



PERCONA
Performance Consulting Experts

Maatkit & Aspersa, MySQL Toolkits

Kenny Gryp

kenny.gryp@percona.com

Principal Consultant @ Percona

LOADAYS 2011

Percona

- MySQL/LAMP Consulting
- MySQL Support
- Percona Server (XtraDB)
- Percona XtraBackup
- InnoDB Recovery Toolkit, tcpstat
- (maatkit, innotop, aspersa, mysql-mmm, mysql-cacti-templates)
- ...
- <http://www.percona.com>
- <http://www.mysqlperformanceblog.com>

Maatkit

- Collection of perl scripts
- More than 26 scripts
- Making life easier for DBAs
- tools start with 'mk-'
- Open Source: GPL
- Created by Baron Schwartz

Aspersa

- Open Source: New BSD License
- Show disk statistics:
 - from strace (ioprofile)
 - from /proc/diskstats (diskstats)
- Simple, but very helpful automation:
 - get summary of OS/hardware configuration (summary)
 - reformat SHOW GLOBAL STATUS (mext)
 - if threads_running=2, start collecting data! (stalk)
- Created by Baron Schwartz

Download

```
wget...
```

```
http://www.maatkit.org/get/TOOLNAME
```

```
http://www.maatkit.org/trunk/TOOLNAME
```

```
http://aspersa.googlecode.com/svn/trunk/TOOLNAME
```

Table Of Contents

- **Maatkit**
 - Replication
 - Query Analysis
 - Common DBA Tasks
- **Aspersa**

Table Of Contents

- **Maatkit**
 - **Replication**
 - Query Analysis
 - Common DBA Tasks
- **Aspersa**

Replication

- **mk-slave-delay**: Make a slave lag
- **mk-heartbeat**: Monitor Replication Delay
- **mk-table-checksum**: Verify data consistency between servers
- **mk-table-sync**: synchronize data between servers

mk-slave-delay

- protect against human errors
- easily promote as master
- have state of data from the past
- built-in in MySQL 5.6
<http://forge.mysql.com/worklog/task.php?id=344>
- To hold slavehost one minute behind its master for ten minutes:

```
mk-slave-delay --delay 1m  
--interval 15s --run-time 10m  
slavehost
```

mk-heartbeat

- Seconds_behind_master not accurate:
 - calculated by amount of seconds elapsed since latest executed query's timestamp
 - fails when IO thread is lagging: SQLthread is waiting on IO thread: seconds_behind_master = 0
 - in multi-tiered replication, no clear view on lag compared to master
- send heartbeats on master:
**mk-heartbeat -D test --update
-h master-server**
- monitor:
**mk-heartbeat -D test --monitor
-h slave-server**

mk-table-checksum

- Checksum tables across servers

- using replication

- ```
mk-table-checksum --replicate=mydb.checksum
master-host
```

- by comparing multiple servers:

- ```
mk-table-checksum --port 4500 master  
h=slave1,P=3306 slave2 slave3 slave4
```

- Many features:

- chunking checksums to impact less

- modulo to spread checksumming over multiple days

- wait for replication to catch up

- only checksum random n% of the data

mk-table-sync

- Fix data inconsistency between 2 or more servers
- When syncing using replication changes happen on the master
- Examples:

```
mk-table-sync --execute  
h=host1,D=db,t=tbl h=host2
```

```
mk-table-sync --execute  
host1 host2 host3
```

```
mk-table-sync --execute  
--replicate test.checksum master1
```

```
mk-table-sync --execute --replicate  
test.checksum --sync-to-master  
slave1
```

Table Of Contents

- **Maatkit**

- Replication

- Query Analysis**

- Common DBA Tasks

- **Aspersa**

Query Analysis

- **mk-query-digest**: analyze various logs and generate reports
- **mk-upgrade**: check for differences in query time and resultset between servers
- **mk-visual-explain**: format EXPLAIN as a tree

mk-query-digest

- generate reports from
 - slow query log
 - mk-query-digest /path/to/slow.log**
 - binlog files
 - processlist
 - postgresql log files
 - general log (not so useful)
 - tcpdump files that captured traffic from: mysql, memcached, http
- group-by& order-by: db/ip/host/query_time:sum/max/min/count

mk-query-digest

- store reports in db: **--review, --review-history**

- enhanced filtering capabilities

```
' $event->{fingerprint} =~ m/^select/ '
```

```
' $event->{Warning_count} > 1 '
```

```
' $event->{InnoDB_IO_r_ops} > 50 '
```

```
' $event->{QC_hit} = "Yes" '
```

```
' $event->{Bytes} >= m/^1_048_576/ '
```


