

# IQ for MySQL Cluster – 眠らないDBMS –

株式会社インサイトテクノロジー

# Agenda.

**I. MySQLとMySQL Clusterとの違い**

**II. IQ for MySQL Clusterとは**

**III. IQ for MySQL Cluster 検証環境**

# I. MySQL と MySQL Cluster との違い

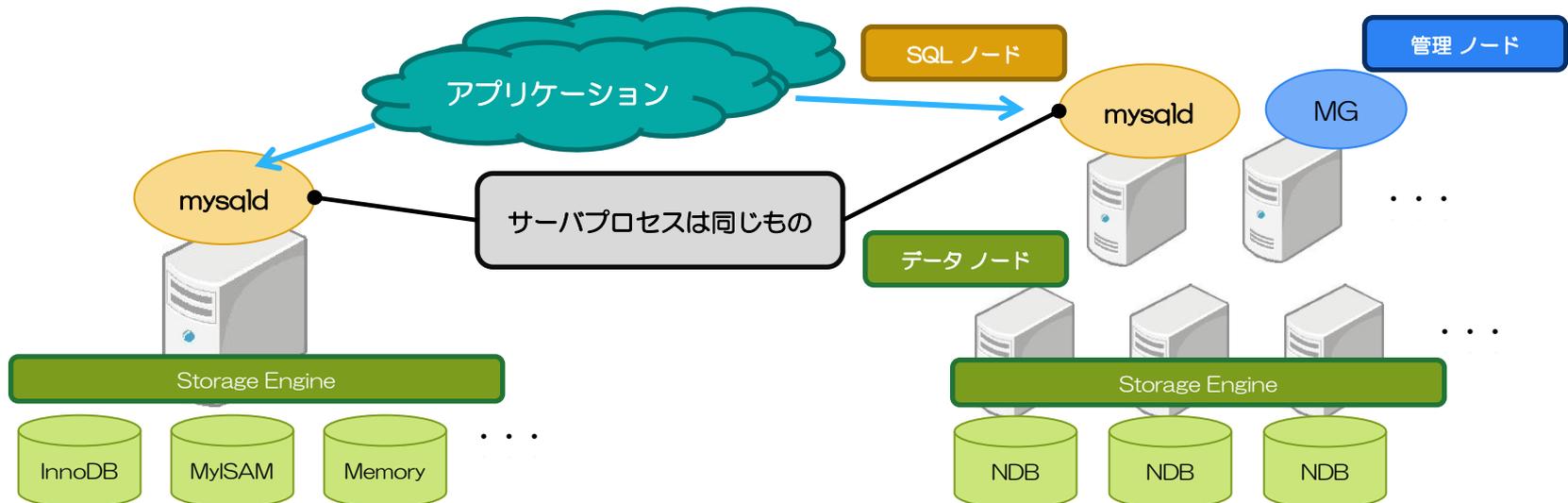
## ■ アーキテクチャ

### MySQL

- 1つのノードにサーバプロセス(mysql)とストレージエンジン(InnoDB、MyISAM、etc...)  
→ データ格納場所はストレージエンジンに依存

### MySQL Cluster

- 管理ノード、SQLノード、データノード  
→ 複数ノードをネットワークで接続
- ストレージエンジンはNDB  
→ データはメモリに保持
- シェアドナッシング・アーキテクチャ



# I. MySQL と MySQL Cluster との違い

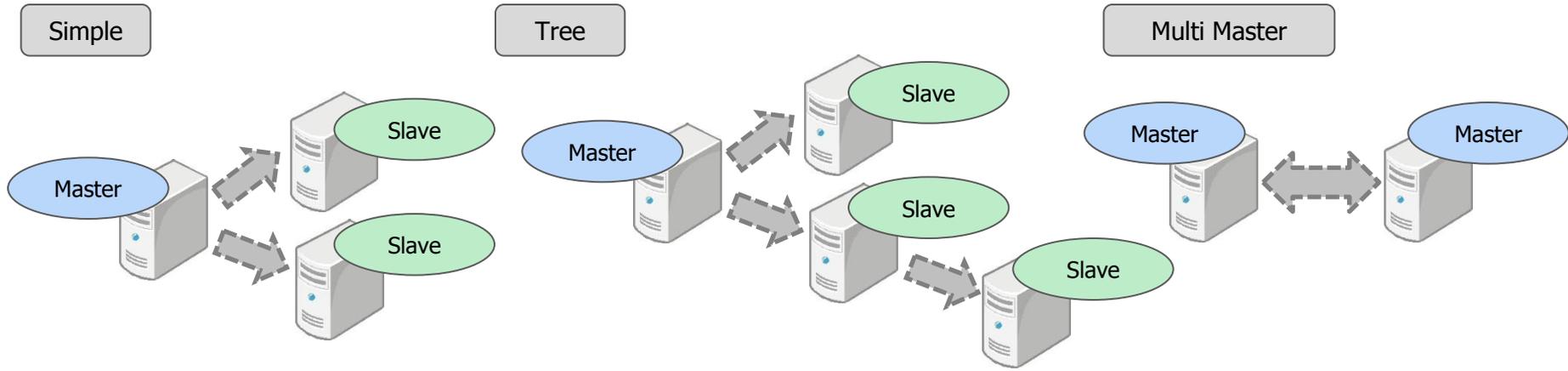
## ■ パフォーマンスアップの方法

### MySQL

- パーツのアップグレード
  - CPU : Clock up、コア追加、etc...
  - メモリ : 容量追加、高速なものに交換、etc...
- レプリケーション
  - データをレプリケーションで分散



