

# Clustering for the Masses

## A Gentle Introduction to Tungsten for MySQL

Robert Hodges

CTO, Continuent, Inc.

# Topics

- / **What is the Problem?**
- / **What is Tungsten and how does it work?**
- / **What can you do with it?**
- / **Prove it!**
- / **Summary and Questions**

# About Continuent

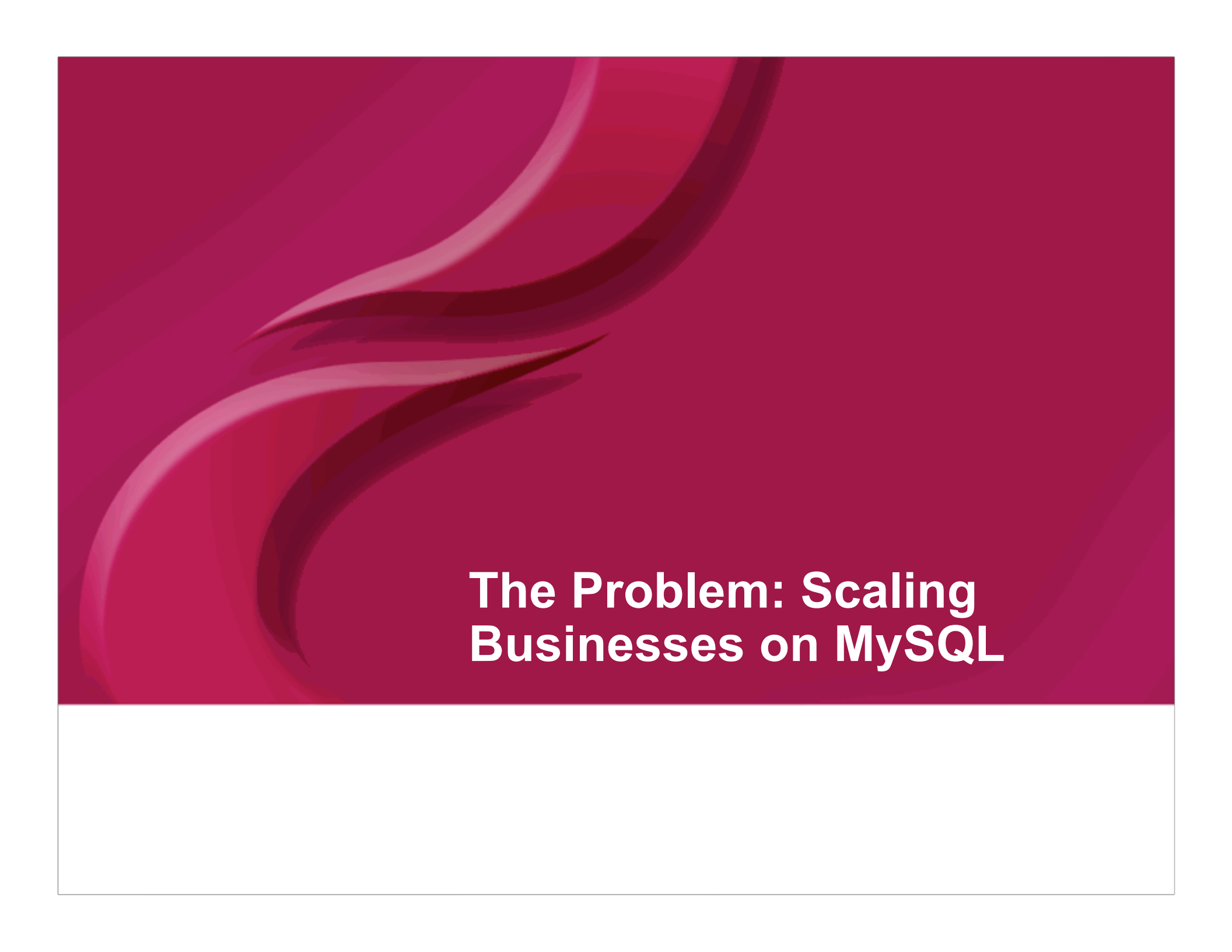
## / **Our Business: Full service database management**

### / **Value:**

- Ensure data are available when and where you need them
- Make open source DBMS scale economically

### / **Technical Expertise**

- Database replication
- Cluster management
- Application connectivity
- SaaS applications

The background of the slide is a deep red color. On the left side, there are several thick, flowing, wavy lines in a lighter shade of red, creating a sense of movement and depth. The right side of the slide is a solid, uniform red. The text is positioned in the lower right quadrant of the red area.

# **The Problem: Scaling Businesses on MySQL**

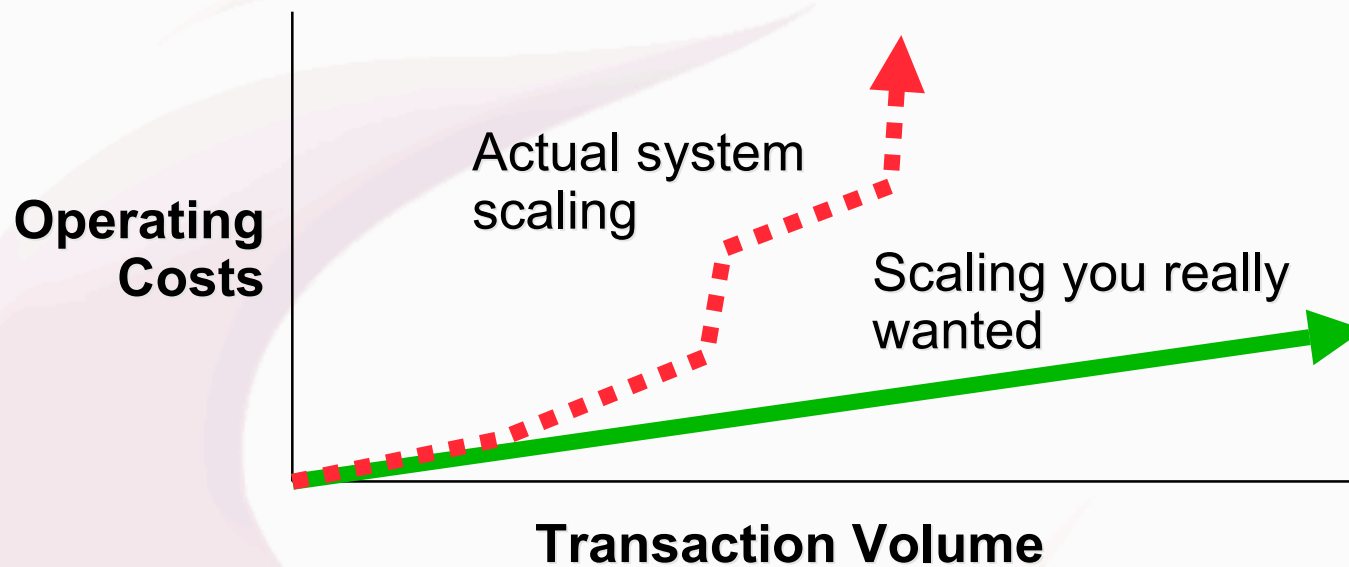
# MySQL Is Not as Cheap as You Think

**We have a system in place based on shared mysql + memcache but it's quickly becoming prohibitively costly (in terms of manpower) to operate. We need a system that can grow in a more automated fashion and be highly available.**

**Ryan King - Twitter**

# MySQL Scaling Problems

- / Failures and maintenance on complex topologies
- / SLAs on shared resources
- / Multi-tenant operations
- / Growing transaction and data volumes
- / Administrative labor costs



# Convert to NoSQL?? NOT!

- / Integrate clusters of off-the-shelf MySQL databases**
- / Hide management complexity behind simple interfaces**
- / Use hardware efficiently**
- / Manage multi-tenant data and resource use**
- / Provide seamless upgrade**
- / Operate across multiple sites**
- / Protect data with decent backup/restore and consistency checking**

