PERCONA
Performance Consulting Experts

10 Percona Toolkit tools every MySQL DBA should know about

Fernando Ipar - Percona Webinar Dec/2012

About me

- Fernando Ipar
- Consultant @ Percona
- fernando.ipar@percona.com

About this presentation

- Introductory level
- Tool selecction based on frequency of use
- Presented by category
  - Replication management
  - Performance optimization
  - Operations
  - Root cause analysis
- Q/A at the end

PERCONA
Performance Consulting Experts

About Percona Toolkit

- Actual customers problems
- Extensive test coverage
- Works with every version since 5.0
  - Some tools with 4.1 too
- Good community
- Covered by our Support services

Before we begin

Knowing your DSNs

PERCONA
Performance Consulting Experts

(This includes all of today's tools except pt-stalk and pt-sift)

PERCONA
Performance Consulting Experts

DSNs in a nutshell

- comma separated K/V list
- h=localhost,u=root,p=s3cr3t
- For multiple hosts, specific DSN inherit from others
  - So your life will be easier if you keep your credentials consistent across hosts
- Full story: http://bit.ly/percona-toolkit-dsn-spec

Before running a tool

- Read the manual carefully
- Test
- Have a tested backup available.

Replication management tools

pt-table-checksum

Determines if a master and its replicas have a consistent copy of the dataset

Why would a replica **not** be consistent?

PERCONA
Performance Consulting Experts

- Writing directly to it
- Using SBR
- Bad coordinates after server crash
- <your reason here>

PERCONA
Performance Consulting Experts

How does it work?

- Uses STATEMENT based replication
  - Does not change other sessions
- Runs checksum queries against master
  - Waits for them to replicate to slaves
  - Checks for differences in the results

The scenario

Master A, replicas B and C

PERCONA
Performance Consulting Experts

```
pt-table-checksum --replicate
    percona.checksums h=A
```


PERCONA
Performance Consulting Experts

Yes, it's **that** simple!

Some considerations:

- Schema differences may break replication
- Will self-throttle checking replica lag
  - But don't leave unattended

pt-table-sync

If pt-table-checksum gives you the bad news, pt-table-sync helps you go back to a sane state.

