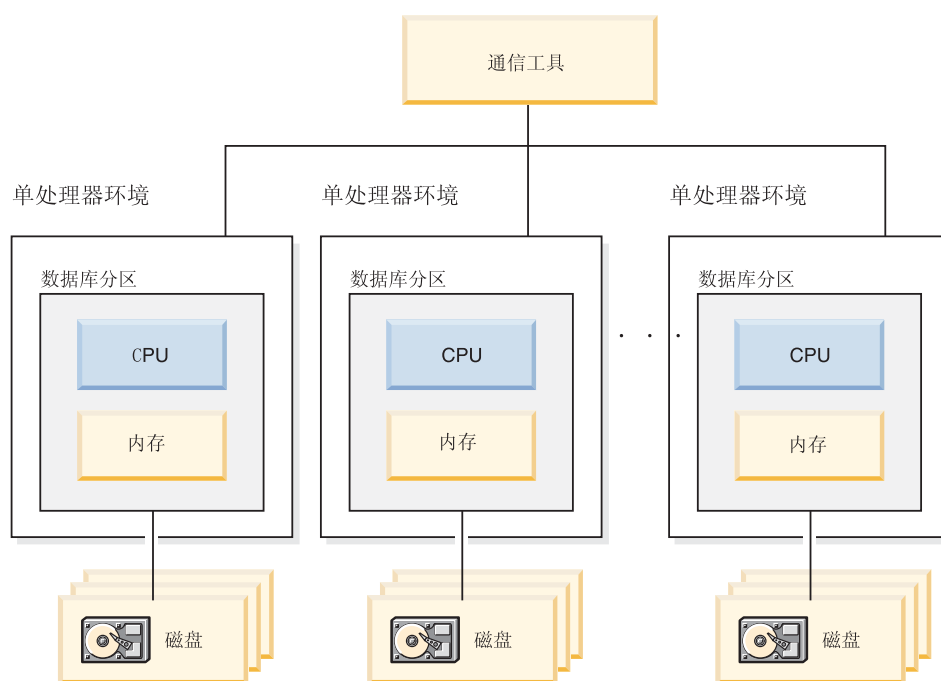


图 33. 大规模并行处理（MPP）环境

### 容量和可伸缩性

在此环境中可以向您的配置添加更多的数据库分区（节点）。在某些平台上（例如，RS/6000® SP），最大数目是 512 个节点。然而，在管理很多机器和实例时，可能存在实际的限制。

如果已达到最大容量或最大可伸缩性，那么可考虑移到每个数据库分区都有多个处理器的系统中。



## 具有多处理器的数据库分区

每个数据库分区具有单处理器的替代配置是，每个数据库分区具有多处理器的配置。这称为 *SMP* 集群（图 34）。

此配置结合了 *SMP* 和 *MPP* 并行性的优点。这表示一个查询可以在跨多处理器的单个数据库分区中执行。亦即一个查询可以用并行方式在多个数据库分区中执行。

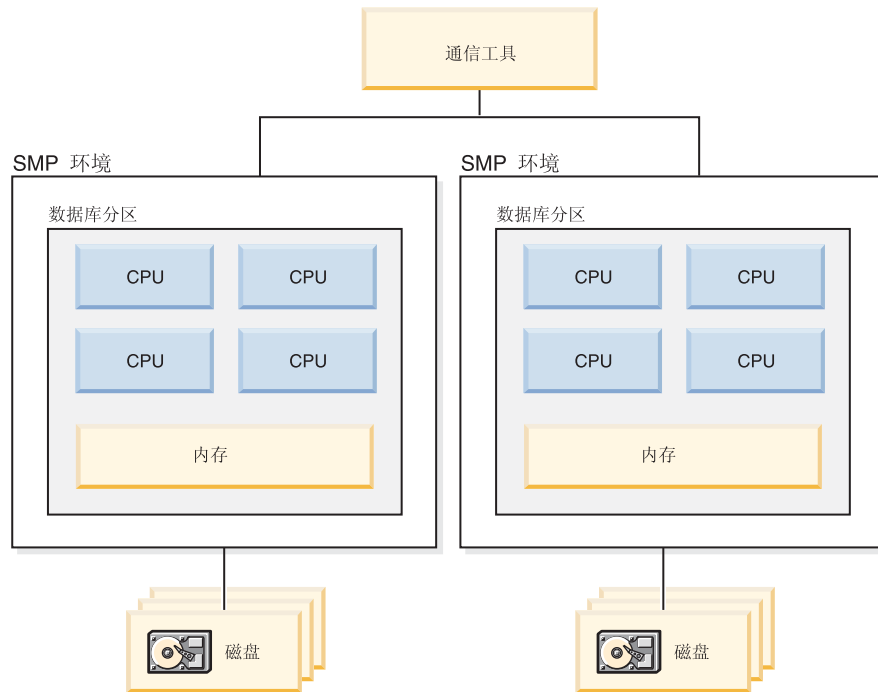


图 34. 集群中的几个对称多处理器（*SMP*）环境

### 容量和可伸缩性

在此环境中，可以添加更多的数据库分区，并可以向现有数据库分区添加更多的处理器。

### 逻辑数据库分区

逻辑数据库分区与物理分区的不同之处在于逻辑分区未被授予对整台机器的控制权。虽然机器已共享资源，但是数据库分区不共享资源。处理器是共享的，但磁盘和内存却不共享。

逻辑数据库分区提供了可伸缩性。在多个逻辑分区上运行的多个数据库管理器可以比一个单数据库管理器更能充分利用可用的资源。第 76 页的图 35 说明了通过添加更多的数据库分区可以在一台 *SMP* 机器上获得更大的可伸缩性；对于那些具有许多处理器的机器而言更是如此。通过分配数据库，可以分别对每个数据库分区进行管理和恢复。

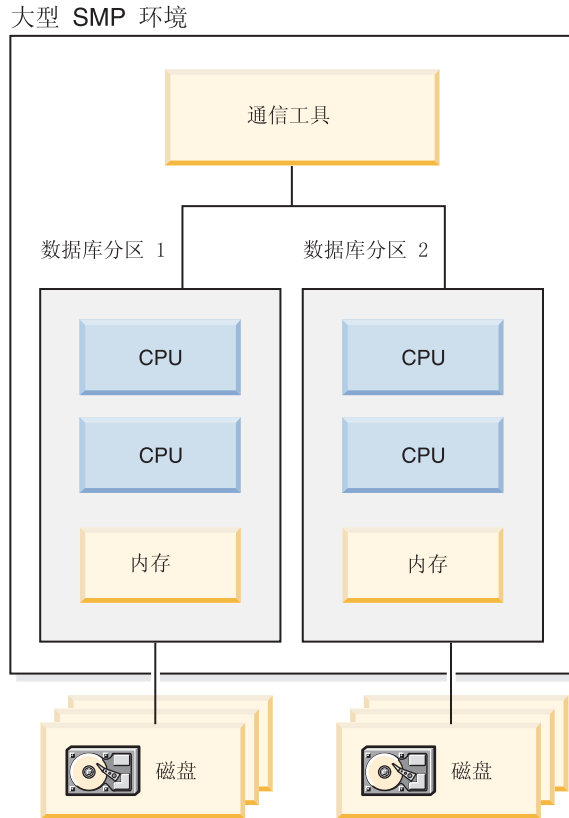


图 35. 带有对称多处理器环境的分区数据库

第 77 页的图 36 举例说明可以扩大图 35 中显示的配置以增强处理能力。

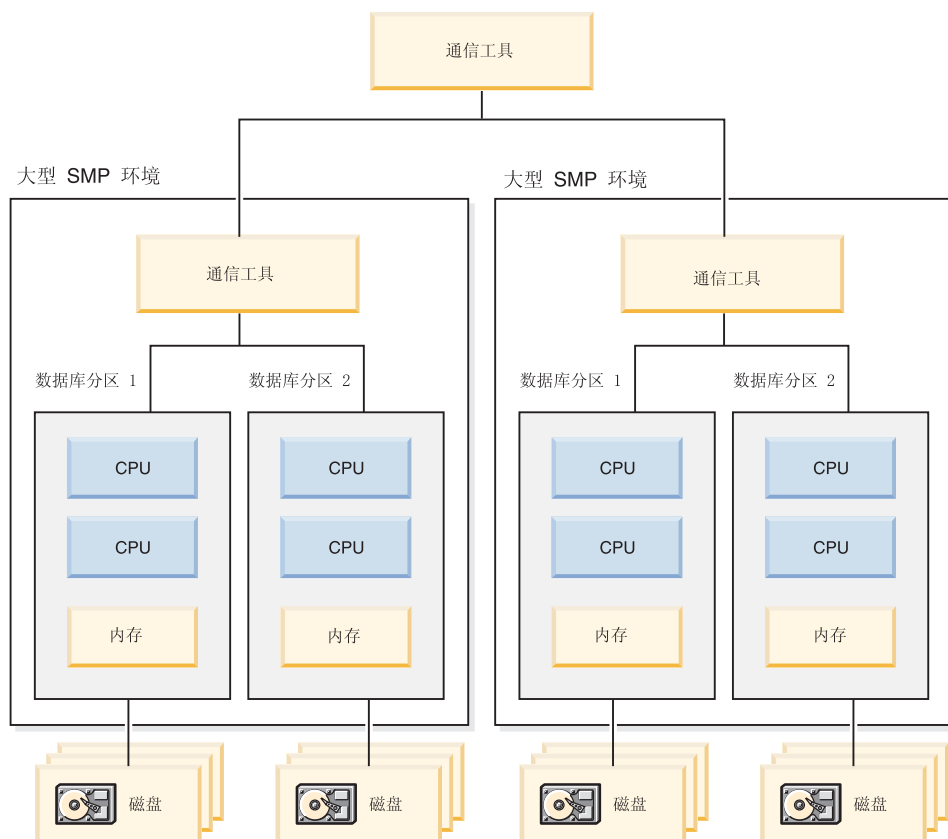


图 36. 带有集中在一起的对称多处理器环境的分区数据库

注：两个以上的数据库分区同时存在于同一台机器上的能力（不考虑处理器的数量），使得可以更灵活地设计高可用性配置和故障转移策略。机器发生故障之后，可以将一个数据库分区自动移至已包含同一数据库的另一个数据库分区的另一台机器上，然后重新启动该分区。

### 最适合每个硬件环境的并行性总结

下表总结了最适合于利用各种硬件环境的并行性的类型。

表 8. 每种硬件环境中的可能并行性类型

| 硬件环境                  | I/O 并行性 | 查询内并行性 |        |
|-----------------------|---------|--------|--------|
|                       |         | 分区内并行性 | 分区间并行性 |
| 单个数据库分区，单处理器          | 是       | 否 (1)  | 否      |
| 单个数据库分区，多处理器 (SMP)    | 是       | 是      | 否      |
| 多个数据库分区，单处理器 (MPP)    | 是       | 否 (1)  | 是      |
| 多个数据库分区，多处理器 (SMP 集群) | 是       | 是      | 是      |
| 逻辑数据库分区               | 是       | 是      | 是      |

表 8. 每种硬件环境中的可能并行性类型 (续)

| 硬件环境  | I/O 并行性 | 查询内并行性 |        |
|---|---------|--------|--------|
|   |         | 分区内并行性 | 分区间并行性 |
| 注: (1) 甚至是在单处理器系统中, 使用一个配置参数将并行度设置为大于 1 的值也可能获得好处, 特别是在执行的查询并未充分利用 CPU 时 (例如, 若它们受 I/O 约束)。 |         |        |        |

---

## 第 2 部分 安装注意事项



---

## 第 5 章 安装先决条件

---

### 安装 DB2 服务器 ( Windows )

此任务描述如何在 Windows 上启动“DB2 安装”向导。将使用“DB2 安装”向导来定义安装，并将 DB2 产品安装到系统上。

#### 先决条件

在启动“DB2 安装”向导之前:

- 如果您打算设置分区数据库环境，请参阅『设置分区数据库环境』主题。
- 确保系统满足安装、内存和磁盘要求。
- 如果打算在 Windows 上使用 LDAP 以在 Active Directory 中注册 DB2 服务器，那么在安装之前应扩展目录模式。
- 必须有带有建议的用户权限的本地管理员用户帐户才能执行安装。在 LocalSystem 可以用作 DAS 和 DB2 实例用户且您未使用数据库分区功能的 DB2 服务器中，具有提升物权的非管理员用户可以执行此安装。
- 建议您关闭所有的程序（虽然不是强制），这样安装程序可以更新计算机上的任意文件而不需要重新引导。

#### 限制

- DB2 副本名称和实例名不能以数字值开始。
- DB2 副本名称和实例名在所有的 DB2 副本中必须唯一。
- 只能对仅具有一个数据库分区的数据库使用 XML 功能。
- 如果已安装下列其中一项，那么其他 DB2 产品均不能安装在同一路径中：
  - IBM® 数据服务器运行时客户机
  - IBM 数据服务器 ODBC、CLI 和 .NET 驱动程序
  - DB2 信息中心。
- “DB2 安装”向导的字段不接受非英文字符。
- 如果在 Windows Vista 操作系统上启用扩展安全性，那么用户必须属于 DB2ADMNS 或 DB2USERS 组才能运行本地 DB2 命令和应用程序，这是因为有一个额外的安全性功能（用户访问控制）在缺省情况下会限制本地管理员所具有的特权。如果用户不属于其中任何一个组，那么他们将没有对本地 DB2 配置或应用程序数据的读访问权。

要启动“DB2 安装”向导:

1. 使用已为 DB2 安装定义的本地本地管理员帐户来登录系统。
2. 如果您具有 DB2 产品 DVD，那么将其插入驱动器中。如果启用了自动运行功能，它将自动启动“DB2 安装启动板”。如果不能自动运行，那么使用 Windows 资源管理器来浏览 DB2 产品 DVD 并双击安装图标以启动 DB2 安装启动板。
3. 如果您从 passport advantage 下载 DB2 产品，那么运行该可执行文件来解压缩 DB2 产品安装文件。使用 Windows 资源管理器来浏览 DB2 安装文件并双击安装图标来启动 DB2 安装启动板。

4. 从“DB2 安装启动板”中，可以查看安装先决条件和发行说明，也可以直接进行安装。您可能想查看安装先决条件和发行说明以获取最新信息。
5. 单击**安装产品**，**安装产品**窗口将显示可供安装的产品。

如果计算机上尚未安装任何 DB2 产品，那么通过单击**安装新产品**来启动安装。遵循“DB2 安装”向导的提示逐步完成安装。

如果计算机上至少已经安装了一个 DB2 产品，那么可以：

- 单击**安装新产品**以创建新的 DB2 副本。
  - 单击**使用现有产品**以升级现有的 DB2 副本，将功能添加到现有的 DB2 副本，迁移现有的 DB2 版本 8 或版本 9.1 副本，或者安装附件产品。
6. “DB2 安装”向导将确定系统语言，并启动该语言的安装程序。联机帮助可指导您完成其余步骤。要调用联机帮助，请单击**帮助**或按 **F1** 键。可随时单击**取消**来结束安装。

缺省情况下，DB2 产品将安装在 <Program Files>\IBM\sqllib 目录中，其中 <Program Files> 表示 Program Files 目录的位置。

如果您安装到一个该目录已经在使用的系统上，DB2 产品安装路径会添加 `_xx`，其中 `_xx` 是数字，从 01 开始并以您已安装的 DB2 副本数量递增。

您还可以指定自己的 DB2 产品安装路径。

有关安装期间遇到的错误的信息，请查看位于 My Documents\DB2LOG\ 目录中的安装日志文件。该日志文件使用以下格式：DB2-ProductAbbrrev-DateTime.log，例如 DB2-ESE-Tue Apr 04 17\_04\_45 2006.log。

如果想要 DB2 产品能够访问本地计算机或网络中的另一计算机上的 DB2 文档，那么必须安装 DB2 信息中心。“DB2 信息中心”包含 DB2 数据库系统和 DB2 相关产品的文档。缺省情况下，如果没有在本地安装 DB2 信息中心，那么可以从 Web 访问 DB2 信息。

### **DB2 易捷版和 DB2 工作组服务器版内存限制**

如果要安装 DB2 易捷版，那么允许用于实例的最大内存为 4GB。

如果要安装 DB2 工作组服务器版，那么允许用于实例的最大内存为 16GB。

分配给实例的内存量由 **INSTANCE\_MEMORY** 数据库管理器配置参数确定。

**从版本 9.1 迁移时的重要注释：**

- 如果版本 9.1 DB2 产品的内存配置超出允许的上限，那么在您迁移到当前版本之后，DB2产品可能无法启动。
- 自调整内存管理器不会将总的实例内存上限增大到超出许可证限制。

## **为分区 DB2 服务器准备环境 ( Windows )**

本主题描述了在准备 Windows 环境以便对 DB2 产品进行分区安装时需要执行的步骤。

参与的每台计算机都必须使用相同的操作系统。

要准备 Windows 环境以进行安装：



1. 确保主计算机和参与的计算机属于同一 Windows 域。使用可以通过“控制面板”访问的“系统属性”对话框来检查计算机所属的域。
2. 确保主计算机和参与的计算机上的时间与日期设置是一致的。要使时间和日期设置被认为是一致的，所有计算机之间的 GMT 时差一定不能超过一个小时。

使用可以通过“控制面板”访问的“日期/时间”属性对话框，可以修改系统日期与时间。可以使用 `max_time_diff` 配置参数来更改此限制。缺省值是 `max_time_diff = 60`，这允许时差小于 60 分钟。

3. 确保对每个参与分区数据库环境的计算机对象都标记了“信任计算机以委任”特权。您可以验证“Active Directory 用户和计算机”控制台中每台计算机的帐户“属性”对话框的“常规”选项卡上的“信任计算机以委任”复选框是否已选中。
4. 确保所有参与的计算机都可以使用 TCP/IP 相互通信：
  - a. 在一台参与的计算机上，输入 `hostname` 命令，它将返回该计算机的主机名。
  - b. 在另一台参与的计算机上，输入以下命令：

```
ping hostname
```

其中 `hostname` 表示主计算机的主机名。如果测试成功，您将接收到类似如下的输出：

```
Pinging ServerA.ibm.com [9.21.27.230] with 32 bytes of data:
```

```
Reply from 9.21.27.230: bytes=32 time<10ms TTL=128
Reply from 9.21.27.230: bytes=32 time<10ms TTL=128
Reply from 9.21.27.230: bytes=32 time<10ms TTL=128
```

重复执行这些步骤，直到确定所有参与的计算机都能使用 TCP/IP 相互进行通信为止。每台计算机必须具有静态 IP 地址。

如果正打算使用多个网络适配器，可以指定要使用哪个适配器在数据库分区服务器之间进行通信。安装完成后，使用 `db2nchg` 命令在 `db2nodes.cfg` 文件中指定“网络名”字段。

5. 安装期间，系统将要求您提供 DB2 管理服务器用户帐户。这是将由 DB2 管理服务器 (DAS) 使用的本地或域用户帐户。DAS 是用于支持 GUI 工具和协助完成管理任务的管理服务。现在可以定义一个用户，也可以让“DB2 安装”向导创建一个用户。如果想要使用“DB2 安装”向导创建新的域用户，用来执行安装的帐户必须具有创建域用户的权限。
6. 在将安装实例拥有的分区的主计算机上，必须具有属于本地 *Administrators* 组的域用户帐户。安装 DB2 时，您将作为此用户登录。必须将同一用户帐户添加至参与的每台计算机上的本地 *Administrators* 组。此用户必须具有以操作系统方式操作用户权限。
7. 确保实例中的所有计算机都具有同一个本地驱动器盘符上的数据库目录。您可以通过运行 `GET DATABASE CONFIGURATION` 命令并验证 `DFTDBPATH` DBM 配置参数值来检查此情况。
8. 安装期间，会要求您提供与 DB2 实例相关联的域用户帐户。每个 DB2 实例都指定了一个用户。当启动实例时，DB2 将使用此用户名进行登录。现在可以定义一个用户，也可以让“DB2 安装”向导创建新的域用户。

当添加新节点到分区环境中时，DB2 副本名称必须在所有的计算机上相同。

如果想要使用“DB2 安装”向导创建新的域用户，用来执行安装的帐户必须具有创建域用户的权限。实例用户域帐户在所有参与的计算机上必须属于本地 *Administrators* 组，并且将被授予下列用户权限：

- 以操作系统方式操作
- 创建标记对象
- 锁定内存中的页
- 作为服务登录
- 增加限额
- 替换进程级别标记

如果选择了扩展安全性，该帐户还必须是 DB2ADMNS 组的成员。DB2ADMNS 组已经具有这些特权，因此，不需要将这些特权显式添加至该帐户。

## 快速通信管理器 ( Windows )

快速通信管理器 (FCM) 为属于同一实例的 DB2 服务器产品提供了通信支持。每个数据库分区服务器都有一个 FCM 发送方守护程序和一个 FCM 接收方守护程序，用于在数据库分区服务器之间提供通信，以处理代理程序请求和传递消息缓冲区。FCM 守护程序在您启动实例时启动。

如果数据库分区服务器之间的通信失败，或者它们重新建立通信，那么 FCM 线程将更新信息（可以用数据库系统监视器进行查询），并导致执行适当的操作（例如，回滚受影响的事务）。可以使用数据库系统监视器来帮助您设置 FCM 配置参数。

可以使用 *fcm\_num\_buffers* 数据库管理器配置参数来指定 FCM 消息缓冲区数，以及使用 *fcm\_num\_channels* 数据库管理器配置参数来指定 FCM 通道数。数据库管理器配置参数 *fcm\_num\_buffers* 和 *fcm\_num\_channels* 设置为 AUTOMATIC（缺省值）。当这些参数中的任何设置为 AUTOMATIC 时，FCM 就会监视资源使用情况，并逐渐释放资源。建议将这些参数设置为 AUTOMATIC。

---

## 安装 DB2 服务器产品的概述 ( Linux 和 UNIX )

此主题概述了在 AIX®、HP-UX、Linux 和 Solaris 上安装 DB2 服务器产品的步骤。

要安装 DB2 服务器产品：

1. 查看 DB2 产品先决条件。
2. 查看 DB2 迁移信息（如果适用的话）。
3. 在 HP-UX、Linux 和 Solaris 上修改内核参数。在除 Linux on x86\_32 之外的所有平台上，您必须安装 64 位内核，然后才可以继续进行安装，否则安装将失败。
4. 准备安装介质：

### 产品 DVD

如果 DB2 产品 DVD 未自动安装，那么安装 DB2 产品 DVD。

### 安装映像

如果已下载安装映像，那么将该文件解压缩。

5. 使用其中一种可用方法来安装 DB2 产品：
  - “DB2 安装”向导

- db2\_install 命令
- 使用响应文件进行静默安装
- 有效内容文件部署

对于 DB2 服务器，您可以使用“DB2 安装”向导来执行安装和配置任务，例如：

- 选择 DB2 安装类型（典型、精简或定制）。
  - 选择 DB2 产品安装位置。
  - 安装您稍后可以指定的语言作为产品界面和消息的缺省语言。
  - 安装或升级 IBM Tivoli® System Automation for Multiplatforms Base Component（Linux 和 AIX）。
  - 设置 DB2 实例。
  - 设置 DB2 管理服务器（包括 DAS 用户设置）。
  - 设置 DB2 文本搜索服务器。
  - 设置管理联系人和运行状况监视器通知。
  - 设置和配置实例（包括实例用户设置）。
  - 设置 Informix 数据源支持。
  - 准备 DB2 工具目录。
  - 指定 DB2 信息中心端口。
  - 创建响应文件。
6. 如果使用除“DB2 安装”向导之外的其他方法安装 DB2 服务器，那么需要执行安装后配置步骤。

## DB2 安装方法

本主题提供有关 DB2 安装方法的信息。下表显示操作系统可用的安装方法。

表 9. 操作系统可用的安装方法

| 安装方法           | Windows | Linux 或 UNIX |
|----------------|---------|--------------|
| “DB2 安装”向导     | 是       | 是            |
| 响应文件安装         | 是       | 是            |
| db2_install 命令 | 否       | 是            |
| 有效内容文件部署       | 否       | 是            |

以下列表描述了 DB2 安装方法。

### “DB2 安装”向导

“DB2 安装”向导是可在 Linux、UNIX 和 Windows 操作系统上使用的一个 GUI 安装程序。“DB2 安装”向导提供了易于使用的界面，用于安装 DB2 产品和执行初始设置与配置任务。

“DB2 安装”向导还可以用来创建 DB2 实例和响应文件，它们可用于在其他机器上复制此安装。

**注：**对于 Linux 和 UNIX 平台上的非 root 用户安装，只能存在一个 DB2 实例。“DB2 安装”向导自动创建 root 用户实例。

在 Linux 和 UNIX 平台上，需要 X 服务器来显示“DB2 安装”向导。

## 响应文件安装

响应文件是一个包含设置和配置值的文本文件。DB2 安装程序将读取该文件，并根据已指定的值来执行安装。

响应文件安装也称为静默安装。

此安装方法的一个好处是：无需执行安装，即可创建响应文件。此功能在捕获安装 DB2 产品所需的选项时非常有用。根据您的指定的精确选项，响应文件在稍后可用于安装 DB2 产品。

响应文件的另一个优点是：它们提供了对那些不能使用“DB2 安装”向导设置的参数的访问。

在 Linux 和 UNIX 操作系统上，如果将 DB2 安装映像嵌入您自己的应用程序中，那么您的应用程序有可能从安装程序中以计算机可读的格式接收安装进度信息和提示。此行为由 INTERACTIVE 响应文件关键字控制。

可以采用下列方法来创建响应文件：

### 使用响应文件生成器（Windows 平台）

在 Windows 上，您可以使用响应文件生成器来创建一个复制现有安装的响应文件。例如，可以安装 IBM 数据服务器客户机、完整配置该客户机，然后生成响应文件以将该客户机的安装和配置复制到其他计算机。

### 使用“DB2 安装”向导

“DB2 安装”向导可以根据您在完成“DB2 安装”向导过程中所作的选择来创建响应文件。您的选择会记录在一个响应文件中，可以将该响应文件保存至系统上的某个位置。如果选择分区数据库安装，那么会生成两个响应文件，一个是为拥有实例的计算机生成的，一个是为参与的计算机生成的。

可以使用 `db2cfexp` 命令来导出客户机或服务器概要文件，以便保存客户机或服务器配置，而后就可以很容易地使用 `db2cfimp` 命令来导入概要文件。使用 `db2cfexp` 命令导出的客户机或服务器概要文件同样可以在使用 `CLIENT_IMPORT_PROFILE` 关键字的响应文件安装期间导入。

在执行安装和编目数据源之后，您应该导出客户机或服务器概要文件。

### 定制为每个 DB2 产品提供的样本响应文件

除了使用响应文件生成器或“DB2 安装”向导以外，还可以手工修改样本响应文件。样本响应文件是在 DB2 产品 DVD 上提供的。样本响应文件提供了关于每个产品的所有有效关键字的详细信息。

### db2\_install 命令（仅适用于 Linux 和 UNIX 平台）

`db2_install` 命令将安装您指定的具有“英语”界面支持的 DB2 产品的所有组件。通过使用 `-L` 参数就可以选择要支持的其他语言。您不能选择或者取消选择组件。

尽管 `db2_install` 命令会安装您指定的 DB2 产品的所有组件，但它不会执行用户和组创建、实例创建或配置。在安装之后执行配置时，此安装方法可能是首选。如果您宁愿在安装 DB2 产品时配置它，那么请考虑使用“DB2 安装”向导。

在 Linux 和 UNIX 操作系统上，如果将 DB2 安装映像嵌入您自己的应用程序中，那么您的应用程序有可能从安装程序中以计算机可读的格式接收安装进度信息和提示。

此安装方法在您部署产品文件之后需要手工配置。

#### 有效内容文件部署（仅适用于 Linux 和 UNIX）

此方法是一种高级安装方法，对于大多数用户，建议不要使用此方法。它要求用户以物理方式安装有效内容文件。有效内容文件是一个压缩的 tarball，它包含可安装的组件的所有文件和元数据。

此安装方法在您部署产品文件之后需要手工配置。

**注：**DB2 产品安装不再是 Linux 和 UNIX 平台上的操作系统程序包。因此，可以不再使用操作系统命令来进行安装。将需要更改您用来与 DB2 安装交互和查询 DB2 安装的任何现有脚本。

## 使用“DB2 安装”向导来安装 DB2 服务器（Linux 和 UNIX）

此任务描述如何在 Linux 和 UNIX 系统上启动“DB2 安装”向导。“DB2 安装”向导用来定义安装首选项以及将 DB2 产品安装到系统上。

在启动“DB2 安装”向导之前：

- 如果您打算设置分区数据库环境，那么请通过进入此主题底部的相关链接来了解如何执行此操作。
- 确保系统满足安装、内存和磁盘要求。
- 您可以使用 root 用户或非 root 用户权限来安装 DB2 服务器。有关非 root 用户安装的更多信息，请参阅相关链接。
- DB2 产品映像必须可用。通过购买物理 DB2 产品 DVD 或从 Passport Advantage 下载安装映像，您可以获取 DB2 安装映像。
- 如果要安装非英文版本的 DB2 数据库产品，那么必须具有适当的本地语言程序包。
- “DB2 安装”向导是一个图形安装程序。必须具有能够提供图形用户界面的 X Windows 软件，才能使“DB2 安装”向导在机器上运行。确保 X windows 服务器正在运行。确保正确导出了显示内容。例如，`export DISPLAY=9.26.163.144:0`。
- 如果在您的环境中使用了安全性软件，那么在启动“DB2 安装”向导之前必须手工创建所需的 DB2 用户。

**注：**

- 只能对使用代码集 UTF-8 定义的并且只有一个数据库分区的数据库使用 XML 功能。
- “DB2 安装”向导的字段不接受非英文字符。

要启动“DB2 安装”向导：

1. 如果具有物理 DB2 产品 DVD，那么通过输入下列命令来切换至安装 DB2 产品 DVD 的目录：

```
cd /dvdrom
```

其中 `/dvdrom` 表示 DB2 产品 DVD 的安装点。

2. 如果您下载了 DB2 产品映像，那么必须对产品文件进行解压缩和解压。
  - a. 解压缩产品文件：

```
gzip -d product.tar.gz
```

其中 *product* 是下载的产品的名称。

- b. 解压产品文件:

在 **Linux** 操作系统上

```
tar -xvf product.tar
```

在 **AIX、HP-UX 和 Solaris** 操作系统上

```
guntar -xvf product.tar
```

其中 *product* 是下载的产品的名称。

- c. 更改目录:

```
cd ./product
```

其中 *product* 是下载的产品的名称。

**注:** 如果下载了本地语言程序包, 那么将其解压缩至同一个目录中。这样就会在同一目录中创建子目录 (例如, *./nlpack/disk1*), 并且使安装程序能够自动查找安装映像而不需要提示。

3. 通过从产品映像所在目录中输入 *./db2setup* 命令来启动“DB2 安装”向导。
4. “IBM DB2 安装启动板”将打开。在此窗口中, 可以查看安装先决条件和发行说明, 也可以直接进行安装。您可能想查看安装先决条件和发行说明以获取最新信息。
5. 单击**安装产品**, **安装产品**窗口将显示可供安装的产品。

通过单击“**安装新产品**”启动安装。遵循“DB2 安装”向导的提示逐步完成安装。

一旦启动了安装, 就通过“DB2 安装”向导的安装面板来继续进行安装并作出选择。安装帮助可用来指导您完成其余步骤。要调用安装帮助, 单击**帮助**或按 F1 键。可随时单击**取消**来结束安装。

对于非 root 用户安装, DB2 产品始终安装在 *\$HOME/sqllib* 目录中, 其中 *\$HOME* 表示非 root 用户的主目录。

对于 root 用户安装, DB2 产品在缺省情况下安装在下列其中一个目录中:

**AIX、HP-UX 和 Solaris**

```
/opt/IBM/db2/V9.5
```

**Linux** /opt/ibm/db2/V9.5

如果您安装到一个该目录已经在使用的系统上, DB2 产品安装路径会添加 *\_xx*, 其中 *\_xx* 是数字, 从 01 开始并以您已安装的 DB2 副本数量递增。

您还可以指定自己的 DB2 产品安装路径。

DB2 安装路径具有下列规则:

- 可以包含小写字母 (a-z)、大写字母 (A-Z) 和下划线字符 ( \_ )
- 不能超过 128 个字符
- 不能包含空格
- 不能包含非英文字符

在安装了 DB2 数据库产品之后，通过从本地语言包所在的目录中运行 `./db2setup` 命令，也可以安装本地语言包。

安装日志文件是：

- DB2 安装日志文件。此文件将捕获包括错误在内的所有 DB2 安装信息。
  - 对于 root 用户安装，DB2 安装日志文件名是 `db2setup.log`。
  - 对于非 root 用户安装，DB2 安装日志文件名是 `db2setup_username.log`，其中 `username` 是用于执行安装的非 root 用户标识。
- DB2 错误日志文件。此文件会捕获由 Java™ 返回的任何错误输出（例如，异常和陷阱信息）。
  - 对于 root 用户安装，DB2 错误日志文件名是 `db2setup.err`。
  - 对于非 root 用户安装，DB2 错误日志文件名是 `db2setup_username.err`，其中 `username` 是用于执行安装的非 root 用户标识。

缺省情况下，这些日志文件位于 `/tmp` 目录中。可以指定日志文件的位置。

不再存在 `db2setup.his` 文件。但是，DB2 安装程序会将 DB2 安装日志文件的副本保存在 `DB2_DIR/install/logs/` 目录中，并将该文件重命名为 `db2install.history`。如果该名称已存在，那么 DB2 安装程序会将它重命名为 `db2install.history.xxxx`，其中 `xxxx` 的范围是 0000-9999，这取决于机器上具有的安装数。

每个安装副本都有一个单独的历史记录文件列表。如果除去了一个安装副本，那么此安装路径下的历史记录文件也将被除去。此复制操作是在安装快要结束时执行的，如果程序在完成之前就已停止或者异常中止，那么不会创建历史记录文件。

在 Linux x86 上，如果想要 DB2 产品能够访问本地计算机或网络上的另一计算机上的 DB2 文档，那么必须安装 DB2 信息中心。“DB2 信息中心”包含 DB2 数据库系统和 DB2 相关产品的文档。

### DB2 易捷版和 DB2 工作组服务器版内存限制

如果要安装 DB2 易捷版，那么允许用于实例的最大内存为 4GB。

如果要安装 DB2 工作组服务器版，那么允许用于实例的最大内存为 16GB。

分配给实例的内存量由 **INSTANCE\_MEMORY** 数据库管理器配置参数确定。

从版本 9.1 迁移时的重要注释：

- 如果版本 9.1 DB2 产品的内存配置超出允许的上限，那么在您迁移到当前版本之后，DB2 产品可能无法启动。
- 自调整内存管理器不会将总的实例内存上限增大到超出许可证限制。

### 快速通信管理器 (Linux 和 UNIX)

快速通信管理器 (FCM) 为使用“数据库分区功能” (DPF) 的 DB2 服务器产品提供了通信支持。

对于多分区实例，每个数据库分区服务器都有一个 FCM 发送方守护程序和一个 FCM 接收方守护程序，用于在数据库分区服务器之间提供通信，以处理代理程序请求和传递消息缓冲区。FCM 守护程序在您启动多分区实例时启动。

如果数据库分区服务器之间的通信失败，或者它们重新建立通信，那么 FCM 守护程序将更新信息。您可以使用数据库系统监视器查询此信息。FCM 守护程序还会触发相应的操作。例如回滚受到影响的事务。可以使用数据库系统监视器来帮助您设置 FCM 配置参数。

可以使用 *fcm\_num\_buffers* 数据库管理器配置参数来指定 FCM 消息缓冲区数。此外，还可以使用 *fcm\_num\_channels* 数据库管理器配置参数来指定 FCM 通道数。数据库管理器配置参数 *fcm\_num\_buffers* 和 *fcm\_num\_channels* 设置为 AUTOMATIC（缺省值）。当任何这些参数设置为 AUTOMATIC 时，FCM 就会监视资源使用情况，并以递增方式释放资源。建议将这些参数设置为 AUTOMATIC。



---

## 第 6 章 安装之前

---

### 其他分区数据库环境预安装任务 (Linux 和 UNIX)

#### 更新用于分区 DB2 安装的环境设置 (AIX)

此任务描述了将参与分区数据库系统的每台计算机上需要更新的环境设置。

要更新 AIX 环境设置:

1. 作为具有 root 用户权限的用户登录计算机。
2. 通过输入以下命令, 将 AIX maxuproc (每个用户的最大进程数) 设备属性设置为 4096:

```
chdev -l sys0 -a maxuproc='4096'
```

注: 如果另一映像正在运行, 那么可能需要将 bosboot/reboot 切换为 64 位内核。

3. 在参与分区数据库系统的所有工作站上, 将 TCP/IP 网络参数设置为下列值。这些值都是这些参数的最小值。如果任何网络相关参数都已设置为较高的值, 那么不要更改它。

```
thewall      = 65536
sb_max       = 1310720
rfc1323      = 1
tcp_sendspace = 221184
tcp_recvspace = 221184
udp_sendspace = 65536
udp_recvspace = 65536
ipqmaxlen    = 250
somaxconn    = 1024
```

要列示所有网络相关参数的当前设置, 输入以下命令:

```
no -a | more
```

要设置参数, 输入以下命令:

```
no -o parameter_name=value
```

其中:

- *parameter\_name* 表示想要设置的参数。
- *value* 表示想要对此参数设置的值。

例如, 要将 tcp\_sendspace 参数设置为 221184, 输入以下命令:

```
no -o tcp_sendspace=221184
```

4. 如果正在使用高速互连, 那么必须将 css0 的 spoolsize 和 rpoolsize 设置为下列值:

```
spoolsize    16777216
rpoolsize    16777216
```

要列示这些参数的当前设置, 输入以下命令:

```
lsattr -l css0 -E
```

要设置这些参数, 输入下列命令:

```
/usr/lpp/ssp/css/chgcss -l css0 -a spoolsize=16777216
/usr/lpp/ssp/css/chgcss -l css0 -a rpoolsize=16777216
```

如果没有使用 `/tftpboot/tuning.cst` 文件来调整系统，那么在安装之后可以使用 `DB2DIR/misc/rc.local.sample` 样本脚本文件，其中 `DB2DIR` 是 `DB2` 产品安装的路径，来更新网络相关参数。要在安装之后使用样本脚本文件来更新网络相关参数，执行下列步骤：

- a. 通过输入下列命令，将此脚本文件复制到 `/etc` 目录，并使它可由 `root` 用户执行：

```
cp /usr/opt/db2_09_01/misc/rc.local.sample /etc/rc.local
chown root:sys /etc/rc.local
chmod 744 /etc/rc.local
```

- b. 查看 `/etc/rc.local` 文件并对其进行更新（如果有必要的话）。
- c. 向 `/etc/inittab` 文件添加一个条目，以便每当机器重新引导时执行 `/etc/rc.local` 脚本。可使用 `mkitab` 命令来向 `/etc/inittab` 文件添加一个条目。要添加此条目，输入以下命令：

```
mkitab "rclocal:2:wait:/etc/rc.local > /dev/console 2>&1"
```

- d. 通过输入以下命令，确保 `/etc/inittab` 文件包括 `/etc/rc.nfs` 条目：

```
lsitab rcnfs
```

- e. 通过输入以下命令，在不重新引导系统的情况下更新网络参数：

```
/etc/rc.local
```

5. 确保具有足够的调页空间来运行 `DB2 ESE` 的分区安装。如果没有足够的调页空间，那么操作系统将停止正在使用大部分虚拟内存（这可能是其中一个 `DB2` 进程）的进程。要检查可用的调页空间，输入以下命令：

```
lsps -a
```

此命令将返回类似于以下的输出：

| Page Space | Physical Volume | Volume Group | Size | %Used | Active | Auto | Type |
|------------|-----------------|--------------|------|-------|--------|------|------|
| paging00   | hdisk1          | rootvg       | 60MB | 19    | yes    | yes  | lv   |
| hd6        | hdisk0          | rootvg       | 60MB | 21    | yes    | yes  | lv   |
| hd6        | hdisk2          | rootvg       | 64MB | 21    | yes    | yes  | lv   |

可用调页空间应是在计算机上安装的物理内存量的两倍。

6. 如果正在创建小型到中型的分区数据库系统，那么拥有实例的计算机上的网络文件系统守护程序（`NFSD`）的数目应该接近于：

每台计算机上的 `biod` 数 × 实例中的计算机数

理想情况下，应该在每台计算机上运行 10 个 `biod` 进程。根据以上公式，在由四台计算机组成的系统（每台计算机上有 10 个 `biod` 进程）上，将使用 40 个 `NFSD`。

如果正在安装大型系统，那么计算机上最多可以安装 120 个 `NFSD`。

有关 `NFS` 的其他信息，参阅 `NFS` 文档。

## 建立工作集合以将命令分发至 **ESE** 工作站（**AIX**）

在 `AIX` 上的分区数据库环境中，可以设置工作集合以将命令分发至参与分区数据库系统的一组 `RS/6000 SP™` 工作站。使用 `dsh` 命令可以将命令分发至工作站。

在 AIX 上安装或管理分区数据库系统时这可能会很有用，它使您能够对环境中的所有计算机快速地执行相同命令并且减少了出错的可能性。

必须知道想要包括在工作集中的每台计算机的主机名。

必须作为具有 root 用户权限的用户登录控制工作站。

具有列示所有将要参与分区数据库系统的 RS/6000 SP 工作站的主机名的文件。要设置工作集合以将命令分发至此工作站列表：

1. 创建名为 `eeelist.txt` 的文件，它将列示将参与工作集合的所有工作站的 *hostnames*。

例如，假定您想要借助名为 `workstation1` 和 `workstation2` 的两个 SP 节点创建工作集合。`eeelist.txt` 的内容将是：

```
workstation1
workstation2
```

2. 更新工作集合环境变量。要更新此列表，输入以下命令：

```
export WCOLL=path/eeelist.txt
```

其中 *path* 是创建 `eeelist.txt` 的位置，而 `eeelist.txt` 是创建的文件名称，该文件列示工作集中的 RS/6000SP 工作站。

3. 通过输入以下命令，验证工作集合中的名称是否确实是想要的工作站：

```
dsh -q
```

您将接收到与下列内容类似的输出：

```
Working collective file /eeelist.txt:
workstation1
workstation2
Fanout: 64
```

## 验证 NFS 是否在运行 (Linux 和 UNIX)

在设置数据库分区环境之前，应验证“网络文件系统”（NFS）在将参与分区数据库系统的每台计算机上是否正在运行。

NFS 必须在每台计算机上都正在运行。

要验证 NFS 是否在每台计算机上都正在运行：

### AIX 操作系统

在每台计算机上输入以下命令：

```
lssrc -g nfs
```

NFS 进程的状态字段应该指示活动。

验证了 NFS 是否在每个系统上都正在运行之后，应检查 DB2 必需的特定 NFS 进程。必需的进程有：

```
rpc.lockd
rpc.statd
```

### HP-UX 和 Solaris 操作系统

在每台计算机上输入以下命令：

```
showmount -e hostname
```

输入不带 *hostname* 参数的 `showmount` 命令来检查本地系统。

如果 NFS 不活动，那么您将接收到类似如下的一条消息：

```
showmount: ServerA: RPC: Program not registered
```

验证了 NFS 是否在每个系统上都正在运行之后，应检查 DB2 必需的特定 NFS 进程：

```
rpc.lockd  
rpc.statd
```

可以使用下列命令来检查这些进程：

```
ps -ef | grep rpc.lockd  
ps -ef | grep rpc.statd
```

## Linux 操作系统

在每台计算机上输入以下命令：

```
showmount -e hostname
```

输入不带 *hostname* 参数的 `showmount` 命令来检查本地系统。

如果 NFS 不活动，那么您将接收到类似如下的一条消息：

```
showmount: ServerA: RPC: Program not registered
```

验证了 NFS 是否在每个系统上都正在运行之后，应检查 DB2 必需的特定 NFS 进程。必需的进程是 `rpc.statd`。

可以使用 `ps -ef | grep rpc.statd` 命令来检查此进程。

如果这些进程没有运行，那么参阅您的操作系统文档。

## 验证参与的计算机上的可用端口范围（Linux 和 UNIX）

此任务描述了验证参与的计算机上的可用端口范围时需要执行的步骤。端口范围由“快速通信管理器”（FCM）使用。FCM 是 DB2 用来处理数据库分区服务器之间的通信的功能部件。

应该在安装实例拥有的数据库分区服务器之后，并在安装任何参与的数据库分区服务器之前验证参与的计算机上的可用端口范围。

当在主计算机上安装实例拥有的数据库分区服务器时，DB2 根据指定的参与分区数据库环境的逻辑数据库分区服务器数目保留端口范围。缺省范围为四个端口。对于每个参与分区数据库环境的服务器，您必须为 FCM 端口手工配置 `/etc/services` 文件。FCM 端口范围取决于要在参与的计算机上使用的逻辑分区数目。至少需要两个条目：**DB2\_<instance>** 和 **DB2\_<instance>\_END**。在参与的计算机上指定的 FCM 端口的其他要求：

- 起始端口号必须与主计算机的起始端口号相匹配
- 后续端口必须按顺序进行编号
- 指定的端口号必须是空闲的
- 

要更改 `services` 文件，需要具有 `root` 用户权限。

要验证参与的计算机上的可用端口范围:

1. 打开位于 `/etc/services` 目录中的 `services` 文件。
2. 查找为 DB2 的“快速通信管理器”(FCM)保留的端口。这些条目应类似如下:

```
DB2_db2inst1      60000/tcp
  DB2_db2inst1_1  60001/tcp
  DB2_db2inst1_2  60002/tcp
  DB2_db2inst1_END 60003/tcp
```

DB2 将保留 60000 之后的前四个可用端口。

3. 在参与的每台计算机上,打开 `services` 文件,并验证在主计算机的 `services` 文件中为 DB2 FCM 保留的端口是否未使用。
4. 如果参与的某台计算机正在使用需要的端口,那么应确定所有计算机的可用端口范围,并更新每个 `services` 文件(包括主计算机上的 `services` 文件)。

在主计算机上安装实例拥有的数据库分区服务器之后,必须在参与的数据库分区服务器上安装 DB2 产品。您可以使用为分区服务器生成的响应文件(缺省名称为 `db2ese_addpart.rsp`),需要为 FCM 端口手工配置 `/etc/services` 文件。FCM 端口范围取决于要在当前机器上使用的逻辑分区数目。最小条目为 `DB2_` 和 `DB2__END` 这两个具有连续可用端口号的条目。在参与的每台机器上使用的 FCM 端口号均必须具有相同的起始端口号,并且后续端口必须按顺序进行编号。

## 为分区 DB2 服务器创建文件系统 (Linux)

此任务是设置分区数据库系统这一任务的一部分。此任务描述如何完成下列操作:

- 创建 DB2 主文件系统
- NFS 导出主文件系统
- NFS 在参与的每台计算机中都安装主文件系统

您的文件系统必须可用于将参与分区数据库系统的所有机器。此文件系统将用作实例主目录。

对于对单个数据库实例使用多台机器的配置,使用 NFS(网络文件系统)来共享此文件系统。通常,集群中的一台机器用于使用 NFS 导出文件系统,而集群中的其余机器将从此机器安装 NFS 文件系统。对于导出文件系统的机器,将以本地方式安装该文件系统。

有关更多命令信息,请参阅 Linux 分发文档。

要创建此文件系统:

1. 在一台机器上,选择一个磁盘分区或使用 `fdisk` 来创建一个磁盘分区。
2. 通过使用诸如 `mkfs` 之类的实用程序,在此分区上创建文件系统。该文件系统应该足够大以包含必需的 DB2 程序文件以及数据库需要的足够空间。
3. 以本地方式安装您刚刚创建的文件系统,并向 `/etc/fstab` 文件添加一个条目,以便系统每次重新引导时都安装此文件系统。例如:

```
/dev/hda1 /db2home ext3 defaults 1 2
```

4. 要在 Linux 上在引导时自动导出 NFS 文件系统,应向 `/etc/exports` 文件添加一个条目。务必包括参与集群的所有主机名以及机器可能具有的所有名称。并且,还应该通过使用“root”选项来确保集群中每台机器对导出的文件系统都具有 root 用户权限。

/etc/exports 文件是包含以下类型的信息的 ASCII 文件:

```
/db2home machine1_name(rw) machine2_name(rw)
```

要导出 NFS 目录, 运行

```
/usr/sbin/exports -r
```

5. 在集群中的其余每台机器上, 向 /etc/fstab 文件添加一个条目, 以便在引导时以 NFS 方式自动安装该文件系统。如下例所示, 当指定安装点选项时, 确保该文件系统是可读写的并是在引导时硬安装的, 它包括了 bg (后台) 选项, 且 setuid 程序可以正常运行。

```
fusion-en:/db2home /db2home nfs rw,timeo=7,  
hard,intr,bg,suid,lock
```

其中, *fusion-en* 表示机器名。

6. 通过输入以下命令, 以 NFS 方式在集群中的其余每台机器上安装导出的文件系统:

```
mount /db2home
```

如果安装命令失败, 那么使用 showmount 命令来检查 NFS 服务器的状态。例如:

```
showmount -e fusion-en
```

其中, *fusion-en* 表示机器名。

此 showmount 命令应列示从机器 *fusion-en* 导出的文件系统。如果此命令失败, 那么 NFS 服务器可能尚未启动。要手工启动 NFS 服务器, 在该服务器上作为 root 用户运行以下命令:

```
/etc/rc.d/init.d/nfs restart
```

假定目前的运行级别为 3, 则可以通过将目录 /etc/rc.d/rc3.d 下的 K20nfs 重命名为 S20nfs, 以便在引导时自动运行此命令。

7. 确保下列步骤成功完成:
  - a. 在集群中的单一机器上, 已经创建了要用作实例和主目录的文件系统。
  - b. 如果配置对单个数据库实例使用多台机器, 那么通过使用 NFS 导出此文件系统。
  - c. 已在集群中的其余每台机器上安装了导出的文件系统。

## 创建分区数据库系统 (AIX) 的 DB2 本地文件系统

此任务是设置分区数据库系统这一任务的一部分。此任务描述如何完成下列操作:

- 创建 DB2 主文件系统
- NFS 导出主文件系统
- NFS 在参与的每台计算机中都安装主文件系统

建议创建一个大小与 DB2 产品 DVD 上的内容相同的主文件系统。可以使用下列命令来检查大小 (以 KB 为单位显示):

```
du -sk <DVD mounting point>
```

一个 DB2 实例至少需要 50 MB 空间。如果没有足够的可用空间, 那么可以从参与的每台计算机安装 DB2 产品 DVD, 这是将内容复制到磁盘的另一种方法。

您必须具有:

- 创建文件系统的 root 用户权限
- 创建了要将文件系统真正放置于的卷组。

要创建以 NFS 方式导出并以 NFS 方式安装的 DB2 主文件系统，执行下列步骤：

### 创建 DB2 主文件系统

作为具有 root 用户权限的用户登录分区数据库系统的主计算机（ServerA），并为分区数据库系统创建名为 /db2home 的主文件系统。

1. 输入 **smit jfs** 命令。
2. 单击添加日志文件系统图标。
3. 单击添加标准日志文件系统图标。
4. 从卷组名列表中选择想要将此文件系统真正放置于的现有卷组。
5. 设置文件系统的大小（文件系统大小（以 512 字节块计）（数字）字段）。此大小按 512 字节块计数缩放，如果只需要为实例主目录创建文件系统，那么可以使用 180 000，大约 90 MB。如果需要完整地复制产品 DVD 映像以运行安装，那么可以创建它为值 20 000 000，大约 1 GB。
6. 在安装点字段中输入此文件系统的安装点。在此示例中，安装点为 /db2home。
7. 将在系统重新启动时自动安装字段设置为是。

其余字段可保留缺省设置。

8. 单击**确定**。

### 导出 DB2 主文件系统

1. 以 NFS 方式导出 /db2home 文件系统，使它可用于将参与分区数据库系统的所有计算机：
  - a. 输入 **smit nfs** 命令。
  - b. 单击**网络文件系统（NFS）**图标。
  - c. 单击**将目录添加至导出列表**图标。
  - d. 在**要导出的目录的路径名字段**中，输入要导出的路径名和目录（例如，/db2home）。
  - e. 在**允许具有 root 用户访问权的主机**字段中输入将要参与分区数据库系统的每个工作站的名称。使用逗号（,）作为每个名称之间的定界符。例如，ServerA, ServerB, ServerC。如果是在使用高速互连，我们建议您在此字段中还要指定每个工作站的高速互连名称。其余字段可保留缺省设置。
  - f. 单击**确定**。
2. 注销。

### 从参与的每台计算机安装 DB2 主文件系统

登录参与的每台计算机（ServerB、ServerC 和 ServerD），通过执行下列步骤来以 NFS 方式安装已导出的文件系统：

1. 输入 **smit nfs** 命令。
2. 单击**网络文件系统（NFS）**图标。
3. 单击**添加文件系统以进行安装**图标。

4. 在**安装点的路径名（路径）**字段中输入安装点的路径名。

安装点的路径名就是应创建 DB2 主目录的位置。对于此示例，使用 /db2home。

5. 在**远程目录的路径名字段**中输入远程目录的路径名。

对于本示例，应输入您在**安装点的路径名（路径）**字段中输入的值。

6. 在**远程目录所在的主机**字段中输入导出了文件系统的机器的主机名。

此值是创建了要安装的文件系统的机器的主机名。

为了提高性能，您可能想通过高速互连以 NFS 方式安装您创建的文件系统。如果想要使用高速互连来安装此文件系统，那么必须在**远程目录所在的主机**字段中输入其名称。

您应该知道，如果高速互连由于某种原因而变得不可用，那么参与分区数据库系统的每个工作站都将无法访问此 DB2 主目录。

7. 将**立即安装**，将条目添加至 **/etc/filesystems** 还是**全部**字段设置为全部。
8. 将 **/etc/filesystems** 条目将在系统重新启动时**安装目录**字段设置为是。
9. 将此 **NFS** 文件系统的方式字段设置为读写。
10. 将对文件系统**进行软安装还是硬安装**字段设置为软安装。

软安装表示计算机将不会无限期地尝试以远程方式安装该目录。硬安装表示机器将不停地尝试安装该目录。这可能会在系统崩溃时产生问题。建议将此字段设置为软安装。

其余字段可保留缺省设置。

11. 确保此文件系统是在**是否允许在此文件系统中执行 SUID 和 sgid 程序**字段设置为是的情况下安装的。这是缺省设置。
12. 单击**确定**。
13. 注销。

## 在分区数据库环境中为 DB2 服务器安装创建必需的用户（Linux）

运行 DB2 数据库需要三个用户和组。下表说明在下列指令中使用的用户名和组名。可以指定您自己的用户名和组名，但是它们必须符合系统命名规则和 DB2 命名规则。

如果您打算使用“DB2 安装”向导来安装 DB2 产品，那么“DB2 安装”向导将创建这些用户。

表 10. 必需的用户和组

| 必需的用户       | 用户名      | 组名       |
|-------------|----------|----------|
| 实例所有者       | db2inst1 | db2iadm1 |
| 受防护的用户      | db2fenc1 | db2fadm1 |
| DB2 管理服务器用户 | dasusr1  | dasadm1  |

如果 DB2 管理服务器用户是现有用户，那么此用户必须存在于所有参与的计算机上，然后才能安装。如果使用“DB2 安装”向导在拥有实例的计算机上创建了新的 DB2 管理



服务器用户，那么在安装响应文件期间也会在参与的计算机上创建该新用户（如果必要的话）。如果该用户在参与的计算机上已存在，那么它必须具有相同的主组。

### 先决条件

- 必须具有 root 用户权限才能创建用户和组。
- 如果使用安全性软件来管理用户和组，那么在定义 DB2 用户和组时可能还需要执行其他步骤。

**限制** 您创建的用户名必须同时符合操作系统命名规则和 DB2 命名规则。

要创建全部这三个用户，执行下列步骤：

1. 登录主计算机。
2. 通过输入下列命令，为实例所有者创建一个组（例如，db2iadm1），为将要执行 UDF 或存储过程的组创建一个组（例如，db2fadm1），并为将拥有 DB2 管理服务器的组创建一个组（例如，dasadm1）：

```
groupadd -g 999 db2iadm1
groupadd -g 998 db2fadm1
groupadd -g 997 dasadm1
```

确保正在使用的特定号码当前不存在于任何机器上。

3. 通过使用下列命令，为前一步骤中创建的每个组创建一个用户。每个用户的主目录将是您先前创建且共享的 DB2 主目录（db2home）。

```
useradd -u 1004 -g db2iadm1 -m -d /db2home/db2inst1 db2inst1 -p password1
useradd -u 1003 -g db2fadm1 -m -d /db2home/db2fenc1 db2fenc1 -p password2
useradd -u 1002 -g dasadm1 -m -d /home/dasusr1 dasusr1 -p password3
```

4. 通过输入下列命令，为创建的每个用户设置初始密码：  

```
passwd db2inst1    passwd db2fenc1    passwd dasusr1
```
5. 注销。
6. 作为已创建的每个用户（db2inst1、db2fenc1 和 dasusr1）登录主计算机。因为这是这些用户第一次登录系统，所以可能会提示您更改每个用户的密码。
7. 注销。
8. 在将参与分区数据库环境的每台计算机上创建完全相同的用户和组帐户。

## 在分区数据库环境中为 DB2 服务器安装创建必需的用户（AIX）

运行 DB2 数据库需要三个用户和组。下表说明在下列指令中使用的用户名和组名。可以指定您自己的用户名和组名，但是它们必须符合系统命名规则和 DB2 命名规则。

如果您打算使用“DB2 安装”向导来安装 DB2 产品，那么“DB2 安装”向导将创建这些用户。

表 11. 必需的用户和组

| 必需的用户       | 用户名      | 组名       |
|-------------|----------|----------|
| 实例所有者       | db2inst1 | db2iadm1 |
| 受防护的用户      | db2fenc1 | db2fadm1 |
| DB2 管理服务器用户 | dasusr1  | dasadm1  |

如果 DB2 管理服务器用户是现有用户，那么此用户必须存在于所有参与的计算机上，然后才能安装。如果使用“DB2 安装”向导在拥有实例的计算机上创建了新的 DB2 管理服务器用户，那么在安装响应文件期间也会在参与的计算机上创建该新用户（如果必要的话）。如果该用户在参与的计算机上已存在，那么它必须具有相同的主组。

#### 先决条件

- 必须具有 root 用户权限才能创建用户和组。
- 如果使用安全性软件来管理用户和组，那么在定义 DB2 用户和组时可能还需要执行其他步骤。

**限制** 您创建的用户名必须同时符合操作系统命名规则和 DB2 命名规则。

要创建全部这三个用户，执行下列步骤：

1. 登录主计算机。
2. 通过输入下列命令，为实例所有者创建一个组（例如，db2iadm1），为将要执行 UDF 或存储过程的组创建一个组（例如，db2fadm1），并为将拥有 DB2 管理服务器的组创建一个组（例如，dasadm1）：

```
mkgroup id=999 db2iadm1
mkgroup id=998 db2fadm1
mkgroup id=997 dasadm1
```

3. 通过使用下列命令，为前一步骤中创建的每个组创建一个用户。每个用户的主目录将是您先前创建且共享的 DB2 主目录（db2home）。

```
mkuser id=1004 pgrp=db2iadm1 groups=db2iadm1 home=/db2home/db2inst1
core=-1 data=491519 stack=32767 rss=-1 fsize=-1 db2inst1
mkuser id=1003 pgrp=db2fadm1 groups=db2fadm1 home=/db2home/db2fenc1
db2fenc1
mkuser id=1002 pgrp=dasadm1 groups=dasadm1 home=/home/dasusr1
dasusr1
```

4. 通过输入下列命令，为创建的每个用户设置初始密码：

```
passwd db2inst1
passwd db2fenc1
passwd dasusr1
```

5. 注销。
6. 作为已创建的每个用户（db2inst1、db2fenc1 和 dasusr1）登录主计算机。因为这是这些用户第一次登录系统，所以可能会提示您更改每个用户的密码。
7. 注销。
8. 在将参与分区数据库环境的每台计算机上创建完全相同的用户和组帐户。

---

## 第 7 章 安装 DB2 服务器产品

---

### 设置分区数据库环境

本主题描述如何使用数据库分区功能 (DPF) 来设置分区数据库环境。将使用“DB2 安装”向导来安装实例拥有的数据库服务器和创建响应文件，然后再使用这些响应文件来创建参与的数据库服务器。

如果您购买了 DB2 企业服务器版 (ESE)，那么可以创建分区实例。此选项称为 DPF。

**注：**DPF 在非 root 用户安装中不受支持。

数据库分区是数据库的一部分，它有自己的数据、索引、配置文件和事务日志组成。分区数据库是有两个或更多个分区的数据库。

要利用 ESE 的 DPF，除 ESE 许可证密钥以外，您还必须具有 DPF 许可证密钥。如果要在安装期间自动添加许可证密钥，那么需要在启动“DB2 安装”向导之前将许可证密钥复制到安装映像的 db2 目录的 license 子目录。

#### 先决条件

- 将要参与 DPF 环境的每台计算机上相同数量的连续端口必须空闲。例如，如果 DPF 环境将由四台计算机组成，那么这些计算机上相同的四个连续端口都必须空闲。在创建实例期间，将在下列文件中保留端口，而且其数目与当前服务器上的逻辑分区数目相等：在 Linux 和 UNIX 上，将保留在 /etc/services 中；而在 Windows 上，将保留在 %SystemRoot%\system32\drivers\etc\services 中。这些端口将由“快速通信管理器”使用。保留端口将使用以下格式：

```
DB2_InstanceName
DB2_InstanceName_1
DB2_InstanceName_2
DB2_InstanceName_END
```

唯一必填的条目是起始端口 (DB2\_InstanceName) 和结束端口 (DB2\_InstanceName\_END)。其他条目都保留在 services 文件中，以使其他应用程序不使用这些端口。

- 为了能支持多个参与 DB2 数据库服务器，要安装 DB2 的计算机必须属于一个可访问域。但是，可以将本地分区添加至该计算机，即使该计算机不属于某个域。
- 在 Linux 和 UNIX 系统上，分区数据库系统需要远程 shell 实用程序。DB2 支持下列远程 shell 实用程序：

- rsh
- ssh

缺省情况下，当执行远程 DB2 节点上的命令时（例如，启动远程 DB2 数据库分区时），DB2 就会使用 rsh 实用程序。要使用 DB2 缺省值，必须安装 rsh-server 程序包。有关安装 DB2 产品时安全性问题的更多信息，请参阅相关链接。

如果选择使用远程 shell 实用程序 rsh，那么还必须安装并运行 inetd（或 xinetd）。如果选择使用远程 shell 实用程序 ssh，那么需要在完成 DB2 安装之后立即设置 DB2RSHCMD 注册表变量。如果不设置此注册表变量，那么使用 rsh 实用程序。

- 在 Linux 和 UNIX 操作系统上，确保 etc 目录中的 hosts 文件不包含“127.0.0.2”的条目（如果该 IP 地址映射到机器的标准主机名）。

**注：**使用 XML 功能部件之后，您将无法使用数据库分区功能部件（DPF）。

要设置分区数据库环境：

1. 使用“DB2 安装”向导来安装实例拥有的数据库服务器。有关详细的指示信息，请参阅适合您的平台的相应的『安装 DB2 服务器』主题。
  - 在选择安装和/或响应文件创建窗口中，确保选择将安装设置保存在响应文件中选项。完成安装之后，下面两个文件将被复制至在“DB2 安装”向导中指定的以下目录中：PROD\_ESE.rsp 和 PROD\_ESE\_addpart.rsp。PROD\_ESE.rsp 文件是实例拥有的数据库服务器的响应文件。PROD\_ESE\_addpart.rsp 文件是参与的数据库服务器的响应文件。
  - 在为 **DB2 实例设置分区选项**窗口上，务必选择多个分区实例并输入最大逻辑分区数。
2. 使 DB2 安装映像可用于分区数据库环境中的所有分区计算机。
3. 分发参与数据库服务器响应文件（PROD\_ESE\_addpart.rsp）。
4. 在参与的每台计算机上，使用 db2setup 命令（在 Linux 和 UNIX 上）或者 setup（在 Windows 上）来安装 DB2 数据库服务器：

#### Linux 和 UNIX

转至提供了 DB2 产品代码的目录，然后运行：

```
./db2setup -r /responsefile_directory/response_file_name
```

#### Windows

```
setup -u x:\responsefile_directory\response_file_name
```

例如，下面就是一个使用 PROD\_ESE\_addpart.rsp 作为响应文件的命令：

#### Linux 和 UNIX

转至提供了 DB2 产品代码的目录，然后运行：

```
./db2setup -r /db2home/PROD_ESE_addpart.rsp
```

其中 /db2home 是已经复制了响应文件的目录。

#### Windows

```
setup -u c:\resp_files\PROD_ESE_addpart.rsp
```

其中 c:\resp\_files\ 是已经复制了响应文件的目录。

5. （仅适用于 Linux 和 UNIX）配置 db2nodes.cfg 文件。DB2 安装仅保留您要对当前计算机使用的最大逻辑分区数量，但不配置 db2nodes.cfg 文件。如果不配置 db2nodes.cfg 文件，那么实例仍是单一分区实例。
6. 更新参与服务器上的 services 文件，以便为 DB2 实例定义相应的 FCM 端口。该 services 文件位于下列位置：
  - /etc/services（在 Linux 和 UNIX 上）

- %SystemRoot%\system32\drivers\etc\services (在 Windows 上)

---

## 使用响应文件在参与的计算机上安装数据库分区服务器 (Windows)

在此任务中, 将使用您用“DB2 安装”向导创建的响应文件, 在参与的计算机上安装数据库分区服务器。

### 先决条件

- 已经在主计算机上使用“DB2 安装”向导安装了 DB2 副本。
- 您已经创建响应文件以安装在参与的计算机和将其复制到参与的计算机。
- 您在参与的计算机上必须具有管理权限。
- 您必须将 DB2 产品 DVD 的内容复制到参与的计算机。

要使用响应文件安装其他数据库分区服务器:

1. 使用已为 DB2 安装定义的本地管理员帐户来登录将参与分区数据库环境的计算机。
2. 切换至在其中复制 DB2 产品 DVD 的内容的目录。例如:

```
cd c:\db2dvd
```

其中 db2dvd 表示在其中复制 DB2 产品 DVD 内容的目录的名称。

3. 从命令提示符处输入 setup 命令, 如下所示:

```
setup -u responsefile_directory\response_file_name
```

在下列示例中, 响应文件 Addpart.file 可在 c:\responsefile 目录中找到。此示例的命令将是:

```
setup -u c:\responsefile\Addpart.file
```

4. 完成安装之后, 检查日志文件中的消息。您可以在 My Documents\DB2LOG\ 目录中找到该日志文件。您应该在日志文件的末尾看到与下列信息类似的输出:

```
=== Logging stopped: 5/9/2007 10:41:32 ===  
MSI (c) (C0:A8) [10:41:32:984]: Product: DB2  
Enterprise Server Edition - DB2COPY1 -- Installation  
operation completed successfully.
```

5. 当在主计算机上安装实例拥有的数据库分区服务器时, DB2 产品根据指定的参与分区数据库环境的逻辑数据库分区服务器数目保留端口范围。缺省范围为四个端口。对于每个参与分区数据库环境的服务器, 您必须为 FCM 端口手工配置 /etc/services 文件。FCM 端口范围取决于要在参与的计算机上使用的逻辑分区数目。至少需要两个条目: **DB2\_<instance>** 和 **DB2\_<instance>\_END**。在参与的计算机上指定的 FCM 端口的其他要求:
  - 起始端口号必须与主计算机的起始端口号相匹配。
  - 后续端口必须按顺序进行编号。
  - 指定的端口号必须是空闲的。

必须登录参与的每台计算机并重复这些步骤。

如果想要 DB2 产品能够访问本地计算机或网络中的另一计算机上的 DB2 文档, 那么必须安装 DB2 信息中心。“DB2 信息中心”包含 DB2 数据库系统和 DB2 相关产品的文档。

---

## 使用响应文件在参与的计算机上安装数据库分区服务器 (Linux 和 UNIX)

在此任务中，将使用您用“DB2 安装”向导创建的响应文件，在参与的计算机上安装数据库分区服务器。

### 先决条件

- 已经在主计算机上使用“DB2 安装”向导安装了 DB2，并在参与的计算机上创建了用于安装的响应文件。
- 您在参与的计算机上必须具有 root 用户权限。

要使用响应文件安装其他数据库分区服务器：

1. 作为 root 用户登录将参与分区数据库环境的计算机。
2. 切换至在其中复制 DB2 产品 DVD 的内容的目录。例如：

```
cd /db2home/db2dvd
```

3. 输入 db2setup 命令，如下所示：

```
./db2setup -r /responsefile_directory/response_file_name
```

在本示例中，已将响应文件 AddPartitionResponse.file 保存到 /db2home 目录中。本示例的命令将为：

```
./db2setup -r /db2home/AddPartitionResponse.file
```

4. 完成安装之后，检查日志文件中的消息。

必须登录参与的每台计算机并执行响应文件安装。

如果想要 DB2 产品能够访问本地计算机或网络中的另一计算机上的 DB2 文档，那么必须安装 DB2 信息中心。“DB2 信息中心”包含 DB2 数据库系统和 DB2 相关产品的文档。

---

## 第 8 章 安装之后

---

### 验证安装

#### 验证分区数据库环境安装 ( Windows )

要验证是否成功安装了 DB2 服务器，创建样本数据库并运行 SQL 命令来检索样本数据并验证数据是否已分发至所有参与的数据库分区服务器。

已完成所有安装步骤。

要创建 SAMPLE 数据库：

1. 作为具有 SYSADM 权限的用户登录主计算机 ( ServerA )。
2. 输入 db2sampl 命令来创建 SAMPLE 数据库。

处理此命令可能要花几分钟。当返回命令提示符时，该过程完成。

创建 SAMPLE 数据库时，该数据库自动以数据库别名 SAMPLE 进行编目。

3. 输入 db2start 命令来启动数据库管理器。
4. 在 DB2 命令窗口中输入下列 DB2 命令来连接至 SAMPLE 数据库，并检索在部门 20 工作的所有职员的列表：

```
db2 connect to sample
db2 "select * from staff where dept = 20"
```

5. 要验证是否已将数据分发至数据库分区服务器，在 DB2 命令窗口中输入下列命令：

```
db2 "select distinct dbpartitionnum(empno) from employee"
```

输出将列示 employee 表使用的数据库分区。特定输出将取决于数据库中的数据库分区数以及创建 employee 表的表空间所使用的数据库分区组中的数据库分区数。

在验证安装后，可除去 SAMPLE 数据库以释放磁盘空间。但是，如果打算利用样本应用程序，那么保留样本数据库很有用。

输入 db2 drop database sample 命令以删除 SAMPLE 数据库。

#### 验证分区数据库服务器安装 ( Linux 和 UNIX )

要验证是否成功安装了 DB2 服务器，创建样本数据库并运行 SQL 命令来检索样本数据并验证数据是否已分发至所有参与的数据库分区服务器。

在执行这些步骤之前，确保已完成所有安装步骤。

要创建 SAMPLE 数据库：

1. 作为拥有实例的用户登录主计算机 ( ServerA )。对于此示例，db2inst1 是拥有实例的用户。
2. 输入 db2sampl 命令来创建 SAMPLE 数据库。缺省情况下，将在实例所有者的主目录中创建样本数据库。在本示例中，/db2home/db2inst1/ 是实例所有者的主目录。实例所有者的主目录是缺省数据库路径。

处理此命令可能要花几分钟。没有完成消息；当返回命令提示符时，该过程完成。

创建 SAMPLE 数据库时，该数据库自动以数据库别名 SAMPLE 进行编目。

3. 输入 `db2start` 命令来启动数据库管理器。
4. 在 DB2 命令窗口中输入下列 DB2 命令来连接至 SAMPLE 数据库，并检索在部门 20 工作的所有职员列表：

```
db2 connect to sample
db2 "select * from staff where dept = 20"
```

5. 要验证是否已将数据分发至数据库分区服务器，在 DB2 命令窗口中输入下列命令：  
`db2 "select distinct dbpartitionnum(empno) from employee"`

输出将列示 `employee` 表使用的数据库分区。特定输出将取决于：

- 数据库中数据库分区的数目
- 创建 `employee` 表的表空间所使用的数据库分区组中的数据库分区数

在验证安装后，可除去 SAMPLE 数据库以释放磁盘空间。输入 `db2 drop database sample` 命令以删除 SAMPLE 数据库。



---

## 第 3 部分 实施和维护



---

## 第 9 章 创建数据库之前

---

### 设置分区数据库环境

必须在创建数据库之前决定创建多分区数据库。在作出数据库设计决定时，必须确定是否应利用数据库分区可以提供的性能提高。

在分区数据库环境中，仍然使用 `CREATE DATABASE` 命令或 `sqlcrea()` 函数来创建数据库。无论使用哪种方法，都可以通过 `db2nodes.cfg` 文件中列示的任何分区来发出请求。

在创建多分区数据库之前，必须选择将作为数据库的目录分区的数据库分区。然后，可以直接从该数据库分区创建数据库，也可以从连接至该数据库分区的远程客户机创建数据库。您要连接并对其执行 `CREATE DATABASE` 命令的数据库分区成为该特定数据库的目录分区。

目录分区是用于存储所有系统目录表的数据库分区。对系统表的所有访问都必须通过此数据库分区进行。所有联合数据库对象（例如，包装器、服务器和昵称）都存储在此数据库分区上的系统目录表中。

若可能的话，应该在独立的实例中创建每个数据库。若不可能做到此点（即，必须在每个实例中创建多个数据库），应该将目录分区分布至可用的数据库分区中。这样可以减少在单个数据库分区中对目录信息的争用。

**注：**应该定期备份目录分区，同时，因为其他数据会增加备份所需的时间，所以要避免将用户数据置于该节点上（任何可能的時候）。

当创建数据库时，它会在 `db2nodes.cfg` 文件中定义的所有数据库分区之间自动创建。

创建系统中的第一个数据库时，就会形成一个系统数据库目录。并追加有关您创建的任何其他数据库的信息。在 UNIX 上工作时，系统数据库目录是 `sqlbdir`，位于主目录下的 `sqllib` 目录中或安装 DB2 数据库的目录下面。在 UNIX 上工作时，由于组成分区数据库环境的所有数据库分区只有一个系统数据库目录，所以此目录必须位于共享文件系统（例如，UNIX 平台上的 NFS）上。在 Windows 上工作时，系统数据库目录位于实例目录中。

位于 `sqlbdir` 目录中的还有系统意向文件。它称为 `sqlbins`，用于确保数据库分区保持同步。该文件也必须位于共享文件系统中，因为所有数据库分区中只有一个目录。该文件由组成数据库的所有数据库分区共享。

必须修改配置参数，才能利用数据库分区。使用 `GET DATABASE CONFIGURATION` 和 `GET DATABASE MANAGER CONFIGURATION` 命令以了解特定数据库或数据库管理器配置文件中的个别条目的值。要修改特定数据库或数据库管理器配置文件中的个别条目，可分别使用 `UPDATE DATABASE CONFIGURATION` 和 `UPDATE DATABASE MANAGER CONFIGURATION` 命令。

影响分区数据库环境的数据库管理器配置参数包括 **conn\_elapse**、**fcm\_num\_buffers**、**fcm\_num\_channels**、**max\_connretries**、**max\_coordagents**、**max\_time\_diff**、**num\_poolagents** 和 **stop\_start\_time**。

---

## 创建节点配置文件

如果数据库要在分区数据库环境中运行，那么必须创建一个名为 `db2nodes.cfg` 的节点配置文件。

此文件必须在实例主目录的 `sqllib` 子目录中，这样才能在跨多个数据库分区启用并行能力的情况下启动数据库管理器。该文件包含一个实例中所有数据库分区的配置信息，并且它由该实例的所有数据库分区共享。

### Windows 注意事项

如果正在 Windows 上使用 DB2 企业服务器版，那么节点配置文件是在创建实例时创建的。您不应尝试手动创建或修改节点配置文件。可使用 `db2nrcrt` 命令来将数据库分区服务器添加至实例。可使用 `db2ndrop` 命令从实例中删除数据库分区服务器。可使用 `db2nchg` 命令来修改数据库分区服务器配置，包括将数据库分区服务器从一台计算机移至另一台计算机；更改 TCP/IP 主机名；或选择另一逻辑端口或网络名。

**注：** 不应该在不是数据库管理器创建的 `sqllib` 子目录下创建文件或目录，以防止删除实例时丢失数据。但有两个例外情况。如果系统支持存储过程，那么将该存储过程应用程序置于 `sqllib` 子目录下的 `function` 子目录中。另一个例外是在已创建用户定义的函数 (UDF) 的情况下。允许 UDF 可执行程序位于同一个目录中。

对于属于一个实例的每个数据库分区该文件都包含一行。每行的格式如下：

```
dbpartitionnum hostname [logical-port [netname]]
```

记号由空格定界。这些变量是：

#### **dbpartitionnum**

数据库分区号唯一地定义数据库分区，可在 0 到 999 之间。数据库分区号必须以升序顺序排序。该顺序中可以有空隙。

一旦指定了数据库分区号，就不能对其进行更改。否则，分布图（它指定数据分布方式）中的信息可能不正确。

如果删除一个数据库分区，那么它的数据库分区号可以再次用于添加的任何新数据库分区。

数据库分区号用于在数据库目录中生成数据库分区名。它的格式为：

```
NODE nnnn
```

`nnnn` 是数据库分区号，其左边以零填充。CREATE DATABASE 和 DROP DATABASE 命令也使用此数据库分区号。

#### **hostname**

用作分区间通信的 IP 地址的主机名。使用主机名的标准名称。`/etc/hosts` 文件也应该使用标准名称。如果未在 `db2nodes.cfg` 文件和 `/etc/hosts` 文件中使用标准名称，那么可能接收到错误消息 `SQL30082N RC=3`。

（指定 `netname` 时例外。在此情况下，`netname` 用于大多数通信，而 `hostname` 仅用于 `db2start`、`db2stop` 和 `db2_all`。）

### logical-port

此参数是可选的，它指定该数据库分区的逻辑端口号。此号码与数据库管理器实例名一起用来标识 `etc/services` 文件中的 TCP/IP 服务名称条目。

IP 地址和逻辑端口的组合被用作熟知地址，且在所有支持数据库分区间通信连接的应用程序中必须是唯一的。

对于每个主机名，一个逻辑端口必须为 0（零）或空白（缺省为 0）。与此逻辑端口相关联的数据库分区是与客户机连接的主机上的缺省节点。可以使用 `db2profile` 脚本中的 `DB2NODE` 环境变量或 `sqlsetc()` API 来将其覆盖。

### netname

此参数是可选的，并且用于支持有多个活动 TCP/IP 接口的主机，每个接口有其自己的主机名。

以下示例显示了一个 RS/6000 SP 系统的可能节点配置文件，在该系统上，SP2EN1 有多个 TCP/IP 接口和两个逻辑分区，并且使用 SP2SW1 作为 DB2 数据库接口。此示例还显示了从 1 开始（而不是从 0 开始）的数据库分区号以及 `dbpartitionnum` 序列中的间隙：

表 12. 数据库分区号示例表。

| dbpartitionnum | hostname             | logical-port | netname |
|----------------|----------------------|--------------|---------|
| 1              | SP2EN1.mach1.xxx.com | 0            | SP2SW1  |
| 2              | SP2EN1.mach1.xxx.com | 1            | SP2SW1  |
| 4              | SP2EN2.mach1.xxx.com | 0            |         |
| 5              | SP2EN3.mach1.xxx.com |              |         |

可以使用选择的编辑器更新 `db2nodes.cfg` 文件。（例外情况：不应在 Windows 上使用编辑器）。但是，必须小心保护文件中信息的完整性，这是因为数据库分区要求当您发出 `db2start` 时使节点配置文件处于锁定状态，而在发出 `db2stop` 来结束数据库管理器之后又将它解锁。文件被锁定时，`db2start` 命令可在必要时更新该文件。例如，可发出带有 `RESTART` 选项或 `ADDNODE` 选项的 `db2start`。

**注：**如果 `db2stop` 命令不成功而又未解锁该节点配置文件，那么发出 `db2stop FORCE` 来将其解锁。

## DB2 节点配置文件的格式

本主题提供有关节点配置文件（`db2nodes.cfg`）格式的信息。`db2nodes.cfg` 文件用来定义参与 DB2 实例的数据库分区服务器。如果想要将高速互连用于数据库分区服务器通信，那么还可以使用 `db2nodes.cfg` 文件来指定高速互连的 IP 地址或主机名。

Linux 和 UNIX 操作系统上的 `db2nodes.cfg` 文件的格式如下：

```
nodenumber hostname logicalport netname resourcesetname
```

在下一节中定义了 `nodenumber`、`hostname`、`logicalport`、`netname` 和 `resourcesetname`。

Windows 操作系统上的 `db2nodes.cfg` 文件的格式如下：

```
nodenumber hostname computername logicalport netname resourcesetname
```

在 Windows 操作系统上，db2nodes.cfg 的这些条目由 db2ncrt 或 db2 add db partition 命令添加。您不应该直接添加这些行或编辑此文件。

### **nodenumber**

这是一个唯一号码，其范围是 0 到 999，用来标识分区数据库系统中的数据库分区服务器。

要扩充分区数据库系统，对每个数据库分区服务器，向 db2nodes.cfg 文件添加一个条目。必须按升序为其他的数据库分区服务器选择 *nodenumber* 值，但是在此序列中可以存在间隔。如果您打算添加逻辑分区服务器，并且希望使节点在此文件中保持按逻辑分组，那么可以选择使 *nodenumber* 值之间保持一定间隔。

此条目是必需的。

### **hostname**

供快速通信管理器（FCM）使用的数据库分区服务器的 TCP/IP 主机名。

此条目是必需的。

### **logicalport**

指定数据库分区服务器的逻辑端口号。此字段用来在正在运行逻辑数据库分区服务器的工作站上指定特定数据库分区服务器。

在安装时，DB2 将在 /etc/services 文件中保留一定范围内的端口（例如，60000 到 60003 之间）用于分区间通信。db2nodes.cfg 中的 *logicalport* 字段指定您要将该范围中的哪个端口分配给特定的逻辑分区服务器。

如果此字段中无任何条目，那么缺省值为 0。但是，如果对 *netname* 字段添加一个条目，那么必须对 *logicalport* 字段输入一个数字。

如果正在使用逻辑数据库分区，那么指定的 *logicalport* 值必须从 0 开始，并按升序依次递增（例如，0、1、2）。

此外，如果为一个数据库分区服务器指定 *logicalport* 条目，那么必须为 db2nodes.cfg 文件中列示的每个数据库分区服务器指定 *logicalport*。

仅当未使用逻辑数据库分区或高速互连时，此字段才是可选的。

### **netname**

指定用于 FCM 通信的高速互连的主机名或 IP 地址。

如果为此字段指定了一个条目，那么数据库分区服务器之间的所有通信（除了由于 db2start、db2stop 和 db2\_all 命令而进行的通信之外）都是通过高速互连来处理的。

仅当您要使用高速互连来进行数据库分区通信时，才需要此参数。

### **resourcesetname**

*resourcesetname* 定义应在其中启动节点的操作系统资源。*resourcesetname* 用于进程相似性支持、用于多个逻辑节点（MLN）。此支持与先前被称为 *quadname* 的字符串类型字段一起提供。

此参数仅在 AIX、HP-UX 和 Solaris 操作系统上受支持。

在 AIX 上，此概念被称作“资源集”，而在 Solaris 操作系统上，它被称为“项目”。有关资源管理的更多信息，请参阅操作系统文档。

在 HP-UX 上，*resourcesetname* 参数是 PRM 组的名称。请参阅 HP 的“HP-UX Process Resource Manager User Guide (B8733-90007)”文档以获取更多信息。

在 Windows 操作系统上，可以通过 DB2PROCESSORS 注册表变量来定义逻辑节点的进程相似性。

在 Linux 操作系统上，resourcesetname 列定义与系统上的“非一致性内存访问”（NUMA）节点相对应的一个数字。除了具有 NUMA 策略支持的 2.6 内核之外，系统实用程序 numactl 必须可用。

如果使用 resourcesetname 参数，那么必须指定 netname 参数。

## 示例配置

使用下面的示例配置来确定环境的相应配置。

### 一台计算机，四个数据库分区服务器

如果未使用集群环境，且想要在一台名为 ServerA 的物理工作站上具有四个数据库分区服务器，那么应对 db2nodes.cfg 文件作如下更新：

```
0          ServerA      0
1          ServerA      1
2          ServerA      2
3          ServerA      3
```

### 两台计算机，每台计算机一个数据库分区服务器

如果想要让分区数据库系统包含两个物理工作站：ServerA 和 ServerB，那么应对 db2nodes.cfg 文件作如下更新：

```
0          ServerA      0
1          ServerB      0
```

### 两台计算机，一台计算机上有三个数据库分区服务器

如果想要分区数据库系统包含两个物理工作站：ServerA 和 ServerB，并且 ServerA 要运行 3 个数据库分区服务器，那么应对 db2nodes.cfg 文件作如下更新：

```
4          ServerA      0
6          ServerA      1
8          ServerA      2
9          ServerB      0
```

### 两台计算机，带有高速交换机的三个数据库分区服务器

如果想要分区数据库系统包含两台计算机：ServerA 和 ServerB（且 ServerB 运行两个数据库分区服务器），并且使用名为 switch1 和 switch2 的高速互连，那么应对 db2nodes.cfg 文件作如下更新：

```
0          ServerA      0          switch1
1          ServerB      0          switch2
2          ServerB      1          switch2
```

## 使用 resourcename 的示例

这些限制适用于以下示例：

- 本示例说明在配置中没有高速互连时 resourcename 的使用。
- netname 是第四列，在该列没有交换机名而您却想要使用 resourcename 的情况下，您还可以在该列上指定 hostname。第五个参数是 resourcename（如果已定义的话）。此资源组规范只可显示为 db2nodes.cfg 文件中的第五列。这意味着，要指定资源组，还必须输入第四列。第四列打算用于高速交换机。

- 如果没有高速交换机或者不想使用它，那么必须输入 `hostname`（与第二列相同）。换句话说，DB2 数据库管理系统不支持 `db2nodes.cfg` 文件中存在列间隔（或列交换）。此限制以前适用于前三列，现在它适用于所有五列。

## AIX 示例

此处是如何为 AIX 操作系统设置资源集的示例。

在此示例中，有一个具有 32 个处理器和 8 个逻辑数据库分区（MLN）的物理节点。此示例说明如何为每个 MLN 提供进程相似性。

1. 在 `/etc/rset` 中定义资源集:

```
DB2/MLN1:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00000,sys/cpu.00001,sys/cpu.00002,sys/cpu.00003
```

```
DB2/MLN2:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00004,sys/cpu.00005,sys/cpu.00006,sys/cpu.00007
```

```
DB2/MLN3:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00008,sys/cpu.00009,sys/cpu.00010,sys/cpu.00011
```

```
DB2/MLN4:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00012,sys/cpu.00013,sys/cpu.00014,sys/cpu.00015
```

```
DB2/MLN5:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00016,sys/cpu.00017,sys/cpu.00018,sys/cpu.00019
```

```
DB2/MLN6:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00020,sys/cpu.00021,sys/cpu.00022,sys/cpu.00023
```

```
DB2/MLN7:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00024,sys/cpu.00025,sys/cpu.00026,sys/cpu.00027
```

```
DB2/MLN8:
owner      = db2inst1
group      = system
perm       = rwr-r-
resources  = sys/cpu.00028,sys/cpu.00029,sys/cpu.00030,sys/cpu.00031
```

2. 通过输入以下命令来启用内存相似性:

```
vmo -p -o memory_affinity=1
```

3. 允许实例使用资源集:



```
chuser capabilities=
    CAP_BYPASS_RAC_VMM,CAP_PROPAGATE,CAP_NUMA_ATTACH db2inst1
```

4. 将资源集名称作为第五列添加到 db2nodes.cfg:

```
1 regatta 0 regatta DB2/MLN1
2 regatta 1 regatta DB2/MLN2
3 regatta 2 regatta DB2/MLN3
4 regatta 3 regatta DB2/MLN4
5 regatta 4 regatta DB2/MLN5
6 regatta 5 regatta DB2/MLN6
7 regatta 6 regatta DB2/MLN7
8 regatta 7 regatta DB2/MLN8
```

## HP-UX 示例

此示例说明如何使用 PRM 组在使用 4 个 CPU 和 4 个 MLN 的机器上分享 CPU 资源，为每个 MLN 设置 24% 的 CPU 份额，为其他应用程序留下 4%。DB2 实例名为 db2inst1。

1. 编辑 /etc/prmconf 的 GROUP 段:

```
OTHERS:1:4::
db2prm1:50:24::
db2prm2:51:24::
db2prm3:52:24::
db2prm4:53:24::
```

2. 向 /etc/prmconf 添加实例所有者条目:

```
db2inst1:::OTHERS,db2prm1,db2prm2,db2prm3,db2prm4
```

3. 通过输入以下命令来初始化组并启用 CPU 管理器:

```
prmconfig -i
prmconfig -e CPU
```

4. 将 PRM 组名作为第五列添加到 db2nodes.cfg:

```
1 voyager 0 voyager db2prm1
2 voyager 1 voyager db2prm2
3 voyager 2 voyager db2prm3
4 voyager 3 voyager db2prm4
```

可以使用交互式 GUI 工具 xprm 执行 PRM 配置（步骤 1-3）。

## Linux 示例

在 Linux 操作系统上，resourcesetname 列定义与系统上的“非一致性内存访问”（NUMA）节点相对应的一个数字。作为对支持 NUMA 策略的 2.6 内核的补充，还必须具有 numactl 系统实用程序。有关 Linux 操作系统上的 NUMA 支持的更多信息，请参阅关于 numactl 的联机帮助页。

本示例说明如何设置一个具有四个 NUMA 节点的计算机，并使每个逻辑节点都与一个 NUMA 节点相关联。

1. 确保系统上具备 NUMA 功能。
2. 发出下列命令:

```
$ numactl --hardware
```

将显示与以下内容相似的输出:

```
available: 4 nodes (0-3)
node 0 size: 1901 MB
node 0 free: 1457 MB
```

```
node 1 size: 1910 MB
node 1 free: 1841 MB
node 2 size: 1910 MB
node 2 free: 1851 MB
node 3 size: 1905 MB
node 3 free: 1796 MB
```

3. 在此示例中，系统上有四个 NUMA 节点。按如下所示编辑 db2nodes.cfg 文件，以使每个 MLN 都与系统上的一个 NUMA 节点相关联：

```
0 hostname 0 hostname 0
1 hostname 1 hostname 1
2 hostname 2 hostname 2
3 hostname 3 hostname 3
```

## Solaris 示例

此处是如何为 Solaris V9 设置项目的示例。

在此示例中，有一个带有 8 个处理器的物理节点：其中 1 个 CPU 将用于缺省项目，3 个 CPU 由应用程序服务器使用，4 个 CPU 用于 DB2。实例名为 db2inst1。

