

Ежедневная работа с MySQL и Percona Tools

Как понять что происходит с сервером под высокой нагрузкой?

Ответы дает Percona Toolkit

Владимир Федорков, Frédéric Descamps



<http://www.devconf.ru>

О чем будем говорить

- Как упростить себе жизнь
 - Как узнать что происходит на сервере
 - Как проверить MySQL сервер
 - В том числе репликацию
 - В том числе консистентность слейвов
 - И многое другое
- Какие утилиты когда нужны
 - И что они показывают

Минута саморекламы

- Связался высокими нагрузками в 2006-м году и покатился.
 - Успел поработать в Перконе
- Performance geek
 - Твиты читать тут: twitter.com/vfedorkov
 - Фоточки смотреть здесь: fb.com/vfedorkov
 - Пытаюсь писать в блог <http://astellar.com>

Percona Toolkit

- Изначально Maatkit и Aspersa
 - Придуманы и разработаны Baron Schwartz и Daniel Nichter
- Opensource GNU GPL v2
- <https://launchpad.net/percona-toolkit>
- <http://www.percona.com/software/percona-toolkit>

Другие тулкиты



Percona toolkit



А конкретнее

```
$ ls
pt-align
pt-archiver
pt-collect
pt-config-diff
pt-deadlock-logger
pt-diskstats
pt-duplicate-key-checker
pt-fifo-split
pt-find
pt-fk-error-logger
pt-heartbeat
pt-index-usage
pt-ioprofile
pt-kill
pt-log-player
pt-mext
pt-mysql-summary
pt-online-schema-change
pt-pmp
pt-query-advisor
pt-query-digest
pt-show-grants
pt-sift
pt-slave-delay
pt-slave-find
pt-slave-restart
pt-stalk
pt-summary
pt-table-checksum
pt-table-sync
pt-tcp-model
pt-trend
pt-upgrade
pt-variable-advisor
pt-visual-explain
```

Установка

- Требования:
 - Perl, DBI, DBD::mysql, Term::ReadKey
- Загрузка
 - `wget percona.com/get/percona-toolkit.rpm`
 - Доступен через репозиторий Перконы
- Все утилиты доступны по отдельности
 - `wget percona.com/get/pt-diskstats`

Открываем для себя железо и
операционную систему

Знание — сила!

- Сколько памяти свободно? Чем занята? Можно добавить еще?
- Сколько на машине ядер? Какая у них скорость?
- Какая битность у операционки? Совпадает ли с железом?
- Включен ли SELinux?
- Рейд стоит? Сколько дисков? Каких? Память? VBU?
- Что с сетью? `ip_local_port_range`? Какие порты слушают?
- Версия операционки? Ядро?
- Какой стоит IO scheduler?
- Какие типы файл системы используется?

Summary!

Name: pt-summary

Function: показывает всю основную информацию о системе

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-summary.html>

pt-summary

```
$ wget percona.com/get/pt-summary
$ chmod +x pt-summary
$ ./pt-summary
# Percona Toolkit System Summary Report #####
    Date | 2012-02-02 12:25:18 UTC (local TZ: EET +0200)
  Hostname | db01
    Uptime | 42 days, 16:14, 1 user, load average: 1.15, 1.22, 1.08
    System | HP; ProLiant DL580 G7; vNot Specified (Rack Mount Chassis)
Service Tag | FOSDEM2012
  Platform | Linux
    Release | CentOS release 6.1 (Final)
    Kernel | 2.6.32-131.21.1.el6.x86_64
Architecture | CPU = 64-bit, OS = 64-bit
  Threading | NPTL 2.12
  Compiler | GNU CC version 4.4.5 20110214 (Red Hat 4.4.5-6).
  SELinux | Disabled
Virtualized | No virtualization detected
# Processor #####
  Processors | physical = 4, cores = 40, virtual = 80, hyperthreading = yes
    Speeds | 80x2397.393
    Models | 80xIntel(R) Xeon(R) CPU E7- 4870 @ 2.40GHz

[ Further output omitted ]
```

Открываем для себя MySQL

Что ты знаешь о своей базе?

- Версия? Билд? Сколько баз? Где лежат?
- Что с репликацией? Сколько реплик? Какие?
- Как настроена InnoDB?
- Какие еще Storage Engines есть?
- Какие типы полей используются? Какие индексы?
- Внешние ключи? Views? Триггеры?
- Полнотекстовый поиск? Partitioning?
- У кого какие привелегии?

