



ORACLE®

MySQLの高可用性構成

日本オラクル 株式会社

MySQL Global Business Unit

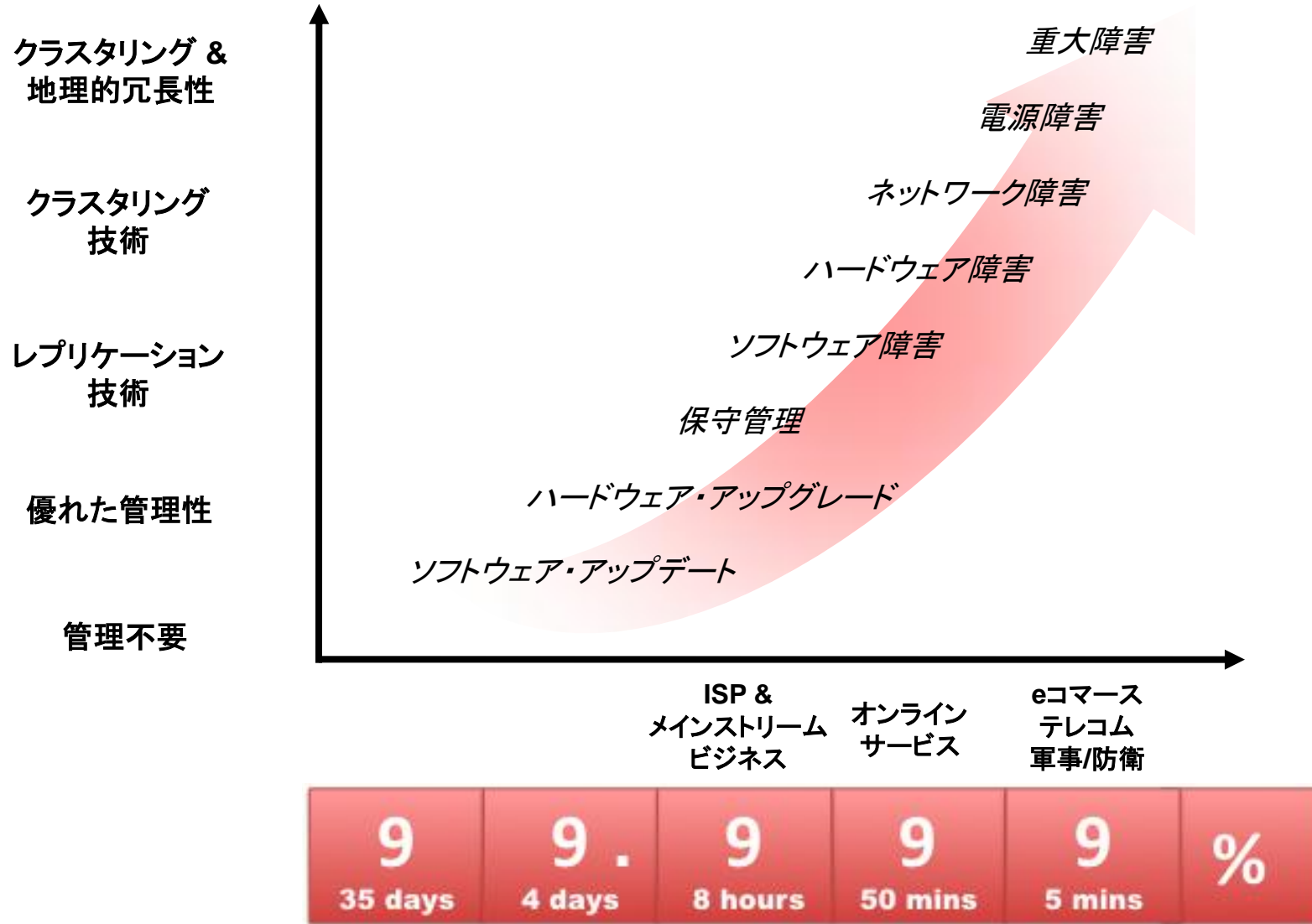
以下の事項は、弊社の一般的な製品の方向性に関する概要を説明するものです。また、情報提供を唯一の目的とするものであり、いかなる契約にも組み込むことはできません。以下の事項は、マテリアルやコード、機能を提供することをコミットメント(確約)するものではないため、購買決定を行う際の判断材料になさらないで下さい。オラクル製品に関して記載されている機能の開発、リリースおよび時期については、弊社の裁量により決定されます。

OracleとJavaは、Oracle Corporation 及びその子会社、関連会社の米国及びその他の国における登録商標です。文中の社名、商品名等は各社の商標または登録商標である場合があります。

