

リーダーシップカンパニー デル・テクノロジーズ、Cloudera、レッドハットの3社が共同でサポート

今、データ活用を進めようとする企業の間で「データマネジメント」に対する関心がにわかに高まっている。DX時代の「データマネジメント」は、それ以前と、言葉の指し示す範囲が変わってきているのだ。デル・テクノロジーズ、Cloudera、レッドハットの3社は、企業の「データマネジメント」にまつわる課題を解決するための共同ソリューションを提供しているという。今、「データマネジメント」が注目される背景と、その課題に対し、3社が共同でソリューションを提供する意義について聞いた。

社会のあらゆる領域にデジタル技術が浸透した現在、企業が市場での競争力を高め、成長を続けて行く上で「デジタルトランスフォーメーション (DX)」が必須であるという認識は、多くの企業が持っている。DXのポイントは、テクノロジーを積極的に活用することで、さまざまなシステムから生みだされる「データ」を、新たなビジネス価値の創出や、ビジネスプロセスの変革に生かして行くところにある。近年、「AI」技術に対する社会的な期待が高まっているが、その背景にある機械学習やディープラーニングといった手法を実施する上でも「データ」は必要不可欠だ。

DXの時代に、「データ」が持つビジネス上の価値や重要性は、かつてないほどに高まっている。こうした背景のもと、企業が関心を寄せているのが「データマネジメント」だ。

「データを継続的にビジネスへ生かす 営み」がDX時代のデータマネジメント

デル・テクノロジーズのDCW (Data Centric Workload) ソリューション本部でデータ分析主幹を務める堀田鋭二郎氏は「DX時代において“データマネジメント”という言葉が持つ意味は、以前のものから大きく変化している」と話す。



デル・テクノロジーズ株式会社
DCWソリューション本部 データ分析 主幹
堀田 鋭二郎 氏

かつて、IT運用における「データマネジメント」と言えば、データを蓄積するストレージやデータベース、あるいは、データを障害などから守るためのバックアップ/リストアのシステムをどう構築し、運用していくかといった観点を中心だった。しかし、DX時代の「データマネジメント」は、「データをビジネスに生かすことができる状態を継続的に維持し、さらに進化させていくための組織的な営み」という、データ活用のアーキテクチャを含んだ、より広い範囲を指す言葉になっていると堀田氏は言う。

この、より広い意味での「データマネジメント」には、DXを推進する世界各国の企業が、直近に取り組むべき課題として関心を寄せている。デル・テクノロジーズでは、世界各国の大規模、中規模企業のビジネスリーダーを対象とした調査 ([Digital Transformation Index](#)) を定期的に行い、その調査レポートを公に公開している。「今後1~3年で重点的にIT投資を行いたい領域」として「データマネジメント」は、欧米およびアジアのどの地域においても、最も関心の高い「サイバーセキュリティ」に次ぎ、トップ3以内にランクインしているという。

「世界的にデータマネジメントへのIT投資意欲が高い一方、日本単独での調査結果を見ると、実はトップ5にも入っていないという現実がある。日本では、AI/データ分析への期待が高まっているが、検討や概念実証、部分利用に留まり、ビジネスの中で本格的に活用する段階まで至っていない企業が未だ多いことが、その背景にあるのではないか。対照的に、既にAI/データ分析を本格的に活用し運用している欧米、中国などの企業では、AI/データ分析を取り入れたシステムを活用し、データを継続的にビジネスへ生かす運用管理手法、つまり“データマネジメント”への重要性を理解し、データマネジメントへのIT投資が、より高まっている点は必然と考えられる」(堀田氏)

