redhat.

OpenShiftで加速するコンテナによるGPU活用

HPE HPC & AI フォーラム 2018

2018/09/07

Red Hat K.K.
Cloud Solution Architect, Technical Sales
Shingo Kitayama

Agenda

GPU Acceleration on OpenShift

1. オープンハイブリッドクラウド戦略
2. コンテナによるAI/MLアプリ開発
3. GPU スケジューリングの詳細

オープンハイブリッドクラウド戦略

オープンハイブリッドクラウド

クラウドの選択に自由を

ハイブリッドクラウド基盤

RED HAT® OPENSTACK® PLATFORM

クラウドネイティブアプリケーション基盤

RED HAT® OPENSHIFT Container Platform

クラウドに対応した管理と自動化

RED HAT® ANSIBLE® Automation

Red Hat Open Innovation Labs

redhat.

Enterprise Kubernetes

Cloud-native Applications

AI & Machine Learning

Blockchain

Internet of Things

Innovation Culture

Application Services	Cluster Services	Developer Services
Middleware, Service Mesh, Functions, ISV	Metrics, Chargeback, Registry, Logging	Dev Tools, Automated Builds, CI/CD, IDE

Automated Operations*

Kubernetes

Red Hat Enterprise Linux or Red Hat CoreOS

Best IT Ops Experience ⟷ CaaS PaaS Best Developer Experience

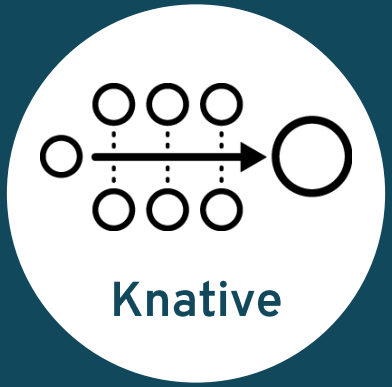
*coming soon

Kubernetesで広がる新世界

Service Mesh

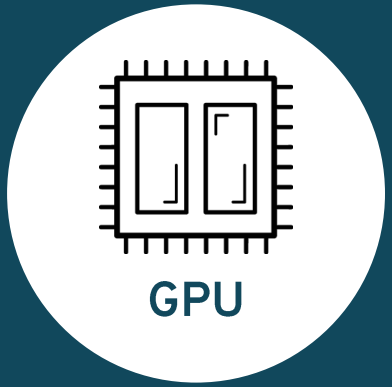
マイクロサービス間の通信を統一的な仕組みで制御。
複雑化したマイクロサービスの課題を解決。

Serverless



「非常駐型プロセス」をイベントによって制御。
プロセスのオートスケールを提供。

AI / ML



学習/推論に利用するGPUリソースを制御。
GPUリソースの効率化を提供。

コンテナによるAI/MLアプリ開発

redhat.

AI/MLアプリ開発におけるトレンド

1. 精度競争は落ち着きつつある
   - 基本的な学習タスクの効率化に注力
   - クラウドサービスの利用に注目
   - より難しい学習による、精度競争はいまも継続中
2. 応用技術への展開が加速
   - 画像＋自然言語といった組み合わせの複雑なモデル開発
   - 高速な並列データ処理が可能なGPUの利用

AI/MLアプリ開発におけるコンテナの活用

Portability
バージョン管理からの開放

Composability
AI/MLアプリ開発のパイプライン

Scalability
GPUリソースの自由な選択

AI/MLアプリ開発を行う際の課題

アプリケーションを作成する以前に、GPUドライバ、ライブラリ、DLフレームワークなどの依存関係を正しく管理するための高度な専門性が必要

Deep Learning Framework

Application Language

GPU-accelerated Libraries

CUDA Toolkit

Nvidia Driver

GPU Hardware (Nvidia)

