

INGRES 数据库管理系统

资 料 汇 编

第 二 卷

1. INGRES的存贮结构和存贮方法
2. 一个数据库系统的回顾
3. 分解——一种查询处理策略
4. EQUOL/C用户指南
5. INGRES VIFRED用户指南
6. INGRES RBF用户指南
7. INGRES QBF用户指南
8. INGRES 的设计和实现

中国科学院新技术开发局培训中心
中国科学院计算技术研究所十七室

一 九 八 四 年 十 二 月

INGRES数据库管理系统

資料匯編

第二卷

1. INGRES的存贮结构和存贮方法
2. 一个数据库系统的回顾
3. 分解——一种查询处理策略
4. EQUUEL/C用户指南
5. INGRES VIFRED用户指南
6. INGRES RBF用户指南
7. INGRES QBF用户指南
8. INGRES 的设计和实现

中国科学院新技术开发局培训中心
中国科学院计算机技术研究所十七室

一九八四年十二月

前 言

INGRES (INteractive Gaphics and REtrieval System) 是一个著名的关系数据库管理系统。它最早由美国加州大学的克利分校电子研究实验室于1973年开始研制, 1975年投入运行。经过几年的改造, 形成了一系列版本。以后由美国 Relational Technology Inc. 改制成为商品推向市场。

早期的INGRES是建立在PDP-11系列机的UNIX系统基础上的, 后来在VAX-11系列机的VMS和UNIX系统上也可以运行。目前它又移植到了16位微机上, 在UNIX系统及其变种支持下运行, 在我国国内引进的Codata 68000和Dual 68000上都有微型化的INGRES。INGRES几乎全部用C语言描述。

INGRES具有关系数据库管理系统所必须具备的基本思想。比如, 数据独立性, 多用户共享, 数据安全性、完整性的管理以及恢复功能等等。它功能齐全, 用户可以通过终端监控命令在一种交互环境中有效地描述、创建、修改、检索以及删除数据。系统也提供了一种非过程化的查询语言QUEL, 对数据库中的数据进行查询和更新。在QUEL语句内允许嵌入多种运算, 包括算术、逻辑和关系比较以及它特有的计数、求平均值、最大最小值等聚合运算。QUEL嵌入到C语言中就称为EQUEL, 嵌入到C中的EQUEL语句经过预处理, 再统一由C编译程序处理成为可执行代码。这样, 就为用户非交互地使用INGRES功能提供了方便。

INGRES是最早实现的为数不多的关系系统之一。它除了具有一般关系数据库管理系统的思想外, 在实现时还采用了很多新颖而有效的技术。因此, 不论在研究领域还是在许多使用范围内, 它都具有相当深刻的影响。许多数据库管理系统的专著和论文中都引用了这个系统, 而在以后实现的许多关系数据库管理系统也借鉴了它的经验。随着国内16位微机的大量引进和生产, 许多用户都迫切希望掌握数据库管理系统的基本思想, 利用所提供的数据库管理系统来开展计算机管理和开发工作。

我们觉得, 无论从深入掌握一个具有代表性的关系数据库管理系统的设计思想和实现技术以充实教学内容, 以及进一步开展关系数据库管理系统的研究和开展工作, 还是从推广数据库管理系统的应用来说, 系统地解剖INGRES的实现版本都是具有重要价值的。为此, 从1983年起, 我们三个单位的有关同志先后对第6、第7两个版本的INGRES的全部代码进行了系统分板, 同时组织了一些研究生和大学毕业生一起收集有关INGRES方面的许多专著和论文, 进行了深入讨论和报告, 在此基础上, 我们翻译了有关资料。

目前, INGRES的解剖工作已告一段落, 我们将把这一段时间工作情况公诸于众。同时, 翻译整理了INGRES的外部资料。在中国科学院计算技术研究所应用和人才开发研究室的大力协助下, 我们将分卷出版所有的资料, 包括INGRES代码的详细流程图。第一卷、第二卷刊登目前我们收集到的所有INGRES的外部材料, 包括INGRES概述, 演示、参考手册, 自学指南、报表编写程序参考手册、EQUEL/C用户指南、QBF用户指南、VIFRED用户指南、RBF用户指南, 以及若干具有代表性的有关INGRES的文章。这两

