

クラウド環境構築システムVCPによる Open OnDemand環境の簡易構築

2023年10月11日
大江 和一

国立情報学研究所 クラウド基盤研究開発センター

PCCC Open OnDemandワークショップ

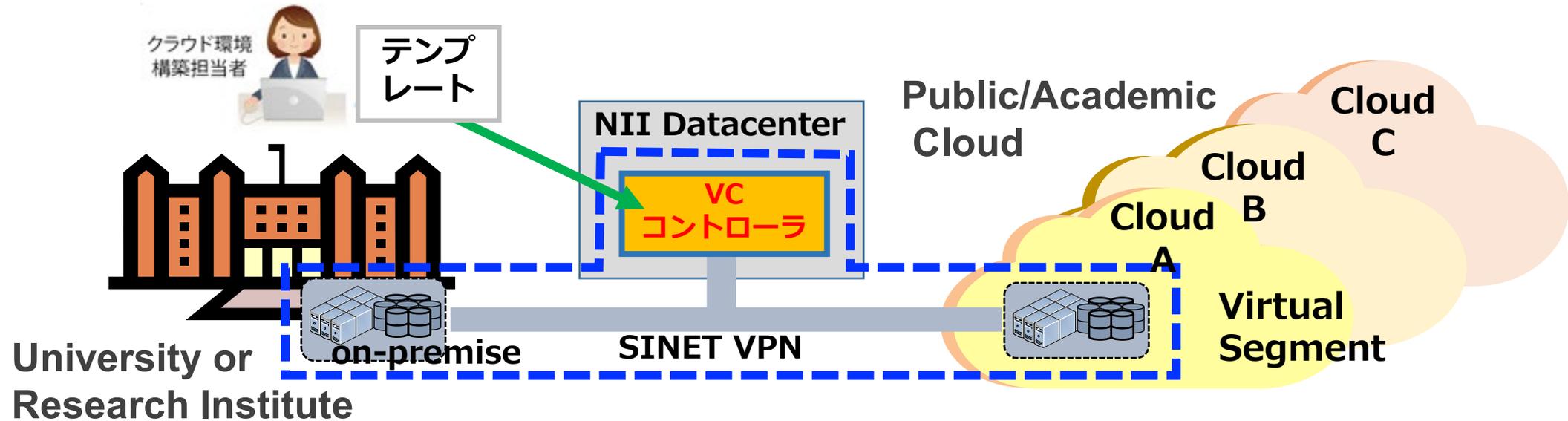
はじめに

- 初心者がHPCクラスタの前提知識なしでアプリケーション実行が可能なOpen OnDemandの利用が進んでいる。しかしながら、その構築・運用は初心者にとって容易ではない。
- NIIでは、SINETとクラウドを活用して再現性のある研究教育用アプリケーション環境の構築・運用を支援する学認クラウドオンデマンド構築サービス（OCS）を提供しており、複数の学術機関に利用頂いている。VCP（Virtual Cloud Provider）は、OCSを実現するための中核となるシステムである。
- 本講演では、VCPを用いることでOpenHPC環境とOpen OnDemand環境を容易に構築できることを紹介する。

Outline

1. 基盤ソフトウェアVCP、及び学認クラウドオンデマンド構築サービス（OCS）の概要
2. VCPを用いたOpenHPC環境構築/Open OnDemand環境構築のご紹介
3. 操作例（ハンズオンセミナーのビデオ（約3分））
4. アプリケーションテンプレートのご紹介
5. まとめ

VCP(Virtual Cloud Provider)とは



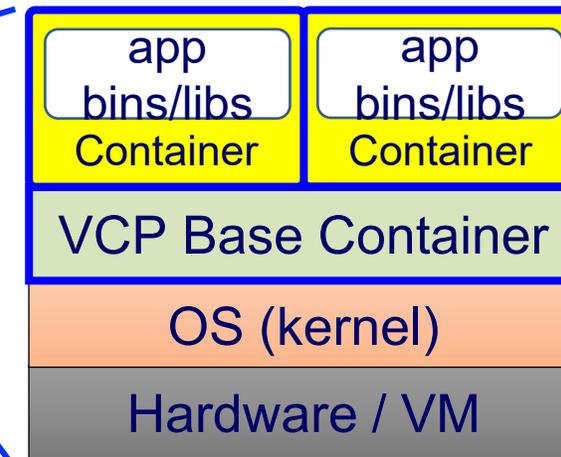
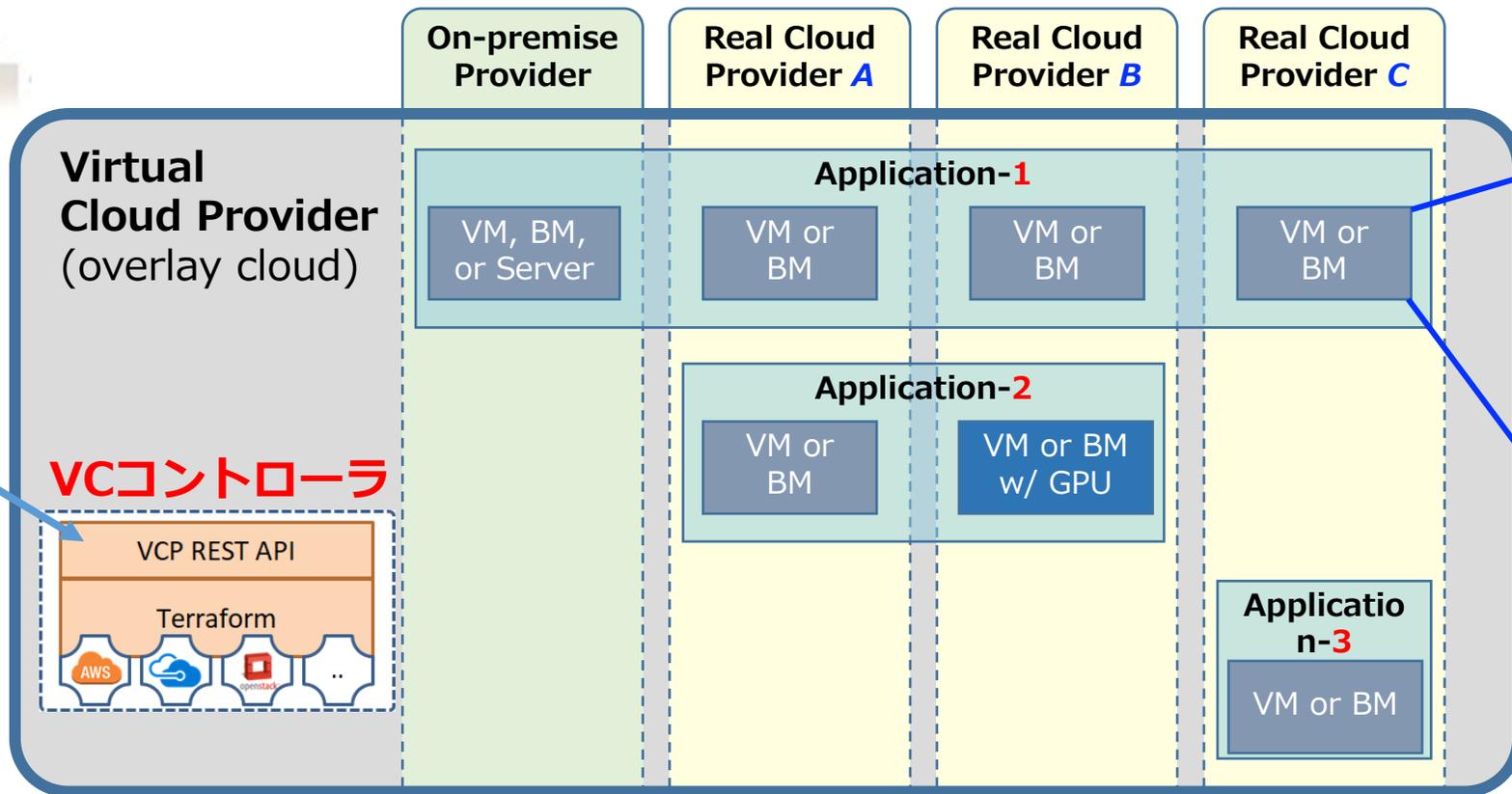
- VPNで接続された1つ以上のクラウドとオンプレミスシステム（仮想クラウド）でアプリ環境構築を支援
 - クラウドの異なる制御APIの差異を吸収する**VCコントローラ**を提供
 - アプリケーションテンプレートによるアプリケーション環境構築支援
 - アプリケーションの構築・運用手順を記述したJupyterNotebook形式のファイル