mk-query-digest

```
# 834.7s user time, 9.1s system time, 302.78M rss, 392.96M vsz
# Current date: Mon Nov 29 09:47:43 2010
# Hostname: servername
# Files: STDIN
# Overall: 670.66k total, 1.73k unique, 955.33 QPS, 3.08x concurrency ____
# Time range: 2010-11-29 09:14:29.955239 to 09:26:11.979320
# Attribute          total          min          max          avg          95%          stddev         median
# =====
# Exec time          2163s           0           3s           3ms           2ms           29ms           89us
# Rows affected      18.58k           0           146          0.03           0             0.49           0
# Query size         121.29M          6          21.55k       189.64         363.48         328.74         97.36
# Warning coun       438.18k          0          25.60k       0.67           0             122.19          0
# Boolean:
# No good inde       0% yes, 99% no
# No index use       10% yes, 89% no
```

mk-query-digest

Profile

#	Rank	Query ID	Response time	Calls	R/Call	Apdx	V/M	Item
#	====	=====	=====	=====	=====	=====	=====	=====
#	1	0x3928FBFF36663F33	1349.6240 62.4%	11976	0.1127	1.00	0.03	SELECT loan_officer_states leads office_assigned refinance_leads assigned
#	2	0x8A539A15CDC891EB	114.9014 5.3%	437	0.2629	1.00	0.50	SELECT processing_assigned task_history history tasks task_history history closing_assigned
#	3	0xFA5D75AB1925777C	92.9441 4.3%	791	0.1175	1.00	0.06	SELECT security_dashboard
#	4	0x6F1DB5CAB019DB16	77.5712 3.6%	43	1.8040	0.65	0.73	SELECT new_rate_lock_bi_view new_rate_locks leads office_assigned loan_officers contact_info lenders loan_officers rate_lock_exempt_branches
#	5	0xDFEC78D47187A0CD	67.1673 3.1%	296	0.2269	1.00	0.17	SELECT history assigned loan_officers history statuses
#	6	0x5D51E5F01B88B79E	49.0330 2.3%	15630	0.0031	1.00	0.00	ADMIN CONNECT
#	7	0xD704F6F4D36804AB	43.4990 2.0%	274	0.1588	1.00	0.12	SELECT user_agents
#	8	0x7EC8CF8EAF26907	30.0898 1.4%	416	0.0723	1.00	0.07	SELECT security_dashboard

mk-query-digest

Query 1: 17.06 QPS, 1.92x concurrency, ID 0x3928FBFF36663F33 at byte 1417466467

This item is included in the report because it matches --limit.

Scores: Apdex = 1.00 [1.0], V/M = 0.03

Time range: 2010-11-29 09:14:30.052415 to 09:26:11.914796

# Attribute	pct	total	min	max	avg	95%	stddev	median
-------------	-----	-------	-----	-----	-----	-----	--------	--------

# =====	===	=====	=====	=====	=====	=====	=====	=====
---------	-----	-------	-------	-------	-------	-------	-------	-------

# Count	1	11976						
---------	---	-------	--	--	--	--	--	--

# Exec time	62	1350s	25ms	395ms	113ms	219ms	54ms	91ms
-------------	----	-------	------	-------	-------	-------	------	------

# Rows affecte	0	39	0	35	0.00	0	0.32	0
----------------	---	----	---	----	------	---	------	---

# Query size	23	28.75M	2.46k	2.46k	2.46k	2.38k	0	2.38k
--------------	----	--------	-------	-------	-------	-------	---	-------

# Warning coun	11	51.51k	0	12.80k	4.40	0	233.99	0
----------------	----	--------	---	--------	------	---	--------	---

Boolean:

No index use 99% yes, 0% no

String:

Databases

Errors none (273/99%), #1064 (1/0%)

Hosts 172.20.101.178

mk-query-digest

```
# Query_time distribution
# 1us
# 10us #####
# 100us #####
# 1ms ##
# 10ms #
# 100ms #####
# 1s
# 10s+
# Tables
# SHOW TABLE STATUS LIKE 'user_agents'\G
# SHOW CREATE TABLE `user_agents`\G
# EXPLAIN /*!50100 PARTITIONS*/
SELECT user_agent_id, search_engine FROM user_agents WHERE
       user_agent='Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
```

