

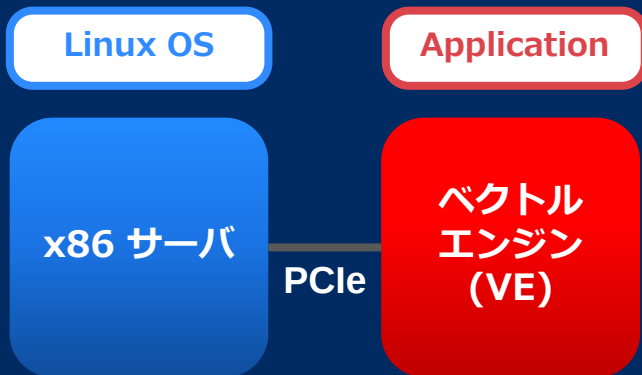
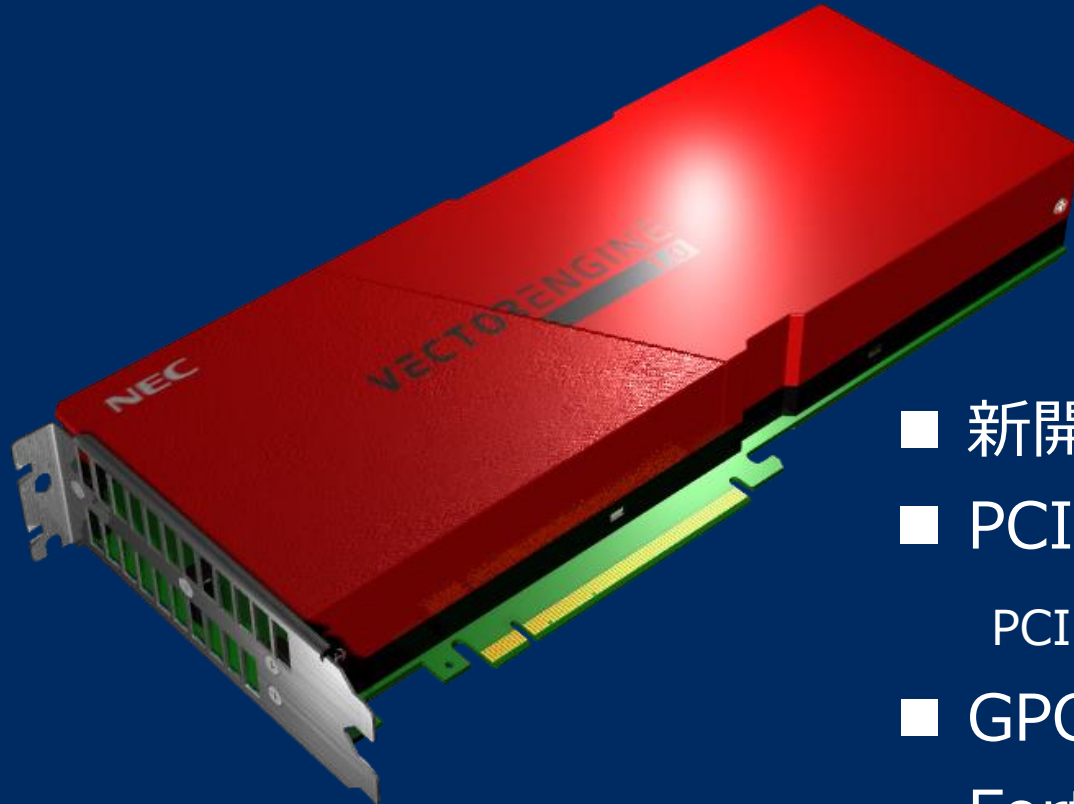
新ベクトルスーパーコンピュータ

SX-Aurora TSUBASAのAI/BD適用と事例について

2018年 12月13日

日本電気株式会社

PCIeカード型ベクトルエンジン (VectorEngine)



