巨大表形式データ
活用上の3課題を解決した
写像表形式データのデモ

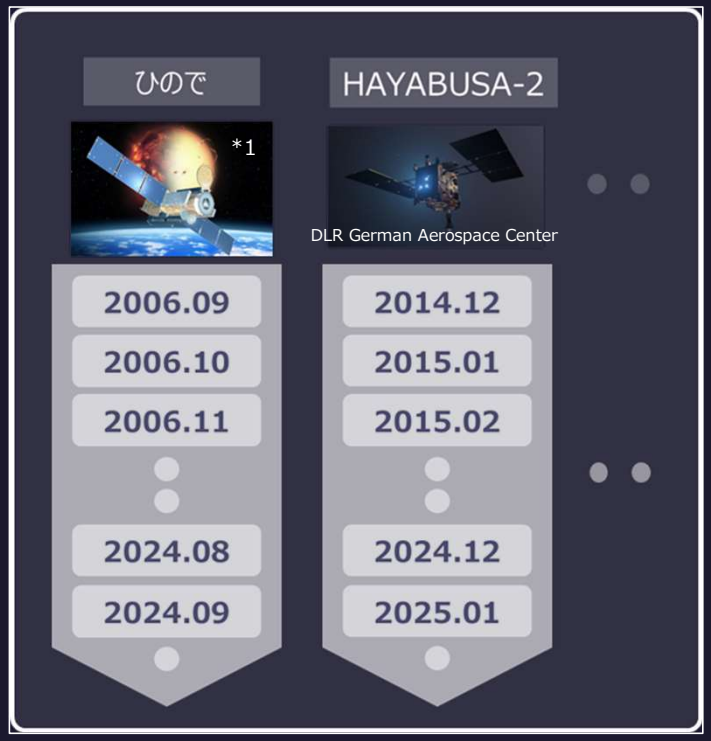
NNIテクノロジーズ株式会社

古庄 晋二

2025.02.28 rev13

本 編

ビッグアーカイブ（BA）とは、



*1. https://www.jaxa.jp/projects/sas/solar_b/images/solar_b_main_001.jpg

ビッグアーカイブ（BA）とは、

日々蓄積されることで巨大になる表形式データです。例えば、「ひので」や「HAYABUSA-2」の HK*2 Data は ビッグアーカイブ（BA）です。

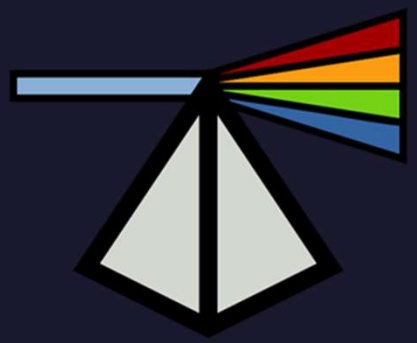
このような BA は、全体として数兆レコード*3 あったり、数万カラム*4 あったりする巨大な表形式データです。そのため、コストや運用上の制約から、多数のファイルに分割して保管されることが普通です。その結果、その利用は非常に厄介になりがちです。

しかし、JAXA に限られず、BA は現代社会のあらゆる活動を記録する貴重な情報資産であり、AI と連携して適切に活用することができれば、多様な分野で画期的なイノベーションを生み出す可能性を秘めています。

*2. HK: House Keeping
*3. 数兆レコード：電話会社の通話記録、巨大IoTでの例
*4. 数万カラム：半導体製造装置、最近の宇宙機

BA の 3課題 を解決する
写像表形式データ実現までの歩み

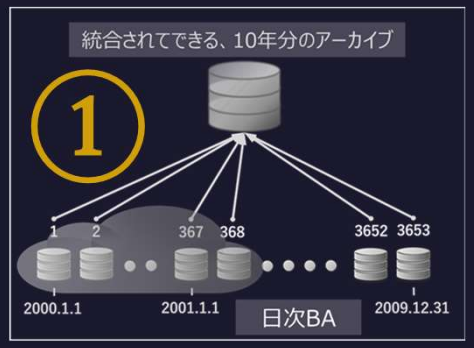
1998年



すべての表形式データは、値の情報（文字列や数値など）と位置の情報（自然数）で構成されています。前者は１つ、後者は異なる特性を持った複数の成分に一意に分解することができます。

位置の成分はさまざまに組み合わせることが可能で、多様な自然数ならではの性質を利用した、高速化アルゴリズム群を生成できます。この成分の構成は全てのカラムで共通ですから、これらの高速化アルゴリズム群は全てのカラムで利用できます。同様に全てのカラムの組合せ、全ての部分集合に対しても高速化アルゴリズムが存在します。自然数の性質を使うことで多様な処理の高速化を全ての場所で行うことが可能になります。