mk-query-digest

```
# Item 1: 3.41 QPS, 0.97x concurrency, ID 0xABCE5AD2A2DD1BA1 at byte 288124661
# This item is included in the report because it matches --limit.
# Scores: Apdex = 0.97 [1.0], V/M = 19.02
# Query_time sparkline: | ^_____|
# Time range: 2011-04-05 16:12:13 to 16:14:45
# Attribute      pct      total      min      max      avg      95%      stddev     median
# =====      ==      =====      =====      =====      =====      =====      =====
# Count          0        519
# Exec time      2        148s      11us     33s     285ms    53ms      2s         26us
# Lock time      0         5ms       0        334us   9us      66us     32us       0
# Rows sent      0         41        0         1       0.08     0.99     0.27       0
# Rows examine   1        4.97M     0        445.49k 9.80k    5.73k    49.33k     0
# Rows affecte   0         2         0         1       0.00     0        0.06       0
# Rows read      1        2.01M     0        250.47k 3.96k    1.96     27.94k     0.99
# Bytes sent     0        241.20k   11       8.01k   475.89   918.49   689.98     258.32
# Merge passes   0         0         0         0        0         0         0          0
# Tmp tables     0         15        0         1       0.03     0        0.17       0
# Tmp disk tbl   0         3         0         1       0.01     0        0.08       0
```

mk-query-digest

```
# Tmp tbl size    0   4.78k      0   4.78k    9.43      0  211.60      0
# Query size     0 100.95k     19   2.71k   199.17   363.48   206.60   151.03
# 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   1  67.99k     0  10.64k   1.26k    3.88k    2.47k    31.70
# queue wait     0      0      0      0      0      0      0      0
# rec lock wai   0      0      0      0      0      0      0      0
# Boolean:
# Filesort       0% yes, 99% no
# Full scan      7% yes, 92% no
# QC Hit        78% yes, 21% no
# Tmp table      2% yes, 97% no
# Tmp table on   0% yes, 99% no
```

mk-upgrade

- runs individual queries on multiple servers
- compares
 - response time
 - resultset
- reports queries that have
 - different response time between servers
 - other resultset output
- useful for verifying upgrades (versions, hardware, engines...)
- not useful for load test!

mk-visual-explain

- Make a tree representation of explain output
- Might be easier to read with complex queries

```
mysql -e "explain select * from  
sakila.film_actor join sakila.film  
using(film_id);"  
| mk-visual-explain
```


mk-visual-explain

JOIN

+ - Bookmark lookup

| +- Table

| | table film_actor

| | possible_keys idx_fk_film_id

| +- Index lookup

| key film_actor->idx_fk_film_id

| possible_keys idx_fk_film_id

| key_len 2

| ref sakila.film.film_id

| rows 2

+ - Table scan

rows 952

+ - Table

table film

possible_keys PRIMARY

Table Of Contents

- **Maatkit**
 - Replication
 - Query Analysis
 - Common DBA Tasks**
- **Aspersa**

Common DBA Tasks

- **mk-deadlock-logger**: log deadlock information
- **mk-show-grants**: print grants
- **mk-kill**: kill queries
- **mk-find**: find tables and do ...
- **mk-duplicate-key-checker**: report duplicate indexes/
foreign keys

mk-deadlock-logger

- SHOW ENGINE INNODB STATUS show latest deadlock
- This tool fetches those deadlocks and logs them

-to STDOUT

```
mk-deadlock-logger SOURCE_DSN
```

-to another table

```
mk-deadlock-logger SOURCE_DSN
```

```
--dest D=test,t=deadlocks
```

```
--daemonize --run-time 4h
```

```
--interval 30s
```

mk-show-grants

- shows grants
- makes sure the order of grants entries and permissions is constant to be able to version control them

mk-show-grants

- generate REVOKE and DROP USER statements
mk-show-grants --drop --revoke

mk-kill

- Kill queries running longer than 60s:
mk-kill --busy-time 60 --kill
- Print, do not kill, queries running longer than 60s:
mk-kill --busy-time 60 --print
- Find sleeping processes and kill them all every 10s:
**mk-kill --match-command Sleep
--kill --victims all --interval 10**
- Print all login processes:
**mk-kill --match-state login --print
--victims all**

mk-find

- Examples:

```
mk-find --ctime +1 --engine MyISAM
```

```
mk-find --mtime +30 --engine InnoDB
```

```
--exec
```

```
"ALTER TABLE %D.%N ENGINE=MyISAM"
```

```
mk-find --empty junk test
```

```
--exec-plus "DROP TABLE %s"
```

```
mk-find --tablesize +5G
```

mk-duplicate-key-checker

```
# tezt.media_pictures
#####
# subject_node_id is a left-prefix of INDEX
# Key definitions:
#   KEY `subject_node_id` (`subject_node_id`)
#   KEY `INDEX` USING BTREE
#     (`subject_node_id`, `frame_id`, `file_id`, `source_id`),
# Column types:
# `subject_node_id` int(11) unsigned default null
# `frame_id` smallint(6) unsigned not null
# `file_id` int(11) unsigned not null
# `source_id` int(11) not null
# To remove this duplicate index, execute:
ALTER TABLE `tezt`.`media_pictures` DROP INDEX
  `subject_node_id`;
```


mk-duplicate-key-checker

```
# Db.system_transaction
#####
# FOREIGN KEY A (`table_id`) REFERENCES `Db`.`table` (`id`)
  is a duplicate of FOREIGN KEY B (`table_id`) REFERENCES
  `Db`.`table` (`id`)

# Key definitions:
#   CONSTRAINT `A` FOREIGN KEY (`table_id`) REFERENCES
#   `table` (`id`)
#   CONSTRAINT `B` FOREIGN KEY (`table_id`) REFERENCES
#   `table` (`id`)

# Column types:
# `table_id` bigint(20) default null
# To remove this duplicate foreign key, execute:
ALTER TABLE `Db`.`system_transaction` DROP FOREIGN KEY `A`;
# MySQL uses the A index for this foreign key constraint
```

