

AI や機械学習プロジェクトに積極的に取り組み、「全ての人の お金の課題を解決し、チャレンジできる社会」の実現へ。



「お金を前へ。人生をもっと前へ。」を ミッションに、金融サービスを提供

株式会社マネーフォワードは、2012年5月に創業で、「お金を前へ。人生をもっと前へ。」をミッションに掲げ、全てのお金の課題解決をめざす SaaS/Fintech 企業。個人向けお金の見える化サービス『マネーフォワード ME』や、事業者向けバックオフィス SaaS『マネーフォワード クラウド』など、50 を超える多様なサービスを提供している。昨今では、SDGs の目標達成に向けて、「User Forward」「Society Forward」「Talent Forward」の3つの重点テーマを設定し、「全ての人のお金の課題を解決し、チャレンジできる社会」の実現に貢献している。

データ量の増加に伴い データ分析環境の課題に直面

ビジネスポートフォリオは年々拡大しており、比例して分析対象となるデータソースの量や種類も増加している。実際のデータ分析シーンとしては、データから得たインサイトから経済学における学術論文作成や、CRM/デジタルマーケティング、管理会計向けのデータマネジメントなどが挙げられる。従前のデータ分析環境では、いくつかの課題に直面していた。

1つ目は、ETL 処理の高度化の必要性である。AWS EMR の場合、データフローが煩雑であり、エラー発生時の対応に工数がかかっていた。また、過去の任意の時点のデータセットを参照することもできなかった。2つ目は、分析パフォーマンスの改善である。AWS Athena での分析運用において、パフォーマンスに改善の余地があると感じていた。3つ目が、柔軟なデータ分析環境の実現である。分析環境の細分化による属人化やデータのサイロ化が進んでおり、データセットやノウハウの共有が困難であったことに加え、一貫したデータセキュリティやガバナンスの担保にも課題があった。



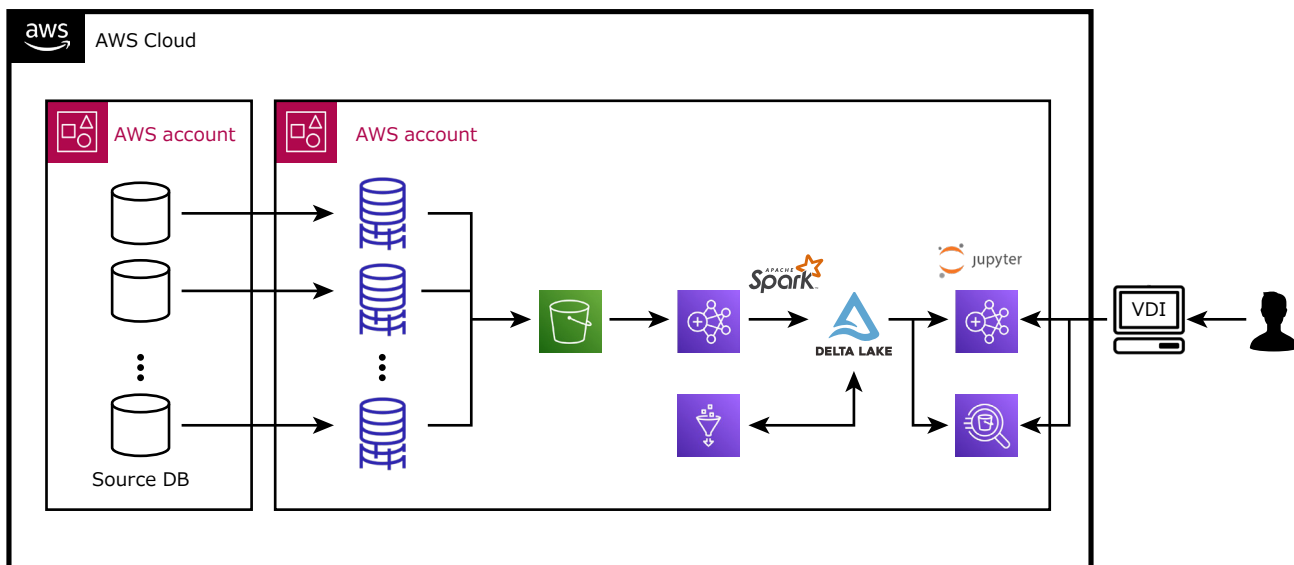
(左)CTO室 副室長 松本 裕也 氏
(中)CTO室 分析基盤部 リーダー 長谷川 慎一郎 氏
(右)CTO室 分析基盤部 山崎 隼也 氏

Databricks Delta の活用で バッチ処理時間を 1/2~1/3 に短縮

データブリックスのデータ・インテリジェンス・プラットフォームの主要コンポーネントが馴染みのある Delta Lake や Spark であることもあり、データブリックスの採用を前提に、POC (Proof of Concept) を実施した。Databricks Delta を活用し、ETL 処理の POC を実施した結果、1 テーブル数テラバイトのデータセット + 数十ギガバイトのバッチ処理に関して、以前は 2~3 時間かかっていたものが、1 時間以内に短縮できることが確認できた。加えて、データ・インテリジェンス・プラットフォームの大きな特徴として、過去の任意の時点のデータセットを参照可能なタイムトラベル機能が備わっていたり、複数のデータソースから抽出するデータを Delta 化することによって一元的にデータ管理ができるため、データサイロを解消することも可能である。また、データガバナンスに関しても、将来的に Unity Catalog を活用することによって、ワークスペースにおけるユーザーやデータへのアクセスを集中的に管理できる。

今後は、データ・インテリジェンス・プラットフォーム採用によるデータ基盤のモダナイズ化が実現したため、これをもとに、バックオフィス業務の自動化などの分野において、AI や機械学習プロジェクトに積極的に取り組む予定である。

アーキテクチャー図 (導入前)



アーキテクチャー図 (導入後)

