

Big Dataにより新たな価値を創出する 次世代ITプラットフォーム

大量情報の集積・活用で新たな価値を 創出する次世代ITプラットフォーム

—One Platform for All Data—

Hitachi's Next-generation Platform Solution Creating Innovative Value by Utilizing Big Data

藤井 啓明

Fujii Hiroaki

変化する IT の潮流

クラウドコンピューティングの普及に伴って、クラウド上を飛び交い、収集・蓄積される大量のデータ、いわゆる「Big Data（ビッグデータ）」に注目が集まっている。また、従来は特定の目的で収集・蓄積・処理・生成されていたそれらのデータを、多面的に活用する「知識化」が、顧客のビジネスや社会インフラの新たな付加価値を創出する鍵となっている。

日立グループは、そうした新たな価値創出を支援するITプラットフォームの実現に向け、取り組みを加速している。

IT 活用の新たなトレンド

インターネットの登場以降、広くエンドユーザーはITを使いこなし、その発信力あるいはニーズが、ビジネスや社会のあり方をおかしてないほど大きく左右する時代になっている。例えば、スマートフォンなどの高機能な携帯端末の登場によって、ブログやTwitter^{※1)}といったソーシャルネットワークワーキングサービスの活用が、いっそうの広がりを見せている。そうしたサービスでは、日々ペタバイト規模のデータを処理するようになってきている。コンシューマ向けビジネスを生業としている企業を中心に、このようなサービスを介して得られるデータの活用をマーケティング活動に積極的に取り入れることで、新たなビジネスにつなげていく試みも多く始まっている。

また、昨今、RFID（Radio-frequency Identification）やセンサーなどを活用することで、さまざまな種類のデジタルデータを大量に取得することが可能である。例えば、流通などのモノの状態、工場の温度・湿度などの状況、プローブカーやETC（Electronic Toll Collection System）を活用した交通の状況、IC（Integrated Circuit）カード乗車券の履歴情報、電子マネーの購買情報、電力使用量などの自動検針、衛星による地球観測情報など、さまざまな形で現場の「見

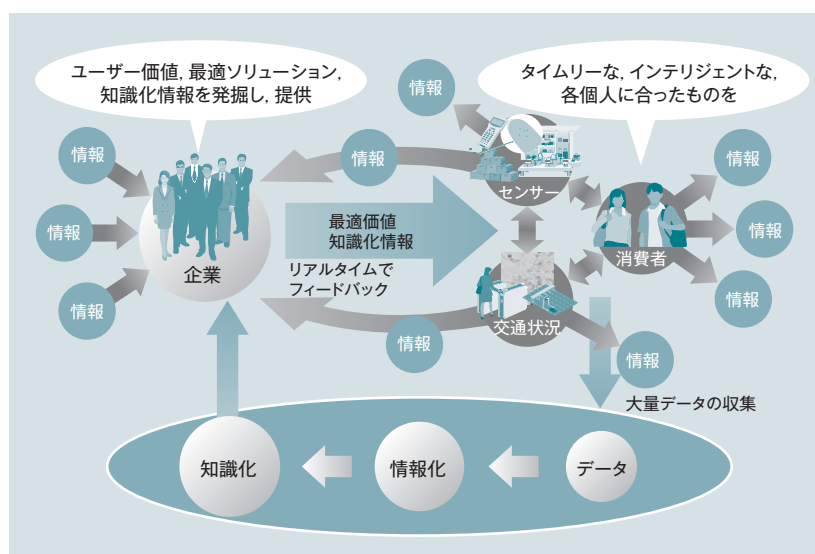


図1 | コンシューマ向けビジネスでの大量データ活用のフレームワーク

消費者の購買行動に伴って発生するデータを活用して、より適切な商品やサービスの提供が可能となる。

※1) Twitterは、米国およびその他の国におけるTwitter, Inc.の登録商標である。

える化」が可能となっている。

このような爆発的に増加するデータを活用することが、今後のITにとっての中核的な要件になってくると考えられる。

ビジネスの現場でのデータ活用のフレームワークを整理したものを図1に示す。

例えば、コンシューマ向けのビジネスにおいては、消費者の購買行動に伴ってさまざまなデータが発生し、また、発信されている。これらのデータを情報化し、さらに、知識化することで、商品やサービスを提供している企業は、ユーザーニーズの把握が容易となり、ニーズに応じた最適な価値提供が可能になる。

