

ハイブリッドクラウドにおける データベース同期システムの一検討

細谷 柚子[†]三島 健[‡]小口 正人[†][†]お茶の水女子大学[‡]NTT ソフトウェアイノベーションセンタ

1 はじめに

近年、パブリッククラウドやプライベートクラウドが普及しつつあり、その2つをシームレスに結合するハイブリッドクラウドが注目されつつある。パブリッククラウドでは、スケールアウト/スケールダウンできることによりコスト削減や技術面のリスク軽減に繋がる。一方で、社外のサービスを利用することによるセキュリティへの不安も残るために、パブリッククラウドを使いながらも、1部の個人情報や社外秘情報などは社内システムであるプライベートクラウドを利用するというようにハイブリッドクラウドの導入により両方のメリットを得ることができる。

しかし、実社会においては、ハイブリッドクラウドの導入はあまり進んでいない。これは、2つのクラウドを併用するには、データの一貫性をとらなければならないという技術的な問題があるためである。データベースサーバは企業の基幹を構成しているため、クラウドで動作させるべき重要度の高いシステムである。従って、本研究ではハイブリッドクラウドでデータベースを同期させることに注目した。

LAN環境を前提としてデータベースサーバを同期する Pangea[?] という既存のミドルウェアがある。これをハイブリッドクラウドに適用し、TPC-W ベンチマークを用いて評価実験を行った。実験では東京一大阪間を想定した RTT16ms と日米間や日欧間を想定した RTT256ms の場合を測定した。高遅延の測定は、外資系の企業や、日本企業の海外進出による海外支店が増えたことで、日本と海外でのデータの同期が必要になってくると考えたためである。また、日本のように自然災害が多い国では、近隣にバックアップを置いておくと、大規模災害が起こった際には両方とも失ってしまう恐れもあるため、遠隔バックアップの需要が高くなっていることも考えられる。その評価をもとに、ハイブリッドクラウドでも有用な新しい方式を検討する。

2 Pangea

Pangea は、LAN環境を前提として複数のデータベースサーバ間でスナップショット分離を保証するデータベースレプリケーションミドルウェアである。クライア

ントからサーバ側に直接アクセスするのではなく、ミドルウェアを介してデータベースにアクセスする。データベースを改造することなく、サーバを増やすことで性能を向上させることが可能である。Pangea では照会処理は1台のレプリカで、更新処理は全レプリカで処理される。クライアントからミドルウェアまでの処理を Global transaction、ミドルウェアからサーバまでの処理を Local transaction といい、全レプリカでの Local transaction がコミットされ次第、Global transaction のコミット処理が完了する。また、レプリカの中の1台を Leader、その他は Follower とされ、更新処理の場合は Leader に対して更新をした後に、他の Follower に対しても同様に処理を行う。本研究では Pangea をハイブリッドクラウド環境に適用し、Pangea の拡張を検討していく。

3 基本性能評価

3.1 実験環境

Pangea の基本性能評価を行った。データベースサーバ用にマシン2台を用いた。どちらのマシンも表1に示すスペックである。データベースサーバには PostgreSQL9.2.6 を使用した。クライアントとデータベースサーバの間に Pangea を接続して同期を行う。Webサーバとアプリケーションサーバには Tomcat6.0.37 を用いた。性能評価は TPC-W ベンチマークを使用した。TPC-W は仮想的なブラウザ(以下 EB とする)が、データベースにトランザクションを発行する。TPC-W には3種類ワークロードがあるが、それらの違いは図1に示すように照会処理と更新処理の割合が異なる。性能評価指標は、スループット(1秒あたりのWEB画面表示:(WIPS))とレスポンス時間(1画面データの転送時間:(秒))とした。

また、パブリッククラウドは遠隔地にあることを想定して dummy net を使用して人工的に遅延を起こした。今回は、東京一大阪間を想定した 16ms と日米間や日欧間を想定した高遅延の 256ms の場合で実験を行った。実験環境を図2に示す。

OS	Linux 2.6.32
CPU	Intel(R) Xeon(R) CPU @ 1.60GHz
Memory	2GByte

表 1: 2 台のデータベースサーバのスペック

A Study on Database Replication in the Hybrid Cloud

[†] Yuzuko Hosoya, Masato Oguchi[‡] Takeshi Mishima

Ochanomizu University (†)

NTT Software Innovation Center (‡)

