MySQL BREAK/FIX LAB
Significant Performance Issues

Presented by: Miklos Szel, Gillian Gunson, Brian Cain

April 13, 2015

Pythian®
love your data

About Us

- Miklos Szel
  - Database Consultant - Pythian
- Gillian Gunson
  - Currently: Senior DBA - Okta
- Brian Cain
  - Currently: Lead Database Consultant - Pythian

Tutorial Agenda

1. Fix standalone MySQL instance
2. Replication issues
3. Performance issues

GETTING STARTED

Each attendee has its own instance (not shared)

we break it, you fix it :-)

One standalone MySQL instance

Several MySQL instances using MySQL sandbox

GETTING STARTED

Access:
Username/password: user-lab / pythianlab123
ssh user-lab@hostname
http://tinyurl.com/plsc2015-demo

Pre-requirement: ssh client

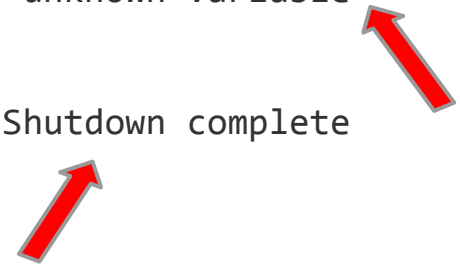
Pastebin:
http://tinyurl.com/plsc2015-pastebin

AGENDA - PART 1

- Recover a mysql instance unable to start
  - misconfiguration
  - files permission
  - corrupted files
- Connectivity issues
  - misconfiguration
  - recover password
- Learn how to troubleshoot crash
  - read error log
  - fix misconfiguration
  - tune variables

START MYSQLD

```
[root@hostdb ~]# service mysql start
Initializing MySQL database:  Installing MySQL system tables...
2014-10-30 16:29:23 10826 [ERROR] /usr/sbin/mysqld: unknown variable
    'tmpd1r=/var/tmp'
2014-10-30 16:29:23 10826 [ERROR] Aborting
2014-10-30 16:29:23 10826 [Note] /usr/sbin/mysqld: Shutdown complete


[root@hostdb ~]# !ps
ps aux | grep mysql
root      2185  0.0  0.0 103424   828 pts/2    S+   13:19   0:00 grep mysql
```



WHICH CONFIG FILE?

```
[root@hostdb ~]# grep tmpd /etc/my.cnf
[root@hostdb ~]# grep tmpd /etc/mysql/my.cnf
grep: /etc/mysql/my.cnf: No such file or directory
```

Multiple configuration file(s) ??

The easy way:

https://dev.mysql.com/doc/refman/5.6/en/option-files.html

The hard way:

strace

STRACE

```
[root@hostdb ~]# strace /usr/sbin/mysqld

[root@hostdb ~]# strace -e trace=open,stat /usr/sbin/mysqld
...
stat("/etc/my.cnf", {st_mode=S_IFREG|0644, st_size=243, ...}) = 0
open("/etc/my.cnf", O_RDONLY)           = 3
stat("/etc/mysql/my.cnf", 0x7fffea4c0d80) = -1 ENOENT (No such file or directory)
stat("/usr/etc/my.cnf", {st_mode=S_IFREG|0644, st_size=25, ...}) = 0
open("/usr/etc/my.cnf", O_RDONLY)       = 3
stat("/root/.my.cnf", {st_mode=S_IFREG|0644, st_size=33, ...}) = 0
open("/root/.my.cnf", O_RDONLY)         = 3
...
```

STRACE: MYSQLD --PRINT-DEFAULTS

```
# strace -e stat64 /usr/sbin/mysqld --print-defaults
/usr/sbin/mysqld would have been started with the following arguments:
--datadir=/var/lib/msql --innodb_data_file_path=ibdata1:18M --
    innodb_buffer_pool_size=100G --innodb_log_file_size=64M --sort_buffer_size=60M
    --tmpd1r=/var/tmp


stat64("/etc/my.cnf", 0xbfb9d750)       = -1 ENOENT (No such file or directory)
stat64("/etc/mysql/my.cnf", {st_mode=S_IFREG|0644, st_size=3564, ...}) = 0
stat64("/usr/etc/my.cnf", 0xbfb9d750)   = -1 ENOENT (No such file or directory)
stat64("/root/.my.cnf", 0xbfb9d750)     = -1 ENOENT (No such file or directory)
```

FIX TMPDIR

```
[root@hostdb ~]# cat /usr/etc/my.cnf
[mysqld]
tmpd1r=/var/tmp
```

```
[root@hostdb ~]# sed -i -e 's/tmpd1r/tmpdir/' /usr/etc/my.cnf
[root@hostdb ~]# cat /usr/etc/my.cnf
[mysqld]
tmpdir=/var/tmp
```

START AGAIN

```
[root@hostdb ~]# service mysql start
MySQL Daemon failed to start.
Starting MySQL... ERROR! The server quit without updating PID file
    (/var/lib/msql/ip-10-87-0-19.pid).
```

CHECK ERROR LOG

```
[root@hostdb ~]# tail -n 100 /var/log/mysqld.log
1141028 11:40:32 mysqld_safe Starting mysqld daemon with databases from /var/lib/msql
2014-10-28 11:40:34 0 [Warning] TIMESTAMP with implicit DEFAULT value is deprecated. Please use --
    explicit_defaults_for_timestamp server option (see documentation for more details).
2014-10-28 11:40:34 32549 [Note] Plugin 'FEDERATED' is disabled.
/usr/sbin/mysqld: Table 'mysql.plugin' doesn't exist
2014-10-28 11:40:34 32549 [ERROR] Can't open the mysql.plugin table. Please run mysql_upgrade to create
    it.
2014-10-28 11:40:34 32549 [Note] InnoDB: Using atomics to ref count buffer pool pages
2014-10-28 11:40:34 32549 [Note] InnoDB: The InnoDB memory heap is disabled
2014-10-28 11:40:34 32549 [Note] InnoDB: Mutexes and rw_locks use GCC atomic builtins
2014-10-28 11:40:34 32549 [Note] InnoDB: Memory barrier is not used
2014-10-28 11:40:34 32549 [Note] InnoDB: Compressed tables use zlib 1.2.3
2014-10-28 11:40:34 32549 [Note] InnoDB: Using Linux native AIO
2014-10-28 11:40:34 32549 [Note] InnoDB: Using CPU crc32 instructions
/usr/sbin/mysqld: Can't create/write to file '/var/tmp/ib3qf0c3' (Errcode: 13- Permission denied )
2014-10-28 11:40:34 7fda23d8b740  InnoDB: Error: unable to create temporary file; errno: 13
2014-10-28 11:40:34 32549 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-28 11:40:34 32549 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-28 11:40:34 32549 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-28 11:40:34 32549 [ERROR] Aborting
```

FIX PERMISSIONS

```
[root@hostdb ~]# ls -l /var/lib/mysql/mysql/plugin.*
-rw-rw---- 1 root root 8586 Mar 13 12:30 /var/lib/mysql/mysql/plugin.frm
-rw-rw---- 1 root root  0 Mar 13 12:30 /var/lib/mysql/mysql/plugin.MYD
-rw-rw---- 1 root root 1024 Mar 13 12:30 /var/lib/mysql/mysql/plugin.MYI


[root@hostdb ~]# chown -R mysql:mysql /var/lib/mysql/mysql/


[root@hostdb ~]# service mysql start
Starting MySQL... ERROR! The server quit without updating PID file
    (/var/lib/msql/ip-10-87-0-19.pid).
```

CHECK ERROR LOG

```
141030 16:44:55 mysqld_safe Starting mysqld daemon with databases from /var/lib/msql
2014-10-30 16:44:57 0 [Warning] TIMESTAMP with implicit DEFAULT value is deprecated. Please use --
    explicit_defaults_for_timestamp server option (see documentation for more details).
2014-10-30 16:44:57 11576 [Note] Plugin 'FEDERATED' is disabled.
/usr/sbin/mysqld: Table 'mysql.plugin' doesn't exist
2014-10-30 16:44:57 11576 [ERROR] Can't open the mysql.plugin table. Please run mysql_upgrade to create
    it.
2014-10-30 16:44:57 11576 [Note] InnoDB: Using atomics to ref count buffer pool pages
2014-10-30 16:44:57 11576 [Note] InnoDB: The InnoDB memory heap is disabled
2014-10-30 16:44:57 11576 [Note] InnoDB: Mutexes and rw_locks use GCC atomic builtins
2014-10-30 16:44:57 11576 [Note] InnoDB: Memory barrier is not used
2014-10-30 16:44:57 11576 [Note] InnoDB: Compressed tables use zlib 1.2.3
2014-10-30 16:44:57 11576 [Note] InnoDB: Using Linux native AIO
2014-10-30 16:44:57 11576 [Note] InnoDB: Using CPU crc32 instructions
^G/usr/sbin/mysqld: Can't create/write to file '/var/tmp/ibvm13fg' (Errcode: 13 - Permission denied)
2014-10-30 16:44:57 7fd985ef8740  InnoDB: Error: unable to create temporary file; errno: 13
2014-10-30 16:44:57 11576 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-30 16:44:57 11576 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-30 16:44:57 11576 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-30 16:44:57 11576 [ERROR] Aborting
```

