

ORACLE

MySQL Cluster CGE 製品概要

高可用性インメモリデータベース、RDBMSとNoSQLのハイブリッド

日本オラクル株式会社

MySQL Global Business Unit

注意事項

「MySQL Cluster = MySQL NDB Cluster」です

- MySQL InnoDB Clusterとの混同することを避けるために、マニュアル等の表記が「MySQL Cluster」から「MySQL NDB Cluster」に変更されました
- 但し、「MySQL Cluster」の表記が使われる場面も残っています
 - 例) 製品名は「MySQL Cluster CGE(Carrier Grade Edition)」です

アジェンダ

1. MySQL NDB Clusterの特徴、事例
2. MySQL NDB Clusterのアーキテクチャー
3. MySQL NDB Clusterの性能
4. MySQL Cluster CGEの特徴



アジェンダ

1. MySQL NDB Clusterの特徴、事例
2. MySQL NDB Clusterのアーキテクチャー
3. MySQL NDB Clusterの性能
4. MySQL Cluster CGEの特徴



MySQL NDB Clusterの特徴

分散イン・メモリ・データ管理

リアルタイム・パフォーマンス、
自動パーティショニング、データ分散、
ビルトインのレプリケーション

ハイブリッド・データストア

キー・バリュ型 NoSQL
+ 複雑なリレーショナルな処理 RDBMS

高い拡張性と堅牢性の両立

マルチ・マスター、スケール・アウト可能、
テラバイト級データへの対応、
ACID準拠トランザクション



99.9999%の可用性

ミッション・クリティカル・システム向けの
設計、自律型、単一障害点のないシェアー
ド・ナッシング構成

一貫性のあるトランザクション

分散かつ分割配置されたデータセットに
対するトランザクションの一貫性

オープン・ソース

リアルタイム・パフォーマンス、
自動パーティショニング、データ分散、
ビルトインのレプリケーション

MySQL NDB Cluster導入メリット

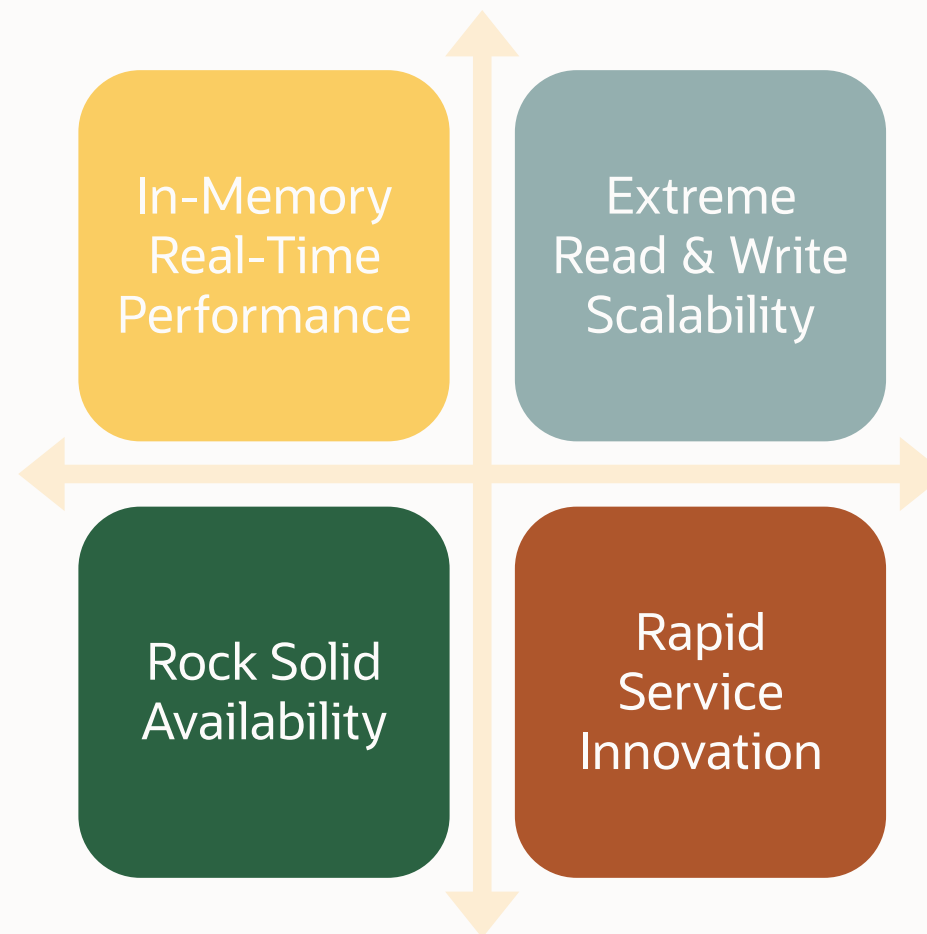
新しいタイプのデータベース利用要件に対応

インメモリデータベースとしての
リアルタイム性能と高い堅牢性の両立

お客様ビジネスの成長を支える
拡張性のあるデータ管理

RDBMSとNoSQLのハイブリッド構成による
柔軟性とビジネスの展開に対応したアジリティ

運用管理を自動化させたことによる
運用コストの低減



NewSQL:

NoSQLの拡張性と柔軟性を持ち、
SQL文やACIDトランザクションをサポートするデータストア

Aslett, Matthew (451 Group)

How Will The Database Incumbents Respond To NoSQL And NewSQL?

<http://cs.brown.edu/courses/cs227/archives/2012/papers/newsq1/aslett-newsq1.pdf>

MySQL "NDB" Clusterの歴史

MySQL Clusterの基礎となっている技術は、通信機器ベンダのエリクソンで携帯通信網の加入者データベース向けに開発されたEricsson Network Data Base(NDB)と呼ばれていた技術が起源

Original design paper:

[Design and Modelling of a Parallel Data Server for Telecom Application](#)

90年代後期
Ericsson (Alzato) にて設計/開発

2003年
MySQL ABが NDB Cluster Projectを
Ericssonから引き継ぎ開発 (MySQL 4.1)

2008年
SunがNDB Cluster Projectを引き継
ぎ開発(MySQL 5.1/Cluster 6.x～)

2010年
OracleがNDB Cluster Projectを引き継
ぎ開発(MySQL 5.1 / Cluster 7.x～)

MySQL NDB Clusterが向いているシステム

高可用性が求められるシステム

- 携帯電話の通信インフラを支えるために開発された技術がベース
- 単一障害点が無い構成
- 障害発生時に、アプリケーションは処理をリトライすれば存続したノードで処理を継続可能