Chainer TensorFlow mxnet Caffe2

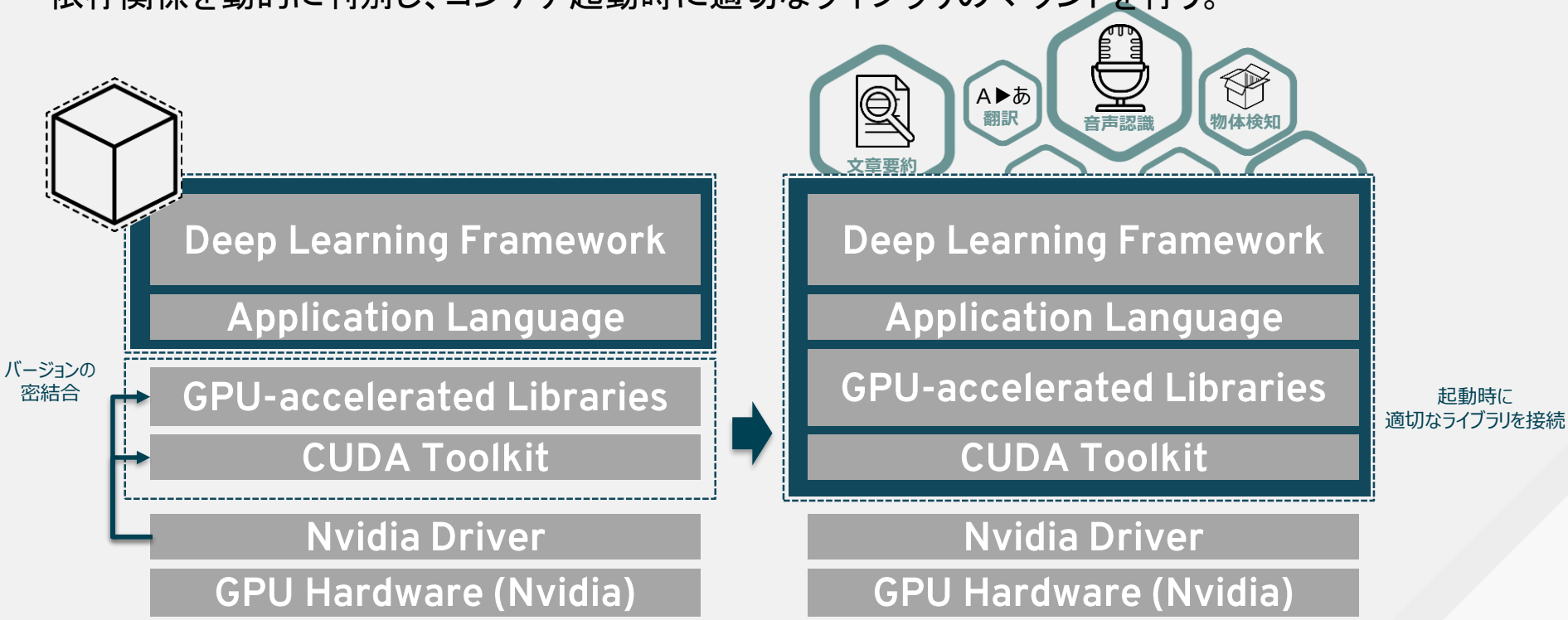
cuDNN / cuBLAS / cuFFT / cuSPARSE

バージョンの依存関係を正しく管理

NVIDIA

Nvidia-Docker

依存関係を動的に判別し、コンテナ起動時に適切なライブラリのマウントを行う。



NVIDIA GPU CLOUD(NGC)

