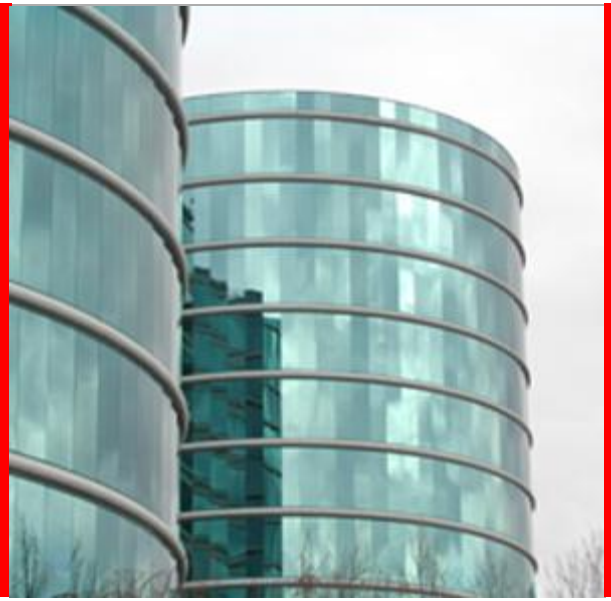


ORACLE®

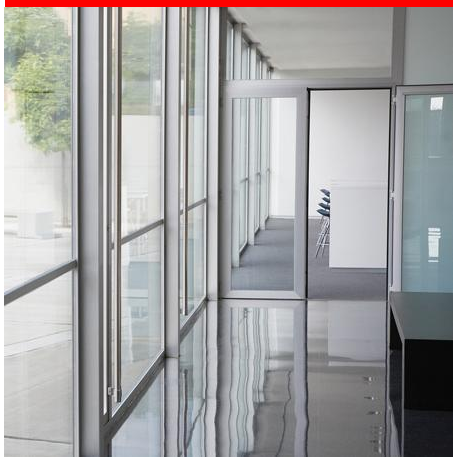
MySQL™

MySQL Web Reference Architectures

Learning from the Giants of the Web

Luca Olivari
MySQL Sales Consulting Manager, EMEA

Session Agenda



- Requirements for Innovating on the Web
- Reference Architectures
  - Sizing & Topologies
  - Enabling Technology
  - Best Practices
- Value Added Components
- Services
- Resources to Get Started

MySQL:

The World's Leading On-Line Database

Powering 9 of the top 10 most trafficked sites on the web*

* http://www.alexa.com/topsites

Innovating on the Web

Translating Business Requirements to Technology Implementation

- Top Challenges
  - Fast Time to Market
  - High availability
  - High performance
  - Cost-effectively scale to meet rapidly growing demands
  - Open, customizable and repeatable
  - Protection of customer data
  - Low total cost of ownership
- Leverage “Best Practices”
  - Accelerating Time to Market with greater agility
  - Reducing cost, risk and complexity
  - Eliminate expensive “trial & error”

MySQL Web Reference Architectures

- Repeatable best practices developed with leading web properties
  - Recommended architectures & topologies
- 4 x Reference Architectures
  - Small
  - Medium
  - Large
  - Extra Large (Social Networking)
- 4 x common platform components
  - User Authentication & Session Management
  - Content Management
  - eCommerce
  - Analytics

Reference Architecture Sizing

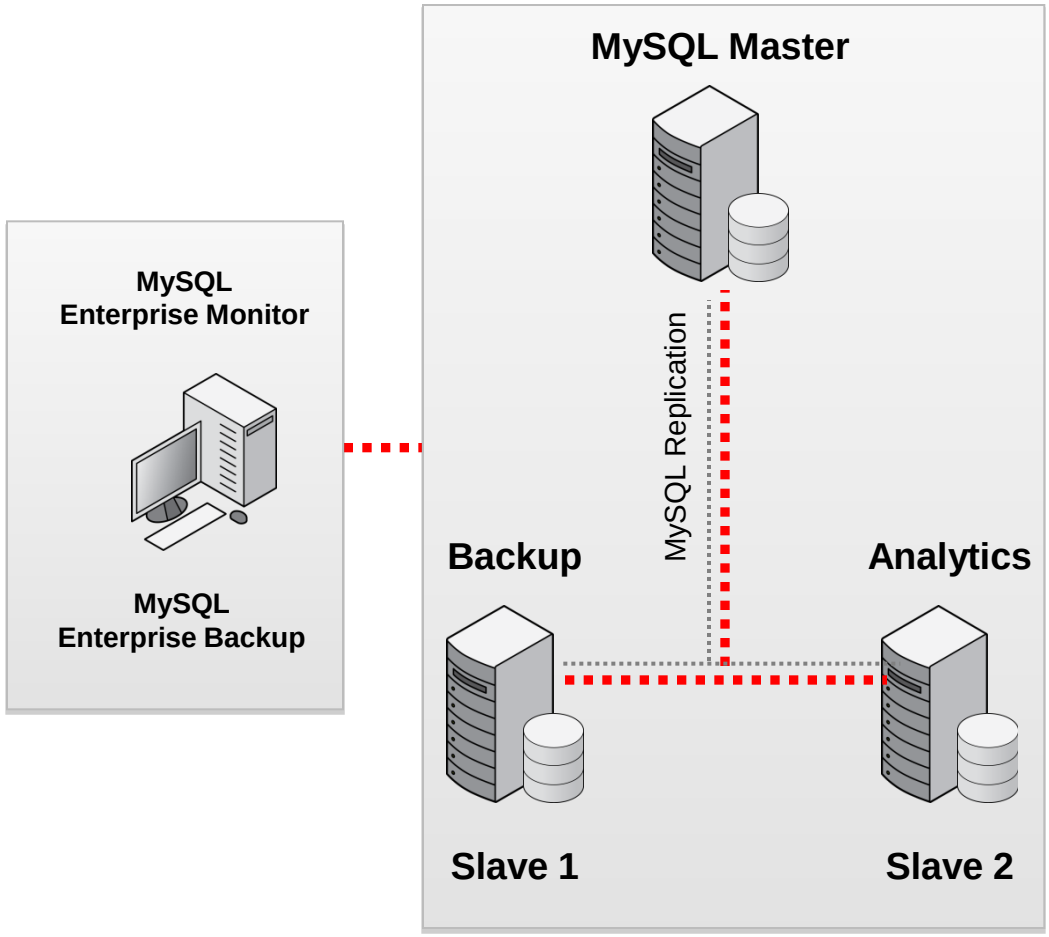
				Social Network
	Small	**Medium**	**Large**	**Extra Large**
Queries/Second	**<500**	**<5,000**	**10,000+**	**25,000+**
Transactions/Second	**<100**	**<1,000**	**10,000+**	**25,000+**
Concurrent Read Users	**<100**	**<5,000**	**10,000+**	**25,000+**
Concurrent Write Users	**<10**	**<100**	**1,000+**	**2,500+**
Database Size				
Sessions	**<2 GB**	**<10 GB**	**20+ GB**	**40+ GB**
eCommerce	**<2 GB**	**<10 GB**	**20+ GB**	**40+ GB**
Analytics	**<10 GB**	**<500 GB**	**1+ TB**	**2+ TB**
Content Management	**<10 GB**	**<500 GB**	**1+ TB**	**2+ TB**

Reference Architectures

Small: Web Reference Architecture

				Social Network
	Small	Medium	Large	Extra Large
Queries/Second	**<500**	<5,000	10,000+	25,000+
Transactions/Second	**<100**	<1,000	10,000+	25,000+
Concurrent Read Users	**<100**	<5,000	10,000+	25,000+
Concurrent Write Users	**<10**	<100	1,000+	2,500+
Database Size				
Sessions	**<2 GB**	<10 GB	20+ GB	40+ GB
eCommerce	**<2 GB**	<10 GB	20+ GB	40+ GB
Analytics	**<10 GB**	<500 GB	1+ TB	2+ TB
Content Management	**<10 GB**	<500 GB	1+ TB	2+ TB

