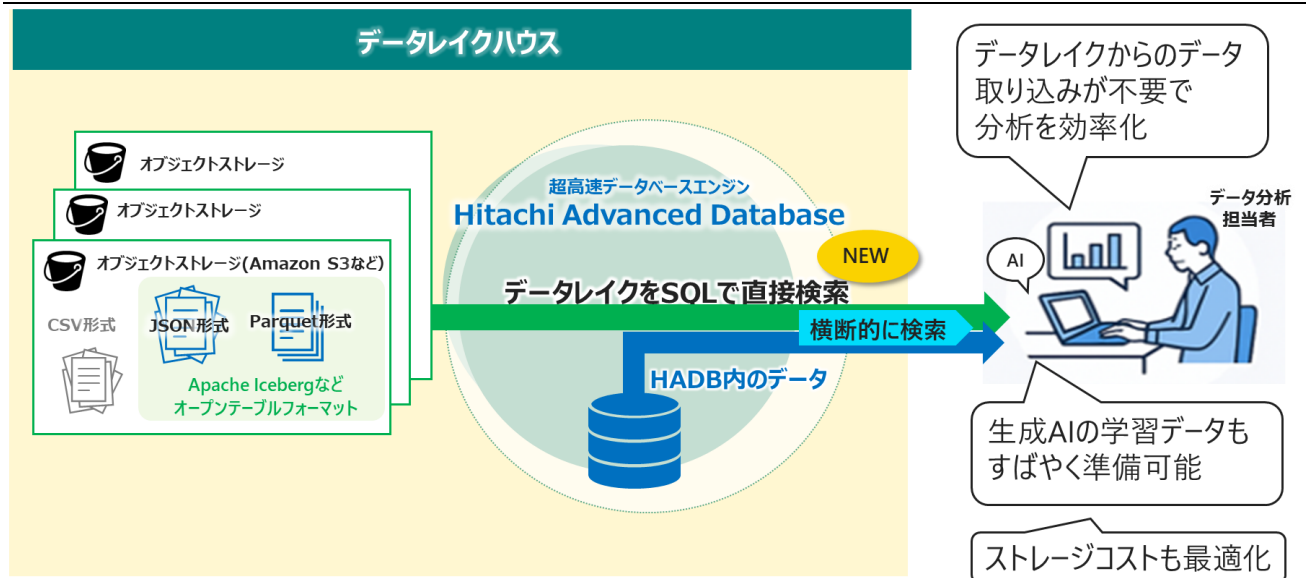


多様なデータ活用を容易にするデータレイクハウス向けに進化した 超高速データベースエンジン「HADB」の最新版を提供開始

取引履歴データなどの分析による新サービスの創出や、生成 AI 学習データの整備の加速、
分析データ保管コストの最適化などを支援



HADB 最新版の利用イメージ

株式会社日立製作所(以下、日立)は、多様なデータの活用を容易にするデータレイクハウス向けに進化した、超高速データベースエンジン「Hitachi Advanced Database」^{*1}(以下、HADB)の最新版 Version 6 を 10 月 1 日から提供開始します。

具体的には、データレイクとして一般的に利用される Amazon S3^{*2}などのオブジェクトストレージ上で、各種データ分析や生成 AI の学習などに用いられるオープンテーブルフォーマット^{*3}のデータを、HADB から SQL で直接検索できるようになりました。これにより、データ分析担当者は、データレイク内で分析対象となるデータの事前加工や HADB への取り込み作業が不要となり、HADB 内のデータと合わせた横断的な検索・分析が可能となります。例えば、新サービス創出に向け、さまざまなデータを組み合わせることで多角的に分析でき、サービスや取引履歴の分析業務の効率化が図れます。加えて、生成 AI を業務に活用する際にも、データレイク内のデータ検索をすばやく行えるため、大量の学習データを早く整備でき、生成 AI の学習効率の向上にもつながります。

また、HADB が備える非順序型実行原理^{*4}による並列処理で、一般的に、オブジェクトストレージがブロックストレージに比べてデータの読み出し速度が遅いという点をカバーし、高速なデータ検索が可能です。これにより、データ保管コストの低いオブジェクトストレージを利用しながらも、データウェアハウス水準の高速なデータ分析を実現できます。

さらに、一般的なクラウドサービスは検索データ量に応じた従量課金であるのに対し、HADB 最新版では、従来のデータ容量ベースのライセンスに加え、データレイク上のデータを容量制限なく定額で検索できる新しいライセンスも選択できます。これらにより、データ保管コストの最適化にも寄与します。

なお、日立では、2025 年度中に、オンプレミス環境や拠点ごとのデータレイクに適した Amazon S3 互換オブジェクトストレージ「Hitachi Virtual Storage Platform One Object」を提供開始予定です。これにより、HADB のオンプレミス向けデータレイクソリューションの強化を進めるとともに、データレイク検索のさらなる高度化や今後の AI 技術の進化を

見据え、ベクトルデータやグラフ構造など、データの意味や属性まで考慮した類似検索技術への対応も行うなど HADB の機能強化を図っていきます。これらの取り組みにより、HADB の高速処理能力をさまざまな分野で活用し、企業の DX の加速を支援していきます。

- *1 今回の Version 6 提供開始とあわせて、製品名称を「Hitachi Advanced Data Binder」から「Hitachi Advanced Database」に変更。
- *2 Amazon Web Services が提供するクラウドストレージサービス。
- *3 Parquet、Apache Iceberg など各種のデータ形式、および、それらを対象に、SQL など直接検索・更新・分析可能とするためのメタデータ管理とトランザクション制御の仕組み。Parquet は列指向でデータ圧縮効率が高く、Apache Iceberg はデータの効率的な管理とクエリ処理が可能な点が特長。データ分析や生成 AI の学習データなどに活用される。
- *4 喜連川 情報・システム研究機構 機構長／東大特別教授、合田 東大教授が考案した原理。