Nvidiaのコンテナレジストリサービスにより DLやHPCの複雑な環境構築や構築工数から開発者を開放

GPUフレームワークがすぐに利用可能

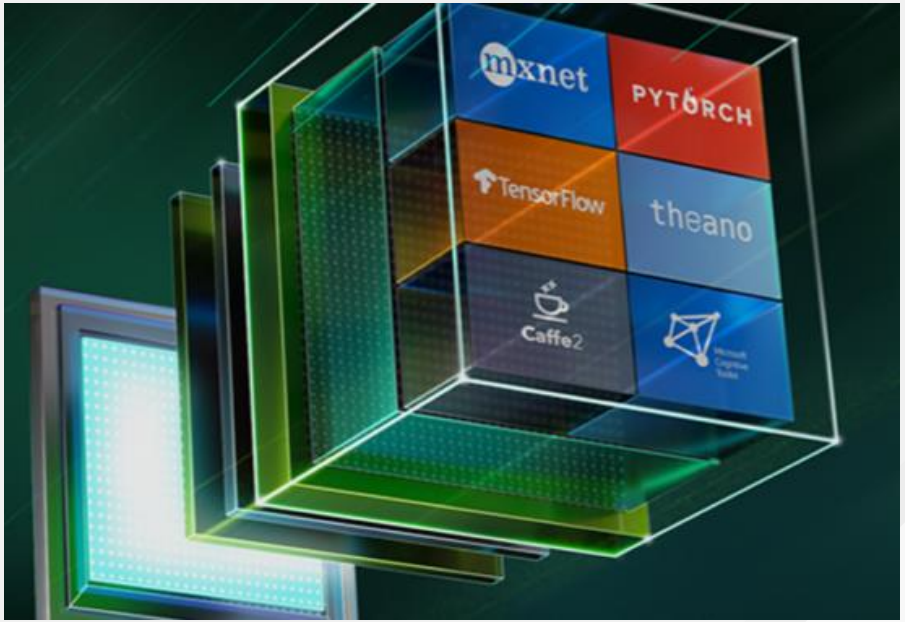
NVCaffe/Caffe2/Microsoft Cognitive Toolkit(CNTK)/DIGITS/MXNet/PyTorch/TensorFlow/Theano/Torch

NVIDIAがチューニング、テスト、動作確認済み

DLフレームワークは、最新のNVIDIA GPU上で高速に学習が行えるようチューニング。

最新の環境

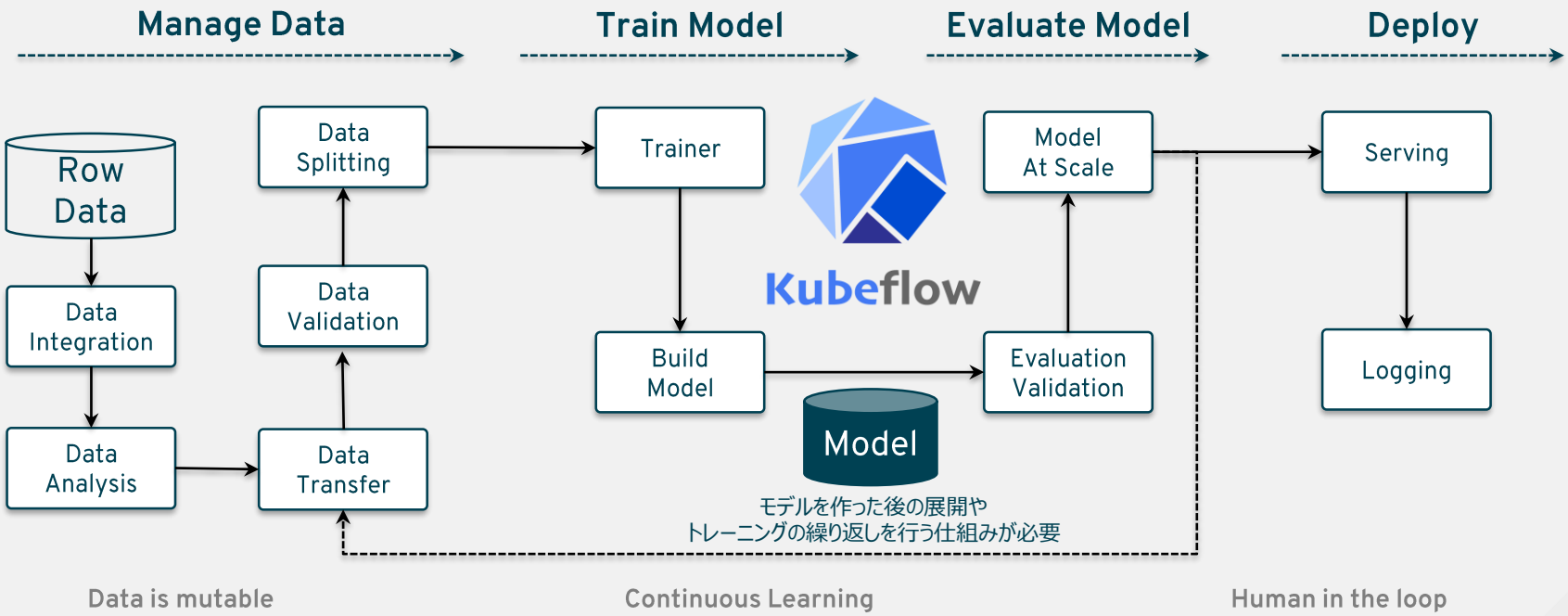
NVIDIAはライブラリ、ドライバ、コンテナを継続的に最適化し、更新プログラムを毎月提供。



