

TRS 全文检索网关

TRS Gateway 4.6

用户手册

北京拓尔思信息技术股份有限公司
Beijing TRS Information Technology Co., Ltd.

版权说明

本手册由北京拓尔思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 是北京拓尔思信息技术股份有限公司的注册商标。

关于本手册

本手册介绍的主要内容可以分为五部分，第一部分为系统概述，产品特点，系统需求以及产品变化；第二部分为客户端使用说明，详细介绍产品的界面，功能以及使用；第三部分为服务器端使用说明，详细介绍产品的界面和使用；第四部分为辅助工具使用说明，介绍日志分析工具，配置文件转换工具的使用；第五部分为附录。

读者对象

本手册的读者为 TRS 和 RDBMS 的使用者。

手册组织

本手册的内容组织如下：

第 1 章 TRS 全文检索网关系统介绍

概述了 TRS 全文检索网关系统，介绍了产品的特点，描述了产品的系统需求以及与旧版本相比发生的变化。

第 2 章 客户端使用说明

详细描述 Windows 操作系统上 TRS 全文检索网关客户端的使用方法。

第 3 章 服务端使用说明

详细描述 Windows、Linux 平台下 TRS 全文检索网关服务端的使用方法。

第 4 章 辅助工具使用说明

详细介绍日志分析工具和配置文件转换工具使用方法。

附录 1 字段对应表

附录 2 Windows 环境 RDBMS 客户端配置

附录 3 Linux 环境 RDBMS 客户端配置

附录 4 扩展功能插件编写规范

附录 5 正则表达式

附录 6 常见问题解答

相关手册

《TRS DataBase 全文数据库系统使用指南》

《TRS 集群服务器（2x）使用手册》

《TRS CKM 功能手册》

用户反馈

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。

目 录

第 1 章 TRS 全文检索网关系统介绍	1
1.1 概述.....	1
1.2 主要特点.....	2
1.3 系统需求.....	2
1.4 产品变化.....	2
第 2 章 客户端使用说明	4
2.1 启动界面.....	4
2.2 服务器交互.....	5
2.3 任务的视图.....	7
2.4 创建任务.....	7
2.5 字段扩展.....	29
2.6 高级设置.....	31
2.7 查看和修改任务属性.....	34
2.8 选项菜单.....	37
2.9 版本信息查看.....	39
第 3 章 服务端使用说明	41
3.1 WINDOWS 网关服务注册与启动.....	41
3.2 LINUX 网关服务启动.....	42
3.3 网关服务的停止.....	43
3.4 网关注册码升级.....	43
3.5 扩展插件 CKM 的使用.....	45
3.6 服务器部分文件及参数说明.....	45
3.7 服务端支持的命令参数.....	47
第 4 章 辅助工具使用说明	48
4.1 日志分析工具.....	48
4.2 配置文件转换工具.....	49
附录 1 字段对应表	50
1.1 RDBMS 和 TRS 字段对应表.....	50
附录 2 WINDOWS 环境 RDBMS 客户端配置	55
2.1 ORACLE 客户端的配置.....	55
2.2 SQL SERVER 客户端的配置.....	58
2.3 SYBASE 的客户端配置.....	59
2.4 DB2 的客户端配置.....	60
2.5 MySQL 的客户端配置.....	65
附录 3 LINUX 环境 RDBMS 客户端配置	66
3.1 ORACLE 客户端的配置.....	66
3.2 DB2 客户端的配置.....	70
3.3 MySQL 客户端的配置.....	70

3.4 SYBASE 客户端的配置.....	70
3.5 SQL SERVER 客户端的安装和配置.....	71
3.6 其他 ODBC 数据源配置.....	73
附录 4 扩展功能插件编写规范.....	74
附录 5 正则表达式.....	78
5.1 正则表达式部分语法.....	78
5.2 正则匹配举例.....	78
附录 6 常见问题解答.....	80
6.1 创建任务.....	80
6.2 运行任务.....	80
6.3 FAQ.....	81

第1章 TRIS 全文检索网关系统介绍

1.1 概述

TRIS 全文检索网关 (TRIS Gateway) 是由北京拓尔思 (TRIS) 信息技术股份有限公司与国内、外各大数据库厂商鼎力合作, 推出的实现关系型数据库向 TRIS 全文检索数据库进行数据迁移的工具, 该工具实现了关系型数据库与 TRIS 全文检索数据库或集群之间的数据共享, 使用户在享有关系型数据库系统卓越的数据处理功能的同时, 并拥有 TRIS 优秀的全文检索功能。当前支持的关系型数据库系统有 Oracle, Sybase, DB2, SQL Server, MySQL, 人大金仓 (KBE) 等 (以下统称为 RDBMS)。

TRIS 全文检索网关最新版本采用 C/S 模式架构, 将原有的单机版本剥离开来, 网关产品服务端可运行于多种平台 (Windows/Linux), 更有针对不同位数 (32 位/64 位) 操作系统的单独版本。

注意: 当前网关的客户端产品和服务器端的日志分析工具等只能运行在 Windows 平台上。

TRIS 全文检索网关组成部分为:

服务端:

- 服务主体程序 TrsAgent.exe(Windows)或 TrsAgent.sh(Linux)
- 服务关闭程序 gw_shutdown.exe(Windows)或 shutdown.sh(Linux)
- 服务监控工具 GatewayMonitor.exe(Windows)
- 产品注册工具 register.exe(Windows)或 register.sh(Linux)
- 配置升级工具 Upgrade.exe(Windows)或 Upgrade.sh(Linux)
- 日志分析工具 LogParse.exe(Windows)
- Oracle, Sybase, DB2, SQL Server, MySQL, KBE 的数据库组件 rdb_*.dll (Windows) 或 librdb_*.so(Linux) (64 位网关仅提供 Oracle、DB2、SQL Server、MySQL 的数据库组件)
- 服务端依赖组件*.dll(Windows)或 lib*.so(Linux)

客户端:

- 客户端主体程序 TrsAgentClient.exe
- 客户端依赖组件*.dll

TRIS 全文检索网关是运行在相应平台上的高度自动化、智能化和定制化的任务处理工具, 它可以独立运行一个 RDBMS 数据库组件, 也可以同时运行多个 RDBMS 数据库组件。它与数据库组件共同完成:

- TRIS 格式文件到 TRIS 的数据迁移
- TRIS 数据库的优化
- RDMBS 到 TRIS 数据库的数据迁移
- RDBMS 到 TRIS 格式文件数据迁移

以上各种数据迁移与优化任务的管理、维护和定时执行等操作。

1.2 主要特点

TRIS Gateway 系列软件是辅助完成数据索引同步更新的自动化工具，它具有如下特点：

- 自动化程度高：用户只需要按步跟随任务创建向导的提示就可以创建更新任务。
- 可设置的定时执行任务：对创建好的任务，用户可以设置其自动定时执行。如：用户可以设置一个增量更新任务每隔 30 分钟执行一次，意即每隔 30 分钟将用户对 RDBMS 数据库表中数据的修改向 TRIS 数据库中进行一次索引的更新。
- 实时查看和修改任务的各种属性：对任何一个创建好的任务，用户可以查看其属性，并可修改属性。
- 美观友好的界面风格：界面友好，简单易用。

1.3 系统需求

应用程序运行的硬件环境要求：

- CPU：主频 > 200MH（主频越高性能越好）。
- 内存：> 256M
- 硬盘：请确保系统临时目录的可用空间在 500MB 以上。
- CDROM：16X 以上的光盘驱动器。

客户端应用程序运行的软件环境要求：

- 操作系统：Microsoft Windows XP/NT/2000/2003/2008/Vista/7

服务端应用程序运行的软件环境要求：

- 操作系统：Microsoft Windows XP/NT/2000/2003/2008/Vista/7、Redhat Linux、SuSE Linux
- 软件环境：相应关系数据库的客户端工具或数据库服务。

1.4 产品变化

TRS Gateway 系列软件的最新版本号为 4.6 Build4004, 较 4.6 Build4003 有以下改进:

- 支持对配置相同的多个集群服务器同时发送读写分离命令 (IDLE 和 ORDER)。配置相同的集群, 上层是通过硬件做到负载均衡。
- 支持 TRS 集群扩展, 通过配置, 多个任务共用临时表、触发器, 达到一定规则的数据入一个 TRS 集群, 另一个规则的数据入另一个 TRS 集群。
- 支持注册网关服务进程到系统服务, 开放服务名, 支持用户自定义, 并提供更完善的服务管理工具。
- 支持低版本网关配置文件通过客户端上传, 服务程序自动识别配置文件版本并进行转换, 为当前版本所用。
- 提供命令行安装介质。

新的全文检索网关使您在享有各种主要 RDBMS 卓越的数据处理功能的同时, 拥有 TRS 优秀的全文检索功能。

第2章 客户端使用说明

2.1 启动界面

启动 TRS 全文检索网关客户端后（运行 TrsAgentClient.exe），显示客户端控制台的界面并自动弹出连接服务器的窗口。

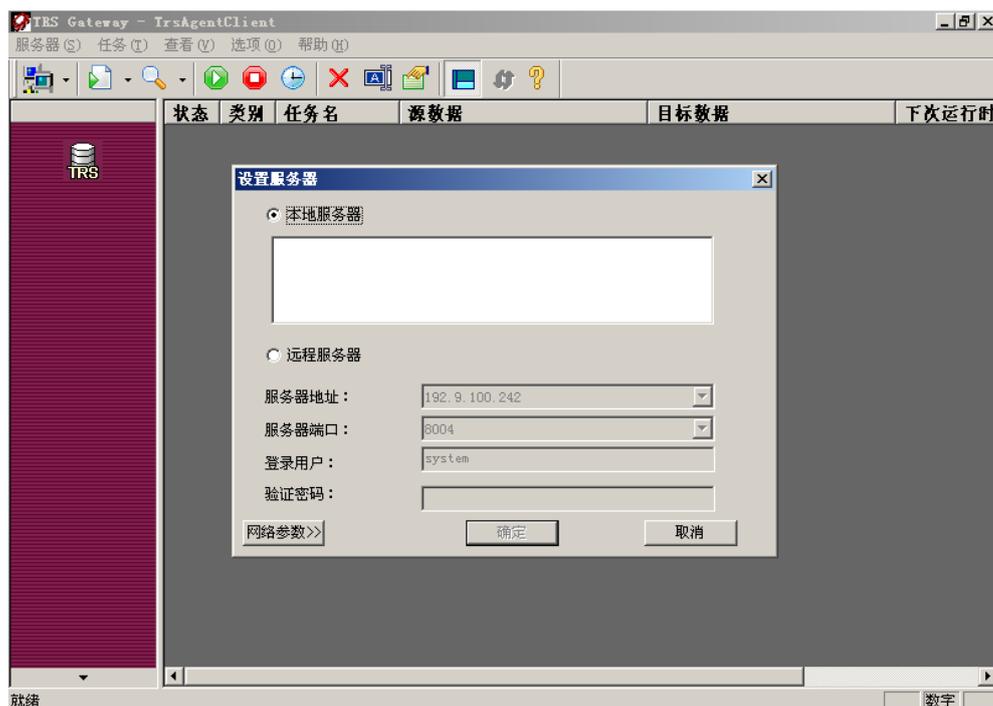


图 2.1 启动界面

TRS 全文检索网关客户端控制台的界面由以下几部分组成：

1) 菜单

- 服务器菜单

服务器菜单包括连接服务器、断开服务器、刷新服务器以及（控制台）退出四个功能，其中成功连接上后台服务器是客户端进行后续操作的基础。

- 任务菜单

任务菜单包括以下功能：新建、执行、中断执行、删除、重命名和属性。其中新建功能，在任务组为 RDBMS 时支持四种任务的创建：RDBMS=>TRS、RDBMS=>TRS 格式文件、TRS 格式文件=>TRS、TRS 优化。执行任务包括：立即执行和定时执行。

- 查看菜单

查看菜单包括以下几种功能：查看（任务）和显示执行信息。

- 选项菜单

选项菜单功能包括加载扩展功能插件、快速创建增量更新任务、上传单一任务以及日志设置。

- 帮助菜单

帮助菜单包括：目录和索引、关于本产品。

2) 工具条

工具条上包含了服务器设置、新建、查看任务、执行任务、中断任务执行、定时、删除、重命名、属性、显示执行信息窗口、刷新服务器以及目录和索引等按钮。其中服务器设置包括了连接、断开操作，新建按钮可以打开下拉菜单提供四种任务的创建。查看按钮可以打开下拉菜单提供五种选项。

3) 窗口

- 左上部：任务组窗口

任务组窗口反映的是客户端控制台当前所连接的后台网关服务支持的有效的数据库组件，其中在最多的情况下，任务组窗口将包含六个组件：ORACLE, SQL SERVER, DB2, SYBASE, MySQL 和 TRIS，未连接状态时只显示 TRIS 图标。任何时候有且只有一个任务组为选中状态，选中任务组图标为彩色显示，其他的为灰色显示。

- 右上部：任务窗口

任务窗口是一用来显示当前已创建任务信息的列表窗口，该列表在任务组为 RDBMS 时共有八列，分别为：状态、类别、任务名、源数据、目标数据、下次运行时间、次数和创建时间。

- 下部：执行信息窗口

执行信息窗口是反映了当前任务组指定任务的历次执行信息。该窗口可以显示或隐藏，点击“查看”菜单的“显示执行信息”或工具条上的“显示执行信息”按钮将控制该窗口的显示或隐藏状态。

2.2 服务器交互

客户端控制台上所有的操作，只有在连上后台网关服务时才有意义。因此启动客户端控制台后的第一件事，应该是连接后台网关服务器。通过服务器菜单或工具栏服务器连断，选择连接服务器。



图 2.2 设置服务器

客户端所在 PC 上如存在网关服务正在运行，客户端会自动发现并列出，用户只需选中目标连接的服务，确定即可。如需连接其他机器上的网关服务，则需要在远程服务器中进行设置，输入网关服务所在机器的 IP 地址，网关服务端口以及登录验证信息。其中网关默认登录用户为 system（不能修改），用于以后的扩展，验证密码为安装网关服务端介质时设定的验证密码，用户可在网关服务端 gwserver[_x64] 目录下查看 validatecode.txt 文件，验证信息记录其中，如果该文件为空，按默认值 12345678 处理。

有些情况下，服务器与客户端所处的网络环境不好，客户端对服务端的操作以及数据接收容易超时，用户可以先退出连接，选择“网络参数>>”，调整发送超时以及接收超时，重新连接，情况将会得到改善。



图 2.3 网络参数设置

“使用 SSL 连接服务器”，如果网关 Server 端目录下，TRSAGENT.INI 文件的 USE_SSL 项值为 1，那么需要选择该项来连接服务器。

客户端连接上网关服务后，会在客户端安装目录 config、log 子目录下分别创建以服务器地址命名的文件夹，存放客户端与服务器交互的临时信息，主要包括任务配置文件和运行推送日志。

用户在不需要使用客户端配置任务或查看任务执行状况时，可直接关闭客户端控制台，客户端将自

已断开与服务器之间的连接。用户也许需要连接另一个网关服务，这时需要先手动断开服务器连接，才能连接另一个网关服务。

注意：客户端控制台一次只能连接一个网关服务进行相关操作。

2.3 任务的视图

TRS 全文检索网关客户端提供按任务类型查看任务的方法，默认情况下，控制台显示所有后台网关服务器中任务列表。

在当前任务组，点击“查看”菜单的“查看任务”子项并选择不同的任务的方式，任务窗口将显示当前任务组指定任务类型的任务。

针对 RDBMS 任务组，可查看的任务类型包括：

- 所有任务
- RDBMS=>TRS 任务
- RDBMS=>TRS 格式文件任务

针对 TRS 任务组，可以查看的任务类型包括：

- TRS 格式文件=>TRS 任务
- 优化 TRS 数据库任务

2.4 创建任务

2.4.1 创建“RDBMS=>TRS”任务

所谓“RDBMS=>TRS”任务就是将保存在 RDBMS（ORACLE, SYBASE, DB2, SQL SERVER, MySQL, KBE）数据库中的数据向 TRS 数据库进行迁移的任务。

创建“RDBMS=>TRS”任务可以单击“任务”菜单的“新建”->“RDBMS=>TRS”菜单项，或单击工具条上的“创建任务”的“RDBMS=>TRS”项，又或先选中 RDBMS 图标，鼠标右键选择新建。创建“RDBMS=>TRS”任务共有八大步组成：

第一步：登录 RDBMS 数据库

创建“RDBMS=>TRS”任务的第一步是登录 RDBMS 数据库。

登录 RDBMS 数据库需要填写的参数包括：

- 服务器地址/服务名/数据库名：指安装 RDBMS 的服务器名，如果 RDBMS 是 ORACLE 的话，请填写本机 CLIENT 端设置的有效数据库名。
- 用户名：指登录 RDBMS 服务器的用户名。
- 密码：指登录 RDBMS 服务器的密码。
- 客户端字符集：指 RDBMS 客户端设置的字符集，需要正确选择。

客户端字符集的选择会影响高级设置中原数据库和目标数据库的字符设置。如果在登录过程中出现异常，请在下边的状态框中查看出错的原因。



图 2.4 登录数据库

第二步：指定数据来源于 RDBMS 数据库表或视图

创建“RDBMS=>TRS”任务的第二步是指定数据来源。

源为 RDBMS 数据库时，数据来源可以是 RDBMS 的一个表或一个视图。

- 选择源数据库

此选项只对 SQL Server、Sybase、MySQL 有意义，对其他数据库将自动屏蔽。

- 源为表

需要选择表名。表名是由两部分组成，第一部分是表空间，第二部分是表的名字。对于 Sybase、MySQL 而言，表名没有第一部分。

- 源为视图

需要选择源视图，以及指定视图所对应的主表。对于 Sybase、MySQL 而言，表名没有第一部分。主表的含义是如果针对视图中的记录选择多值合并则根据主表中的主键来合并视图中的记录。



图 2.5 指定数据库表或视图

“选择条件”栏是指用户可增加适当的约束条件，用于检索或过滤一定条件的记录，在任务执行时使用。内容为 where 子句部分，但不需加上“where”字符，如只迁移公司名称（对应字段名 COMPANY）为 TRS 的数据记录，建立任务时填写 COMPANY='TRS'即可。需要注意的是选择条件中的 SQL 语句正确性要由用户来保证。

最后状态框中显示的 SQL 语句是由系统内部处理获得，用户一般不需要关心，当用户配置任务到这一步时，如果出现异常，如获取表或视图错误，用户可拷贝出状态框中的 SQL 语句，用相同的用户名密码在 RDBMS 客户端登录数据库，执行该语句验证是否出现同样错误，避免由用户权限引起的获取异常。

第三步：指定 RDBMS 表或视图的主键，并可进行字段扩展属性设置

创建“RDBMS=>TRS”任务的第三步是指定 RDBMS 表或视图的主键，该页面在所有情况下均出现。

- 数据源为 RDBMS 表

如果指定的数据源来自 RDBMS 表时，系统将自动检测该表在创建时是否设置了主键，如果设置了主键，则在主键下拉框中缺省显示主键。

在系统提示的指定主键列表中列出了指定的源数据库表中所有可能成为主键的字段。可能成为主键的字段是指字段类型为 NUMBER, CHAR, DATE, VARCHAR2 的字段。

- 数据源为 RDBMS 视图

在系统提示的指定主键列表中列出了指定的源数据库视图中所有可能成为主键的字段。可能成为主键的字段是指字段类型为 NUMBER, CHAR, DATE, VARCHAR2 的字段。同时也要选择与

此对应的主表的主键，因为有些视图可能会把主表的字段名重新命名，所以这里要有一个对应关系存在。

需要注意的是：上一步操作，如果选择操作对象为视图时，必须为视图指定源表，也就是视图的主表，当建立增量更新任务时，记录任务更新的 TRIGGER 就创建在主表的主键上。在指定视图主键时，视图的主键应与视图对应的表列名相对应，否则会出现问题。另外如果要做全部更新的多线程任务则必须保证主键为 NUMBER 类型，任务的分段处理是以主键是数值为基础的。

如果建成多值处理的任务，主键可能不唯一，此时应该选中“合并记录集中主表（主键）重复的非主表字段内容”，其合并逻辑为，相邻的两条记录如果主键相同，检查非主表字段，如果其对应到 TRS 数据库的“DOCUMENT”或“BIT”字段，则直接取出内容进行合并。如果对应到 TRS 数据库其他字段，先检查字段的值是否相同，相同不进行合并，不同才合并。另外源字段若是格式文档（rtf, pdf, doc, xls 等），能否取出内容取决于安装的红缨枫插件（具体内容参照《TRS DataBase 全文数据库系统使用指南》）。



图 2.6 指定视图或表的主键

- 字段扩展

点击字段拷贝或字段合并按钮可以对字段内的数据进行加工处理，具体设置，请参考字段扩展属性设置。

第四步：设置目标数据库

创建“RDBMS=>TRS”任务的第四步是“设置目标数据库”。



图 2.7 设置目标数据库

需要设置的参数包含两个部分：

- 设定目标数据库

设定目标数据库可采取两种方式，

第一种是通过虚拟视图的方式，用户首先需要登录 TRS 集群服务器，程序自动获取服务器上的虚拟视图列表，选择目标虚拟视图后，点击解析按钮得到后台数据库或视图，再以统一登录或逐个登录的方式登录，确定后即完成设置。如果不进行登录设置操作，默认情况下程序以“system/manager”登录。