VCPの概要



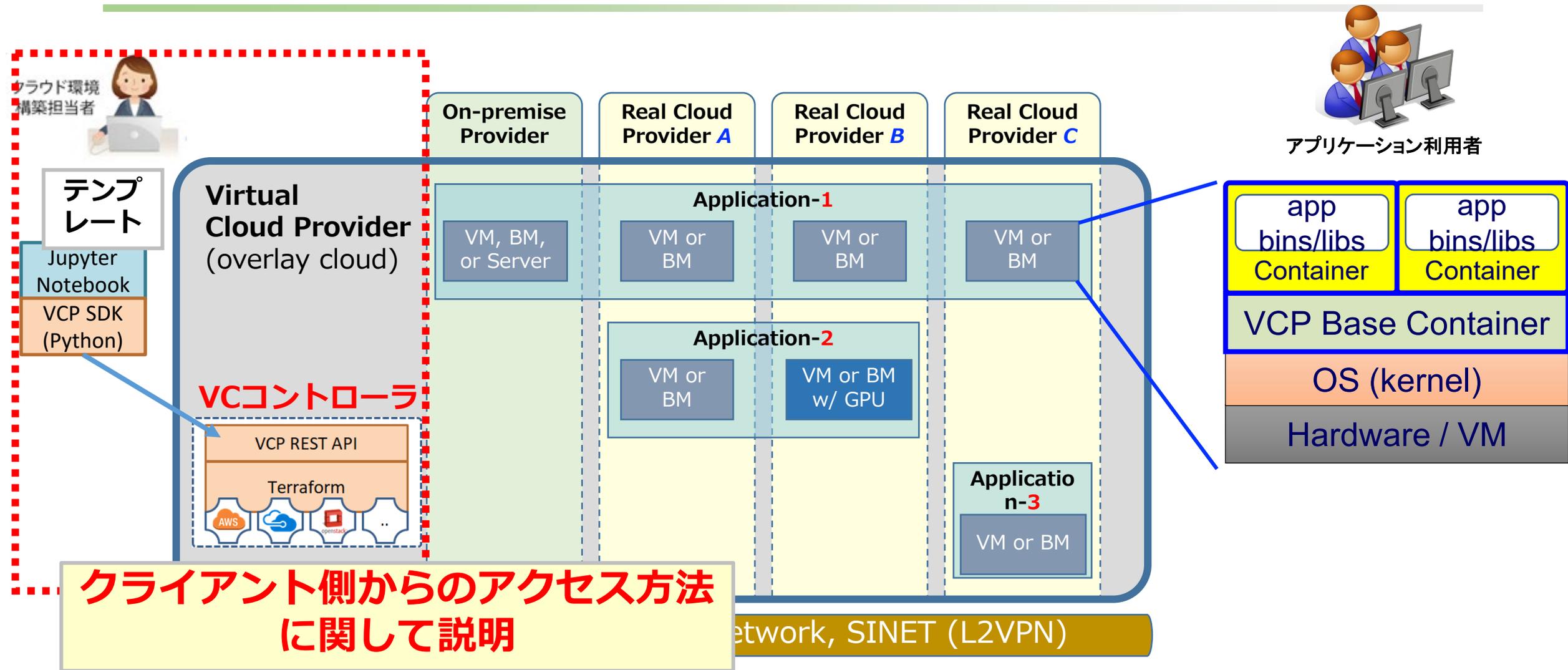
テンプレート

- Jupyter Notebook
- VCP SDK (Python)



Academic Hi-Speed Backbone Network, SINET (L2VPN)