卷每卷约35万字左右。后面几卷是有关INGRES的内部材料，主要是我们分析工作的总结。

由于INGRES是一个较大型的数据库管理系统，代码总量约有8万行C语句，加上我们水平有限，所以对它的认识和分析可能有不当之处，在翻译和撰写的文章中也会有错误，肯切地希望广大用户和有关专家对此提出批评。

如果这套资料的出版对国内INGRES数据库系统的推广应用以及数据库管理系统本身的研究、开发和教学工作起某种推动作用，那么我们将会感到十分欣慰。

最后对杨芙清、仲萃豪、萨师煊诸位老师及有关同志的指导、帮助以及计算所应用和人才开发研究室的大力协助表示衷心的感谢。

译编者

1984年7月

总 目 录

1. 关系数据库管理系统INGRES的存贮结构和存贮方法	(1)
1. 引言	(3)
2. QUEL	(3)
3. 分解	(6)
4. 存取方法接口	(6)
5. 存贮结构的选择	(8)
6. 随机函数	(9)
7. 保护计算函数	(9)
8. 目录结构	(10)
9. 广义目录	(10)
10. 静态目录与动态目录的比较	(12)
11. 压缩存贮结构	(13)
12. 总结	(13)
2. 一个数据库系统的回顾	(17)
1. 引言	(19)
2. 历史	(19)
2.1 早期.....	(19)
2.2 第一个实现.....	(20)
2.3 使系统投入实用.....	(22)
3. 经验教训.....	(23)
3.1 目标.....	(23)
3.2 结构化设计.....	(23)
3.3 编码规定.....	(24)
3.4 用户支持.....	(24)
4. 失误	(25)
4.1 命令的解释执行.....	(25)
4.2 有效性检查.....	(25)
4.3 进程问题.....	(26)
4.4 存取方法.....	(26)
4.5 静态索引.....	(27)
4.6 分解.....	(27)
4.7 保护.....	(27)

4.8 律师.....	(27)
5. 评论.....	(28)
5.1 UNIX.....	(28)
5.2 PDP-11	(28)
5.3 数据模型.....	(28)
6. INGRES项目的计划.....	(29)
6.1 分布式INGRES.....	(29)
6.2 分布式数据库机器.....	(29)
6.3 新的数据库程序设计语言.....	(29)
6.4 数据输入机制.....	(29)
6.5 改进的完整性控制.....	(29)
3. 分解——一种查询处理策略	(33)
1. 引言	(35)
2. QUEL.....	(35)
3. 分解	(36)
4. 算法的描述	(42)
4.1 归约算法.....	(42)
4.1 子查询序列.....	(44)
4.1 元组替换.....	(44)
5. 变元选择.....	(44)
6. 总结.....	(47)
附录A 系统组织	(48)
附录B 数据结构	(48)
4. EQUOL/C用户指南	(53)
前 言.....	(55)
0.1 手册的目的.....	(55)
0.2 针对的读者对象.....	(55)
0.3 这本文献资料的组织.....	(55)
0.4 关联的文献资料.....	(55)
0.5 约 定.....	(55)
第一章 绪 论.....	(56)
1.1 EQUOL是什么?	(56)
1.2 EQUOL的设计特点.....	(56)
第二章 EQUOL语言.....	(57)
2.1 语法规定.....	(57)
2.2 EQUOL说明.....	(57)
2.3 存取一个INGRES数据库	(58)

2.4	嵌入的QUEL语句	(59)
2.5	变量替换	(62)
2.6	使用sort (分类) 子句	(65)
2.7	参数化的QUEL语句	(66)
2.8	性能所需考虑的问题	(68)
第三章	运行时系统 (Run-time system)	(69)
3.1	类型转换	(69)
3.2	运行时错误处理	(70)
3.3	中断处理	(71)
3.4	保留符号	(71)
第四章	嵌入表格结构	(72)
4.1	嵌入表格结构概述	(72)
4.2	起动表格	(72)
4.3	初始化一个表格	(73)
4.4	显示一个表格	(74)
4.5	检索关于表格的信息	(84)
4.6	杂项命令	(87)
第五章	予处理程序操作	(89)
5.1	命令行	(89)
5.2	输入/输出约定	(89)
5.3	包含文件处理	(89)
5.4	编译和连接一个EQUEL程序	(90)
5.5	装配和连接予编译的格式	(91)
附录A	QUEL的语法小结	(91)
附录B	使用QUERY方式的程序例子	(92)
附录C	使用FORMDATA的程序例子	(96)
附录D	使用多种表格的程序例子	(104)

