



MySQL High Availability with Pacemaker

Percona Live NYC May 2011

Yves Trudeau

Percona Inc

MySQLPerformanceBlog.com

DIAMOND SPONSORSHIPS

- THANK YOU
 - TO OUR
- DIAMOND SPONSORS



About Me



- <http://www.percona.com>
- <http://www.mysqlperformanceblog.com/>
- <http://www.bigdbahead.com>

Yves Trudeau, Ph.D.
Principal Consultant

MySQL HA Landscape



- Replication
- Pacemaker/Heartbeat/DRBD
- Pacemaker/Heartbeat/SAN
- NDB Cluster

Why not replication?



- Setting up master – slave is trivial
- Handling failover looks easy but is it?
 - Stop writes on the current master
 - Pick the most up to date slave as a new master
 - Wait for new master to be fully in sync
 - Reconfigure all the slaves to the new master
 - Allow writes on new master

What about MMM?



- Supposed to do that
- Many issues and bugs, easy to get broken replication
- Communication layer is weak, easily freezes
- Very basic fencing support
- Project is stagnant

Using Pacemaker?

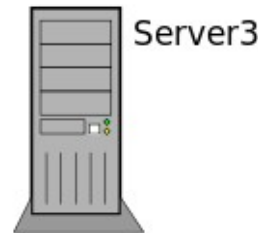
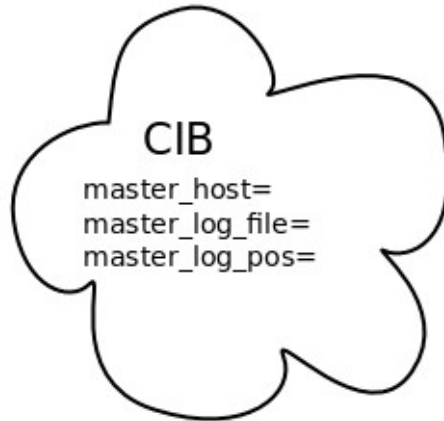


- Amazingly flexible tool, an idea enabler
- Rock solid communication layer
- Can share and distribute information in CIB
- Fencing
- Master-slave resource type
- Monitoring
- Master scores

The principle



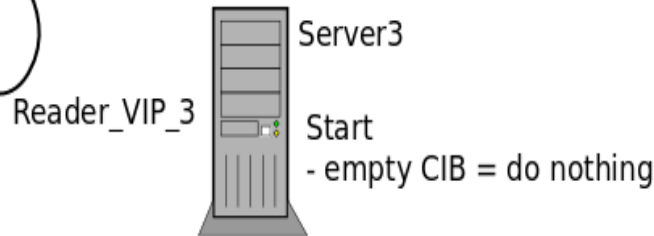
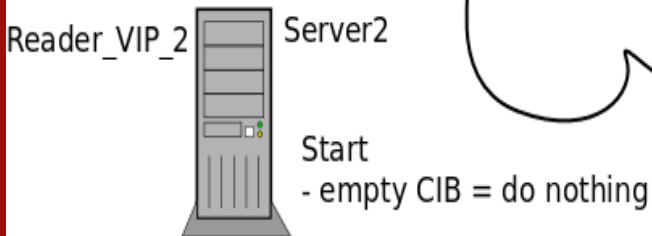
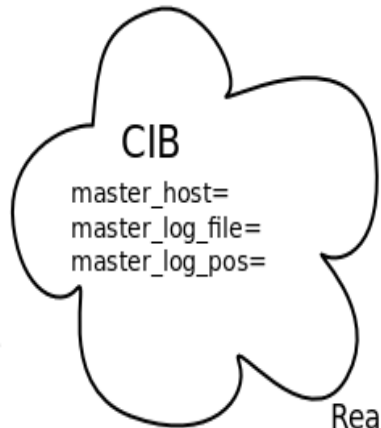
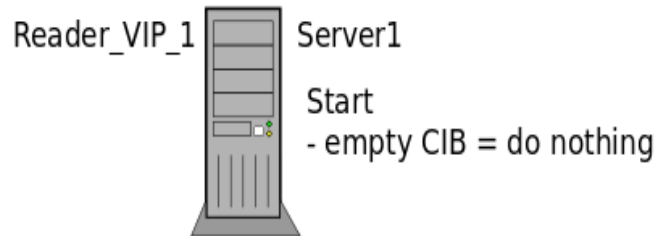
Initial state



