



**ORACLE®**

Pavan Venkatesh  
Senior Sales Consultant

## ***MySQL Replication and Scalability***



# AGENDA

- Replication and Scalability basics
- Read and write scalability options
- Replication use cases
- Semi-sync, replication heartbeat
- Customer success stories
- Enterprise tools
- Highlights of 5.5 and 5.6
- Q&A



# What is Replication?

- “The duplication of data to one or more locations”
- Asynchronous
- Synchronous
- Statement-Based
- Row-Based
- Mixed-Format



# What is Database Scale-out?

Scale-Out is a modern computing architecture that enables organizations to improve application performance and scalability on an incremental, as-needed basis by adding multiple replicated database servers on low-cost commodity hardware.

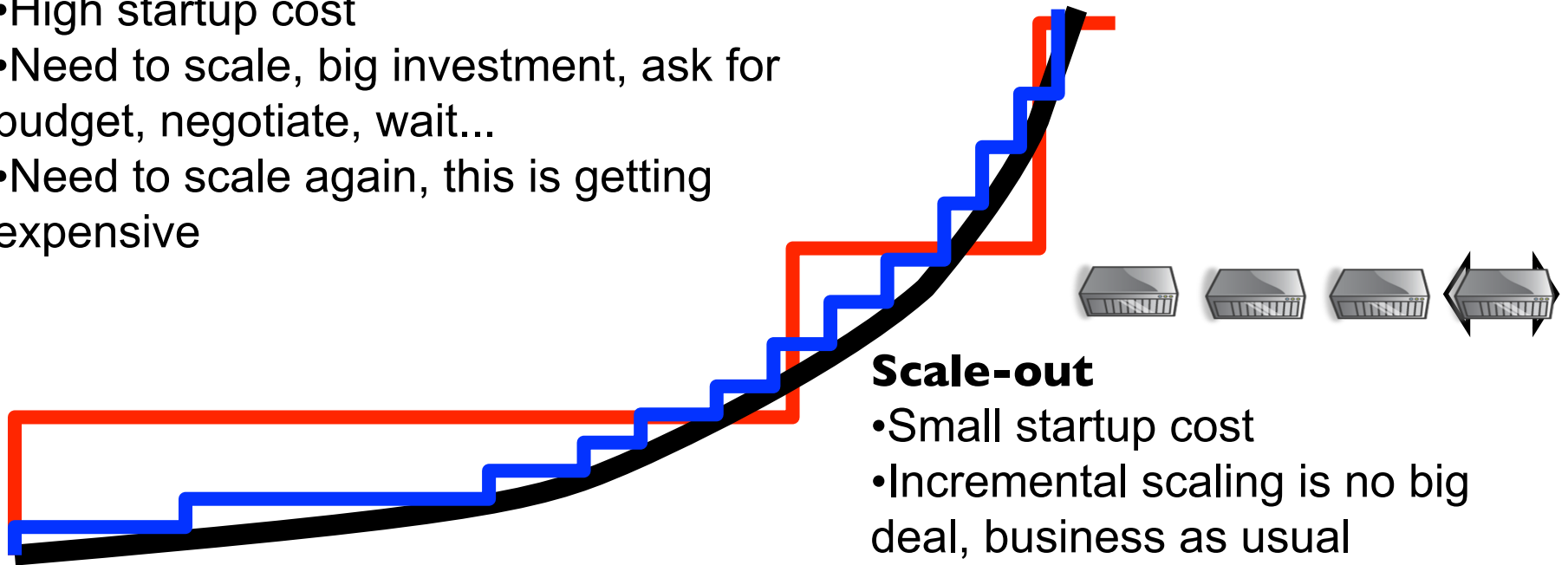
In the online world, many of the largest and fastest-growing companies use MySQL to cost-effectively Scale-Out their successful businesses including Google, Yahoo, Craigslist, Ticketmaster, Wikipedia, YouTube.

# Scalability Strategies



## Scale-up

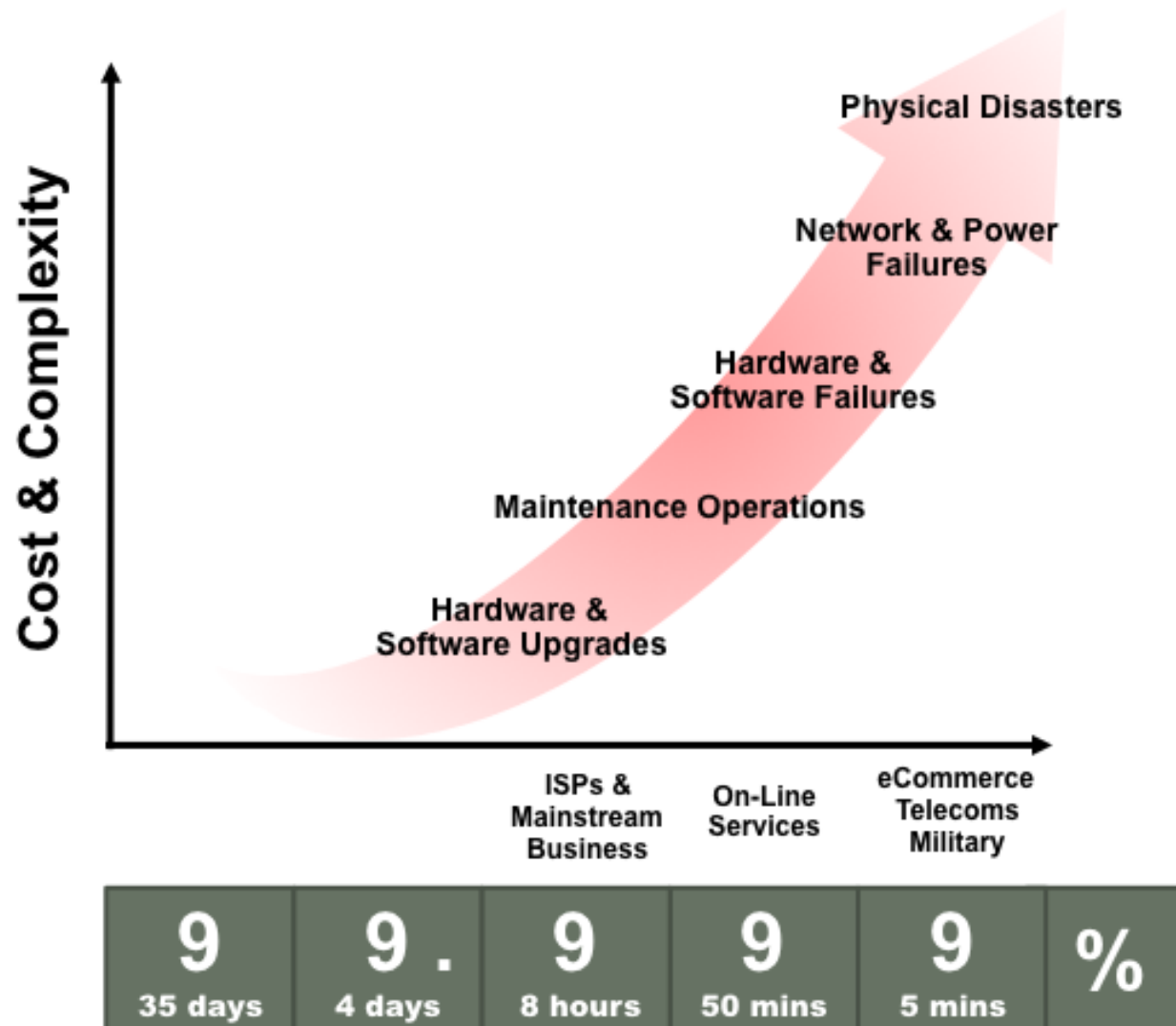
- High startup cost
- Need to scale, big investment, ask for budget, negotiate, wait...
- Need to scale again, this is getting expensive



## Scale-out

- Small startup cost
- Incremental scaling is no big deal, business as usual
- No limit to scaling
- Btw, what is "commodity hardware"?
  - 8 cores? 16? 32?

