

MySQL Clusterによる 高可用システム運用ガイド

山崎 由章 = 著

インストールからバックアップ/リストア、
サイジングやレプリケーションまで、
MySQL Clusterの運用ノウハウが身につく!

インプレス

- 本書は、インプレスが運営するWebメディア「Think IT」で、「MySQL Cluster 運用ガイド」として連載された技術解説記事を電子書籍およびオンデマンド書籍として加筆・再編集したものです。
- 本書の内容は、執筆時点までの情報を基に執筆されています。紹介したWebサイトやアプリケーション、サービスは変更される可能性があります。
- 本書の内容によって生じる、直接または間接被害について、著者ならびに弊社では、一切の責任を負いかねます。
- 本書中の会社名、製品名、サービス名などは、一般に各社の登録商標、または商標です。なお、本書では©、®、TMは明記していません。

はじめに

本書では「MySQL Cluster」を利用するためのチュートリアルとなるように、その特徴と基本的なアーキテクチャからインストール方法、基本的な操作などをコマンド付きで解説します。

表 .1 に Think IT 連載時の解説内容を一覧にまとめました。本書をきっかけに、MySQL Cluster のユーザーが 1 人でも増えれば、とても嬉しく思います。

表 1 本書の Think IT 連載時における解説内容

連載回数	解説内容
第 1 回	MySQL Cluster の概要（特徴とアーキテクチャ） https://thinkit.co.jp/story/2015/07/21/6191
第 2 回	MySQL Cluster のインストール方法、基本的な設定&操作 https://thinkit.co.jp/story/2015/07/30/6247
第 3 回	MySQL Cluster の主要な設定（パラメータ）、設定変更方法 https://thinkit.co.jp/story/2015/08/17/6279
第 4 回	MySQL Cluster のバックアップ/リストアの仕組み、関連するパラメータ https://thinkit.co.jp/story/2015/08/25/6337
第 5 回	MySQL Cluster のバックアップ/リストア的具体例 https://thinkit.co.jp/story/2015/09/09/6395
第 6 回	MySQL Cluster のサイジング方法 https://thinkit.co.jp/story/2015/09/30/6430
第 7 回	MySQL Cluster のレプリケーションの基礎 https://thinkit.co.jp/story/2015/10/29/6523
第 8 回	MySQL Cluster のレプリケーション環境構築例 https://thinkit.co.jp/article/8419
第 9 回	MySQL Cluster のチューニングの基礎 https://thinkit.co.jp/article/8462

本書において示されている見解は、私自身の見解であって、私の所属するオラクルの見解を必ずしも反映したものではありません。

目 次

はじめに	iii
第 1 章 MySQL Cluster の特徴とアーキテクチャ	1
1.1 MySQL Cluster とは？	1
1.2 MySQL Cluster の用途や事例	2
1.3 MySQL Cluster のアーキテクチャ	2
1.4 MySQL Cluster の特徴	6
第 2 章 MySQL Cluster のインストールと基本的な設定および操作	9
2.1 MySQL Cluster のインストール	9
2.2 Linux 環境でのインストール手順	10
2.3 基本的な操作	14
第 3 章 MySQL Cluster の主要な設定、設定変更時の注意点	19
3.1 パラメータ変更時の注意点とパラメータ変更例	19
3.2 主要なパラメータ設定	23
第 4 章 MySQL Cluster のバックアップ／リストアの仕組み	33
4.1 MySQL Cluster がデータを永続化する仕組み	33
4.2 バックアップ／リストアの仕組み	35
4.3 バックアップ、リストア関連のパラメータ	37

