

ビッグデータ利活用を支える ITプラットフォーム

IT Platform for Big Data Repurposing

山本 康友

Yamamoto Yasutomo

社会や企業活動へのビッグデータ利活用の機運が高まる中、ITプラットフォームには、膨大かつ多様なデータの安全・確実な収集・蓄積、高速・簡便な抽出・分析、高い信頼性と拡張性、高度な運用管理自動化などが求められている。

日立グループは、ITプラットフォーム事業ビジョンOne Platform for All Dataを掲げ、今後必要となる技術を先進的に開発し、顧客のビジネスを支える製品・サービスを提供している。

1. はじめに

ビッグデータ時代が到来し、企業システムやインターネット上に存在する膨大なデータや、実世界の企業・個人・社会の活動から生まれるあらゆる情報をビジネスに利活用する動きが大きな潮流となっている。

日立グループは、このビッグデータ利活用と社会イノベーションへの貢献を支えるIT (Information Technology)

プラットフォームの事業ビジョンとして、「One Platform for All Data」を掲げている(図1参照)。

このビジョンでは、まず、日立グループのサーバ、ストレージ、ネットワーク、オープンミドルウェアを活用した統合プラットフォームに膨大かつ多様なデータが収集・蓄積される。検索・集計・分析を通じて、それらが社会インフラや企業情報システムで活用され、そこからまた新しいデータが生み出される。このデータの循環を、高速データ処理や仮想化をはじめとする先進機能を駆使した高信頼で統合管理可能な1つの基盤によって支援していく。

ここでは、ビッグデータ利活用と社会イノベーションへの貢献を支えるITプラットフォームソリューションについて述べる。

2. ビッグデータ利活用を支えるミドルウェア

ビッグデータ利活用においては、実世界で発生する膨大なデータを「収集」・「蓄積」・「配信」し、業務に「活用」するプロセスが重要となる。

「収集」では、膨大かつ多様なデータを収集し、リアルタイムに監視して即時処理すること、「蓄積」では、収集した膨大なデータを効率的に蓄積すること、「配信」では、必要なデータを素早く簡便に抽出・配信できることが求められる。最後の「活用」では、配信されたデータを高速かつ容易に集計・分析・活用できることが求められる。集計・分析結果に基づいた「知識」や「気づき」を企業や社会の現場にリアルタイムにフィードバックすることで、新たなイノベーションが実現可能となる。

日立グループは、データ利活用の課題とその解決技術を「Field to Future Technology」として体系化し、これらの技術を実装したオープンミドルウェア製品群として実現している(図2参照)。



図1 | ITプラットフォーム事業ビジョンOne Platform for All Data
日立グループのハードウェア製品を活用した統合プラットフォームにさまざまなデータが収集・蓄積され、検索・分析・活用される。

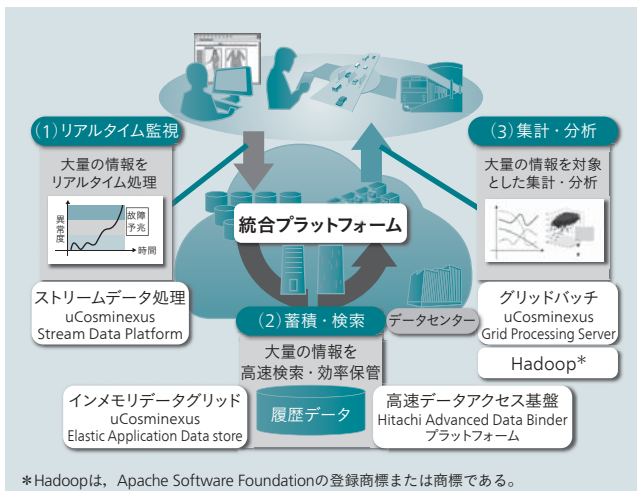


図2 | ビッグデータ活用を支えるオープンミドルウェア
ビッグデータ活用要件に対応したオープンミドルウェア群を示す。

「収集」要件に対して、発生する情報を即時処理することが可能なストリームデータ処理基盤 uCosminexus Stream Data Platform と、大量のデータを高速に処理可能なインメモリデータグリッドを構築できる uCosminexus Elastic Application Data store を提供している。また、「蓄積」・「配信」要件に対しては、高速データアクセス基盤 Hitachi Advanced Data Binder プラットフォーム^{*1)} を提供している。さらに「活用」要件に対しては、既存プログラム資産を活用したバッチ高速化を実現するグリッドバッチソリューションを提供している。

2.1 高速データアクセス基盤

日立グループは、ハードウェア性能を最大限に引き出す超高速データベースエンジンの研究開発を東京大学と共同で推進しており、これまでの研究開発成果を製品に採用した。

超高速データベースエンジンは、データ入出力要求の発生順序とは無関係な順序で非同期にデータを処理する非順序型実行原理^{*2)}に基づき、データベース検索処理を並列実行単位に自動分割して高多重で実行する。これにより、サーバやストレージの能力を最大限に使いきることができ、自社従来比100倍の高速検索性能を実現する。