# Considerations for HA





# MySQL Replication

- Asynchronous
- Manual Failover
- Easy to setup, configure & maintain
- Very Popular for Scale Out
- Fast Log shipping
- Can replicate from one engine to another

## Disadvantages

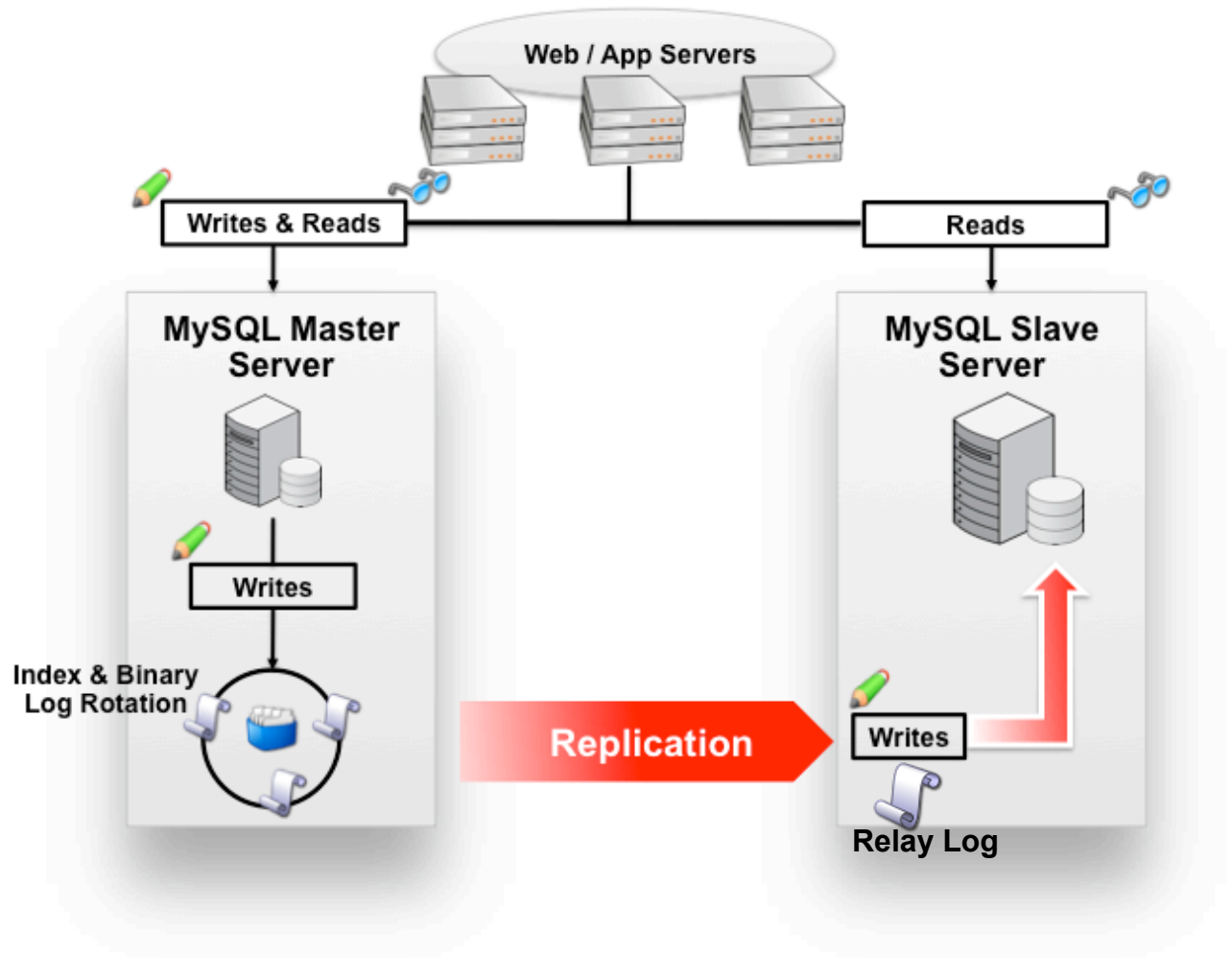
- Not true HA - Data can be lost on system failure
- Complex fail-over/fail-back with more than 1 slave
- Slaves can fall behind



# High Availability & Scalability

## MySQL Replication

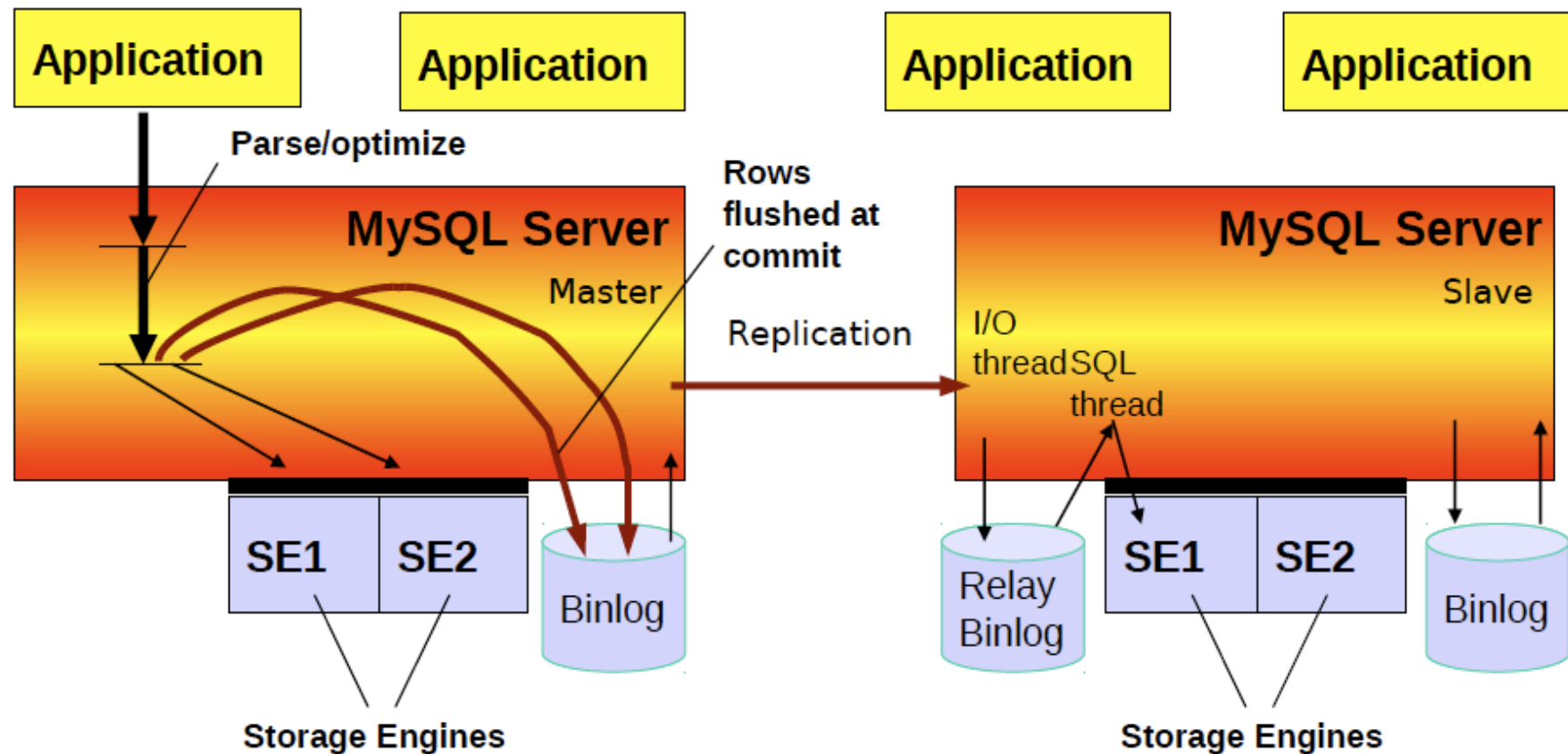
- Native in MySQL
- Used for Scalability and HA
- Asynchronous as standard
- Semi-Synchronous support added in MySQL 5.5
- Each slave adds minimal load on master





