

MySQL server replication, optimization and monitoring

Igor Ajdišek

Lead System Operations Engineer
Central European Media Enterprises



OSVAJALEC

osvajalec.si

Brief introduction...

- Leading media company in CEE region w/ operations in 6 countries
- US funded; NASDAQ: CETV
- Owns 20+ TV stations
- Operates POPTV, Kanal A and TV Pika in Slovenia
- Combined audience reach = 50 Million people



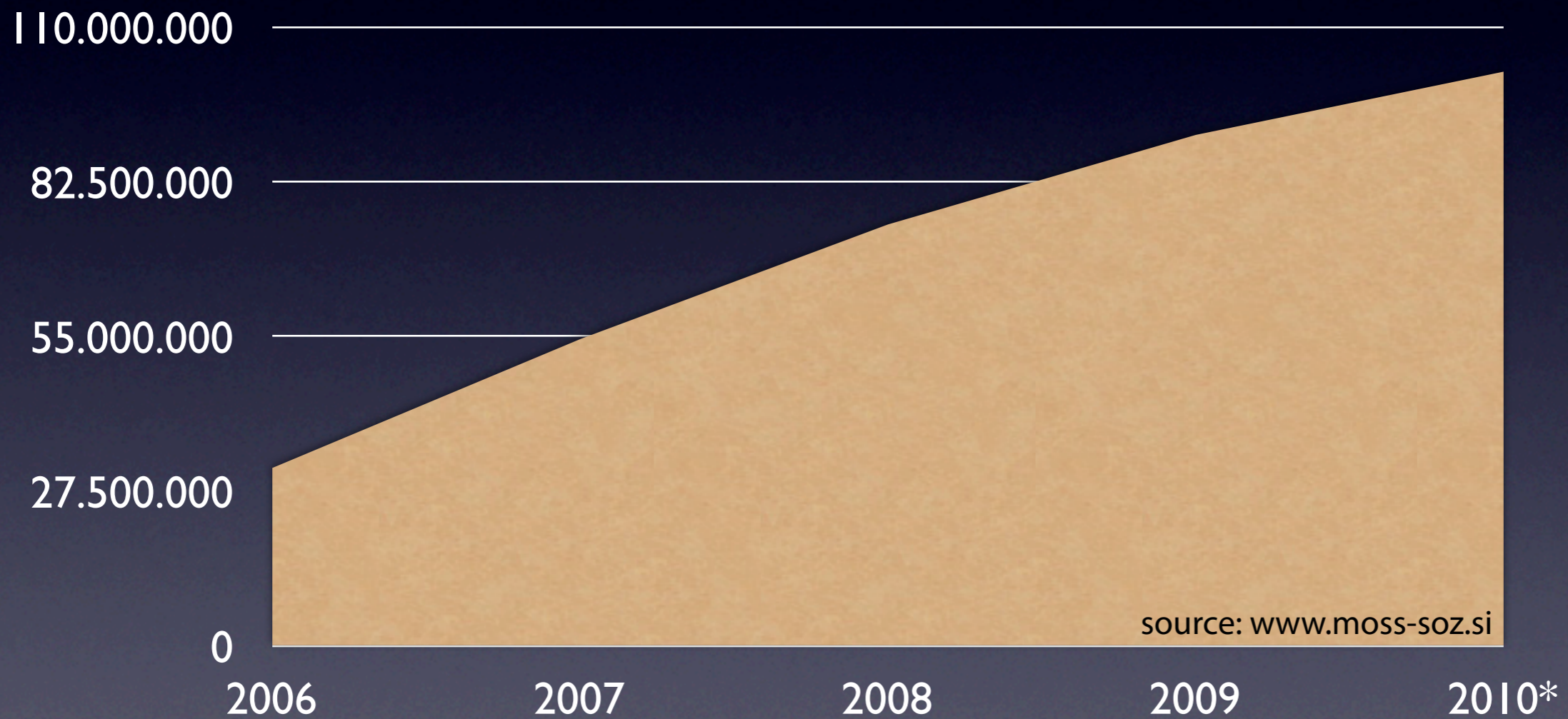
Well, what about
Internet?