FIX DATADIR PATH

```
[root@hostdb ~]# grep datadir /etc/my.cnf
datadir=/var/lib/msql

[root@hostdb ~]# sed -i -e 's/datadir=\/var\/lib\/msql/datadir=\/var\/lib\/mysql/'
    /etc/my.cnf

[root@hostdb ~]# grep datadir /etc/my.cnf
datadir=/var/lib/mysql
```

CHECK ERROR LOG

```
[root@hostdb~]# service mysql start
Starting MySQL... ERROR! The server quit without updating PID file (/var/lib/msql/ip-10-87-0-19.pid).


141028 11:46:25 mysqld_safe Starting mysqld daemon with databases from /var/lib/mysql
2014-10-28 11:46:26 0 [Warning] TIMESTAMP with implicit DEFAULT value is deprecated. Please use --
    explicit_defaults_for_timestamp server option (see documentation for more details).
2014-10-28 11:46:27 641 [Note] Plugin 'FEDERATED' is disabled.
2014-10-28 11:46:27 641 [Note] InnoDB: Using atomics to ref count buffer pool pages
2014-10-28 11:46:27 641 [Note] InnoDB: The InnoDB memory heap is disabled
2014-10-28 11:46:27 641 [Note] InnoDB: Mutexes and rw_locks use GCC atomic builtins
2014-10-28 11:46:27 641 [Note] InnoDB: Memory barrier is not used
2014-10-28 11:46:27 641 [Note] InnoDB: Compressed tables use zlib 1.2.3
2014-10-28 11:46:27 641 [Note] InnoDB: Using Linux native AIO
2014-10-28 11:46:27 641 [Note] InnoDB: Using CPU crc32 instructions
/usr/sbin/mysqld: Can't create/write to file '/var/tmp/ibYhjMI4' (Errcode: 13 - Permission denied)
2014-10-28 11:46:27 7f61a2bde740  InnoDB: Error: unable to create temporary file; errno: 13
2014-10-28 11:46:27 641 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-28 11:46:27 641 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-28 11:46:27 641 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-28 11:46:27 641 [ERROR] Aborting
```

FIX TMPDIR

```
[root@hostdb ~]# ls -ld /var/tmp
drwxrwx--T 2 root root 4096 Mar 13 13:23 /var/tmp
```

```
[root@hostdb ~]# chmod a+rwx /var/tmp
[root@hostdb ~]# ls -ld /var/tmp
drwxrwxrwt 2 root root 4096 Mar 13 13:23 /var/tmp
```

```
[root@hostdb ~]# service mysql start
Starting MySQL ... ERROR! The server quit without updating PID file
    (/var/lib/mysql/ip-10-87-0-19.pid).
```

CANNOT ALLOCATE MEMORY

```
141028 11:47:42 mysqld_safe Starting mysqld daemon with databases from /var/lib/mysql
2014-10-28 11:47:44 0 [Warning] TIMESTAMP with implicit DEFAULT value is deprecated. Please use
    --explicit_defaults_for_timestamp server option (see documentation for more details).
2014-10-28 11:47:44 898 [Note] Plugin 'FEDERATED' is disabled.
2014-10-28 11:47:44 898 [Note] InnoDB: Using atomics to ref count buffer pool pages
2014-10-28 11:47:44 898 [Note] InnoDB: The InnoDB memory heap is disabled
2014-10-28 11:47:44 898 [Note] InnoDB: Mutexes and rw_locks use GCC atomic builtins
2014-10-28 11:47:44 898 [Note] InnoDB: Memory barrier is not used
2014-10-28 11:47:44 898 [Note] InnoDB: Compressed tables use zlib 1.2.3
2014-10-28 11:47:44 898 [Note] InnoDB: Using Linux native AIO
2014-10-28 11:47:44 898 [Note] InnoDB: Using CPU crc32 instructions
2014-10-28 11:47:44 898 [Note] InnoDB: Initializing buffer pool, size = 100.0G
InnoDB: mmap(13736345600 bytes) failed; errno 12
2014-10-28 11:47:44 898 [ERROR] InnoDB: Cannot allocate memory for the buffer pool
2014-10-28 11:47:44 898 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-28 11:47:44 898 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-28 11:47:44 898 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-28 11:47:44 898 [ERROR] Aborting
```

FIX INNODB_BUFFER_POOL_SIZE

```
[root@hostdb ~]# perror 12
OS error code  12:  Cannot allocate memory
[root@hostdb ~]# grep 100 /etc/my.cnf
innodb_buffer_pool_size=100G
```

```
[root@hostdb ~]# sed -i -e 's/100G/256M/' /etc/my.cnf
[root@hostdb ~]# grep innodb_buffer_pool_size /etc/my.cnf
innodb_buffer_pool_size=256M
```

```
[root@hostdb ~]# service mysql start
StartingMySQL... ERROR! The server quit without updating PID file
    (/var/lib/mysql/ip-10-87-0-19.pid).
```

AGAIN ERROR 13

```
141028 11:49:50 mysqld_safe Starting mysqld daemon with databases from /var/lib/mysql
2014-10-28 11:49:51 0 [Warning] TIMESTAMP with implicit DEFAULT value is deprecated. Please use
    --explicit_defaults_for_timestamp server option (see documentation for more details).
2014-10-28 11:49:51 1169 [Note] Plugin 'FEDERATED' is disabled.
2014-10-28 11:49:51 1169 [Note] InnoDB: Using atomics to ref count buffer pool pages
2014-10-28 11:49:51 1169 [Note] InnoDB: The InnoDB memory heap is disabled
2014-10-28 11:49:51 1169 [Note] InnoDB: Mutexes and rw_locks use GCC atomic builtins
2014-10-28 11:49:51 1169 [Note] InnoDB: Memory barrier is not used
2014-10-28 11:49:51 1169 [Note] InnoDB: Compressed tables use zlib 1.2.3
2014-10-28 11:49:51 1169 [Note] InnoDB: Using Linux native AIO
2014-10-28 11:49:51 1169 [Note] InnoDB: Using CPU crc32 instructions
2014-10-28 11:49:51 1169 [Note] InnoDB: Initializing buffer pool, size = 256.0M
2014-10-28 11:49:51 1169 [Note] InnoDB: Completed initialization of buffer pool
2014-10-28 11:49:51 1169 [ERROR] InnoDB: ./ibdata1 can't be opened in read-write mode
2014-10-28 11:49:51 1169 [ERROR] InnoDB: The system tablespace must be writable!
2014-10-28 11:49:51 1169 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-28 11:49:51 1169 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-28 11:49:51 1169 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-28 11:49:51 1169 [ERROR] Aborting
```

FIX PERMISSIONS

```
[root@hostdb ~]# ls -l /var/lib/mysql/ibdata1
-rw-rw---- 1 27 27 18874368 Mar 13 12:34 /var/lib/mysql/ibdata1
[root@hostdb ~]# ls -l /var/lib/mysql
total 83980
-rw-rw---- 1  27   27 18874368 Mar 13 12:34 ibdata1
-rw-rw---- 1  27   27 33554432 Mar 13 12:34 ib_logfile0
-rw-rw---- 1  27   27 33554432 Mar 13 12:34 ib_logfile1
drwx------ 2 mysql mysql    4096 Mar 13 12:30 mysql
drwx------ 2 root  root     4096 Mar 13 12:30 performance_schema
drwx------ 2 root  root     4096 Mar 13 12:30 test
[root@hostdb ~]# chown -R mysql:mysql /var/lib/mysql
[root@hostdb~]# service mysql start
Starting MySQL......... SUCCESS!
```