The principle



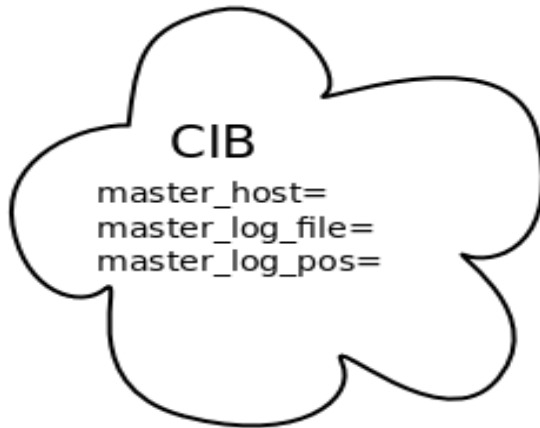
Start resource



The principle



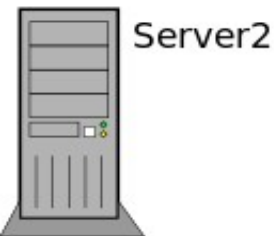
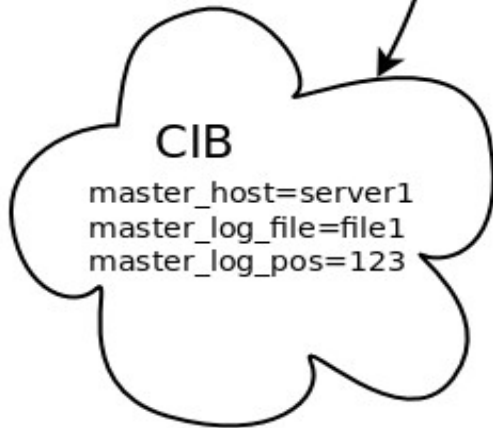
Promote Server1



The principle



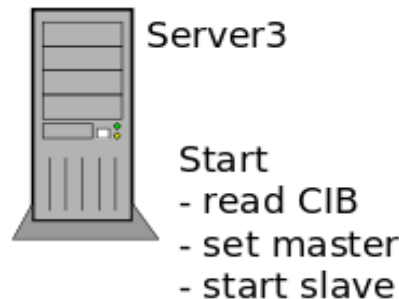
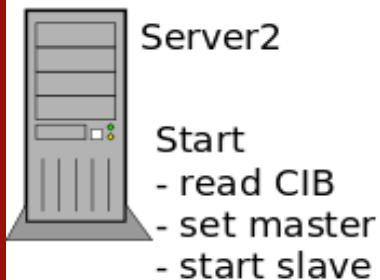
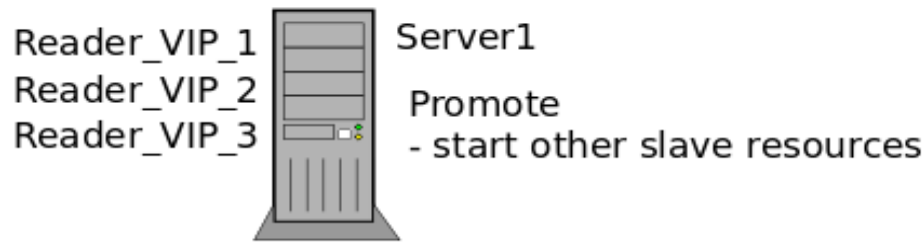
Server1 set CIB



