

ビッグデータを利活用するための ITプラットフォームとは

2012/11/9

株式会社 日立製作所 情報・通信システム社
ソフトウェア本部
ビッグデータソリューション部

山口俊朗

**Human Dreams.
Make IT Real.**

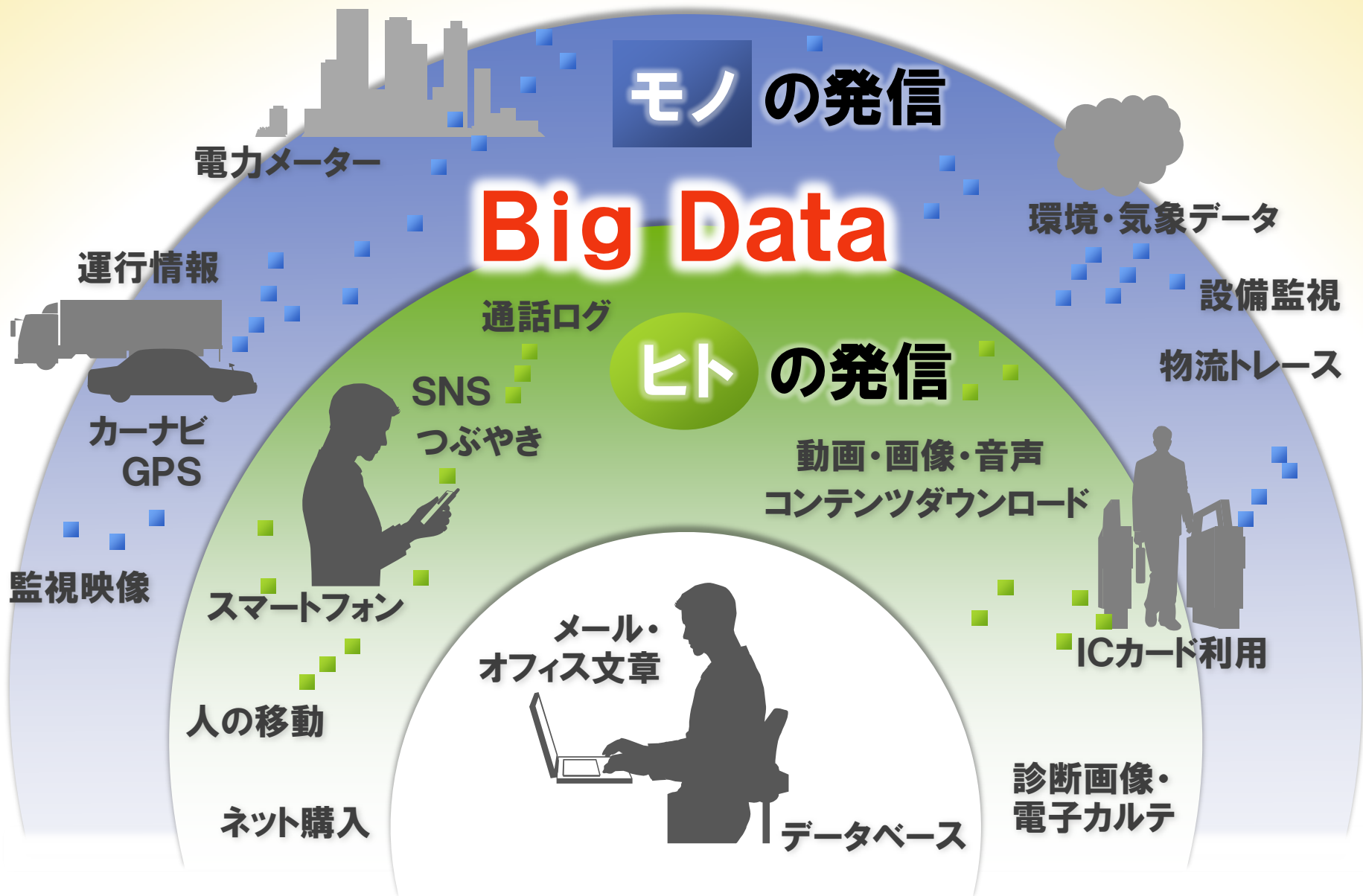
ビッグデータを利活用するためのITプラットフォームとは

Contents

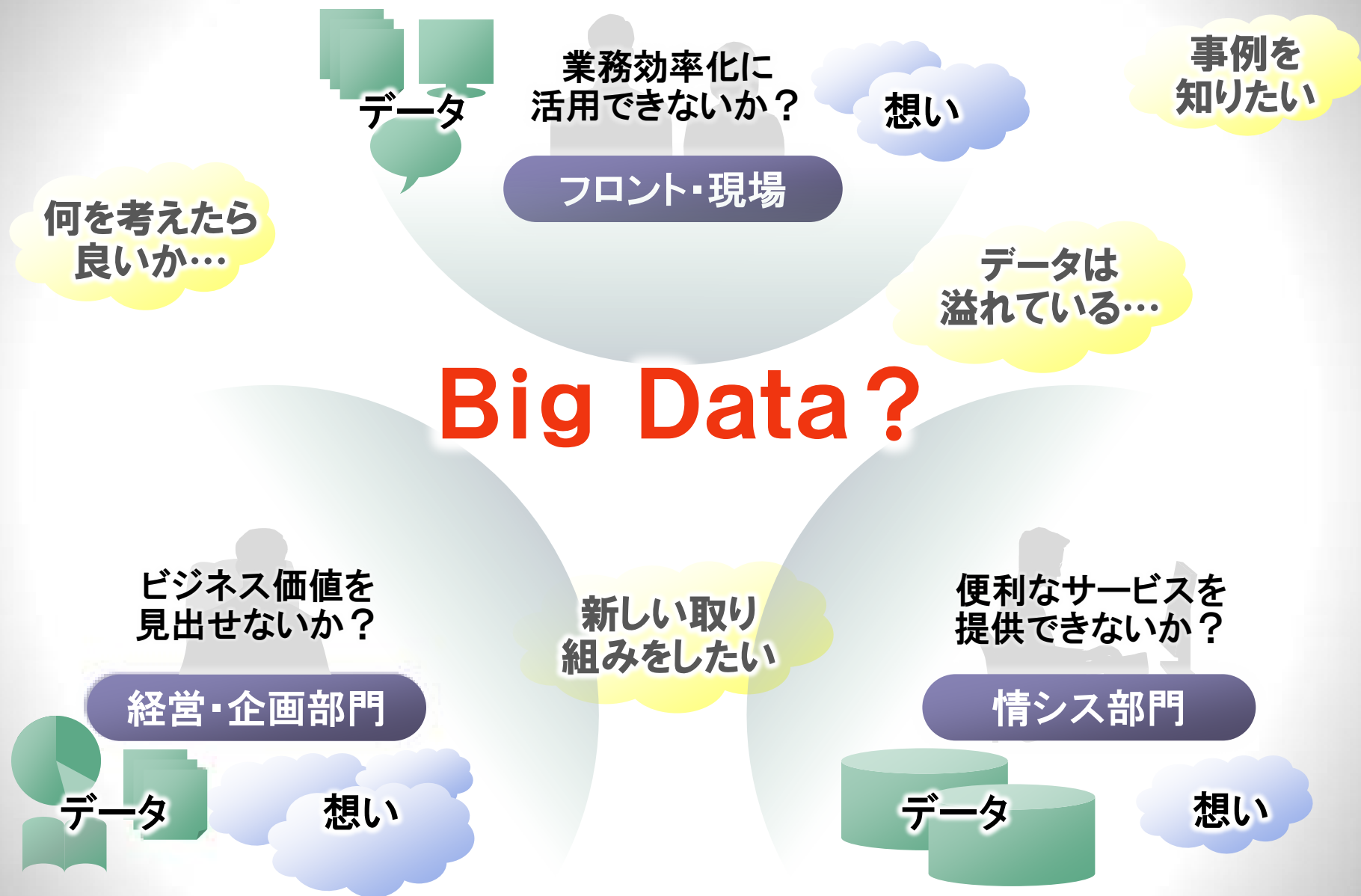
- 1.ビッグデータ利活用とは何か？
- 2.ビッグデータ利活用への取り組み
- 3.ビッグデータ利活用プラットフォーム
- 4.ビッグデータ利活用事例
- 5.まとめ

ビッグデータを利活用するためのITプラットフォームとは

1. ビッグデータ利活用とは何か？



1-2 ビッグデータを使って何をするか？



■ 例えば…

従来型動物園

=動物の姿、形が分かる
形態展示

新しい動物園

=動物の動きが分かる
行動展示

今までと同じ＝うーん。。

利用者の
受け止め方

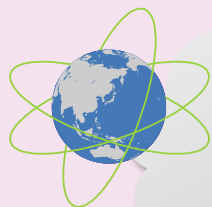
今までには見られなかった
「面白さ」「楽しさ」
⇒わざわざ行ってみたいと思う。

企業は…

新たな提案・イノベーション・顧客の創造をしなければならない。

新しい価値を見出すために、ビッグデータが活路を開く

インターネット黎明期



初期のインターネット

ノイズの山

例えば **Google**

PageRank
(リンクを人気投票とみなす
画期的アイデア)

一気に整理

宝の山

今日の企業内情報



ノイズの山

- 部門間・システム間のサイロ
- アクセスコントロールの複雑さ
- データ構造の多様さ
などが利活用を困難に

未収集・未整理

?

新たな発見・価値創造に活用したいが...

