

# MySQL Multi-Master Replication Failover

A step-by-step explanation



***Walter Heck***

walter@openquery.com



# Presumed knowledge

- Basic Linux Skills
- Basic MySQL Skills
- Basic Understanding of MySQL Replication
- Basic Networking knowledge



# Overview

- Introduction
- How MMM works
- Overview of our setups
- MMM 1.2.3
- MMM 2.0.X

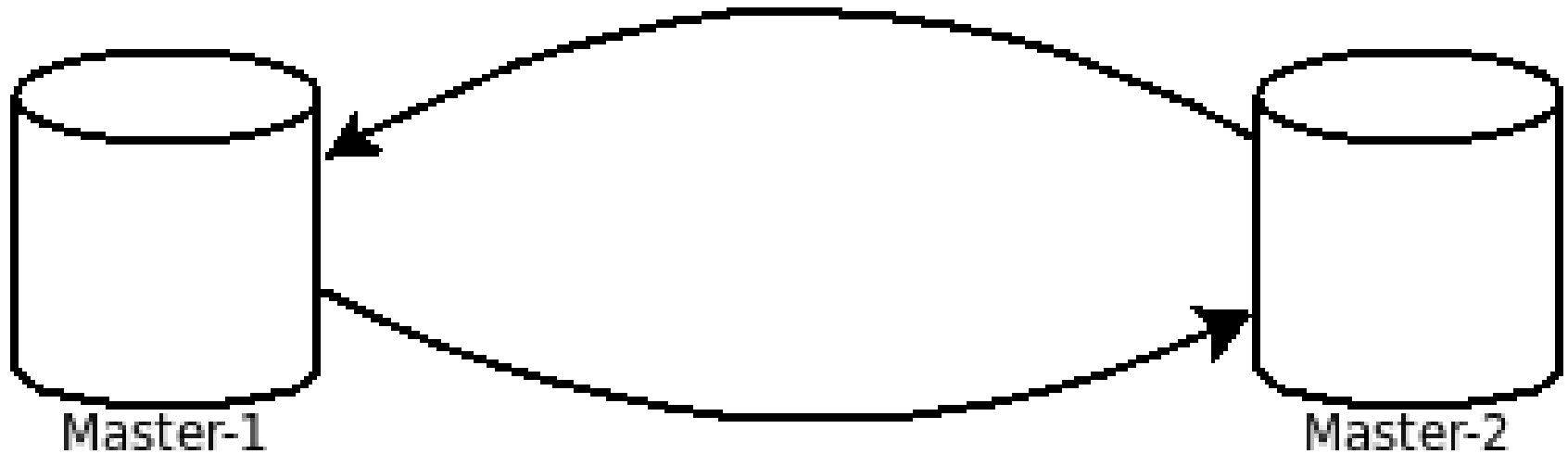


# Introduction

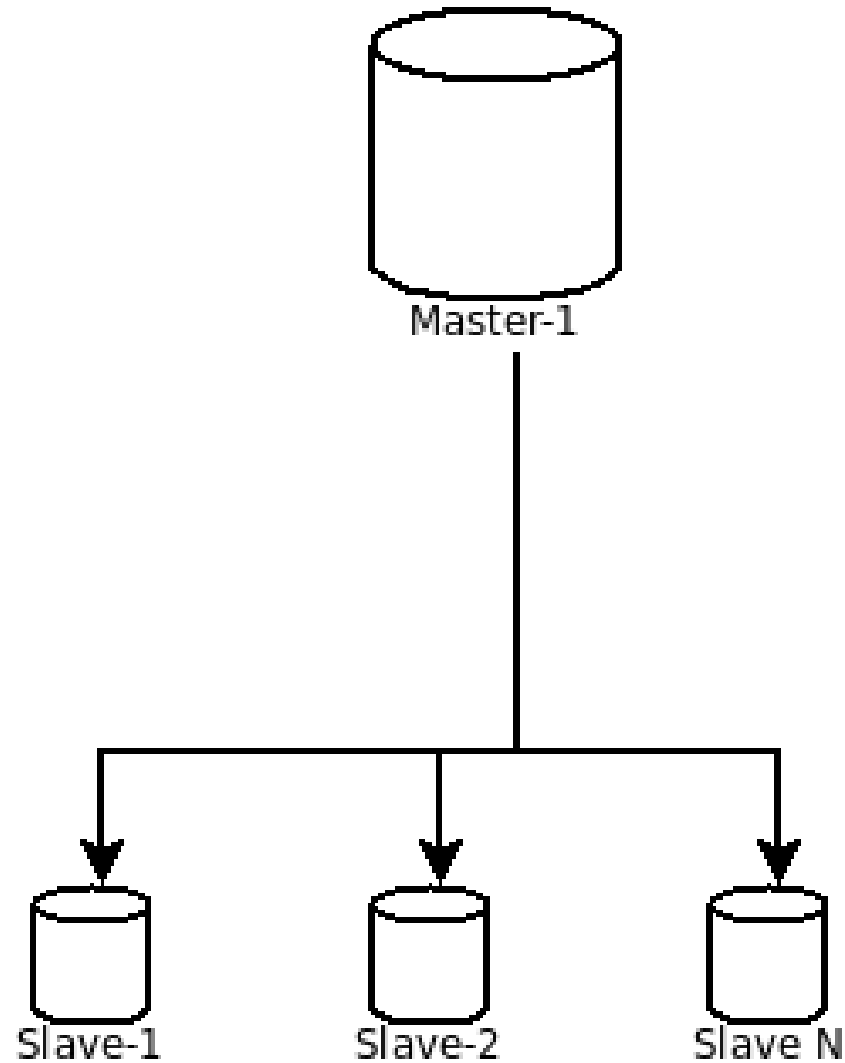
- What is MMM?
  - HA solution for *near* immediate failover
  - Automatic failover for slaves to another host
- What is MMM not?
  - Load balancer
  - 100% Data Reliability (Replication is not perfect!)



