Large scale deployment PMM

Santa Clara, California | April 23th – 25th, 2018
Johan Nilsson, Kristofer Grahn – Verisure Innovation
Why are we here? - PMM sucks!! :-) 
(and it's really cool to talk at Percona Live... )

or at least, in large scale environment, it does...

Default configuration is optimized for small scale deployments. To get decent performance, we've had to tweak, and tweak a lot...

- We are going to look at (finding) tweaking
  - Memory parameters – MySQL, Prometheus
  - IO parameters – Prometheus
  - Database schema and data life cycle management – Query analyzer
Code of conduct

• No snoring!
  • Should the person next to you snore, please poke (gently)
• Questions
  • Please, ask at anytime
What is Verisure

... it's a human right to feel safe and secure ...
Growing Leadership Positions in Europe & Latin America

Continued growth through increasing penetration of existing geographies & repeated, successful expansion into new geographies

Latin America

Europe
Best security solution: attractive product and service offering

<table>
<thead>
<tr>
<th>Product Lineup</th>
<th>Service Lineup</th>
</tr>
</thead>
</table>
| ![Product Illustration](image1.png) | - 24/7 prof. monitored intruder alarm handling  
- Professionally installed and maintained  
- Smoke & fire, water leakage, SOS alarm handling  
- Alarm centre can remotely see pictures taken by cameras  
- Direct user access via App, Web:  
  - Monitor temperature and control heatpump  
  - Smart plugs |

Europe's most widely installed home alarm

| ![Europe's Illustration](image2.png) | - 24/7 prof. monitored intruder alarm handling  
- Professionally installed and maintained with yearly customer visits + 24/7 technical service  
- Smoke & fire, water leakage, SOS alarm handling  
- Alarm centre, user can remotely see pictures taken by cameras  
- Alarm centre calls back via alarm panel  
- Direct user access via App, Web:  
  - See pictures from security cameras |
Innovation is embedded in our culture

- Research & Development centers in Madrid and Malmö
- Innovations for a breakthrough Product & Service Proposition: identify, research, develop, test & refine advanced security solutions
- +300 R&D and IT leading experts
- Our Company invests more than 2% of revenue in new innovation
- Develop exclusive hardware and software features
Who are we?

Kristofer Grahn (kristofer.grahn@verisure.com)
- Senior Systems Specialist
  - But mostly Dba :)
  - Cassandra
  - Mysql
- Missing Netware (Things where better..)
- Sysadmin from 2001
- Dba from 2010

Johan Nilsson (johan.nilsson@verisure.com)
- Unix/Linux/Network admin (since 1999)
- MySQL DBA (since 2000-ish...)
- Oracle 11g DBA OCP (since 2008)
Our environment

... one server more ..
Production environment
What do we monitor with PMM

• Mysql
  • 100+ instances
  • 5.5,6,7
  • Oracle / Percona
• ProxySQL
  • 20+ instances
  • Connection pooling
  • Firewall / Query rewrite (Soon)
Production environment

- Core application
  - Sharding
  - AA/MM
  - Vm's
- 3-party / Legacy
  - AP/MM
  - Hw/Flash
  - ProxySql
- CentOS
- On Prem
PMM setup

... first there was an old server under a desk ...
Specs PMM v1

- Old hardware
- 2x6-core Intel Xeon X5675 @ 3.07 GHz
- 142 GB RAM
- 2x 300 GB SAS for OS
- NetApp mounted via NFSv3 (32k rsize/wsize) for pmm-server-data

running PMM 1.2.2 in Docker, with MySQL in host OS
Performance / bottlenecks PMM v1

Ineffective memory parameters in Prometheus – generating loads of disk IOs
Loads of disk I/O on non-NVMe – leading to high cpu-load
Specs PMM v2

- 2x8-core Intel Xeon E5-2667 v4 @ 3.20GHz
- 256 GB RAM
- 2x 300 GB SSD for OS
- 2x 1.6T NVMe for pmm-server-data

Moved tuned PMM 1.2.2 to new hardware
- Load avg 20-30 → 5-10
- IO-wait 30% → 5%
Tuning with sledgehammer and axe

... when all you have is a hammer, every problem is a nail ...
Broken default values...