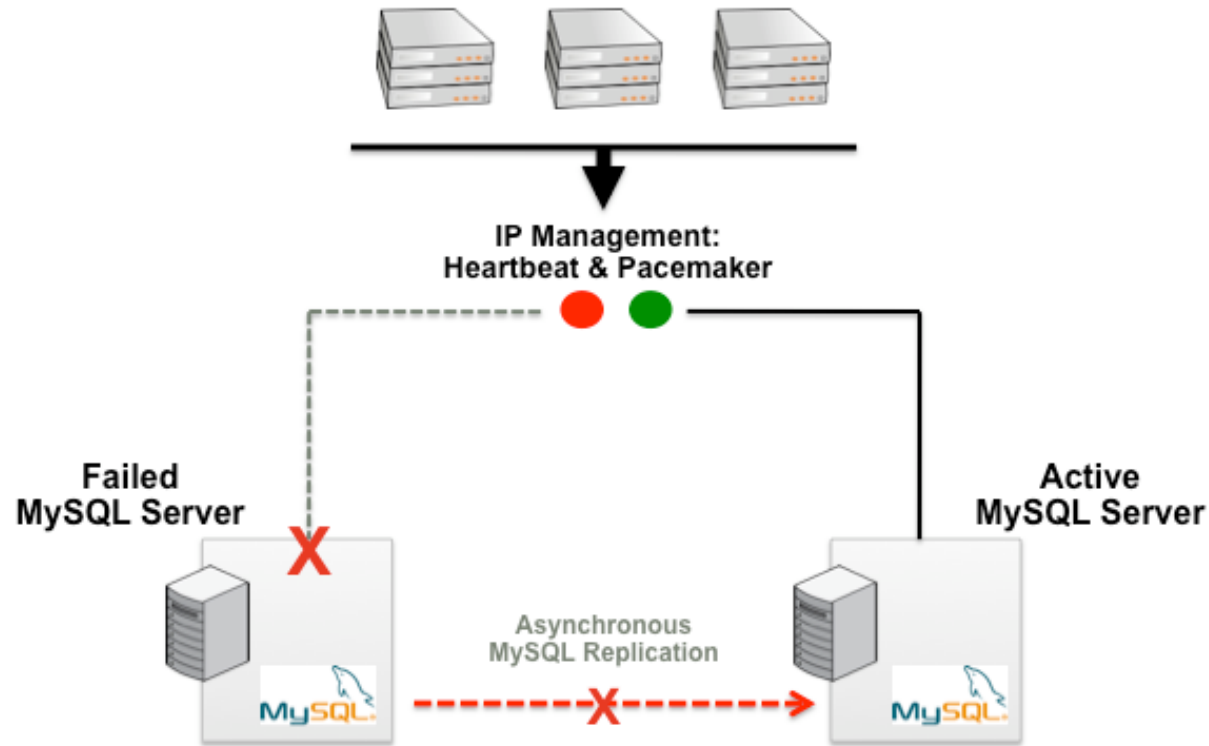
# MySQL Replication Architecture

## MySQL 5.1: Row-based replication



# Building on Replication

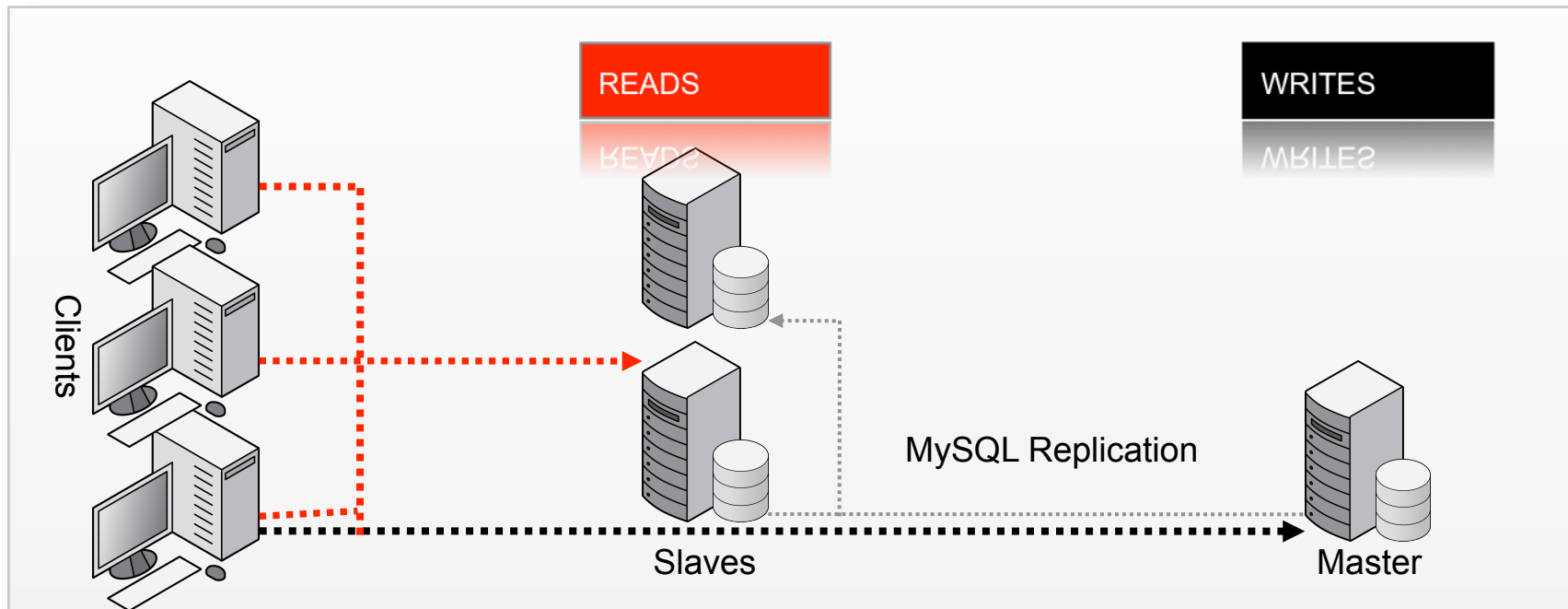
## Failure Detection & Failover



- Linux Heartbeat implements heartbeat protocol between nodes
- Failover initiated by Cluster Resource Manager (Pacemaker) if heartbeat message is not received
- Virtual IP address failed over to ensure failover is transparent to apps

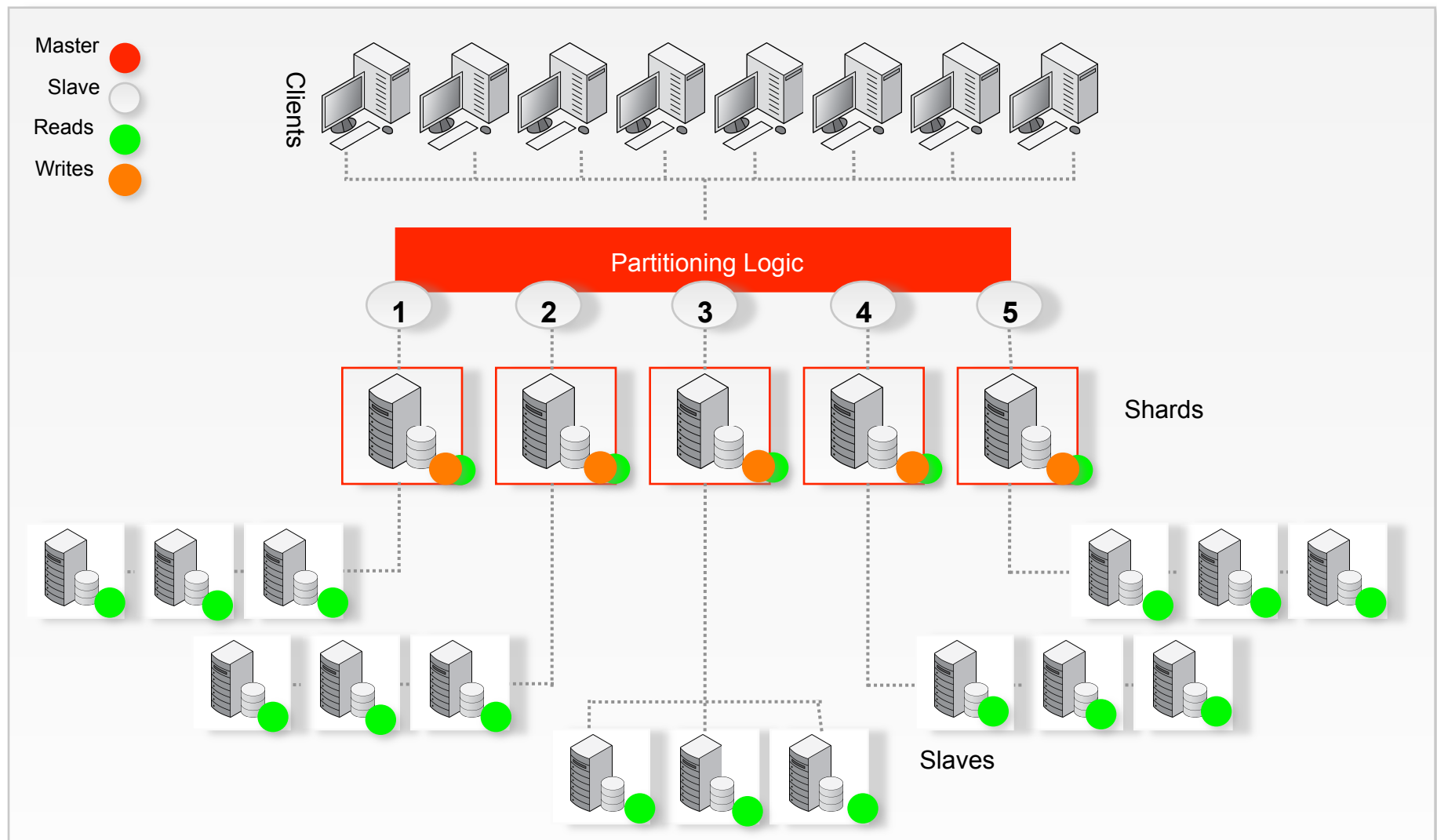
# MySQL Replication

## Read Scalability



- Used by leading web properties for scale-out
- Reads are directed to slaves, writes to master
- Delivers higher performance & scale with efficient resource utilization