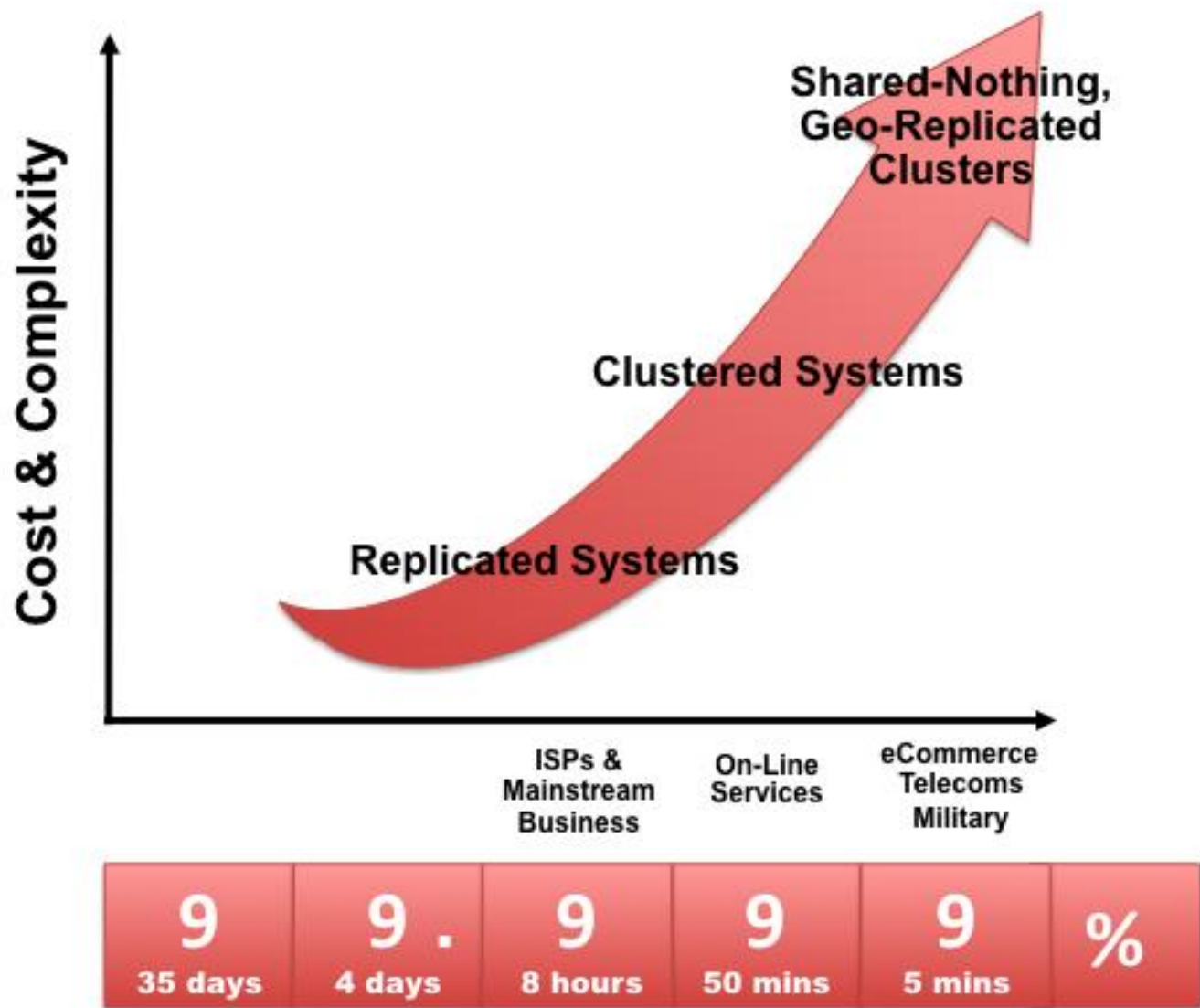


The world's most popular open source database
世界で最も普及しているオープンソース データベース

99.999%の可用性

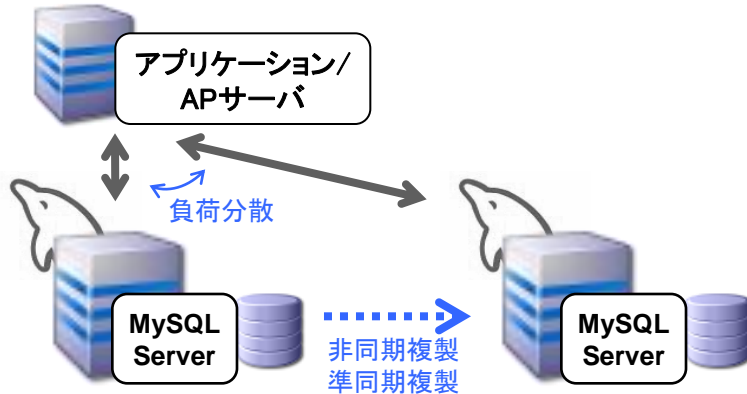


高可用アーキテクチャと可用性レベル

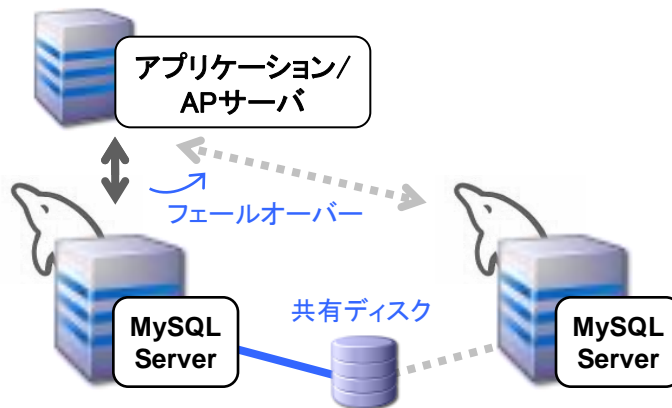


MySQLの高可用性構成

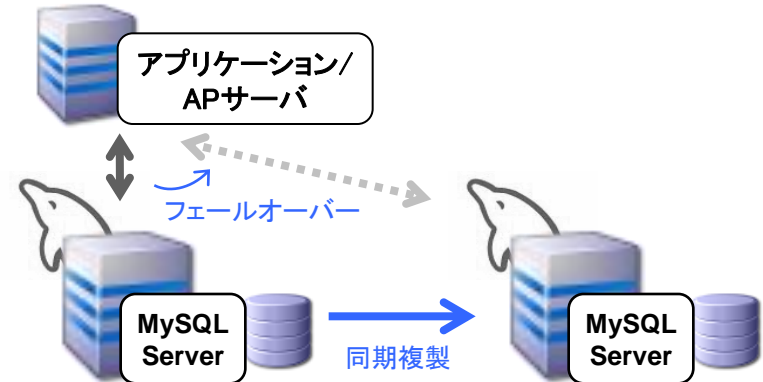
- レプリケーション(標準機能)
非同期&準同期データレプリケーション



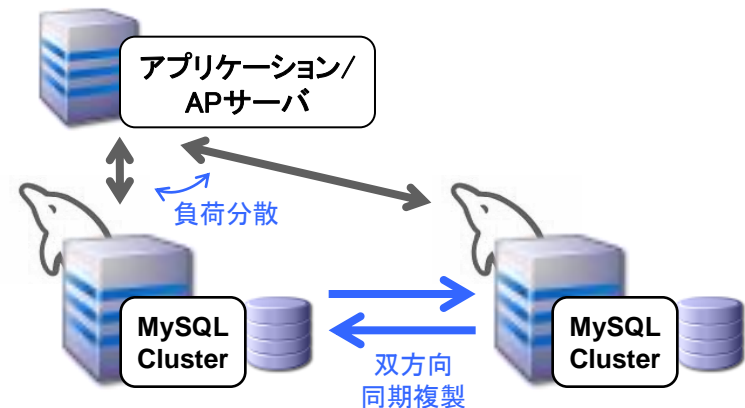
- 3rdベンダ製HAソフト利用
共有ディスクにデータを格納



- MySQL+DRBD
Linux用のノード間データコピー

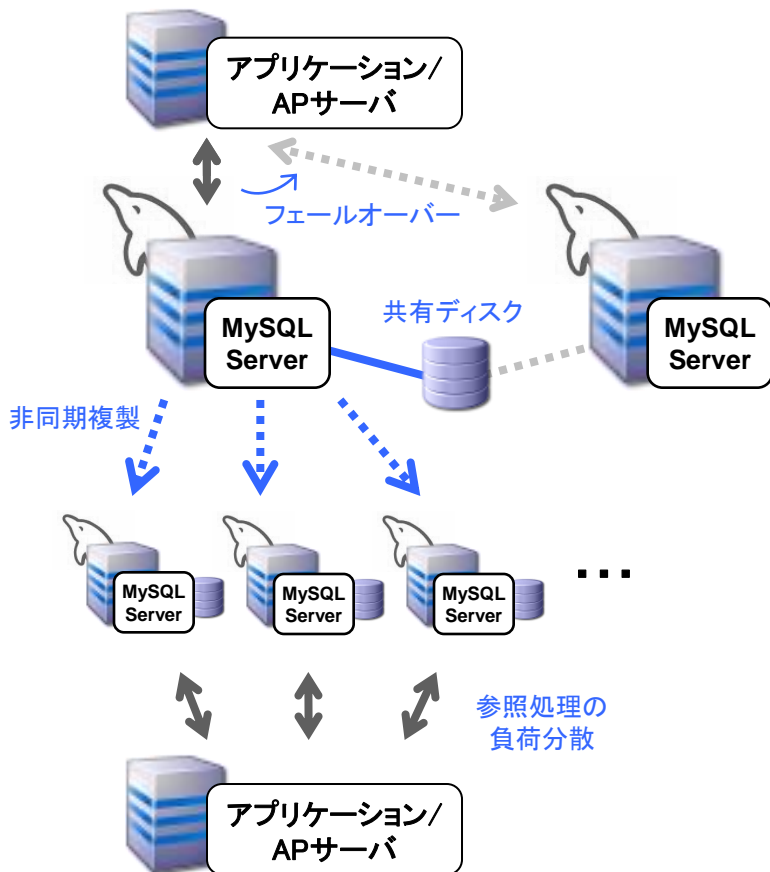


- MySQL Cluster
シェアードナッシング型高性能クラスタ