24ur.com, poptv.si, **zadovoljna.si**, bibaleze.si, cekin.si, vizita.si,
maxtv.si, moskismet.com, dominvrt.si, **dnevnik.hr**, **blog.hr**,
novatv.hr, zadovoljna.hr, gle.to, **markiza.sk**, **tvnoviny.sk**,
nova.cz, **tn.cz**, **blog.cz**, **galerie.cz**, **doma.cz**, **dobyvatel.cz**,
vybereme.cz, mtv.cz, **protv.ro**, **sport.ro**, **acessatv.ro**,
stirileprotv.ro, **mtv.ro**, **profm.ro**, **kombat.ro**, **conquizardor.ro**,
btv.bg, **probg.bg**, **ring.bg**, **profm.bg**, ...

A few facts...

- Over 40 online products and services
- 2.5 Million users / month
- 450 Million page views / month
- Substantial YoY growth
- Dedicated infrastructure in 4 data centers (Ljubljana/Prague/Sofia/Bucharest)
- Built upon LAMP stack

Average PV/month (24ur.com)



* data only from Q1/2010

Adding 30 Million new
PV/month every year

How to keep up?

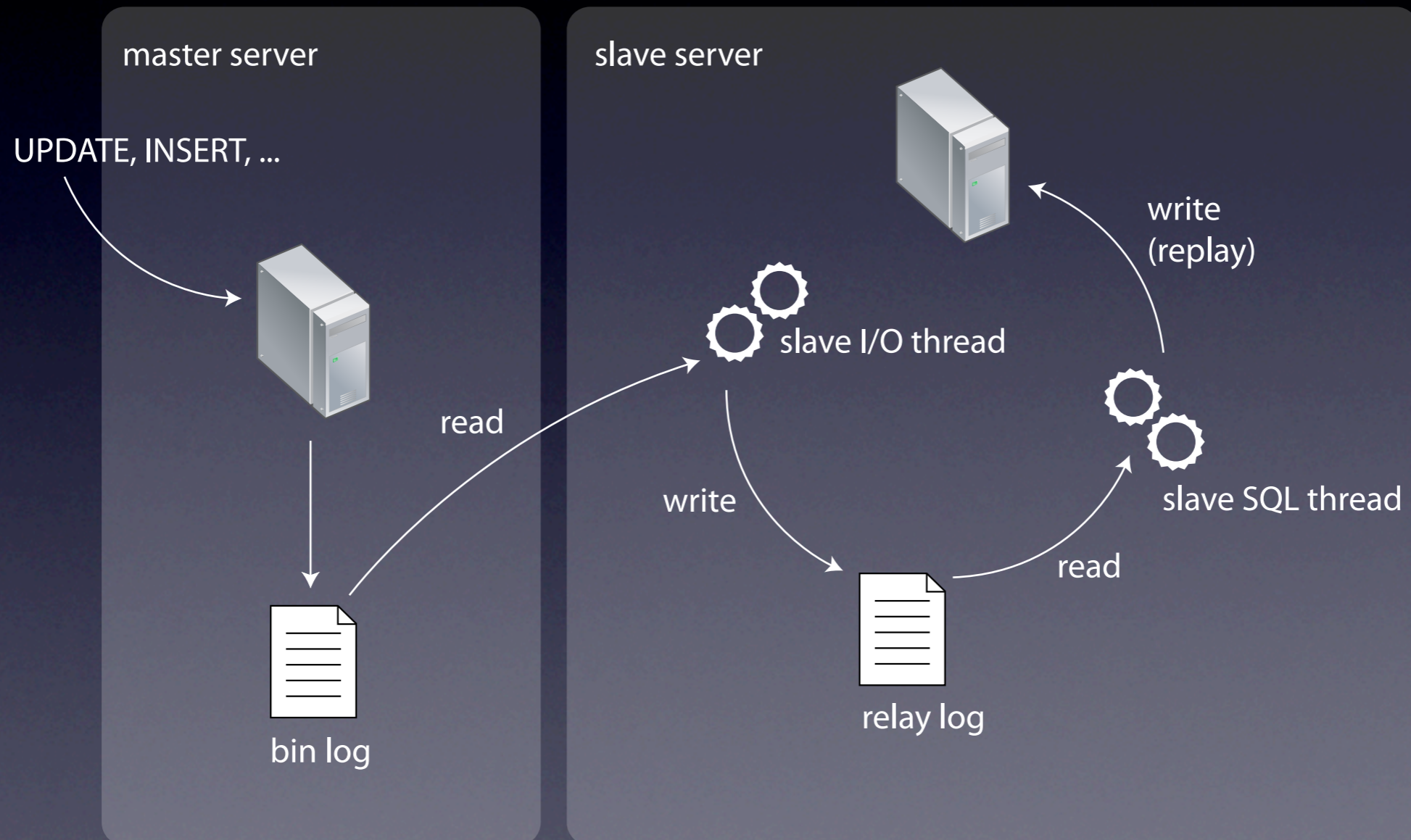
Scale.