其中统一登录的条件是，解析得到的后台服务器具有相同的用户名密码，用户即可使用此方式一次性全部登录。如不满足统一登录的条件，用户只能逐个选择解析得到的后台服务器一一输入用户名密码进行登录。

另外，如果选择的虚拟视图经过解析得到的不是某后台服务器上的表，而是普通视图（建立在一个或多个表上，同关系型数据库中视图概念，不同于 TRS 自动分裂视图，详细定义参考《TRS DataBase 全文数据库系统使用指南》，此时设置目标数据库失败，需要重新选择符合要求的虚拟视图。



图 2.8 虚拟视图配置数据库集群

对于 Cluster2.5 以上的版本集群，其采用多级集群架构技术，支持在原有的多个集群之上再架设上层集群。这就增加了一种可能，在网关解析视图时得到后台数据库列表，其中又包含下一级集群的虚拟视图，网关在统一登录或确定时会给出相应的提示信息，这时，我们需要单独对每一个虚拟视图进行登录，网关会再解析其后台数据库列表，直到列表中不再包含未解析的虚拟视图为止。



图 2.9 完整的虚拟视图解析列表

IDLE&ORDER（通常所说的读写分离）是集群服务器提供的另一功能，IDLE 使后台服务器进入空闲状态，ORDER 使空闲的后台服务器继续服务。选择网关的 IDLE&ORDER 选项将使网关在装载数据之前会给后台集群服务器发送 IDLE 命令，暂时使集群下待装数据库处于空闲状态，等待

装载完成后，再发送 ORDER 命令，恢复其服务。网关支持向配置相同的多个集群服务器同时发送 IDLE&ORDER，通过“添加配置相同的集群服务器根节点”完成设置。IDLE、ORDER 应用场景和功能详情具备参阅《TRS 集群服务器（2x）使用手册》。

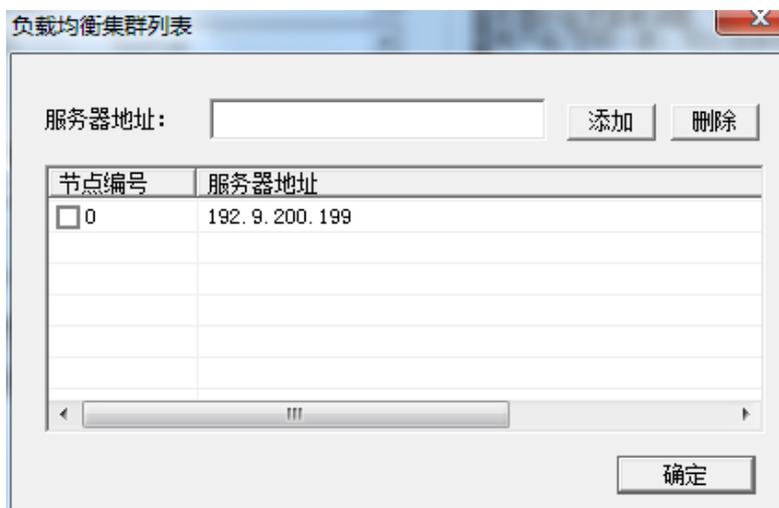


图 2.10 添加负载均衡集群服务器根节点

第二种是通过手工配置的方式，该方式下用户首先需要逐一登录后台 TRS 服务器，选择表或自动分裂视图进行增加。相对第一种方式较为繁琐，但优点在于可以动态创建 TRS 数据库表或自动分裂视图。

另外 TRS 集群服务器同一节点（对应节点编号）下可能存在多个机器（对应机器编号），可通过设置是否作为备份库，让同一节点下的多台服务器相互备份，相互备份的机器数据将保存一致。

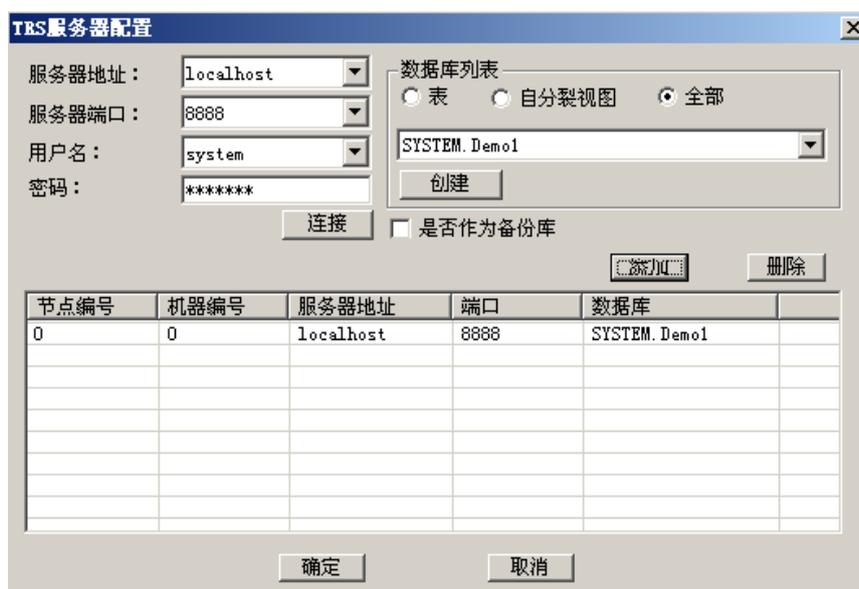


图 2.11 多服务器配置

通过手工配置的目标数据库将会自动生成一个虚拟视图，命名为 gw_auto_view，该虚拟视图并不真实存在，只是套用集群的概念，方便网关管理，任务列表中如想查看 gw_auto_view 的真实组成，请点击任务右键属性常规选项卡中查看。

- 数据同步

TRS 全文检索网关提供对 TRS 数据库集群的支持。网关可选择以哪个字段为基础进行数据同

步，默认情况下使用网关配置的主键，用户可根据应用修改。网关提供了三种方案配置，以便用户根据系统的数据特点选择合适的方案达到数据均衡的目标。

第一种方案是主键取模，即对主键进行取模运算，模数为数据库集群的节点数。采用这种方案配置の場合为，记录的主键字段是整数数值类型，且数值按照取模运算比较均匀的分布。

第二种方案是先哈希再取模，即先对主键值进行哈希运算，再根据哈希结果进行取模运算。采用这种方案配置の場合为，记录的主键字段是字符类型（数值类型的字段值在 TRS 全文检索网关内部也是按照字符类型存储），且字符串经过哈希函数运算转化的数值结果应该比较随机。数据的均衡，一方面需要字符串比较随机，另一方面需要哈希函数能够充分利用字符串的所有字母，生成的结果均匀。

第三种方案是正则匹配，即为 TRS 数据库集群的节点设置正则表达式，能够匹配正则表达式的记录，同步到对应的 TRS 数据库节点服务器上。具体的逻辑是，记录主键依次与正则表达式组合比较。如果符合该正则表达式，则将该记录同步到对应的 TRS 数据库集群节点中。该方案可应用于主键字段为任意类型，但用户需要充分考虑记录主键的特点、规则，配置不同的正则表达式组合，把数据均衡分布到 TRS 数据库集群中。正则设置步骤如下：

- 1、用户分析记录主键值的特征，根据目标数据库设置中的节点数 N，将主键记录的值分为 N-1 堆
- 2、逐一书写正则表达式进行增加，其中用户可选择堆中具有代表性的主键记录写入“测试数据”进行测试，通过后再增加，以此保证表达式的正确

说明：用户指定正则表达式可能无法覆盖全部数据，因此，在 N 个节点的 TRS 数据库集群中，用户只需也只能指定 N-1 个正则表达式，最后一个节点存放不满足所有正则表达式的数据。

关于正则表达式的使用参见“[附录 5 正则表达式](#)”。

正则表达式添加完成之后，用户也可以对选定节点进行修改和删除操作，如下图。

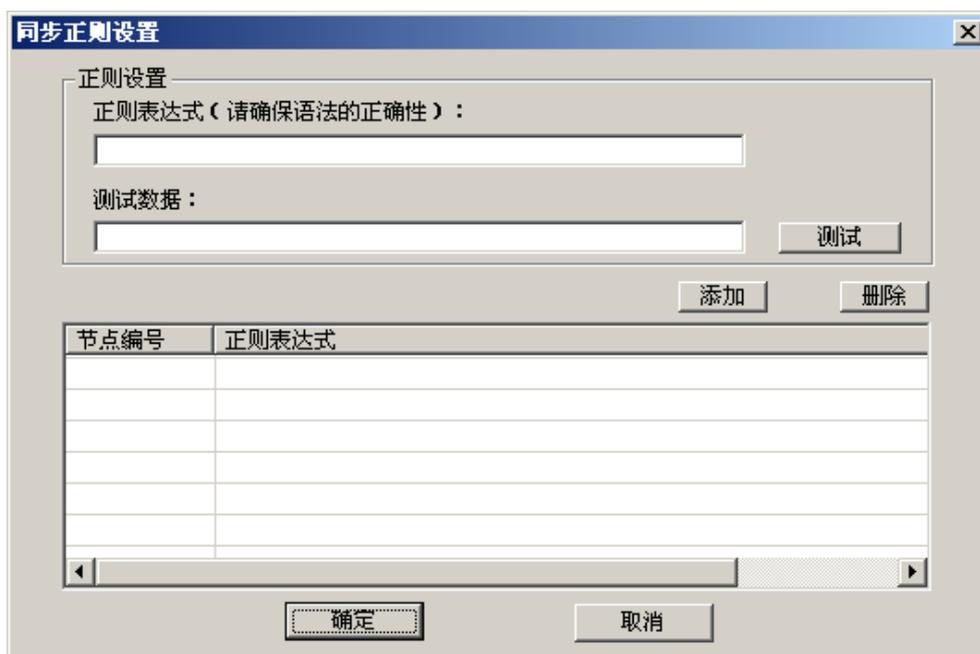


图 2.12 同步正则设置

注意：该方案存在限制，只适用于通过手工配置目标数据库的情况下使用。另外用户需要确保正则表达式的语法正确，避免 Linux 环境服务宕机。

第五步：设置目标数据库参数

控制文件，即根据控制文件装库，该文件路径相对于网关服务端所在机器，用户设置后可选择“校验”按钮检查文件是否存在，

注意：在 Unix 系统下，目录也是一种特殊文件，校验后会返回相应的文件类型信息。

格式文件临时目录即从 RDBMS 导出的 TRS 装库文件存放的路径，该路径也是相对于网关服务端所在机器，缺省情况为网关安装目录下的 temp 文件夹，不建议更改。在网关目录下空间不足时，用户可填写其他路径并使用“新建”按钮创建。

注意：新建 TRS 格式文件目录需要输入全路径，且只能逐级创建，控制文件、格式文件临时目录都是相对网关服务端所在机器，与网关客户端无关。

生成格式文件最大记录数，即每个 TRS 格式文件内可以存放的记录数目，缺省为 5000，用户也可以根据需要重新设定。

批量更新装库，该功能限于 TRS Server 6.50.4005 以后版本。选中此选项网关将自动创建批量更新装库所需要的控制文件，存储路径为网关安装目录 ctrl 子目录下。同时该选项的选择将屏蔽全部更新时选择删除再追加的入库方式。有关批量更新装库的应用场景、功能说明请参阅《TRS DataBase 全文数据库系统使用指南》

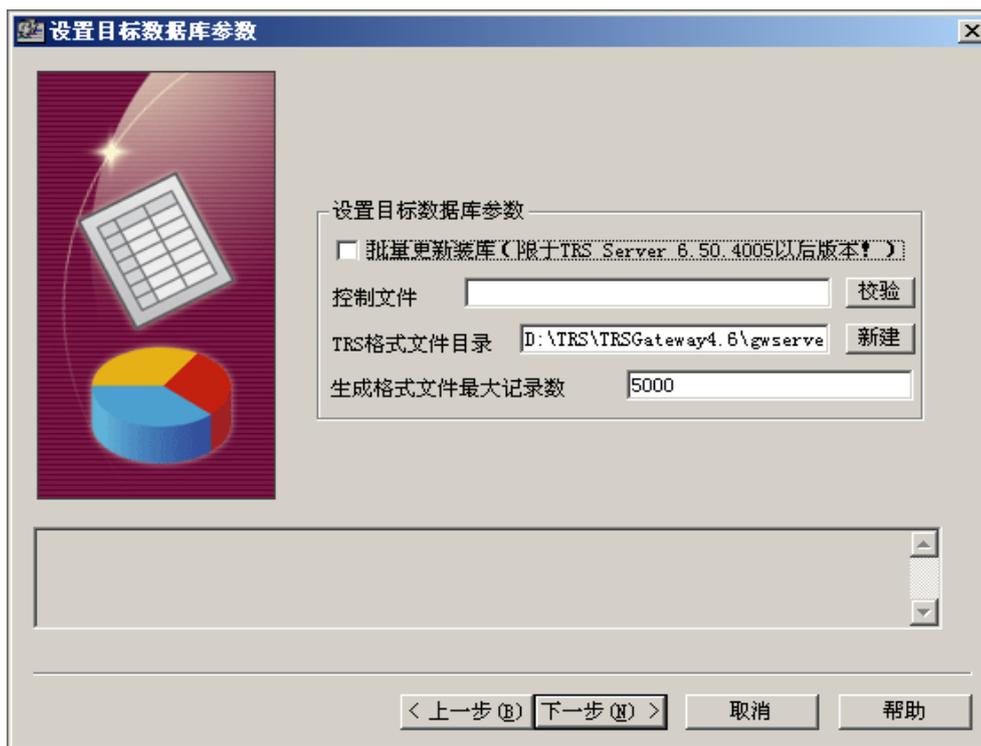


图 2.13 设置目标数据库参数

第六步：字段匹配

创建“RDBMS=>TRS”任务的第六步是“字段匹配”。

字段匹配页面由三部分组成：

- RDBMS 源数据库表字段

源数据库字段显示的是指定的 RDBMS 源数据库表/视图的所有字段以及它们的类型、长度、主键和其他属性。如果以上字段中包含了大对象字段，则用户还可以设置大对象的属性和类型。

● TRS 目标数据库字段

目标数据库字段显示的是指定的 TRS 目标数据库的所有字段以及它们的类型和长度。当字段类型不为字符型 (CHAR) 时，所有的长度均为 0。

● 大对象的设置

大对象的处理缺省采用不处理的方式，可以指定大字段的扩展名唯一、由其他字段指定、内容分析或者文件路径等。

- ◇ 当扩展名唯一的时候，它可以是以下几种类别：如果映射到的目标为 BIT 则可以为以下几种：

"bit","txt","doc","ws","wp","xls","wps","s2","aud","img","vid","avi","mpg","wav","mid","cda","ps","gif","tif","pcx","bmp","jpg","html","pdf","ppt","rtf","xml","mp3",但是数据在库中不建立索引，即不能够查询字段内存储的内容；如果映射到的目标字段为 DOCUMENT，则可以为以下几种："txt","html","rtf","pdf","doc","xls","xml",这些字段允许建立索引，如果建立了索引可以查询字段内的内容，（对于 rtf, pdf, doc, xls 类型文档，前提是安装了红缨枫插件，具体内容参照《TRS DataBase 全文数据库系统使用指南》）。

- ◇ 如果是类型由其他字段指定，意味着大字段的扩展名由相同行里面其他字段内容来指定，指定扩展名的那个字段里的内容是同上述的一样。
- ◇ 如果选择是内容分析，将对大对象中存储的内容进行文件格式分析，抽取内容。支持的格式文件类型由红缨枫插件决定，具体内容参照《TRS DataBase 全文数据库系统使用指南》。
- ◇ 如果是 varchar, char, 或者 nvarchar, nchar 等字符类型导入到 DOCUMENT 字段时，可以指定字段内容是文件路径或者非文件路径，如果指定了文件路径，则字段中存放的内容必须是[@字符]+文件路径（注意如果是本地文件路径，则相对网关服务所在机器，如果是 FTP 文件路径，则需要在高级中设置 FTP 连接信息，总之需要确保路径可访问）。如果指定了非文件路径（缺省情况），则会把字段内容一致的导入到 DOCUMENT 字段中。注意这种情况只有在用户建立了 varchar, char 等字段与 DOCUMENT 字段的映射以后才可以单击第五列出现设置的下拉框。
- ◇ 文件路径与内容分析的组合，该选项结合以上两种设置，获取路径的文件并进行内容分析，最终结果入库。

● 字段的匹配

源数据字段与目标数据字段的匹配是通过鼠标的拖放进行的。您可以选中源数据列表的一行，按住鼠标左键，然后拖放到目标数据库列表中。如果选中的源数据列表的指定字段的类型能够和目标数据库列表的鼠标所在行的字段类型进行匹配，则光标将显示为 ，否则为 。如果您希望该源指定字段和某一支持匹配的目标字段进行匹配，松开鼠标左键即可。

另外，通过字段拷贝和字段合并得到的扩展字段在字段匹配时不进行字段类型检查，可匹配任意 TRS 数据类型，用户需要进行适当检查。

功能按钮：

● 按目标匹配

所谓“按目标匹配”就是系统自动根据源数据字段名和目标数据字段名进行匹配，也就是说如

果某一源字段名与某一目标字段名相同且上述两个字段的类型与长度均符合匹配条件，则自动匹配。

- 删除匹配

在源或目标字段中选中一行已经进行匹配的字段，点击“删除匹配”将删除指定字段以及它的对应字段的匹配信息。

- 还原

取消所有已经设置的字段匹配信息，并将源字段信息恢复到初始状态。



图 2.14 字段匹配

网关在默认情况下的字段匹配是进行类型检查的，如日期类型不能对应到 NUMBER 类型，但有些情况下需要突破这种限制，此时勾选不进行字段类型检查，即可进行字段对应。

报告出错信息的对话框,点击“下一步”时，如果判断当前的匹配信息不正确，报告出错原因。

第七步：设置更新方式

创建“RDBMS=>TRS”任务的第七步是“设置更新方式”。

RDBMS=>TRS 支持的更新方式包括两种：

- 全部更新

全部更新可以进行参数设置，也有三种选择：数据追加、临时表清空再追加和直接清空再追加。

数据追加是直接将所有源数据库表中符合条件的数据迁移到目标数据库中，对目标数据库原有数据不做任何操作。

临时表清空再追加，通过一个临时表进行中间转换，过程对用户透明，最后的结果相当于先清空目标数据库，再将所有源数据库表中符合条件的数据迁移到目标数据库中。需要注意的是清空再追加建立临时表是以原数据库名为基础的，格式为：原数据库名+ “__temp”，对于原数据库名称过长，加上后缀后更长进而超过 TRS Server 数据库命名长度时，会报出错误，此时需要用户创建一个名称稍短的数据库，避免此种错误。

直接清空再追加，顾名思义，迁移数据前先将目标数据库清空，再将符合条件的数据迁移到目标数据库中。

- 增量更新

就是自该任务创建成功时起，通过记录 RDBMS 源数据库表或视图记录的更新状态，在执行任务时，将所有更新了源数据库表中符合条件的数据迁移到目标数据库中。增量更新的时候会创建一个临时表和相关的 TRIGGER，临时表的名字缺省采用：任务的名字+ “\$_TEMP”，在某些类型的数据库中表名字的长度有限制，而任务为了容易记忆名字又比较长，可以在这里更改，则临时表的名字采用：自定义临时表+ “\$_TEMP”。

根据临时表表名构造特点，用户在建立任务或自定义临时表表名时尽量少用特殊字符，代之使用下划线作为连接或分隔的符号，以避免相关结构的名称的语法不符合数据库规定，导致建立增量更新任务失败。

点击“高级”按钮可以进行多线程和字符转码的设置，具体内容，请参考文档中的“高级设置”部分。



图 2.15 设置更新方式

第八步：定时

创建“RDBMS => TRS”任务的最后一步是定时。

有关任务定时的配置请参见“[2.7.3 任务的定时](#)”章节中的内容。

2.4.2 创建“RDBMS=>TRS 格式文件”任务

所谓“RDBMS=>TRS 格式文件”任务就是将保存在 RDBMS (ORACLE, SYBASE, DB2, SQL SERVER, MySQL) 数据库中的数据向 TRS 格式文件进行迁移的任务。

创建“RDBMS=>TRS 格式文件”任务可以单击“任务”菜单的“新建RDBMS=>TRS 格式文件”菜单项，或单击工具条上的“创建任务”的“RDBMS=>TRS 格式文件”项，又或先选中 RDBMS 图标，鼠标右键选择新建。创建“RDBMS=>TRS 格式文件”任务与“RDBMS=>TRS”任务基本类似但只需要七步：

第一步：登录 RDBMS 数据库

创建“RDBMS=>TRS 格式文件”任务的第一步是登录 RDBMS 数据库。

登录 RDBMS 数据库需要填写的参数包括：

- 服务器地址/服务名/数据库名：指安装 RDBMS 的服务器名，如果 RDBMS 是 ORACLE 的话，请填写本机 CLIENT 端设置的有效数据库名。
- 用户名：指登录 RDBMS 服务器的用户名。
- 密码：指登录 RDBMS 服务器的密码。
- 客户端字符集：指 RDBMS 客户端设置的字符集，需要正确选择。