Tuned 1.2.2 vs 1.8.1 on the new server
**Docker dis-assembled**

Most configuration found in supervisord-config
– also useful for stopping/starting/restarting individual services

**Moving MySQL out from Docker**
- Percona server 5.7.21-20 instead of 5.5.59-38
- Changing all services to use host MySQL
- Partitioned pmm.query_class_metrics – inserting ~15M rows/24h
- Added partitioned archive-table for query_class_metrics, and moved both to TokuDB - to hold 60 days query statistics

**Adding Apache as reverse proxy (for LDAP-auth)**

**Modified memory parameters for Prometheus**
– target heap size, checkpoint interval, dirty series etc
Broken default values – MySQL

Any guess as to when we restarted MySQL with better parameter values?
Broken default values – Prometheus
PMM 1.2.2 vs 1.8.1 after tuning-session
Bonus features

Query statistics queries
TokuDB for disk saving
Integration with other data sources for Grafana
MySQL-replication / Percona XtraDB Cluster
Separation of services – "scale out"
Pulling PMM apart – limb for limb...

Pros:
• Better / simpler performance optimization
• Freedom in upgrading / tweaking components
• Modified Grafana-pages / templates not overwritten
• Added data sources

Cons:
• Unsupported from Percona (officially)
• Difficult to upgrade PMM
• All component configuration must be reverse-engineered
Finding problems

… that should not happen ?…
Someone running a nasty query?
Finding top-n queries

```sql
mysql> select i.name 'Server',
        > c.name 'Summary',
        > sum(c.query_count) 'Count',
        > min(c.query_time_min) 'Min runtime (s)',
        > avg(c.query_time_avg) 'Avg runtime (s)',
        > max(c.query_time_max) 'Max runtime (s)',
        > avg(cROWS_sent_avg) 'Avg rows sent',
        > avg(c.Rows_examined_avg) 'Avg row examined'
        > from query_class_metrics a
          inner join query_classes c on m.query_class_id = c.query_class_id
          inner join instances i on m.instance_id = i.instance_id
        > where i.name like 'Myhotel%' and m.start_ts > date_sub(now(), interval 1 hour)
        > group by 1.m.query_class_id (i.e. check issues having avg(c.query_time_avg) > 1)
        > order by avg(c.query_time_avg) desc limit 10;
```