# Sharding aka Application Partitioning



# Replication Topologies

**Single**



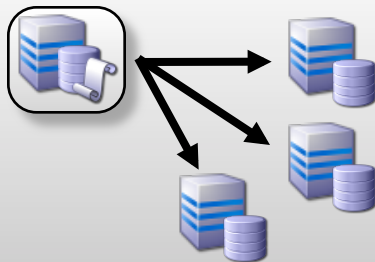
**Chain**



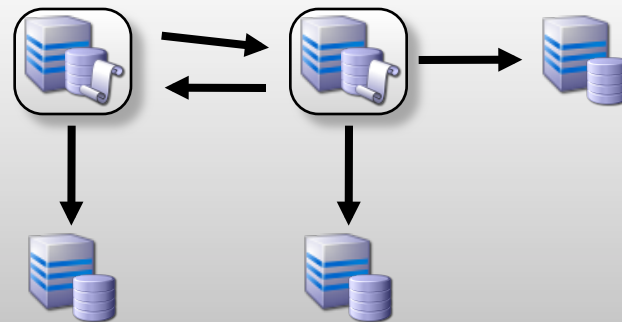
**Circular**



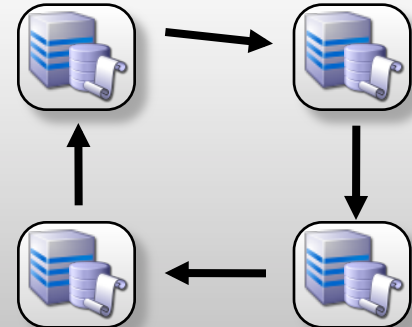
**Multiple**



**Multi - Master**



**Multi - Circular**



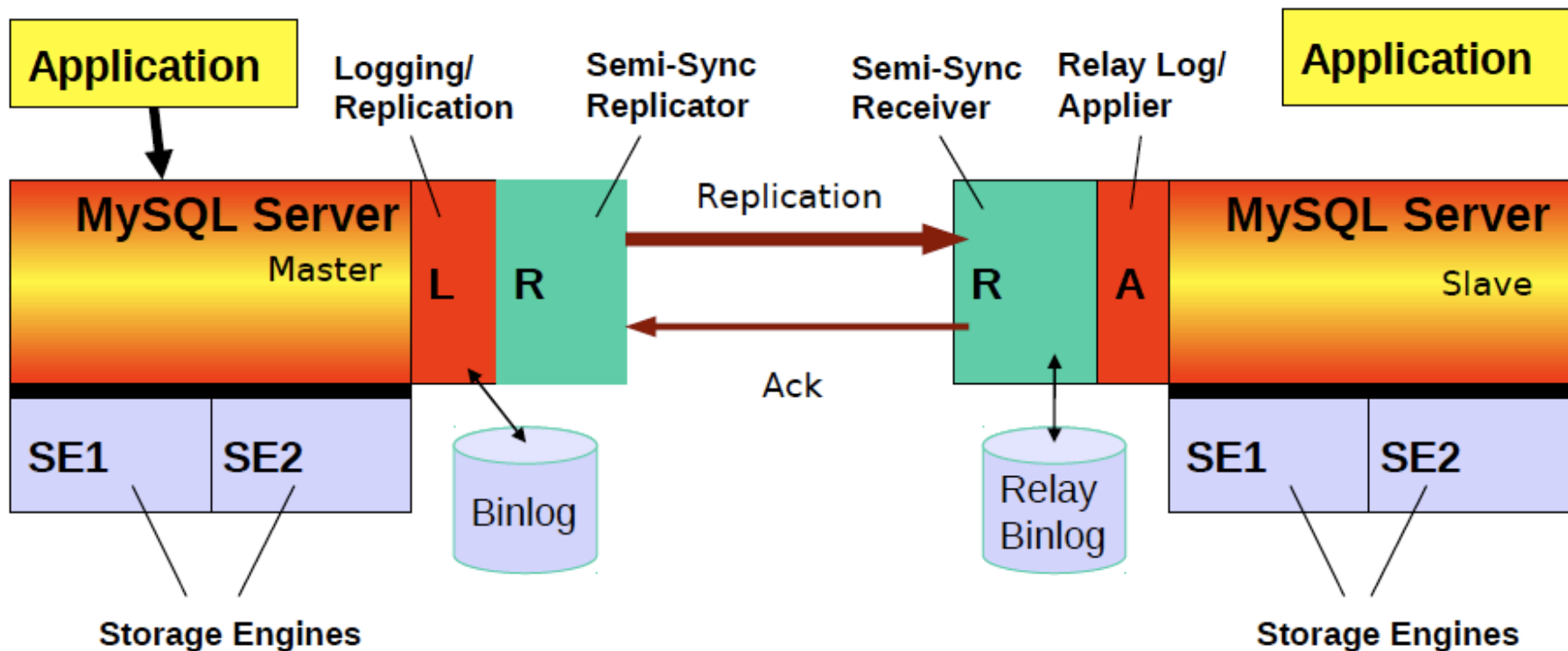


# Replication use cases

- Scale-Out
  - Partition “read” workload across slave servers
  - Compare to “Scale-Up”
- High Availability
  - Failover from master to stand-by slave
- Backup
  - Perform backups on slave to limit impact
- Analytics
  - Execute long-running or resource intensive queries on slaves
- Geographic Replication
  - Replicate data over large distances

# Semi-Sync Replication

Originally developed by Mark Callaghan and Wei Li, Google  
Modularized, tested, and bug fixed by Zhenxing He, MySQL





# Semi-Synchronous replication

- **To enable semi-synchronous replication on master**
  - `INSTALL PLUGIN 'rpl_semi_sync_master' SONAME 'semisync_master.so';`
  - `SET rpl_semi_sync_master_enabled=1;`
  - `SET rpl_semi_sync_master_timeout=1000; (1s, default 10s)`
- **To enable semi-synchronous replication on slave/slaves**
  - `INSTALL PLUGIN 'rpl_semi_sync_slave' SONAME 'semisync_slave.so';`
  - `SET rpl_semi_sync_slave_enabled=1;`
  - `START SLAVE;`

## On Master:

- `Rpl_semi_sync_master_status` - indicates status of when master is using asynchronous or semi-synchronous replication.
- `Rpl_semi_sync_master_clients` - shows how many slaves are configured for semi-synchronous replication.
- `Rpl_semi_sync_master_yes_tx` - shows number of successfully acknowledged commits by slaves.
- `Rpl_semi_sync_master_no_tx` - shows number of unsuccessfully acknowledged commits by slaves.

## On Slave:

- `Rpl_semi_sync_slave_status` - indicates if semi-synchronous replication is enabled on slave.





## Difference between these two types of Replication

- **MySQL replication is asynchronous**
  - A recovered master may have *non-binlogged changes*.  
This can cause master and slave to diverge.
  - A failed master may have *binlogged but non-replicated changes*  
If a slave is promoted to master some information can be lost.
- **Semisynchronous replication ensure redundancy**
  - At least one slave acknowledge relay logging the transaction.
  - Master waits for slave acknowledgement before commit returns.  
(On time-out, master temporarily switch to async replication.)