The principle



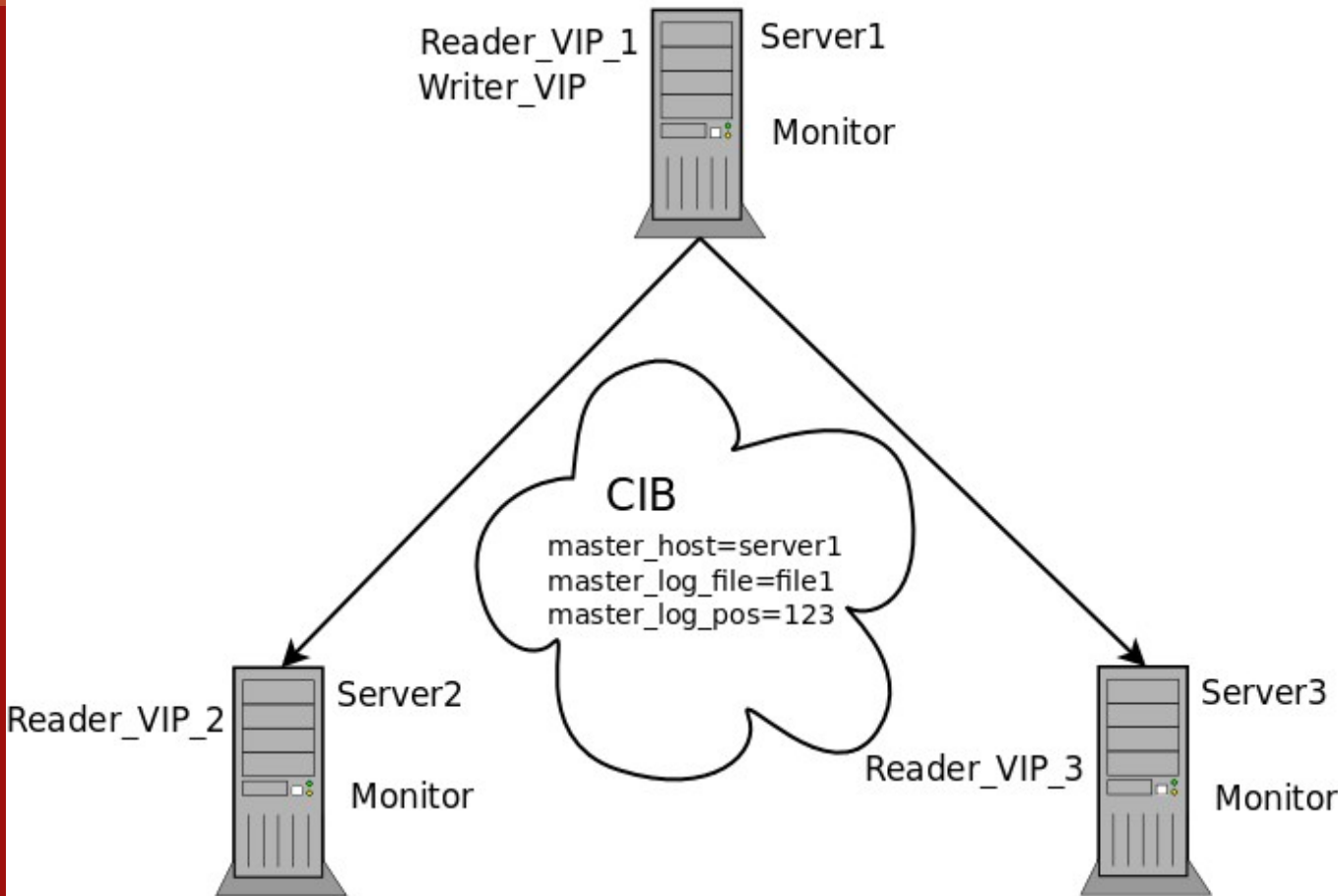
Server1 start other slave



The principle

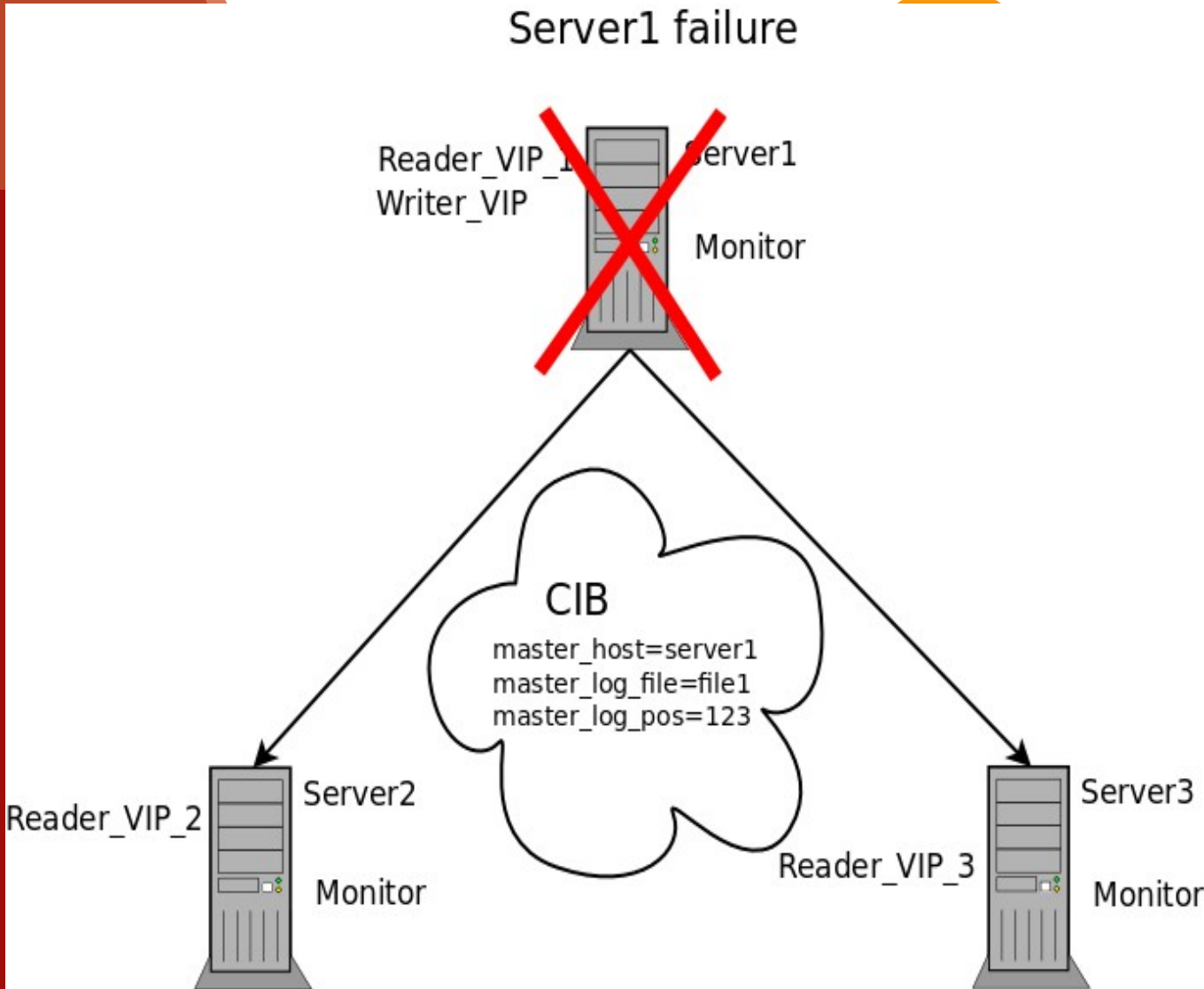