第 5 章	MySQL Cluster のバックアップ／リストアの具体例	41
5.1	前提となる環境	41
5.2	オンラインバックアップの取得	44
5.3	リストアの実行	46
第 6 章	MySQL Cluster のサイジング	53
6.1	前提となる環境	53
6.2	はじめに	53
6.3	MySQL Cluster が使用するインデックスの種類	54
6.4	メモリ使用量のサイジング	56
6.5	ディスク使用量のサイジング	60
第 7 章	MySQL Cluster におけるレプリケーションの基礎	63
7.1	前提となる環境	63
7.2	レプリケーションのメリットと主な用途	63
7.3	MySQL Cluster のレプリケーションの仕組み	64
7.4	レプリケーションの注意事項	69
第 8 章	MySQL Cluster におけるレプリケーション環境構築例	73
8.1	前提となる環境	73
8.2	レプリケーション環境構築の流れ	73
8.3	レプリケーション環境構築の具体例	74
第 9 章	MySQL Cluster におけるチューニングの基礎	83
9.1	スループットを向上させる	83
9.2	SQL をチューニングする	85
9.3	その他のチューニング TIPS	87
第 10 章	Distribution Awareness とパーティショニングテーブルを活用したチューニング	89
10.1	Distribution Awareness とは？	89
10.2	パーティショニングテーブルを使用してデータの分散方法を変更する	91

10.3	Distribution Awareness とパーティショニングテーブルを活用した、複数テーブルにアクセスするトランザクションのチューニング	97
10.4	パーティショニングテーブル使用時の注意事項	98

第1章 MySQL Clusterの特徴とアーキテクチャ

第1章では、MySQL Cluster の特徴とアーキテクチャについて解説します。

1.1 MySQL Cluster とは？

「MySQL Cluster」は「MySQL Server」とは開発ツリーの異なる製品で、共有ディスクを使わずにアクティブ-アクティブのクラスタ構成が組めるリレーショナルデータベースです。カラムやインデックス、ノードの追加・削除といった各種メンテナンス処理をオンラインで実行できる、単一障害点がなく可用性が非常に高い、などの特徴があります。そのため、米国海軍の航空母艦における航空機管制システムなど、ミッションクリティカルな分野でも多く利用されています。

また、基本的にはデータとインデックスを全てメモリ上に持つインメモリデータベースであり、トランザクションを高速に処理できるため、リアルタイム性が求められるアプリケーションにも向いています。インメモリデータベースとして使用する場合でも、データが更新された時には更新内容をディスクに保存しているため、データの永続性は担保されています（データベースを再起動した時に更新したデータが消失するといったことはない）。

MySQL Cluster の基礎となる技術は、通信機器ベンダであるエリクソンの携帯通信網加入者データベース向けに開発された「Ericsson Network Database (NDB)」です。携帯通信網加入者データベースは、大半の処理が加入者 ID に紐づく処理であるため、1つ1つの処理は比較的シンプルです（RDBMS 視点では主キーに紐づく処理、NoSQL 視点では KVS (Key-Value Store：キーバリューストア) 的な処理）。

しかし、携帯電話の加入者増加に合わせて多重実行される処理数も増えるため、システムには

拡張性が求められます。また、通話履歴などの書き込み処理が大量に発生するため、読み込み処理に対してだけでなく、書き込み処理に対する拡張性も求められます。さらに通信障害が起きれば収益機会を逃してしまう上に社会インフラとしても大きな問題となるため、高可用性も求められます。このような“大量のトランザクション処理”や“高可用性”が求められる環境向けに最適化された NDB と MySQL を統合したものが、現在の MySQL Cluster です。

また、元々 NDB は SQL を使わずに C++ の API でデータを処理していましたが、MySQL と統合されたことにより SQL も使えるようになりました。そのため、現在は NoSQL (KVS) と SQL の両方が使用できます。そして、NoSQL による処理でも ACID 準拠のトランザクションに対応しています。さらに、現在は C++ の API 以外にも Java、memcached、Node.js など各種の NoSQL API が使用できるため、NoSQL の形態で Web アプリケーションのバックエンドとしても利用しやすくなっています。