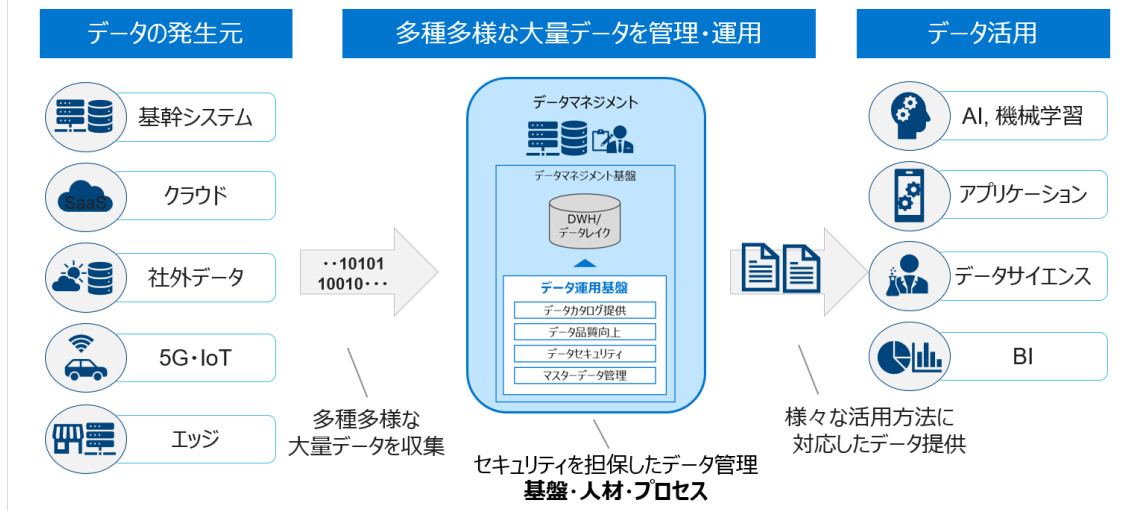
デル・テクノロジーズが考える データマネジメントに必要な「3つの“P”」

堀田氏は、国内の企業で多く見受けられる、DX推進やデータ活用に関する課題として「情報が分散していて、データ探求が難しい」「リアルタイムでのデータ把握が困難」「組織の中でデータの利活用に関して、制約が多い」「分析用データの作成に時間が掛かる」「外部にあるデータの取込や活用が難しい」などがあると話す。

デル・テクノロジーズでは、グローバルで展開しているベストプラクティスを、日本でデータマネジメントに取り組もうとしている、あるいは取り組みの中で前出のような課題を感じている企業に向けて提供している。同社がホワイトペーパー [Data Management for Research](#) を公に公開しているのもその一例である。

デル・テクノロジーズが考えるデータマネジメント

多種多様な大量データを収集し、継続的に、データ活用できる仕組みが必要



「われわれは、データマネジメントを実現するための要素として『3つの“P”』が必要だと考えている。“P”は、Platform（プラットフォーム、基盤）、People（人材）、Process（プロセス）を指す。これらが整備されてこそ、きちんとしたデータが、それを必要とするビジネスの現場へ、滞りなく提供できる。こうした環境を継続的に維持することで、データを、組織の生産性を高め、ビジネス上の価値を高めるために活用していくことができる」（堀田氏）

同社は、データ活用のための基盤を構築するためのハードウェアと、その基盤を組織のビジネスプロセスに組み込んで機能させていくためのコンサルテーション、DXトレーニングなどを提供する。主なハードウェアコンポーネントとしては、インテル® Xeon® スケーラブル・プロセッサを搭載した高性能サーバ「PowerEdge シリーズ」、マルチプロトコル対応スケールアウト型NAS「PowerScale シリーズ」をはじめ、多様なアプリケーションのパフォーマンスニーズやOS環境を、ベアメタル、仮想化環境を問わずシンプルに統合できるSDS（ソフトウェア・デファインド・ストレージ）製品「PowerFlex シリーズ」、ワークロードリソースとストレージとを統合し、柔軟な拡張性と高い管理性を提供するHCIソリューション「VxRail シリーズ」などが挙げられる。

データマネジメントを実現するデータ基盤の構築においては、これらのハードウェア上で動く、データ管理やアプリケーション展開のための、信頼性の高いミドルウェアが不可欠だ。デル・テクノロジーズでは、グローバルで多くの実績を持つ「Cloudera」「レッドハット」との共同によるデータマネジメントソリューションを、日本においてもより積極的に推進していく計画だ。

「Cloudera」で実現するクラウド混在環境でのデータマネジメント

「ビッグデータ」と呼ばれる、多種多様かつ大規模なデータを収集、管理、処理するための基盤として、10年以上前から企業での利用が進められているOSS（オープンソースソフトウェア）に「Apache Hadoop」がある。「Cloudera（クラウデラ）」は、企業がHadoop および Spark、Flink などのHadoopエコシステムを活用していく上で必要な周辺環境やサポートをパッケージと

