



Oracle Export und Oracle Data Pump

Ein Vergleich



Lange haben die Anwender darauf gewartet, ihre Exports und Imports zu beschleunigen, ab der Version 10g hat Oracle sich dieses Themas angenommen und ein neues Konzept entwickelt.

Basierend auf dem Advanced Queuing Feature wurde der Export ca. 60% und der Import bis Faktor 20 (lt. Oracle) beschleunigt.

Dieser Vortrag soll sowohl die Unterschiede zum alten Export/Import wie auch die neuen Funktionen des DATA PUMP erklären.

➤ Data Pump – EXP / IMP Unterschiede



- ▶ Export wird immer auf dem Server und nicht mehr auf dem Client durchgeführt
- ▶ Bessere Performance
- ▶ Man kann mit einer Gruppe von Dateien auch parallel arbeiten
- ▶ Der COMPRESS Parameter wird nicht mehr unterstützt, Extents werden nach den normalen Speicherparametern der Tabelle angelegt
- ▶ Sollte beim Laden der Zeilen ein Fehler auftreten, wird jetzt abgebrochen
- ▶ FLASHBACK_SCN und FLASHBACK_TIME ersetzen CONSISTENT !
(siehe auch ORACLE SUPPORT Note ID: 377218.1)
- ▶ Beim Schema Import wird nun auch der Benutzer angelegt
- ▶ Logfile wird automatisch im gleichen Verzeichnis mit Namen export.log angelegt
- ▶ Beim Tabellenimport kann nun REPLACE/TRUNCATE/SKIP oder APPEND benutzt werden.
- ▶ Es können nicht einzelne Tabellen von verschiedenen Benutzern exportiert werden
- ▶ Fast alle Objekttypen können jetzt einzeln exportiert werden
- ▶ Es existiert keine benutzergeführte (interaktive) Variante nur Kommandozeile



Gemeinsamkeiten zum alten Export/Import

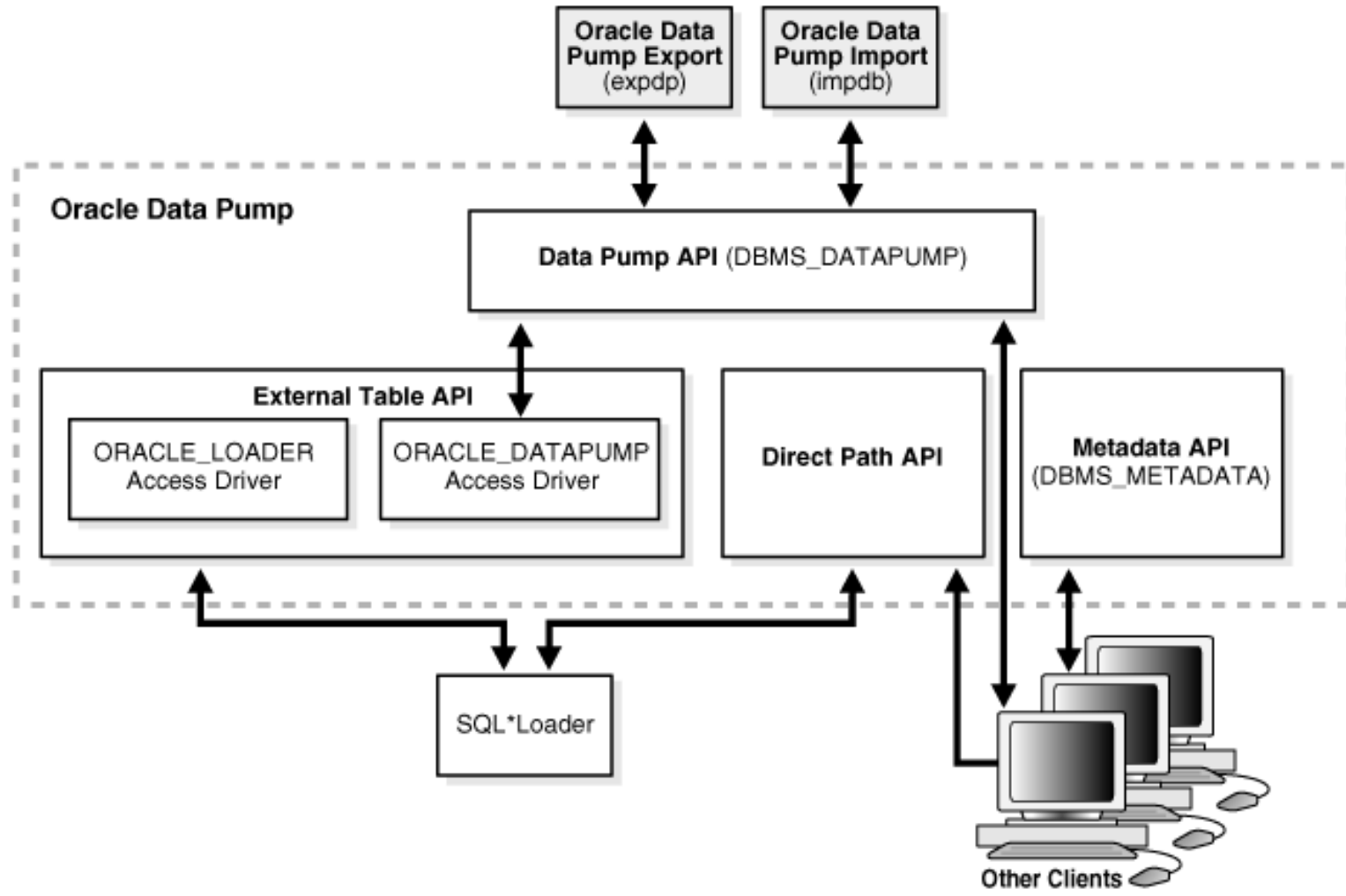
robotron[®]