Server1 is master

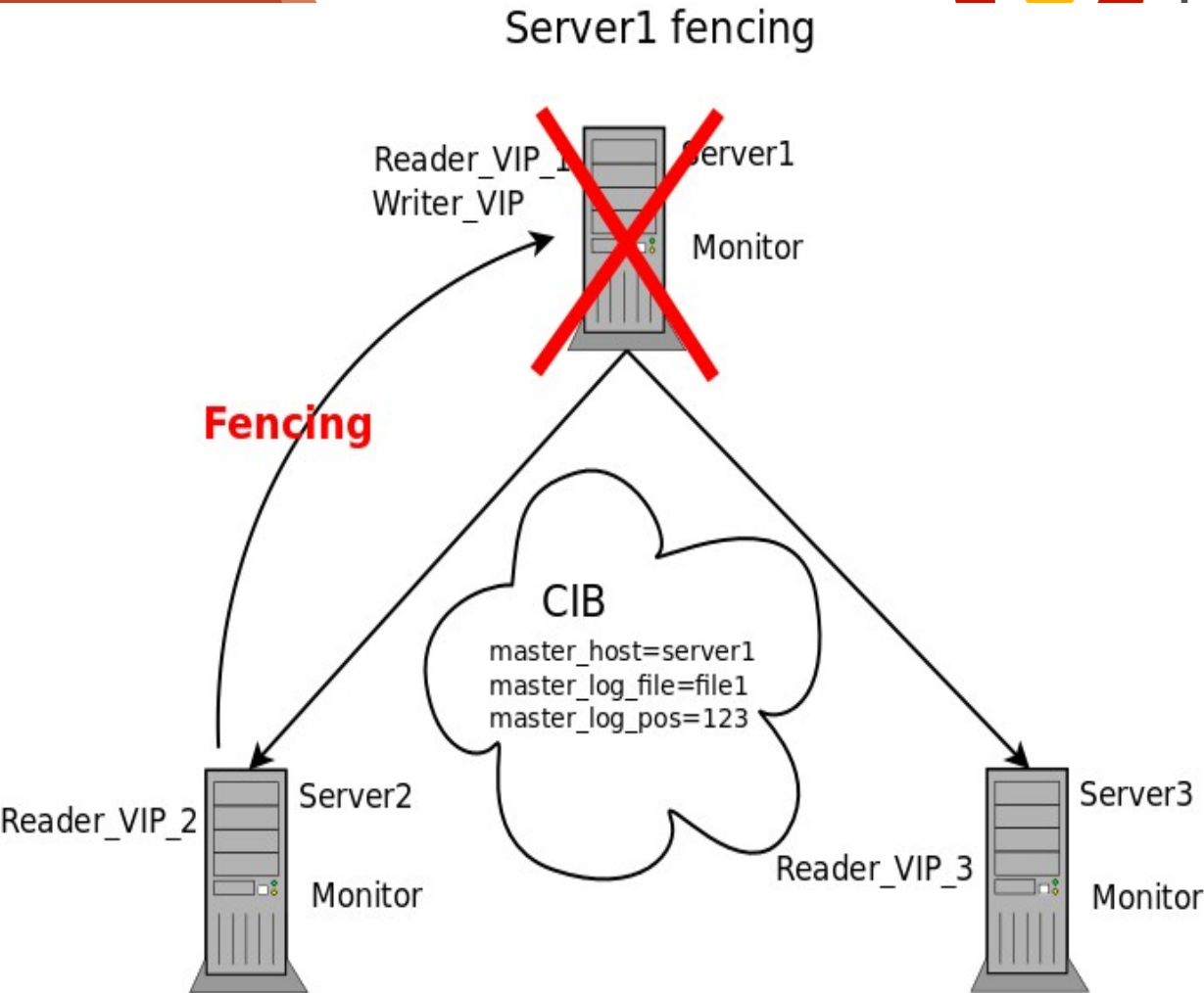


The principle

R C O N A