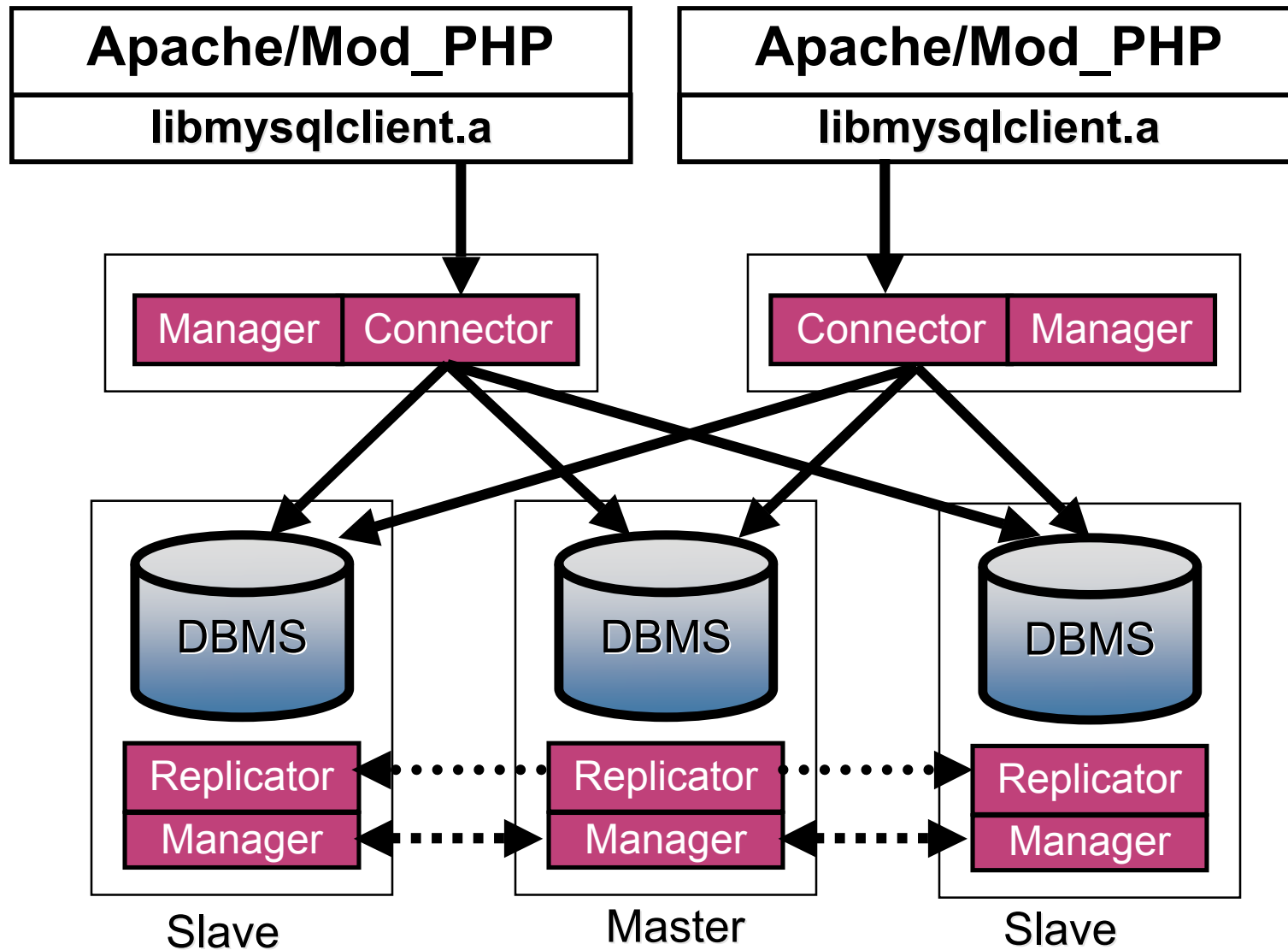


# **What is Tungsten and How Does It Work?**

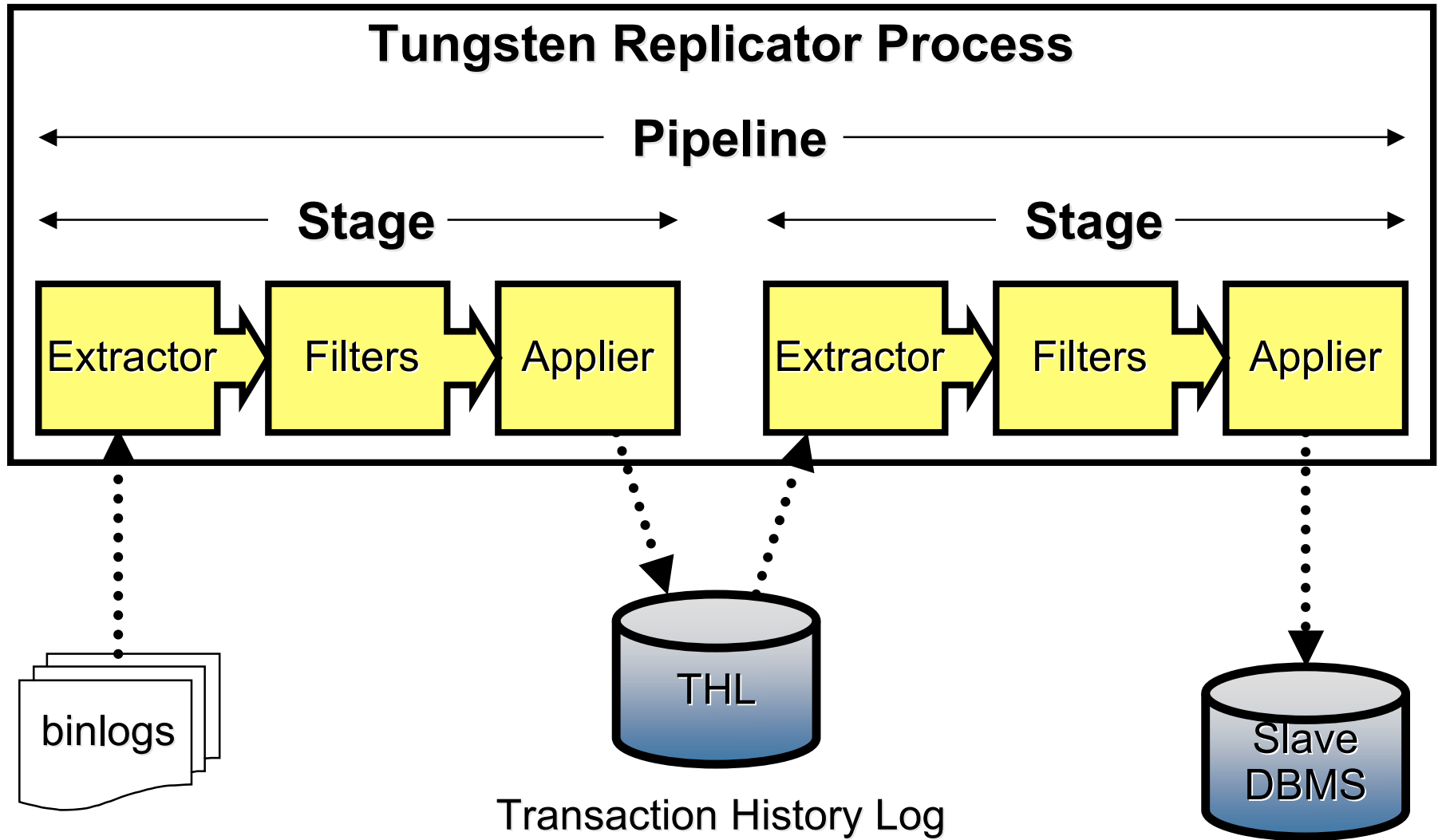
# What Is Tungsten?