企業内の情報活用は現在まだ黎明期

従来

これから

企業競争力の源泉

- 業務の効率化
- 資源の独占

- イノベーション
- 新たな価値に裏打ちされたサービス・製品・顧客の創造

情報活用の目的・範囲

業務効率化

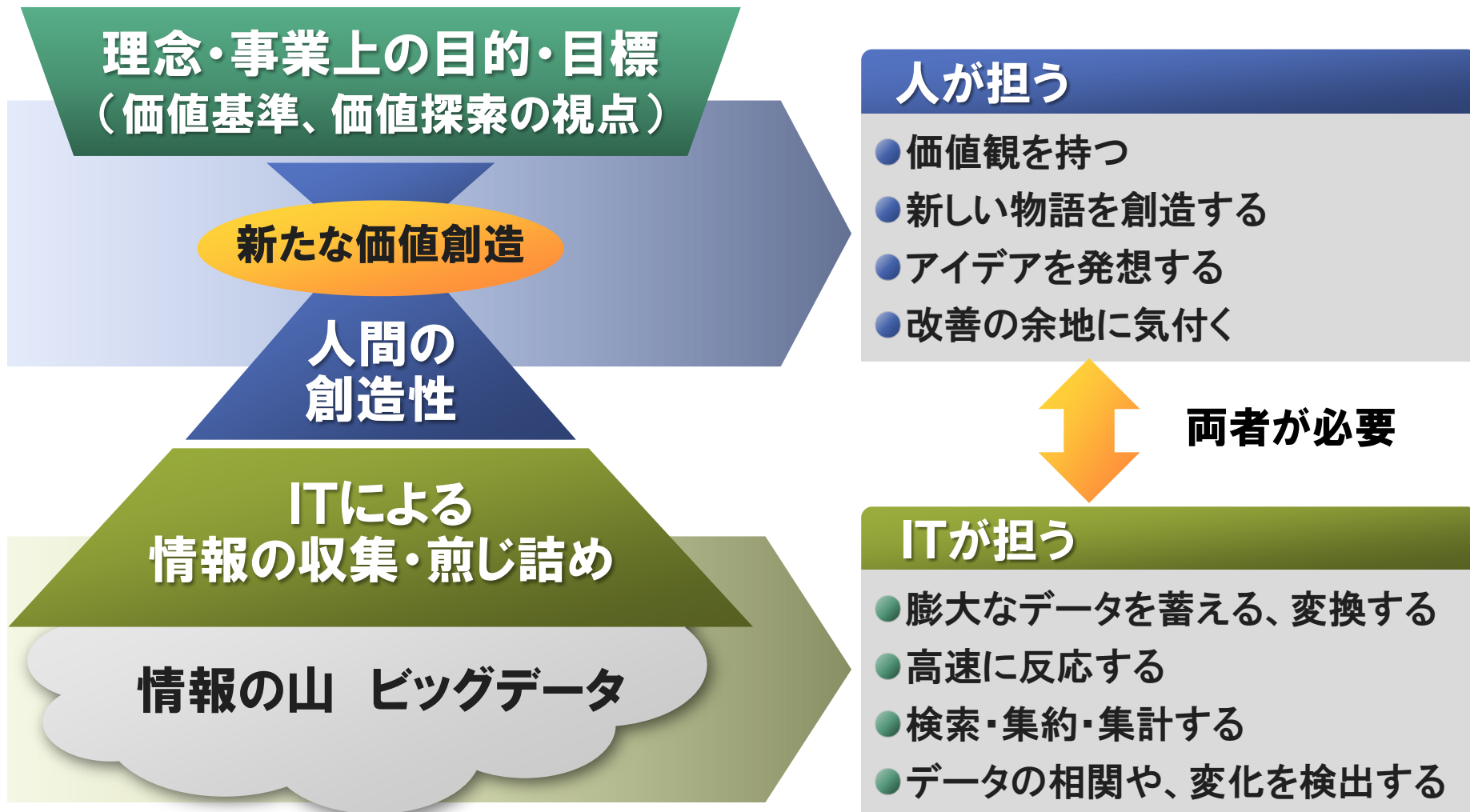
目的志向の
情報活用

新たな価値創造

隠れた
関係性の発見

扱うデータの
種類・範囲が拡大

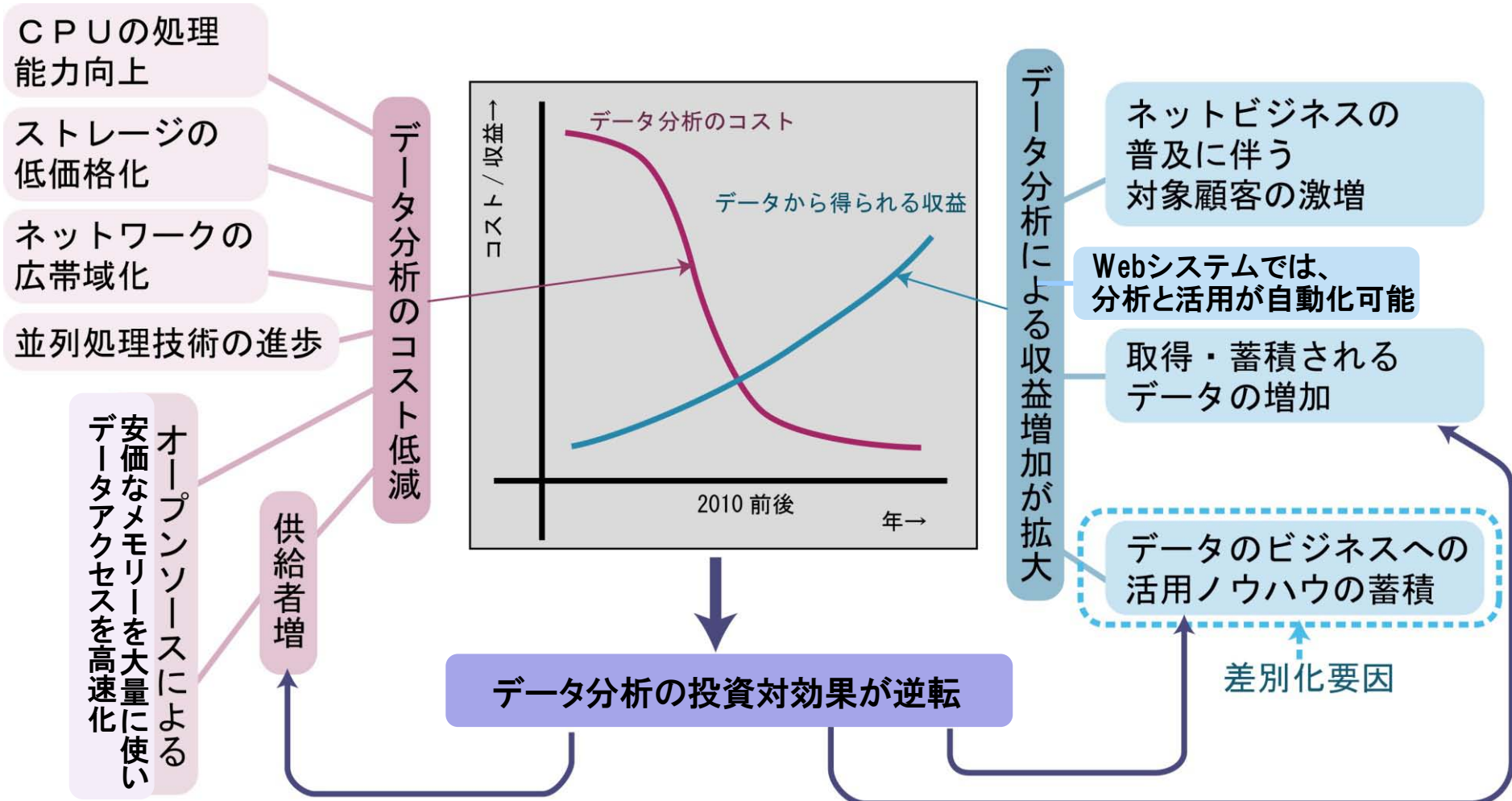
ビッグデータ利活用による新たな価値創造へ期待



価値創造は人間が行う — ITは道具である

1-7 ビッグデータを支える大幅な性能と経済性向上

ITリソースを並列活用する方法や安価なメモリーを大量に活用する方法が出現
投資対効果が逆転し、大量データの蓄積や分析、利活用でビジネスに貢献



第3のデータ

用途が定まらず、大規模なITシステム化の決断はできない一方で、捨てるしまうと将来のイノベーションが失われてしまうようなデータ

例) PageRank以前、ネットのリンク構造のデータは検索エンジンにとって第3のデータだった

従来

これから

第1のデータ
「価値を生む」

活用

第2のデータ
「価値を生まない」

集めない
or
捨てる

第1のデータ
「価値を生む」

活用

第3のデータ
現時点で用途なしだが、後になって価値を生む可能性

将来の競争力の源泉

第2のデータ
「価値を生まない」

捨てる

「第3のデータ」のジレンマ

価値が確信できない



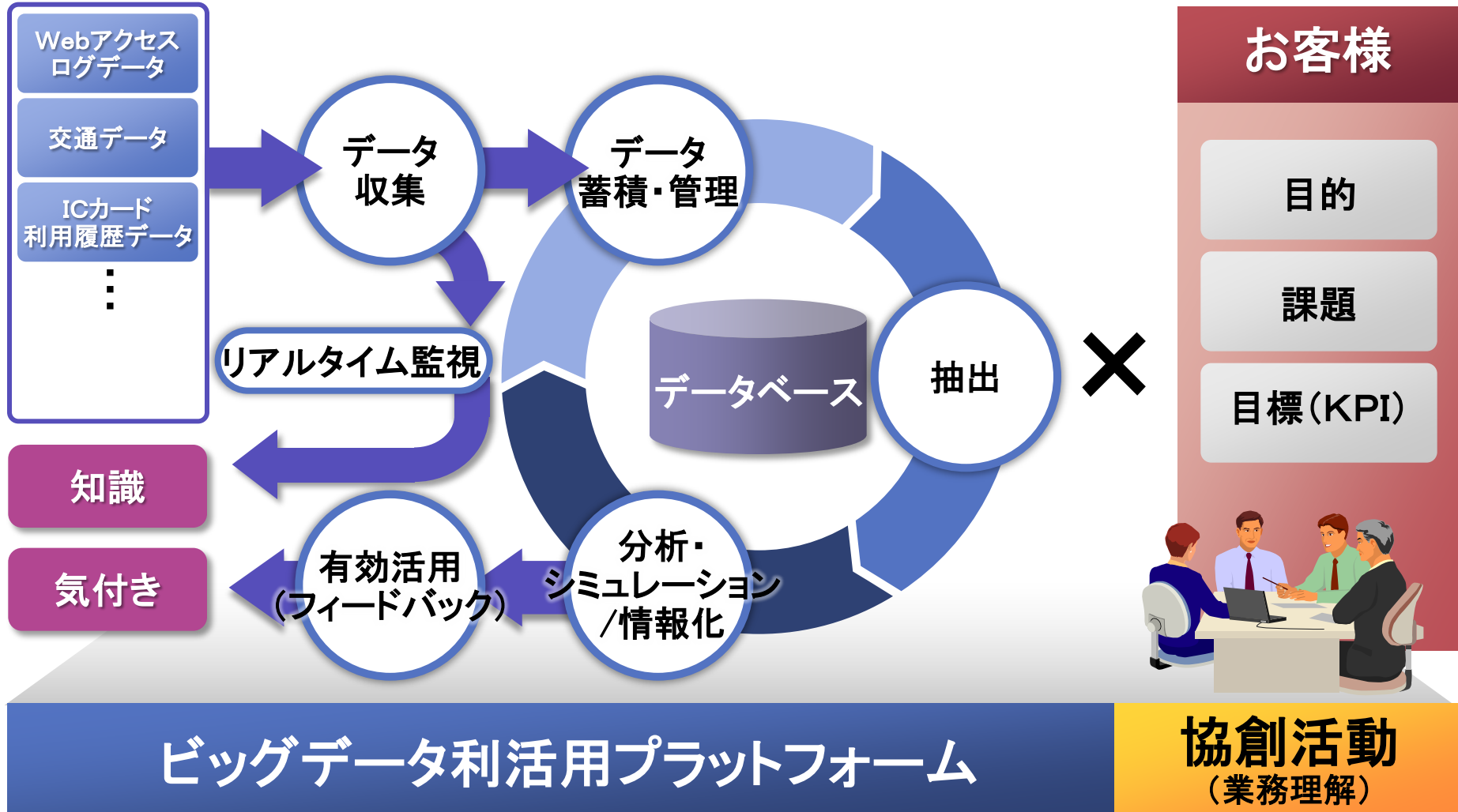
貯めない・価値探索しない

ビッグデータを利活用するためのITプラットフォームとは

2. ビッグデータ利活用への取り組み

2-1 日立のビッグデータ利活用への取り組み

- ビッグデータ利活用プラットフォーム提供
- ビッグデータ分析活動(お客さまとの協創活動の推進)



ビッグデータ利活用プラットフォーム

協創活動
(業務理解)