Table Of Contents

- **Maatkit**

- Replication

- Query Analysis

- Common DBA Tasks

- **Aspersa**

Aspersa

- **diskstats**: show IO statistics
- **ioprofile**: use strace and lsof to show IO activity
- **mext**: output SHOW GLOBAL STATUS side-by-side

diskstats

```
percona@machine ~ $ ./diskstats -d sdb1
```

#ts	dev	rd_mb_s	rd_cnc	rd_rt	wr_mb_s	wr_cnc	wr_rt	busy	in_prg
{1}	sdb1	0.3	1.0	14.9	1.6	2.2	7.2	96%	0
{1}	sdb1	0.4	1.3	13.9	6.3	2.5	7.4	95%	0
{1}	sdb1	0.5	0.9	6.8	2.7	4.9	11.9	96%	0

c

```
Enter a column pattern: wr.*
```

#ts	device	wr_s	wr_avkb	wr_mb_s	wr_mrg	wr_cnc	wr_rt
{1}	sdb1	208.5	21.5	2.2	63%	1.1	5.4
{1}	sdb1	208.5	21.5	2.2	63%	1.1	5.4
{1}	sdb1	216.2	25.9	2.7	69%	2.0	9.2

c

```
Enter a column pattern: rd.*
```

#ts	device	rd_s	rd_avkb	rd_mb_s	rd_mrg	rd_cnc	rd_rt
{1}	sdb1	28.5	16.3	0.2	0%	1.6	55.0
{1}	sdb1	104.8	9.4	0.5	0%	1.6	14.8
{1}	sdb1	85.0	11.0	0.5	0%	1.5	17.6
{1}	sdb1	105.7	10.3	0.5	0%	2.1	19.7

ioprofile

```
$ ioprofile t/samples/ioprofile-001.txt
```

	total	pread	read	pwrite	write	filename
10.094264	10.094264	0.000000	0.000000	0.000000	0.000000	aia_227_228.ibd
8.356632	8.356632	0.000000	0.000000	0.000000	0.000000	aia_227_223.ibd
0.048850	0.046989	0.000000	0.001861	0.000000	0.000000	aia_instances.ibd
0.035016	0.031001	0.000000	0.004015	0.000000	0.000000	vo_difuus.ibd
0.013360	0.000000	0.001723	0.000000	0.011637	0.000000	mysql-relay.002113
0.008676	0.000000	0.000000	0.000000	0.008676	0.000000	master.info
0.002060	0.000000	0.000000	0.002060	0.000000	0.000000	ibdata1
0.001490	0.000000	0.000000	0.001490	0.000000	0.000000	ib_logfile1
0.000555	0.000000	0.000000	0.000000	0.000555	0.000000	mysql-relay.info
0.000141	0.000000	0.000000	0.000141	0.000000	0.000000	ib_logfile0
0.000100	0.000000	0.000000	0.000100	0.000000	0.000000	9fvus.ibd

ioprofile

```
$ ioprofile -c count t/samples/ioprofile-001.txt
```

tot	prd	rd	pwr	wt	filename
4282	4282	0	0	0	aia_227_223.ibd
2713	2713	0	0	0	aia_227_228.ibd
390	0	47	0	343	mysql-relay.002113
343	0	0	0	343	master.info
30	8	0	22	0	vo_difuus.ibd
19	7	0	12	0	aia_instances.ibd
16	0	0	16	0	ib_logfile1
16	0	0	0	16	mysql-relay.info
6	0	0	6	0	ibdata1
1	0	0	1	0	ib_logfile0
1	0	0	1	0	9fvus.ibd

mext

```
percona@machine ~ $ ./mext -r -- mysqladmin ext -i 10 -c 3
```

Aborted_clients	428	0	0
Aborted_connects	80886	0	0
Binlog_cache_disk_use	0	0	0
Binlog_cache_use	0	0	0
Bytes_received	2875788973602	1738235	346057
Bytes_sent	863929033790	588078	536398
Com_begin	6298644573	3516	5102
Com_call_procedure	0	0	0
Com_delete	23721852	26	51
Com_insert	4454794705	1518	3287
Com_replace	527848577	197	121
Com_select	6993291133	8114	7594
Com_set_option	5112076	250	262
Connections	7331059	250	262
Created_tmp_disk_tables	113568	0	0
Created_tmp_files	7803	0	0
Created_tmp_tables	729281259	1816	479