同時多発的に大量のトランザクションが発生するシステム

- サーバー台数を増加することで負荷分散可能

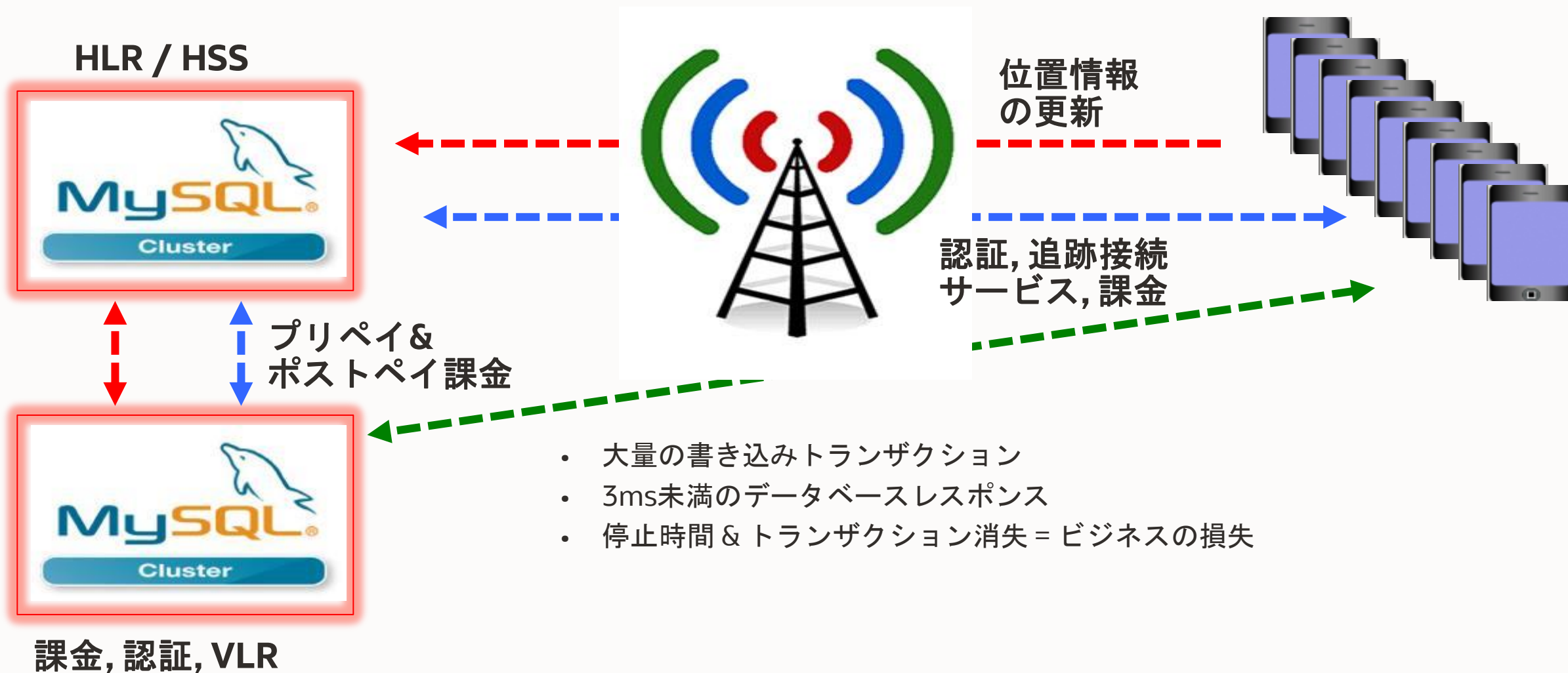
参照処理だけでなく、更新処理に対しても拡張性が求められるシステム

- 自動的にデータをシャーディングし、更新処理についても負荷分散可能

MySQL NDB Clusterを導入済みのお客様



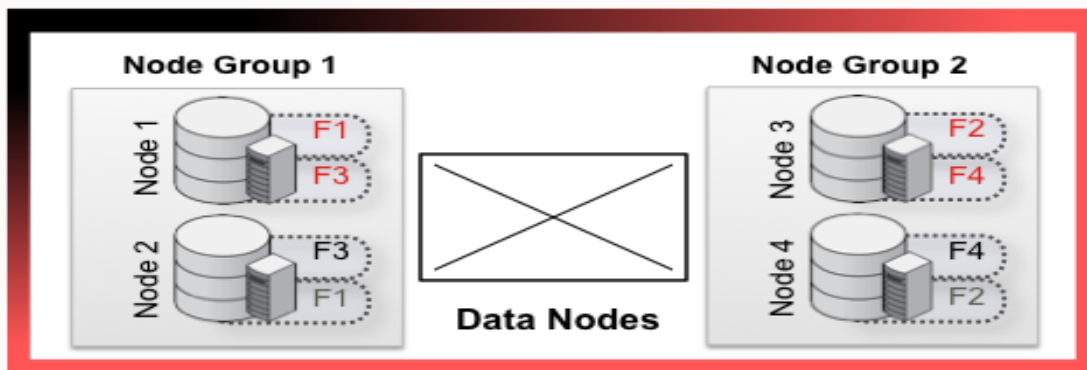
導入事例: 携帯電話ネットワーク



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 & 過酷な利用環境での利用
- 4台のMySQL Clusterノード LinuxおよびWindows



アジェンダ

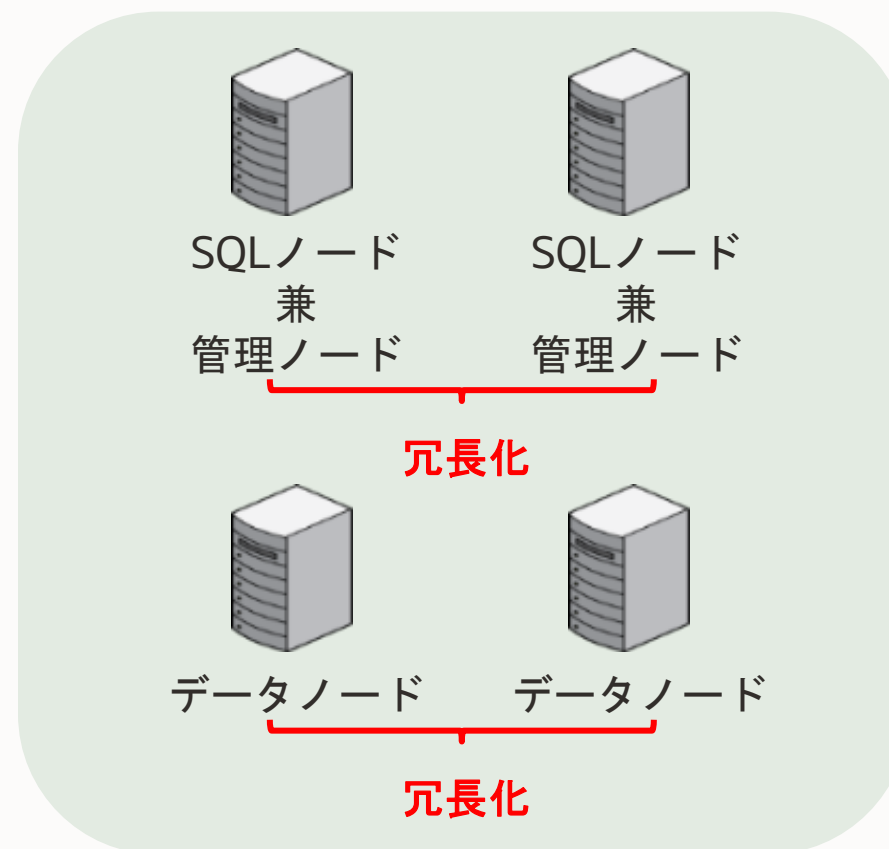
1. MySQL NDB Clusterの特徴、事例
2. MySQL NDB Clusterのアーキテクチャー
3. MySQL NDB Clusterの性能
4. MySQL Cluster CGEの特徴



MySQL NDB Clusterは複数台のサーバーでDBを構成する(推奨最低構成は4台)



MySQL Server



MySQL NDB Cluster



MySQL Clusterを構成する3種類のノード

データノード

- データ、インデックスを管理し、トランザクションを制御するノード
- データのレプリカを別のノードで持つことで冗長性を確保
- MySQL Clusterの中核となるノード

アプリケーションノード（SQLノード）

- アプリケーションとデータノードをつなぐインターフェースとなるノード
- SQLで処理する場合は、MySQL Serverが稼働する
- NoSQLで処理する場合は、MySQL Serverは不要