- / Tungsten implements data services consisting of off-the-shelf databases linked by replication to:**
  - Protect data
  - Maintain high availability
  - Improve resource utilization
  - Raise performance
- / Install and set up in a few minutes**
- / Integrated backup/restore and data integrity checks**
- / Rule-driven management with automated failover**
- / Simple procedures for updates and maintenance**
- / SQL scaling for multi-tenant systems**

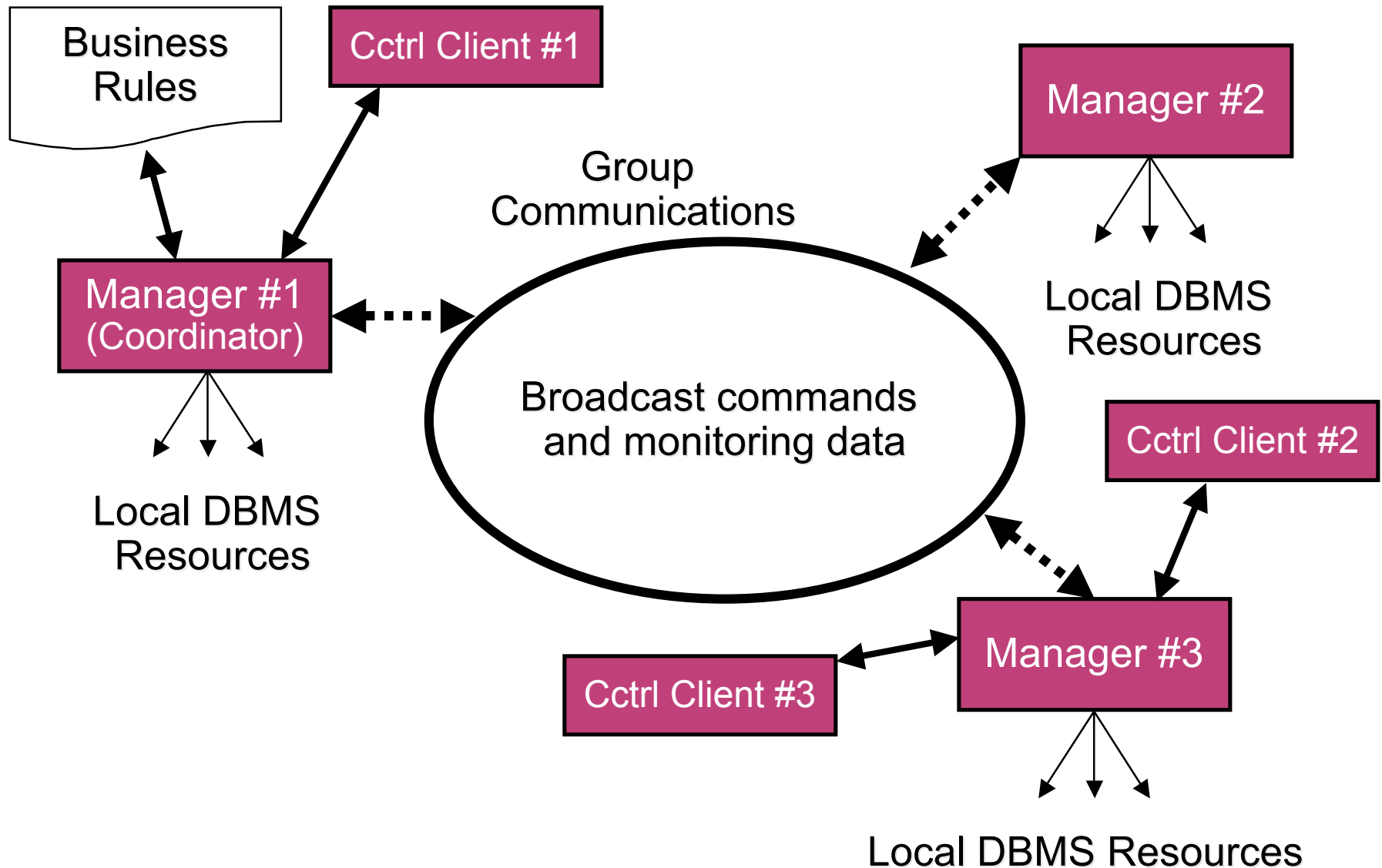
# Tungsten Data Service Architecture



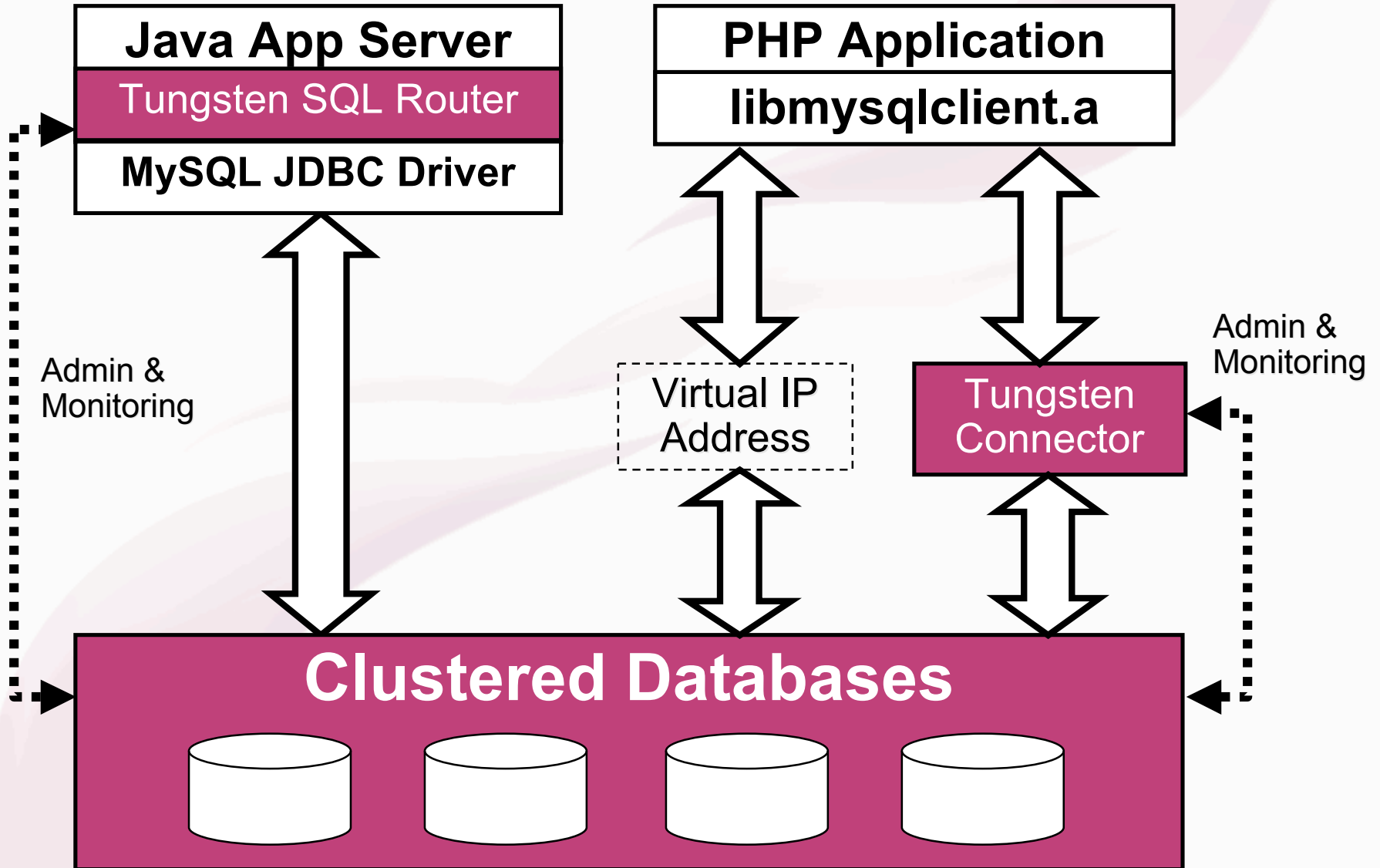
# Flexible Replication Pipelines



# Distributed Rule-Based Management



# Managed Routes to Databases



# How Do You Set Up A Cluster?

## / **Fulfill host pre-reqs**

- MySQL, Ruby, Java, network hygiene

## / **Download Tungsten software to host and un-tar**

- Or build it yourself from source!