Small: Web Reference Architecture



- Single server supporting all workloads

  Members/Authentication
  eCommerce
  Content Management
  Search

- Data replicated to slaves for back-up & analysis

Only deploy when future traffic growth is very limited

Best Practices

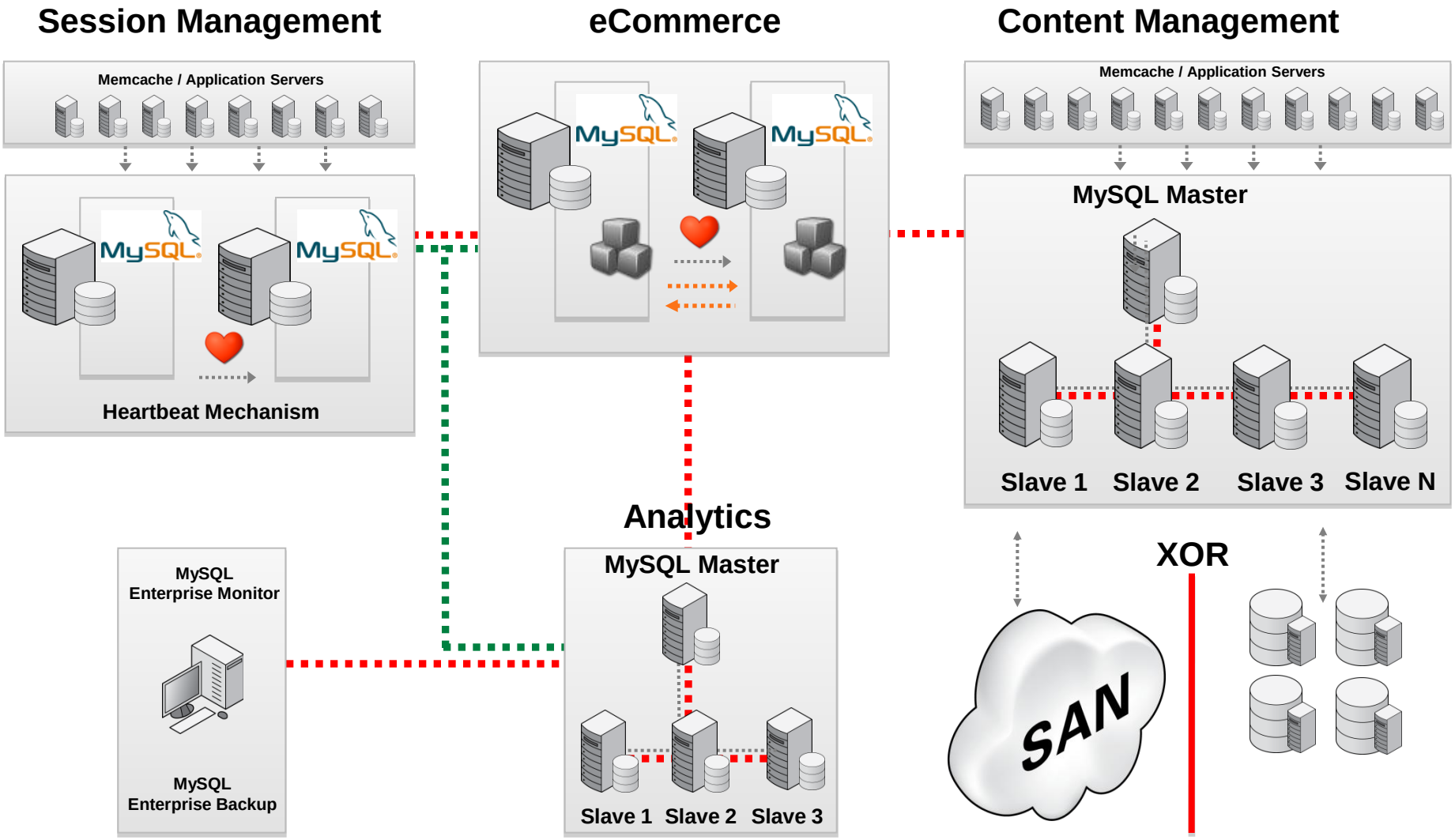
Small Web Reference Architecture

- Use default InnoDB storage engine for all workloads
  - Default MySQL storage engine
  - ACID Compliant, Transactional
  - MVCC & Row-Level Locking
  - Foreign Keys & constraints
- If traffic volumes increase, scale session management first
  - Migrate Session Management to a dedicated MySQL server
- If future scalability is required, start with the Medium Reference Architecture
  - Complex to tune multiple applications on shared hardware

Medium: Web Reference Architecture

				Social Network
	Small	**Medium**	Large	Extra Large
Queries/Second	<500	**<5,000**	10,000+	25,000+
Transactions/Second	<100	**<1,000**	10,000+	25,000+
Concurrent Read Users	<100	**<5,000**	10,000+	25,000+
Concurrent Write Users	<10	**<100**	1,000+	2,500+
Database Size				
Sessions	<2 GB	**<10 GB**	20+ GB	40+ GB
eCommerce	<2 GB	**<10 GB**	20+ GB	40+ GB
Analytics	<10 GB	**<500 GB**	1+ TB	2+ TB
Content Management	<10 GB	**<500 GB**	1+ TB	2+ TB

Medium: Web Reference Architecture



Best Practices (1)

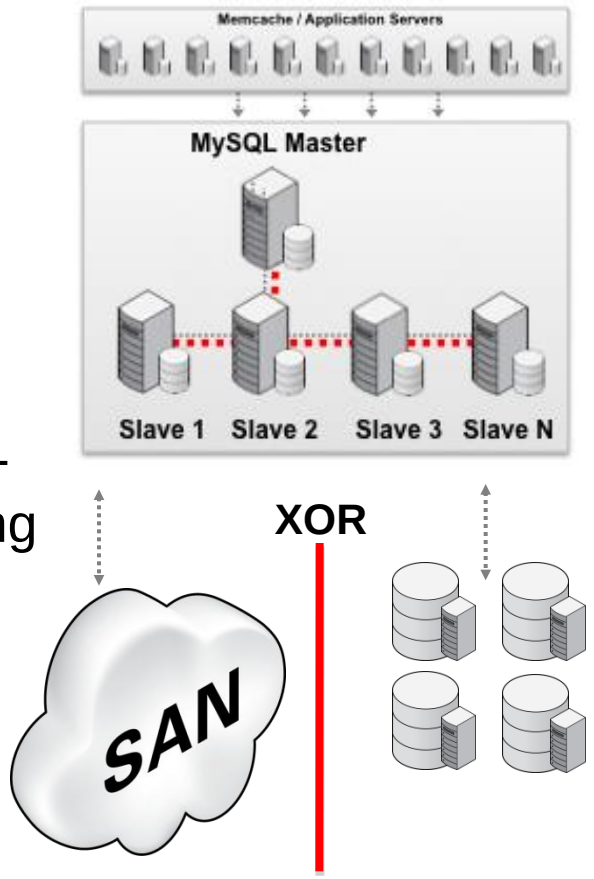
Medium Web Reference Architecture

- Each component deployed onto dedicated server & storage infrastructure
  - Deployed, managed and scaled independently
- Server ratio: 8 application servers to each MySQL Server
  - More for PHP applications, less for Java
  - Add more slaves as the application tier scales
- Memcached deployed in session & content management components
  - Distributed memory caching layer
  - Reads fulfilled from cache, relieving load on the source database servers

