

2019年7月1日

IoT データなど大規模データの分析のさらなる高速化により お客さまのビジネス価値創出の迅速化を支援する 超高速データベースエンジン「Hitachi Advanced Data Binder」 最新バージョン 05-00 を提供開始

株式会社日立製作所(以下、日立)は、このたび、IoT データなど大規模データの分析のさらなる高速化により、お客さまのビジネス価値創出を支援する超高速データベースエンジン「Hitachi Advanced Data Binder」^{*1} の最新バージョン 05-00(以下「HADB V5」)を、7月1日から提供開始します。今回、IoT センサーや取引のログなど大規模データの集計や大規模データどうしの関連付けといった分析で多く用いられる処理の高速化を図りました。あわせて、ライセンス体系の変更により、増え続けるデータ量に対応するシステム拡張時のコストの抑制も実現します。これらにより、HADB は、今後ますます大規模化するデータの高速分析を可能とし、デジタル変革や新たなデジタルサービスの開発などビジネス価値創出の迅速化を支援します。

近年、企業においては、市場における競争力を強化し続けるため、センサーログなどの IoT データやオープンデータ、取引履歴など、さまざまな種類の大規模データを AI 技術などを活用して分析し、デジタル変革や新たなサービス創出につなげる取り組みが活発化しています。デジタル化が進展し、取り扱うデータの規模がますます大きくなる中、複数の大規模データを関連付けながら高速な分析を将来にわたり最適なコストで行うことが求められています。

日立は、2012 年に非同期的にデータを処理しハードウェアの処理性能を最大限に引き出す「非順序型実行原理」^{*2}を利用した超高速なリレーショナルデータベースとして HADB を市場に投入して以来、データ分析業務に影響を与えずにデータベース内にバックグラウンドでデータを高速に取りこむ機能や、データ特性にあわせた格納形式^{*3}の拡充などの強化を続けてきました。また、標準規格の SQL に対応し、Pentaho などの ETL^{*4} や BI^{*5}など各種ツールとオープンに連携可能な HADB は、クラウド環境での構築も可能です。これらの特長を生かし、HADB は、金融、流通、小売、公共、交通、ヘルスケアなど、さまざまな分野で活用されています。

今回提供開始する「HADB V5」では、IoT データなど大規模データの分析で多く用いられる処理の高速化に加え、増え続けるデータ量に対応するシステム拡張時のコストの抑制も実現しました。

さらに、今後、HADB では、クラウド環境上でのさらなる処理高速化など、大規模データの利活用のさまざまな要望に対応する進化を続け、お客さまのビジネス価値創出に貢献していきます。

*1 内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的サービスの実証・評価」(中心研究者：喜連川 東大教授/国立情報学研究所所長)の成果を利用。

*2 喜連川 東大教授/国立情報学研究所所長・合田 東大准教授が考案した原理。

*3 HADB はローストア形式とカラム形式の2つのデータ格納形式をサポート。

ローストア形式は、データ表の行単位でデータを格納する形式。商品マスター在庫情報など参照検索処理が速い。

カラムストア形式は、データ表の列単位でデータを格納する形式。各種ログデータなど列単位の集計や検索処理が速く、データ圧縮しやすい。

*4 ETL:Extract Transform Load

*5 BI : Business Intelligence

■「HADB V5」の特長

1. ログデータの集計やデータどうしの関連付けなど、大規模データ分析で多く用いられる処理を高速化

HADB では、データ格納形式の拡充とあわせて、センサーのログや取引履歴など時系列データの集計や検索処理に適したカラムストア形式の参照処理の高速化などに取り組んできました。カラムストア形式では、データベースのインデクスのチューニングが不要であり、データの圧縮によるストレージアクセスの削減効果も期待でき、柔軟かつ迅速な分析が可能です。

今回の「HADB V5」ではさらに、カラムストア形式でのデータベース更新処理における効率化を進めました。具体的には、データの追加や更新後に、性能維持のために必要であった追加・更新部分のデータ格納形式の変換を自動的に行なう機能を追加したことにより、データベースの運用負荷を軽減し、かつ、集計処理の高速性を維持します。

また、大規模データどうしを関連付けて分析する場合に利用されるデータ表の結合機能も強化しました。ユーザーが指定した結合の条件に該当しないデータの読み込みを自動的にスキップする機能の追加により、結合処理を高速化^{*6} することができます。これらの処理の高速化により、IoT データなどの大規模データをさらに高速に分析可能となりました。

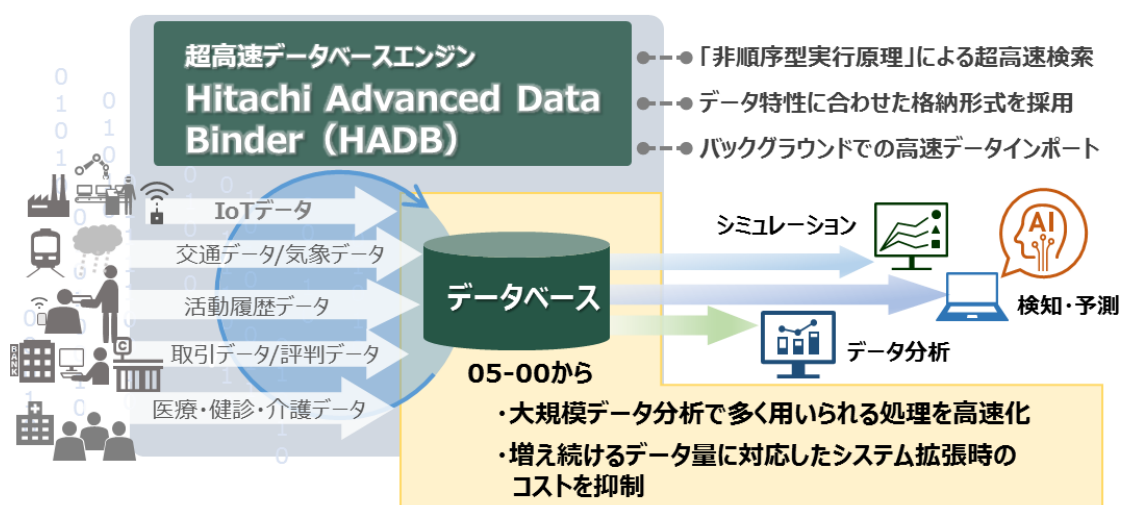
*6:センサーデータなどの分析で数千億レコードの表の結合を行なう業務を想定した例において、読み込み量を従来の 150 億レコードから 30 億レコードまで削減。(日立調べ)