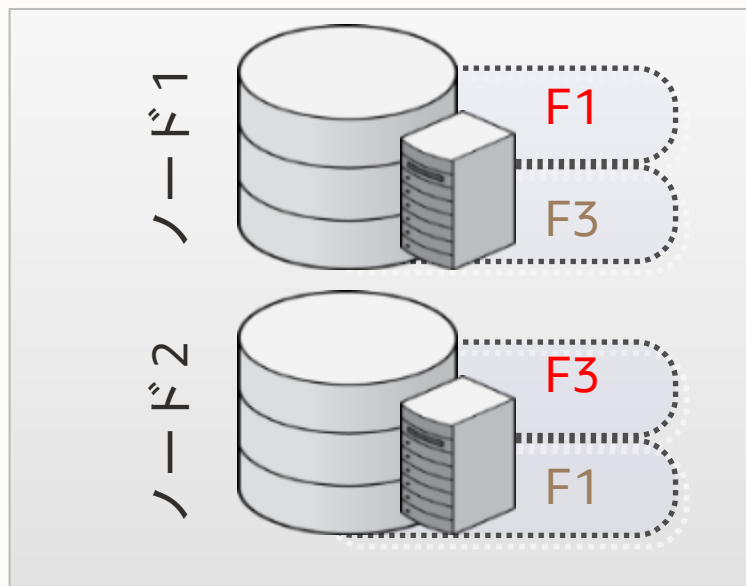
管理ノード

- MySQL Cluster全体を管理するノード
- 各ノードの設定管理、クラスタ全体の起動/停止、オンラインバックアップ、など

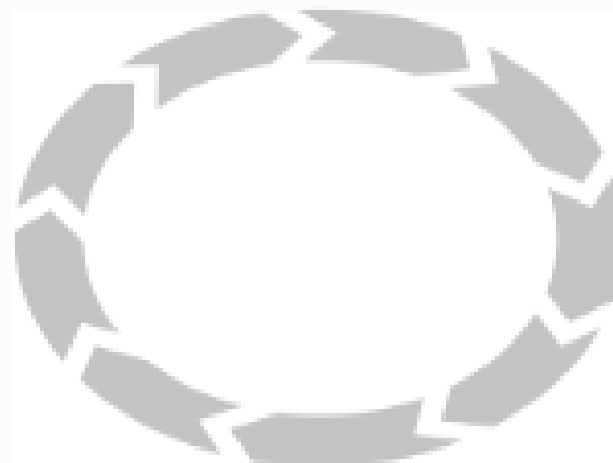
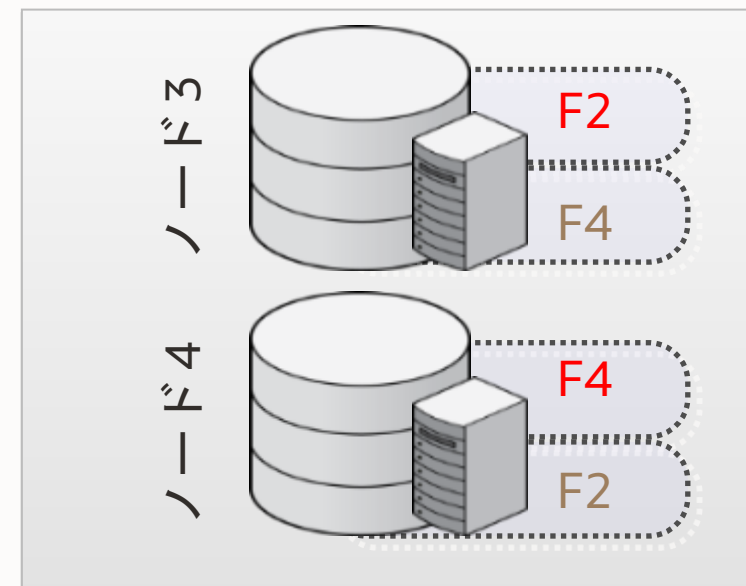
データノードでのデータ保持方法

データ・シャーディング

ノード・グループ1



ノード・グループ2

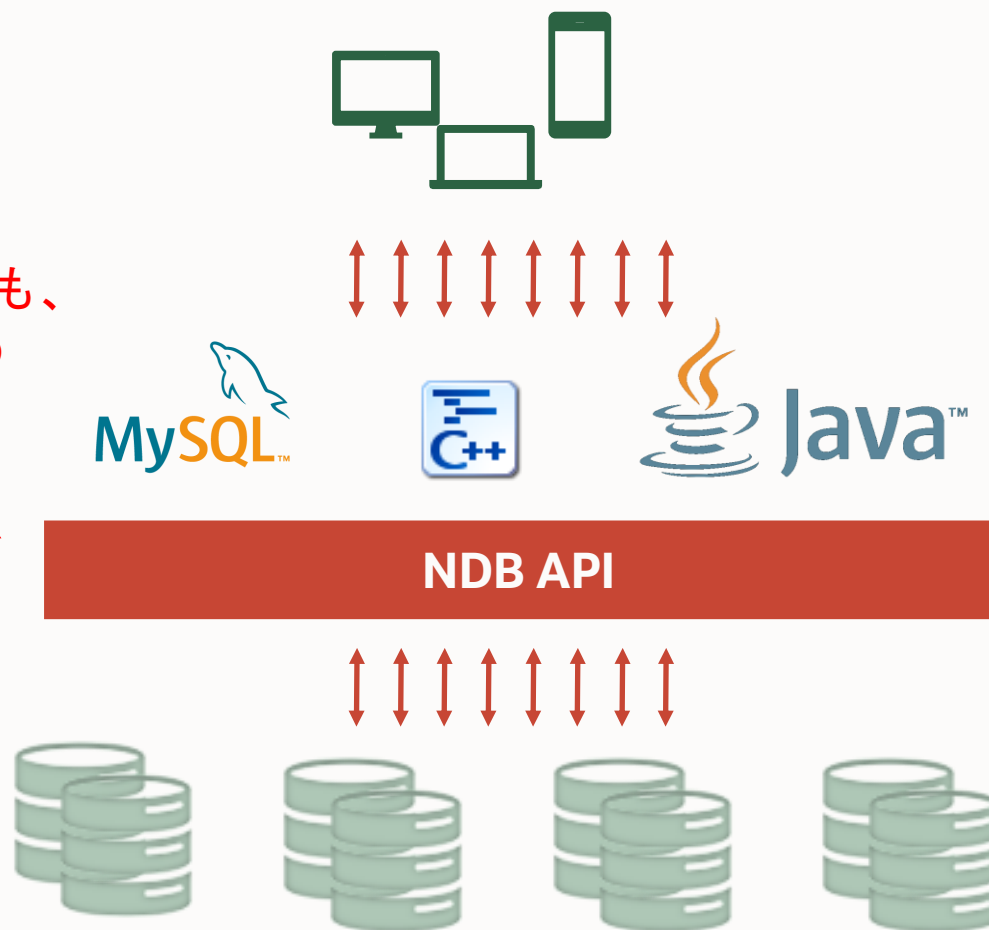
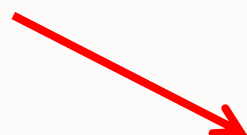


データ・ノード

- ・ 主キーのハッシュ値で水平分割して、各データノードにデータを保持
⇒ 負荷分散による性能向上（台数追加によりスケールアウト可能）
- ・ レプリカを別のデータノードと持ち合う
⇒ 冗長化による耐障害性

