

JB プロジェクト活動報告

山本 成一 (yama@wide.ad.jp) 島 慶一 (shima@wide.ad.jp)

関谷 勇司 (sekiya@wide.ad.jp)

平成 24 年 12 月 26 日

1 概要

JGN-X は、独立行政法人 情報通信機構 (NICT) が運営する超高速・高機能研究開発テストベッドネットワークである。JGN-X ではさまざまな組織がプロジェクトの単位で実験利用を行なっている。WIDE プロジェクトでは JB プロジェクトの名前で JGN-X 資源を活用し、次世代インターネット基盤として必須となる、仮想化技術を中心に研究開発活動を行った。

2 研究内容

JGN-X は広く国内外に設置された拠点に、分散コンピューティング環境を構築しており、ネットワークテストベッドに付随した資源として、これらのコンピューティング資源 (VM, ストレージ等) の貸し出しを行なっている。WIDE プロジェクトでは、次世代のインターネット基盤として、今後も高度化していく計算機の仮想化を重視し、クラウドワーキンググループを設立している。JB プロジェクトではクラウドワーキンググループでの研究利用に先立つ、試験的な環境基盤を構築し、知見収集をしている。以下に代表的な活動を記す。

3 広域分散ストレージの利用実験

仮想化技術の貢献により、これまでは現実的と考えられていた数の計算機資源をインターネットサービスに導入運用できるようになってきた。計算機の仮想化の目的は、計算機機能の重畳による物理計算機資源の有効活用である。そのため、仮想計算機がデータセンターの負荷や拡張状況に応じて、高い自由度で再配置できることが重要な課題となる。今後、異なる場所に配置されたデータセンター間での仮想計算機移動なども考慮すると、広域運用可能な仮想計算機基盤の構築が必要になる。

仮想計算機の再配置において、ストレージ資源の透過的アクセス技術は重要な要素のひとつである。本章では、分散ストレージ技術のひとつである GlusterFS の、広域ネットワークでの運用可能性について検討した結果を報告する。

3.1 実験トポロジ

GlusterFS は、ストレージクラスタを構成するノードの台数にも依存するが、ストライピング、レプリケーションの柔軟な組み合わせ構成が可能となっている。今回は、4 台の

表 1: 拠点間の RTT (単位:ms)

	岡山	堂島	東京
堂島	3.2	—	—
東京	11.5	8.6	—
小松	18.1	15.2	7.2

ノードを用いて、単純なレプリケーション（複製数 2）の構成で実験を実施した。

実験を実施したネットワークポロジを図 1 に示す。構成の意図としては、なんらかのサービスを運用している状況で、そのデータストレージを広域に分散配置し、どちらかの拠点到障害が発生した場合でも継続運用が可能になる仕組みの実現性を検討するためである。サービスの例としては、仮想計算機のディスクイメージストレージとしての利用や、Dropbox に代表されるクラウドストレージサービスなどのバックエンドを念頭に置いている。

それぞれの GlusterFS サーバーノード (vm1.okayama、vm5.dojima、vm5.nezu、vm1.komatsu) を図に示した通り地理的に分散配置し、GlusterFS がデータ保管用として利用するローカルディレクトリを設定した (/var/glusterfs-brick/ディレクトリ)。この 4 箇所のディレクトリを GlusterFS のレプリケーション機能を用いてひとつの GlusterFS ボリュームとして設定し、それぞれのサーバー上の /var/glusterfs/ディレクトリにマウントして運用している。拠点間の RTT を表 1 示す。

3.2 性能計測

構築した GlusterFS で Bonnie++¹ を用いた性能計測を実施した。Bonnie++ は東京のノード (vm1.nezu) で実行した。比較のため、

¹<http://www.coker.com.au/bonnie++/>

vm1.nezu と同じセグメント上に配置してある NFS サーバを vm1.nezu でマウントしたファイルシステム上での計測、および小松ノードと同じセグメントに配置してある NFS サーバを vm1.nezu でマウントしたファイルシステム上での計測も実行した。図 2 に結果を示す。なお、東京の NFS サーバは PC ベースの NFS サーバ、小松の NFS サーバは EqualLogic を用いている。

図から、GlusterFS は NFS と比較して書き込み性能が劣っていることがわかる。原因のひとつは、今回の GlusterFS の設定が複製モード（複製数 2）になっていることであると考えられるが、それを考慮しても特にキャラクター書き込みの際の性能の低下が著しい。これに比較して、読み込み性能は NFS と遜色ない。キャラクター読み込みに至っては、ローカルセグメントに配置された NFS サーバよりも高い性能がでている。