2. ライセンス体系の変更により、増え続けるデータ量に対応するシステム拡張時のコストを抑制

2019 年 1 月より HADB は、サーバなどのハードウェアと一体化したプラットフォームとしての提供からソフトウェアでの提供へと変更し、柔軟なシステム構成を可能にしてきました。

今回、「HADB V5」において、ソフトウェア提供におけるライセンス体系をさらに変更^{*7} し、システム拡張時のコストを抑制します。HADB は処理するデータ量の増大に対してプロセッサを増強するスケールアップで性能を拡張できますが、今回の変更により、プロセッサを増やしても新たなライセンス費用は発生しないため、システム拡張時のコストを抑制できます。また、カラムストア形式では連続するデータをまとめて圧縮可能なため、ストレージ容量の増加を抑制できます。これらの改善により、導入時だけでなく将来の拡張にわたりシステムコストの抑制を実現します。

*7:インストールするサーバのプロセッサ数とストレージのデータ容量に基づくソフトウェアライセンス体系から、インストールするサーバ台数に基づくライセンスとユーザーデータ領域のデータ容量に基づくライセンス体系に変更。



図：HADB の特長と適用イメージ

■ HADB の適用イメージ

HADB は、金融業での不正取引検知、流通・小売業での顧客嗜好解析、公共分野での防災シミュレーション、ヘルスケア分野での医療・健診・介護データの分析など、さまざまな分野へ適用可能です。

HADB の高速性を生かした具体的な業務への適用例は以下の通りです。

(1) IoT データなどの高速な反映および AI による検知・予測への適用例

IoT データを活用した交通インフラの混雑予測や、エネルギー分野での広域送電網分析によるブラックアウトの予測、AI によるインフラ設備の故障予兆検知や、製造ラインの人員配置の最適化など、鮮度の高いデータの活用が望まれるシステムへの適用が可能です。

(2) パーソナルデータ活用時のプライバシー保護のための匿名加工の高速化への適用例

個人の特定が危惧されるパーソナルデータ活用において、事業者間でデータを受け渡す場合には、個人を識別できないよう匿名加工処理を施す必要があります。データ活用の視点で有用な情報量を確保するためには、データを抽出する範囲や加工する単位などを調整しながらデータの検証を繰り返す必要があります。データが大規模になるほど分析者の工数が膨大になるという課題がありました。こうした業務に HADB を適用することで、匿名加工処理にかかる時間を大幅に短縮^{*8}し、対話的に調整を行いながらデータの抽出・加工・検証を行うことが可能です。

^{*8} 2018 年 11 月 20 日 ニュースリリース「大規模データの匿名加工処理を高速化する技術を開発」
<https://www.hitachi.co.jp/New/cnews/month/2018/11/1120.html>

新製品の価格と提供開始時期

名称	新製品の特長	価格(税別)	提供開始時期
超高速データベースエンジン Hitachi Advanced Data Binder 05-00	<ul style="list-style-type: none">ログデータの集計やデータどうしの関連付けなど、大規模データ分析で多く用いられる処理の高速化システム拡張時のコストの抑制	サーバインストール ライセンス ^{*9} : 700 万円 容量ライセンス ^{*10} : 300 万円～	7 月 1 日

^{*9} クラスタ構成時には、サーバインストールライセンスの追加が必要(プロセッサ数制限なし)。同時に接続する最大数に基づく。

^{*10} 容量ライセンスは、ユーザーデータを格納する DB エリアに基づく。10TB あたり 300 万円、100TB あたり 2000 万円。

Hitachi Advanced Data Binder(HADB)について

超高速データベースエンジン「Hitachi Advanced Data Binder」は、リレーショナルデータベースの検索処理はレコードの処理順序に依存しないという点に着目し、データ処理開始が可能なところから非同期的にデータを処理しハードウェアの処理性能を最大限に引き出す「非順序型実行原理」を利用し、自社従来比約 100 倍にデータ検索性能を実現しています。超高速検索性能に加え、データ特性に合わせた格納形式を採用。また、データ分析処理に影響を与えずに、バックグラウンドでデータを高速に取り込む機能により、鮮度の高いデータの活用が可能です。

[超高速データベースエンジン「Hitachi Advanced Data Binder」](https://www.hitachi.co.jp/data-binder/)

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

商標に関する表示

- ・ Pentaho は、Hitachi Vantara Corporation の日本及びその他の国における商標または登録商標です。
- ・ 記載の会社名、製品名は、それぞれの会社の商標もしくは登録商標です。

お問い合わせ先

株式会社日立製作所 サービスプラットフォーム事業本部 IoT・クラウドサービス事業部

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

以上