

COBOLコンソーシアムセミナー

ビッグデータへの挑戦

～基幹システムも融合して新しい価値の創造へ～

2013/04/19

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 ソフトウェア開発本部
ビッグデータソリューション部

山口 俊朗

Human Dreams.
Make IT Real.

ビッグデータへの挑戦

～基幹システムも融合して新しい価値の創造へ～

Contents

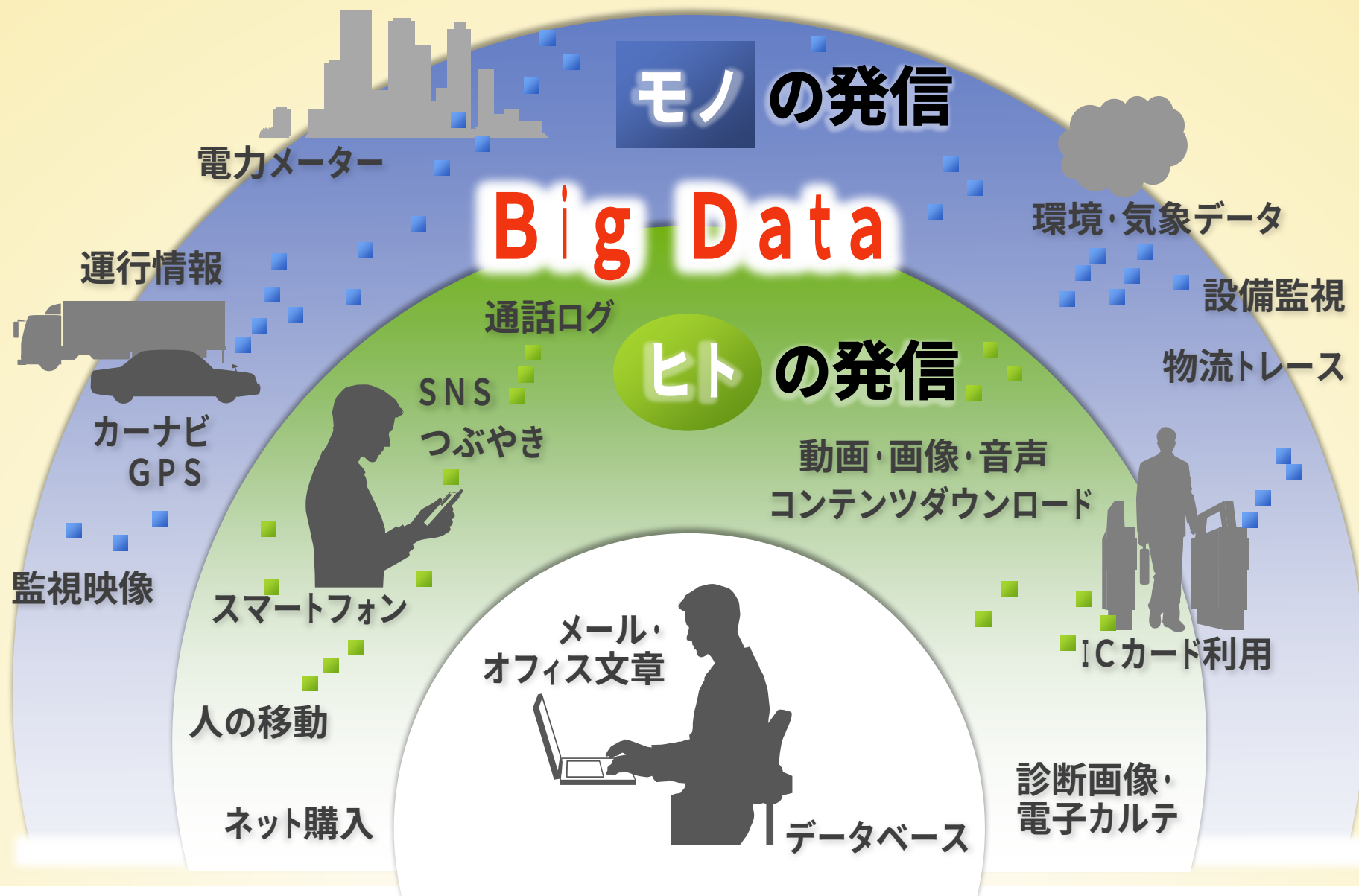
- 1. ビッグデータの時代がやってきた**
- 2. 日立のビッグデータ利活用事業**
- 3. ビッグデータの利活用**
- 4. まとめ**