構成変更時サービスは停止

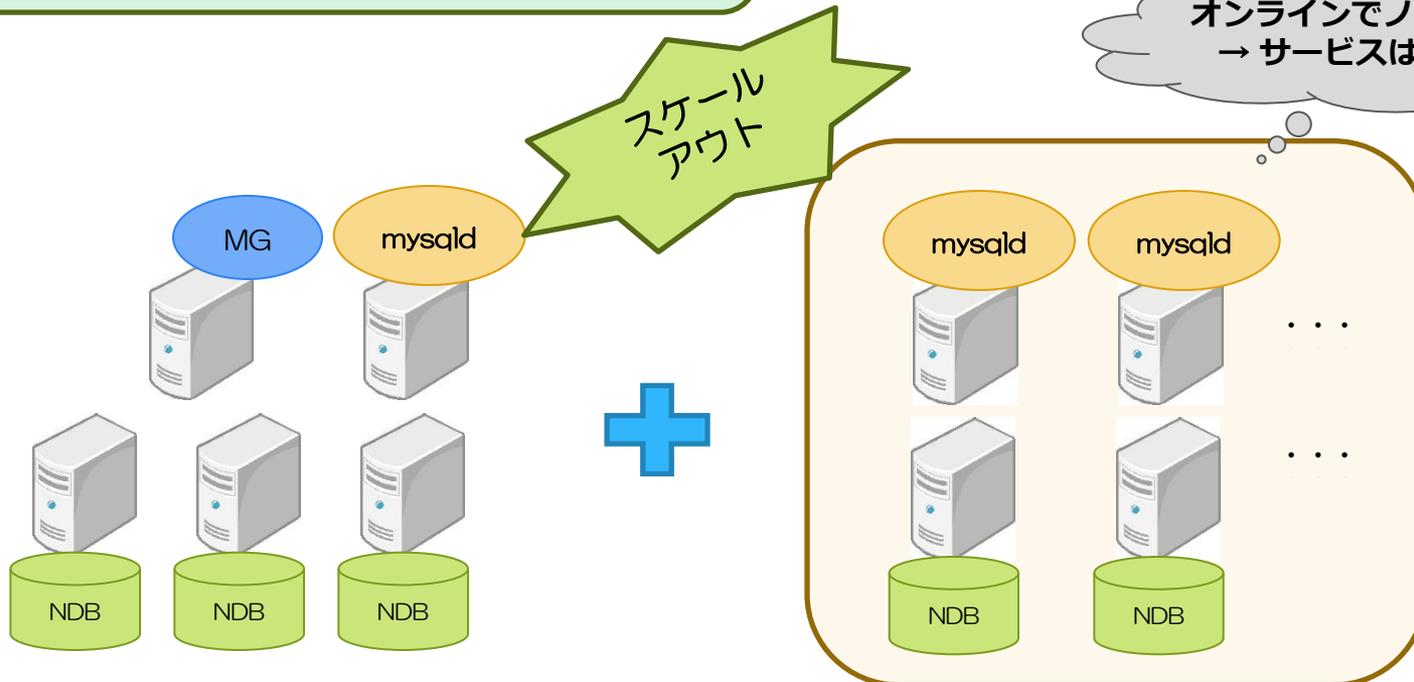


# I. MySQL と MySQL Cluster との違い

## ■ パフォーマンスアップの方法

### MySQL Cluster

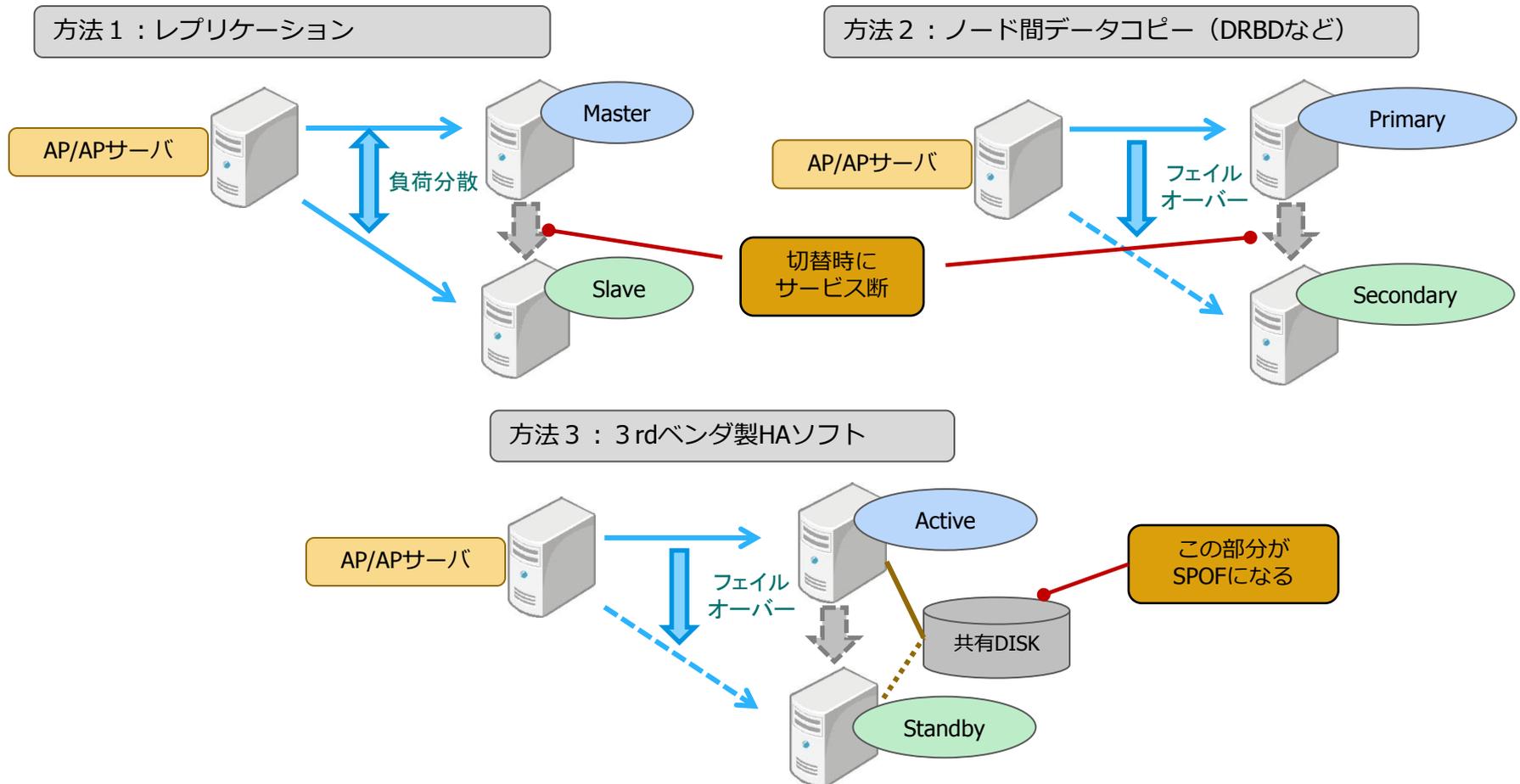
- SQLノードまたはデータノードを追加  
→ 高スペックのマシンでなくてもよい  
(同一のアーキテクチャである必要があるが・・・)



