

業務ソリューションとクラウドデータウェアハウスによるデータ駆動ビジネスの実現

Building Data-driven Business with Cloud Data Warehouse and Our Business Solutions

DX を推進しデータ駆動ビジネスを実現するには、それを支えるデータ活用基盤が欠かせない。本稿では、従来のオンプレミスのシステムで構築されたデータ活用基盤の課題を述べ、その解決策としてクラウドデータウェアハウスの活用を提案する。また、その活用イメージを通し、業務ソリューションがクラウドデータウェアハウスを中心に BI や統計解析・機械学習の分析ツールと連携することで、データ駆動ビジネスが実現可能となることを示す。(株)日立ソリューションズ東日本が提供する AppSQUARE を例に、そのような連携が実際に可能であることを試作により検証した。今後は、業務ソリューションのラインナップ拡充などにより、さらに顧客価値を高めていく。

| | |
|--------|----------------|
| 飯塚 新司 | Iizuka Shinji |
| 佐々木 和也 | Sasaki Kazuya |
| 齋藤 彰大 | Saito Akihiro |
| 及川 慎也 | Oikawa Shinya |
| 後藤 秀樹 | Goto Hideki |
| 橋田 真人 | Hashida Masato |

1. はじめに

IT 技術の利用により企業の業務プロセスやビジネスを変革していくデジタルトランスフォーメーション(DX)では、データの活用と分析が重要となる。収集したデータに基づいた分析・予測の結果を踏まえてビジネスの意思決定や課題の解決を迅速に行っていくアプローチはデータ駆動と呼ばれ、DX を推進する上でのデータ活用のあり方として注目されている。このデータ駆動に基づくビジネスを実現するには、企業内外の様々なデータを収集し、利用しやすい形態に加工・蓄積するためのデータ活用基盤が必要となる。データ活用基盤には、多様なデータソースや分析ツール・サービスとの柔軟な連携、データのセキュリティの確保とユーザへの開放の両立、データ量の増加や機能追加に耐えうる拡張性と柔軟性などが要求される。従来のオンプレミスのデータウェアハウス(以下、DWH)ではこれらの要件は実現が困難なため、(株)日立ソリューションズ東日本(以下、HSE)では、クラウドサービスとして提供されているDWHであるクラウドDWHをデータ活用基盤として利用することを提案している。本稿では、HSEが提供する業務ソリューションや分析ソリューションとクラウドDWHを連携させることで、どのようにデータ駆動ビジネスを実現できるか、クラウドDWHの活用イメージを述べる。

2. データ駆動ビジネスに要求されるデータ活用基盤の課題

情報処理推進機構(IPA)が公開する「DX実践手引書 ITシステム構築編」¹⁾では、データ駆動ビジネスを実現するためにデータ活用基盤に求められるITシステムの要件として、以下の5つのポイントが挙げられている。①データ活用基盤が社内外の様々なソースからの柔軟なデータ収集・蓄積が可能である。②収集・蓄積されるデータの品質が担保され、かつ活用しやすい状態に整理されている。③堅守すべきデータセキュリティの確保と、積極的なデータ活用のユーザへの開放が両立できている。④データ量の増加に対応しうるキャパシティーや機能追加に耐えうる拡張性・柔軟性を持っている。⑤疎結合で拡張性が高く、周辺システムとの連携が可能である。

近年のDXの浸透により、企業競争力を確保・維持するために、上記の要件を満たすデータ活用基盤をより速く構築し、臨機応変に運用したいというニーズが増えている。従来はオンプレミスのDWHでデータ活用基盤の構築が行われてきたが、以下のような課題が生じている。