データノードへのアクセスは、NDB API経由で行われる

SQLで処理する場合でも、
その他のAPI経由で処理する場合でも、
バックグラウンドでNDB API経由の
処理に変換されている



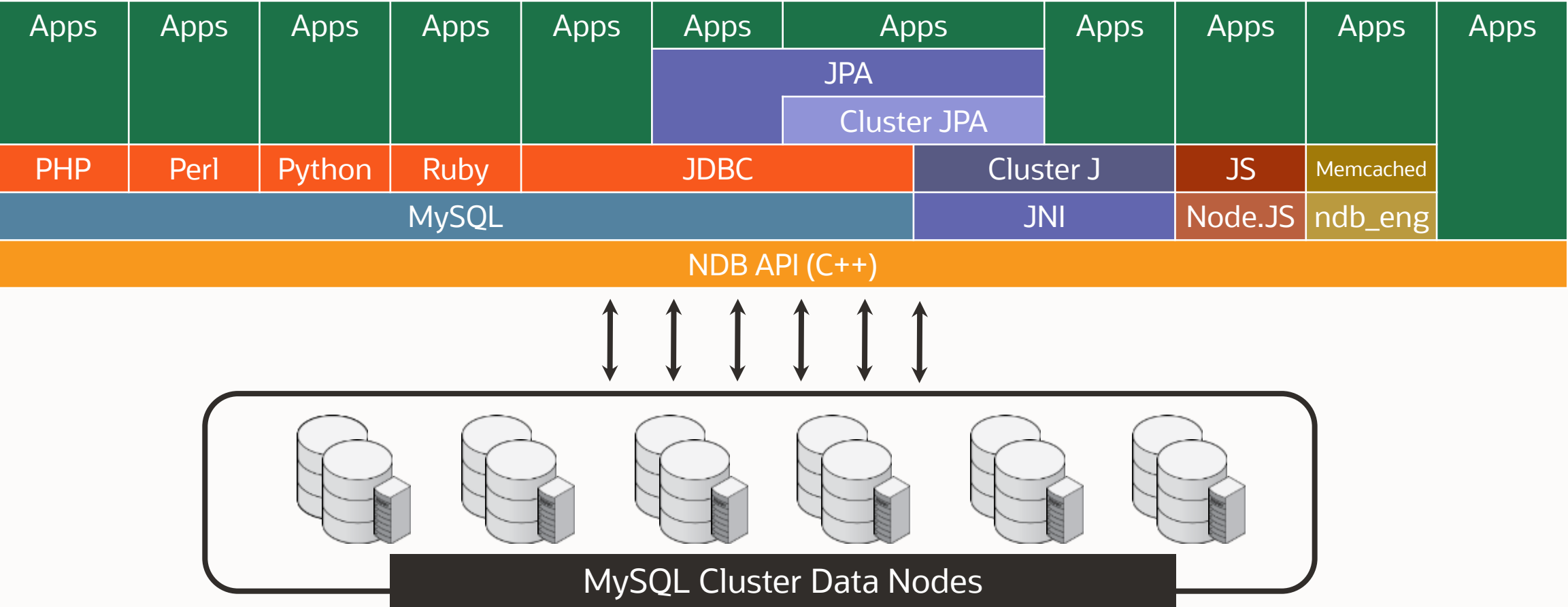
Clients

MySQL Server &
Custom Applications

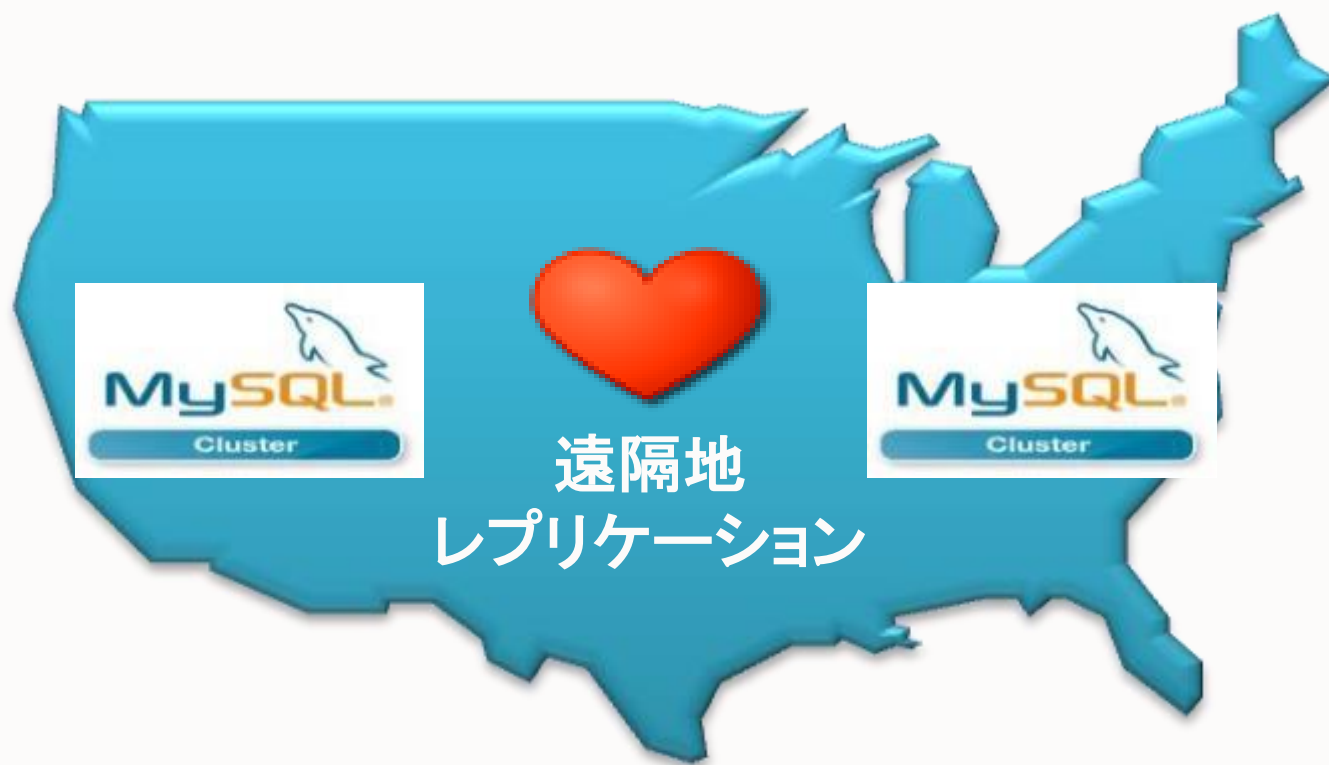
Common NDB API

Data Nodes

データノードへのアクセスイメージ

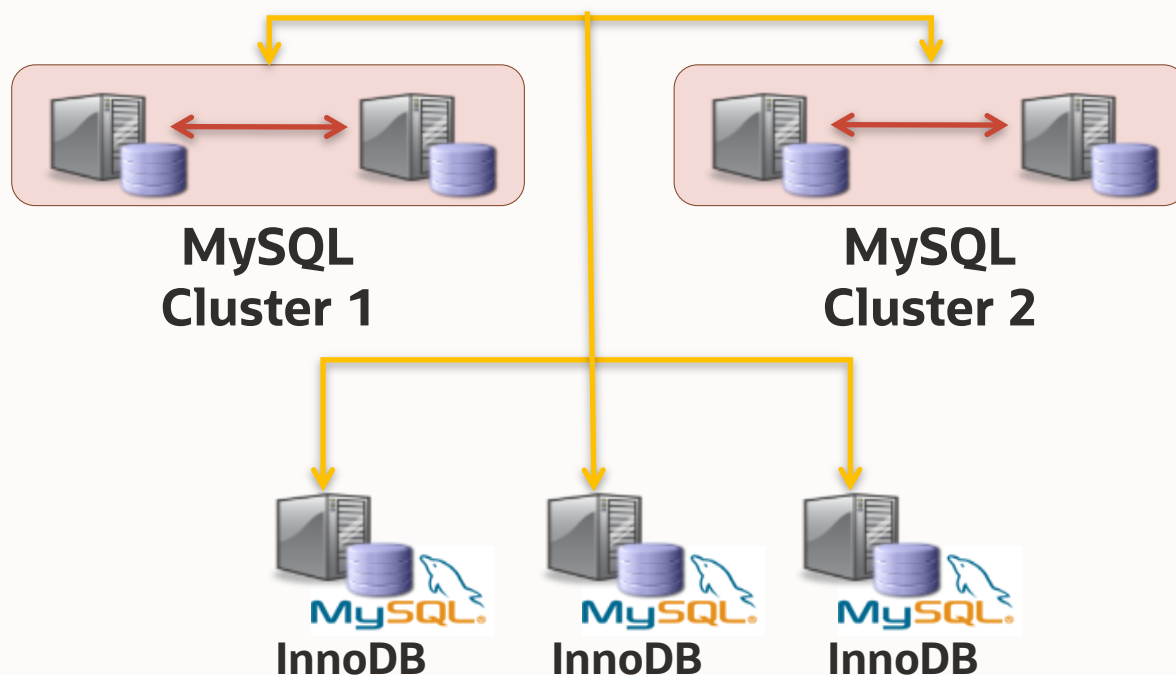


遠隔地レプリケーション



- データセンター間でクラスタを完全に複製
- DR & データの局所性
パッシブ・リソースなし
- アクティブ / アクティブ
レプリケーションの簡素化
- グローバルにスケーラブルで冗長性をもつサービスの運用がよりシンプルに

MySQL Serverも組み合わせた複合型の構成



←→ 同期型
レプリケーション

←→ 非同期型
レプリケーション

- 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、**サイト内の冗長性を確保**
- 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
- NoSQL経由の処理もレプリケーション可能

アジェンダ

1. MySQL NDB Clusterの特徴、事例
2. MySQL NDB Clusterのアーキテクチャー
3. MySQL NDB Clusterの性能