# I. MySQL と MySQL Cluster との違い

## ■ 高可用性構成

### MySQL

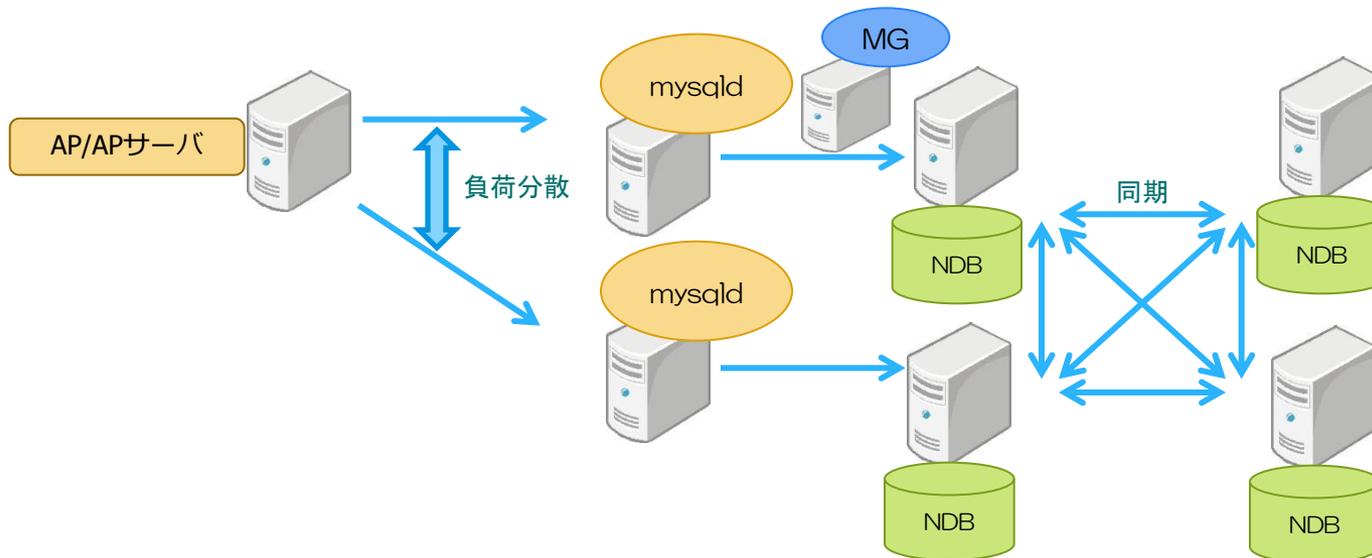


# I. MySQL と MySQL Cluster との違い

## ■ 高可用性構成

### MySQL Cluster

- SQLノードまたはデータノードを追加  
→ 高スペックのマシンでなくてもよい  
(同一のアーキテクチャである必要がありますが・・・)

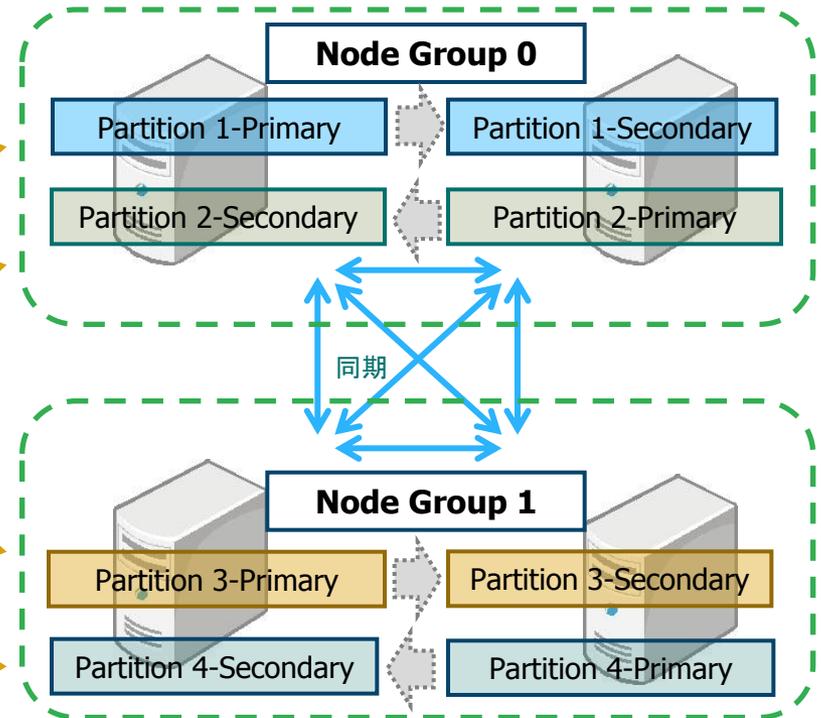
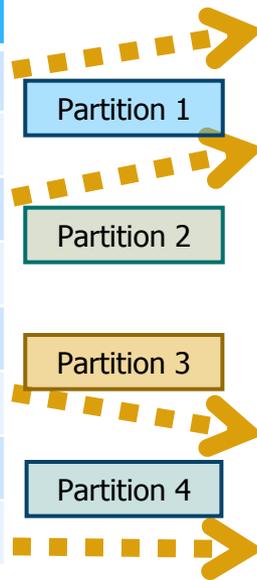


# I. MySQL と MySQL Cluster との違い

## ■ 高可用性構成

- データのパーティションへの振り分け  
→ 主キー値をハッシュ関数で振り分け

C_ID	C_CITY	C_STATE	C_ZIP
001	aaaa	xxxxx	1234567
002	bbbb	yyyyy	2345678
003	cccc	zzzzz	3456789
004	dddd	vvvvv	4567890
005	eeee	wwwww	5678901
006	ffff	ooooo	6789012
007	gggg	ppppp	7890123
008	hhhh	qqqqq	8901234