FIX PERMISSIONS

```
[root@hostdb ~]#  tail -n 100 /var/log/mysqld.log
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_history_long' has the wrong structure
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_summary_by_thread_by_event_name' has the wrong
    structure
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_summary_by_account_by_event_name' has the wrong
    structure
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_summary_by_user_by_event_name' has the wrong
    structure
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_summary_by_host_by_event_name' has the wrong
    structure
2014-11-01 21:17:17 8895 [ERROR] Native table
    'performance_schema'.'events_statements_summary_global_by_event_name' has the wrong
    structure
```

ACCESS TO MYSQLD

```
[root@hostdb ~]# mysql
ERROR 2002 (HY000): Can't connect to local MySQL server through socket
    '/tmp/mysql.sock' (2)
```

```
[root@hostdb ~]# perror 2
OS error code   2:  No such file or directory
[root@hostdb ~]# ls -l /tmp/mysql.sock
ls: cannot access /tmp/mysql.sock: No such file or directory
```

ACCESS TO MYSQLD

```
[root@hostdb ~]# grep socket /var/log/mysqld.log | tail -n 1
Version: '5.5.34'  socket: '/var/lib/mysql/mysql.sock'  port: 3306  MySQL
    Community Server (GPL)
[root@hostdb ~]# lsof -n | grep mysql | grep unix
mysqld    21737     mysql    12u    unix 0xffff880002e0dd40      0t0       22829
    /var/lib/mysql/mysql.sock
```

```
[root@hostdb ~]# grep -B 1 socket /etc/my.cnf

[client]
socket=/tmp/mysql.sock
```

```
[root@hostdb ~]# sed -i -e 's/\/tmp\/mysql.sock/\/var\/lib\/mysql\/mysql.sock/'
    /etc/my.cnf
```

ACCESS TO MYSQLD

```
[root@hostdb ~]# mysql
ERROR 1045 (28000): Access denied for user 'root'@'localhost' (using password:
    YES)
```

```
[root@hostdb ~]# strace -e trace=open mysql
...
open("/etc/my.cnf", O_RDONLY)           = 3
open("/usr/etc/my.cnf", O_RDONLY)       = 3
open("/root/.my.cnf", O_RDONLY)         = 3
```

```
[root@hostdb ~]# cat ~/.my.cnf
[client]
password=adummypassword
```

ACCESS TO MYSQLD

```
[root@hostdb ~]# mysql --no-defaults
```

```
[root@hostdb ~]# mysql -p
```

CHANGE ROOT PASSWORD

```
[root@hostdb ~]# echo "SET PASSWORD=PASSWORD('$RANDOM$RANDOM')" | mysql
[root@hostdb ~]# mysql
ERROR 1045 (28000): Access denied for user 'root'@'localhost' (using password: NO)
```

Add skip-grant-tables to the [mysqld] section

```
[root@hostdb ~]# sed -i 's/\[mysqld\]/&\nskip-grant-tables/' /etc/my.cnf

[root@hostdb ~]# service mysql restart
Shutting down MySQL.. SUCCESS!
Starting MySQL... SUCCESS!
```

CHANGE ROOT PASSWORD

```
[root@hostdb ~]# mysql

mysql> UPDATE mysql.user SET password=PASSWORD('newpass') WHERE user='root';
mysql> FLUSH PRIVILEGES;
```

Remove `skip-grant-tables` from /etc/my.cnf

```
[root@hostdb ~]# sed -i 's/skip-grant-tables//' /etc/my.cnf

[root@hostdb ~]# service mysql restart

[root@hostdb ~]# sed -i 's/password=adummypassword/password=newpass/' ~/.my.cnf
```

MYSQL_UPGRADE

```
[root@hostdb~]# grep "ERROR" /var/log/mysqld.log |tail -n 3
2014-11-01 21:24:09 9336 [ERROR] Native table 'performance_schema'.'socket_summary_by_event_name' has
    the wrong structure
2014-11-01 21:24:09 9336 [ERROR] Native table 'performance_schema'.'session_connect_attrs' has the wrong
    structure
2014-11-01 21:24:09 9336 [ERROR] Native table 'performance_schema'.'session_account_connect_attrs' has
    the wrong structure
[root@hostdb ~]# mysql_upgrade
Looking for 'mysql' as: mysql
Looking for 'mysqlcheck' as: mysqlcheck
Running 'mysqlcheck with default connection arguments
Running 'mysqlcheck with default connection arguments
mysql.columns_priv                                 OK
...
mysql.user                                         OK
Running 'mysql_fix_privilege_tables'...
Running 'mysqlcheck with default connection arguments
Running 'mysqlcheck with default connection arguments
OK
[root@hostdb ~]# service mysql restart
```

LOCAL_INFILE = ON

http://dev.mysql.com/doc/refman/5.6/en/load-data-local.html

"In a Web environment where the clients are connecting from a Web server, a user could use LOAD DATA LOCAL to read any files that the Web server process has read access to (assuming that a user could run any command against the SQL server)."

```
[root@hostdb~]# mysql -e "SHOW GLOBAL VARIABLES LIKE 'local_infile'"
+---------------+-------+
| Variable_name | Value |
+---------------+-------+
| local_infile  | ON    |
+---------------+-------+
```

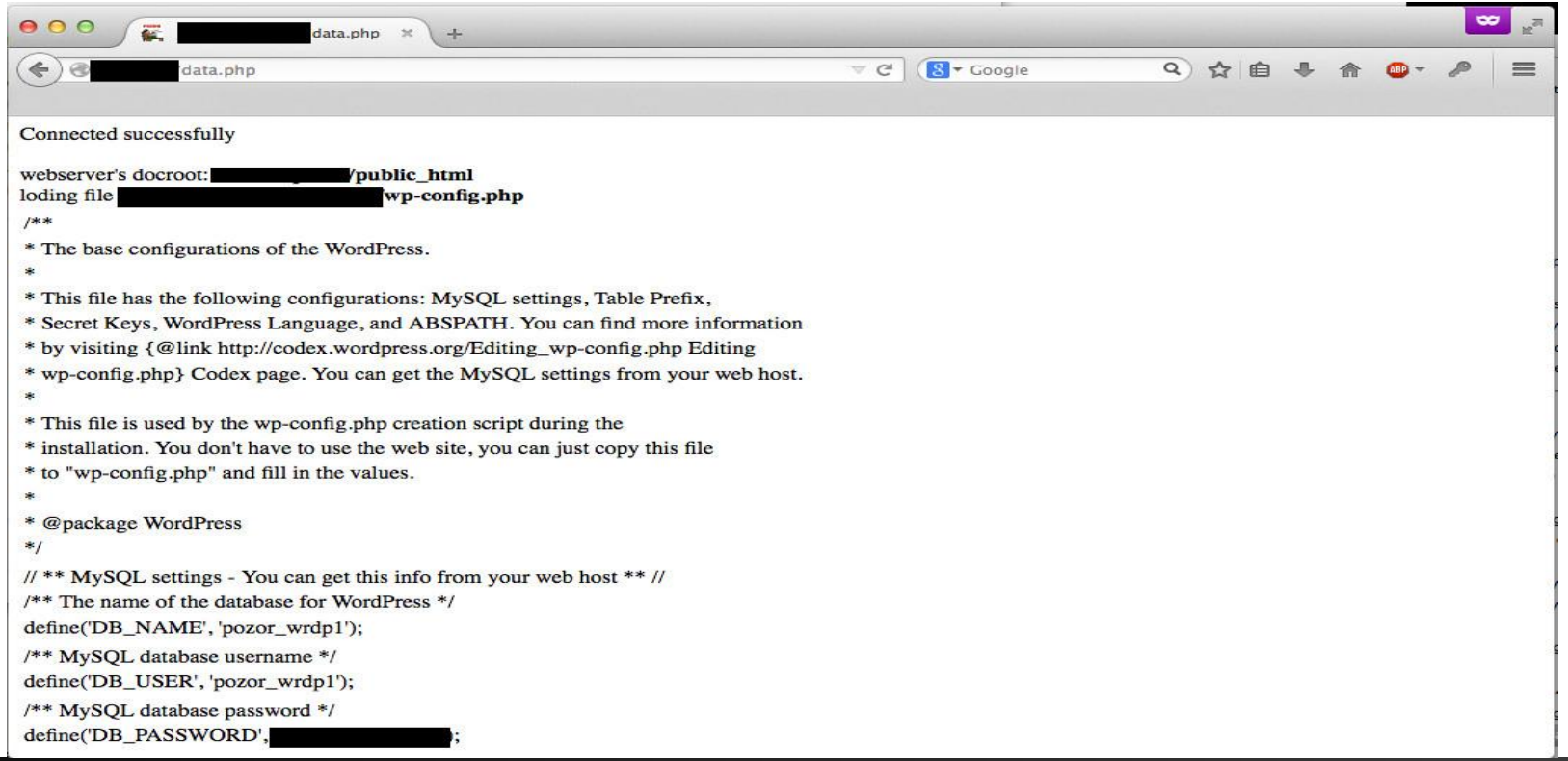
LOCAL_INFILE = ON

```
[root@hostdb ~]# mysql -e "show grants for evil@localhost”
GRANT USAGE ON *.* TO 'evil'@'localhost' IDENTIFIED BY PASSWORD ‘xxx’
GRANT SELECT, INSERT, CREATE ON `test`.* TO 'evil'@'localhost’

[root@hostdb ~]# mysql -u evil -p4242 test

mysql> select @@datadir;
+-----------------+
| @@datadir       |
+-----------------+
| /var/lib/mysql/ |
+-----------------+

mysql> LOAD DATA LOCAL INFILE  '/var/lib/mysql/mysql/user.MYD' INTO TABLE store
    LINES TERMINATED BY '*';
```

