Mysql Innodb Hotcopy

- Written in Python
- Source code available under the Apache software license here : http://bitbucket.org/david415/innodb-hotcopy/

Mysql Innodb Hotcopy in Python project goals

- restore a mysql replica from another replica in the same shard
  - Transactional model: if any exception is thrown then rollback() to the original state using my Python closure queue
  - Originally atomicity was guarantied by the innodb_freeze I isolated from Google's monolithic Mysql patch. Currently it uses LVM but I could write modules to use ZFS or whatever else can take snapshots...
  - So far only rsyncd is used to transfer files but other mechanisms could be added (ssh, netcat???)

General Procedure of Innodb Hotcopy

- Source replica: Gradually lock up the source database (and stop replication) before taking the snapshot...
- mount snapshot and transfer to target replica
- Unmount and Destroy snapshot
- Configure and startup mysqld on target

Precarious state changes on the source replica database server...

- If the hotcopy operation were to fail without “rolling back the transaction” then the systems administrator would be stuck with cleaning up state changes on the souce replica such as:
  - An LVM snapshot exists causing the logical volume to copy-on-write degrading performance
  - snapshot is mounted
  - Mysql is in single user mode
  - replication is stopped etc...

Restore/Rollback queue

- With each state change we push the undo operation onto the queue
- Originally implemented as a closure queue...
- using a lambda like this:
- queue.push(
- lambda:dbh_source.start_replication(), 'start_replication')
- The id string is pushed onto the queue as well so that we can exec_remove() items from the queue in an arbitrary order

```
def rollback(self):
        traceback.print_exc()
        while len(self.queue) != 0:
            func, name = self.queue.pop()
            print("Rollback: %s" % name)
            func()
```

Future project additions...

- Automatically invoke replica hotcopy via a highly available distributed network voting protocol
- Patch Innodb storage engine: make crash recovery faster (Spinn3r is probably going to throw down some dow to get this done)
- A distributed/highly available persistent storage mechanism for the restore queue to allow a rollback even after the server running the hotcopy program, crashes.

A related project... Slave promotion/Master demotion

- This is a related project and will soon be open-sourced
- This program can handle multiple replicas in shard, but in order to re-parent slaves (e.g. in a shard with 3 or more replicas) that are behind in replication, Google's Mysql Global Transaction Id code is needed...
  - Global Transaction Ids hasn't yet been isolated from from Google's monolithic V2 Mysql patch