

Hewlett Packard  
Enterprise

実践ガイド著者が語る！

# 機械学習/AIのためのプラットフォーム ～GPUやコンテナの活用～

日本ヒューレットパカード株式会社  
HPE認定オープンソース・Linuxテクノロジーエバンジェリスト

古賀政純 @masazumi\_koga

2018年11月6日(火)

# 古賀 政純(こが まさずみ)

# オープンソース歴: 22年



## 保有認定資格

- Docker Technical Sales Prof.
- Docker Sales Prof.
- Hadoop (CCAH)
- Red Hat OpenStack
- Red Hat Certified Engineer
- Red Hat Certified Virt. Admin.
- Novell Certified Linux Prof.
- HPE ASE Cloud Architect
- HPE Server Sol. Architect
- HPE Product Certified – OneView
- EXIN Cloud
- 等

## 略歴

- ‘94~’00: AI, ニューラルネット(大学院)
- ‘00: 入社、SE部配属(名古屋)
- ‘02: 科学技術計算システムSI担当
- ‘03: Linux技術SE
- ‘08: ブレード/仮想化技術SE
- ‘09: オープンソース系SE
- ‘10: HPパートナー企業担当SE
- ‘11: オープンソース系SE
- ‘14: HPエバンジェリスト

## 主な受賞歴

- 日本HP社長賞(FY05)
- 日本HP Presales特別賞(FY05)
- 米国HP Linux技術Hall-of-Fame(FY06)
- 米国HP Linux技術Hall-of-Fame(FY07)
- 日本HP Presales特別賞(FY11)
- 日本HP Presales MVP(FY11,12,14, 14Q4)
- 日本HP TCE貢献賞(FY14)

HPE Web  
技術情報執筆

OSS/Linux  
書籍出版

インプレス  
連載記事

寄稿・取材対応  
ITmedia, ITPro

セミナー講師  
出展、実演

趣味



Racing  
Kart

Billiard

Game  
Programming

動物園  
水族館  
めぐり





最先端オープンソース書籍出版の取り組み

# 古賀政純の実践ガイドシリーズ



OS部門 **2**位

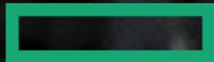


OS部門 **1**位



社内クラウド構築





Hewlett Packard  
Enterprise

2020年のIT

# ビッグデータを軸にした AI・機械学習基盤

# 今は、第3次人工知能ブーム！

## Hadoop+機械学習の書籍



悲しき第2次人工知能ブーム世代  
古賀政純

大学時代の古賀の研究は、  
人工知能(AI)、  
ニューラルネットワーク  
でした...(90年代)

20年の  
時を越えて



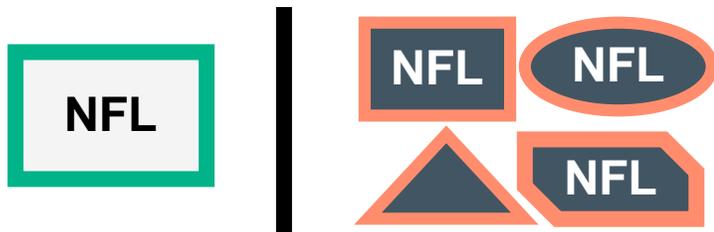
フライトデータ分析、  
迷惑メール分類、  
おすすめ映画タイトルの  
表示など具体例掲載！

2018年5月17日発売

# 知的情報処理の活用例

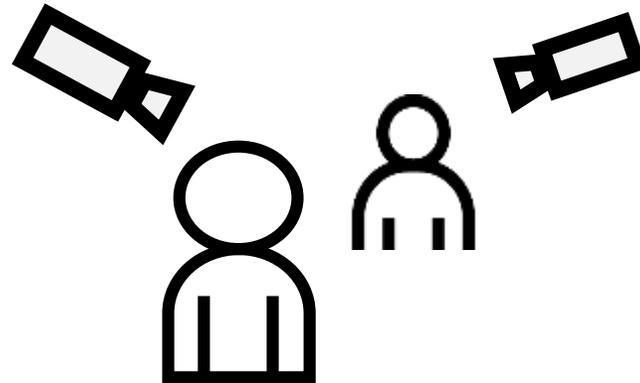
## スポーツチームの ロゴ認識

- 適正使用の識別
- 未許可のロゴの  
使用識別



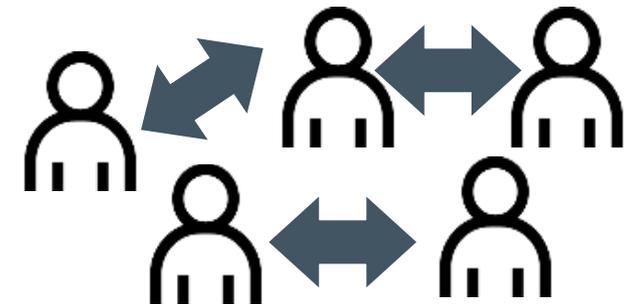
## 22,000台のビデオ 監視システム

- 注意人物の認識
- 不審行動の認識



## ソーシャル解析

- 自分の事を誰が  
話しているか識別

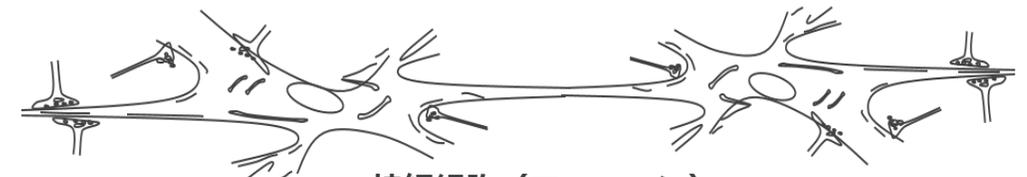


# なぜAI・深層学習では、GPUなのか？



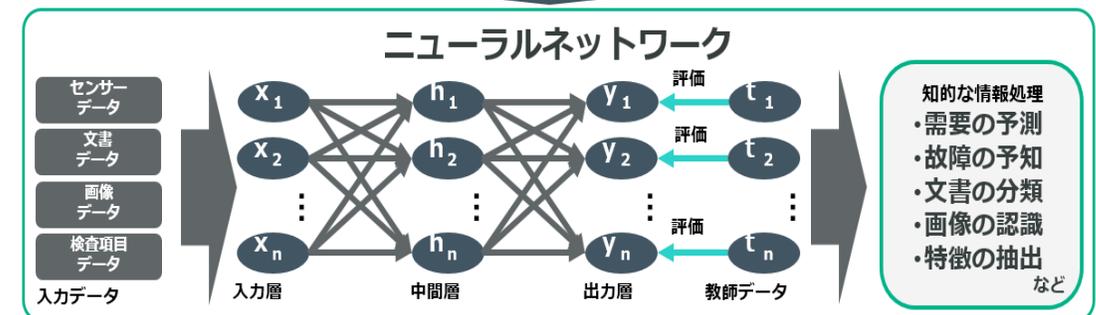
- 行列の積和演算処理
- 3Dグラフィックスの行列演算と全く同じ処理！
- GPUは、CPUの何百倍もの速度計算！

$$D = \begin{pmatrix} A_{0,0} & A_{0,1} & A_{0,2} & A_{0,3} \\ A_{1,0} & A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,0} & A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,0} & A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix} + \begin{pmatrix} B_{0,0} & B_{0,1} & B_{0,2} & B_{0,3} \\ B_{1,0} & B_{1,1} & B_{1,2} & B_{1,3} \\ B_{2,0} & B_{2,1} & B_{2,2} & B_{2,3} \\ B_{3,0} & B_{3,1} & B_{3,2} & B_{3,3} \end{pmatrix} + \begin{pmatrix} C_{0,0} & C_{0,1} & C_{0,2} & C_{0,3} \\ C_{1,0} & C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,0} & C_{2,1} & C_{2,2} & C_{2,3} \\ C_{3,0} & C_{3,1} & C_{3,2} & C_{3,3} \end{pmatrix}$$

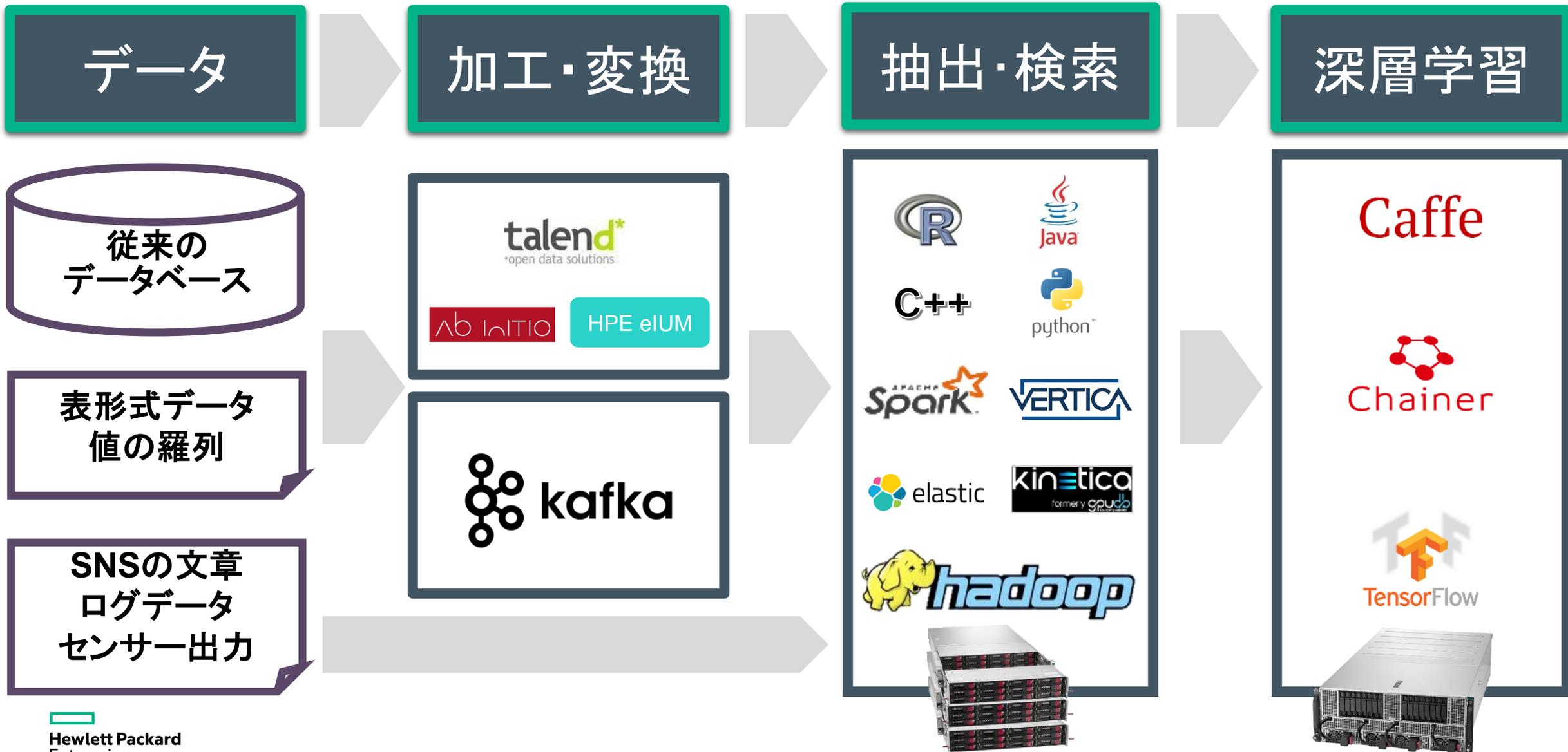


神経細胞（ニューロン）

モデル化

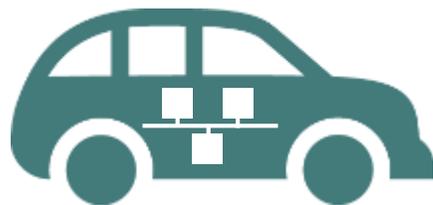


# ビッグデータ＋深層学習の処理の流れ



# センサー+ビッグデータ= データボリュームは100倍！

従来の手動運転



従来のバスシステム

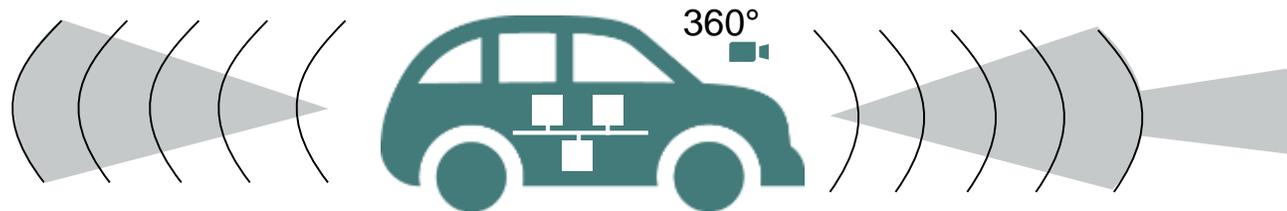
- CAN :Controller Area Network
- LIN :Local Interconnect Network
- FlexRay :車載LAN
- MOST :Media Oriented Systems Transport

15MB/s

比較的単純なデータ形式

~ 500GB / 1台(1日)

自動運転



- 交通状況を認知 (Traffic situation perception)
- フロント、リア、フィールド、リアフィールドの状況認識
  - HD サラウンドビューカメラ
  - 超音波センサー
  - ラダーセンサー
  - レーザースキャナー等

2GB/s

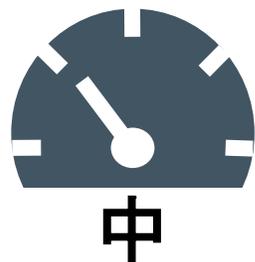
複雑なデータ形式 例) HDビデオストリーム動画

~ 50TB / 1台(1日)

もはや、従来のITツールや古い手法では、  
巨大データボリュームの検索、分析ができない！

# 機械学習の活用例：自動運転

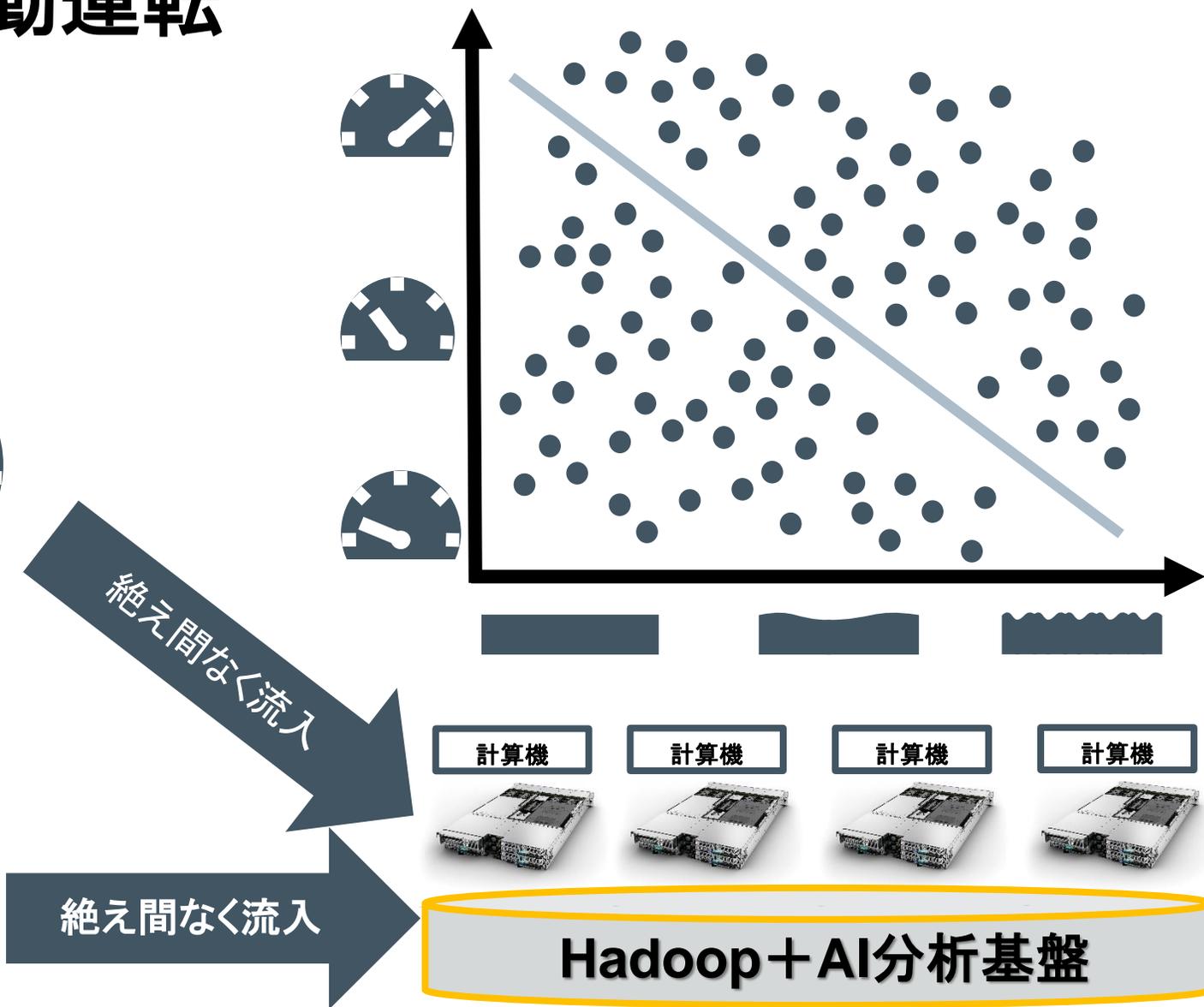
- 車から分析基盤にデータが次々と絶え間なく流入
- 地形データと速度データで学習



