



# 软件产品说明

## (*Software Product Description*)

---

产品名称	TRS 全文数据库系统
产品简称	TRS Database Server 或 TRS Server
发布版本	6.8
发布日期	2010 年 5 月 14 日

### 一、产品描述

一般地，信息可分为结构化和非结构化两大类。对于结构化数据，关系数据库管理系统可以很好的管理，而非结构化数据包括网页、电子邮件、Word 文档、PDF 文件以及没有以行列格式保存的任何内容，关系数据库很难进行有效的管理和查询。TRS 全文检索系统是一款兼顾结构化和非结构化数据、兼顾查全和查准需求、适合于企业数据搜索的全文数据库系统，它针对信息内容管理和资源建设的新需求，发展了包括 Native XML、集群、UNICODE 支持、自然语言及智能扩展检索等在内的众多新功能，结合 TRS 领先的结构化和非结构化数据联合查询技术，从而满足了用户对企业搜索的广泛需求。TRS 全文数据库系统 6.8 版继承了得到业界已经广泛使用的 TRS 全文检索系统的全部功能和性能，重新设计的检索和相关性算法更加高效快速。

TRS 全文数据库系统 6.8 主要特点：

- 结构化与非结构化数据的统一管理：实现结构化和非结构化数据的统一检索，回答基于元数据（meta-data）和全文（full-text）的联合查询。
- 实时动态索引：数据增删改时快速同步更新索引，无需重建整个索引，也无需局部重建索引。即数据增删改后立即能够被检索。
- 自动分库（Partitioning）：充分利用多库并行检索技术，进一步提高了检索速度；使得数据加载速度保持稳定，进一步提高了数据更新速度。
- 支持 GB18030 和 UTF8 大字符集：便于对中文偏僻字和世界各国文字的支持。
- Native XML 支持：能够为更精确的检索提供存储和检索手段。XML 文件无需关系数据库一样分解后存储，支持 XML 的全息索引，即任何标记，以及标记的任何属性，都可以用来描述检索目标。
- 中文字词混合索引：由于中文的特殊性，存在字词索引的问题，按字索引保证百分之百的查全率；按词检索保证一定的查准率，并且可以维护分词词典，通过对词典的训练，

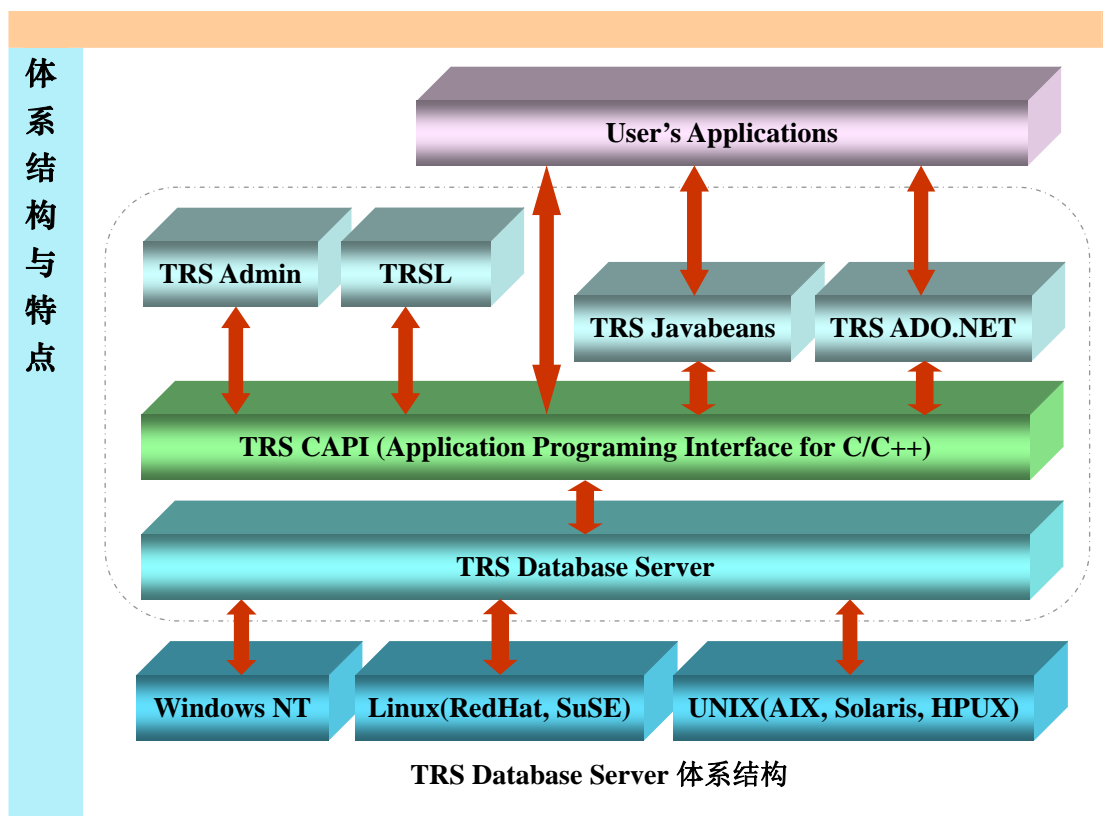
逐步提高检索查准精度。所以中文检索系统需要提供一种索引方式，以保证一定查准率的同时提供百分之百查全的手段。

