

大規模データ処理に対する OSS Hadoopの活用

高橋 千恵子・世良 直彦
宍元 賢治・大崎 寛達

要旨

近年、ネット企業を中心に「ビッグデータ」に対する処理技術の開発が進められてきました。Apache Hadoopはそれらの技術を基に、ペタバイト級の巨大データの蓄積とそれに対する分散処理をスケールアウトで実現するOSSとして注目を集めています。NECでは、Apache Hadoopについて検証を行い、その特性に合わせた構築を行っています。また一般に難しいといわれるApache Hadoopのサイジングについて、シミュレーション技術により予測可能とする技術を開発しており、その内容についても紹介します。

キーワード

- ビッグデータ ●分散処理 ●スケールアウト ●オープンソース ●OSS ●Hadoop
- シミュレーション ●サイジング ●CASSI

1. まえがき

「ビッグデータ」、「オープンソースソフトウェア(OSS)」は、ここ最近IT分野で注目されているキーワードです。NECでは、ビッグデータの領域で注目を集めているOSSの1つ「Apache Hadoop(以下、Hadoop)」に関して、かねてより技術検証を実施してきました。これまで世の中で使用されてきた分散処理基盤や大量データを扱うDWH(Data Warehouse)などの技術とどこが違い、どのようなメリットがあるのかを明確にすることが目的の1つでした。

本稿では、まず第2章でHadoopの概要と活用できる領域を明確にします。続く第3章で、Hadoopの適用事例と実際の構築時に考慮する点などを説明します。そして最後に、Hadoopの性能を最大限に引き出すことを可能としたシミュレータを紹介します。

2.1 Hadoopの設計思想

Hadoopは大規模分散処理を実現するために、次のような設計思想で開発されています。

・スケールアウトによるリニアな性能拡大

Hadoopは複数のノードでの並列分散処理により性能拡大を行うスケールアウト型の処理システムとして設計されています。一般的にスケールアウト型の処理系の特長としては、サーバ増設により性能と価格がリニアに向かし、性能限界の克服と費用の急激な上昇を抑えられる点が挙げられます(図1)。

しかし、現実には増加していくサーバ間の通信量がボトルネックになり、いずれ増設サーバ台数の限界を迎えます。対してHadoopのフレームワークは、サーバ間通信量の分散、マスタサーバの負荷軽減と集中排除などによりボトルネックを回避し、より巨大なクラスタ(サーバの集合体)の構築を可能とします。

・分散処理を容易に実現するフレームワーク

大量のデータ処理をスケールアウト型のサーバクラスタで効率的に行うには、サーバ台数を生かした分散並列処理が有効です。Hadoopのフレームワークでは「MapReduce」と呼ばれる処理方式を用いて開発を行うことにより、自動的に適切な分散処理をサーバクラスタ内で実行します。処理プログラムはHadoopを構成するス

2. Hadoopとその実力

Hadoopは2004年にGoogle社が発表した分散処理技術「MapReduce」及び「Google File System」の論文を元にOSSとして開発されたフレームワークです。ビッグデータの現実的な処理手段として世界的に注目を集めています。

レーブサーバに自動的に分配され、分散処理を行い集計が行われます。

・非構造化データ処理可能

RDB (Relational Database) などの構造化したデータを取り扱う処理プログラムに比べて、分散処理のフレームワークであるHadoopは格納したデータ群に対してさまざまなプログラムを実行できるため、ドキュメントやログのマイニング、画像データの形式変換など非構造化データの変換/解析処理に適しています。

・汎用サーバの大量採用を可能とする耐障害性

数千台にも及ぶサーバクラスタを構成した場合、サーバやディスク装置の故障は確率的に「例外」ではなく常に発生し得るものと考えるべきです。Hadoopはデータ格納時、自動的にレプリケーション（複製）を作成しており、サーバやデータブロック、割り当てた処理プログラムに問題が生じた場合、別のサーバにあるデータレプリケーションに処理を割り当てます。これにより、サーバやディスク装置の障害によるシステムの停止を回避しています。またサーバやディスク装置の故障により欠落したデータについては、別の正常なレプリケーションからのコピーが自律的に他の正常なサーバに作成され、復元されます（図2）。

