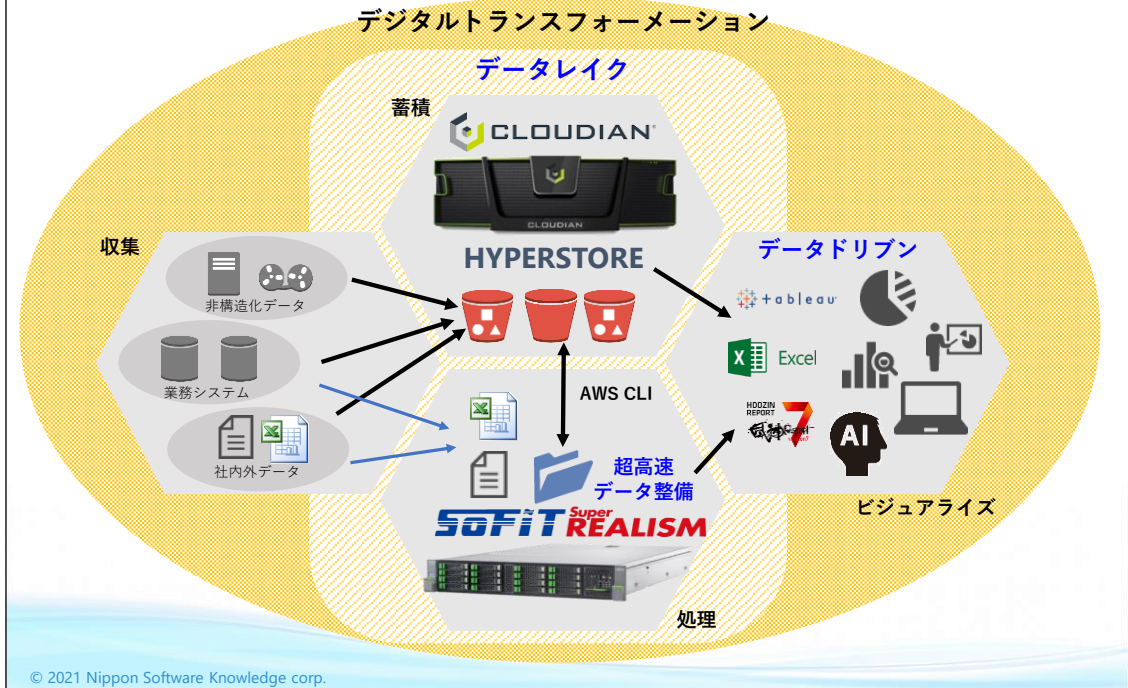


統合DWHからデータレイクへ



統合DWHからデータレイクに注目が集められた理由や、弊社がどのような形でデータレイクを利用しているのかをお話します。

社内にある構造化データを取り扱うのであれば、予めどのようなデータ分析が必要かを予測し、分析し易い形式に加工してから蓄積するデータウェアハウス（DWH）とBIツールの活用で十分でした。

ところが、IoTの発展に伴い、取り扱うデータはセンサーログ、画像、音声など様々な非構造化データが増え、データレイクという概念が注目されるようになりました。

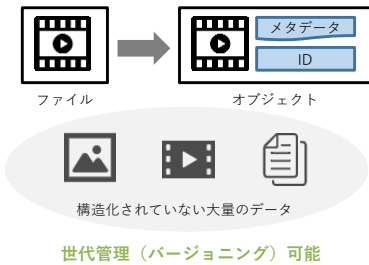
データレイクを実現するためには、様々な場所で発生する様々なデータを格納することに適したオブジェクトストレージが有効となります。

ただ、単にデータを蓄積するだけでは、使えないデータの沼地（データスワンプ）になってしまいます。データを整理・可視化し、いつでも統合・活用できる状態 = 澄み切った湖（データレイク）に整備しておくことで、データをビジネス資源として活用できるようになります。

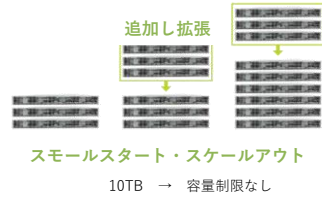
そこで、弊社ではクラウドファン株式会社のオブジェクトストレージ製品HYPERSTORE（ハイパーストア）とREALISMをAWS CLIで連携することでデジタルトランスフォーメーションを実現します。

HYPERSTOREとは

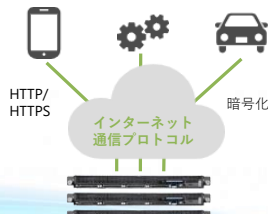
オブジェクトというフラットなデータ管理



最小10TBからのスモールスタート



インターネット経由のアクセス



複数拠点機能 (災害対策)



© 2021 Nippon Software Knowledge corp.

クラウドのオブジェクトストレージ製品HYPERSTOREについて説明します。

データレイクの実現には、ファイル文書、写真、動画、ログなど、大量のデータ保存が求められます。HYPERSTORE (オブジェクトストレージ) は、ファイルとメタデータをセットにしたオブジェクトというフラットな単位で格納するため、大量データを扱うことができます。また、同名のファイルを格納する時、ファイルストレージでは上書きされますが、HYPERSTOREでは別IDを付与し別オブジェクトとして格納することで、世代管理が可能です。

データ量の増加とともにサーバの台数を追加するだけで、ストレージシステム全体のデータ容量を拡張できます。(スケールアウト)
最小10TBから無制限に拡張可能です。

DXでは、世界中のデバイスやセンサーがデータを集めるIoTや、移動先、自宅、遠隔拠点からのデータアクセスも重要です。HYPERSTOREは、HTTP/HTTPSといったプロトコルで、インターネットを経由したデータの読み書き、ファイルアクセスが可能となります。

DXを実現するストレージには、冗長性や高可用性も求められます。HYPERSTOREは、ファイルを保存すると複数のサーバに分散保存します。複数データセンター機能も標準で備えていますので、ノード3台構成で3複製とした場合、1台は本社、1台は支店、そして残り1台はクラウド事業者 (データセンター) という設置が可能です。複数のデータセンターにHYPERSTOREを設置することで、いずれかのデータセンターに被害があった場合も無停止で運用を継続できます。さらに、Amazon S3のAPIに完全準拠しているため、多くのクライアントやアプリを改修せず利用できます。

以上のように、オブジェクトストレージは、データレイク/DXの実現に必要な不可欠なストレージです。