- 分布式与负载均衡：提供分布式检索和负载均衡集群服务器，实现分布式检索和负载均衡功能，满足大数据量和高并发的检索要求。
- 快速返回结果：重新设计的检索算法更加高效快速，同时还具有快速返回部分检索结果的能力，使应用展现更加灵活方便。
- 集成性：提供运行版，支持与其它产品的集成；支持第三方提供的文档过滤插件，以满足用户对非文本数据的检索要求。
- 优化的 C/S 连接：不再需要复杂的连接池处理就能满足应用的扩展性，具有和 MySQL 类似的特点（MySQL does not require complex connection pooling to deliver scalability）。

TRS 全文数据库系统由服务器、管理工具和开发接口三部分组成。

服务器	管理工具	开发接口
<ul style="list-style-type: none"> <li>● TRS 全文数据库服务器 6.8 (TRS Database Server6.8)</li> <li>● TRS 全文数据库运行版 6.8 (TRS Database Runtime6.8)</li> </ul>	<ul style="list-style-type: none"> <li>● TRS 管理员工具 6.5 (TRS Admin6.5)</li> <li>● TRS 命令工具 6.5 (TRSL6.5)</li> </ul>	<ul style="list-style-type: none"> <li>● C/C++应用开发接口 6.5 (TRS CAPI6.5)</li> <li>● Java 应用开发接口 6.1 (TRS Javabeans6.1)</li> <li>● ADO.Net 应用开发接口 6.5 (TRS ADO.NET6.5)</li> </ul>

## 二、主要特点



	<ul style="list-style-type: none"> <li>● 支持各种主流操作系统和主流开发平台，不依赖其它任何数据库。</li> <li>● 支持分布式检索和负载均衡集群服务器的调度。</li> <li>● 多线程设计，充分发挥 SMP 机器的优势，支持大量并发用户访问。</li> <li>● 32 位平台也支持 64 位文件读写，即支持大于 2G 或 4G 的大文件。</li> <li>● 提供运行版（TRS Database Runtime），降低 OEM 成本。</li> </ul>
自然语言处理与智能检索	<ul style="list-style-type: none"> <li>● 核心支持以下三种字符集：GB18030，BIG5 和 UTF8。GB18030 是继 GB2312 之后的关于汉字编码的国家标准，它能表示所有的 UNICODE 字符，对于一般应用来说，GB18030 比 UNICODE 优越之处在于它与 GB2312/GBK 全面兼容，具有更短的编码。</li> <li>● 客户端支持以下三种字符集：GB18030 编码，BIG5 编码，UTF8 编码；开发接口支持 UCS2/UTF16。方便了多语言检索应用程序的开发。</li> <li>● 内嵌汉语分词：统计建立了大量歧义排除规则，有效提高了分词准确性，同时在不肯定情况下采取冗余切分，极大地提高了查全率和查准率。</li> <li>● 除汉语外，还支持藏文（含扩展集 A 和 B）、蒙文、维文、彝文等少数民族语言的切词，能够对包括拉丁文、俄文（西里尔文、斯拉夫语）、阿拉伯文和希腊文等几乎所有字母文字的识别与分词。</li> <li>● 支持英文、法文、德文和俄文等 10 多种语言的词根处理。在不进行词根索引的情况下也支持英文同根词的扩展检索。</li> <li>● 支持按词索引、按字索引、字词混合索引，追求高查准率的同时提供百分之百查全的手段，实现按需检索。</li> <li>● 扩展检索：包括主题词典自动扩展检索、同义词和反义词自动扩展检索、全半角自动扩展检索、简繁体自动扩展检索。扩展检索不仅支持 PHRASE 型字段、DOCUMENT 型字段，同时也支持 CHAR 型字段。</li> <li>● 检索相关性计算：输入一个检索串或一句话，分词后进行高频词过滤、词语个数限定、词语之间的位置关系等处理，根据词语在文章中的位置和频度，以及词语的倒文档频度，计算相关度，按相关度高低排序后输出结果。</li> <li>● 自然语言检索和扩展检索结合使用，能够实现一定程度的智能检索（使用 LIKE 函数和 INCLUDE 函数等）。</li> </ul>
系统与数据库管理	<ul style="list-style-type: none"> <li>● 支持 Text、HTML、XML、RTF、MS Office 文档（Word/Execl/Powerpoint）和 PDF 文件自动建立索引。</li> <li>● 提供字段的唯一值（Unique）特性，用于数据的一致性检查。</li> <li>● 管理员可设置“定时优化”的时间，降低系统管理的成本。</li> <li>● 支持 NUMBER 型字段的五种子类型，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）以及 64 位二进制有符号整数（18 位十进制有符号整数）。</li> <li>● 支持 CHAR 型字段的冗余索引，避免模糊检索以提高检索速度。</li> <li>● 支持检索优先机制：在数据库的维护过程中，如果有检索请求在执行，则在检查点“睡眠”指定的秒数，以让出 CPU 时间来尽快地完成检索请求。</li> <li>● 支持选择一些常用字段集中存储，提高排序、统计分类和概览的速度。</li> <li>● 支持分词词典、停用词典、主题词表、同义词、反义词典、禁用词典的维护。</li> </ul>