・スループット重視

一般的なRDBの設計思想は、ロックやキューなど複雑な機構による一貫性の保持を行なうながら低レイテンシの確保を目指します。一方Hadoopは、複雑な処理機構を廃するためデータの更新は行わず、常に新しいデータの書き込みを前提とします（Write Once Read Many）。また、データを格納する大量のスレーブサーバから並行してデータの入出力をを行うため、巨大データの入出力に関してスループットの提供が可能です。

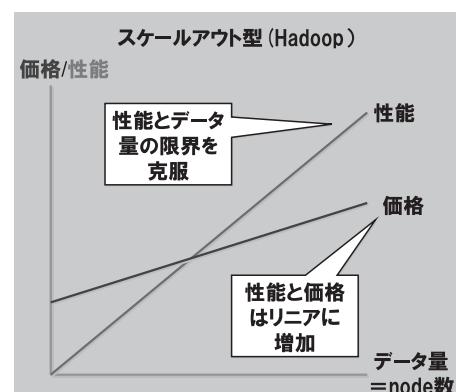


図1 スケールアウト型処理系の特性

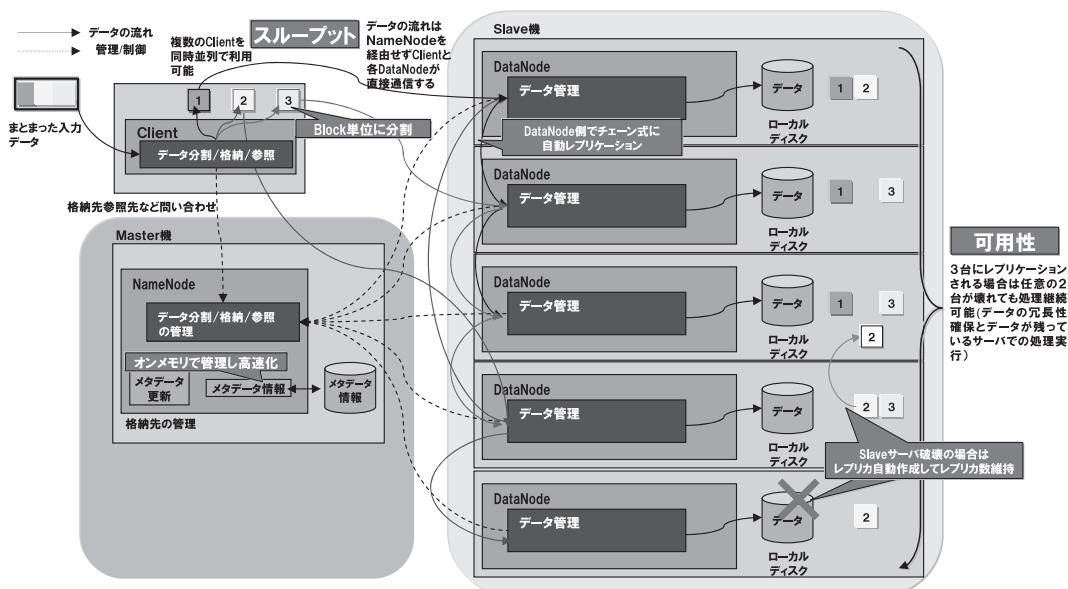


図2 データレプリケーションと自動復旧

2.2 Hadoopの活用分野

Hadoopは主に次の3つの分野で活用され始めました。

- **DWH強化**

Webサイトに関するアクセスログ分析、データマイニング、リスク分析などDWH処理に適用可能です。この分野では既存のDWHの前段で大量データ蓄積やデータ加工、処理対象データのフィルタリングを行うことにより、既存のDWHソリューションが扱えなかった膨大なデータを扱うことが可能となり、分析精度の向上が図られた事例が出始めています。

- **ビッグデータ対応**

これまでのシステムでは処理できなかった大量のデータの処理に適用可能です。ライフログやM2M (Machine to Machine) 技術の進展と相まってセンサデータ、全世界でのカードの利用記録など既存技術では解析が難しかった分野での利用が始まっていきます。