1. 使用编辑器创建资源池配置文件。对于此示例，该文件将被称为 pool.db2。其内容如下：

```
create system hostname
create pset pset_default (uint pset.min = 1)
create pset db0_pset (uint pset.min = 1; uint pset.max = 1)
create pset db1_pset (uint pset.min = 1; uint pset.max = 1)
create pset db2_pset (uint pset.min = 1; uint pset.max = 1)
create pset db3_pset (uint pset.min = 1; uint pset.max = 1)
create pset appsrv_pset (uint pset.min = 3; uint pset.max = 3)
create pool pool_default (string pool.scheduler="TS";
    boolean pool.default = true)
create pool db0_pool (string pool.scheduler="TS")
create pool db1_pool (string pool.scheduler="TS")
create pool db2_pool (string pool.scheduler="TS")
create pool db3_pool (string pool.scheduler="TS")
create pool appsrv_pool (string pool.scheduler="TS")
associate pool pool_default (pset pset_default)
associate pool db0_pool (pset db0_pset)
associate pool db1_pool (pset db1_pset)
associate pool db2_pool (pset db2_pset)
associate pool db3_pool (pset db3_pset)
associate pool appsrv_pool (pset appsrv_pset)
```

2. 编辑 /etc/project 文件以添加 DB2 项目和 appsrv 项目，如下所示：

```
system:0::::
user.root:1::::
noproject:2::::
default:3::::
group.staff:10::::
appsrv:4000:App Serv project:root::project.pool=appsrv_pool
db2proj0:5000:DB2 Node 0 project:db2inst1,root::project.pool=db0_pool
db2proj1:5001:DB2 Node 1 project:db2inst1,root::project.pool=db1_pool
db2proj2:5002:DB2 Node 2 project:db2inst1,root::project.pool=db2_pool
db2proj3:5003:DB2 Node 3 project:db2inst1,root::project.pool=db3_pool
```

3. 创建资源池：# poolcfg -f pool.db2。
4. 激活资源池：# pooladm -c
5. 将项目名作为第五列添加到 db2nodes.cfg 文件：

```
0 hostname 0 hostname db2proj0
1 hostname 1 hostname db2proj1
2 hostname 2 hostname db2proj2
3 hostname 3 hostname db2proj3
```

## 指定分区数据库环境中的机器列表

缺省情况下，从配置文件 `db2nodes.cfg` 中获取计算机列表。

可以通过以下方法覆盖此设置：

- 通过导出（在 Linux 和 UNIX 平台上）或设置（在 Windows 上）环境变量 `RAHOSTFILE`，指定包含计算机列表的文件的名称。
- 通过导出（在 Linux 和 UNIX 平台上）或设置（在 Windows 上）环境变量 `RAHOSTLIST`，明确指定该列表为由空格分隔的一连串名称。

**注：**如果这两个环境变量都被指定，那么 `RAHOSTLIST` 具有优先顺序。

**注：**在 Windows 上，要避免将不一致引入该节点配置文件，不要以手动方式对其进行编辑。要获取实例中的计算机列表，可使用 `db2nlist` 命令。

## 除去分区数据库环境中机器列表中的重复条目

如果正在一台计算机上运行具有多个逻辑数据库分区服务器的 DB2 企业服务器版，那么 `db2nodes.cfg` 文件将包含该计算机的多个条目。

在这种情况下，`rah` 命令需要知道您是希望该命令在每台计算机上只执行一次，还是对 `db2nodes.cfg` 文件中列示的每个逻辑数据库分区均执行一次。使用 `rah` 命令来指定计算机。使用 `db2_all` 命令来指定逻辑数据库分区。

**注：**在 Linux 和 UNIX 平台上，如果指定计算机，那么 `rah` 将会正常地从计算机列表中删除重复的项，例外情况是：如果指定逻辑数据库分区，那么 `db2_all` 会将下列赋值附加到您的命令之前：

```
export DB2NODE=nnn （对于 Korn shell 程序语法）
```

其中 `nnn` 取自 `db2nodes.cfg` 文件的相应行中的数据库分区号，以便将命令路由至所希望的数据库分区服务器。

当指定逻辑数据库分区时，可以使用 `<<-nnn<` 和 `<<+nnn<` 前缀顺序来限制列表包括除某个逻辑数据库分区外的所有逻辑数据库分区，或只指定一个逻辑数据库分区。如果想要首先运行对数据库分区进行编目的命令，并在该命令完成时，在所有其他数据库分区服务器上运行同一命令（可能以并行方式），那么可能要执行此操作。当运行 `db2 restart database` 命令时，它常常是必需的。需要知道目录分区的数据库分区号才能执行此操作。

如果使用 `rah` 命令来执行 `db2 restart database`，重复条目就会从计算机列表中被删除。然而，如果指定 `”` 前缀，那么不删除重复项，因为认为使用 `”` 前缀会隐含发送至每个数据库分区服务器，而不是发送至每台计算机。

## 更新节点配置文件（Linux 和 UNIX）

此任务提供了一些步骤，以将 `db2nodes.cfg` 文件更新为包括参与的计算机的条目。

节点配置文件 (db2nodes.cfg) 位于实例所有者的主目录中, 它包含一些配置信息, 告诉 DB2 有哪些服务器参与分区数据库环境的实例。分区数据库环境中的每个实例都有一个 db2nodes.cfg 文件。

对于每个参与实例的服务器, db2nodes.cfg 文件必须包含一个条目。当创建实例时, 会自动创建 db2nodes.cfg 文件并对拥有实例的服务器添加条目。

例如, 在拥有实例的服务器 ServerA 上使用“DB2 安装”向导创建了 DB2 实例时, db2nodes.cfg 文件将被更新为如下所示:

```
0      ServerA      0
```

#### 先决条件

- 必须在所有参与的计算机上安装 DB2 应用程序。
- DB2 实例必须存在于主计算机上。
- 您必须是具有 SYSADM 权限的用户。
- 如果以下条件均适用, 查看在 DB2 节点配置文件主题的格式中提供的配置示例和文件格式信息:
  - 您计划使用高速交换机在数据库分区服务器之间进行通信
  - 您的分区配置将具有多个逻辑分区

**限制** 在“过程”部分的步骤中使用的主机名必须是标准主机名。

要更新 db2nodes.cfg 文件:

1. 作为实例所有者登录 (在本示例, db2inst1 是实例所有者)。
2. 通过输入以下命令确保已停止 DB2 实例:

```
INSTHOME/sqllib/adm/db2stop
```

其中 *INSTHOME* 是实例所有者的主目录 (db2nodes.cfg 文件在实例运行时被锁定, 并且仅当实例停止时才可以编辑该文件)。

例如, 如果实例主目录为 /db2home/db2inst1, 那么输入以下命令:

```
/db2home/db2inst1/sqllib/adm/db2stop
```

3. 对于每个 DB2 实例, 向 .rhosts 文件添加一个条目。通过添加以下项来更新文件:

```
<hostname> <db2instance>
```

其中 <hostname> 是数据库服务器的 TCP/IP 主机名, <db2instance> 是用来访问数据库服务器的实例的名称。

4. 向每个参与的服务器的 db2nodes.cfg 文件添加一个条目。当第一次查看 db2nodes.cfg 文件时, 它应该包含类似于以下内容的条目:

```
0      ServerA      0
```

此条目包括数据库分区服务器号 (节点号)、数据库分区服务器驻留的服务器的 TCP/IP 主机名以及数据库分区服务器的逻辑端口号。

例如, 如果正在对分区配置 (有四台计算机, 每台计算机上安装一个数据库分区服务器) 进行安装, 那么应该出现类似如下的已更新 db2nodes.cfg:

```
0          ServerA          0
1          ServerB          0          2      ServerC          0
3          ServerD          0
```

5. 当完成更新 `db2nodes.cfg` 文件时，输入 `INSTHOME/sqllib/adm/db2start` 命令，其中 `INSTHOME` 是实例所有者的主目录。例如，如果实例主目录为 `/db2home/db2inst1`，那么输入以下命令：

```
/db2home/db2inst1/sqllib/adm/db2start
```

6. 注销。

## 设置多逻辑节点

通常，将 DB2 企业服务器版配置为对每台计算机分配一个数据库分区服务器。然而，在几种情况下，在同一计算机上运行数个数据库分区服务器很有好处。

这意味着配置包含的数据库分区数可以多于计算机数。在这些情况下，我们称该计算机正在运行多逻辑分区或多逻辑节点（如果这些逻辑节点参与同一实例的话）。如果它们参与不同的实例，那么此计算机不主管多逻辑分区。

借助多逻辑分区支持，您可以从三种类型的配置中进行选择：

- 标准配置，即每台计算机只有一个数据库分区服务器。
- 多逻辑分区配置，即计算机有多个数据库分区服务器。
- 在数台计算机的每一台上运行数个逻辑分区的配置。

当系统在具有对称多处理器（SMP）体系结构的计算机上运行查询时，使用多逻辑分区的配置非常有用。万一计算机出现故障，在计算机上配置多逻辑分区的能力也很有用。如果一台计算机出现故障（导致其上的数据库分区服务器失败），那么可以使用 `DB2START NODENUM` 命令在另一机器上重新启动数据库分区服务器。这确保用户数据保持可用。

另一个好处是多逻辑分区可利用 SMP 硬件配置。另外，因为数据库分区较小，所以当执行诸如备份和复原数据库和表空间以及创建索引之类的任务时，可以获得较佳的性能。

## 配置多逻辑分区

可使用两种方法来配置多逻辑分区。

- 在 `db2nodes.cfg` 文件中配置逻辑分区（数据库分区）。然后，可使用 `DB2START` 命令或它的相关 API 启动所有逻辑分区和远程分区。

**注：**对于 Windows(R)，如果系统中没有数据库，那么必须使用 `db2ncrt` 来添加数据库分区；否则，如果有一个或多个数据库，那么应使用 `DB2START ADDNODE` 命令。在 Windows 中，绝对不应手动编辑 `db2nodes.cfg` 文件。

- 在另一个处理器上重新启动一个逻辑分区，其他逻辑分区（节点）已在该处理器上运行。这允许您覆盖在 `db2nodes.cfg` 中为逻辑分区指定的主机名和端口号。

要在 `db2nodes.cfg` 中配置一个逻辑分区（节点），您必须在该文件中建立一个条目以便为该数据库分区分配一个逻辑端口号。以下是应使用的语法：

```
nodenumber hostname logical-port netname
```

**注：**对于 Windows，如果系统中没有数据库，那么必须使用 `db2ncrt` 来添加数据库分区；否则，如果有一个或多个数据库，那么应使用 `DB2START ADDNODE` 命令。在 Windows 中，绝对不应手动编辑 `db2nodes.cfg` 文件。

Windows 上 db2nodes.cfg 文件的格式与 UNIX 上同一文件的格式不同。在 Windows 上，列格式为：

```
nodenumber hostname computername logical_port netname
```

使用主机名的标准名称。/etc/hosts 文件也应该使用标准名称。如果未在 db2nodes.cfg 文件和 /etc/hosts 文件中使用标准名称，那么可能接收到错误消息 SQL30082N RC=3。

必须确保在 etc 目录的 services 文件中为 FCM 通信定义了足够的端口。

---

## 启用分区内查询并行性

根据数据库分区的数目以及数据在这些数据库分区上的分布，分区间并行性自动生效。

**注：**必须修改配置参数，以利用数据库分区内或非分区数据库内的并行性。例如，可以使用分区内并行性来利用对称多处理器（SMP）机器上的多个处理器。

### 启用并行性以装入数据

Load 实用程序自动利用并行性，也可在 LOAD 命令上使用下列参数：

- CPU\_PARALLELISM
- DISK\_PARALLELISM

在分区数据库环境中，当对多个数据库分区定义目标表时，用于数据装入的分区间并行性就会自动发生。用于数据装入的分区间并行性可通过指定 OUTPUT\_DBPARTNUMS 来覆盖。load 实用程序还会根据目标数据库分区大小来智能地启用数据库分区并行性。可使用 MAX\_NUM\_PART\_AGENTS 来控制 load 实用程序选择的最大并行度。可在指定 ANYORDER 的同时指定 PARTITIONING\_DBPARTNUMS 以覆盖数据库分区并行性。

### 在创建索引时启用并行性

要在创建索引时启用并行性：

- *intra\_parallel* 数据库管理器配置参数必须为 ON
- 表必须足够大，以便能从并行性受益
- 在 SMP 计算机上必须启用多个处理器。

### 在备份数据库或表空间时启用 I/O 并行性

要在备份数据库或表空间时启用 I/O 并行性：

1. 使用多个目标媒体。
2. 通过定义多个容器为表空间配置并行 I/O，或将单个容器与多个磁盘配合使用，并适当地使用 DB2\_PARALLEL\_IO 注册表变量。如果想要利用并行 I/O，那么必须考虑在定义任何容器之前必须完成的操作。不能在您认为需要时才执行此操作；必须在需要备份数据库或表空间之前进行计划。
3. 在 BACKUP 命令上使用 PARALLELISM 参数以指定并行度。
4. 在 BACKUP 命令上使用 WITH num-buffers BUFFERS 参数以保证提供足够的缓冲区来满足该并行度。缓冲区数应比已有的目标媒体数与选择的并行度之和略大。

同时使用满足以下条件的备份缓冲区大小：

- 尽可能大。4 MB 或 8 MB（1024 或 2048 页）比较合适。

- 至少等于要备份的表空间（扩展数据块大小 \* 容器数之积）的最大者。

### 在复原数据库或表空间时启用 I/O 并行性

要在复原数据库或表空间时启用 I/O 并行性:

- 使用多个源媒体。
- 配置用于并行 I/O 的表空间。在定义容器之前，必须决定是否实用此选项。不能在您认为需要时才执行此操作；必须在需要复原数据库或表空间之前进行计划。
- 在 RESTORE 命令上使用 PARALLELISM 参数以指定并行度。
- 在 RESTORE 使用上使用 WITH num-buffers BUFFERS 参数以保证提供足够的缓冲区来满足该并行度。缓冲区数应比已有的目标媒体数与选择的并行度之和略大。

同时使用满足以下条件的复原缓冲区大小:

- 尽可能大。4 MB 或 8 MB (1024 或 2048 页) 比较合适。
- 至少等于要复原的表空间中（扩展数据块大小 \* 容器数之积）的最大者。
- 等于备份缓冲区大小或是其偶数倍。

---

## 对查询启用分区内并行性

还可以使用 GET DATABASE CONFIGURATION 和 GET DATABASE MANAGER CONFIGURATION 命令来了解特定数据库或数据库管理器配置文件中的个别条目的值。要修改特定数据库或数据库管理器配置文件中的个别条目，可分别使用 UPDATE DATABASE CONFIGURATION 和 UPDATE DATABASE MANAGER CONFIGURATION 命令。

影响分区内并行性的配置参数包括 *max\_querydegree* 和 *intra\_parallel* 数据库管理器参数以及 *dft\_degree* 数据库参数。

要使分区内查询并行性可用，必须修改一个或多个数据库配置参数、数据库管理器配置参数、预编译或绑定选项或专用寄存器。

### *intra\_parallel*

用来指定数据库管理器是否可使用分区内并行性的数据库管理器配置参数。缺省值是不使用分区内并行性。

### *max\_querydegree*

指定用于在此实例上运行的任何 SQL 语句的最大程度分区内并行性的数据库管理器配置参数。当在数据库分区中运行并行操作时，SQL 语句将不会使用大于此参数的数值。要使用 *max\_querydegree* 中的值，还必须将 *intra\_parallel* 配置参数设置为“YES”。此配置参数的缺省值为 -1。此值表示系统使用优化器确定的并行度；否则，使用用户指定的值。

### *dft\_degree*

提供 DEGREE 绑定选项和 CURRENT DEGREE 专用寄存器的缺省值的数据库配置参数。缺省值为 1。值 ANY 表示系统使用优化器确定的并行度。

## DEGREE

静态 SQL 的预编译或绑定选项。

## CURRENT DEGREE

动态 SQL 的专用寄存器。

---

## 数据服务器容量的管理

如果数据服务器的容量不能满足目前或将来的要求，那么可通过增大磁盘空间和创建其他容器或通过增加内存来扩充其容量。如果这些简单措施不能增加所需的容量，还可以考虑添加处理器或物理分区。当通过更改环境来调整系统时，应意识到这类更改可给数据库过程（如装入数据、备份和复原数据库）带来的影响。

### 添加处理器

如果具有单个处理器的单一分区数据库配置已被最大限度地使用，那么可能需要添加处理器或数据库分区。添加处理器可以获得更大的处理能力。在 SMP 系统中，多个处理器共享内存和存储系统资源。由于所有处理器均在一个系统内，因此不存在附加的开销费用（如铺设系统间的通信线路），也不会增加系统间的协调任务。一些实用程序（例如，Load、Backup 和 Restore）可以利用其他处理器。

**注：**某些操作系统（例如，Solaris 操作系统）可以动态地使处理器联机和脱机。

如果添加了处理器，请检查并修改那些决定处理器使用数目的数据库配置参数。下列数据库配置参数决定处理器的使用数目，可能需要进行更新：

- 缺省级别（dft\_degree）
- 最大并行度（max\_querydegree）
- 启用分区内并行性（intra\_parallel）

此外，还应评估那些决定应用程序如何执行并行处理的参数。

在使用 TCP/IP 进行通信的环境中，应检查 DB2TCPCONNMGRS 注册表变量的值。

### 添加物理分区

如果数据库管理器当前在分区数据库环境中，那么可以通过添加不同的单处理器物理分区或多处理器物理分区来增加数据存储空间和处理能力。每个数据库分区上的内存和存储系统资源不与其他数据库分区共享。虽然添加数据库分区可能会导致通信和任务协调问题，但这一选择所带来的好处是可在多个系统上平衡数据和用户访问。数据库管理器支持此环境。

无论数据库管理器系统处于运行状态还是停止状态，均可添加数据库分区。但是，如果在系统运行时添加数据库分区，那么必须在停止并重新启动系统后，才能将数据库迁移到新的数据库分区。

当添加一个新的数据库分区时，在该过程完成并将新服务器成功集成到系统之前，不能删除或创建使用新数据库分区的数据库。



---

## 快速通信管理器

### 快速通信管理器 ( Windows )

快速通信管理器 (FCM) 为属于同一实例的 DB2 服务器产品提供了通信支持。每个数据库分区服务器都有一个 FCM 发送方守护程序和一个 FCM 接收方守护程序，用于在数据库分区服务器之间提供通信，以处理代理程序请求和传递消息缓冲区。FCM 守护程序在您启动实例时启动。

如果数据库分区服务器之间的通信失败，或者它们重新建立通信，那么 FCM 线程将更新信息（可以用数据库系统监视器进行查询），并导致执行适当的操作（例如，回滚受影响的事务）。可以使用数据库系统监视器来帮助您设置 FCM 配置参数。

可以使用 *fcm\_num\_buffers* 数据库管理器配置参数来指定 FCM 消息缓冲区数，以及使用 *fcm\_num\_channels* 数据库管理器配置参数来指定 FCM 通道数。数据库管理器配置参数 *fcm\_num\_buffers* 和 *fcm\_num\_channels* 设置为 AUTOMATIC（缺省值）。当这些参数中的任何设置为 AUTOMATIC 时，FCM 就会监视资源使用情况，并逐渐释放资源。建议将这些参数设置为 AUTOMATIC。

### 快速通信管理器 ( Linux 和 UNIX )

快速通信管理器 (FCM) 为使用“数据库分区功能” (DPF) 的 DB2 服务器产品提供了通信支持。

对于多分区实例，每个数据库分区服务器都有一个 FCM 发送方守护程序和一个 FCM 接收方守护程序，用于在数据库分区服务器之间提供通信，以处理代理程序请求和传递消息缓冲区。FCM 守护程序在您启动多分区实例时启动。

如果数据库分区服务器之间的通信失败，或者它们重新建立通信，那么 FCM 守护程序将更新信息。您可以使用数据库系统监视器查询此信息。FCM 守护程序还会触发相应的操作。例如回滚受到影响的事务。可以使用数据库系统监视器来帮助您设置 FCM 配置参数。

可以使用 *fcm\_num\_buffers* 数据库管理器配置参数来指定 FCM 消息缓冲区数。此外，还可以使用 *fcm\_num\_channels* 数据库管理器配置参数来指定 FCM 通道数。数据库管理器配置参数 *fcm\_num\_buffers* 和 *fcm\_num\_channels* 设置为 AUTOMATIC（缺省值）。当任何这些参数设置为 AUTOMATIC 时，FCM 就会监视资源使用情况，并以递增方式释放资源。建议将这些参数设置为 AUTOMATIC。

### 使用 FCM 通信启用数据库分区之间的通信

在分区数据库环境中，数据库分区之间的大多数通信都是由快速通信管理器 (FCM) 来处理。

要在一个数据库分区上启用 FCM 并允许与其他数据库分区通信，必须在该数据库分区的 `etc` 目录的 `services` 文件中创建一个服务条目，如下所示。FCM 将使用指定的端口来通信。如果已在同一主机上定义了多个数据库分区，那么必须定义一个端口范围，如下所示。

在尝试为快速通信管理器（FCM）手动配置内存之前，建议从 FCM 缓冲区数（*fcm\_num\_buffers*）和 FCM 信道数（*fcm\_num\_channels*）的自动设置开始（缺省设置）。使用 FCM 活动的系统监视器数据来判定此设置是否合适。

### Windows 注意事项

如果正在 Windows 环境中使用 DB2 企业服务器版，那么会通过下列程序自动将 TCP/IP 端口范围添加至 *services* 文件：

- 安装程序，在创建实例或添加新的数据库分区时
- *db2icrt* 实用程序，在创建新实例时
- *db2ncrt* 实用程序，在计算机上添加第一个数据库分区时

服务条目的语法如下所示：

```
DB2_instance port/tcp #comment
```

#### DB2\_instance

*instance* 的值是数据库管理器实例的名称。该名称的所有字符必须为小写。假定实例名为 *db2puser*，那么应指定 *DB2\_db2puser*。

#### port/tcp

要为该数据库分区保留的 TCP/IP 端口。

#### #comment

想要与该条目关联的任何注释。注释之前必须加 # 符号。

如果 *etc* 目录的 *services* 文件是共享的，那么必须确保在该文件中分配的端口的数目大于或等于该实例中多个数据库分区的最大数目。当分配端口时，还要确保考虑了可以用作备份的任何处理器。

如果 *etc* 目录的 *services* 文件不是共享的，那么注意事项基本相同，但它有一个附加注意事项：必须确保为 DB2 数据库实例定义的条目在 *etc* 目录的所有 *services* 文件中都是相同的（而不适用于分区数据库环境的条目不必相同）。

如果在一个实例中的相同主机上有多个数据库分区，那么必须定义多个端口以供 FCM 使用。为此，在 *etc* 目录的 *services* 文件中包括两行，以指示正分配的端口的范围。第一行指定第一个端口，而第二行指示端口块的结束。在下列示例中，为实例 *sales* 分配了五个端口。这意味着该实例中不会有处理器具有多于五个的数据库分区。例如，

```
DB2_sales      9000/tcp
DB2_sales_END  9004/tcp
```

注：只能用大写字母来指定 END。还必须确保包括了两个下划线（*\_*）字符。

由于 FCM 基础结构利用 TCP 套接字和定向网络流量的方式，AIX 5.x 上的 FCM 用户应将内核参数“*tcp\_nodelayack*”设置为 1。

## 启用数据库分区服务器之间的通信（Linux 和 UNIX）

此任务描述了如何启用参与分区数据库系统的数据库分区服务器之间的通信。数据库分区服务器之间的通信由“快速通信管理器”（FCM）处理。要启用 FCM，必须在分区数据库系统中的每台计算机上的 */etc/services* 文件中保留一个端口或端口范围。

您的用户标识必须具有 root 用户权限。

您必须在参与实例的所有计算机上执行此任务。

为 FCM 保留的端口数目等于实例中由任何计算机主管或可能主管的数据库分区的最大数目。

在下列示例中，db2nodes.cfg 文件包含这些条目：

```
0 server1 0
1 server1 1
2 server2 0
3 server2 1
4 server2 2
5 server3 0
6 server3 1
7 server3 2
8 server3 3
```

假设 FCM 端口从 60000 开始编号。在此情况下：

- server1 对它的两个数据库分区使用两个端口（60000 和 60001）
- server2 对它的三个数据库分区使用三个端口（60000、60001 和 60002）
- server3 对它的四个数据库分区使用四个端口（60000、60001、60002 和 60003）

所有计算机均必须保留 60000、60001、60002 和 60003，因为这是实例中的任何计算机所需的最大端口范围。

如果使用诸如 High Availability Cluster Multi-Processing (HACMP™) 或 Tivoli System Automation (TSA) 之类的高可用性解决方案来将数据库分区从一台计算机故障转移到另一台计算机，那么必须解决潜在的端口需求。例如，如果计算机通常主管四个数据库分区，但另一计算机的两个数据库分区有可能故障转移到该计算机，那么您必须为该计算机规划六个端口。

当创建实例时，将在主计算机上保留某个端口范围。主计算机也称为实例拥有的计算机。但是，如果最初添加到 /etc/services 文件的端口范围不够满足需要，那么需要手工添加其他条目来扩展保留端口的范围。

要使用 /etc/services 在分区数据库环境中的服务器之间启用通信：

1. 作为具有 root 用户权限的用户登录主计算机（拥有实例的计算机）。
2. 创建实例。
3. 查看 /etc/services 文件中已保留的缺省端口范围。除了基本配置外，FCM 端口应类似于以下所示：

```
db2c_db2inst1      50000/tcp
#Add FCM port information
DB2_db2inst1      60000/tcp
  DB2_db2inst1_1   60001/tcp
  DB2_db2inst1_2   60002/tcp
  DB2_db2inst1_END 60003/tcp
```

缺省情况下，第一个端口（50000）保留给连接请求使用，而 60000 以上的前四个可用端口保留给 FCM 通信使用。一个端口用于实例拥有的数据库分区服务器，三个端口用于逻辑数据库分区服务器，可以在完成安装后选择将其添加至计算机。

端口范围必须包括起始条目和结束条目。中间条目为可选项。显式（包括中间值）对于防止其他应用程序使用这些端口很有用，但这些条目未经数据库管理器验证。

DB2 端口条目的格式如下：

```
DB2_instance_name_suffix port_number/tcp # comment
```

其中:

- *instance\_name* 是分区实例的名称。
  - *suffix* 不用于第一个 FCM 端口。中间条目是介于最低和最高端口之间的端口。如果端口范围包括介于第一个 FCM 端口和最后一个 FCM 端口之间的中间条目,那么 *suffix* 包含一个整数,其他每个端口依据此整数递增。例如,第二个端口编号为 1,那么第三个端口编号为 2,以此类推以确保唯一性。END 一词必须用作最后条目的 *suffix*。
  - *port\_number* 是为数据库分区服务器通信保留的端口号。
  - *comment* 是用于描述条目的可选注释。
4. 确保您保留了足够的端口供 FCM 通信使用。如果保留端口范围不够用,那么将新的条目添加至该文件。
  5. 作为 root 用户登录参与实例的每台计算机,并将相同的条目添加至 `/etc/services` 文件。

---

## 第 10 章 创建和管理分区数据库环境

---

### 初始数据库分区组

最初创建数据库时，会为 db2nodes.cfg 文件中指定的所有数据库分区创建数据库分区。可以使用 ADD DBPARTITIONNUM 和 DROP DBPARTITIONNUM VERIFY 命令来添加或除去其他数据库分区。

定义了三个数据库分区组：

- 用于容纳 SYSCATSPACE 表空间的 IBMCATGROUP，它保存系统目录表
- 用于容纳 TEMPSPACE1 表空间的 IBMTEMPGROUP，它保存数据库处理期间创建的临时表
- 用于容纳 USERSPACE1 表空间的 IBMDEFAULTGROUP，缺省情况下它保存用户表和索引

---

### 创建数据库分区组

可使用 CREATE DATABASE PARTITION GROUP 语句创建数据库分区组。此语句指定将来存放表空间容器和表数据的一组数据库分区。

计算机和系统必须可用且能够处理分区数据库环境。您已经购买并安装了 DB2 企业服务器版。该数据库必须存在。

此语句还可以：

- 为数据库分区组创建分布图。
- 生成分布图标识。
- 将记录插入下列目录表：
  - SYSCAT.DBPARTITIONGROUPS
  - SYSCAT.PARTITIONMAPS
  - SYSCAT.DBPARTITIONGROUPDEF

要使用控制中心创建数据库分区组：

1. 展开对象树，直到您看到**数据库分区组**文件夹。
2. 右键单击**数据库分区组**文件夹，并从弹出菜单中选择**创建**。
3. 在“创建数据库分区组”窗口中，填写信息，使用箭头将节点从“可用的节点”框移至“选择的数据库分区”框，然后单击**确定**。

要使用命令行来创建数据库分区组，请输入：

```
CREATE DATABASE PARTITION GROUP <name> ON PARTITIONS (<value>,<value>)
```

例如，假定要在数据库中的一部分数据库分区上装入一些表。可使用以下命令创建一个数据库分区组，它包含至少由三个（0 至 2）数据库分区组成的数据库中的两个数据库分区（1 和 2）：

```
CREATE DATABASE PARTITION GROUP mixng12 ON PARTITIONS (1,2)
```

CREATE DATABASE 命令或 sqlcrea() API 还会创建缺省系统数据库分区组 IBMDEFAULTGROUP、IBMCATGROUP 和 IBMTEMPGROUP。

---

## 数据库分区组中的表空间

通过将表空间放在多分区数据库分区组中，就将该表空间内的所有表划分或分区到该数据库分区组的每个数据库分区中。


由此该表空间被创建到了一个数据库分区组中。一旦位于某个数据库分区组中，该表空间就必须保留在该处；而不能更改至另一数据库分区组。CREATE TABLESPACE 语句用于将表空间与数据库分区组关联。

---

## 管理数据库分区

可以使用控制中心中的“分区”视图来启动和停止分区、删除和跟踪分区以及显示诊断日志。

要使用数据库分区，将需要连接至实例的权限。具有 SYSADM 或 DBADM 权限的任何人都可以授予您访问特定实例的权限。要查看 DB2 日志，将需要连接至实例的权限。具有 SYSADM 或 DBADM 权限的任何人都可以授予您访问特定实例的权限。

 如果 IBM 支持机构要求这样做，那么使用他们指示的选项运行 trace 实用程序。trace 实用程序记录关于 DB2 操作的信息并将此信息格式化为可读格式。有关更多信息，请参阅 db2trc - 跟踪: DB2 主题。

**注意:** 只应在 DB2 客户服务人员或技术支持代表指导您使用时才使用跟踪工具。

使用“诊断日志”窗口来查看 DB2 trace 实用程序所记录的文本信息。

“分区”视图显示下列信息:

**节点号** 此列包含图标和节点号。节点号是唯一编号，可以是 0 到 999。这些号码存储在 db2nodes.cfg 文件中。节点号按递增顺序显示，虽然序列中可能有间隔。

节点号一经指定，就不能更改。此保护确保分布映射（它详细描述数据是如何分区的）中的信息不受损害。

**主机名** 主机名是快速通信管理器（FCM）用于内部通信的 IP 地址。（但是，如果指定了转换名，那么 FCM 使用转换名。在此情况下，主机名仅用于 DB2START、DB2STOP 和 db2\_all。主机名存储在 db2nodes.cfg 文件中。

**端口号** 端口号是节点的逻辑端口号。此号码与数据库管理器实例名一起用来标识 etc/services 文件中的 TCP/IP 服务名称条目。此号存储在 db2nodes.cfg 文件中。

IP 地址（主机名）和逻辑端口的组合用作熟知地址，并且必须在支持节点间通信连接的所有应用程序中是唯一的。

对于每个显示的主机名，一个端口号将为 0。端口号 0 指示客户机所连接的主机上的缺省节点。（要覆盖此行为，可使用 db2profile 脚本中的 DB2NODE 环境变量。）

**转换名** 转换名（或网络名）用来支持具有多个活动的 TCP/IP 接口的主机，每个接口都具有其自己的主机名。它还用于共享高速交换机的节点之间的快速通信（也称

为 FCM)。转换名存储在 db2nodes.cfg 文件中。如果在 db2nodes.cfg 文件中未指定转换名，那么转换名与主机名相同。

### 资源集名称

资源集名称定义应启动节点的操作系统的资源。资源集名称用于进程亲缘关系支持、用于多逻辑节点 (MLN) 并随字符串类型字段 (以前称为 quadname) 提供。

1. 打开“分区”视图：从控制中心展开对象树，直到找到想要查看的分区实例为止。右键单击想要的实例并从弹出菜单中选择 **打开 -> 分区**。“分区”视图将打开。
2. 要启动分区：突出显示一个或多个分区并选择 **分区 -> 启动**。这就启动了选择的分区。
3. 要停止分区：突出显示一个或多个分区并选择 **分区 -> 停止**。这就停止了选择的分区。
4. 要在分区上运行 trace 实用程序：
  - a. 打开“DB2 跟踪”窗口：突出显示分区并选择 **分区 -> 服务 -> 跟踪**。“DB2 跟踪”窗口将打开。
  - b. 指定跟踪选项。
  - c. 单击 **开始** 以开始记录信息，单击 **停止** 以停止记录信息，单击 **另存为** 以将信息保存到文件中。
  - d. 可选：查看日志。
  - e. 如果要求这样做，请将跟踪输出发送至 IBM 支持机构。

## 在分区数据库环境中添加数据库分区

可在分区数据库系统运行或停止时，向其添加数据库分区。因为添加新服务器可能要花费较长时间，因此应在数据库管理器已经运行时来添加。

使用 `ADD DBPARTITIONNUM` 命令来将数据库分区添加到系统。可以按下列方式调用此命令：

- 作为 `db2start` 中的一个选项
- 使用命令行处理器 `ADD DBPARTITIONNUM` 命令
- 使用 API 函数 `sqlleaddn`
- 使用 API 函数 `sqllepstart`

如果您的系统已停止，使用 `db2start`。如果系统在运行，可以使用任何其他选项。

当使用 `ADD DBPARTITIONNUM` 命令将一个数据库分区添加至系统时，会将实例中的所有现有的数据库扩充到新数据库分区。还可指定要将临时表空间的哪些容器用于数据库。这些容器可能有以下特征：

- 与为每个数据库的目录分区定义的那些容器相同。（这是缺省情况。）
- 与为另一个数据库分区定义的那些容器相同。
- 根本没有创建。必须使用 `ALTER TABLESPACE` 语句来将临时表空间容器添加至每个数据库之后，才能使用数据库。

在改变一个或多个数据库分区组以包括新数据库分区之前，不能使用新数据库分区上的数据库来包含数据。

不能简单地通过对系统添加数据库分区来将单分区数据库更改为多分区数据库。这是因为跨数据库分区再分发数据需要每个受影响的表上的分布键。分布键是在多分区数据库中创建表时自动生成的。在单分区数据库中，可使用 CREATE TABLE 或 ALTER TABLE SQL 语句显式地创建分布键。

**注：**如果系统中没有定义任何数据库且您正在 UNIX 操作系统上运行“企业服务器版”，那么编辑 db2nodes.cfg 文件以添加新的数据库分区定义；不要使用所述的任何一个过程，因为它们只有在存在数据库时才适用。

**Windows 注意事项：**如果在 Windows 上使用“企业服务器版”，并且该实例没有数据库，那么应使用 DB2NCRT 命令来调整该数据库系统。但是，如果您已有数据库，使用 DB2START ADDNODE 命令，以确保在调整系统时为每个现有的数据库创建一个数据库分区。在 Windows 上，永远都不应手动编辑节点配置文件（db2nodes.cfg），因为这会给该文件带来不一致性。

## 将数据库分区添加到正在运行的数据库系统

当分区数据库环境正在运行且应用程序已与数据库连接时，可以将新数据库分区添加至该环境。但是，只有将数据库管理器关闭并重新启动之后，新添加的服务器才能供所有数据库使用。

要使用命令行将数据库分区添加到运行的数据库管理器：

1. 在任何现有的数据库分区上，运行 DB2START 命令。

在所有平台上，为 DBPARTITIONNUM、ADD DBPARTITIONNUM、HOSTNAME、PORT 和 NETNAME 参数指定新数据库分区值。在 Windows 平台上，也指定 COMPUTER、USER 和 PASSWORD 参数。

您还可为需要在该数据库中创建的任何“临时表空间容器定义”指定源。如果未提供表空间信息，那么从每个数据库的目录分区检索临时表空间容器的定义。

当完成 DB2START 命令时，停止新服务器。

2. 通过运行 DB2STOP 命令来停止所有数据库分区上的数据库管理器。

当停止系统中的所有数据库分区时，会更新节点配置文件以包括新的数据库分区。在执行 DB2STOP 之前，不使用新服务器信息更新节点配置文件。这确保 ADD DBPARTITIONNUM 命令（当对 DB2START 命令指定 ADDNODE 参数时调用的命令）在正确的数据库分区上运行。当该实用程序结束时，停止新服务器分区。

3. 通过运行 DB2START 命令来启动数据库管理器。

现在，新添加的数据库分区与系统的其余部分一起启动。

当系统中的所有数据库分区正运行时，可执行系统范围内的活动，如创建或删除数据库。

**注：**可能必须对所有数据库分区服务器发出 DB2START 命令两次，才能访问新的 db2nodes.cfg 文件。

4. 可选：在新数据库分区上备份所有数据库。
5. 可选：将数据再分发至新数据库分区。



## 将数据库分区添加至已停止的数据库系统（Windows）

当分区数据库系统停止时，可以将新数据库分区添加至其中。当再次启动数据库管理器时，新添加的数据库分区可用于所有数据库。

必须先安装新服务器，然后才可以在该服务器上创建数据库分区。

要使用命令行将数据库分区添加到已停止的分区数据库服务器：

1. 发出 `DB2STOP` 来停止所有的数据库分区。
2. 在新服务器上运行 `ADD DBPARTITIONNUM` 命令。

存在于系统中的每个数据库在本地创建一个数据库分区。将新数据库分区的数据库参数设置为缺省值，并且在将数据移动至其中之前，每个数据库分区保持为空。更新数据库配置参数值，以便与其他数据库分区上的值相匹配。

3. 运行 `DB2START` 命令启动数据库系统。注意，新服务器安装期间，节点配置文件（`cfg`）已更新为包括新服务器。
4. 按如下所示更新新数据库分区上的配置文件：
  - a. 在任何现有数据库分区上，运行 `DB2START` 命令。

为 `DBPARTITIONNUM`、`ADDDBPARTITIONNUM`、`HOSTNAME`、`PORT` 和 `NETNAME` 参数以及 `COMPUTER`、`USER` 和 `PASSWORD` 参数指定新数据库分区值。

您还可为需要在该数据库中创建的任何“临时表空间容器定义”指定源。如果未提供表空间信息，那么从每个数据库的目录分区检索临时表空间容器的定义。

当完成 `DB2START` 命令时，停止新服务器。

- b. 通过运行 `DB2STOP` 命令来停止整个数据库管理器。

当停止系统中的所有数据库分区时，会更新节点配置文件以包括新的数据库分区。在执行 `DB2STOP` 之前，不使用新服务器信息更新节点配置文件。这确保 `ADD DBPARTITIONNUM` 命令（当对 `DB2START` 命令指定 `ADDDBPARTITIONNUM` 参数时调用的命令）在正确的数据库分区上运行。当该实用程序结束时，停止新服务器分区。

5. 通过运行 `DB2START` 命令来启动数据库管理器。

现在，新添加的数据库分区与系统的其余部分一起启动。

当系统中的所有数据库分区正运行时，可执行系统范围内的活动，如创建或删除数据库。

**注：**可能必须对所有数据库分区服务器发出 `DB2START` 命令两次，才能访问新的 `db2nodes.cfg` 文件。

6. 可选：在新数据库分区上备份所有数据库。
7. 可选：将数据再分发至新数据库分区。

## 将数据库分区添加到已停止的数据库系统（UNIX）

当分区数据库系统停止时，可以将新的数据库分区添加到该系统中。当再次启动数据库管理器时，新添加的数据库分区可用于所有数据库。

如果新服务器不存在，那么必须安装该服务器，包括下列任务：

- 使可执行文件可访问（使用共享文件系统安装或本地副本）
- 使操作系统文件与现有处理器上的操作系统文件同步。
- 确保 `sqllib` 目录可作为共享文件系统访问。
- 确保将相关的操作系统参数（如最大进程数）设置为适当的值。

向名称服务器注册该主机名，或者在所有数据库分区的 `etc` 目录中的主机文件中注册该主机名。

要使用命令行将数据库分区添加到已停止的分区数据库服务器：

1. 发出 `DB2STOP` 来停止所有的数据库分区。
2. 在新服务器上运行 `ADD DBPARTITIONNUM` 命令。

存在于系统中的每个数据库在本地创建一个数据库分区。将新数据库分区的数据库参数设置为缺省值，并且在将数据移动至其中之前，每个数据库分区保持为空。更新数据库配置参数值，以便与其他数据库分区上的值相匹配。

3. 运行 `DB2START` 命令启动数据库系统。注意，新服务器安装期间，节点配置文件（`cfg`）已更新为包括新服务器。
4. 按如下所示更新新数据库分区上的配置文件：
  - a. 在任何现有的数据库分区上，运行 `DB2START` 命令。

为 `DBPARTITIONNUM`、`ADDDBPARTITIONNUM`、`HOSTNAME`、`PORT` 和 `NETNAME` 参数以及 `COMPUTER`、`USER` 和 `PASSWORD` 参数指定新数据库分区值。

您还可为需要在该数据库中创建的任何“临时表空间容器定义”指定源。如果未提供表空间信息，那么从每个数据库的目录分区检索临时表空间容器的定义。

当完成 `DB2START` 命令时，停止新服务器。

- b. 通过运行 `DB2STOP` 命令来停止整个数据库管理器。

当停止系统中的所有数据库分区时，会更新节点配置文件以包括新的数据库分区。在执行 `DB2STOP` 之前，不使用新服务器信息更新节点配置文件。这确保 `ADD DBPARTITIONNUM` 命令（当对 `DB2START` 命令指定 `ADDDBPARTITIONNUM` 参数时调用的命令）在正确的数据库分区上运行。当该实用程序结束时，停止新服务器分区。

5. 通过运行 `DB2START` 命令来启动数据库管理器。

现在，新添加的数据库分区与系统的其余部分一起启动。

当系统中的所有数据库分区正运行时，可执行系统范围内的活动，如创建或删除数据库。

**注：**可能必须对所有数据库分区服务器发出 `DB2START` 命令两次，才能访问新的 `db2nodes.cfg` 文件。

6. 可选：在新数据库分区上备份所有数据库。
7. 可选：将数据再分发至新数据库分区。

也可以手动更新配置文件，如下所示：

1. 编辑 `db2nodes.cfg` 文件，并将新数据库分区添加至该文件。
2. 发出以下命令来启动新数据库分区：`DB2START DBPARTITIONNUM partitionnum`

将您分配给该新数据库分区的编号指定为 `nodenum` 的值。

3. 如果新服务器是一个逻辑分区（即，它不是数据库分区 0），那么使用 `db2set` 命令来更新 `DBPARTITIONNUM` 注册表变量。指定要添加的数据库分区的号码。
4. 在新数据库分区上运行 `ADD NODE` 命令。

此命令还为已存在于系统中的每个数据库在本地创建一个数据库分区。将新数据库分区的数据库参数设置为缺省值，并且在将数据移动至其中之前，每个数据库分区保持为空。更新数据库配置参数值，以便与其他数据库分区上的值相匹配。

5. 当 `ADD DBPARTITIONNUM` 命令完成时，发出 `DB2START` 命令来启动系统中的其他数据库分区。

在成功启动所有数据库分区之前，不要执行任何系统范围内的活动，如创建或删除数据库。

## 添加数据库分区时的错误恢复

因为数据库管理器将创建系统缓冲池以对所有缓冲池页大小提供缺省自动支持，所以添加数据库分区不会因不存在的缓冲池而失败。但是，因为系统缓冲池非常小，所以如果使用其中一个系统缓冲池，可能会严重影响性能。如果使用系统缓冲池，管理通知记录中写入一条消息。

在下列情况的数据库分区添加方案中使用系统缓冲池：

- 将数据库分区添加到分区数据库环境中，该数据库具与缺省值 4KB 不同的页大小的一个或多个系统临时表空间。当创建数据库分区时，仅 `IBMDEFAULTDP` 缓冲池存在，此缓冲池页大小为 4 KB。

请考虑以下示例：

1. 使用 `db2start` 命令将数据库分区添加到当前的多分区数据库：

```
DB2START DBPARTITIONNUM 2 ADD DBPARTITIONNUM HOSTNAME newhost PORT 2
```

2. 在使用新的数据库分区描述手动更新 `db2nodes.cfg` 文件之后，使用 `ADD DBPARTITIONNUM` 命令。

防止这些问题发生的一种方法是在 `ADD NODE` 或 `db2start` 命令上指定 `WITHOUT TABLESPACES` 子句。此后，需要使用 `CREATE BUFFERPOOL` 语句来创建要使用的缓冲池，并使用 `ALTER TABLESPACE` 语句将系统临时表空间与缓冲池相关联。

- 将数据库分区添加到现有的数据库分区组，该数据库分区组具与缺省页大小 4KB 不同的页大小的一个或多个表空间。发生此情况的原因是尚未对表空间激活在新数据库分区上创建的非缺省页大小缓冲池。

**注：**在先前版本中，此命令使用了 `NODEGROUP` 关键字，而不是 `DATABASE PARTITION GROUP` 关键字。

请考虑以下示例：

- 可以使用 `ALTER DATABASE PARTITION GROUP` 语句来向数据库分区组添加数据库分区，如下所示：

```
DB2START
CONNECT TO mpp1
ALTER DATABASE PARTITION GROUP ng1 ADD NODE (2)
```

防止此问题发生的一种方法是在发出 ALTER DATABASE PARTITION GROUP 语句之前，为每个页大小创建缓冲池，然后重新连接至数据库：

```
DB2START
CONNECT TO mpp1
CREATE BUFFERPOOL bp1 SIZE 1000 PAGESIZE 8192
CONNECT RESET
CONNECT TO mpp1
ALTER DATABASE PARTITION GROUP ng1 ADD NODE (2)
```

注：如果数据库分区组具有页大小为缺省值的表空间，那么返回消息 SQL1759W。

## 删除数据库分区

您可以删除任何数据库都未在使用的数据库分区，并为其他的用户空出计算机。

通过发出 DROP DBPARTITIONNUM VERIFY 命令或 sqledrpn API，验证是否未使用此数据库分区。

- 如果接收到消息 SQL6034W（未在任何数据库中使用数据库分区），那么可以删除该数据库分区。
- 如果接收到消息 SQL6035W（数据库正在使用数据库分区），那么使用 REDISTRIBUTE NODEGROUP 命令，将要删除的数据库分区中的数据再分发到数据库别名不同的其他数据库分区。

还应确保所有事务（此数据库分区是其协调程序）已成功落实或回滚。这可能需要在其他服务器上执行崩溃恢复。例如，如果删除协调程序分区，且参与事务的另一个数据库分区在删除协调程序分区之前崩溃，那么该崩溃的数据库分区将不能查询协调程序分区来获取任何不确定事务的输出。

要使用命令行删除数据库分区，发出带 DROP NODENUM 参数的 DB2STOP 命令来删除该数据库分区。在成功完成命令后，停止系统。然后，使用 DB2START 命令启动数据库管理器。

## 列示实例中的数据库分区服务器

在 Windows 上，使用 db2nlist 命令来获取参与实例的数据库分区服务器的列表。

按如下所示使用该命令：

```
db2nlist
```

当按以上所示使用此命令时，缺省实例是当前实例（由 DB2INSTANCE 环境变量设置）。要指定特定的实例，可使用以下命令指定该实例：

```
db2nlist /i:instName
```

其中 instName 是想要的特定实例名。

（可选）也可使用以下命令请求每个数据库分区服务器的状态：

```
db2nlist /s
```

每个数据库分区服务器的状态可能是：正在启动、正在运行、正在停止或已停止。

## 将数据库分区服务器添加至实例 ( Windows )

在 Windows 上, 使用 `db2ncrt` 命令来将数据库分区服务器添加至实例。

**注:** 如果此实例已包含数据库, 那么不要使用 `db2ncrt` 命令。而使用 `db2start addnode` 命令。这确保可正确地将该数据库添加至新的数据库分区服务器。**不要编辑** `db2nodes.cfg` 文件, 因为更改文件可能导致分区数据库环境中的不一致性。

该命令具有下列必需参数:

```
db2ncrt /n:node_number
        /u:username,password
        /p:logical_port
```

**/n:** 用于标识数据库分区服务器的唯一数据库分区号。该号码可以是按递升顺序排列的 1 到 999 中的任何一个值。

**/u:** DB2(R) 服务的登录帐户名和密码。

**/p:logical\_port**

用于数据库分区服务器的逻辑端口号 ( 若逻辑端口不是零 ( 0 ) )。如果不指定, 那么将逻辑端口号指定为 0。

仅当在计算机上创建第一个数据库分区时, 逻辑端口参数才是可选的。如果创建逻辑数据库分区, 那么必须指定此参数并选择一个没有在使用的逻辑端口号。有几项限制:

- 在每台计算机上, 都必须要有有一个逻辑端口为 0 的数据库分区服务器。
- 端口号不能超过 `%SystemRoot%\system32\drivers\etc` 目录中的 `services` 文件中为 FCM 通信保留的端口范围。例如, 如果为当前实例保留四个端口, 那么最大端口号将是 3 ( 端口 1、2 和 3; 端口 0 用于缺省逻辑数据库分区)。端口范围是在将 `db2icrt` 命令与 `/r:base_port, end_port` 参数配合使用时定义的。

还有几个可选参数:

**/g:network\_name**

指定数据库分区服务器的网络名。如果不指定此参数, 那么 DB2 使用它在系统上检测到的第一个 IP 地址。

如果计算机上有多个 IP 地址, 且您想要对数据库分区服务器指定特定 IP 地址, 那么使用此参数。可输入使用网络名或 IP 地址的 `network_name` 参数。

**/h:host\_name**

FCM 用于内部通信的 TCP/IP 主机名 ( 如果该主机名不是本地主机名 )。如果在远程计算机上添加数据库分区服务器, 那么此参数是必需的。

**/i:instance\_name**

实例名; 缺省值为当前实例。

**/m:computer\_name**

该数据库分区所在的 Windows 工作站的计算机名称; 缺省名称是本地计算机的计算机名称。

**/o:instance\_owning\_computer**

作为实例拥有的计算机的计算机名称; 缺省值是本地计算机的计算机名称。当在任何非实例拥有的计算机上调用 `db2ncrt` 命令时, 此参数是必需的。

例如，如果要向实例拥有的计算机 MYMACHIN 上的实例 TESTMPP 添加新的数据库分区服务器（以便运行多逻辑数据库分区），且您想要让这个新数据库分区成为使用逻辑端口 1 的数据库分区 2，那么输入：

```
db2ncrt /n:2 /p:1 /u:my_id,my_pword /i:TESTMPP
/M:TEST /o:MYMACHIN
```

## 使用“添加分区”向导将数据库分区添加至实例

使用“添加分区”向导来创建分区并将其添加到一个或多个数据库分区组中。首先向实例添加新分区并将该分区指定给一个或多个数据库分区组，然后进行更高级的选择。

要使用数据库分区组，您必须具有 SYSADM 或 DBADM 权限。

1. 打开“添加分区”向导：
  - a. 从“控制中心”展开对象树，直到找到要使用的实例对象为止。右键单击该对象，并在弹出菜单中单击“添加分区”。“添加分区”启动板将打开。
  - b. 单击添加分区按钮。“添加分区”向导将打开。
2. 完成每个适用的向导页。单击第一页上的向导概述链接以获取更多信息。在对向导填写了添加分区所需的足够信息后，完成按钮才可用。

## 更改数据库分区 ( Windows )

在 Windows 上，可使用 db2nchg 命令来更改数据库分区。

- 将数据库分区从一台计算机移至另一台计算机。
- 更改计算机的 TCP/IP 主机名。

如果打算使用多个网络适配器，那么必须使用此命令为 db2nodes.cfg 文件中的“netname”字段指定 TCP/IP 地址。

- 使用另一逻辑端口号。
- 对数据库分区服务器使用另一名称。

该命令具有以下必需参数：

```
db2nchg /n:node_number
```

参数 /n: 是您想要更改的数据库分区服务器的配置的数目。此参数是必需的。

可选参数包括：

### **/i:instance\_name**

指定此数据库分区服务器所参与的实例。如果未指定此参数，那么缺省值是当前实例。

### **/u:username,password**

更改 DB2(R) 数据库服务的登录帐户名和密码。如果未指定此参数，那么登录帐户和密码保持不变。

### **/p:logical\_port**

更改数据库分区服务器的逻辑端口。如果将数据库分区服务器移至另一台计算机，那么必须指定此参数。如果未指定此参数，那么逻辑端口号保持不变。

### **/h:host\_name**

更改 FCM 用于内部通信的 TCP/IP 主机名。如果未指定此参数，那么主机名不更改。

### **/m:computer\_name**

将数据库分区服务器移至另一台计算机。仅当实例中没有现有数据库时，才可移动数据库分区服务器。

### **/g:network\_name**

更改数据库分区服务器的网络名。

如果计算机上有多个 IP 地址，且您想要对数据库分区服务器使用特定 IP 地址，那么使用此参数。可使用网络名或 IP 地址输入 `network_name`。

例如，要将指定给分区 2（它参与实例 TESTMPP）的逻辑端口更改为使用逻辑端口 3，那么输入以下命令：

```
db2nchg /n:2 /i:TESTMPP /p:3
```

DB2 数据库管理器提供访问远程计算机上实例级的 DB2 数据库系统注册表变量的能力。当前，以三种不同的级别存储 DB2 数据库系统注册表变量：机器或全局级、实例级和数据库分区级。对于以实例级（包括数据库分区级）存储的注册表变量，可使用 DB2REMOTEPEG 将其重定向到其他计算机。如果设置了 DB2REMOTEPEG，那么 DB2 数据库管理器将从 DB2REMOTEPEG 所指向的计算机访问 DB2 数据库系统注册表变量。db2set 命令会显示为：

```
db2set DB2REMOTEPEG=<remote workstation>
```

其中 <remote workstation> 是远程工作站名称。

#### **注：**

- 应该小心设置此选项，因为所有 DB2 数据库实例概要文件和实例列表都将位于指定的远程计算机名称上。
- 如果您的环境包括域中的用户，那么确保与 DB2 实例服务相关联的登录帐户是域帐户。这样可以确保该 DB2 实例具有适当的特权来枚举域级别的组。

此功能可与设置 DBINSTPROF 结合起来使用，以指向包含该注册表的同一台计算机上的远程 LAN 驱动器。

## **向数据库分区上的 SMS 表空间添加容器**

只能向数据库分区上当前没有任何容器的 SMS 表空间添加容器。

要使用命令行来向 SMS 表空间添加容器，请输入以下内容：

```
ALTER TABLESPACE <name>
  ADD ('<path>')
  ON DBPARTITIONNUM (<database partition_number>)
```

按编号指定的数据库分区以及数据库分区范围内的每个分区都必须存在于定义表空间的数据库分区组中。database partition\_number 只能显式地出现或者只出现在语句的某个 *db-partitions* 子句的范围内。

以下示例显示在基于 UNIX 的操作系统上，如何将表空间“plans”所使用的数据库分区组的 3 号数据库分区添加新容器：

```
ALTER TABLESPACE plans
  ADD ('/dev/rhdisk0')
  ON DBPARTITIONNUM (3)
```

## 从实例中删除数据库分区（Windows）

在 Windows 上，使用 `db2ndrop` 命令从没有数据库的实例中删除数据库分区服务器。如果删除一个数据库分区服务器，那么它的数据库分区号可以再次用于新的数据库分区服务器。

当从实例中删除数据库分区服务器时务必小心。如果从该实例中删除实例拥有的数据库分区服务器 0，那么该实例将变得不可用。如果要删除该实例，那么使用 `db2idrop` 命令。

**注：**如果此实例包含数据库，那么不要使用 `db2ndrop` 命令。而使用 `db2stop drop nodenum` 命令。这确保可正确地数据库分区中除去该数据库。**不要编辑** `db2nodes.cfg` 文件，因为更改文件可能导致分区数据库环境中的不一致性。

如果要从运行多个逻辑数据库分区的计算机中删除分配了逻辑端口 0 的数据库分区，那么在删除分配了逻辑端口 0 的数据库分区之前，必须删除分配给其他逻辑端口的所有其他数据库分区。每个数据库分区服务器都必须带有一个分配给逻辑端口 0 的数据库分区。

该命令具有下列参数：

```
db2ndrop /n:node_number /i:instance_name
```

**/n:** 用于标识数据库分区服务器的唯一数据库分区号。这是一个必需参数。该号码可以是按递升顺序排列的 0 到 999 中的任何一个值。记住数据库分区 0 表示实例拥有的计算机。

**/i:instance\_name**

实例名。这是一个可选参数。如果未提供此参数的值，那么缺省值是当前实例（由 `DB2INSTANCE` 注册表变量设置）。

## 使用“删除分区”启动板从实例中删除数据库分区

使用“删除分区”启动板来指导您完成从数据库分区组删除数据库分区、在数据库分区组中再分发数据和从实例删除分区所必需的任务。

**注：**从数据库分区组删除数据库分区时，数据库分区不会立即被删除。而是对想要删除的数据库分区标上标志，以便可以在数据库分区组中再分发数据时从这些分区中除掉数据。

建议在数据库分区组中再分发数据之前和之后备份实例中的所有数据库。如果不备份数据库，那么可能会在数据库受到毁坏后无法对其进行恢复。

要使用“删除分区”启动板来删除分区：

1. 可选：使用“备份”向导来备份数据。
2. 打开“删除分区”启动板：
  - a. 打开“分区”窗口：从“控制中心”中展开对象树，直到找到要查看的分区的实例为止。右键单击想要的实例并从弹出菜单中选择**打开数据库分区服务器**。“分区”窗口对选择的实例打开。
  - b. 选择想要删除的分区。
  - c. 右键单击选择的分区，并单击弹出菜单中的**删除**。“删除分区”启动板将打开。
3. 从数据库分区组中删除数据库分区：



- a. 确认您要从数据库分区组中删除的数据库分区。

**注:**

- 必须从数据库分区组删除数据库分区，才能从实例删除分区。
  - 此操作并不会立即删除数据库分区。而是对想要删除的数据库分区标上标志，以便可以在数据库分区组中再分发数据时从这些分区中除掉数据。
4. 在数据库分区组中再分发数据。
  5. 从实例中删除分区:
    - a. 打开“从实例中删除分区确认”窗口:
      - 打开如上所述的“分区”窗口。
      - 选择想要删除的分区。
      - 右键单击所选分区并单击弹出菜单中的“删除”。“删除分区”启动板将打开。
      - 单击**从实例中删除分区**按钮。“从实例中删除分区确认”窗口将打开。
    - b. 在**删除**列中，验证是否想要删除选择的实例的分区。
    - c. 单击**确定**以打开一个窗口，在该窗口中可以安排想要何时删除分区。
  6. 可选：使用“备份”向导来备份数据。

---

## 方案：对数据库中的数据进行分区

此方案说明如何将数据库分区添加至数据库并在数据库分区之间再分发数据。演示了 REDISTRIBUTE 命令的变体以说明如何使用新选项将数据再分发到数据库分区组内的不同表集中。

**方案：** 数据库 DBPG1 有 2 个指定为 (0, 1) 的数据库分区和一个数据库分区组定义 (0, 1)。

在数据库分区组 DBPG\_1 上定义了下列表空间:

- 表空间 TS1 - 此表空间有两个表 T1 和 T2
- 表空间 TS2 - 此表空间定义了三个表 T3、T4 和 T5

### 在 DBPG1 中的数据库分区间分发数据:

要将三个新数据库分区添加至该数据库，请发出下列命令:

```
DB2START DBPARTITIONNUM 3 ADD DBPARTITIONNUM HOSTNAME <HOSTNAME3>  
PORT <PORT3>;
```

```
DB2START DBPARTITIONNUM 4 ADD DBPARTITIONNUM HOSTNAME <HOSTNAME4>  
PORT <PORT4>;
```

```
DB2START DBPARTITIONNUM 5 ADD DBPARTITIONNUM HOSTNAME <HOSTNAME5>  
PORT <PORT5>;
```

```
DB2STOP;
```

```
DB2START;
```

以下再分发命令将 DBPG\_1 定义从 (0, 1) 更改为 (0, 1, 3, 4, 5) 并且还再分发数据:

```
DB2 REDISTRIBUTE DATABASE PARTITION GROUP DBPG_1 NOT ROLLFORWARD  
RECOVERABLE  
UNIFORM ADD DBPARTITIONNUM (3 TO 5) STOP AT 2006-03-10-07.00.00.000000;
```

假定该命令对表 T1、T2 和 T3 成功运行，然后由于指定了 STOP AT 选项而停止。

要异常终止数据库分区组的数据再分发并还原对表 T1、T2 和 T3 所作的更改，请发出以下命令：

```
DB2 REDISTRIBUTE DATABASE PARTITION GROUP DBPG_1 NOT ROLLFORWARD  
RECOVERABLE ABORT;
```

如果在数据再分发期间发生错误或出现中断，并且您不希望继续执行再分发操作，那么您可以异常终止数据再分发。对于此方案，假定此命令运行成功并且表 T1 和 T2 已还原为它们的原始状态。

要在只有 5000 个 4K 页面作为数据缓冲区的情况下再分发 T5 和 T4：

```
DB2 REDISTRIBUTE DATABASE PARTITION GROUP DBPG_1 NOT ROLLFORWARD  
RECOVERABLE  
UNIFORM ADD DBPARTITIONNUM (3 TO 5) TABLE (T5, T4) ONLY DATA BUFFER 5000  
TABLE PARALLEL TABLE 2;
```

如果该命令运行成功，那么说明已成功再分发表 T4 和 T5 中的数据。

要按指定顺序完成表 T1、T2 和 T3 中数据的再分发，请发出：

```
DB2 REDISTRIBUTE DATABASE PARTITION GROUP DBPG_1 NOT ROLLFORWARD  
RECOVERABLE  
CONTINUE TABLE (T1) FIRST;
```

指定 TABLE (T1) FIRST 将强制数据库管理器首先处理表 T1，以便该表可以在其他表之前返回到联机（只读）状态。所有其他表按数据库管理器确定的顺序处理。

#### 注：

- 不需要指定 ADD DATABASE PARTITION 选项和 DROP DATABASE PARTITION 选项。然而，在执行 REDISTRIBUTE DATABASE PARTITION GROUP 命令之前，可以使用 ALTER DATABASE PARTITION GROUP 命令来添加或删除数据库分区。在这种情况下，REDISTRIBUTE DATABASE PARTITION GROUP 命令不会添加或删除分区，将只是根据指定的选项来再分发数据。
- 强烈建议用户在执行 REDISTRIBUTE DATABASE PARTITION GROUP 命令之前对数据库进行脱机数据库备份。在上述示例中未显示此操作。
- REDISTRIBUTE DATABASE PARTITION GROUP 命令是不可前滚恢复的。有关此问题的充分讨论，请参阅 REDISTRIBUTE DATABASE PARTITION GROUP 主题。
- 完成 REDISTRIBUTE DATABASE PARTITION GROUP 命令之后，它所访问的所有表空间都将保持 BACKUP PENDING 状态。在可访问这些表空间包含的表以执行写入活动之前必须备份这些表空间。

上面的步骤说明如何使用再分发命令的变体在数据库分区之间再分发数据。

---

## 在分区数据库环境中发出命令

在分区数据库环境中，您可能想要发出将在实例中的计算机或在多个数据库分区服务器（节点）上运行的命令。为此，您可以使用 `rah` 命令或 `db2_all` 命令。`rah` 命令允许您发出将在实例中的计算机上运行的命令。

如果想要命令在实例中的多个数据库分区服务器上运行，那么运行 `db2_all` 命令。本章节概述了这些命令。以下信息仅适用于分区数据库环境。

注：

1. 在 Linux 和 UNIX 平台上，您的登录 shell 程序可以是 Korn shell 程序或任何其他 shell；但是，不同 shell 处理包含特殊字符的命令所用的方式不同。
2. 另外，在 Linux 和 UNIX 平台上，`rah` 使用 `DB2RSHCMD` 注册表变量指定的远程 shell 程序。可以在两个远程 shell 程序之间选择：`ssh`（用于更高的安全性要求）或 `rsh`（或者对于 HP-UX 是 `remsh`）。`ssh` 远程 shell 程序用来防止在 UNIX 操作系统环境中以明文形式传输密码。如果不设置此注册表变量，那么使用 `rsh`（或者对于 HP-UX 是 `remsh`）。
3. 在 Windows 上，要运行 `rah` 命令或 `db2_all` 命令，您必须使用 Administrators 组成员的用户帐户来登录。

要确定命令的范围，请参阅 *Command Reference*，它指示命令是在单个的还是所有的数据库分区服务器上运行。如果命令在一个数据库分区服务器上运行，而您想该命令在其所有上面运行，那么使用 `db2_all`。例外情况是 `db2trc` 命令，它在一台计算机的所有逻辑节点（数据库分区服务器）上运行。如果要在所有计算机的所有逻辑节点上运行 `db2trc`，那么使用 `rah`。

---

## rah 和 db2\_all 命令概述

可以依次在各个数据库分区服务器上按顺序运行命令，或可以用并行的方式运行命令。在 Linux 和 UNIX 平台上，如果以并行方式运行这些命令，那么可以选择将输出发送至缓冲区并收集该输出以便显示（缺省行为），或者可在发出该命令的计算机上显示该输出。

在 Windows 上，如果以并行方式运行这些命令，那么在发出该命令的计算机上显示输出。

要使用 `rah` 命令，输入：

```
rah command
```

要使用 `db2_all` 命令，输入：

```
db2_all command
```

有关 `rah` 语法的帮助，输入

```
rah "?"
```

该命令几乎可以是在交互式提示符下输入的任何内容，例如，要按顺序运行的多个命令。在 Linux 和 UNIX 平台上，使用分号（`;`）将多个命令分开。在 Windows 上，使用 `&` 符号将多个命令分隔开。不要在最后一个命令之后使用分隔符。

以下示例显示如何使用 `db2_all` 命令来更改在节点配置文件中指定的所有数据库分区上的数据库配置。因为 `;` 字符在双引号之内，所以将同时运行请求：

```
db2_all ";DB2 GET DB CFG FOR sample USING LOGFILSIZ 100"
```

## 指定 `rah` 和 `db2_all` 命令

可以在命令行中将 `rah` 命令指定为参数，或者在未指定任何参数时，为响应提示而指定该命令。

若命令包含下列特殊字符则应该使用提示方法：

```
| & ; < > ( ) { } [ ] unsubstituted $
```

若在命令行上指定命令作为参数，而且若它包含刚才列示的任何特殊字符，必须以双引号将命令括起来。

**注：**在 Linux 和 UNIX 平台上，将该命令添加至命令历史记录中，就像您在提示符处输入它一样。

可以正常输入命令中的所有特殊字符（除 `\` 外，不必以双引号括起来）。如果需要在命令中包括 `\`，那么必须输入两个反斜杠（`\\`）。

**注：**在 Linux 和 UNIX 平台上，如果未使用 Korn shell 程序，那么可以正常输入命令中的所有特殊字符（除 `"`、`\`、不可替换的 `$` 和单引号 `'`）外，其他字符都不需要用引号引起来）。如果需要在命令中包括这些字符之一，那么必须在字符前加三个反斜杠（`\\\`）。例如，如果需要在命令中包括一个 `\`，那么必须输入四个反斜杠（`\\\\`）。

如果需要在命令中包括双引号（`"`），那么必须在双引号前加三个反斜杠，例如，`\\\"`。

**注：**

1. 在 Linux 和 UNIX 平台上，除非命令 `shell` 提供了在被括上单引号的字符串中输入单引号的某种方法，否则您不能在命令中包括单引号（`'`）。
2. 在 Windows 上，除非命令窗口提供了在被括上单引号的字符串中输入单引号的某种方法，否则您不能在命令中包括单引号（`'`）。

当运行任何 Korn shell 程序的 shell 脚本，而该脚本包含从后台中的标准输入进行读取的逻辑时，您应该明确地将标准输入重定向到一个源，在此源上进程可以读取而不必在终端上停止（SIGTTIN 消息）。要重定向标准输入，可以用下列格式运行脚本：

```
shell_script </dev/null &
```

若没有提供输入。

同样，当在后台运行 `db2_all` 时，应该始终指定 `</dev/null`。例如：

```
db2_all ";run_this_command" </dev/null &
```

通过执行此操作，可以重定向标准输入，并避免在终端上停止。

当您不关心来自远程命令的输出时，此方法的一种替代方法是在 `db2_all` 前缀中使用“`daemonize`”选项：

```
db2_all ";daemonize_this_command" &
```

## 以并行方式运行命令 (Linux 和 UNIX)

缺省情况下，命令在每台计算机上顺序运行，但通过在命令前加上某些前缀序列，可以指定使用后台 `rshell` 以并行方式运行命令。如果 `rshell` 在后台运行，那么每个命令将输出放置在其远程计算机的缓冲区文件中。

注：本节中的信息只适用于 Linux 和 UNIX 平台。

此进程分两个部分检索输出：

1. 在远程命令完成后。
2. 在 `rshell` 终止后，如果某些进程仍在运行，那么 `rshell` 可能会过一段时间才终止。

缺省情况下，缓冲区文件的名称为 `/tmp/$USER/rahout`，但可以通过环境变量 `$RAHBUFDIR/$RAHBUFNAME` 指定该名称。

当指定想要命令同时运行时，缺省情况下，此脚本将附加命令作为前缀加到发送至所有主机的命令上，以检查 `$RAHBUFDIR` 和 `$RAHBUFNAME` 是否可用于缓冲区文件。这会创建 `$RAHBUFDIR`。为避免此类情况，导出环境变量 `RAHCHECKBUF=no`。如果知道目录存在且可用，那么可以执行此操作以节省时间。

在使用 `rah` 同时在多台计算机上运行命令之前：

- 确保对于您的用户标识，目录 `/tmp/$USER` 在每台计算机上存在。要在目录尚不存在的情况下创建目录，运行：

```
rah ")mkdir /tmp/$USER"
```

- 将下一行添加至 `.kshrc`（对于 Korn shell 程序语法）或 `.profile`，并将它输入到当前会话中：

```
export RAHCHECKBUF=no
```

- 确保您运行远程命令的每台计算机的标识在其 `.rhosts` 文件中有一个条目对应于运行 `rah` 的标识；并且运行 `rah` 的标识在其 `.rhosts` 文件中有一个条目对应于运行远程命令的每台计算机的标识。

## 扩展 rah 命令以使用树逻辑 (AIX 和 Solaris)

为了增强性能，在大型系统上扩展了 `rah` 以使用 `tree_logic`。也就是说，`rah` 将检查该列表包含多少个节点，若该数目超过阈值，它会构造列表的一个子集，并将它自己的递归调用发送到那些节点。

在那些节点上，递归调用的 `rah` 遵循相同的逻辑，直到该列表小得能够符合将该命令发送到列表中所有节点的标准逻辑（现在称为“树叶”逻辑）为止。该阈值可由环境变量 `RAHTREETHRESH` 指定，缺省值为 15。

对于每个物理节点存在多逻辑节点的系统，`db2_all` 比较愿意将递归调用发送到各个不同的物理节点，然后 `rsh` 到同一个物理节点上的其他逻辑节点，这样可减少物理节点间的通信量。（这种方法只适合 `db2_all`，不适合 `rah`，因为 `rah` 始终只发送到不同的物理节点。）

## rah 和 db2\_all 命令

本主题包括对 `rah` 和 `db2_all` 命令的描述。

命令 描述

**rah** 在所有计算机上运行该命令。

#### **db2\_all**

在指定的所有数据库分区服务器上运行该命令。

#### **db2\_kill**

突然停止正在多个数据库服务器上运行的所有进程，并清除所有数据库分区服务器上的所有资源。此命令使数据库变得不一致。除非有 IBM 服务中心指导，否则不要发出此命令。

#### **db2\_call\_stack**

在 Linux 和 UNIX 平台上，使在所有数据库分区服务器上运行的所有进程将调用回溯写入 syslog。

在 Linux 和 UNIX 平台上，这些命令执行带特定隐式设置的 rah，例如：

- 以并行方式在所有计算机上运行
- 将命令输出分别缓存到 /tmp/\$USER/db2\_kill 和/tmp/\$USER/db2\_call\_stack 中。

命令 db2\_call\_stack 在 Windows 上不可用。请改为使用 db2pd -stack 命令。

## rah 命令前缀序列

前缀序列是一个或多个特殊字符。

在命令字符前输入一个或多个前缀序列而不插入任何空格。若想指定多个序列，可以任何顺序输入它们，但任何多字符序列中的输入字符必须按顺序输入。如果您输入任何前缀序列，那么您必须将整个命令（包括该前缀序列）置于双引号内，如下例所示：

- 在 Linux 和 UNIX 平台上：

```
rah "};ps -F pid,ppid,etime,args -u $USER"
```

- 在 Windows 上：

```
rah "||db2 get db cfg for sample"
```

前缀序列有：

#### 序列 用途

**I** 在后台按顺序运行命令。

**I&** 在后台按顺序运行这些命令，并在所有远程命令完成之后，终止该命令，即使有一些进程仍在运行。例如，如果子进程（在 Linux 和 UNIX 平台上）或后台进程（在 Windows 上）仍在运行，那么可能会造成延迟。在此情况下，该命令启动独立的后台进程来检索命令终止之后生成的任何远程输出，并将该输出写回至源计算机。

**注：**在 Linux 和 UNIX 平台上，指定 & 会降低性能，因为需要运行更多的 rsh 命令。

**II** 在后台以并行方式运行命令。

**II&** 在后台以并行方式运行命令并在所有远程命令完成之后终止命令，如以上 I& 的情况所述。

**注：**在 Linux 和 UNIX 平台上，指定 & 会降低性能，因为需要运行更多的 rsh 命令。

- ;
- 与以上的 `||&` 相同。这是一个较短的替代格式。  
**注:** 在 Linux 和 UNIX 平台上, 指定 `;` 会降低性能 (相对于 `||` 而言), 因为需要运行更多的 `rsh` 命令。
- ] 在执行命令之前预先暂挂用户概要文件的点执行。  
**注:** 仅在 Linux 和 UNIX 平台上可用
- } 在执行命令之前预先暂挂在 `$RAHENV` 中命名的文件的点执行 (可能是 `.kshrc`)。  
**注:** 仅在 Linux 和 UNIX 平台上可用
- ] 在执行命令之前, 预先暂挂用户概要文件的点执行, 然后执行在 `$RAHENV` 中命名的文件 (可能是 `.kshrc`)。  
**注:** 仅在 Linux 和 UNIX 平台上可用
- ) 停止执行用户概要文件和 `$RAHENV` 中命名的文件。  
**注:** 仅在 Linux 和 UNIX 平台上可用
- ' 将命令调用回传至计算机。
- < 发送至除此计算机外的所有计算机。
- <<-nnn< 发送至除数据库分区服务器 `nnn` 外的所有数据库分区服务器 (`db2nodes.cfg` 中除节点号为 `nnn` 之外的所有数据库分区服务器, 请参阅本表中最后一个前缀序列后的第一段)。
- <<+nnn< 仅发送至数据库分区服务器 `nnn` (`db2nodes.cfg` 中数据库分区号为 `nnn` 的数据库分区服务器, 请参阅本表中最后一个前缀序列后的第一段)。
- (空白字符)  
 在后台运行远程命令, `stdin`、`stdout` 和 `stderr` 全部关闭。此选项仅当在后台运行命令时才有效, 即仅在还包括 `\` 或 `;` 的前缀序列中有效。它允许命令尽快完成 (远程命令一启动就完成)。如果在 `rah` 命令行上指定此前缀序列, 那么将该命令用单引号括起来, 或用双引号括起该命令并在前缀字符之前加 `\`。例如,  

```
rah ' ; mydaemon'
```

 或者  

```
rah " ; \ mydaemon"
```

 当作为后台进程运行时, `rah` 命令从不会等待任何要返回的输出。
- > 用计算机名称替换找到的 `>`。
- " 用计算机索引替换找到的 `()`, 用数据库分区号替换找到的 `##`。
- 注:**
  1. 计算机索引是与数据库系统中的计算机关联的号码。如果没有在运行多逻辑分区, 那么计算机的计算机索引对应于节点配置文件中该计算机的数据库分区号。要在多逻辑分区数据库环境中获取计算机的计算机索引, 不要计算那些运行多逻辑分区的计算机的重复条目。例如, 如果 `MACH1` 正在运

行两个逻辑分区，MACH2 也正在运行两个逻辑分区，那么节点配置文件中 MACH3 的数据库分区号为 5。但是，MACH3 的计算机索引应是 3。

在 Windows 上，不要编辑节点配置文件。要获取计算机索引，可使用 db2nlist 命令。

2. 当指定了 " 时，不会从计算机列表中删除重复项。

当使用 <<-nnn< 和 <<+nnn< 前缀序列时，*nnn* 是任何 1 位、2 位或 3 位的数据库分区号，该分区号必须与 db2nodes.cfg 文件中的 *nodenum* 值匹配。

**注：**前缀序列被认为是命令的一部分。若指定前缀序列作为命令的一部分，必须将整个命令，包括前缀序列，括在双引号内。

## 控制 rah 命令

本主题列示用于控制 rah 命令的环境变量。

表 13. 控制 rah 命令的环境变量

| 名称   | 含义                                  | 缺省值                                    |
|--|-------------------------------------|--|
| \$RAHBUFDIR<br>注：仅在 Linux 和 UNIX 平台上可用                       | 缓冲区目录                               | /tmp/\$USER                            |
| \$RAHBUFNAME<br>注：仅在 Linux 和 UNIX 平台上可用                      | 缓冲区文件名                              | rahout                                 |
| \$RAHOSTFILE（在 Linux 和 UNIX 平台上）；RAHOSTFILE（在 Windows 上）     | 包含主机列表的文件                           | db2nodes.cfg                           |
| \$RAHOSTLIST（在 Linux 和 UNIX 平台上）；RAHOSTLIST（在 Windows 上）     | 字符串形式的主机列表                          | 取自 \$RAHOSTFILE                        |
| \$RAHCHECKBUF<br>注：仅在 Linux 和 UNIX 平台上可用                     | 如果设置为“no”，那么绕过检查                    | 未设置                                    |
| \$RAHSLEEPTIME（在 Linux 和 UNIX 平台上）；RAHSLEEPTIME（在 Windows 上） | 以秒计的时间，此脚本将在此时间内等待来自以并行形式运行的命令的初始输出 | 对于 db2_kill 为 86400 秒，对于其他所有命令则为 200 秒 |



表 13. 控制 rah 命令的环境变量 (续)

| 名称  | 含义   | 缺省值    |
|---|--|--------|
| \$RAHWAITTIME (在 Linux 和 UNIX 平台上); RAHWAITTIME (在 Windows 上) | 在 Windows 上, 连续检查远程作业是否仍在运行的时间间隔 (以秒计)。在 Linux 和 UNIX 平台上, 连续检查远程作业是否仍在运行以及 rah: 正在等待 pid> ... 消息的时间间隔 (以秒计)。对于所有平台, 指定任何正整数。具有前导零的前缀值可抑制消息, 例如, 导出 RAHWAITTIME=045。不必指定较小的值, 因为 rah 不依靠这些检查来检测作业是否完成。 | 45 秒   |
| \$RAHENV<br>注: 仅在 Linux 和 UNIX 平台上可用                          | 如果 \$RAHDOTFILES=E 或 K 或 PE 或 B, 那么指定要执行的文件名   | \$ENV  |
| \$RAHUSER (在 Linux 和 UNIX 平台上); RAHUSER (在 Windows 上)         | 在 Linux 和 UNIX 平台上, 运行远程命令将使用的用户标识。在 Windows 上, 与 DB2“远程命令服务”相关的登录帐户   | \$USER |

注: 在 Linux 和 UNIX 平台上, 将使用运行 rah 的 \$RAHENV 的值, 而不是远程 shell 设置的值 (如果有的话)。

## 指定与 rah 一起运行的 . 文件 (Linux 和 UNIX)

本主题列示在未指定前缀序列时运行的 . 文件。

注: 本节中的信息只适用于 Linux 和 UNIX 平台。

- P** .profile
- E** 在 \$RAHENV 中命名的文件 (可能是 .kshrc)
- K** 与 E 相同
- PE** 后跟 \$RAHENV 中命名的文件的 .profile (可能是 .kshrc)
- B** 与 PE 相同
- N** 无

注: 如果登录 shell 程序不是 Korn shell 程序, 那么将在 Korn shell 程序进程中执行您指定要执行的任何点文件, 因此, 必须遵守 Korn shell 程序语法。例如, 如果登录 shell 程序是 C shell, 要对 rah 执行的命令设置 .cshrc 环境, 应该创建等效于 .cshrc 的 Korn shell 程序 INSTHOME/.profile, 并在 INSTHOME/.cshrc 中指定:

```
setenv RAHDOTFILES P
```

或应该创建等效于 .cshrc 的 Korn shell 程序 INSTHOME/.kshrc, 并在 INSTHOME/.cshrc 中指定:

```
setenv RAHDOTFILES E
setenv RAHENV INSTHOME/.kshrc
```

此外，如果没有 tty（由 rsh 调用时），那么 .cshrc 必须不写入标准输出。通过将写至标准输出的任何行用引号括起来，可确保这一点，例如，

```
if { tty -s } then echo "executed .cshrc";
endif
```

## 确定 rah 的问题（Linux 和 UNIX）

本主题提供了一些建议，告诉您如何处理在运行 rah 时可能遇到的某些问题。

**注：**本节中的信息只适用于 Linux 和 UNIX 平台。

### 1. rah 挂起（或运行很长的时间）

此问题可能是由下列原因引起的：

- rah 已确定它需要缓冲输出，而您未导出 RAHCHECKBUF=no。因此，在运行命令之前，rah 向所有计算机发出一条命令以检查缓冲区目录是否存在，如果该目录不存在，那么会创建该目录。
- 发送命令的一台或多台计算机不响应。rsh 命令最终将超时，但超时时间间隔相当长，通常为 60 秒左右。

### 2. 已接受到下列各类消息：

- 登录不正确
- 拒绝许可权

其中一台计算机没有在其 .hosts 中正确定义运行 rah 的标识，或运行 rah 的标识没有在其 .rhosts 文件中正确定义其中一台计算机。如果已将 DB2RSHCMD 注册表变量配置为使用 ssh，那么可能是未正确配置每台计算机上的 ssh 客户机和服务器。

**注：**对于以明文方式在数据库分区之间传输密码的情况，可能需要使用更高的安全性。这将取决于正在使用的远程 shell 程序。rah 使用 DB2RSHCMD 注册表变量指定的远程 shell 程序。可以在两个远程 shell 程序之间选择：ssh（用于更高的安全性要求）或 rsh（或者对于 HP-UX 是 remsh）。如果不设置此注册表变量，那么使用 rsh（或者对于 HP-UX 是 remsh）。

### 3. 当使用后台远程 shell 以并行方式运行命令时，虽然这些命令在期望的时间内在计算机上运行并完成，但是 rah 要耗费较长的时间来检测它并设置 shell 提示符。

运行 rah 的标识没有在其 .rhosts 文件中正确定义其中一台计算机，或者如果已将 DB2RSHCMD 注册表变量配置为使用 ssh，那么可能是未正确配置每台计算机上的 ssh 客户机和服务器。

### 4. 虽然 rah 从 shell 命令行运行时运行情况良好，但是如果使用 rsh 远程运行 rah，例如，

```
rsh somewhere -l $USER db2_kill
```

，那么 rah 将永不会完成。

这是正常的。rah 会启动后台监视进程，这些进程在 rah 退出后继续运行。那些进程通常将在与您运行的命令关联的所有进程将它们终止之后才结束。在 db2\_kill 的情况下，这意味着终止所有数据库管理器。可通过查找其命令是 rahwaitfor 和 kill process\_id> 的进程来终止监视进程。不要指定信号编号。而是要使用缺省值（15）。

### 5. 当在同一 \$RAHUSER 下发出多个 rah 命令时，rah 的输出未正确显示，或 rah 错误地报告 \$RAHBUFNAME 不存在。

这是因为，并行执行多个 `rah` 正在尝试使用同一缓冲区文件（例如 `$RAHBUFDIR/$RAHBUFNAME`）对输出进行缓冲。要避免此问题，将不同的 `$RAHBUFNAME` 用于每个并行 `rah` 命令，例如，在下列 `ksh` 中：

```
export RAHBUFNAME=rahout
rah ";$command_1" &
export RAHBUFNAME=rah2out
rah ";$command_2" &
```

或使用使 `shell` 自动选择唯一名称的方法，如：

```
RAHBUFNAME=rahout.$$ db2_all "....."
```

如果磁盘空间有限，那么无论使用哪种方法，都必须确保在某个时间清除缓冲区文件。`rah` 不会在执行结束时擦除缓冲区文件，不过，下次您指定与现有文件相同的缓冲区文件时，它将擦除并复用现有文件。

## 6. 输入

```
rah 'print from ()'
```

并接收到消息：

```
ksh: syntax error at line 1 : (' unexpected
```

替代 `()` 和 `##` 的先决条件是：

- 使用 `db2_all`，而不是使用 `rah`。
- 确保通过导出 `RAHOSTFILE` 或缺省为您的 `/sqllib/db2nodes.cfg` 文件来使用 `RAHOSTFILE`。没有这些先决条件，`rah` 将照原样保持 `()` 和 `##`。您接收到错误，因为命令 `print from ()` 无效。

要获取以并行方式运行命令时的性能提示，除非确实需要由 `&` 提供的功能，否则，使用 `|` 代替 `|&`，并使用 `||` 代替 `||&` 或 `;`。指定 `&` 需要更多远程 `shell` 命令，因此会降低性能。

## 监视 `rah` 进程 (Linux 和 UNIX)

当任何远程命令仍在运行或仍在累加缓冲输出时，由 `rah` 启动的进程将监视活动并将指示哪些活动尚未运行的消息写入终端并检索缓冲输出。

**注：**本节中的信息只适用于 Linux 和 UNIX 平台。

按环境变量 `RAHWAITTIME` 所控制的时间间隔写出参考消息。请参阅帮助信息以获取如何指定时间间隔的详细信息。通过导出 `RAHWAITTIME=0`，可以完全不显示所有参考消息。

主监视过程是一个命令，其命令名（由 `ps` 命令显示）为 `rahwaitfor`。第一条参考消息告诉您此进程的 `pid`（进程标识）。所有其他监视进程将表现为运行 `rah` 脚本（或符号链接的名称）的 `ksh` 命令。若愿意，可通过以下命令停止所有监视进程：

```
kill <pid>
```

其中 `<pid>` 是主监视进程的进程标识。不要指定信号编号。让它为缺省值 15。这根本不会影响远程命令，但会阻止自动显示缓冲输出。注意，在执行单个 `rah` 时，可能有两组或更多组不同监视进程在不同时间执行。然而，如果任何时候停止当前组监视进程，那么不会再启动其他监视进程。

如果正规登录 shell 程序不是 Korn shell 程序（例如，`/bin/ksh`），那么可以使用 `rah`，但如何输入包含下列特殊字符的命令的规则稍微有些差异：

```
" unsubstituted $ '
```

有关更多信息，请输入 `rah "?"`。而且，在 Linux 和 UNIX 环境中，如果执行远程命令的标识的登录 shell 程序不是 Korn shell 程序，那么以该标识执行 `rah` 的登录 shell 程序也不得是 Korn shell 程序。（`rah` 根据本地标识来决定远程标识的 shell 是否为 Korn shell 程序）。该 shell 不得对单引号内的字符串执行任何替换或特殊处理。必须照原样保持。

## 在 Windows 上为 rah 设置缺省环境概要文件

要为 `rah` 命令设置缺省环境概要文件，可使用文件 `db2rah.env`，该文件应在实例目录中创建。

**注：**本节中的信息只适用于 Windows。

该文件应该有以下格式：

```
; This is a comment line
DB2INSTANCE=instancename
DB2DBDFT=database
; End of file
```

可以指定为 `rah` 进行环境初始化所需要的所有环境变量。

---

## 第 11 章 创建表和其他相关表对象

---

### 分区数据库环境中的表

在分区数据库环境中跨几个数据库分区创建表在性能上有几个优点。与检索数据相关的工作可分成几部分在各个数据库分区中进行。

在创建将以物理方式划分或分布的一个表之前，需要考虑下列事宜：

- 表空间可以横跨多个数据库分区。它们所跨的数据库分区数取决于数据库分区组中的数据库分区数。
- 可以通过如下方法来并置表：将表置于同一个表空间中，或置于另一个表空间中，该表空间与第一个表空间一起，与同一个数据库分区组相关联。

在创建表时，指定创建的表将成为若干数据库分区的一部分。当在分区数据库环境中创建表时，有一个附加选项：**分布键**。分布键是作为一个表定义的一部分的键。它确定用于存储每行数据的数据库分区。

若不显式指定分布键，会使用下列缺省值。确保缺省分布键适合。

- 如果在 **CREATE TABLE** 语句中指定了主键，那么该主键的第一列会用作分布键。
- 对于多分区数据库分区组，如果不存在主键，那么使用非长整型字段的第一列。
- 如果没有列满足缺省分布键的需求，那么会不带关键字创建该表（这只在单一分区数据库分区组中允许）。

必须小心地选择适当的分布键，因为以后再也不能更改它。再者，必须将任何唯一索引（因此也是唯一键或主键）定义为分布键的一个超集。即，如果定义了分布键，那么唯一键和主键必须包括所有与分布键相同的列（它们可能有多列）。

数据库分区的表大小是与使用的表空间和页大小的类型关联的特定限制和可用的磁盘空间大小这两者中的较小者。例如，假定有一个 4 KB 页大小的大型 DMS 表空间，表大小是 2 TB 与数据库分区数目的乘积和可用的磁盘空间大小这两者中的较小者。请参阅相关链接，以获取数据库管理器页大小限制的完整列表。

要使用命令行在分区数据库环境中创建一个表，请输入：

```
CREATE TABLE name>
    (<column_name> <data_type> <null_attribute>)
IN <table_space_name>
    INDEX IN <index_space_name>
    LONG IN <long_space_name>          DISTRIBUTE BY HASH (<column_name>)
```

以下是一个示例：

```
CREATE TABLE MIXREC (MIX_CNTL INTEGER NOT NULL,
                      MIX_DESC CHAR(20) NOT NULL,
                      MIX_CHR CHAR(9) NOT NULL,
                      MIX_INT INTEGER NOT NULL,
                      MIX_INTS SMALLINT NOT NULL,
                      MIX_DEC DECIMAL NOT NULL,
                      MIX_FLT FLOAT NOT NULL,
                      MIX_DATE DATE NOT NULL,
                      MIX_TIME TIME NOT NULL,
```

```
MIX_TMSTMP TIMESTAMP NOT NULL)
      IN MIXTS12
DISTRIBUTE BY HASH (MIX_INT)
```

在上一个示例中，表空间是 MIXTS12，而分布键是 MIX\_INT。如果未显式指定分布键，那么它是 MIX\_CNTRL。（如果未指定主键且未定义分布键，那么分布键是该列表中的第一个非长型列。）

表的一行和有关该行的所有信息始终位于同一个数据库分区上。

---

## 分区表中的大对象行为

分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。根据 CREATE TABLE 语句的 PARTITION BY 子句中指定的内容，给定表的数据被划分到多个存储对象中。这些存储对象可以在不同的表空间中，也可以在相同表空间中。