このアルゴリズム群は検索をはじめとした処理を高速化することができるのでインデックスと呼べます。およそ20年後に自然数インデックス（NNI、Natural Number Index）と命名されることになる表形式データの多機能・広範囲の高速化のアプローチがここに始まりました。このあと約20年の間、主にインメモリー技術として発展しました。



2019年

JAXAとの共同研究として、「巨大時系列データの高速アクセス」に関する取り組みを開始しました。そこでの研究対象が後に BA という概念に繋がる表形式データでした。そのBAを効果的に活用するには、次の **3つの課題** を解決する必要があることが分かりました。

① **目的別BA** (後述) を作る 遅さ、困難さ
② 目的別BA へのアクセスの遅さ
③ 目的別BA の保管・転送の困難さ

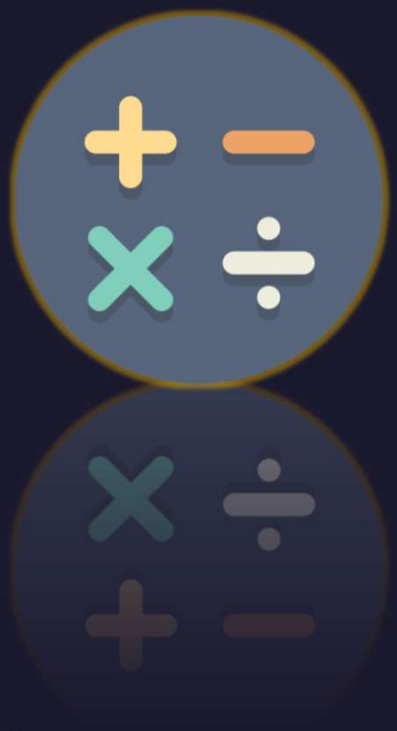
加えて、コスト上の制約から BA は低速な環境に保管せざるを得ないことも問題でした。

（①）**目的別BA** とは期間や地域などで分割された表形式データを格納する多数のファイルから、必要な期間や地域のファイルを選んで結合、多数のカラムの中から必要なカラムを所望の順番に抽出、単位などの変換、さらにレコード位置の調整を行って生成される 新たな BA です。

（②）目的別BA は巨大であることが多く、その場合、例え高速なクラウド環境で処理しても効率的な研究や AI との連携に不可欠なリアルタイムのレスポンスは望めません。

（③）目的別BA は保管や転送が困難です。保管できないと追試や研究継続に支障が出ます。転送できないと複数の人や組織で共同で研究することが難しくなります。

2023年-1



NNI代数

NNI(Natural Number Index)代数の開発に成功しました。これにより、次に述べる実表形式データ（D5A）を基盤に据え、その上に構築される写像表形式データ（MTD）という革新的な表形式データが実現されました。

実表形式データ（D5A）

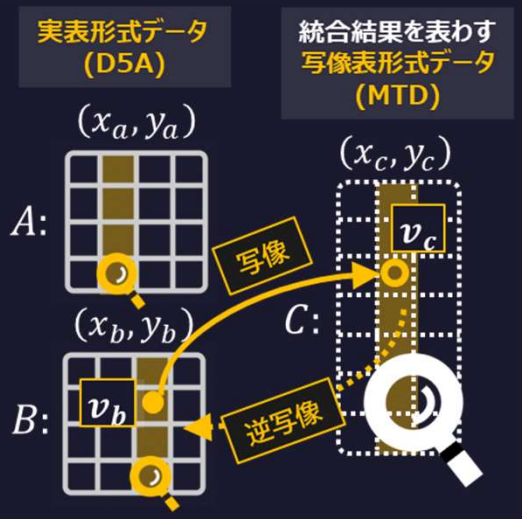
オンディスク環境での使用に適した成分群に分解された表形式データを、後述する写像表形式データと区別して実表形式データと呼びます。成分群に分解されていますから、あらゆる場所にインデックスが存在しています。

この実表形式データは、D5Aというフォーマットに格納され、ファイルとして保持されます。そしてこのファイル（以降、D5A）はいわば、あらゆる場所にインデックスが付いた CSVです。ファイルサイズは CSVの 1/3 程度であることが多く、ディスク容量の消費は大きくはありません。

ファイルですからさまざまな環境（PC、スマホ、サーバー、クラウド）に置くことができ、置かれた環境でそのまま利用することができます。

このような特長を持つD5Aは、分割された BAの保持に適しています。

2023年-2



写像表形式データ（MTD）

1つまたは複数のソースデータ（D5Aや他のMTD）からの写像で定義された写像表形式データ（MTD[*1]）は 目的別BAを実現します。値や位置の情報を持たないため非常にコンパクトで、D5A同様にファイルとして保持されます。