mext

Handler_commit	4002481284	5295	4911
Handler_delete	7256841	10	25
Handler_discover	0	0	0
Handler_prepare	0	0	0
Handler_read_first	47274	0	0
Handler_read_key	42993091324	34920	27522
Handler_read_next	19633194815	16911	10142
Handler_read_prev	2440127	0	0
Handler_read_rnd	488760449	40	12
Handler_read_rnd_next	2731205271	268	231
Handler_rollback	5781	0	0
Handler_savepoint	0	0	0
Handler_savepoint_rollback	0	0	0
Handler_update	7022320034	10047	3329
Handler_write	7334430104	1945	3638

mext

Qcache_free_blocks	2899	100	-15
Qcache_free_memory	519642808	164104	-8080
Qcache_hits	325634530	0	0
Qcache_inserts	978847229	194	104
Qcache_lowmem_prunes	19158357	0	0
Qcache_not_cached	211301010	806	798
Qcache_queries_in_cache	3677	-112	9
Qcache_total_blocks	10277	-131	6
Threads_cached	9	1	0
Threads_connected	11	-1	0
Threads_created	294	0	0
Threads_running	5	-3	0
Uptime	21912350	10	10

Aspersa

- **mysql-summary**: summarize status, configuration of database
- **summary**: summarize status, configuration, hardware of a server

mysql-summary

```
# Aspersa MySQL Summary Report #####
  System time | 2011-04-12 21:17:52 UTC (local TZ: CDT -0500)
# Instances #####
Port  Data Directory          Socket
=====
 3306 /disk1/data/              /var/run/mysqld/mysqld.sock
# Report On Port 3306 #####
      User | root@localhost
      Time | 2011-04-12 16:17:53 (CST)
  Hostname | xxxxx
   Version | 5.1.43-60.5.0.3.20-log (Percona
  Built On | debian-linux-gnu x86_64
   Started | 2010-11-30 20:02 (up 132+20:15:18)
 Databases | 5
   Datadir | /disk1/data/
 Processes | 2 connected, 2 running
Replication | Is a slave, has 1 slaves connected
   Pidfile | /var/run/mysqld/mysqld.pid (exists)
```

mysql-summary

```
# Processlist #####
```

Command	COUNT(*)	Working	SUM(Time)	MAX(Time)
Binlog Dump	1	1	3000	3000
Connect	2	2	12500000	12500000
Query	1	1	0	0

User	COUNT(*)	Working	SUM(Time)	MAX(Time)
replication	1	1	3000	3000
root	1	1	0	0
system user	2	2	12500000	12500000

mysql-summary

```
# Status Counters (Wait 10 Seconds) #####
...
Handler_commit                125000000          1500
Handler_delete                 125000             1
Handler_prepare                125000000          1500
Handler_read_first              25
Handler_read_key                70000000           800
Handler_read_next              20000000           225
Handler_read_prev               35000
Handler_read_rnd               20000000           225
Handler_read_rnd_next          3500000            40
    25
Handler_rollback                350
Handler_update                 20000000           225
Handler_write                   22500000           250
    25
...
Uptime                          90000              1
    45
```

mysql-summary

```
# Table cache #####  
      Size | 2048  
      Usage | 1%  
# Key Percona Server features #####  
Table & Index Stats | Not Supported  
Multiple I/O Threads | Enabled  
Corruption Resilient | Not Supported  
Durable Replication | Disabled  
Import InnoDB Tables | Disabled  
Fast Server Restarts | Not Supported  
  Enhanced Logging | Not Supported  
Replica Perf Logging | Not Supported  
  Response Time Hist. | Not Supported  
  Smooth Flushing | Enabled  
HandlerSocket NoSQL | Not Supported  
Fast Maatkit Hashes | Unknown
```

mysql-summary

```
# Query cache #####
  query_cache_type | OFF
        Size      | 0.0k
        Usage      | 0%
  HitToInsertRatio | 0%
# Schema #####
Would you like to mysqldump -d the schema and analyze it? y/n n
Skipping schema analysis
# Noteworthy Technologies #####
  Full Text Indexing | No
  Geospatial Types   | No
    Foreign Keys     | No
    Partitioning     | No
        SSL          | No
Explicit LOCK TABLES | Yes
  Delayed Insert     | No
  XA Transactions    | No
    NDB Cluster      | No
Prepared Statements  | No
```