参考: http://www.atmarkit.co.jp/ait/articles/1710/30/news058.html

Workflow for Kubeflow

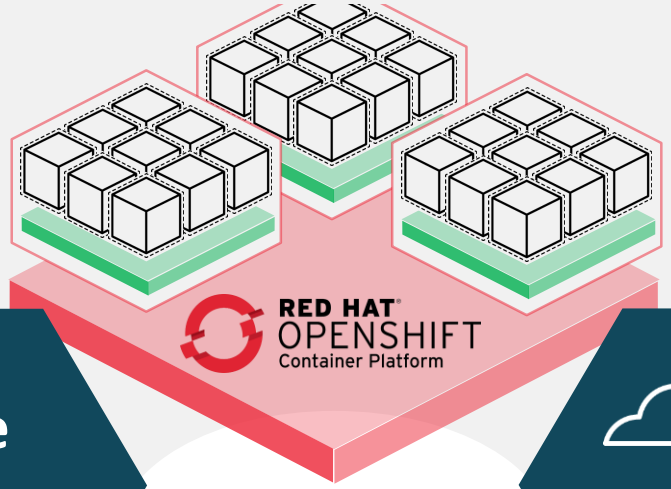
AI/MLアプリ開発におけるパイプラインの構築もコンテナによる制御が主流に



https://medium.com/intuitionmachine/google-and-ubers-best-practices-for-deep-learning-58488a8899b6

コンテナ化による自由なGPUリソースの選択

ユーザーは環境が異なるごとにアプリケーションをインストールする必要はなく、どの環境においても同等のワークロードを実行し、そのシミュレーション結果を得ることを実現する。