(課題1) 多様なデータソースへの対応や、AI・BIなど周辺の分析ツールやサービスとの柔軟な連携が求められるようになり、システムの改修負荷が高くなっている。

(課題2) データセキュリティとユーザへの開放を両立

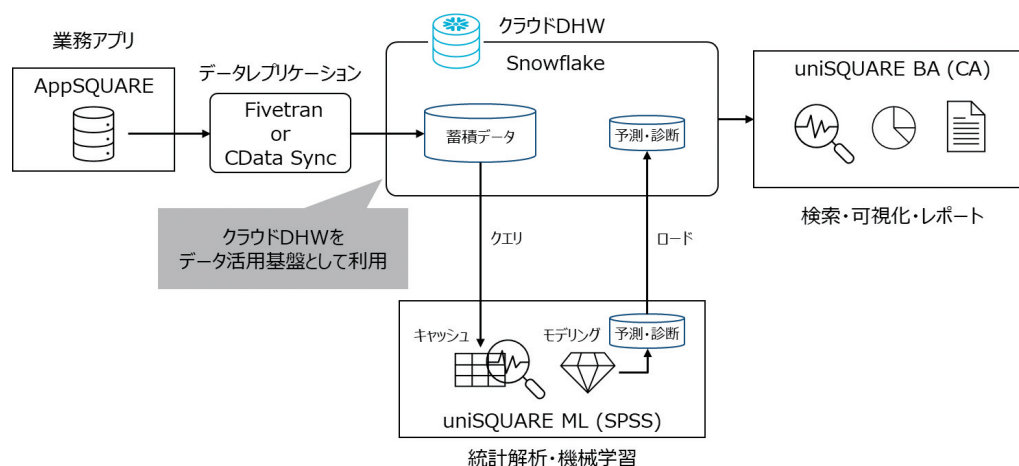


図 1 クラウド DWH 活用イメージ

させるため、サイバー攻撃などシステムへの攻撃の監視や脆弱性への対応が常に求められるようになり、運用負荷が高くなっている。

(課題3) データ駆動を実現するために、社内の業務システムだけでなく、IoTデータや顧客向けサービスのログデータなど、増加し続ける大量データに対応する必要があり、システムの拡張性・柔軟性が求められる。しかし、オンプレミスのシステムではハードウェアの調達も含め事前に計画する必要があり、迅速な対応ができない。

3. クラウドDWHによるデータ活用基盤

2章で述べた課題に対応するため、HSEではクラウドDWHによるデータ活用基盤の構築を提案している。クラウドDWHは、クラウド上でマネージドサービスとして提供されているDWHであり、各クラウドベンダから Amazon Redshift, Microsoft Azure Synapse Analytics, Google BigQueryなどが提供されている。また、マルチクラウドのサービスであるSnowflakeが急成長して注目されており、2021年11月から日立ソリューションズグループでも取り扱いを開始している²⁾。

(課題1) については、クラウドDWHの各サービスベンダが、数多くのETL、データレプリケーション、AI、BIの製品やサービスとの連携に取り組み、自社のサービスを中心としたエコシステムを形成している。これにより多くの場合、利用者はシステムを改修することなく、多様なデータソースを利用でき、様々なAI・BIの分析ツールやサービスと柔軟な連携が可能である。(課題2) についても、クラウドDWHはサービスベンダが攻撃の監視や脆弱性への対応などのセキュリティ対策を担うマネージドサービスとして提供されている。そのため、利用

者はセキュリティ対策の運用負荷を軽減でき、ユーザの利便性向上のための作業に注力できる。(課題3) については、クラウドのメリットである柔軟な拡張性を活かし、データ量やリクエスト数の増加に合わせて動的にストレージやVMを追加してスケールアップ・スケールアウトを行うことで、ユーザはハードウェアを意識することなく即座にシステムを拡張できる。このように、クラウドDWHは従来のオンプレミスのDWHの持っている複数の課題を解決でき、データ駆動ビジネスを実現するデータ活用基盤としてより適した特長を備えている。

4. データ駆動ビジネスの実現に向けたクラウドDWH活用イメージ

HSE では自社の業務ソリューションと連携するデータ活用基盤として、新たなクラウドDWHのサービスである Snowflake を提供できるようにすることを検討している。図1にHSEの業務ソリューションとSnowflakeを例としたクラウドDWHの活用イメージを示す。業務アプリ構築プラットフォームAppSQUARE³⁾で構築した業務アプリのデータを、Fivetran⁴⁾やCData Sync⁵⁾などのデータレプリケーションのサービスやツールで、データベースごと Snowflake に複製して定期的な同期を実行する。データソースはAppSQUAREに限らず、社内の既存の業務システムやIoTデータ、あるいは顧客向けサービスのログデータなどであってもよい。様々なデータをクラウドDWHに取り込み、一元的に集約して分析することで、データからより深いインサイトが得られるようになる。Snowflakeに蓄積された業務データは、クエリにより統計解析・機械学習の分析ツールであるuniSQURE MLにキャッシュされ、機械学習用の訓練