客户端字符集的选择会影响高级设置中原数据库和目标数据库的字符设置。如果在登录过程中出现异常，请在下面的状态框中查看出错的原因。

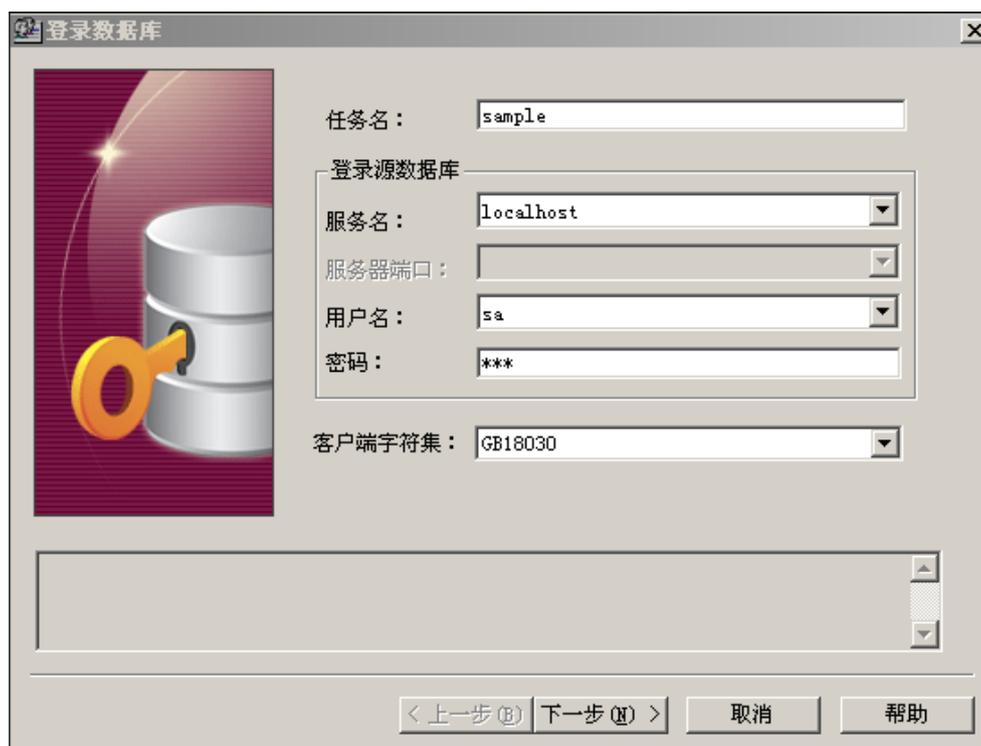


图 2.16 登录数据库

第二步：指定数据来源于 RDBMS 数据库表或视图

创建“RDBMS=>TRS”任务的第二步是指定数据来源。

源为 RDBMS 数据库时，数据来源可以是 RDBMS 的一个表或一个视图。

- 选择源数据库

此选项只对 SQL Server、Sybase、MySQL 有意义，对其他数据库将自动屏蔽。

- 源为表

需要选择表名。表名是由两部分组成，第一部分是表空间，第二部分是表的名字。对于 Sybase、MySQL 而言，表名没有第一部分。

- 源为视图

需要选择源视图，以及指定视图所对应的主表。对于 Sybase、MySQL 而言，表名没有第一部分。主表的含义是如果针对视图中的记录选择多值合并则根据主表中的主键来合并视图中的记录。



图 2.17 指定数据库表或视图

“选择条件”栏是指用户可增加适当的约束条件，用于检索或过滤一定条件的记录，在任务执行时使用。内容为 where 子句部分，但不需加上“where”字符，如只迁移公司名称（对应字段名 COMPANY）为 TRS 的数据记录，建立任务时填写 COMPANY=‘TRS’即可。需要注意的是选择条件中的 SQL 语句正确性要由用户来保证。

最后状态框中显示的 SQL 语句是由系统内部处理获得，用户一般不需要关心，当用户配置任务到达这一步时，如果出现异常，如获取表或视图错误，用户可拷贝出状态框中的 SQL 语句，用相同的用户名密码在 RDBMS 客户端登录数据库，执行该语句验证是否出现同样错误，避免是由用户权限引起的获取异常。

第三步：指定 RDBMS 表或视图的主键，并可进行字段扩展属性设置

创建“RDBMS=>TRS 格式文件”任务的第三步是指定 RDBMS 表或视图的主键。

- 数据源为 RDBMS 表

如果指定的数据源来自 RDBMS 表时，系统将自动检测该表在创建时是否设置了主键，如果设置了主键，则在主键下拉框中缺省显示主键。

在系统提示的指定主键列表中列出了指定的源数据库表中所有可能成为主键的字段。可能成为主键的字段是指字段类型为 NUMBER, CHAR, DATE, VARCHAR2 的字段。

- 数据源为 RDBMS 视图

在系统提示的指定主键列表中列出了指定的源数据库视图中所有可能成为主键的字段。可能成为主键的字段是指字段类型为 NUMBER, CHAR, DATE, VARCHAR2 的字段。同时也要选择与此对应的主表的主键，因为有些视图可能会把主表的字段名重新命名，所以这里要有一个对应关系存在。

需要注意的是：上一步操作，选择操作对象为视图时，必须为视图指定了源表，也就是视图的主表，当建立增量更新任务时，记录任务更新的 TRIGGER 就创建在主表的主键上。指定视图主键时，视图的主键应与视图对应的表列名相对应，否则会出现问题。另外如果要做全部更新的多线程任务则必须保证主键为 NUMBER 类型，任务的分段处理是以主键是数值为基础的。

如果建成多值处理的任务，主键可能不唯一，此时应该选中“合并记录集中主表（主键）重复的非主表字段内容”，其合并逻辑为，相邻的两条记录如果主键相同，检查非主表字段，如果其对应到 TRS 数据库的“DOCUMENT”或“BIT”字段，则直接取出内容进行合并。如果对应到 TRS 数据库其他字段，先检查字段的值是否相同，相同不进行合并，不同才合并。另外源字段若是格式文档(rtf, pdf, doc, xls 等)，能否取出内容取决于安装的红缨枫插件(具体内容参照《TRS DataBase 全文数据库系统使用指南》)。



指定源数据库主键

指定视图主键
视图主键：
与视图对应的表的列名：
 合并记录集中主表(主键)重复的非主表字段内容

指定表主键
主键：
emp_id

字段扩展
字段拷贝 字段合并

```
SELECT COLUMN_NAME, DATA_TYPE, CHARACTER_MAXIMUM_LENGTH FROM
INFORMATION_SCHEMA.COLUMNS WHERE TABLE_NAME = 'EMPLOYEE' ORDER BY
COLUMN_NAME
SELECT COLUMN_NAME FROM INFORMATION_SCHEMA.KEY_COLUMN_USAGE WHERE
```

< 上一步(B) 下一步(N) > 取消 帮助

图 2.18 指定数据库表或视图主键

- 字段扩展

点击字段拷贝或字段合并按钮可以对字段内数据的进行加工处理，具体设置，请参考字段扩展属性设置。

第四步：选择字段并指定存储目录

创建“RDBMS=>TRIS 格式文件”任务的第四步是“选择字段并指定存储目录”。

需要设置的参数包含以下三个部分：

- 字段列表

选择字段：选择要将哪些 RDBMS 的字段和扩展字段导入 TRIS 格式文件中。其中可以修改“对应”列和目标“类型”列，“对应”列指明在生成的 TRIS 格式文件中显示的字段名，目标“类型”列指明对应 TRIS 表或自动分裂视图中字段的类型。

注意一般只对字段扩展得到的新字段进行修改。对于目标类型为“DOCUMENT”或“BIT”的字段，生成 TRIS 格式文件时其字段的值存到外部文件。

- 格式文件目录

指定用于存放 TRIS 格式文件的目录，该路径相对网关服务所在的机器，默认为网关服务端安装目录下的 temp 文件夹，不建议更改。在网关目录下空间不足时，用户可填写其他路径并使用“新建”按钮创建。

- 生成格式文件的最大记录数目

TRIS 格式文件存放的最多的记录条目，缺省为 5000 条，用户可以根据需要进行调整。

- 是否生成 ok 文件

选中该选项，每生成一份 TRIS 格式文件时都会生成一个相应的.ok 文件，如生成 TRIS 格式文件文件名为 1196238732_3136_1.tris，相应的.ok 文件名为 1196238732_3136_1.tris.ok。

如未选中该选项，文件则不生成.ok 文件。

- 数据同步目标节点数

用于设置数据同步节点个数，缺省值为 1。



图 2.19 选择字段并指定存储目录

第五步：设置数据同步

TRS 全文检索网关提供对 TRS 数据库集群的支持。网关可选择以哪个字段为基础进行数据同步，默认情况下使用网关配置的主键，用户可根据应用修改。网关提供了三种方案配置，以使用户根据系统的数据特点选择合适的方案达到数据均衡的目标。



图 2.20 数据同步方式

第一种方案是主键取模，即对主键进行取模运算，模数为数据同步目标的节点数。采用这种方案配置の場合为，记录的主键字段是整数数值类型，且数值按照取模运算比较均匀的分布。

第二种方案是先哈希再取模，即先对主键值进行哈希运算，再根据哈希结果进行取模运算。采用这种方案配置の場合为，记录的主键字段是字符类型（数值类型的字段值在 TRS 全文检索网

关内部也是按照字符类型存储), 且字符串经过哈希函数运算转化的数值结果应该比较随机。数据的均衡, 一方面需要字符串比较随机, 另一方面需要哈希函数能够充分利用字符串的所有字母, 生成的结果均匀。

第三种方案是正则匹配, 即为 TRS 数据库集群的节点设置正则表达式, 能够匹配正则表达式的记录, 同步到对应的 TRS 数据库节点服务器上。具体的逻辑是, 记录主键依次与正则表达式组合比较。如果符合该正则表达式, 则将该记录同步到对应的 TRS 数据库集群节点中。该方案可应用于主键字段为任意类型, 但用户需要充分考虑记录主键的特点、规则, 配置不同的正则表达式组合, 把数据均衡分布到 TRS 数据库集群中。正则设置步骤如下:

- 1、用户分析记录主键值的特征, 根据目标数据库设置中的节点数 N, 将主键记录的值分为 N-1 堆
- 2、逐一书写正则表达式进行增加, 其中用户可选择堆中具有代表性的主键记录写入“测试数据”进行测试, 通过后再增加, 以此保证表达式的正确

说明: 用户指定正则表达式可能无法覆盖全部数据, 因此, 在 N 个节点的 TRS 数据库集群中, 用户只需也只能指定 N-1 个正则表达式, 最后一个节点存放不满足所有正则表达式的数据。

关于正则表达式的使用参见“[附录 5 正则表达式](#)”。

正则表达式添加完成之后, 用户也可以对选定节点进行修改和撤销操作, 如下图。

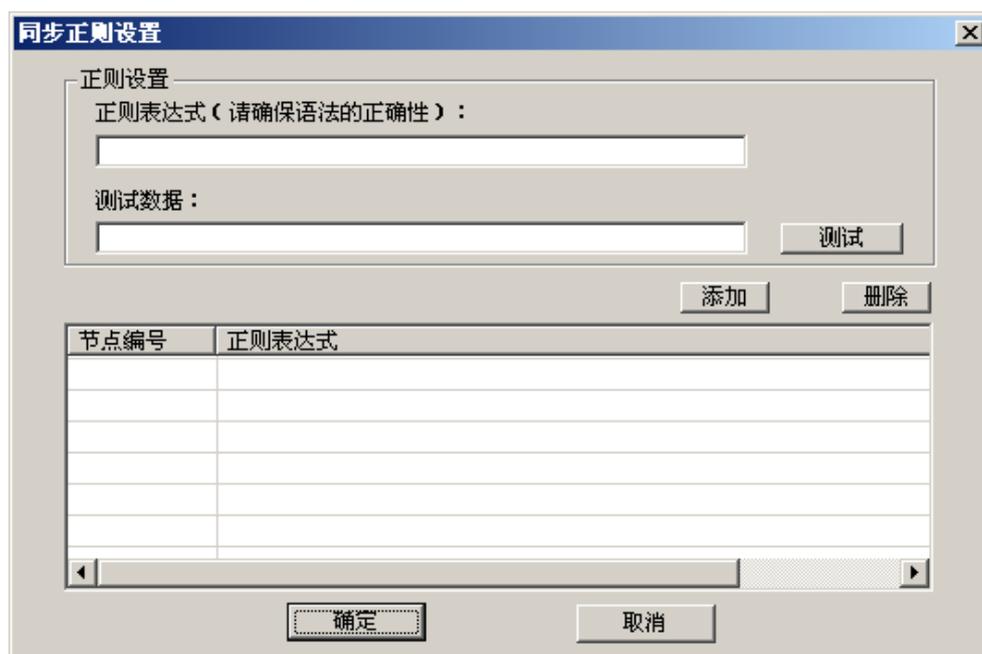


图 2.21 同步正则设置

第六步: 设置更新方式

创建“RDBMS=>TRS 格式文件”任务的第五步是“设置更新方式”。

RDBMS=>TRS 格式文件支持的更新方式包括两种:

- 全部更新

就是在每次执行任务时, 直接将所有源数据库表中符合条件的数据生成 TRS 格式文件。

- 增量更新

就是自该任务创建成功时起，通过记录 RDBMS 源数据库表或视图记录的更新状态，在执行任务时，将所有更新了的源数据库表中符合条件的数据生成 TRS 格式文件。

增量更新的时候会创建一个临时表和相关的 TRIGGER，临时表的名字缺省采用：任务的名字 + “\$_TEMP”，在某些类型的数据库中表名字的长度有限制，而任务为了容易记忆名字又比较长，可以在这里更改，则临时表的名字采用：自定义临时表+ “\$_TEMP”。

根据临时表表名构造特点，用户在建立任务或自定义临时表表名时尽量少用特殊字符，代之使用下划线作为连接或分隔的符号，以避免相关结构的名称的语法不符合数据库规定，导致建立增量更新任务失败。

此外还可以点击高级，进行多线程和字符转码的设置，具体内容请参考高级设置。

注意：有一点与 RDBMS=>TRS 任务不同的是，RDBMS=>TRS 文件的增量更新任务对于表或视图只有增加记录逻辑，网关运行结果正确。对于还包含删除或更新记录，存在问题。建议用户不要建立此类任务。



图 2.22 设置更新方式

第七步：定时

创建“RDBMS =>TRS 格式文件”任务的最后一步是定时。

有关任务定时的配置请参见“[2.7.3 任务的定时](#)”章节中的内容。

2.4.3 创建“TRS 格式文件=>TRS”任务

所谓“TRS 格式文件=>TRS”任务就是将保存在 TRS 格式文件或保存在一个指定目录下的所有 TRS

格式文件数据向 TRS 数据库进行迁移的任务。

创建“TRS 格式=>TRS”任务可以单击“任务”菜单的“新建[TRS 格式文件=>TRS]”菜单项，或单击工具条上的“创建任务”的“TRS 格式文件=>TRS”项，又或先选中 TRS 图标，鼠标右键选择新建。创建“TRS 格式文件=>TRS”任务共有三步组成：

第一步：TRS 服务器设置

创建“TRS 格式文件=>TRS”的第一步是 TRS 服务器设置，包括登录 TRS 数据库和添加备份库。

登录 TRS 数据库需要填写的参数包括：

- 服务器地址：TRS 服务器的 IP 地址。
- 服务器端口：TRS 服务器的端口号，缺省为 8888。
- 用户名：登录 TRS 服务器的用户名。
- 密码：登录 TRS 服务器的密码。

添加备份库：用户可以重新登录另一 TRS 数据库，也可在同一 TRS 数据库中选择表或自动分裂视图进行添加，这些表或自动分裂视图被视为主备关系。



图 2.23 登录数据库

第二步：设定装载参数

创建“TRS 格式文件=>TRS”的第二步是设定装载参数。

- 控制文件

如装库需要，请填写控制文件路径，该路径对网关服务所在机器，可使用“校验”按钮检查书

写路径文件是否存在。

- TRS 格式文件目录

手动填写格式文件目录，任务执行的时候会从指定的目录搜索所有的以“.trs”为扩展名的文件，然后把所有文件装入指定的库。该路径相对网关服务所在机器，默认为网关服务端安装目录下temp文件夹，用户可以通过“新建”按钮创建一个新文件夹并进行监控。

- 入库后格式文件处理

网关装载完 TRS 格式文件后，默认的操作是删除。用户可以选择其他方式，如直接备份装库文件到某一目录，还可以选择在备份到某一目录后按日期目录组织装库文件，日期目录格式为YYYYMMDD，例如 20110410。

- 是否检测 ok 文件

选中该选项，网关在装载 TRS 格式文件时会检测该 TRS 格式文件是否存在相应.ok 文件，如果存在则正常装载，如果不存在，跳过该文件继续下一个 TRS 格式文件检测装载。

如未选中该选项，网关直接装载 TRS 格式文件。

- 入库前先清空目标库

正常的 TRS 格式文件入库操作是将记录直接追加到目标数据库，用户也可以选择入库前先清空目标数据库再追加记录。其中清空方式有两种：直接清空和临时表清空。直接清空就是迁移数据前先将目标数据库清空；临时表清空就是通过一个临时表进行中间转换，先把 trs 文件导入目标库的临时表，再用这个临时表替换原来的目标库的表。用户需要注意在使用临时表清空时，确保目标数据库的表名不要过长，以防止创建临时表失败。

- 数据库字符集

指定要装入的数据库的字符集，必须指出的是，生成的 TRS 格式文件必须已经是相同的字符集。即如果要用 UTF-8 入库，则生成的 TRS 格式文件必须已经是 UTF-8，否则入库不正确。

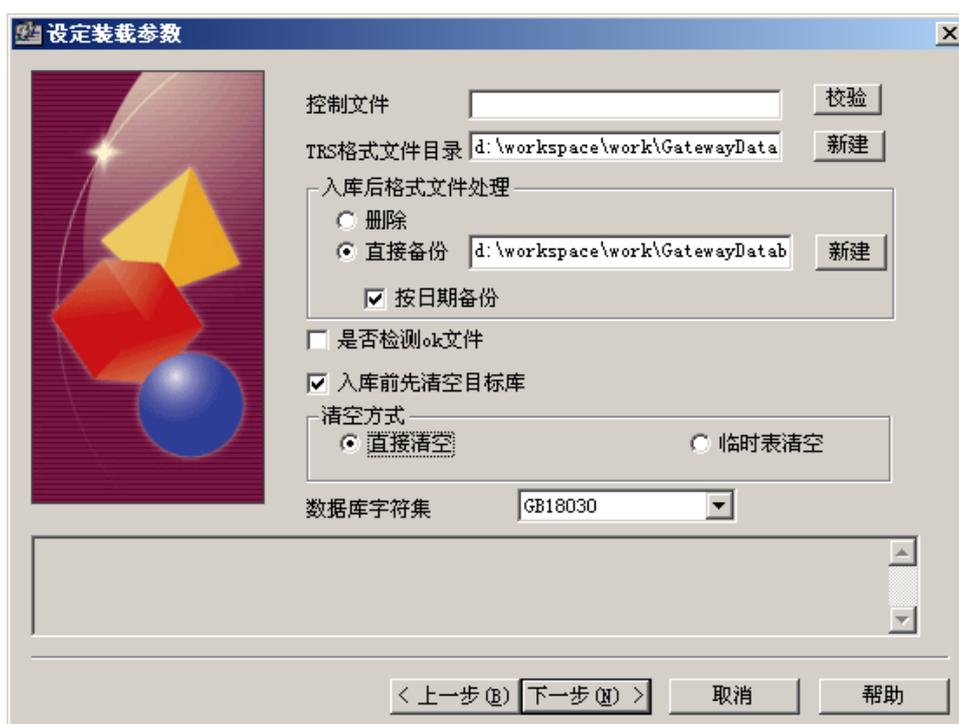


图 2.24 设定装载参数

第三步：定时

创建“TRIS 格式文件=>TRIS”任务的最后一步是定时。

有关任务定时的配置请参见“[2.7.3 任务的定时](#)”章节中的内容。

2.4.4 创建“优化 TRIS 数据库”任务

所谓“优化 TRIS 数据库”任务就是对指定的 TRIS 数据库进行优化的任务。

创建“优化 TRIS 数据库”任务可以单击“任务”菜单的“新建|优化 TRIS 数据库”菜单项，或单击工具条上的“创建任务”的“优化 TRIS 数据库”项，又或先选中 TRIS 图标，鼠标右键选择新建。创建“优化 TRIS 数据库”任务共有两步组成：

第一步：TRIS 服务器设置：

创建“TRIS 格式文件=>TRIS”的第一步是 TRIS 服务器设置，包括登录 TRIS 数据库和添加优化库。

登录 TRIS 数据库需要填写的参数包括：

- 服务器地址：TRIS 服务器的 IP 地址。
- 服务器端口：TRIS 服务器的端口号，缺省为 8888。
- 用户名：登录 TRIS 服务器的用户名。
- 密码：登录 TRIS 服务器的密码。

