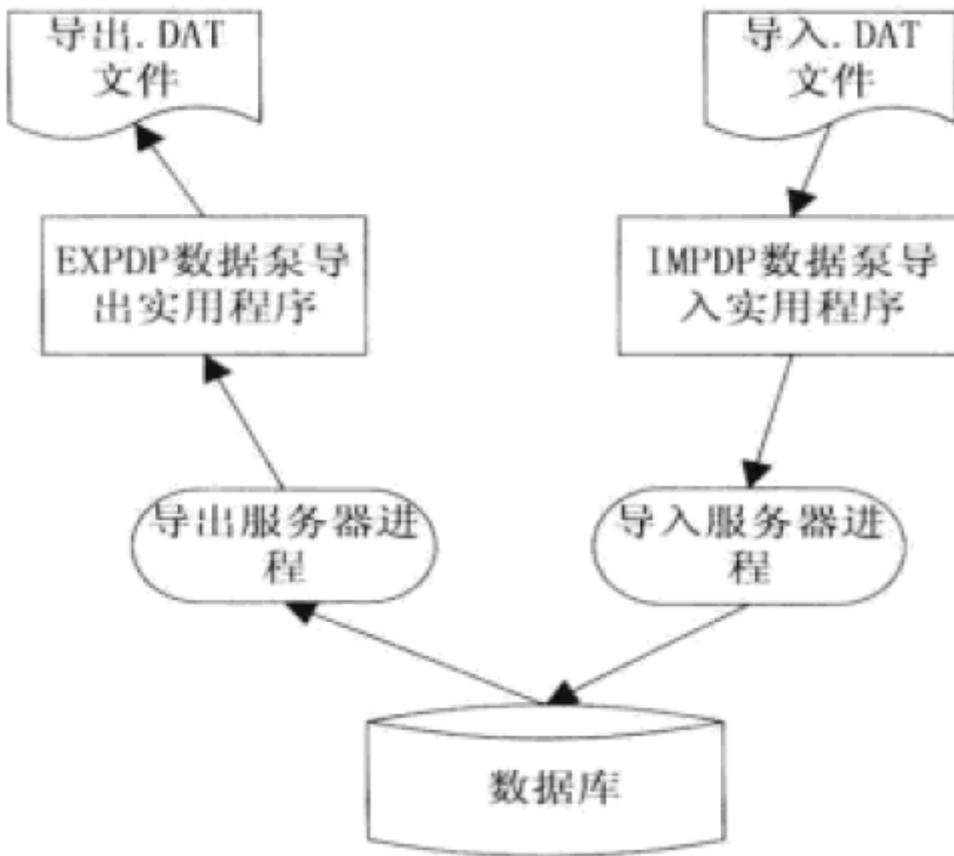


# 1 数据泵EXPDP和IMPDP技术

## 1.1 数据泵技术概览

在数据泵导出导入技术中，涉及导入程序expdp和导出程序impdp，当启动数据泵导入或导出程序时，在数据库服务器端启动相应的服务器进程，完成数据的导入和导出任务，所以也称数据泵技术是基于Oracle数据库服务器的，导入及导出的数据文件也保存在数据库服务器端。数据泵导入导出结构图如下：



## 1.2 数据泵技术优点

Oracle 11g及以上版本提供了数据泵技术，可以用来进行数据库的逻辑备份和恢复，Oracle建议使用数据泵来代替EXP和IMP技术。数据泵技术基于Oracle数据库服务器，导入导出的数据文件也保存在数据库服务器端，而EXP/IMP基于客户机运行。利用数据泵技术，Oracle要求必须使用目录对象，以防止用户误操作服务器上特定目录下的操作系统文件。

使用数据泵技术的优点有：

- 断点记录：在数据泵运行过程中，主表（Master Table）起到了关键作用，在MT表中，保存了整个数据泵运行过程中的相关信息，包括要处理的所有对象信息、正在处理的对象信息等。启动数据泵时，如果我们同时指定了任务名称，则该MT表的名称等于任务名称，如果用户下面不存在这个名称的对象，则失败。如果没有指定任务名称，数据泵自动为MT表生成一个名称，因为MT表的存在，可以重新启动被终止的任务；
- 可以只处理某些对象，或者不处理某些对象，或者只处理某些对象下面满足条件的数据等；
- 不实际导出的情况下，估计整个导出工作需要占用的磁盘空间；
- 通过数据库链接dblink，将远程数据库导出到转储文件；
- 通过数据库链接dblink，直接将远程的、位于其他主机的数据库里的数据导入到当前数据库中，从而实现跨平台的数据迁移；
- 在导入时，可以修改导入数据所在的schema名称，表空间名称以及数据文件名称；
- 通过采样，导出部分数据；
- 只导出元数据，不导出实际的数据，可以知道是否进行压缩；
- 可以进行并行操作；

### 1.3 数据泵的目录对象

#### 1) 查询数据字典dba\_directories

```
SQL> SELECT * FROM DbA_Directories WHERE directory_name='DATA_PUMP_DIR';
```

OWNER	DIRECTORY_NAME	DIRECTORY_PATH	ORIGIN_CON_ID
SYS	DATA_PUMP_DIR	/u01/app/oracle/admin/orcl/dpdump/	0

#### 2) 向hr用户授权并创建目录对象

```
SQL> conn sys/system@orcl as sysdba
```

```
Connected.
```

```
SQL> grant create any directory to hr;
```

```
Grant succeeded.
```

```
SQL> conn hr/hr@orcl
```

```
Connected.
```

```
SQL> create directory hr_dump as '/home/oracle/hrdump';
```

```
Directory created.
```

```
SQL> host mkdir /home/oracle/hrdump
```