# I. MySQL と MySQL Cluster との違い

## ■ 高可用性構成

- 継続可能なサービス  
→ ノードグループ内の全ノードが破損しない限りサービスの継続が可能

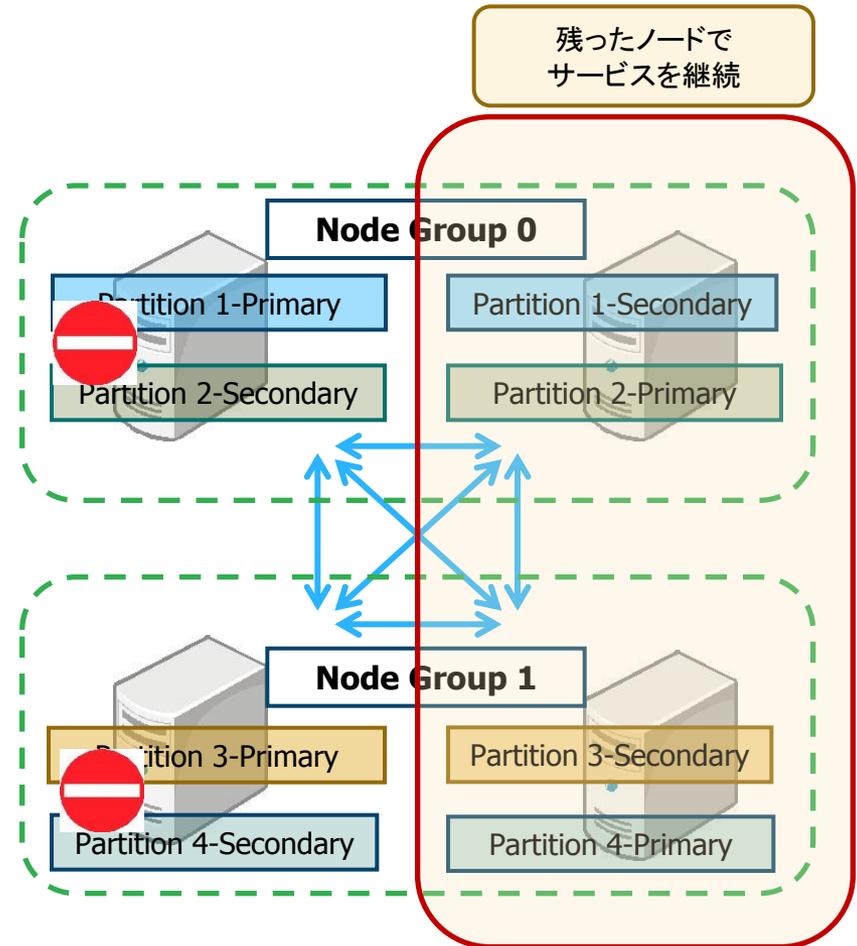
C_ID	C_CITY	C_STATE	C_ZIP
001	aaaa	xxxxx	1234567
002	bbbb	yyyyy	2345678
003	cccc	zzzzz	3456789
004	dddd	vvvvv	4567890
005	eeee	wwwww	5678901
006	ffff	ooooo	6789012
007	gggg	ppppp	7890123
008	hhhh	qqqqq	8901234

Partition 1

Partition 2

Partition 3

Partition 4



# I. MySQL と MySQL Cluster との違い

## ■ 高可用性構成

- データの自動復旧  
→ ノードが復旧したら自動的にデータを再同期

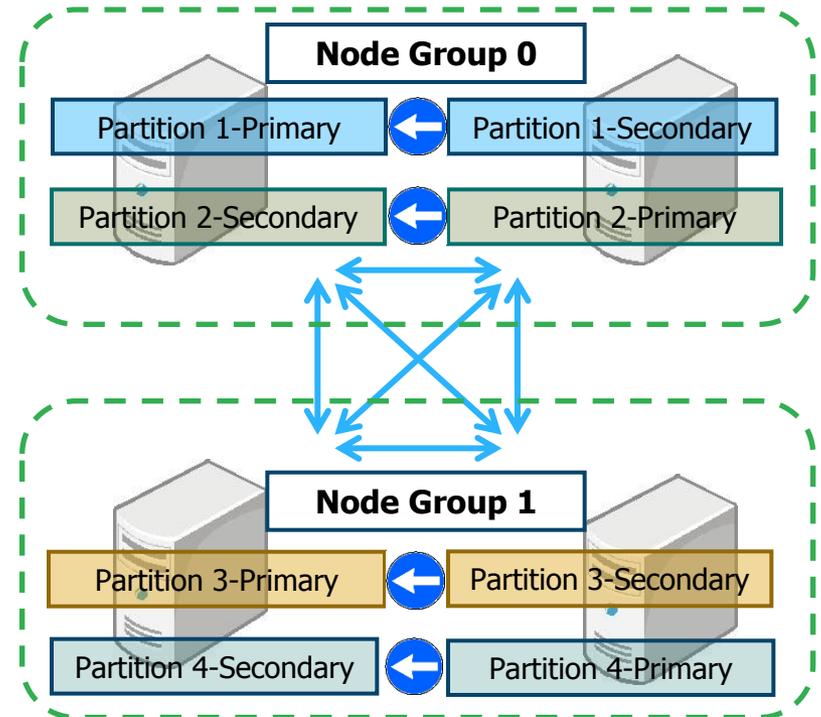
C_ID	C_CITY	C_STATE	C_ZIP
001	aaaa	xxxxx	1234567
002	bbbb	yyyyy	2345678
003	cccc	zzzzz	3456789
004	dddd	vvvvv	4567890
005	eeee	wwwww	5678901
006	ffff	ooooo	6789012
007	gggg	ppppp	7890123
008	hhhh	qqqqq	8901234

Partition 1

Partition 2

Partition 3

Partition 4



# I. MySQL と MySQL Cluster との違い

## ■ 向いている用途

- 大量の書込みトランザクション  
→ シェアドナッシング・アーキテクチャの利点を生かした分散処理
- トランザクション消失が許されない処理  
→ 複数データノードの同期処理によるデータ保全
- レスポンスが要求される処理  
→ データをメモリに保持するため、高速なデータ処理が可能

OLTP 得意!

