

TRS[®] DATABASE 全文数据库系统

使用指南 版本 6.80

北京拓尔思(TRS)信息技术股份有限公司 版权所有
Beijing TRS Information Technology Co., Ltd.



版权

本手册由北京拓尔思 (TRS) 信息技术股份有限公司 (以下简称 TRS 公司) 出版, 版权属 TRS 公司所有。未经出版者正式书面许可, 不得以任何方式复制本文档的部分或全部内容。

©北京拓尔思(TRS)信息技术股份有限公司 版权所有。保留所有权利。

TRS是北京拓尔思 (TRS) 信息技术股份有限公司的注册商标

前言

本手册分为六部分。第一部分介绍了 TRS 的主要特点、TRS 中各对象的定义和作用、TRS 的体系结构；第二部分介绍了在 TRS 数据库所支持的丰富的检索运算符以及检索表达式的构造规则；第三部分详细介绍了在 TRS 数据库加载前组织各种数据文件所必须遵循的格式规范；第四部分介绍了 TRS 服务器在安装前的准备工作、具体的安装步骤，以及安装后的工作；第五部分是系统附录；第六部分是错误信息。

TRS 公司感谢您使用 TRS 产品。如果您发现本手册中有错误或者产品运行不正确，或者您对本手册有任何意见和建议，请及时与 TRS 公司联系。您的意见将是我们做版本修订时的重要依据。

公司总部

北京市朝阳区大屯路科学园南里 风林西奥中心 B 座 16 层 100101

电话：010-64848899

传真：010-64879084

总部营销中心

北京市朝阳区大屯路科学园南里 风林西奥中心 A 座 17 层 100101

电话：010-64848899

传真：010-64889088

总部研发中心

北京市北四环中路 35 号 北京信息科技大学图书馆三层 100101

电话：010-64859900

网站与邮箱

主页：<http://www.trs.com.cn>

论坛：<http://forum.trs.com.cn>

上海：trs.sh@trs.com.cn

广州：trs.gz@trs.com.cn

支持：support@trs.com.cn

资料：info@trs.com.cn

咨询：trs@trs.com.cn

目 录

第一部分 数据库管理	1
第 1 章 基本概念	2
1.1 用户组	2
1.2 用户	3
1.3 数据库	5
1.4 数据库字段	9
1.4.1 数值的输出格式	15
1.4.1.1 半角数值输出	15
1.4.1.2 全角数值输出	16
1.4.2 日期的输出格式	16
1.4.2.1 数字日期输出	16
1.4.2.2 英文日期输出	18
1.4.2.3 中文日期输出	19
1.4.3 字符型字段的冗余索引	21
1.4.4 虚拟逻辑字段索引属性	21
1.5 视图	21
1.6 视图字段	23
1.7 数据格式	24
1.8 词典	25
第 2 章 词典的分类	28
2.1 分词词典	28
2.2 附加分词词典	29
2.3 停用词典	29
2.4 附加停用词典	30
2.5 稀疏词典	30
2.6 主题词典	31
2.7 同义词典	33
2.8 反义词典	33
2.9 实体词典	34
2.10 其它词典	34
第 3 章 字段与索引	36

3.1	字段类型	36
3.1.1	日期型字段	36
3.1.2	数值型字段	36
3.1.3	字符型字段	37
3.1.4	短语型字段	37
3.1.5	全文型字段	38
3.1.6	二进制型字段	38
3.2	索引类型	39
3.3	索引策略	39
3.3.1	按词索引策略	40
3.3.2	按字索引策略	40
3.3.3	关键词索引策略	41
3.3.4	二元组索引策略	41
3.4	索引属性	42
3.4.1	位置索引属性	42
3.4.2	格式索引属性	43
3.4.3	定义索引属性	44
第 4 章	权限控制	49
4.1	权限级别	49
4.1.1	用户级权限	49
4.1.2	数据库级权限	50
4.1.3	字段级权限	52
4.1.4	记录级权限	52
4.2	权限管理	53
4.3	权限检查	53
4.3.1	用户组操作	54
4.3.2	用户操作	54
4.3.3	数据库操作	54
4.3.4	数据库字段操作	56
4.3.5	视图操作	57
4.3.6	视图字段操作	59
4.3.7	数据格式操作	59
4.3.8	词典操作	59
第 5 章	系统配置	61
5.1	物理内存数	61
5.2	用户实例最大空闲时间	61
5.3	用户实例最小空闲时间	62

5.4	检索超时时间	62
5.5	最大排序记录数	62
5.6	最大检索历史步数	63
5.7	空闲状态起止时间	63
5.8	缺省数据库属性	64
5.9	缺省知识词典集	64
5.10	数据库路径集	64
5.11	缺省数据库备份设备	64
5.12	临时文件存放路径	65
第二部分 数据库检索		67
第 6 章 检索运算符		68
6.1	统计运算符	68
6.2	检索运算符	68
6.2.1	比较运算符	68
6.2.2	逻辑运算符	69
6.2.3	重复逻辑运算符	69
6.2.4	属性运算符	69
6.2.5	限定运算符	72
6.2.6	加权运算符	74
6.2.7	运算符的等价符号	74
6.2.8	运算符与字段的关系	74
第 7 章 函数与键值		77
7.1	统计函数	77
7.2	检索函数	77
7.3	检索键值	82
7.3.1	日期型键值	82
7.3.2	数值型键值	85
7.3.3	字符型键值	85
7.3.4	短语和全文型键值	86
7.3.5	特殊键值	86
第 8 章 表达式构造		88
8.1	统计表达式	88
8.2	检索表达式	89
8.3	缺省字段检索	92

第 9 章 检索结果处理	93
9.1 排序.....	93
9.1.1 排序向量.....	93
9.1.2 相关性排序.....	93
9.1.2.1 相关度的加权.....	94
9.1.2.2 按命中词数量排序.....	94
9.1.2.3 多字段相关性排序.....	95
9.1.2.4 焦点排序.....	95
9.1.3 TOP N排序.....	95
9.1.4 中文的排序.....	96
9.2 按字段值分类统计.....	96
9.3 快速检索与估算.....	97
第三部分 数据库加载	99
第 10 章 准备加载	100
10.1 数据库的准备.....	100
10.2 准备加载文件.....	101
第 11 章 数据文件	103
11.1 TRS格式文件.....	103
11.1.1 标准格式.....	104
11.1.2 字段内部标识号标记格式.....	108
11.1.3 字段顺序号标记格式.....	110
11.1.4 无字段标记格式.....	111
11.1.5 全文格式.....	113
11.1.6 传输格式.....	114
11.1.7 注释信息.....	115
11.1.8 标记引导符前缀.....	115
11.2 国际标准格式.....	117
11.2.1 ISO2709 记录格式.....	117
11.2.2 XML记录格式.....	118
11.3 批处理数据文件.....	122
第 12 章 多媒体数据的加载	124
12.1 加载DOCUMENT字段的数据.....	124
12.1.1 数据格式控制符.....	125
12.1.2 存储方式控制符.....	125
12.1.3 加载格式.....	127

12.2	加载BIT字段的数据	129
12.3	应用举例	130
第 13 章	日期的加载	131
13.1	通用格式日期的加载	131
13.2	英文格式日期的加载	131
13.3	无分隔格式日期的加载	132
13.4	中文格式日期的加载	132
第 14 章	词典的加载	134
14.1	加载主题词典	134
14.2	加载同义词典	136
14.3	加载反义词典	136
14.4	加载其它词典	137
第 15 章	控制文件	138
15.1	控制文件的格式	138
15.2	控制文件的变量	138
15.2.1	RECORD_FILE_FORMAT	140
15.2.2	BIT_FORMAT_DEFAULT	140
15.2.3	HTML_BASE_DEFAULT	141
15.2.4	DATA_PATH_DEFAULT	141
15.2.5	DATE_CENTURY_DEFAULT	141
15.2.6	IGNORE_RECORD_CRYPT	141
15.2.7	SYNC_CREATE_INDEX	142
15.2.8	FILE_SUFFIX_PRIOR	142
15.2.9	CLASS_BIT_STORE	143
15.2.10	CLASS_DOCUMENT_STORE	143
15.2.11	WAIVE_DOCUMENT_REFER	144
15.2.12	WAIVE_OEMFILTER_TEXT	144
15.2.13	KEEP_TEXT_ACTUAL	144
15.2.14	CHECK_TEXT_ITERATIVE	145
15.2.15	MULTI_VALUE_SEPARATER	145
15.2.16	SECTION_TAG_PREFIX	146
15.2.17	START_RECORD_FROM	146
15.2.18	PROCESS_RECORD_NUMBER	146
15.2.19	MAX_COMMIT_NUMBER	146
15.2.20	MAX_ERROR_NUMBER	147
15.2.21	COLUMN_NAME_SEQUENCE	147

15.2.22	XML_SEGMENT_MARK	147
15.2.23	XML_NEWLINE_MARK	148
15.2.24	XML_HITSHOW_MARK	149
15.2.25	XML_FILTER_CDATA	150
15.2.26	XML_OVERWRITE_PI	151
15.2.27	COLUMN_AUTO_MATCH	152
15.2.28	NAMIZE_ISO2709_FIELD	153
15.2.29	THROUGHT_ACCESS	153
15.2.30	XML_SENTENCE_MARK	154
15.2.31	MASK_OEMFILTER_ERROR	154
15.2.32	XML_KEEP_CDATA	155
15.3	控制文件实例	156

第四部分 服务器安装..... 159

第 16 章 在WINDOWS下安装..... 160

16.1	目录结构	160
16.1.1	bin目录	160
16.1.2	system目录	161
16.1.3	dict目录	162
16.1.4	ibase目录	162
16.1.5	oem目录	163
16.1.6	temp目录	163
16.2	准备工作	163
16.2.1	软件及硬件环境	163
16.2.2	确认安装方式	163
16.3	安装步骤	164
16.3.1	向导式安装	164
16.3.2	命令行安装	177
16.4	运行与维护	180
16.4.1	启动服务器	180
16.4.2	关闭服务器	181
16.4.3	服务器升级	183
16.4.4	服务器卸载	185
16.5	失败后的处理	185
16.5.1	原因及解决方法	185
16.5.2	清理工作	185
16.6	内存映射文件	186

16.7 安装多套系统	186
第 17 章 在UNIX下安装.....	187
17.1 目录结构	187
17.1.1 bin目录	187
17.1.2 system目录	187
17.1.3 dict目录	189
17.1.4 ibase目录	189
17.1.5 oem目录	190
17.1.6 temp目录	190
17.2 准备工作	190
17.2.1 软件及硬件环境	190
17.2.2 服务器的用户帐号	192
17.2.3 确认安装方式	192
17.3 安装步骤	192
17.4 运行与维护	195
17.4.1 启动服务器	195
17.4.2 关闭服务器	197
17.4.3 服务器升级	198
17.4.4 服务器卸载	200
17.5 失败后的处理	200
17.5.1 原因及解决方法	200
17.5.2 清理工作	201
17.6 安装多套系统	201
第 18 章 数据库镜像.....	202
18.1 功能描述	202
18.2 实施准备	202
18.2.1 主节点的准备工作	202
18.2.2 子节点的准备工作	203
18.3 镜像初始化	203
18.3.1 手工初始化	203
18.3.2 自动初始化	203
18.4 镜像的运行	203
18.5 新增镜像的子节点	204
18.6 新增镜像的数据库	204
18.7 当镜像数据库被破坏时	204
第 19 章 服务器工具.....	205

19.1	服务器管理员	205
19.1.1	指定服务器的端口	206
19.1.2	列出所有的用户实例	206
19.1.3	强制清除指定的用户实例	206
19.1.4	刷新日志缓冲区	207
19.1.5	在线备份服务器系统	207
19.1.6	恢复服务器系统	207
19.1.7	结束系统自动维护操作	207
19.1.8	查看系统超级用户口令	208
19.1.9	查看系统许可证	208
19.1.10	查看系统字符集	208
19.1.11	优化数据库索引与数据	208
19.1.12	检查数据库索引和视图	208
19.1.13	增量创建数据库的索引	208
19.1.14	删除指定的数据库或视图	209
19.1.15	查看用户对象的信息	209
19.1.16	查看数据库对象的信息	209
19.1.17	查看数据库的字段信息	209
19.1.18	查看视图对象的信息	209
19.1.19	修改系统注册码	210
19.1.20	备份数据库	210
19.1.21	恢复或查看数据库备份	210
19.1.22	添加数据库存储目录	210
19.2	数据库加载工具	211
19.2.1	指定服务器的端口	211
19.2.2	指定源数据的字符集	211
19.2.3	命令开关选项	212
19.2.4	指定目标数据库	212
19.2.5	指定源数据文件	212
19.2.6	指定控制文件	213
第 20 章	配置文件	214
20.1	运行参数配置文件	214
20.1.1	配置域IndexCache	215
20.1.2	配置域Maintain	216
20.1.3	配置域Retrieval	220
20.1.4	配置域Network	223
20.1.5	配置域Concurrent	224

20.1.6	配置域AutoView	225
20.1.7	配置域AutoBatch	225
20.1.8	配置域AutoReboot	226
20.1.9	配置域Background	227
20.1.10	配置实例	227
20.2	数据镜像配置文件	242
20.2.1	主节点配置项	243
20.2.2	主节点配置实例	243
20.2.3	子节点配置项	244
20.2.4	子节点配置实例	244
20.3	插件配置文件	245
20.3.1	配置域VAREnviron	245
20.3.2	配置域DOCEXTRACT	245
20.3.3	配置域URLACQUIRE	247
20.3.4	配置域DLLPRODUCT	247
20.3.5	配置实例	248
20.4	日志配置文件	250
20.4.1	方式与规模配置	251
20.4.2	内容配置	251
20.4.3	日志记录	251
20.4.4	配置实例	253
第五部分	系统附录	257
附录一	字符集	257
附录二	保留字	259
附录三	正则表达式	260
第六部分	错误信息	263
一	TRS错误	263
1.	虚拟平台错误（17400 ~ 17499, 17900 ~ 17999）	263
2.	网络扩展错误（17800 ~ 17899）	265
3.	内存文件错误（19800 ~ 19899）	267
4.	数据库维护错误（18000 ~ 18499）	268
5.	数据库检索错误（18500 ~ 18899）	270
6.	分词系统错误（18900 ~ 18999）	272

7. 远过程调用错误 (19000 ~ 19499)	272
8. 系统注册码错误 (19500 ~ 19599)	273
9. 数据词典错误 (20000 ~ 20499)	274
10. 集群系统错误 (20500 ~ 20999)	279
11. 镜像系统错误 (30001 ~ 30099)	284
12. 应用程序接口错误 (21000 ~ 21999)	285
13. 命令语言错误 (22000 ~ 22999)	286
二 操作系统错误.....	288
1. 基本系统错误 (17000 ~ 17199, 17300 ~ 17399)	288
2. 网络系统错误 (17200 ~ 17299, 17500 ~ 17799)	292
3. WINDOWS错误 (1 ~ 16999)	294

修 订

【 修订 001 ： 6.00.3001】	85
【 修订 002 ： 6.00.3002】	207
【 修订 003 ： 6.00.3003】	219
【 修订 004 ： 6.00.3003】	233
【 修订 005 ： 6.00.3005】	63
【 修订 006 ： 6.00.3005】	163
【 修订 007 ： 6.00.3005】	190
【 修订 008 ： 6.00.3005】	192
【 修订 009 ： 6.00.3005】	192
【 修订 010 ： 6.10.3100】	13
【 修订 011 ： 6.10.3100】	14
【 修订 012 ： 6.10.3100】	14
【 修订 013 ： 6.10.3100】	74
【 修订 014 ： 6.10.3100】	218
【 修订 015 ： 6.10.3100】	13
【 修订 016 ： 6.10.3100】	10
【 修订 017 ： 6.10.3100】	27
【 修订 018 ： 6.10.3100】	92
【 修订 019 ： 6.10.3100】	6
【 修订 020 ： 6.10.3100】	95
【 修订 021 ： 6.10.3100】	96
【 修订 022 ： 6.10.3120】	221
【 修订 023 ： 6.10.3120】	236
【 修订 024 ： 6.10.3130】	101
【 修订 025 ： 6.10.3130】	43
【 修订 026 ： 6.10.3190】	217
【 修订 027 ： 6.10.3190】	230
【 修订 028 ： 6.10.3190】	217
【 修订 029 ： 6.10.3190】	230
【 修订 030 ： 6.10.3190】	217
【 修订 031 ： 6.10.3190】	230
【 修订 032 ： 6.10.3190】	217
【 修订 033 ： 6.10.3190】	230
【 修订 034 ： 6.10.3194】	208
【 修订 035 ： 6.10.3194】	208
【 修订 036 ： 6.10.3194】	209
【 修订 037 ： 6.10.3194】	217
【 修订 038 ： 6.10.3194】	231
【 修订 039 ： 6.10.3200】	220
【 修订 040 ： 6.10.3200】	235
【 修订 041 ： 6.10.3200】	219
【 修订 042 ： 6.10.3210】	233
【 修订 043 ： 6.10.3210】	6
【 修订 044 ： 6.10.3210】	219
【 修订 045 ： 6.10.3210】	233

【 修订 046 : 6.10.3210】	222
【 修订 047 : 6.10.3210】	236
【 修订 048 : 6.10.3210】	97
【 修订 049 : 6.10.3211】	262
【 修订 050 : 6.10.3211】	227
【 修订 051 : 6.10.3300】	225
【 修订 052 : 6.10.3300】	22
【 修订 053 : 6.10.3300】	216
【 修订 054 : 6.10.3300】	222
【 修订 055 : 6.10.3300】	236
【 修订 056 : 6.10.3300】	222
【 修订 057 : 6.10.3300】	236
【 修订 058 : 6.10.3300】	222
【 修订 059 : 6.10.3300】	236
【 修订 060 : 6.10.3300】	222
【 修订 061 : 6.10.3300】	237
【 修订 062 : 6.10.3300】	208
【 修订 063 : 6.10.3360】	115
【 修订 064 : 6.10.3361】	154
【 修订 065 : 6.10.3362】	222
【 修订 066 : 6.10.3362】	237
【 修订 067 : 6.10.3370】	63
【 修订 068 : 6.10.3370】	96
【 修订 069 : 6.10.3382】	224
【 修订 070 : 6.10.3382】	239
【 修订 071 : 6.10.3383】	209
【 修订 072 : 6.10.3383】	209
【 修订 073 : 6.10.3383】	209
【 修订 074 : 6.10.3383】	209
【 修订 075 : 6.50.4000】	8
【 修订 076 : 6.50.4000】	13
【 修订 077 : 6.50.4000】	13
【 修订 078 : 6.50.4000】	21
【 修订 079 : 6.50.4000】	21
【 修订 080 : 6.50.4000】	14
【 修订 081 : 6.50.4000】	14
【 修订 082 : 6.50.4000】	78
【 修订 083 : 6.50.4000】	78
【 修订 084 : 6.50.4000】	9
【 修订 085 : 6.50.4000】	78
【 修订 086 : 6.50.4000】	43
【 修订 087 : 6.50.4000】	154
【 修订 088 : 6.50.4000】	149
【 修订 089 : 6.50.4003】	210
【 修订 090 : 6.50.4003】	213
【 修订 091 : 6.50.4010】	104
【 修订 092 : 6.50.4010】	80
【 修订 093 : 6.50.4010】	219
【 修订 094 : 6.50.4010】	223
【 修订 095 : 6.50.4010】	251

【 修订 096 : 6.50.4011】	223
【 修订 097 : 6.50.4021】	179
【 修订 098 : 6.50.4021】	184
【 修订 099 : 6.50.4021】	194
【 修订 100 : 6.50.4021】	199
【 修订 101 : 6.50.4022】	259
【 修订 102 : 6.50.4022】	24
【 修订 103 : 6.50.4022】	251
【 修订 104 : 6.50.4022】	206
【 修订 105 : 6.50.4050】	207
【 修订 106 : 6.50.4080】	169
【 修订 107 : 6.50.4080】	178
【 修订 108 : 6.50.4080】	194
【 修订 109 : 6.50.4100】	210
【 修订 110 : 6.50.4100】	163
【 修订 111 : 6.50.4100】	190
【 修订 112 : 6.50.4100】	94
【 修订 113 : 6.50.4101】	22
【 修订 114 : 6.50.4101】	223
【 修订 115 : 6.50.4104】	80
【 修订 116 : 6.50.4104】	80
【 修订 117 : 6.50.4140】	248
【 修订 118 : 6.50.4140】	78
【 修订 119 : 6.50.4140】	21
【 修订 120 : 6.80.5030】	72
【 修订 121 : 6.80.5030】	73
【 修订 122 : 6.80.5030】	80
【 修订 123 : 6.80.5030】	78
【 修订 124 : 6.80.5030】	78
【 修订 125 : 6.80.5030】	80
【 修订 126 : 6.80.5030】	95
【 修订 127 : 6.80.5030】	96
【 修订 128 : 6.80.5030】	80
【 修订 129 : 6.80.5030】	96
【 修订 130 : 6.80.5030】	219
【 修订 131 : 6.80.5030】	220
【 修订 132 : 6.80.5030】	97
【 修订 133 : 6.80.5030】	94
【 修订 134 : 6.80.5030】	223
【 修订 135 : 6.80.5133】	220
【 修订 136 : 6.80.5133】	223
【 修订 137 : 6.80.5140】	210
【 修订 138 : 6.80.5140】	210
【 修订 139 : 6.80.5141】	222
【 修订 140 : 6.80.5141】	43
【 修订 141 : 6.80.5141】	95
【 修订 142 : 6.80.5141】	223
【 修订 143 : 6.80.5143】	186
【 修订 144 : 6.80.5147】	26
【 修订 145 : 6.80.5147】	34

【修订 146: 6.80.5152】	6
【修订 147: 6.80.5152】	81
【修订 148: 6.80.5200】	210
【修订 149: 6.80.5227】	221
【修订 150: 6.80.5227】	221
【修订 151: 6.80.5230】	182
【修订 152: 6.80.5230】	197
【修订 153: 6.80.5230】	227
【修订 154: 6.80.5230】	156
【修订 155: 6.80.5230】	154
【修订 156: 6.80.5230】	220
【修订 157: 6.80.5238】	211
【修订 158: 6.80.5264】	81

第一部分 数据库管理

本部分介绍了 TRS 的主要特点、TRS 中各对象的定义和作用、TRS 的体系结构。其内容不涉及任何 TRS 工具的使用方法，是一个了解 TRS 概念和 TRS 全文数据库系统的概念性说明。

读者对象

本部分的读者为 TRS 全文数据库系统管理员，以及任何希望对 TRS 全文数据库系统有一个全面深入详细了解的 TRS 用户。

内容组织

本部分按如下结构组织：

- 第 1 章 基本概念
- 第 2 章 词典的分类
- 第 3 章 字段与索引
- 第 4 章 权限控制
- 第 5 章 系统配置

第 1 章 基本概念

TRS 全文数据库系统中的对象包括：

- 用户组
- 用户
- 数据库
- 数据库字段
- 视图
- 视图字段
- 数据格式
- 词典

所有这些对象都具有一致的命名规则：

- 对象名只能包括 A-Z, a-z, 0-9, _，或多字节编码字符。
- 对象名不能是 TRS 保留字(关键词)。
- 对象名在同一属域内不能重复。
- 对象名（除数据格式名外）最长为 31 个字符，且其中英文字母不区分大小写。
- 数据格式名最长为 15 个字符，且其中英文字母不区分大小写。

1.1 用户组

TRS 具有独立于操作系统的用户组管理机制。用户组的设立，使得系统对数据的安全控制更加完善，也更加简单方便。每个 TRS 全文数据库系统可支持多达 65535 个用户组。

用户组对象具有下列属性：

- 名称
名称是 TRS 用户组对象名，即用户组名。用户组名与用户组名之间，以及用户组名与用户名之间不能重复。
- 注释信息
注释信息是对 TRS 用户组对象进行说明的信息。

系统安装时，将自动创建两个用户组：ADMINISTRATOR 和 GUEST。系统还有一个匿名用户组。

1.2 用户

TRS 具有独立于操作系统的用户管理机制。要访问 TRS 全文数据库系统，必须首先申请一个用户帐号。用户帐号是实现系统及数据安全的主要手段，也是系统计费的主要依据。每个 TRS 全文数据库系统可支持多达 65535 个用户。

用户对象具有下列属性：

- 名称
名称是 TRS 用户对象名，即用户帐号名，也就是用户名。用户名与用户名之间，以及用户名与用户组名之间不能重复。
- 登录口令
登录口令是该用户登录到 TRS 全文数据库系统时必须输入的口令。设置登录口令的目的主要是为了保证用户帐号的私用性。可以不设置登录口令，否则登录口令中只允许出现字母或数字的 ASCII 字符。
- 登录互斥开关
登录互斥开关决定该用户是否允许同时从多个不同的 IP 地址进行登录。但来自本机 IP 地址的用户登录不计算在内。
- 密级
密级定义了该用户在 TRS 全文数据库系统的接触机密数据的级别。数值越小，级别越高。对于数据库中使用密级加密的数据记录，只有当用户的密级值不大于记录的密级值时，才能看到该记录的内容。0 密级是一个特殊密级，表示该用户不能查看任何用密级加密的记录。
- 类型
用户的类型实际上是用户在 TRS 全文数据库系统中的权限级别。TRS 全文数据库系统有四种用户类型：
 - ◇ 系统数据库管理员(DBA)
 - ◇ 用户组数据库管理员(GDA)
 - ◇ 系统资源用户(RESOURCE)
 - ◇ 系统登录用户(CONNECT)
- 所属用户组
即该用户所在的用户组。用户也可属于匿名用户组。但一个用户只能属于一个组。
- 最大检索历史步数
最大检索历史步数是指用户在进行数据库查询时，需要保存的最大检索历史步数。保存的检索历史可以在以后检索中直接引用，而不需要重新检索。
- 缺省附加分词词典

该用户在创建数据库时，如果数据库的附加分词词典指定为缺省值，则系统自动为数据库指定这里所给出的附加分词词典。

- 缺省附加停用词典

该用户在创建数据库时，如果数据库的附加停用词典指定为缺省值，则系统自动为数据库指定这里所给出的附加停用词典。

- 缺省稀疏词典

该用户在创建数据库时，如果数据库的稀疏词典指定为缺省值，则系统自动为数据库指定这里所给出的稀疏词典。

- 注释信息

注释信息是对 TRS 用户对象进行说明的信息。

系统安装时，将自动创建两个用户：SYSTEM 和 PUB。

SYSTEM 是 TRS 全文数据库系统的超级用户，其主要特征有：

- 是 DBA 用户。
- 属于 ADMINISTRATOR 用户组。
- 不能被注销。
- 不能修改名称。
- 可由自己修改口令。
- 可创建和注销其它任何用户。
- 可清除任何用户的口令。
- 可修改任何用户除口令以外的任何属性。
- 可授予或收回任何用户的任何权限。

PUB 是 TRS 全文数据库系统的一个公共用户，其主要特征有：

- 是 CONNECT 用户。
- 属于 GUEST 用户组。
- 不能修改名称。
- 没有口令，也不允许设置口令。
- 可被 SYSTEM 用户注销。
- 对 TRS 全文数据库系统数据对象的操作，必须通过相关的权限检查。

TRS 用户还有其它一些特点：

- 每一个用户组中允许有多个 DBA 用户。
- 每一个用户组中允许有多个 GDA 用户。
- 任何用户可修改“自己”的某些属性，这些属性包括：
 - ◇ 登录口令
 - ◇ 登录互斥开关

- ◇ 最大检索历史步数
- ◇ 缺省附加分词词典
- ◇ 缺省附加停用词典
- ◇ 缺省稀疏词典
- ◇ 注释信息

其它项只能由 SYSTEM 用户管理。

1.3 数据库

数据库是 TRS 全文数据库系统中的主要数据对象，它物理地存储了用户加载到系统中的所有数据资料。每个 TRS 全文数据库系统可管理多达数十亿个数据库。数据库实际上是一个表，表的每一行是一个数据记录，每一列则是一个数据字段，行与列的交叉点即为字段值。

数据库的规模主要指以下两个方面：能够容纳的最大记录数和数据库文件的最大容量。一个数据库的最大记录数取决于操作系统的位数和 TRS 所采用的文件系统的位数：在 32 位文件系统中，最多可容纳 4 亿多条记录，而在 64 位文件系统中，如果是 32 位操作系统，则最多可容纳 20 多亿条记录，如果是 64 位操作系统，则最多可容纳 40 多亿条记录。制约数据库文件最大容量的唯一因素是文件系统的位数，如果 TRS 采用了 32 位文件系统，则单个文件的大小不能超过 2G 字节（在 Windows 平台上不能超过 4G 字节），但如果采用 64 位文件系统，则单个文件的容量基本没有限制。操作系统与文件系统的位数没有必然的联系，许多 32 位操作系统都支持 64 位文件系统，如 WIN32、Linux、Solaris 等，所以在 32 位操作系统上，TRS 也尽可能地采用 64 位文件系统。

每个记录的最大长度为 256 兆字节，每个字段值的最大长度为 16 兆字节（但以独立方式，即 ALONE 方式存储的字段值的长度不受此限制）。但在实际中，如果一个记录（或字段值）太长，会引起操作系统忙于虚拟内存的交换，而使机器性能急剧下降，所以系统人为地限制了一个记录的最大长度为以兆为单位的物理内存大小除以 32 得到的结果数大小（最小为 1）所对应的兆字节数，如物理内存为 128 兆字节，则允许的最大记录长度为 4 兆字节。为了方便一些应用存储大文件，系统增加了一点灵活性：以 CLASS 方式（即缺省方式）存储的 DOCUMENT 字段的非文本值（包括附件）和 BIT 字段的所有类型值，其长度也受这种内存容量控制方式的限制，但最多还是不能超过 128 兆字节。这里的“记录长度”是指集中存储在数据库主文件中的数据块的长度（包含各字段的存储控制信息，以及以集中方式存储的字段值。对于字段值，如果需要并允许压缩，则计算压缩后的字段值长度），或者是各个全文型和二进制型字段中的文本数据的总长度（如果需要并允许压缩，则计算压缩后的数据长度）；“物

理内存”是指 TRS 全文数据库系统配置中的值，而不是实际机器所拥有的硬件内存，调节 TRS 全文数据库系统配置就可以改变当前系统中一个记录所允许的最大长度，但最大不能超过 256 兆字节。^{【修订 146】}

数据库对象具有下列属性：

- 名称
名称是 TRS 数据库的对象名，即数据库名，是访问、维护和管理该数据库的入口。数据库名不能与属于同一用户的其它数据库、视图以及各种词典同名。
- 别名
别名是该数据库名的别称，一般是为了方便使用而给出的数据库名的缩写，或其英文名称。数据库别名与数据库名具有同等的地位。一个数据库可以定义多个别名。
- 所有者
所有者表明哪个用户拥有该数据库。所有者与名称（别名）唯一确定一个数据库对象。所有者对该数据库能够进行任何操作，而不受其它数据安全控制的制约。
- 缺省检索字段
缺省检索字段是指在该数据库进行查询时，检索表达式中没有明确指定检索目标字段的子表达式所默认的目标字段。^{【修订 019】}在缺省情况下，缺省检索字段也是系统在进行相关性排序的目标字段^[1]，如果没有在该字段上进行检索，则就不能进行相关性排序。这个缺省检索字段是系统的默认值，虽然在检索时可以指定一个或多个其它的“缺省检索字段”，但这些“缺省检索字段”与相关性排序没有关系，相关性排序仍然取决于这个缺省检索字段。请参见 8.3 节。
- 分词词典
分词词典是该数据库按词索引时所引用的词典。如果不为数据库指定分词词典，则该数据库将按字索引。
- 附加分词词典
附加分词词典是该数据库按词索引时所引用的扩充词典。
- 停用词典
停用词典是该数据库按词索引时所引用的不允许索引的词典。^{【修订 043】}虽然停用词典不是必须的，但停用词典里一般包含了许多没有实际检索意义的符号（如标点符号），如果没有停用词典，或这些符号没有包含在停用词典里，那么这些符号不但会建立索引，影响索引性能，而且在检索时如果包括了这些符号，尤其还有一些扩展检索（如全/半角）时，会一定程度地影响检索性能。所以除非必要，建议数据库都要引用一个相应的停用词典。^[2]
- 附加停用词典
附加停用词典是该数据库按词索引时所引用的不允许索引的扩充词典。
- 稀疏词典
稀疏词典是该数据库按词索引时所引用的允许索引的词典。如果不为数据库指定

稀疏词典，则该数据库索引所有不属于停用词典和附加停用词典中的词汇。

- 公共访问权限

公共访问权限是指 TRS 全文数据库系统中所有用户对该数据库所拥有的权限。对数据库可设置的公共访问权限共有五种：

- ◇ 记录检索权
- ◇ 数据更新权
- ◇ 数据索引权
- ◇ 数据库结构修改权
- ◇ 数据库删除权

- 访问状态锁

数据库的访问状态锁用于指示数据库的当前状态，以对数据库的维护操作进行同步，保证数据及索引的一致性。数据库共五种访问状态锁：

- ◇ 共享状态锁
这种锁表明数据库未处于维护状态，可以进行查询操作，也可以立即使用其它锁对数据库进行锁定。
- ◇ 访问互斥锁
这种锁表明数据库处于禁止访问状态，其它用户不能进行任何操作，也不能使用其它锁对数据库进行锁定。除解锁操作外，所有操作将被拒绝。
- ◇ 修改互斥锁
这种锁表明数据库处于维护状态，其它用户不能进行维护操作，也不能使用其它锁对数据库进行锁定。数据修改操作将被拒绝，但允许数据库的查询操作。
- ◇ 修改等待锁
这种锁表明数据库处于维护状态，其它用户如果要进行维护操作，或使用其它锁对数据库进行锁定，系统将自动等待，直到锁的所有者释放该锁。该锁状态下允许数据库的查询操作。由于这种锁在两个数据库之间存在“死锁”的可能，因此只供系统内部使用。
- ◇ 拒绝访问锁
这种锁表明数据库处于禁止访问状态，其它用户不能进行任何操作，也不能使用其它锁对数据库进行锁定。所有操作（包括解锁操作）将被拒绝。这种锁也只供系统内部使用。

访问互斥锁与拒绝访问锁的唯一区别就是前者允许被手工解除，而后者只能由系统内部自动解除。

- 词根索引开关

词根索引开关决定数据库的短语型和全文型字段数据中的字母单词是否按词根进行索引。如果设置该开关，则短语型和全文型字段的索引将忽略“大小写敏感开关”，不区分大小写。

- 大小写敏感开关

大小写敏感开关决定数据库的字符型、短语型和全文型字段数据中的字母串是否区分大小写进行索引。

- 数字索引开关

数字索引开关决定数据库的短语型和全文型字段数据中的数字是否进行索引。

- 二元组索引开关

二元组索引开关决定数据库的短语型和全文型字段是否建立二元组索引。

- 禁止压缩开关

禁止压缩开关决定数据库是否对数据和索引进行压缩。对于 CPU 速度慢而 IO 速度较快的系统，不压缩数据与索引对数据库的入库与检索速度会略有提高，但空间膨胀率会显著增大，所以在实际使用时总是压缩索引，以获得较好的时空比。

- 禁止停用词典开关 【修订 075】

禁止停用词典开关的作用是允许建立停用词的索引（即创建数据的索引不启用停用词典），检索时停用词是否参与运算由应用进行选择。

- 数据字符集

数据字符集是指存储在数据库中的数据所使用的字符集编码。TRS 全文数据库系统支持的字符集有：

- ◇ 简体中文(GB18030)

- ◇ 繁体中文(BIG5)

- ◇ 纯英文(ENGLISH)

- ◇ 统一码(UTF8)

特别注意：当没有指定系统字符集时，字符集不支持统一码。

- 数据宿主系统

有一类应用：TRS 全文数据库系统只管理和维护索引，而数据存储在其他系统中。数据宿主系统就是用来描述存储数据的系统的信息。描述信息对 TRS 没有实际的意义，如果数据库没有数据宿主系统，则 TRS 全文数据库系统负责管理数据库的数据。不存储数据不针对日期型、数值型和字符型字段，因一些特殊应用和性能要求，这三种字段值始终存储。

- 数据存放路径

数据存放路径指明数据库数据文件的存储目录。

- 注释信息

注释信息是对 TRS 数据库对象进行说明的信息。

在定义数据库时，对每一个属性都必须给出其正确取值。数据库类型始终不允许修改。在数据库加载之前，对数据库的属性可以进行任意修改。

在加载了数据后，下列属性不再允许修改：

- 数据字符集

- 数据宿主系统

- 数据存放路径

当创建了索引后，下列属性不再允许修改：

- 分词词典
- 附加分词词典
- 停用词典
- 附加停用词典
- 稀疏词典
- 缺省检索字段
- 词根索引开关
- 大小写敏感开关
- 数字索引开关
- 二元组索引开关
- 禁止压缩开关
- 禁止停用词典开关【修订 084】

1.4 数据库字段

数据库字段是 TRS 全文数据库系统对不同数据类型进行存储和管理的主要依据，也是对数据库进行查询的唯一入口，所有的查询操作都是在一定的字段上进行的。每个数据库最多可定义 1023 个字段。

数据库字段对象具有下列属性：

- 名称
名称是 TRS 数据库字段的对象名，即字段名，是访问、维护和管理该数据库字段的入口。字段名不能与同一数据库中的其它字段同名。
- 别名
别名是该字段名的别称，一般是为了方便使用而给出的字段名的缩写，或其英文名称。字段别名与字段名具有同等的地位。一个字段可以定义多个别名。
- 数据类型
数据类型即字段类型，它决定了该字段所能存储的数据类型。TRS 能够存储和索引六种类型的数据：
 - ◇ 日期型(DATE)
 - ◇ 数值型(NUMBER)
 - ◇ 字符型(CHAR)
 - ◇ 短语型(PHRASE)
 - ◇ 全文型(DOCUMENT)

◇ 二进制型(BIT)

● 公共查询开关

公共查询开关决定是否允许系统中的所有用户在该字段上进行查询。

● 禁止索引开关

禁止索引开关决定该字段是否建立索引。BIT 字段始终不建索引。

● 最大长度或最大精度

该属性对于不同的字段类型有不同的意义：

◇ 对于 NUMBER 型字段，该属性表示字段值的小数位精度，同时决定整数位的位数。支持五种子类型，具体如下：

限定值	说明
0 ~ 6	数据为小数位精度为 0 到 6 位的实数。限定值即为小数位的最大精度，超出精度的小数部分被忽略并遵循四舍五入的规则。0 表示只取整数部分，直接忽略小数部分。 表示范围是 32 位二进制单精度浮点数。表示整数时可确保 7 位十进制有符号整数 ，即：±9999999，在这个范围之外的整数部分，不能保证精确地表示；表示带小数的实数时只能保证 7 位有效数字是精确的，大于 7 位数字的实数是近似值。
100 ~ 114	数据为小数位精度为 0 到 14 位的实数。限定值减 100，即为小数位的最大精度，超出精度的小数部分被忽略并遵循四舍五入的规则。100 表示只取整数部分，直接忽略小数部分。 表示范围是 64 位二进制双精度浮点数。表示整数时可确保 15 位十进制有符号整数 ，即：±999999999999999，在这个范围之外的整数部分，不能保证精确地表示；表示带小数的实数时只能保证 15 位有效数字是精确的，大于 15 位数字的实数是近似值。
9	表示范围是 32 位二进制有符号整数。可确保 9 位十进制有符号整数 。即：±999999999。
18	表示范围是 64 位二进制有符号整数。可确保 18 位十进制有符号整数 。即：±9999999999999999。
32	表示范围是 32 位十进制有符号整数 。 表示范围是：±99999999999999999999999999999999。

◇ 对于 CHAR 型字段，该属性表示单个字段值的最大取值长度，超过规定长度的字符将被忽略，并以不超过最大长度的最后一个完整的数据库字符集编码为截断点。**字符型单个字段值的最大长度是 255 个字节，0 表示以最大长度自动截断。**

◇ 对于其他类型的字段，该属性没有意义。

● 字段值格式【修订 016】

该属性对于不同的字段类型有不同的意义，限定格式本身不区分大小写。

◇ 对于 CHAR 型字段，字段值格式是指数据库加载时，给出的字段值所必须遵

循的限定格式。限定格式串首尾不允许出现空格。限定格式的组成规则为：

- A 表示在相应位置上必须出现一个英文字母。
- 9 表示在相应位置上必须出现一个数字。
- 空格 表示在相应位置上必须出现一个空格。
- X 表示在相应位置上可出现一个任意字符。

如果该字段设置了“最大长度”限制，则“限定格式”的长度必须与“最大长度”相一致。一个字段设置了“限定格式”，就意味着每个记录的字段值必须与“限定格式”具有相同的字节长度。

- ◇ 对于 NUMBER 型字段，字段值格式是指输出字段值时使用的格式。请参见 1.4.1 节的说明。
- ◇ 对于 DATE 型字段，字段值格式是指输出字段值时使用的格式。请参见 1.4.2 节的说明。
- ◇ 对于其他类型的字段，不支持字段值格式。

- 缺省字段值

缺省字段值是指在数据库加载时，对于未给出字段值的 DATE、NUMBER 或 CHAR 型字段所对应的缺省取值。DATE 型字段的缺省值只支持绝对年份，而不支持缺省世纪数的短格式年份，即总是认为缺省世纪数为 0。缺省字段值中只允许指定一个字段值。DATE 字段如果取 \$DATE（不区分大小写）作为缺省值，则表示字段值是记录入库日期的“年月日”部分，如果取 \$TIME（不区分大小写）作为缺省值，则表示字段值是记录入库日期的“年月日时分秒”；NUMBER 字段如果取 \$CODE（不区分大小写）作为缺省值，则表示字段值是该记录入库时的唯一自然数顺序码，记录之间以加载时间先后为序，**自然数顺序码的最大位数与该字段的子类型所能支持的最大十进制整数位数相同**；CHAR 字段如果取 \$CODE（不区分大小写）作为缺省值，则表示字段值是该记录入库时的非数字唯一顺序码，记录之间以加载的时间先后为序，如果取 \$USER（不区分大小写）作为缺省值，则表示字段值是当前进行操作的用户名，当该 CHAR 字段定义了字段值格式时，其缺省字段值还必须符合这一格式限定。

- 字段值范围

字段值范围是指在加载数据时，DATE、NUMBER 或 CHAR 型字段的合法取值的集合。取值范围的定义格式包括：

- ◇ 枚举定义 格式为“V1,V2,V3,Vn”，即枚举该字段上所有可能取值。
- ◇ 区间定义 格式为“V1,...,V2”，表示该字段上所有可能取值必须在最小值 V1 和最大值 V2 之间。
- ◇ 混合定义 格式为“V1,V2,...,V3,Vn”，即枚举定义与区间定义的混合定义。

以上各格式中的逗号和点号都必须是半角符号。CHAR 型字段的取值范围所给出的每个值串首尾不允许空格，当该 CHAR 字段定义了字段值格式时，其字段值范围定义中的字段值还必须符合这一格式限定。DATE 型字段的取值范围所给出的

每个日期中只支持绝对（完全）年份，而不支持缺省世纪数的短格式年份，即总是认为缺省世纪数为0。只有 DATE 型字段的取值范围支持模糊格式，即允许日期格式中缺省一些成份，但对于区间定义定义的范围，其中的 V1 和 V2 “前模糊”的成份必须是相同的，即在“年月日时分秒”的顺序下，V1 和 V2 的第一个确定的日期成份的位置必须是相同的。对于 DATE 型字段还必须注意：其日期的书写格式中不要使用逗号 ‘,’ 和点号 ‘.’ 分隔各成分，以免与“取值范围”本身的表示格式中的逗号和点号发生冲突。

- 多值允许开关

多值允许开关决定在一个记录中该字段（DATE、NUMBER、CHAR 型或 BIT 型字段）是否允许存在多个值。

- 值唯一性开关

值唯一性开关决定在数据库的所有记录中，该字段（DATE、NUMBER 或 CHAR 型字段）是否允许出现重复的字段值。该开关将自动关闭“禁止索引开关”。

- 空值禁止开关

空值禁止开关决定在一个记录中该字段（DATE、NUMBER 或 CHAR 型字段）是否允许没有字段值。

- 混合索引模式

混合索引模式只对具有“按词”索引属性的数据库的 PHRASE 和 DOCUMENT 字段有效。字词混合索引可以细分成三部分：只属于“按词”的索引；只属于“按字”的索引；属于“按词”和“按字”的公共索引。混合索引模式是指明其中字词公共索引部分的存储方式，共有四种：

- ◇ 全冗余模式 将公共部分分别在“按词”和“按字”索引中各存一份。优点是检索时直接取各自的索引即可；缺点是索引空间膨胀大。

- ◇ 半冗余模式 将公共部分只存储在“按词”的索引中。优点是“按词”检索时可直接取“按词”的索引即可，且索引空间膨小；缺点是“按字”检索时要取“按字”的索引和“按词”的索引进行合并，多一次“或”运算。

- ◇ 零冗余模式 将公共部分单独存储。优点是索引空间膨小；缺点是“按字”检索时要取“按字”的索引和“按词”的索引进行合并，多一次“或”运算，逻辑上，“按词”检索也多一次“或”运算。

- ◇ 禁止词模式 这种模式实际不是“混合索引”，而是为了实现在其它未指明混合索引模式的字段“按词”索引的情况下，该字段“按字”索引。

一般地，如果一个字片是“按词”和“按字”的公共部分，则它不可能存在只属于“按词”的索引。因此零冗余模式与半冗余模式的实际效果可能是一样的，没有明显的差别。如果一个 PHRASE 或 DOCUMENT 字段未指明任何混合索引模式，则该字段是“按词”，还是“按字”索引，取决于数据库的属性。

- 全字索引开关

全字索引开关只有在混合索引模式有效或者数据库按字索引的前提下才有效。如果是全字索引，则“按字”检索能够进行单个字的查询。参见 3.3.2 和 3.3.4 节。

- 词根索引语种
如果数据库指明进行词根索引，则所有的字母文字的单词都要进行词根处理。词根索引语种就是用来指定数据库中字母文字的语种，目前支持 12 种语言：英语、俄语、法语、西班牙语、葡萄牙语、意大利语、德语、荷兰语、瑞典语、挪威语、丹麦语、芬兰语。显然当一个字段中有多种字母文字时，就无法进行正确的词根处理。如果不指明词根索引语种，则系统将根据字母的种类默认一个语种：拉丁字母默认为英语；西里尔字母默认为俄语。
- 数据缓冲开关
数据缓冲开关决定该字段的“数据”是否^{【修订 015】}集中存储，并在“检索缓冲”启动的情况下时，允许进入内存缓冲。虽然所有类型的字段都允许进行缓冲，但 DOCUMENT 和 BIT 类型的字段只存储一个标志，表示字段是否有值，而不是存储实际的字段值。而将一些关键的（如用于排序、统计的）小字段进行集中存储的一个好处是：集中存储的文件会小于常规存储的文件，当只读取这些字段的值时，从小文件里读效率更高^{【4】}。
- 相关性开关^{【修订 010】}
相关性开关决定检索命中该字段时，该字段是否允许参与相关性排序。除 BIT 类型的字段外，其它类型的字段都允许参与相关性排序。
- 虚拟逻辑字段开关^{【修订 076】}
如果虚拟逻辑字段开关被打开，则表示该字段被设置成一个虚拟字段。虚拟字段不能存储数据和索引，而只能在该字段的索引属性中定义一些它所对应的实际字段及其逻辑关系，在检索时作为单一的字段使用。只有 BIT 字段才能设置为虚拟逻辑字段，且必须设置“索引属性”。
- 缺省虚拟字段开关^{【修订 077】}
如果缺省虚拟字段开关被打开，则表示在检索时该虚拟逻辑字段的字段名可以用特殊的字符串“\$DUMMY”来替代，以便于检索表达式的书写或选项的指定。
- 索引属性
 - ◇ 对 DOCUMENT 字段来说，索引属性是指在检索时，可以用来限定检索的目标词或短语的位置和格式的属性，包括位置索引属性和格式索引属性。位置索引属性是传统的索引属性，包括：
 - 段落(SEG)
 - 句子(SEN)
 - 位置(POS)位置索引属性是必须的，在指定索引属性时，可以使用其中之一，或者多个的组合。但格式索引属性则不是必须的。请参见 3.4 节的说明。
 - ◇ 对于 CHAR 字段来说，索引属性是用来定义冗余索引模式的。请参见 1.4.3 节的说明。
 - ◇ 对 BIT 字段来说，只有它被设置成虚拟逻辑字段时才支持索引属性。请参见 1.4.4 节的说明。

◇ 其它类型字段不支持索引属性。

- 索引存放路径
索引存放路径指明该字段的索引文件的存储目录。
- 注释信息
注释信息是对 TRS 数据库字段对象进行说明的信息。

CHAR、PHRASE 和 DOCUMENT 三种字段类型不易区别,其最大的差别是:CHAR 和 PHRASE 字段支持按字段值进行排序; PHRASE 和 DOCUMENT 字段则支持全文检索; DOCUMENT 字段能存储非文本数据。

在定义数据库字段时,对每一个属性都必须给出其正确取值。在数据库加载之前,对数据库字段的属性可以进行任意修改。

在加载了数据后,下列属性不再允许修改:

- 数据类型
- 多值允许开关
- 值唯一性开关
- 空值禁止开关
- 最大长度或精度
- 数据缓冲开关【修订 011】
- 相关性开关【修订 012】

仅针对 PHRASE 和 DOCUMENT 类型字段而言,DATE、NUMBER 和 CHAR 类型字段在任何时候都允许修改。

- 虚拟逻辑字段开关【修订 080】

当创建了索引后,下列属性不再允许修改:

- 索引存放路径
- 索引属性
 - ◇ 对 DOCUMENT 字段来说,如果原来已有格式索引属性,则对格式索引属性的修改有些例外:允许添加新的格式索引属性;允许为格式索引属性添加新的特征;允许添加/修改格式索引属性的别名;允许添加/修改格式索引属性的特征别名。请参见 3.4 节的说明。
 - ◇ 对于虚拟逻辑字段来说,创建索引后仍然可以修改索引属性。【修订 081】
- 混合索引模式
- 全字索引开关
- 词根索引语种

除了可定义的字段外,每个数据库都有一个固定的名为 DOCID 的逻辑字段。该

1.4.1.2 全角数值输出

- L. DP (全角符号)
输出效果：左对齐；全角；保留小数点后的无效0。
例如：1 2 3 4 5 6 . 7 8 9 0 0 0
1 2 3 4 5 . 0 0 0 0 0 0
- L. D (全角符号)
输出效果：左对齐；全角；不保留小数点后的无效0。
例如：1 2 3 4 5 6 . 7 8 9
1 2 3 4 5 . 0
- nR. DP (全角符号)
其中 n 是一个十进制数（半角），表示对齐的基准宽度，即数字串的总宽度（字节数）。如：50R. DP。如果宽度不够，则没有对齐效果。
输出效果：右对齐；全角；保留小数点后的无效0。
例如：1 2 3 4 5 6 . 7 8 9 0 0 0
1 2 3 4 5 . 0 0 0 0 0 0
- nR. D (全角符号)
其中 n 是一个十进制数（半角），表示对齐的基准宽度，即数字串的总宽度（字节数）。如：30R. D。如果宽度不够，则没有对齐效果。
输出效果：右对齐；全角；不保留小数点后的无效0。
例如：1 2 3 4 5 6 . 7 8 9
1 2 3 4 5 . 0
- nM. D (全角符号)
其中 n 是一个十进制数（半角），表示对齐的基准宽度，即数字串的总宽度（字节数）。如：30M. D。如果宽度不够，则没有对齐效果。
输出效果：实数按小数点对齐，整数居中对齐；全角；不保留小数点后的无效0。
例如：1 2 3 4 5 6 . 7 8 9 或者 1 2 3 4 5 6 7 8 9
1 2 3 4 5 . 0 1 2 3 4 5

1.4.2 日期的输出格式

系统支持 36 种格式的日期输出，分别包括 14 种数字输出格式、12 种英文输出格式和 10 种中文输出格式。

1.4.2.1 数字日期输出

- Y.MM.DD HH:MM:SS (半角符号，缺省格式)
输出效果：半角；保留“时分秒”；用固定两位表示“月日时分秒”；其中“年月

日”的分隔符是“.”。

例如：2006.12.09 16:09:35

- Y.M.D H:M:S (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“月日时分秒”；其中“年月日”的分隔符是“.”。
例如：2006.12.9 16:9:35
- Y.MM.DD (半角符号)
输出效果：半角；不保留“时分秒”；用固定两位表示“月日”；其中“年月日”的分隔符是“.”。
例如：2006.12.09
- Y.M.D (半角符号)
输出效果：半角；不保留“时分秒”；用实际位数表示“月日”；其中“年月日”的分隔符是“.”。
例如：2006.12.9
- Y/MM/DD HH:MM:SS (半角符号)
输出效果：半角；保留“时分秒”；用固定两位表示“月日时分秒”；其中“年月日”的分隔符是“/”。
例如：2006/12/09 16:09:35
- Y/M/D H:M:S (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“月日时分秒”；其中“年月日”的分隔符是“/”。
例如：2006/12/9 16:9:35
- Y/MM/DD (半角符号)
输出效果：半角；不保留“时分秒”；用固定两位表示“月日”；其中“年月日”的分隔符是“/”。
例如：2006/12/09
- Y/M/D (半角符号)
输出效果：半角；不保留“时分秒”；用实际位数表示“月日”；其中“年月日”的分隔符是“/”。
例如：2006/12/9
- Y-MM-DD HH:MM:SS (半角符号)
输出效果：半角；保留“时分秒”；用固定两位表示“月日时分秒”；其中“年月日”的分隔符是“-”。
例如：2006-12-09 16:09:35
- Y-M-D H:M:S (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“月日时分秒”；其中“年月日”的分隔符是“-”。
例如：2006-12-9 16:9:35

- Y-MM-DD (半角符号)
输出效果: 半角; 不保留“时分秒”; 用固定两位表示“月日”; 其中“年月日”的分隔符是“-”。
例如: 2006-12-09
- Y-M-D (半角符号)
输出效果: 半角; 不保留“时分秒”; 用实际位数表示“月日”; 其中“年月日”的分隔符是“-”。
例如: 2006-12-9
- YMMDDHHMMSS (半角符号)
输出效果: 半角; 无分隔符; 保留“时分秒”; 用固定两位表示“月日时分秒”; 用固定四位表示“年”。
例如: 20061209160930
- YMMDD (半角符号)
输出效果: 半角; 无分隔符; 不保留“时分秒”; 用固定两位表示“月日”; 用固定四位表示“年”。
例如: 20061209

1.4.2.2 英文日期输出

- MonthDay,Y HH:MM:SS (半角符号)
输出效果: 半角; 保留“时分秒”; 用固定两位表示“日时分秒”; 格式最前面是用英文全称表示的“月”。
例如: December 09,2006 16:09:35
- Mon.Day,Y H:M:S (半角符号)
输出效果: 半角; 保留“时分秒”; 用实际位数表示“日时分秒”; 格式最前面是用英文缩写表示的“月”。
例如: Dec.9,2006 16:9:35
- MonthDay,Y (半角符号)
输出效果: 半角; 不保留“时分秒”; 用固定两位表示“日”; 格式最前面是用英文全称表示的“月”。
例如: December 09,2006
- Mon.Day,Y (半角符号)
输出效果: 半角; 不保留“时分秒”; 用实际位数表示“日”; 格式最前面是用英文缩写表示的“月”。
例如: Dec.9,2006
- Day-Month-Y HH:MM:SS (半角符号)
输出效果: 半角; 保留“时分秒”; 用固定两位表示“日时分秒”; 格式第二项是用英文全称表示的“月”; 其中“年月日”的分隔符是“-”。

例如：09-December-2006 16:09:35

- Day-Mon.Y H:M:S (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“日时分秒”；格式第二项是用英文缩写表示的“月”；其中“月日”的分隔符是“-”。
例如：9-Dec.2006 16:9:35
- Day-Month-Y (半角符号)
输出效果：半角；不保留“时分秒”；用固定两位表示“日”；格式第二项是用英文全称表示的“月”；其中“年月日”的分隔符是“-”。
例如：09-December-2006
- Day-Mon.Y (半角符号)
输出效果：半角；不保留“时分秒”；用实际位数表示“日”；格式第二项是用英文缩写表示的“月”；其中“月日”的分隔符是“-”。
例如：9-Dec.2006
- DayMonthY HH:MM:SS (半角符号)
输出效果：半角；保留“时分秒”；用固定两位表示“日时分秒”；格式第二项是用英文全称表示的“月”；其中“年月日”的分隔符是空格。
例如：09 December 2006 16:09:35
- DayMon.Y H:M:S (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“日时分秒”；格式第二项是用英文缩写表示的“月”；其中“月日”的分隔符是空格。
例如：9 Dec.2006 16:9:35
- DayMonthY (半角符号)
输出效果：半角；不保留“时分秒”；用固定两位表示“日时分秒”；格式第二项是用英文全称表示的“月”；其中“年月日”的分隔符是空格。
例如：09 December 2006
- DayMon.Y (半角符号)
输出效果：半角；不保留“时分秒”；用实际位数表示“日”；格式第二项是用英文全称表示的“月”；其中“月日”的分隔符是空格。
例如：9 Dec.2006

1.4.2.3 中文日期输出

- Y年MM月DD日HH时MM分SS秒 (半角符号)
输出效果：半角；保留“时分秒”；用固定两位表示“月日时分秒”；使用中文单位“年月日时分秒”。
例如：2006年12月09日16时09分35秒
- Y年M月D日H时M分S秒 (半角符号)
输出效果：半角；保留“时分秒”；用实际位数表示“月日时分秒”；使用中文单

位“年月日时分秒”。

例如：2006年12月9日16时9分35秒

- Y年MM月DD日 (半角符号)
输出效果：半角；不保留“时分秒”；用固定两位表示“月日”；使用中文单位“年月日”。
例如：2006年12月09日
- Y年M月D日 (半角符号)
输出效果：半角；不保留“时分秒”；用实际位数表示“月日”；使用中文单位“年月日”。
例如：2006年12月9日
- Y年MM月DD日HH时MM分SS秒 (全角符号)
输出效果：全角；保留“时分秒”；用固定两位表示“月日时分秒”；使用中文单位“年月日时分秒”。
例如：2006年12月09日16时09分35秒
- Y年M月D日H时M分S秒 (全角符号)
输出效果：全角；保留“时分秒”；用实际位数表示“月日时分秒”；使用中文单位“年月日时分秒”。
例如：2006年12月9日16时9分35秒
- Y年MM月DD日 (全角符号)
输出效果：全角；不保留“时分秒”；用固定两位表示“月日”；使用中文单位“年月日”。
例如：2006年12月09日
- Y年M月D日 (全角符号)
输出效果：全角；不保留“时分秒”；用实际位数表示“月日”；使用中文单位“年月日”。
例如：2006年12月9日
- 年月日时分秒 (中文)
输出效果：使用中文表示法；保留“时分秒”。
例如：二00六年十二月九日十六时九分三十五秒
- 年月日 (中文)
输出效果：使用中文表示法；不保留“时分秒”。
例如：二00六年十二月九日

特别说明：除第一种格式（缺省格式）外，对于其它的所有格式，如果格式包含了“时分秒”，但实际数据缺少其中的一个或两个，则按缺省格式输出；如果格式包含了“年月日”，但实际数据缺少其中的一个或两个，则按缺省格式输出。

1.4.3 字符型字段的冗余索引

要使用 CHAR 字段的冗余索引功能，必须首先在该字段的“索引属性”里定义冗余索引模式。CHAR 字段的冗余索引是从原字段值中按规则抽取部分值的组合并进行索引，允许从字段值中抽取多个冗余值。抽取规则采用正则表达式语法，每个正则表达式对应一个冗余值，正则表达式之间用换行符分隔，每行一个正则表达式。表达式语法规则请参见附录三。【修订 078】

1.4.4 虚拟逻辑字段索引属性

如果一个 BIT 字段的打开了“虚拟逻辑字段”开关，则该 BIT 字段就成为了“虚拟逻辑字段”，同时还必须在虚拟逻辑字段的“索引属性”里定义一个它所对应的实际字段的列表（至少包含一个字段），以及这些字段的逻辑关系、运算优先模式和相关性排序加权。【修订 079】

例如属性：“字段 A/10,字段 B/100,*&”。其中：

- 符号‘&’是运算优先模式，表示“字段逻辑优先”，即虚拟逻辑在每个检索原子（检索键值）上进行扩展。缺省时表示“表达式逻辑优先”，即虚拟逻辑在虚拟逻辑字段所限定的最大子表达式上进行扩展。
- 符号‘*’是‘字段 A’和‘字段 B’之间的逻辑关系（AND 运算）。也可以是其它运算，如：‘+’（OR 运算）、‘^’（XOR 运算）以及‘-’（NOT 运算）。逻辑关系缺省时默认为 OR 运算。
- 符号‘,’是字段之间、以及字段与逻辑关系之间的分隔符。逻辑关系符和运算优先模式必须放在最后。
- 数值‘10’和‘100’是定义字段在参与相关性排序时的权重加权数。这些值将分别与表达式中该虚拟字段上的权重值相乘，其积作为各字段的最后权重。所以表达式中该虚拟字段上的权重也起到“是否参与相关性排序”的开关作用。【修订 119】

1.5 视图

视图是 TRS 全文数据库系统中的一种辅助数据对象。TRS 全文数据库系统支持对数据库记录和字段进行直接的授权机制来实现数据的安全保密性控制，但有时显得有些烦琐。通过视图来限制某些用户对数据库字段和记录的查询，就可以用一种简便的方式间接地完成对数据库记录和字段查询权的控制。

视图实际上是一个逻辑数据表，它是逻辑地从一个或多个数据库中抽取一个或多个字段并满足指定条件的记录。视图并不物理地存储数据，只是描述了组成该视图的

数据所遵循的逻辑。每个 TRS 全文数据库系统可管理多达数十亿个视图，每个视图中可包含 1 到^{【修订 113】}512^{【4】} 个数据库。

视图对象具有下列属性：

- 名称
名称是 TRS 视图的对象名，即视图名，是访问、维护和管理该视图的入口。视图名不能与属于同一用户的其它视图、数据库以及各种词典同名。
- 别名
别名是该视图名的别称，一般是为了方便使用而给出的视图名的缩写，或其英文名称。视图别名与视图名具有同等的地位。一个视图可以定义多个别名。
- 所有者
所有者表明哪个用户拥有该视图。所有者与名称（别名）唯一确定一个视图对象。所有者对该视图能够进行任何操作，而不受其它数据安全控制的制约。
- 缺省检索字段
缺省检索字段是指在该视图进行查询时，检索表达式没有明确指定检索目标字段的子表达式所默认的目标字段。
- 公共访问权限
公共访问权限是指 TRS 全文数据库系统中所有用户对该视图所拥有的权限。对视图可设置的公共访问权限共有三种：
 - ◇ 记录检索权
 - ◇ 视图结构修改权
 - ◇ 视图删除权
- 所含数据库
所含数据库是指该视图所包含的数据库。视图中不能包含视图。
- 记录选取条件
记录选取条件实际上是一个对该视图所包含的每个数据库进行检索的检索表达式。记录选取条件中的检索目标字段（包括缺省的目标字段）动态地对应于该视图所包含的每个数据库中的字段，而不是该视图的字段。**如果从多个数据库中选取数据的条件不一致，则必须在各个数据库中建立不存储任何数据的空字段进行相互之间的映射，以保证选取条件中的字段在各个数据库中都存在。**^{【修订 052】}对于自动模式视图，该“条件”对应于分裂数据库的记录数，其表示方法为： $\$RECORD=n$ ， n 是一个 10000 到 20000000 之间的任意数，当视图中最后一个数据库的记录数不小于 n 时将分裂除一个新的数据库。如果“条件”为空，则分裂数据库的记录数由系统配置确定（参见 20.1.6 节说明）。^{【4】}
- 注释信息
注释信息是对 TRS 视图对象进行说明的信息。

一般情况下，视图只能用于对数据的查询，而不支持对数据的“增、删、改”等

操作, 但 TRS 全文数据库系统支持一种特殊的视图: 数据库自动分裂模式视图, 简称为自动模式视图。这种视图不但支持对数据记录的维护, 而且随着数据记录的增加, 能够自动分裂数据库, 以维持单个数据库的规模。自动模式视图除了普通视图的三种访问权限外, 还具有数据索引权限和数据更新权限的管理。

1.6 视图字段

视图字段是对视图进行查询的唯一入口, 所有的查询操作都是在一定的字段上进行的。每个视图最多可定义 1023 个字段, 每个字段可映射 1 到 64 个数据库中的某个字段。

视图字段对象具有下列属性:

- 名称
名称是 TRS 视图字段的对象名, 即字段名, 是访问、维护和管理该视图字段的入口。字段名不能与同一视图中的其它字段同名。
- 别名
别名是该字段名的别称, 一般是为了方便使用而给出的字段名的缩写, 或其英文名称。字段别名与字段名具有同等的地位。一个字段可以定义多个别名。
- 数据类型
数据类型即字段类型, 它表明了该字段所映射的数据库字段所存储的数据类型。对应于数据库字段, 视图字段也有六种类型:
 - ◇ 日期型(DATE)
 - ◇ 数值型(NUMBER)
 - ◇ 字符型(CHAR)
 - ◇ 短语型(PHRASE)
 - ◇ 全文型(DOCUMENT)
 - ◇ 二进制型(BIT)
- 公共查询开关
公共查询开关决定是否允许 TRS 全文数据库系统中的所有用户在该字段上进行查询。
- 数据库字段映射
数据库字段映射决定了该字段对应于哪些数据库中的哪个字段。
- 注释信息
注释信息是对 TRS 视图字段对象进行说明的信息。

1.7 数据格式

数据格式用来说明多媒体数据的格式类型。在数据库加载时，TRS 全文数据库系统将根据数据格式确定相应的数据存储方式和文本抽取方法；在数据输出时，TRS 全文数据库系统将根据数据格式对数据进行相应的处理，便于浏览。每个 TRS 全文数据库系统可管理多达 1023 个数据格式对象。

数据格式对象具有下列属性：

- 名称
名称是 TRS 数据格式的对象名，即数据格式名，是访问、维护和管理该数据格式的入口。数据格式名不能重复，并且**只能由英文字母和数字组成**。
- 所属类
即该数据格式属于哪种类型。TRS 全文数据库系统把数据格式划分成六种类型：
 - ◇ 纯文本(TEXT)
 - ◇ 格式文档(DOCS)
 - ◇ 图像(IMAGE)
 - ◇ 音频(AUDIO)
 - ◇ 视频(VIDEO)
 - ◇ 自定义(BIT)
- 文件后缀
文件后缀是指该数据格式的数据文件通常使用的文件名后缀，**只能由英文字母和数字组成**。
- 压缩开关
压缩开关决定该数据格式的数据是否需要无损压缩。
- 注释信息
注释信息是对数据格式对象进行说明的信息。

TRS 全文数据库系统在安装时，将自动创建下列数据格式对象(名称)：

数据格式	格式说明	文件后缀
BIT	不确定格式的二进制数据	bit
TEXT	文本数据格式	txt
WORD	Microsoft Word 文件格式	doc;docx;docm
WS	WordStar 文件格式	ws
WP	WordProfect 文件格式	wp
XCL	Microsoft Excel 文件格式	xls;xlsx;xlsm
WPS	金山 WPS 文件格式	wps
S2	方正、华光二扫文件格式	s2
DOCS	不确定格式的文档文件格式 <small>【修订 102】</small>	dcs

数据格式	格式说明	文件后缀
AUDIO	不确定格式的音频文件格式	aud
IMAGE	通用图象文件格式	img
VIDEO	不确定格式的视频文件格式	vid
AVI	AVI 动画文件格式	avi
MPEG	MPEG 压缩动画文件格式	mpg
WAVE	Wave 音频文件格式	wav
MIDI	MIDI 音频文件格式	mid
CDA	CD 音频文件格式	cda
PS	PostScript 文件格式	ps
GIF	GIF 图象文件格式	gif
TIF	TIF 图象文件格式	tif
PCX	PCX 图象文件格式	pcx
BMP	BMP 图象文件格式	bmp
JPEG	JPEG 图象文件格式	jpg
HTML	HTML 超文本文件格式	htm;html
PDF	Adobe PDF 文件格式	pdf
PPT	Microsoft Power Piont 文件格式	ppt;pps;pptx;ppsm
RTF	Rich Text 文件格式	rtf
XML	XML 文件格式	xml
MP3	MP3 文件格式	mp3
MP4	MP4 文件格式	mp4
ZIP	ZIP 压缩文件格式	zip;z;Z
RAR	RAR 压缩文件格式	rar
TAR	TAR 文件归档格式	tar

TRS 全文数据库系统允许用户自己定义新的数据格式,或者修改已有的数据格式,但系统定义格式的名称和所属类不允许修改。

需要特别注意: 数据格式一旦定义,就不允许删除,所以创建必须谨慎,只有超级用户(即 SYSTEM 用户)有权创建新的数据格式。

1.8 词典

TRS 的全文检索不单是一种快速的字串匹配系统,要获得良好的检索效果,必须使用一系列知识词典。词典对象是 TRS 全文数据库系统中非常重要的基础资源,是按词索引和检索技术的根本依据。

词典对象具有下列属性:

- 名称
名称是 TRS 词典的对象名，即词典名，是访问、维护和管理该词典的入口。词典名不能与属于同一用户的其它视图、数据库以及各种词典同名。
- 所有者
所有者表明哪个用户拥有该词典。所有者与名称（别名）唯一确定一个词典对象。所有者对该词典能够进行任何操作，而不受其它数据安全控制的制约。
- 类型
按用途划分，TRS 全文数据库系统有九种类型的词典：
 - ✧ 分词词典(SEGMENT)
 - ✧ 附加分词词典(SEGMENTEX)
 - ✧ 停用词典(STOP)
 - ✧ 附加停用词典(STOPEX)
 - ✧ 稀疏词典(SPARSE)
 - ✧ 主题词典(THESAURUS)
 - ✧ 同义词典(SYNONYM)
 - ✧ 反义词典(ANTONYM)
 - ✧ 实体词典(ENTITY) 【修订 144】其中，分词词典、附加分词词典、停用词典、附加停用词典和稀疏词典是用来建立数据库的索引并进行查询的词典，统称为**索引词典**；主题词典、同义词典、反义词典和实体词典是用于智能概念扩展检索的词典，统称为**辅助知识词典**。
- 公共访问权限
公共访问权限是指 TRS 全文数据库系统中所有用户对该词典所拥有的权限。对词典可设置的公共访问权限共有三种：
 - ✧ 引用权（检索权）
 - ✧ 维护权（更新权）
 - ✧ 删除权其中引用权总是开放的。
- 字符集
字符集是指该词典语言版本。TRS 全文数据库系统支持的字符集有：
 - ✧ 简体中文(GB18030)
 - ✧ 繁体中文(BIG5)
 - ✧ 纯英文(ENGLISH)
 - ✧ 统一码(UTF8)特别注意：当没有指定系统字符集时，字符集不支持统一码。
- 数据存放路径
数据存放路径指明词典数据文件的存储目录。词典创建后，其存储目录任何时候都不允许修改。
- 注释信息

注释信息是对 TRS 词典对象进行说明的信息。

系统安装时，将自动创建以下几部系统缺省词典：分词词典、停用词典、主题词典、同义词典、反义词典和实体词典。

需要特别注意：对索引词典进行维护后^{【修订 017】}不必要重新启动服务器就会即时生效，但引用了这些词典的数据库必须重建索引才会有正确的检索结果^{【2】}。

第 2 章 词典的分类

如 1.8 节所述，按用途划分，TRS 全文数据库系统主要有八种类型的词典是：分词词典、附加分词词典、停用词典、附加停用词典、稀疏词典、主题词典、同义词典和反义词典。除此之外，还存在一些其它类型的词典，下面将分别予以说明。

2.1 分词词典

按词索引和检索是 TRS 的主要特点之一，内嵌的分词系统采用以词典为基础的分词算法。分词词典用于对表意文字（中文、日文、韩文）的自动分词，由若干个词汇组成，词典中由字母和数字组成的词汇不起作用，每个词汇的最大长度为 48 个字节。

TRS 全文数据库系统提供的缺省分词词典的每个词汇具有语法属性，以提高分词的准确性。用户定义的分词词典或增加的词汇则没有属性。该分词词典是经过加密处理的，用户可以浏览词典的词汇，但不能浏览其属性。

分词词典的使用要点：

- 在创建数据库时引用分词词典。分词词典被引用后，一般不允许再对该词典进行维护，除非重新创建数据库的索引。
- 在一般情况下均使用系统提供的缺省分词词典。
- 对英文数据库不使用分词词典。
- 一般不对系统缺省分词词典进行维护，当需要加入用户专业词汇时，可通过创建附加分词词典来实现。
- 分词词典中由字母和数字组成的词汇不起作用，因此不要加入这些词汇。
- 生僻词并不一定要加入到分词词典，这些词即使不在词典中，也能检索。

TRS 全文数据库系统提供一个特殊空分词词典，其特殊性在于：系统安装或升级时，该词典将被自动创建，不能进行维护，也不能以其它任何手工方式创建（否则不具有自动创建时的功能）；词典中没有任何词项，当数据库引用该词典时，将按字建立索引，并用来指定“所有的单字都建索引”，以便能够进行单个字的查询（请参见 3.3.2 和 3.3.4 节）。因此也把该词典叫做“全字词典”。

2.2 附加分词词典

附加分词词典是分词词典的补充。TRS 提供的缺省分词词典是一部通用的分词词典，在多数情况下没有包括用户应用的特殊词汇。在这种情况下，一般不是对系统缺省分词词典进行维护，而是通过建立新的附加分词词典来定义新的词汇，系统在自动分词时将同时参考分词词典和附加分词词典中的词汇。

附加分词词典由一系列词汇组成，词典中由字母和数字组成的词汇不起作用，每个词汇的最大长度为 48 个字节。

附加分词词典的使用要点：

- 在创建数据库时引用附加分词词典。附加分词词典被引用后，一般不允许再对该词典进行维护，除非重新创建数据库的索引。
- 在一般情况下均使用系统提供的缺省分词词典。即使在数据库不引用任何附加分词词典，用户应用的特殊词汇也能进行检索。
- 对英文数据库不使用附加分词词典。
- 附加分词词典中由字母和数字组成的词汇不起作用，因此不要在加入这些词汇。
- 附加分词词典一般包含了某个领域的专业词汇。

2.3 停用词典

停用词典又称为禁用词典，它是由一系列没有检索意义的高频词组成的，如英文文献中的“the”、“of”、“and”、“to”等，中文文献中的“的”、“关于”、“但是”、“而且”等。从相关性方面讲，文献中的这些词没有检索意义，因为这些词会出现在每篇文献中。在检索系统中，通常使用停用词典来过滤掉文献中没有检索意义的词，以最大限度地减少数据库的空间膨胀率、加快查询速度。

停用词典中的词汇可以是字母、标点符号、高频词等，每个词汇的最大长度为 48 个字节。

停用词典的使用要点：

- 在创建数据库时引用停用词典。
- 一般不对系统缺省停用词典进行维护，当需要过滤更多的无意义词时，可通过创建附加停用词典来实现。
- 无论是何种语言的数据库，均可引用停用词典。
- 并不是语言中的所有的高频词都需要作为停用词，例如“家”、“世界”是高频词，但对大多数社会科学资料数据库来说，它们可能是重要的词汇。一个词是不是需

要作为停用词，与数据库的领域特点有关，如在计算机科学文献中，“计算机”可作为停用词，因为它几乎出现在每一篇文献中，没有检索意义。

2.4 附加停用词典

附加停用词典是停用词典的补充。TRS 提供的缺省停用词典是一部通用的停用词典，不包括特殊领域的无检索意义的词汇。在这种情况下，一般不是对系统缺省停用词典进行维护，而是通过建立新的附加停用词典来满足特殊要求。

附加停用词典中的词汇可以是字母、标点符号、高频词等。每个词汇的最大长度为 48 个字节。

附加停用词典的使用要点：

- 在创建数据库时引用附加停用词典。
- 无论是何种语言的数据库，均可引用附加停用词典。
- 附加停用词典一般包含了某个领域的高频词。

2.5 稀疏词典

在某些应用领域中，用户并不需要进行全文检索，而希望从文献中自动抽取一些词汇作为检索的关键词，也就是说并不需要索引文献中的所有词汇，而只需索引其中一些关键词。稀疏词典就是存储这些关键词汇的词典。

关键词可以是任何词汇，包括由字母和数字组成的词汇。每个词汇的最大长度为 48 个字节。

稀疏词典的使用要点：

- 在创建数据库时引用稀疏词典。稀疏词典被引用后，一般不允许再对该词典进行维护，除非重新创建数据库的索引。
- 稀疏词典一般单独使用。
- 稀疏词典中的词汇既可以是任何词汇，包括由字母和数字组成的词汇。

2.6 主题词典

主题词典又称为叙词表，它是一种语义词典，由词及其各种关系组成，能反映某学科领域的语义相关概念。

TRS 全文数据库系统支持 ANSI Thesaurus 标准 (Z39.19-1980) 所规定的全部 13 种词间关系，即：族首词、上位词、下位词、等同词、替代词（用代词）、相关词、缩略词、被缩略词、组合概念、历史注释、范围注释、外文等同词、后组配概念。这些关系完全包括了我国《汉语主题词表》的“用、代、属、分、参”结构。

下表描述了 ANSI 主题词表的结构：

关系	说明	举例
族首词 (Lead Term)	概念上外延最广的主题词	“木刻”的族首词是“美术”
上位词 (Broader Term)	概念上外延较广的主题词	“木刻”的上位词是“版画”
下位词 (Narrower Term)	概念上内涵更窄的主题词	“木刻”的下位词有“水印木刻”和“套色木刻”
等同词 (Equivalence, Preferred Term)	也称正式主题词或规范主题词	“电子计算机”是“电脑”的等同词
替代词(用代词) (Used For, Non-Preferred Term)	也称非正式主题词或非规范主题词	“电脑”、“计算机”是“电子计算机”的替代词
相关词 (Related Term)	概念相近但不是等关系的主题词	“计算机应用”和“自动控制”是相关关系
缩略词 (Abbreviation)	主题词的缩略词	“Database Management System”的缩略词为“DBMS”
被缩略词 (Abbreviation For)	主题词被缩略之前的正式名称	“DBMS”的正式名称为“Database Management System”
组合概念 (Combined Concepts)	为将来 TRS 版本预留	
历史注释 (History Note)	对主题词的历史进行说明	
范围注释 (Scope Note)	对主题词的适用范围进行说明	
外文等同词 (Language Equivalent)	其它语种的等同词	“计算机”的英文等同词是“Computer”
后组配概念 (Post-coordinated Concept)	为将来 TRS 版本预留	

图 1-1 是一个“下位词”网络的局部示意图：

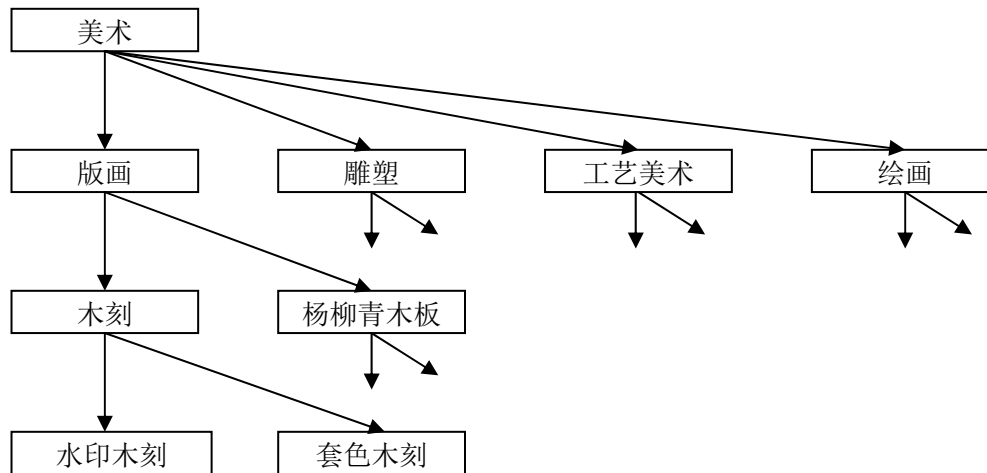


图 1-1

主题词典的使用要点：

- 通过浏览主题词典进行词汇的交互式选择完成检索。
- 利用主题词典函数，或自动扩展功能进行多语种和智能化概念检索。此时必须保持词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。

在 TRS 全文数据库系统中，主题词典是一种特殊的数据库，其最大规模与一般的数据库是一致的。主题词典主要用于检索时的后控制和标引时的自动或辅助选择索引词，是提高查全率和查准率、实现多语种检索和智能化概念检索的重要途径。为此 TRS 全文数据库系统定义了与主题词关系相对应的函数以使能在检索表达式中引用词典中的各种关系。下表是与主题词典有关的检索函数：

函数名	含义和作用
BT(主题词[,词典])	获得某个主题词的上位词
NT(主题词[,词典])	获得某个主题词的下位词
PT(主题词[,词典])	获得某个主题词的等同词
UF(主题词[,词典])	获得某个主题词的替代词
AB(主题词[,词典])	获得某个主题词的缩略词
ABF(主题词[,词典])	获得某个主题词的被缩略词
RT(主题词[,词典])	获得某个主题词的相关词
LE(主题词[,词典])	获得某个主题词的外文等同词
ALLBT(主题词[,词典])	获得某个检索词的所有层次的上位词
ALLNT(主题词[,词典])	获得某个检索词的所有层次的下位词

2.7 同义词典

同义词典是对主题词典中替代关系的扩充，以实现范围更大和意义更广的同义词扩展检索，以提高查全率，是 TRS 全文数据库系统完成智能化检索的又一种手段。比如：假设在同义词典中“计算机”的同义词为“电脑”，那么在使用该同义词典对“计算机”进行扩展检索时，所有包含“计算机”或“电脑”的文献将被检索出来。

在主题词典中，有一些关系（如替代、缩略等）也具有“同义词”的含义，这些关系与同义词典的区别在于：

1. 主题词典中的主题词及其关系是经过严格控制的，一般不能随意地加以变动。而同义词典则没有这么严格，可以比较随意地进行增加、删除和修改等维护。
2. 主题词典中一般只收录概念词，不收录或很少收录专用名词，如人名、机构名等，这些词应在同义词典中定义。

同义词典的使用要点：

- 通过浏览同义词典进行词汇的交互式选择完成检索。
- 利用同义词典函数，或自动扩展功能进行智能化概念检索。此时必须保持词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。

在 TRS 全文数据库系统中，同义词典是一种特殊的数据库，所以其规模能够达到一般数据库的所能达到的规模。同义词典的内部结构非常简单，一般不需要了解。为了便于引用同义词典，TRS 全文数据库系统定义了一个同义词典函数，通过该函数可在指定的同义词典中获取给定词汇的同义词集。下表是与同义词典有关的检索函数：

函数名	含义和作用
ST(词汇[,词典])	获得某个词的同义词

2.8 反义词典

反义词典主要用来进行反义词的联想检索，也是 TRS 全文数据库系统实现智能化检索的手段之一。比如：假设在反义词典中“拥护”的反义词为“反对”，那么在使用该反义词典对“拥护”进行扩展检索时，所有包含“拥护”或“反对”的文献将被检索出来。

反义词典的使用要点：

- 通过浏览反义词典进行词汇的交互式选择完成检索。

- 利用反义词典函数，或自动扩展功能进行智能化概念检索。此时必须保持词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。

在 TRS 全文数据库系统中，反义词典是一种特殊的数据库，所以其规模能够达到一般数据库的所能达到的规模。反义词典的内部结构非常简单，一般不需要了解。为了便于引用反义词典，TRS 全文数据库系统定义了一个反义词典函数，通过该函数可在指定的反义词典中获取给定词汇的反义词集。下表是与反义词典有关的检索函数：

函数名	含义和作用
AT(词汇[,词典])	获得某个词的反义词

2.9 实体词典

实体词典是 TRS 全文数据库系统新增加的一种词典^[修订 145]，一个系统只允许一部实体词典，并且只能在系统安装或升级时由安装程序或升级程序自动创建，外部工具和接口不能创建和删除，但允许维护其内容。

与其他词典一样，实体词典也是由一系列词条组成的，词条一般包括人名、组织机构名称及简称、专有名词等。词条中不能含有标点符号。

实体词典的使用要点：

- 实体词典不分字符集，可在各字符集之间通用。
- 实体词典主要用于提升检索的相关性效果，由检索函数 LIKE 和 INCLUDE 内部引用，不需要专门的函数显式调用。检索时，实体词将以“整体”的形式参与检索，数据中“被分隔”出现的形式不会被命中。
- 实体词的长度限制比较宽松，但不建议使用很长的实体词。
- 门牌地址等一般不建议作为实体词。
- 无“专指”意义的词组不建议作为实体词。

2.10 其它词典

在 2.1 至 2.5 节中，已经分别说明了分词词典、附加分词词典、停用词典、附加停用词典以及稀疏词典等五种索引词典。实际上，在 TRS 全文数据库系统中还有另外三种索引词典，即歧义词典、二元组词典（高频词词典）和单字词典，统称为辅助索引词典。它们是系统的固定资源，具有唯一性，用户不可维护，只能进行系统级的维护。

歧义词典是用于按词索引时决定分词歧义串的正确分词方法的词典。歧义词典一旦改变，所有使用按词索引策略（见 3.3 节）的数据库必须重新建立索引。

二元组词典是用来决定哪些字、词（或字串）需要与其前后相邻的字、词（或字串）进行二元组合的词典。对于按词和按字索引，TRS 使用不同的二元组词典。当按词二元组词典被改变时，所有使用按词的二元组索引策略（见 3.3.2 节）的数据库必须重新建立索引。当按字二元组词典被改变时，所有使用按字的二元组索引策略的数据库必须重新建立索引。

单字词典是用来决定按字索引时，哪些字需要单独建立索引的词典。在 3.3.2 将会说明：在缺省情况下，除了那些存储在单字词典中的具有单独查询意义的字外，TRS 按字索引时排除了其它单个字的索引，也就是说，在查询单个字时，如果这个字在单字词典中，则可能会有查询结果，否则就肯定没有查询结果。单字词典一旦改变，所有使用按字索引策略的数据库都必须重建索引。如果一个数据库引用了系统中“全字词典”，则单字词典将不起作用。

第3章 字段与索引

索引是实现快速检索的基础，所有对数据库的查询操作都是在相应的数据库索引上进行的。在 TRS 全文数据库系统中，作为检索入口的数据库字段必须建立一定的索引，否则该字段上的值不能被检索。

3.1 字段类型

正如前章所述，TRS 数据库有 6 种不同类型的字段，它们是：

1. DATE 日期型字段
2. NUMBER 数值型字段
3. CHAR 字符型字段
4. PHRASE 短语型字段
5. DOCUMENT 全文型字段
6. BIT 二进制型字段

3.1.1 日期型字段

日期型字段只能用来存储“年月日时分秒”这类表示日期的数据。这类字段的主要特点是：

- 允许设置缺省值。
- 允许限制取值范围。
- 允许多值。
- 允许加载时字段值的唯一性检查。
- 允许加载时空值的检查。
- 允许建立索引。
- 不支持倒排索引属性。
- 字段值必须是一个有效的日期。
- 数据集中存储。
- 支持对字段值的排序。

3.1.2 数值型字段

数值型字段只能用来存储整数或实数。对于实数，可以限制其小数位的长度，在

数据库加载时将忽略多余的小数位。整数或实数不能超出机器的单精度浮点数所能表示的范围，可以使用科学计数法表示这些数据。这类字段的主要特点是：

- 允许设置小数位的精度。
- 允许设置缺省值。
- 允许限制取值范围。
- 允许多值。
- 允许加载时字段值的唯一性检查。
- 允许加载时空值的检查。
- 允许建立索引。
- 不支持倒排索引属性。
- 字段值必须是一个有效的整数或实数。
- 数据集中存储。
- 支持对字段值的排序和统计运算。

3.1.3 字符型字段

字符型字段可以用来存储任何形式的字符串。单个字符串的最大长度可以在建库时指定，但最大不能超过 255 个字符，在数据库加载时将截断超过规定长度的字符串，而只保留字符串的前部。这类字段的主要特点是：

- 允许设置字符串的最大长度。
- 允许设置缺省值。
- 允许限制取值范围。
- 允许多值。
- 允许加载时字段值的唯一性检查。
- 允许加载时空值的检查。
- 允许建立索引。
- 不支持倒排索引属性。
- 单个字段值是一个长度不超过 255 个字符的任意字符串。
- 数据集中存储。
- 支持对字段值的排序。如果是多值，则只对第一个字段值进行排序。

3.1.4 短语型字段

短语型字段可以用来存储任何形式、任意长度的文本数据，最适合作为“标题”这样的字段。短语型字段与字符型字段的最大差别在于短语型字段支持自动分词，而字符型字段不支持自动分词，其字段值整体是一个索引键值。这类字段的主要特点是：

- 没有缺省值。
- 没有取值范围。

- 不允许多值。
- 加载时不进行字段值的唯一性检查。
- 加载时允许空值。
- 允许建立索引。
- 支持自动分词，倒排索引属性固定（只有位置索引属性中“位置”）。
- 字段值必须是文本数据。
- 数据集中存储。
- 支持对字段值的排序。进行排序时，只有字段值的前 256 个字符是有效的。

3.1.5 全文型字段

全文型字段既可以用来存储任何形式、任意长度的文本数据，也可以用来存储其它格式类型的多媒体数据。全文型字段区别于短语型字段的关键所在有四点：

1. 全文型字段的数据有多种存储方式。系统通常是集中存储短语型字段的数据，而分类存储全文型字段的数据，这样有助于提高大规模数据库的检索结果的概览速度（概览不包括全文型字段和二进制型字段的内容）。
2. 全文型字段的字段值可以存储一个源格式文档附件。
3. 全文型字段的倒排索引属性可以任意指定。
4. 全文型字段不支持对字段值的排序。

综合起来，这类字段的主要特点是：

- 没有缺省值。
- 没有取值范围。
- 不允许多值，但允许一个源格式文档附件。
- 加载时不进行字段值的唯一性检查。
- 加载时允许空值。
- 允许建立索引。
- 支持自动分词，倒排索引属性可任意定义。
- 字段值可以是多媒体数据。
- 数据有多种存储方式。

3.1.6 二进制型字段

二进制型字段可以用来存储任何格式类型的多媒体数据，主要特点是：

- 没有缺省值。
- 没有取值范围。
- 允许多值。
- 加载时不进行字段值的唯一性检查。

- 加载时允许空值。
- 不建立索引。
- 字段值可以是任何多媒体数据。
- 数据有多种存储方式。

3.2 索引类型

TRS 全文数据库系统有两种类型的数据库索引： B^+ 树索引和倒排（INVERT）索引。这两种类型的索引相辅相成，互为补充。

B^+ 树索引的突出特点是很好地解决了键值的插入、溢出、删除和空间回收等问题，从而使 B^+ 树索引可以适应成批插入、文件易变的情况； B^+ 树索引在操作中可“动态地”进行维护，无须周期性地重新组织文件；还可以采用压缩索引项的办法，提高扇出率，降低树的高度，减少读取次数，加快查找速度。因此 B^+ 树索引是组织大型索引文件的一种最主要的方法。

B^+ 树索引实现了索引键值的唯一化，即一个键值在 B^+ 树索引中只有一个唯一的入口，而倒排索引则用来存储索引键值在同一个记录和不同记录中出现的情况。倒排索引的建立大大提高了重复键值的查询速度和空间使用率，同时也简化了 B^+ 树索引的维护操作。

倒排索引由一系列倒排索引项组成，每个倒排索引项又可包含多个倒排索引属性，但不同的字段类型，所支持的倒排索引属性也不尽相同：

- DATE、NUMBER 和 CHAR 字段不支持任何倒排索引属性。
- PHRASE 字段上所支持的倒排索引属性是固定的，由系统自动定义。
- DOCUMENT 字段上的倒排索引属性则可在创建数据库时根据实际需要由用户自行定义。

3.3 索引策略

对于字母文字，TRS 全文数据库系统总是按“单词”进行索引，但对于表意文字（中文、日文、韩文）TRS 全文数据库系统则有四种类型的索引策略：按词检索的索引策略、按字检索的索引策略、按用户自定义关键词的索引策略，以及二元组索引策略。在实际应用中，采用何种索引策略主要取决于数据的特点、检索的需求和系统的资源等。不同的策略，所需的索引词典也不尽相同。

索引策略与词典的关系:

策略	稀疏词典	分词词典	附加分词词典	停用词典	附加停用词典
按词	置空	必需	可有可无	可有可无	可有可无
按字	置空	置空或全字词典	忽略	可有可无	可有可无
关键词	必需	忽略	忽略	忽略	忽略
二元组	置空	可有可无	可有可无	可有可无	可有可无

3.3.1 按词索引策略

TRS 内嵌的自动分词系统是 TRS 的领先技术,使得 TRS 在国内外同类全文检索系统中独具特色。对大多数全文检索应用来说,按词索引是首选的索引策略。

按词索引策略的优越性在于:

- 检索的准确率即查准率高。如果采用按词索引策略,则在检索“华人”时,那些仅包含“中华人民共和国”的记录不会被检索出来。
- 查询的速度快。测试表明,在大型数据库上,按词索引的查询速度要比按单字索引(不进行二元索引)的查询速度快 5-10 倍。
- 空间膨胀率低。由于一个词往往包含了多个字,这样按词索引就大大减少了索引量,再加上停用词典的使用,索引文件的空间膨胀率比按字索引的小 50-100%。
- 按词索引是实现智能概念检索的基础。由于所有辅助知识词典(如主题词典、同义词典、反义词典等)都是基于词的,而不是基于字的,所以只有按词索引,才能进行概念检索。

按词索引策略的不足在于:

- 词典维护后数据库一般需要重新进行索引。
- 由于表意文字的词与词之间没有语法上的分隔符,即使自动分词的准确率再高,也不可能达到 100%的正确,再加上自然语言本身的复杂性,有些歧义人工也难以在单个句子中进行正确分词(必须根据上下文句子的语义才能判断),因此对检索会有一定的负面影响。

使用按词索引策略的方法是:为数据库指定相应的分词词典(全字词典除外)、附加分词词典,或者一部稀疏词典。

3.3.2 按字索引策略

如果数据是文言文或数据库规模很小,可采用按字索引策略。按字索引策略的主要优点是:

- 不需要自动分词系统，实现相对简单。
- 没有自动分词的错误，查全率较高。

然而，按字索引策略的不足也是显而易见的，按词索引策略的所有优点均是按字索引策略的致命缺点，所以除非需要百分之百的查全率，而不惜牺牲查准率，一般情况下不建议使用按字索引策略。由于传统意义上的按字索引策略的查询速度太慢，这里的按字索引实际上是一种二元组索引（参见 3.3.4 节）。

使用“不需要检索单字”的按字索引策略的方法是：

- 方法 1. 数据库不指定任何分词词典、附加分词词典和稀疏词典。
- 方法 2. 在数据库使用按词索引策略的前提下，将需要按字索引的字段的混合索引模式指定为“禁止词模式”（参见 1.4 节）。

使用“允许检索单字”的按字索引策略的方法是：

- 方法 1. 为数据库指定分词词典为全字词典。
- 方法 2. 在数据库使用按词索引策略的前提下，将需要按字索引的字段的混合索引模式指定为“禁止词模式”，并开启该字段的全字索引开关（参见 1.4 节）。

3.3.3 关键词索引策略

按关键词索引实际上是一种抽词索引技术，是在按词索引策略的基础上只抽取其中的某些词进行索引。所有需要索引的关键词构成一部稀疏词典。一般对专用领域的文献可采用这种索引策略。

按关键词索引策略兼有按词索引策略的优点，并且具有更低的空间膨胀率和更快的检索速度。但由于按关键词索引策略只对“关键词”进行了索引，这样就不能象按词索引策略那样可以用任意的自由词进行全文检索，而只能对“关键词”进行检索。

3.3.4 二元组索引策略

二元组索引策略是指在按词索引时对高频词进行二元组合，或者在按字索引时对高频字串进行二元组合。有分词词典或附加分词词典的二元组索引称为按词二元组索引；反之，如果既没有分词词典，也没有附加分词词典，则称为按字二元组索引。由此可见二元组索引策略并不是一种独立的索引策略，它是按词索引策略和按字索引策略的扩展。二元组索引还依赖于二元组词典（高频词词典），不同的数据库语言对应不同字符集的词典，如果数据库语言与系统主字符集不一致，则必须通过一次与数据库语言一致的系统升级来自动创建所需要的词典。

二元组索引策略的最大优势是在大数据库中对含有高频词（字串）的短语的查询时，速度有显著提高。但这种索引策略的缺点也同样明显：一是空间膨胀率较高，尤其是对于小数据库；二是装库速度要比单独的按词或按字索引策略慢得多。对于小数据库，即使没有二元组索引，其查询速度也很快，所以完全没有必要使用这种策略。

特别地，按字二元组索引有两个级别：字之间的二元组索引（一级二元索引）和高频字串（即一级二元组）与字之间的二元组索引（二级二元索引，实际上是三元索引，但由于不是所有的字都进行三元索引，所以这里只把它看成是二元索引的一种优化扩展）。第一级按字二元组索引策略就是按字索引策略（因此，并不是所有的字都能单独查询，如果需要能够检索所有的单字，可以给数据库指定系统中的“所有汉字词典”作为分词词典），而这里所说的按字二元组索引就是指第二级按字二元组索引策略。二元组索引策略的选择不是通过词典的选取来决定的，而是通过数据库的二元组索引开关来决定的。

3.4 索引属性

在创建数据库时，可以为一个 DOCUMENT 字段定义索引属性（参见 1.4 节）。索引属性分为两类，即：位置索引属性和格式索引属性。如果索引属性中包含了格式索引属性，则此种索引就称之为**全息索引**。

3.4.1 位置索引属性

位置索引属性是 TRS 的传统索引属性，用来标明一个索引项在文本数据中的自然位置，通常包括：所在的段落；所在段落中的句子；所在句子中的位置。位置索引属性隐含在文本中，不需要显式的标注，是通过分词系统自动提取的，都是从 0 开始起始计数。

位置索引属性的目的之一就是位置限定运算，如在检索“国际大师”时，“国际艺术大师”中“国际”与“大师”不会命中，因为这里隐含了“国际”与“大师”在位置上必须相邻的限定要求。这只是其中的一种隐式的位置限定运算，位置限定运算还可以通过相关的运算符显式地指定（参见 6.2 节）。位置索引属性的另一个目的是为检索结果提供准确的命中点。

TRS 全文数据库系统只支持三种位置索引属性，即：SEG（段落）、SEN（句子）和 POS（位置）。这些属性都有各自的生成规则。

- “段落”的提取规则是：换行符后面如果还是换行符，或者是空白符，则认为是一个段落的结束。

- “句子”的提取规则是：当遇到“句号”、“分号”、“问号”或者“叹号”，或者段落结束时，认为是一个句子的结束。^{【修订 140】}在以前的系统里，“逗号”也是一个句子的结束，但现在已不再是句子的结束。^{【1】}
- “位置”的提取规则是：一个表意字符占一个位置；一个字母单词占一个位置；一个数字串占一个位置；所有的空白符被忽略；其它符号占一个位置。

对于句子，TRS 全文数据库系统还支持对文本数据按标志进行人工切分句子，即在文本中加入特定的标志来告知分词系统进行分句，同时忽略上述以标点符号为依据的自然分句规则。具体规则是：

1. 在文本的起始位置添加一个由三字节组成的“人工分句标志”序列，分别为十六进制的“1B”、“21”和“08”（对应键盘的“ESC”、“!”和“Backspace”键的 ASCII 码）。这个标志必须出现在文本的起始位置，其它位置的出现不起作用，它表示文本的分句要依据标志进行切分。
2. 在文本里需要分句的位置插入由两个字节组成的“分句标志”序列，分别为十六进制的“21”和“08”（对应键盘的“!”和“Backspace”键的 ASCII 码）。如果在文本的起始位置没有“人工分句标志”，则“分句标志”即使在文中出现，也不会起作用，仍然还是采用自然分句规则进行分句。^{【修订 086】}

在位置索引属性中，除 POS 外，SEG 和 SEN 都不是必须的，这样可能的组合有：<SEG, SEN, POS>、<SEG, POS>、<SEN, POS>和<POS>。在存储的索引中，SEG、SEN 和 POS 都用两个字节表示，对于前三种组合基本足够，但在只有 POS 属性的情况下，对于超长文本则会丢失一些索引信息，从而导致漏检，因此增加一种位置属性 POS4，表示“位置”用四字节存储，但 POS4 与 SEG 或 SEN 组合时等价于 POS。

必须特别注意的是：对于一些不规则的文本，一个“字片”可能会被“段落”规则所打断，无论是采用按“词”索引还是采用按“字”索引，都无法检索到，从而影响查全率。^{【修订 025】}要使这样的“字片”也能够检索到，其唯一方法是位置索引属性不包含 SEG 和 SEN。当没有定义 SEG 索引属性时，如果定义了 SEN 索引属性，则“段落”的结束将按“句子”的结束处理，如果 SEN 索引属性也没有定义，则“段落”的结束将被按空白处理。^{【1】}

3.4.2 格式索引属性

格式索引属性是用来标明一个索引项在格式文本中除位置索引属性外的其它所有属性。格式索引属性对纯文本没有意义，必须进行显式地标注，也就是说格式索引属性只对格式化的文档有用，如：XML、HTML、RTF 格式等。

格式索引属性的唯一目的就是在检索时用来限定检索目标，以保证所命中的检索

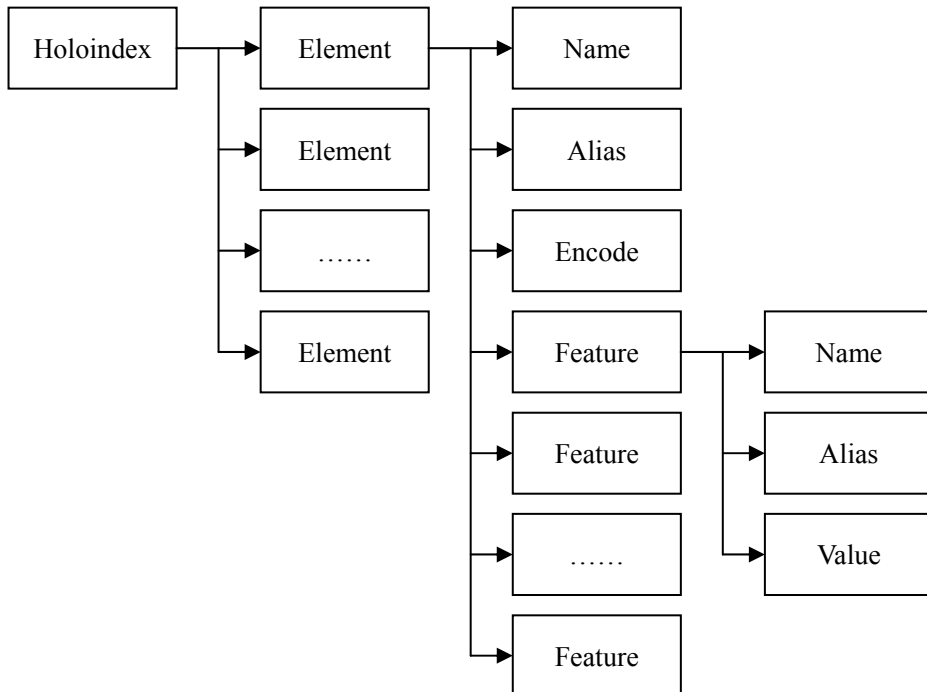
结果能够满足用户对检索目标所具有的格式要求。

与位置索引属性的提取需要分词系统一样，格式索引属性的提取也需要有相应格式文档的分析器，如：对 XML 文档，需要 XML 分析器；对 RTF 文档，需要 RTF 分析器。目前系统只提供了 XML 和 HTML 分析器，因此 TRS 只支持对 XML 和 HTML 文档进行全息索引。

3.4.3 定义索引属性

位置索引属性的名称和数量都是固定的，格式索引属性则不同，没有固定的名称和数量，必须由用户在创建数据库时按照规范自行定义。

格式索引属性的拓扑结构为：



格式索引属性的定义则要采用类 XML 文档结构，如下所示：

```

<Holoindex Structure = BITS | BRIEF | ENTIRE Sensitive = TRUE | FALSE>
  <Element Rank = ID>
    <Name> NAME </Name>
    <Alias> ALIAS </Alias>
    <Encode> BOOLEAN | SEQUENCE | COUNTER | FEATURE </Encode>
    <Feature Rank = ID>
  
```

```
<Name> NAME </Name>
<Alias> ALIAS </Alias>
<Value> SYMBOL | SENSITIVE | DIGITAL </Value>
</Feature>
<Feature>
.....
</Feature>
</Element>
<Element>
.....
</Element>
</Holoindex>
```

Holoindex 是该结构的类型元素,该元素的内容就是一个完整的“格式索引属性集”定义。在 Holoindex 元素下,可以用多个 Element 元素定义所有需要建立索引的“格式索引属性”。Holoindex 元素允许定义下列属性:

- Structure 属性。
 - ◇ 取值 BITS 表示:所有 Element 元素定义的“格式索引属性”都将采用 BOOLEAN 型编码,且“格式索引属性”的总数不超过 32;
 - ◇ 取值 BRIEF 表示:如果在所有 Element 元素定义的“格式索引属性”之间具有嵌套层次,则只有处在最里层的“格式索引属性”才进行索引,而忽略所有外层“格式索引属性”,检索时不能同时使用多个“格式索引属性”同时进行限定;
 - ◇ 取值 ENTIRE 表示:采用完全索引存储结构,这种存储结构支持层次检索的要求。

缺省取 ENTIRE 值。

- Sensitive 属性。
 - ◇ 取值 TRUE 表示:该结构所定义的“格式索引属性”的名字和别名,以及“格式索引属性”特征的名字和别名是区分大小写的。
 - ◇ 取值 FALSE 表示:该结构所定义的“格式索引属性”的名字和别名,以及“格式索引属性”特征的名字和别名是不区分大小写的。

缺省取 TRUE 值。

Element 元素用来定义和描述“格式索引属性”的索引方式。一个 Element 元素只能定义一个“格式索引属性”。(Element 元素的 Rank 属性仅供内部使用)

- Name 元素用来指定“格式索引属性”的名字。这个名字必须是格式化文档中用以进行格式修饰的真实“格式”的名称。对于 XML 文档来说,“格式”就是标记。

- Alias 元素用来定义“格式索引属性”的别名，别名之间不能重名，也不能与系统内部的特定限定运算符重名，并且必须遵循 TRS 全文数据库系统的命名规则（参见第1章）。在通常情况下，“格式索引属性”的名字就可作为属性限定运算符，而无须定义 Alias 元素，但有两种情况必须定义：

- ◇ 全文检索系统本身有固定的位置索引属性 SEG、SEN、POS，另外还规定了几种特定的限定运算符，如 TXT、TIT、FREC、FRE、REC 等。如果一个“格式索引属性”的名字恰好与这些限定运算符同名，则就会发生冲突，而不能作为限定运算符使用。
- ◇ 为了便于检索表达式的书写与分析，尽量避免歧义的产生，TRS 全文数据库系统要求限定运算符的命名必须遵循其对象的命名规则，当“格式索引属性”的名字不符合这一规则时，也不能作为限定运算符使用。

Alias 元素一旦定义，在检索时将只允许使用这个别名，而“格式索引属性”的名字则只在创建文档的索引时使用。

- Encode 元素用来指定“格式索引属性”的编码方式：

- ◇ BOOLEAN 型编码。在检索时，如果只要求验证该“格式索引属性”是否修饰了检索对象，而总是不要求验证它的其它任何特征，则该“格式索引属性”就可使用 BOOLEAN 型编码。
- ◇ SEQUENCE 型编码。如果某些或全部“格式索引属性”（包含该“格式索引属性”）在修饰目标文本时，位置上具有层次树状结构，在检索时，不但要求验证该“格式索引属性”是否修饰了检索对象的同时，而且要求验证在同一个父结点下它重复出现的次序（从 0 开始记数），则该“格式索引属性”就可使用 SEQUENCE 型编码。检索时允许比较运算。
- ◇ COUNTER 型编码。在检索时，如果只要求验证该“格式索引属性”在检索目标出现之前，相对于整个文档出现的次数（从 0 开始记数），则该“格式索引属性”就可使用 COUNTER 型编码。如分页标记。检索时允许比较运算。
- ◇ FEATURE 型编码。在检索时，如果要求验证该“格式索引属性”是否修饰了检索对象的同时，还要求验证它是否具有指定的特征值，则该“格式索引属性”必须使用 FEATURE 型编码。这种编码方式需要用 Feature 元素来进一步描述需要索引的“格式索引属性”的特征。

缺省为 BOOLEAN 型编码。

- Feature 元素用来描述在 FEATURE 型编码方式下（其他编码方式将忽略所有 Feature 元素），该“格式索引属性”所有具有检索需求的特征，而其它没有检索需求的特征将被忽略。一个 Feature 元素只能定义一个“格式索引属性”特征，多个“格式索引属性”特征需要用多个 Feature 元素来定义。需要注意的是：如果 Holoindex 元素的 Structure 属性取值为 BRIEF，则只允许定义一个特征。（Feature 元素的 Rank 属性仅供内部使用）
- ◇ Name 元素用来指定“格式索引属性”特征的名字。这个名字必须是格

式化文档中用以进行格式修饰的真实“格式”的属性的名称。对于 XML 文档来说，“格式索引属性”特征就是标记的属性。

- ◇ **Alias** 元素用来定义“格式索引属性”特征的别名，特征的别名之间不能重名，且必须遵循 TRS 全文数据库系统的命名规则。在通常情况下，“格式索引属性”特征的名字就可以在属性/限定运算符中用以给定“格式索引属性”的具体特征，而无须定义 **Alias**，但当“格式索引属性”特征的名字不符合 TRS 全文数据库系统的命名规则时，必须定义 **Alias** 元素。特征的别名一旦定义，在检索时将只允许使用这个别名，而“格式索引属性”特征的名字则只在创建文档的索引时使用。
- ◇ **Value** 元素用来指定“格式索引属性”特征值的类型：
 - **SYMBOL** 型特征值。这种类型的特征值可以是不区分大小写的任意字符串，系统内部对其进行自动编码，检索时只能进行“相等”或“不等”运算，而不能使用其它比较运算。
 - **SENSITIVE** 型特征值。与 **SYMBOL** 型唯一不同的是 **SENSITIVE** 型特征值是区分大小写的。
 - **DIGITAL** 型特征值。这种属性值必须是一个数字串，其数值就是特征值编码。检索时允许所有比较运算。

缺省为 SENSITIVE 型。

将位置索引属性与格式索引属性放在一起就构成了全息索引属性，当定义较长时可以考虑存放在文件中。位置索引属性和格式索引属性放置顺序没有特别要求，但一般将位置索引属性放在前面，把格式索引属性当成一个整体，各个索引属性之间用单字节的分号“;”分隔。

下面是一个全息索引属性定义的例子：

SEG;SEN;POS;

<Holoindex Structure = ENTIRE Sensitive = TRUE>

<Element>

<Name> claims-item </Name>

<Alias> 要求项 </Alias>

<Encode> FEATURE </Encode>

<Feature>

<Name> id </Name>

<Value> SENSITIVE </Value>

</Feature>

</Element>

<Element>


```
<Name> paragraph </Name>
<Alias> 段落 </Alias>
<Encode> SEQUENCE </Encode>
</Element>
<Element>
  <Name> dp </Name>
  <Alias> 页次 </Alias>
  <Encode> COUNTER </Encode>
</Element>
<Element>
  <Name> claims </Name>
  <Alias> 要求书 </Alias>
  <Encode> BOOLEAN </Encode>
</Element>
<Element>
  <Name> national-description </Name>
  <Alias> 说明书 </Alias>
  <Encode> BOOLEAN </Encode>
</Element>
<Element>
  <Name> title-of-invention </Name>
  <Alias> 专利标题 </Alias>
  <Encode> BOOLEAN </Encode>
</Element>
<Element>
  <Name> description </Name>
  <Alias> 描述 </Alias>
  <Encode> BOOLEAN </Encode>
</Element>
</Holoindex>
```

第 4 章 权限控制

TRS 全文数据库系统的安全控制是通过权限管理来实现的。

4.1 权限级别

TRS 全文数据库系统的权限有四个级别，即：用户级权限、数据库级权限、字段级权限和记录级权限。

4.1.1 用户级权限

用户级权限对应于用户的类型。TRS 全文数据库系统有五种类型的用户级权限：

- 超级用户(SYSTEM)。
- 系统数据库管理员(DBA)。
- 用户组数据库管理员(GDA)。
- 系统资源用户(RESOURCE)。
- 系统登录用户(CONNECT)。

超级用户所拥有的权限包括：

- 维护系统配置。
- 创建和注销用户组，维护所有用户组的属性。
- 创建和注销用户，维护所有用户的属性。
- 创建数据格式，维护数据格式的属性。
- 蕴涵系统数据库管理员的所有权限。

系统数据库管理员所拥有的权限包括：

- 不经授权就可对系统中的所有数据库、视图，以及词典对象进行所有操作。
- 蕴涵用户组数据库管理员的所有权限。

用户组数据库管理员所拥有的权限包括：

- 不经授权就可对系统中与该用户处在同一用户组的用户所拥有的所有数据库、视图，以及词典对象进行所有操作。
- 蕴涵系统资源用户的所有权限。

系统资源用户所拥有的权限包括：

- 有权在系统中创建数据库、视图，以及词典对象。
- 蕴涵系统登录用户的所有权限。

系统登录用户所拥有的权限包括：

- 有权登录到系统中。
- 不经授权就可对系统中自己所拥有的数据库、视图，以及词典对象进行所有操作。
- 经过授权，可对系统中其他用户所拥有的数据库、视图，以及词典对象进行相应的操作。

4.1.2 数据库级权限

数据库权限、视图权限和词典权限统称数据库级权限。

数据库权限共有五种：

- 记录检索权(SELECT)。拥有这种权限的用户，能够对数据库进行检索操作。
- 数据更新权(UPDATE)。拥有这种权限的用户，能够对数据库进行数据加载、记录插入、修改和删除操作。
- 数据索引权(INDEX)。拥有这种权限的用户，能够对数据库的索引进行创建、删除、验证和优化操作。
- 数据库结构修改权(ALTER)。拥有这种权限的用户，能够对数据库的属性及结构进行修改操作，包括增加和删除字段，以及修改字段的属性。
- 数据库删除权(DROP)。拥有这种权限的用户，能够删除数据库。

这五种数据库权限的蕴涵关系是：ALTER 和 DROP 蕴涵 INDEX，INDEX 蕴涵 UPDATE，UPDATE 蕴涵 SELECT。当收回被蕴涵的权限时，系统将自动收回蕴涵该权限的所有权限，如收回 SELECT 权限，则实际将收回所有的权限。需要特别说明的是：虽然数据库属性中的公共访问权限也有这种蕴涵关系，但去掉其中的某个公共访问权限时，不会自动去掉蕴涵该公共访问权限的公共访问权限。

视图权限共有三种：

- 记录检索权(SELECT)。拥有这种权限的用户，能够对视图进行检索操作。
- 视图结构修改权(ALTER)。拥有这种权限的用户，能够对视图的属性及结构进行修改操作，包括增加和删除字段，以及修改字段的属性。
- 视图删除权(DROP)。拥有这种权限的用户，能够删除视图。

这三种视图权限的蕴涵关系是：ALTER 和 DROP 蕴涵 SELECT。当收回 SELECT 权限时，系统将自动收回所有的权限。需要特别说明的是：虽然视图属

性中的公共访问权限也有这种蕴涵关系，但去掉其中的某个公共访问权限时，不会自动去掉蕴涵该公共访问权限的公共访问权限。

自动模式视图权限共有五种：

- 记录检索权(SELECT)。拥有这种权限的用户，能够对自动模式视图进行检索操作。
- 数据更新权(UPDATE)。拥有这种权限的用户，能够对自动模式视图进行数据库加载、记录插入、修改和删除操作。
- 数据索引权(INDEX)。拥有这种权限的用户，能够对自动模式视图的索引进行创建、删除、验证和优化操作。
- 视图属性修改权(ALTER)。拥有这种权限的用户，能够对自动模式视图及其字段的属性进行修改操作，但下列各项不允许修改：视图所包含的数据库；视图的创建条件；视图的缺省检索字段；添加新的字段；删除字段；修改字段中所包含的数据库字段。
- 视图删除权(DROP)。拥有这种权限的用户，能够删除自动模式视图。

这五种视图权限的蕴涵关系是：ALTER 和 DROP 蕴涵 INDEX，INDEX 蕴涵 UPDATE，UPDATE 蕴涵 SELECT。当收回被蕴涵的权限时，系统将自动收回蕴涵该权限的所有权限，如收回 SELECT 权限，则实际将收回所有的权限。需要特别说明的是：虽然视图属性中的公共访问权限也有这种蕴涵关系，但去掉其中的某个公共访问权限时，不会自动去掉蕴涵该公共访问权限的公共访问权限。

为了方便系统内部进行权限的管理与验证，自动模式视图所包含的数据库只保留 UPDATE、INDEX、ALTER 和 DROP 数据库级权限，但它们不能进行单独的权限管理（授予与收回），只能通过自动模式视图的 UPDATE、INDEX、ALTER 和 DROP 权限由系统自动授予与收回。这些数据库也不保留其它数据库级及字段级权限。在自动模式视图创建时，继承源数据库以及所有字段的全部权限，而源数据库只保留 UPDATE、INDEX、ALTER 和 DROP 权限。自动模式视图分裂新数据库，或向自动模式视图中添加一个数据库时，新数据库或添加的数据库将继承其 UPDATE、INDEX、ALTER 和 DROP 权限，而添加的数据库及其字段的其它所有权限将被清除。

词典权限共有三种：

- 检索引用权(SELECT)。拥有这种权限的用户，能够对词典进行浏览操作，能够在创建数据库或检索时引用词典。该权限对所有用户总是开放的。
- 更新维护权(UPDATE)。拥有这种权限的用户，能够对词典进行维护操作，包括增加和删除词汇。
- 删除权(DROP)。拥有这种权限的用户，能够删除词典。

词典的这三种权限的蕴涵关系是：DROP 权蕴涵 UPDATE 权，UPDATE 权蕴涵 SELECT 权。当收回被蕴涵的权限时，系统将自动收回蕴涵该权限的权限。需要特别说明的是：虽然词典属性中的公共访问权限也有这种蕴涵关系，但去掉其中的某个公共访问权限时，不会自动去掉蕴涵该公共访问权限的公共访问权限。

4.1.3 字段级权限

字段级权限包括数据库字段权限和视图字段权限。字段级权限只有一种，即：字段检索权(SELECT)，拥有这种权限的用户，能够在该字段上进行检索操作。

4.1.4 记录级权限

TRS 全文数据库系统中的记录级权限是一种逻辑上的权限，没有显式授权。记录级权限有两种形式，即：记录检索权和记录浏览权，分别通过视图和对记录进行加密来实现。数据库和无“记录选取条件”视图中的记录的检索权是不受限制的，记录的检索权只能控制该记录不能被检索，而不能控制记录能被检索。记录的浏览权只控制检索到该记录的用户是否有权浏览其内容，而不能控制用户能否检索到该记录。

用视图实现记录级的检索权限：

1. 在数据库中确定一个能够分类记录的字段。如果没有这样的字段，则在创建数据库时必须增加一个辅助字段，用以分类记录，并在数据库加载时给出记录的分类值。
2. 按记录的分类值，创建一系列视图，每个视图的“记录选取条件”属性是记录分类的检索表达式。
3. 将这些视图或其字段的记录检索权分别授予相应的用户或用户组，同时收回这些用户和用户组在数据库和其字段上的所有权限。

记录的浏览权则是通过加密的方式实现的。TRS 全文数据库系统有三种记录加密方式：

- 密级。通过密级加密的记录，只有当用户的密级值不大于记录的密级值时才允许输出。但用户的密级如果是 0，则也不允许输出记录。请参见 1.2 节的说明。
- 用户组。通过用户组加密的记录，只有当用户所在的用户组是记录所要求的用户组时才允许输出。当该用户组被删除后，加密还是有效，但原用户组中普通用户已无法输出记录，除非更新记录取消用户组加密。
- 口令。通过口令加密的记录，只有当用户提供了正确的口令时才允许输出。

这三种加密方式可组合使用，这时必须满足每种方式的所有条件才允许输出记录。加密方式即可在数据库加载、记录插入和记录更新时与数据一起指定，也可以对已存在的记录进行加密。请参见《数据库加载》部分。

需要说明的是：对记录的加密只有在读取数据库记录时才会起作用，并且根据 TRS 全文数据库系统的权限管理规则，还必须满足下列所有条件：

1. 当前用户不是超级用户。
2. 当前用户没有 DBA 权。
3. 当前用户不是数据库的拥有者。
4. 当前用户没有数据库拥有者所在用户组的 GDA 权。
5. 数据库没有 PUBLIC 检索属性。

4.2 权限管理

权限管理包括权限授予和收回权限两方面的操作。权限授予是指改变用户级权限，或在数据库、视图、词典和字段上授予用户或用户组一定的操作权限。收回权限是指改变用户级权限，或收回用户或用户组在数据库、视图、词典和字段上的操作权限。权限的授予与收回不能管理记录级权限。

用户级权限的授予与收回：

- 只有超级用户才能改变用户级权限。

数据库级权限的授予与收回：

- 超级用户和 DBA 用户可管理系统中所有数据库、视图和词典的操作权限。
- GDA 用户可管理该用户组的用户所拥有的所有数据库、视图和词典的操作权限。
- 任何用户可管理自己所拥有的数据库、视图和词典的操作权限。

字段级权限的授予与收回：

- 超级用户和 DBA 用户可管理系统中所有数据库和视图的字段的操作权限。
- GDA 用户可管理该用户组的用户所拥有的所有数据库和视图的字段的操作权限。
- 任何用户可管理自己所拥有的数据库和视图的字段的操作权限。

4.3 权限检查

权限检查是权限管理的最终手段，是安全控制的必然要求。对 TRS 全文数据库系统中各个对象的操作都必须有相应的权限才能进行。

4.3.1 用户组操作

只有超级用户有权创建、删除和修改用户组。

当删除一个用户组时，该用户组下的所有用户将自动迁移到系统的匿名用户组，并同时删除与该用户组有关的所有权限。

4.3.2 用户操作

只有超级用户有权创建、删除和修改用户。另外，用户“自己”也可修改一些基本属性：

- 登录口令（超级用户只能删除用户的口令，而不能修改）
- 登录互斥开关
- 最大检索历史步数
- 缺省附加分词词典
- 缺省附加停用词典
- 缺省稀疏词典
- 注释信息

当删除一个用户时，该用户所拥有的所有数据库、视图，以及词典对象的拥有者将自动归为 SYSTEM 用户所拥有，并同时删除与该用户有关的所有权限。

4.3.3 数据库操作

- 创建数据库及恢复数据库。如果满足下列条件之一，则有权创建数据库或恢复数据库：
 1. 超级用户。
 2. DBA 用户。
 3. GDA 用户。
 4. RESOURCE 用户。
- 删除数据库。如果满足下列条件之一，则有权删除数据库：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库的拥有者。
 4. 数据库具有公共 DROP 属性。
 5. 数据库拥有者所在用户组的 GDA 用户。
 6. 拥有该数据库的 DROP 权的用户。

当删除一个数据库时，系统将删除所有用户在该数据库上的权限，同时自动修改映射到该数据库的所有视图的结构，如果这时视图中不再包含任何数据库，则系统将自动删除该视图。

- 修改数据库、锁定数据库及转成自动模式视图。如果满足下列条件之一，则有权修改数据库或锁定数据库或转成自动模式视图：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库所有者所在用户组的 GDA 用户。
 4. 数据库的拥有者。并且对于转成自动模式视图的操作，当前用户还必须具有创建视图的权限。
 5. 数据库具有公共 ALTER 属性。并且对于转成自动模式视图的操作，当前用户还必须具有创建视图的权限。
 6. 拥有该数据库的 ALTER 权的用户。并且对于转成自动模式视图的操作，当前用户还必须具有创建视图的权限。

当删除一个数据库字段时，系统将删除所有用户在该数据库字段上的权限，同时自动修改映射到该数据库字段的所有视图的字段结构，如果这时视图字段中不再包含任何数据库字段，则系统将自动删除该视图字段，当视图中不再包含任何字段时，也将被自动删除。

- 解锁数据库。如果满足下列条件之一，则有权解锁数据库：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库所有者所在用户组的 GDA 用户。
 4. 数据库锁的拥有者，并且数据库具有公共 ALTER 属性或者操作者拥有该数据库的 ALTER 权。

但如果一个数据库被“拒绝访问锁”锁定，则任何用户都不能解锁，只能由系统内部自动解除，或者重新启动系统。

- 备份数据库。如果满足下列条件之一，则有权备份数据库：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库的拥有者。
 4. 数据库所有者所在用户组的 GDA 用户。

- 维护数据库索引及优化数据库。如果满足下列条件之一，则有权维护数据库索引或优化数据库：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库的拥有者。
 4. 数据库具有公共 INDEX 属性。
 5. 数据库拥有者所在用户组的 GDA 用户。
 6. 拥有该数据库的 INDEX 或 ALTER 或 DROP 权的用户。

- 维护数据库记录（数据库作为加载、插入、删除、修改和复制记录的目标数据库）。如果满足下列条件之一，则有权维护数据库记录：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库的拥有者。
 4. 数据库具有公共 UPDATE 属性。
 5. 数据库拥有者所在用户组的 GDA 用户。
 6. 拥有该数据库的 UPDATE 或 INDEX 或 ALTER 或 DROP 权的用户。

- 从数据库复制记录。如果满足下列条件之一，则有权从数据库复制记录：
 1. 超级用户。
 2. DBA 用户。
 3. 源数据库的拥有者。
 4. 源数据库具有公共 UPDATE 属性。
 5. 源数据库拥有者所在用户组的 GDA 用户。
 6. 拥有该源数据库的 UPDATE 或 INDEX 或 ALTER 或 DROP 权的用户。

- 检索数据库记录。如果满足下列条件之一，则有权检索数据库记录：
 1. 超级用户。
 2. DBA 用户。
 3. 数据库的拥有者。
 4. 数据库具有公共 SELECT 属性。
 5. 数据库拥有者所在用户组的 GDA 用户。
 6. 拥有该数据库的 SELECT 或 UPDATE 或 INDEX 或 ALTER 或 DROP 权限的用户。

4.3.4 数据库字段操作

如果满足下列条件之一，则有权在数据库字段上进行检索：

1. 数据库字段具有公共 SELECT 属性。
2. 拥有该字段所在数据库的 SELECT 权的用户。
3. 拥有该字段的 SELECT 权的用户。

4.3.5 视图操作

- 创建视图及恢复自动模式视图。如果满足下列条件之一，则有权创建视图或恢复自动模式视图：
 1. 超级用户。
 2. DBA 用户。
 3. GDA 用户。
 4. RESOURCE 用户。

如果创建者是 GDA 或 RESOURCE 用户，则该用户必须对视图字段所映射的数据库字段具有 SELECT 权。

- 删除视图。如果满足下列条件之一，则有权删除视图：
 1. 超级用户。
 2. DBA 用户。
 3. 视图的拥有者。
 4. 视图具有公共 DROP 属性。
 5. 视图拥有者所在用户组的 GDA 用户。
 6. 拥有该视图的 DROP 权的用户。

当删除一个视图时，系统将自动删除所有用户在该视图上的权限。

- 修改视图及自动模式视图中添加数据库。如果满足下列条件之一，则有权修改视图或在自动模式视图中添加数据库：
 1. 超级用户。
 2. DBA 用户。
 3. 视图的拥有者。
 4. 视图具有公共 ALTER 属性。
 5. 视图拥有者所在用户组的 GDA 用户。
 6. 拥有该视图的 ALTER 权的用户。

如果修改者是 GDA 或 RESOURCE 用户，则该用户必须对新增的视图字段所映射的数据库字段具有 SELECT 权。

当删除一个视图字段时，系统将自动删除所有用户在该视图字段上的权限。

- 检索视图记录。如果满足下列条件之一，则有权检索视图记录：
 1. 超级用户。
 2. DBA 用户。
 3. 视图的拥有者。
 4. 视图具有公共 SELECT 属性。
 5. 视图拥有者所在用户组的 GDA 用户。
 6. 拥有该视图的 SELECT 或 ALTER 或 DROP 权的用户。

- 备份自动模式视图。如果满足下列条件之一，则有权备份自动模式视图：
 1. 超级用户。
 2. DBA 用户。
 3. 自动模式视图的拥有者。
 4. 自动模式视图拥有者所在用户组的 GDA 用户。

- 维护自动模式视图索引及优化自动模式视图。如果满足下列条件之一，则有权维护自动模式视图索引或优化自动模式视图：
 1. 超级用户。
 2. DBA 用户。
 3. 自动模式视图的拥有者。
 4. 自动模式视图具有公共 INDEX 属性。
 5. 自动模式视图拥有者所在用户组的 GDA 用户。
 6. 拥有该自动模式视图的 INDEX 或 ALTER 或 DROP 权的用户。

- 维护自动模式视图记录（自动模式视图作为加载、插入、删除、修改和复制记录的目标视图）。如果满足下列条件之一，则有权维护自动模式视图记录：
 1. 超级用户。
 2. DBA 用户。
 3. 自动模式视图的拥有者。
 4. 自动模式视图具有公共 UPDATE 属性。
 5. 自动模式视图拥有者所在用户组的 GDA 用户。
 6. 拥有该自动模式视图的 UPDATE 或 INDEX 或 ALTER 或 DROP 权的用户。

- 从视图复制记录。如果满足下列条件之一，则有权从视图复制记录：
 1. 超级用户。
 2. DBA 用户。
 3. 视图中的源数据库的拥有者。

4. 视图中的源数据库具有公共 UPDATE 属性。
5. 视图中的源数据库所有者所在用户组的 GDA 用户。
6. 拥有该视图中源数据库的 UPDATE 或 INDEX 或 ALTER 或 DROP 权的用户。

4.3.6 视图字段操作

如果满足下列条件之一，则有权在视图字段上进行检索：

1. 视图字段具有公共 SELECT 属性。
2. 拥有该字段所在视图的 SELECT 权的用户。
3. 拥有该字段的 SELECT 权的用户。

4.3.7 数据格式操作

只有超级用户有权创建和修改数据格式。

4.3.8 词典操作

- 创建词典。如果满足下列条件之一，则有权创建词典：
 1. 超级用户。
 2. DBA 用户。
 3. GDA 用户。
 4. RESOURCE 用户。
- 删除词典。如果满足下列条件之一，则有权删除词典：
 1. 超级用户。
 2. DBA 用户。
 3. 词典的拥有者。
 4. 词典具有公共 DROP 属性。
 5. 词典拥有者所在用户组的 GDA 用户。
 6. 拥有该词典的 DROP 权的用户。
- 修改词典的属性。如果满足下列条件之一，则有权修改词典的属性：
 1. 超级用户。
 2. DBA 用户。
 3. 词典拥有者所在用户组的 GDA 用户。
 4. 词典的拥有者。
- 更新维护词典。如果满足下列条件之一，则有权更新维护词典：

1. 超级用户。
 2. DBA 用户。
 3. 词典的拥有者。
 4. 词典具有公共 UPDATE 属性。
 5. 词典拥有者所在用户组的 GDA 用户。
 6. 拥有该词典的 UPDATE 或 DROP 权的用户。
- 引用词典。词典的 SELECT 权限始终是开放的，任何用户都能引用所有的词典。

第 5 章 系统配置

TRS 全文数据库系统运行时有许多参数,这些参数就构成了 TRS 的系统配置数据词典。具体包括:物理内存数、用户实例最大空闲时间、用户实例最小空闲时间、检索超时时间、最大排序记录数、最大检索历史步数、最大下载结果记录数、空闲状态起止时间、缺省数据库属性、缺省知识词典集、数据库路径集、缺省数据库备份设备以及临时文件存放路径等。

5.1 物理内存数

物理内存数是指 TRS 全文数据库系统能够使用的物理内存的总量。既可以是机器的实际内存大小,也可以是一个虚拟的内存大小。这个参数将影响一个数据库记录的最大长度(参见 1.3 节),同时数据库的加载和检索过程所使用的缓冲区的大小,也将依此进行自动调节。当设置小于 64M 时,系统内部默认为 64M。

对于大多数操作系统平台来说,TRS 全文数据库系统安装或升级时会自动检测机器的实际物理内存量来确定或更新这个参数,但升级时如果原来的值已大于实际测得的值,则会保持原来的值。

5.2 用户实例最大空闲时间

在访问 TRS 全文数据库系统之前,需要通过 TRS 全文数据库系统中的一个有效的用户进行登录,TRS 全文数据库系统为每一次登录都分配一个独立的用户实例号。

在大多数应用中,当决定不再访问 TRS 全文数据库系统时,总是主动退出 TRS 全文数据库系统,即返还所得的用户实例号。但有的应用(如 Web 应用)却没有机会这样做。这种情况将导致系统会累积许多已经无人使用的用户实例,与此相关联的一些系统资源也得不到及时释放。

所以 TRS 全文数据库系统设计了一种自动清除无效用户实例的机制:每当有新的用户登录,并且已有的用户实例数正处在“检查点”(256 的倍数)时,系统将自动检查所有用户实例是否过期,即检查每一个用户实例没有被用来访问系统的时间是否超过了一定的时间间隔,如果超过了这一时间间隔,则该用户实例将会被强制取消。“检

查点”的设置意味着即使用户实例已经过期，但如果未达到检查点，则不会被强制取消。这里所规定的“时间间隔”就是用户实例的最大空闲时间。用户实例被强制取消后，不能再用来访问系统，必须重新登录。特别说明：来自集群系统的用户实例始终是有有效的，其退出完全有集群系统控制。

5.3 用户实例最小空闲时间

如 5.2 节所述，当一个用户实例没有被用来访问系统的时间超过了一定的时间间隔时，有可能会被强制取消。与此对应，当一个用户实例未被用来访问系统的时间没有超过某一规定的时间间隔时，在任何情况下都不允许被强制取消。这里所规定的“时间间隔”就是用户实例的最小空闲时间。这个配置参数在使用外部命令强制清除系统中的用户实例时将用到。

5.4 检索超时时间

完成任何检索操作都需要一定的时间。检索条件越复杂，数据库的规模越大，检索时间就越长。在多用户系统中，有时需要限制每个检索操作所占用的系统时间，以牺牲个别要求来提高其整体的吞吐量。检索超时时间正是应此需求而设置的。当一个检索操作的运行时间超过了所规定的时间时，系统就会强行终止这次检索操作。

5.5 最大排序记录数

TRS 全文数据库系统的每一次检索操作都会产生一个检索结果集。在应用中，可以根据需要对这一检索结果集所对应的数据库记录进行排序。但是，如果检索结果集太大，即结果记录太多，排序起来需要花费可观的时间，使得检索响应时间延长。然而在一般意义下，检索结果集太大说明检索需求还不够明确，往往需要增加检索条件以缩小检索结果集，所以对大检索结果集的排序也就没有太大的意义。

基于上述理由，TRS 全文数据库系统提供了一个“最大排序记录数”的配置项来决定是否拒绝对检索结果记录的排序要求：如果检索结果集的数据库记录数超过了该配置项所规定的数目，则 TRS 全文数据库系统始终不会对检索结果记录进行排序。

这个配置项只是常规检索的缺省限制，且最大取值只能是 65533，取 0 时表示不作限制。如果需要突破 65533 的限制，这只能采用下面的方法：首先按《服务器安装》部分的配置文件中有关配置域 Retrieval 的配置项 SortAlwaysMax 的说明，设置

SortAlwaysMax 成所希望的目标值，再在检索时指定“总是允许排序”的选项。

如果对检索结果进行 TOP N 剪裁排序，且实际检索结果记录数超过“最大排序记录数”时，则不会对全部结果进行排序，而只对先出现的前“最大排序记录数”个记录（即不排序时结果集里前面一部分记录）进行排序，然后再截取前 N 记录保存检索结果。【修订 067】

5.6 最大检索历史步数

检索操作所产生的检索结果集必须保存，以备后续操作的使用。对于同一个用户实例，每保存一个检索结果，就是一步检索历史。“最大检索历史步数”配置项是 TRS 全文数据库系统在创建新的用户时，用户属性“最大检索历史步数”的缺省值（参见 1.2 节）。TRS 全文数据库系统总是保留最新的若干步检索历史，此前的检索历史不能再被引用。

5.7 空闲状态起止时间

空闲状态起止时间包括空闲状态起始时间和空闲状态终止时间。空闲状态起始时间是指允许 TRS 全文数据库系统进入空闲状态的起始时间，空闲状态终止时间是指允许 TRS 全文数据库系统进入空闲状态的截止时间。【修订 005】“空闲状态起始时间”的取值范围是 0 到 23，其它值表示禁止服务器的自动维护功能；“空闲状态终止时间”的取值范围也是 0 到 23，其它值表示没有限制，但服务器自动维护操作的最大工作时间为 23 小时；当“空闲状态起始时间”和“空闲状态终止时间”相等时也表示禁止服务器的自动维护功能。请参见《服务器安装》部分的配置文件中有关配置域 Maintain 的有关配置项的说明。^[4]这个两个配置项改变后，需要重新启动 TRS 全文数据库系统才有效。

起始时间和终止时间之间的小时数为 TRS 全文数据库系统进入空闲状态的检测窗口，如果不限定终止时间，则检测窗口为一小时。如果机器的当前系统时间在检测窗口内，并且已有一段时间（大约 10 分钟）没有任何用户访问 TRS 全文数据库系统，则 TRS 全文数据库系统进入空闲状态。

在空闲状态下，系统可以进行一系列的自动维护任务，如优化数据库，整理用户实例等。当系统完成一个自动维护任务时，如果机器的当前系统时间仍处在“空闲状态起始时间”和“空闲状态终止时间”之间，则系统会启动一个新的自动维护任务。

如果启动系统的当前时间到设定的空闲起始时间之间不足 4 个小时，则服务器会忽略第一次到达的空闲起始时间，而必须等到第二个空闲起始时间到达时才有可能进入空闲状态。但系统自动重启时（请参见 20.1.8 节的说明），或者运行带 ‘-reboot’ 选项的关闭服务器的命令（如 ‘tdbshutdown -reboot’）重启系统时，没有这个限制。

特别说明：对于 6.80.5230 之前的版本，如果服务器作为集群的后台运行，则该配置无效，即服务器不支持自动维护功能；从 6.80.5230 版本及之后的版本，如果服务器作为集群的后台运行，则只有当 trssearcher 也启动时，服务器才支持维护功能，否则该配置无效。

5.8 缺省数据库属性

缺省数据库属性是指在创建新的数据库时部分数据库属性的缺省值。包括：字母单词词根索引开关、字母串大小写（敏感）索引开关、数字索引开关以及数据库中数据的语言字符码制。参见 1.3 节。

5.9 缺省知识词典集

在 TRS 全文数据库系统中，一个知识词典集由索引词典和辅助知识词典组成，包括分词词典、停用词典、主题词典、同义词典和反义词典，请参见 1.8 节的说明。每一种 TRS 所支持的语言都可以定义一个知识词典集。其中，分词词典和停用词典是 TRS 全文数据库系统在创建新的数据库时的缺省词典，而主题词典、同义词典和反义词典则是在调用检索函数进行数据库检索时，检索函数的缺省词典（参见 7.2 节）。

5.10 数据库路径集

数据库路径集是允许 TRS 全文数据库系统存放数据库的数据和索引文件的目录的集合，这些目录必须对 TRS 全文数据库系统开放其读/写权限。在创建新的数据库，并确定其数据和索引的存放路径时，只能从中选取一个。

5.11 缺省数据库备份设备

缺省数据库备份设备是 TRS 全文数据库系统进行数据库备份时的缺省目标设备名或目录。

5.12 临时文件存放路径

临时文件存放路径指定了 TRS 全文数据库系统运行时，临时文件的存放目录。在这个目录下至少会有以下几个子目录：日志目录 LOGS、检索历史目录 HIST、二级检索缓冲目录 SWAP 和网络传输临时文件 TRAN。对于这些子目录而言，当该配置项改变后，需要重新启动 TRS 全文数据库系统才有效。

第二部分 数据库检索

本部分详细介绍了在 TRS 数据库丰富的检索运算符、检索函数、检索表达式的构造规则，以及检索结果的处理。其内容不涉及任何 TRS 数据库检索工具的使用方法，不同应用工具的具体检索过程请参见相应的使用手册。如果未加特别说明，以下章节中的[]是一个语法符号，表示其中的内容可以省略（取缺省值），或是一个可选项。

读者对象

本部分的读者为数据库管理员、具有检索数据库权限的 TRS 用户。在阅读本手册之前，建议您先阅读《数据库管理》部分。

内容组织

本手册按如下结构组织：

- 第 6 章 检索运算符
- 第 7 章 函数与键值
- 第 8 章 表达式构造
- 第 9 章 检索结果处理

第6章 检索运算符

运算符是表达式的不可缺少的组成部分，负责将多个检索目标按照一定的逻辑连接起来。TRS的运算符分为统计运算符和检索运算符两大类，不区分大小写。

6.1 统计运算符

统计运算符就是算术运算符，在统计表达式中用于对检索结果的统计，包括：加、减、乘、除和取余。

符号	含义和作用	举例
*	数值型数据相乘	sum(salary) * 0.10
/	数值型数据相除	sum(salary) / 20
%	数值型数据求余数	sum(salary) % 20
+	数值型数据相加	avg(age) + 10
-	数值型数据相减或求负数	avg(age) - 10

6.2 检索运算符

检索运算符又分为6类，分别是：比较运算符、逻辑运算符、重复逻辑运算符、属性运算符、限定运算符和加权运算符。检索运算符只能在检索表达式中使用，并且都不能作用于BIT字段。

6.2.1 比较运算符

符号	含义和作用	举例
=	等于	作者 = 李力
!=	不等于	版面 != 国际
>	大于	工资 > 1000
<	小于	工资 < 1000
>=	大于等于	工资 >= 1000
<=	小于等于	工资 <= 1000

- 比较运算符作用于NUMBER、DATE、CHAR字段。但‘=’和‘!=’可作用于PHRASE和DOCUMENT字段。
- 比较运算符的左值必须是一个检索入口（字段名或其别名）。
- 只有当字段类型为DATE、NUMBER和CHAR时，比较运算符的右值才可以是一个键值列表，并且只有比较运算符‘=’和‘!=’支持右值为一个键值列表。
- 如果检索条件不是单个词，而是一个表达式，则该表达式应该用圆括号“(”和“)”括起来。

6.2.2 逻辑运算符

符号	含义和作用	举例
and	两者必须同时满足	中国 and 美国
xor	两者只能满足其一	中国 xor 美国
not	两者中只能出现前者	中国 not 美国
or	两者至少满足其一	中国 or 美国

- 逻辑运算符可作用于除BIT之外的所有字段。
- 逻辑运算符不仅能够连接两个检索词，也可以连接两个检索子表达式，从而构造出非常复杂的检索表达式。

6.2.3 重复逻辑运算符

符号	含义和作用	举例
*=	检索各个字段上满足同一条件的记录	c1,c2 *= (中国 and 香港)
^=	检索只有一个字段上满足条件的记录	c1,c2 ^= (中国 and 香港)
-=	检索只有第一个字段满足条件的记录	c1,c2 -= (中国 and 香港)
+=	检索任意字段上同时满足条件的记录	c1,c2 += (中国 and 香港)

- 重复逻辑运算符实际上是在各个指定的字段上用相同的条件式进行检索后，再在各个字段之间进行指定的逻辑运算，可作用于除BIT之外的所有字段。
- 重复逻辑运算符的左值必须是由单字节的逗号‘,’分隔的两个或两个以上的检索入口（字段名或其别名）。如果右值不是单个词，而是一个表达式（该运算需要在这个表达式上进行），则该表达式应该用圆括号“(”和“)”括起来。

6.2.4 属性运算符

符号	含义和作用	举例
adj	两者邻接先后同时出现	中国 adj 香港
equ/n	两者正好相距n个位置先后同时出现	中国 equ/10 香港

符号	含义和作用	举例
pre/n	两者相距至多n个位置先后同时出现	中国 pre/10 香港
pre/n#	两者正好相距n个位置先后同时出现	中国 pre/10# 香港
pre/pos=n		中国 pre/pos=10 香港
pre/nL	两者相距至少n个位置先后同时出现	中国 pre/10L 香港
pre/sen	两者在同一句中先后同时出现	中国 pre/sen 香港
pre/sen=n	两者在指定的句中先后同时出现	中国 pre/sen=1 香港
pre/seg	两者在同一段中先后同时出现	中国 pre/seg 香港
pre/seg=n	两者在指定的段中先后同时先出现	中国 pre/seg=1 香港
pre/tit	两者在标题中先后同时出现	中国 pre/tit 香港
pre/txt	两者在正文中先后同时出现	中国 pre/txt 香港
pre/att	两者在指定属性的节点先后同时出现	中国 pre/文 香港
pre/att=v	两者在指定属性值的节点先后同时出现	中国 pre/章=2 香港
pre/att\$sub=v	两者在指定特征值的节点先后同时出现	中国 pre/文\$色=红 香港
and/n	两者相距至多n个位置同时出现	中国 and/10 香港
and/n#	两者正好相距n个位置同时出现	中国 and/10# 香港
and/pos=n		中国 and/pos=10 香港
and/nL	两者相距至少n个位置同时出现	中国 and/10L 香港
and/sen	在同一句中两者同时出现	中国 and/sen 香港
and/sen=n	在指定的句中两者同时出现	中国 and/sen=2 香港
and/seg	在同一段中两者同时出现	中国 and/seg 香港
and/seg=n	在指定的段中两者同时出现	中国 and/seg=2 香港
and/tit	在标题中两者同时出现	中国 and/tit 香港
and/txt	在正文中两者同时出现	中国 and/txt 香港
and/att	在指定属性的节点两者同时出现	中国 and/文 香港
and/att=v	在指定属性值的节点两者同时出现	中国 and/章=2 香港
and/att\$sub=v	在指定特征值的节点两者同时出现	中国 and/文\$色=红 香港
and/col	在同一字段中两者同时出现	中国 and/col 香港
xor/n	同一句中不能在n个位置之内同时出现	中国 xor/10 香港
xor/n#	同一句中不能正好相距n个位置同时出现	中国 xor/10# 香港
xor/pos=n		中国 xor/pos=10 香港
xor/nL	同一句中不能在n个位置之外同时出现	中国 xor/10L 香港
xor/sen	在同一句中只能出现其一	中国 xor/sen 香港
xor/sen=n	在指定的句中只能出现其一	中国 xor/sen=1 香港
xor/seg	在同一段中只能出现其一	中国 xor/seg 香港
xor/seg=n	在指定的段中只能出现其一	中国 xor/seg=1 香港
xor/tit	在标题中只能出现其一	中国 xor/tit 香港

符号	含义和作用	举例
xor/txt	在正文中出现且只能出现其一	中国 xor/txt 香港
xor/att	在指定属性的节点只能出现其一	中国 xor/文 香港
xor/att=v	在指定属性值的节点只能出现其一	中国 xor/章=2 香港
xor/att\$sub=v	在指定特征值的节点只能出现其一	中国 xor/文\$色=红 香港
not/n	同一句中不能在n个位置之内出现后者	中国 not/10 香港
not/n#	同一句中不能正好相距n个位置出现	中国 not/10# 香港
not/pos=n	后者	中国 not/pos=10 香港
not/nL	同一句中不能在n个位置之外出现后者	中国 not/10L 香港
not/sen	在同一句中只能出现前者	中国 not/sen 香港
not/sen=n	在指定的句中出現且只能出现前者	中国 not/sen=1 香港
not/seg	在同一段中只能出现前者	中国 not/seg 香港
not/seg=n	在指定的段中出现且只能出现前者	中国 not/seg=1 香港
not/tit	在标题中出现且只能出现前者	中国 not/tit 香港
not/txt	在正文中出现且只能出现前者	中国 not/txt 香港
not/att	在指定属性的节点只能出现前者	中国 not/文 香港
not/att=v	在指定属性值的节点只能出现前者	中国 not/章=2 香港
not/att\$sub=v	在指定特征值的节点只能出现前者	中国 not/文\$色=红 香港
or/sen=n	在指定的句中至少出现一个	中国 or/sen=1 香港
or/seg=n	在指定的段中至少出现一个	中国 or/seg=1 香港
or/tit	在标题中至少出现一个	中国 or/tit 香港
or/att	在指定属性的节点至少出现一个	中国 or/文 香港
or/att=v	在指定属性值的节点至少出现一个	中国 or/章=2 香港
or/att\$sub=v	在指定特征值的节点至少出现一个	中国 or/文\$色=红 香港

- 属性运算符只作用于DOCUMENT字段。但adj、equ/n、X/n、X/n#、X/nL和X/pos=n运算符可作用于PHRASE字段，其中的‘X’分别代表‘pre’、‘and’、‘or’、‘not’和‘xor’，以下同，不再特别说明。
- 在属性运算符equ/n、X/n、X/n#、X/nL和X/pos=n中的距离是指两个字/词之间所间隔的“位置”的个数，而不是其首“位置”的差值。一个字/词在句中的位置的计算规则请参见3.4.1节的说明。
- 上述列出的属性运算符只是一个基本集合，实际上除特别说明外，相同运算符后面的不同类别的属性可以组合使用。属性的类别包括：位（X/n、X/n#、X/nL和X/POS=n）、句（X/sen和X/sen=n）、段（X/seg、X/seg=n、X/tit和X/txt），每个格式属性都是不同的类别。例如：“中国 and/10/sen>=1/txt/文\$色=红/章=2 香港”。
- 属性运算符X/att、X/att=v以及X/att\$sub=v中的att对应一个具体的格式索引属性别名（如果没有别名，则为格式索引属性名，其本身不是一个法定的语法符号，而只是用来说明语法的合法属性的表示符）；sub对应一个具体的格式索引属性的特

征别名（如果特征没有别名，则为特征名，其本身不是一个法定的语法符号，而只是用来说明语法的合法特征的表示符），如果这个特征是该格式索引属性所定义的唯一特征，或者是第一个特征，则\$sub部分可以省略；v是一个格式索引属性或其特征的值，值的形式与具体的编码方式有关。如果att之前加上“~”符，则表示是属性限定条件的“非”运算，如“and/~att\$sub=v”的意义是“两者同时出现，且不满足‘att的sub特征值为v’的条件”。

- 运算符pre/att、pre/att=v和pre/att\$sub=v可以单独使用，默认任意位置限定。att、att=v和att\$sub=v也可与pos、sen、seg、tit、txt等一起使用。
- 属性运算符中的“=”可以用其他任意比较运算符替代。但X/att=v和X/att\$sub=v中的“=”可以用哪些比较运算符替代，完全取决于格式索引属性或其特征的编码方式。
- 属性运算符X/sen=n和X/seg=n中的n，都从0开始计数。
- 属性运算符X/n[#/L]的隐含条件是它们在同一句中。N的计算方法是位置在前的检索词的词尾到位置在后的检索词的词首之间的字母单词或中文字个数。其中字母L大小写无关，表示Leastwise。这里X/n等价于X/pos<=n；X/n#等价于X/pos=n；X/nL等价于X/pos>=n。
- 属性运算符adj等价于pre/0；equ/n等价于pre/n#。运算符adj和equ/n只是为了向下兼容而设计的，应尽量避免使用。
- 属性运算符X/tit等价于X/seg=0。
- 属性运算符X/txt不等价于X/seg>0(或X/seg!=0)。X/txt不要求满足“段号相等”的条件。
- 属性运算符and/col在两种情况下会有实际效果。一种情况是：指定了多个缺省检索字段，字段之间的逻辑关系为“或”，且按字段逻辑优先的模式进行扩展（请参见8.3节）；另一种情况是：在虚拟逻辑字段上检索，且虚拟逻辑字段所对应的实际字段之间的逻辑关系为“或”。由于虚拟逻辑字段在检索时按单一字段处理的，所以第二种情况与第一种情况实际是相同的。这样综合起来，两种情况的表达式的形式就是：(c1,c2 += T1) and/col (c1,c2 += T2)，因‘/col’的限定，它实际等价于：(c1 = (T1 and T2)) or (c2 = (T1 and T2))。该运算符的主要用途是便于从某些已有应用系统的运算符到TRS运算符的功能性兼容转换。【修订120】

6.2.5 限定运算符

符号	含义和作用	使用举例
/pos=n	检索词在指定的位置出现	中国/pos=0
/sen=n	检索词在指定的句中出現	中国/sen=12
/sen	LIKE函数中规定的词在同一句中出現	LIKE(中国国家剧院)/sen
/seg=n	检索词在指定的段中出現	中国/seg=12

符号	含义和作用	使用举例
/seg	LIKE函数中规定的词在同一段中出现	LIKE(中国国家剧院)/seg
/tit	检索词在标题中	中国/tit
/txt	检索词在正文中	中国/txt
/att	检索词具有att属性	中国/文
/att=v	检索词具有att属性, 且值为v	中国/章=2
/att\$sub=v	检索词具有att属性, 且sub特征值为v	中国/文\$色=红
/span=n	检索词或短语必须在跨度范围内出现	LIKE(中国国家剧院)/span=20
/frec=n	检索词在记录中的频率必须满足要求	中国/frec>5
/fre=n	输出满足词频数条件的索引词(值)	%/fre>=1000000
/rec=n	输出满足命中数条件的索引词(值)	%/rec>=1000000

- 上述列出的限定运算符只是一个基本集合, 实际上除特别说明外, 不同类别的限定符可以组合使用。限定符类别包括: 位(/POS=n)、句(/sen和/sen=n)、段(/seg、/seg=n、/tit和/txt)、频率(/frec=n), 以及每个格式索引属性都是不同的限定符。例如检索表达式: “中国/pos=10/sen>=1/seg<4/文\$色=红/章=2/frec>5”。
- 限定符/rec=n和/fre=n只用于输出索引词(或字段值), 且n的值不能为0, 而对于以记录为结果的检索则无效。可作用于除BIT之外的所有字段。
- 限定符/pos=n、frec=n和/span=n可作用于DOCUMENT和PHRASE字段, 其它检索限定符则只对DOCUMENT字段上的检索有效。限定符/frec=n中n的值不能为0。
- 限定符/sen和/seg只能用于LIKE函数结果的限定。并且需要特别注意的是: 当限定符/pos=n与限定符/sen或/sen=n同时使用时, 限定符/pos=n的意义发生改变, 表示词之间的距离, 而不是词本身的位置, 此时函数LIKE(...)/sen/pos=n的效果相当于使用属性运算符and/pos=n。请参见LIKE函数的说明。
- 限定符/att、/att=v以及/att\$sub=v中的att对应一个具体的格式索引属性别名(如果没有别名, 则为格式索引属性名, 其本身不是一个法定的语法符号, 而只是用来说明语法的合法属性的表示符); sub对应一个具体的格式索引属性的特征别名(如果特征没有别名, 则为特征名, 其本身不是一个法定的语法符号, 而只是用来说明语法的合法特征的表示符), 如果这个特征是该格式索引属性所定义的唯一特征, 或者是第一个特征, 则\$sub部分可以省略; v是一个格式索引属性或其特征的值, 值的形式与具体的编码方式有关。如果att之前加上“~”符, 则表示是限定条件的“非”运算, 如“/~att\$sub=v”的意义是“检索词不满足‘具有att属性, 且sub特征值为v’的条件”。
- 跨度限定符/span=n只能用于LIKE和INCLUDE函数的限定(请参见7.2节), 其含义是: 所有满足条件的检索词或短语都必须出现在(同段)同句, 而且在第一个命中词(或短语)的首字与最后一个命中词(或短语)的尾字之间的位置跨度必须满足该限定所规定的条件。【修订121】
- 限定符中的“=”可以用其他任意比较运算符替代。但/att=v和/att\$sub=v中的“=”

可以用哪些比较运算符替代，取决于格式索引属性或其特征的编码方式。

- 限定符中的n，都从0开始计数。其中/pos=n中的n是指检索词在句中出现的起始位置，一个字/词在句中的位置的计算规则请参见3.4.1节的说明。限定符/pos=n与属性运算符X/pos=n是有差别的。
- 限定符/txt等价于/seg>0(或/seg!=0)，并且可以与/sen=n和/pos=n同时使用，这一点与运算符X/txt不同。但限定符/tit等价于/seg=0。

6.2.6 加权运算符

表示符号	含义和作用	使用举例
/n	指定某个检索词（短语）或某个字段在检索结果中的权重	计算机/10 * 科学标题/10,正文/5 +=计算机

- 权重n是一个不小于1的【修订013】正整数。权重只有在对检索结果进行相关性排序时才能体现出来。
- 检索词在查询中的缺省权重是1；字段在查询中的缺省权重是0，但对于缺省检索字段，如果其它字段都未指定权重（即缺省为0），则缺省检索字段的缺省权重是1；只要有一个字段指定了非0权重，则缺省检索字段的缺省权重就是0。
- 权重为0的字段实际不参与相关性排序；只有参与了检索的字段才有可能参与相关性排序；只有字段允许参与相关性排序时（请参见1.4节的详细说明），对该字段的加权才有效^[4]。
- 检索词的权重只对允许参与相关性排序的DOCUMENT或PHRASE字段。在进行相关性排序前，系统需要根据固定的算法自动计算每个检索词对一个命中记录的相关度的作用因子（即贡献率），权重的作用就是用来倍增放大这个词的“作用因子”的。

6.2.7 运算符的等价符号

用英文单词书写运算符，有时觉得比较麻烦，可用下列等价符号进行替代：

常规运算符	!=	ADJ	PRE	AND	XOR	NOT	OR
等价运算符	<>	空格	&	*	^	-	+

6.2.8 运算符与字段的关系

运算符	DATE	NUMBER	CHAR	PHRASE	DOCUMENT
=	✓	✓	✓	✓	✓
!=	✓	✓	✓	✓	✓
>	✓	✓	✓		

运算符	DATE	NUMBER	CHAR	PHRASE	DOCUMENT
<	✓	✓	✓		
>=	✓	✓	✓		
<=	✓	✓	✓		
*=	✓	✓	✓	✓	✓
^=	✓	✓	✓	✓	✓
-=	✓	✓	✓	✓	✓
+=	✓	✓	✓	✓	✓
adj				✓	✓
equ/n				✓	✓
pre/n[#/L]				✓	✓
pre/pos=n				✓	✓
pre/sen[=n]					✓
pre/seg[=n]					✓
pre/tit					✓
pre/txt					✓
pre/[~]att[[\$sub]=v]					✓
and	✓	✓	✓	✓	✓
and/n[#/L]				✓	✓
and/pos=n				✓	✓
and/sen[=n]					✓
and/seg[=n]					✓
and/tit					✓
and/txt					✓
and/[~]att[[\$sub]=v]					✓
and/col	✓	✓	✓	✓	✓
xor	✓	✓	✓	✓	✓
xor/n[#/L]				✓	✓
xor/pos=n				✓	✓
xor/sen[=n]					✓
xor/seg[=n]					✓
xor/tit					✓
xor/txt					✓
xor/[~]att[[\$sub]=v]					✓
not	✓	✓	✓	✓	✓
not/n[#/L]				✓	✓
not/pos=n				✓	✓
not/sen[=n]					✓

运算符	DATE	NUMBER	CHAR	PHRASE	DOCUMENT
not/seg[=n]					✓
not/tit					✓
not/txt					✓
not/[~]att[[\$sub]=v]					✓
or	✓	✓	✓	✓	✓
or/sen=n					✓
or/seg=n					✓
or/tit					✓
or/txt					✓
or/[~]att[[\$sub]=v]					✓
/pos=n				✓	✓
/sen[=n]					✓
/seg[=n]					✓
/tit					✓
/txt					✓
/[~]att[[\$sub]=v]					✓
/span=n				✓	✓
/frec=n				✓	✓
/rec=n	✓	✓	✓	✓	✓
/fre=n	✓	✓	✓	✓	✓
/n	✓	✓	✓	✓	✓

- 表中的“✓”表示该字段支持这种运算符。
- 算术运算符不能用于检索表达式，所以任何字段都不支持。
- 所有运算符都不能作用于BIT字段。
- 运算符中的‘[’和‘]’（如and/[~]att[[\$sub]=v]）不是运算符的一部分，表示被包括的部分可以省略。但省略与否，其运算符的含义是有差别的。

第7章 函数与键值

TRS全文数据库系统所支持的函数分为两种：统计函数和检索函数。统计函数是统计表达式的一部分，用于对检索结果的统计；而检索函数则是检索表达式的一部分，以函数的形式描述检索结果必须满足的条件。检索键值是最基本的检索表达式，用来直接描述检索结果必须满足的条件，相当于关系型数据库中的关键字。

7.1 统计函数

函数名	参数	含义和作用
AVG()	字段名	求数据库中满足条件的记录中某个数值型字段的平均值
MAX()	字段名	求数据库中满足条件的记录中某个数值型字段的最大值
MIN()	字段名	求数据库中满足条件的记录中某个数值型字段的最小值
SUM()	字段名	求数据库中满足条件的记录中某个数值型字段的和
COUNT()	*	用于统计数据库中满足条件的记录有多少

- 统计函数只能用于统计表达式。
- 函数参数中的“字段名”也可以使用字段的别名。
- 只有NUMBER型字段作为统计函数的参数时才有结果，否则函数的结果总为0。

7.2 检索函数

函数名	参数	含义和作用
BT()	主题词[, 词典名]	获得某个主题词的上位词
NT()	主题词[, 词典名]	获得某个主题词的下位词
PT()	主题词[, 词典名]	获得某个主题词的等同词
UF()	主题词[, 词典名]	获得某个主题词的替代词
AB()	主题词[, 词典名]	获得某个主题词的缩略词
ABF()	主题词[, 词典名]	获得某个主题词的被缩略词
RT()	主题词[, 词典名]	获得某个主题词的相关词

函数名	参数	含义和作用
LE()	主题词[, 词典名]	获得某个主题词的外文等同词
ALLBT()	主题词[, 词典名]	获得某个检索词的所有层次的上位词
ALLNT()	主题词[, 词典名]	获得某个检索词的所有层次的下位词
ST()	主题词[, 词典名]	获得某个检索词的同义词。因词典也可包含非同义的相关词, 所以也称广义同义词
AT()	主题词[, 词典名]	获得某个检索词的反义词
BETWEEN()	值1, 值2	获取在指定范围内的字段值
BEYOND()	值1, 值2	获取在指定范围之外的字段值
EXACT()	值	对值不进行任何处理, 直接匹配索引
INCLUDE()	词串	对词串不进行分词处理, 并加上前后模糊
INCLUDE()	短语串组, 匹配度	以短语串组中给定的短语为单位, 忽略短语之间的位置关系, 进行‘与’、‘或’逻辑运算【修订124】
RANGE()	词串 = 值	在同一句进行数值范围检索(一般不鼓励使用)
RANGE()	数值1, 数值2	在文本中检索指定范围内的数值【修订082】
DATA()	键值列表	直接扫描数据进行字段值匹配, 无参数时表示空值查找
LIKE()	词串[, 开关]	对词串进行分词处理, 忽略词之间的位置关系, 进行‘与’、‘或’逻辑运算
SPELL()	规则串	对规则串进行分拆组合, 形成词或短语的集合后进行检索【修订083】
NEAR()	目标词, 数量	在指定的“目标词”附近读取一定“数量”的索引词【修订085】

- BETWEEN和BEYOND函数只能作用于DATE、NUMBER和CHAR字段; EXACT和NEAR检索函数可作用于除BIT之外的所有字段; DATA函数可作用于包括BIT字段在内的所有字段; SPELL函数可作用于CHAR、PHRASE和DOCUMENT字段, 但作用于PHRASE和DOCUMENT字段时不支持非字母文字;【修订123】INCLUDE函数可作用于PHRASE和DOCUMENT字段, 但两个参数的INCLUDE函数还可作用于CHAR字段^[4];【修订118】RANGE和LIKE函数只能作用于PHRASE和DOCUMENT字段; 上述函数之外的其他函数则可作用于CHAR、PHRASE和DOCUMENT字段。^[4]
- TRS全文数据库系统支持辅助知识自动扩展检索, 这一功能涵盖了大部分的检索函数, 包括: BT、NT、PT、UF、AB、ABF、RT、LE、ALLBT、ALLNT、ST、AT。自动扩展检索可以对检索表达式里的所有词自动施加这些函数调用, 无须显式调用。
- 检索函数BETWEEN和BEYOND的参数表的圆括号‘(’和‘)’可以分别使用方括号‘[’和‘]’替换。

1. 对于BETWEEN函数，‘(’表示值1不包含在目标值范围之内，‘[’表示值1包含在目标值范围之内；‘)’表示值2不包含在目标值范围之内，‘]’表示值2包含在目标值范围之内。
 2. 对于BEYOND函数，‘(’表示值1不排除在目标值范围之外，‘[’表示值1排除在目标值范围之外；‘)’表示值2不排除在目标值范围之外，‘]’表示值2排除在目标值范围之外。
 3. 如果值1和值2都是精确值，或者都不是以模糊匹配开头的键值，则内部会自动按大小调换位置，小的作为值1，大的作为值2；否则内部不调换值1和值2的位置。检索把第一个与值1匹配的字段值，以及第一个与值2匹配的字段值分别作为函数检索范围的两个控制端点。
- 函数BT、NT、PT、UF、AB、ABF、RT、LE、ALLBT、ALLNT、ST、AT的参数中的‘[’和‘]’不是参数的一部分，表示词典可以省略。当词典省略时，系统将根据被检索的数据的语言类型，获取系统缺省的相应类型的词典作为被引用的词典，其中函数ST所引用的词典必须是同义词典，函数AT所引用的词典必须是反义词典，其它函数所引用的词典必须是主题词典。如果指定词典，则无论是函数调用，还是自动扩展检索，都要求词典的字符集必须与所检索的数据库的字符集保持一致，否则无效。
 - 在传统用法的RANGE函数中，比较运算符‘=’可以用其他任意比较运算符替代。RANGE函数用于检索文本中的数值范围时，参数表的圆括号‘(’和‘)’可以分别使用方括号‘[’和‘]’替换，其意义与BETWEEN函数相同。
 - DATA函数的参数形式与8.2节中“检索键值列表”的形式是一致的。参数表的圆括号对‘(’和‘)’可以使用方括号对‘[’和‘]’替换，表示：作用于CHAR型字段时，不区分字母的大小写；作用于DATE型字段时，进行精确匹配。如果DATA函数作用于PHRASE、DOCUMENT和BIT字段，则只允许无参数调用。
 - 函数EXACT和DATA检索函数可用于进行字段值大小比较的表达式中，但如果DATA函数用于除“=”或“!=”之外的表达式中时，其参数只能是单个值，且认为是精确检索键值，例如表达式：“字段名>function(...)”。除上述函数之外的其它函数则只能用于“=”或“!=”的表达式中。
 - LIKE函数主要用来改变全文检索的方式。常规的全文检索方式是：对于给定的短语（词串），经过分词处理后，各词之间要进行严格的位置运算。而LIKE函数则是对参数给定的短语（词串），经过分词处理后，各词之间不进行位置运算，而只按下列四种模式进行简单的逻辑运算：
 1. 『完全与』模式。参数表以圆括号对‘(’和‘)’表示，即LIKE(...)；第二个参数表示是否忽略二元索引：取值为‘TRUE’（大小写无关）或者‘1’、或者缺省时，表示忽略二元索引，以提高查全率，尽可能地取得更多的结果；取值为‘FALSE’（大小写无关）或者‘0’时，表示不忽略二元索引，以提高检索速度。这种模式允许附加所有合法的限定运算符和加权运算符。
 2. 『有限与』模式：参数表以圆括号对‘(’和‘)’表示，即LIKE(...)；第二

个参数，如果是2到^{【修订115】}19^{【9】}之间的数字，则表示检索结果限定在“忽略二元索引，并且在限定范围内出现的词数不能少于该参数所对应的数目”，如果是^{【修订116】}20^{【9】}到100之间的数值，则表示检索结果限定在“忽略二元索引，并且在限定范围内出现的词数与第一个参数所含总词数的百分比不能少于该参数所对应的数值”。其中的限定范围由附加的属性运算决定，缺省限定范围是记录。这种模式允许附加所有合法的限定运算符和加权运算符。

3. 『完全或』模式：参数表以方括号对‘[’和‘]’表示，即LIKE[...]；第二个参数的意义与『完全与』模式相同。除了“/seg”和“/sen”外，这种模式允许附加其它合法的限定运算符和加权运算符。
4. 『有限或』模式：参数表以方括号对‘[’和‘]’表示，即LIKE[...]；第二个参数的意义与『有限与』模式相同。这种模式允许附加所有合法的限定运算符和加权运算符。

这里，“与”模式与“或”模式的区别在于限定运算符“/sen>n”和“/seg>n”的解释不同（其中的“>”可用“>=”、“<”、“<=”、“=”和“!=”替代）。“与”模式要求相关的词必须出现同一句/段中，并且其句/段号必须满足给定的条件；但“或”模式不要求相关的词出现在同一句/段中，而只要求其句/段号都满足给定的条件。对于某些限定运算，『有限与』模式和『有限或』模式是等价的，例如：“LIKE(词串, 80)/sen”等价于“LIKE[词串, 80]/sen”。**这里需要特别指出的是：对于限定运算符“/pos=n”，如果有“/sen”显式的同句限定，或者在“与”模式下有“/sen=n”隐式的同句限定，则意义发生改变，表示词之间的相对距离；否则仍然仍维持原意，表示词自身的位置。**

『有限与』模式的作用是扩大『完全与』模式的结果集，以满足更宽泛的查询需求；而『有限或』模式的作用则是缩小『完全或』模式的结果集，以满足更精确的查询需求。

LIKE函数支持span限定，即“跨度检索”，例如：“LIKE(词串, 80)/span<=20”，其含义是：对于“LIKE(词串, 80)”得到的一组命中词进行进一步限定，即这些命中词必须是出现在（同段）同句中，且第一个命中词与最后一个命中词的位置跨度不能大于20。^{【修订122】}

LIKE函数与TOP N剪裁排序（请参见9.1.3节）配合使用时，还有一个很实用的优化策略，即“匹配度自适应调节”，这一策略的实现过程是：LIKE函数的匹配度从一个较低的规定值（第二个参数）开始，在运算过程中，当符合当前匹配度的命中记录数达到TOP N所要求的最少记录数时，系统将提高自动提高匹配度以尽早地淘汰只有较低匹配度的记录，保留具有较高匹配度的记录，如此迭代直到检索完成，得到一个最小的且具有最高匹配度的结果集。^{【修订128】}

另外，如果不希望单个字参与检索运算，则可在第二个参数前加一个‘w’字母，如“LIKE(词串, w80)”，表示分词后忽略单个字的项。^{【修订092】}

- INCLUDE函数如果只有一个参数，则是传统用法；^{【修订125】}如果有两个参数，则是扩展模式，即：INCLUDE(短语串组, P)。在使用上，扩展模式的INCLUDE函数与

LIKE函数只有以下五点区别:

1. INCLUDE函数可以作用于CHAR字段, 而LIKE函数则不能作用于CHAR字段, 原因是CHAR字段不进行分词处理。
2. INCLUDE函数没有『完全与』和『完全或』模式, 只支持『有限与』和『有限或』模式。
3. 函数INCLUDE(短语串组, P)的参数“短语串组”是由两个以上的短语串组成的, 短语串之间用空格分隔, 如果短语串本身包含空格, 则该短语串必须用单引号括起来。
4. 函数INCLUDE(短语串组, P)的参数P, 即“匹配度”的范围是1~100, 其中: 1~19表示结果记录中出现的短语串数目不能少于该参数值; 20~100表示结果记录中出现的短语串数目所占参数“短语串组”里包含的短语串数的百分比不能少于该参数值。这是常规意义的“匹配度”参数, 它以短语串为单位, 而LIKE函数的“匹配度”则以“词”为单位。
5. INCLUDE函数^{【修订147】}支持更复杂的“匹配度”参数, “匹配度”参数的格式包括: N#M:P, 或者N:M:P, 或者N:M#P, 或者N#M#P, 或者N#M, 或者N:M。例如: INCLUDE(短语串组, N:M:P)。

这些格式的“匹配度”参数中的P仍对应于常规意义的“匹配度”参数意义, 也就是上述INCLUDE(短语串组, P)中的“匹配度”参数P, 它可以省略, 此时默认为100, 即所有短语串都必须同时有命中。

复杂化的“匹配度”参数的作用是要改变INCLUDE函数的运算逻辑: 把“短语串组”参数中的短语串分成两类, 即短串和长串, 以采用不同处理逻辑, 优化查询结果。如果一个短语串的字数不超过参数里的N值, 则认为是短串, 否则为长串。短串的查询处理逻辑将更注重词之间的位置关系, 而长串则按函数LIKE(长串, M)或LIKE[长串, M]的逻辑处理(『与』『或』模式继承于INCLUDE函数^{【1】}。

特别强调: 在任何情况下, 无论是短串还是长串, 只要用单引号括起来, 则总是严格按“短语”逻辑处理, 即分词且要求字词的位置必须相邻。

相对来说, 短串的处理逻辑要复杂些。^{【修订158】}对于N#M:P, 或N#M#P, 或N#M, 短串则严格按“短语”逻辑处理; 而对于参数N:M:P, 或N:M#P, 或N:M, 则会采用下面兼顾结果集与相关性的优化逻辑进行处理:

- 如果N大于或等于4, 且短串的字数小于4, 则严格按“短语”逻辑处理。
- 如果N大于或等于4, 且如果短串的字数等于4, 则按带位置限定的LIKE函数处理, 即‘LIKE(短串, X)/SPAN<12’。其中X是处理短串时LIKE函数的“匹配度”, 如果参数是N:M#P, 则X等于M, 其它参数时则X等于100; LIKE函数的『与』『或』模式继承于INCLUDE。
- 如果N大于4, 且短串的字数是5~N, 则按带位置限定的LIKE函数处理, 即‘LIKE(短串, X)/SPAN<50’。与上述情形相比, 只是扩大了位置的限定范围, 其它逻辑与字数为4的短串处理一致。

➤ 如果N小于4，则短串严格按“短语”逻辑处理。^[4]

实现扩展模式的INCLUDE函数的目的是为了更方便地实现主流Web方式的检索。通过设置合适的参数，可以实现如下Web高级搜索中的常见功能：

✧ 包含全部字词：第二个参数取100。

✧ 包含至少一个（或几个）字词：第二个参数取需要的命中词数。

✧ 不包括字词：% - INCLUDE(短语串组, 1)。^[4]

- SPELL函数是用于拼写组合的检索函数。拼写组合有以下三种规则：
 1. 连接规则“{X}” 该规则描述规则符号“{X}”前后字符串的连接扩展。规则符号中的花大括号“{ }”（单字节ASCII码）是固定的语法符号；“X”是一个或多个非语法符号（包括空白符），表示规则符号前后的字符串可由“X”连接（当X不是空白符时）、或由一个空格分隔、或没有任何连接符号。例如：SPELL(Vitamin{-}B)将扩展组合成“Vitamin-B”、“Vitamin B”和“VitaminB”三个字符串。
 2. 排除规则“{%,X}” 该规则描述规则符号“{%,X}”前后字符串之间应该排除的符号。规则符号中的花括号“{ }”和逗号“,”（单字节ASCII码）是固定的语法符号；“%”表示模糊匹配，也可以是多个“%”和“?”（单字节ASCII码）的组合；“X”是一个或多个非语法符号组成的需要被排除的组合。例如：SPELL(Dat{%,e})表示“以Dat开头的，但不包括Date的单词”。
 3. 选择规则“{X,Y}” 该规则描述规则符号“{X,Y}”前后字符串之间可以进行的多种符号组合的选择。规则符号中的花大括号“{ }”和逗号“,”（单字节ASCII码）是固定的语法符号；“X”和“Y”都是一个或多个非语法符号组成的供选择的组合，“X”和“Y”其中之一可为空白，但不能两个都为空白。例如：SPELL(Colo{,u}ri{s,z}ing)将组合成“Colorising”、“Colorizing”、“Colouising”和“Colourizing”四个单词。
- NEAR函数用于索引词的非精确检索。该函数以第一个参数给出的“目标词”为中心，按索引的组织内部顺序进行搜索，输出“目标词”前后总数由第二个“数量”参数指定的一组索引词。

7.3 检索键值

TRS 既支持在 DATE、NUMBER 和 CHAR 字段上通过字段值进行数据库查询，也支持在 PHRASE 和 DOCUMENT 字段上通过任意的自由字、词、短语及数值进行检索。通过 DATA 函数还可以进行空值检索。

7.3.1 日期型键值

在检索表达式中，DATE 键值的格式有通用格式、英文格式、无分隔符格式和中

文格式四种类型。下列格式中，用‘[’和‘]’包括的部分表示可以省略。

● 通用格式

Year[/Month[/Day[/Hour[/Minute[/Second]]]]][A[M]]
[Year]/Month[/Day[/Hour[/Minute[/Second]]][A[M]]
[Year]/[Month]/Day[/Hour[/Minute[/Second]]][A[M]]
[Year]/[Month]/[Day]/Hour[/Minute[/Second]][A[M]]
[Year]/[Month]/[Day]/[Hour]/Minute[/Second][A[M]]
[Year]/[Month]/[Day]/[Hour]/[Minute]/Second[A[M]]

其中：

- ◇ Year为一到四位表示的完全年份(如“2008”)。
- ◇ Month、Day、Hour、Minute、Second分别为两位或一位数字表示的月、日、时、分、秒。
- ◇ A/M可用P/PM替换。A/AM表示上午，P/PM表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能缺省。
- ◇ 单字节除号‘/’分隔符可分别用单字节的逗号‘,’、句号‘.’、冒号‘:’以及减号‘-’任意替代，在两个明确(不省略)的日期成份之间，空格也可作为分隔符。
- ◇ 用方括号引起的部分表示可以省略(模糊日期)。
- ◇ Year、Month、Day、Hour、Minute、Second都可以用一个或多个单字节的问号‘?’、百分号‘%’替代，表示模糊日期。

● 英文格式

Month[Day[,Year[Hour[:Minute[:Second]]]]][A[M]]
Month[Day],Year[Hour[:Minute[:Second]]][A[M]]
Month[Day],Year[Hour]:Minute[:Second][A[M]]
Month[Day],Year[Hour]:[Minute]:Second[A[M]]
[Day-]Month[-Year[Hour[:Minute[:Second]]]]][A[M]]
[Day-]Month[-Year[Hour]:Minute[:Second]][A[M]]
[Day-]Month[-Year[Hour]:[Minute]:Second][A[M]]
DayMonthYear[Hour[:Minute[:Second]]][A[M]]
DayMonthYear[[Hour]:Minute[:Second]][A[M]]
DayMonthYear[[Hour]:[Minute]:Second][A[M]]

其中：

- ◇ Year为一到四位数字表示的完全年份(如“2008”)。
- ◇ Month为英文表示的月份名：Jan. Feb. Mar. Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec. (其最后的点号可以省略)；或：January February March April May June July August September October November December。
- ◇ Day、Hour、Minute、Second分别为两位或一位数字表示的日、时、分、秒。

- ◇ A/M可用P/PM替换。A/AM表示上午，P/PM表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能缺省。
- ◇ Year与Hour之间用单字节的空格分隔；‘,’ ‘:’ ‘-’ 分别为单字节的逗号、冒号、减号；除Year与Hour之间的空格外，其它空格不是分隔符，被忽略，但一个日期成份中间不能出现空格；Day与Year之间的逗号‘,’可用减号‘-’替代。
- ◇ 用方括号引起的部分表示可以省略（模糊日期）。

● 无分隔格式

YearMonth[Day[Hour[Minute[Second]]]]

其中：

- ◇ Year为固定四位表示的完全年份(如“2008”)。
- ◇ Month、Day、Hour、Minute、Second分别为两位数字表示的月、日、时、分、秒。
- ◇ 用方括号引起的部分表示可以省略（模糊日期）。

● 中文格式

中文格式包括所有用全角字符表示的上述几种格式，另外还支持如下的格式：

[Year年][Month月][Day日][Hour时][Minute分][Second秒][上午]

其中：

- ◇ Year为一到四位数字表示的完全年份(如“2008”或“二〇〇八”或“贰零零捌”)。其中数字既可为半角，也可为全角。
- ◇ Month、Day、Hour、Minute、Second分别为两位或一位数字表示的月、日、时、分、秒。
- ◇ 格式中“上午”可用“下午”替换。如果“上午”或“下午”出现在格式中，则“Hour时”不能缺省。
- ◇ 用方括号引起的部分表示可以省略（模糊日期）。

另外，还支持\$DATE 和\$TIME 特殊日期（不区分大小写）的检索，其中前者表示当前日期的“年月日”，后者表示当前日期的“年月日时分秒”。这里\$DATE 和\$TIME 也可以用作函数来表示一个相对于“现在”的日期和时间，即：\$DATE(...)和\$TIME(...)，它们的参数都是一个正整数（表示未来）、或者负整数（表示过去），\$DATE(...)的参数单位是“天”，\$TIME(...)的参数单位是“秒”。如：\$DATE(7)、\$DATE(-7)、\$DATE(+7)，其中\$DATE(+7)等价于\$DATE(7)；\$TIME(60)、\$TIME(-60)、\$TIME(+60)，其中\$TIME(+60)等价于\$TIME(60)；\$DATE(0)、\$DATE(-0)、\$DATE(+0)和\$DATE()等价于\$DATE；\$TIME(0)、\$TIME(-0)、\$TIME(+0)和\$TIME()等价于\$TIME。

对于一个不完全的日期，TRS 全文数据库系统缺省以模糊日期处理，如日期“2008”将检索到“2008.1.1”到“2008.12.31”以及“2008”中没有确定的“月”和“日”的所有日期。如果需要将“2008”当作一个精确日期来处理，即不要命中有确定的“月”

或“日”的日期，则可使用“EXACT”函数进行检索，如“EXACT(2008)”。

在阅读《数据库加载》部分时会发现，其日期格式中的“年份”与这里所要求的不一样：加载时的“年份”可以省略“世纪数”，而检索时的“年份”没有缺省“世纪数”，无论多少位都表示的是完全年份，如“08”表示“公元8年”，而不是“公元2008年”。在检索时，无分隔格式中的“年份”必须是四位，如日期“080808”表示“808年8月”，但在数据库加载时则表示“2008年8月8日”（假设缺省世纪数为20）。之所以存在这种差别，主要原因是在数据库加载时，往往是大批量的日期数据，通过加载控制文件指定年份的缺省世纪数可以支持按习惯方法表示的日期，而检索时则没有必要使用这种方法指定年份的缺省世纪数，直接写出完全的年份更为简捷方便。

7.3.2 数值型键值

在检索表达式中，NUMBER型键值是一个实数，不支持模糊检索。

7.3.3 字符型键值

在检索表达式中，CHAR键值是一个字符串。这个字符串可以单引号引起来，也可以不用。对于普通的键值，是否被引号引起来，或是否被转义，都具有相同的含义。

一般情况下，给出的键值必须是一个完整的字段值，而不是字段值的一部分，但TRS也支持对键值的模糊匹配。用通配符“?”表示一个字母字符或汉字，用“%”表示0到多个字母字符或汉字。通配符可出现在键值中的任意位置。^{【修订001】}这两个通配符的组合‘%??...?’表示在字符串中至多可以通配n个（即0~n个）字母字符或汉字；而其它组合，例如‘??...?%’，则表示在字符串中至少需要通配n个（即≥n个）字母字符或汉字，其中的n都是‘?’的个数^{【1】}。但是如果一个给定的键值只包含有‘?’和‘%’，而没有其它任何字符，则无论其格式如何都表示全模糊查找。

例如：

- 表达式“作者=%国庆”将检索出所有作者为“张国庆”、“欧阳国庆”、“国庆”等的记录；
- 表达式“作者=张%”将检索出所有作者为“张国庆”、“张三”、“张”等的记录；
- 表达式“作者=张%庆”将检索出所有作者为“张国庆”、“张大庆”、“张庆”等的记录；
- 表达式“作者=张?庆”将检索出所有作者为“张国庆”、“张大庆”等的记录，而作者为“张庆”的记录不会被检索出来。

另外，还支持\$USER特殊键值（不区分大小写）的检索，表示检索字段值等于当

前操作的用户名。

7.3.4 短语和全文型键值

在检索表达式中, PHRASE 和 DOCUMENT 的键值是一个字、词、短语, 或者句子。它们可以单引号引起来, 也可以不用。对于普通的键值, 是否被引号引起来, 或是否被转义, 都具有相同的含义, 并且是经过分词处理后再进行检索。

对于含有“?”或“%”的模糊键值, 也需要先分词再检索, 因此这类键值的检索要比 CHAR 键值的检索复杂得多, 其中“?”和“%”的意义可能不完全等同于 CHAR 键值中规定的意义, 比如: “张?庆”将检索出包含“张国庆”、“张大庆”等的记录, 同时也将检索出包含“张庆”的记录。

但是, 如果模糊键值在分词词典中有对应的词(即用该键值查询分词词典时有结果), 则“?”和“%”的意义与 CHAR 键值的相同。例如: “中国”、“中立国”和“中华民国”都是词, 则“中?国”不能命中“中国”和“中华民国”。

7.3.5 特殊键值

TRS 检索表达式中有四类语法符号:

1. 第一类符号: ‘(’、‘)’、‘[’、‘]’、‘,’、‘/’、‘@’、‘=’、‘>’、‘<’、‘!’、‘&’、‘*’、‘^’、‘-’、‘+’。
2. 第二类符号: ‘ADJ’、‘EQU’、‘PRE’、‘AND’、‘XOR’、‘NOT’、‘OR’、‘TO’、空格、函数名。
3. 第三类符号: 模糊匹配符 ‘%’、‘?’。
4. 第四类符号: 单引号 ‘’ 和转义符 ‘\’。

检索表达式的键值必须对这些符号进行特殊处理:

- 如果键值中含有第一类语法符号, 则必须将该键值用单引号引起来, 或者用转义符将这些语法符号进行转义。如: 日期键值 ‘,8’、\\,8。
- 如果第二类语法符号是一个键值, 则必须将该键值用单引号引起来, 或者用转义符将这些语法符号进行转义。如: ‘AND’、\AND。
- 如果第三类或第四类语法符号是键值的一部分, 则必须使用转义符将这些语法符号进行转义。如: 10\%、\\。

需要注意的是:

1. 对于其它一些非文字符号, 如果需要在 PHRASE 或 DOCUMENT 字段中能够被检索, 则应该将这些符号从数据库所引用的停用词典中去掉。

2. 特别地，单字节的“惊叹号”、“逗号”、“分号”、“句号”和“问号”，以及各语种自己的多字节编码的“惊叹号”、“逗号”、“分号”、“句号”和“问号”是分词系统法定的“分句”符号，一般情况下会被忽略。如果需要在 PHRASE 字段中能够被检索，则也应该将这些符号从数据库所引用的停用词典中去掉；如果需要在 DOCUMENT 字段中能够被检索，则不但应该将这些符号从数据库所引用的停用词典中去掉，而且 DOCUMENT 字段的索引属性中不能包含有位置属性“SEN”。请参见 3.4.1 节。
3. 如果不是特别需要，则即使是按字索引，也建议使用停用词典停用词典，以阻止哪些没有检索意义的符号。