添加优化库：用户可以重新登录另一 TRIS 数据库，也可在同一 TRIS 数据库中选择表或自动分裂视图进行添加，这些表或自动分裂视图将在任务执行时顺序被优化。



图 2.25 登录数据库

第二步：定时

有关任务定时的配置请参见“[2.7.3 任务的定时](#)”章节中的内容。

2.5 字段扩展

字段扩展功能为 TRS 全文检索网关增加了对数据加工处理的能力，包括字段拷贝和字段合并两大项功能，其原理都是拷贝源字段，在此基础上进行加工处理。

对于字段扩展列表，判断某个字段属于拷贝还是合并的依据是：字段名称前面的两个字符。如果字段名的前两个字符是“+C”，则该字段是拷贝字段；如果是“+M”，则该字段是合并字段。

2.5.1 字段拷贝

字段拷贝根据源数据库的字段属性，细分为六类操作：

- 一、 不做处理，该操作实现对源数据库字段内容的直接拷贝，值存入新字段。应用举例，一个 Oracle 的 blob 字段，既要存储到 Document 字段又要存储到 BIT 字段；
- 二、 日期类型，此功能主要针对源数据库字段为日期类型（date,time,datetime），需要获得日期字段中的 year、month 或 day 的值。如数据库字段内容为“2007-10-11 00:00:00”，定义一个抽取 year 的字段拷贝，则此字段的内容为 2007，如果定义抽取 month 的字段拷贝，则此字段的内容为 10，如果定义抽取 day 的字段拷贝，则此字段的内容为 11；
- 三、 XML 字段，该操作实现对源数据库字段内容的拷贝，如字段中的内容是 xml 格式，可根据应用，指定抽取置标，得到该置标内的信息，值存入新字段。抽取的格式为<tag></tag>形式的封闭区域。举例说明：如字段内容为<student><name>张三</name><sex>男</sex></student>可抽取出 student 置标，值为：<name>张三</name><sex>男</sex>；抽取

name 置标, 值为: 张三; 抽取 sex 置标, 值为: 男。

四、 字符替换, 该操作实现对源数据库字段内容拷贝, 然后用指定的字符替换字符串中的某个字符, 作为拷贝的新字段的值。举例说明: 如需将字段内容“2007-10-11”变成“2007/10/11”, 只需使用“/”替换“-”。

五、 网关排重扩展, 为网关自己定义的一种排重操作。可根据“主键”+“内容”进行, 也可只根据“内容”进行, 其中“内容”可以是单个字段的值(字段拷贝中), 也可能是多个字段值合并的结果(字段合并中)。扩展出的字段存放根据“内容”计算出的唯一值, 计算接口开放, 默认情况下使用广义 MD5, 也可以使用 CKM 的指纹 MD5, 或其他算法。处理目标各种取值的含义如下:

- compare_pk_load: 代表先比较主键, 重复则结果入库, 如果主键不重复再比较内容, 重复则结果入库;
- compare_pk_not_load: 代表先比较主键, 重复则结果不入库, 如果主键不重复再比较内容, 重复则结果不入库;
- not_compare_pk_load: 代表不比较主键, 只比较内容, 重复则结果入库;
- not_compare_pk_not_load: 代表不比较主键, 只比较内容, 重复则结果不入库。

注意: 这里提到的“主键”是指任务配置过程中设置的主键。网关排重功能只能设置一次。

六、 扩展功能, 该操作实现对源数据库字段内容按照用户定义的处理方式进行加工, 标准产品集成了 CKM 的大部分功能, 包括: 文本分类、相似性检索、文本摘要、文本分词、常识校对、规则分类、实体识别、案件名识别、数字识别、房屋信息识别、简历信息抽取(需二次开发)、拼音检索、短语检索。相关功能的说明请参阅《TRS CKM 功能手册》。

注意: 扩展功能需成功加载扩展功能组件后方可使用。



图 2.26 字段拷贝

2.5.2 字段合并

字段合并操作两个或两个以上字段,按照用户指定的先后顺序,将其内容以字符串的形式进行相加,其中用于合并的字段可以是源数据库的字段,也可以是通过字段合并生成的新字段以及通过字段拷贝生成的新字段。“序号”列用于设置合并字段的先后顺序。另外“类型”和“长度”列的值对于新生成的字段(即扩展字段)没有影响,扩展字段默认显示的“类型”和“长度”是用户指定序号为1的合并字段的类型和长度。

字段合并也可选择处理目标,范围为字段拷贝中网关排重和扩展功能的并集,将合并后的值进行相应的处理,默认合并不进行任何操作,处理目标为“----”(空的意思)。



图 2.27 字段合并

2.6 高级设置

2.6.1 多线程设置



图 2.28 高级设置

勾选了多线程分段处理以后，就可以对任务进行多线程处理了。默认情况下，线程总数为 1，每个线程处理的最大记录数为 10000，可以指定分段字段名或者采用配置任务时选定的主键。

其中线程总数指的是如果记录数目足够多，运行任务的线程数目的上限。而每个线程处理的最大记录数为每个线程针对分段的列（比如主键）取一个初始值，然后在初始值的基础上加最大记录数，对于落在此区间的记录，由一个线程来处理。比如，一个任务针对列 ID（为数字类型）分段，线程总数为 3，每个线程处理的最大记录为 1000，则如果表中的记录 ID 最小为 500，最大为 15000。则系统会安排三个线程取数据，第一个线程负责 ID 的值在 500 -- (500+1000) 的数据，第二个线程负责 ID 在 (500+1000) -- (500+2*1000) 的数据，第三个线程负责 ID 的值在 (500+2*1000) -- (500+3*1000) 的数据，每个线程执行完之后会推出，接着下一轮调度中再分配三个线程（有可能仍然是刚才的三个线程）继续往下执行，直到某个线程处理完所有的数据。

也可以自己指定另一个非主键的字段对数据进行分段，要保证字段的值尽可能的分散，而且是数值类型。

对于 DB2，由于系统限制一个连接取大对象的数目有一个上限，因此，建议设置成线程总数为 1，每个线程处理的数目小于 10000 条。

全部更新断点续传，是针对全部更新任务的一种特殊处理，勾选之后将自动选中多线程处理，其中线程总数只能是 1，处理的最大记录数可修改。用户可理解为使用一个线程进行操作的分段任务。该选择会在全部更新出错后，表现不同的处理逻辑，分为以下三类：

1)、如果是任务运行的段数超过一段后，对于关系型数据库（如：网关与关系型数据库通信中断）出现错误导致任务失败，网关将每隔 10 秒自动重试，尝试 10 次，自动进行错误续传工作，无须人工干预。尝试 10 次均失败后，任务标记状态运行失败，下次执行任务时将提示“重新执行”和“从上次错误开始执行”，供用户选择。

2)、如果是 TRS Server 错误（如：Server 正在优化，排它独占或无法连接）导致任务失败，网关将每隔 10 秒自动重试，尝试 10 次，自动进行错误续传工作，无须人工干预。尝试 10 次均失败后，任务标记状态运行失败，下次执行任务时将提示“重新执行”和“从上次错误开始执行”，供用户选择。

3)、如果任务运行的段数超过一段后，网关异常退出（可能的原因：操作系统关机或网关进程被强

制杀死) 导致任务未运行完成, 下次启动执行时将提示“重新执行”和“从上次错误开始执行”, 供用户选择。

特别注意的是, 全部更新断点续传的分段字段也是要求数值类型, 并且对应到 TRIS Server 的字段是 NUMBER 类型, 否则将可能出现结果不一致。

2.6.2 字符集设置

勾选了字符集指定之后, 可以指定字符集设置, 源数据库字符集指关系型数据库, 此处的字符集指的是客户端字符集而不是服务器中数据库的字符集, 例如 Oracle 服务器中字符集为 UTF-8, 客户端为 GBK 则此处要指定为 GB18030。目标字符集指的是 TRIS 中库的字符集, 此处与客户端无关。如果源数据库字符集与目标数据库字符集字符不同则进行字符转码, 如果相同则按照指定的字符集格式装库。比如指定两者都是 UTF-8, 则不进行字符集转码, 而采用 UTF-8 接口装库, 如果不进行指定, 则缺省采用 GB18030 装库。

对于关系型数据库到 TRIS 格式文件, 如果指定的源数据库字符集与目标数据库字符集字符不同, 仍然会对记录进行转码。即如果指定了源数据库字符集为 GB18030, 目标数据库为 UTF-8 则产生的 TRIS 格式文件为 UTF-8。

2.6.3 引导符前缀

用户在创建任务时可指定引导符前缀, 定义请查阅《TRIS DataBase 全文数据库系统使用指南》。

2.6.4 FTP 参数设置

应用中, 存在源数据库某字段记录的不是文件的具体内容, 而是文件存在于文件服务器上的路径, 此时就需要设置 FTP 参数, 程序运行时会根据设置的参数登录 FTP, 并按字段记录的路径下载文件。

其中:

连接模式有主动被动之分, 默认选择被动模式。如果网关在下载文件时报出类似“550 Passive mode not allowed on this server”提示, 此时应选择主动模式, 并在网关服务安装目录下创建文件 localhost.txt, 写入网关服务所在机器的 IP 地址。

注意: 如果网关服务和 FTP 服务器处于不同网段, 并有防火墙设置, 此时需要主动、被动均做尝试, 以找到合适的方式下载文件。强烈建议用户将网关服务与 FTP 服务器部署在同一网段。

FTP 服务器字符集选择 FTP 服务器所在操作系统的字符集, 默认为 GB18030。

FTP 服务器包含大量大文件 (如 20M 以上) 默认未勾选, 请根据实际情况设置, 该选项只是改变 FTP 下载文件的处理逻辑。



图 2.29 FTP 设置

2.6.5 CKM 参数设置

对于选择了字段拷贝=>扩展功能中 CKM 功能，需要设置相关功能的属性。注意，所有的路径设置都是相对网关服务所在的机器。



图 2.30 CKM 设置

2.7 查看和修改任务属性

2.7.1 常规属性

任务的属性包含了该任务的任务名、任务类型、登录的源数据库和目标数据库基本信息。其中对于手工配置目标数据库，网关自动生成一个 gw_auto_view 的虚拟视图，该视图并非真实存在，只是网关用于管理手工添加的后台数据库。



图 2.31 常规属性

点击修改登录设置之后，会显示更改登录设置页面。用户只可以修改源数据库或目标数据库的用户名密码。点击“确定”后，如果新的密码正确则进入属性查看，否则显示登录错误，要求重新登录。



图 2.32 更改登录设置

2.7.2 字段匹配

任务属性的“字段对应”页面显示的是当前任务的所有的字段匹配情况，该页面由两部分组成：

- 字段对应列表

字段对应列表是由源数据和目标数据已匹配字段组成，左侧为源数据字段列表，右侧为目标数据字段列表，两个列表均包含了“字段名”、“类型”、“长度”三列描述信息，左右同一行的字段为一组匹配字段。

- 应用字段变化

如果按照任务配置任务的时候对字段对应进行修改，或者对大字段的处理方式进行了修改，点击应用或者确定按钮以后，字段对应关系将发生变化。



图 2.33 字段匹配

2.7.3 任务的定时

设定任务运行的时间：

- 仅运行一次：

任务配置完以后，需要点击运行按钮，手工运行。

- 每天运行多次：

需要设置任务运行的起始时间，以及相互之间的时间间隔，时间间隔的单位可以为“小时”，“分钟”或“秒”。系统支持的最小时间间隔是 30 秒，当用户设置间隔小于 30 秒时，系统默认设置为 30 秒。

- 每天运行一次：

需要设置运行的具体时间，时间的形式为“XX 小时:YY 分钟:ZZ 秒”。(X,Y,Z 为合理数字)

- 每周执行一次：

执行方式为星期几:DD， 执行时间形式为" XX 小时:YY 分钟:ZZ 秒" (X,Y,Z 为合理数字)



图 2.34 定时

2.8 选项菜单

2.8.1 加载扩展功能插件

客户端控制台向网关服务发送加载扩展功能插件请求，后台网关服务进行相应操作，成功与否都会返回相应的信息。只有成功加载扩展功能插件，才能在字段拷贝和字段合并中使用相关的扩展。用户可加载任意按照《扩展功能插件编写规范》编写的组件，对源数据进行加工。

2.8.2 快速创建增量任务

该菜单项的功能是根据一个全部更新任务，创建与之对应的增量更新任务，源库，目标库，字段对应，字段扩展等完全等同全部更新任务。以此建立的增量更新任务任务名为：全部更新任务名+“_INC”。只有在选中全部更新任务时，此菜单项才可用。快速创建增量任务如失败，将提示错误信息，并将增量更新的配置文件保存在客户端机器上，命名为：任务名.single。

2.8.3 上传单一任务

对于客户端已存在的任务配置文件，可通过上传的方式提交到服务器，如果一切正常，服务上将成功添加，并在客户端控制台上显示。此操作一次只能上传一个任务的配置。

对于增量更新任务，用户可以选择提交后不创建临时表触发器。该功能用于任务已创建好临时表触发器，或共用临时表触发器。完全更新任务勾选与否无实质意义。

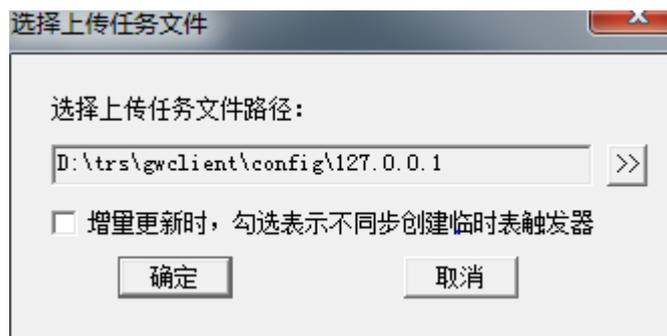


图 2.35 上传任务

一个的应用场景：客户端在创建完任务，然后提交网关服务器，因为种种原因都有可能失败，比如任务重名，创建临时表触发器失败，连接超时，网络断开等。此时客户端会将配置文件保存在 config 目录以网关服务地址命名的文件夹下，命令为：任务名.single。用户在找出任务失败的原因后，可手动修改相应配置，重新上传，以节省重新创建任务的时间。

注意：该方法不推荐一般用户使用，如需不了解网关配置文件请咨询产品人员。

2.8.4 上传或升级配置文件

网关升级到最新版本后，一般会带来原任务配置文件无法直接使用的问题。此时可以通过配置文件转换工具升级。也可以使用客户端上传或升级配置文件，网关服务程序自动识别配置文件的版本，再进行相应的转化，达到可用。此操作可以一次上传多个任务的配置。

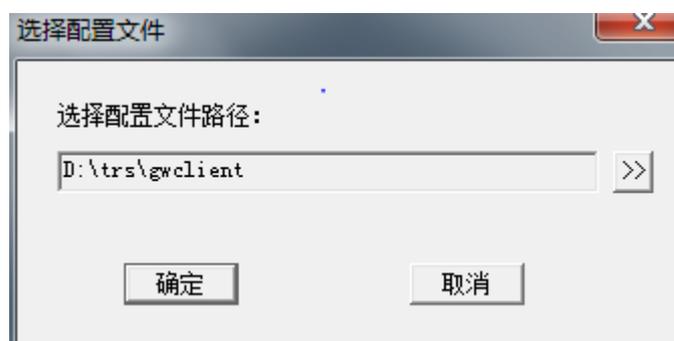


图 2.36 上传或升级配置文件

注意：上传或升级配置文件要求在网关服务端任务列表为空时才能进行，最大限度的避免任务重名。

2.8.5 日志设置

参数设定中用户可选择按文件大小备份和按日期间隔备份两者之一，日志文件保存的份数必填。

选择按文件大小备份时，需要设置日志文件最大值（单位为：M）。程序在写日志文件时，先检查当前文件大小是否超过了设定的值，如果否，继续在当前文件中写日志，如果是，程序就将当前日志文件进行备份，再写日志。

选择按日期间隔备份时，需要设置日志文件时间间隔（单位为：天）。程序在写日志文件时，先检查程序已运行的天数是否超过了设定的值，如果否，继续在当前文件中写日志，如果是，程序就将当前日志文件进行备份，再写日志。

日志文件保存的份数的设定，约束着 Log 文件夹下同一类型文件的最大个数，日志的份数如果超过这个值，新备份的日志文件将覆盖最旧的日志文件，形成循环。

精简模式表示只输出开始，错误和结束的日志，完全模式表示输出所有日志。

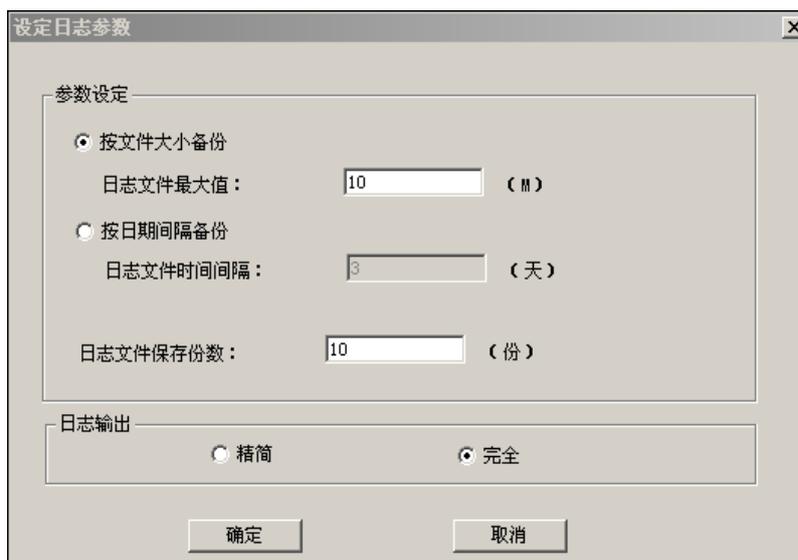


图 2.37 设定日志参数

2.9 版本信息查看

2.9.1 版本说明

TRS 全文检索网关客户端，服务器分别有自己的版本号。其中高版本客户端将有可能连接不上低版本的服务器，或存在一些操作未定义异常。如 TRS Gateway Client 4.6 版 Build4004 介质存在无法连接 Linux 下 TRS Gateway4.5 的服务。

详细的客户端版本，服务器版本可通过帮助菜单查看。也可以通过服务端命令参数查看，Windows 和 Linux 下可以通过 `TrsAgent -version` 查看版本信息。

```
D:\TRS\TRSGateway4.6sp1\gwserver>TrsAgent.exe -version
TRS(R) TRSGateway server 4.6 B4004, build time: Mar 1 2011 13:15:12
```

图 2.38 Windows 下使用命令查看版本信息

```
[wanghan@localhost gwserver_x64]$ ./TrsAgent -version
TRS(R) TRSGateway server 4.6 B4004, build time: Mar 11 2011 11:09:26
```

图 2.39 Linux 下使用命令查看版本信息

服务端具体的命令参数可以通过 `-help` 查看。

```
D:\TRS\TRSGateway4.6sp1\gwserver>TrsAgent.exe -help
Usage:
  -install install to the windows service
  -remove remove from the windows service
  -version check TRSGateway version information
```

图 2.40 Windows 下 TrsAgent 命令参数

```
[wanghan@localhost gwserver_x64]# ./TrsAgent -help
Usage:
    -version check TRSGateway version information
    -daemon run gateway in the background
```

图 2.41 Linux 下 TrsAgent 命令参数

具体参数说明参见“[3.7 服务端支持的命令参数](#)”。

2.9.2 注册信息查看

网关客户端的帮助菜单，关于 TRS Gateway Client 项包含完整的客户端信息以及服务器信息，其中服务器信息中包含网关服务所运行的平台，版本信息，注册信息等，用户可查看网关服务的使用期限以及注册组件情况。

“重新注册”功能方便用户在客户端注册网关服务。



图 2.42 注册信息查看

第3章 服务端使用说明

3.1 Windows 网关服务注册与启动

网关服务程序默认情况下可直接运行(点击 TrsAgent.exe)，也可先注册成服务，交于操作系统管理。注册服务的步骤如下：

首先命令行下切换到网关服务安装目录：

```
C:\TRS\TRSGateway4.6\gwserver_x64>dir
驱动器 C 中的卷没有标签。
卷的序列号是 8C01-24A0

C:\TRS\TRSGateway4.6\gwserver_x64 的目录

2010/07/20  17:57    <DIR>          .
2010/07/20  17:57    <DIR>          ..
2010/07/20  17:57    <DIR>          base2
2010/07/20  17:57    <DIR>          CKM
2010/04/01  13:46                175 ckm_sample
2010/06/09  17:02           2,332,672 convert_64.dll
2010/06/09  17:02           168,448 dmcjava.dll
2010/06/09  17:02           193,024 dmc_ahgralzwmp.dll
2010/06/09  17:02           591,872 dmc_comm.dll
2010/06/09  17:02            65,024 dmc_conf.dll
2010/06/09  17:02           287,232 dmc_dtct.dll
2010/06/09  17:02            76,288 dmc_oscomm.dll
```

图 3.1 切换目录

其次查看服务注册说明：

```
C:\TRS\TRSGateway4.6\gwserver_x64>TrsAgent.exe -help
Usage :
    -install install to the windows service
    -remove remove from the windows service
    -version check TRSGateway version information

C:\TRS\TRSGateway4.6\gwserver_x64>
```