この高速検索性により、例えば、データウェアハウスに蓄積された企業データの分析業務において、データマートを作ることなく業務データをアドホックに分析できる。そのため、バッチ運用のコスト削減と、分析軸の柔軟な追加によるデータ利活用が可能となる (図3 参照)。

高速データアクセス基盤の製品化第1弾として、超高速データベースエンジン、日立アドバンストサーバ HA8000 シリーズ、およびデータアクセス性能に優れた SSD (Solid State Drive) を搭載した日立のストレージ製品を組み合わ

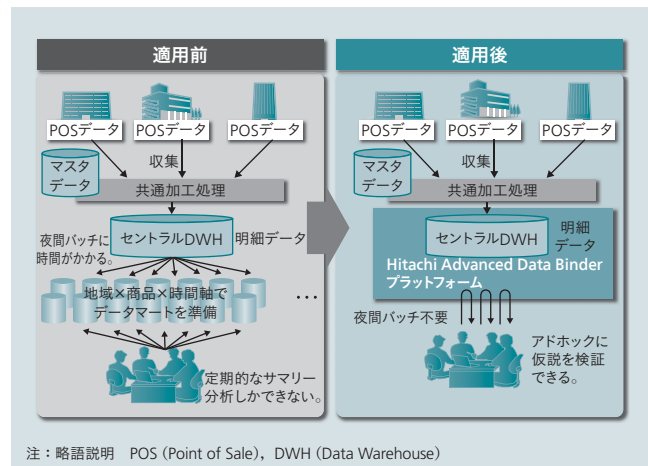


図3 | 高速データアクセス基盤の適用例

データマートを作ることなく、データをアドホックに分析できる。夜間バッチ運用のコストを削減でき、分散軸を自由に追加可能である。

せた Hitachi Advanced Data Binder プラットフォームを2012年6月に販売開始した。

さらに、第2弾として、株式会社DTS との共同でBI (Business Intelligence) 用アプライアンス製品「DaTa SuperExpress^{*3)}」を開発し、2013年6月に販売開始した。この製品は、DTS社の操作性に優れたBIツール「BI NavigationStudio^{*4)}」と Hitachi Advanced Data Binder プラットフォームを超高速度大容量のデータウェアハウス処理基盤として連携させたもので、迅速かつ容易にビッグデータの戦略的な分析・活用を実現する。

今後も、この高速データアクセス基盤を基に、データ利活用PoC (Proof of Concept) などを通じて、顧客およびパートナー企業と共同でのベストプラクティスマodelの拡充を図るとともに、流通分析ソリューションをはじめ、業務課題を解決するための各種ソリューション・サービスを展開する予定である。

2.2 新データ統合基盤

ビッグデータ時代では、事業の変化に対応した、膨大かつ多様なデータの収集・蓄積・配信によるビッグデータ利活用が事業発展のキーとなる。社会や企業の現場における「今」を知り、迅速に業務へフィードバックするためには、活用するデータの鮮度が重要である。しかし、分析・抽出といった、ビッグデータから新たな価値を創出する活動の生産性は低く、分析の観点を変更するたびに分析対象の再モデリングやデータ変換処理の再設計、データマートの作

*1) 内閣府が創設した最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会サービスの実証・評価」(中心研究者: 喜連川東京大学教授/国立情報学研究所所長)の成果を利用している。

*2) 喜連川東京大学教授/国立情報学研究所所長、合田東京大学特任准教授が考案した原理。

*3) DaTa SuperExpressは、株式会社DTSの登録商標である。

*4) BI NavigationStudioは、株式会社DTSの登録商標である。

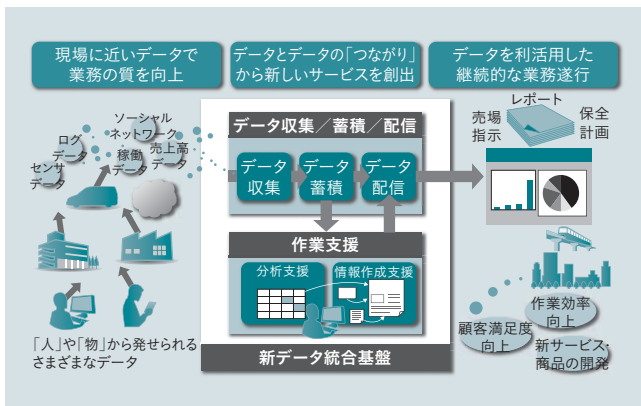


図4 | 新データ統合基盤
あらゆるデータの収集から配信までを支援する基盤を提供する。

り直しが発生し、時間と工数がかかる。また、企業内外に散在する多様なデータの収集には、個々のデータ処理に対応した新規のソフトウェア／ハードウェアの導入やアプリケーション開発などシステムの再構築が必要となる。その結果、データの鮮度を上げることが難しいという課題があった。