VCPの概要



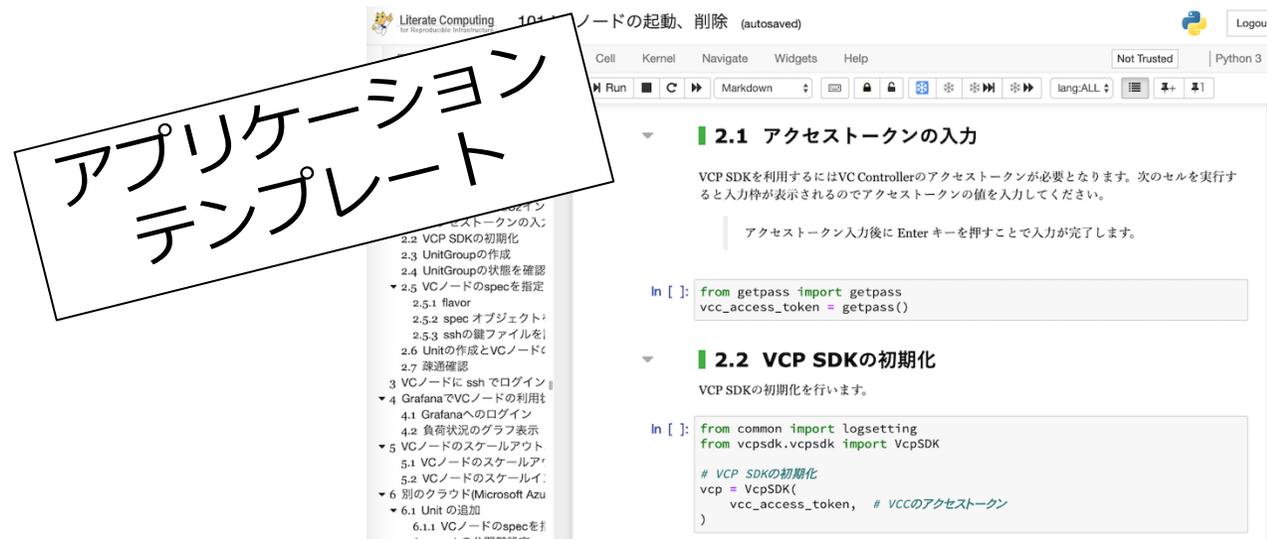
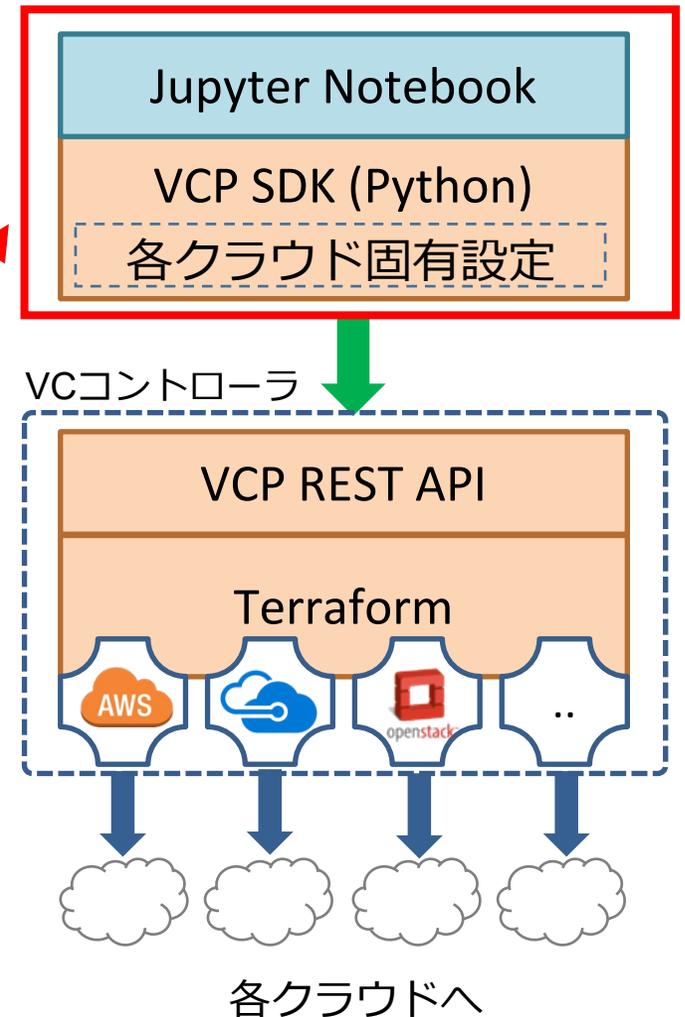
クライアント側からのアクセス方法

■VC利用者

- アプリケーションテンプレート*からVCP SDKを利用してVCP REST APIを操作

■VCコントローラ

- VCP REST API (プロバイダIFを抽象化) からTerraformを介して各クラウドを操作



*: Jupyter Notebook (+NII拡張) で記述

アプリケーションテンプレート

他者が作ったテンプレートの流用も可能

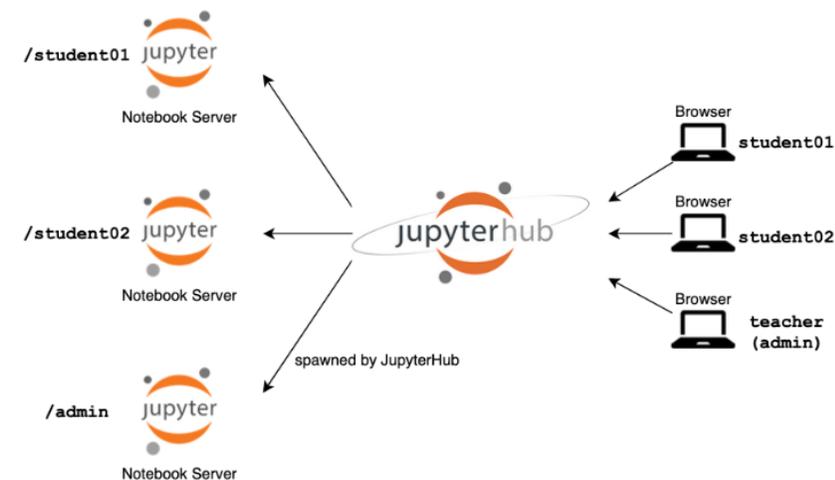
The Littlest JupyterHub による軽量Python実習環境の構築

JupyterHub は、Webブラウザからアクセス可能なマルチユーザ対応の認証機能付きJupyterNotebookサーバです。

JupyterHubを利用して管理者が用意したNotebookをユーザがブラウザからすぐに実行可能な環境を提供できるため、Pythonによるプログラミング研修やワークショップを開催したり、講義演習環境として活用したりするのに適しています。

本ハンズオンでは、JupyterHubを小規模なグループで手軽に利用することを想定し、単一のサーバで実行するために開発された「The Littlest JupyterHub」(以下「TLJH」と略)をVCPを用いて構築します。

ハンズオンご参加の皆様には、このテンプレートでTLJHによるVCPアプリケーション環境を構築していただきます。



構築環境情報の入力

TLJH環境の構築情報を入力します。必要に応じ、下記の情報を修正してください。

★ハンズオンでは以下のパラメータを変更しないでください★

```
#####
### ハンズオンでは以下のパラメータを変更しないでください。###
#####

# UnitGroup名
ugroup_name = 'handson403'

# プロバイダ
vc_provider = 'aws'
```

図表を組み合わせた説明を挿入できる

VCノードのspecを指定する

TLJHを利用するのに十分な性能・容量のノードspecを指定します。
固定割当IPアドレスは、ハンズオン環境のNAT Proxyサーバに予め設定されているIPアドレスを使います。

```
In [ ]: # UnitGroup の作成
unit_group = vcp.create_vgroup(ugroup_name)

# VCノード spec
spec = vcp.get_spec(vc_provider, vonode_flavor)

# spec オプション (ディスクサイズ 単位:GB)
spec.volume_size = volume_size

# spec オプション (固定割当IPアドレス)
spec.ip_addresses = [fixed_ipaddress]

# ssh keyfiles
import os
ssh_public_key = os.path.expanduser("~/ssh/id_rsa.pub")
spec.set_ssh_pubkey(ssh_public_key)
```