## / **Run 'configure' script**

## / **Services start automatically and join cluster as soon as replicator goes online for the first time**

## / **Rinse and repeat on each host**

# How Do You Administer Clusters?

/ **'cctrl'** is the Tungsten management client

/ **Connect to any manager to administer cluster**

/ **List data sources in the cluster**

```
[LOGICAL] /cluster/demo/> ls
```

/ **Check cluster liveness**

```
[LOGICAL] /cluster/demo/> cluster heartbeat
```

/ **Switch master and slave**

```
[LOGICAL] /cluster/demo/> switch to prod01
```

/ **Start and stop operating system services**

```
[LOGICAL] /cluster/demo/> service */mysql restart
```

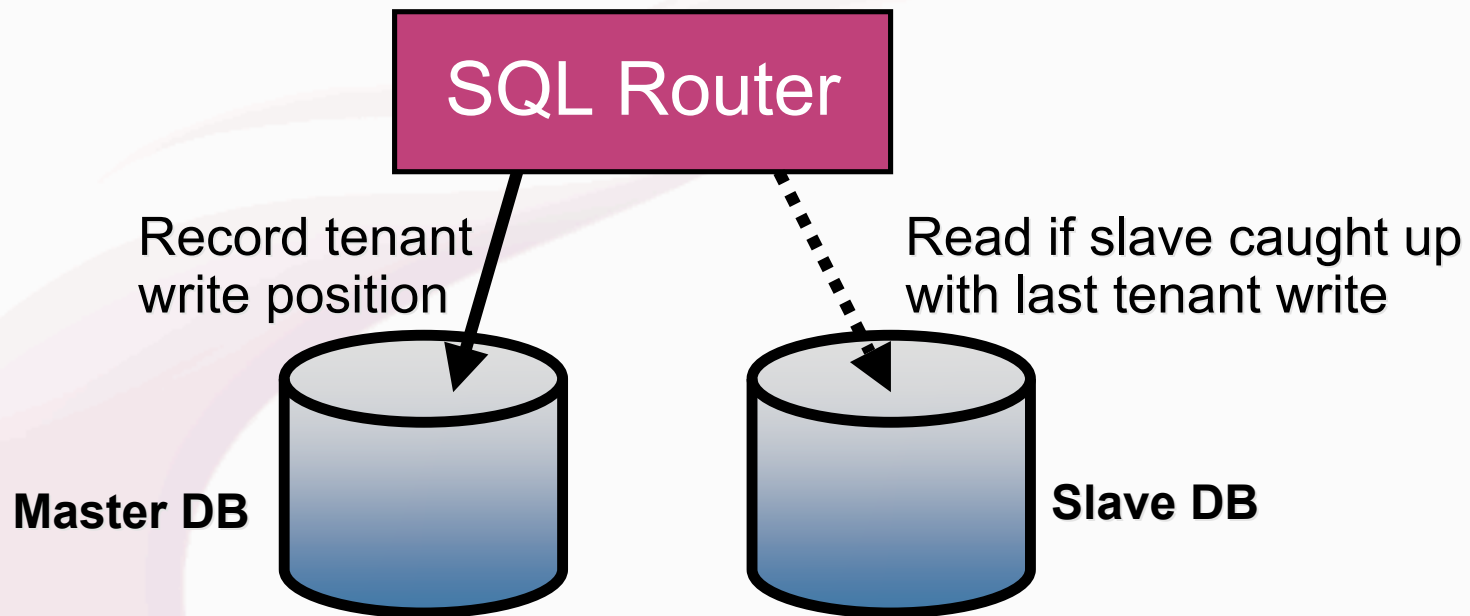
/ **Backup and restore databases**

```
[LOGICAL] /cluster/demo/> datasource prod02 backup
```

```
[LOGICAL] /cluster/demo/> datasource prod02 restore
```

# How Do You Scale Reads?

- / Tungsten “session-consistency” distributes tenant reads to slaves if caught up with last tenant write
- / Works with tenants sharded by database or with independent database connections



# How Do You Handle Failures?

/ **Automatic policy mode performs master failover automatically whenever:**

- A database crashes
- A replicator fails or crashes
- A server reboots or drops off the network

/ **Manual policy mode lets you failover manually**

```
[LOGICAL] /cluster/demo/> failover
```

/ **Failover procedure promotes most advanced slave**

- And points other slaves to the slave

/ **Recover command brings back failed master**

```
[LOGICAL] /cluster/demo/> datasource centos5a recover
```

/ **Slaves recover without admin intervention in automatic mode**

# How Do You Check Data Consistency?

## / Tungsten incremental checks verify data without stopping replication

- Implemented as replication events inside Tungsten Replicator
- Compute MD5 on all or part of tables
- Can do single tables or all tables in a database

## / Run checks from manager in combination with flush command

```
[LOGICAL] /cluster/demo/> cluster check tenant1.*  
[LOGICAL] /cluster/demo/> cluster flush
```

## / Slaves can either fail or warn of failures

# How Do You Perform Maintenance?

- / Tungsten supports “daylight maintenance”
- / To perform maintenance on a slave, just take it offline

```
[LOGICAL] /cluster/demo/> set policy maintenance
[LOGICAL] /cluster/demo/> datasource centos5b offline
[LOGICAL] /cluster/demo/> replicator centos5b offline
```
- / To perform maintenance on a master, first do a switch to promote a new master

```
[LOGICAL] /cluster/demo/> switch
```
- / Then perform maintenance using the slave procedure
- / Switch the master back when you are done