- 新開発ベクトルプロセッサ
- PCIe規格準拠
 - PCI Express Gen3. x16
- GPGPUと異なる実行モデル、
Fortran/C/C++の標準環境
- 演算性能：
 - 2.45TF(倍精度), 4.91TF(単精度)
- メモリ帯域 1.22TB/s
- メモリ容量 48GB

SX-Aurora TSUBASA

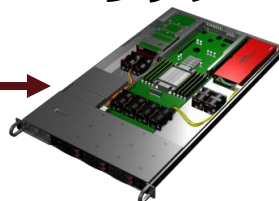
SX-Aurora
(従来のスパコン)



タワー



ラック



POINT
1

高性能

ベクトル技術で大量データの一括処理を実現し、Xeon比10倍以上の性能を実現

POINT
2

使いやすさ

C/C++/Fortranで記述されたプログラムをコンパイルして実行するだけで高い性能を享受可能

POINT
3

フレキシブル

お客様ニーズに従い、サーバ/エンジン構成選択可能、C/P最適化、適応市場拡大

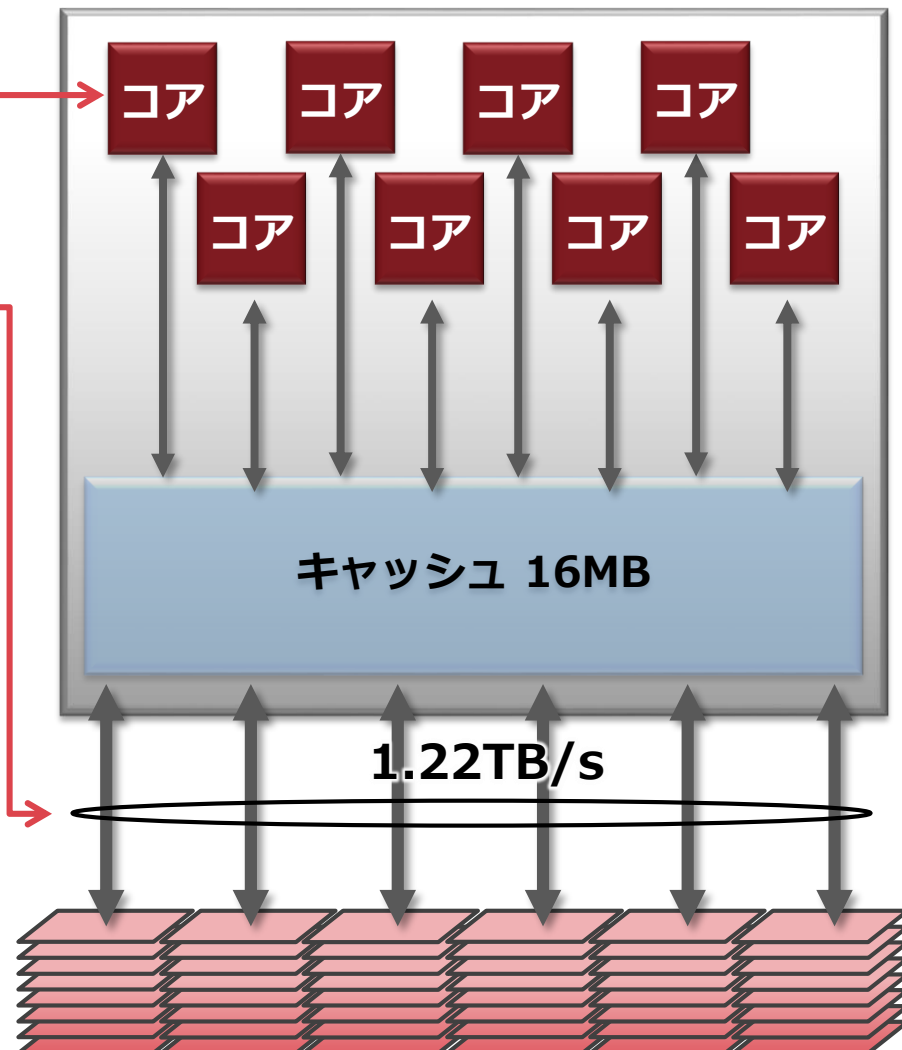
ベクトルプロセッサ

- ✓ **世界最速コア**
307GFlops (DP)
614GFlops (SP)
- ✓ **世界最速のデータアクセス性能**
1.22TB/s
- ✓ **テクノロジー**
世界初HBM2 x6実装