5. INGRES VIFRED直观表格编辑程序用户指南 (113)

前 言	(115)
第一章 引言	(116)
1.1 表格与计算机系统	(116)
1.2 VIFRED的性质	(116)
1.3 VIFRED简介	(117)
第二章 表格	(117)
2.1 表格和它们的组成部分	(118)
2.2 表格成分	(119)
2.3 数据格式	(120)
2.4 属性	(122)

第三章 运行VIFRED	(124)
3.1 定义终端	(124)
3.2 VIFRED命令	(124)
3.3 引用省缺表格来编辑	(125)
第四章 表格编辑	(126)
4.1 表格中的移动	(126)
4.2 移动成分	(127)
4.3 插入空行	(130)
4.4 加入新的点缀成分	(131)
4.5 编辑已有的成分	(131)
4.6 行和成分的删除	(135)
4.7 输入新的字段	(136)
4.8 UNDO命令	(137)
4.9 HELP命令	(138)
第五章 保存编辑过的表格	(139)
5.1 写出表格定义	(139)
5.2 继续	(141)
5.3 退出VIFRED	(141)
5.4 再启动	(142)
第六章 高级编辑	(143)
6.1 其他编辑任选项	(143)
6.2 为多个表创建一个表格	(143)
6.3 编辑一个现存表格	(144)
6.4 创建一个高表格	(144)
6.5 维护表格目录	(144)
6.6 对表格编译	(145)
6.7 印出你的表格	(147)
第七章 用VIFRED (INSERT方式) 编辑正文	(148)
7.1 INSERT方式编辑	(148)
7.2 在错误正文之上按键	(148)
7.3 在新正文中改正错误	(149)
7.4 移到一个错误处	(149)
7.5 扞入方式命令要览	(149)
第八章 用VIFRED (编辑方式) 辑编正文	(151)
8.1 EDIT方式编辑	(151)
8.2 进入EDIT方式	(151)
8.3 来自INSERT方式的命令	(152)
8.4 命令重复	(152)
8.5 在正文内移动	(152)

8.6	编辑命令	(154)
8.7	EDIT方式命令要览	(157)
附录A	VIFRED命令总结	(158)
A.1	启动菜单	(158)
A.2	光标菜单	(158)
A.3	移动菜单点缀	(158)
A.4	移动菜单字段	(156)
A.5	移动菜单	(159)
A.6	字段编辑菜单	(159)
A.7	字段输入菜单	(159)
A.8	写菜单	(160)
A.9	写任选菜单	(160)
A.10	目录菜单	(161)
A.11	INSERT方式编辑程序	(161)
A.12	EDIT方式编辑程序	(161)
A.13	VIFRED光标移动命令	(162)

6. INGRES RBF (按表格产生的报表) 用户指南 (163)

前 言	(165)
0.1 本资料的用途	(165)
0.2 读者	(165)
0.3 概述	(165)
0.4 其它资料	(166)
第一章 引论	(166)
1.1 计算机生成的报表	(166)
1.2 INGRES中的报表	(168)
第二章 表格、缺省报表及RBF	(171)
2.1 表格	(171)
2.2 缺省报表	(172)
第三章 选择一个报表的内容	(175)
3.1 选择数据	(175)
3.2 定义视图	(175)
第四章 启动RBF	(177)
4.1 启动RBF	(177)
4.2 RBF主菜单	(178)
4.3 主RBF编辑菜单	(180)
第五章 编辑一个报表的布局	(181)
5.1 RBF布局的表格	(181)
5.2 用RBF编辑	(182)