第 8 章 表达式构造

表达式包括统计表达式和检索表达式。统计表达式告诉数据库如何统计检索，而检索表达式则是告诉数据库如何挑选数据记录以得到检索结果。

8.1 统计表达式

统计表达式分为两类：一类是计算总的命中结果记录数，即缺省形式，其表示非常简单，只有一个单字节的星号 ‘*’；另一类是对数值，或利用统计函数对 NUMBER 型字段在命中结果集上进行算术运算，这类统计表达式较为复杂。

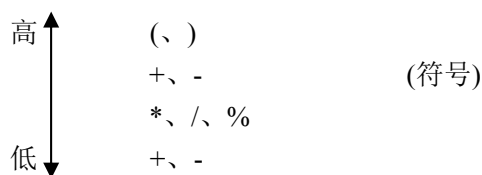
- “统计表达式”的组成：
 - *
 - 子统计表达式
 - 子统计表达式, 子统计表达式, …… , 子统计表达式

其中，单字节的逗号 ‘,’ 是一种语法符号，是“统计表达式”的一部分，用于分隔多个“子统计表达式”。省略号 “……” 表示可有多个“子统计表达式”。每个“子统计表达式”都会有一个统计结果。

- “子统计表达式”的组成：
 - 数值
 - 统计函数
 - -子统计表达式
 - (子统计表达式)
 - 子统计表达式 ArithOP 子统计表达式

“子统计表达式”的组成具有自递归的特点。其中“ArithOP”代表所有的算术运算符。单字节的 ‘-’ 是一元“负”运算。单字节的括号 ‘(’ 和 ‘)’ 用以改变运算优先级，括号中的“子统计表达式”优先计算。

- 统计表达式中算术运算符的优先级为：



注：同级运算遵循左结合律。

8.2 检索表达式

检索表达式的构成非常复杂，总的来说包含这样一些基本要素：检索键值、检索函数、检索历史引用、物理记录号、检索键值列表、物理记录号列表以及各种运算符。

- “检索表达式”的组成：
 - 逻辑运算表达式
 - 限定运算表达式

- “逻辑运算表达式”的组成：
 - ENTRY = (字段逻辑表达式)
 - ENTRY != (字段逻辑表达式)
 - ENTRY CompOP 检索键值
 - ENTRY, ENTRY, ……, ENTRY RepeatOP (字段逻辑表达式)
 - DOCID = 记录号列表
 - DOCID != 记录号列表
 - DOCID CompOP 记录号
 - @检索历史号
 - 限定运算表达式
 - (逻辑运算表达式)
 - 逻辑运算表达式 LogicOP 逻辑运算表达式

“逻辑运算表达式”的组成具有自递归的特点。其中，“ENTRY”是字段名或字段别名，单字节的逗号‘,’是语法符号，是“检索表达式”的一部分，用于分隔多个“ENTRY”，省略号“……”表示可有多个“ENTRY”。“CompOP”代表所有的比较运算符。“RepeatOP”代表所有的重复逻辑运算符。“LogicOP”代表所有的逻辑运算符。单字节的‘@’是检索历史引用符。单字节的括号‘(’和‘)’用以改变运算优

优先级，括号中的表达式优先计算。

● “字段逻辑表达式”的组成：

- 检索键值
- 检索函数
- 检索键值列表
- 限定运算表达式
- (字段逻辑表达式)
- 字段逻辑表达式 LogicOP 字段逻辑表达式

“字段逻辑表达式”的组成具有自递归的特点。其中，“LogicOP”代表所有的逻辑运算符。单字节的括号‘(’和‘)’用以改变运算优先级，括号中的表达式优先计算。

● “检索键值列表”的组成：

- 检索键值
- 检索键值 TO 检索键值
- 检索键值列表, 检索键值列表, …… , 检索键值列表

“检索键值列表”的组成具有自递归的特点。其中，单字节的逗号‘,’是语法符号，是“检索键值列表”的一部分，用于分隔多个“检索键值列表”，省略号“……”表示可有多个“检索键值列表”。“TO”表示一个连续的检索键值区间，等价于检索函数 BETWEEN[]。

对于多个键值的检索，应该优先使用“检索键值列表, 检索键值列表, …… , 检索键值列表”或者“检索键值 TO 检索键值”或者 BETWEEN 函数的高效形式，而不要采用将多个“ENTRY = 检索键值”子表达式进行逻辑运算的低效形式。

● “记录号列表”的组成：

- 记录号
- 记录号 TO 记录号
- 记录号列表, 记录号列表, …… , 记录号列表

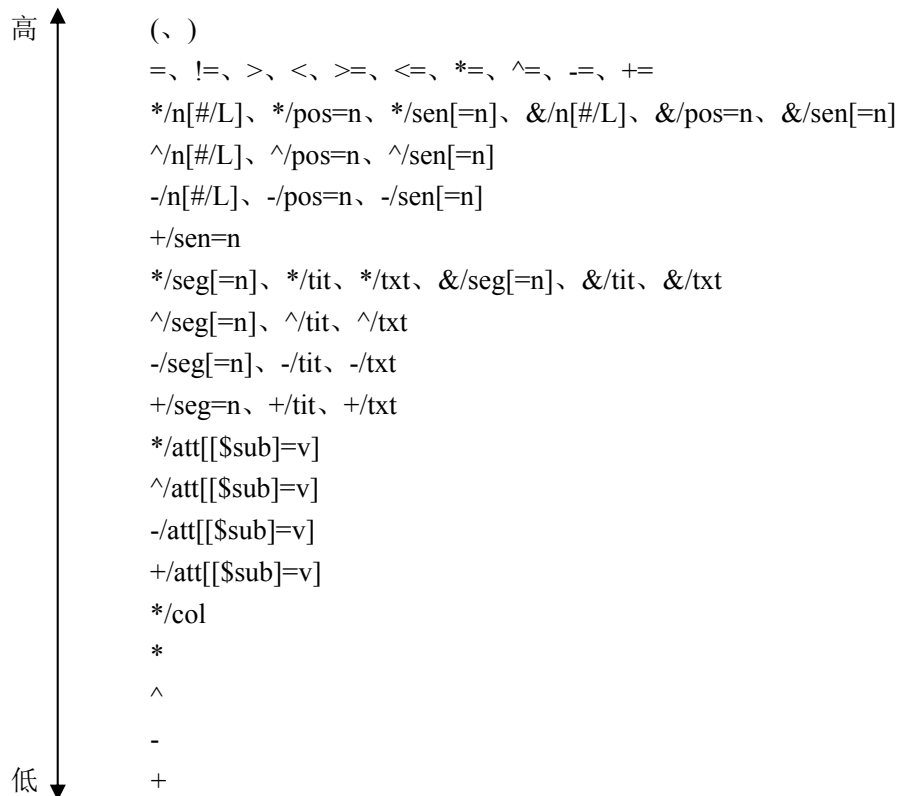
“记录号列表”的组成具有自递归的特点。其中，单字节的逗号‘,’是语法符号，是“记录号列表”的一部分，用于分隔多个“记录号列表”，省略号“……”表示可有

多个“记录号列表”。“TO”表示一个连续的记录号区间。

- “限定运算表达式”的组成：
 - 检索键值
 - 检索函数
 - 检索键值/Limit
 - 检索函数/Limit
 - 限定运算表达式 LimitOP 限定运算表达式
 - (限定运算表达式)

“限定运算表达式”的组成具有自递归的特点。其中，“LimitOP”代表所有的属性运算符。“/Limit”代表所有用于检索的限定运算符和加权运算符。单字节的括号‘(’和‘)’用以改变运算优先级，括号中的表达式优先计算。

- 检索表达式中各种运算符的优先级为：



注：同级运算遵循左结合律。限定运算符和加权运算符与检索词是不可分割，因

此与其它运算符不存在优先级的问题，限定运算符之间也没有运算优先级。

这里需要特别提醒的是：^[修订 018]比较运算符（‘=’、‘!=’、‘>’、‘<’、‘>=’、‘<=’）和重复逻辑运算符（‘*=’、‘^=’、‘-=’、‘+=’）的优先级仅次于括号 ‘(’ 和 ‘)’ 的优先级，因此，如果这些运算符的右边是一个表达式，则必须用括号括起来，否则产生的结果就不是你需要的结果^[4]。

8.3 缺省字段检索

在“逻辑运算表达式”的组成中，子表达式“ENTRY = (字段逻辑表达式)”可以省略成“(字段逻辑表达式)”，这时表示在系统默认的缺省检索字段上（参见 1.3 节的说明），或者在检索接口中由外部指定的一个或多个“缺省字段”上进行检索。

如果外部指定了多个“缺省字段”，则必须同时指定这些字段上的命中结果之间的逻辑关系，系统自动在这些“缺省字段”上进行扩展检索。这种扩展有两种方式：一种方式是指定的缺省字段检索的检索逻辑具有最高运算优先级，即缺省检索字段检索逻辑在每个检索原子（检索键值）上进行扩展，称为“**字段逻辑优先**”扩展；另一种方式是优先处理检索表达式中的逻辑，即缺省检索字段检索逻辑将在尽可能大的子检索表达式上进行扩展，称为“**表达式逻辑优先**”扩展，系统缺省进行这种扩展。

这两种扩展在原理上是不等价的，其检索结果也完全有可能不一样。“字段逻辑优先”扩展的效果是：所有的缺省检索字段被捆绑成了一个整体，检索结果遵循在单个字段上检索时的规律。举例说明：假设指定的缺省检索字段是 c1 和 c2，c1 和 c2 检索结果之间的逻辑是“or”，且检索的目标表达式为“A*B”，则“字段逻辑优先”扩展后的检索表达式为“(c1,c2+=A)*(c1,c2+=B)”，而“表达式逻辑优先”扩展后的检索表达式为“c1,c2+=(A*B)”。

第 9 章 检索结果处理

检索结果的处理是指按照检索表达式给出的逻辑完成数据库检索并形成检索结果集后,对结果集的进一步操作。本章主要介绍检索结果的排序和按字段值统计分类两方面的内容。

9.1 排序

在缺省情况下,检索结果集里的记录是按记录入库的先后顺序存放的,如果要改变这种顺序,则必须对检索结果进行排序。TRS 全文数据库系统支持字段值排序、相关性排序和 LIFO 排序。

9.1.1 排序向量

按字段值排序时,不仅可以指定一个字段作为排序对象,也可以同时指定多个字段作为排序对象,它们按照一定的顺序排列后就构成了排序向量。支持按字段值排序的字段类型是 DATE、NUMBER、CHAR 和 PHARSE,其它类型不支持。

在排序向量中,排序优先级顺序是从左到右的,最左边的字段最先排序,最右边的字段最后排序。只有当一个字段有多个记录具有相同的值时,才会对该字段右边的字段继续排序。

相关度(相关性)可以当作一个字段放在排序向量中,但对于一个查询来说,各个记录之间很少会有相同的相关度,所以排在相关度右边的字段在排序时所起的作用可能不大。相关性的排序固定采用降序,而字段值则可以指定是升序排列,还是降序排列。

LIFO 排序与其它排序是互斥的,也就是说要么按 LIFO 排序,要么按指定的排序向量排序。LIFO 排序后输出结果的顺序与缺省情况下(不排序时)输出结果的顺序正好相反,即最后入库的记录最先输出。

9.1.2 相关性排序

除按字段值和 LIFO 排序外,TRS 全文数据库系统还支持按相关性对检索结果进

行排序。相关性排序的前提是记录的相关度计算。TRS 全文数据库系统是采用基于向量余弦相似性模型的改进算法来计算一个文档的相关度的，其基本要素有三点：TF（词频，词在文档出现的次数）、IDF（逆文档频率，总文档数与命中文档数之比），以及文档的规范化长度。这也就是通常所说的 TF*IDF 相关度加权模型。

IDF 的加入使得相关度与数据库相关，即同样的数据在不同的数据库中，对于相同的检索条件，其表现出的相关度会不一样。有的应用可能需要消除这种特性，目前唯一的方法是计算相关度时忽略 IDF，请参见 20.1.3 节有关配置项 RelevanceMode 的说明。

9.1.2.1 相关度的加权

前面所介绍的 TF*IDF 模型是最基本的相关度加权模型，在实际应用中，这种加权的不是特别的理想，往往满足不了应用的需求；一些应用还希望特别强调“文章标题”部分的权重比例等。^{【修订 133】}因此，TRS 系统在计算相关度时考虑了多方面的加权因素，以其得到更好的排序效果：

- 标题加权。这里的标题不是一个字段，而是指文章的开始部分，一般就是第一句话。对出现在标题里的命中点，应用可以指定相应的加权系数。
- 相对位置加权。考虑命中点之间的距离因素，使得来自于短语的命中点具有更高的相关性权重。
- 分支加权。对表达式中的分支（叶子节点）命中情况，以及 LIKE 和 INCLUDE 函数的“符合度”进行加权。这种加权有可能使得输出检索结果时，其最大相关度不是 1。^{【2】}

相关度计算还支持外部指定的记录级加权：每个记录都事先赋予了一个相关性权值，以一个字段值的形式和记录的其它字段数据一起存入数据库中，并要求在进行相关性排序时，将这个权值^{【修订 112】}乘以^{【2】}相关度模型的计算结果后作为该记录的相关度。

9.1.2.2 按命中词数量排序

在一般情况下，如果检索条件涉及缺省检索字段，且指定按相关性排序，则系统将通过上述方法得到记录的相关度后进行缺省方式的相关性排序。在 TRS 全文数据库系统里，这是一种缺省的相关性模式。

除此之外，TRS 全文数据库系统还支持一种扩展的相关性模式，即命中词的单位向量长度相关性模式，这种模式的实质是在进行相关性排序时，用一个记录中命中词的单位向量长度（每个命中词只计算一次）作为记录的相关度，含命中词的个数（注意：不是词出现的次数）越多的记录，其相关度越大。在此基础上还有一种衍生相关

性模式，那就是当多个记录的命中词的单位向量长度相等时，这些记录将再按命中词在命中点出现的次数总和的降序排列。

9.1.2.3 多字段相关性排序

前一节所述扩展是在一个字段上对相关性模式进行的扩展，对相关性模式扩展的另一个方向是支持多字段的相关性排序，不仅支持 PHRASE 和 DOCUMENT 字段，而且还支持 DATE、NUMBER 和 CHAR 字段。多字段的相关性排序不是分别对多个字段计算相关度后每个字段分别对应排序向量的一个分量进行排序，而是将多个字段的相关度进行加权求和后的综合值对应排序向量中的唯一（相关性）分量进行排序。除缺省检索字段（参见 1.3 节）外，其它字段要参与相关性排序，必须打开字段的“相关性开关”（参见 1.4 节），并在检索时指定该字段的权重（参见 6.2.6 节），即使检索只涉及一个字段时，也必须指定。但如果检索时所有涉及的字段都未显式地指定权重，且缺省检索字段又是检索所涉及的字段之一，则只计算缺省检索字段上的相关度，并进行相关性排序。【修订 020】

对于多字段的相关性排序，仅仅通过对字段进行加权的方法有时还不足以满足所有应用的需求，因此系统提供了另一种选择，即允许规定字段的排序优先级，实现按字段分别进行相关性排序的应用需求。这一功能的使用方法是：在提交检索请求时指定“按字段分别进行相关性排序”的参数（具体参数需要参见所使用的接口说明），同时按上述方式指定字段的“权重”。但这里的“权重”仅代表字段之间相对的排序优先顺序，不再有绝对值意义，“权重”越大，排序越优先。【修订 126】

9.1.2.4 焦点排序

对于新闻数据，一个不能忽视的特殊性就是时效性。一条新闻的重要性会随着时间的推移而逐渐降低，因此其相关性排序必须考虑结果记录的时间因素，这就是通常意义的按“焦点排序”，TRS 全文数据库系统也支持了这种排序方式。【修订 141】

9.1.3 TOP N 排序

TRS 全文数据库系统支持对检索结果进行 TOP N 排序。TOP N 排序分为两种模式：一种是 **TOP N 部分排序**，即排序后的结果集只有前 N 个记录是有序的（指定的序），其余的记录虽然也被保留，但是无序的；另一种 **TOP N 剪裁排序**，即排序后的结果集只保留（存储）前 N 个有序的（指定的序）记录，其余的记录被丢弃。

TOP N 部分排序总是受系统的“最大排序记录数”控制，即当检索结果记录数超出“最大排序记录数”时，TOP N 部分排序不起作用；而当检索结果记录数小于“最

大排序记录数”时，排序结果只保证前 N 个记录是有序的（记录数不发生变化）。

TOP N 剪裁排序的功能要稍微复杂些，分为如下两种情况：第一种剪裁排序是与检索操作一起提交的排序，此时如果检索结果记录数超出“最大排序记录数”，则只在先检索到的“最大排序记录数”个记录中剪裁并保留前 N 个有序（指定的序）的记录，其余的记录根本不参与排序过程（因此在这种情况下，不同的“最大排序记录数”就有可能得到不同的排序结果，不同的检索完成顺序也有可能得到不同的排序结果）；第二种剪裁排序是在检索操作之后的单独排序，这种情况和 TOP N 部分排序一样受系统的“最大排序记录数”控制，即当检索结果记录数超出“最大排序记录数”时，TOP N 剪裁排序将不起作用，否则排序后的结果集只保留前 N 个有序的记录（记录数将发生变化）。【修订 068】

如果在检索时指定 TOP N 剪裁排序，这时表达式中的 LIKE 和 INCLUDE 函数会自动启用“匹配度自适应调节”的检索优化策略（请参见 7.2 节），最大限度地提高系统性能。【修订 129】

9.1.4 中文的排序

在缺省情况下，作为字段内容的中文，其排序是按内码确定顺序的。这种方式在 GB2312 字符集里问题不大，但对于其它编码（GBK、GB18030、UTF8）汉字，按内码顺序的排序结果实际不太符合人们的日常习惯，尤其是 UTF8 编码，排序结果实际是非常的“凌乱”。为此系统实现了以汉语拼音和汉字笔画为基础的中文排序方法，并且支持多音字的分辨。要使用这些方法只需在提交检索请求或排序请求时指定相应的参数（具体参数需要参见所使用的接口说明）即可。【修订 127】

9.2 按字段值分类统计

按字段值对检索结果进行分类，并不是采用文本分类技术，而是按照一个指定的字段的字段值，把检索结果区分为不同的类别。这其中有三个基本功能：一是只统计每一类别中的结果记录数；二是在统计每一类别的结果记录数的同时按类别进行排序后输出；三是每一类别中只存储指定数量的结果记录并按类别进行排序后输出。

上面已经提到，类别是一个指定的字段的字段值。即可以是所有的字段值，也可以只指定其中的某些字段值作为关注的类别，而忽略其它字段值，被关注的类别值之间通过分行符分隔，每行一个值。【修订 021】与此同时，TRS 全文数据库系统还支持通过正则表达式从所有字段值中选取符合条件的字段值作为目标类别，并容许抽取某些匹配项进行组合而形成新的类别，一个正则表达式对应一个被关注的类别。表达式语法规

则请参见附录三。^[4]

正则表达式一般用于 CHAR 型字段。^{【修订 048】}用于对日期字段进行分类统计时，日期将采用缺省格式，即：YYYY.MM.DD HH:MM:SS（如：2007.12.08 12:00:00，都是半角 ASCII 码），“DD”（天）和“HH”（时）之间是空格，正则表达式要基于这种格式进行构造。除此之外，NUMBER 型字段也允许使用正则表达式进行匹配。^[4]

分类统计也支持类别名的“区分大小写”和“不区分大小写”的匹配与输出选择，当类别名选择“区分大小写”时，则严格按“大小写”进行匹配，输出也区分大小写。当类别名选择“不区分大小写”时，则匹配时忽略“大小写”的区别，而输出分为两种情况：如果选取条件是正则表达式，或者没有任何选取条件，则输出总是大写字母；如果选取条件是（被关注的）字段值列表，则输出以这些值的大小写为准。

9.3 快速检索与估算

常规的 TRS 检索是在指定的数据库中检索出所有符合条件的记录，并将完整的结果集返回给应用。^{【修订 132】}然而，有的应用希望先在最新的数据里检索出一部分记录以进行快速展现，并在需要的时候可继续浏览其它的记录。

为了满足上述应用需求，TRS 系统实现了一种快速检索与估算的机制。应用在使用这种机制时，可通过参数指定需要快速返回的最少记录数等要求，然后系统根据这些参数，并按照 LIFO 的原则首先在最后入库的记录中进行检索。当结果记录数达到最低数量要求后，系统将暂停检索，同时根据所处的检索状态进行结果集的估算，最后将估算结果集返回给应用。应用不仅能够从结果集中读取已检索到的记录，而且也能够通过“翻页”的方式获取后续不断补充的检索所得到的记录，并进行更为“精确”的结果估算。

在快速检索与估算的机制下也支持检索结果的排序功能，但因为一次得不到完整的结果集，这种模式下的结果集排序结果不具有全局性，记录之间只是部分有序的。^[4]

第三部分 数据库加载

本部分详细介绍了在 TRS 数据库加载前组织各种数据文件所必须遵循的格式规范。其内容不涉及任何 TRS 数据库加载工具的使用方法，不同应用工具的具体加载过程请参见相应的使用手册。

读者对象

本部分的读者为数据库管理员、具有创建数据库权限以及准备进行数据库加载的 TRS 用户。在阅读本手册之前，建议您先阅读《数据库管理》部分。

内容组织

本手册按如下结构组织：

- 第 10 章 准备加载
- 第 11 章 数据文件
- 第 12 章 多媒体数据的加载
- 第 13 章 日期的加载
- 第 14 章 词典的加载
- 第 15 章 控制文件

第 10 章 准备加载

10.1 数据库的准备

在进行数据加载前，必须选定一个目标数据库。如果数据库还未建立，则必须借助相应的应用工具创建一个 TRS 数据库。下面就 TRS 数据库的一些基本知识作一些简单的描述，详细的内容请参见有关的应用工具用户手册。

TRS 数据库有 6 种不同类型的字段，它们是：

1. DATE 日期型字段
2. NUMBER 数值型字段
3. CHAR 字符型字段
4. PHRASE 短语型字段
5. DOCUMENT 全文型字段
6. BIT 二进制型字段

DATE 型字段只能用来存储“年月日时分秒”这类表示日期的数据。在数据加载时，系统所支持的日期格式请参见第 13 章，如果遇到非日期格式的数据，则系统会进行出错处理。DATE 型字段允许多值。在数据加载时，DATE 型字段有两个特殊的字段值，即\$DATE 和\$TIME（不区分大小写），\$DATE 表示当前日期的“年月日”，\$TIME 表示当前日期的“年月日时分秒”。将\$DATE 和\$TIME 作为字段的缺省值，即可进行相应意义字段值的自动加载。

NUMBER 型字段只能用来存储整数或实数。对于实数，可以限制其小数位的长度，在数据加载时将忽略多余的小数位。整数或实数不能超出机器的单精度浮点数所能表示的范围。NUMBER 型字段允许多值，并且支持五种子类型，即：32 位二进制单精度浮点数（7 位十进制有符号整数）、64 位二进制双精度浮点数（15 位十进制有符号整数）、32 位二进制有符号整数（9 位十进制有符号整数）、64 位二进制有符号整数（18 位十进制有符号整数）以及 32 位十进制有符号整数，详细说明请参见 1.4 节的内容。在数据加载时，NUMBER 型字段接受一个特殊的字段值，即\$CODE（不区分大小写），表示该记录加载到数据库中的唯一自然数顺序码，记录之间以加载的时间先后为序。将\$CODE 作为字段的缺省值，即可完成自动加载。但自然数顺序码不是无限增长的，它的最大位数与该 NUMBER 型字段的子类型所支持的最大十进制整数位数相同。

CHAR 型字段可以用来存储任何形式的字符串。单个字符串的最大长度可以在建库时指定，但不能超过 255 个字符，在数据加载时将截断超过规定长度的字符串，而只保留字符串的前部。CHAR 型字段允许多值。在数据加载时，CHAR 型字段有两个特殊的字段值，即\$USER 和\$CODE（不区分大小写），\$USER 表示当前进行操作的用户名，\$CODE 表示该记录加载到数据库中后的唯一顺序码，记录之间以加载的时间先后为序。将\$USER 和\$CODE 作为字段的缺省值，即可进行相应意义字段值的自动加载。

PHRASE 型字段可以用来存储任何形式、任意长度的文本数据。这种类型的字段最适合作为“标题”这样的字段，对于长文本数据（一般的界定是 64K 个汉字，或 64K 个英文单词，包括标点符号），建议还是使用 DOCUMENT 字段来存储，主要原因是检索时，对超出长度的文本中的命中点不能准确定位。系统对 PHRASE 型字段上的数据进行分词处理，并建立倒排索引，这是 PHRASE 型字段与 CHAR 型字段的关键区别之一，但其倒排索引只有“位置”索引属性，而不允许通常倒排索引的“段、句”等其它索引属性。PHRASE 型字段与 CHAR 型字段的另一个重要区别就是 PHRASE 型字段不允许多值。

DOCUMENT 型字段既可以用来存储任何形式、任意长度的文本数据（但受软硬件环境的制约，请参见 1.3 节的有关说明），也可以用来存储其它格式类型的多媒体数据。系统对 DOCUMENT 型字段上的数据进行分词处理，并建立倒排索引。和 PHRASE 型字段一样，DOCUMENT 型字段也不允许多值。DOCUMENT 型字段与 PHRASE 型字段还有以下几点区别：

- ◆ 对 DOCUMENT 型字段的数据的存储方式可以进行控制。
- ◆ DOCUMENT 字段可以引用外挂的字段值文件以及文本值文件。
- ◆ DOCUMENT 字段能够对 XML、HTML 以及其它一些格式化文档自动抽取文本并建立索引。

BIT 型字段可以用来存储任何格式类型的多媒体数据。系统对 BIT 型字段的数据不建立任何索引；允许多值；在数据加载时，字段值可以引用文件。

10.2 准备加载文件

准备加载文件，是完成数据加载的一个关键步骤。加载文件有两类：一类是数据文件；另一类是控制文件。

数据文件是用来存放待加载的数据的文件。数据文件有两种形式，即数据记录文件和批处理文件，将在第 11 章中详细说明这两种文件。^{【修订 024】}各词典的加载请参见第

14 章；DATE 的加载格式请参见第 13 章；DOCUMENT 和 BIT 字段的加载格式请参考第 12 章；对 NUMBER 型字段的数据加载需要特别说明的是：虽然支持用科学记数法表示的数值，但不支持用逗号作为千分位分隔符的数值表示法，例如 “1,234.567” 的写法不支持，而应该写成 “1234.567”，或者 “1.234567E3”。^[4]

控制文件用来控制数据的加载过程的配置文件，通过配置控制文件的变量，可以触发系统的内部功能，或改变系统的运行方式。第 15 章详细说明了控制文件的组织和各个变量的意义与配置方法。

值得特别注意的是，如果当前应用设置了本地码（如 UTF-8，请参见附录一），则数据加载所涉及到的全部文本数据及控制文件都必须是采用本地码（如 UTF-8）进行编码的，但批处理文件（参见 11.3 节）的字符集需与文件系统（或操作系统）保持一致，否则会因文件名或路径错误而无法地完成数据加载过程。对于外挂（参见第 12 章）和内嵌（参见 11.1.1 节）的 XML 与 HTML 格式不进行编码转换，当本地码与入库目标数据库的字符集不一致时，格式中必须指定相应的字符集，否则就会出现“乱码”的问题。

第 11 章 数据文件

11.1 TRS格式文件

数据记录文件实际上是由一个或多个记录数据组成的 TRS 标记文件, 统称为 **TRS 格式文件**, 包括: 标准格式、字段内部标识号标记格式、字段顺序号标记格式、无字段标记格式、全文格式和传输格式。单个 TRS 格式文件的大小不能超过 2G 字节。

在 TRS 格式文件中, 每个记录都以标记 “<REC>” 开始, 即标记 “<REC>” 是记录之间的分隔符。第一个记录前和最后一个记录后的该标记可以省略; 如果文件中只有一个记录, 则该标记可以省略。

数据记录文件的一般格式是:

```
<REC>  ATTRIBUTES
      字段 1
      字段 2
      .....
      字段 n
<REC>  ATTRIBUTES
      字段 1
      字段 2
      .....
      字段 n
```

其中:

- ◆ 标记<REC>必须顶格书写。
- ◆ 字段 n 的描述共有 5 种不同的格式, 分别为“标准格式”、“字段内部标识号标记格式”、“字段顺序号标记格式”、“无字段标记格式”和“全文格式”。
- ◆ 标记<REC>后面的 ATTRIBUTES 是记录的属性描述, 它必须与标记<REC>写在同一行。ATTRIBUTES 不是必须的。

ATTRIBUTES 由下列一个或多个分项组成:

- ◆ SecurityCode = CodeValue。其中“SecurityCode”可使用缩写“SC”, 不区分大小

写。这里“CodeValue”是一个有效的用户密级（数字，0 表示没有密级加密，不允许为空串），表示只有用户的密级不大于该值时才能读出该记录。

- ◆ UserGroup = GroupName。其中“UserGroup”可使用缩写“UG”，不区分大小写。这里“GroupName”是一个有效的用户组名（不区分大小写，不允许为空串），表示只有该用户组的用户才能读出该记录。
- ◆ Password = PasswordString。其中“Password”可使用缩写“PWD”，不区分大小写。这里“PasswordString”是一个不超过 8 个字节的字符串（大小写敏感，不允许为空串，首尾的空白字符无效），表示该记录读出时必须提供相同的口令。
- ◆ DOCID = RecordNumber。其中“RecordNumber”表示该记录所对应的库中记录的物理记录号。属性名“DOCID”可以用“ID”代替。该属性一般用于批量更新记录时提交的数据记录文件中。【修订 091】

ATTRIBUTES 中包含多个分项时，分项之间用单字节的逗号“,”或分号“;”或空格分隔，分项之间没有顺序限制，但如果用空格分隔，则必须给出全部分项。如：

```
<REC> SecurityCode = 10; UserGroup = Administrator;
.....
<REC> PWD = OnlyMe
.....
<REC>
.....
```

有关记录加密的使用规则，请参阅 4.1.4 节“记录级权限”的描述。

11.1.1 标准格式

标准格式的 TRS 数据记录文件的组织格式为：

```
<REC>
< 字段名 1>= 字段值 1.....
.....
< 字段名 2>= 字段值 2.....
.....
< 字段名 n>= 字段值 n.....
.....
<REC>
< 字段名 1>= 字段值 1.....
```

.....
<字段名 2>= 字段值 2.....

.....
<字段名 n>= 字段值 n.....

.....

其中:

- ◆ 标记引导符“<”必须顶格书写，标记结束符“>=”是一个整体，中间不能有任何其他字符。这两个标记符号都是 ASCII 码，必须成对出现。
- ◆ 标记引导符“<”之前可以加上一个前缀，以更加准确地指明一个 TRS 标记的开始。请参见 11.1.8 节的说明。
- ◆ 标记引导符“<”和标记结束符“>=”之间是一个字段名，或者是一个字段的别名。“<字段名 n>=”必须书写在同一文本行，不能跨行。**如果使用字段的别名，则必须在控制文件中用 COLUMN_NAME_SEQUENCE 变量列出所有需要加载数据的字段所对应的别名，除非设置 COLUMN_AUTO_MATCH 变量为 TRUE。**
- ◆ 标记结束符“>=”与下一个标记引导符“<”之间的一行或多行文本为该字段的值。文本前后的空字符不计入字段值，但 PHRASE 和 DOCUMENT 字段可通过控制文件的设置（KEEP_TEXT_ACTUAL 变量）保持其文本原样。字段值可以为空。
- ◆ 字段之间没有顺序限制。多值字段的值可以用一个“<字段名 n>=字段值 n”的方式书写，多值之间用特定的字符或字符串分隔（分隔符必须与控制文件中 MULTI_VALUE_SEPARATER 变量的值相一致）；也可以用多个“<字段名 n>=字段值 n”的方式书写；这两种方式也可以同时使用。但是如果一个字段不允许多值，则该字段的所有“<字段名 n>=字段值 n”中，只有第一个有效，其它将被忽略。
- ◆ 对于 BIT 字段，如果要在一个“<字段名 n>=字段值 n”格式中存放多个值，则其中的每个值都必须用文件方式给出。
- ◆ 对于 PHRASE 和 DOCUMENT 字段，如果在一个记录中出现多次，则根据控制文件的设置（CHECK_TEXT_ITERATIVE 变量）进行处理，详情请参见 15.2.14 节的说明。

下面是一个包含两条记录的“专利数据库”的标准格式的数据记录文件:

```
<REC>
<公告日>=920101
<分类号>=A01B49/02
<公告号>=CN_2091520U
```

<申请号>=91208300.X

<申请日>=910514

<申请人>=宝鸡县机电设备维修厂

<地址>=721304 陕西省宝鸡县钓渭乡

<设计人>=赵志栋

<专利代理机构>=宝鸡市专利事务所

<代理人>=翟中平

<摘要>=耕耙犁

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>

<公告日>=920101

<分类号>=A01B49/04

<公告号>=CN_2091521U

<申请号>=91206137.5

<申请日>=910411

<申请人>=张岳

<地址>=475000 河南省开封市南太平街3号楼2单元付3号

<设计人>=张岳

<设计人>=张龙

<设计人>=范文增

<设计人>=石建峡

<设计人>=阎唯一

<专利代理机构>=开封市专利事务所

<代理人>=黄永武

<代理人>=乔玉萍

<摘要>=牵引式多功能农农机

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农农机。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、

造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

对于 DOCUMENT 字段，其字段标记还支持标记属性，用以描述说明字段值的特性，这种方式常称为**内嵌格式化数据方式**。但是如果字段值以外挂文件的方式给出（参见 12.1 节），则字段标记属性不起作用。用这种方式来加载数据库记录，其数据的传输效率要比外挂文件方式的效率稍高一些。带标记属性的字段标记的格式为：

<字段名 ATTRIBUTES>= 字段值.....

ATTRIBUTES 由下列一个或多个分项组成：

- ◆ Format = *FormatName*。其中属性名“Format”可以使用缩写“FMT”，且不区分大小写。“*FormatName*”必须是一个有效的数据格式名（不区分大小写），表示字段值的数据库格式。目前只支持 TEXT、HTML 和 XML 三种格式。
- ◆ Storage = *StorageMode*。其中属性名“Storage”可使用缩写“STG”，且不区分大小写。“*StorageMode*”必须是一个有效的存储方式控制符（不区分大小写），表示字段值在数据库中的存储方式。有关“存储方式控制符”的说明请参见 12.1.2 节。

当 ATTRIBUTES 中包含多个分项时，分项之间用单字节的逗号“,”或分号“;”或空格分隔，分项之间没有顺序限制，但如果用空格分隔，则给出的分项必须是完整的，即分项的“=”后面必须有值。下面是一个例子：

```
$$STRSTAG$$<正文 FORMAT=HTML; STORAGE=Class>=
```

```
<HTML>
```

```
<HEAD>
```

```
<TITLE>
```

```
Internet 数据库连接程序示例
```

```
</TITLE>
```

```
</HEAD>
```

```
<BODY BACKGROUND="backgrnd.gif" BGCOLOR="FFFFFF">
```

该页提供了使用 Internet 数据库连接程序的几个示例的链接。尽管 Internet 数据库连接程序能使用任何 32 位开放数据库连接 (ODBC) - 兼容驱动器，该驱动器支持系统数据源名称(DSN)，但这些例子使用了 Microsoft SQL 服务器。

```
<P>
```

要不加改动地运行这些示例，您的计算机必须能访问 Microsoft SQL Server <I>sa</I> 账号，而且此账号不要求密码（该账号是在默认安装

Microsoft SQL Server 时产生的)。如果不能或者不愿在这些示例中使用 `<I>sa</I>` 账号, 可在示例使用的 .idc 文件中编辑 “Username field”(如 果必要, 增加 “Password field”), 使其成为合法的用户账号。

`<P>`

要安装这些例子, 须在连接到 SQL 服务器的计算机上创建名为 `Web SQL` 的 ODBC 系统数据源。使用控制面板上的 ODBC 程序创建系 统 DSN, 或者单击 `` 此处 ``, 用随 Microsoft Internet Information Server 或 Peer Web Services 一 起提供的 ``ODBC 数据源工具`` 来创建它。

`</BODY>`

`</HTML>`

.....

在这个例子中, 我们在字段标记“`<正文 FORMAT=XML; STORAGE=Class>=`”前 插入了一串特殊字符“`$$STRSTAG$$`”, 它是一个标记引导符前缀, 这一点将在 11.1.8 节中进行详细介绍。

数据记录文件采用“标准格式”进行数据加载时, 在控制文件中可以指定变量 `RECORD_FILE_FORMAT` 为 `STANDARD`, 或者缺省。请参见 15.2.1 节的说明。

11.1.2 字段内部标识号标记格式

字段内部标识号标记格式与标准格式的唯一区别在于: 标准格式中用标记“`<字 段名>=`”来指示一个字段, 而这种格式则用标记“`<字段内部标识号>=`”来指示一个 字段。字段内部标识号是指该字段在数据词典中的唯一标识号。

下面是这个“专利数据库”的字段内部标识号标记格式的数据记录文件组织:

`<REC>`

`<1>=920101`

`<2>=A01B49/02`

`<3>=CN_2091520U`

`<4>=91208300.X`

`<5>=910514`

`<7>=宝鸡县机电设备维修厂`

`<8>=721304 陕西省宝鸡县钓渭乡`

<9>=赵志栋

<10>=宝鸡市专利事务所

<11>=翟中平

<12>=耕耙犁

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>

<1>=920101

<2>=A01B49/04

<3>=CN_2091521U

<4>=91206137.5

<5>=910411

<7>=张岳

<8>=475000 河南省开封市南太平街3号楼2单元付3号

<9>=张岳

<9>=张龙

<9>=范文增

<9>=石建峡

<9>=阎唯一

<10>=开封市专利事务所

<11>=黄永武

<11>=乔玉萍

<12>=牵引式多功能农用车

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农用车。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

这里假设各个字段的内部标识号为：

字段名	标识号	字段名	标识号	字段名	标识号
公告日	1	申请日	5	设计人	9
分类号	2	颁证日	6	专利代理机构	10
公告号	3	申请人	7	代理人	11
申请号	4	地址	8	摘要	12

数据记录文件采用“字段内部标识号标记格式”进行数据加载时，在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 IDENTISTD。请参见 15.2.1 节的说明。

11.1.3 字段顺序号标记格式

字段顺序号标记格式与字段内部标识号标记格式非常相似，仅有两点差别：

1. 字段内部标识号标记格式则用标记“<字段内部标识号>=”来指示一个字段，而这种格式是用标记“<字段顺序号>=”来指示一个字段。
2. 字段顺序号是字段在控制文件中（COLUMN_NAME_SEQUENCE 变量）给出的字段列表中的次序。第一个字段的顺序号为 1，依次类推。

下面是“专利数据库”的字段顺序号标记格式的数据记录文件组织：

```

<REC>
<1>=920101
<2>=A01B49/02
<3>=CN_2091520U
<4>=91208300.X
<5>=910514
<6>=宝鸡县机电设备维修厂
<7>=721304 陕西省宝鸡县钓渭乡
<8>=赵志栋
<9>=宝鸡市专利事务所
<10>=翟中平
<11>=耕耙犁

```

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时

不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>

<1>=920101

<2>=A01B49/04

<3>=CN_2091521U

<4>=91206137.5

<5>=910411

<6>=张岳

<7>=475000 河南省开封市南太平街3号楼2单元付3号

<8>=张岳

<8>=张龙

<8>=范文增

<8>=石建峡

<8>=阎唯一

<9>=开封市专利事务所

<10>=黄永武

<10>=乔玉萍

<11>=牵引式多功能农用机

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农用机。其特征是：设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔的市场。

这里假设控制文件中给出的字段列表为：公告日;分类号;公告号;申请号;申请日;申请人;地址;设计人;专利代理机构;代理人;摘要。

数据记录文件采用“字段顺序号标记格式”进行数据加载时，在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 SIMPLESTD。请参见 15.2.1 节的说明。

11.1.4 无字段标记格式

无字段标记格式是指在 TRS 数据记录文件的一个记录中，字段之间不是用显式的标记来分隔，而是通过自然的文本行来标示不同的字段。这种格式的特点是：

- ◆ 字段的顺序由控制文件（COLUMN_NAME_SEQUENCE 变量）定义。

- ◆ 除最后一个字段外，每个字段的值只能占用一行。没有值的字段，如果不是最后一个字段，则也必须用一空行来表示。
- ◆ 最后一个字段的值可以是多行，直到下一个记录的标记引导符“<”为止。也就是说，字段值多于一行的字段必须是每个记录的最后一个字段。
- ◆ 如果有多个字段的值多于一行，则不能采用这种格式组织数据记录文件。

下面是“专利数据库”的无字段标记格式的数据记录文件组织：

<REC>

920101

A01B49/02

CN_2091520U

91208300.X

910514

宝鸡县机电设备维修厂

721304 陕西省宝鸡县钓渭乡

赵志栋

宝鸡市专利事务所

翟中平

耕耙犁

耕耙犁是用于农田深耕、碎土的农机具，它是由主动力输入总成、动力传输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农机具具有结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。

<REC>

920101

A01B49/04

CN_2091521U

91206137.5

910411

张岳

475000 河南省开封市南太平街3号楼2单元付3号

张岳;张龙;范文增;石建峡;阎唯一

开封市专利事务所

黄永武;乔玉萍

牵引式多功能农用车

本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打畦以及碾压封土的多功能农用车。其特征是:设有筑埂打畦器和带圆锥形轮齿的主动轮及采用塔式链轮变速,并结合容器轮的容器槽及连接于容器槽内的容器,实现了不同株距的要求,通过成套部件还可将单机构成全机和并机,可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、造价低、操作维修方便,适用于个体户、农场、农机站的需要,具有广阔的市场。

这里假设控制文件中给出的字段列表为:公告日;分类号;公告号;申请号;申请日;申请人;地址;设计人;专利代理机构;代理人;摘要。

数据记录文件采用“无字段标记格式”进行数据加载时,在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 SEQUENCE。请参见 15.2.1 节的说明。

11.1.5 全文格式

全文格式的特点是:

- ◆ 数据记录文件的每个记录中只有一个字段。
- ◆ 记录中没有字段标记。
- ◆ 如果数据库有多个字段,则目标字段由控制文件(COLUMN_NAME_SEQUENCE 变量)定义,或按其缺省规则自动确定目标字段。如果指定目标字段,则该字段必须是一个 DOCUMENT 或 PHRASE 字段,并且只能指定一个字段。

下面是一个包含两条记录的“红楼梦数据库”的全文格式的数据记录文件:

<REC>

第一回 甄士隐梦幻识通灵 贾雨村风尘怀闺秀

此开卷第一回也。作者自云:因曾历过一番梦幻之后,故将真事隐去,而借“通灵”之说,撰此《石头记》一书也。故曰“甄士隐”云云。但书中所记何事何人?自又云“今风尘碌碌,一事无成,忽念及当日所有之女子,一一细考较去,觉其行止见识,皆出于我之上。何我堂堂须眉,诚不若彼裙钗哉?实愧则有余,悔又无益之大无可如何之日也!当此,则自欲将已往所赖天恩祖德,锦衣纨裤之时,饫甘餍肥之日,背父兄教育之恩,负师友规谈之德,以至今日一技无成,半生潦倒之罪,编述一集,以告天

下人：我之罪固不免，然闺阁中本自历历有人，万不可因我之不肖，自护己短，一并使其泯灭也。虽今日之茅椽蓬牖，瓦灶绳床，其晨夕风露，阶柳庭花，亦未有妨我之襟怀笔墨者。虽我未学，下笔无文，又何妨用假语村言，敷演出一段故事来，亦可使闺阁昭传，复可悦世之目，破人愁闷，不亦宜乎？故曰“贾雨村”云云。

.....

<REC>

第二回 贾夫人仙逝扬州城 冷子兴演说荣国府

诗云

一局输赢料不真，香销茶尽尚逡巡。
欲知目下兴衰兆，须问旁观冷眼人。

却说封肃因听见公差传唤，忙出来陪笑启问。那些人只嚷：“快请出甄爷来！”封肃忙陪笑道：“小人姓封，并不姓甄。只有当日小婿姓甄今已出家一二年了，不知可是问他？”那些公人道：“我们也不知什么‘真’‘假’，因奉太爷之命来问，他既是你女婿，便带了你去亲见太爷面禀，省得乱跑。”说着，不容封肃多言，大家推拥他去了。封家人个个都惊慌，不知何兆。

那天约二更时，只见封肃方回来，欢天喜地。众人忙问端的。他乃说道：“原来本府新升的太爷姓贾名化，本贯胡州人氏，曾与女婿旧日相交。方才在咱门前过去，因见娇杏那丫头买线，所以他只当女婿移住于此。我一一将原故回明，那太爷倒伤感叹息了一回，又问外孙女儿，我说看灯丢了。太爷说：‘不妨，我自使番役务必探访回来。’说了一回话，临走倒送了我二两银子。”甄家娘子听了，不免心中伤感。一宿无话。

.....

这里假设控制文件中给出的字段列表为：COLUMN_NAME_SEQUENCE=正文。

数据记录文件采用“全文格式”进行数据加载时，在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 FULLTEXT。请参见 15.2.1 节的说明。

11.1.6 传输格式

传输格式是数据加载时用于传输一批记录的内部格式文件。这种格式的最大优点是能够将文本与非文本数据打包在一个文件里，避免大量外挂文件可能引起的性能问

题，并且也是唯一支持数据加密与压缩的记录文件格式。鉴于此特点，这种格式也开放给外部使用，但生成这种格式文件需要调用专门提供的接口。【修订 063】

11.1.7 注释信息

为了便于浏览，有时需要在数据文件中加上一些说明信息，而这些信息又不能作为数据加载到数据库中。为此，TRS 全文数据库系统支持在“标准格式”、“字段内部标识号标记格式”、“字段顺序号标记格式”、“无字段标记格式”和“全文格式”这五种 TRS 格式文件中可以插入注释信息。注释信息实际上是一个特殊的 TRS 标记。

注释标记的格式为：<*>=注释信息

其中，“<*>=”标记注释信息的开始，直到当前行的结束。注释标记可以出现在 TRS 格式文件中任何位置。

11.1.8 标记引导符前缀

在缺省情况下，TRS 格式文件中的标记是以引导符“<”开始的，通常条件下这不会遇到麻烦。但有的数据中会含有符号“<”，并且恰好是某行的第一个字符，这是该记录就无法加载到数据库中。为了适应这种情况，TRS 全文数据库系统允许在标记（包括注释标记）的引导符前加上一个前缀，以特别地指明一个 TRS 标记的开始。具体采用什么前缀，由控制文件中的 SECTION_TAG_PREFIX 变量给出，一个引导符前缀的最大长度为 30 个字符。

下面是一个包含两条记录的“专利数据库”的带有标记引导符前缀的标准格式的数据记录文件：

```
$$TRSTAG$$<REC>  
$$TRSTAG$$<公告日>=920101  
$$TRSTAG$$<分类号>=A01B49/02  
$$TRSTAG$$<公告号>=CN_2091520U  
$$TRSTAG$$<申请号>=91208300.X  
$$TRSTAG$$<申请日>=910514  
$$TRSTAG$$<申请人>=宝鸡县机电设备维修厂  
$$TRSTAG$$<地址>=721304 陕西省宝鸡县钓渭乡  
$$TRSTAG$$<设计人>=赵志栋  
$$TRSTAG$$<专利代理机构>=宝鸡市专利事务所
```

```

$$TRSTAG$$<代理人>=翟中平
$$TRSTAG$$<摘要>=耩耙犁
耩耙犁是用于农田深耕、碎土的农具，它是由主动力输入总成、动力传
输箱、左、右传动轴总成、左、右动力输出箱及左、右碎土器等构成。主
动力输入总成与动力传输箱连接左、右传动轴总成与动力传输箱连接且位
于主动力输入总成的两侧，左、右动力输出箱分别与左、右传动轴总成连
接，左、右碎土器与左、右动力输出箱的动力输出轴连接，该农具具有
结构设计新颖、简单、实用的特点，特别适应于在小块田地作业且作业时
不留死角，并且能同时完成深耕和碎土两道作业，具有工效高的优点。
$$TRSTAG$$<REC>
$$TRSTAG$$<公告日>=920101
$$TRSTAG$$<分类号>=A01B49/04
$$TRSTAG$$<公告号>=CN_2091521U
$$TRSTAG$$<申请号>=91206137.5
$$TRSTAG$$<申请日>=910411
$$TRSTAG$$<申请人>=张岳
$$TRSTAG$$<地址>=475000 河南省开封市南太平街3号楼2单元付3号
$$TRSTAG$$<设计人>=张岳
$$TRSTAG$$<设计人>=张龙
$$TRSTAG$$<设计人>=范文增
$$TRSTAG$$<设计人>=石建峡
$$TRSTAG$$<设计人>=阎唯一
$$TRSTAG$$<专利代理机构>=开封市专利事务所
$$TRSTAG$$<代理人>=黄永武
$$TRSTAG$$<代理人>=乔玉萍
$$TRSTAG$$<摘要>=牵引式多功能农具
本实用新型是一种具有施肥、追肥、播种、点种、刨茬、锄草、筑埂、打
畦以及碾压封土的多功能农具。其特征是：设有筑埂打畦器和带圆锥形
轮齿的主动轮及采用塔式链轮变速，并结合容器轮的容器槽及连接于容器
槽内的容器，实现了不同株距的要求，通过成套部件还可将单机构成全机
和并机，可分别适用于人力、畜力、拖拉机牵引。本实用新型结构简单、
造价低、操作维修方便，适用于个体户、农场、农机站的需要，具有广阔
的市场。

```

在这个例子的控制文件中必须给出：SECTION_TAG_PREFIX=\$\$TRSTAG\$\$。

11.2 国际标准格式

TRS 全文数据库系统的数据加载,除了支持以上五种 TRS 格式外,还支持 ISO2709 和 XML 格式文件的直接入库。

11.2.1 ISO2709 记录格式

ISO2709 格式是一种国际标准,这里就不再进行说明。

下面是一个包含两条记录的“专利数据库”的 ISO2709 格式的数据文件:

```
006990000000001550005500001000110000000200010000110030001200021  
004000110003300500011000440060001100055007000070006600800033000  
7300900007001060120043000113#1994.01.05#A01B49/02#CN_2151607Y#9  
2243586.3#1992.12.08#1993.10.03#刘殿富#154552 黑龙江省勃利县双河镇  
永平村#刘殿富#
```

铲耧锄草耕作机

本实用新型公开了一种耧帮、苗间锄草,起垄深松破茬一次完成的铲耧锄草耕作机,是通过牵引带动向位传轮的伞形齿与苗间锄草器立式转动轴的伞型齿相啮合,使苗间锄草器在苗间水平旋转,同时通过动力带动安装在顺梁上的耧帮锄草器、双犁头深松器和齐刃犁头一起工作,就可达到耧帮、苗间锄草,垄沟深松、破茬起垄一次完成的目的。苗间和耧帮锄草可通过调节杆调节入土深度,苗间锄草部可单独与各种牵引,悬挂式三铧犁配套使用,也可用畜力牵引。

```
##0072500000000018100055000010001100000002000100001100300012000  
210040001100033005000110004400600011000550070000700066008000350  
00730090000700108010000210011501100007001360120040000143#1994.01  
.05#A01C23/04#CN_2151608Y#93202243.X#1993.01.20#1993.09.12#刘忠厚  
#161005 黑龙江省齐齐哈尔市龙沙区农委#刘忠厚#齐齐哈尔市专利事务  
所#叶仲刚#
```

恒压等浓度施肥罐

一种恒压等浓度施肥罐，由罐体、输水管、进水管、施肥管及各种控制阀、压力表和水表组成。在罐体内设置一隔板，将罐体分为上下罐，隔板上有一个气体连通管将上下罐连通。当上罐加入化肥或农药溶液密封后，由进水管向下罐注入清水，罐中产生气压，使上罐内的化肥或农药溶液经过施肥管注入输水管。本装置具有施肥均匀度好，施肥浓度变化可控制，便于自动化实施，成本低且运行费用少，可制成固定式或移动式。

##

这里假设控制文件中给出的字段列表为：公告日;分类号;公告号;申请号;申请日;颁证日;申请人;地址;设计人;专利代理机构;代理人;摘要。

需要注意的一点：格式中的“目录项标记”域是该字段在控制文件字段列表中的序号（从 1 开始计数）。参照上例，如“目录项标记”是“012”，则该字段就是数据库中的“摘要”字段。

还有一点：由于 TRS 不支持子字段，所以 ISO2709 数据域的各个分项将被当作该字段的多值处理；如果该字段不支持多值，则只有第一个数据分项是有效，其它项将被忽略。

TRS 也支持使用格式中的“目录项标记”域作为数据库的字段或别名对数据进行映射，这时需要将控制文件中的 NAMIZE_ISO2709_FIELD 变量设置成 TRUE。请参见 15.2.28 节的说明。

数据记录文件采用 ISO2709 格式进行数据加载时，在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 ISO2709。请参见 15.2.1 节的说明。

11.2.2 XML 记录格式

和 ISO2709 一样，XML 格式也是一种国际标准，具体说明请参考有关资料。

TRS 以两种方式支持 XML 数据的直接入库：一种是记录级的支持，即 XML 文件的结构对应于一个 TRS 数据库的结构，该 XML 文件由许多记录组成，这时可以将 XML 文件理解成符合 XML 规范的 TRS 格式文件；另一种是字段级的支持，即把 XML 数据作为 TRS 数据库中的一个字段的值进行加载，自动完成文本过滤并建立索引，在处理上，与 HTML 没有本质的区别。对于字段级的支持，在加载数据时还可以通过控

制文件设置一些选项，请参见 15.2.22 ~ 15.2.26 节的说明。

这一节所描述的内容都是针对于 XML 记录级支持的。在这种方式下，不解析字段值中的子标记，用来标记数据的“字段标记”，既可以是字段名，也可以是字段的别名。如果使用字段别名，则必须在控制文件中用 **COLUMN_NAME_SEQUENCE** 变量列出所有需要加载数据的字段所对应的别名，除非设置 **COLUMN_AUTO_MATCH** 变量为 **TRUE**。

下面是一个包含两条记录的“专利数据库”的 XML 格式的数据文件：

```
<?xml version="1.0" encoding="gb2312"?>
<!DOCTYPE TRS [
<!ELEMENT TRS (REC)>
<!ELEMENT REC (公告日,分类号,公告号,申请号,申请日,颁证日,申请人,
地址,设计人,专利代理机构,代理人,摘要)>
<!ELEMENT 公告日 (#PCDATA)>
<!ELEMENT 分类号 (#PCDATA)>
<!ELEMENT 公告号 (#PCDATA)>
<!ELEMENT 申请号 (#PCDATA)>
<!ELEMENT 申请日 (#PCDATA)>
<!ELEMENT 颁证日 (#PCDATA)>
<!ELEMENT 申请人 (#PCDATA)>
<!ELEMENT 地址 (#PCDATA)>
<!ELEMENT 设计人 (#PCDATA)>
<!ELEMENT 专利代理机构 (#PCDATA)>
<!ELEMENT 代理人 (#PCDATA)>
<!ELEMENT 摘要 (#PCDATA)>
]>
<TRS>
<REC>
<公告日>1994.01.05</公告日>
<分类号>A01B49/02</分类号>
<公告号>CN_2151607Y</公告号>
<申请号>92243586.3</申请号>
<申请日>1992.12.08</申请日>
<颁证日>1993.10.03</颁证日>
<申请人>刘殿富</申请人>
```


<地址>154552 黑龙江省勃利县双河镇永平村</地址>
 <设计人>刘殿富</设计人>
 <摘要><![CDATA[
 铲耧锄草耕作机

本实用新型公开了一种垄帮、苗间锄草，起垄深松破茬一次完成的铲耧锄草耕作机，是通过牵引带动向位传轮的伞形齿与苗间锄草器立式转动轴的伞型齿相啮合，使苗间锄草器在苗间水平旋转，同时通过动力带动安装在顺梁上的垄帮锄草器、双犁头深松器和齐刃犁头一起工作，就可达到垄帮、苗间锄草，垄沟深松、破茬起垄一次完成的目的。苗间和垄帮锄草可通过调节杆调节入土深度，苗间锄草部可单独与各种牵引，悬挂式三铧犁配套使用，也可用畜力牵引。

]]></摘要>

</REC>

<REC>

<公告日>1994.01.05</公告日>

<分类号>A01C23/04</分类号>

<公告号>CN_2151608Y</公告号>

<申请号>93202243.X</申请号>

<申请日>1993.01.20</申请日>

<颁证日>1993.09.12</颁证日>

<申请人>刘忠厚</申请人>

<地址>161005 黑龙江省齐齐哈尔市龙沙区农委</地址>

<设计人>刘忠厚</设计人>

<专利代理机构>齐齐哈尔市专利事务所</专利代理机构>

<代理人>叶仲刚</代理人>

<摘要><![CDATA[

恒压等浓度施肥罐

一种恒压等浓度施肥罐，由罐体、输水管、进水管、施肥管及各种控制阀、压力表和水表组成。在罐体内设置一隔板，将罐体分为上下罐，隔板上有一个气体连通管将上罐连通。当上罐加入化肥或农药溶液密封后，由进水管向下罐注入清水，罐中产生气压，使上罐内的化肥或农药溶液经过施肥管注入输水管。本装置具有施肥均匀度好，肥浓度变化可控制，便于自动化实施，成本低且运行费用少，可制成固定式或移动式。

]]></摘要>

</REC>

</TRS>

对于任何一个 XML 的 DTD 定义, TRS 目前只检查辨认其中的 DOCTYPE 声明(定义)以及 DOCTYPE 的子元素声明(定义)的内容, 并且 DOCTYPE 只允许有一个子元素, 同时忽略 DTD 中的所有其它声明(定义)。这里的 DOCTYPE 即是 XML 的顶层标记, DOCTYPE 的子元素即是 XML 的第一层子标记。

如果 DOCTYPE 为 TRS, 且 DOCTYPE 的子元素为 REC, 则 XML 的 DTD 部分可以省略, 否则必须定义文档的 DTD (也只需 DOCTYPE 声明和 DOCTYPE 的子元素声明)。对 XML 字段级的支持, 则 DTD 部分总是可以省略。

如果在 XML 文件中定义了 DTD, 且 DOCTYPE 为 TRS, DOCTYPE 的子元素为 REC, 则有以下的特殊支持:

- ◆ 元素 REC 支持三个属性: SC 或 SecurityCode; UG 或 UserGroup; PWD 或 Password。这三个属性用于对记录的加密, 详细说明请参见 11.1 节。其它属性将被忽略。
- ◆ 如果将“br”定义成一个“空”元素, 即定义: <!ELEMENT br EMPTY>, 则此时 br 标记的意义固定为“换行”, 而 XML 数据中原来的所有“自然文本换行”将被忽略。对于字段级的支持, TRS 也是这么处理的, 除非另有设置(参见 15.2.23 节的说明)。例如:

```
<!DOCTYPE TRS [  
<!ELEMENT TRS (REC)>  
<!ELEMENT REC (日期,版次,版名,栏目,标题,作者,正文,br)>  
<!ELEMENT 日期 (#PCDATA | br)*>  
<!ELEMENT 版次 (#PCDATA | br)*>  
<!ELEMENT 版名 (#PCDATA | br)*>  
<!ELEMENT 栏目 (#PCDATA | br)*>  
<!ELEMENT 标题 (#PCDATA | br)*>  
<!ELEMENT 作者 (#PCDATA | br)*>  
<!ELEMENT 正文 (#PCDATA | br)*>  
<!ELEMENT br EMPTY>  
>
```

对于 XML 数据中的实体引用(实体定义被忽略), TRS 接受绝大多数 ISO 标准的实体名, 而自己定义的非标准实体则不能识别。不能识别的 ISO 标准实体和自定义的非标准实体引用, TRS 将保持原样, 不进行替换。

TRS 支持大多数流行的字符集编码格式的 XML 文件, 这些字符集包括: GB2312、GBK、GB18030、BIG5、UTF-8、ISO8859-1 ~ ISO8859-16 以及 WINDOWS-1250 ~ WINDOWS-1258。

数据记录文件采用 XML 格式进行数据加载时, 在控制文件中必须指定变量 RECORD_FILE_FORMAT 为 XML。请参见 15.2.1 节的说明。

11.3 批处理数据文件

对数据库进行加载时, 数据记录文件既可以是单个 TRS 格式(“标准格式”, 或“字段内部标识号标记格式”, 或“字段顺序号标记格式”, 或“无字段标记格式”, 或“全文格式”之一)文件, 或者单个 ISO2709 格式文件, 或者单个 XML 格式文件; 也可以是一批相同格式(上述 7 种数据格式文件之一)的数据文件。

批处理数据文件实际上是存储一批数据记录文件名的文件。其基本规则是:

- ◆ 批处理数据文件是一个文本文件。文件的每一自然行存储一个数据记录文件名。文件名的字符集需与文件系统(或操作系统)保持一致。
- ◆ 数据记录文件名中既可以包含路径, 也可以不包含路径。如果没有路径, 则默认文件放在当前工作目录下; 如果当前工作目录下不存在指定的数据记录文件, 则系统将取批处理数据文件的路径作为数据记录文件的缺省路径。其中路径的子目录分隔符没有 ‘\’ 和 ‘/’ 之分, TRS 能够自动适应这两种符号系统。
- ◆ 文件中所列的全部数据记录文件必须具有相同数据记录格式。

下面是一个批处理数据文件的例子:

```
D:\DATA\NEWS\1990.TRS  
D:\DATA\NEWS\1991.TRS  
D:\DATA\NEWS\1992.TRS  
D:\DATA\NEWS\1993.TRS  
D:\DATA\NEWS\1994.TRS  
D:\DATA\NEWS\1995.TRS  
D:\DATA\NEWS\1996.TRS  
D:\DATA\NEWS\1997.TRS  
D:\DATA\NEWS\1998.TRS  
D:\DATA\NEWS\1999.TRS  
D:\DATA\NEWS\2000.TRS
```

D:\DATA\NEWS\2001.TRS

D:\DATA\NEWS\2002.TRS

D:\DATA\NEWS\2003.TRS

D:\DATA\NEWS\2004.TRS

D:\DATA\NEWS\2005.TRS

第 12 章 多媒体数据的加载

12.1 加载 DOCUMENT 字段的数据

DOCUMENT 字段除了接受直接放在第 11 章所描述 7 种数据记录文件中的 TEXT 数据以及以内嵌格式化数据方式给出的 HTML 和 XML 格式的数据之外，该字段还可以接受其他数据格式类型的数据作为字段值，诸如 PDF、WORD 等等。由于这些类型的数据可能是二进制的，不能直接放在数据记录文件中，而需要将它们置于单独的文件中，然后在数据记录文件中引用这些数据文件名。引用的方法是在文件名的前面加上字符 '@' 后再作为字段的值。这种加载数据的方式一般称为**外挂文件方式**。

DOCUMENT 字段的值一般是需要对其进行索引的，但是 TRS 全文数据库系统目前只能对 TEXT、HTML 和 XML 格式的数据进行索引。其它格式类型的数据，如果需要建立索引，则必须事先将其转换成 TEXT 格式的数据文件，并与原数据文件一起提供给数据记录文件引用，或者安装一个 OEM 插件程序，由系统自动完成这一过程。通过对原数据文件进行格式转换得到的文件，或外部通过其它方式获取的与原数据文件相关联的文件，称为 DOCUMENT 字段的索引参考文件，主要是用来建立查询索引的，因该文件一般都是文本文件，所以有时也称为文本值文件；而原格式的数据文件则称为源格式文档，习惯上称为附件，但它才是真正的字段值文件。

综合起来，DOCUMENT 字段的多媒体数据的加载有以下三种格式：

简单格式： @Source_Name^Format^Store

典型格式： @Source_Name | Refer_Name^Format^Store

完全格式： @Source_Name^Format^Store | Refer_Name^Format^Store

其中：

- ◆ 符号 '@' 为单字节 ASCII 码，是字段值文件引用的引导符。
- ◆ 符号 '|' 为单字节 ASCII 码，是字段值文件与其索引参考文件之间的分隔符。
- ◆ 符号 '^' 为单字节 ASCII 码，是字段值文件及其索引参考文件的控制符引导符。字段值文件及其索引参考文件的控制符有两种类型：数据格式控制符，即格式中的 Format 符号；存储方式控制符，即格式中的 Store 符号。

在被引用的字段值文件及其索引参考文件的文件名中，既可以包含路径，也可以不包含路径，系统能够自动确定其正确的路径。方法是：

步骤1: 直接检查文件名所对应的文件是否存在，如果存在，则不对文件名做任何修改，检查结束，否则继续。

- 步骤2: 检查文件名中是否存在确定的设备定位 (以 ‘\’ 或 ‘/’ 开头, 或以磁盘驱动器符 “X:” 开头的文件名), 如果存在, 则进行错误处理, 否则继续。
- 步骤3: 检查控制文件中是否设置了 DATA_PATH_DEFAULT 变量, 如果没有, 则进行错误处理, 否则继续。
- 步骤4: 检查控制文件的 DATA_PATH_DEFAULT 变量所指定的目录下是否存在这个文件, 如果不存在, 则进行错误处理, 否则继续。
- 步骤5: 确定 DATA_PATH_DEFAULT 变量值为文件的路径, 检查结束。

12.1.1 数据格式控制符

数据格式控制符用来描述一个文件的具体数据格式, TRS 全文数据库系统预定义了一些基本的数据格式, 请参见 1.7 节的说明。

对于同一个文件, 不能同时使用多个数据格式控制符来控制。数据格式控制符可以省略 (其引导符也同时省略), 在这种情况下, 确定一个字段值文件或其索引参考文件的数据格式的步骤如下:

- 步骤1: 检查控制文件中 FILE_SUFFIX_PRIOR 变量的值是否被设置成 TRUE, 如果是, 则继续步骤 2, 否则转到步骤 4。
- 步骤2: 检查文件是否有文件名后缀, 如果有, 则继续步骤 3, 否则转到步骤 4。
- 步骤3: 通过文件名后缀, 获取其对应的数据格式控制符。如果文件名后缀存在对应的数据格式控制符, 则该数据格式控制符就确定了字段值文件或其索引参考文件的数据格式, 否则进行出错处理。
- 步骤4: 默认该字段值文件或其索引参考文件为文本数据。

12.1.2 存储方式控制符

存储方式控制符用来控制字段值文件及其索引参考文件在数据库中的存储方式, 如 CLASS 和 DUMB 等等 (不区分大小写)。

TRS 全文数据库系统所支持的存储方式控制符有:

控制符	存储方式
ALONE	以独立的文件形式存储在数据库的数据目录下
CLASS	将文件内容集中存储到数据库的按字段类型分类的文件中
URL	文件名为一个 URL, 在数据库中存储这个 URL
LINK	文件名为一个 URL, 在数据库中存储这个 URL (与 URL 相同)
BASE	说明一个 URL 为 HTML 文件中 BASE 元素的 HREF 属性值
DUMB	只存储文件的数据格式等信息, 不存储数据

对于同一个文件，不能同时使用多个存储方式控制符来控制。除此之外，系统还有一种内部存储方式，即集中存储方式，这种方式是将字段值存储在数据库的主数据文件中。DATE、NUMBER、CHAR 和 PHRASE 字段始终采用集中存储方式，DOCUMENT 和 BIT 字段缺省采用分类存储方式，而只有在下列情况下，那些未明确指定存储方式的 DOCUMENT 和 BIT 字段字段值才会采用集中存储方式：

- ◆ 加载记录时在控制文件中将 CLASS_DOCUMENT_STORE(CLASS_BIT_STORE) 变量设置为 FALSE。
- ◆ 在插入/修改记录时数据库中已有记录，并且所有记录的字段值都没有采用分类存储方式。

如果 HTML 文件的内容存储在系统中，则输出时肯定不在它原来的位置，这时它引用的其他一些处在相对路径下的资源就无法找到，BASE 控制符的目的就是要在 HTML 中设置一个绝对路径，以使它在输出时还能正确引用那些资源。显然 BASE 控制符不能与 URL（或 LINK）控制符控制同一个对象。

BASE 控制符只能用在 DOCUMENT 字段的“典型格式”和“完全格式”两种加载格式中，并且只对 HTML 数据格式的索引参考文件有效，它表示将指定的 URL 自动插入到 HTML 文件中作为其 BASE 元素（如果原文件中已经存在 BASE 元素，则不作任何改动），被控制的 URL 本身不存入数据库。控制文件中 HTML_BASE_DEFAULT 变量指定了缺省 BASE 控制符时的情况。

对于一个数据库记录，如果某个 DOCUMENT 字段的用于建立索引的数据部分采用了 URL（或 LINK）或 DUMB 存储方式，则存在以下限制，必须予以注意：

- ◆ 重建索引时，在这个记录的字段上没有索引。但对 URL 模式，如果此时服务器能够直接访问这些 URL，则仍然能够创建索引。
- ◆ 复制记录时，这个记录不能作为源数据复制到其它数据库中。但对 URL 模式，如果此时服务器能够直接访问这些 URL，则仍然能够被复制。
- ◆ 修改记录时，要么这个记录所有的 PHRASE 和 DOCUMENT 字段都不修改，要么这个记录的以这些方式存储的字段也提供新的字段值。

存储方式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的存储方式的步骤如下：

- 步骤1: 检查控制文件中 CLASS_DOCUMENT_STORE 变量的值是否为 TRUE，如果是，则按分类方式存储字段值文件，否则转到步骤 2。
- 步骤2: 按系统缺省的存储方式来存储字段值文件。

在省略存储方式控制符的情况下，确定一个索引参考文件的存储方式的步骤是：

- 步骤1: 检查控制文件中 WAIVE_DOCUMENT_REFER 变量的值是否为 TRUE，如

果是，则按 DUMB 方式存储字段值索引参考文件，否则转到步骤 2。

步骤2: 检查控制文件中 CLASS_DOCUMENT_STORE 变量的值是否为 TRUE，如果是，则按分类方式存储字段值索引参考文件，否则转到步骤 3。

步骤3: 按系统缺省的存储方式来存储字段值索引参考文件。

12.1.3 加载格式

在 12.1 节已经给出了 DOCUMENT 字段多媒体数据的三种格式，下面对这三种格式进行详细说明。

◆ 简单格式 @Source_Name^Format^Store

在这种格式中，Format 可以是 12.1.1 节所描述的所有数据格式控制符，Store 可以是 12.1.2 节所描述的所有存储方式控制符。

这种格式是唯一能够使用自动文本抽取功能的格式，也就是说：如果 Source_Name 是一个格式化文档，TRS 全文数据库系统能够自动抽取这个文档中的文本数据以建立索引。但系统提供的内嵌功能只能抽取 HTML 和 XML 文档中的文本，其它格式文档中的文本抽取则需要 OEM 插件的支持才能实现。TRS 全文数据库系统是否存储由 OEM 插件所抽取的文本数据，由控制文件的 WAIVE_OEMFILTER_TEXT 变量决定。

自动抽取文本时，TRS 支持大多数流行的字符集编码格式的 HTML 和 XML 文件，这些字符集包括：GB2312、GBK、GB18030、BIG5、UTF-8、ISO8859-1 ~ ISO8859-16 以及 WINDOWS-1250 ~ WINDOWS-1258。

当控制符为 BASE 时，实际上相当于一个空字段，将不会存储任何内容。与控制符 DUMB 不同的是，DUMB 存储方式是不存储数据但要创建索引。

当存储方式控制符为 URL 时，如果 Source_Name 对应的数据格式是 TEXT，或者 HTML，或者 XML，或者安装了相应格式文档的文本抽取 OEM 插件，则系统将自动从远程获取该 URL 的所对应的实际文件（如果需要，则同时完成文本抽取）作为其索引参考文件，以便建立索引。系统以 URL 的方式存储 Source_Name 本身，但并不存储这个临时的索引参考文件。

系统能够自动识别一个 URL，即系统默认以大小写无关的字符串“//”或“http://”或“www.”或“ftp://”引导的 Source_Name，或者在 Win32 下以“\”引导的“网络邻居”资源 Source_Name，都是一个 URL，另外以大小写无关的“file:”引导的 Source_Name 也认为是一个 URL。这样，当存储方式控制符不是 URL，而系统自动识别 Source_Name 是一个 URL 时，系统将会自动获取该 URL 所对应的实际文件来替换

Source_Name, 然后完成数据的加载。

需要说明的是: 系统自动获取 URL 文件的内嵌功能仅限于以“file:”引导的 URL 资源, 或者在 Win32 下以“\\”引导的“网络邻居”上的 URL 资源, 而其它类型的 URL 资源, 则同样需要通过 OEM 插件的支持才能实现。在以下的典型格式和完全格式中的 URL 资源, 以及 BIT 字段值中的 URL 资源, 都是这样处理的。

有关获取 URL 资源所对应的文件和抽取文档中的文本的 OEM 插件的说明, 请参见《服务器安装》部分的配置文件中的说明。

◆ **典型格式 @Source_Name | Refer_Name^Format^Store**

在这种格式中, Format 可以是 12.1.1 节所描述的所有数据格式控制符, Store 可以是 12.1.2 节所描述的所有存储方式控制符。

其中数据格式控制符控制的对象是 Source_Name, 而不是 Refer_Name。索引参考文件 Refer_Name 的数据格式的确定规则是: 如果 Refer_Name 没有文件名后缀, 则默认其数据格式为 TEXT, 否则需要进一步判断这个文件名后缀是否对应了一个数据格式, 如果有这样的数据格式, 则 Refer_Name 的数据格式即为该格式, 否则进行错误处理。

与数据格式控制符不同, 存储方式控制符控制的对象是不一致的: LINK、URL 和 BASE 控制的是 Source_Name; DUMB 控制的是 Refer_Name; 其它存储方式控制符则既控制 Refer_Name, 又控制 Source_Name。

对于 BASE 存储方式控制符, 由 Refer_Name 给出的索引参考文件必须是 HTML 格式文件, 否则不能加载该记录。

如果存储方式控制符为 URL, 则系统存储 Source_Name 本身, 否则当系统自动识别 Source_Name 为一个 URL 时, 将自动从远程获取该 URL 的所对应的实际文件来替换 Source_Name, 然后完成数据的加载。

与 Source_Name 一样, 如果系统自动识别 Refer_Name 为一个 URL 时, 将自动从远程获取该 URL 的所对应的实际文件来替换 Refer_Name, 然后完成数据的加载。

◆ **完全格式 @Source_Name^Format^Store | Refer_Name^Format^Store**

在这种格式中, Source_Name 和 Refer_Name 具有各自的数据格式控制符和存储方式控制符。控制 Source_Name 和 Refer_Name 的 Format 可以是 12.1.1 节所描述的所有数据格式控制符; 控制 Source_Name 的 Store 可以是 12.1.2 节所描述的所有存储方

式控制符；但控制 Refer_Name 的 Store 只能是 ALONE、CLASS 和 DUMB 三种存储方式控制符，BASE、URL 和 LINK 不能用来控制 Refer_Name。

同样对于 BASE 存储方式控制符，由 Refer_Name 给出的索引参考文件必须是 HTML 格式文件，否则不能加载该记录。

如果 Source_Name 的存储方式控制符为 URL，则系统存储其本身，否则当系统自动识别 Source_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Source_Name，然后完成数据的加载。

与 Source_Name 一样，如果系统自动识别 Refer_Name 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Refer_Name，然后完成数据的加载。

12.2 加载BIT字段的数据

对于 BIT 字段，虽然可以接受直接放在 TRS 格式文件中的文本作为字段值，但作为字段值的其它格式化的数据是不能直接放在 TRS 格式文件中的，所以 DOCUMENT 字段引用文件名作为字段值的方法同样适合加载 BIT 字段的值。

BIT 字段对文件的引用格式要比 DOCUMENT 字段简单得多，只有一种形式：

```
@Source1_Name^Format^Store; Source2_Name^Format^Store; ...
```

其中：

- ◆ 符号 '@'、'^' 'Format' 以及 'Store' 与 DOCUMENT 字段的引用格式中的相应符号具有相同的意义。
- ◆ 分号 ';' 为单字节 ASCII 码，是多个引用文件之间的分隔符，即多值之间的分隔符，也可以用其他字符或者字符串来代替。但不管用什么作为多值之间的分隔符，都要求控制文件中 MULTI_VALUE_SEPARATER 变量的值必须与其相一致。

格式中 Format 可以是 12.1.1 节所描述的所有数据格式控制符，但 BASE 和 DUMB 控制的数据将不会被存储。数据格式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的数据格式的步骤与 DOCUMENT 字段基本相同，只是字段值文件的默认数据格式为 BIT 格式，而不是 TEXT 格式。

格式中 Store 可以是 12.1.2 节所描述的所有存储方式控制符。存储方式控制符可以省略（其引导符也同时省略），在这种情况下，确定一个字段值文件的存储方式的方法是：如果控制文件中 CLASS_BIT_STORE 变量的值不为 TRUE，则按系统缺省的存

储方式来存储字段值文件，否则按分类方式存储字段值文件。

如果 Source_NameX 的存储方式控制符为 URL，则系统存储其本身，否则当系统自动识别 Source_NameX 为一个 URL 时，将自动从远程获取该 URL 的所对应的实际文件来替换 Source_NameX，然后完成数据的加载。

与 DOCUMENT 字段一样，在被引用的字段值文件名中，既可以包含路径，也可以不包含路径，系统能够自动确定其正确的路径，确定方法与 DOCUMENT 字段完全一样（请参见 12.1 节的说明）。

12.3 应用举例

下面是一个加载多媒体数据的例子：

```
<REC>
<正文>= @file1.txt
<REC>
<正文>= @\\server\data\file2.htm ^URL
<REC>
<正文>= @http://server/data/ | file3.html ^BASE
<REC>
<正文>= @http://server/data/file4.doc ^WORD^CLASS
<REC>
<正文>= @file5.doc | file5.txt ^WORD^CLASS
<REC>
<正文>= @file6.doc ^CLASS | file6.txt ^DUMB
<REC>
<正文>= @//server/data/file7.doc | file7.txt ^URL
<REC>
<图片>= @file8.doc^CLASS; file9.txt
<REC>
<图片>= @\\server\data\file10.doc ^URL; www.server.com\data\file11.html
```

这里，文件路径中的子目录分隔符没有 ‘\’ 和 ‘/’ 之分，这两种符号系统能够自动纠正。但对于一个 URL，其路径中的子目录分隔符系统不会进行纠正，必须使用实际的目录分隔符。

第 13 章 日期的加载

日期的加载格式有通用格式、英文格式、无分隔符格式和中文格式四种类型。虽然检索时的日期键值也是这四种类型，但加载格式与检索键值还是有些不同。

13.1 通用格式日期的加载

```
Year[/Month[/Day[/Hour[/Minute[/Second]]]]][A[M]]  
[Year]/Month[/Day[/Hour[/Minute[/Second]]][A[M]]  
[Year]/[Month]/Day[/Hour[/Minute[/Second]]][A[M]]  
[Year]/[Month]/[Day]/Hour[/Minute[/Second]][A[M]]  
[Year]/[Month]/[Day]/[Hour]/Minute[/Second][A[M]]  
[Year]/[Month]/[Day]/[Hour]/[Minute]/Second[A[M]]
```

其中：

- ◇ Year为四位表示的年(如“2008”)或者二位表示的年(如“08”)。
- ◇ Month、Day、Hour、Minute、Second分别为二位或一位数字表示的月、日、时、分、秒。
- ◇ A/M可用P/PM替换。A/AM表示上午，P/PM表示下午，大小写无关。如果它们出现在格式中，则“Hour”不能缺省。
- ◇ 单字节除号‘/’分隔符可分别用单字节的逗号‘,’、句号‘.’、冒号‘:’、减号‘-’以及空格任意替代。
- ◇ 用方括号引起的部分表示可以省略(模糊日期)。
- ◇ Year、Month、Day、Hour、Minute、Second都可以用一个或多个单字节的问号‘?’、百分号‘%’替代，表示模糊日期。

13.2 英文格式日期的加载

```
Month[Day[,Year[ Hour[:Minute[:Second]]]]][A[M]]  
Month[Day],Year[ Hour[:Minute[:Second]]][A[M]]  
Month[Day],Year[ Hour]:Minute[:Second][A[M]]  
Month[Day],Year[ Hour]:[Minute]:Second[A[M]]  
[Day-]Month[-Year[ Hour[:Minute[:Second]]]]][A[M]]  
[Day-]Month[-Year[ Hour]:Minute[:Second]][A[M]]
```

[Day-]Month[-Year[Hour]:[Minute]:Second][A[M]]

DayMonthYear[Hour[:Minute[:Second]]][A[M]]

DayMonthYear[[Hour]:Minute[:Second]][A[M]]

DayMonthYear[[Hour]:[Minute]:Second][A[M]]

其中:

- ◇ Year为四位表示的年(如“2008”)或者二位数字表示的年(如“08”)。
- ◇ Month为英文表示的月份: Jan. Feb. Mar. Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec.; 或: January February March April May June July August September October November December。
- ◇ Day、Hour、Minute、Second分别为二位或一位数字表示的月、日、时、分、秒。
- ◇ A/M可用P/PM替换。A/AM表示上午, P/PM表示下午, 大小写无关。如果它们出现在格式中, 则“Hour”不能缺省。
- ◇ Year与Hour之间用单字节的空格分隔; ‘,’ ‘.’ ‘-’ 分别为单字节的逗号、冒号、减号。
- ◇ 用方括号引起的部分表示可以省略(模糊日期)。

13.3 无分隔格式日期的加载

Year[MonthDay[Hour[Minute[Second]]]]

YY[MonthDay]

其中:

- ◇ Year为四位表示的年(如“2008”); YY为二位表示的年份(如“08”)。
- ◇ Month、Day、Hour、Minute、Second分别为二位数字表示的月、日、时、分、秒。
- ◇ 用方括号引起的部分表示可以省略(模糊日期)。

13.4 中文格式日期的加载

中文格式包括所有用全角字符表示的上述几种格式, 另外还支持如下的格式:

[Year年][Month月][Day日][Hour时][Minute分][Second秒][上午]

其中:

- ◇ Year为四位表示的年(如“2008”或“二〇〇八”或“贰零零捌”)或者二位数字表示的年(如“08”或“〇八”或“零捌”)。其中数字既可为半角, 也可为全角。
- ◇ Month、Day、Hour、Minute、Second分别为二位或一位数字表示的月、日、时、分、秒。
- ◇ 格式中“上午”可用“下午”替换。如果“上午”或“下午”出现在格式中, 则“Hour时”不能缺省。

◇ 用方括号引起的部分表示可以省略（模糊日期）。

另外，还支持\$DATE 和\$TIME 特殊日期（不区分大小写）的入库，其中前者表示当前日期的“年月日”，后者表示当前日期的“年月日时分秒”。这里\$DATE 和\$TIME 还可以用作函数来表示一个相对于“现在”的日期，即：\$DATE(...)、\$TIME(...)，它们的参数都是一个正整数、或者负整数，\$DATE(...)的参数单位是天数，\$TIME(...)的参数单位是秒数。如：\$DATE(7)、\$DATE(-7)、\$DATE(+7)等价于\$DATE(7)；\$TIME(60)、\$TIME(-60)、\$TIME(+60)等价于\$TIME(60)；\$DATE(0)、\$DATE(-0)和\$DATE()等价于\$DATE，\$TIME(0)、\$TIME(-0)和\$TIME()等价于\$TIME。

第 14 章 词典的加载

按用途划分，TRS 全文数据库系统有八种类型的词典，即：分词词典 SEGMENT、附加分词词典 SEGMENTEX、停用词典 STOP、附加停用词典 STOPEX、稀疏词典 SPARSE、主题词典 THESAURUS、同义词典 SYNONYM、反义词典 ANTONYM。其中，分词词典、附加分词词典、停用词典、附加停用词典和稀疏词典我们统称为索引词典，主题词典、同义词典和反义词典则统称为辅助知识词典。另外还有三种辅助索引词典，即歧义词典、二元组词典和单字词典。

词典的内容也能象普通数据库一样从一个外部数据文件中加载，但不支持 11.3 节所说明的批处理数据文件方式进行加载，每次只能加载一个数据文件。辅助知识词典也是由字段组成的，所以它们的数据文件格式与普通数据库的完全一致。这里需要说明的是各种辅助知识词典的字段结构。

14.1 加载主题词典

主题词典的字段结构：

名称	类型	允许多值	值唯一
CTM (主题词)	CHAR	否	否
CLT (族首词)	CHAR	是	否
CBT (上位词)	CHAR	是	否
CNT (下位词)	CHAR	是	否
CPT (等同词)	CHAR	是	否
CUF (替代词)	CHAR	是	否
CCC (组合概念)	CHAR	是	否
CAB (缩略词)	CHAR	是	否
CABF (被缩略词)	CHAR	是	否
CRT (相关词)	CHAR	是	否
CHN (历史注释)	CHAR	是	否
CSN (范围注释)	CHAR	是	否
CLE (外文等同词)	CHAR	是	否
CPPC (后组配概念)	CHAR	是	否

下面是一个包含两条记录的主题词典的标准格式的数据记录文件：

<REC>
<CTM>=关键词
<CLT>=检索语言
<CBT>=检索词
<CNT>=
<CPT>=Guanjianci
<CUF>=
<CCC>=
<CAB>=
<CABF>=
<CRT>=
<CHN>=
<CSN>=07G
<CLE>=Keyword
<CPCC>=
<REC>
<CTM>=检索词
<CLT>=
<CBT>=检索语言
<CNT>=关键词; 叙词
<CPT>=Jiansuoci
<CUF>=
<CCC>=
<CAB>=
<CABF>=
<CRT>=
<CHN>=
<CSN>=07G
<CLE>=Searching words
<CPCC>=

如果一个字段值是多值，则用“多值分隔符”（如‘;’）分隔。“多值分隔符”的确定请参见 15.2.15 节的有关说明。

14.2 加载同义词典

同义词典的字段结构:

名称	类型	允许多值	值唯一
STM	CHAR	是	否

下面是一个包含两条记录的同义词典的标准格式的数据记录文件:

```
<REC>  
<STM>=微机; 计算机; 电脑  
<REC>  
<STM>=单车; 脚踏车; 自行车
```

同义词之间用“多值分隔符”(如‘;’)分隔。“多值分隔符”的确定请参见 15.2.15 节的有关说明。

14.3 加载反义词典

反义词典的字段结构:

名称	类型	允许多值	值唯一
ATM	CHAR	是	否

下面是一个包含两条记录的反义词典的标准格式的数据记录文件:

```
<REC>  
<ATM>=高 ;; 矮; 低  
<REC>  
<ATM>=赞同; 赞成; 同意; 拥护 ;; 反对
```

反义词典数据文件中字段值的格式有点特殊,即由两组意义相反的词组成,这两组词之间用两个连续的“多值分隔符”(如‘;;’)进行分隔,组内的词彼此为同义词,用一个“多值分隔符”(如‘;’)分隔。“多值分隔符”的确定请参见 15.2.15 节的有关说明。

14.4 加载其它词典

除主题词典、同义词典和反义词典外，所有索引词典、辅助索引词典，以及实体词典的数据文件具有统一的、非常简单的格式：

- ◆ 数据文件是一个文本文件。
- ◆ 每一个文本行对应一个词条。
- ◆ 允许空行。

第 15 章 控制文件

15.1 控制文件的格式

控制文件是指存放那些用以控制数据加载过程的变量配置文件。这个文件是在进行数据加载前由用户编写的。

控制文件的一些基本特性是：

- ◆ 控制文件是一个文本文件。
- ◆ 以单字节字符 ‘#’ 开始的行为注释行。注释行可以放在任何位置。
- ◆ 一行只能指定一个变量，其格式为：变量名 = 变量值。其中“变量名”、“=”不能省略，“变量名”和“变量值”的字符串不包括其首尾的空白字符。如果“变量值”省略，则等同于该变量未出现在控制文件中。
- ◆ 变量在文件中的顺序无关。
- ◆ 变量名不区分大小写。除了特别说明的外，变量值也不区分大小写。
- ◆ 在控制文件中，并不是所有的控制变量都必须出现，实际上只需给出那些实际变量值与其缺省值不一致的控制变量。如果用户在数据加载时未指定控制文件，或者只给出了部分控制变量，而那些未出现的控制变量存在缺省值，则系统将自动取其缺省值。

15.2 控制文件的变量

控制文件所支持的变量包括：

变量名	说明	缺省值
RECORD_FILE_FORMAT	数据记录文件的格式	标准
BIT_FORMAT_DEFAULT	BIT 字段值文件的缺省数据格式名	BIT
HTML_BASE_DEFAULT	HTML 数据文件中 BASE 元素的缺省 HREF 属性值	无
DATA_PATH_DEFAULT	字段值所引用的数据文件的缺省路径	当前数据文件路径
DATE_CENTURY_DEFAULT	DATE 字段值的年份缺省世纪数（从 0 开始计数，即 0 表示 1 世纪）	当前世纪数
IGNORE_RECORD_CRYPT	是否忽略记录加密描述符的开关	FALSE

变量名	说明	缺省值
SYNC_CREATE_INDEX	在加载数据的同时建立索引的开关	TRUE
FILE_SUFFIX_PRIOR	所引用的字段值文件未指定数据格式控制符时，用文件名的后缀来确定其数据格式的开关	TRUE
CLASS_BIT_STORE	是否分类存储 BIT 字段值的开关	TRUE
CLASS_DOCUMENT_STORE	是否分类存储 DOCUMENT 字段值的开关	TRUE
WAIVE_DOCUMENT_REFERER	是否丢弃 DOCUMENT 字段的索引参考文件的开关	FALSE
WAIVE_OEMFILTER_TEXT	是否丢弃由 OEM 过滤器所生成的文本数据的开关	TRUE
KEEP_TEXT_ACTUAL	是否保持 PHRASE 和 DOCUMENT 字段值文本中首尾空白行的开关	TRUE
CHECK_TEXT_ITERATIVE	是否需要检查和处理 PHRASE 或 DOCUMENT 字段在同一个记录中出现重复字段值的开关	见 15.2.14 节的说明
MULTI_VALUE_SEPARATER	多值字段 DATE、NUMBER、CHAR、BIT 的各个字段值之间的分隔符	单字节 ‘;’
SECTION_TAG_PREFIX	TRS 标记引导符前缀	无
START_RECORD_FROM	在数据记录文件中数据加载的起始逻辑记录号（序号）	1
PROCESS_RECORD_NUMBER	需要加载的最大记录数	全部
MAX_COMMIT_NUMBER	加载时分批提交的最大记录数	4000
MAX_ERROR_NUMBER	加载时允许错误记录的最大个数	最大
COLUMN_NAME_SEQUENCE	需要加载数据的字段名列表	全部
XML_SEGMENT_MARK	XML 数据作为一个字段值入库时，过滤器进行自动分段的标记	无
XML_NEWLINE_MARK	XML 数据作为一个字段值入库时，过滤器进行自动换行的标记	无
XML_HITSHOW_MARK	XML 数据作为一个字段值入库后，检索命中点的指示标记	无
XML_FILTER_CDATA	XML 数据作为一个字段值入库时，过滤器是否处理标记中的内容	FALSE
XML_OVERWRITE_PI	XML 数据作为一个字段值入库时，是否需要覆盖数据中已有的关于过滤器的处理指令	FALSE
COLUMN_AUTO_MATCH	是否允许自动匹配入库字段	FALSE

变量名	说明	缺省值
NAMIZE_ISO2709_FIELD	是否以 ISO2709 格式的目录标记(以及数据标识符)映射入库字段	FALSE
THROUGHT_ACCESS	是否允许以“直通模式”加载记录的开关	FALSE
MASK_OEMFILTER_ERROR	是否允许屏蔽 OEM 过滤器(包括内嵌的 HTML 和 XML 过滤器)的错误。如果过滤器报出错误, 但被屏蔽后, 系统对过滤失败的文档不建索引。	FALSE

15.2.1 RECORD_FILE_FORMAT

该变量指明数据记录文件的组织格式。TRS 支持的数据记录文件格式包括:

名称	数据记录格式
STANDARD	标准格式
IDENTISTD	字段内部标识号标记格式
SIMPLESTD	字段顺序号标记格式
SEQUENCE	无字段标记格式
FULLTEXT	全文格式
TRANSPORT	传输格式
ISO2709	ISO2709 国际标准格式
XML	XML 国际标准格式

如果控制文件中未包含该变量, 则系统取其缺省值, 即 STANDARD。TRANSPORT 格式允许将包括字段值文件(即其它格式中的外挂文件)在内的所有装库数据打包成一个文件。

15.2.2 BIT_FORMAT_DEFAULT

该变量指明数据记录文件中, BIT 字段引用的字段值文件的缺省数据格式控制符。该变量的值必须是一个当前系统中数据格式词典的一个合法数据格式控制符。该变量只有当 FILE_SUFFIX_PRIOR 变量的值为 FALSE, 或者当引用的字段值文件名没有后缀时, 才有效。有关 BIT 字段值文件的数据格式的确定方法详见 12.2 节。

如果控制文件中未包含该变量, 则系统取其缺省值, 即 BIT。

15.2.3 HTML_BASE_DEFAULT

该变量指明 DOCUMENT 字段引用的 HTML 格式的字段值文件中 BASE 元素的缺省 HREF 属性值。它只对对在 DOCUMENT 字段上以“简单格式”加载 HTML 数据时有效，其作用相当于在“典型格式”和“完全格式”中使用显式的 BASE 控制符。相关内容请参见 12.1.2 节的说明。

该变量系统没有缺省值。

15.2.4 DATA_PATH_DEFAULT

该变量指明数据记录文件中，DOCUMENT 和 BIT 字段引用的字段值文件的缺省路径。该变量的值必须是一个当前用户有权访问的合法的目录。缺省路径名是大小写敏感的，并且路径的子目录分隔符没有‘\’和‘/’之分，这两种符号系统能够自动纠正。有关字段值文件所在路径的确定方法详见 12.1 节。

如果控制文件中未包含该变量，则系统取其缺省值，即数据记录文件所在的路径。对于批处理数据加载，其缺省值总是当前正在被处理的数据记录文件所在的路径。因此，该变量总是设置了有效值。

