

リアルタイム・アナリティクスを実現する デジタルデータ活用基盤の構築に向けて

— ． ．
— まずは始めるべきことは何か? —

PART 1

「あるべき姿」実現へは3段階のステップがあった

日本アイ・ビー・エム株式会社 グローバルビジネスサービス事業本部
ビジネス・コンサルティング 佐藤 俊 氏

PART 2

Can-BeからTo-Beの実現へ

リアルタイム・アナリティクスを実現するデータ基盤「SAP HANA」

SAPジャパン株式会社 ソリューション統括本部 新久保 浩二 氏

PART 3

選ぶべきITインフラの最適解とは?

日本アイ・ビー・エム株式会社 グローバルISV 鳥谷 健史 氏

多種多様なデータ活用が成果を生む

デジタル時代である。データの重要性は高まっている。しかし、その背景には何があるのだろうか。

IBMが実施しているCxOを対象にした調査結果では、「自社に最も影響を及ぼす外部要因」について、CEOの回答のトップが2012年から継続してテクノロジーになっている。佐藤氏は「UberやAWSなどに見られるように、異業種から新たなプレイヤーがITを武器に参入してくる時代です。テクノロジーをどう活かすのかが企業経営のポイントになっているのです」と説く。

一方、IT技術の進化自体がデータの重要性を後押ししている。「ネットに接続された“モノ”の数は100億を超え、2020年には300億、2050年には1,000億になると言われています」と佐藤氏。自動車、家電、スマートフォン、ウェアラブルデバイス。あらゆる機器がネットにつながり、そこから発信・生み出されるデータは爆発的に増え続け、2020年には40ゼタバイトにもなるとみられる。「しかも、データの大半は画像や動画を含めた非構造化データであり、これらをどう活用していくのかが課題になります」（佐藤氏）。

例えば、ホテルなどの身近なところにもその応用例があるだろう。ホテルマンの代わりにフロントにカメラが置かれ、そこから得られる画像をベースに接客する接客形態も考えられるのではないか。データ活用の重要性は、あらゆる業界で急速に高まっている。

理想的な姿を描きながらまずは現実の落とし所を追求する

こうした時代に対応するにはどんな戦略が必要となるのだろうか。佐藤氏が最初に指摘するのがデータ統合の重要性だ。

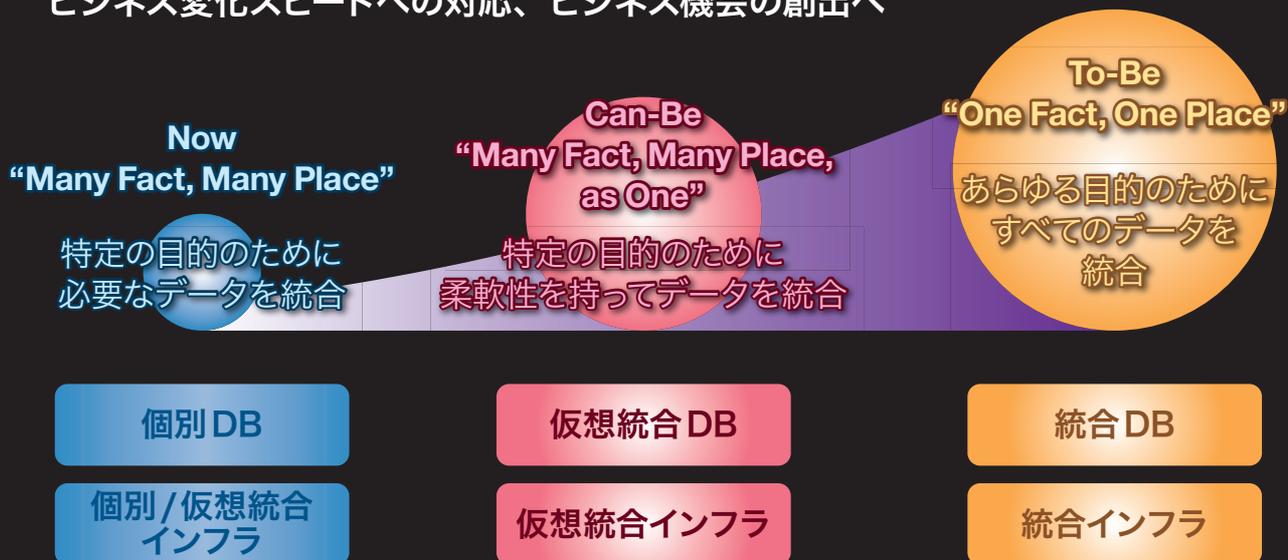
「今はいろいろなデータをさまざまなシステムが個別に保有しています。これでは限定的なデータ分析しかできません。バッチ処理でデータを統合、エクセルなどで加工するといった従来型の方法もありますが、予めデータの分析軸を決めておかなければならないため、効果も限定的になります」と佐藤氏は語る。