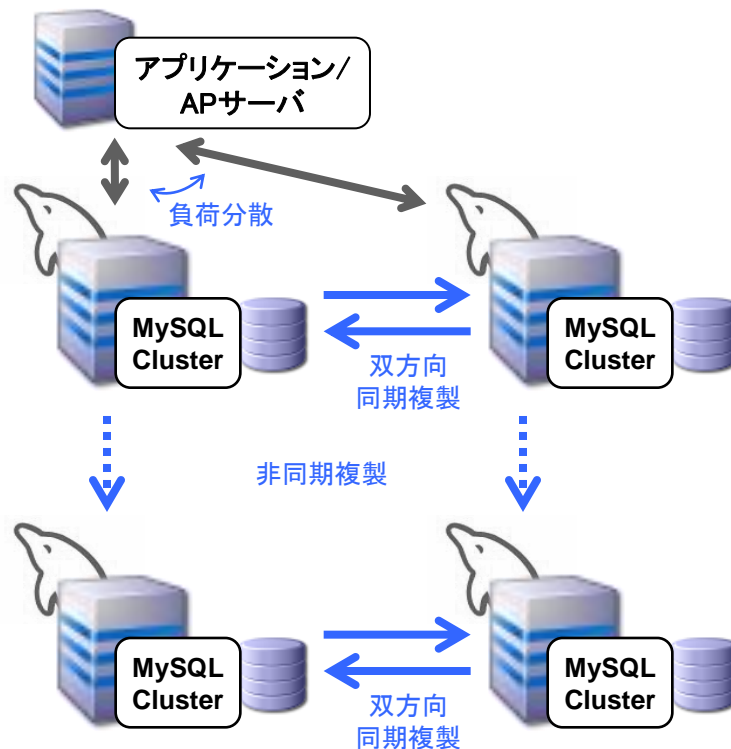


複合型の高可用性構成例

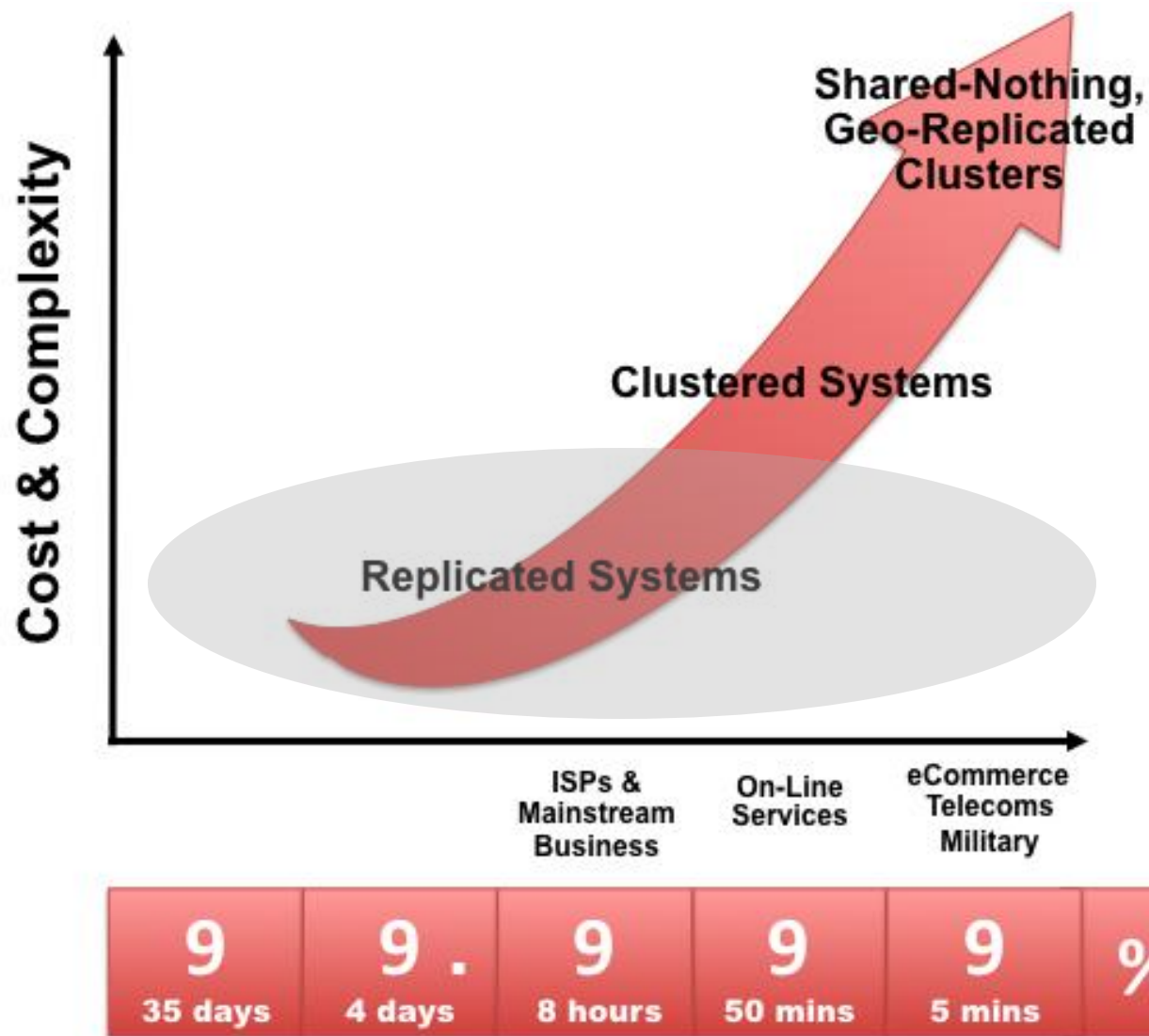
- 共有ディスク型構成
+レプリケーション



- MySQL Cluster
+レプリケーション



レプリケーション



レプリケーションとは？

データの変更点を1つ以上の場所に複製すること

非同期レプリケーション - Asynchronous Replication



同期レプリケーション - Synchronous Replication

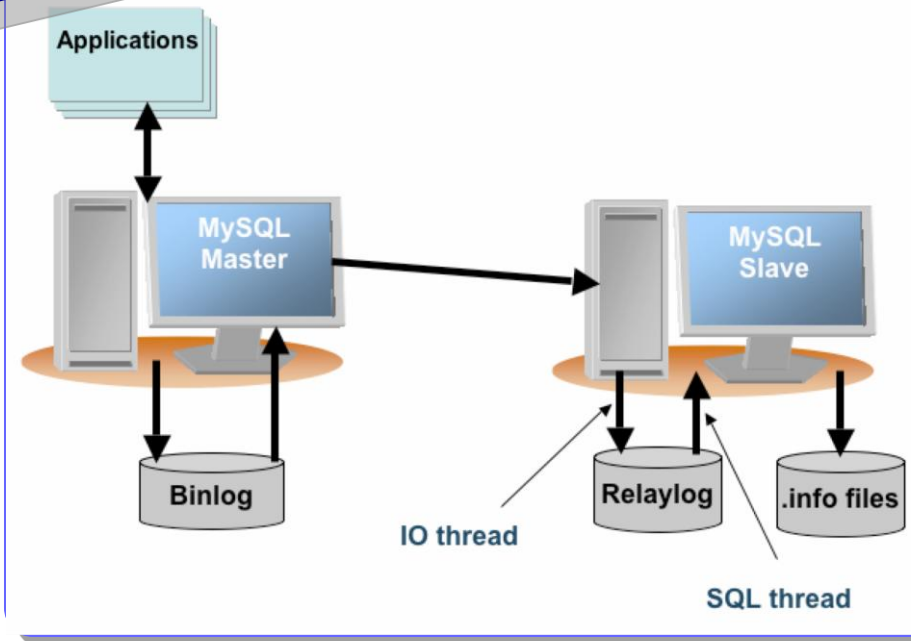


ORACLE

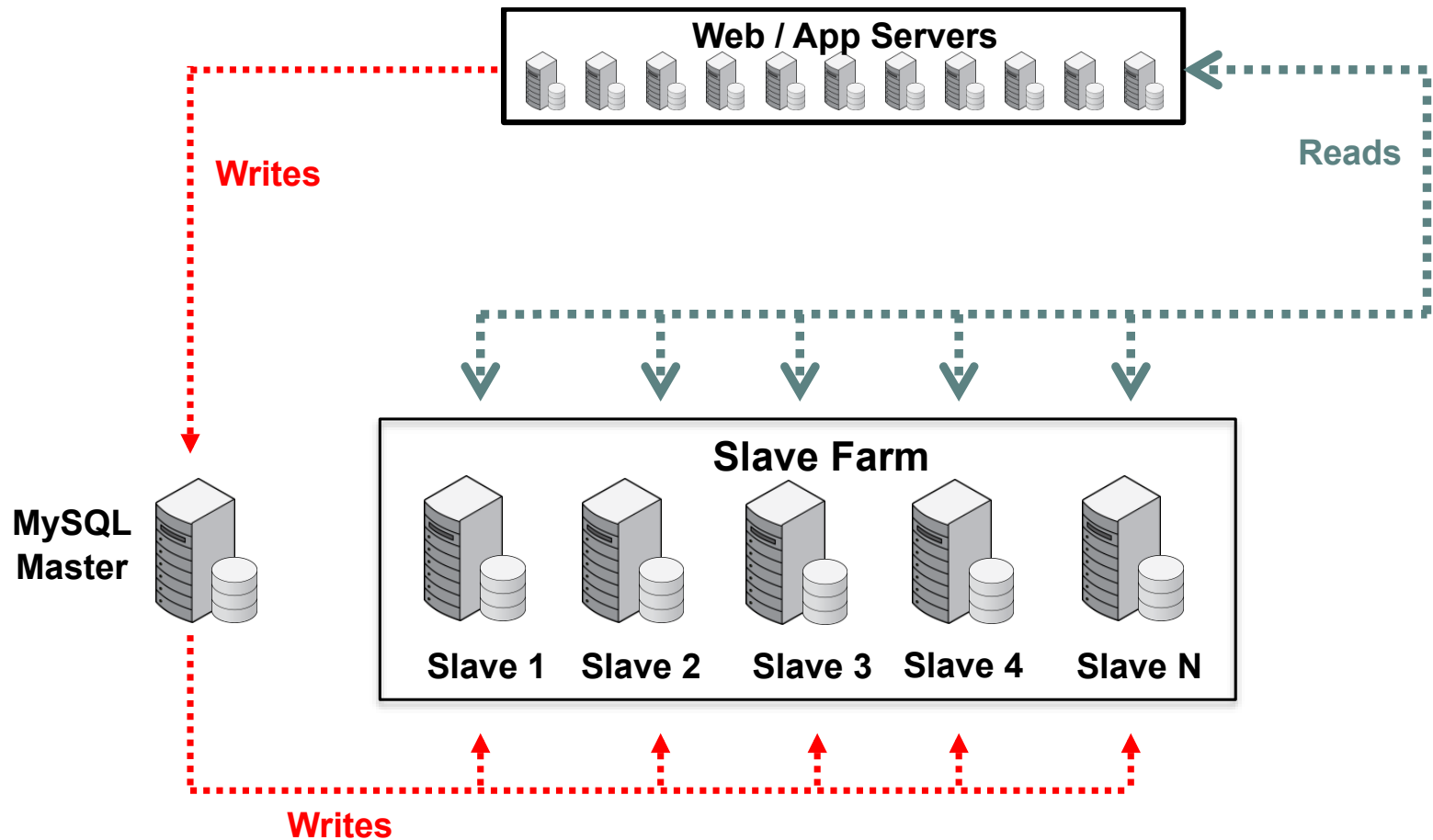
レプリケーション

- MySQLの標準機能
 - シンプルな設定
 - マスタ→スレーブ
 - 多数Webでの実績
- 非同期型&準同期型
- 特徴
 - 参照性能を向上させる構成
 - バックアップ用途での利用も
 - 基本は一方向でのデータコピーだが、双方向や循環型での利用も可能（データの更新には注意が必要）
 - 更新ログ(bin-log)を利用