<table>
<thead>
<tr>
<th>Server</th>
<th>Summary</th>
<th>Count</th>
<th>Min runtime (s)</th>
<th>Avg runtime (s)</th>
<th>Max runtime (s)</th>
<th>Avg rows sent</th>
<th>Avg row examined</th>
</tr>
</thead>
<tbody>
<tr>
<td>segum-myhotel-01</td>
<td>SELECT sdi100_panels sdi100_actions</td>
<td>4</td>
<td>72.4098</td>
<td>3.0277939354248</td>
<td>84.2289</td>
<td>3170.0000</td>
<td>53546506.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT sqllog_acceps hemosend.sqllog</td>
<td>3</td>
<td>62.6721</td>
<td>62.6721</td>
<td>62.6721</td>
<td>13539.0000</td>
<td>10339.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT sqllog_acceps voderfone.sqllog</td>
<td>5</td>
<td>36.5774</td>
<td>26.5774</td>
<td>36.5774</td>
<td>0.0000</td>
<td>1984803.0000</td>
</tr>
<tr>
<td>sys-myhotel-01</td>
<td>SELECT op_m_parent_log</td>
<td>24</td>
<td>24.4352</td>
<td>24.4352</td>
<td>24.4352</td>
<td>0.0000</td>
<td>3436300.0000</td>
</tr>
<tr>
<td>sys-myhotel-01</td>
<td>SELECT op_m_parent_log</td>
<td>9</td>
<td>24.4352</td>
<td>24.4352</td>
<td>24.4352</td>
<td>0.0000</td>
<td>3436300.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT sqllog_acceps voderfone.sqllog</td>
<td>7</td>
<td>21.0061</td>
<td>21.0061</td>
<td>21.0061</td>
<td>1.0000</td>
<td>6494325.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT op_m_parent_log</td>
<td>7</td>
<td>21.0061</td>
<td>21.0061</td>
<td>21.0061</td>
<td>1.0000</td>
<td>6494325.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT UNION eventum.eventum_issue eventum.eventum_note eventum.eventum_time_tracking eventum.eve...</td>
<td>7</td>
<td>19.7853</td>
<td>19.7853</td>
<td>19.7853</td>
<td>0.0000</td>
<td>33070502.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>SELECT UNION eventum.eventum_issue eventum.eventum_note eventum.eventum_time_tracking eventum.eve...</td>
<td>24</td>
<td>19.1329</td>
<td>19.1329</td>
<td>19.1329</td>
<td>0.0000</td>
<td>33070502.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>UNION eventum.eventum_issue eventum.eventum_note eventum.eventum_time_tracking eventum.eve...</td>
<td>3</td>
<td>5.67066</td>
<td>5.67066</td>
<td>5.67066</td>
<td>2513.0000</td>
<td>822275.0000</td>
</tr>
<tr>
<td>secol-myhotel-06</td>
<td>UNION eventum.eventum_issue eventum.eventum_note eventum.eventum_time_tracking eventum.eve...</td>
<td>24</td>
<td>4.68229</td>
<td>4.68229</td>
<td>4.68229</td>
<td>0.0000</td>
<td>4315823.0000</td>
</tr>
<tr>
<td>18 rows in set (4.90 sec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What's next?

... improvise – adapt – overcome ...
Where do we go from here?

Adding more servers / databases / services to PMM as we grow
Prometheus 2.0
MySQL replication / XtraDB Cluster
Separate PMM-servers for prod and test
Adding development environment to test-installation
Continuous performance improvement (tweaking)
Support for Cassandra ?
We are hiring!

https://www.verisure.se/jobb.html
Open positions

Application Security Lead
Backend Developer within Business Systems
Cloud Infrastructure and Collaboration Specialist – Corporate Systems

**Database Specialist - 24x7 Core Systems**
Delivery Lead IT Operations
Frontend Software Developer - Malmö
Information Security Analysts
Leader within Software Development - Backend Services Manager
Manager Core Systems IT Operations
Network Specialist - IP Communications & Infrastructure
Planning & Supply Manager
Senior Perimeter Security Engineer
Senior Project Manager R&D
Senior Software Developer
Software Project Manager
System Specialist - Core Systems
Test Project Leader
Questions?

Good questions get a gift :)

PERCONA LIVE
Conclusions

... tuning stuff is fun ...
PMM is great!

The functionality PMM provides is well designed and really useful!

- but in large-scale implementations it really needs to be tweaked

Docker / Virtual Appliance is an "easy" and well-functioning way to distribute / provide support for the server-part

- but we'd rather see individually supplied packages and templates, and installation guidelines
- configuration isn't easy to find / tweak, but the gain might be huge
Rate My Session

Large-Scale Percona Monitoring and Management Installation

April 24, 14:20 - 15:10
Room M3

Description:

The presentation will be a real-life study on how we use PMM for monitoring of 120+ MySQL and ProxySQL-servers, as well as query optimisation. During the project, we found a few caveats, that

See More

Speakers

Johan Nilsson
Systems Specialist
Verisure Innovation

Kristofer Grahn
Senior Systems Specialist
Verisure Innovation

Tap a star to rate

Review (optional)

Submit
Thank You!

See you next year!