- 分词词典、停用词典维护后无需重新启动将立即生效。
- 支持优化的索引结构，使得检索时的相关性计算更加方便高效。
- 支持基于 BIT 字段的虚拟逻辑字段。通过虚拟逻辑字段，能够方便地在异构数据库之间构建统一的检索应用。

## 检索功能

- 全方位检索手段：提供了多达 40 多种检索运算符。包括外部特征与文本内容的各种逻辑组合检索、位置检索、二次检索（检索历史引用）、词根检索、大小写敏感检索。
- 多种检索入口：属性字段检索（如分类检索）、关键词检索、位置检索（如同段检索）、多字段“与或非”组合检索、二次检索（渐进检索）、大小写敏感检索、中文简繁体扩展检索、支持超长检索表达式等。
- 知识扩展检索：检索时能够应用同义词典、反义词典和主题词典进行自动或函数式的扩展检索。
- 基于成本优化的查询算法：高效的索引压缩技术、SKIP 索引跳跃式扫描技术、多库并行检索技术、表达式优化技术、结果共享技术、CACHE 技术。CACHE 技术支持更多的并发用户访问，大大提高了综合查询速度。
- 支持对检索结果的各种排序：检索结果与检索词的相关度排序；基于字段的排序；后进先出的快速排序。同时对多库检索结果可以进行混排。LIFO（后进先出的快速排序）对实际运行系统的性能提高有很大作用。TOP N 排序可通过结果记录数的裁减以减少检索历史的存储来提高检索速度。
- 相关性排序考了多种加权因素：命中点的数量、词出现的概率、记录的文本长度、命中点的绝对位置、命中点之间的相对位置和次序等。
- 支持多个文本字段（PHRASE 型、DOCUMENT 型）和非文本字段（CHAR 型、NUMBER 型、CHAR 型）参与相关性排序。支持字段之间的加权以及排序优先次序。
- 支持检索结果的分类统计，支持用正则表达式来筛选分类统计的目标字段值，并支持抽取部分子串组成新的类别。
- 支持精确检索，准确报告检索记录数；支持估算检索，快速返回部分结果，并对结果集进行估算；同时支持对结果进行补充检索与重新估算。
- 支持短语级别（INCLUDE 函数）和词级别（LIKE 函数）的“相似性”检索，INCLUDE 函数支持 CHAR 字段的运算。
- TOP N 剪裁排序时，LIKE 和 INCLUDE 函数具有“匹配度自适应调节”功能，以最大限度地提高检索效率。
- 支持中文按拼音或笔画排序，避免中文按内码排序时的“无序”问题。
- 支持英文的同根词检索。在没有词根索引的情况下，也能检索与输入的词具有相同词根的词。
- 支持拼写组合的检索，允许基于“连接规则”、“排除规则”与“选择规则”的三种拼写组合。
- 最大检索时间和最大排序记录数限制，保护系统资源的使用。并支持每个检索请求的独立的“超时”设置，允许检索被“中断”后返回已经得到的结果。
- 支持基于 BIT 字段的虚拟逻辑字段的检索，并支持其实体字段之间的逻辑关系与排序加权。
- 实现“同字段”的限定运算，支持复杂条件下的检索需求。

**安全管理**

- 系统提供多种权限级别的用户管理。具有系统级、数据库级、记录级和字段级四级安全控制机制。
- 支持用户组的概念，以及用户组管理员（GDA）的概念，适应大型应用中复杂的系统管理和应用。
- 支持数据的加密与压缩传输。支持广域网的用户登录。
- 除了可对数据库进行联机备份外，还提供系统的联机备份。
- 较完善的日志管理，可以根据需要灵活配置。
- 运行版提供了数据备份时的加密功能（需要定制），可基本保证数据在发行环节的安全性。

**开放性**

- 支持主流开发平台，提供 CAPI、Javabeans 和 ADO.NET 二次开发接口，通过这些接口，应用可以完成对系统和数据的管理与操作。
- 支持第三方提供的文档过滤插件，以满足用户对非文本数据的检索要求。（需要按规范提供接口与动态链接库）
- 很多全文检索系统只针对网页，缺乏其他格式文档的灵活性。TRS 数据库没有定义具体的数据源，而是一个开放的数据库。支持多种入库格式。
- 很多全文检索系统往往是把数据当成一个单一的文档进行整体索引。TRS 数据库系统按记录组织数据，每个记录又由多个字段组成，可以控制哪些字段建立索引，哪些字段不建立索引。更进一步地，需要索引的字段还可分为需要分词后索引的字段和不需要分词直接索引的字段：需要进行分词后索引的字段，如：标题、正文字段；直接按值索引的字段，如：作者、日期字段。