第六章 编辑数据的显示	(186)
6.1 显示格式	(186)
6.2 编辑显示的格式	(189)
6.3 数字的样板	(190)
第七章 报表任选项	(192)
7.1 进一步的报表任选项	(192)
7.2 页长度和走纸	(192)
7.3 下划线的特征	(192)
第八章 编辑报表的结构	(194)
8.1 报表结构表格	(194)
8.2 报表中的排序特征	(195)
8.3 间断列	(195)
第九章 选择列的任选项	(196)
9.1 列任选表格	(196)
9.2 对所有列的任选项	(197)
9.3 对间断列的任选项	(198)
第十章 保存报表规格说明	(199)
10.1 保存 报表	(199)
10.2 Write (写) 命令	(199)
第十一章 输出报表	(200)
11.1 输出选择	(200)
11.2 交互的终端输出	(201)
第十二章 rbf和report命令	(202)
12.1 rbf命令	(202)
12.2 report 命令	(204)
附录A RBF命令摘要	(206)
A.1 RBF主菜单	(206)
A.2 RBF编辑主菜单	(206)
A.3 布局菜单	(206)
A.4 移动菜单	(206)
A.5 列编辑菜单	(207)
A.6 报表结构菜单	(207)
A.7 保存报表菜单	(207)
A.8 报表目录菜单	(207)
附录B QUEL检索命令摘要	(208)
附录C 适于RBF的终端名	(208)

7. INGRES QBF (用表格的询问) 用户指南 (213)

前 言	(215)
-----	-------

0.1	这份文档的用途	(215)
0.2	读者	(215)
0.3	概貌	(215)
0.4	其它文档	(216)
第一章	引言	(216)
第二章	运行QUERY BY FORMS	(217)
2.1	定义终端	(217)
2.2	使用Q B F	(217)
2.3	表格	(218)
第三章	用QBF录入数据 (INSERT方式)	(220)
3.1	在一个字段中填充数据	(220)
3.2	通过QBF作错误检查	(221)
3.3	纠正前面的字段	(221)
3.4	在一个字段中纠正错误	(221)
3.5	多行字符字段	(222)
3.6	录入最后的字段	(222)
3.7	在一个完整的表格中纠正错误	(222)
3.8	在一字段内移动	(223)
3.9	复制前面录入的数据	(224)
3.10	INSERT方式的命令提要	(224)
第四章	QBF中的编辑命令 (EDIT方式)	(225)
4.1	进入EDIT方式	(226)
4.2	从INSERT方式来的命令	(226)
4.3	命令的重复	(226)
4.4	在一字段内移动	(226)
4.5	编辑命令	(228)
4.6	EDIT方式命令的提要	(233)
第五章	Query By Forms功能	(234)
5.1	用命令行标志指定一个功能	(234)
5.2	用主菜单指定一个功能	(234)
5.3	TABLE命令	(235)
第六章	APPEND功能	(235)
6.1	显示	(235)
6.2	将数据添加到一个表上	(236)
6.3	APPEND菜单	(236)
第七章	UPDATE功能	(237)
7.1	QUERY状态——指定行	(238)
7.2	GO状态——编辑行	(241)
第八章	RETRIEVE功能	(243)

8.1	QUERY状态——指定行.....	(243)
8.2	GO状态	(244)
第九章	HELP命令.....	(245)
第十章	对高级用户的提示	(246)
10.1	屏蔽冗余的消息	(246)
10.2	规定输入方式	(246)
10.3	加快询问	(246)
10.4	在数据库应用中使用QBF.....	(247)
第十一章	改变QBF表格	(247)
11.1	在QBF中使用“VIFRED表格”	(247)
11.2	改变表格的外表	(248)
11.3	改变字段属性	(248)
附录A	QBF命令提要	(249)
附录B	用于QBF的终端名	(250)
8.	INGRES 的设计和实现.....	(255)
1.	引言	(257)
2.	INGRES进程结构.....	(264)
3.	数据结构和存取方法... ..	(267)
4.	进程2的结构.....	(272)
5.	进程3.....	(277)
6.	进程4中的实用程序.....	(282)
7.	结论和将来的发展	(284)