お客様のビジネス



集積された知見

イノベティブ・アナリティクス



サービス
パートナー

テクノロジー
パートナー



ビッグデータ利活用プラットフォーム

データ可視化

データ仮想化

データ並列化

データ抽象化



Big Data

1 お客様/ビジネスパートナーとの協創の推進



2 ビッグデータ利活用に関するサービスの提供



データ利活用専任組織

「スマート・ビジネス・イノベーション・ラボ」設立

- データ・アナリティクス・マイスターサービス 提供開始

3 ビッグデータ利活用を支えるプラットフォーム製品、技術の開発



データ利活用基盤技術群「Field to Future Technology」体系化

- 「uCosminexus Elastic Application Data store」 提供開始
- 「Hitachi Advanced Data Binder^{※1} プラットフォーム」 提供開始

※1:内閣府の最先端研究開発支援プロジェクト「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社会サービスの実証・評価」(中心研究者:東大喜連川教授)の成果を利用。

現場の事実から未来の業務に不可欠な情報を生成する技術

Field to Future Technology

課題

データの所在特定・収集・連携・関連付け
(見つける、集める、関連づける)

データの保護・管理
(格納する、検索する、守る、統一的に扱う)

処理の高速化
(あらゆるデータ処理の速度を高める)

モデル生成、統計解析
(分析する、圧縮する、トレースする)

解決技術

データ可視化

データ仮想化

データ並列化

データ抽象化

データの
利活用

データの の 利活用

データ可視化

現場のデータを収集し、
既存のデータと連携を図り、
業務とデータの関連を可視化

データ仮想化

データの格納場所、構造、
関連性、内容など違いを隠ぺいし
統一的なデータ管理を実現

データ並列化

データの並列分割と実行を
記憶デバイスの並列性に
自動的に整合し実行

データ抽象化

データの全体像、相互関係、
潜在構造など分析&抽出する
ことでデータを情報へ昇華

2-6 Field to Future Technologyを実装する商品

技術		商品カテゴリ	商品名
データ可視化	ストリームデータ処理技術	ストリーム処理	uCosminexus Stream Data Platform
		サービスレベル管理	JP1/IT Service Level Management
		モニタリング	HiRDB RealTime Monitor
データ仮想化	分散データ管理技術	インメモリーデータグリッド	uCosminexus Elastic Application Data store
データ並列化	並列データ処理技術	データベース	HiRDB
	超高速データアクセス技術	大量データ分析プラットフォーム	Hitachi Advanced Data Binder ^{※1} プラットフォーム
	データ分割処理制御技術	分散グリッド実行基盤	uCosminexus Grid Processing Server
	分散データ管理技術	インメモリーデータグリッド	uCosminexus Elastic Application Data store
データ抽象化	時系列圧縮格納技術	時系列ストア	uCosminexus Stream Data Platform/Data Store
	障害予兆検知技術	サービスレベル管理	JP1/IT Service Level Management
	モデル化機械学習融合技術	サービス	ビッグデータ関連サービス

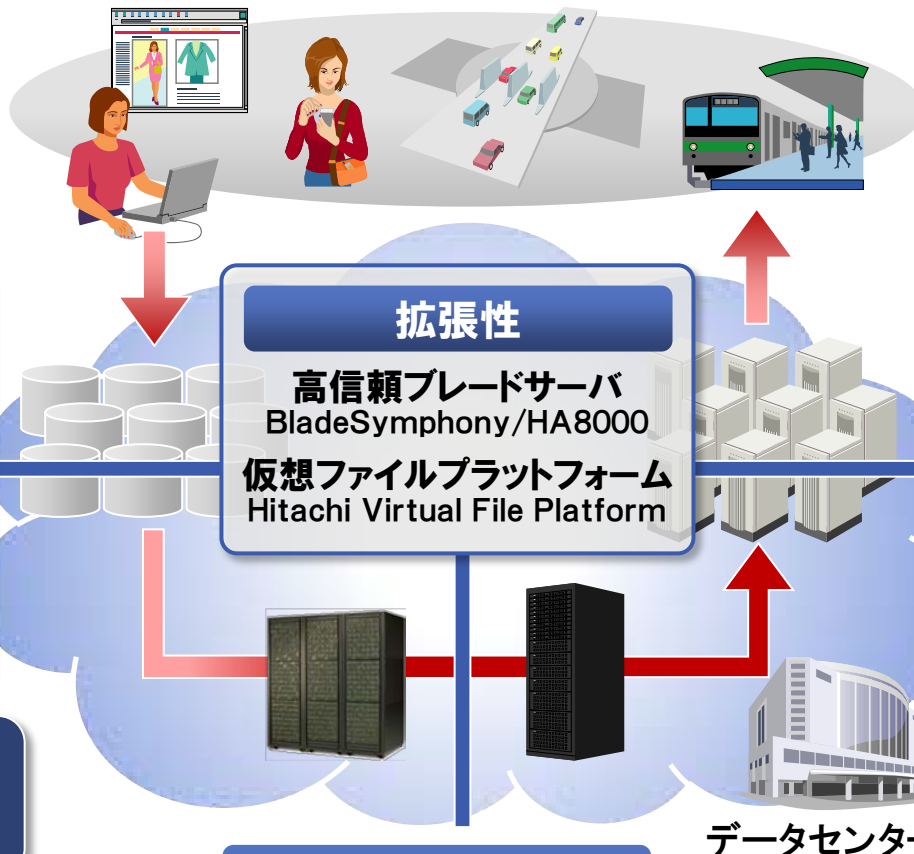
※1:内閣府の最先端研究開発支援プロジェクト「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会的サービスの実証・評価」(中心研究者:東大喜連川教授)の成果を利用。

ビッグデータを利活用するためのITプラットフォームとは

3. ビッグデータ利活用プラットフォーム

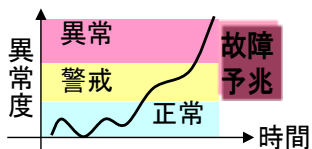
3-1 ビッグデータ利活用プラットフォーム

ビッグデータ



① リアルタイム監視

大量の情報を
リアルタイム処理



Cosminexus

uCosminexus Stream
Data Platform

Cosminexus

uCosminexus Elastic
Application Data store

拡張性

高信頼ブレードサーバ
BladeSymphony/HA8000
仮想ファイルプラットフォーム
Hitachi Virtual File Platform

② 蓄積・検索

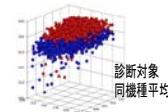
大量の情報を高速検索・効率保管

履歴データ

③ 集計・分析

大量の情報を
対象とした分析

稼働データ
P1
P2
P3
P4



時間

Cosminexus

uCosminexus Grid
Processing Server

Hadoop

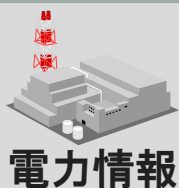
Hitachi Advanced Data
Binder※¹ プラットフォーム

※1: 内閣府の最先端研究開発支援プロジェクト「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: 東大喜連川教授)の成果を利用。

① リアルタイム監視

● ITの瞬発力 = 反射神経系の処理を持つITシステム

実世界から膨大なデータが発生



電力情報



センサー情報



設備情報



稼働情報

今までのITシステム

データの蓄積

「何が起きた」を知る

DB・DWH

「過去」の集計・分析結果

蓄積した膨大なデータを集計・分析 = 大脳系処理

ストリームデータ処理

Cosminexus
uCosminexus Stream
Data Platform

データを逐次入力

「今」の集計・分析結果

POINT

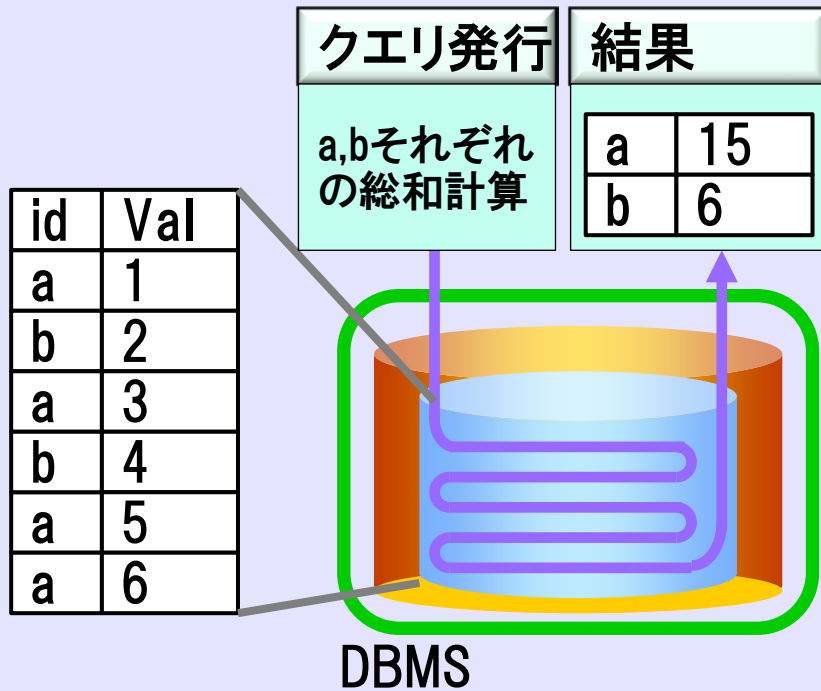
データ発生と同時に、蓄積せずに、集計・分析 = 反射神経系処理

「何が起きてる」「何が起きそう」を知る

DBMS [ストック型データ処理]

データをDBMSに格納し一括処理

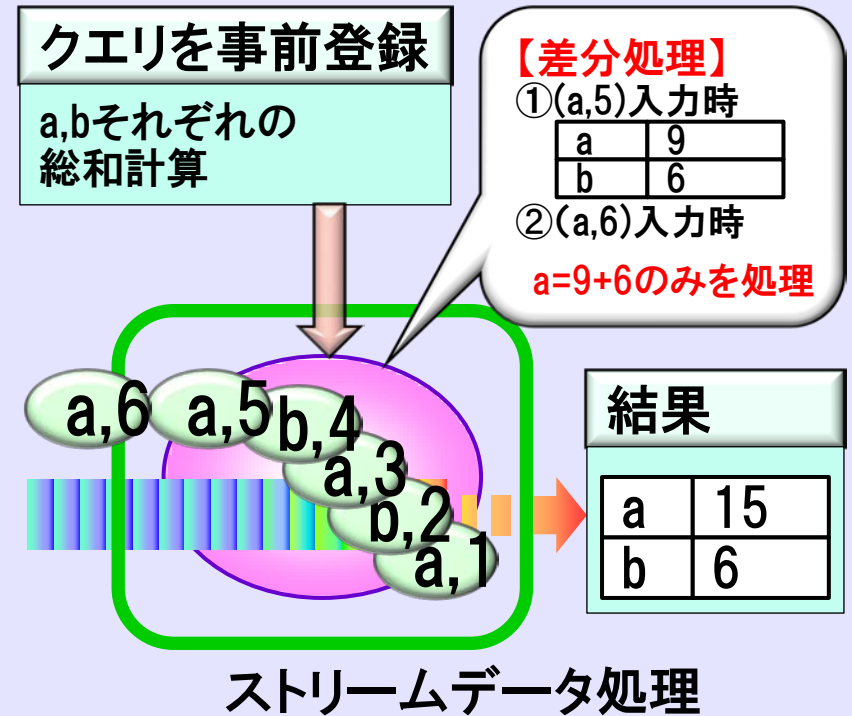
クエリ発行時にすべてのデータを参照しすべての結果を抽出



ストリームデータ処理 [フロー型データ処理]

