MySQL in the cloud, as a service

Stewart Smith
Percona

Hi!
I work for

PERCONA

Performance Consulting Experts

which is a MySQL support, consulting and development company

Joined May 2011

Director of Software Architecture

Current title, 2011 onwards

MySQL Cluster

Joined MySQL AB in 2004 hacking on NDB

Drizzle

Core developer, 2008 onwards. For Sun and Rackspace.
somewhat know what I'm talking about.
internals rather than Ops.
You don't want me as your DBA.

```
mysqld --cloud-mode=ON
```

There is no one magic solution to having MySQL As A Service work well.

```
mysqld --this --that --the-other
```

it's a lot of small moving parts and options that need to be set, monitored and configured, and some things aren't going to happen.

Greener grass

We may wish it was different, or look at other database technologies, but there is a lot of legacy code that talks to MySQL, with all it's idiosyncrasies - and we need to be able to support this code.

Magic drop-in replacement

Ignorant, only special cases, or Liars and Charlatans.

What is "MySQL"?

Oracle MySQL
Percona Server
MariaDB
For all intents and purposes, pretty similar.
MariaDB is going their own way, do not expect 10.0 or later to be so interchangeable or compatible.

Drizzle

Is different
DML is similar
DDL mostly similar
managemnet is *different*

DBaaS

DataBase As A Service
typically no SSH to host, just TCP to MySQL
Not just for cloud providers.
can make sense for deploying DB apps internally.
or just staying slightly less insane
knowledge how things work help for those writing/deploying apps

Bare Metal

1 tenant per metal is much easier
close to traditional tuning.
costly though.

Thank you.

all is good in the world
Eat My Data: how everybody gets POSIX file IO wrong

"A Bit of A Whiny Bitch-fest, 4 stars"

- OSCON attendee feedback

Multi-tenancy

one tenant per machine is easy.
just like regular DB servers
relatively solved problem.

Shared hosting

an account on a MySQL DB
This is a BAD idea.

HOWTO: DoS Shared Hosted MYSQL

START TRANSACTION WITH CONSISTENT SNAPSHOT

and then wait. for a long time.
Assuming InnoDB
ENOSPC and then everything explodes.

Replication?

create larger table, run ALTER TABLE. In a loop.

So... not shared hosting.

Virtual Machines

Because they're cool.
and you can use them to handle the fundamental resource constraint
True parallel replication

Damage control

DBaaS is basically this.
database servers are resource hogs
MySQL codebase is not going to be clean any time soon
or ever, that's what Drizzle is for.
So, to support legacy apps, we have to properly
partition resources. Hopefully your hypervisor helps.

IOPs

if you can partition RAM, disk space and IOPs between VMs,
this gets you the best MySQL isolation out there.

IOPs not iHOP

This joke works better in USA
MySQL eats IOPs for breakfast.
single COMMIT = *MULTIPLE* fsync()s

group commit

MariaDB with group-commit
Percona Server with group-commit
possibly increases number of tenants/machine

MyISAM: Just Say No

Not crash safe
replication will be problematic
not high perf
dead

InnoDB

this is what you should pretty much enforce.
MyISAM recovery/consistency steps going to be
a support nightmare.
set the default engine

mysqld --enforce_storage_engine=InnoDB

in Percona Server
error if the NO_ENGINE_SUBSTITUTION SQL mode is enabled
warning if NO_ENGINE_SUBSTITUTION is disabled

ENOSPC

Never, ever, ever, ever, ever hit ENOSPC on MySQL
Remember DoS? It gets worse.
binlogs for replication

InnoDB and ENOSPC

You're probably using InnoDB
You could use file per table
or could statically allocate disk space
cannot dynamically add innodb data files, so file-per-table
just have to monitor, currently no good solution.

InnoDB UNDO

This is for MVCC
transactions get consistent read view
cannot purge rows that can still be seen
UNDO can grow forever

kill_idle_transactions

percona server
hold locks, prevent purge
doesn't help non-idle txns
(maybe doesn't matter)
write a script yourself.

binlog and ENOSPC

world of hurt if you ENOSPC
rotating binlogs

max-binlog-size

size of individual file
not total.
useful for working around ext3

max-binlog-files

in Percona Server
set to 20, 20 * max-binlog-size is max.
server rotates files on other conditions, so this is upper limit.

Temporary files

many types of temp files in MySQL
query execution, ALTER TABLE, CREATE TEMP TABLE
and replication
Can be in different places.
Don't assume /tmp is just it.
unlink, so du isn't going to work.

Users

If doing own VM, you can relatively safely give them root.
BUT: password recovery.

Password recovery for root

you probably want this
AWS approach of enforcing use of API instead of GRANT
local daemon that does the manual foo

Utility User

again, Percona Server feature.
A user that has root-like qualities (in some ways) but doesn't really exist (and can't be deleted)

Replication

Standard awkwardness applies

Backup

for bootstrapping slaves
also for, well, backup.
mysqldump longer, bad for huge data
XtraBackup also good.
probably want to hide all the details.
need disk space of db size free, or stream

Spawn slaves

for scale out
often from backups
how much control over replication to users?

GOOD LUCK