Part 1: MySQL server replication

About MySQL replication...

- Foundation for building large, scalable web application on top of MySQL
- It's simple. Has almost no overhead
- Great for web applications that have lots of reads and little writes

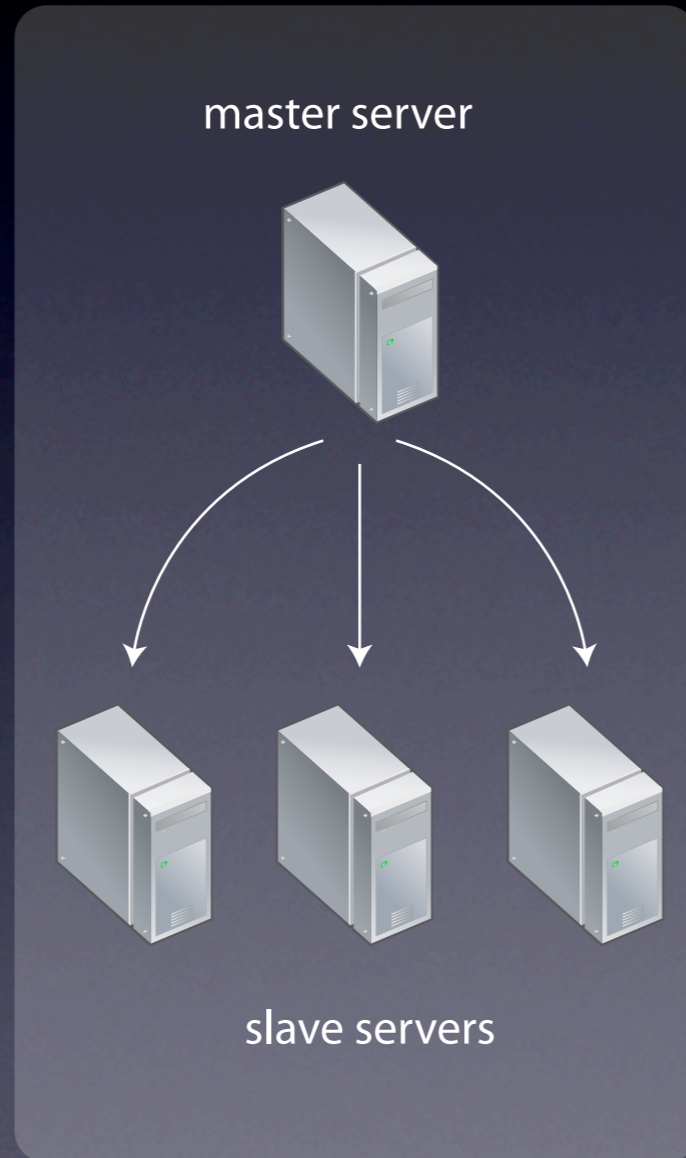
How replication works?