MTDは3つの課題を一気に解決します。

第1の課題の解決：遅延転送（後述）機能により、目的別BAをリアルタイムに生成でき、どんなに巨大なデータでも軽快に表示できます。

第2の課題の解決：インデックス継承（後述）機能により、MTDのあらゆる場所にインデックスがありアクセスを高速化します。

第3の課題の解決：MTDはコンパクトであるため、保管・転送が容易で、第3の課題も解決します。

※ **遅延転送**とは、値が必要になったセルの値だけを、必要になった時に、逆写像を辿ってD5Aから取得する機能です。

※ **インデックス継承**とは、遅延転送と同様に、インデックスを D5Aから取得する機能です。

*1. MTD: Mapped Table formatted Data

実現可能になった
新 BA システム

分散テーブルファイル機構 を構成する D5A と MTD

写像表形式データ（MTD）

役割：利用目的に沿った形のBAを提供

機能：ソース（D5A もしくは 他のMTD）から継承した、
値とインデックス*1 を提供する

形態：ファイルとして存在
（値もインデックスも保持しないため、非常にコンパクト）

実表形式データ（D5A）

役割：分割されたBAを保持

機能：値とインデックス*1 を提供する

形態：ファイルとして存在

*1. どのカラム、どのカラムの組合せ、どの部分集合にも有効なインデックス

分散テーブルファイル機構 として

実表形式データ（D5A）および写像表形式データ（MTD）はそれぞれファイルであり、これらのファイルをアクセスするライブラリーがあれば利用することができます。そのため、その稼働環境はPCやスマホでも良く、大規模な設備（クラウドやサーバー）を必要としません[*1]。このような D5A+MTD からなるシステムは、「**フルインデックスな分散テーブルファイル機構**」と考えられます。

	BigQuery	実＋写像 表形式データ
仕組み	分散並列コンピューティング	フルインデックスな分散テーブルファイル機構
動作環境	クラウド上（限定的な実行環境）	ファイルシステム上（幅広い実行環境）
可能な処理	幅広い	インデックスで加速可能な処理とその応用に限定
コスト	高い（多数のCPUを使う）	低い（SSD上に格納するだけ）
レスポンス	多くの処理が何秒、それ以上	多くの処理が1秒以下、しばしばミリ秒単位
データ更新	可能	追記のみに限定
データ統合（結合・抽出・変換）	長時間	リアルタイム
データの保管・転送	長時間	リアルタイム

クラウド上のシステム（例えば BigQuery）に比べ、インデックスによる加速機構を用いるため高速であり、ファイル機構であるため低コストでポータビリティーが高いというメリットがあります。