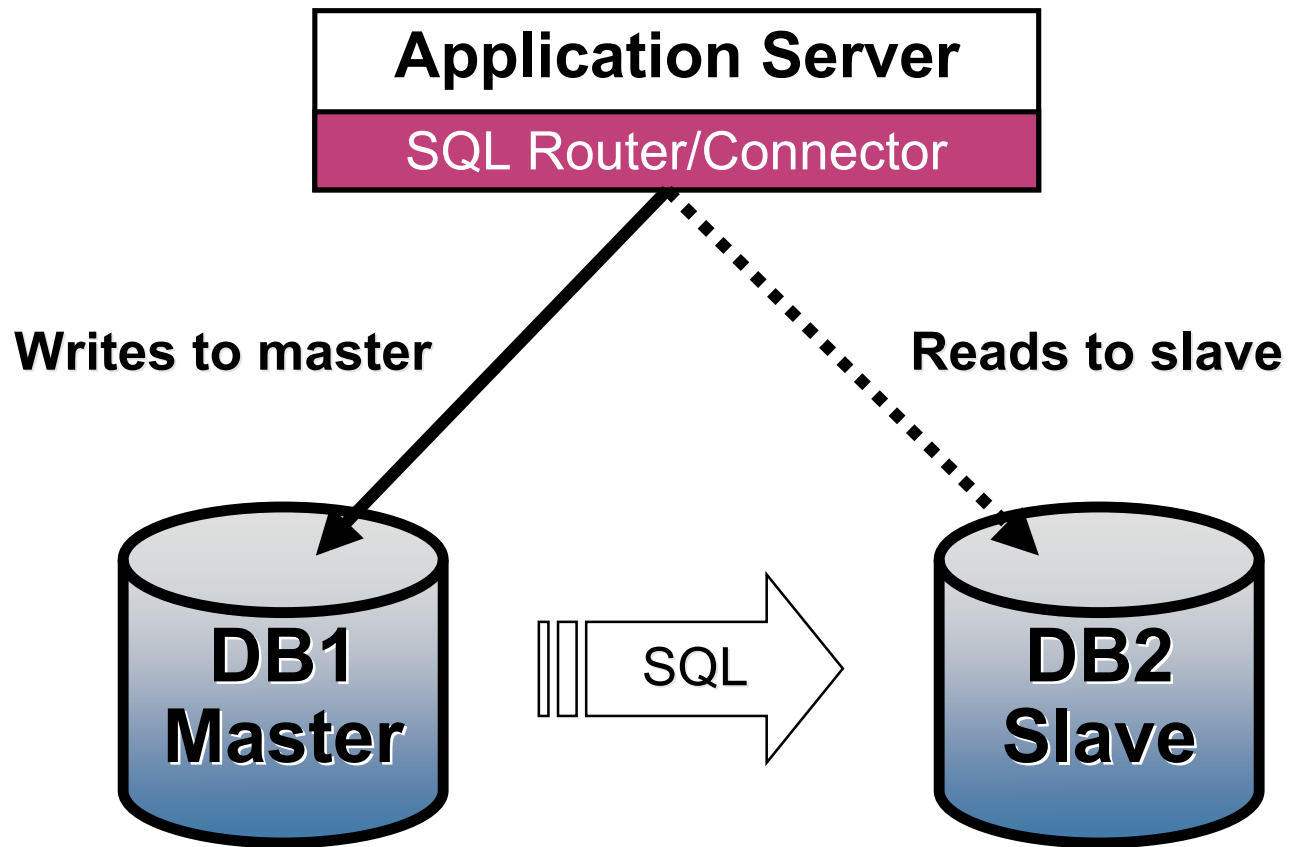
# What Can I Do with Tungsten?

# Lots of Things...

- / Raise application availability without restarts**
- / Scale applications using reads on slaves**
- / Perform daylight maintenance on hosts**
- / Archive selected data off-site**
- / Ensure applications are online**
- / Replicate from databases into files, queues, etc.**
  - Some assembly required

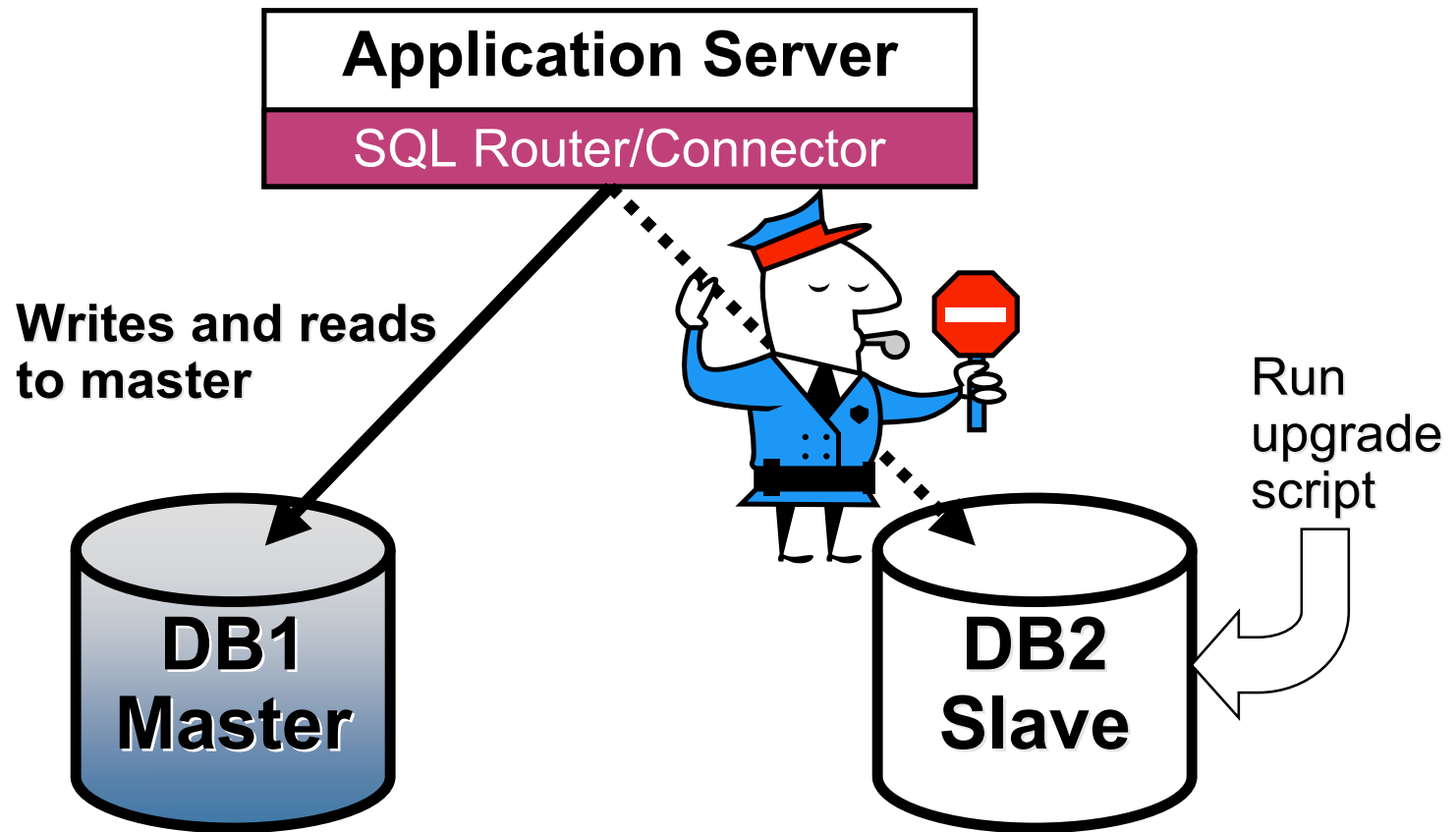
## How about online schema upgrade?