して提供する、商用Hadoopベンダーだ。同社は、2008年の創業以来15年近くにわたって、多くの企業におけるデータ基盤の構築を支援してきた。

「Clouderaでは、今のようにデータマネジメントの重要性が叫ばれる以前から、企業がデータ活用で直面する課題に向き合い、その解決に有効なプラクティスをプロダクトへ反映してきた」

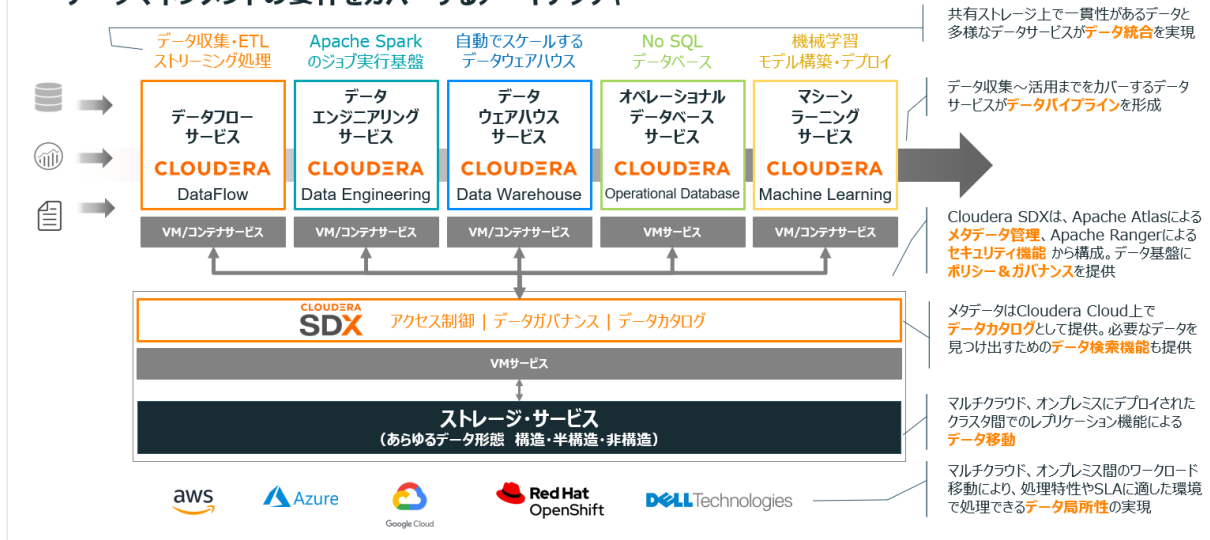
そう話すのは、Clouderaのソリューションエンジニアである杉山朋広氏だ。杉山氏によれば、この10年ほどの間に、多くの企業ではデータ活用に関連して「サイロ化したデータの単一基盤への集約に伴って考慮すべき事項の多さ」「グループ企業や部署間でのデータ基盤共有におけるリソース競合」「データへのアクセス制限の範囲設定や、その管理をいかに適切に行うか」「利用するデータが、もともと、どのシステムによって、いつ生成されたものかを正確に把握したい」「複数サービスのインテグレーションに掛かる手間を軽減したい」「データ基盤をスケールアウト、スケールインする際の運用負荷の増大を避けたい」といった多様な課題に直面し、解決に苦心してきたとする。



Clouderaでは、企業のこうした課題に対応できるHadoop、Spark、Flinkなど50以上のデータ処理ソフトウェアがフィーチャーされた商用パッケージソリューション「Cloudera Data Platform(CDP)」を提供している。このCloudera Data Platformのアーキテクチャが、データマネジメントにおける各考慮事項をカバーしている。

Cloudera Data Platform (CDP)

データマネジメントの要件をカバーするアーキテクチャ



「Clouderaでは、データの収集から、加工、分析、提供、プロダクトへの活用といった、ライフサイクル全体でのマネジメントを網羅できる機能をプラットフォームとして提供してきた。近年では、データの分析や活用において、オンプレミスのストレージだけでなく、パブリッククラウドを利用するハイブリッドクラウド環境も増えている。ハイブリッドクラウド、マルチクラウドを前提とした、データ統合基盤に必要なアーキテクチャと機能を備えている」(杉山氏)