付録

1-1 社会にあふれる有用なデータ

1. ビッグデータの時代がやってきた

HITACHI
Inspire the Next



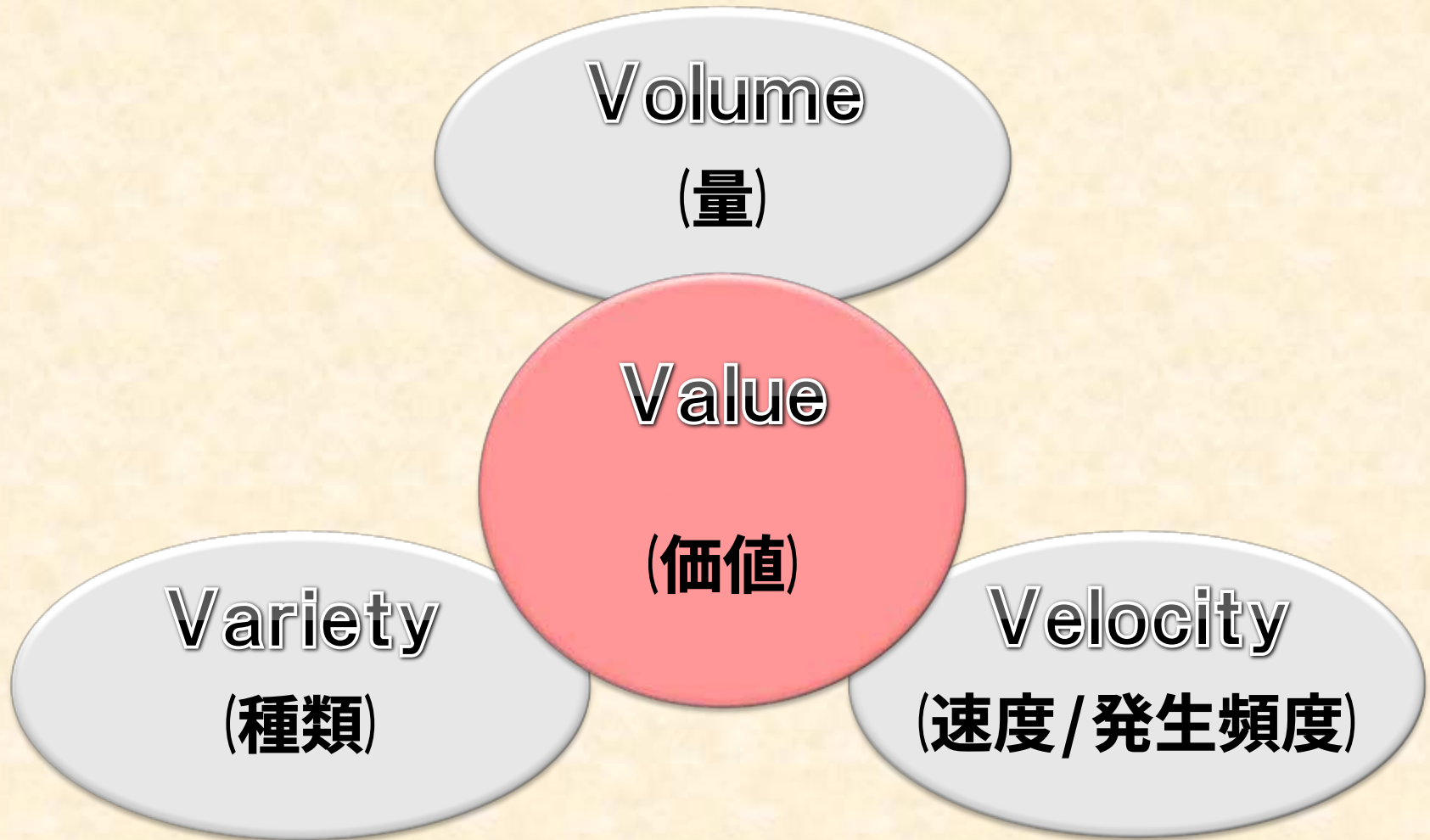
1-2 ビッグデータ利活用の課題



1-3 Big Data = 3V + V

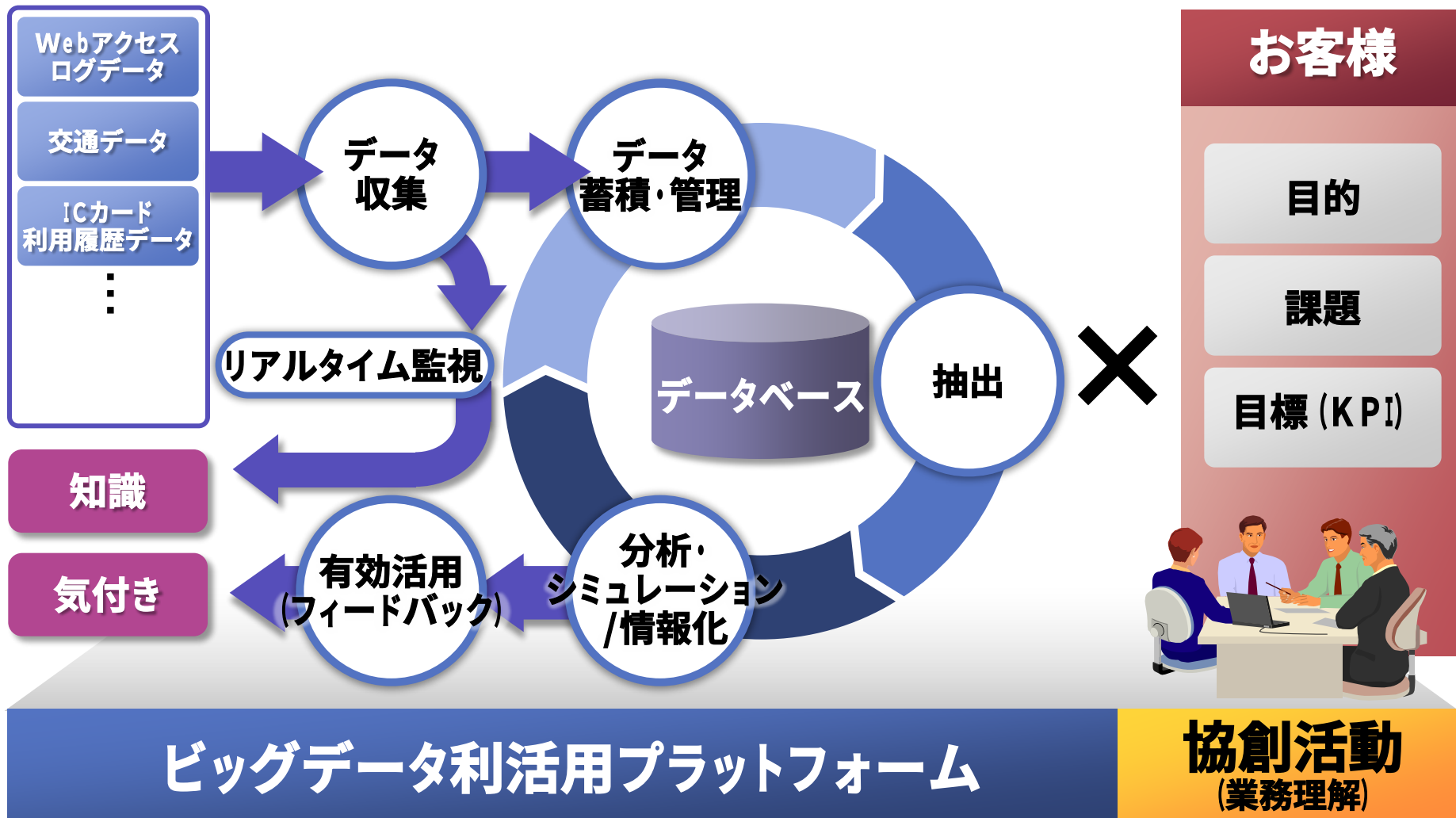
1. ビッグデータの時代がやってきた

HITACHI
Inspire the Next



2-1 日立のビッグデータ対応

- 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)