3.3 まとめ

今回、広域環境におけるストレージシステムの運用可能性を検討するため、GlusterFS を用いた実地検証を実施した。その結果、GlusterFS はキャラクター書き込みにおいて著しい性能低下が見られるものの、読み込み性能に関しては NFS と同等以上の性能が得られることがわかった。すなわち、小さな単位での書き込みを抑制し、適切なバッファリングを採用すれば広域での GlusterFS の運用は現実的だと考えられる。今回、特定のアプリケーションを想定せずに、Bonnie++ による純粋なディスク I/O 性能を計測したが、上記の挙動から、アプリケーションによって大きく体感品質が変わることが予想される。今後、様々なアプリケーションを試用することで、性能の低下を押さえつつ、広域で信頼性の高いファイルシステムの運用方法を模索してい

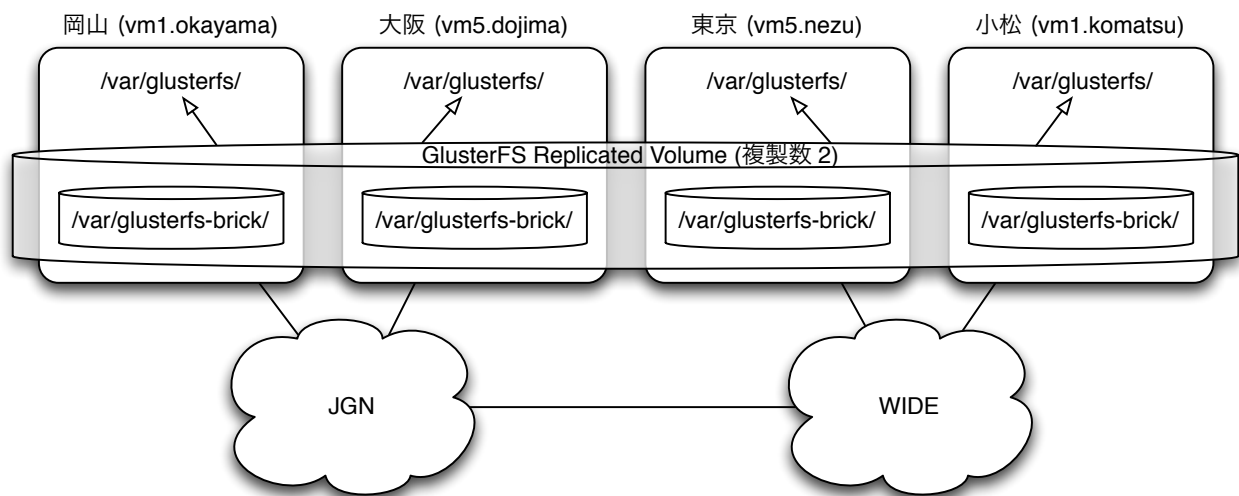


図 1: 広域 GlusterFS 運用実験トポロジ

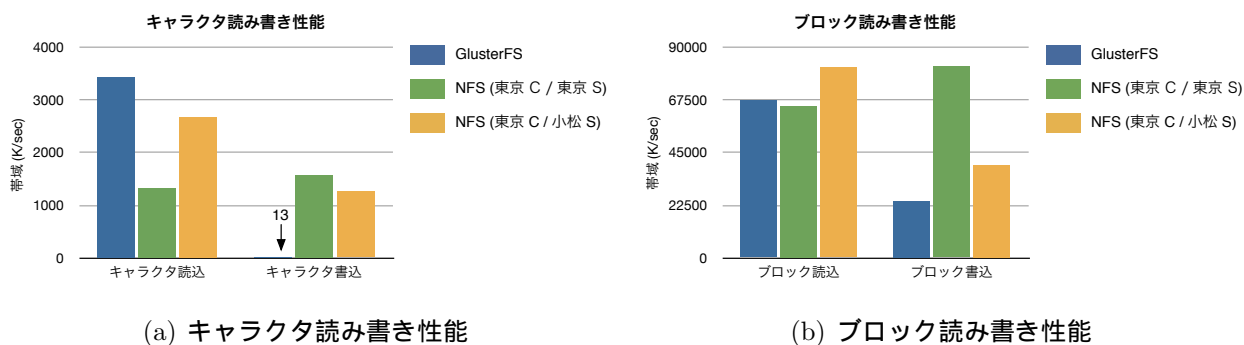


図 2: GlusterFS と NFS の性能比較

く予定である。

4 今後の予定

JGN-X を利用した大規模テストベッド環境を展開することで、多様な研究者や利用者が所属する研究組織ならびに教育組織を相互接続し、総合的で実践的なネットワーク環境における次世代インターネット技術の検証と評価を行うことができる。特に本プロジェクトで対象としている仮想化技術は他に展開される技術の基盤となる重要な要素である。引き続き、JGN-X 等のテストベッド環境を利用して、各分野における次世代の研究開発

リーダの育成・養成を視野に入れ、総合的かつ実践的なテストベッド環境を構築し、グローバルな協調関係の維持をしながら研究活動を推進させたい。