データ発生時に継続的に逐次処理

データ発生時にそのデータを参照し関係する処理だけを実行



② 蓄積・検索

- インメモリデータグリッドを提供するデータ処理基盤
 - システム内のデータキャッシュ/データバッファ/データハブとして利用
- 適用アプリケーションのメリットは、高レスポンス・高スループット、高可用性、開発容易性

データキャッシュ



必要なデータをリポジトリから
インメモリデータグリッドに事前
読み込み

データバッファ



インメモリデータグリッドで
蓄積したデータをリポジトリに
定期的／一括書き込み

Cosminexus
uCosminexus Elastic
Application Data store



データハブ

アプリケーション／サービス間の
データ受け渡しをインメモリデータ
グリッドで仲介

特徴

シンプルなデータ構造

「キー」と「値」から構成されるKVS

アクセス高速化

メモリ内配置

耐障害性

複製作成によるデータ多重化

位置透過性

データ配置を隠蔽

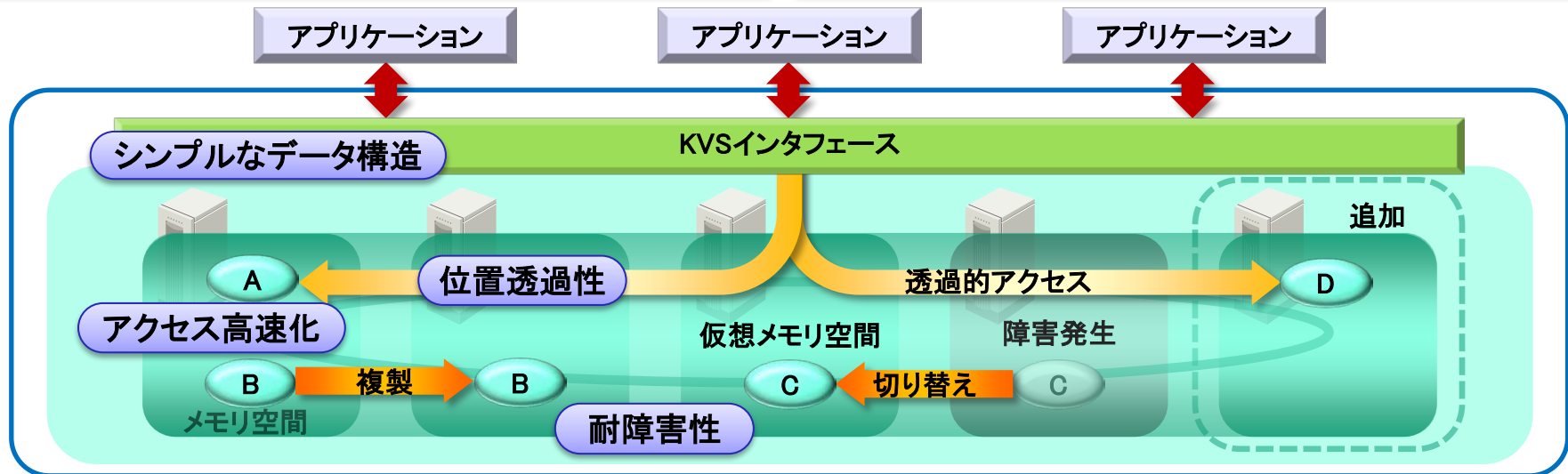
適用アプリケーションのメリット

高レスポンス・
高スループット

高可用性

開発容易性

●		●
●		
	●	
●	●	●



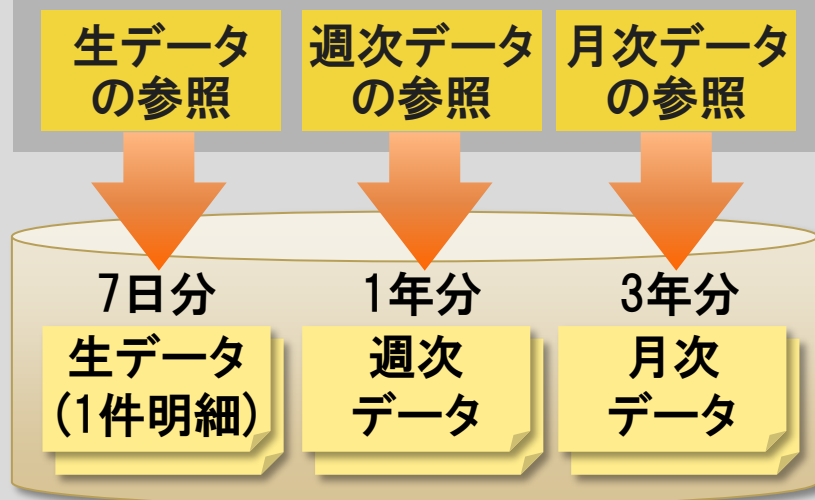
② 蓄積・検索

大幅な性能向上がされ、数億件以上の生データをそのまま扱うことができる。
DB技術は、各社それぞれの状態(インメモリ、カラムストア、Key-Value、...)

従来

- 企業活動での基本的なデータ活用とITリソースの制約から、データの持ち方を事前に設計していた

事前に決めた使い方

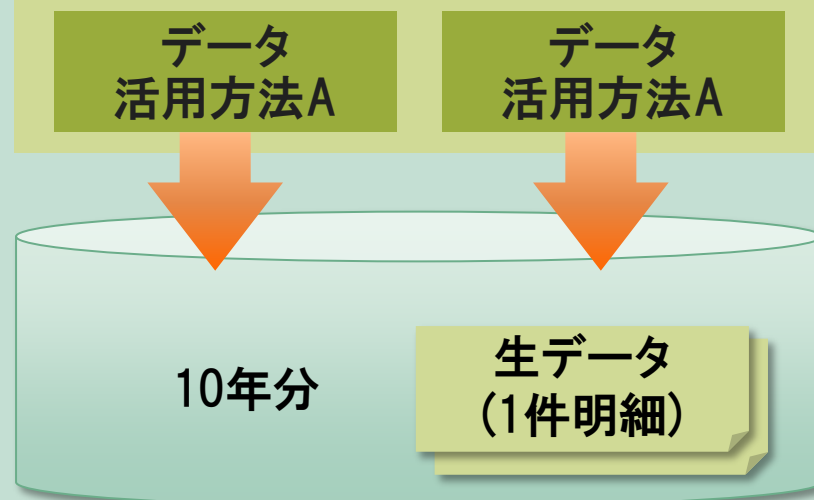


企業活動の基本データ

今後

- プラス α の使い方として、制約を排し、データ活用方法を後で容易に追加できる

試行錯誤しながら追加



付加価値を生むデータ

② 蓄積・検索

Hitachi Advanced Data Binder^{※1} プラットフォーム

● 高速なデータアクセス

「非順序型実行原理」^{※2}に基づき、従来比最大約100倍^{※3}のデータ検索性能を発揮する超高速データベースエンジンを採用。ビッグデータの高速処理を求める市場ニーズに対応。

● シンプルな構成でのシステム構築が可能

シンプルなシングルサーバ構成でありながら、サーバのマルチコアプロセッサ^{※4}およびストレージシステムの利用効率を最大限に高めることで、処理性能の大幅な向上を実現。

● ベストプラクティスモデルにより導入が容易

超高速データベースエンジンと、高信頼・高性能な日立サーバおよび日立ストレージとを組み合わせたベストプラクティスモデルを高速データアクセス基盤として製品化し提供。

高速

従来型データベースエンジン

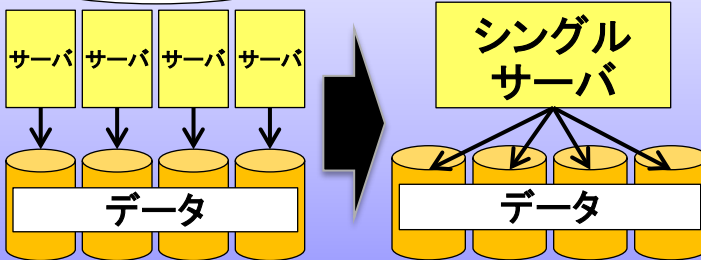
検索時間

最大100倍

超高速データベースエンジン

シンプル

並列RDB



ベストプラクティスモデル

超高速データベースエンジン

Red Hat Enterprise Linux

HA8000/RS440 (サーバ)

BR1200 (ストレージ)

超高速データベースエンジンには内閣府が創設した最先端研究開発支援プログラムで採択された「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(略)」(東大、日立)で技術開発された成果が反映されています。

※1: 内閣府の最先端研究開発支援プロジェクト「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: 東大喜連川教授)の成果を利用。

※2: 喜連川教授・合田特任准教授(東京大学)が考案した

※3: 解析系データベースに関する標準的なベンチマークを元に作成した、各種のデータ解析要求の実行性能を計測。データ解析要求の種類によって高速化率に差は見られるが、データベースにおいて特定の条件を満たす一定量のデータを絞り込んで解析を行うデータ解析要求を対象とした結果。

※4: 多数のコアを集積したプロセッサ。

- ビッグデータの処理では、並列での全件検索が求められる
- 全件検索をすることで、データから価値を探し出す

ATM, POS, 証券取引など
金融系業務にみられる
即時性の求められる処理

より情報系に

特定url検索業務、ドリルダウン分析、
仮説検証などの試行錯誤を伴う検索

【条件抽出】
(得意領域)