# Replication Heartbeat

"Heartbeat" is a message sent at regular intervals from a master node to the slave nodes.

You can configure the heartbeat period. If the message is not received, the slave knows that the master node has failed. You can now avoid the spurious relay log rotation when the master is idle, rely on a more precise failure detection mechanism, and have an accurate estimation for seconds behind master.

**STOP SLAVE;**

**CHANGE MASTER TO master\_heartbeat\_period=  
milliseconds;**

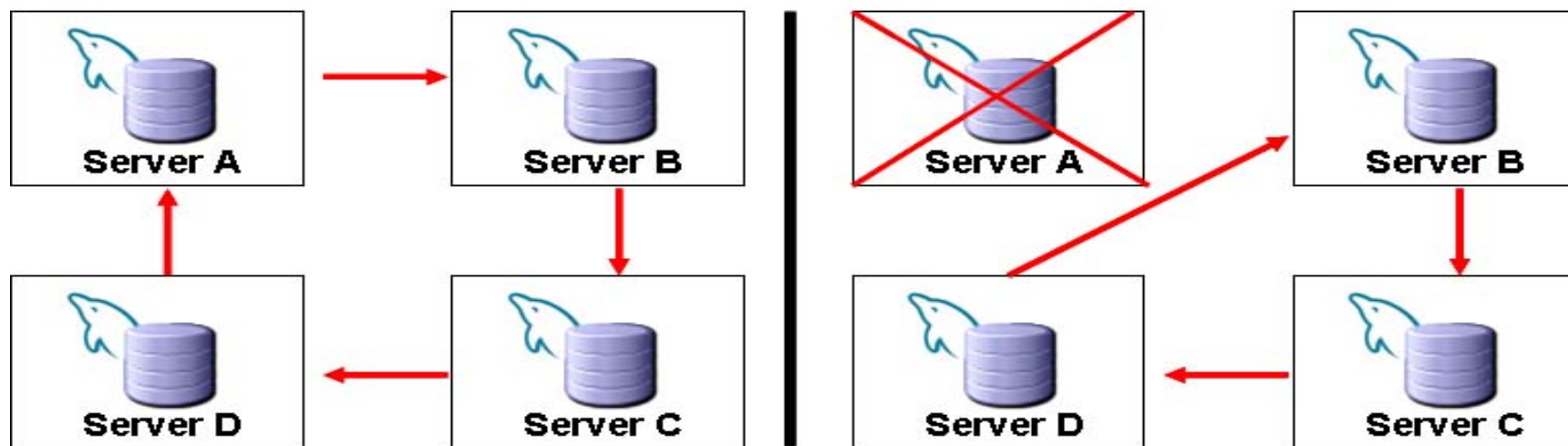
**START SLAVE;**

The following status variables can then be monitored to easily detect when a master is idle and to get a finer-grained estimate on slave seconds behind master for recovery purposes:

**SHOW STATUS like 'slave\_heartbeat period'**

**SHOW STATUS like 'slave\_received\_heartbeats'**

# Replication Server Filtering



When Server A is removed from the topology, users can now easily filter any Server A related events by entering the following command on the next server in the calling chain:

Server B> CHANGE MASTER TO MASTER\_HOST=D ...

IGNORE\_SERVER\_IDS=(A)



# Replication Slave Side Data Type Conversions

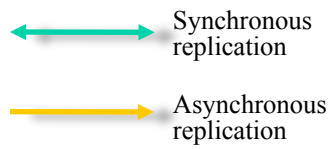
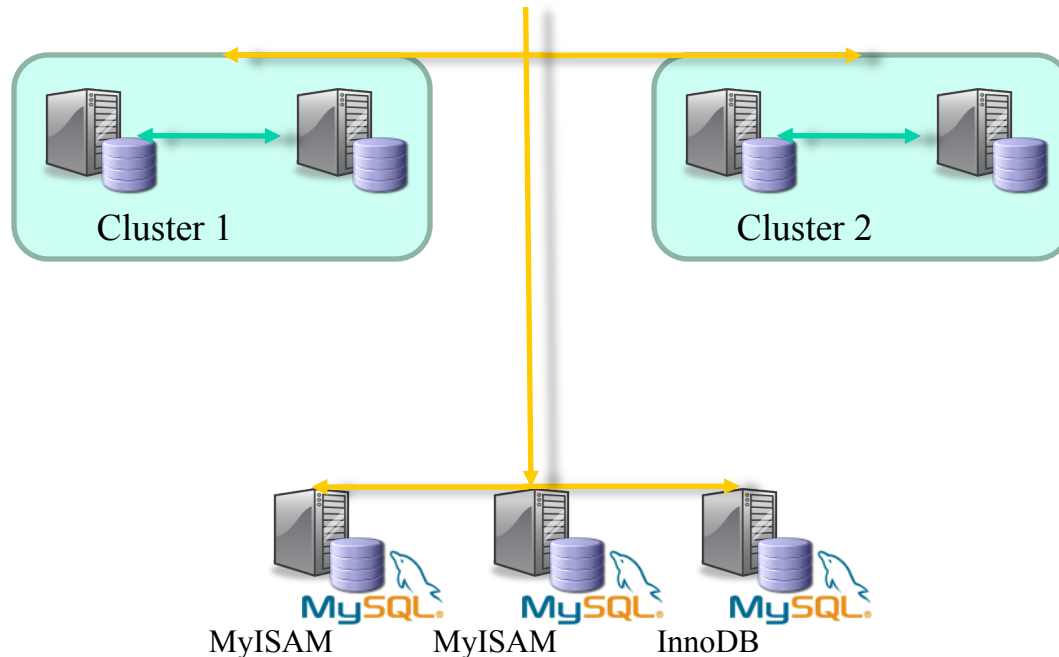
MySQL 5.5 now provides precise data type conversions between master and slave for both statement-based and row-based operations.

Conversions within integer, decimal, string, binary, BIT, ENUM and SET domains are supported.

- **SET SLAVE\_TYPE\_CONVERSIONS="ALL\_LOSSY"** - enables conversions to types with smaller domain (INT to TINY for example)

- **SET SLAVE\_TYPE\_CONVERSION="ALL\_NON\_LOSSY"** - enables conversions to types with larger domain (TINY to INT for example)

# Replication Flexibility



- Synchronous replication within a Cluster node group for HA
- Bi-Direction asynchronous replication to remote Cluster for geographic redundancy
- Asynchronous replication to non-Cluster databases for specialised activities such as report generation
- Mix and match replication types

# MySQL In Action on the Web



"In my opinion, MySQL is the only database we would ever trust to power the Zappos.com website."



"As a leader in our field, we are committed to providing the best service to our users, and a web experience that meets members expectations and that starts with IT"

The Ticketmaster logo, featuring the word "ticketmaster" in a blue, lowercase, sans-serif font, with a yellow underline beneath the word.

"On any given day we can sell close to 300,000 tickets on the Web site using MySQL as the database to search for events. It is amazing."

The Facebook logo, featuring the word "facebook" in a white, lowercase, sans-serif font, centered within a blue rectangular background.

"We are one of the largest MySQL web sites in production"

The Craigslist logo, featuring the word "craigslist" in a blue, lowercase, sans-serif font.

"craigslist infrastructure could not have handled the exponential growth in traffic without MySQL."



"They have a master server for all writes and slave servers for most Reads. The secret truth they claim behind configuring the master and slave machines is to make sure the slave machines are faster than the masters"

ORACLE



Booking.com is Europe's largest online hotel travel reservations agency attracting over 30 million unique visitors each month.

"Booking.com has been growing significantly every year. That is why we designed a database architecture that we believe will scale up to ten times over our current requirements. MySQL's open source structure offers us opportunities for growth and integration because data replication can be introduced and managed from an early stage."

