

# 超高速データベースエンジン

## Hitachi Advanced Data Binder(HADB)のご紹介

2024/07/11

株式会社 日立製作所  
クラウドサービスプラットフォームビジネスユニット  
マネージド&プラットフォームサービス事業部 DB部

# Contents

---

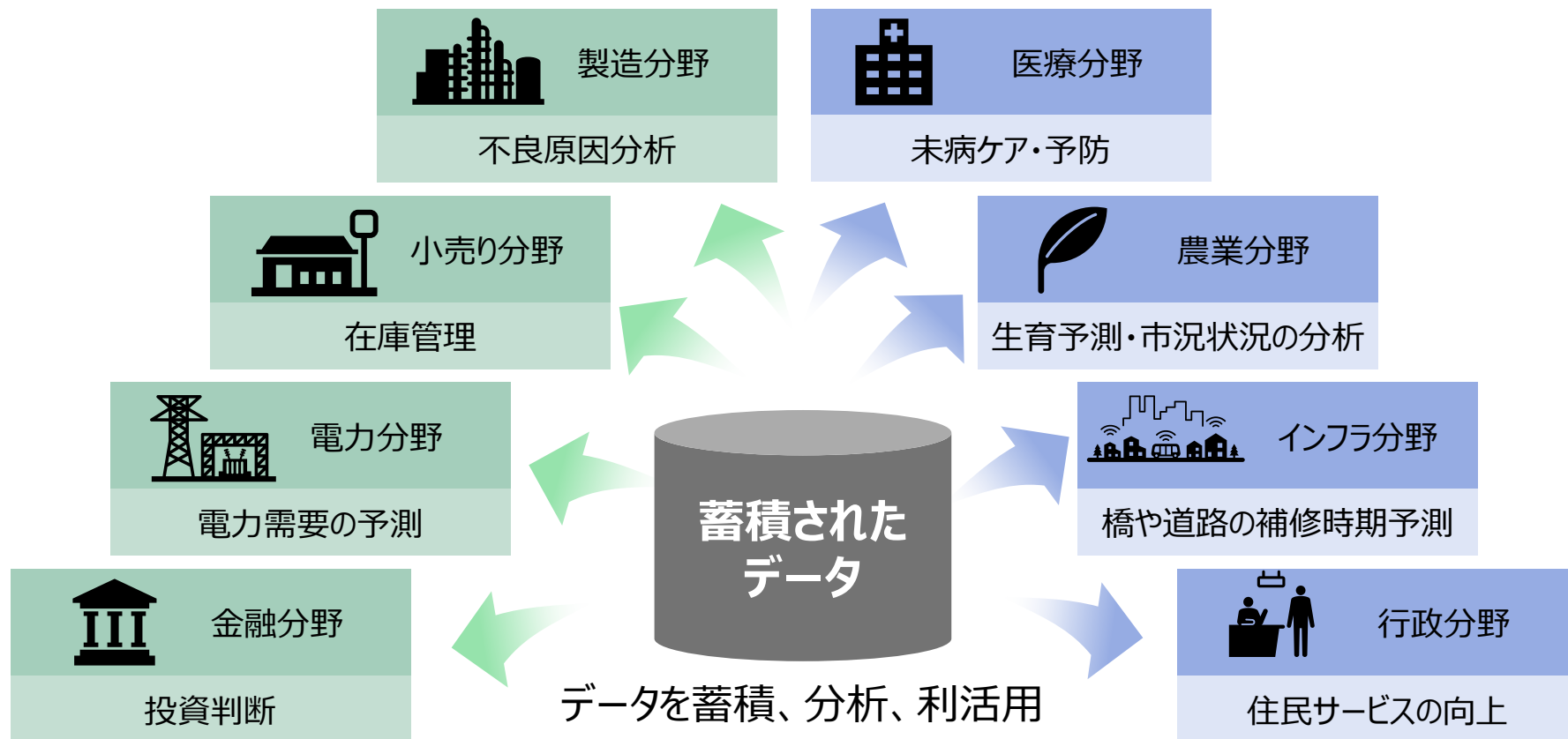
1. 背景
2. HADBの性能面での特長
3. クラウド対応における特徴
4. その他の特徴
5. ライセンス・サポート体系と価格
6. 導入事例

# Contents

---

- 1. 背景**
2. HADBの性能面での特長
3. クラウド対応における特徴
4. その他の特徴
5. ライセンス・サポート体系と価格
6. 導入事例

近年、サービスの質や生産性の向上、事業に掛かるコストの最適化などを目指し、製造分野から金融分野や、医療分野から行政分野など、さまざまな分野でデータの利活用が次々と広がっています。さらに、企業活動や公共事業などでのサービスの高度化やDX※の推進などにより、今後も活用するデータの多様化・大規模化が進んでいきます。



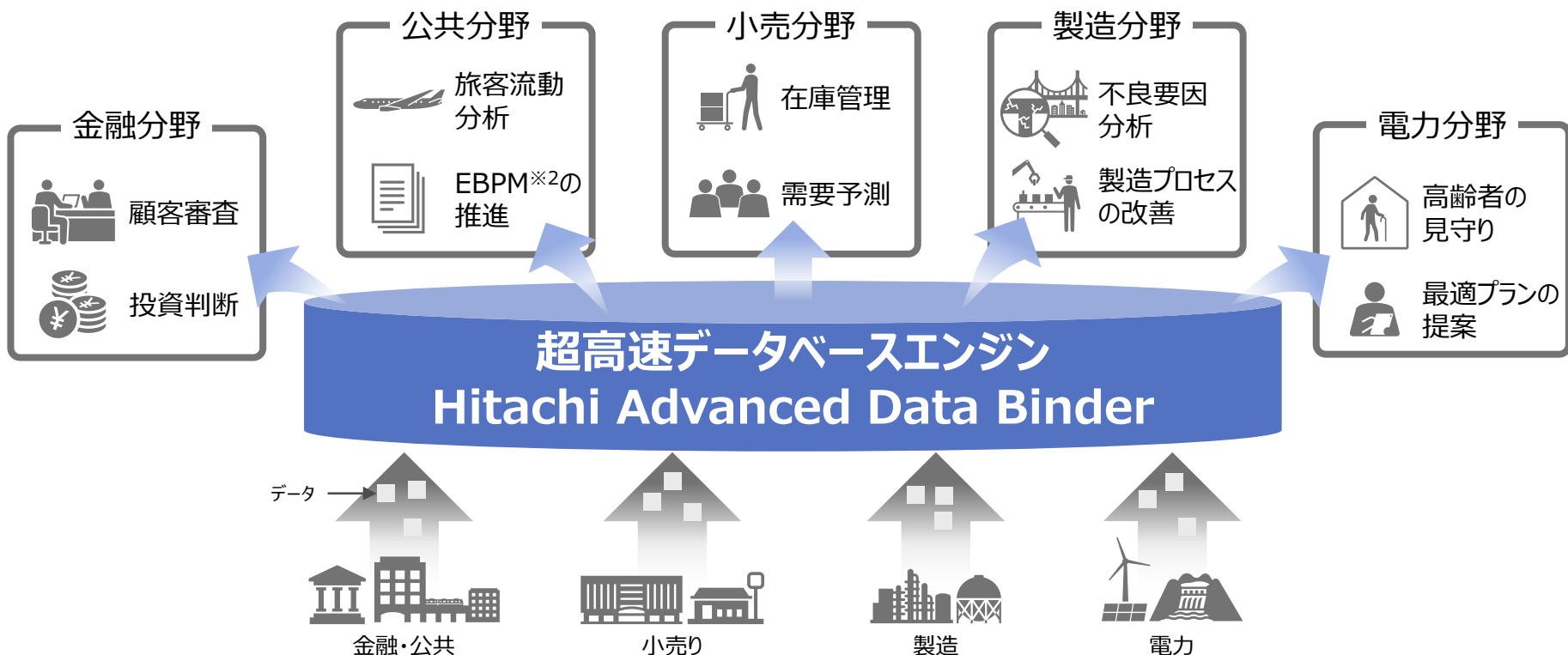
データ利活用の広がりにより、企業活動で活用するデータが多様化・大規模化

# Hitachi Advanced Data Binderの誕生

変化の激しい時代で、より多くの時間を施策・立案にかけるには、大量のデータをタイムリーに活用して、個々のニーズに素早く対応する必要があります。

データをタイムリーに活用するためには、データを素早く抽出して分析することが必要です。

超高速データベースエンジン Hitachi Advanced Data Binder<sup>※1</sup>（以降、HADB）は、独自技術による超高速検索、柔軟なデータ格納方式、高速データ取り込みなどの機能により、タイムリーな分析を実現します。



※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 情報・システム研究機構 機構長/東大特別教授)の成果を利用。

【社外サイトURL】: <https://www.hitachi.co.jp/Prod/comp/soft1/data-binder/index.html>

※2) EBPM (Evidence Based Policy Making) : 証拠に基づき政策立案

HADBは大量データの分析・集計処理に特化した分析系システム用のリレーショナルデータベースです。参照を中心として、過去に蓄積した大量データから広範囲に分析する用途に適しています。

## 基幹系システム

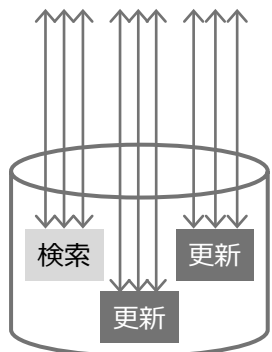
(Oracle、HiRDB、PostgreSQLが得意)



### 具体例

- ATM
- 旅券予約システム

- ✓ 細かい更新中心の定型クエリ
- ✓ 多くの小さなクエリを短時間で処理
- ✓ 同時接続数も多い。
- ✓ 当日の取引データが中心
- ✓ 数GB～数10GB/日のデータ



## 分析系システム

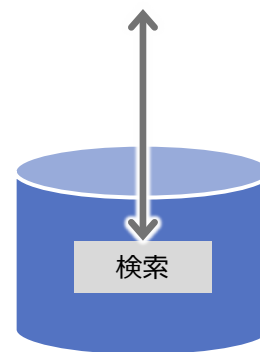
(DB2 Warehouse、Teradata、**HADBが得意**)



### 具体例

- 優良顧客分析
- 趣味嗜好分析

- ✓ 参照中心の非定型を含むクエリ
- ✓ 広範囲のデータを対象とするクエリを処理
- ✓ 同時接続数は少ない。
- ✓ 過去蓄積されたデータが中心
- ✓ 数100GB/日以上 of データ



分析系システムのこのような特性を踏まえ  
高速に処理できるのがHADBです

- 時系列に発生する大量データが発生しており、今後、蓄積、活用により価値を見いだすことを考えているお客さま

データ例：取引データ、購入履歴（POS<sup>※1</sup>データ）、センサーログ、稼働ログ、SNS<sup>※2</sup>、通話ログ など

- 蓄積した大量データから広い範囲（数億件レベル）を対象とする参照を行うお客さま

- 大量データどうしを掛け合わせて、傾向分析・予兆検知・不正検知といったデータ処理を行うお客さま

※1) POS (Point Of Sales)

※2) SNS (Social Networking Service)

さらなるデータの多様化、大規模化に向け、HADBは今後もお客さまの期待に応えていきます。

本資料で紹介している内容は、今後予告なく変更させていただく場合がございます。

2024/7  
New

## ■ HADB 05-00～05-03

- ・ カラムストア表運用容易化
- ・ サポートSQL構文の拡充
- ・ 大規模データ処理機能強化
- ・ IoTセンサーなどへの対応データ形式拡充
- ・ 対応BI<sup>※2</sup>ツールの拡大

## ■ HADB V1-V4

- ・ 基本機能の拡充
- ・ SQL機能強化
- ・ 大規模データ対応 (マルチノード、チャンクアーカイブ)
- ・ カラムストア表サポート
- ・ セキュリティ対応