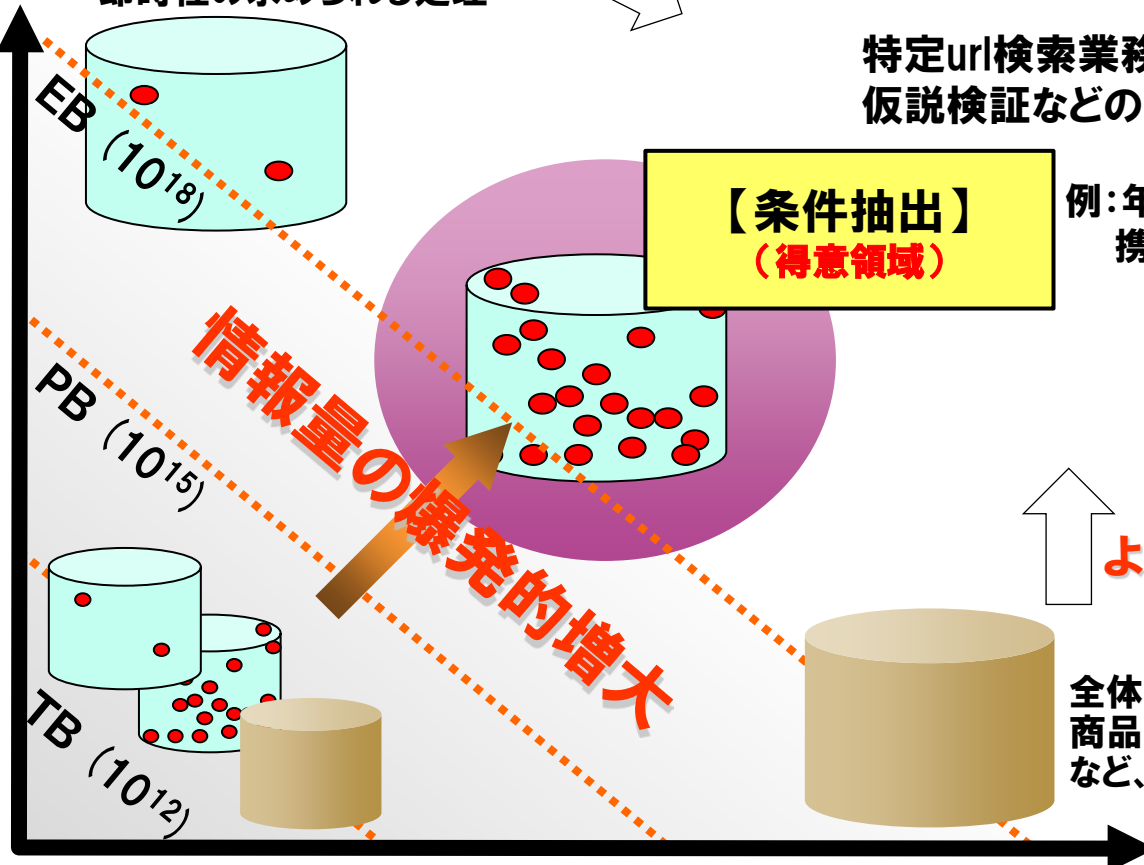
例:年代別(20代)×場所別(新宿)で
携帯電話の使用状況

情報量の爆発的増大

よりリアルタイムに

全体統計分析、ログ全体管理、
商品売筋分析、障害や動向の調査
など、バッチ型の処理

問合せ処理
の実行頻度
(1件処理)

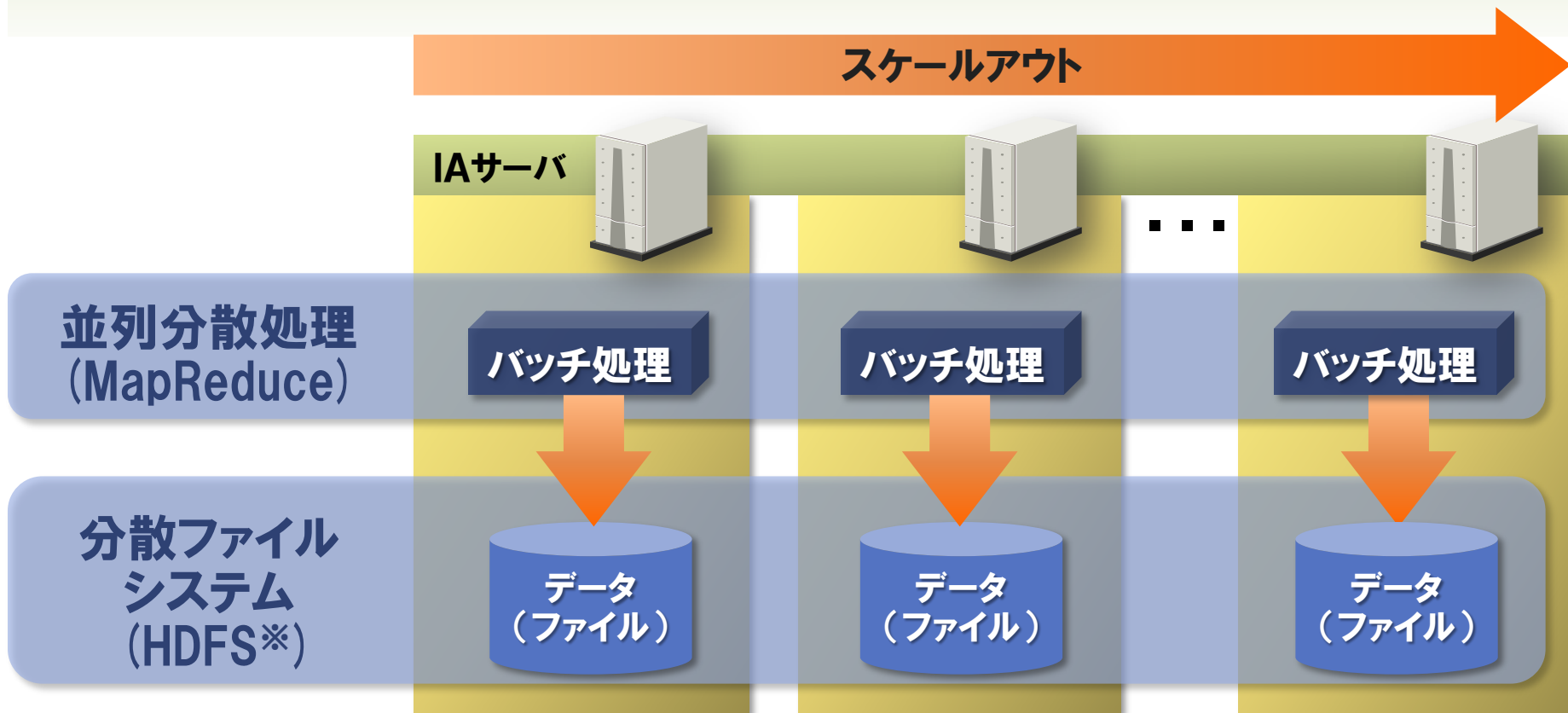


※1:内閣府の最先端研究開発支援プロジェクト「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的サービスの実証・評価」(中心研究者:東大喜連川教授)の成果を利用。

③ 集計・分析

Hadoop

- Googleの公開論文を参考にして開発されたオープンソース
- 安価なIAサーバを並べることで、バッチ処理を高速化
- 並列分散処理フレームワーク、分散ファイルシステムで構成

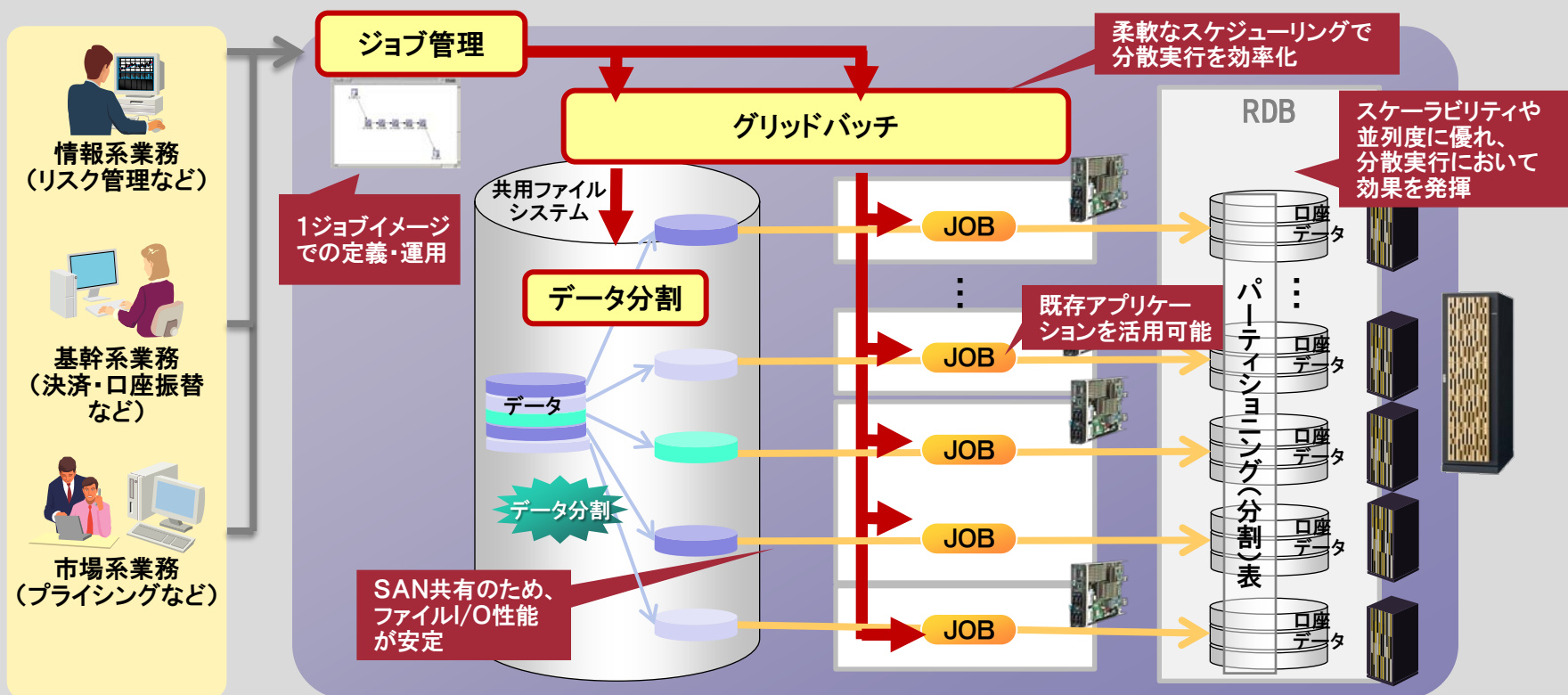


※HDFS:Hadoop Distributed File System

③ 集計・分析

Cosminexus
uCosminexus Grid
Processing Server

- 既存資産(COBOL・Java等)を活用した並列処理技術
- 並列処理を一元管理でき運用が容易
- データ量に応じてスケールアウト&障害箇所の局所化



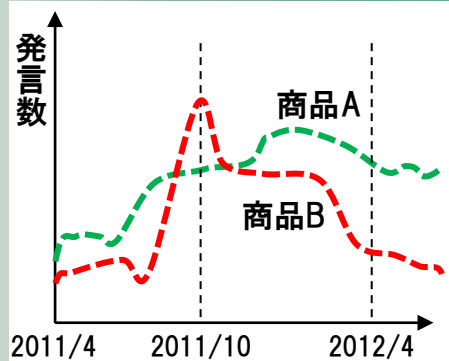
ビッグデータを利活用するためのITプラットフォームとは

4. ビッグデータ利活用事例

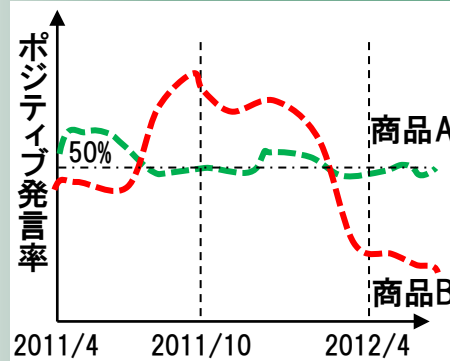
ソーシャルメディアデータによる特定キーワードの傾向分析



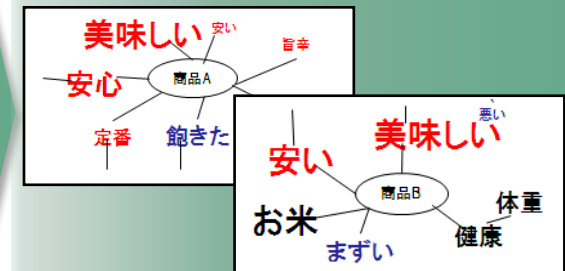
頻度分析処理
(特定キーワードの頻度分析)