また、社会インフラにおけるデータ活用のフレームワークを整理したものを図2に示す。

電力インフラ、鉄道・道路などの交通インフラ、流通システムなどの社会インフラは、従来、あらかじめ決められたことを正確・確実に実施するということを第一義とする制御システムに支えられて、利用者向けのサービスを提供している。しかし、今後の社会インフラは、情報システムがかかわることで新たな価値を提供できるようになる。具体的には、社会インフラ自体の稼働状況や、利用者による社会インフラサービスの利用状況にかかわる各種の情報を、センサーやRFIDなどの新たなデバイス技術を活用して吸い上げ、今現場で起きていることを把握する。それを需要予測や需要と供給の調整といった情報に変換し、制御システムが活用できるようにすることで、社会インフラの制御をこれまで以上にきめ細かくすることができる。

これによって、利用者に提供される社会インフラサービスには、適度さや快適さといった付加価値が加味される。すなわち、社会インフラによって提供されるサービスに利用者の思いを還元することができるようになる。結果として、低炭素社会や、安全・安心・快適な社会といった目標の実現に寄与していくことができる。日本国内では今後、東日本大震災からの復興に向けて、さまざまな形で社会インフラの再構築

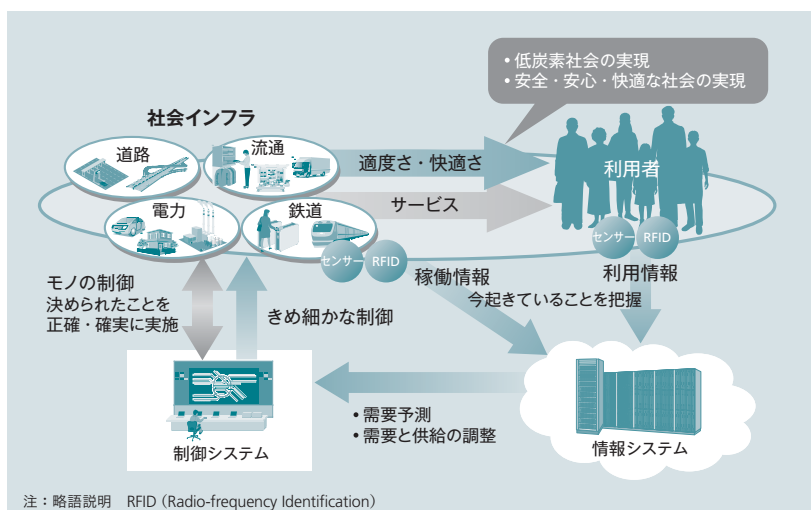


図2 | 社会インフラにおける大量データ活用のフレームワーク

社会インフラにかかわるさまざまな現場データを活用することで、より安全・安心・快適な社会の実現につなげることができる。

が進む。その際、データ活用による高付加価値化は必須の要件になると考えられる。日立グループもそうしたソリューションの提供によって、復興への貢献を果たしていく考えである。

このようなデータの情報化、知識化のプロセスを支えるのが、これからのITの新たな役割となる。従来は、一定範囲のデータの入力を受け、それを定型的に処理することがITの役割であり、これによって業務、特に定型業務の効率化を図っていた。今後は、図3に示すとおり、デジタル化された多量の現場データの活用により、単なる業務の効率化だけでなく、ビジネスの革新につながる「知」を生み出して