## 1.4 数据泵和EXP/IMP比较

Oracle数据泵技术是对传统的EXP和IMP的扩展，使得在数据库服务器端快速的移动数据，二者的主要区别为：

- 数据泵技术比传统的EXP/IMP技术更快速的移动大量数据，因为数据泵技术采用并行流技术实现快速的并行处理；
- 数据泵技术基于数据库服务器，在启动数据泵导入导出程序时在数据库服务器端产生服务器进程负责备份或导入数据，并且将备份的数据备份在数据库服务器端。而且服务器进程与EXPDP客户机建立的会话无关；
- 传统的EXP/IMP是类似于普通的用户进程，执行类似select、insert、create等的SQL语句，而数据泵技术类似于启动作业的控制进程，不但启动客户端进程建立会话，还控制整个导入或导出过程，如重启作业；
- 使用传统的EXP/IMP程序导出的数据格式与数据泵技术导出的数据格式不兼容；
- 数据泵技术与传统的导入导出程序不同，它使用目录和目录对象存储数据泵导出文件，使用数据泵导出数据前必须先创建目录对象，否则无法使用数据泵导入和导出作业；

## 2 数据泵导出EXPDP

数据泵提供了不同的模式可以导出数据库的不同部分，这些模式分别为：

- 全库导出模式；
- 用户模式；
- 表模式；
- 表空间模式；
- 可传输的表空间模式；

### 2.1 EXPDP参数

```
[oracle@strong ~]$ expdp help=y
```

## 2.2 导出整库

使用system登录数据库，限制备份的数据文件的大小为200M，一旦备份数据文件满，则自动创建一个新的备份文件，使用替换变量%U实现。

```
[oracle@strong ~]$ expdp system/system@orcl dumpfile=hr_dump:db_%u.dat  
filesize=200m job_name=db full=y
```

注：一些系统模式不能被导出，因为它们不是用户模式，它们包含Oracle管理的数据和元数据，系统模式中不能导出的有SYS、ORDSYS和MDSYS。

## 2.3 导出一个模式

```
[oracle@strong ~]$ expdp system/system@orcl schemas=hr  
dumpfile=hr_dump:schema_by_system.dat
```

```
[oracle@strong ~]$ expdp hr/hr@orcl directory=hr_dump  
dumpfile=schema_by_hr.dat
```

## 2.4 导出特定的表

```
[oracle@strong ~]$ expdp system/system@orcl  
tables=hr.employees,hr.jobs,hr.departments directory=hr_dump  
dumpfile=table_by_system.dat
```

```
[oracle@strong ~]$ expdp hr/hr@orcl tables=hr.employees,hr.jobs,hr.departments  
directory=hr_dump dumpfile=table_by_hr.dat
```

## 2.5 导出指定的表空间

```
[oracle@strong ~]$ expdp system/system@orcl tablespaces=sysaux parallel=2  
directory=hr_dump dumpfile=tablespace_by_system.dat
```

## 2.6 只导出数据行

通过content参数，可以指定导出的内容：

- ALL：导出表数据和元数据；
- DATA\_ONLY：只导出表行数据；
- METADATA\_ONLY：只导出元数据，即对象的定义

```
[oracle@strong ~]$ expdp system/system@orcl schemas=hr content=data_only  
directory=hr_dump dumpfile=schema_by_system_dataonly.dat
```

## 2.7 只计算导出作业所需空间

```
[oracle@strong ~]$ expdp system/system@orcl full=y estimate_only=y
```

# 3 数据泵导入IMPDP

## 3.1 IMPDP参数

```
[oracle@strong ~]$ impdp help=y
```

## 3.2 导入整库

导入整库至少需要两个参数，一个是FULL，设置FULL=Y，说明是导入全库，一个是DUMPFILE，说明要导入的备份文件的目录和名称。

```
[oracle@strong ~]$ impdp system/system@orcl full=y directory=hr_dump  
dumpfile=db_01.dat
```

## 3.3 导入表空间

```
[oracle@strong ~]$ impdp system/system@orcl tablespaces=sysaux  
directory=hr_dump dumpfile=tablespace_by_system.dat
```

table\_exists\_action=replace

### 3.4 导入指定的表

```
[oracle@strong ~]$ impdp hr/hr@orcl tables=employees directory=hr_dump  
dumpfile=table_by_hr.dat table_exists_action=replace
```

### 3.5 导入指定的数据库对象

```
[oracle@strong ~]$ impdp system/system@orcl tables=hr.employees  
dumpfile=hr_dump:schema_by_system.dat include=table_data
```

### 3.6 重新映射模式

```
[oracle@strong ~]$ impdp system/system@orcl  
dumpfile=hr_dump:schema_by_system.dat remap_schema=hr:t2
```

### 3.7 提取SQL脚本

```
[oracle@strong ~]$ impdp system/system@orcl  
dumpfile=hr_dump:schema_by_hr.dat directory=hr_dump sqlfile=schema_by_hr.sql
```

## 4 传输数据

### 4.1 迁移过程

五步走（源、目标端字符集相同）：

- 1) 源端表空间read write
- 2) 目标端建立用户
- 3) 检查自包含
- 4) 源端表空间设为read only, expdp导出元数据, 并scp元数据和数据文件到目标主机
- 5) impdp目标主句, 登录并检查是否成功

具体参照官方网站