15.2.5 DATE_CENTURY_DEFAULT

TRS 支持多种日期字段的入库格式。有关日期字段的入库格式请参见第 13 章。

有的日期格式允许只给出两位数表示年份（如 08）。在这种情况下，需要确定该年份所在的世纪数，通过该变量即可设置这一世纪数。如将变量的值设置为 20，则年份 08 表示 2008 年。如果设置变量的值为 0，则表示日期格式中的任何年份都为绝对年份，如 08 即表示公元 8 年。该变量的值必须是一个数字。

如果控制文件中未包含该变量，则系统取其缺省值，即当前世纪数，**但必须注意的是：如果年份（如 80）大于当前实际年份（如 08），则默认位是上一世纪的年份（如 1980），而不是本世纪的年份（如 2080）。**该变量只对那些小于 100 的年份有效。

15.2.6 IGNORE_RECORD_CRYPT

TRS 支持对数据记录进行加密。有关加密方法详见 11.1 节。

如果数据记录文件中有加密描述的记录，而又想在本次入库时忽略这些加密描述，

即不对记录进行加密，则可通过设置这一变量来完成。如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示忽略所有记录的加密描述；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示对那些有加密描述的记录进行加密。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

15.2.7 SYNC_CREATE_INDEX

TRS 有两种方式来创建数据库的索引：一种是在加载数据本身的同时建立数据的索引，即同步索引方式；另一种是先只加载数据本身，以后再单独建立数据的索引，既异步索引方式。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示以同步方式建立索引；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示以异步方式建立索引。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。

特别注意：如果在向数据库中加载、插入或复制新记录时，指定了“不同步建立索引”，则以后加载、插入或复制的新记录也不会建立索引，无论是否指定了“同步建立索引”；甚至在以后修改数据库记录时，如果因修改了 PHRASE 或 DOCUMENT 字段的内容，而需要更换物理记录号时，则更新后的整个记录也不会建立索引。除非使用“创建索引”功能为这些尚未建立索引的记录创建索引，以后的加载、插入、复制（“同步建立索引”方式）或修改后的新记录才会建立索引。

15.2.8 FILE_SUFFIX_PRIOR

在 TRS 格式文件中，支持以文件名的形式给出 DOCUMENT 和 BIT 字段值，同时支持对这些文件显式地指定其数据格式。有关内容详见第 12 章。

当未指定数据格式，而又需要系统通过文件名的后缀来自动识别其文件的数据格式时，可通过设置该变量来触发系统的自动识别功能。如果变量值为 TRUE（或 YES、或 ON、或 1），则表示可由文件后缀来确定其数据格式；如果变量值为 FALSE（或 NO、或 OFF、或 0），则表示按字段类型确定其缺省数据格式。BIT 字段值的缺省数据格式为 BIT_FORMAT_DEFAULT 变量的值，DOCUMENT 字段值的缺省数据格式为 TEXT。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。该变量只对那些没有显式指定其数据格式的字段值文件有效。

15.2.9 CLASS_BIT_STORE

由于 BIT 字段的数据一般都是数据量较大的诸如声音、图象以及动画之类的多媒体数据，所以正如 12.2 节所描述的那样，对于 BIT 字段的数据，TRS 允许将其存储到数据库的 BIT 分类文件中，这样就可以缩小数据库主文件的规模，从而有利于提高检索结果的概览速度。

在数据记录文件中，当以引用文件的形式给出字段值时，如果对这些文件没有显式地定义相应的存储方式控制符，而又要求系统按分类方式存储这些字段值，则可通过设置该变量实现这一功能。该变量同样能够控制那些直接放在数据记录文件中文本类型的字段值的存储方式。请参见 12.2 节的说明。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示将字段值存储到数据库的 BIT 分类文件中；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示将字段值集中存储到数据库的主文件中。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。该变量只对那些没有明确指定存储方式的 BIT 字段的字段值有效。

15.2.10 CLASS_DOCUMENT_STORE

虽然 DOCUMENT 字段的数据的数据量一般不是很大，但通常不用于概览，所以从有利于提高检索结果的概览速度的角度出发，TRS 也允许将其存储到数据库的 DOCUMENT 分类文件中，以缩小数据库主文件的规模，达到加速概览的目的。

在数据记录文件中，当以引用文件的形式给出的字段值时，如果没有对这些文件显式地定义相应的存储方式控制符，而又要求系统按分类方式存储这些字段值，则可通过设置该变量实现这一功能。该变量同样能够控制那些直接放在数据记录文件中文本类型的字段值的存储方式。请参见 12.1 节的说明。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示将字段值存储到数据库的 DOCUMENT 分类文件中；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示将字段值集中存储到数据库的主文件中。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。该变量只对那些没有明确指定存储方式的 DOCUMENT 字段的字段值有效。

15.2.11 WAIVE_DOCUMENT_REFERER

为了完成对非 TEXT、HTML 和 XML 格式的数据进行索引，在没有配置 OEM 过滤器，或者 OEM 过滤器不支持所给出的文件的格式情况下，TRS 要求外部为其提供一个 TEXT、或者 HTML、或者 XML 格式的索引参考文件。对于那些没有明确指定存储方式的索引参考文件，其存储方式将参考该变量和 CLASS_DOCUMENT_STORE 变量的设置情况综合确定。具体方法详见 12.1 节。

如果将该变量的值为 TRUE（或 YES、或 ON、或 1），则表示数据库中不存储这个索引参考文件；如果变量值为 FALSE（或 NO、或 OFF、或 0），则表示数据库中存储这个索引参考文件，存储方式由 CLASS_DOCUMENT_STORE 变量的值确定。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。该变量只对那些没有明确指定存储方式的 DOCUMENT 字段的索引参考文件有效。

15.2.12 WAIVE_OEMFILTER_TEXT

TRS 支持配置 OEM 过滤器自动抽取 OFFICE 文档的文本数据以建立索引。该变量确定将这些文本数据是否存储到数据库中。

如果将该变量的值设置为 TRUE（或 YES、或 ON、或 1），则表示数据库中不存储这些文本数据；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示数据库中存储这些文本数据。

在该变量的值设置为 FALSE（即存储自动抽取所得到的文本数据）的情况下，如果源格式文档（待抽取文本的格式化文档）的存储方式为 URL 或 LINK，则文本的存储方式取决于变量 CLASS_DOCUMENT_STORE 的值；如果源格式文档的存储方式为 DUMB，则文本的存储方式取决于变量 CLASS_DOCUMENT_STORE 的值，并且可达到只存储文本数据，而不存储源格式文档数据的应用效果；如果源格式文档采用其它存储方式，则文本的存储方式与源格式文档的相同。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。

15.2.13 KEEP_TEXT_ACTUAL

对于 PHRASE 和 DOCUMENT 的文本类型字段值而言，位于其首尾的空白行是无用的，如果需要在数据加载时将它们去掉，则可通过设置该变量来触发系统完成对首尾空白行的过滤。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示保持字段值文本的原样；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示过滤字段值文本的首尾空白行。

如果控制文件中未包含该变量，则系统取其缺省值，即 TRUE。

15.2.14 CHECK_TEXT_ITERATIVE

PHRASE 和 DOCUMENT 字段不支持多值。但在数据记录文件中，可能会有各种原因，使得在一个记录中同一个 PHRASE 或 DOCUMENT 字段出现一次以上。对于这种情况，系统缺省的处理方式是：PHRASE 只保留最后出现的值；DOCUMENT 只保留最后出现的两个值，最后一个值作为该字段的文本值文件（索引参考文件）的内容，而倒数第二个值则作为该字段的源格式文档（附件）的内容。通过设置该变量可以强制系统进行检查，当发生这种情况时就进行错误处理，或者将其合并成一个字段值。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示一个记录中如果多次出现同一个 PHRASE 或 DOCUMENT 字段，则认为该记录的数据是错误的，进行错误处理；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示按系统的缺省方式进行处理。

如果将变量值设置为 JOIN0、或 JOIN1、或 JOIN2，则表示将多次出现的 PHRASE 或 DOCUMENT 字段值进行合并，并在这些值之间分别插入 0 到 2 个（取决于 JOIN 后面的数字）换行符。需要注意的是：只要 DOCUMENT 字段有一个值是引用的外部文件，则就不能进行字段值的合并，而总是进行错误处理。

如果控制文件中未包含该变量，则系统取其缺省值：如果数据记录文件的格式是 ISO2709 或 XML，则缺省值为 JOIN1，否则为 FALSE。

15.2.15 MULTI_VALUE_SEPARATER

该变量用来给出数据记录文件中，DATE、NUMBER、CHAR 以及 BIT 多值字段的各个值之间所采用的分隔符。多值分隔符既可以是单个字符，也可以是多个字符（如全角字符 ‘;’），但最多为 8 个字符，并且大小写敏感。空格不能作为多值分隔符。

如果控制文件中未包含该变量，则系统取其缺省值，即单字节分号 ‘;’。

15.2.16 SECTION_TAG_PREFIX

该变量指定 TRS 格式文件中 TRS 标记引导符的前缀。TRS 支持标记引导符前缀的目的就是要排除数据中的干扰，所以标记引导符前缀应该选择一个在数据中不会出现的字符序列，至少不会在一行的开头出现。一个引导符前缀的最大长度为 30 个字符，并且其首尾的空白字符将被忽略，即首尾的空白字符不是引导符前缀的有效部分。

该变量系统没有缺省值。

15.2.17 START_RECORD_FROM

该变量指定从数据记录文件中第几个记录开始加载。对于批处理方式，该变量只对第一个数据记录文件有效，对于其他数据记录文件，总是从第一个记录开始加载。

如果控制文件中未包含该变量，则系统默认从第 1 个记录开始加载。如果设置该变量，则其值必须大于 0。

15.2.18 PROCESS_RECORD_NUMBER

该变量指定要处理的总记录数，包括加载成功的记录和错误记录。如果控制文件中未包含该变量，则系统默认加载给出的所有数据记录文件中的全部记录。如果设置该变量，则其值必须大于 0。

15.2.19 MAX_COMMIT_NUMBER

该变量就是通常所说“最大提交数”。一个数据记录文件往往包含许多记录，系统总是将这些记录分批地加载到数据库中，该变量就是用来设定每批加载的最大记录数的。每一批实际加载多少个记录，系统内部还需要根据机器内存的大小综合考虑，但不会超过这一数。

最大提交数与数据的加载速度有一定的关系，合适的值会提高加载速度，反之则减慢加载速度，尤其是在数据库采用按字索引策略或二元组索引策略的情况下更为突出。对不同的数据有不同的最佳值，一般选在 2000 到 6000 之间比较合适，以 5K 字节大小为参考，数据记录较大时使用较小的数，数据记录较小时使用较大的数。

如果控制文件中未包含该变量，则每一批实际加载的记录数由系统完全根据机器内存的大小自动确定。如果设置该变量，则其值必须大于 0。

15.2.20 MAX_ERROR_NUMBER

该变量指定：如果在加载过程中所遇到的错误记录数达到了这里指定的最大记录数，则结束加载过程。如果控制文件中未包含该变量，则系统默认加载过程不受错误记录数的控制。如果设置该变量，则其值必须大于 0。

15.2.21 COLUMN_NAME_SEQUENCE

该变量用来指定需要加载数据的字段名（或别名）列表，字段名之间用单字节的分号 ‘;’、逗号 ‘,’ 或空格分隔。在下列几种情况下需要设置该变量：

- ◆ 只需从数据记录文件中加载部分数据库字段的值。
- ◆ 数据记录文件的记录中存在不属于当前加载的数据库的字段。
- ◆ 数据记录文件的格式为 STANDARD 或 XML 或 ISO2709（“目录项标记”与字段别名的映射方式），而数据记录文件中的某些或所有的字段标记使用了字段别名。
- ◆ 数据记录文件的格式为 SIMPLESTD 或 SEQUENCE 或 ISO2709（其中的“目录项标记”与字段顺序的映射方式），而数据记录文件的记录中字段的顺序与数据库字段的物理顺序（即输出时的顺序）不完全一致。
- ◆ 数据记录文件的格式为 FULLTEXT，而加载的目标字段不是数据库的最后一个 DOCUMENT 字段；或者数据库没有 DOCUMENT 字段，而加载的预期字段不是数据库的最后一个 PHRASE 字段；或者数据库有 DOCUMENT 字段，而加载的预期字段是其他 PHRASE 字段。

对于同一个字段，如果这里给出的是字段名，则数据文件中只能用字段名作为“字段标记”；如果这里给出的是字段别名，则数据文件中不能用其它字段别名作为“字段标记”，但字段名仍可以作为“字段标记”。请参见变量 COLUMN_AUTO_MATCH。

该变量的系统缺省值要依数据格式而定：当数据格式为 FULLTEXT 时，如果数据库存在 DOCUMENT 字段，则取数据库的最后一个 DOCUMENT 字段；如果数据库不存在 DOCUMENT 字段，但是存在 PHRASE 字段，则取数据库的最后一个 PHRASE 字段；其他情况则不能进行数据加载。对于其他数据格式，则取数据库的所有字段，并按其物理顺序排列。

15.2.22 XML_SEGMENT_MARK

该变量用来指定 XML 数据作为一个字段值入库，自动抽取文本进行索引时，通知过滤器进行自动分段的标记，它必须是这个 XML 数据中的一个有效的以成对方式出现的标记。在这一标记对之间的数据为一自然段落。一般情况下，为了控制 XML 数据的显示效果，通常会在 XML 数据中插入必要的标记以使相应的解释器进行分段

处理, 这样的标记就可作为变量 XML_SEGMENT_MARK 的值, 注意该标记的大小写是敏感的, 最大长度是 60 个字符。

在一个段落中, 过滤器将忽略所有的回车换行符, 以避免一个词被 TRS 分词系统中的换段规则所打断而无法检索。

该变量没有缺省值, 它控制的是过滤器输出的文本, 而不是浏览时输出的 XML 文件。在没有分段标记的情况下, 系统将根据内部规则, 按“换行”状态自动确定文本的段落(请参见 3.4.1 的说明)。对于那些游离于所有分段标记对之外的内容也按这种方式处理。

下面举一个例子:

```
<paragraph id="d11">  
    存储器可以存放最后一次接受的参考生物特征图像。存储器可以  
<br/>  
    存放过去输入的多个的参考生物特征图像, 其包括以前接受的生物特  
<br/>  
    <!-- 1 -->  
    <dp n="d1"/>  
    征图像。确定器可以计算输入生物特征图像和各参考生物特征图像之  
<br/>  
    间的相似程度, 当所有计算的相似程度小于预定阈值时, 确定器确定<br/>  
    输入的生物特征图像可以接受。<br/>  
</paragraph>
```

在这个例子中, 标记“paragraph”是 XSL 用来进行换段控制的, 所以我们就将 XML_SEGMENT_MARK 变量指向这个标记。如果不设置该变量, 过滤器无论如何处理, 其中第三个“特征”将不能被检索。

15.2.23 XML_NEWLINE_MARK

该变量用来指定 XML 数据作为一个字段值入库, 自动抽取文本进行索引时, 通知过滤器进行自动换行的标记, 它必须是该 XML 数据中的一个有效标记, 否则过滤器将不会输出任何换行符(因为不可能遇到这个标记)。一般情况下, 为了控制 XML 数据的显示效果, 通常会在 XML 数据中插入必要的标记以使相应的 XML 浏览器进行换行, 这样的标记就可作为变量 XML_NEWLINE_MARK 的值, 注意该标记的大小写

是敏感的，最大长度是 60 个字符。如 15.2.22 节例子中的“br”标记。

如果设置了变量 XML_NEWLINE_MARK，则过滤器将忽略 XML 数据中的回车换行符，直至遇到由该变量所指定的标记时才进行自动换行。若未设置该变量，则将以 XML 数据中原有的回车换行符作为过滤器的换行符。但正如 15.2.22 节的说明，过滤器将忽略处在变量 XML_SEGMENT_MARK 所指定的段落标记中的换行标记，以及 XML 数据中原有的回车换行符。

与指定分段标记一样，指定换行标记也有避免一个词被 TRS 分词系统中的换段规则所打断而无法检索的作用。

该变量没有缺省值，它控制的是过滤器输出的文本，而不是浏览时输出的 XML 文件。但是，如果在 XML 文件中定义了 DTD，且 DOCTYPE 为 TRS，DOCTYPE 的子元素为 REC 的前提下，将“br”定义成一个“空”元素，即定义：<!ELEMENT br EMPTY>，则取 br 标记作为该变量的缺省值（请参见 11.2 节的说明）。除此之外，系统将以 XML 文件中的自然换行符作为文本的换行符。

15.2.24 XML_HITSHOW_MARK

该变量用来指定 XML 数据作为一个字段值入库后，指示检索命中点的标记。这个标记有以下几个特点：

- ◆ 其大小写是敏感的，并且被浏览器所识别，最大长度是 60 个字符。
- ◆ 在入库前，XML 数据中并不存在这个标记。
- ◆ 在检索时，由 TRS 全文数据库系统自动在输出的 XML 数据中插入。
- ◆ 标记以成对方式出现：在命中点之前插入起始标记，在命中点之后插入结束标记。
- ◆ 在起始标记中可能会出现一个名为“ID”的属性，其值为命中点的从“1”开始计数的序号。没有出现这个属性的起始标记，说明是一个命中点被其它标记所打断后的重新开始。属性值也有为“-1”的情况，请参见 15.2.25 节的说明。
- ◆ 如果指定了段落标记（请参见 15.2.22），则在命中点起始标记中将会有有一个标记属性“SEG”，其值为截止到命中点所在位置，段落结束标记出现的次数（从 0 开始计数）；同样，如果指定了分句标记（请参见 15.2.30），则在命中点起始标记中将会有有一个标记属性“SEN”，其值为截止到命中点所在位置，分句结束标记出现的次数（从 0 开始计数）。【修订 088】

该变量没有缺省值，它控制的是浏览时输出的 XML 文件中的检索命中点，与过滤器输出的文本没有关系。

下面举一个例子：

```

<paragraph id="d11">
    存储器可以存放最后一次接受的参考生物<TRSHL ID="1"> 特征
    </TRSHL> 图像。存储器可以<br/> 存放过去输入的多个的参考生物
    <TRSHL ID="2"> 特征</TRSHL> 图像，其包括以前接受的生物<TRSHL
    ID="3"> 特</TRSHL><br/>
        <!-- 1 -->
        <dp n="d1"/>
    <TRSHL> 征</TRSHL> 图像。确定器可以计算输入生物<TRSHL ID="4">
    特征</TRSHL> 图像和各参考生物<TRSHL ID="5"> 特征</TRSHL> 图像
    之<br/> 间的相似程度，当所有计算的相似程度小于预定阈值时，确定器
    确定<br/> 输入的生物<TRSHL ID="6"> 特征</TRSHL> 图像可以接受。
    <br/>
</paragraph>

```

在这个例子中，“TRSHL”是由 XML_HITSHOW_MARK 变量指定标记。注意其中第三个“特征”命中点被其它标记所打断。“TRSHL”标记在 XSL 中可以用下面的方式进行处理（用红色反显）：

```

<xsl:template match="TRSHL">
    <font color="red"><xsl:value-of select="."/></font>
</xsl:template>

```

15.2.25 XML_FILTER_CDATA

该变量用来指定 XML 数据作为一个字段值入库，自动抽取文本进行索引时，过滤器是否处理 CDATA 标记部分的内容。一般情况下，CDATA 数据不需要（或者说不能）进行索引，但有的 XML 数据中的 CDATA 部分也要求能被检索，有时还要求标记命中点。

- ◆ 如果将变量 XML_FILTER_CDATA 设置成 FALSE（或 NO、或 OFF、或 0），则表示不过滤 CDATA 数据中的任何标记，所有内容及标记都建立索引。检索结果输出时，如果其中有命中点，则由 XML_HITSHOW_MARK 变量所指定的命中点标记不是直接在 CDATA 数据中插入，而是在其结束符后插入一个空标记（标记之间没有任何数据），其“ID”属性的值为“-1”。
- ◆ 如果设置成 TRUE（或 YES、或 ON、或 1），则表示过滤 CDATA 数据中的所有标记，只索引 CDATA 数据中非标记部分。检索命中点的输出方式与取值为 FALSE

时相同。

- ◆ 如果设置成 HIT，则表示过滤 CDATA 数据中的所有标记，只索引 CDATA 数据中非标记部分。检索结果输出时，如果其中有命中点，则由 XML_HITSHOW_MARK 变量所指定的命中点标记直接在 CDATA 数据中的准确位置插入。
- ◆ 如果设置成 ROW，则表示不过滤 CDATA 数据中的任何标记，但去掉所有回车换行符和多余的空白符，把全部的内容及标记都看作处在同一文本行进行索引。检索命中点的输出方式与取值为 FALSE 时相同。
- ◆ 如果设置成 ANY，则表示去掉 CDATA 数据中的全部内容，即任何内容（标记和非标记）都不建立索引。

该变量缺省值为 FALSE，它控制的是过滤器输出的文本，而不是浏览时输出的 XML 文件。

15.2.26 XML_OVERWRITE_PI

该变量用来指定 XML 数据作为一个字段值入库，自动抽取文本进行索引时，是否需要覆盖 XML 数据中已有的关于过滤器的处理指令。

为了能够在 XML 中准确指示检索命中点的位置，命中点跟踪器也需要使用过滤器完全一致的参数，所以 TRS 全文数据库系统会根据控制文件给出的各个相关变量，在 XML 数据入库时，自动插入一条 XML 处理指令。XML 过滤器以及跟踪器都将引用这条处理指令，其格式如下：

```
<?trs-parser ATTRIBUTES ?>
```

ATTRIBUTES 由下列一个或多个分项组成：

- ◆ SegmentMark = “Mark”。其中 “Mark” 是由变量 XML_SEGMENT_MARK（请参见 15.2.22 节的说明）指定的段落标记，大小写有关，最大长度是 60 个字符。没有缺省值。
- ◆ SentenceMark = “Mark”。其中 “Mark” 是由变量 XML_SENTENCE_MARK（请参见 15.2.30 节的说明）指定的分句标记，大小写有关，最大长度是 60 个字符。没有缺省值。
- ◆ NewLineMark = “Mark”。其中 “Mark” 是由变量 XML_NEWLINE_MARK（请参见 15.2.23 节的说明）指定的换行标记，大小写有关，最大长度是 60 个字符。没有缺省值。
- ◆ HitShowMark = “Mark”。其中 “Mark” 是由变量 XML_HITSHOW_MARK（请参见 15.2.24 节的说明）指定的命中点指示标记，大小写有关，最大长度是 60 个字符。没有缺省值。

- ◆ FilterCDATA = “Mode”。其中 “Mode” 的取值包括：“no”、“yes”、“hit”、“row”以及 “any”，分别对应于变量 XML_FILTER_CDATA（请参见 15.2.25 节的说明）的取值：“FALSE”、“TRUE”、“HIT”、“ROW” 和 “ANY”。缺省值为 “no”。其中 “no” 可用 “FALSE”、“OFF”、“0” 替代；“yes” 可用 “TRUE”、“ON”、“1” 替代。

你也可以手工创建或修改这条指令，但必须注意 ATTRIBUTES 的书写要遵循 XML 的规范，分项之间用空格分隔，给出的每个分项都必须是完整的。指令中，作为属性值的标记名是区分大小写的（可以用双引号或者单引号引起来），其他字符则是大小写无关的。指令必须出现在 XML 的 DTD 部分之前，并只允许出现一次。下面是一个例子：

```
<?xml version="1.0" encoding="GB2312"?>
<?xml-stylesheet type="text/xsl" href="xsl0117.xsl"?>
<?trs-parser segmentmark="paragraph" newlinemark="br" hitshowmark="trshl"?>
<!DOCTYPE patent SYSTEM "dtd0117.dtd">
.....
```

在一般情况下，如果过滤器发现 XML 数据中已经有了这条处理指令，无论其属性如何，都会优先使用指令中给出的属性参数，而忽略控制文件给出的变量值。但有的情况可能需要用控制文件中给出的变量值重新构造这条处理指令。

如果将变量 XML_OVERWRITE_PI 设置成 FALSE（或 NO、或 OFF、或 0），则表示不覆盖 XML 数据中已有的过滤器处理指令；如果将变量的值设置成 TRUE（或 YES、或 ON、或 1），则表示重新构造并覆盖 XML 数据中已有的过滤器处理指令。

该变量缺省值为 FALSE。

15.2.27 COLUMN_AUTO_MATCH

如 15.2.21 所述，一般地，如果数据文件中的“字段标记”对应于数据库的字段别名，则可以设置 COLUMN_NAME_SEQUENCE 变量，以指明这些字段别名。使用字段别名进行加载数据，更简便的方法是将 COLUMN_AUTO_MATCH 变量设置成 TRUE（或 YES、或 ON、或 1），让系统根据数据文件中实际给出的“字段标记”（数据库的字段名或字段别名）自动地与数据库字段进行匹配。但是，当数据文件的记录中存在无法“自动匹配”的字段时，仍需根据 COLUMN_NAME_SEQUENCE 变量的设置以忽略这些字段。必须指出：这个功能在下面这些儿情况下会影响系统性能：在

数据中存在错误的“字段标记”；同一个字段有多个别名同时作为“字段标记”。因此一般情况不建议使用该功能。

如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示不允许 TRS 全文数据库系统进行字段自动匹配，而是按照 COLUMN_NAME_SEQUENCE 变量的设置情况进行处理。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

15.2.28 NAMIZE_ISO2709_FIELD

对于 ISO2709 格式的入库，缺省情况是：ISO2709 格式中的“目录项标记”对应于 COLUMN_NAME_SEQUENCE 变量（或缺省值）给出的字段列表中的字段所在的序号（列表中的第一个字段的序号为“001”）。通过 NAMIZE_ISO2709_FIELD 的设置可以改变这种映射关系。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示将 ISO2709 格式中的“目录项标记”（如“010”）将对应于数据库字段的字段名或别名（如字段名或别名为“010”的字段）；同时还支持抽取数据域的各个子项分别单独入库，即将“目录项标记”（如“010”）与数据域的各个子项标识符（如“a”）进行结合，形成一个“子字段名”（如“010a”），如果在入库目标字段列表中存在该字段名，则对应的数据子项将单独加载到这个数据库字段中。另外，由于数据库的字段名是不区分大小写的，当子项标识符是一个大写字母时，必须按特殊规则处理：将“目录项标记”（如“010”）与数据子项标识符（如“A”）进行结合时用下划线进行连接（如“010_A”），因此其对应的数据库字段名或别名也必须遵循这一规则。**如果使用字段的别名，则必须在控制文件中用 COLUMN_NAME_SEQUENCE 变量列出所有需要加载数据的字段所对应的别名，除非设置 COLUMN_AUTO_MATCH 变量为 TRUE。**

如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示格式中的“目录项标记”域是该字段在控制文件字段列表中的序号（从 1 开始计数）。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

15.2.29 THROUGH_ACCESS

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示允许以“直通模式”加载数据。如果数据加载所涉及到所有原始数据文件、批处理文件、控制文件，服务器的启动帐号都有读取权限，则可使用直通模式进行数据加载。由于直通模式下服务

器直接读取原始数据文件，从而减少了中间环节，所以效率会更高，但要求客户端应用与服务器必须运行在同一机器中。如果将变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示以常规方式加载数据。^{【修订 064】}

有一种特殊情况也可以使用直通模式加载数据，即客户端应用与服务器通过网络盘等共享的方式，将加载过程需要访问的文件放在共享盘上供双方共享。在这种情况下使用直通模式，必须保证双方访问这些文件时的路径是一致的，并且将该变量的值必须设置成 SHARE，即 THROUGHT_ACCESS=SHARE。^{【修订 155】}

该变量缺省值为 FALSE。

15.2.30 XML_SENTENCE_MARK

该变量用来指定 XML 数据作为一个字段值入库时，通知过滤器进行人工“切分句子”的标记（即分句标记），它必须是该 XML 数据中的一个有效标记，否则过滤器在一个段落内将不会进行任何分句（一段即是一句）。注意该标记的大小写是敏感的，最大长度是 60 个字符，并且以结束标记为切分点。

该变量系统没有缺省值。在没有分句标记的情况下，系统将根据内部规则，按文中的标点符号自动确定文本的句子（请参见 3.4.1 的说明）。对于那些游离于所有分句标记对之外的内容也按这种方式处理。

如果设置了变量 XML_SENTENCE_MARK，则过滤器将输出人工“切分句子”所要求的格式文本（请参见 3.4.1 的说明），文中的自然标点符号将不起作用。^{【修订 087】}

15.2.31 MASK_OEMFILTER_ERROR

该变量用来指定 OFFICE 文档及 XML 和 HTML 文件作为一个字段值入库时，通过配置的 OEM 过滤器或内嵌的 XML 和 HTML 格式过滤器抽取文本进行索引时是否忽略过滤器报出的错误。

如果将变量值设置为 TRUE（或 YES、或 ON、或 1），则表示忽略（屏蔽）过滤器报出的错误；如果变量值设置为 FALSE（或 NO、或 OFF、或 0），则表示不要忽略过滤器报出的错误。如果过滤器检测到了数据格式等方面的错误，但按设置要求被忽略后，系统对过滤失败的文档不建索引。

如果控制文件中未包含该变量，则系统取其缺省值，即 FALSE。

15.2.32 XML_KEEP_CDATA

该变量用来指定 XML 数据作为一个字段值入库，自动抽取文本进行索引时，过滤器对 CDATA 标记部分的处理方式的例外情况。对于大多数情况，都可以使用变量 XML_FILTER_CDATA 来进行控制，但有时在同一个 XML 文件中，某些 CDATA 中的标记和内容需要全部保留，与 XML_FILTER_CDATA 指定的方式相冲突，这时需要使用 XML_KEEP_CDATA 变量来指定需要例外处理的 CDATA 的标记和内容。

变量 XML_KEEP_CDATA 的值是一个或多个标记名，如果是多个标记，则用单字节的分号 ‘;’ 或逗号 ‘,’ 分隔，例如：XML_KEEP_CDATA=Math-img,Math-pic。这些标记必须是 XML 格式文件中的合法标记，表示当这些标记出现在 CDATA 中时，该标记以及它包含的子标记和内容，都将以文本方式输出，但去掉所有回车换行符和多余的空白符，把全部的内容及标记都看作处在同一文本行进行索引；检索命中点的输出方式与取值为 FALSE 时相同。除此之外的标记和内容，仍按 XML_FILTER_CDATA 指定的方式处理，即使是在同一个 CDATA 下。下面是一个例子：

```
<?xml version="1.0" encoding="UTF-8" ?>
<?trs-parser SegmentMark="sipoParagraphTag" NewLineMark="br"
HitShowMark="FONT" KeepCDATA="Math-img" FilterCDATA="any"?>
<r>
<sipoParagraphTag num="1">
<tag><![CDATA[ - /<Math-img><math-img id="math62" file="p62"
img-content="drawing" img-format="PNG" /></Math-img>" :]]></tag>
</sipoParagraphTag>

<sipoParagraphTag num="2">
<tag><![CDATA[ - ADVANTAGE :]]></tag>
<advantageAscii xmlns="http://schemas.thomson.com/ts/20041221/tsiz">
<paraAscii>Worktables can be combined into units in a versatile manner. The
units thus produced are simply and quickly modifiable. The system can be
produced cost-effectively and with low material consumption.</paraAscii>
</advantageAscii>
</sipoParagraphTag>
```

过滤后建立索引的文本为：

```
<Math-img><math-img id="math62" file="p62" img-content="drawing"
img-format="PNG" /></Math-img>
```

Worktables can be combined into units in a versatile manner. The units thus produced are simply and quickly modifiable. The system can be produced cost-effectively and with low material consumption.

该变量没有缺省值，它控制的是过滤器输出的文本，而不是浏览时输出的 XML 文件。【修订 154】

15.3 控制文件实例

下面是一个控制文件的例子：

```
#This is the control file of loading News database.  
RECORD_FILE_FORMAT = STANDARD  
BIT_FORMAT_DEFAULT = BIT  
HTML_BASE_DEFAULT =  
DATA_PATH_DEFAULT =  
DATE_CENTURY_DEFAULT =  
IGNORE_RECORD_CRYPT = TRUE  
FILE_SUFFIX_PRIOR = TRUE  
SYNC_CREATE_INDEX = TRUE  
CLASS_BIT_STORE = TRUE  
CLASS_DOCUMENT_STORE = TRUE  
WAIVE_DOCUMENT_REFERER = TRUE  
WAIVE_OEMFILTER_TEXT = TRUE  
KEEP_TEXT_ACTUAL = TRUE  
CHECK_TEXT_ITERATIVE =  
MULTI_VALUE_SEPARATER = ;  
SECTION_TAG_PREFIX = $$NONDATA$$  
START_RECORD_FROM = 1  
PROCESS_RECORD_NUMBER =  
MAX_COMMIT_NUMBER = 4000  
MAX_ERROR_NUMBER =  
COLUMN_NAME_SEQUENCE = 日期;版次;版名;标题;作者;正文;图像  
XML_SEGMENT_MARK =  
XML_SENTENCE_MARK =
```

XML_NEWLINE_MARK =
XML_HITSHOW_MARK =
XML_FILTER_CDATA =
XML_KEEP_CDATA =
XML_OVERWRITE_PI =
COLUMN_AUTO_MATCH = TRUE
NAMIZE_ISO2709_FIELD =
THROUGHT_ACCESS =

第四部分 服务器安装

本部分介绍了 TRS 服务器在安装前的准备工作、具体的安装步骤，以及安装后的工作。

读者对象

本部分的使用对象是 TRS 系统安装人员。在阅读之前，建议您先阅读《数据库管理》部分，以便了解一些基本概念。

内容组织

本部分按如下结构组织：

- 第 16 章 在 Windows 下安装
- 第 17 章 在 UNIX 下安装
- 第 18 章 数据库镜像
- 第 19 章 服务器工具
- 第 20 章 配置文件

第 16 章 在 Windows 下安装

16.1 目录结构

图 16-1 是安装好以后 TRS 服务器在操作系统中的目录结构图。

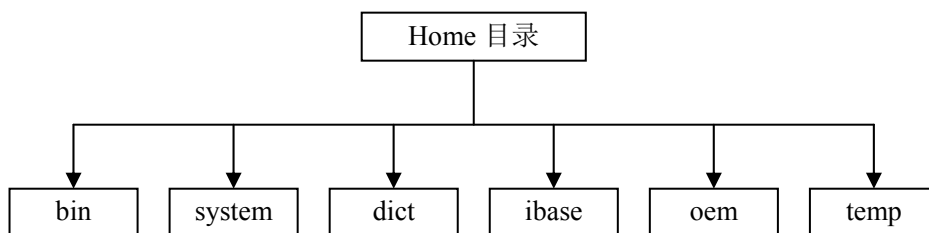


图 16-1

16.1.1 bin 目录

bin 目录存放的是服务器的所有可执行程序 and 动态链接库，以及版本信息文件，包括：

- ✧ readme.txt 这是一个文本文件，用来说明当前服务器版本所做的改进。
- ✧ trsidenter.exe 获得服务器的机器码，用于生成运行时所必须的注册码。
- ✧ trsserver.exe 启动服务器的程序。
- ✧ trssearcher.exe 启动服务器检索缓冲模块的程序。
- ✧ trsmirror.exe 启动服务器数据镜像模块的程序。
- ✧ tdbservice.exe TRS 服务器作为一个 NT 服务时的命令行管理程序。
- ✧ tdbconsole.exe TRS 服务器作为一个 NT 服务时的图形界面控制台程序。
- ✧ tdbshutdown.exe 关闭服务器的程序。
- ✧ tdbdict.exe 诊断与修复服务器的数据词典的工具。
- ✧ tdbmaster.exe 服务器管理员。
- ✧ tdbload.exe 数据库加载工具。
- ✧ tdbdaemon.dll 连接响应与服务调度的动态链接库。
- ✧ tdbshell.dll 服务器数据词典管理与服务请求响应的动态链接库。
- ✧ tdbsearch.dll 数据库检索动态链接库。
- ✧ tdbkernel.dll 数据库维护管理动态链接库。
- ✧ tdbcmkplo.dll 数据库检索 CKM 辅助功能动态链接库。（可选）
- ✧ trskit32.dll 基本通用接口动态链接库。

- ◇ trsapi32.dll 客户端访问服务器的接口动态链接库。
- ◇ trsrpc32.dll 远过程调用动态链接库。

16.1.2 system 目录

system 目录存放服务器的数据词典、配置文件以及 KEY.INI 文件。该目录下的文件是服务器的核心文件，不能破坏，否则服务器将无法启动和运行。除配置文件以及 KEY.INI 文件外，其他文件均由服务器自动维护。

在 system 目录下有四个配置文件，分别是：运行日志配置文件 LOG.INI、数据镜像模块配置文件 MIRROR.INI、插件配置文件 OEM.INI，以及运行参数配置文件 PROFILE.INI。其详细说明请参见配置文件。

KEY.INI 文件记载了服务器运行时所需要的一系列键值，它是一个文本文件，每一行一个键值，其形式是：“键名 = 键值”，除个别外，其中的键值都可以进行修改。

- ◇ PRODUCT = 1.IBQI4CBD-GHA4CD21-PTKLPHSD
键 PRODUCT 的值是 TRS 信息技术股份有限公司提供给用户的产品注册码，用户不能随意改动。如果服务器由于该键值不正确而不能正常启动或运行时，请与 TRS 信息技术股份有限公司的技术支持联系。
- ◇ TOKEN = 40000
键 TOKEN 的值是服务器的动态信号量系统的初始信号量的键值，同时 TOKEN+1 (40001)，TOKEN+2 (40002) 也将被占用，必须保留。在确保唯一性前提下，该键值可以被修改。
- ◇ CONFIG = 10000
键 CONFIG 的值是服务器存储系统配置数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。
- ◇ GROUP = 10001
键 GROUP 的值是服务器存储用户组数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。
- ◇ USER = 10002
键 USER 的值是服务器存储用户数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。
- ◇ BASE = 10003
键 BASE 的值是服务器存储数据库数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。
- ◇ BASECOL = 10004
键 BASECOL 的值是服务器存储数据库字段数据词典的共享内存映射文件的

键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。

✧ VIEW = 10005

键 VIEW 的值是服务器存储视图数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。

✧ VIEWCOL = 10006

键 VIEWCOL 的值是服务器存储视图字段数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。

✧ RIGHT = 10007

键 RIGHT 的值是服务器存储权限管理数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。

✧ FORMAT = 10008

键 FORMAT 的值是服务器存储数据格式数据词典的共享内存映射文件的键值。这个键值不能与以上各键值冲突。在确保唯一性前提下，该键值可以被修改。

✧ INSTANCE = 10009

键 INSTANCE 的值是服务器存储用户登录信息的共享内存映射文件的键值。这个键值不能与以上各键值冲突，同时 INSTANCE+1 (10010) 也将被占用，必须保留。在确保唯一性前提下，该键值可以被修改。

在 system 目录里，除了存放的一些系统文件外，还包含三个子目录：etc、inherit 和 shared，这三个子目录里的文件由系统自己维护，在非特殊情况下，外部不需要也不能进行修改或删除，否则系统将有可能被破坏，不能正常运行。

16.1.3 dict 目录

dict 目录存放的是服务器的各种知识词典，包括分词词典、附加分词词典、停用词典、附加停用词典、主题词典、同义词典和反义词典等。TRS 全文数据库系统允许用户自己创建这些知识词典，在没有指定词典存放目录的前提下，这些词典也存放在该目录。

16.1.4 ibase 目录

ibase 目录是服务器存放数据库的缺省目录，包括数据库的全部数据和索引。只要磁盘空间允许，用户创建的任何数据库都可以存放在该目录下。TRS 全文数据库系统允许用户设置并使用其它的目录来存储数据库。

16.1.5 oem目录

oem 目录将主要用来存放为 TRS 全文数据库系统配置的 OEM 产品。

16.1.6 temp目录

temp 目录是 TRS 全文数据库系统的临时文件公用区。服务器在运行时要产生许多临时文件，如操作日志文件、检索结果历史文件、输出检索结果记录时产生的临时文件、数据库加载时网络传输产生的临时文件等等，这些文件都存储在该目录中。有时还要在该目录下自动细分更多的子目录，特别地，LOGS 子目录是用来存放客户端请求的日志文件^{【修订 110】}以及系统启动、关闭操作的日志文件（system.log）的。^{【9】}

16.2 准备工作

16.2.1 软件及硬件环境

服务器的运行需要一定的软硬件环境，下面所列的各项是最基本的要求，在安装之前请确认一下您的机器是否满足这些条件：

- 运行中文版 Microsoft Windows NT Server 4.0(Service Pack3)以上版本操作系统的小型机、服务器、工作站、Pentium 及其兼容机。
- 内存^{【修订 006】} 1GB 以上，建议使用 4GB 内存。
- 一个可用空间在 1GB 以上的独立硬盘或分区用于安装服务器。
- 操作系统可使用的对换空间最好在 1GB^{【9】} 以上。
- 局域网连接设备。

如果你要求 TRS 服务器支持大文件（即 largefile，有的系统称超过 2G 的文件为大文件，而有的系统则认为超过 1G 的文件就是大文件），则必须使用支持大文件的安装包进行安装，并选择合适的文件系统，如 FAT32 支持的文件最大只能是 4G。

还必须注意：虽然在所有支持 Microsoft Win32 系统的操作系统平台上（如 Microsoft Windows 95/98/NT）都能安装 TRS 服务器，但只推荐使用 Windows NT 以及后续版本（如 Windows 2000）的系统，因为其他版本的 Windows 系统不能保证与 Windows NT 完全兼容。

16.2.2 确认安装方式

TRS 服务器的安装有三种方式：

- 从硬盘上安装。
- 从光盘上安装。这种方式要求机器具有光盘驱动器，并安装了该设备。
- 从网络上安装。这种方式需要先将机器连接到存放安装包的机器上，并保证能访问到安装包，或者将安装包从网络上复制下载到机器的本地硬盘上，然后再从硬盘上安装。

此外，安装包也有两种形式：使用 InstallShield 制作的安装包，以及使用命令行的安装包。这两种安装包的安装程序是不同的，前者的安装程序是 InstallShield 提供的通用程序 `setup.exe`，后者则是 TRS 的专用程序 `tdbinstall.exe`。

16.3 安装步骤

16.3.1 向导式安装

(1) 进入安装包所在目录，并运行 ‘`setup.exe`’ 程序，这时显示如图 16-2 所示的界面。

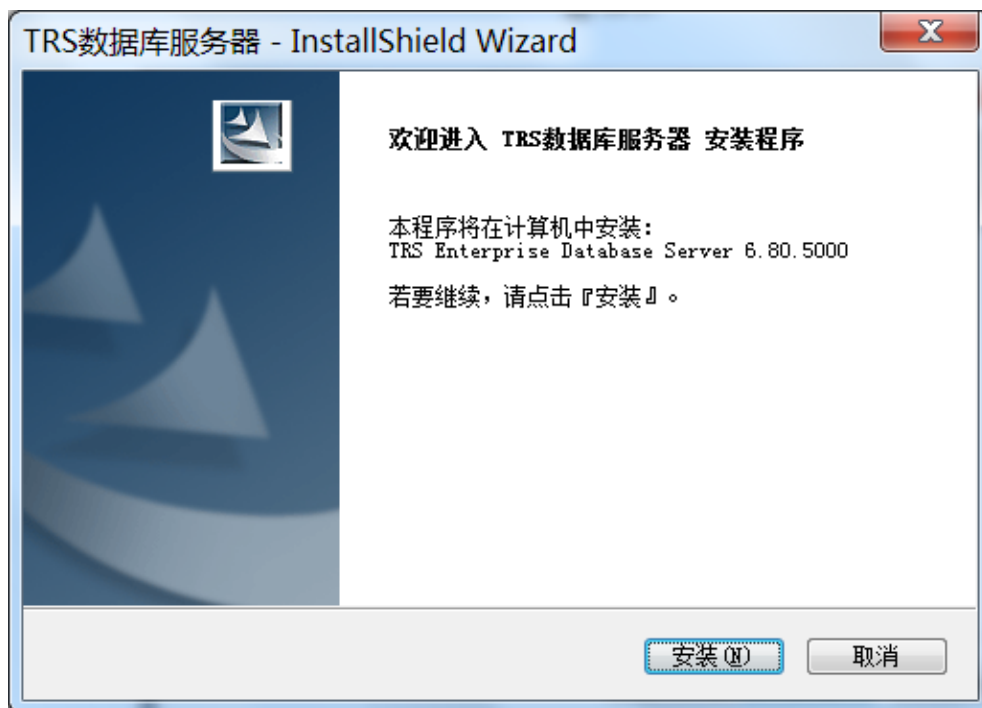


图 16-2

(2) 点击『安装』按钮，显示如图 16-3 所示的对话框。

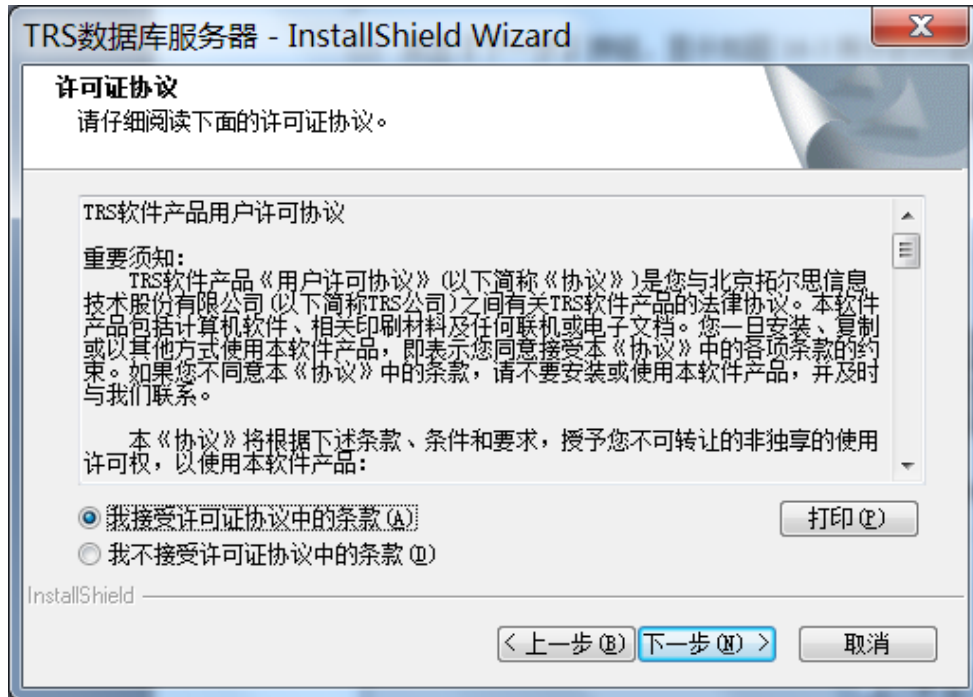


图 16-3

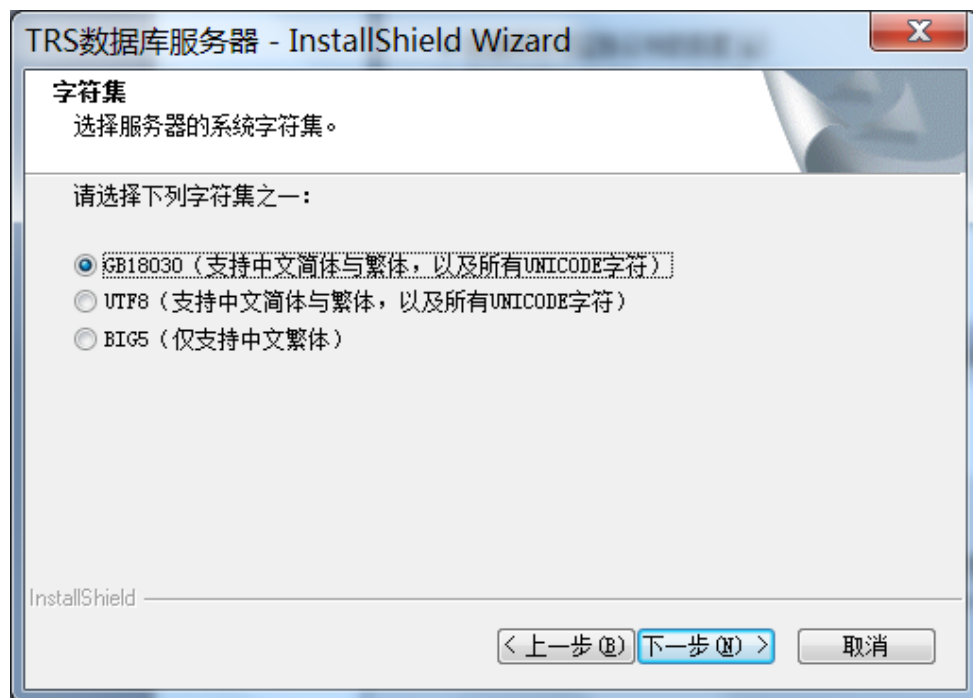


图 16-4

- (3) 如果选择“我接受许可协议中的条款”，并点击『下一步』按钮，显示如图 16-4 所示的对话框。
- (4) 为系统选择合适的字符集。一般情况选择 GB18030 就可以了，GB18030 和 UTF8 能够支持所有 UNICODE 所收录的字符。点击『下一步』按钮，显示如图 16-5 所示的对话框。

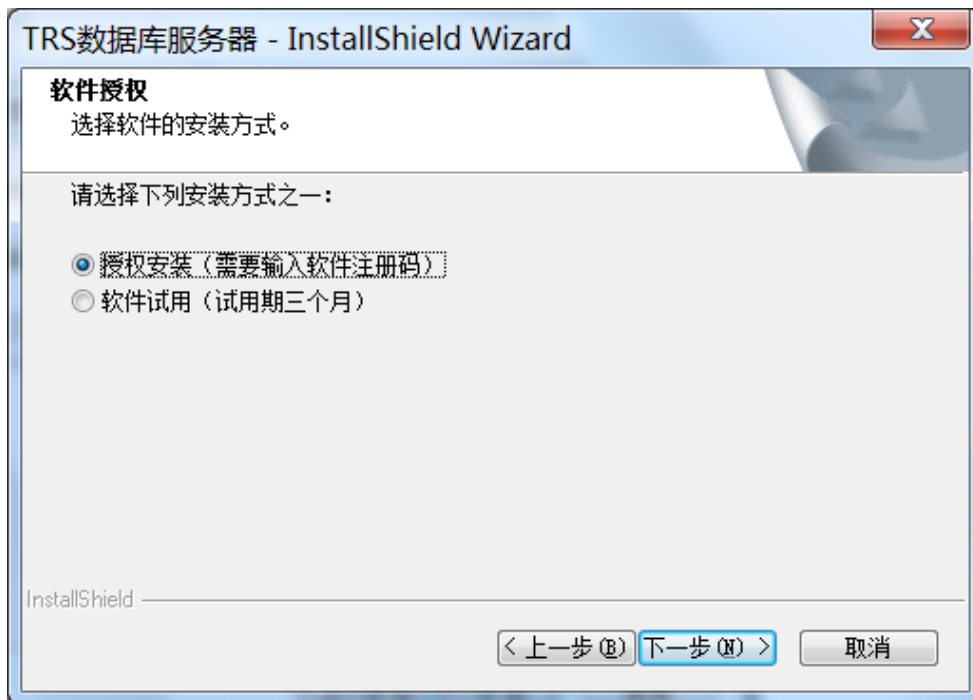


图 16-5

- (5) 选择安装的版本类型。如果您已购买产品的 LICENCE，则可以进行“授权安装”正式版，正式版需要从厂商申请产品注册码。如果您选择“软件试用”安装试用版，则系统只有三个月的有效期，届时您必须向厂商申请延长试用期，或者购买正式的产品 LICENCE，否则系统不能工作。选择“正式版”，并点击『下一步』按钮后会显示如图 16-6 所示的对话框；选择“试用版”，并点击『下一步』按钮后会显示如图 16-7 所示的对话框。
- (6) 如果是安装正式版，则请在图 16-6 所示对话框的[注册码]中输入产品注册码，产品注册码是由厂家根据该对话框中显示的机器码生成的，所以这时您需要与厂家联系，并告知您的机器码。因字体原因，所显示的数字 0 和字母 O 可能会难以区分（但在命令行窗口下的差别如图 16-6 所示一样明显），建议您直接从对话框中复制出来，并发送给负责生成注册码的人员，或者在厂家提供的网站上生成入产品注册码。如果是安装试用版，则图 16-7 所示的对话框不需要输入产品注册码。



图 16-6

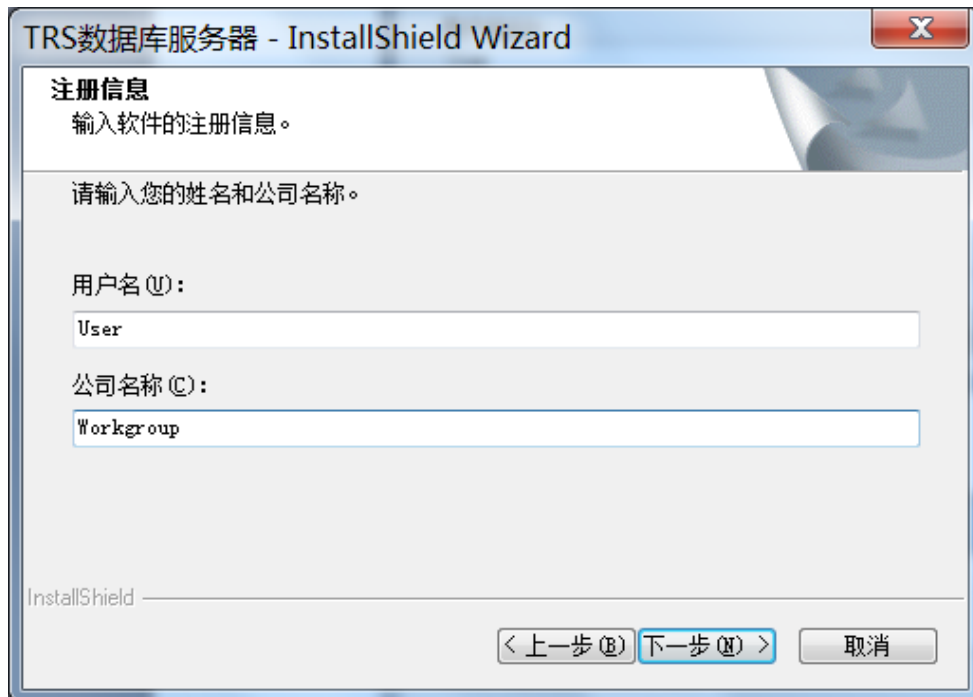


图 16-7

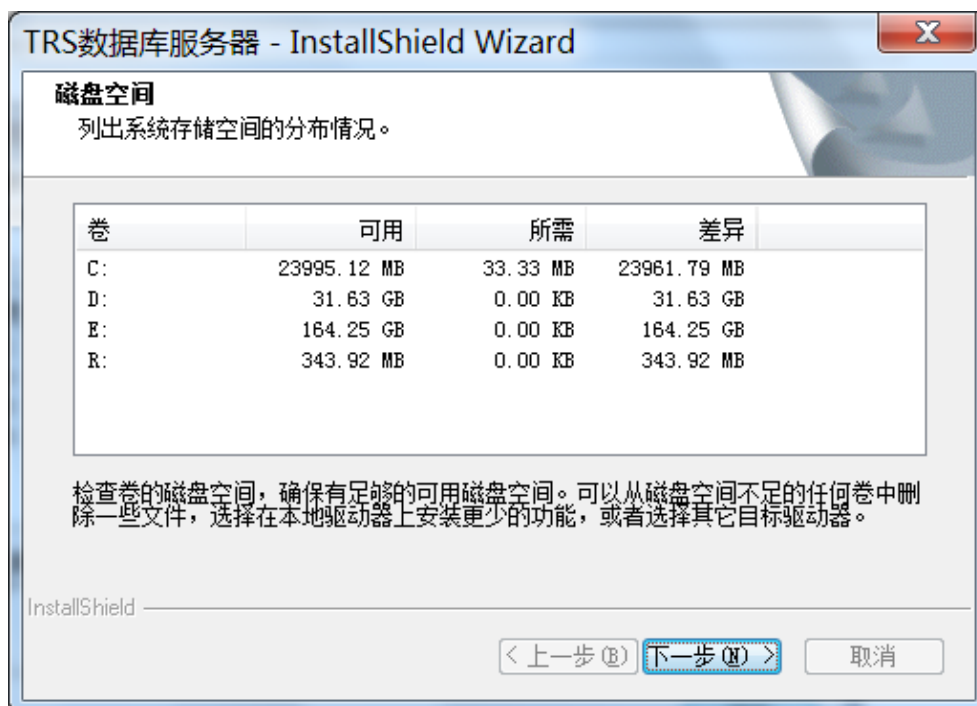


图 16-8

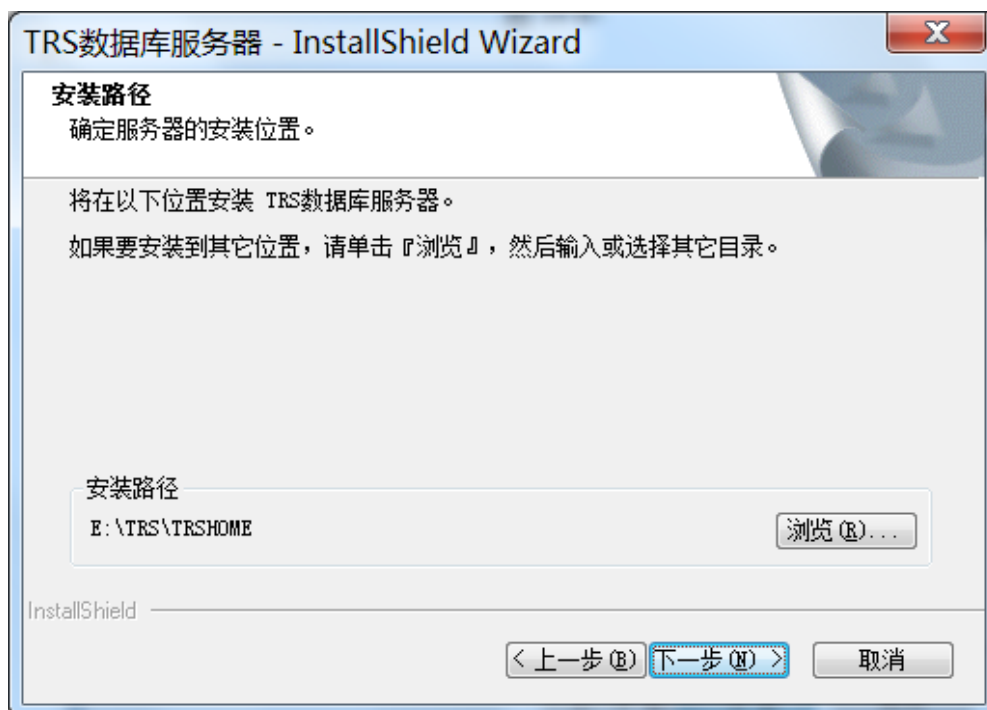


图 16-9

- (7) 在图 16-6 或 16-7 所示的对话框里点击『下一步』按钮后，显示如图 16-8 所示的对话框。该对话框列出了您机器中所有分区的剩余空间，以便您确定存放服务器的目标文件夹的位置。点击『下一步』按钮后，显示如图 16-9 所示的对话框。
- (8) 在图 16-9 所示的对话框里，首先给出了一个缺省的安装路径。如果需要改变目标文件夹位置或名称，则可通过点击『浏览』功能进行选择或编辑。^{【修订 106】}如果指定的安装目录已存在，并且其中或其任一子目录里存在文件，则将显示如图 16-10 所示的对话框，要您选择一种处理方式：安装时要删除所有文件；或者安装时只覆盖同名文件。如果目标文件夹存在，但其中或其所有子目录里都只包含有空的子目录，而不包含任何文件，则会事先自动清空目标文件夹，以保持一个“干净”的环境。^{【1】}

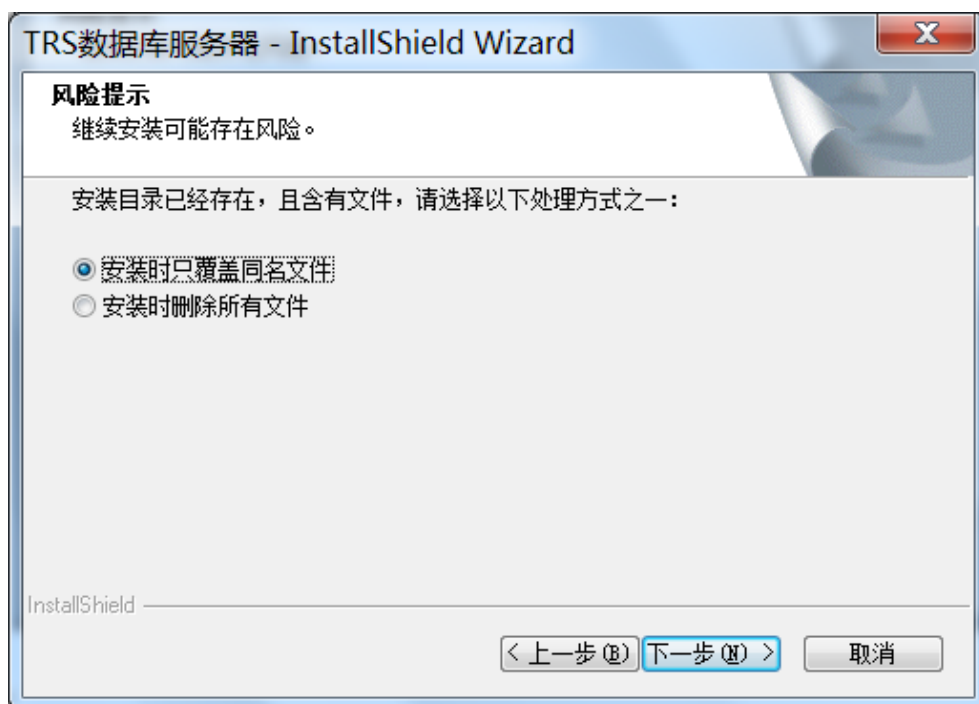


图 16-10

- (9) 确定好安装服务器的目标文件夹后，在图 16-9 或 16-10（目标文件夹里已存在文件的情况）所示的对话框里点击『下一步』按钮将显示如图 16-11 所示的对话框。该对话框中给出了是否需要安装示例数据库的选择。如果选择“是”，则安装程序会在系统里创建三个示例数据库：DEMO1、DEMO2 和 DEMO3。安装示例数据库的好处是完成安装后马上就能检验系统是否能够正常运行。
- (10) 在选择“是否安装示例数据库”后，点击『下一步』按钮将显示如图 16-12 所示的对话框，以在启动菜单里设置服务器的“程序文件夹”。

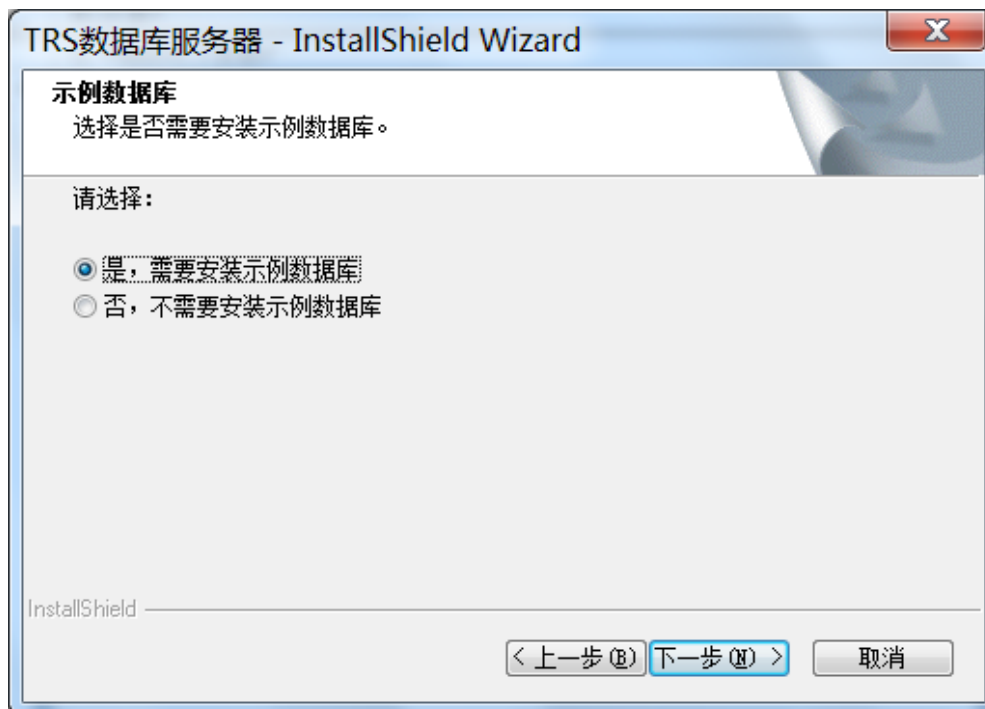


图 16-11

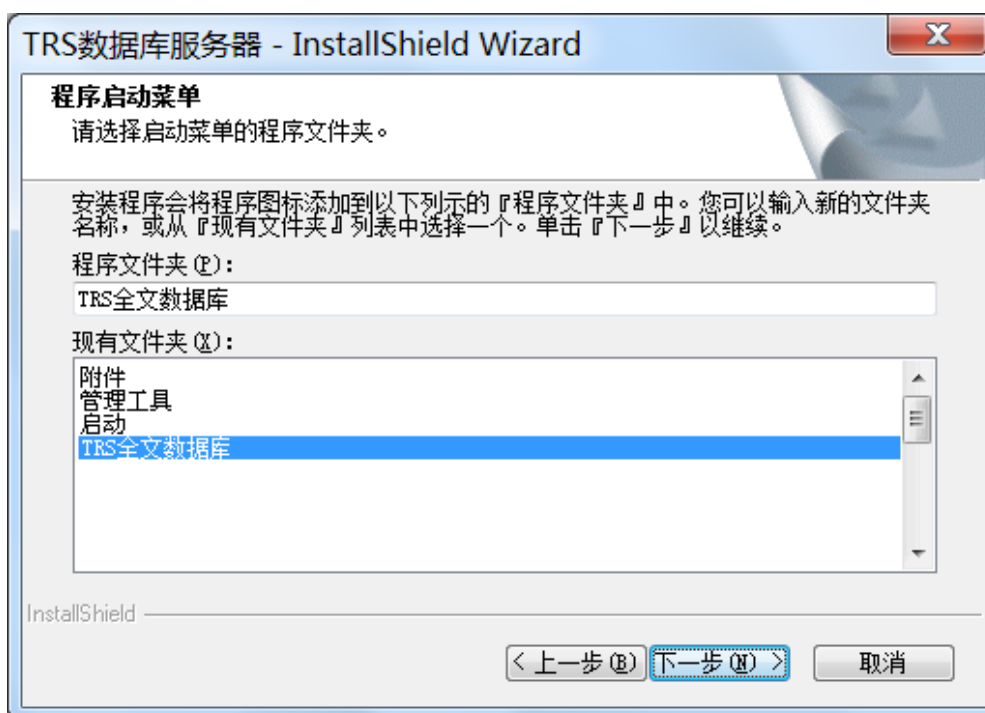


图 16-12

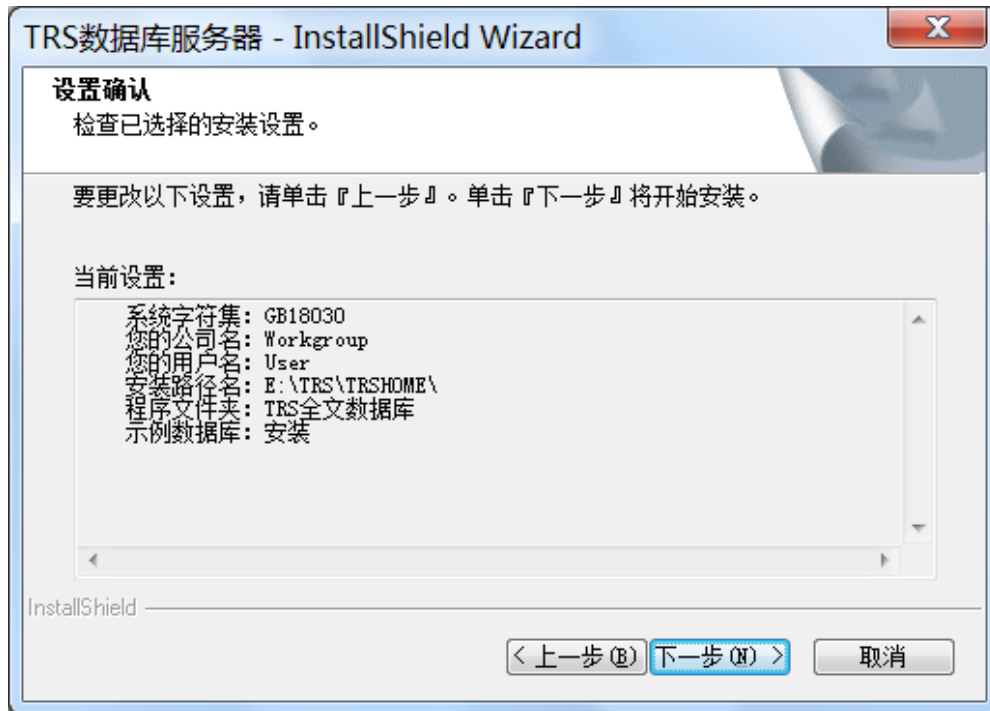


图 16-13

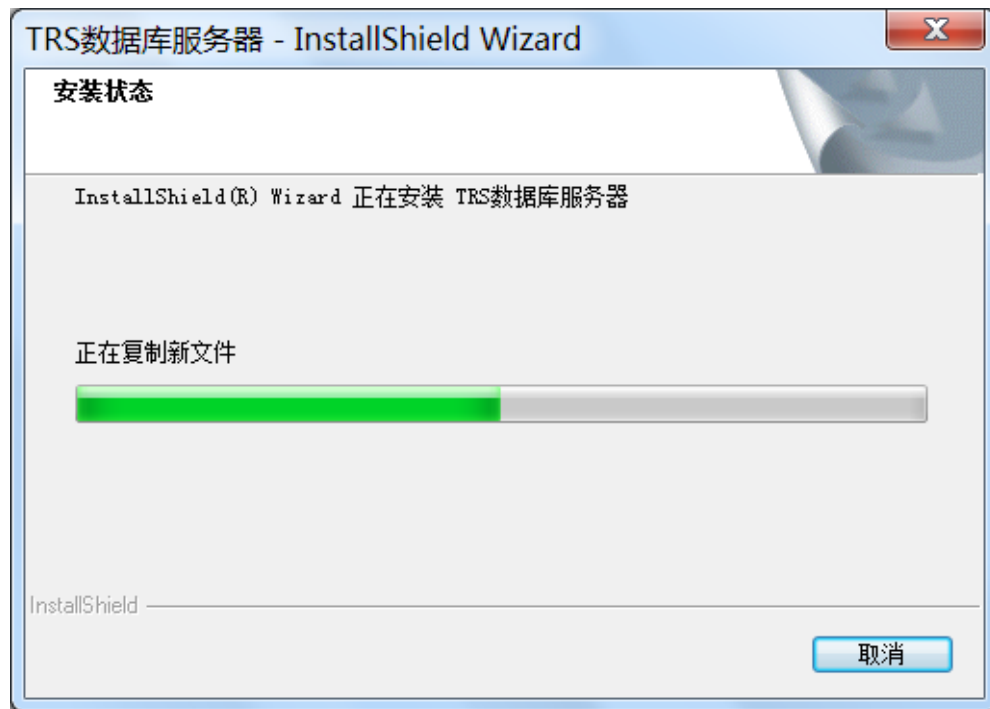


图 16-14

- (11) 确定“程序文件夹”后点击『下一步』按钮将显示如图 16-13 所示的对话框。该对话框表示安装服务器的一些基本选择已经完成。点击『下一步』按钮就开始正式安装，并显示如图 16-14 和 16-15 所示的对话框，报告安装状态和过程。

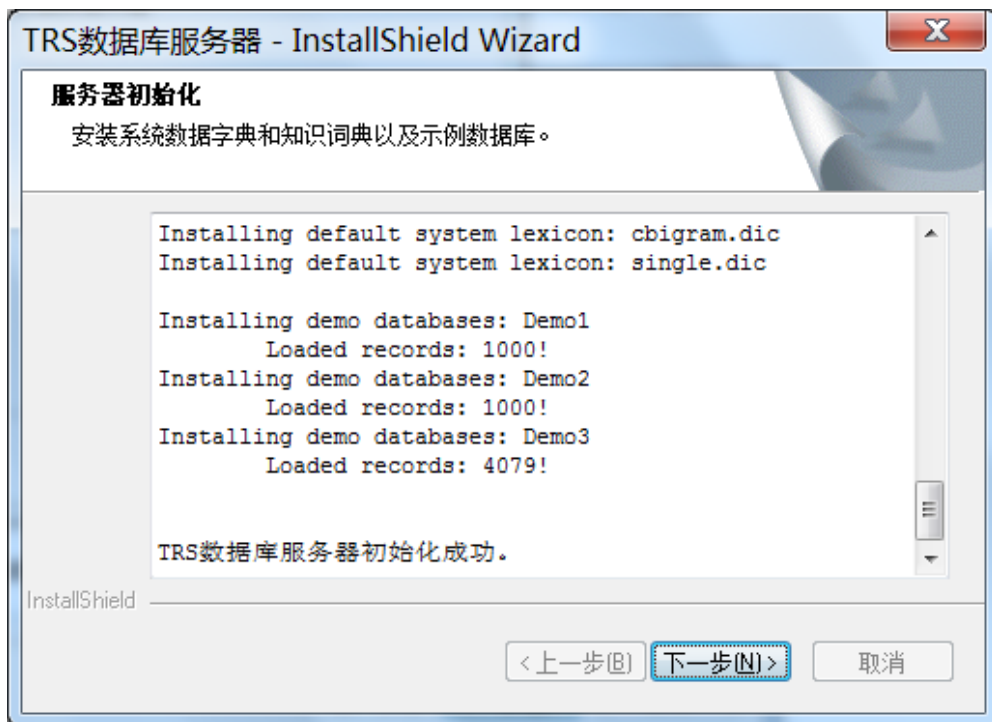


图 16-15

- (12) 到这一步，服务器的基本系统已经安装完成，在图 16-15 所示的对话框里点击『下一步』按钮，将显示如图 16-16 所示的对话框，以完成将 TRS 服务器作为 Windows 服务器进行注册的步骤。
- (13) 如果您在图 16-16 所示的对话框里选择“否，不将 TRS 服务器注册为 Windows NT 服务”，则点击『下一步』按钮后将显示如图 16-17 所示的对话框，以便您为 TRS 服务器设定一个缺省的网络端口号。TRS 服务器默认的网络端口名为“trserver”，如果这个端口名在 Windows 的 services 系统文件里已经存在的话，则不能再进行设置缺省的网络端口号，此时将给出如图 16-18 所示的提示。在图 16-17 或 16-18 所示的对话框里点击『下一步』按钮，安装过程将跳到第(18)步。
- (14) 如果您在图 16-16 所示的对话框里选择“是，将 TRS 服务器注册为 Windows NT 服务”，则点击『下一步』按钮后将显示如图 16-19 所示的对话框。与 16-17 对话框不同的是，16-19 对话框不仅需要输入 TRS 服务器的网络端口号，而且还要为 TRS 服务指定一个名字，该名字将是 TRS 服务在 Windows 系统的“服务管理器”里的显示名。

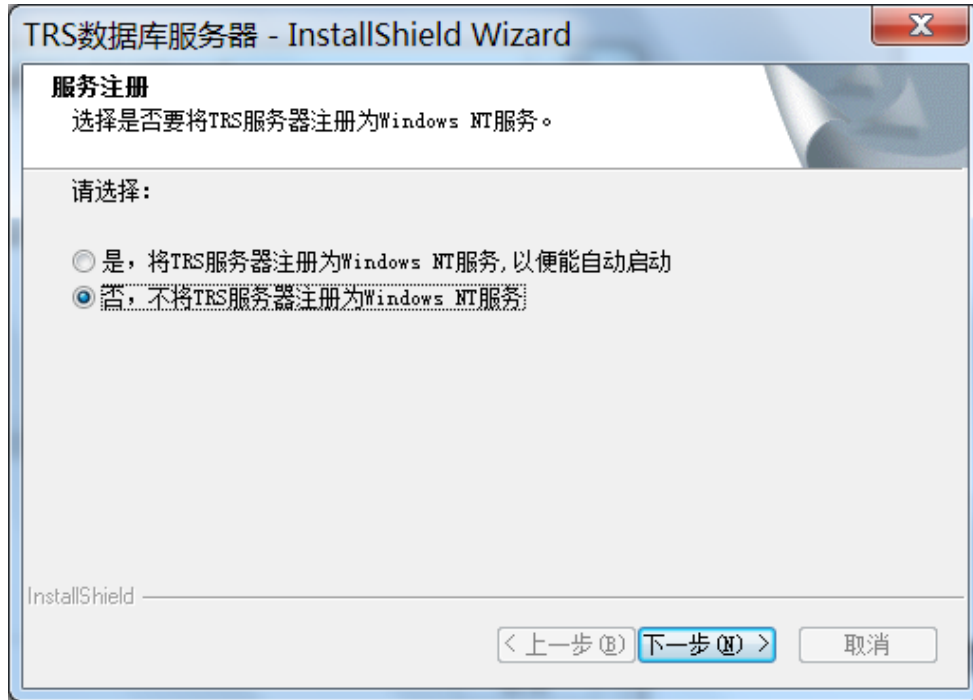


图 16-16

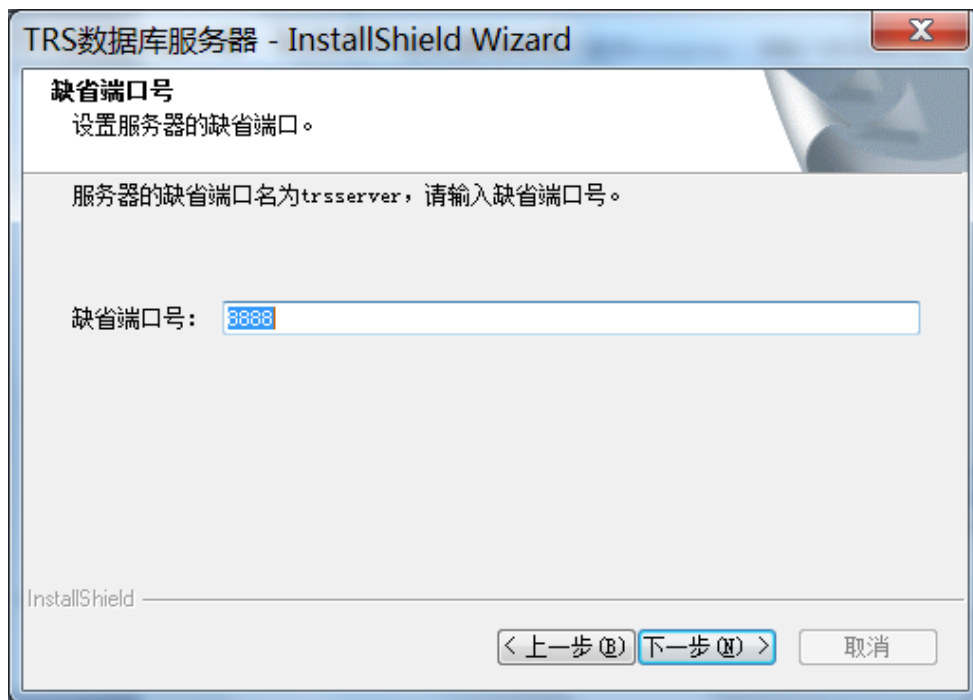


图 16-17

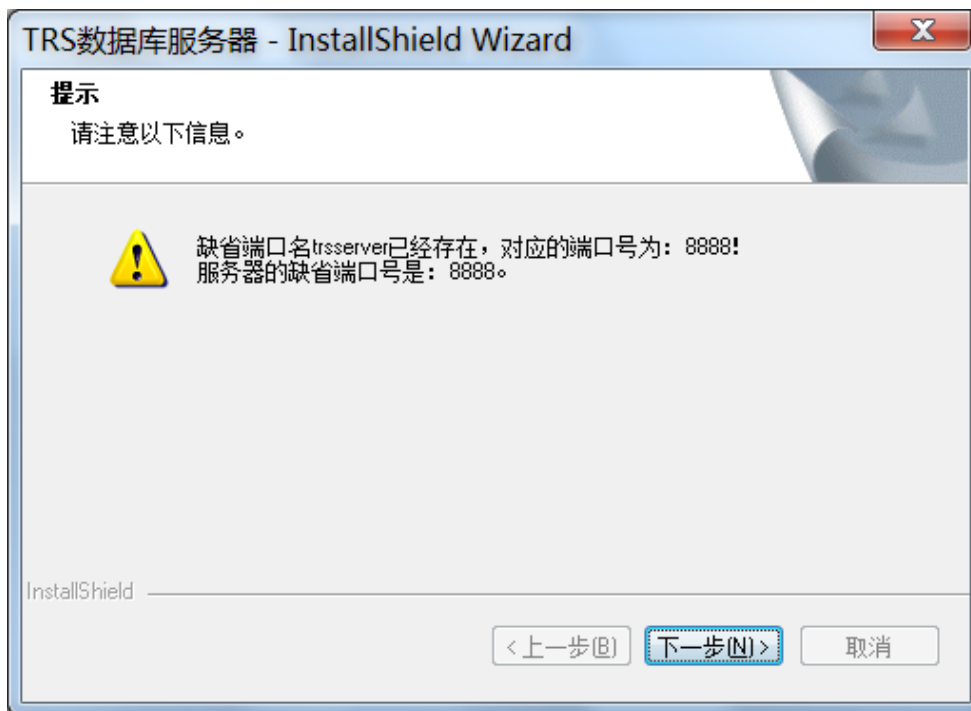


图 16-18

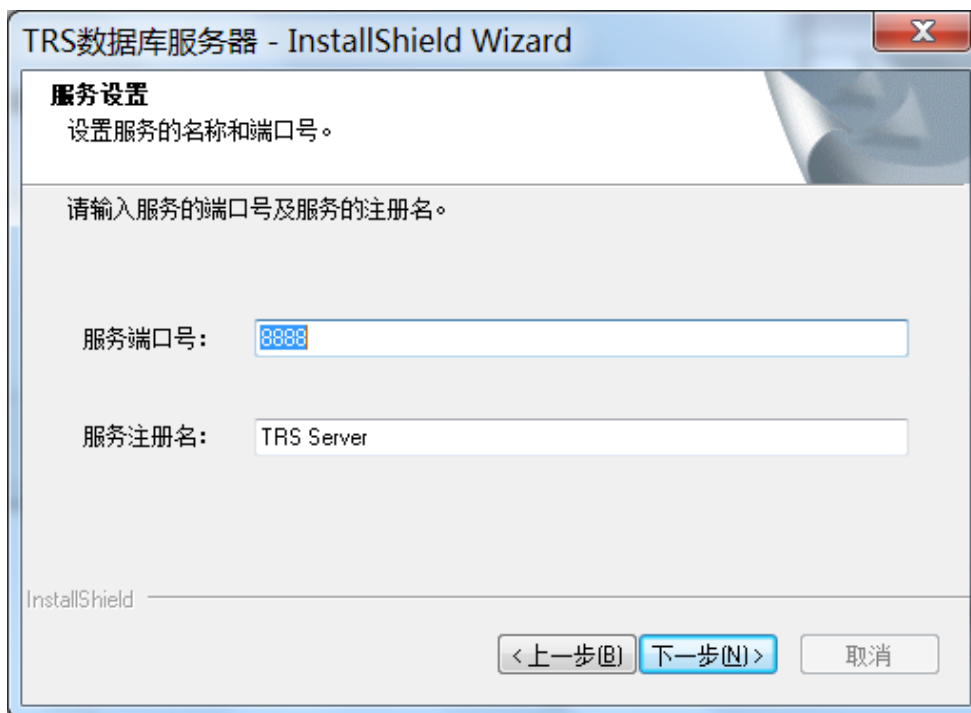


图 16-19

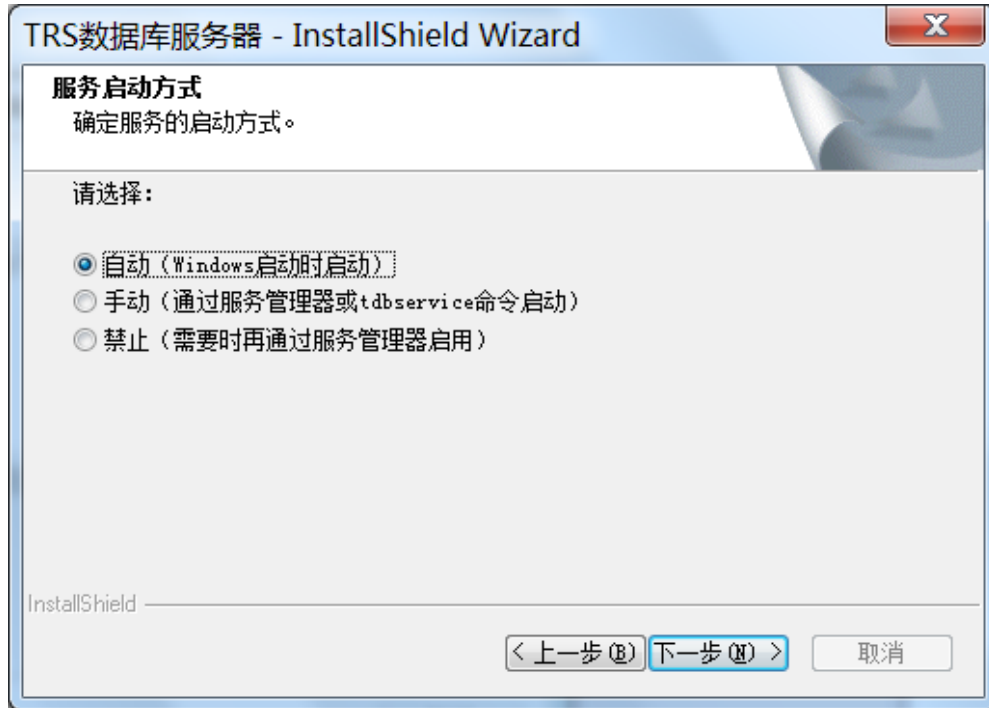


图 16-20

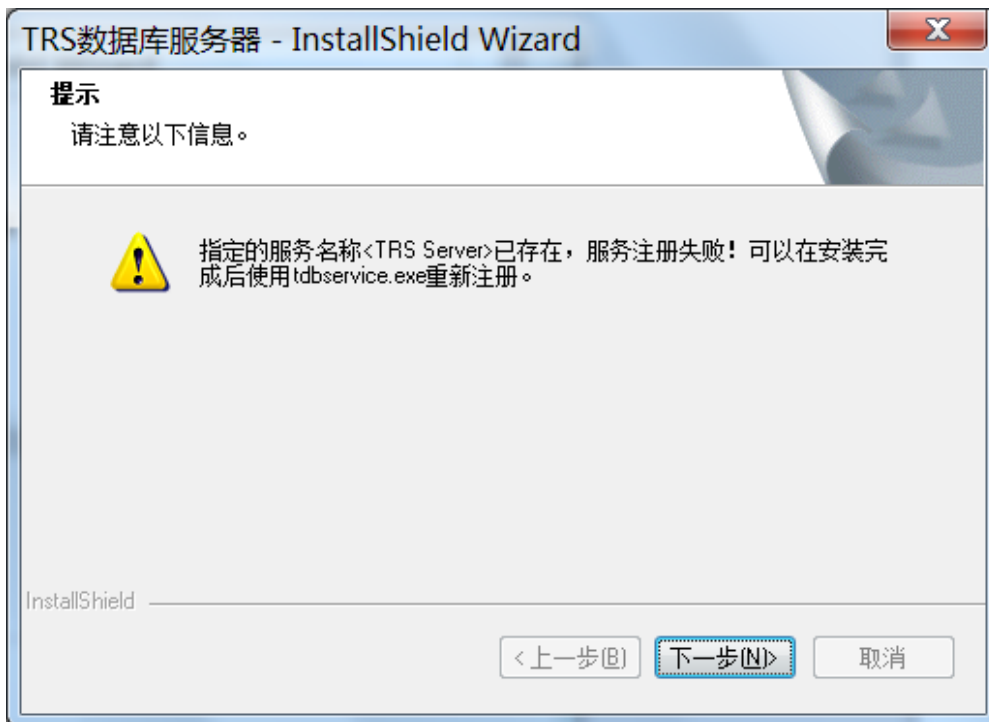


图 16-21

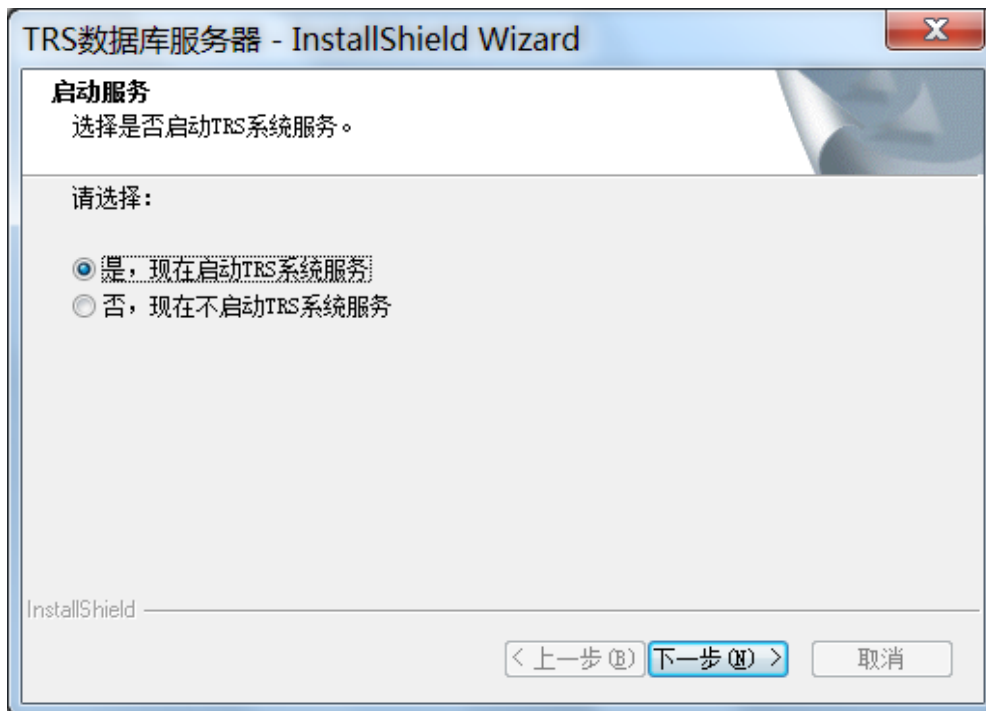


图 16-22

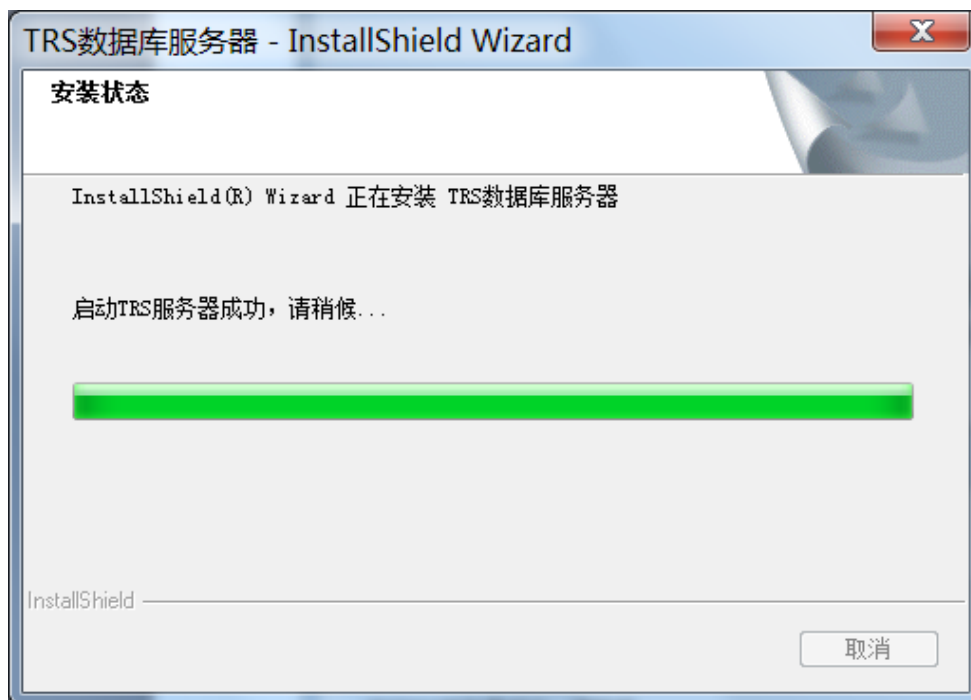


图 16-23

- (15) 在图 16-19 对话框里输入服务器的网络端口号以及服务名称后，点击『下一步』按钮，则安装程序显示如图 16-20 所示的对话框，当您选择了一种合适的启动方式并点击『下一步』按钮后，安装程序将试图将 TRS 服务器注册为 Windows NT 服务。
- (16) 如果 16-19 对话框里指定“服务名称”在 Windows 系统里已经存在，则不能再次以注册同名的服务，以保证不覆盖以前的服务注册信息。此时显示如图 16-21 所示的提示信息，点击『下一步』按钮安装过程将跳到第(18)步。
- (17) 如果服务注册成功，则显示如图 16-22 所示的对话框，询问您是否现在启动刚注册的服务。如果您选择现在就启动服务，则安装程序显示如图 16-23 所示的对话框显示服务的启动状态。启动成功后自动进入下一步。

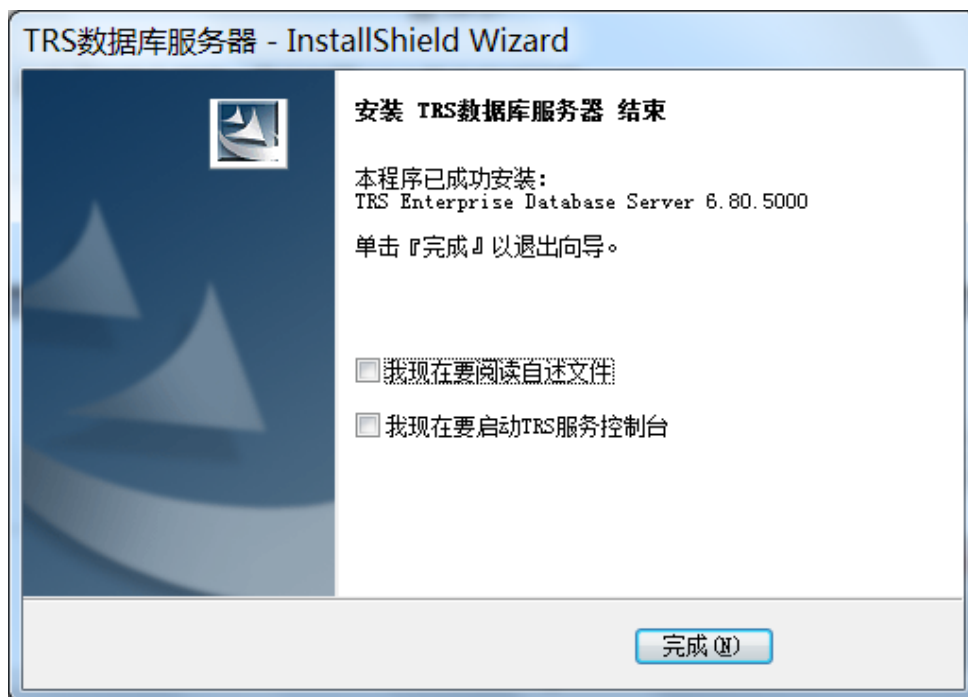


图 16-24

- (18) 至此，安装显示如图 16-24 所示的对话框，表示安装过程已全部结束。

16.3.2 命令行安装

- (1) 进入命令行方式安装包所在目录，运行安装程序 `tdbinstall`。安装程序将询问你是否安装授权版本：“Would you install TRS Server with KEY authorization? (Y/N)”。回答“Y”表示安装授权（正式）版本，否则表示安装试用版本。
- (2) 如果安装试用版本，则跳过这一步，否则安装程序将提示并等待你输入产品注册

码，即“Input PRODUCT KEY”。产品注册码是由厂家根据机器码生成的，所以您需要与厂家联系，并将您的机器码告诉我们。您的机器码已经在屏幕上打印出来：“Your machine code is xxxxxxxx zzzzzzzz”，其中“xxxxxxx”和“zzzzzzz”所对应的 8 个字母数字串就是机器码，其中可能没有后面的“zzzzzzz”，而只有前面的“xxxxxxx”。也可以通过运行安装包目录下的程序‘trsidenter’得到同样的机器码。

- (3) 输入服务器的安装目录名，即“Input a target dir for TRS Server(e.g.: d:\trshome):”，如：‘d:\trshome’，则服务器将安装在‘d:’盘的‘trshome’目录下。注意：如果目录名中有空格，则需用双引号将其括起来。

- (4) 如果输入的安装目录不存在，或不包含任何文件与子目录，则跳过这一步。

如果目录已存在，^[修订 107]且含有文件，则终止安装。

如果目录已存在，且含有空的子目录，但不含有任何文件，则安装程序提示你确定是否需要清空（删除并重建）这个目录并继续安装，即：

WARNING: The inputted dir (d:\trshome) already existed!