速度データ

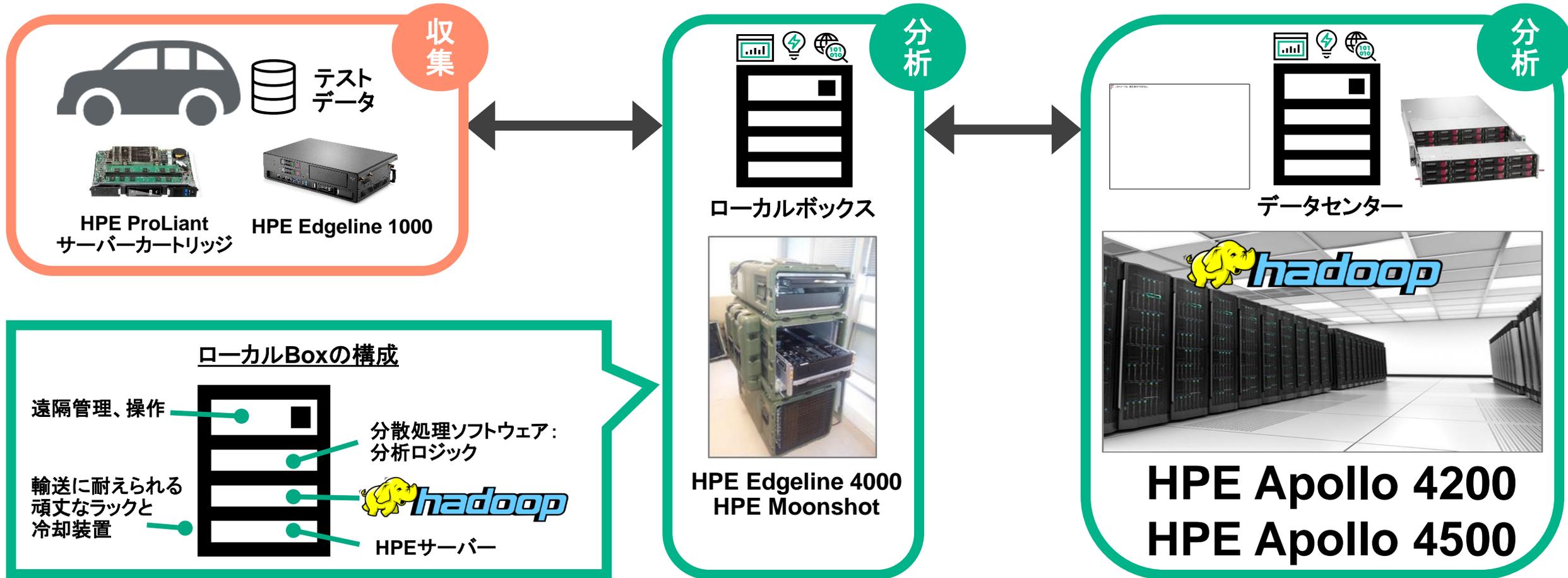


地形データ



# HPEのコンセプト:「Master Data Management (MDM) in a Box」

– データ収集とリアルタイム解析を車内で



ワークステーション1台の処理性能 : 3日

「MDM in a Box」の処理性能 : 1時間!

# 深層学習の活用例



データセンターの消費電力削減！

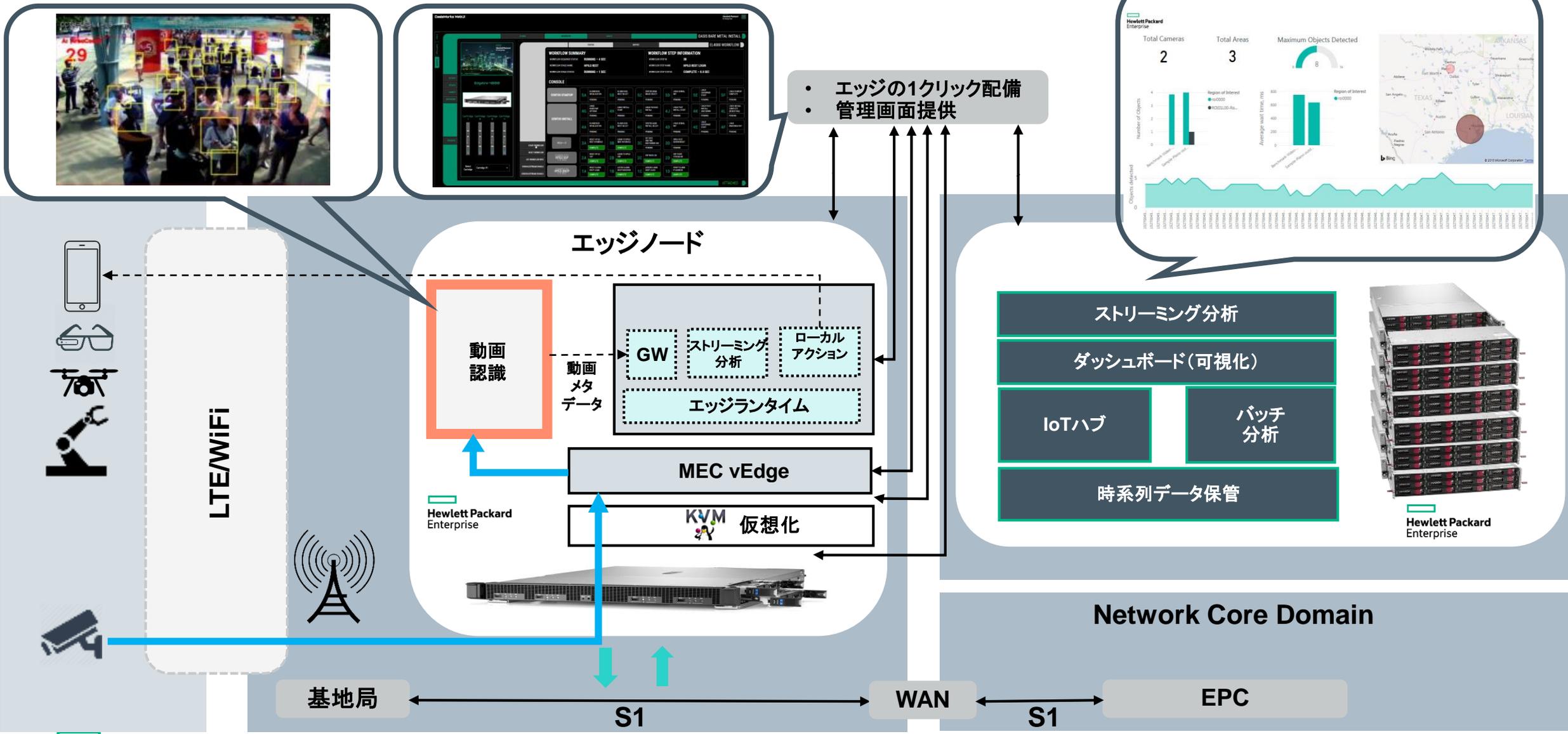
• 気温、電力、冷却ポンプ内のクーラントの流速等のデータを深層学習で解析し、効率化のフレームワークを作成

• 冷却システムの消費電力**40%削減**

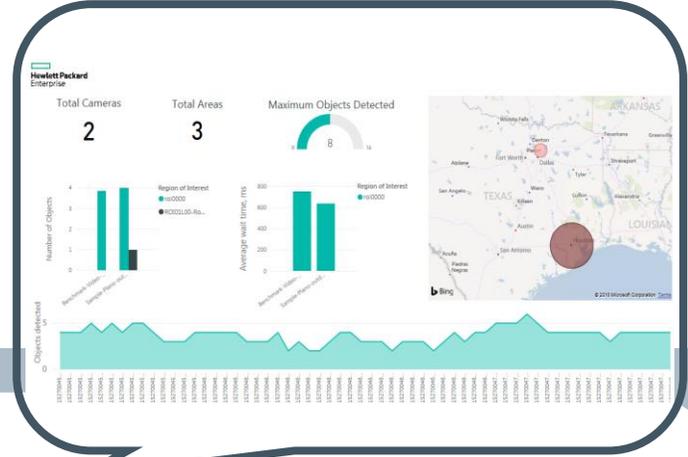


<http://gigazine.net/news/20160721-deepmind-reduce-google-data-center-energy/>

# IoTシステム例：画像認識、動画分析



- エッジの1クリック配備
- 管理画面提供



HPE Customers, HPE Partners and HPE Internal Use Only: 本ドキュメントはHPEのお客様およびHPEのパートナーとHPE社内での利用に限られます。HPEの競合他社への開示、配布は絶対に行わないでください。

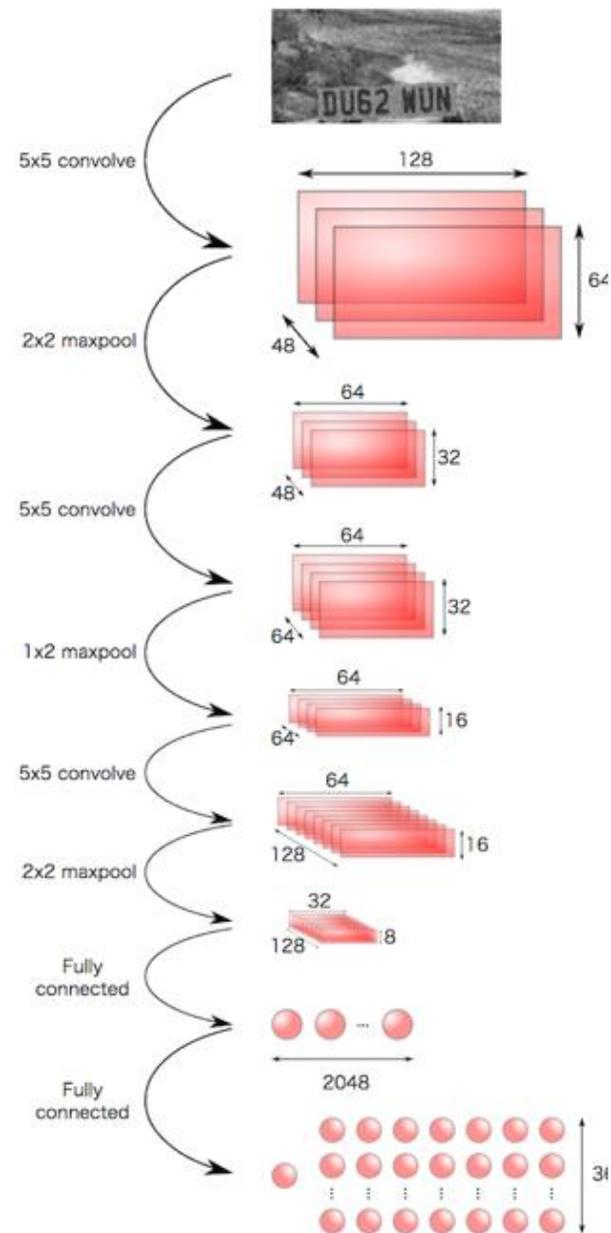
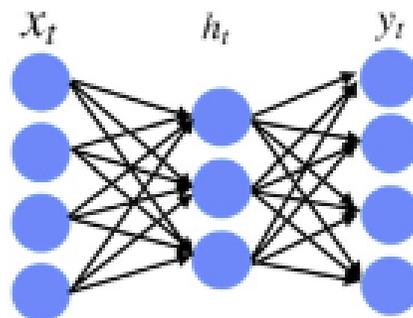
# CNN (Convolutional Neural Network)の活用例

## ナンバープレートの認識



# CNN

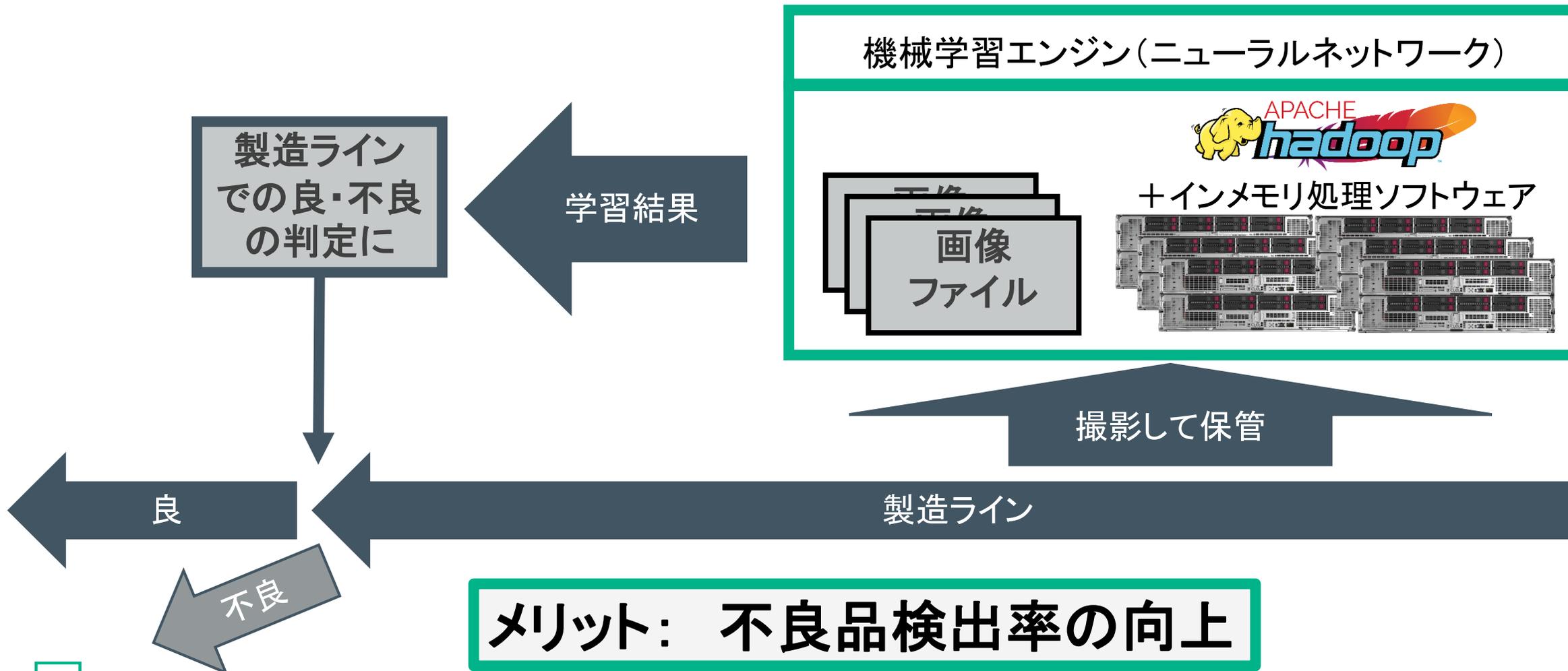
## 特徴を抽出 高精度の認識能力



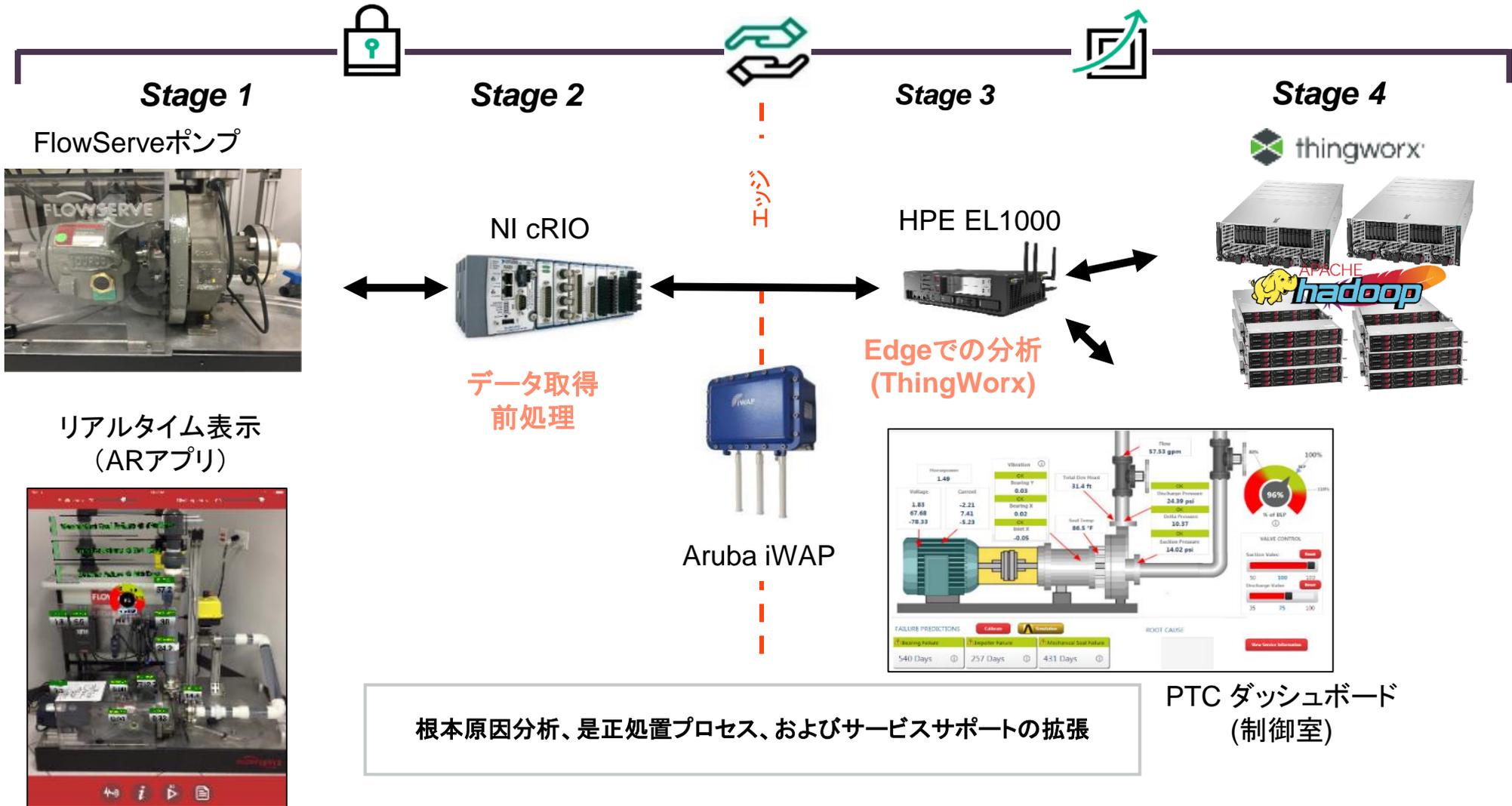
<https://matthewearl.github.io/2016/05/06/cnn-anpr/>