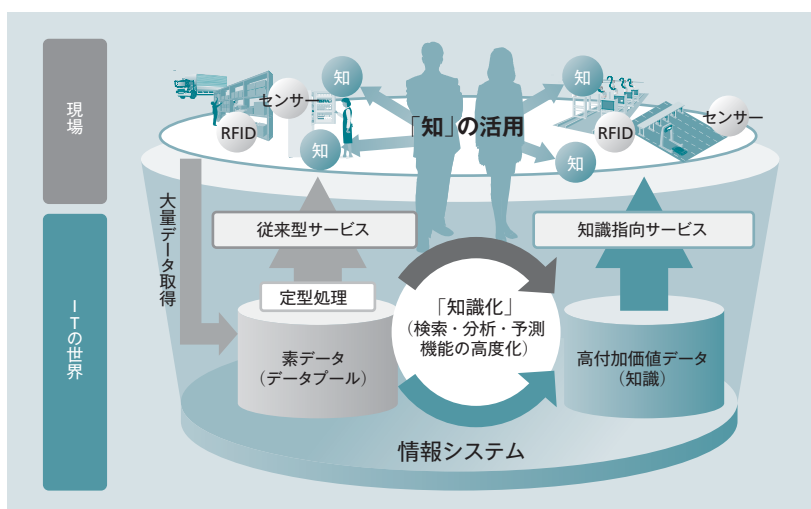


図3 | 大量データ活用の将来像

現場から素データを集め、ITで分析・活用して知識化を図り、その知識化された高付加価値データを使用し、新規サービス提供につなげる。

いくこともITの役割となる。集められた膨大な量の素データは、検索や分析、予測といった高度なIT機能によって、より付加価値の高いデータへと知識化され、これが新たな知として、さまざまな現場にフィードバックされ、そこでさらに質の高いサービスや新商品を生み出していく。以上がデータ活用を支えるITの大まかな仕組みである。

新たな価値創出に向けたITプラットフォームの進化

大量データ処理を支えるインフラの姿

日立グループは、大量のデータを活用し、知的生産性を向上させるという新たな役割を担うITプラットフォームの将来形を定義し、One Platform for All Dataと通称している(図4参照)。

多様で大量のデータを集積し、それをさまざまな形で処理し、活用するためには、データの一元的な管理・処理のための機能や、システムとしての柔軟性、スケーラビリティが重要となる。そうした要件を満たすため、日立グループは、以下の二つの方向性でプラットフォーム製品を強化・整備している。

(1) ITインフラ層

仮想化と自動化を追求し、サーバやストレージといった複数の装置から成るシステ

ムの構成や処理能力、データの格納容量などを意識せずに運用可能とする。

(2) コンテンツ管理層

ITインフラ層の上で、各データ、コンテンツの属性や種類にかかわらず、一元的、かつインテリジェントな管理・処理が可能とする。

One Platform for All Dataでは、このITインフラ層とコンテンツ管理層が有機的に統合され、大量データ活用による新たな価値創出を実現する。

ITインフラ層

前述のとおり、ITインフラ層を支える基盤要素技術は、仮想化と自動化である。

日立グループのストレージは、これまで一貫して先進的な仮想化技術を追求してきた。まず、異機種・複数のストレージデバイスの仮想化技術Hitachi Universal Volume Managerによって、ストレージの一元管理を実現し、運用を容易化している。さらに、ボリューム容量の仮想化機能Hitachi Dynamic Provisioningで、複雑なボリューム容量設計・管理を簡略化できる。また、Hitachi Virtual Storage Platformにおいては、さまざまなコストパフォーマンスの要件に最適に応えるべく、HDD(Hard Disk Drive)やSSD(Solid State Drive)など複数種類のデバイスで構成されるストレージ階層の仮想化機能Hitachi Dynamic Tieringにより、その運用管理を自動化している。

日立グループは、統合サービスプラットフォーム「BladeSymphony」を、メインフレームで培った仮想化技術を生かした独自の仮想化機構「Virtage」によって、高信頼なクラウド基盤に適したプラットフォームへと進化させてきた。さらに今後は、I/O(Input/Output)の仮想化や、サーバのリソースプールを実現し、より柔軟なプラットフォームをめざす。

一方、ITインフラ層の運用自動化の核となるのは、運用管理ミドルウェアである。運用管理ミドルウェア「JP1」のITリソース管理「JP1/IT Resource Management」