缺省情况下，分区表的大对象与其相应的数据对象存储在相同表空间中。这适用于只使用一个表空间或使用多个表空间的分区表。当分区表的数据存储在多个表空间中时，大对象数据也存储在多个表空间中。

使用 CREATE TABLE 语句的 LONG IN 子句来覆盖此缺省行为。可以为表指定将存储长型数据的表空间列表。如果选择覆盖缺省行为，那么在 LONG IN 子句中指定的表空间必须是大型表空间。如果指定将一个或多个数据分区的长型数据存储在单独的表空间中，那么必须对表的所有数据分区都如此。也就是说，不能将某些数据分区的数据存储在远地，而将其他数据分区的数据存储在本地。无论是使用缺省行为还是使用 LONG IN 子句来覆盖缺省行为，都会创建一个与每个数据分区相对应的长型对象。对于 SMS 表空间，长型数据必须与它所属的数据对象位于相同表空间中。用于存储与每个数据分区相对应的长型数据对象的所有表空间必须具有相同的页大小、扩展数据块大小、存储机制（DMS 或 SMS）和类型（常规或大型）。远程大型表空间的类型必须是 LARGE，不能是 SMS。

例如，以下 CREATE TABLE 语句在 CLOB 数据所在的表空间中为每个数据分区的 CLOB 数据创建对象：

```
CREATE TABLE document(id INT, contents CLOB)
PARTITION BY RANGE(id)
(STARTING FROM 1 ENDING AT 100 IN tbsp1,
 STARTING FROM 101 ENDING AT 200 IN tbsp2,
 STARTING FROM 201 ENDING AT 300 IN tbsp3,
 STARTING FROM 301 ENDING AT 400 IN tbsp4);
```

可以使用 LONG IN 将 CLOB 数据放在一个或多个与该数据所在的大型表空间不同的大型表空间中。

```
CREATE TABLE document(id INT, contents CLOB)
PARTITION BY RANGE(id)
(STARTING FROM 1 ENDING AT 100 IN tbsp1 LONG IN large1,
 STARTING FROM 101 ENDING AT 200 IN tbsp2 LONG IN large1,
 STARTING FROM 201 ENDING AT 300 IN tbsp3 LONG IN large2,
 STARTING FROM 301 ENDING AT 400 IN tbsp4 LONG IN large2);
```

注：对于每个数据分区，在表级别只允许使用一个 LONG IN 子句。

## 创建分区表

分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。根据 `CREATE TABLE` 语句的 `PARTITION BY` 子句中指定的内容，给定表的数据被划分到多个存储对象中。这些存储对象可以在不同的表空间中，也可以在相同表空间中。

要创建表，语句授权标识拥有的特权必须至少包括下列其中一项权限或特权：

- 对数据库的 `CREATETAB` 权限、对该表使用的所有表空间的 `USE` 特权以及下列其中一项权限或特权：
  - 对数据库的 `IMPLICIT_SCHEMA` 权限（如果该表的隐式或显式模式名不存在的话）
  - 对模式的 `CREATEIN` 特权（如果该表的模式名引用现有模式的话）
- `SYSADM` 或 `DBADM` 权限

可以使用 `CREATE TABLE` 语句创建分区表。

要使用命令行来创建分区表，请发出 `CREATE TABLE` 语句。

```
CREATE TABLE <NAME> (<column_name> <data_type> <null_attribute>) IN
<table space list> PARTITION BY RANGE (<column expression>)
STARTING FROM <constant> ENDING <constant> EVERY <constant>
```

例如，以下语句将创建一个表，在该表中，满足  $a \geq 1$  且  $a \leq 20$  的行位于 `PART0`（第一个数据分区）中，满足  $21 \leq a \leq 40$  的行位于 `PART1`（第二个数据分区）中，而满足  $81 \leq a \leq 100$  的行位于 `PART4`（最后一个数据分区）中。

```
CREATE TABLE foo(a INT)
PARTITION BY RANGE (a) (STARTING FROM (1)
ENDING AT (100) EVERY (20))
```

## 定义分区表的范围

在创建分区表时，可以为每个数据分区指定范围。分区表使用了数据组织方案，即，表数据根据该表中表分区键列的值分布到多个数据分区中。

根据 `CREATE TABLE` 语句的 `PARTITION BY` 子句中指定的内容，给定表的数据被划分到多个存储对象中。范围由 `PARTITION BY` 子句的 `STARTING FROM` 和 `ENDING AT` 值指定。

要全面地定义每个数据分区的范围，必须指定足够的边界。以下是定义分区表的范围时要考虑的一系列准则：

- `STARTING` 子句指定数据分区范围的下界。对于最低数据分区范围来说，此子句是必需的（尽管可以将边界定义为 `MINVALUE`）。最低数据分区范围是具有最低指定边界的数据分区。
- `ENDING`（或 `VALUES`）子句指定数据分区范围的上界。对于最高数据分区范围来说，此子句是必需的（尽管可以将边界定义为 `MAXVALUE`）。最高数据分区范围是具有最高指定边界的数据分区。
- 如果未对某个数据分区指定 `ENDING` 子句，那么下一个更大数据分区就必须指定 `STARTING` 子句。否则，如果未指定 `STARTING` 子句，那么上一个数据分区就必须指定 `ENDING` 子句。

- MINVALUE 指定一个值，该值小于所用列类型的任何可能值。不能将 MINVALUE 与 INCLUSIVE 或 EXCLUSIVE 一起指定。
- MAXVALUE 指定一个值，该值大于所用列类型的任何可能值。不能将 MAXVALUE 与 INCLUSIVE 或 EXCLUSIVE 一起指定。
- INCLUSIVE 表示将所有等于指定值的值都包括在包含此边界的数据分区中。
- EXCLUSIVE 表示所有等于指定值的值都不包括在包含此边界的数据分区中。
- CREATE TABLE 语句的 NULL 子句指定考虑数据分区布置时是将空值安排在高位置还是低位置。缺省情况下，将空值安排在高位置。在此情况下，将把表分区键列中的空值视为正无穷并放到以 MAXVALUE 结尾的范围中。如果未定义这样的数据分区，就会将空值视为超出范围的值。如果要排除表分区键列中的空值，请使用 NOT NULL 约束。LAST 指定让空值在排序的值列表中最后出现。FIRST 指定让空值在排序的值列表中最先出现。
- 当使用长语法格式时，必须对每个数据分区至少指定一个边界。

**提示：** 在开始对表定义数据分区之前，您应该了解表分区是否能使表受益以及影响分区列选择的那些因素，这一点十分重要。

可以自动生成对每个数据分区指定的范围，也可以手动生成这些范围。

### 自动生成

自动生成方法十分简单，它使您能够快速方便地创建许多数据分区。此方法适合于创建基于日期或数值并且大小相等的范围。

示例 1 和 2 演示如何使用 CREATE TABLE 语句来自动定义和生成对每个数据分区指定的范围。

示例 1:

发出定义了下列范围的 CREATE TABLE 语句:

```
CREATE TABLE lineitem (
  l_orderkey    DECIMAL(10,0) NOT NULL,
  l_quantity    DECIMAL(12,2),
  l_shipdate    DATE,
  l_year_month  INT GENERATED ALWAYS AS (YEAR(l_shipdate)*100 +
                                         MONTH(l_shipdate))
  PARTITION BY RANGE(l_shipdate)
  (STARTING ('1/1/1992') ENDING ('12/31/1992') EVERY 1 MONTH);
```

此语句生成 12 个数据分区，每个数据分区包含 1 个键值：(l\_shipdate) >= ('1/1/1992'), (l\_shipdate) < ('3/1/1992'), (l\_shipdate) < ('4/1/1992'), (l\_shipdate) < ('5/1/1992'), ..., (l\_shipdate) < ('12/1/1992'), (l\_shipdate) < ('12/31/1992')。

由于整体起始界限 ('1/1/1992') 包括端值 (缺省情况)，所以第一个数据分区的起始值包括端值。同样，由于整体结束界限 ('12/31/1992') 包括端值 (缺省情况)，所以最后一个数据分区的结束界限包括端值。其余 STARTING 值都包括端值，并且其余 ENDING 值也都包括端值。每个数据分区都存放 n 个键值，其中 n 由 EVERY 子句指定。使用公式 (start + every) 来确定每个数据分区的范围末端。如果 START 到 END 的范围无法整除 EVERY 值，最后一个数据分区包含的键值就会较少。

示例 2:



发出定义了下列范围的 CREATE TABLE 语句:

```
CREATE TABLE t(a INT, b INT)
PARTITION BY RANGE(b) (STARTING FROM (1) EXCLUSIVE ENDING AT (1000) EVERY (100))
```

此语句生成 10 个数据分区, 每个数据分区包含 100 个键值: (1 < b <= 101, 101 < b <= 201, ..., 901 < b <= 1000)。

由于整体起始界限 (1) 不包括端值, 所以第一个数据分区 (b > 1 and b <= 101) 的起始值不包括端值。同样, 由于整体结束界限 (1000) 包括端值, 所以最后一个数据分区 (b > 901 b <= 1000) 的结束界限包括端值。其余 STARTING 值都不包括端值, 并且其余 ENDING 值全都包括端值。每个数据分区都存放 n 个键值, 其中 n 由 EVERY 子句指定。最后, 如果整个子句的起始和结束界限都不包括端值, 那么由于整体起始界限 (1) 不包括端值, 所以第一个数据分区的起始值不包括端值。同样, 由于整体结束界限 (1000) 不包括端值, 所以最后一个数据分区的结束界限不包括端值。其余 STARTING 值都不包括端值, 并且 ENDING 值全都包括端值。每个数据分区 (最后一个数据分区除外) 都存放 n 个键值, 其中 n 由 EVERY 子句指定。

### 手动生成

手动生成方法为 PARTITION BY 子句中列示的每个范围创建一个新数据分区。这种语法格式提高了定义范围时的灵活性, 从而增加了数据和 LOB 布置选项。示例 3 和 4 演示如何使用 CREATE TABLE 语句来以手动方式定义和生成对数据分区指定的范围。

示例 3:

此语句对两个日期列进行分区, 这两个日期列都是生成列。请注意自动生成的 CREATE TABLE 语法格式的使用, 并注意每个范围都只指定了一端。另一端由相邻数据分区隐式确定并且要使用 INCLUSIVE 选项:

```
CREATE TABLE sales(invoice_date date, inv_month int NOT NULL
GENERATED ALWAYS AS (month(invoice_date)), inv_year INT NOT
NULL GENERATED ALWAYS AS ( year(invoice_date)), item_id int NOT NULL,
cust_id int NOT NULL) PARTITION BY RANGE (inv_year, inv_month)
(PART Q1_02 STARTING (2002,1) ENDING (2002, 3) INCLUSIVE,
PART Q2_02 ENDING (2002, 6) INCLUSIVE,
PART Q3_02 ENDING (2002, 9) INCLUSIVE,
PART Q4_02 ENDING (2002,12) INCLUSIVE,
PART CURRENT ENDING (MAXVALUE, MAXVALUE));
```

在范围之间允许存在间隔。对于未紧贴上一数据分区 ENDING 值的范围, CREATE TABLE 语法允许您对该范围指定 STARTING 值, 从而支持间隔。

示例 4:

创建一个表, 并且在值 101 与 200 之间存在间隔。

```
CREATE TABLE foo(a INT)
PARTITION BY RANGE(a)
(STARTING FROM (1) ENDING AT (100),
STARTING FROM (201) ENDING AT (300))
```

使用允许添加或除去数据分区的 ALTER TABLE 语句还会导致范围中出现间隔。

把行插入分区表时, 根据该行的键值以及它所处的范围自动将其放入正确的数据分区。如果该行处于对该表定义的所有范围之外, 插入就会失败, 并且将把以下错误返回给应用程序:

SQL0327N 由于该行处于已定义数据分区范围的界限外部，  
所以无法将其插入到表 <tablename> 中。  
SQLSTATE=22525

### 限制

- 表级别限制:
  - 使用自动生成的语法格式（包含 EVERY 子句）创建的表在表分区键中只能使用数字或日期时间类型。
- 语句级限制:
  - 在自动生成的语法格式中，不支持 MINVALUE 和 MAXVALUE。
  - 范围按升序排列。
  - 在自动生成的语法格式中，只能指定一列。
  - EVERY 子句中的增量必须大于零。
  - ENDING 值必须大于或等于 STARTING 值。

## 将现有表和视图迁移到分区表

要将数据从版本 9.1 的表迁移到空分区表，请使用 LOAD 命令。

可以通过三种方法来将现有的表或视图迁移到分区表：

- 迁移正规表时，创建一个新的空分区表并使用 LOAD from CURSOR 来将数据从旧表直接移到分区表中，而不执行任何中间步骤。
- 迁移常规表时，使用 export 实用程序或高性能卸载来卸载源表、创建一个新的空分区表并使用 LOAD 命令来填充空的分区表。
- 迁移 UNION ALL 视图时，创建一个带有单一虚拟数据分区的分区表，然后连接所有的表。

示例 1: 假定您有一个常规表 t1:

```
CREATE TABLE t1 (c1 int, c2 int);
```

创建新的空分区表:

```
CREATE TABLE sales_dp (c1 int, c2 int)
PARTITION BY RANGE (c1)
(STARTING FROM 0 ENDING AT 10 EVERY 2);
```

填充表 t1:

```
INSERT INTO t1 VALUES (0,1), (4, 2), (6, 3);
```

要避免在平面文件中创建第三份数据副本，请发出 LOAD 命令以将数据从 SQL 查询直接放入新分区表。

```
SELECT * FROM t1;
DECLARE c1 CURSOR FOR SELECT * FROM t1;
LOAD FROM c1 OF CURSOR INSERT INTO sales_dp;SELECT * FROM sales_dp;
SELECT * FROM sales_dp;
```

删除旧表:

```
DROP TABLE t1;
```

### 转换 UNION ALL 视图

可以将 UNION ALL 视图中的版本 9.1 数据转换到分区表中。UNION ALL 视图用于管理大型的表，它简化了表数据的转入和转出，并且提供了分支消除的性能优势。表分区功能实现了上述全部目标，并且更易于管理。通过使用 ALTER TABLE ...ATTACH 操作，可以完成转换，而不需要移动基本表中的数据。转换后，必须重新创建索引、从属视图或具体化查询表（MQT）。

建议的策略是，创建带有单一虚拟数据分区的分区表，然后连接 UNION ALL 视图的所有表。务必在处理过程中尽早删除虚拟数据分区以避免范围重叠问题。

示例 2:

UNION 中第一个表的 CREATE TABLE 语法:

```
CREATE TABLE sales_0198(  
  sales_date DATE NOT NULL,  
  prod_id INTEGER,  
  city_id INTEGER,  
  channel_id INTEGER,  
  revenue DECIMAL(20,2),  
  CONSTRAINT ck_date  
  CHECK  
  (sales_date BETWEEN '01-01-1998' AND '01-31-1998'));
```

UNION ALL 视图的 CREATE VIEW 语法:

```
CREATE VIEW all_sales AS  
(  
  SELECT * FROM sales_0198  
  WHERE sales_date BETWEEN '01-01-1998' AND '01-31-1998'  
  UNION ALL  
  SELECT * FROM sales_0298  
  WHERE sales_date BETWEEN '02-01-1998' AND '02-28-1998'  
  UNION ALL  
  ...  
  UNION ALL  
  SELECT * FROM sales_1200  
  WHERE sales_date BETWEEN '12-01-2000' AND '12-31-2000'  
);
```

创建带有单一虚拟分区的分区表。选择范围时，应该确保它不会与所要连接的第一个数据分区重叠:

```
CREATE TABLE sales_dp (  
  sales_date DATE NOT NULL,  
  prod_id INTEGER,  
  city_id INTEGER,  
  channel_id INTEGER,  
  revenue DECIMAL(20,2))  
PARTITION BY RANGE (sales_date)  
(PART dummy STARTING FROM '01-01-1900' ENDING AT '01-01-1900');
```

连接第一个表:

```
ALTER TABLE sales_dp ATTACH PARTITION  
STARTING FROM '01-01-1998' ENDING AT '01-31-1998'  
FROM sales_0198;
```

删除虚拟分区:

```
ALTER TABLE sales_dp DETACH PARTITION dummy  
INTO dummy;  
DROP TABLE dummy;
```

连接其余的分区:

```
ALTER TABLE sales_dp ATTACH PARTITION STARTING
FROM '02-01-1998' ENDING AT '02-28-1998' FROM sales_0298;
```

...

```
ALTER TABLE sales_dp ATTACH PARTITION STARTING
FROM '12-01-2000' ENDING AT '12-31-2000' FROM sales_1200;
```

发出 SET INTEGRITY 语句以使连接的数据分区联机。

```
SET INTEGRITY FOR sales_dp IMMEDIATE CHECKED
FOR EXCEPTION IN sales_dp USE sales_ex;
```

根据情况, 创建索引。

### 转换注意事项

对于特定的列, 除非源列和目标列的 SYSCAT.COLUMNS IMPLICITVALUE 字段值都是非空值, 并且该值不匹配, 否则允许连接数据分区。在不允许连接数据分区的情况下, 必须将源表删除, 然后重新创建该表。

如果符合下列其中一个条件, 列的 SYSCAT.COLUMNS IMPLICITVALUE 字段就可以具有非空值:

- 该列是作为 ALTER TABLE ...ADD COLUMN 语句的结果创建的
- IMPLICITVALUE 字段是在连接期间从源表传播而来的
- IMPLICITVALUE 字段是在拆离期间从源表继承的
- IMPLICITVALUE 字段是从 V8 迁移到 V9 期间设置的, 在 V9 中, 确定该字段是添加的列, 或者可能是添加的列。如果数据库无法确定该列是否是添加的列, 那么将其视为如此。添加的列就是作为 ALTER TABLE ...ADD COLUMN 语句的结果创建的列。

为了避免产生这些不一致, 建议您始终要创建连接操作中涉及到的源表和目标表, 并且这两个表定义了相同的列数。特别是, 千万不要使用 ALTER TABLE 语句来向连接操作的目标表添加列。

要了解有关使用分区表时避免出错的最佳示例, 请参阅第 176 页的『将数据分区连接至分区表的准则』。

---

## 已分区的具体化查询表 (MQT) 行为

分区表支持所有类型的具体化查询表 (MQT)。使用已分区的 MQT 时, 可使用一些准则来帮助您更有效地管理已连接的数据分区和已拆离的数据分区。

当使用已分区的 MQT 或带有已拆离的从属表的分区表时, 下列准则和限制适用:

- 如果发出 DETACH PARTITION 操作, 并且有任何需要根据已拆离数据分区进行递增维护的从属表 (这些从属表称为已拆离的从属表), 那么新拆离的表最初不可访问。对于该表, 在 SYSCAT.TABLES 目录视图的 TYPE 列中将标记 L。此表称为已拆离表。这样, 在运行 SET INTEGRITY 语句以递增方式维护已拆离的从属表之前, 无法读取、修改或删除这个已拆离表。在对所有已拆离的从属表运行 SET INTEGRITY 语句后, 已拆离表就会转换为可供完全访问的正规表。

- 要检测已拆离表是否仍不可访问，请查询 SYSCAT.TABDETACHEDDEP 目录视图。如果检测到任何不可访问的已拆离表，那么对所有已拆离的从属表运行带有 IMMEDIATE CHECKED 选项的 SET INTEGRITY 语句，以将已拆离表转换为可访问的常规表。如果在维护已拆离表的所有已拆离从属表之前试图访问该已拆离表，那么将返回错误代码 SQL20285N。
- 不能在具体化查询表（MQT）定义中使用 DATAPARTITIONNUM 函数。尝试使用此函数来创建 MQT 将返回错误（SQLCODE SQL20058N 和 SQLSTATE 428EC）。
- 当对带有已拆离数据分区的表创建索引时，索引将不包括已拆离数据分区中的数据，除非已拆离数据分区具有需要根据该数据分区进行递增刷新的从属具体化查询表（MQT）。在后面这种情况下，索引将包括此已拆离数据分区的数据。
- 不允许将带有已连接数据分区的表改变为 MQT。
- 不支持已分区的登台表。
- 不直接支持与 MQT 连接。有关详细信息，请参阅示例 1。

### 示例 1: 将已分区的 MQT 转换为一个普通表

虽然不直接支持对已分区的 MQT 执行 ATTACH 操作，但可以通过下列操作获得相同的效果：将已分区的 MQT 转换为一个普通表，对表数据执行期望的转入和转出，然后将该表转换回 MQT。下列 CREATE TABLE 和 ALTER TABLE 语句演示了该效果：

```
CREATE TABLE lineitem (
  l_orderkey    DECIMAL(10,0) NOT NULL,
  l_quantity    DECIMAL(12,2),
  l_shipdate    DATE,
  l_year_month  INT GENERATED ALWAYS AS (YEAR(l_shipdate)*100 +
                                          MONTH(l_shipdate))
  PARTITION BY RANGE(l_shipdate)
  (STARTING ('1/1/1992') ENDING ('12/31/1993') EVERY 1 MONTH);
CREATE TABLE lineitem_ex (
  l_orderkey    DECIMAL(10,0) NOT NULL,
  l_quantity    DECIMAL(12,2),
  l_shipdate    DATE,
  l_year_month  INT,
  ts            TIMESTAMP,
  msg           CLOB(32K));

CREATE TABLE quan_by_month (
  q_year_month, q_count) AS
  (SELECT l_year_month AS q_year_month, COUNT(*) AS q_count
   FROM lineitem
   GROUP BY l_year_month)
  DATA INITIALLY DEFERRED REFRESH IMMEDIATE
  PARTITION BY RANGE(q_year_month)
  (STARTING (199201) ENDING (199212) EVERY (1),
   STARTING (199301) ENDING (199312) EVERY (1));
CREATE TABLE quan_by_month_ex(
  q_year_month  INT,
  q_count       INT NOT NULL,
  ts            TIMESTAMP,
  msg           CLOB(32K));

SET INTEGRITY FOR quan_by_month IMMEDIATE CHECKED;
CREATE INDEX qbm ON quan_by_month(q_year_month);

ALTER TABLE quan_by_month DROP MATERIALIZED QUERY;
ALTER TABLE lineitem DETACH PARTITION part0 INTO li_reuse;
ALTER TABLE quan_by_month DETACH PARTITION part0 INTO qm_reuse;

SET INTEGRITY FOR li_reuse OFF;
ALTER TABLE li_reuse ALTER l_year_month SET GENERATED ALWAYS AS
```

```

(YEAR(l_shipdate)*100 + MONTH(l_shipdate));
LOAD FROM part_mqt_rotate.del OF DEL MODIFIED BY GENERATEDIGNORE
MESSAGES load.msg REPLACE INTO li_reuse;

DECLARE load_cursor CURSOR FOR
  SELECT l_year_month, COUNT(*)
    FROM li_reuse
   GROUP BY l_year_month;
LOAD FROM load_cursor OF CURSOR MESSAGES load.msg
REPLACE INTO qm_reuse;

ALTER TABLE lineitem ATTACH PARTITION STARTING '1/1/1994'
  ENDING '1/31/1994' FROM li_reuse;

SET INTEGRITY FOR lineitem ALLOW WRITE ACCESS IMMEDIATE CHECKED
FOR EXCEPTION IN lineitem USE lineitem_ex;

ALTER TABLE quan_by_month ATTACH PARTITION STARTING 199401
  ENDING 199401 FROM qm_reuse;

SET INTEGRITY FOR quan_by_month IMMEDIATE CHECKED
FOR EXCEPTION IN quan_by_month USE quan_by_month_ex;

ALTER TABLE quan_by_month ADD MATERIALIZED QUERY
  (SELECT l_year_month AS q_year_month, COUNT(*) AS q_count
    FROM lineitem
   GROUP BY l_year_month)
  DATA INITIALLY DEFERRED REFRESH IMMEDIATE;

SET INTEGRITY FOR QUAN_BY_MONTH ALL IMMEDIATE UNCHECKED;

```

使用带有 IMMEDIATE CHECKED 选项的 SET INTEGRITY 语句，以检查已连接的数据分区是否存在完整性违例。在将表更改回 MQT 之前，需要执行此步骤。使用带有 IMMEDIATE UNCHECKED 选项的 SET INTEGRITY 语句来绕过对 MQT 进行必需的完全刷新。要获得最佳性能，MQT 的索引是必需的。建议您在适当的时候将异常表与 SET INTEGRITY 语句配合使用。

通常，在同样进行分区了的大型事实上创建已分区的 MQT。如果在大型事实上转入或转出表数据，那么必须手动调整已分区的 MQT，如示例 2 中所示。

## 示例 2: 手动调整已分区的 MQT

改变 MQT (quan\_by\_month) 以将它转换为一个普通的分区表:

```
ALTER TABLE quan_by_month DROP MATERIALIZED QUERY;
```

从事实表 (lineitem) 和 MQT 中拆离要转出的数据，并使用要转入的新数据重新装入登台表 li\_reuse:

```
ALTER TABLE lineitem DETACH PARTITION part0 INTO li_reuse;
LOAD FROM part_mqt_rotate.del OF DEL MESSAGES load.msg REPLACE INTO li_reuse;
ALTER TABLE quan_by_month DETACH PARTITION part0 INTO qm_reuse;
```

进行插入之前修剪 qm\_reuse。这将在插入子查询数据之前删除已拆离的数据。这是通过替换装入 MQT 来实现的，其中装入的数据文件是子查询的内容。

```
db2 load from datafile.del of del replace into qm_reuse
```

可以使用 INSERT INTO ... (SELECT ...) 来手动刷新表。这仅对新数据来说是必需的，因此应该在连接之前发出该语句:

```
INSERT INTO qm_reuse
  (SELECT COUNT(*) AS q_count, l_year_month AS q_year_month
   FROM li_reuse
   GROUP BY l_year_month);
```

现在，可以转入事实表的新数据：

```
ALTER TABLE lineitem ATTACH PARTITION STARTING '1/1/1994'
ENDING '1/31/1994' FROM TABLE li_reuse;
SET INTEGRITY FOR lineitem ALLOW WRITE ACCESS IMMEDIATE CHECKED FOR
EXCEPTION IN li_reuse USE li_reuse_ex;
```

接着，转入 MQT 的数据：

```
ALTER TABLE quan_by_month ATTACH PARTITION STARTING 199401
ENDING 199401 FROM TABLE qm_reuse;
SET INTEGRITY FOR quan_by_month IMMEDIATE CHECKED;
```

在连接数据分区之后，必须验证新数据以确保它在范围内。

```
ALTER TABLE quan_by_month ADD MATERIALIZED QUERY
  (SELECT COUNT(*) AS q_count, l_year_month AS q_year_month
   FROM lineitem
   GROUP BY l_year_month)
  DATA INITIALLY DEFERRED REFRESH IMMEDIATE;
SET INTEGRITY FOR QUAN_BY_MONTH ALL IMMEDIATE UNCHECKED;
```

直到 SET INTEGRITY 语句验证了这些数据之后，才能对它们进行访问。虽然支持 REFRESH TABLE 操作，但此情况演示了通过 ATTACH PARTITION 和 DETACH PARTITION 操作对已分区的 MQT 进行手动维护。用户通过 SET INTEGRITY 语句的 IMMEDIATE UNCHECKED 子句将数据标记为已验证。

---

## 创建范围集群表

### 用于范围集群表的算法

使用了一种算法来使记录的键值与表中特定行的位置相等。基本算法是相当简单的。采用其最基本的格式（使用单列而不是使用两列或多列来组成键），这种算法将序号映射至逻辑行号。

这种算法还使用记录的键来确定逻辑页号和槽号。此过程能够特别快速地访问记录；即，访问表中的特定行。

这种算法不涉及到散列，这是因为散列不会保持键值排序。必须保持键值排序，原因是它经过一段时间之后不需要重组表数据。

表中的每个记录键都应该具有以下特征：

- 唯一
- 非空
- 整数（SMALLINT、INTEGER 或 BIGINT）
- 单调增大
- 在根据键中的每一列预先确定的一组范围内

当创建允许键值超出所定义的范围的表时，使用 `ALLOW OVERFLOW` 选项。当创建键值将不超出所定义的范围的表时，使用 `DISALLOW OVERFLOW` 选项。在此情况下，如果将一条记录插入到由范围指示的边界外部，那么会返回一条 SQL 错误消息。

分布有紧密集群（密集）的序列键的应用程序很可能是范围集群表的最佳候选应用程序。当使用这种类型的键来创建范围集群表时，该键用来生成表中的一行的逻辑位置。此过程不需要具有独立的索引。

## 范围集群表索引

在范围集群表中，索引根据记录中的键找到该记录、应用启动和停止扫描并以垂直方式分布数据。通过使用 `RCT`，唯一不考虑的索引属性是垂直数据分布。

## 与常规表的差别

当决定使用范围集群表时应考虑一些特征，这些特征使范围集群表与常规表有所不同。

- 范围集群表没有可用空间控制记录（`FSCR`）。
- 预分配了空间。

预分配了表的空间并保留给表使用（即使没有填充表的记录）。创建表时，表中没有记录；但是，预分配了整个页范围。预分配是根据记录大小和要存储的最大记录数来进行的。

- 如果在每条记录中使用了变长字段（例如，`VARCHAR`），那么将使用该字段的最大长度，并且整个记录大小的长度是固定的。将每条记录的整个固定长度与最大记录数配合使用来确定需要的空间。
  - 这可能会导致分配不能有效利用的附加空间。
  - 如果键值稀疏，那么存在未使用的空间和不良范围扫描性能。
  - 即使尚未将包含这些键值的那些行插入数据库中，范围扫描也必须访问范围内的所有可能的记录。
- 不允许修改模式。

如果需要对范围集群表修改模式，那么必须重新创建表以包括它的新模式名称和旧表中的所有数据。尤其是：

- 不支持改变键范围。

这是很重要的，原因是如果需要改变表的范围，那么必须创建具有期望范围的新表，并且必须使用旧表中的数据来填充新表。

- 不允许重复键值。
- 不允许超出定义的范围的键值。

这仅适用于定义为 `DISALLOW OVERFLOW` 的范围集群表。

- 明确不允许存在 `NULL` 值。
- 范围集群索引未具体化

在系统目录中指示了带有 `RCT` 键属性的索引，并且可以由优化器来选择这些索引，但是索引并不在磁盘上具体化。对于常规表，还需要为与表相关联的每个索引提供



空间。对于 RCT，RCT 索引不需要空间。优化器使用引用此 RCT 索引的系统目录中的信息来确保可以选择对表的正确访问方法。

- 不允许根据范围集群表索引的定义创建主键或唯一键，因为它将是重复的。
- 范围集群表保留了原始键值排序，此功能部件将保证表中的行的集群。

## 关于使用范围集群表的准则

使用范围集群表（RCT）时，要遵循一些准则。

- 当定义键值的范围时，最小值是可选的；如果没有指定最小值，那么缺省值是（1）。允许最小值和最大值为负值。当使用负值时，必须明确声明最小值。例如，ORGANIZE BY KEY SEQUENCE (F1 STARTING FROM -100 ENDING AT<sup>®</sup> -10)
- 不允许对用来定义范围集群表的相同键值创建常规索引。
- 某些 ALTER TABLE 选项不可对范围集群表使用。此时选项不会影响表的物理结构，允许使用该选项。
- 因为创建范围集群表的过程中就预分配了必需的磁盘空间，所以该空间必须是可用的，否则创建表就会失败。

## SQL 编译器如何使用范围集群表

SQL 编译器处理范围集群表（RCT）的方式与处理具有辅助 B+ 树索引的常规表的方式相似。RCT 并不是通过 B+ 树索引来确定记录的位置或记录标识（RID），而是使用涉及到范围定义中的记录键值和算法的功能查找。这与具有索引的情况相似，因为可以使用键值来快速获得 RID。

当确定必需数据的最佳访问路径时，SQL 编译器使用有关表的统计信息。索引统计信息是在发出 RUNSTATS 命令时在扫描表期间收集的。对于 RCT，将表模型化为常规表，将索引模型化为基于函数的索引。

当创建允许溢出的范围集群表时，不能保证表中的记录顺序。

## 方案：范围集群表

这些方案很简单，它们演示了创建范围集群表的方法。它们说明了可以如何使用单列或多列作为一个表的键。另外，它们还说明了如何创建一个允许数据溢出的表和不允许数据溢出的表。

### 方案 1：创建范围集群表

此方案说明了一个范围集群表，它使用 STUDENT\_ID 来查找学生。每一条学生记录都包括下列信息：

- 学校标识
- 程序标识
- 学生编号
- 学生标识
- 学生的名
- 学生的姓
- 学生的学年平均成绩（GPA）

在这种情况下，学生记录只基于 STUDENT\_ID。STUDENT\_ID 将用来添加、更新和删除学生记录。

注：其他索引可以另外单独添加。但是，对于此示例，创建表时就定义了表的组织以及如何访问表的数据。

以下是此表所需要的语法：

```
CREATE TABLE STUDENTS
(SCHOOL_ID      INT NOT NULL,
 PROGRAM_ID     INT NOT NULL,
 STUDENT_NUM    INT NOT NULL,
 STUDENT_ID     INT NOT NULL,
 FIRST_NAME     CHAR(30),
 LAST_NAME      CHAR(30),
 GPA            FLOAT)
ORGANIZE BY KEY SEQUENCE
(STUDENT_ID     STARTING FROM 1 ENDING AT 1000000)
ALLOW OVERFLOW
;
```

每条记录的大小是所有列的总计。在本例中，包括 10 个字节的头 + 4 + 4 + 4 + 4 + 30 + 30 + 8 + 3（用于可空列），总共为 97 个字节。页大小为 4 KB（或者 4096 个字节），除去开销之后还有 4038 个字节，具有足够的空间用于每页 42 条记录。如果允许具有一百万条学生记录，那么将需要 23809.5 页（一百万条记录除以每页 42 条记录）。实际上将需要 23810 页（舍入）。添加四页用于表开销，添加三页用于扩展数据块映射。因此，需要预分配 23817 页（每页大小为 4 KB）。（扩展数据块映射假定使用单个容器来保存此表。因此，每个容器应该具有三页。）

## 方案 2: 创建范围集群表（不允许溢出）

在此方案中（它是第一个方案的变体）假定了教育局这样一个概念。教育局负责管理 200 所学校，每所学校有 20 间教室，每间教室可容纳 35 个学生。此教育局最多可以招收 140,000 名学生。

在这种情况下，学生记录基于三个因素：SCHOOL\_ID、CLASS\_ID 和 STUDENT\_NUM 值。这三列中的每一列都将具有唯一值，并将这三列一起来添加、更新和删除学生记录。

注：与前一个示例一样，可以另外单独添加其他索引。

以下是此表所需要的语法：

```
CREATE TABLE STUDENTS
(SCHOOL_ID      INT NOT NULL,
 CLASS_ID       INT NOT NULL,
 STUDENT_NUM    INT NOT NULL,
 STUDENT_ID     INT NOT NULL,
 FIRST_NAME     CHAR(30),
 LAST_NAME      CHAR(30),
 GPA            FLOAT)
ORGANIZE BY KEY SEQUENCE
(SCHOOL_ID     STARTING FROM 1 ENDING AT 200,
 CLASS_ID      STARTING FROM 1 ENDING AT 20,
 STUDENT_NUM   STARTING FROM 1 ENDING AT 35)
DISALLOW OVERFLOW
;
```

在本例中，不允许溢出。这样做是有道理的，因为教育局的政策可能会限制每个班级允许招收的学生人数。在此方案中，每个班级最多只能有 35 名学生。如果将此因素与教室数和学校数实际存在的限制联系起来考虑，就不难理解教育局为何不允许学生超员。

一些学校的教室数目可能会发生一些变化。如果是这种情况，当（使用 CLASS\_ID）定义教室数的范围时，上限应该是将所有学校考虑在内时所得到的最大教室数。这可能意味着某些较小的学校（与最大的学校比较起来教室数更少的那些学校）将具有用于从未使用的学生记录的空间（除非，例如，为学校添加移动教室）。

与前一个示例中一样，使用相同的 4 KB 页大小和相同的学生记录大小，每一页可以具有 42 条记录。因此，140,000 条学生记录就将需要 3333.3 页，进行舍入之后就是 3334 页。另外，表信息还需要两页，扩展数据块映射需要三页。因此，需要预分配 3339 页（每页大小为 4 KB）。

---

## 创建 MDC 表时的注意事项

当创建 MDC 表时应该考虑许多因素。下列各节讨论了您当前所处的数据库环境（例如，是否具有分区数据库）和为 MDC 表选择的维对于您创建、布置和使用 MDC 表将产生的影响。还讨论了“DB2 设计顾问程序”以及可以如何使用它来提供有关这样一些问题的建议。

### 将数据从现有表移至 MDC 表

要提高数据仓库或大型数据库环境中的查询性能和减少数据维护操作的开销，可以将数据从常规表移至多维集群（MDC）表。要将数据从现有表移至 MDC 表：导出数据，删除原始表（可选），创建多维集群（MDC）表（使用带有 ORGANIZE BY DIMENSIONS 子句的 CREATE TABLE 语句），并将您自己的数据装入 MDC 表。

可以使用称为 SYSPROC.ALTOBJ 的 ALTER TABLE 过程来将现有表中的数据转换为 MDC 表中的数据。可以从“DB2 设计顾问程序”中调用此过程。在这两个表之间转换数据所需要的时间可能很长，这取决于表的大小以及需要转换的数据量。

当改变表时，ALTOBJ 过程将执行下列操作：

- 删除表的所有从属对象
- 重命名表
- 使用新定义来创建表
- 重新创建表的所有从属对象
- 将表中的现有数据变换为新表中需要的数据。也就是说，从旧表中选择数据，然后将该数据装入新表中，可以使用列函数来将旧的数据类型变换为新的数据类型。

### SMS 表空间中的 MDC 表

如果打算将 MDC 表存储在 SMS 表空间中，那么需要使用多页文件分配。（对于在版本 8.2 和更新版本中新创建的数据库，多页文件分配是缺省值。）这是因为 MDC 表总是按整个扩展数据块来扩展的，并且这些扩展数据块中的所有页在物理上保持连续很重要。因此，禁用多页文件分配不利于利用空间；而且，启用多页文件分配将显著提高每个扩展数据块中的页在物理上连续的可能性。

### DB2 设计顾问程序上的 MDC 顾问程序功能部件

“DB2 设计顾问程序”（db2advis）（以前称为“索引顾问程序”）具有 MDC 功能部件。此功能部件建议用于 MDC 表中的集群维（包括基本列的粗糙度）以便提高工作负载性

能。粗糙度这个术语表示用来减小集群维的基数（单值的数目）的一个数学表达式。粗糙度的常见示例是日期，日期的粗糙度可为日期、日期所在的星期、日期所在的月份或一年中的季度。

使用“DB2 设计顾问程序”的 MDC 功能部件要求数据库中至少存在几个扩展数据块的数据。“DB2 设计顾问程序”使用数据来对数据密度和基数建立模型。

如果数据库的表中没有数据，那么“DB2 设计顾问程序”不会建议使用 MDC，即使该数据库包含空表，但有一组虚假的统计信息来表示它是一个已填充的数据库。

建议还标识了用来定义维的粗糙度的潜在生成列。建议中不包括可能的块大小。在为 MDC 表提供建议时，使用表空间的扩展数据块大小。假定将在现有表所在的表空间中创建建议的 MDC 表，因此该表将具有相同的扩展数据块大小。对 MDC 维的建议将根据表空间的扩展数据块大小变化而变化，这是因为扩展数据块大小将影响可以填充到块或单元中的记录数。这将直接影响单元的密度。

只考虑单列维而不考虑组合列维，尽管可以为表建议单个维或多个维。MDC 功能部件将建议大多数受支持的数据类型使用的粗糙度，目标是减小所采用的 MDC 解决方案中的单元的基数。异常的数据类型包括：CHAR、VARCHAR、GRAPHIC 和 VARGRAPH 数据类型。所有受支持的数据类型都将被强制类型转换为 INTEGER 并通过生成的表达式来设置粗糙度。

“DB2 设计顾问程序”的 MDC 功能部件的目标是选择可提高性能的 MDC 解决方案。另一个目标是将数据库的存储器扩充限制在适当的级别。使用统计方法来确定每个表的最大存储器扩充。

顾问程序中的分析操作既会利用块索引访问的优点，也会受到 MDC 对表的维执行插入、更新和删除操作的影响。对表执行这些操作时可能会导致在各个单元之间移动记录。分析操作还会模拟对特定 MDC 维上的数据进行组织时产生的任何表扩充而对性能造成的潜在影响。

通过对 db2advise 实用程序使用 -m <advise type> 标志来启用 MDC 功能部件。使用“C”建议类型来指示多维集群表。建议类型为：“I”表示索引、“M”表示具体化查询表、“C”表示 MDC，而“P”表示分区数据库环境。建议类型可以相互组合使用。

注：“DB2 设计顾问程序”将不会处理小于 12 个扩展数据块大小的表。

当提出建议时，顾问程序将同时分析 MQT 和常规基本表。

MDC 功能部件的输出包括：

- 每个表的生成列表表达式（用于对 MDC 解决方案中出现的维设置粗糙度）。
- 为每个表建议的 ORGANIZE BY 子句。

对标准输出和作为说明工具的一部分的 ADVISE 表报告了建议。

## MDC 表和分区数据库环境

多维集群可以与分区数据库环境联合使用。实际上，MDC 可以作为分区数据库环境的补充。分区数据库环境用来将一个表中的数据分配到多个物理节点或逻辑节点上，以便达到下列目的：

- 利用多台机器来并行增加处理请求。

- 增大表的物理大小（超过单个数据库分区的限制）。
- 提高数据库的可伸缩性。

分发表的原因与该表是 MDC 表还是常规表无关。例如，选择用来组成分布键的列的规则是相同的。MDC 表的分布键可以包括任何列，无论这些列是否组成表的维的一部分。

如果分布键与表的某个维完全相同，那么每个数据库分区都将包含该表的不同部分。例如，如果作为示例的 MDC 表按颜色分发给两个数据库分区上，那么将使用 Color 列来划分数据。因此，可能在一个数据库分区上找到 Red 和 Blue 片，而在另一个数据库分区上找到 Yellow 片。如果分布键与表中的维不完全相同，那么每个数据库分区都将具有每个片的数据的子集。当选择维和估计单元占用率时（请参阅『单元密度』一节），注意，平均来说，每个单元的数据总量是通过用所有数据除以数据库分区数来确定的。

## 具有多个维的 MDC 表

如果知道将在查询中大量使用特定谓词，那么可以使用 ORGANIZE BY DIMENSIONS 子句根据涉及到的列对表进行集群。

示例 1:

```
CREATE TABLE T1 (c1 DATE, c2 INT, c3 INT, c4 DOUBLE)
      ORGANIZE BY DIMENSIONS (c1, c3, c4)
```

示例 1 中的表根据形成逻辑立方体（即，具有三个维）的三个本地列中的值集群。现在可以在查询处理期间根据一个或多个维对表进行逻辑分片，以便涉及的关系运算符仅处理相应的片或单元中的块。注意，块的大小（页数）将是表的扩展数据块大小。

## 具有基于多列的维的 MDC 表

每个维可以由一列或多列组成。作为一个示例，可以创建一个根据包含两列的一个维来集群的表。

示例 2:

```
CREATE TABLE T1 (c1 DATE, c2 INT, c3 INT, c4 DOUBLE)
      ORGANIZE BY DIMENSIONS (c1, (c3, c4))
```

在示例 2 中，表将根据两个维 c1 和 (c3, c4) 来进行集群。这样，在查询处理期间，表可以根据 c1 维或组合 (c3, c4) 维逻辑分片。该表将与示例 1 中的表具有相同数目的块，但是少一个维块索引。在示例 1 中，将有三个维块索引，列 c1、c3 和 c4 各一个。在示例 2 中，将有两个维块索引，一个针对列 c1，而另一个针对 c3 和 c4。这两个方法的主要不同之处在于，在示例 1 中，仅涉及 c4 的查询可以使用 c4 的维块索引来快速直接地访问相关数据块。在示例 2 中，c4 是维块索引中的辅助键部分，因此仅涉及 c4 的查询涉及更多的处理。但是，在示例 2 中，将少维护和存储一个块索引。

“DB2 设计顾问程序”将不对包含多列的维提供建议。

## 将列表表达式作为维的 MDC 表

列表表达式也可用于集群维。根据列表表达式集群的功能对于将维上滚至更低的详细程度非常有用，例如，将地址上滚为地理位置或区域，或者将日期上滚为星期、月份或

年。要以此方式实现维的上滚，可以使用生成列。此类型的列定义将允许使用可表示维的表达式创建列。在示例 3 中，该语句创建根据一个基本列和两个列表表达式进行集群的表。

示例 3:

```
CREATE TABLE T1(c1 DATE, c2 INT, c3 INT, c4 DOUBLE,  
                c5 DOUBLE GENERATED ALWAYS AS (c3 + c4),  
                c6 INT GENERATED ALWAYS AS (MONTH(C1)))  
    ORGANIZE BY DIMENSIONS (c2, c5, c6)
```

在示例 3 中，c5 列是基于 c3 和 c4 列的表达式，而 c6 列会将 c1 列上滚至更低的详细程度。此语句将根据 c2、c5 和 c6 列中的值集群该表。

## 对生成列维的范围查询

对生成列维的范围查询需要单调列函数。表达式必须是单调的才能为生成列的维派生范围谓词。如果对生成列创建维，那么对基本列的查询将能够利用生成列的块索引来提高性能（有一种情况例外）。要使基本列（例如，日期）的范围查询对维块索引使用范围扫描，用来在 CREATE TABLE 语句中生成列的表达式必须是单调的。尽管列表表达式可以包括任何有效表达式（包括用户定义的函数（UDF）），但是如果表达式不是单调的，那么当等价谓词或 IN 谓词都在基本列上时，它们才能够使用块索引来满足查询。

作为一个示例，假定使用生成列 month 的维来创建 MDC 表，其中 month = INTEGER (date)/100。对于该维（month）的查询，可以执行块索引扫描。对于基本列（date）的查询，也可以执行块索引扫描来缩小要扫描的块的范围，然后只将日期的谓词应用于这些块中的行。

编译器将生成要在块索引扫描中使用的其他谓词。例如，对于以下查询：

```
SELECT * FROM MDCTABLE WHERE DATE > "19999/03/03" AND DATE < "2000/01/15"
```

编译器将生成以下谓词：“month >= 199903”和“month < 200001”，它们可以用作维块索引扫描的谓词。当对获得的块进行扫描时，会将原始谓词应用于这些块中的行。

非单调表达式将只允许对该维应用等价谓词。非单调函数的一个较好的示例是 MONTH()，如示例 3 中的 c6 列的定义所示。如果 c1 列是日期、时间戳记或日期或时间戳记的有效字符串表示法，那么函数将返回范围是 1 到 12 的整数值。尽管函数的输出是确定的，但是实际上它生成的输出与阶跃函数（即，循环模式）相似：

```
MONTH(date('99/01/05')) = 1  
MONTH(date('99/02/08')) = 2  
MONTH(date('99/03/24')) = 3  
MONTH(date('99/04/30')) = 4  
...  
MONTH(date('99/12/09')) = 12  
MONTH(date('00/01/18')) = 1  
MONTH(date('00/02/24')) = 2  
...
```

尽管此示例中的日期是连续增加的，但是 MONTH(date) 不会增加。更具体地说，每当 date1 大于 date2，并不能保证 MONTH(date1) 大于或等于 MONTH(date2)。这是单调性所要求的。此非单调性是允许的，但是它限制了维，基本列的范围谓词不能生成维的范围谓词。但是，表达式的范围谓词是可以的，例如，where month(c1) between 4 and 6。这可以采用常规方式使用维的索引，起始键为 4 而停止键为 6。

要使此函数单调，必须将年包括为月份的高位部分。版本 9.2 提供对 INTEGER 内置函数的扩展以帮助根据日期定义单调表达式。INTEGER(date) 返回日期的整数表示法，可以分开查找年和月份的整数表示法。例如，INTEGER(date('2000/05/24')) 返回 20000524，因此  $\text{INTEGER}(\text{date}('2000/05/24'))/100 = 200005$ 。函数  $\text{INTEGER}(\text{date})/100$  是单调的。

相似的，内置函数 DECIMAL 和 BIGINT 也具有扩展，所以可以派生单调函数。DECIMAL(timestamp) 返回时间戳记的十进制表示法，可以在单调表达式中使用它来派生月份、天、小时或分钟等等的增加的值。BIGINT(date) 返回日期的大整数表示法，类似于 INTEGER(date)。

只要可能，数据库管理器将在为表创建生成列或者根据维子句中的表达式创建维时确定表达式的单调性。特定函数被识别为保留单调性，例如，DATENUM( )、DAYS() 和 YEAR( )。并且，列和常量的各种算术表达式，例如，除法、乘法或加法是保留单调性的。当 DB2 确定表达式不保留单调性时，或者如果它不能确定这一点，该维将仅支持对其基本列使用等价谓词。





---

## 第 12 章 改变数据库

---

### 改变实例

#### 更改多个数据库分区中的数据库配置

当数据库分布在多个数据库分区上时，数据库配置文件在所有数据库分区上应该相同。

一致性是必需的，因为 SQL 编译器根据节点配置文件中的信息来编译分布式 SQL 语句，并创建一个访问方案以满足 SQL 语句的需要。维护数据库分区上的不同配置文件可能产生不同的访问方案，这取决于预编译该语句所在的数据库分区。使用 db2\_all 来跨所有数据库分区维护配置文件。

---

### 改变数据库

#### 改变数据库分区组

要使用命令行处理器改变数据库分区组：使用 REDISTRIBUTE DATABASE PARTITION GROUP 命令。一旦添加或删除了数据库分区，就必须在数据库分区组中新的一组数据库分区之间再分发当前数据。

#### 从“控制中心”管理数据库分区

可以通过“控制中心”的“数据库分区”视图来使用数据库分区。

要使用数据库分区，将需要具有连接至实例的权限。具有 SYSADM 或 DBADM 权限的任何人都可以授予您访问特定实例的权限。

要配置数据库分区或使数据库分区脱离前滚暂挂状态，您必须具有 SYSADM、SYSCTRL 或 SYSMANT 权限。

可以使用“数据库分区”视图来重新启动数据库分区、使数据库分区退出前滚暂挂状态、备份数据库分区、复原数据库分区或使用“配置顾问程序”配置数据库分区。

要从“控制中心”打开“数据库分区”视图：

1. 从“控制中心”中展开对象树，直到找到要查看其数据库分区的分区数据库为止。
2. 右键单击想要的分区数据库并从菜单列表中选择“打开数据库分区”。
3. 所选分区数据库的“数据库分区”视图将打开。

要配置数据库分区：

1. 从“数据库分区”视图中选择想要的数据库分区。
2. 选择“数据库分区”，单击右键，然后从列表中选择“配置顾问程序”。
3. “配置顾问程序”将打开。使用“配置顾问程序”来指定数据库配置参数的值。



---

## 第 13 章 改变表和其他相关表对象

---

### 改变分区表

分区表支持 ALTER TABLE 语句的所有相关子句。此外，ALTER TABLE 语句允许添加 (ADD) 新数据分区、转入 (ATTACH) 新数据分区以及转出 (DETACH) 现有数据分区。

要改变分区表以拆离数据分区，用户必须拥有下列权限或特权：

- 执行 DETACH 操作的用户必须有权对源表执行 ALTER、SELECT 和 DELETE 操作。
- 该用户还必须有权创建目标表。因此，要改变一个表以拆离数据分区，语句授权标识拥有的特权必须至少包括对目标表的下列其中一项权限或特权：
  - SYSADM 或 DBADM 权限
  - 对数据库的 CREATETAB 权限、对该表所使用表空间的 USE 特权以及下列其中一项权限或特权：
    - 对数据库的 IMPLICIT\_SCHEMA 权限（如果该表的隐式或显式模式名不存在的话）
    - 对模式的 CREATEIN 特权（如果该表的模式名引用现有模式的话）

要改变分区表以连接数据分区，语句授权标识拥有的特权必须至少包括对源表的下列其中一项权限或特权：

- 对源表的 SELECT 特权以及对源表模式的 DROPIN 特权
- 对源表的 CONTROL 特权
- SYSADM 或 DBADM 权限。

要改变分区表以添加数据分区，语句授权标识必须有权使用要添加新分区的表空间，并且其拥有的特权必须至少包括对源表的下列其中一项权限或特权：

- ALTER 特权
- CONTROL 特权
- SYSADM
- DBADM
- 对表模式的 ALTERIN 特权

#### 用法准则

- 在指定了 PARTITION 子句情况下发出的每个 ALTER TABLE 语句都必须在单独的 SQL 语句中。
- 在包含 ALTER TABLE...PARTITION 操作的 SQL 语句中，不允许执行任何其他 ALTER 操作。例如，在单个 SQL 语句中，不能同时连接数据分区和对表添加列。
- 可以执行多个 ALTER 语句，接着执行一个 SET INTEGRITY 语句。

要使用命令行来改变分区表，请发出 ALTER TABLE 语句。

---

## 有关改变分区表的准则和限制

本主题标识了在存在已连接和已拆离的数据分区的情况下最常用的改变表操作和注意事项。

### 添加或改变检查约束或外键约束

已连接和已拆离的数据分区支持添加检查约束或外键约束。

**添加列** 在对带有已连接数据分区的表添加列时，还将对已连接数据分区添加该列。由于已拆离数据分区在物理上不再与该表相关联，所以在对带有已拆离数据分区的表添加列时，不会对已拆离数据分区添加该列。

**改变列** 当改变带有已连接数据分区的表中的列时，还将对已连接数据分区改变该列。当改变带有已拆离数据分区的表中的列时，由于已拆离数据分区在物理上不再与该表相关联，不会改变已拆离数据分区上的该列。

### 添加生成列

当将生成列添加到带有已连接或已拆离数据分区的分区表时，必须遵循任何其他列类型的添加规则。

### 添加或修改索引

当对带有已连接数据分区的表创建、重新创建或重组索引时，该索引不包括已连接数据分区中的数据，这是因为 `SET INTEGRITY` 语句维护所有已连接数据分区的所有索引。当对带有已拆离数据分区的表创建、重新创建或重组索引时，除非已拆离数据分区带有需要根据该数据分区进行递增刷新的已拆离从属表或登台表，否则索引不包括已拆离数据分区中的数据。在后面这种情况下，索引将包括此已拆离数据分区的数据。

### WITH EMPTY TABLE

不能清空带有已连接数据分区的表。

### ADD MATERIALIZED QUERY AS

不允许将带有已连接数据分区的表改变为 MQT。

### 改变数据分区中存储的其他表属性

在数据分区中还存储了下列表属性。对这些属性所作的更改将反映到已连接数据分区中，但不会反映到已拆离数据分区中。

- DATA CAPTURE
- VALUE COMPRESSION
- APPEND
- COMPACT/LOGGED FOR LOB COLUMNS
- ACTIVATE NOT LOGGED INITIALLY (WITH EMPTY TABLE)

---

## 连接数据分区

表分区功能提高了表数据的转入和转出效率。这是通过使用 `ALTER TABLE` 语句的 `ATTACH PARTITION` 和 `DETACH PARTITION` 子句实现的。通过转入分区表数据，可以方便地将新范围作为附加数据分区合并到分区表中。

要改变一个表以连接数据分区，语句授权标识拥有的特权必须至少包括对源表的下列其中一项权限和特权：

- 对源表的 `SELECT` 特权以及对源表模式的 `DROPIN` 特权

- 对源表的 CONTROL 特权
- SYSADM 或 DBADM 权限。

ATTACH PARTITION 子句使用现有的表（源表）并将其作为新数据分区与目标表连接。最初，查询无法访问新连接的数据分区。该表的其余部分处于联机状态。需要调用 SET INTEGRITY 语句以使连接的数据分区转入联机状态。

### 限制和用法准则

要连接数据分区，必须符合下列条件：

- 要连接新数据分区的表（即目标表）必须是现有的分区表。
- 源表必须是现有的非分区表，或者是只有单个数据分区且没有 ATTACHED 或 DETACHED 数据分区的分区表。要连接多个数据分区，有必要发出多个 ATTACH 语句。
- 源表不能为分层表（类型表）。
- 源表不能是范围集群表（RCT）。
- 源表的表定义必须与目标表匹配。
- 源表与目标表在列数目、类型和顺序方面必须匹配。
- 对于这两个表来说，列在是否包含缺省值方面必须匹配。如果源列是使用 ALTER TABLE ADD COLUMN 创建的，即 SYSCOLUMNS.ADD\_DEFAULT = 'Y'，那么 existDefault 值（SYSCOLUMNS.ADDED\_DEFAULT）必须与目标列的该值匹配。
- 对于这两个表来说，列在是否允许 NULL 方面必须匹配。
- 对于源表和目标表来说，同时包含 VALUE COMPRESSION 和 SYSTEM COMPRESSION DEFAULT 值的 Compression 子句必须匹配。
- 带有 data capture 选项和 not logged initially 选项的 APPEND 子句的使用必须匹配。
- 即使在目标列是生成列而源列不是生成列时，也允许连接数据分区。语句 SET INTEGRITY FOR T ALLOW WRITE ACCESS IMMEDIATE CHECKED FORCE GENERATED 将生成所连接行的生成列值。与生成列匹配的列在类型和可空性方面必须匹配。此列没有必需的缺省值。建议的方法是保证 ATTACH 的源表的生成列中包含正确的生成值。这样，就不需要使用 FORCE GENERATED 选项。可以使用以下语句：
 

```
SET INTEGRITY FOR T GENERATED COLUMN IMMEDIATE UNCHECKED
（指示系统不检查生成列）
SET INTEGRITY FOR T ALLOW WRITE ACCESS IMMEDIATE CHECKED FOR EXCEPTION
IN T USE T_EX（对连接的分区执行完整性检查，但不检查生成列是否正确）
```
- 即使在目标列是标识列而源列不是标识列时，也允许连接数据分区。语句 SET INTEGRITY IMMEDIATE CHECKED 不为连接的行生成标识值。语句 SET INTEGRITY FOR T GENERATE IDENTITY ALLOW WRITE ACCESS IMMEDIATE CHECKED 填写所连接的行的标识值。与标识列匹配的列在类型和可空性方面必须匹配。对此列的缺省值没有要求。建议的方法是在登台表中填写正确的标识值。于是，由于已在源表中保证了标识值，所以不需要在执行 ATTACH 后使用 GENERATE IDENTITY 选项。
- 对于数据分布在多个数据库分区中的表来说，源表必须也使用同一分布键和同一个分布图分布在同一个数据库分区组中。
- 源表必须是可删除的（即，不能对其设置 RESTRICT DROP）。

- 如果指定了 DATAPARTITIONNAME, 那么它在目标表中不能已存在。
- 如果目标表是多维集群 (MDC) 表, 那么源表也必须是 MDC 表。
- 源表的数据表空间在类型 (即 DMS 或 SMS)、页大小、扩展数据块大小和数据库分区组方面必须与目标表的数据表空间匹配。如果预取大小不匹配, 就会返回警告给用户。源表的大型表空间在类型、数据库分区组和页大小方面必须与目标表的大型表空间匹配。

要使用 DB2 命令行来改变分区表以及将某个数据分区与该表连接, 请发出 ALTER TABLE 语句。

---

## 将数据分区连接至分区表的准则

本主题提供了一些准则, 可用于指导通过发出 ALTER TABLE ...ATTACH PARTITION 语句来更正尝试将数据分区连接至数据库时出现的各种类型的不匹配情况。通过将源表修改为与目标表特征匹配, 或者通过将目标表修改为与源表特征匹配, 可以使这两个表匹配。

源表是想要连接至目标表的现有表。目标表是想要将新数据分区连接至的表。

要成功执行连接, 一种建议的方法是对源表使用与目标表完全相同的 CREATE TABLE 语句, 但不带 PARTITION BY 子句。在难以修改源表或目标表特征以实现兼容的情况下, 可以创建与目标表兼容的新源表。有关创建新源表的详细信息, 请参阅创建类似于现有表的表。

为了避免出现不匹配情况, 请参阅第 174 页的『连接数据分区』中的“限制和用法准则”一节。该节概述了成功连接数据分区之前必须满足的条件。如果无法满足所列示的条件, 就会返回错误 SQL20408N 或 SQL20307N。

下列各节描述了可能出现的各种类型的不匹配情况, 并提供了一些建议步骤来使两个表匹配:

**(值) 压缩子句 (SYSCAT.TABLES 的 COMPRESSION 列) 不匹配。(SQL20307N 原因码 2)**

要使值压缩一致, 请使用下列其中一个语句:

```
ALTER TABLE... ACTIVATE VALUE COMPRESSION
或者
ALTER TABLE... DEACTIVATE VALUE COMPRESSION
```

要使行压缩值匹配, 请使用下列其中一个语句:

```
ALTER TABLE... COMPRESS YES
或者
ALTER TABLE... COMPRESS NO
```

**表的 APPEND 方式不匹配。(SQL20307N 原因码 3)**

要使追加方式匹配, 请使用下列其中一个语句:

```
ALTER TABLE ... APPEND ON
或者
ALTER TABLE ... APPEND OFF
```

**源表与目标表的代码页不匹配。(SQL20307N 原因码 4)**

创建一个新的源表

源表是包含多个数据分区或包含已连接或已拆离数据分区的分区表。（**SQL20307N** 原因码 5）

使用以下语句来从源表中拆离数据分区，直到只剩下一个可视数据分区为止：

```
ALTER TABLE ... DETACH PARTITION
```

包括任何必需的 **SET INTEGRITY** 语句。如果源表具有索引，那么可能无法立即连接源表。已拆离的数据分区保持已拆离状态，直到所有索引清除了已拆离的键为止。如果想要立即执行连接，那么删除源表上的索引。否则，创建新的源表。

源表是系统表、视图、类型表、按键序列组织的表或者已声明的全局临时表。（**SQL20307N** 原因码 6）

创建一个新的源表。

目标表与源表是同一个表。（**SQL20307N** 原因码 7）

不能将表连接到它自身。确定正确的表以用作源表或目标表。

对源表或目标表指定了 **NOT LOGGED INITIALLY** 子句，但未对两个表都指定此子句。（**SQL20307N** 原因码 8）

通过发出 **COMMIT** 语句，对最初未进行日志记录的表进行日志记录，或者通过输入以下语句来将进行了日志记录的表更改为最初不进行日志记录：

```
ALTER TABLE .... ACTIVATE NOT LOGGED  
INITIALLY
```

对源表或目标表指定了 **DATA CAPTURE CHANGES** 子句，但未对两个表都指定此子句。（**SQL20307N** 原因码 9）

要对未打开数据捕获更改功能的表启用数据捕获更改功能，请运行以下语句：

```
ALTER TABLE ... DATA CAPTURE CHANGES
```

要对已打开数据捕获更改功能的表禁用数据捕获更改功能，请运行以下语句：

```
ALTER TABLE ... DATA CAPTURE NONE
```

表的分布子句不匹配。源表与目标表的分布键必须相同。（**SQL20307N** 原因码 10）

建议您创建一个新的源表。对于跨多个数据库分区的表，不能更改其分布键。要对单分区数据库中的表更改分布键，请运行下列语句：

```
ALTER TABLE ... DROP DISTRIBUTION;  
ALTER TABLE ... ADD DISTRIBUTION(key-specification)
```

仅对其中一个表指定了 **ORGANIZE BY DIMENSIONS** 子句，或者组织维不同。（**SQL20307N** 原因码 11）

创建一个新的源表。

列的数据类型（**TYPENAME**）不匹配。（**SQL20408N** 原因码 1）

要更正数据类型的不匹配，可发出以下语句：

```
ALTER TABLE ... ALTER COLUMN ... SET DATA TYPE...
```

列的可空性 (**NULLS**) 不匹配。 (**SQL20408N** 原因码 2)

要改变其中一个表中不匹配的列的可空性，发出下列语句：

```
ALTER TABLE... ALTER COLUMN...  
                                DROP NOT NULL
```

或者

```
ALTER TABLE... ALTER COLUMN...  
                                SET NOT NULL
```

列的隐式缺省值 (**SYSCAT.COLUMNS IMPLICITVALUE**) 不兼容。 (**SQL20408N** 原因码 3)

创建一个新的源表。如果目标表列和源表列都具有隐式缺省值 (且 **IMPLICITVALUE** 不为 **NULL**)，那么这两个隐式缺省值必须完全匹配。

如果对于目标表中的一列，**IMPLICITVALUE** 不是 **NULL**，并且对于源表的相应列，**IMPLICITVALUE** 不是 **NULL**，那么会在对表执行原始的 **CREATE TABLE** 语句之后添加每一列。在此情况下，存储在 **IMPLICITVALUE** 中的值必须对于此列是相匹配的。

有一种情况是，通过从版本低于 V9.1 的表进行迁移或者从版本低于 V9.1 的表连接数据分区之后，**IMPLICITVALUE** 不是 **NULL**，因为系统不知道在执行原始 **CREATE TABLE** 语句之后是否添加了列。如果数据库无法确定该列是否是添加的列，那么将其视为如此。添加的列就是作为 **ALTER TABLE ...ADD COLUMN** 语句的结果创建的列。在此情况下，不允许执行该语句，因为如果允许继续进行连接的话，该列可能会被破坏。必须将源表中的数据复制到新表中 (此列的 **IMPLICITVALUE** 为 **NULL**)，并将新表用作连接操作的源表。

列的代码页 (**COMPOSITE\_CODEPAGE**) 不匹配。 (**SQL20408N** 原因码 4)

创建一个新的源表。

系统压缩缺省子句 (**COMPRESS**) 不匹配。 (**SQL20408N** 原因码 5)

要改变列的系统压缩，发出下面的其中一个语句来更正不匹配的情况：

```
ALTER TABLE ... ALTER COLUMN ...  
                                COMPRESS SYSTEM DEFAULT
```

或者

```
ALTER TABLE ... ALTER COLUMN ...  
                                COMPRESS OFF
```

---

## 拆离数据分区

表分区功能提高了表数据的转入和转出效率。这是通过使用 **ALTER TABLE** 语句的 **ATTACH PARTITION** 和 **DETACH PARTITION** 子句实现的。

要从分区表拆离数据分区，您必须具有下列权限或特权：

- 执行 **DETACH** 操作的用户必须有权对源表执行 **ALTER**、**SELECT** 和 **DELETE** 操作。
- 该用户还必须有权创建目标表。因此，要改变一个表以拆离数据分区，语句授权标识拥有的特权必须至少包括对目标表的下列其中一项权限或特权：
  - **SYSADM** 或 **DBADM** 权限



- 对数据库的 `CREATETAB` 权限、对该表所使用表空间的 `USE` 特权以及下列其中一项权限或特权：
  - 对数据库的 `IMPLICIT_SCHEMA` 权限（如果该表的隐式或显式模式名不存在的话）
  - 对模式的 `CREATEIN` 特权（如果该表的模式名引用现有模式的话）

**注：**拆离数据分区时，语句授权标识将会有效地执行 `CREATE TABLE` 语句，因此必须具有执行该操作所必需的特权。`ALTER TABLE` 语句的授权标识将成为新表的定义者（具有 `CONTROL` 权限），就像是该用户发出了 `CREATE TABLE` 语句一样。不会将所改变的表的任何特权传递至新表。只有 `ALTER TABLE` 语句的授权标识以及 `DBADM` 或 `SYSADM` 才有权在 `ALTER TABLE ...DETACH PARTITION` 操作执行后立即访问数据。

通过转出分区表数据，可以方便地从分区表中分离出某些范围的数据。一旦将数据分区拆离成单独的表，就可以通过多种方法处理该表。可以废弃单独的表（这样就破坏了数据分区中的数据）；可以对它进行归档或者以别的方式将它作为单独的表使用；将它连接到另一个分区表（例如历史表）；也可以对它进行操纵、清理和变换以及将它重新连接到原始分区表或另一分区表。

### 限制

如果源表是多维集群表（MDC），那么不允许在 `ALTER TABLE ...DETACH` 操作所在的工作单元中访问新拆离的表。在落实 `ALTER TABLE ...DETACH` 操作之后，根据对表的第一次访问创建块索引。创建块索引后，访问时间将减少。

要执行 `DETACH` 操作，必须符合下列条件：

- 要从中拆离数据分区的表（源表）必须存在，并且必须是分区表。
- 要拆离的数据分区必须存在于源表中。
- 源表必须有多个数据分区。分区表必须至少有一个数据分区。只有已连接的可视数据分区才能用于此上下文。已连接数据分区就是已连接但尚未使用 `SET INTEGRITY` 语句进行验证的数据分区。
- `DETACH` 操作将要创建的表（目标表）的名称不能已存在。
- 不允许对强制实施的引用完整性（RI）关系的父表执行 `DETACH`。
- 如果有任何需要根据已拆离数据分区进行递增维护的从属表（这些从属表称为已拆离的从属表），那么新拆离的表最初不可访问。对于该表，在 `SYSCAT.TABLES` 目录视图的 `TYPE` 列中将标记 `L`。此表称为已拆离表。这样，在运行 `SET INTEGRITY` 语句以递增方式维护已拆离的从属表之前，无法读取、修改或删除这个已拆离表。在对所有已拆离的从属表运行 `SET INTEGRITY` 语句后，已拆离表就会转换为可供完全访问的正规表。

要改变分区表以及从该表中拆离数据分区，发出带 `DETACH PARTITION` 子句的 `ALTER TABLE` 语句。

---

## 已拆离数据分区的属性

使用 ALTER TABLE 语句的 DETACH PARTITION 子句从分区表拆离数据分区后，该数据分区将成为独立目标表。新目标表的许多属性是从源表继承的。对于任何未从源表继承的属性来说，它们的设置就像是执行 DETACH 操作的用户创建了该目标表一样。

**注：**如果有已拆离的从属表，那么已拆离数据分区在拆离时不会变成独立表。在这种情况下，必须发出 SET INTEGRITY 语句以完成拆离并使该表可访问。

### 目标表继承的属性

目标表继承的属性包括：

- 下列列定义：
  - 列名
  - 数据类型（对于具有长度和精度的数据类型（例如 CHAR 和 DECIMAL）来说，还包括长度和精度）
  - 可空性
  - 列缺省值
  - 代码页（SYSCAT.COLUMNS 目录视图的 CODEPAGE 列）
  - LOB 日志记录（SYSCAT.COLUMNS 目录视图的 LOGGED 列）
  - LOB 压缩（SYSCAT.COLUMNS 目录视图的 COMPACT 列）
  - 压缩（SYSCAT.COLUMNS 目录视图的 COMPRESS 列）
  - 隐藏列类型（SYSCAT.COLUMNS 目录视图的 HIDDEN 列）
  - 列顺序
- 如果源表是多维集群表（MDC），那么目标表也是 MDC 表，并且是使用相同的维列定义的。当源表是 MDC 时，不允许在拆离操作所在的工作单元中访问新拆离的表。
- 块索引定义。在落实 DETACH 操作后第一次访问新拆离的独立表时，将重建索引。
- 表空间标识和表对象标识继承自数据分区，而不是继承自源表。这是因为在 DETACH 操作期间未移动任何表数据。在目录方面，源数据分区中 SYSCAT.DATAPARTITIONS 目录视图的 TBSPACEID 列将成为 SYSCAT.TABLES 目录视图的 TBSPACEID 列。当转换成表空间名称时，目标表中有 SYSCAT.TABLES 目录视图的 TBSPACE 列。源数据分区中 SYSCAT.DATAPARTITIONS 目录视图的 PARTITIONOBJECTID 列将成为目标表中 SYSCAT.TABLES 目录视图的 TABLEID 列。
- 源数据分区中 SYSCAT.DATAPARTITIONS 目录视图的 LONG\_TBSPACEID 列将转换为表空间名并成为目标表中 SYSCAT.TABLES 的 LONG\_TBSPACE 列。
- 表空间位置
- 多分区数据库的分布图标识（SYSCAT.TABLES 目录视图的 PMAP\_ID 列）
- 可用百分比（SYSCAT.TABLES 目录视图的 PCTFREE 列）
- 追加方式（SYSCAT.TABLES 目录视图的 APPEND\_MODE 列）
- 首选锁定详细程度（SYSCAT.TABLES 目录视图的 LOCKSIZE 列）
- 数据捕获（SYSCAT.TABLES 目录视图的 DATA\_CAPTURE 列）
- VOLATILE（SYSCAT.TABLES 目录视图的 VOLATILE 列）

- DROPRULE (SYSCAT.TABLES 目录视图的 DROPRULE 列)
- 压缩 (SYSCAT.TABLES 目录视图的 COMPRESSION 列)
- 搜索最大可用空间 (SYSCAT.TABLES 目录视图的 MAXFREESPACESEARCH 列)

注: 不支持分区的分层表或临时表、范围集群表和分区视图。

#### 不从源表继承的属性

不从源表继承的属性包括:

- 不继承目标表类型。目标表始终是常规表。
- 特权和权限
- 模式
- 生成列、标识列、检查约束和引用约束。在源列是生成列或标识列的情况下, 相应的目标列没有显式的缺省值, 这表示它的缺省值为 NULL。
- 索引表空间 (SYSCAT.TABLES 目录视图的 INDEX\_TBSPACE 列)。值设置为 NULL。拆离 (DETACH) 操作产生的表的索引与该表在同一个表空间中。
- 触发器
- 主键
- 统计信息
- 未包括在从源表继承的属性列表中的所有其他属性。

---

## 对分区表添加数据分区

创建分区表后, 可以使用 ALTER TABLE 语句来修改该表。确切地说, 可以使用 ADD PARTITION 子句来对现有分区表添加新数据分区。在下列情况下, 对分区表添加数据分区要比连接数据分区更为合适: 数据是随着时间的推移而添加到数据分区中的; 数据是缓慢移动的而不是从外部源转入的; 您直接将数据插入或装入到分区表中。特定的示例包括每天将数据装入到一月份数据的数据分区中或者持续不断地插入各行。

#### 限制和用法准则

- 不能对非分区表添加数据分区。要了解有关将现有表迁移到分区表的详细信息, 请参阅第 156 页的『将现有表和视图迁移到分区表』。
- 每个新数据分区的值范围是由 STARTING 和 ENDING 子句确定的。
- 必须至少提供 STARTING 和 ENDING 子句两者中的一个。
- 新范围不能与现有数据分区的范围重叠。
- 在第一个现有数据分区前面添加新数据分区时, 必须指定 STARTING 子句。使用 MINVALUE 来使此范围的末端是开放的。
- 同样, 如果要在最后一个现有数据分区后面添加新数据分区, 那么必须指定 ENDING 子句。使用 MAXVALUE 来使此范围的末端是开放的。
- 如果省略了 STARTING 子句, 那么数据库将在上一数据分区结束边界的紧后面建立开始边界。同样, 如果省略 ENDING 子句, 数据库就会在下一数据分区起始边界的紧前面建立结束边界。
- 起始子句和结束子句的语法与 CREATE TABLE 语句中的指定语法相同。
- 如果未对 ADD PARTITION 指定 IN 或 LONG IN 子句, 那么用来存放数据分区的表空间是使用 CREATE TABLE 语句所使用的方法选择的。

- 包在 ALTER TABLE ...ADD PARTITION 操作期间会失效。
- 一旦落实 ALTER TABLE 语句，新添加的数据分区就会变为可用。

如果在 ADD 或 ATTACH 操作中省略了 STARTING 或 ENDING 边界，就可以填充范围值间隔。以下示例使用 ADD 操作来填充间隔，在该操作中只指定了起始边界：

```
CREATE TABLE hole (c1 int) PARTITION BY RANGE (c1)
(STARTING FROM 1 ENDING AT 10, STARTING FROM 20 ENDING AT 30);
DB20000I 已成功完成 SQL 命令。

ALTER TABLE hole ADD PARTITION STARTING 15;
DB20000I 已成功完成 SQL 命令。

SELECT SUBSTR(tabname, 1,12) tabname,
SUBSTR(datapartitionname, 1, 12) datapartitionname,
seqno, SUBSTR(lowvalue, 1, 4) lowvalue, SUBSTR(highvalue, 1, 4) highvalue
FROM SYSCAT.DATAPARTITIONS WHERE TABNAME='HOLE' ORDER BY seqno;

TABNAME DATAPARTITIONNAME SEQNO LOWVALUE HIGHVALUE
-----
HOLE PART0 0 1 10
HOLE PART2 1 15 20
HOLE PART1 2 20 30
```

选择了 3 个记录。

要改变分区表以及将新数据分区添加至该表，在命令行中发出带 ADD PARTITION 子句的 ALTER TABLE 命令。

示例 1: 对现有分区表添加数据分区，该分区存放 901 到 1000 之间的值（包含端值）。假定 sales 表存储了 9 个范围（0-100 和 101-200 等等），最大值为 900。本示例在表的末尾添加一个附加范围，此范围由未包括的 STARTING 子句指示：

```
ALTER TABLE sales ADD PARTITION dp10
(ENDING AT 1000 INCLUSIVE)
```

## 删除数据分区

可以使用带 DETACH PARTITION 子句的 ALTER TABLE 语句来删除数据分区。但是，可以使用 DROP TABLE 语句来废弃单独的表。

要从分区表拆离数据分区，用户必须拥有下列权限或特权：

- 执行 DETACH 操作的用户必须有权对源表执行 ALTER、SELECT 和 DELETE 操作。
- 该用户还必须有权对目标表执行 CREATE 操作。因此，要改变一个表以拆离数据分区，语句授权标识拥有的特权必须至少包括对目标表的下列其中一项权限或特权：
  - SYSADM 或 DBADM 权限
  - 对数据库的 CREATETAB 权限、对该表所使用表空间的 USE 特权以及下列其中一项权限或特权：
    - 对数据库的 IMPLICIT\_SCHEMA 权限（如果该表的隐式或显式模式名不存在的话）
    - 对模式的 CREATEIN 特权（如果该表的模式名引用现有模式的话）

要废弃表，用户必须拥有下列权限或特权：

- 您必须是 SYSCAT.TABLES 的 DEFINER 列中记录的定义者，或者至少拥有下列其中一项特权：

- SYSADM 或 DBADM 权限
- 对表模式的 DROPIN 特权
- 对表的 CONTROL 特权

**注：** 拆离数据分区情况的含义是，语句授权标识实际上将要发出 CREATE TABLE 语句，所以必须拥有执行该操作所必需的权限。表空间就是正在拆离的数据分区所在的表空间。ALTER TABLE 语句的授权标识将成为新表的定义者（具有 CONTROL 权限），就像是该用户发出了 CREATE TABLE 语句一样。不会将所改变的表的任何特权传递至新表。只有 ALTER TABLE 语句的授权标识以及 DBADM 或 SYSADM 才有权在 ALTER TABLE ...DETACH PARTITION 操作执行后立即访问数据。

要拆离分区表的数据分区，请在命令行中发出带 DETACH PARTITION 子句的 ALTER TABLE 语句。

可以从 DB2 命令行处理器（CLP）废弃表。

要废弃表，请在命令行中发出 DROP TABLE 语句。在以下示例中，将数据分区 dec01 从表 stock 中拆离并放到表 junk 中。然后，可以废弃表 junk，从而有效地删除相关数据分区。

```
ALTER TABLE stock DETACH PART dec01 INTO junk;  
DROP TABLE junk;
```

**注：** 为了尽快地执行 DETACH 操作，对源表执行的索引清除操作是使用后台异步索引清除进程自动完成的。如果有已拆离的从属表，那么已拆离数据分区在拆离时不会变成独立表。在这种情况下，必须发出 SET INTEGRITY 语句以完成拆离并使该表可访问。

---

## 方案：旋转分区表中的数据

旋转 DB2 中的数据指的是一种复用数据分区中的空间的方法，它是通过除去表中的旧数据然后添加新数据来实现的。表分区功能允许先拆离包含旧数据的数据分区，然后连接包含最新数据的新数据分区。

要从分区表拆离数据分区，用户必须拥有下列权限或特权：

- 执行 DETACH 操作的用户必须有权对源表执行 ALTER、SELECT 和 DELETE 操作。
- 该用户还必须有权对目标表执行 CREATE 操作。因此，要改变一个表以拆离数据分区，语句授权标识拥有的特权必须至少包括对目标表的下列其中一项权限或特权：
  - SYSADM 或 DBADM 权限
  - 对数据库的 CREATETAB 权限、对该表所使用表空间的 USE 特权以及下列其中一项权限或特权：
    - 对数据库的 IMPLICIT\_SCHEMA 权限（如果该表的隐式或显式模式名不存在的话）
    - 对模式的 CREATEIN 特权（如果该表的模式名引用现有模式的话）

要改变表以连接数据分区，用户必须拥有下列权限或特权：

- 执行连接的用户必须有权对目标表执行 ALTER 和 INSERT 操作

- 用户还必须能够对源表执行 `SELECT` 和 `DROP` 操作。因此，要改变表以连接数据分区，语句的授权标识所具有的特权必须至少包括对源表的下列其中一种权限或特权：
  - 对源表的 `SELECT` 特权以及对源表模式的 `DROPIN` 特权
  - 对源表的 `CONTROL` 特权
  - `SYSADM` 或 `DBADM` 权限。

要旋转分区表中的数据，请在命令行中发出 `ALTER TABLE` 命令。以下示例演示如何通过除去 2001 年 12 月的数据并将其替换为 2003 年 12 月的最新数据来更新 `stock` 表。

1. 从 `stock` 表中除去旧数据。

```
ALTER TABLE stock DETACH PARTITION dec01 INTO newtable;
```

2. 装入新数据。使用指定了 `REPLACE` 选项的 `LOAD` 来覆盖现有数据。

```
LOAD FROM data_file OF DEL REPLACE INTO newtable
```

**注意：** 如果有已拆离的从属表，那么必须先对已拆离的从属表运行 `SET INTEGRITY` 语句，然后才能装入已拆离表。

3. 如果需要的话，可以执行数据清理操作。数据清理活动包括：

- 填充缺少的值
- 删除不一致的和不完整的数据
- 除去来自多个来源的冗余数据
- 变换数据
  - 规范化（对于来自不同来源并且以不同方式表示同一个值的数据来说，在将这些数据转入仓库时必须对其进行协调。）
  - 聚集（在转入期间，必须对由于太详细而无法存储在仓库中的原始数据进行预聚集。）

4. 将新数据作为新范围进行连接。

```
ALTER TABLE stock ATTACH PARTITION dec03  
STARTING '12/01/2003' ENDING '12/31/2003'  
FROM newtable;
```

如果连接数据分区，就会使查询为空，并且会使程序包失效。

5. 使用 `SET INTEGRITY` 语句来更新索引和其他从属对象。在执行 `SET INTEGRITY` 语句期间，允许进行读写访问。

```
SET INTEGRITY FOR stock ALLOW WRITE ACCESS  
IMMEDIATE CHECKED FOR EXCEPTION IN stock USE stock_ex;
```

---

## 方案：转入和转出分区表数据

这些方案描述的是数据仓库中的一个常见管理操作，在每月初转入新数据，而根据一个特定日期潜在地转出旧数据。

方案 1 通过从表中除去旧数据来说明 `DETACH` 操作（转出）。操作变体包括删除数据以及将数据移入另一个表。这些方案通过将新数据装入到表中来说明 `ADD` 操作和 `ATTACH` 操作（转入）。包括以下操作：

1. 变换数据、将数据装入到非分区表中，然后连接数据分区（传统抽取、变换和装入（ETL））。
2. 将数据装入到非分区表中并变换数据
3. 连接数据分区。

**方案 1: 通过使用分区表, 转出操作在适当的数据分区上就只是一个 DETACH 操作:**

```
ALTER TABLE stock DETACH PART dec01 INTO stock_drop;
COMMIT WORK
```

为了加快 DETACH 操作的执行速度, 对源表执行的索引清除操作是通过后台异步索引清除进程自动完成的。如果源表上没有定义已拆离的从属表, 那么不需要发出 SET INTEGRITY 语句就可以完成 DETACH 操作。

无需像上面的方案中描述的那样废弃表, 同样可以将表连接至另一个表, 或者截断表, 并在重新连接它之前将它用作要装入新数据的表。即使异步索引清除尚未完成也可以立即执行这些操作, 但在 stock 表有已拆离的从属表时除外。

要检测已拆离表是否仍不可访问, 请查询 SYSCAT.TABDETACHEDDEP 目录视图。如果检测到任何不可访问的已拆离表, 那么对所有已拆离的从属表运行带有 IMMEDIATE CHECKED 选项的 SET INTEGRITY 语句, 以将已拆离表转换为可访问的常规表。如果在维护已拆离表的所有已拆离从属表之前试图访问该已拆离表, 那么将返回错误代码 SQL20285N。

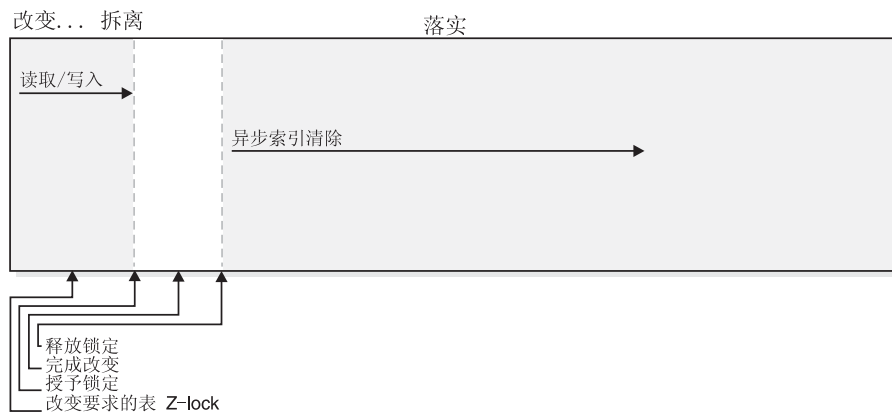


图 37. 此图演示 DETACH 操作期间数据可用性的各个阶段。如果没有拆离从属项, 异步索引清除在落实了 DETACH 操作后立即开始。否则, 异步索引清除在落实了维护拆离从属项之后开始。

**方案 2: 创建一个新的空范围**

以下方案说明了将数据装入到非分区表中, 然后将该数据分区添加到其余表的步骤。

```
ALTER TABLE stock ADD PARTITION dec03
STARTING FROM '12/01/2003' ENDING AT '12/31/2003';
```

此 ALTER TABLE ...ADD 操作使对 stock 表运行的查询为空并使程序包失效。也就是说, 现有查询通常在 ADD 操作继续之前完成。在发出 ADD 操作之后, 访问 stock 表的任何新查询将在锁定处阻塞。

将数据装入到表中:

```
LOAD FROM data_file OF DEL INSERT  
INTO stock ALLOW READ ACCESS;
```

使用 SET INTEGRITY 语句来验证约束并刷新从属具体化查询表 (MQT):

```
SET INTEGRITY FOR stock ALLOW READ  
ACCESS IMMEDIATE CHECKED FOR EXCEPTION IN stock USE stock_ex;  
COMMIT WORK;
```

**提示:** 使用 ALTER TABLE ...ADD PARTITION 后接着使用 LOAD 操作与使用 LOAD 操作后接着使用 ALTER TABLE ...ATTACH 相比有一个优点: 在表未定义约束或 MQT 时, 不需要 SET INTEGRITY 语句来使新数据可用。添加新数据分区或将数据直接装入到表中有一些缺点。使用 ALTER TABLE ...ADD PARTITION 语句的主要缺点是在执行 Load 操作和后续任何 SET INTEGRITY 语句的过程中无法更新表。虽然 ALTER TABLE ...ADD PARTITION 语句和 ALTER TABLE ...ATTACH PARTITION 语句都能使程序包失效, 但 LOAD 命令后跟 ALTER ...ATTACH PARTITION 操作使数据可用性更好。然而, 在数据不是转入较大的块中, 而是慢慢进入的情况下, ALTER TABLE ...ADD PARTITION 语句后跟 IMPORT 命令或常规 INSERT 语句更有意义。在转入的数据与数据分区边界不匹配的情况下, 添加数据分区也有意义。

### 方案 3: 连接数据分区

在此方案中, 使用 ATTACH 帮助将新范围的数据装入到现有分区表中。通常将数据装入到一个新的空表中, 以对数据执行任何必需的清理和检查, 而不会影响目标表。在准备数据之后, 连接新装入的数据分区。

```
CREATE TABLE dec03{.....};  
LOAD FROM data_file OF DEL REPLACE INTO dec03;
```

在转入表数据之前, 可能需要在连接数据之前进行数据清理。数据清理活动包括:

- 填充缺少的值
- 删除不一致的和不完整的数据
- 除去来自多个来源的冗余数据
- 变换数据
  - 规范化 (对于来自不同来源并且以不同方式表示相同值的数据来说, 在将这些数据转入仓库时必须对其进行协调。)
  - 聚集 (在转入期间, 必须对由于太详细而无法存储在仓库中的原始数据进行预聚集。)

接着, 转入数据:

```
ALTER TABLE stock ATTACH PARTITION dec03  
STARTING FROM '12/01/2003' ENDING AT '12/31/2003'  
FROM dec03;
```

在 ATTACH 操作期间, 必须提供 STARTING 和/或 ENDING 子句, 并且下边界 (STARTING) 必须小于或等于上边界 (ENDING)。此外, 新连接的数据分区不能与目标表中的现有数据分区范围重叠。如果已将最高范围定义为 MAXVALUE, 那么任何尝试连接新高范围的操作都将失败, 因为它与现有高范围重叠。此限制也适用于 MINVALUE。除非新数据分区落在现有范围内, 否则不能中途添加或连接新数据分区。创建表时将确定用户未指定的边界。



ALTER TABLE ...ATTACH 操作执行完所有的查询并使从属于 stock 表的程序包失效。也就是说，现在查询通常在 ATTACH 操作继续之前完成。在发出 ATTACH 操作之后，访问 stock 表的任何新查询将在锁定处阻塞。在此转换过程中，stock 表是 z 锁定的（完全不可访问）。已连接的数据分区中的数据仍不可视，因为尚未经 SET INTEGRITY 语句验证。提示：在 ATTACH 操作之后立即发出 COMMIT WORK 语句以使表可供使用。

**COMMIT WORK;**

SET INTEGRITY 语句是验证新连接的数据是否在范围内所必需的。它还对索引和其他从属对象（例如，MQT）进行任何必需的维护。在落实 SET INTEGRITY 语句之前，新数据不可视。如果使用联机 SET INTEGRITY，那么可以对 stock 表中的现有数据进行完全读写访问。当 SET INTEGRITY 正在运行时，缺省方式是 ALLOW NO ACCESS 方式。

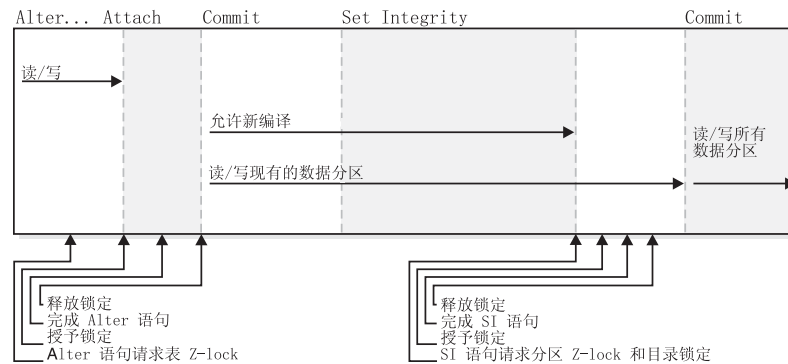


图 38. 此图演示 ATTACH 操作期间数据可用性的各个阶段。

注：当 SET INTEGRITY 正在运行时，不能对表执行 DDL 或实用程序类型操作。操作包括但不限于 LOAD、REORG、REDISTRIBUTE、ALTER TABLE（例如，使用 ALTER 以“not logged initially”的方式执行的添加列以及 ADD、ATTACH、DETACH 和 TRUNCATE 等操作）和 INDEX CREATE。

```
SET INTEGRITY FOR stock ALLOW WRITE ACCESS
IMMEDIATE CHECKED FOR EXCEPTION IN stock USE stock_ex;
```