### 三、性能指标

#### 3.1 功能指标

项目	性能指标	备注
数据加载和索引速度	单机可达每小时 3G（单进程装库）。 单机可达每小时 6G（双进程装库）。	测试环境：Dell2850 或 IBM X346, 2CPUs, 4G 内存, 10000 转硬盘, Redhat AS4.0 (64 位)。除此之外，索引速度还与具体的索引方式、字段数量以及数据本身等因素有关有关。
用户组/用户数	多达 65535 个用户组/用户。	
用户类型	(1) 超级用户（SYSTEM）。 (2) 系统数据库管理员（DBA）。 (3) 用户组数据库管理员（GDA）。 (4) 系统资源用户（RESOURCE）。 (5) 系统登录用户（CONNECT）。	

项目	性能指标	备注
支持的字符集	(1) 简体中文 (GB18030)。 (2) 繁体中文 (BIG5)。 (3) 纯英文 (ENGLISH)。 (4) 统一码 (UTF8)。	支持中英文混合。
镜像功能	增量镜像。	不支持自动模式视图的镜像。
数据库/视图数	多达数十亿个数据库/视图。	
数据库大小	每个数据库可加载 42 亿条记录，数据库大小没有限制。	对于 32 位字长的机器或操作系统，也能支持大于 4G 的数据库；如果数据库中删除过记录，则该数据所能容纳的记录数就会减少。
视图包含的数据库个数	每个视图中可包含 1 到 128 个数据库。	
数据库/视图包含的字段个数	每个数据库可定义 1 到 1023 个字段。	
同时检索的数据库个数	没有限制。	
字段类型	(1) DATE (日期型字段)。 (2) NUMBER (数值型字段)。 (3) CHAR (字符串型字段)。 (4) PHRASE (短语型字段)。 (5) DOCUMENT (全文型字段)。 (6) BIT (二进制型字段)。	

## 3.2 查询性能

TRIS 查询速度很快，在千万级记录的数据库上，也能获得亚秒级查询速度。实际查询速度和很多因素有关，包括硬件平台、并发用户数、数据库规模、检索条件的复杂性等，下表给出的查询速度测试结果，仅供参考。

(1)新闻数据库(1000万/2000万/4000万篇)分别在 1/2/4 台 Dell2850 的**集群环境**下的查询速度（未启动 Cache）测试结果：

数据量	Dell2850 台数	速度(秒)	
1000 万	1	关键词检索	0.075
		逻辑表达式	0.309
2000 万	2	关键词检索	0.076
		逻辑表达式	0.314
4000 万	4	关键词检索	0.103
		逻辑表达式	0.377



(2)新闻数据库(1000 万篇) 分别在 1/10/50/100 并发数情况时**单机环境**下的查询速度(未启动 Cache) 测试结果:

数据量	并发数	速度(秒)	
1000 万	1	关键词检索	0.08
		逻辑表达式	0.31
	10	关键词检索	0.25
		逻辑表达式	0.57
	50	关键词检索	1.26
		逻辑表达式	4.75
	100	关键词检索	2.50
		逻辑表达式	13.12

说明:

- 关键词检索: 指非结构化数据的全文检索。
- 逻辑表达式: 指非结构化数据和结构化数据的联合检索。
- Dell2850: 2 CPUs, 4G 内存, 10000 转硬盘, 64 位 Redhat AS4.0。

## 四、产品组件简介

### 4.1 TRS全文数据库服务器 6.8

简称 TRS Database Server6.8, 它是 TRS 全文数据库系统的核心, 以服务器方式运行, 为全文数据的管理和检索提供原动力。

**主要功能:**

- 系统控制: 连接和登录管理, 权限检查, 设置服务器主语言, 服务器和客户端通讯, 无用资源回收等。
- 对象和权限管理: 创建和维护数据库、视图、词、用户组、用户和文件格式; 用户级, 数据库级, 字段级以及记录级的权限管理。
- 数据维护: 加载数据 (Load), 插入数据 (Insert), 修改数据 (Update) 以及删除数据 (Delete), 包括数据库和词典的数据维护。
- 检索和结果处理: 全文检索及扩展检索; 输出检索结果。
- 操作日志。

**新增特点**

- 优化存储结构, 提高相关排序性能。TRS Server 6.8 对存储结构进行了优化与改进, 使得在检索过程中能够即时得到计算相关度所需要的数据, 避免了额外的文件访问和计算的时间开销, 从而有效地提高了相关排序性能。
- 估算检索与快速返回。TRS Server 6.8 采用广度优先的计算策略, 实现了估算检索与快速返回的检索机制, 在不需要对整个库进行完全检索的情况下就能返回给应用部分结果, 以达到快速展现检索结果的应用需求。