- **基幹系バッチ処理**

これまでメインフレームなどで処理されてきたバッチ処理の一部を、スケールアウト型に置き換えることができます。スケールアップ型性能の限界により、これ以上データ量の拡張や処理時間の短縮が不可能であった処理をHadoopに切り替えることで、スケールアウト型の処理として更に拡張やバッチ時間の短縮が可能です。

3. 大規模分散処理システムを実現する Hadoopソリューション

海外においては、大規模分散処理システムを実現するためのソリューションとして、既にいたるところで商用利用されてきたHadoopですが、日本国内においても試験運用の段階から商用ソリューションとして活用される段階に移行しています。

3.1 Hadoopソリューション適用事例

- **大規模ログ検索システム**

ネット事業者や通信キャリアでのアクセス記録、通信記録などのログデータは日々刻々と動的に発生し、膨大な量となります。その膨大なデータを保存して検索を行うことは、従来のシステムでは難しい処理でした。このよ

うなシステムにHadoopを適用することによって、スケールアウトによるデータ量増大への対応と、分散処理(MapReduce)による検索時間の大幅短縮を実現しました。データ量は導入当初から増大を続けており、定期的にマシンを追加してスケールアウトを行っています。

- **センサデータ集計システム**

日本全国各地に設置したセンサからの入力情報を元にセンサデータを集計、分析して有用なデータを短時間で生成し、提供可能にするシステムです。このシステムはHadoop上にHBaseというカラム型データベースを搭載しており、従来のRDBでは難しかったスケール可能なデータベースとHadoopのMapReduceによる大規模分散処理を同時に実現しています。

HadoopとHBaseの組み合わせは自由度が高いため、センサ数の増加、収集するデータ種別の追加、分析観点の追加などに対して、データ構造再構成なしに柔軟に対応することができます。

- **統合認証管理システム**

従来、RDBで管理していた統合認証管理システムの一部の処理を、HadoopとHBaseの組み合わせにより実現しました。

RDBでは構造化データを前提としているため、ユーザー情報の追加などの変更を苦手とする場合が多いですが、HBaseはカラム型データベースのため、ユーザー情報の変更に強いというメリットがあります。また、Hadoop上にデータベースを構築しているため、ユーザー情報に関する分析や抽出、組織の体制変更や異動による一斉変更をMapReduceにより短い時間で実現することもできます。

3.2 Hadoopソリューション構築の実際

Hadoopソリューションの構築に当たって考慮すべき点、実際に改善を行った点について紹介します。

- **Single Point of Failure対策**

一般にHadoopでは、スレーブサーバと呼ぶ大多数のマシンの冗長性を考慮しており、サービスを停止することなくハードウェア故障などの障害に対応しています。しかし、Hadoopシステム全体を管理するマスタサーバは、Single Point of Failureであり、マスタサーバを冗長化するための仕組みが必要になります。弊社ではクラスタリ

ングソフトウェア「CLUSTERPRO」と組み合わせることにより、マスタサーバダウン時にもサーバを切り替えてサービスを継続するシステムを構築しています（Hadoopの最新バージョンではマスタサーバの冗長化について改善しつつあります）。

・運用監視の考え方

Hadoopは多くのマシンを使用するため、ハードウェア故障などの障害が起きる可能性も高くなります。個々のサーバが正常に動作しているか、またシステム全体のリソース消費傾向がどうなっているかなどの情報は逐次監視を行い、必要な対処を早めに取る必要があります。監視項目にはOSに関する一般的なリソース情報の他に、Hadoopが独自に出力する監視項目があります。

・自動構築ツールの活用

マシン数が数十台規模に及ぶ場合、それぞれのマシンの構築を手動で行うと構築期間の長期化や、作業ミスが発生しやすくなるため、対策が必要となります。そのため、OSSのPuppetやChefといった自動構築用の仕組みを使用します。

・チューニング

システムが扱うデータの特性や、データに対して実行する処理、その他システムに要求される要件によっては、パラメタチューニングにより、より効果的な運用が可能となります。例えば、ハードウェアの持つ資源を最大限に利用して性能を引き出すためのチューニングや、ジョブへの影響を低減してシステム全体の安定稼働を助けるためのチューニングなどがあります。