一方、実質的に処理できることはインデックスで加速可能な処理に限られ、データ更新も実質的に追記に限定されます[*2]。

このような特性により、**既存のデータベースシステムを補完し、**これまで難しかった課題を解決することができます。

*1. 大規模なD5Aファイルを生成する場合は十分なメモリーおよびSSD が必要です。

*2. 更新はD5Aファイルを再作成することで可能です。

デモムービー

1. 目的別 BA を 作る

- **（結合）**1,000個結合し 31.5兆レコードに
- **（抽出）**全カラムを抽出している
- **（変換）**各ソースの日付時刻を変換している

以上全てを行っても 1秒未満

2. 目的別 BA の 表示・検索 を行う

- **（検索）**100～110の範囲を検索し、
  307,999,653,000 ヒット

上記の検索 30 mSec

3. 目的別 BA の 保管 を行う

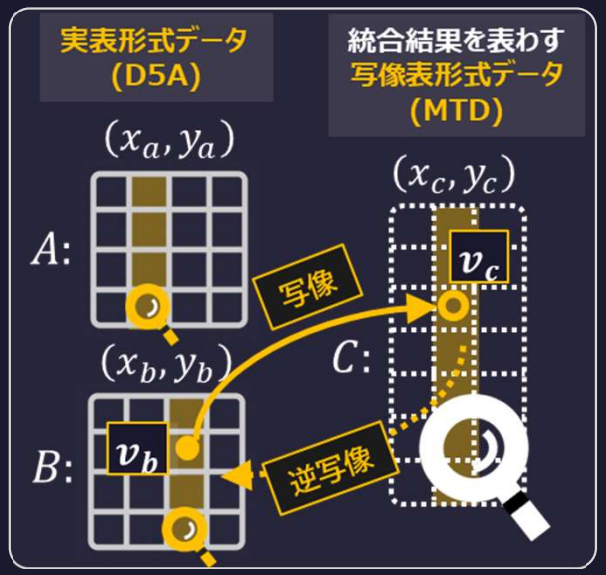
- **（写像表形式データのサイズ）** 512.4 KB

https://youtu.be/rFROcVKHutA

注. 録画が古く、D5Aファイルの拡張子は .bcf となっています。
原則は .d5a です。

https://youtu.be/vDsjX5QRVMQ

上記「目的別 BA を作る」を参照
https://youtu.be/rFROcVKHutA



BA の 3課題を解決した D5A+MTD

1. **ビッグアーカイブ**（BA）とは日々蓄積され、巨大になる表形式データです。
2. **実表形式データ**（D5A）は、値とインデックスを格納した表形式データで、ファイルとして存在します。
3. **写像表形式データ**（MTD）は、ソース（他の MTD または D5A）からの写像の定義を格納したファイルとして存在します。値とインデックスを持たず、それらは D5A から必要なタイミングで継承します。
4. **MTD** は、目的別BA として使うことができ、BA の3課題を解決できます。
    1. 目的別BA をリアルタイムに生成できます。
    2. 目的別BA の全ての場所にインデックスが存在し、効率良くアクセスできます。
    3. 目的別BA は極めてコンパクトであり、保管・転送が容易です。
5. **D5A と MTD** で、「**フルインデックスな分散テーブルファイル機構**」を構成でき、既存のデータベースシステムを補完することが可能になります。

資 料 編

1. MTDを便利にする
Peta Book（PB）

Peta Book (PB) とは、

これまでの説明で、D5A+MTD によって ビッグアーカイブ（BA）の利用における3つの課題を解決できることを示しました。
しかし、それだけでは実際の運用はできず、以下の機能が必要になります。

1. **メタ情報の管理**
   各カラムの属性、関連ドキュメント、他データとの関係性を一元管理し、データの意味を明確化。
2. **データ表示・可視化**
   グリッド表示やグラフ表示で直感的にデータを確認し、CSVやJSON形式でのエクスポートをサポート。
3. **アカウント・セキュリティ管理**
   ユーザーごとのアクセス制御、変更履歴の管理、データの暗号化による安全な共有。
4. **業務特化アプリの統合**
   サプライチェーンデータのBOM展開や異常検知など、BAに応じたアプリを提供。

これらの機能を MTD と統合して利用可能にし、PB は BA の統合、検索から活用までをシームレスに支援するプラットフォームとなります。

PB-Copilot

これまでの説明で、Peta Book（PB）を活用することで、ビッグアーカイブ（BA）へのアクセスが飛躍的に容易になることを示しました。しかし、単にアクセスが可能になっただけでは、データから有意義な知見を導き出すには十分ではありません。BAを的確に活用するためには、以下のような情報を理解しておく必要があるからです。

- **BA の成り立ちと背景**
- **データの意義と特性**
- **過去に得られた知見**

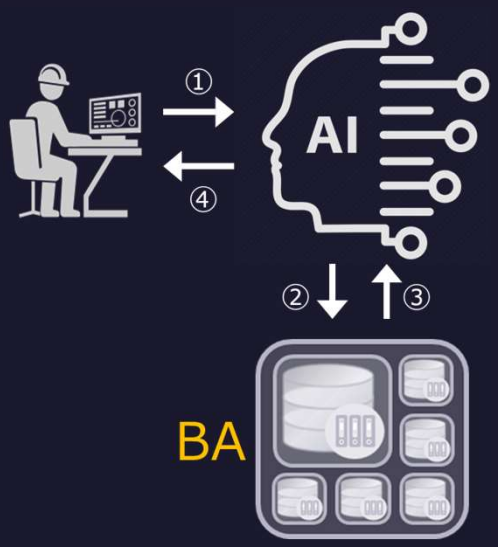
これらを事前に把握していないと、PB を活用しようとしても、どのようにデータを読み解き、どの観点から分析すべきかが分からず、適切なインサイトを得ることが難しくなります。この課題を解決するのが LLMで実現される PB-Copilot です。PB-Copilot は、各 PBに紐づく詳細情報や関連するレポート・論文を事前に学習し、ユーザーが PBを利用する際に、適切なアドバイスを提供します。これにより、以下のような利点が生まれます。

1. **PB の利用ハードルを大幅に低減**
2. **より高度なデータ活用が可能に**
3. **知見の蓄積と共有を促進**

PB-Copilot によって、PBの利便性が大幅に向上し、PBは単なるデータ参照ツールではなく、継続的な知見の発展を支えるプラットフォームとしての役割を果たせるようになります。

2. BA の RAG を実現する Peta Book（PB）

BA の RAG を実現する Peta Book



ビッグアーカイブ（BA）は、センサーログ、販売履歴、アクセスログなど、日々蓄積される膨大なデータ群です。これらのデータには貴重な知見が含まれていますが、適切に検索・活用することが難しいという課題がありました。

LLM（大規模言語モデル） が BA にリアルタイムでアクセスし、必要な情報を的確に抽出できるようになれば、意思決定の迅速化と対応力の向上が可能になります。

この課題を解決するのが Peta Book です。
Peta Book は、

- 目的別BA をリアルタイムに生成
- 目的別BA のどこでも迅速に検索

を実現することで、BAのRAGを実現できます。

例えば、コールセンターでは、顧客から問い合わせがあった際に、PBを活用することで、

1. 該当商品のトレーサビリティデータ（製造・流通履歴）を即時取得
2. 関連する製造・流通プロセスの異常や過去の対応履歴 を即座にリストアップ
3. 過去の類似事例とその対応結果 を参照し、最適な回答を自動提案