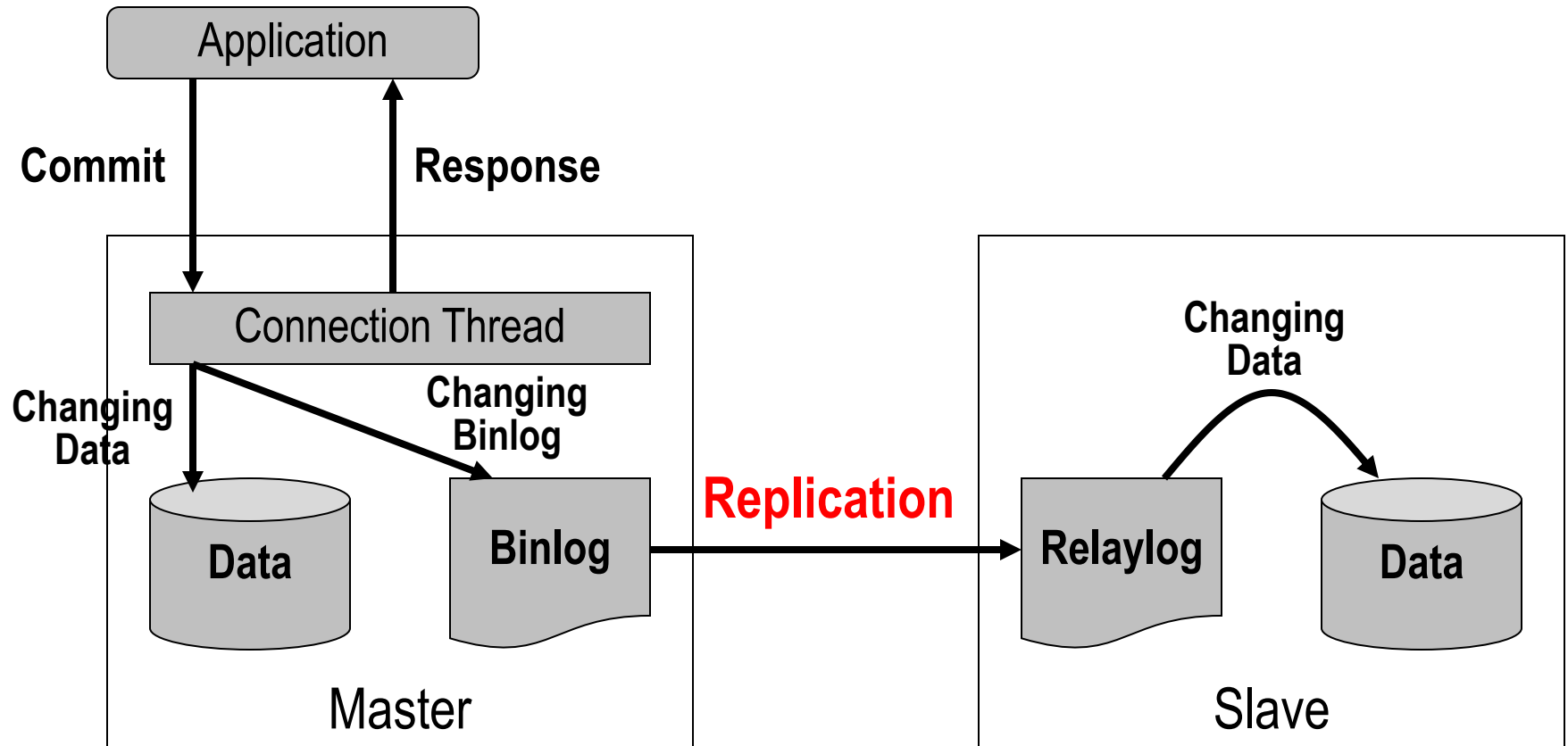
Webアプリケーションでは参照が95%、
更新が5%というケースも
-> シンプルなスケールアウト構成によって
簡単に大幅な性能向上が図れる



MySQLのレプリケーションを利用したスケールアウト構成

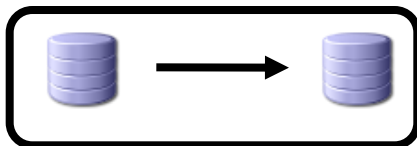


非同期レプリケーション

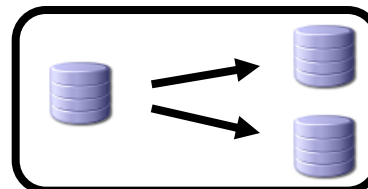


MySQL レプリケーションの構成パターン

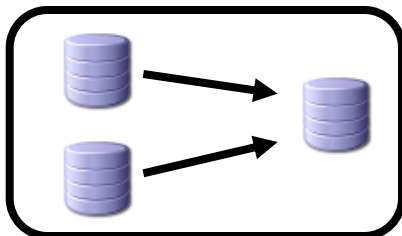
マスタ > スレーブ



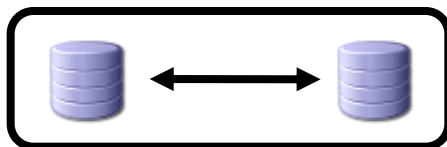
マスタ > マルチスレーブ



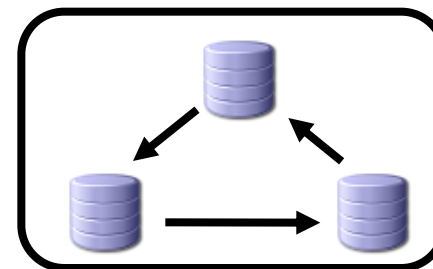
マルチマスタ > スレーブ (マルチソース)



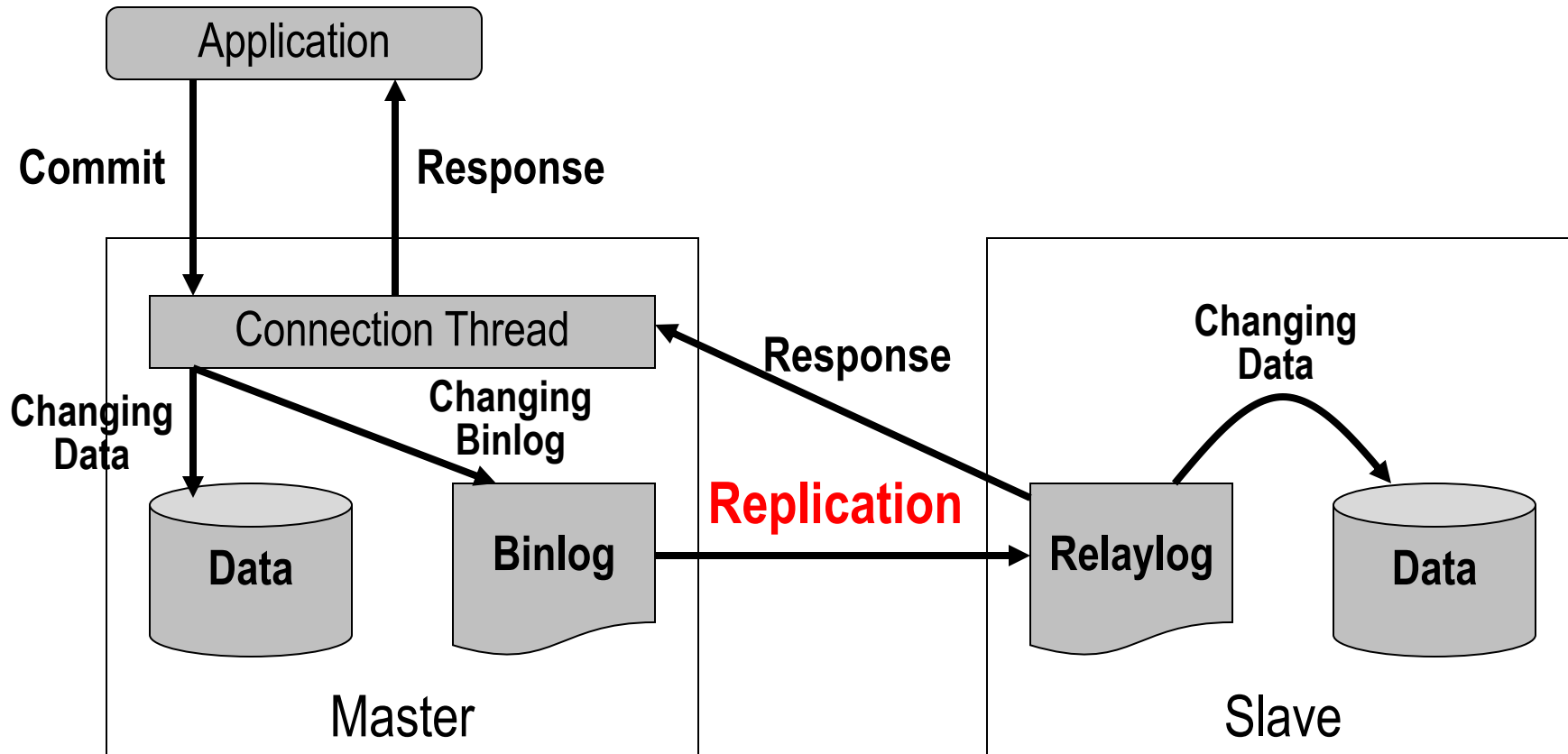
マスタ < > マスタ (マルチマスタ)



