

株式会社ドワンゴ

MariaDB Galera Cluster でマイクロサービス化に対応 高可用性の実現と運用コストの削減に成功

概要



ドワンゴ株式会社（以下：ドワンゴ社）はユーザの多様なニーズに応えるため、数年前よりマイクロサービス化を推進している。それに伴いサービスリリースの早期化やコスト削減などといった課題を解決するべく、各サービスで利用可能なデータベースシステム基盤の構築を決定。従来使用していた MySQL のマスタ / スレーブ構成ではなく、MariaDB Galera Cluster を利用したマルチマスタ構成を新たに採用し、以前よりも高い可用性を担保しながら運用体制の効率化を実現した。

課題

マイクロサービス化に対応するため、データベース運用の効率化と更なる安定的な運用を目指す

解決策

MariaDB Galera Cluster を使用した新データベース基盤の構築

結果

■ ダウンタイムの大幅削減によりデータベースの安定した運用を実現

以前ではフェイルオーバー時に1時間以上かかっていたものが、Galera Cluster を利用したマルチマスタ構成に変更することでダウンタイムがほぼ“0”に。

■ データベース基盤の集約化に伴い、運用コストの削減に成功

1つの基盤から各サービスに必要なデータベースを払い出すことで、包括的な運用管理（メンテナンスの簡略化など）が可能になり、運用にかかるトータルコストの削減に成功。

バージョン

MariaDB 10.1 系

対象システム

マイクロサービス化に伴う新サービスに対する新規データベース基盤

導入インタビュー

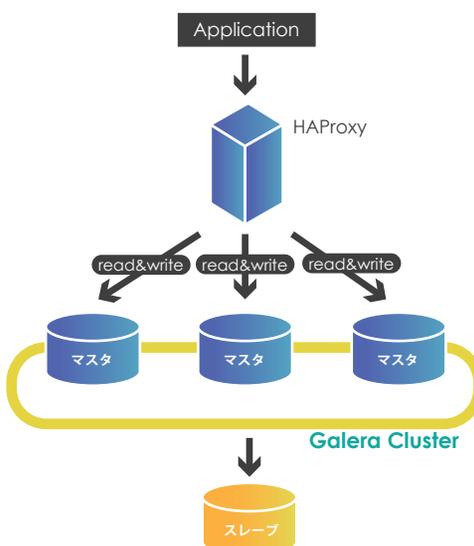
導入の経緯

MariaDB Galera Cluster を導入するに至ったきっかけは、当社全体で推進しているマイクロサービスへの対応のためです。本システムの構築以前は、MySQL を使用したマスタ / スレーブ構成のデータベースシステムを利用していましたが、今回サービスごとに専用の MySQL を構築するのではなく、運用管理の効率化や構築コストの削減のためデータベースを提供する基盤をひとつ用意しその中から新規サービスに必要な容量を払い出していくという方針を取るようになりました。

その際にいくつかの方法（NoSQL, Oracle など）を検討しましたが、最終的に「MySQL との互換性」「マスタ / スレーブ構成ではない点」「ダウンタイムが少ない」「導入コスト」などの面から MariaDB Galera Cluster の導入を決定しました。また、今までは自社内で運用保守を行ってきたのですが、基盤化に伴いより一層安定したシステム稼働を目指したいという考えの下、MariaDB 社の公式サポートが付随した MariaDB Enterprise Cluster※1 を利用するに至りました。

スマートスタイル社から購入した理由としては、提案を受けた際にエンジニアの方から直接詳しい技術的な要素まで聞くことができたので、今後のやりとりもスムーズに進めることができると感じたからです。

システム構成



本システムは、当社の各サービスのデータベースの集約化を目的として新しく構築した新データベースシステム基盤となります。

現在はおおよそ 1000qps (Query Per Second) くらいの規模で、その大半は SELECT です。更新クエリは 20 ~ 30qps くらい、データ量は全体で約 60GB となり、使いたいサービスごとにデータベースを払い出して使っています。ひとつひとつのインスタンスは 10GB 程度ですが、サービスごとに使われ方は異なります。MariaDB は 10.1 系を使用、ノード数はマスタ 3 台で、そのうちの 1 台にバックアップ用のスレーブ 1 台をぶら下げています。スレーブはバックアップ目的です。また、マスタサーバの前段に HAProxy があります。Galera Cluster のおかげで全ノードに対し read と write が可能なので、HAProxy で read と write の両方を全ノードに分散させるマルチマスタ構成にしています。サーバは全て物理サーバ、ディスクは SSD。ただし、特別スペックが高いわけではなく、汎用的なサーバとなっています。

開発体制

基盤システム全体での開発期間は 4 ヶ月程度です。そのうち構築にかかった期間は検証なども含め 1 ヶ月程度です。それぞれデータベースに関連する業務に就いていたメンバー 5 人の体制で行いましたが、いわゆる DB 専門の「スペシャリスト」はいませんでした。