Do you want to clear this dir, and continue? (Y/N) ^[4]

回答“Y”表示将自动清空该目录，否则结束安装过程。**特别提醒：“清空该目录”的涵义是删除并重建该目录，所有空子目录都不保留。**所以用于安装的安装包不能放在这个目录下。

- (5) 安装程序提示：

Please select the character set for TRS Server.

1: GB18030

2: UTF8

3: BIG5

You select:

如果输入‘1’，表示以 GB18030 编码为主字符集进行安装；如果输入‘2’，表示以 UTF8 编码为主字符集进行安装；如果输入‘3’，表示以 BIG5 编码为主字符集进行安装。

- (6) 输入以上各项后，安装程序进入自动运行状态，加载服务器所必须的各种知识词典和几个示例数据库。

- (7) 知识词典和示例数据库加载完成后，安装程序提示你输入服务器使用的网络端口号，即“Input port number”，如输入“8888”，并“回车”。

- (8) 安装程序提示：

Install TRS Server into Windows service panel?

0: Not Install

1: Automatic Mode

2: Manual Mode

You select:

如果你输入‘0’，则安装程序不将 TRS Server 安装到 Windows 的服务“控制面板”

中（以后还可以通过 `tdbservice` 程序进行安装），TRS 服务器只能在命令行下用 `trsserver` 直接启动，而其它任何方法都不能启动（参见 16.4.1 节的“用命令行直接启动”方式）；如果你输入 ‘1’，则安装程序将 TRS Server 以“自动”模式安装到 Windows 的服务“控制面板”；如果你输入 ‘2’，则安装程序将 TRS Server 以“手工”模式安装到 Windows 的服务“控制面板”。**注意如果以前在系统的服务控制管理器中登记过，并且未删除，则不能重新登记，以保证不覆盖以前的登记信息。**

- (9) 安装成功之后，安装程序将进一步提示你“Would you test file system for large file exceeding 4G? (Y/N)”如果回答“Y”，则安装程序开始测试服务器所在的“文件系统设备”是否支持大文件的读写。
- (10) 结束。

上述安装过程是交互式的。^[83] 097]安装程序 `tdbinstall` 也支持通过命令行的参数方式进行输入。参数与交互式的输入项是一一对应的，其定义如下：

- (1) `KEY:code` 其中 KEY 是参数名，code 是一个实际的 KEY 串。该参数对应于安装步骤 5 和 6 的输入。例如：`KEY:1.IBQI4CBD-GHA4CD21-PTKLPDSD`。如果 code 对应的是“DEMO”，则表示安装试用版。
- (2) `DIR:path` 其中 DIR 是参数名，path 是服务器要安装的目录。该参数对应于安装步骤 3 的输入。例如：`"DIR:d:\server"`。
- (3) `REW:yesorno` 其中 REW 是参数名，yesorno 是对“是否需要覆盖已经存在的目录”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于安装步骤 4 的输入。例如：`REW:YES`。
- (4) `CS:charset` 其中 CS 是参数名，charset 是系统字符集的选择（1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于安装步骤 5 的输入。例如：`CS:1`。
- (5) `PN:port` 其中 PN 是参数名，port 是服务器的缺省端口号。该参数对应于安装步骤 7 的输入。例如：`PN:8888`。
- (6) `SM:mode` 其中 SM 是参数名，mode 是 TRS Server 作为 NT 服务的启动模式（0 表示不安装成 NT 服务；1 表示“自动启动”；2 表示“手工启动”）。该参数对应于安装步骤 8 的输入。例如：`SM:1`。
- (7) `LF:yesorno` 其中 LF 是参数名，yesorno 是对“是否进行大文件支持测试”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于安装步骤 9 的输入。例如：`LF:NO`。

各参数名不区分大小写，参数之间没有顺序限定。参数名与参数值之间必须用冒号 ‘:’ 分隔，参数名与 ‘:’ 之间不能有空格，参数值与 ‘:’ 之间的空格将看成参数值的一部分。**如果参数值包含了数字和字母以外的符号，则必须将参数名、‘:’，以及参数值作为整体用双引号括起来。**如果输入项没有出现对应的参数，则仍然要通过外部输入。^[4]

16.4 运行与维护

16.4.1 启动服务器

- 用命令行直接启动

- (1) 在命令行下进入服务器所在的文件夹，如 ‘d:\trshome’。
- (2) 进入下一级子目录 bin。
- (3) 运行 trsserver。如果这时打印出 “Start up TRS Server successfully!”，则表示服务器已经启动成功；如果这时打印出没有端口号的错误（一般是 17803 号错误），则可以使用 ‘trsserver 端口号’（如 ‘trsserver 8888’）的方式启动服务器，也可以在系统文件 ‘Windows NT 根目录\system32\drivers\etc\services’ 或 Windows 9x 的 ‘Windows 根目录\services’ 中添加 ‘trsserver 端口号/tcp’ 这一行后再运行 trsserver；如果出现其他错误而不能启动服务器，则请参看错误信息表后进行相应处理，或者与相关的技术支持联系。
- (4) 在缺省情况下，trsserver 启动时会自动启动 trssearcher，不需要单独启动。但如果在配置文件 PROFILE.INI 里没有将 trssearcher 设置成自动启动，或其它一些原因导致 trssearcher 未启动，则可以单独运行 trssearcher 启动服务器的检索缓冲模块。与 trsserver 一样，它也可以带一个端口号参数，但必须与 trsserver 启动所使用的端口号相同。在一般情况下，trssearcher 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取 trsserver 启动所使用的端口号，而无须特别指定。
- (5) 运行 trsmirror 启动服务器的数据镜像模块。与 trsserver 一样，它也可以带一个端口号参数，但必须与 trsserver 启动所使用的端口号相同。在一般情况下，trsmirror 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取 trsserver 启动所使用的端口号，而无须特别指定。

- 用命令行启动作为 NT 系统服务的 TRS 服务器

- (1) 在命令行下进入服务器所在的文件夹，如 ‘d:\trshome’。
- (2) 进入下一级子目录 bin。
- (3) 运行 tdbservice -start。

- 用 TRS 控制台启动作为 NT 系统服务的 TRS 服务器

- (1) 启动 TRS 控制台 (tdbconsole)。
- (2) 单击位于屏幕右下角任务图标区的 TRS 控制台图标。
- (3) 在 TRS 控制台的菜单中选取 “启动服务” 功能。

- 在操作系统启动时自动启动作为 NT 系统服务的 TRS 服务器
 - (1) 如果 TRS 服务器还没有登记到 NT 系统服务控制管理器中, 则先进行登记工作 (tdbservice 和 TRS 控制台都能完成这种工作)。
 - (2) 将服务的启动方式设置为“自动”(通过 NT 控制面板或 TRS 控制台可完成设置工作)。
 - (3) 重新启动时, 操作系统将自动启动 TRS 服务器。

需要说明的是: 除“用命令行直接启动”方式外, TRS 服务器作为 NT 服务的三种启动方式实质和效果是完全相同的, 只是所使用的工具不同而已。这三种方式启动必须满足三个前提条件:

1. 操作系统必须是 Windows NT 及后续版本 (Windows 9x 不支持服务);
2. TRS 服务器已经作为 Windows 系统的服务进行了注册 (请参见 16.3.1 节的第 10 步和 16.3.2 节的第 7 步; tdbservice 也能够完成服务的注册, 不带参数运行 tdbservice 时, 程序将打印出它的具体用法)。
3. 在操作系统文件 ‘Windows NT 根目录\system32\drivers\etc\services’ 中已经加入了 ‘trserver xxxx/tcp’ 这个服务端口, 或者安装程序 (或 tdbservice) 在进行服务注册时指定了服务端口。

借助 TRS 全文数据库系统的运行参数配置文件 PROFILE.INI 的 AutoBatch 配置域, 可以将服务器的数据镜像模块 (trsmirror) 和服务器的检索缓冲模块 (trssearcher) 与 trserver 一起自动启动, 请参见配置文件中的详细说明。

特别注意: 如果 TRS 服务器作为 Windows 服务启动时, 若要启动 trssearcher, 则必须对其进行以上的配置, 以便与 TRS 服务器一起自动启动, 而不能用命令行单独手工启动 trssearcher。

16.4.2 关闭服务器

- 用命令行直接关闭
 - (1) 在命令行下进入服务器所在的文件夹, 如 ‘d:\trshome’。
 - (2) 进入下一级子目录 bin。
 - (3) 运行 tdbshutdown。如果这时打印出错误信息, 并提示你指定端口号, 则可以使用 ‘tdbshutdown 端口号’ 的方式关闭服务器, 其实这种方式可在任何情况下使用, 如 ‘tdbshutdown 8888’。在一般情况下, tdbshutdown 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取 trserver 启动所使用的端口号, 而无须特别指定。
 - (4) 如果只需 tdbshutdown 向服务器发送“关闭”命令后会立即结束, 而不管服务器进程是否已经退出, 则可使用 ‘-nowait’ 参数。

- (5) 如果运行 `tdbshutdown` 时带上 `-cache` 选项参数, 则只关闭服务器的检索缓冲模块。如 `'tdbshutdown -cache'`。
 - (6) 如果运行 `tdbshutdown` 时带上 `-reboot` 选项参数, 则表示关闭并重新启动服务器。如 `'tdbshutdown -reboot'`。该选项的功能与“先关闭, 再手工启动”的区别在于系统在“空闲状态起止时间”下的自动维护功能的启动时间, 请参见 5.7 节的描述。^{【修订 151】}
 - (7) 运行 `trsmirror -shutdown`, 关闭服务器的数据镜像模块。它需要与 `tdbshutdown` 完全一致的‘端口号’参数, 如 `'trsmirror -shutdown 8888'`。在一般情况下, `trsmirror` 可以从与程序所在目录平级的 `system` 子目录中的 `START.INF` 文件中自动获取 `trserver` 启动所使用的端口号, 而无须特别指定。
- 用命令行关闭作为 Windows 系统服务的 TRS 服务器
 - (1) 在命令行下进入服务器所在的文件夹, 如 `'d:\trshome'`。
 - (2) 进入下一级子目录 `bin`。
 - (3) 运行 `tdbservice -stop`。
 - 用 TRS 控制台启动作为 Windows 系统服务的 TRS 服务器
 - (1) 启动 TRS 控制台 (`tdbconsole`)。
 - (2) 单击位于屏幕右下角任务图标区的 TRS 控制台图标。
 - (3) 在 TRS 控制台的菜单中选取“关闭服务”功能。

需要说明的是: 除“用命令行直接关闭”方式外, TRS 服务器作为 NT 服务的三种关闭方式实质和效果是完全相同的, 只是所使用的工具不同而已。这三种方式关闭的前提条件是: TRS 服务器是作为 NT 服务启动的 (参见 16.4.1 节)。

借助 TRS 全文数据库系统的运行参数配置文件 `PROFILE.INI` 的 `AutoBatch` 配置域, 可以将服务器的数据镜像模块 (`trsmirror`) 在 `trserver` 关闭时自动关闭, 请参见配置文件中的详细说明。

无论是用哪一种方式关闭服务器, 所使用的端口号必须与服务器启动时的端口号一致。在服务器关闭时并不是立即退出, 它必须等待所有正在执行的服务请求完成后才彻底结束, 但是这时服务器不再会响应新的服务请求了。如果系统中已没有任何名为 `trserver` 的进程时, 表示服务器已正式关闭。

有时候需要强制“杀死” TRS 服务器 (例如在不能正常关闭的情况下), 虽然有时可以使用任务管理器, 或者通过命令 `'tdbshutdown -kill:ProcessID'` 完成, 但是还要手工清除 TRS 服务器所创建的共享内存映射文件等。用命令 `'tdbshutdown -kill 端口号'` 可自动完成所有这些操作。

16.4.3 服务器升级

与命令行方式安装一样，服务器的升级也有命令行方式的标准升级介质包。这个服务器的标准升级包中必须包含有下列文件：readme.txt、trserver.exe、trsearcher.exe、trsmirror.exe、tdbdict.exe、tdbmaster.exe、tdbload.exe、tdbshutdown.exe、tdbservice.exe、trsidenter.exe、tdbdaemon.dll、tdbshell.dll、tdbkernel.dll、trskit32.dll、trsapi32.dll、trsrpc32.dll、tdbupdate.exe、LOG.INI、MIRROR.INI、OEM.INI、PROFILE.INI，以及新词典的资源文件 segment.dic、stop.dic、ambiguous.dic、wbigram.dic、cbigram.dic 和 single.dic，这些文件必须放在同一目录下。tdbconsole.exe 必须按字符集分别放在对应的 GB、BIG5 和 UTF8 子目录下。运行 tdbupdate 即可完成 TRS 服务器系统文件的自动更新。

升级步骤：

- (1) 在进行服务器升级前，必须首先关闭服务器，然后运行 tdbupdate。
- (2) 如果所运行的 tdbupdate 处在服务器的 bin 目录下，则只进行数据词典的更新，而不会进行服务器系统文件的复制。
- (3) 更新过程将有输入提示：“Input the target dir of TRS Server(e.g.: d:\trshome):”，即要求你输入已经安装的服务器所在的绝对路径。请参见命令行安装方式的第 3 步。
- (4) 接下来，升级程序提示：

Please select the character set for added dictionaries:

- 0: NONE
- 1: GB18030
- 2: UTF8
- 3: BIG5

You select:

输入 ‘0’，表示不增加新的词典；输入 ‘1’，表示增加 GB18030 字符集的新词典；输入 ‘2’，表示增加 UTF8 字符集的新词典；输入 ‘3’，表示增加 BIG5 字符集的新词典。

升级程序也支持修改更新现有的词典，要启用这个功能，则必须在命令行中给出参数 “DIC:REW” 或 “DIC:RIX”，参数的意义请参见下面的“参数方式”描述部分。此时：

Please select the character set for modified dictionaries:

- 0: NONE
- 1: GB18030
- 2: UTF8
- 3: BIG5

You select:

输入 ‘0’，表示不修改更新已有词典；输入 ‘1’，表示修改更新 GB18030 字符集的已有词典；输入 ‘2’，表示修改更新 UTF8 字符集的已有词典；输入 ‘3’，表示修改更新 BIG5 字符集的已有词典。

- (5) 如果原系统未设置主字符集（系统字符集），则可能会要求你回答是否需要设置系统的主字符集。

Please select the character set for TRS Server:

0: don't set the native charset.

1: set GB18030 as the native charset.

2: set UTF8 as the native charset.

3: set BIG5 as the native charset.

You select:

如果不设置主字符集，则系统不会自动完成客户端加载和读取数据时的字符集转换，也不能创建 UTF8 字符集的数据库。需要注意的是：你指定要设置系统的主字符集时，则要求当前系统中已有的全部数据库/词典的字符集都必须与所指定的主字符集一致，否则设置无效。

- (6) 如果原系统是 5.0 之前的版本，则要求你回答是否需要复制一套旧名称的程序集，即 “Do you want to keep the old names? (Y/N)”。如果不复制，则有可能影响你在其它地方对旧名称的程序的引用，如：在机器启动时自动启动设置等（不包括 NT 系统中的服务）。

上述升级过程是交互式的。^{【修订 098】}升级程序 `tdbupdate` 也支持通过命令行的参数方式进行输入。参数与交互式的输入项是一一对应的，其定义如下：

- (1) `DIR:path` 其中 `DIR` 是参数名，`path` 是服务器所在的目录。该参数对应于升级步骤 3 的输入。例如：“`DIR:d:\server`”。
- (2) `NEW:charset` 其中 `NEW` 是参数名，`charset` 是需要新增或更新的词典的字符集的选择（0 表示不更新词典；1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于升级步骤 4 的输入。例如：`NEW:1`。
- (3) `DIC:action` 其中 `DIC` 是参数名，`action` 是针对需要更新的词典的动作选择（`REW` 表示只强制更新指定字符集的词典内容；`RIX` 表示强制更新指定字符集的词典内容，且删除相关数据库的索引）。该参数没有对应的升级步骤，不支持外部输入。例如：`DIC:RIX`。
- (4) `CS:charset` 其中 `CS` 是参数名，`charset` 是设置系统字符集的选择（0 表示不设置系统字符集；1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于升级步骤 5 的输入。例如：`CS:1`。
- (5) `OLD:yesorno` 其中 `OLD` 是参数名，`yesorno` 是对“是否需要复制一套旧名称的程序集”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于升级步骤 6 的输入。例如：`OLD:NO`。

各参数名不区分大小写，参数之间没有顺序限定。参数名与参数值之间必须用冒号‘:’分隔，参数名与‘:’之间不能有空格，参数值与‘:’之间的空格将看成参数值的一部分。如果参数值包含了数字和字母以外的符号，则必须将参数名、‘:’，以及参数值作为整体用双引号括起来。如果输入项没有出现对应的参数，则仍然要通过外部输入。^[4]

16.4.4 服务器卸载

TRS 服务器没有自动的卸载程序，卸载系统时需要手工操作，其步骤如下：

- (1) 如果注册了 Windows 服务，则需使用相应的 ‘tdbservice -remove [-n:name]’ 命令先删除该服务，否则要使用 ‘tdbshutdown’ 命令关闭服务器。
- (2) 一般情况下安装程序会在操作系统的 services 文件里写入 “trsserver” 对应的端口号，可打开该文件直接删除。services 文件在 Windows NT 中是 ‘Windows 根目录\system32\drivers\etc\services’，在 Windows 9x 中是 ‘Windows 根目录\services’。
- (3) 如果使用的 InstallShield 安装方式，则在“开始”菜单的相应程序文件夹里有卸载程序链接，可以启动卸载程序进行卸载。
- (4) 删除服务器所在的目录。如果为服务器添加过其它的存储目录，则这些目录也需手工删除。

16.5 失败后的处理

16.5.1 原因及解决方法

下面几个原因可能导致 TRS 服务器安装失败：

- (1) 安装键值不合法。在此种情况下，仔细检查所输入的键值是否正确，如果输入键值确实是正确的，请与 TRS 信息技术公司联系。
- (2) 所输入的安装目标文件夹已经存在，或者驱动器不存在。
- (3) 写盘失败。检查目标文件夹是否可‘写’，空间是否足够。

16.5.2 清理工作

TRS 服务器安装失败后，请将服务器安装目标文件夹删除，以便于以后重新进行 TRS 服务器的安装。

16.6 内存映射文件

用于存放数据词典共享信息的共享内存映射文件（文件名称为 TxxxxxxSH，其中的 xxxxxx 对应于这里给出键值）的存放目录【修订 143】的缺省位置不再是以前的 Windows 系统的临时目录，而是此 system 目录的 shared 子目录^[4]。

共享内存映射文件的存放目录也可以由你自己确定：以 NT 系统为例，通过控制面板，在“系统变量”中新增环境变量 SHARE_MEMORY_PATH，该变量的值就是你要指定的共享内存映射文件的存放目录，之后必须重新启动机器，因为只有这样，新增环境变量才能有效。

16.7 安装多套系统

在同一台机器上需要安装多套 TRS 全文数据库系统，一般可以直接进行，但遇特殊情况或发生冲突时，需要关闭正在运行的 TRS 全文数据库系统，并修改这些系统的 system 目录下 KEY.INI 文件中从 TOKEN 到 INSTANCE 的键值，以保证与缺省取值以及各套系统之间互不相同，**一般建议在增加或减少键值的位数，或修改高位数上的数字，以在各套系统的键值之间保持较大的间隔**，如分别改成 5000、1001、1002、1003、1004、1005、1006、1007、1008 和 1009。在启动不同的 TRS 全文数据库系统时，当然还需要指定不同的网络服务端口，而不能使用同一端口。

另外如果还需要将多套 TRS 全文数据库系统都作为 NT 服务登记到 Windows 系统的服务控制管理器中，则需要用命令 tdbservice 来完成登记操作，同时指定不同的服务名和网络服务端口。

第 17 章 在UNIX下安装

17.1 目录结构

图 17-1 是安装好以后 TRS 服务器在操作系统中的目录结构图。

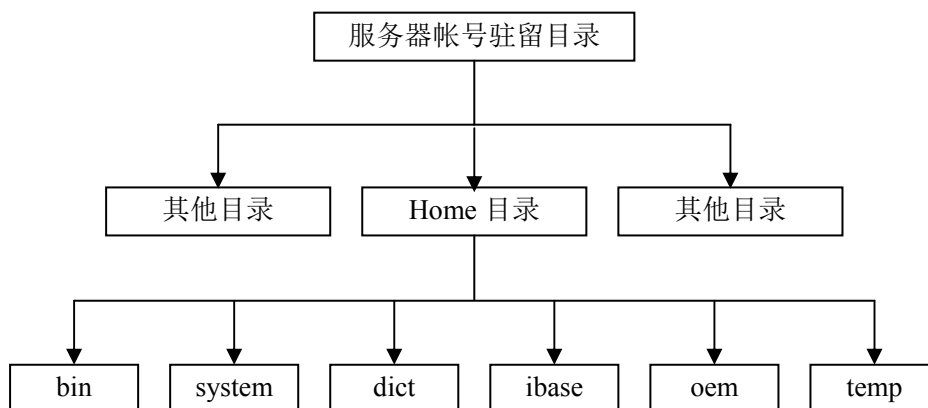


图 17-1

17.1.1 bin目录

bin 目录存放的是服务器的所有可执行程序以及版本信息文件，包括：

- ◇ readme.txt 这是一个文本文件，用来说明当前服务器版本所做的改进。
- ◇ trsidenter 获得服务器的机器码，用于生成运行时所必须的注册码。
- ◇ trserver 启动服务器的程序。
- ◇ trsearcher 启动服务器检索缓冲模块的程序。
- ◇ trsmirror 启动服务器数据镜像模块的程序。
- ◇ tdbshutdown 关闭服务器的程序。
- ◇ tdbdict 诊断与修复服务器的数据词典的工具。
- ◇ tdbmaster 服务器管理员。
- ◇ tdbload 数据库加载工具。

17.1.2 system目录

system 目录存放服务器的数据词典、配置文件以及 KEY.INI 文件。该目录下的文件是服务器的核心文件，不能破坏，否则服务器将无法正常运行。除配置文件

以及 KEY.INI 文件外，其他文件均由服务器自动维护。

在 system 目录下有四个配置文件，分别是：运行日志配置文件 LOG.INI、数据镜像模块配置文件 MIRROR.INI、插件配置文件 OEM.INI，以及运行参数配置文件 PROFILE.INI。其详细说明请参见配置文件。

KEY.INI 文件记载了服务器运行时所需要的一系列键值，它是一个文本文件，每一行一个键值，其形式是：“键名 = 键值”，其中的键值除个别外都可以进行编辑修改。

✧ **PRODUCT = 1.IBQI4CBD-GHA4CD21-PTKLPHSD**

键 PRODUCT 的值是 TRS 信息技术股份有限公司提供给用户的产品注册码，用户不能随意改动。如果服务器由于该键值不正确而不能正常启动或运行时，请与 TRS 信息技术股份有限公司的技术支持联系。

✧ **TOKEN = 40000**

键 TOKEN 的值是服务器的动态信号量系统的初始信号量 IPC 键值，同时 TOKEN+1 (40001)，TOKEN+2 (40002) 也将被使用。这三个键值不能与 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

✧ **CONFIG = 10000**

键 CONFIG 的值是服务器存储系统配置数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

✧ **GROUP = 10001**

键 GROUP 的值是服务器存储用户组数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

✧ **USER = 10002**

键 USER 的值是服务器存储用户数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

✧ **BASE = 10003**

键 BASE 的值是服务器存储数据库数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

✧ **BASECOL = 10004**

键 BASECOL 的值是服务器存储数据库字段数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。

- ◇ VIEW = 10005
键 VIEW 的值是服务器存储视图数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。
- ◇ VIEWCOL = 10006
键 VIEWCOL 的值是服务器存储视图字段数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。
- ◇ RIGHT = 10007
键 RIGHT 的值是服务器存储权限管理数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。
- ◇ FORMAT = 10008
键 FORMAT 的值是服务器存储数据格式数据词典的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突。在确保唯一性的前提下，该键值可以被修改。
- ◇ INSTANCE = 10009
键 INSTANCE 的值是服务器存储用户登录信息的共享内存段的 IPC 键值。这个键值不能与以上各键值，以及 UNIX 系统中的其他 IPC 键值冲突，同时 INSTANCE+1 (10010) 也将被占用，必须保留。在确保唯一性的前提下，该键值可以被修改。

在 system 目录里，除了存放的一些系统文件外，还包含三个子目录：etc、inherit 和 socket，这三个子目录里的文件由系统自己维护，在非特殊情况下，外部不需要也不能进行修改或删除，否则系统将有可能被破坏，不能正常运行。

17.1.3 dict目录

dict 目录存放的是服务器的各种知识词典，包括分词词典、附加分词词典、停用词典、附加停用词典、主题词典、同义词典和反义词典等。TRS 全文数据库系统允许用户自己创建这些知识词典，在没有指定词典存放目录的前提下，这些词典也存放在该目录。

17.1.4 ibase目录

ibase 目录是服务器存放数据库的缺省目录，包括数据库的全部数据和索引。只要磁盘空间允许，用户创建的任何数据库都可以存放在该目录下。TRS 全文数据库系统允许用户设置并使用其它的目录来存储数据库。

17.1.5 oem目录

oem 目录将主要用来存放为 TRS 全文数据库系统配置的 OEM 产品。

17.1.6 temp目录

temp 目录是 TRS 全文数据库系统的临时文件公用区。服务器在运行时要产生许多临时文件，如历史日志文件、检索结果历史文件、输出检索结果记录时产生的临时文件、数据库加载时网络传输产生的临时文件等等，这些文件都存储在该目录中。有时还要在该目录下自动细分更多的子目录，特别地，LOGS 子目录是用来存放客户端请求的日志文件^{【修订 111】}以及系统启动、关闭操作的日志文件（system.log）的。^{【2】}

17.2 准备工作

17.2.1 软件及硬件环境

服务器的运行需要一定的软硬件环境，下面所列的各项是最基本的要求，在安装之前请确认一下您的机器是否满足这些条件：

- 运行 UNIX 或 LINUX 操作系统的小型机、服务器、工作站、Pentium 及其兼容机。
- 内存^{【修订 007】} 1GB 以上，建议使用 4GB 内存。
- 一个可用空间在 1GB 以上的独立硬盘或分区用于安装服务器。
- 操作系统可使用的对换空间最好在 1GB^{【4】} 以上。
- 局域网连接设备。

如果你要求 TRS 服务器支持大文件（即 largefile，有的系统称超过 2G 的文件为大文件，而有的系统则认为超过 1G 的文件就是大文件），则必须使用支持大文件的安装包进行安装，同时还必须保证用来存储数据库文件的文件系统允许大文件的读写操作，即清除文件系统本身对大文件的“人为”限制，例如：SUN Solaris 要用 mount_ufs 命令的 largefiles 选项安装目标文件系统（无论操作系统是 64 位，还是 32 位的）；IBM AIX 要将系统的/etc/security/limits 配置文件的 fsize 和 fsize_hard 参数值设置为-1，并且用 smit 工具将文件系统的属性设置成“Large File Enabled”（无论是 64 位的操作系统，还是 32 位的操作系统）。HPUX 可通过系统管理工具 sam 进行配置管理。

文件系统的性能对 TRS 的影响极大，通过一些基本的测试我们发现：在 Linux 中使用 ext3 文件系统（主要在于缺省的日志模式）；在 Solaris 中使用 zfs 文件系统要比使用 ufs 时的性能好得多。

TRS 服务器使用了 UNIX 系统 IPC（进程间通讯）机制中的“共享内存”和“信号量”，它们在系统内有一定的限制，有的操作系统的缺省限制比较小，以至 TRS 服务器在安装时，或者在安装完后的运行过程中，报告 17004、17439、17448~17451 号错误，这需要修改操作系统的缺省配置。遇到这种情况比较多的是 SUN Solaris 系统，其修改方法是在系统配置文件“/etc/system”中加入（或修改）下列各项配置（参考值）：

配置	意义
set shmsys:shminfo_shmmax=67108864	一个共享内存段的最大字节数
set shmsys:shminfo_shmmin=1	一个共享内存段的最小字节数
set shmsys:shminfo_shmmni=512	系统中共享内存段标识符的最大个数
set shmsys:shminfo_shmseg=64	单进程同时使用的共享内存段最大数目
set semsys:seminfo_semmni=128	系统中信号量集（KEY）的最大个数
set semsys:seminfo_semmns=8192	系统信号量的最大数目（semmni*semmsl）
set semsys:seminfo_semmsl=64	单个信号量集所含信号量的最大数目
set semsys:seminfo_semopm=32	单次调用所操作信号量的最大数目
set semsys:seminfo_semmap=64	系统中信号量映射入口的最大数目
set semsys:seminfo_semmnu=4096	系统中信号量的 UNDO 结构最大数目
set semsys:seminfo_semume=4096	单进程的信号量 UNDO 结构最大数目

HPUX 的相关配置项与 Solaris 非常相似，但需要在系统管理工具 sam 中进行配置管理，而不是直接修改配置文件。

Linux 有关信号量的限制在“/etc/sysctl.conf”文件中，对应于变量“kernel.sem”。这个变量包含四个值，按顺序分别是 SEMMSL SEMMNS SEMOPM SEMMNI，例如：250 32000 32 128。

SEMMSL 单个信号量集所含信号量的最大数目。

SEMMNS 系统信号量的最大数目（SEMMNI * SEMMSL）。

SEMOPM 单次调用所操作信号量的最大数目。

SEMMNI 系统信号量集（KEY）的最大个数。

可用“/sbin/sysctl -a”查看“kernel.sem”当前的系统设置。如果需要修改时，可直接编辑“/etc/sysctl.conf”文件，添加或修改“kernel.sem”变量，也可用下面的命令之一进行操作：

```
sysctl -w kernel.sem="250 32000 100 128"
echo "kernel.sem=250 32000 100 128" >> /etc/sysctl.conf
```

Linux 共享内存的限制也是在“/etc/sysctl.conf”文件中，但分别有三个变量：

kernel.shmall 系统中共享内存的最大页数。
kernel.shmmax 一个共享内存段的最大字节数。
kernel.shmmni 系统中共享内存段标识符的最大个数。
这些变量值的查看和修改，与上述 kernel.sem' 变量相同，参照操作即可。

AIX 使用不同的方法，在 AIX 中定义了 IPC 机制的上限，是不可配置的，按需分配和释放 IPC 数据结构。

有的操作系统（例如 AIX、Linux 等）还有一些其它配置，如：cpu、data、stack、rss、nofiles 等，这些配置的限制也有可能影响 TRS 服务器的正常运行（如客户端报告 17813 或 17004 号错误），所以建议进行适当的修改，但无法给出一个参考值，或者都简单地设置成 unlimited。这里需要注意的是，有的系统用 -1 表示^{【修订 008】}unlimited，而有的系统设成 -1 时，相当于设置了一个巨大的整数，对于堆栈（stack）也是这样^{【2】}。

改变操作系统的配置后，必须重新启动机器才有效。

17.2.2 服务器的用户帐号

TRS 服务器的各种文件必须驻留在操作系统的用户帐号之下，因此在安装之前要首先为 TRS 服务器确定一个操作系统用户，以下简称为服务器帐号。操作系统中已有的任意一个用户帐号都可作为服务器帐号，但最好还是专门创建一个用户帐号作为服务器帐号。服务器帐号所在的硬盘或分区的可用空间应在^{【修订 009】}1GB 以上，对服务器帐号在硬盘空间使用上的限制不能小于 1GB，具体应与实际的数据规模相适应^{【2】}。

17.2.3 确认安装方式

TRS 服务器的安装有三种方式：

- 从硬盘上安装。这种方式直接从硬盘上打开安装包进行安装。
- 从光盘上安装。这种方式要求机器具有光盘驱动器，并安装了该设备。
- 从网络上安装。这种方式需要先将安装包从网络上传到机器的硬盘上，然后打开硬盘上的安装包进行安装。

17.3 安装步骤

- (1) 确定进行安装操作的用户帐号，即服务器帐号，其所在目录服务器帐号驻留目录。
TRS 服务器的安装既可以在服务器帐号下进行，也可以在 root 帐号下进行，一般建议使用 root 帐号，因为安装程序需要将 TRS 服务器的网络端口号加入到系统文

件 '/etc/services' 中, 而这个文件只有 root 帐号具有 '写' 权限, 如果使用其他帐号, 则需要系统管理员在 root 帐号下手工加入一行 "trserver 端口号/tcp"。不管使用哪个用户帐号, 都必须满足以下几个条件:

- 该用户能够打开安装包。
- 该用户能够对服务器帐号下的所有目录及文件具有 '读写' 权限。
- 该用户能够对服务器帐号下的所有目录及文件成功实施 'chown'、'chgrp' 和 'chmod' 三种操作。

这里需要澄清的是: 建议在 root 下安装, 并不等于建议将 TRS Server 安装在 root 帐号下, 而还是需要为 TRS Server 分配一个帐号便于文件的权限管理。

- (2) 登录到安装用户帐号或 root 帐号。
- (3) 进入到一个适当的目录, 并用命令 'tar -zxvf 安装包名' 解开安装包。安装包的命名形式为: TRS6.OS.MACHINE.nnBxxxx.tar.Z
其中: 'TRS6' 表示该安装包是一个 TRS6 服务器的安装介质 (即可以安装正式版本, 也可以安装试用版本; 'OS' 表示该服务器介质生成时的软件环境, 即操作系统名称和版本号; 'MACHINE' 表示该服务器介质生成时的硬件环境, 一般指机器的 CPU 名称; 'nn' 是编译环境的机器字位数, 如 '64', 'L32' 表示在 32 位环境下支持 64 位文件; 'Bxxxx' 是该服务器的生产批号, 如 'B4000'; 'tar' 表示该包是使用 'tar' 命令归档而成的格式文件; 'Z' 表示该安装包已进行了压缩, 需要在 tar 命令里加上 'z' 选项才能解开安装包, 如果操作系统里的 tar 命令不支持该选项, 则需要先使用系统的 'uncompress' 命令进行解压, 如果操作系统里没有安装 'uncompress' 命令, 则可以在其它的系统里使用支持该压缩格式的工具进行解压。如安装包: TRS6.AIX51.POWER4.64B4000.tar.Z。
- (4) 安装包被打开后, 在当前目录下会自动产生一个 'package' 子目录 (这时可以删除安装包了), 进入该子目录。
- (5) 运行安装程序 tdbinstall。安装程序将询问你是否安装授权版本: "Would you install TRS Server with KEY authorization? (Y/N)". 回答 "Y" 表示安装授权 (正式) 版本, 否则表示安装试用版本。
- (6) 如果安装试用版本, 则跳过这一步, 否则安装程序将提示并等待你输入产品注册码, 即 "Input PRODUCT KEY"。产品注册码是由厂家根据机器码生成的, 所以您需要与厂家联系, 并将您的机器码告诉我们。您的机器码已经在屏幕上打印出来: "Your machine code is xxxxxxxx zzzzzzzz", 其中 "xxxxxxx" 和 "zzzzzzz" 所对应的 8 个字母数字串就是机器码, 其中可能没有后面的 "zzzzzzz", 而只有前面的 "xxxxxxx"。也可以通过运行安装包目录下的程序 'trsidenter' 得到同样的机器码。
- (7) 输入服务器帐号, 即 "Input the UNIX account for TRS Server(e.g.: trs):"。这里要输入服务器帐号就是在 17.2.2 节所确定的用户帐号。
- (8) 输入服务器的安装目录名, 即 "Input a target dir in the account(e.g.: trshome):"。这个目录不是文件系统下的绝对路径, 而是相对于服务器帐号的驻留目录下的相对

路径，但如果输入内容的第一个字符为 ‘/’，则认为输入是文件系统下的绝对路径。比如输入的是 ‘trshome’，则服务器将安装在 ‘服务器帐号驻留目录/trshome’ 目录下。注意：如果目录名中有空格，则需用双引号将其括起来。

- (9) 如果输入的安装目录不存在，或不包含任何文件与子目录，则跳过这一步。

如果目录已存在，^[修订 108]且含有文件，则终止安装。

如果目录已存在，且含有空的子目录，但不含有任何文件，则安装程序提示你确定是否需要删除（删除并重建）这个目录并继续安装，即：

```
WARNING: The inputted dir (trshome) already existed!
```

```
Do you want to clear this dir, and continue? (Y/N) [4]
```

回答“Y”表示将自动清空该目录，否则结束安装过程。**特别提醒：“清空该目录”的涵义是删除并重建该目录，所有空子目录都不保留。**所以用于安装的安装包不能放在这个目录下。

- (10) 安装程序提示：

```
Please select the character set for TRS Server.
```

```
1: GB18030
```

```
2: UTF8
```

```
3: BIG5
```

```
You select:
```

如果输入 ‘1’，表示以 GB18030 编码为主字符集进行安装；如果输入 ‘2’，表示以 UTF8 编码为主字符集进行安装；如果输入 ‘3’，表示以 BIG5 编码为主字符集进行安装。

- (11) 输入以上各项后，安装程序进入自动运行状态，加载服务器所必须的各种知识词典和几个示例数据库。
- (12) 知识词典和示例数据库加载完成后，如果当前操作用户不是 root 帐号，则安装程序打印提示信息 “You aren't super user, so you can't add a service port automatically to the SERVICES file of system!” 之后结束；否则安装程序提示你输入服务器使用的网络端口号，即 “Input port number”，如输入 “8888”，并 “回车”。
- (13) 安装成功之后，安装程序将进一步提示你 “Would you test file system for large file exceeding 4G? (Y/N)” 如果回答 “Y”，则安装程序开始测试服务器所在的 “文件系统设备” 是否支持大文件的读写。
- (14) 结束。（这时可删除 ‘package’ 子目录了）

上述安装过程是交互式的。^[修订 099]安装程序 `tdbinstall` 也支持通过命令行的参数方式进行输入。参数与交互式的输入项是一一对应的，其定义如下：

- (6) KEY:code 其中 KEY 是参数名，code 是一个实际的 KEY 串。该参数对应于安装步骤 5 和 6 的输入。例如：KEY:1.IBQI4CBD-GHA4CD21-PTKLPHSD。如果 code 对应的是 “DEMO”，则表示安装试用版。
- (7) USR:account 其中 USR 是参数名，account 是一个实际的用户帐号。该参数对应

于安装步骤 7 的输入。例如：USR:trs。

- (8) DIR:path 其中 DIR 是参数名，path 是服务器要安装的目录。该参数对应于安装步骤 8 的输入。例如：“DIR:/home/server”。
- (9) REW:yesorno 其中 REW 是参数名，yesorno 是对“是否需要覆盖已经存在的目录”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于安装步骤 9 的输入。例如：REW:YES。
- (10) CS:charset 其中 CS 是参数名，charset 是系统字符集的选择（1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于安装步骤 10 的输入。例如：CS:1。
- (11) PN:port 其中 PN 是参数名，port 是服务器的缺省端口号。该参数对应于安装步骤 12 的输入。例如：PN:8888。
- (12) LF:yesorno 其中 LF 是参数名，yesorno 是对“是否进行大文件支持测试”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于安装步骤 13 的输入。例如：LF:NO。

各参数名不区分大小写，参数之间没有顺序限定。参数名与参数值之间必须用冒号‘:’分隔，参数名与‘:’之间不能有空格，参数值与‘:’之间的空格将看成参数值的一部分。如果参数值包含了数字和字母以外的符号，则必须将参数名、‘:’，以及参数值作为整体用双引号括起来。如果输入项没有出现对应的参数，则仍然要通过外部输入。^[4]

17.4 运行与维护

17.4.1 启动服务器

- 在服务器帐号下启动
 - (1) 登录到服务器帐号。
 - (2) 进入服务器的安装目录，如‘trshome’。
 - (3) 进入下一级子目录 bin。
 - (4) 运行 trsserver。如果这时打印出“Start up TRS Server successfully!”，则表示服务器已经启动成功；如果这时打印出没有端口号的错误（一般是 17803 号错误），则可以使用‘trsserver 端口号’（如‘trsserver 8888’）的方式启动服务器，也可以在系统文件‘/etc/services’中添加‘trsserver 端口号/tcp’这一行后再运行 trsserver；如果出现其他错误而不能启动服务器，则请参看错误信息表后进行相应处理，或者与相关的技术支持联系。
 - (5) 在缺省情况下，trsserver 启动时会自动启动 trssearcher，不需要单独启动。但如果在配置文件 PROFILE.INI 里没有将 trssearcher 设置成自动启动，或其它一些原因导致 trssearcher 未启动，则可以单独运行 trssearcher 启动服务器的

检索缓冲模块。与 trsserver 一样，它也可以带一个端口号参数，但必须与 trsserver 启动所使用的端口号相同。在一般情况下，trssearcher 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取 trsserver 启动所使用的端口号，而无须特别指定。

- (6) 运行 trsmirror 启动服务器的数据镜像模块。与 trsserver 一样，它也可以带一个端口号参数，但必须与 trsserver 启动所使用的端口号相同。在一般情况下，trsmirror 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取 trsserver 启动所使用的端口号，而无须特别指定。

- 在 root 账号下启动

- (1) 登录到操作系统的 root 帐号。
- (2) 进入服务器帐号的驻留目录，如 '/usr1/trs'。
- (3) 进入服务器的安装目录，如 'trshome'。
- (4) 进入下一级子目录 bin。
- (5) 运行 trsserver 启动服务器。
- (6) 运行 trsmirror 启动服务器的数据镜像模块。
- (7) 运行 trssearcher 启动服务器的检索缓冲模块。

需要说明的是：虽然 TRS 服务器是由 root 启动的，但服务器运行时的实际用户帐号和用户组仍然是服务器帐号和它所在的用户组，而不是 root。另外，对于**核心版本在 2.2 以下的 Linux 操作系统**，服务器要求在 root 账号下启动，虽然在服务器帐号也能成功启动服务器，但由于无法使用进程池技术，其运行效率要比由 root 启动时低。

- 在操作系统启动时自动引导

- (1) 假设你已经知道如何在操作系统启动时，自动启动你自己的应用程序。如果你不知道怎么做，则请咨询一下操作系统厂家的售后服务机构。
- (2) 将 '服务器帐号的驻留目录/服务器的安装目录/bin/trsserver 端口号' 加入到操作系统的启动文件中。当系统文件 'etc/services' 中存在 'trsserver xxxx/tcp' 时，其端口号可以省略。如加入：'/usr1/trs/trshome/bin/trsserver'。
- (3) 重新启动时，操作系统将自动启动 TRS 服务器。

借助 TRS 全文数据库系统的运行参数配置文件 PROFILE.INI 的 AutoBatch 配置域，可以将服务器的数据镜像模块 (trsmirror) 和服务器的检索缓冲模块 (trssearcher) 与 trsserver 一起自动启动，请参见配置文件中的详细说明。

17.4.2 关闭服务器

- 在服务器帐号下关闭
 - (1) 登录到服务器帐号。
 - (2) 进入服务器的安装目录，如 ‘trshome’。
 - (3) 进入下一级子目录 bin。
 - (4) 运行 `tdbshutdown`。如果这时打印出错误信息，并提示你指定端口号，则可以使用 ‘`tdbshutdown 端口号`’ 的方式关闭服务器，其实这种方式可在任何情况下使用，如 ‘`tdbshutdown 8888`’。在一般情况下，`tdbshutdown` 可以从与程序所在目录平级的 `system` 子目录中的 `START.INF` 文件中自动获取 `trserver` 启动所使用的端口号，而无须特别指定。
 - (5) 如果运行 `tdbshutdown` 时带上 ‘`-cache`’ 选项参数，则只关闭服务器的检索缓冲模块。如 ‘`tdbshutdown -cache`’。
 - (6) 如果运行 `tdbshutdown` 时带上 ‘`-reboot`’ 选项参数，则表示关闭并重新启动服务器。如 ‘`tdbshutdown -reboot`’。该选项的功能与“先关闭，再手工启动”的区别在于系统在“空闲状态起止时间”下的自动维护功能的启动时间，请参见 5.7 节的描述。【修订 152】
 - (7) 运行 `trsmirror -shutdown`，关闭服务器的数据镜像模块。它需要与 `tdbshutdown` 完全一致的‘端口号’参数，如 ‘`trsmirror -shutdown 8888`’。在一般情况下，`trsmirror` 可以从与程序所在目录平级的 `system` 子目录中的 `START.INF` 文件中自动获取 `trserver` 启动所使用的端口号，而无须特别指定。

- 在 root 账号下关闭
 - (1) 登录到操作系统的 root 帐号。
 - (2) 进入服务器帐号的驻留目录，如 ‘`/usr1/trs`’。
 - (3) 进入服务器的安装目录，如 ‘trshome’。
 - (4) 进入下一级子目录 bin。
 - (5) 运行 `tdbshutdown`，关闭服务器。
 - (6) 如果运行 `tdbshutdown` 时带上 ‘`-cache`’ 选项参数，则只关闭服务器的检索缓冲模块。
 - (7) 运行 `trsmirror -shutdown`，关闭服务器的数据镜像模块。

借助 TRS 全文数据库系统的运行参数配置文件 `PROFILE.INI` 的 `AutoBatch` 配置域，可以将服务器的数据镜像模块 (`trsmirror`) 在 `trserver` 关闭时自动关闭，请参见配置文件的详细说明。

无论是用哪一种方式关闭服务器，所使用的端口号必须与服务器启动时的端口号一致。在服务器关闭时并不是立即退出，它必须等待所有正在执行的服务请求完成后

才彻底结束，但是这时服务器不再会响应新的服务请求了。如果系统中已没有任何名为 `trserver` 的进程时，表示服务器已正式关闭。如果只要求 `tdbshutdown` 向服务器发送“关闭”命令后立即退出，而不管服务器进程是否已经退出结束，则可使用‘`-nowait`’参数。

有时候需要强制“杀死”TRS 服务器（例如在不能正常关闭的情况下），虽然可以使用操作系统的‘`kill`’命令，或通过命令‘`tdbshutdown -kill:ProcessID`’完成，但是比较烦琐，并且还要手工清除 TRS 服务器所创建的 IPC 资源（如共享内存、信号量等）。用命令‘`tdbshutdown -kill 端口号`’可自动完成所有这些操作。

17.4.3 服务器升级

与服务器的安装一样，服务器的升级也有标准的升级包。升级包的命名以 `TRS6UPG` 开头，其余与安装包一致，如：`TRS6UPG.AIX51.POWER.64B4000.tar.Z`。升级包被打开后，在当前目录下会自动产生一个‘`upgrade`’子目录（这时可以删除升级包了），进入该子目录进行服务器的自动升级（升级完成后，删除该子目录）。

服务器升级包中必须包含下列文件：`readme.txt`、`trserver`、`trsearcher`、`trsmirror`、`tdbdict`、`tdbmaster`、`tdbload`、`trsidenter`、`tdbshutdown`、`tdbupdate`、`LOG.INI`、`MIRROR.INI`、`OEM.INI`、`PROFILE.INI`，以及新词典的资源文件 `segment.dic`、`stop.dic`、`ambiguous.dic`、`wbigram.dic`、`cbigram.dic` 和 `single.dic`。这些文件必须放在同一目录下。运行 `tdbupdate` 即可完成 TRS 服务器系统文件的自动更新。

升级步骤：

- (1) 在进行服务器升级前，必须首先关闭服务器，然后运行 `tdbupdate`。
- (2) 如果所运行的 `tdbupdate` 处在服务器的 `bin` 目录下，则进行数据词典的更新，而不会进行 TRS 服务器系统文件的复制。
- (3) 更新过程输入提示：“Input the UNIX account owned TRS Server(e.g.: trs):”，即要求你输入服务器所在驻留帐号。请参见安装过程的第 7 步。
- (4) 更新过程输入提示：“Input the target dir in the account (e.g.: trshome):”，即要求你输入已经安装的服务器所在的路径。请参见安装过程的第 8 步。
- (5) 接下来，升级程序提示：

Please select the character set for added dictionaries:

- 0: NONE
- 1: GB18030
- 2: UTF8
- 3: BIG5

You select:

输入 ‘0’，表示不增加新的词典；输入 ‘1’，表示增加 GB18030 字符集的新词典；如果输入 ‘2’，表示增加 UTF8 字符集的新词典；输入 ‘3’，表示增加 BIG5 字符集的新词典。

升级程序也支持修改更新现有的词典，要启用这个功能，则必须在命令行中给出参数 “DIC:REW” 或 “DIC:RIX”，参数的意义请参见下面的“参数方式”描述部分。此时：

Please select the character set for modified dictionaries:

0: NONE

1: GB18030

2: UTF8

3: BIG5

You select:

输入 ‘0’，表示不修改更新已有词典；输入 ‘1’，表示修改更新 GB18030 字符集的已有词典；输入 ‘2’，表示修改更新 UTF8 字符集的已有词典；输入 ‘3’，表示修改更新 BIG5 字符集的已有词典。

- (6) 如果原系统未设置主字符集（系统字符集），则可能会要求你回答是否需要设置系统的主字符集。

Please select the character set for TRS Server:

0: don't set the native charset.

1: set GB18030 as the native charset.

2: set UTF8 as the native charset.

3: set BIG5 as the native charset.

You select:

如果不设置主字符集，则系统不会自动完成客户端加载和读取数据时的字符集转换，也不能创建 UTF8 字符集的数据库。需要注意的是：你指定要设置系统的主字符集时，则要求当前系统中已有的全部数据库/词典的字符集都必须与所指定的主字符集一致，否则设置无效。

- (7) 如果原系统是 5.0 之前的版本，则要求你回答是否需要复制一套旧名称的程序集，即 “Do you want to keep the old names? (Y/N)”。如果不复制，则有可能影响你在其它地方对旧名称的程序的引用，如：在机器启动时自动启动设置等。

上述升级过程是交互式的。【修订 100】升级程序 `tdbupdate` 也支持通过命令行的参数方式进行输入。参数与交互式的输入项是一一对应的，其定义如下：

- (1) `USR:account` 其中 `USR` 是参数名，`account` 是服务器的用户帐号。该参数对应于升级步骤 3 的输入。例如：`USR:trs`。
- (2) `DIR:path` 其中 `DIR` 是参数名，`path` 是服务器所在的目录。该参数对应于升级步骤 4 的输入。例如：`"DIR:d:\server"`。
- (3) `NEW:charset` 其中 `NEW` 是参数名，`charset` 是需要新增或更新的词典的字符集的

选择（0 表示不更新词典；1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于升级步骤 5 的输入。例如：NEW:1。

- (4) DIC:action 其中 DIC 是参数名，action 是针对需要更新的词典的动作选择（REW 表示只强制更新指定字符集的词典内容；RIX 表示强制更新指定字符集的词典内容，且删除相关数据库的索引）。该参数没有对应的升级步骤，不支持外部输入。例如：DIC:RIX。
- (5) CS:charset 其中 CS 是参数名，charset 是设置系统字符集的选择（0 表示不设置系统字符集；1 表示 GB18030；2 表示 UTF8；3 表示 BIG5）。该参数对应于升级步骤 6 的输入。例如：CS:1。
- (6) OLD:yesorno 其中 OLD 是参数名，yesorno 是对“是否需要复制一套旧名称的程序集”的回答（“yes”或“y”或“no”或“n”，不分大小写）。该参数对应于升级步骤 7 的输入。例如：OLD:NO。

各参数名不区分大小写，参数之间没有顺序限定。参数名与参数值之间必须用冒号‘:’分隔，参数名与‘:’之间不能有空格，参数值与‘:’之间的空格将看成参数值的一部分。如果参数值包含了数字和字母以外的符号，则必须将参数名、‘:’，以及参数值作为整体用双引号括起来。如果输入项没有出现对应的参数，则仍然要通过外部输入。^[4]

17.4.4 服务器卸载

TRS 服务器没有自动的卸载程序，卸载系统时需要手工操作，其步骤如下：

- (1) 使用‘tdbshutdown’命令关闭服务器。
- (2) 一般情况下安装程序会在操作系统的 services 文件里写入“trserver”对应的端口号，可打开该文件直接删除（需要 root 权限）。services 文件是‘/etc/services’。
- (3) 删除服务器所在的目录。如果为服务器添加过其它的存储目录，则这些目录也需手工删除。

17.5 失败后的处理

17.5.1 原因及解决方法

下面几个原因可能导致 TRS 服务器安装失败：

- (1) 安装键值不合法。在此种情况下，仔细检查所输入的键值是否正确，如果输入键值确实是正确的，请与 TRS 信息技术公司联系。
- (2) 所输入的服务器帐号不存在。请仔细确定服务器帐号。
- (3) 写盘失败。检查安装用户是否对服务器帐号具有‘读写’权限，空间是否足够。

17.5.2 清理工作

TRS 服务器安装失败后，请将服务器的安装目录从服务器帐号的驻留目录下删除，以便于以后重新进行 TRS 服务器的安装。

17.6 安装多套系统

在同一台机器上需要安装多套 TRS 全文数据库系统，一般可以直接进行，但遇特殊情况或为保险起见，需要关闭正在运行的 TRS 全文数据库系统，并修改这些系统的 system 目录下 KEY.INI 文件中从 TOKEN 到 INSTANCE 的键值，以保证与缺省取值以及各套系统之间互不相同，**一般建议在增加或减少键值的位数，或修改高位数上的数字，以在各套系统的键值之间保持较大的间隔**，如分别改成 5000、1001、1002、1003、1004、1005、1006、1007、1008 和 1009。在启动不同的 TRS 全文数据库系统时，当然还需要指定不同的网络服务端口，而不能使用同一端口。

第 18 章 数据库镜像

18.1 功能描述

- 单向一对多完全镜像。
- 可设置多个数据库做完全镜像，但最多不能超过 100 个数据库。
- 无论是主节点还是子节点，在启动镜像模块之前必须正确设置 MIRROR.INI 配置文件中的各个配置项。
- 主节点在镜像数据库上进行的任何记录增删改，将定期地复制到子节点上（增量式复制，每天一次）。
- 镜像初始化有两种方式：手工方式和自动方式。
- 当主节点的镜像数据库被删除时，应从 MIRROR.INI 中删除该镜像数据库的描述行，并重新启动 TRS Server；如果删除后，又重新创建该镜像数据库，则应重新启动 TRS Server。
- 子节点的镜像数据库，一般只有 SYSTEM 用户才有修改数据的权限。原则上可以进行任何数据库记录的增删改操作。但由于镜像是单向的，所以子节点应尽量少修改；镜像不会影响增加的记录（除非主节点有清空镜像数据库的操作）；删除记录后，该记录将失去镜像；在修改记录后，如果主节点上也对该记录进行了修改，则主节点上的修改将覆盖子节点上的修改。
- 不要修改子节点上的镜像数据库结构，也不要修改数据库名，否则可能无法顺利进行数据的镜像同步。

18.2 实施准备

各个节点在安装完 TRS 服务器后，就可以开始镜像的实施准备工作了。

18.2.1 主节点的准备工作

1. 启动 TRS 服务器。
2. 创建镜像用户组（mirror）。
3. 在镜像用户组下为所有子节点创建用户。
4. 创建镜像数据库。
5. 把主节点的 IP 地址和 TRS 服务器的服务端口号，以及子节点的用户名/口令通知给相应的子节点。

6. 关闭 TRS 服务器。
7. 组织好配置文件。
8. 启动 TRS 服务器。

18.2.2 子节点的准备工作

1. 组织好配置文件。
2. 启动 TRS 服务器。

注：子节点上的镜像数据库是自动创建的。

18.3 镜像初始化

各个节点在完成镜像的实施准备工作之后，接下来的工作就是进行节点的初始化。

18.3.1 手工初始化

1. 主节点上运行 ‘trsmirror -init=% [port]’，在 ‘../mirror’ 目录下将产生一个初始镜像数据包 (.INIT)。这里[port]为 TRS 服务器的服务端口号，以下同，不再说明。
2. 将生成好的初始镜像数据包发送给各个子节点。（发送完后可以删除）。
3. 各子节点接收到初始镜像数据包后运行 ‘trsmirror -load=filename [port]’。这里 filename 为初始镜像数据包的全路径及文件名。

18.3.2 自动初始化

1. 在主节点上运行 ‘trsmirror -renew=% [port]’ 即可完成，在子节点上不需要做任何工作。
2. 子节点上的数据库初始化在同步时间到达时才进行。
3. 这种方式只适合于存在少量数据的情况或局域网。

18.4 镜像的运行

1. 完成镜像的初始化工作之后，分别启动主节点和各个子节点上的 trsmirror。启动方法请参见 16.4.1 和 17.4.1 节。
2. 主节点定期产生镜像数据包 (.PKG)。

3. 子节点定期从主节点读取镜像数据包，并以此更新镜像数据库的内容。

18.5 新增镜像的子节点

1. 在镜像用户组（mirror）下为新的子节点创建用户，并把主节点的 IP 地址和 TRS 服务器的服务端口号，以及子节点的用户名/口令通知给新的子节点。
2. 按照 18.3 的步骤初始化新的子节点。但是，在新增镜像子节点时，手工方式的主节点需运行命令 ‘trsmirror -init=username [port]’，即用新的子节点的用户名替代 18.3.1 对应命令中的 ‘%’；自动方式也需要用新的子节点的用户名替代 18.3.2 对应命令中的 ‘%’

18.6 新增镜像的数据库

1. 在主节点的 MIRROR.INI 中增加该镜像数据库的 DatabaseName 描述行和 NewDatabaseName 描述行。
2. 重新启动 TRS 服务器。

18.7 当镜像数据库被破坏时

1. 当某个子节点的镜像数据库被破坏时，则相当于新增一个子节点。这时可按照新增镜像子节点（18.5）中的方法初始化该子节点。
2. 当主子节点的镜像数据库被破坏时，则需要先从数据库备份中恢复镜像数据库，然后从镜像初始化（18.3）开始重新做。

第 19 章 服务器工具

19.1 服务器管理员

服务器管理员对应 TRS 安装目录的 bin 子目录下的 tdbmaster 命令。在操作系统的 SHELL 命令行上运行命令 ‘tdbmaster’ 或 ‘tdbmaster -help’，则屏幕上将打印出该命令的使用说明（方括号中的参数表示可以省略）：

Usage: tdbmaster [-P:Port] [-Command [Parameters]]

Port Server's port number(default in START.INF).

The command includes followings:

help Display this help.
licence Get licence of system.
charset Get charset of system.
password Get password of system.
endidle End idle state of system.
flush Flush log information.
list List all clients.
logout Logout clients by timeout.
logout all Logout all clients.
logout INSTANCE Logout client by instance ID.
user [NAMES] Show all or some users.
database [NAMES] Show all or some databases.
column [DATABASES] Show columns of all or some databases.
view [NAMES] Show all or some views.
backup PACKAGE Backup system to package path.
resume PACKAGE Resume system from package path.
save NAMES BACKUP Save some databases as backup.
restore BACKUP NAMES Restore the backup of databases.
restore BACKUP Show databases in the backup.
optimize [DATABASES] Optimize all or some databases.
check [DATABASES] Check all or some databases.
index [DATABASES] Index all or some databases.

recover DATABASES Recover records of some databases.
delete database#ID Delete a database by specified ID.
delete view#ID Delete a view by specified ID.
register CODE Register a new system License.
pathset DIR [NUMBER] Create and add config's paths.

19.1.1 指定服务器的端口

服务器管理员运行时需要与服务器进行连接，所以需要知道服务器启动时所使用的网络端口号。命令行中的 ‘-P:Port’ 参数就是服务器管理员用来指定连接服务器的端口号的。其中：‘-P:’ 是一个引导符，其后紧跟的是服务器端口号。如果该参数省略，则服务器管理员默认为服务器启动时的端口号。在一般情况下，tdbmaster 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取服务器启动所使用的端口号，而无须特别指定。

19.1.2 列出所有的用户实例

使用命令 ‘tdbmaster [-P:Port] -list’ 可列出当前连接到服务器上的所有客户端的信息。内容如下：

Instance	UserID	ClientIP	LastAccess
1	1	192.9.200.168	2009.04.17 12:58:05
2	1	192.9.200.169	2009.04.17 13:58:05
3	2	192.9.200.170	2009.04.17 14:58:05
4	3	192.9.200.171	2009.04.17 15:58:05

Total Online: 4

Total Active: 0

其中，‘Instance’ 是 TRS 的登录用户实例号，也可以说是客户端访问服务器的标识号；‘UserID’ 是客户端登录时所使用的 TRS 用户的 ID；‘ClientIP’ 是客户端所在的 IP 地址；‘LastAccess’ 是客户端使用该用户实例最近一次访问服务器的时间（年-月-日 时:分:秒）。^[104] ‘Total Online’ 是登录用户实例总数，‘Total Active’ 是输出该列表时正在执行请求操作的用户实例总数。^[105]

19.1.3 强制清除指定的用户实例

使用命令 ‘tdbmaster [-P:Port] -logout [INSTANCE | all]’ 可强制清除服务器中已有

的用户实例。其中 INSTANCE 是一个用户实例号，对应于命令 ‘tdbmaster [-P:Port] -list’ 输出结果中的 Instance，表示清除该用户实例；^{【修订 002】}all 表示清除所有当前未访问服务器的用户实例^{【9】}。如果 ‘-logout’ 没有任何其它参数，则表示在服务器中：如果存在未访问服务器的时间已超过系统配置数据词典中规定的最大时间间隔的用户实例，则只清除这些用户实例，否则清除所有未访问服务器的时间已超过系统配置数据词典中规定的最小时间间隔的用户实例。参见 5.2 和 5.3 节。

该命令的输出结果是实际被清除的用户实例数。

19.1.4 刷新日志缓冲区

命令 ‘tdbmaster [-P:Port] -flush’ 用于刷新客户端访问服务器的日志缓冲区，将缓冲区中的内容写入文件。

19.1.5 在线备份服务器系统

命令 ‘tdbmaster [-P:Port] -backup PACKAGE’ 用于将整个 TRS 服务器系统备份到指定的目录下。其中 PACKAGE 是用以存放备份包的还未创建的目录，不能是一个已存在的目录；并且该目录不能是服务器的安装目录的任何层次的子目录；如果目录中含有空格，则必须使用双引号将目录扩起来。进行服务器系统备份时需要注意的是：服务器必须已经启动。

19.1.6 恢复服务器系统

命令 ‘tdbmaster [-P:Port] -resume [PACKAGE]’ 用于从以前的一个备份包中恢复 TRS 服务器系统。其中 PACKAGE 对应命令 ‘tdbmaster [-P:Port] -backup PACKAGE’ 中的参数 PACKAGE，^{【修订 105】}省略时默认为当前 tdbmaster 所在目录的上一级目录^{【9】}。进行服务器系统恢复时需要注意的是：必须首先关闭现在运行的服务器（参数 ‘Port’ 用于自动检查服务器是否在运行）；必须运行备份包中的 bin 子目录下的 ‘tdbmaster’ 命令；TRS 服务器的驻留目录是由备份包自动确定的。

19.1.7 结束系统自动维护操作

使用命令 ‘tdbmaster [-P:Port] -endidle’ 可强制结束 TRS 全文数据库系统在进入空闲状态时启动的正在执行的自动维护操作。

19.1.8 查看系统超级用户口令

使用命令 ‘tdbmaster [-P:Port] -password’ 可获取 TRS 全文数据库系统的超级用户，即 SYSTEM 的口令，主要是为了帮助 TRS 的系统管理员在忘记口令的情况下获取口令。该命令的输出结果就是超级用户的口令。

注：Ruintime 版本的服务器不具有此项功能。

19.1.9 查看系统许可证

使用命令 ‘tdbmaster [-P:Port] -licence’ 可获取 TRS 全文数据库系统许可证的数目【修订 062】以及系统许可运行的有效日期【*】。该命令的输出结果如果是 “unlimited”，则表示 “无限用户数”。

19.1.10 查看系统字符集

使用命令 ‘tdbmaster [-P:Port] -charset’ 可获取 TRS 全文数据库系统的主字符集。该命令的输出结果是：GB18030、BIG5、UTF8 和 NONE（表示没有主字符集）。

19.1.11 优化数据库索引与数据

使用命令 ‘tdbmaster [-P:Port] -optimize [DATABASES]’ 可优化指定的一个或多个数据库。其中 DATABASES 是指定的数据库名，允许模糊匹配，也可以是一个数据库名的列表（用逗号或分号分隔，不允许出现空格，但允许模糊匹配）。如果数据库省略，即命令 ‘tdbmaster [-P:Port] -optimize’，则表示优化所有的数据库。【修订 034】

19.1.12 检查数据库索引和视图

使用命令 ‘tdbmaster [-P:Port] -check [DATABASES]’ 可检查指定的一个或多个数据库的索引是否完整。其中 DATABASES 是指定的数据库名，允许模糊匹配，也可以是一个数据库名的列表（用逗号或分号分隔，不允许出现空格，但允许模糊匹配）。如果数据库省略，即命令 ‘tdbmaster [-P:Port] -check’，则表示检查所有的数据库，同时还检查所有视图的完整性。如果检查发现存在不完整的索引（或有错误的视图），则输出相关的数据库名以及字段信息（或输出相关错误信息）。【修订 035】

19.1.13 增量创建数据库的索引

使用命令 ‘tdbmaster [-P:Port] -index [DATABASES]’ 可以增量创建（补齐）指定

的一个或多个数据库的索引，使其完整。其中 DATABASES 是指定的数据库名，允许模糊匹配，也可以是一个数据库名的列表（用逗号或分号分隔，不允许出现空格，但允许模糊匹配）。如果数据库省略，即命令 ‘tdbmaster [-P:Port] -check’，则表示补齐所有数据库的索引。无论索引是否完整，该命令都不影响已经创建的索引。【修订 036】

19.1.14 删除指定的数据库或视图

使用命令 ‘tdbmaster [-P:Port] -delete database#ID’ 和 ‘tdbmaster [-P:Port] -delete view#ID’ 可以分别删除指定的数据库和视图。参数 “database#ID” 中的 ID 是待删除的数据库的 ID；参数 “view#ID” 中的 ID 是待删除的视图的 ID。【修订 071】

19.1.15 查看用户对象的信息

命令 ‘tdbmaster [-P:Port] -user [NAMES]’ 用于查看指定的一个或多个用户对象的信息。其中 NAMES 是指定的用户名，允许模糊匹配，也可以是用户名的列表（用逗号或分号分隔用户名，不允许出现空格，但允许模糊匹配）。如果用户名省略，则表示查看所有用户对象的信息。【修订 072】

19.1.16 查看数据库对象的信息

命令 ‘tdbmaster [-P:Port] -database [NAMES]’ 用于查看指定的一个或多个数据库对象的信息。其中 NAMES 是指定的数据库名，允许模糊匹配，也可以是数据库名的列表（用逗号或分号分隔数据库名，不允许出现空格，但允许模糊匹配）。如果数据库名省略，则表示查看所有数据库对象的信息。【修订 073】

19.1.17 查看数据库的字段信息

使用命令 ‘tdbmaster [-P:Port] -column [DATABASES]’ 可以查看指定的一个或多个数据库的全部字段的信息。其中 DATABASES 是指定的数据库名（特别注意这里不是字段名），允许模糊匹配，也可以是数据库名的列表（用逗号或分号分隔数据库名，不允许出现空格，但允许模糊匹配）。如果数据库名省略，则表示查看所有数据库的全部字段的信息。【修订 074】

19.1.18 查看视图对象的信息

命令 ‘tdbmaster [-P:Port] -view [NAMES]’ 用于查看指定的一个或多个视图对象的信息。其中 NAMES 是指定的视图名，允许模糊匹配，也可以是视图名的列表（用逗号或分号分隔数据库名，不允许出现空格，但允许模糊匹配）。如果视图省略，则表

示查看所有视图对象的信息。【修订 089】

19.1.19 修改系统注册码

命令 ‘`tdbmaster [-P:Port] -register CODE`’ 用于修改替换系统的注册码。其中 CODE 是一个新的注册码，它必须是该系统能够识别的合法注册码。系统注册码就是我们常说的 KEY，它存在于系统的 `system` 子目录下的 `KEY.INI` 文件中，对应于其中的配置项 ‘`Product`’ 的值，所以也可通过手工操作完成修改。该命令将只保存一个备份在文件 `KEY.BAK` 中【修订 109】。【修订 148】在修改注册码的同时，将验证其有效性和合法性。如果注册码不是合法有效的，则不会替换原注册码，`KEY.INI` 保持不变。【4】

19.1.20 备份数据库

命令 ‘`tdbmaster [-P:Port] -save NAMES BACKUP`’ 用于备份数据库。其中：NAMES 是需要备份的目标数据库名，允许模糊匹配，也可以是数据库名的列表（用逗号或分号分隔数据库名，不允许出现空格，但允许模糊匹配）；BACKUP 是数据库备份的存放路径（其中包含“备份包名”），BACKUP 必须是一个不存在的路径，但“备份包名”前的目录则必须存在，且当前用户和服务器运行的用户有“写”操作权。【修订 137】

19.1.21 恢复或查看数据库备份

命令 ‘`tdbmaster [-P:Port] -restore BACKUP NAMES`’ 用于恢复数据库，如果省略 NAMES 参数，则表示查看数据库备份包里所包含的数据库。其中：BACKUP 是数据库备份的存放路径（其中包含“备份包名”），当前用户和服务器运行的用户对 BACKUP 必须有“读”操作权；NAMES 是需要恢复的目标数据库名，允许模糊匹配，也可以是数据库名的列表（用逗号或者分号分隔数据库名，不允许出现空格，但允许模糊匹配），省略时，表示只查看“备份包名”所包含的数据库，但不进行恢复操作。【修订 138】

19.1.22 添加数据库存储目录

命令 ‘`tdbmaster [-P:Port] -pathset DIR NUMBER`’ 用于创建新的数据库存储路径并自动添加到系统配置。如果 NUMBER 省略，或者为 0，或者为 1，则表示创建由 DIR 指定的单个目录，并添加到系统配置。如果 NUMBER 的值为 2~1024（最大为 1024），则又分为两种情况：第一种情况是当 DIR 指定的目录不存在时，则要求其上级目录必须是已存在的、可读写的目录，该命令将自动创建 DIR 目录，同时在 DIR 目录下创建 NUMBER 个子目录，并把这些子目录（即：DIR/1; DIR/2; DIR/3...）添加到系统配置；第二种情况是当 DIR 是一个已存在的目录时，则它必须是读写的，且 DIR 目录下不存在该命令所默认的 `ibases` 子目录，该命令将自动创建 `DIR/ibases` 目录，同

时在 DIR/ibases 目录下创建 NUMBER 个子目录，并把这些子目录（即：DIR/ibases/1; DIR/ibases/2; DIR/ibases/3...）添加到系统配置。^{【修订 157】}

19.2 数据库加载工具

数据库加载工具对应 TRS 安装目录的 bin 子目录下的 tdbload 命令。在操作系统的 SHELL 命令行上运行命令 ‘tdbload’ 或 ‘tdbload -help’，则屏幕上将打印出该命令的使用说明（方括号中的参数表示可以省略）：

Usage: tdbload [-P:Port] [-CS:Charset] [-Options] [Database Source [Control]]

Usage: tdbload [-P:Port] [-CS:Charset] [-Options] [Database Source [-F:Name]]

Port The server's port number(default in START.INF).
Charset The charset of source file(default is same as Database's).
Format The record format of source file(default by control file).
Database The database's name with owner(default owner is SYSTEM).
Source The record file or batch file.
Control The controll file for process.

The options include followings:

help Display help.
t Throught mode(server accesses the Source directly).
b Batch mode(Source is a batch file).
r Output record IDs file.
w Output wrong data file.

19.2.1 指定服务器的端口

数据库加载工具运行时需要与服务器进行连接，所以需要知道服务器启动时所使用的网络端口号。命令行中的 ‘-P:Port’ 参数就是服务器管理员用来指定连接服务器的端口号的。其中：‘-P:’ 是一个引导符，其后紧跟的是服务器端口号。如果该参数省略，则数据库加载工具默认为服务器启动时的端口号。在一般情况下，tdbload 可以从与程序所在目录平级的 system 子目录中的 START.INF 文件中自动获取服务器启动所使用的端口号，而无须特别指定。

19.2.2 指定源数据的字符集

数据库加载工具允许源数据的字符集与目标数据库的字符集不一致，但这时需要

使用命令参数 ‘-CS:Charset’ 来指定源数据的字符集。其中：‘-CS:’ 后紧跟的是源数据的字符集。如果该参数省略，则数据库加载工具默认源数据的字符集为目标数据库的字符集。

19.2.3 命令开关选项

命令行中的 ‘-Options’ 参数可以为数据库加载工具设置开关选项。该工具包括 4 个主要的开关选项，即：b、t、r、w。这些选项既可单独书写（如：-b、-t、-r、-w），也可集中书写（如：-btrw）。

批处理模式开关 ‘b’。如果设置该选项，则表示命令行中 ‘SourceFile’ 所对应的文件是一个存放所有原始数据文件名列表的文件，否则 ‘SourceFile’ 对应一个原始数据文件。

直通模式开关 ‘t’。这个参数与控制文件中直通模式变量（THROUGHT_ACCESS）是完全相同的，这里只是为在没有指定控制文件的情况下提供一种简洁的指定加载模式的方法。如果该选项未设置，控制文件中也未设置相应变量，则以常规方式进行数据加载，即由 tdbload 读取原始数据文件，与一般的客户端一样通过 RPC 通讯机制将数据传递给服务器，这时只要求 tdbload 对数据加载所涉及到所有原始数据文件、批处理文件、控制文件有读取权限。

记录号输出开关 ‘r’。如果设置该选项，则要求命令输出原始数据记录加载到数据库中以后所对应的记录号的列表文件，否则命令不输出该文件。如果目标数据库实际对应一个自动分裂模式视图，则忽略该开关，不输出该文件。

错误记录日志文件输出开关 ‘w’。如果设置该选项，则要求命令输出因数据存在错误而未加载到数据库中的原始数据记录的日志文件，否则命令不输出该文件。

19.2.4 指定目标数据库

命令行中的 ‘Database’ 用以指定数据加载时的目标数据库名，其书写格式为 ‘Owner.Name’ 或 ‘Name’。如果没有指定数据库的拥有者，即以 ‘Name’ 格式给出，则默认数据库的拥有者为 SYSTEM。

19.2.5 指定源数据文件

命令行中的 ‘SourceFile’ 用以指定数据加载时的源数据文件。请参见 19.2.3 节中有关批处理模式开关的说明。

19.2.6 指定控制文件

命令行中的‘ControlFile’用以指定数据加载时的控制文件。控制文件可以省略。请参见《数据库加载》部分的第 15 章。

如果控制文件只需简单地指定源数据文件的记录格式，则可不用控制文件，而在命令中可以直接给出格式名，方法是将命令中原控制文件的位置输入‘-F:Name’代替，其中的‘Name’就是要指定的记录格式。请参见《数据库加载》部分的 15.2.1 的说明。【修订 090】

第 20 章 配置文件

20.1 运行参数配置文件

文件 ‘PROFILE.INI’ 是服务器运行参数配置文件，处在服务器主目录（安装目录）的 system 子目录下。该文件的格式是一种比较通用的配置文件格式。其一般形式如下：

```
# this is comment line.
```

```
[SectionName1]
```

```
ItemName1=ItemValue1
```

```
ItemName2=ItemValue2
```

```
.
```

```
.
```

```
.
```

```
ItemNameN=ItemValueN
```

```
[SectionName2]
```

```
ItemName1=ItemValue1
```

```
ItemName2=ItemValue2
```

```
.
```

```
.
```

```
.
```

```
ItemNameN=ItemValueN
```

```
.
```

```
.
```

```
.
```

```
[SectionNameN]
```

```
.
```

```
.
```

```
.
```

- 以 ‘#’ 开头的行是注释行。允许空白行。
- [SectionNameX]表示一个配置域的开始，其中 SectionNameX 是域名。

- 一个配置域可以包括 0 个或多个配置项。
- 配置项的格式是：项名=项值。
- 一个配置项必须在一行内完成。

在 PROFILE.INI 文件中，目前支持八种配置域，分别是：IndexCache、Maintain、Retrieval、Network、Concurrent、AutoView、AutoBatch 和 AutoReboot。对配置文件的任何修改，必须在重新启动服务器以后才有效。

20.1.1 配置域IndexCache

这个配置域用来配置服务器检索缓冲模块的运行参数。其配置项包括：

- 配置项 ‘LowLimit=number’。‘number’ 是一个十进制数字，表示数据库的低频词的词频的上限，无配置时的缺省值是 100000。小于该词频的词不参与缓冲。如果其值大于或等于 MediumLimit 的值，则将被自动调整为 MediumLimit 的二分之一与 100000 之间的最小值。
- 配置项 ‘MediumLimit=number’。‘number’ 是一个十进制数字，表示数据库的中频词的词频的下限，无配置时的缺省值是 400000。小于该词频的所有词只参与一级缓冲（内存），不参与二级缓冲（文件）。如果其值大于或等于 HighLimit 的值，则将被自动调整为 HighLimit 的二分之一与 400000 之间的最小值。
- 配置项 ‘HighLimit=number’。‘number’ 是一个十进制数字，表示数据库的高频词的词频的下限，无配置时的缺省值是 1600000。小于该词频的词既参与一级缓冲，又参与二级缓冲，处于二级缓冲中的内容还能再次被调入一级缓冲；大于该词频的词属于超高频词，既参与一级缓冲，又参与二级缓冲，但一旦进入二级缓冲，则不会再次被调入一级缓冲。二级缓冲使用磁盘文件，即二级缓冲文件（参见配置项 DiskSize）。如果 MemorySize 的值小于 1024，则其最小值为 800000。
- 配置项 ‘MemorySize=number’。‘number’ 是一个十进制数字，表示一级缓冲能够使用的最大内存量，以兆字节计算。无配置时缺省将自动检测机器的物理内存量，检测失败时，则缺省取 TRS 全文数据库系统配置数据词典，请参见 20.1 节中的有关说明。因为有的 32 位操作系统在地址空间上有一定限制，所以在 32 位操作系统上，MemorySize 的最大取值为 1024。
- 配置项 ‘DiskSize=number’。‘number’ 是一个十进制数字，表示二级缓冲文件能够使用的最大磁盘空间，以兆字节计算。如果没有配置或者设置成 0，则表示不进行二级缓冲，当一级缓冲满时，将直接全部清空。如果取值为-1，则表示根据磁盘余空间自动调节。二级缓冲文件存放在系统配置的“临时文件存放路径”的 SWAP 子目录下（参见 5.12 节）。
- 配置项 ‘MemoryFS=number’。‘number’ 是一个十进制数字，表示内存文件系统能够使用的最大物理内存空间，以兆字节计算。如果设置成 0，或不设置，则表示不使用内存文件系统；如果设置成-1，则表示将自动检测并自动调节内存文件

系统的物理内存使用量。在 ‘number’ 后还可以指定内存文件系统创建的方式，格式是 ‘number@device’，如 ‘1024@/shm/trsmfs’ 表示在文件系统（unix/linux）‘/shm/’ 设备下以 ‘trsmfs’ 文件创建 1G 容量的内存文件系统；在 Windows 系统下可以用 WAE 的方式创建内存文件系统（device 是 WAE，这也是 Windows 下的缺省创建方式），如 ‘1024@WAE’；还可以在常规内存中创建内存文件系统（device 是 RAM），如 ‘1024@RAM’，这时内存文件系统将占用进程的地址空间。内存文件系统目前主要用于检索结果集的缓冲以实现延迟（异步）写检索历史文件。

- 配置项 ‘WordLoad=switch’。‘switch’ 是一个开关，取 ‘ASY’、‘SYN’ 和 ‘OFF’ 值之一，无配置时缺省为 ‘OFF’。‘ASY’ 表示系统在空闲的时候加载已经访问过的词的索引，‘SYN’ 表示在查找词索引的缓冲时，如果词的索引还未进缓冲，则立即加载到缓冲，‘OFF’ 表示不缓冲任何词的索引。
- 配置项 ‘Expression=switch’。‘switch’ 是一个开关，取 ‘ANY’、‘TOP’ 和 ‘OFF’ 值之一，无配置时缺省为 ‘ANY’。‘ANY’ 表示按系统内部既定的策略缓冲各级表达式的检索结果，‘TOP’ 则表示只允许顶层表达式的检索结果，‘OFF’ 表示不缓冲任何表达式的结果。
- 配置项 ‘Preload=DatabaseName(ColumnName,ColumnName...)’，表示服务器检索缓冲模块启动时需预加载的索引。字段省略时表示只预加载该数据库的缺省全文检索字段的索引。一行只能指定一个数据库，如果需要预加载多个数据库的索引，则可用多行分别说明这些数据库。^{【修订 053】}如果 ColumnName 是 “DATA”，则表示预加载数据库的“允许缓存”的字段值和“允许相关排序”的文本字段的规范长度。这里 DatabaseName 也可以对应一个视图，表示对视图中的每个数据库进行相应的预加载。^{【9】}

20.1.2 配置域 Maintain

这个配置域用来配置服务器在进行数据库维护时的运行参数。其配置项包括：

- 配置项 ‘BTCacheEnable=switch’。‘switch’ 是一个开关，取 ‘TRUE’、‘FALSE’ 和 ‘WRITE’ 值之一，无配置时的缺省值为 ‘WRITE’。‘TRUE’ 表示服务器进行索引维护时，允许将 B⁺树索引的读写操作缓冲在内存，‘FALSE’ 则表示不允许 B⁺树索引缓冲在内存，‘WRITE’ 表示服务器进行索引维护时，只允许将 B⁺树索引的写操作缓冲在内存，这样可以避免因 B⁺树索引与倒排索引的更新不同步而导致检索操作的失败（比如读倒排索引文件失败等）。对于慢磁盘系统上的大数据库，缓冲 B⁺树索引虽然会大大提高索引维护的速度，但由于需要较多的内存空间，所以系统的并发性将受到影响。该配置项只有在大于 512 兆内存的机器上才有效，并且对于高速磁盘系统上小数据库，效果也不明显。
- 配置项 ‘IDXAsyncLimit=number’。‘number’ 是一个十进制数字，表示允许缓存异步关闭的索引描述符的最大个数，如配置为 1024，即允许缓存异步关闭的索引

描述符数最大为 1024。“异步关闭索引描述符”的功能只在 Windows 平台起作用，用以提高索引维护的性能，但在 Windows 系统被异常关闭时（如断电），如果仍有“异步关闭索引描述符”未关闭，则有可能丢失索引数据。如果取 0 值（无配置时的缺省值），则表示禁止“异步关闭索引描述符”这项功能。【修订 026】

- 配置项 ‘IDXAsyncDelay=number’。‘number’ 是一个十进制数字，表示一个索引描述符放入“异步关闭”缓存之后，至少要过多长时间（秒）没有被重用的情况下允许被异步关闭。无配置时缺省为 300 秒，最大值是 3600 秒。【修订 028】
- 配置项 ‘IDXAsyncIdled=number’。‘number’ 是一个十进制数字，表示对“异步关闭”缓存在多长时间（秒）内没有进行任何访问的情况下就可认为“异步关闭索引描述符”功能已经进入空闲状态，此时可关闭所有缓存的索引描述符。如果取值为 0，则表示无空闲状态。无配置时缺省为 30 秒，最大值是 3600 秒。【修订 030】
- 配置项 ‘FileSyncFlush=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘FALSE’。取值 ‘TRUE’ 表示在关闭已经修改过的数据库文件（包括数据和索引文件）之前，需要将文件系统缓存中的数据同步写入磁盘；取值 ‘FALSE’ 表示不主动同步写盘，而是由文件系统内部的缓冲机制决定何时同步磁盘数据，以提高性能，但如果操作系统在完成磁盘数据同步之前被异常关闭（如断电），则有可能丢失数据。如果只需同步数据文件，而不同步索引文件，则可通过配置上述“异步关闭索引描述符”的功能实现：只要配置项 IDXAsyncLimit 或 IDXAsyncDelay 之一不为 0 即可。【修订 032】
- 配置项 ‘BitsMapFilter=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时缺省为 ‘FALSE’。取值 ‘TRUE’ 表示在优化数据库时创建和使用位图来标识已经被删除的记录，并以此来完成索引的过滤。如果一个数据库删除了大量的记录（比如删除几百万），位图可能会大大提高优化的速度。【修订 037】
- 配置项 ‘IdleTaskRules=rules’。‘rules’ 是用来描述系统在空闲状态下的自动维护任务的一个或多个规则串。规则串规定系统在空闲状态下：需要执行哪些自动维护任务；一次自动维护任务需要维护哪些数据库；这个自动维护任务在哪天执行。一个规则串的格式是：NAME%DATABASE_MOD@TIME%TIME_MOD。规则串之间用单字节的分号或逗号分隔。其中：
 1. NAME 是任务名。系统支持下列任务：optimax、optimin 和 index。optimax 任务是完整地优化数据库（包括所有字段的索引和数据库的数据）；optimin 任务是部分地优化数据库（只优化 DATE、NUMBER 和 CHAR 三种类型的所有字段的索引）；index 任务是对已有部分记录建立索引的字段创建其它所有记录的索引（没有任何记录建立索引的字段不在其列）。
 2. DATABASE_MOD 是任务将数据库分组后的组数。系统内的所有数据库通过对数据的 ID 用 DATABASE_MOD 取模进行分组，一次自动维护任务只维护其中的一组，所有的组被轮流维护。如果 DATABASE_MOD 的值为 0，则表示这个自动维护任务被禁止。
 3. TIME 是自动维护任务的执行时间。是星期几，或是一个周期内的某一天，

具体意义取决于后面 TIME_MOD 的值。

4. TIME_MOD 是自动维护任务的执行时间的模式。如果是 0，则表示前面的 TIME 是指星期几（从 1 到 7）；否则 TIME_MOD 表示一个周期，它把一年内所有的“天”分成许多时间段，这时 TIME 表示每个周期内的某一天（从 0 到 TIME_MOD - 1）。
- 配置项 ‘IdleTimeUsers=number’。‘number’ 是一个十进制数字，表示当时间到达系统配置的空闲状态起始时间时，允许系统进入“空闲状态”执行自动维护任务时的最大活动用户实例数。-1 表示只要系统总是在空闲状态起始时间立即无条件进入“空闲状态”执行自动维护任务。
 - 配置项 ‘IdleTaskAbort=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘FALSE’。‘TRUE’ 表示服务器进入空闲状态并正在执行一个自动维护任务时，如果此时的系统时间已经到达空闲状态终止时间，或者已经收到外部请求，则终止当前的自动维护任务，服务器立即退出空闲状态而全力响应外部请求。‘FALSE’ 则表示服务器按正常的方式执行自动维护任务。请参见 5.7 节中的有关说明。
 - 配置项 ‘IndexRuleWord=switch’。‘switch’ 是一个开关，取 ‘TRUE’、‘FALSE’ 和 ‘DOT’ 值之一，无配置时缺省为 ‘FALSE’。服务器分词系统支持用 ‘#&/-.’ 字符之一连接而成的规则串（如 “AT&T”）。‘TRUE’ 值表示将这些规则串作为一个整体（如 “AT&T”）进行索引和检索；‘FALSE’ 值表示将规则串自动拆分成多个独立的“词”（如 “AT”、“&” 和 “T”）进行索引和检索；^[修订 014]而 ‘DOT’ 值则表示将规则串自动拆分成多个独立的“词”，但由 ‘.’ 连接的串仍然作为一个整体进行索引和检索（如规则串 “service.trs.com.cn/portal/default-index.jsp” 在此种情况下将自动拆分成：“service.trs.com.cn”、“/”、“portal”、“/”、“default”、“-” 以及 “index.jsp”），并且认为 “://” 之后的规则串是一个 URL，对其中的网络域名将冗余索引剥离最小子域后的上一级域名（如对域名 “service.trs.com.cn” 将冗余索引 “trs.com.cn”）^[9]。该配置项的修改只对以后新建的数据库有效，新建的数据库将记住当前设置，而不受以后对该配置项修改的影响。
 - 配置项 ‘IndexCompress=switch’。‘switch’ 是一个开关，取 ‘ANY’、‘MAX’ 以及 ‘OFF’ 值之一，无配置时表示其值取决于数据库的“禁止压缩开关”属性设置：如果禁止压缩，则 IndexCompress 取值为 ‘OFF’，否则 IndexCompress 取值为 ‘ANY’。该配置项只对新创建的数据库（包括自动分裂的数据库），或者重建字段的索引时起作用：取值 ‘ANY’ 表示按系统内部确定的方式压缩索引；取值 ‘MAX’ 表示按具有最大压缩比的方式对索引进行压缩；取值 ‘OFF’ 表示对索引不进行压缩。
 - 配置项 ‘LastPathPrior=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘TRUE’。该配置项只对新创建的数据库（包括自动分裂的数据库）没有指定明确的数据存储路径时起作用：取值 ‘FALSE’ 表示

取系统配置的路径集中的第一个路径作为数据库的路径；取值‘TRUE’表示取系统配置的路径集中除 ibase 和 dict 之外的（即新增加的）最后一个路径作为数据库的路径。