设置完整性验证新连接的数据分区中的数据。

接着，落实事务以使表可供使用。

**COMMIT WORK;**

任何超出范围或违反其他约束的行都被移至异常表 stock\_ex。可以查询 stock\_ex 以检查违例的行，并可以清除这些行，然后再将它们重新插入到表中。



## 第 14 章 装入

### 并行性和装入

load 实用程序利用使用多个处理器或多个存储设备的硬件配置，如对称多处理器（SMP）环境。

通过使用 load 实用程序，有多种方法可用来并行处理大量数据。一种方法是通过使用多个存储设备，这允许在装入操作期间利用 I/O 并行性（请参阅图 39）。另一种方法涉及在 SMP 环境中使用多个处理器，这允许利用分区内并行性（请参阅图 40）。两种方法可一起使用以提高数据装入速度。

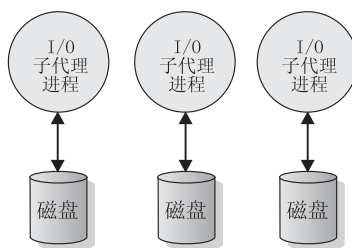


图 39. 在装入数据时利用 I/O 并行性

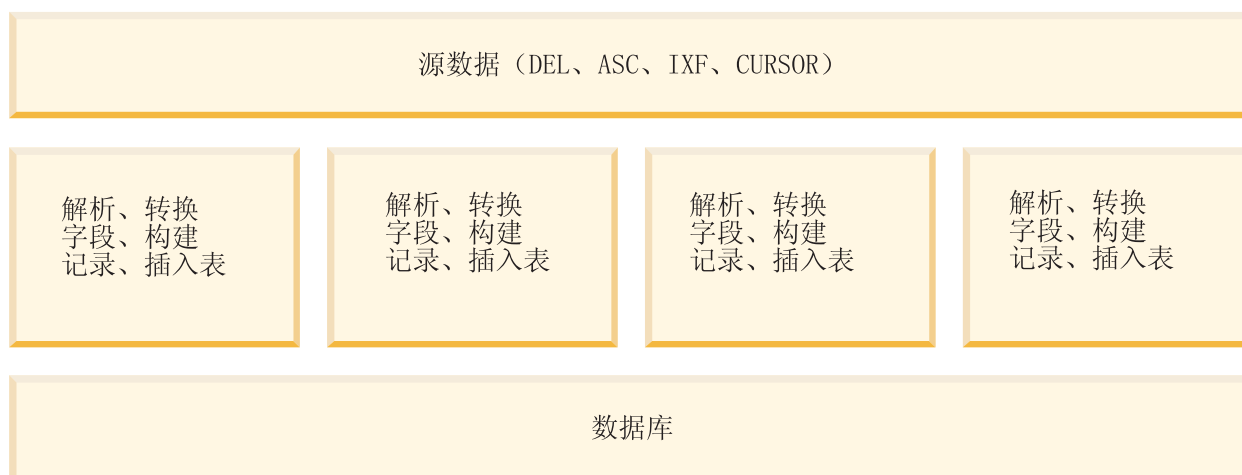


图 40. 在装入数据时利用分区内并行性

---

## 的多维集群注意事项

在将数据装入多维集群（MDC）表时，下列限制适用：

- 不支持 LOAD 命令的 SAVECOUNT 选项。
- 由于这些表管理它们自己的可用空间，所以不支持 totalfreespace 文件类型修饰符。
- MDC 表需要 anyorder 文件类型修饰符。如果对 MDC 表执行装入，但未指定 anyorder 修饰符，那么实用程序将显式启用该修饰符。

对 MDC 表使用 LOAD 命令时，将按以下方式处理唯一约束违例：

- 如果执行装入操作前该表包含唯一键，并且将重复记录装入该表，那么将保留原始记录，并且在删除阶段删除新记录。
- 如果执行装入操作前该表未包含唯一键，并且将唯一键和重复记录都装入该表，那么将只装入其中一个带有唯一键的记录，并且在删除阶段删除其他记录。

注：没有确切的技术可用来确定将要装入的记录以及将要删除的记录。

### 性能注意事项

为了提高 load 实用程序在装入 MDC 表时的性能，应该增大 *util\_heap\_sz* 数据库配置参数值。当有更多内存可供该实用程序使用时，MDC 装入算法的性能会显著提高。这将减少在装入阶段执行数据集群时的磁盘 I/O 次数。当指定了 LOAD 命令的 DATA BUFFER 选项时，还应增大它的值。如果使用 LOAD 命令来并发地装入若干个 MDC 表，那么应相应地增大 *util\_heap\_sz*。

由于所有 MDC 表都有块索引，所以 MDC 装入操作始终包括构建阶段。

在装入阶段，将执行附加的记录以维护块映射。对于分配的每个扩展数据块，大约有两个附加的日志记录。为了确保性能良好，在设置 *logbufsz* 数据库配置参数地值时应该考虑此情况。

在将数据装入 MDC 表时，将使用一个带有索引的系统临时表。该表的大小与装入的单个单元数成正比。该表中每一行的大小与 MDC 维键的大小成正比。为了最大程度地减少装入操作期间处理此表时执行的磁盘 I/O 次数，请确保临时表空间的缓冲池足够大。

---

## 分区表的装入注意事项

对目标表进行分区时，将支持所有现有装入功能，但存在以下常规限制：

- 当分区代理程序数大于 1 时，不支持一致点。
- 不支持将数据装入到数据分区子集中的同时保持其余数据分区完全联机。
- 装入操作使用的异常表不能分区。
- 当 load 实用程序以插入方式或重新启动运行并且装入目标表具有任何已拆离的从属时，那么不能重建唯一索引。
- 与装入 MDC 表相同，装入分区表时将不会保留输入数据记录的精确排序。只有在单元或数据分区中才保留排序。
- 在每个数据库分区上利用多个格式化程序的装入操作仅保留输入记录的大致排序。在每个数据库分区上运行单个格式化程序，将输入记录按单元或表分区键进行分组。要在每个数据库分区上运行单个格式化程序，应显式请求 CPU\_PARALLELISM 为 1。

## 一般装入行为

load 实用程序将数据记录插入到正确的数据分区中。在装入之前，不需要使用外部实用程序（如分割程序）来对输入数据进行分区。

load 实用程序不访问任何拆离的或相连的数据分区。数据仅插入到可视数据分区中。可视数据分区不会拆离，也不会相连。此外，装入替换操作不会截断拆离或相连的数据分区。因为 load 实用程序获取针对目录系统表的锁定，所以它将等待任何未落实的 ALTER TABLE 事务。这些事务将获取针对目录表中的相关行的互斥锁定，并且互斥锁定必须先终止，装入操作才能继续。这意味着装入操作运行期间，可能没有未落实的 ALTER TABLE ...ATTACH、DETACH 或 ADD PARTITION 事务。将拒绝目标为拆离或相连的数据分区的所有输入源记录，但可从异常表中检索它们（如果指定了异常表）。会有一条参考消息写入消息文件，以指示某些目标表数据分区处于相连或拆离状态。针对对应于目标表的相关目录表行的锁定使得用户无法通过在 load 实用程序运行时发出 ALTER TABLE ...ATTACH、DETACH 或 ADD PARTITION 操作来更改目标表的分区。

## 处理无效行

当 load 实用程序遇到的记录不属于任何可视数据分区时，将拒绝该记录并且 load 实用程序继续进行处理。因为范围限制违例而拒绝的记录个数不会显式显示出来，但会包括在拒绝的记录的总行数中。因为范围违例而拒绝记录不会增加行警告的数目。会有一条消息（SQL0327N）写入 load 实用程序消息文件，指示发现范围违例，不会对每一个记录来记录消息。除了目标表中的所有列之外，异常表还包括用于描述特定行发生的类型违例的列。包含无效数据的行（包括未分区的数据）将写至转储文件。

因为异常表插入成本很高，所以可以控制插入到异常表中的约束违例。例如，load 实用程序的缺省行为是将本来有效但因为范围限制或唯一约束违例而拒绝的行插入到异常表中。通过对 FOR EXCEPTION 子句分别指定 NORANGEEXC 或 NOUNIQUEEXC 可以关闭此行为。如果指定不应将这些约束违例插入到异常表中，或者未指定异常表，那么有关违反范围限制或唯一约束的行的信息将会丢失。

## 历史记录文件

如果目标表已分区，那么相应的历史记录文件条目不会包括目标表跨越的表空间列表。另一操作详细程度标识（“R”而不是“T”）指示对分区表运行了装入操作。

## 终止装入操作

终止装入替换操作将完全截断所有可视数据分区，而终止装入插入操作会将所有可视数据分区截断至装入前的长度。如果联机装入操作在装入复制阶段失败，那么在终止该操作期间，索引会变得无效。在终止接触索引的脱机装入操作时，索引也会变得无效（因为重建索引方式或者增量维护期间插入了键而使得索引处于不一致状态，从而导致索引无效）。将数据装入到多个目标中不会影响装入恢复操作，但将无法从装入阶段期间获取的一致点重新启动装入操作。在此情况下，如果对目标表进行分区，那么将忽略 SAVECOUNT 装入选项。此行为与将数据装入到 MDC 目标表中的行为一致。

## 生成列

如果生成列在任何分区、维或分布键中，那么会忽略 generatedoverride 文件类型修饰符并且 load 实用程序会生成值，就像指定了 generatedignore 文件类型修饰符一样。在此情况下，装入错误的生成列值可能导致将记录放置在错误的物理位置上，例如，错误的数据分区、MDC 块或数据库分区。例如，一旦记录放在错误的数据分区上，设置完整性就必须将其移至另一物理位置，这不能在联机设置完整性操作期间完成。

### 数据可访问性

当前联机装入算法扩展至分区表。在 `LOAD` 命令上指定的联机装入 (`ALLOW READ ACCESS`) 允许阅读器并行访问整个表，包括装入和非装入数据分区。

### 数据分区状态

在某些情况下，成功装入后可视数据分区可能切换至“设置完整性暂挂”状态和/或“只读访问”表状态。如果该表存在装入操作不能保留的约束，那么数据分区可能会置于这些状态。这种约束可能包括检查约束和拆离的具体化查询表。失败的装入操作会导致所有可视数据分区处于“装入暂挂”表状态。

### 错误隔离

不支持在数据分区级别进行错误隔离。隔离错误意味着在运行时未出现错误的数据库分区上继续装入，而在运行时出现错误的数据库分区上停止装入。可在不同数据库分区间隔离错误，但 `load` 实用程序不能在一个可视数据分区子集上落实事务，而回滚其余可视数据分区。

### 其他注意事项

- 如果有任何索引标记为无效，那么不支持递增。如果索引需要重建或拆离的从属项需要使用 `SET INTEGRITY` 语句进行验证，那么认为索引无效。
- 同时支持装入到分区表中，这些表使用按范围分区、按散列分布或按维算法组织的任何组合进行分区。
- 对于包括受装入影响的对象和表空间标识列表的日志记录，这些日志记录的大小 (`LOAD START` 和 `COMMIT (PENDING LIST)`) 可能增长得很快，并且因此而降低可供其他应用程序使用的活动日志空间量。
- 当表同时进行了分区和分布时，分区数据库装入可能不会影响所有数据库分区。只有输出数据库分区上的对象才会更改。
- 在装入操作期间，分区表的内存消耗会随表数的增加而增加。注意，总增加量不是线性的，因为仅总内存需求的一小部分与数据分区数成正比。

---

## 第 15 章 在分区数据库环境中装入数据

---

### 装入概述 - 分区数据库环境

在多分区数据库环境中，大量的数据放在多个数据库分区中。分布键用来确定每部分数据所在的数据库分区。必须先分布数据，然后才能将该数据装入到正确的数据库分区中。

在多分区数据库中装入表时，load 实用程序可以：

- 并行地分布输入数据
- 同时各个相应数据库分区中装入数据
- 将数据从一个系统传输到另一个系统

将数据装入到多分区数据库中分两阶段完成：第一阶段为设置阶段，在此阶段获取数据库分区资源（如表锁定）；第二阶段为装入阶段，在此阶段将数据装入到数据库分区中。可以使用 LOAD 命令的 ISOLATE\_PART\_ERRS 选项来选择这些阶段的错误处理方式，并可以选择一个或多个数据库分区上的错误对未发生错误的数据库分区上的装入操作的影响。

在将数据装入多分区数据库时，可以使用下列其中一种方式：

#### **PARTITION\_AND\_LOAD**

对数据进行分布（有可能以并行方式进行分布），并且同时各个相应数据库分区上装入数据。

#### **PARTITION\_ONLY**

对数据进行分布（有可能以并行方式进行分布），并将输出写入每个装入数据库分区上指定位置中的文件。每个文件都包含分区头，该分区头指定数据在数据库分区上的分布方式，并指定可以使用 LOAD\_ONLY 方式将该文件装入到数据库中。

#### **LOAD\_ONLY**

假定数据已分布在数据库分区上；将跳过分布过程，并且在相应的数据库分区上同时装入数据。

#### **LOAD\_ONLY\_VERIFY\_PART**

假定数据已分布在数据库分区上，但数据文件未包含分区头。将跳过分布过程，并且在相应的数据库分区上同时装入数据。在装入操作期间，将检查每一行以验证它是否在正确的数据库分区中。如果指定了 dumpfile 文件类型修饰符，那么会将发生数据库分区违例的行放到转储文件中。否则会删除那些行。如果特定装入数据库分区上存在数据库分区违例，那么会将一条有关该数据库分区的警告写至装入消息文件。

#### **ANALYZE**

生成最佳分布图（在所有数据库分区之间均匀地分布数据）。

#### 概念和术语

在讨论 load 实用程序在带有多个数据库分区的分区数据库环境中的行为和操作时，将使用以下术语：

- 协调程序分区是一个数据库分区，用户连接到该分区以执行装入操作。在 PARTITION\_AND\_LOAD、PARTITION\_ONLY 和 ANALYZE 方式下，除非指定了 LOAD 命令的 CLIENT 选项，否则假定数据文件在此数据库分区上。如果指定 CLIENT，那么表示要装入的数据在连接的远程客户机上。
- 在 PARTITION\_AND\_LOAD、PARTITION\_ONLY 和 ANALYZE 方式下，预分区代理程序读取用户数据并以循环方式将其分发给分区代理程序，后者将分布该数据。此过程始终是在协调程序分区上执行的。对于任何装入操作，每个数据库分区最多允许一个分区代理程序。
- 在 PARTITION\_AND\_LOAD、LOAD\_ONLY 和 LOAD\_ONLY\_VERIFY\_PART 方式下，在每个输出数据库分区上都运行装入代理程序，它协调该数据库分区上的数据装入操作。
- 在 PARTITION\_ONLY 装入操作期间，在每个输出数据库分区上运行“装入到文件”代理程序。它们从分区代理程序接收数据并将该数据写入所在数据库分区上的文件中。
- SOURCEUSEREXIT 选项提供了一种工具，load 实用程序可通过该工具执行定制脚本或可执行文件（此处称为用户出口）。

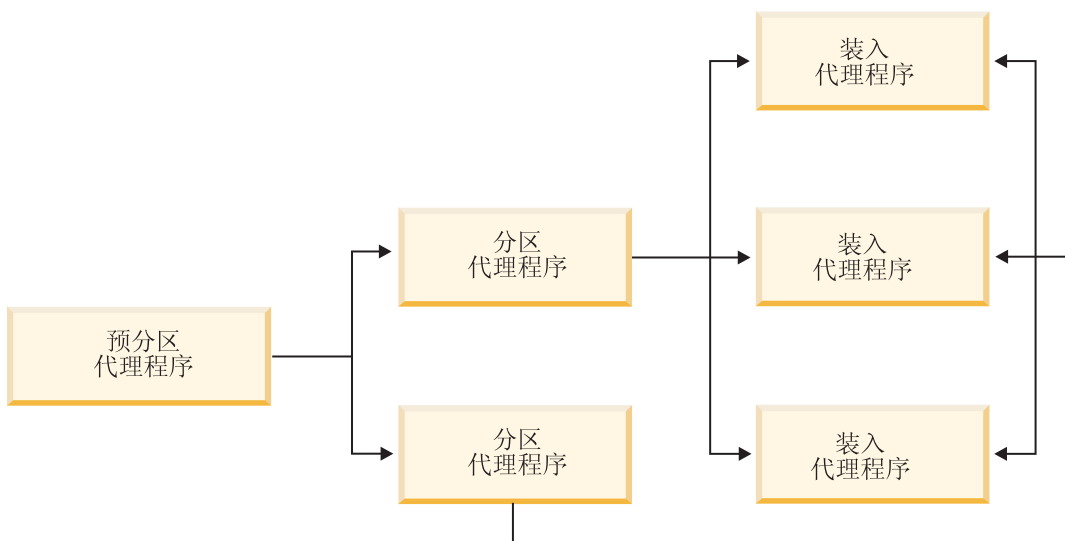


图 41. 分区数据库装入概述。预分区代理程序读取源数据，然后向两个分区代理程序中的每个分区代理程序各发送接近一半的数据，这两个分区代理程序分布数据并将该数据发送给三个数据库分区中的一个数据库分区。每个数据库分区上的装入代理程序装入数据。

## 在分区数据库环境中装入数据 - 提示与技巧

以下是在多分区数据库中装入表前要考虑的一些信息：

- 对少量数据使用 load 实用程序，熟悉装入配置选项。



- 如果输入数据已排序或者具有某种选择的顺序，并且要在装入过程中维护该顺序，那么只应该将一个数据库分区用于分布。并行分布无法保证按照数据的接收顺序来装入该数据。缺省情况下，如果在 LOAD 命令中未指定 anyorder 修饰符，那么 load 实用程序就会选择单分区代理程序。
- 如果正在从不同的文件装入大对象（LOB）（即，如果使用 load 实用程序时指定了 lobsinfile 修饰符），那么所有执行装入的数据库分区都必须能够对所有包含 LOB 文件的目录进行读访问。处理 LOB 时，LOAD lob-path 参数必须是标准路径。
- 通过将 ISOLATE\_PART\_ERRS 选项设置为 SETUP\_ERRS\_ONLY 或 SETUP\_AND\_LOAD\_ERRS，可以强制在多分区数据库中运行的作业继续运行，即使装入操作在启动时检测到某些装入数据库分区或相关表空间或表处于脱机状态亦如此。
- 使用 STATUS\_INTERVAL 装入配置选项来监视在多分区数据库中运行的作业的进度。装入操作按指定的时间间隔生成消息，以指示预分区代理程序已读取的数据的兆字节数。这些消息将被转储到预分区代理程序消息文件中。要在装入操作期间查看此文件的内容，请连接到协调程序分区并对目标表发出 LOAD QUERY 命令。
- 如果参与分布过程的数据库分区（由 PARTITIONING\_DBPARTNUMS 选项定义）与装入数据库分区（由 OUTPUT\_DBPARTNUMS 选项定义）不同，就会由于 CPU 周期争用情况减少而提高性能。将数据装入到多分区数据库中时，对未参与分布或装入操作的数据库分区调用 load 实用程序。
- 如果在 LOAD 命令中指定了 MESSAGES 参数，就会将预分区代理程序、分区代理程序和装入代理程序生成的消息文件保存下来，以供装入操作完成后参考。要在装入操作期间查看这些文件的内容，请连接到期望的数据库分区并对目标表发出 LOAD QUERY 命令。
- load 实用程序仅选择一个输出数据库分区以便在该分区中收集统计信息。可以使用 RUN\_STAT\_DBPARTNUM 数据库配置选项来指定该数据库分区。
- 在多分区数据库中装入数据之前，请运行设计顾问程序以确定每个表的最佳分区。有关更多信息，请参阅《调整数据库性能》中的『设计顾问程序』。

### 故障诊断

如果 load 实用程序挂起，您可以：

- 使用 STATUS\_INTERVAL 参数来监视多分区数据库装入操作的进度。将把状态时间间隔信息转储到协调程序分区上的预分区代理程序消息文件中。
- 检查分区代理程序消息文件，了解每个数据库分区上的分区代理进程状态。如果执行装入操作时未出错，并且设置了 TRACE 选项，那么这些消息文件应该会包含许多记录的跟踪消息。
- 检查装入消息文件以了解是否有任何装入错误消息。

**注：**必须指定 LOAD 命令的 MESSAGES 选项，这样这些文件才能存在。

- 如果有错误指示某个装入进程出错，请中断当前装入操作。

---

## 在分区数据库环境中装入数据

使用 load 实用程序将数据装入到分区数据库环境中。

在多分区数据库中装入表之前：

1. 确保正确设置了 *svcename* 数据库管理器配置参数和 **DB2COMM** 概要文件注册表变量。由于 load 实用程序使用 TCP/IP 来将数据从预分区代理程序传输至分区代理程序以及从分区代理程序传输至装入数据库分区，所以正确进行上述设置非常重要。
2. 在调用 load 实用程序前，您必须连接至（或者能够隐式地连接至）要装入数据的数据库。由于 load 实用程序将发出 COMMIT 语句，所以，在开始执行装入操作前，应该通过发出 COMMIT 或 ROLLBACK 语句完成所有事务并释放任何锁定。如果使用的是 PARTITION\_AND\_LOAD、PARTITION\_ONLY 或 ANALYZE 方式，那么装入的数据文件必须在此数据库分区上，但在下列情况下除外：
  - a. 指定了 CLIENT 选项，在这种情况下，数据必须在客户机上；
  - b. 输入源类型为 CURSOR，在这种情况下，没有输入文件。
3. 运行设计顾问程序以确定每个表的最佳数据库分区。有关更多信息，请参阅《调整数据库性能》中的『设计顾问程序』。

在使用 load 实用程序以在多分区数据库环境中装入数据时，下列限制适用：

- 装入操作的输入文件位置不能是磁带设备。
- 除非使用 ANALYZE 方式，否则不支持 ROWCOUNT 选项。
- 如果目标表带有进行分布所需的标识列，并且未指定 identityoverride 文件类型修饰符，或者如果正在使用多个数据库分区来分布数据并接着装入该数据，那么不支持在 LOAD 命令上使用大于 0 的 SAVECOUNT。
- 如果分布键包含标识列，那么只支持 PARTITION\_AND\_LOAD 方式。
- LOAD\_ONLY 和 LOAD\_ONLY\_VERIFY\_PART 方式不能与 LOAD 命令的 CLIENT 选项配合使用。
- LOAD\_ONLY\_VERIFY\_PART 方式不能与 CURSOR 输入源类型配合使用。
- 分布错误隔离方式 LOAD\_ERRS\_ONLY 和 SETUP\_AND\_LOAD\_ERRS 不能与 LOAD 命令的 ALLOW READ ACCESS 和 COPY YES 选项配合使用。
- 如果 OUTPUT\_DBPARTNUMS 和 PARTITIONING\_DBPARTNUMS 选项指定的数据库分区不重叠，那么多个装入操作可以同时将数据装入到同一个表中。例如，如果表是在数据库分区 0 至 3 上定义的，那么一个装入操作可以将数据装入到数据库分区 0 和 1 中，而另一个装入操作可以将数据装入到数据库分区 2 和 3 中。
- 对于跨多个数据库分区的表来说，只能将非定界 ASCII (ASC) 和定界 ASCII (DEL) 文件分布到这些表中。不能对 PC/IXF 文件进行分布。要将 PC/IXF 文件装入到跨多个数据库分区的表中，可以首先设置环境变量 **DB2\_PARTITIONEDLOAD\_DEFAULT=NO** 并使用 LOAD\_ONLY\_VERIFY\_PART 方式，以将该表装入到位于单个数据库分区上的表中。然后，可以使用 CURSOR 文件类型来执行装入操作，以将该数据移到分布在多个数据库分区中的表中。也可以使用 LOAD\_ONLY\_VERIFY\_PART 方式的装入操作将 PC/IXF 文件装入到分布在多个数据库分区中的表中。

下列示例说明如何使用 LOAD 命令来启动各种类型的装入操作。下列示例中使用的数据库有 5 个数据库分区：0、1、2、3 和 4。每个数据库分区都有一个本地目录 /db2/data/。在数据库分区 0、1、3 和 4 上定义了两个表 TABLE1 和 TABLE2。从客户机装入数据时，用户能够访问除数据库分区以外的远程客户机。

## 从服务器分区中装入

### 分布和装入示例

在此示例中，您连接到一个数据库分区，该数据库分区可能是也可能不是用来定义 TABLE1 的数据库分区。数据文件 load.del 在此数据库分区的当前工作目录中。要将 load.del 中的数据装入到所有定义了 TABLE1 的数据库分区中，请发出以下命令：

```
LOAD FROM LOAD.DEL of DEL REPLACE INTO TABLE1
```

注：在此示例中，将对分区数据库环境的所有配置参数使用缺省值：MODE 参数将缺省为 PARTITION\_AND\_LOAD，OUTPUT\_DBPARTNUMS 选项将缺省为所有定义 TABLE1 的数据库分区，而 PARTITIONING\_DBPARTNUMS 将缺省为在未指定数据库分区时根据 LOAD 命令规则选择的一组数据库分区。

要在数据分布在数据库分区 3 和数据库分区 4 上的位置执行装入操作，请发出以下命令：

```
LOAD FROM LOAD.DEL of DEL REPLACE INTO TABLE1
PARTITIONED DB CONFIG PARTITIONING_DBPARTNUMS (3,4)
```

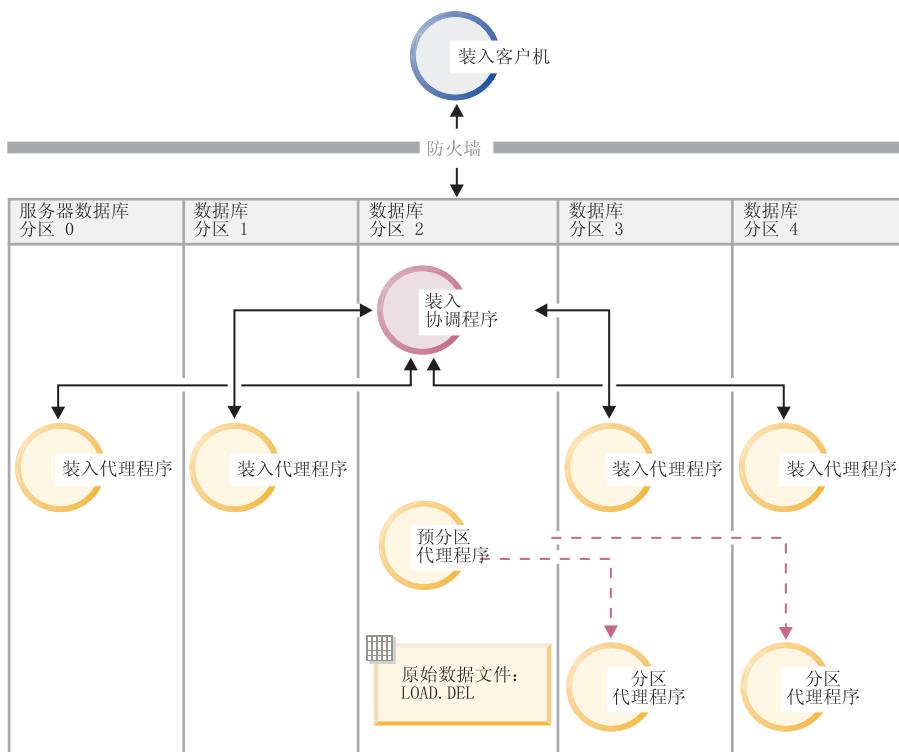


图 42. . 此图说明发出以上命令后产生的行为。数据将装入到数据库分区 3 和 4 中。

### 仅分布示例

在此示例中，您连接到一个数据库分区，该数据库分区可能是也可能不是用来定义 TABLE1 的数据库分区。数据文件 load.del 在此数据库分区的当前工作目录中。在使用数据库分区 3 和数据库分区 4 的情况下，要将 load.del 分布（而不装入）到所有定义 TABLE1 的数据库分区中，请发出以下命令：

```
LOAD FROM LOAD.DEL of DEL REPLACE INTO TABLE1
PARTITIONED DB CONFIG MODE PARTITION ONLY
PART_FILE_LOCATION /db2/data
PARTITIONING_DBPARTNUMS (3,4)
```

这将导致把文件 load.del.xxx 存储在每个数据库分区上的 /db2/data 目录中，其中 xxx 是 3 位的数据库分区号。

在仅使用数据库分区 0（缺省 PARTITIONING\_DBPARTNUMS 值）上运行的 1 个分区代理程序的情况下，要将 load.del 文件分布到数据库分区 1 和 3，请发出以下命令：

```
LOAD FROM LOAD.DEL OF DEL REPLACE INTO TABLE1
      PARTITIONED DB CONFIG MODE PARTITION_ONLY
      PART_FILE_LOCATION /db2/data
      OUTPUT_DBPARTNUMS (1,3)
```

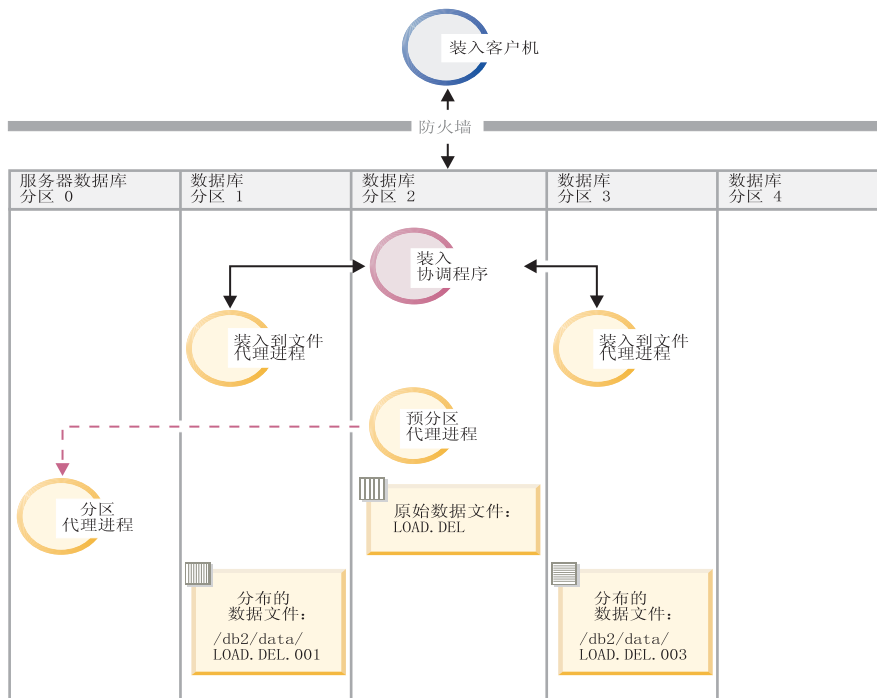


图 43. . 此图说明发出以上命令后产生的行为。将使用数据库分区 0 上运行的 1 个分区代理程序将数据装入到数据库分区 1 和 3 中。

### 仅装入示例

如果已经以 PARTITION\_ONLY 方式执行了装入操作，并且要将每个装入数据库分区的 /db2/data 目录中的分区文件装入到所有定义了 TABLE1 的数据库分区中，请发出以下命令：

```
LOAD FROM LOAD.DEL OF DEL REPLACE INTO TABLE1
      PARTITIONED DB CONFIG MODE LOAD_ONLY
      PART_FILE_LOCATION /db2/data
```

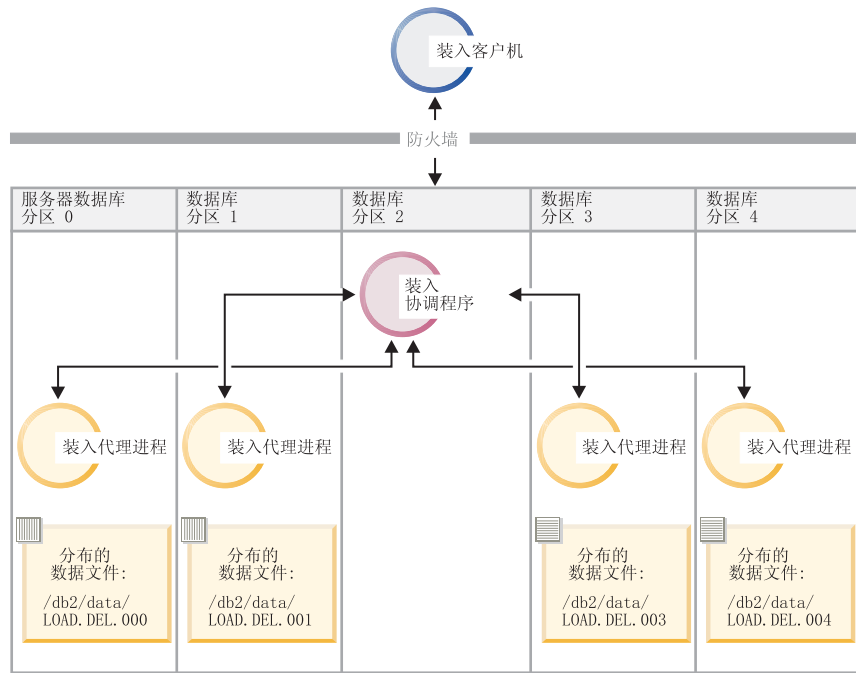


图 44. 此图说明发出以上命令后产生的行为。将把分布式数据装入到所有用来定义 TABLE1 的数据库分区中。

要仅装入到数据库分区 4 中，请发出以下命令：

```
LOAD FROM LOAD.DEL OF DEL REPLACE INTO TABLE1
PARTITIONED DB CONFIG MODE LOAD_ONLY
PART_FILE_LOCATION /db2/data
OUTPUT_DBPARTNUMS (4)
```

#### 装入不带分布图头的预分布文件

可以使用 LOAD 命令来将不带分布头的装入数据文件直接装入到数个数据库分区中。如果数据文件在每个用来定义 TABLE1 的数据库分区上的 /db2/data 目录中，并且名为 load.del.xxx（其中 xxx 是数据库分区号），那么可以通过发出以下命令来装入那些文件：

```
LOAD FROM LOAD.DEL OF DEL modified by dumpfile=rejected.rows
REPLACE INTO TABLE1
PARTITIONED DB CONFIG MODE LOAD_ONLY_VERIFY_PART
PART_FILE_LOCATION /db2/data
```

要仅将数据装入到数据库分区 1 中，请发出以下命令：

```
LOAD FROM LOAD.DEL OF DEL modified by dumpfile=rejected.rows
REPLACE INTO TABLE1
PARTITIONED DB CONFIG MODE LOAD_ONLY_VERIFY_PART
PART_FILE_LOCATION /db2/data
OUTPUT_DBPARTNUMS (1)
```

注：如果指定了转储文件，那么将拒绝装入不属于源数据库分区的行并将它们放入转储文件。

#### 从远程客户机装入至多分区数据库

要将远程客户机上的文件中的数据装入到多分区数据库中，必须指定 `LOAD` 命令的 `CLIENT` 选项以指示数据文件不在服务器分区上。例如：

```
LOAD CLIENT FROM LOAD.DEL OF DEL REPLACE INTO TABLE1
```

注：不能将 `LOAD_ONLY` 或 `LOAD_ONLY_VERIFY_PART` 方式与 `CLIENT` 选项配合使用。

### 从游标装入

与在单一分区数据库中一样，可以从游标装入到多分区数据库中。在此示例中，对于 `PARTITION_ONLY` 和 `LOAD_ONLY` 方式，`PART_FILE_LOCATION` 选项必须指定标准文件名。此名称是在每个输出数据库分区上创建或装入的分布文件的标准基本文件名。如果目标表包含 `LOB` 列，那么可以使用指定的基本名称来创建多个文件。

要将语句 `SELECT * FROM TABLE1` 的应答集中的所有行分布至名为 `/db2/data/select.out.xxx`（其中 `xxx` 是数据库分区号）的每个数据库分区上的文件，以便将来装入到 `TABLE2` 中，请发出以下命令：

```
DECLARE C1 CURSOR FOR SELECT * FROM TABLE1

LOAD FROM C1 OF CURSOR REPLACE INTO TABLE2
      PARTITIONED DB CONFIG MODE PARTITION_ONLY
      PART_FILE_LOCATION /db2/data/select.out
```

然后，可以通过发出以下 `LOAD` 命令来装入以上操作生成的数据文件：

```
LOAD FROM C1 OF CURSOR REPLACE INTO TABLE2
      PARTITIONED CB CONFIG MODE LOAD_ONLY
      PART_FILE_LOCATION /db2/data/select.out
```

---

## 使用 `LOAD QUERY` 命令在分区数据库环境中监视装入操作

在分区数据库环境中执行装入操作期间，某些装入进程会在它们执行时所在的数据库分区上创建消息文件。

这些消息文件存储装入执行期间生成的所有参考消息、警告消息和错误消息。用户可以查看以下三个能生成的消息文件的装入进程：装入代理程序、预分区代理程序和分区代理程序。只有在装入操作完成后，消息文件的内容才可用。

在装入操作期间，可连接至各个数据库分区并对目标表发出 `LOAD QUERY` 命令。从 `CLP` 中发出此命令时，此命令将显示由 `LOAD QUERY` 命令指定的表的数据库分区上当前存在的所有消息文件的内容。

例如，在数据库 `WSDB` 中，表 `TABLE1` 是在数据库分区 0 至 3 上定义的。您将连接至数据库分区 0 并发出以下 `LOAD` 命令：

```
load from load.del of del replace into table1 partitioned db config
      partitioning_dbpartnums (1)
```

此命令将启动装入操作，该操作包括：在数据库分区 0、1、2 和 3 上运行的装入代理程序；在数据库分区 1 上运行的分区代理程序；在数据库分区 0 上运行的预分区代理程序。

数据库分区 0 将包含预分区代理程序的消息文件以及该数据库分区上装入代理程序的消息文件。要同时查看这些文件的内容，请启动新会话并从 `CLP` 中发出下列命令：

```

set client connect_node 0
connect to wsdB
load query table table1

```

数据库分区 1 将包含装入代理程序的消息文件和分区代理程序的消息文件。要查看这些文件的内容，请启动新会话并从 CLP 中发出下列命令：

```

set client connect_node 1
connect to wsdB
load query table table1

```

**注：**STATUS\_INTERVAL 装入配置选项生成的消息将出现在预分区代理程序消息文件中。要在装入操作期间查看这些消息，必须连接至协调程序分区并发出 LOAD QUERY 命令。

### 保存消息文件内容

如果通过 db2Load API 启动装入操作，那么必须指定消息选项（piLocalMsgFileName），将消息文件从服务器传输到客户机并存储下来以供查看。

对于从 CLP 启动的多分区数据库装入操作来说，不会在控制台上显示或保留消息文件。要在多分区数据库装入完成后保存或查看这些文件的内容，必须指定 LOAD 命令的 MESSAGES 选项。如果使用了此选项，那么装入操作一旦完成，就会将每个数据库分区上的消息文件传输到客户机并使用 MESSAGES 选项中指示的基本名称存储在文件中。对于多分区数据库装入操作，下面列示了各个装入进程所生成的消息文件名：

| 进程类型    | 文件名   |
|---------|---|
| 装入代理程序  | message-file-name>.LOAD.dbpartition-number> |
| 分区代理程序  | message-file-name>.PART.dbpartition-number> |
| 预分区代理程序 | message-file-name>.PREP.dbpartition-number> |

例如，如果 MESSAGES 选项指定 /wsdb/messages/load，那么数据库分区 2 的装入代理程序消息文件将是 /wsdb/messages/load.LOAD.002。

**注：**强烈建议从 CLP 中启动的多分区数据库装入操作使用 MESSAGES 选项。

---

## 在分区数据库环境中继续、重新启动或终止装入操作

如果装入操作在分区数据库环境中失败，那么接下来需要执行的操作取决于出现故障的时间。

多分区数据库中的装入过程由两个阶段组成：

1. 设置阶段，在该阶段中获取数据库分区级别的资源，例如，输出数据库分区上的表锁定

通常，如果设置阶段发生故障，不必执行重新启动和终止操作。您需要执行的操作取决于对失败的装入操作指定的错误隔离方式。

如果装入操作指定不隔离设置阶段错误，那么整个装入操作将被取消并且每个数据库分区上的表状态都将回滚到该表在装入操作前所处的状态。

如果装入操作指定要隔离设置阶段错误，那么装入操作将在成功完成设置阶段的数据库分区上继续，但每个失败数据库分区上的表都将回滚到该表在装入操作前所处的状态。这意味着如果某些分区在设置阶段失败，而其他分区在装入阶段失败，那么单个装入操作可能在不同阶段失败。

## 2. 装入阶段，在该阶段中格式化数据并将它们装入到数据库分区上的表中。

在分区数据库装入操作的装入阶段，如果装入操作在至少一个数据库分区上失败，那么必须发出 `load RESTART` 或 `load TERMINATE` 命令。因为分区数据库中的数据装入操作将通过单个事务完成，所以有必要执行此操作。

如果您可以解决导致装入操作失败的问题，那么应该选择 `load RESTART`。这样可以节省时间，因为在启动了装入重新启动操作时，装入操作会在所有数据库分区上的中断位置继续执行。

如果您希望表返回到该表在初始装入操作前所处的状态，那么应该选择 `load TERMINATE`。

过程:

### 确定装入失败的时间

如果装入操作在分区环境中失败，那么您需要做的第一件事是确定操作在哪个分区上失败以及每个操作在哪个阶段失败。这可通过查看分区总结来完成。如果从 `CLP` 发出 `LOAD` 命令，那么分区总结将显示在每个 `LOAD` 命令的末尾（请参阅下面的示例）。如果从 `db2Load API` 发出 `LOAD` 命令，那么分区总结包含在 `db2PartLoadOut` 结构的 `poAgentInfoList` 字段中。

如果对于给定分区，“代理程序类型”有一个“LOAD”条目，那么该分区已到达装入阶段，否则在设置阶段出现故障。SQL 代码为负数表示失败。在以下示例中，装入操作在装入阶段在分区 1 上失败。

| 代理程序类型 | 节点  | SQL 代码    | 结果                          |
|--------|-----|-----------|-----------------------------|
| LOAD   | 000 | +00000000 | Success.                    |
| LOAD   | 001 | -00000289 | Error. May require RESTART. |
| LOAD   | 002 | +00000000 | Success.                    |
| LOAD   | 003 | +00000000 | Success.                    |
| .      | .   | .         | .                           |

### 继续、重新启动或终止失败的装入操作

在设置阶段，只有使用值为 `SETUP_ERRS_ONLY` 或 `SETUP_AND_LOAD_ERRS` 的 `ISOLATE_PART_ERRS` 选项的装入应失败。对于在此阶段在至少一个输出数据库分区上失败的装入来说，可以发出 `LOAD REPLACE` 或 `LOAD INSERT` 命令。使用 `OUTPUT_DBPARTNUMS` 选项来仅指定装入操作在其上失败的那些数据库分区。

对于在装入阶段在至少一个输出数据库分区上失败的装入来说，发出 `load RESTART` 或 `load TERMINATE` 命令。



对于在设置阶段在至少一个输出数据库分区上失败并且在装入阶段在至少一个输出数据库分区上失败的装入来说，需要执行两个装入操作以继续失败的装入 - 一个装入操作用于设置阶段故障，另一个用于装入阶段故障，如上所述。要有效地撤销此类失败的装入操作，发出 `load TERMINATE` 命令。但是，在发出该命令后，您必须说明所有分区，因为没有对在设置阶段失败的分区上的表进行任何更改，并且已撤销在装入阶段失败的分区的所有更改。

例如，在数据库 `WSDB` 中，`TABLE1` 是在数据库分区 0 至 3 上定义的。发出以下命令：

```
load from load.del of del insert into table1 partitioned db config
isolate_part_errs setup_and_load_errs
```

在设置阶段，输出数据库分区 1 上出现故障。由于隔离了设置阶段错误，所以装入操作继续，但在装入阶段在分区 3 上出现故障。要继续装入操作，应发出下列命令：

```
load from load.del of del replace into table1 partitioned db config
output_dbpartnums (1)
```

```
load from load.del of del restart into table1 partitioned db config
isolate_part_errs setup_and_load_errs
```

**注：**对于装入重新启动操作，将使用在 `LOAD RESTART` 命令中指定的选项，因此，这些选项应该与原始 `LOAD` 命令中指定的选项完全相同，这一点非常重要。

---

## 为分区数据库环境装入配置选项

### MODE X

指定装入多分区数据库时装入操作采用的方式。`PARTITION_AND_LOAD` 是缺省值。有效的值为：

- `PARTITION_AND_LOAD`。对数据进行分布（有可能以并行方式进行分布），并且同时在各个相应数据库分区上装入数据。
- `PARTITION_ONLY`。对数据进行分布（有可能以并行方式进行分布），并将输出写入每个装入数据库分区上指定位置中的文件。对于 `CURSOR` 以外的文件类型来说，每个数据库分区上的输出文件名格式都是 `filename.xxx`，其中 `filename` 是 `LOAD` 命令中指定的输入文件名，`xxx` 是 3 位的数据库分区号。对于 `CURSOR` 文件类型来说，每个数据库分区上的输出文件名由 `PART_FILE_LOCATION` 选项确定。请参阅 `PART_FILE_LOCATION` 选项以了解有关如何指定每个数据库分区的分布文件位置的详细信息。

#### 注：

1. 此方式不能用于 `CLI` 装入操作。
  2. 如果表包含进行分布时所需的标识列，那么除非指定了 `identityoverride` 文件类型修饰符，否则不支持此方式。
  3. 为文件类型 `CURSOR` 生成的分布文件在 `DB2` 发行版之间不兼容。这意味着不能使用 `LOAD_ONLY` 方式装入在先前发行版中生成的文件类型为 `CURSOR` 的分布文件。同样，不能使用 `LOAD_ONLY` 方式在将来发行版中装入在当前发行版中生成的文件类型为 `CURSOR` 的分布文件。
- `LOAD_ONLY`。假定已对数据进行分布；将跳过分布过程，并且在相应的数据库分区上同时装入数据。对于 `CURSOR` 以外的文件类型来说，每个数据库分区的输入文件名格式都应该是 `filename.xxx`，其中 `filename` 是 `LOAD` 命

令中指定的文件名，xxx 是 3 位的数据库分区号。对于 CURSOR 文件类型来说，每个数据库分区上的输入文件名由 PART\_FILE\_LOCATION 选项确定。请参阅 PART\_FILE\_LOCATION 选项以了解有关如何指定每个数据库分区的分布文件位置的详细信息。

**注:**

1. 此方式不能用于 CLI 装入操作；或者当指定了 LOAD 命令的 CLIENT 选项时，也不能使用此方式。
  2. 如果表包含进行分布时所需的标识列，那么除非指定了 identityoverride 文件类型修饰符，否则不支持此方式。
- **LOAD\_ONLY\_VERIFY\_PART.** 假定已对数据进行分布，但数据文件未包含分区头。将跳过分布过程，并且在相应的数据库分区上同时装入数据。在装入操作期间，将检查每一行以验证它是否在正确的数据库分区中。如果指定了 dumpfile 文件类型修饰符，就会将发生数据库分区违例的行放到转储文件中。否则会删除那些行。如果特定装入数据库分区上存在数据库分区违例，那么会将一条有关该数据库分区的警告写至装入消息文件。每个数据库分区的输入文件名格式都应该是 filename.xxx，其中 filename 是 LOAD 命令中指定的文件名，xxx 是 3 位的数据库分区号。请参阅 PART\_FILE\_LOCATION 选项以了解有关如何指定每个数据库分区的分布文件位置的详细信息。

**注:**

1. 此方式不能用于 CLI 装入操作；或者当指定了 LOAD 命令的 CLIENT 选项时，也不能使用此方式。
  2. 如果表包含进行分布时所需的标识列，那么除非指定了 identityoverride 文件类型修饰符，否则不支持此方式。
- **ANALYZE.** 生成最佳分布图（在所有数据库分区之间均匀地分布数据）。

### **PART\_FILE\_LOCATION X**

在 PARTITION\_ONLY、LOAD\_ONLY 和 LOAD\_ONLY\_VERIFY\_PART 方式下，此参数可用来指定分布文件的位置。在 OUTPUT\_DBPARTNUMS 选项指定的每个数据库分区上，此位置必须存在。如果指定的位置是相对路径名，那么会将该路径追加至当前目录以创建分布式文件位置。

对于 CURSOR 文件类型来说，必须指定此选项，并且位置必须引用标准文件名。在 PARTITION\_ONLY 方式下，此名称是在每个输出数据库分区上创建的分布式文件的标准基本文件名，或者在 LOAD\_ONLY 方式，此名称是对于每个数据库分区要读取的文件的位置。使用 PARTITION\_ONLY 方式时，如果目标表包含 LOB 列，那么可以使用指定的基本名称来创建多个文件。

对于 CURSOR 以外的文件类型来说，如果未指定此选项，那么将使用当前目录来存储分布式文件。

### **OUTPUT\_DBPARTNUMS X**

X 表示数据库分区号列表。数据库分区号表示要执行装入操作的数据库分区。数据库分区号必须是定义了该表的数据库分区的子集。缺省情况下，选择了所有数据库分区。必须将此列表括在圆括号中，并且列表项必须由逗号分隔。允许指定范围（例如，(0, 2 to 10, 15)）。

### **PARTITIONING\_DBPARTNUMS X**

X 表示分布过程中使用的数据库分区号列表。必须将此列表括在圆括号中，并且

列表项必须由逗号分隔。允许指定范围（例如，(0, 2 to 10, 15)）。对分布过程指定的数据库分区可能与要装入的数据库分区不同。如果未指定 `PARTITIONING_DBPARTNUMS`，那么 `load` 实用程序会确定需要的数据库分区数以及为获得最优性能而需要使用的数据库分区。

如果在 `LOAD` 命令中未指定 `anyorder` 文件类型修饰符，那么在装入会话中将只使用一个分区代理程序。此外，如果仅对 `OUTPUT_DBPARTNUMS` 选项指定了一个数据库分区，或者装入操作的协调程序分区不是 `OUTPUT_DBPARTNUMS` 的元素，那么会在分布过程中使用装入操作的协调程序分区。否则，在分布过程中使用 `OUTPUT_DBPARTNUMS` 中的第一个数据库分区（不是协调程序分区）。

如果指定了 `anyorder` 文件类型修饰符，那么按以下方式确定分布过程中使用的数据库分区数：(`OUTPUT_DBPARTNUMS` 中的分区数/4 + 1)。

### **MAX\_NUM\_PART\_AGENTS X**

指定装入会话中要使用的最大分区代理程序数。缺省值为 25。

### **ISOLATE\_PART\_ERRS X**

指示装入操作如何对各个数据库分区上发生的错误作出反应。除非同时指定了 `LOAD` 命令的 `ALLOW READ ACCESS` 和 `COPY YES` 选项（在这种情况下，缺省值为 `NO_ISOLATION`），否则缺省值为 `LOAD_ERRS_ONLY`。有效的值为：

- `SETUP_ERRS_ONLY`。设置期间在数据库分区上发生的错误（例如，访问数据库分区时发生的问题，或者访问数据库分区上的表空间或表时发生的问题）将导致装入操作在发生故障的数据库分区上停止运行，但在其余数据库分区上继续运行。装入数据时在数据库分区上发生的错误将导致整个操作失败。
- `LOAD_ERRS_ONLY`。设置期间在数据库分区上发生的错误将导致整个装入操作失败。如果在装入数据时发生错误，那么装入操作将在出错的数据库分区上停止运行。装入操作将在其余数据库分区上继续运行，直到发生故障或者装入了所有数据为止。在执行装入重新启动操作并成功完成之前，新装入的数据将不可视。

**注：**在同时指定了 `LOAD` 命令的 `ALLOW READ ACCESS` 和 `COPY YES` 选项时，不能使用此方式。

- `SETUP_AND_LOAD_ERRS`。在此方式下，设置期间或装入数据期间发生的数据库分区级别错误将导致仅在受影响的数据库分区上停止处理装入操作。对于 `LOAD_ERRS_ONLY` 方式，如果在装入数据时发生分区错误，那么在执行装入重新启动操作并成功完成之前，新装入的数据将不可视。

**注：**在同时指定了 `LOAD` 命令的 `ALLOW READ ACCESS` 和 `COPY YES` 选项时，不能使用此方式。

- `NO_ISOLATION`。装入操作期间发生的任何错误都会导致装入操作失败。

### **STATUS\_INTERVAL X**

X 表示读取多少数据量时发出通知。计量单位是兆字节（MB）。缺省值是 100 MB。有效值是 1 到 4000 的整数。

### **PORT\_RANGE X**

X 表示用来创建内部通信套接字的 TCP 端口的范围。缺省范围是 6000 到

6063。如果在调用时定义了 **DB2ATLD\_PORTS** 注册表变量的值，那么该值将替换 **PORT\_RANGE** 装入配置选项的值。对于 **DB2ATLD\_PORTS** 注册表变量，应该使用以下格式来提供范围：

```
<lower-port-number:higher-port-number>
```

在 CLP 中，格式为：

```
( lower-port-number, higher-port-number )
```

### **CHECK\_TRUNCATION**

指定程序应该在输入/输出时检查数据记录截断情况。缺省行为是：输入/输出时不检查数据截断情况。

### **MAP\_FILE\_INPUT X**

X 指定分布图的输入文件名。由于此参数指向包含定制分布图的文件，所以，如果使用定制分布图，就必须指定此参数。通过使用 **db2gpmmap** 程序从数据库系统目录表中抽取映射，或者使用 **LOAD** 命令的 **ANALYZE** 方式来生成最佳映射，可以创建定制分布图。必须先将使用 **ANALYZE** 方式生成的映射移至数据库中的每个数据库分区，这样装入操作才能继续运行。

### **MAP\_FILE\_OUTPUT X**

X 表示分布图的输出文件名。将在发出 **LOAD** 命令的数据库分区上创建输出文件（假定执行分区操作的数据库分区组包含该数据库分区）。如果在未参与分区的数据库分区（由 **PARTITIONING\_DBPARTNUMS** 定义）上调用 **LOAD** 命令，那么会在使用 **PARTITIONING\_DBPARTNUMS** 参数定义的第一个数据库分区上创建输出文件。考虑以下分区数据库环境设置：

```
1 serv1 0
2 serv1 1
3 serv2 0
4 serv2 1
5 serv3 0
```

在 **serv3** 上运行以下 **LOAD** 命令将在 **serv1** 上创建分布图。

```
LOAD FROM file OF ASC METHOD L ( ...) INSERT INTO table CONFIG
MODE ANALYZE PARTITIONING_DBPARTNUMS(1,2,3,4)
MAP_FILE_OUTPUT '/home/db2user/distribution.map'
```

指定了 **ANALYZE** 方式时，应该使用此参数。生成最佳分布图（在所有数据库分区之间均匀地分布数据）。如果未指定此参数但指定了 **ANALYZE** 方式，那么程序将出错并退出。

### **TRACE X**

当您要求复查数据转换过程转储和散列值输出时，指定要跟踪的记录个数。缺省值为 0。

### **NEWLINE**

当输入数据文件是 **ASC** 文件（各个记录由换行符定界），并且在 **LOAD** 命令中指定了 **reclen** 文件类型修饰符时，使用此选项。当指定了此选项时，将对每个记录检查换行符。还将检查 **reclen** 文件类型修饰符中指定的记录长度。

### **DISTFILE X**

如果指定了此选项，那么 **load** 实用程序将生成具有给定名称的数据库分区分布文件。数据库分区分布文件包含 4096 个整数；目标表分布图中的每个条目都有一个对应的整数。此文件中的每个整数都表示所装入输入文件中被分散到相应分布图条目的行数。此信息可以帮助您标识数据偏差，并且还可以帮助您确定

是否应该使用实用程序的 ANALYZE 方式来生成表的新分布图。如果未指定此选项，那么 load 实用程序的缺省行为是不生成分布文件。

**注：**当指定了此选项时，对装入操作最多使用一个分布代理程序。即使您显式请求多个分布代理程序，也只使用一个。

### OMIT\_HEADER

指定在分布文件中不应包括分布图头。如果未指定，那么生成头。

### RUN\_STAT\_DBPARTNUM X

如果在 LOAD 命令中指定了 STATISTICS YES 参数，那么将只在一个数据库分区上收集统计信息。此参数指定要收集统计信息的数据库分区。如果值为 -1，或者根本未指定值，那么将在输出数据库分区列表中的第一个数据库分区上收集统计信息。

---

## 分区数据库环境中的装入会话 - CLP 示例

下列示例说明如何在多分区数据库中装入数据。

数据库有四个数据库分区，其编号从 0 到 3。数据库 WSDDB 是在所有数据库分区上定义的，表 TABLE1 在缺省数据库分区组中，该数据库分区组也是在所有数据库分区上定义的。

### 示例 1

要将用户数据文件 load.del 中的数据装入到 TABLE1 中（该文件在数据库分区 0 上），请连接到数据库分区 0，然后发出以下命令：

```
load from load.del of del replace into table1
```

如果装入操作成功，那么输出将如下所示：

| 代理程序类型        | 节点           | SQL 代码    | 结果       |
|---------------|--------------|-----------|----------|
| LOAD          | 000          | +00000000 | Success. |
| LOAD          | 001          | +00000000 | Success. |
| LOAD          | 002          | +00000000 | Success. |
| LOAD          | 003          | +00000000 | Success. |
| PARTITION     | 001          | +00000000 | Success. |
| PRE_PARTITION | 000          | +00000000 | Success. |
| 结果:           | 成功完成了 4 个装入。 |           |          |

分区代理程序的总结:  
读取的行数 = 100000  
拒绝的行数 = 0  
分区的行数 = 100000

装入代理程序的总结:  
读取的行数 = 100000  
跳过的行数 = 0  
装入的行数 = 100000  
拒绝的行数 = 0  
删除的行数 = 0  
落实的行数 = 100000

输出指示在每个数据库分区上有一个装入代理程序，并且每个装入代理程序都运行成功。输出还显示在协调程序分区上运行了一个预分区代理程序，在数据库分区 1 上运行了一个分区代理程序。这些进程都成功完成并返回正常 SQL 返回码 0。统计总结显示预分区代理程序读取了 100,000 行，分区代理程序分布了 100,000 行，装入代理程序装入的总行数为 100,000。

## 示例 2

在以下示例中，以 PARTITION\_ONLY 方式将数据装入到 TABLE1 中。分布式输出文件存储在每个输出数据库分区上的 /db/data 目录中：

```
load from load.del of del replace into table1 partitioned db config mode
partition_only part_file_location /db/data
```

以上 LOAD 命令的输出如下所示：

| 代理程序类型        | 节点  | SQL 代码    | 结果       |
|---------------|-----|-----------|----------|
| LOAD_TO_FILE  | 000 | +00000000 | Success. |
| LOAD_TO_FILE  | 001 | +00000000 | Success. |
| LOAD_TO_FILE  | 002 | +00000000 | Success. |
| LOAD_TO_FILE  | 003 | +00000000 | Success. |
| PARTITION     | 001 | +00000000 | Success. |
| PRE_PARTITION | 000 | +00000000 | Success. |

```
分区代理程序的总结:
读取的行数           = 100000
拒绝的行数           = 0
分区的行数           = 100000
```

输出指示在每个输出数据库分区上都运行了“装入到文件”代理程序，这些代理程序运行成功。在协调程序分区上运行了一个预分区代理程序，在数据库分区 1 上运行了一个分区代理程序。统计总结显示预分区代理程序成功读取了 100,000 行，分区代理程序成功分布了 100,000 行。由于未将任何行装入到表中，因此未显示已装入行数总结。

## 示例 3

要装入在以上 PARTITION\_ONLY 装入操作期间生成的文件，请发出以下命令：

```
load from load.del of del replace into table1 partitioned db config mode
load_only part_file_location /db/data
```

load 命令的输出如下所示：

| 代理程序类型 | 节点           | SQL 代码    | 结果       |
|--------|--------------|-----------|----------|
| LOAD   | 000          | +00000000 | Success. |
| LOAD   | 001          | +00000000 | Success. |
| LOAD   | 002          | +00000000 | Success. |
| LOAD   | 003          | +00000000 | Success. |
| 结果:    | 成功完成了 4 个装入。 |           |          |

```

装入代理程序的总结:
读取的行数           = 100000
跳过的行数           = 0
装入的行数           = 100000
拒绝的行数           = 0
删除的行数           = 0
落实的行数           = 100000

```

此输出显示每个输出数据库分区上的装入代理程序都运行成功，并且所有装入代理程序装入的总行数为 100,000。由于未执行分布操作，因此未显示分布行数总结。

#### 示例 4 - 失败的装入操作

如果发出以下 LOAD 命令:

```
load from load.del of del replace into table1
```

并且其中一个装入数据库分区在装入操作期间耗尽表空间，那么将返回以下输出:

```

SQL0289N 不能在表空间"DMS4KT"中分配新页。
SQLSTATE=57011

```

| 代理程序类型        | 节点                 | SQL 代码    | 结果                          |
|---------------|--------------------|-----------|-----------------------------|
| LOAD          | 000                | +00000000 | Success.                    |
| LOAD          | 001                | -00000289 | Error. May require RESTART. |
| LOAD          | 002                | +00000000 | Success.                    |
| LOAD          | 003                | +00000000 | Success.                    |
| PARTITION     | 001                | +00000000 | Success.                    |
| PRE_PARTITION | 000                | +00000000 | Success.                    |
| 结果:           | 成功完成了 4 个装入中的 3 个。 |           |                             |

```

分区代理程序的总结:
读取的行数           = 0
拒绝的行数           = 0
分区的行数           = 0

```

```

装入代理程序的总结:
读取的行数           = 0
跳过的行数           = 0
已装入的行数         = 0
拒绝的行数           = 0
删除的行数           = 0
已落实的行数         = 0

```

输出指示装入操作返回了错误 SQL0289。数据库分区总结指示数据库分区 1 耗尽空间。如果对数据库分区 1 上的表空间容器添加了更多空间，那么可以按如下方式重新启动装入操作:

```
load from load.del of del restart into table1
```

## 迁移和版本兼容性

在多分区数据库中，**DB2\_PARTITIONEDLOAD\_DEFAULT** 注册表变量可用来还原为 DB2 通用数据库™ 版本 8 之前的装入行为。

**注：**从版本 9.5 起，就不推荐使用 DB2\_PARTITIONEDLOAD\_DEFAULT 注册表变量并且在以后的发行版中可能会将其除去。

通过在多分区数据库中还原为 LOAD 命令的 DB2 UDB 版本 8 之前的行为，可以将带有有效分布头的文件装入到单一数据库分区中，而不必指定任何其他分区数据库配置选项。将 **DB2\_PARTITIONEDLOAD\_DEFAULT** 的值设置为 NO 可以达到此目的。如果要避免修改对单一数据库分区发出 LOAD 命令的现有脚本，那么可以选择使用此选项。例如，要将分布文件装入到一个表的数据库分区 3 中（而该表所在的数据库分区组包含 4 个数据库分区），请发出以下命令：

```
db2set DB2_PARTITIONEDLOAD_DEFAULT=NO
```

然后，从 DB2 命令行处理器中发出下列命令：

```
CONNECT RESET

SET CLIENT CONNECT_NODE 3

CONNECT TO DB MYDB

LOAD FROM LOAD.DEL OF DEL REPLACE INTO TABLE1
```

在多分区数据库中，当未指定多分区数据库装入配置选项时，将在所有用来定义该表的数据库分区上执行装入操作。输入文件不需要分布头，并且 MODE 选项缺省为 PARTITION\_AND\_LOAD。要装入单一数据库分区，必须指定 OUTPUT\_DBPARTNUMS 选项。



---

## 第 16 章 迁移分区数据库环境

---

### 迁移分区数据库

迁移分区数据库环境需要您在所有数据库分区服务器中安装 DB2 版本 9.5，迁移实例，然后迁移数据库。

可以从目录数据库分区服务器或任何其他数据库分区服务器中迁移数据库分区服务器。如果迁移过程失败，可以从目录数据库分区服务器或者其他任何数据库分区服务器中重试迁移。

由于这种迁移是一项非常重要的任务，因此，对迁移过程及其先决条件和限制的描述超出了本书的范围。《迁移指南》中的『迁移分区数据库环境』主题中不仅提供了详细描述，还提供许多其他在执行迁移之前应查看的参考主题。



---

## 第 17 章 使用快照和事件监视器

---

### 使用快照监视器数据来监视分区表的重组

下列信息描述了一些最有用的全局表重组状态监视方法。

没有单独的用于指示分区表整体表重组状态的数据组。分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。但是，可以根据所重组的各个数据分区数据组中的元素值来推断全局表重组状态。下列信息描述了一些最有用的全局表重组状态监视方法。

#### 确定重组的数据分区数

通过计算表名和模式名相同的表数据监视器数据块数，可以确定表中重组的数据分区总数。此值指示启动了重组的数据分区数。示例 1 和 2 指示正在对 3 个数据分区进行重组。

#### 标识所重组的数据分区

可以根据阶段开始时间（`reorg_phase_start`）来推断当前正在重组的数据分区。在 `SORT/BUILD/REPLACE` 阶段，与正在重组的数据分区相对应的监视器数据显示了最新阶段开始时间。在 `INDEX_RECREATE` 阶段，所有数据分区的阶段开始时间都是相同的。在示例 1 和 2 中，指示了 `INDEX_RECREATE` 阶段，因此所有数据分区的开始时间都是相同的。

#### 标识索引重建需求

通过获取与任何一个正在重组的数据分区相对应的最大重组阶段数（`reorg_max_phase`）元素值，可以确定是否需要重建索引。如果 `reorg_max_phase` 值为 3 或 4，那么表示需要重建索引。示例 1 和 2 报告的 `reorg_max_phase` 值为 3，即表示需要重建索引。

以下样本输出来自一台 3 节点服务器，该服务器包含一个带有 3 个数据分区的表：

```
CREATE TABLE sales (c1 INT, c2 INT, c3 INT)
PARTITION BY RANGE (c1)
(PART P1 STARTING FROM (1) ENDING AT (10) IN parttbs,
PART P2 STARTING FROM (11) ENDING AT (20) IN parttbs,
PART P3 STARTING FROM (21) ENDING AT (30) IN parttbs)
DISTRIBUTE BY (c2)
```

执行的语句：

```
REORG TABLE sales ALLOW NO ACCESS ON ALL DBPARTITIONNUMS
```

示例 1:

```
GET SNAPSHOT FOR TABLES ON DPARTDB GLOBAL
```

已将输出修改为仅包括相关表的表信息。

## 表快照

```

第一个数据库连接时间戳记      = 06/28/2005 13:46:43.061690
上次复位时间戳记              = 06/28/2005 13:46:47.440046
快照时间戳记                  = 06/28/2005 13:46:50.964033
数据库名称                    = DPARTDB
数据库路径                    = /work/sales/NODE0000/SQL00001/
输入数据库别名                = DPARTDB
已访问的表的数目              = 5
    
```

```

表列表
表模式                        = NEWTON
表名                          = SALES
表类型                        = 用户
数据分区标识                  = 0
数据对象页                    = 3
读取的行数                    = 12
写入的行数                    = 1
溢出数                        = 0
重组的页数                    = 0
表重组信息:
  节点号                      = 0
  重组类型                    =
    
```

```

回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
    
```

```

重组索引数                    = 0
重组表空间数                  = 3
长临时空间标识                = 3
  开始时间                    = 06/28/2005 13:46:49.816883
  重组阶段                    = 3 - 索引重建
  最大阶段                    = 3
  阶段开始时间                = 06/28/2005 13:46:50.362918
  状态                        = 已完成
  当前计数器                  = 0
  最大计数器                  = 0
  完成                        = 0
  结束时间                    = 06/28/2005 13:46:50.821244
    
```

```

表重组信息:
  节点数                      = 1
  重组类型                    =
回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
    
```

```

重组索引数                    = 0
重组表空间数                  = 3
长临时空间标识                = 3
  开始时间                    = 06/28/2005 13:46:49.822701
  重组阶段                    = 3 - 索引重建
  最大阶段                    = 3
  阶段开始时间                = 06/28/2005 13:46:50.420741
  状态                        = 已完成
  当前计数器                  = 0
  最大计数器                  = 0
  完成                        = 0
  结束时间                    = 06/28/2005 13:46:50.899543
    
```

```

表重组信息:
  节点数                      = 2
  重组类型                    =
    
```

回收

```

表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数 = 0
重组表空间数 = 3
长临时空间标识 = 3
开始时间 = 06/28/2005 13:46:49.814813
重组阶段 = 3 - 索引重建
最大阶段 = 3
阶段开始时间 = 06/28/2005 13:46:50.344277
状态 = 已完成
当前计数器 = 0
最大计数器 = 0
完成 = 0
结束时间 = 06/28/2005 13:46:50.803619

```

```

表模式 = NEWTON
表名 = SALES
表类型 = 用户
数据分区标识 = 1
数据对象页 = 3
读取的行数 = 8
写入的行数 = 1
溢出数 = 0
重组的页数 = 0
表重组信息:
节点号 = 0
重组类型 =

```

```

回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数 = 0
重组表空间数 = 3
长临时空间标识 = 3
开始时间 = 06/28/2005 13:46:50.014617
重组阶段 = 3 - 索引重建
最大阶段 = 3
阶段开始时间 = 06/28/2005 13:46:50.362918
状态 = 已完成
当前计数器 = 0
最大计数器 = 0
完成 = 0
结束时间 = 06/28/2005 13:46:50.821244

```

```

表重组信息:
节点数 = 1
重组类型 =

```

```

回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数 = 0
重组表空间数 = 3
长临时空间标识 = 3
开始时间 = 06/28/2005 13:46:50.026278
重组阶段 = 3 - 索引重建
最大阶段 = 3
阶段开始时间 = 06/28/2005 13:46:50.420741
状态 = 已完成
当前计数器 = 0
最大计数器 = 0
完成 = 0

```

```

      结束时间          = 06/28/2005 13:46:50.899543
表重组信息:
  节点数              = 2
  重组类型            =
    回收
    表重组
    不允许访问
    通过表扫描重新集群
    仅重组数据
  重组索引数          = 0
  重组表空间数        = 3
长临时空间标识      = 3
  开始时间            = 06/28/2005 13:46:50.006392
  重组阶段            = 3 - 索引重建
  最大阶段            = 3
  阶段开始时间        = 06/28/2005 13:46:50.344277
  状态                = 已完成
  当前计数器          = 0
  最大计数器          = 0
  完成                = 0
  结束时间            = 06/28/2005 13:46:50.803619

```

```

表模式              = NEWTON
表名                = SALES
表类型              = 用户
数据分区标识        = 2
数据对象页          = 3
读取的行数          = 4
写入的行数          = 1
溢出数              = 0
重组的页数          = 0
表重组信息:
  节点号              = 0
  重组类型            =

```

```

    回收
    表重组
    不允许访问
    通过表扫描重新集群
    仅重组数据
  重组索引数          = 0
  重组表空间数        = 3
长临时空间标识      = 3
  开始时间            = 06/28/2005 13:46:50.199971
  重组阶段            = 3 - 索引重建
  最大阶段            = 3
  阶段开始时间        = 06/28/2005 13:46:50.362918
  状态                = 已完成
  当前计数器          = 0
  最大计数器          = 0
  完成                = 0
  结束时间            = 06/28/2005 13:46:50.821244

```

```

表重组信息:
  节点数              = 1
  重组类型            =
    回收
    表重组
    不允许访问
    通过表扫描重新集群
    仅重组数据
  重组索引数          = 0
  重组表空间数        = 3
长临时空间标识      = 3
  开始时间            = 06/28/2005 13:46:50.223742
  重组阶段            = 3 - 索引重建

```

```

最大阶段 = 3
阶段开始时间 = 06/28/2005 13:46:50.420741
状态 = 已完成
当前计数器 = 0
最大计数器 = 0
完成 = 0
结束时间 = 06/28/2005 13:46:50.899543

```

```

表重组信息:
节点数 = 2
重组类型 =
回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数 = 0
重组表空间数 = 3
长临时空间标识 = 3
开始时间 = 06/28/2005 13:46:50.179922
重组阶段 = 3 - 索引重建
最大阶段 = 3
阶段开始时间 = 06/28/2005 13:46:50.344277
状态 = 已完成
当前计数器 = 0
最大计数器 = 0
完成 = 0
结束时间 = 06/28/2005 13:46:50.803619

```

示例 2:

#### GET SNAPSHOT FOR TABLES ON DPARTDB AT DBPARTITIONNUM 2

已将输出修改为仅包括相关表的表信息。

#### 表快照

```

第一个数据库连接时间戳记 = 06/28/2005 13:46:43.617833
上次复位时间戳记 =
快照时间戳记 = 06/28/2005 13:46:51.016787
数据库名称 = DPARTDB
数据库路径 = /work/sales/NODE0000/SQL00001/
输入数据库别名 = DPARTDB
已访问的表的数目 = 3

```

```

表列表
表模式 = NEWTON
表名 = SALES
表类型 = 用户
数据分区标识 = 0
数据对象页 = 1
读取的行数 = 0
写入的行数 = 0
溢出数 = 0
重组的页数 = 0
表重组信息:
节点数 = 2
重组类型 =
回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数 = 0
重组表空间数 = 3
长临时空间标识 = 3

```

```

开始时间          = 06/28/2005 13:46:49.814813
重组阶段          = 3 - 索引重建
最大阶段          = 3
阶段开始时间      = 06/28/2005 13:46:50.344277
状态              = 已完成
当前计数器        = 0
最大计数器        = 0
完成              = 0
结束时间          = 06/28/2005 13:46:50.803619

```

```

表模式            = NEWTON
表名              = SALES
表类型            = 用户
数据分区标识      = 1
数据对象页        = 1
读取的行数        = 0
写入的行数        = 0
溢出数            = 0
重组的页数        = 0
表重组信息:
  节点数          = 2
  重组类型        =
回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数        = 0
重组表空间数      = 3
长临时空间标识    = 3
  开始时间          = 06/28/2005 13:46:50.006392
  重组阶段          = 3 - 索引重建
  最大阶段          = 3
  阶段开始时间      = 06/28/2005 13:46:50.344277
  状态              = 已完成
  当前计数器        = 0
  最大计数器        = 0
  完成              = 0
  结束时间          = 06/28/2005 13:46:50.803619

```

```

表模式            = NEWTON
表名              = SALES
表类型            = 用户
数据分区标识      = 2
数据对象页        = 1
读取的行数        = 4
写入的行数        = 1
溢出数            = 0
重组的页数        = 0
表重组信息:
  节点数          = 2
  重组类型        =
回收
表重组
不允许访问
通过表扫描重新集群
仅重组数据
重组索引数        = 0
重组表空间数      = 3
长临时空间标识    = 3
  开始时间          = 06/28/2005 13:46:50.179922
  重组阶段          = 3 - 索引重建
  最大阶段          = 3
  阶段开始时间      = 06/28/2005 13:46:50.344277
  状态              = 已完成

```



```

当前计数器      = 0
最大计数器      = 0
完成            = 0
结束时间        = 06/28/2005 13:46:50.803619

```

示例 3:

```
SELECT * FROM SYSIBMADM.SNAPLOCK WHERE tabname = 'SALES';
```

已将输出修改为仅包括相关表的表信息的子集。

| ... | TBSP_NAME | TABNAME | LOCK_OBJECT_TYPE | LOCK_MODE | LOCK_STATUS | ... |
|-----|-----------|---------|------------------|-----------|-------------|-----|
| ... | PARTTBS   | SALES   | ROW_LOCK         | X         | GRNT        | ... |
| ... | -         | SALES   | TABLE_LOCK       | IX        | GRNT        | ... |
| ... | PARTTBS   | SALES   | TABLE_PART_LOCK  | IX        | GRNT        | ... |
| ... | PARTTBS   | SALES   | ROW_LOCK         | X         | GRNT        | ... |
| ... | -         | SALES   | TABLE_LOCK       | IX        | GRNT        | ... |
| ... | PARTTBS   | SALES   | TABLE_PART_LOCK  | IX        | GRNT        | ... |
| ... | PARTTBS   | SALES   | ROW_LOCK         | X         | GRNT        | ... |
| ... | -         | SALES   | TABLE_LOCK       | IX        | GRNT        | ... |
| ... | PARTTBS   | SALES   | TABLE_PART_LOCK  | IX        | GRNT        | ... |

选择了 9 个记录。

此查询（已继续）的输出。

| ... | LOCK_ESCALATION | LOCK_ATTRIBUTES | DATA_PARTITION_ID | DBPARTITIONNUM |
|-----|-----------------|-----------------|-------------------|----------------|
| ... | 0               | INSERT          | 2                 | 2              |
| ... | 0               | NONE            | -                 | 2              |
| ... | 0               | NONE            | 2                 | 2              |
| ... | 0               | INSERT          | 0                 | 0              |
| ... | 0               | NONE            | -                 | 0              |
| ... | 0               | NONE            | 0                 | 0              |
| ... | 0               | INSERT          | 1                 | 1              |
| ... | 0               | NONE            | -                 | 1              |
| ... | 0               | NONE            | 1                 | 1              |

示例 4:

```
SELECT * FROM SYSIBMADM.SNAPTAB WHERE tabname = 'SALES';
```

已将输出修改为仅包括相关表的表信息的子集。

| ... | TABSCHEMA | TABNAME | TAB_FILE_ID | TAB_TYPE   | DATA_OBJECT_PAGES | ROWS_WRITTEN | ... |
|-----|-----------|---------|-------------|------------|-------------------|--------------|-----|
| ... | NEWTON    | SALES   | 2           | USER_TABLE | 1                 | 1            | ... |
| ... | NEWTON    | SALES   | 4           | USER_TABLE | 1                 | 1            | ... |
| ... | NEWTON    | SALES   | 3           | USER_TABLE | 1                 | 1            | ... |

选择了 3 个记录。

此查询（已继续）的输出。

| ... | OVERFLOW_ACCESSES | PAGE_REORGS | DBPARTITIONNUM | TBSP_ID | DATA_PARTITION_ID |
|-----|-------------------|-------------|----------------|---------|-------------------|
| ... | 0                 | 0           | 0              | 3       | 0                 |
| ... | 0                 | 0           | 2              | 3       | 2                 |
| ... | 0                 | 0           | 1              | 3       | 1                 |

示例 5:

```
SELECT * FROM SYSIBMADM.SNAPTAB_REORG WHERE tabname = 'SALES';;
```

已将输出修改为仅包括相关表的表信息的子集。

| REORG_PHASE    | REORG_MAX_PHASE | REORG_TYPE                                    | ... |
|----------------|-----------------|---|-----|
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |
| INDEX_RECREATE | 3               | RECLAIM+OFFLINE+ALLOW_NONE+TABLESCAN+DATAONLY | ... |

选择了 9 个记录。

此查询（已继续）的输出。

| REORG_STATUS | REORG_TBSPC_ID | DBPARTITIONNUM | DATA_PARTITION_ID |
|--------------|----------------|----------------|-------------------|
| COMPLETED    | 3              | 2              | 0                 |
| COMPLETED    | 3              | 2              | 1                 |
| COMPLETED    | 3              | 2              | 2                 |
| COMPLETED    | 3              | 1              | 0                 |
| COMPLETED    | 3              | 1              | 1                 |
| COMPLETED    | 3              | 1              | 2                 |
| COMPLETED    | 3              | 0              | 0                 |
| COMPLETED    | 3              | 0              | 1                 |
| COMPLETED    | 3              | 0              | 2                 |

## 分区数据库系统上的全局快照

在分区数据库系统上，可获取当前分区、指定分区或所有分区的全局快照。对分区数据库的所有分区获取全局快照时，会先聚集数据，然后返回结果。

对不同元素类型聚集数据的方式如下所示：

- **计数器、时间和标尺**

包含从实例中的每个分区收集的所有可能值的总和。例如，`GET SNAPSHOT FOR DATABASE XYZ ON TEST GLOBAL` 对分区数据库实例中的所有分区返回从数据库读取的行数（`rows_read`）。

- **水位标记**

返回分区数据库系统中的任何分区的最高（高水位）或最低（低水位）值。如果返回的值值得关注，那么可以获取各个分区的全局快照以确定特定分区是否使用过度或者问题是否为实例范围内的问题。

- **Timestamp**

设置为连接快照监视器实例代理程序的分区的时间戳记值。注意，所有时间戳记值都在 `timestamp` 监视开关控制之下。

- **信息**

返回可能妨碍工作的分区的最重要信息。例如，对于元素 `appl_status`，如果一个分区上的状态为“正在执行 UOW”，而另一个分区上的状态为“等待锁定”，那么返回“等待锁定”，原因是这是挂起应用程序的执行的执行的状态。

您也可以复位计数器，设置监视开关，以及检索分区数据库中的个别分区或所有分区的监视开关设置。

**注：**获取全局快照时，如果一个或多个分区遇到错误，那么将从成功获取快照的分区收集数据，同时返回一个警告（`sqlcode 1629`）。如果以全局方式获取或更新监视开关，或者计数器在一个或多个分区上复位失败，那么不会设置这些分区的监视开关或进行数据复位。

---

## 为分区数据库创建事件监视器

在分区数据库系统上创建文件或管道事件监视器时，需要确定想要收集的监视数据的作用域。

您将需要 `DBADM` 权限来为分区数据库创建事件监视器。

事件监视器使用操作系统进程或线程来写入事件记录。运行此进程或线程的数据库分区称为监视器分区。如果文件和管道事件监视器在监视器分区上以局部方式运行，或者在运行 `DB2` 数据库管理器的任何分区上以全局方式运行，那么这些事件监视器可充当监视事件。全局事件监视器将在监视分区上写入单个跟踪，它包含所有分区的活动。不管事件监视器是局部的还是全局的，都被称为监视作用域。

监视器分区和监视器作用域都将使用 `CREATE EVENT MONITOR` 语句指定。

仅当监视器分区处于活动状态时，才能激活事件监视器。如果 `SET EVENT MONITOR` 语句用于激活事件监视器但监视器分区尚未处于活动状态，那么将在下一次启动监视器分区时激活事件监视器。而且，在显式释放事件监视器或显式释放实例之后，将自动激活事件监视器。例如，在数据库分区 0 上：

```
db2 connect to sample
db2 create event monitor foo ... on dbpartitionnum 2
db2 set event monitor foo state 1
```

在运行上述命令后，每次数据库 `sample` 在数据库分区 2 上激活时都将自动激活事件监视器 `foo`。发出 `db2 set event monitor foo state 0` 或停止分区 2 之前，会一直进行此自动激活。

局部或全局作用域的注释不适用于写至表事件监视器。激活写至表事件监视器时，事件监视器在所有分区上运行。（更具体地说，事件监视器进程将在属于数据库分区组的分区上运行，而目标表位于这些数据库分区组上。）运行事件监视器进程的每个分区具有同一组目标表。因为是从各个分区的角度表示监视数据，所以这些表中的数据将会有所不同。可通过发出 `SQL` 语句来访问每个分区的事件监视器目标表中的期望值，以获取来自所有分区的聚集值。

每个目标表的第一列称为 `PARTITION_KEY`，并且被用作表的分区键。此列的值将被选中，以便每个事件监视器进程将数据插入到运行该进程的数据库分区中；即，插入操作将在运行事件监视器进程的数据库分区本地执行。在任何数据库分区上，`PARTITION_KEY` 字段都将包含相同的值。这意味着，如果删除数据分区并且再分发数

据，那么被删除数据库分区上的所有数据都将转至另一数据库分区而不是平均分布。因此，在除去数据库分区之前，考虑删除该数据库分区上的所有表行。

此外，可对每个表定义名为 `PARTITION_NUMBER` 的列。此列包含插入数据的分区的编号。与 `PARTITION_KEY` 列不同，`PARTITION_NUMBER` 列不是必需的。

用来定义目标表的表空间必须在所有分区上存在，事件监视器数据将写至这些分区。如果不遵守此规则，那么不会使用事件监视器将记录写至不存在表空间的登录分区。事件仍将写至存在表空间的分区，并且不返回任何错误。此行为允许用户通过创建只在特定分区上存在的表空间，以选择要监视的分区子集。

在写至表事件监视器激活期间，`FIRST_CONNECT` 和 `EVMON_START` 的 `CONTROL` 表行将仅插入至目录数据库分区。这要求目录数据库分区上存在控制表的表空间。如果目录数据库分区上不存在该表空间，那么不会执行这些插入。

如果激活写至表事件监视器时分区尚未激活，那么将在下一次激活该分区时激活事件监视器。

**注：**详细的死锁连接事件中的锁定列表将仅包含这样的锁定，这些锁定是应用程序在等待死锁的分区上挂起的。例如，如果涉及死锁的应用程序正在节点 20 上等待锁定，那么只有该应用程序在节点 20 上挂起的锁定才会包括在列表中。

1. 指定要监视的分区。

```
CREATE EVENT MONITOR dlmon FOR DEADLOCKS
                                WRITE TO FILE '/tmp/dlevents'
                                ON PARTITION 3
```

`dlmon` 表示事件监视器的名称。

`/tmp/dlevents` 是目录路径（在 UNIX 上）的名称，事件监视器会将事件文件写至该目录。

3 表示要监视的分区号。

2. 指定是在局部作用域还是全局作用域收集事件监视器数据。要从所有分区收集事件监视器报告，那么发出以下语句：

```
CREATE EVENT MONITOR dlmon FOR DEADLOCKS
                                WRITE TO FILE '/tmp/dlevents'
                                ON PARTITION 3 GLOBAL
```

只有死锁和与带有详细信息事件监视器的死锁才能定义为 `GLOBAL`。所有分区会将与死锁有关的事件记录报告至分区 3。

3. 要仅从局部分区收集事件监视器报告，那么发出以下语句：

```
CREATE EVENT MONITOR dlmon FOR DEADLOCKS
                                WRITE TO FILE '/tmp/dlevents'
                                ON PARTITION 3 LOCAL
```

这是分区数据库中的文件和管道事件监视器的缺省行为。对于写至表事件监视器，将忽略 `LOCAL` 和 `GLOBAL` 子句。

4. 可复查现有事件监视器的监视器分区和作用域值。为此，请使用以下语句查询 `SYSCAT.EVENTMONITORS` 表：

```
SELECT EVMONNAME, NODENUM, MONSCOPE FROM SYSCAT.EVENTMONITORS
```

一旦创建并激活事件监视器，该事件监视器就会在指定的事件发生时记录监视数据。



## 第 18 章 开发好的备份和恢复策略

### 崩溃恢复

对数据库执行的事务（也称工作单元）可能被意外中断。如果在作为工作单元一部分的所有更改完成和落实之前发生故障，那么该数据库就会处于不一致和不可用的状态。崩溃恢复是将数据库移动回一致并可用状态的进程。为此，回滚未完成的事务，并完成当发生崩溃时仍在内存中的已落实事务（图 45）。当数据库处于一致和可用状态时，它处于一种称为“一致点”的状态。

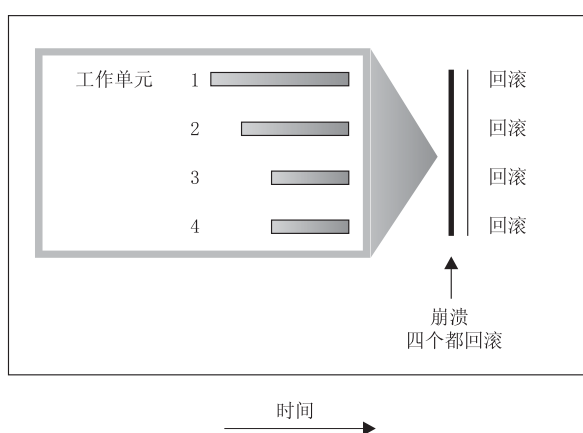


图 45. 回滚工作单元（崩溃恢复）

事务处理失败是由于出现了严重错误或导致数据库或数据库管理器异常结束的情况。部分完成的工作单元或发生故障时未清仓至磁盘中的 UOW 使数据库处于不一致状态。在事务处理故障之后必须恢复数据库。导致事务处理故障的情况有：

- 机器上的断电故障，它会导致使用该机器的数据库管理器和数据库分区崩溃
- 硬件故障，例如内存毁坏、磁盘、CPU 或网络故障。
- 导致 DB2 崩溃的严重操作系统错误
- 应用程序异常终止。

如果您希望不完整工作单元的回滚是由数据库管理器自动完成的，那么应将 *autorestart* 数据库配置参数设置为 ON，以启用该自动重新启动参数。（这是缺省值。）如果不要重新启动行为，那么将 *autorestart* 数据库配置参数设置为 OFF。这样，将需要在数据库故障发生时发出 `RESTART DATABASE` 命令。如果数据库 I/O 在发生崩溃之前已处于暂挂状态，那么必须指定 `RESTART DATABASE` 命令的 `WRITE RESUME` 选项才能使崩溃恢复继续进行。管理通知日志记录数据库重新启动操作开始的时间。

如果对用于正向恢复的数据库应用崩溃恢复（即，未将 *logarchmeth1* 配置参数设置为 OFF），且在崩溃恢复期间因个别表空间而发生错误，那么会让该表空间脱机，直到修复后才能对其进行访问。崩溃恢复继续进行。在崩溃恢复完成时，该数据库中的其他

表空间将是可访问的，并且可以与该数据库建立连接。但是，如果脱机的表空间包含系统目录，那么必须先修复它才允许进行所有连接。

---

## 从分区数据库环境中的事务故障中恢复

如果事务处理失败发生在分区数据库环境中，通常需要对发生了故障的数据库分区服务器和参与了该事务的任何其他数据库分区服务器都进行数据库恢复：

- 对发生了故障的数据库分区服务器的崩溃恢复发生在更正了先前的错误情况后。
- 对其他（仍活动的）数据库分区服务器的数据库分区故障恢复紧接在检测到故障后发生。

在分区数据库环境中，提交应用程序的数据库分区服务器是协调程序分区，而为该应用程序工作的第一个代理程序是协调代理程序。协调代理程序负责将工作分布至其他数据库分区服务器上，并跟踪那些参与了该事务的服务器。当应用程序对一个事务发出 `COMMIT` 语句时，该协调代理程序使用两阶段落实协议来落实该事务。在第一阶段期间，协调程序分区将 `PREPARE` 请求分布至所有其他参与该事务的数据库分区服务器。然后，这些服务器用以下其中一项应答：

### **READ-ONLY**

在此服务器中未发生任何数据更改

**YES** 在此服务器中发生了数据更改

**NO** 由于错误，服务器未准备落实

如果其中一个服务器应答 `NO`，那么回滚该事务。否则，协调程序分区开始第二阶段。

在第二阶段，协调程序分区写入一条 `COMMIT` 日志记录，然后将 `COMMIT` 请求分布至所有应答了 `YES` 的服务器。在所有其他数据库分区服务器都已落实后，它们会将 `COMMIT` 的应答发送至协调程序分区。当协调代理程序从所有参与的服务器接收到所有 `COMMIT` 应答时，该事务完成。在此时间点，协调代理程序会写入一条 `FORGET` 日志记录。

