

クラウド環境下におけるデータベースの管理システムの検討

西出 彩花[†]

小口 正人[†]

[†]お茶の水女子大学

1. はじめに

近年、爆発的に増加する情報量の処理プラットフォームとしてハイブリッドクラウドの利用が注目を集めている。一般にパブリッククラウドとプライベートクラウドは別の場所に構築され、多くの場合、両者の距離は離れている。従って両者間でデータベースの遠隔同期や遠隔バックアップが行われることになる。一方で、日本は自然災害の影響を受けることが多い。災害発生時には迅速にデータを保護する必要がある。この際に安全にデータを転送するためには、帯域確保のための工夫が必要である。そこで本研究では、データベースの冗長的な遠隔バックアップと、災害時の外部情報をトリガとするマイグレーションによって継続的なデータアクセスを実現するシステムを提案する。

2. 関連研究

2.1 Pangea

ここで本研究で用いる Pangea[1] について説明する。Pangea は、NTT 研究所で開発されている、LAN 環境を前提としたデータベース同期ミドルウェアである。コンシステンシを保ちながら、複数台のサーバでデータベースのクエリ処理を並列実行する事により、性能向上を実現した。これら複数台のサーバは、同一のデータベースイメージを保持しており、同一 LAN 上に接続されている。サーバの 1 台を Leader、その他は Follower としており、クライアントからこのミドルウェアを介してサーバにアクセスをして同期をとる。全てのクエリは照会処理と更新処理に分類され、照会処理はどれか 1 台のサーバで、更新処理は全てのサーバで実行される。更新処理の場合は Leader に対して更新をした後に、Follower に対しても同様に処理を行う。

3. 提案システム

OpenStack[2] を用いて仮想環境上にクラウド基盤を構築し、データベースの遠隔バックアップを動作させて、インスタンスマイグレーションを行うことにより、提案システムを実現する。OpenStack とは、クラウドを構成する仮想マシンや物理サーバの運用管理を実行し、それを効率的に行うためのオープンソースのクラウド構築ソフトウェアである。複数のコンポーネントから構成され、これらのコンポーネントが連携することで IaaS のサービスを提供する。OpenStack の利用者は、KVM など構成されるハイパーバイザ上で動作する仮想マシンに外部ネットワークからアクセスし、CPU、メモリ、HDD、IP アドレス等の計算資源を利用することができる。

4. 実験

4.1 OpenStack への Pangea の導入

クラウド上で Pangea を用いる性能考察を行うために、構築したクラウド環境内に仮想マシンを立ち上げる。OpenStack においては、コントローラノードから指示を出し、この指示に従ってコンピュータード上でインスタンスが起動される。本実験では OpenStack のコンピュータード 4 台のうち、1 台をクライアントサーバ、1 台を Pangea 配置サーバ、2 台をデータベース配置サーバとする。この際、仮想マシンのインスタンスとしては、ダウンロードした Ubuntu14.04 のインストールイメージを元に、80G のディスク領域を使用した Linux OS が立ち上がるように設定する。

データベース配置用のサーバに PostgreSQL サーバをインストールし、クライアントサーバでは TPC-W 用の Tomcat を動かし、Pangea 配置用のサーバ上で Pangea を立ち上げる。今回はこの環境下でデータベース間での転送を行い、スループット値やレスポンス時間を計測することで、Pangea の利用による転送の性能の低下が起こらず Pangea の導入が有用であることを確認する。作成した仮想マシンのネットワークイメージを図 1 に示す。

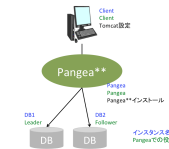


図 1: 仮想マシンイメージ図

4.2 実験結果

前章で紹介した環境下で性能検証を行う。TPC-W における仮想ブラウザ数 EB（クライアントからの負荷量）を変化させた時のスループット値は EB 数 600 の時、78.24 であった。インスタンスを介さず物理サーバ上で直接同じ実験を行った結果である EB 数 1000 の時の 126.78WIPS と比較すると、最大スループット値が 39%低下するなどの大きな性能の低下がみられる。この要因として、今回の測定環境ではメモリサイズの大きさが仮想マシンと物理サーバで大きく異なることが考えられる。これは今後の測定の際に仮想マシンと物理サーバのメモリサイズを調整することで、正確な比較ができるようになる事が期待される。

また、仮想化によるオーバーヘッドも性能低下の要因の一つであると考えられる。

そこで、本研究ではシステム中のどの部分を物理サーバに残すべきかを検討した。クライアントサーバ用のサーバは、実際に企業で使われる際にほとんどの場合でクラウド

A proposal for database managing on the cloud computing system

[†] Sayaka NISHIDE, [†] Masato OGUCHI
Ochanomizu University ([†])