PERCONA
Performance Consulting Experts

The scenario

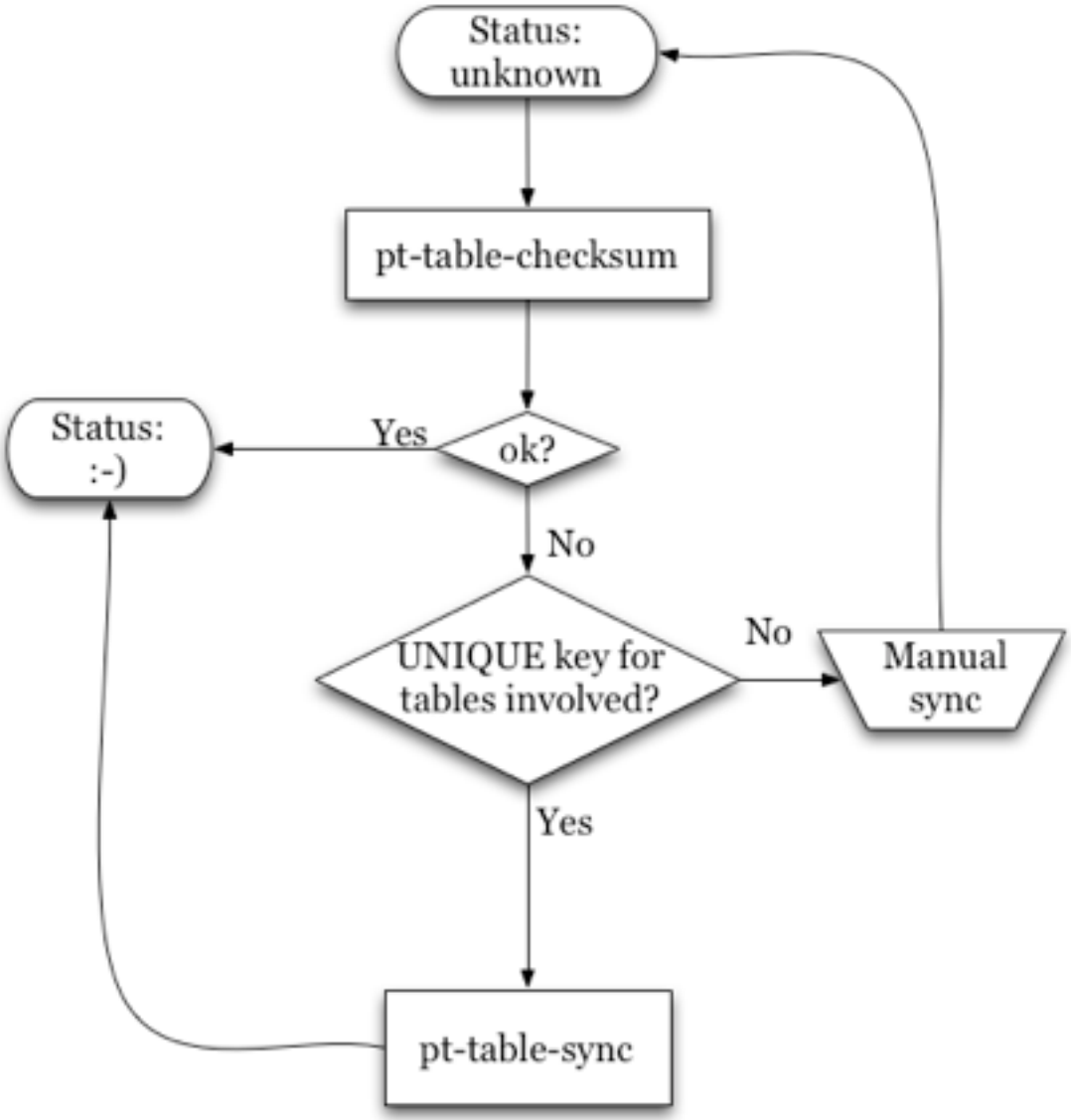
Master A, replicas B and C
pt-table-checksum found differences on C.

```
pt-table-sync --replicate
percona.checksums --print h=A
```