On-Premise

オンプレに専有した高性能なGPUリソース

NVIDIA

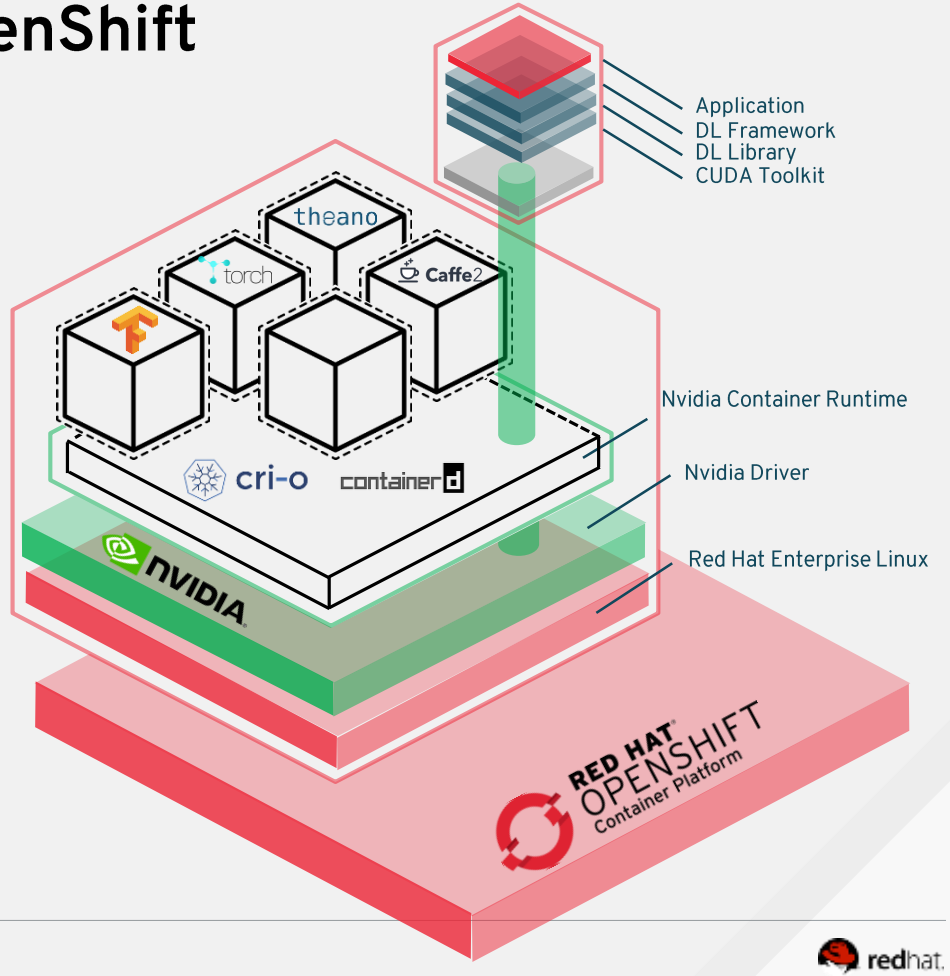
Public Cloud

クラウドプロバイダーによるGPUインスタンス

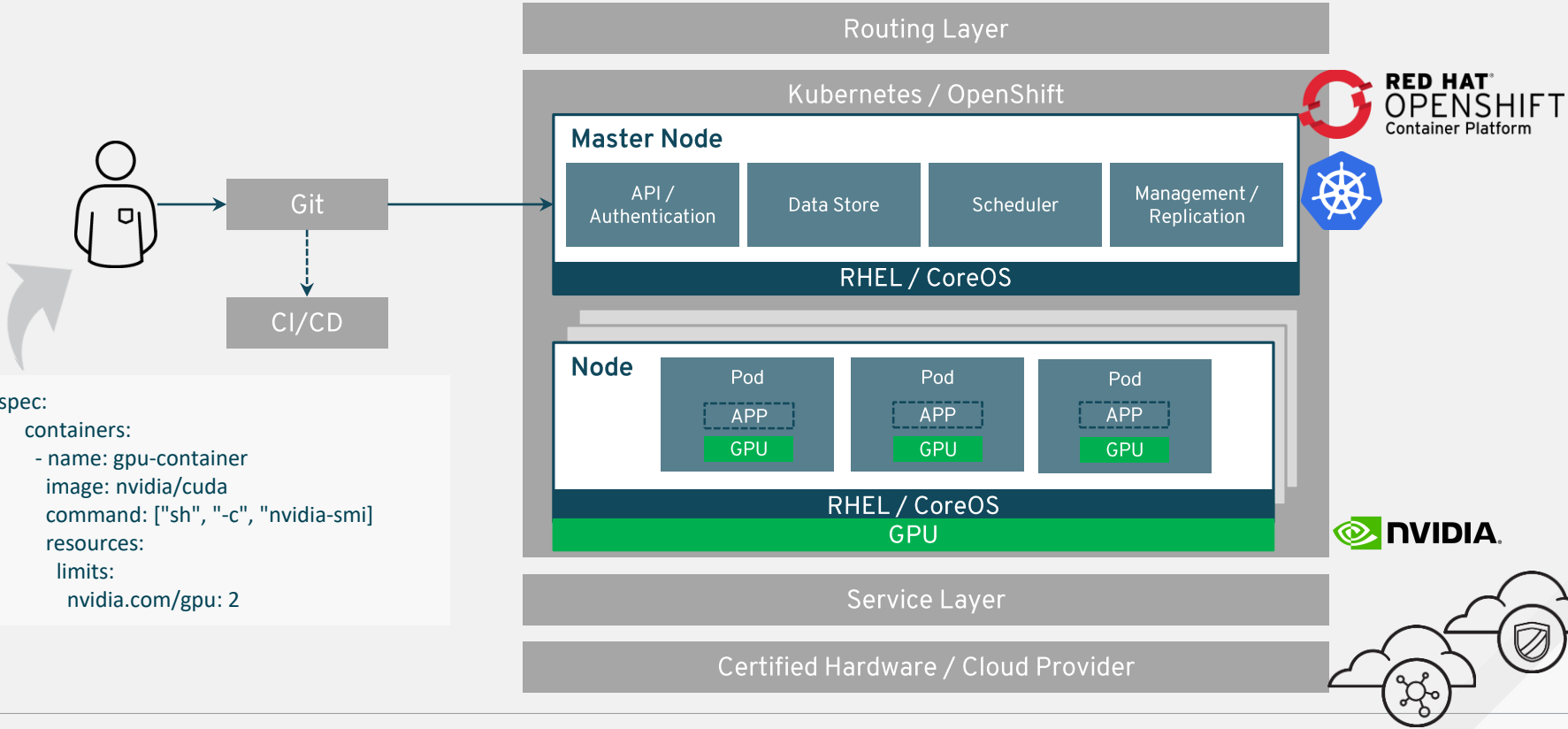
GPU スケジューリングの詳細

Device Plugin Support on OpenShift

- ✓ Joint collaboration with strategic partners for drivers, plugins and container images
- ✓ Device Manager GA
- ✓ Scheduler: Priority and preemption
- ✓ Seamless install experience of drivers, plugins and dependencies
- ✓ Container images in RHCC/ISV Registry
- ✓ Certifications and support

