

特集

「ビッグデータ活用を支えるOSS」 特集号について

濱野 賢一朗^{†1} 吉野 松樹^{†2}^{†1} (株) NTTデータ ^{†2} (株) 日立製作所

ビジネス領域でチャレンジが進むビッグデータ活用は、従来のBIやBAでの取り組みの延長にとどまらず、新技術やアイディアに裏付けられたデータ活用の流れである。特に、この流れを支える技術要素にはさまざまなものがあるが、並列分散処理技術やスケールアウト型のアーキテクチャ等では、オープンソースソフトウェア (OSS) を中心に取り組みが進んでいるといっても過言ではない状況である。Hadoopに代表されるようにサービス企業が主体となって実用化されたOSSが広く利用されている。OSSの観点では、従来、商用製品の代替として進化・普及するものが多かったが、現在では商用製品に追随するものではなく、クラウドやビッグデータ活用を実現する新しい種のソフトウェアが増えている。

このような背景のもと、本特集号は、ビッグデータ活用のサービス現場で活用されているOSSの実例をその手法やノウハウを中心に紹介するものである。

「**社会の頭脳システム**」におけるドコモの**Hadoop** クラスターの活用事例」は、携帯電話サービスで用いる運用データから時々刻々と変化する人口推移の推計などを行う「社会の頭脳システム」で採用されたHadoopの構築・運用に関する知見が述べられている。コモディティなハードウェアとOSSを組み合わせて、可用性や運用性を向上・維持させた方式と実績が報告されている。

「**業務系バッチ処理へのHadoop導入についての課題と解決—Asakusa Frameworkの導入によるバッチの高速化—**」は、大手小売流通会社における管理会計、財務会計へのデータ生成を担うシステムでのHadoop導入について紹介している。著者らが開発するフレームワークでは、提供されるDSLでデータフローとデータモデルを記述するとコンパイラが適切なMapReduceプログラムを生成し、プログラムやデータの品質を確保するとともに、業務系バッチ処理で要求される諸条件を満たしたシステムを構成できるというものである。RDBMSとHadoopを密結合したアーキテクチャをとることで、可用性を向上させるとともに、業務システムで要求される処理の中断・

再開の仕組みを実現している点が興味深い。

「**OSSを活用したTwitterデータ提供システムの構築**」では、つぶやきデータ全量を受信、蓄積、配信するシステムをOSSを中心に構築した構成と工夫が紹介されている。大量の流量を取りこぼすことなく受信するために、スケールアウト構成にするとともに、フェイルソフトな設計を実現している。大量データを効率的にバッチ処理するために、データ格納方法と処理方式の両面から取り組んだ工夫が報告されている。

「**NetCOBOLのHadoop連携機能の開発と実践事例**」では、基幹系処理で動作実績のある既存アプリケーションをそのまま流用してHadoopで動作させるための手法を提案・実装している。既存バッチ処理の基本パターンに基づき並列処理できる仕組みを実現するとともに、Hadoopの標準ではサポートされないCOBOLのファイルフォーマットに対応する。既存資産を最大限に活かしつつ、Hadoopの並列分散処理の仕組みを活用できるようにした取り組みは応用が期待される内容である。

「**Jubatusの機能を利用した二者択一型不動産賃貸物件推薦サービスの開発**」は、不動産賃貸物件を題材として、利用者の嗜好を反映させながら絞り込みを行い最終的に物件を推薦するサービスのPoCに基づく報告である。バッチ処理によって事前にデータ変換や探索空間の構築を行うのとあわせて、機械学習エンジンJubatusを用いて利用者の操作に基づく逐次処理で、オンラインでの推薦を実現している。賃貸物件を探す利用者の多くが会員登録を行わず検索することから会員プロフィールや過去の行動履歴を反映させることが難しいこと、スマートフォンでは細かい希望条件を入力するのが難しいこと、利用者に負担の少ない選択回数にとどめたいなどの制約条件を満たすサービスが実現されている。Jubatusという新領域のOSSを活用した観点にとどまらず、新しいサービスの形を提案するものといえるだろう。本特集のインタビュー「**ビッグデータテクノロジーでサービスをドライブする～リクルートテクノロジーズのOSS活用～**」

でも、経緯や今後の思いも含めて話を伺った。

本特集の論文は、チャレンジが進むビッグデータの領域において、技術的な観点でも難しい課題に果敢に取り組んだものであり、今後の技術開発・サービス開発で活用・応用されることを期待している。

最後に、本特集の招待論文を執筆いただいた著者各位、原稿推敲に協力いただいた編集委員各位の皆様に深く感謝いたします。
