



Importing and Exporting Data with MySQL

Presented by:
Sheeri K. Cabral

At ODTUG Kaleidoscope 2010



Pythian
love your data

Who I Am

- MySQL DBA
- MySQL User Group
- First Oracle ACE Director for MySQL
- Lots of community stuff (videos, blog, podcast on hiatus)

How

- Universal import/export
- Storage engines
- Batched INSERTs
- LOAD DATA INFILE

Storage Engines

- Different ways of storing data
 - Different load speeds
- How much does the technique matter?
- What else are you doing?

CSV storage engine

- Just what it sounds like
- No indexes
- No NULLs
- Drop-in table data – sort of

MyISAM

- `concurrent_insert=2`
- Key cache only, no data cache

InnoDB

- Consider the buffer pool
- Loads in memory first, then flushed to disk
- But data is then already in memory
 - But MyISAM has an fs cache

Replication Topologies

- Different table type on the slave
- Loading directly to the slave
 - Slave and master have different data!

Regular INSERT statement

```
mysql> INSERT INTO tbl (fld1, fld2, ...)
VALUES (val1, val2, ...);
```

- Import a file

```
$ mysql < file.sql
```

- with a mysql client command

```
mysql> source file.sql
```

Batched INSERT statement

```
INSERT INTO tbl (fld1, fld2...)  
VALUES (val1a, val2a...),  
(val1b, val2b), ....
```

- Faster than single inserts
- “Extended insert”
- Any storage engine

Other tricks for any storage engine

- Turn off foreign key checking
- Lock the table
- Build the indexes only at the end

Turn Off Foreign Key Checking

- Only if you are 100% sure your data is good!
 - `SET SESSION FOREIGN_KEY_CHECKS=OFF;`
 - [load data]
 - `SET SESSION FOREIGN_KEY_CHECKS=ON;`
- Does not EVER test loaded data!

Lock the Table

- Other writes/reads may interfere with your load
 - LOCK TABLES tbl WRITE;
 - [load data]
 - UNLOCK TABLES;

Build the Indexes at the End

- Disable keys, load, enable keys
- ALTER TABLE tbl DISABLE KEYS;
- [load data]
- ALTER TABLE tbl ENABLE KEYS;
- Wait as index builds all at once

Exporting with mysqldump

- mysqldump commandline export tool by default:
 - Extended insert (--skip-extended-insert)
 - Disables keys (--skip-disable-keys)
 - Locks the table for writes (--skip-add-locks)
 - Drops the table and re-adds (--no-create-info --skip-add-drop-table)
 - Sets foreign key checking off
 - Assumptive statements (--complete-insert)

INSERT DELAYED

- Online batches with MyISAM, MEMORY, ARCHIVE, BLACKHOLE
- Delay queue
- Ignored on slaves

INSERT DELAYED

- `delayed_insert_limit`
 - 100 by default
- Handlers
- Terminates after waiting `delayed_insert_limit`
 - 300 seconds by default
- `delayed_queue_size`
 - 1000 by default

Delayed Key Write

- MyISAM only
- Index buffers not flushed after each write
- Flushed (keys written) only on table close
 - FLUSH TABLES closes and re-opens
 - Only for speeding up writes
 - In general, not a good idea

DELAY_KEY_WRITE

```
CREATE TABLE ... () ENGINE=MYISAM  
                DELAY_KEY_WRITE=1;
```

- `delay_key_write` system variable
 - ON by default, honors table definition
 - OFF is never delay key writes
 - ALL is always delay key writes

LOAD DATA INFILE

- Table must already exist
- Table name does not need to relate to file name:

```
mysql> LOAD DATA INFILE  
'/tmp/importme.sql' INTO TABLE  
rental;
```

- Complement of SELECT INTO OUTFILE

LOAD DATA INFILE

- On an empty MyISAM table, non-unique indexes are created in a separate batch
- Possibly made even faster by
 - ALTER TABLE tbl DISABLE KEYS
 - LOAD DATA INFILE...
 - ALTER TABLE tbl ENABLE KEYS

LOAD DATA INFILE

LOAD DATA

[LOW_PRIORITY | CONCURRENT]

- **less/no blocking**

[LOCAL]

- **from client machine**

INFILE 'file_name'

- **can be full path**

LOAD DATA INFILE

```
LOAD DATA INFILE 'file_name'  
  [REPLACE | IGNORE]
```

- What to do if there are duplicates
- Default is error

```
  INTO TABLE tbl_name
```

LOAD DATA INFILE

```
LOAD DATA INFILE 'file_name' INTO  
TABLE tbl_name  
[CHARACTER SET charset_name]
```


LOAD DATA INFILE

```
LOAD DATA INFILE 'file_name' INTO  
TABLE tbl_name
```

```
[FIELDS
```

```
  [TERMINATED BY 'string'] (comma)
```

```
  [[OPTIONALLY] ENCLOSED BY 'char']
```

```
  [ESCAPED BY 'char']]
```

```
[LINES
```

```
  [STARTING BY 'string']
```

```
  [TERMINATED BY 'string']] (newline)
```

SELECT INTO OUTFILE

```
SELECT * INTO OUTFILE 'file_name'
```

```
[CHARACTER SET charset_name]
```

```
[FIELDS
```

```
  [TERMINATED BY 'string']
```

```
  [[OPTIONALLY] ENCLOSED BY 'char']
```

```
  [ESCAPED BY 'char']]
```

```
[LINES
```

```
  [STARTING BY 'string']
```

```
  [TERMINATED BY 'string']]
```

SELECT INTO DUMPFIELD

```
SELECT * INTO DUMPFIELD 'file_name';
```

- No other options
- Good for BLOBs

Faster inserts/imports

- Batched inserts
- Lock tables
- Turn off foreign key checking
- Build index at end of import
- `DELAY_KEY_WRITE`
- `INSERT DELAYED`

Import/export

- Import
 - LOAD DATA INFILE
 - Different storage engines (ie, CSV)
- Export
 - mysqldump
 - SELECT ... INTO OUTFILE
 - SELECT ... INTO DUMPFILE