

# クラウドと実世界ビッグデータ処理

日本電気株式会社(NEC)  
情報・ナレッジ研究所  
福島 俊一

# 本日のトピック 「クラウドと実世界ビッグデータ処理」

## 「クラウド」「ビッグデータ」というと？

Google Amazon Facebook …

ネット上の大量データ      Hadoop (MapReduce)

## NECが狙っている「実世界ビッグデータ」 を中心に話します

**監視カメラの映像解析**

**大量のセンサーデータのリアルタイム処理**

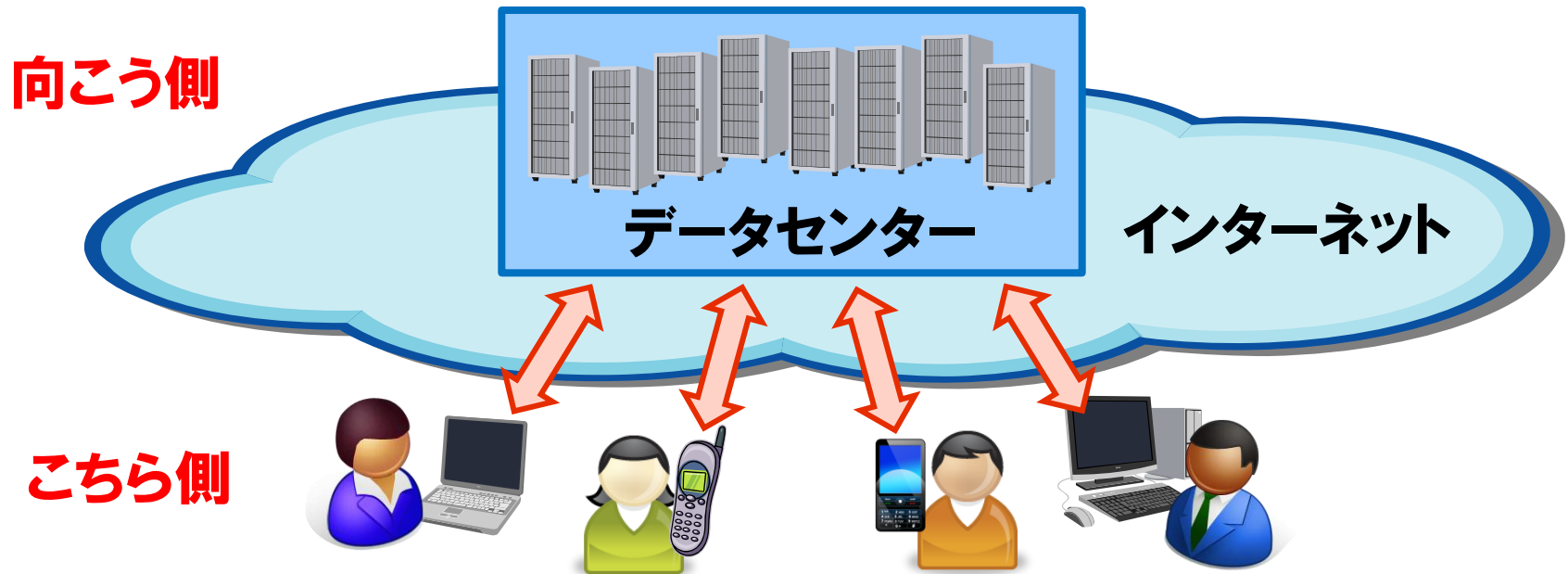
**最新技術では、こんなことまでできるようになった！**

## 第1部

# クラウドとビッグデータ

# クラウドコンピューティングとは？

インターネット側( **向こう側** )にあるITリソース(サーバ、ストレージ、データ、アプリケーション等)を、**こちら側**からサービスとして利用するコンピューティングモデル



ITリソースを「所有」するのではなく「利用」する

# ローカルアプリケーションとクラウドサービス（例）

## クラウドサービス



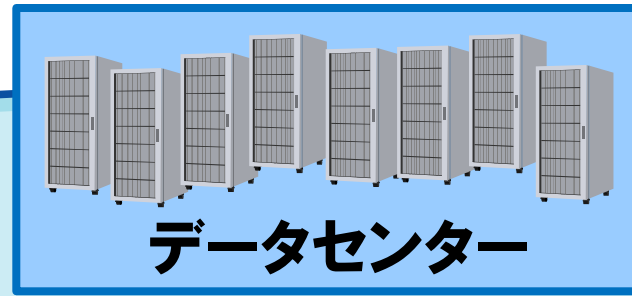
## ローカルアプリケーション／デバイス

# クラウドの5つの特徴

④ ITリソースをプールしておいて  
複数ユーザで共有

⑤ 迅速なシステム  
拡張・縮退

向こう側



こちら側

① ネットワークアクセス  
(どこからでも何からでも)

② すぐに  
利用できる

③ 使った分だけ払う  
(従量課金)

# クラウドサービスのメリットは？



ピーク時に備えた  
過剰な資産保有

手配から構築までの  
待ち時間

バージョンアップの手間

バックアップの手間



# クラウドサービスはどんどん普及する？ 懸念はないの？

利用者の手間が減る

ITサービスの利用が  
安く早くなる

コラボレーションが  
しやすくなる



セキュリティや信頼性に  
不安を感じる

利用する側の  
自由度・統制性が減る





# クラウドからビッグデータ活用へ

要するに  
効率化・コスト削減

利用者の手間が減る

ITサービスの利用が  
安く早くなる

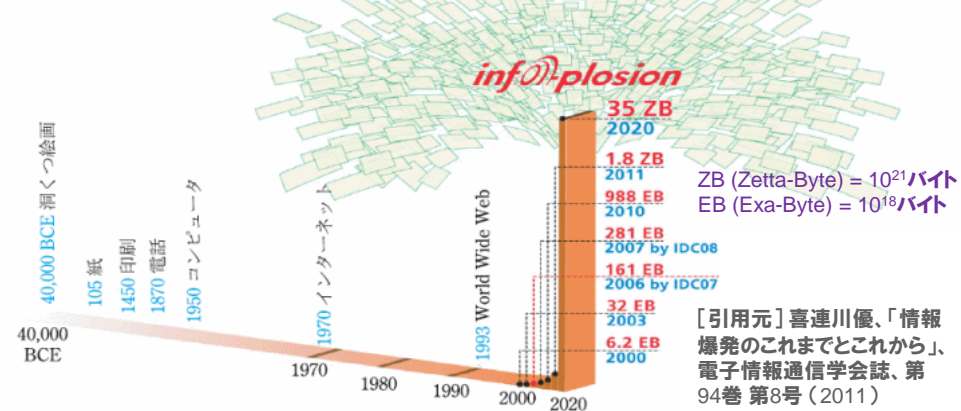
コラボレーションが  
しやすくなる



ビジネス創出・差異化の源泉は  
「ビッグデータ活用」

- クラウドはビッグデータ化を促進  
(クラウドにデータが集積)
- クラウドはビッグデータ活用を下支え  
(クラウドの強力な計算パワー)

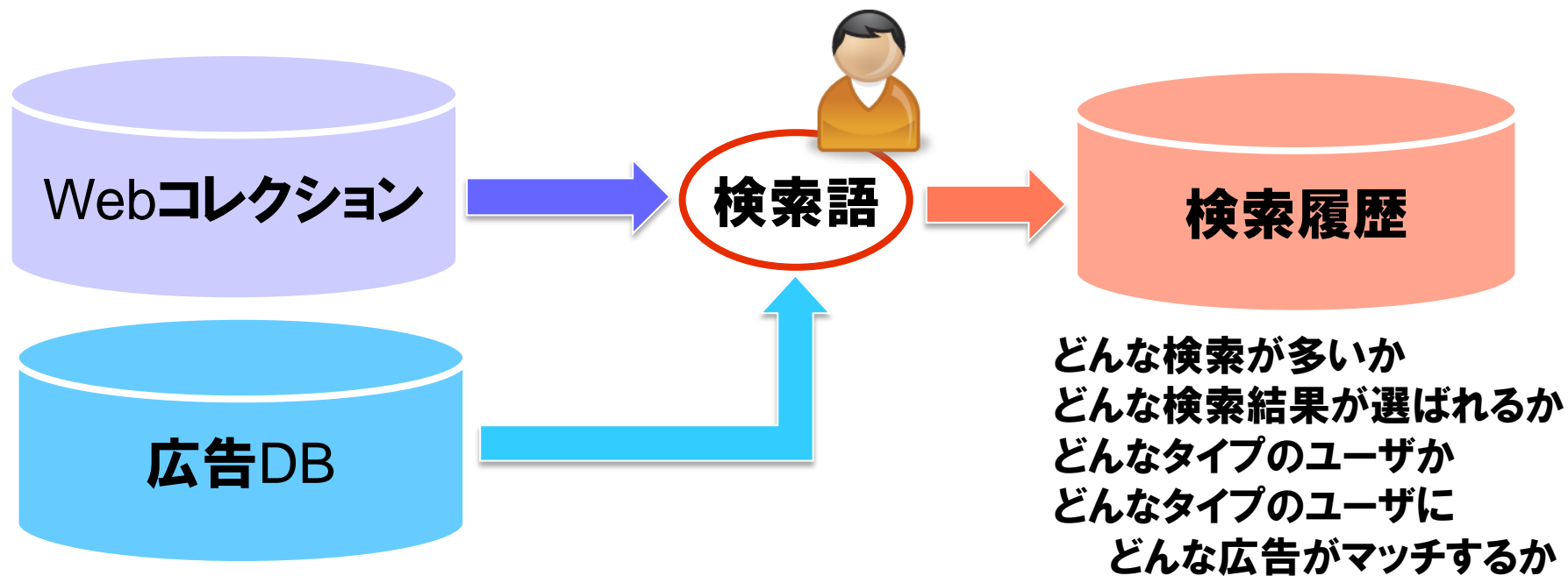
情報爆発をどのように価値に変えるか



情報爆発 (出典: Horison Information Strategies, cited from Storage New Game New Rules, p. 34 (www.horison.com), IDC, The Diverse and Exploding Digital Universe 2020 (http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm))

# ビッグデータ活用の例(1)Webサーチエンジン

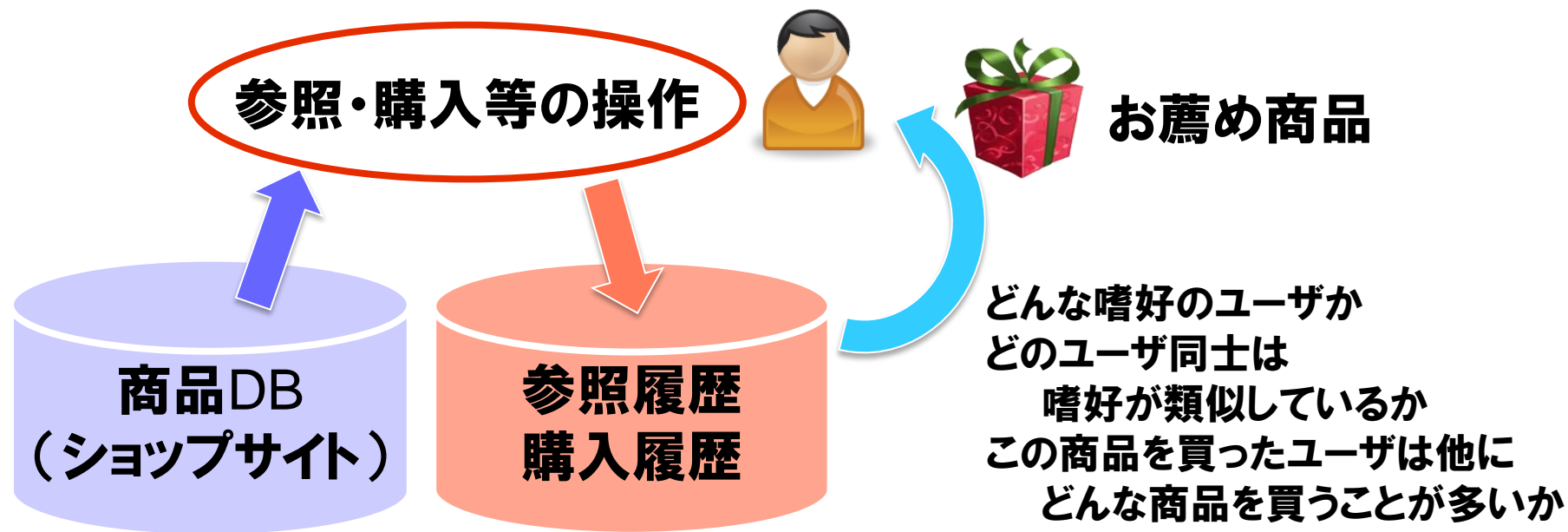
検索語を入力すると、それにマッチする検索結果と広告が提示される



- Webを大量に集める(欲しい情報が見つかる)ほど、ユーザが集まる
- ユーザが集まるほど、広告も集まる(売上が増える)
- たくさん使われるほど、検索や広告の精度が高くなる