Best Practices (2)

Medium Web Reference Architecture

- Content Management
  - Each slave can handle around 3,000 concurrent users
  - Each master can handle up to 30 slaves
  - MySQL Replication for high availability
    - Can include Heartbeat, depending on application failover requirements
  - Meta data of content assets managed by MySQL
  - Distributed File System (i.e. MogileFS for indexing content assets)
  - Physical storage for content assets
    - High quality SAN (redundancy for HA)
    - Distributed across local storage with Semi-Synchronous replication of indexing and meta data



Best Practices (3)

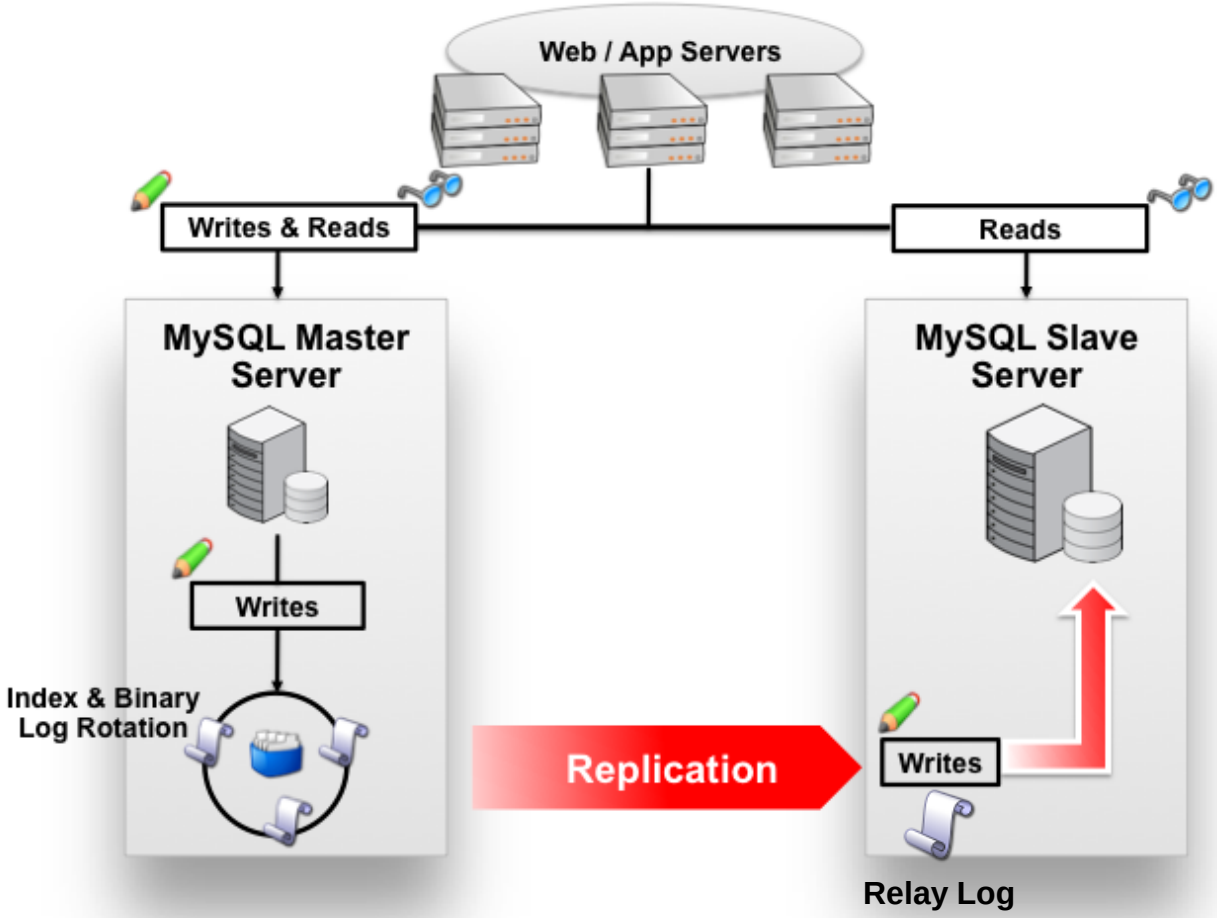
Medium Web Reference Architecture

- Session Management & eCommerce
  - Deployed onto InnoDB storage engine
  - Session data maintained for up to 1 hour in a dedicated partition, rolling partitions used to delete aged data
  - Data is captured in Analytics Datababase
  - MySQL Replication with Heartbeat for HA
    - Configure semi-synchronous replication or OS-level DRBD for eCommerce
  - If web traffic grows, move Session Management to MySQL Cluster
    - Persist session data for real-time personalization of user experience
    - 99.999% availability and in-memory data management can reduce need for DRBD & memcached