データとして使用される。学習済みの機械学習モデルは予測や診断に用いられ、その結果は再び Snowflake にロードされる。この予測・診断結果は業務データとともに BI から参照でき、例えば uniSQUARE BA では検索や可視化、レポートの作成に予測・診断結果を利用できる。これにより、ユーザはデータに基づいた分析・予測の結果を利用してビジネスの意思決定や課題の解決を行うことができ、データ駆動ビジネスの実現が可能になる。

5. AppSQUARE とクラウド DWH の連携試作

4 章のクラウド DWH 活用イメージを実装し、HSE の業務ソリューションとクラウド DWH が実際に連携できるか評価するため、AppSQUARE と Snowflake を例に業務ソリューションとクラウド DWH の連携試作を実施した。業務アプリデータとして PC 修理受付業務での PC 故障データと、部署ごとの出張申請データの 2 種類のデモデータを用いた。また、BI ツールとして Power BI Desktop を使用した。AppSQUARE の既存のユーザが自身の参照権限でこれらのデータを分析するために必要な機能要件を、以下の 3 ステップで検証した。(Step1) クラウド DWH と BI の連携。(Step2) 業務アプリのアクセス権を考慮したクラウド DWH のアクセス制御。(Step3) クラウド DWH へのデータ連携。各検証項目の結果を表 1 に記す。

表 1 連携試作の検証結果

| 検証項目 | 検証結果 |
|--|--|
| (Step1) クラウドDWH とBIの連携 | ・ AppSQUAREの業務アプリデータにマスタデータや他の業務システムの情報を統合した情報をBIで分析可能 |
| (Step2) クラウドDWH のアクセス制御 | ・ ロールベースのアクセス制御で、AppSQUAREと同じユーザ名・同じ参照権限で業務アプリデータを分析可能 |
| (Step3) クラウドDWH へのデータ連携 ※AppSQUARE以外の業務ソリューションも可能 | ・ FivetranやCData Syncなどのデータレプリケーションにより業務アプリのローデータをそのままクラウドDWHに複製可能 |

(Step1) では、AppSQUARE のデータベースに含まれる業務アプリデータのテーブルを Snowflake のデータベースにコピーし、Power BI Desktop から Snowflake に接続して、それらのテーブルを参照した。各業務アプリデータの BI レポートの画面例を図 2 と図 3 に示す。PC 故障データに製品マスタ情報を統合することで、PC の製品別故障件数のトレンド分析をしたり、出張申請データに旅費精算情報や住所情報を統合することで、グラフやマップによる出張回数や出張費の比較分析を行うことができる。このようにデータ活用基盤としてクラウド DWH を用いることで、これらの多様なデータソースとの連携を容易に実現でき、AppSQUARE の業務アプリデータと統合した情報を BI で分析できるようになる。

(Step2) における AppSQUARE とクラウド DWH のアクセス制御の連携方式を図 4 に示す。AppSQUARE と同じユーザを自動的に Snowflake 上にも定義し、