## ビッグデータ活用の例(2)レコメンデーション

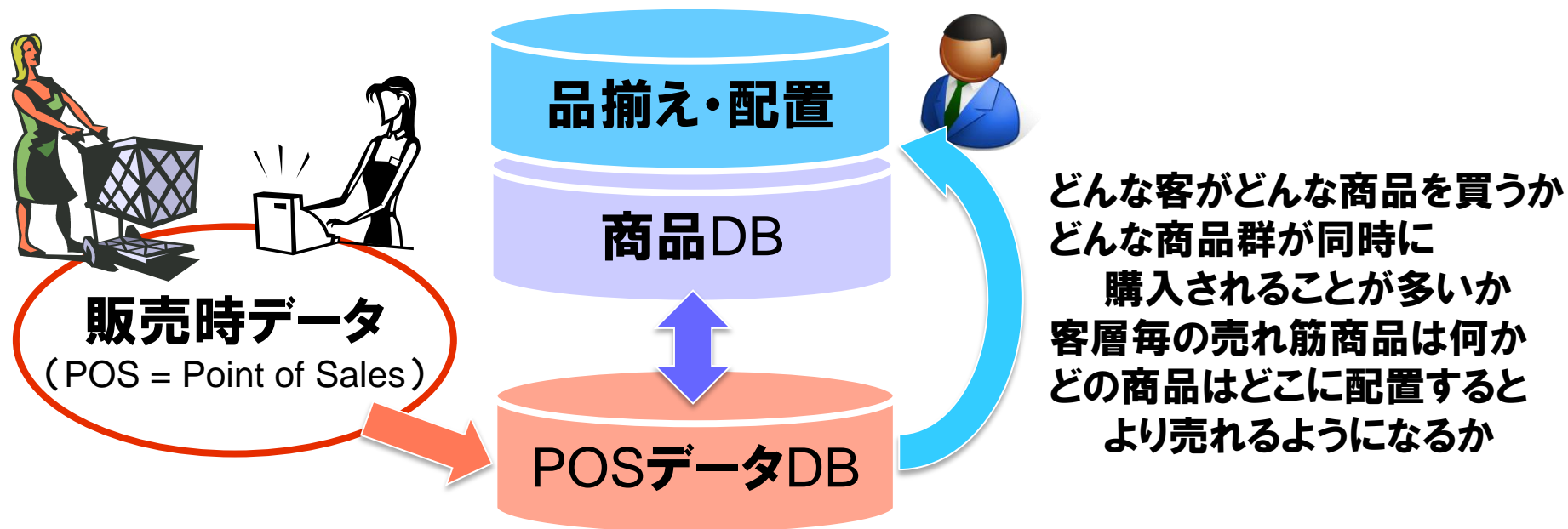
ショッピングサイトにおいて、ユーザの購入履歴やサイト参照履歴等に基づき、お薦め商品が提示される



- 様々な商品を集める(欲しい商品が何でも買える)ほど、ユーザが集まる
- 多数のユーザと多種の商品の関係がたまるほど、お薦め精度が高まる
- お薦め精度が高まると、商品売上が向上する

# ビッグデータ活用の例(3)POSデータ分析

リアル店舗において、販売時データ(購入商品、購入者属性、天候等)を集めて分析し、各店舗の品揃えや配置を計画



- その土地、その時期の売れ筋商品をそろえるほど、来店者が集まる
- 多数のPOSデータ(いろいろな店舗、いろいろな時期、いろいろな客層)が集まるほど、分析の精度向上やバリエーション充実につながる

# ビッグデータ処理を支える技術

## データを読み解く技術

大量データから価値を生み出す  
(ゴミの山を宝の山に変える)



例えば代表的なものとして

データマイニング技術

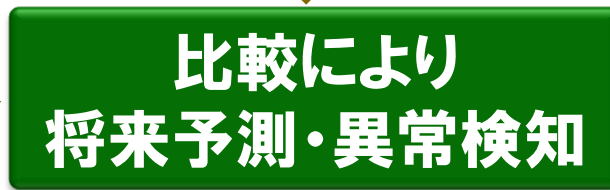
## 大量データの高速処理技術

データが大量でも止まらない、  
さらに増え続けても破綻しない

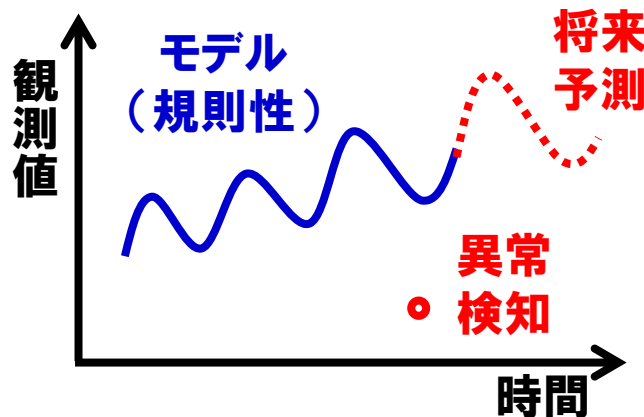
# データマイニング技術

観測されたデータ中の**隠れた規則性**を見つけ出し、  
将来予測、異常検知、最適アクション決定等を可能にする

データから規則性を学習  
(あるいは 理論・原理からモデル構築)



- POSデータ
- 参照・購買履歴
- 検索履歴
- メモリ使用量
- 株価データ
- 銀行・クレジット履歴  
など

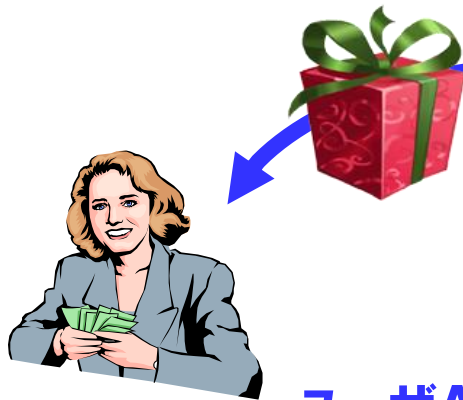


- 売れ筋商品予想
- 売上予測
- お薦め商品・広告
- システム異常検知
- 株価推移予想
- 信用判定  
など

# データ蓄積に基づく適応的なレコメンデーション

嗜好の類似する別ユーザのお気に入りアイテムを推薦する  
**協調フィルタリング方式**（クチコミの自動化）

⇒データがたまるほどお薦め精度が向上



嗜好が  
類似

ユーザA  
ユーザB  
ユーザC  
ユーザD

|      | アイテム1 | アイテム2 | アイテム3 | アイテム4 | アイテム5 | アイテム6 |
|------|-------|-------|-------|-------|-------|-------|
| ユーザA | 5     | 1     | —     | 4     | 1     | —     |
| ユーザB | 2     | 2     | 1     | 5     | 4     | —     |
| ユーザC | 5     | —     | 5     | 4     | 1     | 2     |
| ユーザD | —     | —     | —     | 2     | 4     | 4     |
|      |       |       |       |       |       |       |
|      |       |       |       |       |       |       |

マトリックス中の値は  
ユーザ評価:

好き(5) ⇔ 嫌い(1)  
「—」は未知

マトリックス中の  
既知の値をもとに  
未知の値を推定

ユーザ評価情報(ユーザ×アイテムのマトリックス)

# ビッグデータ処理を支える技術

## データを読み解く技術

大量データから価値を生み出す  
(ゴミの山を宝の山に変える)