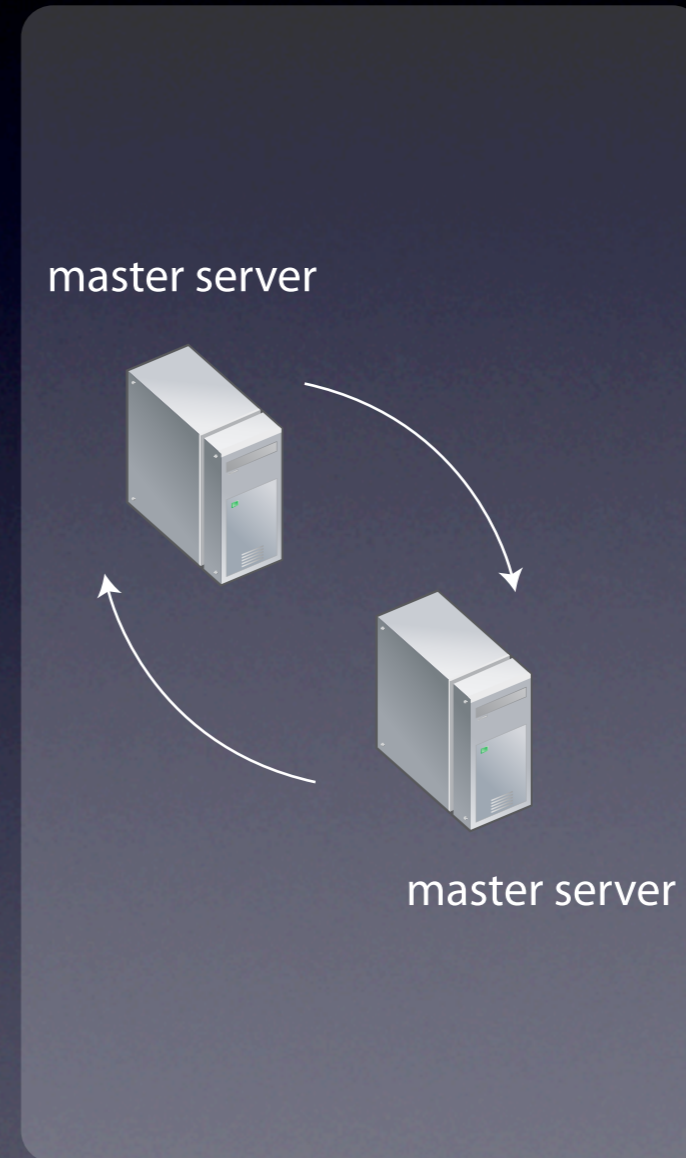
Demo

(Setting up basic replication)

Replication Topologies



Master and Multiple Slaves



Master-Master

Demo

(Going for a Master-Master setup)

Distributing read queries across Slave servers

- Round-robin through DNS
- Linux Virtual Server (LVS)
<http://www.linuxvirtualserver.org>
- MySQL Proxy
<http://dev.mysql.com/doc/refman/5.1/en/mysql-proxy.html>
- Proprietary solutions
- Own custom solution

Using PHP as you load-balancer

- Load-balancing is spread across all our webserver, making every webserver a load-balancer -> **no single point of failure!**
- `mysql.connect_timeout`
Connect timeout in **seconds!!!**
default = 60 sec
minimum = 1 sec

Is slave alive?

- Overcome `mysql.connect_timeout` issue with PHP Socket functions

<http://www.php.net/manual/en/intro.sockets.php>

```
function is_slave_alive( $host ) {  
    sock = @socket_create(AF_INET, SOCK_STREAM, SOL_TCP);  
    socket_set_nonblock($sock);  
    socket_connect($sock, $host, 3306);  
    $error = socket_last_error();  
    if ($error != SOCKET_EINPROGRESS && $error != SOCKET_EALREADY)  
    {  
        socket_close( $sock );  
        return FALSE;  
    }  
    socket_close( $sock );  
}
```

Part 2: Server-side optimization

Inspect Server Status Variables

- The best way to tune MySQL server:
`SHOW GLOBAL STATUS`
- There are tools that help you with that:
 - `mysqltuner.pl`
<http://blog.mysqltuner.com>
 - `tuning-primer.sh`
<http://www.day32.com/MySQL/>
- Predefined config files come with MySQL
`my-huge.cnf`, `my-large.cnf`, `my-small.cnf`

Demo

“There is no way to optimize MySQL
enough to compensate for poor
application design.”