High Availability & Scalability

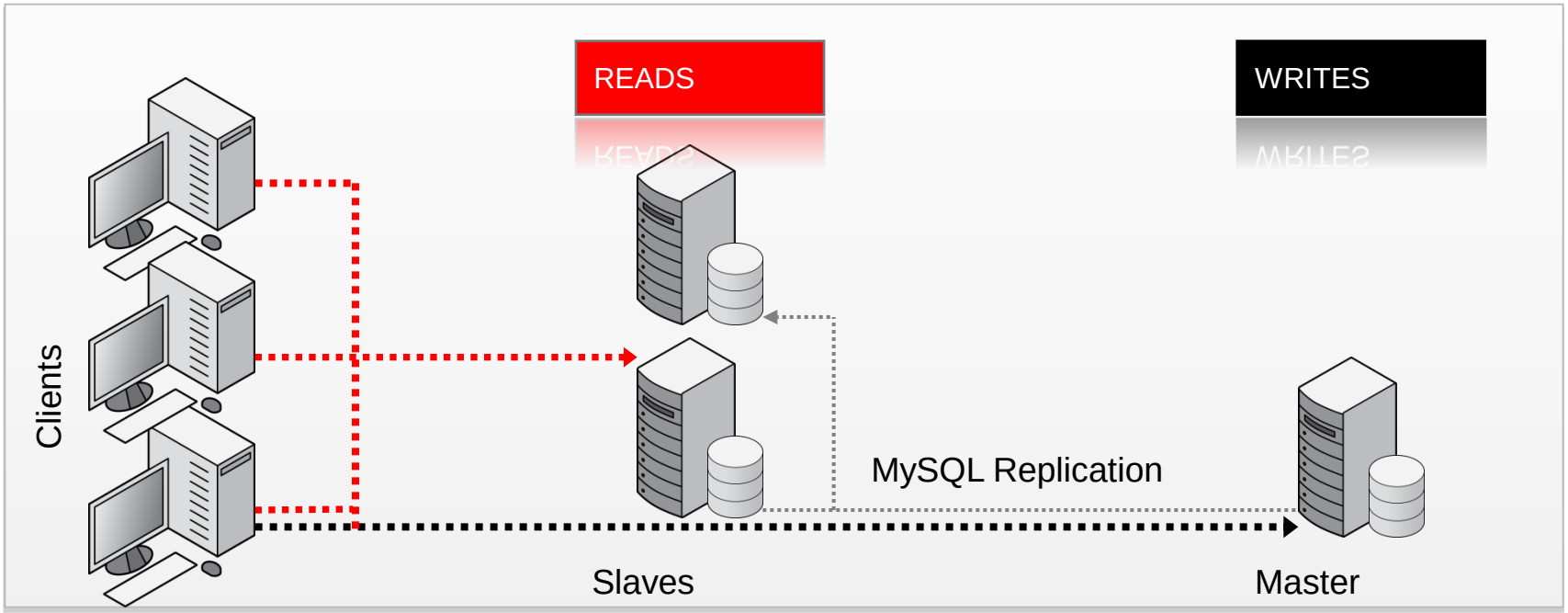
MySQL Replication

- 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



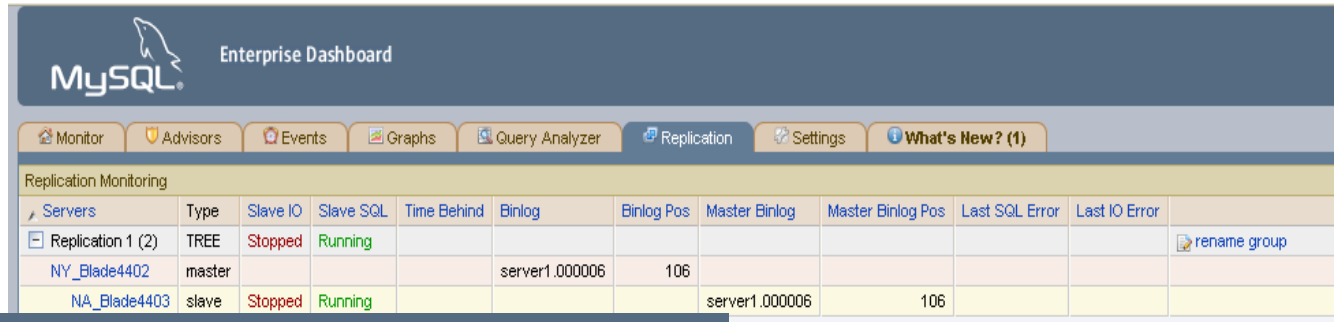
MySQL Replication

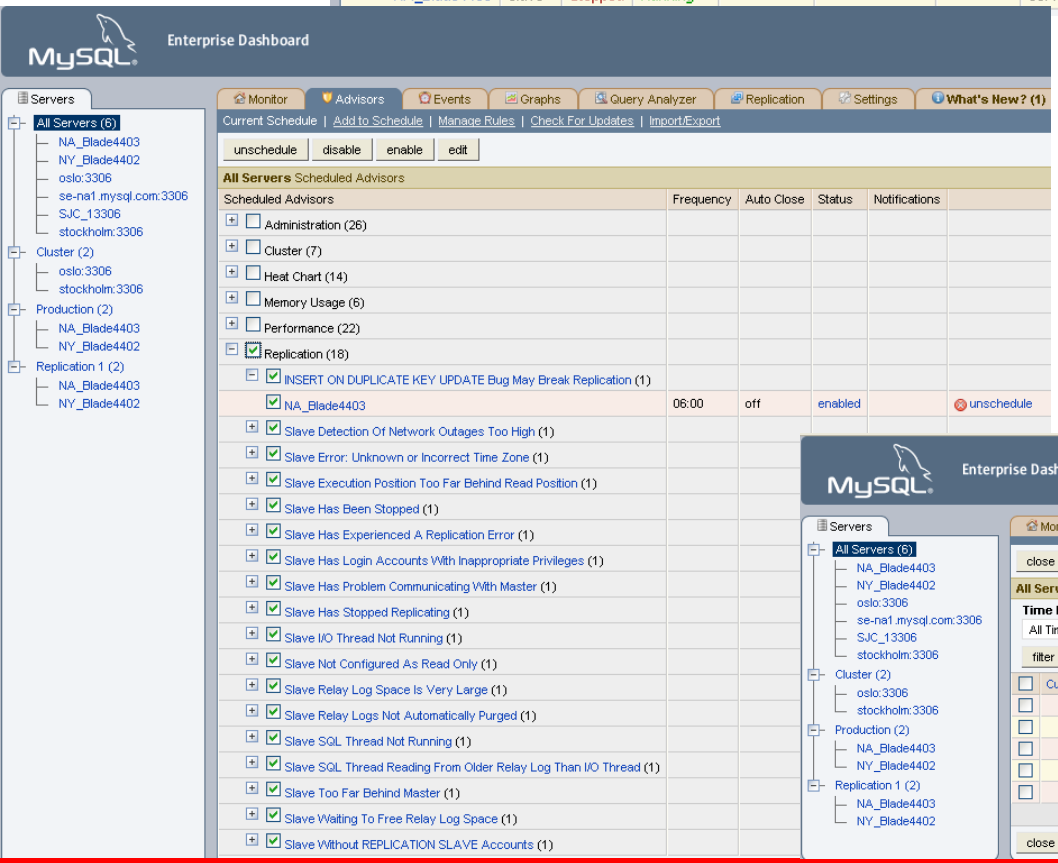
Read Scalability

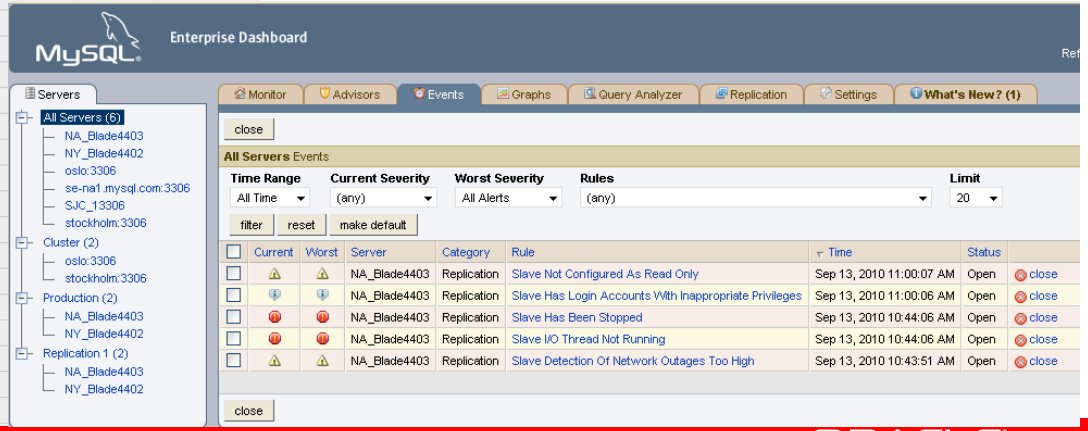


- Used by leading web properties for scale-out
- Reads are directed to slaves, writes to master
- Delivers higher performance & scale with efficient resource utilization