## ■ 向かない用途？

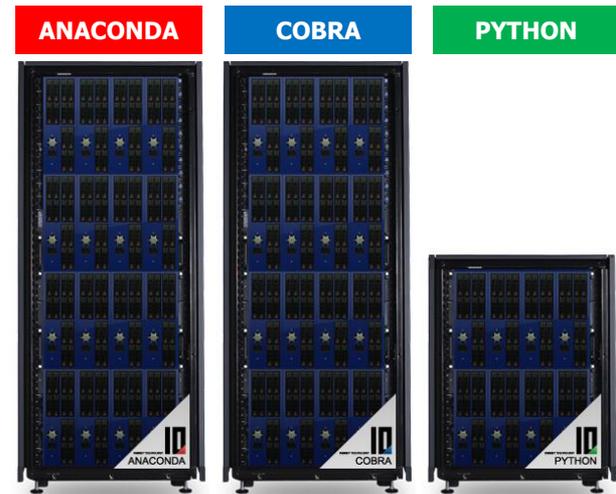
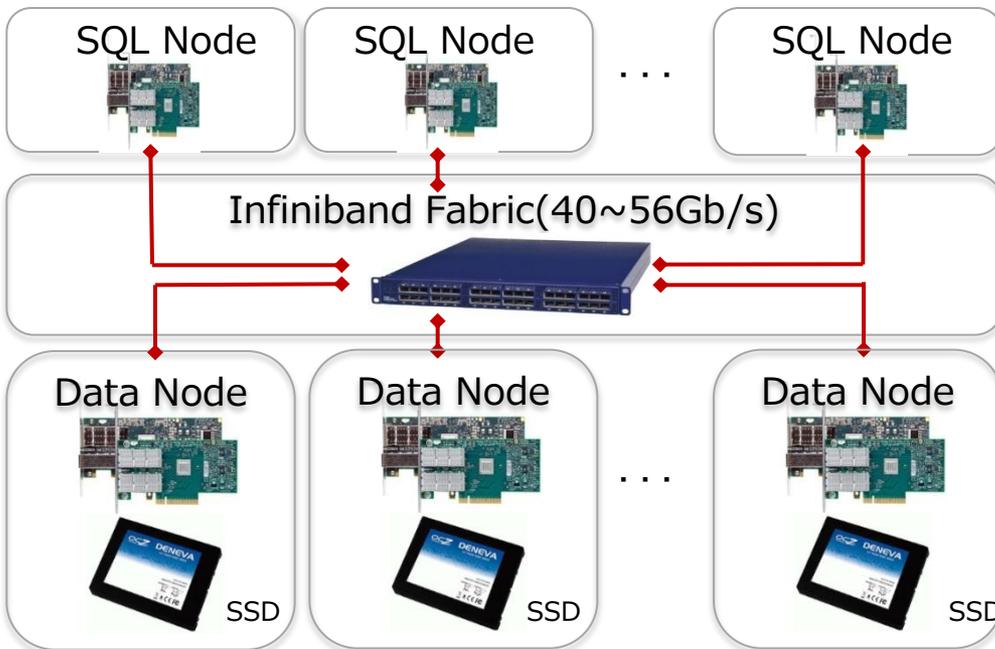
- 集計処理  
→ group by などの集計処理はシェアドナッシング・アーキテクチャが苦手な処理

DWH 苦手?

現在検証中

## II. IQ for MySQL Clusterとは

高性能インメモリ型クラスターデータベース  
と  
Super I/O Tunedシステム の融合



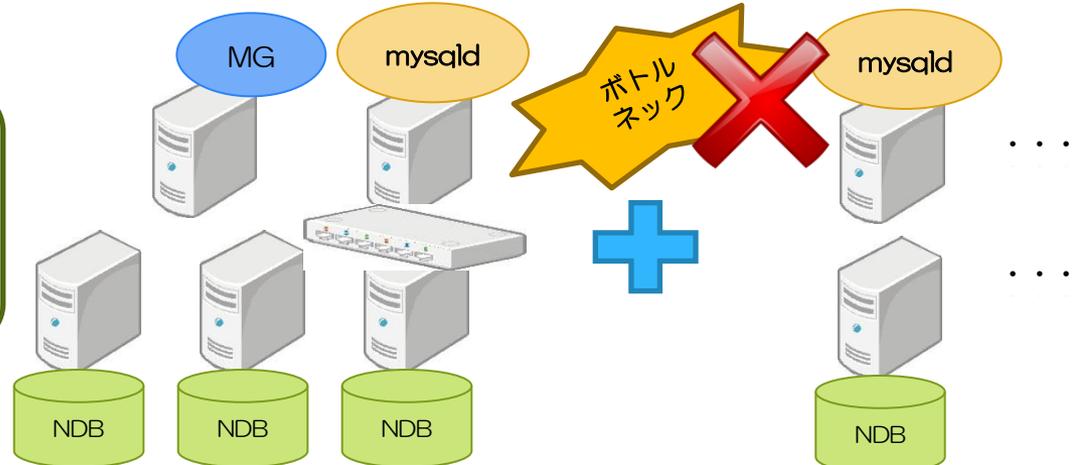
<http://www.insight-tec.com/products/IQ/index.html>

- SSDとInfinibandを利用したIQ  
複数データノードからの同時読込は  
**10GB/s**を記録!!!
- CPU, Memory, Storage構成まで含めて  
**MySQL Cluster**にベストチューンされたモデル  
を提供