日立グループは、Field to Future Technologyの「収集・蓄積・配信を支援する技術」の個々に対応するミドルウェア製品に対して、これらの技術に対応するオペレーションを実行・管理するためのフレームワークである新データ統合基盤Hitachi Advanced Data Integratorを製品化した。今後、Global e-Service on TWX-21などのSaaS (Software as a Service) を支えるPaaS (Platform as a Service) として提供していく予定である。

Hitachi Advanced Data Integratorでは、構造型／非構造型／準構造型／非リアルタイム／ニアリアルタイムなどの多様な形式の収集データを、高速データアクセス基盤を活用して一元管理する。これにより、逐次発生するデータを高速かつ容易に扱え、現場で今起こっていることについて漏れのない計画策定と計画の確実な伝達を実現する(図4参照)。

3. ビッグデータ基盤を支える統合プラットフォーム

各企業や組織は、データベース、メール、画像、センサデータなど、多様なデータを保持している。そのデータ量は爆発的に増大し、現在のペタバイト (PB)、エクサバイト (EB) 規模から、今後はゼタバイト (ZB) 規模に達すると言われている。一方、各企業のIT投資は抑制傾向にあるため、こうした多様かつ膨大なデータを効率的に処理可能な、高性能・高機能・低TCO (Total Cost of Ownership) のITプラットフォームによって顧客の実業務を加速し、顧客のIT投資効果を最大化することが求められる。

3.1 高速処理を支えるハードウェアソリューション

取り扱うデータ量の増加に伴い、サーバ、ストレージ製品の処理性能向上、メモリ容量強化、省電力化など製品エンハンスが求められている。

サーバについては、大量データの高速度演算やサーバ仮想化による大量業務集約への対応のため、最新プロセッサやOS (Operating System) などを採用した新製品をタイムリーに提供していく。2012年11月に最新OS「Windows^{※5)} Server^{※5)} 2012」に対応し、処理性能やメモリ容量を強化したブレードサーバ「BladeSymphony」、2013年7月に最新のインテル^{※6)} Xeon^{※6)} プロセッサを搭載し、従来比約1.1倍の処理性能を実現したPC (Personal Computer) サーバHA8000シリーズの新製品をそれぞれ開発し、販売を開始している。

ストレージについては、大量データ高速処理の観点で注目されているフラッシュ媒体の製品展開を強化した。2012年8月には、ハイエンドアレイ Hitachi Virtual Storage Platformで Flash acceleration機能を提供し、ストレージコントローラをフラッシュによるデータ処理に適した方式へ強化することで、100万IOPS (Input/Output per Second) 以上のデータ読み込み速度を達成した。

また、2012年11月には、低価格・高性能を実現する独自の技術によって開発したフラッシュモジュールHitachi Accelerated Flash (以下、HAFと記す。)を全世界で販売開始した。HAFは、日立独自開発のフラッシュメモリコントローラの搭載などにより、従来比約2倍の高速データ処理と導入コストの約30%削減を実現する(図5参照)。さらに、2013年7月にはユニファイドストレージHitachi Unified Storage VM向けHAFを製品化し、オールフラッ

※5) Windows, Windows Serverは、米国Microsoft Corporationの米国およびその他の国における登録商標または商標である。
 ※6) インテル, Xeonは、米国およびその他の国におけるIntel Corporationまたはその子会社の商標または登録商標である。

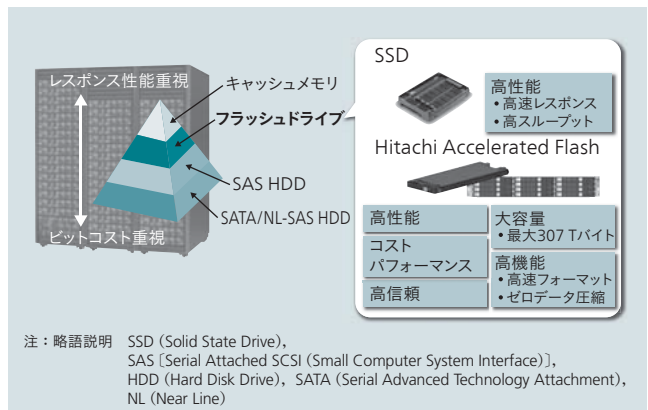


図5 | フラッシュ活用高速化ソリューション
日立独自技術に基づくフラッシュモジュールHitachi Accelerated Flashを活用し、高速データ処理が可能なプラットフォームを実現する。