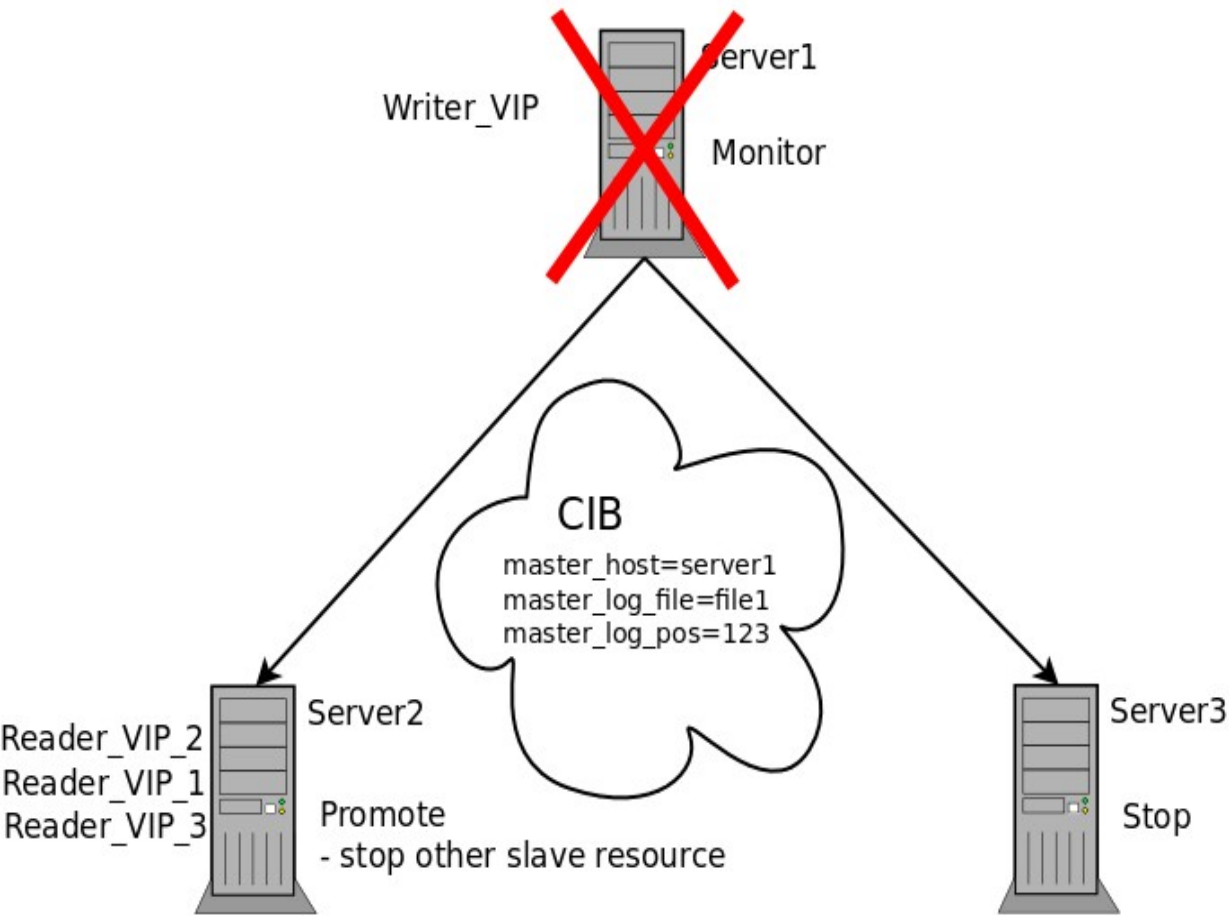
The principle



The principle



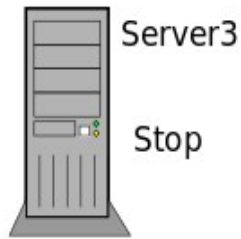