# 機械学習の例：製造ラインでの不良品検出率向上

画像分析：良品の画像と不良品の画像の特徴を抽出し学習

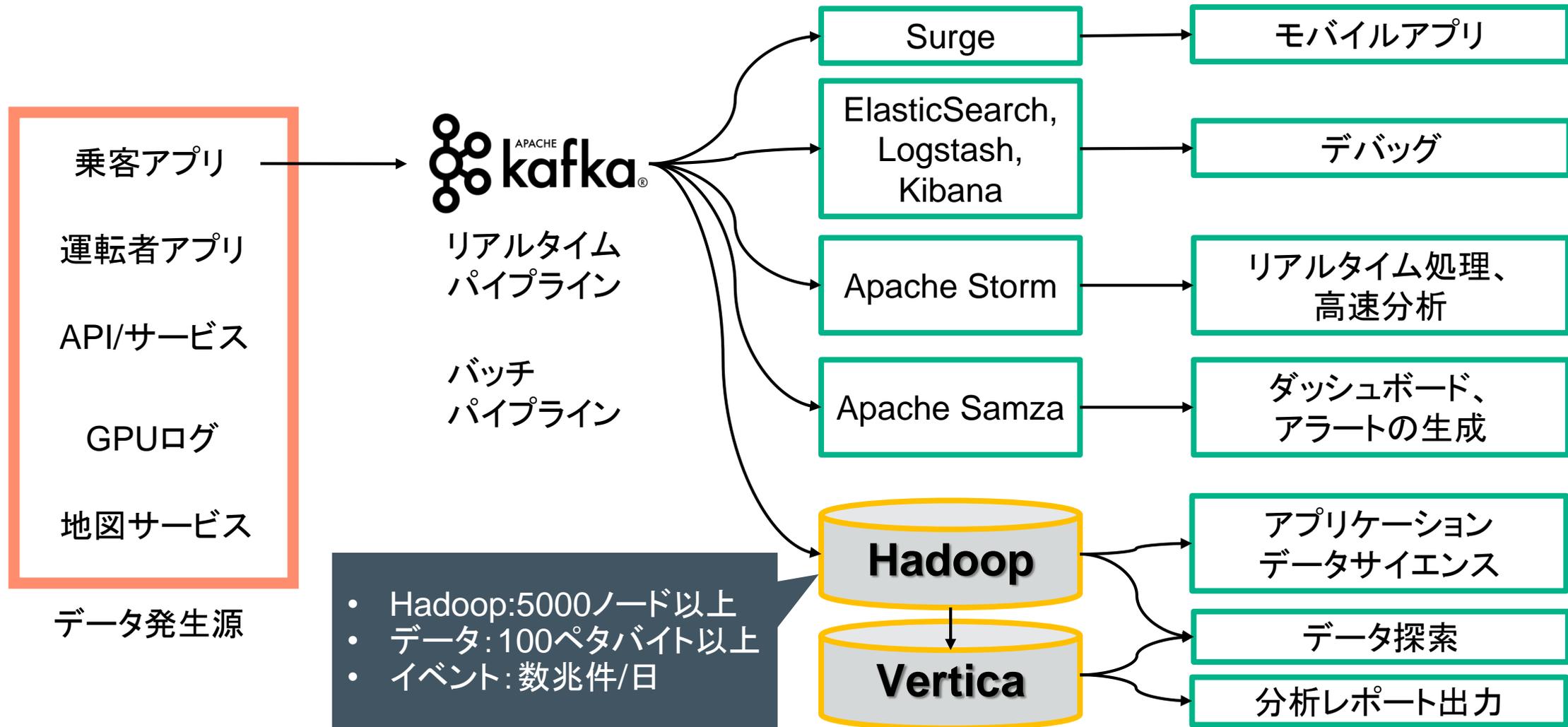


# 工場における設備の予知保全とリアルタイム監視



# 配車サービス会社「Uber」のデータ処理基盤

## リアルタイムでの高速ビッグデータ解析



# Libratus: 不完全情報ゲームを制したAI

- カーネギーメロン大学 (CMU) が開発
- ピッツバーグスパコンセンター (PSC) で稼働

ポーカー: 手札の情報が見えない

不完全情報 = 隠された情報

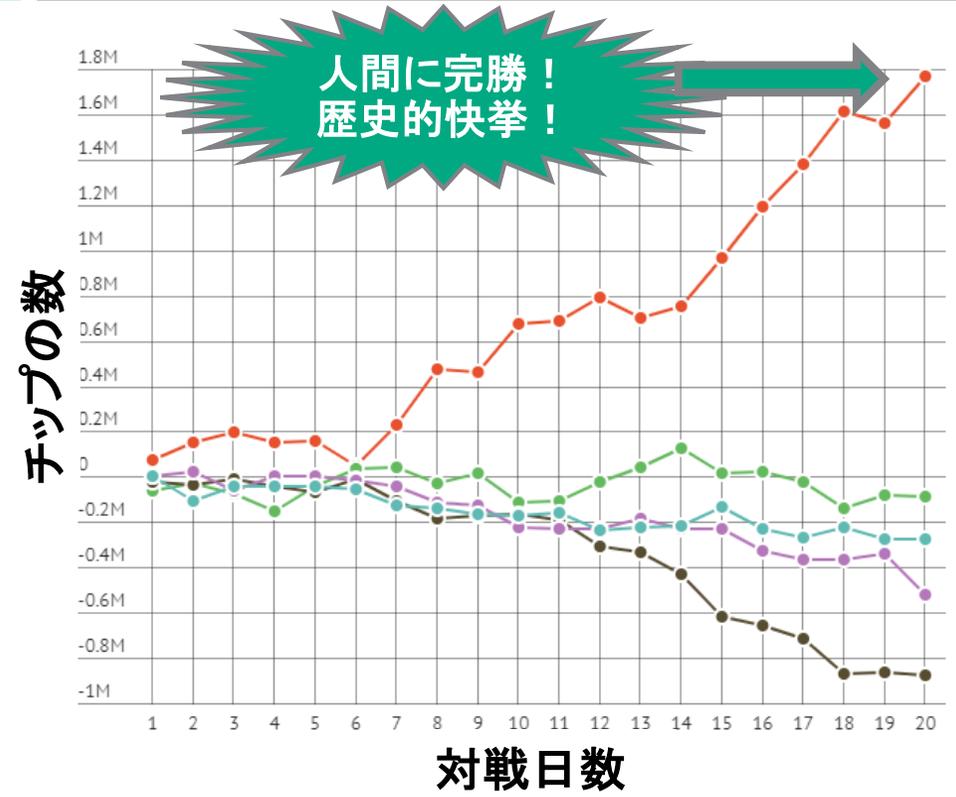
## 不完全情報ゲーム

ポーカーにおける分岐点の数:

### 10の161乗 = 計算が不可能

# • HPEマシン + GPUで構成!

- 戦略の計算:
  - 膨大な計算パワー (1,900万コア時間) が必要だった
  - プレイ終了後の毎夜、人間の新たな戦略に対応

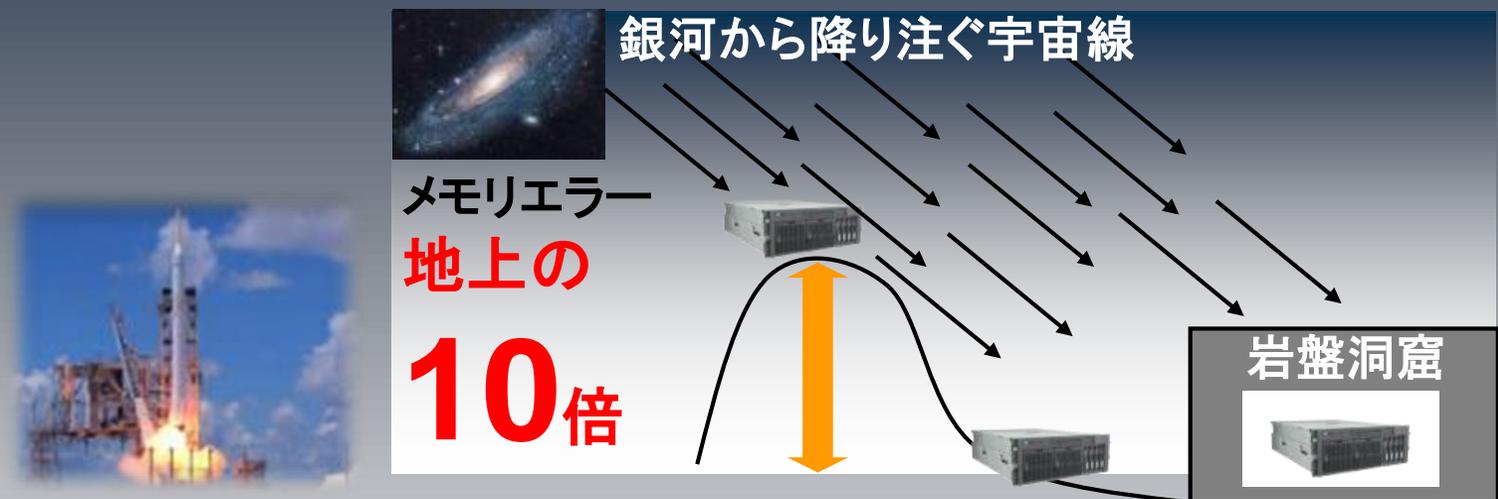


● Libratus ● Dong Kim ● Jason Les ● Jimmy Chou ● Daniel McAulay  
<https://triblive.com/business/technology/13087657-74/cmu-team-publishes-paper-on-how-their-poker-playing-ai-beat-the-best>



PSCのBridgesスーパーコンピューター (HPEサーバーで構成)

# 火星上で稼動するコンピューター Spaceborne Computer



火星上の探査機の様子を認識、分析して  
指令を出すコンピューターを地球上に置くと、

指令が届くまでに掛かる時間：**40**分



火星で稼動する  
スーパーコンピューターが必要に

地球上においても、常に銀河から降り注ぐ宇宙線をコンピューターは浴びている。宇宙線によりメモリ上のデータが勝手に変更されてしまう。

火星までの道程、火星上では大量の宇宙線を浴びる

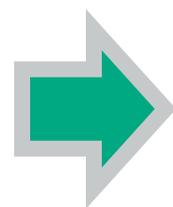
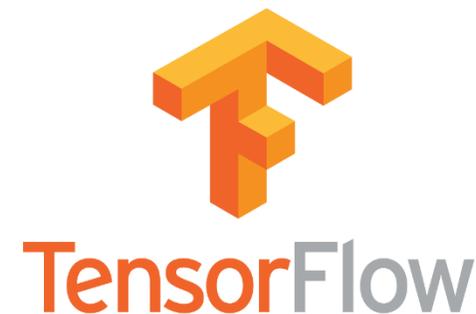


宇宙ステーションにて1年間  
(地球から火星まで往復にかかる時間)の  
稼動実験開始 (2017年8月より)

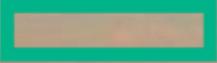


# 深層学習のオープンソース:「TensorFlow」

- Googleが2015/11に公開
- GPUが利用可能
- 実際の利用例
  - 音声認識
  - 翻訳
  - 写真の被写体認識、顔認識
  - ウェブ検索結果の最適化
  - メール仕分け、自動返信文作成
  - 動画配信、広告事業



HPEの技術コンサルティング部隊による  
**TensorFlow構築サービス**もご用意！

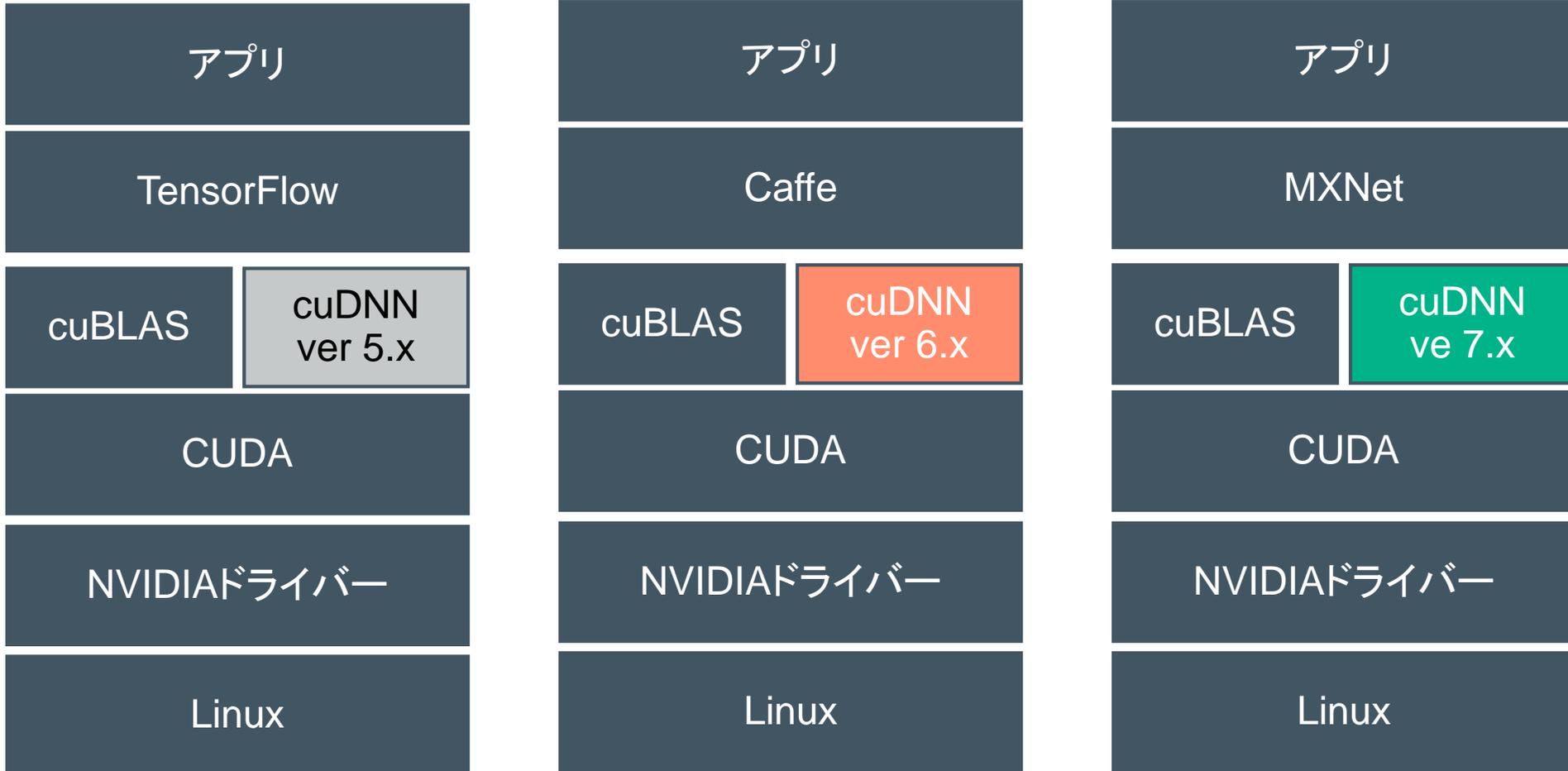


Hewlett Packard  
Enterprise

# 分析基盤における GPUとDockerの利用

# フレームワークの利用

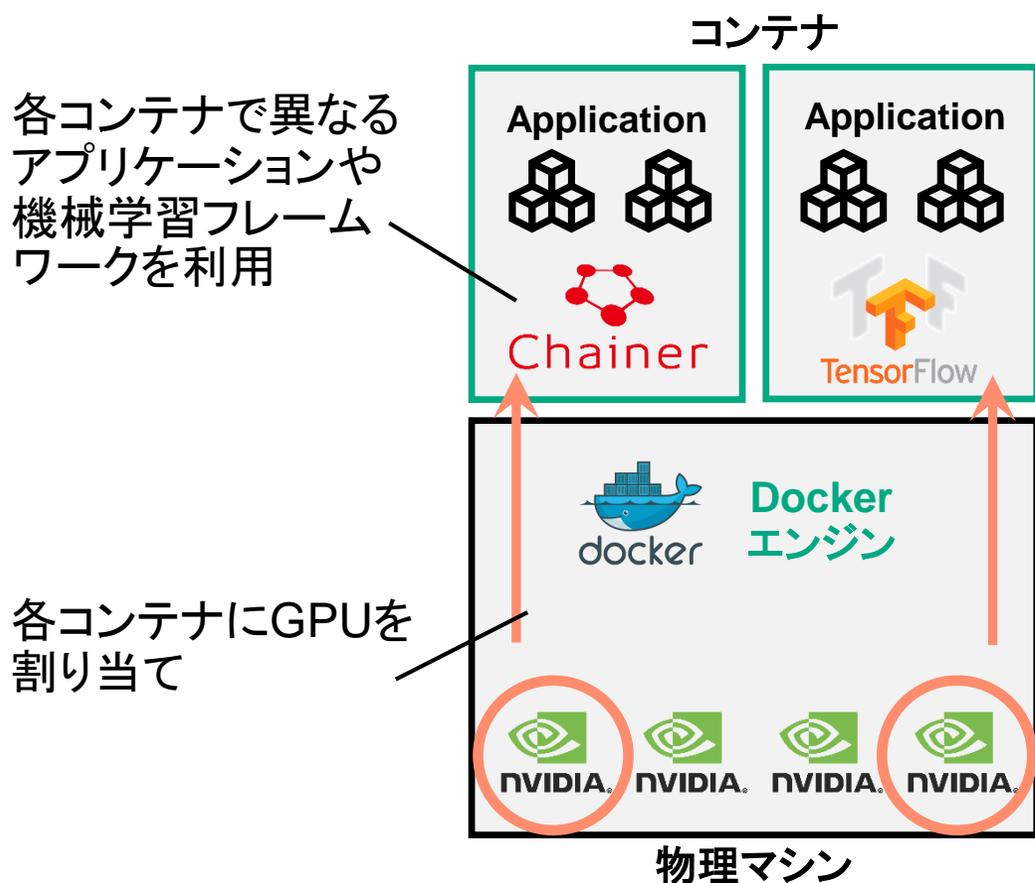
機械学習フレームワークの種類ごとにGPUの対応可否が異なる → 複数バージョンが必要になる。