例えば代表的なものとして

データマイニング技術

## 大量データの高速処理技術

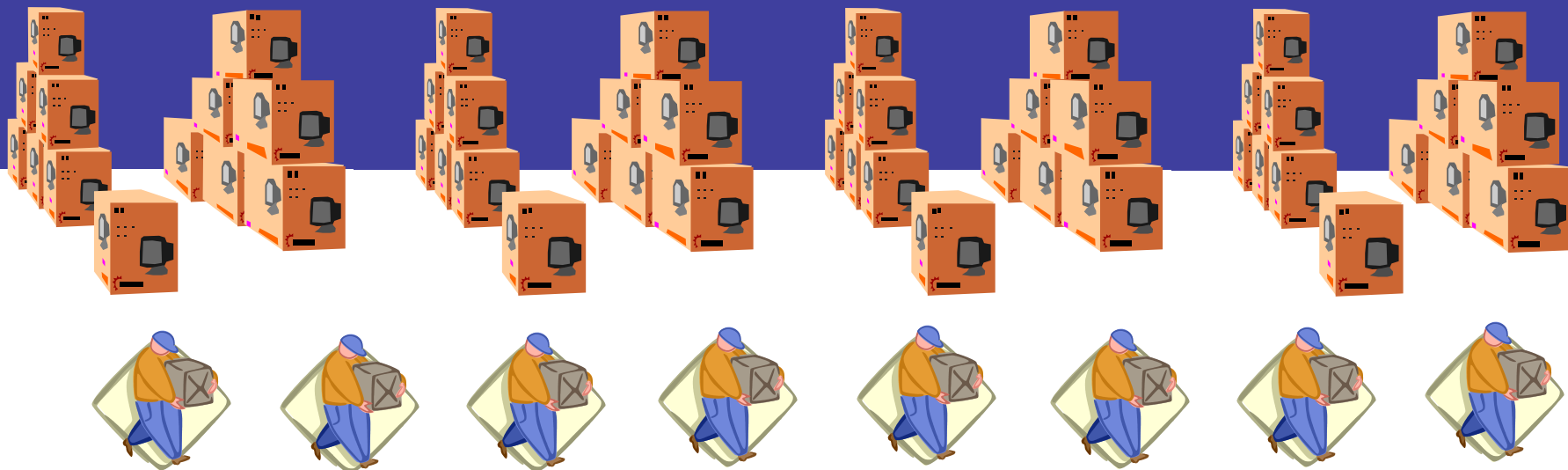
データが大量でも止まらない、  
さらに増え続けても破綻しない

例えば代表的なものとして

分散並列処理による  
スケールアウト技術



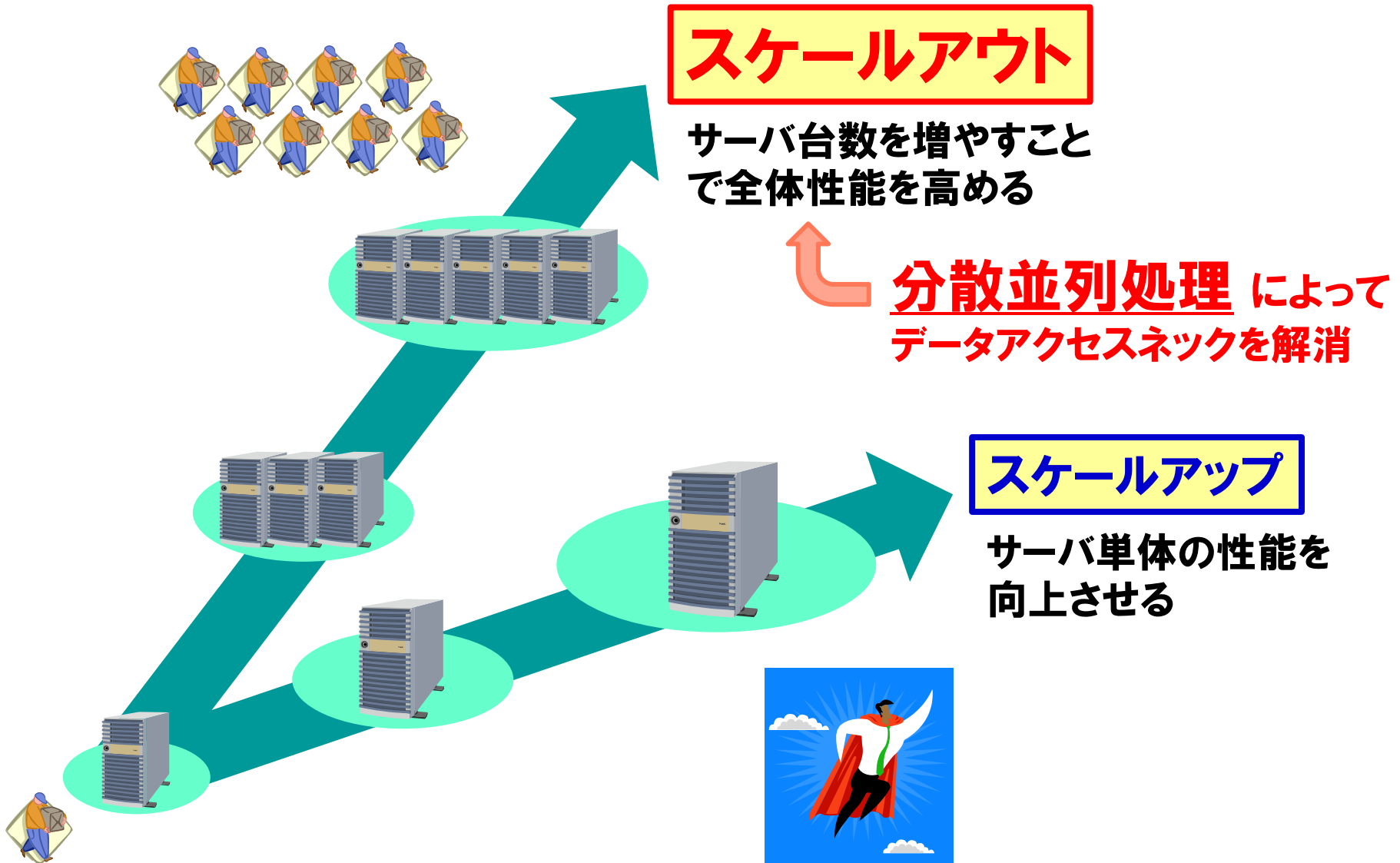
# 大量の荷物(データ)を急いでさばくには？



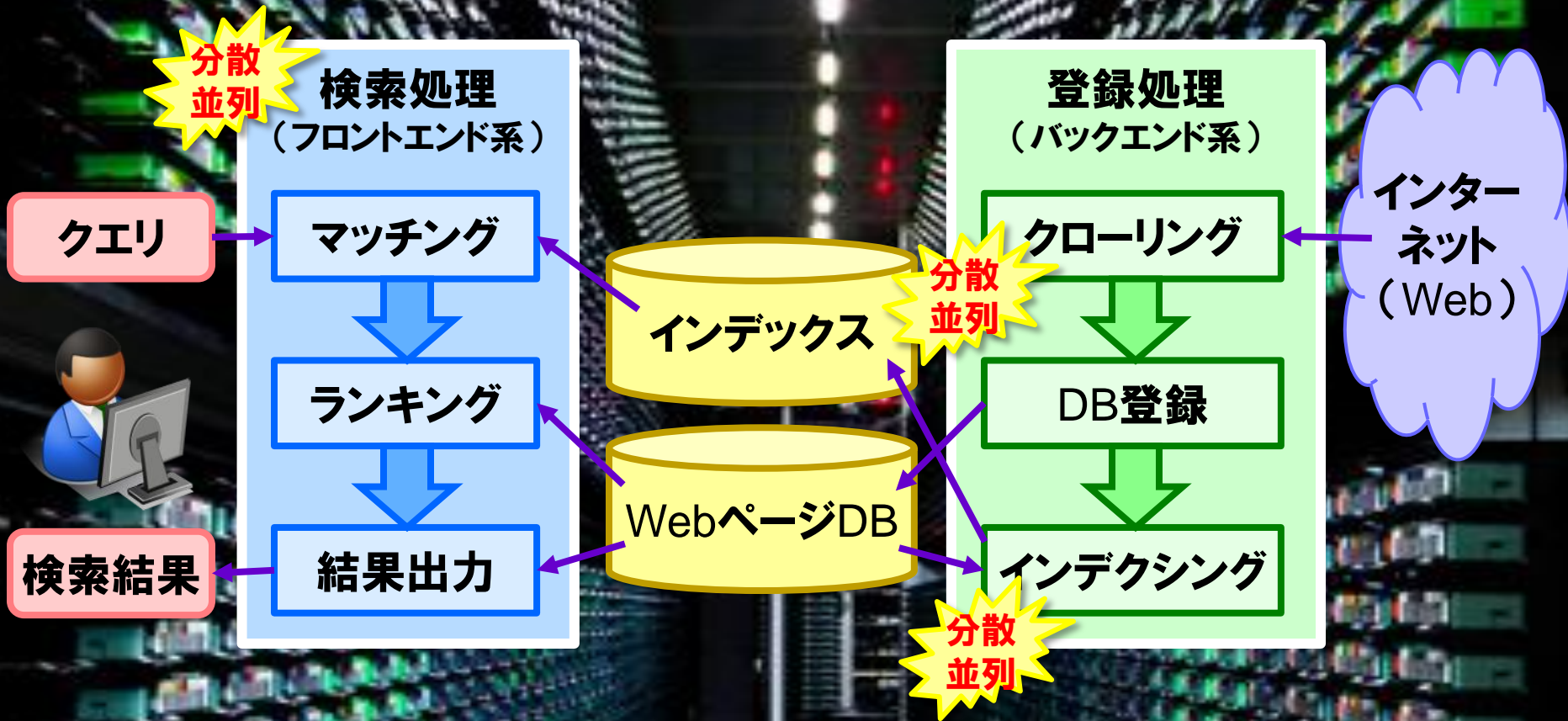
一人では大変

**多人数で分担すればすぐ片付けられる！**

# 大量データの処理にはスケールアウトが有効



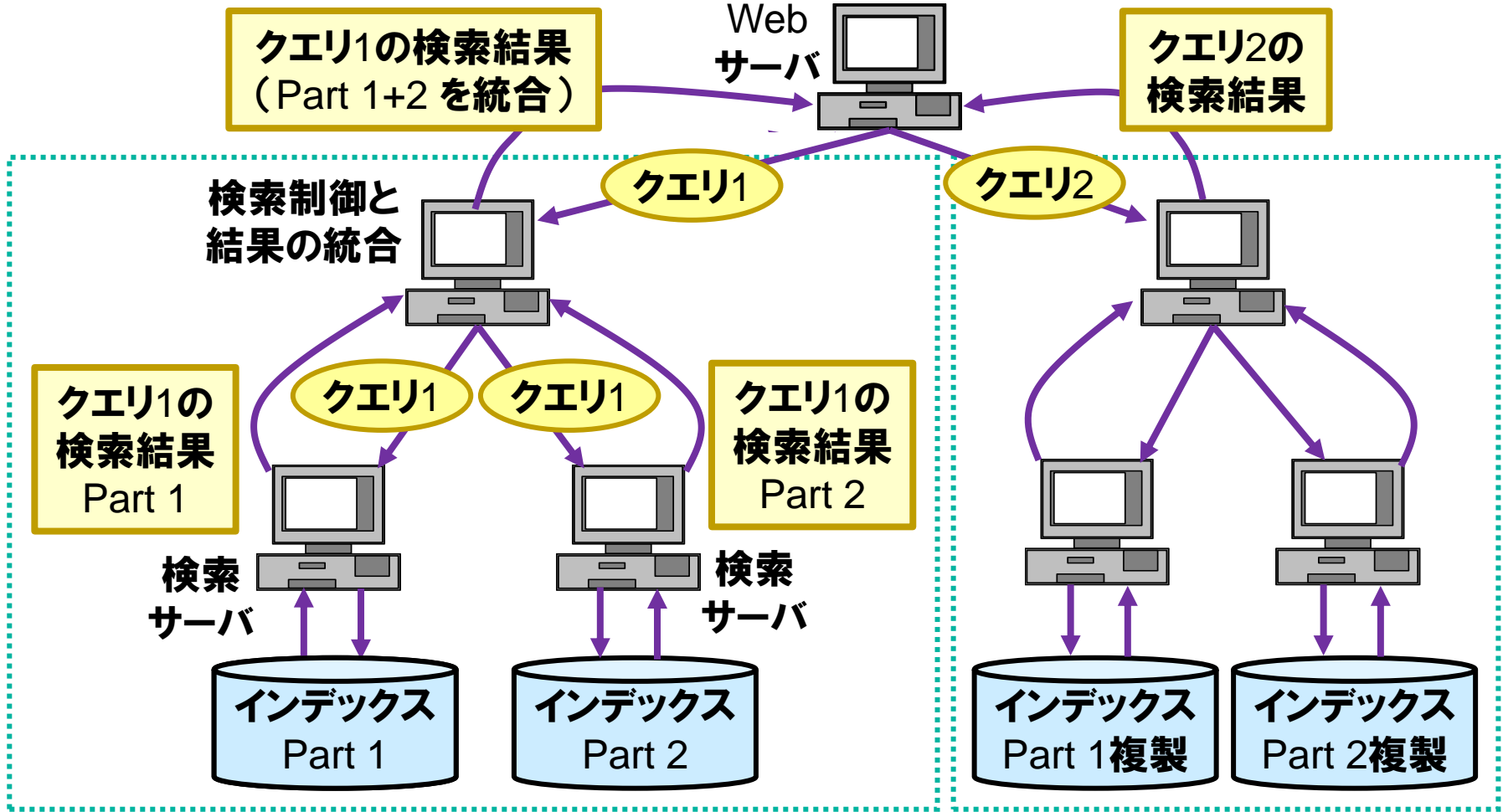
# スケールアウトの例：Webサーチエンジン



Webページ規模が年々増大しても、**スケールアウト**の考え方で、サーバ台数を増やすことで、高速な処理性能・応答時間を維持

[背景画像の出典] <http://www.google.com/about/datacenters/gallery/>

# 検索処理の分散並列化



**インデックス分割による検索レスポンス向上**  
(データ量に対するスケーラビリティ)

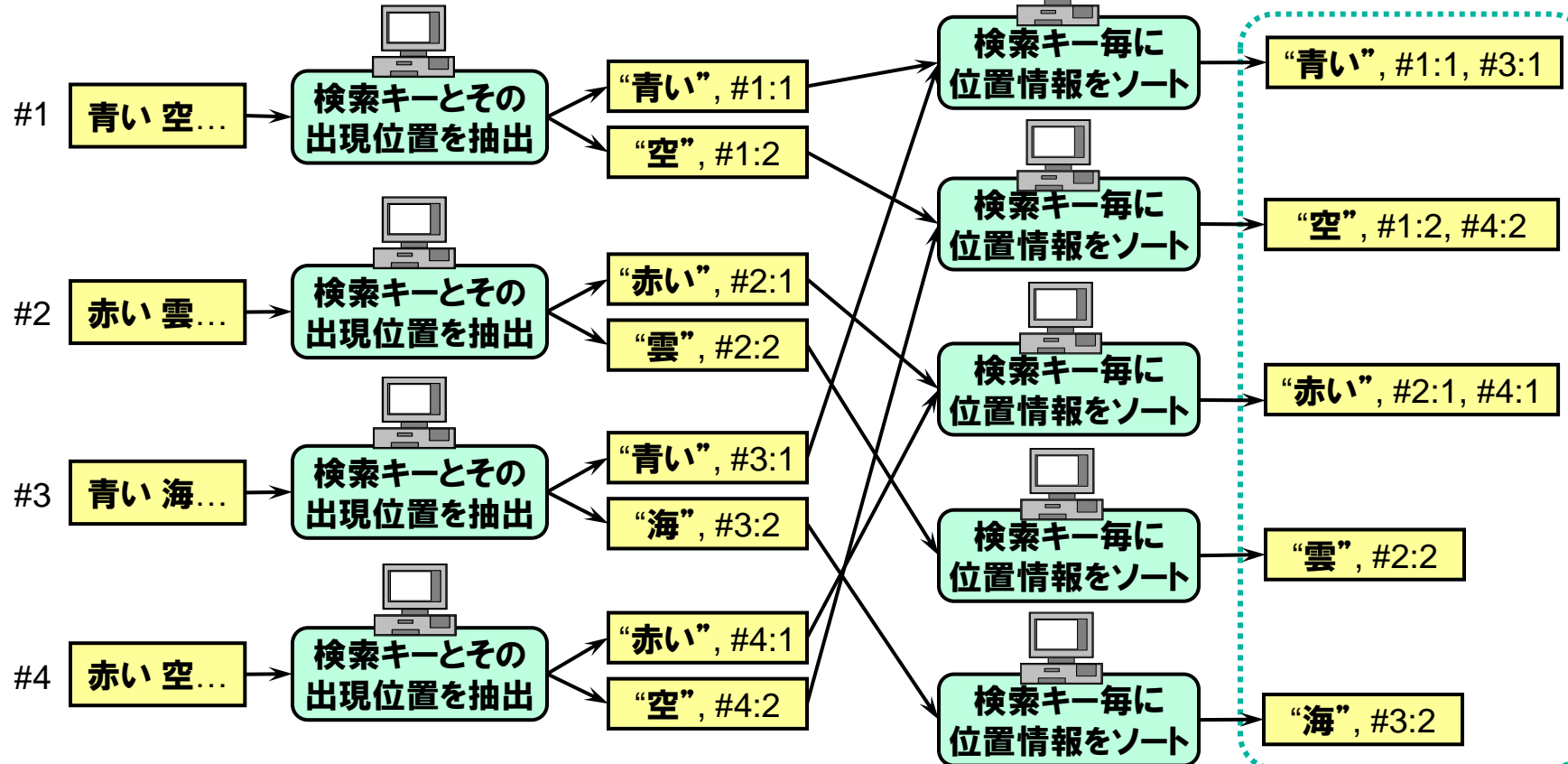
**複製によるスループット向上**  
(アクセス頻度に対するスケーラビリティ)

# インデクシング(登録処理)の分散並列化:MapReduce

Webページ

検索キーと  
出現位置のペア

検索キー毎に集約・ソートされた  
位置情報群



インデックスに登録

Webページ毎に処理が独立で  
単純に分散並列化が可能な「前段」

並列処理する前に、前段の結果を  
いったん集約することが必要な「後段」

# 第1部のまとめ

クラウドは効率化・コスト削減が中心



**ビジネス創出・差異化の源泉はビッグデータ活用**

(クラウドはビッグデータ化を促進、ビッグデータ活用を下支えする役割)

ビッグデータ処理のための2通りの技術

データを読み解く技術

データマイニング技術

大量データの高速処理技術

分散並列処理による  
スケールアウト技術

## 第2部

# 実世界ビッグデータの活用

# ビッグデータ(Big Data)とは？ = ネット上の大量データ？

従来のデータ管理技術では扱うのが困難なほど大規模なデータセット

なぜならば



量だけじゃない

| データ源  |          | データの例                    |
|-------|----------|--------------------------|
| サイバー系 | ネットサービス  | Webコンテンツ、SNS、Webアクセスデータ等 |
|       | 企業アプリ    | トランザクション、インハウスDB・DWH等    |
|       | システムリソース | ITシステムログ、ネットワークログ等       |
| 実世界系  | センサー類    | 監視カメラ映像、センサー(状況、追跡)等     |

3V

Volume  
(量)

Velocity  
(速度)

Variety  
(多様性)

ビッグ  
データ



これも増えている



# 様々なセンサーやカメラが、携帯に、自動車に、街中に・・・

## スマートフォンに内蔵されるセンサー

| タイプ         | センサ    | 主な測定対象           | 主な製品(未発売含む)                    |
|-------------|--------|------------------|--------------------------------|
| 主な内蔵<br>センサ | GPS    | 位置               | 多くの機種で内蔵                       |
|             | 加速度センサ | 加速度(傾き、動き、振動、衝撃) | 多くの機種で内蔵                       |