mysql-summary

```
# InnoDB #####  
Version | 1.0.6-9.1  
Buffer Pool Size | 160.0G  
Buffer Pool Fill | 100%  
Buffer Pool Dirty | 0%  
File Per Table | ON  
Page Size | 16k  
Log File Size | 3 * 750M = 2.2G  
Log Buffer Size | 8M  
Flush Method | O_DIRECT  
Flush Log At Commit | 2  
XA Support | ON  
Checksums | ON  
Doublewrite | ON  
R/W I/O Threads | 8 8  
I/O Capacity | 2000  
Thread Concurrency | 0
```


mysql-summary

```
Concurrency Tickets | 500
Commit Concurrency | 0
Txn Isolation Level | READ-COMMITTED
Adaptive Flushing | OFF
Adaptive Checkpoint | estimate
Checkpoint Age | 0k
InnoDB Queue | 0 queries inside InnoDB, 0 queries in
queue
Oldest Transaction | 0 Seconds
History List Len | 6
Read Views | 1
Undo Log Entries | 2 transactions, 4 total undo, 2 max undo
Pending I/O Reads | 0 buf pool reads, 0 normal AIO, 0 ibuf
AIO, 0 preads
Pending I/O Writes | 0 buf pool (0 LRU, 0 flush list, 0 page);
0 AIO, 0 sync, 0 log IO (0 log, 0 chkp); 0 pwrites
Pending I/O Flushes | 0 buf pool, 0 log
Transaction States | 2xnot started
```

mysql-summary

```
# MyISAM #####
      Key Cache | 16.0M
      Pct Used  | 20%
      Unflushed | 0%
# Security #####
      Users     | 108 users, 2 anon, 5 w/o pw, 92 old pw
      Old Passwords | OFF
# Binary Logging #####
      Binlogs   | 7
      Zero-Sized | 0
      Total Size | 1.0k
      binlog_format | ROW
      expire_logs_days | 5
      sync_binlog  | 0
      server_id   | 1000
      binlog_do_db | production
      binlog_ignore_db | mysql,test
```

mysql-summary

```
# Noteworthy Variables #####
Auto-Inc  Incr/Offset | 1/1
default_storage_engine | 0
      flush_time | 0
      init_connect | 0
      init_file | 0
      sql_mode | 0
      join_buffer_size | 128k
      sort_buffer_size | 8M
      read_buffer_size | 128k
read_rnd_buffer_size | 8M
      bulk_insert_buffer | 0k
max_heap_table_size | 32M
      tmp_table_size | 32M
max_allowed_packet | 32M
      thread_stack | 192k
              log | OFF
      log_error | /var/log/mysql/mysql.log
log_warnings | 1
```

summary

```
# Aspersa System Summary Report #####
    Date | 2010-12-28 10:14:56 UTC (local TZ: CET +0100)
    Hostname | xxx
Uptime|134 days,10:41, 2 users, load average: 0.86, 0.65, 0.62
System|Dell Inc.; PowerEdge R410; vNot Specified (<OUT OF SPEC>)
    Service Tag | xxx
        Release | CentOS release 5.5 (Final)
        Kernel | 2.6.18-194.11.1.el5
Architecture | CPU = 64-bit, OS = 64-bit
    Threading | NPTL 2.5
        Compiler | GNU CC version 4.1.2 20080704 (Red Hat 4.1.2-48).
        SELinux | Disabled
# Processor #####
Processors | physical = 2, cores = 8, virtual = 16,
    hyperthreading = yes
        Speeds | 16x2261.059
        Models | 16xIntel(R) Xeon(R) CPU L5520 @ 2.27GHz
        Caches | 16x8192 KB
```

summary

```
# Memory #####  
      Total | 7.79G  
      Free  | 379.47M  
      Used  | physical = 7.42G, swap = 236.00k, virtual = 7.42G  
  Buffers  | 232.73M  
  Caches   | 4.74G  
    Used   | 1.80G  
  Swappiness | vm.swappiness = 60  
  DirtyPolicy | vm.dirty_ratio = 40, vm.dirty_background_ratio =  
    10
```

Locator	Size	Speed	Form Factor	Type	Type Detail
DIMM_A1	2048 MB	1066 MHz	DIMM	DDR3	Synchronous
DIMM_A2	2048 MB	1066 MHz	DIMM	DDR3	Synchronous
DIMM_B1	2048 MB	1066 MHz	DIMM	DDR3	Synchronous
DIMM_B2	2048 MB	1066 MHz	DIMM	DDR3	Synchronous
DIMM_A3	{EMPTY}	Unknown	DIMM	DDR3	Synchronous
DIMM_A4	{EMPTY}	Unknown	DIMM	DDR3	Synchronous
DIMM_B3	{EMPTY}	Unknown	DIMM	DDR3	Synchronous
DIMM_B4	{EMPTY}	Unknown	DIMM	DDR3	Synchronous