シユ構成が可能な新モデルの提供を開始した。

3.2 コンテンツクラウドソリューション

データベースなどの構造型データ、メールや書類といった非構造型データなど、企業内で利用されるデータが多数の拠点や部門、社員が持つPCやモバイルデバイスなどに散在し、日々増加している。日立グループは、さまざまなファイルやコンテンツを一元管理し、簡単に参照できるようにするコンテンツクラウドソリューションを提供している。

拠点や部門に設置した仮想ファイルプラットフォーム Hitachi Virtual File Platform がクラウド環境の入り口 (Cloud on-Ramp) として機能し、拠点データのデータセンターに設置したコンテンツストレージ Hitachi Contents Platform への自動集約・一元管理を実現する。ファイル仮想化により、拠点側のユーザーはデータが拠点側・データセンター側のどちらにあるかを意識せずに利用することができる。これにより、拠点ごとの運用管理や、データ増大に伴う煩雑なストレージ装置・容量増設/データ管理といった課題を解決し、TCO削減やROI (Return on Investment) 向上に寄与する。

さらに、海外では、日立データシステムズ社 (Hitachi Data Systems Corporation) が2013年5月に、企業向けの安全なモバイル端末ファイル同期/共有ソリューション Hitachi Content Platform Anywhere の提供を開始した。登録されたモバイル端末間でどこからでも簡単にファイル共有でき、モバイル端末の安全性と信頼性を高め、従業員のよりよいITエクスペリエンスの提供を支援する。

3.3 運用管理を効率化する統合システム運用管理

情報システムの規模拡大、管理業務の複雑化、運用負担の増大、IT投資の抑制などが、企業のIT部門の喫緊の課題となっている。

日立グループは、統合システム運用管理JP1 Version10で、IT運用自動化のJP1/Automatic Operation, JP1で培った運用管理の技術・ノウハウを応用したJP1システム監視サービス、企業内でのスマートデバイス資産管理・セキュリティ管理をSaaS型で提供するJP1スマートデバイス管理サービスなど、ITシステムの運用管理コストを低減するソリューションを提供している (図6参照)。

ITシステムの導入コスト低減の観点では、事前検証済みのサーバ、ストレージ、ネットワーク機器、管理ミドルウェアをオールインワンで提供する統合プラットフォーム Hitachi Unified Compute Platform を提供している。事前

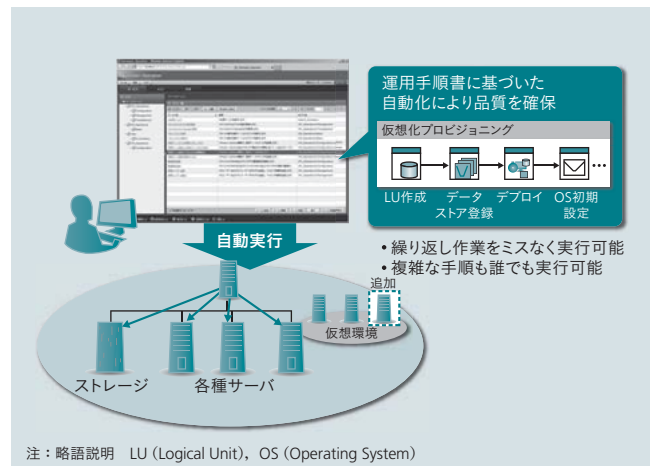


図6 | IT運用を自動化するJP1/Automatic Operation

運用手順書に従ってIT運用の操作を自動実行できるため、運用管理者の負荷を軽減する。

検証済み製品のため、製品導入後、ユーザーが製品を利用可能になるまでの時間 (サービスイン時間) を短縮することができる。また、同梱 (こん) した管理ソフトウェアによって仮想化環境における運用の容易化・自動化を支援し、運用コストを削減することができる。

さらに、中小規模システム向けのサーバ仮想化環境の迅速かつ容易な構築を実現するため、サーバ仮想化ソフトウェアを搭載したHitachi Unified Compute Platform かんたん仮想化モデルを2013年7月から販売している。

今後も、利用シーンに応じてソフトウェアとハードウェアを組み合わせた構成の拡充により、幅広い顧客のITシステムのコスト低減に貢献する。

4. おわりに

ここでは、ビッグデータ利活用と社会イノベーションへの貢献を支えるITプラットフォームソリューションについて述べた。

これらのソリューションは、現場のデータをITで処理することによって、現場の「過去」と「今」を知り、人間の創造力を上げ、「未来」を知ることを目指したものである。現場から収集した情報を正しく処理し、社会や企業活動プロセスにフィードバックすることによって、これまでにない価値を生み出すことができる。

執筆者紹介



山本 康友

1992年日立製作所入社、情報・通信システム社 経営戦略室 企画本部 所属

現在、ITプラットフォーム全体の事業戦略策定取りまとめに従事