cuBLASライブラリ:  
NVIDIA®CUDA™ランタイムの上にBLAS (Basic Linear Algebra Subprograms) を実装したもの

cuDNN : NVIDIAのCUDAディープニューラルネットワーク向けのGPUアクセラレーションライブラリ

# 個々のコンテナ要件に従って、適切なGPUリソースを割り当てる



## Non-Dependency

- 異なる機械学習フレームワークを利用可能
- 作成したDockerイメージを別環境にコピーでき、環境依存による不具合発生リスクも低減

## Nvidia Support

- Nvidia社提供のライブラリにより、GPUをコンテナ内で使用/分離可能
- CUDA対応のDockerイメージが利用可能

## Easy to Start

- Docker Hub上で提供される機械学習フレームワーク入りDockerイメージを使って、すぐにアプリ開発が可能
- 機械学習フレームワーク自体のビルドの手間を大幅削減

# nvidia-docker

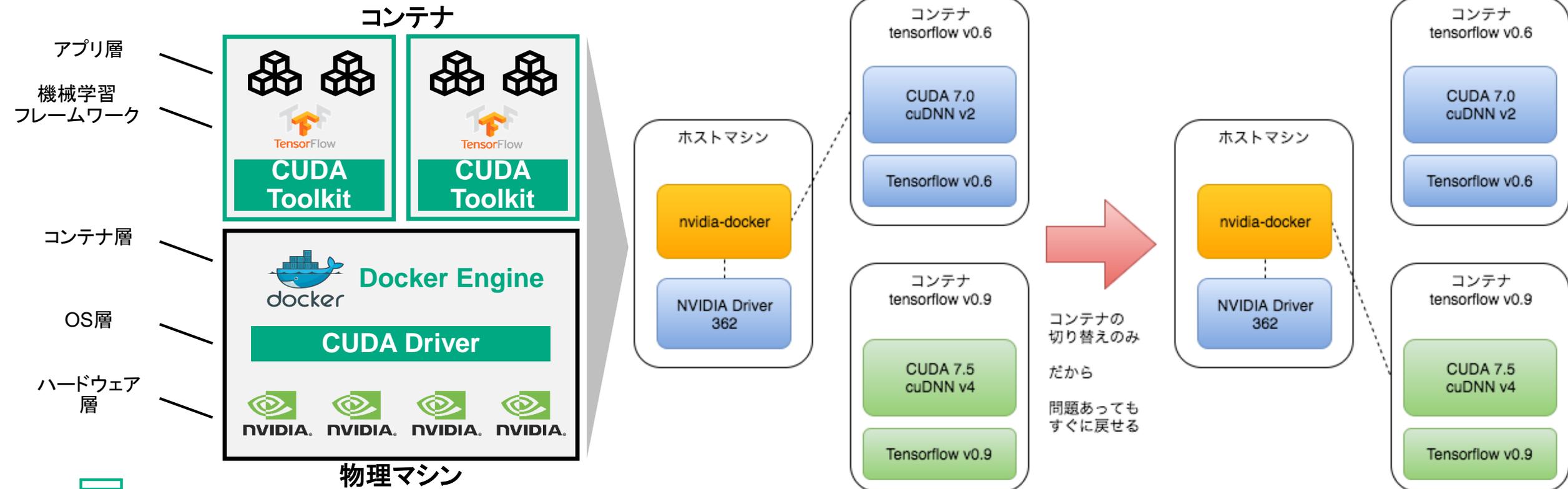
NVIDIAが開発しているDockerプラグイン

Dockerホスト側: NVIDIA GPUドライバーが稼働

Dockerコンテナ側: CUDA Toolkit(+cuDNN)が稼働

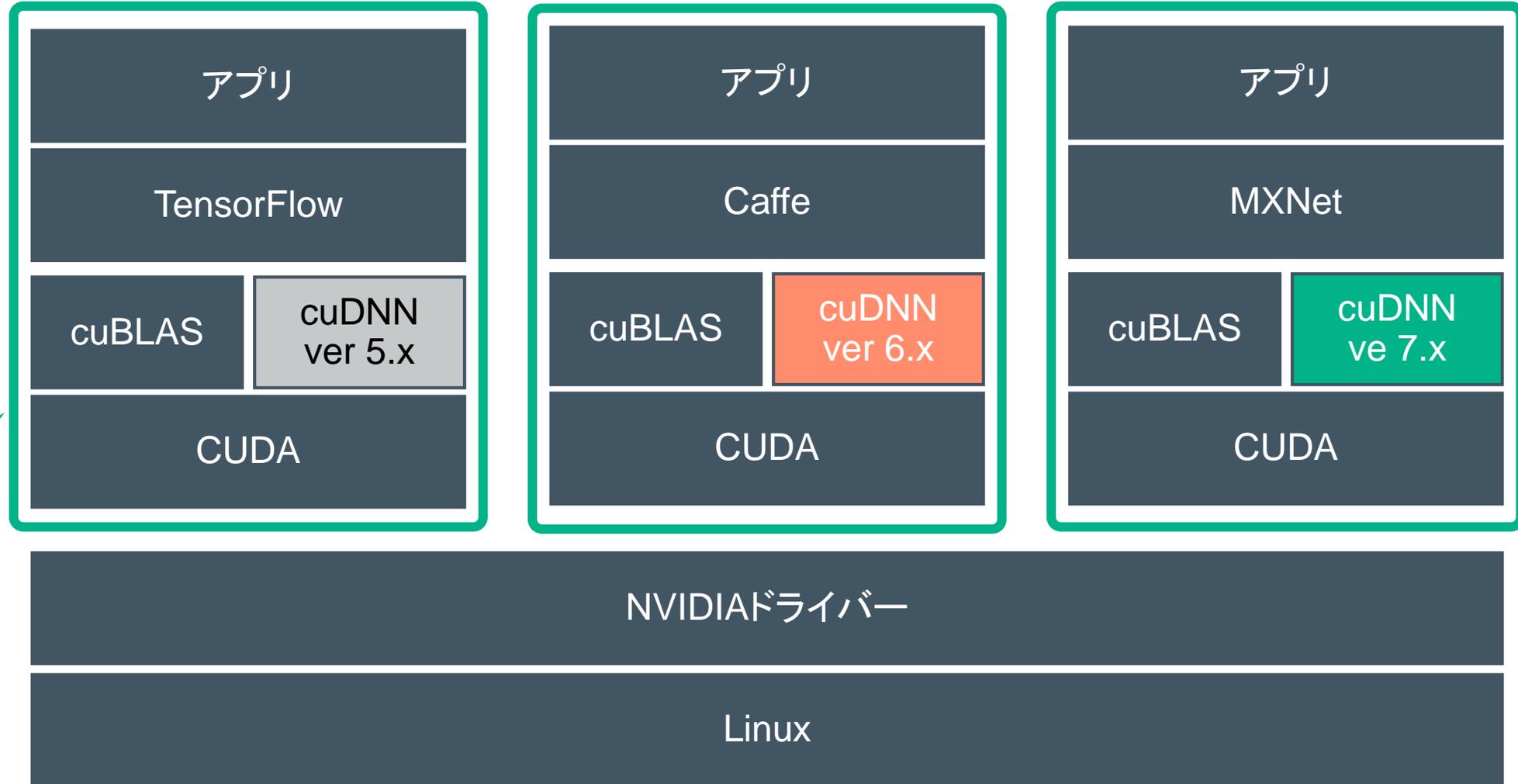
1台の物理マシン上の複数コンテナで、バージョンの異なるCUDAツールキットが混在、利用可能

nvidia-dockerを使う場合



# フレームワークの環境をコンテナ化

コンテナ化すれば、単一のLinux環境に複数バージョンが同時に利用可能

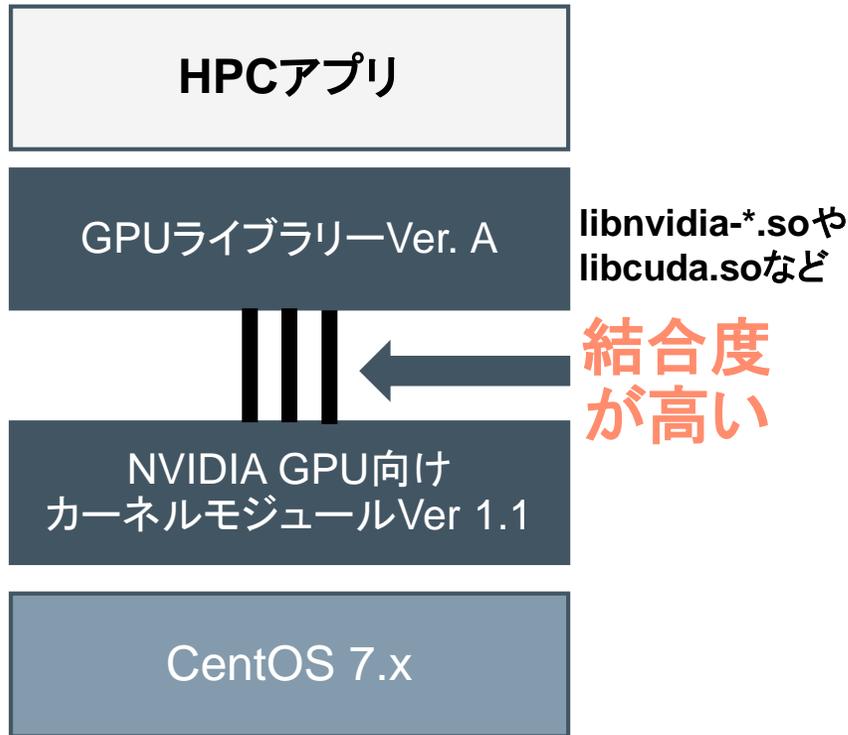


コンテナ化

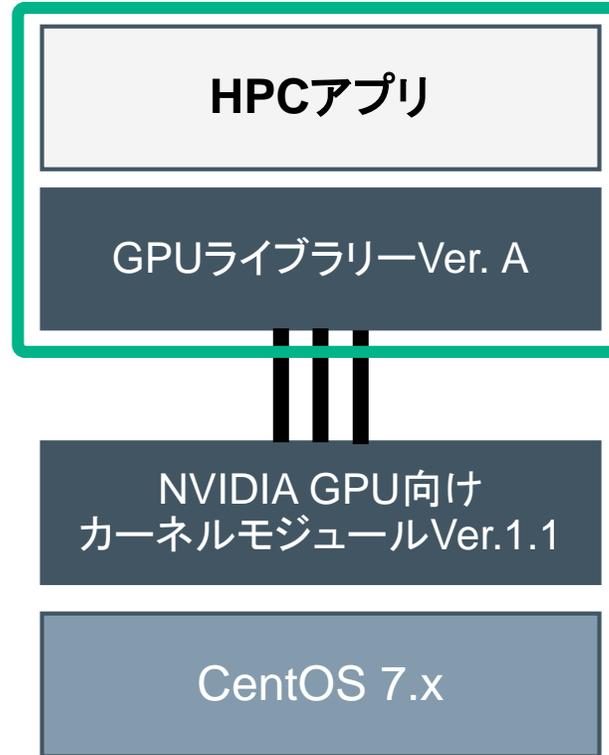
# ドライバーとライブラリーの関係

GPU向けカーネルモジュール(=ドライバー)とGPUライブラリーは、バージョンの整合性が厳密に求められる！

非コンテナ環境

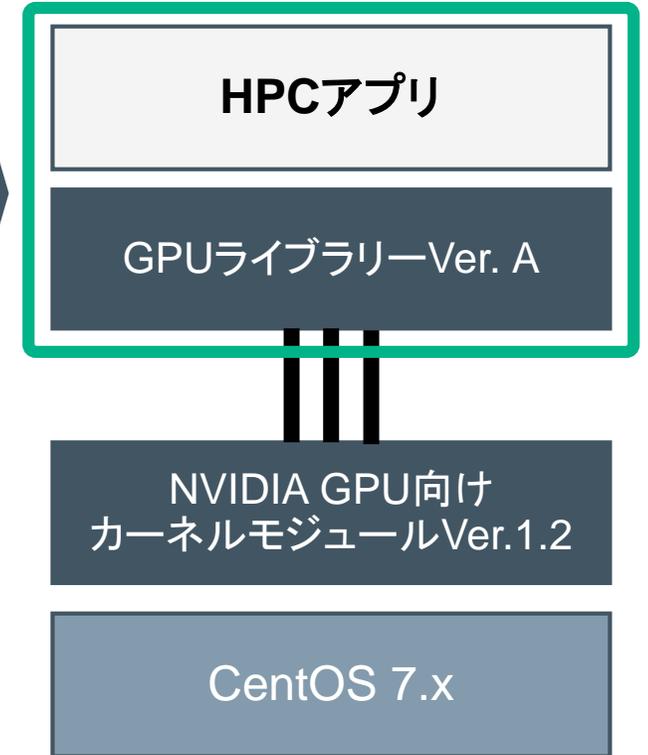


GPUライブラリーを含んだ  
コンテナを作るという発想だが...