といった、リアルタイムかつ的確なサポートが可能になります。Peta Book により、LLMがリアルタイムにBAを活用できるようになると、その結果、さまざまな分野で高度な知識活用と意思決定の革新が実現されます。

3. Peta Book（PB）を推進する NNIアライアンス

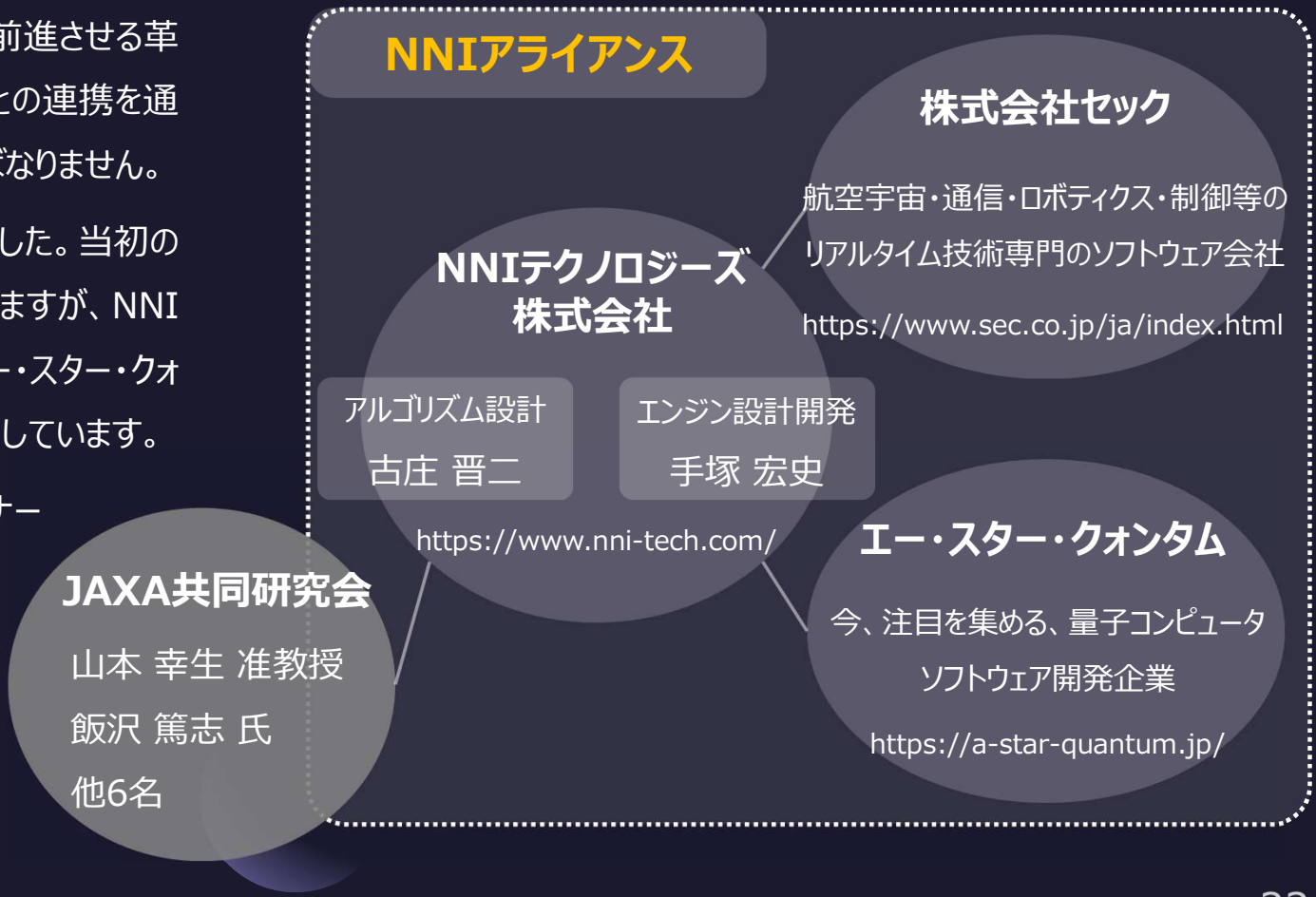
Peta Book を推進する NNIアライアンス と パートナー

Peta Bookは次世代のデータ活用を大きく前進させる革新的なソリューションであるため、多くの企業との連携を通じたグローバル展開を是非とも推進しなければなりません。

その中心となるNNIアライアンスが結成されました。当初のNNIアライアンスは、スモールスタートではありますが、NNIテクノロジーズを中心に、株式会社セック、エー・スター・クォンタムなど、先端技術をリードする企業が参加しています。