LOCAL_INFILE = ON

```
mysql> select * from store\G
...
*************************** 4. row ***************************
col1: D8DECEC305209EEFEC43008E1D420E1AA06B19E0   Uûü::1root
col2: NULL
col3: NULL
*************************** 6. row ***************************
col1: 89C6B530AA78695E257E55D63C00A6EC9AD3E977mysql_native_password   jûü%webapp
col2: NULL
col3: NULL

mysql>  LOAD DATA LOCAL INFILE  '/root/.my.cnf' INTO TABLE store   LINES
    TERMINATED BY '\n';
mysql> select * from store\G
...
*************************** 35. row ***************************
col1: password=newpass
col2: NULL
col3: NULL
```

LOCAL_INFILE = ON - POC



REMOVED IB_LOGFILE (DON'T!)

Free space was low: I deleted some log files...

```
[root@ ~]# cd /var/lib/mysql
[root@hostdb mysql]# rm -rf ib_logfile*

[root@hostdb mysql]# lsof -n | grep ib_logfile
mysqld     12076 mysql 8uW   REG   202,1 33554432  403503 /var/lib/mysql/ib_logfile0 (deleted)
mysqld     12076 mysql 9uW   REG   202,1 33554432  403505 /var/lib/mysql/ib_logfile1 (deleted)

Or
[root@hostdb mysql]# ll /proc/`pidof mysqld`/fd/
total 0
[..]
lrwx------ 1 root root 64 Oct 30 17:19 8 -> /var/lib/mysql/ib_logfile0 (deleted)
lrwx------ 1 root root 64 Oct 30 17:19 9 -> /var/lib/mysql/ib_logfile1 (deleted)
```

REMOVED IB_LOGFILE (DON'T!)

mysqld and InnoDB continue working normally

```
mysql> use test;
mysql> create table tbl1 (id int auto_increment primary key, v varchar(100))
    engine=innodb;
mysql> insert into tbl1 values(null,'aa');
mysql> insert into tbl1 select null, v from tbl1;
mysql> SELECT * FROM tbl1;
```

```
[root@hostdb ~]# service mysql restart
```

REMOVED IB_LOGFILE (DON'T!)

InnoDB Redo Log are automatically recreated

```
2014-10-30 17:22:24 14206 [Note] InnoDB: Setting log file ./ib_logfile101 size to 64 MB
2014-10-30 17:22:28 14206 [Note] InnoDB: Setting log file ./ib_logfile1 size to 64 MB
2014-10-30 17:22:32 14206 [Note] InnoDB: Renaming log file ./ib_logfile101 to ./ib_logfile0
2014-10-30 17:22:32 14206 [Warning] InnoDB: New log files created, LSN=1605224
```

REMOVED IBDATA1 (REALLY, DON'T!)

Free space was low: I deleted a large file...

```
[root@hostdb mysql]# rm -f ibdata1

[root@hostdb mysql]# lsof -n | grep ibdata1
mysqld   7478 mysql 3uW  REG  202,1  27262976     135257 /var/lib/mysql/ibdata1
   (deleted)
```

REMOVED IBDATA1 (REALLY, DON'T!)

mysqld and InnoDB continue working normally … :

```
mysql> use test;
mysql> insert into tbl1 select null, v from tbl1;
mysql> SELECT * FROM tbl1;
```

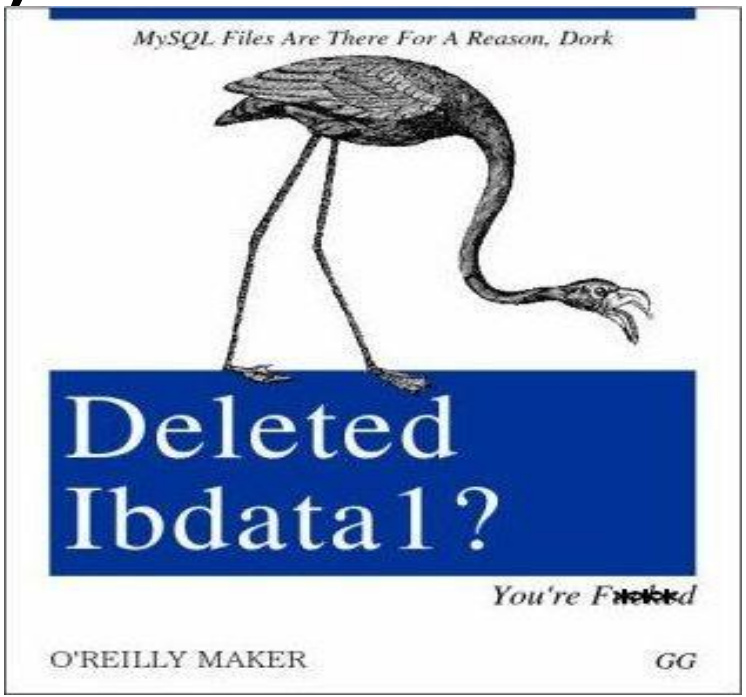
… until restart:

```
[root@hostdb ~]# service mysql restart
service mysql restart
Shutting down MySQL.. SUCCESS!
Starting MySQL..... ERROR! The server quit without updating PID file
(/var/lib/msql/ip-10-87-0-19.pid).
```

REMOVED IBDATA1 (REALLY, DON'T!)

```
2014-10-30 17:34:33 14786 [Note] InnoDB: Initializing buffer pool, size = 256.0M
2014-10-30 17:34:33 14786 [Note] InnoDB: Completed initialization of buffer pool
2014-10-30 17:34:33 14786 [Note] InnoDB: Restoring page 0 of tablespace 0
2014-10-30 17:34:33 14786 [Warning] InnoDB: Doublewrite does not have page_no=0 of space: 0
2014-10-30 17:34:33 14786 [ERROR] InnoDB: space header page consists of zero bytes in data file
    ./ibdata1
2014-10-30 17:34:33 14786 [ERROR] InnoDB: Could not open or create the system tablespace. If you
    tried to add new data files to the system tablespace, and it failed here, y
ou should now edit innodb_data_file_path in my.cnf back to what it was, and remove the new
    ibdata files InnoDB created in this failed attempt. InnoDB only wrote those files
 full of zeros, but did not yet use them in any way. But be careful: do not remove old data
    files which contain your precious data!
2014-10-30 17:34:33 14786 [ERROR] Plugin 'InnoDB' init function returned error.
2014-10-30 17:34:33 14786 [ERROR] Plugin 'InnoDB' registration as a STORAGE ENGINE failed.
2014-10-30 17:34:33 14786 [ERROR] Unknown/unsupported storage engine: InnoDB
2014-10-30 17:34:33 14786 [ERROR] Aborting
```

REMOVED IBDATA1 (REALLY, DON'T!)



You can now recover from backup!

MySQL BREAK/FIX LAB
Replication

Presented by: Miklós Szél, Gillian Gunson, Brian Cain

April 13, 2015

Pythian®
love your data

AGENDA - PART 2

- Replication overview and tools
- Bad server-id
  - slave id identical to master
  - slave id identical to other slave
- Incorrect slave data
  - Duplicate key error
  - Data drift

REPLICATION OVERVIEW

- asynchronous: based on copying and executing of binary logs
  - allows for delayed slaves, incremental backups, point-in-time recovery
- binary log statements have two formats
  - statement: SQL queries
  - row events: row changes; starting with 5.1
- statement binlog_format will be used in this tutorial

REPLICATION OVERVIEW (cont.)