一方、データを統合して、活用シーンを高度化するテクノロジーは進化し続けている。佐藤氏は「夢だと思われていたことは、すでに実現されている」と指摘する。全てのあらゆるデータを一か所に統合し、統合的、横断的に見れるようにすることで、データ活用の可能性は飛躍的に拡大するというのだ。

「リアルタイムでの分析や自由な角度からの分析、あらゆるデータを組み合わせた分析など、やりたかったことができる時代がきています。例えば、個人の売上データと残業時間などの人事データを分析することで、本当の生産性がわかったり、顧客の顔認証と対話型で好みの食べ物を分析・把握できるといったように」（佐藤氏）。求められているのは、これまで別々に存在していたデータを全て統合し、必要な軸で分析する仕組みなのである。では実現のためにはどんなアプローチが考えられるのか。佐藤氏は3段階のステップからなる、To-Be/Can-Beアプローチを推奨する。個別のサーバ上にそれぞれのデータベースがある現状から、インフラとデータベースを仮想統合する次の段階に進み、さらに物理的にインフラとデータベースを統合するというベストの状態(To-Be)を目指す。この状態を目標としながら、まず既存環境に大きな変更を加えることなく、仮想的なデータ統合を実現すること(Can-Be)がまず最初に実施すべきことだ。

データ統合、そのためのプラットフォーム

To-Be/Can-Beアプローチ：目標とする姿、現実的なFeasibilityを踏まえた姿
ビジネス変化スピードへの対応、ビジネス機会の創出へ



©2016 IBM Business Consulting

「まずは既存環境に大きな変更を加えることなく仮想的に統合することで、データ分析のリアルタイム化と投資/TCOの投資対効果最大化を両立させることができます。そして最終的には物理的にインフラとデータベースを統合、あらゆる目的に活用できるデータ基盤を構築するのです。」(佐藤氏)