感情分析処理
(特定キーワードのネガポジ分析)



グラフ分析処理
(特定キーワードのグラフ分析※)



※赤字はポジティブ、青字はネガティブ、黒字は中立、文字の大きさは発言数を示す

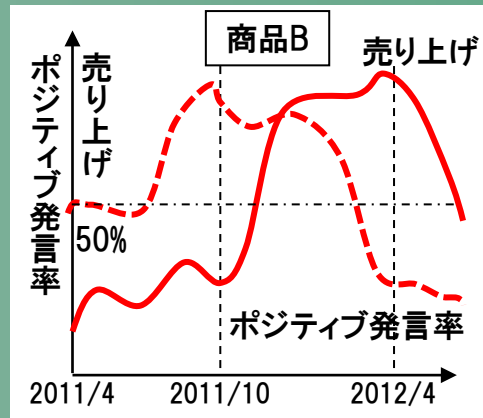
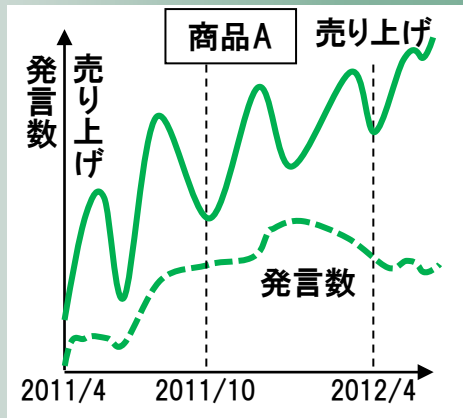
キーワードの選択は困難 → POSの特徴データから対象キーワードを自動抽出による容易化

キーワードの出現頻度/ネガポジ意見/関係性に基づくトレンド商品抽出

ソーシャルメディア×POSデータによる需要予測の補正

業務データに基づくトレンド分析結果および特徴量/キーワード

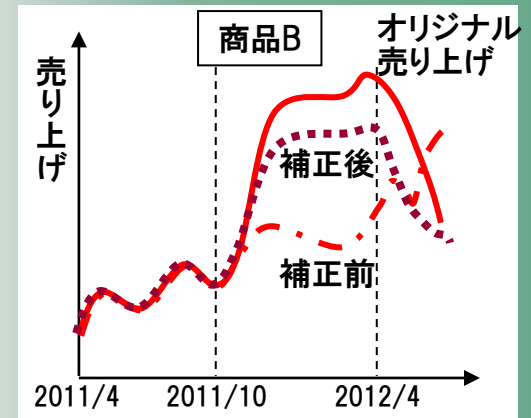
POS/ソーシャル相関分析処理



- ・商品Aは発言数の変化少 → 需要予測が容易
- ・商品Bはポジティブ発言増加後販売増。その後ネガティブ発言増加と共に販売少

需要予測モデルとの誤差の原因分析

需要予測補正および詳細化

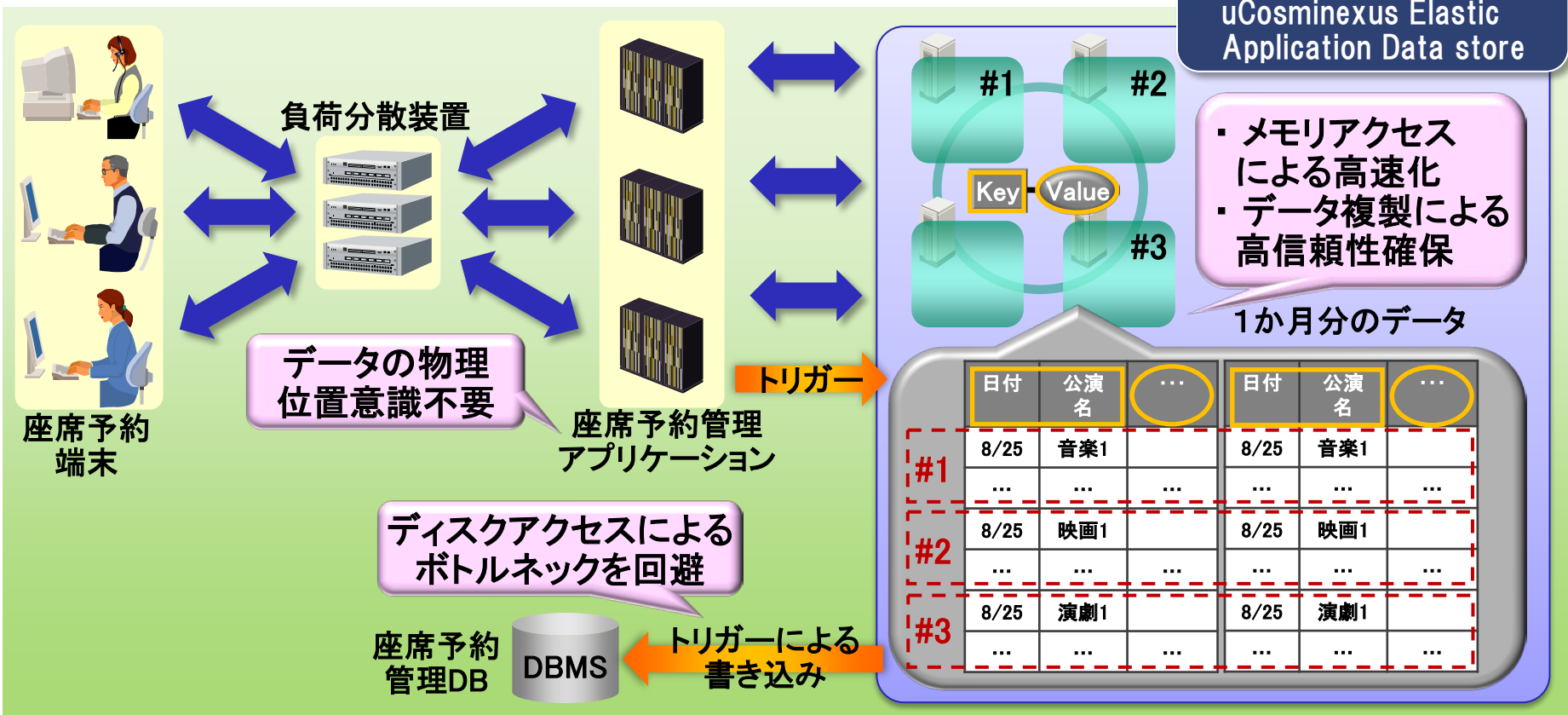


POS特徴データとソーシャルメディア分析の多変量分析による需要予測補正

ソーシャルメディアデータに基づくトレンド分析結果およびキーワード出現頻度/ネガポジ意見/関係性

座席予約システムの処理高速化

- 対象座席データのメモリ保持によるフロント側の高速処理
- メモリ空間の統合管理によるスケールアウトでの処理性能確保
- データの冗長配置によるデータ損失の防止



メールシステムの処理高速化

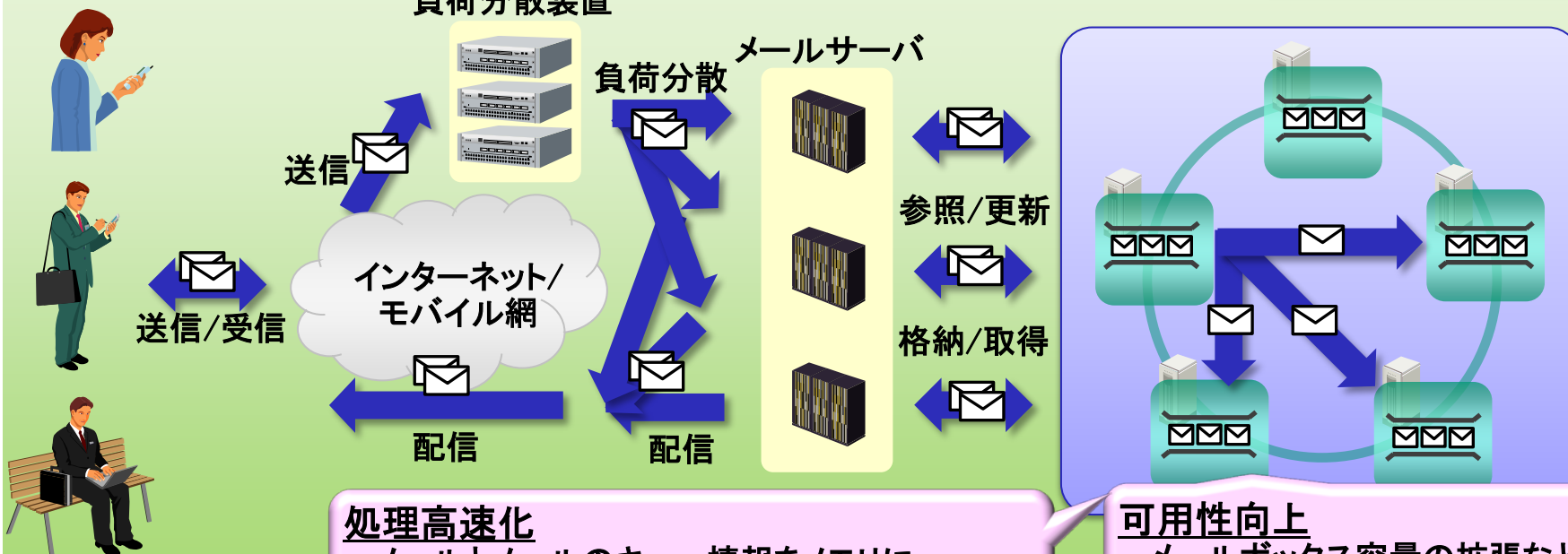
- 送受信処理高速化(1,000通以上/秒、平均レスポンスタイム10m秒以下)
- メモリ空間の統合管理によりスケールアウトでの可用性向上
- 障害発生時にメールシステムへの影響を局所化

Cosminexus
uCosminexus Elastic
Application Data store

携帯端末

負荷分散装置

メールサーバ



処理高速化

メールとメールのキュー情報をメモリにキャッシュすることで処理性能を飛躍的に向上

可用性向上

メールボックス容量の拡張など、スケールアウトで柔軟に対応可能

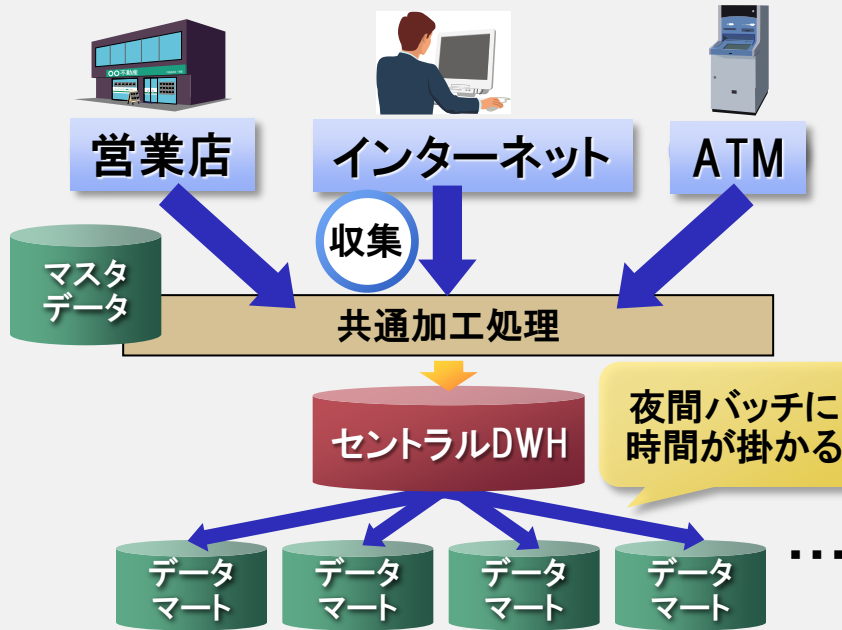
4-4 データマートレスによる高速分析（1）

課題

- データ量の増加で夜間バッチが朝までギリギリ...
- 本当はもっと分析軸を増やしたいが
あきらめている...



