



MIC Associates, Inc.

グリッドストレージとデータマネージメント

エムアイシー・アソシエーツ株式会社

2009/07/22

ここに記載された内容は更新される可能性があります。この文書に記載されている内容はこの文書の発行時点におけるエムアイシー・アソシエーツ株式会社の見解を述べたものです。エムアイシー・アソシエーツ株式会社が、この文書に記載された内容の実現に関して確約するものではありません。また発行日以降については、この文書に記載された内容の正確さは保証しません。

この文書は情報の提供のみを目的としており、明示的または黙示的に関わらず、この文書の内容について エムアイシー・アソシエーツ株式会社はいかなる保証をするものでもありません。

エムアイシー・アソシエーツ株式会社は、本書に記載してあるすべて、または、一部の記載内容に関し、許可なく転載、または、引用することを禁じます。

バージョン	作成日付	旧バージョンからの 変更点	総ページ数
1.0	2009/07/22	新規	8

本書作成、編集、管理



エムアイシー・アソシエーツ株式会社
〒103-0004 東京都中央区東日本橋3-
12-12
櫻正宗東日本橋ビル9F
Tel. 03-5614-3757 Fax. 03-5614-3752

目次

クラウド時代のストレージグリッド化

データマネージメントの必要性

データのバックアップの課題

データのリプリケーション

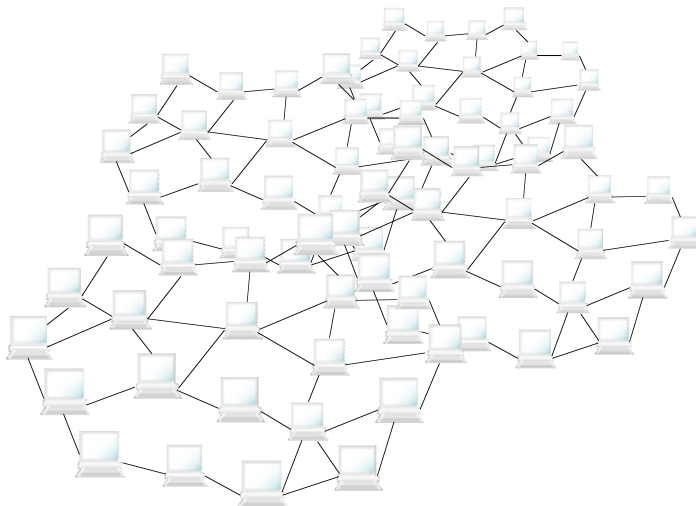
データマイグレーション

階層化ストレージ管理

まとめ

クラウド時代のストレージグリッド化

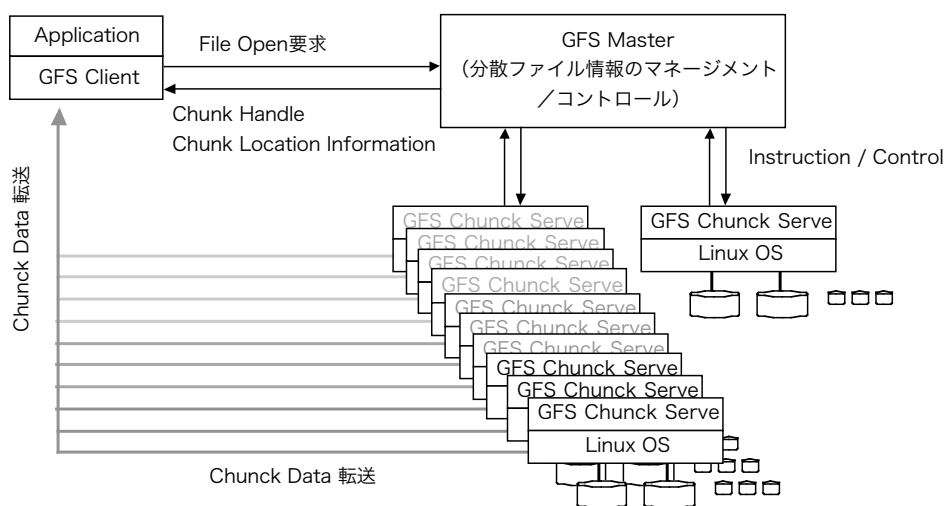
近年、ITの話題でコンピュータのクラウド化という言葉をよく聞くようになりました。また、これらの技術に関する多くの文献を目にするようになりました。今日、クラウドコンピューティングはまさに雲の様に
つかみ所の無いネットワークの中で、ユーザは最低限のコンピュータ資源により、リモートサイトにあるコンピュータ群とインターネットで接続し、その資源を利用して必要業務を行ない、場合により、データもそのまま預けるコンピュータサービスの形態を
指す様になりつつあります。