Unitの作成とVCノードの起動

Unitを作成します。Unitを作成すると同時にVCノード（ここではAmazon EC2インスタンス）が起動します。処理が完了するまで1分半~2分程度かかります。

```
In [ ]: # Unitの作成 (同時に VCノードが作成される)
unit = unit_group.create_unit('tljh-node', spec)
```

疎通確認

まず、sshのknown_hostsの設定を行います。

その後、VCノードに対してuname -aを実行し、ubuntu x86_64 linux が起動していることを確認します。起動していない場合は、spec.image に誤りがあります。本テンプレート下部にある「環境の削除」を実行し、spec.image を修正、全てのセルを unfreeze してから、最初から再実行してください。

```
In [ ]: # unit_group.find_ip_addresses() は UnitGroup内の全VCノードのIPアドレスのリストを返します
ip_address = unit_group.find_ip_addresses(node_state='RUNNING')[0] # 今は1つのVCノードのみ起動しているので [0] で最初の要素を取り出す
print(ip_address)

# ssh 設定
!touch ~/.ssh/known_hosts
!ssh-keygen -R $ip_address # ~/.ssh/known_hosts から古いホストキーを削除する
!ssh-keyscan -H $ip_address >> ~/.ssh/known_hosts # ホストキーの登録

# システムの確認
!ssh $ip_address uname -a
```

TLJH (The Littlest JupyterHub)

VCノード上に、本ハンズオン

TLJHコンテナイ

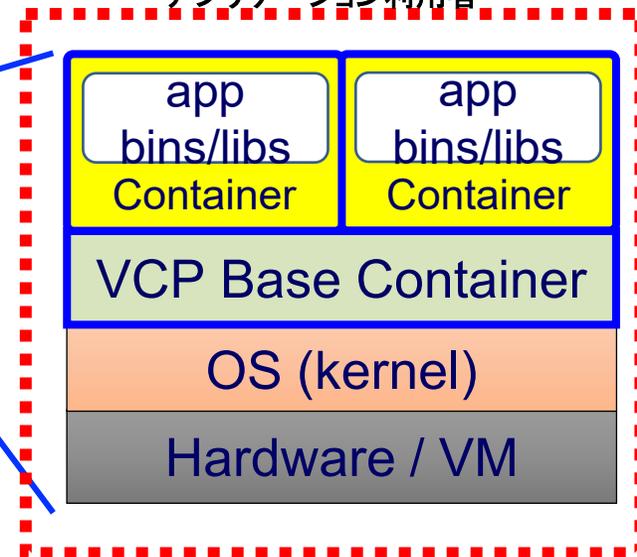
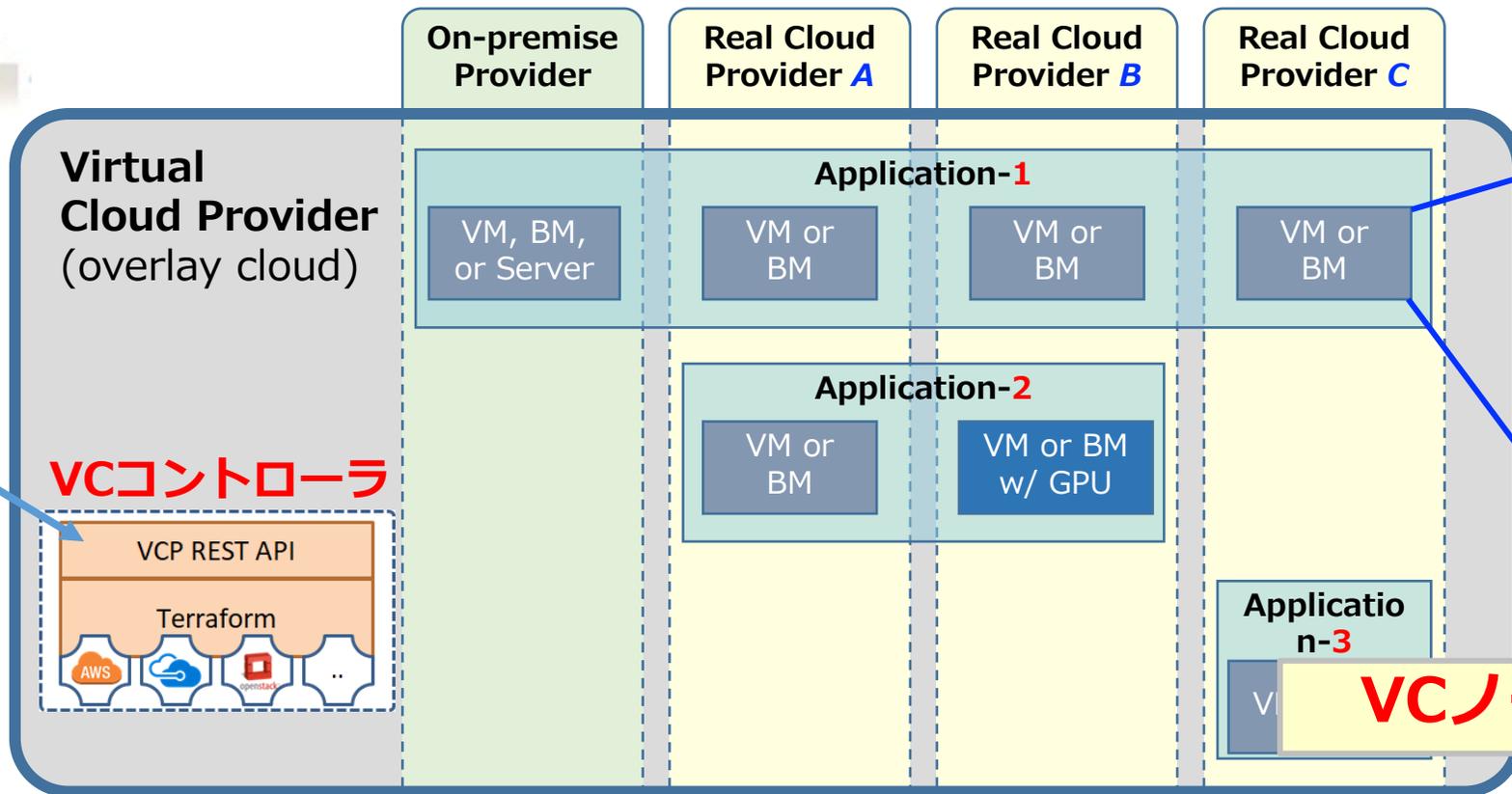
VCノード上にコンテナイメージを取得するために docker pull を実行します。

スクリプトを組み込むことができ、ここから実行できる。実行結果を残すことも出来る。

VCPの概要



テンプレート
Jupyter Notebook
VCP SDK (Python)