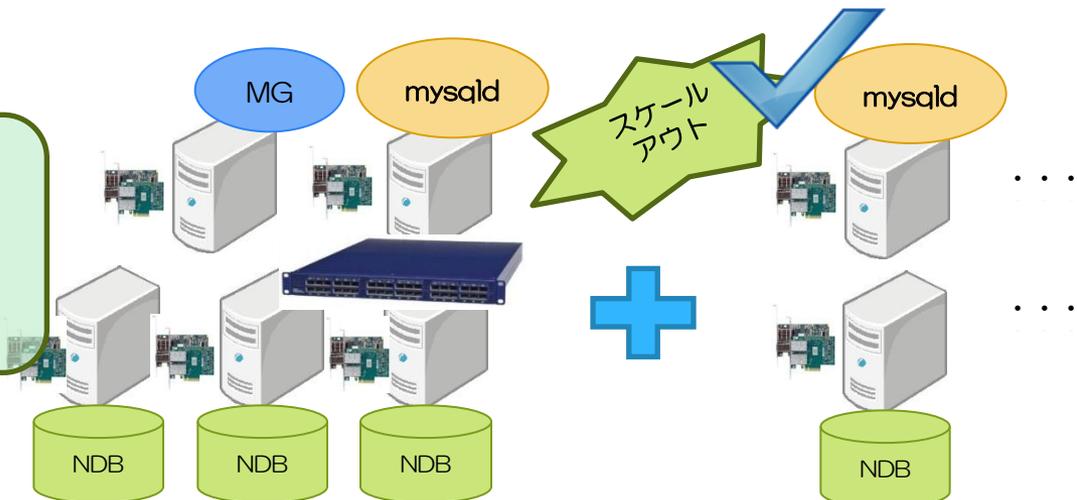
## II. IQ for MySQL Clusterとは

### ■ (前出) MySQL Clusterパフォーマンスアップの方法

- スケールアウトモデルはネットワークがボトルネックになってしまう

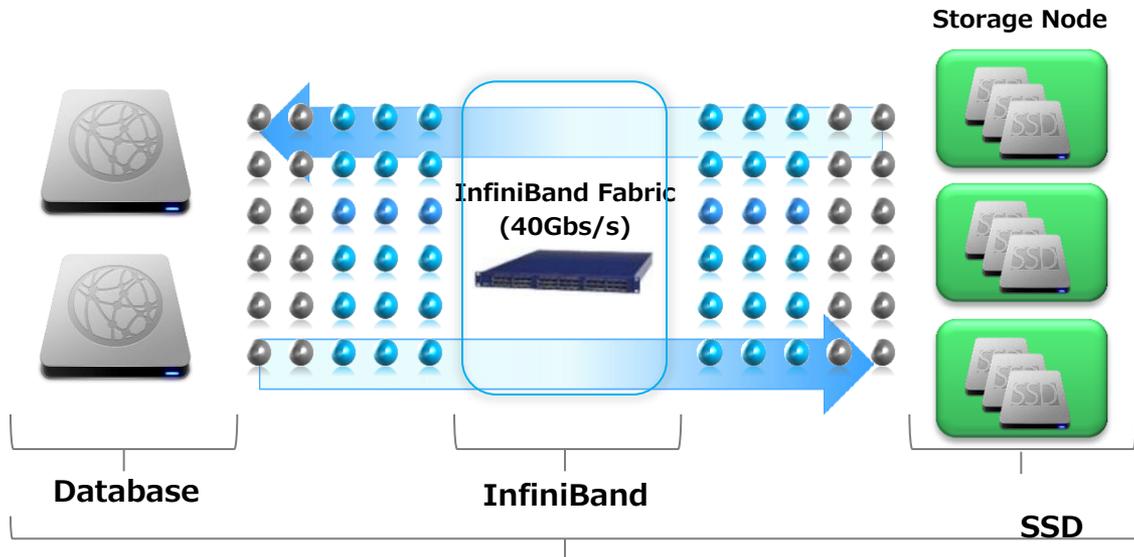


- 伝送遅延の少ないInfinibandならノードを追加してもネットワークがボトルネックになりにくい



## II. IQ for MySQL Clusterとは

### HPCとコモディティがオープンな規格で融合したハードウェア



#### HPC

ハイパフォーマンスコンピューティング (HPC)で長年実績のあるInfiniband Fabric

#### コモディティ

現在、I/O性能と低電力からHDDに代わりつつあるSSD

#### オープン

SSDで高速化された各ストレージノードをオープン規格SRP(SCSI RDMA Protocol)で結んだクラスターストレージ

Insight Knowledge

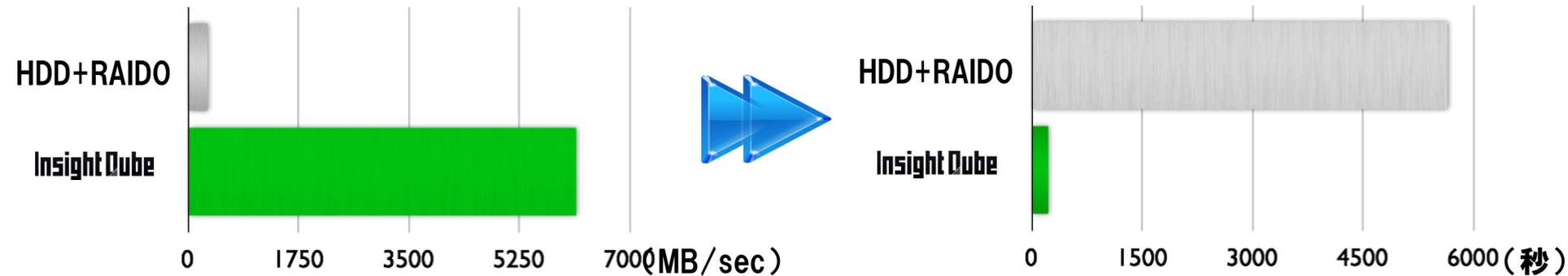
## II. IQ for MySQL Clusterとは

大量検索処理は従来システムと比べて 20倍以上高速に

select \* from ...