Monitoring Replication

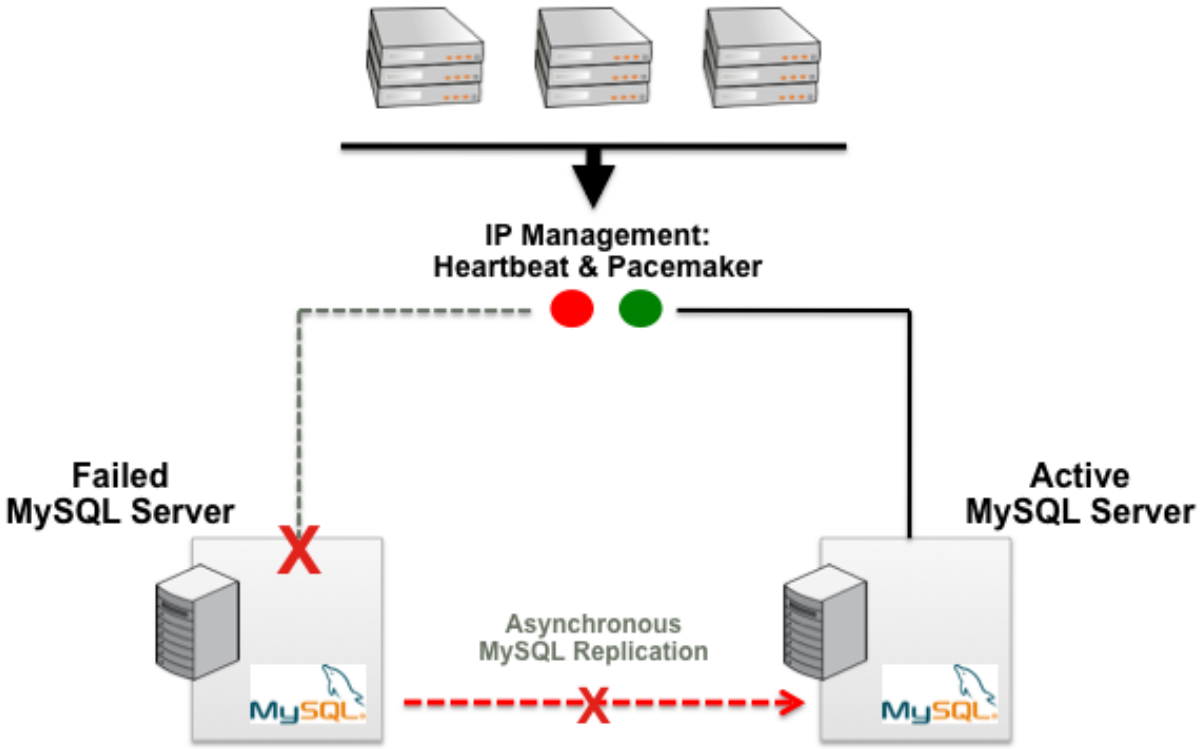






Building on Replication

Failure Detection & Failover

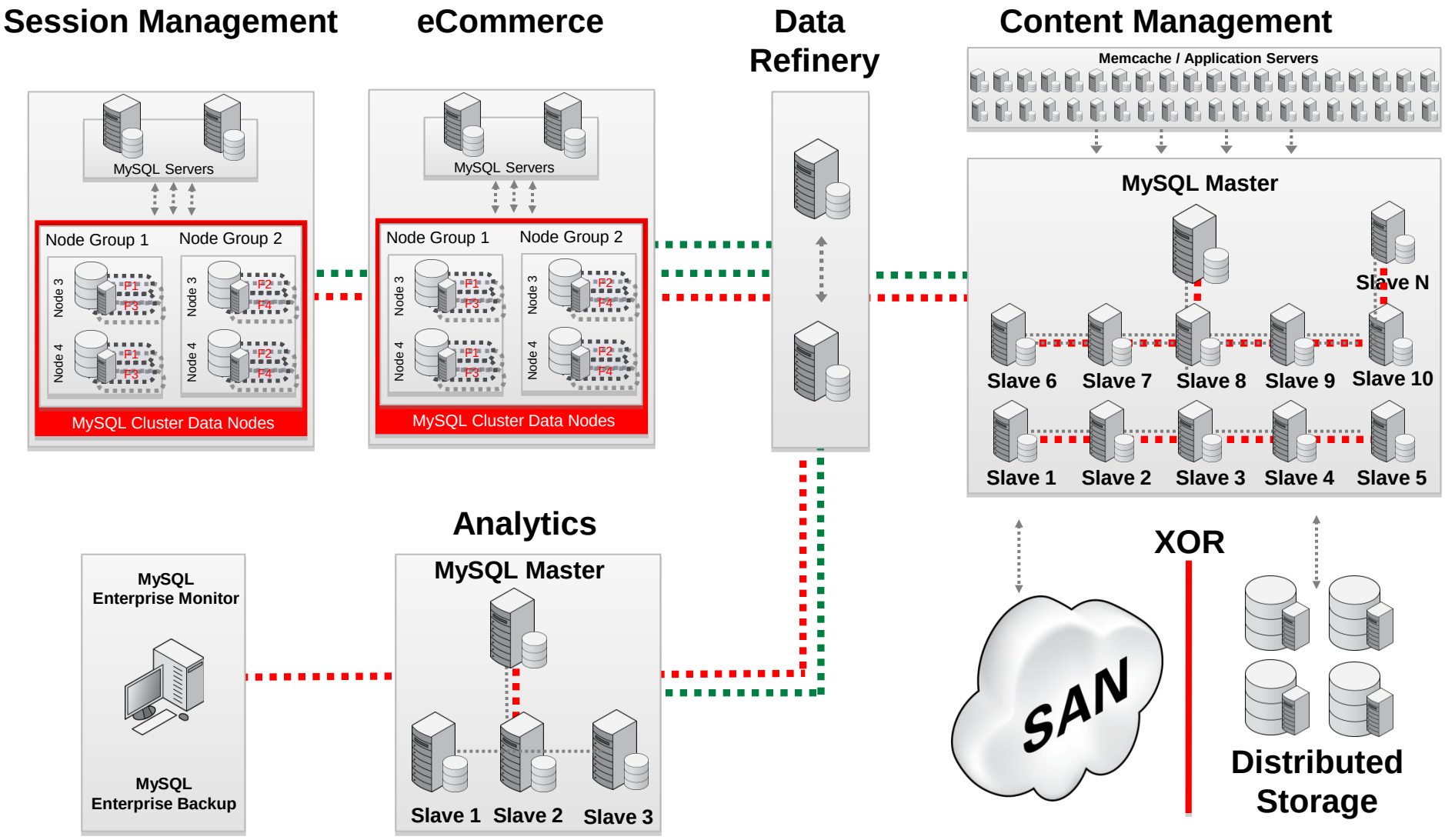


- Linux Heartbeat implements heartbeat protocol between nodes
- Failover initiated by Cluster Resource Manager (Pacemaker) if heartbeat message is not received
- Virtual IP address failed over to ensure failover is transparent to apps

