

高速で高機能な MapReduce システム SSS

■ 研究担当：中田秀基 / 小川宏高

hide-nakada@aist.go.jp

■ 情報技術研究部門 インフラウェア研究グループ

■ 連携担当：伊藤智 satoshi.ito@aiist.go.jp

つくば中央

研究のポイント

- 大量データ処理の基本技術である MapReduce を高速化
- 分散キーバリューストアを基盤とし Map と Reduce を任意に組み合わせた処理を可能に
- 処理フローの最適化により従来の処理系に対してほぼ 10 倍の高速化を実現

研究のねらい

SNS や買い物サイトなどでは、ユーザの行動が生成する大量のログデータを解析し、よりよいサービスの提供に役立っています。このような解析には大容量のデータの処理に適した MapReduce が広く用いられています。しかし、従来の MapReduce 処理系は、繰り返し処理が遅く、また Map と Reduce を自由に組み合わせることができないなどの問題があり、プログラミングが困難でした。本研究ではこれらの問題を解決した MapReduce システム SSS を開発しています。

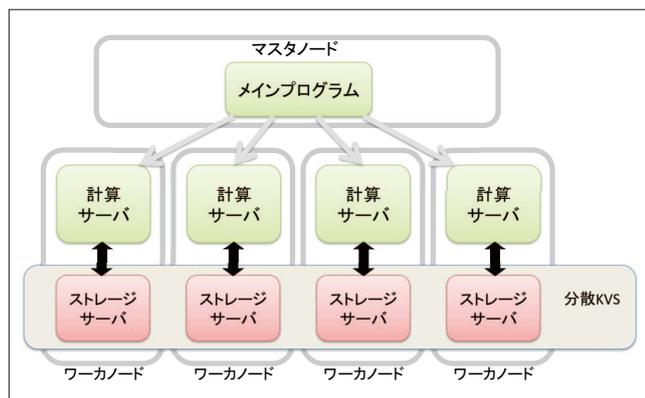
研究内容

SSS システムのワーカノードはそれぞれ、計算機能とストレージ機能の双方を持っており、ストレージは分散キーバリューストアを構成します。入力データも、Map の結果と Reduce の結果も、すべてこの分散キーバリューストアに書き出されます。データは書き込み時にノード間に分散され、各ノードは自ノード上のデータの処理のみを受け持ちます。このようにデータを配置することによって、Map と Reduce が独立したキーバリューストアへの読み書きとなるため、これらを自由に組み合わせた繰り返しを含む複雑なワークフローの実行が可能となります。

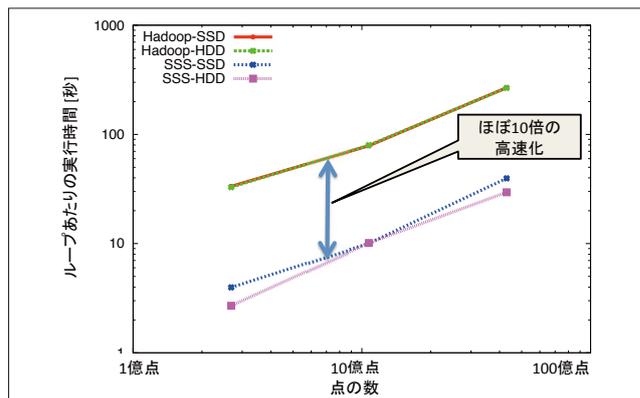
連携可能な技術・知財

- 大規模ログデータの高速な解析
- クラウドデータセンターの低消費電力運用
- バイオデータベースなどの大容量データの解析

謝辞：本研究の一部は、独立行政法人新エネルギー・産業技術総合開発機構 (NEDO) の委託業務「グリーンネットワーク・システム技術研究開発プロジェクト (グリーン IT プロジェクト)」の成果を活用しています。



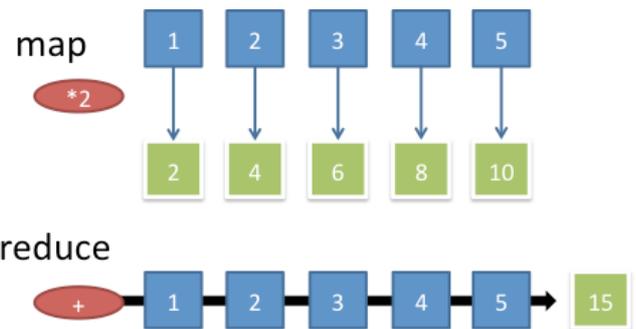
SSS の構成



K-means クラスタリングによる Hadoop との性能比較

MapReduceとは

- 高階関数Map, Reduceにちなんで命名された並列計算の枠組み
- Google が文献[1]で提案
- Map は複数のデータに対してまったく同じ処理を行う
- Reduceは複数のデータに対して集計処理を行う

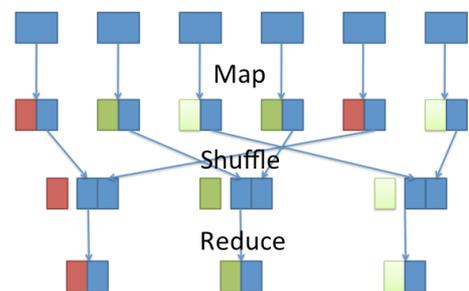


[1] Jeffrey Dean and Sanjay Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters", OSDI'04: Sixth Symposium on Operating System Design and Implementation, 2004

MapReduceの動作

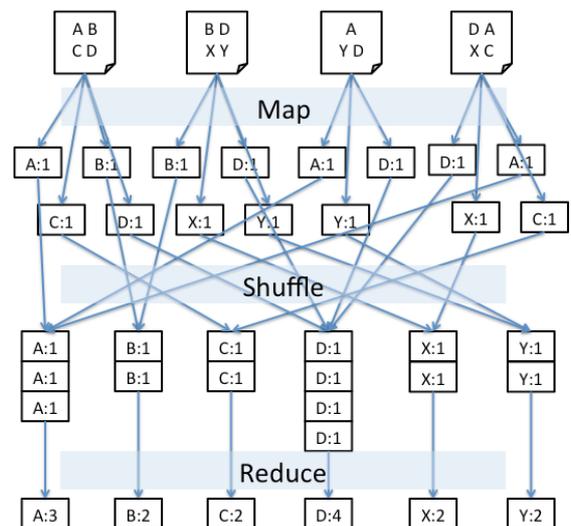
- Mapはキーとバリューを出力
- Shuffle でキー別を集める
- Reduceは同じキーの集合ごとに行う

→ Reduceも並列に動作する



MapReduceの例

- 複数の文書に登場する文字の数を数える
- Mapでは、文書を単語に分割、単語をキー、1をバリューとして出力
- Reduceでは、単語ごとに出力された1の数を積算



他の並列記述方式との比較

- マルチスレッド
一つのプログラムのなかで、複数の実行スレッドが並列に動作する。マルチコアに適する
- メッセージパッシング
独立したプログラムがお互いにメッセージを投げ合って通信する。大規模な並列動作が可能

	マルチスレッド	メッセージパッシング	MapReduce
プログラミングの容易さ	×	△	◎
高速性	◎	◎	×
高並列適応性	×	○	◎
効率	◎	◎	×
適応範囲	◎	◎	△