Cloudera Data Platform の各機能を実現するコンポーネントは、VM (仮想マシン) やコンテナの形式で提供されており、オンプレミスのプライベートクラウド環境、あるいはパブリッククラウド上に展開することが可能だ。プライベートクラウド環境では本稿で紹介する「Red Hat OpenShift」にも対応している。「データマネジメント」の観点に基づき、これらの多様なインフラ上に展開された Cloudera によるデータ基盤を、統合的に管理するのが

「Cloudera SDX (Shared Data Experience) 」と呼ばれるツールだ。

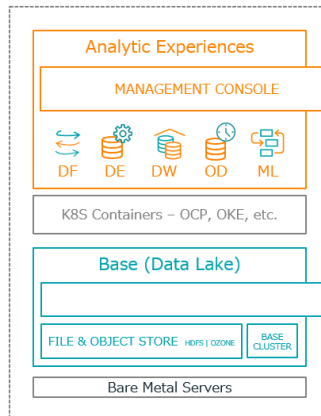
Cloudera SDX では、データのライフサイクル全体を通じ、データの格納先を問わず、一括したセキュリティ、ガバナンスの確保が可能となる。中心となるのは「アクセス制御」「データガバナンス」「データカタログ」の各機能だ。アクセス制御では、ポリシーベースの設定で、権限ごとにアクセス可能なデータベースを制御できる。カラム単位、ロー (行) 単位といった詳細なレベルでのアクセス制御、特定項目のマスキングなども可能となっている。

「データカタログ」は、企業が管理しているデータ資産全体の情報を取得し、一覧形式での把握や検索を可能にする機能。Cloudera SDX では、マルチクラウド環境においても、各環境からデータ資産の状況をリアルタイムに収集し、ガバナンスの確保や現場での活用の助けとなる情報を集約する。

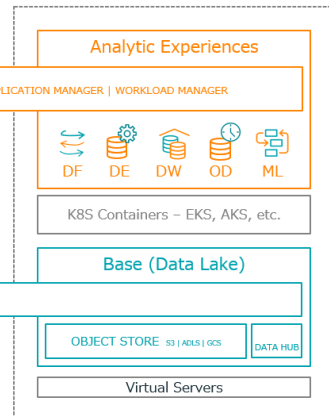
データプラットフォームにおける真のマルチ・ハイブリッドクラウド

- Public Cloud と Private Cloud が全く同じアーキテクチャです
- 管理コンソールで Public Cloud も Private Cloud も一緒に管理することができます
- データだけでなく、データセキュリティ・ガバナンス情報もレプリケーションできます
- もちろんアプリケーションはどちらの Cloud でも実行できます
- 全てにおいてシームレスに連携できるので、これがデータプラットフォームの将来像と言えます

CDP Private Cloud



CDP Public Cloud



SDX

「データマネジメントを実現していく上で求められる要件を、Cloudera SDX は網羅している。ハイブリッドクラウド、マルチクラウドを前提としたデータ基盤の構築、そのマネジメントに課題を感じている企業に、特にお勧めをしたい」(杉山氏)

企業における Kubernetes の活用を支援する「Red Hat OpenShift」

DX を指向したデータ基盤を構築するにあたって、現在、非常に注目度が上がっているのが、コンテナオーケストレーション技術である「Kubernetes」だ。

「コンテナ」とは、アプリケーションとその実行に必要なランタイム、一連のライブラリ、設定情報などをコンパクトにパッケージングする技術だ。また、コンテナ技術でシステムを構築する場合には、コンテナのデプロイ方法から、コンテナ同士の連携、コンテナ自体の死活監視やスケールリングなどさまざまな管理・運用を考える必要がある。こうした機能を自動化するソリューションが「コンテナオーケストレーション」と呼ばれ、この領域でのデファクトスタンダードが OSS のコンテナオーケストレーションツールである「Kubernetes」だ。

システムの開発や運用をコンテナベースで行うことで、オンプレミス、クラウド、マネージドサービス、そしてエッジといった多様な環境での展開や運用において、高い可搬性や運用における俊敏性といったメリットが得られることから、近年急速に活用領域が広がっている。その一方で、これらは現在も日々進化している技術であり、特に開発を迅速に行うためのツールの選定や、本番環境を安定して運用するための各種運用ツール、クラスタのバージョン管理、セキュリティというように、さまざまな要件を整える必要がある。Kubernetes のメリットを享受したいものの、そうした様々な機能要件を満たすコンポーネントや機能の選定であったり、OSS であることからサポートといった非機能要件をどう満たすことができるのか、を考える必要があり、本格的な利用までには大きな壁と感じる企業も多いのではないだろうか。

