Percona Toolkit

Baron Schwartz

Agenda

- Explain each tool's purpose and use

- Answer your questions

- Slides will be online later

- Do live demos? Maybe

Utility

- General-purpose helper tools

pt-find

- Purpose: Find tables and take actions

- Like the unix 'find' command

- Works from SHOW TABLES

  - Does not use INFORMATION_SCHEMA

  - May use SHOW TABLE STATUS

```
pt-find --mtime +30 --engine InnoDB \
  --exec "ALTER TABLE %D.%N ENGINE=MyISAM"
```

pt-kill

- Watch SHOW PROCESSLIST

- Select processes, or groups of processes

- Apply filters

- Take a specified action

-

```
pt-kill –busy-time 60 --kill --interval 30
```

pt-tcp-model

- Ingest tcpdump data

- Analyze and transform it

- Come to my talk on "scalability and performance analysis with TCP data"

Unix-Related

- Not specific to MySQL

- Useful for arbitrary tasks

pt-fifo-split

- Read lines from a file

- Print them to a FIFO file

- "End of file" every N lines

- Useful for loading large files into MySQL

pt-ioprofile

- Dangerous tool!

- Attaches strace to a process

- Captures IO-related system calls

- Generates a profile of IO activity

pt-pmp

- Attaches GDB to a process

- Prints out stack traces

- Aggregates them

- Useful for bottleneck analysis

pt-align

- Aligns sloppy output

- For example: vmstat, iostat

Do Things

- Do useful, important stuff with MySQL

pt-archiver

- Retrieve rows from a table

- Write them to another table, or a file

- Delete them from the first table

- Many safety checks and features

- Extensible via plugins

pt-log-player

- For load simulation, benchmarking

- Splits a query log into several

- Replays them multi-threadedly

- Probably obsoleted by Percona Playback

  - See Stewart's talk

pt-online-schema-change

- Perform ALTER without locking

- Creates a new table

- Copies rows to it

- Uses triggers to track updates to original table

- Swaps original/new table

- Many safety/sanity features

Learn Things

- Inspect your systems and servers

- Analyze VARIABLES, STATUS, schema

pt-summary

- Report on system config

- Goals: diff-able, email-able

- Easily extensible

pt-mysql-summary

- Like pt-summary, but for MySQL

pt-config-diff

- Compare two servers

- Compare a server to its my.cnf

pt-variable-advisor

- "Best practices" analysis of configuration

- Works from SHOW VARIABLES

pt-duplicate-key-checker

- Analyze indexes and foreign keys

- Find duplicate/redundant

- Print SQL that can remove them

pt-mext

- Format SHOW STATUS samples

- Make it easier to see what's changing

Record Things

pt-deadlock-logger

- Keep track of how often deadlocks happen

- Analyze which tables and indexes are involved

pt-fk-error-logger

- Ditto, but for foreign keys instead

pt-show-grants

- Print GRANT statements for all users

pt-diskstats

- A replacement for iostat -dx

Replication-Related

pt-heartbeat

- Measure replication delay accurately

- Place "quasi-global transaction IDs" into binlogs

- Ease recovery on large and complex topologies

pt-slave-delay

- Intentionally delay replication

- Works by starting and stopping slave SQL thread

pt-slave-find

- Automatically discover all servers in a replication relationship

- Print a summary of their status and configuration

pt-slave-restart

- Restart replication when it fails

- Please use with care!

    - Skipping errors is not good practice

    - Useful in dire circumstances

pt-table-checksum

- Determine whether replicas have logically the same data as master

- See Daniel's talk

- See a recording of my webinar

pt-table-sync

- Repair (online!) an out-of-sync replica

- Use with care! It does change data

- Also possible to do multi-server sync with conflict detection, etc

Query-Related

pt-query-advisor

- Analyze a query and point out problems with the SQL

- Analyze an entire log file

- See also tools.percona.com GUI interface

pt-index-usage

- Make a catalog of tables and indexes

- Analyze a log of queries

- Report on indexes that are unused, queries that have several execution plans, etc

pt-query-digest

- Retrieve queries from some source

- Pass them through a "pipeline" of filters/transforms

- Aggregate them and generate a report

- Do lots of other nifty/crazy things

pt-visual-explain

- Analyze EXPLAIN and generate a tree-formatted view

- "Reverse engineer" the execution plan

pt-upgrade

- Execute a log of queries on two servers and compare

- Highlight resultset changes, execution time changes, errors, etc

- See Peter's talk about upgrading MySQL

pt-trend

- Obsolete tool, likely to be removed

- Slices a query log and reports aggregate stats from segments of it

Troubleshooting

pt-stalk

- Essential tool for diagnosing things

- Many customers run it 24x7 to catch unforeseen problems

- Watches the server and collects information when there is a problem

- See my webinar "troubleshooting intermittent problems"

pt-sift

- Important for helping analyze pt-stalk's collected data

- Not complete; just a helper script!

- Useful for analyzing a lot of collected samples