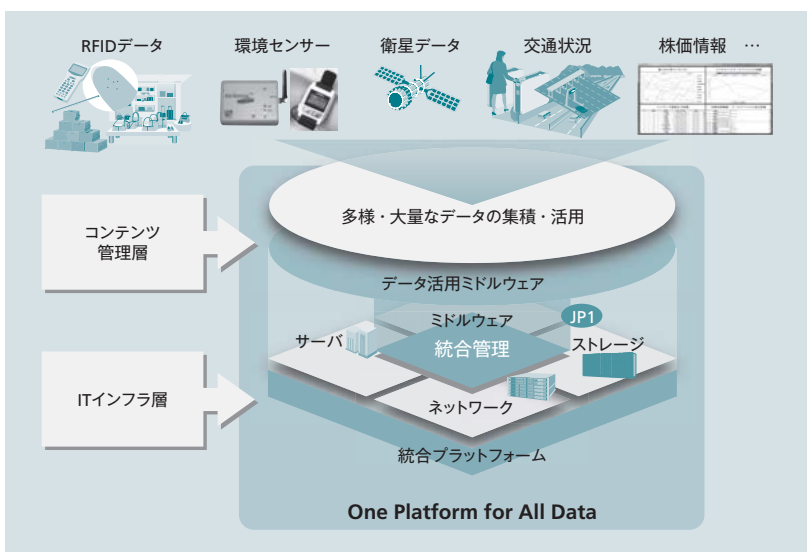


図4 | One Platform for All Data

多様・大量なデータの集積・活用を一元的・効率的に実現する、日立グループがめざすITプラットフォームの新たな形を示す。

は、仮想化・共有化されたITリソースを業務システム横断で最適活用するために、ITリソース割り当ての計画、実行、容量監視といった運用ライフサイクルを実現する。さらに、ITリソース管理は、混在する仮想化アーキテクチャを統一的に管理することも求められる。また、業務システム横断でITリソースが管理されるようになると、プラットフォーム機器それぞれの運用管理から、「JP1/IT Service Level Management」によるサービスレベル管理を絡めたような統合的な運用管理が、仮想化されたプラットフォームを支える基盤としてますます重要となる。

One Platform for All DataのITインフラ層は、ここまで述べてきた、サーバ、ストレージなどの個別のプラットフォーム製品とその運用機能を統合したものとなる(図5参照)。

従来、個々に最適なコンポーネントを調達し、組み合わせる形態が主流であった海外市場でも、最近は垂直統合型の統合プラットフォームに対するニーズが高まっている。この背景には、年々増大するITシステムの運用管理コストの低減が重要課題となっていることが挙げられる。ITシステムの集約・統合・仮想化の取り組みを推し進めていくと、集約・統合によって運用管理業務が軽減される一方で、サーバ、ストレージなど個々のコンポーネントの管理ツールをそれぞれ駆使する運用管理は、仮想化が進展することでその環境の複雑さゆえにコストが増大する結果となっている。

これに対して、日立グループがめざす統合プラットフォームは、こうした仮想化環境を前提としたシステム環境構築やサービスレベル確保などの運用管理機能を自動化し、TCO (Total Cost of Ownership: 総保有コスト) を削減する。例えば、省電力運用の実現、クラウド運用の効率化といったことも統合プラットフォームが提供する価値の一つである。

この統合プラットフォームを具現化した製品である Hitachi Unified Compute Platform は、すでに北米においてパイロット

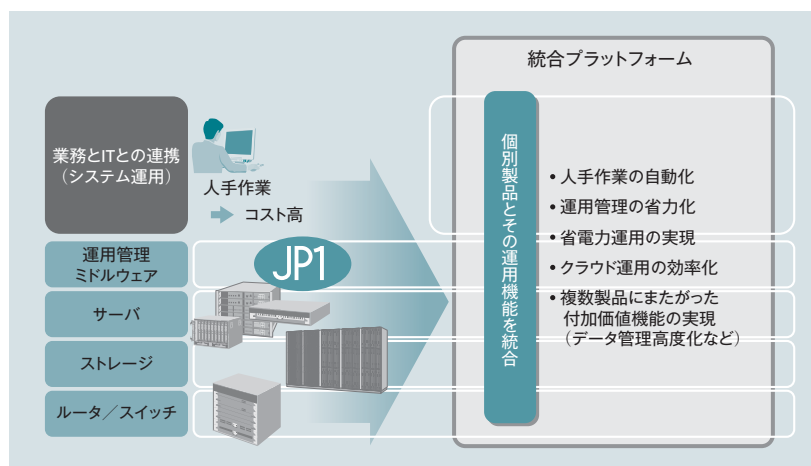


図5 | 統合／仮想化と自動運用によって高付加価値化したプラットフォーム

製品個別提供から垂直統合型プラットフォーム提供への高付加価値化により、顧客のTCO (Total Cost of Ownership) を削減する。

事業の展開を始めている。この製品は、企業内データセンターの**プライベートクラウド**^(a)を構成する基本構成要素としても位置づけることが可能である。詳細は、本特集掲載の「Integrated IT Platform for Highly Efficient Operation in Virtual Server Environment」を参照されたい。