- 中文的拼音与笔画排序。以前对中文的排序都是基于字符集的内码次序的，而人们习以为常的顺序是拼音序，但内码序与拼音序基本一致的只有 GB2312 字符集所涵盖的中文，其它字符集的内码序与拼音序都是不一致的，尤其是 UTF8 字符集。TRS Server 6.8 实现了中文按拼音与笔画排序的算法，彻底解决了中文排序问题。按拼音排序还支持多音字的区分。
- 同根词检索。在没有词根索引的情况下，TRS Server 6.8 也能进行英文同根词的扩展检索，无需创建额外的冗余索引。避免了选择词根索引后，不能区分英文单词的不同形态进行精确检索的问题。
- 按优先次序的多字段相关性排序。TRS Server 6.8 在原有多字段相关性排序的基础上增加了一种新的模式，把字段之间的加权值当作字段之间的优先次序，以保证检索结果严格按照应用对字段优先次序的要求进行排序，避免纯粹的“加权”排序方式所引起的结果顺序不确定的问题。
- 考虑更多的加权因素，改进相关性效果。TRS Server 6.8 对相关性模型进行了微调，增加了更多的调节因子，自动对各种因素进行加权，使排序结果更加合理、更加贴近应用要求，有效改善了在排序效果方面的用户体验。
- 增强 INCLUDE 函数功能。TRS Server 6.8 对 INCLUDE 函数进行了两方面的改进：实现短语级别的“相似性”检索；同时允许在 CHAR 字段上实施“相似性”检索。提高了检索结果的准确性，解决了 LIKE 函数（词级别的“相似性”检索）不能满足有些应用需求的问题。
- 匹配度的自适应调节。TRS Server 6.8 的 INCLUDE 和 LIKE 函数的“相似性”检索在与 TOP N 剪裁排序方式一起使用时，允许应用使用一个较低的“匹配度”，以此为起点，系统在检索过程中将根据记录的命中情况逐步提高“匹配度”，以快速淘汰或否决“匹配度”低记录，为应用筛选并保留那些“匹配度”更高的记录作为最终检索结果集。既降低了应用的构建难度、又提高应用的检索效率。
- 虚拟逻辑字段。TRS Server 6.8 实现了基于 BIT 字段的虚拟逻辑字段，便于复杂的应用在异构数据库之间构造统一的检索需求。一个虚拟逻辑字段可以包含多个实体字段，实体字段之间可以定义逻辑关系、排序权重，以及对表达式的作用方式（是表达式逻辑优先，还是字段逻辑优先）。
- 同字段检索限定。有些复杂的应用的检索表达式，传统的“段、句、位”检索限定还不足以满足运算需求。尤其是支持了虚拟逻辑字段后，有时在虚拟逻辑字段上的检索会认为就是对单个字段的检索，因此会在表达式中加入“段、句、位”这样的限定，这在以往的版本中是不允许的。为此，TRS Server 6.8 增加了同字段检索限定，以便于这种复杂应用的构建。

#### 平台与配置：

- Windows（Windows2000/Windows2003/Windows2008/Windows7； X86/X64 平台）
- Linux（Red Hat AS4.0/ Red Hat AS5.0/SuSE ES10/红旗 DC5.0； X86/X64 平台）
- Solaris（Solaris8/Solaris9/Solaris10； SPARC/X64 平台）
- AIX（AIX5.1/AIX5.3； POWER4 平台）
- HPUX（HPUX11.11； PA-RISC 平台）
- 内存：4G 以上，建议 16G
- 硬盘：根据应用数据量的需要进行配置



## 4.2 TRS全文数据库运行版 6.8

简称 TRS Database Runtime6.8, 是在 TRS 全文数据库服务器的基础上发展出来的, 目的是方便与其它 TRS 产品进行捆绑, 或者满足一些 OEM 项目的需求。它继承了 TRS 全文数据库服务器的性能和绝大部分功能, 区别是:

- 只能创建预先定义的固定结构的数据库。
- 不能修改已经创建了的数据库的结构。
- 不能创建、修改用户组信息。
- 不能创建、修改视图。
- 不能创建新的词典。
- 不支持数据库镜像。
- 无 Licence 限制。
- 支持数据备份的加密。

另外, TRS 全文数据库运行版与 TRS 全文数据库服务器不能互换; TRS 全文数据库服务器的数据库备份不能恢复到 TRS 全文数据库运行版中; 但 TRS 全文数据库运行版的数据库备份则能恢复到 TRS 全文数据库服务器中。

### 新增特点

- 对于有加密要求的运行版, 则其数据库备份只能在相同的应用系统里恢复。
- 同时在导出数据库记录时也支持使用特定的 TRS 装库格式文件加密后存储。

### 平台与配置:

- Windows (Windows2000/Windows2003/Windows2008/Windows7; X86/X64 平台)
- Linux (Red Hat AS4.0/ Red Hat AS5.0/SuSE ES10/红旗 DC5.0; X86/X64 平台)
- Solaris (Solaris8/Solaris9/Solaris10; SPARC/X64 平台)
- 内存: 4G 以上, 建议 16G
- 硬盘: 根据应用数据量的需要进行配置

## 4.3 TRS管理员工具 6.5

简称 TRS Admin6.5, 是操作 TRS 全文数据库服务器的管理工具, 主要用于对各种 TRS 对象进行管理、设计和维护等, 并具备完整的检索、浏览和维护功能。