Large: Web Reference Architecture

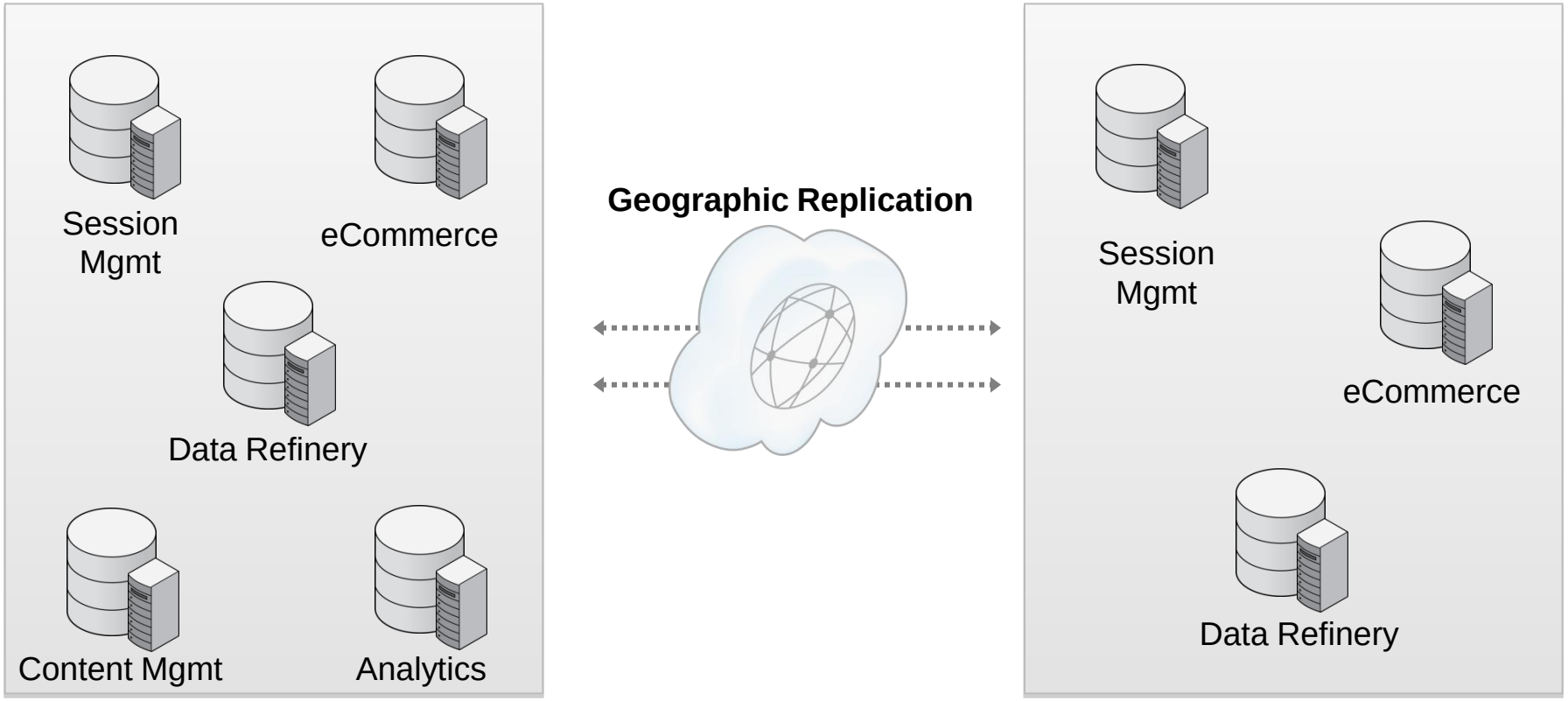
				Social Network
	Small	Medium	**Large**	Extra Large
Queries/Second	<500	<5,000	**10,000+**	25,000+
Transactions/Second	<100	<1,000	**10,000+**	25,000+
Concurrent Read Users	<100	<5,000	**10,000+**	25,000+
Concurrent Write Users	<10	<100	**1,000+**	2,500+
Database Size				
Sessions	<2 GB	<10 GB	**20+ GB**	40+ GB
eCommerce	<2 GB	<10 GB	**20+ GB**	40+ GB
Analytics	<10 GB	<500 GB	**1+ TB**	2+ TB
Content Management	<10 GB	<500 GB	**1+ TB**	2+ TB

Large: Web Reference Architecture



Large: Web Reference Architecture

Conceptual View



East Coast Data Center

West Coast Data Center

Best Practices

Large Web Reference Architecture

- Builds on best practices of Medium Web Ref Arch
  - Dedicated infrastructure for each workload, MySQL Replication, Memcached, etc.
- Introduces Data Refinery
  - Aggregate data across the web components
  - Data cleansing
  - Builds Data Warehouse Dimensions
  - Supports higher volume content management and analytics
- Introduces MySQL Cluster
  - Session Management and eCommerce
  - 4 x Data Nodes support 6k page hits per second
  - Each page hit generating 8 – 12 database operations

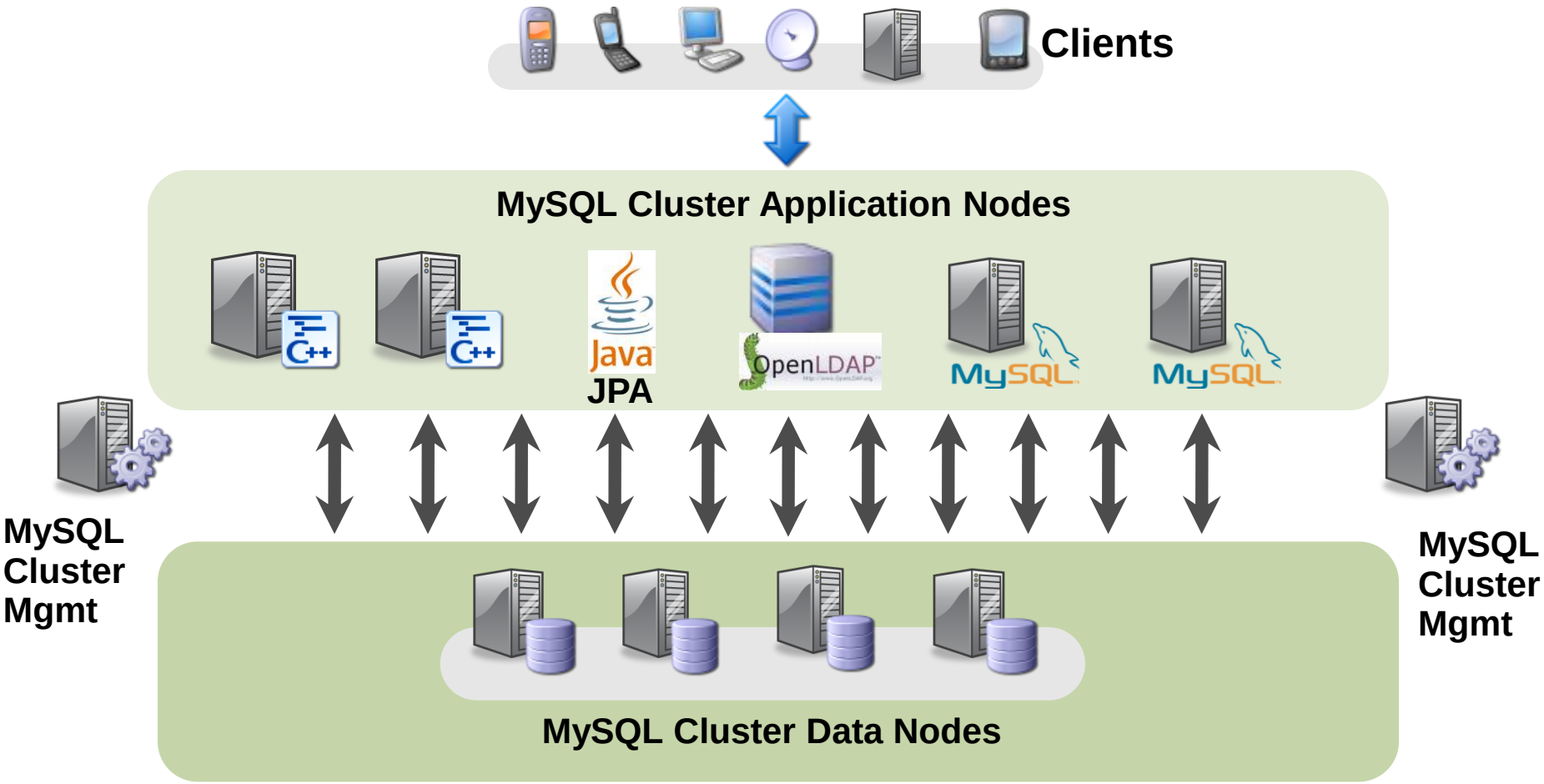
MySQL Cluster Overview

Write-Scalable, Real-Time, 99.999%s Uptime

- distributed hash table backed by an ACID relational model
- shared-nothing architecture, scale out on commodity hardware
- implemented as a pluggable storage engine for the MySQL Server, with additional direct access via embedded APIs.
- automatic or user configurable data partitioning across nodes
- synchronous data redundancy
- sub-second fail-over & self-healing recovery
- geographic replication
- data stored in main-memory or on disk (configurable per-column)
- logging and check pointing of in-memory data to disk
- online operations (i.e. add-nodes, schema updates, maintenance, etc)