- replication is run by two threads on slave
  - IO thread: copies master binary logs to slave relay logs
  - SQL thread: executes relay log events on slave
- binary log position: file name and byte offset from start of file
- MySQL 5.6 has some new features
  - GTIDs: global transaction identifiers
  - multi-threaded slaves
  - (won’t be covered in this tutorial)

REPLICATION TOOLS

- MySQL Sandbox
  - allows you to install and run multiple instances of MySQL on same server as non-root user
  - Linux/FreeBSD/MacOSX
  - not for production instances: plaintext passwords
  - can easily set up replication groups
    - default is master and two slaves

REPLICATION TOOLS (cont.)

- Percona Toolkit
  - Linux only
  - pt-table-checksum/pt-table-sync
  - pt-slave-restart
  - etc.

- mysqlbinlog: binary log parser

- MySQL error log

REPLICATION TOOLS (cont.)

- slave commands
  - SHOW SLAVE STATUS\G
  - STOP SLAVE; START SLAVE;
  - STOP/START SLAVE IO_THREAD;
  - STOP/START SLAVE SQL_THREAD;

- master commands
  - SHOW MASTER STATUS;
  - SHOW SLAVE HOSTS;

REPLICATION SANDBOX

```
[user-lab@hostdb repl_test]$ pwd
/home/user-lab/sandboxes/repl_test

[user-lab@hostdb repl_test]$ ls
check_slaves             m       restart_all    status_all
clear_all                master  s1             stop_all
connection.json          node1   s2             use_all
default_connection.json  node2   send_kill_all
initialize_slaves        README  start_all

[user-lab@hostdb repl_test]$ ./start_all
# executing "start" on /home/user-lab/sandboxes/repl_test
executing "start" on master
... sandbox server started
executing "start" on slave 1
... sandbox server started
executing "start" on slave 2
... sandbox server started
```

SANDBOX MASTER

```
[user-lab@hostdb repl_test]$ cd master

[user-lab@hostdb master]$ ls
change_paths             grants.mysql    proxy_start             status
change_ports             json_in_db      README                  stop
clear                    load_grants     rescue_mysql_dump.sql   tmp
connection.json          msb             restart                 use
data                     my              send_kill               USING
default_connection.json  my.sandbox.cnf  start

[user-lab@hostdb master]$ ls data
ibdata1      msandbox.err      mysql-bin.index         test
ib_logfile0  mysql             mysql_sandbox23992.pid
ib_logfile1  mysql-bin.000001  performance_schema
```

SANDBOX SLAVE

```
[user-lab@hostdb master]$ cd ../node1

[user-lab@hostdb node1]$ ls
change_paths     default_connection.json  my              send_kill  use
change_ports     grants.mysql             my.sandbox.cnf  start      USING
clear            json_in_db               proxy_start     status
connection.json  load_grants              README          stop
data             msb                      restart         tmp

[user-lab@hostdb node1]$ ls data
ibdata1            mysql_sandbox23993.pid
ib_logfile0        mysql_sandbox23993-relay-bin.000001
ib_logfile1        mysql_sandbox23993-relay-bin.000002
master.info        mysql_sandbox23993-relay-bin.index
msandbox.err       performance_schema
mysql              relay-log.info
mysql-bin.000001   test
mysql-bin.index
```

REPLICATION PROCESSLIST THREADS - MASTER

```
[user-lab@hostdb repl_test]$ ./m
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 3

master [localhost] {msandbox} ((none)) > show processlist;
+----+----------+-----------------+------+-------------+------
    +-----------------------------------------------------------------------+------------------+
| Id | User     | Host            | db   | Command     | Time | State
    | Info             |
+----+----------+-----------------+------+-------------+------
    +-----------------------------------------------------------------------+------------------+
|  1 | rsandbox | localhost:44094 | NULL | Binlog Dump |  147 | Master has sent all binlog to slave;
    waiting for binlog to be updated | NULL             |
|  2 | rsandbox | localhost:44095 | NULL | Binlog Dump |  144 | Master has sent all binlog to slave;
    waiting for binlog to be updated | NULL             |
|  4 | msandbox | localhost       | NULL | Query       |    0 | NULL
    | show processlist |
+----+----------+-----------------+------+-------------+------
    +-----------------------------------------------------------------------+------------------+
3 rows in set (0.00 sec)
```

REPLICATION PROCESSLIST THREADS - SLAVE

```
[user-lab@hostdb repl_test]$ ./s1
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 3

slave1 [localhost] {msandbox} ((none)) > show processlist;
+----+-------------+-----------+------+---------+------
     +-----------------------------------------------------------------------------------+------------------+
| Id | User        | Host      | db   | Command | Time | State
     | Info             |
+----+-------------+-----------+------+---------+------
     +-----------------------------------------------------------------------------------+------------------+
|  1 | system user |           | NULL | Connect |   31 | Slave has read all relay log; waiting for the
     slave I/O thread to update it | NULL             |
|  2 | system user |           | NULL | Connect |   31 | Waiting for master to send event
     | NULL             |
|  3 | msandbox    | localhost | NULL | Query   |    0 | NULL
     | show processlist |
+----+-------------+-----------+------+---------+------
     +-----------------------------------------------------------------------------------+------------------+
3 rows in set (0.00 sec)
```

PROBLEM #1: DUPLICATE SERVER-ID VALUES

- the server-id uniquely identifies a server in a replication topology
  - recorded in binary log
  - dynamic variable (changed without restart)
- server-uuid added MySQL 5.6
  - set in auto.cnf
  - static variable
- if a server is cloned (e.g., ec2 snapshot), the server-id/server-uuid need to be edited in new instance

PROBLEM #1: DUPLICATE SERVER-ID VALUES

Slave with identical server_id to master:

```
$ ./s1 -e "show slave status\G"
*************************** 1. row ***************************
               Slave_IO_State:
…
             Slave_IO_Running: No
            Slave_SQL_Running: Yes
…
        Seconds_Behind_Master: NULL
Master_SSL_Verify_Server_Cert: No
                Last_IO_Errno: 1593
                Last_IO_Error: Fatal error: The slave I/O thread stops because master and slave
    have equal MySQL server ids; these ids must be different for replication to work (or the --
    replicate-same-server-id option must be used on slave but this does not always make sense;
    please check the manual before using it).
           Master_Server_Id: 1
```

SOLUTION #1: MODIFY SLAVE SERVER-ID

1. set server_id variable

```
[user-lab@hostdb dupl_server_id1]$ ./s1

slave1 [localhost] {msandbox} ((none)) > select @@server_id;
+-------------+
| @@server_id |
+-------------+
|           1 |
+-------------+
1 row in set (0.00 sec)

slave1 [localhost] {msandbox} ((none)) > stop slave; set global server_id = 102; start slave;

Query OK, 0 rows affected (0.00 sec)

slave1 [localhost] {msandbox} ((none)) > show slave status\G
```

SOLUTION #1: MODIFY SLAVE SERVER-ID

2. edit node1/my.sandbox.cnf to use different server-id

```
[user-lab@hostdb dupl_server_id1]$ cd node1

[user-lab@hostdb node1]$ sed -i 's/server-id=1/server-id=102/g' my.
   sandbox.cnf

[user-lab@hostdb node1]$ grep server-id my.sandbox.cnf
server-id=102
```

PROBLEM #1B: DUPLICATE SERVER-ID VALUES

- with two slaves with identical server-ids and server-uuids, the symptoms and error messages can be misleading
- master will show both slave connections in the processlist, but only 1 slave in “SHOW SLAVE HOSTS”

```
master [localhost] {msandbox} ((none)) > show processlist;
+----+----------+-----------------+------+-------------+------
+-----------------------------------------------------------------------+------------------+
| Id | User     | Host            | db   | Command     | Time | State
| Info             |
+----+----------+-----------------+------+-------------+------
+-----------------------------------------------------------------------+------------------+
|  1 | rsandbox | localhost:42165 | NULL | Binlog Dump |  158 | Master has sent all binlog to slave; waiting for binlog to be
updated | NULL             |
|  2 | rsandbox | localhost:42166 | NULL | Binlog Dump |  156 | Master has sent all binlog to slave; waiting for binlog to be
updated | NULL             |
|  5 | msandbox | localhost       | NULL | Query       |    0 | init
| show processlist |
+----+----------+-----------------+------+-------------+------
+-----------------------------------------------------------------------+------------------+
3 rows in set (0.00 sec)
```