Part 3: Monitoring

mtop

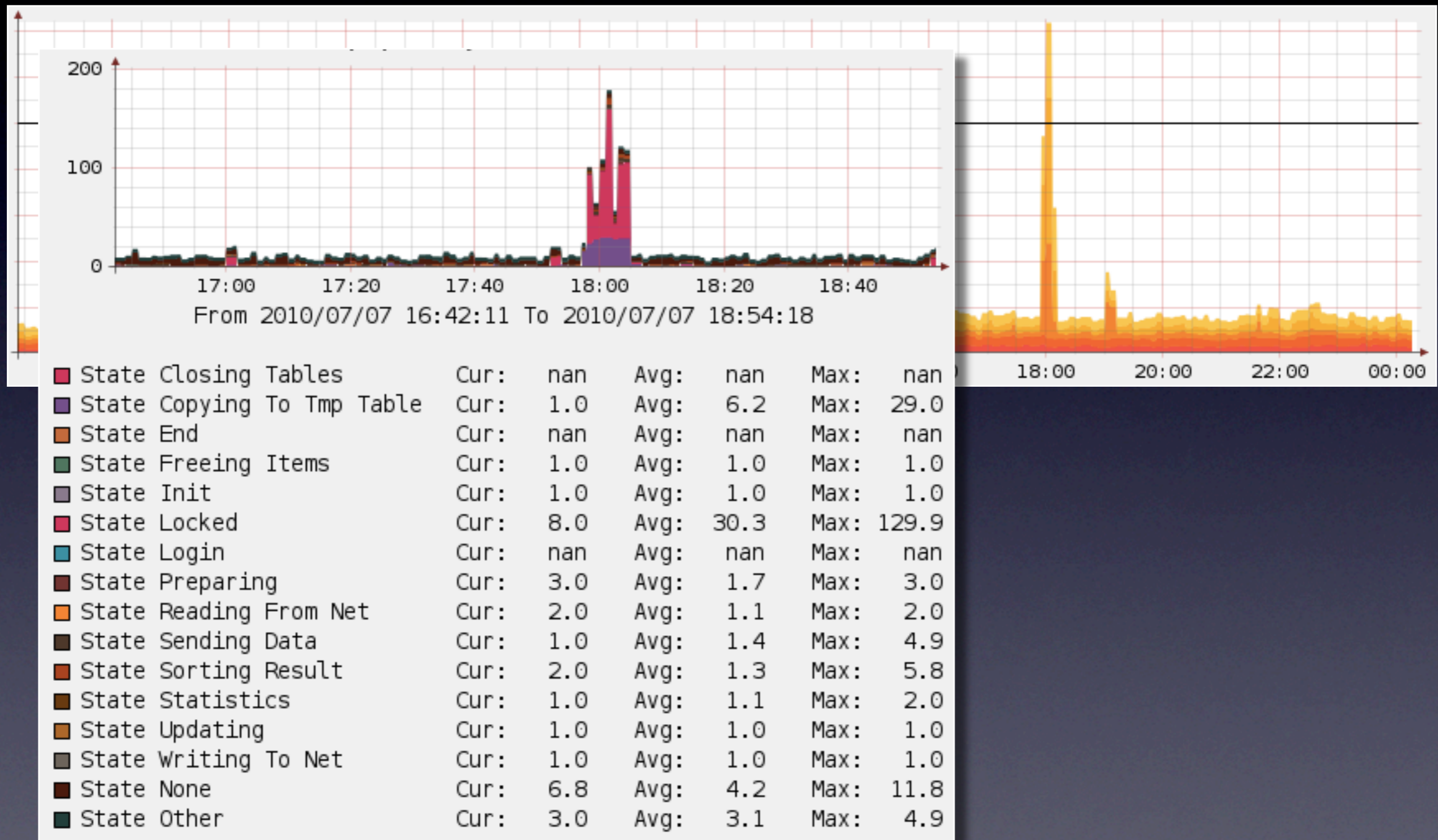
<http://mtop.sourceforge.net>

```
Terminal — ssh — 200x50
load average: 23.96, 20.70, 11.99 mysql 5.0.67-community-log up 67 day(s), 23:59 hrs
157 threads: 147 running, 9 cached. Queries/slow: 21.2G/0 Cache Hit: 99.99%
Opened tables: 0 RRN: 2.7K TLW: 9.2M SFJ: 0 SMP: 0 QPS: 4K

ID      USER      HOST          DB          TIME  COMMAND STATE      INFO
2       system u
98789591 system u
111168011 10.0.0.73:44108 6 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111173481 10.0.0.192:51712 6 Query Sending data select ... from media INNER JOIN article_media ON media.media_id=article_media.media_id LEFT JOIN source ON source.sourc
111173617 10.0.0.192:52962 6 Query Sorting resu select ... from media INNER JOIN article_media ON media.media_id=article_media.media_id LEFT JOIN source ON source.sourc
111173831 10.0.0.71:54113 6 Query Sorting resu select ... from media INNER JOIN article_media ON media.media_id=article_media.media_id LEFT JOIN source ON source.sourc
111175271 10.0.0.86:40355 6 Query Sorting resu select ... from media INNER JOIN article_media ON media.media_id=article_media.media_id LEFT JOIN source ON source.sourc
111175845 10.0.0.191:44411 5 Query Sending data select ... from media INNER JOIN article_media ON media.media_id=article_media.media_id LEFT JOIN source ON source.sourc
111171988 10.0.0.74:52966 4 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111175226 10.0.0.86:40203 4 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168278 10.0.0.192:35048 1 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111172712 10.0.0.191:45659 1 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111167911 10.0.0.73:43104 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111167917 10.0.0.192:32806 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168002 10.0.0.74:51806 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168004 10.0.0.86:57853 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168019 10.0.0.73:44183 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168023 10.0.0.74:52015 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168075 10.0.0.74:52334 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168137 10.0.0.74:52698 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168147 10.0.0.192:34202 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168153 10.0.0.74:52767 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168159 10.0.0.74:52806 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168212 10.0.0.86:34194 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168285 10.0.0.73:47069 Query statistics SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168500 10.0.0.86:36180 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111168516 10.0.0.192:36665 Query Copying to t SELECT ... from @ UNIX_TIMESTAMP() ) INNER JOIN md_value MT ON (M.media_id=MT.item_id and MT.item_type='MEDIA' and MT.m
111172678 10.0.0.191:45383 Query Locked select ... from media where media_id='60338142'
111172684 10.0.0.86:38932 Query Opening tabl SELECT ... from <1278518742
111172716 10.0.0.86:39179 Query Sorting resu SELECT ... FROM article_media WHERE type=1 AND article_id = 131787 ORDER BY (show_at & 1) desc, seqno, media_id
111172762 10.0.0.191:46109 Query statistics SELECT ... from <1278518742 AND (ar.article_id NOT IN (107039,131830)) AND fr.show_as = '15' ORDER BY fr.seqno ASC
111172918 10.0.0.191:47632 Query Locked select ... from article_media t1 inner join media t2 on (t1.media_id=t2.media_id) where (t1.article_id=5639) and (t1.sho
111173042 10.0.0.74:33271 Query Locked select ... from media where media_id='60181063'
111173138 10.0.0.191:49604 Query Locked select ... from media where media_id='60327964'
111173148 10.0.0.191:49655 Query Locked select ... from media where media_id='60324666'
111173197 10.0.0.71:48734 Query
111173199 10.0.0.73:38654 Query Locked SELECT ... FROM article_media INNER JOIN media ON media.media_id = article_media.media_id INNER JOIN media AS picture ON
111173271 10.0.0.71:49256 Query Locked SELECT ... FROM article_media am LEFT JOIN media m ON am.media_id = m.media_id WHERE m.status >= 3 AND m.status <= 9 AND
111173280 10.0.0.191:50531 Query Locked SELECT ... FROM source s LEFT JOIN media m ON s.source_id = m.source_id WHERE m.media_id = 60339690
111173320 10.0.0.71:49790 Query Locked SELECT ... FROM article_media WHERE article_id=86237 AND type IN (3,5)) as am INNER JOIN media vid ON am.media_id=vid.me
111173376 10.0.0.191:51510 Query Locked select ... from media where media_id='60250003'
111173450 10.0.0.191:52046 Query Locked select ... from media where media_id='60334201'
111173500 10.0.0.73:40738 Query Locked select ... from media where media_id='60315695'
111173522 10.0.0.71:51542 Query Locked SELECT ... FROM article_media am LEFT JOIN media m ON am.media_id = m.media_id WHERE m.status >= 3 AND m.status <= 9 AND
111173533 10.0.0.192:52312 Query Locked select ... from article_media t1 inner join media t2 on (t1.media_id=t2.media_id) where (t1.article_id=131883) and (t1.s
```