また、JAXA共同研究会もその重要なパートナーとなっています。

これらにより、今まさに、Peta Bookが最初の一歩を踏み出そうとしています。



付録　1

インデックスの2つのモード

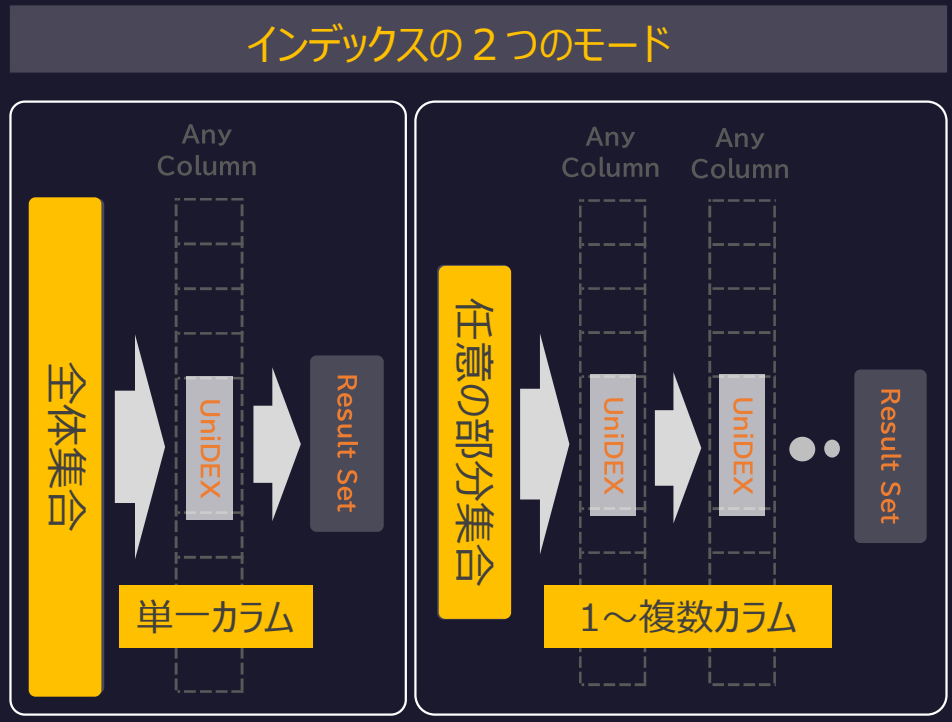
D5A上に存在し、MTD上でもD5Aから継承して使えるインデックスがあり、これを UniDEX と呼びます。

UniDEX は、全体集合や部分集合から、任意の1～複数のカラムに条件を設定して絞り込むことで検索を行います。

1. **高速モード**
   任意の1カラムに条件を設定して、全体集合から絞り込みます。このモードは、単純で高速な検索が必要な場合に適しています。
2. **高自由度モード**[*1]
   任意の1～複数のカラムに条件を設定して、任意の部分集合からさらに絞り込みます。このモードは、柔軟性が求められる複雑な検索に適しています。

これらの2つの検索モードを適切に使い分ける、あるいはミックスして使うことで、様々な検索要求に迅速かつ柔軟に応答することが可能になります。

*1. このモードの拡張として「類似ベクトル検索」が実装されます。



高速モード
任意の1カラムに対し、全体集合からの検索を高速に行えます

高自由度モード
任意の1～多数カラムに対し、任意の部分集合からの自由度の高い検索を行えます

付録 2

Peta Book の
パフォーマンスレポート

1. Peta Book の総合パフォーマンス
2. 配送トラックデータでの UniDEX のパフォーマンス測定(1/3)
3. 配送トラックデータでの UniDEX のパフォーマンス測定(2/3)
4. 配送トラックデータでの UniDEX のパフォーマンス測定(3/3)
5. パフォーマンスのまとめ

1. Peta Book の総合パフォーマンス

3つの高速化機構

MTD は統合時間、保管・転送、検索の3つの高速化を達成しました。
統合および保管・転送の高速化は MTD そのものの仕組みに由来し、検索の高速化は MTD 上のあらゆる場所に対して有効なインデックスである、UniDEX に由来します。一方、データ取得は1つ1つのセル毎に逆写像を辿る仕組み上、レイテンシーはともかく、スループットは低めになります。
ここでは、3つの高速化機構のパフォーマンスを調べていきます。