Еще одно Summary

Name: `pt-mysql-summary`

Function: Основная информация о MySQL сервере

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-mysql-summary.html>

pt-mysql-summary

```
$ wget percona.com/get/pt-mysql-summary
$ chmod +x pt-mysql-summary
$ ./pt-mysql-summary
# Percona Toolkit MySQL Summary Report #####
      System time | 2012-02-02 12:24:37 UTC (local TZ: EET +0200)
# Instances #####
Port  Data Directory          Socket
=====
      /mysql
# Report On Port 3306 #####
      User | root@localhost
      Time | 2012-02-02 14:24:38 (EET)
      Hostname | db01
      Version | 5.5.18-55-log Percona
      Built On | Linux x86_64
      Started | 2011-12-26 21:09 (up 37+17:15:35)
      Databases | 6
      Datadir | /mysql/
      Processes | 10 connected, 4 running
      Replication | Is a slave, has 1 slaves connected
      Pidfile | /mysql/mysqld.pid (exists)

[ Further output omitted ]
```


Мастер? Слейв?

Name: `pt-slave-find`

Function: находит и выдает
иерархию слейвов

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-slave-find.html>

pt-slave-find

```
$ ./pt-slave-find localhost
localhost
Version          5.5.18-55-log
Server ID        11
Uptime           37+17:45:50 (started 2011-12-26T21:09:03)
Replication      Is a slave, has 1 slaves connected, is not read_only
Filters
Binary logging   MIXED
Slave status     0 seconds behind, running, no errors
Slave mode       STRICT
Auto-increment   increment 2, offset 1
InnoDB version   1.1.8-rel23.0
+- 10.0.0.17
  Version          5.5.18-55-log
  Server ID        12
  Uptime           27+21:02:35 (started 2012-01-05T17:52:18)
  Replication      Is a slave, has 1 slaves connected, is not read_only
  Filters
  Binary logging   MIXED
  Slave status     0 seconds behind, running, no errors
  Slave mode       STRICT
  Auto-increment   increment 2, offset 2
  InnoDB version   1.1.8-rel23.0
```

Чем занимаются диски?

- `pt-diskstats` улучшенный `iostat`
 - Интерактивный
 - Понятный
- `pt-ioprofile` меряет кто и как использует диск
 - Какие файлы использует MySQL? Сколько времени проводится в чтении, записи и т.р.
 - Сколько запросов на чтение/запись?
 - Насколько они большие?

ДИСКОВЫЙ I/O

Name: `pt-diskstats`

Function: агрегирует и суммирует

`/proc/diskstats`

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-diskstats.html>

pt-diskstats

```
# ./pt-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_prg io_s qtime stime
{1} sdb 158.0 5.8 0.9 0% 0.8 4.9 392.1 24.3 9.3 51% 0.6 0.7 54% 0 550.1 0.9 0.6
{1} sdb1 158.0 5.8 0.9 0% 0.8 4.9 392.1 24.3 9.3 51% 0.6 0.7 54% 0 550.1 0.9 0.6

{2} sda 0.5 4.0 0.0 0% 0.0 8.0 3.0 59.3 0.2 93% 0.0 0.2 1% 0 3.5 0.0 0.3
{2} sda2 0.0 0.0 0.0 0% 0.0 0.0 1.0 16.0 0.0 75% 0.0 1.0 0% 0 1.0 0.0 1.0
{2} sda5 0.5 4.0 0.0 0% 0.0 8.0 2.0 81.0 0.2 95% 0.0 0.1 1% 0 2.5 0.0 0.2
{2} sdb 150.0 8.7 1.3 3% 1.1 6.9 448.5 40.6 17.8 84% 1.4 0.5 53% 0 598.6 0.6 0.2
{2} sdb1 143.0 8.9 1.2 3% 1.0 6.8 400.5 24.7 9.7 49% 0.7 0.9 51% 0 543.6 1.3 0.5
{2} sdb2 7.0 4.0 0.0 0% 0.1 10.0 48.0 173.0 8.1 98% 0.7 0.3 9% 0 55.0 0.3 0.0

wr_s wr_avkb wr_mb_s wr_mrg wr_cnc wr_rt
392.1 24.3 9.3 51% 0.6 0.7
392.1 24.3 9.3 51% 0.6 0.7

3.0 59.3 0.2 93% 0.0 0.2
1.0 16.0 0.0 75% 0.0 1.0
2.0 81.0 0.2 95% 0.0 0.1
448.5 40.6 17.8 84% 1.4 0.5
400.5 24.7 9.7 49% 0.7 0.9
48.0 173.0 8.1 98% 0.7 0.3
```

IO profile

Name: `pt-ioprofile`

Function: анализирует ввод-вывод.