Docker  
イメージを  
移植

GPUカーネルモジュールの  
バージョンが異なると、動かない！



可搬性が低いコンテナ  
になってしまう

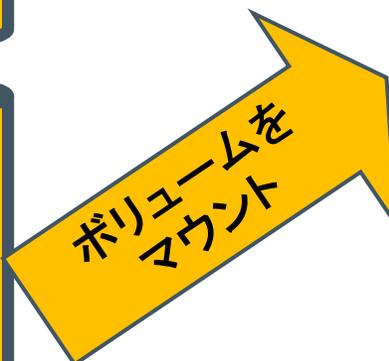
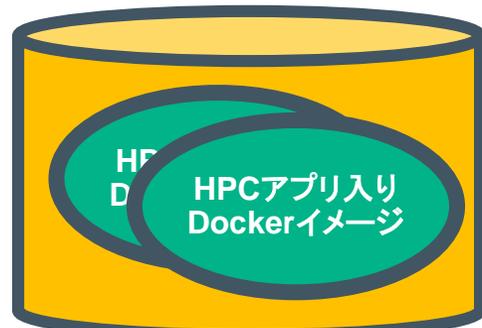
GPUのカーネルモジュール(=ドライバー)とGPUライブラリーのミスマッチによるエラーの例:

```
$ nvidia-smi  
Failed to initialize NVML: Driver/library version mismatch
```

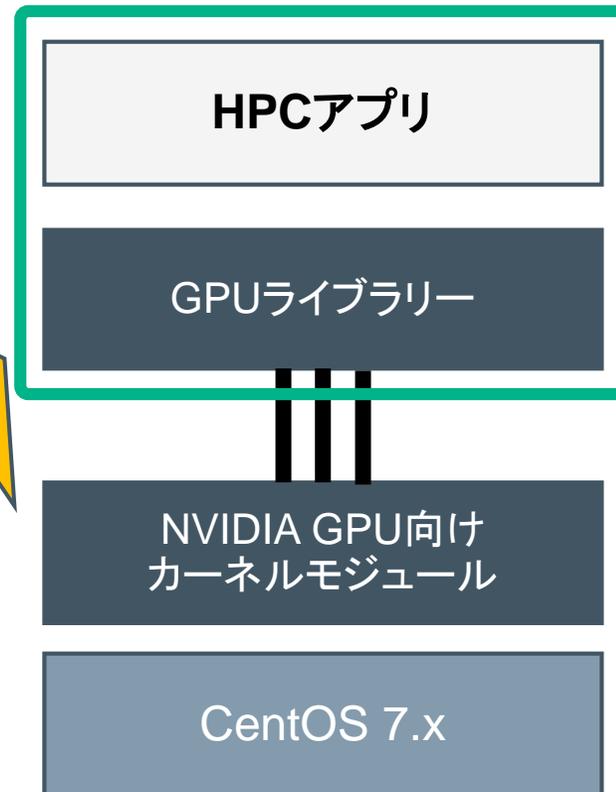
# 可搬性の高いGPU対応コンテナを作るには？

ホストOS上のGPU向けカーネルモジュール(ドライバー)と整合性のとれたGPUライブラリーを Dockerコンテナ向けのボリュームに保管し、コンテナへ提供する

Dockerイメージには  
ユーザーアプリとそのランタイム  
ライブラリーのみを含める



コンテナ起動時に  
GPUライブラリーを含む  
ボリュームをマウント



# 大規模Docker基盤の課題

- 大量コンテナの起動、調整、削除
- 手動実行は、管理が煩雑
- 大量コンテナへのIT資源の割り当て作業が面倒

# コンテナの自動配備

## IT資源管理の自動化

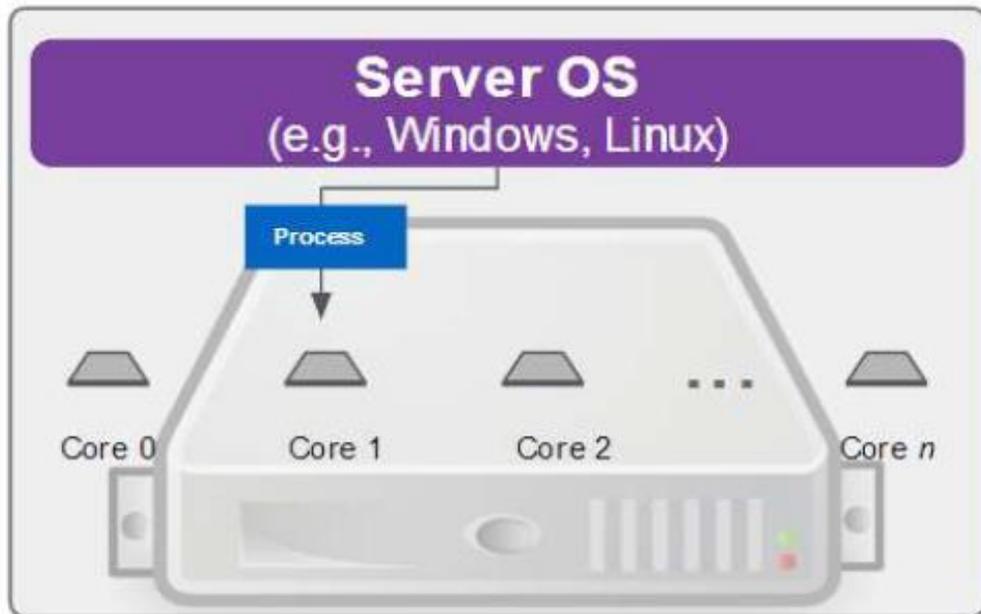


DC全体を管理する

“OSのようなもの”が必要

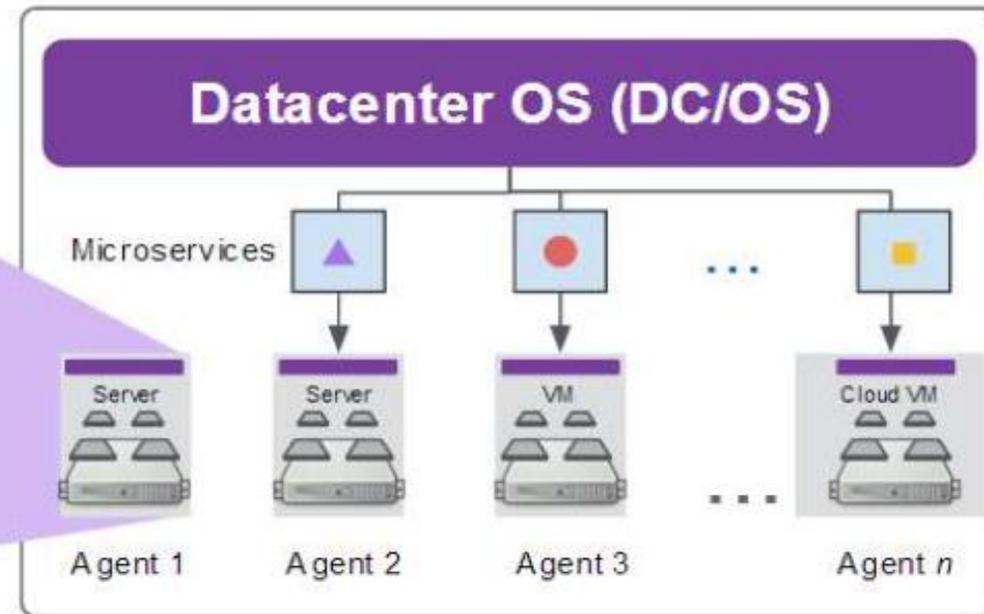
# Datacenter Operating Systemとはなにか？

サーバーOS



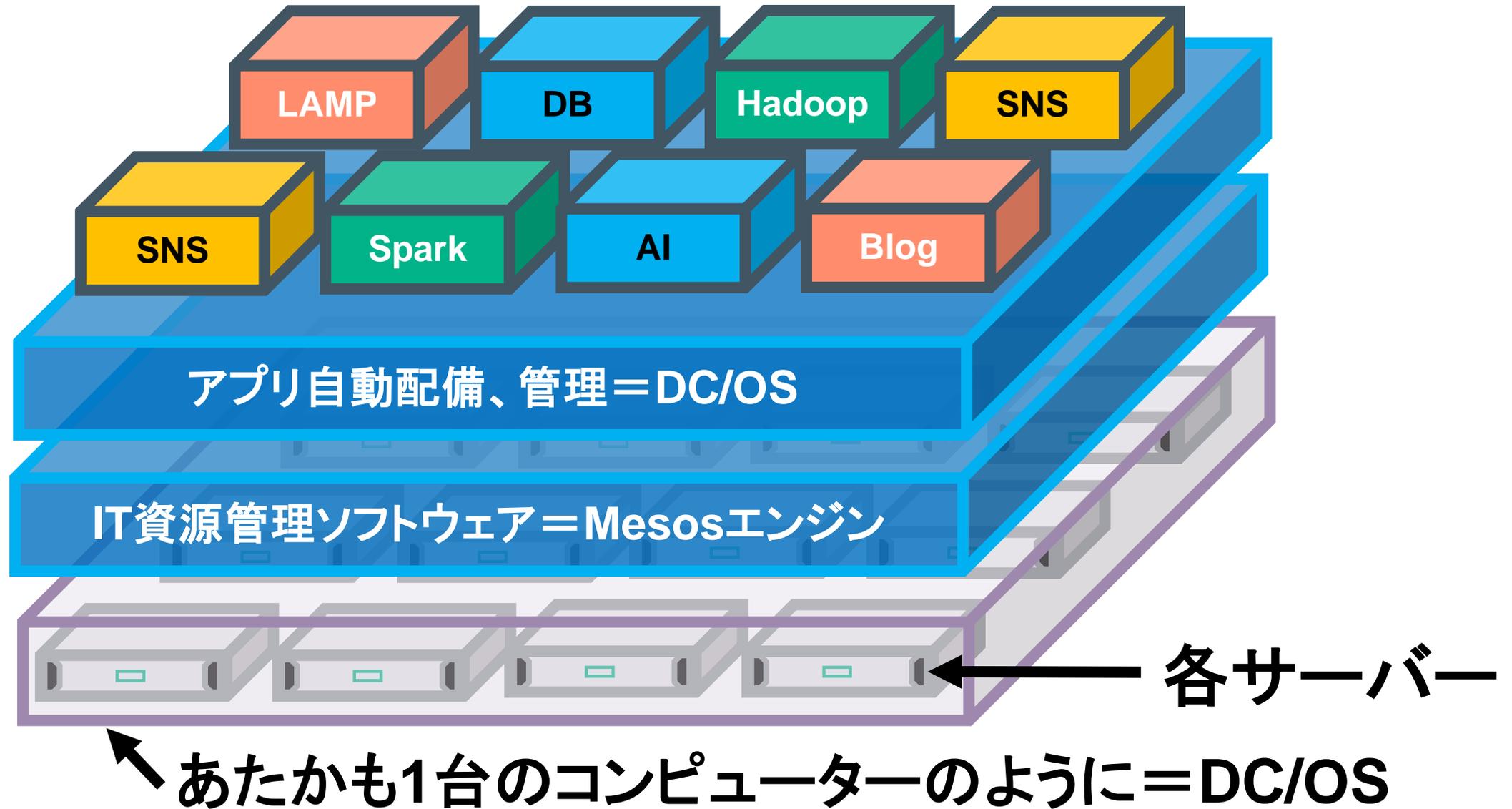
サーバー内の空いているCPUコアにアプリケーションのプロセスが割り当てられる

DC/OS

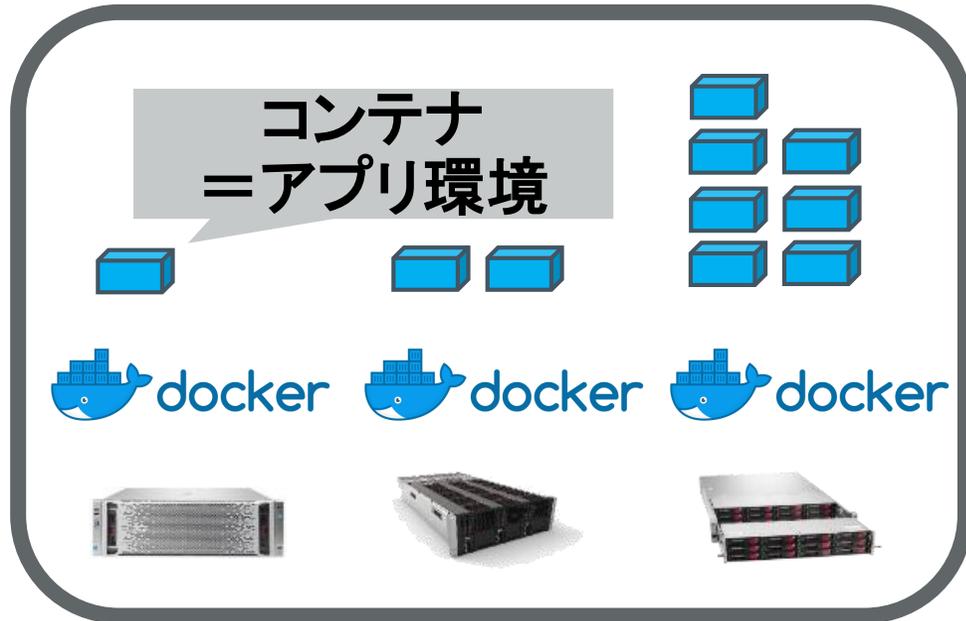


データセンター内の空いているCPUコアに分散アプリケーションのプロセスが割り当てられる

# 複数のマシンを1台のコンピューターとみなし、IT資源をプール化

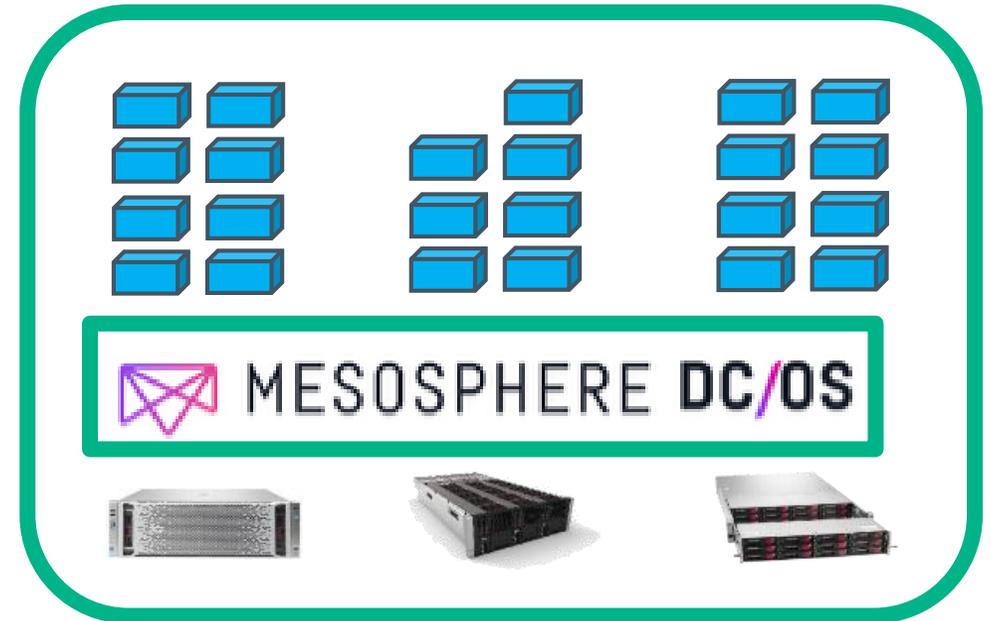


# コンテナ 手動配備



## ムラがある

# 空いたマシン に自動配備



今まで

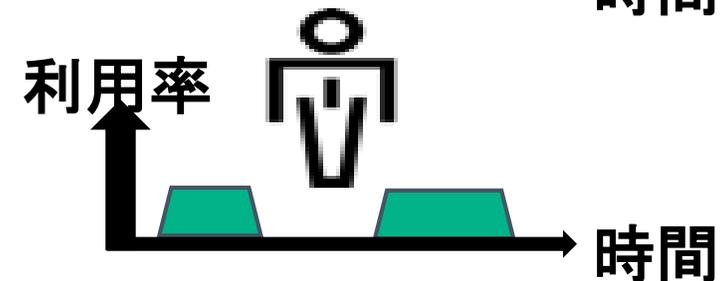
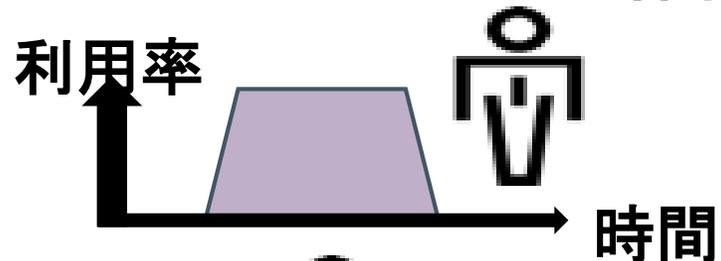
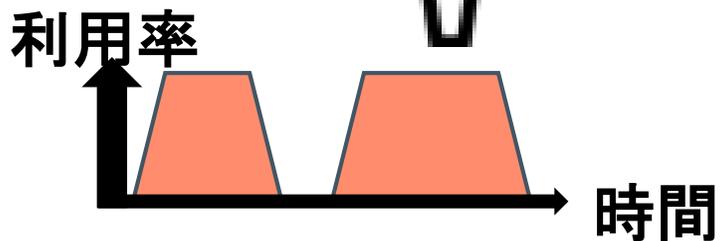
DB