- 配置项 ‘FileIdleReuse=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘TRUE’。取值 ‘TRUE’ 表示回收并重用在维护数据库的数据或索引时产生的空闲文件块，以节省空间；取值 ‘FALSE’ 表示忽略并丢弃在维护数据库的数据或索引时产生的空闲文件块，以节省空闲块的管理时间。
- 配置项 ‘NiceSleepTime=number’。‘number’ 是一个十进制数字，表示“睡眠”的秒数。当对数据库的记录或索引进行维护操作时，系统会自动降低当前过程的调度优先级，以便于更快地执行其它的数据库检索过程。但有时降低调度优先级不足以达到预期效果，还需要维护操作过程进行检测：如果在某个检测点发现有数据库检索请求在执行，则强制当前维护过程在该检测点“睡眠”由这里 NiceSleepTime 指定的秒数。最长有效“睡眠”时间为 300 秒；无配置时的缺省值为 0，表示不“睡眠”。【修订 003】
- 配置项 ‘RelevanceNorm=mode’。‘mode’ 的取值为 ‘QURT’、‘SQRT’、‘LOG’ 和 ‘COS’ 之一，无配置时的缺省值为 ‘QURT’。‘QURT’ 表示用所有词的词频和的四次方根作为该字段相关性规范长度；‘SQRT’ 表示用所有词的词频和的平方根作为该字段相关性规范长度；‘LOG’ 表示用所有词的词频和的对数作为该字段相关性规范长度；【修订 003】‘COS’ 表示用标准余弦公式计算相关度，余弦公式能够体现词之间的词频分布对相关度的影响，其规范长度是所有词的词频平方和的平方根^[4]。规范长度的实际作用是计算命中点的“平均密度”，由于四次方根、对数和平方根对数值的抑制强度是递减的，所以就“字段长度”对相关性的影响程度来说，四次方根方法的影响最小（也就是说对“字段长度”最不敏感），而平方根方法的影响最大（也就是说对“字段长度”最敏感）。【修订 041】
- 配置项 ‘OEMFilterMute=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘FALSE’。取值 ‘TRUE’ 表示如果在数据入库调用 OEM 文档过滤器抽取文本时失败，则忽略错误，按“无文本”处理（不影响其它数据的正常入库与索引）；取值 ‘FALSE’ 表示不忽略 OEM 文档过滤器的错误，当前记录将不能入库。【修订 044】
- 配置项 ‘IDXStoreNorms=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时缺省为 ‘FALSE’。取值 ‘TRUE’ 表示冗余存储字段的相关性规范长度；取值 ‘FALSE’ 表示不冗余存储字段的相关性规范长度。冗余存储字段的相关性规范长度的好处是可以加快相关度的计算，但会增大数据库的存储空间（一般会增大 11% 左右），精度也会有所损失。该配置项改变后，只对那些新建库，或者删除后重建的索引起作用。【修订 130】
- 配置项 ‘IDXStoreFreqs=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时缺省为 ‘FALSE’。取值 ‘TRUE’ 表示冗余存储索引词的词频；

取值 ‘FALSE’ 表示不冗余存储词频。冗余存储词频的好处是可以避免检索时的词频统计而加快相关度的计算，但会增大数据库的存储空间（一般会增大 11% 左右），词频精度也会有所损失。该配置项改变后，只对那些新建库，或者删除后重建的索引起作用。【修订 131】

- 配置项 ‘SplitAlphaNum=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，缺省值为 ‘FALSE’。取值 ‘TRUE’ 表示对于文本中的字母和数字不粘连，并拆分带小数点的数值串，将各部分单独索引（如：字符串 “TRS6.5” 将拆分成 “TRS”、“6”、“.” 和 “5” 四项独立的索引，因此这种模式对检索性能会有一些的影响）；取值 ‘FALSE’ 表示按传统方式处理：允许字母数字粘连成一个整体，带小数点的数值串整体索引（如：字符串 “TRS6” 和 “6.5” 将都是一个整体索引）。一个数据库被创建时，它将记住该配置项，配置项以后的修改不会影响已经创建的数据库。【修订 135】
- 配置项 ‘ParallelIndex=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，缺省值为 ‘TRUE’。取值 ‘TRUE’ 表示创建数据的索引（涉及批量记录加载、记录复制和重建索引三种操作）时，使用并行处理方式，提高系统性能；取值 ‘FALSE’ 时，表示使用老的常规方式创建索引。【修订 156】

20.1.3 配置域 Retrieval

这个配置域用来配置服务器在进行数据库检索时的运行参数。其配置项包括：

- 配置项 ‘ParallelLimit=number’。‘number’ 是一个十进制数字，表示服务器进行并行检索时每个检索过程最多能够并发的线程数。当取值为 ‘0’ 或 ‘1’ 时，表示禁止并行检索功能；取值为 ‘-1’ 时，则表示通过自动检测系统 CPU 的个数来决定。无配置时的缺省值为 0。
- 配置项 ‘RelevanceMode=mode’。‘mode’ 表示对检索结果进行相关性排序的模式，取值为 ‘NORMAL’、‘WORD_V1’、‘WORD_V2’、‘POS_COE’、‘NO_IDF’ 以及 ‘HIT_TF’ 和 ‘LENGTH’ 七者之一。
 1. NORMAL 是无配置时的缺省值，表示相关性为传统意义的相关性。
 2. WORD_V1 表示用一个记录中命中词的单位向量长度（每个命中词只计算一次）作为记录的相关性。
 3. WORD_V2 表示用一个记录中命中词的单位向量长度，以及命中词的词频之和作为记录的相关性，即当多条记录命中词的单位向量长度相等时，这些记录将再按命中词在命中点出现的次数总和的降序排列。
 4. POS_COE 表示传统意义的相关度乘以位置加权系数（位置越靠前，加权系数越大）后的综合值计算记录的相关性。
 5. NO_IDF 表示按传统意义计算相关性时，不考虑词的 IDF 因素，也就是认为所有的词在所有的数据库中都是同等重要的，以避免因一个词在不同数据库中 IDF 的差别而导致相同的数据在跨数据库排序时的一致。【修订 039】

6. HIT_TF 表示简单地以命中点所包含的所有词的在命中点中出现次数的总和作为相关度, 计算记录的相关性。【修订 149】
 7. LENGTH 表示简单地以字段的归一化长度(因子)的倒数作为相关度, 计算记录的相关性。【修订 150】
- 配置项 ‘SortAlwaysMax=number’。‘number’ 是一个十进制数字, 表示在检索时指定“总是允许排序”的情况下, 允许排序的检索结果的最大命中记录数的最高限值。如果取值为 ‘-1’, 则表示不对排序的检索结果记录数进行限制。当取值为 ‘0’ 时, 或者小于 TRS 全文数据库系统配置数据词典的“最大排序记录数”设置时, 则取值无效。无配置时的缺省值为 0。请参见 5.5 节中的有关说明。支持这个配置项的根本原因是在必要时能够突破“最大排序记录数”的最大值只能是 65533 的限制, 但必须要在检索时设置“总是允许排序”选项, 该配置项才会起作用。
 - 配置项 ‘ShowWordLimit=number’。‘number’ 是一个十进制数字, 表示对数据库字段上的索引项(字、词或字段值)进行浏览时, 允许匹配并输出的索引项的最大个数, 如果实际匹配的索引项数超过此限制, 系统将只输出此限制数目的索引项。当取值为 ‘0’ 或 ‘-1’ 时, 表示没有限制。无配置时的缺省值为 100000。
 - 配置项 ‘HFIgnoreLimit=number’。‘number’ 是一个十进制数字, 表示在检索并指定“忽略高频词”选项时, 一个字、词项或字段值在数据库中出现的记录数(即命中数)与总记录数的百分比必须小于这个值, 否则将被忽略, 不参与运算, 也就是说: 对于“与”类运算, 相当于所有的记录中都存在这个“高频词”; 对于“或”和“异或”类运算, 相当于任何记录中都不存在这个“高频词”; 对于“非”类运算, 如果“高频词”是运算的左操作数(前者), 则相当于所有的记录中都有这个词, 如果“高频词”是运算的右操作数(后者), 则相当于任何记录中都不存在这个词。当取值为 ‘0’ 或 ‘-1’ 时, 表示不忽略任何“高频词”, 最大为 100。无配置时的缺省值为 0。
 - 配置项 ‘MaxVirtualSet=number’。‘number’ 是一个十进制数字, 表示检索历史的最大虚拟步数, 当检索历史的个数超过这个数时, 将淘汰最早产生的虚拟历史, 以存储新的虚拟历史。虚拟历史的作用是: 只要一个检索历史在仍在虚拟历史中, 即使实际的“物理”检索历史已被清除, 则当引用该检索历史时, 系统能够自动恢复原来的检索结果集, 并生成一个新的“物理”检索历史。需要注意的是: 检索历史是否存储到虚拟历史中, 不是默认的, 需要在检索时明确指定。
 - 配置项 ‘HitPointExact=switch’。‘switch’ 是一个开关, 取 ‘TRUE’、‘FALSE’ 和 ‘NONE’ 值之一, 无配置时的缺省值为 ‘TRUE’。取值 ‘TRUE’ 表示存储实际命中点的准确位置; 取值 ‘FALSE’ 表示只存储其它必要的信息以便能够恢复非精确的命中点; 取值 ‘NONE’ 表示不存储任何命中点信息, 则意味着不能恢复命中, 也不支持相关性排序。【修订 022】
 - 配置项 ‘IgnorePunct=switch’。‘switch’ 是一个开关, 取 ‘TRUE’ 和 ‘FALSE’ 值之一, 缺省值为 ‘TRUE’。取值 ‘TRUE’ 表示如果在全文字段的检索“短语”

中包含有标点符号（【修订 039】除 ‘\$’、‘%’ 之外的非数字/字母的 ASCII 字符和对应的全角字符，以及中文里的标点符号。以前的配置项是 `IgnoreSenPunct`，只对“句号”、“逗号”、“分号”、“问号”、“感叹号”起作用^[4]），则忽略，不参与检索运算；取值 ‘FALSE’ 表示标点符号需要参与运算，进行精确匹配。通常地，如果数据库使用了停用词典，且停用词典中包含了这些标点符号，则检索时总是会忽略它们的。忽略标点符号可以提高检索性能。【修订 046】

- 配置项 ‘`ClassStatLimit=number`’。‘`number`’ 是一个十进制数字，表示对字段值进行分类统计时能够处理的最大类别数。最大值为 65535，0 也表示取最大值。无配置时缺省取值为 5120。在应用调用分类统计功能时，还可以在参数里指定最终存储的有效类别数，缺省存储的有效类别数也是这里设置的值。【修订 054】
- 配置项 ‘`ExpressHeight=number`’。‘`number`’ 是一个十进制数字，表示对检索表达式进行分析形成运算树时，运算树所允许的最大高度。0 表示不限制，无配置时缺省为 5120。如果一个运算树太高，在递归调用时有可能因堆栈溢出而导致系统异常。在这种情况下，要么限制运算树的高度，要么扩大堆栈的尺寸。【修订 056】
- 配置项 ‘`SortValidSize=number`’。‘`number`’ 是一个十进制数字，表示对检索结果进行排序时保证有序的记录数，即进行 TOP N 部分排序。无配置时的缺省值为 0，表示不进行 TOP N 部分排序。排序后的结果集只有前 N 个记录是有序的（指定的序），其余的记录则是无序的。在应用调用排序的相关功能时，还可以在参数里指定 TOP N 部分排序，以替代该配置。【修订 058】
- 配置项 ‘`SortPruneSize=number`’。‘`number`’ 是一个十进制数字，表示对检索结果进行排序时保证有序的记录数，并且在检索结果集中只存储了这些记录，即对检索结果进行 TOP N 剪裁排序。无配置时的缺省值为 0，表示不进行 TOP N 剪裁排序。在应用调用排序的相关功能时，还可以在参数里指定 TOP N 剪裁排序，以替代该配置。TOP N 剪裁排序总是优先于 TOP N 部分排序，即如果前者有效，则后者就无效。【修订 060】
- 配置项 ‘`PrecacheFiles=DatabaseName(ColumnName,ColumnName...)`’，表示服务器启动时需预读的数据库索引或数据文件。预读的目的只是为了尽早使文件系统的缓存机制发生作用，加快重启操作系统后“第一次”检索的速度。字段省略时表示只预读该数据库的缺省全文检索字段的索引。一行只能指定一个数据库，如果需要预读多个数据库的索引或数据文件，则可用多行分别说明这些数据库。如果 `ColumnName` 是“DATA”，则表示预读数据库的数据文件，如果存在集中存储数据，则只预读集中存储文件。这里 `DatabaseName` 也可以对应一个视图，表示对视图中的每个数据库进行相应的预加载。【修订 065】
- 配置项 ‘`VirtualResult=mode`’。‘`mode`’ 表示检索时对单引号引起的空串和分词后没有词项的短语（如检索的“短语”为停用词）的结果处理模式，共有五种模式：‘ERROR’、‘DEFAULT’、‘BLANK’、‘WHOLE’ 以及 ‘SMART’，无配置时缺省值为 ‘DEFAULT’。‘ERROR’ 表示对于对单引号引起的空串进行报错，即不允许出现单引号中无内容的情况，而分词后没有词项的短语则按无结果的方

式处理；‘DEFAULT’表示兼容老的处理模式，即单引号引起来的空串等价于全模糊串，而分词后没有词项的短语则没有检索结果；^{【修订 114】}‘BLANK’模式表示单引号引起来的空串和分词后没有词项的短语都按无结果的方式处理^{【1】}；‘WHOLE’模式表示单引号引起来的空串和分词后没有任何词项的短语都等价于一个全模糊串；‘SMART’模式表示内部进行更合理的处理，根据参与的运算，按不同的方式忽略单引号引起来的空串和分词后没有词项的短语。^{【修订 094】}

- 配置项 ‘DummyColPrior=mode’。‘mode’表示检索时，虚拟逻辑字段的运算展开模式，可取 ‘TRUE’、‘FALSE’和 ‘DEFINED’三者之一。‘DEFINED’表示虚拟逻辑的展开模式由字段的数据词典定义（参见 1.4 节的说明）；‘TRUE’表示忽略字段的数据词典定义，虚拟逻辑总是按“字段逻辑优先”的模式展开，相当于所有的虚拟逻辑字段都指定了运算优先模式符号 ‘&’；‘FALSE’表示忽略字段的数据词典定义，虚拟逻辑总是按“表达式逻辑优先”的模式展开，相当于所有的虚拟逻辑字段都没有指定了运算优先模式符号 ‘&’。^{【修订 096】}
- 配置项 ‘MaxFuzzyMatch=number’。‘number’是一个十进制数字，表示在全文字段（PHRASE 和 DOCUMENT 字段）上进行模糊检索时，分词并组合后的单个内部模糊词所允许匹配的最大索引项数。取值范围是 0~65535，无配置时的缺省值为 0，表示没有限制。^{【修订 134】}
- 配置项 ‘ExtraWeighted=switch’。‘switch’是一个开关，取 ‘TRUE’和 ‘FALSE’值之一，缺省值为 ‘FALSE’。取值 ‘TRUE’表示允许相关性计算考虑额外的加权，考虑命中分支等对相关性的影响，以期优化排序效果。取值 ‘FALSE’表示相关性计算只考虑传统的加权。^{【修订 136】}
- 配置项 ‘FocusWeighted=parameters’。‘parameters’是一组参数，用于按焦点排序（参见 9.1.2.4 节）时结果记录的权重计算。按焦点排序是将记录分段，分别对应：当天的记录（第一段）、24 小时之内的记录（第二段）、一周之内的记录（第三段）、一月之内的记录（第四段）、一年之内的记录（第五段）及其它记录。这组参数包含六个值，分别是：段之间的权重折扣率、第一段的记录数、第二段的记录数、第三段的记录数、第四段的记录数、第五段的记录数。参数之间要用空格分隔，并且顺序是固定的。假设定义：段之间的权重折扣率为 0.5、第一段的记录数为 10000（权重确定为 1.0）、第二段的记录数为 100000（按上述定义的 0.5 折扣率，权重为 0.5）、第三段的记录数为 500000（权重为 0.25）、第四段的记录数为 2000000（权重为 0.125）、第五段的记录数为 10000000（权重为 0.0625，其它记录的权重为 0.03125），那么该配置项应该定义成：FocusWeighted=0.5 10000 100000 500000 2000000 10000000（缺省值）。^{【修订 142】}

20.1.4 配置域 Network

这个配置域用来配置有关服务器网络通讯的参数。其配置项包括：

- 配置项 ‘ConnectCache=switch’。‘switch’是 ‘TRUE’和 ‘FALSE’之一。‘TRUE’

表示并发用户数较少时，服务器将缓冲与客户端的连接；‘FALSE’是无配置时的缺省值，表示服务器执行完成一个服务请求后立即断开与客户端的连接。缓冲“连接”的优点是：并发用户数较少时，加快客户端访问服务器的速度；缺点是：当并发用户数较多时，服务器端的 TCP 网络层会存在一定数量的资源较长时间地处在 FIN_WAIT_2 状态。不缓冲“连接”的优点是：TCP 网络层没有或很少有资源处在 FIN_WAIT_2 状态；缺点是：当访问频繁时，TCP 网络层会存在一定数量的资源处在 TIME_WAIT 状态（每个 TIME_WAIT 状态大约持续 2 分钟，在操作系统层次也可以设置），当然通过 TimeWaitMust 配置项可以消除这一状态。

- 配置项 ‘TimeWaitMust=value’。‘value’ 是 ‘TRUE’、‘REMOTE’ 和 ‘FALSE’ 三者之一。
 1. TRUE 表示断开服务器与客户端的网络连接时，TCP 网络层将按正常程序进入 TIME_WAIT 状态。
 2. REMOTE 表示对于来自本地（LOCALHOST，即 127.0.0.1，或服务器所在的 IP 地址）的网络连接，断开时允许 TCP 网络层忽略 TIME_WAIT 状态，立即释放资源；而对于来自远程的网络连接，断开时 TCP 网络层则将按正常程序进入 TIME_WAIT 状态。
 3. FALSE 是无配置时的缺省值，表示断开服务器与客户端的网络连接时，TCP 网络层对任何连接都直接跳过 TIME_WAIT 状态，立即释放资源。

严格地讲，无 TIME_WAIT 状态时，TCP 网络层的可靠性存在隐患，尤其在非局域网（如广域网）环境中，在这样的环境中一般不建议使用。

- 配置项 ‘ConnectUserIP= list’。‘list’ 是 IP 地址模板集，每个模板之间用分号分隔。这个地址模板集的作用是对系统中只有“登录”权限的用户，即用户类型为 CONNECT 的用户在登录系统时进行 IP 地址限制：只有通过与模板集中某个模板相匹配的 IP 地址才能登录到系统中（但本地地址总是能登录）。IP 地址模板的格式是：X.X.X.X。其中 X 表示可以是 0 到 255 之间的任意值；而如果某个 X 用一个 0 到 255 之间的任意值替代（如“192.X.X.X”），则表示 IP 地址的那个字节必须是这个指定的值（如第一个字节必须是 192）。
- 配置项 ‘WideAreaLogin=switch’。‘switch’ 是 ‘TRUE’ 和 ‘FALSE’ 之一。‘FALSE’ 是无配置时的缺省值，表示系统在验证后续请求的合法性时要求客户端 IP 地址保持不变，否则拒绝服务；‘TRUE’ 表示把系统登录请求都看成是来自广域网的客户端，系统在验证后续请求的合法性时不再要求客户端 IP 地址保持一致性（广域网客户端的 IP 地址在不同的网络连接时有可能不一样），以避免请求被拒绝。对于运行在广域网的客户端软件，也可以在连接服务器时通过调用参数给定本地 IP 地址，从而避免请求被拒绝，这时就不需要服务器修改缺省配置。^{【修订 069】}

20.1.5 配置域 Concurrent

这个配置域用来配置服务器在响应服务请求时的并发参数。其配置项包括：

- 配置项 ‘RoutineLimit=number’。‘number’ 是一个整数值，表示系统运行时允许并发的最大子进程/线程数。如果是 0，或者大于系统限制，则该值无效。无配置时的缺省值是-1，表示根据系统的位数和物理内存总数自动调节。

20.1.6 配置域AutoView

这个配置域用来配置自动模式视图分裂数据库的有关参数。其配置项包括：

- 配置项 ‘SplitLimit=number’。‘number’ 是一个整数值，表示当一个自动模式视图中的最后一个数据库的记录数不小于该值，并且需要向该自动模式视图中加载、插入或复制新的记录时，自动分裂出一个新的数据库以存放新的记录，也就是说分裂只会发生在向自动分裂视图中加载数据之时，既不是在向被分裂的数据库中加载数据之时，也不是对已加载的数据进行分裂存储。取值范围从 10000 至 20000000（小于 10000 时取 10000，大于 20000000 时取 20000000），无配置时的缺省值为 1000000。^{【修订 051】}每个自动模式视图也可单独设置分裂阈值（参见 1.5 节的说明）。^{【4】}
- 配置项 ‘InheritPath=switch’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时缺省值为 ‘TRUE’。该配置用来控制分裂出来的新数据库的路径：如果取值 ‘TRUE’，则表示继承原数据库的路径；如果取值 ‘FALSE’，则表示新数据库的路径取系统配置中的缺省路径。参见 Maintain 配置域的 LastPathPrior 配置项。

20.1.7 配置域AutoBatch

这个配置域用来配置：在启动 trsserver 时需自动运行的其它模块程序，如 trssearcher 和 trsmirror，以及 tdbshutdown 关闭服务器时需同时运行的其它模块的关闭程序，如 trsmirror 的关闭程序。其配置项有两种类型：ProgramStart 和 ProgramClose，每种配置项可以多次出现。

- 配置项 ‘ProgramStart=command’。‘command’ 是一个字符串，包括：程序名、程序参数以及程序声明，表示在启动 trsserver 之后，trsserver 将自动加载并运行该程序，程序的路径是相对于 trsserver 所在的目录，程序名是一个完全文件名，必须包含文件的后缀（如果有的话），但在 Windows 平台中，如果后缀为 ‘.exe’，则可省略。虽然 trsserver 也能出现在这个配置项中，但实际会被忽略。command 中可包含三种程序声明：&NOHUP、&BLOCK 和 &PORT。&NOHUP 声明只在 UNIX 系统中有效，它表示：该程序需要使用 nohup 命令启动；&BLOCK 声明表示：以阻塞方式启动该程序，即等待程序执行完成并退出；&PORT 声明表示：该程序启动时需要接收与 trsserver 相同的服务器端口号。

例如： ProgramStart = trssearcher &PORT &BLOCK

ProgramStart = trsmirror &PORT &BLOCK

注：trsssearcher 和 trsmirror 最好使用 ‘&BLOCK’ 声明，这样启动过程的输出信息会更加清晰，而不会相互错乱。并且当配置了 trsmirror 的自动启动时，必须同时配置 trsmirror 的自动关闭，否则，除 tdbshutdown 不能关闭 trsmirror 外，有时还会出现 trssserver 不能再启动，以及在未启动 trssserver 的情况下，客户端连接无响应等异常现象。

- 配置项 ‘ProgramClose=command’。该配置项表示 tdbshutdown 程序在关闭 TRS 服务器之前，tdbshutdown 将自动加载并运行该程序。虽然 tdbshutdown 也能出现在这个配置项中，但实际会被忽略。该配置项的格式与 ProgramStart 的完全相同。例如：`ProgramClose = trsmirror -shutdown &PORT &BLOCK。`

20.1.8 配置域AutoReboot

这个配置域用来配置服务器自动重新启动的有关参数。其配置项包括：

- 配置项 ‘PermitClock=number’。‘number’ 是一个整数值，表示允许服务器自动重新启动的起始钟点，取值范围是：-1 至 23，大于 23 时，将取其 24 的模数。-1 表示：如果服务器的空闲状态起始时间有效，则使用空闲状态起始时间的前一小时作为允许服务器自动重新启动的起始钟点，否则表示禁止服务器自动重新启动。**为确保自动重新启动能准时工作，最好将 PermitClock 的值设置在空闲状态起始时间之前。**请参见 5.7 节中的有关说明。
- 配置项 ‘ForbidClock=number’。‘number’ 是一个整数值，表示允许服务器自动重新启动的截止钟点，取值范围是：-1 至 23，大于 23 时，将取其 24 的模数。-1 表示：如果服务器的空闲状态截止时间有效，则使用空闲状态截止时间作为允许服务器自动重新启动的截止钟点，否则表示没有截止钟点。请参见 5.7 节中的有关说明。
- 配置项 ‘LeastPeriods=value’。‘value’ 是一个整数值时，表示服务器自动重新启动的最小周期（单位：天），取值范围是：1 至 28，最大是 28 天。如果为 0，或未指定具体值，则表示禁止服务器自动重新启动。特别地，当 value 为 ‘@n’ 的形式时（其中 n 为 1、2、3、4、5、6 和 7 七个数字之一），表示在每周的“星期 n”定时重新启动。如果当一个周期到达时未找到空闲时间完成重新启动，则将在下一天重新试着自动启动，依次类推（因此是“最小周期”）。
- 配置项 ‘ShutdownCall=command’。该配置项表示服务器自动重新启动之前用来关闭服务器的程序（通常是 tdbshutdown），其格式与配置域 AutoReboot 中的配置项 ProgramStart 完全相同，**需要特别注意的是：用来关闭服务器的程序不能等待服务器完全退出，而必须在发送“关闭”指令后以最快的方式返回（或者进程终止），否则自动重新启动功能就可能运转不正常。**如果不指定具体值，则表示禁止服务器自动重新启动。例如：`ShutdownCall=tdbshutdown &PORT -NOWAIT。`

特别说明：对于 6.80.5230 之前的版本，如果服务器作为集群的后台运行，则配置

域无效，即服务器不支持自动重新启动；从 6.80.5230 版本及之后的版本，如果服务器作为集群的后台运行，则只有当 `trsearcher` 也启动时，服务器才支持自动重新启动的功能，否则配置域无效。^{【修订 153】}该功能启动的前提是系统必须有机会在允许重启的时间区间内进入空闲状态（大约 10 分钟内无服务请求），但有些应用可能因有数据定时入库的功能而使系统无法进入空闲状态，致使自动重启功能无效，在这种情况下，应用需要利用操作系统的定时任务等功能，调用带 ‘-reboot’ 选项的关闭服务器的命令强制系统重启（如 ‘`tdbshutdown -reboot`’）。^{【4】}

20.1.9 配置域Background

这个配置域用来配置服务器作为一个或多个集群系统后台服务器运行时的有关参数。其配置项包括：

- 配置项 ‘`JoinCluster=switch`’。‘switch’ 是一个开关，取 ‘TRUE’ 和 ‘FALSE’ 值之一，无配置时的缺省值为 ‘FALSE’。如果取值 ‘TRUE’，则表示该服务器要加入一个或多个集群系统而作为后台服务器运行；如果取值 ‘FALSE’，则表示服务器将作为普通服务器单独运行。
- 配置项 ‘`ClusterIP=list`’。‘list’ 是 IP 地址集，每个地址之间用分号分隔。IP 地址的格式是：X.X.X.X，其中每个 X 是 0 到 255 的整数。^{【修订 050】}一个 ‘X’（或者 255）表示地址通配。如果集群系统与该服务器是处在同一机器上，则在地址集中只需包含其中一个 IP 地址或者本地地址 “127.0.0.1”，除非通配地址中包含了本机的 IP 地址。^{【4】}这个地址集必须包含该服务器所要加入的集群系统的所有 IP 地址，没有缺省值。
- 配置项 ‘`PortalUser=name`’。‘name’ 是一个用户名，用于所加入的所有集群系统登录该服务器。无配置时的缺省用户名是 ‘system’。
- 配置项 ‘`CPUWeight=value`’。‘value’ 是一个大于 0 的浮点数，表示在所加入的所有集群系统所包含的全部后台服务器中，该服务器的 CPU 的相对速度。无配置时该配置项的缺省取值是 1。
- 配置项 ‘`DiskWeight=value`’。‘value’ 是一个大于 0 的浮点数，表示在所加入的所有集群系统所包含的全部后台服务器中，该服务器的硬盘 IO 的相对速度。无配置时该配置项的缺省取值是 1。

20.1.10 配置实例

```
# this is the configuration file of server system.
#####

[IndexCache]
```

```
# LowLimit is the upper limit of word's low-frequency. Those
# words are not cached, if their frequency are below it.
# e.g: LowLimit=100000          (the default value)
LowLimit=

# MediumLimit is the lower limit of word's medium-frequency.
# Those words are only cached in the first-level, when their
# frequency are between LowLimit and MediumLimit.
# e.g: MediumLimit=400000      (the default value)
MediumLimit=

# HighLimit is the lower limit of word's high-frequency. The
# words can be cached with both first-level and second-level,
# when their frequency are between MediumLimit and HighLimit,
# and those words are only cached in the second-level, while
# their frequency are above the HighLimit.
# e.g: HighLimit=1600000       (the default value)
HighLimit=

# MemorySize is the available maximum memories by mega-bytes
# for cache. Its default value is from Server configurations
# while failing to detect the total physical memory size.
# e.g: MemorySize=1024         (1G bytes)
MemorySize=

# DiskSize is the available maximum disk quota in mega-bytes
# for second-level cache files. It is adjusted automatically
# according to the remanent available space, if its value is
# -1. The zero value (default) will disable the second-level
# cache files.
# e.g: DiskSize=1024           (1G bytes)
DiskSize=

# MemoryFS is the maximum memories in mega-bytes reserved by
# the Memory File System. If its value is zero (or no value),
# the MFS would not be started. And it is allowed to specify
# a RAM file or other given device for MFS running mode.
# e.g: MemoryFS=1024           (1GB in auto-mode)
```

```
#      MemoryFS=1024@WAE          (1GB in Windows' WAE)
#      MemoryFS=1024@RAM          (1GB in normal RAM)
#      MemoryFS=1024@file        (1GB in RAM Device)
#      The -1 means detecting and adjusting automatically.
MemoryFS=

# WordLoad is the mode for loading indices of words accessed
# into cache.
# e.g: WordLoad=OFF              (the default value)
#      The ASY option means loading asynchronously, the SYN
#      option means loading synchronously, and the OFF mode
#      will not allow to cache any word's indices.
WordLoad=

# Expression is the cache option of interim retrieval result.
# e.g: Expression=ANY           (the default value)
#      The ANY option means that it's decided by system how
#      to cache expression's result, the TOP option denotes
#      to only cache the top level expression's result, and
#      the OFF option will not allow to cache any result of
#      expression.
Expression=

# Preload specifies the databases (or views) and cloumns (or
# data) to preload indices (or values) to cache.
# e.g: Preload=database/view
#      Preload=database/view(column[, ...])
#      Preload=database/view(data[, column, ...])
Preload=

#####

[Maintain]

# BTCacheEnable is a switch to decide if the cache of B-Tree
# is enabled or not, or partial, when we maintain it.
# e.g: BTCacheEnable=WRITE      (the default value)
#      The TRUE option will enable to read and write B-Tree
```

```
# file with memory cache. But the WRITE option permits
# only writing operation to use cache. The FALSE value
# will disable to use cache for any operation.
BTCacheEnable=

# IDXAsyncLimit is the maximum number of IDX descriptors for
# asynchronous closing while maintaining indices of database.
# The value 0 (default) denotes to prohibit the asynchronous
# closing function, and which function is effectual only for
# Windows platform to improve performance.
# e.g: IDXAsyncLimit=0 (the default value)
# The maximum value is 2048.
IDXAsyncLimit= 【修订 027】

# IDXAsyncDelay is the least delay time (seconds) before the
# IDX descriptor is closed asynchronously. In the delay time,
# we hope that the IDX descriptor will be reused again.
# e.g: IDXAsyncDelay=300 (the default value)
# The maximum value is 3600.
IDXAsyncDelay= 【修订 029】

# IDXAsyncIdled is the least interval time (seconds) without
# any access to this asynchronous closing function to denote
# that it's idle and can close all IDX descriptors. The zero
# value denotes to forbid idle function.
# e.g: IDXAsyncIdled=30 (the default value)
# The maximum value is 3600.
IDXAsyncIdled= 【修订 031】

# FileSyncFlush switch decides whether or not to synchronize
# (flush) the database file data buffered in the file system
# cache to hard disk before the files are closed.
# e.g: FileSyncFlush=FALSE (the default value)
# The TRUE is another option to enable flushing.
FileSyncFlush= 【修订 033】

# BitmapFilter setting decides whether or not to create and
# use a bitmap to filter the logically deleted records while
```

```
# optimizing database. A bitmap maybe speed up the filtering
# process, if there are a large number of deleted records in
# a single database.
# e.g: BitmapFilter=FALSE          (the default value)
#     The TRUE is another option to enable using bitmap.
BitmapFilter=TRUE    【修订 038】

# IdleTaskRules is the rules of idle-task to describe: which
# will functions be executed? when is the function executed?
# which would databases be processed by a function? A rule's
# format is 'NAME%DATABASE_MOD@TIME%TIME_MOD', the separator
# between rules is the symbol ';' or ','.
#
# The 'NAME' in format is task name, one of the followings:
# 'optimax': optimize a database entirely.
# 'optimin': optimize a database partly, just for indices of
# DATE/CHAR/NUMBER columns.
# 'index': create the indices of column that partial records
# has indices.
#
# The 'DATABASE_MOD' in format is group number of databases.
# The databases are grouped by the modulus of database ID by
# 'DATABASE_MOD'. One group will be processed in a idle-zone,
# and all groups will be processed by turns. If its value is
# zero, this task will be disabled.
#
# The 'TIME' in format is a week-day or periods-day, decided
# by the after 'TIME_MOD' in format, to execute this task.
#
# The 'TIME_MOD' in format has two meaning:
# zero value: the foregoing 'TIME' in format is the week-day,
# denotes that the task will be executed in the day per week.
# non-zero value: it is a periods to split the year-day into
# some slices, and the foregoing 'TIME' in format is a order
# number from 0 to 'TIME_MOD' - 1 of one day in a periods.
#
# e.g: IdleTaskRules=optimax%1@0%1;optimin%1@0%1;index%1@0%1
#     this is the the default setting.
```



```
IdleTaskRules=

# IdleTimeUsers is the maximum number of active instances to
# allow system executing idle-functions when the clock is on
# the idle-zone of system configuration.
# e.g: IdleTimeUsers=0          (the default value)
#     The -1 means that system always enters idle when the
#     idle-zone comes.
IdleTimeUsers=

# IdleTaskAbort specifies whether a idle-task can be aborted,
# when the ending-time of idle status has come.
# e.g: IdleTaskAbort=FALSE     (the default value)
#     The TRUE is another option to enable.
IdleTaskAbort=

# IndexRuleWord specifies whether a rule-word is regarded as
# a unitary word, when creating database's index.
# e.g: IndexRuleWord=FALSE     (the default value)
#     The FALSE means that a rule-word is not unitary word,
#     and the TRUE denotes that it is unitary, and the DOT
#     option does that only the string connected by dot is
#     a unitary word.
IndexRuleWord=

# IndexCompress decides the compressing mode of column index.
# e.g: IndexCompress=ANY
#     The OFF means no compression, the ANY denotes to use
#     a faster compression method, and the MAX does to use
#     the compression method with a maximum ratio. But its
#     default value means to reference database option.
IndexCompress=ANY

# LastPathPrior is a setting to decide whether the last path
# of data path set in syetem configuration is adopted as the
# default path of new database without specific path defined.
# e.g: LastPathPrior=TRUE      (the default value)
#     The FALSE expresses which the first path is default,
```

```
# and the TRUE means that the last path is default.
LastPathPrior=

# FileIdleReuse switch decides whether or not to recycle and
# reuse the idle blocks in database's file while maintaining
# database's records.
# e.g: FileIdleReuse=TRUE (the default value)
# The FALSE means ignoring idle blocks, the TRUE means
# recycling and reusing the idle blocks.
FileIdleReuse=FALSE

# NiceSleepTime is the seconds of nice sleep at check points
# while maintaining data or indices of database to give more
# CPU time to other processes (if there are and) if that are
# executing the requests about database retrieval or result.
# e.g: NiceSleepTime=0 (the default value)
# The 0 means no sleep in any condition.
NiceSleepTime= 【修订 004】

# RelevanceNorm decides the computation method for norm size
# of column's value when we create a database. The different
# norm size will bring different relevance sort effect.
# e.g: RelevanceNorm=QURT (the default value)
# The QURT means using the quartic root of column size
# as norm size, or the SQRT denotes the square root of
# column size, or the LOG denotes the logarithm result
# of column size, or the COS denotes that the result's
# relevance will be computed by cosine formula.
RelevanceNorm=COS 【修订 042】

# OEMFilterMute switch decides whether or not to ignored the
# error, when failure to call the OEM filter to drawout text
# from office documents.
# e.g: OEMFilterMute=FALSE (the default value)
# The TRUE is another option to ignore error.
OEMFilterMute=TRUE 【修订 045】

# IDXStoreFreqs will decide whether or not to save the float
```

```
# value of norm size in the inverted index.
# e.g: IDXStoreNorms=FALSE          (the default value)
#     The FALSE means not to store, the TRUE means storing
#     the float value of norm size.
IDXStoreNorms=TRUE

# IDXStoreFreqs will decide whether or not to save the float
# value of the word frequency in the inverted index.
# e.g: IDXStoreFreqs=FALSE          (the default value)
#     The FALSE means not to store, the TRUE means storing
#     the float value of the word frequency.
IDXStoreFreqs=TRUE

# SplitAlphaNum is the switch for deciding whether or not to
# separate the alphanumeric strings, and the numeric strings
# with decimal point, to index them alone.
# e.g: SplitAlphaNum=FALSE          (the default value)
#     The FALSE means not to separate the strings, and the
#     TRUE means separating the strings.
SplitAlphaNum=

# ParallelIndex is the switch for deciding whether or not to
# use parallel processing mode when system is creating index,
# or loading records, or copying records.
# e.g: ParallelIndex=TRUE           (the default value)
#     The TRUE means to open parallel processing mode, and
#     the FALSE means to do with normally mode.
ParallelIndex=

#####

[Retrieval]

# ParallelLimit specifies the max number of parallel threads
# for each searching process.
# e.g: ParallelLimit=0              (the default value)
#     The -1 denotes detecting system's CPUs automatically.
ParallelLimit=
```

```
# RelevanceMode decides the sorting mode of relevance method.
# e.g: RelevanceMode=NORMAL          (the default value)
#     The NORMAL means the normal or traditional relevance,
#     the WORD_V1 means relevance sorting by the number of
#     words hit, or the WORD_V2 does by both the number of
#     words hit and the sum of their frequency, or POS_COE
#     means the normal mode with a position-coefficient of
#     word, 【修订040】or NO_IDF denotes the relevance computation
#     without regard to IDF of word in a database. 【41】
RelevanceMode=

# SortAlwaysMax is the maximum number of result records that
# the result set is allowed to sort, when the related option
# of searching is specified.
# e.g: SortAlwaysMax=0                (the default value)
#     The -1 means no limit. The 0 means that the limit is
#     form the configuration of system.
SortAlwaysMax=

# ShowWordLimit is the maximum hit number of words or column
# values for a selecting operation. It would report error if
# the actual hit number exceeded this limit.
# e.g: ShowWordLimit=100000          (the default value)
#     The -1 or 0 means no limit.
ShowWordLimit=

# HFIgnoreLimit is the minimum percent rate (between the hit
# records and total records) of word and column's value that
# will be ignored while searching.
# e.g: HFIgnoreLimit=0               (the default value)
#     The range of HFIgnoreLimit is from -1 to 100. The 0
#     or -1 means no ingoring.
HFIgnoreLimit=

# MaxVirtualSet is the maximum size of virtual table to save
# the result set. Its maximum value is 65535.
# e.g: MaxVirtualSet=512             (the default value)
```

```
MaxVirtualSet=

# HitPointExact is a switch to decide whether the hit points
# will be saved exactly in the result histories.
# e.g: HitPointExact=TRUE          (the default value)
#     The TRUE option is for saving all hit points exactly,
#     the FALSE option is for reserving inexact hit points,
#     and the NONE option will not save any hit point.
HitPointExact=    【修订 023】

# IgnorePunct is a switch to decide whether the punctuations
# will be ingored, when they appear in the phrase expression
# for retrieval.
# e.g: IgnorePunct=TRUE           (the default value)
#     the FALSE option is for searching them exactly.
IgnorePunct=    【修订 047】

# ClassStatLimit is the maximum number of valid classes. Its
# maximum value is 65535.
# e.g: ClassStatLimit=5120        (the default value)
#     The 0 means maximum limit(65535).
ClassStatLimit=    【修订 055】

# ExpressHeight is the maximum height/depth of an expression
# tree so as to ensure the stack's safety.
# e.g: ExpressHeight=5120        (the default value)
#     The value 0 means no limit.
ExpressHeight=    【修订 057】

# SortValidSize is the maximum number of result records with
# correct order after sorting.
# e.g: SortValidSize=0           (the default value)
#     The value 0 means sorting all records.
SortValidSize=    【修订 059】

# SortPruneSize is the maximum number of records to be saved
# and sorted with correct order.
# e.g: SortPruneSize=0          (the default value)
```

```
# The value 0 means sorting without pruning records.
SortPruneSize=      【修订 061】

# PrecacheFiles specifies the database (or view) and cloumns
# (or data) to preread into system's file cache.
# e.g: PrecacheFiles=database/view
# PrecacheFiles=database/view(column[, ...])
# PrecacheFiles=database/view(data[, column, ...])
PrecacheFiles=      【修订 066】

VirtualResult decides the mode of reporting virtual result
# of nil string bracketed by single quotes and nil result of
# word segmentation: ERROR means reporting error for the nil
# string; DEFAULT means old compatible mode; BLANK means the
# null set for them; the WHOLE means the operating them as a
# whole fuzzy set; and the SMART means ignoring them as they
# aren't existent in the expression.
# e.g: VirtualResult=DEFAULT      (the default value)
VirtualResult=BLANK

# DummyColPrior is a switch to decide whether the DummyLogic
# will be operated priorly to spread it to the every atom of
# expression such as word and phrase and etc.
# e.g: DummyColPrior=DEFINED      (the default value)
# The DEFINED value denotes the priority is defined in
# the column's dictionary, the TRUE value denotes that
# the DummyLogic is always operated priorly, and FALSE
# option denotes the DummyLogic is always not prior.
DummyColPrior=

# MaxFuzzyMatch is the maxinum number of index terms matched
# with a fuzzy KEY, and the other terms will be ignored.
# e.g: MaxFuzzyMatch=0      (the default value)
# The value 0 means no limit. The valid value range is
# from 0 to 65535.
MaxFuzzyMatch=4096

# ExtraWeighted is a switch to decide whether the extra mode
```

```
# of relevance weighting (for branch and etc.) is adopted.
# e.g: ExtraWeighted=FALSE          (the default value)
#     The TRUE is another option to adopt the extra mode.
ExtraWeighted=TRUE

# FocusWeighted is the setting of weighting parameters about
# Time-Focused relevance sorting. Its format is (weight-rate
# first-segment-records second-segment-records third-segment-
# records fourth-segment-records fifth-segment-records), and
# the default is (0.5 10000 100000 500000 2000000 10000000).
FocusWeighted=

#####

[Network]

# ConnectCache is a switch to decide whether the connections
# between client and server are kept, when the number is not
# large (the limit depends on OS platform).
# e.g: ConnectCache=FALSE          (the default value)
#     The TRUE is another option to turn on ConnectCache.
ConnectCache=

# TimeWaitMust is a switch to decide if the TIME_WAIT, a TCP
# state, can be ignored or not. Perhaps it is unsafe to jump
# over the TIME_WAIT state, but TCP can release the resource
# immediately.
# e.g: TimeWaitMust=FALSE          (the default value)
#     The FALSE option to allow ignoring this state of all
#     connections, and the REMOTE option to allow ignoring
#     this state only for the local connections. The other
#     option TRUE denotes that this state can't be ignored
#     for any connection.
TimeWaitMust=

# ConnectUserIP is a IP set allowed the users with a CONNECT
# type to login system, and separated by semicolon. The IP's
# format is as an 'X.X.X.X'. If some 'X' is a digital number
```

```
# between 0 and 255, it denotes that at that place, the byte
# value of IP must be the number, otherwise an 'X' expresses
# the byte can be any value.
```

```
# e.g: ConnectUserIP=X.X.X.X      (the default value)
```

```
ConnectUserIP=
```

```
# WideAreaLogin is a switch to decide if the login users are
# treated as clients from the Wide Area Network, and it will
# be ignored to check the IP's consistency of WAN clients in
# the sequent accessing.
```

```
# e.g: WideAreaLogin=FALSE      (the default value)
```

```
# The FALSE option to check the IP's consistency for a
# client, and the TRUE option to disable checking.
```

```
WideAreaLogin=TRUE 【修订 070】
```

```
#####
```

```
[Concurrent]
```

```
# RoutineLimit is the maximum number of concurrent processes
# or threads of Server. A zero value or the values above the
# license limit of system are ignored. The default value(-1)
# denotes that it whould be adjusted automatically according
# to the bit number and total physical memories of system.
```

```
# e.g: RoutineLimit=-1      (the default value)
```

```
RoutineLimit=
```

```
#####
```

```
[AutoView]
```

```
# SplitLimit is the minimum number of records of database to
# split a new database to upload or insert records.
```

```
# e.g: SplitLimit=1000000      (the default value)
```

```
# The SplitLimit's range is from 10000 to 20000000.
```

```
SplitLimit=
```

```
# InheritPath is a switch to decide whether the new database
```



```
# splitted will Inherit the path of original database.
# e.g: InheritPath=TRUE          (the default value)
#     The FALSE option expresses the new database will use
#     the default path of system configuration.
InheritPath=

#####

[AutoBatch]

# ProgramStart defines an extended module started along with
# the trsserver, and accepts three kinds of declaration such
# as '&NOHUP', '&BLOCK' and '&PORT'. The '&NOHUP' means that
# the module must be started by 'nohup' command of UNIX; the
# '&BLOCK' means that trsserver must wait for this module to
# finish and exit after starting up it; and the '&PORT' does
# that the module need a same port number as trsserver.
# e.g: ProgramStart=trsserver &PORT &BLOCK
#     ProgramStart=trsmirror &PORT &BLOCK
#     ProgramStart=trssearcher &PORT &BLOCK
#     The module name must be a whole name with suffix (if
#     there is), for example: trsserver.exe, trsmirror.exe,
#     trssearcher.exe. Actually, the trsserver module will
#     be ignored here.
ProgramStart=

# ProgramClose specifies the closing program for an extended
# module corresponding to a ProgramStart definition. And the
# format of ProgramClose is same as ProgramStart item. These
# programs are loaded by tdbshutdown, and executed before it
# one by one.
# e.g: ProgramClose=tdbshutdown &PORT &BLOCK
#     ProgramClose=trsmirror -shutdown &PORT &BLOCK
#     Here, the tdbshutdown module will be ignored.
ProgramClose=

#####
```

[AutoReboot]

```
# PermitClock defines the starting clock of time window that
# server is allowed to be rebooted. When the value is -1, it
# denotes that PermitClock will be assigned with the advance
# one hour of idle's starting clock in system configurations,
# if it is valid, otherwise rebooting is disabled.
# e.g: PermitClock=-1          (the default value)
#     The PermitClock's range is from -1 to 23.
PermitClock=
```

```
# ForbidClock defines the ending clock of time window. While
# the value is -1, it means that ForbidClock's value equals to
# the idle's over clock, if it is valid, otherwise rebooting
# has no ForbidClock.
# e.g: ForbidClock=-1         (the default value)
#     The ForbidClock's range is from -1 to 23.
ForbidClock=
```

```
# LeastPeriods defines the least periods of rebooting server
# counted by days. If the value of LeastPeriods is zero, the
# rebooting function is disabled. Especially, when the value
# is expressed as '@n' (the 'n' is a digit form 1 to 7), the
# meaning is in which day for every week the server would be
# restarted.
# e.g: LeastPeriods=0         (the default value)
#     The LeastPeriods' range is from 1 to 28.
LeastPeriods=
```

```
# ShutdownCall specifies the program to shut down server and
# other extended modules before rebooting server. Ordinarily,
# this program is namely the tdbshutdown, and shutdownCall's
# format is same as ProgramStart item of AutoBatch field. If
# this item is not defined, rebooting is disabled.
# e.g: ShutdownCall=tdbshutdown &PORT -NOWAIT
ShutdownCall=
```

[Background]

```
# JoinCluster is a switch to decide whether this system will
# join some clusters as a database background server.
# e.g: JoinCluster=FALSE          (the default value)
#      The TRUE option denotes to join.
JoinCluster=

# ClusterIP is a table to list all IPs of which clusters the
# system will join. The IP's format is as 'X.X.X.X', and the
# every 'X' must be a decimal integer between 0 and 255. the
# 'X'(or 255) is a wildcard match. There must be a semicolon
# bewteen IPs to separate each other. No default list.
# e.g: ClusterIP=127.0.0.1;192.9.200.X
ClusterIP=

# PortalUser is a user name open for all clusters in list to
# login this system.
# e.g: PortalUser=system          (the default value)
PortalUser=

# CPUWeight is a weight factor about CPU's relative speed of
# the machine among all machines.
# e.g: CPUWeight=1                (the default value)
CPUWeight=

# DiskWeight is a weight factor about IO's relative speed of
# on hard disk in the machine among all machines.
# e.g: DiskWeight=1              (the default value)
DiskWeight=
```

20.2 数据镜像配置文件

文件‘MIRROR.INI’是服务器数据镜像模块的配置文件，处在服务器主目录（安装目录）的system子目录下。该文件是一个TEXT文件，一行是一个配置项，行之间没有顺序关系。对配置文件的任何修改，必须在重新启动服务器以后才有效。镜像主节点与子节点的配置文件各有不同，下面予以分别说明。

20.2.1 主节点配置项

- 以‘#’开头的行是注释行。允许空白行。
- 配置项‘NodeWorkMode=host’。‘host’表示服务器以镜像主节点运行。
- 配置项‘SynchroClock=hour:minute’。表示镜像主节点定期生成同步操作包的触发时间。‘hour’的取值为0到23，‘minute’的取值为0到59。这里有一个特别的格式：24:x，其中‘24:’是这个格式的识别标志，‘x’是一个分钟数（可以是一个多位数），表示镜像主节点定期生成同步操作包的触发周期（间隔分钟数）。
- 配置项‘DatabaseName=owner.name’。‘owner.name’在主节点上定义一个镜像数据库。一行只能指定一个数据库，如果需要定义多个镜像数据库，则可用多行分别进行定义。如果这个数据库从系统中删除后（不是删除这个配置项），再重建一个同名的数据库，或者从数据库备份包中恢复一个同名数据库，或者将别的数据库改成这个名字，则系统不会将后来的那个数据库认为是一个镜像数据库；如果确实需要镜像后来的数据库，则必须重新启动服务器，否则必须在下次重新启动服务器之前，删除这个配置项。
- 配置项‘NewDatabaseName=owner.name’。‘owner.name’在主节点上新增一个镜像数据库。一行只能指定一个数据库，如果需要定义多个镜像数据库，则可用多行分别进行定义。这类配置项是临时配置项，在服务器数据镜像模块启动后将被自动删除。

20.2.2 主节点配置实例

```
# this is the configuration file of databases mirror.

# NodeWorkMode is a paramter of all nodes.
# e.g: NodeWorkMode=host          (or slave)
NodeWorkMode=host

# SynchroClock is a paramter of all nodes.
# e.g: SynchroClock=23:00        (hour:minute)
SynchroClock=0:30

# DatabaseName is a paramter of host node only.
# e.g: DatabaseName=owner.name   (the first database)
#     DatabaseName=owner.name   (the other database)
#     ...                       (the more databases)
DatabaseName=部委规章库
DatabaseName=地方法规与规章库
```

DatabaseName= 各国地区市场惯例库
DatabaseName= 国际条约库
DatabaseName=system. 国家法律法规库
DatabaseName=system. 司法解释库

20.2.3 子节点配置项

- 以 ‘#’ 开头的行是注释行。允许空白行。
- 配置项 ‘NodeWorkMode=slave’。‘slave’ 表示服务器以镜像子节点运行。
- 配置项 ‘SynchroClock=hour:minute’。表示镜像子节点定期主动进行同步操作的触发时间。‘hour’ 的取值为 0 到 23，‘minute’ 的取值为 0 到 59。这里有一个特别的格式：24:x，其中 ‘24:’ 是这个格式的识别标志，‘x’ 是一个分钟数（可以是一个多位数），表示镜像子节点定期主动进行同步操作的触发周期（间隔分钟数）。
- 配置项 ‘NodeUserName=user’。‘user’ 是主节点分配给该子节点的 TRS 用户名。
- 配置项 ‘NodePassword=passwors’。‘passwors’ 是主节点分配给该子节点的 TRS 用户的登录口令。
- 配置项 ‘HostNodeAddr=adrees’。‘adrees’ 是主节点所在的 IP 地址。
- 配置项 ‘HostNodePort=port’。‘port’ 是主节点的 TRS 服务器所使用的服务端口号。

20.2.4 子节点配置实例

```
# this is the configuration file of databases mirror.  
  
# NodeWorkMode is a paramter of all nodes.  
# e.g: NodeWorkMode=host          (or slave)  
NodeWorkMode=slave  
  
# SynchroClock is a paramter of all nodes.  
# e.g: SynchroClock=23:00        (hour:minute)  
SynchroClock=1:30  
  
# NodeUserName is a paramter of slave node only.  
# e.g: NodeUserName=shanghai    (of shanghai slave)  
NodeUserName=shanghai  
  
# NodePassword is a paramter of slave node only.
```

```
# e.g: NodePassword=sh123          (of shanghai slave)
NodePassword=shanghai987

# HostNodeAddr is a paramter of slave node only.
# e.g: HostNodeAddr=192.9.200.1   (of beijing host)
HostNodeAddr=10.1.0.2

# HostNodePort is a paramter of slave node only.
# e.g: HostNodePort=8888          (of beijing host)
HostNodePort=8888
```

20.3 插件配置文件

文件‘OEM.INI’是服务器的 OEM 插件配置文件，处在服务器主目录（安装目录）的 system 子目录下。该文件的格式遵循通用的初始化配置文件格式标准，一般形式请参见服务器运行参数配置文件中的有关说明。**这里特别指出的是：OEM.INI 不是只能应用于服务器端，在客户端也同样支持，但其存放位置取决于客户端应用程序的安排。在客户端使用的优点有两点：一、可减轻服务器的负担，同时还能进行并行处理；二、如果插件运行不稳定，则服务器的稳定性将受到影响。但在服务器端使用也有好处：对于记录的修改、复制以及索引的重建，服务器能自动执行 OEM 功能。**

在 OEM.INI 文件中，目前可以支持四种配置域，即：VAREnviron、DOCExtract、URLAcquire 和 DLLProduct。对配置文件的任何修改，必须在重新启动服务器以后才有效。

20.3.1 配置域VAREnviron

这个配置域（兼容 4140 版本之前的 OEMEnvironment）专门用来为其它配置域中的插件程序运行时指定所需要的环境变量的，它所包含的配置项是不定的，完全取决于各个插件程序的具体需要。但所有环境变量（包括变量名和值）的总长度不能超过 4096 个字节。如果在配置文件中重复出现，则只有第一个有效。

20.3.2 配置域DOCExtract

这个配置域（兼容 4140 版本之前的 OEMFilterDoc）用来配置一个文本抽取插件程序，用于数据库加载时对格式化文档的自动索引。在配置文件中该配置域可重复出现，每个配置域可对应一种或多种文档格式。其中的配置项包括：FORMAT、SUFFIX、

FILTER 和 MODULE。如果这些配置项重复出现，则只取第一个有效的配置项，并且 FORMAT、SUFFIX、FILTER 和 MODULE 这四个配置项必须都有效，否则该配置域将被忽略。

- 配置项 FORMAT。该配置项定义了当前配置域可支持的文档格式名的列表，格式名之间用分号或逗号隔开。这里的文档格式名必须是 TRS 全文数据库系统中数据格式词典的数据格式对象名。当 TRS 格式文件中对格式化文档使用了格式控制符时，系统将首先使用这个格式控制符在该配置项的列表中进行匹配，以查找合适配置域。如果配置项中存在与前面定义的 DOCEExtract 配置域中的 FORMAT 配置项重复的文档格式名，则该文档格式名将被忽略。
- 配置项 SUFFIX。该配置项定义了当前配置域可支持的格式化文档的常用文件名后缀的列表，后缀之间用分号或逗号隔开。这个配置项是 FORMAT 配置项的补充，当通过 FORMAT 无法匹配一个合适的配置域时，可通过文件名后缀再次进行匹配。如果配置项中存在与前面定义的 DOCEExtract 配置域中的 SUFFIX 配置项重复的文件名后缀，则该后缀将被忽略。
- 配置项 FILTER。该配置项指定文本抽取程序在动态链接库中的函数名。程序的接口规范是：

```
#include <unitype.h>
typedef VOID (_STDCALL_ *OEMERRHANDLE) (CHAR* ErrorStr);

_STDAPI_ (BOOL) DrawOutText(SourceDoc, TextFile, CharSet, HandleError)
CHAR* SourceDoc;          /* 源文件名 */
CHAR* TextFile;          /* 文本文件名 */
CHAR* CharSet;           /* GB/GBK/BIG5/UNICODE/ENGLISH */
OEMERRHANDLE HandleError; /* 输出错误信息的回调函数 */
{
    ...
    if (Success) return TRUE;
    if (HandleError != NULL) HandleError(ErrorMessage);
    return FALSE;
}
```

- 配置项 MODULE。该配置项指定插件程序所在的动态链接库，即对应于：Windows 系统下的一个*.DLL 文件；HPUX 系统下的一个*.sl 文件；其它 UNIX 系统下的一个*.so 文件。动态链接库文件可以使用绝对路径，但如果不指明其路径，或者动态链接库又引用了其它动态库，则必须在 VAREnviron 配置域中设置相应的搜索路径的环境变量：Windows 系统变量名为 PATH、AIX 系统变量名为 LIBPATH、HPUX 系统变量名为 SHLIB_PATH、其它 UNIX 系统为 LD_LIBRARY_PATH。如果有多个路径名，则路径名之间 Windows 系统里用分号 ‘;’ 分隔、UNIX 系统里一般用冒号 ‘:’ 分隔（具体请参考实际的操作系统规定）。或者：Windows 系统

可在“系统环境变量”区中设置搜索路径环境变量；而在 UNIX 系统可使用其 shell 的 env 命令启动服务器，也可用其它方法设置搜索路径环境变量（不同的系统和 shell，方法可能都不一样）。

20.3.3 配置域URLAcquire

这个配置域（兼容 4140 版本之前的 OEMGetURL）用来配置一个 URL 资源获取插件程序。在配置文件中该配置域可重复出现，每个配置域可对应一种或多种 URL 协议格式。其中的配置项包括：FORMAT、METHOD 和 MODULE。如果这些配置项重复出现，则只取第一个有效的配置项，并且 FORMAT、METHOD 和 MODULE 这三个配置项必须都有效，否则该配置域将被忽略。

- 配置项 FORMAT。该配置项定义了当前配置域可支持的协议格式名的列表，格式名之间用分号或逗号隔开。这里的协议格式名只能取下列三者之一：FILE、HTTP 和 FTP。如果配置项中存在与前面定义的 URLAcquire 配置域中的 FORMAT 配置项重复的文档格式名，则该协议格式名将被忽略。
- 配置项 METHOD。该配置项指定 URL 资源文件获取程序在动态链接库中的函数名。程序的接口规范是：

```
#include <unitype.h>
typedef VOID (_STDCALL_ *OEMERRHANDLE) (CHAR* ErrorStr);

_STDAPI_(BOOL) GetURLFile(URLName, DestFile, HandleError)
CHAR* URLName;          /* URL 资源 */
CHAR* DestFile;         /* 目标文件名 */
OEMERRHANDLE HandleError; /* 输出错误信息的回调函数 */
{
    ...
    if (Success) return TRUE;
    if (HandleError != NULL) HandleError(ErrorMessage);
    return FALSE;
}
```

- 配置项 MODULE。该配置项与 DOCEExtract 配置域中的 MODULE 配置项具有相同的意义和配置方法。

20.3.4 配置域DLLProduct

这个配置域是用来集成动态库产品的通用配置。在配置文件中该配置域可重复出现，每个配置域对应一种产品。其中的配置项包括：IDENTY、IMPORT 和 MODULE。如果这些配置项重复出现，则只取第一个有效的配置项，并且 IDENTY、IMPORT 和

MODULE 这三个配置项必须都有效，否则该配置域将被忽略。【修订 117】

- 配置项 IDENTITY。该配置项定义了当前配置域用于识别产品的法定产品名（应用程序配合定义）。如果它与前面定义的 DLLProduct 配置域中的 IDENTITY 配置项重复，则该配置域将被忽略。
- 配置项 IMPORT。该配置项指定该产品所支持的接口函数名列表，函数名之间用分号分隔。接口函数没有统一的规范，但必须与应用程序配合，为每个接口制定固定的调用方法和协议。
- 配置项 MODULE。该配置项与 DOCEExtract 配置域中的 MODULE 配置项具有相同的意义和配置方法。

20.3.5 配置实例

```
# this is the configuration file of OEM function.

#####

[VAREnviron]
# The VAREnviron section must be unique, and its
# maximum total size of all VARs is 4 kilo-bytes.
# And the following %trshome% should be replaced
# by the actual installation path of TRS Server.

# Windows configuration for HYFSOFT and LANGWARE
#PATH = %trshome%\oem\hyfsoft;%trshome%\oem\langware

# AIX configuration for HYFSOFT and LANGWARE
#LIBPATH = %trshome%/oem/hyfsoft:%trshome%/oem/langware

# HPUX configuration for HYFSOFT and LANGWARE
#SHLIB_PATH = %trshome%/oem/hyfsoft:%trshome%/oem/langware

# Linux and other UNIXs configuration for HYFSOFT and LANGWARE
#LD_LIBRARY_PATH = %trshome%/oem/hyfsoft:%trshome%/oem/langware

#####

[DOCEExtract]
# Windows configuration for HYFSOFT
```

```
#FORMAT = PDF;WORD;XCL;PPT;RTF;WPS
#SUFFIX = pdf;doc;docx;docm;dot;xls;xlsx;xlsm;xlt;ppt;pptx;pps;ppsm;rtf;wps
#FILTER = GetDocumentText
#MODULE = dmc_trs.dll
```

```
[DOCExtract]
# HPUX configuration for HYFSOFT
#FORMAT = PDF;WORD;XCL;PPT;RTF
#SUFFIX = pdf;doc;docx;docm;dot;xls;xlsx;xlsm;xlt;ppt;pptx;pps;ppsm;rtf
#FILTER = GetDocumentText
#MODULE = libdmc_trs.sl
```

```
[DOCExtract]
# Linux and other UNIXs configuration for HYFSOFT
#FORMAT = PDF;WORD;XCL;PPT;RTF
#SUFFIX = pdf;doc;docx;docm;dot;xls;xlsx;xlsm;xlt;ppt;pptx;pps;ppsm;rtf
#FILTER = GetDocumentText
#MODULE = libdmc_trs.so
```

```
[DOCExtract]
# More configuration
FORMAT =
SUFFIX =
FILTER =
MODULE =
```

#####

```
[URLAcquire]
# Just a DEMO configuration, can't be used actually for you
#FORMAT = HTTP
#METHOD = GetURLFile
#MODULE = robot.dll
```

```
[URLAcquire]
# More configuration
FORMAT =
METHOD =
```

```
MODULE =

#####

[DLLProduct]
# Windows configuration for LANGWARE
#IDENTY = LANGWARE
#IMPORT = OpenDict;CognateWords;WordStem;TextSegment;CloseDict
#MODULE = lang_trs.dll

[DLLProduct]
# HPUX configuration for LANGWARE
#IDENTY = LANGWARE
#IMPORT = OpenDict;CognateWords;WordStem;TextSegment;CloseDict
#MODULE = liblang_trs.sl

[DLLProduct]
# Linux and other UNIXs configuration for LANGWARE
#IDENTY = LANGWARE
#IMPORT = OpenDict;CognateWords;WordStem;TextSegment;CloseDict
#MODULE = liblang_trs.so

[DLLProduct]
# More configuration
IDENTY =
IMPORT =
MODULE =
```

20.4 日志配置文件

文件‘LOGINI’是服务器的运行日志配置文件，处在服务器主目录（安装目录）的 system 子目录下。该文件的组成为两部分，一部分是用来规定日志的方式与规模的配置，另一部分则是用来规定需要进行日志的请求配置。配置的基本格式是：

- 以‘#’开头的行是注释行。允许空白行。
- 配置项的格式是“项名 项值”对。“项名”与“项值”之间用空格或制表符分隔。项值一般是 0 或 1，0 表示“禁止”，1 表示“允许”。
- 每一行只能有一个配置项。

20.4.1 方式与规模配置

方式与规模配置包括四个配置项，分别是 MAX_FILENUM、MAX_FILESIZE、ALL_ALLOW、ALL_PROHIBIT 和 DAY_DIVIDE。

- MAX_FILENUM 规定总的日志文件数。
- MAX_FILESIZE 规定每个日志文件的大小（实际文件大小可能会有出入）。
- ALL_ALLOW 是否允许所有支持日志的请求都写日志。
- ALL_PROHIBIT 是否禁止所有的请求写日志。如果该项的配置值是 1，则将忽略 ALL_ALLOW 的配置值。
- DAY_DIVIDE 是否按天写日志文件，即哪天发生的请求就写入哪天的日志文件，日志文件以日期命名。如果要每天一个文件，则需将 MAX_FILESIZE 设置成适当的大小（-1 表示最大的文件尺寸），此时 MAX_FILENUM 相当于保存日志的天数。【修订 095】

20.4.2 内容配置

这一部分的配置规定了哪些请求允许写日志，哪些请求不要写日志，但这些配置只有当 ALL_ALLOW 和 ALL_PROHIBIT 的值都为 0 时才起作用，否则就被忽略。由于请求数目较多，这里不一一罗列，具体请参见 20.4.4 节的实例。

20.4.3 日志记录

日志以标准的 TRS 格式文件（参见 11.1.1 的说明）进行存储。【修订 103】日志的输出项数与请求及参数相关，不同的请求的日志输出项数不同，不同参数的同一请求的日志输出项数也不相同。下面介绍日志记录的所有项输出项及其内容：

标记	内容
REC	记录标记
Request	输出请求名。
UserID	输出请求的用户ID。
ClientIP	输出请求的源IP地址。
Instance	输出请求的用户实例ID。如果为0，则表示IDLE自动维护。
CostTime	输出请求的耗时毫秒数。
StartTime	输出请求的起始日期与时间，格式是：year.mm.dd hh:mm:ss。
CloseTime	输出请求的结束日期与时间，格式是：year.mm.dd hh:mm:ss。
Object	输出请求的操作对象，具体内容参见20.4.4节的实例注释。
Params	输出请求的主要参数，具体内容参见20.4.4节的实例注释。
Method	输出请求的排序方法，具体内容参见20.4.4节的实例注释。

标记	内容
ResultID	输出请求的对应的检索结果集。
Records	输出请求的主结果数（记录数或词数），或请求操作所涉及的记录数，具体内容参见20.4.4节的实例注释。
SubHits	输出请求的次结果数，具体内容参见20.4.4节的实例注释。
Failure	输出请求失败时的错误号及错误提示信息。

下面是两条日志记录的例子：

```
<REC>
<Request>=SELECT_RECORD
<UserID>=1
<ClientIP>=127.0.0.1
<Instance>=1
<CostTime>=2
<StartTime>=2009.04.18 09:26:14
<CloseTime>=2009.04.18 09:26:14
<Object>=SYSTEM.Demo3
<Params>=字段=(中国 * 美国)
<Method>=RELEVANCE;
<Failure>=18027 ID:0 Name:字段 Next:0
```

```
<REC>
<Request>=SELECT_RECORD
<UserID>=1
<ClientIP>=127.0.0.1
<Instance>=1
<CostTime>=15
<StartTime>=2009.04.18 09:29:22
<CloseTime>=2009.04.18 09:29:22
<Object>=SYSTEM.Demo3
<Params>=正文=(中国 * 美国)
<Method>=RELEVANCE;
<ResultID>=3
<Records>=166
<SubHits>=1402[4]
```

20.4.4 配置实例

```
# this is the configuration file of LOG functions.

MAX_FILENUM      7
MAX_FILESIZE     5000000
ALL_ALLOW        0
ALL_PROHIBIT     0
DAY_DIVIDE       1

#request         log comment
COMMAND          0
LOGIN            1
LOGOUT           1
TEST_LICENSE     0
SET_CONFIG       0
GET_CONFIG       0
CREATE_GROUP     0  #Object=Name ;
ALTER_GROUP      0  #Object=Name ;
GET_GROUP        0  #Object=Names ;
RENAME_GROUP     0  #Object=Old;Params=New ;
DELETE_GROUP     0  #Object=Names ;
CREATE_USER      0  #Object=Name ;Params=Group ;
ALTER_USER       0  #Object=Name ;Params=Group ;
GET_USER         0  #Object=Names ;
RENAME_USER      0  #Object=Old;Params=New ;
DELETE_USER      0  #Object=Names ;
CREATE_BASE      0  #Object=Name ;
ALTER_BASE       0  #Object=Name ;
GET_BASE         0  #Object=Names ;
RENAME_BASE      0  #Object=OldName ;Params=NewName ;
DELETE_BASE      1  #Object=Names ;
GET_BASECOL     0  #Object=BaseName ;Params=Names ;
RENAME_BASECOL  0  #Object=BaseName ;Params=Old ;
CREATE_VIEW      0  #Object=Name ;
ALTER_VIEW       0  #Object=Name ;
GET_VIEW         0  #Object=Names ;
RENAME_VIEW      0  #Object=OldName ;Params=NewName ;
```

```
DELETE_VIEW          0  #Object=Names;
AUTOMATE_VIEW        0  #Object=BaseName;
SHRINK_VIEW          1  #Object=Name;Params=BaseName
EXPAND_VIEW          0  #Object=Name;Params=BaseName;
GET_VIEWCOL          0  #Object=ViewName;Params=Names;
RENAME_VIEWCOL       0  #Object=ViewName;Params=Old;
CREATE_FORMAT         0  #Object=Name;Params=Suffix;
ALTER_FORMAT         0  #Object=Name;Params=Suffix;
GET_FORMAT           0  #Object=Names;Params=Suffixes;
RENAME_FORMAT        0  #Object=Old;Params=New;
GRANT_RIGHT          0  #Object=Users;Params=BaseName;
REVOKE_RIGHT         0  #Object=Users;Params=BaseName;
GET_RIGHT            0  #Object=Users;Params=BaseName;
LOCK_BASE            0  #Object=Names;Params=Mode;
COPY_BASE            0  #Object=Name;Params=SourceCols;
OPTIMIZE_BASE        1  #Object=BaseName;Params=Columns;
BACKUP_BASE          1  #Object=Names;Params=Device;
RESTORE_BASE         1  #Object=Names;Params=Device;
EXTEND_ATTR          0
MOVE_BASE            1  #Object=Name;Params=DestPath;
LOAD_RECORD          1  #Object=BaseName;Params=SourceFile;
                     #Records=SuccessNum;
INSERT_RECORD        1  #Object=BaseName;
UPDATE_RECORD        1  #Object=BaseName;Params=WhereExpr;
                     #Records=SuccessNum;
DELETE_RECORD        1  #Object=BaseName;Params=WhereExpr;
                     #Records=SuccessNum;
COPY_RECORD          1  #Object=BaseName;Params=SourceName;
                     #Records=SuccessNum;
CREATE_INDEX         1  #Object=BaseName;Params=Columns;
                     #Records=StopRecord;
REMOVE_INDEX         1  #Object=BaseName;Params=Columns;
                     #Records=StartRecord;
VALIDATE_INDEX       1  #Object=BaseName;Params=Columns;
INDEX_INFO           0  #Object=BaseName;
EDIT_LEXICON         0  #Object=DictName;
                     #Params=DeleteWords/NewWords;
SELECT_RECORD        1  #Object=Sources;Params=WhereExpr;
```

```
                                #Method=SortMethod;ResultID=ResultID;
                                #Records=TotalRecords;Words=TotalPoints;
SORT_RESULT          1  #ResultID=ResultID;Method=MethodExpr;
LOCKUP_RESULT        0  #ResultID=ResultID;Records=RecordNum;
UNLOCK_RESULT        0  #ResultID=ResultID;
GET_RESULT           0  #ResultID=ResultID;
GET_RECORD           1  #ResultID=ResultID;Params=Columns;
                                #Records=RankNum;
GET_RECORDID         0  #ResultID=ResultID;Records=RankNum;
READ_RECORD          0  #Object=BaseName;Params=Columns;
PICK_RECORDID        0  #Object=BaseName;Params=WhereExpr;
SELECT_WORD          0  #Object=BaseName;Params=TargetExpr;
                                #ResultID=ResultID;Records=TotalHits;
GET_WORD             0  #ResultID=ResultID;Records=RankNum;
GET_SORT             1  #ResultID=ResultID;Records=RecordNum;
                                #Method=MethodExpr;
DELETE_RESULT        0  #ResultID=ResultID;
RESUME_RESULT        0  #ResultID=ResultID;
CLASS_RESULT         1  #Object=ColumnName;Method=SortMethod;
                                #ResultID=ResultID;Records=ValidRecords;
```


第五部分 系统附录

附录一 字符集

TRS 全文数据库系统对于字符集的支持表现在三个方面：系统字符集、数据库字符集和应用字符集。

系统字符集也称系统主语言，用来规定在数据词典中存储字符时所使用的字符编码，在系统安装（并确认设置）时自动确定，不允许更改。可用作系统字符集的编码包括：GB18030、BIG5 和 UTF8。需要特别注意：作为一种兼容模式，虽然一个系统也可以不指定系统字符集，此时内部默认按 GB18030 的编码规则处理字符，但在这种模式下，系统不具有自动转码的功能，此时应用字符集实际上也是无效的，并且不支持以 UTF8 作为数据库字符集。系统字符集不仅决定了存储数据词典所使用的字符集，而且还决定了系统安装后缺省拥有的一系列词典（分词词典、附加分词词典、停用词典、附加停用词典、稀疏词典、主题词典、同义词典和反义词典等）的字符集，通过不同字符集的系统“升级”可以补充新的字符集的上述词典。

数据库字符集对应于数据库的语言属性，用来规定在数据库中存储字符时所使用的编码，在创建数据库时指定。参见 1.3 节的说明。如果一个数据库决定采用按词索引策略，则必须保证该系统中已经存在相应字符集的词典，否则需要通过相应字符集的系统“升级”进行词典补充。

应用字符集也称应用本地码，用来规定应用程序在输入/输出字符（串）时所使用的字符编码，应用程序可随时改变。可用作应用字符集的编码包括：GB18030、BIG5 和 UTF8。需要特别注意：作为一种兼容模式，虽然一个应用也可以不指定应用字符集，此时内部默认按 GB18030 的编码规则处理字符，但在这种模式下，系统不具有自动转码功能。

如果一个应用设置了应用字符集，同时 TRS 全文数据库系统也有确定的系统字符集，则当应用向系统“写入”数据（例如：创建/修改对象，加载/修改数据库记录等）

时，或当应用从系统“读出”数据（如读取对象，读取数据库记录等）时，都将自动完成字符编码的转换。例如：系统字符集采用 UTF8、数据库语言是 BIG5、应用字符集是 GB18030，则伴随数据的“写入”，将发生“GB18030 到 UTF8”和“GB18030 到 BIG5”的转码过程；相反地，伴随数据的“读出”，将发生“UTF8 到 GB18030”和“BIG5 到 GB18030”的转码过程。其中在“GB18030 到 BIG5”和“BIG5 到 GB18030”，以及“UTF8 到 BIG5”和“BIG5 到 UTF8”的转码过程将发生汉字的“简/繁”字形变化。进行“简/繁”字形转换的主要原因是由于 GB18030 和 UTF8 中的许多简体字在 BIG5 中没有编码，转换后可避免绝大多数字符集交叉情况下不能访问的问题。

中文的“简/繁”转换是一个非常复杂的问题。从简体到繁体有一对多的问题，从繁体到简体也有一对多的问题。不同的语境有不同的对应方法。例如：“台阶”对应“臺阶”、“台鉴”对应“台鑒”、“台布”对应“檯布”、“台风”对应“颱風”，“噍血為盟”对应“歃血為盟”、“噍血”对应“喋血”。

但语境的收集是不可能百分之百完备的，再加上还有一些与语境无关的情况（如人名等），所以汉字的“简/繁”转换很难做到完全正确。另一问题是受 BIG5 编码空间的限制，很多 GB18030 和 UTF8 的编码是无法映射到 BIG5 编码中的。

下表说明了各字符集相互转换的支持程度与问题：

目标 源	GB18030 系统或数据库字符集	BIG5 系统或数据库字符集	UTF8 系统或数据库字符集
GB18030 应用字符集	不转换。	支持，但存在问题：需要简繁双向转换，其结果可能不完全正确；目标编码 BIG5 不能容纳所有的 GB18030 码点。	完全支持。
BIG5 应用字符集	支持，但存在问题：需要简繁双向转换，其结果可能不完全正确。	不转换。	支持，但存在问题：需要简繁双向转换，其结果可能不完全正确。
UTF8 应用字符集	完全支持。	支持，但存在问题：需要简繁双向转换，其结果可能不完全正确；目标编码 BIG5 不能容纳所有的 GB18030 码点。	不转换。

附录二 保留字

【修订 101】

AB	ABF	ADJ	ALL	ALLBT
ALLNT	AMB	AND	AT	AVG
BETWEEN	BEYOND	BT		
CASE	COUNT			
DATA	DOCID			
EQU	EXACT			
INCLUDE				
LE	LIFO	LIKE	LOWER	
MAX	MIN			
NEAR	NOT	NT		
OR				
PRE	PT			
RANGE	RELEVANCE	RT		
SPELL	ST	SUM		
TO	TOASCII	TOCHINESE		
UF	UPPER			
XOR				

[*]

附录三 正则表达式

TRS 全文数据库系统正则表达式采用与 Perl 兼容的语法规则。下面是正则表达式中的一些常用模式。

.	匹配除换行符以外的所有字符。
x?	匹配 0 次或一次 x 字符。
x*	匹配 0 次或多次 x 字符，但匹配可能的最少次数。
x+	匹配 1 次或多次 x 字符，但匹配可能的最少次数。
.*	匹配 0 次或一次的任何字符。
.+	匹配 1 次或多次的任何字符。
{m}	匹配刚好是 m 个的指定字符。
{m,n}	匹配在 m 个以上 n 个以下的指定字符。
{m,}	匹配 m 个以上的指定字符。
[]	匹配符合 [] 内的字符。
[^]	匹配不符合 [] 内的字符。
[c]	匹配 c 对应的字符。
[0-9]	匹配所有数字字符。
[a-z]	匹配所有小写字母字符。
[^0-9]	匹配所有非数字字符。
[^a-z]	匹配所有非小写字母字符。
^	匹配字符开头的字符。
\$	匹配字符结尾的字符。
\d	匹配一个数字字符，和 [0-9] 语法一样。
\d+	匹配多个数字字符，和 [0-9]+ 语法一样。
\D	非数字，其他同 \d。
\D+	非数字，其他同 \d+。
\w	匹配字母或数字字符，和 [a-zA-Z0-9] 语法一样。
\w+	和 [a-zA-Z0-9]+ 语法一样。
\W	匹配非英文字母或数字字符，和 [^a-zA-Z0-9] 语法一样。
\W+	和 [^a-zA-Z0-9]+ 语法一样。
\s	空格，和 [\n\t\r\f] 语法一样。
\s+	和 [\n\t\r\f]+ 一样。
\S	非空格，和 [^\n\t\r\f] 语法一样。
\S+	和 [^\n\t\r\f]+ 语法一样。
\b	匹配以字母、数字为边界的字符串。

\B	匹配不以字母、数字为边界的字符串。
a b c	匹配 a 字符或是 b 字符或是 c 字符。
abc	匹配字符串 abc。
(pattern)	() 这个符号会记住所找寻到的字符串，是一个很实用的语法。第一个 () 内所找到的字符串变成 \$1 这个变量，第二个 () 内所找到的字符串变成 \$2 这个变量，以此类推下去。这些变量所对应的字符串按顺序组合在一起就是抽取的结果字符串。
\	如果要在 pattern 模式中找寻一个特殊字符，如 *，则要在这个字符前加上 \ 符号，这样才会让特殊字符失效。

下面给出一些例子：

perl	含有 perl 的字符串。
^perl	开头是 perl 的字符串。
perl\$	结尾是 perl 的字符串。
c g i	含有 c 或 g 或 i 的字符串。
cg{2,4}i	c 后面跟着 2 个到 4 个 g，再跟着 i 的字符串。
cg{2,}i	c 后面跟着 2 个以上 g，再跟着 i 的字符串。
cg{2}i	c 后面跟着 2 个 g，再跟着 i 的字符串。
cg*i	c 后面跟着 0 个或多个 g，再跟着 i 的字符串，如同规则 cg{0,}i 。
cg+i	c 后面跟着一个以上 g，再跟着 i 的字符串，如同规则 cg{1,}i 。
cg?i	c 后面跟着 0 个或是 1 个 g，再跟着 i 的字符串，如同规则 cg{0,1}i 。
c.i	c 后面跟着一个任意字符，再跟着 i 的字符串。
c..i	c 后面跟着二个任意字符，再跟着 i 的字符串。
[cgi]	符合有这三个字符任意一个的字符串。
[^cgi]	没有这三个字符中任意一个的字符串。
\d	符合数字的字符，可以使用 ^d+ 规则来表示一个或是多个数字组成的字符串。
\D	符合不是数字的字符，可以使用 ^D+ 规则来表示一个或是更多个非数字组成的字符串。
*	符合 * 这个字符，因为 * 在常规表达式中有它的特殊意思，所以要在这个特殊符号前加上 \ 符号，这样才会让这个特殊字符失效。

对多字节字符和 Unicode 的支持：

- 1、允许表达式中出现多字节编码的字符。如：**^[张|王|李|赵]**。
- 2、用 “[**u{dddd}**]” 可表示一个 Unicode 字符，其中 “**dddd**” 是这个 Unicode 码的十六进制表示。如：“**^[u{9648}]**” 等价于 “**^[陈]**”。
- 3、用 “[**U]**” 可表示一个任意的单字节或多字节编码字符，对应于大于 0 的所有 Unicode 字符。如：“**^[U]**” 等价于 “**^[u{1}-u{7FFFFFFF}]**”。

- 4、用“[\CH]”可表示一个汉字，实际对应于下列多个 Unicode 区域：3400-4DBF、4E00-9FAF、20000-2A6DF 和 2F800-2FA1F。例如：“^\[CH]+”表示匹配一个以“至少一个汉字开头”的字符串。^{【修订 049】}

第六部分 错误信息

一 TRS错误

1. 虚拟平台错误 (17400 ~ 17499, 17900 ~ 17999)

17400	无效的参数
17401	内存分配失败
17402	创建文件失败
17403	打开文件失败
17404	读文件失败
17405	写文件失败
17406	向以只读方式打开的文件里写数据
17407	文件溢出
17408	获取文件读写偏移量失败
17409	定位文件读写地址失败
17410	冲刷文件缓冲失败
17411	关闭文件失败
17412	修改文件大小失败
17413	创建临时文件名失败
17414	复制文件失败
17415	文件改名失败
17416	移动文件失败
17417	创建目录失败
17418	删除目录失败
17419	文件路径太长
17420	无效的路径名
17421	无效的文件后缀
17422	获取设备自由空间失败
17423	创建一个子进程/线程失败
17424	错误的数值范围表示格式
17425	数值超出了限定的有效范围
17426	创建共享内存失败
17427	打开共享内存失败
17428	追加共享内存失败
17429	访问超出共享内存的地址范围
17430	共享内存键冲突
17431	共享内存不存在

- 17432 错误的共享内存块
- 17433 获取共享内存映像文件路径失败
- 17434 无效的共享内存映像文件路径
- 17435 共享内存映像文件路径太长
- 17436 映射共享内存失败
- 17437 共享内存已被删除
- 17438 删除共享内存失败
- 17439 信号量系统未启动或已关闭
- 17440 创建信号量失败
- 17441 打开信号量失败
- 17442 信号量冲突
- 17443 信号量不存在
- 17444 获取信号量失败
- 17445 信号量嵌套太深
- 17446 信号量名重复
- 17447 删除信号量失败
- 17448 信号量集的数目超出系统限制
- 17449 系统没有足够内存进行信号量的 UNDO 操作
- 17450 共享内存段数目超出系统限制
- 17451 单个进程同时映射的共享内存段数目超出系统限制
- 17452 消息队列数目超出系统限制
- 17453 线程局部存储异常
- 17454 设置定时器失败
- 17455 操作被中断
- 17456 操作系统不支持
- 17457 取 URL 文件失败
- 17458 调用失败但未设置错误
- 17459 文件中的文本行太长
- 17460 错误的文件文本行
- 17461 调用相应功能之前应首先调用多线程应用初始化函数
- 17462 多线程应用初始化失败
- 17463 打开共享动态库失败
- 17464 连接共享动态库失败
- 17465 激活触发器失败
- 17466 设置触发器失败
- 17467 取当前工作目录失败
- 17468 找不到 OEM 功能配置文件
- 17469 错误的 OEM 功能配置文件项
- 17470 OEM 的文档功能函数调用失败
- 17471 OEM 的 URL 功能函数调用失败
- 17472 OEM 功能函数执行时发生异常
- 17473 初始化临界区互斥对象失败
- 17474 锁定临界区互斥对象失败
- 17475 初始化事件消息对象失败
- 17476 发送事件消息失败
- 17477 等待事件消息失败
- 17478 调用线程组线程失败
- 17479 线程组中所有的线程都被占用
- 17480 线程组中存在活动线程或存在未取出的执行结果
- 17481 线程组的线程初始化程序发生异常

- 17482 线程组的线程主程序发生异常
- 17483 信号量的值超出系统限制
- 17484 取文件或目录的绝对路径失败
- 17485 取文件或目录信息失败
- 17486 没有配置 URL 相应格式的 OEM 文件获取模块
- 17487 单个信号量集所包含的信号量数目超出系统限制
- 17488 一个共享内存段的大小超出系统限制范围
- 17489 字符集转换时结果缓冲区太短
- 17490 字符集之间的字符转换失败
- 17491 创建管道失败
- 17492 向管道写数据失败
- 17493 从管道读数据失败
- 17494 信号量系统的键值已被其他进程占用
- 17495 执行命令行程序失败
- 17496 错误的调用参数
- 17497 调用线程池线程失败
- 17498 线程池中所有的线程都被占用
- 17499 线程池中存在活动线程

- 17900 取操作系统信息失败
- 17901 信号量系统已经启动
- 17902 一个进程只能启动或打开一个键值的信号量系统
- 17903 试图从文件的末尾开始读取数据
- 17904 不能识别的字符集
- 17905 操作系统用户帐号不存在
- 17906 Windows 模块快照的创建或查找失败
- 17907 取模块路径及文件名失败
- 17908 没有与指定地址相匹配的模块

2. 网络扩展错误 (17800 ~ 17899)

- 17800 WINSOCK 网络子系统的版本不匹配
- 17801 网络子系统的驱动程序未安装
- 17802 在当前网络子系统的 hosts 文件中未找到指定的主机地址
- 17803 在当前网络子系统的 services 文件中未找到指定的服务端口
- 17804 获取本机 IP 地址失败
- 17805 获取本机名字失败
- 17806 获取网络连接对方端点的 IP 地址失败
- 17807 创建网络套接字失败
- 17808 绑定网络套接字失败
- 17809 设置网络套接字的侦听状态失败
- 17810 在网络套接字上接受连接请求失败
- 17811 网络连接失败
- 17812 在多个网络套接字上巡回侦听消息失败
- 17813 在网络套接字上发送数据失败
- 17814 在网络套接字上接收数据失败
- 17815 创建管道套接字失败
- 17816 向管道套接字写失败
- 17817 从管道套接字读失败

- 17818 客户与服务器间的网络连接模块的版本不匹配
- 17819 网络传输时数据的目的文件名太长
- 17820 网络传输时数据的目的文件名无效
- 17821 连接服务器失败
- 17822 网络访问重入
- 17823 不支持的网络传输数据压缩算法
- 17824 网络传输时的数据压缩失败
- 17825 网络传输时的还原压缩数据失败
- 17826 网络子系统的能力不足以支持服务器的运行
- 17827 服务器在响应服务请求时发生异常
- 17828 服务器已经启动
- 17829 上次的异步服务结果还未取走
- 17830 错误的服务请求数据
- 17831 没有更多的请求等待服务
- 17832 服务请求数据序列不正确
- 17833 错误的服务结果数据
- 17834 没有更多的服务结果数据
- 17835 多余的服务结果数据
- 17836 服务请求被防火墙禁止
- 17837 插件程序已经注册
- 17838 无效的插件程序注册名
- 17839 相同插件程序注册名的类型冲突
- 17840 服务请求进入脱机处理状态
- 17841 服务请求异常终止
- 17842 错误的服务端口号
- 17843 服务器端口不正确或冲突
- 17844 服务器并发新进程失败
- 17845 服务器并发的新进程不能是当前进程的子进程
- 17846 服务请求被后来重入的服务器请求自动取消
- 17847 套接字的个数超出 FD_SETSIZE 的范围
- 17848 套接字的值超出 FD_SETSIZE 的范围
- 17849 服务器关闭时服务请求被终止
- 17850 服务请求失败内部错误
- 17851 无效的网络 IP 地址
- 17852 服务总控模块内部错误
- 17853 应用指定的客户端 IP 地址无效
- 17854 不是 IPv4 网络地址
- 17855 不支持的代理服务器类型
- 17856 服务请求未通过代理服务器的安全检测
- 17857 错误的代理服务器地址或端口
- 17858 客户或服务器句柄不是代理服务器模式
- 17859 所连接的服务器不是代理服务器
- 17860 错误的代理服务器的目的地址或端口
- 17861 代理服务器模式的用法不正确
- 17862 等待期间没有任何结果到达
- 17863 无效的服务端口号
- 17864 UNIX 域套接字地址的路径太长

3. 内存文件错误 (19800 ~ 19899)

- 19801 内存文件系统未启动或已关闭
- 19802 内存文件不存在
- 19803 内存空间不足
- 19812 初始化内存文件系统失败
- 19813 清理内存文件系统失败
- 19814 创建内存文件失败
- 19815 打开内存文件失败
- 19816 关闭内存文件失败
- 19817 删除内存文件失败
- 19818 获取内存文件读写位置失败
- 19819 定位内存文件读写地址失败
- 19820 读内存文件失败
- 19821 写内存文件失败
- 19822 映射内存块失败
- 19823 取消内存块映射失败

4. 数据库维护错误 (18000 ~ 18499)

- 18000 错误的数据库文件类型
- 18001 源字段值太长
- 18002 源记录太大
- 18003 错误的字段值文件
- 18004 记录不存在
- 18005 记录已被删除
- 18006 错误的文件数据块
- 18007 错误的文件空闲块
- 18008 错误的文件数据块大小
- 18009 重新分配文件数据块失败
- 18010 数据记录被密级加密
- 18011 数据记录被用户组加密
- 18012 数据记录被口令加密
- 18013 不支持的字段类型
- 18014 数据库中没有定义字段
- 18015 不能修改已加载了数据的数据库
- 18016 修改数据库字段失败
- 18017 索引属性名或别名不合法
- 18018 没有定义索引属性
- 18019 不支持的索引属性
- 18020 不支持的索引属性特征
- 18021 没有定义索引属性特征
- 18022 不存在给定的数据格式
- 18023 不存在给定的用户组
- 18024 不存在给定的用户
- 18025 不存在给定的数据库
- 18026 不存在给定的视图
- 18027 不存在给定的字段
- 18028 指定的位置属性没有建立索引
- 18029 数据库记录号已经用完
- 18030 错误的数据库优化的临时设备文件
- 18031 错误的数据库加载控制文件
- 18032 错误的数据库加载源数据文件
- 18033 不支持的源数据文件格式
- 18034 错误的源数据记录制标
- 18035 错误的源数据字段制标
- 18036 错误的源数据记录或字段制标
- 18037 缺少源数据记录或字段制标
- 18038 错误的 ISO2709 格式
- 18039 错误的日期格式
- 18040 数据库文件中的字段描述信息与数据词典不一致
- 18041 无效的数字值
- 18042 文本字段值重复
- 18043 错误的 DOCUMENT 字段值加载格式
- 18044 错误的字段值加载控制符格式
- 18045 不支持的字段值存储模式
- 18046 字段不允许空值
- 18047 给定字段上没有建立索引

- 18048 字段值文件的控制符重复
- 18049 对文本字段值进行分词时失败
- 18050 对字符串进行分词时失败
- 18051 不能使用系统保留的索引属性
- 18052 错误的索引属性定义格式
- 18053 系统正在进行还原事物处理
- 18054 错误的数据库记录地址
- 18055 数据库加载时记录号内部错误
- 18056 文件空闲块内部错误
- 18057 错误的反义词项描述
- 18058 取不到系统配置数据词典
- 18059 不支持的压缩算法
- 18060 不完整的倒排索引块
- 18061 倒排索引的顺序性不正确
- 18062 B 树索引的顺序性不正确
- 18063 错误的索引词词频不正确
- 18064 错误的索引词命中记录数不正确
- 18065 被修改或复制的记录中哑数据不能重新建索引
- 18066 设置的文档分词定制函数太多
- 18067 找不到支持给定数据格式的文档分词函数
- 18068 数据库中没有文本类型的字段
- 18069 文本数据压缩失败
- 18070 文本数据解压缩失败
- 18071 索引数据压缩失败
- 18072 索引数据解压缩失败
- 18073 读出的数据块太长
- 18074 复制记录时源字段的类型与目的字段的类型不兼容
- 18075 UNIQUE 字段值重复
- 18076 UNIQUE 字段上有一些记录还未建索引
- 18077 给定索引属性的特征值不是唯一的
- 18078 没有指定记录的镜像 ID
- 18079 源镜像 ID 与新定义的镜像 ID 不配对
- 18080 已建立索引的数据库不允许被修改
- 18081 不合法或不支持的 NUMBER 字段子类型
- 18082 错误的 XML 格式
- 18083 XML 标记太长
- 18084 错误的记录索引过滤文件
- 18085 数据加载时记录号内部冲突错误
- 18086 字段的类型不支持字段标记属性
- 18087 多个字段值之间的字段标记属性不一致
- 18088 错误的字段标记属性格式
- 18089 不合法的字段标记属性
- 18090 索引属性名或别名太长
- 18091 错误的索引属性值格式
- 18092 错误的索引属性编码方式
- 18093 缺少索引属性名或特征名
- 18094 索引属性名或别名重复
- 18095 索引属性特征名或特征别名不合法
- 18096 索引属性特征名或特征别名太长
- 18097 错误的索引属性特征值编码格式