### 主要功能:

- 完全 UNICODE 支持: TRS 管理工具采用了 Win32 的 UNICODE 版本的 GUI 图形界面, 并尽可能与常用应用程序的界面风格 and 操作方法保持一致, 实现完全的 UNICODE 支持。
- 集所有功能于一身: TRS 管理工具把各功能集于一身, 减少了不必要的切换和重复劳动。界面各部分风格一致, 紧密融合, 互相协调。

- 更为方便直观的对象管理方法：新的对象管理工具界面类似于 Windows 的资源管理器，可用于管理和维护各种 TRIS 系统对象，可直接对各种 TRIS 系统对象进行查看、改名、删除等操作。
- 对象和数据的移动与复制：通过剪贴板可方便地实现对象的移动与复制，减少重复劳动，提高工作效率。不仅可以在不同的 TRIS 服务器之间移动或复制 TRIS 系统对象，还可以在同一台服务器上创建对象的副本。
- 方便快捷的数据库/视图的设计界面：新的数据库/视图设计界面可直观地设置或修改数据库/视图的属性和结构，并可方便地管理各用户组/用户对其的操作权限。
- 浏览和维护更为得心应手：新的浏览和维护界面把窗口分为三个部分：导航、概览和细览。导航部分用于快速检索特定的记录，概览部分主要用于浏览日期、数值、字符串等类型的字段内容，细览部分主要用于浏览正文/二进制等大型字段的详细内容。
- 在集成环境里直接浏览格式化文档：TRIS 管理员工具的浏览和维护界面里集成了 Microsoft 的 Internet Explorer，可以在不离开集成环境的情况下直接浏览各种格式化文档，并可快速地在文本和源文件之间进行切换。
- 内嵌文件编辑器：TRIS 管理员工具有不同的文件向导，用于快速地生成 TRIS 的导航文件、装库数据文件和装库控制文件等，并能方便对这些文件及其它文本文件进行编辑。

#### 新增特点：

- 删除对象和删除记录功能分开，防止误操作。
- 数据库字段长度限定、索引属性等增加弹出框进行设置或选择。
- 高级检索和检索选项对话框作了调整，并增加了一些新选项。
- 保存检索结果增加记录引导前缀和每批读取记录数设置；失败后可选择重试。
- 记录维护增加了以文件方式修改记录，以及设置数据加载参数功能。
- 表单设计的“网格设置”改为“布局设置”，增加了标签宽度选项。
- 备份/恢复数据的路径支持目录名，本地服务器下可选择。
- 执行结果报告可查看超长的错误信息。
- 选项对话框缺省登陆属性页，增加代理服务器等相关选项。
- 关于对话框增加“系统信息”按钮，显示服务器系统和设备信息。
- 统一错误报告方式，增加错误号显示。
- 增加了 Server 新功能的支持。

#### 平台与配置：

- 简体中文版 Windows2000/Windows2003/Windows2008/Windows7；X86/X64 平台
- 内存：1G 以上
- 硬盘：32G 以上

## 4.4 TRS命令工具 6.5

简称 TRSL6.5,是操作 TRS 全文数据库服务器的命令语言式工具,主要用于对各种 TRS 对象进行管理、设计和维护等,并具备完整的检索、浏览和维护功能。

### 主要功能:

- 全新设计的界面,风格清新,易用性也得到了很大提升。
- 重新整理和规范了 TRSL 语法,并新增了新的语句和语法元素。
- 采用类似 SQL 的操作命令。
- 提供二次开发接口。
- 支持脚本文件。把需要完成的任务写成 TRS 脚本文件,既可以在 TRSL 命令语言工具里运行,也可以在操作系统的命令行下通过该命令工具直接运行。脚本文件在操作系统的命令行下直接运行的功能配合 Windows 的任务计划可以实现定时完成一系列管理工作的需求。

### 新增特点:

- 采用标签页式多文档管理。
- 主界面采用了分割式窗体,将输入的语句命令和输出的结果显示分开;输出窗口分为执行结果视图、概览视图和细览视图,可以直接浏览检索结果;命令窗口实现了语句中不同类型词法符号的彩色显示。
- 语句输入过程显示选项列表以供选择,并可记住曾经输入过的历史列表。
- 显示语法范式作为提示,并可标示当前输入符号在范式中的位置。
- 修改装入记录等语句的进度显示方式,可实现进度百分数等信息不断刷新显示。
- 一些语句增加了提示确认的功能,防止误操作。
- SHOW 语句输出结果列表过程中增加暂停和中断输出功能。
- 密码输入增加保护措施,防止泄密。
- 检索历史、操作日志和用户变量以窗口的形式展现,更加直观。
- 支持打开多个脚本文件进行编辑,同样提供了命令行中的辅助输入功能。
- 脚本运行过程支持 Ctrl+Break 中断。
- 增加了定时执行脚本功能。