图 3.2 服务注册说明

最后根据说明进行相应操作：

TrsAgent.exe -install ----- 注册服务

TrsAgent.exe -remove ----- 卸载（删除）服务

TrsAgent.exe -version ----- 查看服务端的版本信息

同时 Windows 上网关服务提供一个托盘程序（GatewayMonitor.exe），方便用户操作网关服务的启停。

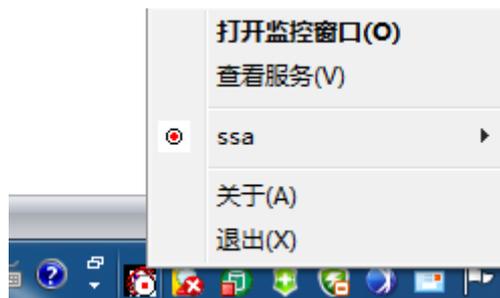


图 3.3 服务托盘

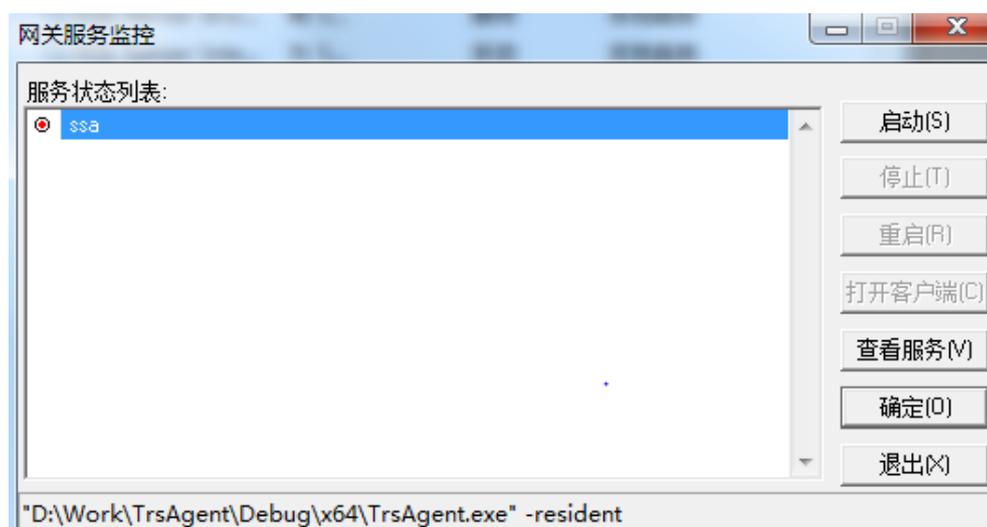


图 3.4 打开监控窗口

注意：在 Windows Vista 或 Windows 7 下注册和删除服务需要管理员权限，用户打开命令行时请选择以管理员身份运行，才能保证后续的操作成功。托盘程序也需要以管理员身份运行，才可操作服务。

另外默认情况下，一台机器只能启动一个网关服务器的实例。

3.2 Linux 网关服务启动

Linux 平台上网关默认只提供命令行方式启动，请运行 startup.sh 脚本。不推荐使用 TrsAgent 直接运行，会导致由于部分环境变量未设置，而加载不了组件的情况。

```
[root@localhost gwserver_x64 ]# pwd
/opt/trs/gwserver_x64
[root@localhost gwserver_x64 ]# ./startup.sh
TRS Gateway 4.6 版 Build4004
版权所有 (C) 2000-2011
北京拓尔思 (TRS) 信息技术股份有限公司
欢迎使用TrsAgent. 当前注册使用截止日期: 2011年9月12日!
该版本未注册使用扩展功能插件!
已注册的数据库组件为:
SQL Server
Oracle
load /opt/trs/gwserver_x64 /librdb_sql.so success!
binding and listening localhost:8899 succeed!
```

图 3.5 启动信息

startup.sh 为 Bash Shell 脚本，记录网关安装目录以及各种数据库需要设置的变量，用户需根据服务器上的情况适当修改信息，保证已注册数据库组件的成功加载。

同样，默认情况下，一台机器只能启动一个网关服务器实例。但有时会报出如下信息：open in check_single_on(): Permission denied，导致一个实例都无法启动。这是因为先前使用了高权限用户启动网关，现在再使用低权限用户启动，权限不足。解决的一个办法是删除/tmp/gw_proc.pid 文件即可。

3.3 网关服务的停止

对于运行的网关服务，用户可通过多种方式进行关闭。

Windows 平台下如已注册成服务，可使用网关服务托盘程序（GatewayMonitor.exe）关闭服务，也可直接进入 Windows 服务管理器操作。也可使用 gw_shutdown.exe 程序进行关闭。对于直接运行的网关服务，用户还可以使用 CTRL+C、CTRL+Break 等方式关闭。

```
D:\TRS\TRSGateway4.6\gwserver>TrsAgent.exe
Create global namespace file mapping success.
欢迎使用TrsAgent。当前注册使用截止日期：2011年10月20日！
该版本未注册使用扩展功能插件！
已注册的数据库组件为：
SQL Server
Oracle
load D:\TRS\TRSGateway4.6\gwserver\rdb_sql.dll success!
load D:\TRS\TRSGateway4.6\gwserver\rdb_ora.dll success!
binding and listening localhost:8004 succeed!
get a signal CTRL-C, server is stopping, please wait...
server is stopped!

D:\TRS\TRSGateway4.6\gwserver>
```

图 3.6 Windows 关闭网关服务

Linux 平台下，用户可使用 shutdown.sh 脚本进行关闭，也可直接在运行着网关服务的终端使用 CTRL+C、CTRL+^等关闭。

```
[root@localhost gwserver_x64 ]# ./startup.sh
TRS Gateway 4.6 版 Build4004
版权所有 (C) 2000-2011
北京拓尔思 (TRS) 信息技术有限公司
欢迎使用TrsAgent。当前注册使用截止日期：2011年9月12日！
该版本未注册使用扩展功能插件！
已注册的数据库组件为：
SQL Server
Oracle
load /opt/trs/gwserver_x64 /librdb_sql.so success!
binding and listening localhost:8899 succeed!
get a signal SIGINT, server is stopping, please wait...
server is stopped!
```

图 3.7 Linux 关闭网关服务

3.4 网关注册码升级

首次安装 TRS 全文检索网关新产品，将有三个月的试用期间，并且注册组件 SQL Server 和 Oracle，如用户使用过期，或组件不满足应用需要，可根据合同向 TRS 公司产品中心申请新的产品注册码。

注册的过程需要使用 register.exe(Windows)或 register.sh (Linux)。将输出的机器码提交给产品中心，并说明需要注册的数据库组件以及是否注册扩展功能插件。再将产品中心返回的注册码输入（粘贴）到注册终端（命令行）下。



图 3.8 Windows 注册码升级

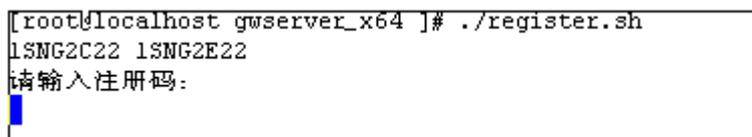


图 3.9 Linux 注册码升级

也可以在试用未过期的情况下，连接上网关服务，使用客户端重新注册。



图 3.10 客户端重新注册

3.5 扩展插件 CKM 的使用

网关的安装包中未打包 CKM 相关文件（字典和动态库），所以即使已注册扩展功能插件，网关服务也无法成功加载插件。用户如需使用扩展功能，首先确认注册码是否支持扩展功能插件功能，如不支持，请先升级注册码（参照网关注册码升级），再从产品服务器上下载 CKM 字典以及相关动态库，拷贝至网关服务器安装目录下 CKM 子目录中，放置结构如下：

文件	大小	类型
DICT		文件夹
cat2ext.dll	160 KB	应用程序扩展
clukit20_c.dll	500 KB	应用程序扩展
convert.dll	2,244 KB	应用程序扩展
crawltool.dll	172 KB	应用程序扩展
libdb43.dll	692 KB	应用程序扩展
pcre.dll	204 KB	应用程序扩展
rtk_32.dll	76 KB	应用程序扩展
sqlite3.dll	508 KB	应用程序扩展
stk_crawl.dll	192 KB	应用程序扩展
t40api32.dll	456 KB	应用程序扩展
trs_abs.dll	124 KB	应用程序扩展
trs_casename.dll	72 KB	应用程序扩展
trs_cat.dll	344 KB	应用程序扩展
trs_ckm_dy.dll	936 KB	应用程序扩展
trs_ckm_py.dll	968 KB	应用程序扩展
trs_ckm_seg.dll	496 KB	应用程序扩展
trs_ckm_version.dll	40 KB	应用程序扩展
trs_clu.dll	396 KB	应用程序扩展
trs_collation.dll	136 KB	应用程序扩展
Trs_Compare.dll	80 KB	应用程序扩展
trs_CorrectEn.dll	64 KB	应用程序扩展
trs_data_contrast.dll	84 KB	应用程序扩展
trs_eval.dll	80 KB	应用程序扩展

图 3.11 网关 CKM 目录结构（Windows 环境）

文件	大小	类型
DICT		Folder
libckm_abs.so	339,189	Apache L...
libckm_casename.so	281,116	Apache L...
libckm_cat.so	152,086	Apache L...
libckm_catrain.so	492,393	Apache L...
libckm_clu.so	246,307	Apache L...
libckm_collation.so	149,249	Apache L...
libckm_contrast.so	3,477,763	Apache L...
libckm_dy.so	299,585	Apache L...
libckm_ext.so	102,929	Apache L...
libckm_filt.so	12,468	Apache L...
libckm_house.so	3,364,928	Apache L...
libckm_keyword.so	261,180	Apache L...

图 3.12 网关 CKM 目录结构（Linux 环境）

注意：32 位的 gateway 需要 32 位 CKM 的动态库。64 位的 gateway 需要 64 位的 CKM 动态库。

3.6 服务器部分文件及参数说明

网关服务器端存在多个文本文件，分别有着不同的含义，并且直接影响网关服务的行为。

KEY.INI、license.dat: 共同维护 TRIS 全文检索网关的注册信息，用户不可手动修改，以免造成服务无法启动。重新注册或升级请参照网关注册码升级一节。

gw_ext.cfg、ckm_sample: 共同维护扩展功能插件功能，用户不可随意更动，如须增加，请咨询产品部门。

port.cfg: 该文件记录网关服务所使用的端口，默认值为：8004。用户可直接修改，变更网关端口，修改后重启网关服务生效。

validatecode.txt: 该文件记录客户端连接网关服务时所需要的验证信息，默认值为：12345678，如果该文件为空，按默认值处理。用户可直接修改，修改后立即生效，无需重启网关服务。

TRSAGENT.INI: 该文件在设置完日志参数时出现，内容如下：

[LOG]

LOG_SIZE=10485760

BAK_SPAN=3

BAK_COUNT=10

LOG_TYPE=0

OUTPUT_TYPE=1

[EXT_DLL]

AUTOLOAD=0

[SOCKET]

MAX_CONNECT=6

KEEP_TIME=0

USE_SSL=0

[HYFSOFT]

ANOTHER_PROCESS=0

[CLUSTER]

SHARE_TEMP=0

其中[LOG]部分配置日志行为，不建议手动修改，请使用客户端日志参数菜单设置。[EXT_DLL]设置扩展功能插件行为，AUTOLOAD=1表示网关服务启动时自动加载扩展功能插件，默认不自动加载。[SOCKET] MAX_CONNECT 设定网关服务监听连接数，默认值为 6，而一个客户端需要两个连接，所以同时允许三个客户端登录，用户可根据实际情况调节值大小。KEEP_TIME 设置服务器主动断开与客户端连接的时间，从连接开始计时，单位是小时，默认值为 0，0 表示不断开。USE_SSL=0 表示登录服务器时客户端不需要勾选 SSL 选项，USE_SSL=1 表示需要勾选 SSL 选项。[HYFSOFT]配置使用其他进

程进行红缨枫抽取，某些情况下由于红缨枫产品的缺陷，抽取文件时可能造成网关的宕机，此时设置 ANOTHER_PROCESS=1，网关将使用额外进程进行抽取处理，保证网关的正常运行。[CLUSTER]配置用于大型项目中的扩容需求，SHARE_TEMP=0 表示每个增量任务单独使用自己的临时表触发器。SHARE_TEMP=1 时，通过选择条件，可以实现多个任务共用临时表触发器，达到一定规则的数据入一个 TRS 集群，另一规则的数据入另一个 TRS 集群。以上所有参数在网关服务重启后生效。

3.7 服务端支持的命令参数

服务端支持参数命令，具体的命令参数可以通过 -help 查看。Windows 下参数有：

```
D:\TRS\TRSGateway4.6sp1\gwserver>TrsAgent.exe -help
Usage:
  -install install to the windows service
  -remove remove from the windows service
  -version check TRSGateway version information
```

图 3.13 Windows 下 TrsAgent 命令参数

其中：

- install 表示注册为系统服务。
- remove 表示卸载（删除）服务。
- version 表示查看服务端的版本信息。

Linux 下支持的参数有：

```
[wanghan@localhost gwserver_x64]$ ./TrsAgent -help
Usage:
  -version check TRSGateway version information
  -daemon run gateway in the background
```

图 3.14 Linux 下 TrsAgent 命令参数

其中：

- version 表示查看服务端的版本信息。
- daemon 表示使网关服务在后台运行。

第4章 辅助工具使用说明

4.1 日志分析工具

日志分析工具是帮助用户在网关任务运行出错时，快速定位问题的一种辅助工具。其工作的原理是将加载的文件内的日志信息按时间先后顺序从小到大的进行排序，当指定查找的任务名时，分析工具会将该任务的日志信息完整输出。可以说，网关的日志功能逻辑上是将日志分类保存，而日志分析工具恰好相反，将各类信息进行汇总。



图 4.1 日志分析工具

点击“添加”按钮选择加载的文件，文件可多选。增加文件后点击“加载日志”按钮，进行日志的汇总，加载完成后在任务名称列表中显示日志文件中存在的所有任务名，“分析日志”按钮可用。然后选择要分析的任务名称，点击“分析日志”按钮进行分析，如果存在该任务的相关日志信息，则在 gwserver 的 LogParserOutput 目录下以任务名称命名生成目录，目录下存储相应的日志分析结果文件。

另外，用户可选择任务运行的时间段，输出则为该时间段内任务运行的信息。缺省情况为完全输出。

“标记错误发生时的 SQL 语句”选项，用来输出错误发生时，上一次执行的 SQL 语句（生成*.sql 文件），方便用户定位问题。省略情况为选中状态。

Agent.inf 保存任务的基本日志，agent.sql 保存任务创建和执行中涉及到的 sql 语句，agent.err 保存网关发生的错误信息，system.err 保存 TRSServer 返回的错误信息或者操作系统返回的错误信息，db.err 保存操作数据库时返回的错误信息。

4.2 配置文件转换工具

TRS 全文检索网关采用 C/S 架构后，配置文件也做了相应转变。为了兼容网关以前版本的配置文件，4.6 提供一个配置文件转换工具，平滑升级，使以前的配置继续可用。

Windows 平台下运行 Upgrade.exe，选择待转换配置的版本，生成 4.6 兼容的配置：agent.cfg.46

```
C:\TRS\TRSGateway4.6\gwserver>Upgrade.exe
本工具为TRS全文检索网关配置转换工具，
其功能是将Windows下低版本网关配置文件转换成TRS Gateway4.6可用配置。

请先将待转换的网关配置文件拷贝到当前目录下。
请按任意键继续. . .
请选择待转换配置文件的网关版本：
1:TRS Gateway 3.0
2:TRS Gateway 4.0
3:TRS Gateway 4.5
你的选择是：3
转换完成，配置文件为当前目录下agent.cfg.46。
请按任意键继续. . .
```

图 4.2 Windows 配置文件转换

Linux 平台下运行 Upgrade，无须选择版本（Linux 下低版本网关只有 TRS Gateway 4.5），

```
[root@localhost Upgrade]# ./Upgrade
本工具为TRS全文检索网关配置转换工具，
其功能是将Linux下低版本网关配置文件转换成TRS Gateway4.6可用配置。

请先将待转换的网关配置文件拷贝到当前目录下。
回车后请继续...

转换完成，配置文件为当前目录下agent.cfg.46。
```

图 4.3 Linux 配置文件转换

转换完成后，请备份原 agent.cfg 文件，再将 agent.cfg.46 更改成 agent.cfg，启动网关服务即可成功加载任务配置。

附录1 字段对应表

1.1 RDBMS 和 TRS 字段对应表

1.1.1 ORACLE 和 TRS 的字段对应表

Oracle 字段类型	TRS 字段类型	说明
CHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
VARCHAR2	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
NCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
NVARCHAR2	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
NUMBER	NUMBER, CHAR	数值比较大对应 CHAR
FLOAT	NUMBER, CHAR	数值比较大对应 CHAR
DATE	DATE, CHAR	
LONG	DOCUMENT, PHRASE	
RAW	BIT	
LONG RAW	BIT	
BLOB	BIT, DOCUMENT	
CLOB	DOCUMENT	
NCLOB	DOCUMENT	
BFILE	BIT, DOCUMENT	

1.1.2 SQL SERVER 和 TRS 的字段对应

SQL Server 字段类型	TRS 字段类型	说明
char	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
nchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
varchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
nvarchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
int	NUMBER, CHAR	数值比较大对应 CHAR
decimal	NUMBER, CHAR	数值比较大对应 CHAR

numeric	NUMBER, CHAR	数值比较大对应 CHAR
tinyint	NUMBER, CHAR	数值比较大对应 CHAR
smallint	NUMBER, CHAR	数值比较大对应 CHAR
real	NUMBER, CHAR	数值比较大对应 CHAR
float	NUMBER, CHAR	数值比较大对应 CHAR
datetime	DATE	
smalldatetime	DATE	
money	NUMBER, CHAR	数值比较大对应 CHAR
smallmoney	NUMBER, CHAR	数值比较大对应 CHAR
bit	NUMBER, CHAR	数值比较大对应 CHAR
uniqueidentifier	CHAR	
timestamp	CHAR	
binary	BIT	
varbinary	BIT	
text	DOCUMENT, PHRASE	
ntext	DOCUMENT, PHRASE	
image	BIT	

1.1.3 SYBASE 和 TRS 的字段对应

Sybase 字段类型	TRS 字段类型	说明
char	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
nchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
varchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
nvarchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
unichar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
univarchar	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
sysname	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
longsysname	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 对应 DOCUMENT 或 PHRASE 或 BIT
int	NUMBER, CHAR	数值比较大对应 CHAR
unsigned int	NUMBER, CHAR	数值比较大对应 CHAR

bigint	NUMBER, CHAR	数值比较大对应 CHAR
unsigned bigint	NUMBER, CHAR	数值比较大对应 CHAR
decimal	NUMBER, CHAR	数值比较大对应 CHAR
numeric	NUMBER, CHAR	数值比较大对应 CHAR
tinyint	NUMBER, CHAR	数值比较大对应 CHAR
smallint	NUMBER, CHAR	数值比较大对应 CHAR
unsigned smallint	NUMBER, CHAR	数值比较大对应 CHAR
real	NUMBER, CHAR	数值比较大对应 CHAR
float	NUMBER, CHAR	数值比较大对应 CHAR
datetime	DATE	
smalldatetime	DATE	
date	DATE	
time	CHAR	
money	NUMBER, CHAR	数值比较大对应 CHAR
smallmoney	NUMBER, CHAR	数值比较大对应 CHAR
bit	NUMBER	
timestamp	CHAR	
binary	BIT	
varbinary	BIT	
text	DOCUMENT, PHRASE	
image	BIT	
unitext	DOCUMENT, PHRASE	

1.1.4 DB2 和 TRS 的字段对应

DB2 字段	TRS 字段	说明
CHARACTER	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
VARCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
LONG VARCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
GRAPHIC	CHAR, DOCUMENT, PHRASE	长度大于 256 时对应 Document 或 Phrase
VARGRAPHIC	CHAR, DOCUMENT, PHRASE	长度大于 256 时对应 Document 或 Phrase
LONG VARGRAPHIC	CHAR, DOCUMENT, PHRASE	长度大于 256 时对应 Document 或 Phrase
INTEGER	NUMBER, CHAR	数值比较大对应 CHAR
SMALLINT	NUMBER, CHAR	数值比较大对应 CHAR
BIGINT	NUMBER, CHAR	数值比较大对应 CHAR
REAL	NUMBER, CHAR	数值比较大对应 CHAR

DOUBLE	NUMBER, CHAR	数值比较大对应 CHAR
DECIMAL	NUMBER, CHAR	数值比较大对应 CHAR
DATE	DATE	
TIME	DATE	
TIMESTAMP	DATE	
BLOB	BIT	
CLOB	DOCUMENT, PHRASE	
DBCLOB	DOCUMENT, PHRASE	
XML	DOCUMENT, PHRASE	