## ■ TPC-Hに登録!

「TPC-H<sup>※1</sup>」の最大クラス (100TB) の性能評価リストに世界で初めて登録

## ■ HADB 05-10

- ・ クラウド対応強化 (低価格ストレージ使用時の性能安定性向上)
- ・ サポートSQL構文の拡充

## ■ HADB 05-11～(2024年以降の予定)

ノード間平行拡張、運用性向上、オートスケール対応、多種多様なデータモデル・データ形式への対応強化、一時表など

2025年

2024年

## ■ HADB 05-09

- ・ Azure対応強化 (Azure Blob Storageへの対応)
- ・ 表定義における各種上限値の拡大

2023年

## ■ HADB 05-08

- ・ AmazonS3対応フルサポート
- ・ チャンク運用改善

2022年

## ■ HADB 05-07

- ・ クラウド対応強化 低価格ストレージ(Amazon S3)への対応
- ・ Python対応

2020年

## ■ HADB 05-04～05-06

- ・ 繰返し構造対応 (配列型サポート)
- ・ スケールアウト対応 (SQL平行実行機能)
- ・ DB暗号化
- ・ 外部認証機構連携

2018年

2013年

※1) TPC-H : DWH (Data Ware House) 分野のデータベースの検索や抽出処理性能のベンチマーク

※2) BI (Business Intelligence) ツール : データに基づいた意思決定の支援ツール



# Contents

---

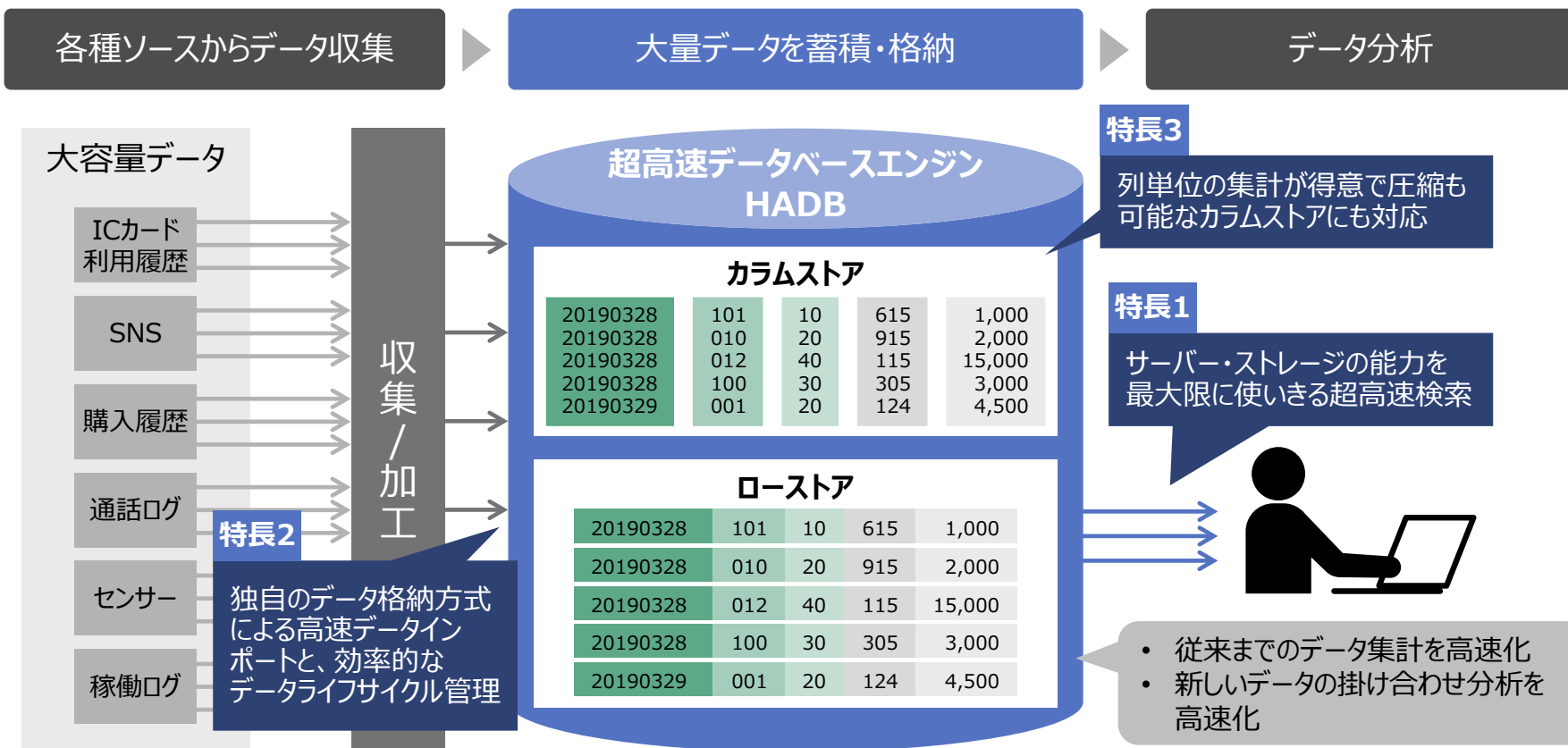
1. 背景
- 2. HADBの性能面での特長**
3. クラウド対応における特徴
4. その他の特徴
5. ライセンス・サポート体系と価格
6. 導入事例

大容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。

**特長1** :サーバー・ストレージの能力を最大限に使いきる**超高速検索**

**特長2** :独自のデータ格納方式による**高速データインポート**と**効率的なデータライフサイクル管理**

**特長3** :ローストアだけでなく集計・分析に多用される**列単位の集計が得意なカラムストア**

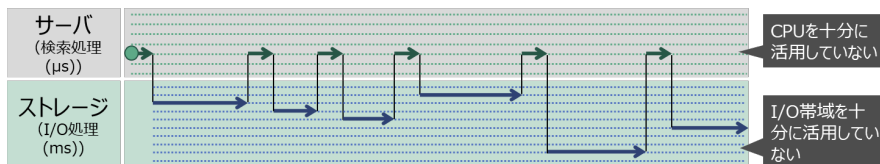


大規模データの超高速検索を「非順序型実行原理※1」で実現しています。

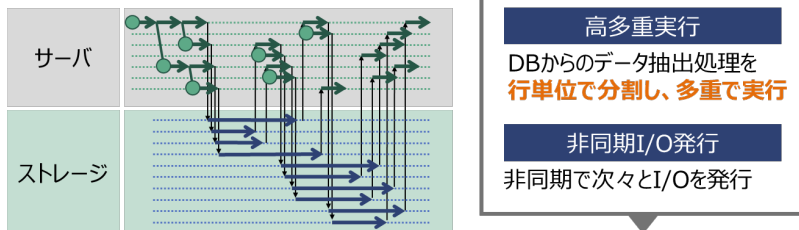
※1) 喜連川 情報・システム研究機構 機構長 / 東大特別教授、合田 東大教授が考案した原理。

専用のハードウェアは不要  
データ処理可能なところから並列に複数タスクを実行  
通常のサーバー・ストレージの能力を最大限に使い  
切り、超高速検索を実現

## 【従来方式】順序実行方式



## 【本方式】非順序実行方式

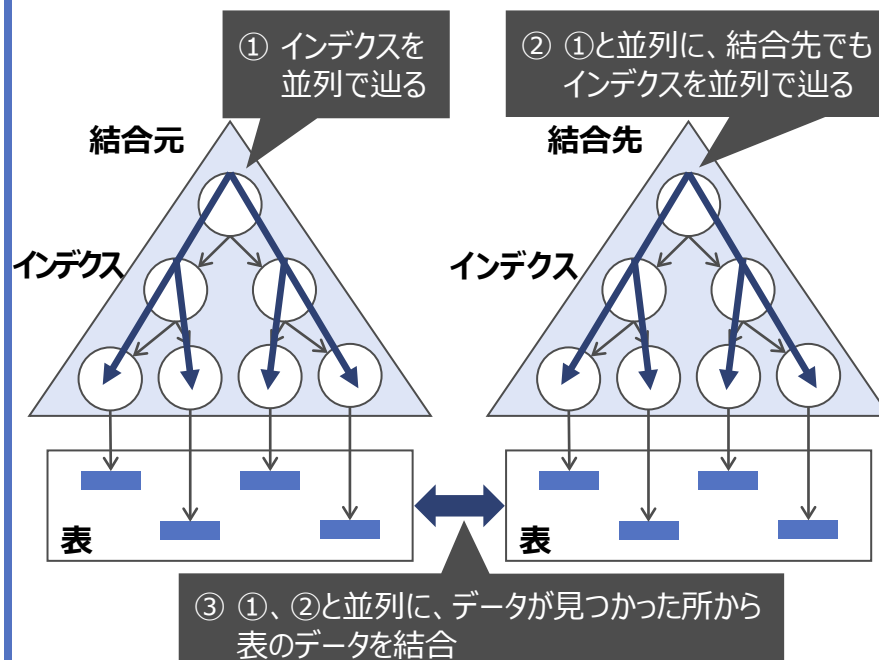


CPU、I/O帯域を  
十分に活用

凡例) ● : タスク割当、→ : 検索処理、→ : ディスクI/O ... : タスク割当待ち、... : I/O発行待ち

非順序型実行原理のイメージ

インデクスを使用した検索、表の結合、  
GROUP BY句によるグループ化処理、  
再帰検索などを、並列処理により高速化



処理高速化例 (インデクスを使用した検索による表の結合)

サーバー・ストレージの能力を使い切る超高速検索のHADBなら、データの増加に対して、ストレージ増設のみで対応可能です。

スケールアウト型のデータベースではデータ量が増加すると台数増加により、運用管理の手間が増加。  
HADBなら**数PB以上**になっても、**ストレージ増設のみでハードウェア性能を引き出し、高速に分析・集計処理。**

HADBは、インメモリDBのように**搭載メモリによる扱えるデータ量の制限は無く**、数100TBでも対応可能。カラムストアを利用すれば分析に必要な列にのみ効率的にアクセス可能。

数10TB

数100TB

数PB以上

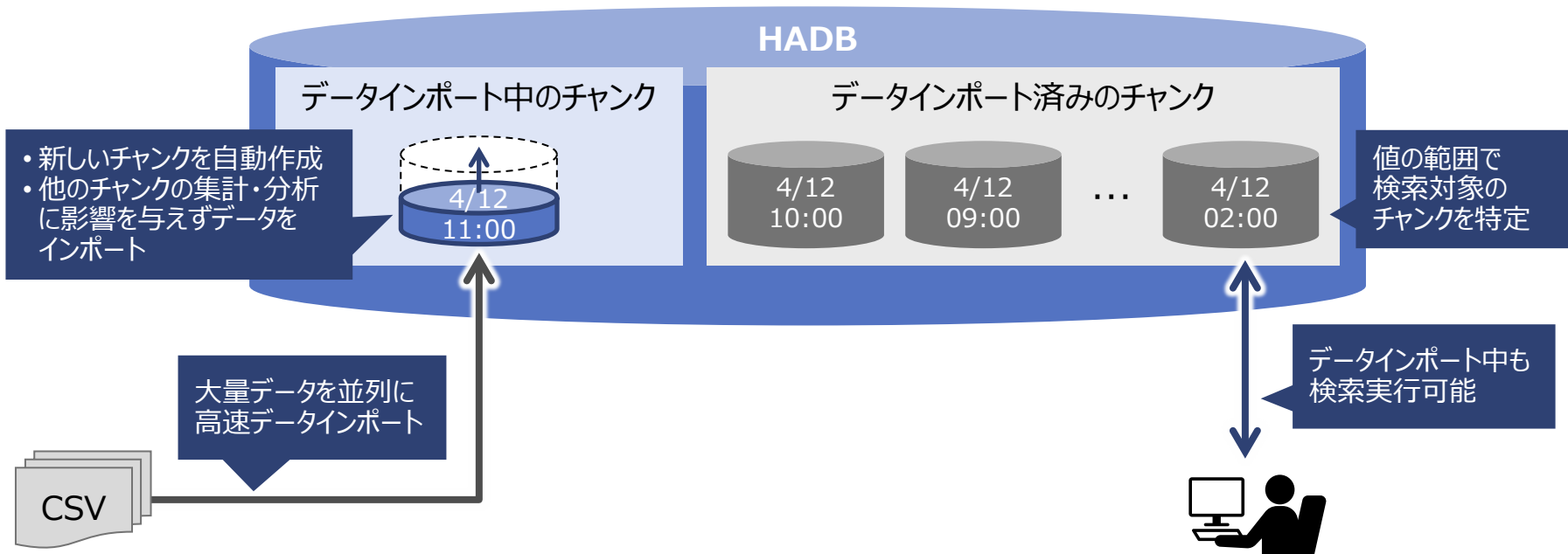
HADBは比較的**安価なライセンス**。  
10TBまではライセンス※1費用も変わらず、費用面でも導入が容易。

※1) ライセンスの詳細については、5章を参照。

**大量データを高速に並列インポートします。** 検索を効率化する為にデータインポートごとに「チャンク」を作成、自動でパーティショニングします。

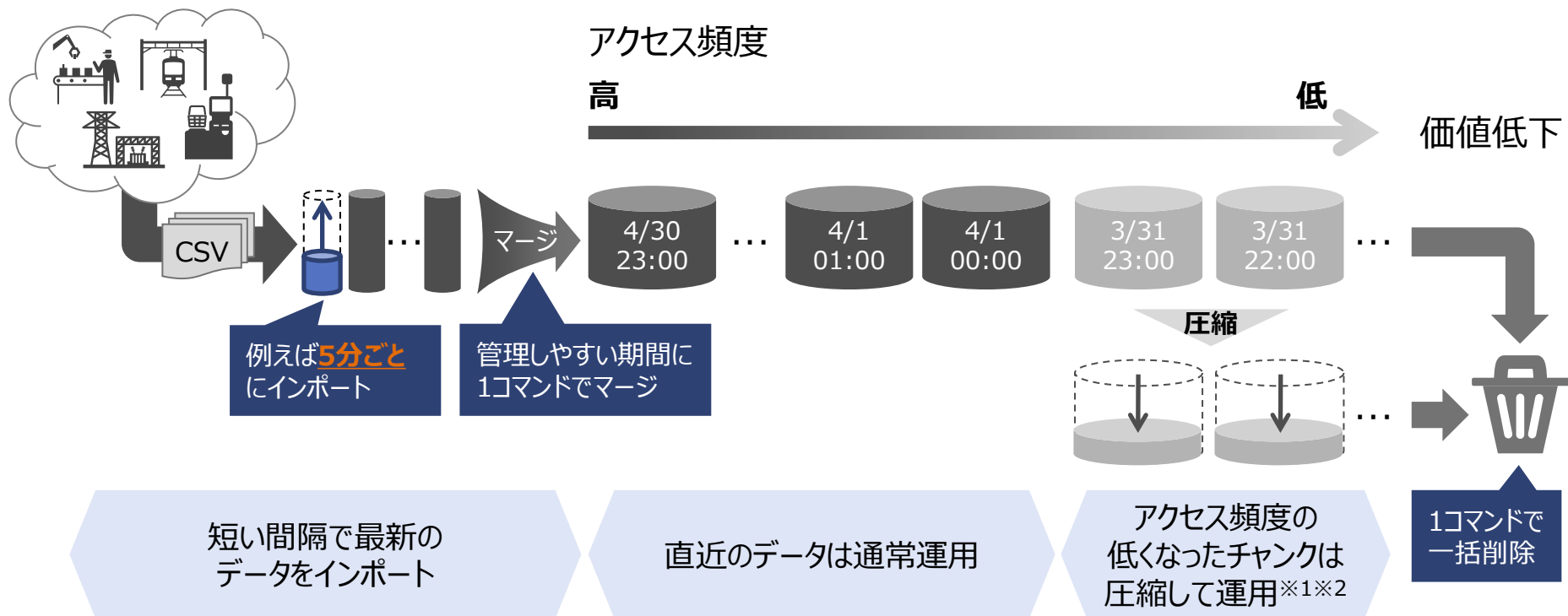
メリット1 : データインポートは、集計・分析に影響を与えないため、**最新データを短い間隔（例えば5分間隔など）でインポートし、集計・分析可能**

メリット2 : レンジ・パーティションのように**値の範囲で検索対象のチャンクを絞り、検索を高速化**



HADBでは、時系列データを独自の「チャンク」という固まりにして利用すると、**データライフサイクルを意識した効率的なデータ管理**が可能です。

- **短い間隔で最新のデータをインポート**し、管理しやすい期間に**コマンド1つでマージ可能**です。
- 不要になったデータのチャンクは、**コマンド1つで削除可能**です。
- 通常圧縮しないローストアでも、アクセス頻度の低くなったチャンクをコマンド1つで圧縮することで容量を削減できます。