### 平台与配置:

- 简体中文版 Windows2000/Windows2003/Windows2008/Windows7; X86/X64 平台
- 内存: 1G 以上。
- 硬盘: 32G 以上。

## 4.5 TRS C/C++应用开发接口 6.5

简称 TRS CAPI6.5,是访问 TRS 全文数据库服务器的 C/C++应用程序开发接口。可在 Visual C++、Visual Basic、Borland C++、Borland Delphi、WEB 和 UNIX 等开发环境下使用。

**主要功能:**

- 系统控制: 连接和登录管理,权限检查, 服务器和客户端通讯等。
- 对象和权限管理: 创建和维护数据库、视图、词典、用户组、用户和文件格式; 用户级、数据库级、字段级以及记录级的权限管理。
- 数据维护: 加载数据 (Load), 插入数据 (Insert), 修改数据 (Update), 以及删除数据 (Delete), 包括数据库和词典的数据维护。
- 检索和结果处理: 全文检索及扩展检索; 读取概览与细览字段的内容。

**新增特点:**

- UNICOD 支持: 原来版本只能识别 GB、BIG5、UTF8 三种字符集, 不能直接处理 UNICOD, 这给 ADO.NET、Javabeans 等接口设计带来不少麻烦。新版本在 API 级统一实现了这一支持, 所有调用参数均可采用 UNICOD (UCS2/UTF16), 返回值也以 UNICOD (UCS2/UTF16) 形式提供, 只须在设置字符集时指定一个标志即可。
- 负载均衡: 可由用户指定一系列的服务器, 由 API 根据轮回方式选择连接相关服务器。此时返回的访问许可证包含服务器地址和端口信息, 可用于恢复连接。
- 属性存取: 以结构形式保存对象属性有不少缺陷, 如长度受限、不易扩展、在某些语言下调用困难等。新的属性机制引入对象概念, 不仅可以直接根据属性名取属性值, 还能将对象所有属性以字符串形式备份, 以便于重新恢复, 同时还保证了更好的可扩展性。
- 命中点标注: 可指定前后添加的标志和截断长度。
- 附件分目录存放: 主要针对保存检索结果时附件过多的情况, 可按指定记录数分组存放附件, 自动创建新目录存放。
- 装库错误文件: 装库失败以后, 根据返回的报告信息, 自动生成装库错误文件, 便于用户查看和修改后重新装库。这原本是在 Admin 的功能, 下移至 API 实现, 以便于其他接口和应用调用。
- 增加了 Server 新功能的支持。

**平台与配置:**

- Windows (Windows2000/Windows2003/Windows2008/Windows7; X86/X64 平台)
- Linux (Red Hat AS4.0/ Red Hat AS5.0/SuSE ES10/红旗 DC5.0; X86/X64 平台)
- Solaris (Solaris8/Solaris9/Solaris10; SPARC/X64 平台)
- AIX (AIX5.1/AIX5.3; POWER4 平台)
- HPUNIX (HPUNIX11.11; PA-RISC 平台)
- 内存: 512M 以上

## 4.6 TRS Java应用开发接口 6.1

简称 TRS Javabeans6.1, 是访问 TRS 全文数据库服务器的 Java 应用程序开发接口, 具有良好的跨平台特性。用户应用界面可以是基于浏览器的 JSP 或 Java Applet, 也可以是独立的 Java 应用程序, TRS Javabeans 为这些应用提供了统一的调用接口。

**主要功能:**

- TRSConnection: 该类的实例表示与 TRS 服务器已建立的连接。可以在该实例上进行数据库备份与恢复、数据增删改、检索、相关配置信息的获取与修改等操作。
- TRSException: TRS 异常处理对象,实现对操作过程中产生的异常进行捕捉与反馈。
- TRSResultSet: 该类用来保存和获取查询结果集的各种信息。
- TRSRecord: 记录对象,表示记录集中的当前记录。
- TRSField: 记录字段对象,表示记录集中当前记录的指定字段。
- TRSHitPoint: 命中点对象,表示命中点的位置信息。
- TRSObject: TRS 对象基类,用来定义 TRS 对象的一些通用方法,如获取设置属性,取消设置,列出属性值等,并以此为基类派生出以下各子类: TRSDataBase (数据库)、TRSDataBaseColumn (数据库字段)、TRSView (视图)、TRSViewColumn (视图字段)、TRSGroup (用户组)、TRSUser (用户)、TRSFormat (数据格式)。

**平台与配置:**

- Windows (Windows2000/Windows2003/Windows2008/Windows7; X86/X64 平台)
- Linux (Red Hat AS4.0/ Red Hat AS5.0/SuSE ES10/红旗 DC5.0; X86/X64 平台)
- Solaris (Solaris8/Solaris9/Solaris10; SPARC/X64 平台)
- AIX (AIX5.1/AIX5.3; POWER4 平台)
- HPUNIX (HPUNIX11.11; PA-RISC 平台)
- 内存: 512M 以上。