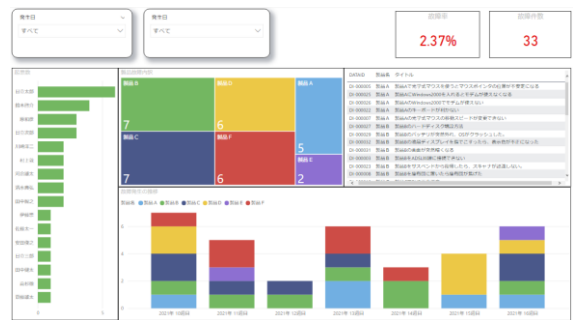


図 2 PC 故障データの BI レポート画面例

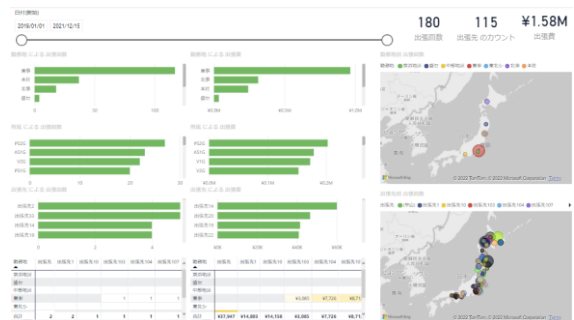


図 3 出張申請データの BI レポート画面例

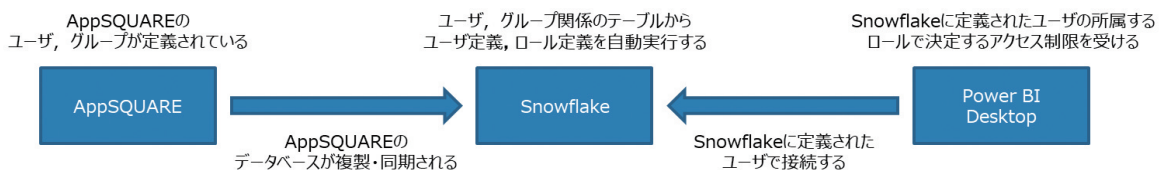


図 4 AppSQUARE とクラウド DWH のアクセス制御の連携

AppSQUARE で利用可能な業務アプリのデータを Snowflake 上でも参照できるようにロールベースのアクセス制御を行う。BI からは Snowflake に定義されたユーザで接続することで、AppSQUARE と同じユーザ名・同じ参照権限で業務アプリデータを分析できる。本試作により、このような制御が可能であることを検証した。

(Step3) の業務アプリデータの連携については、データレプリケーションのサービスである Fivetran とオンプレミスの製品である CData Sync を使用して、AppSQUARE が内部的に使用しているデータベースを Snowflake に複製して定期的に同期を実行する方式を検証した。データの変換や加工を行ってから DWH にロードする ETL と比較して、ローデータにアクセスできるため、データ分析の柔軟性が高い。また、テーブルごとにインタフェース機能を開発しなくてよいため、工数・期間・コストが発生しない。運用中にデータソースの仕様変更があっても DWH 側で一元的に対応でき、維持管理がしやすい。コンピュータリソースを柔軟に確保できるクラウド DWH により適した方式と考える。

6. おわりに

データ駆動ビジネスを実現するためのデータ活用基盤として、拡張性・柔軟性に優れたクラウド DWH は、従来のオンプレミスの DWH と比較して多くのメリットがある。本稿ではデータ活用基盤としてのクラウド DWH の活用イメージや、AppSQUARE を例とした連携試作の成果を通して、その具体的なメリットを述べた。HSE が提供する各種の業務ソリューションにデータ活用基盤としてクラウド DWH を追加して提供することで、付加価値としてこれらのメリットを提供でき、データ駆動ビジネスを促進できるものと考えている。

今後の課題として、顧客価値の高い業務アプリを具体的に選定し、業務アプリの利用から、機械学習によるモデルの作成、予測の実行、BI による可視化まで、図 1 に示した活用イメージをシームレスに実現できるか検証を行う予定である。その成果をソリューションのテンプレートとし、活用イメージに沿ったソリューション全体を短期間で構築して提供できるようにしていく。現在は AppSQUARE での検証を進めているが、工程管理・プロジェクト管理の SynViz S2 や、生産計画の SynPLA など、その他のソリューションも同様の仕組みで実現できる。今後は業務ソリューションのラインナップを広げていくことで、さらに顧客価値を高めていく所存である。

参考文献

- 1) DX 実践手引書 IT システム構築編 (2022/12/5 閲覧)
<https://www.ipa.go.jp/files/000094497.pdf>
- 2) クラウドデータプラットフォーム「Snowflake」を提供開始：(株)日立ソリューションズ (2022/12/5 閲覧)
<https://www.hitachi-solutions.co.jp/company/press/news/2021/1116.html>
- 3) 阿部一雄, 他: 企業の DX を加速させる業務アプリケーションビジネス革新, 日立ソリューションズ東日本 技報 第 26 号, 2021 年
- 4) Fivetran | Automated, reliable, and secure data pipelines (2022/12/5 閲覧)
<https://www.fivetran.com/>
- 5) 自動データレプリケーション | CData Software Japan (2022/12/5 閲覧)
<https://www.cdata.com/jp/sync/>



飯塚 新司 2008 年入社
研究開発部
クラウド DWH, データ分析, AI,
量子コンピュータ関連技術の研究



佐々木 和也 2004 年入社
アナリティクスソリューション部
アナリティクスソリューションの
企画, 設計, 開発, 顧客提供



齋藤 彰大 2018 年入社
アナリティクスソリューション部
Power Platform を活用したシステム
設計, 開発



及川 慎也 1991 年入社
アナリティクスソリューション部
アナリティクス事業の推進, IoT,
AI 活用コンサルティング



後藤 秀樹 1989 年入社
アナリティクスソリューション部
BI/BA システム設計, 構築



橋田 真人 2008 年入社
アナリティクスソリューション部
BI/DWH システムの設計, 開発, 顧客
提供