## 活动数据库分区服务器上的事务故障恢复

如果任何数据库分区服务器检测到另一个服务器当机，那么与该发生故障的数据库分区服务器相关的所有工作都会停止：

- 如果仍处于活动状态的数据库分区服务器是某个应用程序的协调程序分区，且该应用程序在发生故障的数据库分区服务器上运行（尚未准备 `COMMIT`），那么会中断该协调代理程序，以便执行故障恢复。如果该协调代理程序处于 `COMMIT` 处理的第二个阶段，会将 `SQL0279N` 返回给应用程序，应用程序随之会丢失它的数据库连接。否则，协调代理程序将一个 `ROLLBACK` 请求分布至所有其他参与该事务的服务器，并将 `SQL1229N` 返回至该应用程序。
- 如果发生故障的数据库分区服务器是该应用程序的协调程序分区，那么仍在活动服务器上为该应用程序工作的代理程序会被中断，以便执行故障恢复。在事务未处于准备就绪状态的每个数据库分区上本地回滚事务。在事务处于准备就绪状态的那些数据库分区上，事务变得不确定。由于协调程序数据库分区不可用，所以协调程序数据库分区不知道某些数据库分区上的事务处于不确定状态。
- 如果该应用程序与发生故障的数据库分区服务器连接（在它发生故障之前），但是本地数据库分区服务器和发生故障的数据库分区服务器都不是协调程序分区，那么



会中断为此应用程序工作的代理程序。协调程序分区将向其他数据库分区服务器发送 ROLLBACK 或断开连接消息。如果协调程序分区返回 SQL0279, 那么事务将只在仍然活动的数据库分区服务器上是不确定的。

试图向该发生故障的服务器发送请求的任何进程（如，代理程序或死锁检测器）都会得到通知：它不能发送该请求。

## 发生故障的数据库分区服务器上的事务故障恢复

如果事务失败导致数据库管理器异常结束，那么可以发出指定了 RESTART 选项的 db2start 命令，以便在重新启动数据库分区后立即重新启动数据库管理器。如果无法重新启动数据库分区，那么可以发出 db2start, 以便在另一数据库分区上重新启动数据库管理器。

如果数据库管理器异常结束，那么服务器上的数据库分区可能会处于不一致状态。要使它们可用，可以在数据库分区服务器上触发崩溃恢复：

- 通过 RESTART DATABASE 命令显式地触发
- 在 autorestart 数据库配置参数已设置为 ON 后，通过 CONNECT 请求隐式触发

崩溃恢复将重新应用活动日志文件中的日志记录，以确保所有已完成的事务的结果都在数据库中。重新应用了这些更改后，除不确定事务外的所有未落实的事务都将本地回滚。分区数据库环境中有两种类型的不确定事务：

- 在不是协调程序分区的数据库分区服务器上，已就绪但未落实的事务就是不确定的。
- 在协调程序分区上，已落实但还未被记录为完成（即，还未写入 FORGET 记录）的事务是不确定的。当协调代理程序未从为该应用程序工作的所有服务器接收到全部 COMMIT 应答时，会发生此情况。

崩溃恢复试图通过以下其中一项操作解决所有不确定事务。要执行的操作取决于数据库分区服务器是否为应用程序的协调程序分区：

- 如果重新启动的服务器不是该应用程序的协调程序分区，它会将一个查询消息发送至该协调代理程序，以发现该事务的结果。
- 如果重新启动的服务器是该应用程序的协调程序分区，它会将一个消息发送至协调代理程序仍在等待它们的所有其他代理程序（下级代理）。

有可能崩溃恢复不能解决所有不确定事务（例如，某些数据库分区服务器可能会不可用）。在这种情况下，会返回 SQL 警告消息 SQL1061W。由于不确定事务占用了资源（例如锁定和活动日志空间），有可能导致不能对数据库进行任何更改，因为不确定事务占用了活动日志空间。因此，应确定在崩溃恢复之后是否还有不确定事务，并尽快恢复解决这些不确定事务所需的所有数据库分区服务器。

如果解决不确定事务所需的一个或多个服务器不能及时恢复，且需要访问其他服务器上的数据库分区，可以通过作出启发式决策来手工解决这些不确定事务。可以使用 LIST INDOUBT TRANSACTIONS 命令来查询、落实和回滚服务器上的不确定事务。

**注：**LIST INDOUBT TRANSACTIONS 命令还可用于分布式事务环境中。要区分这两种类型的不确定事务，LIST INDOUBT TRANSACTIONS 命令返回的输出中的 *originator* 字段显示以下其中一项：

- DB2 企业服务器版，指示该事务始发于分区数据库环境。

- XA，它指示该事务始发于分布式环境中。

## 标识发生故障的数据库分区服务器

当一个数据库分区服务器发生故障时，应用程序通常会接收到下列其中一个 SQLCODE。检测哪个数据库管理器发生故障的方法取决于接收到的 SQLCODE：

### SQL0279N

当在 COMMIT 处理期间终止的事务中涉及了数据库分区服务器时，会接收到此 SQLCODE。

### SQL1224N

当发生故障的数据库分区服务器是该事务的协调程序分区时，会接收到此 SQLCODE。

### SQL1229N

当发生故障的数据库分区服务器不是该事务的协调程序分区时，会接收到此 SQLCODE。

确定哪个发生了故障的数据库分区服务器是一个两阶段进程。与 SQLCODE SQL1229N 相关的 SQLCA 在 *sqlerrd* 字段的第六个数组位置包含检测到错误的服务器的节点号。（为服务器写入的节点号与 *db2nodes.cfg* 文件中的节点号对应。）在检测到错误的数据库分区服务器上，会将指示故障服务器的节点号的消息写至管理通知日志。

**注：**如果正在一个处理器上使用多个逻辑节点，那么一个逻辑节点发生故障会导致同一个处理器上的其他逻辑节点发生故障。

---

## 从数据库分区服务器的故障恢复

要从数据库分区服务器的故障中恢复，请执行下列步骤。

1. 校正导致该故障的问题。
2. 通过从任何数据库分区服务器发出 `db2start` 命令，重新启动数据库管理器。
3. 通过在有故障的一个或多个数据库分区服务器上发出 `RESTART DATABASE` 命令，重新启动数据库。

---

## 重建分区数据库

要重建分区数据库，分别重建每个数据库分区。对于每个数据库分区，从目录分区开始，首先复原需要的所有表空间。未复原的所有表空间都处于复原暂挂状态。复原了所有数据库分区之后，在目录分区上发出 `ROLLFORWARD DATABASE` 命令以前滚所有数据库分区。

**注：**如果在将来的某一天您需要复原最初未包括在重建阶段中的任何表空间，那么需要确保在后来前滚表空间时 `rollforward` 实用程序使数据库分区上的所有数据保持同步。如果在原始复原和前滚操作中丢失了某个表空间，那么直到尝试访问数据时才会检测到这种情况，将出现数据访问错误。需要复原和前滚丢失的表空间，以使它恢复与其余分区的同步。

要使用表空间级备份映像来重建分区数据库，请考虑以下示例。

在此示例中，有一个称为 `SAMPLE` 的可恢复数据库，它具有三个数据库分区：

- 数据库分区 1 包含表空间 SYSCATSPACE、USERSP1 和 USERSP2，它是目录分区
- 数据库分区 2 包含表空间 USERSP1 和 USERSP3
- 数据库分区 3 包含表空间 USERSP1、USERSP2 和 USERSP3

进行了下列备份，其中 BK<sub>xy</sub> 表示分区 y 上的备份编号 x:

- BK11 是 SYSCATSPACE、USERSP1 和 USERSP2 的备份
- BK12 是 USERSP2 和 USERSP3 的备份
- BK13 是 USERSP1、USERSP2 和 USERSP3 的备份
- BK21 是 USERSP1 的备份
- BK22 是 USERSP1 的备份
- BK23 是 USERSP1 的备份
- BK31 是 USERSP2 的备份
- BK33 是 USERSP2 的备份
- BK42 是 USERSP3 的备份
- BK43 是 USERSP3 的备份

下列过程演示了使用 CLP 发出的 RESTORE DATABASE 和 ROLLFORWARD DATABASE 命令将整个数据库重建至日志末尾。

1. 在数据库分区 1 上，发出带有 REBUILD 选项的 RESTORE DATABASE 命令:

```
db2 restore db sample rebuild with all tablespaces in database
      taken at BK31 without prompting
```

2. 在数据库分区 2 上，发出带有 REBUILD 选项的 RESTORE DATABASE 命令:

```
db2 restore db sample rebuild with tablespaces in database
      taken at BK42 without prompting
```

3. 在数据库分区 3 上，发出带有 REBUILD 选项的 RESTORE DATABASE 命令:

```
db2 restore db sample rebuild with all tablespaces in database
      taken at BK43 without prompting
```

4. 在目录分区上，发出带有 TO END OF LOGS 选项的 ROLLFORWARD DATABASE 命令:

```
db2 rollforward db sample to end of logs
```

5. 发出带有 STOP 选项的 ROLLFORWARD DATABASE 命令:

```
db2 rollforward db sample stop
```

此时，该数据库在所有数据库分区上都是可连接的，并且所有表空间都处于 NORMAL 状态。

---

## 使用 db2adutl 来恢复数据

下列示例说明如何使用 db2adutl 命令以及 *logarchopt1* 和 *vendoropt* 数据库配置参数来执行跨节点恢复。

在下列示例中，计算机 1 名为 bar，它正在运行 AIX。此机器的所有者是 roecken。bar 上的数据库名为 zample。计算机 2 名为 dps。此机器也在运行 AIX 并由 regress9 所拥有。

**PASSWORDACCESS = generate**

## 计算机 1

1. 设置数据库以便将日志归档至 TSM。更新 `zample` 数据库的数据库配置参数 `logarchmeth1`:

```
bar:/home/roecken> db2 update db cfg for zample using LOGARCHMETH1 tsm
```

将返回以下信息:

```
成功完成 DB20000I UPDATE DATABASE CONFIGURATION 命令。
```

**注:** 在更新数据库配置之前, 可能需要对数据库进行脱机备份。

2. 强制应用程序断开连接:

```
db2 force applications all
```

3. 验证是否已强制所有应用程序断开连接:

```
db2 list applications
```

您应该接收到一条消息, 说明未返回任何数据。

**注:** 在分区数据库环境中, 必须对所有数据库分区都执行此步骤。

4. 备份数据库:

```
db2 backup db zample use tsm
```

将返回类似以下的信息:

```
备份成功。此备份映像的时间戳记是: 20040216151025
```

**注:** 在分区数据库环境中, 必须对所有数据库分区都执行此步骤。根据您正在执行联机备份还是脱机备份, 在数据库分区上执行此步骤的顺序有所不同。有关更多信息, 请参阅第 367 页的『使用备份』。

5. 连接至 `zample` 数据库, 然后在该数据库中创建一个表。
6. 将数据装入新表中。在此示例中, 表名为 `a`, 从定界 ASCII 码文件中装入的数据为 `mr`。指定了 `COPY YES` 选项来生成装入的数据的副本, 并且 `USE TSM` 选项指定数据的副本存储在 Tivoli Storage Manager 上。

**注:** 仅当数据库启用了前滚恢复功能时才能指定 `COPY YES` 选项; 即, 必须将 `logarchmeth1` 数据库配置参数设置为 `USEREXIT` 或 `LOGRETAIN`。

```
bar:/home/roecken> db2 load from mr of del modified by noheader replace  
into a copy yes use tsm
```

实用程序返回了一系列消息来指示它的进度:

```
SQL3109N 实用程序正开始从"/home/roecken/mr"文件中装入数据。  
SQL3500W 实用程序在"02/16/2004 15:12:13.392633"时进入"装入"阶段。  
SQL3520W 装入一致点成功。  
SQL3110N 实用程序已完成处理。从输入文件读取了"1"行。  
  
SQL3519W 开始装入一致点。输入记录数 = "1"。 SQL3520W 装入一致点成功。  
SQL3515W 实用程序在"02/16/2004 15:12:13.445718"时已完成"装入"阶段。  
读取的行数 = 1  
跳过的行数 = 0  
装入的行数 = 1  
拒绝的行数 = 0  
删除的行数 = 0  
落实的行数 = 1
```

现在，在 TSM 上应该具有一个备份映像、一个装入副本映像和一个日志文件。可以按如下所示对 zample 数据库运行查询：

```
bar:/home/roecken/sql1lib/adsm> db2adutl query db zample
```

将返回以下信息：

```
正在检索 FULL DATABASE BACKUP 信息。  
1 时间: 20040216151025 最早的日志: S0000000.LOG 数据库分区号: 0  
会话数: 1  
正在检索 INCREMENTAL DATABASE BACKUP 信息。  
找不到 ZAMPLE 的 INCREMENTAL DATABASE BACKUP 映像
```

```
正在检索 DELTA DATABASE BACKUP 信息。  
找不到 ZAMPLE 的 DELTA DATABASE BACKUP 映像
```

```
正在检索 TABLESPACE BACKUP 信息。  
找不到 ZAMPLE 的 TABLESPACE BACKUP 映像
```

```
正在检索 INCREMENTAL TABLESPACE BACKUP 信息。  
找不到 ZAMPLE 的 INCREMENTAL TABLESPACE BACKUP 映像
```

```
正在检索 DELTA TABLESPACE BACKUP 信息。  
找不到 ZAMPLE 的 DELTA TABLESPACE BACKUP 映像
```

```
正在检索 LOAD COPY 信息。  
1 时间: 20040216151213
```

```
正在检索 LOG ARCHIVE 信息。  
日志文件: S0000000.LOG, 链号: 0, 数据库分区号: 0  
生成时间: 2004-02-16-15.10.38
```

7. 要启用跨节点恢复，必须为另一个节点和帐户提供对 bar 计算机上的对象的访问权。在此示例中，对 dps 节点和 regress9 用户提供了访问权。

```
bar:/home/roecken/sql1lib/adsm> db2adutl grant user regress9  
on nodename dps for db zample
```

将返回以下信息：

```
成功添加了 regress9 访问 dps 节点上的 ZAMPLE 的许可权。
```

要查询 db2adutl grant 操作的结果，发出以下命令：

```
bar:/home/roecken/sql1lib/adsm> db2adutl queryaccess
```

将返回以下信息：

| 节点  | 用户名      | 数据库名称  | 类型 |
|-----|----------|--------|----|
| DPS | regress9 | ZAMPLE | A  |

访问类型: B - 备份映像 L - 日志 A - 同时使用这两种访问类型

## PASSWORDACCESS = 生成环境

### 计算机 2

尚未设置第 2 台计算机 dps。在 dps 上对 zample 数据库的 db2adutl 查询返回了下列结果：

```
dps:/home/regress9/sql1lib/adsm> db2adutl query db zample
--- 数据库目录是空的 ---
警告: 在 ADSM 服务器上 DB2 没有创建文件空间
警告: 在 ADSM 中找不到任何别名的 DB2 备份映像。

dps:/home/regress9/sql1lib/adsm> db2adutl query db zample nodename
bar owner roecken
--- 数据库目录是空的 ---
```

对 ZAMPLE 数据库的查询

```
正在检索 FULL DATABASE BACKUP 信息。
1 时间: 20040216151025 最早的日志: S0000000.LOG 数据库分区号: 0
会话数: 1
正在检索 INCREMENTAL DATABASE BACKUP 信息。
找不到 ZAMPLE 的 INCREMENTAL DATABASE BACKUP 映像
```

```
正在检索 DELTA DATABASE BACKUP 信息。
找不到 ZAMPLE 的 DELTA DATABASE BACKUP 映像
```

```
正在检索 TABLESPACE BACKUP 信息。
找不到 ZAMPLE 的 TABLESPACE BACKUP 映像
```

```
正在检索 INCREMENTAL TABLESPACE BACKUP 信息。
找不到 ZAMPLE 的 INCREMENTAL TABLESPACE BACKUP 映像
```

```
正在检索 DELTA TABLESPACE BACKUP 信息。
找不到 ZAMPLE 的 DELTA TABLESPACE BACKUP 映像
```

```
正在检索 LOAD COPY 信息。
1 时间: 20040216151213
```

```
正在检索 LOG ARCHIVE 信息。
日志文件: S0000000.LOG, 链号: 0, 数据库分区号: 0
生成时间: 2004-02-16-15.10.38
```

在 dps 计算机上尚不存在 zample 数据库。

1. 将 zample 数据库复原至 dps 计算机:

```
dps:/home/regress9> db2 restore db zample use tsm options
''-fromnode=bar -fromowner=roecken'' without prompting
```

将返回以下信息:

```
DB20000I 已成功完成 RESTORE DATABASE 命令。
```

**注:** 如果 dps 上已经存在 zample 数据库, 那么将省略 `OPTIONS` 参数, 并且将使用数据库配置参数 `vendoropt`。此配置参数会覆盖备份或复原操作的 `OPTIONS` 参数。

对 zample 数据库的前滚操作将失败, 原因是 `rollforward` 实用程序找不到日志文件。前滚操作如下所示:

```
dps:/home/regress9> db2 rollforward db zample to end of logs and stop
```

将返回以下错误:

SQL4970N 数据库"ZAMPLE"上的前滚恢复不能达到指定的停止点(日志结束或时间点),原因是在节点"0"上丢失了日志文件。

2. 要强制 rollforward 实用程序查找另一台机器上的日志文件,必须配置正确的 logarchopt 值,在此情况下,为 logarchopt1 数据库配置参数:

```
dps:/home/regress9> db2 update db cfg for zample using logarchopt1
''-fromnode=bar -fromowner=roecken''
```

3. 要使 rollforward 实用程序能够使用装入副本映像,还必须设置 vendoropt 数据库配置参数:

```
dps:/home/regress9> db2 update db cfg for zample using VENDOROPT
''-fromnode=bar -fromowner=roecken''
```

4. 现在,可以对 zample 数据库进行前滚了:

```
dps:/home/regress9> db2 rollforward db zample to end of logs and stop
```

将返回以下信息:

| 前滚状态        |                                  |
|-------------|----------------------------------|
| 输入数据库别名     | = zample                         |
| 返回了状态的节点数   | = 1                              |
| 节点号         | = 0                              |
| 前滚状态        | = 未暂挂                            |
| 要读取的下一个日志文件 | =                                |
| 已处理的日志文件    | = S0000000.LOG - S0000000.LOG    |
| 上次落实的事务     | = 2004-02-16-20.10.38.000000 UTC |

DB20000I 已成功完成 ROLLFORWARD 命令。

## PASSWORDACCESS = 提示环境

在“提示”环境中,需要额外的信息,尤其是创建了对象的机器的 TSM 节点名和密码。

对于 db2adutl,更新 dsm.sys 文件(在基于 Windows 的平台上称为 dsm.opt 文件)并将 NODENAME bar(因为 bar 是源计算机的名称)添加至服务器子句:

```
dps:/home/regress9/sql1lib/adsm> db2adutl query db zample nodename bar
owner roecken password *****
```

将返回以下信息:

对 ZAMPLE 数据库的查询

正在检索 FULL DATABASE BACKUP 信息。

1 时间: 20040216151025 最早的日志: S0000000.LOG 数据库分区号: 0  
会话数: 1

正在检索 INCREMENTAL DATABASE BACKUP 信息。

找不到 ZAMPLE 的 INCREMENTAL DATABASE BACKUP 映像

正在检索 DELTA DATABASE BACKUP 信息。

找不到 ZAMPLE 的 DELTA DATABASE BACKUP 映像

正在检索 TABLESPACE BACKUP 信息。

找不到 ZAMPLE 的 TABLESPACE BACKUP 映像

正在检索 INCREMENTAL TABLESPACE BACKUP 信息。

找不到 ZAMPLE 的 INCREMENTAL TABLESPACE BACKUP 映像

正在检索 DELTA TABLESPACE BACKUP 信息。  
找不到 ZAMPLE 的 DELTA TABLESPACE BACKUP 映像

正在检索 LOAD COPY 信息。  
1 时间: 20040216151213

正在检索 LOG ARCHIVE 信息。  
日志文件: S0000000.LOG, 链号: 0, 数据库分区号: 0  
生成时间: 2004-02-16-15.10.38

1. 如果该数据库不存在, 那么请创建空的 `zample` 数据库。如果 `zample` 数据库已存在, 那么可以跳过此步骤和用来更新数据库配置的下面两个步骤。

```
dps:/home/regress9> db2 create db zample
```

2. 更新 `zample` 数据库的数据库配置参数 `tsm_nodename`:

```
dps:/home/regress9> db2 update db cfg for zample using tsm_nodename bar
```

3. 更新 `zample` 数据库的数据库配置参数 `tsm_password`:

```
dps:/home/regress9> db2 update db cfg for zample using  
tsm_password *****
```

4. 复原 `zample` 数据库:

```
dps:/home/regress9> db2 restore db zample use tsm options  
"-fromnode=bar -fromowner=roecken'" without prompting
```

成功完成复原操作, 但是发出了一条警告:

SQL2540W 复原成功, 但是在以"无中断"方式进行处理时, 在"数据库复原"期间遇到了警告"2523"。

此时, `rollforward` 实用程序又找不到正确的日志文件:

```
dps:/home/regress9> db2 rollforward db zample to end of logs and stop
```

将返回以下错误消息:

SQL1268N 由于检索节点"0"上的数据库"ZAMPLE"的日志文件"S0000000.LOG"时发生错误"-2112880618", 前滚恢复已停止。

5. 因为数据库复原操作替换了数据库配置文件, 所以必须将 TSM 数据库配置值设置为正确的值。首先, 必须复位 `tsm_nodename` 配置参数:

```
dps:/home/regress9> db2 update db cfg for zample using tsm_nodename bar
```

6. 必须复位 `tsm_password` 数据库配置参数:

```
dps:/home/regress9> db2 update db cfg for zample using tsm_password *****
```

7. 必须复位 `logarchopt1` 数据库配置参数, 以便 `rollforward` 实用程序可以找到正确的日志文件:

```
dps:/home/regress9> db2 update db cfg for zample using logarchopt1  
"-fromnode=bar -fromowner=roecken'"
```

8. 还必须复位 `vendoropt` 数据库配置参数, 以便还可以使用装入恢复文件:

```
dps:/home/regress9> db2 update db cfg for zample using VENDOROPT  
"-fromnode=bar -fromowner=roecken'"
```

9. 当设置数据库配置参数时, 可以对数据库进行前滚:

```
dps:/home/regress9> db2 rollforward db zample to end of logs and stop
```

对 `zample` 数据库运行 `ROLLFORWARD QUERY STATUS` 命令将显示以下结果:



前滚状态

```
输入数据库别名           = zample
返回了状态的节点数       = 1

节点号                   = 0
前滚状态                 = 未暂挂
要读取的下一个日志文件   =
已处理的日志文件         = S0000000.LOG - S0000000.LOG
上次落实的事务           = 2004-02-16-20.10.38.000000 UTC
```

DB20000I 已成功完成 ROLLFORWARD 命令。

---

## 使分区数据库环境中的时钟同步

应该使所有数据库分区服务器的系统时钟保持相对同步，以确保数据库操作顺利进行以及正向可恢复性不受限制。数据库分区服务器之间的时差加上事务的任何潜在操作和通信延迟，应小于对 *max\_time\_diff*（节点之间的最大时差）数据库管理器配置参数指定的值。

为确保日志记录时间戳记反映分区数据库环境中的事务顺序，DB2 使用每台机器上的系统时钟作为日志记录时间戳记的基准。但是，如果将系统时钟设置得提前，就会自动将日志时钟设置得提前。虽然可以将系统时钟往后调，但是日志时钟却不能这样设置，它会保持相同的超前时间，直至系统时钟与此时间匹配为止。于是，这两个时钟便同步了。这意味着一个数据库节点上的短期系统时钟错误可能会对数据库日志时间戳记产生长期的影响。

例如，假定数据库分区服务器 A 上的系统时钟被错误地设置为 2005 年 11 月 7 日，而当前年份是 2003 年，并假定在该数据库分区服务器上的数据库分区中落实了更新事务之后更正了此错误。如果继续使用该数据库，并且随着时间的推移定期对其进行更新，那么 2003 年 11 月 7 日至 2005 年 11 月 7 日之间的任何时间点实际上是无法通过前滚恢复到达的。当数据库分区服务器 A 上的 COMMIT 完成时，数据库日志中的时间戳记被设置为 2005，而日志时钟会停留在 2005 年 11 月 7 日，直到系统时钟与此时间相匹配为止。如果尝试前滚到这个时间范围内的某个时间点，那么操作将在指定停止点后的第一个时间戳记（即 2003 年 11 月 7 日）处停止。

虽然 DB2 无法控制对系统时钟的更新，但是 *max\_time\_diff* 数据库管理器配置参数减少了发生此类问题的机会：

- 此参数的可配置值的范围是 1 分钟至 24 小时。
- 当对非目录分区发出第一个连接请求时，数据库分区服务器会将它的时间发送至该数据库的目录分区。该目录分区就会检查请求连接的数据库分区上的时间与它自己的时间是否在 *max\_time\_diff* 参数指定的范围之内。如果超出此范围，那么拒绝该连接。
- 涉及到数据库中两个以上数据库分区服务器的更新事务，必须先验证参与的数据库分区服务器上的时钟是否同步，然后才可落实该更新。如果两个或更多个数据库分区服务器的时差超出 *max\_time\_diff* 允许的限制，那么会回滚该事务，以防止将不正确的传播至其他数据库分区服务器。



---

## 第 19 章 故障诊断

---

### DB2 故障诊断

通常，故障诊断要求您先找出并确定问题，然后再寻求解决方法。本节将提供关于 DB2 产品特定功能的故障诊断信息。

对于已知常见问题，找出的解决方法在本节以核对表的形式提供。如果您无法在该核对表中找到解决方法，那么可以自己收集其他诊断数据并进行分析，或者将该数据提交给 IBM 软件支持机构进行分析。

下列问题将引导您执行相应的故障诊断任务：

1. 是否已经应用了所有已知修订包？如果尚未应用，请考虑《DB2 服务器快速入门》中的“应用修订包”。
2. 是否在以下情况下出现问题：
  - 安装 DB2 数据库服务器或客户机时？如果是，请参阅本书其他位置处的“收集有关安装问题的数据”主题。
  - 创建、删除、更新或迁移实例或 DB2 管理服务器（DAS）时？如果是，请参阅本书其他位置处的“收集有关 DAS 和实例管理问题的数据”主题。
  - 使用 EXPORT、IMPORT、LOAD 或 db2move 命令移动数据时？如果是，请参阅本书其他位置处的“收集有关数据移动问题的数据”主题。

如果您的问题不属于上面任一类别并且要与 IBM 软件支持机构联系，那么可能仍需要基本的诊断数据。应。参阅本书其他位置处的“收集有关 DB2 的数据”主题。

---

### 诊断分区数据库环境

#### 在分区数据库环境中发出命令

在分区数据库环境中，您可能想要发出将在实例中的计算机或在多个数据库分区服务器（节点）上运行的命令。为此，您可以使用 rah 命令或 db2\_all 命令。rah 命令允许您发出将在实例中的计算机上运行的命令。

如果想要命令在实例中的多个数据库分区服务器上运行，那么运行 db2\_all 命令。本章节概述了这些命令。以下信息仅适用于分区数据库环境。

注：

1. 在 Linux 和 UNIX 平台上，您的登录 shell 程序可以是 Korn shell 程序或任何其他 shell；但是，不同 shell 处理包含特殊字符的命令所用的方式不同。
2. 另外，在 Linux 和 UNIX 平台上，rah 使用 DB2RSHCMD 注册表变量指定的远程 shell 程序。可以在两个远程 shell 程序之间选择：ssh（用于更高的安全性要求）或 rsh（或者对于 HP-UX 是 remsh）。ssh 远程 shell 程序用来防止在 UNIX 操作系统环境中以明文形式传输密码。如果不设置此注册表变量，那么使用 rsh（或者对于 HP-UX 是 remsh）。

3. 在 Windows 上，要运行 `rah` 命令或 `db2_all` 命令，您必须使用 Administrators 组成员的用户帐户来登录。

要确定命令的范围，请参阅 *Command Reference*，它指示命令是在单个的还是所有的数据库分区服务器上运行。如果命令在一个数据库分区服务器上运行，而您想该命令在其所有上面运行，那么使用 `db2_all`。例外情况是 `db2trc` 命令，它在一台计算机的所有逻辑节点（数据库分区服务器）上运行。如果想要在所有计算机的所有逻辑节点上运行 `db2trc`，那么使用 `rah`。

---

## 第 4 部分 性能问题



---

## 第 20 章 数据库设计中的性能问题

---

### 性能增强功能

#### 表分区和多维集群表

一个表可以同时是多维集群表和分区表。在同时是多维集群表和分区表的表中，可以同时使用表分区的范围分区规范和 MDC 键中使用列。与只单独使用多维集群或分区功能相比，这更有助于获取较详细的数据分区和块消除。在许多应用中，将 MDC 键列指定为不同于对表进行分区的列很有用。应该注意的是，表分区是多列，而 MDC 是多维。

#### 主流 DB2 V9.1 数据仓库的特征

下列建议主要针对对于 DB2 V9.1 来说是新产品的典型主流仓库。假定下列特征：

- 数据库在多台机器或多个 AIX 逻辑分区上运行。
- 使用数据库分区功能 (DPF) (表是使用 DISTRIBUTE BY HASH 子句创建的)。
- 有 4 到 50 个数据分区。
- 考虑其 MDC 和表分区的表是主要事实表。
- 表包含 1 亿到 1000 亿行。
- 在各种时间范围装入新数据：每夜、每周和每月。
- 每日接受量为 1 万到 1 亿条记录。
- 数据量变化：最多的一个月是最少的月的 5 倍。同样，最大维数（生产线，区域）具有 5 倍大小范围。
- 获取 1 到 5 年的详细数据。
- 每月或每个季度转出到期数据。
- 表使用大范围的查询类型。但是，相对于 OLTP 工作负载来说，该工作负载通常是具有下列特征的分析查询：
  - 较大的结果集，最多有 2 百万行
  - 大多数或全部查询都命中视图，而不是基本表
- SQL 子句按范围 (BETWEEN 子句)、列表中的项等选择数据。

#### 主流 DB2 V9.1 数据仓库事实表的特征

一个典型仓库事实表可能采用以下设计：

- 在 Month 列中创建数据分区。
- 为转出的每个时间段（例如，1 个月和 3 个月）定义数据分区。
- 在 Day 和 1 到 4 个其他维的基础上创建 MDC 维。典型的维有：生产线和区域。
- 所有数据分区和 MDC 集群都分布在所有数据库分区中。

MDC 和表分区具有一些相同的好处。下表列示组织中的潜在需求并根据先前确定的特征确定建议的组织方案。

表 14. 将表分区与 MDC 表配合使用

| 问题  | 建议的方案    | 建议  |
|---|----------|---|
| 转出期间的数据可用性  | 表分区      | 使用 DETACH PARTITION 子句来转出大量数据，并且只出现最少中断。  |
| 查询性能  | 表分区和 MDC | MDC 最适合用来查询多个维。表分区通过数据分区消除提高性能。   |
| 最少重组  | MDC      | MDC 维护集群，从而减少进行重组的必要性。  |
| 在传统脱机窗口中转出一个<br>月或更长时间的数据                               | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。  |
| 在短时间脱机窗口（小于 1 分钟）<br>期间转出一个月或更长时间的数据                    | 表分区      | 数据分区可以完全解决此需求。MDC 不起任何作用，并且仅 MDC 并不适合。  |
| 转出一个月或更长时间的数据，<br>并同时在不损失任何服务的情况下使表<br>对于提交查询的企业用户完全可用。 | MDC      | MDC 只能解决一部分此需求。由于表处于脱机状态的时间段太短，表分区并不适合。   |
| 每天装入数据（联机或脱机）   | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。   |
| “连续”装入数据（联机）  | 表分区和 MDC | 此时 MDC 具有很多好处。表分区具有递增的好处。   |
| “传统 BI”查询的查询执行性能  | 表分区和 MDC | MDC 特别适合用来查询立方体/多个维。表分区通过分区消除提高性能。  |
| 通过消除进行重组的必要性或减少执行任务所产生的不良影响，<br>使重组所带来的不良影响降到最低。        | MDC      | MDC 维护集群，从而减少进行重组的必要性。如果使用 MDC，那么数据分区不提供递增好处。但是，如果不使用 MDC，那么范围分区通过在分区级别维护一些粗粒度集群，有助于减少重组的必要性。 |

示例 1:

考虑一个具有键列 YearAndMonth 和 Province 的表。一种合理的规划此表方法是按日期进行分区，每 2 个月添加一个数据分区。此外，还可以按 Province 进行组织，以便任何两个月日期范围内的特定省份的所有行集群在一起，如第 28 页的图 6 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (Province);
```



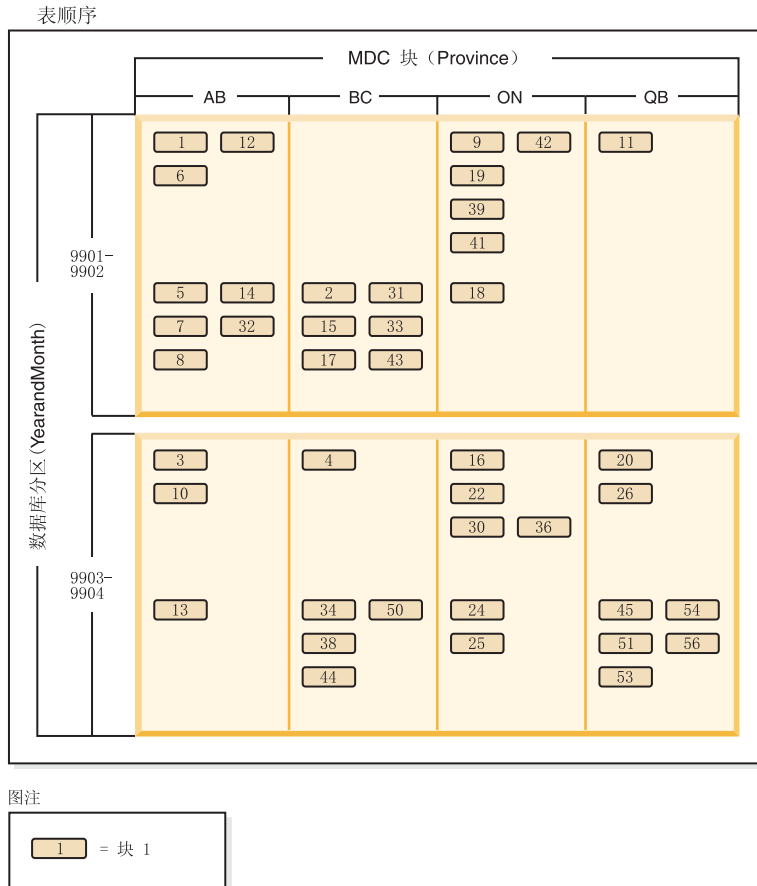


图 46. 按 YearAndMonth 分区并按 Province 组织的表

示例 2:

通过将 YearAndMonth 添加至 ORGANIZE BY 子句, 可以获得较高的详细程度, 如第 29 页的图 7 中所示。

```
CREATE TABLE orders (YearAndMonth INT, Province CHAR(2))
PARTITION BY RANGE (YearAndMonth)
(STARTING 9901 ENDING 9904 EVERY 2)
ORGANIZE BY (YearAndMonth, Province);
```

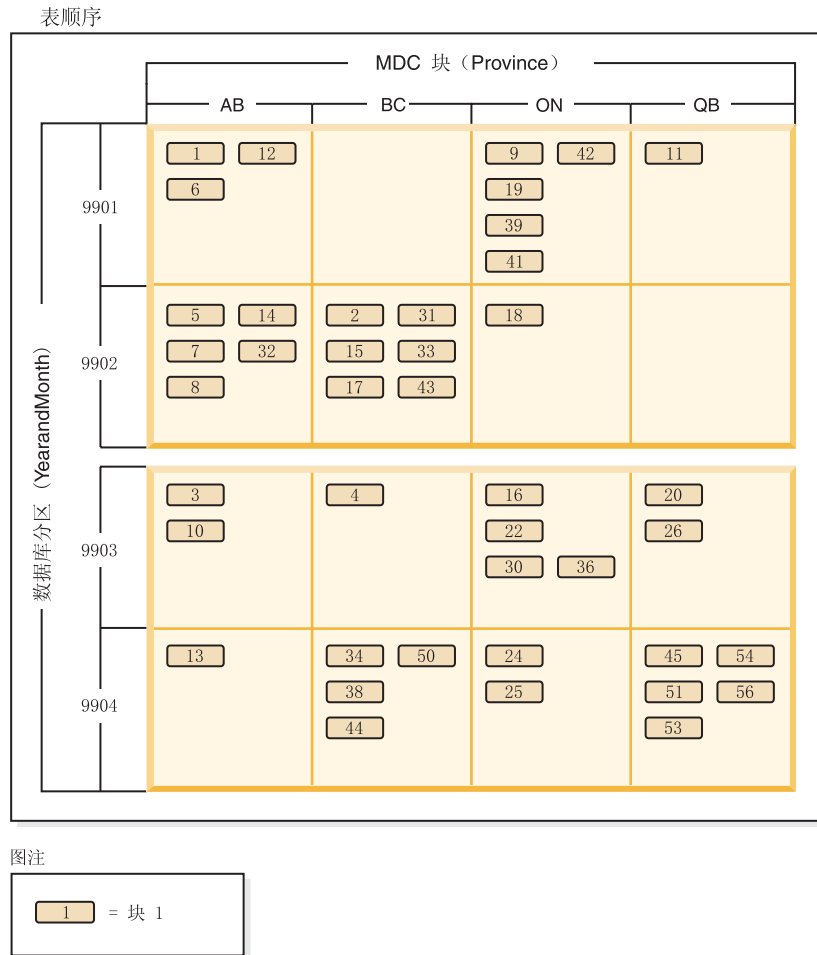


图 47. 按 YearAndMonth 分区并按 Province 和 YearAndMonth 组织的表

如果分区导致每个范围内只有单个值，那么通过在 MDC 键中包括表分区列不能获得任何好处。

### 注意事项

- 与基本表相比，MDC 表和分区表都需要一些存储器。这些存储器需求是附加的，但相对于所带来的好处来说，它们是合理的。
- 如果选择不在分区数据库环境中组合表分区和 MDC 功能，那么在您可以非常自信地预计数据分发情况时（通常也就是此处讨论的系统类型情况），使用表分区是最好的选择。否则，应考虑 MDC。

## 分区表的优化策略

数据分区消除指的是数据库服务器根据查询谓词确定只需要访问表的一部分数据分区就可以实现查询的能力。当对分区表运行决策支持查询时，数据分区消除可以提供特定好处。

分区表使用了数据组织方案，即，表数据根据该表中一个或多个表分区键列中的值分布到多个存储对象（称为数据分区或范围）中。根据 CREATE TABLE 语句的 PAR-

TITION BY 子句中指定的内容，给定表的数据被划分到多个存储对象中。这些存储对象可以在不同的表空间中，也可以在相同表空间中。

以下示例演示了数据分区消除所产生的性能方面的好处。如果发出以下语句：

```
CREATE TABLE custlist(subsdate DATE, Province CHAR(2), AccountID INT)
PARTITION BY RANGE(subsdate)
(STARTING FROM '1/1/1990' IN ts1,
 STARTING FROM '1/1/1991' IN ts1,
 STARTING FROM '1/1/1992' IN ts1,
 STARTING FROM '1/1/1993' IN ts2,
 STARTING FROM '1/1/1994' IN ts2,
 STARTING FROM '1/1/1995' IN ts2,
 STARTING FROM '1/1/1996' IN ts3,
 STARTING FROM '1/1/1997' IN ts3,
 STARTING FROM '1/1/1998' IN ts3,
 STARTING FROM '1/1/1999' IN ts4,
 STARTING FROM '1/1/2000' IN ts4,
 STARTING FROM '1/1/2001' ENDING '12/31/2001' IN ts4);
```

假定您对 2000 年的客户信息感兴趣。如果发出以下查询：

```
SELECT * FROM custlist WHERE subsdate BETWEEN '1/1/2000' AND '12/31/2000';
```

正如图 48 所显示的那样，数据库服务器确定只需要访问表空间 4 (ts4) 中的一个数据分区就可以解决此查询。

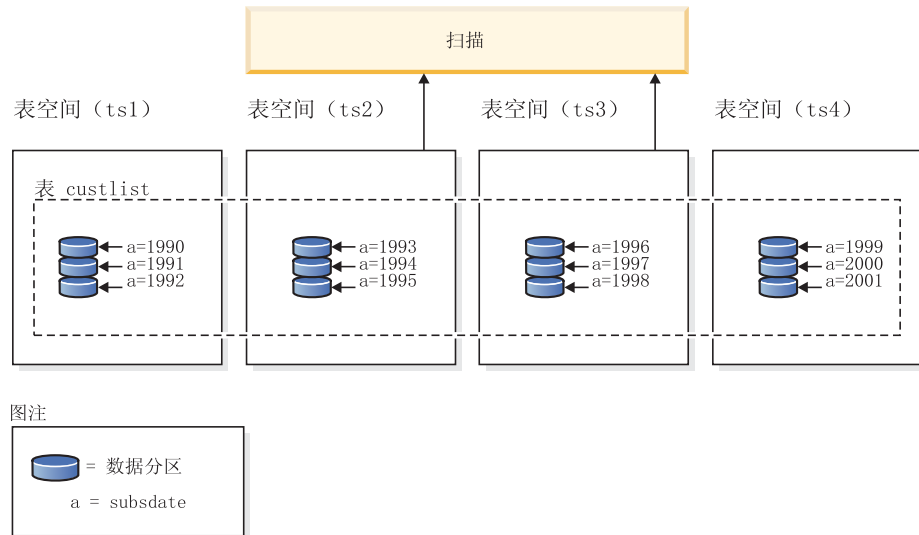


图 48. 分区表上数据分区消除所产生的性能方面的好处

图 246 页的图 49 中显示的另一个数据分区消除示例是索引扫描，它涉及两个索引并根据以下方案进行扫描：

```
CREATE TABLE multi (sale_date date, region char(2))
PARTITION BY (sale_date)
(STARTING '01/01/2005' ENDING '12/31/2005' EVERY 1 MONTH);
CREATE INDEX sx ON multi(sale_date);
CREATE INDEX rx ON multi(region);
```

如果发出以下查询：

```
SELECT * FROM multi WHERE
sale_date BETWEEN '6/1/2005' AND '7/31/2005' AND REGION = 'NW';
```

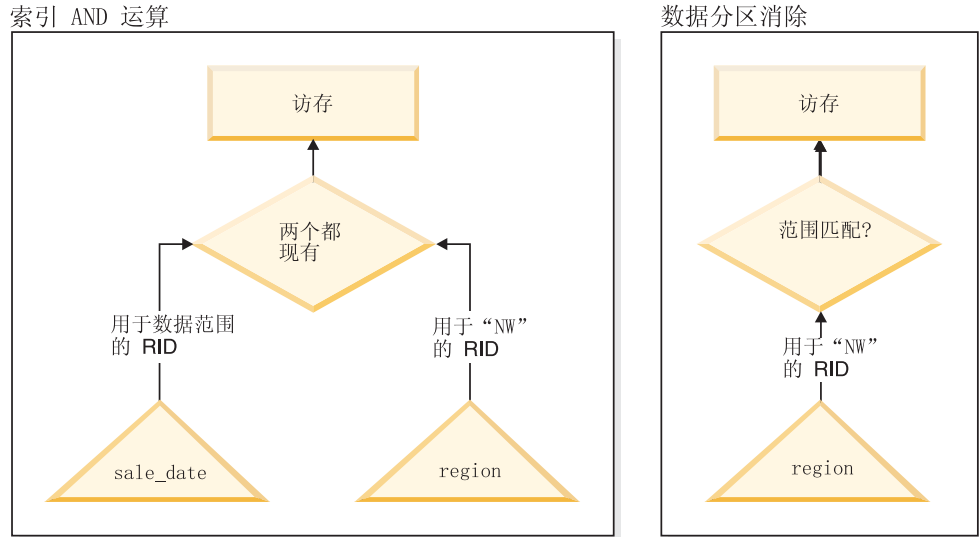


图 49. 表分区和索引“与” (AND) 的优化器决策路径

在不使用表分区时，一种可能的方案是索引“与” (AND)。索引“与” (AND) 执行下列任务：

- 读取每个索引中的所有相关索引条目
- 保存两组行标识 (RID)
- 匹配 RID 以确定哪些 RID 同时出现在这两个索引中
- 使用 RID 来访问行

如图 49 中所示，在使用表分区的情况下，读取索引以查找 region 和 sale\_date 的匹配项，从而允许快速检索匹配行。

## DB2 Explain

还可以使用 DB2 说明来确定 DB2 优化器选择的分区消除。DP Elim Predicates 信息显示扫描了哪些数据分区来解决以下查询：

```
SELECT * FROM custlist WHERE subsdate
BETWEEN '12/31/1999' AND '1/1/2001'
```

Arguments:

```
-----
DPESTFLG: (Number of data partitions accessed are Estimated)
          FALSE
DPLSTPRT: (List of data partitions accessed)
          9-11
DPNUMPRT: (Number of data partitions accessed)
          3
```

DP Elim Predicates:

```
-----
Range 1)
  Stop Predicate: (Q1.A <= '01/01/2001')
  Start Predicate: ('12/31/1999' <= Q1.A)
```

Objects Used in Access Plan:

```
-----  
Schema: MRSRINI  
Name: CUSTLIST  
Type: Data Partitioned Table  
Time of creation: 2005-11-30-14.21.33.857039  
Last statistics update: 2005-11-30-14.21.34.339392  
Number of columns: 3  
Number of rows: 100000  
Width of rows: 19  
Number of buffer pool pages: 1200  
Number of data partitions: 12  
Distinct row values: No  
Tablespace name: <VARIOUS>
```

## 多列支持

在使用多个列作为表分区键的情况下，数据分区消除将起作用。

例如，如果发出以下语句：

```
CREATE TABLE sales (year INT, month INT)  
PARTITION BY RANGE(year, month)  
(STARTING FROM (2001, 1) ENDING AT(2001,3) IN ts1,  
ENDING AT(2001,6) IN ts2,  
ENDING AT(2001,9) IN ts3,  
ENDING AT(2001,12) IN ts4,  
ENDING AT(2002,3) IN ts5,  
ENDING AT(2002,6) IN ts6,  
ENDING AT(2002,9) IN ts7,  
ENDING AT(2002,12) IN ts8)
```

接着，发出以下查询：

```
SELECT * FROM sales WHERE year = 2001 AND month < 8
```

查询优化器推断只需要访问 ts1、ts2 和 ts3 中的数据分区就可以解决此查询。

**注：**在多个列组成表分区键的情况下，只有当拥有组合键的前导列上的谓词时才能实现数据分区消除，因为用于表分区键的非前导列不是独立的。

## 多范围支持

可以使用多个范围在数据分区上实现数据分区消除（即，一起执行“或”（OR）运算）。通过使用上一个示例中创建的表，执行下列查询：

```
SELECT * FROM sales  
WHERE (year = 2001 AND month <= 3) OR (year = 2002 and month >= 10)
```

数据库服务器只访问 2001 年的第一季度和 2002 年的最后一个季度的数据。

## 生成列

可以将生成列用作表分区键。

例如，可以发出以下语句：

```
CREATE TABLE sales(a INT, b INT GENERATED ALWAYS AS (a / 5))
  IN ts1,ts2,ts3,ts4,ts5,ts6,ts7,ts8,ts9,ts10
  PARTITION BY RANGE(b)
  (STARTING FROM (0) ENDING AT(1000) EVERY (50))
```

在此示例中，将生成列上的谓词用于数据分区消除。此外，当用来生成列的表达式是单调的时，数据库服务器会将源列上的谓词转换为生成列上的谓词，从而在生成列上启用数据分区消除。

例如，如果具有以下查询：

```
SELECT * FROM sales WHERE a > 35
```

数据库服务器根据 (a > 35) 在 b (b > 7) 上生成额外谓词，从而允许数据分区消除。

## 连接谓词

如果将连接谓词下推到表访问级别，那么也可以在数据分区消除中使用连接谓词。连接谓词仅在嵌套循环连接 (NLJN) 的内部才下推到表访问级别。

请考虑下列表：

```
CREATE TABLE T1(A INT, B INT)
  PARTITION BY RANGE(A, B)
  (STARTING FROM (1, 1)
  ENDING (1,10) IN ts1, ENDING (1,20) IN ts2,
  ENDING (2,10) IN ts3, ENDING (2,20) IN ts4,
  ENDING (3,10) IN ts5, ENDING (3,20) IN ts6,
  ENDING (4,10) IN ts7, ENDING (4,20) IN ts8)
```

```
CREATE TABLE T2 (A INT, B INT)
```

使用的谓词：

```
P1: T1.A = T2.A
P2: T1.B > 15
```

在此示例中，由于不知道连接的外值，因此不能确定将在编译时访问的额外数据分区。在这种情况下，以及在使用主变量或参数标记的情况下，当绑定必需的值时，就会发生数据分区消除。

在运行时，当 T1 是 NLJN 的内部表时，会根据 T2.A 的每个外值的谓词自动进行数据分区消除。在运行时，对外值 T2.A = 3 应用谓词 T1.A = 3 和 T1.B > 15，这样就限定了访问的表空间 ts6 和 ts7 中的数据分区。

考虑表 T1 和 T2 中的列 A 具有下列值：

| 外部表 T2: 列 A | 内部表 T1: 列 A | 内部表 T1: 列 B | 内部表 T1: 数据分区位置 |
|-------------|-------------|-------------|----------------|
| 2           | 3           | 20          | ts6            |
| 3           | 2           | 10          | ts3            |
| 3           | 2           | 18          | ts4            |
|             | 3           | 15          | ts6            |
|             | 1           | 40          | ts3            |

要执行嵌套循环连接（假定对内部表进行表扫描），数据库管理器执行下列步骤：

1. 读取 T2 中的第一行。A 的值是 2。

2. 在连接谓词  $T1.A = T2.A$  中将  $T2.A$  值（它是 2）绑定到列  $T2.A$ 。该谓词就变成  $T1.A = 2$ 。
3. 使用谓词  $T1.A = 2$  和  $T1.B > 15$  应用数据分区消除。这将限定表空间  $ts4$  和  $ts5$  中的数据分区。
4. 在应用  $T1.A = 2$  和  $T1.B > 15$  之后，扫描表  $T1$  的表空间  $ts4$  和  $ts5$  中的数据分区，直到找到一行为止。找到的第一个合格行是  $T1$  的行 3。
5. 连接匹配的行。
6. 扫描表  $T1$  的表空间  $ts4$  和  $ts5$  中的数据分区，直到找到下一个匹配项（ $T1.A = 2$  和  $T1.B > 15$ ）为止。再也找不到其他行。
7. 对  $T2$  的下一行（将  $A$  的值替换为 3）重复步骤 1 至 6，直到用完  $T2$  中的所有行为止。

## MDC 表的优化策略

如果创建多维集群（MDC）表，因为优化器可以应用附加的优化策略，可以提高许多查询的性能。这些策略主要基于改进的块索引效率，但在多维上进行集群的优点也允许更快速的数据检索。

**注：**MDC 表优化策略也可以实现分区内并行性和分区间并行性的性能优点。

考虑 MDC 表的下列特定优点：

- 维块索引查找可以标识表的所需部分，并仅快速扫描所需的块。
- 因为块索引小于 RID 索引，所以查找更快。
- 可以在块级别执行索引 AND 和 OR 运算，并可以将这些运算与 RID 组合在一起。
- 保证在扩展数据块内集群数据，这使得检索更快。
- 当可以使用转出时，可以更快地删除行。

考虑在 **region** 和 **month** 列上定义了维且命名为 SALES 的 MDC 表的以下简单示例：

```
SELECT * FROM SALES
      WHERE MONTH='March' AND REGION='SE'
```

对于此查询，优化器可以执行维块索引查找来寻找出现在月份为三月且地区为 SE 的块。然后该查询可以仅快速扫描表的结果块来访问结果集。

## 转出删除

当满足允许使用转出以进行删除的条件时，就会使用删除 MDC 表中的行的更有效方法。这些条件是：

- 搜索而不是定位 DELETE 语句（即，不使用“WHERE CURRENT OF”子句）。
- 没有 WHERE 子句（将删除所有行）或 WHERE 子句中仅有的条件是针对维的。
- 未使用 DATA CAPTURE CHANGES 子句来定义表。
- 表在引用完整性关系中不是父表。
- 表没有定义“删除时”触发器。
- 未在任何立即刷新的 MQT 中使用表。
- 如果级联删除操作的外键是其表的维列的子集，那么它可能适合于转出。

- 在触发由 CREATE TRIGGER 语句上的 OLD TABLE AS 子句指定的 SQL 操作之前，DELETE 语句不能出现在对用于标识受影响行集的临时表执行的 SELECT 语句中。

对于转出删除，不会记录被删除的记录。而是会通过重新格式化页的某些部分来使包含这些记录的页看上去是空的。将会记录对重新格式化的部分所做的更改，但不会记录这些记录本身。

立即清除转出这一缺省行为就是在删除时就清除 RID 索引。还可以通过将注册表变量 **DB2\_MDC\_ROLLOUT** 设置为 IMMEDIATE，或者通过对 SET CURRENT MDC ROLLOUT MODE 语句指定 IMMEDIATE 来指定此方式。与标准删除相比，索引更新的日志记录没有变化，因此，性能提高取决于有多少个 RID 索引。RID 索引越少，性能就越好，衡量方法是总时间与日志空间的百分比。

可以通过以下公式来估计日志中剩余的空间量，其中 N 是已删除的记录数，S 是已删除的记录的总大小（包括诸如空指示符和 varchar 长度的开销），P 是包含已删除的记录的块中的页数：

$$S + 38*N - 50*P$$

此图是实际日志数据的简化形式。由于节省了为进行回滚而保留的空间，因此，活动日志空间要求将减小一半。

另外，在落实事务之后，可以使用延迟清除转出来更新 RID 索引。还可以通过将注册表变量 **DB2\_MDC\_ROLL\_OUT** 设置为 DEFER，或者通过对 SET CURRENT MDC ROLLOUT MODE 语句指定 DEFERRED 来指定此方式。在延迟转出方式下，将在落实删除之后在后台异步清除 RID 索引。对于非常大型的删除任务或者表中存在大量 RID 索引时，使用这种转出方法可以非常快速地进行删除。整个清除操作的速度也提高了，这是因为在执行延迟索引清除时将并行清除索引，而在执行立即索引清除时将逐行清除索引中的每一行。另外，DELETE 语句的事务日志空间需求显著降低，这是因为索引按索引页而不是按索引键来更新异步索引清除日志。

**注：**延迟清除转出需要更多内存资源，内存资源是从数据库堆中获取的。如果 DB2 无法分配它需要的内存结构，那么延迟清除转出将失败，并将一条消息写入管理员日志中。

## 何时使用延迟清除转出

如果对于您来说删除性能是最重要的因素，并且已经对表定义了 RID 索引，那么就应使用延迟清除转出。注意，在进行索引清除之前，对已转出的块进行基于索引的扫描会稍微降低性能，这取决于已转出的数据量。下面是在决定执行立即索引清除和延迟索引清除时应考虑的其他问题：

- 删除工作量大小：对于非常大型的删除任务选择延迟清除转出。在对许多小型 MDC 表频繁发出维删除语句的情况下，异步清除索引对象所产生的开销在价值上要超过删除期间节省的时间所带来的好处。
- 索引的数量和类型：如果表中包含大量 RID 索引，且需要对这些索引执行行级别处理，那么应使用延迟清除转出。
- 块可用性：如果您希望由 DELETE 语句释放的块空间在落实 DELETE 语句之后立即可用，那么应使用立即清除转出。
- 日志空间：如果日志空间有限，那么应对大型删除任务使用延迟清除转出。



- 内存约束: 延迟清除转出将在所有具有延迟清除暂挂的表上消耗更多数据库堆。

要在执行删除时禁止转出行为, 可以将注册表变量 **DB2\_MDC\_ROLLOUT** 设置为 OFF, 或者对 SET CURRENT MDC ROLLOUT MODE 语句指定 NONE。



---

## 第 21 章 索引

---

### 分区表中的索引

#### 了解分区表上的索引行为

分区表上的索引与普通表上的索引类似，即每个索引都包含指向表的所有数据分区中的行的指针。然而，一个主要的区别是分区表上的每个索引都是一个独立的对象。在分区数据库环境中，索引采用与表相同的方式分布在各数据库分区上。因为关于分区表的索引可以独立于其他索引操作，所以在分区表上创建索引时需要注意一些关于使用哪个表空间的特殊注意事项。

即使分区表的数据分区跨多个表空间，也只在单个表空间中创建该分区表的索引。DMS 和 SMS 表空间都支持在不同于表的另一个位置使用索引。所有指定的表空间必须位于同一数据库分区组中。可以将每个索引放置在它自己的表空间中，包括大型表空间。每个索引表空间都必须使用与数据分区相同的存储机制 DMS 或 SMS。大型表空间中的索引最多可以包含  $2^{29}$  页。

分区表上的索引的其他好处包括：

- 提高了删除索引和联机索引创建的性能
- 能够针对每个表索引之间的任何表空间特征使用不同的值（例如，为了确保更好的空间利用率，对每个索引使用不同的页大小可能更合适）。
- 减少 IO 争用并提供对表索引数据更有效的并发访问。
- 删除各个索引时，空间将立即可供系统使用，而无需进行索引重组。
- 如果您选择执行索引重组，可以重组单个索引。

第 254 页的图 50 显示了位于单个表空间中的分区表上的非分区索引。

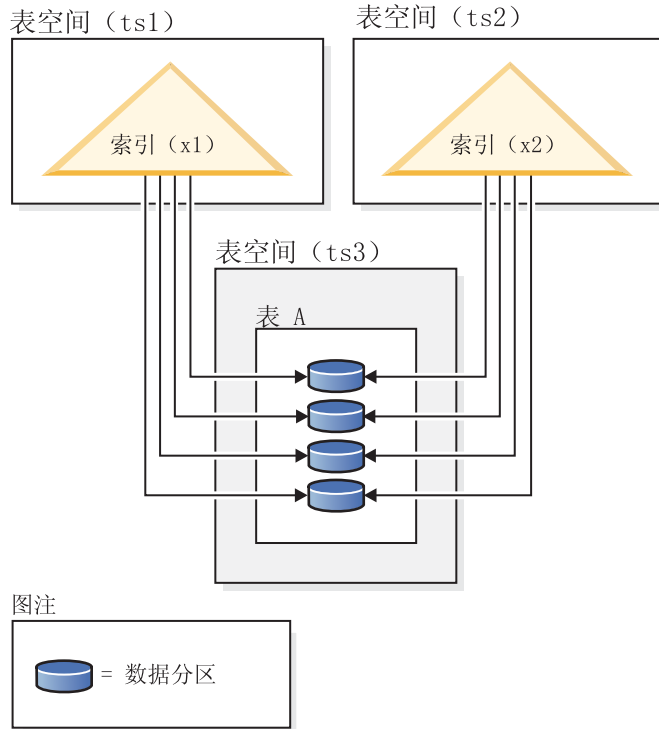


图 50. 分区表上的索引行为

第 255 页的图 51 显示了还分布在多个数据库分区上的分区表的索引行为。

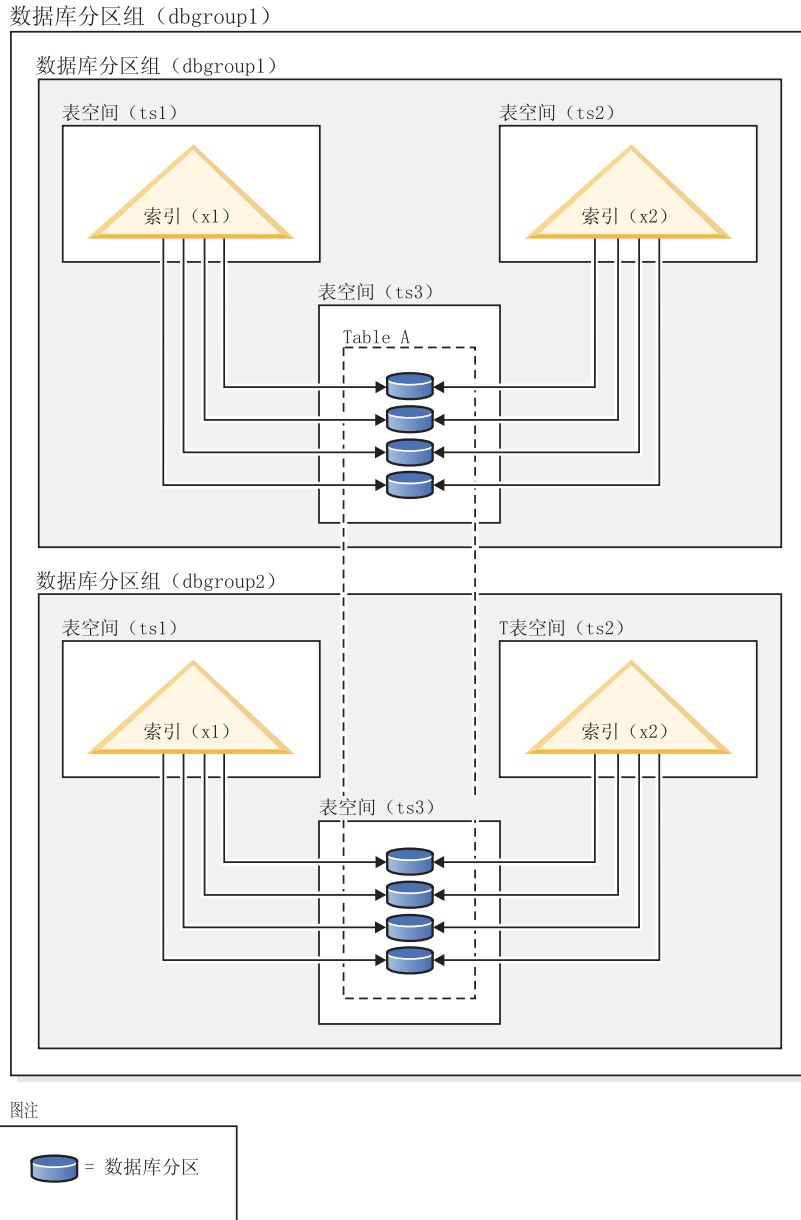


图 51. 分布式分区表的索引行为

可以使用 `CREATE INDEX ...IN <tspace1>` 语句为分区表指定索引表空间，这与使用 `CREATE TABLE .. INDEX IN <tspace2>` 语句指定的索引表空间不同。

只能对分区表使用 `CREATE INDEX` 语句中的 `IN` 子句来覆盖索引位置，这将允许您为索引指定表空间位置。此方法允许您根据需要将一个分区表上的不同索引放在不同表空间中。如果创建分区表时未指定放置它的非分区索引的位置，并且使用未指定特定表空间的 `CREATE INDEX` 语句来创建索引，那么将在第一个已连接的或可视数据分区的表空间中创建索引。按顺序对下面三种可能情况进行求值（从情况 1 开始），以确定创建索引的位置。当与其中一种情况相匹配时，求值将停止：

- 情况 1:  
当在 `CREATE INDEX ... IN <tspace1>` 语句中指定了索引表空间时，将 `<tspace1>` 中指定的表空间用于此索引。
- 情况 2:

当在 `CREATE TABLE .. INDEX IN <tbspace2>` 语句中指定了索引表空间中，将 `<tbspace2>` 中指定的表空间用于此索引。  
情况 3:  
未指定任何表空间时，使用第一个已连接的或可视数据分区使用的表空间。

创建索引的位置取决于在 `CREATE TABLE` 语句期间执行的操作。对于非分区表，如果未指定任何 `INDEX IN` 子句，那么数据库将在该子句中填写您的数据表空间。对于分区表，如果您将它留空，那么它将保留为空白，此时情况 3 适用。

示例 1: 此示例假定存在分区表 `sales (a int, b int, c int)`，并在表空间“ts1”中创建唯一索引“a\_idx”。

```
CREATE UNIQUE INDEX a_idx ON sales ( a ) IN ts1
```

示例 2: 此示例假定存在分区表 `sales (a int, b int, c int)`，并在表空间“ts2”中创建索引“b\_idx”。

```
CREATE INDEX b_idx ON sales ( b ) IN ts2
```

## 了解分区表上的集群索引行为

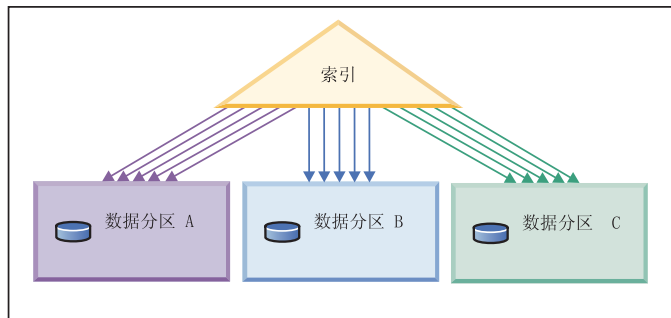
对分区表使用集群索引可获得与对常规表使用集群索引相同的好处。但是，在选择关于表分区键定义的集群索引时必须小心。

可以使用任何集群键在分区表上创建集群索引。数据库服务器尝试使用集群索引来以本地方式集群每个数据分区中的数据。在已划分集群的插入期间，将在索引中进行查询以找到合适的行标识 (RID)。将从此 RID 开始查找表中的空间以插入记录。要获得维护良好且性能优良的集群，索引列与表分区键列之间应该存在关联。确保这种关联的一种方法是使表分区键列成为索引列的前缀，如下面的示例中所示：

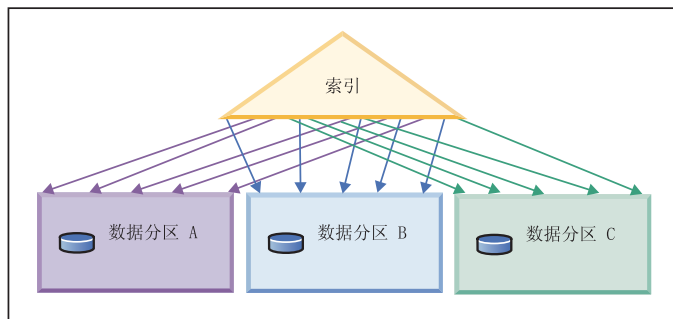
```
PARTITION BY RANGE (Month, Region)  
CREATE INDEX ... (Month, Region, Department) CLUSTER
```

虽然数据库服务器不强制此关联，但还是希望索引中的所有键按分区标识聚集在一起，以获得较好的集群。例如，如果一个表按季度进行分区，而集群索引按日期定义，由于季度与日期之间存在关联，所以可以获得具有良好性能的最佳数据集群，其中任何数据分区的所有键都聚集在索引中。

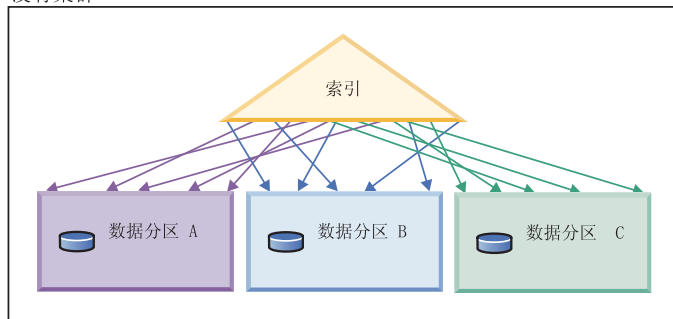
将分区键作为前缀的集群（相关）



集群与分区键不匹配（以本地方式集群）



没有集群



图注



图 52. 集群索引可能对分区表产生的影响。在第一个图中，数据同时以全局和本地方式集群。

如图 52 中所示，在给出每个示例中索引和数据的布局的情况下，当集群与表分区键相关时，可以获得最佳扫描性能。当集群与表分区键不相关时，不可能以本地方式集群索引。因为需要表分区列与索引列之间存在关联，所以根本不可能有一个完美的本地集群方案。

集群的好处包括：

- 在每个数据分区内，行以键的顺序排列。

- 集群索引改善了以键的顺序扫描整张表的性能。这是因为扫描访存第一页的第一行，然后访存同一页上的每一行，在访存了该页上的所有行之后，才移至下一页。这表示任何给定时间内都只需要表的一页位于缓冲池中。相反，如果表未集群，那么访存的每行有可能是不同页中的。除非缓冲池中有空间保存整个表，否则这会导致每页被访存多次，从而极大地减慢扫描速度。

对于分区表，只有当表分区键是集群键的前缀（请参阅第 257 页的图 52 中的第一个图）时，才能保证扫描期间出现每页只访存一次的理想情况。但是，如果像上面描述的那样集群键与表分区键不相关，并且数据是本地集群的，那么当缓冲池中有足够的空间来保存每个数据分区的一页时，仍然可以获得集群索引的全部好处。这是因为对给定数据分区访存的每一行都位于上一次对同一数据分区访存的行的附近。（请参阅第 257 页的图 52 的第二个图）。如前面提及的那样，如果集群键与表分区键不相关，那么不能很好地维护集群，但如果您不想对表进行高级插入、更新和删除活动，使用此方法应该可以受益。

即使缓冲池中并没有足够的空间来保存每个数据分区的一页，定义集群索引仍然可以获得一些好处。



---

## 第 22 章 设计顾问程序

---

### 使用设计顾问程序从单一分区迁移到多分区数据库

可以使用“设计顾问程序”来帮助您从单一分区迁移到多分区数据库。“设计顾问程序”可提供有关分布数据的建议，同时提供有关新索引、具体化查询表（MQT）和多维集群（MDC）表的建议。

1. 使用 `db2licm` 命令注册 DB2 产品或功能部件许可证密钥
2. 在多分区数据库分区组中至少创建一个表空间。

**注：**因为“设计顾问程序”只能建议对现有表空间再分发数据，所以在运行“设计顾问程序”之前，分区数据库中必须存在您想要“设计顾问程序”考虑的表空间。

3. 在使用在“设计顾问程序”GUI 中选择的数据库分区功能部件或者使用为 `db2advis` 命令指定的分区选项的情况下运行“设计顾问程序”。
4. 如果正在“控制中心”中使用“设计顾问程序”，那么可自动实现数据库分区建议。如果正在使用 `db2advis` 命令，那么在运行由“设计顾问程序”生成的 DDL 语句之前需要稍微修改 `db2advis` 输出文件。



## 第 23 章 管理并行性

### 表的锁定方式和 MDC 表的 RID 索引扫描

在多维集群（MDC）环境中，还将使用附加的锁定级别 BLOCK。以下表列示了不同访问方案在各个级别获得的锁定类型。每个条目由三部分组成：表锁定、块锁定和行锁定。虚线表示未使用某个特定级别的锁定。

注：可以使用 SELECT 语句的锁定请求子句来显式更改锁定方式。

表 15. 不使用谓词的表扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |          | 搜索的更新或删除 |        |
|------|----------|---------|----------|----------|--------|
|      |          | 扫描      | 当前位置     | 扫描或删除    | Update |
| RR   | S/-/     | U/-/    | SIX/IX/X | X/-/     | X/-/   |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/U  | IX/X/-   | IX/I/- |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X  | IX/X/-   | IX/X/- |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X  | IX/X/-   | IX/X/- |

表 16. 仅使用维列上的谓词的表扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |          | 搜索的更新或删除 |         |
|------|----------|---------|----------|----------|---------|
|      |          | 扫描      | 当前位置     | 扫描或删除    | Update  |
| RR   | S/-/     | U/-/    | SIX/IX/X | U/-/     | SIX/X/- |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X  | IX/U/-   | X/X/-   |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X  | IX/U/-   | X/X/-   |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X  | IX/U/-   | X/X/-   |

表 17. 使用其他谓词（sargs 和 resids）的表扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |          | 搜索的更新或删除 |          |
|------|----------|---------|----------|----------|----------|
|      |          | 扫描      | 当前位置     | 扫描或删除    | Update   |
| RR   | S/-/     | U/-/    | SIX/IX/X | U/-/     | SIX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X  | IX/IX/U  | IX/IX/X  |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X  | IX/IX/U  | IX/IX/X  |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X  | IX/IX/U  | IX/IX/X  |

以下两个表显示了 MDC 表的 RID 索引的锁定方式。

表 18. 不使用谓词的 RID 索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |        |
|------|----------|---------|---------|----------|--------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update |
| RR   | S/-/     | IX/IX/S | IX/IX/X | X/-/     | X/-/   |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |

表 18. 不使用谓词的 RID 索引扫描的锁定方式 (续)

| 隔离级别 | 只读和模糊扫描 | 游标操作    |         | 搜索的更新或删除 |        |
|------|---------|---------|---------|----------|--------|
|      |         | 扫描      | 当前位置    | 扫描删除     | Update |
| UR   | IN/IN/- | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |

表 19. 使用单个合格行的 RID 索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |        |
|------|----------|---------|---------|----------|--------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update |
| RR   | IS/IS/S  | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | X/X/X    | X/X/X  |

表 20. 仅使用 Start 和 Stop 谓词的 RID 索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IS/IS/S  | IX/IX/S | IX/IX/X | IX/IX/X  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |

表 21. 仅使用索引谓词的 RID 索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IS/S/S   | IX/IX/S | IX/IX/X | IX/IX/S  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

表 22. 使用其他谓词 (sargs 和 resids) 的 RID 索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IS/S/S   | IX/IX/S | IX/IX/X | IX/IX/S  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

注: 在下面的表中, 显示了用于延迟的数据页访问的 RID 索引扫描的锁定方式, 在 UR 隔离级别, 如果具有 1 类索引的 IN 锁定, 或者如果在索引的包含列中具有谓词, 那么隔离级别将升级至 CS 并且锁定将升级至 IS 表锁定、IS 块锁定和 NS 行锁定。

表 23. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 不使用谓词的 RID 索引扫描

| 隔离级别 | 只读和模糊扫描 | 游标操作    |      | 搜索的更新或删除 |        |
|------|---------|---------|------|----------|--------|
|      |         | 扫描      | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/S  | IX/IX/S |      | X/-/-    |        |
| RS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| CS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| UR   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |

表 24. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 在不使用谓词的 RID 索引扫描后延迟的数据页访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IN/IN/-  | IX/IX/S | IX/IX/X | X/-/-    | X/-/-   |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/X  | IX/IX/X |

表 25. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 使用谓词 (sargs 和 resids) 的 RID 索引扫描

| 隔离级别 | 只读和模糊扫描 | 游标操作    |      | 搜索的更新或删除 |        |
|------|---------|---------|------|----------|--------|
|      |         | 扫描      | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/-  | IX/IX/S |      | IX/IX/S  |        |
| RS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| CS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| UR   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |

表 26. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 在使用谓词 (sargs 和 resids) 的 RID 索引扫描后延迟的页数据访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IN/IN/-  | IX/IX/S | IX/IX/X | IX/IX/S  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

表 27. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 仅使用 Start 和 Stop 谓词的 RID 索引扫描

| 隔离级别 | 只读和模糊扫描 | 游标操作    |      | 搜索的更新或删除 |        |
|------|---------|---------|------|----------|--------|
|      |         | 扫描      | 当前位置 | 扫描删除     | Update |
| RR   | IS/IS/S | IX/IX/S |      | IX/IX/X  |        |
| RS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| CS   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |
| UR   | IN/IN/- | IN/IN/- |      | IN/IN/-  |        |

表 28. 延迟的数据页访问使用的 RID 索引扫描的锁定方式: 在仅使用 Start 和 Stop 谓词的 RID 索引扫描后延迟的页数据访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IN/IN/-  | IX/IX/S | IX/IX/X | IX/IX/X  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IS/-/-   | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

## MDC 表的块索引扫描的锁定

以下表列示了不同访问方案在各个级别获得的锁定类型。每个条目由三部分组成: 表锁定、块锁定和行锁定。虚线表示未进行某个特定级别的锁定。

注: 可以使用 SELECT 语句的锁定请求子句来显式更改锁定方式。

表 29. 不使用谓词的索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | S/--/--  | IX/IX/S | IX/IX/X | X/--/--  | X/--/-- |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/--   | X/X/--  |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/--   | X/X/--  |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | X/X/--   | X/X/--  |

表 30. 仅使用维谓词的索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |        |
|------|----------|---------|---------|----------|--------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update |
| RR   | IS/-/-   | IX/IX/S | IX/IX/X | X/-/-    | X/-/-  |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/X/-   | IX/X/- |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/X/-   | IX/X/- |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/X/-   | IX/X/- |

表 31. 仅使用 Start 和 Stop 谓词的索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描 | 游标操作    |         | 搜索的更新或删除 |         |
|------|---------|---------|---------|----------|---------|
|      |         | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IS/S/-  | IX/IX/S | IX/IX/S | IX/IX/S  | IX/IX/S |
| RS   | IX/IX/S | IX/IX/U | IX/IX/X | IX/IX/-  | IX/IX/- |
| CS   | IX/IX/S | IX/IX/U | IX/IX/X | IX/IX/-  | IX/IX/- |
| UR   | IN/IN/- | IX/IX/U | IX/IX/X | IX/IX/-  | IX/IX/- |

表 32. 使用谓词的索引扫描的锁定方式

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IS/S/-   | IX/IX/S | IX/IX/X | IX/IX/S  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IN/IN/-  | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

下表列示延迟的数据页访问所使用的块索引扫描的锁定方式:

表 33. 延迟的数据页访问使用的块索引扫描的锁定方式: 不使用谓词的块索引扫描

| 隔离级别 | 只读和模糊扫描  | 游标操作     |      | 搜索的更新或删除 |        |
|------|----------|----------|------|----------|--------|
|      |          | 扫描       | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/--  | IX/IX/S  |      | X/--     |        |
| RS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| CS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| UR   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |

表 34. 延迟的数据页访问使用的块索引扫描的锁定方式: 在不使用谓词的块索引扫描后延迟的数据页访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |        |
|------|----------|---------|---------|----------|--------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update |
| RR   | IN/IN/-- | IX/IX/S | IX/IX/X | X/--     | X/--   |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/--   | X/X/-- |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | X/X/--   | X/X/-- |
| UR   | IN/IN/-- | IX/IX/U | IX/IX/X | X/X/--   | X/X/-- |

表 35. 延迟的数据页访问使用的块索引扫描的锁定方式: 仅使用谓词的块索引扫描

| 隔离级别 | 只读和模糊扫描  | 游标操作     |      | 搜索的更新或删除 |        |
|------|----------|----------|------|----------|--------|
|      |          | 扫描       | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/--  | IX/IX/-- |      | IX/S/--  |        |
| RS   | IS/IS/NS | IX/--    |      | IX/--    |        |
| CS   | IS/IS/NS | IX/--    |      | IX/--    |        |
| UR   | IN/IN/-- | IX/--    |      | IX/--    |        |

表 36. 延迟的数据页访问使用的块索引扫描的锁定方式: 在仅使用谓词的块索引扫描后延迟的数据页访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IN/IN/-- | IX/IX/S | IX/IX/X | IX/S/--  | IX/X/-- |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/U/--  | IX/X/-- |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/U/--  | IX/X/-- |
| UR   | IN/IN/-- | IX/IX/U | IX/IX/X | IX/U/--  | IX/X/-- |

表 37. 延迟的数据页访问使用的块索引扫描的锁定方式: 仅使用 *Start* 和 *Stop* 谓词的索引扫描

| 隔离级别 | 只读和模糊扫描  | 游标操作     |      | 搜索的更新或删除 |        |
|------|----------|----------|------|----------|--------|
|      |          | 扫描       | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/--  | IX/IX/-- |      | IX/X/--  |        |
| RS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| CS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| UR   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |

表 38. 延迟的数据页访问使用的块索引扫描的锁定方式: 在仅使用 *Start* 和 *Stop* 谓词的块索引扫描后延迟的数据页访问

| 隔离级别 | 只读和模糊扫描  | 游标操作     |      | 搜索的更新或删除 |        |
|------|----------|----------|------|----------|--------|
|      |          | 扫描       | 当前位置 | 扫描删除     | Update |
| RR   | IN/IN/-- | IX/IX/X  |      | IX/X/--  |        |
| RS   | IS/IS/NS | IN/IN/-- |      | IN/IN/-- |        |
| CS   | IS/IS/NS | IN/IN/-- |      | IN/IN/-- |        |
| UR   | IS/--/-- | IN/IN/-- |      | IN/IN/-- |        |

表 39. 延迟的数据页访问使用的块索引扫描的锁定方式: 使用其他谓词 (*sargs* 和 *resids*) 的块索引扫描

| 隔离级别 | 只读和模糊扫描  | 游标操作     |      | 搜索的更新或删除 |        |
|------|----------|----------|------|----------|--------|
|      |          | 扫描       | 当前位置 | 扫描删除     | Update |
| RR   | IS/S/--  | IX/IX/-- |      | IX/IX/-- |        |
| RS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| CS   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |
| UR   | IN/IN/-- | IN/IN/-- |      | IN/IN/-- |        |

表 40. 延迟的数据页访问使用的块索引扫描的锁定方式: 在使用其他谓词 (*sargs* 和 *resids*) 的块索引扫描后延迟的数据页访问

| 隔离级别 | 只读和模糊扫描  | 游标操作    |         | 搜索的更新或删除 |         |
|------|----------|---------|---------|----------|---------|
|      |          | 扫描      | 当前位置    | 扫描删除     | Update  |
| RR   | IN/IN/-- | IX/IX/S | IX/IX/X | IX/IX/S  | IX/IX/X |
| RS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| CS   | IS/IS/NS | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |
| UR   | IN/IN/-- | IX/IX/U | IX/IX/X | IX/IX/U  | IX/IX/X |

## 对分区表的锁定行为

除整个表锁定之外, 还有一个对分区表的每个数据分区的锁定。与非分区表相比, 这允许更高的详细程度并提高了并行性。在 `db2pd` 命令、事件监视器、管理视图和表函数的输出中标识新数据分区锁定。

访问表时, 锁定行为按访问的数据所指示的那样, 首先获取表锁定, 然后获取数据分区锁定。访问方法和隔离级别可能获取不在结果集中的数据分区的锁定。获取这些数据分区锁定之后, 在表锁定期间可能会一直挂起这些锁定。例如, 对索引的游标稳定



性（CS）扫描可能保持对先前访问的数据分区的锁定，以便在后续键中引用该数据分区的情况下，减少重新获取该数据分区锁定的成本。数据分区锁定还承担用于确保对表空间的访问的成本。对于非分区表，表空间访问由表锁定处理。因此，对于分区表，即使在表级别存在互斥或共享锁定，也会进行数据分区锁定。

更高的详细程度允许一个事务具有对给定数据分区的互斥访问权并避免行锁定，而同时其他事务能够访问其他数据分区。这可能是由于为大量更新选择的方案或将锁定升级到数据分区级别而产生的。许多访问方法的表锁定通常是意向锁定，即使是以共享或互斥方式锁定数据分区。这将提高并行性。但是，如果在数据分区级别需要非意向锁定并且方案表明可能会访问所有数据分区，那么可能会在表级别选择非意向锁定，以防止并发事务的数据分区锁定之间出现死锁。

## 对 SQL LOCK TABLE 语句的锁定

对于分区表，在表级别获取对 LOCK TABLE 语句的唯一锁定；不获取数据分区锁定。这将确保没有对后续 DML 语句中的表的行锁定，并避免在行、块或数据分区级别出现死锁。在更新索引时使用 LOCK TABLE IN EXCLUSIVE MODE 还可用于保证互斥访问，这对于在大量更新期间限制 2 类索引的增长很有用。

## ALTER TABLE 语句的 LOCKSIZE TABLE 参数的影响

ALTER TABLE 语句有一个用于设置 LOCKSIZE TABLE 的选项，这确保不使用意向锁定以共享或互斥方式锁定表。对于分区表来说，这种锁定策略既适用于表锁定，也适用于对访问的任何数据分区的数据分区锁定。

## 行锁定和块锁定升级

对于分区表，行锁定和块锁定的升级单元是升级到数据分区级别。这再次表明与其他事务相比，更容易访问表，即使数据分区升级到共享、互斥或超级互斥，并使其他未升级的数据分区不受影响。对给定数据分区进行升级后，事务可能继续对其他数据分区进行行锁定。升级的通知日志消息包括已升级的数据分区和分区表的名称。因此，锁定升级不能确保对索引的互斥访问。语句必须使用互斥表级别锁定、必须发出显式 LOCK TABLE IN EXCLUSIVE MODE 语句，或表必须使用 LOCKSIZE TABLE 属性。由优化器根据数据分区消除选择访问方法的整个表锁定。如果没有发生数据分区消除，那么要大量更新表时可以选择互斥表锁定。

## 解释锁定信息

SNAPLOCK 管理视图中的以下示例可以帮助您解释对分区表返回的锁定信息。

示例 1:

此 SNAPLOCK 管理视图是在脱机索引重组期间捕获的。