PERCONA
Performance Consulting Experts

Happy with what you see? Then use/add --execute

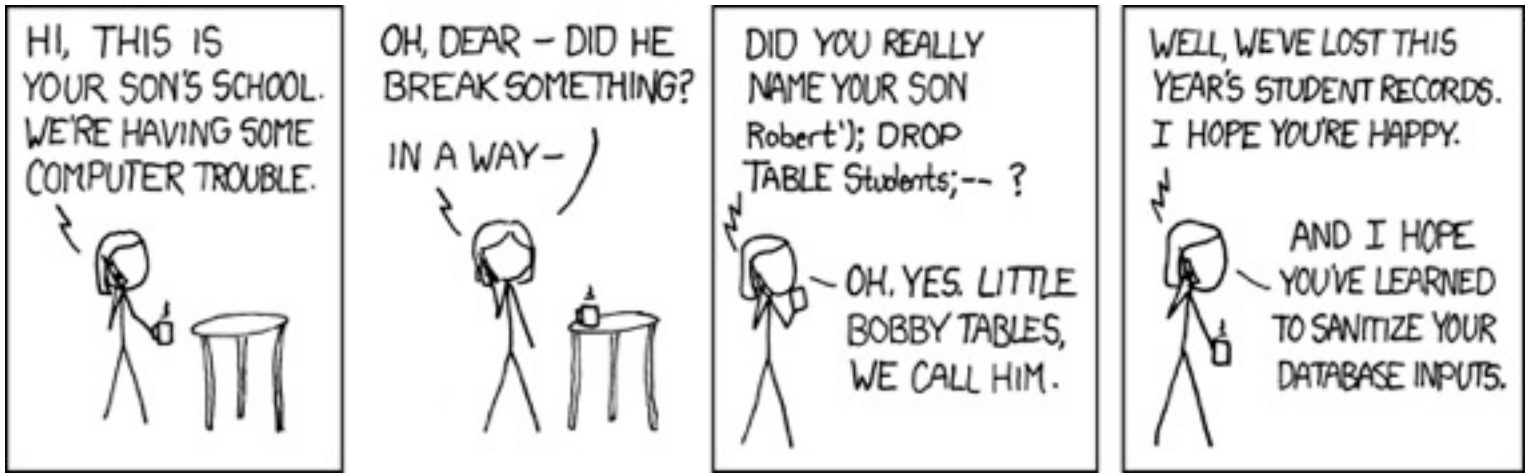
PERCONA
Performance Consulting Experts


Status:
unknown
pt-table-checksum
Status:
:-)
Yes
ok?
No
UNIQUE key for
tables involved?
No
Manual
sync
Yes
pt-table-sync

pt-slave-delay

Intentionally keep a slave behind

PERCONA
Performance Consulting Experts



(http://xkcd.com/327/ is one reason why a slave is **not** a backup)

```
pt-slave-delay --delay 2h h=slave
```