■背景

近年、企業では、新サービスの創出や生成 AI を活用した業務効率化に向け、データ利活用の重要性がさらに高まっています。そのため、データ分析担当者は、新サービス創出につながるアイデアを得るため、多様で大量のデータをさまざまな観点から多角的に繰り返し分析しています。その際に必要となる大量のデータは、保管コストの面からオブジェクトストレージを利用したデータレイクに格納されることが一般的です。しかし、分析のたびにデータレイク内のデータを加工し、データベースに取り込む作業が分析担当者の大きな負担となっていました。

日立はこれまで、ビッグデータ利活用の促進を目的に、大量のデータの複雑な分析を高速に実行する非順序型実行原理を備えた HADB を、製造・流通・金融・社会・公共など、さまざまな分野のデータウェアハウスに適用してきました。今回、企業の DX をさらに加速させるため、データウェアハウスのみならず外部のデータレイクに対しても、SQL により柔軟かつ高速なデータ検索が可能なデータレイクハウスを実現する HADB の最新版を提供します。

■HADB 最新版の特長

今回の最新版では、オブジェクトストレージ上のファイルやオープンテーブルフォーマットに対して、SQL による直接検索が可能になりました。具体的には Amazon S3 などのオブジェクトストレージ上の CSV、JSON、Parquet 形式データや Apache Iceberg テーブルを、データの加工や HADB への取り込み作業なしで、直接 SQL で検索できます。また、オブジェクトストレージのデータ読み出し速度を HADB 独自の非順序型実行原理による並列処理でカバーし、高速な検索が可能です。さらに、検索した結果は、HADB の機能により、オブジェクトストレージ上で Parquet ファイルとして出力できるため、大規模分析ワークロードや機械学習パイプラインに用いられるデータとして手間なく活用できます。これらの特長により、たとえば以下のような複雑な分析やデータ利活用のさまざまな場面で利用可能です。

■HADB の活用例と効果

1：サービスや商品の履歴、購買などのデータ分析の多角化や効率化が容易に

企業が提供したサービスや商品の品質向上などのため、履歴データの追跡・検証や購買データを用いて顧客の傾向分析などを行う際、データレイクから HADB へその都度データを加工して取り込む作業が不要となります。これにより、データレイク内のデータと HADB 内のデータを横断的に検索できるため、過去と現在のデータを掛け合わせるなど、より多角的な観点での分析が容易となり、検証や分析、予測の効率化や高精度化などが図れます。

2：生成 AI の学習データ加工エンジンとしても利用でき、業務への生成 AI 活用を促進

データレイクに蓄積したデータを生成 AI の学習データとして用いる場合には、学習データはあらかじめ ETL ツールなどで構造化データに変換する作業が必要であり、こうした手間を減らすことが大量のデータをすばやく準備したい生成 AI の学習において重要です。

HADB は、データレイク上の学習データをそのまま構造化データとして扱い検索できるため、多くの生成 AI がそのまま学習可能な Parquet 形式へ出力できます。そのため、学習前のデータの加工処理が不要となり、学習データの準備を加速します。これにより、与えるデータの量が品質を左右する生成 AI の学習効率を向上でき、業務への生成 AI 適用を促進します。

3：用途や利用頻度に応じてデータ保管場所を最適化でき、コストを抑えつつ迅速な検索が可能

オブジェクトストレージ上のデータも高速検索でき、デーウェアハウスのデータと横断的な検索が可能です。そのため、期毎や期間別の分析などで用いる利用頻度が低い分析データは、保管コストが安価なデータレイク上に格納し、利用頻度が高く、より高速な検索が求められるデータは、HADB に格納するといった使い分けが可能です。これにより、ストレージコストの最適化が図れます。また、利用頻度が低下したデータをデーウェアハウスからデータレイクへ移した場合でも、監査対応や品質保証などの業務で、急遽データが必要になった際、再度データを取り込まずに HADB から直接 SQL で検索できるため、必要なデータを迅速に分析に活用することが可能です。

新製品の価格および提供時期

名 称	概 要	標準価格 (税別)	提供開始日
Hitachi Advanced Database Version 6	外部のデータレイクに対し SQL で直接検索を実現	サブスクリプション ^{*5} 年額 3,834,000 円 ～	10 月 1 日

*5 買い取り型での提供も可能

Hitachi Advanced Database について

HADB は、大量データの高速かつタイムリーな分析を可能とする標準 SQL に対応した国産のリレーショナルデータベースです。本製品は、データ処理可能なところから並列に複数タスクを実行することで、サーバ・ストレージの能力を最大限に使い切る「非順序型実行原理」に基づき、超高速検索処理を可能としています。また、集計や分析に影響を与えない高速データインポート機能により、日々生み出される多様な大量のデータをほぼリアルタイムで反映できるため、お客さまの新たなビジネス価値の創出に寄与します。これらの特長は、パブリッククラウド上でのデータ検索にも活用でき、より柔軟で迅速なデータ分析を実現します。

関連ウェブサイト

超高速データベースエンジン Hitachi Advanced Database

<https://www.hitachi.co.jp/data-binder/>

商標注記

記載の会社名、製品名などは、それぞれの会社の登録商標もしくは商標です。

お問い合わせ先

株式会社日立製作所 AI & ソフトウェアサービスビジネスユニット マネージド & サービスプラットフォーム事業部

お問い合わせフォーム：<https://www.hitachi.co.jp/it-pf/inq/NR/>