導入手順

MariaDB Galera Cluster の導入に関しては yum でインストールを行いました。手順は MariaDB 社が公開している公式マニュアルを参考にしましたが、設定なども含め簡単に出来たので、特に困ったことはありませんでした。HAProxy については、以前から他のシステムでは利用していたのですが、私達のチームで扱うシステムとしては Galera Cluster に併せて初めて導入する運びとなりました。また、導入後の DB 監視については、「Zabbix」を使用しています。サーバリソースや MySQL ステータスは勿論ですが、Galera 特有のオプション (wsrep_... 変数) も新しく監視対象としました。

メリット

導入時

MySQL からの移行に関しては、ほとんどのパラメータが同じでとてもスムーズに行うことができました。アプリに関しても特に修正の必要がなく、設定ファイル (my.cnf) も流用することができたので、MariaDB は MySQL に近いものだと感じました。

性能面

以前のマスタ / スレーブ構成のものであればサーバが高負荷となりマスタがダウンした際、フェイルオーバーに 1 時間ほど時間がかかり、長いダウンタイムが発生していました。しかし、新しく構築した本システムでは MariaDB Galera Cluster を用いたマルチマスタ構成を採用しているため、問題発生時には参照先のノードを切り替えるだけで済むため、ダウンタイムはほぼなくなりました。

性能面

以前のマスタ / スレーブ構成のものであればサーバが高負荷となりマスタがダウンした際、フェイルオーバーに 1 時間ほど時間がかかり、長いダウンタイムが発生して

いました。しかし、新しく構築した本システムでは MariaDB Galera Cluster を用いたマルチマスタ構成を採用しているため、問題発生時には参照先のノードを切り替えるだけで済むため、ダウンタイムはほぼなくなりました。

運用面

運用に関しては、今のところ安定稼働していて、特に不具合は発生していません。今回の新システムから自動構築系のスクリプトを導入したこともあり、DB サーバにログインして直接コマンドを実行して…、などの作業も減り運用もかなり簡略化できました。また、Galera Cluster では RSU (Rolling Schema Upgrade)※2 が可能なのでこれまで困難だった「マスタ」のメンテナンスも容易になりました。DB のメンテナンスにかかるコストや手間を省くという面でも大きなメリットだと思います。

サポート

本システムの構築時や運用の際に MariaDB 社のサポートを利用する機会が幾度あったのですが、レスポンスも早く回答内容もしっかりしており満足できるものでした。具体的には、構築時にエラーが起こった際の原因特定 (原因はバグでした) や、バックアップに関する戦略、ローリングアップデートに関する質問など幅広く対応してもらうことができ非常に助かりました。

デメリット

本システムは新規サービスに対する新しいデータベース基盤のため、以前のデータベース環境との比較は難しいですが、クラスタという意味で比較すると、Galera 用のクライアントがないので HAProxy 等のプロキシを挟む必要があることが気になりました。

また、レプリケーション遅延が依然として発生する点はデメリットと感じています。以前のマスタ / スレーブ構成では、レプリケーション遅延が起こったときにデータ

を参照することができないので、その部分に気をつけてシステムを組んでいました。Galera のマルチマスタ構成にすれば、このような問題が発生しないと思っていましたが、ノード間の同期の部分で時折データのズレが発生することがあり、やはりレプリケーションの遅延には気を付ける必要があると考えています。

注意した点

データの書き込み (更新クエリ) の際に、CPU の iowait が発生することがありました。いくつかパラメータを調整することで対処することはできたのですが、その後もディスク I/O の負荷が高くなることがあるので、普段から注意して I/O の状況を見るようにしています。特に、本システムの新しい利用申請があった時は、更新クエリ (Write) の qps をしっかり確認しています。参照クエリ (Read) に関しては qps ベースでは特に注意することはないのですが、レスポンスタイムに幅がある時は、調整することもあります。

全体を通しての感想

全体を通して運用面では人員コストも含め改善することができたので、MariaDB Galera Cluster を導入して良かったと感じています。また、新システムを使用して約 1 年経ちましたが、大きな問題もなく安定して運用することができています。今後もサービス向上のため、より高いパフォーマンスと可用性を求めシステムの改善を行っていきたくと考えております。

※1 MariaDB Enterprise Cluster
MariaDB 社が提供する MariaDB Galera Cluster の公式サポート

※2 Rolling Schema Upgrade
1 ノードずつクラスタから切り離して、テーブルのメンテナンス (DDL) やバージョンアップを実施する方法。ノードを切り離している間も、クラスタに残ったノードでサービスが継続可能。

お客様プロフィール

d w a n g o

株式会社ドワンゴ

URL : http://dwango.co.jp
住所 : 〒104-0061 東京都中央区銀座 4-12-15 歌舞伎座タワー
事業内容 : ネットワークエンタテインメントコンテンツ及びシステムの企画 開発
運用、サポート、コンサルティング

関連プロダクト

Galera Cluster



- ・「完全同期レプリケーション」を実現する、オープンソースのクラスタリングソフトウェア
- ・全ノードが「マスタ」となり、更新・参照の受付が可能 (3 ノード以上の構成を推奨)
- ・データベースとしての基本機能は MySQL と同じ (ストレージエンジンは InnoDB のみ)

※ MariaDB Galera Cluster とは、「Galera Cluster」をベースとし商用環境での利用を想定した MariaDB 独自のクラスタリング製品