ビジネスナレッジ



現場力

お客様のビジネス



深い洞察

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



運行情報



人の移動



メールログ

Big Data



業務データ



通話ログ



天候・気象

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

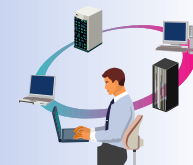


データ利活用専任組織

「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

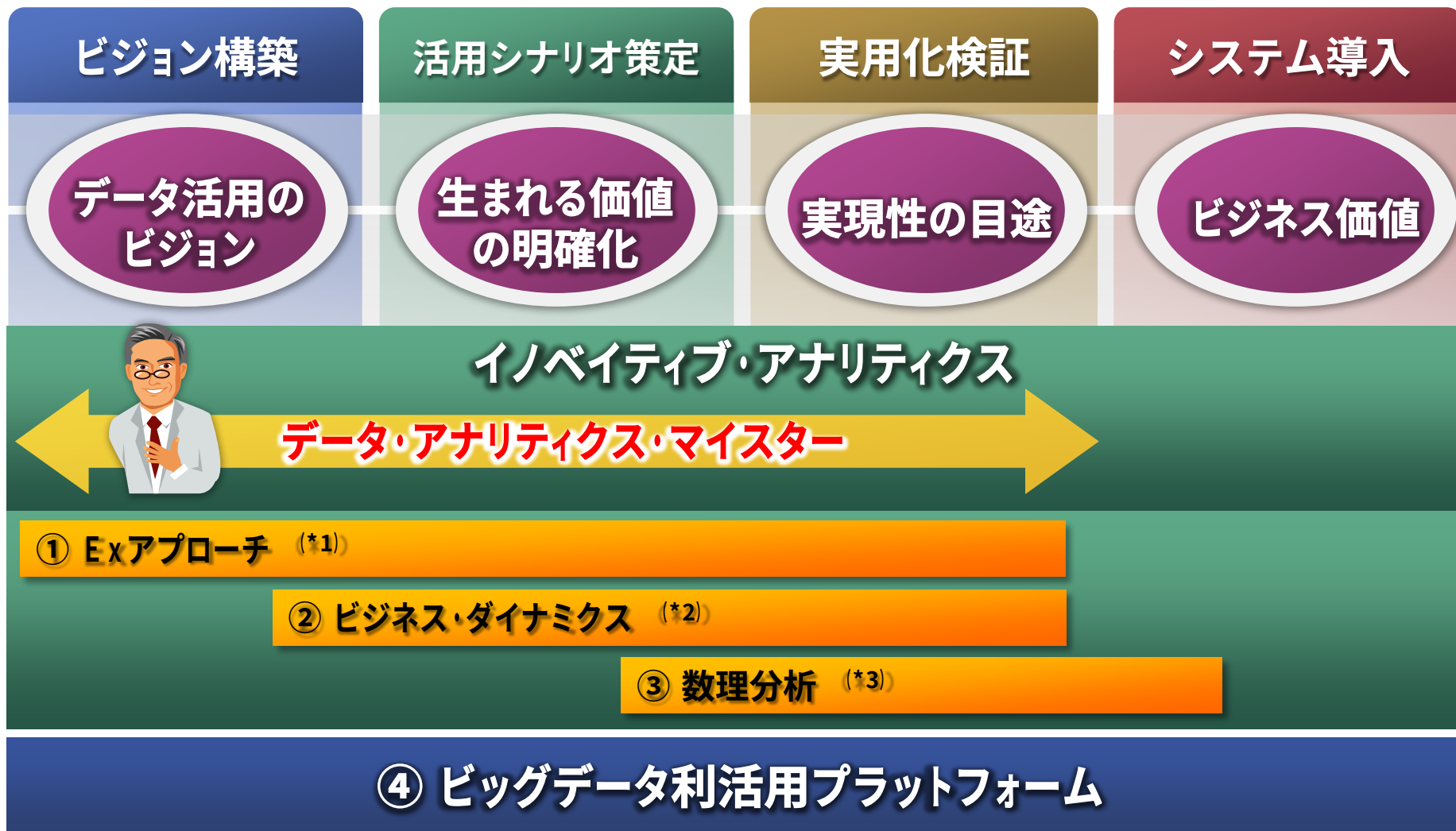
3 ビッグデータ利活用を支えるプラットフォーム製品、技術の開発



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 プラットフォーム※1」 提供開始

2-4 データを価値に換える4つの技術



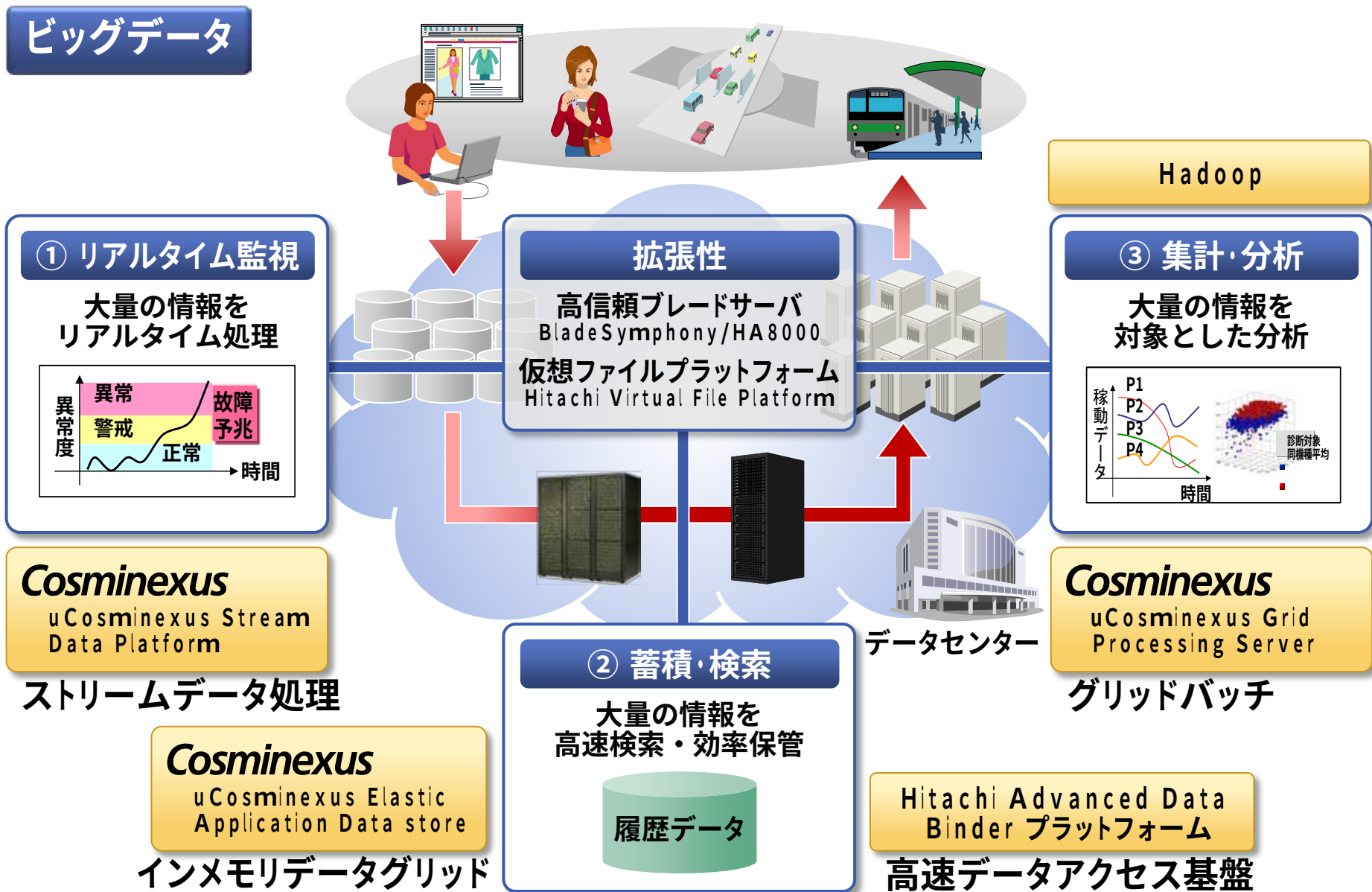
*1: お客さまの経験価値を最適化する観点で、現在の業務から将来像を創出協創によって、改善のための具体的なプロセス構築が可能。

*2: 事業構造をモデル化するMIT(Massachusetts Institute of Technology) 提唱の手法。日立は豊富な実績と、業種・業界毎の独自テンプレートを保有。

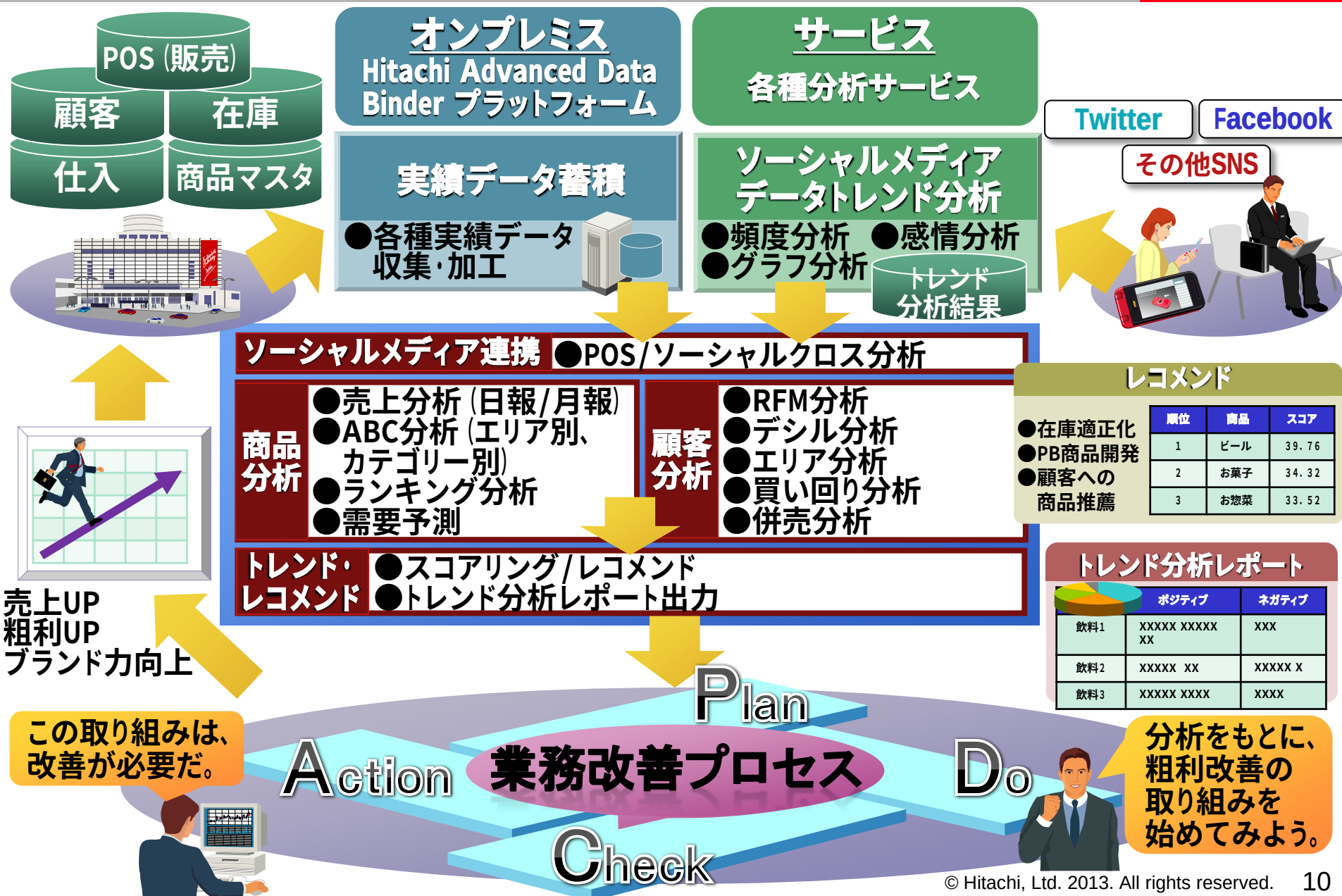
*3: 日立の豊富な分析技術・事例を体系的に整理した“分析テンプレート”(目的指向・手法指向・データ指向)で、あらゆる種類のデータの分析に対応。