レッドハットは、こうした課題感を持っている企業を支援するため、Kubernetes をコアとした商用パッケージソリューション「Red Hat OpenShift」(以下、OpenShift) を提供している。OpenShift では、Kubernetes をはじめ、特にエンタープライズにおけるコンテナ活用に適した周辺機能群をパッケージ化し、サポートと合わせて提供する。

OpenShift によるハイブリッドクラウドの実現

一貫性 - Consistency

イノベーションへの フォーカス

差別化につながらない重労働(製品比較、脆弱性の対応など)は極力減らす必要があります。

ビジネス変化に追随するためには、ユーザはデータ利用や、データを利用するアプリケーションの開発にフォーカスできる基盤を選択することが重要です。

- Kubernetesを補完する付加価値機能
- RHEL CoreOS
- Container Catalog
- Universal Base Image

効率性 - Agility

Time to Valueの 加速化

開発・運用プロセスの最適化には、それぞれのプロセスにおけるセルフサービス化が求められます。

Kubernetesの自動化機能を拡張した Operatorを活用することで、セルフサービス化を促し、開発者主体でプロセスを進めることで生産性を向上できます。

- Operator (セルフサービス化)

柔軟性 - Flexibility

システムの柔軟性 (インフラ選択の自由、ISVエコシステム)

目的や変化に対応できるインフラ環境や機能をタイムリーに用意する必要があります。

オンプレやクラウドといった多様化するプラットフォームへのデプロイの対応、様々なISVパートナーと協創することで、アーキテクチャに豊富な選択肢を与えます。

- 様々なインフラ環境へのデプロイ
- Red Hat ミドルウェア
- ISVパートナーとのエコシステム



運用を自律化し、開発者のイノベーションを加速する基盤

レッドハット、テクニカルセールス本部 パートナーソリューションアーキテクト部のソリューションアーキテクトである鈴木亮氏は「Web アプリケーションといったワークロードでの利用が比較的多いイメージのある Kubernetes ですが、近年の調査では、データ活用に関連するワークロードの利用が多いことを示しているように、ステータフルなコンポーネントへの利用が進んできている」と話す。

「例えば、データベースやデータキャッシュ。そのほかにも、機械学習や AI の基盤、データ収集、クレンジング、アナリティクスといった、データ活用に関連するワークロード。データマネジメントの領域に Kubernetes を活用することは、既に企業にとって現実的な選択肢のひとつになっている」(鈴木氏)

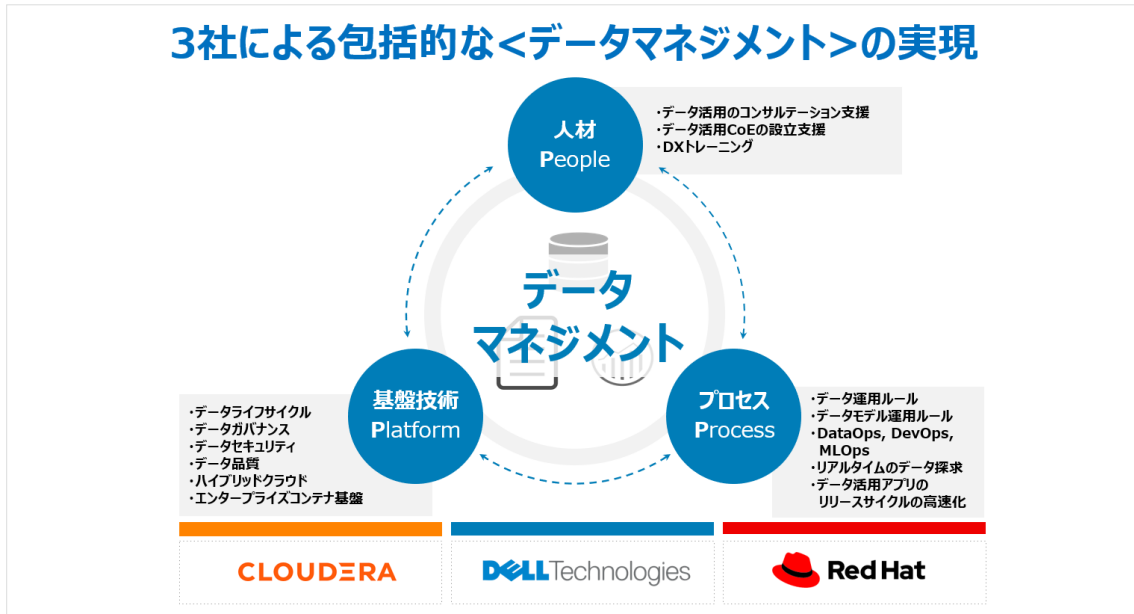
データマネジメントを実現するデータ基盤の構築はハイブリッドクラウドへ進化をしている中で、企業が「OpenShift」を採用