summary

```
# Mounted Filesystems #####
Filesystem      Size Used Type  Opts          Mountpoint
/dev/mapper/Vo0-LVol100 261G 34% ext3 (rw,noatime) /
/dev/sda1       99M 14% ext3 (rw)         /boot
/mnt/tmpfs      3.9G 0% tmpfs (rw)        /mnt/tmpfs
tmpfs           3.9G 0% tmpfs (rw)        /dev/shm

# Disk Schedulers And Queue Size #####
    sda | [deadline] 128
    sr0 | [cfq] 128

# Disk Partitioning #####
Device          Type      Start      End          Size
=====
/dev/sda        Disk          299439751168
/dev/sda1       Part         1          13           98703360
/dev/sda2       Part        14         36404        299317939200

# Kernel Inode State #####
dentry-state | 357302 352501 45 0 0 0
    file-nr | 4590 0 765829
54 inode-nr | 530914 205111
```

summary

```
# LVM Volumes #####
LV          VG          Attr      LSize    Origin Snap%   Move Log Copy
% Convert
LogVol100  VolGroup00 -wi-ao    269.00G
LogVol101  VolGroup00 -wi-ao     9.75G
# RAID Controller #####
Controller | LSI Logic MegaRAID SAS
Model      | PERC 6/i Adapter, PCIE interface, 8 ports
Cache     | 256MB Memory, BBU Present
          | BBU | 89% Charged, Temperature 34C, isSOHGood=Yes
VirtualDev Size  RAID Level Disks SpnDpth Stripe Status  Cache
ODisk      278GB 1 (1-0-0)    2      1-1     64 Optimal WB,no-ra

PhysiclDev Type State   Errors Vendor  Model              Size
Hard Disk  SAS  Online  0/0/0  HITACHI HUC103030CSS600  279.396
Hard Disk  SAS  Online  0/0/0  HITACHI HUC103030CSS600  279.396
```

summary

```
# Network Config #####
Controller | Broadcom Corporation NetXtreme II BCM5716 Gigabit
Ethernet (rev 20)
Controller | Broadcom Corporation NetXtreme II BCM5716 Gigabit
Ethernet (rev 20)
FIN Timeout | net.ipv4.tcp_fin_timeout = 60
Port Range | net.ipv4.ip_local_port_range = 32768 61000
# Interface Statistics #####
int rx_bytes rx_packets rx_err tx_bytes tx_packets tx_err
=== =====
lo 10046316 95424 0 10046316 95424 0
eth 03931185084 140277525 0 550200667 439889576 0
eth1 768740181 3815328653 0 2902824733 1782131420 0
sit0 0 0 0 0 0 0
```


summary

```
# Network Connections #####  
Connections from remote IP addresses  
  172.18.10.5          3  
  172.18.10.6          3  
Connections to local IP addresses  
  172.18.10.7          6  
Connections to top 10 local ports  
  2049                 2  
  3306                 4  
  43183                1  
  43186                1  
States of connections  
  ESTABLISHED         8  
  LISTEN              10  
  TIME_WAIT           7
```

summary

```
# Top Processes #####  
PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMA  
1903 mysql 15 0 4015m 1.7g 3828 S 9.8 21.7 203:23.38 mysqld  
24107 root 15 0 0 0 0 S 2.0 0.0 190:49.40 nfsd  
24109 root 15 0 0 0 0 S 2.0 0.0 189:40.39 nfsd  
24112 root 15 0 0 0 0 S 2.0 0.0 189:19.66 nfsd  
1 root 15 0 10348 632 540 S 0.0 0.0 0:18.59 init
```

```
# Simplified vmstat (wait please) #####  
procs ---swap-- ----io---- ---system---- -----cpu-----  
r b si so bi bo in cs us sy id wa st  
0 0 0 0 5 9 0 0 3 2 95 0 0  
0 0 0 0 0 340 2125 1633 0 0 100 0 0  
0 0 0 0 0 212 2808 2088 1 0 98 0 0  
1 0 0 0 0 0 1813 1618 3 1 96 0 0  
2 0 0 0 0 0 4339 4413 4 2 94 0 0
```