```
SELECT SUBSTR(TABNAME, 1, 15) TABNAME, TAB_FILE_ID, SUBSTR(TBSP_NAME, 1, 15) TBSP_NAME, DATA_PARTITION_ID, LOCK_OBJECT_TYPE, LOCK_MODE, LOCK_ESCALATION FROM SYSIBMADM.SNAPLOCK where TABNAME like 'TP1' and LOCK_OBJECT_TYPE like 'TABLE_%' ORDER BY TABNAME, DATA_PARTITION_ID, LOCK_OBJECT_TYPE, TAB_FILE_ID, LOCK_MODE
```

| TABNAME | TAB_FILE_ID | TBSP_NAME    | DATA_PARTITION_ID | LOCK_OBJECT_TYPE | LOCK_MODE | LOCK_ESCALATION |
|---------|-------------|--------------|-------------------|------------------|-----------|-----------------|
| TP1     |             | 32768 -      | -1                | TABLE_LOCK       | Z         | 0               |
| TP1     |             | 4 USERSPACE1 | 0                 | TABLE_PART_LOCK  | Z         | 0               |
| TP1     |             | 5 USERSPACE1 | 1                 | TABLE_PART_LOCK  | Z         | 0               |
| TP1     |             | 6 USERSPACE1 | 2                 | TABLE_PART_LOCK  | Z         | 0               |
| TP1     |             | 7 USERSPACE1 | 3                 | TABLE_PART_LOCK  | Z         | 0               |
| TP1     |             | 8 USERSPACE1 | 4                 | TABLE_PART_LOCK  | Z         | 0               |

|     |    |            |    |                 |   |   |
|-----|----|------------|----|-----------------|---|---|
| TP1 | 9  | USERSPACE1 | 5  | TABLE_PART_LOCK | Z | 0 |
| TP1 | 10 | USERSPACE1 | 6  | TABLE_PART_LOCK | Z | 0 |
| TP1 | 11 | USERSPACE1 | 7  | TABLE_PART_LOCK | Z | 0 |
| TP1 | 12 | USERSPACE1 | 8  | TABLE_PART_LOCK | Z | 0 |
| TP1 | 13 | USERSPACE1 | 9  | TABLE_PART_LOCK | Z | 0 |
| TP1 | 14 | USERSPACE1 | 10 | TABLE_PART_LOCK | Z | 0 |
| TP1 | 15 | USERSPACE1 | 11 | TABLE_PART_LOCK | Z | 0 |
| TP1 | 4  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 5  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 6  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 7  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 8  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 9  | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 10 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 11 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 12 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 13 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 14 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 15 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |
| TP1 | 16 | USERSPACE1 | -  | TABLE_LOCK      | Z | 0 |

选择了 26 个记录。

在此示例中，类型为 TABLE\_LOCK 且 DATA\_PARTITION\_ID 为 -1 的锁定对象用于控制对分区表 TP1 的访问和并行性。类型为 TABLE\_PART\_LOCK 的锁定对象用于控制对每个数据分区的大多数访问和并行性。

此输出中还捕获了其他类型为 TABLE\_LOCK 的锁定对象 (TAB\_FILE\_ID 4 到 16)，这些锁定对象没有 DATA\_PARTITION\_ID 值。如果某种类型的锁定中对象的 TAB\_FILE\_ID 和 TBSP\_NAME 与分区表中的数据分区或索引相对应，那么可以使用这种类型的锁定来控制脱机备份实用程序的并行性。

---

## 第 24 章 代理程序管理

---

### 分区数据库中的代理程序

对于分区数据库环境和启用分区内并行性的环境，每个数据库分区（即每个数据库服务器或节点）有它自己的代理程序存储池，可从中抽出子代理程序。因为有此缓冲池，子代理程序就不必在每次需要时创建或完成其工作时破坏掉。这些子代理程序仍然可以作为此缓冲池中的相关代理程序，并由数据库管理器使用，以执行与它们相关的应用程序或者新应用程序发出的新请求。

对于分区数据库环境和启用分区内并行性的环境，对系统中性能和内存成本的影响与代理程序存储池是如何调整的有很大关系：

- 代理程序缓冲池大小的数据库管理器配置参数 (*num\_poolagents*) 影响可与一个数据库分区（也称为节点）上的应用程序保持相关的代理程序和子代理程序的总数。如果缓冲池大小太小且该缓冲池已满，那么子代理程序解除它自己与它正在其上工作的应用程序的关联并终止。因为必须常常创建子代理程序，并重新使它们与应用程序相关，所以性能会受损。

缺省情况下，*num\_poolagents* 设置为 AUTOMATIC 并且值为 100。当此参数设置为 AUTOMATIC 时，数据库管理器将自动管理池中的空闲代理程序数。

另外，如果 *num\_poolagents* 的值太小，那么一个应用程序就可能填满有相关子代理程序的缓冲池。因此，当另一个应用程序需要新的子代理程序，并且在其关联的代理程序池中沒有子代理程序时，它将从其他应用程序的代理程序池中重新启动不活动的子代理程序。这种行为可确保充分利用资源。

- 必须根据在任何给定时间允许过多代理程序活动的资源成本权衡这些情况。

例如，如果 *num\_poolagents* 的值太大，使用不能用于其他任务的数据库管理器资源，那么相关的子代理程序也许很长时间都留在缓冲池中未使用。

**注：**当启用连接集中器时，由 *num\_poolagents* 指定的代理程序数不一定会影响任何时候在池中处于空闲状态的代理程序的准确数目。可能会临时超过代理程序数以处理更多的工作负载活动。

#### 其他异步进程和线程

除数据库代理程序外，其他异步数据库管理器活动还作为它们自己的进程或线程运行，包括：

- 数据库 I/O 服务器或 I/O 预取程序
- 数据库异步页清除程序
- 数据库记录器
- 数据库死锁检测器
- 通信和 IPC 侦听器
- 表空间容器重平衡程序。



---

## 第 25 章 优化访问方案

---

### 索引访问和集群比率

当它选择访问方案时，优化器估计将所需要的页从磁盘访存至缓冲池所需要的 I/O 次数。此估计包括预测缓冲池的使用情况，因为不需要其他 I/O 来读取已在缓冲池内一页中的行。

对于索引扫描，系统目录表 (SYSCAT.INDEXES) 中的信息帮助优化器估计将数据页读入缓冲池的 I/O 成本。它使用 SYSCAT.INDEXES 表中以下列中的信息：

- **CLUSTERRATIO** 信息指示表数据关于此索引的集群度。该数越高，按索引键顺序排序的行越好。如果表行接近索引键顺序，那么在数据页位于缓冲区时可以从该页中读取行。如果此列的值为 -1，那么优化器使用 **PAGE\_FETCH\_PAIRS** 和 **CLUSTERFACTOR** 信息（如果它可用的话）。
- **PAGE\_FETCH\_PAIRS** 包含用于模拟将数据页读入各种大小的缓冲池所需的 I/O 数的多对数字以及 **CLUSTERFACTOR** 信息。仅当您使用 **DETAILED** 子句对索引执行 **RUNSTATS** 时，才对这些列收集数据。

如果索引集群统计信息不可用，那么优化器使用缺省值，在这种情况下假定数据与该索引的集群关系不强。

数据相对于索引的集群度可大大影响性能，您应尝试保持该表上的一个索引的集群率接近 100%。

一般情况下，只能有一个索引可达到百分之百的集群，但下列情况除外：其键是集群索引键的超集的那些索引，或者在两个索引的键列之间存在实际相关的那些索引。

当您重组表时，您可指定一个索引，它将用于集群这些行并试图在插入处理期间保留此特征。因为更新和插入可能降低表相对于索引的集群率，所以您可能需要定期重组该表。要降低对由于 **INSERT**、**UPDATE** 和 **DELETES** 而频繁更改的表重组的频率，当改变表时使用 **PCTFREE** 参数。这允许把附加的插入与现有的数据集群在一起。

### MDC 表的表和索引管理

多维集群 (MDC) 表的表和索引组织基于与标准表组织相同的逻辑结构。与标准表类似，MDC 表是按页组织的，这些页包含分为许多列的数据行，每页上的行由行标识 (RID) 标识。但是，MDC 表的各页还分组为按扩展数据块大小计算的块。例如，下图显示扩展数据块大小为 4 的表，前四页（从 0 到 3 编号）是表中的第一个块。下一组页（从 4 到 7 编号）是表中的第二个块。

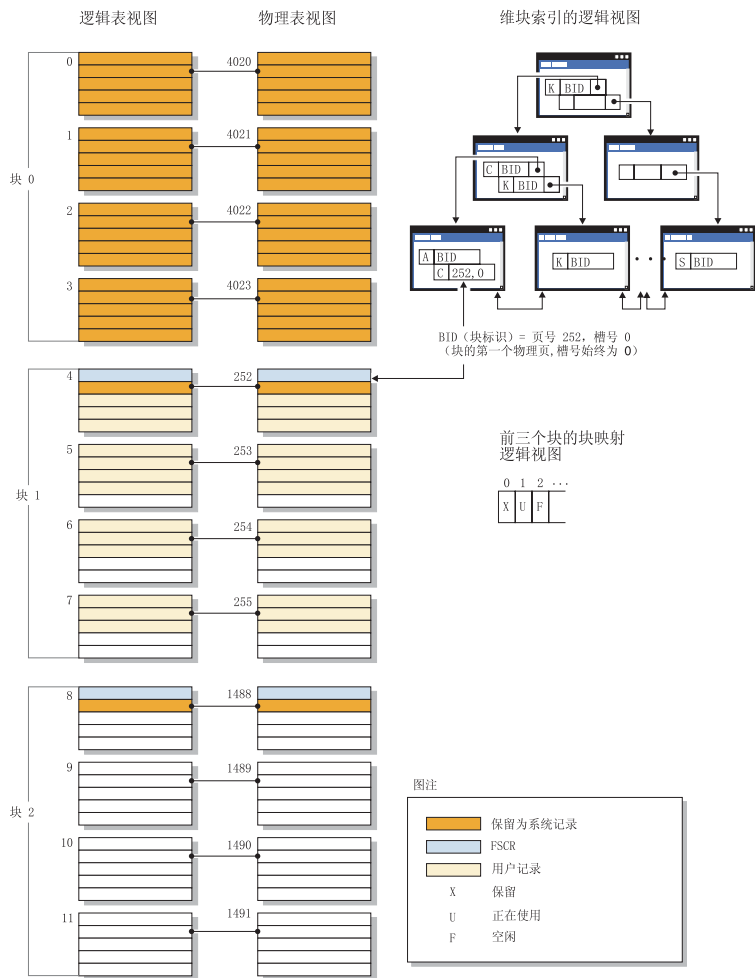


图 53. MDC 表的逻辑表、记录和索引结构

第一个块包含 DB2 用来管理表的特殊内部记录，包括可用空间控制记录（FSCR）。在后续块中，第一页包含 FSCR。FSCR 为块中每页上存在的新记录映射可用空间。当将记录插入表时，将使用这部分可用的可用空间。

顾名思义，MDC 表将对多个维上的数据进行集群。每个维都由您在 CREATE TABLE 语句的 ORGANIZE BY DIMENSIONS 子句中指定一列或一组列确定。当创建 MDC 表时，会自动创建以下两种索引：

- 维块索引，它包含指向单个维的每个已占用块的指针。
- 组合块索引，它包含所有维键列。组合块索引用于在插入和更新活动期间维护集群。

当优化器确定特定查询的最有效的访问方案时，它考虑利用维块索引的访问方案。当查询在维值上具有谓词时，优化器可以使用维块索引来标识包含这些值的扩展数据

块，并从扩展数据块中进行访存。因为扩展数据块是磁盘中物理上相邻的页，所以这会导致更有效的性能并使 I/O 次数减少到最低。

另外，如果数据访问方案的分析指示特定 RID 索引将提高查询性能，那么可以创建这样的索引。

除了维块索引和组合块索引以外，MDC 表还维护包含位图的块映射，该位图指示每个块的可用性状态。在位图列表中编码下列属性：

- X (保留)：第一个块仅包含表的系统信息。
- U (正在使用)：此块已被使用并与维块索引相关联
- L ((已装入)：已通过当前装入操作装入此块
- C (检查约束)：此块由装入操作设置，用来指定装入期间的增量约束检查。
- T (刷新表)：此块由装入操作设置，用来指定需要 AST 维护。
- F (空闲)：如果没有设置其他属性，该块被认为是空闲的。

因为每个块都在块映射文件中具有一个条目，所以文件随表的增长而增长。作为单独对象存储此文件。在 SMS 表空间中，它是一个新的文件类型。在 DMS 表空间中，它在对象表中具有新的对象描述符。

## 集群

有时，由于更新降低了索引与数据页之间存在的集群度，所以会导致数据页上的行更改位置。

重组与所选索引相关的表将对数据重新建立集群。建立了集群的索引对于带有范围谓词的列非常有用，因为它允许对基本表中的数据作更有效的顺序访问。由于相似的值在同一数据页上，从而减少页访存次数。

通常，表中只有一个索引可以具有较高的集群度。

要检查索引的集群度，可双击它的节点以显示“索引统计信息”窗口。集群比率或集群因子的值显示在此窗口中。若该值较低，应考虑重组表数据。

---

## 分区内并行性的优化策略

如果在编译 SQL 语句时指定了并行度，优化器能选择一个访问方案来在单个数据库执行查询。

在执行时，创建称为“子代理程序”的多个数据库代理程序来执行该查询。子代理程序的数目小于或等于编译该 SQL 语句时指定的并行度。

要并行化访问方案，优化器将它划分为两个部分，每个子代理程序运行一部分，协调代理程序运行另一部分。子代理程序通过表队列将数据传送至协调代理程序或其他子代理程序。在分区数据库环境中，子代理程序能通过表队列从其他数据库分区中的子代理程序发送或接收数据。

### 分区内并行扫描策略

关系扫描和索引扫描可在同一个表或索引上并行执行。要进行并行关系扫描，需将表划分为由页或行组成的范围。将一个范围内的页或行分配给一个子代理程序。一个子代理程序扫描分配给它的范围，当它处理完当前范围时，会给它分配另一个范围。

要进行并行索引扫描，需根据索引键值和一个键值的索引条目数，将索引划分为许多记录范围。并行索引扫描的执行方式类似于并行表扫描，将给予代理程序分配一个范围内的记录。当子代理程序处理完当前范围时，会给它分配一个新的范围。

优化器确定扫描单位（页或行）和扫描详细程度。

并行扫描将工作均匀地分布在各子代理程序中。并行扫描的目标是平衡所有子代理程序的负载，并使它们保持相同的繁忙程度。如果繁忙子代理程序数等于可用处理器数，且磁盘没有过度处理 I/O 请求，那么表明机器资源正在得到有效的利用。

当执行查询时，其他访问方案策略可能导致数据不平衡。优化器选择并行策略，以便维护子代理程序间的数据平衡。

### 分区内并行排序策略

优化器能选择下列其中一个并行排序策略：

- **循环排序**

这也称为再分布排序。这种方法有效地使用共享内存，将数据尽可能均匀地再分布到所有子代理程序中。它使用轮询算法来提供均匀分布。它首先为每个子代理程序创建个别排序。在插入阶段期间，子代理程序以循环方式插入每个个别的排序中，以使数据分布更加均匀。

- **分区排序**

这类似于循环排序，在循环排序中，要为每个子代理程序创建一个排序。子代理程序将一个散列函数应用于排序列，以确定应将行插入哪个排序中。例如，如果一个合并连接的内部表和外部表是一个分区排序，那么子代理程序可使用合并连接来连接对应的表部分并且并行执行。

- **复制排序**

如果每个子代理程序需要全部排序输出，那么使用这种排序。创建一个排序，当将行插入排序时使子代理程序同步。当完成排序时，每个子代理程序都读取整个排序。如果行数较小，可使用此排序来重新平衡数据流。

- **共享排序**

此排序与复制排序相同，只是子代理程序要对已排序的结果打开一个并行扫描以使用一种与循环排序相似的方式在子代理程序中分布数据。

### 分区内并行临时表

子代理程序可以协同工作，将行插入同一个表中来产生临时表。这称为共享临时表。子代理程序可以根据数据流是要复制还是要分割，在共享临时表上打开专用扫描或并行扫描。

### 分区内并行聚集策略

子代理程序可以并行执行聚集操作。聚集操作要求根据分组列将数据排序。如果可以保证一个子代理程序接收一组分组列值的所有行，那么该程序可以执行完整的聚集。如果由于先前的分区排序，已在分组列上分割了数据流，那么这种情况可发生。



否则，子代理程序可以执行部分聚集，并使用另一种策略来完成该聚集。其中部分策略如下：

- 通过一个合并表队列将部分聚集的数据发送至协调代理程序。协调程序完成聚集。
- 将部分聚集的数据插入一个分区排序中。在分组列上分割该排序，并保证一组分组的的所有行都包含在一个排序分区中。
- 如果需要复制数据流以便平衡处理，可将部分聚集的数据插入一个复制排序中。每个子代理程序使用复制排序完成聚集，并接收完全相同的聚集结果副本。

### 分区内并行连接策略

子代理程序可以并行执行连接操作。并行连接策略由数据流的特征确定。

通过在连接的内部和外部表上分区和/或复制数据流，可将连接并行化。例如，如果因为并行扫描已将一个嵌套循环连接的外部流分区，而且内部流由每个子代理程序单独重新求值，那么可将该连接并行化。如果一个合并连接的内部和外部流因为分区排序都已按值分区，那么可将该连接并行化。

---

## db2expln 和 dynexpln 输出的示例

以下显示的示例可以帮助您了解 db2expln 和 dynexpln 的输出的布局和格式。这些示例是对 DB2 中提供的 SAMPLE 数据库运行的（除非显示不同）。每个示例都提供了简短的讨论。示例之间的主要区别用**粗体**显示。

### 示例一：无并行性

本示例只请求一个列表，它包含所有职工的姓名、职务、部门名称和位置，以及他们所从事的项目的名称。此访问方案的实质是使用散列连接将每个指定的表中的相关数据连接在一起。因为没有索引可用，故访问方案在连接每个表时对它们进行关系扫描。

```
***** PACKAGE *****
Package Name = "DOOLE"."EXAMPLE" Version = ""
Prep Date = 2002/01/04
Prep Time = 14:05:00

Bind Timestamp = 2002-01-04-14.05.00.415403

Isolation Level          = Cursor Stability
Blocking                  = Block Unambiguous Cursors
Query Optimization Class = 5

Partition Parallel       = No
Intra-Partition Parallel = No

SQL Path                  = "SYSIBM", "SYSFUN", "SYSPROC", "DOOLE"

----- SECTION -----
Section = 1

SQL Statement:
  DECLARE EMPCUR CURSOR
  FOR
    SELECT e.lastname, e.job, d.deptname, d.location, p.projname
    FROM employee AS e, department AS d, project AS p
    WHERE e.workdept = d.deptno AND e.workdept = p.deptno
```



\*\*\*\*\* PACKAGE \*\*\*\*\*

Package Name = "DOOLE"."EXAMPLE" Version = ""  
Prep Date = 2002/01/04  
Prep Time = 14:12:38

Bind Timestamp = 2002-01-04-14.12.38.732627

Isolation Level = Cursor Stability  
Blocking = Block Unambiguous Cursors  
Query Optimization Class = 5

Partition Parallel = No  
Intra-Partition Parallel = **Yes (Bind Degree = 4)**

SQL Path = "SYSIBM", "SYSFUN", "SYSPROC", "DOOLE"

----- SECTION -----

Section = 1

SQL Statement:

```
DECLARE EMPCUR CURSOR  
FOR
```

```
SELECT e.lastname, e.job, d.deptname, d.location, p.projname  
FROM employee AS e, department AS d, project AS p  
WHERE e.workdept = d.deptno AND e.workdept = p.deptno
```

**Intra-Partition Parallelism Degree = 4**

Estimated Cost = 133.934692  
Estimated Cardinality = 221.535980

( 2) Process Using 4 Subagents

```
( 7) | Access Table Name = DOOLE.EMPLOYEE ID = 2,5  
      | | | | | #Columns = 3  
      | | | | | Parallel Scan  
      | | | | | Relation Scan  
      | | | | | | Prefetch: Eligible  
      | | | | | Lock Intents  
      | | | | | Table: Intent Share  
      | | | | | Row : Next Key Share
```

( 7) | | Process Build Table for Hash Join

```
( 3) | Hash Join  
      | | | | | Estimated Build Size: 7111  
      | | | | | Estimated Probe Size: 9457
```

```
( 6) | | Access Table Name = DOOLE.PROJECT ID = 2,7  
      | | | | | #Columns = 2  
      | | | | | Parallel Scan  
      | | | | | Relation Scan  
      | | | | | | Prefetch: Eligible  
      | | | | | Lock Intents  
      | | | | | Table: Intent Share  
      | | | | | Row : Next Key Share
```

( 6) | | | Process Build Table for Hash Join

```
( 4) | | Hash Join  
      | | | | | Estimated Build Size: 5737  
      | | | | | Estimated Probe Size: 6421
```

```
( 5) | | | Access Table Name = DOOLE.DEPARTMENT ID = 2,4  
      | | | | | #Columns = 3  
      | | | | | Parallel Scan  
      | | | | | Relation Scan  
      | | | | | | Prefetch: Eligible  
      | | | | | Lock Intents  
      | | | | | Table: Intent Share  
      | | | | | Row : Next Key Share
```

( 5) | | | | Process Probe Table for Hash Join



Estimated Cost = 118.483406  
Estimated Cardinality = 474.720032

```
Coordinator Subsection:(-----)  Distribute Subsection #2
| Broadcast to Node List
| Nodes = 10, 33, 55
(-----) Distribute Subsection #3
| Broadcast to Node List
| Nodes = 10, 33, 55
(-----) Distribute Subsection #1
| Broadcast to Node List
| Nodes = 10, 33, 55
( 2) Access Table Queue ID = q1 #Columns = 5
( 1) Return Data to Application
| #Columns = 5

Subsection #1:( 8) Access Table Queue ID = q2 #Columns = 2
( 3) Hash Join
| Estimated Build Size: 5737
| Estimated Probe Size: 8015
( 6) | Access Table Queue ID = q3 #Columns = 3
( 4) | Hash Join
| | Estimated Build Size: 5333
| | Estimated Probe Size: 6421
( 5) | | Access Table Name = DOOLE.DEPARTMENT ID = 2,4
| | | #Columns = 3
| | | Relation Scan
| | | | Prefetch: Eligible
| | | | Lock Intents
| | | | Table: Intent Share
| | | | Row : Next Key Share
( 5) | | | Process Probe Table for Hash Join
( 2) Insert Into Asynchronous Table Queue ID = q1
| Broadcast to Coordinator Node
| Rows Can Overflow to Temporary Table

Subsection #2:
( 9) Access Table Name = DOOLE.PROJECT ID = 2,7
| #Columns = 2
| Relation Scan
| | Prefetch: Eligible
| | Lock Intents
| | Table: Intent Share
| | Row : Next Key Share
( 9) | Insert Into Asynchronous Table Queue ID = q2
| | Hash to Specific Node
| | Rows Can Overflow to Temporary Tables
( 8) Insert Into Asynchronous Table Queue Completion ID = q2

Subsection #3:
( 7) Access Table Name = DOOLE.EMPLOYEE ID = 2,5
| #Columns = 3
| Relation Scan
| | Prefetch: Eligible
| | Lock Intents
| | Table: Intent Share
| | Row : Next Key Share
( 7) | Insert Into Asynchronous Table Queue ID = q3
| | Hash to Specific Node
| | Rows Can Overflow to Temporary Tables
( 6) Insert Into Asynchronous Table Queue Completion ID = q3
```

End of section

Optimizer Plan:



```

DECLARE EMPCUR CURSOR
FOR
    SELECT e.lastname, e.job, d.deptname, d.location, p.projname
    FROM employee AS e, department AS d, project AS p
    WHERE e.workdept = d.deptno AND e.workdept = p.deptno

```

**Intra-Partition Parallelism Degree = 4**

Estimated Cost = 145.198898  
 Estimated Cardinality = 474.720032

```

Coordinator Subsection:(-----)  Distribute Subsection #2
| Broadcast to Node List
| Nodes = 10, 33, 55
(-----) Distribute Subsection #3
| Broadcast to Node List
| Nodes = 10, 33, 55
(-----) Distribute Subsection #1
| Broadcast to Node List
| Nodes = 10, 33, 55
( 2) Access Table Queue ID = q1 #Columns = 5
( 1) Return Data to Application
| #Columns = 5

```

```

Subsection #1:( 3) Process Using 4 Subagents
( 10) | Access Table Queue ID = q3 #Columns = 2
( 4) | Hash Join
| | | | Estimated Build Size: 5737
| | | | Estimated Probe Size: 8015
( 7) | | Access Table Queue ID = q5 #Columns = 3
( 5) | | Hash Join
| | | | Estimated Build Size: 5333
| | | | Estimated Probe Size: 6421
( 6) | | | Access Table Name = DOOLE.DEPARTMENT ID = 2,4
| | | | #Columns = 3
| | | | Parallel Scan
| | | | Relation Scan
| | | | Prefetch: Eligible
| | | | Lock Intents
| | | | Table: Intent Share
| | | | Row : Next Key Share
( 6) | | | | Process Probe Table for Hash Join
( 3) | Insert Into Asynchronous Local Table Queue ID = q2
( 3) Access Local Table Queue ID = q2 #Columns = 5
( 2) Insert Into Asynchronous Table Queue ID = q1
| Broadcast to Coordinator Node
| Rows Can Overflow to Temporary Table

```

```

Subsection #2:
( 11) Process Using 4 Subagents
( 12) | Access Table Name = DOOLE.PROJECT ID = 2,7
| | | | #Columns = 2
| | | | Parallel Scan
| | | | Relation Scan
| | | | Prefetch: Eligible
| | | | Lock Intents
| | | | Table: Intent Share
| | | | Row : Next Key Share
( 11) | Insert Into Asynchronous Local Table Queue ID = q4
( 11) Access Local Table Queue ID = q4 #Columns = 2
( 10) Insert Into Asynchronous Table Queue ID = q3
| Hash to Specific Node
| Rows Can Overflow to Temporary Tables

```

```

Subsection #3:
( 8) Process Using 4 Subagents
( 9) | Access Table Name = DOOLE.EMPLOYEE ID = 2,5

```





| Table1 | Table2  |
|--------|---------|
| PROJ   | PROJ_ID |
|        | 2       |
|        | 名称      |
|        | Mike    |

要将标识列具有相同值的 Table1 和 Table2 连接，使用下列 SQL 语句：

```
SELECT PROJ, x.PROJ_ID, NAME
FROM TABLE1 x, TABLE2 y
WHERE x.PROJ_ID = y.PROJ_ID
```

此查询得到下列一组结果行：

| PROJ | PROJ_ID | 名称   |
|------|---------|------|
| A    | 1       | Sam  |
| A    | 1       | Sue  |
| B    | 2       | Mike |
| C    | 3       | Joe  |
| D    | 4       | Mary |

根据连接谓词是否存在，以及所涉及到的、由表和索引统计信息来确定的各种成本，优化器选择下列其中一种连接方法：

- 嵌套循环连接
- 合并连接
- 散列连接

当连接两个表时，将一个表选择为外部表，而将另一个表选为内部表。首先访问外部表，并只扫描它一次。是否扫描内部表多次取决于连接的类型和存在的索引。即使一个查询连接两个以上表，优化器一次也只连接两个表。如果有必要，创建临时表来保存中间结果。

可以提供显式连接运算符（如 INNER 或 LEFT OUTER JOIN）来确定如何在连接中使用表。然而，以这种方式改变查询之前，应允许优化器确定如何连接表。然后分析查询性能来决定是否要添加连接操作程序。

## 数据库分区组对查询优化的影响

在分区数据库环境中，优化器识别表的并置，并在确定查询的最佳访问方案时使用此并置。如果在连接查询中频繁涉及某些表，那么应该将那些表划分到分区数据库环境中的各数据库分区中，以便要连接的每个表中的行位于同一数据库分区中。在执行连接操作期间，两个已连接的表中数据的并置会防止将数据从一个数据库分区移动至另一个数据库分区。将两个表置于同一数据库分区组中以确保并置两个表中的数据。

在分区数据库环境中，根据表的大小，将数据分布在多个数据库分区上可缩短执行查询的估计时间（或成本）。表的数目、表的大小、那些表中数据的位置以及查询类型（例如，是否需要连接）都会影响查询的成本。

## 分区数据库中的连接策略

在某些方面，连接策略在分区数据库环境中与非分区数据库环境中不同。可以将其他技术应用于标准连接方法以提高性能。

对于在一个分区数据库环境中进行的频繁连接所涉及的那些表，要考虑的一点是表并置。表并置提供了一种方法，即在一个分区数据库环境中，根据相同的分布键，可在同一个数据库分区中使用一个表中的数据查找另一个表中的数据。一旦被并置，要连接的数据可参与查询，而不必在执行查询活动时移动至另一个数据库分区。只将连接的答案集移动至协调程序节点。

## 表队列

分区数据库环境中连接技术的描述使用下列术语：

- 表队列

在数据库分区之间或单一分区数据库中的处理器之间传送行的一种机制。

- 定向表队列

一种表队列，其中，将行散列至其中一个接收数据库分区。

- 广播表队列

一种表队列，其中，将行发送至所有接收数据库分区，但不散列这些行。

表队列用于下列情况下：

- 在使用分区间并行性时，将表数据从一个数据库分区传送至另一个
- 在使用分区内并行性时，传送一个数据库分区中的表数据
- 在使用单一分区数据库时，传送一个数据库分区中的表数据。

每个表队列按单方向传送数据。编译器决定何处需要使用表队列，并在方案中包括表队列。当执行方案时，数据库分区之间的连接会启动表队列。一旦进程结束，那么表队列关闭。

有几种类型的表队列：

- *异步表队列*。这些表队列称为异步，是因为它们在应用程序发出的任何 `FETCH` 之前读取行。当发出 `FETCH` 时，从表队列检索该行。

当您在 `SELECT` 语句上指定 `FOR FETCH ONLY` 子句时，使用异步表队列。如果您只读取行，那么异步表队列会更快。

- *同步表队列*。这些表队列称为同步，是因为它们对应用程序发出的每个 `FETCH` 读取一行。在每个数据库分区中，将游标定位在要从该数据库分区读取的下一行上。

当您未在 `SELECT` 语句上指定 `FOR FETCH ONLY` 子句时，使用同步表队列。在一个分区数据库环境中，如果您要更新行，那么数据库管理器将使用同步表队列。

- *合并表队列*。

这些表队列维持顺序。

- *非合并表队列*。

这些表队列也称为“正规”表队列。它们不保持顺序。

- *侦听器表队列*。

这些表队列在关联子查询中使用。使用这种类型的表队列，将关联值向下传送至子查询，然后将结果向上传送回父查询块。

## 分区数据库环境中的连接方法

下列图形举例说明了分区数据库环境中的连接方法。

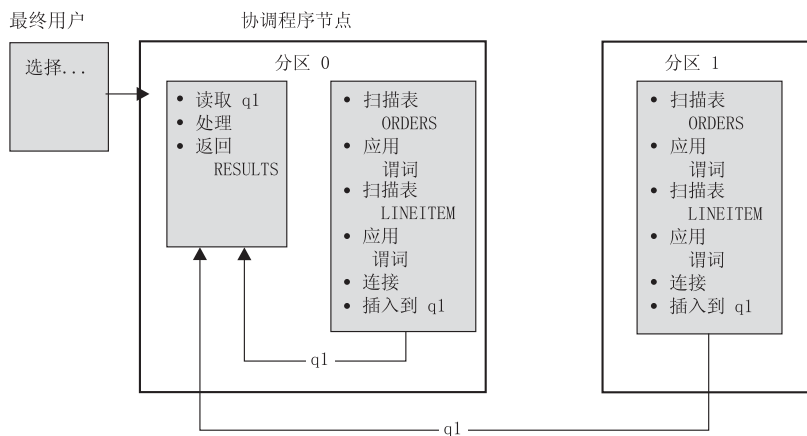
注：在这些图表中，q1、q2、和 q3 表示示例中的表队列。引用的表分布在两个数据库分区中，以实施这些方案。箭头指示发送表队列的方向。协调程序节点是数据库分区 0。

### 并置连接

并置连接以本地方式在数据所在的数据库分区上发生。该数据库分区在完成连接以后将数据发送到其他数据库分区。要使优化器考虑并置连接，必须并置连接的表，且所有相应的分布键对必须参与等式连接谓词。

以下图形提供了一个示例。

注：复制的具体化查询表提高了并置连接的可能性。



在 ORDERKEY 列上对 LINEITEM 表和 ORDERS 表进行分区。

连接在每个数据库分区本地执行。

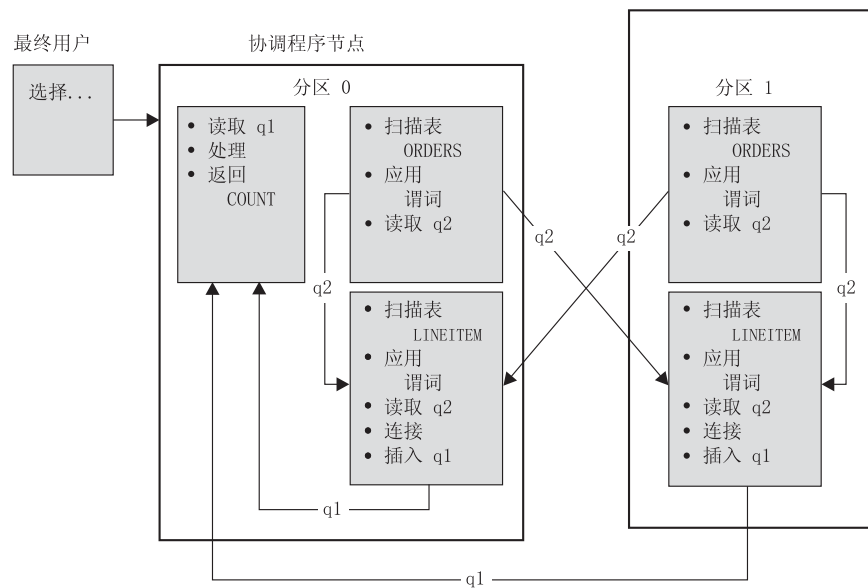
在此例子中，假定该连接谓词为：

`ORDERS.ORDERKEY = LINEITEM.ORDERKEY。`

图 54. 并置连接示例

### 广播外部表连接

广播外部表连接是一个并行连接策略，如果在连接的表之间没有等式连接谓词，可使用该连接。在可以将它用作成本最低的连接方法的其他情况中，也可使用该连接。例如，当有一个很大的表和一个很小的表且没有在连接谓词列上对任何一个表进行分割时，可能会发生广播外部表连接。不分割这两个表，将更小的表广播至更大的表可能成本更低。以下图形提供了一个示例。

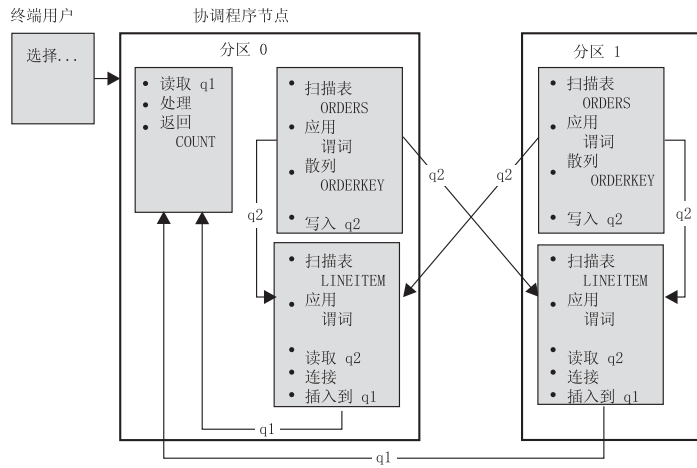


ORDERS 表将发送至所有具有 LINEITEM 表的数据库分区。  
表队列 q2 将广播至内部表的所有数据库分区。

图 55. 广播外部表连接示例

### 定向外部表连接

在定向外部表连接策略中，根据内部表的分割属性将外部表的每一行发送至内部表的一部分。该连接在此数据库分区上进行。以下图形提供了一个示例。

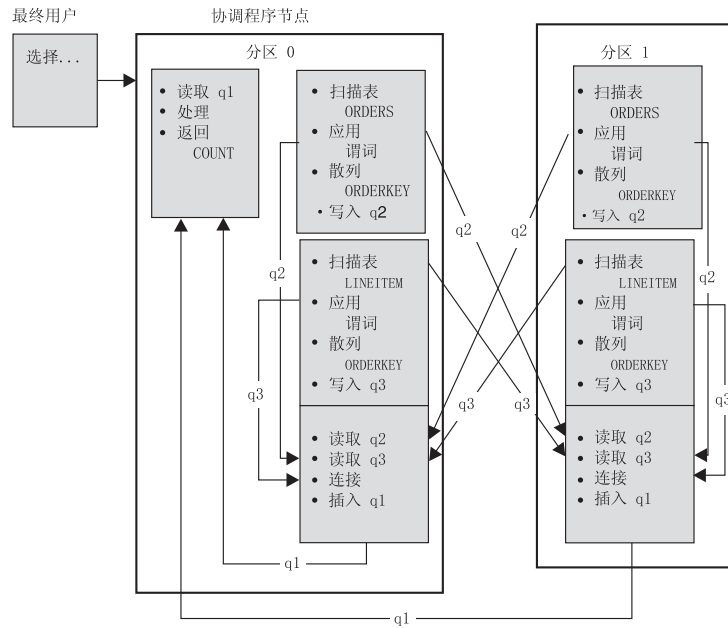


LINEITEM 表在 ORDERKEY 列上进行分区。  
 ORDERS 表在另一列上进行分区。  
 ORDERS 表被散列并发送至正确的 LINEITEM 表数据库分区。  
 在此例子中, 假定连接谓词为:  
 ORDERS.ORDERKEY = LINEITEM.ORDERKEY。

图 56. 定向外部表连接示例

### 定向内部表和外部表连接

在定向内部表和外部表策略中, 根据连接列的值, 将外部表和内部表的行定向到一组数据库分区。该连接在这些数据库分区上进行。以下图形提供了一个示例。以下图形中显示了一个示例。

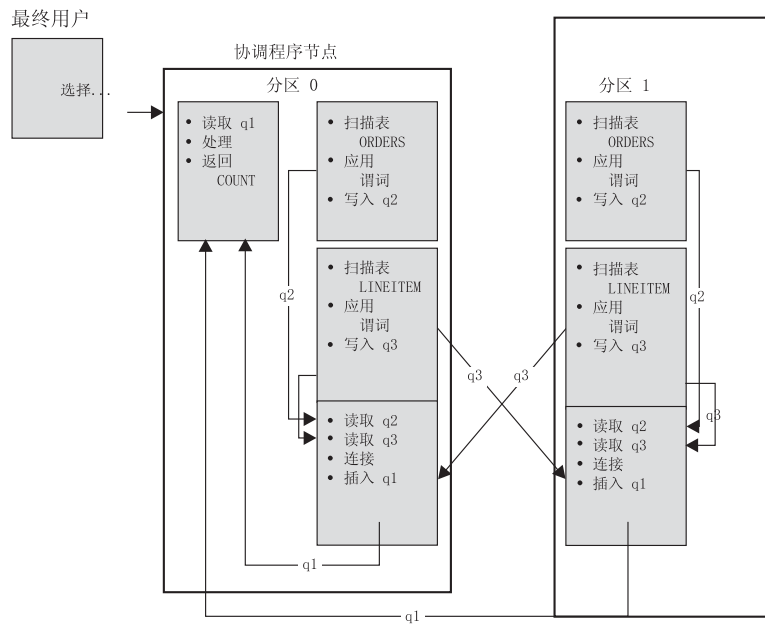


没有任何一个表分区到 ORDERKEY 列。  
 这两个表都被散列且发送至它们所连接的新数据库分区。  
 表队列 q2 和 q3 都已定向。  
 在此例子中,假定连接谓词为:  
`ORDERS.ORDERKEY = LINEITEM.ORDERKEY`

图 57. 定向内部表和外部表连接示例

### 广播内部表连接

在广播内表部连接策略中, 将内部表广播至外连接表的所有数据库分区。以下图形提供了一个示例。

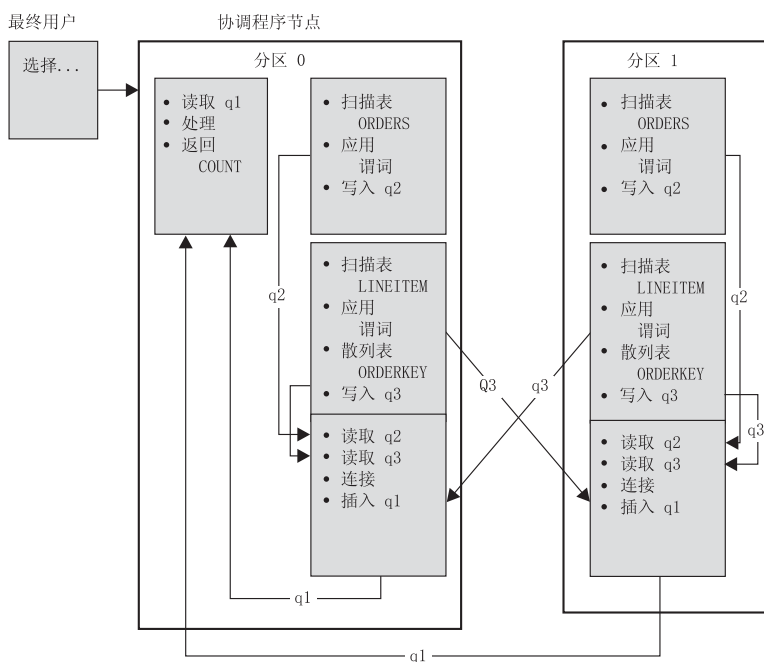


LINEITEM 表将发送至所有具有 ORDERS 表的数据库分区。  
 表队列 q3 将广播至外部表的所有数据库分区。

图 58. 广播内部表连接示例

### 定向内部表连接

使用定向内部表连接策略，根据外部表的分割属性，将内部表的每一行发送至外连接表的一个数据库分区。该连接在此数据库分区上进行。以下图形提供了一个示例。



ORDERS 表在 ORDERKEY 列上进行分区。  
 LINEITEM 表在另一列上进行分区。  
 散列 LINEITEM 表并将其发送至正确的 ORDERS 表数据库分区。  
 在此例子中, 假定连接谓词为:  
 ORDERS.ORDERKEY = LINEITEM.ORDERKE.

图 59. 定向内部表连接示例

## 分区数据库环境中重复的具体化查询表

重复具体化查询表通过允许数据库管理表数据的预先计算的值来提高分区数据库环境中频繁执行连接的性能。

考虑一个查询和重复具体表示例。作下列假定:

- SALES 表在多分区表空间 REGIONTABLESPACE 中, 且在 REGION 列上分割。
- EMPLOYEE 和 DEPARTMENT 表在单分区数据库分区组中。

根据 EMPLOYEE 表中的信息创建复制的具体化查询表。

```
CREATE TABLE R_EMPLOYEE
AS (
    SELECT EMPNO, FIRSTME, MIDINIT, LASTNAME, WORKDEPT
    FROM EMPLOYEE
)
DATA INITIALLY DEFERRED REFRESH IMMEDIATE
IN REGIONTABLESPACE
REPLICATED;
```

要更新重复的具体化查询表的内容, 运行下列语句:

```
REFRESH TABLE R_EMPLOYEE;
```



**注：**使用 REFRESH 语句之后，您应该在复制表上运行 RUNSTATS，正如您在任何其他表上要运行的一样。

以下示例计算职员的销售额、部门总销售额和总计：

```
SELECT d.mgrno, e.empno, SUM(s.sales)
FROM   department AS d, employee AS e, sales AS s
WHERE  s.sales_person = e.lastname
       AND e.workdept = d.deptno
GROUP BY ROLLUP(d.mgrno, e.empno)
ORDER BY d.mgrno, e.empno;
```

如果不使用仅在一个数据库分区上的 EMPLOYEE 表，那么数据库管理器使用 R\_EMPLOYEE 表，该表在存储 SALES 表的每个数据库分区上均有复制。这样性能将会增强，因为职员信息不必通过网络传送到每个数据库分区来计算该连接。

### 并置连接中的重复具体化查询表

重复的具体化查询表也可以有助于并置连接。例如，如果一个星型模式包含一个在 20 个节点上分布的大的事实表，那么在事实表和维表之间的连接是最有效的（如果并置了这些表）。如果所有表位于同一个数据库分区组中，那么对于一个并置连接，最多正确分区一个维表。其他的维表不能在并置连接中使用，因为事实表上的连接列与事实表的分布键不对应。

考虑在 C1 上分割称为 FACT (C1, C2, C3, ...) 的表；在 C1 上分割称为 DIM1 (C1, dim1a, dim1b, ...) 的表；在 C2 上分割称为 DIM2 (C2, dim2a, dim2b, ...) 的表；等等。

在此情况下，您会发现在 FACT 和 DIM1 之间的连接是最好的，因为并置了谓词 DIM1.C1 = FACT.C1。这两个表都在列 C1 上分割。

但是，不能并置 DIM2 和谓词 WHERE DIM2.C2 = FACT.C2 之间的连接，因为在列 C1 上分割 FACT，而不在列 C2 上分割 FACT。在此情况下，可以在事实表的数据库分区组中复制 DIM2，以便连接按本地方式在每个数据库分区上进行。

**注：**此处复制的具体化查询表讨论与分区内数据库复制有关。交互数据库复制与预订、控制表以及位于不同数据库中和不同操作系统上的数据有关。

当创建复制的具体化查询表时，源表可以是数据库分区组中的单节点表或多节点表。在大多数情况下，复制的表不大，并且可以放置在单节点数据库分区组中。可以通过指定只有此表中列的子集，或通过指定使用的谓词行数，或通过使用两种方法来限制要复制的数据。为了复制具体化查询表有效，并不需要数据捕获选项。

也可以在多节点数据库分区组中创建复制具体化查询表，以便在所有数据库分区上创建源表的副本。如果向所有数据库分区广播此源表，那么大型事实表和维表之间的连接在此环境中更可能在本地执行。

未自动创建复制表上的索引。可以创建与源表上的索引不同的索引。但是，为防止未在源表上出现的约束违例，不能创建唯一索引或对复制的表使用约束。即使在源表上发生相同的约束，也禁止这些约束。

可以在查询中直接引用复制的表，但您不可以将 NODENUMBER() 谓词与复制表一起使用，来查看特定分区上的表数据。

使用说明工具来查看查询的访问方案是否使用了复制的具体化查询表。优化器所选择的访问方案是否使用复制的具体化查询表取决于连接所需要的信息。如果优化器确定向数据库分区组中的其他数据库分区广播原始源表比较便宜，那么优化器可能不使用复制具体化查询表。

## 第 4 课 在分区数据库环境中改进访问方案

您将了解当执行各种调整活动时基本查询的访问方案和相关窗口如何更改。

通过使用一系列示例，附带一些图例，您将通过使用 **runstats** 命令并添加适当的索引来了解如何改进访问方案的估计总成本（即使是简单查询的访问方案的估计总成本）。

当有了使用 Visual Explain 的经验时，您将发现调整查询的其他方法。

### 使用访问方案图

通过将四个样本说明快照用作示例，您将了解进行调整对于数据库性能的重要性。

与说明快照相关联的查询被编号为 1 到 4。每个查询都使用同一 SQL 或 XQuery 语句（在第 1 课中描述）：

```
SELECT S.ID,SNAME,O.DEPTNAME,SALARY+COMM
FROM ORG O, STAFF S
WHERE
  O.DEPTNUMB = S.DEPT AND
  S.JOB <> 'Mgr' AND
  S.SALARY+S.COMM > ALL( SELECT ST.SALARY*.9
                        FROM STAFF ST
                        WHERE ST.JOB='Mgr' )
ORDER BY S.NAME
```

但是，查询的每次迭代都比先前执行使用了更多的调整技巧。例如，“查询 1”没有进行性能调整，而“查询 4”使用的调整技巧最多。各个查询之间的区别描述如下：

#### 查询 1

运行没有索引和统计信息的查询

#### 查询 2

收集查询中的表和索引的当前统计信息

#### 查询 3

对用于在查询中连接表的列创建索引

#### 查询 4

对表列创建其他索引

这些示例是在具有 7 个物理节点的 RS/6000 SP 机器上使用分区间并行性创建的。

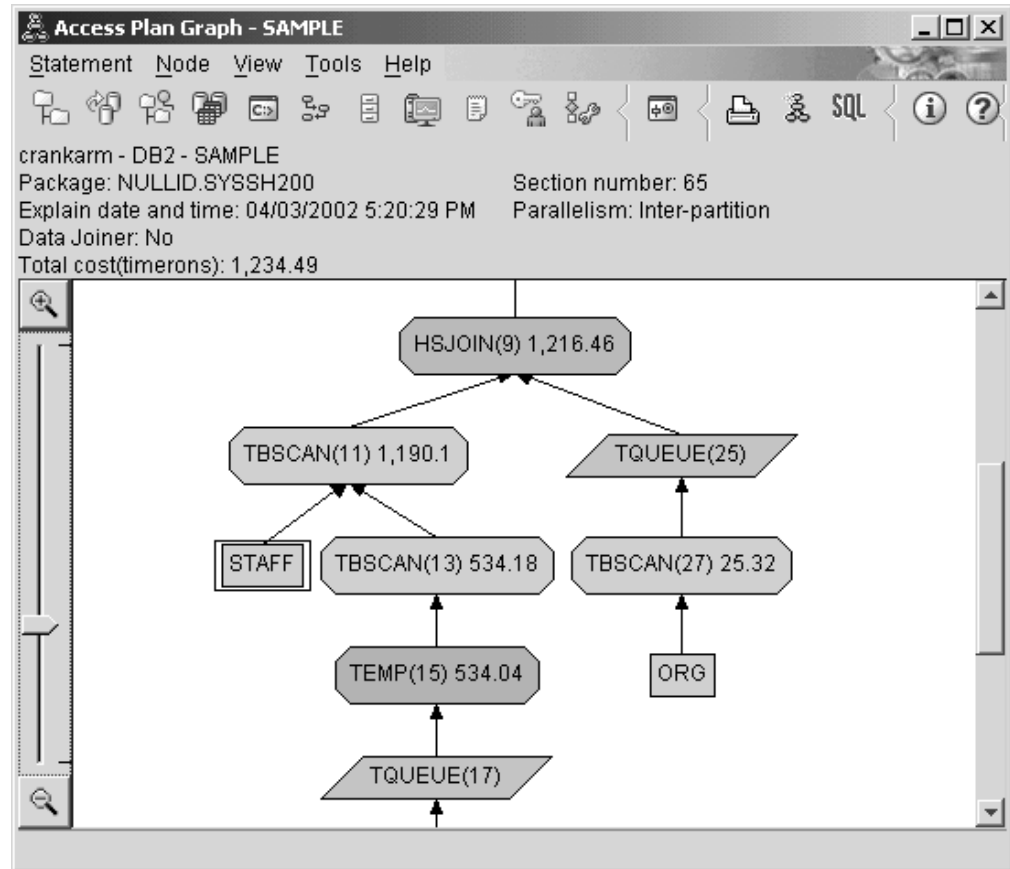
### 在分区数据库环境中运行没有索引和统计信息的查询

在此示例中，为没有索引和统计信息的 SQL 查询创建了访问方案。

要查看此查询（查询 1）的访问方案图：

1. 在控制中心中，展开对象树直到找到 SAMPLE 数据库为止。
2. 右键单击该数据库并从弹出菜单中选择显示说明语句的历史记录。  
“说明语句的历史记录”窗口打开。

3. 双击标识为“查询号 1”的条目（可能需要滚动到右边才能找到查询号列）。  
语句的“访问方案图”窗口打开。



对下列问题的回答将有助于您了解如何改进查询:

1. 查询中的每个表都存在当前统计信息吗?

要检查查询中的每个表的当前统计信息是否存在，请双击访问方案图中的每个表节点。在打开的对应“表统计信息”窗口中，说明列下面的 **STATS\_TIME** 行包含“统计信息未更新”字样，指示创建快照时未收集任何统计信息。

如果当前统计信息不存在，优化器将使用缺省统计信息，该统计信息可能与实际统计信息不同。缺省统计信息由“表统计信息”窗口中说明列下面的“缺省值”字样来标识。

根据“表统计信息”窗口中关于 **ORG** 表的信息，优化器使用了缺省统计信息（在说明值旁边指示）。因为创建快照时实际统计信息不可用，所以使用缺省统计信息（如 **STATS\_TIME** 行中所示）。

| Statistics       | Explained              | Current                |
|------------------|------------------------|------------------------|
| CREATE_TIME      | 03/26/2002 1:35:42 PM  | 03/26/2002 1:35:42 PM  |
| STATS_TIME       | Statistics not updated | Statistics not updated |
| CARD             | 55(default)            | -1                     |
| NPAGES           | 1(default)             | -1                     |
| FPAGES           | 1(default)             | -1                     |
| COLCOUNT         | 5(default)             | 5                      |
| OVERFLOW         | 0(default)             | -1                     |
| TABLESPACE       | USERSPACE1             | USERSPACE1             |
| INDEX_TABLESPACE |                        |                        |
| LONG_TABLESPACE  |                        |                        |
| VOLATILE         | No(default)            | No                     |

2. 此访问方案使用了访问数据的最有效方法吗？

此访问方案包含表扫描而非索引扫描。表扫描显示为八边形并标记为 TBSCAN 运算符。如果已使用了索引扫描，它们会以菱形显示并标记为 IXSCAN。如果抽取少量数据，那么使用为表创建的索引比表扫描的成本更低。

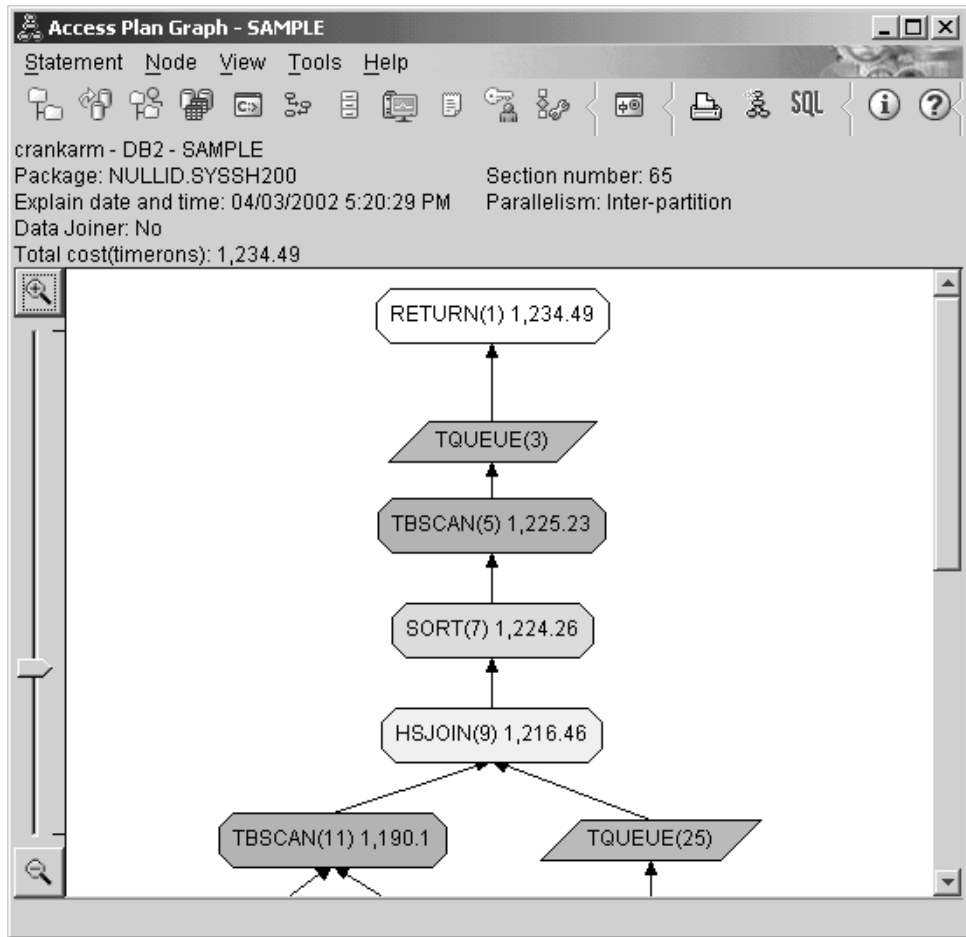
3. 此方案的效果如何？

仅当访问方案基于实际统计信息时才能确定它的效果。由于优化器使用访问方案中的缺省统计信息，所以不能确定该方案的效果如何。

一般来说，应该记下访问方案的总估计成本，以便日后与经过修订的访问方案进行比较。每个节点中列示的成本为累计成本，从查询的第一步开始算起，直到该节点（并包括该节点）。

**注：**对于分区数据库，这是使用最多资源的节点的累计成本。

在“访问方案”窗口中，总成本大约为 1,234 timeron，显示在图顶部的 **RETURN (1)** 中。总估计成本也显示在该窗口的顶部区中。



## 下一步如何操作

移至查询 2。

在运行 `runstats` 之后，“查询 2”查看基本查询的访问方案。使用 `runstats` 命令为优化器提供有关查询访问的所有表的当前统计信息。

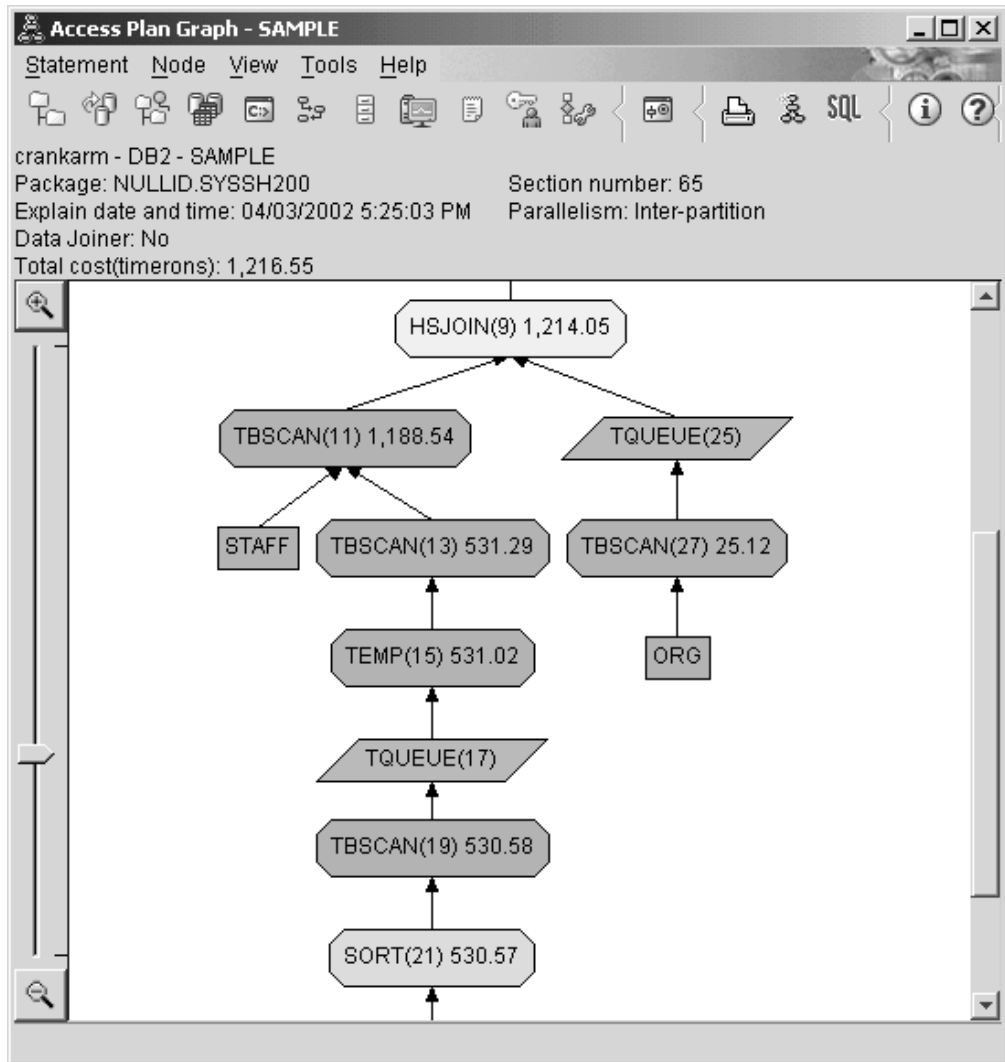
## 在分区数据库环境中使用 `runstats` 来收集表和索引的当前统计信息

通过使用 `runstats` 命令收集当前统计信息，并根据在“查询 1”中描述的访问方案来构建此示例。

强烈建议您使用 `runstats` 命令来收集关于表和索引的当前统计信息，特别当自上次执行 `runstats` 命令以来已经发生了重大的更新活动或已创建了新索引时。这会为优化器提供最准确的信息来确定最佳的访问方案。如果当前统计信息不可用，优化器会根据不准确的缺省统计信息选择一个低效的访问方案。

一定要在更新表之后使用 `runstats`；否则优化器可能会认为该表是空的。若在“运算符详细信息”窗口上的基数等于零，那么此问题很明显。在此情况下，完成表的更新，重新运行 `runstats` 命令并重新创建受影响的表的说明快照。

要查看此查询（查询 2）的访问方案图：在“说明语句历史记录”窗口中，请双击标识为“查询编号 2”的条目。将对此语句执行打开“访问方案图”。



对下列问题的回答将有助于您了解如何改进查询。

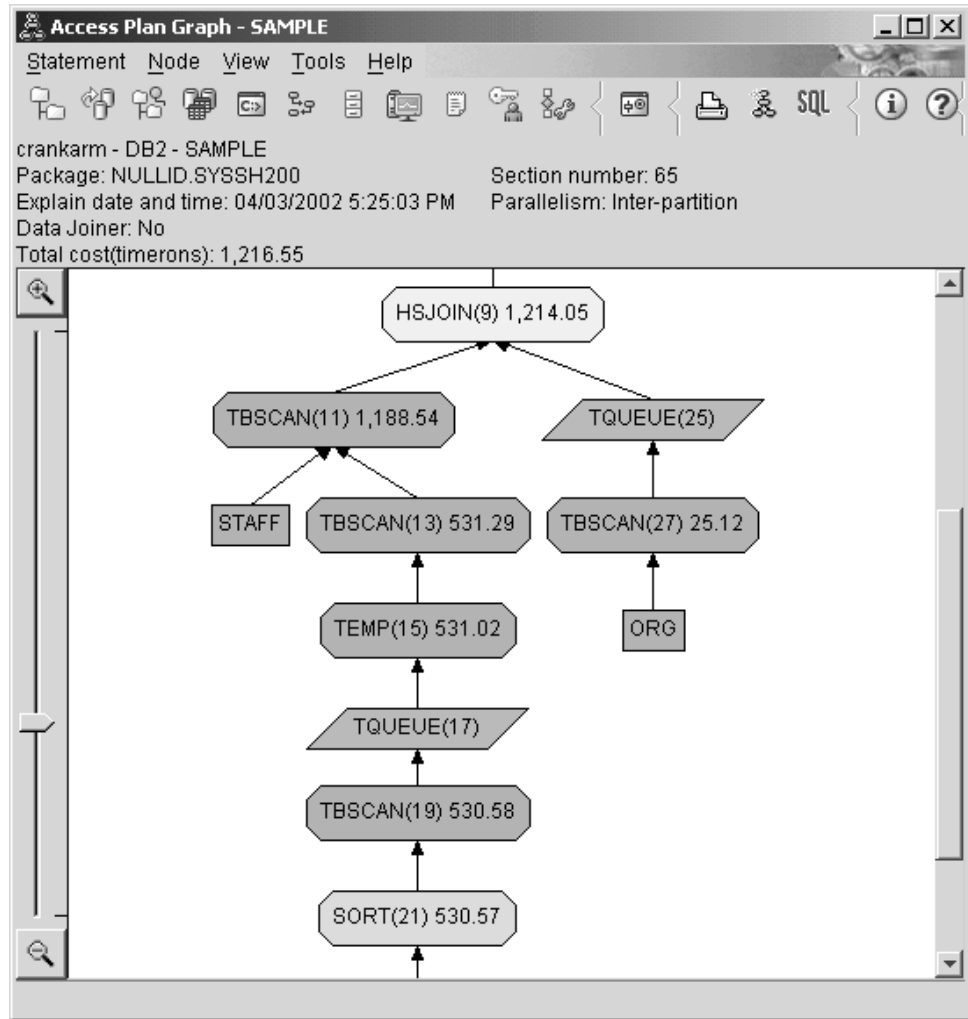
1. 查询中的每个表都存在当前统计信息吗？

ORG 表的“表统计信息”窗口显示优化器使用实际统计信息（STATS\_TIME 值是收集统计信息的实际时间）。统计信息的准确性取决于自从运行 runstats 命令以来是否对表的内容作了重大更改。

| Statistics       | Explained             | Current               |
|------------------|-----------------------|-----------------------|
| CREATE_TIME      | 03/26/2002 1:35:42 PM | 03/26/2002 1:35:42 PM |
| STATS_TIME       | 04/03/2002 5:24:55 PM | 04/03/2002 5:24:55 PM |
| CARD             | 4                     | 8                     |
| NPAGES           | 1                     | 2                     |
| FPAGES           | 1                     | 2                     |
| COLCOUNT         | 5                     | 5                     |
| OVERFLOW         | 0                     | 0                     |
| TABLESPACE       | USERSPACE1            | USERSPACE1            |
| INDEX_TABLESPACE |                       |                       |
| LONG_TABLESPACE  |                       |                       |
| VOLATILE         | No                    | No                    |

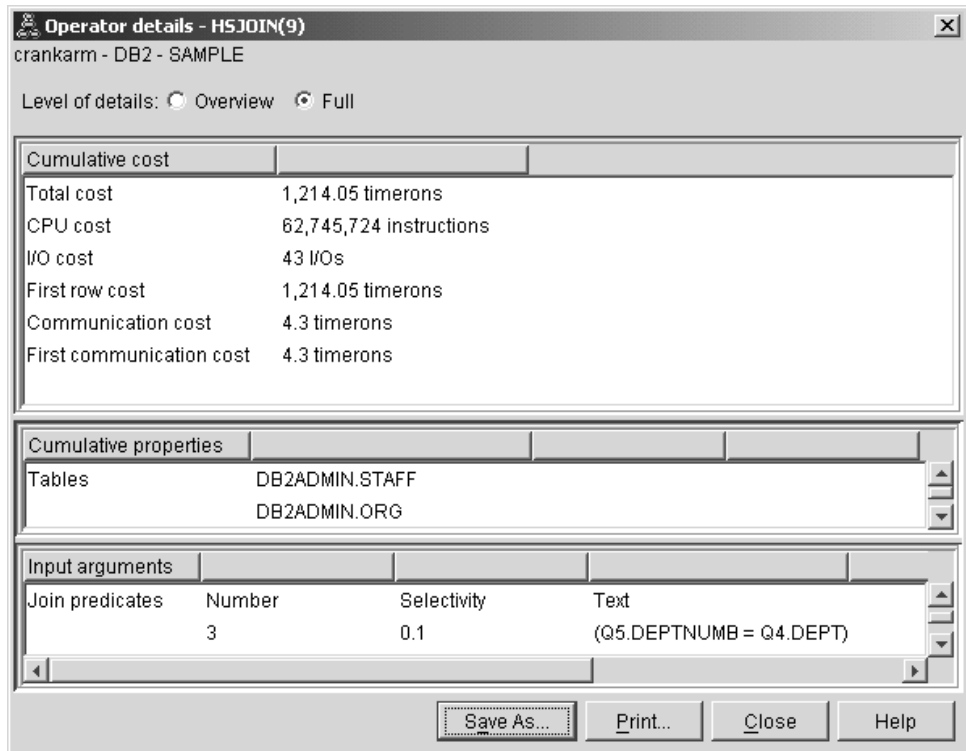
2. 此访问方案使用了访问数据的最有效方法吗？

与“查询 1”一样，“查询 2”中的访问方案使用表扫描（TBSCAN 运算符）而不是索引扫描（IXSCAN）。尽管当前统计信息存在，但由于查询所用的列上没有索引，所以未执行索引扫描。改进查询的一个方法是给优化器提供用来连接表的列上（即连接谓词）。在此示例中，这是第一个合并扫描连接：HSJOIN (9)。



在 HSJOIN (9) 运算符的“运算符详细信息”窗口中，请查看输入自变量下面的连接谓词一节。在此连接操作中所使用的列列示在文本列下面。在此示例中，这些列为 DEPTNUMB 和 DEPT。





### 3. 此访问方案的效果如何？

基于最新统计信息的访问方案始终会产生实际的估计成本（以 timeron 计）。由于“查询 1”中的估计成本基于缺省统计信息，所以不能比较两个访问方案图的成本来确定哪一个访问方案更有效。与成本的高低无关。必须比较基于实际统计信息的访问方案的成本，才能得到准确的有效性评估。

### 下一步如何操作

移至查询 3。

“查询 3”查看对 DEPTNUMB 和 DEPT 列添加索引的效果。对连接谓词中使用的列添加索引可提高性能。

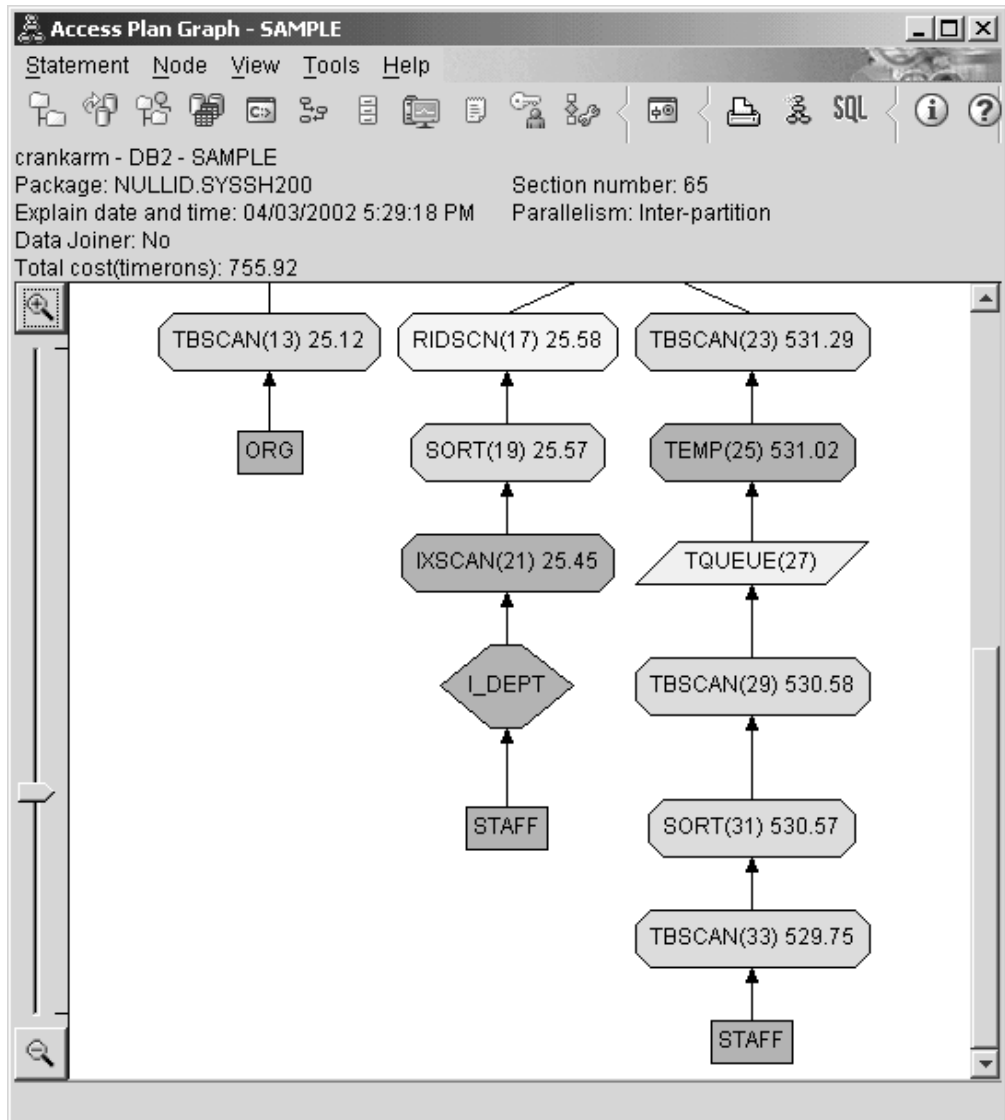
## 对分区数据库环境的查询中用于连接表的列创建索引

此示例通过对 STAFF 表上的 DEPT 列以及 ORG 表上的 DEPTNUMB 列创建索引，并根据“查询 2”中描述的访问方案而构建。

**注：**可使用设计顾问程序来创建建议的索引。

要查看此查询（查询 3）的访问方案图：在“说明语句历史记录”窗口中，请双击标识为“查询编号 3”的条目。将对此语句执行打开“访问方案图”。

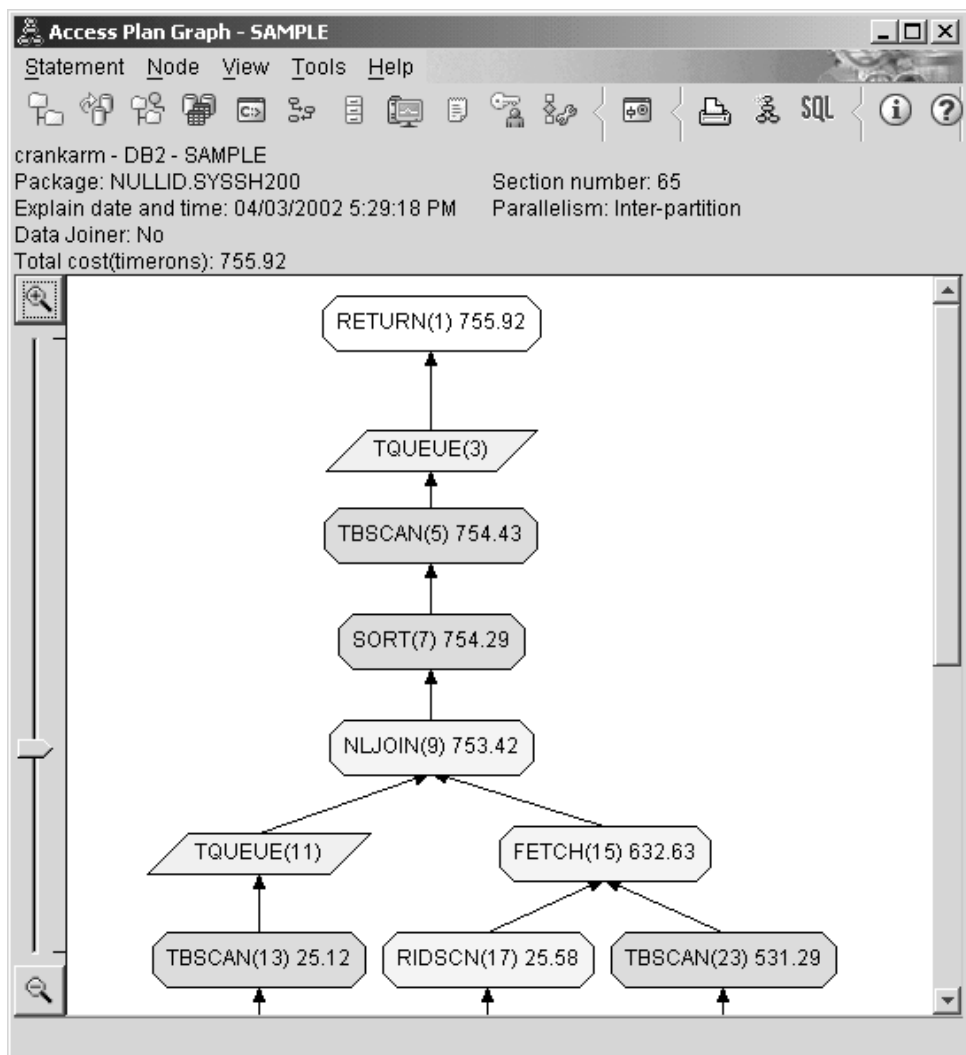
**注：**尽管创建了 DEPTNUM 的索引，但是优化器不会使用它。



对下列问题的回答将有助于您了解如何改进查询。

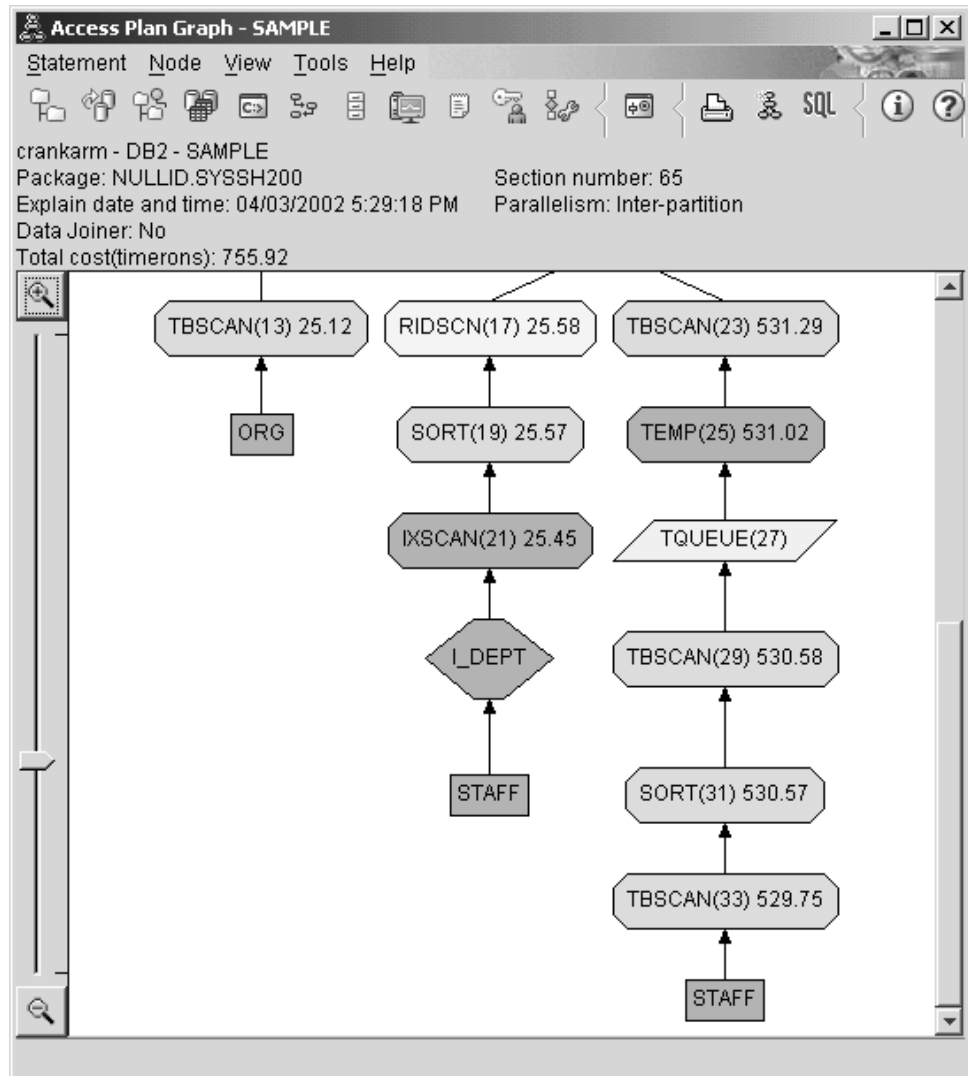
1. 在带有索引的访问方案中更改了什么内容？

已将一个新的菱形节点 **I\_DEPT** 添加在 **STAFF** 表上面。此节点表示对 **DEPT** 创建的索引，它显示优化器使用的是索引扫描而不是表扫描来确定要检索哪些行。

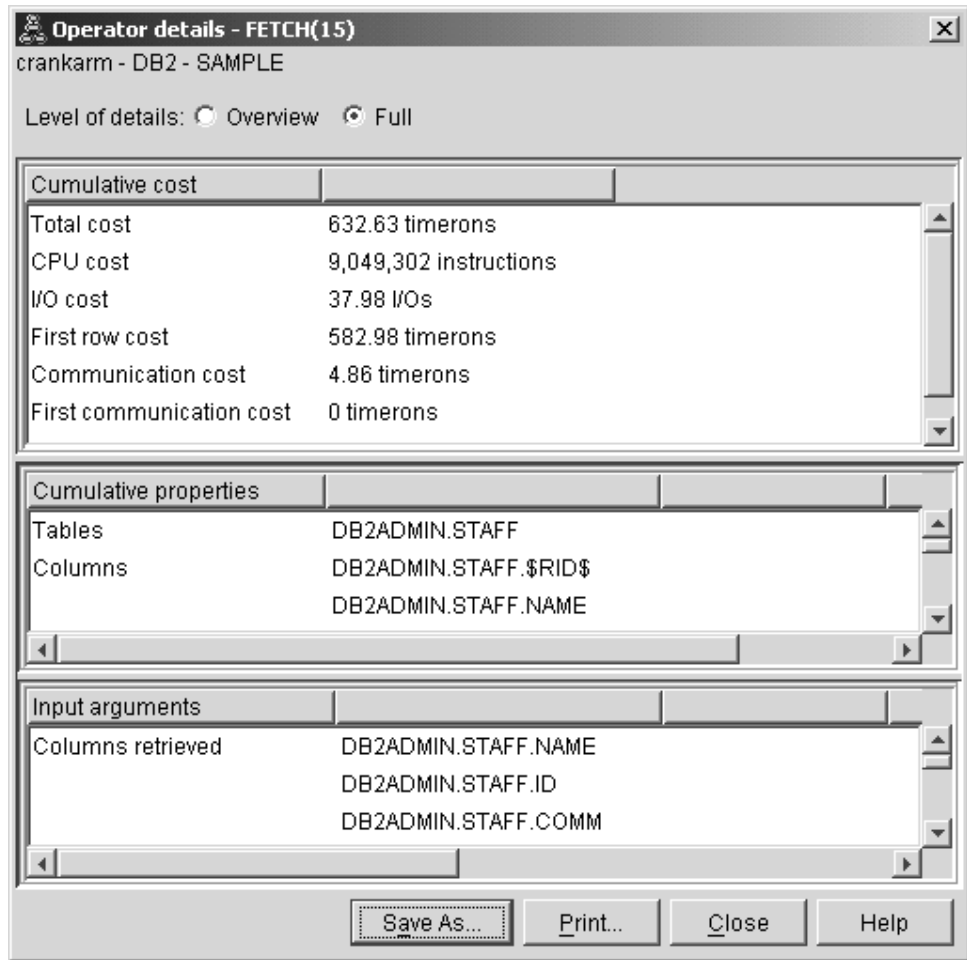


2. 此访问方案使用了访问数据的最有效方法吗？

查询的访问方案显示出创建有关 ORG 表的 DEPTNUMB 列的索引的效果，结果用 FETCH (15) 和 IXSCAN (21) 表示，而在 STAFF 表的 DEPT 列上。“查询 2”没有此索引；因此，在该示例中使用了表扫描。



FETCH (15) 运算符的“运算符详细信息”窗口显示此操作正在使用的列。



计算出索引和访存组合在一起的成本比先前访问方案中使用的完整表扫描的成本要低。

### 3. 此访问方案的效果如何？

此访问方案比先前示例中的访问方案的成本更低。累计成本已从“查询 2”中的大约 1,214 timeron 减少至“查询 3”中的大约 755 timeron。

### 下一步如何操作

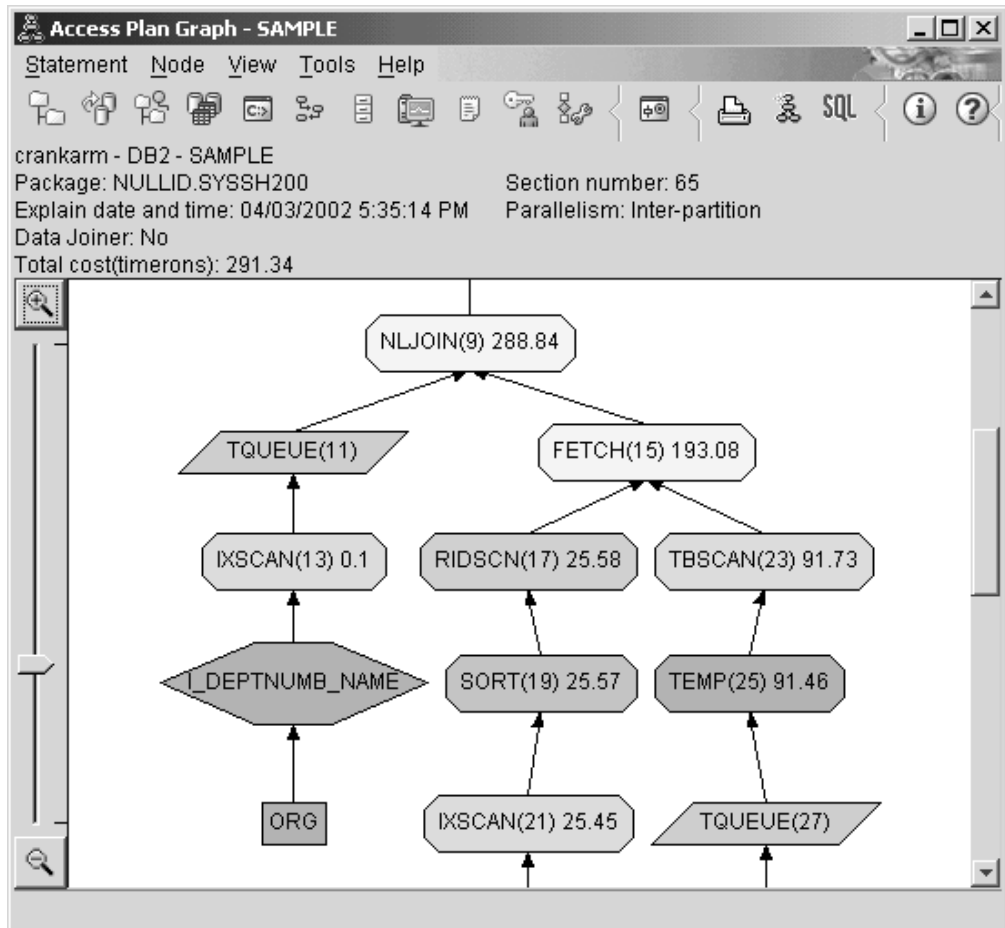
移至查询 4。

“查询 4”将访存和索引扫描减少至没有访存的单个索引扫描。创建其他索引可能会降低访问方案的估计成本。

## 对分区数据库环境中的表列创建其他索引

此示例通过对 STAFF 表中的 JOB 列创建索引并将 DEPTNAME 添加至 ORG 表的现有索引中（添加单独的索引可能会导致额外的访问），并根据在“查询 3”中描述的访问方案而构建。

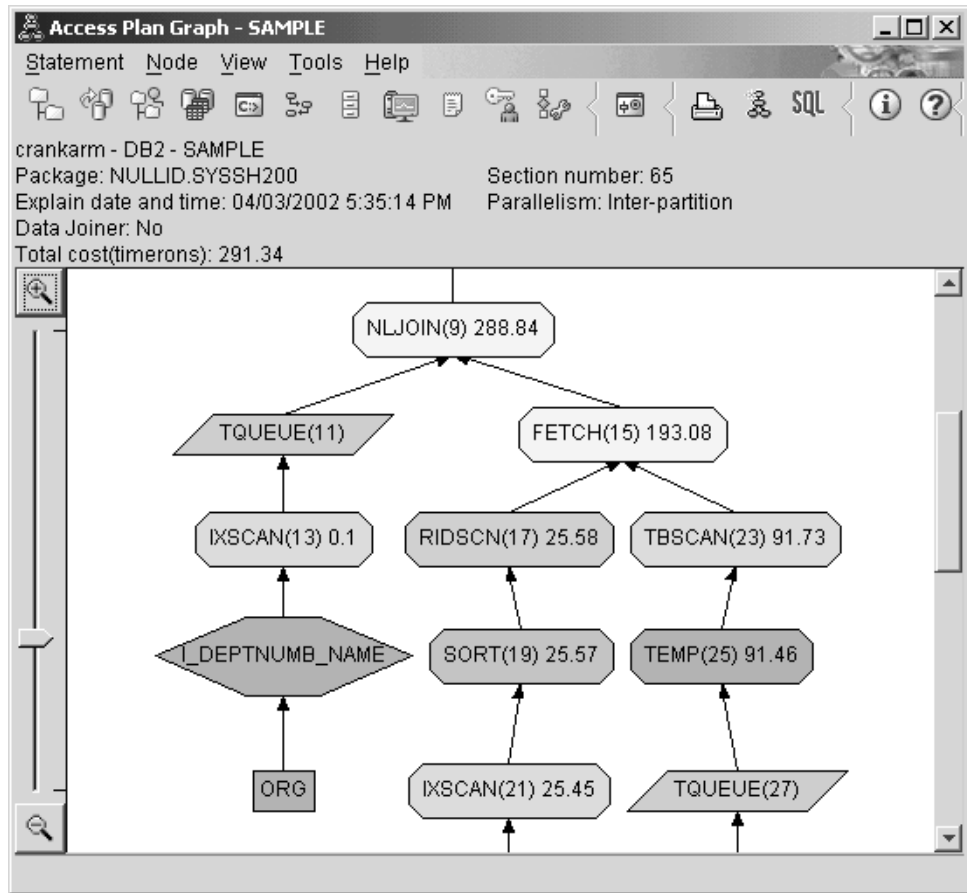
要查看此查询（查询 4）的访问方案图：在“说明语句历史记录”窗口中，请双击标识为“查询编号 4”的条目。将对此语句执行打开“访问方案图”。



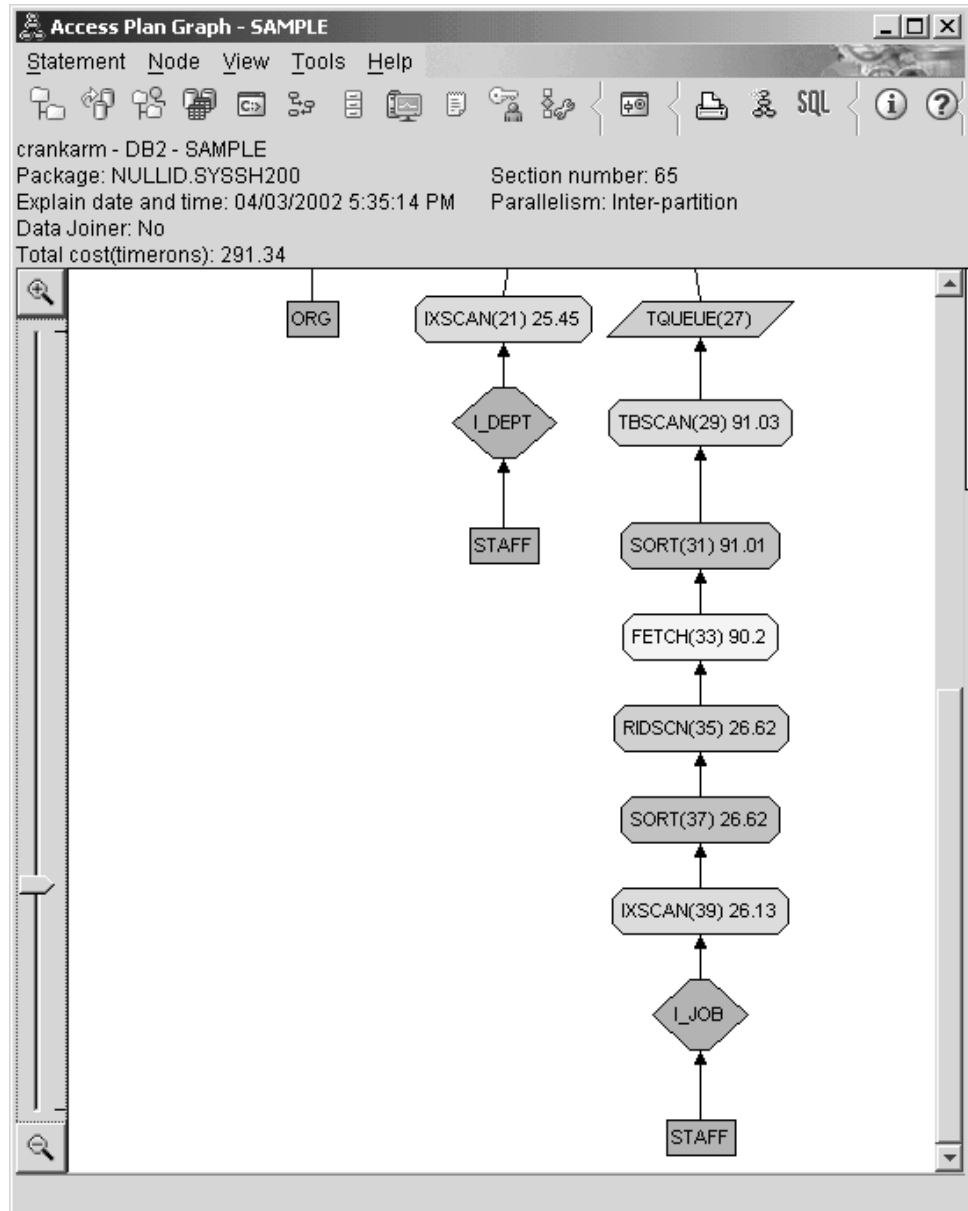
对下列问题的回答将有助于您了解如何改进查询。

1. 此访问方案由于创建附加索引而更改了哪些内容？

注意，在访问方案图的中间部分，对于 `ORG` 表，先前的表扫描已更改为索引扫描 `IXSCAN (7)`。将 `DEPTNAME` 列添加至 `ORG` 表的索引已使优化器可改进涉及表扫描的访问。



注意，在访问方案图的底部，对于 STAFF 表，之前的索引扫描和访存已更改为仅索引扫描 IXSCAN (39)。对 STAFF 表创建 JOB 索引已使优化器不必为进行访存而做额外的访问工作。



## 2. 此访问方案的效果如何？

此访问方案比先前示例中的访问方案的成本更低。累计成本已从“查询 3”中的大约 753 timeron 减少至“查询 4”中的大约 288 timeron。

## 下一步如何操作

提高您自己的 SQL 或 XQuery 语句的性能。

请参阅 *DB2 信息中心* 以了解有关可用于提高性能的其他步骤的详细信息。然后可返回至 *Visual Explain* 以了解操作带来的影响。



---

## 第 26 章 数据再分发

要在分区数据库环境中的各数据库分区中再分发表数据，可使用 `REDISTRIBUTE DATABASE PARTITION GROUP` 命令。

**注：**在先前版本中，此命令使用了 `NODEGROUP` 关键字，而不是 `DATABASE PARTITION GROUP` 关键字。

在分区数据库中，可能由于下列原因而再分发数据：

- 为了在数据库分区之间平衡数据量和处理负载。

如果可以将数据访问分布在多个数据库分区上，那么性能会提高。

- 为了在跨数据库分区的数据分发中引入数据偏差。

如果再分发频繁访问的表中的数据，以便不经常访问的数据位于数据库分区组中的少量数据库分区中，而频繁访问的数据分布在更多的数据库分区中，那么访问和吞吐量性能可能会提高。这将提高最频繁运行的应用程序的访问性能和吞吐量。

要保留表并置，可使用 `REDISTRIBUTE DATABASE PARTITION GROUP` 命令在数据库分区组级别再分发数据。在单个操作中再分发所有表。要完成指定的数据分发，`REDISTRIBUTE DATABASE PARTITION GROUP` 命令在它移动行时将表分到各个数据库分区上。根据您的指定的选项，实用程序可生成目标分发映射或将一个现有的分发映射用作输入。

**注：**在执行 `Redistribute` 实用程序之前进行脱机数据库备份是很重要的。

### 如何在数据库分区之间再分发数据

在数据库的指定数据库分区组中的一组表上执行数据再分发。在执行 `REDISTRIBUTE DATABASE PARTITION GROUP` 命令来调用“数据再分发”实用程序之前，必须连接到目录数据库分区的数据库。该实用程序使用源分发映射和目标分发映射来标识已将哪些散列数据库分区指定给新位置，它是新的数据库分区号。与具有新位置的一个数据库分区对应的所有行从源分发映射中指定的数据库分区移动到目标分发映射中指定的数据库分区。

“数据再分发”实用程序执行下列步骤：

1. 获取目标分发映射的新分发映射标识，并将它插入到 `SYSCAT.PARTITIONMAPS` 目录视图中。
2. 使用新的分发映射标识更新数据库分区组的 `SYSCAT.DBPARTITIONGROUPS` 目录视图中的 `REBALANCE_PMAP_ID` 列。
3. 将任何新的数据库分区添加至 `SYSCAT.DBPARTITIONGROUPDEF` 目录视图。
4. 如果指定了 `DROP` 命令选项，对要删除的任何数据库分区，将 `SYSCAT.DBPARTITIONGROUPDEF` 目录视图中的 `IN_USE` 列设置为“D”。
5. 执行 `COMMIT` 以更新目录。
6. 如果指定了 `ADD` 命令选项，那么为所有新的数据库分区创建数据库文件。
7. 对于数据库分区组中的每个表，按下列步骤再分发每个表中的数据：

- a. 如果该实用程序尚未使表空间进入 BACKUP PENDING 状态，那么就使它们进入此状态。
  - b. 在 SYSTABLES 目录表中锁定关于该表的那一行。
  - c. 使涉及此表的所有程序包失效。因为再分发表行，所以与该表关联的分发映射标识更改。因为程序包无效，所以编译器必须获得有关该表的新的数据库分区信息，并相应地生成程序包。
  - d. 以超级互斥方式锁定表（使用 Z 锁定）。
  - e. 使用批量数据移动操作再分发数据。
  - f. 如果再分发操作成功，那么在 SYSCAT.TABLES 中就更新了表的分发映射标识，并且该实用程序将对表发出 COMMIT 并继续处理数据库分区组中的下一个表。如果在完全再分发表之前该操作失败，那么实用程序失败，并使任何部分再分发的表处于 REDIST\_IN\_PGRS 状态，并且在再分发继续或异常终止前表不可访问。
8. 对于先前标记为要删除的数据库分区，删除数据库文件并删除 SYSCAT.NODEGROUPS 目录视图中的条目。
  9. 更新 SYSCAT.NODEGROUPS 目录视图中的数据库分区组记录，以将 PMAP\_ID 设置为 REBALANCE\_PMAP\_ID 的值，并将 REBALANCE\_PMAP\_ID 设置为空。
  10. 从 SYSCAT.PARTITIONMAPS 目录视图中删除旧的分发映射。
  11. 对所有更改执行 COMMIT。

---

## 对 Redistribute 实用程序的限制

某些限制适用于 Redistribute 实用程序。

- 当该实用程序正在运行时，不能执行下列操作：
  - 由于该实用程序不可前滚恢复，因此强烈建议您不要在执行再分发操作时更新数据库中的表，除非可反复更新此更新源。在这种情况下，“可反复更新”意味着以后发生再分发故障时都可以再次应用更新。当发生再分发故障时，将要求您从运行该实用程序之前所作的数据库备份进行复原。有关此问题的充分讨论，请参阅 REDISTRIBUTE DATABASE PARTITION GROUP 命令。
  - 在数据库分区组上启动另一个再分发操作。
  - 对数据库分区组中的任何表执行 ALTER TABLE 语句。
  - 删除数据库分区组。
  - 改变数据库分区组。
- 对于类型（层次结构）表，如果对 REDISTRIBUTE DATABASE PARTITION GROUP 命令指定了 TABLE ONLY 选项，那么只能在表列表中指定根表。不能显式指定任何子表。
- 对于范围分区表，Redistribute 实用程序不支持在数据分区表的范围之间移动数据。但是，它支持在数据库分区之间移动数据。
- 对于分区表，只有在下列情况下才支持再分发数据：
  - 表在 systables.access\_mode 目录表中的访问方式为完全访问。
  - 表没有任何正在连接或拆离的分区。
- 如果在包含处于暂挂清除状态的已转出块的数据分区组中有 MDC 表，那么 REDISTRIBUTE DATABASE PARTITION GROUP 命令将失败。

- 不允许废弃处于“正在再分发”状态的表。要废弃处于此状态的表，首先对适当的表列表运行附带有 ABORT 或 CONTINUE 选项的 REDISTRIBUTE DATABASE PARTITION GROUP 实用程序，以便完成或异常终止该表的再分发任务。

---

## 确定是否要再分发数据

决定要再分发数据之前，查明数据在数据库分区之间是否分布不均匀。在有了当前分布信息之后，可以使用此信息来创建定制再分发文件或分发映射。

要获取关于数据库分区组中数据库分区的当前数据分发的信息：

1. 确定是否有任何数据库分区具有不相等的行分布。

对于最大的表，使用适当的数据库分区列并输入如下查询：