	Read - only	Update
Browsing mix	95%	5%
Shopping mix	80%	20%
Ordering mix	50%	50%

図 1: ワークロード

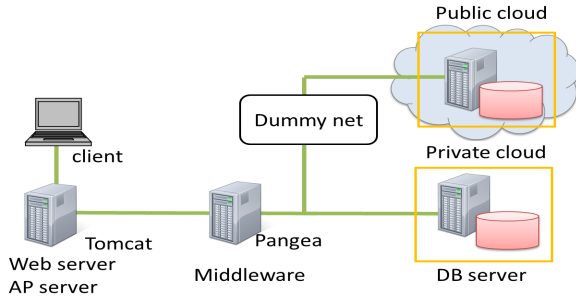


図 2: 実験環境

3.2 基本性能

本章では、レプリケーションに最も負荷が大きい ordering mix での性能を測定した。ローカル環境における WIPS 値と WIRT 値を図 3 に示す。上の折れ線グラフはスループット、下の折れ線グラフはレスポンス時間を表している。EB が 290 のとき最大スループットが 38.9WIPS となる。EB 数を増やしていくにつれ、スループットも上がっていく。レスポンス時間は指数関数的に上昇していく。レスポンス時間が上昇していき、1 より大きくなるあたりでオーバーロード状態になるため、EB 数を増やしてもスループットは下降していく。



図 3: ローカル環境における ordering mix の性能

3.3 クラウド環境における性能

dummy net を使って遅延を入れた場合の ordering mix における測定を行った。近隣の街にバックアップを置く場合と、海外のような遠隔にバックアップを置く場合を想定して RTT16ms と 256ms で評価した。ローカル環境との比較を図 4 に示す。

遅延を入れた場合においても、スループットとレスポンス時間の変化はローカル環境と同じく、EB 数を増やしていくにつれスループットも大きくなり、レスポンス時間は指数関数的に上昇する。ある点において最大のスループットとなり、その後は EB 数を増やしても下降し

ていく。RTT16ms における最大スループットは EB 数が 290 のときで 38.8WIPS であり、RTT256ms における最大スループットは EB 数が 300 のときで 36.1WIPS である。RTT16ms の場合は、ローカル環境とスループット、レスポンス時間もほとんど差がないことから、性能の低下はみられなかった。RTT256ms の場合は、ローカル環境と比較するとスループットが下回り、レスポンス時間がほとんどのケースで 1 秒を超えていることから、高遅延環境では性能低下が無視できない。

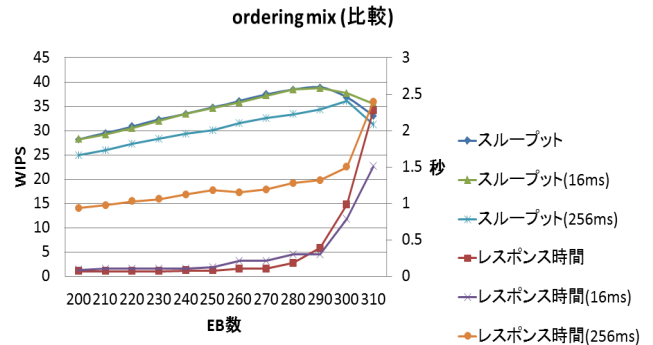


図 4: ordering mix の性能比較

4 まとめと今後の課題

データベース同期システム Pangea についての基本性能評価を行った。ordering mix におけるローカル環境でのスループットを測定し、EB 数が増えればスループットが上がることを確認できた。また、EB 数が 290 のときに最大スループット 38.9 であることが確認できた。EB 数が多すぎればリクエスト数も膨大になり飽和状態になるため、最大スループットを超えた後に下降することは、妥当だといえる。

クラウド環境を想定し、遅延を起こした場合の測定も行った。ローカル環境と比較して、スループットの低下がみられた。これは、遅延を入れることで、距離がボトルネックになることが要因だと考えられる。

今後の課題としては、他の 2 種のワークロードでの測定、比較も行いたい。その結果からクエリの種別やネットワーク遅延を考慮に入れたハイブリッド環境における Pangea の効率的なスケジューリングを検討したい。

参考文献

[1] T.Mishima and H.Nakamura, "Pangea: An Eager Database Replication Middleware guaranteeing Snapshot Isolation without Modification of Database Servers", PVLDB2009.