スループットが20倍に

処理時間は1/25に

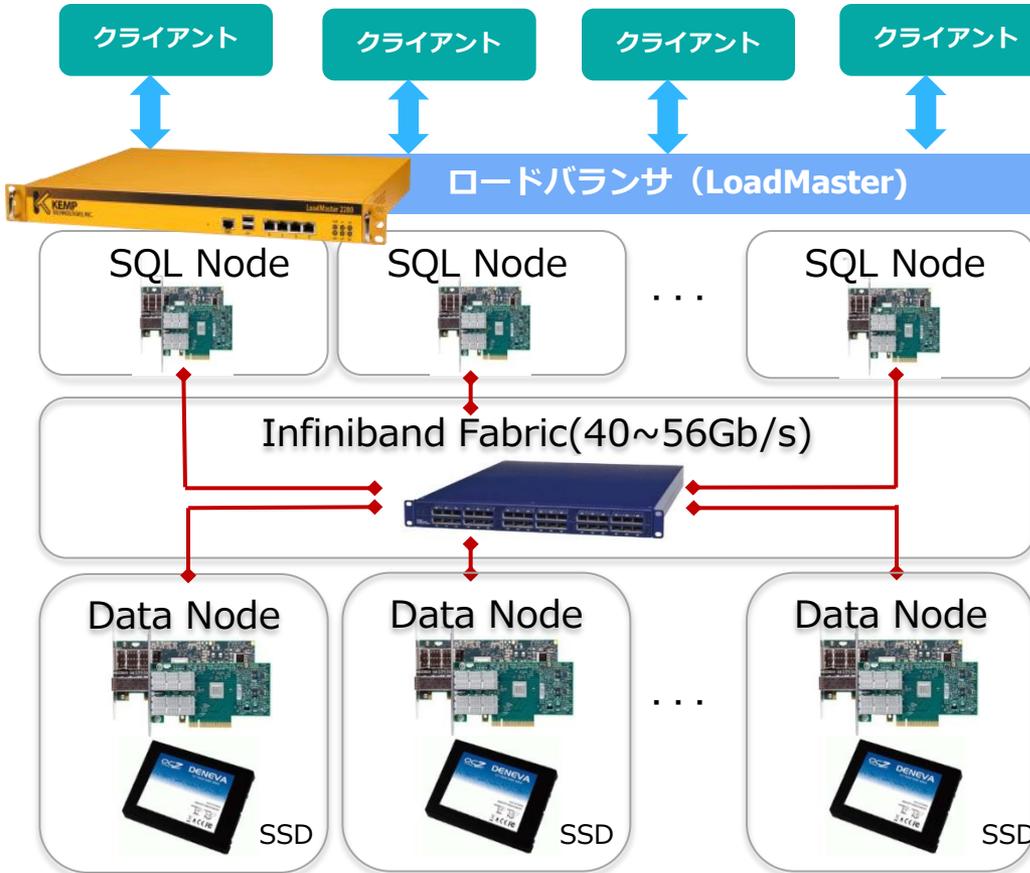


\* Elapsed TimeはTPC-H Power Test実施時の総処理時間(秒)

\* Throughputは、vdbenchによるBlocksize 1MBでのシーケンシャル読み込み時の実測値(MB/s)

## II. IQ for MySQL Clusterとは

更なる可用性向上を目指して・・・



### ロードバランサ

ロードバランサは、SQL Nodeを抽象化と負荷分散を提供します。クライアントはロードバランサを一切意識する必要はありません。

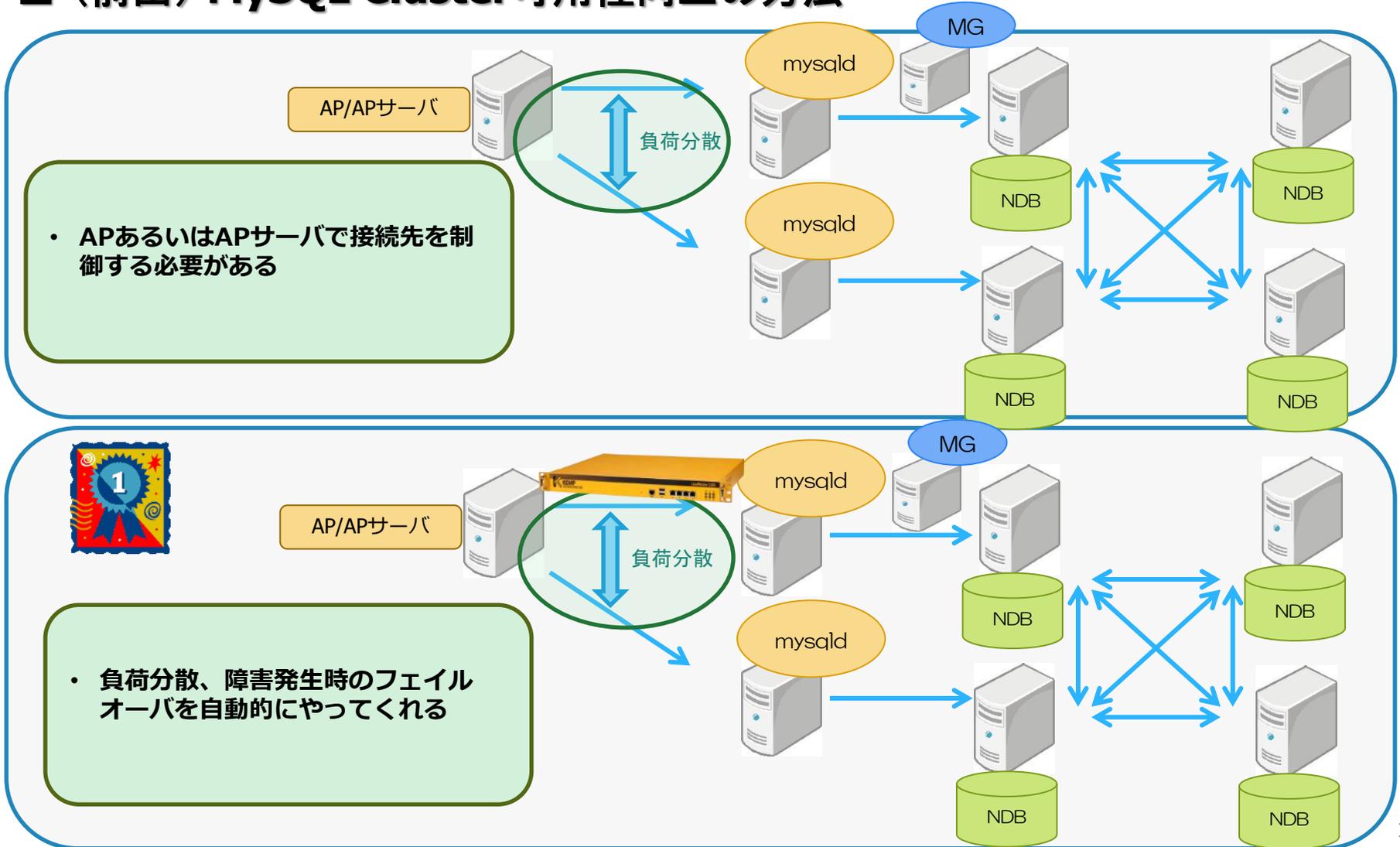
- 負荷の少ないSQL Nodeにクライアントのリクエストを転送します。
- SQL Nodeの障害を自動的に検知し、障害が発生していないSQL Nodeにのみクライアントのリクエストを転送します。
- SQL Nodeが追加された場合でも、クライアントは意識する必要がありません。

### Load Master

- 全世界で1万台以上の稼働実績を持つロードバランサです。
- 中小規模からエンタープライズまでをサポートしたコストパフォーマンスに優れたロードバランサです。
- DSRによりMySQL Clusterからの応答はLoadMasterを経由せず、直接クライアントに戻されます。