することから得られるメリットとして、鈴木氏は「一貫性」「効率性」「柔軟性」を挙げた。



レッドハット株式会社
テクニカルセールス本部 パートナーソリューションアーキテクト部
ソリューションアーキテクト 鈴木 亮氏

3社による包括的な<データマネジメント>の実現



「一貫性」は、ユーザー（開発・運用担当者、データ利用者）が、データ基盤のインフラ運用に煩わされることなく、「データ活用によるビジネスゴールの達成」という、本来の業務に集中できる環境を構築できることを指す。また、OpenShiftでは、脆弱性情報に基づいた迅速なセキュリティパッチ適用を行うことができるように、定期的なアップデートが行われる。またバージョンアップにおいてもワンクリックで実行できる仕組みを提供している。コンテナ基盤を常にセキュリティレベルが高い状態で運用していくためのユーザーにかかる負担は、大きく軽減される。

「効率性」は、データを利用するデータサイエンティストや蓄積されたデータを利用したアプリケーションを開発する開発者にとって、彼らを主体としてタイムリーにツールをデプロイしたり、必要なアプリケーションリソースを自ら調達できることが重要。それにより、社内外のサービス提供を迅速化し、さらにはフィードバックを素早く得て、サービスの改善につなげるサイクルを作ることができる。OpenShift(Kubernetes)では、インフラやアプリケーションの運用における知見を、実行可能なプログラムとして蓄積し、必要な作業を自律的に実行する「Operator」を活用している。こうした仕組みを使うことで、利用者主体でプロセスを回すことができるセルフサービス化の環境を実現している。Operatorの仕組みは、データマネジメントを実現する要素である『3つの“P”』の特にプロセスの改善に貢献することができる。

「柔軟性」は、ビジネスニーズや技術的な要件に伴って変化する、企業のシステムインフラへの要件に、容易に追従できる状態を指す。オンプレからクラウドへの移行および拡張、さらにクラウド間での移行や、同一クラウド内でのスケールアウトなど、データ基盤を支えるインフラには、常に変化が求められる。OpenShiftはそうした変化へ柔軟に対応できるコンテナ環境を提供できることが大きなメリットだ。加えてISVパートナーとのエコシステムも活用することができ、幅広いデータ関連のコンポーネントをオンデマンドに展開できることにより、さらに柔軟性の高いデータ基盤が実現できる。

「DXを視野に入れたITインフラを考える場合、企業がその時々求めるシステムを、スピーディに構築、運用し、必要に応じて変化させられる仕組みには、大きな価値がある。データ基盤とし

て、「インフラストラクチャとしてのKubernetes」と「アプリケーションプラットフォームとしてのKubernetes」の両面で価値を提供できることがOpenShiftの強みであり、OpenShiftを利用することで、企業は、いち早くそうした環境を導入できる。

デル・テクノロジーズ社、Cloudera社、そしてレッドハットの強みを組み合わせることで、DXを視野に入れた「データマネジメント」環境を実現することができると考えている。」（鈴木氏）

グローバルで蓄積したデータマネジメントの成功事例とノウハウを提供

日本企業における「データマネジメント」の取り組みを支援するにあたり、デル・テクノロジーズ、Cloudera、レッドハットの3社が共同でソリューションを推進する理由として、堀田氏はグローバルでの「実績」と「協力体制」を挙げた。

「デル・テクノロジーズのハードウェア群およびDX向け各種コンサルタントと、Clouderaおよびレッドハットのミドルウェア製品群の組み合わせは、グローバルでも事例が多く、相性が良いことが実証されている。特に、Clouderaについては、2021年の“Global Partner of the Year”をデル・テクノロジーズが受賞しており、日本国内でもトップリセラーとなっている。デル・テクノロジーズ自身も、Cloudera利用ユーザーであり、2.5ペタバイト以上のデータを、Cloudera上で管理しながら、データマネジメントのノウハウを自社内で蓄積している」（堀田氏）