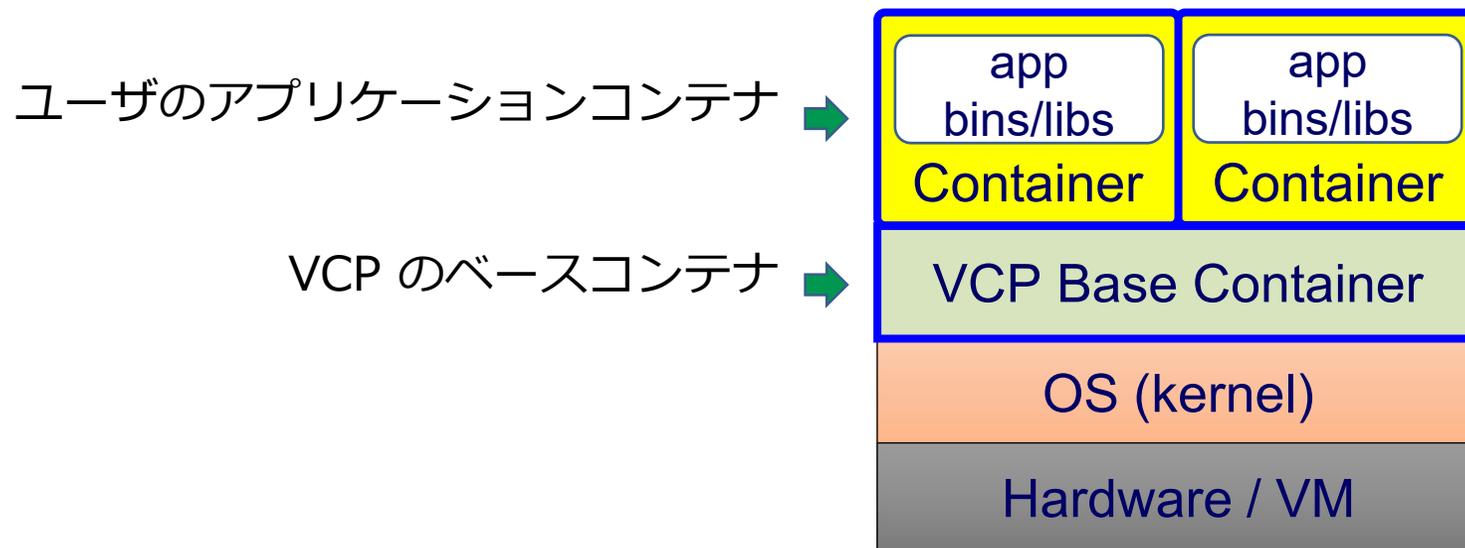
VCノードの構成を説明

Academic Hi-Speed Backbone Network, SINET (L2VPN)

VCノード構成

■ Docker in Docker構成

- ベースコンテナ
 - 死活監視やメトリクス収集などシステムの基本機能
- アプリケーションコンテナ
 - アプリケーションと関連ソフトウェアをベースコンテナ上に起動



VCPとOCSの関係

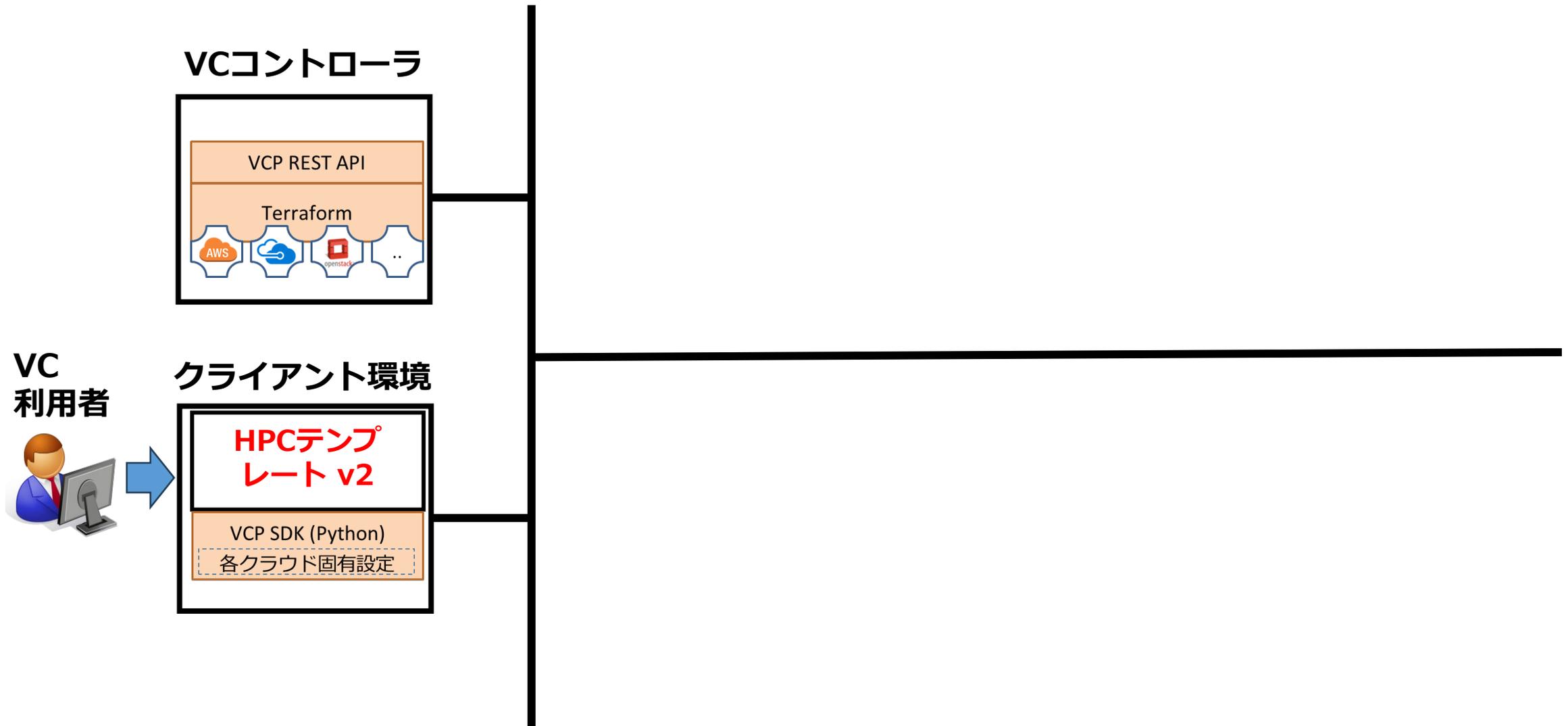
■学認クラウドオンデマンド構築サービス（OCS）

- あらかじめ用意されたテンプレートを指定して実行するだけで、大学・研究機関が契約しているSINETと連携したクラウドに計算資源を確保し、その上で動作するアプリケーション環境のインストールや設定までを自動的に行うことが可能となるNIIのサービス。
- その中核となるソフトウェアがVCP