世界初



世界初となるCPUと6個の3次元積層メモリHBM2
搭載技術をTSMC社と共同開発

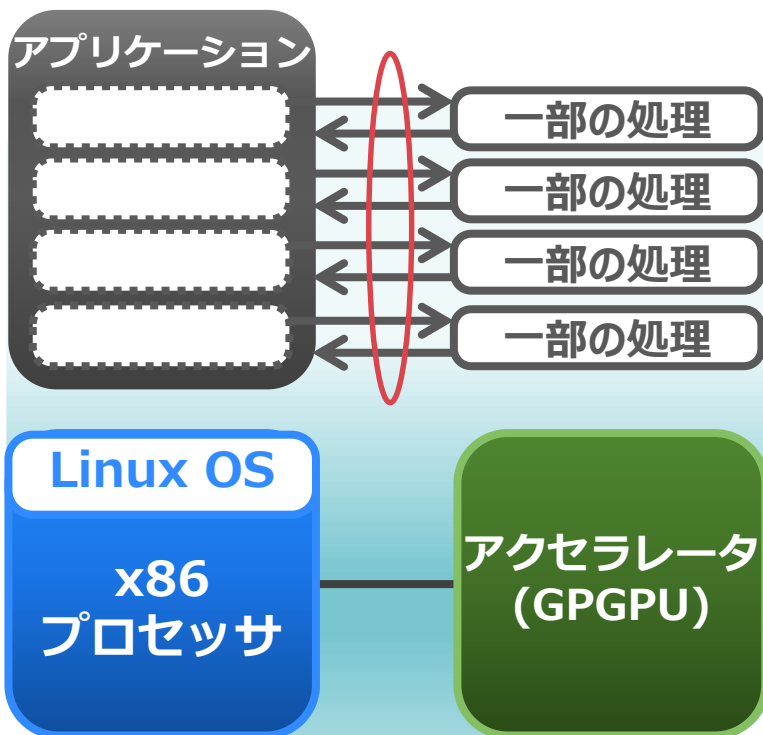


アプリケーションの実行方式

アクセラレータ型（GPGPU）



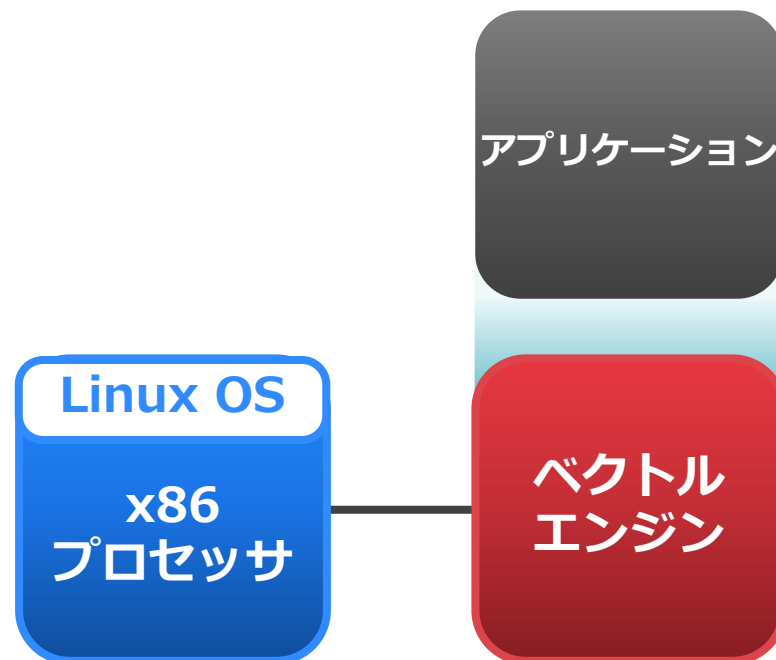
データ移送が頻発すると
性能ボトルネックとなる



SX-Aurora TSUBASA



アプリケーションを丸ごと
ベクトルエンジン上で実行するため
ボトルネックが発生しない



プログラミング環境・アプリケーション実行環境

プログラミング環境



```
$ vi sample.c  
$ ncc sample.c
```

コンパイラ

自動ベクトル化機能

自動並列化機能

Fortran: F2003, F2008 (一部)