1.2 MySQL Cluster の用途や事例

前述したように、MySQL Cluster は携帯通信網加入者データベース向けに開発された技術が基礎になっているということもあり、Alcatel-Lucent や Nokia、NEC などの大手通信機器ベンダでも活用事例があるほか、携帯キャリアのコンテンツ配信プラットフォーム等でも多数利用されています。また、オンラインゲームでも Big Fish Games、Blizzard Entertainment、Zynga といった海外ベンダの他、国内ベンダでも採用が広がっています。

さらに、単一障害点がなく非常に可用性が高いことなどから、前述の航空機管制システムのほか PayPal や国内の証券会社など、ミッションクリティカルなシステムでの採用事例も増えてきています。

1.3 MySQL Cluster のアーキテクチャ

MySQL Cluster は「データノード」「SQL ノード」「管理ノード」という 3 種類のノードから構成されています。図 1.1 は、データノード 4 台、SQL ノード 2 台、管理ノード 2 台で構成した場合のイメージ図です。各ノードの主な役割は表 1.1 の通りですが、順番に解説していきます。

データノード：データ管理とトランザクション制御

データノードは MySQL Cluster の肝となるノードで、データやインデックスを保持し、トランザクションを制御します。データノードでは、挿入されたデータを各テーブルの主キーのハッ

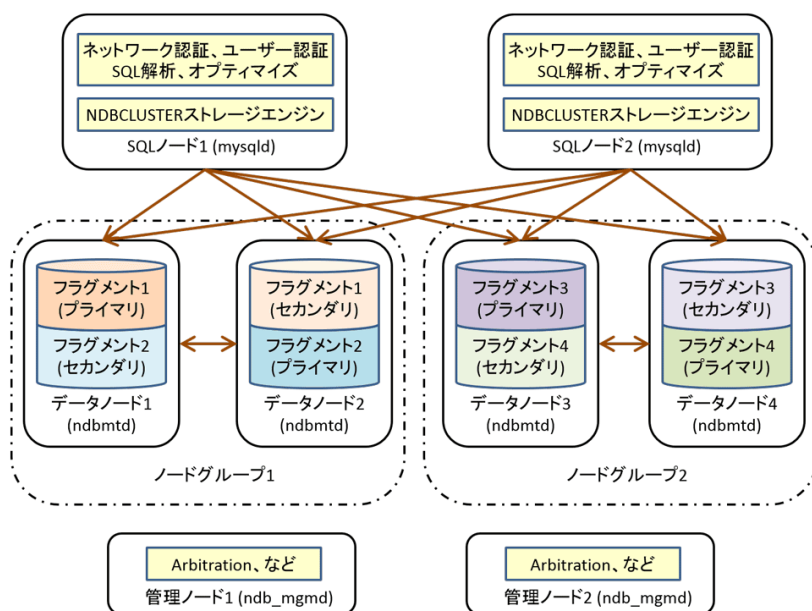


図 1.1 MySQL Cluster アーキテクチャ概要

表 1.1 MySQL Cluster の各ノードの役割

ノード名	プロセス名	主な役割
データノード	ndbmd	・データ、インデックスの管理 ・トランザクションの制御
SQL ノード	mysqld	・アプリケーションとデータノードをつなぐ SQL インターフェース ・ユーザー認証、権限付与
管理ノード	ndb_mgmd	・MySQL Cluster の起動/停止 ・各ノードの設定管理 ・バックアップ/リストア処理の開始 ・Arbitration（スプリットブレインの解消）

シュ値に基づいて水平分割（行単位で分割）して格納します。さらに、分割したデータを同じノードグループ内のデータノードに複製して持つことで、データを冗長化します。データの冗長化はデフォルトでは2重化ですが、最大4重化まで設定できます。

図 1.1 の例ではデータノードが4台存在するため、データは4分割され、4つのデータノードに分散して格納されています。分割された各フラグメント（断片）の複製を別のデータノードが保持するため、データノード1はフラグメント1だけでなくフラグメント2の複製も保持しています。同様に、データノード2はフラグメント2だけでなくフラグメント1の複製も保持して