Herald van der Breggen, Senior Developer

Booking.com

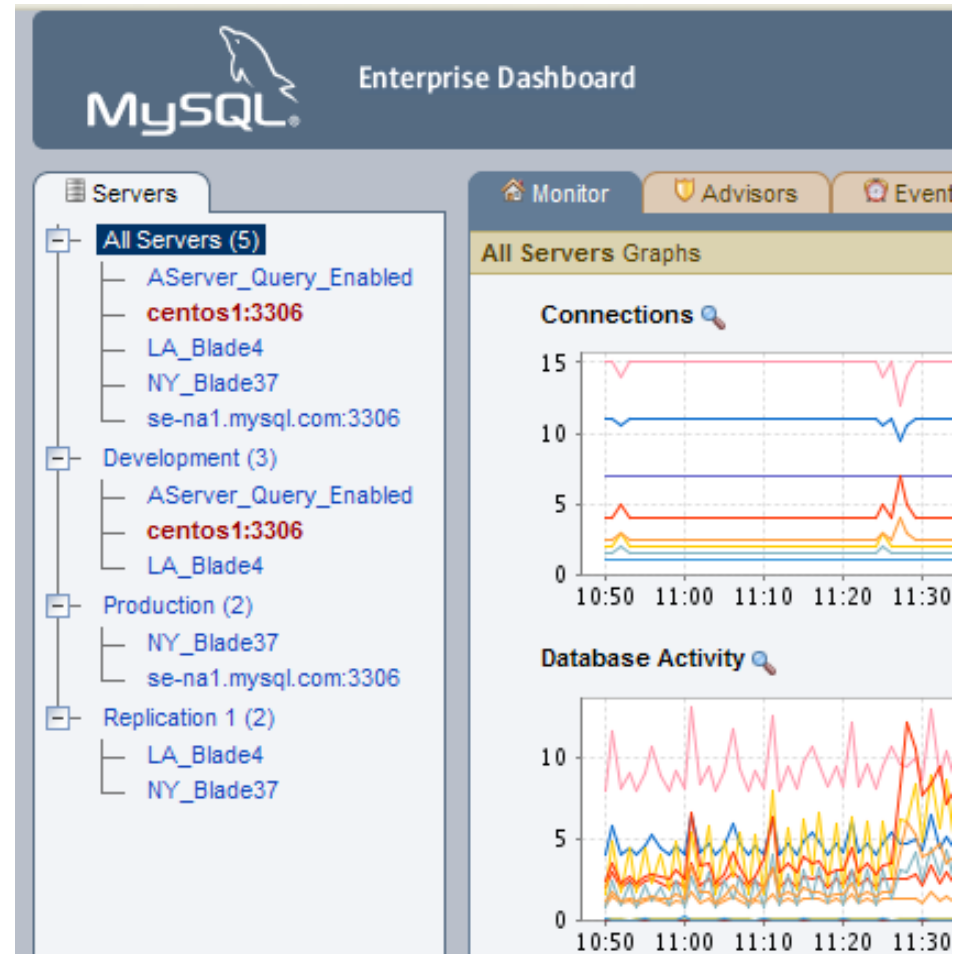


# Enterprise Tools



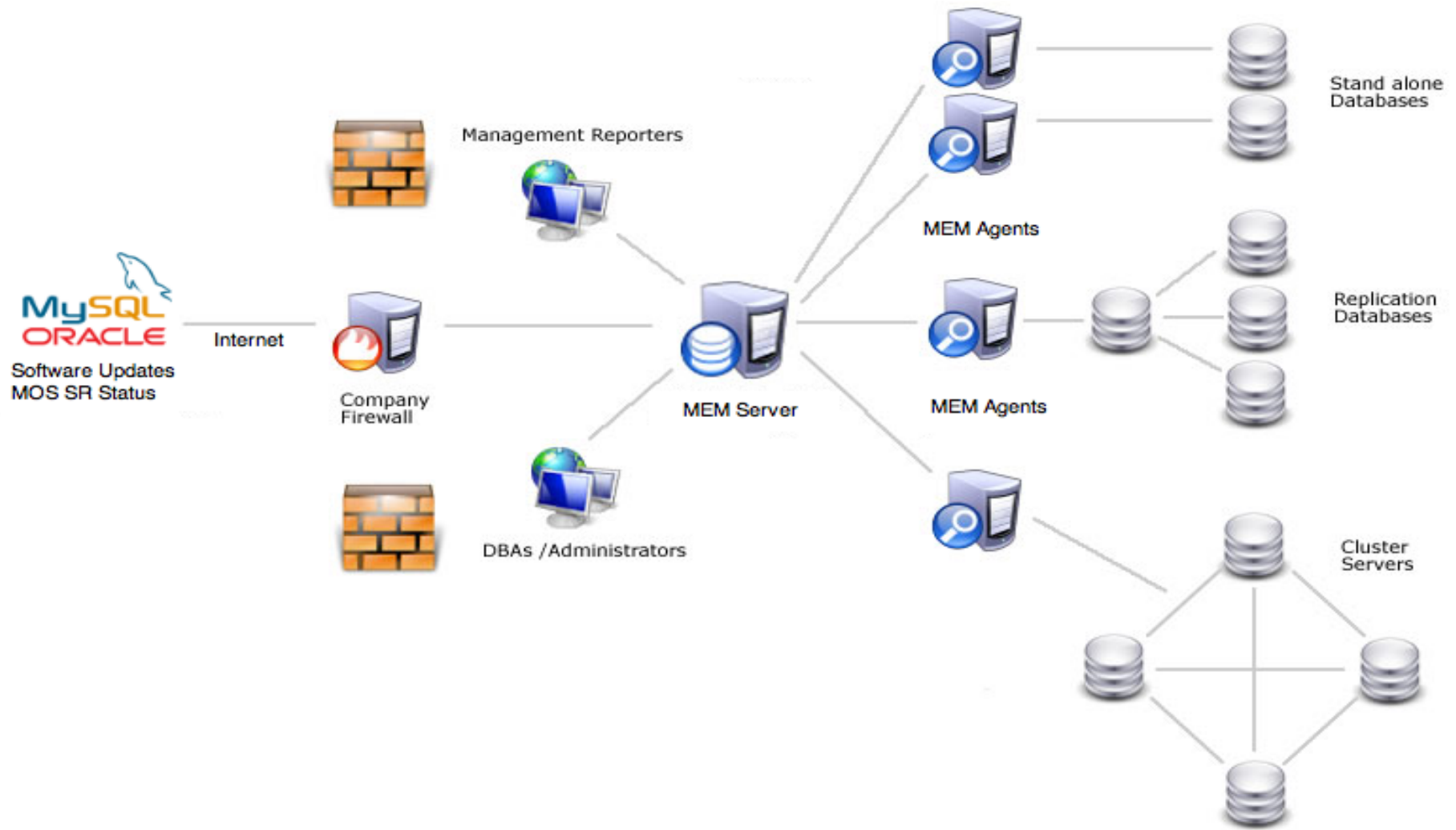
# MySQL Enterprise Monitor

- Single, consolidated view into entire MySQL application development environment
- Auto-discovery of MySQL servers, replication topologies
- Automated, customizable rules-based monitoring, tuning, SNMP/SMTP alerts
- Query Analyzer for query monitoring, analysis, tuning, source code tracing
- Application Tuning during Dev/QA/Roll out
- Reduces risk of problems after apps are deployed