C: C11

C++: C++14

OpenMP: OpenMP4.5

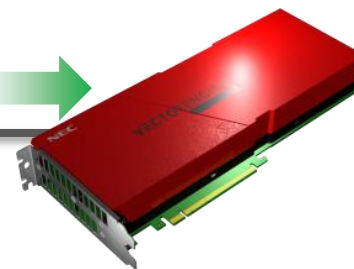
MPI: MPI3.1

アプリケーション実行環境



```
$ ve_exec ./a.out
```

x86



アプリケーション実行

幅広いラインアップ

ベクトルエンジンを搭載したサーバを組み込み用途から大規模データセンターシステムまで幅広く提供

データセンタモデル

データセンタ、計算センタでの巨大処理

例：大規模AI・ビッグデータ、
大規模シミュレーションなど

オンサイトモデル

製品設計のシミュレーション、
AI・ビッグデータ利用

例：流体シミュレーション、
大規模需要予測 など

エッジモデル & 組み込みソリューション

AI・ビッグデータ・システム組み込み用途

例：製品検査、医療画像処理、需要予測、
レーダ解析など

データセンタモデル(水冷)

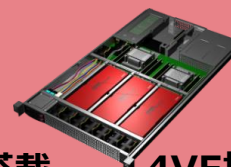
64VE搭載



オンサイトモデル



2VE搭載



4VE搭載



8VE搭載

エッジモデル

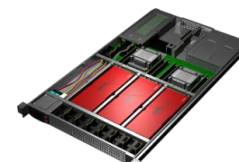
1VE搭載



組み込みソリューション



ユーザー専用システム



SX-Aurora TSUBASA

AI/BD領域への取り組み