コンテンツ管理層

現在、インターネット上で各種のクラウドサービスが展開され、ファイル、メール、画像、映像などのコンテンツと呼ばれる多様な**非構造化データ**^(b)が急激に増加している。現状、それぞれのコンテンツは個別のアプリケーションを介して管理されており、その相互活用には、煩雑な処理が必要である。こうした多様なコンテンツを一元的に管理し、活用できるようにするのが、コンテンツ管理層の主要な役割となる。

このようなデータの一元管理を効率よく実現するためには、従来のファイルサーバやNAS(Network Attached Storage)の統合・集約化、さらに、ライフサイクル管理のようなデータの階層化とその管理・運用の自動化が必要となる。2010年10月に販売を開始したHitachi Virtual File Platform(以下、VFPと記す。)は、そうした機能を提供する新たなソリューションである。VFPでは、大量のコンテンツデータを集約化するために、それに対応した高いスケーラビリティやアクセス性能を実現している。ま

(a) プライベートクラウド

企業などが内部に構築したクラウドコンピューティングのシステム。プライベートクラウドに対し、広くインターネット上で提供される一般的なクラウドの環境をパブリッククラウドと呼ぶ場合もある。

(b) 非構造化データ

テキスト、電子メール、画像、音声、動画など、データベースに納められないタイプのデータ。これに対し、経理データや販売データなどのように、汎用のデータベースシステムに格納して管理できるデータを構造化データという。

た、さまざまなアプリケーションからのアクセスに対応するため、複数のプロトコルでのデータアクセス機能をサポートしている。さらに、効率よく大量のデータを扱うために、ポリシー制御によるデータ管理機能など各種のデータ運用管理機能を提供している（図6参照）。

このVFPを大容量ファイルストレージ

のHitachi Content Platform（以下、HCPと記す。）と組み合わせた、クラウド上でのデータ管理ソリューションの例（Cloud on-Ramp：「クラウドへの入口」という意味）を図7に示す。拠点・部門ごとに配備されたVFPが、データセンターに配備されたHCPと連携し、ユーザーに意識させることなくデータを自動集約することで、バックアップ・アーカイブ、圧縮・重複排除などのデータセンターでの一元的なデータ管理が可能となり、拠点・部門ごとのストレージ容量監視・管理コストが低減できる。

将来的なコンテンツ管理層の実現に向けては、VFPが提供するソフトウェア機能も含めて、より高度なデータ活用を行うための機能を充実させる。これらによって、コンテンツ管理層は、コンテンツの仮想化、すなわち、物理的な場所に依存しないデータの集約・管理と、アプリケーションに依存しないデータ利用を実現する。さらに、コンテンツ管理層では、データの検索・探索、分析・活用（Re-purposing）を可能にするミドルウェアを整備し、アプリケーションとの連携を進めていく（図8参照）。

(c) Hadoop

大量データを効率的に分散処理するためのソフトウェアフレームワーク。オープンソースソフトウェアコミュニティApache Software Foundationによって開発・公開されている。