Webアプリ



人工知能

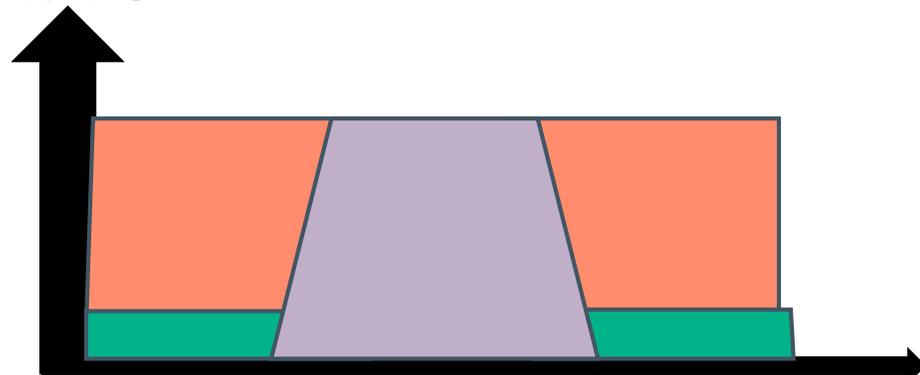


IT基盤の利用効率化:  
人間の手作業

Hewlett Packard  
Enterprise

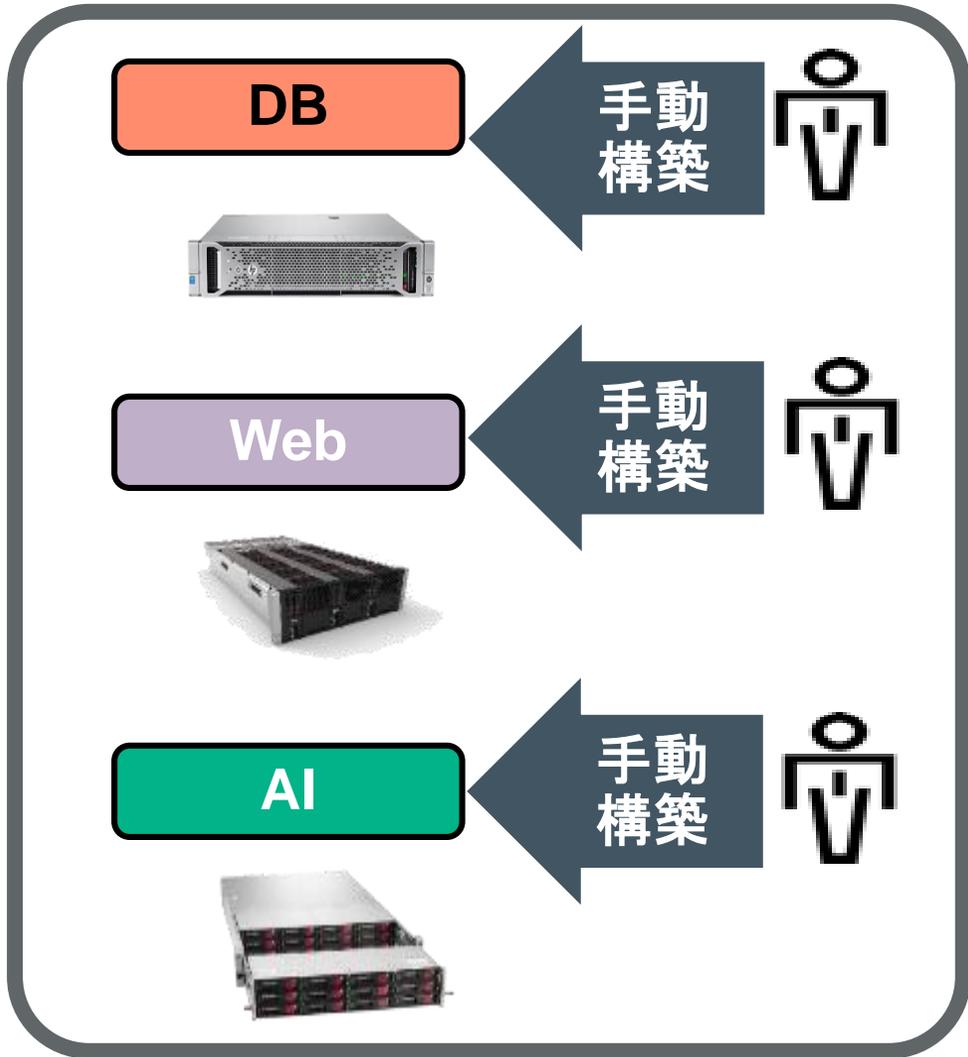
MESOSPHERE DC/OS

利用率

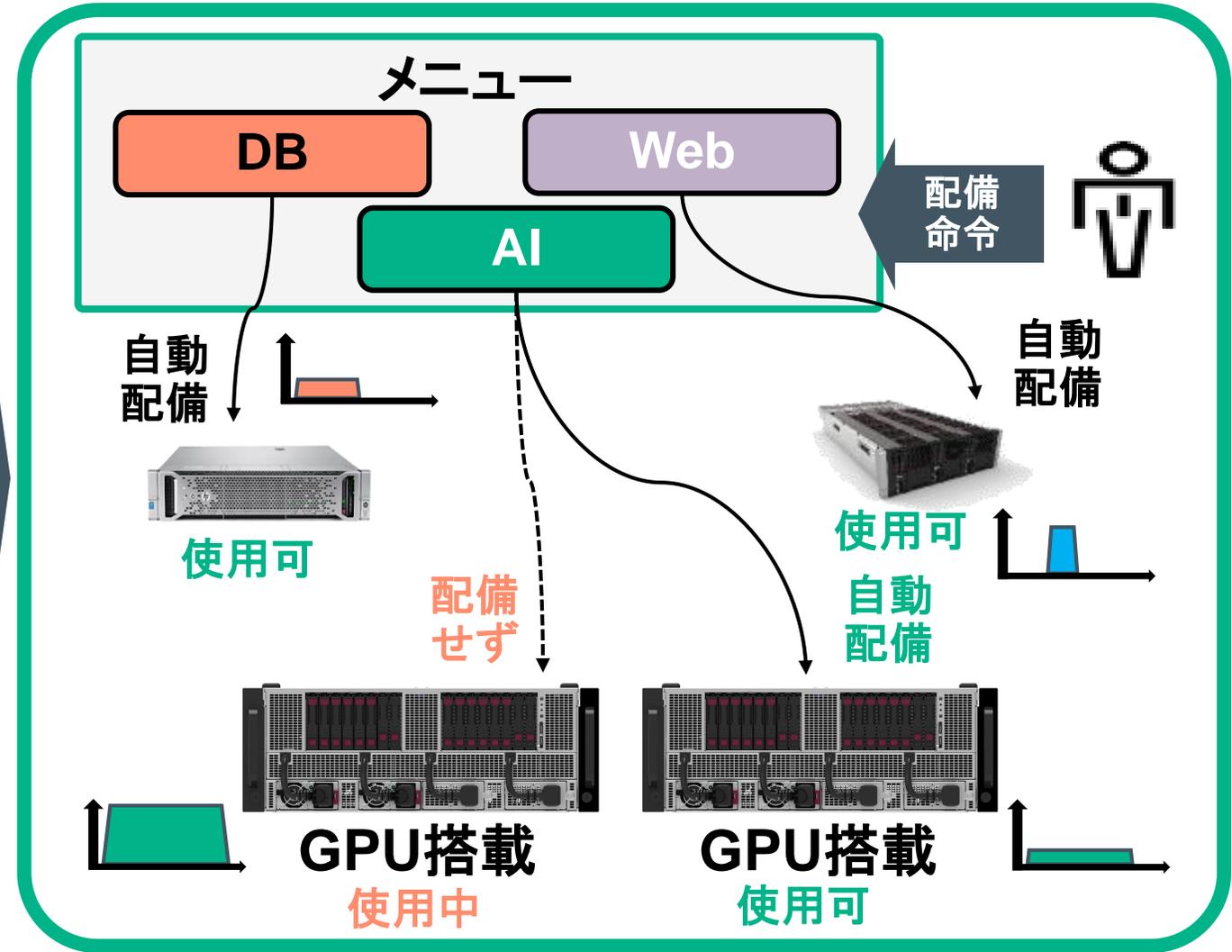


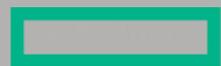
利用効率向上の作業:  
ソフトウェアで  
自動化!

# 今まで



# Mesosphere DC/OS





**Hewlett Packard  
Enterprise**

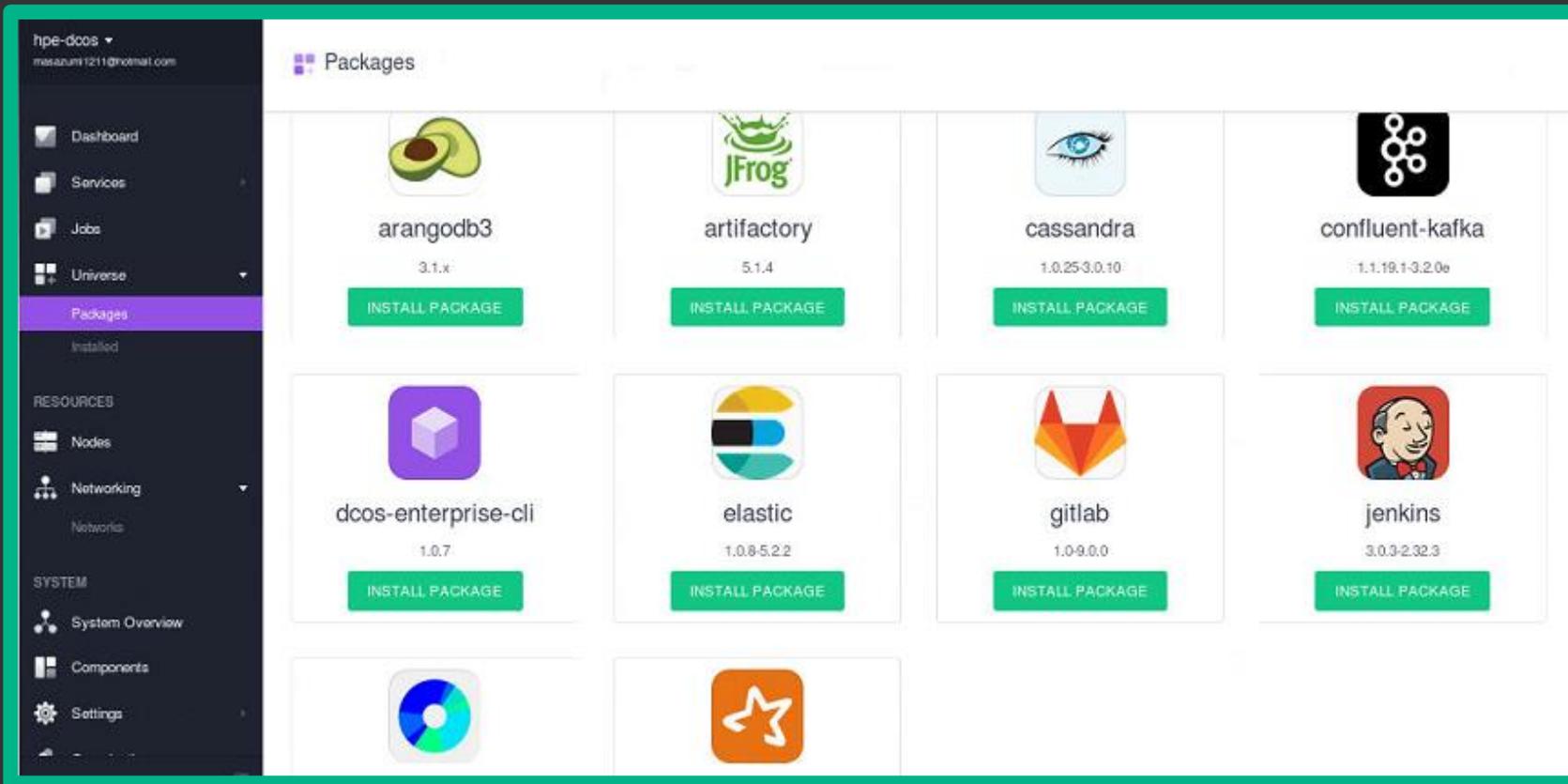


MESOSPHERE

分散IT基盤でのOSS  
構築は、手間ばかり  
かかるから、正直、  
やりたくないな...

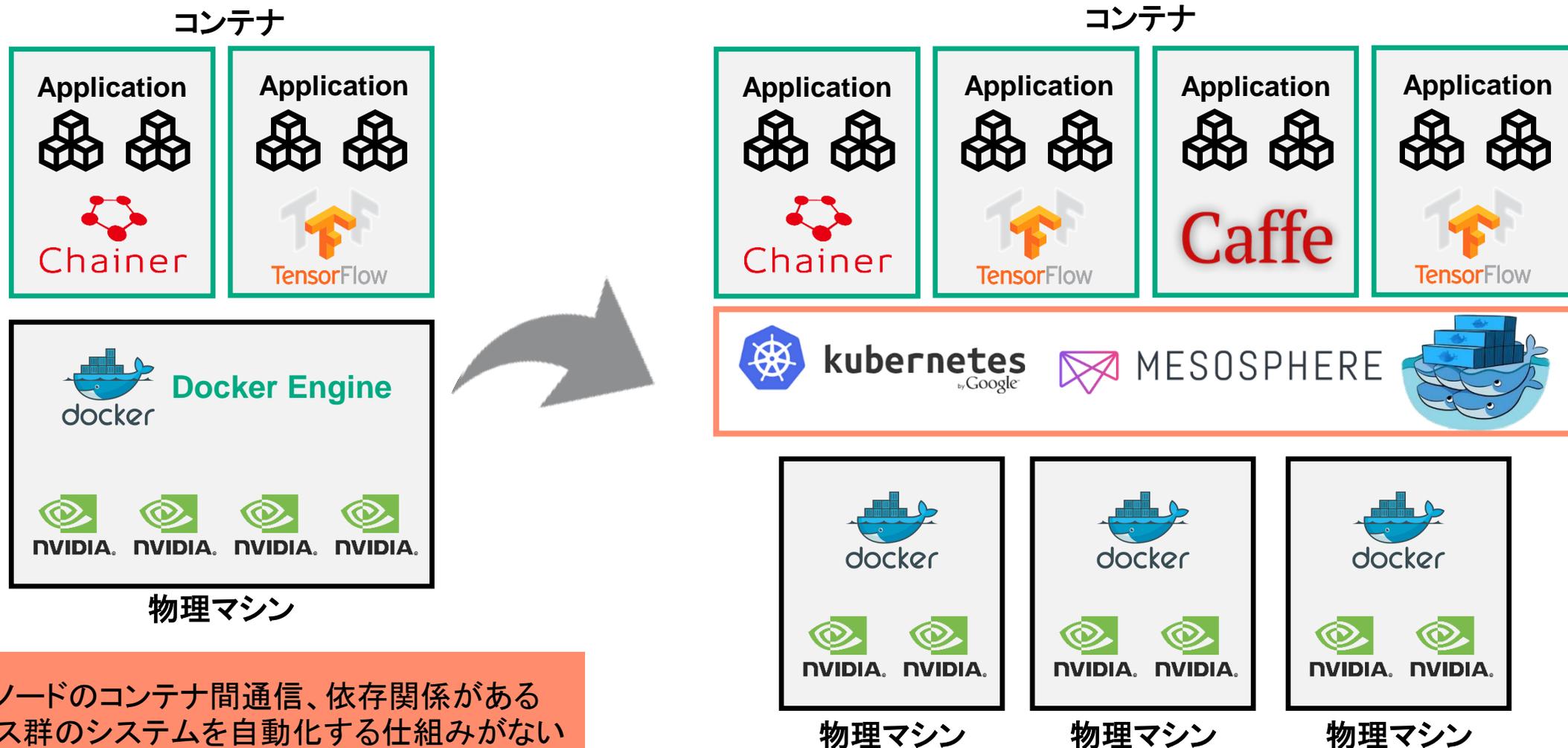


# アプリは、選んで配備する時代！ ～ オープンソースの自販機 ～



コンテナ向けハードウェア  
HPE Synergy

# 複数ノードでの運用では、コンテナ連携ツールが必要



複数ノードのコンテナ間通信、依存関係があるサービス群のシステムを自動化する仕組みがない

# 旧システム: 例えば、仮想通貨の取引所

## 閑散期



VM =  
性能劣化アリ



種

## 繁忙期



手動  
増設



遅い...

種

何個動かせばいいんだよ...

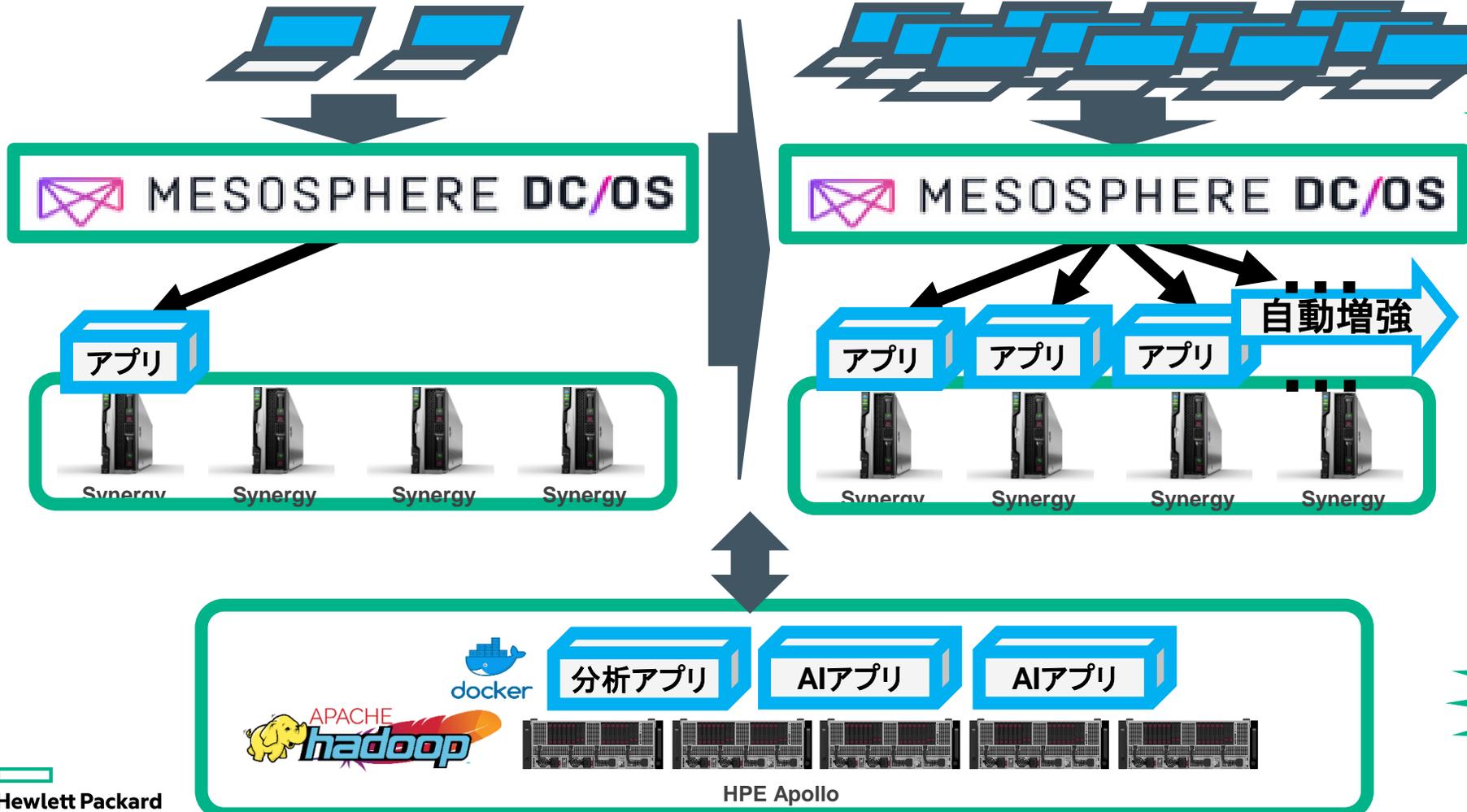
もう夜11時...  
早く帰りたい...