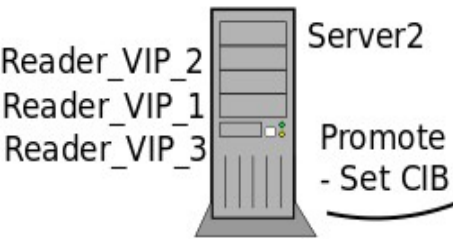
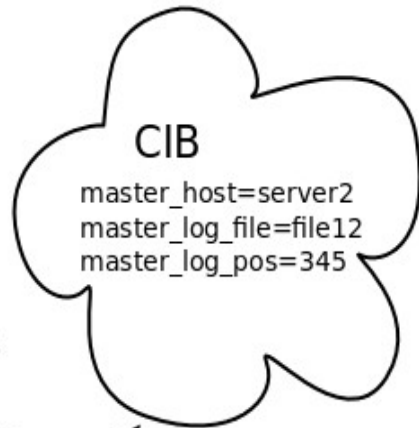
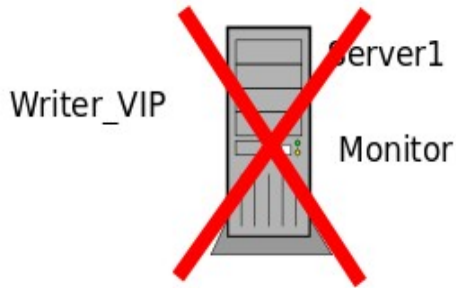
Promote Server2