4. MySQL Cluster CGEの特徴



YCSBベンチマーク

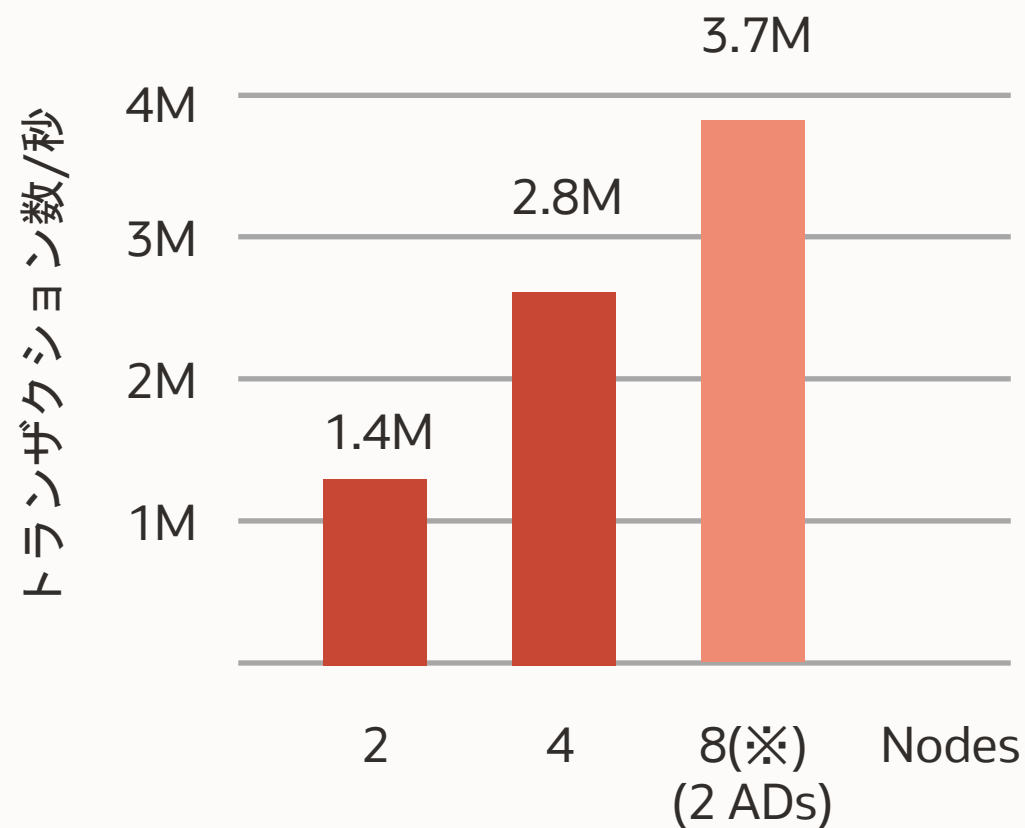
他のプロダクトよりも高いスループットを実現

- YCSB : Yahoo Cloud Serving Benchmark
- Yahooが開発したNoSQL(KVS)やクラウドサービスストア向けのベンチマークツール
- NoSQL(KVS)データベース、(耐久性の無い)インメモリデータグリッド、スケールアウトデータベースの比較に広く使用されている
- NoSQLワークロードで比較した際の前提 :
 - Workload A: 50% 読込み, 50% 書込み
 - YCSBクライアントの変更不可
 - DBベンダーが実装したJDBCクライアントを使用
 - バージョンと正確な設定が重要

Product	Nodes	TPS/OPS
 cassandra	32	227k
 mongodb	2	275k
 SCYLLA	3	715k
 VOLTDB	6	1.6M
 AEROSPIKE	8	1.6M
 MySQL Cluster	4	2.8M

YCSBベンチマーク

データノード数増加により、スループットをほぼリニアに向上



[検証環境]

YCSB 0.15.0 with JDBC / SQL

- レコード数1kB (デフォルト)
- requestdistribution: uniform (均一分布)
- データノード数 : 2、4、8
- NoOfReplicas : 2
- strong consistency
- ACID (read committed)