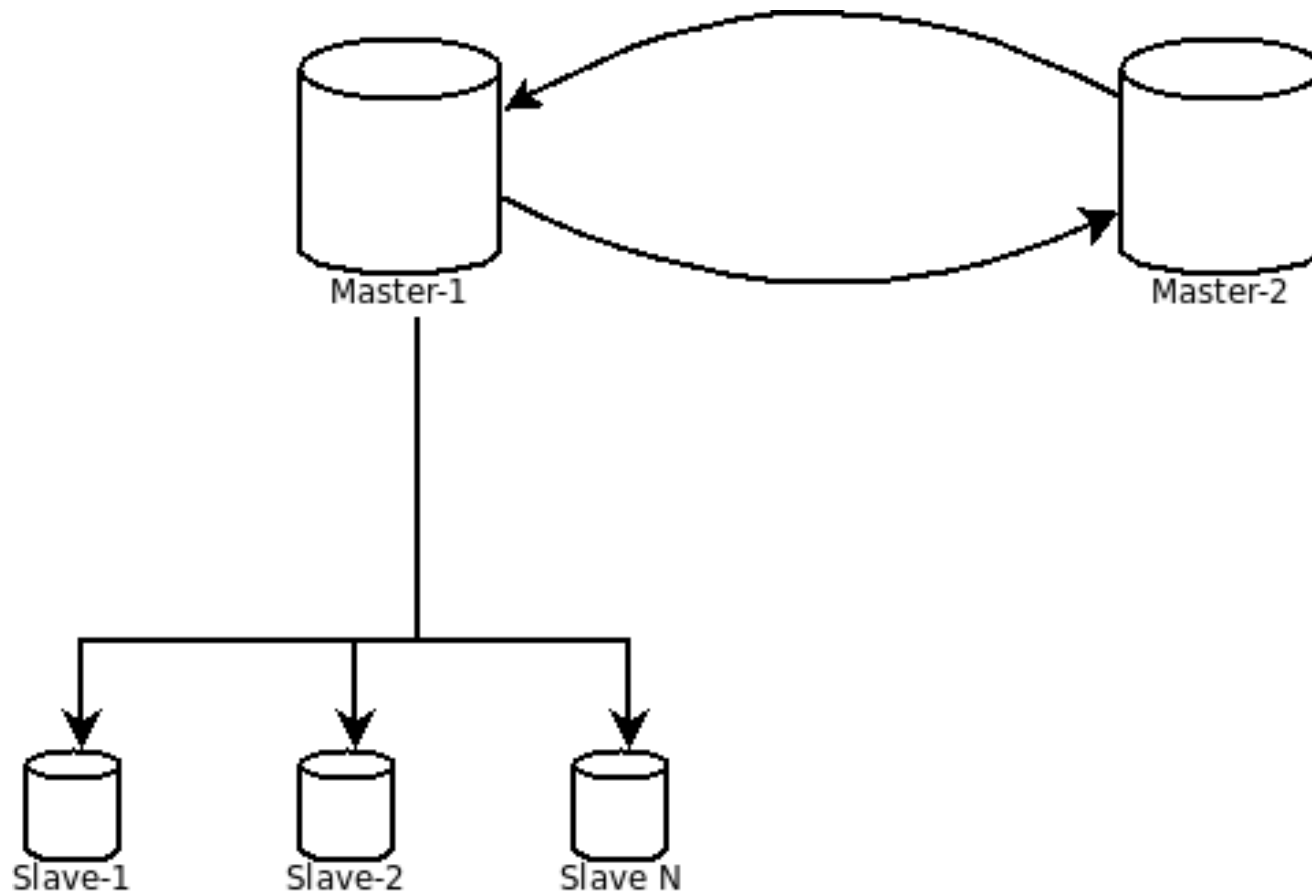
## How MMM works



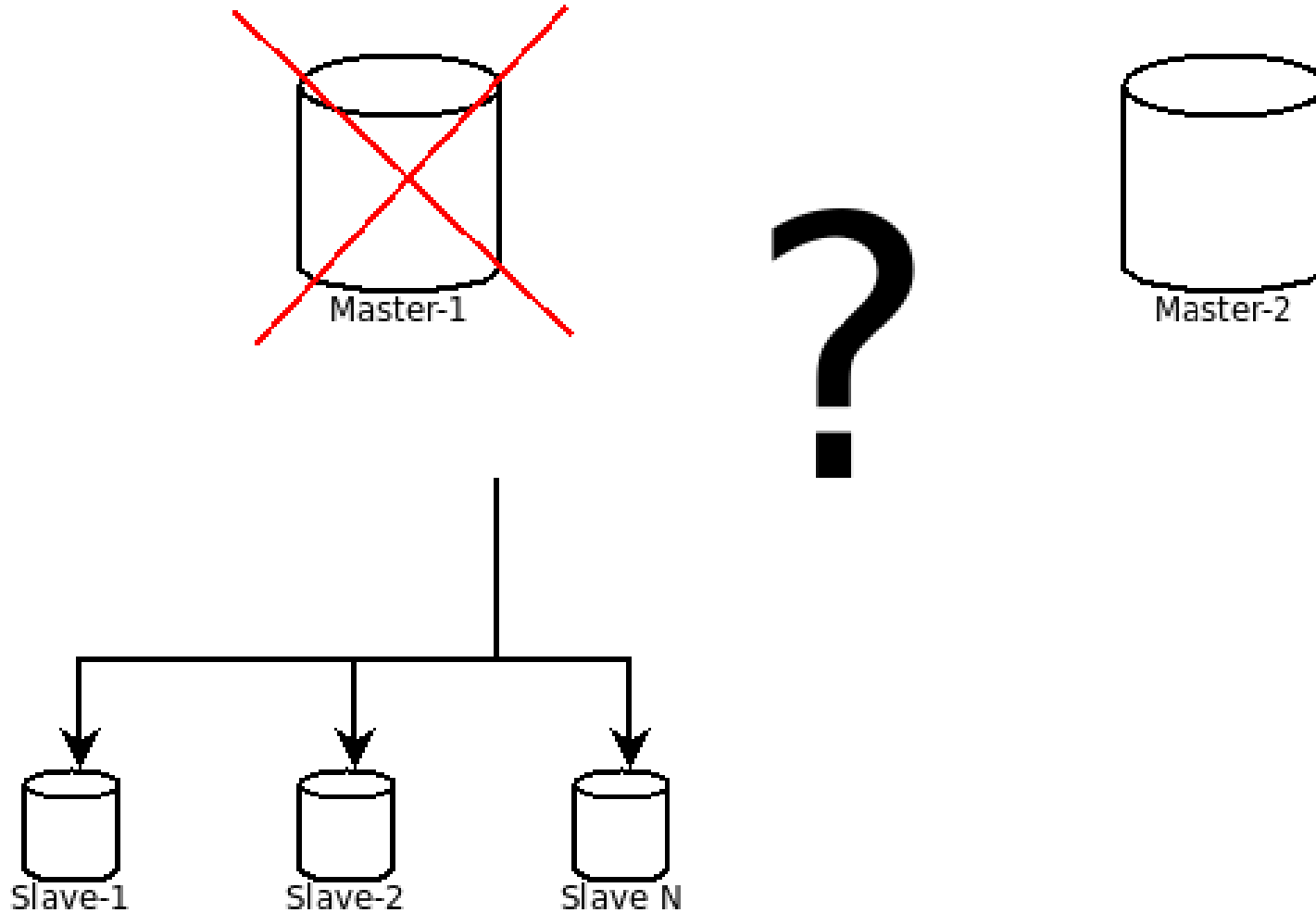
# How MMM works



# How MMM works

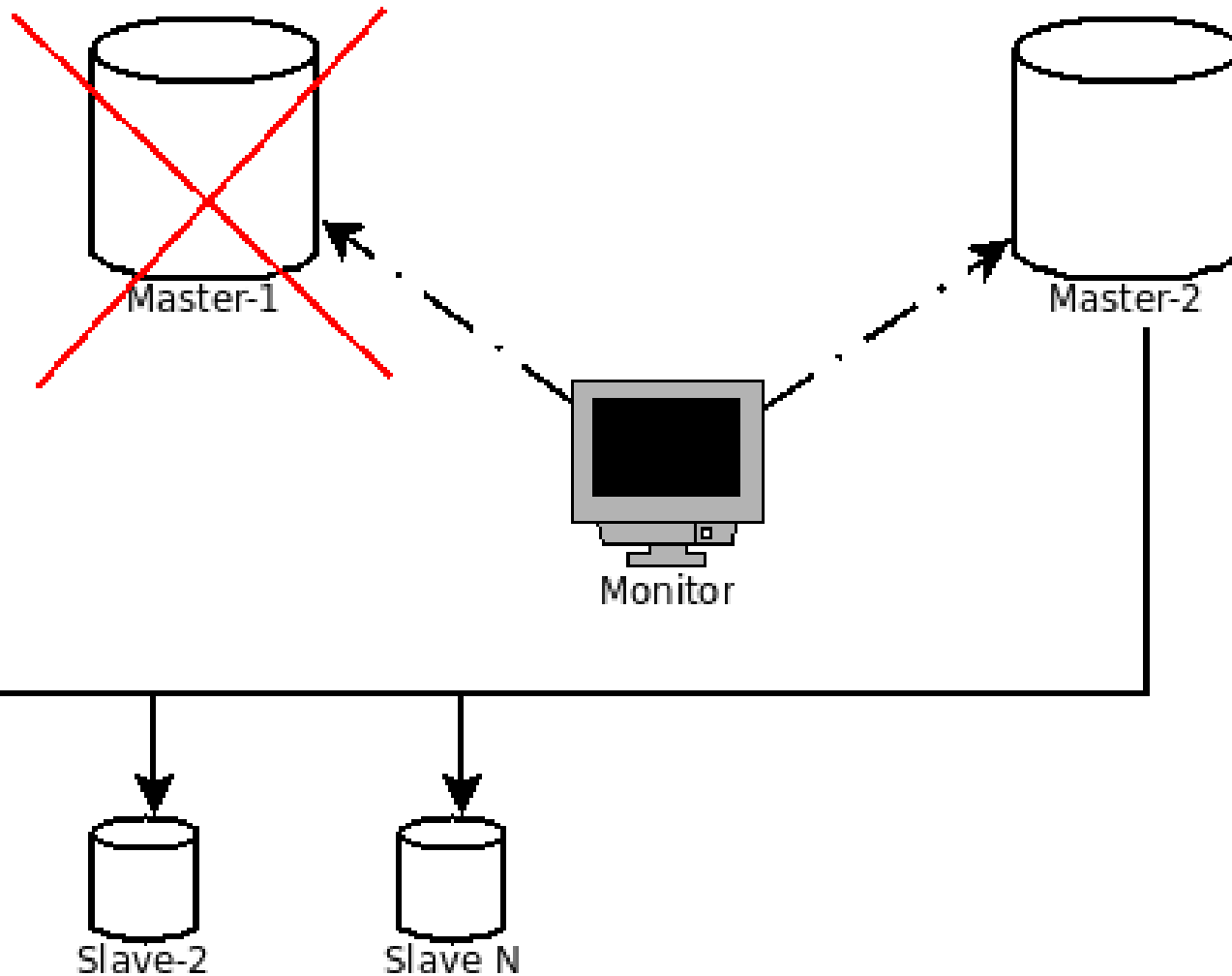


## How MMM works





## How MMM works



# Virtual IP Voodoo

- Exclusive role
  - One virtual IP for multiple machines
  - machine goes down? → ip is moved to another machine
  - Usually used for write-mostly nodes (e.g. masters)
- Balanced role
  - One virtual IP for each machine
  - machine goes down? → ip is moved to another machine
  - Usually used for read-only nodes (e.g. slaves)
  - Caution! Nothing to do with load balancing!
  - One machine can have multiple virtual IPs