現状 業務データを収集し、データマートを作成。ただしバッチに時間が掛かっており、もっと分析軸を増やしたり、1件明細まで見たいが、あきらめている。

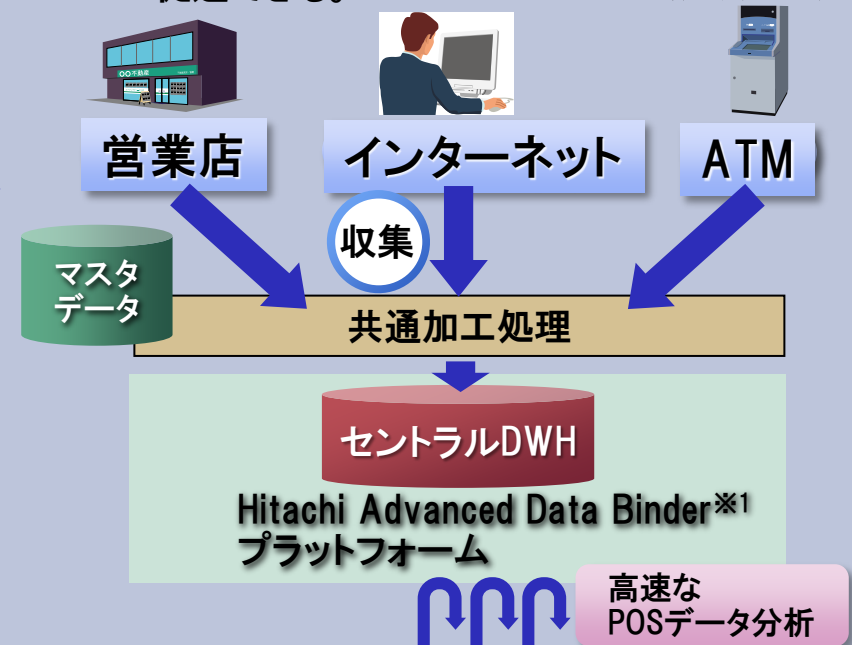


解決策

- 高速データアクセス基盤によって、データマートを削減することが可能に



今後 セントラルなDWH※だけで高速検索ができる。バッチ処理削減による運用コスト削減、分析軸の追加などデータの利活用を促進できる。
※ Data Warehouse(データウェアハウス)



これらの技術には内閣府が創設した最先端研究開発支援プログラムで採択された「最高速データベースエンジンの開発(略)」(東大、日立)で技術開発された成果が反映されています。

※1:内閣府の最先端研究開発支援プロジェクト「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:東大喜連川教授)の成果を利用。

4-4 データマートレスによる高速分析（2）

いままであきらめていたデータ活用が可能に（食品・小売業を例に）

「日別集計」から「時間帯別集計」の拡大

現状 店舗別、地域別、全店の日別集計に留まる

商品A	10月1日	10月2日	10月3日	10月4日	10月5日
	日	月	火	水	木
売上					
値引					

今後 時間帯別集計が可能。品切れや余剰を防ぎ、食品の作る/見切るタイミングを調整できる

商品A		9時	10時	...	20時
10月1日	単価				
	販売個数		品切れ	時間軸で見える	
	在庫		0		
10月2日	単価				
	販売個数				余剰
	在庫		項目が見える		5

顧客の併売分析の実現(バスケット、買い回り)

今後 何らかのターゲット商品を決めて、顧客の併売分析をすることで、セット商品化、プロモーション、店舗改善などに活かせる

レシートa
顧客番号XXXX

x社麻婆豆腐の素
y社豆腐豚挽き肉
...
z社ザーサイ

レシートb
顧客番号YYYY

x社麻婆豆腐の素
中華麺
ナス
...

レシートc
顧客番号ZZZZ

x社麻婆豆腐の素
パスタ
キャベツ
...

麻婆豆腐の付け合わせとして何が購買されているか？

「麻婆豆腐の素」の新しい使い方が流行っていないか？（麻婆麺、麻婆パスタ）

レシートx
顧客番号XXXX

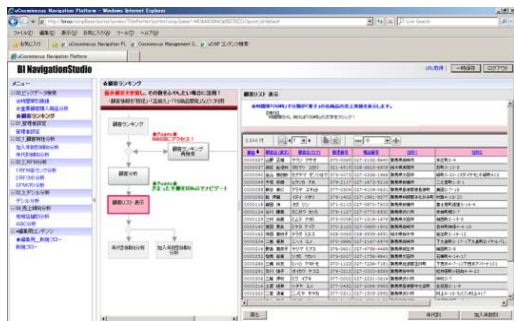
x社ランドセル
...

レシートy
顧客番号XXXX

y社羊羹
...

同一顧客の複数レシートを組み合わせる判断

ランドセルと一緒に購買されているものは？（祖父・祖母へのお返しもあるのか？）



ビッグデータを高速に
集計・分析できます！

Webポータル

集計・分析

BI NavigationStudio

業務データをフローやナビゲーションなどの画面表示に沿って容易に分析できる。
BI利用に関する専門知識の無い現場部門のユーザも容易に利用でき、全社でのデータ利活用を実現できる。

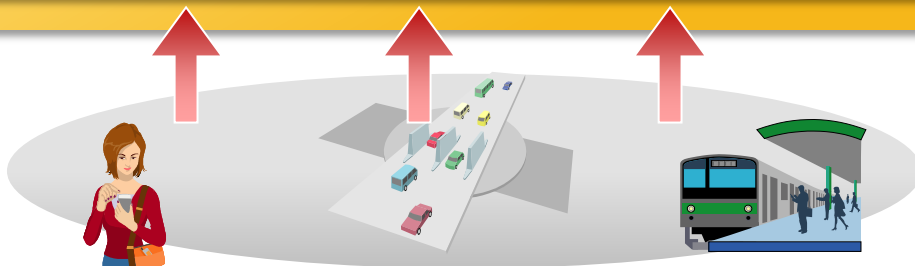
Hitachi Advanced Data Binder※1
プラットフォーム BIアプライアンス

蓄積・検索

Hitachi Advanced Data Binder※1 プラットフォーム



ビッグデータ



大賞を受賞



※1: 内閣府の最先端研究開発支援プロジェクト「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社会サービスの実証・評価」(中心研究者: 東大喜連川教授)の成果を利用。

本製品は(株)DTSが販売します

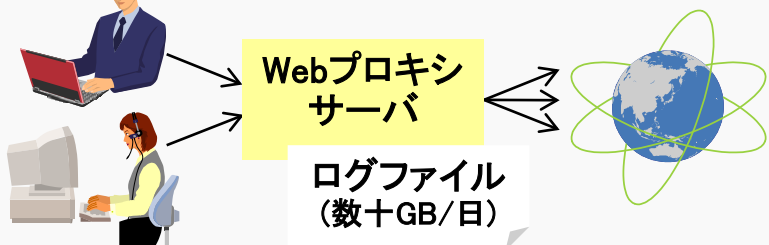
課題

- セキュリティを緩和し、従業員の利便性を高めたい...
- ただし逆に監視は強化したいが、データ量が膨大...



【現状】従来はWeb閲覧を制限してきたが、フィルタを緩和し、利便性を高めたい。逆に監視は強化したいが、システムが追いついていない。

従業員 インターネット



レポート作成に
数十時間かかる

オンデマンドな
検索は不可

レポート



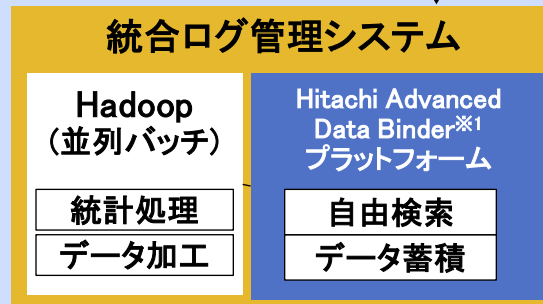
解決策

- 高速データアクセス基盤×Hadoopを用いたデータ蓄積&高速検索&高速バッチ



【今後】今まであきらめていた分析ができ、レポートの種類を増やせる。さらに他のログも併せてコンプライアンスを強化できる。

ログファイル
(数十GB/日)



他のログも
併せて活用

社内メール

社内SNS

入退管理

短時間で生成
(数十時間⇒数十分)

高速な
オンデマンド検索

レポート



結果

- 高速バッチ処理で、レポートの種類を増やせる！
- あきらめていたオンデマンド検索ができる！
- 他のログも併せ、コンプライアンスを強化できる！



評価例

- 特定ユーザの検索
(130GBのデータ量)