循環型 (マルチマスタ)



準同期レプリケーション



MySQL 5.5のレプリケーション改善点

- 1. 準同期(Semisynchronous)レプリケーション**
スレーブに変更点を転送してからアプリケーションに
応答を返すため、信頼性が向上
- 2. スレーブでのfsync改良&リレーログの自動復旧**
fsyncsを改良しスレーブの障害時のログ破損を回避
破損した場合にも自動普及
- 3. レプリケーション・ハードビート**
障害発生検知のメカニズムの精度を向上
- 4. サーバ毎のレプリケーション・フィルタリング**
特定のサーバIDを持つマスターのイベントを無視

MySQL 5.5のレプリケーション改善点

5. スレーブでのデータ型変換の自動化(RBR)

マスターとスレーブ間でデータ型が異なる場合に自動的に型を変換

6. ログを個別にフラッシュ

'FLUSH LOGS'実行時にログを選択してフラッシュ可能に

7. トランザクション非対応に関係なく安全にログに記録

InnoDBとMyISAMのテーブルを同一のトランザクション内で更新した場合、コミット後に全変更点をバイナリログに記録

レプリケーション監視の自動化

MySQL Enterprise Monitor

- 自動でレプリケーション構成、マスタ/スレーブを検出
- リアルタイムでレプリケーションの稼働状況を収集
- 同期に問題があれば通知

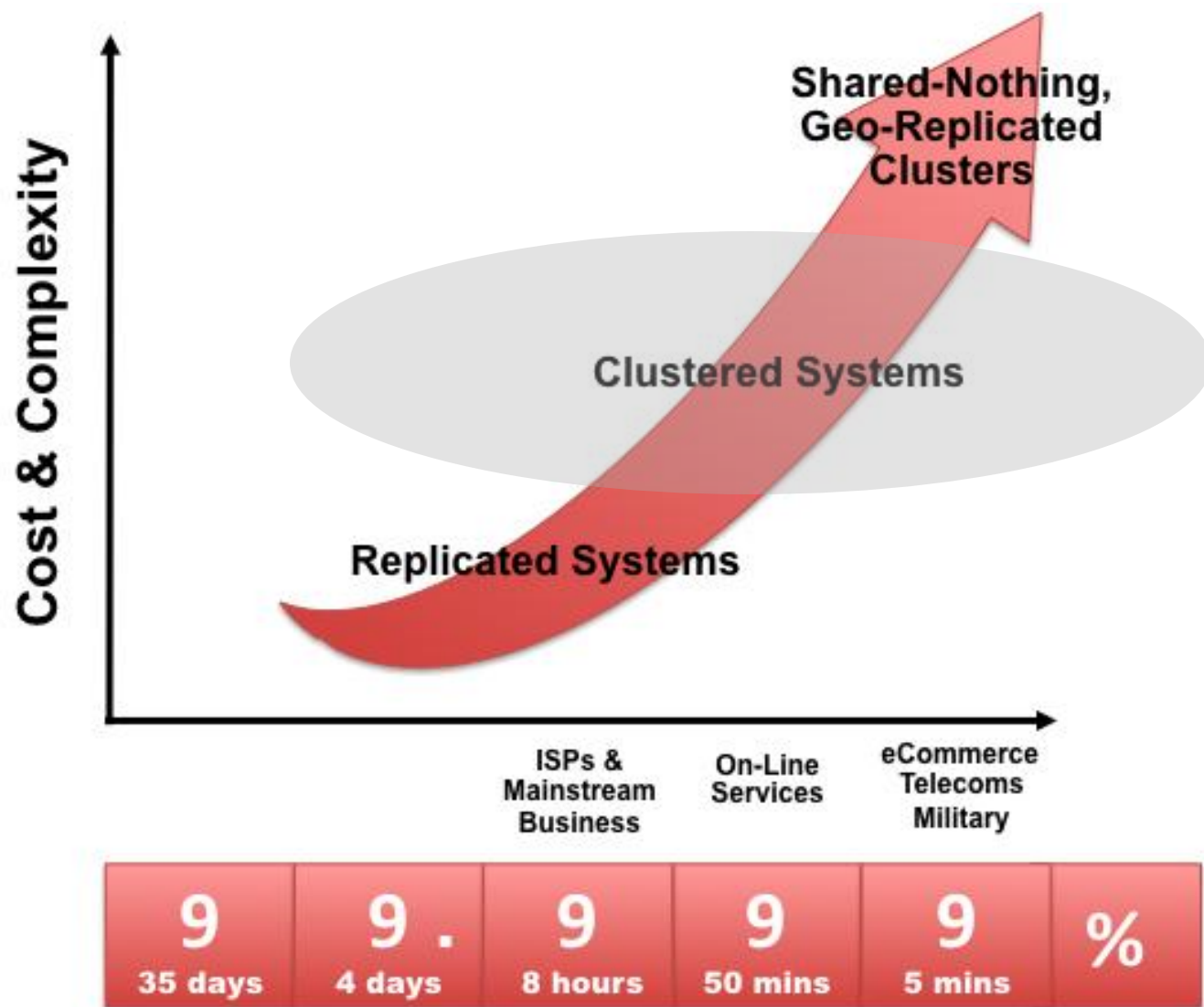
監視作業負荷の軽減:レプリケーションの監視と稼働統計をコマンド無しで



The screenshot shows the 'Replication Monitoring' tab in the MySQL Enterprise Monitor. It displays a table of replication configurations with columns for Servers, Type, Slave IO, Slave SQL, and Seconds Behind. The table is organized into expandable groups: Basic (2), Ringlet (2), RingSpoke (4), and Tree 3 (5). Each group contains details for master and slave nodes, their status (Running), and the time lag (Seconds Behind).

Servers	Type	Slave IO	Slave SQL	Seconds Behind
Basic (2)	TREE	Running	Running	
master:10101	master			
slave:10100	slave	Running	Running	00:00:00
Ringlet (2)	RING	Running	Running	
Yang:10120	master/slave	Running	Running	00:00:00
Yin:10121	master/slave	Running	Running	00:00:00
RingSpoke (4)	MIXED	Running	Running	
ring1:10183	master/slave	Running	Running	00:00:00
ring2:10182	master/slave	Running	Running	00:00:00
ring3:10181	master/slave	Running	Running	00:00:00
ring3slave:10180	slave	Running	Running	00:00:00
Tree 3 (5)	TREE	Running	Running	
master:10153	master			
slave1:10150	slave	Running	Running	00:00:00
slave2master:10152	master/slave	Running	Running	00:00:00
slave2slave:10151	slave	Running	Running	00:00:00
slave3:10154	slave	Running	Running	00:00:00

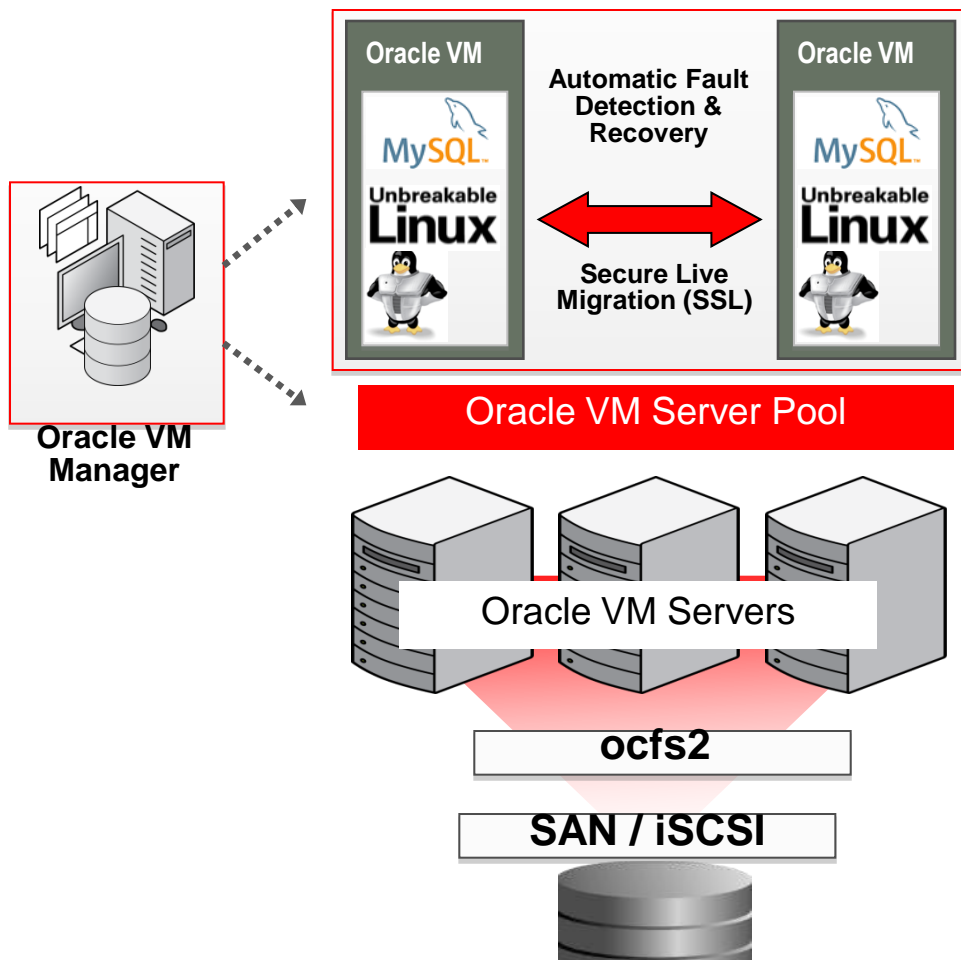
クラスタ・システム



Oracle VM Template for MySQL

設定済 & テスト済のOS, VM, Databaseスタック

- Oracle Linux
- Oracle VM
- Oracle VM Manager
- Oracle Cluster File System 2 (OCFS2)
- MySQL Database 5.5 (Enterprise Edition)