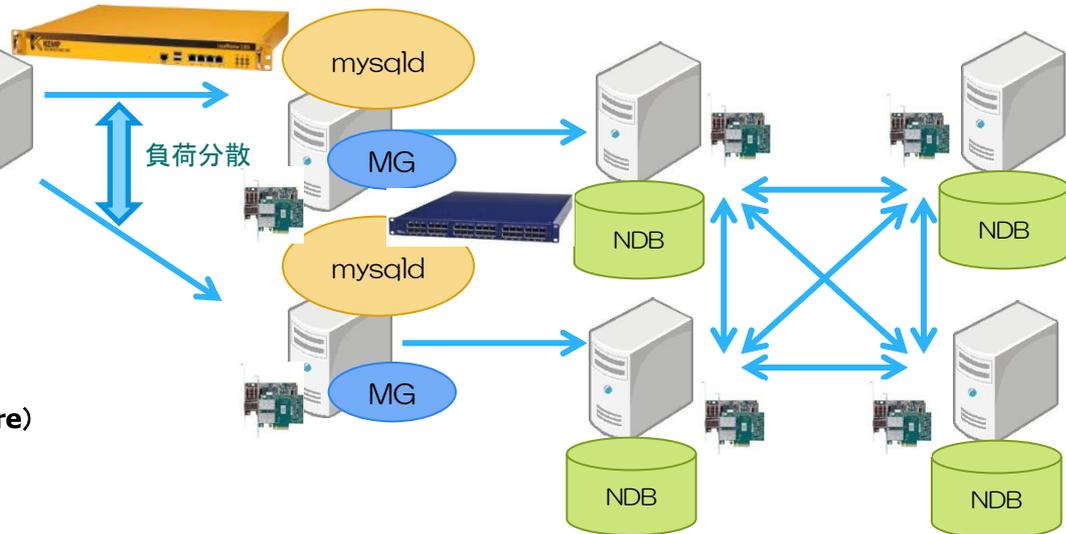
## II. IQ for MySQL Clusterとは

### ■ (前出) MySQL Cluster可用性向上の方法



# III.IQ for MySQL Cluster 検証環境

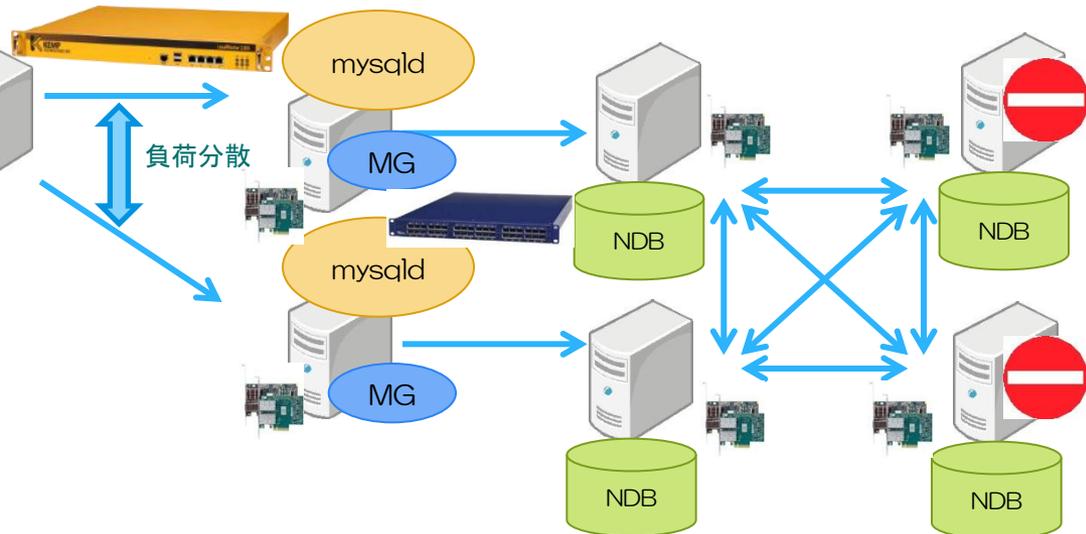
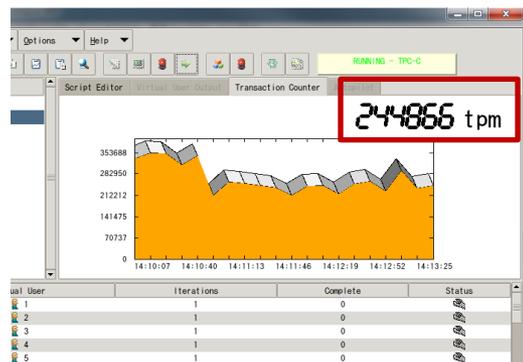
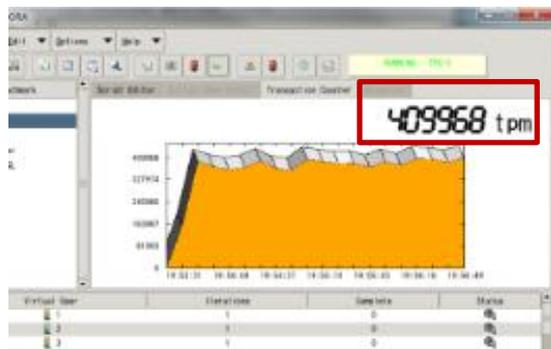
## IQ PYHTON による検証環境



**CPU: AMD A4-3400 × 1 (2.7GHZ, 2core)**  
**Memory: 6GB**  
**Infiniband: QDR(40Gbps)**

# III.IQ for MySQL Cluster 検証環境

試しにデータノードを落としてみました



- ・ ノード数が減るのでパフォーマンスは落ちます。
- ・ サービスは停止していません。  
(実行途中でロールバックされるトランザクションもあります)
- ・ ノードが復旧すると元のパフォーマンスに戻ります。

# III.IQ for MySQL Cluster 検証環境

## IQ PYHTON による構成(実物)



EthernetケーブルとInfiniband ケーブル

Infiniband スイッチ



# III.IQ for MySQL Cluster 検証環境

MySQL Cluster のスケールアウト検証



<http://enterprisezine.jp/dbonline/detail/4315>

現在検証中

ご静聴ありがとうございました。