```
SELECT PARTITION(column_name), COUNT(*) FROM table_name
GROUP BY PARTITION(column_name)
ORDER BY PARTITION(column_name) DESC
FETCH FIRST 100 ROWS ONLY
```

PARTITION 和 DBPARTITIONNUM SQL 函数确定散列分区或数据库分区上的当前数据分布。PARTITION 函数返回表中每行的分布映射索引。DBPARTITIONNUM 函数返回行的数据库分区号。

2. 对在数据库分区组上分发的其他大表执行此查询。
3. 使用此信息来创建一个分布文件和一个目标分布图。

---

## 在数据库分区之间再分发数据

在分区数据库环境中，当某些数据库分区包含的数据比其他分区要多时，或者某些数据库分区被访问的频率比其他分区要高时，可以在数据库分区之间再分发数据以平衡数据访问。

**重复的具体化查询表：** 如果数据库分区组中的数据包含重复的具体化查询表，那么在再分发数据之前必须删除这些表。再分发数据之后，可以重新创建具体化查询表。

由于 Redistribute 实用程序不可前滚恢复，因此，当 Redistribute 实用程序首次访问某个表空间时，该表空间就会进入 BACKUP PENDING 状态，这就要求在再分发操作结束时备份对表空间所作的更改。在执行再分发操作期间，被再分发的数据库分区组中的表是只读的，并且进行再分发的任何表都是完全不可访问的。建议您在执行再分发操作期间不要更新数据库分区组中未进行再分发的表，除非您完全了解这样做将产生的后果。有关此命令的不可前滚恢复特性的充分讨论，请参阅 REDISTRIBUTE DATABASE PARTITION GROUP 主题。

当该实用程序正在运行时，不能执行下列操作：

- 对数据库分区组启动另一个再分发操作
- 对数据库分区组中任何一个表执行 ALTER TABLE 语句
- 删除数据库分区组
- 改变数据库分区组。

除非所有受影响的表都有分布键，否则在将数据库分区添加至单一分区系统之后，不能使用此过程来再分发数据。REDISTRIBUTE DATABASE PARTITION GROUP 命令

依赖于分布键来再分发数据。分布键是在多分区数据库分区组中创建表时自动生成的，也可使用 CREATE TABLE 或 ALTER TABLE SQL 语句显式将其定义。如果表是在单一分区组中创建的，并且没有在 CREATE TABLE SQL 语句中定义分布键，那么不会定义分布键。在再分发数据之前，必须使用 ALTER TABLE SQL 语句来为每个受影响的表创建分布键。

要在数据库分区组中的数据库分区之间再分发数据：

1. 对数据库进行脱机备份。
2. 连接至包含系统目录表的数据库分区。
3. 执行先决条件任务（如果有必要的话）。
4. 发出 REDISTRIBUTE DATABASE PARTITION GROUP 命令。

**注：**在先前版本的 DB2 中，此命令使用了 NODEGROUP 关键字，而不是 DATABASE PARTITION GROUP 关键字。

指定下列自变量：

**数据库分区组名**

必须指定要在其中再分发数据的数据库分区组。

**UNIFORM**

如果数据已均匀分布且要保持均匀分布，那么指定 UNIFORM 或省略任何分布类型的自变量。UNIFORM 是缺省值。

**USING DISTFILE distfile-name**

要指定改正或创建数据偏差的定制分布，请包括分布文件名。“再分发数据”实用程序使用此文件来构造目标分发映射。

**USING TARGETMAP targetmap-name**

“再分发数据”实用程序直接使用指定的目标映射。

有关详细信息，参阅 REDISTRIBUTE DATABASE PARTITION GROUP 命令行实用程序信息。

5. 再分发完成之后：
  - 备份数据库分区组中所有处于 BACKUP PENDING 状态的表空间。也可以执行完整数据库备份。
  - 重新创建在再分发之前废弃的任何重复的具体化查询表。
  - 如果对 REDISTRIBUTE 命令指定了 STATISTICS NONE 选项，并且数据库分区组中有处理统计信息概要文件的表，那么现在就执行 RUNSTATS 命令来收集数据分布统计信息，以便 SQL 编译器和优化器在为查询选择数据访问方案时使用。

**注：**说明表包含关于用来再分发数据的分发映射的信息。

---

## 在数据库分区组中再分发数据

使用“再分发数据”向导为数据库分区组创建有效的再分发计划和再分发数据。首先选择再分发方法和策略，然后进行更高级的选择。

要使用数据库分区组，您必须具有 SYSADM 或 DBADM 权限。

要在数据库分区组中再分发数据：

1. 打开“再分发数据”向导：从控制中心中展开对象树，直到找到**数据库分区组**文件夹为止。所有现有数据库分区组都显示在窗口右边的内容窗格中。右键单击要使用的数据库分区组并从弹出菜单中选择**再分发**。“再分发数据”向导将打开。

还可以从使用“添加分区”启动板添加数据库分区或使用“删除分区”启动板删除数据库分区中打开“再分发数据”向导。

2. 完成每个适用的向导页。单击第一页上的向导概述链接以获取更多信息。指定了向导再分发数据所需的足够信息后，**完成**按钮才会启用。

---

## 数据再分发的日志空间要求

REDISTRIBUTE DATABASE PARTITION GROUP 实用程序的日志空间需求已急剧减少。当要再分发的数据库分区组中不存在长字段 (LF) 或大对象 (LOB) 数据时，实用程序为每个表生成的日志记录数相对于用户数据大小来说比较小，这是因为不会为移动的每一行都写入日志记录。但是，当存在 LF/LOB 数据时，就会为移动的每一行中的每个 LF/LOB 值生成日志记录。在这种情况下，期望每个数据库分区的日志空间需求大约是在该分区上移动的数据量（即，发送和/或接收的数据）的三分之一。无论接收分区上是否存在 LF/LOB 数据，都会写入一种日志记录，这种日志记录的数目取决于移动的数据量：扩展数据块分配日志记录。但是，这些日志记录需要的空间总量还是比较小，绝不会超过移动的用户数据总量的很小一部分。

---

## 再分发事件日志记录和错误恢复

REDISTRIBUTE DATABASE PARTITION GROUP 命令将有关它处理的每个表的信息转储到一个再分发事件日志文件中。对在运行的每个再分发操作都将创建一个再分发事件日志文件。事件日志文件的格式为 *database-name.database-partition-group-name.timestamp*，并且事件日志文件被写入下列目录：

- 在基于 Linux 和 UNIX 的系统上，事件日志文件被写入 `homeinst/sqllib/redistribute` 目录。
- 在 Windows 操作系统上，事件日志文件被写入 `DB2INSTPROF\instance\redistribute` 目录，其中 DB2INSTPROF 是 DB2INSTPROF 注册表变量的值。

例如，名称为 DB819.NG1.2007062419415651 的文件是当地时间 2007 年 6 月 24 日下午 7:41 对数据库分区组 NG1 中的数据库 DB819 执行再分发操作所产生的事件日志文件。

事件日志文件的三个主要用途为如下所示：

- 提供有关再分发操作的一般信息，例如，旧的分发映射和新的分发映射。
- 为用户提供一些信息，以帮助他们对迄今为止实用程序已经再分发了哪些表保持跟踪。
- 提供有关已经再分发的每个表的信息，这些信息包括：用于该表的建立索引方式、指示是否已成功再分发该表以及对表执行再分发操作的开始时间和结束时间。

### 再分发期间的故障恢复

正如上面提到的那样，再分发事件日志文件指示是否已成功再分发某个表。此信息对于从错误进行恢复很有用，这是因为可以查阅再分发事件日志文件来确定哪个表已导致 Redistribute 实用程序失败，当然假定该实用程序失败，直到处理表为止。例如，如

果再分发操作因遇到某个表处于受限制访问状态（例如，LOAD PENDING）而失败，那么很容易通过查阅事件日志文件来确定哪些表发生了问题。在解决该表存在的问题之后，用户可以重新提交具有下列其中一个选项的 REDISTRIBUTE DATABASE PARTITION GROUP 命令：

- CONTINUE - 此选项指示应继续执行再分发操作，直到再分发了原始 REDISTRIBUTE DATABASE PARTITION GROUP 命令中指定的所有表为止。
- ABORT - 此选项指示再分发操作应异常终止，并且迄今为止已再分发或者已部分再分发的所有表都应返回到以下状态：在数据库分区组上首次运行 REDISTRIBUTE DATABASE PARTITION GROUP 命令之前这些表所处的状态。

**注：**在可以使用任何一个选项之前，先前的数据再分发操作必须已经失败或已完成，但是未再分发数据库分区组中的所有表；如果使用了 TABLE 选项并且指定了一部分表，那么可以执行再分发。在这些情况下，SYSCAT.DBPARTITIONGROUPS 表中的 REBALANCE\_PMIID 列的值将不为 -1。

### 再分发事件日志文件条目：

下面是一些常见再分发事件日志文件条目以及对每个条目的简要描述的列表。

#### 示例 1:

```
2007-06-24-19.41.56.538982-240 I1268A835          LEVEL: Event
PID      : 1020784          TID   : 3857          PROC  : db2sysc
INSTANCE: newton          NODE  : 000          DB    : DB819
APPHDL  : 0-120          APPID: *N0.newton.070624234148
AUTHID   : NEWTON
EDUID    : 3857          EDUNAME: db2agent (DB819) 0
FUNCTION: DB2 UDB, relation data serv, sqlrdrin, probe:3126
CHANGE  : DB PART MAP ID : NG1 : FROM "4" : TO "3" : success
IMPACT  : None
DATA #1 : String, 24 bytes
HexDump Old Map Entries:
DATA #2 : Dumped object of size 8192 bytes at offset 83120, 61 bytes
/home4/newton/sqllib/redistribute/1020784.3857.000.dump.bin
DATA #3 : String, 24 bytes
HexDump New Map Entries:
DATA #4 : Dumped object of size 8192 bytes at offset 91432, 61 bytes
/home4/newton/sqllib/redistribute/1020784.3857.000.dump.bin
```

这是在执行再分发操作期间转储的第一个事件日志条目，它提供了有关实用程序将使用的分发映射文件（既包括旧文件也包括新文件）的信息。在这种特殊情况下，分区组 NG1 的旧分发映射文件的标识为 4，而新的分发映射的标识为 3。每个分发映射文件的十六进制转储目标位置是实例路径中的再分发目录，并且转储文件的名称包含在条目中。在这种情况下，可以在 /home4/newton/sqllib/redistribute/1020784.3857.000.dump.bin 文件中找到标识为 4（即，NG1 的旧分发映射文件）的分发映射。可以在 /home4/newton/sqllib/redistribute/1020784.3857.000.dump.bin 文件中找到标识为 3（即，NG1 的新分发映射文件）的分发映射。

#### 示例 2:

```
2007-06-24-19.41.56.596251-240 I2104A507          LEVEL: Event
PID      : 1020784          TID   : 3857          PROC  : db2sysc
INSTANCE: newton          NODE  : 000          DB    : DB819
APPHDL  : 0-120          APPID: *N0.newton.070624234148
AUTHID   : NEWTON
EDUID    : 3857          EDUNAME: db2agent (DB819) 0
FUNCTION: DB2 UDB, relation data serv, sqlrdrinInitLogfileInfo, probe:1425
```

```

START   : REDIST DB PART GROUP : NG1 : success
IMPACT  : None
DATA #1 : String, 11 bytes
Option: 'T'

```

此条目指示已成功启动再分发操作并且即将开始执行表的再分发。还给出了再分发选项。'T' 意味着用户已为再分发操作提供了他们自己的目标分发映射。其他可以使用的值包括: 'U', 此值指示将平均地再分发数据 (这是缺省值); 'C', 此值指示再分发 CONTINUE 操作; 'A', 此值指示再分发 ABORT 操作。

### 示例 3:

```

2007-06-24-19.41.56.700558-240 I2612A593          LEVEL: Event
PID      : 1020784          TID : 8145          PROC : db2sysc
INSTANCE: newton          NODE : 000          DB   : DB819
APPHDL  : 0-129          APPID: *N0.DB2.070624234157
AUTHID   : NEWTON
EDUID    : 8145          EDUNAME: db2agent (DB819) 0
FUNCTION: DB2 UDB, database utilities - Redistribute, sqlurRedistributeTableByID,
probe:9064
START    : REDIST TABLE : "NEWTON"."RDADHC13" : success
IMPACT   : None
DATA #1 : String, 26 bytes
Runstats with profile: '0'
DATA #2 : String, 18 bytes
Indexing Mode: '0'

```

此条目指示成功启动了再分发 "NEWTON"."RDADHC13" 表。Runstats with profile: '0' 条目指示统计信息收集选项为 STATISTICS USE PROFILE, 这是缺省统计信息收集方式, 这意味着如果表具有统计信息概要文件, 那么将根据此概要文件来收集统计信息。否则, 将不会收集统计信息。如果值为 '1', 那么指示指定了 STATISTICS NONE 选项, 这意味着无论是否对表定义了统计信息概要文件, 都不会为该表收集统计信息。此外, Indexing Mode: '0' 这一行指示表的建立索引方式为 AUTOSELECT, 这意味着实用程序将智能决定使用 INCREMENTAL 还是 REBUILD 建立索引方式。值 '1' 指示用户指定了 INCREMENTAL 建立索引方式, 值 '2' 指示用户指定了 DEFERRED 建立索引方式, 值 '3' 指示用户指定了 REBUILD 建立索引方式。

### 示例 4:

```

2007-06-24-19.42.02.370410-240 I3206A493          LEVEL: Event
PID      : 1020784          TID : 8145          PROC : db2sysc
INSTANCE: newton          NODE : 000          DB   : DB819
APPHDL  : 0-129          APPID: *N0.DB2.070624234157
AUTHID   : NEWTON
EDUID    : 8145          EDUNAME: db2agent (DB819) 0
FUNCTION: DB2 UDB, database utilities - Redistribute, sqlurRedistributeTableByID,
probe:9751
STOP     : REDIST TABLE : "NEWTON"."RDADHC13" : success
IMPACT   : None

```

此条目指示成功再分发了 "NEWTON"."RDADHC13" 表。如果在处理此表期间发生了错误, 那么此条目将指示故障。

### 示例 5:

```

2007-06-24-19.42.02.425600-240 I3700A455          LEVEL: Event
PID      : 1020784          TID : 3857          PROC : db2sysc
INSTANCE: newton          NODE : 000          DB   : DB819
APPHDL  : 0-120          APPID: *N0.newton.070624234148
AUTHID   : NEWTON

```

```
EDUID      : 3857                EDUNAME: db2agent (DB819) 0
FUNCTION: DB2 UDB, relation data serv, sqlrdrdt, probe:1034
STOP       : REDIST DB PART GROUP : NG1 : success
IMPACT     : None
```

此条目指示已成功完成再分发数据库分区组 NG1。如果未成功完成此操作，那么此条目将指示故障。

---

## STEPWISE\_REDISTRIBUTE\_DBPG 过程

可以使用 STEPWISE\_REDISTRIBUTE\_DBPG 存储过程安全地再分发数据库分区组。

1. 使用 *ANALYZE\_LOG\_SPACE* 过程 - 检索日志空间分析信息分析与日志空间可用性和数据偏差有关的数据库分区组。

*analyze\_log\_space* 函数返回日志空间分析结果的结果集（开放式游标），并且对于给定的数据库分区组的每个数据库分区，包含字段。

2. 使用 *GENERATE\_DISTFILE* 过程 - 生成数据分发文件为给定的表创建数据分发文件。

*generate\_distfile* 函数针对给定的表生成数据分发文件，并将它以给定的文件名保存。

3. 使用 *STEPWISE\_REDISTRIBUTE\_DBPG* 过程 - 再分发部分数据库分区组为数据库分区组创建并报告按步骤再分发计划的内容。
4. 使用 *GET\_SWRD\_SETTINGS* 过程 - 检索再分发信息和 *SET\_SWRD\_SETTINGS* 过程 - 创建或更改再分发注册表为给定的表创建数据分发文件。

*get\_swrd\_settings* 函数读取给定数据库分区组的现有的再分发注册表记录。

*set\_swrd\_settings* 函数创建或更改再分发注册表。如果该注册表不存在，那么将创建它并向其中添加记录。如果注册表已存在，那么使用 *overwriteSpec* 来标识要覆盖哪些字段值。*overwriteSpec* 字段使此函数能够将 NULL 输入用于不需要更新的字段。

5. 使用 *STEPWISE\_REDISTRIBUTE\_DBPG* 过程 - 再分发部分数据库分区组按照计划再分发数据库分区组。

*stepwise\_redistribute\_dbpg* 函数根据输入和设置文件来再分发数据库分区组的部分。

### 用法示例

以下是 AIX 上的 CLP 脚本的一个示例:

```
# -----
# Set the database you wish to connect to
# -----
dbName="SAMPLE"

# -----
# Set the target database partition group name
# -----
dbpgName="IBMDEFAULTGROUP"

# -----
# Specify the table name and schema
# -----
tbSchema="$USER"
tbName="STAFF"

# -----
```



```

# Specify the name of the data distribution file
# -----
distFile="$HOME/sqllib/function/$dbName.IBMDEFAULTGROUP_swrData.dst"

export DB2INSTANCE=$USER
export DB2COMM=TCPIP

# -----
# Invoke call statements in clp
# -----
db2start db2 -v "connect to $dbName"

# -----
# Analysing the effect of adding a database partition without applying the changes - a 'what if'
# hypothetical analysis
#
# - In the following case, the hypothesis is adding database partition 40, 50 and 60 to the
# database partition group, and for database partitions 10,20,30,40,50,60, using a respective
# target ratio of 1:2:1:2:1:2.
#
# NOTE: in this example only partitions 10, 20 and 30 actually exist in the database
# partition group
# -----
db2 -v "call sysproc.analyze_log_space('$dbpgName', '$tbSchema', '$tbName', 2, ' ',
'A', '40,50,60', '10,20,30,40,50,60', '1,2,1,2,1,2')"

# -----
# Analysing the effect of dropping a database partition without applying the changes
#
# - In the following case, the hypothesis is dropping database partition 30 from the database
# partition group, and redistributing the data in database partitions 10 and 20 using a
# respective target ratio of 1 : 1
#
# NOTE: In this example all database partitions 10, 20 and 30 should exist in the database
# partition group
# -----
db2 -v "call sysproc.analyze_log_space('$dbpgName', '$tbSchema', '$tbName', 2, ' ',
'D', '30', '10,20', '1,1')"

# -----
# Generate a data distribution file to be used by the redistribute process
# -----
db2 -v "call sysproc.generate_distfile('$tbSchema', '$tbName', '$distFile')"

# -----
# Write a step wise redistribution plan into a registry
#
# Setting the 10th parameter to 1, may cause a currently running step wise redistribute
# stored procedure to complete the current step and stop, until this parameter is reset
# to 0, and the redistribute stored procedure is called again.
# -----
db2 -v "call sysproc.set_swr_settings('$dbpgName', 255, 0, ' ', '$distFile', 1000,
12, 2, 1, 0, '10,20,30', '50,50,50')"

# -----
# Report the content of the step wise redistribution plan for the given database
# partition group.
# -----
db2 -v "call sysproc.get_swr_settings('$dbpgName', 255, ?, ?, ?, ?, ?, ?, ?, ?, ?)"

# -----
# Redistribute the database partition group "dbpgName" according to the redistribution
# plan stored in the registry by set_swr_settings. It starting with step 3 and
# redistributes the data until 2 steps in the redistribution plan are completed.
# -----
db2 -v "call sysproc.stepwise_redistribute_dbpg('$dbpgName', 3, 2)"

```



---

## 第 27 章 配置自调整内存

---

### 分区数据库环境中的自调整内存

在分区数据库环境中使用自调整内存功能时，有一些因素决定该功能是否能适当地调整系统。

在分区数据库中启用自调整内存功能时，会将一个数据库分区指定为调整分区，所有内存调整决定都根据该数据库分区的内存和工作负载特征作出。一旦对调整分区作出调整决定，就会将内存调整分布到所有其他数据库分区上，以确保所有数据库分区都维持类似的配置。

单个调整分区模型要求只对具有类似内存需求的数据库分区使用该功能。可以使用下列准则确定您的分区数据库上是否要启用自调整内存功能。

#### 建议在分区数据库中使用自调整的情况

当所有数据库分区都具有类似内存需求并且正在类似硬件上运行时，可以不进行任何修改就启用自调整内存。这些类型的环境共享下列特征：

- 所有数据库分区都在完全相同的硬件上，包括多逻辑节点平均分布在多个物理节点上
- 最佳或者接近最佳地分配数据
- 在数据库分区上运行的工作负载平均分布到各数据库分区上。这意味着没有一个数据库分区的一个或多个堆具有较多的内存需求。

在这种环境中，需要同等配置所有数据库分区，并且自调整内存将正确配置系统。

#### 建议在分区数据库中小心使用自调整的情况

在环境中的大多数数据库分区具有类似内存需求并且正在类似硬件上运行的情况下，可以使用自调整内存功能，但在进行初始配置时要小心。这些系统可能有一组数据库分区用于数据，并且有一组少得多的协调程序分区和目录分区。在这些环境中，将协调程序分区和目录分区配置为与包含数据的数据库分区不同可能会有好处。

在此环境中，如果进行一些较小的设置，仍可以通过使用自调整内存功能获得好处。由于包含数据的数据库分区由大量数据库分区组成，因此应该在所有这些数据库分区上启用自调整，并应将这此数据库分区中的一个分区指定为调整分区。此外，由于目录分区和协调程序分区的配置可能不同，因此应在这些分区上禁用自调整内存。要在目录分区和协调程序分区上禁用自调整，将这些分区上的 *self\_tuning\_mem* 数据库配置参数更新为 OFF。

#### 建议不要在分区数据库中使用自调整的情况

如果环境中的每个数据库分区的内存需求都不同，或者不同的数据库分区正在极不相同的硬件上运行，那么建议禁用自调整内存功能。这可以通过将所有分区上的 *self\_tuning\_mem* 数据库配置参数设置为 OFF 来实现。

## 比较不同数据库分区的内存需求

确定不同数据库分区的内存需求是否非常相似的最佳方法是查看快照监视器。如果所有分区上的下列快照元素相似（差别不超过 20%），那么认为这些分区相似。

通过发出 `get snapshot for database on <dbname>` 命令收集下列数据。

```
已分配的共享排序堆总数           = 0
共享排序堆高水位标记             = 0
后阈值排序数（共享内存）         = 0
排序溢出数                       = 0

程序包高速缓存查询数             = 13
程序包高速缓存插入数             = 1
程序包高速缓存溢出数             = 0
程序包高速缓存高水位标记（以字节计） = 655360

散列连接数                       = 0
散列循环数                       = 0
散列连接溢出数                   = 0
小散列连接溢出数                 = 0
后阈值散列连接数（共享内存）     = 0

当前挂起的锁定数                 = 0
锁定等待数                       = 0
数据库等待锁定的时间（毫秒）     = 0
正在使用的锁定列表内存（以字节计） = 4968
锁定升级数                       = 0
互斥锁定升级数                   = 0
```

通过发出 `get snapshot for bufferpools on <dbname>` 命令收集下列数据:

```
缓冲池数据逻辑读取数             = 0
缓冲池数据物理读取数             = 0
缓冲池索引逻辑读取数             = 0
缓冲池索引物理读取数             = 0
缓冲池总计读取时间（毫秒）       = 0
缓冲池总计写入时间（毫秒）       = 0
```

---

## 在分区数据库环境中使用自调整内存

在分区数据库环境中启用自调整功能之后，就会出现一个单独的数据库分区（称为调整分区），它监视内存配置的情况，并将任何配置更改传播到所有其他数据库分区以使所有参与数据库分区的配置保持一致。

调整分区是根据许多特征选择的，例如分区组中的数据库分区数以及已定义的缓冲池数。

- 要确定当前已指定为调整分区的数据库分区，请使用以下 ADMIN\_CMD:

```
CALL SYSPROC.ADMIN_CMD('get stmm tuning dbpartitionnum')
```

- 要更改调整分区，请使用以下 ADMIN\_CMD:

```
CALL SYSPROC.ADMIN_CMD('update stmm tuning dbpartitionnum <db_partition_num>')
```

发出此命令时，将以异步方式或者在数据库下次启动时更新调整分区。

- 要让内存调整器自动重新选择调整分区，请对 `<db_partition_num>` 值输入 `-1`。

## 在 DPF 系统上启动内存调整器

由于自调整功能要求所有分区都处于活动状态才能正确地调整多分区系统上的内存，所以，在 DPF 环境中，仅当使用显式 `ACTIVATE DATABASE` 命令激活数据库时，才会启动内存调整器。

### 对给定数据库分区禁用自调整功能

- 要对一部分数据库分区禁用自调整功能，请将不想调整的数据库分区的 `self_tuning_mem` 配置参数设置为 `OFF`。

- 

要对特定数据库分区上由配置参数控制的一部分内存使用者禁用自调整功能，请在该数据库分区上将相关配置参数值或缓冲池大小设置为 `MANUAL` 或特定值。但是，建议使自调整配置参数值在所有正在运行的分区中一致。

- 要对某个数据库分区上的特定缓冲池禁用调整功能，请发出指定了大小值的 `ALTER BUFFER POOL` 命令并指定 `PARTITIONNUM` 参数值，以指示要禁用自调整功能的分区。

使用 `PARTITIONNUM` 子句在特定数据库分区上指定大小的 `ALTER BUFFERPOOL` 语句将在 `SYSCAT.SYSBUFFERPOOLNODES` 目录中为给定缓冲池创建例外条目，如果例外条目已存在，那么此语句将更新该条目。如果缺省缓冲池大小设置为 `AUTOMATIC`，那么当此目录中的某个缓冲池存在例外条目时，该缓冲池将不会参与自调整操作。要除去例外条目，以便可以对缓冲池重新启用自调整功能：

1. 通过发出 `ALTER BUFFERPOOL` 语句并将缓冲池大小设置为特定值来对此缓冲池禁用调整功能。
2. 发出另一个 `ALTER BUFFERPOOL` 语句并指定 `PARTITIONNUM` 子句，以便将此数据库分区上缓冲池的大小设置为缺省缓冲池大小。
3. 发出另一个 `ALTER BUFFERPOOL` 语句，将大小设置为 `AUTOMATIC`，从而启用自调整功能。

### 在不均匀的环境中启用自调整内存

理想情况下，数据应该均匀地分布在所有数据库分区中，每个分区上运行的工作负载的内存需求应该比较接近。如果数据分布不均匀，以致一个或多个数据库分区包含的数据显著多于或少于其他数据库分区，就不应该对这些不规则的数据库分区启用自调整功能。这也适用于不同数据库分区上的内存需求不均匀的情况。例如，如果只在一个分区上执行需要大量资源的排序操作，或者某些数据库分区使用的硬件与别的分区不同并且有更多的可用内存，就会发生这种情况。在此类环境中，仍然可以对某些数据库分区启用自调整功能。要在不均匀环境中利用自调整内存功能，请确定一组具有相似数据和内存需求的数据库分区并对它们启用自调整功能。对于其余分区，应该手动进行内存配置。



---

## 第 28 章 DB2 配置参数和变量

---

### 配置跨多个分区的数据库

数据库管理器提供了多个分区中的所有数据库配置元素的单个视图。这意味着您可以更新或复位所有数据库分区中的数据库配置，而不必对每个数据库分区调用 `db2_all` 命令。

通过从数据库所在的任何分区只发出一个 SQL 语句或一个管理命令，即可更新多个分区中的该数据库配置。缺省情况下，用于更新或复位数据库配置的方法是在所有数据库分区上。

要实现命令脚本和应用程序的向后兼容性，您有下面三种选择：

- 使用 `db2set` 命令将 **DB2\_UPDDBCFG\_SINGLE\_DBPARTITION** 注册表变量设置为 `TRUE`，如下所示：

```
DB2_UPDDBCFG_SINGLE_DBPARTITION=TRUE
```

**注：** 设置该注册表变量不适用于使用 `ADMIN_CMD` 过程发出的 `UPDATE DATABASE CONFIGURATION` 或 `RESET DATABASE CONFIGURATION` 请求。

- 对 `UPDATE DATABASE CONFIGURATION` 或 `RESET DATABASE CONFIGURATION` 命令或者 `ADMIN_CMD` 过程使用 **DBPARTITIONNUM** 参数。例如，要更新所有数据库分区上的数据库配置，请按如下所示调用 `ADMIN_CMD` 过程：

```
CALL SYSPROC.ADMIN_CMD  
( 'UPDATE DB CFG USING sortheap 1000' )
```

要更新单个数据库分区，请按如下所示调用 `ADMIN_CMD` 过程：

```
CALL SYSPROC.ADMIN_CMD  
( 'UPDATE DB CFG DBPARTITIONNUM 10 USING sortheap 1000' )
```

- 对 `db2CfgSet` API 使用 **DBPARTITIONNUM** 参数。**db2Cfg** 结构中的标志指示数据库配置的值是否将应用于单个数据库分区。如果设置一个标志，那么还必须提供 **DBPARTITIONNUM** 值，例如：

```
#define db2CfgSingleDbpartition          256
```

如果未设置 `db2CfgSingleDbpartition` 值，那么该数据库配置的值将应用于所有数据库分区，除非对用于设置数据库管理器或数据库配置参数的 `db2CfgSet` API 将 **DB2\_UPDDBCFG\_SINGLE\_DBPARTITION** 注册表变量设置为 `TRUE`，或者将 `versionNumber` 设置为低于版本 9.5 的版本号的任意版本号。

将数据库迁移到版本 9.5 时，由于所有数据库配置参数都保持相同的值，所以不需要迁移数据库配置文件。但是，对已迁移数据库发出的任何后续更新或复位数据库配置请求将使用版本 9.5 方法来更新或复位配置请求。

对于现有更新或复位命令脚本，前面提到的规则同样适用：可以使用版本 9.5 之前的方法，也可以修改脚本以包括 `UPDATE DATABASE CONFIGURATION` 或 `RESET DATABASE CONFIGURATION` 命令的 **DBPARTITIONNUM** 选项，也可以设置 **DB2\_UPDDBCFG\_SINGLE\_DBPARTITION** 注册表变量。

对于调用 db2CfgSet API 的现有应用程序，必须使用版本 9.5 方法。如果要使用版本 9.5 之前的版本的方法，那么可以设置 **DB2\_UPDDBCFG\_SINGLE\_DBPARTITION** 注册表变量，或者修改应用程序以使用版本 9.5 版本号调用该 API，包括新的 db2CfgSingleDbpartition 标志以及用于更新或复位特定数据库分区的数据库配置的新 **dbpartitionnum** 字段。

注：如果您发现数据库配置值不一致，那么可以单独地更新或复位每个数据库分区。

---

## 分区数据库环境变量

### DB2CHGPWD\_EEE

- 操作系统: AIX 和 Windows 上的 DB2 ESE
- 缺省值 = NULL, 值: YES 或 NO
- 此变量指定是否允许其他用户更改 AIX 或 Windows ESE 系统上的密码。必须确保在 Windows 上使用 Windows 域控制器或在 AIX 上使用 LDAP 来集中维护所有数据库分区或节点的密码。如果没有集中维护，在所有数据库分区或节点之间密码可能会不一致。这可能会导致只在用户为了更改而连接的数据库分区中更改密码。

### DB2\_NUM\_FAILOVER\_NODES

- 操作系统: 所有操作系统
- 缺省值 = 2, 值: 0 至需要的数据库分区数
- 设置 **DB2\_NUM\_FAILOVER\_NODES** 以指定在发生故障转移时可能需要在机器上启动的其他数据库分区的数目。

在 DB2 数据库高可用性解决方案中，如果数据库服务器出现故障，那么可以在另一台机器上重新启动故障机器上的数据库分区。快速通信管理器 (FCM) 使用 **DB2\_NUM\_FAILOVER\_NODES** 来计算每台机器上要保留以便于进行此故障转移的内存大小。

例如，考虑以下配置：

- 机器 A 有两个数据库分区: 1 和 2。
- 机器 B 有两个数据库分区: 3 和 4。
- **DB2\_NUM\_FAILOVER\_NODES** 在机器 A 和机器 B 上都设置为 2。

在 DB2START 时，FCM 将在机器 A 和机器 B 上保留足够多的内存，以便管理多达四个数据库分区，这样在一台机器出现故障时，可以在另一台机器上重新启动故障机器上的两个数据库分区。如果机器 A 出现故障，那么可以在机器 B 上重新启动数据库分区 1 和 2。如果机器 B 出现故障，那么可以在机器 A 上重新启动数据库分区 3 和 4。

### DB2\_PARTITIONEDLOAD\_DEFAULT

- 操作系统: 所有受支持的 ESE 平台
- 缺省值 = YES, 值: YES 或 NO
- **DB2\_PARTITIONEDLOAD\_DEFAULT** 注册表变量允许用户在未指定特定于 ESE 的装入选项时更改 ESE 环境中 load 实用程序的缺省行为。缺省值为 YES，这指定在 ESE 环境中，如果未指定特定于 ESE 的装入选项，那么将



在定义了目标表的所有数据库分区上尝试装入。当值为 NO 时，仅在 load 实用程序当前连接指的数据库分区上尝试装入。

**注：**不推荐使用此变量，在以后的发行版中可能会将其除去。LOAD 命令具有各种选项，可使用它们获得相同的行为。通过使用 LOAD 命令指定以下内容，可获得与此变量的 NO 设置相同的结果：PARTITIONED DB CONFIG MODE LOAD\_ONLY OUTPUT\_DBPARTNUMS x，其中 x 是装入数据的分区的分区号。

### DB2PORTRANGE

- 操作系统：Windows
- 值：nnnn:nnnn
- 将此值设置为 FCM 使用的 TCP/IP 端口范围，以便在另一台机器上创建的任何附加数据库分区也有同样的端口范围。

---

## 分区数据库环境配置参数

### 通信

#### conn\_elapse - 连接耗用时间

此参数指定在两个数据库分区服务器之间建立 TCP/IP 连接所用的秒数。

##### 配置类型

数据库管理器

适用于 带有本地和远程客户机的分区数据库服务器

##### 参数类型

可联机配置

传播类 立即

##### 缺省值 [范围]

10 [0-100]

##### 计量单位

秒

如果连接尝试在此参数指定的时间内成功，那么建立了通信。如果失败，那么进行另一次尝试来建立通信。如果尝试连接的次数达到 *max\_connretries* 参数指定的次数且始终超时，那么发出错误。

#### fcm\_num\_buffers - FCM 缓冲区数

此参数指定数据库服务器之间及内部用于内部通信（消息）的 4 KB 缓冲区数。

##### 配置类型

数据库管理器

适用于

- 带有本地和远程客户机的数据库服务器
- 带有本地客户机的数据库服务器
- 带有本地和远程客户机的分区数据库服务器

## 参数类型

可联机配置

传播类 立即

缺省值 [范围]

### 32 位平台

Automatic [128 - 65 300]

### 64 位平台

Automatic [128 - 524 288]

- 带有本地和远程客户机的数据库服务器: 缺省值为 1024
- 带有本地客户机的数据库服务器: 缺省值为 512
- 带有本地和远程客户机的分区数据库服务器: 缺省值为 4096

在单一分区数据库系统上, 如果 *intra\_parallel* 参数不是活动的, 那么不使用此参数。

当此参数设置为 AUTOMATIC 时, FCM 将监视资源使用情况并且如果在 30 分钟内未使用资源, 那么会逐渐将其释放。如果数据库管理器在实例启动时无法分配所指定的资源数, 它就会逐渐减小配置值, 直到可以启动实例为止。

如果同一机器上有多逻辑节点, 那么您可能会发现需要增大此参数的值。如果由于系统上的用户数、系统上的数据库分区服务器数或这些应用程序的复杂程度而使消息缓冲区用尽, 您可能也会发现需要增大此参数的值。

如果您正在使用多逻辑节点, 那么同一机器上所有多逻辑节点将共享一个缓冲区数为 *fcm\_num\_buffers* 的池。通过 *fcm\_num\_buffers* 值的 N 次方 (其中 N 为该物理机器上逻辑节点的数目) 可以确定池大小。复查正在使用的值。考虑在多逻辑节点所在的机器上总共将分配多少 FCM 缓冲区。

## fcm\_num\_channels - FCM 通道数

此参数指定用于每个数据库分区的 FCM 通道的数目。

### 配置类型

数据库管理器

适用于

- 带有本地和远程客户机的数据库服务器
- 带有本地客户机的数据库服务器
- 带有本地和远程客户机的分区数据库服务器
- 带有本地客户机的卫星数据库服务器

### 参数类型

可联机配置

传播类 立即

缺省值 [范围]

### UNIX 32 位平台

Automatic, 起始值为 256, 512, 2 048 [128 - 120 000 ]

### UNIX 64 位平台

Automatic, 起始值为 256, 512, 2 048 [128 - 524 288 ]

### Windows 32 位

Automatic, 起始值为 10 000 [128 - 120 000 ]

### Windows 64 位

Automatic, 起始值为 256, 512, 2 048 [128 - 524 288 ]

- 对于带有本地客户机和远程客户机的数据库服务器，起始值为 512。
- 对于带有本地客户机的数据库服务器，起始值为 256。
- 对于带有本地和远程客户机的分区数据库环境服务器，起始值为 2 048。

在非分区数据库环境上，*intra\_parallel* 参数必须是活动的，然后才能使用 *fcnum\_channels*。

FCM 通道表示在 DB2 引擎中运行的 EDU 之间的逻辑通信端点。控制流（请求和应答）和数据流（表队列数据）都依靠通道在各个分区之间传送数据。

当设置为 AUTOMATIC 时，FCM 将监视通道使用情况，并且会随要求的更改而逐渐分配和释放资源。

## max\_connretries - 节点连接重试次数

此参数指定尝试在两个数据库分区服务器之间建立 TCP/IP 连接的最大次数。

### 配置类型

数据库管理器

适用于 带有本地和远程客户机的分区数据库服务器

### 参数类型

可联机配置

传播类 立即

### 缺省值 [范围]

5 [0-100]

如果试图在两个数据库分区服务器之间建立通信失败（例如，达到 *conn\_elapse* 参数指定的值），那么 *max\_connretries* 指定可对数据库分区服务器进行连接重试的次数。如果超过为此参数指定的值，将返回一个错误。

## max\_time\_diff - 节点间的最大时差

此参数以分钟为单位指定列示在节点配置文件文件中的数据库分区服务器所允许的最大时间差。

### 配置类型

数据库管理器

适用于 带有本地和远程客户机的分区数据库服务器

### 参数类型

可配置

### 缺省值 [范围]

60 [1 - 1 440]

## 计量单位

分钟

每个数据库分区服务器都有自己的系统时钟。如果两个或多个数据库分区服务器与某一事务相关联，并且它们的时钟在由此参数指定的时间内不同步，那么拒绝该事务，并且返回 `SQLCODE`。（仅当数据修改与事务相关时，才拒绝事务。）

DB2 使用全球标准时间（UTC），因此在设置此参数时不考虑不同的时区。“全球标准时间”与“格林威治标准时间”相同。

## **start\_stop\_time** - 启动和停止超时

此参数以分钟为单位指定时间，在该段时间内所有数据库分区服务器必须响应 `DB2START` 或 `DB2STOP` 命令。在 `ADD DBPARTITIONNUM` 操作期间，它也用作超时值。

### 配置类型

数据库管理器

适用于 带有本地和远程客户机的数据库服务器

### 参数类型

可联机配置

传播类 立即

### 缺省值 [范围]

10 [1 - 1 440]

### 计量单位

分钟

在指定的时间内未响应 `DB2START` 命令的数据库分区服务器向此实例的主目录下 `sqllib` 子目录的 `log` 子目录中的 `db2start` 错误日志发送一条消息。应在重新启动这些节点之前对这些节点发出 `DB2STOP` 命令。

在指定的时间内未响应 `DB2STOP` 命令的数据库分区服务器向此实例的主目录下 `sqllib` 子目录的 `log` 子目录中的 `db2stop` 错误日志发送一条消息。您可以对每个未响应的数据库分区服务器或者对所有服务器发出 `db2stop`。（那些已停止的服务器将返回一条信息以表明它们是停止的。）

如果多分区数据库中的 `db2start` 或 `db2stop` 操作未在 `start_stop_time` 数据库管理器配置参数所指定的值内完成，那么超时的数据库分区将内部终止。如果具有许多数据库分区的环境的 `start_stop_time` 值较低，那么可能会遇到此行为。要解决这种情况，增大 `start_stop_time` 的值。

在使用 `DB2START`、`START DATABASE MANAGER` 或 `ADD DBPARTITIONNUM` 命令之一添加新数据库分区时，添加数据库分区操作必须确定实例中的每个数据库是否已启用自动存储器。这是通过与每个数据库的目录分区通信完成的。如果已启用自动存储器，就会在该通信过程中检索存储路径定义。同样，如果要创建带有数据库分区的系统临时表空间，该操作就可能必须与另一数据库分区服务器通信以检索该服务器上数据库分区的表空间定义。在确定 `start_stop_time` 参数值时，应该考虑这些因素。

## 并行处理

### **intra\_parallel** - 启用分区内并行性

此参数指定数据库管理器是否可以使用分区内并行性。

#### 配置类型

数据库管理器

#### 适用于

- 带有本地和远程客户机的数据库服务器
- 带有本地客户机的数据库服务器
- 带有本地和远程客户机的分区数据库服务器

#### 参数类型

可配置

#### 缺省值 [范围]

NO (0) [SYSTEM (-1), NO (0), YES (1)]

值 -1 导致该参数值设置为“YES”或“NO”，这取决于正在运行数据库管理器的硬件。

当此参数设置为“YES”时，一些参数可利用并行性能改进，这些操作包括数据库查询和索引创建。

#### 注:

- 并行索引创建不使用此配置参数。
- 如果更改此参数值，那么可能将程序包重新绑定至数据库，并且可能会使性能有一定下降。

### **max\_querydegree** - 最大查询并行度

此参数指定用于在数据库管理器的此实例上执行的任何 SQL 语句的最大分区内并行度。当执行某条 SQL 语句时，该语句在一个数据库分区内使用的并行操作的数目将不大于此数目。

#### 配置类型

数据库管理器

#### 适用于

- 带有本地和远程客户机的数据库服务器
- 带有本地客户机的数据库服务器
- 带有本地和远程客户机的分区数据库服务器

#### 参数类型

可联机配置

#### 传播类 语句边界

#### 缺省值 [范围]

-1 (ANY) [ANY, 1 - 32 767] (ANY 表示由系统确定)

必须将 *intra\_parallel* 配置参数设置为“YES”，以便允许数据库分区将分区内并行性用于 SQL 语句。创建并行索引不再需要 *intra\_parallel* 参数。

此配置参数的缺省值为 -1。此值表示系统使用优化器确定的并行度；否则，使用用户指定的值。

**注：**可使用 `CURRENT DEGREE` 专用寄存器或 `DEGREE` 绑定选项在编译语句时指定 SQL 语句的并行度。

活动应用程序的最大查询并行度可以使用 `SET RUNTIME DEGREE` 命令来修改。实际使用的运行时并行度是下列值中较小的那一个：

- `max_querydegree` 配置参数
- 应用程序运行时并行度
- SQL 语句编译并行度

此配置参数仅适用于查询。

---

## 第 5 部分 管理 API、命令和 SQL 语句





---

## 第 29 章 管理 API

---

### sqlleaddn - 添加数据库分区服务器到分区数据库环境

添加新的数据库分区服务器到分区数据库环境。此 API 为新数据库分区服务器上实例中当前定义的所有数据库创建数据库分区。用户可以为将要与数据库一起创建的任何系统临时表空间指定源数据库分区服务器，或者指定将不创建系统临时表空间。您必须从将要添加的数据库分区服务器发出该 API，并且它只能在数据库分区服务器上发出。

#### 作用域

此 API 只影响对其执行该 API 的数据库分区服务器。

#### 授权

为下列其中一项：

- sysadm
- sysctrl

#### 必需的连接

无

#### API 包含文件

sqlenv.h

#### API 和数据结构语法

```
SQL_API_RC SQL_API_FN
sqlleaddn (
    void * pAddNodeOptions,
    struct sqlca * pSqlca);
```

```
SQL_API_RC SQL_API_FN
sqlgaddn (
    unsigned short addnOptionsLen,
    struct sqlca * pSqlca,
    void * pAddNodeOptions);
```

#### sqlleaddn API 参数

##### pAddNodeOptions

输入。指向可选的 `sqlc_addn_options` 结构的指针。此结构用于为在“添加节点”操作期间创建的所有数据库分区指定系统临时表空间定义的源数据库分区服务器（如果有）。如果未指定（即，指定了 NULL 指针），那么系统临时表空间定义将与目录分区的那些表空间定义相同。

##### pSqlca

输出。指向 `sqlca` 结构的指针。

## 特定于 sqlgaddn API 的参数

### addnOptionsLen

输入。一个 2 字节的无符号整数，表示可选的 `sqlc_addn_options` 结构的长度（以字节计）。

### 使用说明

在添加新的数据库分区服务器之前，确保具有足够的存储空间用于必须为系统上所有现有数据库创建的容器。

“添加节点”操作在新数据库分区服务器上为存在于实例中的每个数据库创建空的数据库分区。将新数据库分区的配置参数设置为缺省值。

如果在本地创建数据库分区时“添加节点”操作失败，那么它会进入清除阶段，以本地方式删除已创建的所有数据库。这意味着仅从添加的数据库分区服务器（即，本地数据库分区服务器）除去数据库分区。现有数据库分区在所有其他数据库分区服务器上不受影响。如果此操作失败，那么将不执行进一步清除，并且返回错误。

直到使用 `ALTER DATABASE PARTITION GROUP` 语句将数据库分区服务器添加到数据库分区组后，才可以使用新数据库分区服务器上的数据库分区来包含用户数据。

如果创建数据库或删除数据库操作正在进行中，那么此 API 将会失败。一旦操作完成，可再次调用该 API。

在任何时候，如果在系统的数据库中创建具有 XML 列的用户表，或者注册 XSR 对象，不论操作是否成功，此 API 都将失败。

要确定是否已启用数据库以获得自动存储器，`sqlcaddn` API 必须与实例中每个数据库的目录分区通信。如果已启用自动存储器，那么作为该通信的一部分将检索存储路径定义。同样，如果要创建带有数据库分区的系统临时表空间，那么 `sqlcaddn` API 就可能必须与分区数据库环境中的另一数据库分区服务器通信，以便检索表空间定义。`start_stop_time` 数据库管理器配置参数用于指定其他数据库分区服务器必须使用自动存储器和表空间定义响应的时间（以分钟计）。如果超出此时间，那么 API 失败。增大 `start_stop_time` 的值，然后再次调用该 API。

### REXX API 语法

可以通过 `SQLDB2` 接口从 REXX 调用此 API。

---

## sqlcgran - 在数据库分区服务器上创建数据库

仅在调用 API 的数据库分区服务器上创建数据库。此 API 不准备用于一般用途。例如，如果数据库分区服务器上的数据库分区受损且必须重新创建，那么它应该与 `db2Restore` 配合使用。错误使用此 API 可能会导致系统中出现不一致情况，因此应谨慎使用。

**注：**如果此 API 用于重新创建已删除的数据库分区（因为它已损坏），那么此数据库分区服务器上的数据库将处于复原暂挂状态。在重新创建数据库分区后，必须在此数据库分区服务器上立即复原该数据库。

## 作用域

此 API 只影响在其上调用该 API 的数据库分区服务器。

## 授权

为下列其中一项:

- sysadm
- sysctrl

## 必需的连接

实例。要在另一数据库分区服务器上创建数据库，首先必须连接至该数据库分区服务器。此 API 会在处理期间临时建立数据库连接。

## API 包含文件

sqlenv.h

## API 和数据结构语法

```
SQL_API_RC SQL_API_FN
sqlcgran (
    char * pDbName,
    void * pReserved,
    struct sqlca * pSqlca);
```

```
SQL_API_RC SQL_API_FN
sqlgcran (
    unsigned short reservedLen,
    unsigned short dbNameLen,
    struct sqlca * pSqlca,
    void * pReserved,
    char * pDbName);
```

## sqlcgran API 参数

### pDbName

输入。包含要创建的数据库名的字符串。不能为 NULL。

### pReserved

输入。被设置为 null 或指向零的备用指针。保留以备将来使用。

### pSqlca

输出。指向 sqlca 结构的指针。

## 特定于 sqlgcran API 的参数

### reservedLen

输入。保留用于 pReserved 的长度。

### dbNameLen

输入。一个 2 字节的无符号整数，表示数据库名的长度（以字节计）。

## 使用说明

当成功创建数据库时，它被置于复原暂挂状态。在可以使用数据库之前，必须在此数据库分区服务器上复原该数据库。

## REXX API 语法

可以通过 SQLDB2 接口从 REXX 调用此 API。

---

### sqlledpan – 删除数据库分区服务器上的数据库

删除指定的数据库分区服务器上的数据库。只能在分区数据库环境中运行。

#### 作用域

此 API 只影响在其上调用该 API 的数据库分区服务器。

#### 授权

为下列其中一项:

- sysadm
- sysctrl

#### 必需的连接

无。在调用期间建立实例连接。

#### API 包含文件

sqlenv.h

#### API 和数据结构语法

```
SQL_API_RC SQL_API_FN
sqlledpan (
    char * pDbAlias,
    void * pReserved,
    struct sqlca * pSqlca);
```

```
SQL_API_RC SQL_API_FN
sqlgdpan (
    unsigned short Reserved1,
    unsigned short DbAliasLen,
    struct sqlca * pSqlca,
    void * pReserved2,
    char * pDbAlias);
```

#### sqlledpan API 参数

##### pDbAlias

输入。包含要删除的数据库的别名的字符串。此名称用于引用系统数据库目录中的实际数据库名。

##### pReserved

保留参数。应为 NULL。

##### pSqlca

输出。指向 sqlca 结构的指针。

#### 特定于 sqlgdpan API 的参数

##### Reserved1

保留以备将来使用。

### **DbAliasLen**

输入。一个 2 字节的无符号整数，表示数据库别名的长度（以字节计）。

### **pReserved2**

被设置为 null 或指向零的备用指针。保留以备将来使用。

### **使用说明**

错误使用此 API 可能会导致系统中出现不一致情况，因此应谨慎使用。

### **REXX API 语法**

可以通过 SQLDB2 接口从 REXX 调用此 API。

---

## **sqlcdrpn - 检查是否可以删除数据库分区服务器**

验证数据库是否正在使用数据库分区服务器。将会返回消息，指示是否可以删除数据库分区服务器。

### **作用域**

此 API 只影响在其上发出 API 的数据库分区服务器。

### **授权**

为下列其中一项：

- sysadm
- sysctrl

### **API 包含文件**

sqlenv.h

### **API 和数据结构语法**

```
SQL_API_RC SQL_API_FN
sqlcdrpn (
    unsigned short Action,
    void * pReserved,
    struct sqlca * pSqlca);
```

```
SQL_API_RC SQL_API_FN
sqlgdrpn (
    unsigned short Reserved1,
    struct sqlca * pSqlca,
    void * pReserved2,
    unsigned short Action);
```

### **sqlcdrpn API 参数**

**操作** 所请求的操作。有效值为: SQL\_DROPNODE\_VERIFY

### **pReserved**

保留参数。应为 NULL。

### **pSqlca**

输出。指向 sqlca 结构的指针。

## 特定于 sqlgdrpn API 的参数

### Reserved1

保留用于 pReserved2 的长度。

### pReserved2

被设置为 NULL 或指向 0 的备用指针。保留以供将来使用。

## 使用说明

如果返回消息，指示数据库分区服务器未在使用中，那么将 db2stop 命令与 DROP NODENUM 配合使用，以便从 db2nodes.cfg 文件中除去数据库分区服务器的条目，它将从分区数据库环境中除去数据库分区服务器。

如果返回消息，指示数据库分区服务器正在使用中，那么应采取下列操作：

1. 对于实例中的每个数据库，要删除的数据库分区服务器将在其上具有数据库分区。如果任何这些数据库分区包含数据，那么再分发那些使用这些数据库分区的数据库分区组。再分发数据库分区组，以将数据移至存在于数据库分区服务器的尚未被删除的数据库分区。
2. 在再分发数据库分区组之后，从使用数据库分区的每个数据库分区组中将其删除。要从数据库分区组中除去数据库分区，您可以使用 sqludrtd API 的删除节点选项或 ALTER DATABASE PARTITION GROUP 语句。
3. 删除在数据库分区服务器上定义的任何事件监视器。
4. 重新运行 sqlgdrpn 以确保数据库分区服务器上的数据库分区不再处于使用中。

## REXX API 语法

可以通过 SQLDB2 接口从 REXX 调用此 API。

---

## sqlugrpn - 为行获取数据库分区服务器号

基于分布键值返回数据库分区号和数据库分区服务器号。应用程序可以使用此信息来确定将特定的表行存储在哪个数据库分区服务器上。

分区数据结构 sqlupi 是此 API 的输入。该结构可由 sqlugtpi API 返回。另一个输入是相应分布键值的字符表示法。输出是由分布策略以及来自分布图的相应数据库分区服务器号生成的数据库分区号。如果未提供分布图信息，那么只返回数据库分区号。这在分析数据分布时非常有用。

在调用此 API 时，数据库管理器不需要是在运行中。

### 作用域

必须从 db2nodes.cfg 文件中的数据库分区服务器调用此 API。由于客户机与服务器之间在代码页和尾数法方面存在差别，所以您不应该从客户机调用此 API，否则可能导致返回错误的数据库分区信息。

### 授权

无

## API 包含文件

sqlutil.h

## API 和数据结构语法

```
SQL_API_RC SQL_API_FN
sqlugrpn (
    unsigned short num_ptrs,
    unsigned char ** ptr_array,
    unsigned short * ptr_lens,
    unsigned short territory_ctypecode,
    unsigned short codepage,
    struct sqlupi * part_info,
    short * part_num,
    SQL_PDB_NODE_TYPE * node_num,
    unsigned short chklvl,
    struct sqlca * sqlca,
    short dataformat,
    void * pReserved1,
    void * pReserved2);
```

```
SQL_API_RC SQL_API_FN
sqlggrpn (
    unsigned short num_ptrs,
    unsigned char ** ptr_array,
    unsigned short * ptr_lens,
    unsigned short territory_code,
    unsigned short codepage,
    struct sqlupi * part_info,
    short * part_num,
    SQL_PDB_NODE_TYPE * node_num,
    unsigned short chklvl,
    struct sqlca * sqlca,
    short dataformat,
    void * pReserved1,
    void * pReserved2);
```

## sqlugrpn API 参数

### num\_ptrs

ptr\_array 中的指针数目。该值必须与为 part\_info 参数指定的值（即 part\_info->sqld）相同。

### ptr\_array

指向 part\_info 中指定的分布键的每个部分的相应值的字符表示法的指针数组。如果需要空值，那么将相应的指针设置为 null。对于生成列，此函数不会生成行值。用户负责提供将导致行正确分区的值。

### ptr\_lens

包含 part\_info 中指定的分区键的每个部分的相应值的字符表示法的长度的无符号整数的数组。

### territory\_ctypecode

目标数据库的国家/地区代码。您也可以使用 GET DATABASE CONFIGURATION 命令从数据库配置文件中获取此值。

### codepage

目标数据库的代码页。您也可以使用 GET DATABASE CONFIGURATION 命令从数据库配置文件中获取此值。

### part\_info

sqlupi 结构的指针。

**part\_num**

指向用于存储数据库分区号的双字节有符号整数的指针。

**node\_num**

指向用于存储节点号的 SQL\_PDB\_NODE\_TYPE 字段的指针。如果指针为空，那么将不返回节点号。

**chklvl** 指定对输入参数进行的检查级别的无符号整数。如果指定的值为零，那么不进行检查。如果指定了任何非零的值，那么将检查所有输入参数。

**sqlca** 输出。指向 sqlca 结构的指针。

**dataformat**

指定分布键值的表示法。有效的值为：

**SQL\_CHARSTRING\_FORMAT**

所有分布键值由字符串表示。这是缺省值。

**SQL\_IMPLIEDDECIMAL\_FORMAT**

隐含的小数点所在的位置由列定义来确定。例如，如果列定义为 DECIMAL(8,2)，那么值 12345 被处理为 123.45。

**SQL\_PACKEDDECIMAL\_FORMAT**

所有十进制列分布键值采用压缩十进制格式。

**SQL\_BINARYNUMERICS\_FORMAT**

所有数字分布键值采用大尾数法二进制格式。

**pReserved1**

保留以备将来使用。

**pReserved2**

保留以备将来使用。

## 使用说明

操作系统上支持的数据类型与可定义为分布键的那些数据类型相同。

**注：**必须将 CHAR、VARCHAR、GRAPHIC 和 VARGRAPHIC 数据类型转换为数据库代码页，然后才调用此 API。

对于数字和日期时间数据类型，字符表示法必须在调用 API 的相应系统的代码页上。

如果 node\_num 不为空，那么必须提供分布图；即，part\_info 参数中的 pmaplen 字段（part\_info->pmaplen）是 2 或 8192。否则，将返回 SQLCODE -6038。您必须定义分布键；即，part\_info 参数中的 sqld 字段（part\_info->sqld）必须大于零。否则，将返回 SQLCODE -2032。

如果将空值分配给非空分区列，那么将返回 SQLCODE -6039。

对于 CHAR、VARCHAR、GRAPHIC 和 VARGRAPHIC 数据类型，只截掉尾部空格；对于除此之外的其他类型，输入字符串的所有前导空格和尾部空格都会被截掉。



## 第 30 章 命令

### REDISTRIBUTE DATABASE PARTITION GROUP

在数据库分区组中的数据库分区之间再分发数据。当前数据分布不论是均匀分布还是非对称分布，均可被指定。再分发算法将根据当前数据分布来选择要移动的分区。

此命令只能从目录数据库分区发出。使用 LIST DATABASE DIRECTORY 命令，为每个数据库确定哪个数据库分区是目录数据库分区。

#### 作用域

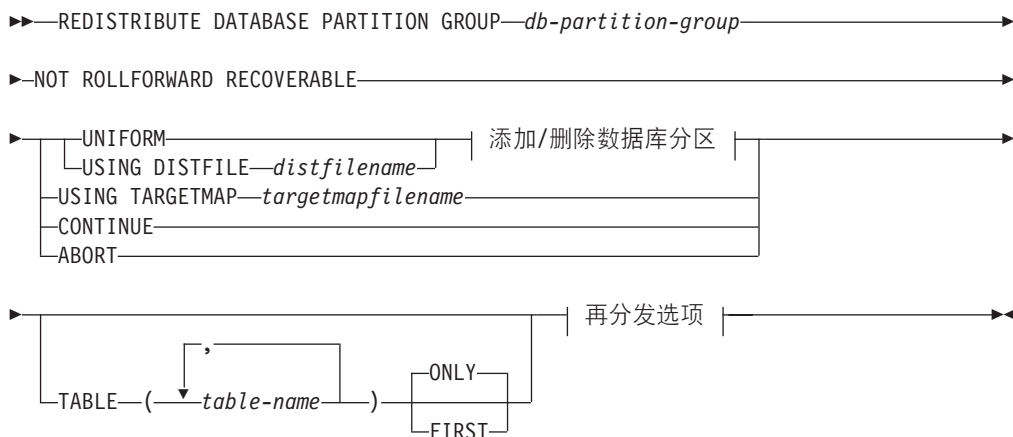
此命令会影响数据库分区组中的所有数据库分区。

#### 授权

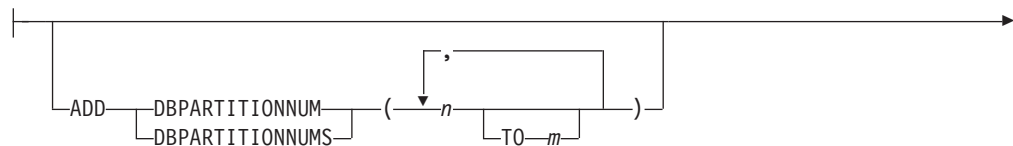
为下列其中一项：

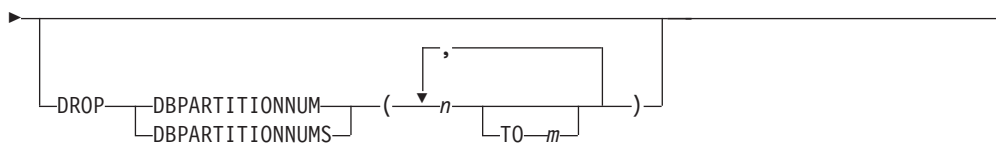
- *sysadm*
- *sysctrl*
- *dbadm*

#### 命令语法

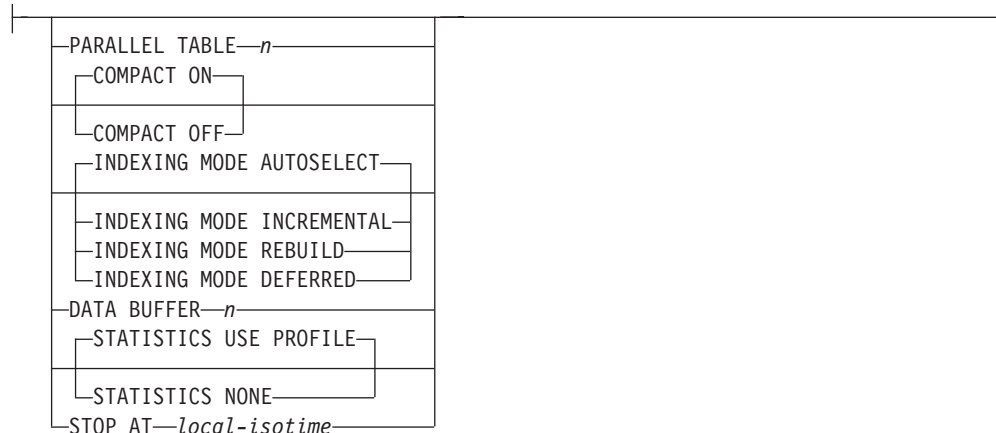


#### 添加/删除数据库分区：





再分发选项:



## 命令参数

### DATABASE PARTITION GROUP *db-partition-group*

数据库分区组的名称。此单一部分名称标识了 SYSCAT.DBPARTITIONGROUPS 目录表中描述的数据库分区组。数据库分区组当前无法进行再分发。

**注:** 无法将 IBMCATGROUP 和 IBMTEMPGROUP 数据库分区组中的表进行再分发。

### NOT ROLLFORWARD RECOVERABLE

提醒您 REDISTRIBUTE DATABASE PARTITION GROUP 命令是不可前滚恢复的重要提示。

### UNIFORM

指定数据均匀分布在散列分区（即，假定每个散列分区均具有相同的行数），但相同数目的散列分区未映射至每个数据库分区。在再分发之后，数据库分区组中的所有数据库分区均具有数目大致相同的散列分区。

### USING DISTFILE *distfilename*

如果分布键值的分布不均匀，那么使用此选项以实现在数据库分区组的数据库分区之间均匀地再分发数据。

使用 *distfilename* 以指示当前在 4,096 个散列分区之间的数据分布。

使用行计数、字节量或任何其他量度标准来指示每个散列分区表示的数据量。实用程序将与分区相关的整数值理解为该分区的权重。当指定 *distfilename* 时，实用程序会生成目标分布图，用于在数据库分区组中的数据库分区之间尽可能均匀地再分发数据。在再分发之后，数据库分区组中每个数据库分区的权重将会大致相同（数据库分区的权重是映射至该数据库分区的所有散列分区的权重总和）。

例如，输入分布文件可能包含以下条目：

```
10223
1345
112000
0
100
...
```

在示例中，散列分区 2 的权重为 112,000，而权重为 0 的分区 3 根本没有映射至该分区的数据。

*distfilename* 应该包含 4,096 个字符格式的正整数值。这些值的总和应小于或等于 4,294,967,295。

如果未指定 *distfilename* 的路径，那么将使用当前目录。

#### **USING TARGETMAP** *targetmapfilename*

使用在 *targetmapfilename* 中指定的文件作为目标分布图。将根据此文件来再分发数据。如果未指定路径，那么将使用当前目录。

如果目标图中包含的数据库分区不在该数据库分区组中，那么将返回错误。发出 ALTER DATABASE PARTITION GROUP ADD DBPARTITIONNUM 语句，然后运行 REDISTRIBUTE DATABASE PARTITION GROUP 命令。

如果从目标图中排除的数据库分区位于数据库分区组中，那么分区时将不包括该数据库分区。在使用 REDISTRIBUTE DATABASE PARTITION GROUP 命令的前后，您都可以使用 ALTER DATABASE PARTITION GROUP DROP DBPARTITIONNUM 语句来删除此类数据库分区。

#### **CONTINUE**

继续先前失败或停止的 REDISTRIBUTE DATABASE PARTITION GROUP 操作。如果什么都未发生，那么将返回错误。

#### **ABORT**

中止先前失败或停止的 REDISTRIBUTE DATABASE PARTITION GROUP 操作。如果什么都未发生，那么将返回错误。

#### **ADD**

##### **DBPARTITIONNUM** *n*

**TO** *m*

*n* 或 *n TO m* 指定将添加至数据库分区组中的数据库分区编号列表。在数据库分区组 (SQLSTATE 42728) 中，任何指定的分区都必须尚未被定义。这相当于执行指定了 ADD DBPARTITIONNUM 子句的 ALTER DATABASE PARTITION GROUP 语句。

##### **DBPARTITIONNUMS** *n*

**TO** *m*

*n* 或 *n TO m* 指定将添加至数据库分区组中的数据库分区编号列表。在数据库分区组 (SQLSTATE 42728) 中，任何指定的分区都必须尚未被定义。这相当于执行指定了 ADD DBPARTITIONNUM 子句的 ALTER DATABASE PARTITION GROUP 语句。

**注：**使用此选项添加数据库分区时，表空间的容器将基于数据库分区组中最低编号的现有分区上相应表空间的容器。如果新分区与现有容

器位于同一物理机器上，那么可能导致容器之间发生命名冲突，倘若是这样，那么您不应使用此选项。而是应该在发出 REDISTRIBUTE DATABASE PARTITION GROUP 命令之前，将 ALTER DATABASE PARTITION GROUP 语句与 WITHOUT TABLESPACES 选项配合使用。然后，可以通过手工指定适当的名称来创建表空间容器。

## DROP

### DBPARTITIONNUM *n*

**TO** *m*

*n* 或 *n TO m* 指定将从数据库分区组中删除的数据库分区编号列表。在数据库分区组（SQLSTATE 42729）中必须已定义所有指定的分区。这相当于执行指定了 DROP DBPARTITIONNUM 子句的 ALTER DATABASE PARTITION GROUP 语句。

### DBPARTITIONNUMS *n*

**TO** *m*

*n* 或 *n TO m* 指定将从数据库分区组中删除的数据库分区编号列表。在数据库分区组（SQLSTATE 42729）中必须已定义所有指定的分区。这相当于执行指定了 DROP DBPARTITIONNUM 子句的 ALTER DATABASE PARTITION GROUP 语句。

### TABLE *tablename*

为进行再分发处理指定表顺序。

**ONLY** 如果表顺序后跟 ONLY 关键字（这是缺省值），那么将仅再分发指定的表。其余的表稍后可以由后续的 REDISTRIBUTE CONTINUE 命令来处理。这是缺省情况。

**FIRST** 如果表顺序后跟 FIRST 关键字，那么指定的表将采用给定顺序进行再分发，而数据库分区组中其余的表将按照随机顺序再分发。

### PARALLEL TABLE *n*

指定要以并行方式再分发的表数目。在获得进一步通知之前，将忽略所输入的任何数目，并且使用缺省值 1。

### INDEXING MODE

此参数指定在再分发期间如何维护索引。有效的值为：

#### AUTOSELECT

再分发将根据要移动的数据量，自动选择使用 INCREMENTAL 和 REBUILD 方式之一。这是缺省方式。

#### INCREMENTAL

以递增方式扩展索引。因此，索引必须有效。

#### REBUILD

将从头开始重建索引。索引不必有效即可使用此选项。使用此选项将导致索引页在磁盘上集群。

#### DEFERRED

再分发将不会试图维护任何索引。索引将被标记为需要刷新。首次访问此类索引可能会强制执行重建，或者在重新启动数据库时可能会重建索引。

**注：**对于非 MDC 表，如果表上存在无效的索引，那么 REDISTRIBUTE DATABASE PARTITION GROUP 命令在您未指定 INDEXING MODE DEFERRED 的情况下将自动重建它们。对于 MDC 表，由于实用程序需要组合索引来处理 MDC 表，所以即使您指定 INDEXING MODE DEFERRED，在开始再分发表之前，仍会重建无效的组合索引。

#### **DATA BUFFER *n***

指定要用作传送实用程序中数据的缓存空间的 4 KB 页数。如果指定值小于实用程序的体系结构最小值，那么将使用需要的最少资源，但不会返回警告。此内存是直接来自实用程序堆中分配的，可通过数据库配置参数 `util_heap_sz` 来修改此内存大小。如果未指定值，那么在处理每个表的开始阶段实用程序将在运行时期间计算智能缺省值。特别是，缺省值将使用在开始再分发表时实用程序堆中可用内存的 50%，并且也将各种表属性考虑在内。

#### **COMPACT**

此选项指定实用程序是否应填充在处理期间遇到的数据页上的孔洞。

**ON** 对于任何给定的表，在开始再分发数据之前包含表数据的数据库分区将在再分发操作进行时填充该页上的孔洞。对于有规则的表，将从该表的逻辑端获取记录，并且将在末尾处截断该表以释放空间。对于 MDC 表，将在同一单元格中获取记录以填充孔洞，并且释放任何空块以供重复使用。对于范围分区表，将从某个范围的逻辑端获取记录，并且每个范围都会发生截断以便释放空间。缺省值是 ON。

**OFF** 实用程序将不填充数据页上的孔洞。REORG INDEXES/TABLE 命令可用于在再分发之后收回空间。

#### **STOP AT *local-isotime***

当指定此选项时，实用程序将 *local-isotime* 与当前的本地时间戳记比较，然后开始处理每个表。如果指定的 *local-isotime* 等于或早于当前的本地时间戳记，那么实用程序将停止，并且出现警告消息。因此，所有先前处理的表将使用重新平衡分布图，其余的表将继续使用数据库分区组中定义的分布图。您可以使用 CONTINUE 选项来再分发其余的表。此 *local-isotime* 值可指定为时间戳记，即一个标识日期和时间的 7 部分字符串。格式为 `yyyy-mm-dd-hh.mm.ss.nnnnnn`（年、月、日、小时、分钟、秒和微秒），以本地时间表示。

#### **STATISTICS**

此选项指定实用程序应该收集具有统计信息概要文件的表的统计信息。在完成数据再分发之后，指定此选项要比单独发出 RUNSTATS 命令更为高效率。

#### **USE PROFILE**

将为具有统计信息概要文件的表收集统计信息。对于没有统计信息概要文件的表，将不执行任何操作。这是缺省情况。

**NONE** 将不为表收集统计信息。

### **再分发操作是不可前滚恢复的**

REDISTRIBUTE DATABASE PARTITION GROUP 命令使用最小日志记录策略，尽可能避免为移动的每个行编写日志记录。因为对于大系统来说，完全记录所有数据移动的方法可能需要的活动和永久日志空间量太大而变得不切合实际，并且通常将造成较差的性能特征，所以此类日志记录对再分发操作的可用性很重要。然而，能够意识到 REDISTRIBUTE DATABASE PARTITION GROUP 命令由于此最小日志记录模型而导致不可前滚恢复，这一点对用户来说很重要。这意味着通过再分发操作造成数据库前

滚的任何操作将导致再分发操作所涉及的全部表均停留在不可用状态。您只能废弃此类表，它意味着没有办法来恢复这些表中的数据。对于可恢复的数据库，这就是为什么 REDISTRIBUTE DATABASE PARTITION GROUP 实用程序将所涉及的所有表空间置于 BACKUP PENDING 状态，同时强制用户在成功的再分发操作结束时备份所有已再分发的表空间的原因。在再分发操作之后执行备份的情况下，用户就应该不需要通过再分发操作本身来进行前滚。

用户应该知道再分发实用程序缺少可前滚恢复性将存在一个非常重要的后果：当再分发操作正在运行时（包括再分发涉及的表空间由用户备份的再分发结束时的期间），如果用户选择允许针对数据库中的表（甚至正在再分发的数据库分区组外部的表）进行更新，那么此类更新在出现严重故障（例如，数据库容器损坏）的情况下可能会丢失。此类更新可能丢失的原因是再分发操作是不可前滚恢复的。如果必须从再分发操作之前所作的备份复原数据库，那么在不按照上面所述通过再分发来进行前滚从而使已再分发的表处于不可用状态的情况下，将不可能通过日志进行前滚以便重放在再分发操作期间所作的更新。因而，在此情况下只能执行一项操作，即从再分发前所执行的尚未前滚的备份来复原数据库。然后，可再次执行再分发操作。不幸的是，将会丢失在最初再分发操作期间发生的全部更新。

这一点的重要性再怎么强调也不为过。为了确定在再分发操作期间没有丢失的更新，下列其中一项必须属实：

- 用户避免在 REDISTRIBUTE DATABASE PARTITION GROUP 命令操作期间进行更新，包括在命令完成之后的期间（正在备份受影响的表空间）也是如此。
- 在再分发操作期间应用的更新来自可重复的来源，这意味着您可随时再次应用它们。例如，如果更新源是存储在文件中的数据，并且在批处理期间应用更新，那么很明显，即使在发生需要数据库复原的故障的情况下，由于可以随时再次简单地应用更新，所以将不会丢失更新。

关于在再分发操作期间允许对数据库进行更新这一方面，如有必要，用户必须根据在数据库复原之后是否重复更新来决定此类更新是否适合其情况。

**注：**并非 REDISTRIBUTE DATABASE PARTITION GROUP 命令操作期间发生的每个故障均会导致此问题。事实上，大多数都不会。REDISTRIBUTE DATABASE PARTITION GROUP 命令完全可重新启动，这意味着如果实用程序在工作期间出现故障，那么您可以使用 CONTINUE/ABORT 选项来轻松地继续或中止它。以上提及的故障是需要用户从再分发操作之前所执行的备份中进行复原的故障。

## 使用说明

- 实用程序首次访问表空间时，它被置于 BACKUP PENDING 状态。该表空间中的所有表将变为只读，一直到表空间被备份为止，只有在表空间中的所有表再分发完毕时，您才可以执行该备份操作。
- 当再分发操作正在运行时，它会生成事件日志文件，其中包含了有关再分发操作的一般信息以及所处理的每个表的开始与结束时间等信息。此事件日志文件被写入：
  - 基于 Linux 和 UNIX 的系统上的 `homeinst/sqllib/redistribute` 目录（对子目录和文件名使用下列格式：`database-name.database-partition-group-name.timestamp`）。
  - Windows 操作系统上的 `DB2INSTPROF\instance\redistribute` 目录（其中 DB2INSTPROF 是 DB2INSTPROF 注册表变量的值）（对子目录和文件名使用下列格式：`database-name.database-partition-group-name.timestamp`）。

- 时间戳记值是发出命令时的时间。

有关再分发事件日志的更多信息，请参阅再分发错误恢复主题。

- 此实用程序在处理期间执行间歇性 COMMITs。
- 使依赖正在进行再分发的表的所有程序包无效。在完成再分发数据库分区组操作之后，建议您以显式方式重新绑定此类程序包。以显式方式重新绑定可消除为无效包执行第一个 SQL 请求时的初始延迟。再分发消息文件包含正在进行再分发的所有表的列表。
- 缺省情况下，再分发实用程序将更新具有统计信息概要文件的那些表的统计信息。对于没有统计信息概要文件的表，建议在完成再分发操作后通过调用 db2Runstats API 或发出 RUNSTATS 命令，为这些表单独更新表和索引统计信息。
- 您无法再分发包含复制的具体化查询表或使用 DATA CAPTURE CHANGES 定义的表的数据库分区组。
- 如果数据库分区组中存在具有现有已声明临时表的用户临时表空间，那么不允许执行再分发操作。
- PARALLEL TABLE 选项允许以并行方式对同一数据库分区组中的多个表进行再分发，然而，不支持对同一数据库分区组并行启动多个再分发请求，将返回 -6056。
- 在表上将忽略诸如 INDEXING MODE 之类不适用的选项，而不会出现警告。例如，在没有索引的表上将忽略 INDEXING MODE。
- 在开始再分发操作之前，确保没有表处于装入暂挂状态。可以使用 LOAD QUERY 命令来检查表状态。如果发现数据因再分发操作而处于错误的数据库分区中，那么有两个选项可供选择。您可以：
  1. 卸载、废弃然后再重新装入表，或者
  2. 再次使用新的目标图来再分发数据库分区组。

## 兼容性

对于与版本 8 之前的版本的兼容性：

- 可以用关键字 NODEGROUP 来替代 DATABASE PARTITION GROUP。

---

## db2nchg - 更改数据库分区服务器配置

修改数据库分区服务器配置。这包括将数据库分区服务器（节点）从一台机器移至另一台机器；更改机器的 TCP/IP 主机名；以及为数据库分区服务器（节点）选择另一逻辑端口号或另一网络名。此命令仅在停止数据库分区服务器时才会使用。

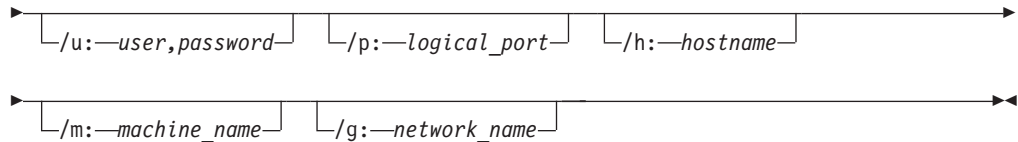
此命令仅在基于 Windows 的操作系统上可用。

### 授权

本地管理员

### 命令语法

```
►► db2nchg -/n:—dbpartitionnum [ /i:—instance_name ]
```



## 命令参数

### **/n:dbpartitionnum**

指定将要更改的数据库分区服务器配置的数据库分区号。

### **/i:instance\_name**

指定此数据库分区服务器所参与的实例。如果未指定参数，那么缺省值是当前实例。

### **/u:username,password**

指定用户名和密码。如果未指定参数，那么将应用现有用户名和密码。

### **/p:logical\_port**

指定数据库分区服务器的逻辑端口。您必须指定此参数以便将数据库分区服务器移至另一机器。如果未指定参数，那么逻辑端口号将保持不变。

### **/h:host\_name**

指定由 FCM 用于内部通信的 TCP/IP 主机名。如果未指定此参数，那么主机名将保留不变。

### **/m:machine\_name**

指定数据库分区服务器将驻留的机器。仅当实例中没有现有数据库时，才可移动数据库分区服务器。

### **/g:network\_name**

更改数据库分区服务器的网络名。当机器上存在多个 IP 地址时，此参数可用于将特定 IP 地址应用于数据库分区服务器。您可以输入网络名或 IP 地址。

## 示例

要将分配给数据库分区 2（它参与实例 TESTMPP）的逻辑端口更改为逻辑端口 3，请输入下列命令：

```
db2nchg /n:2 /i:TESTMPP /p:3
```

---

## db2ncrt - 将数据库分区服务器添加至实例

将数据库分区服务器（节点）添加至实例。

此命令仅在 Windows 操作系统上可用。

### 作用域

如果数据库分区服务器被添加至实例已经存在的计算机，那么数据库分区服务器将作为逻辑数据库分区服务器添加至计算机。如果数据库分区服务器被添加至不存在实例的计算机，那么将会添加实例，并且计算机变成新的物理数据库分区服务器。如果实例中有数据库，那么不应该使用此命令。而是应该发出带有 ADD DBPARTITIONNUM 选项的 START DATABASE MANAGER 命令。这确保可正确地将该数据库添加至新的数据库分区服务器。您也可将数据库分区服务器添加至已创建数据库的实例。由于更



改 db2nodes.cfg 文件可能导致分区数据库环境中出现不一致情况，所以您不应该编辑该文件。

## 授权

对添加了新的数据库分区服务器的计算机的本地管理员权限。

## 命令语法

```
db2nrcrt -/n:—dbpartitionnum—/u:—username,password—  
          /i:—instance_name— /m:—machine_name— /p:—logical_port—  
          /h:—host_name— /g:—network_name— /o:—instance_owning_machine—
```

## 命令参数

### /n:dbpartitionnum

用于标识数据库分区服务器的唯一数据库分区号。所输入的编号介于 1 至 999 之间。

### /u:domain\_name\username,password

指定 DB2 的域名、登录帐户名和密码。

### /i:instance\_name

指定实例名。如果未指定参数，那么缺省值是当前实例。

### /m:machine\_name

指定数据库分区服务器驻留的 Windows 工作站的计算机名。如果在远程计算机上添加数据库分区服务器，那么此参数是必需的。

### /p:logical\_port

指定用于数据库分区服务器的逻辑端口号。如果未指定此参数，那么所分配的逻辑端口号将为 0。当创建逻辑数据库分区服务器时，必须指定此参数，并且必须选择未在使用中的逻辑端口号。但是，应注意下列限制：

- 每台计算机必须具有逻辑端口为 0 的数据库分区服务器。
- 端口号不能超过在 x:\winnt\system32\drivers\etc\ 目录中为 FCM 通信保留的端口范围。例如，如果为当前实例保留 4 个端口的范围，那么最大端口号为 3。端口 0 用于缺省逻辑数据库分区服务器。

### /h:host\_name

指定由 FCM 用于内部通信的 TCP/IP 主机名。当在远程计算机上添加数据库分区服务器时，此参数是必需的。

### /g:network\_name

指定数据库分区服务器的网络名。如果未指定参数，那么将使用系统上检测到的第一个 IP 地址。当计算机上存在多个 IP 地址时，此参数可用于将特定 IP 地址应用于数据库分区服务器。您可以输入网络名或 IP 地址。

### /o:instance\_owning\_machine

指定实例拥有的计算机的计算机名。缺省值是本地计算机。当在任何非实例拥有的计算机上调用 db2nrcrt 命令时，此参数是必需的。

## 示例

要将新的数据库分区服务器添加至实例拥有的计算机 SHAYER 上的实例 TESTMPP（其中新数据库分区服务器被称为数据库分区 2 且使用逻辑端口 1），请输入下列命令：

```
db2ncrt /n:2 /u:QBPAULZ\paulz,g1reeky /i:TESTMPP /m:TEST /p:1 /o:SHAYER /h:TEST
```

---

## db2ndrop - 从实例中删除数据库分区服务器

从没有数据库的实例中删除数据库分区服务器（节点）。如果数据库分区服务器已删除，那么其数据库分区号可重新用于新的数据库分区服务器。此命令仅在停止数据库分区服务器时才会使用。

此命令仅在基于 Windows 的操作系统上可用。

### 授权

对将要删除数据库分区服务器的机器的本地管理员权限。

### 命令语法

```
db2ndrop /n:dbpartitionnum /i:instance_name
```

### 命令参数

#### **/n:dbpartitionnum**

用于标识数据库分区服务器的唯一数据库分区号。

#### **/i:instance\_name**

指定实例名。如果未指定参数，那么缺省值是当前实例。

## 示例

```
db2ndrop /n:2 /i=KMASCI
```

## 使用说明

如果从实例中删除实例拥有的数据库分区服务器（dbpartitionnum 0），那么该实例将变得不可用。要删除实例，请使用 **db2idrop** 命令。

如果在此实例中有数据库，那么不应该使用此命令。而应使用 **db2stop drop nodenum** 命令。这确保正确地分区数据库环境中除去数据库分区服务器。您也可删除数据库存在于其中的实例中的数据库分区服务器。由于更改 **db2nodes.cfg** 文件可能导致分区数据库环境中出现不一致情况，所以您不应该编辑该文件。

要从正在运行多个逻辑数据库分区服务器的机器中删除已分配给逻辑端口 0 的数据库分区服务器，必须先删除分配给其他逻辑端口的所有其他数据库分区服务器。每个数据库分区服务器均必须具有已分配给逻辑端口 0 的数据库分区服务器。

## 第 31 章 SQL 语言元素

### 数据类型

#### 与数据库分区兼容的数据类型

在分布键的对应列的基本数据类型之间，定义了数据库分区兼容性。与数据库分区兼容的数据类型具有以下属性：两个变量分别属于两种类型，它们具有相同的值，由同一数据库分区功能映射至同一分布图索引。

表 41 显示数据库分区中数据类型的兼容性。

数据库分区兼容性具有下列特征：

- 内部格式用于 DATE、TIME 和 TIMESTAMP。它们彼此都不兼容，且都不与 CHAR 兼容。
- 数据库分区兼容性不受具有 NOT NULL 或 FOR BIT DATA 定义的列的影响。
- 同等对待兼容数据类型的空值。对于不兼容的数据类型的空值，可能会产生不同的结果。
- UDT 的基本数据类型可用于分析数据库分区兼容性。
- 对分布键中值相同的小数的处理是完全相同的，即使它们的小数位和精度不同也是如此。
- 字符串（CHAR、VARCHAR、GRAPHIC 或 VARGRAPHIC）中的尾部空格会被系统提供的散列法功能忽略。
- 不同长度的 CHAR 或 VARCHAR 是兼容的数据类型。
- 即使精度不同，但相等的 DECFLOAT 值仍然会被同等对待。即使具有不同的有效数字数目，但数字相等的 DECFLOAT 值仍然会被同等对待。

表 41. 数据库分区兼容性

| 操作数              | 二进制          |              | 浮点           | 十进制浮点        | 字符串            | 图形字符串        |              | 日期           | 时间           | 时间戳          | 单值类型         | 结构化类型 |
|------------------|--------------|--------------|--------------|--------------|----------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|
|                  | 整数           | 十进制数字        |              |              |                | 字符串          | 日期           |              |              |              |              |       |
| 二进制整数            | 是            | 否            | 否            | 否            | 否              | 否            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 十进制数字            | 否            | 是            | 否            | 否            | 否              | 否            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 浮点               | 否            | 否            | 是            | 否            | 否              | 否            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 十进制浮点            | 否            | 否            | 否            | 是            | 否              | 否            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 字符串 <sup>3</sup> | 否            | 否            | 否            | 否            | 是 <sup>2</sup> | 否            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 图形字符串            | 否            | 否            | 否            | 否            | 否              | 是            | 否            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| <sup>3</sup>     |              |              |              |              |                |              |              |              |              |              |              |       |
| 日期               | 否            | 否            | 否            | 否            | 否              | 否            | 是            | 否            | 否            | 否            | <sup>1</sup> | 否     |
| 时间               | 否            | 否            | 否            | 否            | 否              | 否            | 否            | 是            | 否            | 否            | <sup>1</sup> | 否     |
| Timestamp        | 否            | 否            | 否            | 否            | 否              | 否            | 否            | 否            | 否            | 是            | <sup>1</sup> | 否     |
| 单值类型             | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | <sup>1</sup>   | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | <sup>1</sup> | 否     |
| 结构化类型            | 否            | 否            | 否            | 否            | 否              | 否            | 否            | 否            | 否            | 否            | 否            | 否     |
| <sup>3</sup>     |              |              |              |              |                |              |              |              |              |              |              |       |

表 41. 数据库分区兼容性 (续)

| 操作数 | 二进制 |       | 浮点 | 十进制浮点 | 字符串 | 图形字 |    | 时间戳 |   | 单值类型 | 结构化类型 |
|-----|-----|-------|----|-------|-----|-----|----|-----|---|------|-------|
|     | 整数  | 十进制数字 |    |       |     | 字符串 | 日期 | 时间  | 记 |      |       |

注:

- <sup>1</sup> 用户定义的单值类型 (UDT) 值是与源类型的 UDT 或任何其他拥有数据库分区兼容源类型的 UDT 相兼容的数据库分区。
- <sup>2</sup> FOR BIT DATA 属性不影响数据库分区兼容性。
- <sup>3</sup> 请注意, 用户定义的结构化类型和数据类型 LONG VARCHAR、LONG VARGRAPHIC、CLOB、DBCLOB 和 BLOB 不适用于数据库分区兼容性, 因为它们在分布键中不受支持。

## 专用寄存器

### CURRENT DBPARTITIONNUM

CURRENT DBPARTITIONNUM 专用寄存器指定为该语句标识协调程序节点号的 INTEGER 值。对于从应用程序发出的语句, 协调程序是应用程序连接至的数据库分区。对于从例程发出的语句, 协调程序是从中调用例程的数据库分区。

当在例程中的 SQL 语句中使用, CURRENT DBPARTITIONNUM 永远不会继承自调用语句。

如果数据库实例未定义用于支持数据库分区 (换句话说, 如果没有 db2nodes.cfg 文件。), 那么 CURRENT DBPARTITIONNUM 将返回 0。 (对于分区数据库, db2nodes.cfg 文件存在且包含数据库分区定义。)

可以通过 CONNECT 语句来更改 CURRENT DBPARTITIONNUM, 但这仅限于在特定情况下才行。

为了与版本 8 之前的版本相兼容, 可以用关键字 NODE 代替 DBPARTITIONNUM。

示例: 将主变量 APPL\_NODE (整数) 设置为应用程序连接至的数据库分区的编号。