2-5 ビッグデータ利活用プラットフォーム

ビッグデータ



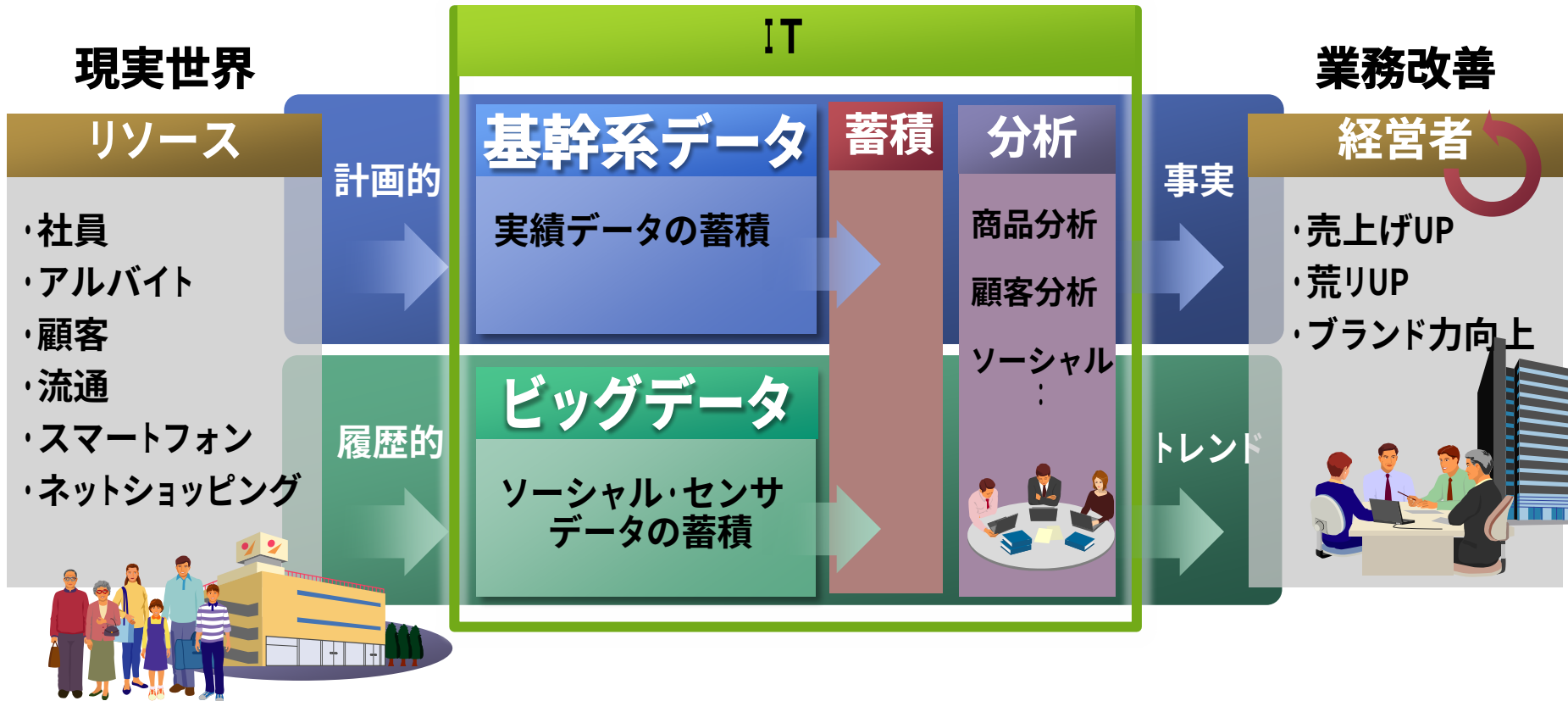
2-6 流通分析ソリューション



3-1 ビッグデータの利活用モデル

現行の基幹系システムを生かしたビッグデータの利活用

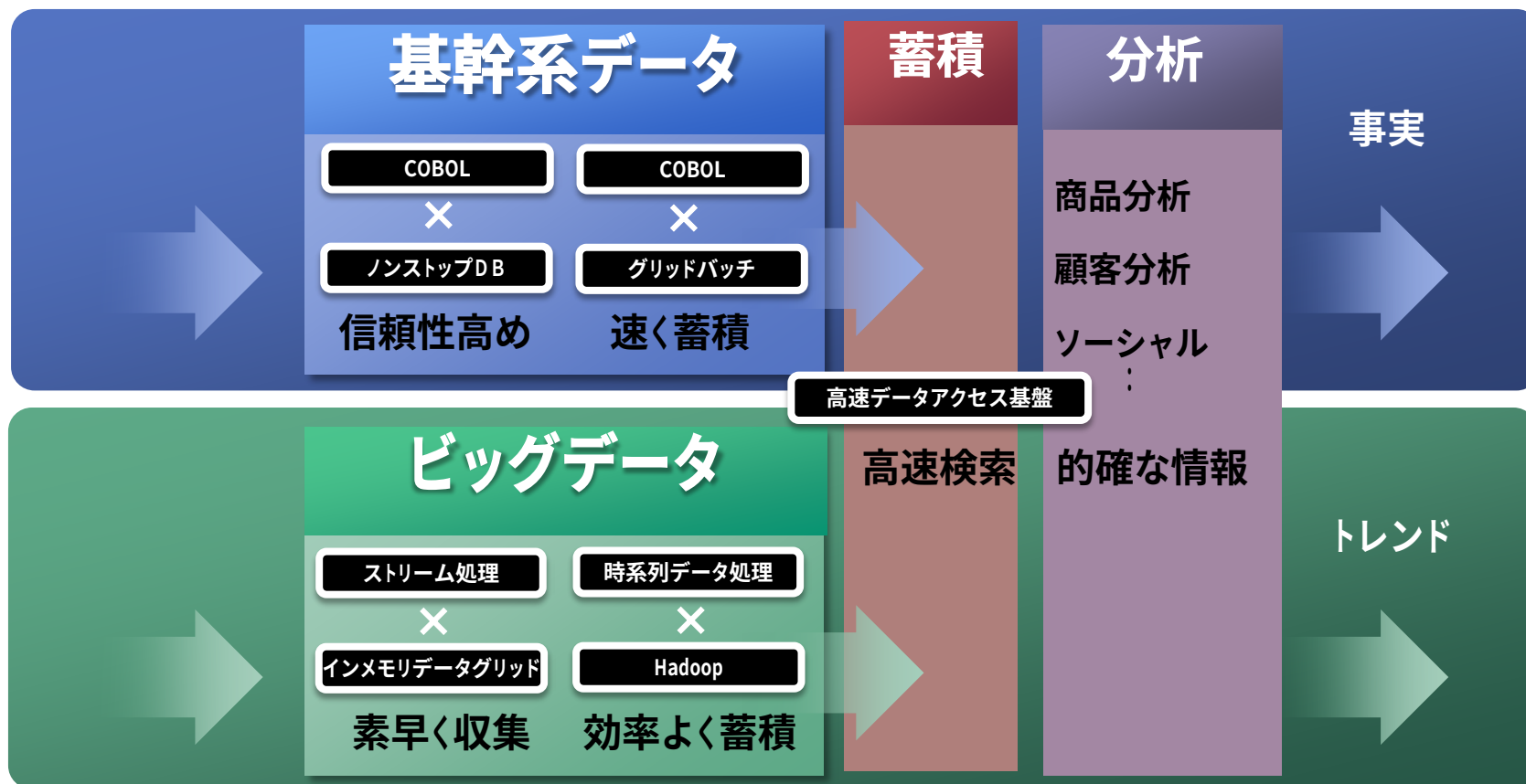
- 業務を改善するためのデータは基幹システムにもある
- さらに新しいデータを蓄積することで色々な分析が可能となる



3-2 ビッグデータ利活用プラットフォームの適用

適材適所で利活用するための製品を提供

- 基幹系のデータはCOBOLで作っているなら、COBOLを更に有効活用
- 新しいデータ処理を付け加えて行くことで、容易にITを拡張



3-3 ビッグデータ利活用へのアプローチ

今あるもの(製品)からスタートして、大きく育てるビッグデータ

仮説

- データが2倍に増加しても、価値は2倍まで増えないこともある。
2倍の投資はリスクがあり、既存資産の多いCOBOLを使うことも検討すべき。
- Bを買ったら、Cも買う確率。これを解く“ベイズの確率”などは算数でも解ける。
マシンが速くなり解けるようになった今、COBOLで分析可能。



不要と判断したデータ
「価値を生まない？」

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{\sum_A P(B|A)P(A)}$$

「Bを買ったらCも買う」ことを知るために

- ・ 2つの商品の関係を知りたい
DBから検索する
- ・ 数千種の商品の組み合わせや、
総当りの計算・集計はバッチ処理が有利
- ・ 一度知った計算結果はいつでも再利用
(検索) 可能にしておき、価値を増やしていく