# Overview of our setups (1)

- MMM 1.2.3 - In production/testlab
  - 2 Masters, 3 Slaves and 1 Monitoring node
  - Only use virtual IP for Masters
  - Slaves still benefit from MMM at failover
  - HW Loadbalancer for slaves
  - CentOS 5.2, OurDelta 5.0.77build8



## Overview of our setups (2)

- MMM 2.0.9 – In production/testlab
  - Setup ex.: 2 Masters, 2 Slaves and 1 Monitoring node
  - Setup ex.: 2 Masters and 1 Monitor
  - Virtual IPs for Masters and Slaves
  - CentOS 5.3, OurDelta 5.0.77build8



# MMM 1.2.3 System requirements

- Linux, OpenSolaris
  - No windows support!
- Perl
  - Must have ithreads enabled
  - Whole bunch of modules
- fping
- arping/net\_send
- MySQL



## MMM 1.2.3 - advantages

- Most used MMM version
- Has some history in production environments
- Stable
- Code is (mostly) easily readable and understandable



## MMM 1.2.3 - disadvantages

- Code is messy
- Poor documentation
- Has some quirks



# MMM 2.0.X – advantages

- Recommended version
- Complete rewrite
- Much more proper perl
- Uses highly customisable Log4Perl
- Built-in angel processes
- Passive mode
- Preferred roles





## MMM 2.0.X - disadvantages

- Not as many production deployments as 1.2.x
- Bit tougher to install than 1.2.3
  - Only tarball or .deb's, no decent .rpm's
- No `auto_setonline` like MMM 1.2.x



# MMM 2.0.X – 'hardware'/network requirements

- 5 (virtual) machines
  - Monitor → very lightweight
  - Data nodes → your choice
  - When using VM's:
    - distribute properly, clone smartly!
- Static IP for each machine
- 3 virtual IP's
  - 1 for master
  - 2 for slaves



# MMM 2.0.X Setup - OS

- CentOS 5.3 minimal
  - Disable all packages at install
  - Minimal != base
- Install yum-priorities
- Install RPMForge repo
  - <http://wiki.centos.org/Repositories/RPMForge>
- Install perl modules
  - Try 'yum install perl-Log-Log4Perl' instead of CPAN → install Log::Log4Perl



# MMM 2.0.X Setup - Security

- Change SSH configuration
  - Different ports
  - Use SSH Keys for authentication
  - Disable root login over SSH
- Setup iptables (or your firewall of choice)



## MMM 2.0.X – Firewall settings

- Port 9989 on agents should be open for traffic from the monitor node
- Port 9988 on monitor optionally to connect to it remotely
- Port 3306 on agents should be open to it's master and to the monitor node
- If no 'public' interface on data nodes, open SSH from the monitor node



## MMM 2.0.X – MySQL

- Install OurDelta MySQL on each data node
  - Use OurDelta's CentOS repository  
<http://ourdelta.org>
  - Better performance, better monitoring, better tuning
- Minimal settings for all data nodes:
  - enable log-bin
  - Set unique server-id
- Settings for masters:
  - Enable log-slave-updates



# MMM 2.0.X – MySQL Permissions

- Monitor user
  - Privileges REPLICATION CLIENT
- Agent user
  - Privileges SUPER, REPLICATION CLIENT, PROCESS
- Replication user
  - Privileges REPLICATION SLAVE



# MMM 2.0.X – MySQL Replication

- Set up MySQL replication
  - Make master-2, slave-1 and slave-2 slaves of master-1
  - Make master-1 slave of master-2
- Start slaves and check if everything works properly
- If all == ok, start loading data into master1 and watch it replicate through the whole cluster





# Download and install MMM 2.0.X

- No RPM/repository for CentOS
  - Use tarball and separate files and folders
- On all machines (data+monitor)
  - Mkdir -p /var/log/mysql-mmm, /etc/mysql-mmm, /usr/local/mysql-mmm/bin/{agent|monitor}
- Create MMM subdir in perl -V:installvendorlib
  - Copy lib/ folder contents there