PROBLEM #1: DUPLICATE SERVER-ID VALUES

```
master [localhost] {msandbox} ((none)) > show slave hosts;
  +-----------+----------+-------+-----------+--------------------------------------+
  | Server_id | Host     | Port  | Master_id | Slave_UUID                           |
  +-----------+----------+-------+-----------+--------------------------------------+
  |       101 | SBslave2 | 21502 |         1 | 6a9e9e41-5453-11e4-ae10-22000af80ef3 |
  +-----------+----------+-------+-----------+--------------------------------------+
  1 row in set (0.00 sec)


$ tail /home/user-lab/sandboxes/dupl_server_id2/node1/data/msandbox.err

2015-04-13 05:52:55 5142 [Note] Slave: received end packet from server due to dump thread being killed
    on master. Dump threads are killed for example during master shutdown, explicitly by a user, or when
    the master receives a binlog send request from a duplicate server UUID <67188b5d-5453-11e4-ae10-
    22000af80ef3> : Error
2015-04-13 05:52:55 5142 [Note] Slave I/O thread: Failed reading log event, reconnecting to retry, log
    'mysql-bin.000009' at position 120
```

PROBLEM #2: DUPLICATE KEY ERRORS

- Error 1062
- very common replication error
- number of possible causes
  - slave created from bad backup
  - slave replication started at wrong binlog file/position
  - mix of transactional and nontransactional tables
  - direct write to the slave
  - “data drift”

PROBLEM #2: DUPLICATE KEY ERRORS

Example

```
[user-lab@hostdb sandboxes]$ cd dupl_key/

[user-lab@hostdb dupl_key]$ ./start_all
# executing "start" on /home/user-lab/sandboxes/dupl_key
executing "start" on master
... sandbox server started
executing "start" on slave 1
... sandbox server started

[user-lab@hostdb dupl_key]$ ./s1
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 3
```

PROBLEM #2: DUPLICATE KEY ERRORS

Example

```
slave1 [localhost] {msandbox} (test) > show slave status\G
*************************** 1. row ***************************
               Slave_IO_State: Waiting for master to send event
                  Master_Host: 127.0.0.1
                  Master_User: rsandbox
...
             Slave_IO_Running: Yes
            Slave_SQL_Running: No
...
                   Last_Errno: 1062
                   Last_Error: Error 'Duplicate entry '3' for key 'PRIMARY'' on query.
    Default database: 'test'. Query: 'insert into dupe_test (val) values ('master3'),
    ('master4')'
...
        Seconds_Behind_Master: NULL
```

PROBLEM #2: DUPLICATE KEY ERRORS

Example (cont.)

```
slave1 [localhost] {msandbox} ((none)) > use test;

slave1 [localhost] {msandbox} (test) > show create table dupe_test\G
*************************** 1. row ***************************
       Table: dupe_test
Create Table: CREATE TABLE `dupe_test` (
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `val` varchar(20) NOT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=4 DEFAULT CHARSET=latin1
1 row in set (0.00 sec)

slave1 [localhost] {msandbox} (test) > select * from dupe_test where id = 3;
+----+--------+
| id | val    |
+----+--------+
|  3 | slave1 |
+----+--------+
1 row in set (0.00 sec)
```

PROBLEM #2: DUPLICATE KEY ERRORS

Example (cont.)

```
[user-lab@hostdb dupl_key]$ cd master

[user-lab@hostdb master]$ mysqlbinlog data/mysql-bin.000002
...
# at 521
#140401  7:05:14 server id 1  end_log_pos 589   Query thread_id=3 exec_time=0 error_code=0
SET TIMESTAMP=1396335914/*!*/;
BEGIN
/*!*/;
# at 589
#140401  7:05:14 server id 1  end_log_pos 617   Intvar
SET INSERT_ID=3/*!*/;
# at 617
#140401  7:05:14 server id 1  end_log_pos 739   Query thread_id=3 exec_time=0 error_code=0
SET TIMESTAMP=1396335914/*!*/;
insert into dupe_test (val) values ("master3"), ('master4')
/*!*/;
...
```

SOLUTION #2: DUPLICATE KEY ERRORS

General solutions:

1. recreate slave from fresh backup of master (preferred)
2. get replication running again, fix data differences later

We will be doing solution #2.

SOLUTION #2: DUPLICATE KEY ERRORS

To check if the error is due to a direct write to slave:

```
[user-lab@hostdb dupl_key]$ cd /home/user-lab/sandboxes/dupl_key/node1

[user-lab@hostdb node1]$ ls data/mysql-bin*
data/mysql-bin.000001  data/mysql-bin.000002  data/mysql-bin.index

[user-lab@hostdb node1]$ mysqlbinlog --server-id=101 data/mysql-bin.00000* | grep
dupe_test
insert into dupe_test (val) values ('slave1')
```

SOLUTION #2: DUPLICATE KEY ERRORS

Do this (if it's a read slave):

```
[mysqld]
read_only
```

Do **NOT** do this:

```
[mysqld]
slave-skip-errors = all
```

SOLUTION #2: DUPLICATE KEY ERRORS

Hack #1: Get slave caught up

```
[user-lab@hostdb dupl_key]$ ../s1

slave1 > stop slave; set global sql_slave_skip_counter = 1; start slave;

slave1 > show slave status\G
```

If you have a lot of events to skip, you can use **pt-slave-restart**.

SOLUTION #2: DUPLICATE KEY ERRORS

Hack #1: Get slave caught up

- note this skips binary log “events”
  - for InnoDB tables: entire transaction

```
slave1 [localhost] {msandbox} (test) > select * from test.dupe_test;
+----+---------+
| id | val     |
+----+---------+
|  1 | master1 |
|  2 | master2 |
|  3 | slave1  |
+----+---------+
3 rows in set (0.00 sec)
```

- can’t easily guarantee that master and slave data are equivalent

PROBLEM #3: DATA DRIFT

- Statement-based binary logging can result in incorrect data on the slave
  - “nondeterministic”
  - updates with LIMIT but no ORDER BY
  - certain system and math functions
  - etc.

- ROW and MIXED binary logging mostly avoids this

PROBLEM #3: DATA DRIFT

```
[user-lab@ip-10-74-3-30 dupl_key]$ ./m

master [localhost] {msandbox} ((none)) > use test;

master [localhost] {msandbox} (test) > insert into dupe_test (val) values (uuid());
Query OK, 1 row affected, 2 warnings (0.04 sec)

master [localhost] {msandbox} (test) > show warnings\G
*************************** 1. row ***************************
  Level: Warning
   Code: 1265
Message: Data truncated for column 'val' at row 1
*************************** 2. row ***************************
  Level: Note
   Code: 1592
Message: Unsafe statement written to the binary log using statement format since BINLOG_FORMAT =
STATEMENT. Statement is unsafe because it uses a system function that may return a different value on
the slave.
2 rows in set (0.00 sec)
```

PROBLEM #3: DATA DRIFT

```
master [localhost] {msandbox} (test) > select * from dupe_test order by id desc limit 1;
+----+----------------------+
| id | val                  |
+----+----------------------+
|  9 | f6fe0011-b932-11e3-a |
+----+----------------------+
1 row in set (0.02 sec)

slave1 [localhost] {msandbox} ((none)) > select * from test.dupe_test order by id desc limit 1;
+----+----------------------+
| id | val                  |
+----+----------------------+
|  9 | f6ff6a8c-b932-11e3-a |
+----+----------------------+
1 row in set (0.00 sec)
```

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Hack #2: Checksum and sync slave data

- pt-table-checksum
  - checksums tables in “chunks” of rows
  - can use replication to compare master and slave
  - can write to checksum table, results on slave

- pt-table-sync
  - can be run without pt-table-checksum
  - can sync data via REPLACE and DELETE statements directly run on master or printed out for review

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Checksum All Tables

```
[user-lab@hostdb dupl_key]$ pt-table-checksum -uroot --ask-pass --replicate test.checksum --host
127.0.0.1 --port 20000
Enter MySQL password: msandbox
...
            TS ERRORS  DIFFS     ROWS  CHUNKS SKIPPED    TIME TABLE
04-01T00:48:17      0      0        0       1       0   0.051 mysql.columns_priv
04-01T00:48:17      0      0        0       1       0   0.052 mysql.db
04-01T00:48:17      0      0        0       1       0   0.049 mysql.event
04-01T00:48:17      0      0        0       1       0   0.050 mysql.func
...
10-22T12:16:57      0      1        5       1       0   0.037 test.dupe_test
```