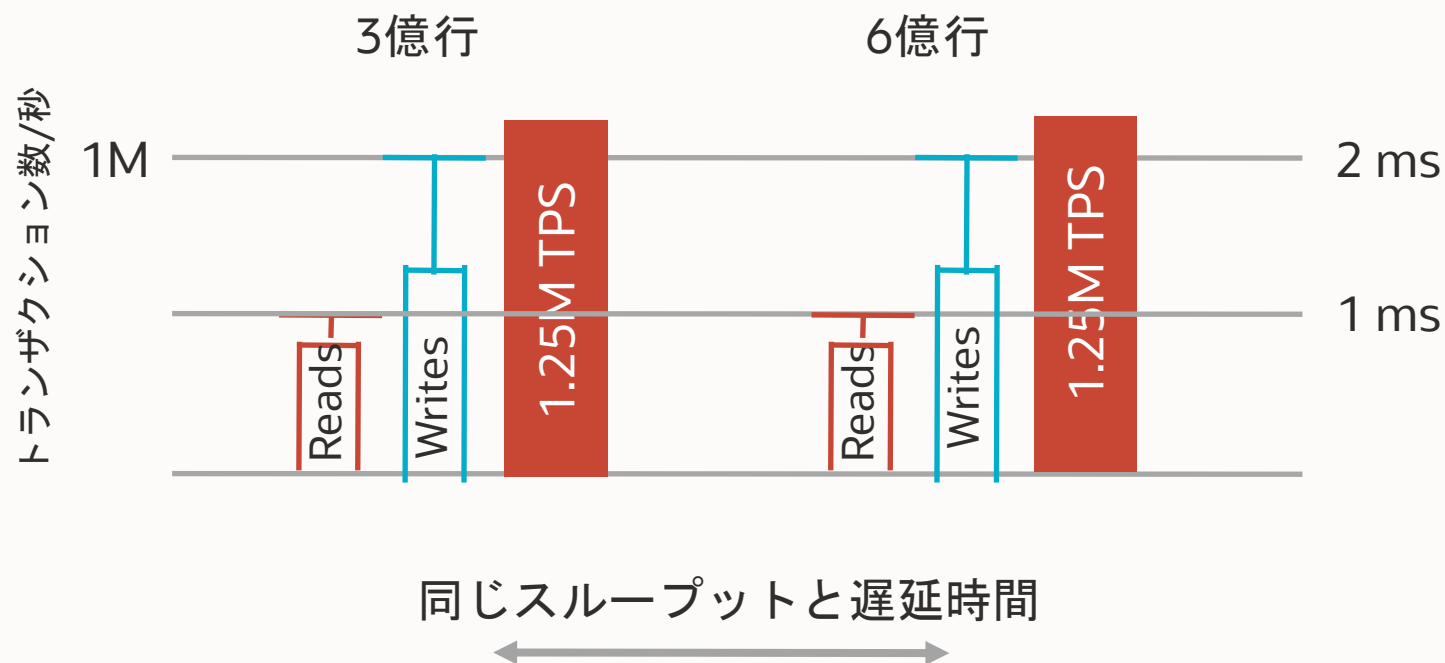
※8ノードでのベンチマーク

DenseIOシェイプを4ノードずつ、
2つのADにまたがって配置

- AD間のネットワークレイテンシは約400us

YCSBベンチマーク

行数が増加しても低遅延のまま



[検証環境]

- ・ 4 データノード
- ・ データ量 : 3億行、6億行
- ・ JDBC使用

[結果]

99% SQL reads < 1ms

- ・ 95% < 0.9ms

99% SQL writes < 2ms

- ・ 95% < 1.7ms

アジェンダ

1. MySQL NDB Clusterの特徴、事例
2. MySQL NDB Clusterのアーキテクチャー
3. MySQL NDB Clusterの性能
4. MySQL Cluster CGEの特徴



MySQL Cluster Carrier Grade Edition

年間サブスクリプション 120万円 (1サーバーあたり)

拡張機能

- 性能拡張性
- 高可用性
- 統合認証
- 監査
- 非対称暗号化
- ファイヤーウォール

管理ツール

- 一括監視
- バックアップ
- 運用効率向上支援
- 開発
- 管理
- アップグレード

サポート

- 技術サポート
- コンサルティングサポート
- オラクル製品との動作保証

**MySQL NDB Clusterの運用効率を向上し
TCO効率化・削減に貢献**

【MySQL NDB Cluster運用にあたっての潜在的な課題】

分散型環境の効率的な監視

- 分散型データベースのため、複数のサーバーが監視対象となる
- インメモリデータベースのため、メモリ利用状況の監視が重要となる

→ MySQL Enterprise Monitor

構成変更などの管理負荷

- 複数のサーバーで構成された環境の変更には多くのコマンドの実行が必要となり、人為的ミスの発生がリスクとなる

→ MySQL Cluster Manager

運用のノウハウや障害解析

- バージョンアップやパフォーマンスチューニング、またトラブルシューティングの際は、MySQLサーバーや他のRDBMSとは異なったアプローチが求められるケースがある

→ Oracle Premier Support for MySQL



【MySQL NDB Cluster運用にあたっての潜在的な課題】

分散型環境の効率的な監視

MySQL NDB Clusterは分散型データベースのため、SQLノード、データノードの複数のプロセスにより構成される

【潜在的な課題】

- 複数のプロセスの動作状況の効率的な監視
- 実行されたSQL文などの性能監視

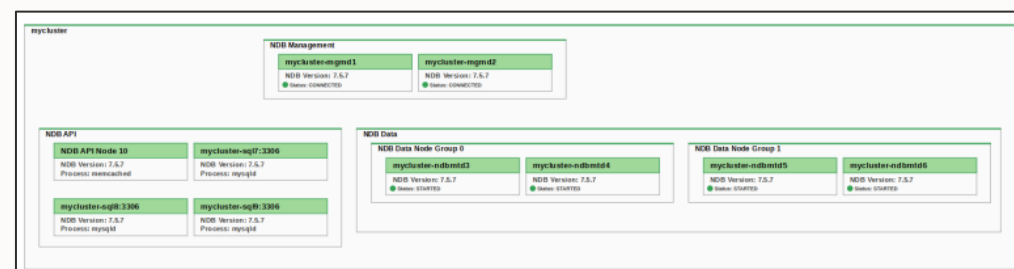
MySQL NDB Clusterインメモリデータベースのため、メモリ利用状況の監視が重要

【潜在的な課題】

- バージョン間の監視対象項目の変化への対応
- 監視スクリプトのメンテナンス負荷

【ご提案】

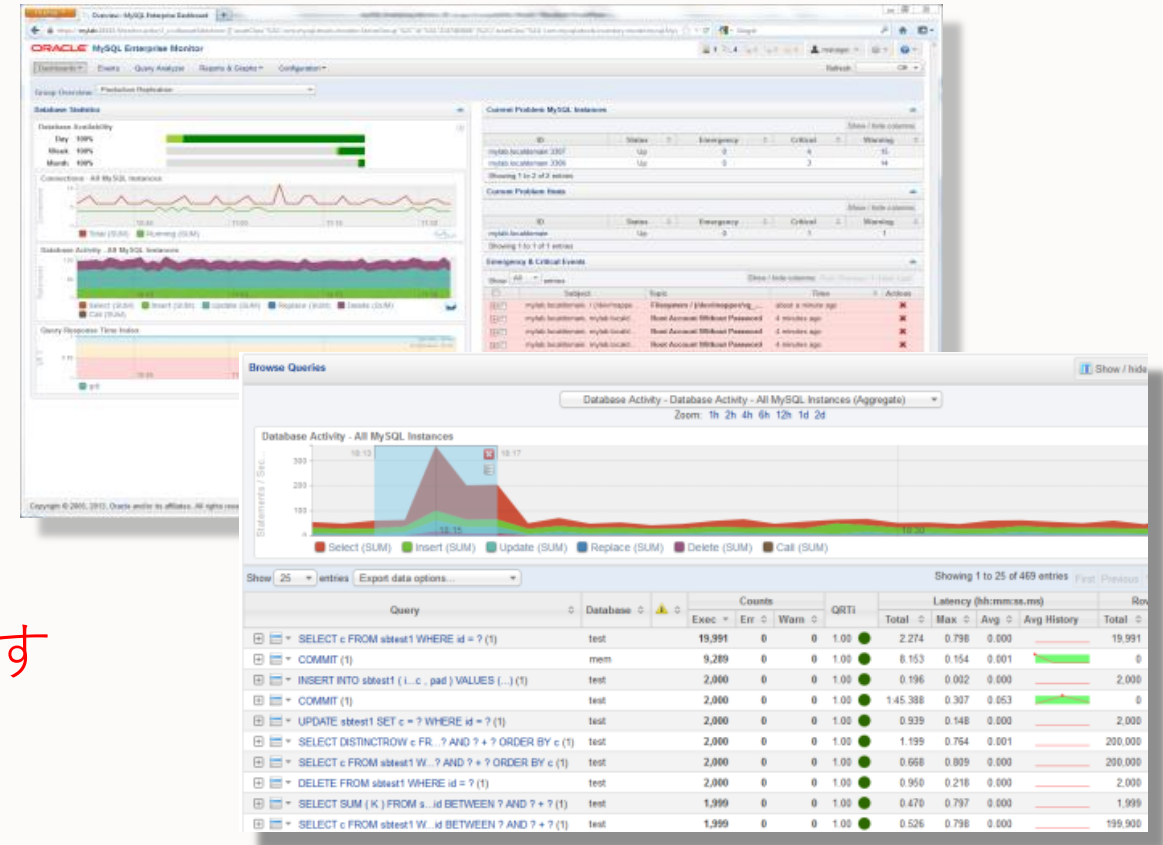
MySQL Enterprise Monitorによる集中監視



MySQL Enterprise Monitor

Enterprise Edition / MySQL Cluster Carrier Grade Edition

- 複数のMySQLサーバも一括監視
- ルールに基づく監視と警告
- 開発部隊とサポート部隊のノウハウを集約
- 問題が発生する前に通知
- MySQL Query Analyzer
 - SQL文を一括監視
 - MySQLの設定変更不要
 - チューニングのための解析作業を省力化