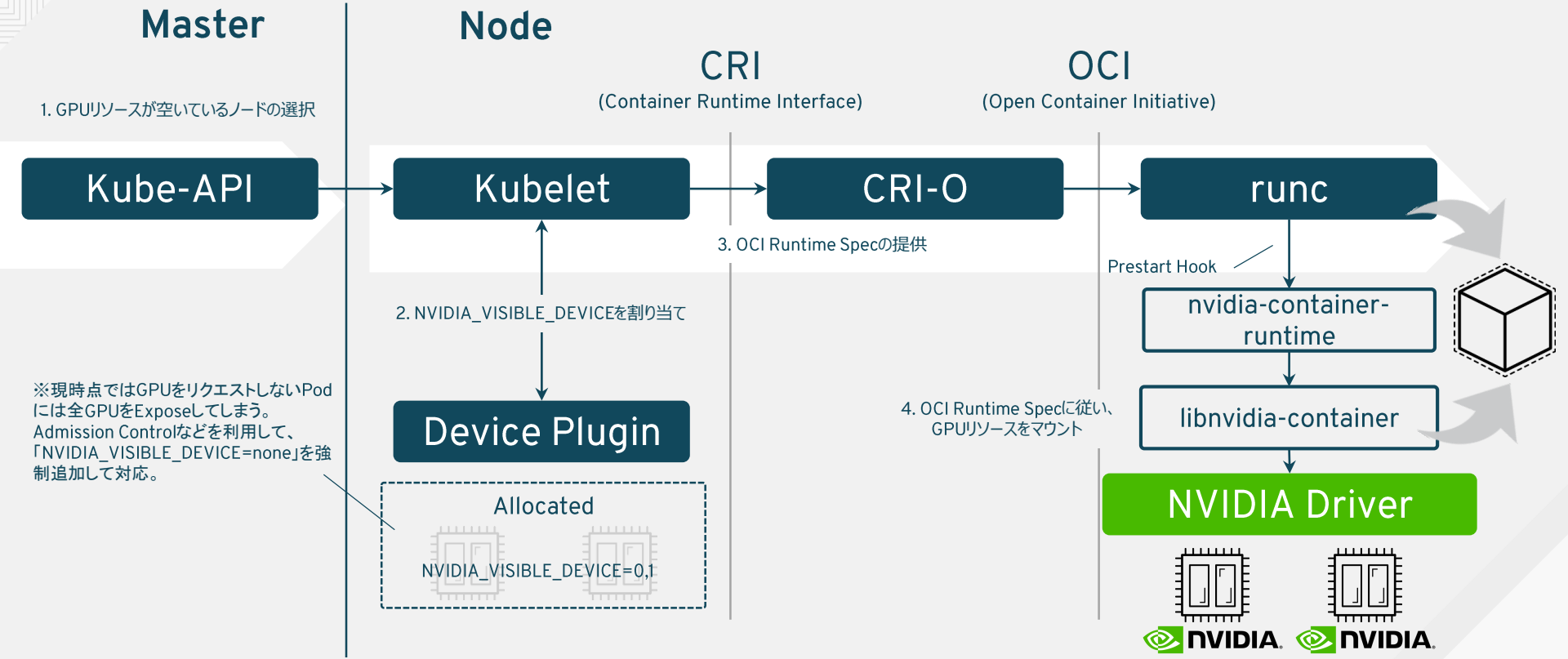


Kubernetes for GPU as a Service




```
spec:
    containers:
     - name: gpu-container
       image: nvidia/cuda
       command: ["sh", "-c", "nvidia-smi]
       resources:
         limits:
           nvidia.com/gpu: 2
```

GPU Allocation Overview



Allocate GPU Resources in Kubernetes

Kubernetesでは、ノードレベルでのExtended Resourceである「Device Plugin」によって、GPUのスケジューリングを実装し、GPUとコンテナをバインドします。

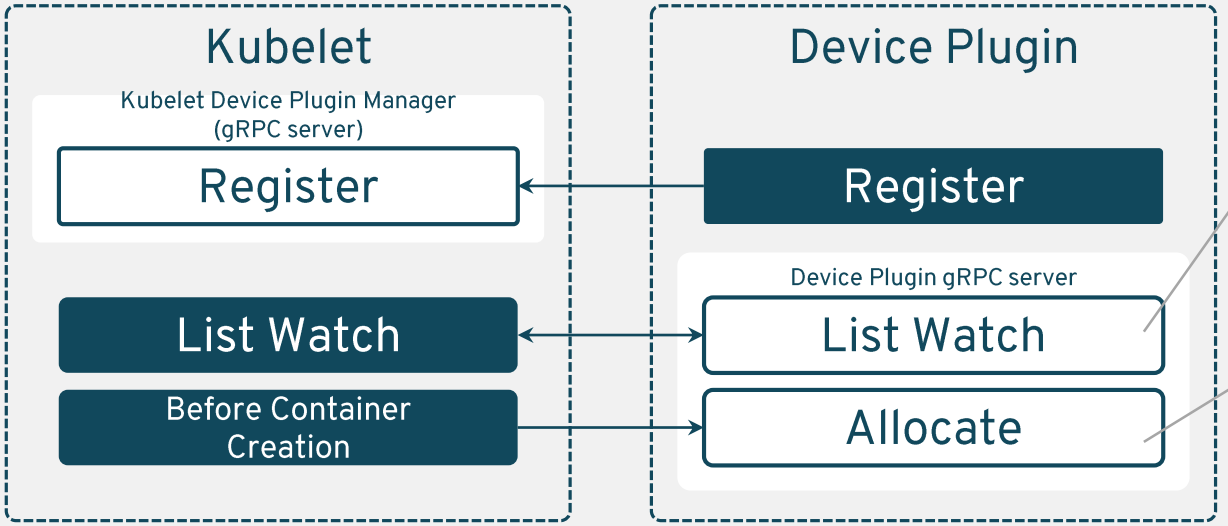
Extended Resource

kubernetes.ioドメイン以外のリソース(GPUやFPGA、Infinibandなど)を登録するもの。
これによりクラスターは外部リソースを提供でき、ユーザーはそのリソースを利用可能となる。

Extended Resourceを提供するためのステップは「Device Plugin」が担う。
1. 拡張リソースの宣言(登録)
2. Pod作成時に拡張リソースを要求

Device Plugin

Device Pluginの実体は、特定のハードウェアリソースを管理するノード(atomic-openshift-node.serviceの外部)上で動作するgRPCサービス。



Device Pluginを利用することによって、カスタムコードを記述することなく、Podに特定のデバイスタイプ(GPU/InfiniBand/ベンダー固有のリソース)を提供