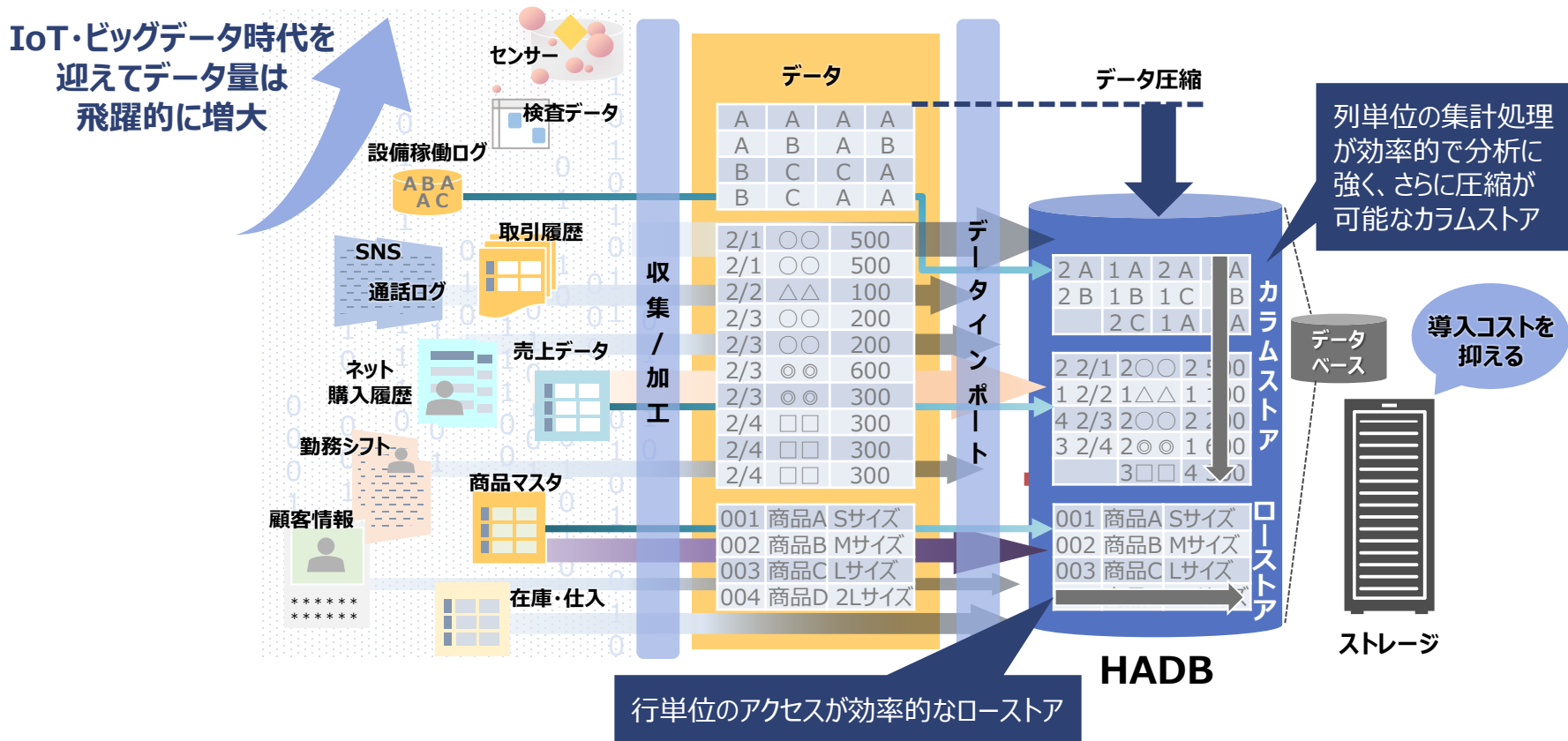
※1) 圧縮しても同じSQLでアクセス可能。

※2) カラムストアはデータ格納時に圧縮済み。

# 特長3 : 列単位の集計が得意なカラムストア (1/3)

データの価値を引き出すためには、さまざまなデータを扱う必要があります。HADBではデータの特성에合わせた最適な格納方式（カラムストア、ローストア）を選択することで効率的な処理を行えます。

- 行単位のアクセスが効率的なローストアだけでなく、**列単位の集計処理が効率的で分析処理に強く、またデータ圧縮も可能なカラムストアにも対応**しています。



ローストア/カラムストアを活用したデータインポートのイメージ

カラムストアとは、列ごとにデータをまとめて格納する方式です。

日付	商品ID	個数	顧客ID	売上
20190328	101	10	615	1,000
20190328	010	20	915	2,000
20190328	012	40	115	15,000
20190328	100	30	305	3,000
20190329	001	20	124	4,500

## ローストア

ローストアでは、1行分のデータがまとめてストレージに格納されます。そのため、行単位でのアクセスに適しています。

(例：顧客マスタ、商品マスタ)

20190328	101	10	615	1,000
20190328	010	20	915	2,000
20190328	012	40	115	15,000
20190328	100	30	305	3,000
20190329	001	20	124	4,500

## カラムストア

カラムストアでは、列ごとのデータがまとめてストレージに格納されます。そのため、列単位でのアクセスに適しています。

(例：設備稼働ログ、取引履歴)

20190328	101	10	615	1,000
20190328	010	20	915	2,000
20190328	012	40	115	15,000
20190328	100	30	305	3,000
20190329	001	20	124	4,500



## メリット1 : 分析処理で多い、列を絞った、大量の行に対する集計処理などが高速

日付	商品ID	個数	顧客ID	売上
20190328	101	10	615	1,000
20190328	010	20	915	2,000
20190328	012	40	115	15,000
20190328	100	30	305	3,000
20190329	001	20	124	4,500

- 例えば、顧客IDごとの売上合計、平均などの集計、最大値/最小値の計算など、列を絞った集計処理が高速です。

※ 商品マスタ、顧客マスタなど、行単位のアクセスが多いものは、従来通り、ロースタアへの配置を推奨します。

ロースタアでは全行全列のデータにストレージアクセスする必要あり。

カラムストアでは集計に必要な列のデータのみストレージアクセスで可能。

## メリット2 : データを圧縮することで、ストレージ容量の削減とデータアクセスの高速化を実現

日付		日付
20190328	4	20190328
20190328	1	20190328
20190328		
20190328		
20190329		

連続するデータをまとめて圧縮

- HADBでは、列ごとのデータ特性をいかして、カラムストアでの高圧縮を実現。これにより、ストレージを削減するだけでなく、ストレージアクセスの低減による処理の高速化を実現しています。

## メリット3 : メリット1、2の高速処理をインデクスのチューニング無しに実現可能

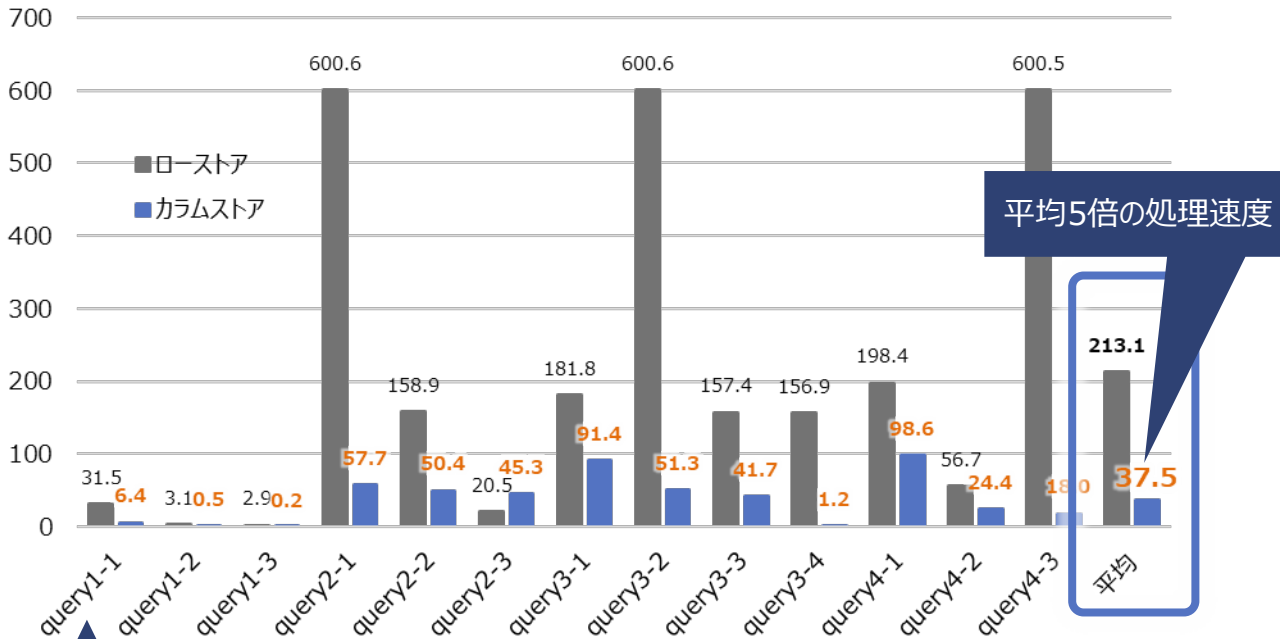
- ロースタアだと、列を絞って効率良くデータにアクセスし、集計処理を行うためには、インデクスのチューニングに非常に労力を要しますが、その労力無しに高速処理が実現できます。

# 【参考情報】カラムストアでのベンチマーク（SSB）測定結果

分析系・DWH※1系DBの一般的なベンチマークであるSSB※2を用いて、SSB13本のクエリについて、HADBのカラムストア環境、およびロースタ環境での実行時間をそれぞれ測定しました（SF※3=1000）。

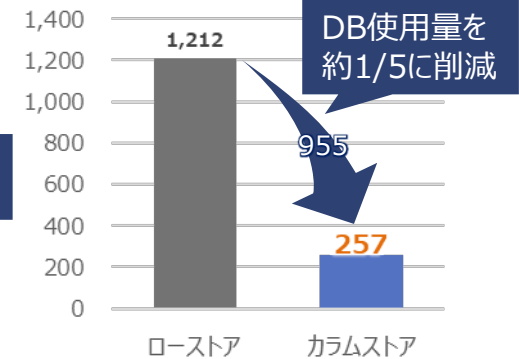
**カラムストアがロースタに対して、平均約5倍処理速度が速く、実行時間も安定しています。また、DB領域の使用量についても、カラムストアによって圧縮され、約1/5になっています。**

SSB各クエリ実行時間（秒）



ロースタ・カラムストア共に、レンジインデックスのみ付与

カラムストアによる圧縮効果  
DB領域の使用量（GB）



- 測定環境
- ①サーバー：HA8000 RS220AN2×1台  
CPU：Intel Xeon E5-2699v4  
(2.2GHz/22コア/44スレッド) × 2CPU  
Mem：192GB
  - ②ストレージ：VSP F400 × 1台  
FMD(HAF DC2) 6.4TB × 16台 (7D1P × 2)  
(オールフラッシュ構成)

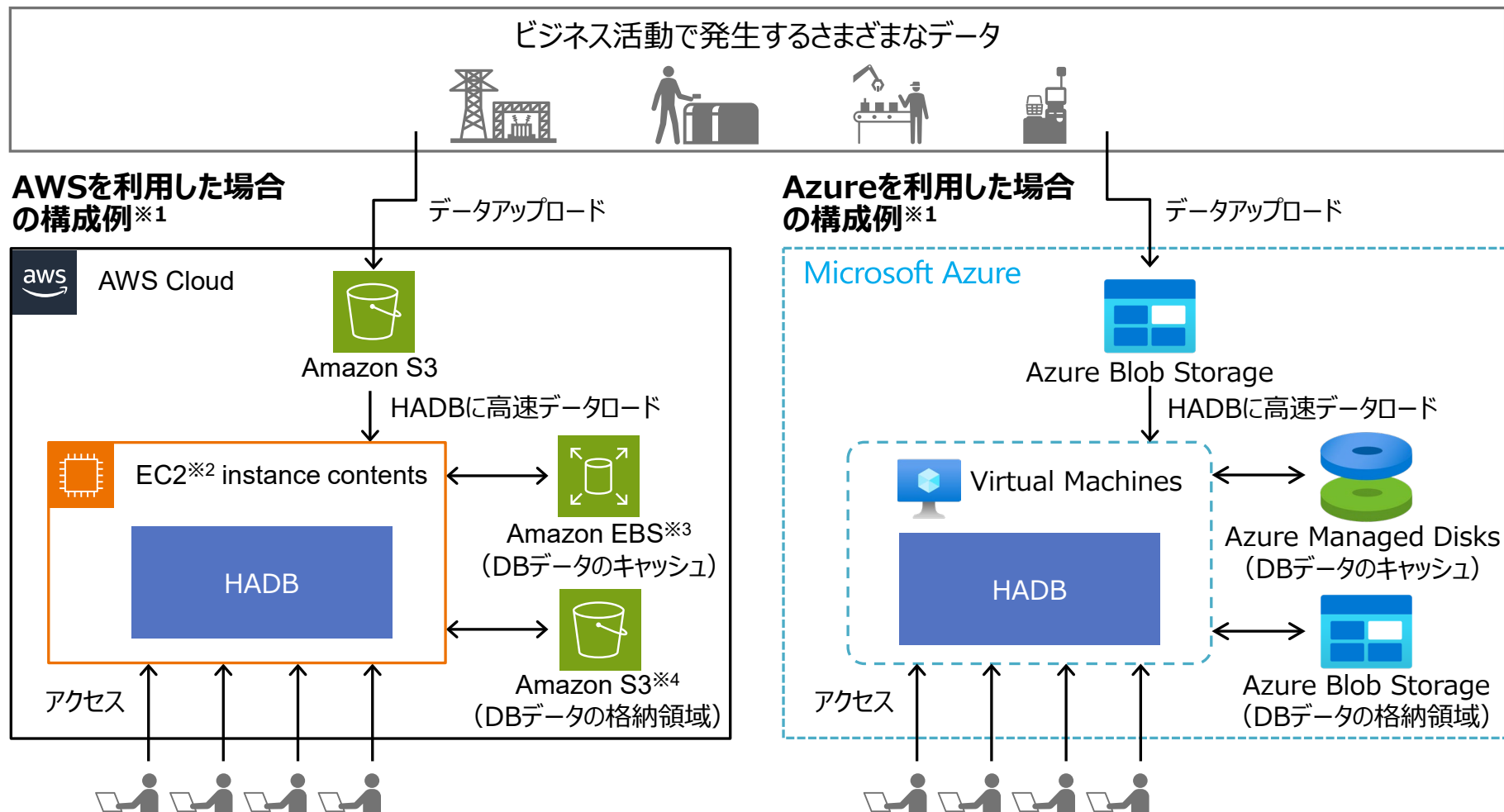
※1) DWH (Data Ware House)  
※2) SSB (Star Schema Benchmark)  
※3) SF (Scale Factor)