|             | 地磁気センサ | 地磁気(方向)          | 多くの機種で内蔵                       |
|             | ジャイロ   | 縦、横、斜めの姿勢        | 多くの機種で内蔵                       |
|             | 近接センサ  | 近接               | 多くの機種で内蔵                       |
|             | 照度センサ  | 明るさ              | 多くの機種で内蔵                       |
|             | 温湿度センサ | 温度、湿度            | 富士通「docomo PRIME series F-09C」 |
|             | 気圧計    | 気圧               | Galaxy Nexus内蔵                 |



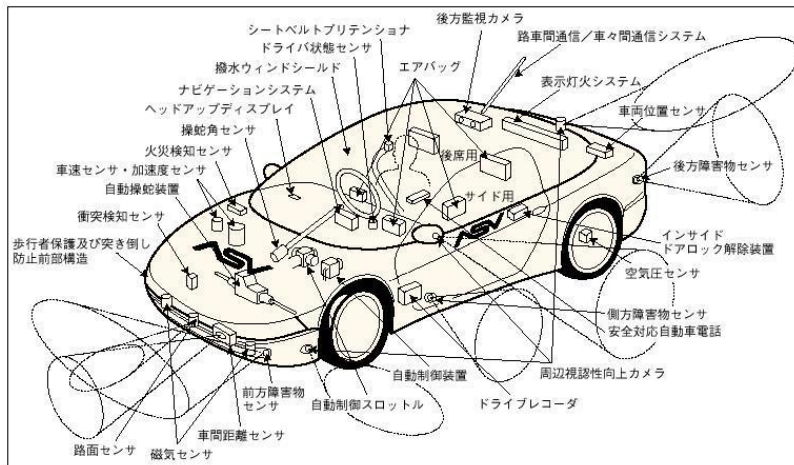
## 屋内外に設置された監視・防犯カメラシステム



## スマートフォンと連携する外付け型センサー

| タイプ           | センサ      | 主な測定対象           | 主な製品(未発売含む)                 |
|---------------|----------|------------------|-----------------------------|
| 外付け:環<br>境測定型 | 放射線センサ   | 放射線              | NTTドコモ「着せ替えセンサジャケット」        |
|               | 紫外線センサ   | 紫外線              | NTTドコモ「着せ替えセンサジャケット」        |
|               | 大気センサ    | 大気(大気の状態、有毒ガスなど) | Synkera Technologies        |
| 外付け:ヒ<br>ト測定型 | 脈派センサ    | 脈拍数、自律神経の状態      | iSiD、立命館大学、ローム「感性アプリ」       |
|               | 呼気センサ    | 口臭・アルコール濃度       | NTTドコモ「着せ替えセンサジャケット」        |
|               | 体脂肪計     | 体脂肪率、筋肉率         | NTTドコモ「着せ替えセンサジャケット」        |
|               | 心拍計センサ   | 心拍数              | アディダスジャパン「mi Coach CONNECT」 |
|               | モーションセンサ | 人の動き、筋肉の動き       | Jawbone「UP」                 |

## 自動車に搭載された様々なセンサー類



[出典] <http://easy.mri.co.jp/20120228.html>

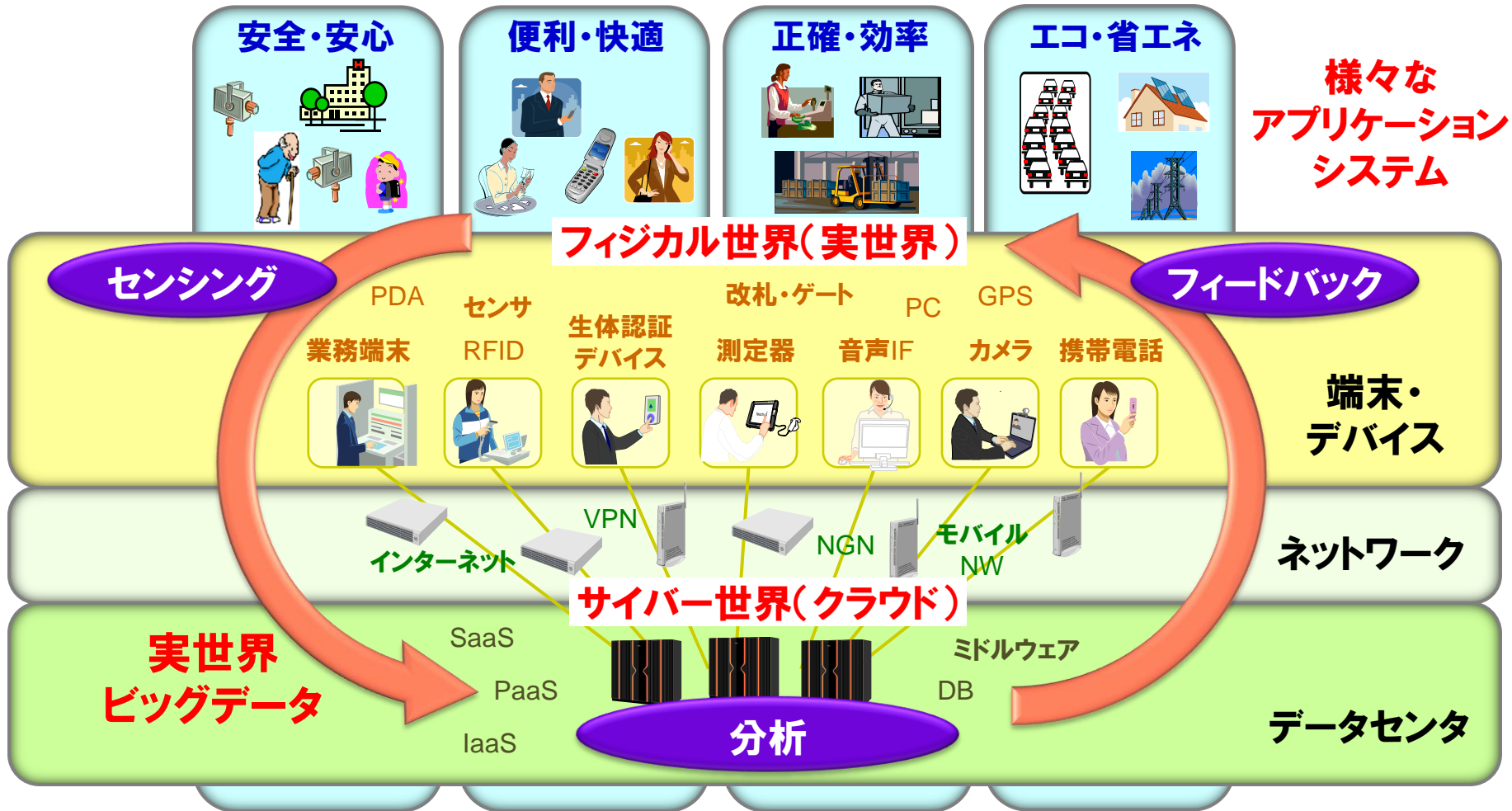
[出典] 国土交通省 先進安全自動車推進検討会資料

# 実世界ビッグデータが注目される背景

- **カメラやセンサーから集まる実世界データが急増**
- **現場を起点としたビジネス創出・差異化**
- **情報通信技術( ICT )が社会インフラとして広く浸透  
( 防災、交通、流通等 )**
- **実世界接点端末やその活用は日本に強み  
( 携帯、自動車、POS等 )**

# サイバーフィジカルシステム (Cyber-Physical Systems)

実世界データもクラウドに取り込み、集まったデータを解釈・分析することで、実世界の活動をより良くするフィードバックを生み出す



# 実世界ビッグデータ処理を支える技術

## データを読み解く技術

大量データから価値を生み出す  
(ゴミの山を宝の山に変える)