上で作成される仮想マシンである。これを物理サーバ上での運用することは非現実的である。

そのため、Pangea 用のサーバかデータベース用のサーバを物理サーバ上で運用することを検討する。

Pangea 用のサーバは、仮想マシン上であるクライアント用サーバからクエリを受け取り、それを自サーバ内で分類し、データベース用サーバに転送する。この動作内で、仮想化によるオーバーヘッドを考えることは難しい。

ここで、データベースの転送の際に最もボトルネックとなるのは、仮想マシンのデータベースがノード自身の物理データベースへアクセスする際にコストがことである。そのため、よく使うデータベース上のデータはバッファプール上にまとめておいた方がアクセスが速くなると考えられる。

これは、ノード 11 に共有用ストレージを置き、NFS マウントを利用して構築した本システムでライブマイグレーションをした際の所要時間が移動先ノードにより大きく異なったことからわかる。これを図 2 に表す。縦軸は転送先ノードである。

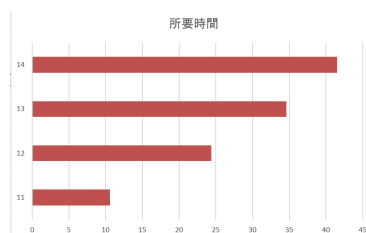


図 2: ライブマイグレーション所要時間

また、Postgresql において、デフォルトテーブルスペースには、通常のデータベースに加えて管理情報など、様々な付随データが格納される。その中でも、トランザクションログはデータベースの全ての更新内容をログとして保存し続けるため、データ量が非常に多い。このトランザクションログがデフォルトテーブルスペースに格納されるため、ハードディスクは I/O 負荷が非常に高い。このため、新たなデータベース用のテーブルスペースとしてシステム上のデータベース部だけを物理サーバに置くことで、デフォルトテーブルスペースにはできるだけデータベースを作らずに、トランザクションログのための I/O 負荷をクラウド上のデータベースに、データベースアクセスのための I/O 負荷は物理サーバ上に配置することで、I/O 負荷の分散が期待される。このイメージを図 3 に示す。

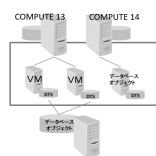


図 3: データベース共有のイメージ

これらより、データベース用のサーバを物理サーバ上におくことが有用であると考えた。これは、データベースを物理サーバ上に置くことで、緊急時以外にクラウド上に重要なデータを置かず済むため、セキュリティ面を考えても有用なシステムだと考えられる。

この時、最大スループット値は EB 数 900 の時に 122.1WIPS となり、システムがすべて物理上に置かれた時と比べても約 4% の性能低下に抑えることができた。本研究で行った 3 種類の環境での実験結果を図 4 で比較した。この結果から、仮想マシンが用いるデータベースを外部の物理サーバ上に置くことで、インスタンスをクラウド上に置く際におこる仮想化によるオーバーヘッドを軽減することができると思うことができる。なお、各データベースをクラウド上に配置するのか物理サーバ上に直接配置するかは、そのデータベースが用いるデータの特性に依存する。

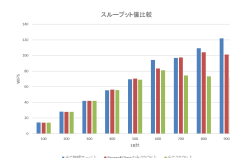


図 4: システム構成によるスループット値の比較

5. まとめと今後の課題

データベースの転送手法について、OpenStack を用いて構築した複数の環境で性能比較を行った。本稿では、データベースサーバ内の全てのデータを同期することを前提としている。しかし、実際にハイブリッドクラウドとしてクラウドを利用する際にはセキュリティの問題などから、パブリッククラウドとプライベートクラウドに保存されるデータは一部異なることが予想される。そのため、要求されたクエリがパブリッククラウドからのアクセスが認められたデータであるかどうか Pangea 内で区別することにより、テーブルごとのデータ管理を可能としていきたい。これは、Pangea がパケットを元にクエリの種類の判断をしていることから、可能であると判断される。

謝辞

本研究を進めるにあたり、N T T 研究所 細谷柚子様に数多くの助言を賜りました。深く感謝いたします。

参考文献

- [1] T.Mishima and H.Nakamura : "Pangea:An Eager Database Replication Middleware guaranteeing Snapshot Isolation without Modication of Database Servers", Proc.VLDB2009,pp.1066-1077, August 2009. PVLDB2009.
- [2] OpenStack : <http://www.oprnstack.org/>