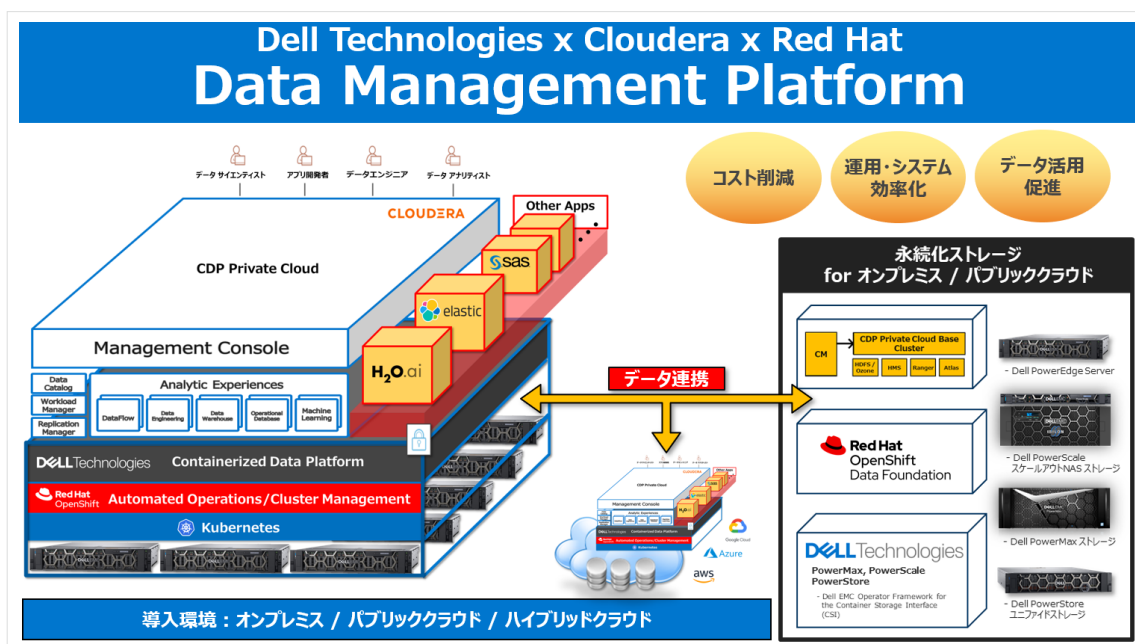


デル・テクノロジーのサーバ製品、ストレージ/ SDS 製品、HCI 製品と、Cloudera およびレッドハットのミドルウェア製品群は、グローバルアライアンスに基づき、事前の動作検証と適合認定が行われている。また、システム導入や運用の際に参照できる「Validated Design Guide」も用意されている。ハードウェアとミドルウェア、双方の提供ベンダーによる「お墨付き」があることで、ユーザー自身がシステム検討時や導入時に割く時間や手間は、大きく削減される。もちろん、運用開始後の信頼性の面でもメリットは大きい。

「この3社は、グローバルでビジネスを展開しており、各社の製品を組み合わせたデータマネジメント事例の知見も多い。DX時代のデータマネジメントを実現し、グローバルでの競争に勝ち

残れるようなデータ活用を目指す企業に対し、海外での成功パターンを積極的にご紹介していきたい」（堀田氏）

DXを視野に入れた「データマネジメント」には、ハードやソフトの導入に留まらず、組織の中でデータをフル活用できるようにするためのワークフローやデータガバナンス・ルールの策定、さらにはデータからビジネスバリューを生みだし、育てていくことができる企業文化の醸成など、さまざまな側面からの取り組みが求められる。その中で、インフラとなるデータ基盤を、いち早く使い始め、運用の手間を減らしながら、データ活用のレベルを高める取り組みに、より多くのリソースを投入したいと考えている企業にとって、デル・テクノロジー、Cloudera、レッドハットを組み合わせたソリューションの価値は大きいのではないだろうか。



Dell PowerEdge R750 サーバー



インテル® Xeon® スケーラブル・プロセッサ
Innovation Built-in