- 18098 索引属性特征名或特征别名重复
- 18099 定义的索引属性或属性特征太多
- 18100 索引属性值的特征值内部编码错误
- 18101 错误的索引属性限定格式
- 18102 索引优化内部错误
- 18103 新的索引属性定义不能兼容原索引属性定义
- 18104 不同字符集的数据库之间不能进行记录复制
- 18105 数据库记录的序列计数器溢出
- 18106 B 树索引块被破坏
- 18107 不支持带通配符的对象名
- 18108 自动分裂视图的记录加载回调函数的返回值不正确
- 18109 错误的索引文件类型
- 18110 属性倒排在倒排块中的位置偏移量不正确
- 18111 B 树索引项错误或者 B 树索引可能被破坏
- 18112 索引维护时的内部错误
- 18113 数据库字段的 ID 超过了 1023 的最大限制
- 18114 视图字段的 ID 超过了 1023 的最大限制
- 18115 错误的数据库文件版本
- 18116 读文件时的起始地址超出了文件尺寸
- 18117 错误的数据库文件头
- 18118 TRANSPORT 格式装库文件错误
- 18119 指定的字段类型不支持索引属性
- 18120 索引属性的正则表达式编译失败
- 18121 装库数据文件中记录没有指定 DOCID
- 18122 索引中含有已被物理删除的记录
- 18123 TRANSPORT 格式装库文件中加密或压缩的描述不正确
- 18124 移动数据库文件时目的路径下已存在同名文件
- 18125 位置索引读取内部错误

5. 数据库检索错误 (18500 ~ 18899)

- 18500 检索内部错误
- 18501 统计表达式语法错误
- 18502 统计表达式中的逗号不能在括号中出现
- 18503 检索表达式语法错误
- 18504 错误的运算符格式
- 18505 错误的限定运算符
- 18506 加权限定运算符重复
- 18507 位置限定运算符重复
- 18508 句限定运算符重复
- 18509 段限定运算符重复
- 18510 索引属性限定运算符重复
- 18511 不支持的限定运算符
- 18512 限定运算符冲突
- 18513 索引属性类型与限定运算符不匹配
- 18514 错误的函数调用格式
- 18515 字段检索表达式不能嵌套
- 18516 检索表达式中的单引号不配对
- 18517 错误的转义符

- 18518 错误的检索结果引用
- 18519 缺少检索运算符
- 18520 缺少检索词
- 18521 检索表达式中的括号不配对
- 18522 括号的嵌套层次太多
- 18523 检索表达式中的存在不允许的限定运算
- 18524 没有权限进行检索
- 18525 数据库处于禁止检索状态
- 18526 文本字段不支持词或短语的范围检索
- 18527 文本字段不支持大小比较运算符
- 18528 错误的 DOCID 检索表达式格式
- 18529 比较运算符的右值不支持所使用的函数
- 18530 函数调用与字段类型不匹配
- 18531 错误的函数参数串
- 18532 打开检索历史出错
- 18533 给出的检索结果记录号超出了有效范围
- 18534 错误的检索结果排序表达式
- 18535 排序目标重复
- 18536 目标字段的数据类型不支持排序
- 18537 在不相似的数据库上不能进行混合排序
- 18538 函数参数中的词典不是一个有效的辅助知识词典
- 18539 函数调用与词典类型不匹配
- 18540 错误的词查询表达式
- 18541 错误的词查询函数
- 18542 错误的缺省字段关系运算符
- 18543 目标数据库或视图中不存在全文检索字段
- 18544 检索操作超过系统配置的最长执行时间
- 18545 目标数据库或视图中没有指定缺省检索字段
- 18546 符合条件的词或字段值太多
- 18547 字段的数据类型不支持键值列表
- 18548 总命中数限定运算符重复
- 18549 总词频数限定运算符重复
- 18550 命中记录的词频数限定运算符重复
- 18551 字段的索引属性定义不支持多个属性同时进行限定运算
- 18552 检索缓冲的外部交换区剩余空间不足
- 18553 存储排序向量时的内部错误
- 18554 错误的排序向量文件
- 18555 字段类型不支持特定字段值的扫描查找
- 18556 在表达式中存在非法字符
- 18557 字段的数据类型不支持所使用的限定运算符
- 18558 字段的数据类型不支持按字段值分类检索结果记录
- 18559 按字段值进行分类的结果集中包含了不同字符集的数据库
- 18560 按字段值进行分类的目标字段在结果集所包含的数据库中的数据类型不一致
- 18561 按字段值进行分类的目标字段在结果集所包含的数据库中的数据类型不一致
- 18562 结果集的历史文件可能被破坏
- 18563 统计表达式中的括号不配对
- 18564 错误的加权运算符
- 18565 错误的函数调用
- 18566 错误的正则表达式格式
- 18567 正则表达式编译失败

- 18568 检索表达式运算树递归太深
- 18569 检索表达式中的单引号没有包含内容
- 18570 词查询表达式不支持逻辑运算
- 18571 检索表达式中的大括号不配对
- 18572 SPELL 函数不支持文本字段的非字母文字单词的排除运算
- 18573 不支持所指定的错误字段的处理模式
- 18574 在虚拟字段上不支持比较运算
- 18575 错误的虚拟字段逻辑属性
- 18576 虚拟字段的逻辑属性中包含了虚拟字段
- 18577 字段限定运算符重复
- 18578 缺少统计表达式运算符
- 18579 缺少统计表达式操作数
- 18580 指定的辅助知识词典与数据库的字符集不一致
- 18581 调用的函数已经被废弃
- 18582 不支持的 SPAN 限定用法
- 18583 不支持的 COL 限定用法
- 18584 未找到指定的辅助知识词典

6. 分词系统错误 (18900 ~ 18999)

- 18900 分词系统未初始化
- 18901 不能识别词典文件的版本
- 18902 不能识别词典数据源文件的版本
- 18903 目标对象不是分词系统所支持的词典
- 18904 目标词典禁止维护
- 18905 无效的词汇
- 18906 词典类型与数据词典不匹配
- 18907 错误的词典类型
- 18908 系统未安装与数据库字符集对应的二元索引所必需的词典
- 18909 给定的分词词典不存在
- 18910 分词失败
- 18911 分词语种不支持
- 18912 同根词语种不支持
- 18913 词根语种不支持
- 18914 给定的附加分词词典不存在
- 18915 给定的停用词典不存在
- 18916 给定的附加停用词典不存在
- 18917 给定的稀疏词典不存在
- 18918 给定的实体词典不存在

7. 远过程调用错误 (19000 ~ 19499)

- 19000 远过程调用协议版本不匹配
- 19001 不支持的服务请求
- 19002 指定的数据库名对应了多个数据库
- 19003 指定的数据格式名对应了多个数据格式
- 19004 无效的词典数据文件

- 19005 没有指定待插入的记录数据
- 19006 没有指定修改记录的新数据
- 19007 没有查询到需要维护的数据库记录
- 19008 数据对象类型不支持所进行的操作
- 19009 没有记录被删除
- 19010 复制数据记录时源数据库和目的数据库不能是同一个数据库
- 19011 结构化服务请求数据流时失败
- 19012 结构化服务结果数据流时失败
- 19013 指定的视图名对应了多个视图
- 19014 指定的视图不是自动分裂模式的视图
- 19015 本地语言和服务器主语言不兼容
- 19016 不支持 64 位数据的流化或结构化
- 19017 从自动视图中复制数据记录时不支持排序
- 19018 用 GB18030 字符集编码时给定的对象名超过了其允许的最大长度
- 19019 用 UTF8 字符集编码时给定的对象名超过了其允许的最大长度
- 19020 数据库中没有记录
- 19021 不存在或没有相应条件的用户组
- 19022 不存在或没有相应条件的用户
- 19023 不存在或没有相应条件的数据库
- 19024 不存在或没有相应条件的数据库字段
- 19025 不存在或没有相应条件的视图
- 19026 不存在或没有相应条件的视图字段
- 19027 不存在或没有相应条件的词典
- 19028 不存在或没有相应条件的数据库或词典
- 19029 不存在或没有相应条件的数据格式
- 19030 远过程调用内部错误
- 19031 服务器版本太低
- 19032 维护的目标对象必须是物理数据库
- 19033 错误的异步请求结果类型
- 19034 修改记录指定的键值字段值太长

8. 系统注册码错误 (19500 ~ 19599)

- 19500 注册码编码失败
- 19501 注册码解码失败
- 19502 不能获取有效的机器码
- 19503 注册码与机器码不匹配
- 19504 注册码已过期失效
- 19505 给定的注册码无效
- 19506 无效的产品标识号
- 19507 无效的产品版本号
- 19508 无效的产品许可证数
- 19509 无效的产品试用期限
- 19510 无效的机器码
- 19511 无效的产品定制码
- 19512 产品定制码太长
- 19513 不能识别的注册码格式
- 19514 不能识别的注册码格式版本
- 19515 给定的注册码可能被篡改

- 19516 试用期间系统时间可能被回调
- 19517 无效的产品签名
- 19518 产品签名与注册码不匹配
- 19519 系统不支持位字长整型
- 19520 给定的注册码与产品不匹配
- 19521 产品版本与注册码不匹配
- 19522 动态库可能被篡改
- 19523 注册码与产品安装时间不匹配
- 19524 无效的机器 MAC 地址码

9. 数据词典错误 (20000 ~ 20499)

- 20000 数据词典共享内存区分配失败
- 20001 数据词典共享内存区清除失败
- 20002 数据词典共享内存区打开失败
- 20003 数据词典共享内存区关闭失败
- 20004 错误的数据词典类型
- 20005 错误的数据词典文件名
- 20006 共享内存打开失败
- 20007 共享内存定位失败
- 20008 共享内存定位失败
- 20009 文件打开失败
- 20010 文件读取失败
- 20011 文件写入失败
- 20012 文件定位失败
- 20013 指定的文件不存在
- 20014 读取数据词典头部失败
- 20015 读取系统配置词典失败
- 20016 分配变长数据区结点失败
- 20017 读取变长数据区结点失败
- 20019 完全互斥信号量打开失败
- 20020 完全互斥信号量获取失败
- 20021 并发控制信号量打开失败
- 20022 并发控制信号量获取失败
- 20023 常规内存不足
- 20024 错误的语言类型
- 20025 未找到指定的数据格式
- 20026 未找到指定的用户组
- 20027 未找到指定的用户
- 20028 未找到指定的数据库
- 20029 未找到指定的视图
- 20030 未找到指定的数据库字段
- 20031 未找到指定的视图字段
- 20032 未找到指定的权限信息
- 20033 未找到视图指定的物理数据库
- 20034 不合法的参数
- 20035 数据库处于进程独占排他状态
- 20036 不是 SYSTEM 用户
- 20037 不是 DBA 用户

- 20038 不是 GDA 用户
- 20039 不是 RESOURCE 用户
- 20040 用户没有 SELECT 权限
- 20041 用户没有 UPDATE 权限
- 20042 用户没有 INDEX 权限
- 20043 用户没有 CREATE 权限
- 20044 用户没有 ALTER 权限
- 20045 用户没有 DROP 权限
- 20046 用户没有 GRANT 权限
- 20047 错误的用户名或口令
- 20048 非法的访问实例
- 20049 对象重名
- 20050 多重名匹配
- 20051 未找到指定的辅助知识词典
- 20052 用户没有权限修改用户名称
- 20053 用户没有权限修改用户组名
- 20054 用户没有权限修改用户类型
- 20055 用户没有权限修改用户口令
- 20056 用户没有权限修改安全密码级
- 20057 用户没有权限修改排它登录属性
- 20058 用户没有权限修改最大检索历史步数
- 20059 用户没有权限修改最大下载记录数
- 20060 用户没有权限修改附加分词词典选项
- 20061 用户没有权限修改附加停用词典选项
- 20062 用户没有权限修改稀疏词典选项
- 20063 用户没有权限修改数据库的锁状态
- 20064 字段不是文本类型
- 20065 不允许引用系统保留字
- 20066 不合法的复合名
- 20067 不合法的数据库文件路径
- 20068 不合法的索引文件路径
- 20069 视图中物理数据库之间的语言类型不一致
- 20070 视图字段对应的物理数据库字段类型不一致
- 20071 线程或进程局部环境未初始化
- 20072 没有符合用户权限的对象
- 20073 系统配置中的数据路径为空
- 20074 试图修改 SYSTEM 用户的名称
- 20075 试图删除 SYSTEM 用户
- 20076 不合法的字段值
- 20077 未指定缺省值
- 20078 字段值不符合字段值格式限定
- 20079 字段值不符合字段值范围限定
- 20080 错误的字段值格式限定符
- 20081 不合法的对象名
- 20082 串中含有除英文字母和数字外的字符
- 20083 配置参数中空闲起始时间无效或超出范围
- 20084 配置参数中空闲结束时间无效或超出范围
- 20085 配置参数中词根索引开关无效
- 20086 配置参数中索引大小写敏感开关无效
- 20087 配置参数中忽略数字索引开关无效

- 20088 配置参数中默认数据库语言不正确
- 20089 配置参数中数据库存储路径无效
- 20090 配置参数中数据库备份路径无效
- 20091 配置参数中临时文件存储路径无效
- 20092 配置参数中缺省主分词词典的类型或语言不匹配, 或名字太长
- 20093 配置参数中缺省主停用词典的类型或语言不匹配, 或名字太长
- 20094 配置参数中缺省主题词典的类型或语言不匹配, 或名字太长
- 20095 配置参数中缺省同义词典的类型或语言不匹配, 或名字太长
- 20096 配置参数中缺省反义词典的类型或语言不匹配, 或名字太长
- 20097 用户组名太长
- 20098 用户组注释信息太长
- 20099 原用户组名太长
- 20100 新用户组名太长
- 20101 用户名太长
- 20102 用户组名太长
- 20103 用户类型无效
- 20104 用户缺省辅助分词词典名字太长
- 20105 用户缺省辅助停用词典名字太长
- 20106 用户缺省稀疏词典名字太长
- 20107 用户注释信息太长
- 20108 用户口令太长或不合法
- 20109 原用户名太长
- 20110 新用户名太长
- 20111 数据库名无效或太长
- 20112 数据库所有者名太长
- 20113 数据库别名总长度或子名长度超出范围
- 20114 数据库类型不正确
- 20115 数据库公共权限不正确
- 20116 数据库语言不正确
- 20117 数据库词根索引开关无效
- 20118 数据库索引大小写敏感开关无效
- 20119 数据库忽略数字索引开关无效
- 20120 数据库二元索引开关无效
- 20121 数据库数据与索引不压缩开关无效
- 20122 数据库缺省检索字段名无效或太长
- 20123 数据库主分词词典名太长
- 20124 数据库辅助分词词典名太长
- 20125 数据库主停用词典名太长
- 20126 数据库辅助停用词典名太长
- 20127 数据库稀疏词典名太长
- 20128 数据库数据存储属主系统名太长
- 20129 数据库数据存储路径太长或不存在
- 20130 数据库注释信息太长
- 20131 数据库字段名无效或太长
- 20132 数据库字段别名总长度或子名长度超出范围
- 20133 数据库字段类型不正确
- 20134 数据库字段公共权限不正确
- 20135 数据库字段显示标签太长
- 20136 数据库字段允许多值开关无效
- 20137 数据库字段不建索引开关无效

- 20138 数据库字段值唯一性开关无效
- 20139 数据库字段不允许空值开关无效
- 20140 数据库字段值格式长度或定义错误
- 20141 数据库字段值取值范围太长或格式错误
- 20142 数据库字段缺省值太长或格式错误
- 20143 数据库字段索引属性太长或格式错误
- 20144 数据库字段索引存储路径太长或不存在
- 20145 数据库字段注释信息太长
- 20146 描述增加数据库字段的参数错误
- 20147 描述修改数据库字段的参数错误
- 20148 描述删除数据库字段的参数错误
- 20149 原数据库名太长或无效
- 20150 数据库原所有者名太长或无效
- 20151 新数据库名太长或无效
- 20152 数据库新所有者名太长或无效
- 20153 原数据库字段名太长或无效
- 20154 新数据库字段名太长或无效
- 20155 视图名太长或无效
- 20156 视图所有者名太长
- 20157 视图别名总长度或子名长度超出范围
- 20158 视图公共权限不正确
- 20159 描述视图所包含的物理数据库的参数错误
- 20160 视图记录的选取条件太长
- 20161 视图注释信息太长
- 20162 视图字段名太长或无效
- 20163 视图字段别名总长度或子名长度超出范围
- 20164 视图字段公共权限不正确
- 20165 视图字段显示标签太长
- 20166 描述视图字段所包含的物理数据库字段的参数错误
- 20167 视图字段注释信息太长
- 20168 描述增加视图字段的参数错误
- 20169 原视图名无效或太长
- 20170 视图原所有者名太长
- 20171 新视图名无效或太长
- 20172 视图新所有者名太长
- 20173 原视图字段名无效或太长
- 20174 新视图字段名无效或太长
- 20175 数据格式名无效或非法或太长
- 20176 数据格式类别名不正确
- 20177 数据格式文件后缀不合法或太长
- 20178 数据格式允许压缩开关不正确
- 20179 数据库(或视图)名无效或太长
- 20180 数据库(或视图)所有者名太长
- 20181 指定的数据库(或视图)权限不正确
- 20182 数据库锁模式不正确
- 20183 KEY.INI 文件中有不合法的键名
- 20184 KEY.INI 文件中有不合法的键值
- 20185 KEY.INI 文件中的键值不完全
- 20186 KEY.INI 文件中的键值重复
- 20187 数据词典版本不匹配

- 20188 不允许修改已装载记录的数据库的属性
- 20189 不允许修改系统数据格式
- 20190 不允许删除系统数据格式
- 20191 机器字长或字节顺序不一致
- 20192 数据库备份文件版本不匹配
- 20193 备份文件中没有找到数据库信息
- 20194 试图在不同 IP 地址上使用同一用户名登录系统
- 20195 试图修改 PUB 用户的名称
- 20196 试图为 PUB 用户设置密码
- 20197 无权创建新用户
- 20198 无权删除用户
- 20199 无权修改用户的属性
- 20200 用户登录互斥开关无效
- 20201 不能修改数据库或词典的类型
- 20202 不能修改词典的结构
- 20203 没有索引信息
- 20204 数据词典块数达到阈值
- 20205 指定的数据库名对应了多个数据库
- 20206 指定的视图名对应了多个视图
- 20207 不是服务器本机客户端的命令式请求
- 20208 错误的命令式请求
- 20209 执行命令请求式失败
- 20210 已经建立索引的数据库的属性不允许修改
- 20211 系统维护与数据库维护操作发生冲突
- 20212 升级程序包与原系统不兼容
- 20213 OEM 版系统不能恢复企业版系统的数据库备份
- 20214 文件系统的位数不一致
- 20215 数据结构的字节对齐方式不一致
- 20216 系统中已没有可用的对象 ID
- 20217 数据库中已没有可用的字段 ID
- 20218 视图中已没有可用的字段 ID
- 20219 自动分裂模式的数据库结构或属性不允许修改
- 20220 自动分裂模式的视图结构或属性不允许修改
- 20221 指定的视图不是自动分裂模式的视图
- 20222 数据库已经是自动分裂模式的数据库
- 20223 视图中所包含的数据库数目已达到系统的最大限制
- 20224 无法为新分裂出的数据库自动命名
- 20225 数据库的扩展属性与视图不一致
- 20226 数据库的字段与视图不一致
- 20227 数据库已经包含在视图中
- 20228 数据路径集中被修改删除或新增加的目录不为空
- 20229 无权备份数据库
- 20230 待备份的数据库重复
- 20231 错误的数据库备份包
- 20232 用户没有权限修改用户注释信息
- 20233 不能直接对自动分裂模式数据库的权限进行授权
- 20234 不能直接收回自动分裂模式数据库的权限
- 20235 系统备份包不能位于服务器的驻留目录之下
- 20236 创建/删除自动模式视图的数据索引时的终止/起始记录号必须为 0
- 20237 数据库字段词根索引时的语种无效

- 20238 数据库字段混合索引模式无效
- 20239 试图通过不允许的 IP 地址进行用户登录
- 20240 操作对象不能同时包含视图和数据库
- 20241 取 OEM 数据库结构模板失败
- 20242 在按字索引模式下不能创建混合索引
- 20243 只有在按字索引或混合索引模式下才允许单字检索
- 20244 词典的语言类型与数据库的语言类型不匹配
- 20245 没有新词汇添加到词典，也没有从词典中删除词汇
- 20246 后台模式配置错误
- 20247 配置参数中辅助知识词典语言不正确
- 20248 词典引用错误或类型不匹配
- 20249 系统没有配置与数据库字符集相对应的缺省辅助知识词典
- 20250 指定的数据路径在数据路径集中已经存在
- 20251 未找到指定的数据库、视图或词典
- 20252 数据格式的文件名后缀重复
- 20253 数据格式注释信息太长
- 20254 自动分裂条件表示错误或太长
- 20255 数据库增量备份开关无效
- 20256 数据库镜像开关无效
- 20257 数据库索引时忽略停用词典开关无效
- 20258 数据库字段的类型不支持虚拟字段
- 20259 数据库虚拟字段没有指定虚拟逻辑属性
- 20260 数据库虚拟字段的逻辑属性中不能包含虚拟字段
- 20261 数据库指定了多个缺省虚拟字段
- 20262 数据库虚拟字段的逻辑属性格式错误
- 20263 数据库虚拟字段的逻辑属性包含了不存在的字段
- 20264 用户实例的客户端 IP 地址发生了变化
- 20265 分词词典的版本不匹配，需要重建数据库索引
- 20266 程序所在的目录不是系统的安装目录
- 20267 存储设备自由空间太小
- 20268 视图缺省检索字段名无效或太长
- 20269 不能恢复被保护的 OEM 数据库备份
- 20270 数据库备份包可能被篡改
- 20271 数据库 CHAR 型字段的长度限制无效
- 20272 数据库 NUMBER 型字段的精度限制无效
- 20273 已装载记录的数据库不能增加带缓存数据和允许相关性排序属性的字段
- 20274 系统实体词典不能改名
- 20275 系统实体词典不能删除
- 20276 在给定的视图里未找到相应的数据库
- 20277 不能删除视图所包含的唯一数据库

10. 集群系统错误 (20500 ~ 20999)

- 20500 参数非法
- 20501 排序最大记录号小于最小记录号
- 20502 目标服务器 ID 为 0
- 20503 参数中没有后台数据对象
- 20504 用户实例号为 0
- 20505 视图指针为空

- 20506 用户组名为空
- 20507 结果集操作句柄为空
- 20508 归并范围的最小记录号为 0
- 20509 归并范围的最大记录号 0
- 20510 归并范围的最大记录号小于最小记录号
- 20511 记录号非法
- 20512 记录号数组为空
- 20513 逻辑排序表为空
- 20514 线程池数目 Num 为 0
- 20515 检索结果号为 0
- 20516 SelectResult 参数为 NULL
- 20517 SelectDesc 参数为 NULL
- 20518 ResultSet 参数为 NULL
- 20519 文件名参数为 NULL
- 20520 负载均衡方法无效
- 20521 用户名不允许包含通配符
- 20522 服务器主目录为空
- 20523 SGetRecordDesc 参数为 NULL
- 20524 要解析检索历史号的表达式为空
- 20525 服务器信息结构指针为空
- 20526 服务器组名为空
- 20527 参数给的连接句柄无效
- 20528 负载均衡系数值越界
- 20529 SLoginDesc 指针为空
- 20530 视图 ID 非法
- 20531 所有者编码为空
- 20532 用于输出信息的地址为空
- 20533 线程入口参数为空
- 20534 配置信息结构指针为空
- 20535 文件句柄为空
- 20536 线程组池结构指针为空
- 20537 用户登录信息指针为空
- 20538 输出参数 Num 的地址为空
- 20539 连接池结构地址为空
- 20540 所有者结构指针为空
- 20541 输出缓冲区空间不够大
- 20542 取映射表的参数 RankNum <= 0
- 20543 更新映射表的参数 RankNum <= 0
- 20544 创建检索结果的后台服务器数目<=0
- 20545 排序最小记录号为 0
- 20546 排序最大记录号为 0
- 20547 用户结构的扩展分词词典不匹配
- 20548 用户结构的停用词词典不匹配
- 20549 用户结构的稀疏词典不匹配
- 20550 用户数不匹配
- 20551 词典结构的数据库数目不匹配
- 20552 词典结构的数据库名不匹配
- 20553 词典结构的字段数不匹配
- 20554 词典结构的公共权限掩码不匹配
- 20555 词典结构的所有者不匹配

- 20556 词典结构的别名不匹配
- 20557 词典结构的类型不匹配
- 20558 词典结构的语言字符集码不匹配
- 20559 词典结构的缺省的全文检索字段不匹配
- 20560 词典结构的分词词典不匹配
- 20561 词典结构的用户扩展分词词典不匹配
- 20562 词典结构的停用词词典不匹配
- 20563 词典结构的用户扩展停用词词典不匹配
- 20564 词典结构的稀疏词典不匹配
- 20565 词典结构的字段数目不匹配
- 20566 词典结构的字段名不匹配
- 20567 词典结构的字段类型不匹配
- 20568 词典结构的字段公共权限掩码不匹配
- 20569 词典结构的字段别名不匹配
- 20570 视图数目不匹配
- 20571 视图结构的名称不匹配
- 20572 视图结构的所有者不匹配
- 20573 视图结构的别名列表不匹配
- 20574 视图结构的公共权限掩码不匹配
- 20575 视图结构的数据库列表不匹配
- 20576 视图结构的数据库数目不匹配
- 20577 视图结构的字段数目不匹配
- 20578 视图结构的缺省的全文检索字段不匹配
- 20579 视图结构的选取记录的检索表达式不匹配
- 20580 视图结构的字段名称不匹配
- 20581 视图结构的字段类型不匹配
- 20582 视图结构的字段别名列表不匹配
- 20583 视图结构的用户组数目不匹配
- 20584 视图结构的用户组名字不匹配
- 20585 用户结构的用户名不匹配
- 20586 用户结构的用户所在组的组名不匹配
- 20587 用户结构的类型不匹配
- 20588 用户结构的最大下载记录数不匹配
- 20589 用户结构的最大历史步数不匹配
- 20590 用户结构的数据安全级别码不匹配
- 20591 用户结构的唯一 IP 登录不匹配
- 20592 服务器不可用
- 20593 组内无可用的服务器
- 20594 服务器不存在
- 20595 服务器的登录信息不存在
- 20596 无法为服务器组指派服务器
- 20597 没有服务器组
- 20598 字符串不是数字串
- 20599 缓冲区内没有内容
- 20600 缓冲区指针为空或者没有缓冲区
- 20601 无法登录后台服务器
- 20602 不是单一关联
- 20603 统一认证失败
- 20604 口令无效或者登录已经超时失效
- 20605 用户不存在

- 20606 登录信息无效
- 20607 还没有登录到后台服务器
- 20608 暂时还没有用户登录到集群
- 20609 客户端 IP 被服务器拒绝访问
- 20610 IP 格式不是*.*.*
- 20611 用户信息处理中途出错
- 20612 用户登录方式不支持
- 20613 视图名称不合法
- 20614 视图名称已经存在
- 20615 视图个数为 0
- 20616 取数的数据词典类型集群不支持
- 20617 视图没有匹配的字段信息
- 20618 视图没有匹配的后台数据对象信息
- 20619 没有对应的所有者名字
- 20620 视图的后台数据对象类型不合法
- 20621 目标视图不存在
- 20622 非法的记录号
- 20623 无法检索多重历史关联
- 20624 后台检索超时
- 20625 后台检索失败
- 20626 检索结果报告信息不存在或者已经被删除
- 20627 记录号越界
- 20628 未归并完的服务器数目为 0
- 20629 无法读取检索结果文件的数据对象
- 20630 检索返回空指针
- 20631 非法的排序串或者排序字段前无'+-'
- 20632 起始记录 ID 无效
- 20633 检索结果的记录数为 0
- 20634 检索结果的记录数为 0，或者用户没有后台对象检索权限
- 20635 记录号为零
- 20636 没有命中任何数据库或视图,可能没有可用后台
- 20637 参数中操作的记录数<=0
- 20638 LIFORecord=0 的 LIFO 排序或者无排序时计算后台 RankID 失败
- 20639 读取后台 RankID 失败，服务器 ID 或后台 Rank 为 0
- 20640 转化后台 RankID 失败，没有从后台取到对应记录
- 20641 归并时从排序向量中取得的 RankID 为 0
- 20642 无法取得后台 Rank，目标记录号还未归并
- 20643 Dimiension 的值为 0，后台实际不需要排序
- 20644 排序表达式相同，不需要重复排序
- 20645 LIFO 轮询计算后台 RankID 失败
- 20646 取到的记录数小于要取的记录数
- 20647 后台检索结果不存在
- 20648 后台排序被中断或失败
- 20649 排序意外中止
- 20650 非法的检索历史号
- 20651 没有读取到期待的数据
- 20652 读取配置文件信息出错
- 20653 配置信息的值未设定
- 20654 配置信息的值非法
- 20655 配置域的名字不存在

-
- 20656 在文件中搜索字符串失败
 - 20657 路径名称非法
 - 20658 排序向量文件中没有可读取的排序向量
 - 20659 二进制配置文件中字符集没有设置或者值非法
 - 20660 二进制配置文件中配置项值的个数错误
 - 20661 未定义该类型的文件
 - 20662 没有取出后台排序排序向量
 - 20663 文件中的数据损坏
 - 20664 没有为服务器负载信息分配空间
 - 20665 没有为服务器内存分配空间
 - 20666 没有为虚拟视图内存分配空间
 - 20667 没有为虚拟视图所有者内存分配空间
 - 20668 没有为服务器连接池内存分配空间
 - 20669 没有为用户信息分配空间
 - 20670 共享内存:服务器词典没有就绪
 - 20671 共享内存:存放原子数信息的空间尚未申请
 - 20672 没有为服务器组内存分配空间
 - 20673 无法设置集群配置信息
 - 20674 无法读取集群配置信息
 - 20675 请求类型不支持
 - 20676 没有可用的线程组用于并发请求
 - 20677 无法创建线程组用于并发请求
 - 20678 取数据操作的数据类型不支持
 - 20679 响应请求的时间超过系统限制
 - 20680 集群正在运行
 - 20681 集群尚未运行
 - 20682 无法清除集群配置信息
 - 20683 要删除的服务器还有未使用完的连接
 - 20684 没有空闲的连接池
 - 20685 查找的连接池不存在
 - 20686 无法连接到后台服务器
 - 20687 引用的互斥量的地址为空
 - 20688 下标出错或者越界
 - 20689 服务器状态表为空
 - 20690 服务器的状态信息不存在
 - 20691 后台服务器响应请求超时
 - 20692 数据词典的版本号无效
 - 20693 无法读取数据词典的版本信息
 - 20694 数据词典的版本信息被破坏
 - 20695 数据词典的版本信息不匹配或无效
 - 20696 共享内存键值非法或错误
 - 20697 共享内存句柄无效
 - 20698 共享内存类别值越界
 - 20699 取后台静态信息失败
 - 20700 参数的指针为空
 - 20701 检索结果未分类或分类数为 0
 - 20702 无法恢复所有后台的检索历史
 - 20703 文件操作过程中出错
 - 20704 共享内存序号值出错
 - 20705 负载均衡模式错误

- 20706 共享内存内容错误
- 20707 没有为实例回收表分配空间
- 20708 没有分配足够的空间
- 20709 未找到目标项
- 20710 发生错误但无法取得错误信息
- 20711 调用失败
- 20712 后台调用失败或超时

11. 镜像系统错误 (30001 ~ 30099)

- 30001 取当前工作目录出错
- 30002 错误的镜像配置文件
- 30003 找不到镜像数据包
- 30004 错误的镜像数据包
- 30005 磁盘空间不够
- 30006 错误的数据库结构
- 30007 不是 TRS 的用户名

12. 应用程序接口错误 (21000 ~ 21999)

- 21000 客户端没有足够的内存
- 21001 非法客户端句柄
- 21002 无效的语言类型
- 21003 非法对象列表句柄
- 21004 非法的对象类型
- 21005 对象下标超界
- 21006 非法的用户类型
- 21007 非法的数据库类型
- 21008 非法的字段类型
- 21009 非法检索句柄
- 21010 必须是记录检索句柄
- 21011 检索结果对象序号越界
- 21012 统计表达式序号越界
- 21013 非法记录号
- 21014 非法字段号
- 21015 读字段失败
- 21016 非法多值号
- 21017 非法命中点类型
- 21018 非法命中点号
- 21019 记录密码不匹配
- 21020 锁定模式不合法
- 21021 没有指定字段值
- 21022 没有指定保存字段名或多个数据库间没有同名的字段
- 21023 非法的记录范围
- 21024 不正确的记录保存格式
- 21025 无效的源文件保存模式
- 21026 记录超过 ISO2709 格式允许的长度
- 21027 字段值超过 ISO2709 格式允许的长度
- 21028 非法的物理记录号列表
- 21029 执行过程被中断
- 21030 恢复自动模式视图中的数据库已存在
- 21031 恢复自动模式视图中的视图已存在
- 21032 连接或登录服务器超时
- 21033 找不到记录对应的物理数据库
- 21034 结构个数不能小于零
- 21035 检索结果分类序号越界
- 21036 上次检索结果为空
- 21037 没有登录到 TRS 服务器
- 21038 对象属性不存在
- 21039 没有指定属性名
- 21040 不能修改只读属性
- 21041 不是知识辅助词典集里包含的词典类型
- 21042 不是词语查询句柄
- 21043 装库错误报告文件格式不正确
- 21044 系统访问许可证格式错误
- 21045 没有指定文件名
- 21046 没有指定装库源文件
- 21047 没有指定服务器列表或服务器地址个数不够

- 21048 无效的字符集类型
- 21049 没有插入新字段
- 21050 读取记录失败
- 21051 检索目标对象不能为空
- 21052 对象名或属性值太长
- 21053 用户组不存在
- 21054 保存记录长度超过文件限制
- 21055 属性文件太大
- 21056 尚未获取服务器所在系统的信息
- 21057 服务器序号超过实际范围
- 21058 设备序号超过实际范围
- 21059 设备类型不合法
- 21060 初始化 ONCE 失败
- 21061 无效函数调用
- 21062 CONNECT 过多
- 21063 SELECT 过多
- 21064 无效 CONNECT
- 21065 无效 SELECT
- 21066 关联 SELECT 过多
- 21067 无效句柄
- 21068 分词句柄不支持该操作
- 21069 记录检索句柄不支持该操作
- 21070 找不到指定用户 ID
- 21071 没有连接到 TRS 服务器
- 21072 文本型字段索引进行词根处理时的语种属性不合法
- 21073 字词混合索引时的存储模式不合法
- 21074 字段值中包含装库文件标记
- 21075 服务器返回检索结果为空

13. 命令语言错误 (22000 ~ 22999)

- 22000 不合法的 TRS SQL 句柄
- 22001 初始化命令语言失败
- 22002 语句中出现非法字符
- 22003 标识符里包含非法字符
- 22004 数字中包含非法字符
- 22005 缺少字符串结束符
- 22006 缺少左括号
- 22007 缺少右括号
- 22008 缺少逗号
- 22009 缺少保留字
- 22010 缺少标识符
- 22011 缺少字符串
- 22012 缺少数字
- 22013 缺少密码
- 22014 缺少文件路径
- 22015 缺少文件名
- 22016 缺少服务器地址
- 22017 缺少电子邮件地址

-
- 22018 缺少表达式
 - 22019 缺少对象名
 - 22020 缺少字段名
 - 22021 没有指定选项
 - 22022 没有找到适合的选项
 - 22023 选项重复
 - 22024 没有指定子句
 - 22025 没有找到适合的子句
 - 22026 子句重复
 - 22027 多余的参数
 - 22028 参数太多
 - 22029 非法的 TRSL 语句
 - 22030 知识辅助词典集太多
 - 22031 字段名和字段值个数不匹配
 - 22032 WHERE 子句和 RECORD 子句不能同时出现
 - 22033 没有指定字段的类型
 - 22034 当前没有检索结果
 - 22035 没有指定服务器地址
 - 22036 不合法的变量名
 - 22037 指定的变量不存在
 - 22038 没有指定脚本文件
 - 22039 词典不能指定字段
 - 22040 参数数值超出限定范围
 - 22041 没有指定保存结果文件名
 - 22042 词法符号序号超过实际个数
 - 22043 找不到合适的语法范式
 - 22044 选项序号超过实际范围
 - 22045 起始记录号不能大于终止记录号
 - 22046 缺少指定参数
 - 22047 创建索引的终止记录号不能为零
 - 22048 删除索引的起始记录号不能为零
 - 22049 没有找到需要删除的源数据库字段
 - 22050 表达式或字符串前缺少空白
 - 22051 更新索引词典不能指定字段名

二 操作系统错误

1. 基本系统错误 (17000 ~ 17199, 17300 ~ 17399)

17000	系统函数调用被中断
17001	错误的文件描述符
17002	无权访问文件或设备
17003	错误的地址
17004	无效的参数
17005	打开的文件太多
17006	无效的功能号
17007	指定的文件或目录不存在
17008	指定的路径不存在
17009	内存块被破坏
17010	没有足够的内存
17011	无效内存块地址
17012	无效的环境变量
17013	无效的格式
17014	无效的访问码
17015	无效的数据
17016	设备不存在
17017	不能删除工作目录
17018	不是同一设备
17019	没有更多的文件
17020	参数表太长
17021	执行文件格式错误
17022	交叉设备连接
17023	文件表溢出
17024	没有子进程
17025	在设备进行的 IO 控制不正确
17026	正文文件被占用
17027	文件太大
17028	设备上已没有足够的剩余空间
17029	在管道上不能进行定位操作
17030	只读文件系统
17031	连接太多
17032	管道已断开
17033	域错误
17034	数值超出允许范围
17035	文件或 IPC 对象已经存在
17036	锁定操作死锁错误
17037	没有操作权限
17038	进程不存在
17039	输入输出错误
17040	没有指定的设备或地址

17041	资源暂时不可用
17042	操作设备不是块设备
17043	资源被占用
17044	指定对象不是一个目录
17045	指定对象是一个目录
17046	结构没有清零
17047	操作将导致死锁
17048	RPC 结构不正确
17049	RPC 版本不正确
17050	RPC 程序不可用
17051	程序版本不正确
17052	错误的程序过程
17053	没有可用的记录锁
17054	不正确的文件类型和格式
17055	没有所期望的那种类型的消息
17056	IPC 对象的 ID 已经被删除
17057	流资源已经用完
17058	系统调用超时
17059	不是一条数据消息
17060	流协议错误
17061	流中没有消息排在读出队列的前面
17062	设备不是一个流设备
17063	通知打开操作克隆设备
17064	安装一个不干净的 w/o 文件系统组件
17065	重复的安装包的名称
17066	版本号不匹配
17067	软件包未安装
17068	未确定的符号名
17069	异步 IO 操作被取消
17070	操作失败
17071	定时器太多
17072	系统功能尚未实现
17073	内部异步 IO 操作完成
17074	无效的字节序列
17075	软件错误
17076	媒介错误
17077	被重新定位
17078	通道号超出范围
17079	2 级没有同步
17080	3 级终止
17081	3 级复位
17082	连接号超出范围
17083	协议驱动器没有连接上
17084	没有可用的 CSI 结构
17085	2 级终止
17086	无效的交换
17087	无效的请求描述符
17088	交换设备满
17089	没有阳极
17090	无效的请求码

- 17091 无效的插槽
- 17092 错误的字体文件格式
- 17093 机器没有连接到网络上
- 17094 链接已被切断
- 17095 广告错误
- 17096 Srmount 错误
- 17097 发送时通讯错误
- 17098 企图多跳
- 17099 值太大不能存储到指定的数据类型中
- 17100 名字在网络中不唯一
- 17101 文件描述符处于错误状态
- 17102 远程地址已被改变
- 17103 不能访问所需要的共享库
- 17104 所访问的共享库已被破坏
- 17105 程序里的 LIB 节被破坏
- 17106 试图连接的库太多
- 17107 不能直接执行一个共享库
- 17108 已中断的系统调用将被重启
- 17109 流式管道操作错误
- 17110 符号连接递归
- 17111 名字太长
- 17112 目录尚未清空
- 17113 创建的进程或线程太多
- 17114 用户太多
- 17115 超出了磁盘的容量

- 17300 块的传送不完整

- 17311 不是一个 XENIX 可识别类型的文件
- 17312 没有可用的 XENIX 信号量
- 17313 是一个 XENIX 可识别类型的文件
- 17314 远程输入输出错误

- 17330 共享数据库没有打开
- 17332 Lnode 层次是不相交的
- 17333 用户未被允许登录
- 17334 达到了用户登录的极限
- 17335 在 lnode 层次里有回路
- 17336 不允许附加到 lnode
- 17338 属性未找到
- 17339 目录被占用

- 17350 控制进程不能被控制
- 17351 进程不是框架控制进程
- 17352 进程处在框架管理程序的控制之下
- 17353 进程没有处在框架管理程序的控制之下
- 17354 进程已经与一个框架管理程序连接
- 17355 进程没有连接到一个框架管理程序
- 17356 进程没有找到
- 17357 进程设置了 `mustrun`
- 17358 没有处在停止状态

-
- 17359 不允许在 CPU 时钟上操作
 - 17360 状态对所需的操作无效
 - 17361 不存在
 - 17362 到达次端点
 - 17363 不正确的缓冲区长度
 - 17364 主或次是空的
 - 17365 没有可用的中断组
 - 17366 无效的模式
 - 17367 非可扩展的定时源
 - 17368 时间超出了范围
 - 17369 已被破坏或正在被破坏

 - 17380 RFS 特定的错误
 - 17381 没有路由到达目标主机
 - 17382 没有找到介质
 - 17383 错误的介质类型
 - 17384 未能提供所需的 KEY
 - 17385 KEY 已过期
 - 17386 KEY 已被吊销
 - 17387 KEY 被服务拒绝

2. 网络系统错误 (17200 ~ 17299, 17500 ~ 17799)

- 17200 操作将被阻塞
- 17201 操作正在进行中
- 17202 操作已经在进行
- 17203 在非套接字上进行套接字操作
- 17204 缺少目的地址
- 17205 消息太长
- 17206 套接字协议类型错误
- 17207 协议不可用
- 17208 不支持的协议
- 17209 不支持的套接字类型
- 17210 操作上不支持传输端点
- 17211 不支持的协议簇
- 17212 所使用的协议簇不支持所给定的地址簇
- 17213 地址被占用
- 17214 不能分配所要求的地址
- 17215 网络子系统已经关闭
- 17216 网络不可能与目的主机连通
- 17217 网络连接被对方关闭
- 17218 网络子系统没有可用的缓冲区空间
- 17219 套接字已经进行了网络连接
- 17220 套接字尚未进行网络连接
- 17221 套接字关闭后不能发送数据
- 17222 不能拼接太多的引用
- 17223 套接字间进行连接时超时
- 17224 网络连接被拒绝
- 17227 主机已经关闭
- 17229 目标是远程对象
- 17230 在主机列表文件中没有找到指定的主机名
- 17231 本地服务器没有收到域名服务器的响应
- 17232 不可恢复的错误
- 17233 没有所需类型的数据记录
- 17234 失效的 NFS 文件句柄

- 17500 启动 WINSOCK 网络子系统失败
- 17501 WINSOCK 网络子系统版本太低
- 17502 WINSOCK 网络子系统未初始化

- 17600 NFS 网络子系统的 RTM 版本不匹配
- 17601 NFS 网络子系统要求 MS-DOS 的版本不低于 3.0
- 17602 NFS 网络子系统检测到 DOS 内存被破坏
- 17603 NFS 网络子系统检测到 INT2F 已经被其他应用所占用
- 17604 NFS 网络子系统的 RTM 文件未找到或被破坏
- 17605 NFS 网络子系统的 RTM 没有被加载
- 17606 NFS 网络子系统没有被加载
- 17607 NFS 网络子系统的版本不匹配
- 17608 NFS 网络子系统的 RTM 操作在 Windows 下不支持
- 17609 NFS 网络子系统的 RNM 版本不匹配

-
- 17610 NFS 网络子系统的 RNM 文件未找到或被破坏
 - 17611 NFS 网络子系统的 RNM 没有被加载
 - 17612 NFS 网络子系统的 RNM 操作在 Windows 下不支持
 - 17613 NFS 网络子系统的 DLL 版本不匹配
 - 17614 访问 NFS 网络子系统的实例太多
 - 17615 NFS 网络子系统 DLL 错误

 - 17700 LAN Workplace 网络子系统的协议栈未安装
 - 17701 LAN Workplace 网络子系统的异步操作尚不支持
 - 17702 LAN Workplace 网络子系统的同步操作尚不支持
 - 17703 LAN Workplace 网络子系统已没有可用的 RCB 资源

3. WINDOWS错误 (1 ~ 16999)

- 1 函数不正确。
- 2 系统找不到指定的文件。
- 3 系统找不到指定的路径。
- 4 系统无法打开文件。
- 5 拒绝访问。
- 6 句柄无效。
- 7 存储控制块被损坏。
- 8 存储空间不足，无法处理此命令。
- 9 存储控制块地址无效。
- 10 环境不正确。
- 11 试图加载格式不正确的程序。
- 12 访问码无效。
- 13 数据无效。
- 14 存储空间不足，无法完成此操作。
- 15 系统找不到指定的驱动器。
- 16 无法删除目录。
- 17 系统无法将文件移到不同的驱动器。
- 18 没有更多文件。
- 19 介质受写入保护。
- 20 系统找不到指定的设备。
- 21 设备未就绪。
- 22 设备不识别此命令。
- 23 数据错误(循环冗余检查)。
- 24 程序发出命令，但命令长度不正确。
- 25 驱动器找不到磁盘上特定区域或磁道。
- 26 无法访问指定的磁盘或软盘。
- 27 驱动器找不到请求的扇区。
- 28 打印机缺纸。
- 29 系统无法写入指定的设备。
- 30 系统无法从指定的设备上读取。
- 31 连到系统上的设备没有发挥作用。
- 32 另一个程序正在使用此文件，进程无法访问。
- 33 另一个程序已锁定文件的一部分，进程无法访问。
- 34 驱动器中的软盘有误。
- 36 用来共享的打开文件过多。
- 38 已到文件结尾。
- 39 磁盘已满。
- 50 不支持请求。
- 51 Windows 无法找到网络路径。请确认网络路径正确并且目标计算机不忙或已关闭。如果 Windows 仍然无法找到网络路径，请与网络管理员联系。
- 52 由于网络上有重名，没有连接。请到“控制面板”中的“系统”更改计算机名，然后重试。
- 53 找不到网络路径。
- 54 网络很忙。
- 55 指定的网络资源或设备不再可用。
- 56 已达到网络 BIOS 命令限制。
- 57 网络适配器硬件出错。
- 58 指定的服务器无法运行请求的操作。
- 59 出现了意外的网络错误。

- 60 远程适配器不兼容。
- 61 打印机队列已满。
- 62 服务器上没有储存等待打印的文件的空间。
- 63 已删除等候打印的文件。
- 64 指定的网络名不再可用。
- 65 拒绝网络访问。
- 66 网络资源类型不对。
- 67 找不到网络名。
- 68 超出本地计算机网络适配器卡的名称限制。
- 69 超出了网络 BIOS 会话限制。
- 70 远程服务器已暂停，或正在启动过程中。
- 71 已达到计算机的连接数最大值，无法再同此远程计算机连接。
- 72 已暂停指定的打印机或磁盘设备。
- 80 文件存在。
- 82 无法创建目录或文件。
- 83 INT 24 上的故障。
- 84 无法取得处理此请求的存储空间。
- 85 本地设备名已在使用中。
- 86 指定的网络密码不正确。
- 87 参数不正确。
- 88 网络上发生写入错误。
- 89 系统无法在此时启动另一个进程。
- 100 无法创建另一个系统信号量。
- 101 另一个进程拥有独占的信号量。
- 102 已设置信号量，无法关闭。
- 103 无法再设置信号量。
- 104 无法在中断时请求独占的信号量。
- 105 此信号量的前一个所有权已结束。
- 107 由于没有插入另一个软盘，程序停止。
- 108 磁盘在使用中，或被另一个进程锁定。
- 109 管道已结束。
- 110 系统无法打开指定的设备或文件。
- 111 文件名太长。
- 112 磁盘空间不足。
- 113 没有更多的内部文件标识符。
- 114 目标内部文件标识符不正确。
- 117 应用程序发出的 IOCTL 调用不正确。
- 118 验证写入的切换参数值不正确。
- 119 系统不支持请求的命令。
- 120 这个系统不支持该功能。
- 121 信号量超时时间已到。
- 122 传递给系统调用的数据区域太小。
- 123 文件名、目录名或卷标语法不正确。
- 124 系统调用级别不正确。
- 125 磁盘没有卷标。
- 126 找不到指定的模块。
- 127 找不到指定的程序。
- 128 没有等候的子进程。
- 130 试图使用操作(而非原始磁盘 I/O)的已打开磁盘分区的文件句柄。
- 131 试图将文件指针移到文件开头之前。

- 132 无法在指定的设备或文件上设置文件指针。
- 133 包含先前加入驱动器的驱动器无法使用 JOIN 或 SUBST 命令。
- 134 试图在已被合并的驱动器上使用 JOIN 或 SUBST 命令。
- 135 试图在已被合并的驱动器上使用 JOIN 或 SUBST 命令。
- 136 系统试图解除未合并驱动器的 JOIN。
- 137 系统试图解除未替代驱动器的 SUBST。
- 138 系统试图将驱动器合并到合并驱动器上的目录。
- 139 系统试图将驱动器替代为替代驱动器上的目录。
- 140 系统试图将驱动器合并到替代驱动器上的目录。
- 141 系统试图替代驱动器为合并驱动器上的目录。
- 142 系统无法在此时运行 JOIN 或 SUBST。
- 143 系统无法将驱动器合并到或替代为相同驱动器上的目录。
- 144 目录不是根目录下的子目录。
- 145 目录不是空的。
- 146 指定的路径已在替代中使用。
- 147 资源不足，无法处理此命令。
- 148 指定的路径无法在此时使用。
- 149 企图将驱动器合并或替代为驱动器上目录是上一个替代的目标的驱动器。
- 150 系统跟踪信息未在 CONFIG.SYS 文件中指定，或不允许跟踪。
- 151 为 DosMuxSemWait 指定的信号量事件数量不正确。
- 152 DosMuxSemWait 没有运行；已设置过多的信号量。
- 153 DosMuxSemWait 列表不正确。
- 154 输入的卷标超过目标文件系统的长度限制。
- 155 无法创建另一个线程。
- 156 接收人进程拒绝此信号。
- 157 段已被放弃且无法锁定。
- 158 段已解除锁定。
- 159 线程 ID 的地址不正确。
- 160 至少有一个参数不正确。
- 161 指定的路径无效。
- 162 信号已暂停。
- 164 无法在系统中创建更多的线程。
- 167 无法锁定文件区域。
- 170 请求的资源在使用中。
- 173 对于提供取消区域进行锁定的请求已完成。
- 174 文件系统不支持锁定类型的最小单元更改。
- 180 系统检测出错误的段号。
- 183 当文件已存在时，无法创建该文件。
- 186 传递的标志不正确。
- 187 找不到指定的系统信号量名称。
- 196 操作系统无法运行此应用程序。
- 197 操作系统当前的配置不能运行此应用程序。
- 199 操作系统无法运行此应用程序。
- 200 代码段不可大于或等于 64K。
- 203 操作系统找不到已输入的环境选项。
- 205 命令子树中的进程没有信号处理程序。
- 206 文件名或扩展名太长。
- 207 第 2 环堆栈已被占用。
- 208 没有正确输入文件名通配符 * 或 ?，或指定过多的文件名通配符。
- 209 正在发送的信号不正确。

- 210 无法设置信号处理程序。
- 212 段已锁定且无法重新分配。
- 214 连到该程序或动态链接模块的动态链接模块太多。
- 215 无法嵌套调用 LoadModule。
- 230 管道状态无效。
- 231 所有的管道范例都在使用中。
- 232 管道正在被关闭。
- 233 管道的另一端上无任何进程。
- 234 有更多数据可用。
- 240 已取消会话。
- 254 指定的扩展属性名无效。
- 255 扩展属性不一致。
- 258 等待的操作过时。
- 259 没有可用的数据了。
- 266 无法使用复制功能。
- 267 目录名无效。
- 275 扩展属性在缓冲区中不适用。
- 276 装在文件系统上的扩展属性文件已损坏。
- 277 扩展属性表格文件已满。
- 278 指定的扩展属性句柄无效。
- 282 装入的文件系统不支持扩展属性。
- 288 企图释放并非呼叫方所拥有的多用户终端运行程序。
- 298 发向信号量的请求过多。
- 299 仅完成部分的 ReadProcessMemory 或 WriteProcessMemory 请求。
- 300 操作锁定请求被拒绝。
- 301 系统接收了一个无效的操作锁定确认。
- 302 此卷太碎，不能完成这个操作。
- 303 不能打开文件，因为它正在被删除。
- 487 试图访问无效的地址。
- 534 算术结果超过 32 位。
- 535 管道的另一端有一进程。
- 536 等候打开管道另一端的进程。
- 994 拒绝访问扩展属性。
- 995 由于线程退出或应用程序请求，已放弃 I/O 操作。
- 996 重叠 I/O 事件不在信号状态中。
- 997 重叠 I/O 操作在进行中。
- 998 内存分配访问无效。
- 999 执行页内操作时的错误。
- 1001 递归太深；堆栈溢出。
- 1002 窗口无法在已发送的消息上操作。
- 1003 无法完成此功能。
- 1004 无效标志。
- 1005 此卷不包含可识别的文件系统。请确定所请求的文件系统驱动程序已加载，且此卷未损坏。
- 1006 文件所在的卷已被外部改变，因此打开的文件不再有效。
- 1007 无法在全屏幕模式下运行请求的操作。
- 1008 试图引用不存在的令牌。
- 1009 配置注册表数据库损坏。
- 1010 配置注册表项无效。
- 1011 无法打开配置注册表项。
- 1012 无法读取配置注册表项。

- 1013 无法写入配置注册表项。
- 1014 注册表数据库中的某一文件必须使用记录或替代复制来恢复。恢复成功完成。
- 1015 注册表损坏。包含注册表数据的某一文件结构损坏，或系统的文件内存映像损坏，或因为替代副本、日志缺少或损坏而无法恢复文件。
- 1016 由注册表启动的 I/O 操作失败并无法恢复。注册表无法读入、写出或清除任意一个包含注册表系统映像的文件。
- 1017 系统试图加载或还原文件到注册表，但指定的文件并非注册表文件格式。
- 1018 试图在标记为删除的注册表项上进行不合法的操作。
- 1019 系统无法分配注册表日志中所需空间。
- 1020 无法在已有子项或值的注册表项中创建符号链接。
- 1021 无法在易变父项下创建稳定子项。
- 1022 正在完成通知更改请求，而且信息没有返回到呼叫方的缓冲区中。当前呼叫方必须枚举文件来查找改动。

- 1051 停止控制被发送到其他正在运行的服务所依赖的服务。
- 1052 请求的控件对此服务无效。
- 1053 服务没有及时响应启动或控制请求。
- 1054 无法创建此服务的线程。
- 1055 服务数据库已锁定。
- 1056 服务的范例已在运行中。
- 1057 帐户名无效或不存在，或者密码对于指定的帐户名无效。
- 1058 无法启动服务，原因可能是已被禁用或与其相关联的设备没有启动。
- 1059 指定了循环服务依存。
- 1060 指定的服务并未以已安装的服务存在。
- 1061 服务无法在此时接受控制信息。
- 1062 服务未启动。
- 1063 服务进程无法连接到服务控制器上。
- 1064 当处理控制请求时，在服务中发生异常。
- 1065 指定的数据库不存在。
- 1066 服务已返回特定的服务错误码。
- 1067 进程意外终止。
- 1068 依存服务或组无法启动。
- 1069 由于登录失败而无法启动服务。
- 1070 启动后，服务停留在启动暂停状态。
- 1071 指定的服务数据库锁定无效。
- 1072 指定的服务已标记为删除。
- 1073 指定的服务已存在。
- 1074 系统当前以最新的有效配置运行。
- 1075 依存服务不存在，或已被标记为删除。
- 1076 已接受使用当前引导作为最后的有效控制设置。
- 1077 上次启动之后，仍未尝试引导服务。
- 1078 名称已用作服务名或服务显示名。
- 1079 此服务的帐户不同于运行于同一进程上的其他服务的帐户。
- 1080 只能为 Win32 服务设置失败操作，不能为驱动程序设置。
- 1081 这个服务所运行的处理和服务控制管理器相同。所以，如果服务处理程序意外中止的话，服务控制管理器无法进行任何操作。
- 1082 这个服务尚未设置恢复程序。
- 1083 配置成在该可执行程序中运行的这个服务不能执行该服务。
- 1084 不能以安全模式开始这项服务
- 1100 已达磁带的实际结尾。
- 1101 磁带访问已达文件标记。

- 1102 已达磁带或磁盘分区的开头。
- 1103 磁带访问已达一组文件的结尾。
- 1104 磁带上不再有任何数据。
- 1105 磁带不能分区。
- 1106 在访问多卷分区的新磁带时，当前的块大小不正确。
- 1107 在加载磁带时找不到磁带分区信息。
- 1108 无法锁定媒体弹出功能。
- 1109 无法卸载介质。
- 1110 驱动器中的介质可能已更改。
- 1111 已复位 I/O 总线。
- 1112 驱动器中没有媒体。
- 1113 在多字节的目标代码页中，没有此 Unicode 字符可以映射到的字符。
- 1114 动态链接库(DLL)初始化例程失败。
- 1115 系统正在关机。
- 1116 因为没有任何进行中的关机过程，所以无法中断系统关机。
- 1117 由于 I/O 设备错误，无法运行此项请求。
- 1118 串行设备初始化不成功。串行驱动程序将卸载。
- 1119 无法打开正在与其他设备共享中断请求(IRQ)的设备。至少有一个使用该 IRQ 的其他设备已打开。
- 1120 序列 I/O 操作已由另一个串行口的写入完成。
- 1121 因为已过超时时间，所以串行 I/O 操作完成。
- 1122 在软盘上找不到 ID 地址标记。
- 1123 软盘扇区 ID 字符域与软盘控制器磁道地址不匹配。
- 1124 软盘控制器报告软盘驱动程序不能识别的错误。
- 1125 软盘控制器返回与其寄存器中不一致的结果。
- 1126 访问硬盘时，重新校准操作失败。重试之后仍然不成功。
- 1127 访问硬盘时，磁盘操作失败。重试之后仍然不成功。
- 1128 当访问硬盘时，即使失败，仍须复位磁盘控制器。
- 1129 已达磁带结尾。
- 1130 服务器存储空间不足，无法处理此命令。
- 1131 检测出潜在的死锁状态。
- 1132 指定的基址或文件偏移量没有适当对齐。
- 1140 改变系统供电状态的尝试被另一应用程序或驱动程序否决。
- 1141 系统 BIOS 改变系统供电状态的尝试失败。
- 1142 试图在文件上创建超过文件系统支持的链接数。
- 1150 指定程序要求更新的 Windows 版本。
- 1151 指定程序不是 Windows 或 MS-DOS 程序。
- 1152 只能启动该指定程序的一个范例。
- 1153 该指定程序是为以前一个版本的 Windows 而写的。
- 1154 运行该应用程序所需的一个库文件已被损坏。
- 1155 没有应用程序与此操作的指定文件有关联。
- 1156 在输送指令到应用程序的过程中出现错误。
- 1157 执行该应用程序所需的库文件之一无法找到。
- 1158 当前程序已使用了 Window 管理器对象的系统允许的所有句柄。
- 1159 消息只能与同步操作一起使用。
- 1160 指出的源元素没有媒体。
- 1161 指出的目标元素已包含媒体。
- 1162 指出的元素不存在。
- 1163 指出的元素是未显示的存储资源的一部分。
- 1164 由于硬件错误，显示设备需要重新初始化。

- 1165 设备指示在尝试进一步操作之前需要清除。
- 1166 设备指示门是打开的。
- 1167 设备没有连接。
- 1168 找不到元素。
- 1169 索引中没有同指定项相匹配的项。
- 1170 在对象上不存在指定的属性集。
- 1171 传递到 `GetMouseMovePoints` 的点不在缓冲区中。
- 1172 跟踪(工作站)服务没有运行。
- 1173 找不到卷 ID。
- 1175 无法删除要被替换的文件。
- 1176 无法将替换文件移到要被替换的文件。要被替换的文件保持原来的名称。
- 1177 无法将替换文件移到要被替换的文件。要被替换的文件已被重新命名为备份名称。
- 1178 正在删除卷更改日志。
- 1179 卷更改日志处于非活动状态。
- 1180 找到一份文件,但是可能不是正确的文件。
- 1181 日志项已从日志中被删除。
- 1200 指定的设备名无效。
- 1201 设备当前未连接上,但其为一个记录连接。
- 1202 本地设备名称已有到另一网络资源的记录连接。
- 1203 无任何网络提供程序接受指定的网络路径。
- 1204 指定的网络提供程序名称无效。
- 1205 无法打开网络连接配置文件。
- 1206 网络连接配置文件损坏。
- 1207 无法枚举空载体。
- 1208 出现了扩展错误。
- 1209 指定的组名格式无效。
- 1210 指定的计算机名格式无效。
- 1211 指定的事件名格式无效。
- 1212 指定的域名格式无效。
- 1213 指定的服务名格式无效。
- 1214 指定的网络名格式无效。
- 1215 指定的共享名格式无效。
- 1216 指定的密码格式无效。
- 1217 指定的消息名格式无效。
- 1218 指定的消息目标格式无效。
- 1219 不允许一个用户使用一个以上用户名与一个服务器或共享资源的多重连接。中断与此服务器或共享资源的所有连接,然后再试一次。
- 1220 试图与网络服务器建立会话,但目前与该服务器建立的会话太多。
- 1221 工作组或域名已由网络上的另一部计算机使用。
- 1222 网络不存在或尚未启动。
- 1223 操作已被用户取消。
- 1224 请求的操作无法在使用用户映射区域打开的文件上执行。
- 1225 远程系统拒绝网络连接。
- 1226 网络连接已被适当地关闭了。
- 1227 网络传输终结点已有与其关联的地址。
- 1228 地址仍未与网络终结点关联。
- 1229 企图在不存在的网络连接上进行操作。
- 1230 企图在使用中的网络连接上进行无效的操作。
- 1231 不能访问网络位置。有关网络排除故障的信息,请参阅 Windows 帮助。
- 1232 不能访问网络位置。有关网络排除故障的信息,请参阅 Windows 帮助。

- 1233 不能访问网络位置。有关网络排除故障的信息，请参阅 Windows 帮助。
- 1234 没有任何服务正在远程系统上的目标网络终结点上操作。
- 1235 请求被终止。
- 1236 由本地系统终止网络连接。
- 1237 操作无法完成。应该重试。
- 1238 因为已达到此帐户的最大同时连接数限制，所以无法连接服务器。
- 1239 试图在这个帐户未被授权的时间内登录。
- 1240 此帐户并未得到从这个工作站登录的授权。
- 1241 请求的操作不能使用这个网络地址。
- 1242 服务器已经注册。
- 1243 指定的服务不存在。
- 1244 因为用户还未被验证，不能执行所要求的操作。
- 1245 因为用户还未登录网络，不能执行所要求的操作。指定的服务不存在。
- 1246 正在继续工作。
- 1247 试图进行初始操作，但是初始化已完成。
- 1248 没有更多的本地设备。
- 1249 指定的站点不存在。
- 1250 具有指定名称的域控制器已经存在。
- 1251 只有连接到服务器上时，该操作才受支持。
- 1252 即使没有改动，组策略框架也应该调用扩展。
- 1253 指定的用户没有一个有效的配置文件。
- 1254 Microsoft Small Business Server 不支持此操作。
- 1255 服务器已关机。
- 1256 远程系统不可用。有关网络疑难解答，请见 Windows 帮助。
- 1257 提供的安全标识符不是来自一个帐户域。
- 1258 提供的安全标识符没有域组件。
- 1259 取消了 AppHelp 对话，导致应用程序无法启动。
- 1260 由于一个软件限制策略的阻止，Windows 无法打开此程序。要获取更多信息，请打开事件查看器或与系统管理员联系。
- 1261 一个程序企图用无效的注册值。通常由未初始化的注册表引起。此错误是 Itanium 特有的。
- 1262 该共享目前处于脱机状态，或者不存在。
- 1263 在智能卡登陆过程中，验证 KDC 证书时 kerberos 协议遇到一个错误。
- 1264 试图使用智能卡子系统时，Kerberos 协议出错。
- 1265 系统检测到危害安全的尝试。请确认您能与对您进行身份验证的服务器联系。
- 1266 用于身份验证的智能卡证书被取消。请与系统管理元联系。事件日志中可能有额外的信息。
- 1267 在处理用于身份验证的智能卡证书时，检测到一个不受信用证书颁发机构。请与系统管理员联系。
- 1268 无法决定用于身份验证的智能卡证书的取消状态。请与系统管理员联系。
- 1269 用于身份验证的智能卡证书不被信任。请与系统管理员联系。
- 1270 用于身份验证的智能卡证书已经过期。请与系统管理员联系。
- 1271 计算机已经锁定而且不使用强制选项无法关机。
- 1273 当调用一个应用程序定义的回叫时，出现无效数据错误。
- 1274 在同步前台策略刷新时，组策略框架应该调用扩展。
- 1275 此驱动程序被阻止加载
- 1276 动态链接库(DLL)引用的模块不是 DLL 也不是过程可执行映像。
- 1300 并非所有被引用的特权都指派给呼叫方。
- 1301 帐户名和安全 ID 间的某些映射未完成。
- 1302 没有为该帐户特别设置系统配额限制。
- 1303 没有可用的加密密钥。返回了一个已知加密密钥。
- 1304 密码太复杂，无法转换成 LAN Manager 密码。返回的 LAN Manager 密码为空字符串。

- 1305 修订级别未知。
- 1306 表明两个修订级别是不兼容的。
- 1307 这个安全 ID 不能指派为此对象的所有者。
- 1308 这个安全 ID 不能指派为对象的主要组。
- 1309 目前未模仿客户端的线程试图在模仿令牌上操作。
- 1310 组可能未被禁用。
- 1311 目前没有可用的登录服务器处理登录请求。
- 1312 指定的登录会话不存在。可能已被终止。
- 1313 指定的特权不存在。
- 1314 客户端没有所需的特权。
- 1315 提供的名称不是正确格式的帐户名。
- 1316 指定的用户已存在。
- 1317 指定的用户不存在。
- 1318 指定的组已存在。
- 1319 指定的组不存在。
- 1320 指定的用户帐户已是指定组的成员，或是因为组包含成员所以无法删除指定的组。
- 1321 指定的用户帐户不是指定组帐户的成员。
- 1322 无法禁用或删除最后剩余的系统管理帐户。
- 1323 无法更新密码。提供作为当前密码的值不正确。
- 1324 无法更新密码。提供给新密码的值包含密码中不允许的值。
- 1325 无法更新密码。为新密码提供的值不符合字符域的长度、复杂性或历史要求。
- 1326 登录失败：未知的用户名或错误密码。
- 1327 登录失败：用户帐户限制。可能的原因包括不允许空密码，登录时间限制，或强制的策略限制。
- 1328 登录失败：违反帐户登录时间限制。
- 1329 登录失败：不允许用户登录到此计算机。
- 1330 登录失败：指定的帐户密码已过期。
- 1331 登录失败：禁用当前的帐户。
- 1332 帐户名与安全标识间无任何映射完成。
- 1333 一次请求过多的本地用户标识符(LUIDs)。
- 1334 无更多可用的本地用户标识符(LUIDs)。
- 1335 对于该特别用法，安全 ID 的次级授权部分无效。
- 1336 访问控制列表(ACL)结构无效。
- 1337 安全 ID 结构无效。
- 1338 安全描述符结构无效。
- 1340 无法创建固有的访问控制列表(ACL)或访问控制项目(ACE)。
- 1341 服务器当前已禁用。
- 1342 服务器当前已启用。
- 1343 所提供的值是无效的标识符颁发机构的值。
- 1344 无更多可用的内存以更新安全信息。
- 1345 指定的特性无效，或不与组的特性兼容。
- 1346 指定的模拟级别无效，或所提供的模拟级别无效。
- 1347 无法打开匿名级安全令牌。
- 1348 请求的验证信息类别无效。
- 1349 令牌的类型对其尝试使用的方法不适当。
- 1350 无法在与安全性无关联的对象上运行安全性操作。
- 1351 未能从域控制器读取配置信息，或者因为机器不可使用，或者是访问被拒绝。
- 1352 安全帐户管理器(SAM)或本地安全颁发机构(LSA)服务器处于运行安全操作的错误状态。
- 1353 域处于运行安全操作的错误状态。
- 1354 该操作只能在主要域控制器执行。

- 1355 指定的域不存在，或无法联系。
- 1356 指定的域已存在。
- 1357 试图超出每服务器域个数的限制。
- 1358 无法完成请求操作，因为磁盘上的严重媒体失败或数据结构损坏。
- 1359 出现了内部错误。
- 1360 通用访问类型包含于已映射到非通用类型的访问掩码中。
- 1361 安全描述符格式不正确(绝对或自相关的)。
- 1362 请求的操作只准登录过程使用。该调用过程并未被记录为登录过程。
- 1363 无法使用已在使用中的标识启动新的会话。
- 1364 特定的身份验证程序包未知。
- 1365 登录会话并非处于与请求操作一致的状态中。
- 1366 登录会话标识已在使用中。
- 1367 登录请求包含无效的登录类型值。
- 1368 在使用命名管道读取数据之前，无法经由该管道模拟。
- 1369 注册表子树的事务处理状态与请求状态不一致。
- 1370 安全性数据库内部出现损坏。
- 1371 无法在内置帐户上运行此操作。
- 1372 无法在内置特殊组上运行此操作。
- 1373 无法在内置特殊用户上运行此操作。
- 1374 无法从组中删除用户，因为当前组为用户的主要组。
- 1375 令牌已作为主要令牌使用。
- 1376 指定的本地组不存在。
- 1377 指定的帐户名不是本地组的成员。
- 1378 指定的帐户名已是本地组的成员。
- 1379 指定的本地组已存在。
- 1380 登录失败: 未授予用户在此计算机上的请求登录类型。
- 1381 已超过在单一系统中可保存机密的最大个数。
- 1382 机密的长度超过允许的最大长度。
- 1383 本地安全颁发机构数据库内部包含不一致性。
- 1384 在尝试登录的过程中，用户的安全上下文积累了过多的安全标识。
- 1385 登录失败: 未授予用户在此计算机上的请求登录类型。
- 1386 更改用户密码时需要交叉加密密码。
- 1387 该成员不存在，因此不能将其添加到本地组中或从本地组删除。
- 1388 无法将新成员加入到本地组中，因为成员的帐户类型错误。
- 1389 已指定过多的安全标识。
- 1390 更改此用户密码时需要交叉加密密码。
- 1391 表明 ACL 未包含任何可承继的组件。
- 1392 文件或目录损坏且无法读取。
- 1393 磁盘结构损坏且无法读取。
- 1394 没有指定登录会话的用户会话密钥。
- 1395 正在访问的服务有连接数目授权限制。这时候已经无法再连接，原因是已经到了可接受的连接数目上限。
- 1396 登录失败: 该目标帐户名称不正确。
- 1397 相互身份验证失败。该服务器在域控制器的密码过期。
- 1398 在客户端和/或服务器之间有一个时间差。
- 1399 无法在当前域上执行此操作。
- 1400 无效的窗口句柄。
- 1401 无效的菜单句柄。
- 1402 无效的光标句柄。
- 1403 无效的加速器表句柄。

- 1404 无效的挂钩句柄。
- 1405 无效的多重窗口位置结构句柄。
- 1406 无法创建最上层子窗口。
- 1407 找不到窗口类别。
- 1408 无效窗口；它属于另一线程。
- 1409 热键已注册。
- 1410 类别已存在。
- 1411 类别不存在。
- 1412 类别仍有打开的窗口。
- 1413 无效索引。
- 1414 无效的图标句柄。
- 1415 使用专用 DIALOG 窗口字词。
- 1416 找不到列表框标识。
- 1417 找不到通配字符。
- 1418 线程没有打开的剪贴板。
- 1419 没有注册热键。
- 1420 窗口不是合法的对话窗口。
- 1421 找不到控件 ID。
- 1422 因为没有编辑控制，所以组合框的消息无效。
- 1423 窗口不是组合框。
- 1424 高度必须小于 256。
- 1425 无效的设备上下文(DC)句柄。
- 1426 无效的挂接程序类型。
- 1427 无效的挂接程序。
- 1428 没有模块句柄无法设置非本机的挂接。
- 1429 此挂接程序只可整体设置。
- 1430 日志挂接程序已安装。
- 1431 挂接程序尚未安装。
- 1432 单一选择列表框的无效消息。
- 1433 LB_SETCOUNT 发送到非被动的列表框。
- 1434 此列表框不支持 Tab 键宽度。
- 1435 无法毁坏由另一个线程创建的对象。
- 1436 子窗口没有菜单。
- 1437 窗口没有系统菜单。
- 1438 无效的消息对话框样式。
- 1439 无效的系统范围内的(SPI_*)参数。
- 1440 已锁定屏幕。
- 1441 多重窗口位置结构中窗口的所有句柄必须具有相同的上层。
- 1442 窗口不是子窗口。
- 1443 无效的 GW_* 命令。
- 1444 无效的线程标识。
- 1445 无法处理非多重文档接口(MDI)窗口中的消息。
- 1446 弹出式菜单已经激活。
- 1447 窗口没有滚动条。
- 1448 滚动条范围不可大于 MAXLONG。
- 1449 无法以指定的方式显示或删除窗口。
- 1450 所需的系统资源不足，无法完成请求的服务。
- 1451 未分页的系统资源不足，无法完成请求的服务。
- 1452 已分页的系统资源不足，无法完成请求的服务。
- 1453 工作集的配额不足，无法完成请求的服务。

- 1454 页面文件的配额不足，无法完成请求的服务。
- 1455 页面文件太小，无法完成操作。
- 1456 找不到菜单项。
- 1457 键盘布局句柄无效。
- 1458 不允许使用挂钩类型。
- 1459 该操作需要交互式窗口工作站。
- 1460 由于超时时间已过，该操作返回。
- 1461 无效监视器句柄。
- 1500 事件日志文件损坏。
- 1501 无法打开事件日志文件，事件日志服务没有启动。
- 1502 事件日志文件已满。
- 1503 事件日志文件已在读取间更改。
- 1601 无法访问 Windows 安装服务。发生这种情况的可能是您在安全模式下运行 Windows，或是没有正确安装 Windows 安装。请与技术支持人员联系以获得帮助。
- 1602 用户取消了安装。
- 1603 安装时发生严重错误
- 1604 安装已挂起，未完成。
- 1605 这个操作只对当前安装的产品有效。
- 1606 功能 ID 未注册。
- 1607 组件 ID 并未注册。
- 1608 未知属性。
- 1609 句柄处于不正确的状态。
- 1610 这个产品的配置数据已损坏。请与技术支持人员联系。
- 1611 组件限制语不存在。
- 1612 这个产品的安装来源无法使用。请验证来源是否存在，是否可以访问。
- 1613 Windows 安装服务无法安装这个安装程序包。您必须安装含有 Windows 安装服务新版本的 Windows Service Pack。
- 1614 没有卸载产品。
- 1615 SQL 查询语法不正确或不被支持。
- 1616 记录字符域不存在。
- 1617 设备已被删除。
- 1618 正在进行另一个安装操作。请在继续这个安装操作之前完成那个操作。
- 1619 未能打开这个安装程序包。请验证程序包是否存在，是否可以访问；或者与应用程序供应商联系，验证这是否是有效的 Windows 安装服务程序包。
- 1620 未能打开这个安装程序包。请与应用程序供应商联系，验证这是否是有效的 Windows 安装服务程序包。
- 1621 启动 Windows 安装服务用户接口时有错误。请与技术支持人员联系。
- 1622 打开安装日志文件的错误。请验证指定的日志文件位置是否存在，是否可以写入。
- 1623 安装程序包的语言不受系统支持。
- 1624 应用变换时的错误。请验证指定的变换路径是否有效。
- 1625 系统策略禁止这个安装。请与系统管理员联系。
- 1626 无法执行函数。
- 1627 执行期间，函数出了问题。
- 1628 指定了无效的或未知的表格。
- 1629 提供的数据类型不对。
- 1630 这个类型的数据不受支持。
- 1631 Windows 安装服务未能启动。请与技术支持人员联系。
- 1632 临时文件夹已满或无法访问。请释放一些空间或验证您是否有写入临时文件夹的权限。
- 1633 这个处理器类型不支持该安装程序包。请与产品供应商联系。
- 1634 组件没有在这台计算机上使用。

- 1635 无法打开修补程序包。请验证修补程序包是否存在，是否可以访问；或者与应用程序供应商联系，验证这是否是 Windows 安装服务的修补程序包。
- 1636 无法打开修补程序包。请与应用程序供应商联系，验证这是否是 Windows 安装服务的修补程序包。
- 1637 Windows 安装服务无法处理这个修补程序包。您必须安装含有 Windows 安装服务新版本的 Windows Service Pack。
- 1638 已安装这个产品的另一个版本。这个版本的安装无法继续。要配置或删除这个产品的现有版本，请用“控制面板”上的“添加/删除程序”。
- 1639 无效的命令行参数。有关详细的命令行帮助，请查阅 Windows 安装服务的 SDK。
- 1640 在终端服务远程会话期间，只有管理员有添加、删除或配置服务器软件的权限。如果您要在服务器上安装或配置软件，请与网络管理员联系。
- 1641 要求的操作已成功结束。要使改动生效，必须重新启动系统。
- 1642 Windows 安装服务无法安装升级修补程序，因为被升级的程序可能会丢失或是升级修补程序可能更新此程序的一个不同版本。请确认要被升级的程序在您的计算机上且您的升级修补程序是正确的。
- 1643 软件限制策略不允许有修补程序包。
- 1644 软件限制策略不允许一个或多个自定义项。
- 1700 字符串绑定无效。
- 1701 绑定句柄的类型不正确。
- 1702 绑定句柄无效。
- 1703 不支持 RPC 协议序列。
- 1704 RPC 协议序列无效。
- 1705 字符串通用唯一标识(UUID)无效。
- 1706 终结点格式无效。
- 1707 网络地址无效。
- 1708 找不到终结点。
- 1709 超时值无效。
- 1710 找不到对象通用唯一标识(UUID)。
- 1711 对象通用唯一标识(UUID)已注册。
- 1712 类型通用唯一标识(UUID)已注册。
- 1713 RPC 服务器已在侦听。
- 1714 未登记任何协议序列。
- 1715 RPC 服务器未在侦听。
- 1716 管理程序类型未知。
- 1717 接口未知。
- 1718 没有绑定。
- 1719 没有协议序列。
- 1720 无法创建终结点。
- 1721 资源不足，无法完成该操作。
- 1722 RPC 服务器不可用。
- 1723 RPC 服务器太忙，不能完成该操作。
- 1724 网络选项无效。
- 1725 该线程中没有远程过程调用活动。
- 1726 远程过程调用失败。
- 1727 远程过程调用失败且未运行。
- 1728 远程过程调用(RPC)协议出错。
- 1730 RPC 服务器不支持传送语法。
- 1732 不支持通用唯一标识(UUID)类型。
- 1733 标识无效。
- 1734 数组绑定无效。

- 1735 绑定不包含登录项名字。
1736 名称语法无效。
1737 不支持该名称语法。
1739 没有可用来创建通用唯一标识(UUID)的网络地址。
1740 终结点重复。
1741 未知的验证类型。
1742 调用的最大个数太小。
1743 字符串太长。
1744 找不到 RPC 协议顺序。
1745 过程号超出范围。
1746 绑定不包含任何验证信息。
1747 未知的验证服务。
1748 未知的验证级别。
1749 安全上下文无效。
1750 未知的授权服务。
1751 项目无效。
1752 服务器终结点无法运行操作。
1753 终结点映射器中没有更多的终结点可用。
1754 未导出任何接口。
1755 项目名不完整。
1756 版本选项无效。
1757 没有其他成员。
1758 没有内容未导出。
1759 接口没有找到。
1760 项目已存在。
1761 找不到项目。
1762 名称服务不可用。
1763 网络地址族无效。
1764 不支持请求的操作。
1765 没有安全上下文可用来允许模仿。
1766 远程过程调用(RPC)中发生内部错误。
1767 RPC 服务器试图以零除整数。
1768 RPC 服务器出现寻址错误。
1769 RPC 服务器上的浮点操作导致以零做除数。
1770 RPC 服务器上发生浮点下溢。
1771 RPC 服务器上发生浮点溢出。
1772 自动句柄绑定的可用 RPC 服务器列表已用完。
1773 无法打开字符翻译表文件。
1774 包含字符转换表的文件小于 512 字节。
1775 在远程过程调用时, 将空的上下文句柄从客户端传递到主机。
1777 在远程过程调用时, 上下文句柄已更改。
1778 传送给远程过程调用的绑定句柄不匹配。
1779 承接体无法获得远程过程调用句柄。
1780 向占位程序传送了空的索引指针。
1781 枚举值越界。
1782 字节计数太小。
1783 占位程序接收到错误数据。
1784 提供给请求操作的用户缓冲区无效。
1785 磁盘媒体无法识别。可能未被格式化。
1786 工作站没有信任机密。

- 1787 服务器上的安全数据库没有此工作站信任关系的计算机帐户。
- 1788 主域和受信域间的信任关系失败。
- 1789 此工作站和主域间的信任关系失败。
- 1790 网络登录失败。
- 1791 此线程的远程过程调用已在进行中。
- 1792 试图登录，但是网络登录服务没有启动。
- 1793 用户帐户已过期。
- 1794 重定向程序正使用中，不能卸载。
- 1795 指定的打印机驱动程序已安装。
- 1796 指定的端口未知。
- 1797 未知的打印机驱动程序。
- 1798 未知的打印机处理器。
- 1799 指定的分隔页文件无效。
- 1800 指定的优先级无效。
- 1801 打印机名无效。
- 1802 打印机已存在。
- 1803 打印机命令无效。
- 1804 指定的数据类型无效。
- 1805 指定的环境无效。
- 1806 没有更多的绑定。
- 1807 所用帐户为域间信任帐户。请使用您的全局用户帐户或本地用户帐户来访问这台服务器。
- 1808 所用帐户是一个计算机帐户。使用您的全局用户帐户或本地用户帐户来访问此服务器。
- 1809 已使用的帐户为服务器信任帐户。使用您的全局用户帐户或本地用户帐户来访问此服务器。
- 1810 指定域的名称或安全标识(SID)与该域的信任信息不一致。
- 1811 服务器在使用中且无法卸载。
- 1812 指定的映像文件不包含资源区域。
- 1813 找不到映像文件中指定的资源类型。
- 1814 找不到映像文件中指定的资源名。
- 1815 找不到映像文件中指定的资源语言标识。
- 1816 配额不足，无法处理此命令。
- 1817 未登记任何接口。
- 1818 远程过程调用被取消。
- 1819 绑定句柄不包含所有需要的信息。
- 1820 在远程过程调用过程中通信失败。
- 1821 请求的验证等级不被支持。
- 1822 未登记任何主名称。
- 1823 指定的错误不是有效的 Windows RPC 错误码。
- 1824 已经分配只在这台计算机上有效的 UUID。
- 1825 出现了特定的安全程序包错误。
- 1826 线程未取消。
- 1827 无效的编码/解码句柄操作。
- 1828 系列程序包的不兼容版本。
- 1829 RPC 占位程序版本不兼容。
- 1830 RPC 管道对象无效或已损坏。
- 1831 试图在 RPC 管道对象上进行无效操作。
- 1832 不支持该版本的 RPC 管道。
- 1898 找不到该组成员。
- 1899 无法创建终结点映射表数据库项。
- 1900 对象通用唯一标识(UUID)为 nil UUID。
- 1901 指定的时间无效。

- 1902 指定的格式名称无效。
- 1903 指定的格式大小无效。
- 1904 指定的打印机句柄正等候在
- 1905 已删除指定的打印机。
- 1906 打印机的状态无效。
- 1907 用户首次登录之前必须更改密码。
- 1908 找不到此域的域控制器。
- 1909 引用的帐户当前已锁定，且可能无法登录。
- 1910 没有发现指定的此对象导出者
- 1911 没有发现指定的对象。
- 1912 没有发现指定的对象解析器。
- 1913 一些待发数据仍停留在请求缓冲区内。
- 1914 无效的异步远程过程调用句柄。
- 1915 这个操作的异步 RPC 调用句柄无效。
- 1916 RPC 管道对象已经关闭。
- 1917 在 RPC 调用完成之前全部的管道都已处理完成。
- 1918 没有其他可用的数据来自 RPC 管道。
- 1919 这个机器没有可用的站点名。
- 1920 系统无法访问此文件。
- 1921 系统无法辨识文件名。
- 1922 项目不是所要的类型。
- 1923 无法将所有对象的 UUID 导出到指定的项。
- 1924 无法将接口导出到指定的项。
- 1925 无法添加指定的配置文件项。
- 1926 无法添加指定的配置文件元素。
- 1927 无法删除指定的配置文件元素。
- 1928 无法添加组元素。
- 1929 无法删除组元素。
- 1930 打印机驱动程序与您计算机上启用的阻止 NT 4.0 驱动程序的策略不兼容。
- 1931 上下文已过期，不能再用了。
- 2000 无效的像素格式。
- 2001 指定的驱动程序无效。
- 2002 窗口样式或类别属性对此操作无效。
- 2003 不支持请求的图元操作。
- 2004 不支持请求的变换操作。
- 2005 不支持请求的剪切操作。
- 2010 指定的颜色管理模块无效。
- 2011 制定的颜色文件配置无效。
- 2012 找不到指定的标识。
- 2013 找不到所需的标识。
- 2014 指定的标识已经存在。
- 2015 指定的颜色文件配置与任何设备都不相关。
- 2016 找不到该指定的颜色文件配置
- 2017 指定的颜色空间无效。
- 2018 图像颜色管理没有启动。
- 2019 在删除该颜色传输时有一个错误。
- 2020 该指定的颜色传输无效。
- 2021 该指定的变换与位图的颜色空间不匹配。
- 2022 该指定的命名颜色索引在配置文件中不存在。
- 2102 没有安装工作站驱动程序。

- 2103 无法定位服务器。
- 2104 发生内部错误，网络无法访问共享内存段。
- 2105 网络资源不足。
- 2106 工作站不支持该操作。
- 2107 设备没有连接。
- 2108 网络连接已成功，但需要提示用户输入一个不同于原始指定的密码。
- 2109 使用默认凭据成功连接网络。
- 2114 没有启动服务器服务。
- 2115 队列空。
- 2116 设备或目录不存在。
- 2117 无法在重定向的资源上执行此操作。
- 2118 名称已经共享。
- 2119 服务器目前无法提供所需的资源。
- 2121 额外请求的项目超过允许的上限。
- 2122 对等服务只支持两个同时操作的用户。
- 2123 API 返回的缓冲区太小。
- 2127 远程 API 错误。
- 2131 打开或读取配置文件时出错。
- 2136 发生一般网络错误。
- 2137 工作站服务的状态不一致。重新启动工作站服务之前，请先重新启动计算机。
- 2138 工作站服务没有启动。
- 2139 所需信息不可用。
- 2140 发生 Windows 2000 内部错误。
- 2141 服务器没有配置事务处理。
- 2142 远程服务器不支持请求的 API。
- 2143 事件名无效。
- 2144 网络上已经有此计算机名。请更名后重新启动。
- 2146 配置信息中找不到指定的组件。
- 2147 配置信息中找不到指定的参数。
- 2149 配置文件中有一行太长。
- 2150 打印机不存在。
- 2151 打印作业不存在。
- 2152 找不到打印机目标。
- 2153 打印机目标已经存在。
- 2154 打印机队列已经存在。
- 2155 无法添加其他打印机。
- 2156 无法添加其他打印作业。
- 2157 无法添加其他打印机目标。
- 2158 此打印机目标处于空闲中，不接受控制操作。
- 2159 此打印机目标请求包含无效的控制函数。
- 2160 打印处理程序没有响应。
- 2161 后台处理程序没有运行。
- 2162 打印目标在当前的状态下无法执行这项操作。
- 2163 打印机队列在当前的状态下无法执行这项操作。
- 2164 打印作业在当前的状态下无法执行这项操作。
- 2165 无法为后台处理程序分配内存。
- 2166 设备驱动程序不存在。
- 2167 打印处理程序不支持该数据类型。
- 2168 没有安装打印处理程序。
- 2180 锁定服务数据库。

- 2181 服务表已满。
2182 请求的服务已经启动。
2183 这项服务没有响应控制操作。
2184 服务尚未启动。
2185 服务名无效。
2186 服务没有响应控制功能。
2187 服务控制处于忙碌状态。
2188 配置文件包含无效的服务程序名。
2189 在当前的状态下无法控制服务。
2190 服务异常终止。
2191 这项服务无法接受请求的 "暂停" 或 "停止" 操作。
2192 服务控制 "计划程序" 在 "计划表" 中找不到服务名。
2193 无法读取服务控制计划程序管道。
2194 无法创建新服务的线程。
2200 工作站已经登录到局域网。
2201 工作站没有登录到局域网。
2202 指定的用户名无效。
2203 密码参数无效。
2204 登录处理程序没有添加消息别名。
2205 登录处理程序没有添加消息别名。
2206 注销处理程序没有删除消息别名。
2207 注销处理程序没有删除消息别名。
2209 暂停网络登录。
2210 中心登录服务器发生冲突。
2211 服务器没有配置有效的用户路径。
2212 加载或运行登录脚本时出错。
2214 没有指定登录服务器, 计算机的登录状态是单机操作。
2215 找不到登录服务器。
2216 此计算机已经有一个登录域。
2217 登录服务器无法验证登录。
2219 找不到安全数据库。
2220 找不到组名。
2221 找不到用户名。
2222 找不到资源名。
2223 组已经存在。
2224 帐户已经存在。
2225 资源使用权限列表已经存在。
2226 此操作只能在该域的主域控制器上执行。
2227 安全数据库没有启动。
2228 用户帐户数据库中的名称太多。
2229 磁盘 I/O 失败。
2230 已经超过每个资源 64 个项目的限制。
2231 不得删除带会话的用户。
2232 找不到上层目录。
2233 无法添加到安全数据库会话高速缓存段。
2234 该操作不能在此特殊的组上执行。
2235 用户帐户数据库会话高速缓存没有缓存此用户。
2236 用户已经属于此组。
2237 用户不属于此组。
2238 此用户帐户尚未定义。

- 2239 此用户帐户已过期。
- 2240 此用户不得从此工作站登录网络。
- 2241 此时不允许用户登录网络。
- 2242 此用户的密码已经过期。
- 2243 此用户的密码无法更改。
- 2244 现在无法使用此密码。
- 2245 密码不满足密码策略的要求。检查最小密码长度、密码复杂性和密码历史的要求。
- 2246 此用户的密码最近才启用，现在不能更改。
- 2247 安全数据库已损坏。
- 2248 不需要更新此副本复制的网络/本地安全数据库。
- 2249 此副本复制的数据库已过时；请同步其中的数据。
- 2250 此网络连接不存在。
- 2251 此 asg_type 无效。
- 2252 此设备目前正在共享中。
- 2270 计算机名无法作为消息别名添加。网络上可能已经有此名称。
- 2271 信使服务已经启动。
- 2272 信使服务启动失败。
- 2273 网络上找不到此消息别名。
- 2274 此消息别名已经被转发。
- 2275 已经添加了此消息别名，但是仍被转发。
- 2276 此消息别名已在本地存在。
- 2277 添加的消息别名已经超过数目上限。
- 2278 无法删除计算机名。
- 2279 消息无法转发回到同一个工作站。
- 2280 域消息处理程序出错。
- 2281 消息已经发送，但是收件者已经暂停信使服务。
- 2282 消息已经发送，但没有收到。
- 2283 消息别名目前正在使用中。请稍后再试。
- 2284 信使服务尚未启动。
- 2285 该名称不在本地计算机上。
- 2286 网络上找不到转发的消息别名。
- 2287 远程通讯站的消息别名表已经满了。
- 2288 此别名的消息目前没有在转发中。
- 2289 广播的消息被截断。
- 2294 设备名无效。
- 2295 写入出错。
- 2297 网络上的消息别名重复。
- 2298 此消息别名会在稍后删除。
- 2299 没有从所有的网络删除消息别名。
- 2300 该操作无法在使用多种网络的计算机上执行。
- 2310 共享资源不存在。
- 2311 设备没有共享。
- 2312 带此计算机名的会话不存在。
- 2314 没有用此识别号打开的文件。
- 2315 执行远程管理命令失败。
- 2316 打开远程临时文件失败。
- 2317 从远程管理命令返回的数据已经被截断成 64K。
- 2318 此设备无法同时共享为后台处理资源和非后台处理资源。
- 2319 服务器列表中的信息可能不正确
- 2320 计算机在此域未处于活动状态

- 2321 在删除共享之前，需要将该共享从分布式文件系统中删除。
- 2331 无法在此设备执行该操作
- 2332 此设备无法共享。
- 2333 此设备未打开。
- 2334 此设备名列表无效。
- 2335 队列优先顺序无效。
- 2337 没有任何共享的通讯设备。
- 2338 指定的队列不存在。
- 2340 此设备列表无效。
- 2341 请求的设备无效。
- 2342 后台处理程序正在使用此设备。
- 2343 此设备已经被当成通讯设备来使用。
- 2351 此计算机名无效。
- 2354 指定的字符串及前缀太长。
- 2356 此路径组成部分无效。
- 2357 无法判断输入类型。
- 2362 类型缓冲区不够大。
- 2370 配置文件不得超过 64K。
- 2371 初始偏移量越界。
- 2372 系统无法删除当前到网络资源的连接。
- 2373 系统无法分析此文件中的命令行。
- 2374 加载配置文件时出错。\
2375 保存配置文件时出错，只部份地保存了配置文件。
- 2378 此日志文件在前后两次读取之间已经发生变化。
- 2380 源路径不可以是目录。
- 2381 源路径无效。
- 2382 目标路径无效。
- 2383 源路径及目标路径分属不同的服务器。
- 2385 请求的 Run 服务器现在暂停。
- 2389 与 Run 服务器通讯时出错。
- 2391 启动后台处理时出错。
- 2392 找不到您连接的共享资源。
- 2400 LAN 适配器号码无效。
- 2401 此网络连接有文件打开或请求挂起。
- 2402 使用中的连接仍存在。
- 2403 此共享名或密码无效。
- 2404 设备正由活动进程使用，无法断开。
- 2405 此驱动器号已在本地使用。
- 2430 指定的客户端已经在指定的事件注册。
- 2431 警报表已满。
- 2432 发出的警报名称无效或不存在。
- 2433 警报接收者无效。
- 2434 用户的登录时间长短不再合法。 所以已经删除用户与该服务器的会话。
- 2440 日志文件中没有请求的记录号。
- 2450 用户帐户数据库没有正确配置。
- 2451 当 Netlogon 服务正在运行时，不允许执行该操作。
- 2452 该操作无法在最新的管理帐户上执行。
- 2453 找不到此域的域控制器。
- 2454 无法设置此用户的登录信息。
- 2455 Netlogon 服务尚未启动。

- 2456 无法添加到用户帐户数据库。
- 2457 此服务器的时钟与主域控制器的时钟不一致。
- 2458 检测到密码不匹配。
- 2460 服务器标识没有指定有效的服务器。
- 2461 会话标识没有指定有效的会话。
- 2462 连接标识没有指定有效的连接。
- 2463 可用服务器表中无法添加其他项。
- 2464 服务器已经到了支持的会话数目上限。
- 2465 服务器已经到了支持的连接数目上限。
- 2466 服务器打开的文件到了上限，无法打开更多文件。
- 2467 这台服务器没有登记替换的服务器。
- 2470 请用低版本的 API (远程管理协议)。
- 2480 UPS 服务无法访问 UPS 驱动程序。
- 2481 UPS 服务配置错误。
- 2482 UPS 服务无法访问指定通讯端口 (Comm Port)。
- 2483 UPS 显示线路中断或电池不足，服务没有启动。
- 2484 UPS 服务无法执行系统关机操作。
- 2500 下面的程序返回一个 MS-DOS 错误码:
- 2501 下面的程序需要更多的内存:
- 2502 下面的程序调用了不支持的 MS-DOS 函数:
- 2503 工作站无法启动。
- 2504 下面的文件已损坏。
- 2505 启动块定义文件中没有指定加载程序。
- 2506 NetBIOS 返回错误: NCB 及 SMB 数据转储。
- 2507 磁盘 I/O 错误。
- 2508 无法替换映像参数。
- 2509 跨越磁盘扇区范围的映像参数太多。
- 2510 不是从用 /S 格式化的 MS-DOS 软盘产生的映像。
- 2511 稍后会从远程重新启动。
- 2512 无法调用远程启动服务器。
- 2513 无法连接到远程启动服务器。
- 2514 无法打开远程启动服务器上的映像文件。
- 2515 正在连接到远程启动服务器...
- 2516 正在连接到远程启动服务器...
- 2517 远程启动服务已经停止，请检测错误记录文件，查明出错的原因。
- 2518 远程启动失败，请检查错误日志文件，查明出错的原因。
- 2519 不允许第二个远程启动 (Remoteboot) 资源连接。
- 2550 浏览服务配置成 MaintainServerList=No。
- 2610 因为没有网卡与这项服务一起启动，所以无法启动服务。
- 2611 因为注册表中的启动信息不正确，所以无法启动服务。
- 2612 因为服务的数据库找不到或损坏，所以无法启动服务。
- 2613 因为找不到 RPLFILES 共享，所以无法启动服务。
- 2614 因为找不到 RPLUSER 组，所以无法启动服务。
- 2615 无法枚举服务记录。
- 2616 工作站记录信息已损坏。
- 2617 找不到工作站记录。
- 2618 其他工作站正在使用此工作站名。
- 2619 配置文件记录信息已损坏。
- 2620 找不到配置文件记录。
- 2621 其他配置文件正在使用此名称。

- 2622 有很多工作站正在使用此配置文件。
2623 配置记录信息已损坏。
2624 找不到配置记录。
2625 适配器 ID 记录信息已损坏。
2626 内部服务出错。
2627 供应商 ID 记录信息已损坏。
2628 启动块记录信息已损坏。
2629 找不到此工作站记录的用户帐户。
2630 找不到 RPLUSER 本地组。
2631 找不到启动块记录。
2632 所选的配置文件与此工作站不兼容。
2633 其他工作站正在使用所选的网卡 ID。
2634 有些配置文件正在使用此配置。
2635 有数个工作站、配置文件或配置正在使用此启动块。
2636 服务无法制作远程启动数据库的备份。
2637 找不到适配器记录。
2638 找不到供应商记录。
2639 其他供应商记录正在使用此供应商名称。
2640 其他启动块记录正在使用启动名称或供应商 ID。
2641 其他配置正在使用此配置名称。
2660 由 DFS 服务所维护的内部数据库已损坏
2661 内部 DFS 数据库中的一条记录已损坏
2662 没有 DFS 名称的项目路径与输入项目路径匹配
2663 给定名称的根或链接已存在
2664 指定的服务器共享已在 DFS 中共享
2665 所显示的服务器共享不支持所显示的 DFS 名称空间
2666 此操作在这一部分的名称空间上无效
2667 此操作在这一部分的名称空间上无效
2668 此操作不明确，因为链接有多个服务器
2669 无法创建链接
2670 该服务器不是 DFS 可识别的
2671 指定的重命名目标路径无效。
2672 指定的 DFS 链接脱机
2673 指定的服务器不是此链接的服务器
2674 检测到 DFS 名中的环路
2675 在基于服务器的 DFS 上不支持该操作
2676 此链接已经受该指定服务器共享支持
2677 无法删除这个根或链接的最新服务器共享支持
2678 DFS 间的链接不支持该操作
2679 DFS 服务的内部状态已经变得不一致
2680 DFS 服务已经安装在指定的服务器上
2681 被协调的 DFS 数据是一样的
2682 无法删除 DFS 根目录 - 如需要请卸载 DFS
2683 该共享的子目录或父目录已经存在在一个 DFS 中
2690 DFS 内部错误
2691 这台机器已经加入域。
2692 这个机器目前未加入域。
2693 这台机器是域控制器，而且无法从域中退出。
2694 目标域控制器不支持在 OU 中创建的机器帐户。
2695 指定的工作组名无效。

- 2696 指定的计算机名与域控制器上使用的默认语言不兼容。
- 2697 找不到指定的计算机帐户。
- 2698 此版本的 Windows 不能加入到域。
- 2999 这是 NERR 范围内的最后一个错误。
- 3000 指定了未知的打印监视器。
- 3001 指定的打印机驱动程序当前正在使用。
- 3002 找不到缓冲文件。
- 3003 未发送 StartDocPrinter 调用。
- 3004 未发送 AddJob 调用。
- 3005 指定的打印处理器已经安装。
- 3006 指定的打印监视器已经安装。
- 3007 该指定的打印监视器不具备所要求的功能。
- 3008 该指定的打印监视器正在使用中。
- 3009 当打印机有作业排成队列时此操作请求是不允许的。
- 3010 请求的操作成功。直到重新启动系统前更改将不会生效。
- 3011 请求的操作成功。直到重新启动服务前更改将不会生效。
- 3012 找不到打印机。
- 3013 已知打印机驱动程序不可靠。
- 3014 已知打印机驱动程序会伤害系统。
- 3023 用户指定的关机命令文件的配置有问题。不过 UPS 服务已经启动。
- 3029 因为用户帐户数据库 (NET.ACC) 找不到或损坏, 而且也没有可用的备份数据库, 所以不能启动本地安全机制。系统不安全!
- 3037 @I *登录小时数
- 3039 已经超过一个目录中文件的副本复制的限制。
- 3040 已经超过副本复制的目录树深度限制。
- 3046 无法登录。用户目前已经登录, 同时参数 TRYUSER 设置为 NO。
- 3052 命令行或配置文件中 没有提供必要的参数。
- 3054 无法满足资源的请求。
- 3055 系统配置有问题。
- 3056 系统出错。
- 3057 发生内部一致性的错误。
- 3058 配置文件或命令行的选项不明确。
- 3059 配置文件或命令行的参数重复。
- 3060 服务没有响应控制, DosKillProc 函数已经停止服务。
- 3061 运行服务程序时出错。
- 3062 无法启动次级服务。
- 3064 文件有问题。
- 3070 内存
- 3071 磁盘空间
- 3072 线程
- 3073 过程
- 3074 安全性失败。
- 3075 LAN Manager 根目录不正确或找不到。
- 3076 未安装网络软件。
- 3077 服务器未启动。
- 3078 服务器无法访问用户帐户数据库 (NET.ACC)。
- 3079 LANMAN 树中安装的文件不兼容。
- 3080 LANMAN\LOGS 目录无效。
- 3081 指定的域无法使用。
- 3082 另一计算机正将此计算机名当作消息别名使用。

- 3083 宣布服务器名失败。
- 3084 用户帐户数据库没有正确配置。
- 3085 服务器没有运行用户级安全功能。
- 3087 工作站配置不正确。
- 3088 查看您的错误日志文件以了解详细信息。
- 3089 无法写入此文件。
- 3090 ADDPAK 文件损坏。请删除 LANMAN\NETPROG\ADDPAK.SER 后重新应用 ADDPAK。
- 3091 因为没有运行 CACHE.EXE, 所以无法启动 LM386 服务器。
- 3092 安全数据库中找不到这台计算机的帐户。
- 3093 这台计算机不是 SERVERS 组的成员。
- 3094 SERVERS 组在本地安全数据库中不存在。
- 3095 此计算机被配置为某个组的成员, 并不是域的成员。此种配置下不需要运行 Netlogon 服务。
- 3096 找不到此域的主域控制器。
- 3098 服务无法与主域控制器进行验证。
- 3099 安全数据库文件创建日期或序列号有问题。
- 3100 因为网络软件出错, 所以无法执行操作。
- 3102 这项服务无法获得网络控制块 (NCB) 段的长期锁定。 错误码就是相关数据。
- 3103 这项服务无法解除网络控制块 (NCB) 段的长期锁定。 错误码就是相关数据。
- 3106 收到意外的网络控制块 (NCB)。NCB 就是相关数据。
- 3107 网络没有启动。
- 3108 NETWKSTA.SYS 的 DosDevIoctl 或 DosFsCtl 调用失败。 显示的数据为以下
式:DWORD 值代表调用 Ioctl 或 FsCtl 的 CS:IP WORD 错误代码 WORD Ioctl 或
FsCtl 号
- 3111 发生意外的 NetBIOS 错误。 错误码就是相关数据。
- 3112 收到的服务器消息块 (SMB) 无效。SMB 就是相关数据。
- 3114 因为缓冲区溢出, 所以错误日志文件中部份的项目丢失。
- 3120 控制网络缓冲区以外资源用量的初始化参数被设置大小, 因此需要的内存太多。
- 3121 服务器无法增加内存段的大小。
- 3124 服务器启动失败。三个 chdev 参数必须同时为零或者同时不为零。
- 3129 服务器无法更新 AT 计划文件。 文件损坏。
- 3130 服务器调用 NetMakeLMFileName 时出错。 错误码就是相关数据。
- 3132 无法长期锁定服务器缓冲区。请检查交换磁盘的可用空间, 然后重启系统以启动服务器。
- 3140 因为多次连续出现网络控制块 (NCB) 错误, 所以停止服务。 最后一个坏的 NCB 以原始数据形式出现。
- 3141 因为消息服务器共享的数据段被锁住, 所以消息服务器已经停止运行。
- 3151 因为 VIO 调用出错, 所以无法显示 POPUP 消息。 错误码就是相关数据。
- 3152 收到的服务器消息块 (SMB) 无效。SMB 就是相关数据。
- 3160 工作站信息段大于 64K。 大小如下(以 DWORD 值的格式):
- 3161 工作站无法取得计算机的名称号码。
- 3162 工作站无法初始化 Async NetBIOS 线程。 错误码就是相关数据。
- 3163 工作站无法打开最前面的共享段。 错误码就是相关数据。
- 3164 工作站主机表已满。
- 3165 收到的邮筒服务器消息块 (SMB) 有问题, SMB 就是相关数据。
- 3166 工作站启动用户帐户数据库时出错。 错误码就是相关数据。
- 3167 工作站响应 SSI 重新验证请求时出错。 函数码及错误码就是相关数据。
- 3174 服务器无法读取 AT 计划文件。
- 3175 服务器发现无效的 AT 计划记录。
- 3176 服务器找不到 AT 计划文件, 所以创建了一个计划文件。
- 3185 因为用户帐户数据库 (NET.ACC) 找不到或损坏, 而且也没有可用的备份数据库, 所以不能启动本地安全机制。 系统不安全!