映像解析技術

データマイニング技術

## 大量データの高速処理技術

データが大量でも止まらない、  
さらに増え続けても破綻しない

分散並列処理による  
スケールアウト技術

イベントストリーム  
処理技術

## 第2部のまとめ ～ 第3部・第4部へのイントロ

ビッグデータはネット上の大量データだけじゃない  
実世界ビッグデータを活用するサイバーフィジカルシステム

そのために新たに必要になる2つの技術

データを読み解く技術

大量データの高速処理技術

映像解析技術

第3部

イベントストリーム  
処理技術

第4部

アプリケーションとともに各技術を紹介

## 第3部

# 実世界で何が起きているかを知る 映像解析技術

# 映像解析技術：実世界で何が起こっているかを知る

## データを読み解く技術

大量データから価値を生み出す  
(ゴミの山を宝の山に変える)

映像解析技術

データマイニング技術

映像データは、その中身を理解し、データマイニングに  
かけられるデータに変えることが必要

ネット上の  
データ

「誰が、何を買ったか」

実世界の  
映像データ

映像解析

「誰が、どこにいるか」

# < 誰なのか？ > 顔認証

## USJ顔認証ゲートシステム

大阪ユニバーサル・スタジオ・ジャパン(USJ)の「年間スタジオ・パス」



入場ゲートでカメラに顔を映すと



登録済みの人であれば“顔パス！”



【NEC Online TV(動画)】[http://www.nec.co.jp/ad/onlinetv/ja/business/facialrecog\\_h.html](http://www.nec.co.jp/ad/onlinetv/ja/business/facialrecog_h.html)

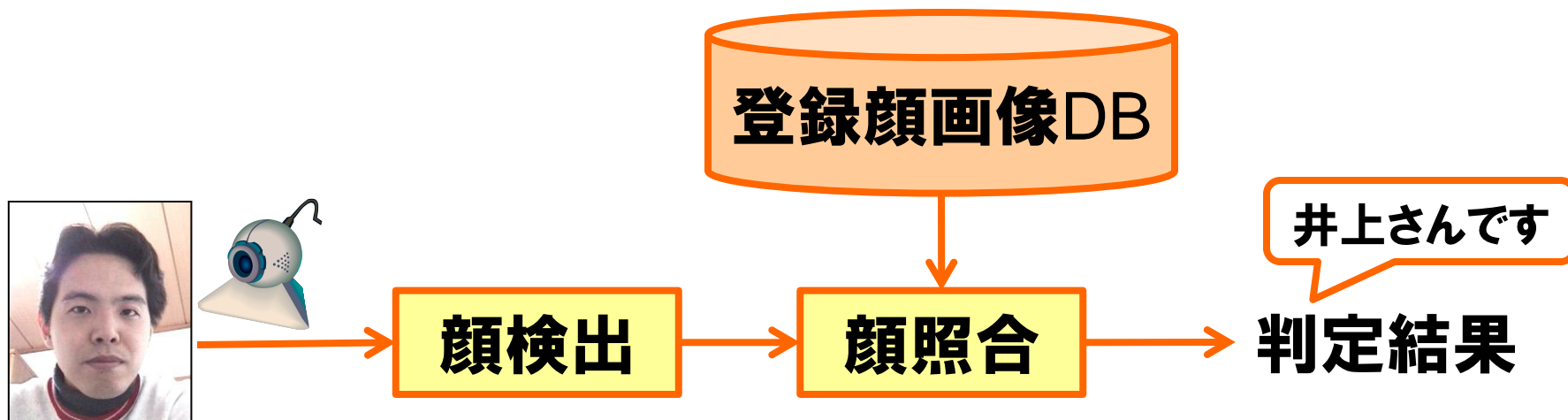
【NEC発見チャンネル】<http://www.nec.co.jp/profile/mitatv/02/index.html>

【製品(NeoFace)紹介】<http://www.nec.co.jp/soft/neoface/index.html>



# 顔認証とは？

顔を写した画像・映像にもとに、それが誰であるかを判断する技術  
(事前にDBに登録してある人の中の誰であるかを見つける)



顔のどんな特徴に着目するか  
(特徴ベクトルの類似性判定)

# 顔認証の難しさは？

眼鏡の有無や  
髪形が異なる

顔の向きが異なる

入力顔画像



照合

登録顔画像

歳をとる

表情が異なる

照明条件が  
異なる

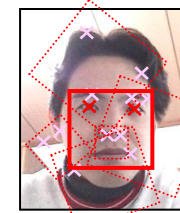
環境や対象物に変動があっても同一人物と判定できる  
(その一方で別人を同一人物と間違わない)



# 変動に強く、高精度な顔認証技術(NEC)

顔検出が高精度で高速  
非正面顔でも検出可能

多重照合  
顔検出法



目の候補を抽出、  
任意のペアを含む  
領域をGLVQにより  
顔/非顔を判定

GLVQ(一般化学習ベクトル量子化)は  
NEC独自の学習アルゴリズム

顔の向きや照明等  
大局的な変動を吸収

摂動空間法

姿勢・照明条件を変動させた  
様々な顔画像を生成



登録画像

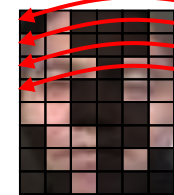
3D顔  
モデル



眼鏡の有無や表情変化等  
局所的な変動を吸収

適応的領域混合  
マッチング法

類似度の高い部分領域に着目



登録画像



照合画像

# <どこにいるのか?> 人物追跡

異なる角度から複数カメラで撮影した映像を対応付けて解析し、人の動きを高精度・リアルタイムに追跡



人物追跡システム

お客様がレジでお待ちです

バックヤード トイレ

コピー機

ATM

飲み物

弁当

本・雑誌

入口

お客様

お客様

店員

店員 1名

お客様 4名

NEC

# マルチカメラ映像解析による人物追跡技術(NEC)

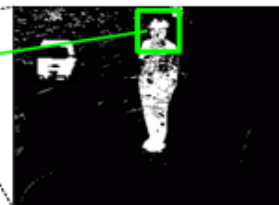
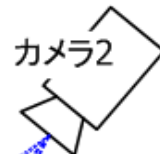
カメラ画像1



シルエット画像1

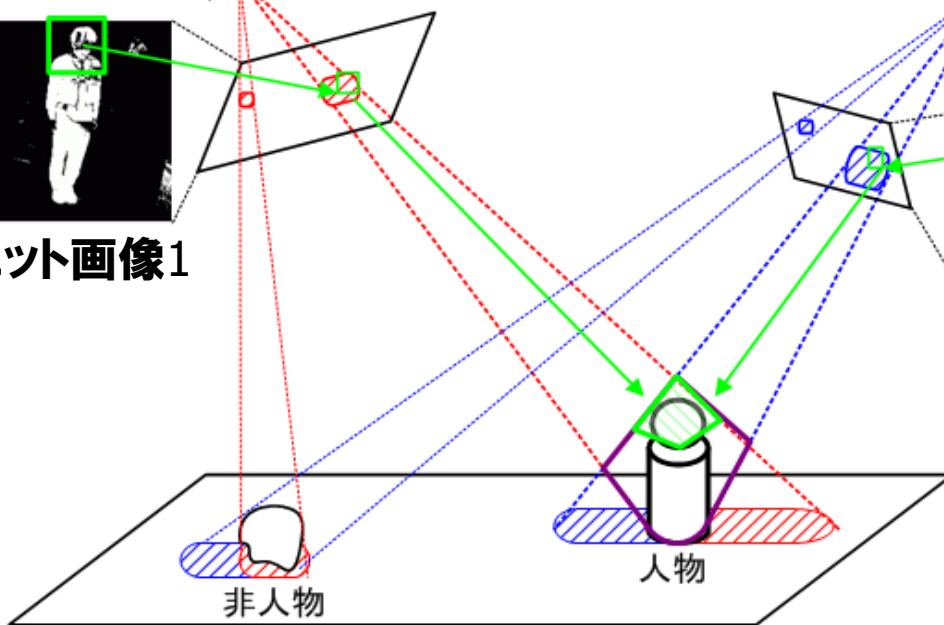
- 頭部検出結果
- 人体ボクセル

カメラ画像2



シルエット画像2

背景差分  
処理



3次元空間に逆投影し、  
視体積の交差部分から  
物体の3次元形状を復元

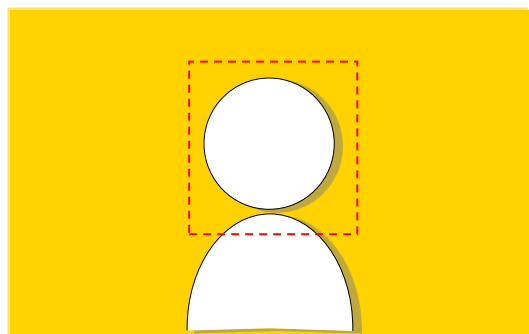
**視体積交差法**

歩行者モデルの当てはまり度から  
人物の位置を推定、時系列追跡

**リアルタイム性を確保、隠れ・重なり問題にも対処**

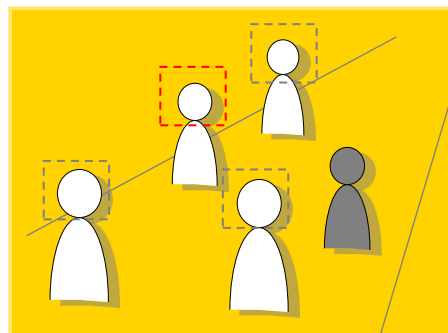
# <誰がどこにいるか？> 群衆中の個人を認証

顔認証

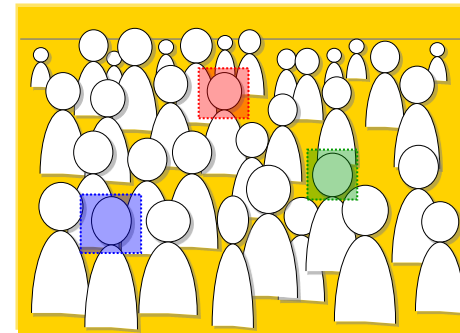


ゲート(1人)

映像監視



通路・施設内(10~100人)



群集(100人~)

## 顔認証に比べて映像監視はどこが難しいか？

顔の位置：固定



顔の位置：追跡が必要

顔の向き：正面



顔の向き：非正面

顔画像の質：高解像



顔画像の質：低解像

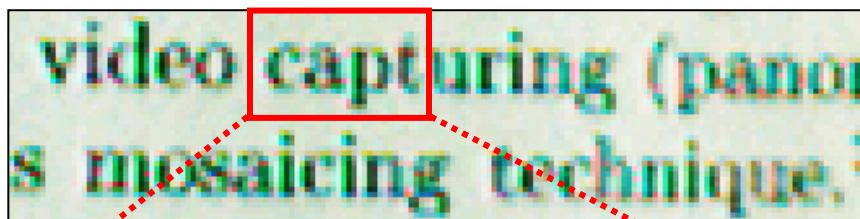
# 超解像技術：低解像画像を高解像画像に自動変換！



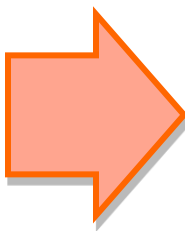
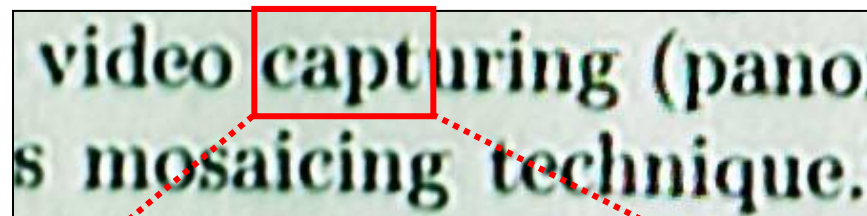
## カメラを変えずに画像処理によって画像を高解像度化