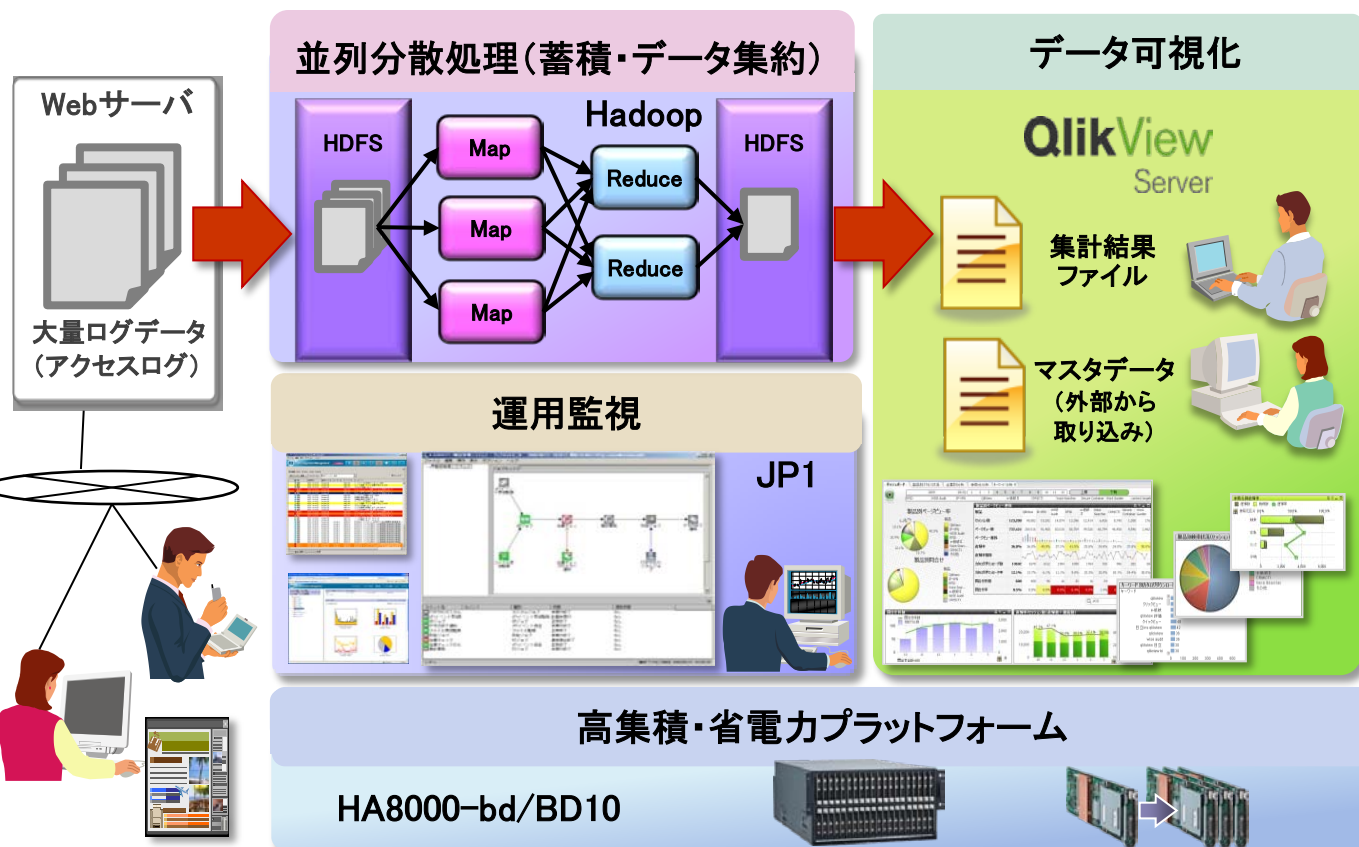
手作業で6時間⇒8秒まで短縮



※1:内閣府の最先端研究開発支援プロジェクト「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:東大喜連川教授)の成果を利用。

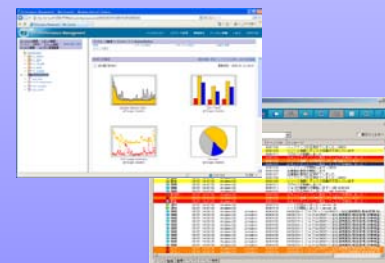
かんたんHadoopソリューション for ログ解析 (QlikView & JP1)

- アクセス集計結果と顧客マスタなどを「QlikView」上で突合せ可視化
- 多ノード構成で複雑になりがちなHadoopシステムの運用を、「JP1」で支援



■ JP1連携による 運用監視の容易化

- (1) Hadoop単体の運用の弱さをJP1で補完
 - (2) 多ノードの運用管理もJP1連携で容易化
- ① バッチジョブの管理
 - ② ログの一元管理
 - ③ 性能や障害状況のモニタリングなど

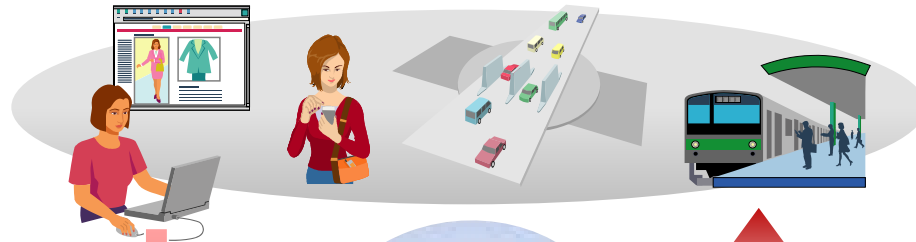


ビッグデータを利活用するためのITプラットフォームとは

5. まとめ

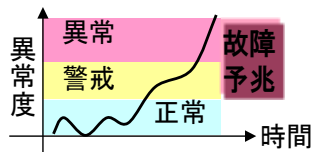
5-1 ビッグデータ利活用プラットフォーム

ビッグデータ



① リアルタイム監視

大量の情報を
リアルタイム処理



Cosminexus

uCosminexus Stream
Data Platform

Cosminexus

uCosminexus Elastic
Application Data store

拡張性

高信頼ブレードサーバ
BladeSymphony/HA8000
仮想ファイルプラットフォーム
Hitachi Virtual File Platform

② 蓄積・検索

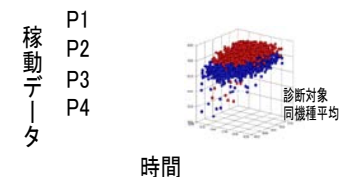
大量の情報を
高速検索・効率保管

履歴データ

データセンター

③ 集計・分析

大量の情報を
対象とした分析



Cosminexus

uCosminexus Grid
Processing Server

Hadoop

Hitachi Advanced Data
Binder※1 プラットフォーム

※1: 内閣府の最先端研究開発支援プロジェクト「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: 東大喜連川教授)の成果を利用。

No.	ニーズ	新技術・製品	概要
1	リアルタイムなデータ活用	ストリームデータ処理基盤 (uCosminexus Stream Data Platform)	<ul style="list-style-type: none"> ◆インメモリ処理と差分計算処理による超高速処理 ◆SQLライクなスクリプト言語(CQL)で分析シナリオを記述可能な開発容易性
2	インメモリデータ量の増加に対して柔軟性や拡張性が欲しい	インメモリデータグリッド (uCosminexus Elastic Application Data store)	<ul style="list-style-type: none"> ◆Java・C言語基盤におけるインメモリキャッシュによるデータアクセス性能(特に参照)高速化を実現 ◆複数のノードに大量データを分散配置し、仮想化してアプリケーションに提供
3	大量データの蓄積、超高速検索	高速データアクセス基盤 (Hitachi Advanced Data Binder ^{※1} プラットフォーム)	<ul style="list-style-type: none"> ◆「非順序型実行原理」^{※2}に基づく超高速データベースエンジンを採用 ◆サーバのマルチコアプロセッサおよびストレージシステムの利用効率を最大限に高めることで、処理性能の大幅な向上 ◆超高速データベースエンジンと、高信頼・高性能な日立サーバおよび日立ストレージとを組合せたベストプラクティスモデルによる導入容易性

※1:内閣府の最先端研究開発支援プロジェクト「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:東大喜連川教授)の成果を利用。

※2:喜連川教授・合田特任准教授(東京大学)が考案した。

No.	ニーズ	新技術・製品	概要
4	バッチ時間の短縮、 バッチ終了時間の遅延防止	グリッドバッチ (uCosminexus Grid Processing Server)	<ul style="list-style-type: none"> ◆データを分割配置しバッチ処理の並列実行により、バッチ業務を高速化 ◆基幹系にも対応出来る高い可用性と、障害の局所化
5	大量データを用いた 新サービスの立上げ	Hadoop (オープンソース)	<ul style="list-style-type: none"> ◆大量ログデータを用いた分析処理の高速化、並列処理を隠蔽した開発容易性 ◆将来性あるOSSであり、世界中で企業システム適用が模索されはじめている ◆日立は「Hadoop MapReduceサポートサービス」を提供 ◆高集積・省電力プラットフォームHA8000-bd/BD10との引合いが多い(環境設定のサービス有り)

※3:OSS(Open Source Software)



■ ストリームデータ処理基盤の利活用

uCosminexus Stream
Data Platform

- 大量に発生する時系列データをリアルタイムに分析して、
即座にアクションする必要があるお客さま

スマートメータなど実世界のさまざまな情報(位置データ、センサデータ等)
を入力源とした監視業務系システムなど

■ インメモリデータグリッドの利活用

uCosminexus Elastic
Application Data store

- 大量データの高速処理で高い信頼性が必要なお客さま

座席照会や証券取引やコールセンタなど、
参照系メインの定型業務システム

■ 高速データアクセス基盤の利活用

Hitachi Advanced Data
Binder^{※1} プラットフォーム

- 大量データの検索・分析業務で高速化が必要なお客さま

特定検索業務、ドリルダウン分析、仮説検証などの試行錯誤を伴う検索を
RDBではあきらめていたり、分析のためのデータマート作成で悩んでいる

※1:内閣府の最先端研究開発支援プロジェクト「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:東大喜連川教授)の成果を利用。

■ グリッドバッチの利活用

UCosminexus Grid
Processing Server

- **高い信頼性が要求される基幹系システムのオープン化**
確実かつ正確な数値計算が要求される
- **業務継続性が重視されるシステムの新規開発**
データの量や更新頻度に応じたスケーラビリティ(柔軟性)が必要
- **業務拡大により夜間バッチの処理限界が予想されるお客さま**
現在バッチ処理に夜間所要時間ぎりぎり、今後業務データが2倍になると夜間での処理は不可能

■ Hadoopの利活用

- **ログ分析等、データ傾向を把握する情報系システム**
リアルタイムでの大量データの傾向分析が必要

ビッグデータはまだまだ発展途上の分野・・・
ただし“**新たな価値創造を起こす**”と期待されている
日立の豊富な事例と経験



イノベイティブ・アナリティクスの実践



ビッグデータ利活用プラットフォームの提供

日立は、お客さまと一緒に
ビッグデータ利活用に取り組みます

<http://www.hitachi.co.jp/bigdata/>

他社商品名、商標等の引用に関する表示

- BI NavigationStudioは、株式会社DTSの日本における登録商標です。
 - Facebookは、Facebook, Inc.の登録商標です。
 - Googleは、Google Incの登録商標です。
 - OracleとJavaは、Oracle Corporation 及びその子会社、関連会社の米国及びその他の国における登録商標です。
 - Red Hatは、米国およびその他の国でRed Hat, Inc. の登録商標もしくは商標です。
 - Linuxは、Linus Torvalds氏の日本およびその他の国における登録商標または商標です。
 - QlikViewは、QlikTech International AB.の登録商標です。
- その他記載の会社名、製品名はそれぞれの会社の商標もしくは登録商標です。

■「ビッグデータ利活用」に関するホームページ

<http://www.hitachi.co.jp/bigdata/>

■「Field to Future Technology」ホームページ

http://www.hitachi.co.jp/soft/data_tech/

■uCosminexus Stream Data Platform

<http://www.hitachi.co.jp/soft/cosminexus/sdp/index.html>

■uCosminexus Grid Processing Server

<http://www.hitachi.co.jp/soft/cosminexus/gps/index.html>

■uCosminexus Elastic Application Data store

<http://www.hitachi.co.jp/soft/cosminexus/uceads/index.html>

■Hitachi Advanced Data Binder^{※1}プラットフォーム

http://www.hitachi.co.jp/soft/big_data/data-binder.html

■かんたんHadoopソリューション for ログ解析(QlikView & JP1)

<http://www.hitachi.co.jp/ha8000-bd/>

■BI NavigationStudio

<http://datastudio.jp/sol/binas.html>

※1:内閣府の最先端研究開発支援プロジェクト「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社会サービスの実証・評価」(中心研究者:東大喜連川教授)の成果を利用。

END

**ビッグデータを利活用するための
ITプラットフォームとは**

2012/11/9

株式会社 日立製作所 情報・通信システム社

ITプラットフォーム事業本部 開発統括本部 ソフトウェア本部 ビッグデータソリューション部

山口 俊朗