<https://cloud.gakunin.jp/ocs/>

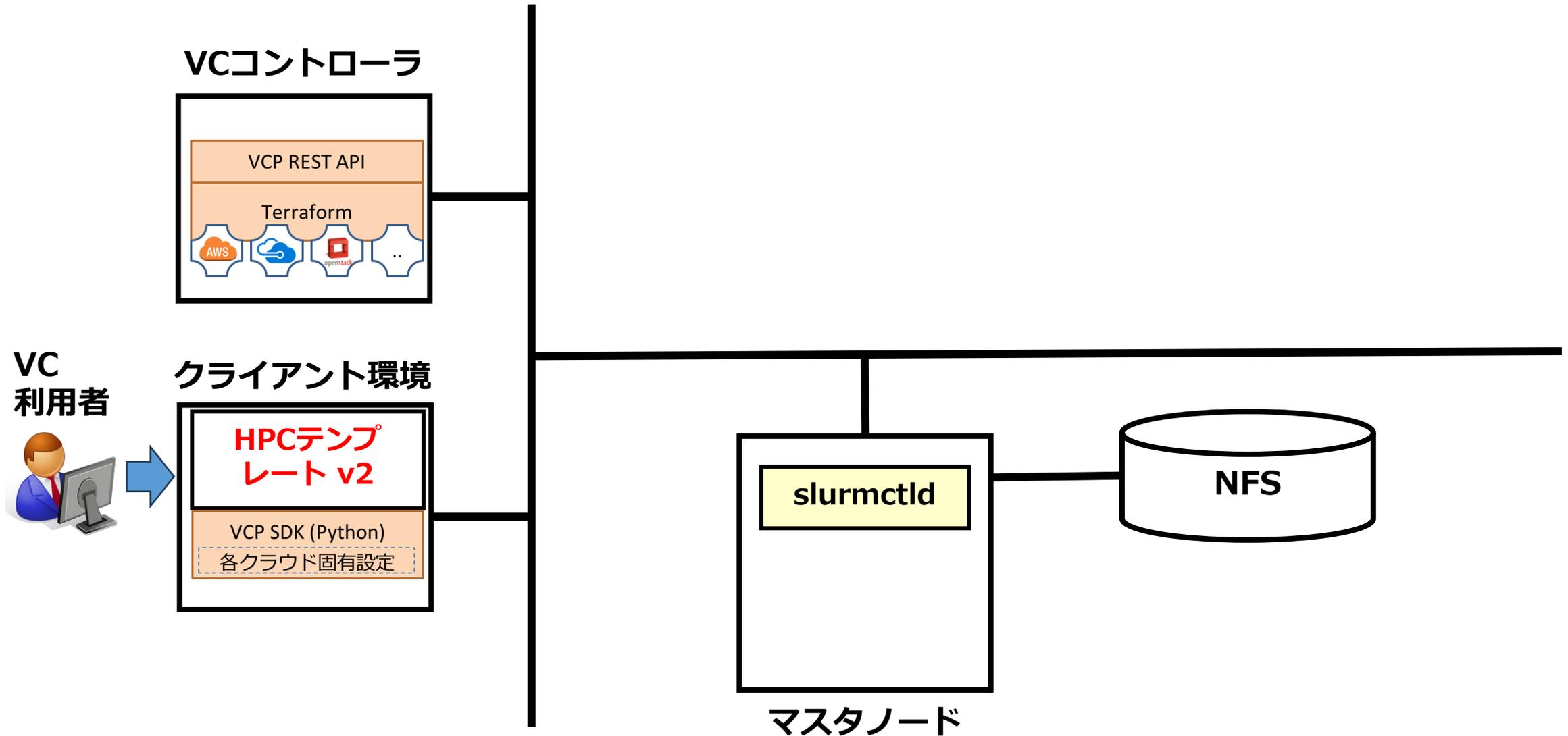
Open OnDemand環境の構築

初期状態



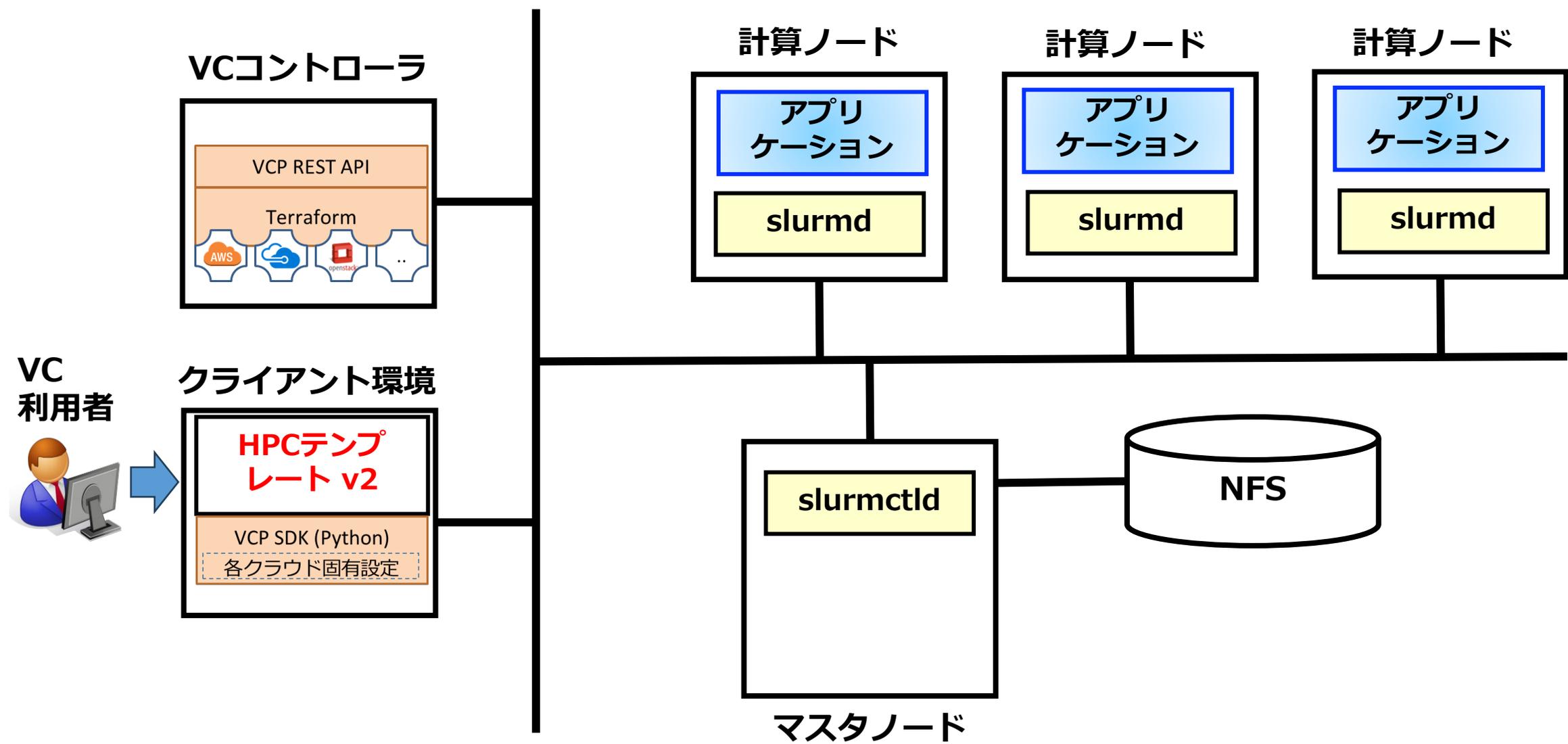
Open OnDemand環境の構築

マスタノード、NFSサーバ起動



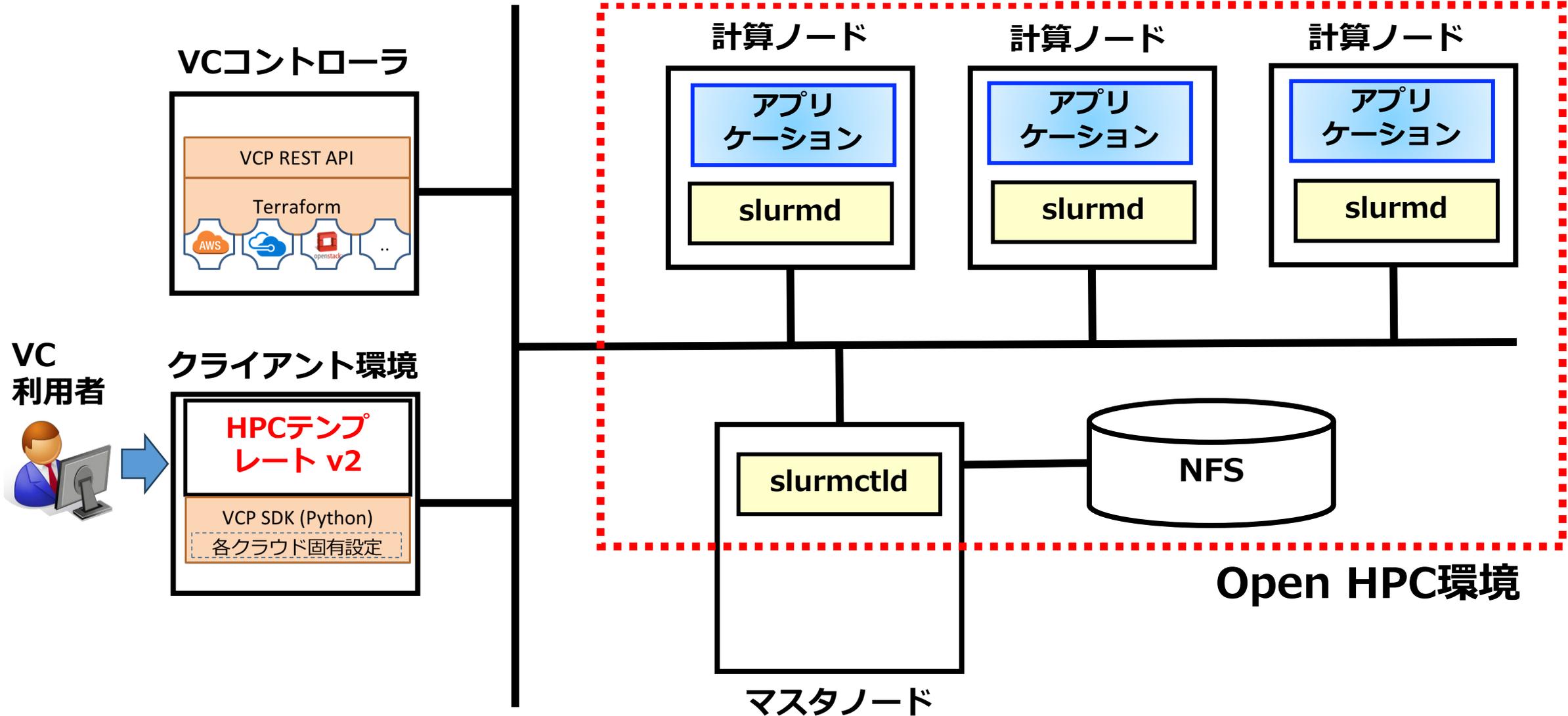
Open OnDemand環境の構築

計算ノード起動



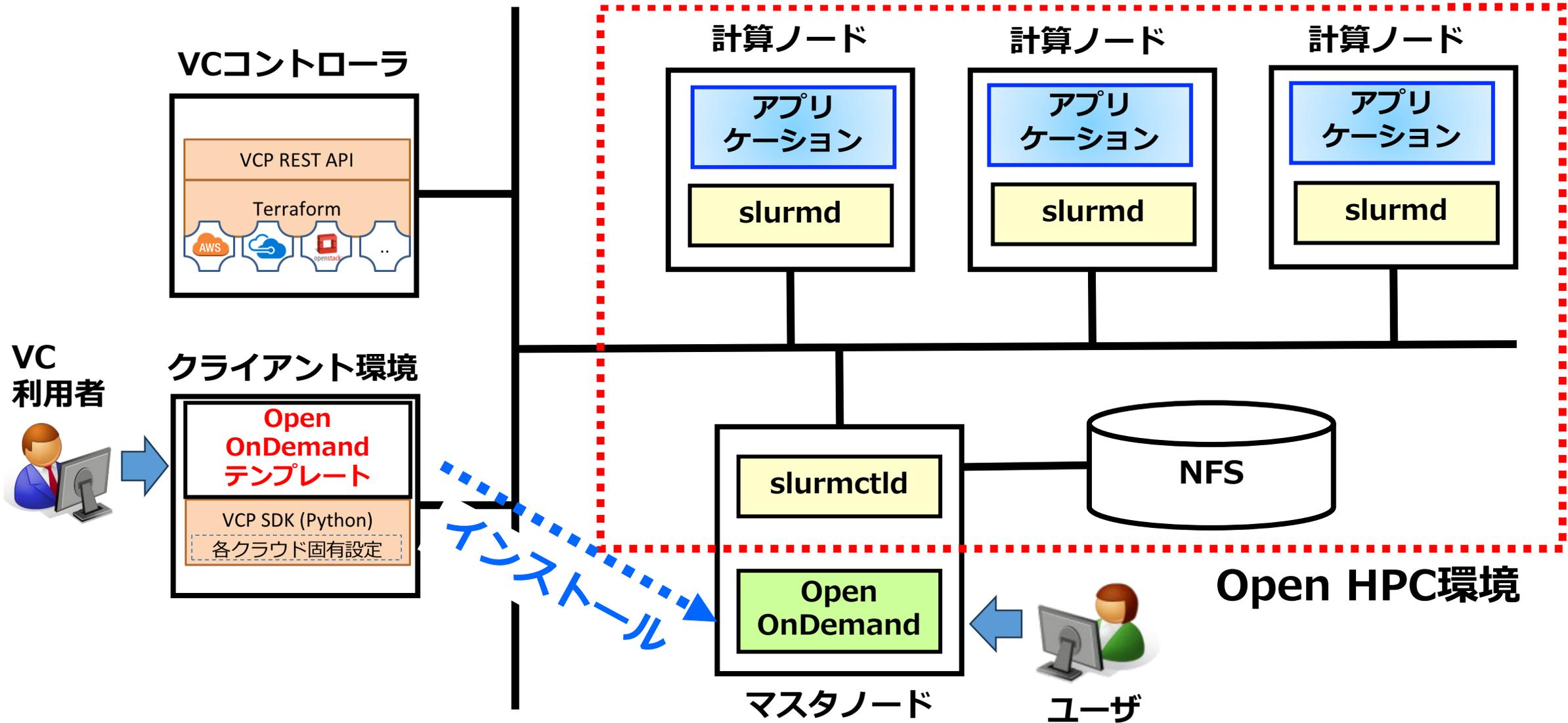
Open OnDemand環境の構築

環境設定後にOpenHPC環境の利用が可能に



Open OnDemand環境の構築

Open OnDemandテンプレートを用いて、Open OnDemand環境を構築



操作例（ハンズオンセミナーのビデオ（約3分））

- HPCテンプレートv2を用いたOpenHPC環境構築
- Open OnDemandのインストール
- Open OnDemand フロントエンド構築とアクセス確認
- Open OnDemand ジョブ実行
- Open OnDemand JupyterNotebookサーバ起動
- Open OnDemand Linpack起動

アプリケーションテンプレート

- HPCテンプレート v2 (VCP SDK v21.04対応 (AWS、Azure、Oracle Cloud、mdx で動作確認済み))
 - OpenHPC v2.xで配布されているパッケージを利用して、クラウド上にHPC環境を構築するテンプレート。v1 の機能に加え、GPUノードの利用とNVIDIA社のNGCカタログのコンテナの実行が可能
- Open OnDemand構築テンプレート (VCP SDK v21.04以降対応 (mdxで動作確認済み))
 - HPCテンプレートv2で構築したOpenHPC環境上にOpen OnDemand環境を構築する



アプリケーションテンプレート (続)

- 手書き文字認識システム構築テンプレート (VCP SDK v21.04以降対応 (AWS、Azureで動作確認済み))
 - Open HPC v2テンプレートをベースにGPU ベースの学習システム (Tensorflowを使用) の構築とCPUベースの認識システム (独自仕様) の構築を行い、フロントエンドとしてJupyterNotebook上に手書き数字認識システムを動作させる
- 計算資源補完テンプレート (VCP SDK v20.04以降対応 (AWS、Azureで動作確認済み))
 - オンプレミスのバッチ型計算機システムの計算ノード不足時に、クラウド上に同じソフトウェア構成を持つ計算ノードを自動的に立ち上げ、バッチシステムに組み込むクラウドバースト機能を提供。Torque等クラウドに対応していないバッチシステムでも、簡単なプラグインを作成することでクラウドバーストが可能。なお、本テンプレートはipynb形式ではなく、Pythonならびにbashスクリプトで記述されている