- カメラ自体の解像度を超えた画質に、撮影済みの画像でも高品質化可能
- 低解像の監視カメラで、群衆中の個人の認証を可能に

元画像



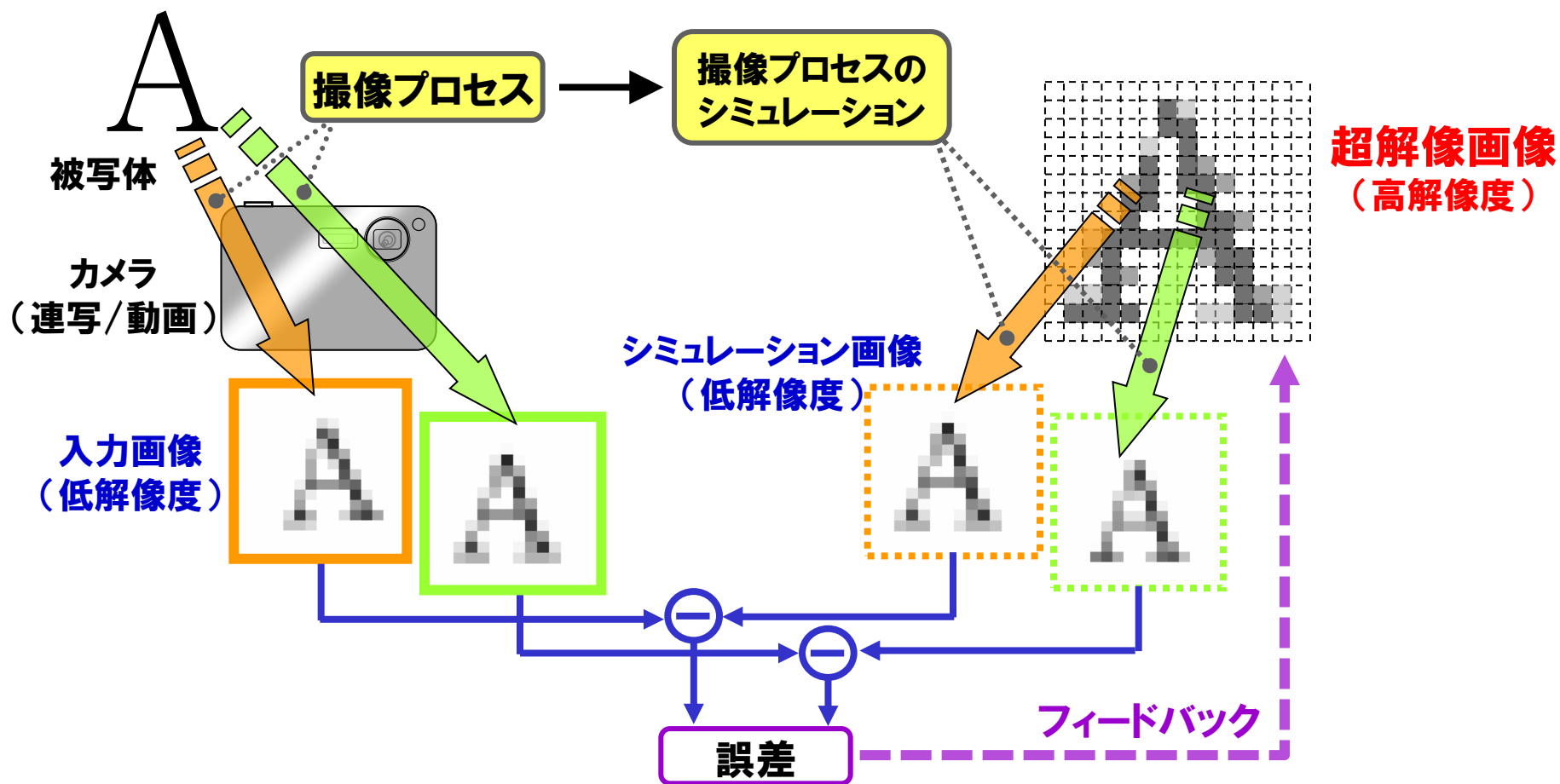
超解像画像



# 超解像の方式(その1) 複数枚超解像方式

## 複数枚の画像(動画もしくは連写)から高解像画像を復元

撮像プロセスをシミュレートした低解像画像と入力画像との誤差を最小にする高解像度画像を推定





# 超解像の方式(その2) 学習型超解像方式(NEC)

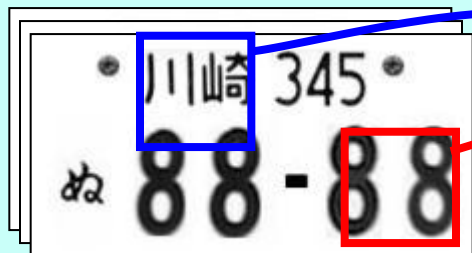
あらかじめ学習した知識を使って高解像画像を復元

学習フェーズ

復元(超解像)フェーズ

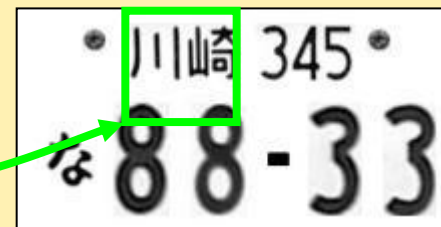
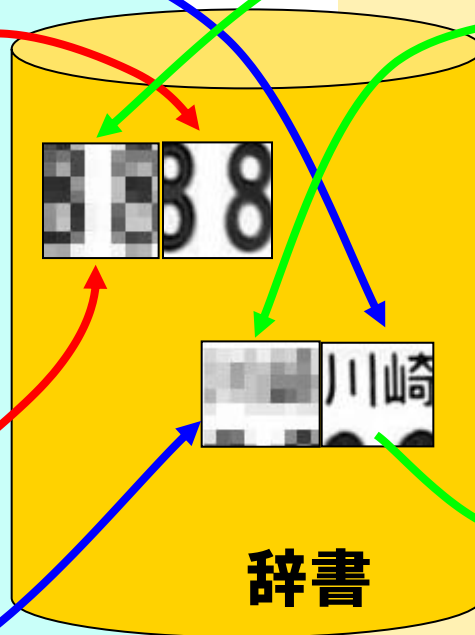
学習用画像(高解像度)

入力画像(低解像度)



画像縮小で  
低解像画像を生成

最も似ている  
パッチを探索



学習用画像中の小領域(パッチ)の  
高解像・低解像画像ペアを多数辞書登録

入力画像に対して最適なパッチを辞書から探索し、  
合成処理によって高解像度画像を復元

# 超解像方式の特徴比較

|             | 複数枚超解像方式                   | 学習型超解像方式                |
|-------------|----------------------------|-------------------------|
| 方式          | 複数枚の画像(動画もしくは連写)から高解像画像を復元 | あらかじめ学習した知識を使って高解像画像を復元 |
| 入力画像枚数      | 複数枚<br>(カメラまたは被写体の動きが必要)   | 1枚                      |
| 対象          | 限定なし                       | 特定の被写体<br>(顔、ナンバープレート等) |
| 倍率<br>(画素数) | ~3倍<br>(~9倍)               | 4倍~<br>(16倍~)           |

# 映像解析のアプリケーション例(1)犯罪監視

監視カメラ映像(過去の映像の蓄積 or 今のリアルタイム映像)から、**服装・人相など特定の条件に合致する人物** あるいは **不審者・不審行動** 等を見つける

- 目撃情報に基づき、犯人の服装・人相に合致する人物の足取りを追跡する
- 指名手配犯等の顔写真をあらかじめ登録しておき、見つけたらアラート
- 立入禁止区域に入る人物や、不審行動(置き引き、万引き、共連れ等)を検知してアラート



人間だと、**うっかり見落と**してしまったり、**目が行きとどかない**ケースが発生

**機械の目は、一度に沢山を見張ったり探したりすることができ、また、疲れを知らない**

# 映像解析のアプリケーション例(2)来店者の行動分析

店舗内をカメラで監視し、**来店者の性別・年齢層や動線を抽出**することで、来店者の傾向を把握し、今後の集客・売上向上に結び付ける

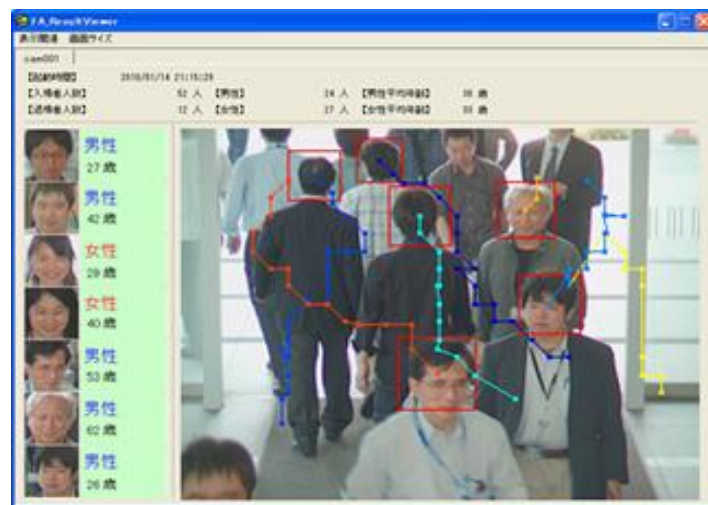
- 人の位置変化を追跡し、来店者が店に入ってから出るまでの動線を把握
- 映像中の人の顔や服装を解析することで性別・年齢層を判別



従来も店員がレジで購入者の性別・年齢層を入力し、客層別売れ筋を分析し、品揃えに反映させている



購入に至るまでの来店者の動線や、買わずに帰った来店者の動線も把握することができるようになり、**商品配置の改善**や、**売り逃しの機会・要因把握**も可能になる



NECソフト FieldAnalyst

<http://www.necsoft.co.jp/vtc/image.html>

## 第3部のまとめ

**映像解析によって実世界で何が起こっているかを把握**

**映像はますます増える、実世界ビッグデータの中心**

**データ量の問題もあるが、むしろ、  
広域の環境・認識対象の多様性が技術的難しさの主要因**

**SFや映画の世界だと思っていたことが  
現実になりつつある(と思いませんか?)**

## 第4部

# リアルタイムアクションを可能にする イベントストリーム処理技術

# イベントストリーム処理技術:リアルタイムアクションを可能に

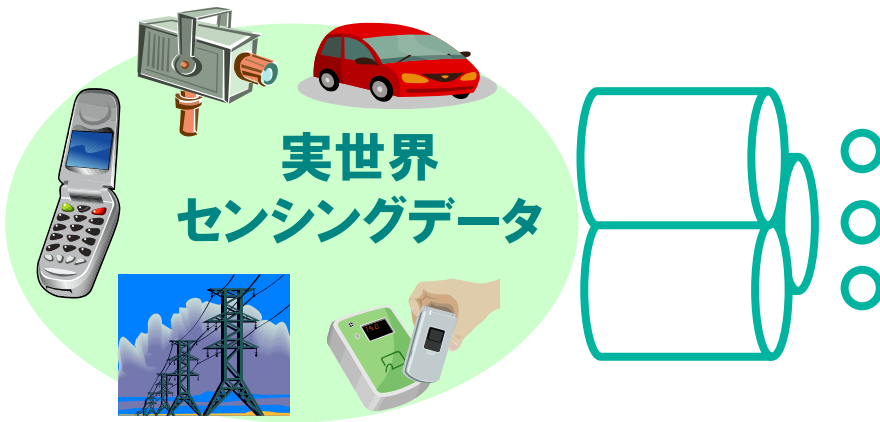
## 大量データの高速処理技術

データが大量でも止まらない、  
さらに増え続けても破綻しない

分散並列処理による  
スケールアウト技術

イベントストリーム  
処理技術

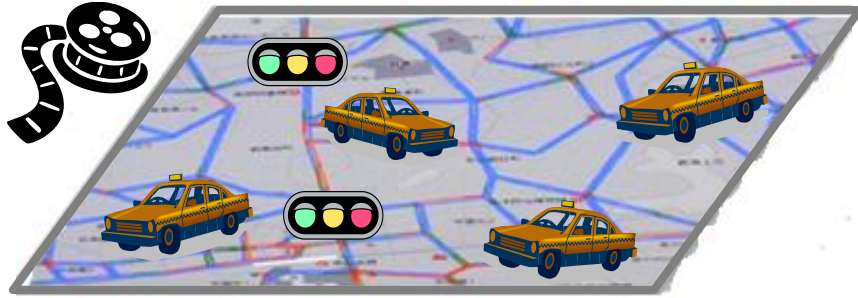
蓄積型データだけでなく、  
ストリーム型データに対しても、処理のスケールアウトが必要



1つのデータサイズは  
比較的小さい

時系列に次々と

# リアルタイム系アプリケーション例(1)交通情報サービス



多数の車の位置情報(GPS)をリアルタイムにモニタリングして渋滞状況を判断、今後の状況も推定して車毎の最適経路も案内(あるいは信号機を制御)