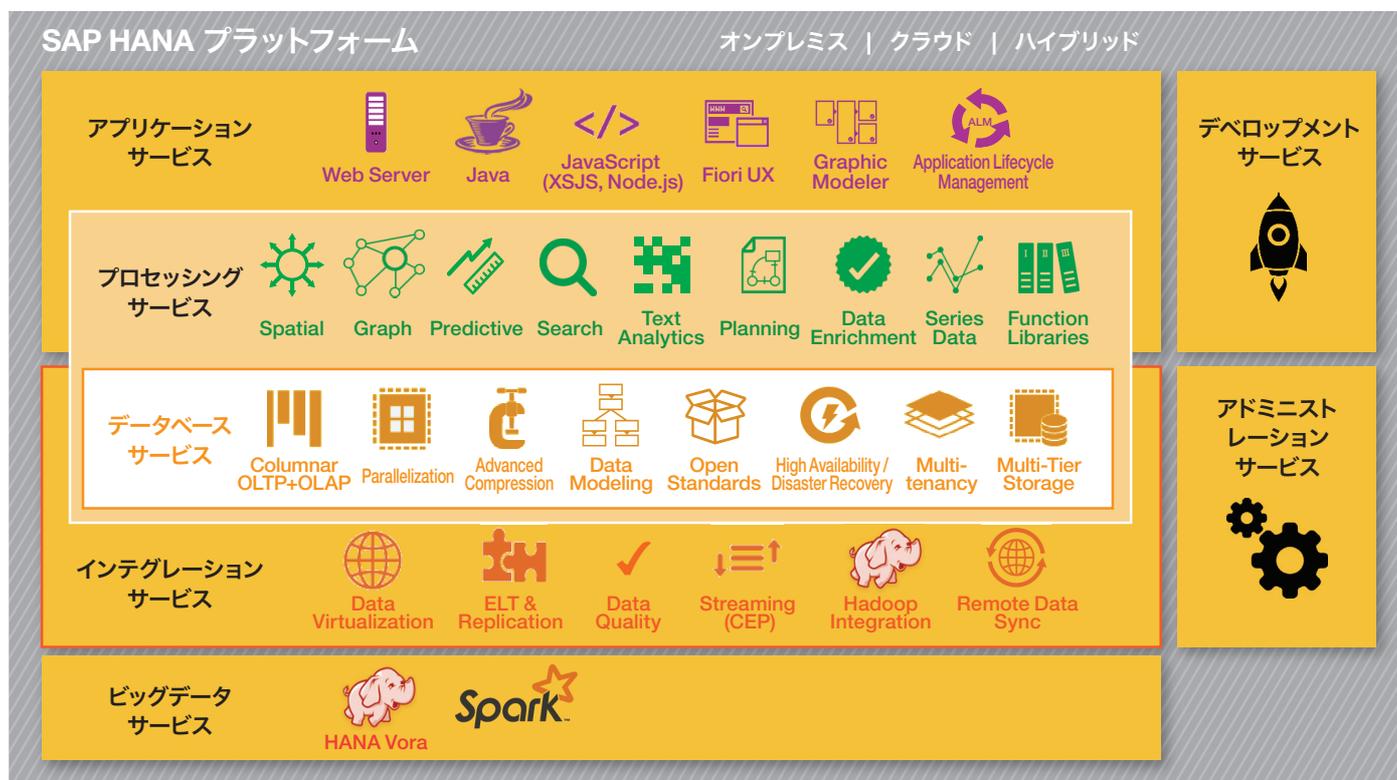
5つの要素から構成される SAP HANAプラットフォーム

このデータ統合を実現するソリューションとして注目されているのがSAP HANAプラットフォームだ。2010年、商業用インメモリ・データベースとして登場したSAP HANAは、IT業界に大きな衝撃を与えた。一般的なデータベースの10万倍以上という圧倒的なパフォーマンスが発揮できるとされたからだ。「しかし、イメージが先行して、誤解されているところも多い」と新久保氏は話す。SAPというERPベンダーが開発したために、ERP専用のデータベースだと思われることもあるという。

SAP HANAと言うと、OLTPとOLAPの両方に対応したインメモリ・データベースという印象が強い。しかし、実際には5種類のSAP HANAが存在する。一般用のデータベースであるPureHANA、SAP Business Suiteの選択肢としてのBusiness Suite on HANA、HANA上で最適化されたERPであるS/4HANA、データウェアハウス構築基盤であるBusiness Warehouse on HANA、そして次世代データウェアハウスであるBW/4HANAだ。

「HANAは登場以来進化し続けています。まずOLAPのためのデータウェアハウスとして提供され、その後OLTPにも対応、そしてHANAに最適化されたERPが開発され、さらにHANAの上で動くBusiness Warehouseが登場しました」(新久保氏)。

ここで現在のSAP HANAプラットフォーム・ソリューションの全体像を見ておきたい。



©2017 SAP SE or an SAP affiliate company. All rights reserved.

大きく5つの部分から構成されている。データベースとしての機能を提供するデータベースサービス、グラフモデルなど非リレーショナルなデータの検索といった用途ごとに提供される処理エンジンであるプロセッシングサービス、必要なアプリケーションを開発するアプリケーションサービス。周辺のシステムからリアルタイムにデータを仮想統合するインテグレーションサービス。そして最後に、Sparkなどに対応するために最近追加されたビッグデータサービスである。

独自の仮想データモデルでOLTPとOLAPを両立

このSAP HANAプラットフォームの中で根幹となる部分が、データベースサービスだ。「SAPが目指しているのはワンファクト・ワンプレイス。それを実現するためには、単一データで実行系のOLTPと分析系のOLAPの処理を実行できる必要があります。これはリアルタイム性の観点からも重要です。また、クラウド環境に対応し運用を簡素化する機能や、高度なデータ圧縮機能などもデータベースサービスとして提供されています」(新久保氏)。

SAP HANAが単一データで実行系と分析系処理を実行できる理由は、単にインメモリーで高速に処理しているからだけではない。独自の仮想データモデルの存在が大きい。新久保氏は「従来のデータベースでは、分析するためにデータ形式を転換してデータベースに入れ直す必要がありました。SAP HANAでは、最終的にどんなデータ形式にしたいかを定義するだけで分析処理を進めることができます」と解説する。