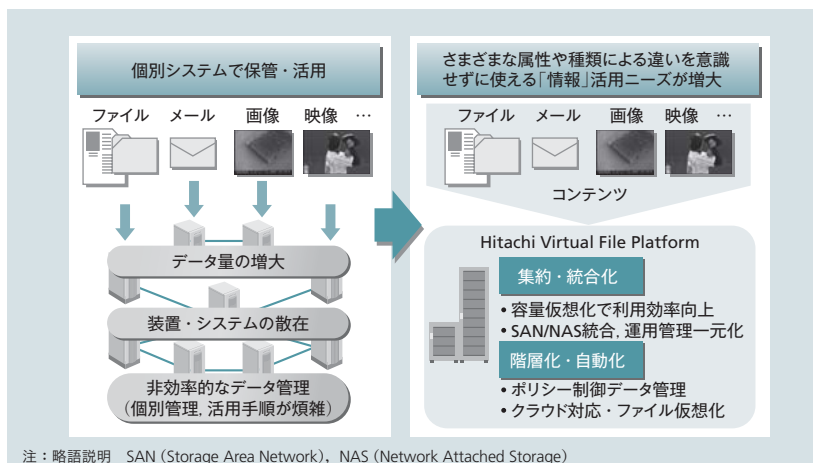


図6 | 大量データ集積を支援するソリューション

Hitachi Virtual File Platformは、大量のコンテンツデータを集約・統合して、データを一元的に管理・活用する。

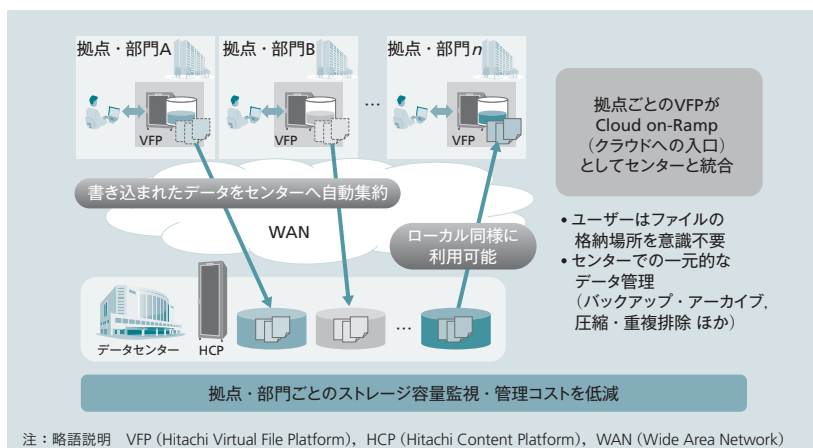


図7 | Hitachi Virtual File Platformを活用したソリューションの例

ファイル仮想化機能によって、クラウド上でのデータの一元的な管理を実現する。

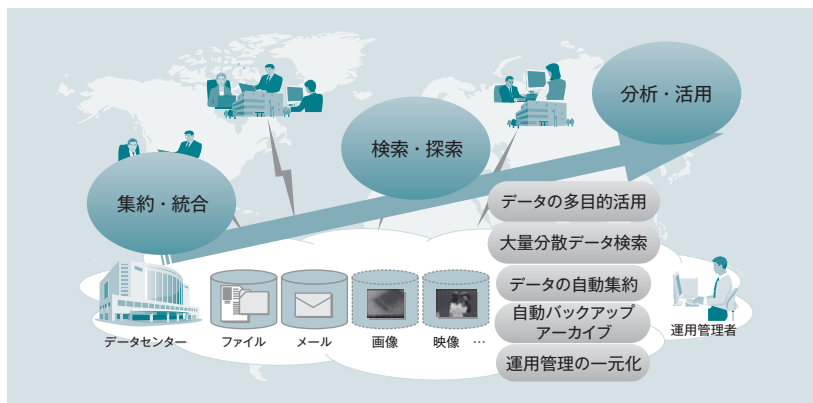


図8 | コンテンツ管理層が提供する機能の発展段階

コンテンツ管理層では、より高度なデータ活用を実現するための機能拡充を継続的に図っていく。

大量データ活用の実現に向けて

企業や社会が大量データを活用してその恩恵を得るには、データが効率よく処理され、現実的な時間で結果が提示される必要がある。そのためには、前述したプラットフォームのITインフラ層がスケーラブルに性能キャパシティを提供するだけでなく、その性能を効率よく生かすソフトウェア技術が不可欠である。

その要素技術の一つが、並列分散処理技術である。近年、大量データの並列分散処理の基盤としてオープンソースHadoop^(c)※2)に注目が集まっている。Hadoopには、消費者向けWebサイトのアクセスログから購買リコメンド情報を生成するような大量データ活用への適用実績がすでにある。日立グループは、このHadoopのサポート

※2) Hadoopは、Apache Software Foundationの商標である。

サービスを提供している。また、Hadoop適用が難しい分野に向けて、独自の並列分散処理基盤「uCosminexus Grid Processing Server」を提供している。

また、絶え間なく発生する大量のセンサデータなどを用いた機器・設備の異常検知などの目的に向けては、リアルタイム性が求められる。ストリームデータ処理基盤「uCosminexus Stream Data Platform」は、こうした課題に対応するもので、時々刻々発生する大量のデータを、その発生と同時にメモリ上でリアルタイムに処理できる。