```
VALUES CURRENT DBPARTITIONNUM
INTO :APPL_NODE
```

---

## 第 32 章 SQL 函数

---

### DATAPARTITIONNUM

►►—DATAPARTITIONNUM—(—*column-name*—)—————◄◄

模式为 SYSIBM。

DATAPARTITIONNUM 函数返回该行驻留在其中的数据分区的序号 (SYSDATAPARTITIONS.SEQNO)。数据分区按范围排序, 序号从 0 开始。例如, DATAPARTITIONNUM 函数为驻留在具有最低范围的数据分区中的行返回 0。

参数必须是表中任何列的标准名称或非标准名称。由于返回行级别信息, 所以无论指定哪一列, 结果都是相同的。该列可以具有任何数据类型。

如果 *column-name* 引用视图中的列, 那么视图中的列的表达式必须引用底层的基本表的列, 并且视图必须是可删除的。嵌套的或公共表表达式与视图一样遵循相同的规则。

结果的数据类型是 INTEGER 且永不为空。

当创建用户定义的函数时, 此函数不能用作源函数。因为该函数可接受任何数据类型作为参数, 所以不必创建其他特征符来支持用户定义的单值类型。

不能在检查约束或生成列的定义中使用 DATAPARTITIONNUM 函数 (SQLSTATE 42881)。不能在具体化查询表 (MQT) 定义中使用 DATAPARTITIONNUM 函数 (SQLSTATE 428EC)。

示例:

- ```
SELECT DATAPARTITIONNUM (EMPNO)
FROM EMPLOYEE
```

要将由 DATAPARTITIONNUM 返回的序号 (例如 0) 转换为可在其他 SQL 语句 (例如, ALTER TABLE...DETACH PARTITION) 中使用的数据分区名称, 您可以查询 SYSCAT.DATAPARTITIONS 目录视图。包括从 WHERE 子句中的 DATAPARTITIONNUM 获取的 SEQNO, 如下列示例中所示。

```
SELECT DATAPARTITIONNAME
FROM SYSCAT.DATAPARTITIONS
WHERE TABNAME = 'EMPLOYEE' AND SEQNO = 0
```

导致值 'PART0'。

---

### DBPARTITIONNUM

►►—DBPARTITIONNUM—(—*column-name*—)—————◄◄

模式为 SYSIBM。

DBPARTITIONNUM 函数为行返回数据库分区号。例如，如果在 SELECT 子句中使用，那么它会为结果集中的每个行返回数据库分区号。

参数必须是表中任何列的标准名称或非标准名称。由于返回行级别信息，所以无论指定哪一行，结果都是相同的。该行可以具有任何数据类型。

如果 *column-name* 引用视图中的列，那么视图中的列的表达式必须引用底层的基本表的列，并且视图必须是可删除的。嵌套的或公共表表达式与视图一样遵循相同的规则。

由 DBPARTITIONNUM 函数为其返回了数据库分区号的特定行（和表）将根据使用该函数的 SQL 语句的上下文来确定。

在转换变量和表上返回的数据库分区号从分布键列的当前转换值得来。例如，在“插入前”触发器中，该函数根据新转换变量的当前值返回映射的数据库分区号。然而，分布键列的值可能被后续的“插入前”触发器修改。因此，在将行插入数据库后，该行的最终数据库分区号可能与映射的值不同。

结果的数据类型是 INTEGER 且永不为空。如果没有 db2nodes.cfg 文件，那么结果为 0。

当创建用户定义的函数时，此函数不能用作源函数。因为该函数可接受任何数据类型作为参数，所以不必创建其他特征符来支持用户定义的单值类型。

不能在复制的表上、检查约束中或生成列的定义中使用 DBPARTITIONNUM 函数（SQLSTATE 42881）。

为了与先前版本的 DB2 相兼容，可以指定 NODENUMBER 来代替 DBPARTITIONNUM。

示例:

- 计算实例的数目，其中 EMPLOYEE 表中给定职员的行为于与 DEPARTMENT 表中职员部门的描述不同的数据库分区上。

```
SELECT COUNT(*) FROM DEPARTMENT D, EMPLOYEE E
WHERE D.DEPTNO=E.WORKDEPT
AND DBPARTITIONNUM(E.LASTNAME) <> DBPARTITIONNUM(D.DEPTNO)
```

- 连接 EMPLOYEE 表和 DEPARTMENT 表，以便两个表的行均在同一数据库分区上。

```
SELECT * FROM DEPARTMENT D, EMPLOYEE E
WHERE DBPARTITIONNUM(E.LASTNAME) = DBPARTITIONNUM(D.DEPTNO)
```

- 在 EMPLOYEE 表上使用前触发器，在名为 EMPINSERTLOG1 的表中记录职员编号和 EMPLOYEE 表中任何新行的映射的数据库分区号。

```
CREATE TRIGGER EMPINSLOGTRIG1
BEFORE INSERT ON EMPLOYEE
REFERENCING NEW AS NEWTABLE
FOR EACH ROW
INSERT INTO EMPINSERTLOG1
VALUES(NEWTABLE.EMPNO, DBPARTITIONNUM
(NEWTABLE.EMPNO))
```

## 第 33 章 SQL 语句

### ALTER DATABASE PARTITION GROUP

ALTER DATABASE PARTITION GROUP 语句可用于:

- 将一个或多个数据库分区添加至数据库分区组
- 从数据库分区组中删除一个或多个数据库分区。

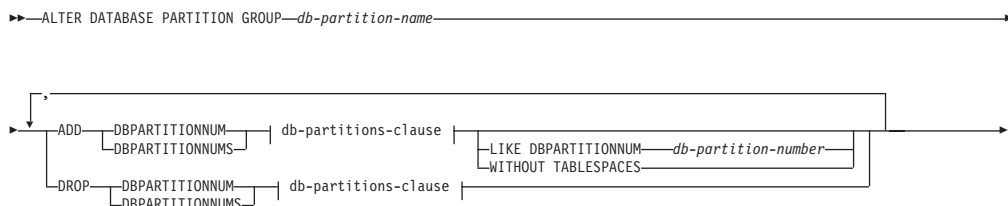
#### 调用

此语句可嵌入应用程序中或者以交互方式发出。它是一个可执行语句，仅当 DYNAMICRULES 运行行为对于程序包有效时才能动态编译该语句 (SQLSTATE 42509)。

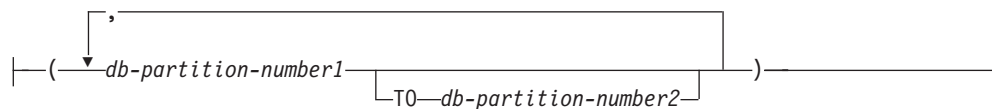
#### 授权

语句的授权标识必须具有 SYSCTRL 或 SYSADM 权限。

#### 语法



#### db-partitions-clause:



#### 描述

##### *db-partition-name*

为数据库分区组命名。这是单一部分名称。它是普通或定界 SQL 标识。它必须是目录中描述的数据库分区组。您不能指定 `IBMCATGROUP` 与 `IBMTEMPGROUP` (SQLSTATE 42832)。

##### ADD DBPARTITIONNUM

指定要添加至数据库分区组的特定数据库分区。DBPARTITIONNUMS 是 DBPARTITIONNUM 的同义词。在数据库分区组中，任何指定的数据库分区均必须尚未被定义 (SQLSTATE 42728)。

## DROP DBPARTITIONNUM

指定要从数据库分区组中删除的特定数据库分区。DBPARTITIONNUMS 是 DBPARTITIONNUM 的同义词。在数据库分区组中必须已定义所有指定的数据库分区 (SQLSTATE 42729)。

### *db-partitions-clause*

指定要添加或删除的数据库分区。

### *db-partition-number1*

指定特定的数据库分区号。

### **TO** *db-partition-number2*

指定数据库分区号的范围。*db-partition-number2* 的值必须大于或等于 *db-partition-number1* 的值 (SQLSTATE 428A9)。

## LIKE DBPARTITIONNUM *db-partition-number*

指定数据库分区组中现有表空间的容器将与指定的 *db-partition-number* 上的容器相同。指定的数据库分区必须是在此语句之前就存在于数据库分区组中且未包括在同一语句的 DROP DBPARTITIONNUM 子句中的分区。

对于定义为使用自动存储器的表空间 (即使用 CREATE TABLESPACE 语句的 MANAGED BY AUTOMATIC STORAGE 子句创建的表空间或根本未为其指定 MANAGED BY 子句的表空间), 容器将不必与来自指定分区的容器相匹配。数据库管理器将根据与数据库相关联的存储路径来自动指定容器, 这可能会也可能不会导致使用同一个容器。每个表空间的大小基于创建表空间时指定的初始大小, 它可能与指定分区上的表空间的当前大小不匹配。

## WITHOUT TABLESPACES

指定不在新添加的数据库分区上创建数据库分区组中现有表空间的容器。使用 *db-partitions-clause* 的 ALTER TABLESPACE 语句必须用于定义要与在此数据库分区组上定义的表空间配合使用的容器。如果未指定此选项, 那么在新添加的数据库分区上为在数据库分区组上定义每个表空间指定缺省容器。

对于已定义为使用自动存储器的表空间 (即使用 CREATE TABLESPACE 语句的 MANAGED BY AUTOMATIC STORAGE 子句创建的表空间或根本未为其指定 MANAGED BY 子句的表空间), 将忽略此选项。没有办法来推迟为这些表空间创建容器。数据库管理器将根据与数据库相关联的存储路径来自动指定容器。每个表空间的大小将基于创建表空间时指定的初始大小。

## 规则

- 必须在 db2nodes.cfg 文件中定义由编号指定的每个数据库分区 (SQLSTATE 42729)。
- *db-partitions-clause* 中列示的每个 *db-partition-number* 必须仅用于唯一的数据库分区 (SQLSTATE 42728)。
- 有效的数据库分区号介于 0 和 999 之间 (其中包括 0 和 999) (SQLSTATE 42729)。
- 数据库分区不能出现在 ADD 和 DROP 子句中 (SQLSTATE 42728)。
- 数据库分区组中必须至少保留一个数据库分区。您不能从数据库分区组中删除最后一个数据库分区 (SQLSTATE 428C0)。
- 如果在添加数据库分区时未指定 LIKE DBPARTITIONNUM 子句或 WITHOUT TABLESPACES 子句, 那么缺省情况下将使用数据库分区组中现有数据库分区的最小数据库分区号 (例如 2), 然后继续进行操作, 就好像已指定了 LIKE



DBPARTITIONNUM 2 一样。对于要用作缺省值的现有数据库分区，它必须具有为数据库分区组中的所有表空间定义的容器（SYSCAT.DBPARTITIONGROUPDEF 的列 IN\_USE 不是“T”）。

## 注意

- 当数据库分区被添加至数据库分区组时，将为数据库分区建立目录条目（请参阅 SYSCAT.DBPARTITIONGROUPDEF）。分布图会立即更改为包括新的数据库分区，同时出现指示器（IN\_USE），表明在下列情况下数据库分区位于分布图中：
  - 数据库分区组中未定义表空间，或者
  - 数据库分区组中定义的表空间中未定义表，并且未指定 WITHOUT TABLESPACES 子句。

分布图未更改，指示器（IN\_USE）已设置为指示在下列任何一种情况下分布图中未包括数据库分区：

- 表存在于数据库分区组中的表空间中，或者
- 表空间存在于数据库分区组中，并且 WITHOUT TABLESPACES 子句已被指定（除非所有表空间已定义为使用自动存储器，在此情况下，将忽略 WITHOUT TABLESPACES 子句）

要更改分布图，必须使用 REDISTRIBUTE DATABASE PARTITION GROUP 命令。这将再分发数据、更改分布图和更改指示器。如果指定了 WITHOUT TABLESPACES 子句，那么在试图再分发数据之前，需要添加表空间容器。

- 当从数据库分区组中删除数据库分区时，将会更新该数据库分区的目录条目（请参阅 SYSCAT.DBPARTITIONGROUPDEF）。如果在数据库分区组中定义的表空间中未定义表，那么分布图将立即更改为排除已删除的数据库分区，并且会删除该数据库分区组中数据库分区的条目。如果表存在，那么不会更改分布图，并且指示器（IN\_USE）被设置为指示数据库分区正在等待被删除。REDISTRIBUTE DATABASE PARTITION GROUP 命令必须用于再分发数据，并且从数据库分区组中删除数据库分区的条目。
- **兼容性**
  - 为了与先前版本的 DB2 相兼容：
    - 可指定 NODE 来代替 DBPARTITIONNUM
    - 可指定 NODES 来代替 DBPARTITIONNUMS
    - 可指定 NODEGROUP 来代替 DATABASE PARTITION GROUP

## 示例

假设您具有的六分区数据库拥有下列数据库分区：0、1、2、5、7 和 8。两个数据库分区（3 和 6）被添加至系统。

- 假设您想要将数据库分区 3 和 6 添加至名为 MAXGROUP 的数据库分区组，并且具有表空间容器（如数据库分区 2 上的表空间容器）。该语句如下所示：

```
ALTER DATABASE PARTITION GROUP MAXGROUP
ADD DBPARTITIONNUMS (3,6)LIKE DBPARTITIONNUM 2
```

- 假设您想要删除数据库分区 1，并且要将数据库分区 6 添加至数据库分区组 MEDGROUP。您将使用 ALTER TABLESPACE 为数据库分区 6 单独定义表空间容器。该语句如下所示：

```
ALTER DATABASE PARTITION GROUP MEDGROUP
ADD DBPARTITIONNUM(6)WITHOUT TABLESPACES
DROP DBPARTITIONNUM(1)
```

## CREATE DATABASE PARTITION GROUP

CREATE DATABASE PARTITION GROUP 语句定义数据库中的新数据库分区组，将数据库分区指定给数据库分区组，以及在系统目录中记录数据库分区组定义。

### 调用

此语句可嵌入应用程序中或者以交互方式发出。它是一个可执行语句，仅当 DYNAMICRULES 运行行为对于程序包有效时才能动态编译该语句（SQLSTATE 42509）。

### 授权

语句的授权标识所拥有的特权必须包括 SYSCTRL 或 SYSADM 权限。

### 语法

```

--CREATE DATABASE PARTITION GROUP--db-partition-group-name--
ON ALL DBPARTITIONNUMS
ON DBPARTITIONNUMS ( db-partition-number1 TO db-partition-number2 )
DBPARTITIONNUM

```

### 描述

#### *db-partition-group-name*

为数据库分区组命名。这是单一部分名称。它是普通或定界 SQL 标识。*db-partition-group-name* 必须未标识已存在于目录中的数据库分区组（SQLSTATE 42710）。*db-partition-group-name* 不得以字符“SYS”或“IBM”开头（SQLSTATE 42939）。

#### ON ALL DBPARTITIONNUMS

指定当创建数据库分区组时在定义给数据库的所有数据库分区上定义数据库分区组（db2nodes.cfg 文件）。

如果将数据库分区添加至数据库系统，那么应该发出 ALTER DATABASE PARTITION GROUP 语句，以将这个新的数据库分区包括在数据库分区组（包括 IBMDEFAULTGROUP）。此外，必须发出 REDISTRIBUTE DATABASE PARTITION GROUP 命令，以将数据移至数据库分区。

#### ON DBPARTITIONNUMS

指定位于数据库分区组中的数据库分区。DBPARTITIONNUM 是 DBPARTITIONNUMS 的同义词。

#### *db-partition-number1*

指定数据库分区号。（可以指定格式 NODEnnnnn 的 *node-name*，以便与先前版本兼容。）

#### TO *db-partition-number2*

指定数据库分区号的范围。*db-partition-number2* 的值必须大于或等于 *db-partition-*

*number1* 的值 (SQLSTATE 428A9)。介于指定的数据库分区号之间且包括这些分区号的所有数据库分区均包含在数据库分区组中。

## 规则

- 必须在 `db2nodes.cfg` 文件中定义由编号指定的每个数据库分区 (SQLSTATE 42729)。
- `ON DBPARTITIONNUMS` 子句中列示的每个 *db-partition-number* 都必须最多仅出现一次 (SQLSTATE 42728)。
- 有效的 *db-partition-number* 介于 0 和 999 之间 (包括 0 和 999) (SQLSTATE 42729)。

## 注意

- 此语句为数据库分区组创建分布图。为每个分布图生成了分布图标识 (PMAP\_ID)。此信息记录在目录中, 并且可以从 `SYSCAT.DBPARTITIONGROUPS` 和 `SYSCAT.PARTITIONMAPS` 中检索。分布图中的每个条目指定所有散列的行驻留的目标数据库分区。对于单一分区数据库分区组, 相应的分布图只有一个条目。对于具有多个分区的数据库分区组, 相应的分布图具有 4096 个条目, 其中数据库分区号在缺省情况下按循环法被分配给该图的条目。
- **兼容性**
  - 为了与先前版本的 DB2 相兼容:
    - 可指定 `NODE` 来代替 `DBPARTITIONNUM`
    - 可指定 `NODES` 来代替 `DBPARTITIONNUMS`
    - 可指定 `NODEGROUP` 来代替 `DATABASE PARTITION GROUP`

## 示例

假设您的分区数据库拥有已定义为 0、1、2、5、7 和 8 的六个数据库分区。

- 假设您想要在所有六个数据库分区上创建名为 `MAXGROUP` 的数据库分区组。该语句如下所示:

```
CREATE DATABASE PARTITION GROUP MAXGROUP ON ALL DBPARTITIONNUMS
```

- 假设您想要在数据库分区 0、1、2、5 和 8 上创建名为 `MEDGROUP` 的数据库分区组。该语句如下所示:

```
CREATE DATABASE PARTITION GROUP MEDGROUP  
ON DBPARTITIONNUMS( 0 TO 2, 5, 8)
```

- 假设您想要在数据库分区 7 上创建单一分区数据库分区组 `MINGROUP`。该语句如下所示:

```
CREATE DATABASE PARTITION GROUP MINGROUP  
ON DBPARTITIONNUM (7)
```



---

## 第 34 章 受支持的 SQL 管理例程和视图

---

### ADMIN\_CMD 存储过程和关联的管理 SQL 例程

#### 使用 ADMIN\_CMD 过程的 GET STMM TUNING DBPARTITIONNUM 命令

用于读取目录表以报告用户首选的自调整内存管理器（STMM）调整数据库分区号和当前 STMM 调整数据库分区号。

##### 授权

SYSADM 或 DBADM 权限

##### 必需的连接

数据库

##### 命令语法

▶▶ GET—STMM—TUNING—DBPARTITIONNUM ◀◀

##### 示例

```
CALL SYSPROC.ADMIN_CMD( 'get stmm tuning dbpartitionnum' )
```

以下是出自此查询的输出的示例。

Result set 1

```
-----  
USER_PREFERRED_NUMBER CURRENT_NUMBER  
-----  
2 2
```

1 record(s) selected.

Return Status = 0

##### 使用说明

用户首选的自调整内存管理器（STMM）调整数据库分区号（USER\_PREFERRED\_NUMBER）由用户设置，它指定用户想要在其上运行内存调整器的数据库分区。当数据库正在运行时，每小时将异步更新调整分区几次。结果，返回的 CURRENT\_NUMBER 和 USER\_PREFERRED\_NUMBER 可能会在您更新用户首选的 STMM 分区号之后处于不同步状态。要解决此问题，等待异步更新 CURRENT\_NUMBER，或者停止和启动数据库以强制更新 CURRENT\_NUMBER。

## 使用 ADMIN\_CMD 过程的 UPDATE STMM TUNING DBPARTITIONNUM 命令

更新用户首选的自调整内存管理器（STMM）调整数据库分区。

### 授权

SYSADM 或 DBADM 权限

### 必需的连接

数据库

### 命令语法

►►—UPDATE—STMM—TUNING—DBPARTITIONNUM—*partitionnum*—►►

### 命令参数

#### **partitionnum**

*partitionnum* 是整数。如果使用 -1 或不存在的数据库分区号，那么 STMM 调整数据库分区将使用缺省的数据库分区，正如为 STMM 功能所定义的那样。如果使用 -1 或不存在的数据库分区号，那么 DB2 将自动选择要在其上运行 STMM 内存调整器的相应的数据库分区。

### 示例

将用户首选的自调整内存管理器（STMM）调整数据库分区更新为数据库分区 3。

```
CALL SYSPROC.ADMIN_CMD( 'update stmm tuning dbpartitionnum 3' )
```

### 使用说明

STMM 调整进程定期检查用户首选的 STMM 调整数据库分区号值的变化。如果 *partitionnum* 存在，并且是活动数据库分区，那么 STMM 调整进程将移至用户首选的 STMM 调整数据库分区。一旦此命令更改 STMM 调整数据库分区号，当前 STMM 调整数据库分区号就会立即更改。

在 SQLCA 中返回因 CALL 语句生成的命令执行状态。

此命令落实在 ADMIN\_CMD 过程中的更改。

---

## 配置管理 SQL 例程和视图

### DB\_PARTITIONS

DB\_PARTITIONS 表函数以表格形式返回 db2nodes.cfg 文件的内容。

### 语法

►►—DB\_PARTITIONS—(—)—►►

模式为 SYSPROC。

## 授权

对 DB\_PARTITIONS 表函数的 EXECUTE 特权。

## 表函数参数

函数没有输入参数。

## 示例

检索来自 3 逻辑分区数据库的信息。

```
SELECT * FROM TABLE(DB_PARTITIONS()) AS T
```

以下是出自此查询的输出的示例。

| PARTITION_NUMBER | HOST_NAME                | PORT_NUMBER | SWITCH_NAME |
|------------------|--------------------------|-------------|-------------|
| 0                | jessicae.torolab.ibm.com | 0           | jessicae    |
| 1                | jessicae.torolab.ibm.com | 1           | jessicae    |
| 2                | jessicae.torolab.ibm.com | 2           | jessicae    |

选择了 3 个记录。

## 返回的信息

表 42. DB\_PARTITIONS 表函数返回的信息

| 列名               | 数据类型         | 描述                                       |
|------------------|--------------|------------------------------------------|
| PARTITION_NUMBER | SMALLINT     | 介于 0 和 999 之间的唯一编号，标识了分区数据库环境中的数据库分区服务器。 |
| HOST_NAME        | VARCHAR(128) | 数据库分区服务器的 TCP/IP 主机名。                    |
| PORT_NUMBER      | SMALLINT     | 数据库分区服务器的端口号。                            |
| SWITCH_NAME      | VARCHAR(128) | 用于数据库分区通信的高速互连、交换机的名称。                   |

## 按步骤再分发管理 SQL 例程

### STEPWISE\_REDISTRIBUTE\_DBPG 过程 - 再分发部分数据库分区组

STEPWISE\_REDISTRIBUTE\_DBPG 过程根据为该过程指定的输入以及由 SET\_SWRD\_SETTINGS 过程创建或更新的设置文件来再分发部分数据库分区组。

#### 语法

```
▶▶ STEPWISE_REDISTRIBUTE_DBPG (—inDBPGroup—, —inStartingPoint—, —————▶)
```

模式为 SYSPROC。

## 过程参数

### *inDBPGroup*

指定目标数据库分区组的名称的 VARCHAR (128) 类型的输入参数。

### *inStartingPoint*

指定要使用的起始点的 SMALLINT 类型的输入参数。如果该参数被设置为正整数且不为空，那么 STEPWISE\_REDISTRIBUTE\_DBPG 过程使用此值，而不是使用在设置文件中指定的 *nextStep* 值。当您想要从特定步骤重新运行 STEPWISE\_REDISTRIBUTE\_DBPG 过程时，这是非常有用的选项。如果该参数被设置为 NULL，那么将使用 *nextStep* 值。

### *inNumSteps*

指定要运行的步骤编号的 SMALLINT 类型的输入参数。如果该参数被设置为正整数且不为空，那么 STEPWISE\_REDISTRIBUTE\_DBPG 过程使用此值，而不是使用在设置文件中指定的 *stageSize* 值。当您想要使用与设置中指定的编号不同的步骤来重新运行 STEPWISE\_REDISTRIBUTE\_DBPG 过程时，这是非常有用的选项。例如，如果在预定的阶段中有五个步骤，并且再分发过程在步骤 3 失败了，那么可以在纠正错误情况后调用 STEPWISE\_REDISTRIBUTE\_DBPG 过程来运行其余的三个步骤。如果该参数被设置为 NULL，那么将使用 *stageSize* 值。在此过程中可以使用值“-2”来指示该编号是不受限制的。

## 授权

- 对 STEPWISE\_REDISTRIBUTE\_DBPG 过程的 EXECUTE 特权
- SYSADM、SYSCTRL 或 DBADM

## 示例

根据 SET\_SWRD\_SETTINGS 过程存储在注册表中的再分发计划来再分发数据库分区组“IBMDEFAULTGROUP”。它从步骤 3 开始再分发数据，直至完成再分发计划中的 2 个步骤。

```
CALL SYSPROC.STEPWISE_REDISTRIBUTE_DBPG('IBMDEFAULTGROUP', 3, 2)
```

有关按步骤再分发过程的完整使用示例，请参阅 STEPWISE\_REDISTRIBUTE\_DBPG 过程

## 使用说明

如果在开始执行 STEPWISE\_REDISTRIBUTE\_DBPG 过程后使用 SET\_SWRD\_SETTINGS 过程将 *processState* 的注册表值更新为 1，那么该进程会在下一步骤的开始阶段停止，并且返回警告消息。

由于再分发进程调用 SQL COMMIT 语句，所以在 2 类连接下运行再分发进程不受支持。



---

## 第 6 部分 附录



---

## 附录 A. 作为非 root 用户安装

---

### 作为非 root 用户来安装 DB2 产品

可作为非 root 用户安装大多数 DB2 产品。

在作为非 root 用户安装任何 DB2 产品之前，您应该了解 root 用户安装和非 root 用户安装之间的差别以及非 root 用户安装的局限性。请参阅本主题末尾的“相关链接”以了解详细信息。

作为非 root 用户安装 DB2 产品的先决条件:

- 您必须能够安装该安装 DVD 或者自动安装。
- 您必须具有可用作 DB2 实例的所有者的有效用户标识。

用户标识具有下列限制和要求:

- 必须具有除 guests、admins、users 和 local 之外的主组
  - 可以包含小写字母 (a-z)、数字 (0-9) 和下划线字符 ( \_ )
  - 长度不能超过八个字符
  - 不能以 IBM、SYS、SQL 或数字开头
  - 不能是 DB2 保留字 (USERS、ADMINS、GUESTS、PUBLIC 或 LOCAL) 或 SQL 保留字
  - 不能包含重音字符
  - 不能是由 NIS 或 NIS+ 管理的标识
  - 如果已指定现有用户标识，而不是创建新用户标识，那么确保该用户标识:
    - 未锁定
    - 不具有到期的密码
- 对于非 root 用户和 root 用户，您要安装的产品硬件和软件先决条件都相同。
  - 在 AIX V5.3 上，必须启用异步 I/O (AIO)。
  - 您的主目录必须是有效的 DB2 路径。

DB2 安装路径具有下列规则:

- 可以包含小写字母 (a-z)、大写字母 (A-Z) 和下划线字符 ( \_ )
- 不能超过 128 个字符
- 不能包含空格
- 不能包含非英文字符

作为非 root 用户安装 DB2 产品应该对非 root 用户透明。换言之，除了作为非 root 用户登录之外，非 root 用户不需要执行特殊的操作来安装 DB2 产品。为了执行非 root 用户安装:

1. 作为非 root 用户登录
2. 使用任何可用的方法来安装 DB2 产品。选项包括:
  - “DB2 安装”向导 (GUI 安装)

- db2\_install 命令
- db2setup 命令与响应文件（静默安装）

**注：**由于非 root 用户无法选择安装 DB2 产品的目录，所以响应文件中的任何 FILE 关键字将被忽略。

请参阅本主题末尾的“相关链接”以了解详细信息。

3. 在安装 DB2 产品后，您需要打开新登录会话以使用非 root 用户 DB2 实例。另外，如果将 DB2 实例环境的来源设置为 \$HOME/sqlllib/db2profile（对于 Bourne shell 和 Korn shell 用户）或 \$HOME/sqlllib/db2chsrc（对于 C shell 用户），其中 \$HOME 是非 root 用户的主目录，那么也可以使用同一登录会话。

在安装 DB2 产品后，您应该验证操作系统用户进程资源限制（ulimit）。如果不符合最小 ulimit 值，那么 DB2 引擎可能遇到意外的操作资源不足错误。这些错误会导致 DB2 停止运行。

---

## 附录 B. 使用备份

---

### 使用备份

使用 `BACKUP DATABASE` 命令来复制数据库数据并将其存储在另一介质上，以防原始介质发生故障或损坏。可以备份整个数据库，可以备份数据库分区，也可以只备份选择的表空间。

#### 开始前

不必连接到将要备份的数据库：备份数据库实用程序自动建立与指定数据库的连接，而此连接会在备份操作完成时终止。如果已连接到要备份的数据库，当发出 `BACKUP DATABASE` 命令时将断开连接，备份操作将继续进行。

数据库可以是本地数据库或远程数据库。备份映像保留在数据库服务器上，除非您使用的是存储管理产品，如 Tivoli Storage Manager (TSM) 或 DB2 高级副本服务 (ACS)。

如果要执行脱机备份并且已使用 `ACTIVATE DATABASE` 语句激活数据库，那么在运行脱机备份之前必须取消激活该数据库。如果存在与该数据库的活动连接，为了成功取消激活该数据库，具有 `SYSADM` 权限的用户必须连接至该数据库并发出下列命令：

```
CONNECT TO database-alias
QUIESCE DATABASE IMMEDIATE FORCE CONNECTIONS;
UNQUIESCE DATABASE;
TERMINATE;
DEACTIVATE DATABASE database-alias
```

在分区数据库环境中，可以使用 `BACKUP DATABASE` 命令来逐个备份数据库分区，使用 `ON DBPARTITIONNUM` 命令参数来一次性备份多个数据库分区或使用 `ALL DBPARTITIONNUMS` 参数来同时备份所有数据库分区。可以使用 `LIST NODES` 命令来标识具有想要备份的用户表的数据库分区。

除非您在使用单一系统视图 (SSV) 备份，否则，如果您要在分区数据库环境中执行脱机备份，那么应将目录分区独立于所有其他数据库分区进行备份。例如，可以先备份目录分区，然后备份所有其他数据库分区。因为备份操作可能需要在目录分区上进行独占数据库连接（在此期间不能连接其他数据库分区），所以必须执行上述操作。如果要执行联机备份，那么可以同时或以任何顺序备份所有数据库分区（包括目录分区）。

也可以使用“命令编辑器”来备份数据库分区。因为此方法不支持前滚恢复，所以应定期备份这些节点上的数据库。还应与您建立的任何备份副本一起，保留一份 `db2nodes.cfg` 文件，以保护可能对此文件造成的损坏。

在分布式请求系统上，备份操作适用于分布式请求数据库和存储在数据库目录（包装器、服务器和昵称等等）中的元数据。不备份数据源对象（表和视图），除非它们也存储在分布式请求数据库中。

如果一个数据库是使用数据库管理器的前发行版创建的且该数据库还未迁移，那么必须在将该数据库迁移之后，才能对其进行备份。

## 关于此任务

以下限制适用于 `backup` 实用程序:

- 表空间备份操作和表空间复原操作不能同时运行，即使涉及的是不同的表空间。
- 如果想要能够在分区数据库环境中执行前滚恢复，那么必须在节点列表上定期备份数据库且必须具有系统中余下节点的至少一个备份映像（即使不包含该数据库的用户数据）。下列两种情况都需要在不包含数据库的用户数据的数据库分区服务器中存在数据库分区的备份映像：
  - 在建立上一个备份之后已将一个数据库分区服务器添加到数据库系统，因此需要在此数据库分区服务器上执行正向恢复。
  - 使用时间点恢复，它要求系统中的所有数据库分区都处于前滚暂挂状态。
- DMS 表空间的联机备份操作与下列操作不兼容：
  - 装入
  - 重组（联机和脱机）
  - 删除表空间
  - 截断表
  - 创建索引
  - 最初未记录任何内容（与 `CREATE TABLE` 和 `ALTER TABLE` 语句配合使用）
- 如果尝试对当前活动的数据库执行脱机备份，那么将接收到错误。在运行脱机备份之前，可以通过发出 `DEACTIVATE DATABASE` 命令来确保数据库未处于活动状态。

通过命令行处理器（CLP）、控制中心的“备份数据库”向导、运行带 `BACKUP DATABASE` 参数的 `ADMIN_CMD` 过程或 `db2Backup` 应用程序编程接口（API），可以调用备份实用程序。

以下是通过 CLP 发出的 `BACKUP DATABASE` 命令的示例:

```
db2 backup database sample to c:\DB2Backups
```

## 过程

要打开“备份数据库”向导:

1. 从“控制中心”中，展开对象树，直到找到要备份的数据库或表空间对象为止。
2. 右键单击该对象，然后从弹出菜单中选择“备份”。这就打开了“备份数据库”向导。

## 下一步任务

在“控制中心”中通过上下文帮助工具提供详细信息。

如果执行了脱机备份，那么在备份完成后，必须重新激活该数据库:

```
ACTIVATE DATABASE sample
```

---

## 附录 C. 分区数据库环境目录视图

---

### SYSCAT.BUFFERPOOLDBPARTITIONS

每一行表示缓冲池和数据库分区的组合，其中该分区上的缓冲池大小与同一数据库分区组中其他分区的缓冲池缺省大小不同（如 SYSCAT.BUFFERPOOLS 中所述）。

表 43. SYSCAT.BUFFERPOOLDBPARTITIONS 目录视图

| 列名             | 数据类型     | 可空 | 描述                |
|----------------|----------|----|-------------------|
| BUFFERPOOLID   | INTEGER  |    | 内部缓冲池标识。          |
| DBPARTITIONNUM | SMALLINT |    | 数据库分区号。           |
| NPAGES         | INTEGER  |    | 此数据库分区上的此缓冲池中的页数。 |

---

### SYSCAT.DATAPARTITIONEXPRESSION

每一行均表示用于该部分表分区键的表达式。

表 44. SYSCAT.DATAPARTITIONEXPRESSION 目录视图

| 列名                      | 数据类型          | 可空 | 描述                                                                                                               |
|-------------------------|---------------|----|------------------------------------------------------------------------------------------------------------------|
| TABSCHEMA               | VARCHAR (128) |    | 分区表的模式名。                                                                                                         |
| TABNAME                 | VARCHAR (128) |    | 分区表的非标准名称。                                                                                                       |
| DATAPARTITIONKEYSEQ     | INTEGER       |    | 从 1 开始的表达式关键部分序列标识。                                                                                              |
| DATAPARTITIONEXPRESSION | CLOB (32K)    |    | SQL 语法中该序列中此条目的表达式。                                                                                              |
| NULLSFIRST              | CHAR (1)      |    | <ul style="list-style-type: none"><li>• N = 此表达式 compare high 中的空值</li><li>• Y = 此表达式 compare low 中的空值</li></ul> |

---

### SYSCAT.DATAPARTITIONS

每一行均表示数据分区。

表 45. SYSCAT.DATAPARTITIONS 目录视图

| 列名                | 数据类型          | 可空 | 描述                                  |
|-------------------|---------------|----|-------------------------------------|
| DATAPARTITIONNAME | VARCHAR (128) |    | 数据分区的名称。                            |
| TABSCHEMA         | VARCHAR (128) |    | 此数据分区所属的表的模式名。                      |
| TABNAME           | VARCHAR (128) |    | 此数据分区所属的表的非标准名称。                    |
| DATAPARTITIONID   | INTEGER       |    | 数据分区的标识。                            |
| TBSPACEID         | INTEGER       | Y  | 存储此数据分区的表空间的标识。STATUS 为“T”时则为 Null。 |
| PARTITIONOBJECTID | INTEGER       | Y  | 表空间中数据分区的标识。                        |
| LONG_TBSPACEID    | INTEGER       | Y  | 存储长数据的表空间的标识。STATUS 为“T”时则为 Null。   |

表 45. SYSCAT.DATAPARTITIONS 目录视图 (续)

| 列名            | 数据类型         | 可空 | 描述                                                                                                                                                                                                                                          |
|---------------|--------------|----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ACCESS_MODE   | CHAR (1)     |    | <p>数据分区的访问限制状态。这些状态仅适用于处于集合完整性暂挂状态的对象或被 SET INTEGRITY 语句处理的对象。可能的值包括:</p> <ul style="list-style-type: none"> <li>• D = 没有数据移动</li> <li>• F = 完全访问</li> <li>• N = 无访问</li> <li>• R = 只读访问</li> </ul>                                         |
| STATUS        | VARCHAR(32)  |    | <ul style="list-style-type: none"> <li>• A = 新近连接数据分区</li> <li>• D = 已拆离数据分区</li> <li>• I = 其条目在目录中的已拆离的数据分区仅在异步索引清除期间才会被维护; 当引用已拆离分区的所有索引记录删除后, STATUS 值为“T”的行将会被除去</li> <li>• 空字符串 = 数据分区可视 (正常状态)</li> </ul> <p>字节 2 至 32 被保留以供将来使用。</p> |
| SEQNO         | INTEGER      |    | 数据分区序号 (从 0 开始)。                                                                                                                                                                                                                            |
| LOWINCLUSIVE  | CHAR (1)     |    | <ul style="list-style-type: none"> <li>• N = 不包括低键值</li> <li>• Y = 包括低键值</li> </ul>                                                                                                                                                         |
| LOWVALUE      | VARCHAR(512) |    | 用于此数据分区的低键值 (SQL 值的字符串表示法)。                                                                                                                                                                                                                 |
| HIGHINCLUSIVE | CHAR (1)     |    | <ul style="list-style-type: none"> <li>• N = 不包括高键值</li> <li>• Y = 包括高键值</li> </ul>                                                                                                                                                         |
| HIGHVALUE     | VARCHAR(512) |    | 用于此数据分区的高键值 (SQL 值的字符串表示法)。                                                                                                                                                                                                                 |

## SYSCAT.DBPARTITIONGROUPDEF

每一行均表示数据库分区组中包含的数据库分区。

表 46. SYSCAT.DBPARTITIONGROUPDEF 目录视图

| 列名             | 数据类型          | 可空 | 描述                                                    |
|----------------|---------------|----|-------------------------------------------------------|
| DBPGNAME       | VARCHAR (128) |    | 包含数据库分区的数据库分区组的名称。                                    |
| DBPARTITIONNUM | SMALLINT      |    | 数据库分区组中包含的数据库分区的分区号。有效的分区号介于 0 和 999 之间 (其中包括 0 和 9)。 |



表 46. SYSCAT.DBPARTITIONGROUPDEF 目录视图 (续)

| 列名     | 数据类型     | 可空 | 描述                                                                                                                                                                                                                                                                                                         |
|--------|----------|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| IN_USE | CHAR (1) |    | 数据库分区的状态。<br><ul style="list-style-type: none"> <li>• A = 新添加的数据库分区不在分布图中，但该数据库分区组中的表空间的容器已创建；当再分发数据库分区组操作成功完成后，会将数据库分区添加至分布图</li> <li>• D = 当再分发数据库分区组操作成功完成后，将删除该数据库分区</li> <li>• T = 新添加的数据库分区不在分布图中，并且它已使用 <b>WITHOUT TABLESPACES</b> 子句进行添加；您必须将容器添加至数据库分区组中的表空间</li> <li>• Y = 数据库分区在分布图中</li> </ul> |

## SYSCAT.DBPARTITIONGROUPS

每一行均表示数据库分区组。

表 47. SYSCAT.DBPARTITIONGROUPS 目录视图

| 列名                   | 数据类型          | 可空 | 描述                                                                                     |
|----------------------|---------------|----|----------------------------------------------------------------------------------------|
| DBPGNAME             | VARCHAR (128) |    | 数据库分区组的名称。                                                                             |
| OWNER                | VARCHAR (128) |    | 创建数据库分区组时所采用的授权标识。                                                                     |
| OWNERTYPE            | CHAR (1)      |    | <ul style="list-style-type: none"> <li>• S = 系统的所有者</li> <li>• U = 所有者是用户个体</li> </ul> |
| PMAP_ID              | SMALLINT      |    | SYSCAT.PARTITIONMAPS 目录视图中分布图的标识。                                                      |
| REDISTRIBUTE_PMAP_ID | SMALLINT      |    | 当前进行再分发所使用的分布图的标识；如果当前未在进行再分发，那么它为 -1。                                                 |
| CREATE_TIME          | TIMESTAMP     |    | 数据库分区组的创建时间。                                                                           |
| DEFINER <sup>1</sup> | VARCHAR (128) |    | 创建数据库分区组时所采用的授权标识。                                                                     |
| REMARKS              | VARCHAR(254)  | Y  | 用户提供的注释，或者为空值。                                                                         |

注:

1. 包括 DEFINER 列以便实现向后兼容。请参阅 OWNER。

## SYSCAT.PARTITIONMAPS

每一行均表示可用于根据散列表的分布键在数据库分区组中的数据库分区之间分布表行的分布图。

表 48. SYSCAT.PARTITIONMAPS 目录视图

| 列名      | 数据类型     | 可空 | 描述      |
|---------|----------|----|---------|
| PMAP_ID | SMALLINT |    | 分布图的标识。 |

表 48. SYSCAT.PARTITIONMAPS 目录视图 (续)

| 列名           | 数据类型        | 可空 | 描述                                                                |
|--------------|-------------|----|-------------------------------------------------------------------|
| PARTITIONMAP | BLOB (8192) |    | 分布图，用于多分区数据库分区组的 4096 个双字节整数的向量。对于单分区的数据数据库分区组，存在有一个条目，表明单分区的分区号。 |

---

## 附录 D. DB2 技术信息概述

可以通过下列工具和方法获取 DB2 技术信息:

- DB2 信息中心
  - 主题 (任务、概念和参考主题)
  - DB2 工具的帮助
  - 样本程序
  - 教程
- DB2 书籍
  - PDF 文件 (可下载)
  - PDF 文件 (在 DB2 PDF DVD 中)
  - 印刷版书籍
- 命令行帮助
  - 命令帮助
  - 消息帮助

**注:** DB2 信息中心主题的更新频率比 PDF 书籍或硬拷贝书籍的更新频率高。要获取最新信息, 请安装可用的文档更新, 或者参阅 [ibm.com](http://www.ibm.com)<sup>®</sup> 上的 DB2 信息中心。

可以在线访问 [ibm.com](http://www.ibm.com) 上的其他 DB2 技术信息, 如技术说明、白皮书和 IBM Redbooks<sup>®</sup> 出版物。访问位于以下网址的 DB2 信息管理软件库站点: <http://www.ibm.com/software/data/sw-library/>。

### 文档反馈

我们非常重视您对 DB2 文档的反馈。如果您想就如何改善 DB2 文档提出建议, 请将电子邮件发送至 [db2docs@ca.ibm.com](mailto:db2docs@ca.ibm.com)。DB2 文档小组会阅读您的所有反馈, 但不能直接答复您。请尽可能提供具体的示例, 这样我们才能更好地了解您所关心的问题。如果您要提供有关具体主题或帮助文件的反馈, 请加上标题和 URL。

请不要用以上电子邮件地址与 DB2 客户支持机构联系。如果您遇到文档不能解决的 DB2 技术问题, 请与您当地的 IBM 服务中心联系以获得帮助。

---

## 硬拷贝或 PDF 格式的 DB2 技术库

下列各表描述 IBM 出版物中心 (网址为 [www.ibm.com/shop/publications/order](http://www.ibm.com/shop/publications/order)) 提供的 DB2 资料库。可以从 [www.ibm.com/support/docview.wss?rs=71&uid=swg2700947](http://www.ibm.com/support/docview.wss?rs=71&uid=swg2700947) 下载 PDF 格式的英文 DB2 版本 9.5 手册和已翻译的版本。

尽管这些表标识书籍有印刷版, 但可能未在您所在国家或地区提供。

表 49. DB2 技术信息

| 书名                                  | 书号           | 是否提供印刷版 |
|-------------------------------------|--------------|---------|
| <i>Administrative API Reference</i> | SC23-5842-00 | 是       |

表 49. DB2 技术信息 (续)

| 书名                                                           | 书号           | 是否提供印刷版 |
|--------------------------------------------------------------|--------------|---------|
| <i>Administrative Routines and Views</i>                     | SC23-5843-00 | 否       |
| <i>Call Level Interface Guide and Reference, Volume 1</i>    | SC23-5844-00 | 是       |
| <i>Call Level Interface Guide and Reference, Volume 2</i>    | SC23-5845-00 | 是       |
| <i>Command Reference</i>                                     | SC23-5846-00 | 是       |
| 《数据移动指南和参考》                                                  | S151-0617-00 | 是       |
| 《数据恢复及高可用性指南与参考》                                             | S151-0619-00 | 是       |
| 《数据服务器、数据库和数据库对象指南》                                          | S151-0612-00 | 是       |
| 《数据库安全性指南》                                                   | S151-0614-00 | 是       |
| <i>Developing ADO.NET and OLE DB Applications</i>            | SC23-5851-00 | 是       |
| <i>Developing Embedded SQL Applications</i>                  | SC23-5852-00 | 是       |
| <i>Developing Java Applications</i>                          | SC23-5853-00 | 是       |
| <i>Developing Perl and PHP Applications</i>                  | SC23-5854-00 | 否       |
| <i>Developing User-defined Routines (SQL and External)</i>   | SC23-5855-00 | 是       |
| <i>Getting Started with Database Application Development</i> | GC23-5856-00 | 是       |
| 《Linux 和 Windows 上的 DB2 安装和管理入门》                             | G151-0623-00 | 是       |
| 《国际化指南》                                                      | S151-0616-00 | 是       |
| 《消息参考, 第 1 卷》                                                | G151-0632-00 | 否       |
| 《消息参考, 第 2 卷》                                                | G151-0633-00 | 否       |
| 《迁移指南》                                                       | G151-0622-00 | 是       |
| 《Net Search Extender 管理和用户指南》                                | S151-0761-00 | 是       |
| 注: DB2 信息中心中并不包含此文档的内容                                       |              |         |
| 《分区和集群指南》                                                    | S151-0615-00 | 是       |
| <i>Query Patroller Administration and User's Guide</i>       | SC23-8507-00 | 是       |
| 《IBM 数据服务器客户机快速入门》                                           | G151-0625-00 | 否       |
| 《DB2 服务器快速入门》                                                | G151-0624-00 | 是       |
| 《Spatial Extender 和地理数据管理功能部件用户指南和参考》                        | S151-0760-00 | 是       |
| <i>SQL Reference, Volume 1</i>                               | SC23-5861-00 | 是       |
| <i>SQL Reference, Volume 2</i>                               | SC23-5862-00 | 是       |

表 49. DB2 技术信息 (续)

| 书名                                          | 书号           | 是否提供印刷版 |
|---------------------------------------------|--------------|---------|
| 《系统监视器指南和参考》                                | S151-0618-00 | 是       |
| 《文本搜索指南》                                    | S151-0620-00 | 是       |
| 《故障诊断指南》                                    | G151-0621-00 | 否       |
| 《调整数据库性能》                                   | S151-0613-00 | 是       |
| 《Visual Explain 教程》                         | S151-0634-00 | 否       |
| 《新增内容》                                      | S151-0629-00 | 是       |
| <i>Workload Manager Guide and Reference</i> | SC23-5870-00 | 是       |
| 《pureXML 指南》                                | S151-0630-00 | 是       |
| 《XQuery 参考》                                 | S151-0631-00 | 否       |

表 50. 特定于 DB2 Connect 的技术信息

| 书名                    | 书号           | 是否提供印刷版 |
|-----------------------|--------------|---------|
| 《DB2 Connect 个人版快速入门》 | G151-0627-00 | 是       |
| 《DB2 Connect 服务器快速入门》 | G151-0628-00 | 是       |
| 《DB2 Connect 用户指南》    | S151-0626-00 | 是       |

表 51. Information Integration 技术信息

| 书名                                                                                            | 书号           | 是否提供印刷版 |
|-----------------------------------------------------------------------------------------------|--------------|---------|
| <i>Information Integration: Administration Guide for Federated Systems</i>                    | SC19-1020-01 | 是       |
| <i>Information Integration: ASNCLP Program Reference for Replication and Event Publishing</i> | SC19-1018-02 | 是       |
| <i>Information Integration: 《联合数据源配置指南》</i>                                                   | S151-0468-00 | 否       |
| <i>Information Integration: 《SQL 复制指南和参考》</i>                                                 | S151-0475-00 | 是       |
| <i>Information Integration: Introduction to Replication and Event Publishing</i>              | SC19-1028-01 | 是       |

## 订购印刷版的 DB2 书籍

如果您需要印刷版的 DB2 书籍，可以在许多（但不是所有）国家或地区在线购买。无论何时都可以从当地的 IBM 代表处订购印刷版的 DB2 书籍。请注意，DB2 PDF 文档 DVD 上的某些软拷贝书籍没有印刷版。例如，DB2 消息参考的任何一卷都没有提供印刷版书籍。

只要支付一定费用，就可以从 IBM 获取 DB2 PDF 文档 DVD，该 DVD 包含许多 DB2 书籍的印刷版。根据您下订单的位置，您可能能够从 IBM 出版物中心在线订购书籍。如果在线订购在您所在国家或地区不可用，您总是可以从当地的 IBM 代表处订购印刷版 DB2 书籍。注意，并非 DB2 PDF 文档 DVD 上的所有书籍都有印刷版。

注：最新最完整的 DB2 文档保留在网址如下的 DB2 信息中心中：<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5>。

要订购印刷版的 DB2 书籍：

- 要了解您是否可从所在国家或地区在线订购印刷版的 DB2 书籍，可查看 IBM 出版物中心站点，网址为：<http://www.ibm.com/shop/publications/order>。必须先选择国家、地区或语言才能访问出版物订购信息，然后再按照针对您所在位置的订购指示信息进行订购。
- 要从当地的 IBM 代表处订购印刷版的 DB2 书籍：
  1. 从下列其中一个 Web 站点找到当地代表处的联系信息：
    - IBM 全球联系人目录，网址为 [www.ibm.com/planetwide](http://www.ibm.com/planetwide)
    - IBM 出版物 Web 站点，网址为 <http://www.ibm.com/shop/publications/order>。必须先选择国家、地区或语言才能访问对应您的所在地的出版物主页。在此页面中访问“关于此站点”链接。
  2. 请在致电时说明您想订购 DB2 出版物。
  3. 请向您当地的代表处提供想要订购的书籍的书名和书号。有关书名和书号的信息，请参阅第 373 页的『硬拷贝或 PDF 格式的 DB2 技术库』。

---

## 从命令行处理器显示 SQL 状态帮助

DB2 返回描述 SQL 语句执行结果的 SQLSTATE。SQLSTATE 帮助说明 SQL 状态和 SQL 状态类代码的含义。

要调用 SQL 状态帮助，打开命令行处理器并输入：

```
? sqlstate or ? class code
```

其中，*sqlstate* 表示有效的 5 位 SQL 状态，*class code* 表示该 SQL 状态的前 2 位。

例如，? 08003 显示 08003 SQL 状态的帮助，而 ? 08 显示 08 类代码的帮助。

---

## 访问不同版本的 DB2 信息中心

对于 DB2 版本 9.5 主题，DB2 信息中心 URL 为 <http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/>

对于 DB2 版本 9 主题，DB2 信息中心 URL 为 <http://publib.boulder.ibm.com/infocenter/db2luw/v9/>

对于 DB2 版本 8 主题，请访问以下版本 8 信息中心 URL：<http://publib.boulder.ibm.com/infocenter/db2luw/v8/>

---

## 在 DB2 信息中心中以您的首选语言显示主题:

DB2 信息中心尝试以您在浏览器首选项中指定的语言显示主题。如果未提供主题的首选语言翻译版本,那么 DB2 信息中心将显示该主题的英文版。

- 要在 Internet Explorer 浏览器中以您的首选语言显示主题:

1. 在 Internet Explorer 中,单击工具 → Internet 选项 → 语言...按钮。“语言首选项”窗口打开。
2. 确保您的首选语言被指定为语言列表中的第一个条目。
  - 要将新语言添加至列表,单击添加... 按钮。

**注:** 添加语言并不能保证计算机具有以首选语言显示主题所需的字体。

- 要将语言移至列表顶部,选择该语言并单击上移按钮直到该语言成为语言列表中的第一个条目。
3. 清除浏览器高速缓存然后刷新页面以便以首选语言显示 DB2 信息中心。
- 要在 Firefox 或 Mozilla 浏览器中以首选语言显示主题:
  1. 在工具 → 选项 → 高级对话框中的语言部分中选择按钮。“语言”面板将显示在“首选项”窗口中。
  2. 确保您的首选语言被指定为语言列表中的第一个条目。
    - 要将新语言添加至列表,单击添加... 按钮以从“添加语言”窗口中选择一种语言。
    - 要将语言移至列表顶部,选择该语言并单击上移按钮直到该语言成为语言列表中的第一个条目。
  3. 清除浏览器高速缓存然后刷新页面以便以首选语言显示 DB2 信息中心。

在某些浏览器和操作系统组合上,可能还必须将操作系统的区域设置更改为您选择的语言环境和语言。

---

## 更新安装在您的计算机或内部网服务器上的 DB2 信息中心

如果已经在本地安装了 DB2 信息中心,您可以下载并安装 IBM 提供的更新。

更新在本地安装的 DB2 信息中心要求您:

1. 停止计算机上的 DB2 信息中心,然后以独立方式重新启动信息中心。如果以独立方式运行信息中心,那么网络上的其他用户将无法访问信息中心,因而您可以下载和应用更新。
2. 使用“更新”功能部件来查看可用的更新。如果有您希望安装的更新,那么请使用“更新”功能部件来下载并安装这些更新。

**注:** 如果您的环境要求在一台未连接至因特网的机器上安装 DB2 信息中心更新,那么必须使用一台已连接至因特网的机器将更新站点镜像至本地文件系统并安装 DB2 信息中心。如果网络中有许多用户将安装文档更新,那么可以通过在本地也为更新站点建立镜像并为更新站点创建代理来缩短每个人执行更新所需要的时间。

如果提供了更新包,那么使用更新功能来下载这些更新包。但是,只有在单机方式下才能使用更新功能。

3. 停止独立信息中心,然后在计算机上重新启动 DB2 信息中心。

**注：**在 Windows Vista 上，必须以管理员身份才能运行下面所列示的命令。要启动具有所有管理员特权的命令提示符或图形工具，右键单击快捷方式，然后选择**以管理员身份运行**。

要更新安装在您的计算机或内部网服务器上的 DB2 信息中心：


1. 停止 DB2 信息中心。
  - 在 Windows 上，单击**开始** → **控制面板** → **管理工具** → **服务**。右键单击 **DB2 信息中心** 服务，并选择**停止**。
  - 在 Linux 上，输入以下命令：  

```
/etc/init.d/db2icdv95 stop
```
2. 以独立方式启动信息中心。
  - 在 Windows 上：
    - a. 打开命令窗口。
    - b. 浏览至信息中心的安装位置。缺省情况下，DB2 信息中心安装在 <Program Files>\IBM\DB2 Information Center\Version 9.5 目录中，其中 <Program Files> 表示 Program Files 目录的位置。
    - c. 从安装目录浏览至 doc\bin 目录。
    - d. 运行 help\_start.bat 文件：  

```
help_start.bat
```
  - 在 Linux 上：
    - a. 浏览至信息中心的安装位置。缺省情况下，DB2 信息中心安装在 /opt/ibm/db2ic/V9.5 目录中。
    - b. 从安装目录浏览至 doc/bin 目录。
    - c. 运行 help\_start 脚本：  

```
help_start
```

系统缺省 Web 浏览器将启动以显示独立信息中心。

3. 单击“更新”按钮 (  )。在信息中心的右边面板上，单击**查找更新**。将显示现有文档的更新列表。
4. 要启动下载进程，请检查您想要下载的选项，然后单击**安装更新**。
5. 在完成下载和安装进程后，单击**完成**。
6. 停止独立信息中心。
  - 在 Windows 上，浏览至安装目录的 doc\bin 目录并运行 help\_end.bat 文件：  

```
help_end.bat
```

**注：**help\_end 批处理文件包含安全终止用 help\_start 批处理文件启动的进程所需的命令。不要使用 Ctrl-C 或任何其他方法来终止 help\_start.bat。
  - 在 Linux 上，浏览至安装目录的 doc/bin 目录并运行 help\_end 脚本：  

```
help_end
```

**注：**help\_end 脚本包含安全终止用 help\_start 脚本启动的进程所需的命令。不要使用任何其他方法来终止 help\_start 脚本。
7. 重新启动 DB2 信息中心。



- 在 Windows 上，单击开始 → 控制面板 → 管理工具 → 服务。右键单击 **DB2 信息中心** 服务，并选择启动。
- 在 Linux 上，输入以下命令：  
`/etc/init.d/db2icdv95 start`

更新后的 DB2 信息中心将显示新的主题和更新后的主题。

---

## DB2 教程

DB2 教程帮助您了解 DB2 产品的各个方面。这些课程提供了逐步指示信息。

### 开始之前

可从信息中心查看 XHTML 版的教程：<http://publib.boulder.ibm.com/infocenter/db2help/>。

某些课程使用了样本数据或代码。有关其特定任务的任何先决条件的描述，请参阅教程。

### DB2 教程

要查看教程，请单击标题。

《*pureXML 指南*》中的『**pureXML™**』

设置 DB2 数据库以存储 XML 数据以及如何对本机 XML 数据存储执行基本操作。

《*Visual Explain 教程*》中的『**Visual Explain**』

使用 Visual Explain 来分析、优化和调整 SQL 语句以获取更好的性能。

---

## DB2 故障诊断信息

很多故障诊断和问题确定信息可帮助您使用 DB2 产品。

### DB2 文档

故障诊断信息可在 DB2 信息中心的“DB2 故障诊断指南”或“支持和故障诊断”部分找到。可在该处找到有关如何使用 DB2 诊断工具和实用程序隔离和找出问题的信息、某些最常见问题的解决方案以及有关如何解决使用 DB2 产品时可能遇到的问题建议。

### DB2 技术支持 Web 站点

如果您遇到了问题并且想要获取查找可能的原因和解决方案的帮助，请参阅 DB2 技术支持 Web 站点。该“技术支持”站点具有指向最新 DB2 出版物、技术说明、授权程序分析报告（APAR 或错误修订）、修订包和其他资源的链接。可搜索此知识库并查找问题的可能解决方案。

访问位于以下网址的 DB2 技术支持 Web 站点：<http://www.ibm.com/software/data/db2/udb/support.html>

---

## 条款和条件

如果符合以下条款和条件，那么授予您使用这些出版物的准用权。

**个人使用:** 只要保留所有的专有权声明, 您就可以为个人、非商业使用复制这些出版物。未经 IBM 明确同意, 您不可以分发、展示或制作这些出版物或其中任何部分的演绎作品。

**商业使用:** 只要保留所有的专有权声明, 您就可以仅在企业内复制、分发和展示这些出版物。未经 IBM 明确同意, 您不可以制作这些出版物的演绎作品, 或者在您的企业外部复制、分发或展示这些出版物或其中的任何部分。

除非本准用权中有明确授权, 不得把其他准用权、许可或权利(无论是明示的还是暗含的)授予其中包含的出版物或任何信息、数据、软件或其他知识产权。

当使用这些出版物损害了 IBM 的利益, 或者根据 IBM 的规定, 未正确遵守上述指导说明时, 那么 IBM 保留自主决定撤销本文授予的准用权的权利。

您不可以下载、出口或再出口本信息, 除非完全遵守所有适用的法律和法规, 包括所有美国出口法律和法规。

IBM 对这些出版物的内容不作任何保证。这些出版物“按现状”提供, 不附有任何种类的(无论是明示的还是暗含的)保证, 包括但不限于暗含的关于适销和适用于某种特定用途的保证。

---

## 附录 E. 声明

本信息是为在美国提供的产品和服务编写的。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务，那么由用户自行负责。

IBM 公司可能已拥有或正在申请与本文档内容有关的各项专利。提供本文档并未授予用户使用这些专利的任何许可。您可以用书面方式将许可查询寄往：

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

有关双字节（DBCS）信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

IBM World Trade Asia Corporation  
Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106, Japan

**本条款不适用英国或任何这样的条款与当地法律不一致的国家或地区：** International Business Machines Corporation“按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某种特定用途的保证。某些国家或地区在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息中可能包含技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将编入本资料的新版本中。IBM 可以随时对本资料中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

本文档可提供非 IBM Web 站点和资源的链接和引用。IBM 对本文中引用、提供或链接到的任何非 IBM Web 站点或第三方资源不承担任何责任、保证或其他义务。对非 IBM Web 站点的链接并不意味着 IBM 认可这些 Web 站点或其拥有者的内容或用途。此外，IBM 并不是您使用第三方可进入的任何事务的一方，也不会为这些事务负责，即使您是通过 IBM 站点了解或链接到第三方。因此，您承认并认可 IBM 不会为这些外部站点或资源的可用性负责，也不会为它们提供的任何内容、服务、产品或其他事务承担责任或义务。第三方提供的任何软件都遵守与该软件相关的许可条款和条件。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：（i）允许在独立创建的程序和其他程序（包括本程序）之间进行信息交换，以及（ii）允许对已经交换的信息进行相互使用，请与下列地址联系：

IBM Canada Limited  
Office of the Lab Director  
8200 Warden Avenue  
Markham, Ontario  
L6G 1C7  
CANADA

只要遵守适当的条件和条款，包括某些情形下的一定数量的付费，都可获得这方面的信息。

本资料中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际软件许可协议或任何同等协议中的条款提供。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其他操作环境中获得的数据可能会有明显的不同。有些测量可能是在开发级的系统上进行的，因此不保证与一般可用系统上进行的测量结果相同。此外，有些测量是通过推算而估计的，实际结果可能会有差异。本文档的用户应当验证其特定环境的适用数据。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

所有关于 IBM 未来方向或意向的声明都可随时更改或收回，而不另行通知，它们仅仅表示了目标和意愿而已。

本信息可能包含在日常业务操作中使用的数据和报告的示例。为了尽可能完整地说明这些示例，示例中可能会包括个人、公司、品牌和产品的名称。所有这些名称都是虚构的，与实际商业企业所用的名称和地址的任何雷同纯属巧合。

版权许可：

本信息可能包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口（API）进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。

凡这些样本程序的每份拷贝或其任何部分或任何衍生产品，都必须包括如下版权声明：

©（贵公司的名称）（年）。此部分代码是根据 IBM 公司的样本程序衍生出来的。© Copyright IBM Corp.（输入年份）。All rights reserved.

## 商标

DB2 版本 9.5 文档库的各个文档中标识的公司、产品或服务名称可能是 International Business Machines Corporation 或其他公司的商标或服务标记。有关 IBM Corporation 在美国和 / 或其他国家的商标的信息在 <http://www.ibm.com/legal/copytrade.shtml> 中。

下列各项是其他公司的商标或注册商标, 且已在 DB2 文档库中的至少一份文档中使用:

Microsoft<sup>®</sup>、Windows、Windows NT<sup>®</sup> 和 Windows 徽标是 Microsoft Corporation 在美国和 / 或其他国家或地区的商标。

Intel<sup>®</sup>、Intel 徽标、Intel Inside<sup>®</sup> 徽标、Intel Centrino<sup>®</sup>、Intel Centrino 徽标、Celeron<sup>®</sup>、Intel Xeon<sup>®</sup>、Intel SpeedStep<sup>®</sup>、Itanium<sup>®</sup> 和 Pentium<sup>®</sup> 是 Intel Corporation 在美国和 / 或其他国家或地区的商标。

Java 和所有基于 Java 的商标是 Sun Microsystems,Inc. 在美国和 / 或其他国家或地区的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

Linux 是 Linus Torvalds 在美国和其他国家或地区的注册商标。

Adobe<sup>®</sup>、Adobe 徽标、PostScript<sup>®</sup> 和 PostScript 徽标是 Adobe Systems Incorporated 在美国和 / 或其他国家或地区的注册商标或商标。

其他公司、产品或服务名称可能是其他公司的商标或服务标记。



# 索引

## [ A ]

- 安装
  - 方法
    - UNIX 85
    - Windows 85
  - 数据库分区服务器
    - 响应文件 104
- 安装数据库分区服务器 103
- 安装 DB2 产品
  - 作为非 root 用户 365
- 按表达式进行的分段
  - 与表分区比较 17
- 按范围分区的表
  - 分区表 10

## [ B ]

- 帮助
  - 显示 377
  - 有关 SQL 语句 376
- 本书的结构 vii
- 本书的使用对象 vii
- 崩溃恢复 225
- 边界范围
  - 指定 153
- 标准输入 142
- 表
  - 并置 4, 9
  - 创建
    - 在分区数据库中 151
  - 队列, 用于分区数据库中的连接策略 283
  - 多维集群 271
  - 多维集群 (MDC) 35
  - 范围集群 31, 163
  - 分区 10
  - 分区表 10, 26, 61, 158, 241
  - 改变分区表 181, 182
  - 具体化查询表 158
  - 迁移到分区表中 156
  - 转换 156
  - MDC (多维集群) 表 26, 61, 241
- 表的比较
  - 常规和多维集群 35
- 表分区
  - 描述 10
  - 优点 10
- 表空间
  - 创建
    - 在数据库分区组中 26, 128

- 并行性
  - 查询 67
  - 分区间 67
  - 分区内
    - 描述 67
    - 启用 121
    - 优化策略 273
  - 概述 71
  - 和不同的硬件环境 72
  - 和索引创建 67
  - 启用分区内并行性配置参数 327
  - 实用程序 67
  - 数据库 backup 和 restore 实用程序 67
  - 最大查询并行度配置参数 327
  - I/O 67
  - load 实用程序 67, 189
- 并集
  - 所有视图, 转换 156
- 并置
  - 表 4, 9
- 部分分区 4, 71

## [ C ]

- 查询
  - 并行性 67
  - 多维集群 37
- 查询间并行性 67
- 查询内并行性 67
- 查询优化
  - 数据库分区组的影响 283
- 重新启动装入操作
  - 多分区数据库装入操作 201
- 处理器
  - 添加 122
- 创建
  - 多维表 44, 165
- 从实例中删除数据库分区服务器命令 348
- 错误消息
  - 将节点添加至分区数据库时 133

## [ D ]

- 大对象 (LOB)
  - 行为 152
  - 使用分区表 152
- 代理程序
  - 在分区数据库中 269
- 单处理器环境 72
- 单分区
  - 单处理器环境 72

- 单一分区 (续)
  - 多处理器环境 72
- 单一性 44, 165
- 调整分区
  - 确定 318
- 订购 DB2 书籍 375
- 端口范围
  - 可用性
    - UNIX 94
  - 缺省 (UNIX) 124
- 端口号
  - 范围
    - 定义 135
- 对数据进行分区
  - 表分区 10
  - 分区兼容性 349
  - 管理 3, 109
  - 跨多个分区 4
  - 示例 139
- 多分区配置 72
- 多分区数据库
  - 从单分区数据库进行迁移
    - 设计顾问程序 259
  - 数据库分区组 5
- 多逻辑节点
  - 配置 119
- 多维集群表 51
- 多维集群 (MDC) 表
  - 创建 44, 165
  - 将列表达式用作维 44, 165
  - 将数据移至 44, 165
  - 选择维 37
  - 在 SMS 表空间中 44, 165
  - 值的密度 37
  - 装入注意事项 190

## [ F ]

- 范围
  - 生成 153
  - 为数据分区定义 153
  - 限制 153
- 范围分区
  - 请参阅表分区 10
  - 请参阅数据分区 10, 12
- 范围集群表
  - 表锁定 31, 33
  - 不兼容性 32
  - 超出范围的记录键 31, 32
  - 方案 163
  - 访问路径确定 31, 163
  - 描述 31
  - 算法 161
  - 索引 162
  - 优点 31
  - 与常规表的差别 162

- 范围集群表 (续)
  - 准则 31, 163
- 方案 51
  - 范围集群表 163
- 访问方案
  - 创建其他索引
    - 在分区数据库环境中 303
  - 对用于连接表的列创建索引
    - 在分区数据库环境中 299
  - 改进
    - 在分区数据库环境中 292
  - 没有索引和统计信息的查询
    - 在分区数据库环境中 292
  - 收集当前统计信息
    - 在分区数据库环境中 295
  - 输出示例
    - 无并行性 275
- 非 root 用户安装 365
- 分布键
  - 表注意事项 151
  - 描述 8
  - 装入数据 193
- 分布图
  - 描述 7
- 分发数据
  - 兼容性表 349
  - 描述 71
  - 装入数据 193
- 分区
  - 部分 4, 71
  - 兼容性 9, 349
  - 具有单处理器 72
  - 具有多处理器 72
  - 删除 134
  - 数据
    - 拆离 183
    - 拆离, 转出数据 173
    - 创建 153
    - 定义范围 153
    - 连接, 转入数据 173
    - 描述 10, 12
    - 删除 182
    - 添加 173, 181
    - 已拆离 174, 180
    - 已连接 174
    - 正在连接 183
    - 转出, 拆离 184
    - 转入, 连接 174, 184
  - 添加 122
- 分区表
  - 避免不匹配 176
  - 表分区 10
  - 拆离 178
  - 重组 213
  - 创建 153, 156
  - 改变 173, 174, 181, 182



- 分区表 (续)
  - 描述 10
  - 迁移 176
  - 使用大对象 (LOB) 152
  - 使用具体化查询表 158
  - 使用 MDC (多维集群) 表 26, 61, 241
  - 数据旋转 183
  - 锁定 266
  - 索引行为 253
  - 限制 10, 174
  - 移动数据 184
  - 优化 244
  - 正在连接 174
  - 转出数据 184
  - 转换 176
  - 转入数据 184
  - 装入 156
  - 装入到 23, 190
- 分区间并行性
  - 与分区内并行性配合使用 67
  - db2expln 工具
    - 输出示例 280
    - 输出样本 278
- 分区间的查询并行性
  - 启用 120
- 分区键
  - 表
    - 描述 21
    - 准则 21
- 分区内并行性
  - 启用 121
  - 优化策略 273
  - 与分区间并行性配合使用 67
  - db2expln 工具
    - 输出示例 280
    - 输出样本 276
- 分区数据库
  - 复制的具体化查询表 290
  - 连接策略 283
  - 连接方法 285
  - 描述 71
  - 事件监视 221
  - 数据库重建 228
  - 数据再分发, 错误恢复 311
  - 数据再分发, 事件日志记录 311
  - 添加节点时出错 133
  - 再分发数据 309
  - 装入限制 195
- 分区数据库服务器
  - 安装 105
  - 验证安装
    - Windows 105
- 分区数据库环境
  - 重复的机器条目, 删除 117
  - 服务器安装, 在 UNIX 上验证 105
  - 监视装入操作 200

- 分区数据库环境 (续)
  - 全局快照 220
  - 设置 101
  - 事务故障恢复 226
  - 指定机器列表 117
  - 装入数据 194, 210
  - 自调整内存 317, 318
- 分区数据库系统上的全局快照 220
- 分区映射
  - 为数据库分区组创建 356
- 复制具体化查询表 25

## [ G ]

- 改变
  - 数据库分区组 171
- 更改数据库分区服务器配置命令 345
- 更新
  - 节点配置文件 118
  - 信息中心 377
  - DB2 信息中心 377
  - db2nodes.cfg (UNIX) 118
- 故障事务 225
- 故障诊断
  - 教程 379
  - 联机信息 379
  - 描述 237
- 关于本书 vii
- 管理通知日志 225
- 过程
  - STEPWISE\_REDISTRIBUTE\_DBPG 314, 361

## [ H ]

- 函数
  - 标量
    - DBPARTITIONNUM 351
    - NODENUMBER (请参阅 DBPARTITIONNUM) 351
  - 表函数
    - DB\_PARTITIONS 360
- 环境变量
  - rah 146
  - RAHDOTFILES 147
- 恢复
  - 崩溃 225
  - 从数据库分区服务器的故障 228
  - 跨节点示例 229
  - 两阶段落实协议 226
- 获取行分布号 API 336

## [ J ]

- 机器列表
  - 分区数据库环境 117

- 集群
  - 定义 273
  - 数据 35
- 集群索引
  - 使用分区表 256
  - 使用分区表的益处 256
- 兼容性
  - 分区 9
- 监视
  - 数据分区 213
  - rah 进程 149
- 键
  - 表分区 21
  - 分布 8
- 将数据库分区服务器添加至实例命令 346
- 教程
  - 故障诊断和问题确定 379
  - Visual Explain 379
- 节点 71
  - 连接耗用时间 323
  - 最大时差 325
  - FCM 守护程序 (UNIX) 89, 123
- 节点间的最大时差配置参数 325
- 节点连接重试次数配置参数 325
- 节点配置文件
  - 创建 110
  - 更新 (UNIX) 118
  - 描述 111
- 节点同步 235
- 具体化查询表 (MQT)
  - 复制 25
  - 复制, 在分区数据库中 290
  - 行为 158
  - 使用分区表 158

## [ K ]

- 可伸缩性 72
- 可用的端口范围
  - UNIX 124
- 可用空间控制记录 (FSCR)
  - 在 MDC 表中 271
- 控制中心
  - 管理数据库分区 171
- 控制 rah 命令 146
- 跨节点数据库恢复示例 229
- 快速通信管理器 (fast communication manager, FCM)
  - 端口号 124
  - 服务条目语法 123
  - 概述 84, 123
  - 启用数据库分区服务器之间的通信 124
  - 消息缓冲区 84, 123
  - Windows 84, 123
- 快照监视
  - 解释数据分区的输出 213
  - 在分区数据库系统上 220

- 快照监视 (续)
  - 在数据分区上 213

## [ L ]

- 连接
  - 并置的 285
  - 定义 282
  - 分区数据库中的表队列策略 283
  - 广播内部表 285
  - 广播外部表 285
  - 耗用时间 323
  - 类型
    - 定向内部表 285
    - 定向外部表 285
    - 在分区数据库中 285
- 连接耗用时间配置参数 323
- 连接集中器
  - 在分区数据库中使用代理程序 269
- 连接数据分区
  - 描述 174
- 两阶段落实
  - 协议 226
- 列表表达式
  - 多维表 44, 165
- 逻辑节点
  - 数据库分区服务器 117, 119
- 逻辑数据库分区 72

## [ M ]

- 命令
  - 以并行方式运行 143
  - db2adutl
    - 跨节点恢复示例 229
  - db2nchg 82, 345
  - db2ncrt 346
  - db2ndrop 348
  - GET STMM TUNING DBPARTITIONNUM 359
  - REDISTRIBUTE DATABASE PARTITION GROUP 339
  - UPDATE STMM TUNING DBPARTITIONNUM 360
- 目录表
  - 存储在数据库目录节点上 3, 109
- 目录节点 3, 109
- 目录视图
  - BUFFERPOOLDBPARTITIONS 369
  - DATAPARTITIONEXPRESSION 369
  - DATAPARTITIONS 369
  - DBPARTITIONGROUPDEF 370
  - DBPARTITIONGROUPS 371
  - PARTITIONMAPS 371
- 目录统计信息
  - 索引集群比率 271

## [ N ]

内存调整器

分区数据库环境 318

## [ P ]

配置

多分区 72

配置参数

分区数据库 3, 109

conn\_elapse 323

fcm\_num\_buffers 84, 123, 323

fcm\_num\_channels 324

intra\_parallel 327

logarchopt1

跨节点恢复示例 229

max\_connretries 325

max\_querydegree 327

max\_time\_diff 325

start\_stop\_time 326

vendoropt

跨节点恢复示例 229

片段消除

请参阅数据分区消除 244

## [ Q ]

启动与停止超时配置参数 326

启用

分区内并行性配置参数 327

数据库分区功能 (UNIX) 124

迁移

分区数据库环境 211

前缀

序列 144

全球标准时间 325

## [ R ]

认证

分区数据库注意事项 5

日志文件空间

数据再分发需要 311

容量

对于每个环境 72

容器

添加至 SMS 表空间 137

## [ S ]

散列分区 4

删除

使用启动板的数据库分区 138

删除重复的机器条目 117

删除节点验证 API 335

设计

数据库分区组 6

设计顾问程序

迁移至分区数据库 259

设置

rah 的缺省环境概要文件 150

声明 381

时间

节点之间的差异, 最大值 325

实例

分区服务器

更改 136

删除 138

列示数据库分区服务器 134

添加分区服务器 135

实用程序并行性 67

事件监视器

创建

分区数据库 221

事务

故障恢复

降低故障的影响 225

在活动的数据库分区服务器上 226

在有故障的数据库分区服务器上 226

数据

表分布 12

分布 71

分区 4

更改分布 171

组织方案

比较 17

描述 17

组合 12

数据分区

拆离 183

转出 173, 178, 184

创建 153

定义范围 153

描述 10, 12

请参阅表分区 10

删除 182

属性 180

添加 181

消除 244

已拆离 174

已连接 174

正在连接 183

转入 173, 174, 184

指定 153

数据服务器

扩充容量 122

数据库

分区环境 101

改变数据库分区组 171

更改数据的分布 171

## 数据库 (续)

- 配置, 跨多个分区 321
- 启用数据分区 3, 109
- 在所有数据库分区中创建 3, 109

## 数据库重建

- 分区数据库 228

## 数据库分区

- 编目 3, 109
- 创建数据库, 在所有 3, 109
- 从实例删除 138
- 更改 136
- 更改数据库配置 171
- 管理 128
  - 从控制中心 171
- 描述 71
- 使用向导添加 136
- 数据库 12, 71
- 数据库分区号 110
- 添加 129
  - 至已停止的系统 132
  - 至正在运行的系统 130
  - 至 NT 系统 131
- 同步 235
- 在数据库分区组中更改 171

## 数据库分区服务器

- 安装
  - 响应文件 104
- 发出命令 141, 237
- 故障恢复 228
- 描述 119
- 启用通信 (UNIX) 124
- 删除 138
- 指定 117

## 数据库分区功能 71

## 数据库分区功能 (DPF)

- 启用通信 (UNIX) 124

## 数据库分区组

- 表注意事项 151
- 并置 6
- 初始定义 127
- 创建 127, 356
- 创建分布图 356
- 对查询优化的影响 283
- 改变 171
- 描述 5
- 确定数据位置 7
- 删除分区 353
- 设计 6
- 添加分区 353
- IBMDEFAULTGROUP 缺省表 151

## 数据库管理器

- 启动超时 326
- 停止超时 326

## 数据库配置

- 参数
  - 自动重新启动 225

## 数据库配置 (续)

- 在分区间更改 171

## 数据类型

- 分区兼容性 349

## 数据偏差

- 在数据库分区组中再分发数据 339

## 数据再分发 310

- 错误恢复 311
- 进程描述 307
- 确定需要 309
- 日志空间需求 311
- 事件日志记录 311
- 指示信息 309
- 准则 307

## 数据组织

- 表 12

## 锁定

- 分区表上的行为 266
- 块索引扫描方式 264
- 离散 31, 33

## 锁定方式

- MDC (多维集群) 表
  - 表和 RID 索引扫描 261

## 索引

- 对分区数据库环境中的表列 303
- 分区表上的行为 253
- 管理
  - 对于 MDC 表 271
- 集群 256
- 集群比率 271
- 块
  - 扫描锁定方式 264
- 使用分区表 253
- 索引的比较
  - 基于集群和块 35

## [ T ]

## 添加节点 API 331

## 条款和条件

- 出版物的使用 379

## 通信

- 地址 89, 123
- 快速通信管理器 89, 123
- 连接耗用时间 323

## 同步

- 恢复注意事项 235
- 节点 235
- 数据库分区 235

## 突出显示约定 x

## [ W ]

## 网络文件系统 (NFS)

- 验证操作 93

- 维
  - 多维表 37
- 文档
  - 使用条款和条件 379
  - PDF 或印刷的 373
- 文档概述 373
- 文件集
  - 描述 89, 123
  - db2fcmr 守护程序 89, 123
  - db2fcms 守护程序 89, 123
- 文件系统
  - 为分区 DB2 服务器创建
    - Linux 95
- 问题确定
  - 教程 379
  - 联机信息 379

## [ X ]

- 响应文件
  - 安装
    - 数据库分区服务器 104
- 响应文件安装 103
- 向导
  - 添加数据库分区向导 136
- 消息缓冲区
  - 快速通信管理器 (FCM) 84, 123
- 协调程序分区 71
- 信息中心
  - 版本 376
  - 更新 377
  - 以各种语言查看 377
- 性能
  - 目录信息, 减少争用 3, 109
- 许可证 71
- 选择
  - 多维表维 37

## [ Y ]

- 验证
  - 端口范围
    - 在 UNIX 上可用 94
- 一致点
  - 数据库 225
- 移动数据
  - 到多维表中 44, 165
- 已拆离数据分区 180
  - 描述 178
- 印刷版书籍
  - 订购 375
- 硬件
  - 环境 72
    - 并行性类型 72
    - 单一分区, 单处理器 72

- 硬件 (续)
  - 环境 (续)
    - 单一分区, 多处理器 72
    - 具有单处理器的分区 72
    - 具有多处理器的分区 72
    - 逻辑数据库分区 72
- 用户
  - 为分区 DB2 服务器创建
    - AIX 99
    - Linux 98
- 优化
  - 分区表 244
  - 分区内并行性 273
  - 连接
    - 定义 282
    - 分区数据库 285
    - MDC 表的策略 249
  - 有故障的数据库分区服务器 226

## [ Z ]

- 再分发 311
- 再分发数据 310
  - 必要性 309
  - 过程 361
  - 跨数据库分区 171
  - 使用按步骤的再分发过程 314
- 在节点 API 创建数据库 332
- 在节点 API 删除数据库 334
- 终止
  - 装入操作
    - 在多分区数据库中 201
- 注册表变量
  - DB2CHGPWD\_ESE 322
  - DB2PORTRANGE 322
  - DB2\_NO\_MPFA\_FOR\_NEW\_DB 44, 165
  - DB2\_NUM\_FAILOVER\_NODES 322
  - DB2\_PARTITIONEDLOAD\_\_DEFAULT 322
- 专用寄存器
  - CURRENT DBPARTITIONNUM 350
  - CURRENT NODE (请参阅 CURRENT DBPARTITIONNUM) 350
- 装入
  - 多维集群 (MDC) 表 190
  - 分区数据库环境 203
  - 配置选项 203
  - 示例 207
    - 在分区数据库环境中 207
  - 数据库分区 194
  - 装入到分区表 23, 190
  - 装入到数据库分区 193
- 准则
  - 范围集群表 31, 163
- 自调整内存
  - 分区数据库环境 317, 318

自调整内存 (续)  
  启用  
    非统一环境 318  
自动重新启动 225  
组织数据  
  方法 12  
最大查询并行度配置参数 327

## A

ADMIN\_CMD 过程  
  受支持的命令  
    GET STMM TUNING DBPARTITIONNUM 359  
    UPDATE STMM TUNING DBPARTITIONNUM 360  
AIX  
  安装  
    DB2 服务器 84  
  创建  
    必需的用户, 用于分区 DB2 服务器安装 99  
    DB2 主文件系统, 用于分区数据库系统 96  
  更新环境设置 91  
  将命令分发给 ESE 工作站 93  
  验证  
    NFS 正在运行 93  
ALTER DATABASE PARTITION GROUP 语句 353  
ALTER NODEGROUP 语句  
  请参阅 ALTER DATABASE PARTITION GROUP 353  
API  
  sqleaddn 331  
  sqlecran 332  
  sqledpan 334  
  sqledrpn 335  
  sqlugrpn 336

## B

backup 实用程序  
  限制 367  
BACKUP DATABASE 命令 367

## C

conn\_elapse 配置参数 323  
CREATE DATABASE PARTITION GROUP 语句 356  
CREATE NODEGROUP 语句  
  CREATE DATABASE PARTITION GROUP 语句 356  
CREATE TABLE 语句  
  OVERFLOW 子句 31, 32  
CURRENT DBPARTITIONNUM 专用寄存器 350

## D

DATAPARTITIONNUM 标量函数  
  值和参数 351

DB2  
  安装  
    UNIX 85  
    Windows 85  
  安装方法 85  
  分区安装  
    更新 AIX 环境设置 91  
DB2 服务器  
  分区  
    准备 Windows 环境 82  
  在 Linux 或 UNIX 上安装 84  
  在 Windows 上安装 81  
DB2 企业服务器版  
  安装  
    Windows 82  
DB2 信息中心  
  版本 376  
  更新 377  
  以各种语言查看 377  
DB2 主文件系统  
  分区数据库系统  
    AIX 96  
DB2 Connect  
  安装  
    UNIX 85  
    Windows 85  
  安装方法 85  
db2adutl 命令  
  跨节点恢复示例 229  
db2Backup API 367  
DB2CHGPWD\_EEE 322  
db2empfa 实用程序 44, 165  
db2exfmt 工具  
  输出示例  
  描述 275  
db2expln 工具  
  输出示例  
    具有分区间并行性的多分区方案 278  
    具有分区内并行性的单分区方案 276  
    具有完全并行性的多分区方案 280  
  描述 275  
  无并行性 275  
db2fcmr 守护程序  
  描述 89, 123  
db2fcms 守护程序  
  描述 89, 123  
db2nchg 命令 136, 345  
db2ncrt 命令 135, 346  
db2ndrop 命令 138, 348  
db2nlist 命令 134  
DB2NODE  
  添加服务器时已导出 130, 131, 132  
db2nodes.cfg 文件 110, 111, 118  
  指定网络名字段 82  
  ALTER DATABASE PARTITION GROUP 353  
  CREATE DATABASE PARTITION GROUP 356

db2nodes.cfg 文件 (续)  
    DBPARTITIONNUM 函数 351  
DB2PORTRANGE 322  
db2start ADDNODE 135  
db2\_all 命令 141, 142, 143, 237  
    概述 141  
db2\_call\_stack 143  
db2\_kill 143  
DB2\_NO\_MPFA\_FOR\_NEW\_DB 44, 165  
DB2\_NUM\_FAILOVER\_NODES 322  
DB2\_PARTITIONEDLOAD\_\_DEFAULT 322  
DBPARTITIONNUM 函数  
    描述 351  
    值和参数 351  
DB\_PARTITIONS 表函数 360  
DPF 71

## E

ESE 71  
ESE 工作站 93  
ESE, 企业服务器版 71

## F

FCM (快速通信管理器)  
    服务条目语法 123  
    快速通信管理器 84, 123  
    通道 324  
fcm\_num\_buffers 配置参数 84, 123, 323  
fcm\_num\_channel  
    配置参数 324

## G

GET STMM TUNING DBPARTITIONNUM 命令  
    使用 ADMIN\_CMD 359

## H

HP-UX  
    安装  
        DB2 服务器 84  
    验证  
        NFS 正在运行 93

## I

IBMCATGROUP 数据库分区组 127  
IBMDEFAULTGROUP 数据库分区组 127  
IBMTEMPGROUP 数据库分区组 127  
intra\_parallel 配置参数 327  
I/O  
    并行性 67

## L

Linux  
    安装  
        DB2 服务器 84  
        DB2 服务器, 使用“DB2 安装”向导 87  
    创建  
        分区 DB2 服务器的必需用户 98  
        分区 DB2 服务器的文件系统 95  
    验证  
        NFS 正在运行 93  
LOAD 命令  
    在分区数据库环境中 195, 210  
load 实用程序  
    并行性 189  
LOAD QUERY 命令  
    在分区数据库环境中 200  
logarchopt1 配置参数  
    跨节点恢复示例 229

## M

max\_connretries 配置参数 325  
max\_querydegree 配置参数 327  
max\_time\_diff 配置参数 325  
MDC 51  
MDC (多维集群) 表 44, 165  
    表和索引的管理 271  
    描述 35  
    使用分区表 26, 61, 241  
    锁定方式  
        表和 RID 索引扫描 261  
    选择维 37  
    优化策略 249  
    转出删除 249  
MDC 表  
    更新 61  
    块索引 49  
    块索引和查询性能 54  
    块索引注意事项 48  
    块映射 59  
    删除, 从 61  
    装入注意事项 48  
    自动维护集群 57  
MPP 环境 72  
MQT (具体化查询表)  
    复制 25  
    复制, 在分区数据库中 290

## N

NFS (网络文件系统)  
    验证操作 93  
nodegroups (数据库分区组)  
    创建 127

NODENUMBER 函数  
DBPARTITIONNUM 351

## R

rah 命令  
递归调用的 143  
概述 141  
环境变量 146  
监视进程 149  
简介 141, 237  
控制 146  
描述 143  
前缀顺序 144  
确定问题 148  
设置缺省环境概要文件 150  
以并行方式运行命令 143  
指定  
数据库分区服务器列表 117  
作为参数或响应 142  
RAHCHECKBUF 环境变量 143  
RAHDOTFILES 环境变量 147  
RAHOSTFILE 环境变量 117  
RAHOSTLIST 环境变量 117  
RAHWAITTIME 环境变量 149  
RAHCHECKBUF 环境变量 143  
RAHDOTFILES 环境变量 147  
RAHOSTFILE 环境变量 117  
RAHOSTLIST 环境变量 117  
RAHTREETHRESH 环境变量 143  
RAHWAITTIME 环境变量 149  
Redistribute 实用程序  
限制 308  
REDISTRIBUTE DATABASE PARTITION GROUP 命令 339  
RESTART DATABASE 命令 225

## S

SIGTTIN 消息 142  
SMP 集群环境 72  
SMS (系统管理的空间)  
表空间  
添加容器 137  
Solaris  
安装  
DB2 服务器 84  
验证  
NFS 正在运行 93  
SQL 语句  
显示帮助 376  
ALTER DATABASE PARTITION GROUP 353  
ALTER NODEGROUP (请参阅 ALTER DATABASE PARTITION GROUP) 353  
CREATE DATABASE PARTITION GROUP 356

SQL 语句 (续)  
CREATE NODEGROUP (请参阅 CREATE DATABASE PARTITION GROUP) 356  
sqleaddn API 331  
sqlecran API 332  
sqledpan API 334  
sqledrpn API 335  
sqlugrpn API 336  
start\_stop\_time 配置参数 326  
STEPWISE\_REDISTRIBUTE\_DBPG 过程 361  
用于再分发数据 314

## U

UNIX  
安装  
DB2 服务器, 使用“DB2 安装”向导 87  
更新节点配置文件 118  
缺省端口范围 124  
验证分区数据库服务器安装 105  
UPDATE STMM TUNING DBPARTITIONNUM 命令  
使用 ADMIN\_CMD 360

## V

vendoropt 配置参数  
跨节点恢复示例 229  
Visual Explain  
教程 379

## W

Windows 操作系统  
安装  
使用“DB2 安装”向导的 DB2 服务器 81  
快速通信管理器 (FCM) 84, 123  
添加数据库分区 131  
验证  
分区数据库服务器安装 105

## [ 特别字符 ]

\$RAHBUFDIR 143  
\$RAHBUFNAME 143  
\$RAHENV 146  
“删除数据库分区”启动板 138  
“DB2 安装”向导  
安装 DB2 服务器  
Linux 87  
UNIX 87







中国印刷

S151-0615-00



Spine information:

**DB2 版本 9.5 Linux 版、UNIX 版和 Windows 版**

**分区和集群指南**