# Contents

---

1. 背景
2. HADBの性能面での特長
- 3. クラウド対応における特徴**
4. その他の特徴
5. ライセンス・サポート体系と価格
6. 導入事例

HADBは、オンプレミスだけでなくAmazon Web Services(以降、AWS)やMicrosoft Azure(以降、Azure)などのパブリッククラウド上でも稼働します。



※1) Amazon EBS (ブロックストレージサービス)、Azure Managed Disks (ブロックストレージサービス) をDBデータの格納領域として利用する構成もご提案可能です。

※2) EC2 (Amazon Elastic Compute Cloud)

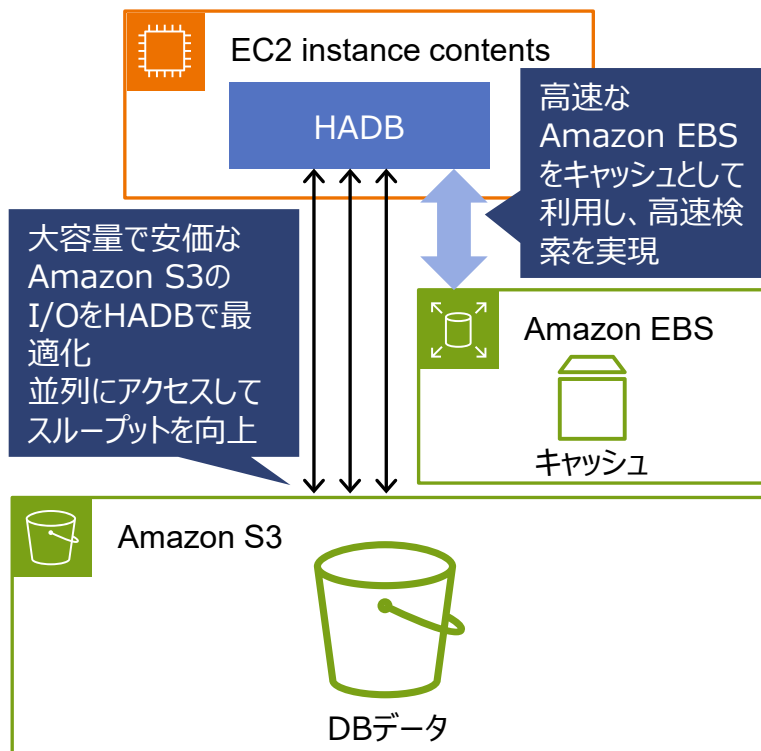
※3) Amazon S3 (Amazon Simple Storage Service)

※4) Amazon EBS (Amazon Elastic Block Store)

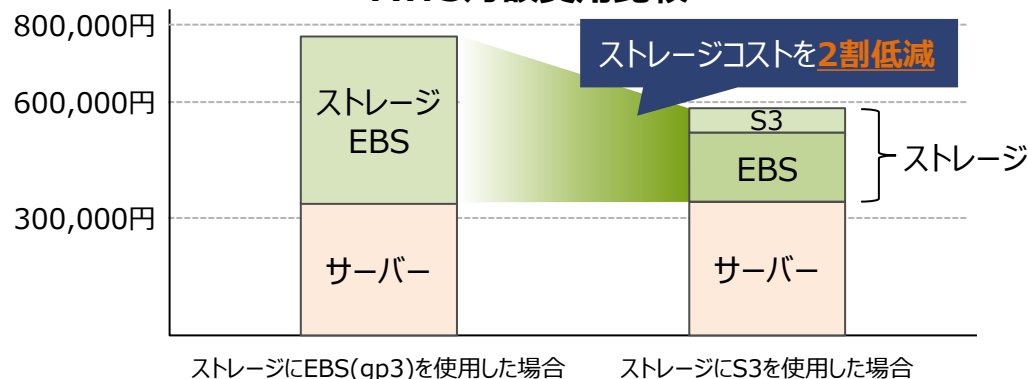
安価なオブジェクトストレージ（Amazon S3、Azure Blob Storage）にDBデータを格納することでストレージコストを低減できます。

また、高速なブロックストレージ（Amazon EBS、Azure Managed Disks）をキャッシュとして利用することで、オブジェクトストレージのコストメリットを享受しつつ大量データの高速検索が実現できます。

## オブジェクトストレージを活用した構成 (AWSを利用した場合の例)



## Amazon EBS (EBS) のみの構成と Amazon S3 (S3) を活用した構成の AWS月額費用比較※1



	ブロックストレージ (EBS) のみを ストレージとして活用した構成	オブジェクトストレージ (S3) を ストレージとして活用した構成
ストレージ	EBS(20TB + 12TB※2)	S3(20TB※3) + EBS(12TB※2 + 2TB※4)
サーバー	EC2 (r6i.12xlarge 48vCPU Mem384GiB)	EC2 (r6i.12xlarge 48vCPU Mem384GiB)

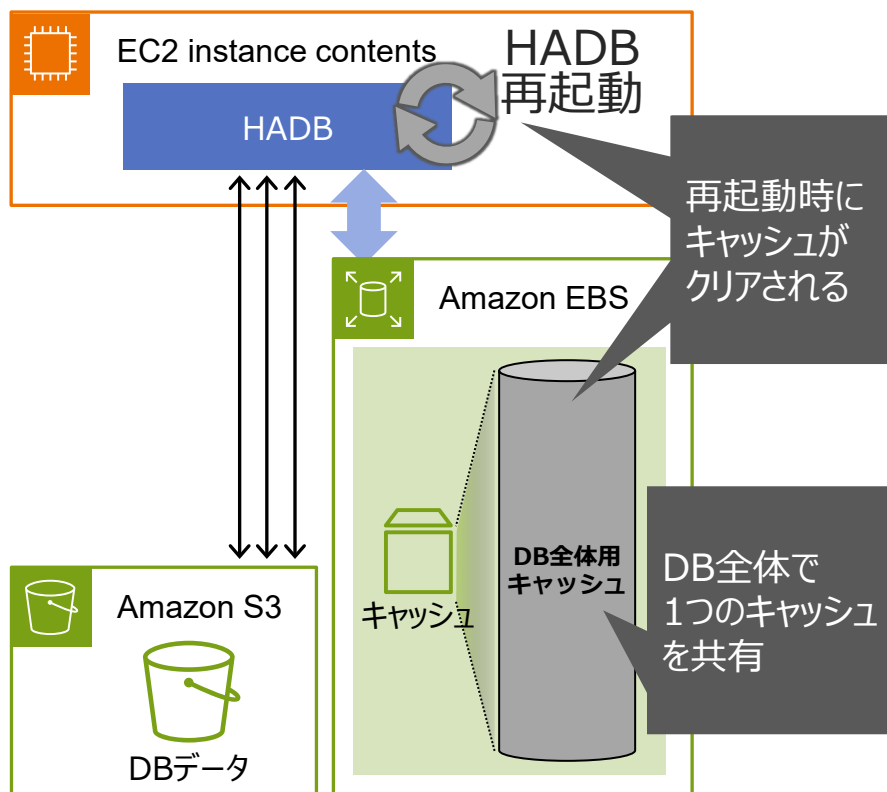
※1) 本資料で記載している費用は当社で試算した一例となります。システム構成などにより変動します。  
 ※2) HADBが使用する作業領域です。  
 ※3) S3の費用については、初回のデータロードで予定のDB容量 (20TB) を使い切ることを前提に計算しています。  
 (S3は使用量に比例した課金であるため、格納データが少しずつ増える場合はこの試算よりも安くなります)  
 ※4) キャッシュとして利用します。

# 【HADB 05-10 新機能】キャッシュ利用の最適化による オブジェクトストレージ活用時の検索性能安定性向上

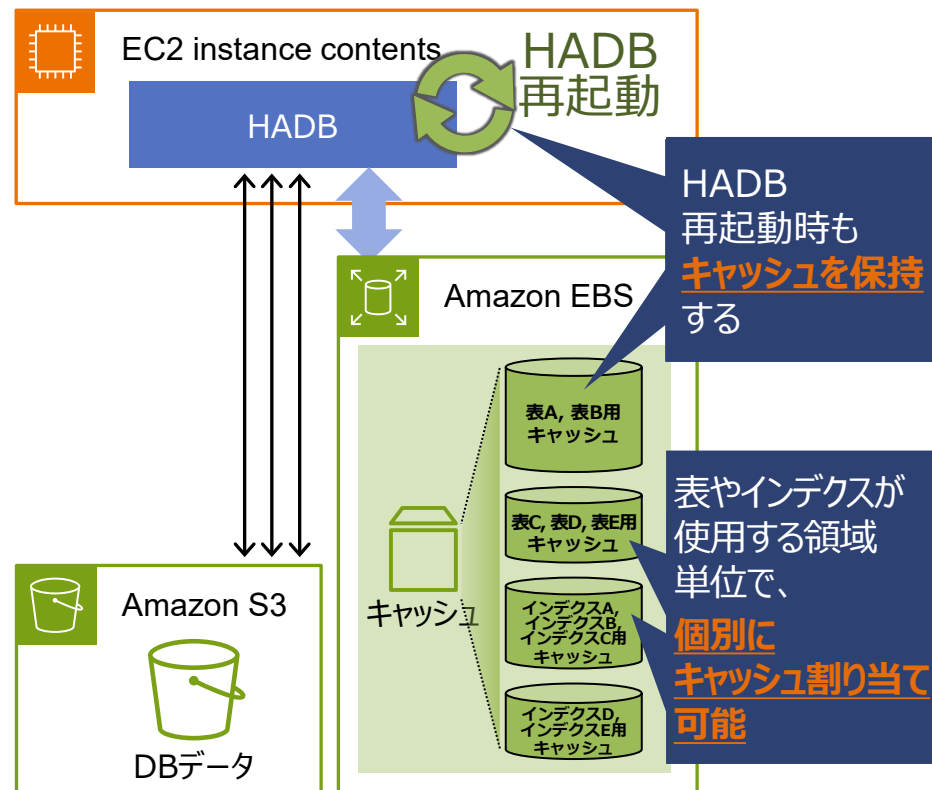
HADB 05-10では下記のエンハンスにより、検索処理でキャッシュがより活用されるようになるため、オブジェクトストレージ活用時に検索性能の安定性が更に向上します。

- 表やインデクスが使用する領域単位で**個別のキャッシュ割り当て**、**サイズ設定が可能**になりました。
- HADB**再起動時にキャッシュを保持**するようになりました。

## HADB 05-09以前

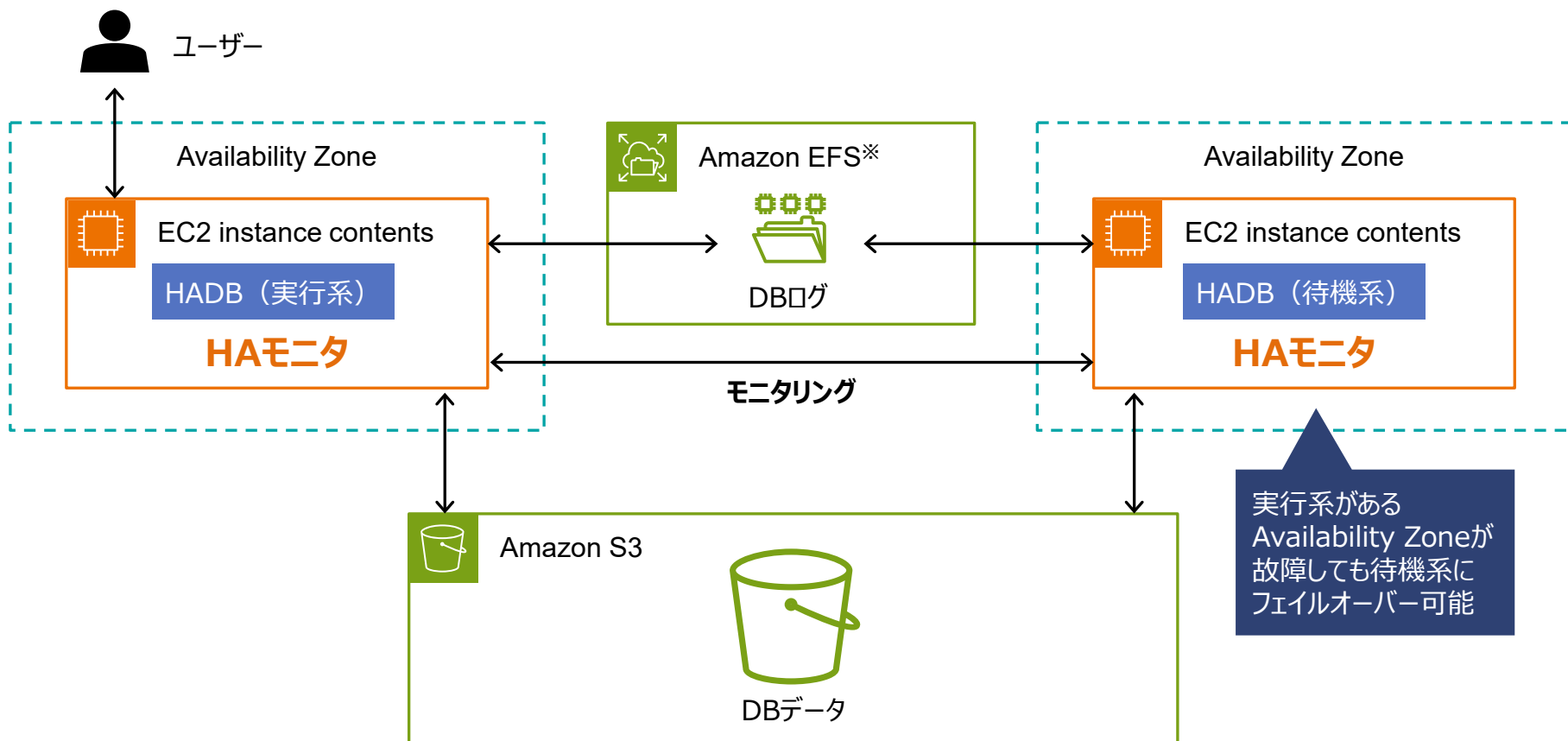


## HADB 05-10



HADBでは、クラスタソフトウェア製品HAモニタを利用することにより、オンプレミスはもちろん、AWSやAzureなどのパブリッククラウド上でもHA構成が可能です。