- 統計型機械学習ミドルウェア : Frovedis
- エコシステムの構築
- ユースケース

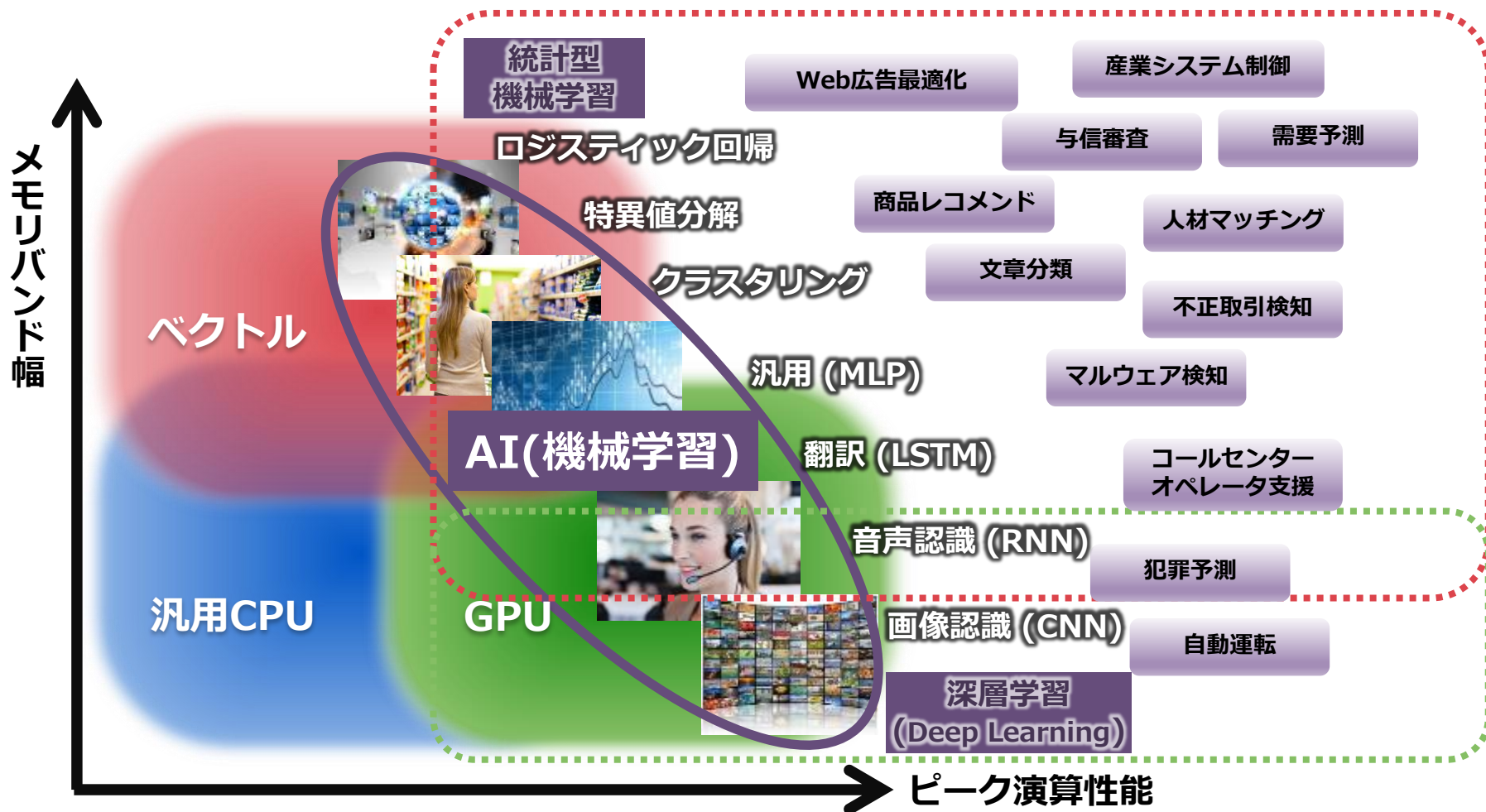
利用領域の広がり

ベクトル技術を核にした高性能シミュレーション(HPC)や AI・ビッグデータ解析技術
が科学技術の進展と産業の高度化を加速し、新しい社会価値を創造する



SX-Aurora TSUBASAの適応領域と応用例

統計型機械学習はメモリバンド幅の広いSX-Aurora TSUBASAが得意とする分野



Sparkの機械学習を高速化するミドルウェア : Frovedis*

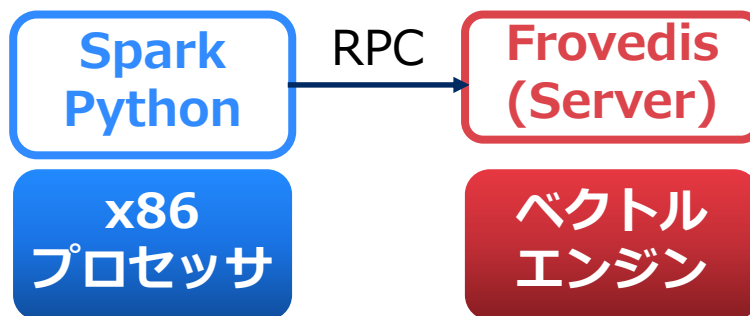
- Spark/Pythonからベクトルを意識せず利用可能なミドルウェア
- SX-Aurora TSUBASAのメモリ性能を生かし、統計型機械学習を高速化

※ FReamework Of VEctorized and DIStributed data analytics

ベクトル化済みの機械学習/行列演算機能を用意(Spark Mllib, 前処理機能)

- 機械学習ライブラリ (Spark Mllib、scikit-learn 互換)
- Data Frame (前処理機能)
- NECでOSS化しており、下記のリンク先で公開中
<https://github.com/frovedis>