- ▶ Defaultname der Exportdatei weiterhin expdat.dmp
- ▶ Dateityp weiterhin Binary
- ▶ Plattformunabhängigkeit
- ▶ Es besteht keine Möglichkeit, die Daten im ASCII-Format zu exportieren
- ▶ **EXPDP** kann nur mit **IMPDP** und **EXP** nur mit **IMP** zusammenarbeiten



- ▶ Folgende Modi können weiterhin verwendet werden:
- ▶ **Full** Export Modus (alle Daten der DB werden exportiert)
- ▶ Schema Modus (**owner**=<Schemaname>)
- ▶ Table Modus (**tables**=<Tabellenname1>, <Tabellenname2>...)
- ▶ Tablespace Modus (**Tablespace**=
- ▶ **Transportable Tablespace** Modus

Data Pump Architektur



Parameter Vergleich (I)

exp

- **buffer**
- **compress**
- **Consistent**
-
- **direct**
- **feedback**
- **file**
- **filesize**
- **flashback_scn**
- **flashback_time**
- **full**
- **grants**
- **indexes**
- **log**
- **object_consistent**
- **owner**
- **parfile**
- **query**
- **recordlength**

expdp

- **compression**
- **flashback_scn** und **flashback_time** (ab 11.2.0.2.0. auch wieder consistent)
- **access_method** (direct_path oder external_tables)
- **status**
- **Dumpfile** (ab 11.2 auch wieder File im Legacy-Mode)
- **filesize**
- **flashback_scn**
- **flashback_time**
- **full**
- **include/exclude=grant**
- **include/exclude=index**
- **Logfile** (ab 11.2 auch wieder Log im Legacy-Mode)
- **--**
- **schema**
- **parfile**
- **query**
- **automatisch**

Parameter Vergleich (II)

exp

- **resumable**
- **resumable_name**
- **resumable_timeout**
- **rows=n**
- **rows=y**
- **statistics**
- **tables**
- **tablespaces**
- **transport_tablespace**
- **triggers**
- **tts_full_check**
- **userid**
- **volsize**

expdp

- **--**
- **--**
- **--**
- **content=metadata_only**
- **content=all**
- **wird immer durchgeführt**
- **tables**
- **tablespaces**
- **transport_tablespace**
- **include/exclude=trigger**
- **transport_full_check**
- **z.B. scott/tiger**



Data Pump - neue Parameter

- **ATTACH**
- **CLUSTER**
- **DATA_OPTIONS**
- **DIRECTORY**
- **ENCRYPTION**
- **ENCRYPTION_ALGORITHM**
- **ENCRYPTION_MODE**
- **ENCRYPTION_PASSWORD**
- **ESTIMATE**
- **ESTIMATE_ONLY**
- **INCLUDE/EXCLUDE**
- **JOB_NAME**
- **NETWORK_LINK**
- **REUSE_DUMPFILES**