## AWSを利用する場合のHA構成（Azure版の構成も可能です。）

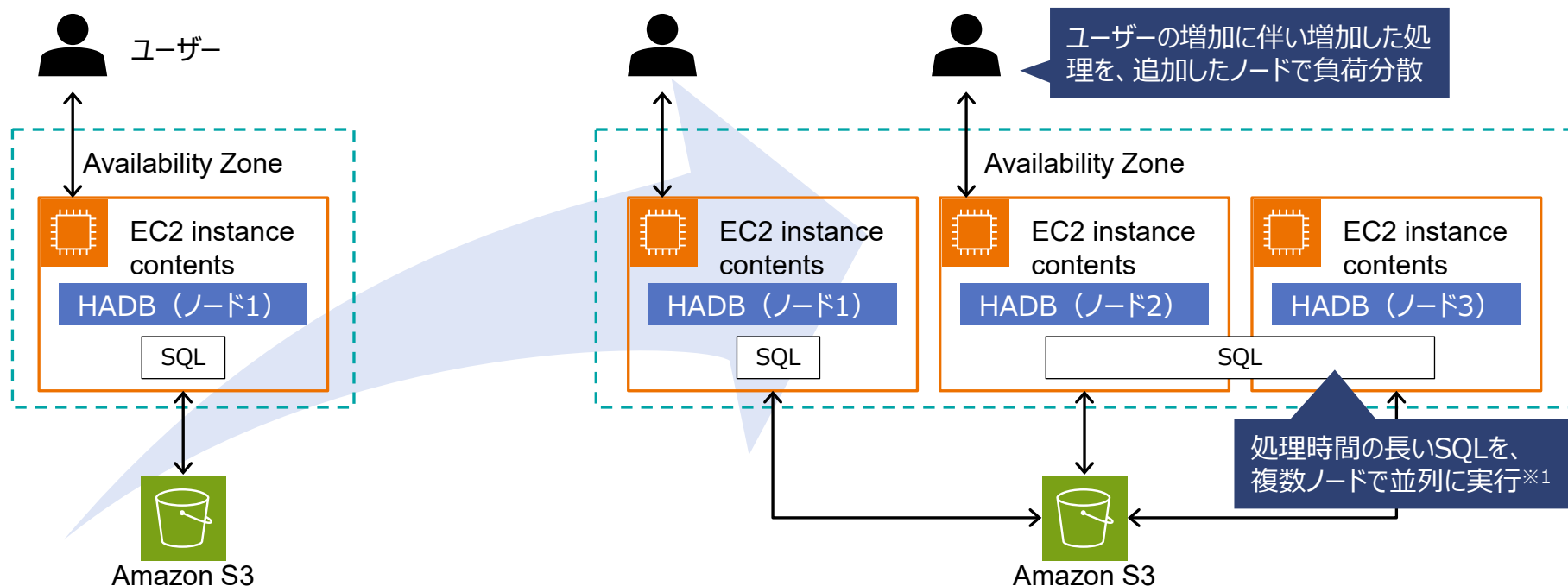


※) Amazon EFS (Amazon Elastic File System)

マルチノード構成とすることで可用性を向上し、ユーザー数やデータ処理の増加に対応し、スケールアウトできます。

- 利用ユーザー数の増加に応じてサーバーを追加していただくことで、複数ノードでユーザーの処理（SQL）を負荷分散し、多数のユーザーにも対応できるようになります。
- 処理時間の長い処理（SQL）を複数ノードで並列に実行し※1、スケールアウトすることができます。
- クラスタウェアを導入すれば、フェイルオーバー（系切り替え）をすることも可能です。

**AWSを利用する場合のマルチノード構成（Azureやオンプレミスでの構成も可能です。）**



※1) 並列実行するSQLには、グループ化を含む単表検索であることなどの制限があります。



AWSやAzureのパブリッククラウド上での構成について、以下の構成を含むシステム構成を実機検証し、コスト・性能面で最適な構成をベストプラクティス構成としてパターン化してまとめています。これらの構成により、高速検索に最適なシステム構成を提案可能です。

- オブジェクトストレージを活用する構成
- マルチAZでのHA構成
- マルチノード構成

### 1-2. HADB環境をS3を活用してAWS上に構築する場合の4つのポイント

HITACHI Inspire the Next

① EC2インスタンス ② EBSボリュームタイプ ③ ロードデータの格納先 ④ バックアップ方法を考慮し、HADBの特性を最大限に引き出します。

**ポイント① EC2インスタンス**  
CPU性能・ネットワーク帯域、I/O帯域のパランスが良いR6iインスタンスを使用し、HADBの検索性能を引き出す  
ポイント①、②でEC2-S3間とEC2-EBS間のスループットを最大化！

**ポイント② EBSボリュームタイプ**  
DBデータのキャッシュとして使用するEBSは、安価なgp3ボリュームを8本準備することで、EC2側のEBS帯域を使い切る

**ポイント③ ロードデータの格納先**  
S3バケットに格納したデータをHADBにデータロードする  
EBSより安価なS3にDBデータを格納

**ポイント④ バックアップ方法**  
DBの整合性を保つため、DBディレクトリとS3バケットの2箇所でのバックアップを取得する

© Hitachi, Ltd. 2023. All rights reserved. 7

### ポイント② EBSボリュームタイプ 1-4. EBSボリュームタイプの選定は？ (2/2)

HITACHI Inspire the Next

EC2インスタンスのI/O性能をLVMで束ねたgp3ボリュームで使い切れるよう、gp3ボリュームの数と各ボリュームのスループット、IOPSを設定します。  
gp3ボリュームの数、I/O設定を決定するポイント

- ✓ ボリューム数は、8個をお勧めします。8個を束ねることで、EC2側のI/O性能を使い切れます。また、将来的なスケールアップにも十分に対応できます。
- ✓ ボリューム数×単一ボリュームのI/O性能がEC2インスタンスのI/O性能に達するように単一ボリュームのスループットとIOPSを設定します。

表2: EBSボリュームを8個併用した場合のI/O性能設定例

EC2インスタンス (r6iクラス)	EBS標準化			単一ボリューム (gp3ボリュームタイプ)		合計スループット (MB/s)	合計IOPS (16KB I/O)
	インスタンスサイズ	vCPU	メモリ (GB)	ベースラインスループット (16KB I/O) **	ベースラインIOPS (16KB I/O)		
r6i.xlarge	4	32	156,25	6,000	8	125 **	3,000 **
r6i.2xlarge	8	64	312,5	12,000	8	125 **	3,000 **
r6i.4xlarge	16	128	625	20,000	8	125 **	3,000 **
r6i.8xlarge	32	256	1,250	40,000	8	160	5,000
r6i.12xlarge	48	384	1,875	60,000	8	240	7,500
r6i.16xlarge	64	512	2,500	80,000	8	320	10,000
r6i.24xlarge	96	768	3,125	120,000	8	400	15,000

① EBS標準化 (Mbps)を128 KiB I/Oを基準として (MB/s)単位で書き換えた場合の表記。  
② 無償設定範囲内の最大値を指す。合計したI/O性能がEC2のI/O性能を使い切れるように、各EBSボリュームの性能を設定

© Hitachi, Ltd. 2023. All rights reserved. 12

### 3-3. 検証1: S3活用によるHADB管理性能 (結果)

HITACHI Inspire the Next

DBデータのキャッシュにヒットする場合は、EBSのみの構成と同等程度の処理時間であることを確認しました。

1日目の実行 (HADB管理後のDBSQL実行)では、アクセスするデータがキャッシュにヒットし、DBデータのキャッシュとしてEBSにヒットするよりも、EBSの構成に比べて、処理時間-10%程度の短縮が確認されました。性能面での優位性が期待できます。

2日目のDBSQL実行では、アクセスするデータがEBSのDBデータキャッシュにヒットするよりも、実行時間-15%程度の短縮が確認されました。

④ EC2側のDBデータのキャッシュにヒットしない場合は、処理時間への影響が確認されました。

⑤ DBデータのキャッシュにヒットする場合は、処理時間への影響が確認されました。

© Hitachi, Ltd. 2023. All rights reserved. 11

### 3-5. 検証2: スケールアップ時のHADB管理性能 (結果)

HITACHI Inspire the Next

EC2インスタンスのスケールアップとEBSボリュームの構成による処理性能の向上を確認しました。

EC2インスタンスのスケールアップ、I/O性能を向上させるEBSボリューム構成による、スケールアップ時の処理性能の向上を確認しました。

① CPU負荷が高い場合 : 4m02 (100%実行)  
② I/Oとネットワークの負荷が高い場合 : 4m06 (100%実行)

© Hitachi, Ltd. 2023. All rights reserved. 15

### 2-1. HADB on AWSベストプラ構成の具体例 (1/2)

HITACHI Inspire the Next

ここまでの4つのポイントを押さえたHADBの構成例です (DBデータ量は20TBを想定)。

① EC2インスタンス  
vCPU: 48  
メモリ: 384GB  
EBS標準化: 15,000Mbps  
EBS標準化 (EBS標準化)  
ベースラインスループット: 1,875MB/s (128KB I/O)  
ベースラインIOPS: 60,000 (16KB I/O)

② Amazon EBS  
DBデータのキャッシュ領域の選定 (DBデータ量の10%程度)  
DBデータ量: 20TB

③ EBSおよびHADBのインストール環境  
OS: RHEL 8.6  
DBデータ量: 2,000GB  
ロードデータ: 1,800GB  
DBデータ量: 6,000GB  
DBデータ量: 10GB  
DBデータ量: 1,600GB

④ HADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリュームをLVMで束ねてI/O性能を確保  
使用量: 20TB

⑤ HADBが使用するVGI (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 14,000GB (LVM)

⑥ HADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑦ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑧ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑨ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑩ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑪ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑫ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑬ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑭ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑮ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑯ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑰ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑱ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑲ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

⑳ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉑ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉒ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉓ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉔ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉕ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉖ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉗ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉘ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉙ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉚ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉛ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉜ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉝ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉞ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㉟ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊱ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊲ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊳ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊴ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊵ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊶ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊷ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊸ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊹ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊺ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊻ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊼ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊽ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊾ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

㊿ S3がHADBが使用するDBデータ (Amazon S3)  
EBSのgp3ボリューム (gp3ボリューム) × 8  
使用量: 20TB

© Hitachi, Ltd. 2023. All rights reserved. 16

### 2-2. DBサイズごとのベストプラ構成例

HITACHI Inspire the Next

DBサイズに応じて、EC2インスタンスタイプやgp3ボリュームの設定を変更してください。

DBサイズごとのHADBベストプラ構成例を次の表に示します。  
DBサイズが20TBの場合の構成例は、「2-1. HADB on AWSベストプラ構成の具体例」に記載しています。

表3: DBサイズごとのHADB on AWSベストプラ構成例

DBデータ量 (TB)	EC2インスタンスの構成			HADBが使用するVGI (EBSのgp3ボリュームをLVMで束ねる)		HADBが使用するDBデータ (Amazon S3)	
	インスタンスタイプ	vCPU	メモリ (GB)	ベースラインスループット (16KB I/O)	ベースラインIOPS (16KB I/O)	スループット (MB/s)	IOPS (16KB I/O)
1	r6i.xlarge	4	32	156,25	6,000	8 × 100	300
2	r6i.2xlarge	8	64	312,5	12,000	8 × 200	600
3	r6i.4xlarge	16	128	625	20,000	8 × 400	1,200
4	r6i.8xlarge	32	256	1,250	40,000	8 × 800	2,400
5	r6i.12xlarge	48	384	1,875	60,000	8 × 1,200	3,600
6	r6i.16xlarge	64	512	2,500	80,000	8 × 1,600	4,800
7	r6i.24xlarge	96	768	3,125	120,000	8 × 2,400	7,200

① EC2インスタンスの構成  
② HADBが使用するVGI (EBSのgp3ボリュームをLVMで束ねる)  
③ HADBが使用するDBデータ (Amazon S3)

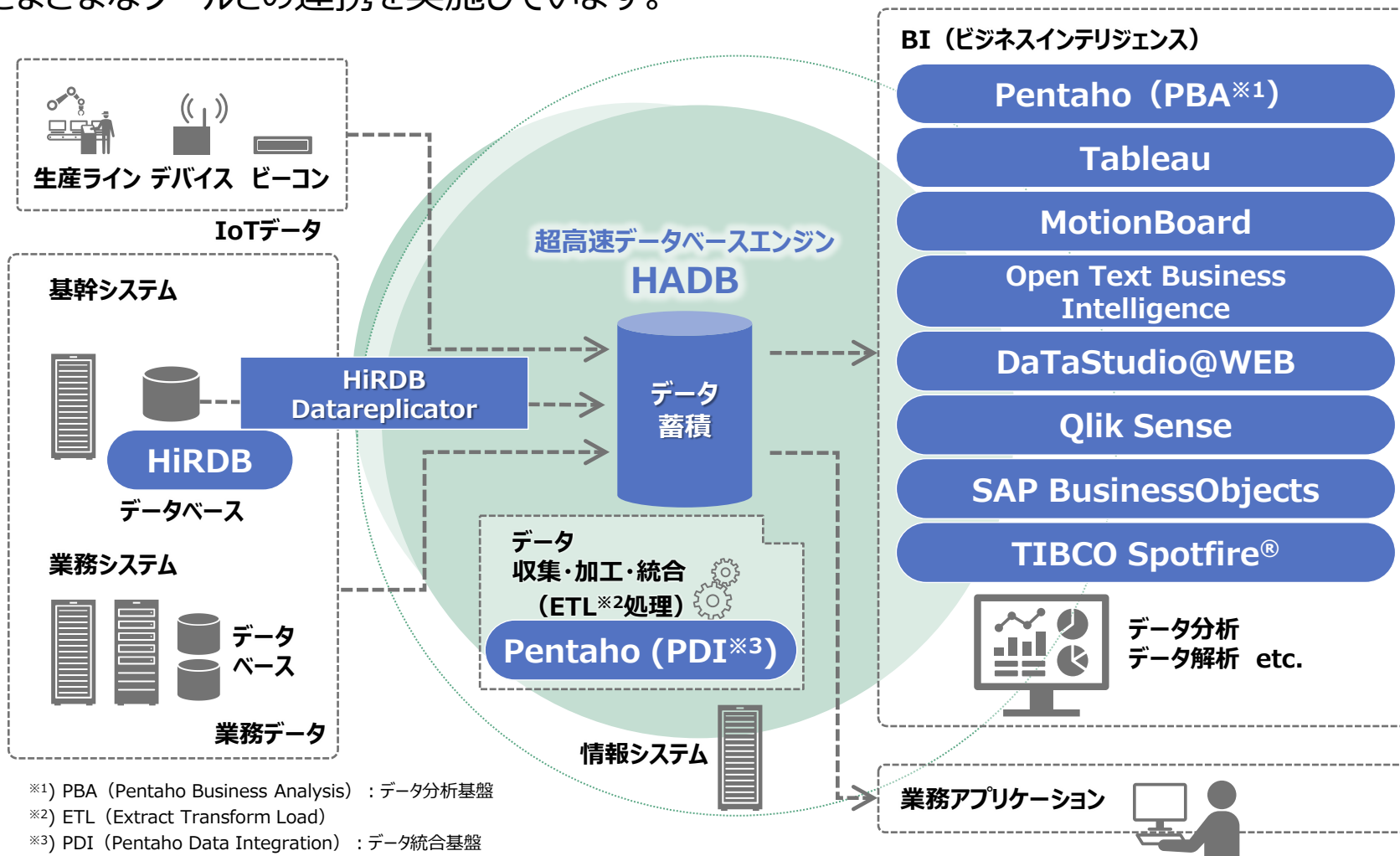
© Hitachi, Ltd. 2023. All rights reserved. 18