#1 - 統合のパフォーマンス

ソースデータ群を統合して作られる MTD は、ソースからの写像で作られています。そのため、MTDが行う統合は、以下になります。

1. 写像の定義を読み込む。
2. ソースとのコネクションを確立。

デモムービー***1**では、1000個のソースを統合し、31.5兆レコードにするのに 0.3秒かかりました。

#2 - 保管・転送のパフォーマンス

MTD はソースから MTD への写像の定義でできています。そのため、どんなに巨大な表形式データを表現しているときでも極めてコンパクトになります。コンパクトであるため、保管・転送がリアルタイムに行えます。
デモムービー***1**では 756PB の表形式データを僅か 512KB で表現しています。この場合、Saveや転送は 0.1秒程度と見積もることができます。

#3 - UniDEXのパフォーマンス

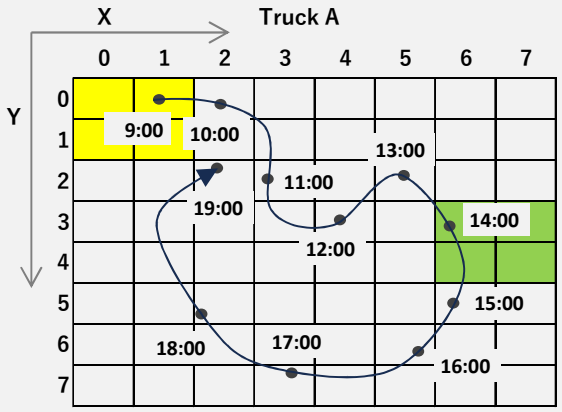
次項以降で解説します。

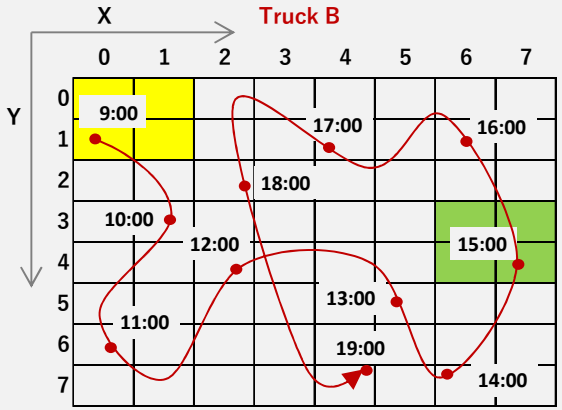
*1 https://youtu.be/wkH49DmSPSQ

2. 配送トラックデータでの UniDEX のパフォーマンス測定（1／3）

	Time		X	Y	kg
0	9:00	A	1	0	2000
1	9:00	B	0	1	1000
2	10:00	A	2	0	1700
3	10:00	B	1	3	800
4	11:00	A	3	2	1600
5	11:00	B	0	6	300
6	12:00	A	4	3	1200
7	12:00	B	2	4	1100
8	13:00	A	5	2	1100
9	13:00	B	5	5	900
10	14:00	A	6	3	800
11	14:00	B	6	7	800
12	15:00	A	6	5	700
13	15:00	B	7	4	500
14	16:00	A	5	6	600
15	16:00	B	6	1	400
16	17:00	A	3	7	400
17	17:00	B	4	1	300
18	18:00	A	2	5	250
19	18:00	B	2	2	200
20	19:00	A	2	2	0
21	19:00	B	4	7	0

10億レコード
13カラム





ベンチマークに用いたデータ

ー 10億レコードのトラック配送記録 ー
（プログラムで生成した疑似データ）

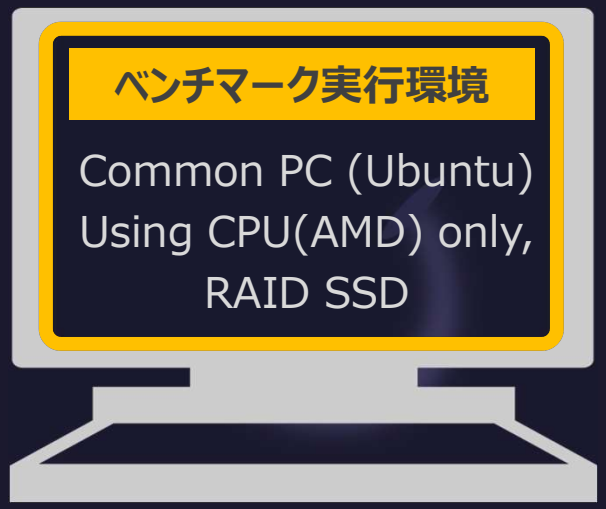
1. 日本全国に5000台の配送トラックを運用する大手運輸を想定。
2. 1分ごとに、各トラックから、緯度・経度・積載量など13カラムのデータを収集したとの想定。
3. 1年366日で総レコード数: **989,297,868**
4. 13カラム
   ①支店名, ②日付時刻, ③経度, ④緯度, ⑤トラックID,
   ⑥ドライバーID, ⑦助手人数, ⑧積載個数, ⑨積載個数率,
   ⑩積載重量kg, ⑪積載重量率, ⑫積載体積m3, ⑬積載体積率