- 3204 服务器无法创建线程。CONFIG.SYS 中的 THREADS 参数必须加大。
- 3213 已经超过一个目录中文件的副本复制的限制。
- 3214 已经超过副本复制的目录树深度限制。
- 3215 邮筒中收到的消息无法识别。
- 3217 无法登录。用户当前已经登录，同时参数 TRYUSER 设置为 NO。
- 3230 检测到服务器的电源中断。
- 3231 UPS 服务已经关掉服务器。
- 3232 UPS 服务没有完成执行用户指定的 关机命令文件。
- 3233 无法打开 UPS 驱动程序。 错误码就是相关数据。
- 3234 电源已经恢复。
- 3235 用户指定的关机命令文件 有问题。
- 3256 该服务的动态链接库 发生无法修复的错误。
- 3257 系统返回意外的错误码。 错误码就是相关数据。
- 3258 容错错误日志文件 - LANROOTLOGS\FT.LOG 超过 64K。
- 3259 容错错误日志文件 - LANROOTLOGS\FT.LOG, 在被打开时 就已设置更新进度位, 这表示上次使用错误日志时, 系统死机。

- 3301 Remote IPC
- 3302 Remote Admin
- 3303 Logon server share
- 3304 网络出错。
- 3400 内存不足，无法启动工作站服务。
- 3401 读取 LAMAN.INI 文件的 NETWORKS 项目出错。
- 3404 LAMAN.INI 文件中的 NETWORKS 项目太多。
- 3408 程序无法用在此操作系统。
- 3409 已经安装重定向程序。
- 3411 安装 NETWKSTA.SYS 时出错。 请按 ENTER 继续。
- 3412 求解程序链接问题。
- 3419 您已经打开文件或设备， 强制断开会造成数据丢失。
- 3420 内部用的默认共享
- 3421 信使服务
- 3500 命令成功完成。
- 3501 使用的选项无效。
- 3503 命令包含无效的参数个数。
- 3504 命令运行完毕，但发生一个或多个错误。
- 3505 使用的选项数值不正确。
- 3510 命令使用了冲突的选项。
- 3512 软件需要新版的操作系统。
- 3513 数据多于 Windows 所能够返回的。
- 3515 此命令只能用于 Windows 域控制器。
- 3516 此命令不能用于 Windows 域控制器。
- 3520 已经启动以下 Windows 服务:
- 3525 停止工作站服务也会同时停止服务器服务 。
- 3526 工作站有打开的文件。
- 3533 服务正在启动或停止中，请稍候片刻后再试一次。
- 3534 服务没有报告任何错误。
- 3535 正在控制设备时出错。
- 3660 下列工作站在这台服务器上有会话:
- 3661 下列工作站在这台服务器上有会话并打开了文件:
- 3666 消息别名已经转发出去。
- 3670 您有以下的远程连接:
- 3671 继续运行会取消连接。

- 3676 会记录新的网络连接。
3677 不记录新的网络连接。
3678 保存配置文件时出错, 原先记录的网络连接状态没有更改。
3679 读取配置文件时出错。
3682 没有启动任何网络服务。
3683 列表是空的。
3689 工作站服务已经在运行中, Windows 会忽略工作站的命令选项。
3694 在打印作业正在后台处理到队列时, 无法删除共享的队列。
3710 打开帮助文件时出错。
3711 帮助文件是空的。
3712 帮助文件已经损坏。
3714 这是专为那些安装旧版软件的系统 提供的操作。
3716 设备类型未知。
3717 日志文件已经损坏。
3718 程序文件名后必须以 .EXE 结束。
3719 找不到匹配的共享, 因此没有删除。
3720 用户记录中的“单位/星期”字段的值不正确。
3725 删除共享时出错。
3726 用户名无效。
3727 密码无效。
3728 密码不匹配。
3729 永久连接没有完全还原。
3730 计算机名或域名无效。
3732 无法设置该资源的默认权限。
3734 没有输入有效的密码。
3735 没有输入有效的名称。
3736 该命名资源无法共享。
3737 权限字符串包含无效的权限。
3738 您只能在打印机或通讯设备上执行这项操作。
3743 服务器没有配置远程管理的功能。
3752 这台服务器上没有用户的会话。
3756 响应无效。
3757 没有提供有效的响应。
3758 提供的目标列表与打印机队列目标列表不匹配。
3761 指定的时间范围中结束的时间比开始的时间早。
3764 提供的时间不是整点。
3765 12 与 24 小时格式不能混用。
3767 提供的日期格式无效。
3768 提供的日期范围无效。
3769 提供的时间范围无效。
3770 NET USER 的参数无效。请检查最短的密码长度 和/或提供参数。
3771 ENABLESCRIPT 的值必须是 YES。
3773 提供的国家(地区)代码无效。
3774 用户已经创建成功, 但是无法添加到 USERS 本地组中。
3775 提供的用户上下文无效。
3777 文件发送功能已不再支持。
3778 您可能没有指定 ADMIN\$ 及 IPC\$ 共享的路径。
3784 只有磁盘共享可以标记为可以缓存。
3802 此计划日期无效。
3803 LANMAN 根目录不可用。

- 3804 SCHED.LOG 文件无法打开。
- 3805 服务器服务尚未启动。
- 3806 AT 作业 ID 不存在。
- 3807 AT 计划文件已损坏。
- 3808 因为 AT 计划文件发生问题，所以无法运行删除操作。
- 3809 命令行不得超过 259 个字符。
- 3810 因为磁盘已满，所以 AT 计划文件无法更新。
- 3812 AT 计划文件无效。请删除此文件并创建新的文件。
- 3813 AT 计划文件已经删除。
- 3814 此命令的语法是:AT [id] [/DELETE]AT 时间 [/EVERY:日期 |/NEXT:日期] 命令 AT 命令会在以后的 指定日期及时间，安排程序在服务器上运行。它也会显示安排运行的程序及命令的清单。您可以将日期指定为 M、T、W、Th、F、Sa、Su 或 1-31 的格式。您可以将时间指定为 HH:MM 的二十四小时格式。
- 3815 AT 命令已经超时。请稍后再试一次。
- 3816 用户帐户的密码最短使用期限不得 大于密码最长使用期限。
- 3817 指定的数值与安装下层软件的服务器不兼容。请指定较小的值。
- 3901 ****
- 3902 **** 意外到达消息的结尾 ****
- 3905 请按 ESC 退出
- 3906 ...
- 3912 找不到时间服务器。
- 3915 无法判断用户的主目录。
- 3916 没有指定用户的主目录。
- 3920 已经没有可用的驱动器号。
- 3936 这台计算机目前没有配置成使用一个指定的 SNTP 服务器。
- 3953 语法错误。
- 3960 指定的文件号码无效。
- 3961 指定的打印作业号码无效。
- 3963 指定的用户或组帐户找不到。
- 3965 已添加用户，但 NetWare 的文件和打印服务无法启用。
- 3966 没有安装 NetWare 的文件和打印服务。
- 3967 无法为 NetWare 的文件和打印服务设置用户属性。
- 3969 NetWare 兼容登录
- 4000 WINS 在处理命令时遇到错误。
- 4001 本地的 WINS 不能删除。
- 4002 文件导入操作失败。
- 4003 备份操作失败。是否先前已作过完整备份？
- 4004 备份操作失败。请检查您备份数据库的目录。
- 4005 WINS 数据库中没有一个名称。
- 4006 不允许复制一个尚未配置的伙伴。
- 4100 DHCP 客户端获得一个在网络上已被使用的 IP 地址。直到 DHCP 客户端可以获得新的地址前，本地接口将被禁用。
- 4200 无法识别传来的 GUID 是否为有效的 WMI 数据提供程序。
- 4201 无法识别传来的范例名是否为有效的 WMI 数据提供程序。
- 4202 无法识别传来的数据项目标识符是否为有效的 WMI 数据提供程序。
- 4203 无法完成 WMI 请求，应该重试一次。
- 4204 找不到 WMI 数据提供程序。
- 4205 WMI 数据提供程序引用到一个未注册的范例组。
- 4206 WMI 数据块或事件通知已启用。
- 4207 WMI 数据块不再可用。
- 4208 WMI 数据服务无法使用。

4209	WMI 数据提供程序无法完成要求。
4210	WMI MOF 信息无效。
4211	WMI 注册信息无效。
4212	WMI 数据块或事件通知已禁用。
4213	WMI 数据项目或数据块为只读。
4214	不能更改 WMI 数据项目或数据块。
4300	媒体标识符没有表示一个有效的媒体。
4301	库标识符没有表示一个有效的库。
4302	媒体缓冲池标识符没有表示一个有效的媒体缓冲池。
4303	驱动器和媒体不兼容或位于不同的库中。
4304	媒体目前在脱机库中，您必须联机才能运行这个操作。
4305	操作无法在脱机库中运行。
4306	库、驱动器或媒体缓冲池是空的。
4307	库、磁盘或媒体缓冲池必须是空的，才能运行这个操作。
4308	在这个媒体缓冲池或库中目前没有可用的媒体。
4309	这个操作所需的资源已禁用。
4310	媒体标识符没有表示一个有效的清洗器。
4311	无法清洗驱动器或不支持清洗。
4312	对象标识符没有表示一个有效的对象。
4313	无法读取或写入数据库。
4314	数据库已满。
4315	媒体与设备或媒体缓冲池不兼容。
4316	这个操作所需的资源不存在。
4317	操作标识符不正确。
4318	媒体未被安装，或未就绪。
4319	设备未就绪。
4320	操作员或系统管理员拒绝了请求。
4321	驱动器标识符不代表一个有效的驱动器。
4322	程序库已满。没有可使用的插槽。
4323	传输程序不能访问媒体。
4324	无法将媒体加载到驱动器中。
4325	无法检索有关驱动器的状态。
4326	无法检索有关插槽的状态。
4327	无法检索传输的状态。
4328	因为传输已在使用中，所以无法使用。
4329	无法打开或关闭弹入/弹出端口。
4330	因为媒体在驱动器中，无法将其弹出。
4331	清洗器插槽已被保留。
4332	没有保留清洗器插槽。
4333	清洗器墨盒已进行了最大次数的驱动器清洗。
4334	意外媒体标识号。
4335	在这个组或源中最后剩下的项目不能被删除。
4336	提供的消息超过了这个参数所允许的最大尺寸。
4337	该卷含有系统和页面文件。
4338	由于库中至少有一个驱动器可以支持该媒体类型，不能从库中删除媒体类型。
4339	由于没有可以使用的已被启动的驱动器，无法将该脱机媒体装入这个系统。
4340	清洗磁带在磁带库中。
4341	(Y/N) [N]
4342	错误
4343	OK
4344	Y

- 4345 N
- 4346 任何
- 4347 A
- 4348 P
- 4349 (找不到)
- 4350 远程存储服务无法撤回文件。
- 4351 远程存储服务此时不可操作。
- 4352 远程存储服务遇到一个媒体错误。
- 4354 请键入密码:
- 4358 请键入用户的密码:
- 4359 请键入共享资源的密码:
- 4360 请键入您的密码:
- 4361 请再键入一次密码以便确认:
- 4362 请键入用户的旧密码:
- 4363 请键入用户的新密码:
- 4364 请键入您的新密码:
- 4365 请键入复制程序服务密码:
- 4368 请键入您的用户名:
- 4372 打印作业详细信息
- 4378 控制下列正在运行的服务:
- 4379 统计数据可用于下列正在运行的服务:
- 4381 此命令的语法是:
- 4382 此命令的选项是:
- 4383 请键入主域控制器的名称:
- 4385 Sunday
- 4386 Monday
- 4387 Tuesday
- 4388 Wednesday
- 4389 Thursday
- 4390 此文件或目录不是一个重解析点。
- 4391 重解析点的属性不能被设置,因为它与已有的属性冲突。
- 4392 在重解析点缓冲区中的数据无效。
- 4393 在重解析点缓冲区中的标签无效。
- 4394 请求中指定的标签和重解析点中的不匹配。
- 4395 W
- 4396 Th
- 4397 F
- 4398 S
- 4399 Sa
- 4401 组名
- 4402 注释
- 4403 成员
- 4406 别名
- 4407 注释
- 4408 成员
- 4411 用户名
- 4412 全名
- 4413 注释
- 4414 用户的注释
- 4415 参数
- 4416 国家(地区)代码
- 4417 权限等级

4418	操作员权限
4419	帐户启用
4420	帐户到期
4421	上次设置密码
4422	密码到期
4423	密码可更改
4424	允许的工作站
4425	磁盘空间上限
4426	无限制
4427	本地组成员
4428	域控制器
4429	登录脚本
4430	上次登录
4431	全局组成员
4432	可允许的登录小时数
4433	全部
4434	无
4436	主目录
4437	需要密码
4438	用户可以更改密码
4439	用户配置文件
4440	已锁定
4450	计算机名
4451	用户名
4452	软件版本
4453	工作站正运行于
4454	Windows NT 根目录
4455	工作站域
4456	登录域
4457	其它域
4458	COM 打开超时 (秒)
4459	COM 发送计数 (字节)
4460	COM 发送超时 (毫秒)
4461	DOS 会话打印超时 (秒)
4462	错误日志文件大小上限 (K)
4463	高速缓存上限 (K)
4464	网络缓冲区数
4465	字符缓冲区数
4466	域缓冲区大小
4467	字符缓冲区大小
4468	计算机全名
4469	工作站域 DNS 名称
4470	Windows 2002
4481	服务器名称
4482	服务器注释
4483	发送管理警报到
4484	软件版本
4485	对等服务器
4486	Windows NT
4487	服务器等级
4488	Windows NT Server

- 4489 服务器正运行于
- 4492 服务器已隐藏
- 4500 零备份存储在这个卷上不可用。
- 4506 登录的用户数量上限
- 4507 同时可并存的管理员数量上限
- 4508 共享的资源数量上限
- 4509 资源连接数量上限
- 4510 服务器打开的文件数量上限
- 4511 每个会话打开的文件数量上限
- 4512 文件锁定数量上限
- 4520 空闲的会话时间 (分)
- 4526 共享等级
- 4527 用户等级
- 4530 未限制的服务器
- 4570 强制用户在时间到期之后多久必须注销?:
- 4571 多少次密码不正确后锁住帐户?:
- 4572 密码最短使用期限 (天):
- 4573 密码最长使用期限 (天):
- 4574 密码长度下限:
- 4575 保持的密码历史记录长度:
- 4576 计算机角色:
- 4577 工作站域的主域控制器:
- 4578 锁定阈值:
- 4579 锁定持续时间(分):
- 4580 锁定观测窗口(分):
- 4600 统计开始于
- 4601 接受的会话
- 4602 会话超时
- 4603 会话出错
- 4604 发送的 KB
- 4605 接收的 KB
- 4606 平均响应时间 (毫秒)
- 4607 网络错误
- 4608 访问的文件
- 4609 后台处理的打印作业
- 4610 系统出错
- 4611 密码违规
- 4612 权限违规
- 4613 访问的通讯设备
- 4614 启动的会话
- 4615 重新连接的会话
- 4616 启动失败的会话
- 4617 断开的会话
- 4618 网络 I/O 执行
- 4619 访问的文件及管道
- 4620 耗尽的时间缓冲区
- 4621 大缓冲区
- 4622 请求缓冲区
- 4626 已做连接
- 4627 失败的连接

4630	接收的字节数
4631	接收的服务器消息块 (SMB)
4632	传输的字节数
4633	传输的服务器消息块 (SMB)
4634	读取操作
4635	写入操作
4636	拒绝原始读取
4637	拒绝原始写入
4638	网络错误
4639	已做连接
4640	重新连接
4641	服务器断开
4642	已启动的会话
4643	会话挂起
4644	失败的会话
4645	失败的操作
4646	使用计数
4647	使用计数失败
4655	消息名称转发已经取消。
4661	密码已经更改成功。
4664	消息已经发给网络上所有的用户。
4666	消息已经送到服务器上的所有用户。
4696	Windows NT Server
4697	Windows NT Workstation
4698	MS-DOS 增强型工作站
4700	服务器名称 注释
4702	(UNC)
4703	...
4704	Domain
4706	无效的网络提供程序。可用的网络为:
4710	Disk
4711	Print
4712	Comm
4713	IPC
4714	状态 本地 远程 网络
4715	OK
4716	休止
4717	已暂停
4718	已断开
4719	错误
4720	正在连接
4721	正在重新连接
4722	状态
4723	本地名称
4724	远程名称
4725	资源类型
4726	# 打开
4727	# 连接
4728	不可用
4730	共享名 资源 注释
4731	共享名
4732	资源

4733	后台处理
4734	权限
4735	用户上限
4736	无限制
4737	用户
4739	正在缓存
4740	ID 路径 用户名 # 锁定
4741	文件 ID
4742	锁定
4743	权限
4744	共享名
4745	类型
4746	使用为
4747	注释
4750	计算机 用户名 客户类型 打开空闲时间
4751	计算机
4752	会话时间
4753	空闲时间
4754	资源共享名 类型 # 打开
4755	客户类型
4756	来宾登录
4770	手动缓存文档
4771	自动缓存文档
4772	自动缓存程序文档
4773	缓存被停用
4774	自动
4775	手动
4776	文档
4777	程序
4778	无
4800	名称
4801	转发到
4802	已经从下列位置转发给您
4803	这台服务器的用户
4804	用户已经按 Ctrl+Break 中断 Net Send。
4810	名称 作业编号 大小 状态
4811	作业
4812	打印
4813	名称
4814	作业 #
4815	大小
4816	状态
4817	分隔文件
4818	注释
4819	优先顺序
4820	打印后于
4821	打印直到
4822	打印处理程序
4823	附加信息
4824	参数

4825	打印设备
4826	打印机活动中
4827	打印机搁置
4828	打印机出错
4829	正在删除打印机
4830	打印机状态未知
4841	作业 #
4842	正在提交用户
4843	通知
4844	作业数据类型
4845	作业参数
4846	正在等候
4847	搁置于队列
4848	正在后台处理
4849	已暂停
4850	脱机
4851	错误
4852	缺纸
4853	需要干预
4854	正在打印
4855	on
4862	驱动程序
4930	用户名 类型 日期
4931	锁定
4932	服务
4933	服务器
4934	服务器已启动
4935	服务器已暂停
4936	服务器已继续操作
4937	服务器已停止
4938	会话
4939	登录来宾
4940	登录用户
4941	登录管理员
4942	正常注销
4943	登录
4944	注销错误
4945	注销自动断开
4946	注销管理员断开
4947	注销受登录限制
4948	服务
4957	帐户
4964	已修改帐户系统设置
4965	登录限制
4966	超过限制: 未知
4967	超过限制: 登录时间
4968	超过限制: 帐户过期
4969	超过限制: 工作站 ID 无效
4970	超过限制: 帐户停用
4971	超过限制: 帐户已删除
4972	资源

- 4978 密码不正确
- 4979 需要管理员特权
- 4980 访问
- 4984 拒绝访问
- 4985 未知
- 4986 其它
- 4987 持续时间:
- 4988 持续时间: 不可用
- 4989 持续时间: 1 秒以下
- 4990 (无)
- 4994 访问结束
- 4995 登录到网络
- 4996 拒绝登录
- 4997 程序 消息 时间
- 4999 管理员已解除帐户的锁定状态
- 5000 注销网络
- 5001 因为其他资源需要它, 不能将群集资源移到另一个组。
- 5002 找不到此群集资源的依存。
- 5003 因为已经处于依存状态, 此群集资源不能依存于指定的资源。
- 5004 此群集资源未联机。
- 5005 此操作没有可用的群集节点。
- 5006 没有群集资源。
- 5007 找不到群集资源。
- 5008 正在关闭群集。
- 5009 除非节点处于关闭状态或是最后一个节点, 群集节点无法从群集中脱离。
- 5010 对象已存在。
- 5011 此对象已在列表中。
- 5012 新请求没有可用的群集组。
- 5013 找不到群集组。
- 5014 因为群集组未联机, 此操作不能完成。
- 5015 群集节点不是此资源的所有者。
- 5016 群集节点不是此资源的所有者。
- 5017 群集资源不能在指定的资源监视器中创建。
- 5018 群集资源不能通过资源监视器来联机。
- 5019 因为群集资源联机, 此操作不能完成。
- 5020 由于是仲裁资源, 群集资源不能被删除或脱机。
- 5021 由于没有能力成为仲裁资源, 此群集不能使指定资源成为仲裁资源。
- 5022 群集软件正关闭。
- 5023 组或资源的状态不是执行请求操作的正确状态。
- 5024 属性已被存储, 但在下次资源联机前, 不是所有的修改将生效。
- 5025 由于不属于共享存储类别, 群集不能使指定资源成为仲裁资源。
- 5026 由于是内核资源, 无法删除群集资源。
- 5027 仲裁资源联机失败。
- 5028 无法成功创建或装入仲裁日志。
- 5029 群集日志损坏。
- 5030 由于该日志已超出最大限量, 无法将记录写入群集日志。
- 5031 群集日志已超出最大限量。
- 5032 群集日志没有发现检查点记录。
- 5033 不满足日志所需的最小磁盘空间。
- 5034 群集节点未能控制仲裁资源, 因为它被另一个活动节点拥有。

- 5035 这个操作的群集网络无效。
- 5036 此操作没有可用的群集节点。
- 5037 所有群集节点都必须运行才能执行这个操作。
- 5038 群集资源失败。
- 5039 该群集节点无效。
- 5040 该群集节点已经存在。
- 5041 一个节点正在加入该群集。
- 5042 找不到群集节点。
- 5043 找不到群集本地节点信息。
- 5044 群集网络已经存在。
- 5045 找不到群集网络。
- 5046 群集网络接口已经存在。
- 5047 找不到群集网络接口。
- 5048 群集请求在这个对象中无效。
- 5049 群集网络提供程序无效。
- 5050 群集节点坏了。
- 5051 无法连接到群集节点。
- 5052 该群集节点不是群集的一个成员。
- 5053 目前没有执行群集加入操作。
- 5054 该群集网络无效。
- 5055 **Mar**
- 5056 该群集节点可以使用。
- 5057 该群集 IP 地址已在使用中。
- 5058 该群集节点没有中止。
- 5059 没有有效的群集安全上下文。
- 5060 该群集网络不是为内部群集通讯配置的。
- 5061 群集节点已经开始。
- 5062 群集节点已经坏了。
- 5063 群集网络已经联机。
- 5064 群集网络已经脱机。
- 5065 群集节点已经是该群集的成员。
- 5066 该群集网络是为两个或更多活动群集节点进行内部群集通讯的唯一配置。不能从网络上删除内部通讯能力。
- 5067 一个或更多的群集资源依靠网络来向客户端提供服务。不能从网络上删除客户端访问能力。
- 5068 该操作不能在群集资源上作为仲裁资源执行。您不能将仲裁资源脱机或修改其所有者名单。
- 5069 该群集仲裁资源不允许有任何依存关系。
- 5070 该群集节点暂停。
- 5071 群集资源不能联机。所有者节点不能在这个资源上运行。
- 5072 群集节点没有准备好，不能执行所请求的操作。
- 5073 群集节点正在关闭。
- 5074 放弃群集节点加入操作。
- 5075 由于加入节点和支持者之间的软件版本不兼容，该群集加入操作失败。
- 5076 由于该群集已经到达其所能监督的资源限制，不能创建这个资源。
- 5077 系统配置在群集加入或形成操作时已更改。放弃加入或形成操作。
- 5078 找不到指定的资源类型。
- 5079 指定的节点不支持这类资源，这也许是由于版本不一致或是由于在这个节点上没有资源 DLL。
- 5080 该资源 DLL 不支持指定的资源名称。这可能是由于一个提供给源 DLL 名称是错误的(或经过更改的)。
- 5081 不能在 RPC 服务器上注册任何身份验证包。
- 5082 由于组的所有者不在组的首选列表中，不能将组联机。要改变组的所有者节点，请移动组。

- 5083 群集数据库的系列号已改变, 或者与锁定程序节点不相容, 因此加入操作没有成功。如果在加入操作期间群集数据库有任何改动, 这都可能发生。
- 5084 资源在其当前状态下, 资源监视器不允许执行失败操作。资源处于挂起状态时, 这都可能发生。
- 5085 非锁定程序代码收到一个为全局更新保留锁定的请求。
- 5086 群集服务找不到仲裁磁盘。
- 5087 已备份的群集数据库可能已损坏。
- 5088 DFS 根目录已在这个群集节点中。
- 5089 由于与另一个现有属性冲突, 未能修改资源属性。
- 5090 西班牙
- 5091 丹麦
- 5092 瑞典
- 5093 挪威
- 5094 德国
- 5095 澳大利亚
- 5096 日本
- 5097 韩国
- 5098 中国
- 5099 台湾
- 5100 亚洲
- 5101 葡萄牙
- 5102 芬兰
- 5103 阿拉伯
- 5104 希伯来
- 5153 UPS 服务即将执行最后的关机操作。
- 5170 工作站必须用 NET START 才能启动。
- 5175 远程 IPC
- 5176 远程管理
- 5177 默认共享
- 5291 从不
- 5292 从不
- 5293 从不
- 5295 NETUS.HLP
- 5296 NET.HLP
- 5300 网络控制块 (NCB) 请求运行成功。NCB 是相关数据。
- 5301 SEND DATAGRAM、SEND BROADCAST、ADAPTER STATUS 或 SESSION STATUS 的网络控制块 (NCB) 缓冲区长度无效。NCB 是相关数据。
- 5302 网络控制块 (NCB) 指定的数据描述数组无效。NCB 是相关数据。
- 5303 网络控制块 (NCB) 指定的命令无效。NCB 是相关数据。
- 5304 网络控制块 (NCB) 指定的消息交换码无效。NCB 是相关数据。
- 5305 网络控制块 (NCB) 命令超时。会话可能异常终止。NCB 是相关数据。
- 5306 接收的网络控制块 (NCB) 消息不完整。NCB 是相关数据。
- 5307 网络控制块 (NCB) 指定的缓冲区地址无效。NCB 是相关数据。
- 5308 网络控制块 (NCB) 指定的会话号码不活动。NCB 是相关数据。
- 5309 网卡没有任何资源可用。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5310 网络控制块 (NCB) 指定的会话已经关闭。NCB 是相关数据。
- 5311 网络控制块 (NCB) 命令已经取消。NCB 是相关数据。
- 5312 网络控制块 (NCB) 指定的消息段不合逻辑。NCB 是相关数据。
- 5313 该名称已经存在于本地适配器名称表中。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5314 网卡名称表已满。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。

- 5315 网络名称已经有活动的会话,现在取消注册。网络控制块 (NCB) 命令运行完毕。NCB 是相关数据。
- 5316 先前发出的 Receive Lookahead 命令对此会话仍然活动。网络控制块 (NCB)命令被拒绝。NCB 是相关数据。
- 5317 本地会话表已满。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5318 拒绝打开网络控制块 (NCB) 会话, 远程计算机上没有侦听命令在执行。NCB 是相关数据。
- 5319 网络控制块 (NCB) 指定的名称号码无效。NCB 是相关数据。
- 5320 网络控制块 (NCB) 中指定的调用名称找不到, 或者没有应答。NCB 是相关数据。
- 5321 网络控制块 (NCB) 中指定的名称找不到。无法将 “*” 或 00h 填入 NCB 名称。NCB 是相关数据。
- 5322 网络控制块 (NCB) 中指定的名称正用于远程适配器。NCB 是相关数据。
- 5323 网络控制块 (NCB) 中指定的名称已经删除。NCB 是相关数据。
- 5324 网络控制块 (NCB) 中指定的会话异常终止。NCB 是相关数据。
- 5325 网络协议在网络上检测到两个或数个 相同的名称。网络控制块 (NCB) 是相关数据。
- 5326 收到意外的协议数据包。远程设备可能不兼容。网络控制块 (NCB) 是相关数据。
- 5333 NetBIOS 界面正忙。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5334 未完成的网络控制块 (NCB) 命令太多。NCB 请求被拒绝。NCB 是相关数据。
- 5335 网络控制块 (NCB) 中指定的适配器号无效。NCB 是相关数据。
- 5336 网络控制块 (NCB) 命令在取消的同时运行完毕。NCB 是相关数据。
- 5337 网络控制块 (NCB) 指定的名称已经保留。NCB 是相关数据。
- 5338 网络控制块 (NCB) 命令无法取消。NCB 是相关数据。
- 5351 同一个会话有多个网络控制块 (NCB)。NCB 请求被拒绝。NCB 是相关数据。
- 5352 网卡出错。唯一可能发出的 NetBIOS 命令是 NCB RESET。网络控制块(NCB)是相关数据。
- 5354 超过应用程序数目上限。网络控制区 (NCB) 请求被拒绝, NCB 是相关数据。
- 5356 请求的资源不可用。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5364 系统出错。网络控制块 (NCB) 请求被拒绝。NCB 即为数据。
- 5365 “ROM 校验和”失败。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5366 RAM 测试失败。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5367 数字式环回失败。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5368 模拟式环回失败。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5369 接口失败。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5370 收到的网络控制块 (NCB) 返回码无法识别。NCB 是相关数据。
- 5380 网卡故障。网络控制块 (NCB) 请求被拒绝。NCB 是相关数据。
- 5381 网络控制块 (NCB) 命令仍然处于挂起状态。NCB 是相关数据。
- 5509 Windows 2000 无法按指定的配置启动, 将换用先前可工作的配置。
- 5600 无法共享用户或脚本路径。
- 5601 计算机的密码在本地安全 数据库中找不到。
- 5602 访问计算机的本地或网络安全数据库时, 发生内部错误。
- 5705 Netlogon 服务用于记录数据库更改数据的 日志高速缓存已损坏。Netlogon 服务正在复位更改日志文件。
- 5728 无法加载任何传输。
- 5739 此域的全局组数目超过可以复制到 LanMan BDC 的限制。请删除部分的全局组或删除域中的 LanManBDC。
- 5742 服务无法检索必要的消息, 所以无法运行远程启动客户端。
- 5743 服务发生严重的错误, 无法从远程启动 3Com 3Start 远程启动客户端。
- 5744 服务发生严重的系统错误, 即将关机。
- 5760 服务在分析 RPL 配置时出错。
- 5761 服务在创建所有配置的 RPL 配置文件时出错。
- 5762 服务在访问注册表时出错。
- 5763 服务在替换可能过期的 RPLDISK.SYS 时出错。

- 5764 服务在添加安全帐户或设置文件权限时出错。这些帐户是独立 RPL 工作站的 RPLUSER 本地组以及用户帐户。
- 5765 服务无法备份它的数据库。
- 5766 服务无法从它的数据库初始化。数据库可能找不到或损坏。服务会试图从备份数据库还原该数据库。
- 5767 服务无法从备份数据库还原它的数据库。服务 将不启动。
- 5768 服务成功从备份数据库还原它的数据库。
- 5769 服务无法从它还原的数据库初始化。服务 将不启动。
- 5771 远程启动数据库采用 NT 3.5 / NT 3.51 格式。NT 正在转换其为 NT 4.0 格式。完成转换后，JETCONV 转换器将写出应用事件日志。
- 5781 由于 DNS 服务器不可用，一个或更多 DNS 记录的动态注册和注销失败。
- 5890 尝试的操作与该节点当前成员身份状态不兼容。
- 5891 仲裁资源不含仲裁日志。
- 5892 成员身份引擎申请关闭该节点上的群集服务。
- 5893 加入操作失败，因为要加入的节点的群集范例 ID 与监护人节点上的群集范例 ID 不匹配。
- 5894 找不到指定的 IP 地址的匹配网络。请指定子网掩码和群集网络。
- 5895 此属性的实际数据类型与它预计的数据类型的不匹配。
- 5896 此群集节点已被成功退出该群集。没有清理此节点。缺少解释此节点没有清理原因的扩展状态信息。
- 5897 为资源属性指定的两个或更多参数值有冲突。
- 5898 此计算机无法成为群集成员。
- 5899 因为此计算机没有安装 Windows 的正确版本，它不能成为一个群集的成员。
- 5900 无法创建指定群集名字的群集，该群集名字已经使用。请指定其他群集名字。
- 6000 无法加密指定的文件。
- 6001 指定的文件无法解密。
- 6002 指定的文件已加密，而且用户没有能力解密。
- 6003 这个系统没有有效的加密恢复策略配置。
- 6004 所需的加密驱动程序并未加载到系统中。
- 6005 文件加密所使用的加密驱动程序与目前加载的加密驱动程序不同。
- 6006 没有为用户定义 EFS 关键字。
- 6007 指定的文件并未加密。
- 6008 指定的文件不是定义的 EFS 导出格式。
- 6009 指定的文件是只读文件。
- 6010 已为加密而停用目录。
- 6011 不信任服务器来进行远程加密操作。
- 6012 为此系统配置的恢复策略包含无效恢复证书。
- 6013 用在此源文件上的加密算法需要比目标文件上更大的密钥缓冲区。
- 6014 磁盘分区不支持文件加密。
- 6015 这台机器已为文件加密而停用。
- 6016 需要一个更新的系统来解密此加密文件。
- 6118 此工作组的服务器列表当前无法使用
- 6200 要正常运行，任务计划程序服务的配置必须在系统帐户中运行。单独的任务可以被配置成在其他帐户中运行。
- 7001 指定的会话名称无效。
- 7002 指定的协议驱动程序无效。
- 7003 在系统路径上找不到指定的协议驱动程序。
- 7004 在系统路径上找不到指定的终端连接。
- 7005 不能为这个会话创建一个事件日志的注册表项。
- 7006 同名的一个服务已经在系统中存在。
- 7007 在会话上一个关闭操作挂起。

- 7008 没有可用的输出缓冲器。
- 7009 找不到 MODEM.INF 文件。
- 7010 在 MODEM.INF 中没有找到调制解调器名称。
- 7011 调制解调器没有接受发送给指令。验证配置的调制解调器与连接的调制解调器是否匹配。
- 7012 调制解调器没有接受发送给指令。验证该调制解调器是否接线正确并且打开了电源开关。
- 7013 由于连接被中断，载波检测没有成功，或者载波已被放弃。
- 7014 在要求的时间内没有检测到拨号音。确定电话线连接正确并可使用。
- 7015 在远程站点回拨时检测到了占线信号。
- 7016 在回叫时远程站点上检测到了声音。
- 7017 传输驱动程序错误
- 7022 找不到指定的会话。
- 7023 指定的会话名称已处于使用中。
- 7024 由于终端连接目前正在忙于处理一个连接、断开连接、复位或删除操作，无法完成该请求的操作。
- 7025 试图连接到其视频模式不受当前客户端支持的会话。
- 7035 应用程序试图启动 DOS 图形模式。不支持 DOS 图形模式。
- 7037 您的交互式登录权限已被禁用。请与您的管理员联系。
- 7038 该请求的操作只能在系统控制台上执行。这通常是驱动程序或系统 DLL 要求直接控制台访问的结果。
- 7040 客户端未能对服务器连接消息作出响应。
- 7041 不支持断开控制台会话。
- 7042 不支持重新将一个断开的会话连接到控制台。
- 7044 远程控制另一个会话的请求被拒绝。
- 7045 拒绝请求的会话访问。
- 7049 指定的终端连接驱动程序无效。
- 7050 不能远程控制该请求的会话。这也许是由于该会话被中断或目前没有一个用户登录。
- 7051 该请求的会话没有配置成允许远程控制。
- 7052 拒绝连接到这个终端服务器。终端服务器客户许可证目前正在被另一个用户使用。请与系统管理员联系，获取一份唯一的许可证号码。
- 7053 拒绝连接到这个终端服务器。还没有为这份终端服务器客户输入您的终端服务器客户许可证号码。请与系统管理员联系。
- 7054 系统已达到其授权的登录限制。请以后再试一次。
- 7055 您正在使用的客户端没有使用该系统的授权。您的登录请求被拒绝。
- 7056 系统许可证已过期。您的登录请求被拒绝。
- 7057 由于当前没有远程控制指定的会话，无法终止远程控制。
- 7058 由于更改了显示模式，控制台的远程控制已终止。不支持在远程控制会话中更改显示模式。
- 8001 没有正确调用文件复制服务 API。
- 8002 无法启动文件复制服务。
- 8003 无法停止文件复制服务。
- 8004 文件复制服务 API 终止了请求。事件日志可能有详细信息。
- 8005 文件复制服务终止了该请求。事件日志可能有详细信息。
- 8006 无法联系文件复制服务。事件日志可能有详细信息。
- 8007 由于该用户没有足够特权，文件复制服务不能满足该请求。事件日志可能有详细信息。
- 8008 由于验证的 RPC 无效，文件复制服务不能满足该请求。事件日志可能有详细信息。
- 8009 由于该用户在域控制器上没有足够特权，文件复制服务不能满足该请求。事件日志可能有详细信息。
- 8010 由于在域控制器上的验证的 RPC 无效，文件复制服务不能满足该请求。事件日志可能有详细信息。
- 8011 该文件复制服务无法与在域控制器上的文件复制服务通讯。事件日志可能有详细信息。
- 8012 在域控制器上的文件复制服务无法与这台计算机上的文件复制服务通讯。事件日志可能

- 有详细信息。
- 8013 由于内部错误，该文件复制服务不能进入该系统卷中。事件日志可能有详细信息。
- 8014 由于内部超时，该文件复制服务不能进入该系统卷中。事件日志可能有详细信息。
- 8015 该文件复制服务无法处理此请求。该系统卷仍在忙于前一个请求。
- 8016 由于内部错误，该文件复制服务无法停止复制该系统卷。事件日志可能有详细信息。
- 8017 该文件复制服务检测到一个无效参数。
- 8200 在安装目录服务时出现一个错误。有关详细信息，请查看事件日志。
- 8201 目录服务在本地评估组成员身份。
- 8202 指定的目录服务属性或值不存在。
- 8203 指定给目录服务的属性语法无效。
- 8204 指定给目录服务的属性类型未定义。
- 8205 指定的目录服务属性或值已经存在。
- 8206 目录服务忙。
- 8207 该目录服务不可用。
- 8208 目录服务无法分配相对标识号。
- 8209 目录服务已经用完了相对标识号池。
- 8210 由于目录服务不是该类操作的主控，未能执行操作。
- 8211 目录服务无法初始化分配相对标识号的子系统。
- 8212 该请求的操作没有满足一个或多个与该对象的类别相关的约束。
- 8213 目录服务只可以在一个页状对象上运行要求的操作。
- 8214 目录服务不能在一个对象的 RDN 属性上执行该请求的操作。
- 8215 目录服务检测出修改对象类别的尝试。
- 8216 不能执行请求的域间移动操作。
- 8217 无法联系全局编录服务器。
- 8218 策略对象是共享的并只可在根目录上修改。
- 8219 策略对象不存在。
- 8220 请求的策略信息只在目录服务中。
- 8221 正在进行域控制器升级。
- 8222 目前没有进行域控制器升级
- 8224 出现了一个操作错误。
- 8225 出现了一个协议错误。
- 8226 已经超过这个请求的时间限制。
- 8227 已经超过这个请求的大小限制。
- 8228 已经超过这个请求的管理限制。
- 8229 比较响应是假的。
- 8230 比较响应是真的。
- 8231 这个服务器不支持请求的身份验证方式。
- 8232 这台服务器需要一个更安全的身份验证方式。
- 8233 不适当的身份验证。
- 8234 未知的身份验证机制。
- 8235 从服务器返回了一个参照。
- 8236 该服务器不支持该请求的关键扩展。
- 8237 这个请求需要一个安全的连接。
- 8238 不恰当的匹配。
- 8239 出现一个约束冲突。
- 8240 在服务器上没有这样一个对象。
- 8241 有一个别名问题。
- 8242 指定了一个无效的 dn 语法。
- 8243 该对象为叶对象。
- 8244 有一个别名废弃问题。

- 8245 该服务器不愿意处理该请求。
- 8246 检查到一个循环。
- 8247 有一个命名冲突。
- 8248 结果集太大。
- 8249 该操作会影响到多个 DSA。
- 8250 该服务器不可操作。
- 8251 出现了一个本地错误。
- 8252 出现了一个编码错误。
- 8253 出现了一个解码错误。
- 8254 无法识别搜索筛选器。
- 8255 一个或多个参数非法。
- 8256 不支持指定的方式。
- 8257 没有返回结果。
- 8258 该服务器不支持该指定的控制。
- 8259 客户端检测到一个参照循环。
- 8260 超过当前的参照限制。
- 8261 搜索要求 SORT 控制。
- 8262 搜索结果超过指定的偏移范围。
- 8301 根目录对象必须是一个命名上下文的头。该根目录对象不能有范例父类。
- 8302 不能执行添加副本操作。命名上下文必须可写才能创建副本。
- 8303 出现一个对架构中未定义的一个属性的参考。
- 8304 超过了一个对象的最大尺寸。
- 8305 试图给目录添加一个名称已在使用中的对象。
- 8306 试图添加一个对象，该对象的类别没有在架构中定义的 RDN。
- 8307 试图添加一个使用 RDN 的对象，但该 RDN 不是在架构中定义的 RDN。
- 8308 在对象中找不到任何请求的属性。
- 8309 用户缓冲区太小。
- 8310 在操作中指定的属性不在对象上。
- 8311 修改操作非法。不允许该修改的某个方面。
- 8312 指定的对象太大。
- 8313 指定的范例类型无效。
- 8314 操作必须在主控 DSA 执行。
- 8315 必须指定对象类别属性。
- 8316 找不到所需的属性。
- 8317 试图修改一个对象，以便包括对其类别非法的属性。
- 8318 指定的属性已在对象上。
- 8320 指定的属性不存在或没有值。
- 8321 为只能有一个值的属性指定了多个值。
- 8322 属性值不在接受范围内。
- 8323 指定的值已存在。
- 8324 由于不在对象上，不能删除该属性。
- 8325 由于不在对象上，不能删除该属性值。
- 8326 指定的根对象不能是子参照。
- 8327 不允许链接。
- 8328 不允许链接的评估。
- 8329 由于对象的父类不是未范例化就是被删除了，所以不能执行操作。
- 8330 不允许使用是别名的父对象。别名是叶对象。
- 8331 对象和父对象必须是同一类型，不是都是原件就是都是副本。
- 8332 由于存在子对象，无法执行操作。这个操作只能在叶对象上执行。
- 8333 没有找到目录对象。

- 8334 找不到别名对象。
- 8335 对象名语法不正确。
- 8336 不允许一个别名参考另一个别名。
- 8337 别名不能解除参考。
- 8338 操作超出范围。
- 8339 由于正在删除对象，操作无法继续。
- 8340 不能删除 DSA 对象。
- 8341 出现了一个目录服务错误。
- 8342 操作只能在内部主控 DSA 对象上执行。
- 8343 对象必须为 DSA 类别。
- 8344 访问权不够，不能执行该操作。
- 8345 由于父类不在可能的上级列表上，不能添加该对象。
- 8346 由于该属性处于“安全帐户管理器”(SAM)，不允许访问该属性。
- 8347 名称有太多部分。
- 8348 名称太长。
- 8349 名称值太长。
- 8350 目录服务遇到了一个错误分列名称。
- 8351 目录服务找不到一个名称的属性类型。
- 8352 该名称不能识别一个对象；该名称识别一个幻象。
- 8353 安全描述符太短。
- 8354 安全描述符无效。
- 8355 未能为删除的对象创建名称。
- 8356 一个新子参考的父项必须存在。
- 8357 该对象必须是一个命名上下文。
- 8358 不允许添加一个属于系统的属性。
- 8359 对象的类别必须是有结构的；您不能范例化一个抽象类别。
- 8360 找不到架构的对象。
- 8361 有这个 GUID (非活动的或活动的)的本地对象已经存在。
- 8362 操作不能在一个后部链接上执行。
- 8363 找不到指定的命名上下文的互交参考。
- 8364 由于目录服务关闭，操作不能执行。
- 8365 目录服务请求无效。
- 8366 无法读取角色所有者属性。
- 8367 请求的 FSMO 操作失败。不能连接当前的 FSMO 盒。
- 8368 不允许跨过一个命名上下文修改 DN。
- 8369 由于属于系统，不能修改该属性。
- 8370 只有复制器可以执行这个功能。
- 8371 指定的类别没有定义。
- 8372 指定的类别不是一个子类别。
- 8373 名称参照无效。
- 8374 交叉参照已经存在。
- 8375 不允许删除一个主控交叉参照。
- 8376 只在 NC 头上支持子目录树通知。
- 8377 通知筛选器太复杂。
- 8378 架构更新失败: 重复的 RDN。
- 8379 架构更新失败: 重复的 OID。
- 8380 架构更新失败: 重复的 MAPI 识别符。
- 8381 架构更新失败: 复制架构 id GUID。
- 8382 架构更新失败: 重复的 LDAP 显示名称。
- 8383 架构更新失败: 范围下部少于范围上部。

- 8384 架构更新失败: 语法不匹配。
- 8385 架构更新失败: 属性在必须包含中使用。
- 8386 架构更新失败: 属性在可能包含中使用。
- 8387 架构更新失败: 可能包含中的属性不存在。
- 8388 架构更新失败: 必须包含中的属性不存在。
- 8389 架构更新失败: 在辅助类别列表中的类别不存在或不是一个辅助类别。
- 8390 架构更新失败: `poss-superior` 中的类别不存在。
- 8391 架构更新失败: 在 `subclassof` 列表中的类别不存在或不能满足等级规则。
- 8392 架构更新失败: `Rdn-Att-Id` 语法不对。
- 8393 架构更新失败: 类别作为辅助类别使用。
- 8394 架构更新失败: 类别作为子类别使用。
- 8395 架构更新失败: 类别作为 `poss superior` 使用。
- 8396 架构更新在重新计算验证缓存时失败。
- 8397 目录树删除没有完成。要继续删除目录树, 必须再次发出请求。
- 8398 不能执行请求的删除操作。
- 8399 不能读取架构记录管理类别识别符。
- 8400 属性架构语法不对。
- 8401 不能缓存属性。
- 8402 不能缓存类别。
- 8403 不能从缓存删除属性。
- 8404 无法从缓存中删除类别。
- 8405 无法读取辨别名称的属性。
- 8406 找不到一个所需的子参考。
- 8407 不能检索范例类型属性。
- 8408 出现了一个内部错误。
- 8409 出现了一个数据错误。
- 8410 找不到属性 `GOVERNSID`。
- 8411 找不到所需属性。
- 8412 指定的命名上下文丢失了一个交叉参考。
- 8413 出现了一个安全检查错误。
- 8414 没有加载架构。
- 8415 架构分配失败。请检查机器内存是否不足。
- 8416 未能为属性架构获得所需语法。
- 8417 全局编录验证失败。全局编录不可用或不支持操作。目录的某些部分目前不可用。
- 8418 由于有关服务器之间的架构不匹配, 复制操作失败。
- 8419 找不到 `DSA` 对象。
- 8420 找不到命名上下文。
- 8421 在缓存中找不到命名上下文。
- 8422 无法检索子对象。
- 8423 由于安全原因不允许修改。
- 8424 操作不能替换隐藏的记录。
- 8425 等级文件无效。
- 8426 未能建立等级表。
- 8427 注册表中缺少目录配置参数。
- 8428 尝试计算地址簿索引失败。
- 8429 等级表的分配失败。
- 8430 目录服务遇到一个内部故障。
- 8431 目录服务遇到一个未知故障。
- 8432 根对象需要一个 `'top'` 类别。
- 8433 这个目录服务器已关闭, 并且不能接受新的浮动单主机操作角色的所有权。

- 8434 目录服务没有必需的配置信息，并且不能决定新的浮动单主机操作角色的所有权。
- 8435 该目录服务无法将一个或多个浮动单主机操作角色传送给其他服务器。
- 8436 复制操作失败。
- 8437 为这个复制操作指定了一个无效的参数。
- 8438 目录服务太忙，目前无法完成这个复制操作。
- 8439 为这个复制操作指定的辨别名称无效。
- 8440 为这个复制操作所指定的命名上下文无效。
- 8441 为这个复制操作指定的辨别名称已经存在。
- 8442 复制系统遇到一个内部错误。
- 8443 复制操作遇到数据库不一致问题。
- 8444 不能连接到为这个复制操作指定的服务器上。
- 8445 复制操作遇到一个有无效范例类型的对象。
- 8446 复制操作无法分配内存。
- 8447 复制操作遇到一个邮件系统错误。
- 8448 目标服务器的复制参考信息已经存在。
- 8449 目标服务器的复制参考信息不存在。
- 8450 由于是由另一台服务器上复制的，因此不能删除命名上下文。
- 8451 复制操作遇到一个数据库错误。
- 8452 命名上下文要被删除或没有从指定的服务器上复制。
- 8453 复制访问被拒绝。
- 8454 这个版本的目录服务不支持请求的操作。
- 8455 取消了复制远程过程呼叫。
- 8456 源服务器目前拒绝复制请求。
- 8457 目标服务器目前拒绝复制请求。
- 8458 由于对象名冲突，复制操作失败。
- 8459 复制源已被重新安装。
- 8460 由于一个所需父对象丢失，复制操作失败。
- 8461 复制操作被抢先。
- 8462 由于缺乏更新，放弃复制同步尝试。
- 8463 由于系统正在关闭，复制操作被中断。
- 8464 由于目标部分属性集不是一个源部分属性集的子集，复制同步尝试失败。
- 8465 由于主复制尝试从部分副本同步，复制同步尝试失败。
- 8466 已经与为这个复制操作的指定的服务器联系，但是该服务器无法与完成这个操作所需的另外一个服务器联系。
- 8467 源目录林的 Active Directory 架构与这台计算机上的 Active Directory 的版本不兼容。您必须将源目录林里域控制器上的操作系统升级，这台计算机才能作为域控制器加到源目录林上。
- 8468 架构更新失败：有同一链接标识符的属性已经存在。
- 8469 名称转换：常见处理错误。
- 8470 名称转换：不能找到该名称；或权限不够，无法看到名称。
- 8471 名称转换：输入名称映射到多个输出名称。
- 8472 名称转换：找到了输出名称，但是找不到相应的输出格式。
- 8473 名称转换：不能完全解析，只找到了域。
- 8474 名称转换：不接到线上，无法在客户端执行纯粹的语法映射。
- 8475 不允许一个构造 att 修改。
- 8476 指定的 OM 对象类别对具有指定语法的属性是不正确的。
- 8477 复制请求已暂停；等待回答。
- 8478 要求的操作需要一个目录服务，但没有可用的。
- 8479 类别或属性的 LDAP 显示名称含有非 ASCII 字符。
- 8480 请求的搜索操作只支持基本查找。

- 8481 搜索未能从数据库检索属性。
- 8482 架构更新操作试图添加一个反向链接，但该反向链接没有相应的正向链接。
- 8483 跨域移动的来源和目标在对象日期上不一致。或者是来源，或者是目标没有对象最新版本。
- 8484 跨域移动的来源和目标在对象当前的名称上不一致。或者是来源，或者是目标没有对象的最新版本。
- 8485 跨域移动的来源和目标是一样的。调用程序应该使用本地移动操作，而不是跨域移动操作。
- 8486 跨域移动的来源和目标与目录林中的命名上下文不一致。来源或目标没有分区容器的最新版本。
- 8487 跨域移动的目标不是目标命名上下文的权威。
- 8488 跨域移动的来源和目标提供的来源对象身份不一样。来源或目标没有来源对象的最近版本。
- 8489 跨域移动的对象应该已经被目标服务器删除。来源服务器没有来源对象的最近版本。
- 8490 要求对 PDC FSMO 的专门访问权的另一个操作正在进行中。
- 8491 跨域移动没有成功，导致被移动对象有两个版本 - 一个在来源域，一个在目标域。需要删除目标对象，将系统还原到一致状态。
- 8492 因为不允许跨域移动这个类别，或者对象有一些特点，如：信任帐户或防止移动的受限制的 RID；所以不能将该对象跨域移动。
- 8493 一旦移动，不能将带有成员身份的对象跨域移动，这会侵犯帐户组的成员身份条件。从帐户组成员身份删除对象，再试一次。
- 8494 命名上下文标题必须是另一个命名上下文标题的直接子标题，而不是一个内节点的子标题。
- 8495 因为目录没有提议的命名上下文上面的命名上下文的副本，所以无法验证所提议的命名上下文的名称。请保证充当域命名主机的服务器已配置成全局编录服务器，并且服务器及其复制伙伴是最新的。(仅适用于 Windows 2000 域命名主机)
- 8496 目标域必须在本机模式中。
- 8497 因为服务器在指定域中没有基础结构容器，所以无法执行操作。
- 8498 不允许跨域移动不是空的帐户组。
- 8499 不允许跨域移动不是空的资源组。
- 8500 属性的搜索标志无效。ANR 位只在 Unicode 或 Teletex 字符串的属性上有效。
- 8501 不允许在将 NC 头作为子体的对象开始删除目录树。
- 8502 因为目录树在使用中，目录服务未能为删除目录树而将其锁定。
- 8503 删除目录树时，目录服务未能识别要删除的对象列表。
- 8505 只有管理员才能修改管理组的成员列表。
- 8506 不能改变域控制器帐户的主要组 ID。
- 8507 试图修改基础架构。
- 8508 不允许进行下列操作：为现有类别添加新的强制属性；从现有类别删除强制属性；为没有向回链接属性的特殊类别 "Top" 添加可选属性，向回链接属性指的是直接或通过继承。例如：添加或删除附属类别。
- 8509 由于 DC 不是架构 FSMO 角色所有者，该域控制器上不允许架构更新。
- 8510 无法在架构容器下创建这个类别的对象。在架构容器下，您只能创建属性架构和类别架构对象。
- 8511 副本/子项安装未能获取源 DC 上的架构容器的 objectVersion 属性。架构容器上的属性不存在，或者提供的凭证没有读取属性的权限。
- 8512 副本/子项安装未能读取 system3 目录中的文件 schema.ini 的 SCHEMA 段中的 objectVersion 属性。
- 8513 指定的组类型无效。
- 8514 如果域是安全启用的，在混合型域中不能嵌套全局组。
- 8515 如果域是安全启用的，在混合型域中不能嵌套本地组。
- 8516 全局组不能将本地组作为成员。
- 8517 全局组不能将通用组作为成员。
- 8518 通用组不能将本地组作为成员。
- 8519 全局组不能有跨域成员。

- 8520 本地组不能将另一个跨域本地组作为成员。
- 8521 包含主要成员的组不能改变为安全停用的组。
- 8522 架构缓冲加载未能转换类架构对象上的字符串默认值 SD。
- 8523 只有配置成全局编录服务器的 DSAs 才能充当域命名主机 FSMO 的角色。
- 8524 由于 DNS 查找故障, DSA 操作无法进行。
- 8525 处理一个对象的 DNS 主机名改动时, 服务主要名称数值无法保持同步。
- 8526 未能读取安全描述符属性。
- 8527 没有找到请求的对象, 但找到了具有那个密钥的对象。
- 8528 正在添加的链接属性的语法不正确。正向链接只能有语法 2.5.5.1、2.5.5.7 和 2.5.5.14, 而反向链接只能有语法 2.5.5.1
- 8529 安全帐户管理员需要获得启动密码。
- 8530 安全帐户管理员需要从软盘获得启动密钥。
- 8531 目录服务无法启动。
- 8532 目录服务未能启动。
- 8533 客户端和服务端之间的连接要求数据包保密性。
- 8534 来源域跟目标域不在同一个目录林中。
- 8535 目标域必须在目录林中。
- 8536 该操作要求启用目标域审核。
- 8537 该操作无法找到来源域的 DC。
- 8538 来源对象必须是一个组或用户。
- 8539 来源对象的 SID 已经在目标目录林中。
- 8540 来源对象和目标对象必须属于同一类型。
- 8542 在复制请求中不能包括架构信息。
- 8543 由于架构不兼容性, 无法完成复制操作。
- 8544 由于前一个架构的不兼容性, 无法完成复制操作。
- 8545 因为源或目标没有收到有关最近跨域启动操作的信息, 所以无法应用复制更新。
- 8546 因为还有主控这个域的域控制器, 所以无法删除请求的域。
- 8547 只能在全局编录服务器上执行请求的操作。
- 8548 本地组只能是同一个域中其他本地组的成员。
- 8549 外部安全主要成员不能是通用组的成员。
- 8550 出于安全, 无法将属性复制到 GC。
- 8551 由于目前正在处理的修改太多, 无法采取 PDC 的检查点。
- 8552 操作需要启用那个源域审核。
- 8553 安全主要对象仅能在域命名环境菜单中创建。
- 8554 服务主要名称(SPN)无法建造, 因为提供的主机名格式不适合。
- 8555 筛选器已传递建造的属性。
- 8556 unicodePwd 属性值必须括在双引号中。
- 8557 您的计算机无法加入域。已超出此域所允许创建的计算机帐户的最大值。请跟系统管理员联系, 以便重设或增加此限制。
- 8558 由于安全原因, 操作必须在目标 DC 上运行。
- 8559 由于安全原因, 源 DC 必须是 NT4SP4 或更新版本。
- 8560 在树目录删除的操作中不能删除“关键目录服务系统”对象。数目录删除操作可能只进行了一部分。
- 8563 此 Windows 版本太老, 不支持当前目录林行为。您必须将此服务器上的操作系统升级, 它才能成为该林的域控制器。
- 8564 此 Windows 版本太老, 不支持当前域行为。您必须将此服务器上的操作系统升级, 它才能成为该域的域控制器。
- 8565 此 Windows 版本不再支持这个目录林里用的行为版本。您必须将此林行为版本升到高级, 此服务器才能成为林里的域控制器。
- 8566 此 Windows 版本不再支持这个域里用的行为版本。您必须将此域行为版本升到高级, 此

- 服务器才能成为域里的域控制器。
- 8567 此 Windows 版本与域或目录林行为版本不兼容。
- 8568 行为版本不能增加到申请的值, 因为域控制器仍以低于申请的值存在。
- 8569 如果域仍处于混合域模式, 行为版本值就不能增加。您必须在增加行为版本前先将域改成本机模式。
- 8570 申请的排序顺序不受支持。
- 8571 发现一个非唯一名称。
- 8572 机器帐户是由 NT4 之前的版本创建的。需要重新创建。
- 8573 数据库在版本存储之外。
- 8574 无法继续操作, 因为用了多个冲突控制。
- 8575 无法找到这个分区的有效安全描述符参考域。
- 8576 架构升级失败: 链接标识符被保留。
- 8577 架构升级失败: 链接标识符不可用。
- 8578 帐户组无法将通用组作为成员。
- 8579 不允许在命名上下文标题或只读对象上重新命名或移动操作。
- 8580 不允许在架构命名上下文里的对象上移动操作。
- 8581 已在对象上设置系统标记并不允许移动或重新命名该对象。
- 8582 不允许此对象改变它的祖容器。在此对象上禁止移动, 但限制于移到兄弟容器。
- 8583 无法完全分解, 生成对另一目录林的检索。
- 8584 申请的操作在标准服务器上不受支持。
- 8585 不能访问远程服务器上的 Active Directory 的分区。请确保至少有一个服务器在为该分区运行。
- 8586 目录不能验证推荐的命名上下文(或分区)名称, 因为它没有副本并且不能联系推荐的命名上下文之上的命名上下文副本。请确保该父命名上下文在 DNS 里的配置正确, 并且域命名主机至少可以到达这个命名一个副本。
- 8587 已超出该请求线程的限制。
- 8588 全局编录服务器不在最近的站点。
- 8589 DS 不能派生用以相互给目标服务器验证身份的服务主要名称(SPN), 这是因为在本地 DS 数据库中与之相对应的服务器对象没有 serverRefernce 属性。
- 8590 目录服务无法进入单一用户模式。
- 8591 由于语法错误, 目录服务无法分解此脚本。
- 8592 由于发生错误, 目录服务无法分解此脚本。
- 8593 因为涉及的服务器处于不同的复制时期(通常与一个正在进行域的重新命名有关), 目录服务无法执行请求的操作。
- 8594 由于服务器扩展名信息改变, 目录服务绑定 必须重新协商。
- 8595 不允许在禁用的交叉引用上操作。
- 8596 架构更新失败: 没有 msDS-IntId 的值。
- 8597 架构更新失败: msDS-INtId 已存在。重新尝试此操作。
- 8598 架构删除失败: 属性在 rDNAAttID 中使用。
- 8599 目录服务无法对此请求进行身份验证。
- 8600 脚本无效, 目录服务无法分解此脚本。
- 8601 在域命名主 FSMO 上的远程创建交叉引用的操作失败。操作的错误在扩展数据中。
- 9001 DNS 服务器无法解释格式。
- 9002 DNS 服务器失败。
- 9003 DNS 名称不存在。
- 9004 名称服务器不支持 DNS 请求。
- 9005 拒绝 DNS 操作。
- 9006 不应该存在的 DNS 名称仍然存在。
- 9007 不应该存在的 DNS RR 集仍然存在。
- 9008 应该存在的 DNS RR 集不存在。

- 9009 DNS 服务器对区域没有权威。
- 9010 在更新或 prereq 中的 DNS 名称不在区域中。
- 9016 DNS 签名验证失败。
- 9017 DNS 不正确密钥。
- 9018 DNS 签名验证过期。
- 9501 为 DNS 查询找不到记录。
- 9502 不正确的 DNS 包。
- 9503 没有 DNS 包。
- 9504 DNS 错误, 请检查 rcode。
- 9505 没有保险的 DNS 包。
- 9551 无效的 DNS 类型。
- 9552 无效的 IP 地址。
- 9553 无效的属性。
- 9554 稍后再试一次 DNS 操作。
- 9555 提供的名称和类型的记录不是唯一的。
- 9556 DNS 名称不符合 RFC 说明。
- 9557 DNS 名称是一个完全合格的 DNS 名称。
- 9558 DNS 名称以“.”分隔(多标签)。
- 9559 DNS 名称是单一部分名称。
- 9560 DNS 名称含有无效字符。
- 9561 DNS 名称完全是数字的。
- 9562 DNS 根目录服务器不允许请求的操作。
- 9601 DNS 区域不存在。
- 9602 DNS 区域信息无效。
- 9603 DNS 区域无效操作。
- 9604 无效 DNS 区域配置。
- 9605 DNS 区域没有颁发机构起始(SOA)的记录。
- 9606 DNS 区域没有名称服务器(NS)的记录。
- 9607 DNS 区域已锁定。
- 9608 DNS 区域创建失败。
- 9609 DNS 区域已经存在。
- 9610 DNS 自动区域已经存在。
- 9611 无效的 DNS 区域类型。
- 9612 次要 DNS 区域需要主 IP 地址。
- 9613 DNS 区域不是次要的。
- 9614 需要一个次要 IP 地址。
- 9615 WINS 初始化失败。
- 9616 需要 WINS 服务器。
- 9617 NBTSTAT 初始化呼叫失败。
- 9618 颁发机构起始(SOA)删除无效
- 9619 已有一个同名的条件转发区。
- 9620 此区域必须用一个或多个主 DNS 服务器的 IP 地址配置。
- 9621 因为此区域已经关闭, 此操作无法执行。
- 9651 主要 DNS 区域需要数据文件。
- 9652 DNS 区域的无效数据文件名称。
- 9653 未能打开 DNS 区域的数据文件。
- 9654 未能写入 DNS 区域的数据文件。
- 9655 读取 DNS 区域的数据文件时出现故障。
- 9701 DNS 记录不存在。
- 9702 DNS 记录格式错误。

- 9703 DNS 中节点创建失败。
- 9704 未知 DNS 记录类型。
- 9705 DNS 记录超时。
- 9706 名称不在 DNS 区域。
- 9707 检测到 CNAME 循环。
- 9708 节点为一个 CNAME DNS 记录。
- 9709 指定名称的 CNAME 记录已经存在。
- 9710 记录不在 DNS 区域根目录。
- 9711 DNS 记录已经存在。
- 9712 次要 DNS 区域数据错误。
- 9713 不能创建 DNS 缓存数据。
- 9714 DNS 名称不存在。
- 9715 不能创建指针(PTR)记录。
- 9716 DNS 域没有被删除。
- 9717 该目录服务不可用。
- 9718 DNS 区域已经在目录服务中存在。
- 9719 DNS 服务器没有为目录服务集合 DNS 区域创建或读取启动文件。
- 9751 DNS AXFR (区域复制)已完成。
- 9752 DNS 区域复制失败。
- 9753 添加了本地 WINS 服务器。
- 9801 安全更新呼叫需要继续更新请求。
- 9851 TCP/IP 没有安装网络协议。
- 9852 没有为本地系统配置 DNS 服务器。
- 9901 指定的目录分区不存在。
- 9902 指定的目录分区已存在。
- 9903 DS 没被列入指定的目录分区。
- 9904 DS 已被列入指定的目录分区。
- 10004 一个封锁操作被对 WSACancelBlockingCall 的调用中断。
- 10009 提供的文件句柄无效。
- 10013 以一种访问权限不允许的方式做了一个访问套接字的尝试。
- 10014 系统检测到在一个调用中尝试使用指针参数时的无效指针地址。
- 10022 提供了一个无效的参数。
- 10024 打开的套接字太多。
- 10035 无法立即完成一个非阻挡性套接字操作。
- 10036 目前正在执行一个阻挡性操作。
- 10037 在一个非阻挡套接字上尝试了一个已经在进行的操作。
- 10038 在一个非套接字上尝试了一个操作。
- 10039 请求的地址在一个套接字中从操作中忽略。
- 10040 一个在数据报套接字上发送的消息大于内部消息缓冲器或其他一些网络限制, 或该用户用于接收数据报的缓冲器比数据报小。
- 10041 在套接字函数调用中指定的一个协议不支持请求的套接字类型的语法。
- 10042 在 getsockopt 或 setsockopt 调用中指定的一个未知的、无效的或不受支持的选项或层次。
- 10043 请求的协议还没有在系统中配置, 或者没有它存在的迹象。
- 10044 在这个地址家族中不存在对指定的插槽类型的支持。
- 10045 参考的对象类型不支持尝试的操作。
- 10046 协议家族尚未配置到系统中或没有它的存在迹象。
- 10047 使用了与请求的协议不兼容的地址。
- 10048 通常每个套接字地址(协议/网络地址/端口)只允许使用一次。
- 10049 在其上下文中, 该请求的地址无效。
- 10050 套接字操作遇到了一个已死的网络。

- 10051 向一个无法连接的网络尝试了一个套接字操作。
- 10052 当该操作在进行中，由于保持活动的操作检测到一个故障，该连接中断。
- 10053 您的主机中的软件放弃了一个已建立的连接。
- 10054 远程主机强迫关闭了一个现有的连接。
- 10055 由于系统缓冲区空间不足或队列已满，不能执行套接字上的操作。
- 10056 在一个已经连接的套接字上做了一个连接请求。
- 10057 由于套接字没有连接并且(当使用一个 `sendto` 调用发送数据报套接字时)没有提供地址，发送或接收数据的请求没有被接受。
- 10058 由于以前的关闭调用，套接字在那个方向已经关闭，发送或接收数据的请求没有被接受。
- 10059 对某个内核对象的引用过多。
- 10060 由于连接方在一段时间后没有正确答复或连接的主机没有反应，连接尝试失败。
- 10061 由于目标机器积极拒绝，无法连接。
- 10062 无法转换名称。
- 10063 名称组件或名称太长。
- 10064 由于目标主机坏了，套接字操作失败。
- 10065 套接字操作尝试一个无法连接的主机。
- 10066 不能删除目录，除非它是空的。
- 10067 一个 Windows 套接字操作可能在可以同时使用的应用程序数目上有限制。
- 10068 限额不足。
- 10069 磁盘限额不足。
- 10070 文件句柄引用不再可用。
- 10071 项目在本地不可用。
- 10091 因为它使用提供网络服务的系统目前无效，`WSAStartup` 目前不能正常工作。
- 10092 不支持请求的 Windows 套接字版本。
- 10093 应用程序没有调用 `WSAStartup`，或者 `WSAStartup` 失败。
- 10101 由 `WSARecv` 或 `WSARecvFrom` 返回表示远程方面已经开始了关闭步骤。
- 10102 `WSALookupServiceNext` 不能返回更多的结果。
- 10103 在该调用还处于处理中时，就调用了 `WSALookupServiceEnd`。该调用被取消。
- 10104 过程调用无效。
- 10105 请求的服务提供程序无效。
- 10106 无法加载或初始化请求的服务提供程序。
- 10107 从来不应失败的系统调用失败了。
- 10108 此服务不存在。在指定的名称空间中找不到这个服务。
- 10109 找不到指定的类别。
- 10110 `WSALookupServiceNext` 不能返回更多的结果。
- 10111 在该调用还处于处理中时，就调用了 `WSALookupServiceEnd`。该调用被取消。
- 10112 由于被拒绝，数据查询失败。
- 11001 不知道这样的主机。
- 11002 这是在主机名解析时通常出现的暂时错误，它意味着本地服务器没有从权威服务器上收到响应。
- 11003 在数据查找中出现一个不可恢复的错误。
- 11004 请求的名称有效并且在数据库中找到，但是它没有相关的正确的数据来被解析。
- 11005 至少到达了一个保留。
- 11006 至少到达了一个路径。
- 11007 没有发送方。
- 11008 没有接受方。
- 11009 保留已经确认。
- 11010 错误是由于资源不足造成。
- 11011 由于管理原因被拒绝 - 无效凭证。
- 11012 未知或有冲突类型。

- 11013 某一部分的 filterspec 或 providerspecific 缓冲区有问题。
- 11014 flowspec 的某部分有问题。
- 11015 一般性 QOS 错误。
- 11016 在流程规格中发现一个无效的或不可识别的服务类型。
- 11017 在 QOS 结构中发现一个无效的或不一致的流程规格。
- 11018 无效的 QOS 提供程序特定缓冲区。
- 11019 使用了无效的 QOS 筛选器样式。
- 11020 使用了无效的 QOS 筛选器类型。
- 11021 FLOWDESCRIPTOR 中指定的 QOS FILTERSPEC 数量不正确。
- 11022 在 QOS 提供程序特定缓冲区中指定了一个 ObjectLength 字符域无效的对象。
- 11023 QOS 结构中指定的流程描述符数量不正确。
- 11024 在 QOS 提供程序特定缓冲区中发现一个不可识别的对象。
- 11025 在 QOS 提供程序特定缓冲区中发现一个无效的策略对象。
- 11026 在流程描述符列表中发现一个无效的 QOS 流程描述符。
- 11027 在 QOS 提供程序特定缓冲区中发现一个无效的或不一致的流程规格。
- 11028 在 QOS 提供程序特定缓冲区中发现一个无效的 FILTERSPEC。
- 11029 在 QOS 提供程序特定缓冲区中发现一个无效的波形丢弃模式对象。
- 11030 在 QOS 提供程序特定缓冲区中发现一个无效的成形速率对象。
- 11031 在 QOS 提供程序特定缓冲区中发现一个保留的策略因素。
- 13000 指定的快速模式策略已存在。
- 13001 找不到指定的快速模式策略。
- 13002 指定的快速模式策略正在使用中。
- 13003 指定的主要模式策略已存在。
- 13004 找不到指定的主要模式策略。
- 13005 指定的主要模式策略正在使用中。
- 13006 指定的主要模式筛选器已存在。
- 13007 找不到指定的主要模式筛选器。
- 13008 指定的传输模式筛选器已存在。
- 13009 指定的传输模式筛选器不存在。
- 13010 指定的主要模式身份验证列表存在。
- 13011 找不到指定的主要模式身份验证列表。
- 13012 指定的快速模式策略正在使用中。
- 13013 找不到指定的主要模式策略。
- 13014 找不到指定的快速模式策略
- 13015 清单文件包含多个语法错误。
- 13016 应用程序试图激活禁用的激活上下文。
- 13017 在活动的激活上下文中没有找到任何查找密钥。
- 13018 主要模式筛选器正搁置删除。
- 13019 传输筛选器正搁置删除。
- 13020 隧道筛选器正搁置删除。
- 13021 主要模式策略正搁置删除。
- 13022 主要模式身份验证绑定正搁置删除。
- 13023 快速模式策略正搁置删除。
- 13800 ERROR_IPSEC_IKE_NEG_STATUS_BEGIN
- 13801 IKE 验证凭证不可接受
- 13802 IKE 安全属性不可接受
- 13803 IKE 协商正在进行中
- 13804 一般处理错误
- 13805 协商超时
- 13806 IKE 未能找到有效的机器证书
- 13807 完成建立之前, IKE SA 已被对等客户删除

13808	完成建立之前, IKE SA 已被删除
13809	协商请求在队列中时间太长
13810	协商请求在队列中时间太长
13811	协商请求在队列中时间太长
13812	协商请求在队列中时间太长
13813	对等客户没有响应
13814	协商所花时间太长
13815	协商所花时间太长
13816	出现了未知错误
13817	证书吊销检查失败
13818	无效的证书密钥用法
13819	无效的证书类型
13820	没有跟机器证书相关链的私钥
13822	Diffie-Helman 计算中的失误
13824	无效的标头
13825	没有配置策略
13826	未能确认签名
13827	未能用 kerberos 验证
13828	对等客户的证书没有公钥
13829	处理 error 负载时的错误
13830	处理 SA 负载时的错误
13831	处理 Proposal 负载时的错误
13832	处理 Transform 负载时的错误
13833	处理 KE 负载时的错误
13834	处理 ID 负载时的错误
13835	处理 Cert 负载时的错误
13836	处理 Certificate Request 负载时的错误
13837	处理 Hash 负载时的错误
13838	处理 Signature 负载时的错误
13839	处理 Nonce 负载时的错误
13840	处理 Notify 负载时的错误
13841	处理 Delete 负载时的错误
13842	处理 VendorId 负载时的错误
13843	收到了无效负载
13844	加载了软 SA
13845	软 SA 被拆散
13846	收到了无效 cookie。
13847	对等客户未能发送有效的机器证书
13848	对等客户证书的证书吊销检查没有成功
13849	新策略使旧策略形成的 SA 无效
13850	没有可用的主模式 IKE 策略。
13851	启用 TCB 特权失败。
13852	加载 SECURITY.DLL 失败。
13853	未能从 SSPI 取得安全函数表调度地址。
13854	未能查询 Kerberos 程序包以取得最大令牌尺寸。
13855	未能为 ISAKMP/ERROR_IPSEC_IKE 服务获取 Kerberos 服务器凭证。Kerberos 验证无法运行。很可能是因为缺少域成员身份。如果您的计算机是工作组的成员, 这是正常的。
13856	未能确定 ISAKMP/ERROR_IPSEC_IKE 服务(QueryCredentialsAttributes)的 SSPI 主要名称。
13857	未能从 Ipsec 驱动程序为入站的 SA 获取新的 SPI。造成这种现象的最常见原因是, 驱动程序没有正确的筛选器。请检查策略来验证筛选器。

- 13858 给出的筛选无效
- 13859 内存分配失败。
- 13860 未能将安全关联添加到 IPSec 驱动程序。造成这种现象的最常见原因是，IKE 协商花的时间太长。如果问题继续存在，请减少造成错误的机器上的负荷。
- 13861 无效策略
- 13862 无效 DOI
- 13863 无效情况
- 13864 Diffie-Hellman 故障
- 13865 无效的 Diffie-Hellman 组
- 13866 加密有效负载的错误
- 13867 解密有效负载的错误
- 13868 策略匹配错误
- 13869 不受支持的 ID
- 13870 散列验证没有成功
- 13871 无效散列算法
- 13872 无效散列大小
- 13873 无效的加密算法
- 13874 无效的验证算法
- 13875 无效的证书签名
- 13876 加载没有成功
- 13877 通过 RPC 调用删除的
- 13878 为执行重新初始化而创建的临时状态。这不是一个真正的故障。
- 13879 在请求器寿命通知里收到的寿命值低于 Windows2000 配置的最小值。请整理对等机器上的策略。
- 13881 对于配置的安全要求来说，证书里的密钥长度太小。
- 13882 超过了对等机器的已建立的 MM SA 最大数值。
- 13883 IKE 收到一个停用协商的策略。
- 13884 ERROR_IPSEC_IKE_NEG_STATUS_END
- 14000 申请的节不存在于激活上下文。
- 14001 由于应用程序配置不正确，应用程序未能启动。重新安装应用程序可能会纠正这个问题。
- 14002 应用程序绑定数据格式无效。
- 14003 参照的汇编没有安装在系统上。
- 14004 清单文件没有以要求的标记和格式信息开始。
- 14005 清单文件包含多个语法错误。
- 14006 应用程序试图激活禁用的激活上下文。
- 14007 在活动的激活上下文中没有找到任何查找密钥。
- 14008 应用程序要求的组件版本同另一个活动的组件有冲突。
- 14009 要求激活上下文部分的类型与使用的查询 API 不匹配。
- 14010 由于系统资源缺乏，要求为当前执行的线程禁用独立的激活上下文。
- 14011 已设置进程默认激活上下文，因此设置失败。
- 14012 不能识别指定的编码组标识符。
- 14013 不能识别申请的编码。
- 14014 清单中参考一个无效的 URI。
- 14015 应用程序的清单中参考一个没安装的从属汇编
- 14016 应用程序所用的汇编的清单参考一个没安装的从属汇编
- 14017 清单中包含无效汇编身份的属性。
- 14018 清单中没有所要求的汇编元素的默认名称空间规格。
- 14019 清单中有汇编元素的默认名称空间，但是它的值不是"urn:schemas-microsoft-com:asm.v1".
- 14020 要查明的专用清单已越过与再次分析点相关联的路径
- 14021 至少有两个被应用程序清单直接或间接引用的组件具有同名的文件。
- 14022 至少有两个被应用程序清单直接或间接引用的组件具有同名的窗口类别。

- 14023 至少有两个被应用程序清单直接或间接引用的组件具有相同的 COM 服务器 CLSID。
- 14024 至少有两个被应用程序清单直接或间接引用的组件具有相同 COM 接口 IID 的代理程序。
- 14025 至少有两个被应用程序清单直接或间接引用的组件具有相同的 COM 类型库 TLBID。
- 14026 至少有两个被应用程序清单直接或间接引用的组件具有相同的 COM ProgID。
- 14027 至少有两个被应用程序清单直接或间接引用的组件是不被允许的同一个组件的不同版本。
- 14028 组件的文件跟组件清单中的验证信息不匹配。
- 14029 策略清单含有至少一个语法错误。
- 14030 清单分析错误：要求字符串程序文字，但没有找到开始双引号字符。
- 14031 清单分析错误：注释中所用的语法不正确。
- 14032 清单分析错误：名称的起始字符无效。
- 14033 清单分析错误：名称含有无效字符。
- 14034 清单分析错误：字符串程序文字含有无效字符。
- 14035 清单分析错误：xml 声明的语法无效。
- 14036 清单分析错误：在文字内容中找到一个无效字符。
- 14037 清单分析错误：找不到必需的空白区域。
- 14038 清单分析错误：要求带字符 '>'。
- 14039 清单分析错误：要求带一个分号字符。
- 14040 清单分析错误：括号不平衡。
- 14041 清单分析错误：内部错误。
- 14042 清单分析错误：此处不允许有空白区域。
- 14043 清单分析错误：在当前加密的无效状态下达到了文件结尾。
- 14044 清单分析错误：找不见括号。
- 14045 清单分析错误：找不见收尾的单引号或双引号(' 或 \')字符。
- 14046 清单分析错误：一个名称中不允许多个冒号。
- 14047 清单分析错误：对于十进制数字无效的字符。
- 14048 清单分析错误：对于十六进制数字无效的字符。
- 14049 清单分析错误：对于这个平台无效的 unicode 字符值。
- 14050 清单分析错误：要求带空白区域或 '?'。
- 14051 清单分析错误：此处不允许结束标记。
- 14053 清单分析错误：重复属性。
- 14054 清单分析错误：在 XML 文档中只允许一个最高层元素。
- 14055 清单分析错误：文档的最高层无效。
- 14056 清单分析错误：无效的 xml 声明。
- 14057 清单分析错误：XML 文档必须有一个最高层元素。
- 14058 清单分析错误：文件结尾出乎意料。
- 14059 清单分析错误：无法在内部子集的标记声明内使用参数实体。
- 14060 清单分析错误：没有关闭元素。
- 14061 清单分析错误：结尾元素没有字符 '>'。
- 14062 清单分析错误：字符串程序文字没有关闭。
- 14063 清单分析错误：注释没有关闭。
- 14064 清单分析错误：声明没有关闭。
- 14065 清单分析错误：CDATA 节没有关闭。
- 14066 清单分析错误：名称空间前缀不能以保留字符串 "xml" 打头。
- 14067 清单分析错误：系统不支持指定的加密。
- 14068 清单分析错误：不支持从当前加密切换到指定的加密。
- 14069 清单分析错误：名称 'xml' 是保留的，必须是小写。
- 14070 清单分析错误：独立属性必须带有 'yes' 或 'no' 值。
- 14071 清单分析错误：独立属性不能用在外部实体中。
- 14072 清单分析错误：无效的版本编号。
- 14073 清单分析错误：属性和属性值之间没有等于号。

- 14074 汇编保护错误：无法恢复指定的汇编。
- 14075 汇编保护错误：汇编的公钥太短，不允许。
- 14076 汇编保护错误：汇编的编录无效，或者跟汇编的清单不匹配。
- 14077 无法将 HRESULT 转换成相应的 Win32 错误代码。
- 14078 汇编保护错误：找不到汇编的编录。
- 14079 提供的汇编身份丢失在上下文中必须存在的一个或多个属性。
- 14080 提供的汇编身份有一个或多个属性名称包含在 XML 名称中不允许使用的字符。