Summary

Summary

Portable: コンテナ化による依存関係の保証

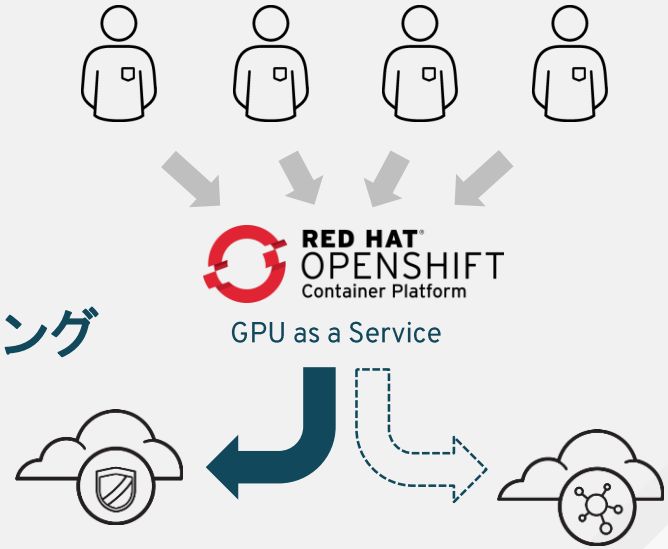
Deep Learningに必要なGPUリソースは、コンテナ化することによって、
ドライバやフレームワークのバージョン依存から開放される。

Composability: パイプラインの構築

AI/MLアプリ開発におけるプラットフォーム管理の煩わしさを排除

Scalability: 柔軟なGPUのスケジューリング

オンデマンドでGPUリソースの提供を行い、AI/MLアプリ開発を加速する



Announcement

Deep Learning開発用コンテナ環境構築サービス

概要

本サービスは、Deep Learning開発環境としてNVIDIA GPUを搭載したHPE Apollo 6500 Gen10 サーバーをRed Hat OpenShift Container Platform（以降OpenShift）のNodeサーバーとして構成し、その上でTensorFlowやChainer等のDeep Learningアプリケーションコンテナを稼働する環境を導入します。OpenShiftの特長でもある堅牢なテナント分離を実現するだけでなく、特定のGPUリソースを柔軟に各利用者のコンテナに割り当てることが可能となる、リソース、開発環境構築時間、コスト等あらゆる面で開発効率が飛躍的に向上しうる環境を提供いたします。

HPEが提供するDeep Learning開発用コンテナ基盤環境

HPE Apollo 6500 serversの特徴

HPE Apollo Systemは、ラックあたり最高レベルのパフォーマンスと効率性を実現する、Deep Learning 向けに最適化されたスケールアウト型GPUシステムです。

最高のGPU密度
NVIDIA Tesla GPU (PCIe もしくは NVLINK 2.0) を最大8基搭載可能

柔軟なストレージ構成
16本までのSATA/SAS/SSDもしくは4本までのNVMEを構成可能

GPUパフォーマンスを活かす
1または2CPUあたり最大8GPUを実現し、アプリケー ションに最適化

優れた管理性
iLO等のHPE ProLiant Gen10と 共通の管理環境を提供

Hewlett Packard Enterprise

OpenShiftによる開発環境のメリット

エンタープライズ向けコンテナオーケストレーション
Dockerと採用実績豊富なKubernetesをネイティブに統合、エンタープライズ向け認証・SDN・Webコンソール・運用管理等の機能も充実

GPUリソースのマルチテナント毎割り当てが可能
最大8基のNVIDIA GPUのワークロードに合わせた柔軟な割り当て、テナント毎のリソース分離とアクセス制御を実現。セキュリティ、パフォーマンスの双方において、開発効率が飛躍的に向上します。

異種・複数バージョンDeep Learningフレームワークの混在可
Deep LearningフレームワークとCUDA Toolkitライブラリをコンテナ化。複数フレームワークの開発環境を同一プラットフォーム上で利用可能です。

redhat.

Red Hat Forum Tokyo 2018 開催決定！

IDEAS WORTH EXPLORING

東京｜11月8日（木）

ウエスティンホテル東京

〒153-8580 東京都目黒区三田1-4-1

https://redhat-forum.jp/

本年度は大阪での開催も予定しております。

Red Hat Forum Osaka 2018

大阪｜12月12日（水）ヒルトン大阪

皆さまのご参加をお待ちしております。

redhat.

redhat.

THANK YOU