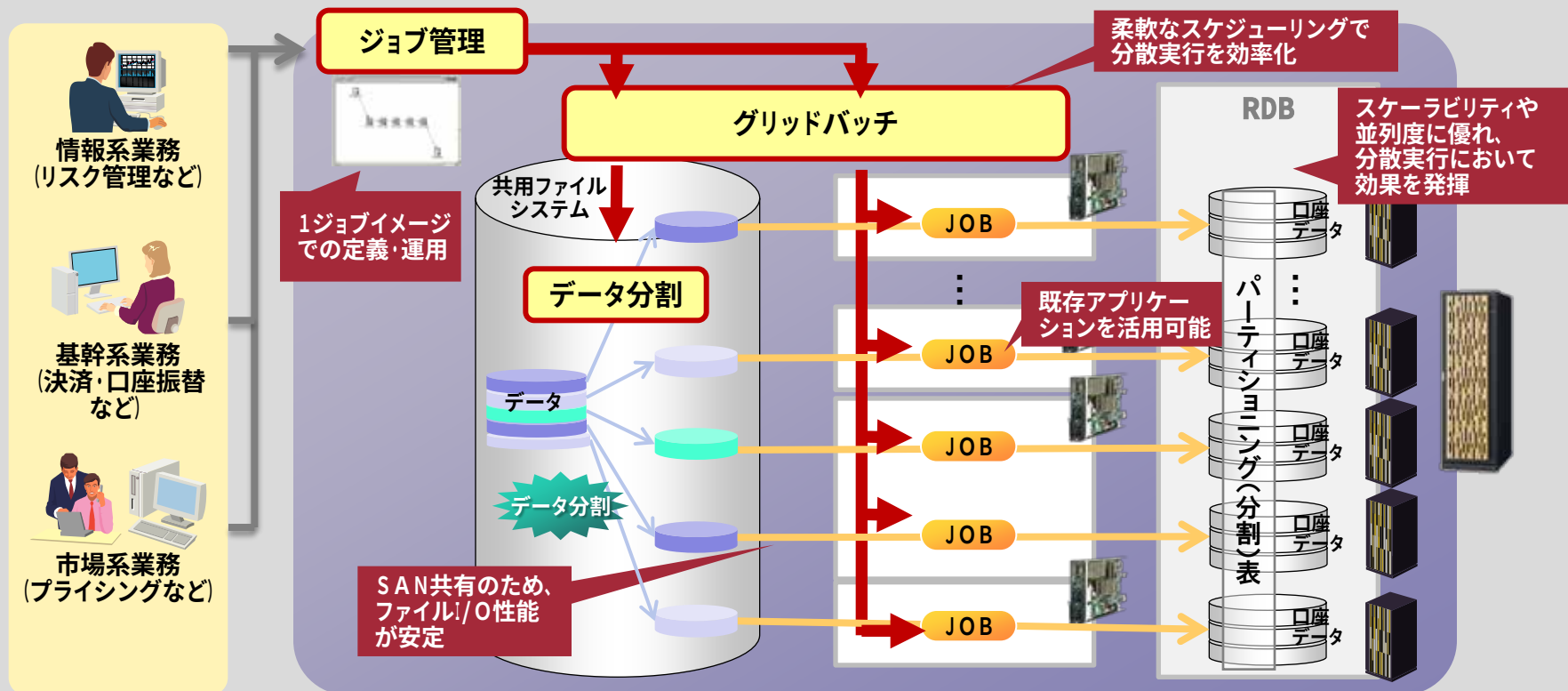
計算処理が多い場合は、COBOL処理を分散処理化

3-4 グリッドバッチ

③ 集計・分析

Cosminexus
 uCosminexus Grid
 Processing Server

- 既存資産 (COBOL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



3-5 グリッドバッチ適用事例 (ホストデータ活用)

ポイント

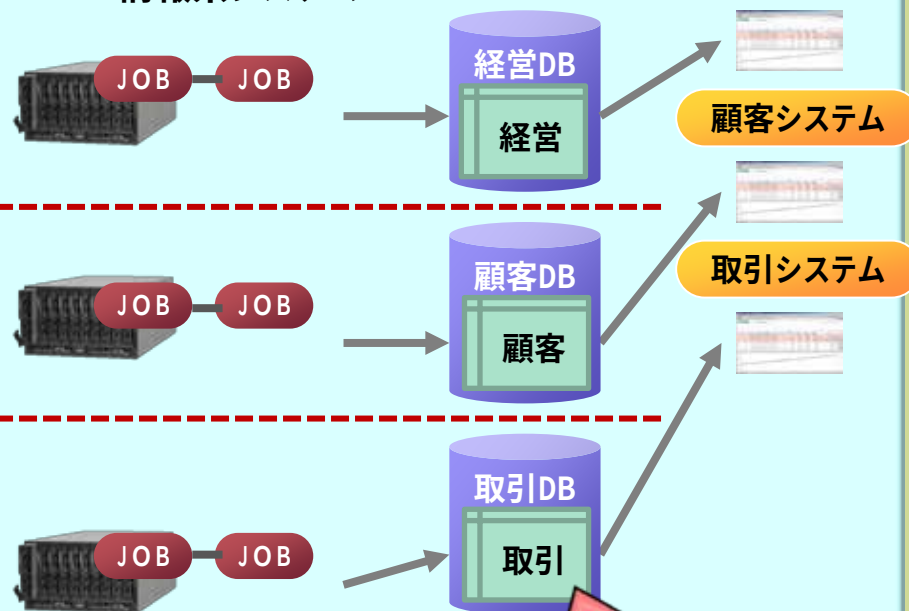
- メインフレーム上のジャーナルデータを、情報系システムへ集計する処理へグリッドバッチを適用し、**集計処理時間を短縮**
- 処理時間短縮により、従来目的別に分離していたDBの、**統合を実現**

適用前



JNL: ジャーナルデータ

一定時間間隔で配信
配信間隔内で集計処理
を完了する必要あり

オープンサーバ
情報系システム

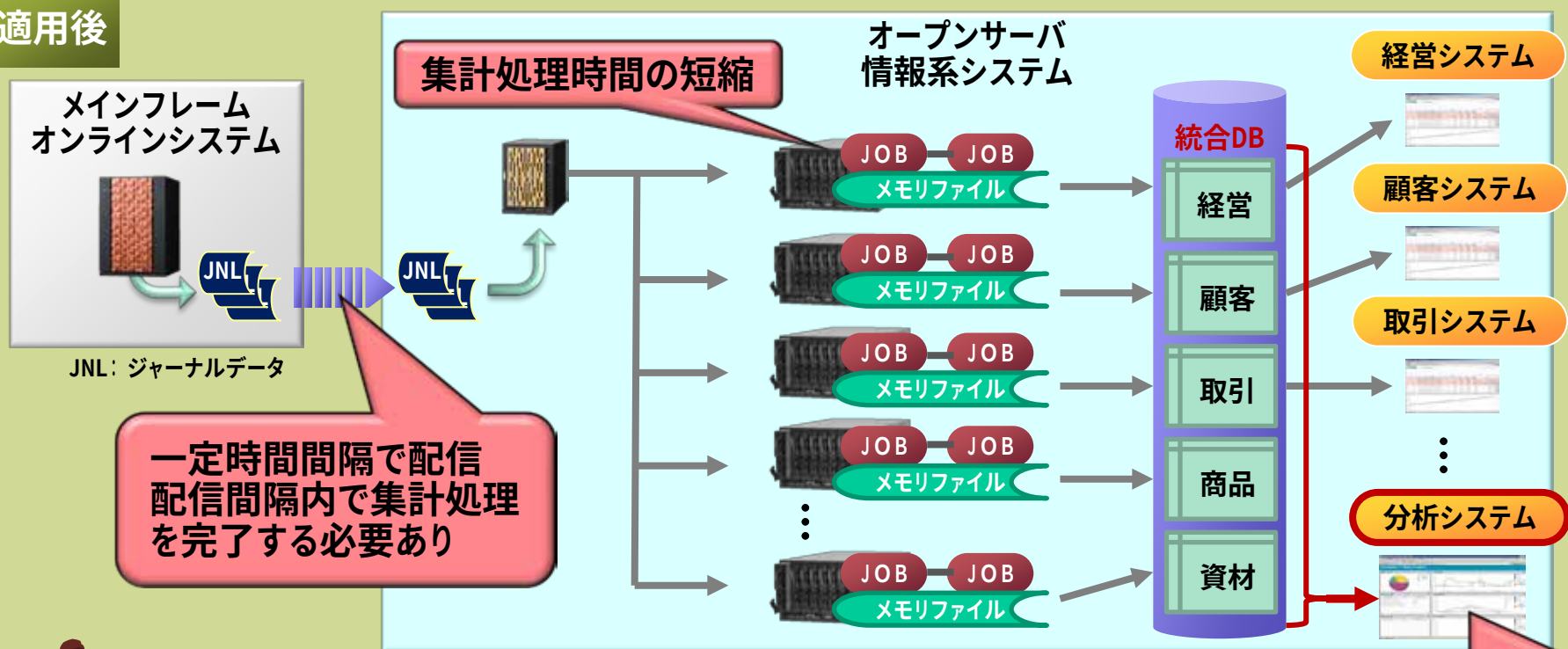
システム毎に個別に集計

3-5 グリッドバッチ適用事例 (ホストデータ活用)

ポイント

- メインフレーム上のジャーナルデータを、情報系システムへ集計する処理へグリッドバッチを適用し、**集計処理時間を短縮**
- 処理時間短縮により、従来目的別に分離していたDBの、**統合を実現**

適用後



お客様の導入メリット

- 集計処理時間の制約で**廃棄していたジャーナル中の情報の有効活用**
- 新情報活用と、DB統合化により、**多角的な分析が可能な新業務システムを構築**

DB統合による多角的
分析新業務システム
の構築

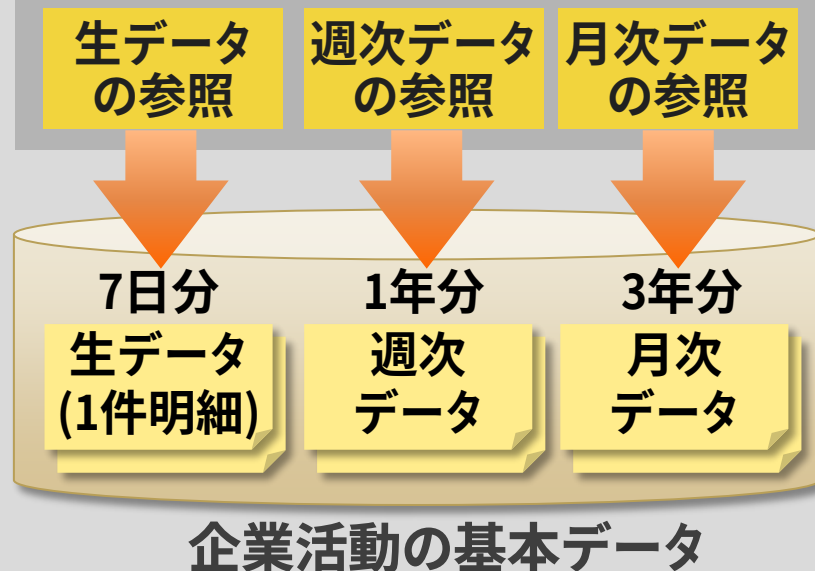
3-6 各種データベース技術

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、...)