➤ Data Pump - Interaktive Parameter

robotron[®]

- ▶ **ADD_FILE**= [directory_object:]<datei> [, [dir:]<datei>...]
- ▶ Fügt eine weitere Export-Datei hinzu
- ▶ **CONTINUE_CLIENT**
- ▶ Wechselt vom interaktiven Modus in den Logging Modus
- ▶ **EXIT_CLIENT**
- ▶ Stoppt die Client-Session, der Job läuft aber weiter
- ▶ **HELP**
- ▶ Zeigt alle möglichen Export- bzw. Import-Parameter an
- ▶ **KILL_JOB**
- ▶ Beendet den aktuellen Job
- ▶ **PARALLEL**
- ▶ Ermöglicht die aktuelle Anzahl der Working-Prozesse zu erhöhen oder zu senken
- ▶ **START_JOB**
- ▶ Startet den aktuellen Job mit dem man verbunden ist
- ▶ **STATUS=<n>**
- ▶ Zeigt den aktuellen Status mit voraussichtlicher Endzeit an
- ▶ **STOP_JOB[=IMMEDIATE]**
- ▶ Beendet den aktuellen Job entweder sofort oder nach einem ordentlichen Shutdown



DataPump - Voraussetzungen



Die beiden neuen Tools EXPDP und IMPDP erfordern ein Directory Objekt mit einem gültigen Verzeichnis für ihre Arbeit. Den Pfad und Dateinamen wie in der früheren Version anzugeben, ist nicht mehr möglich.

Es gibt ein Standard-Directory namens DATA_PUMP_DIR, das auf das Verzeichnis %ORACLE_BASE%\admin\\dpdump verweist. Dies wird bereits bei der Installation von Oracle 10g erzeugt.

Neue Directories anlegen:

```
SQL> CREATE OR REPLACE DIRECTORY expimp_dir AS 'D:\oracle\expimp';  
SQL> CREATE OR REPLACE DIRECTORY expimp_log_dir AS 'D:\oracle\expimp\logs';
```

Überprüfen, auf welche Verzeichnisse man Rechte besitzt:

```
SQL> SELECT privilege, directory_name FROM user_tab_privs t,  
all_directories d WHERE t.table_name(+) = d.directory_name ORDER BY 2,1;
```

Lese- und Schreibrechte vergeben (Wer CREATE ANY DIRECTORY besitzt, hat sie automatisch):

```
SQL> GRANT READ,WRITE ON DIRECTORY expimp_dir TO scott;
```



DataPump-Parameter erklärt I

- **ATTACH**
- Verknüpft die Client Session mit dem interaktiven Export-Modus. So kann man auch während der Export bereits läuft, sich mit einer weiteren Session nochmals anmelden, um Wartungsaufgaben durchzuführen, wie z.B. die Anzahl der aktiven Prozesse erhöhen oder eine neue Datei anzulegen. Auch lässt sich der Export- oder Import-Vorgang unterbrechen.
- **CONTENT [ALL] DATA_ONLY | METADATA_ONLY]**
- ALL Sowohl die Daten als auch die Metadaten werden exportiert/importiert
- DATA_ONLY Nur die Nutzdaten werden exportiert/importiert
- METADATA_ONLY Nur die CREATE/GRANT-Befehle werden exportiert/importiert
- **DIRECTORY**
- Directory Objekt (das mit dem Befehl CREATE DIRECTORY erzeugt wurde) gibt an, wohin das Dump-File geschrieben wird. Hier darf kein Verzeichnis wie z.B. c:\temp angegeben werden!
- **DUMPFIL** (ab 11.2.0.2.0 auch wieder FILE)
- (Default expdat.dmp) Name der Export- bzw. Import-Datei directory_1:file1, directory_2:file2
- Es kann die Variable %U verwendet werden, wenn eine Dateigröße oder der Parameter angegeben wurde. Dann wird eine fortlaufende Nummer im Dateinamen verwendet (_%D – es wird das Tagesdatum an das Dumpfile_<Tagesdatum> angehängt)
- **PARALLEL**
- Ermöglicht die aktuelle Anzahl der Working-Prozesse zu erhöhen oder zu senken (mehr als 1 nur in Enterprise Edition möglich)
- **ESTIMATE [BLOCKS|STATISTICS]**
- Schätzt den Speicherbedarf für die zu exportierende Tabelle aufgrund der aktuellen Blockanzahl (Blocks aus dba_segments, Statistics aus dba_tables). Die Schätzung wird ausgegeben im Logfile und auf STDOUT. Beim Import nur gültig, wenn auch der NETWORK_LINK Parameter angegeben wurde.