# MMM 2.0.X - configuration

```
active_master_role      writer
<host default>
cluster_interface       eth0
pid_path                /var/run/mmm_agent.pid
bin_path                /usr/local/mysql-mmm/bin
replication_user        replication
replication_password    openquery
agent_user              mmm_agent
agent_password          openquery
</host>

<host master-1>
ip      192.168.2.206
mode    master
peer    master-2
</host>

<host slave-1>
ip      192.168.2.208
mode    slave
</host>

<role writer>
hosts   master-1, master-2
ips     192.168.2.201
mode    exclusive
</role>

<role reader>
hosts   slave-1, slave-2
ips     192.168.2.202, 192.168.2.203
mode    balanced
</role>

===

this master-1
```



## MMM 2.0.X – checks

- Checks allow mmm\_mon to monitor health of the cluster
- 4 default checks included
  - Ping → server reachable?
  - rep\_backlog → replication behind?
  - rep\_threads → replication running
  - Mysql → mysql server reachable
- Option for custom checks



# MMM 2.0.X – Agent states

- ONLINE
  - All is peachy, only state in which a node can have a role assigned
- REPLICATION\_DELAY
  - replication backlog is too big (Check rep\_backlog failed)
- REPLICATION\_FAIL
  - replication threads are not running (Check rep\_threads and rep\_backlog failed)
- AWAITING\_RECOVERY
  - Host is awaiting recovery. Entered after HARD\_OFFLINE → all is ok



## MMM 2.0.X – Agent states (2)

- **HARD\_OFFLINE**
  - Host is offline (Check ping and/or mysql failed)
- **ADMIN\_OFFLINE**
  - host was set to offline manually
- **UNKNOWN**
  - Host is in unknown state



## MMM 2.0.X – mmm\_control

- mmm\_control is used to control the cluster
- Needs root privileges (as opposed to MMM 1.X) because of reading config files



# MMM 2.0.X – common operations

- Move masters so you can do maintenance
  - `mmm_control move_role writer my-master-1`
- Set a slave offline for maintenance
  - `mmm_control set_offline my-slave-1`
  - `mmm_control set_online my-slave-1`
- Put MMM in passive state so it doesn't interfere
  - `mmm_control set_passive/set_active`



## MMM 2.0.X – cluster state

- 'mmm\_control show' shows the current state of all agents

```
[openquery@mmm2-monitor ~]$ sudo mmm_control show
master-1(192.168.2.206) master/ONLINE. Roles: writer(192.168.2.201)
master-2(192.168.2.207) master/ONLINE. Roles:
slave-1(192.168.2.208) slave/ONLINE. Roles: reader(192.168.2.202)
slave-2(192.168.2.209) slave/ONLINE. Roles: reader(192.168.2.203)
```

- 'mmm\_control mode' shows whether the cluster is active or passive





## Optional Extras (1)

- Make configuration go through puppet
  - Makes adding new slaves extremely easy
- Install MySQL/MMM monitoring
  - Open Query uses Zabbix
- Preferred hosts
- When using LVM for MySQL, use MMM tools to clone/backup nodes
- Use SSL for MMM connections



## Optional Extras (2)

- Mix and match solutions!
  - Use HW loadbalancer for reads
- Setup cron-jobs with mk-table-checksum / mk-table-synch
  - Make sure schema is suitable



# Ideas for improvement of MMM 2.0.X

- Add 'real' loadbalancer
  - MySQL Proxy?
  - Customised haProxy
- Remove monitor as a SPOF
- Proper packaging



# Pitfalls

- MMM has no 'real' loadbalancing
- Some failover situations will \_not\_ be handled by MMM and need manual intervention!

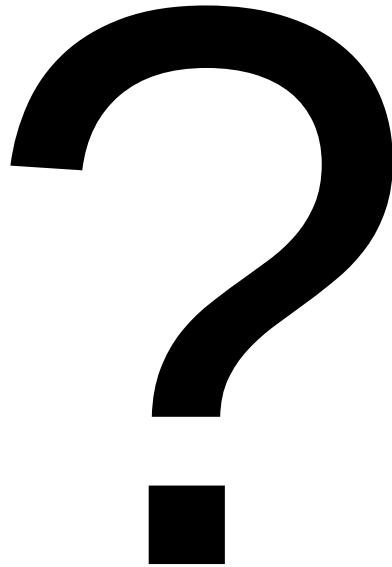


# Who is using it in production?

- Big broadcasting company
  - 2 masters, 3 slaves
- One of top 5 Adult entertainment sites
  - In top 60 sites in Alexa.com
- Big Streaming video provider
  - Never less then 10K people online



# Questions?



## Credits / Links

<http://mysql-mmm.org>

<http://ourdelta.org>

<http://openquery.com>

<http://www.zabbix.com>



***Thank you!***

Walter Heck

walter@openquery.com