4. 実測補正型性能評価シミュレーション技術による Hadoop性能評価シミュレーション

4.1 Hadoop性能評価シミュレータ導入の背景

Hadoopの特長の1つに、サーバ台数に対してリニアに性能を向上させることができる点がありますが、第3章第2節で説明したとおり、その特長を最大限に引き出すためには、Hadoopを適切にチューニングすることが必要です。しかし、Hadoopの設定は複雑で、サーバ台数も多くなるため、机上計算で適切な設定を導出することは大変困難です。また、複雑で大規模なHadoopシステムを実環境で検証するためには、多大なコストと時間が必要となります。

この問題を解決するため、SI作業をコンピュータ支援するCASSI（Computer Aided System model based SI environment）を用いて、Hadoopの性能評価シミュレータを構築しました。この性能評価シミュレータを用いることで、Hadoopの性能を最大限に引き出す設定を、効率よく導き出せるようになりました。

4.2 CASSIの実測補正型性能評価シミュレーション技術

一般に、システムの性能評価シミュレーションではシミュレーション精度と処理時間が問題となります。CASSIでは、実測補正型性能評価シミュレーション技術^②を研究開発し、性能モデルに基づき算出した理論値を実環境で測定した実測値で補正することでシミュレーション精度を向上させるとともに、実測値を利用した計算簡略化によるシミュレーションの処理時間の高速化を実現しています。

4.3 Hadoop性能評価シミュレーションの評価

Hadoop性能評価シミュレーションでは、HadoopのMapReduceの応答時間、Map処理などのスループット、CPU・ネットワークなどのリソース使用率を算出することができます。

シミュレーションの精度については、MapReduce処理中のCPU使用率のシミュレーション時間と実測値の比較（図3）から分かるとおり、CPU使用率の変動を高精度にシミュレーションできており、全体の応答時間の誤差も約7分間の処理に対して2.5%（約10秒）程度に収まっています。

4.4 現状の成果と今後の予定について

以上のとおり、実測補正型性能評価シミュレーション技術を研究開発することで、Hadoopのような複雑で大規模なシステムにおいても、高精度の性能評価シミュレーションを机上で容易に実施できるようになりました。このHadoop性能評価シミュレータを活用することで、従来、コスト・時間の制約で実施が難しかったシステムの性能限界値の確認など高度な性能評価も、低コスト・短時間で容易に実施できるようになります。

実測補正型性能評価シミュレーションは、Hadoopだけでなく、Web3層システムなどへの適用も既に実現しており、Hadoopと同様の成果を上げています。今後も、さまざまなシ