➔ **SX-Aurora TSUBASAの性能を、直ぐにフル活用可能**



Frovedisでサポートしている処理

Web広告最適化

需要予測

与信審査

商品レコメンド

人材マッチング

日報/アンケート
分類

不正取引検知

マルウェア検知

統計型機械学習

ロジスティック
回帰

線形回帰

Lasso回帰

Ridge回帰

線形
SVM

特異値
分解

固有値
分解

主成分
分析

協調フィルタリング
(ALS)

ナイーブ
ベイズ

K-means

決定木

word2vec

Factorization
Machines

Data Frame

Filter

Sort

Join

Group
by

基本行列演算

Solve

逆行列

Gemv

SpMV

LU分解

Backed by ScaLAPACK,
LAPACK, BLAS

Least
square

行列積

ビッグデータ向け高速分析プラットフォームを共同開発



2018年10月15日

報道関係各位

日本電気株式会社
Hortonworks, Inc

NECとHortonworks、SX-Aurora TSUBASAを活用した ビッグデータ向け高速分析プラットフォームを共同開発 ～Apache Hadoop/Sparkによる高速なAI処理に対応～

日本電気株式会社(本社:東京都港区、代表取締役執行役員社長 兼 CEO:新野 隆、以下 NEC)と Hortonworks, Inc. (本社:米国サンタクララ、CEO:ロブ・ベアデン (Rob Bearden)、以下 Hortonworks)は、NECの次世代イノベーションプラットフォーム「SX-Aurora TSUBASA」上で、AIにも利用可能なビッグデータ向け高速分析プラットフォームを共同開発します。

2017年9月にビッグデータ活用における大規模分散処理プラットフォーム分野での戦略的協業の強化を発表(注1)した両社は、より幅広い業種や分野でビッグデータの高速かつ高度な分析を可能にするため、両社の強みを活かしたソリューション基盤を、2018年度中に提供予定です。また、今回の共同開発の成果はオープンソース(OSS)として公開を予定しており、多様な環境や様々な業種ごとのユースケースで利用できるようにします。

【共同開発の内容】

- 「SX-Aurora TSUBASA」が、Hadoop/Spark領域での代表的なリソース管理/分散アプリケーション管理基盤である「YARN(ヤーン、注2)」に対応することで、従来のHadoop/Sparkによるビッグデータ解析プログラムが「SX-Aurora TSUBASA」上で管理可能となるため、容易なアプリケーション連携とシステム管理者の負荷軽減を実現します。「YARN」を利用している一般的なシステム構成では、従来(注3)同等の解析を、最大99%少ないサーバーリソースで実行できます。
- NECのミドルウェア「Frovedis(フロベディス、注4)」をSparkに組み込むことにより、「SX-Aurora TSUBASA」上でのSparkを使用したAI(機

1

Hortonworks



NEC

共同発表内容

SX-Aurora TSUBASAのYARNへの対応



- ・**Apache Hadoop/Spark上で「SX-Aurora TSUBASA」のリソースが利用可能**
- ・**NECのAIミドルウェア「Frovedis」と「Data Platform for Hadoop」を組み合わせたソリューションの提供**

2018年10月15日共同プレスリリース

ユースケース：（1）レコメンデーション

クロスレコメンデーション

ECサイト

ショッピング
サイトA

ショッピング
サイトB

閲覧・購入
履歴

サイトAを
訪問、商品検索、
購入

サイトB訪問時に
同じ／類似商品を
おすすめ

顧客

Auroraにより 56.8倍高速化

分析手法：特異値分解（SVD）
Aurora：1VE 8core
比較対象：Xeon Gold 6126
1socket 12core

保険商品販売

顧客データベース

A	10	...				
B	30	...				
C	5	...				
D	50	...				
E	15	...				
F	20	...				