車数8000万台(国内登録車数)から位置情報を毎分1回取得するとしたら、**要求される処理性能は133万イベント/秒**



リアルタイム処理できれば、細街路まで数十秒毎に更新できるのに(上図)、処理が遅いと、幹線道路だけでも数分毎にしか更新できない(下図)



# リアルタイム系アプリケーション例(2)携帯向け情報配信



携帯ユーザの位置情報(GPS)をリアルタイムにモニタリングしてユーザの状況や嗜好を推定、マッチしたお薦め情報や広告・クーポン等をユーザの携帯に配信

ユーザ5000万人(国内携帯ユーザ数の半分)から位置情報を毎分1回取得するとしたら、**要求される処理性能は83万イベント/秒**

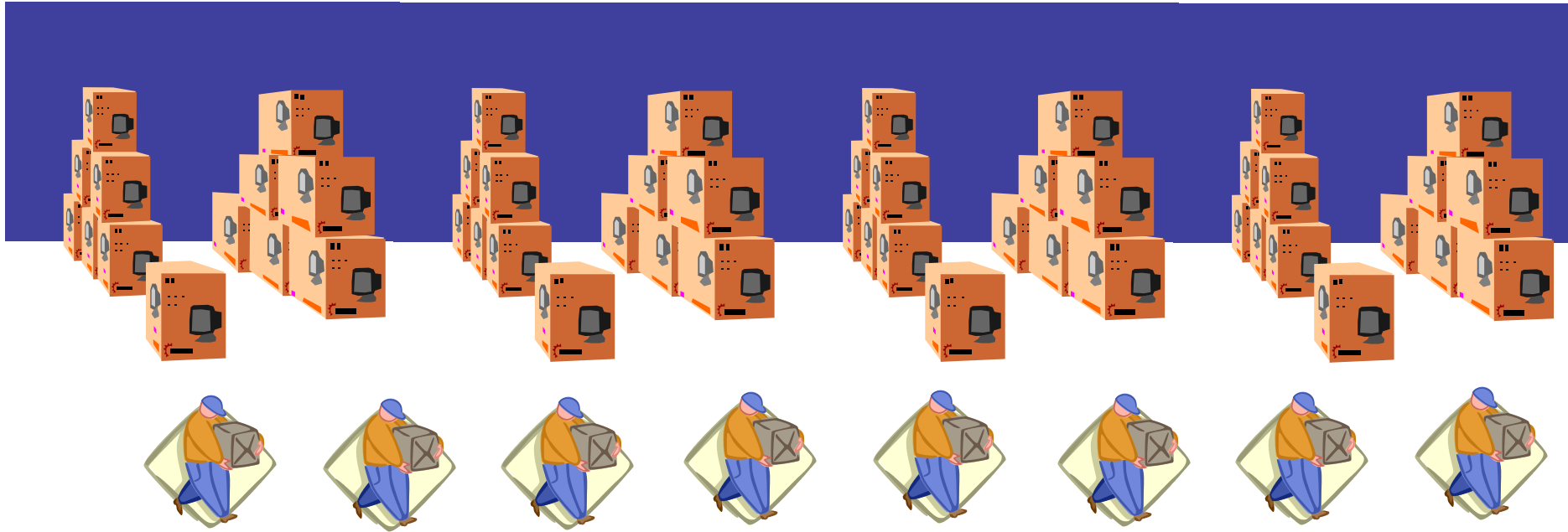


処理が遅いと、終電の知らせが終電が行ってしまった後に届く・・・等

家庭の電力モニタリングの場合： 13万イベント/秒

農作地の状況センシング： 20万イベント/秒

# イベントストリーム処理のスケールアウトが必要



第1部で紹介したのは蓄積型データに対する分散並列処理技術



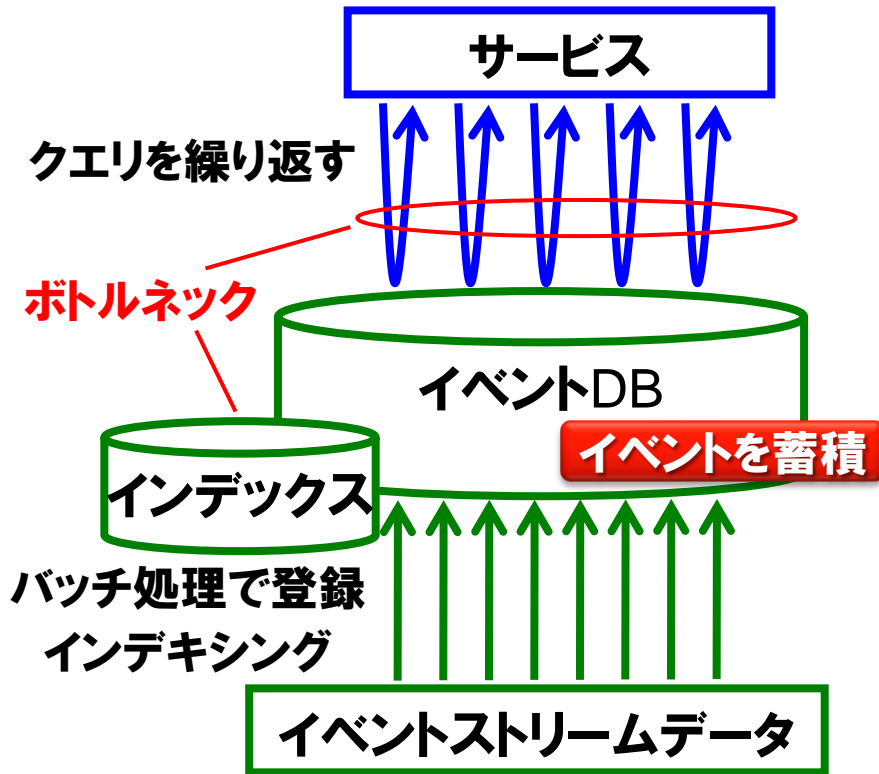
イベントストリーム(ストリーム型データ)に対する分散並列処理が必要

要求される性能 > 100万イベント/秒

# イベントストリーム処理には発想の転換が必要！

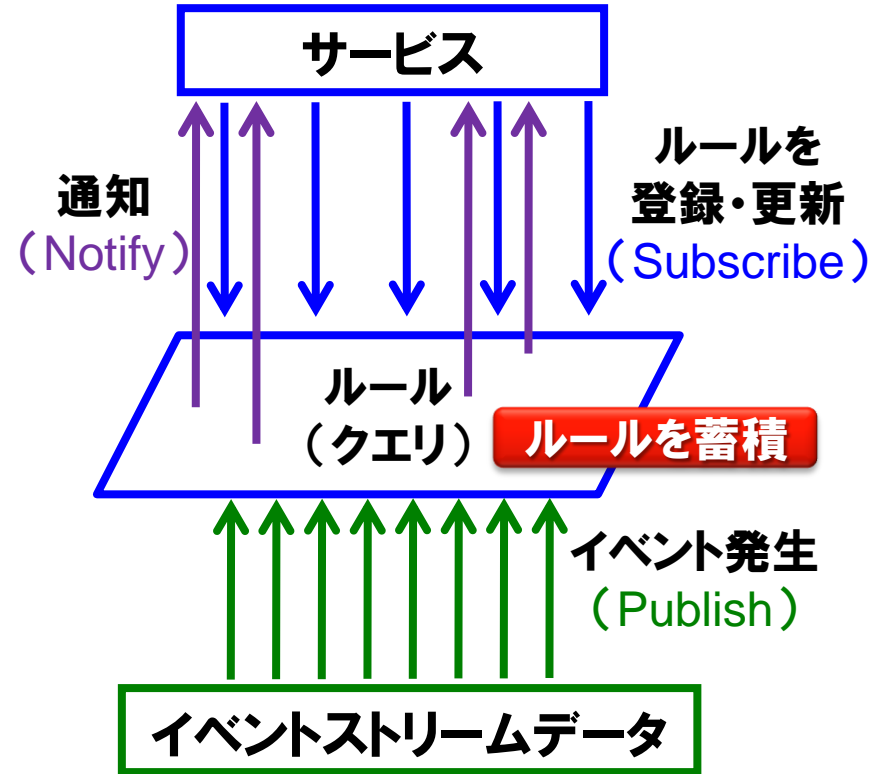
## 蓄積型処理方式(従来)

ストリーム処理ではボトルネックが発生



## Pub/Sub型処理方式

リアルタイムストリーム処理が可能



# ビッグデータストリーム処理基盤(NEC)

複合ルールとイベントの  
マッチング処理を  
スケールアウト化  
(ボトルネック解消)

サーバ16台規模で  
250万イベント/秒を達成

オン・ザ・フライ処理だけでは  
高度な分析に限界  
(深い分析と迅速な  
アクションの両立)

並列CEP  
(CEP=複合イベント処理)

単純ルール(1段目)と複合  
ルールを(2段目)多段に分解

複合ルールは共通イベントを  
含むものが集まるように分散



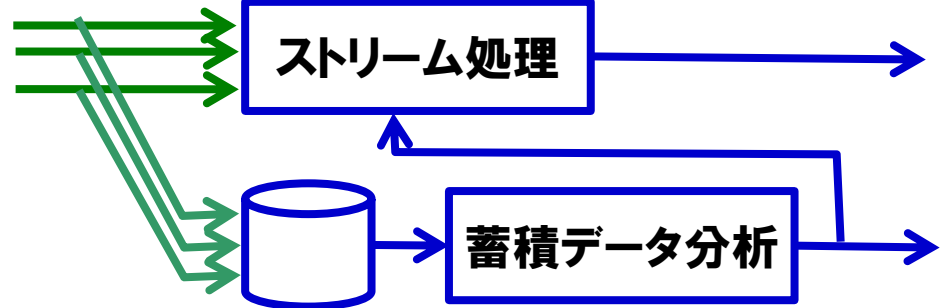
Rule 1 (A & B, Action-X)

Rule 2 (B → C, Action-Y)

Rule 3 (D → E, Action-Z)

ストリーム型・  
蓄積型統合

イベントのストリーム処理と  
並行してイベントを蓄積、  
過去から現在まで蓄積した  
イベント系列からパターン学習



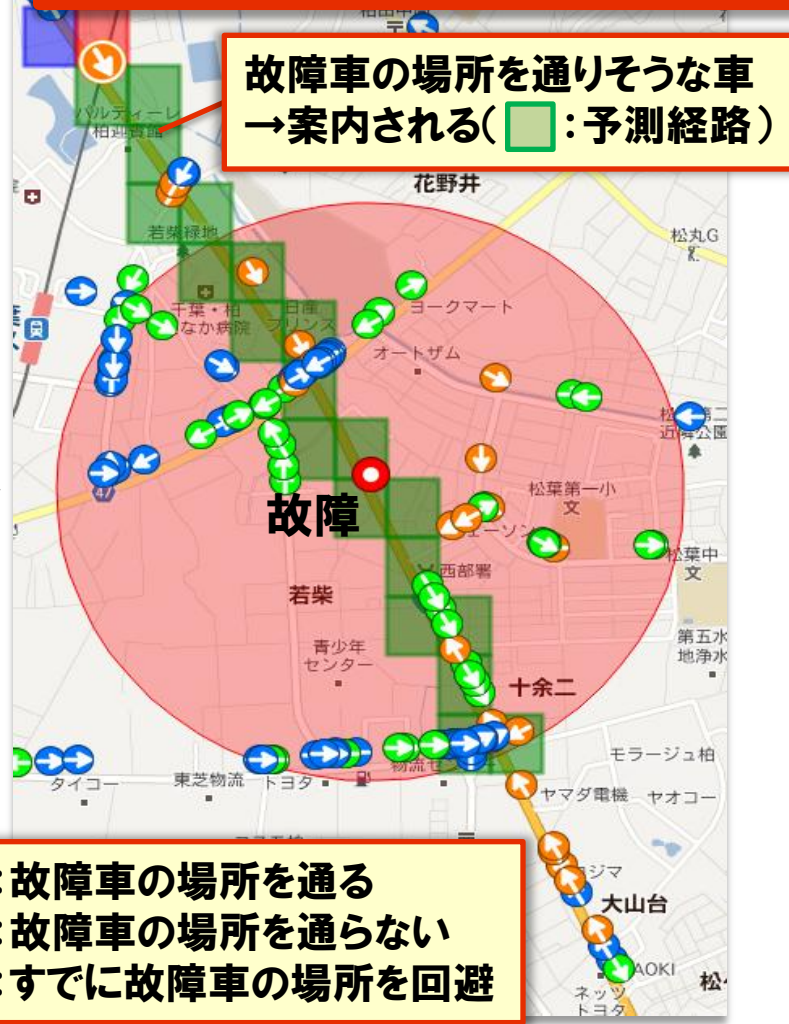
# ビッグデータストリーム処理基盤(NEC)の応用



**Before** 位置情報のみで交通案内  
(故障車から一定範囲内の車に一律に案内)






**After** 予測経路を使って交通案内  
(故障車場所を通る車にのみ案内)



複雑な  
処理を  
高速に

故障車の場所を通りそうな車  
→案内される(□:予測経路)