アプリケーションテンプレート (続)

- LMS (Learning Management System), JupyterHubなどのアプリケーションテンプレートも公開中です。

URL: <https://github.com/nii-gakunin-cloud/ocs-templates/>

まとめ

- VCP (Virtual Cloud Provider) を用いることで、OpenHPC環境とOpen OnDemand環境を容易に構築することをご紹介
- 学認クラウドオンデマンド構築サービス（OCS）にお申込みいただければ、VCPを利用することができます！

<https://cloud.gakunin.jp/ocs/>

- 各種お問い合わせはいかにお願いいたします。

NIIクラウド支援室 cld-office-support@nii.ac.jp

よろしければ、OCSユーザ登録をお願い致します！

OCSユーザ登録フォーム / OCS User Registration Form

OCSは、クラウドやネットワークに関する知識や設定方法を熟知していないユーザでも研究教育用のアプリケーション環境を主要なクラウドでオンデマンドに容易に再現できるサービスです。国立情報学研究所では、OCSの管理システムプログラムをOCSポータブル版として公開しています。ポータブル版をご利用される場合は、以下よりユーザ登録をお願いいたします。ユーザ登録していただくこと、ユーザ用メーリングリストに登録され、OCSに関するリリース、バグフィックス情報、FAQのご紹介等の情報提供をさせていただきます。

OCS is a service that allows users to easily build a research and educational application environment on demand in major clouds, even if the users are not familiar with cloud computing and networking configuration methods.

The National Institute of Informatics has released the OCS management system program as the OCS portable version. To use the portable version, please register using the form below. You can subscribe to the OCS user mailing list to receive information such as releases, bug fixes, FAQs, etc.