The principle



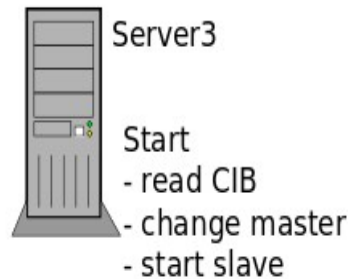
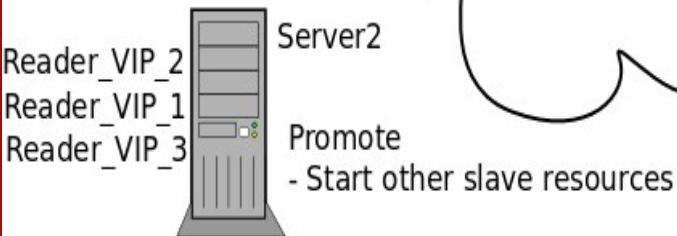
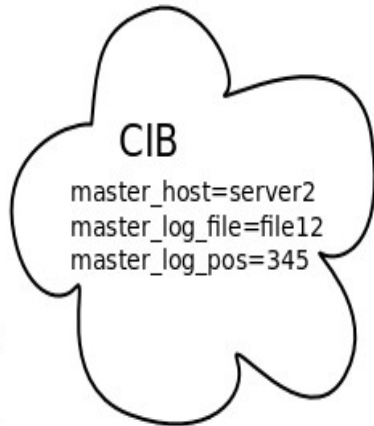
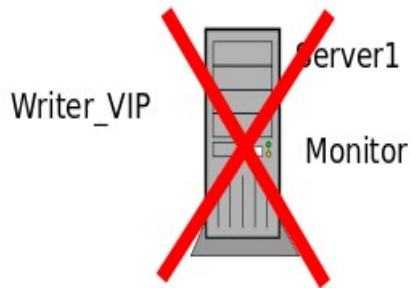
Server2 set CIB



The principle



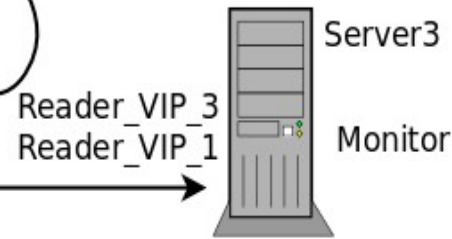
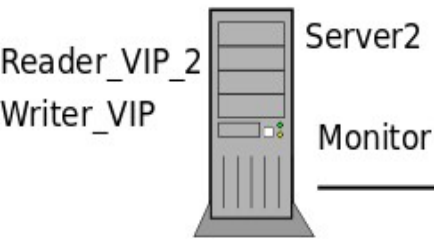
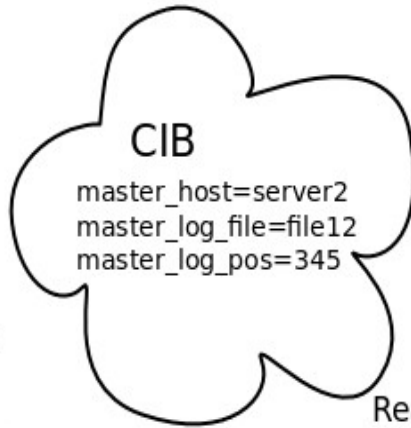
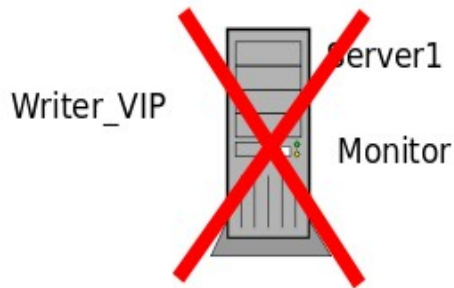
Server2 start replication



The principle



Server2 start replication



Things that can go wrong



- Transactions not replicated before crash
 - semi-sync?
- Chosen new master is not the most up to date
 - Master scores
- Master switch while slave is down

The project



- Percona replication manager
- Hosted on LP
 - <https://launchpad.net/percona-prm>
- Still young but performs the basic
- Lot of things to improve
- 2 parts: the Pacemaker resource script and Pacemaker configuration file

Todo



- Helper tool to generate Pacemaker configuration
- Add mk-heartbeat functionality for more precise replication delay
- Go beyond master-scores to pick up the best slave to promote
- Trap more errors like “Too many connections”

Questions



Questions?

Yves Trudeau: yves@percona.com

<http://www.percona.com/>

<http://www.mysqlperformanceblog.com/>