事故歴、
傷病履歴、
など

学習

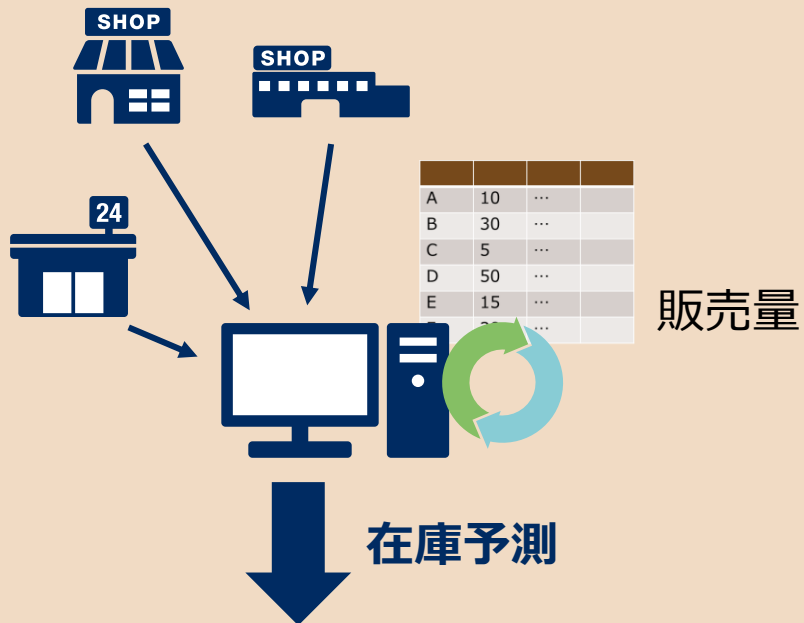
保険商品を
検索

個人に最適な
保険商品を
おすすめ

顧客

ユースケース：（２）需要予測

販売予測

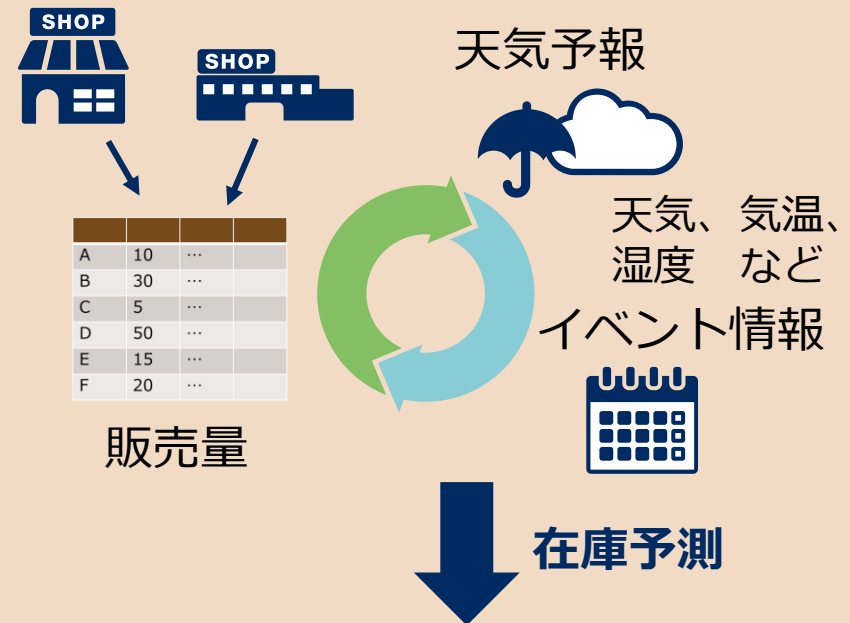


在庫管理責任者

Auroraにより 10.8倍高速化

分析手法：回帰木（Regression tree）
Aurora：1VE 8core
比較対象：Xeon Gold 6126
1socket 12core

販売予測（廃棄ロス削減）



在庫管理責任者

 **Orchestrating a brighter world**

NEC

- ・本資料に記載の社名、製品名、ロゴは各社の商標または登録商標です。
- ・本資料に記載の製品仕様等については予告なく変更する可能性があります。