Cacti

<http://www.cacti.net>



dbmon.pl

<http://faemalia.net/mysqlUtils>

```
usage: ./dbmon.pl {--daemon|--terminal} [--help] [--rotate] [--killlong=N]
[--killvolatile] [--killconn] [--host=host[:port]] [--license]
```

Required argument:

--daemon run as a daemon, logging into
--terminal run in terminal, STDOUT is statistics, STDERR is queries

Optional arguments:

--host=h[:p] connect to a specific host and (optional) :port
--killlong=N kill >N-second-running things ('thing'==query by default, but can be connection)
--killvolatile also kill volatile (insert/update/delete) long-running queries
--killconn kill long-connected connections, not long-running queries

Other useful tools...

- innotop (like mtop)
<http://code.google.com/p/innotop>
- maatkit (advanced tools for MySQL databases)
<http://code.google.com/p/maatkit>
- and many others... :)
<http://www.mysqlperformanceblog.com/tools>

Further reading

- High Performance MySQL book (O'Reilly)
- MySQL Performance Blog
<http://mysqlperformanceblog.com>
- The Twitter Engineering Blog
<http://engineering.twitter.com>
- Digg Technology Blog
<http://about.digg.com/blog/technology>

Thank you.

twitter.com/igorajdisek
twitter.com/labs24ur