SAP HANAでは、実際にデータを移動させることなく、分析用のビューからデータを見て、OLAP処理専用のエンジン上で分析処理を実行している。バッチ処理やETLツール、集計テーブルなどを必要としないので、高速処理が可能だ。この分析用のビューはインフォメーションビューと呼ばれる。

「インフォメーションビューはHANAビューとも言われ、通常のOLAPのキューブに相当するもので、複数の分析軸の相関を知るためのデータモデルを提供しています」と新久保氏。売上や業績などの数値データを、仮想的に定義された顧客や店舗、商品などの切り口で集計し、ダイシング、スライシング、ドリルダウンなどの多次元分析を行うことができるという。

既存のデータベースを活用し、仮想的にデータを統合

SAP HANAが仮想的に分析を実行する上では、周辺システムのデータベースとの連携が必要になる。その部分を担うのがインテグレーションサービスである。そこではデータを統合するための様々な機能が提供されており、前述のCan-Beを実現する上で大きな役割を果たす。

各システムのデータに横断的にアクセスできる仮想的なアクセス機能、物理的にデータを移動したり複製する機能、リモートデータベースと同期をとる機能、そしてHadoopデータに対するアクセスに対応する機能だ。「長期間使わないデータをHadoopに保管するといったことも簡単に行えます」と新久保氏は語る。

特に、必要な時に必要なデータにアクセスできる機能は、仮想的にデータを統合するために重要であり、既存資産を有効活用する観点でも大きな価値がある。それを実現するのがSAP HANAのスマート・データアクセス(Data Virtualization)である。

「スマート・データアクセスによって、既存のデータベースにデータを置いたまま仮想テーブルを構築できます。リアルタイムにデータにアクセスでき、SAP HANA上で仮想テーブル間の結合なども可能で、OracleやSQL Server、DB2などの比較的新しいバージョンであればサポートされています」(新久保氏)。

ユーザは、まずSAP HANAに問い合わせを実行。するとSAP HANAが実データを格納する既存のデータベースにSQLを最適化したのち問い合わせを実行します。「この最適化がポイントです。SAP HANA及び既存のデータベースのどちらで結合などの操作を実行するのがベストなのか、SAP HANAが判断します。さらに仮想テーブルの更新もサポートされており、HANA及び既存のデータベース双方向でリアルタイムなデータ連携ができるようになっていきます」SAP HANA を活用することで、これまでとは異なる手法でスムーズなデータ活用を実現できるかもしれない。

ここまで紹介してきたデータ活用基盤を実際に構築するには、どのようなITインフラを選択すればよいのだろうか。多くのお客様が置かれている環境について鳥谷氏は次のように語る。「システムパフォーマンス悪化、システム障害、システム更改など既存のITにおける問題への対応に追われ、新たな取り組みを始めることができないIT部門が多いことが問題として挙げられます」つまり、既存システムの課題解決とデータ統合環境構築の2つの課題を同時に解決できるITインフラが求められているのだ。鳥谷氏はその解として、IBM Power Systemsを活用したプライベートクラウドの活用を紹介した。「全てを満たすことができるプラットフォームがPower Systems なのです」(鳥谷氏)。

IBM Power Systemsは他システムとどう違うのだろうか。ここでは鳥谷氏の挙げた要素から検証したい。

「高度なパフォーマンス」

「スマート・データアクセスによる仮想データベース統合環境では、既存データベースとSAP HANAの高速データ連携が、パフォーマンス面の重要なポイントになります。既存データベースとHANAを1台のPower Systemsに統合することで、Power Systemsの仮想ネットワークを活用した高速データ連携が可能となり、現行のデータベース環境と比較したパフォーマンスの向上も見込めます。」(鳥谷氏)

また、高速CPUと高帯域メモリーが、インメモリーデータベースSAP HANAのパフォーマンスを最適化でき、その効果はSAP HANA が公開しているベンチマークで他社製品と比べて2倍強の性能を示していることにも現れている。