属人化を排除し、最適な環境を提供します

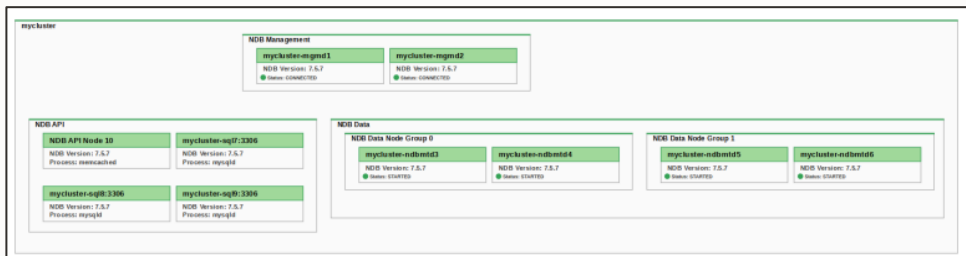


MySQL Enterprise Monitor

MySQL NDB Cluster関連項目

Topology Viewによる構成監視

- SQLノード(APIノード)、データノード、管理ノードの構成と稼働状況を自動的に可視化



<https://dev.mysql.com/doc/mysql-monitor/8.0/en/mem-ui-topology.html>

NDB Cluster Advisorsによる運用支援

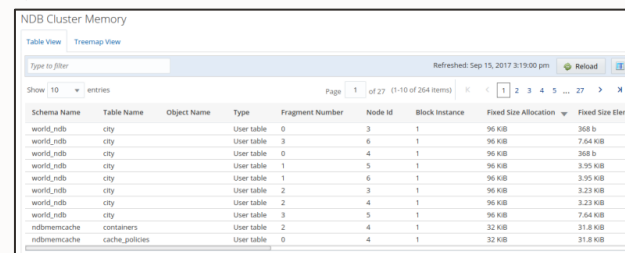
- データノードのコンポーネント別のメモリ空き容量監視と警告

<https://dev.mysql.com/doc/mysql-monitor/8.0/en/cluster-advisors-ref.html>

NDBメモリ利用状況レポート

ndbinfo.memory_per_fragmentのデータを元にテーブルごとのメモリ利用状況を可視化

- テーブルビュー



Schema Name	Table Name	Object Name	Type	Fragment Number	Node Id	Block Instance	Fixed Size Allocation	Fixed Size Element
world_ndb	city	User table	0	3	1	96 KB	368 b	
world_ndb	city	User table	3	6	1	96 KB	7.64 KB	
world_ndb	city	User table	0	4	1	96 KB	368 b	
world_ndb	city	User table	1	5	1	96 KB	3.95 KB	
world_ndb	city	User table	1	6	1	96 KB	3.95 KB	
world_ndb	city	User table	2	3	1	96 KB	3.23 KB	
world_ndb	city	User table	2	4	1	96 KB	3.23 KB	
world_ndb	city	User table	3	5	1	96 KB	7.64 KB	
ndbmemcache	containers	User table	2	4	1	32 KB	31.8 KB	
ndbmemcache	cache_policies	User table	0	4	1	32 KB	31.8 KB	

- ツリービュー



<https://dev.mysql.com/doc/mysql-monitor/8.0/en/mem-reports-ndb-cluster-memory-usage.html>



MySQL Enterprise Monitor

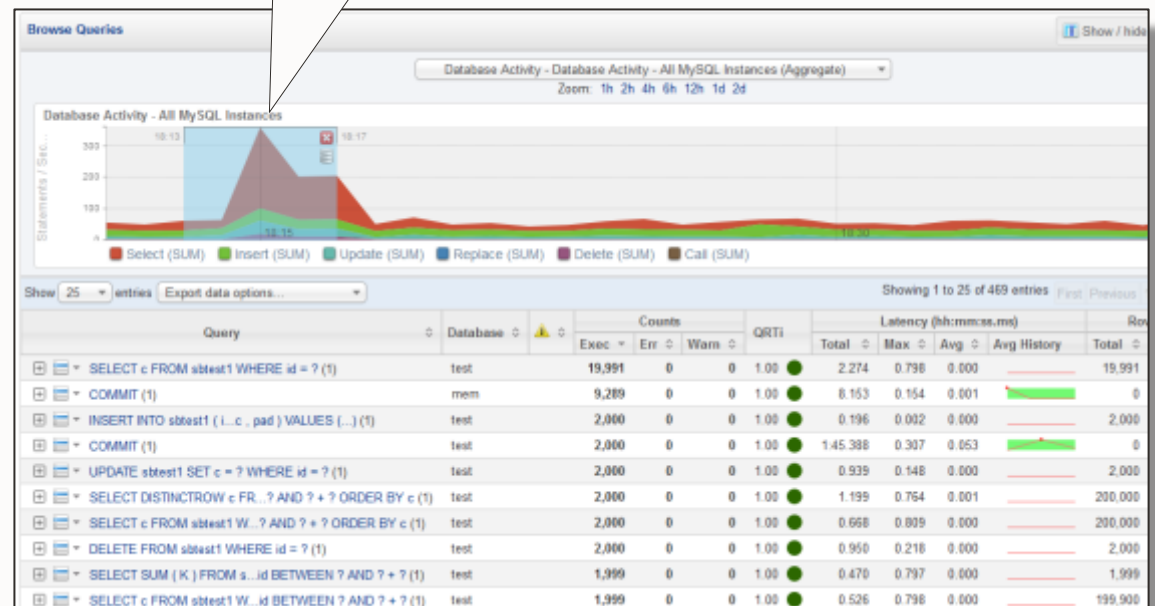
MySQL Query Analyzer

- 全てのMySQLサーバの
全てのSQL文を一括監視
- vmstatなどのOSコマンドやMySQLの
SHOWコマンドの実行、
ログファイルの個別の監視は不要
- クエリの実行回数、エラー回数、
実行時間、転送データ量などを
一覧表示
- チューニングのための
解析作業を省力化

Temp Tables		Sorting		初回実行
Disk	Disk %	Merge Passes		
1	100	0		22:59:26
0	0	0		2016/04/14 21:08:33
0	0	0		2016/04/14 21:08:34
0	0	0		2016/04/14 21:07:57

