Subject: Re: Batch Script connecting two databases in an Extract Tranform Load (ETL) Design Pattern

Posted by rafs on Wed, 03 Jun 2015 10:32:12 GMT

View Forum Message <> Reply to Message

Both the old and new db schemas have two tables parent --< child representing the same concept: an SID. (So, OLDDB/IDP.IDP\_ID == NEWDB/SIDS.SID). At any time, all I care is to synchronize only the parents that are enabled, and the current\_version of their respective child tables.

```
OLDDB (MSSQL)
CREATE TABLE idp {
 idp_id int PRIMARY KEY,
                                # IDP.IDP ID = SIDS.SID
                            # IDP.CURRENT_VERSION = IDP_S.IDP_VERSION
 current_version int,
                          # bitmap where 0x2 is disable flag
 idp_status,
CREATE TABLE idp s {
 idp id int,
 idp version int,
 PRIMARY KEY (idp_id, idp_version) # composite primary key
}
NEWDB (POSTGRESQL)
======
CREATE TABLE sids {
 id int PRIMARY KEY.
                               # not equivalent to idp id in OLDDB
 current signature id int,
                              # points to current verion
 sid int.
                       # IDP.IDP ID = SIDS.SID
 deployment_status_flags int,
                                # similar to OLDDB/idp.idp_status
}
CREATE TABLE signatures {
 id int PRIMARY KEY,
                               # not equivalent to anything in OLDDB
                        # foreign key, signatures.sid id = sids.id
 sid id int,
}
```

So, OLDDB and NEWDB each have a set of currently-enabled (parent, child) pairs represented by a single SID concept

(OLDDB/IDP\_ID == NEWDB/SIDS.SID). Data could be synchronized in either direction: old to new, or vice versa. When synching

from one to the other, I do set operations to yeild two sets of sids: (1) those that need to have a new child inserted, and then enable parent; (2) those that need to have the parent disabled in the target db.

In case (2), I would to do a simple update on multiple record sets, but one at a time should be fine; In case (1) I think I

need to iterate through each pair to extract, transform, then load, inserting a new child, then setting the child's id to parent.current\_version (or parent.current\_signature\_id).