下記の必要事項をご入力ください。

*メールアドレスは、所属機関ドメイン名のメールアドレスをご入力ください。

Please enter the required information below.

* Please enter the e-mail address of your institution domain name.

* の項目は必ずご入力ください。The fields marked by * must be filled.

お名前 / Full Name*	<input type="text"/> 例) 字認 クラウド / e.g. Gakunin Cloud
ふりがな / Last Name*	<input type="text"/> 例) がくにんくらうど / e.g. Gakunin
ご所属 / Affiliation*	<input type="text"/> 例) 国立情報学研究所 / e.g. National Institute of Informatics
メールアドレス / E-mail Address*	<input type="text"/> 半角のみ：所属機関ドメイン名のメールアドレス / The e-mail address of your institution domain name.
備考 / Remarks	<input type="text"/>

ご入力頂いた個人情報につきましては、SINETStreamに関する情報提供等、国立情報学研究所クラウド支援室が適切に管理・利用いたします。法令等により要求された場合を除き、本人の同意なく第三者に提供することはありません。同意される場合は、下部の「同意します」にチェックを入れて次へお進み下さい。

Regarding your personal information, the Cloud Support Office of the National Institute of Informatics properly manages and uses for the provision of SINETStream information. Except as required by law, we will not provide user information to third parties without your consent. If you agree, check "同意します / I agree" at the bottom and proceed to the next.

同意します / I agree

https://reg.nii.ac.jp/m/ocs_user_registration

- ご登録頂ければ、OCSに関する各種情報（OCSハンズオンセミナー、ユーザズミーティング、など）をお送りさせていただきます！
- OCSの利用を検討中の状態でも登録頂けます！

Thank You.