さらに、大量データに対する商用データベースの性能限界に対処するための研究開発なども進めている。

コンテンツ管理層では、ここで述べたような機能層のソフトウェア基盤を土台とし、これにデータの検索・探索、分析・活用を実現する付加価値的なミドルウェアを組み合わせる形で整備を進めていく。

一方、こうした大量データ活用というソリューションの提供方法を考えると、大量データそのものの置き場所として、また、大量データ処理に必要なとなるコンピューティングパワーを供給するという側面からも、さまざまな形でクラウドサービスを活用していくことになると考えている。また、今般の東日本大震災を機に、クラウド活用への注目度はいっそうの高まりを見せている。こうした背景からも、大量データ活用を指向したクラウドサービスの整備にも注力していく。

Data Drives Our World

ここでは、今後の企業／社会イノベーションの核となる「大量データ活用」を指向した日立グループの次世代プラットフォームOne Platform for All Dataがめざ

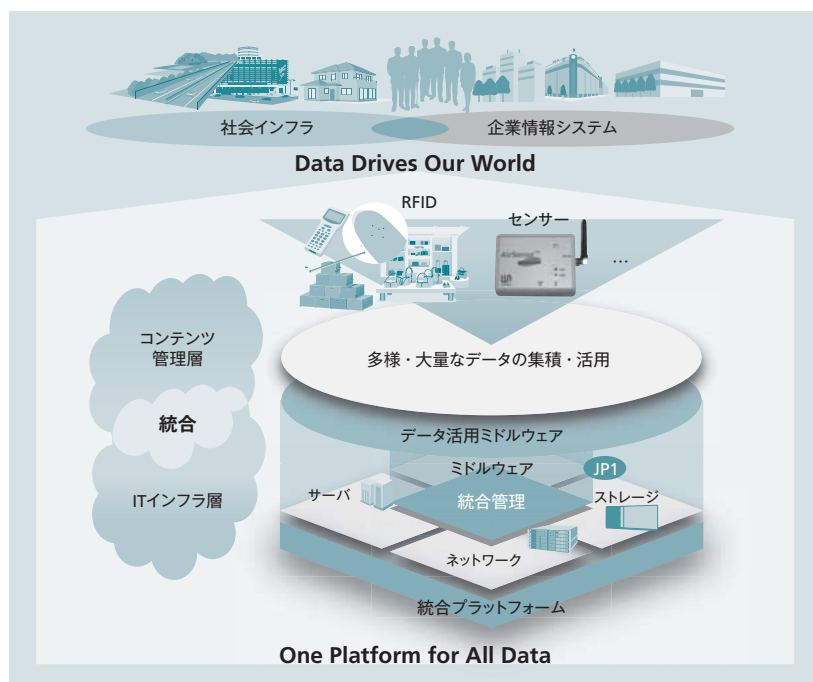


図9 | 日立グループがめざすプラットフォームソリューションの方向性
大量データ活用を指向した次世代プラットフォームOne Platform for All Dataをベースに、ビジネスや社会における新たな価値創出、イノベーションの実現をめざしていく。

す姿について述べてきた。それをベースとした今後の日立プラットフォームソリューションの方向性を図9に示す。Data Drives Our Worldという言葉のとおり、今後はデータが世界を動かす力を持つ「知的創造社会」を迎える。このようなデータの活用によって、ビジネスや社会において新たな価値創出、イノベーションの実現が可能となる。

日立グループは、このような価値創出を支援するプラットフォームソリューションの提供を通して、顧客のベストソリューションパートナーであり続けたいと考えている。その実現に向けて、これまでに提供してきた、サーバ、ストレージを核としたプラットフォームソリューションで培ってきたノウハウ、技術を結集し、発展させることで、One Platform for All Dataを具現化するITプラットフォームの整備・開発を進めていく考えである。

執筆者紹介



藤井 啓明
1991年日立製作所入社、情報・通信システム社 経営戦略室 事業戦略本部 プラットフォーム統括部 所属
現在、情報・通信システム社のプラットフォーム事業戦略検討に従事
ACM会員、IEEE会員、情報処理学会会員