DataPump-Parameter erklärt II

- **ESTIMATE_ONLY = [Y|N]**
- Es wird nur eine Abschätzung durchgeführt, wie viel Platz benötigt wird, ohne den eigentlichen Export auszuführen. DUMPFILERE ist hier nicht zulässig. Ist nur beim Export verfügbar.
- **EXCLUDE**
- Einige Objekte können bei Bedarf vom Export ausgeschlossen werden. Achtung: Die davon abhängigen Objekte sind dann auch ausgeschlossen.
- Beispiele:
- EXCLUDE=CONSTRAINT (Alle Constraints werden ausgeschlossen)
- EXCLUDE=REF_CONSTRAINT (FK wird ausgeschlossen)
- EXCLUDE=VIEW, PACKAGE, FUNCTION
- EXCLUDE=PROCEDURE
- EXCLUDE=PACKAGE:"LIKE '%DUMMY%'"
- EXCLUDE=STATISTICS
- Durch einen Bug in 10.1.0.2 können mit europäischen Zeichensatzinstellungen nur fehlerfreie Imports durchgeführt wenn EXCLUDE=STATISTICS verwendet wird.
- **FILESIZE** (Default 0 = unbegrenzt)
- Größe für die Exportdatei in B|K|M|G Ist nur beim Export verfügbar
- **FLASHBACK_SCN**
- Der Export kann zu einem Zeitpunkt in der Vergangenheit durchgeführt werden, wenn die SCN bekannt ist
- Beim Import nur gültig, wenn auch der NETWORK_LINK Parameter angegeben wurde.
- **FLASHBACK_TIME** (ab 11.2.0.2.0 auch wieder CONSISTENT=Y)
- Der Export kann zu einem Zeitpunkt in der Vergangenheit durchgeführt werden, hier mit Zeitangabe:
- ="TO_TIMESTAMP('21.09.04 17:00:06', 'DD.MM.YY HH24:MI:SS')"
- Beim Import nur gültig, wenn auch der NETWORK_LINK Parameter angegeben wurde.

DataPump-Parameter erklärt III

- ▶ **FULL = [Y|N]**
 - ▶ Hier kann angegeben werden, ob die gesamte Datenbank exportiert werden soll. Dazu ist die Rolle exp_full_database notwendig.
- ▶ **HELP = Y**
 - ▶ Ausgabe aller verfügbaren Parameter
- ▶ **INCLUDE**
 - ▶ Ermöglicht das Filtern der zu exportierenden Daten Es kann nur entweder die INCLUDE oder die EXCLUDE Klausel verwendet werden.
 - ▶ Beispiele:
 - ▶ INCLUDE=TABLE:\ " IN ('EMP','DEPT')\ "
 - ▶ INCLUDE=FUNCTION
 - ▶ INCLUDE=INDEX:\ "LIKE 'PK_%' \ "
 - ▶ Hinweis: Alle Tabellen müssen in einem Schema liegen. Die Syntax schema.tabelle ist nicht mehr erlaubt.
- ▶ **LOGFILE** (ab 11.2.0.2.0 auch wieder LOG)
 - ▶ Name der Logdatei,
 - ▶ Beispiel: LOGFILE= export.log oder LOGFILE=expimp_dir:export.log
- ▶ **JOB_NAME**
 - ▶ Name des Export- bzw. Import- Jobs
 - ▶ Es darf nicht der Name einer Tabelle verwendet werden, die sich im Schema des Exportierenden befindet.
- ▶ **NETWORK_LINK**
 - ▶ DB-Link Name zu einer Remote-Datenbank
 - ▶ Wenn der Parameter in Verbindung mit dem EXPDP verwendet wird, legt Oracle die Daten der Remote-Datenbank in ein lokales Dumpfile.
 - ▶ In Verbindung mit dem IMPDP wird kein Export-File erzeugt, sondern die Daten aus einer Remote-Datenbank sofort in die eigene Datenbank eingespielt. Der Parameter DIRECTORY muss trotzdem für das Logfile verwendet werden.