MySQL Cluster Architecture

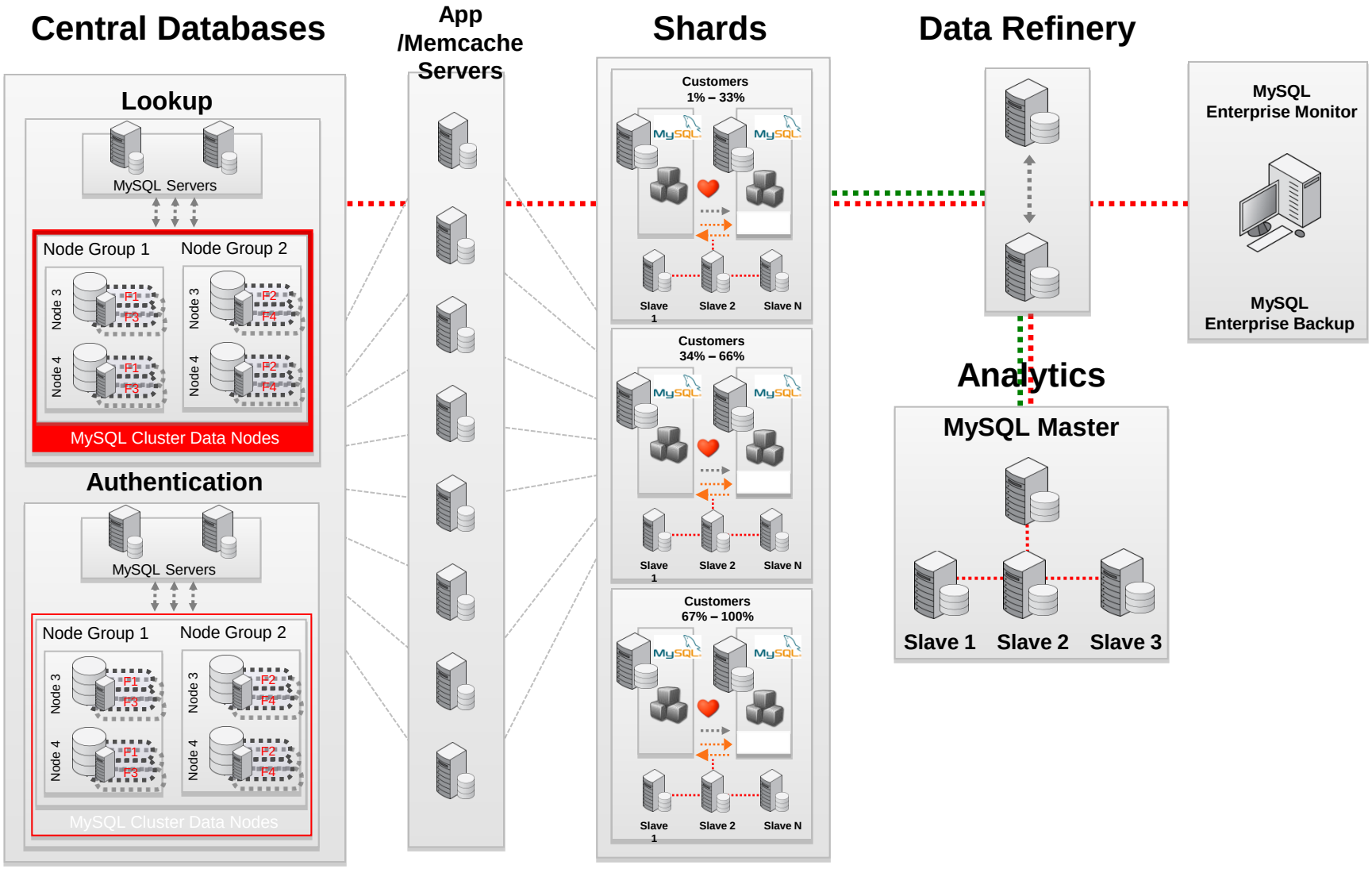
Parallel Database with no SPOF: High Read & Write Performance & 99.999% uptime



Extra Large: Social Network Reference Architecture

				Social Network
	Small	Medium	Large	**Extra Large**
Queries/Second	<500	<5,000	10,000+	**25,000+**
Transactions/Second	<100	<1,000	10,000+	**25,000+**
Concurrent Read Users	<100	<5,000	10,000+	**25,000+**
Concurrent Write Users	<10	<100	1,000+	**2,500+**
Database Size				
Sessions	<2 GB	<10 GB	20+ GB	**40+ GB**
eCommerce	<2 GB	<10 GB	20+ GB	**40+ GB**
Analytics	<10 GB	<500 GB	1+ TB	**2+ TB**
Content Management	<10 GB	<500 GB	1+ TB	**2+ TB**