高速インターネットが当たり前になってきた今日、Google等が提供する各種のサービスアプリケーションを使用して日常業務を行なっても、それほどストレスを感じることはありません。また、同時にこれらのサービスには一定の料金体系でデータそのものを預けるストレージサービスが提供されている場合が殆どです。これらのクラウドコンピューティング

でのデータ保存の場合も、グリッドストレージと呼ばれる複数のネットワークストレージを仮想化した論理的ボリューム空間をクライアントアプリケーションに提供するAPIを介して使用されるサービスが一般化しつつあります。

この背景の技術として有名なのはGoogleですが、独自開発のGoogle FS(GFS)という分散型のファイルシステムにより、ストレージのメタデータベースを管理するマスターと呼ばれるネットワーク上のサーバ群と、データ（チャンク）を保存するネットワークストレージから構成されています。(右図参照)



グリッド、クラウドと呼ばれるコンピュータ技術の進化は大容量のデータを水平的なネットワーク内の仮想的なボリュームに保存し、サーバが提供するメタデータ情報に基づき、データアクセスサービスを提供しま

す。例えば、YouTubeは毎日10TBのデータがアップロードされるということです。この様に、インターネットに接続されたネットワークストレージを仮想的に論理的なストレージボリュームとして利用する技術は、今後数年間でデータ量が現在の10倍以上に増大することが予測される中、ストレージ技術において現在考えられるの唯一のソリューションと言えます。

このGoogleのGFSは一種のストレージ仮想化技術にあたります。SNIA(Storage Network Industry Association)では仮想化技術を以下の様に定義しています。

1. ストレージやサービスの内部構造を普通のコンピュータや、ネットワーク、アプリケーションから抽象化したり、隠蔽したり、隔離したりすることで、ストレージやデータをネットワークやアプリケーションから独立した管理ができることを目的にする。
2. 機能やデバイスを統合し、それらの複合体を隠蔽し、下位レベルのストレージ資源を自由に追加することができるストレージサービスやデバイスに対する仮想化アプリケーションを指します。

以上のストレージの仮想化技術はストレージやコンピュータ、ネットワークといった通常使用されるITハードウェアから切り離され、クラウド、グリッドとして抽象化され、更にはサービスプロバイダが提供するストレージサービスそのものもその一部と見なされつつあります。

データマネージメントの必要性

このようにネットワークに接続されたストレージは今後数年間で現在のデータ量の10倍を超える勢いで増大するデータを保存する上で不可欠な存在であり、今後、これらのストレージをグリッド化する技術はますます進化するものと予想されます。このようにネットワーク上で水平に分散された大容量データをその保存と利用、さらには、データの保存媒体の移動等を管理するデータマネージメントが重要な意味をもっと来ることにも充分予想されることです。たとえば、非構造化データの場合、データ容量そのものは構造化データと比較して非常に大きくなります。このデータをネットワークストレージに収納し、もし、再利用されことなく長期間そのまま保存するとなると、ストレージ資源の有効利用という観点からは大きな問題といえます。そこでデータそのものをどのように取り扱うべきかのポリシーを決め、管理するデータマネージメントが必要になります。このデータマネージメントに課せられた課題として、データの安全な保存、万一の場合のデータ修復、更に、使用頻度、更新と更新の間の期間に応じたデータの保存場所、期間の管理、データとしてオンラインアクセス可能期間終了後のストレージからのデータ削除、オフライン化した場合の収納メディアのラベルデータベース管理、収納ロケーション管理等々があり、大容量のストレージ、大容量データを管理する上でこれらの課題を解決する必要があります。

データのバックアップの課題