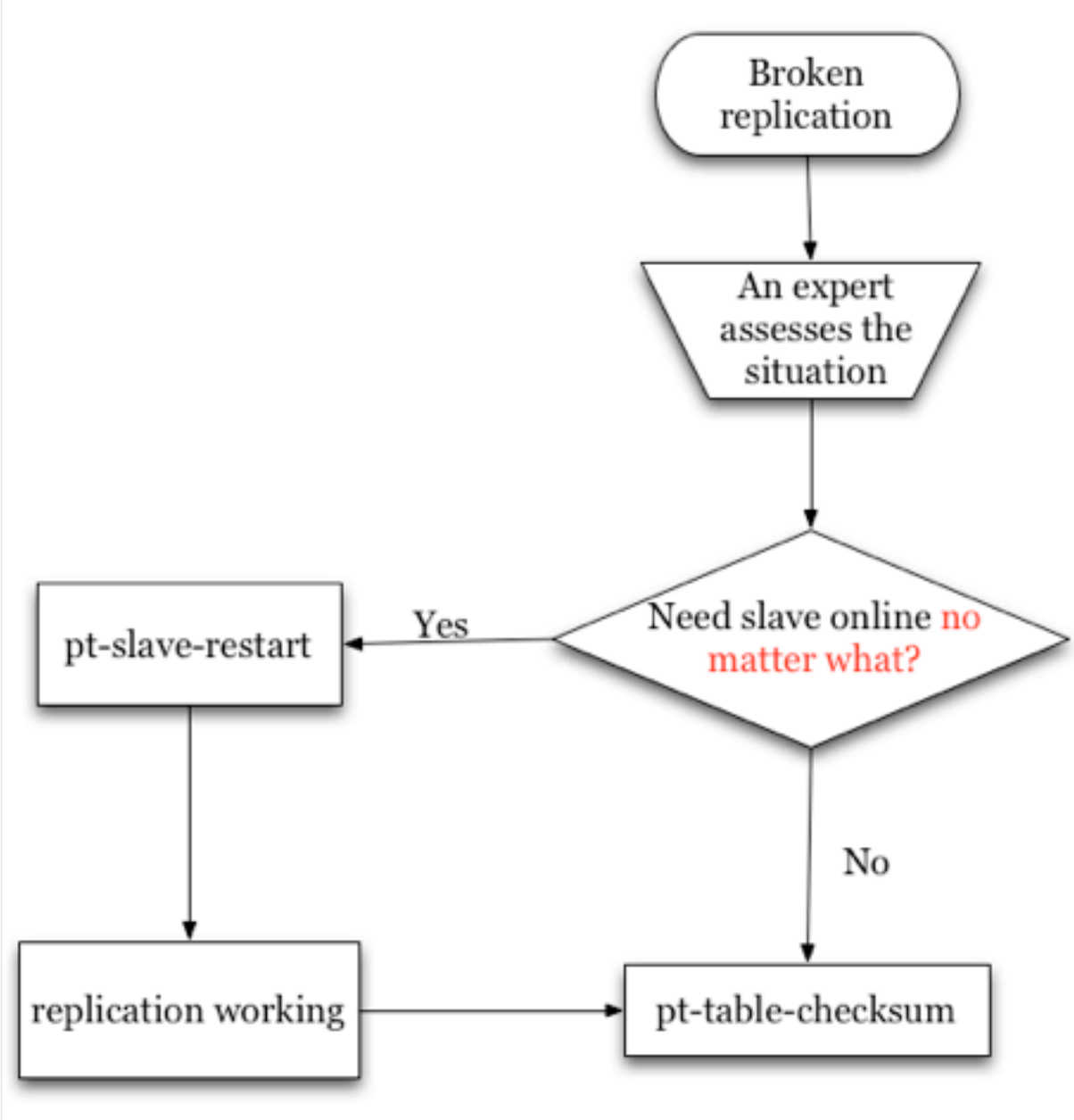

PERCONA
Performance Consulting Experts

pt-slave-restart

- Automatically skip replication errors
- Only as a last resort!
  - Or when you know what you're doing

```
pt-table-restart h=slave
```


PERCONA
Performance Consulting Experts


Broken replication
An expert assesses the situation
Need slave online no matter what?
Yes
pt-slave-restart
No
replication working
pt-table-checksum

Important

When using SBR, skipping errors will usually make a bad situation worse.

pt-heartbeat

- Reliably measure replication lag
- Works with Percona Monitoring Plugins
  - http://www.percona.com/software/percona-monitoring-plugins

On the master

```
pt-heartbeat -D percona --create-table
             --update h=mater
```


PERCONA
Performance Consulting Experts

From nagios/etc

```
pt-heartbeat -D percona --check
h=replica
```

Performance optimization tools

pt-query-digest

- Analyze MySQL queries
- Discover optimization opportunities
- Prevent scalability bottlenecks

```
pt-query-digest <path-to-log>
```


PERCONA
Performance Consulting Experts

```
telecaster:rsandbox_5_5_12 fernandoipar$ pt-query-digest master/data/telecaster-slow.log
master/data/telecaster-slow.log:  85% 00:04 remain

# 34.9s user time, 280ms system time, 21.54M rss, 2.34G vsz
# Current date: Mon Nov 12 13:36:58 2012
# Hostname: telecaster.local
# Files: master/data/telecaster-slow.log
# Overall: 246.84k total, 8 unique, 2.81k QPS, 0.40x concurrency _________
# Time range: 2012-11-12 13:34:47 to 13:36:15
# Attribute          total     min     max     avg     95%  stddev  median
# ============     ======= ======= ======= ======= ======= ======= =======
# Exec time            35s     1us     6ms   141us   568us   201us    20us
# Lock time             4s       0   850us    15us    60us    22us       0
# Rows sent          6.41M       0     195   27.24  192.76   63.74       0
# Rows examine       6.75M       0     195   28.67  192.76   63.31       0
# Query size        31.67M      14     869  134.55  833.10  269.74   31.70

# Profile
# Rank Query ID           Response time Calls R/Call Apdx V/M   Item
# ==== ================== ============= ===== ====== ==== ===== ==========
#    1 0xC69B6ED2C47380A4 18.8741 54.0% 31568 0.0006 1.00  0.00 SHOW VARIABLES
#    2 0xA2750AF24EA2AEE6 10.4547 29.9% 31568 0.0003 1.00  0.00 SHOW COLLATION
#    3 0xAC397D6B2A75526D  3.0548  8.7% 31567 0.0001 1.00  0.00 SELECT test
#    4 0xE4CF7146873CCC28  0.6794  1.9% 31567 0.0000 1.00  0.00 SET
#    5 0x38B3D80280BBFA2A  0.6432  1.8% 31566 0.0000 1.00  0.00 SET
# MISC 0xMISC              1.2615  3.6% 89007 0.0000   NS   0.0 <3 ITEMS>
```