Aspersa

- **collect**: collect data
- **stalk**: wait for some condition to become true and execute a script
- **sift**: run through 'collect' data

collect

```
collected/2010_09_29_20_46_14-df
collected/2010_09_29_20_46_14-diskstats
collected/2010_09_29_20_46_14-innodbstatus1
collected/2010_09_29_20_46_14-innodbstatus2
collected/2010_09_29_20_46_14-iostat
collected/2010_09_29_20_46_14-iostat-overall
collected/2010_09_29_20_46_14-lsof
collected/2010_09_29_20_46_14-mpstat
collected/2010_09_29_20_46_14-mysqldadmin
collected/2010_09_29_20_46_14-netstat
collected/2010_09_29_20_46_14-processlist1
collected/2010_09_29_20_46_14-processlist2
collected/2010_09_29_20_46_14-procstat
collected/2010_09_29_20_46_14-ps
collected/2010_09_29_20_46_14-stacktrace
collected/2010_09_29_20_46_14-top
collected/2010_09_29_20_46_14-variables
collected/2010_09_29_20_46_14-vmstat
collected/2010_09_29_20_46_14-vmstat-overall
```

stalk

```
# This is the max number of <whatever> we want to tolerate.
THRESHOLD=${THRESHOLD:-100}
# This is the thing to check for.
VARIABLE=${VARIABLE:-Threads_connected}
# How many times must the condition be met before the script
  will fire?
CYCLES=${CYCLES:-1}
# Collect GDB stacktraces?
GDB=${GDB:-no}
# Collect oprofile data?
OPROFILE=${OPROFILE:-yes}
# Collect strace data?
STRACE=${STRACE:-no}
# Collect tcpdump data?
TCPDUMP=${TCPDUMP:-yes}
```

sift

```
bash-3.2$ ./sift collected/
```

```
2010_09_29_20_25_09 2010_09_29_20_35_41 2010_09_29_20_46_14  
2010_09_29_20_56_48 2010_09_29_21_07_21 2010_09_29_21_17_53  
2010_09_29_21_28_27 2010_09_29_21_39_00 2010_09_29_21_49_38
```

```
Select a timestamp from the list [2010_09_29_21_49_38]
```

sift

==== unknown at 2010_09_29_21_49_38 DEFAULT (9 of 9) =====

--diskstats--

dev	rd_mb_s	rd_cnc	rd_rt	wr_mb_s	wr_cnc	wr_rt	busy	in_prg
sdb1	3.3	2.5	7.3	0.3	0.5	48.5	100%	0
sdb1	0%	5%

--vmstat--

r	b	swpd	free	buff	cache	si	so	bi	bo	in	cs	us	sy	id	wa	st
6	4	324	172636	1208008	10861412	0	0	79	305	1	2	0	0	96	3	0
1	3	324	159260	1210192	10866144	0	0	120	850	3213	4550	1	0	90	9	0
wa	0%	5%	.	.	.	0%	5%	.	.	.	10%	.	.	.	5%	.

--innodb--

txns: 2xACTIVE (406s) 6xnot (0s)

0 queries inside InnoDB, 0 queries in queue

Main thread: flushing buffer pool pages, pending reads 2, writes 34, flush 0

Log: lsn = 491, chkp = 491, chkp age = 0

Threads are waiting at:

Threads are waiting on:

sift

```
--processlist--
```

```
State
```

```
2 NULL
```

```
1 Waiting for master to send event
```

```
1 Has sent all binlog to slave; waiting for binlog to be updated
```

```
1 Has read all relay log; waiting for the slave I/O thread to update it
```

```
Command
```

```
5 Sleep
```

```
3 Query
```

```
2 Connect
```

```
1 Binlog Dump
```

```
--stack traces--
```

```
17 pthread_cond_wait,end_thread,handle_one_connection,start_thread,clone
```

```
10 pthread_cond_wait,os_event_wait_low,os_aio_simulated_handle,  
fil_aio_wait,io_handler_thread
```

```
5 read,read,vio_read,my_real_read,my_net_read
```

```
1 select,os_thread_sleep,srv_lock_timeout_and_monitor_thread,start_threa
```

```
1 select,os_thread_sleep,srv_error_monitor_thread,start_thread,clone
```

```
--oprofile--
```

```
No oprofile file exists
```


sift

- **'j'**: next, **'k'**: previous
- **'d'**: diskstats
- **'i'**: innodbstatus
- **'m'**: mext
- **'n'**: netstat

Other Toolkits

- <http://code.google.com/p/innotop/>
- <http://www.percona.com/software/mysql-innodb-data-recovery-tools/>
- <http://www.percona.com/docs/wiki/tcprstat:start>
- <http://code.openark.org/forge/openark-kit>
- <https://launchpad.net/mysql-utilities>
- <http://code.openark.org/forge/mycheckpoint>
- <http://code.google.com/p/mysql-cacti-templates/>
- <http://code.google.com/p/securich/>
- <http://mysql-mmm.org/>

Maatkit & Aspersa, MySQL Toolkits

- Maatkit: <http://www.maatkit.org>
- Aspersa: <http://code.google.com/p/aspersa/>
- IRC: Freenode #percona, #maatkit
- Mailinglist: <http://groups.google.com/group/percona-discussion>
- <http://www.percona.com>
- <http://www.mysqlperformanceblog.com>

Percona Live MySQL Conference NYC

May 26, 2011

<http://www.perconalive.com>



PERCONA
LIVE