DataPump-Parameter erklärt IV

- **NOLOGFILE = [Y|N]**
- Die Erzeugung eines Logfiles kann unterdrückt werden.
- **PARALLEL = <n>**
- DEFAULT = 1, sollte der Anzahl der zu erzeugenden Dateien entsprechen Kann nur für die EE benutzt werden
- **PARFILE = <datei>**
- Name der Parameterdatei
- **QUERY = [schema.][table:] query** (nicht verwendbar mit CONTENT=METADATA_ONLY, ESTIMATE_ONLY, TRANSPORT_TABLESPACES)
- Damit kann noch eine WHERE-Klausel mit angegeben werden: QUERY=scott.emp:' " WHERE deptno>10 " '
- **REMAP_DATAFILE=<alter pfad/datei>:<neuer pfad/datei>**
- Wenn die Zieldatenbank nicht im gleichen Verzeichnis liegt, kann hier ein Mapping angegeben werden:
- 'c:\oracle\oradata\o10g\prod01.dbf':d:\o10g\prod01.dbf'
- Ist nur beim Import und nur mit IMP_FULL_DATABASE Rolle verfügbar.
- **REMAP_SCHEMA='<schema_alt>':'<schema_neu>'**
- Damit kann der Eigentümer eines Objekts geändert werden. Wenn das Schema nicht bereits vorhanden ist, wird es unter der Voraussetzung angelegt, dass ein CREATE USER Statement im Export-Dumpfile verfügbar ist. Es wird jedoch ein unbrauchbares Passwort angelegt, das geändert werden kann mittels: ALTER USER <user> IDENTIFIED BY <pwd>
- Entspricht im alten Import-Tool den Parametern FROMUSER=... TOUSER=. Ist nur beim Import und nur mit IMP_FULL_DATABASE Rolle verfügbar.
- **REMAP_TABLESPACE='<tbs_alt>':'<tbs_neu>'**
- Damit kann der Tablespace der Objekte beim Import verändert werden. Ist nur beim Import und nur mit IMP_FULL_DATABASE Rolle verfügbar.

DataPump-Parameter erklärt V

- **REUSE_DATAFILE = [Y|N]**
- Gibt an, ob beim Import bereits existierende Datendateien überschrieben werden können. (ab 11.2.0.2.0 wenn Dumpfile bereits vorhanden, dann automatisch reuse_dumpfiles=true)
- **SCHEMAS = schema [,schema,...]**
- Damit können alle Objekte eines Schemas exportiert/importiert werden.
- **SQLFILE = <directory>:<datei>**
- Alle DDL-Befehle, die beim Import abgesetzt würden, werden in der angegebenen Datei im Klartext eingetragen, ohne dass ein Import stattfindet (Vergleichbar mit alten Import-Parameter: show=y log=x.log)
- **STATUS = <n>**
- Anzeige von Statusinformationen auf STDOUT
- Default 0 (Sekunden), nur beim Beenden Status anzeigen
- **TABLES = [schema.]<tabelle>[:partition] [, [schema.]<tabelle>[:partition],...]**
- Hier können die Tabellennamen angegeben werden
- **TABLESPACES = <tbsname> [,<tbs..>]**
- Exportiert die Objekte eines Tablespace
- **TABLE_EXISTS_ACTION=[SKIP|APPEND|TRUNCATE|REPLACE]**
- SKIP (bereits bestehende Tabellen überspringen, entspricht der alten Einstellung IGNORE=no)
- APPEND (Datensätze an bestehende Tabellen hinten anhängen. Ergibt einen Fehler, wenn z.B. ein Verstoß gegen einen eindeutigen Schlüssel vorliegt.)
- TRUNCATE (erst auf Tabelle truncate durchführen, dann füllen. Dies ist nicht möglich, wenn die Tabelle Ziel einer referenziellen Integrität ist.)
- REPLACE (erst drop, neu erstellen, dann füllen)
- Nur beim Import verfügbar
- **VERSION= [COMPATIBLE|LATEST|<version>]**
- Damit kann man Exports auch für eine ältere Datenbank durchführen.

Datapump DBA Views

robotron[®]

- **SQL> dba_datapump_sessions**
- **SQL> dba_datapump_jobs**