

ora2udw:把 Oracle 表格和数据导入到 UDW

一、 ora2udw 简介

ora2udw 是参考 <https://ora2pg.darold.net/>实现的把 oracle 的表格和数据导入到 udw 的工具。

二、 ora2udw 的安装

2.1 下载地址:

<http://ora2udw.ufile.ucloud.com.cn/ora2udw.tar>

2.2 安装

ora2udw 安装依赖:

- perl (perl 5.10 及以上版本)
- perl-devel
- postgresql-devel
- perl-Test-Simple
- ExtUtils-MakeMaker
- Time-HiRes
- Oracle packages: instant-client, SDK and SQLplus
- DBI
- DBD::Oracle
- DBD::Pg

首先我们需要安装好 perl 5.10 或以上版本，然后解压下载的 ora2udw.tar，进入 ora2udw 目录，tools 里面包含了 ora2udw 依赖工具。

如果是 centos 系统，并且已经有 perl 环境了，可以执行 ora2udw 目录下的 install.sh 脚本可以帮助我们自动完成 ora2udw 的安装。

如果是其他的 Linux 系统

1. 安装 perl、perl-devel、postgresql-devel
2. 执行 ora2udw/tools 目录下 install.sh 脚本安装其他依赖

三、 ora2udw 的使用

3.1 ora2udw 的使用简介

使用 ora2udw 可以选择在配置文件里配置源端 oracle 的信息、目的端 udw 的信息以及一些其他配置项，也可以通过执行命令 ora2pg 时指定参数来实现。

ora2udw 的默认配置文件是在/etc/ora2pg 目录下的 ora2pg.conf.dist。

输入 ora2pg --help 我们就可以看到该工具的使用说明：

```
[root@ora2pg ~]# ora2pg --help
Usage: ora2pg [-dhpqv --estimate_cost --dump_as_html] [--option value]
-a | --allow str : coma separated list of objects to allow from export.
                  Can be used with SHOW_COLUMN too.
-b | --basedir dir: Used to set the default output directory, where files
                  resulting from exports will be stored.
-c | --conf file : Used to set an alternate configuration file than the
                  default /etc/ora2pg/ora2pg.conf.
-d | --debug      : Enable verbose output.
-e | --exclude str: coma separated list of objects to exclude from export.
                  Can be used with SHOW_COLUMN too.
-h | --help      : Print this short help.
-i | --input file : File containing Oracle PL/SQL code to convert with
                  no Oracle database connection initiated.
-j | --jobs num   : number of parallel process to send data to PostgreSQL.
-J | --copies num : number of parallel connection to extract data from Oracle.
-l | --log file   : Used to set a log file. Default is stdout.
-L | --limit num  : number of tuples extracted from Oracle and stored in
                  memory before writing, default: 10000.
-m | --mysql      : Export a MySQL database instead of an Oracle schema.
-n | --namespace schema : Used to set the Oracle schema to extract from.
-o | --out file   : Used to set the path to the output file where SQL will
                  be written. Default: output.sql in running directory.
-p | --plsql      : Enable PLSQL to PLPSQL code conversion.
-P | --parallel num: Number of parallel tables to extract at the same time.
-q | --quiet      : disable progress bar.
-s | --source DSN : Allow to set the Oracle DBI datasource.
-t | --type export: Used to set the export type. It will override the one
                  given in the configuration file (TYPE).
-T | --temp_dir DIR: use it to set a distinct temporary directory when two
```

其中，比较常用的参数有：

-c: 指定配置文件

-d: 显示详细的输出信息

-j: 导入数据的并发数

-o: 指定输出文件的位置（默认在执行的目录下生成 output.sql）

例如，ora2udw.conf.sample 配置文件如下所示，指定 TYPE 为 table。

```
ORACLE_HOME /usr/lib/oracle/11.2/client64
ORACLE_DSN dbi:Oracle:host=10.9.90.52;sid=orcl
ORACLE_USER oracle_user
ORACLE_PWD oracle_password
TYPE TABLE
NLS_LANG AMERICAN_AMERICA.UTF8
PG_DSN dbi:Pg:dbname=dev;host=10.9.185.79;port=5432
PG_USER pg_user
PG_PWD pg_password
OUTPUT output.sql
```

ora2udw.conf.sample（注意把 ORACLE 和 PG 的相关配置修改自己的 ORACLE 和 PG 地址）配置完成之后：

- 1) 执行 `ora2udw -c ora2udw.conf.sample` 导出表结构。导出的表结构将输出到 `output.sql` 文件中。
- 2) 使用 `psql` 连接 `udw`，连接命令如下：
`psql -h hostname -U username -d dbname -p port`
连接成功之后执行 `\i /root/output.sql (output.sql 的路径)`，创建表。
- 3) 修改 `ora2udw.conf.sample` 配置文件，指定 TYPE 为 DATA，然后执行 `ora2udw -c ora2udw.conf.sample` 导入数据到 `udw`。

配置说明：

ORACLE_DSN	Oracle 数据源信息
ORACLE_USER	Oracle 用户名
ORACLE_PWD	Oracle 用户密码
TYPE	导出类型
PG_DSN	udw 数据源信息

PG_USER	udw 用户名
PG_PWD	udw 用户密码
OUTPUT	输出的 sql 脚本文件名

3.2 ora2udw 使用说明

3.2.1 支持的导出对象

这是允许导出的不同的格式，默认是TABLE:

TABLE: 提取所有包括索引、主键、唯一键、外键和检查约束的表。

VIEW: 提取视图。

FUNCTION: 提取函数。

DATA: 提取数据，生成INSERT语句。

COPY: 提取数据，生成 COPY 语句。

PARTITION: 提取范围和列表分区。

TYPE: 提取 Oracle 用户自定义的格式。

3.2.2 限制

ora2Pg 尽力将 Oracle 数据库转换到 udw 中，但是仍需一部分的手动工作。

Oracle 特定的 PL/SQL 代码生成函数时必须进行审查，以便匹配 udw 的语法。