1.1.5 MySQL 和 TRS 的字段对应

MySQL 字段类型	TRS 字段类型	说明
CHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
VARCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
NCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
NVARCHAR	CHAR, DOCUMENT, PHRASE, BIT	长度大于 256 时对应 DOCUMENT 或 PHRASE 或 BIT
TINYINT	NUMBER, CHAR	数值比较大对应 CHAR
SMALLINT	NUMBER, CHAR	数值比较大对应 CHAR
MEDIUMINT	NUMBER, CHAR	数值比较大对应 CHAR
INT	NUMBER, CHAR	数值比较大对应 CHAR
BIGINT	NUMBER, CHAR	数值比较大对应 CHAR
FLOAT	NUMBER, CHAR	数值比较大对应 CHAR
DOUBLE	NUMBER, CHAR	数值比较大对应 CHAR
DECIMAL	NUMBER, CHAR	数值比较大对应 CHAR
DATE	DATE, CHAR	
DATETIME	DATE, CHAR	
TIMESTAMP	DATE, CHAR	
TIME	DATE, CHAR	
YEAR	DATE, CHAR	
BIT	NUMBER	
BINARY	BIT	
VARBINARY	BIT	
TINYTEXT	DOCUMENT, PHRASE, BIT	
TEXT	DOCUMENT, PHRASE, BIT	
MEDIUMTEXT	DOCUMENT, PHRASE, BIT	
LONGTEXT	DOCUMENT, PHRASE, BIT	

TINYBLOB	DOCUMENT, PHRASE, BIT	
BLOB	DOCUMENT, PHRASE, BIT	
MEDIUMBLOB	DOCUMENT, PHRASE, BIT	
LOB	DOCUMENT, PHRASE, BIT	

附录2 Windows 环境 RDBMS 客户端配置

2.1 ORACLE 客户端的配置

在 Oracle8i 中, Oracle 服务器和客户端的交互可以通过 IPC(本机)、TCP/IP、SPX、Named Pipes 等多种协议。在 TRS for Oracle Gateway 中通过 TCP/IP 协议与 Oracle 服务器通信, 因此在安装 Oracle 客户端之后, 需要配置 Oracle Net8。

Oracle Net8 可以在安装 Oracle 客户端的过程中利用安装向导进行配置, 也可以在 Oracle 客户端安装完成之后运行:

- Net8 Assistant
- Net8 Configuration Assistant
- Net8 Easy Config

进行配置, 两者的配置过程和参数都是一样的, 以下以 Net8 Assistant 为例说明如何配置 Oracle Net8。

1) 启动 Oracle Net8 Assistant

在 Oracle 程序组 Network Administration 中选择 Net8 Assistant 运行 Oracle Net8 Assistant。

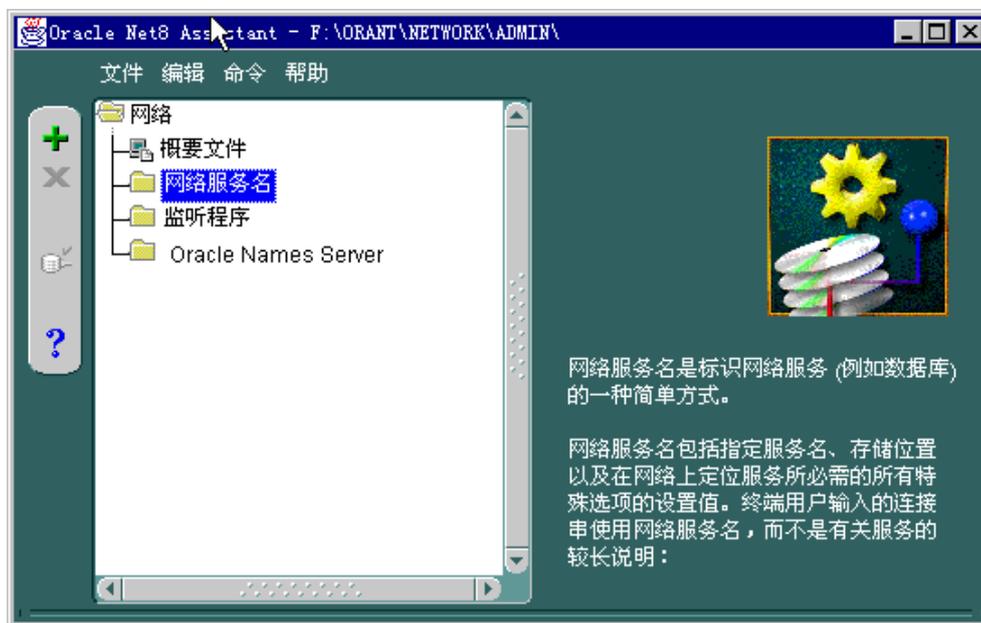


图 5.1 启动 Oracle Net8

2) 创建一个网络服务名

选择网络服务名, 在菜单或工具条中选择创建, 出现创建网络服务名的向导。向导分为 5 个步骤。

3) 配置网络服务名

此网络服务名可以任意指定，此名称在 TRS for Oracle Gateway 中设置 Oracle 服务名时需要使用，假设指定网络服务名为 ora166，进行下一步。

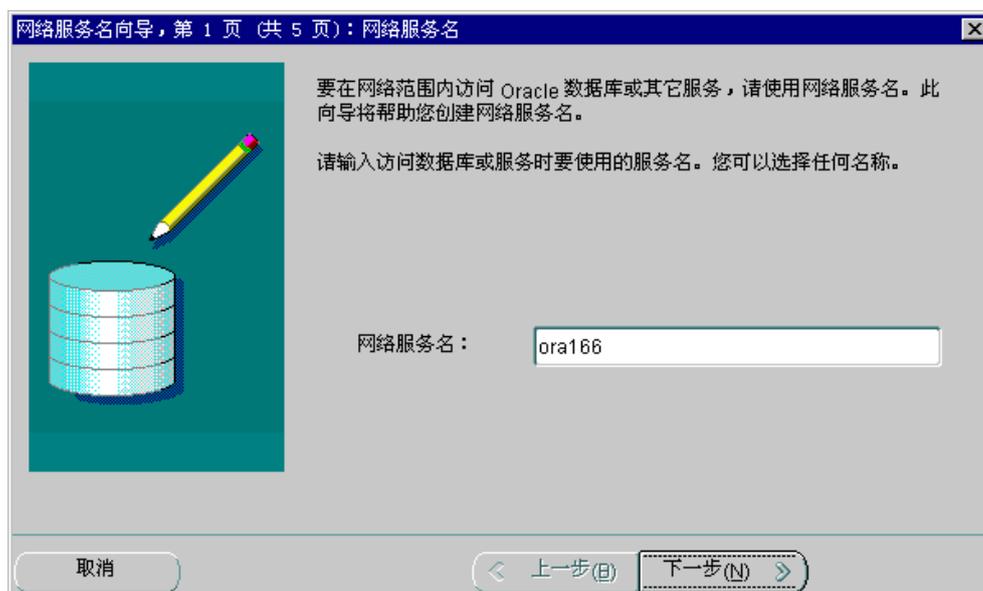


图 5.2 指定网络服务名

4) 选择协议

选择 TCP/IP 协议，进入下一步。

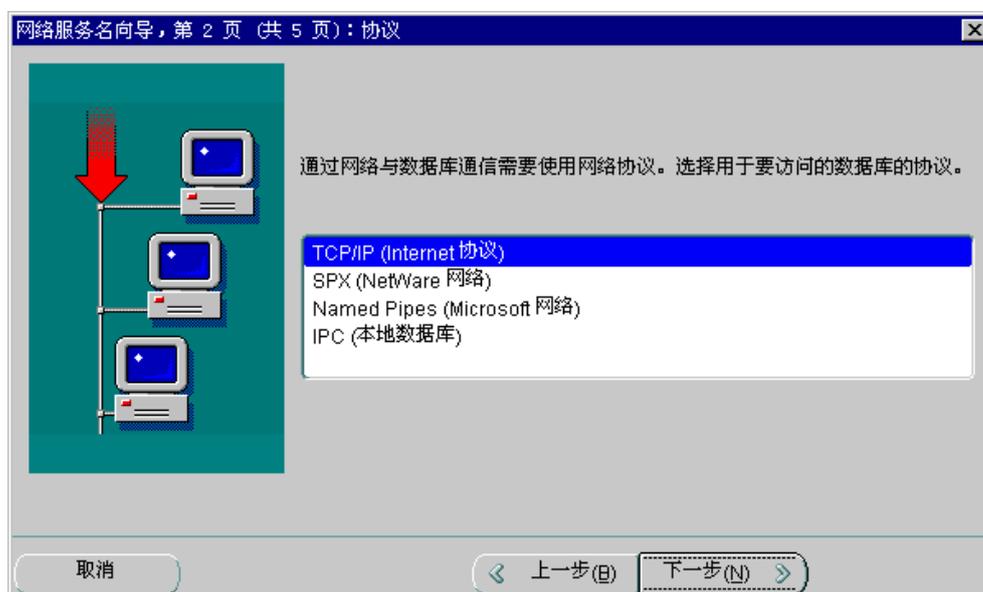


图 5.3 指定协议

5) 协议设置

需要设置 TCP/IP 协议的主机名和端口号。主机名为 Oracle 服务器的主机名或 IP 地址。端口号为 Oracle 数据库的端口号，通常为 1521。进行下一步。

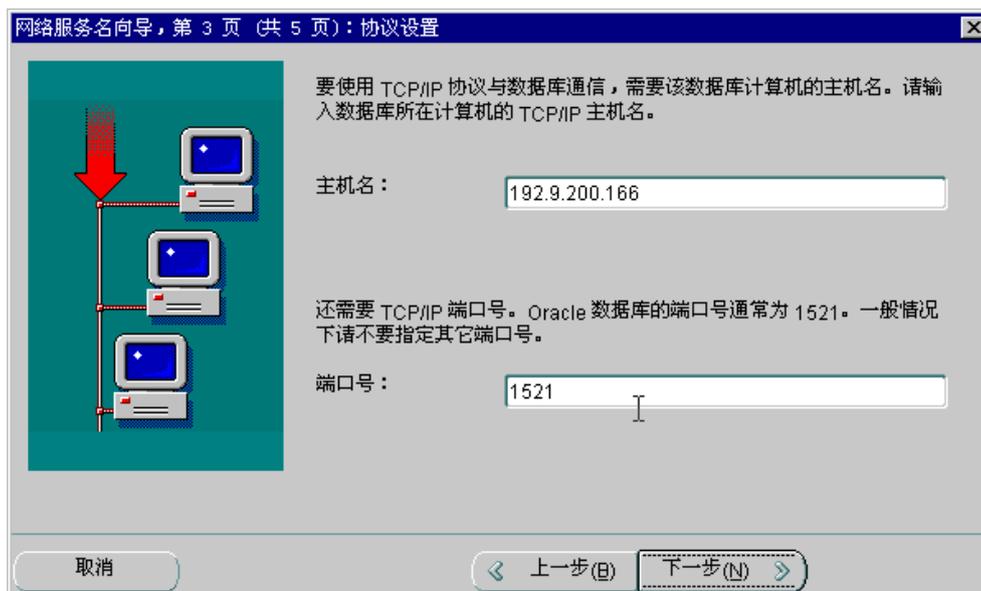


图 5.4 协议配置

6) 设置 Oracle 数据库的服务名

此服务名为 Oracle 数据库的全局数据库名, 如果不清楚此服务名, 请向数据库管理员(DBA)了解相关信息。进行下一步。

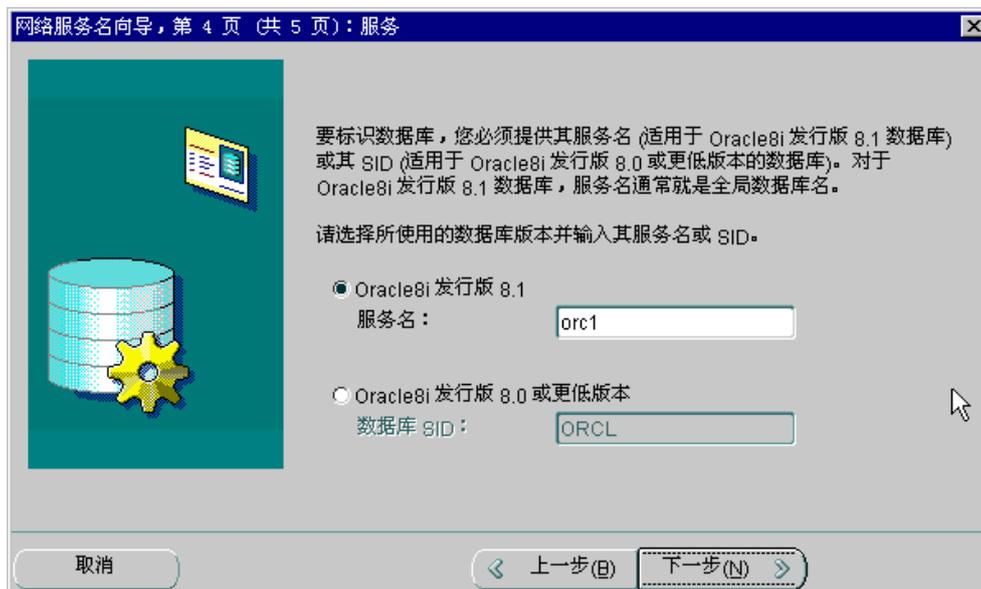


图 5.5 设置 Oracle 服务名

7) 测试以上配置

点击测试按钮, 测试以上配置是否正确, 如果连接成功则表明上述配置正确, 否则根据出错提示, 选择上一步逐个检查以上各个参数是否正确, 直到测试连接成功为止。退出 Net8 Assistant 之前, 请注意先保存网络配置。

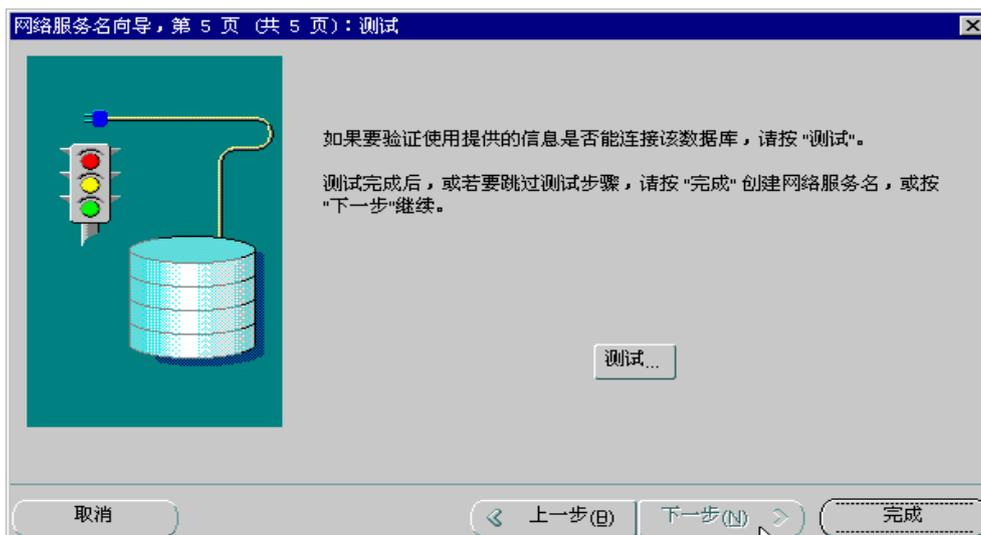


图 5.6 测试配置

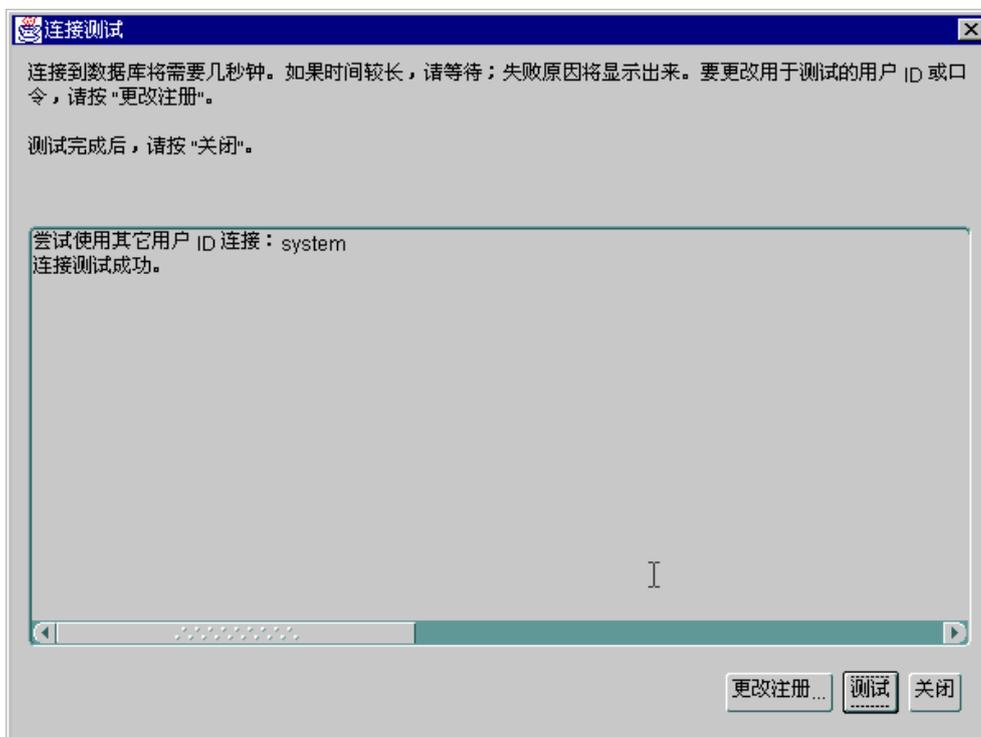


图 5.7 查看测试结果

2.2 SQL SERVER 客户端的配置

运行 TRS 全文检索网关 for Sql Server 必须安装 SQL SERVER 客户端。

- 配置客户端的网络环境

从 SQL Server 的程序组中运行 Client Network Utility 对 SQL Server 2000 客户端协议进行配置。

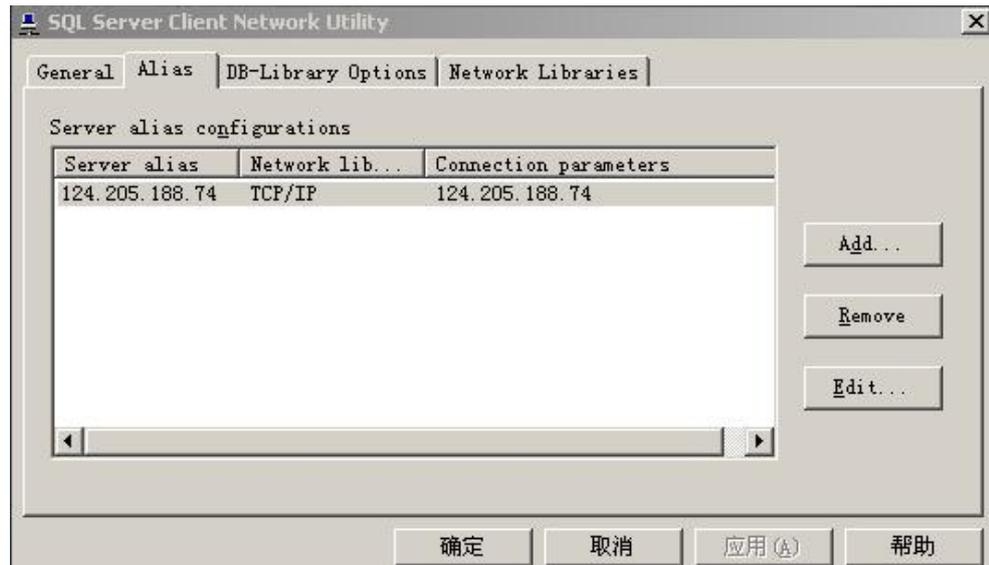


图 5.8 配置网络环境

首先，将缺省的网络协议设置成 TCP/IP；

其次，可以配置 SQL Server 2000 服务器的别名，配置好别名之后，在登录 SQL Server 2000 服务器时，服务器名就可以使用此别名了。如果没有配置别名，则登录 SQL Server 2000 服务器时，服务器名必须使用主机名或 IP 地址。