# Contents

---

1. 背景
2. HADBの性能面での特長
3. クラウド対応における特徴
- 4. その他特徴**
5. ライセンス・サポート体系と価格
6. 導入事例

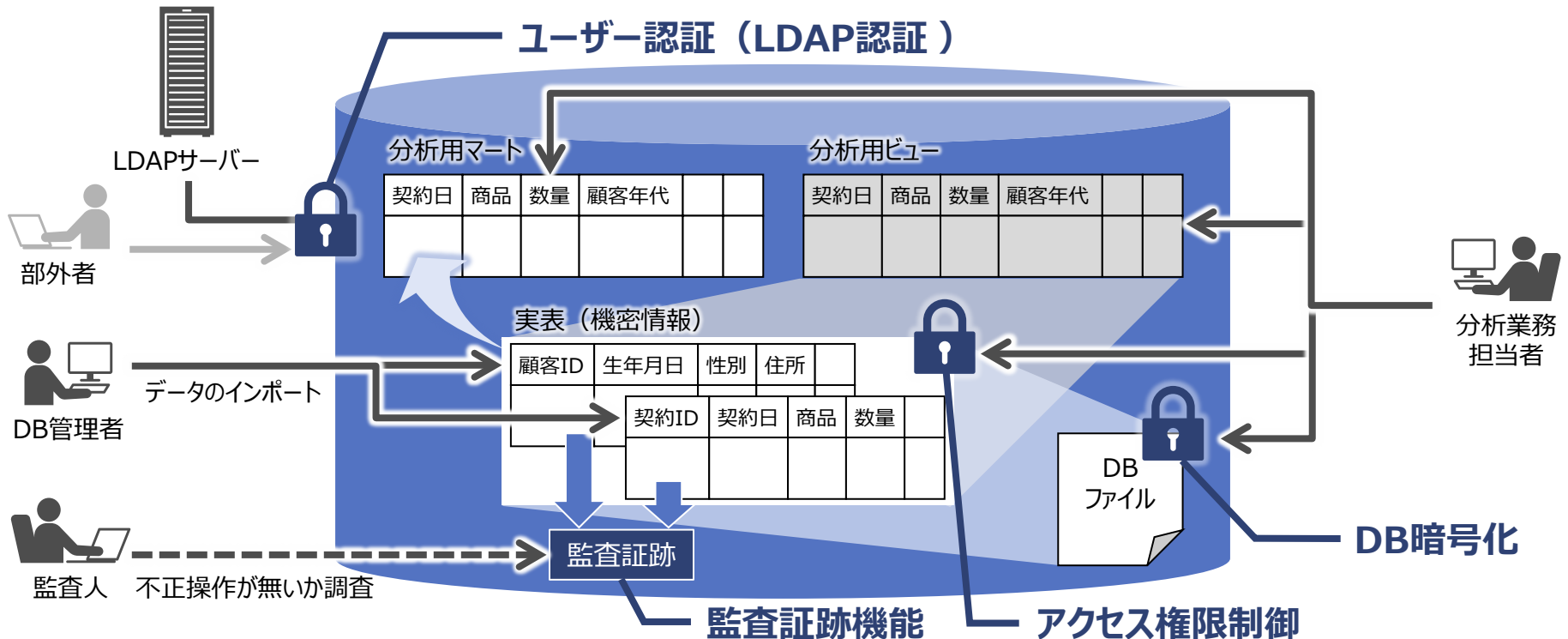
さまざまなツールとの連携を実施しています。



これら以外のツールについては、ご相談ください

HADBでは、超高速検索だけでなく、セキュリティ機能も標準で備えており、安心してお使いいただけます。

- ユーザー認証 : パスワード認証を行います。LDAP認証を利用することも可能です。
- アクセス権限制御 : 各ユーザーにアクセス権限を設定することで、表やビュー表へのアクセス範囲を制限できます。
- 監査証跡機能 : DBへの接続・操作などの証跡を取得し、不正な操作が無いか、チェックができます。
- DB暗号化 : DBファイルを直接参照して機密情報にアクセスすることが防げます。



国産DBだからこそできる開発者による直接的なサポートが受けられます。お客さまからの質問や障害に迅速に対応いたします。

また、バージョンアップ時は上位互換性を保持しているため、アプリケーションの書き換えなしでそのまま最新バージョンをお使いいただけます。

**標準で10年に渡る長期間のサポートを毎年定額で提供します。**

## 開発者直結

### 国産DBだからこそできる開発者によるサポート

- 開発者に直結した**迅速かつ的確なサポートを提供**。

## 上位互換性

### APおよびDBの上位互換性を保証

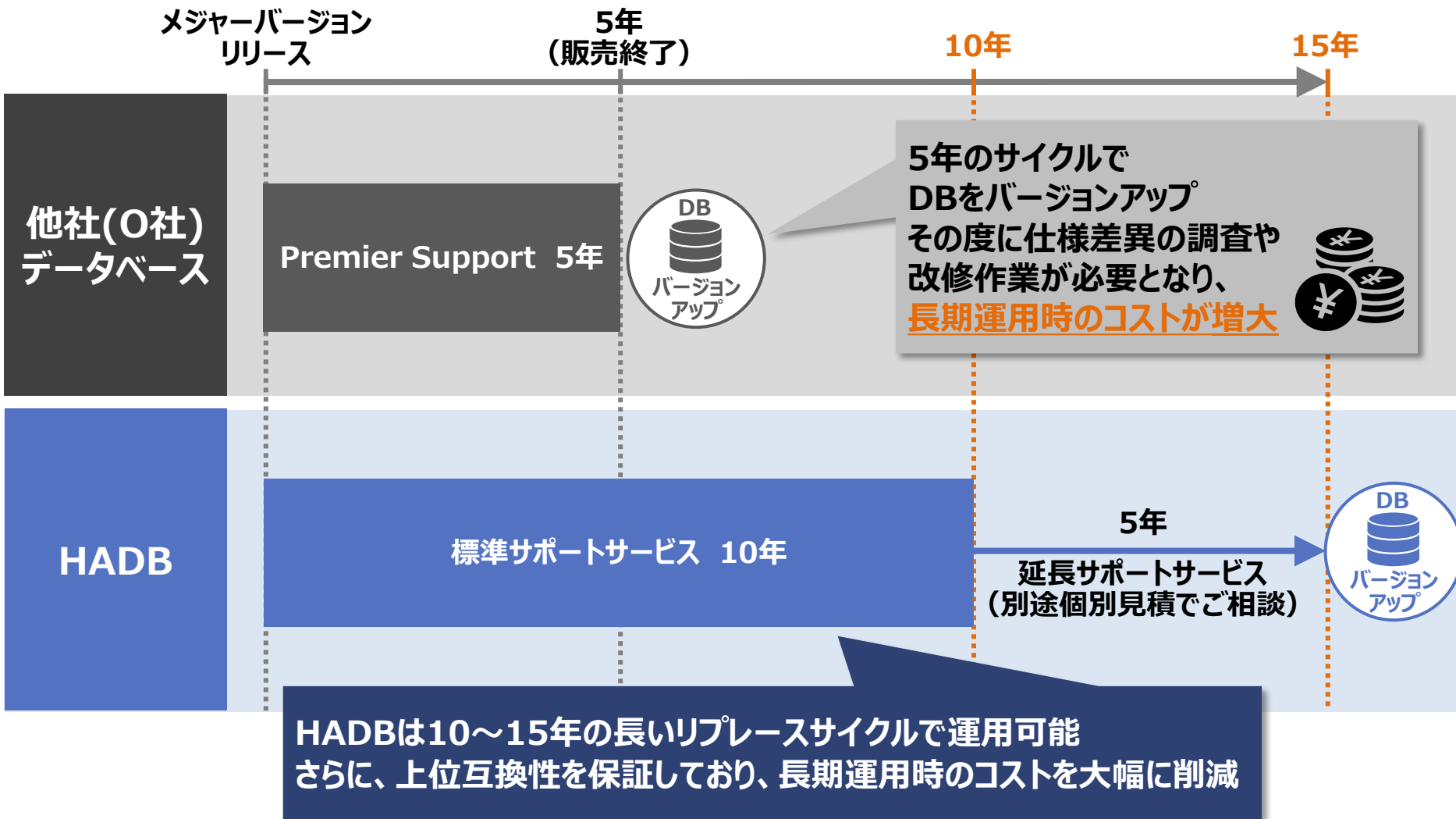
- バージョンアップ時に、アプリケーションプログラム（AP）の書き換え不要。
- バージョンアップしてもデータベースはそのまま使用可能。

## 改良版・ 障害時対応

### 発生した問題を解決するための徹底した対応

- 原則として、**ご使用のバージョンの改良版を提供**。
- ソースを保持し開発者が近くにいるため、迅速な対応が可能。
- 日立サポート360によりワンストップでハードウェアからソフトウェアまでトータルにサポート。

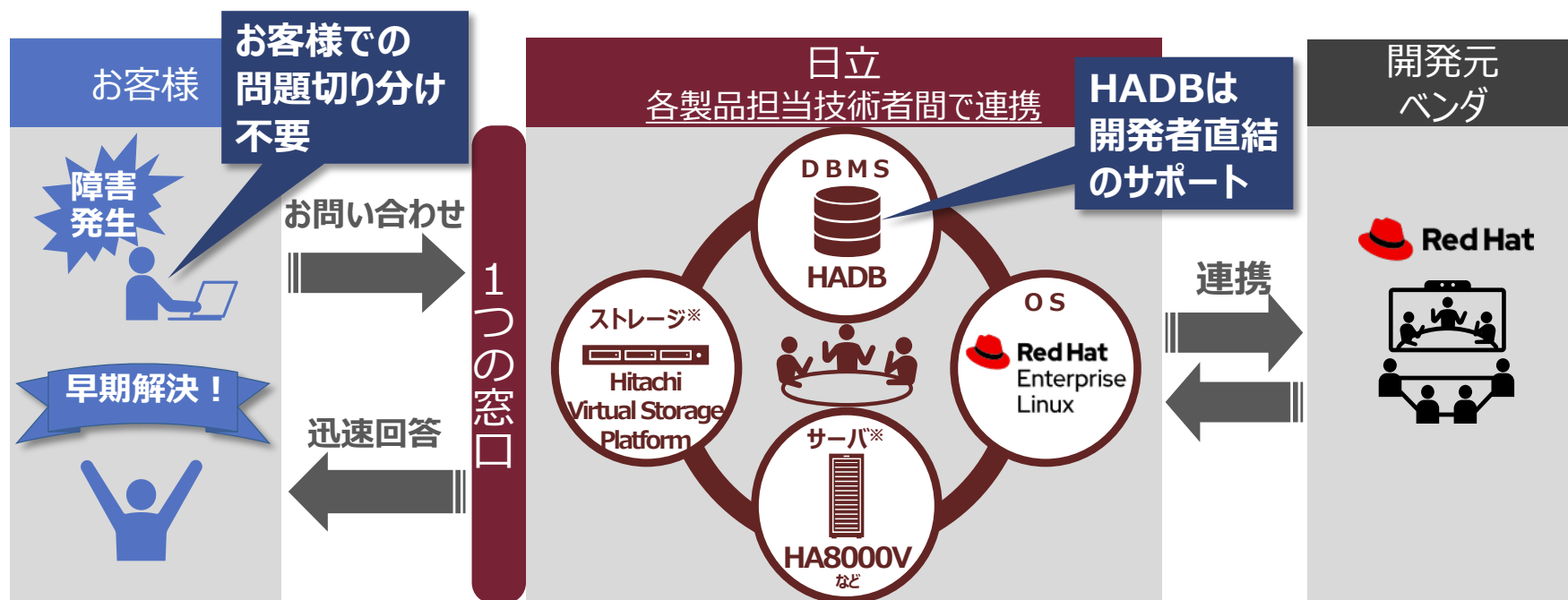
HADBは、標準で10年、最大15年に渡る長期間のサポートを毎年定額で提供します。



# HADBのサポート（日立サポート360）の特徴

HADBは**開発者直結のサポート**を提供します。  
そのため、**障害発生時、開発元へのエスカレーションに時間が掛かりません。**  
また、日立サポート360は以下の特徴があり、HADBのみならずお客様のシステム全体を  
ワンストップでサポート可能です。

- 日立の総合力によりシステム全体でサポートを提供します。
- 障害発生時もお客さまは**問題を切り分けることなく**お問い合わせいただけます。
- 日立が提供するシステム内の問題をアプライアンス製品のように**1つの窓口でワンストップ対応**します。
- 原因究明が難しい問題も各製品技術者、導入元ベンダと密に連携し、問題解決にあたります。