-  :故障車の場所を通る
-  :故障車の場所を通らない
-  :すでに故障車の場所を回避

## 第4部のまとめ

実世界データに対するリアルタイムアクションを  
可能にするには、**100万イベント/秒を超える  
高速なイベントストリーム処理が必要**

蓄積型データだけでなく、ストリーム型データも扱う  
(分散並列処理によるスケールアウトは共通アプローチ)

**深い分析とリアルタイムアクションを両立させるように統合**

# おわりに ～ ビッグデータって、どの企業もやっているのでは？

**企業の研究所の役割： 他社にはないダントツ技術で、  
自社事業を大きく成長させる、大きな成長事業を創り出す**

- 皆が同じような技術&事業で成長したのは昔話
- グローバル競争下で生き残るのはトップ企業
- いまトップの事業でも永遠には続かない

⇒ **企業毎に狙いどころは自ずと異なってくる**

## NECのダントツ技術と狙いどころ

- 映像解析(顔認証・指紋認証等でトップ性能)
  - 高速イベントストリーム処理(業界トップクラス)
  - 異種混合学習、インバリアンツ解析  
(複雑な実世界データのマイニング技術) 等々
- 今回紹介
- 今回省略  
(付録参照)

# 【参考】顔認証技術(NEC): 性能ベンチマーク結果

MBE (Multiple Biometrics Evaluation) 2010

- 米国NIST(国立標準技術研究所)が4年毎に開催する世界最高権威のベンチマークテスト、顔画像認証技術の実用性能評価
- 世界中の有力ベンダーが参加、NECがダントツ1位の性能達成

|   | NEC<br>(ダントツ1位) | 2位企業      | 各性能が重視される応用例                          |
|---|-----------------|-----------|---------------------------------------|
| <b>認証精度</b><br>(エラー率)                       | 0.3%以下          | 1/10 2~3% | 出入国システム、利用者ログイン<br>(高精度が要求される用途)      |
| <b>経年変化</b><br>(8年でのエラー率)                   | 3%              | 1/4 12%   | パスポート認証、ブラックリスト照合<br>(10年前の写真との照合)    |
| <b>多人種対応</b><br>(黒人・白人・アジア・<br>アメリカンインディアン) | 平均4%<br>全人種でトップ | 1/2 平均8%  | 出入国システム、国際空港内監視<br>(多様な人種が通過する場所での照合) |
| <b>顔向き</b><br>(25度でのエラー率)                   | 7%              | 1/3 20%   | 街頭・公共施設監視システム<br>(監視カメラによる照合)         |
| <b>処理時間</b><br>(160万名)                      | 0.3秒            | 1/6 1.7秒  | 国民ID、ブラックリスト照合<br>(リアルタイム性が要求される用途)   |



---

# 付録

# ガジル 画像認識サービス GAZIRU (NEC)

スマートフォン等のカメラ付き端末で撮影すると、  
その被写体に関連する情報が提示される



【NEC Online TV(動画)】 [http://www.nec.co.jp/ad/onlinetv/ja/society/gaziru\\_h.html](http://www.nec.co.jp/ad/onlinetv/ja/society/gaziru_h.html)

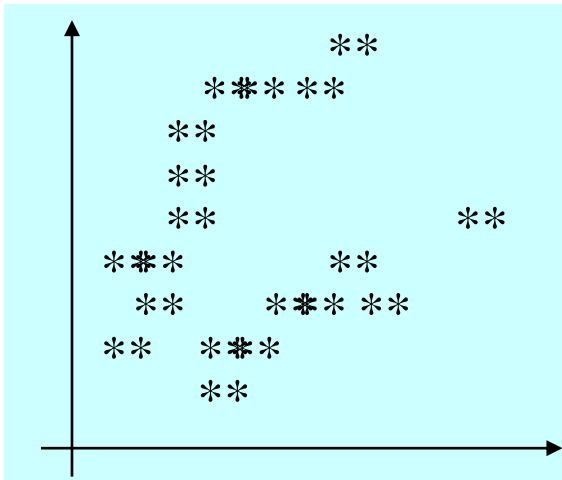
【NEC発見チャンネル】 <http://www.nec.co.jp/profile/mitatv/iexpo2012/01/index.html>

【サービス紹介】 <http://jpn.nec.com/solution/cloud/gazou/>

# 異種混合学習技術 (NEC)

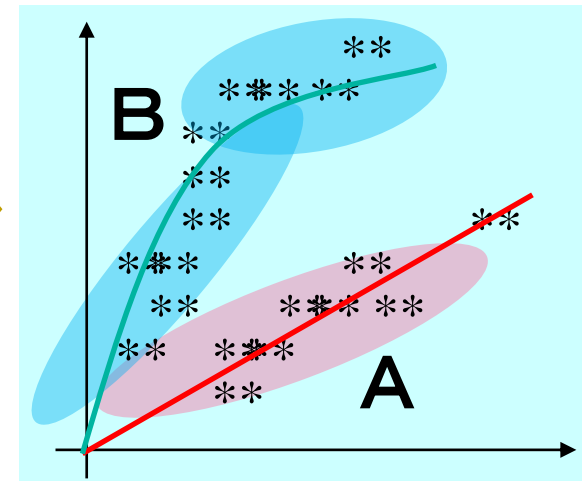
ビッグデータ中に隠れた多数の規則性を自動で発見

従来は複雑すぎて困難だったケースでも高精度な予測・異常検出が可能



異種混合データ

異種混合学習技術では  
こんなケースでも自動的&  
高精度に規則性を発見!



- ①何種類が混ざっているか?
- ②どのように分けるとよいか?
- ③それぞれどんなモデルか?
- ・・・これらのどれもわからない

- ✓ 従来技術では、膨大な組み合わせを延々と探索し続けることになる
- ✓ そのため従来は、データ分割を専門家の勘(最良解とは限らない)に任せるしかなかった

↑  
解決!

# インバリアンツ解析技術SIAT (NEC)

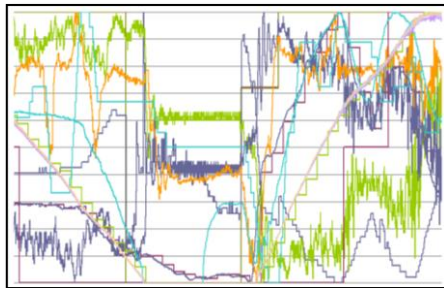
データ間の関係の**インバリアンツ**(Invariants: **不変性**)に着目し、  
観測データのみから自動解析によって、システムをモデル化

- 複雑な事前設定やドメイン知識が不要(従来は専門家の知識が必要だった)

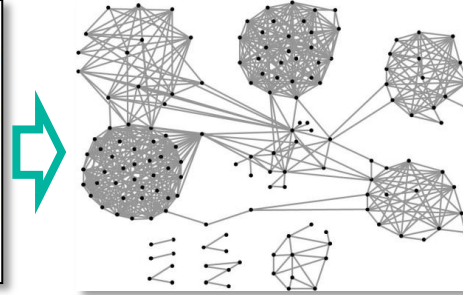
現在の情報とモデルからの予想値を比較し、「**いつもと違う**」動きをしているかどうかに基づき、異常や予兆を早期に発見

- 「いつもと違う」挙動の発生箇所も特定可能

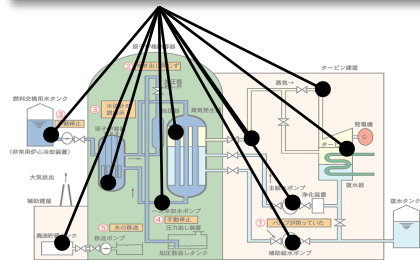
観測点からの  
時系列データ群



システム全体の  
インバリアンツ・モデル  
(依存性モデル)

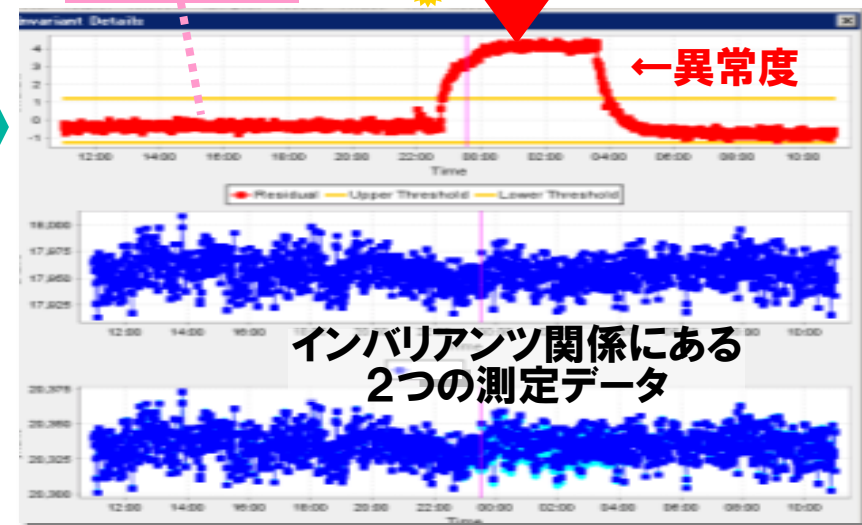


数百～数千の観測点  
例: プラントのセンサー群)



いつもと  
同じ

いつもと違う!



# より詳しく知るための参考資料(1)

■ 城田真琴 「クラウドの衝撃」 東洋経済新報社(2009)

■ 城田真琴 「ビッグデータの衝撃」 東洋経済新報社(2012)

■ 伊藤直也・田中慎司 「Web開発者のための大規模サービス技術入門」 技術評論社(2010)

■ 西田圭介 「Googleを支える技術 巨大システムの内側の世界」 技術評論社(2008)

■ 「情報処理」(情報処理学会誌) 第51巻 第12号 (2010年12月)  
特集:画像認識技術の実用化への取り組み

■ 「NEC技報」第63巻 第3号 (2010年9月) パブリックセーフティを支える要素技術・ソリューション特集号

- <http://www.nec.co.jp/techrep/ja/journal/g10/n03/g1003mo.html>
- 特に「人物行動を把握する画像解析技術と適用例」、「顔認証技術とその応用」

## より詳しく知るための参考資料(2)

### 「NEC技報」第64巻 第3号 (2011年9月) 映像ソリューション特集号

- <http://www.nec.co.jp/techrep/ja/journal/g11/n03/g1103mo.html>
- 特に「NECの映像技術への取り組み」、「人の行動を「見える化」する動線解析技術と活用例」

### 「NEC技報」第64巻 第4号 (2011年11月) Network of Things特集号

- <http://www.nec.co.jp/techrep/ja/journal/g11/n04/g1104mo.html>
- 特に「M2Mサービスプラットフォームにおける大規模リアルタイム処理技術」

### 「NEC技報」第65巻 第2号 (2012年9月) ビッグデータ活用を支える基盤技術・ソリューション特集号

- <http://www.nec.co.jp/techrep/ja/journal/g12/n02/g1202mo.html>
- 特に「高倍率・高精細を実現する事例ベースの学習型超解像方式」、「WebSAMの分析技術と応用例～インバリエント分析の特長と適用領域～」、「ビッグデータ時代の最先端データマイニング」

Empowered by Innovation

**NEC**