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Checksum Results Check

```
[user-lab@hostdb dupl_key]$ ./s1

slave1 [localhost] {msandbox} (test) > SELECT db, tbl, SUM(this_cnt) AS total_rows, COUNT(*) AS chunks
FROM test.checksum
WHERE (
  master_cnt <> this_cnt
  OR master_crc <> this_crc
  OR ISNULL(master_crc) <> ISNULL(this_crc))
  GROUP BY db, tbl;
+------+-----------+------------+--------+
| db   | tbl       | total_rows | chunks |
+------+-----------+------------+--------+
| test | dupe_test |          4 |      1 |
+------+-----------+------------+--------+
1 row in set (0.00 sec)
```

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Sync Output

```
[user-lab@hostdb dupl_key]$ pt-table-sync -uroot --ask-pass --print --replicate test.checksum --
sync-to-master --socket=/tmp/mysql_sandbox20001.sock D=test,t=dupe_test
Enter password for DSN D=test,S=/tmp/mysql_sandbox20001.sock,t=dupe_test,u=root:
msandbox

REPLACE INTO `test`.`dupe_test`(`id`, `val`) VALUES ('3', 'master3') /*percona-toolkit src_db:
test src_tbl:dupe_test src_dsn:D=test,P=23992,S=/tmp/mysql_sandbox20001.sock,h=127.0.0.1,p=...,
t=dupe_test,u=root dst_db:test dst_tbl:dupe_test dst_dsn:D=test,S=/tmp/mysql_sandbox20001.sock,
p=...,t=dupe_test,u=root lock:1 transaction:1 changing_src:test.checksum replicate:test.checksum
bidirectional:0 pid:26192 user:user-lab host:hostdb*/;
...
```

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Sync to Master Data

```
[user-lab@hostdb dupl_key]$ pt-table-sync -uroot --ask-pass --execute --replicate
test.checksum --sync-to-master --socket=/tmp/mysql_sandbox20001.sock D=test,
t=dupe_test
Enter password for DSN D=test,S=/tmp/mysql_sandbox20001.sock,t=dupe_test,u=root:
msandbox
```

Note: pt-table-sync is connecting to the slave socket to start, as that's the location of the test.checksum table results.

SOLUTION #3: FIXING DATA DRIFT/ INCORRECT SLAVE

Quick Sync Check

```
[user-lab@ip-10-252-38-35 dupl_key]$ ./m -e "select * from test.dupe_test"
+----+---------+
| id | val     |
+----+---------+
|  1 | master1 |
|  2 | master2 |
|  3 | master3 |
|  4 | master4 |
+----+---------+
[user-lab@ip-10-252-38-35 dupl_key]$ ./s1 -e "select * from test.dupe_test"
+----+---------+
| id | val     |
+----+---------+
|  1 | master1 |
|  2 | master2 |
|  3 | master3 |
|  4 | master4 |
+----+---------+
```

Thanks

LinkedIn: https://ca.linkedin.com/in/gilliangunson

Twitter: https://twitter.com/shebang_the_cat

Instagram: https://instagram.com/gillicious

SHARDING MYSQL VIA CELLULAR MITOSIS AT OKTA
Will Gunty
Wednesday 3:30 pm
Ballroom B

MySQL BREAK/FIX LAB
Significant Performance Issues

Presented by: Miklos Szel, Gillian Gunson, Brian Cain

April 13, 2015

Pythian®
love your data

Agenda - Part 3

- System Bottlenecks
  - Examine Operating System metrics
  - Run diagnostics

- MySQL Bottlenecks
  - MySQL client (CLI)
  - MySQL tools
  - External tools

- Configuration Setting Changes
  - Dynamic
  - Static

System Bottlenecks Explained

- Trends
  - Memory Utilization
  - CPU Utilization
  - Disk Utilization
  - Network Utilization
- Current Status
  - High Load
  - Swapping
  - I/O Wait

What to look at first?

- Sudden performance issues - what changed?
  - Software release
  - Drive failure, temperature warnings
  - Database schema and configuration changes
  - OS patches, packages, updates
- Issues over time - study the graphs
  - Application servers added or user traffic increase
  - Swap space usage
  - Drive rebuilding
  - Memory leaks
  - Table growth
  - Buffer pool size + overhead

Where to dig

- Operating System Diagnostics
  - top
  - iostat
  - ps
  - dmesg
  - ulimit

OS stats – top

- nice top - defer priority to the issue
  - 1 - what are the cores doing
  - -u mysql - what is mysql doing
  - H - show threads
  - < > - change sort order
  - R - reverse sort order

OS stats – iostat

- iostat -y -x 3
  - -y - throw away stats from the last system boot
  - -x - extended statistics
  - 3 - interval seconds

```
[CPU Utilization Report]
avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           2.65    0.00    0.50    0.14    0.00   96.71

[Device Utilization Report]
Device:         rrqm/s   wrqm/s   r/s   w/s    rsec/s    wsec/s avgrq-sz avgqu-sz   await  svctm %util
sda               0.00    63.67  0.00 46.00      0.00    877.33    19.07     0.04    0.87   0.30  1.40
sda1              0.00     0.00  0.00  0.00      0.00      0.00     0.00     0.00    0.00   0.00  0.00
sda2              0.00    63.67  0.00 46.00      0.00    877.33    19.07     0.04    0.87   0.30  1.40
sdb               0.00     1.00  4.00 126.33   128.00  41238.00   317.39     1.95   14.95   0.50  6.53
```

OS stats – ps

- ps -efww
  - -e - all processes
  - -f - full format listing with command arguments
  - -ww - unlimited width

```
[root@localhost ~]# ps -efww | grep -i mysql | grep -v grep
root      2147     1  0 07:14 ?        00:00:00 /bin/sh /usr/bin/mysqld_safe --datadir=/var/lib/mysql --pid-
file=/var/lib/mysql/localhost.localdomain.pid
mysql     2349  2147  0 07:14 ?        00:00:00 /usr/sbin/mysqld --basedir=/usr --datadir=/var/lib/mysql --
plugin-dir=/usr/lib64/mysql/plugin --user=mysql --log-error=/var/log/mysqld.log --pid-
file=/var/lib/mysql/localhost.localdomain.pid --socket=/var/lib/mysql/mysql.sock
```

OS stats – dmesg

- dmesg [-T]
  - -T - if available prints a human readable timestamp

```
[root@localhost ~]# dmesg
Initializing cgroup subsys cpuset
Initializing cgroup subsys cpu
Linux version 2.6.32-504.8.1.el6.x86_64 (mockbuild@c6b9.bsys.dev.centos.org) (gcc version 4.4.7 20120313
(Red Hat 4.4.7-11) (GCC) ) #1 SMP Wed Jan 28 21:11:36 UTC 2015
Command line: ro root=/dev/mapper/VolGroup-lv_root rd_NO_LUKS LANG=en_US.UTF-8 rd_NO_MD
rd_LVM_LV=VolGroup/lv_swap SYSFONT=latarcyrheb-sun16 crashkernel=auto rd_LVM_LV=VolGroup/lv_root
KEYBOARDTYPE=pc KEYTABLE=us rd_NO_DM rhgb quiet
…….
8021q: adding VLAN 0 to HW filter on device eth0
vboxsf: Successfully loaded version 4.3.10 (interface 0x00010004)
eth0: no IPv6 routers present
```

OS stats – ulimit

- ulimit
  - /etc/security/limits.conf and /etc/security/limits.d/
  - open_files_limit
  - echo -n 'Max processes=low_value:high_value'

```
[root@localhost ~]# ulimit -a
...
open files                      (-n) 1024
...
max user processes              (-u) 3766
...

[root@localhost ~]# for pid in $(ps -e|grep mysql|grep -v safe|awk '{print $1}'); do cat /proc/$pid/limits;
done
```

MySQL Tools (built-in)

- mysql client (CLI)
- mysqladmin
- mysqlshow
- mysqlbinlog

MySQL Tools (built-in) - CLI

- mysql -e
  - “show processlist” | grep -v Sleep
  - “show engine innodb status\G” | tail -40
  - “\s”
  - “select @@max_connections”
  - “show variables like ‘%lock%’”

MySQL Tools (built-in) - mysqladmin

- mysqladmin
  - status
  - extended-status
  - variables
  - shutdown

MySQL Tools (built-in) - mysqlshow

- mysqlshow
  - no parameters - list of schemas
  - schema name - list of tables
  - schema and table name - desc
  - --status
  - --count
  - --keys