1. 关系数据库管理系统

INGRES的存贮结构和存贮方法

内 容 提 要

本文是关于 INGRES 中文件组织形式的专论。作者讨论了存储结构的一般性问题，进而比较了静态的索引顺序结构 (ISAM) 对动态的索引顺序结构 (B树)。最后给出了 INGRES 采用静态结构的理由。尽管本文的结论不一定正确 (这一点作者后来也承认了)，但其中的论述和分析仍有其参考价值，至少可以使我们对 INGRES 的下层文件组织有所了解。

本文由武志学译，潘升庆校。

关系数据库管理系统

INGRES的存贮结构和存取方法

Gerald Held和Michael stonebraker

INGRES是伯克利分院开发的一个关系数据库管理系统，它实现了一个高级非过程化查询语言（QUEL），本文描述用来实现关系的存贮结构的选择。此外，还描述了给所有存贮结构提供的单一关系化视图——存取方法界面。

1. 引言

INGRES (INTERactive Graphics and REtrieval System) 是一个关系数据库管理系统，它是在基于伯克利的硬件配置的PDP-11/40上实现的。INGRES作为一个用户作业在UNIX操作系统〔RITC73〕的上面运行，UNIX系统是贝尔电话实验室开发的，INGRES的实现是用“C”语言〔RITC74〕编码的，该语言是一个高级语言，UNIX本身就是用“C”来编写的。

为了实现高级的、非过程化的查询语言，最好给数据存取定义一个低级的、过程化的语言。已经实现了这样一种语言，该语言在支持为了有效的数据存取而提供的各种实际存取结构的同时，给高级软件提供了一个数据的关系视图。我们将描述这个存取方法界面语言，并且讨论它现在支持的五种存贮结构。讨论主要集中在选择这组存贮结构所作的考虑上。

为了引起讨论，我们首先给出一些查询语言（QUEL）所允许的对话的例子，然后指出用来把对话分解成存取方法调用的机制。

2. QUEL

QUEL (QUErg Language) 与 Data Language /ALPHA [Codd 71]，SQUARE [BOYC73]是SEQUENCE[CHAM74]有许多共同点，它是一个完全的〔Codd72〕查询语言，该语言使得程序员不需再关心数据结构如何实现以及在存贮数据上的操作算法是什么。这样就容易高度的数据独立性〔STON74b〕。由于QUEL的基本实体是关系，因此我们对关系进行定义并指出在本文的例题中使用的样本关系。

已知集合 $D_1, \dots, D_N$ (可以相同)，关系 $R(D_1, \dots, D_N)$ 是笛卡尔积 $D_1 * \dots * D_N$ 的一个子集。换句话说， R 是一个 N 元组 $X = (X_1, \dots, X_N)$ 的集合。其中， X_i 是 D_i 的元素， $i \in \{1, \dots, N\}$ ，集合 $D_1, \dots, D_N$ 叫作 R 的“域” (DOMAINS)， R 的“度” (DEGREE) 为 N 。QUEL中对关系的唯一限制是它们必须是规范化的。因此，每一个域必须是“简单的” (SIMPLE)。也就是说，它的元素不能再是关系。

显然， R 可以看成是一个把 R 的元素看成行，以及用域名来标记列的表，如下例所

示。

NAME	DEPT	SALARY	MANAGER	AGE
Smith	toy	10000	Jones	25
Jones	toy	15000	Johnson	32
Adams	candy	12000	Baker	36
EMPLOYEE				
Johnson	toy	14000	Harding	29
Baker	admin	20000	Harding	47
Harding	admin	40000	none	58

这就给出了具有域NAME, DEPT, SALARY, MANAGER和AGE的EMPLOYEE关系。每一个雇员有一个经理(除了Harding,他可能是公司的总经理)。工资, 年龄, 并且在一个部门工作。

R的表格式表示中的每一列可以看成是R到Di的函数映射。这些函数称作“属性”。属性不能单独地被标识, 而是由定义它的域来识别。

QUEL对话至少包含一个RANGE语句, 其形式为:

RANGE OF Variable-List Is relation-name

在范围语句中说明的符号是用来作为元组参数的变量。它们叫作“元组变量”。这个语句的目的用来说明每个变量所涉及的关系。

此外, 一个对话还包括一个或多个具有下述形式的语句:

Command Result-name (Target-List)

WHERE Rualification

其中, 命令(Command)可以是: RETRIEVE, APPEND, REPLACE, DELETE。对于RETRIEVE和APPEND来讲, Result-name是满足条件的元组将要检索到或添加到其中的关系的名字。对于REPLACE和DELETE来讲, Result-name是用来通过查询条件来标识要被更改或删除的元组的元组变量的名字。Target-List是具有如下形式的表:

Result-domain-Function, ...