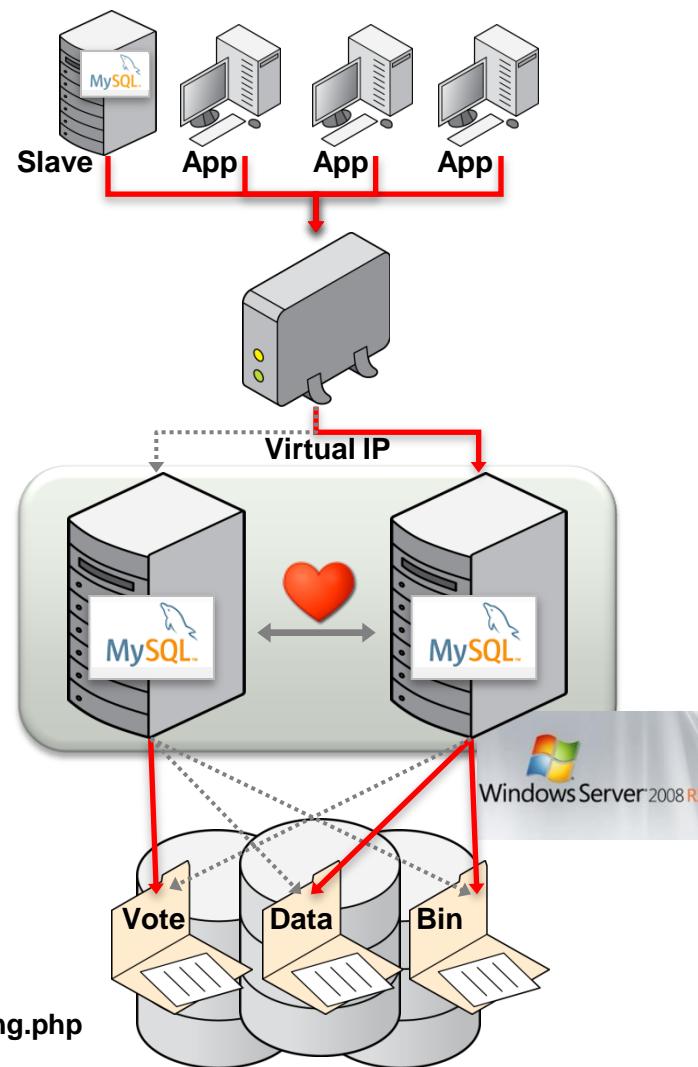
インストール&設定済み
製品統合 & テスト済み
サポートもワンストップ

http://mysql.com/why-mysql/white-papers/mysql_wp_oracle-vm-template-for-mee.php

ORACLE

Windows Server Failover Clustering

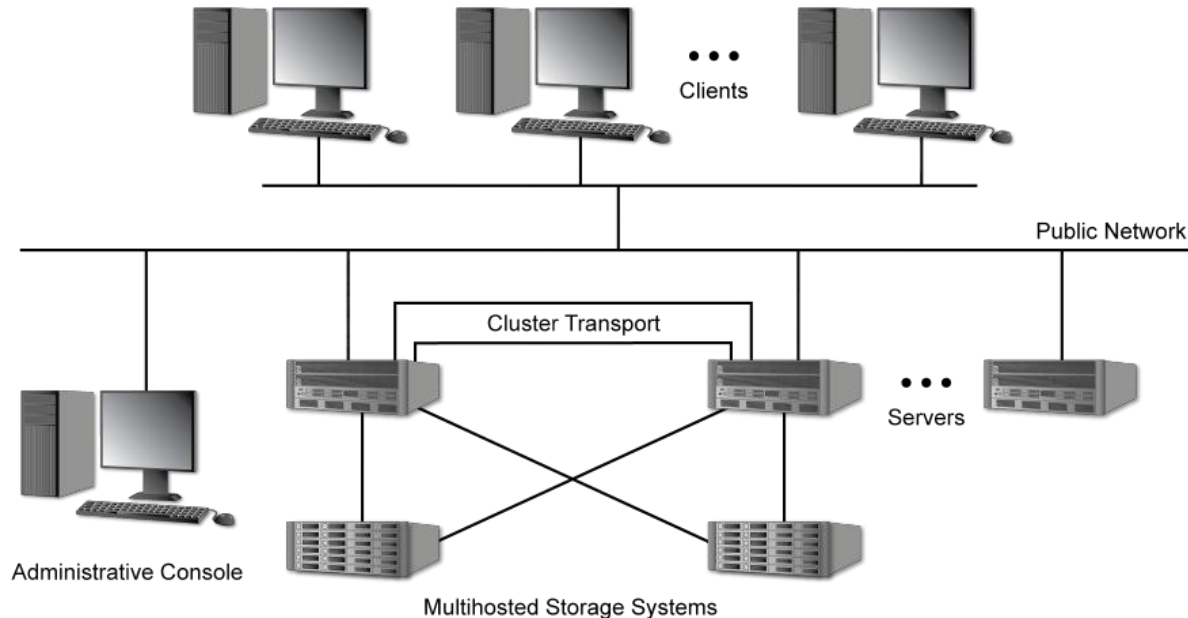
- Windowsサーバのクラスタリング機能を活用
- 障害対応や定期メンテナンス
- クォーラム、データ (InnoDB + スキーマ)、バイナリを共有ディスクに (iSCSI & FCAL)
- サービス停止時間
= 数秒 + InnoDBのリカバリ時間
 - アプリケーションは一時的に接続断
→ 同じIPアドレスに再接続
 - レプリケーションスレーブは自動的に復旧
- Failover Cluster ManagementのGUIからクラスタ環境を管理
- 追加ソフトウェアやスクリプト作成不要



ガイドをダウンロード: MySQL with Windows Server 2008 R2 Failover Clustering
http://www-jp.mysql.com/why-mysql/white-papers/mysql_wp_windows_failover_clustering.php

**** Windows Server Failover Clusteringのサポートは別途Microsoftからの購入が必要**

Solaris Clustering



- カーネル組み込みのハートビートと監視
- MySQLエージェントはSolaris Clusteringに組み込み
– レプリケーションの監視
- 詳しくはこちら: <http://www.oracle.com/technetwork/server-storage/solaris-cluster/overview/index.html>

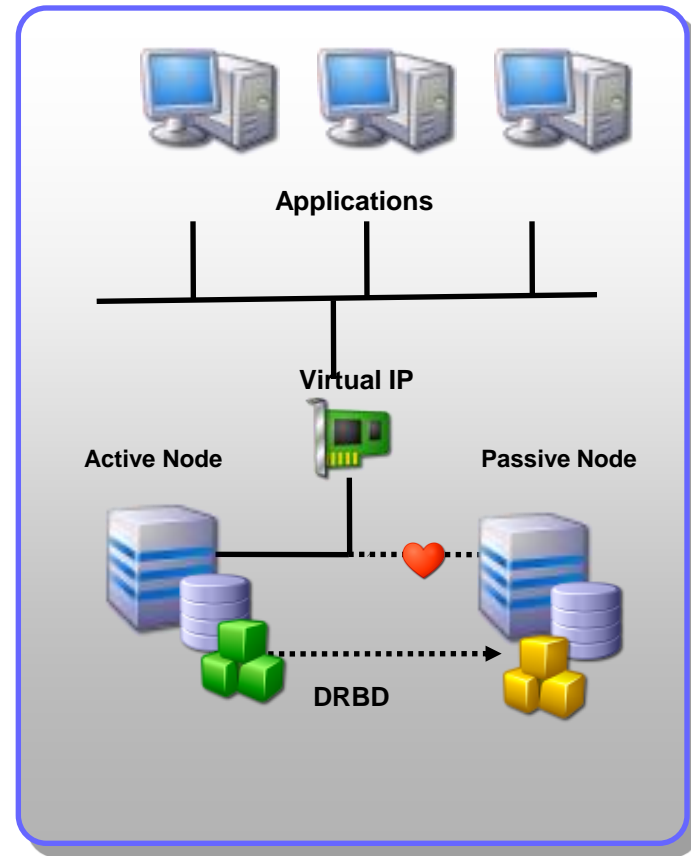
サードパーティ製のHAソリューション

- オラクルはOSSコミュニティやサードパーティの高可用性ソリューション上のMySQLをサポート
 - 高可用性ソリューションそのもののサポートは該当のベンダから
- Linux Heartbeat / Corosync with Pacemaker
 - DRBDによるカーネルレベルのブロックデバイスレプリケーション
- Symantec Veritas Cluster
- Red Hat Cluster Suite