このデータセットを用いて、次に UniDEX の実際のパフォーマンスを検証します

3. 配送トラックデータでの UniDEX のパフォーマンス測定（2／3）

Benchmark-1

高速モード



16通りの 高速モード のパフォーマンス

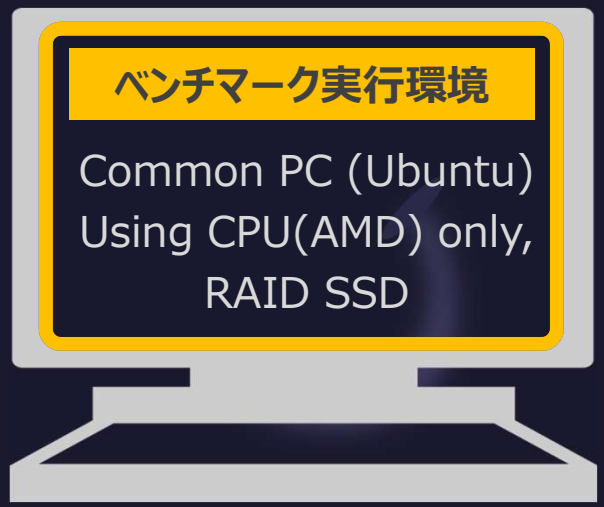
検索条件式の [A, B) は A≦ x ＜ B 、[A, B] は A≦ x ≦ B を表わす

測定#	カラム	検索条件	ヒット数	検索時間(msec)
#1	支店名	札幌市支店	19,779,180	0.62
#2	支店名	札幌市支店|千葉市支店|さいたま市支店|横浜市支店		
			118,712,212	0.14
#3	日付時刻	[2024-03-01,2024-06-01)	248,718,584	0.75
#4	日付時刻	[10:00:00,11:00:00)	109,800,000	5.49
#5	緯度	[30.00,41.00)	946,076,633	2.12
#6	緯度	[35.00,36.00)	284,216,014	1.84
#7	経度	[130.00,140.00]	777,949,877	1.95
#8	経度	[139.00,140.00]	177,995,757	1.39
#9	ドライバーID	[1000000,1030000]	435,283,772	0.22
#10	ドライバーID	[1030000,1040000]	217,679,924	0.08
#11	積載重量kg	[100.0,1300.0)	794,703,211	0.36
#12	積載重量kg	[10.0,20.0)	9,110,234	0.1
#13	積載個数	[50,150)	534,256,412	0.16
#14	積載個数	[100,110)	54,764,041	0.06
#15	助手人数	[0,2]	989,297,868	0.17
#16	助手人数	1	494,656,740	0.06

4. 配送トラックデータでの UniDEX のパフォーマンス測定（2／3）

Benchmark-2

高自由度モード



2通りの 高自由度モード のパフォーマンス

検索条件式の [A, B) は　A≦ x ＜ B 、[A, B] は　A≦ x ≦ B を表わす

測定#	カラム	検索条件	単純検索ヒット数
# 1	支店名	札幌市支店	19,779,180
	日付時刻	[08:00:00,09:00:00)	109,800,000
	緯度	[43.05,43.08)	1,544,843
	経度	[141.32,141.35)	2,014,679
	ドライバーID	[1010010,1010020)	1,979,120
	積載重量kg	[900.0,1100.0)	88,771,362
	積載個数	[80,130)	268,968,867
	助手人数	1	494,656,740
複合検索のヒット数:　351　実行時間(msec): 194.82			

測定#	カラム	検索条件	単純検索ヒット数
# 2	支店名	東京本店|千葉市支店|さいたま市支店|横浜市支店	
			118,712,212
	日付時刻	[08:00:00,09:00:00)	3,100,000
	緯度	[43.05,43.08)	211,029,583
	経度	[141.32,141.35)	64,082,133
	ドライバーID	[1010010,1010020)	118,712,212
	積載重量kg	[900.0,1100.0)	66,292,133
	積載個数	[80,130)	268,968,867
	助手人数	1	494,656,740
複合検索のヒット数:　303　実行時間(msec): 614.60			

5. パフォーマンスのまとめ

これまでのパフォーマンステスト結果を総括すると以下となります。

- 1000個のビッグアーカイブの統合が **1秒未満**、
- 1兆レコードを超えるデータの保管・転送が **1秒未満**、
- 10億レコードに対する、「高速モード」での検索が **1ミリ秒前後**、
- 10億レコードに対する、「高自由度モード」での検索が **1秒未満**、

以上のパフォーマンスの結果から、MTD と MTD上の UniDEX を用いると、ビッグアーカイブにリアルタイムにアクセスできることが分かります。

これにより、**ビッグアーカイブへのリアルタイムアクセス** と、その結果として **AIの活用領域の大幅な拡大** が可能になることは明らかです。

Thank you

https://www.nni-tech.com/

contact@nni-tech.com