これまでの“基幹系DB”

- 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

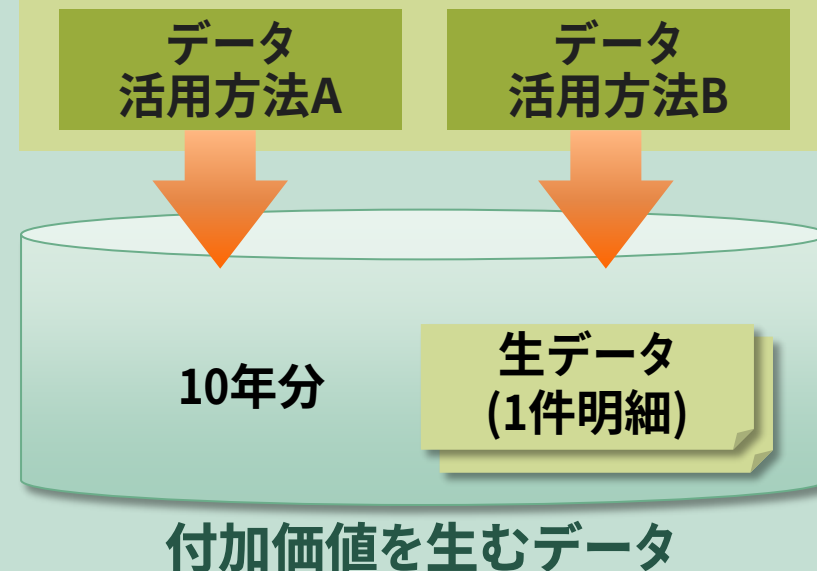
事前に決めた使い方



これからの“情報系DB”

- 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



3-7 高速データアクセス基盤とは

② 蓄積・検索

- 大規模データベースを対象としたデータ分析に求められる、高速なデータ検索を実現

- ・最先端研究開発支援プログラム^{(*)1}において、国立大学法人東京大学と共同で推進している超高速データベースエンジンの研究開発^{(*)2}成果を製品化
- ・超高速データベースエンジンの性能を最大限に発揮すべく、高信頼・高性能な日立ハードウェアと組合せた事前検証済みのベストプラクティスモデルとして ご提供

「Hitachi Advanced Data Binder プラットフォーム」

第一弾として、超高速データベースエンジンおよび、日立サーバと、データアクセス性能に優れたSSDを搭載した日立ストレージとを組み合わせた構成で提供開始。



Hitachi Advanced Data Binder プラットフォーム

超高速データベースエンジン

日立ラックサーバ 「HA8000」

日立ストレージ 「BR1200」

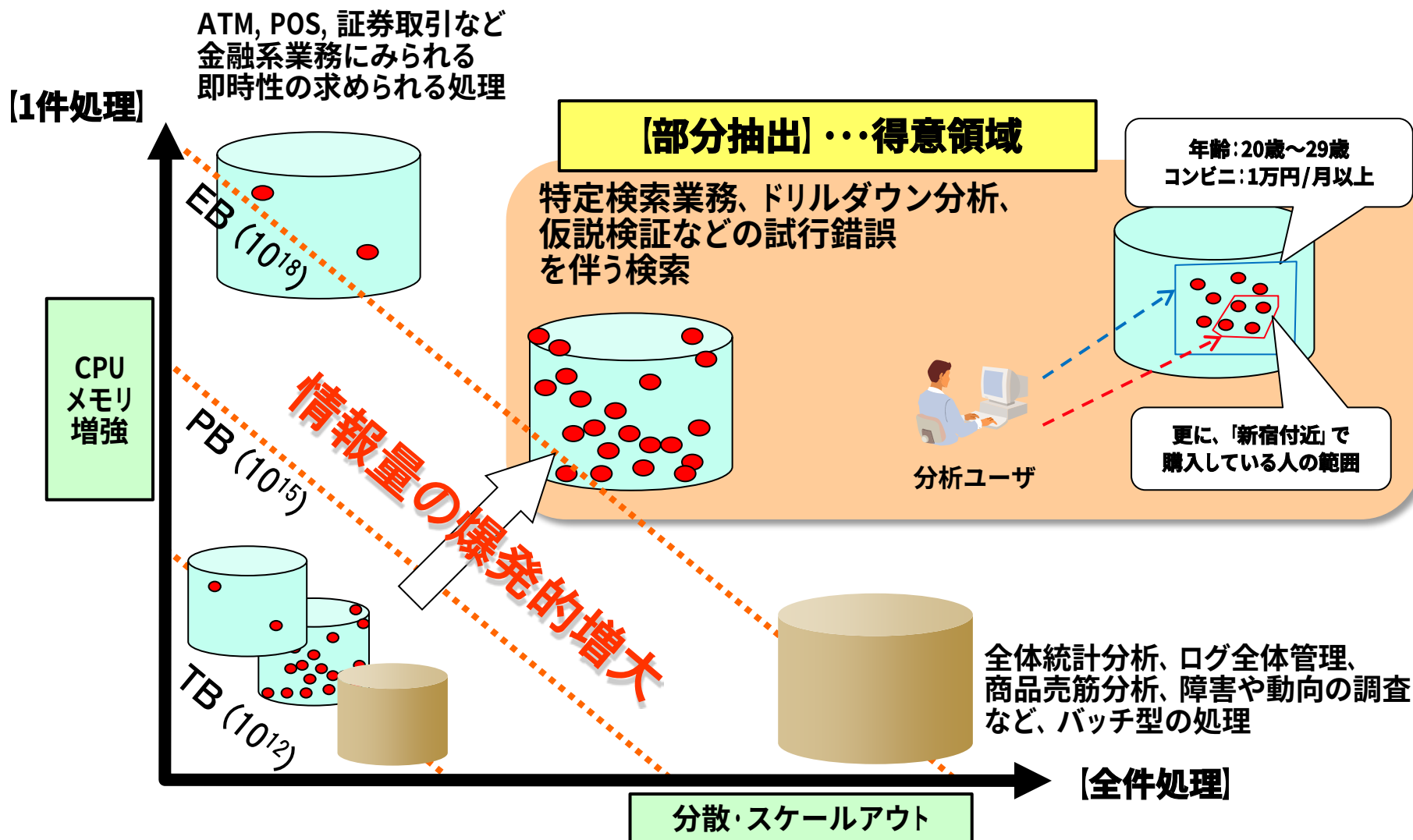
^{(*)1} 世界のトップを目指した先端的研究を推進することで、産業、安全保障等の分野における我が国の中長期的な国際競争力、底力の強化を図るとともに、研究開発成果の国民および社会への確かな還元を図ることを目的として創設された国の研究開発プログラム。
^{(*)2} 内閣府の最先端研究開発支援プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会サービスの実証・評価」(中心研究者:喜連川 東大教授/国立情報学研究所所長)の成果を利用。

3-8 高速データアクセス基盤の得意分野

3. ビッグデータの利活用

HITACHI
Inspire the Next

● 1件検索や全件検索ではなく、「部分抽出」に強みを発揮



3-9 高速データアクセス基盤 デモ

データベース

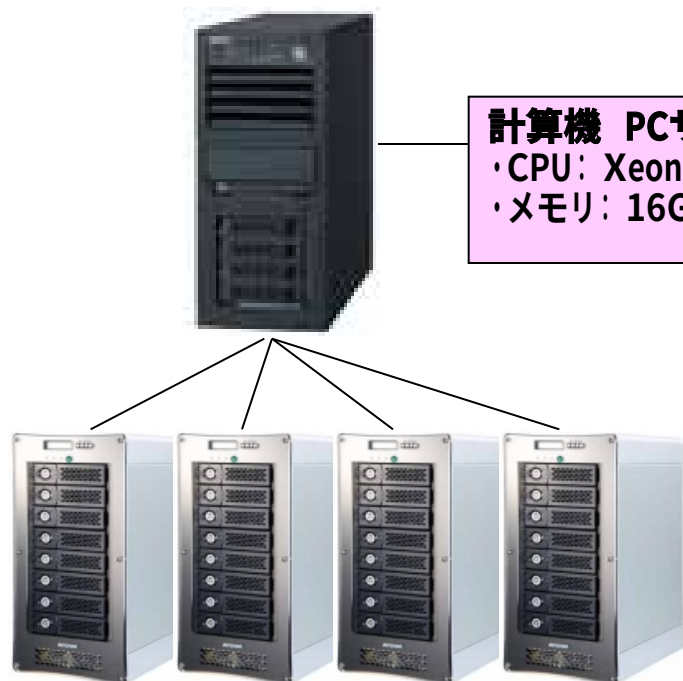
企業活動をモデル化したベンチマーク

約200GB (SF=100)