# Online Upgrade



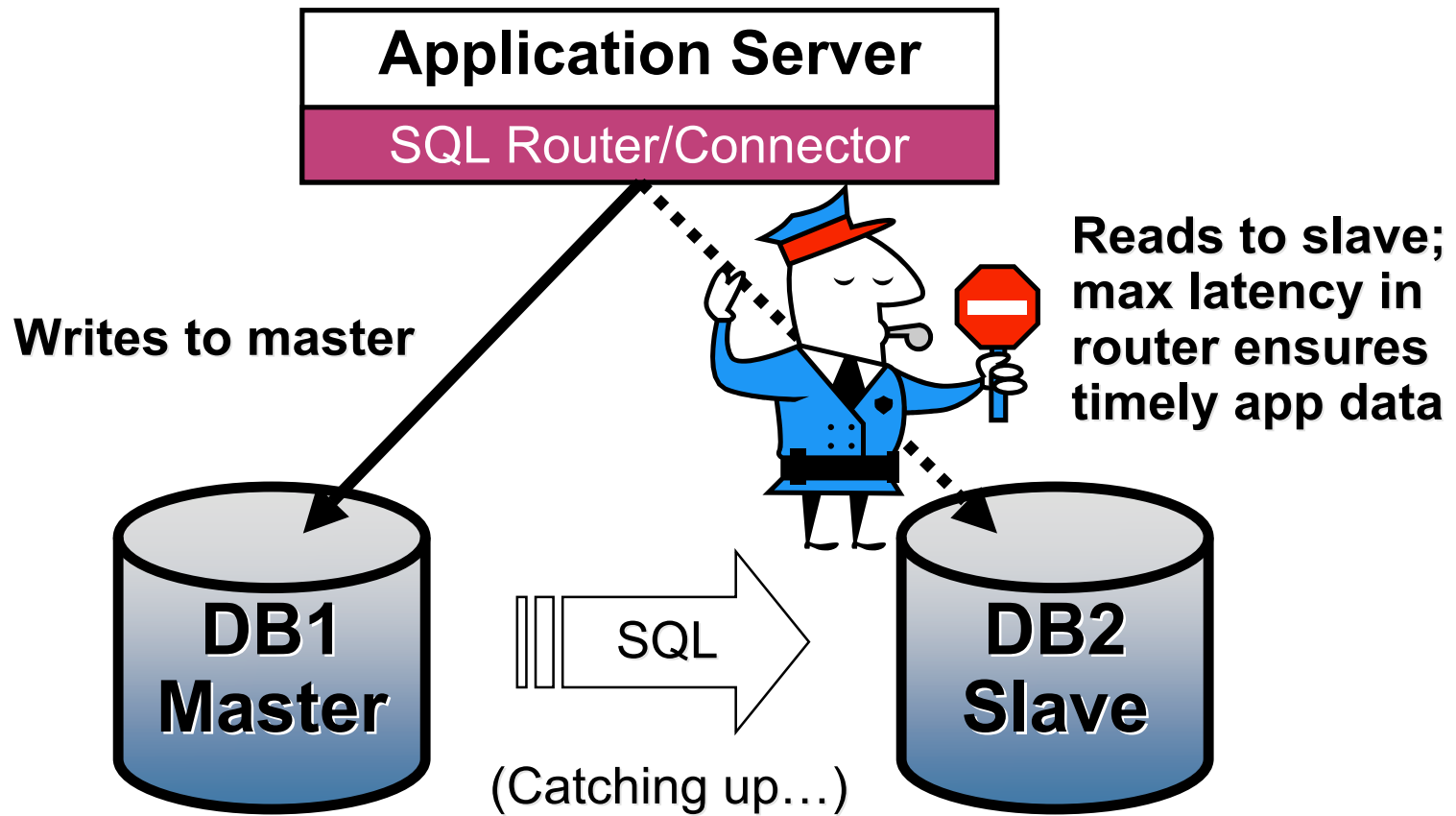
**Step 0: Master and slave both online**