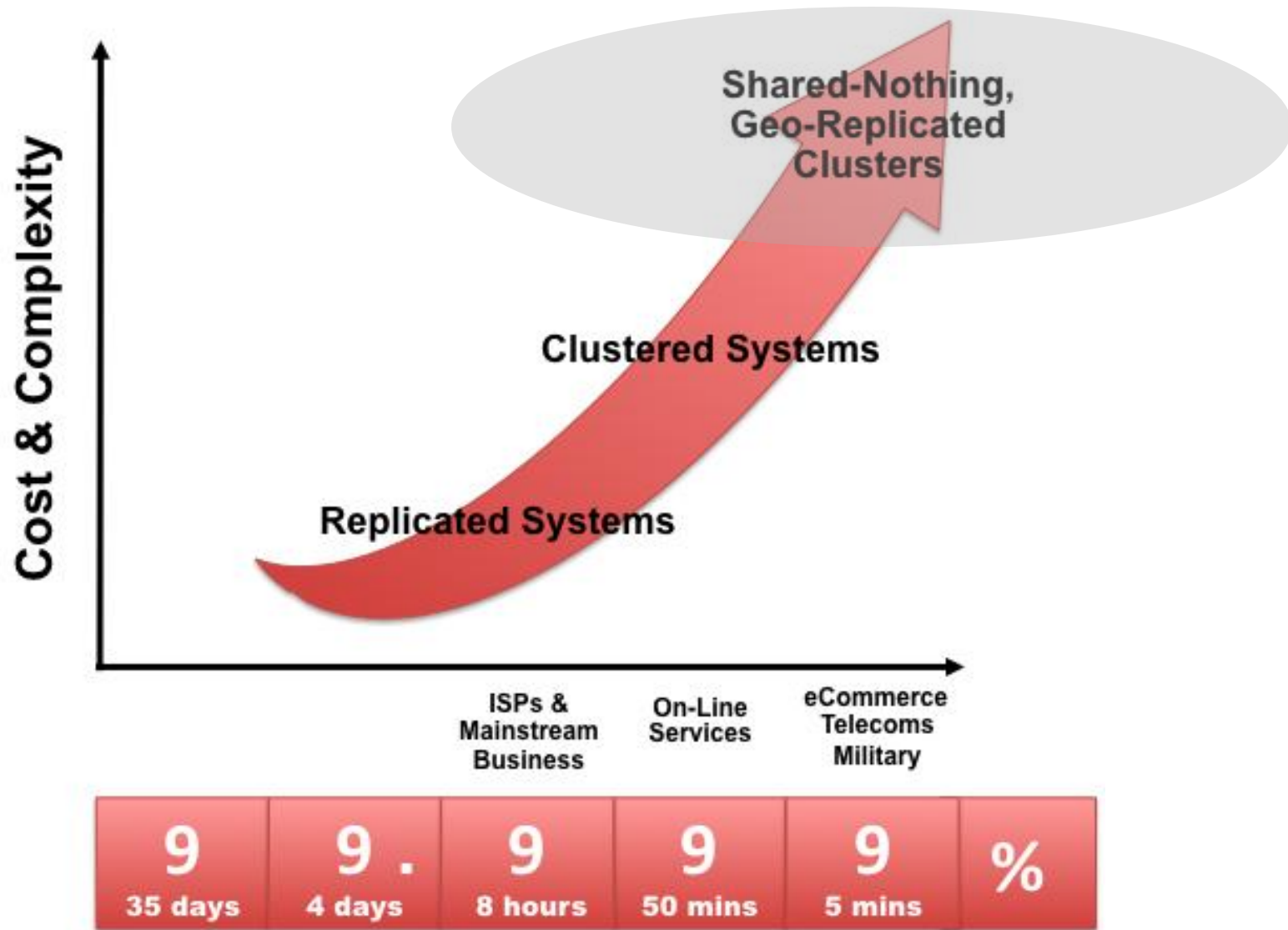


DRBD

- Distributed Replicated Block Device
 - 一般的なIPネットワーク上で動作
 - 分散ストレージ
 - ネットワークRAID
- 同期型
- 特徴
 - Linux上のみで利用可能
 - 特殊なネットワーク関連コンポーネント不要 (Heartbeatを除く)
 - 高いパフォーマンス (blocks vs. rows of data)
 - 障害発生時のデータの不整合を管理・解決
 - Linux Heartbeat が、フェイルオーバーおよび仮想 IP を管理