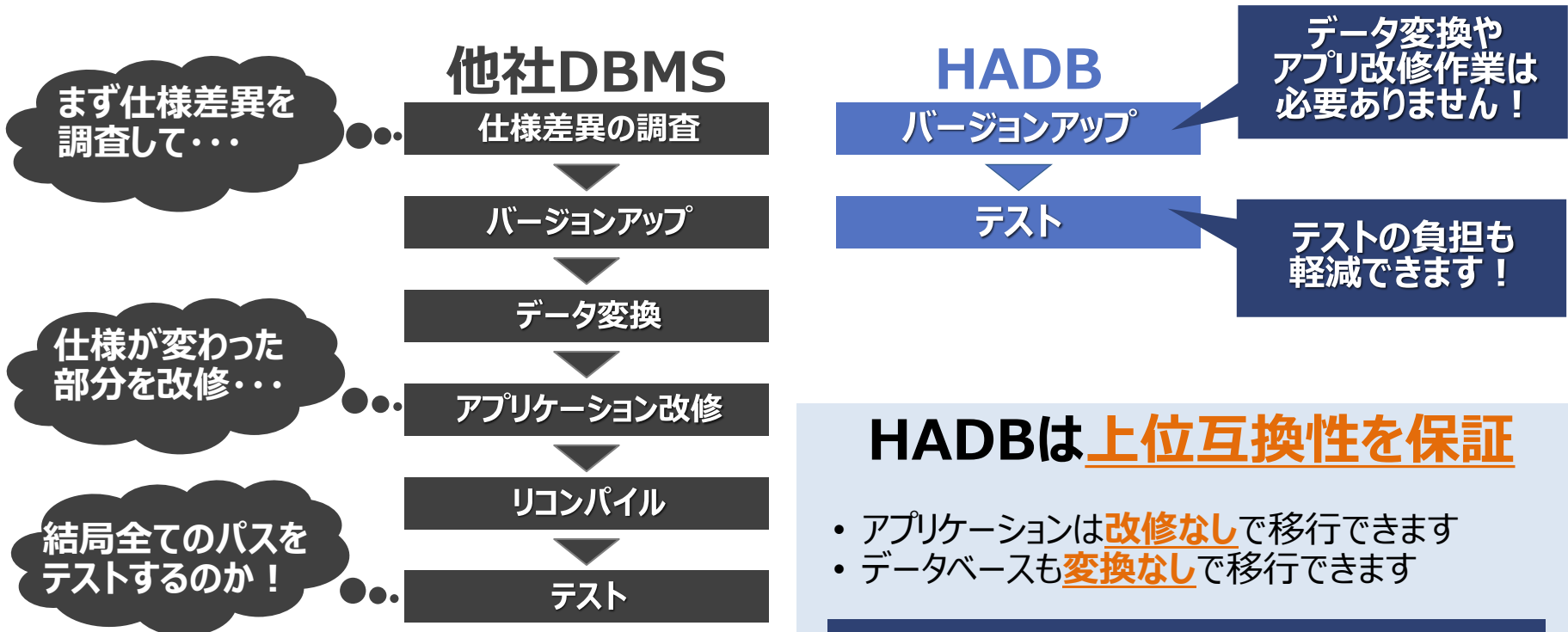


※) ストレージ、およびサーバの機種名は一例。

【参考URL】：日立サポート360 <https://www.hitachi.co.jp/soft/support360/>

HADBは上位互換性を保持しているため、バージョンアップ時に**アプリケーションの書き換えなし**でそのまま**最新バージョン**をお使いいただけます。

## ■ DBのバージョンアップに伴う作業



## HADBは**上位互換性を保証**

- アプリケーションは**改修なし**で移行できます
- データベースも**変換なし**で移行できます

バージョンアップに掛かる  
移行コストを最小限に抑えます。

バージョンごとに仕様が異なるので、仕様差異の調査、アプリケーションの改修、データベースの変換、これらに伴うテストが必要...



# Contents

---

1. 背景
2. HADBの性能面での特長
3. クラウド対応における特徴
4. その他特徴
- 5. ライセンス・サポート体系と価格**
6. 導入事例

HADBには、3種類のライセンス体系があります。

## ■ HADBのライセンス・サポート体系

- 売り切りライセンス
- サブスクリプションライセンス
- 従量課金サービス（オンデマンド・ミドルウェアサービス）

サブスクリプションライセンスと従量課金サービスには、サポートサービスも含まれます。  
従量課金サービスは AWS/Azureのみでご提供可能です。



データ量に合わせて容量ライセンスを追加できる売り切りライセンスなら、データ量の増加に合わせたスモールスタートが可能です。

## ■ 売り切りライセンス※1

インストールライセンス + 容量ライセンスの合算

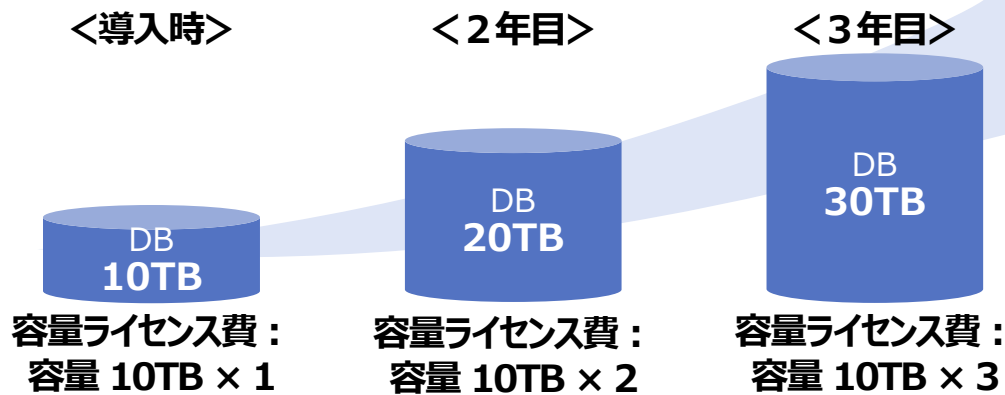
種別	種類	説明	ライセンス数	標準価格（税別）
インストールライセンス	基本ライセンス	1サーバー（コア数制限無）	1（サーバー）	7,700,000円
	追加ライセンス	クラスタ構成時の追加サーバー1台ごとに必要	1（サーバー）	7,700,000円
容量ライセンス※2	容量 10TB	容量10TB	10（TB）	3,300,000円
	容量 100TB	容量100TB	100（TB）	22,000,000円

※1) サポートサービスをご利用の場合、次ページのサポートサービス料金が別途必要です。

※2) ユーザーデータを格納するDBエリアの容量がライセンス対象になります。

## ■ 例) 導入時DB容量10TB、毎年10TBずつ増加するシステム

導入時に最大容量に合わせた容量ライセンスを購入するのではなく、データ量の増加に合わせて必要なライセンスを追加購入することができるため、コストを最適化できます。



HADBでは、売り切りライセンス※において、インストールライセンス分と容量ライセンス分のサポートサービス料金により、サポートサービスを利用いただけます。  
サポートサービスは提供時間により、平日8時-19時のサポートサービスと、24時間週7日サポートサービスの2種類があります。

## (1) 平日8時-19時サポートサービス

＜サポートサービス料金＞

種別	種類	説明	ライセンス数	標準価格（税別）
インストールライセンス	基本ライセンス	1サーバー（コア数制限無）	1（サーバー）	1,155,600円
	追加ライセンス	クラスタ構成時の追加サーバー1台ごとに必要	1（サーバー）	1,155,600円
容量ライセンス※2	容量 10TB	容量10TB	10（TB）	495,600円
	容量 100TB	容量100TB	100（TB）	3,300,000円

## (2) 24時間週7日サポートサービス

＜サポートサービス料金＞

種別	種類	説明	ライセンス数	標準価格（税別）
インストールライセンス	基本ライセンス	1サーバー（コア数制限無）	1（サーバー）	1,618,800円
	追加ライセンス	クラスタ構成時の追加サーバー1台ごとに必要	1（サーバー）	1,618,800円
容量ライセンス※2	容量 10TB	容量10TB	10（TB）	694,800円
	容量 100TB	容量100TB	100（TB）	4,620,000円

※) サブスクリプションライセンスおよび従量課金サービス(オンデマンド・ミドルウェアサービス)によるライセンスの場合、サポートサービスが付帯したライセンス体系になります。

サブスクリプションライセンスなら、オンプレミスや任意のクラウドサービス上で、年契約で使用している間だけの定額でのお支払いにより、HADBをサポートサービス付帯でご利用いただけます。ライセンス体系はサポートサービスの提供時間により、平日8時-19時のサポートサービスと、24時間週7日サポートサービスから選択可能です。

## ■サブスクリプションライセンス(平日8時-19時サポートサービス付帯)

年契約、インストールライセンス+容量ライセンスの合算

種別	種類	説明	ライセンス数	標準価格 (年額)
インストールライセンス	インストールライセンス	1サーバー (コア数制限無)	1 (サーバー)	2,439,600円
容量ライセンス※1	容量 10TB	容量10TB	10 (TB)	1,046,400円
	容量 100TB	容量100TB	100 (TB)	6,967,200円

※1) ユーザーデータを格納するDBエリアの容量がライセンス対象になります。

## ■サブスクリプションライセンス(24時間週7日サポートサービス付帯)

年契約、インストールライセンス+容量ライセンスの合算

種別	種類	説明	ライセンス数	標準価格 (年額)
インストールライセンス	インストールライセンス	1サーバー (コア数制限無)	1 (サーバー)	2,902,800円
容量ライセンス※1	容量 10TB	容量10TB	10 (TB)	1,245,600円
	容量 100TB	容量100TB	100 (TB)	8,287,200円

※1) ユーザーデータを格納するDBエリアの容量がライセンス対象になります。

HADBは、パブリッククラウド上※1で必要なときに必要な間だけ、日立ミドルウェアとそのサポートサービスをセットにして従量課金で提供するオンデマンド・ミドルウェアサービスにも対応しています。

## ■ 従量課金サービス※2（オンデマンド・ミドルウェアサービス for AWS）

1時間あたりの従量課金サービス

種別	種類	説明	ライセンス数	標準価格（税別）
基本サービス	基本サービス	オンデマンド・ミドルウェアサービスの基本料金	—	2,200円/月
HADB 従量課金サービス	基本ライセンス	1プロセッサ（2vCPU）	1（プロセッサ）	85円/時間
	追加ライセンス	追加1プロセッサ（2vCPU）	1（プロセッサ）	37円/時間

## ■ 従量課金サービス※2（オンデマンド・ミドルウェアサービス for Microsoft Azure）

1分あたりの従量課金サービス

種別	種類	説明	ライセンス数	標準価格（税別）
基本サービス	基本サービス	オンデマンド・ミドルウェアサービスの基本料金	—	2,200円/月
HADB 従量課金サービス	基本ライセンス	1プロセッサ（2vCPU）	1（プロセッサ）	1.416円/分
	追加ライセンス	追加1プロセッサ（2vCPU）	1（プロセッサ）	0.616円/分

オンデマンド・ミドルウェアサービスでは、日立ミドルウェア製品をお好きな期間、使った分の料金でご利用いただけます

オンデマンド・ミドルウェアサービスURL：<https://www.hitachi.co.jp/Prod/comp/soft1/odm/>

※1) AWS、Azureにのみ対応しています(2024年7月現在)。

※2) ご利用になるパブリッククラウドの料金が別途必要になります。

このページに記載されているライセンス費は、2024年7月現在の標準価格（税別）です。

# Contents

---

1. 背景
2. HADBの性能面での特長
3. クラウド対応における特徴
4. その他特徴
5. ライセンス・サポート体系と価格
- 6. 導入事例**

## 【医療】医療系ビッグデータ分析

### 顧客課題

蓄積された6年分の全国民のレセプトデータ（約2,000億レコード）全件の活用のために、さまざまな視点で繰り返し解析を進める必要があるが、解析時間の制約で十分な分析ができない。

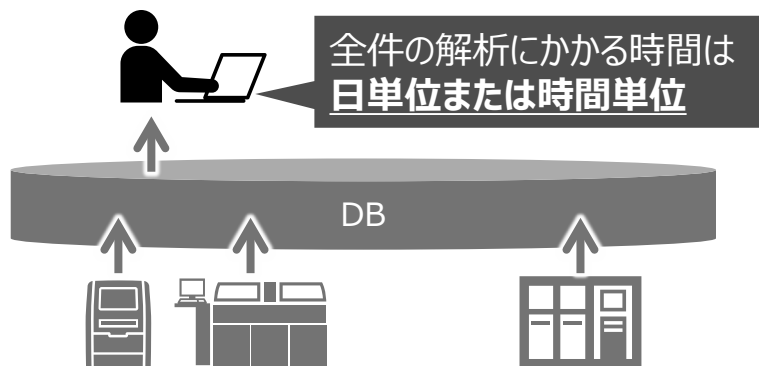
全国民のレセプト情報・特定健診等情報（以下、**NDB**<sup>※1</sup>データ）や自治体が保有する健診・医療・介護データといった医療系ビッグデータの活用のため、さまざまな視点で繰り返し研究を進める必要があるが、**6年分の全国民のレセプトデータ（約2,000億レコード）全件の解析にかかる時間は、日単位または時間単位**となっていた。



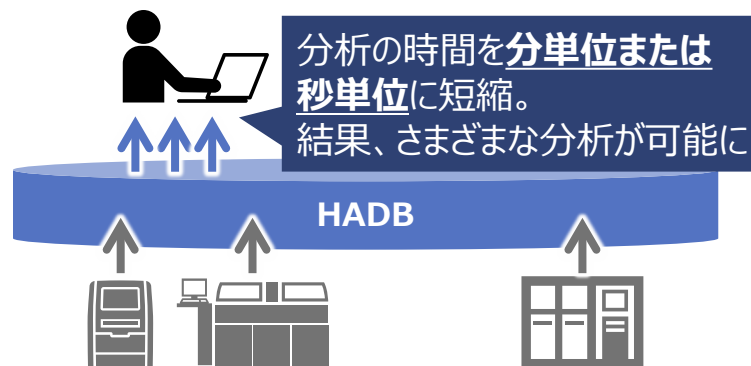
### 解決策

**HADBを導入し、分析の時間を分単位または秒単位に短縮**<sup>※2</sup>。

#### HADB適用前



#### HADB適用後



従来は時間的制約により実現が困難であった悉皆調査<sup>※3</sup>や深掘り調査などが可能に

※1) NDB (National Database of Health Insurance Claims and Specific Health Checkups of Japan)