## 4.6 TRS ADO.Net应用开发接口 6.5

简称 TRS ADO.NET6.5,是访问 TRS 全文数据库服务器的 DOT Net 应用程序开发接口,可在 DOT NET 开发环境下为 TRS 全文数据库服务器开发应用。它建立在 TRS CAPI6.5 之上,实现了与 TRS Javabeen 一致的功能和方法接口,使用户能把主要精力集中在实现应用的功能上。

**主要功能:**

- TRSConnection: 该类的实例表示与 TRS 服务器已建立的连接。可以在该实例上进行数据库备份与恢复、数据增删改、检索、相关配置信息的获取与修改等操作。
- TRSException: TRS 异常处理对象,实现对操作过程中产生的异常进行捕捉与反馈。
- TRSResultSet: 该类用来保存和获取查询结果集的各种信息。
- TRSRecord: 记录对象,表示记录集中的当前记录。
- TRSField: 记录字段对象,表示记录集中当前记录的指定字段。
- TRSHitPoint: 命中点对象,表示命中点的位置信息。
- TRSObject: TRS 对象基类,用来定义 TRS 对象的一些通用方法,如获取设置属性,取消设置,列出属性值等,并以此为基类派生出以下各子类: TRSDataBase (数据库)、TRSDataBaseColumn (数据库字段)、TRSView (视图)、TRSViewColumn (视图字段)、TRSGroup (用户组)、TRSUser (用户)、TRSFormat (数据格式)。

**平台与配置:**

- Windows (Windows2000/Windows2003/Windows2008/Windows7; X86/X64 平台)
- 内存: 512M 以上

## 五、相关产品及说明

### 5.1 相关产品

<b>TRS Cluster</b>	用于将多个 TRS 全文数据库系统构造一个集群系统,实现分布式检索和负载均衡。利用对 TRS 全文数据库的分组架构,不仅解决了负载均衡问题,同时也提高了整个系统的可用性。
<b>TRS Gateway</b>	负责关系型数据库与 TRS 全文数据库之间的数据自动迁移和同步更新,解决了关系数据库大对象的检索和分析效率问题,而上层应用则可以构架在由关系数据库和 TRS 全文数据库共同组成的数据库平台之上。支持 Oracle、Sybase、DB2、SQL Server、MySQL。
<b>TRS WAS</b>	用于以 Web 网页形式发布展示后台数据的内容。以 Web 应用方式提供对后台数据对象的管理与内容的维护功能。
<b>TRS Inforadar</b>	负责 Web 网页实时监控和采集目标网站的内容,对采集到的信息进行过滤和自动分类处理,最终将最新内容及时发布出来,实现统一的信息导航功能,同时提供包括全文检索、日期(范围)检索、标题检索、URL 检索等在内的全方位信息查询手段。
<b>TRS CIS</b>	负责将反映自身、竞争对手和外部环境的时间状态和变化的数据、信息、情报进行收集、存储、处理、分析,并以适当的方式将分析结果发布给有关战略管理人员。
<b>TRS CKM</b>	<p>文本挖掘是进行知识管理的基础和关键技术。TRS CKM 是国内外第一套成熟的中文文本挖掘技术,主要功能包括:</p> <ol style="list-style-type: none"> <li>(1) 文本分类: 自动分类/规则分类/混合分类;</li> <li>(2) 文本聚类;</li> <li>(3) 文本相似性检索和文章自动排重;</li> <li>(4) 文本摘要和关键词标引;</li> <li>(5) 文本信息过滤;</li> <li>(6) 拼音检索、同音检索;</li> <li>(7) 相关短语检索(用于检索词提示);</li> <li>(8) 信息抽取等。</li> </ol> <p>在检索系统中的主要用途:</p> <ol style="list-style-type: none"> <li>(1) 检索词提示和相关文章提示;</li> <li>(2) 预处理时的文章分类和文章去重;</li> <li>(3) 检索结果的聚类。</li> </ol>



## 5.2 主要平台

<b>Windows</b>	OS: Windows2000/Windows2003/Windows2008/Windows7。 CPU: X86/X64。
<b>Linux</b>	OS: Red Hat AS4.0/ Red Hat AS5.0/SuSE ES10/红旗 DC5.0。 CPU: X86/X64。
<b>Solaris</b>	OS: Solaris8/Solaris9/Solaris10。 CPU: SPARC/X64。
<b>AIX</b>	OS: AIX5.1/ AIX 5.3。 CPU: POWER4。
<b>HPUX</b>	OS: HPUX11.11。 CPU: PA-RISC。
<b>推荐平台</b>	PC 服务器: 2X4 核 CPU, 16G 内存, 64 位 Linux2.6 (Red Had AS5.3 或 SuSE Enterprise Server10 SP2)。

## 5.3 运行许可证

产品运行有许可证 (License) 控制, License 即指瞬间 TRS 全文数据库服务器的能够同时执行的最大请求数, 也就是允许在同一时刻并发与 License 相适应的进程与线程数并行地对数据库进行相关操作 (查询、修改、下载、删除等功能)。超过 License 所规定的最大请求数时, 将排队等待服务。License 与同时登录到 TRS 服务器的用户数实例数没有直接的关系, 但登录请求的服务过程还是受 License 的控制。