負荷の高い
処理を特定

早期対応と改善



【MySQL NDB Cluster運用にあたっての潜在的な課題】

構成変更やバージョンアップ

MySQL NDB Clusterは分散型データベースのため、SQLノード、データノードの複数のプロセスにより構成される

【潜在的な課題】

- バージョンアップの際に多数のコマンドの実行を手動で実行することとなりえる
- 操作ミスなどによる手戻りの発生

【ご提案】

MySQL Cluster Managerによる構成管理

```
mcm> show status -r newcluster;
+-----+-----+-----+-----+-----+-----+
| NodeId | Process | Host           | Status | Nodegroup | Package |
+-----+-----+-----+-----+-----+-----+
| 50     | ndb_mgmd | 198.51.100.102 | import |           | newpackage |
| 2      | ndbd     | 198.51.100.103 | import | n/a       | newpackage |
| 3      | ndbd     | 198.51.100.104 | import | n/a       | newpackage |
+-----+-----+-----+-----+-----+-----+
3 rows in set (0.05 sec)
```

```
mcm> upgrade cluster --package=7.6.13 mycluster;
+-----+-----+
| Command result |
+-----+-----+
| Cluster upgraded successfully |
+-----+-----+
1 row in set (3 min 17.00 sec)
```



```
mcm> autotune --writeload=high realtime mycluster;
+-----+-----+
| Command result |
+-----+-----+
| Cluster successfully autotuned to template realtime |
+-----+-----+
1 row in set (2 min 58.09 sec)
```

MySQL Cluster Manager

DevOpsの柔軟性と効率の向上、停止時間の削減



自動管理

- 個別のノードやクラスタ全体の起動/停止
- オンラインでのノード追加
- オンラインでの設定変更
- オンラインでのアップグレード
- オンラインバックアップ & リカバリ
- 既存のクラスタのインポート

Self-Healing

- ノード監視
- SQLノードや管理ノードを含めた自動リカバリ

高可用性の管理

- クラスタ全体の設定の一貫性を管理
- 設定の永続化
- HA エージェント

MySQL Cluster Managerの導入効果

MySQL Cluster アップグレード



MySQL Cluster Manager 導入前

- 1 x クラスタの事前状態チェック
- 8 x ssh コマンド/サーバー
- 8 x 停止コマンド/プロセス
- 4 x 構成ファイルの転送 - scp
(2 x mgmd & 2 x mysqld)
- 8 x プロセスごとの開始コマンド
- 8 x 開始または再参加プロセスの確認
- 8 x 完了確認処理
- 1 x クラスタ全体の完了確認
- 各構成ファイルの手動による編集を除く

合計: 46 コマンド -
2.5 時間の作業

MySQL Cluster Manager 導入後

```
mcm> upgrade cluster --package=8.0.28  
mycluster;
```

合計: 1 コマンド -
完全自動処理

MCMコマンド例



```
mcm> START CLUSTER mycluster;  
mcm> BACKUP CLUSTER mycluster;  
mcm> LIST BACKUPS mycluster;  
mcm> RESTORE CLUSTER --backupid 2 mycluster;  
mcm> UPGRADE CLUSTER --package=8.0.28 mycluster;
```

参照 : <https://dev.mysql.com/doc/mysql-cluster-manager/8.0/en/>

MCMコマンド例 (Data Node追加)



```
mcm> add hosts --hosts=192.168.0.14,192.168.0.15 mysite;  
mcm> add package --basedir=/usr/local/mysql_7_0_9 --hosts=192.168.0.14,192.168.0.15 7.0.9;  
mcm> add process --processhosts=mysqlld@192.168.0.10,mysqlld@192.168.0.11,ndbd@192.168.0.14,  
    ndbd@192.168.0.15,ndbd@192.168.0.14,ndbd@192.168.0.15 -s port:mysqlld:52=3307,  
    port:mysqlld:53=3307 mycluster;  
mcm> start process --added mycluster;  
mysql> ALTER ONLINE TABLE <table-name> REORGANIZE PARTITION;  
mysql> OPTIMIZE TABLE <table-name>;
```

【MySQL NDB Cluster運用にあたっての潜在的な課題】

バージョンアップやパフォーマンスチューニング、トラブルシューティングのノウハウ

分散型インメモリデータベースクラスタであるMySQL NDB Clusterの利用において、バージョンアップやパフォーマンスチューニング、またトラブルシューティングの際はMySQLサーバーや他のRDBMSとは異なったアプローチが求められるケースがある

【潜在的な課題】

- 最適な構成や手順の確立に時間を要する
- システム障害の解決に要した時間によるビジネス機会の損失

【ご提案】

サポートサービスOracle Premier Support for MySQLによるご支援

- 開発チームと連携した広範囲なサポートサービスのご提供
- 「コンサルティングサポート」でのバージョンアップやパフォーマンスチューニングに関するノウハウのご提供

MySQL Premier Support for MySQL

- 最大のMySQLのエンジニアリングおよびサポート組織
- MySQL開発チームによるサポート
- 29言語で世界クラスのサポートを提供
- メンテナンス・リリース、バグ修正、パッチ、アップデートの提供
- 24時間x365日サポート
- MySQL コンサルティング・サポート

～リモートDBAとして、是非ご活用ください！！～



Get immediate help for any MySQL issue, plus expert advice

MySQLサポートの特徴

「パフォーマンス・チューニング」や「SQLチューニング」まで通常サポートの範囲内

- コンサルティングサポートが含まれており、「クエリ・レビュー」、「パフォーマンス・チューニング」、「レプリケーション・レビュー」、「パーティショニング・レビュー」などに対応可能
- 詳細はこちらを参照下さい
<http://www-jp.mysql.com/support/consultative.html>

ソースコードレベルでサポート可能

- ほとんどのサポートエンジニアがソースを読めるため、対応が早い開発エンジニアとサポートエンジニアも密に連携している

サポート、
コンサルテーティブ・サポート
はリモートDBAとして活用頂く
事で、自社内の調査・検証工数
を大幅に削減する事が可能。
TCO削減が可能です。

MySQLサポートの特徴

物理サーバー単位課金

- CPU数、コア数に依存しない価格体系
- 4CPUまで（コア数は制限無し）同一料金、5CPU以上の価格は営業問合せ

コミュニティ版バイナリに対してもサポートを提供可能

- サブスクリプションを契約することで、バイナリを入れ替えずにサポートを受けられる（バイナリはオラクルが提供しているものをご使用ください）
- 商用版の機能を使用する場合のバイナリ入れ替えの必要性については次ページ参照
- Oracle CloudのMySQL Cloud Service以外のDBaaSはサポート対象外

オラクルのライフタイムサポート

- 詳細はこちらを参照下さい
<http://www.oracle.com/jp/support/lifetime-support/index.html>
<http://www-jp.mysql.com/support/>

“MySQL Cluster Carrier Grade Editionを導入したことで、高い処理能力をシステム基盤に構築できました。また、高可用性を持ったインメモリデータベースであるMySQL Clusterを導入したことにより、加盟店に関する膨大なデータの取り込み時間を1週間から30分程度まで短縮できました。”

川島 浩一氏

情報システム部 部長, ビットキャッシュ株式会社



MySQL NDB Cluster CGEが解決するお客様の課題

分散型環境の効率的な監視

- 分散型データベースのため、複数のサーバーが監視対象となる
- インメモリデータベースのため、メモリ利用状況の監視が重要となる

→ MySQL Enterprise Monitor

環境内のすべてのノードの稼働状況、メモリ利用状況、SQL性能を一括監視により運用負荷軽減によるコスト低減

構成変更などの管理負荷

- 複数のサーバーで構成された環境の変更には多くのコマンドの実行が必要となり、人為的ミスの発生がリスクとなる

→ MySQL Cluster Manager

管理作業を自動化することで、管理者の負荷軽減と人為的ミスからのシステム障害による機会損失の削減

運用のノウハウや障害解析

- バージョンアップやパフォーマンスチューニング、またトラブルシューティングの際は、MySQLサーバーや他のRDBMSとは異なったアプローチが求められるケースがある

→ Oracle Premier Support for MySQL

障害の発生原因の分析のみならず未然に防止するアドバイスを受けることでの運用効率改善