*A Virtual MySQL Tuning Assistant!*

# MySQL Enterprise Monitor Deployment



100

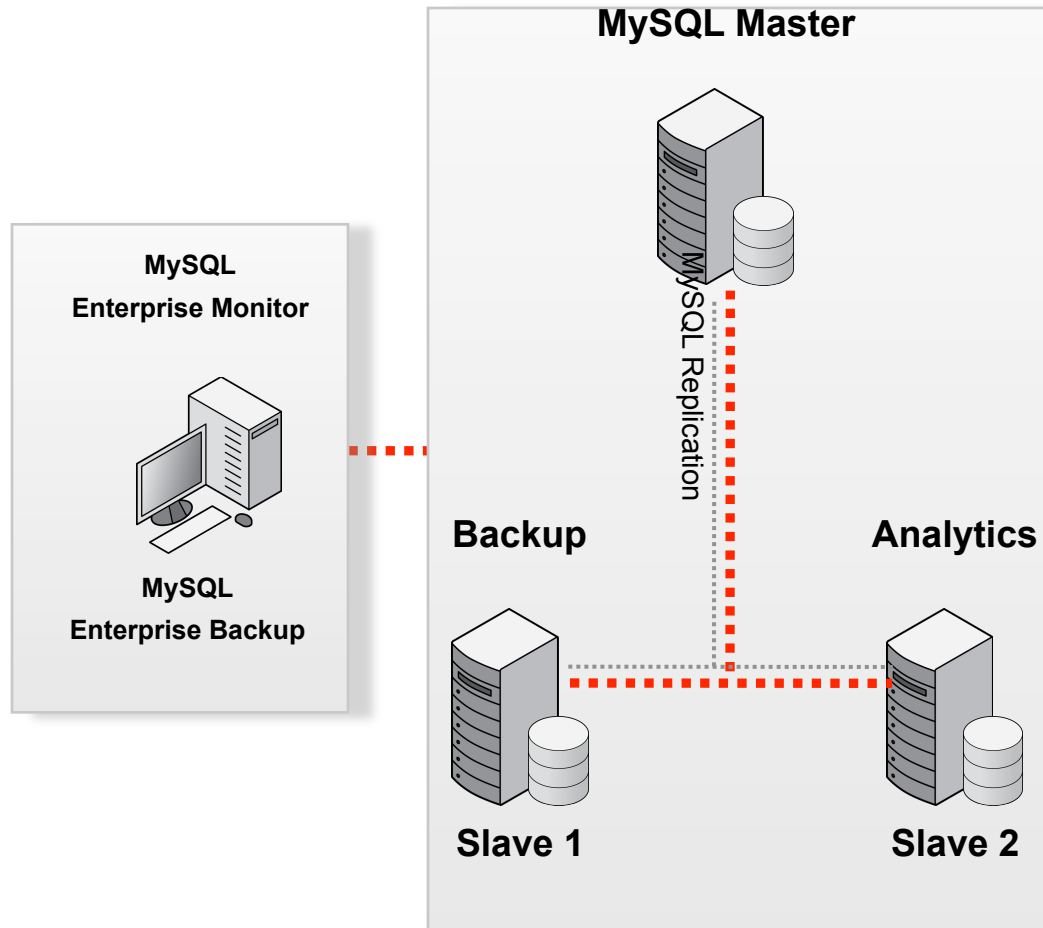
ORACLE®



# MySQL Enterprise Backup

- Formerly “InnoDB Hot Backup”
- InnoDB Hot Backup is rebranded as MySQL Enterprise Backup
- Online, non-locking backup & recovery
  - Tables, Indexes
  - Server, database
- Incremental backup
- Point-in-time recovery
- Compressed backups
- Also provides backup & recovery for MyISAM
- Cross-Platform (Windows, Linux, Unix)

# Small: Web Reference Architecture



- Single server supporting all workloads
- Data replicated to slaves for back-up & analysis

## Applications

- Members/Authentication
- eCommerce
- Content Management
- Search

**Only deploy when future traffic growth is very limited**

# MySQL 5.5 – Highest Quality Release Ever



## InnoDB becomes default storage engine

- ACID Transactions, FKs, Crash Recovery

## Improved Performance

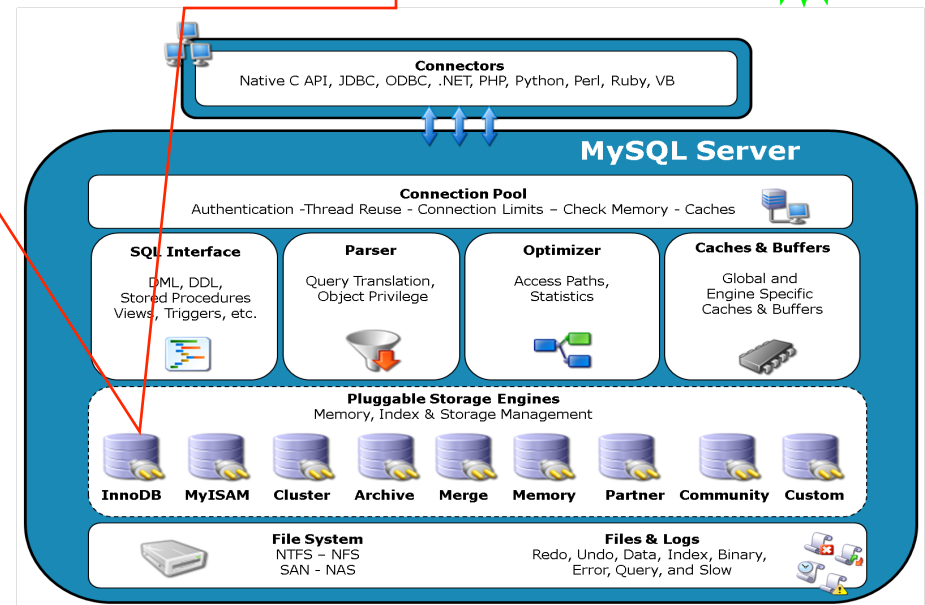
- Enhancements in MySQL DB
- Enhancements in InnoDB
- + 360% over 5.1 on Linux
- + 1500% over 5.1 on Windows

## Improved Availability

- Semi-synchronous Replication
- Replication Heartbeat

## Improved Usability

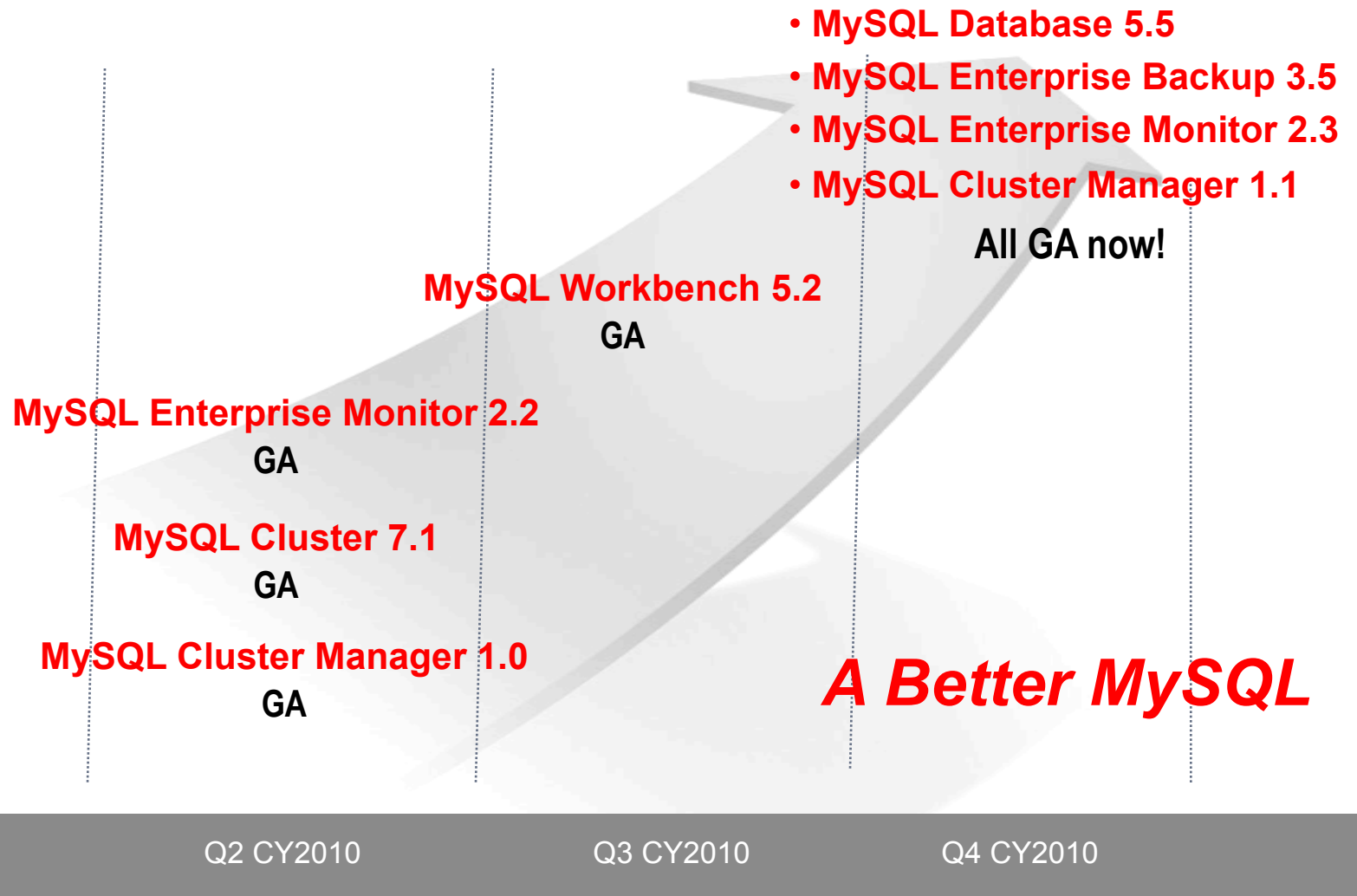
- SIGNAL/RESIGNAL
- More Partitioning Options
- New PERFORMANCE\_SCHEMA