Extra Large: Social Network

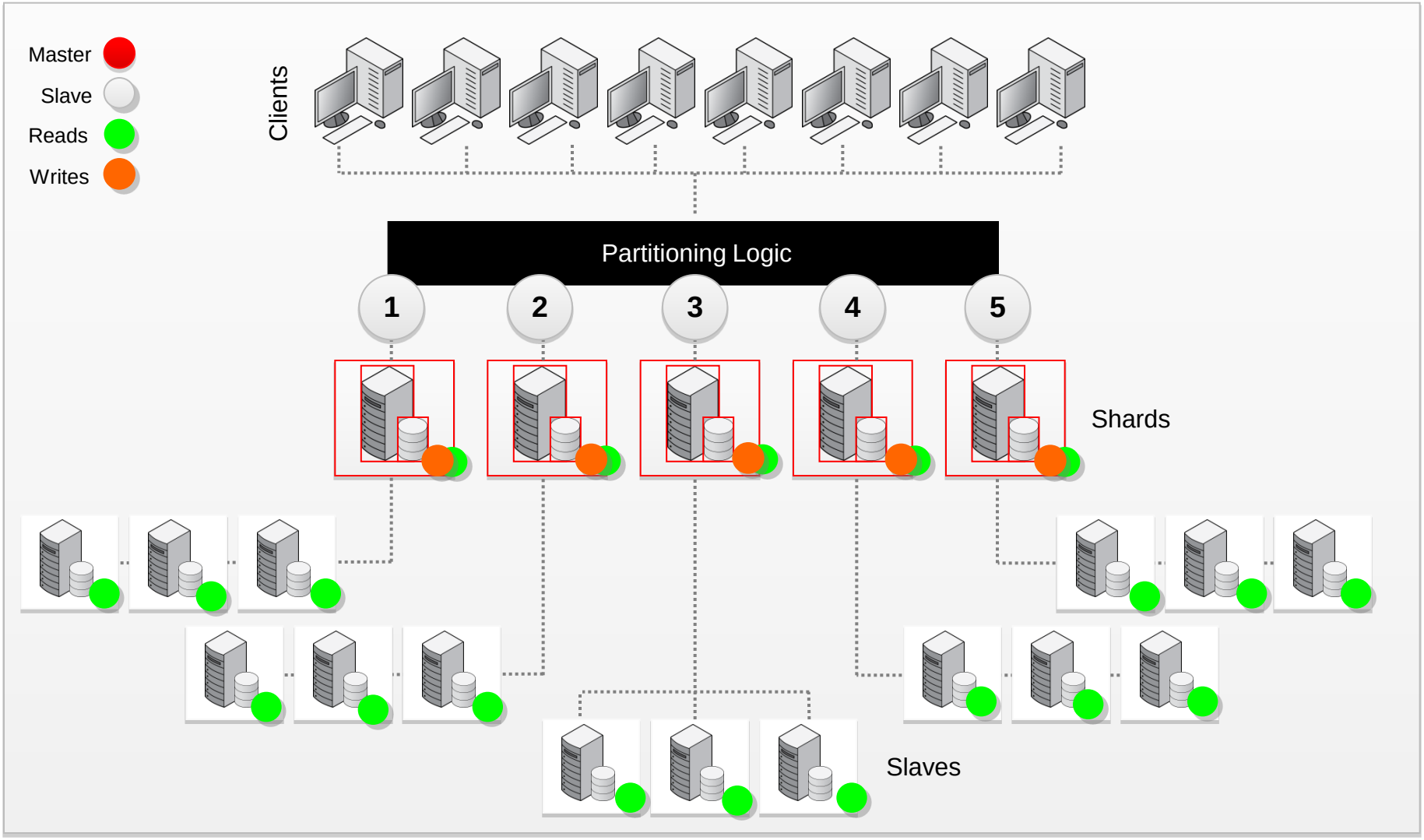


Best Practices

Social Networking Reference Architecture

- Builds on best practices of Web Ref Archs
- MySQL Cluster used for authentication & Look-Up table (Shard Catalog)
- Introduces Sharding
  - Implemented at the application layer for scaling very high volume of writes
  - Data divided into smaller sets, distributed across low-cost hardware
  - Shards based on Hash of a single column – ie. User ID
- Sharding is complex
  - Recommend the Architecture and Design Consulting Engagement
- Sharding only used in a small percentage of workloads
  - Most Web 2.0 workloads are still read-intensive, ie record is read before updates applied

Sharding aka Application Partitioning

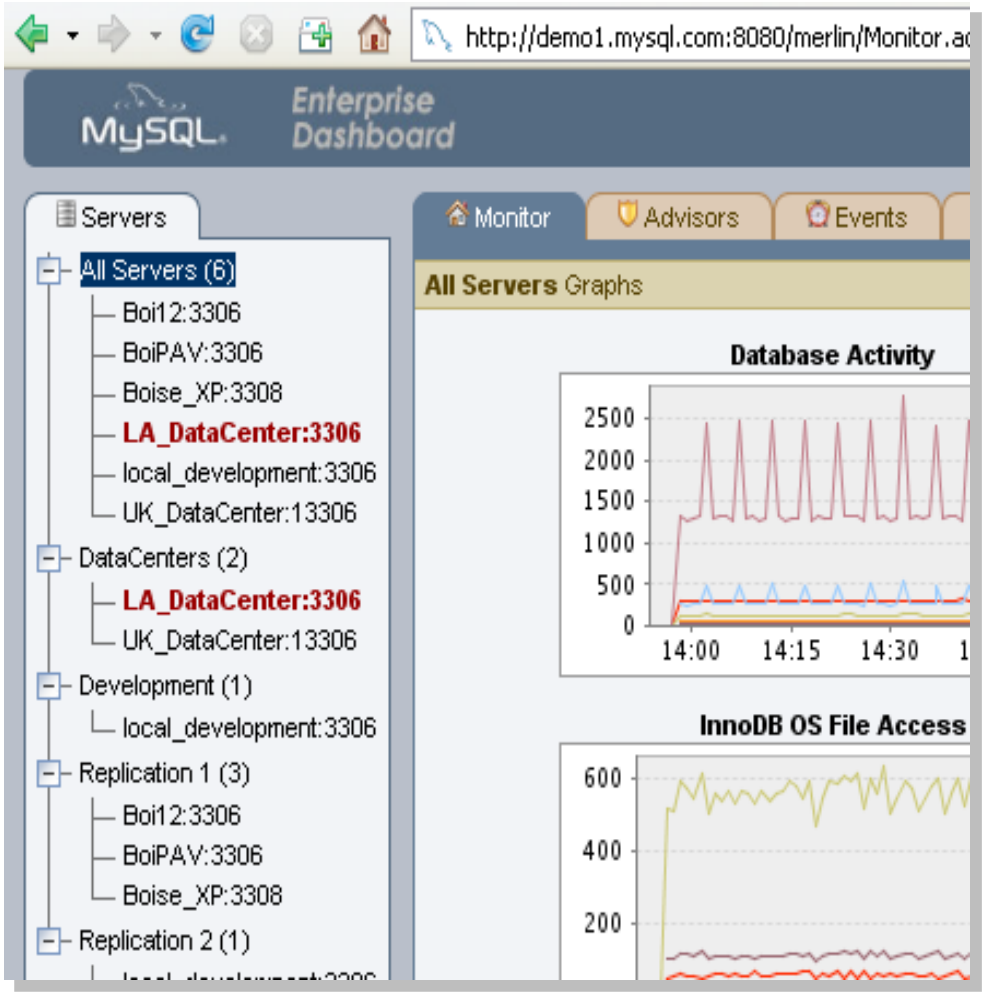


Value-Added Components & Services

Value-Add Components

MySQL Enterprise Monitor

- Single, consolidated view into entire MySQL environment
- Auto-discovery of MySQL servers, replication topologies
- Customizable rules-based monitoring and alerts
- Query monitoring and analysis
- Identifies problems *before* they occur
- Reduces risk of downtime
- Makes it easier to scale out without requiring more DBAs



A Virtual MySQL DBA Assistant!

Value-Add Components

MySQL Enterprise Back-Up

- Online Backup for InnoDB & MyISAM
- High Performance Backup & Restore
- Compressed Backup
- Full Backup
- Incremental Backup
- Partial Backups
- Point in Time Recovery
- Unlimited Database Size
- Cross-Platform
  - Windows, Linux, Unix

Value-Add Components

MySQL Cluster Manager

Automated Management

- Cluster-Wide Management
- Process Management
- On-Line Operations (Add Nodes / Upgrades / Reconfiguration)

Monitoring

- Status Monitoring & Recovery

HA Operations

- Disk Persistence
- Configuration Consistency
- HA Agent Operation

How Does MySQL Cluster Manager Help ?

Example: Initiating upgrade from MySQL Cluster 6.3 to 7.1

Before MySQL Cluster Manager

- 1 x preliminary check of cluster state
- 8 x ssh commands per server
- 8 x per-process stop commands
- 4 x scp of configuration files (2 x mgmd & 2 x mysqld)
- 8 x per-process start commands
- 8 x checks for started and re-joined processes
- 8 x process completion verifications
- 1 x verify completion of the whole cluster.
- Excludes manual editing of each configuration file.

Total: 46 commands -
2.5 hours of attended operation

With MySQL Cluster Manager

upgrade cluster --package=7.1 mycluster;

Total: 1 Command -
Unattended Operation

- Results
  - Reduces the overhead and complexity of managing database clusters
  - Reduces the risk of downtime resulting from administrator error
  - Automates best practices in database cluster management

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

Summary

- Reference Architectures designed as a springboard to innovating on the web
- Based on insight from most successful web properties
- Best practices & repeatable technologies for Scale & HA
- Next Step: Engage with MySQL Consulting from Oracle

Resources

- MySQL Web Reference Architectures Whitepaper
  http://www.mysql.com/why-mysql/white-papers/mysql_wp_high-availability_webrefarchs.php

- Designing and Implementing Scalable Applications with MySQL & Memcached http://www.mysql.com/why-mysql/white-papers/mysql_wp_cluster_ScalingWebServices.php

- MySQL Enterprise Whitepaper http://www.mysql.com/why-mysql/white-papers/mysql_wp_enterprise_ready.php

- MySQL Cluster for Web & eCommerce Applications, Whitepaper http://www.mysql.com/why-mysql/white-papers/mysql_wp_Cluster_For_OnlineApps.php

Thank you!

The preceding 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.


ORACLE®