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-ioprofile.html>

pt-ioprofile

```
# ./pt-ioprofile
Sat Feb 4 13:44:08 CET 2012
Tracing process ID 24179
  total    read    write    open    close    lseek filename
0.007547 0.000583 0.004245 0.002181 0.000268 0.000270 /tmp/#sql_5e73_1.MYI
0.000686 0.000000 0.000000 0.000420 0.000266 0.000000 /tmp/#sql_5e73_1.MYD
0.000065 0.000000 0.000024 0.000029 0.000012 0.000000 /var/lib/mysql/fred/db.opt
0.000055 0.000000 0.000000 0.000055 0.000000 0.000000 /var/lib/mysql/mysql/db.opt
0.000046 0.000000 0.000000 0.000046 0.000000 0.000000 /var/lib/mysql/test/db.opt
```

Не надо стесняться своей базы данных!

Задумайся

- Как используются индексы
 - Существуют ли они вообще?
- Есть ли дублирующиеся ключи?
- Какие запросы потребляют больше всего ресурсов?

Смотрим структуру и запросы!

- Лишние индексы или внешние ключи?
 - `pt-duplicate-key-checker`
- Какие индексы не используются?
 - `pt-index-usage`
 - Также показывает какие запросы меняют план исполнения, какие альтернативные индексы есть и многое другое
- Плохие запросы? Запускаем `pt-query-advisor`

Изучаем запросы

- Используем `pt-query-digest` для анализа `slow-query.log`
- Работает из коробки
- Может работать без `slow-log`'а!
- Может вообще по `tcp-dump` работать!

Что делает?

- Разбирает запросы по типам
- Считает для них статистику
- Помогает бороться с медленными запросами

Сводный отчет

```

# 13903.6s user time, 39.9s system time, 1.78G rss, 1.92G vsz
# Current date: Mon Jan 30 20:36:26 2012
# Hostname: hp-db01
# Files: /var/lib/mysql/db01-slow.log
# Overall: 32.98M total, 1.32k unique, 1.73k QPS, 2.06x concurrency _____
# Time range: 2012-01-30 11:08:22 to 16:25:40
# Attribute          total          min          max          avg          95%          stddev         median
# =====
# Exec time          39309s         1us         161s         1ms         515us        109ms         30us
# Lock time          4471s         0          161s        135us        89us         90ms         0
# Rows sent          118.72M        0        1010.29k     3.77         1.96        963.89         0
# Rows examine       15.54G         0         3.51M        505.92        1.96        11.61k         0
# Rows affecte       1.64M          0         2.85k         0.05         0           1.82          0
# Rows read          118.72M        0        1010.29k     3.77         1.96        963.89         0
# Bytes sent         52.40G         0         45.18M        1.67k         5.20k        20.30k        537.02
# Merge passes       128            0           1           0.00         0           0.00          0
# Tmp tables         957.76k        0           4           0.03         0           0.17          0
# Tmp disk tbl       297.92k        0           2           0.01         0           0.10          0
# Tmp tbl size       5.21T          0         2.22G        169.59k         0           1.78M         0
# Query size         8.52G          6        130.40k     277.38        833.10        709.99        174.84
# InnoDB:
# IO r bytes         672.00k        0        32.00k         0.06         0           35.90          0
# IO r ops           42             0           2           0.00         0           0.00          0
# IO r wait          19ms           0           2ms          0             0           1us           0
# pages distin       329.98M        0        32.08k        28.18         42.48        307.62         4.96
# queue wait         0              0           0           0             0           0             0
# rec lock wai       3402s         0          161s        277us         0           147ms         0
# Boolean:
# Filesort           2% yes, 97% no
# Filesort on        0% yes, 99% no
# Full join          0% yes, 99% no
# Full scan          2% yes, 97% no
# QC Hit             55% yes, 44% no
# Tmp table          2% yes, 97% no
# Tmp table on       0% yes, 99% no

```

Все запросы

```
# Profile
# Rank Query ID          Response time    Calls    R/Call  Apdx V/M  Item
# =====
# 1 0x844F3DA35A4A426D 8695.6994 22.1%    41470  0.2097 1.00  0.12 SELECT catalog
# 2 0xD566217EB8EB1A27 4098.9154 10.4%    48740  0.0841 1.00  0.11 SELECT geoip
# 3 0xB0A48B410B3CC151 2684.9671  6.8%      55 48.8176 0.00 34.18 UPDATE SELECT sales_
# 4 0x60CDD6CAE6B487BC 2541.5271  6.5%     352  7.2202 0.20  4.01 SELECT catalog catalog_category_product
# 5 0xD9DED53EF74C877F 1285.6118  3.3%   5530829  0.0002 1.00  0.00 SELECT op_image
# 6 0x1F3D58843008ACEC 1249.1639  3.2%     182  6.8635 0.12  1.65 SELECT catalog_product
# 7 0x50175ED6D1B1C8B1 1188.1541  3.0%   153964  0.0077 1.00  0.00 SELECT enterprise
# 8 0x70E2BE84868444A4 1035.3475  2.6%     1786  0.5797 0.98  0.16 SELECT sales_order
# 9 0x125DFE84FFA50675  891.3327  2.3%      382  2.3333 0.97 11... INSERT UPDATE product
# 10 0xBB968DD9BF27E8E5  797.3765  2.0%   153964  0.0052 1.00  0.00 SELECT enterprise
```