※2) 2019年6月時点、蓄積されたNDBデータ（約2,000億レコード）全件の解析にかかる時間について、当社従来品と比較

※3) 悉皆（しっかい）調査：調査探究対象を全て漏れなく、また重複することなく調査する方法

#### 【参考URL】

-お知らせ-医療系ビッグデータ解析システムにおける医療経済研究機構・東大生研とのオープンイノベーションの成果の紹介：超高速データベースエンジン Hitachi Advanced Data Binder：ソフトウェア：日立  
[https://www.hitachi.co.jp/Prod/comp/soft1/data-binder/topics/hadb\\_20190712.html](https://www.hitachi.co.jp/Prod/comp/soft1/data-binder/topics/hadb_20190712.html)



## 【流通】顧客売上分析

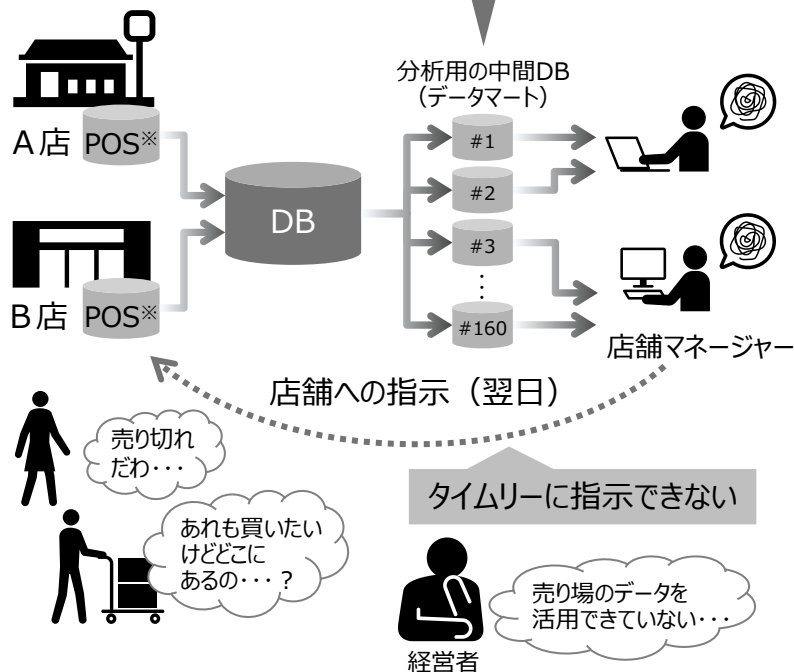
### 顧客課題

- ・分析軸ごとの中間DB（データマート）作成に時間を要し、適切な商品を提供するための指示に時間がかかる。
- ・定型分析しかできないため、詳細なお客さま状況は各店舗に問い合わせる必要があった。

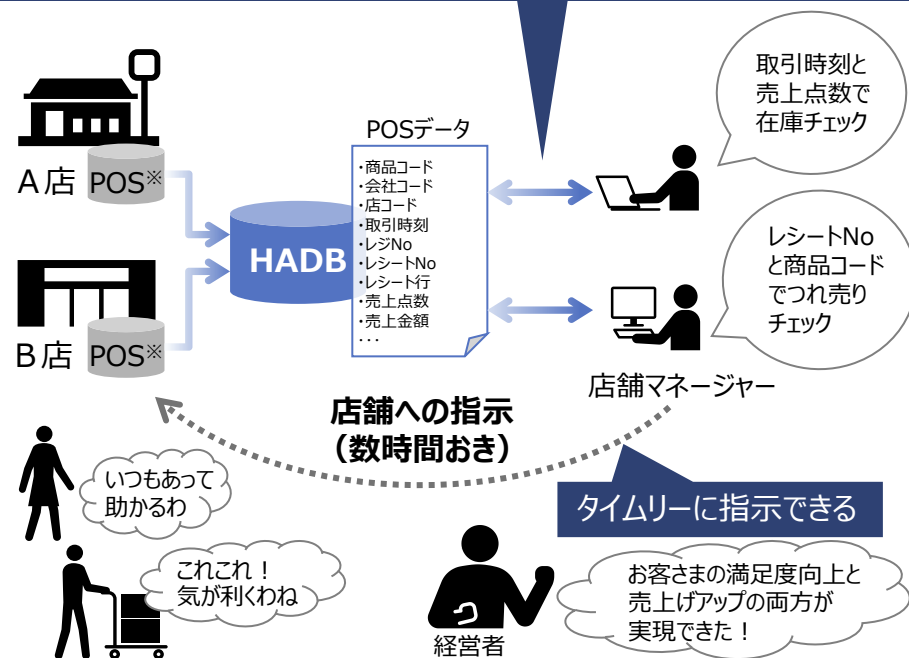
### 解決策

**HADBを導入し、データマートを削減、非定型の分析を可能にすることで、素早い状況把握とタイムリーな各店舗への指示で売り上げアップにつなげる。**

店舗軸×商品軸×時間軸で分析するために  
**160個**のデータマート作成に半日以上を要する。



**1時間以内**で目的に応じた分析用データの抽出を実現



- ・ 売り場への指示を素早くタイムリーに出し、**チャンスロス**を削減
- ・ お客さまの欲しい品物をまとめて置くなど、**気の利いたサービス**を提供
- ・ マート関連の**運用コスト、バッチ保守コスト**の削減

※) POS (Point Of Sales)

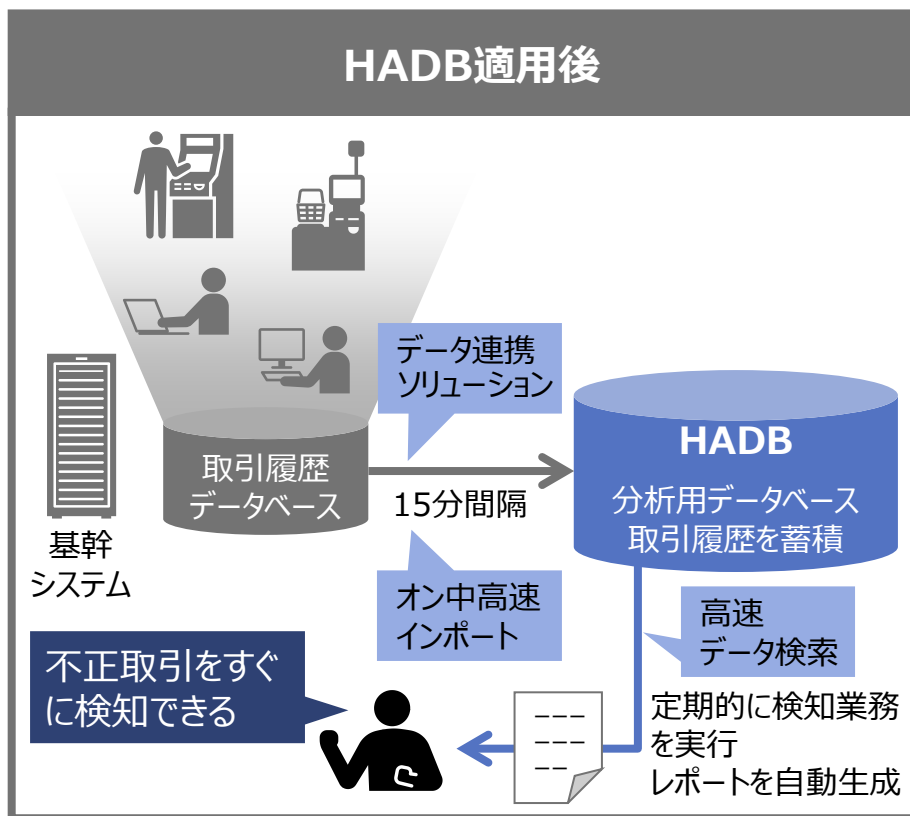
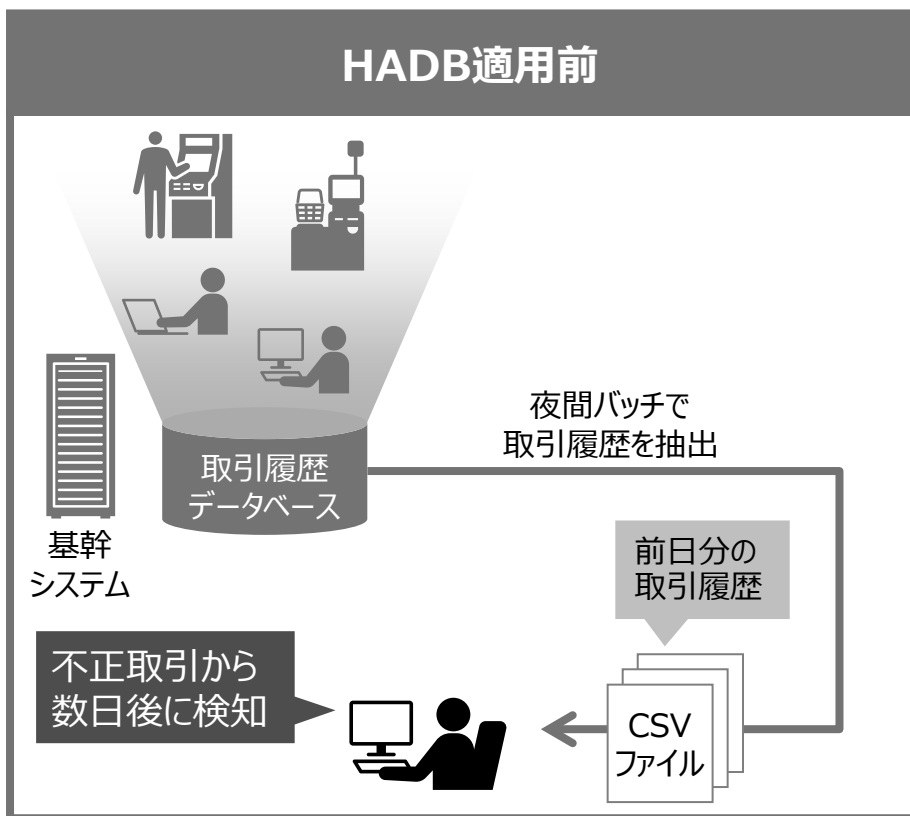
## 【銀行】不正取引検知

### 顧客課題

- ・夜間バッチ処理の関係で**前日分の取引履歴**までの調査しかできない。
- ・取引履歴の検索に時間がかかる。

### 解決策

データの鮮度を保ったままの大量の取引履歴を高速に検索・活用して不正取引の検知が可能に  
 （**15分間隔で取引履歴をインポート/1時間ごとに過去数億件のデータと突き合わせ**）。



#	分類	目的	実施内容	効果
1	電力	● 大規模停電の未然防止 (局所的な系統事故や系統動揺の波及による大規模停電が増加)	● 過去の類似対策事例や、各地に設置したPMU※の計測データ <b>(数千億件/年のデータ) をHADBに蓄積し活用</b> することで、送電網の大規模停電となる予兆を迅速に把握。	● 問題特定や過去の類似対策事例提示などの意思決定時間の削減
2	製造	● 製造実績トレースによる品質管理 (データ活用基盤が整備されていないため、可視化などによる品質向上ができていない)	● <b>各工程実績データをHADBに蓄積・関連付け</b> し、製造情報を可視化／検索に活用。	● 製造から納品までのトレース実現により、影響範囲の特定を実現
3	金融	● 振込データより営業先の分析	● 営業先の分析のため、過去の <b>振込データ数億件より夜間バッチで集計</b> 。	● 夜間バッチ実行時間の短縮
4	社会	● 交通センサーデータの分析	● <b>1日1千万件以上のセンサーデータを5分間隔でインポート、数100TB蓄積し、分析に活用</b> 。	● 準リアルタイムでの大量データインポートと分析
5	社会	● 設備データの状態解析	● 高頻度なデータインポート、 <b>1日最大数十億レコード/日、100TB以上のセンサー情報を蓄積し、分析</b> 。状態確認、傾向分析、保守点検などに活用。	● 対象設備の拡大・新型化による監視データの増加(数百倍)に対応。

※) PMU (Phasor Measurement Unit.) : フェーズ情報計測装置

【参考URL】

#1 : 【活用イメージ】電力 (大規模停電の未然防止) : 事例 : 超高速データベースエンジン Hitachi Advanced Data Binder : ソフトウェア : 日立  
<https://www.hitachi.co.jp/Prod/comp/soft1/data-binder/case/electric/index.html>

## ■ 商標類

- Amazon Web Services、AWS、Powered by AWS ロゴ、Amazon Elastic Block Store、Amazon Elastic Compute Cloud、Amazon Elastic File System、Amazon Redshift、Amazon S3、Amazon Simple Storage Serviceは、Amazon.com, Inc. またはその関連会社の商標です。
- Azureは、マイクロソフト 企業グループの商標です。
- Intel Xeonは、Intel Corporation またはその子会社の商標です。
- Db2、Netezza、SPSSは、世界の多くの国で登録されたInternational Business Machines Corporationの商標です。
- Oracle®、Java及びMySQLは、Oracle、その子会社及び関連会社の米国及びその他の国における登録商標です。
- Pentahoは、Hitachi Vantara LLCの商標または登録商標です。
- Qlik Senseは、米国および他の国々で登録されたQlikTech International ABの商標です。
- Red Hat、Red Hat Enterprise Linux、Shadowmanロゴ、およびJBossは、米国およびその他の国におけるRed Hat, Inc.の登録商標です。Linux®は、米国およびその他の国におけるLinus Torvalds氏の登録商標です。
- RHELは、米国およびその他の国におけるRed Hat, Inc.の商標または登録商標です。
- SAP、および本文書に記載されたその他の SAP 製品、サービス、ならびにそれぞれのロゴは、ドイツおよびその他の国々における SAP SE の商標または登録商標です。
- その他記載の会社名、製品名等は、それぞれの会社の商標もしくは登録商標です。

## ■ 対象となる製品

記載の仕様は、Hitachi Advanced Data Binder 05-10です。  
製品の改良により予告なく記載されている仕様が変更になることがあります。

END

---

**超高速データベースエンジン  
Hitachi Advanced Data Binder(HADB)のご紹介**

2024/07/11

株式会社 日立製作所  
クラウドサービスプラットフォームビジネスユニット  
マネージド&プラットフォームサービス事業部 DB部



Hitachi Social Innovation is  
**POWERING GOOD**