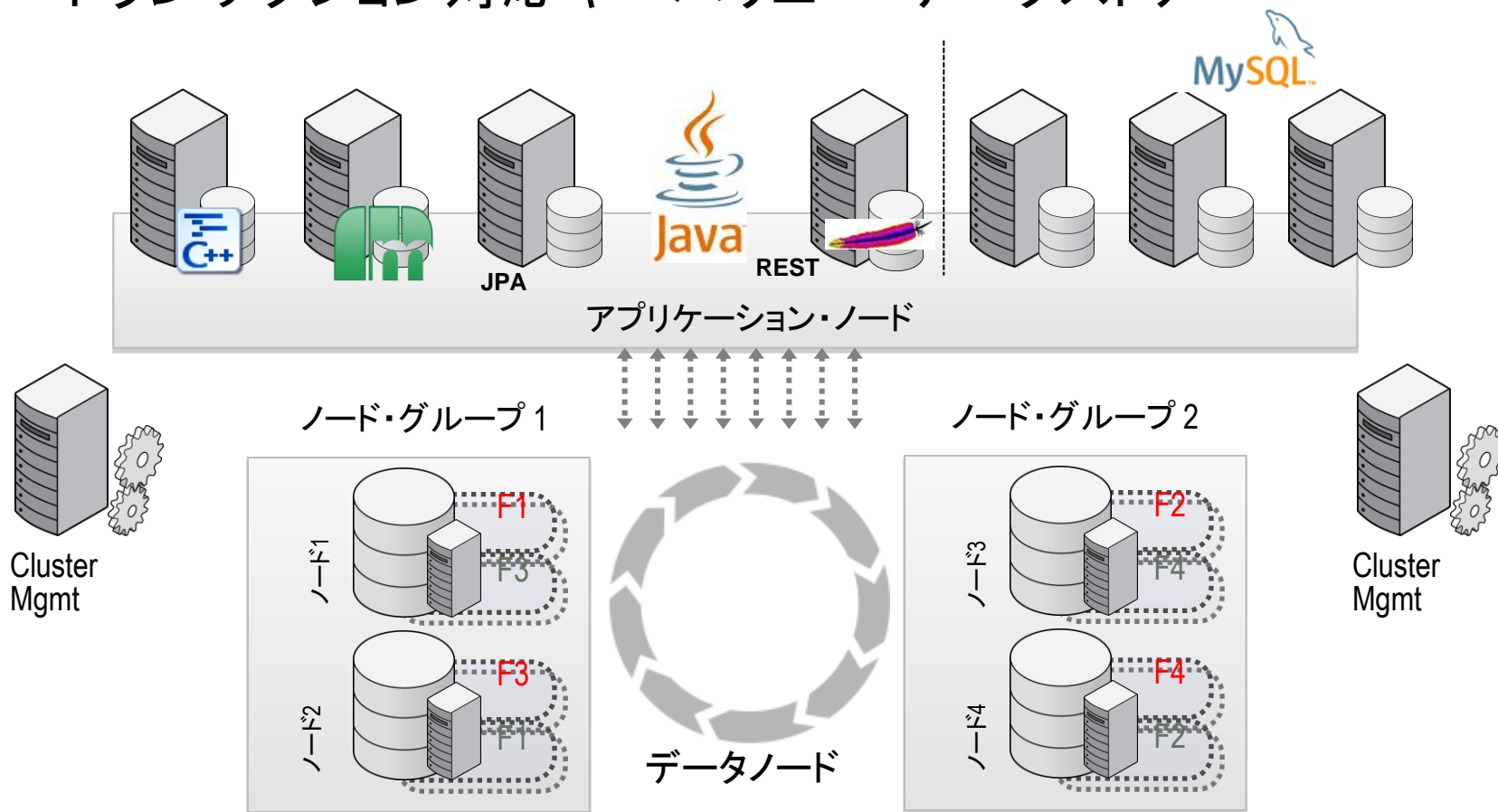
シェアードナッシング型クラスタ



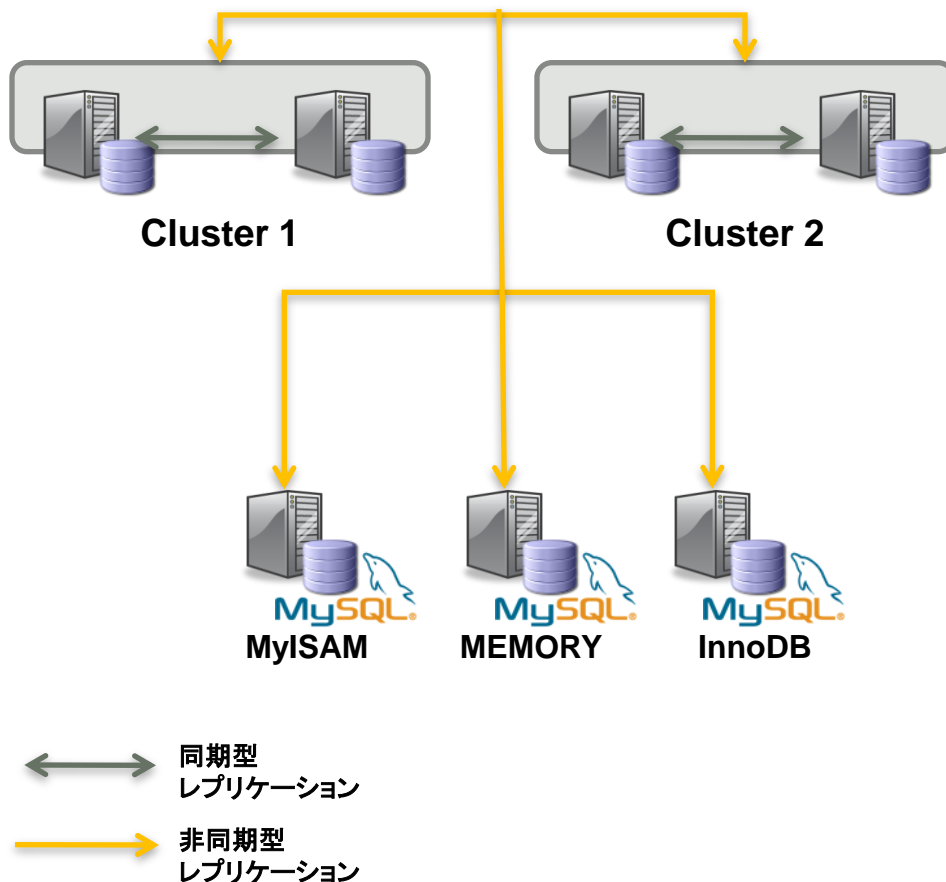
ORACLE

MySQL Cluster

- シェアードナッシング型Active-Activeデータベースクラスタ
- トランザクション対応キーバリュー・データストア



Geographical Replication - 地理的冗長性の確保



- クラスターのノードグループ間では、同期型レプリケーションで冗長性を確保
- 地理的に離れたクラスター間で、双方向の非同同期型レプリケーションを行い、地理的冗長性を確保
- (MySQL Clusterではない)通常のMySQLサーバへ非同同期型のレプリケーションを行い、レポート生成や課金処理などのアプリケーションを実行

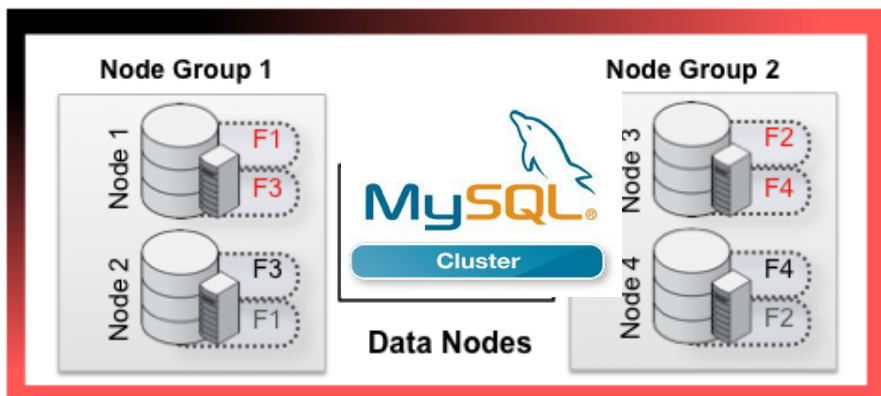
導入事例: 携帯電話ネットワーク



導入事例: 航空機管制システム



- 米国海軍航空母艦
- 包括的航空機運用管制システム
 - メンテナンス記録
 - 燃料搭載量管理
 - 気象状況
 - 飛行甲板管理
- システム要件
 - 単一障害点無し
 - 完全な冗長性
 - 小さなフットプリント & 過酷な環境での利用
- 4台のMySQL Clusterノード
LinuxおよびWindows

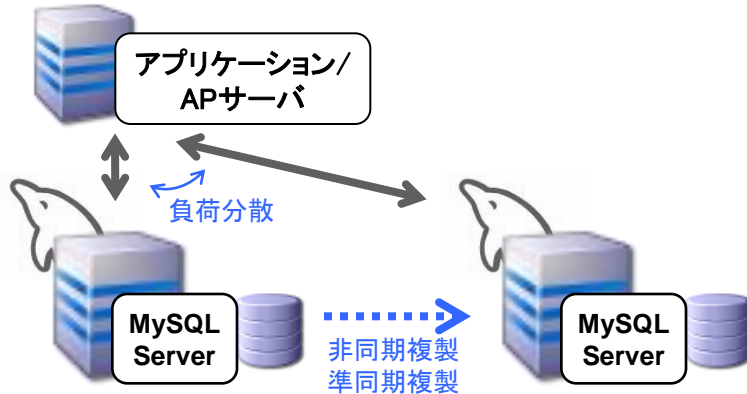


MySQL User Conference Session: <http://bit.ly/ogeid3>

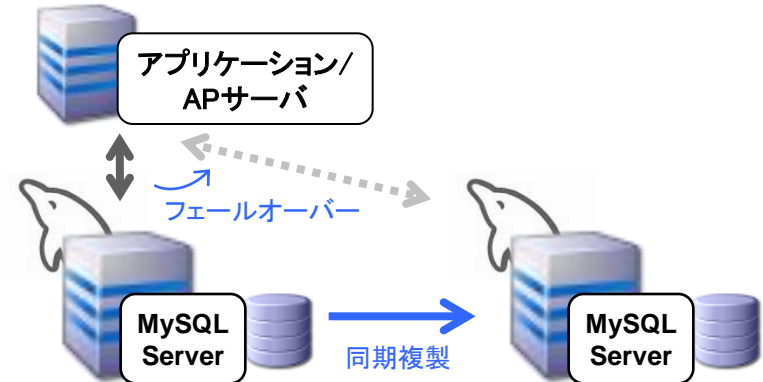
ORACLE

MySQLの高可用性構成

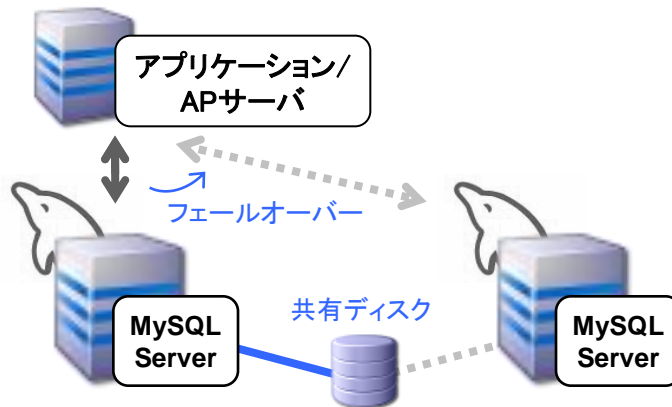
- レプリケーション(標準機能)
非同期&準同期データレプリケーション



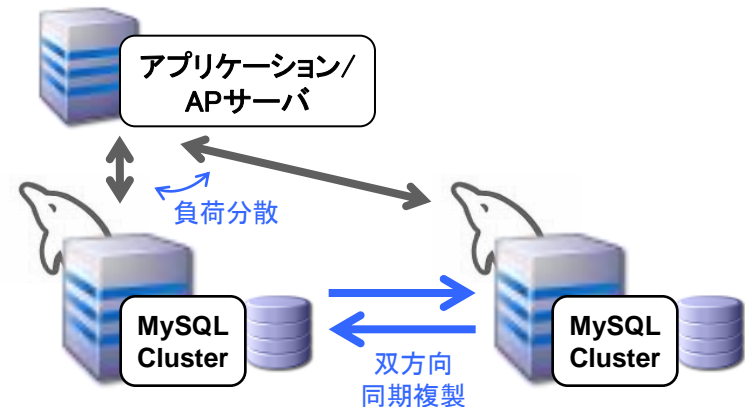
- MySQL+DRBD
Linux用のノード間データコピー



- 3rdベンダ製HAソフト利用
共有ディスクにデータを格納



- MySQL Cluster
シェアードナッシング型高性能クラスタ



Hardware and Software **Engineered to Work Together**

ORACLE®