Статистика запроса

```

# Query 1: 2.25 QPS, 0.47x concurrency, ID 0x844F3DA25A4A426D at byte 4779840203
# Attribute      pct      total      min      max      avg      95%      stddev  median
# =====
# Count          0      41470
# Exec time      22      8696s      97ms     3s      210ms   455ms   159ms   155ms
# Lock time      0         7s      107us    50ms   177us   247us   272us   159us
# Rows sent      1      1.60M      26       50      40.45   49.17   10.74   46.83
# Rows examine   16      2.63G     66.42k   66.60k  66.49k  65.68k  0       65.68k
# Rows affecte   0         0         0         0         0         0         0         0
# Rows read     1      1.60M      26       50      40.45   49.17   10.74   46.83
# Bytes sent     0      200.43M    3.39k    6.14k   4.95k    6.01k    1.08k    5.45k
# Merge passes   0         0         0         0         0         0         0         0
# Tmp tables     4      40.50k     1         1         1         1         0         1
# Tmp disk tbl   0         0         0         0         0         0         0         0
# Tmp tbl size  12      647.83G   16.00M   16.00M  16.00M  16.00M  0       16.00M
# Query size    0      71.78M     1.77k    1.77k   1.77k    1.77k  0       1.77k
# InnoDB:
# IO r bytes     0         0         0         0         0         0         0         0
# IO r ops       0         0         0         0         0         0         0         0
# IO r wait      0         0         0         0         0         0         0         0
# pages distin  47      158.13M    3.90k    3.92k   3.90k    3.88k    0.00    3.88k
# queue wait     0         0         0         0         0         0         0         0
# rec lock wai   0         0         0         0         0         0         0         0
# Boolean:
# Filesort      100% yes,  0% no
# Full scan     100% yes,  0% no
# Tmp table     100% yes,  0% no
# String:
# Databases     prod
# Hosts
# InnoDB trxID  3CB60AED (1/0%), 3CB60B2E (1/0%)... 41468 more
# Last errno    0
# Users        prod
# Query_time distribution
# 1us
# 10us
# 100us
# 1ms
# 10ms
# 100ms #####
# 1s #
# 10s+

```

Анализ ошибок

- `pt-deadlock-logger` анализ дедлоков
- `pt-fk-error-logger` проверка ошибок
внешних ключей
- Можно записать результаты прямо в базу

ALTER TABLE без блокировок

`pt-online-schema-change` выполняет изменение структуры таблицы без ее блокировки.

- Делает “shadow copy”
- Очень полезен для больших таблиц
- Facebook использует ту же технику

Обновляем MySQL

- А вдруг сломается?
 - А вдруг запросы станут работать **не так**?
 - Появятся ли новые ошибки?
 - Станет ли быстрее
- Работает — не трогай!!!
 - Нужен ли вообще этот апгрейд? Вот бы знать!

Как узнать?

- Снять снимок
- Взять лог запросов
- Поднять новую версию в другом месте
- Запустить `pt-upgrade`

pt-upgrade

Name: `pt-upgrade`

Function: Выполняет запросы на нескольких серверах и смотрит разницу

Url: <http://www.percona.com/doc/percona-toolkit/2.0/pt-upgrade.html>

Репликация

- Проверить слейв (`pt-table-checksum`)
- Починить слейв (`pt-table-sync`)
- Померять отставание (`pt-heartbeat`)
- Тормознуть слейв (`pt-slave-delay`)

Что еще?

- `pt-stalk`
- `pt-sift`
- `pt-mext`

pt-stalk

- Собирает статистику по условию в том числе
 - GDB stacktrace
 - oprofile data
 - strace data
 - делает tcpdump
 - запускает shell скрипт

pt-mext

Name: pt-mext

Function: показывает в динамике

SHOW GLOBAL STATUS

Url: <http://www.percona.com/doc/percona-toolkit/pt-mext.html>

pt-mext

- Показывает в динамике (в т.ч.)
 - Created_tmp_*
 - Handler_read_*
 - Innodb_buffer_pool_read*
 - Innodb_os_log_written
 - Key_read_requests
 - Key_reads
 - Qcache*
 - Select_full_*join
 - Select_range
 - Select_scan
 - Threads_running

Убить Плохие Запросы!

- `pt-kill`
- Ждет и Убивает
 - Может убивать транзакции

Времени все равно не хватит.

- Интересное бывает в twitter @vfedorkov
- Задавайте вопросы
 - Завтра тоже можно, но уже сложнее
- В сложных, срочных и запущенных случаях
 - vlad@astellar.com

Спасибо!