MySQL Tools (built-in) - mysqlbinlog

- mysqlbinlog binlog_file
  - --base64-output=DECODE_ROWS and --verbose (-v)
  - --start-position and --stop-position
  - --start-datetime and --stop-datetime
  - --server-id

External Tools

- Percona tools
  - pt-config-diff
  - pt-query-digest
  - pt-kill
  - pt-stalk
- tcpdump
- Bash commands

External Tools - pt-config-diff

- pt-config-diff
  - Compare the running settings (localhost) to my.cnf

```
[root@localhost ~]# pt-config-diff localhost /etc/my.cnf
2 config differences
Variable                  localhost.localdomain /etc/my.cnf
========================= ===================== ===========
join_buffer_size          131072                209715200
sort_buffer_size          2097152               209715200
```

External Tools - pt-query-digest

- pt-query-digest - first things first
  - Where is the slow query log
  - Is it logging
  - What is the minimum time

```
[root@localhost ~]# mysql -e "select @@slow_query_log, @@slow_query_log_file, @@long_query_time\G"
*************************** 1. row ***************************
     @@slow_query_log: 1
@@slow_query_log_file: /var/lib/mysql/localhost-slow.log
    @@long_query_time: 10.000000
```

External Tools - pt-query-digest (cont)

- pt-query-digest
  - --since --until - bound the start and end to examine
  - --filter - limit values further such as by host or user

```
[root@localhost ~]# pt-query-digest --since '2015-03-13' --until '2015-03-14' --filter '$event->{host} =~
/localhost/ && $event->{user} =~ /root/' /var/lib/mysql/localhost-slow.log

# 140ms user time, 10ms system time, 24.68M rss, 202.52M vsz
# Current date: Fri Mar 13 14:07:09 2015
# Hostname: localhost.localdomain
# Files: /var/lib/mysql/localhost-slow.log
# Overall: 1 total, 1 unique, 0 QPS, 0x concurrency _____________________
# Time range: all events occurred at 2015-03-13 13:57:48
# Attribute          total     min     max     avg     95%  stddev  median
# ============     ======= ======= ======= ======= ======= ======= =======
# Exec time            11s     11s     11s     11s     11s       0     11s
# Lock time              0       0       0       0       0       0       0
…
#  10s+  ##############################################################
# EXPLAIN /*!50100 PARTITIONS*/
select sleep(11)\G
```

External Tools - pt-kill

- pt-kill
  - --match-command - Query, Sleep
  - --print - show what pt-kill would kill
  - --kill or --kill-query - drop the connection or just the query
  - --victims - by default just the oldest
  - --interval - by default 30 seconds or half of --busy-time

```
[root@localhost ~]# pt-kill --match-command Query --print --victims all --interval 20
# 2015-03-13T14:16:31 KILL 13 (Query 4 sec) select sleep(100)
# 2015-03-13T14:16:51 KILL 13 (Query 24 sec) select sleep(100)
# 2015-03-13T14:17:11 KILL 13 (Query 44 sec) select sleep(100)
```

External Tools - pt-stalk

- pt-stalk
  - --variable - what global status variable to monitor
  - --threshold - trigger value for the variable

```
[root@localhost stalk]# pt-stalk --variable Threads_running --threshold 2
Overwriting PID file /var/run/pt-stalk.pid because its PID (2788) is not running
2015_03_13_14_34_52 Starting /usr/bin/pt-stalk --function=status --variable=Threads_running --threshold=2 --
match= --cycles=5 --interval=1 --iterations= --run-time=30 --sleep=300 --dest=/var/lib/pt-stalk --prefix= --
notify-by-email= --log=/var/log/pt-stalk.log --pid=/var/run/pt-stalk.pid --plugin=
2015_03_13_14_34_58 Check results: status(Threads_running)=3, matched=yes, cycles_true=1
...
2015_03_13_14_35_02 Check results: status(Threads_running)=3, matched=yes, cycles_true=5
2015_03_13_14_35_02 Collect 1 triggered
2015_03_13_14_35_02 Collect 1 PID 5928
2015_03_13_14_35_02 Collect 1 done
2015_03_13_14_35_02 Sleeping 300 seconds after collect
^C2015_03_13_14_35_13 Caught signal, exiting
2015_03_13_14_35_13 Waiting up to 90 seconds for subprocesses to finish...
2015_03_13_14_35_33 Exiting because OKTORUN is false
2015_03_13_14_35_33 /usr/bin/pt-stalk exit status 1
```

External Tools - pt-stalk (cont)

- Volumes of data to examine

```
[root@localhost stalk]# ls /var/lib/pt-stalk/
2015_03_13_14_35_02-df                2015_03_13_14_35_02-netstat
2015_03_13_14_35_02-disk-space        2015_03_13_14_35_02-netstat_s
2015_03_13_14_35_02-diskstats         2015_03_13_14_35_02-opentables1
2015_03_13_14_35_02-hostname          2015_03_13_14_35_02-opentables2
2015_03_13_14_35_02-innodbstatus1     2015_03_13_14_35_02-output
2015_03_13_14_35_02-innodbstatus2     2015_03_13_14_35_02-pmap
2015_03_13_14_35_02-interrupts        2015_03_13_14_35_02-processlist
2015_03_13_14_35_02-iostat            2015_03_13_14_35_02-procstat
2015_03_13_14_35_02-iostat-overall    2015_03_13_14_35_02-procvmstat
2015_03_13_14_35_02-log_error         2015_03_13_14_35_02-ps
2015_03_13_14_35_02-lsof              2015_03_13_14_35_02-slabinfo
2015_03_13_14_35_02-meminfo           2015_03_13_14_35_02-sysctl
2015_03_13_14_35_02-mpstat            2015_03_13_14_35_02-top
2015_03_13_14_35_02-mpstat-overall    2015_03_13_14_35_02-trigger
2015_03_13_14_35_02-mutex-status1     2015_03_13_14_35_02-variables
2015_03_13_14_35_02-mutex-status2     2015_03_13_14_35_02-vmstat
2015_03_13_14_35_02-mysqladmin        2015_03_13_14_35_02-vmstat-overall
```

External Tools - tcpdump

- tcpdump
  - capture 3306 traffic for processing with pt-query-digest

```
[root@localhost tcpdump]# time tcpdump -s 65535 -x -n -q -tttt -i any -c 50000 port 3306 > mysql.tcp.txt
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on any, link-type LINUX_SLL (Linux cooked), capture size 65535 bytes
^C36 packets captured
72 packets received by filter
0 packets dropped by kernel

real    0m15.514s
user    0m0.006s
sys     0m0.002s
```

External Tools - tcpdump (cont)

- Process the results with pt-query-digest as a slow-query-log

```
[root@localhost tcpdump]# pt-query-digest --type tcpdump --no-report --timeline --output slowlog mysql.tcp.
txt
...
use mysql_native_password;
select @@version_comment limit 1;
# Time: 150313 14:57:53.042467
# User@Host: root[root] @ 127.0.0.1 []
# Client: 127.0.0.1:39732
# Thread_id: 462
# Query_time: 0.000039  Lock_time: 0.000000  Rows_sent: 0  Rows_examined: 0
use mysql_native_password;
show grants;
# Time: 150313 14:57:53.042551
# User@Host: root[root] @ 127.0.0.1 []
# Client: 127.0.0.1:39732
# Thread_id: 462
# Query_time: 0.000000  Lock_time: 0.000000  Rows_sent: 0  Rows_examined: 0
use mysql_native_password;
# administrator command: Quit;
...
```

External Tools - Bash commands

- Space left

```
[root@localhost ~]# echo $(($(stat -f --format="%a*%s" /var/lib/mysql)/1024/1024))M
14446M
```

- Release swap

```
[root@localhost ~]# swapoff -a && swapon -a
```

- Unused InnoDB tables

```
[root@localhost ~]# find /var/lib/mysql/ -name *.ibd -mtime +60 | sort | xargs du -hc
```

Configuration Setting Changes

- Dynamic
  - innodb_io_capacity
  - innodb_lock_wait_timeout
  - lock_wait_timeout
  - query_cache_size
  - table_open_cache
- Static
  - innodb_buffer_pool_size
  - innodb_buffer_pool_instances
  - open_files_limit
  - skip_name_resolve
  - tmpdir

Thank you!

Follow me at

- Linked in: www.linkedin.com/in/dmnscrg