問合せ

注目するパーツの個別売上げレコードを抽出

ハードウェア

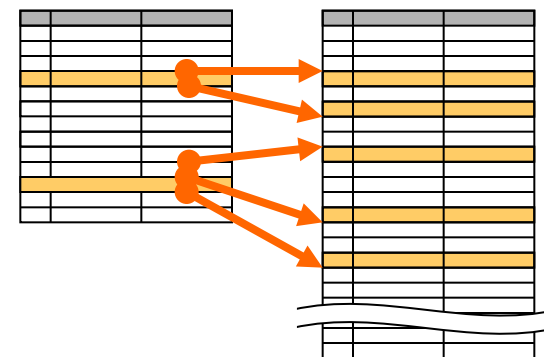


計算機 PCサーバ

- ・CPU: Xeon X3430(4コア, 2.4GHz)
- ・メモリ: 16GB

パーツデータ
2千万件

個別売上げデータ
6億件



検索結果: 65,000件

HDDストレージ装置 (HDD 32台)

ディスク回転時間: 15,000 rpm
平均シーク時間: 3.4 ms
インタフェース: SAS
容量: 600G x 32 = 19.2T

3-10 高速データアクセス基盤適用事例 (小売業)

3. ビッグデータの利活用

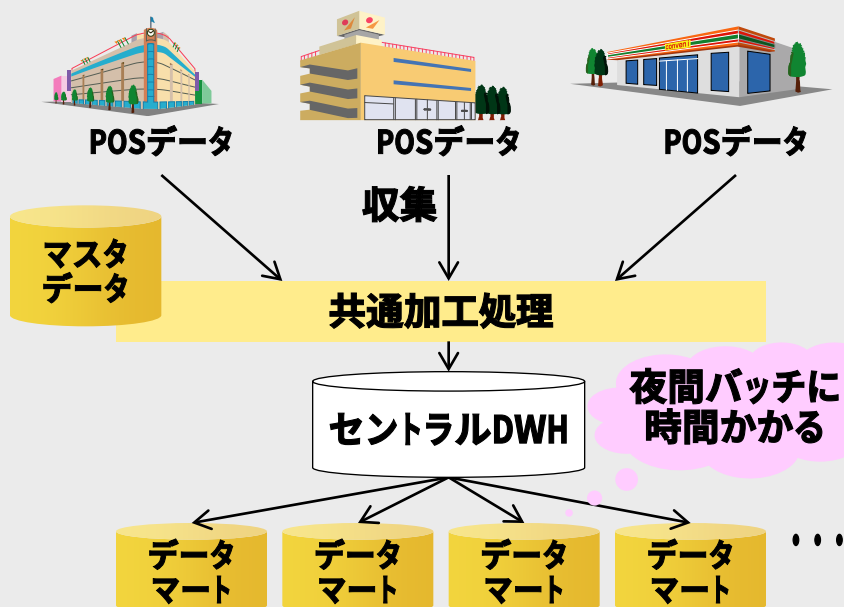
HITACHI
Inspire the Next

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: POS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かかっており、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。