其中, Result-domain是结果关系的域名, 其中存放相应函数的值。

下面提出有效的QUEL对话。该语言的完整描述见 [HELD 75a]

例2.1 找出雇员Jones的生日。

RANGE OF E IS EMPLOYEE

RETRIEVE INTO W (BDATE=1975-E.AGE)

WHERE E.NAME='JONES'

其中, E是涉及EMPLOYEE关系的元组变量, 在该关系中的所有满足限定条件E.NAME='JONES'的元组都将被找出来。查询的结果是一个新关系, W, 它只有一个属性, BDATE, 对于每一个满足条件的元组都要计算其生日。如果省略了结果关系, 那么将在用户终端上显示满足条件的元组。此外, 在Target-List中, 当Function的形式为Variable.Attribute时Result-domain='也可以被省略。(也就是说, NAME=E.NAME可以写为E.NAME——见例2.6)。

例2.2 将元组 (Jackson, candy, 13000, Baker, 30) 插入到 EMPLOYEE 中去。

```
APPEND TO EMPLOYEE (NAME = 'Jackson', DEPT = 'candy',  
    SALARY = 13000, MGR = 'Baker',  
    AGE = 30)
```

这样，结果关系EMPLOYEE通过加入给定的元组到该关系中而得到了修改。

例2.3 删除雇员Jackson的信息。

```
RANGE OF E IS EMPLOYEE  
DELETE E WHERE E.NAME = 'Jackson'
```

这样，对应于所有名字为Jackson的雇员的元组都从EMPLOYEE中删除了。

例2.4 给Jones增加10%的工资。

```
RANGE OF E IS EMPLOYEE  
REPLACE E (SALARY BY 1.1 * E.SALARY)  
WHERE E.NAME = 'Jones'
```

这样，在EMPLOYEE中的E.NAME = 'Jones'的那些元组的 E.SALARY 将由 $1.1 * E.SALARY$ 代替。（注意，关键字IS和BY在任何QUEL对话中都可以用 '=' 来替换。）

此外，QUEL还包含有聚合操作：COUNT, SUM, MAX, MIN, AVG和集合操作SET。下面是两个使用聚合操作的例题。

例2.5 把所有的玩具部的雇员的工资用玩具部的平均工资来代替。

```
RANGE OF E IS EMPLOYEE  
REPLACE E (SALARY BY AVG (E.SALARY  
    WHERE E.DEPT = 'toy')  
    WHERE E.DEPT = 'toy')
```

这样，AVG被看成了满足查询条件E.DEPT = 'toy'的那些元组的工资属性。注意，AVG (E.SALARY WHERE E.DEPT = 'toy') 是一个计算出来的数值。因此称作“聚合”。下面是一个更一般的聚合的例题。

例2.6 找出其平均工资大于公司的平均工资的部门，平均工资均是指其工资超过\$10000的雇员的工资的平均值。

```
RANGE OF E IS EMPLOYEE  
RETRIEVE INTO HIGHPAY (E.DEPT)  
WHERE AVG (E.SALARY BY E.DEPT)  
    WHERE E.SALARY > 10000)  
> AVG (E.SALARY WHERE E.SALARY > 10000)
```

其中，AVG (E.SALARY BY E.DEPT WHERE E.SALARY > 10000) 是一个“聚合函数”，并且是对每一个E.DEPT值取一个值。这个值就是AVG (E.SALARY WHERE E.SALARY > 10000 AND E.DEPT = Value)。对于那些该聚合函数的值超过聚合AVGCE.SALARY WHERE E.SALARY > 10000) 的部门。该语句