# Online Upgrade



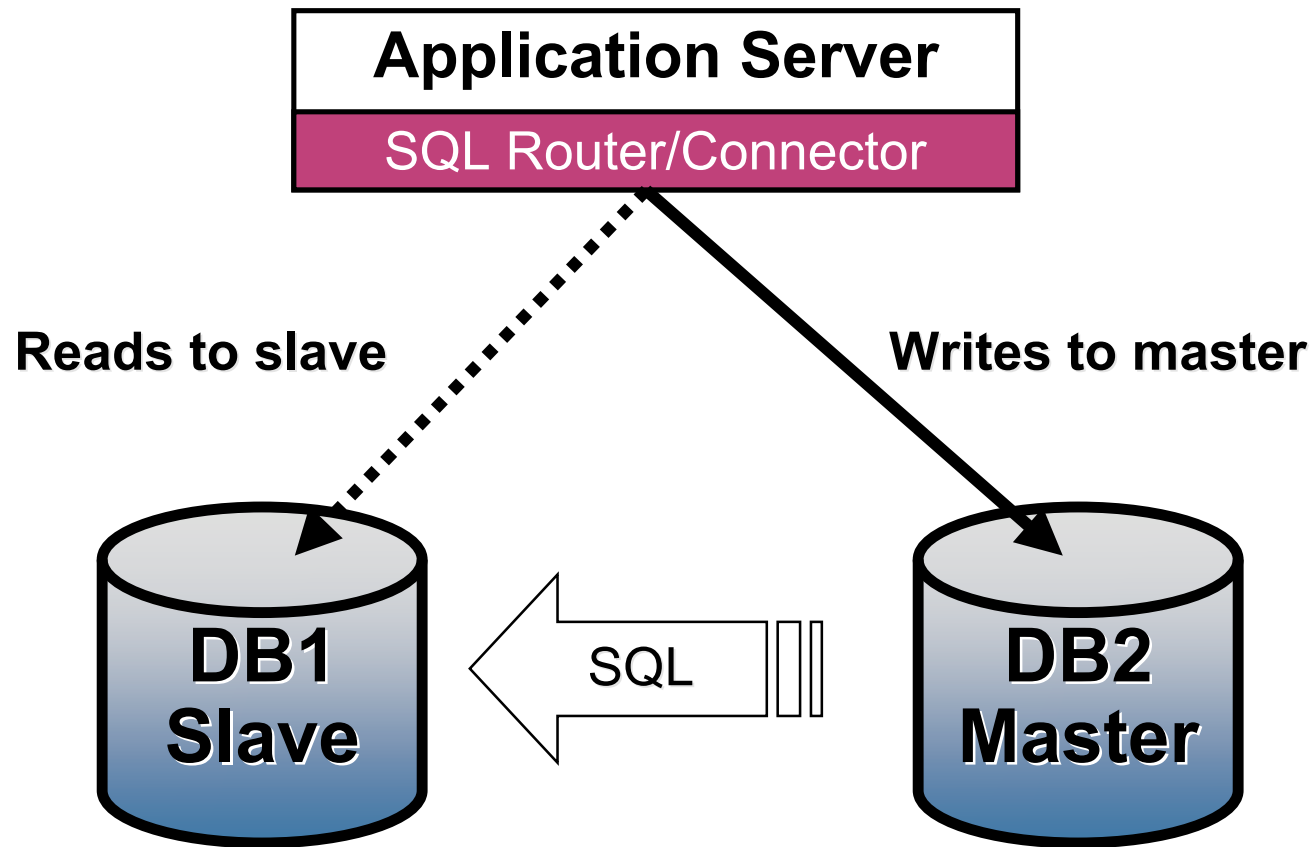
## Step 1: Shun Slave and Upgrade

# Online Upgrade



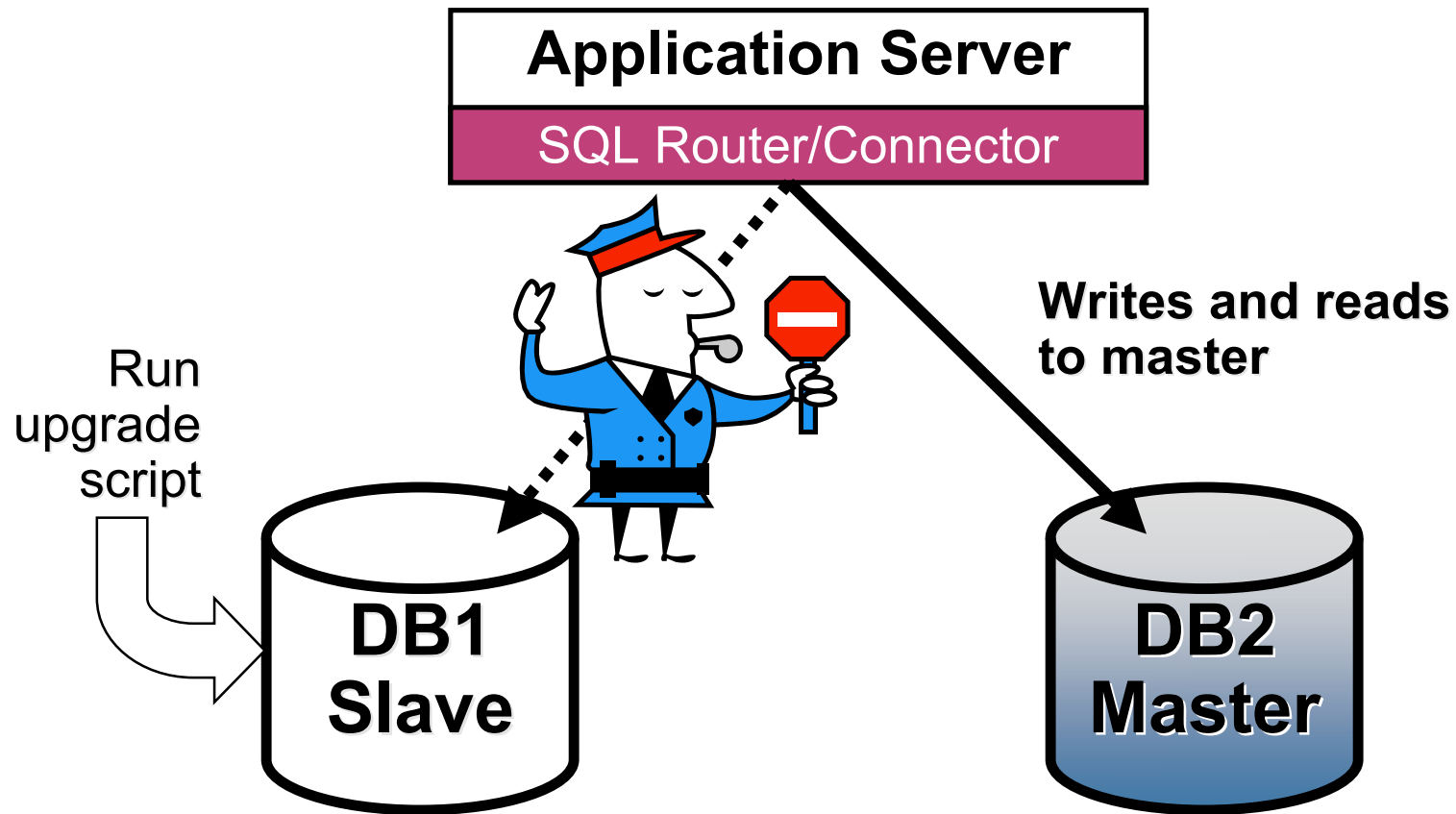
## Step 2: Welcome Slave Back

# Online Upgrade



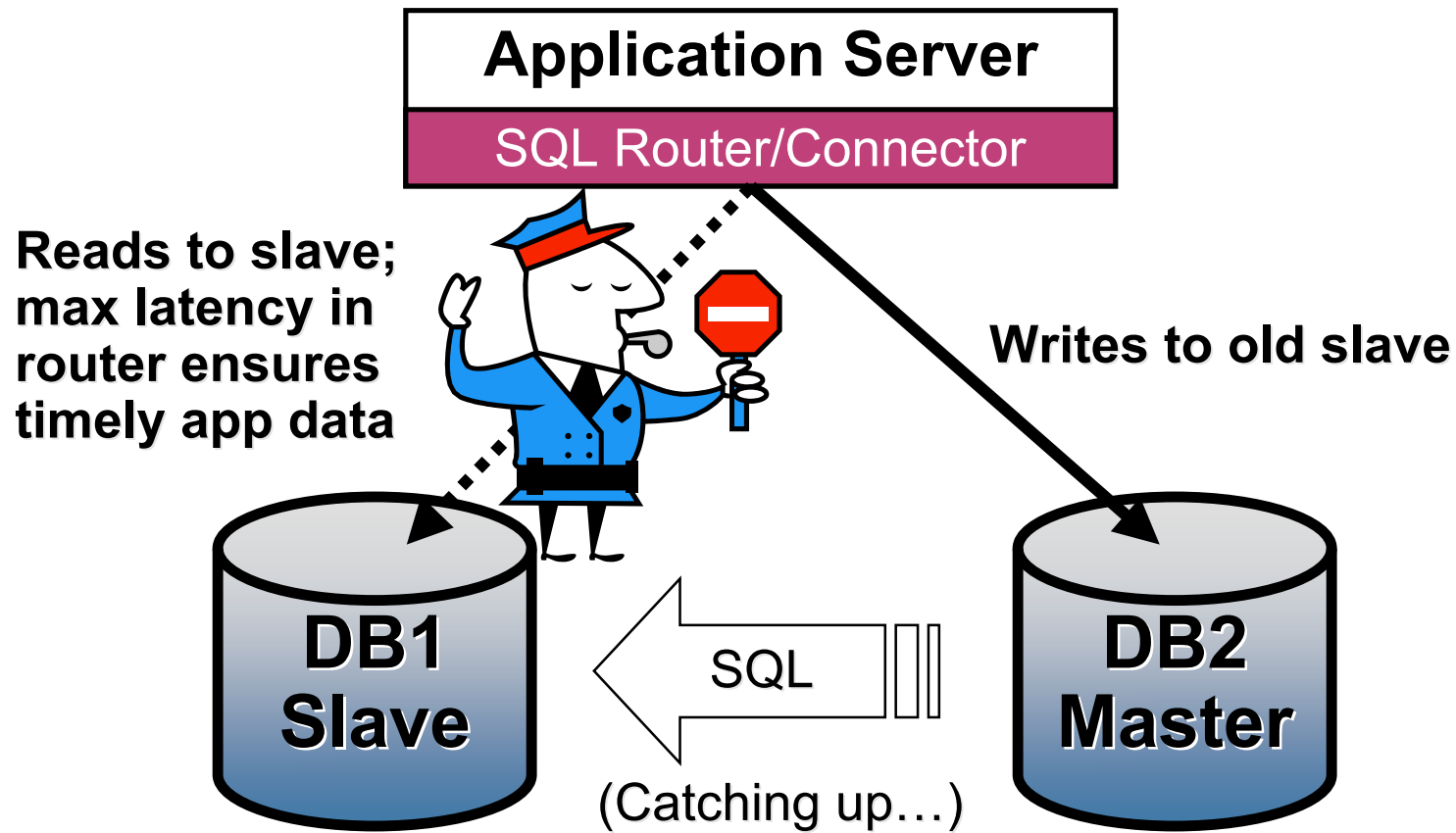
## Step 3: Switch master

# Online Upgrade



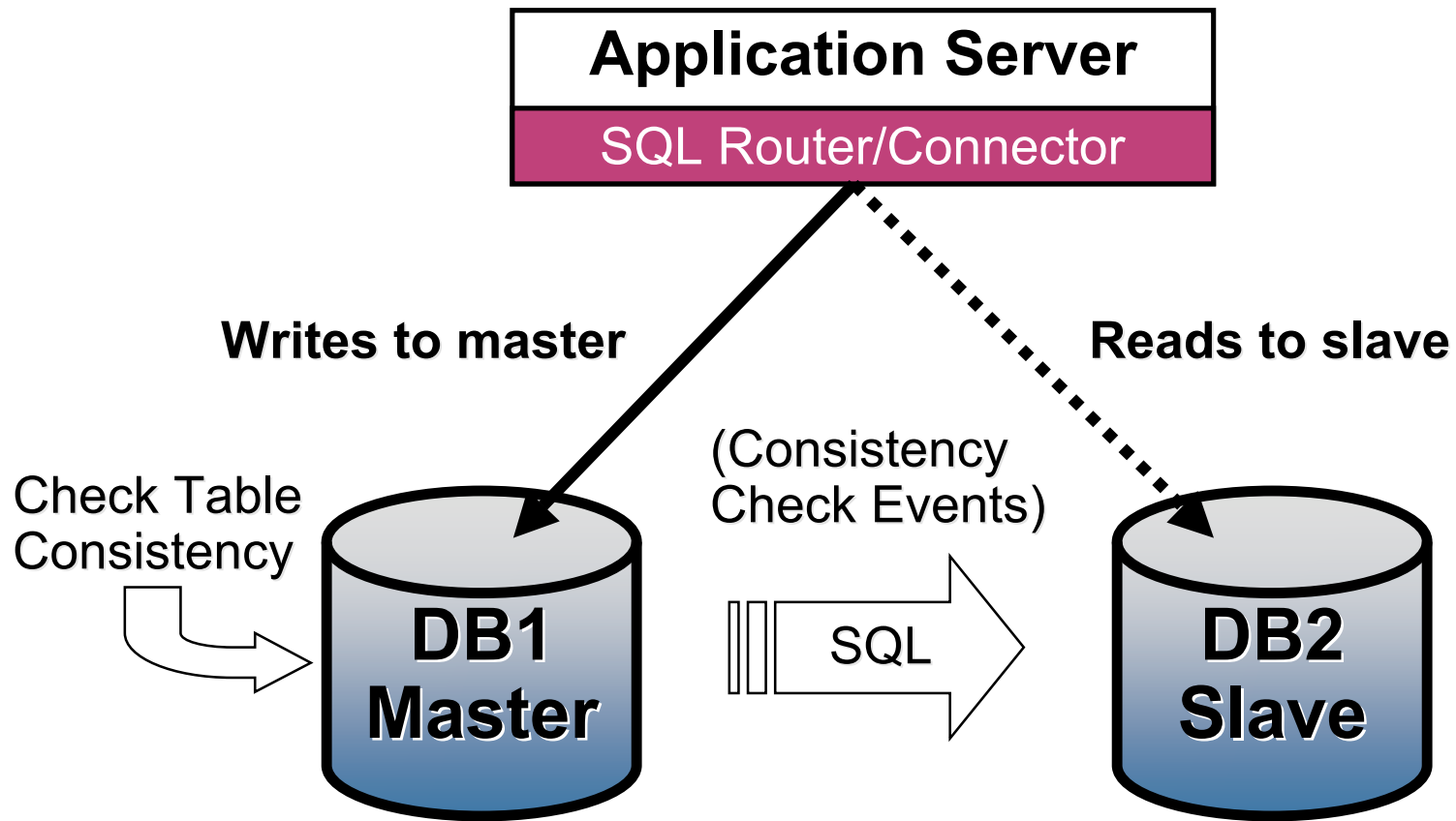
## Step 4: Shun Slave and Upgrade

# Online Upgrade



## Step 5: Welcome slave back

# Online Upgrade



**Step 5: Switch back to original master and check consistency**

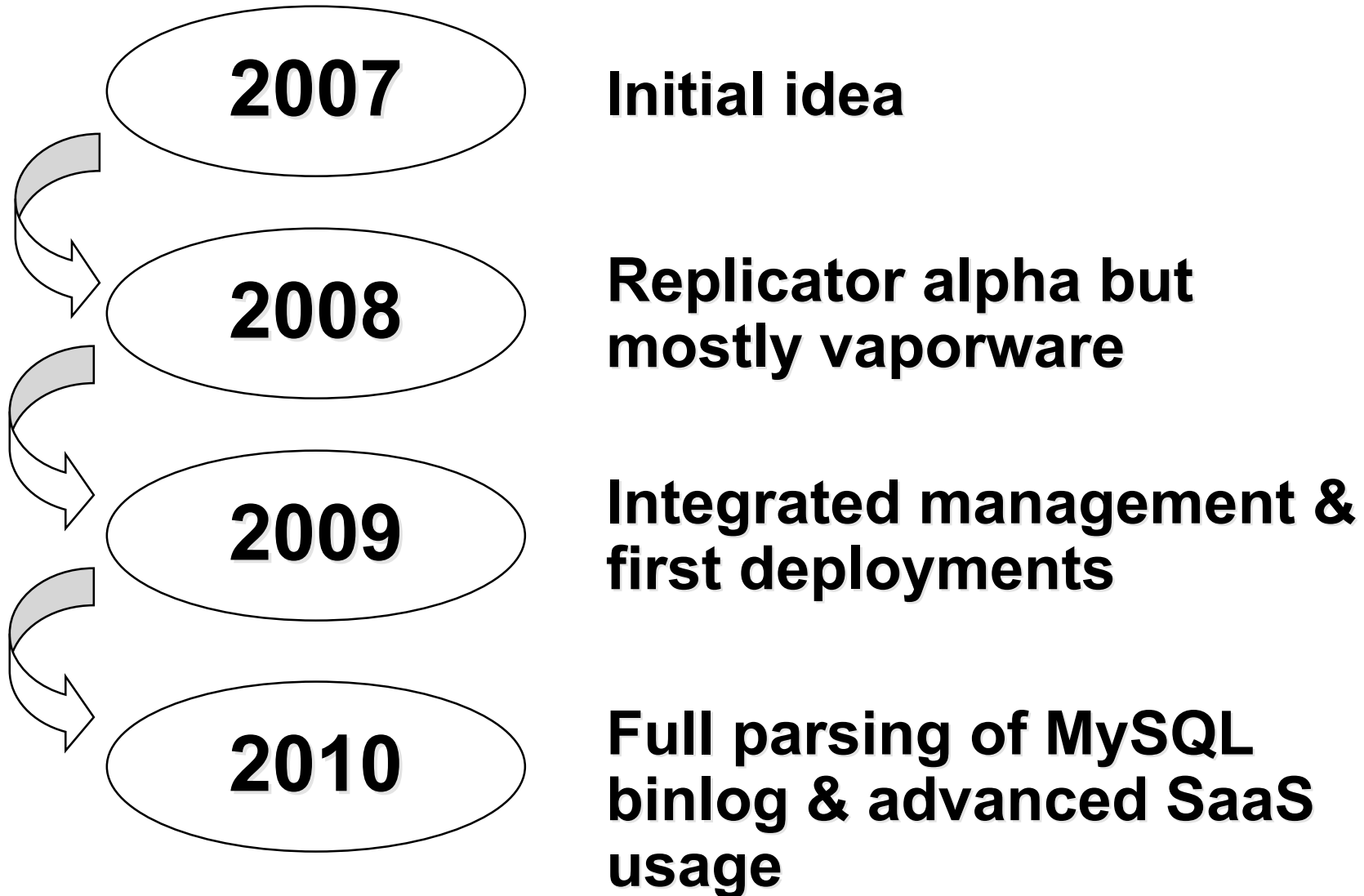
# Does This Really Work?

## Time for a demo!



# Conclusion

# Tungsten Development History



# 2010 Development Roadmap

- / Tungsten is adding features rapidly for clusters based on MySQL 5.0/5.1

<b>Feature</b>	<b>Beta</b>	<b>Prod</b>
<b>Flexible replication pipelines</b>	<b>Q1</b>	<b>NOW</b>
<b>100% of binlog replicated</b>	<b>Q1</b>	<b>NOW</b>
<b>Shard-based parallel replication</b>	<b>Q2</b>	<b>Q2-3</b>
<b>Fast event logging</b>	<b>Q2</b>	<b>Q2-3</b>
<b>Low-latency WAN replication</b>	<b>Q2</b>	<b>Q2-3</b>

- / PostgreSQL 8 warm standby support and adding features to manage PostgreSQL 9
- / Official Drizzle support as soon as we get customers

# Conclusion

- / **Scaling MySQL economically is harder than it looks**
- / **Tungsten solves the problem through easy-to-deploy data services that integrate:**
  - Flexible replication
  - Unitary management
  - Transparent connectivity
- / **Check out Tungsten at <http://www.continuent.com>**

# Contact Information

---

## HQ and Americas

560 S. Winchester Blvd., Suite 500  
San Jose, CA 95128  
Tel (866) 998-3642  
Fax (408) 668-1009

## EMEA and APAC

Lars Sonckin kaari 16  
02600 Espoo, Finland  
Tel +358 50 517 9059  
Fax +358 9 863 0060

**e-mail:** [sales@continuent.com](mailto:sales@continuent.com)

**[robert.hodges at continuent.com](mailto:robert.hodges@continuent.com)**  
**<http://scale-out-blog.blogspot.com>**

**Continuent Web Site:**  
**<http://www.continuent.com>**