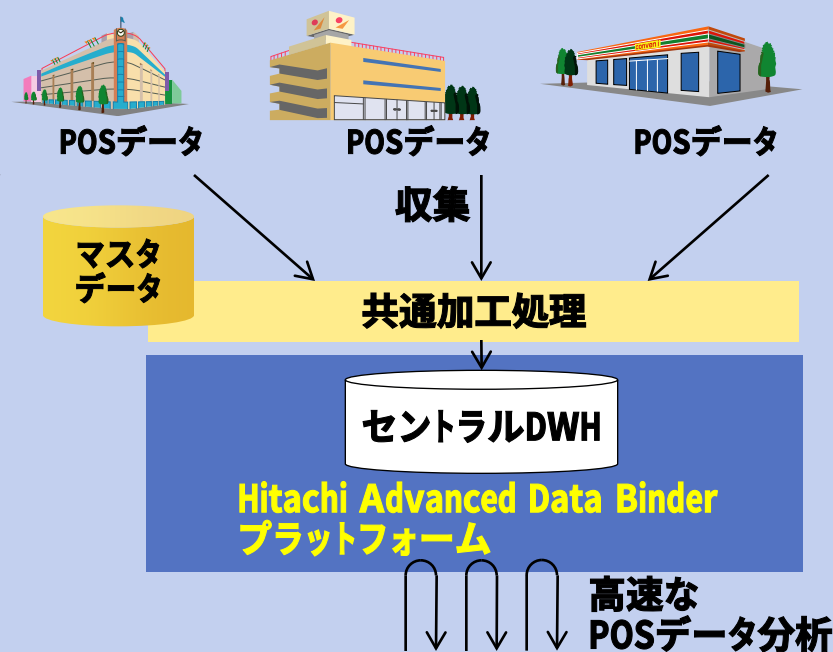


解決策

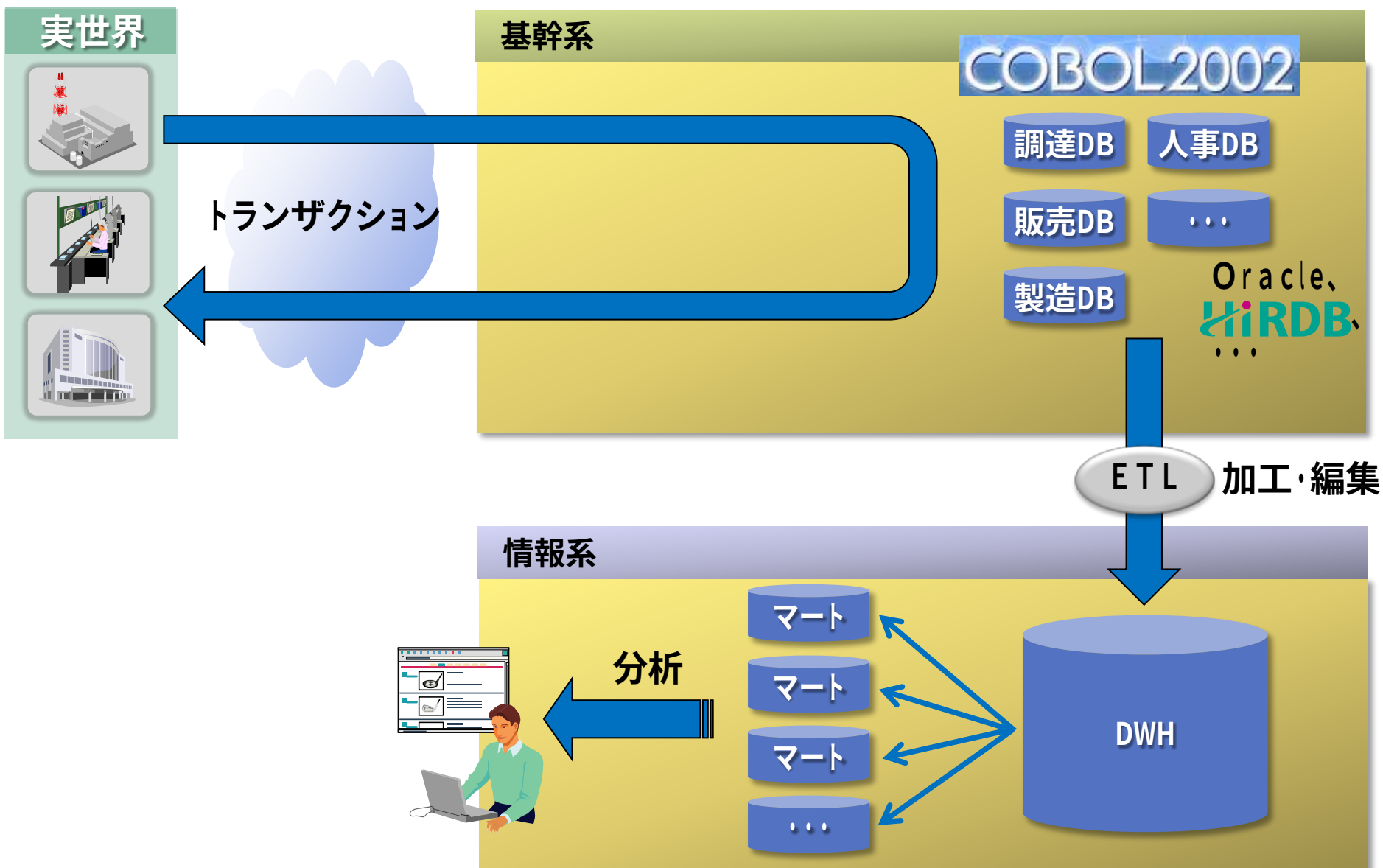
- 高速データアクセス基盤により、データマートを削減することが可能に



今後: セントラルなDWHのみ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。



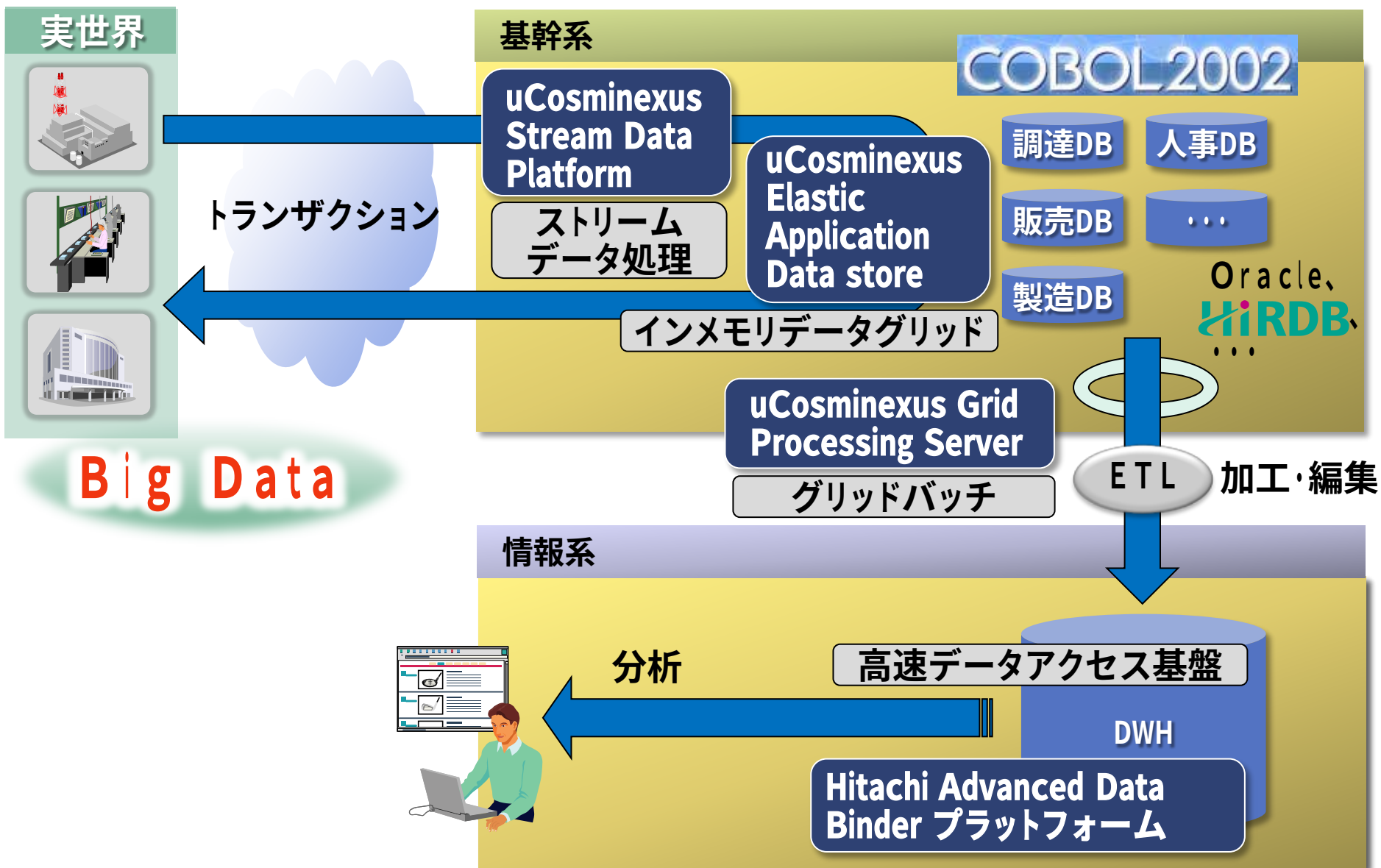
3-11 ITシステムの現状



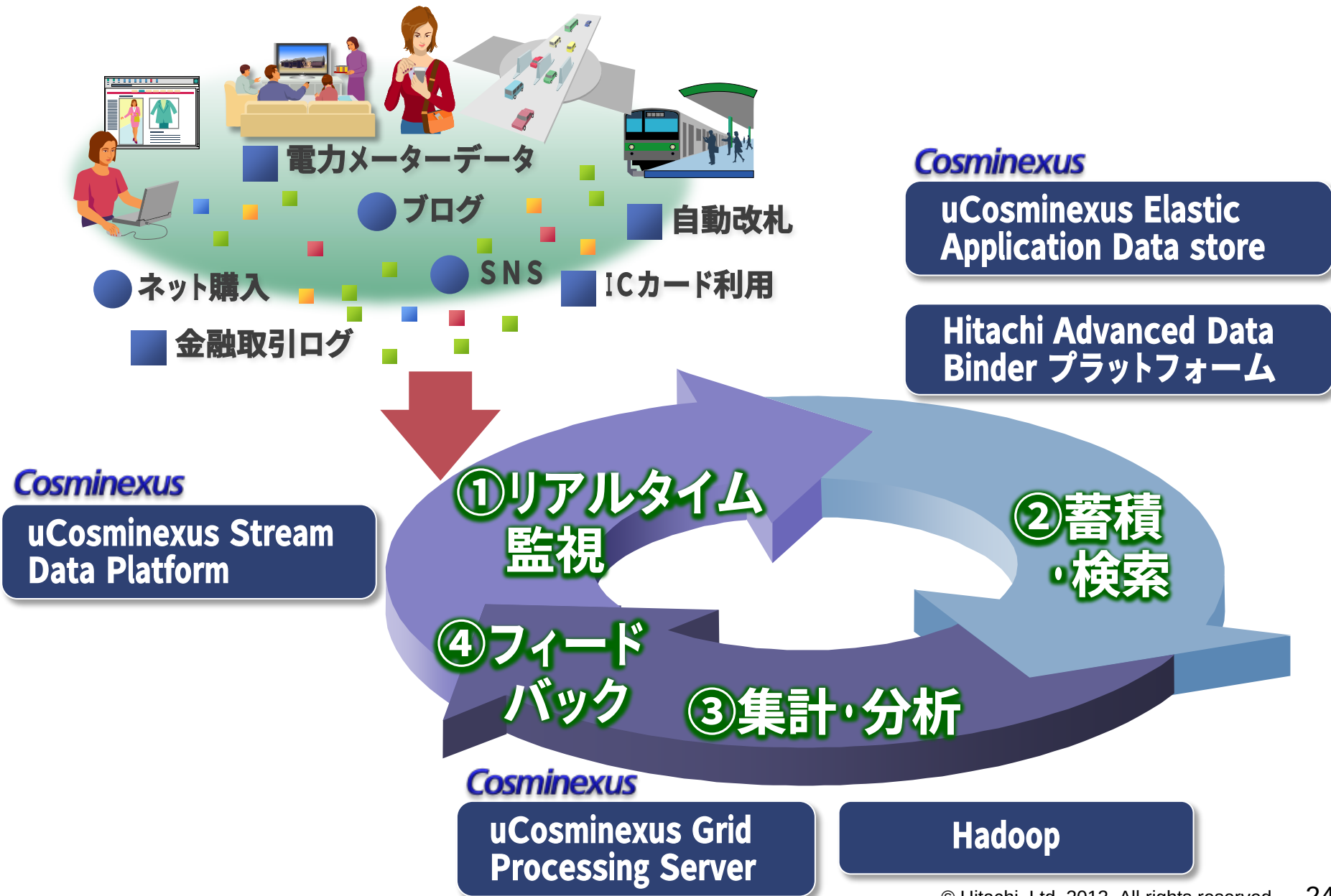
3-12 ITシステムの将来像

3. ビッグデータの利活用

HITACHI
Inspire the Next



4-1 ビッグデータのライフサイクル



4-2 COBOL資産とビッグデータ

ü既存COBOL資産や基幹データの有効活用が、
ビッグデータ利活用への第一歩

COBOL2002

ü基幹系データの増大の対応には、既存資産を
有効活用して高速化を図る「グリッドバッチ」

Cosminexus

uCosminexus Grid
Processing Server

üビッグデータの蓄積・超高速なデータ検索を
実現する「高速データアクセス基盤」

Hitachi Advanced Data
Binder プラットフォーム

4-3 まとめ

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、経験・人材・技術の組合せで
ビッグデータ利活用に取り組みます**

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

付-1 ビッグデータ利活用技術 (1)

SQL (Structured Query Language) 、 CQL (Continuous Query Language)

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)	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)	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 プラットフォーム)	「非順序型実行原理*」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

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

付-2 ビッグデータ利活用技術 (2)

OSS (Open Source Software)

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)	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)	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)



付-3 利活用ポイント (1)

■ 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 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

付-4 利活用ポイント (2)

■ 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

■他社商品名、商標等の引用に関する表示

- ・Hadoop は、Apache Software Foundationの商標です。
- ・Intel Xeonは、アメリカ合衆国およびその他の国におけるIntel Corporationの商標です。
- ・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。
- ・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>

END

ビッグデータへの挑戦
～基幹システムも融合して新しい価値の創造へ～

2013/04/19

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 ソフトウェア開発本部 ビッグデータソリューション部

山口 俊朗

Human Dreams. Make IT Real.

私たちは、ITと制御技術、そして社会インフラシステムで
人々の夢をかなえるイノベーションを起こしていきます。

HITACHI
Inspire the Next