遅い!  
面倒!  
ストレス!

# 繁忙期の作業を自動化！

閑散期

繁忙期



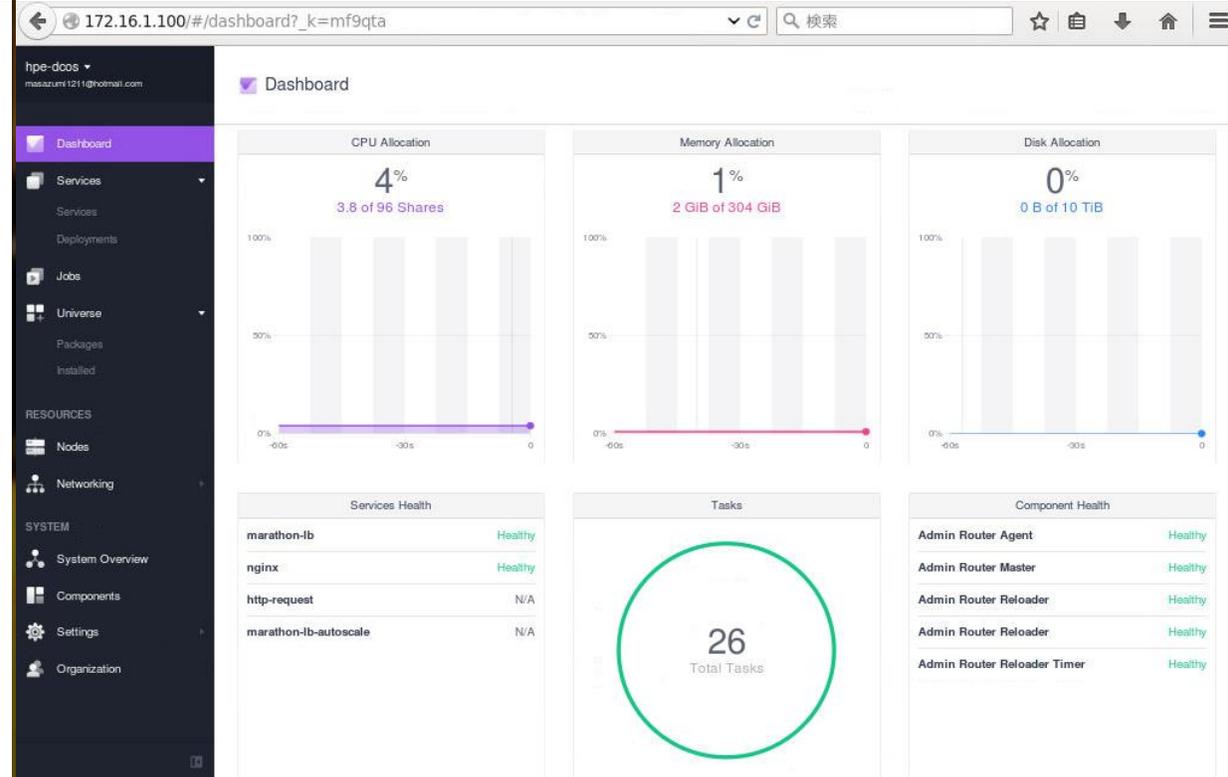
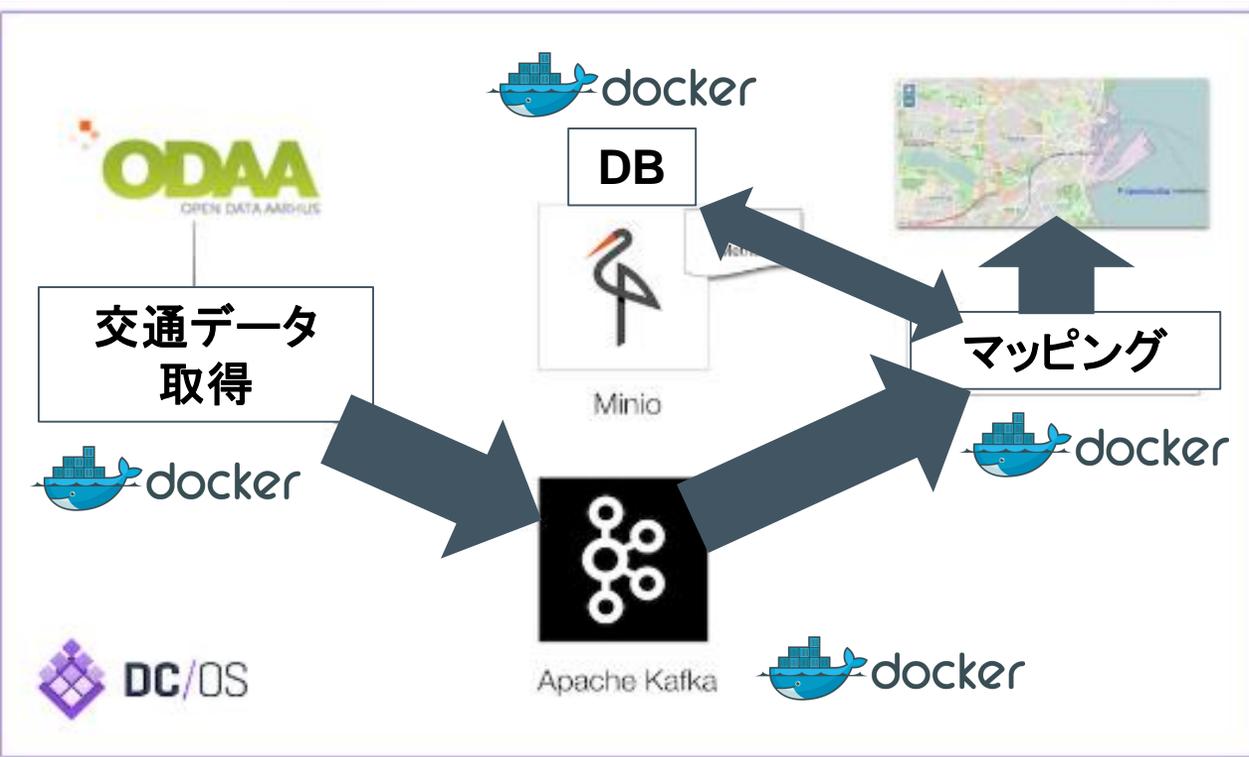
- ✓ 大量提供
- ✓ 迅速稼働
- ✓ 自動増強
- ✓ 手作業無し

傾向データから知見獲得！  
顧客動向把握に活用！

コンテナ  
による業務改革

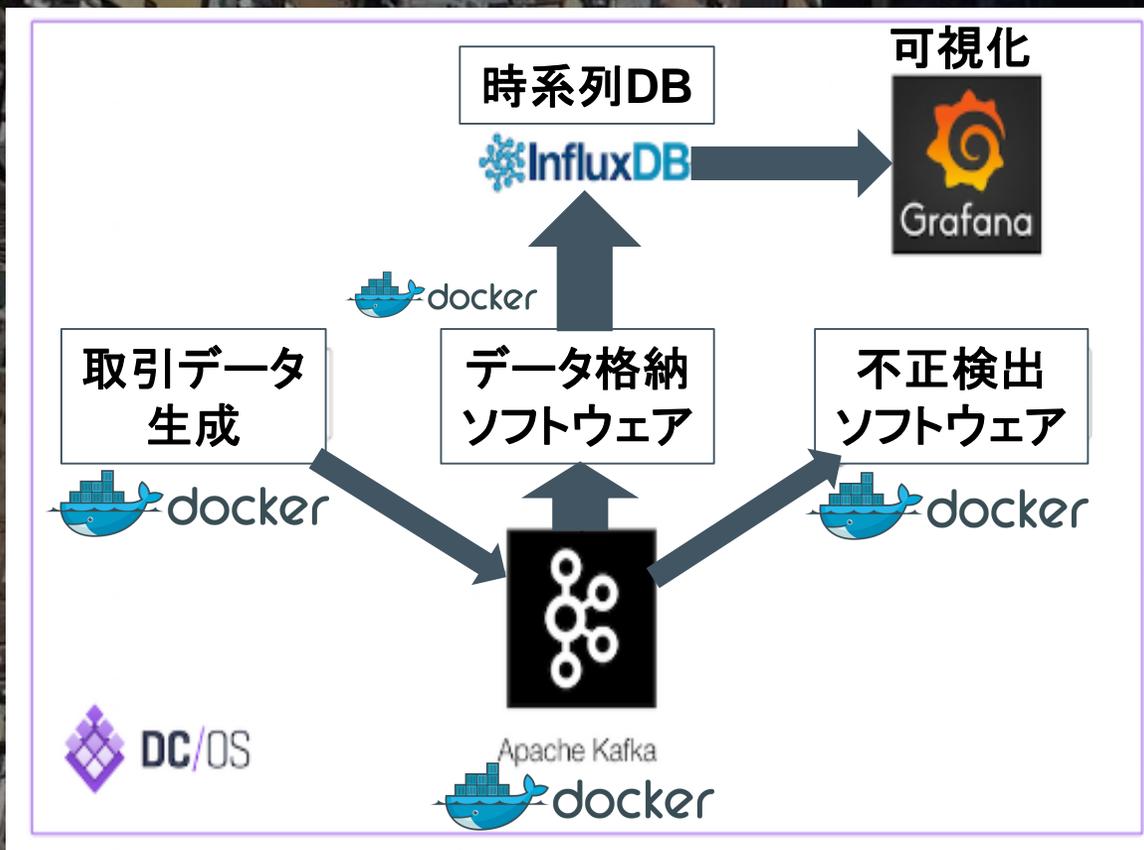
# Mesosphere DC/OSで楽しもう

## すでにあるものを入手して、組み合わせる



# Mesosphere DC/OSで楽しもう

すでにあるものを入手して、組み合わせる



The top screenshot shows the Mesosphere DC/OS 'Services' page. It lists several services: 'intrans', 'grafana', 'influxdb', 'kafka', and 'marathon-lb'. Each service has a status indicator (Running) and resource usage (CPU, MEM, DISK). The bottom screenshot shows a 'Financial Transaction Processing' dashboard with a line chart titled 'Overall Transactions'. The chart displays transaction volume over time for five cities: London, Moscow, New York City, Tokyo, and San Francisco. A legend on the right provides summary statistics for each city.

City	avg	total
London	\$6.8K	\$394.2K
Moscow	\$6.8K	\$347.4K
New York City	\$7.0K	\$358.1K
Tokyo	\$7.5K	\$422.2K
San Francisco	\$7.0K	\$315.7K

# 先進企業は、なぜ Docker + Mesosphere DC/OSなののか？

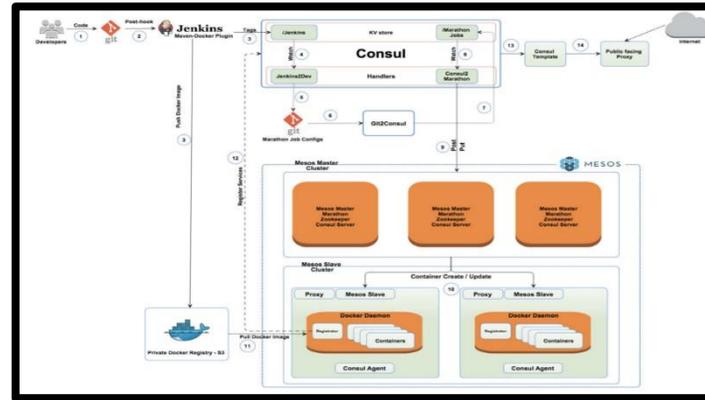
## Verizon



### ビッグデータ基盤の配備

<https://www.youtube.com/watch?v=WXX77IX-zl8w>

## Samsung



### IoT分析ジョブ実行、CI/CD

<https://mesosphere.com/blog/2015/12/21/samsung-is-powering-the-internet-of-things-with-mesos-and-marathon/>

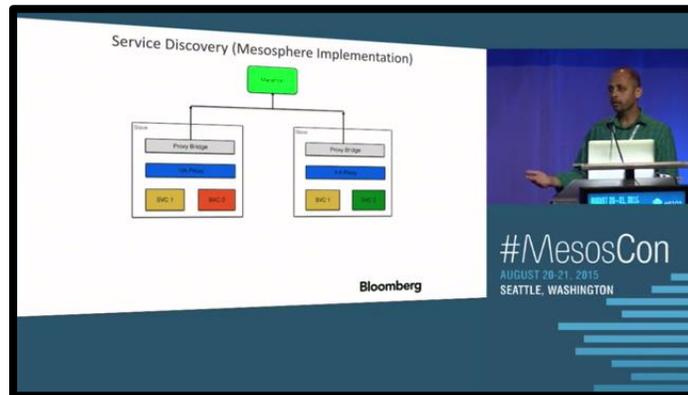
## Apple



### Siriの基盤(Hadoop)

<https://mesosphere.com/blog/2015/04/23/apple-details-j-a-r-v-i-s-the-mesos-framework-that-runs-siri/>

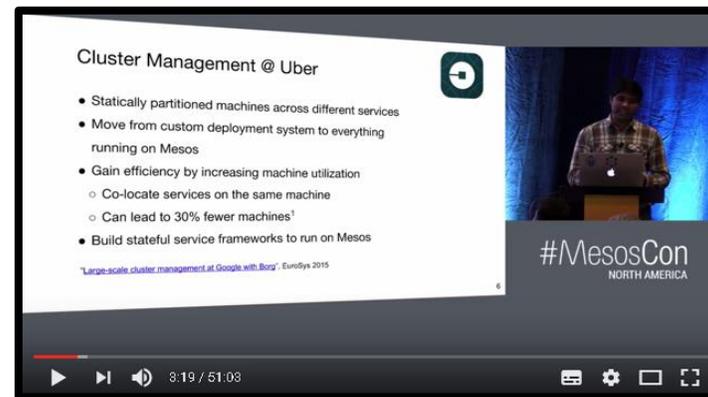
## Bloomberg



### CEP、分析、検索基盤統合

<https://www.youtube.com/watch?v=6w8PXklrC5I>

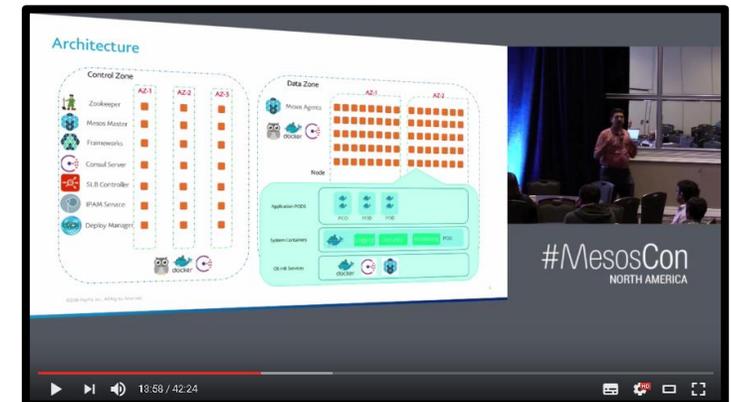
## Uber



### Cassandra基盤

<https://www.youtube.com/watch?v=U2jFLx8NNro>

## PayPal



### CI/CD開発環境の提供

<https://www.youtube.com/watch?v=OJgmfog0e1w>



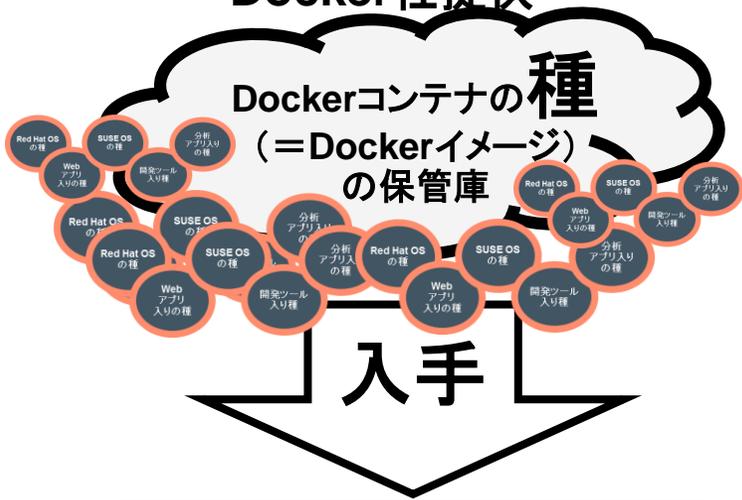
Hewlett Packard  
Enterprise

# 2020年を見据えた 次世代ビッグデータ基盤

# コンテナを活用した次世代型ビッグデータ・AI基盤

Docker社提供

Dockerコンテナの種  
(= Dockerイメージ)  
の保管庫

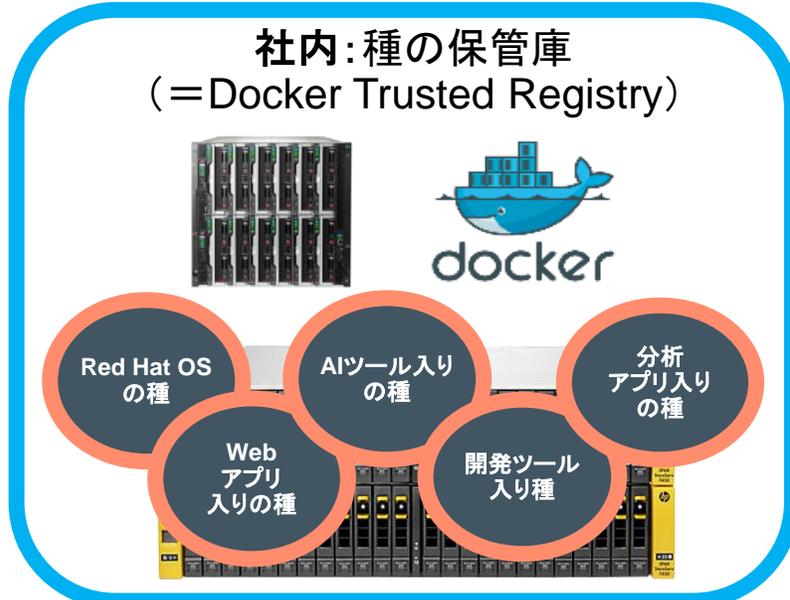


## 特長・メリット:

- ✓ 物理1台で複数・異種Linux環境が同時稼働
- ✓ VMに比べ、性能劣化なく稼働、**超軽量**
- ✓ コンテナの元となる種(= Dockerイメージ)が豊富
- ✓ アプリ入りの種を入手し、すぐに稼働
- ✓ ビッグデータ用アプリの入れ替えも簡単
- ✓ 他のDocker環境に種を簡単にコピー、すぐに稼働