「従量課金型モデル」でコスト最適化

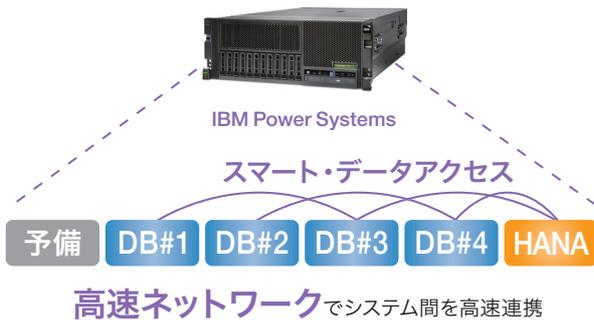
Power Systems には高度な仮想化技術が実装されており、柔軟なシステム基盤の提供によって、結果的にコスト削減効果が期待できる。高度な集約統合によって、管理するサーバーの数が減れば、運用コストの削減につながり、既存システムの運用問題が解決できる。さらにCPUを複数の区画で共有できるため、ソフトウェアライセンスも使用しているCPUやメモリ分だけの課金となり、ライセンスコストの適正化もできるという。SAP HANAは本番機のメモリサイズで課金されるため、まさにこのメリットが享受できるというのだ。「Power Systemsでは、最初は使用する一部の資源だけを活性化し、追加の必要があればCPUやメモリを必要な分だけ追加しながらシステムを拡張することが可能です。これはまさに従量課金型モデルであり、手ごろな価格でスモールスタートできることを意味します」。(鳥谷氏)

Power Systems は、「必要な時に必要なだけ使うというクラウドの一番のメリットを引き出せるプラットフォーム」といえるだろう。

高度なパフォーマンス

高速データ連携

スマート・データアクセス機能とPower Systemsの仮想ネットワークによる高速データ連携で、既存データベースとSAP HANAの高速連携を実現

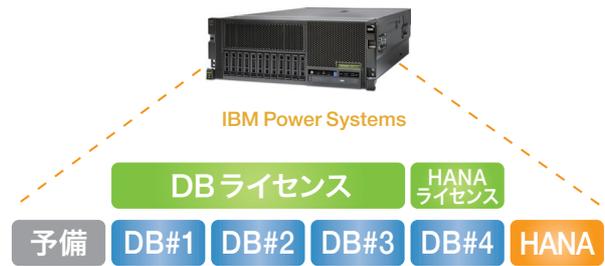


コスト最適化

ソフトウェアライセンス最適化

区画に割り当てられたリソースに応じたライセンス課金でライセンスコスト適正化を実現

- ・Oracleライセンス区画に割り当てられたCPUコア数
- ・HANAライセンス本番区画に割り当てられたメモリ量



©2016 IBM Corporation

「高可用性」

Power Systemsはメインフレーム譲りのテクノロジーを搭載し、障害の発生率を極小化するテクノロジーを実装している。「システムの停止時間としてはLinuxマシンでの最小時間を誇り、セキュリティの脆弱性報告はゼロです」(鳥谷氏)

Power Systemsによるプライベート・クラウドを活用することで、既存システム環境の課題を解決すると同時に、新たなビジネスニーズに迅速に対応できる戦略的データ活用基盤を構築できるのではないだろうか。

無償の簡易診断サービスで導入効果を検証できる

日本IBMでは、現在、IBM Power Systemsについての簡易診断サービスを無料で提供している。既存のシステム環境を何台のPower Systemsに統合できるのか、ライセンスコストやファシリティコストの試算値などを無償で検証できる。また、SAP HANA on Power Systems 検証センターでは、SAP HANA on Power Systems の開発とテスト基盤を提供しているという。気軽に試してみてもいいだろう。

【お問い合わせ】
SAP HANA、IBM Power Systemsの詳細情報や本誌掲載情報についてのご質問、具体的なお相談についてのご相談は下記までお問い合わせください。
■Web お問い合わせフォーム <http://ibm.biz/firstpower>
■お電話でのお問い合わせ先 IBM アクセスセンター 0120-550-210

IBM、IBMロゴ、ibm.com、IBM Power Systems、は、世界の多くの国で登録された International Business Machines Corporationの商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBMの商標リストについては、www.ibm.com/legal/copytrade.shtml をご覧ください。

本冊子の情報は 2016 年 10 月現在のものです。仕様は予告なく変更される場合があります。本事例に記載の肩書きや数値、固有名称等は初掲載当時のものであり、閲覧される時点では変更されている可能性があることをご了承ください。事例は特定のお客様での事例であり、すべてのお客様について同様の効果を実現することが可能なわけではありません。効果はお客様の環境その他の要因によって異なります。製品、サービスの詳細については、弊社もしくは IBM ビジネス・パートナーの営業担当員にご相談ください。