This is the only business query in this capture!

```
# Query 1: 358.73 QPS, 0.21x concurrency, ID 0xC69B6ED2C47380A4 at byte 48156807
# This item is included in the report because it matches --limit.
# Scores: Apdex = 1.00 [1.0], V/M = 0.00
# Query_time sparkline: |   ^_     |
# Time range: 2012-11-12 13:34:47 to 13:36:15
# Attribute    pct   total     min     max     avg     95%  stddev  median
# ============ === ======= ======= ======= ======= ======= ======= =======
# Count         12   31568
# Exec time     53     19s   542us     6ms   597us   725us    91us   568us
# Lock time     52      2s    54us   537us    64us    89us    13us    57us
# Rows sent      7 524.08k      17      17      17      17       0      17
# Rows examine   7 524.08k      17      17      17      17       0      17
# Query size    82  26.16M     869     869     869     869       0     869
# String:
# Databases    test
# Hosts        localhost
# Users        msandbox
# Query_time distribution
#   1us
#  10us
# 100us  ################################################################
#   1ms  #
#  10ms
# 100ms
#    1s
#  10s+
/* mysql-connector-java-5.1.6 ( Revision: ${svn.Revision} ) */SHOW VARIABLES WHERE Variable_name ='language'
= 'net_write_timeout' OR Variable_name = 'interactive_timeout' OR Variable_name = 'wait_timeout' OR Variable
_set_client' OR Variable_name = 'character_set_connection' OR Variable_name = 'character_set' OR Variable_na
t_server' OR Variable_name = 'tx_isolation' OR Variable_name = 'transaction_isolation' OR Variable_name = 'c
ts' OR Variable_name = 'timezone' OR Variable_name = 'time_zone' OR Variable_name = 'system_time_zone' OR Va
er_case_table_names' OR Variable_name = 'max_allowed_packet' OR Variable_name = 'net_buffer_length' OR Varia
de' OR Variable_name = 'query_cache_type' OR Variable_name = 'query_cache_size' OR Variable_name = 'init_con
```

Filtering

```
pt-query-digest --filter filter.pl <path-
to-log>
```

filter.pl:

```
return ($event->{fingerprint} =~ m/users/)
```

Reviews

```
pt-query-digest --create-review-table
--review D=percona,t=reviews  <path-to-
log>
```

http://github.com/box/Anemometer makes good use of this feature

```
# Review information
#    first_seen: 2012-11-14 10:35:33
#     last_seen: 2012-11-14 10:35:46
#   reviewed_by: fipar
#   reviewed_on: 2012-11-14 10:38:45
#      comments: table scan, looks like someone was running mysqlslap against production
# Tables
#    SHOW TABLE STATUS FROM `mysqlslap` LIKE 't1'\G
#    SHOW CREATE TABLE `mysqlslap`.`t1`\G
# EXPLAIN /*!50100 PARTITIONS*/
SELECT intcol1,charcol1 FROM t1\G
```

If this is not NULL, the tool won't display this query again unless --report-all is used

Operations tools

pt-upgrade

- Compare query results & run time against different instances
- Part of proper version upgrade testing

```
pt-upgrade h=host1 h=host2 queries.txt
```


PERCONA
Performance Consulting Experts

```
# Query 1: ID 0xC479001956B2A7BE at byte 0
# host1: 127.0.0.1:5527
# host2: 127.0.0.1:18967
# Found 1 differences in 1 samples:
#   checksums       0
#   column counts   0
#   column types    0
#   query times     1
#   row counts      0
#   warning counts  0
#   warning levels  0
#   warnings        0
#             host1 host2
# Errors          0     0
# Warnings        0     0
# Query_time
#   sum         16ms 100ms
#   min         16ms 100ms
#   max         16ms 100ms
#   avg         16ms 100ms
#   pct_95      16ms 100ms
#   stddev         0     0
#   median      16ms 100ms
# row_count
#   sum           10    10
#   min           10    10
#   max           10    10
#   avg           10    10
#   pct_95        10    10
#   stddev         0     0
#   median        10    10
use `sakila`;
select country from sakila.city join sakila
```