データ管理/処理基盤

大規模データ処理に対する OSS Hadoopの活用

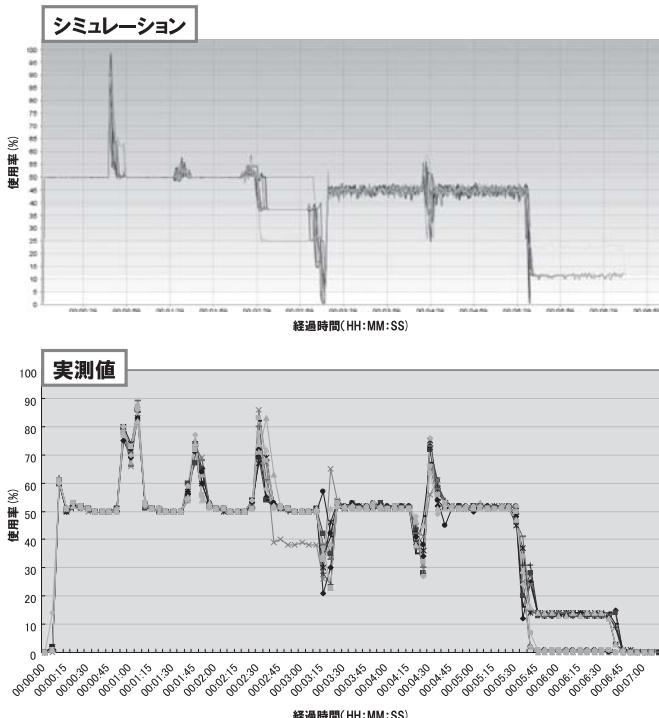


図3 CPU使用率のシミュレーション値と実測値の比較

システムへの実践適用を推し進めながら、シミュレータのいっそうの改良を図り、SI品質の向上とコスト低減の実現を目指します。

5. むすび

Apache Hadoopはまだ完成したものではなく、今後OSSコミュニティで、更に発展していくでしょう。Hadoopと同様のフレームワークは、現在商用製品では見当たりません。オープンソースの特性を生かしながら、継続した検証と更なる適用事例の推進、そしてHadoopの良さを更に最大限に引き出すことが可能なシミュレータを提供していく予定です。

*Apache、Hadoop、HBaseは、The Apache Software Foundationの登録商標または商標です。

*Googleは、Google Inc.の登録商標です。

参考文献

- 1) Tom White (著), 玉川竜司 ほか(訳): “Hadoop,” オライリー・ジャパン, 2010.1
- 2) 伊豆倉さやか ほか: “Evaluation of Non-functional Requirements based on Model-based System Integration Environment,” Proceedings of the 5th International Conference on Project Management, pp. 1271-1282, 2010.10

執筆者プロフィール

高橋 千恵子
ITソフトウェア事業本部
ITプラットフォームソリューション事業部
担当部長

世良 直彦
ITソフトウェア事業本部
ITプラットフォームソリューション事業部
エキスパート

突元 賢治
NECシステムテクノロジー
プラットフォーム事業本部
第三ソフトウェア事業部
主任

大崎 寛達
ITサービスビジネスユニット
先端技術開発グループ
主任

関連URL

Hadoopサポートサービス:
http://www.nec.co.jp/oss/middle_support/hadoop/

NEC 技報のご案内

NEC技報の論文をご覧いただきありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報(日本語)

NEC Technical Journal(英語)

Vol.65 No.2 ビッグデータ活用を支える 基盤技術・ソリューション特集

ビッグデータ活用を支える基盤技術・ソリューション特集によせて
ビッグデータを価値に変えるNECのITインフラ

◇ 特集論文

データ管理/処理基盤

- 超高速データ分析プラットフォーム「InfoFrame DWH Appliance」
- SDN技術で通信フローを制御する「UNIVERGE PFシリーズ」
- 大量データをリアルタイムに処理する「InfoFrame Table Access Method」
- 大量データを高速に処理する「InfoFrame DataBooster」
- ビッグデータの活用に最適なスケールアウト型新データベース「InfoFrame Relational Store」
- 高い信頼性と拡張性を実現したExpress5800/スケーラブルHAサーバ
- 大規模データ処理に対するOSS Hadoopの活用
- 大容量・高信頼グリッドストレージiStorage HSシリーズ(HYDRAstorr)

データ分析基盤

- ファイルサーバのデータ整理・活用を支援する「Information Assessment System」
- 超大規模バイオメトリック認証システムとその実現
- WebSAMの分析技術と応用例～インパリアント分析の特長と適用領域～

データ収集基盤

- スマートな社会を実現するM2Mとビッグデータ
- 微小な振動を検知する超高感度振動センサ技術開発とその応用

ビッグデータ処理を支える先進技術

- 多次元範囲検索を可能とするキーバリューストア「MD-HBase」
- 高倍率・高精細を実現する事例ベースの学習型超解像方式
- ビッグデータ活用のためのテキスト分析技術
- ビッグデータ時代の最先端データマイニング
- ジオタグ付きデータをクラウドでスケーラブルに処理するジオフェンシングシステム
- 柔軟性と高性能を備えたビッグデータ・ストリーム分析プラットフォーム「Blockmon」とその使用事例

◇ 普通論文

- 地デジTVを活用した「まちづくりコミュニティ形成支援システム」

◇ NEC Information

NEWS

- スケールアウト型新データベース「InfoFrame Relational Store」が2つの賞を受賞



Vol.65 No.2 (2012年9月)

特集TOP