物理と同じ速度で  
素早く動くアプリ環境が、  
すぐに手に入り、すぐに  
動く！複数同時に！

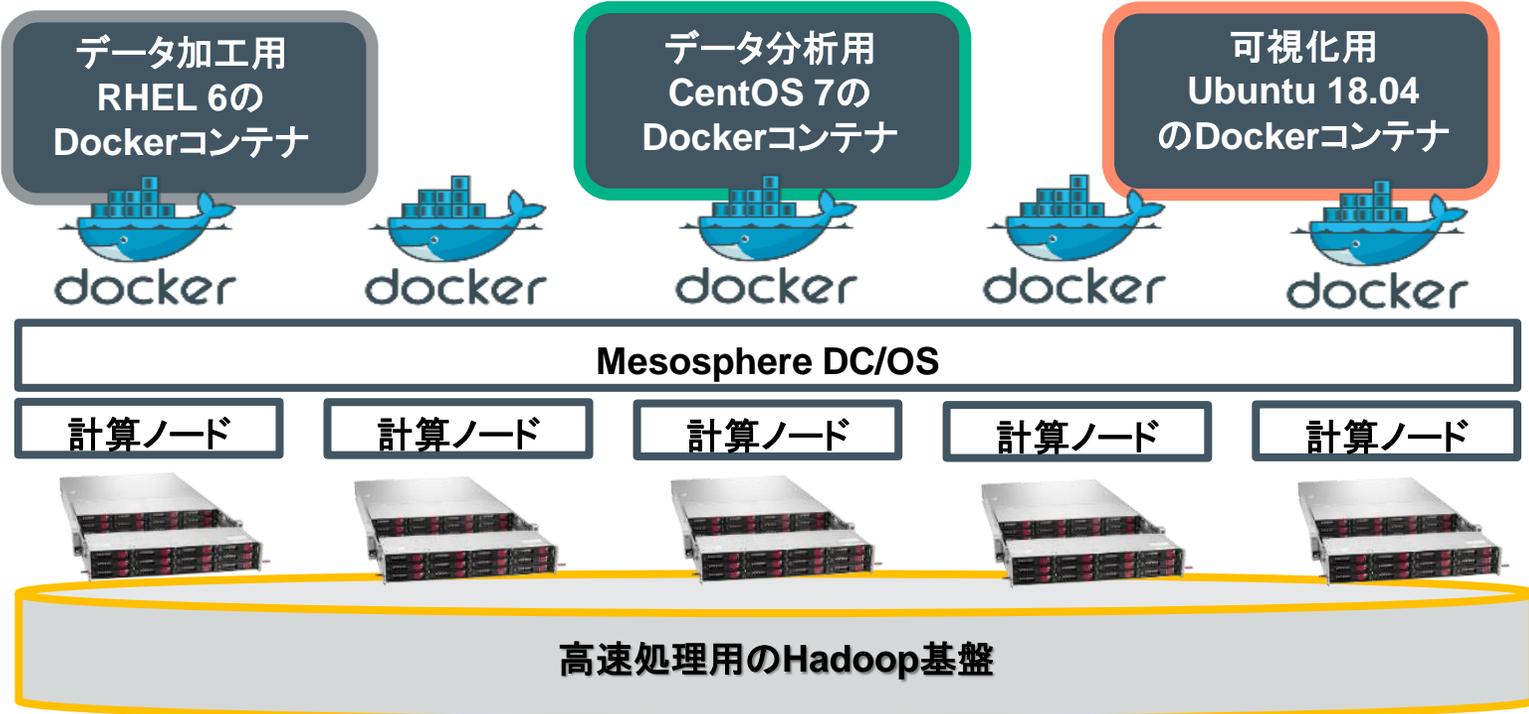
社内:種の保管庫  
(= Docker Trusted Registry)



データ加工用  
RHEL 6の  
Dockerコンテナ

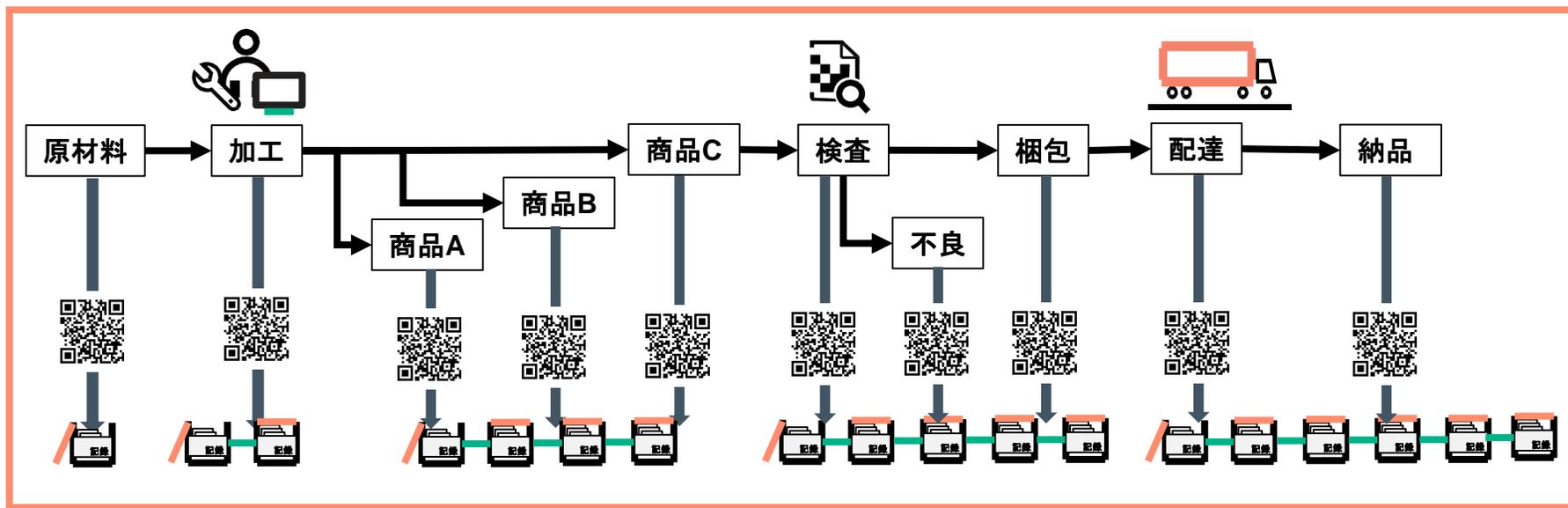
データ分析用  
CentOS 7の  
Dockerコンテナ

可視化用  
Ubuntu 18.04  
のDockerコンテナ

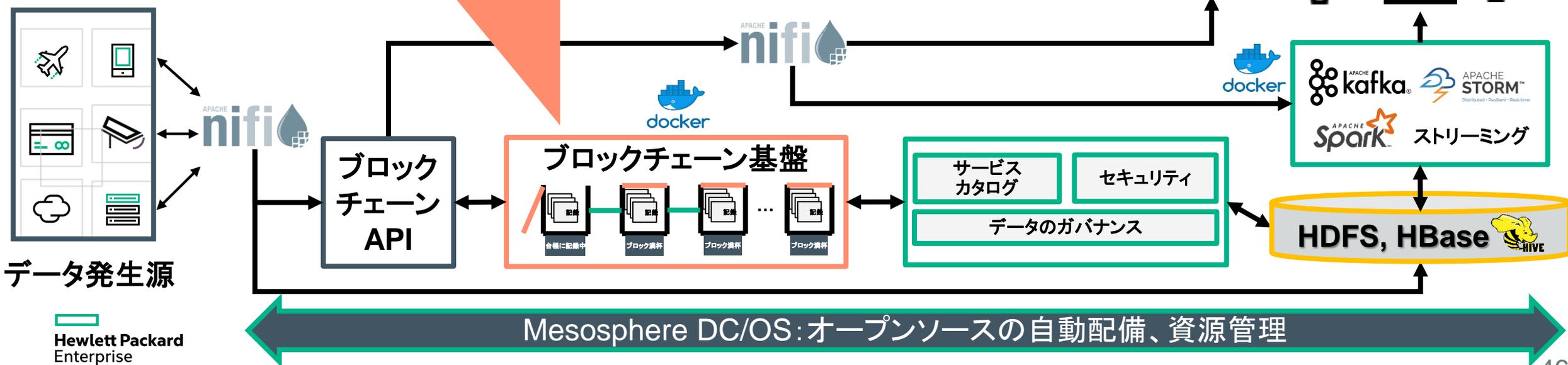




# ブロックチェーン+Hadoop+GPU+Docker on DC/OS



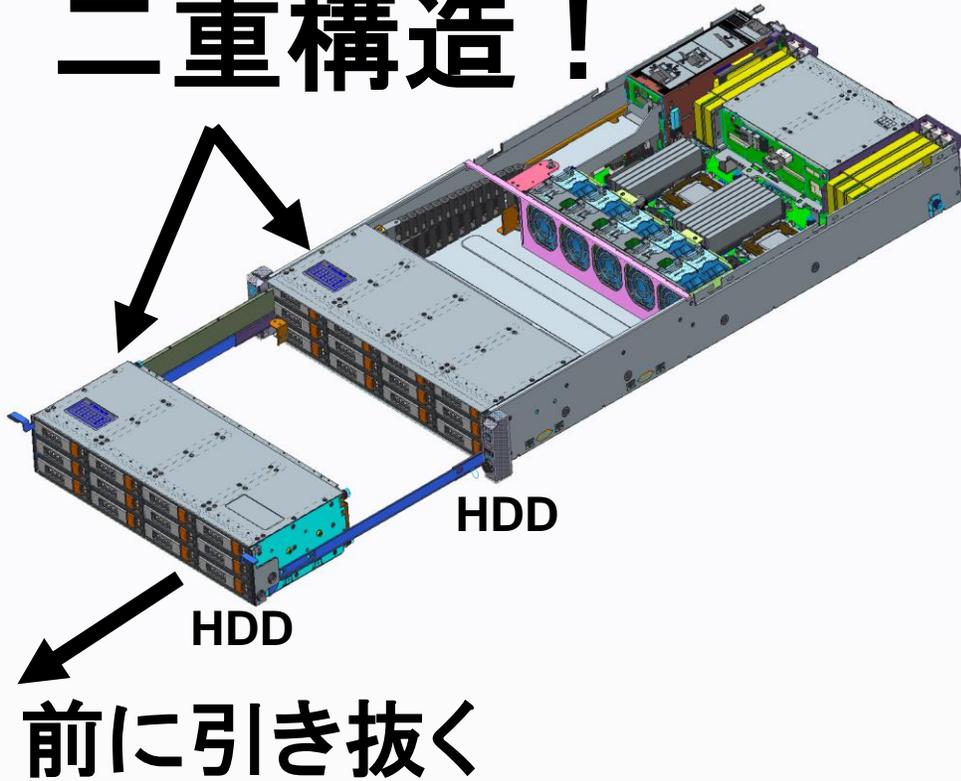
- ブロックチェーン: GPUが必須ではない
- (仮想通貨マイニングならばGPUが有用)
- HadoopはDC/OSで自動配備
- AIアプリ開発は、Docker
- AIアプリ実行は、GPUを駆使



日本で人気の理由とは？

# 人気のHadoop向けサーバーHPE Apollo 4200

なんとHDDケースが  
**二重構造！**



- Xeonプロセッサ :2基(=最大4\*コア)
- メモリ :512GB(最大1TB)

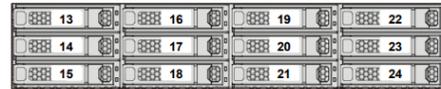
- LFF SATA HDD : **28**本
- 追加NIC:10GbE or 25GbE(2ポート)

前面(標準)

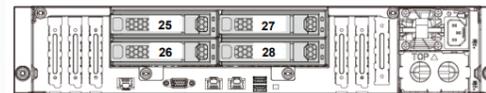
フロント ドライブケース1



フロント ドライブケース2



28LFF 構成時 背面(標準)



\* CTOモデル

なぜ28本も必要なの？

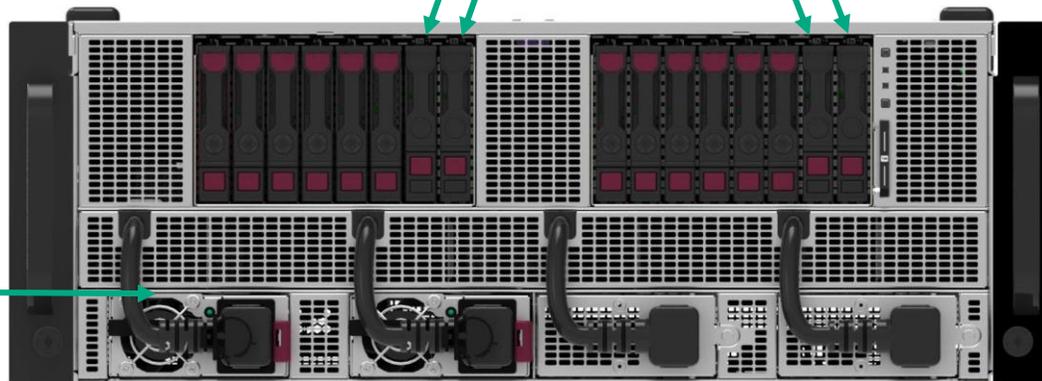
パソコンのグラボと何が違う？

# 爆速GPUスーパーコンピューターHPE Apollo 6500

AI処理時間の大幅短縮  
→ GPUが必要

NVMeドライブ(最大4本)

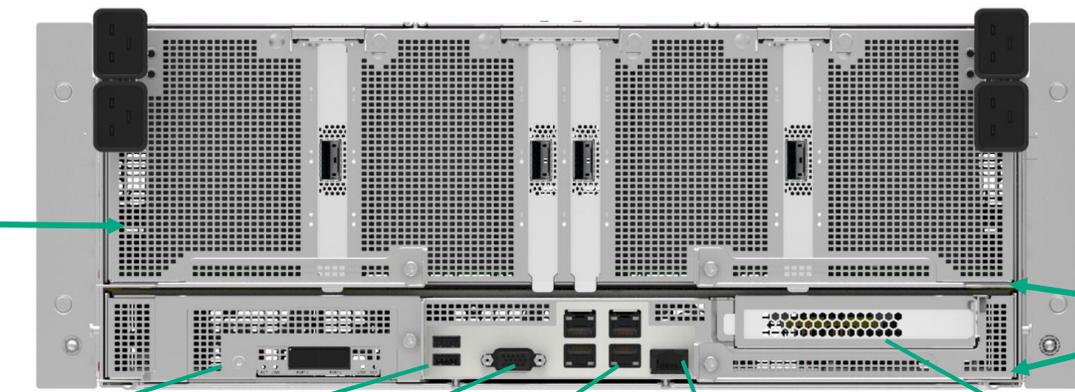
冗長電源  
2200W 80plus Platinum



最大  
8 GPU

高さ4U

ラッチ  
GPUとマザーボードに  
アクセス



FlexLOM 2x USB VGA 4 x 1Gbe iLO5 FH/HL PCIe Gen3 x16

<sup>1</sup> Embedded SATA or m.2 for boot and NVME for high speed cache enabled for early shipments. Smart Array for SAS and SATA SSD will be enabled shortly.

HPE Customers, HPE Channel Partners and HPE Internal Use Only: 本ドキュメントはHPEのお客様とHPEパートナー及びHPE社内での利用に限られます。HPEの競合他社への開示、配布は絶対に行わないでください。

機械学習の導入・構築は簡単？



Hewlett Packard  
Enterprise

ご清聴ありがとうございました

@masazumi\_koga