pt-online-schema-change

- Minimize impact of ALTERing tables
- Be careful with foreign keys!
  - They are handled, but do read the manual first

```
pt-online-schema-change --alter-foreign-keys-method auto --alter "add key actor_last_update (last_update)" --execute h=localhost,D=sakila,t=actor
```

```
Child tables:
  `sakila`.`film_actor` (approx. 5525 rows)
Will automatically choose the method to update foreign keys.
Altering `sakila`.`actor`...
Creating new table...
Created new table sakila._actor_new OK.
Altering new table...
Altered `sakila`.`_actor_new` OK.
Creating triggers...
Created triggers OK.
Copying approximately 200 rows...
Copied rows OK.
Max rows for the rebuild_constraints method: 57078
Determining the method to update foreign keys...
  `sakila`.`film_actor`: 5525 rows; can use rebuild_constraints
Swapping tables...
Swapped original and new tables OK.
Rebuilding foreign key constraints...
Rebuilt foreign key constraints OK.
Dropping old table...
Dropped old table `sakila`.`_actor_old` OK.
Dropping triggers...
Dropped triggers OK.
Successfully altered `sakila`.`actor`.
```

Root cause analysis tools

pt-stalk

- Helps diagnose hard-to-catch problems
- 'Random' stalls

```
pt-stalk --function processlist \
    --variable State \
    --match statistics --threshold 10
```


PERCONA
Performance Consulting Experts

```
pt-stalk --function processlist \
    --variable Command \
    --match Sleep --threshold 155 \
    --cycles 0
```

```
pt-stalk --threshold 40 \
    --cycles 6
```


PERCONA
Performance Consulting Experts

```
pt-stalk --function custom-check.sh \
         --threshold 12
```

`custom-check.sh` has to provide a `trg_plugin` function, which must output a number.

pt-stalk --no-stalk

pt-sift

- High level overview of pt-stalk data

```
--diskstats--
  #ts device    rd_s rd_avkb rd_mb_s rd_mrg rd_cnc   rd_rt    wr_s wr_avkb wr_mb_s wr_mrg wr_cnc   wr_rt busy in_p
 {29} sdd1       0.0     0.0     0.0     0%    0.0     0.0    39.6     4.1     0.2     0%    0.1     3.3  13%
 sdd1  0% 25% . . 15% 0% . . . . . . . . . . . . . . . . . . . . . . . .
--vmstat--
  r b swpd      free   buff   cache si so bi  bo   in    cs us sy  id wa st
100 1 5216 177085184 346908 6741672  0  0 28 162    0     0  0  0 100  0  0
  0 0 5216 177490208 346956 6750640  0  0  0 642 1143 10521  1  2  96  1  0
wa 0% . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
--innodb--
    txns: 96xACTIVE (3s)
    18 queries inside InnoDB, 0 queries in queue
    Main thread: sleeping, pending reads 0, writes 0, flush 0
    Log: lsn = 737633786, chkp = 737388727, chkp age = 245059
    Threads are waiting at:
    216 trx/trx0trx.c line 1711
      3 trx/trx0trx.c line 807
      1 trx/trx0trx.c line 432
    Threads are waiting on:
--processlist--
    State
    556  update
     84  Sending data
     29  freeing items
      2
      1  NULL
    Command
    671  Query
      2  Sleep
--stack traces--
    No stack trace file exists
--oprofile--
    No opreport file exists
```

mutex_enter(&kernel_mutex);

The take home message

- Don't reinvent the wheel
- We've been burned, so you don't have to

Thank you!

PERCONA
Performance Consulting Experts

Percona Live

Percona Live MySQL Conference & Expo

April 22-25, 2013

Santa Clara Convention Center & Hyatt Regency Santa Clara

Join the MySQL community for 4 days of breakout sessions, tutorials, and keynotes

Mingle with the MySQL community at the Welcome Reception and the Community Networking Reception

Visit www.percona.com/live for more information

PERCONA
Performance Consulting Experts