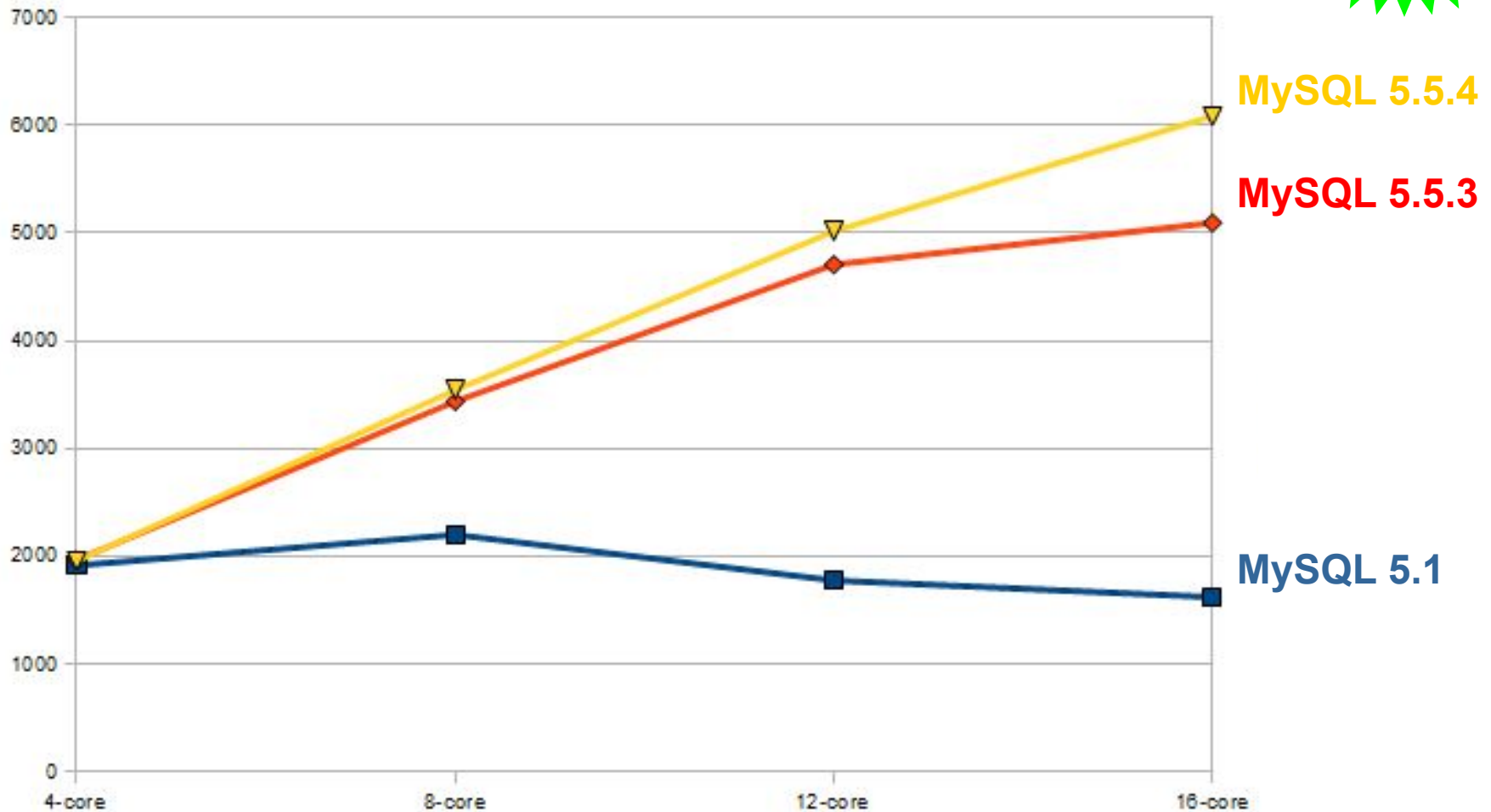
# MySQL Product Releases

## Continuous Innovation



# MySQL 5.5 Scales on multi core SysBench Read Write

GA



AMD Opteron 7160 (Magny-Cours) @2100 MHz  
64 GB memory  
2 x Intel X25E SSD drives  
OS is Oracle Enterprise Linux with the Enterprise Kernel  
4 sockets with a total of 48 cores.





The following is intended to outline our general product direction.  
It is intended for information purposes only, and may not be incorporated into any contract.  
It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.

The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.



# MySQL 5.6 – A Better MySQL.



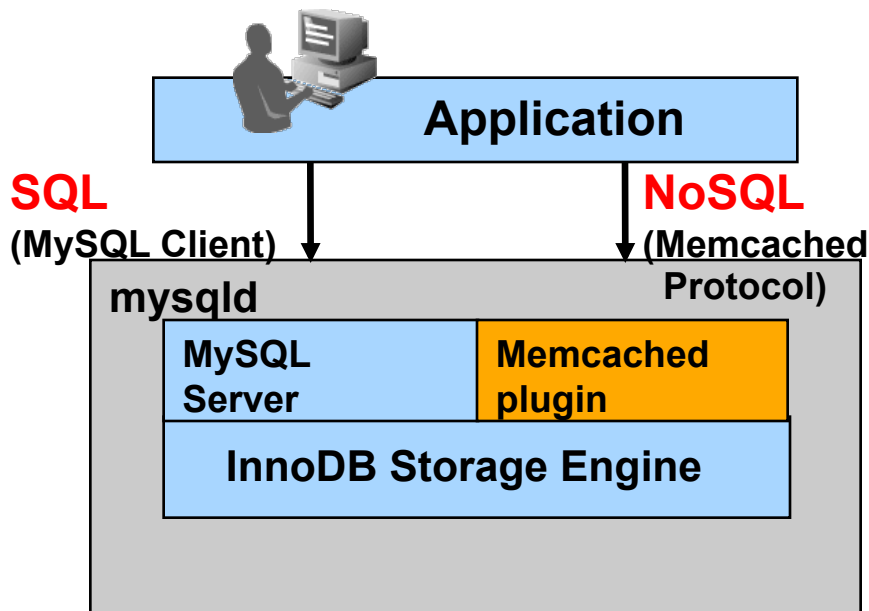
## Better Replication

- Crash-Safe Slaves
- Multi-threaded Slaves
- Replication Checksums
- Time-Delayed Replication
- Remote Binlog Backups
- Server UUIDs

[dev.mysql.com/downloads/mysql](http://dev.mysql.com/downloads/mysql)

# MySQL 5.6:

## NotOnlySQL: Memcached API



- Fast, simple access to InnoDB
  - Accessed via Memcached API
  - Use existing Memcached clients
  - Bypasses SQL transformations
- NotOnlySQL access
  - Memcached for key-value operations
  - SQL for rich queries, JOINS, foreign keys, etc.
- Implementation
  - Memcached daemon plug-in to mysqld
  - Memcached protocol mapped to the native InnoDB API
  - Shared process space for ultra-low latency
  - Additional implementations in future DMs



# MySQL Services from Oracle

- MySQL Support
  - Global, 24 x 7 support coverage
- MySQL Consulting
  - Architecture and Design
  - Performance Tuning
  - High Availability
  - Migration
  - Remote DBAs
- MySQL Training
  - DBAs & Developers of all levels
  - Database and applications
  - Developing Dynamic Web Applications





# Resources

## MySQL Replication Whitepaper

<http://www.mysql.com/why-mysql/white-papers/mysql-wp-replication.php>

## MySQL Newsletter - Special Edition: Scaling with MySQL

<http://www.mysql.com/news-and-events/newsletter/2010/2010-09sp.html>

## Customer's Scale-out success stories

<http://www.mysql.com/why-mysql/scaleout>

## Read MySQL 5.5 Replication Docs

<http://dev.mysql.com/doc/refman/5.5/en/replication.html>

## MySQL Enterprise Whitepaper

[http://www.mysql.com/why-mysql/white-papers/mysql\\_wp\\_enterprise\\_ready.php](http://www.mysql.com/why-mysql/white-papers/mysql_wp_enterprise_ready.php)



# Key Takeaways

- MySQL is important to Oracle and our customers
- 5.5 is GA and scales upto 32 cores. Download and test it
- Use MySQL replication for scale out
- Use semi-sync for better data integrity
- Use our Enterprise tools to monitor and backup MySQL Dbs.
- Need more help??
  - Purchase MySQL Enterprise
  - MySQL Consulting