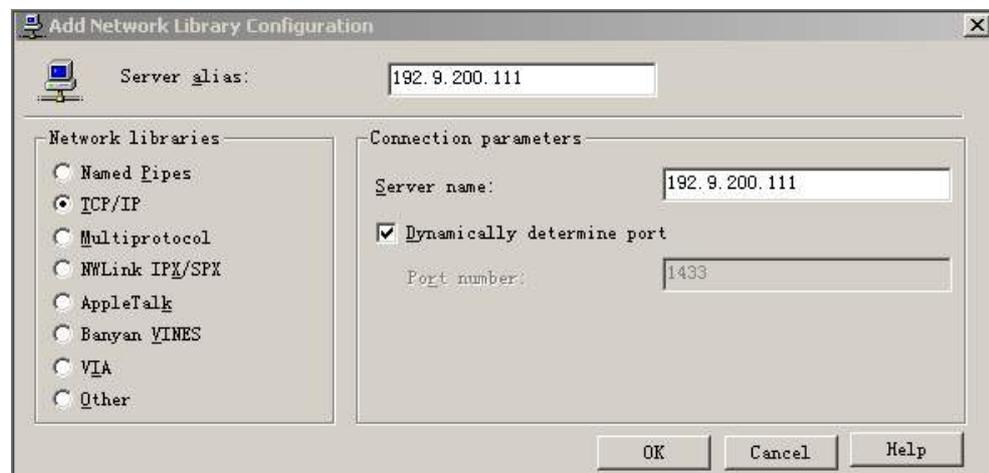


图 5.9 配置服务器别名

2.3 SYBASE 的客户端配置

运行 TRS 全文检索网关 for Sybase 必须安装 Sybase 客户端。

- 配置客户端的网络环境

从 Sybase 的程序组中运行 Dsedit 对 Sybase 客户端协议进行配置。

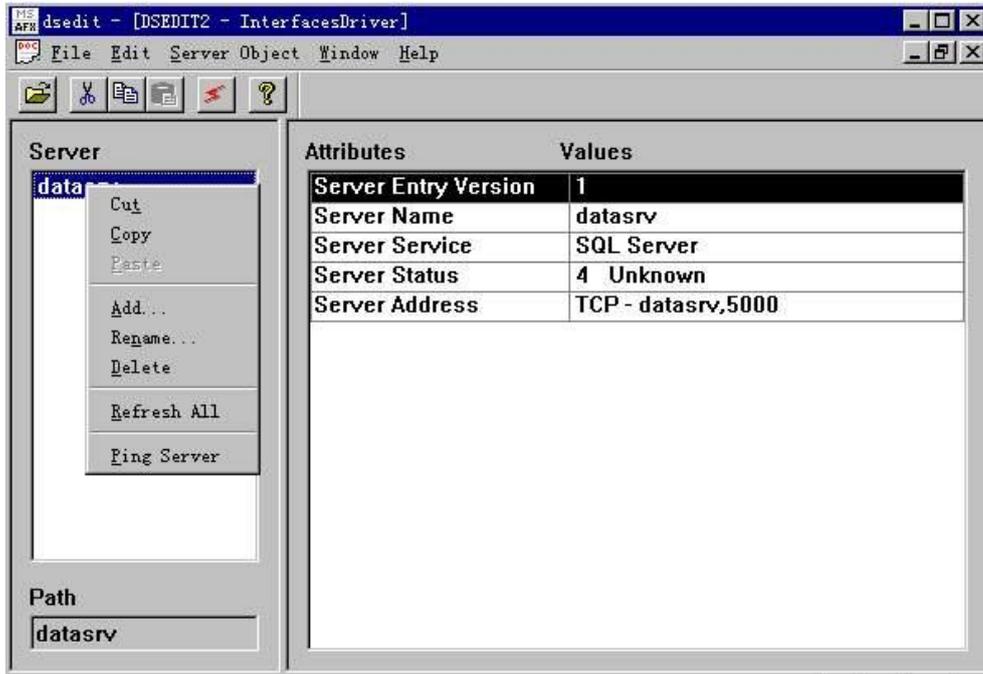


图 5.10 配置 Sybase 网络环境

首先，可以配置 Sybase 服务器的名字,再双击 Server Address 来修改网络协议；其次，可以配置 Sybase 使用多种协议，建议使用 TCP/IP 协议。

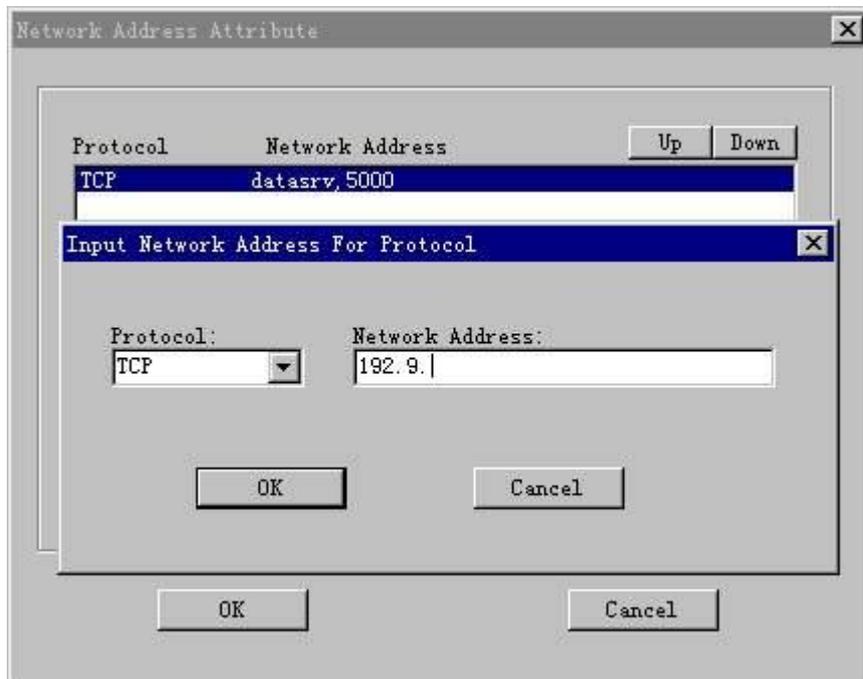


图 5.11 配置 Sybase 网络协议

2.4 DB2 的客户端配置

运行 TRS 全文检索网关 for DB2 必须安装 DB2 客户端。

- 启动 DB2 配置助手，选择“使用向导添加数据库”

- 选择要如何设置连接，选择手工配置
- 选择通信协议，建议使用 TCP/IP
- 设定参数，指定数据库，注册数据源，指节点选项，指定系统选项，指定安全项
- 最后测试连接

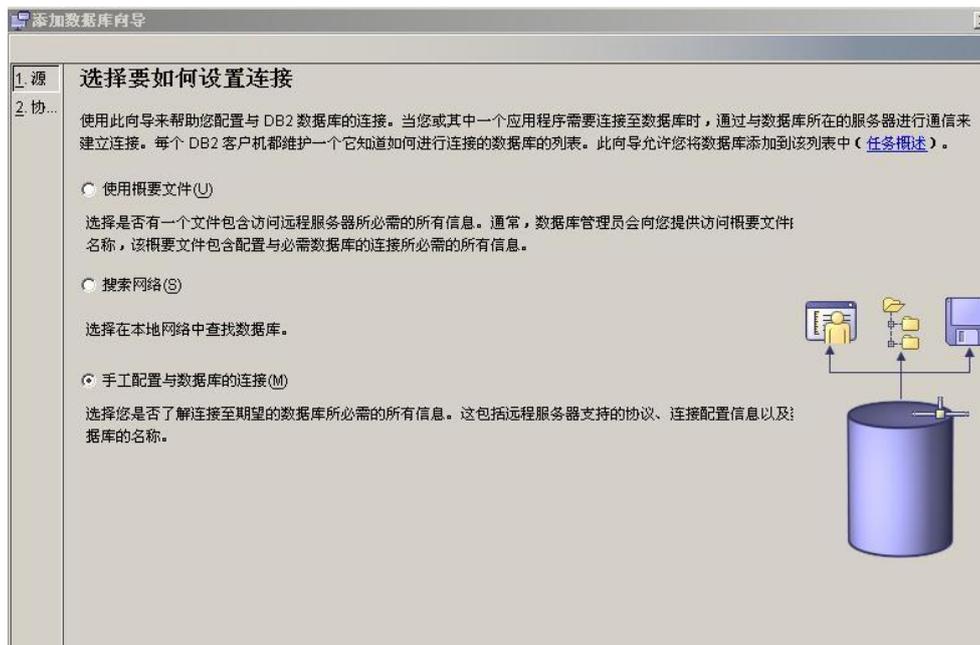


图 5.12 配置 DB2 客户端 1



图 5.13 配置 DB2 客户端 2



图 5.14 配置 DB2 客户端 3

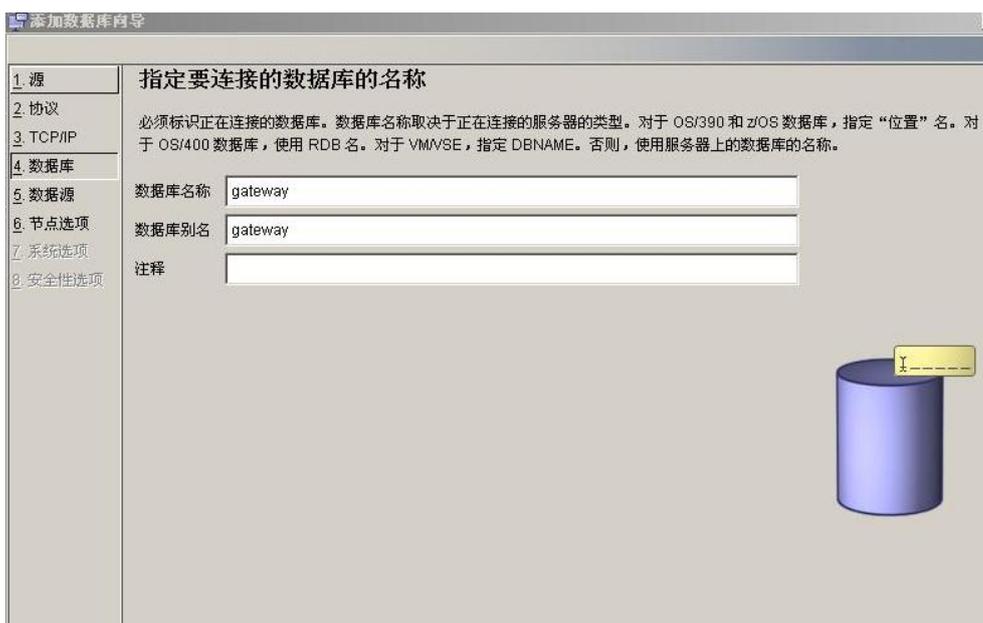


图 5.15 配置 DB2 客户端 4



图 5.16 配置 DB2 客户端 5



图 5.17 配置 DB2 客户端 6



图 5.18 配置 DB2 客户端 7



图 5.19 配置 DB2 客户端 8

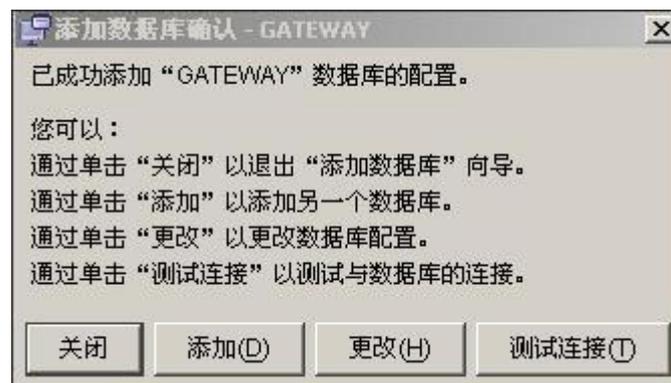


图 5.20 配置 DB2 客户端 9

2.5 MySQL 的客户端配置

网关服务自带 MySQL 连接所需要的动态库，用户只需要注册 MySQL 组件即可直接使用，无须额外配置。

对于 MySQL 数据库的任务，某些情况下需要手工增加 `msq.ini` 文件，以解决非默认端口的问题和灵活使用字符集。`msq.ini` 文件的增加路径为当前网关运行的工作目录，内容为：

```
<服务名>
<port>
...
</port>
<charset>
...
</charset>
</服务名>
```

举例说明：客户端配置的 MySQL 服务名为 `localhost`，端口为 `7788`，数据库的字符集为 `gbk`。`msq.ini` 文件应写入：

```
<localhost>
<port>7788</port>
<charset>gbk</charset>
</localhost>
```

说明：默认情况下，未配置 `msq.ini` 文件，任务执行时使用的端口为 `3306`，字符集为 `utf-8`。

附录3 Linux 环境 RDBMS 客户端配置

Linux 环境下 RDBMS 客户端配置分为两种方式，图形方式和命令行方式，图形方式形同 Windows 环境，请参照附录 2 内容，以下介绍命令行方式下客户端的安装和配置。

注意：对于 64 位的操作系统，建议用户安装与当前系统上数据库相同位数的网关介质，如若不同，配置过程中遇到困难请咨询开发人员。

3.1 ORACLE 客户端的配置

情况一：网关服务所在机器已安装 Oracle 的数据库或者客户端

通过图形化界面配置本地 Net 实例名，然后将以下环境变量设置到网关服务 startup.sh 中：

ORACLE_HOME、PATH、LD_LIBRARY_PATH、TNS_ADMIN、NLS_LANG 中，例如：

网关服务和 Oracle 数据库在同一台机器上：

```
export ORACLE_HOME=/oracle/app/oracle/product/10.2.0/db_1
export PATH=$PATH:/oracle/app/oracle/product/10.2.0/db_1/bin
export LD_LIBRARY_PATH=$ LD_LIBRARY_PATH :/oracle/app/oracle/product/10.2.0/db_1/lib
export TNS_ADMIN=/oracle/app/oracle/product/10.2.0/db_1
export NLS_LANG="SIMPLIFIED CHINESE_CHINA.ZHS16GBK"
```

网关服务和 Oracle 客户端在同一台机器上：

```
export ORACLE_HOME=/oracle/app/oracle/product/10.2.0/client_1
export PATH=$PATH:/oracle/app/oracle/product/10.2.0/ client_1/bin
export LD_LIBRARY_PATH=$ LD_LIBRARY_PATH :/oracle/app/oracle/product/10.2.0/ client_1/lib
export TNS_ADMIN=/oracle/app/oracle/product/10.2.0/ client_1
export NLS_LANG="SIMPLIFIED CHINESE_CHINA.ZHS16GBK"
```

注意：如果 Oracle 客户端或数据库服务安装的是 11g 产品，以上方式设置后网关还无法成功加载 librdbr_ora.so，需要去 Oracle 的 lib 目录下找到 libclntsh.so.11.1 和 libnnz11.so 文件分别将其软链成 libclntsh.so.10.1 和 libnnz10.so，即可。

```
ln -s libclntsh.so.11.1 libclntsh.so.10.1
```

```
ln -s libnnz11.so libnnz10.so
```

情况二：网关服务所在的机器未安装任务 Oracle 产品

A、在 oracle 官网或其他地方下载 instantclient，包含以下 3 个包，需要注意的是以下列出的为 32 位机器上所用介质，64 位机器请下载对应的 64 位介质：

instantclient-basic-linux32-10.2.0.3-20061115.zip

instantclient-sdk-linux32-10.2.0.3-20061115.zip

instantclient-sqlplus-linux32-10.2.0.3-20061115.zip

B、unzip 到/opt/instantclient_10_2/目录

C、配置环境变量

```
export ORACLE_HOME=/opt/instantclient_10_2
export TNS_ADMIN=$ORACLE_HOME/NETWORK/ADMIN
export LD_LIBRARY_PATH=$ORACLE_HOME
export SQLPATH=$ORACLE_HOME
export NLS_LANG=AMERICAN_AMERICA.ZHS16GBK
PATH=$PATH:$ORACLE_HOME
export PATH
```

D、在/opt/instantclient_10_2/新建子目录 NETWORK/ADMIN

E、在\$ TNS_ADMIN 目录下配置 tnsnames.ora 文件

示例：

```
GATEWAY =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (ADDRESS = (PROTOCOL = TCP)(HOST = 192.168.10.253)(PORT = 1521))
    )
    (CONNECT_DATA =
```

```
(SERVICE_NAME = orcltrs)
)
)
```

F、测试连接：./sqlplus user/password@GATEWAY，如果成功则转到步骤七，如果失败则检查以上配置

G、把步骤三的环境变量添加到 Gateway 启动文件中

```
[root@localhost gwserver_x64]# vi startup.sh
#!/bin/sh

echo "TRS Gateway 4.6 版 Build4004"
echo "版权所有 (C) 2000-2011"
echo "北京拓尔思(TRS)信息技术股份有限公司"

GW_HOME=/opt/TRS/TRSGateway4.6/gwserver_x64

LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GW_HOME:$GW_HOME/hysoft:$GW_HOME/CK
M

#config Oracle
#ORACLE_HOME=
#PATH=$PATH:$ORACLE_HOME/bin
#LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$ORACLE_HOME/lib
#TNS_ADMIN=$ORACLE_HOME
#NLS_LANG="SIMPLIFIED CHINESE_CHINA.ZHS16GBK"
#NLS_LANG="SIMPLIFIED CHINESE_CHINA.AL32UTF8"
#export ORACLE_HOME PATH LD_LIBRARY_PATH TNS_ADMIN NLS_LANG
export ORACLE_HOME=/opt/instantclient_11_2
```

```
export TNS_ADMIN=$ORACLE_HOME/NETWORK/ADMIN  
  
export LD_LIBRARY_PATH=$ORACLE_HOME  
  
export SQLPATH=$ORACLE_HOME  
  
export NLS_LANG=AMERICAN_AMERICA.ZHS16GBK  
  
PATH=$PATH:$ORACLE_HOME  
  
export PATH
```

H、启动测试，如果看到红色信息(binding and listening localhost:8004 succeed!)则说明配置成功

```
[root@localhost gwserver_x64]# ./startup.sh  
  
TRS Gateway 4.6 版 Build4004  
  
版权所有 (C) 2000-2011  
  
北京拓尔思(TRS)信息技术股份有限公司  
  
欢迎使用 TrsAgent。当前注册使用截止日期：2011 年 10 月 07 日！  
  
该版本未注册使用扩展功能插件！  
  
已注册的数据库组件为：  
  
SQL Server  
  
Oracle  
  
load /opt/TRS/TRSGateway4.6/gwserver_x64/librdb_ora.so success!  
  
binding and listening localhost:8004 succeed!
```

I、注意：如果下载的 InstantClient 是 Oracle 11g 产品，以上方式设置后网关无法成功加载 librdb_ora.so，需要压缩包解压目录下找到 libclntsh.so.11.1 和 libnnz11.so 文件，分别将其软链成 libclntsh.so.10.1 和 libnnz10.so。

```
cd /opt/instantclient_11_2/  
  
ln -s libclntsh.so.11.1 libclntsh.so.10.1  
  
ln -s libnnz11.so libnnz10.so
```

3.2 DB2 客户端的配置

安装好 DB2 数据库服务器或者客户端后，设置以下环境变量：DB2DIR、DB2INSTANCE、LD_LIBRARY_PATH。例如：

```
export DB2DIR=/opt/IBM/db2/V8.1  
export DB2INSTANCE=brent  
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/opt/IBM/db2/V8.1/lib
```

配置网络：

首先需要知道服务器监听的端口，默认为 50000，可以先用 telnet 服务器 ip 50000 检查是否能够连通。

其次，采用如下命令把服务器连接方式加入到节点目录中：

```
DB2 CATALOG TCPIP NODE NODENAME REMOTE IP SERVER 50000  
DB2 CATALOG DB DBNAME AT NODE NODENAME  
DB2 TERMINATE
```

其中第一行中 IP 需要替换为 DB2 服务器 IP，nodename 可以自己设置，第二行和第一行一致。

然后就可以用 db2 list node directory 进行检查是否设置好，如果需要删除某个节点信息，可以采用

```
Db2 uncatalog node nodename
```

然后可以用 db2 命令进行测试连接了：

```
Db2 => connect to dbname user 用户名 using 密码
```

连接成功后将以上配置设置到网关服务 startup.sh 中。

3.3 MySQL 客户端的配置

Linux 网关服务默认打包 MySQL 相关驱动，无须额外配置，对于端口和字符集设置请参阅 Windows 环境。

3.4 SYBASE 客户端的配置

安装好 Sybase 数据库服务器或者客户端后，设置以下环境变量：SYBASE、SYBASE_OCS、LD_LIBRARY_PATH。例如：

```
SYBASE=/opt/sybase
```

```
SYBASE_OCS=OCS-15_0
```

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/opt/sybase/OCS-15_0/lib
```

```
export SYBASE_OCS SYBASE LD_LIBRARY_PATH
```

进入 sybase 安装目录的 bin 子目录下，执行 ./dsedit 进行服务器注册，同 Windows 环境。

测试命令为 isql -S syb -U sa -P

连接成功后将以上配置设置到网关服务 startup.sh 中。

3.5 SQL SERVER 客户端的安装和配置

SQL Server 是 Microsoft 公司的数据库产品，其未提供 Linux 环境下安装介质，因此我们只能使用第三方软件。

下载软件：

Linux 系统的 ODBC

unixODBC-2.2.8.tar.gz (<http://www.unixodbc.org>)

连接 SQLServer 或 Sybase 的驱动

freetds-0.62.4.tar.gz (<http://www.freetds.org>)

具体两个地址：

<http://www.unixodbc.org/unixODBC-2.2.12.tar.gz>

<http://ibiblio.org/pub/Linux/ALPHA/freetds/>

3.5.1 安装 unixODBC

```
tar zxvf unixODBC-2.2.12.tar.gz
```

```
cd unixODBC-2.2.12
```

```
./configure --prefix=/usr/local/unixODBC --includedir=/usr/include --libdir=/usr/lib -bindir=/usr/bin  
--sysconfdir=/etc #可以不写
```

```
make
```

```
make install
```

安装成功后，unixODBC 所需的头文件都被安装到了 /usr/include 下，编译好的库文件安装到了 /usr/lib 下，与 unixODBC 相关的可执行文件安装到了 /usr/bin 下，配置文件放到了 /etc 下。

3.5.2 安装 freetds

```
# tar vxzf freetds-0.62.4.tar.gz
```

```
# cd freetds-0.62.4
#./configure --prefix=/usr/local/freetds
--with-unixodbc=/usr/local/unixODBC --with-tdsver=8.0
# make
# make install
```

配置 freetds

```
# cd /usr/local/freetds
# vi etc/freetds.conf
```

修改以下的一段，并把;去掉

改好之后的内容:

```
[MyServer2k]
host=192.9.200.119
port=1433
tds version=8.0
```

其中的[MyServer2k]代表在客户端使用的服务器名称，host 代表 SQLServer 服务器的 IP 地址，port 代表端口，SQL Server2000 tds version 设置为 8.0，SQL Server2005 tds version 设置为 7.0。

测试连接:

```
# bin/tsql -S MyServer2k -H 192.168.0.32 -p 1433 -U sa -P password
1> use gameDB
2> go
1> select count(*) from t_ip
4
1> quit
```

以下说明这些参数的作用，-S MyServer2k 是在 freetds.conf 中定义好的，-H 192.168.0.32 服务器 IP 地址，和 freetds.conf 中的一致，如果使用了-S 参数，这个参数可以不用，-p 1433 用来指定端口是 1433 ，-U sa 用来指定连接数据库的用户名，-P password 用来指定连接数据库的密码。

配置 unixODBC

```
# cd /usr/local/unixODBC/bin
```

如果安装 UnixODBC 时写了 **-bindir=/usr/bin --sysconfdir=/etc**

则应该#cd /usr/bin

```
./ODBCConfig
```

Driver 标签页 ->Add。在出来的窗口中，Name 填 MyServer2k，Driver 填 /usr/local/freetds/lib/libtdsodbc.so（如果这个文件不存在请 find 一下），Setup 填 /usr/lib/libtdsS.so（如果不存在也请 find 一下）。确定即可（注意确定是对话框左上角的那个图标）。

选择 System DSN 标签页，点 Add，Driver 选刚才创建的 MyServer2k，点 OK。在弹出的对话框的中，name 填 hejinhu，Port 填 1433，Servername、UID、PWD 分别为 MyServer2k、用户名、密码，请如实填写。再确定即可。可以只填 Servername 和 Port

测试连接：

```
isql -v hejinhu 用户名 密码
```

或者：

```
./DataManager
```

SQL Sever 客户端字符集

查看 Linux 环境变量

```
#env
```

```
LANG = UTF-8 或者 GB2312 或者 GB18030 等
```

```
#cd /usr/local/freetds/ect
```

```
#vi freetds.conf
```

增加红线部分

```
[MyServer2k]
```

```
host=192.9.200.119
```

```
port=1433
```

```
tds version=8.0
```

```
client charset = UTF-8 或者 GB2312 或者 GB18030 等
```

注意：网关服务不推荐在 Linux 迁移 SQL Server 数据库数据。

3.6 其他 ODBC 数据源配置

Linux 所有 ODBC 数据源配置均相同，首先确认已安装 unixODBC，然后安装数据库的驱动，再使用 ODBCConfig 配置，具体参照 3.5 节。

附录4 扩展功能插件编写规范

TRIS 全文检索网关提供一种动态加载扩展功能插件的机制，支持用户外围定制功能，网关动态加载后即可使用所定制的功能。标准产品的扩展功能插件集成了 CKM 的大部分接口功能，包括：文本分类、相似性检索、文本摘要、文本分词、常识校对、规则分类、实体识别、案件名识别、数字识别、房屋信息识别、简历信息抽取、拼音检索和短语检索。

从理论上说，开发人员只要按照规定的格式编写扩展功能插件，网关均可加载使用。扩展功能插件名称可任意定义，标准产品定义为：gw_ext.dll。

头文件定义：

```
#ifndef _GW_EXT_H_
```

```
#define _GW_EXT_H_
```

```
#ifdef __cplusplus
```

```
extern "C" {
```

```
#endif
```

```
typedef void *ext_func_process;
```

```
RDB_DECLARE(ext_func_process) gw_ext_func_init(char *ext_func_string, char *task_name, long *status);
```

```
RDB_DECLARE(int) gw_ext_func_term(ext_func_process param);
```

```
RDB_DECLARE(int) gw_ext_func_process(ext_func_process param, char *keyword, char *string, int tag, char **result);
```

```
RDB_DECLARE(int) gw_ext_func_init_buffer(char **buffer);
```

```
RDB_DECLARE(int) gw_ext_func_term_buffer(char **buffer);
```

```
#ifdef __cplusplus
```

```
}
```

```
#endif
```

```
#endif
```

实现定义:

```
typedef void *ext_func_process;
```

功能: 扩展功能结构句柄, 用户自己定义结构。

```
RDB_DECLARE(ext_func_process) gw_ext_func_init(char *ext_func_string, char *task_name, long *status);
```

功能: 初始化扩展功能, 返回扩展功能句柄。

参数:

ext_func_string

[IN] 字符串指针, 网关调用时输入。

网关传入的内容示例:

```
<ext_func_list>
```

```
<ext_func_field>
```

```
<name>摘要</name><value>21</value><is_use>1</is_use><path></path>
```

```
</ext_func_field>
```

```
<ext_func_field>
```

```
<name>关键字</name><value>22</value><is_use>0</is_use><path></path>
```

```
</ext_func_field>
```

```
.....
```

```
</ext_func_list>
```

实际使用时先转换成链表, 链表中存储结构为:

指针中存储的数据结构为:

```
typedef struct ext_func_field_t
```

```
{
```

```
    char *name;
```

```
    int    value;
```

```
int is_use;  
  
char *path;  
  
}ext_func_field_t;
```

注意：当存在多个扩展功能，用户需要遍历加载 is_use 值为 1 的扩展功能初始化函数。

task_name

[IN]字符串指针，网关调用时会输入任务名。

status

[IN][OUT]错误号，网关调用时输入，初始化失败后将错误号输出。

```
RDB_DECLARE(int) gw_ext_func_term(ext_func_process param);
```

功能：释放扩展功能句柄。

参数

param

[IN]扩展功能句柄。

```
RDB_DECLARE(int) gw_ext_func_process(ext_func_process param, char *keyword, char *string, int tag, char **result);
```

功能：扩展功能实现函数。网关调用前先调 gw_ext_func_init_buffer，申请空间存放结果。gw_ext_func_process 处理过程中结果空间不够，可动态增加空间大小。结束后网关调用 gw_ext_func_term_buffer 释放存放结果空间。

参数：

param

[IN]扩展功能句柄。

keyword

[IN]当前记录主键，可以为空。其中 51-150, keyword 赋值为空, 151-250, 网关内部会遍历出 keyword 的真实值。

string

[IN]扩展功能处理的内存文本。

tag

[IN]扩展功能标识号。取值范围：51-250。其中 51-150，keyword 赋值为空，151-250，网关内部会遍历出 keyword 的真实值。用户合理选择扩展标识号。

result

[IN][OUT]存放结果字符串二级指针（使用二级指针目的是允许动态调节空间大小）

```
RDB_DECLARE(int) gw_ext_func_init_buffer(char **buffer);
```

功能：申请结果空间。网关传入为空指针，初始化时可先不进行空间申请，置*buffer = 0 即可，实际使用时再分配。

```
RDB_DECLARE(int) gw_ext_func_term_buffer(char **buffer);
```

功能：释放结果空间。释放后置*buffer = 0

附录5 正则表达式

5.1 正则表达式部分语法

- ^ 匹配行的开始
- \$ 匹配行的结束
- .
- 匹配任意字符
- * 重复零次或更多,例如 a*b 可匹配 b,ab,aab,aaaaaab
- + 重复一次以上,例如 a+b 可匹配 ab,aab,aaaaaab。但不能匹配 b 了
- ? 零次或一次,例如 ca?b 匹配 cb,cab 但不匹配 caab
- a{n} 匹配字符'a'重复 n 次
- a{n,} 字符 a 重复 n 次以上 (含 n 次)
- a{n,m} a 重复 n 到 m 次 (含)
- | 或操作,例如 ab(d|ef)匹配 abd 或 abef
- [] 字符集操作,例如[abc]将匹配任何单个字符'a', 'b', 'c'
- [a-d] 表示 a、b、c、d
- ^ 否操作,例如[^a-c]表示 a 至 c 之外的所有字符

5.2 正则匹配举例

- ^d+\$ 匹配非负整数 (正整数 + 0)
- ^[0-9]*[1-9][0-9]*\$ 匹配正整数
- ^((-d+)|(0+))\$ 匹配非正整数 (负整数 + 0)
- ^-[0-9]*[1-9][0-9]*\$ 匹配负整数
- ^-?d+\$ 匹配整数
- ^d+(\.d+)?\$ 匹配非负浮点数 (正浮点数 + 0)
- ^(((0-9)+\.0-9)*[1-9][0-9]*|((0-9)*[1-9][0-9]*\.[0-9]+)|((0-9)*[1-9][0-9]*))\$ 匹配正浮点数
- ^((-d+(\.d+)?|(0+(\.0+)?))\$ 匹配非正浮点数 (负浮点数 + 0)
- ^-(((0-9)+\.0-9)*[1-9][0-9]*|((0-9)*[1-9][0-9]*\.[0-9]+)|((0-9)*[1-9][0-9]*))\$ 匹配负浮点数

`^(?\\d+)(\\.\\d+)?$` 匹配浮点数

`^[A-Za-z]+$` 匹配由 26 个英文字母组成的字符串

`^[A-Z]+$` 匹配由 26 个英文字母的大写组成的字符串

`^[a-z]+$` 匹配由 26 个英文字母的小写组成的字符串

`^[A-Za-z0-9]+$` 匹配由数字和 26 个英文字母组成的字符串

`^\\w+$` 匹配由数字、26 个英文字母或者下划线组成的字符串

`^http[s]{0,1}:\\/\\.+` 匹配网址

附录6 常见问题解答

6.1 创建任务

6.1.1 如何创建 RDBMS 视图到 TRIS 数据库的增量更新任务

创建 RDBMS 视图到 TRIS 数据库的增量更新任务可以根据 RDBMS 视图的不同划分为以下两种情况：

- 基于多表[一对一]的 RDBMS 视图到 TRIS 数据库的增量更新任务

在该种情况下，创建任务时需保证 RDBMS 视图存在主键，任务创建成功后，不再需要其他的手工工作。

Gateway 会在 RDBMS 数据库中根据指定的视图主表的主键把主键的内容存入临时表中，等到任务运行的时候，会根据临时表的主键到视图中取出相应的行，如果设置了根据主键合并则把视图中对应于相同主键的多条记录合并为 TRIS 库中单一的一条记录。

- 基于多表[含一对多]的 RDBMS 视图到 TRIS 数据库的增量更新任务

目前尚未提供该种视图到 TRIS 数据库增量更新任务的解决方案。

6.1.2 如何创建字符集转换的任务

目前支持的字符集包括 GB18030 和 UTF-8 相互转化，以及多种其他字符集向 GB18030 和 UTF-8 的单向转化，具体设置请参考高级设置中字符集设置。

6.2 运行任务

6.2.1 运行大数据量的 RDBMS 任务时有什么需要注意的

运行大数据量的 RDBMS 任务（一次迁移数据>100,000 条）时，用户需要监控：

- RDBMS 的数据库表空间是否足够
- RDBMS 的 LOG 文件空间是否足够

RDBMS 的数据库表空间或 LOG 文件空间不足可能导致本应用程序运行过程中出现长时间无工作进度，在系统中所占 CPU 为 0%的情况。LOG 文件空间不足的情况更为普遍，尤其是 SQL SERVER 数据库是更是如此。

- 安装本应用程序的机器上的指定文件目录是否足够

该种情况在 RDBMS=>TRIS 且 TRIS 目标字段中包含大量 DOCUMENT/BIT 类型字段时尤其需要注意，通常在此种情况下，系统临时目录所在光驱需要保证有>500M 的空间。

- 如果任务运行的结果与预期的结果不一致：

首先，请查看该任务的工作日志，确定任务出错的具体原因；

其次，请查看联机帮助，确定任务创建的是否符合要求；

最后，仍不能解决问题时，请与我们联系。

6.3 FAQ

6.3.1 全部更新任务、增量更新任务的区别

定义：

1、全部更新任务，是从设定的关系型数据库视图或者表中获取数据，直接追加到 TRIS 数据库集群，或者清空目标数据库再追加。

2、增量更新任务，是依靠在关系型数据库中创建的临时表、触发器等结构，把用户指定表或者视图中发生增删改等操作的主键缓存到临时表中，然后根据主键关联临时表与视图或者表，获得数据，同步到 TRIS 数据库集群。

错误处理逻辑：

1、全部更新任务没有错误恢复机制，出错后需要人工参与继续任务的运行。

2、增量更新存在错误恢复，一般情况下，即使出现异常，也无须人工参与，任务在后续运行中会试着恢复上次运行错误造成的数据同步不正确。

RDB 错误处理：

1、全部更新，连接 RDB 数据库异常，或者获取记录过程中 RDB 连接异常，任务停止。如果整个运行过程中连接正常，获取记录时出现错误，网关会试着先获取该记录主键，正常获取则输出主键，任务继续运行。获取失败，任务将停止运行。

2、增量更新，连接 RDB 数据库异常，或者获取记录过程中 RDB 连接异常，任务停止，临时表记录不删除。如果整个运行过程中连接正常，获取记录时出现错误，网关会试着先获取该记录主键，正常获取则输出主键，任务继续运行。获取主键失败，任务将停止运行，临时表记录不删除。

TRIS 错误处理：

1、全部更新，连接 TRIS 数据库异常，任务停止运行。连接正常装库失败，任务停止运行。

2、增量更新，连接 TRIS 数据库异常，进入错误处理，缓存删除文件，装库文件，任务继续运行，最后删除临时表记录。连接正常装库失败，同上。

IO 错误处理：

1、全部更新，写 TRIS 格式文件错误，任务停止。

- 2、增量更新，写 TRS 格式文件错误，任务停止，临时表记录不删除。

6.3.2 TRS 全文检索网关功能限制

网关：

1、建立关系型数据库视图到 TRS 文件或数据库的增量更新任务，如果视图是多表联立，非主表字段发生变化，网关无法进行同步。用户可根据项目特点，适当修改触发器得以解决。

2、不能建立多个任务同时同步到 TRS Server 的相同数据库中的应用，任务之间可能造成互相删除数据，结果不一致。遇到这种，要求充分考虑应用，通过建立视图，或者增加字段注明数据来源等方式解决。

3、关系型数据库到 TRS 文件的增量更新任务，对于表或视图只有增加记录逻辑，网关运行结果正确。对于还包含删除或更新记录，存在问题。建议用户不要建立此类任务。

4、网关建立增量更新任务失败后，关系型数据库中可能已生成的临时表和触发器，用户需要进行检查，手工予以删除。

5、对于数据库中存储的大文件（几百 M 到上 G），网关运行时如申请不到足够的内存，会导致该条记录获取失败或任务执行失败。

6、网关不保证手工创建和修改的配置文件能正确加载，如果有需求，需要与研发部门联系。

7、网关建议任务名称按照 C 语言变量命名规范来命名，否则在创建任务时有可能发生错误。

8、创建目录时（比如创建 TRS 格式文件目录），需要输入全路径创建，且只能逐级创建。

关系型数据库：

1、网关任务设置分段处理，如果设置条件“order by 某个字段”，对于 SQL Server 和 DB2 任务执行会失败，原因是 SQL Server 和 DB2 不支持此类查询语句“select min(主键), max(主键) from table order by 某个字段”。此类分段任务用户可以取消分段设置，或者取消“order by 某个字段”的条件设置。用户也可以通过建立视图，在视图中指定 order by 条件。

2、Sybase 数据库，当字符型字段长度大于 255 时，网关条件设置时不能使用 order by 语句。如任务执行中报错，请检查是否设置了 order by 条件。

3、Sybase 数据库，建立触发器时，同一类型的触发器会覆盖旧的触发器，因此网关在建立增量更新任务时，会判断主表上是否存在触发器，如果存在建立任务失败。

4、增量更新任务，网关通过建立临时表，来记录增加、修改和删除的记录的主键，其中临时表主键是一个自增长字段。对于 Sybase 数据库，用户请尽量避免服务器异常宕机和使用 shutdown with nowait 操作，这两种情况下有可能造成自增长字段（SEQ_ID）取值变得异常大，网关无法处理。遇到此类问题的一种处理办法是删除临时表，重新建立。

5、MySQL 数据库，只支持一个表上建立一类相同操作的触发器，导致同一表上的增量更新任务只能建立一次。

6、MySQL 数据库在获取大字段（如 longblob、longtext）时，需要预先申请足够的内存，如申请内存失败会导致任务终止。遇到此类问题请与 TRS 公司联系，协助解决。

TRIS Server:

1、RDB 到 TRIS 数据库的任务，做任务对应时，需要了解 TRIS Server 每个字段类型的属性。如最大长度或最大精度设置。CHAR 类型单个字段值的最大长度是 255 个字节，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。NUMBER 类型精度设置不同，可以保存的值范围不同。详细内容请参阅《TRIS DataBase 全文数据库系统使用指南》。

其他:

1、如果网关采用 ctr 控制文件控制入库，且 ctr 文件中设置了引导符前缀，但是在网关的界面中没有设置引导符前缀，该引导符前缀不能生效。

2、对一个全部更新任务进行快速创建增量，如果全部更新任务名称过长，可在属性中修改名称后再快速创建。

6.3.3 FAQ

1、问：任务执行结束后，网关 temp 目录下有多余文件夹未删除，且含有装载文件，这是什么原因？

答：temp 目录下的文件和文件夹是网关运行过程中产生的中间文件，记录着从关系型数据库获取的数据，生成 TRIS 格式文件的信息。任务结束后，temp 目录下不空，可能的原因是任务执行过程中装载 TRIS 数据库失败，网关对装库文件做的备份。此时用户可根据任务的类型做如下处理：如果是全部更新，用户可以直接删除以该任务命名的文件夹，检测连接 TRIS 数据库正常后，重新执行任务即可。如果是增量更新任务，用户不必删除此文件夹，下次执行任务，网关会自动检测连接 TRIS 数据库是否正常，如果正常将进行错误恢复，装载这些文件，异常则继续备份装库文件。

2、问：全部更新分段任务，报错“分段的最小值为：m；最大值为：n 每段处理的最多数目为：l 本段从主键值：a 开始，运行错误！”，该如何处理？

答：全部更新分段任务，某段任务执行失败后，则报以上错误，任务运行中止。用户可根据该提示信息，入 TRIS 数据库删除主键大于等于 a 的记录，然后设置条件，形如“主键>=a”，再次运行网关即可。

3、问：网关错误日志中包含 17007 错误，什么原因？

答：可能的原因是附件中的文件不存在。例如 TRIS 格式文件为：

```
<REC>
```

```
<URLBODY>=@1205386087_2392_1\1_2.txt
```

而附件中 1_2.txt 文件不存在。为什么会不存在？有可能是该文件内容是脚本，当成病毒被杀了，解决办法是去掉杀毒软件对 temp 文件夹的实时跟踪功能。

4、问：网关注册了 MySQL 数据库组件，机器上也安装了 MySQL 的客户端，但图标无法显示。什么原因？（TRIS Gateway4.6 之前版本存在这个问题）

答：首先检查客户端工具安装目录是否写入系统环境变量中，没有即增加。再检查安装目录下 libMySQL.dll 文件是否存在，如果存在还不能显示图标，请从数据库安装目录下拷贝该文件做替换，即可。

5、网关 Linux 版本在取 SQL Server 中 smallmoney 字段和 money 字段最小边界值存在问题，什么原因？

答：请检查 Linux 下 SQL Server 客户端，使用 isql 查询确认结果。该问题一般是由 SQL Server 客户端(freeTDS)引起的。

6、网关服务所在的机器 Oracle 成功配置了本地 Net 服务名，使用 Oracle 自带工具也能正常连接，但网关连接时为什么总是报出“Log: 登录源数据库错误！ORA-12154: TNS: 无法解析指定的连接标识符”？

答：请检查网关的安装目录是否包含一些特殊字符，比如括号（）之类的东西，Oracle 驱动对于特殊字符的目录会无法识别，导致连接服务失败。建议用户安装的目录尽量只以字母和数组成。

7、Oracle 数据库中的表没有 ROWID 字段，为什么网关建任务时，获取的的字段列表，多出 ROWID？

答：Oracle 数据库的表有些是在建立时没有创建主键字段的，或者根本就无法选择一个可以充当主键的字段。应用中多次遇到这样的情况，需要将多个结构相似但无主键的表或者有主键合并之后主键又变得不唯一的视图数据迁移到同一个 TRS 数据库中，此时通过增加 Oracle 内部 ROWID 作为他们的主键，标识记录的唯一性。用户在迁移单个表时，如不需要，字段对应时无须对应 ROWID，则任务运行时将不迁移该字段值。

8、网关对表或视图建增量更新任务失败，什么原因？

答：网关在建立增量更新任务时，会附带创建一个临时表，三个触发器(Oracle 数据库还会创建一个序列对象)。触发器将表或视图主表记录的增删改情况记录到临时表中，反应到临时表中即为记录主键的值。网关创建增量任务失败，要根据网关返回的错误信息，以及网关日志信息定位原因。比如常见的因为权限原因无法创建上述对象而失败，或者因为拼接的创建临时表语句（如 SQL Server 下选择定义为 nvarchar(-1)的字段为主键），触发器语句错误等原因失败。此时就需要用户适当授权和调整表或视图的定义。

9、网关二级集群是否支持虚拟视图与 TRS Server 同组？

答：网关 4.6 集群配置中，不支持虚拟视图与 TRS Server 同组的情况，如果在配置中遇到，解析视图时会给出提示信息。