



openstack

5

周年特別企画：
OpenStack Summitの歩き方

Summitで人・ひと・ヒト。
いま熱く&深く語ろうコンテナとNeutron DVR
– Container Deep Dive –

2015/07/13

日立製作所 研究開発グループ
情報通信イノベーションセンタ クラウド研究部

西島 直

■ 西島 直 (にしじま なお)

- 株式会社日立製作所 研究開発グループ
情報通信イノベーションセンター クラウド研究部 研究員
兼 OSSテクノロジー・ラボラトリ員
- OSSの評価・検証、upstream活動、現場へのOSS適用を担当
 - Linux Kernel, KVM, OpenStack, ...



■ 発表、寄稿

- OpenStackユーザ会 第17回勉強会発表
「仮想マシンのリソース制限機能」
 - <http://openstack.jp/news/20140120.html>
- 学会誌「情報処理」 Vol.56 No.3 特集 OSSにみるITの最新動向
「OSS/OpenStackに見るクラウド基盤技術」
 - https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=113099&item_no=1&page_id=13&block_id=8

Container Deep Dive

- Introduction
 - コンテナ
 - Docker
 - コンテナマネージャ
- OpenStack Summitのコンテナセッション
- OpenStackのコンテナの取り組み
 - コンテナ対応状況
 - 新しいプロジェクト

Container Deep Dive

- **Introduction**
 - コンテナ
 - Docker
 - コンテナマネージャ
- OpenStack Summitのコンテナセッション
- OpenStackのコンテナの取り組み
 - コンテナ対応状況
 - 新しいプロジェクト

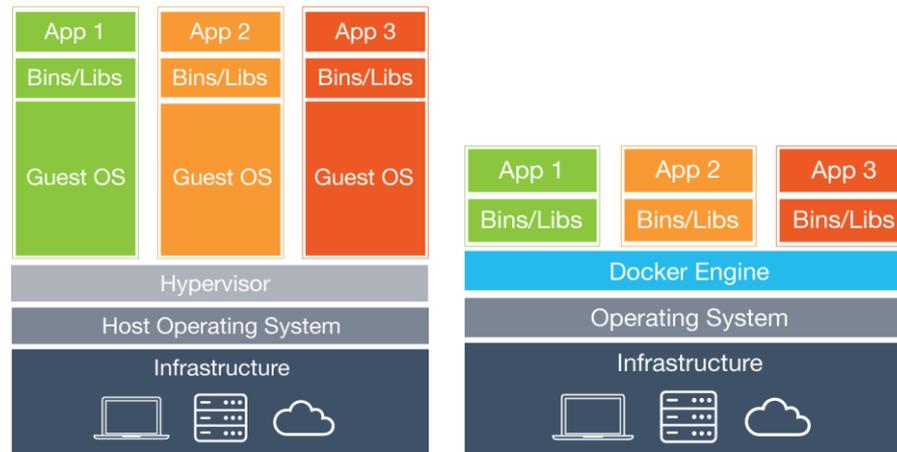
■ コンテナは隔離環境であり昔からある技術

- FreeBSD Jail (1998), Solaris Zones (2005), OpenVZ (2005), LXC (2008), Docker (2013), Rocket (2014)
- Dockerも隔離環境を作るためにコンテナを利用している

5

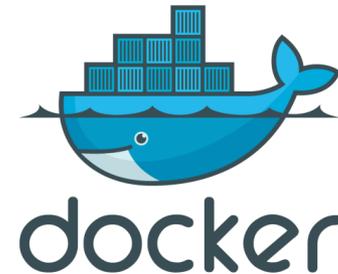
■ コンテナと仮想マシンの違いは？

- 同じところ：リソースの分離と割り当て
- 違うところ：それらの実現方法
 - VM：アプリ、バイナリ、ライブラリ、OSすべてを隔離
 - コンテナ：アプリとその依存関係を隔離、カーネルは共有する



■ Dockerはアプリケーションに関連する全てのものをパッケージ化

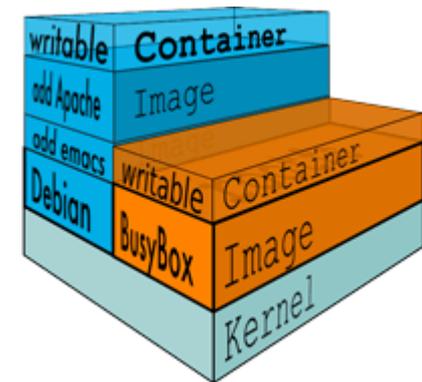
- Build, Ship and Run Any App, **Anywhere**
- 1回ビルドすればどこでも動く
- Dockerfileでインフラをコードとして管理



5

■ どうやって実現しているのか？

- Union file system(aufs, overlayfs, btrfs)、DM
複数のディスクを透過的に使用できるので、
read-onlyなファイルシステムの上に
書き込み可能なファイルシステムを重ねている
- コミット
コミットすると最上段のファイルシステム
(read+write) をread-onlyにして、その上に
ファイルシステム (read+write) をmountする



<https://www.docker.com/whatisdocker>
<https://blog.docker.com/2014/03/docker-0-9-introducing-execution-drivers-and-libcontainer/>

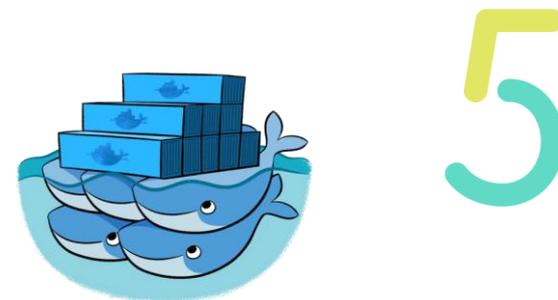
■ Kubernetes

- Linuxコンテナのクラスタを管理する
- コンテナのオーケストレーション
 - コンテナのスケジューリング
 - コンテナクラスタのワークロードの管理
- Google内部で利用していた**Borg**というツールを簡略化したもの
 - Googleは10年前から自社のサービスにコンテナを利用している
 - 1週間で20億のコンテナが起動している
- Kubernetes自体はオートスケーリング機能はないがGCE上だとオートスケールや監視サービスが提供されている
- Pre-Production Beta
 - 7/21にOregon Convention CenterでKubernetes 1.0のイベントあり
 - Event Sponsorsには、Intelや**OpenStack**, Mirantis, Red Hat, Hitachi,...



■ Docker Swarm

- Docker社が公開したDockerコンテナのクラスタを管理するツール
- 複数のDockerコンテナを管理する
- リソースマネジメントとスケジューリング
- 将来は3rdパーティのツール(Mesos等)もプラグインできるようになる予定



■ Mesos

- データセンタOSというコンセプトのクラスタマネージャ
- リソース管理とスケジューリング
- Dockerも扱える
- もともとはMapReduceのスケジューラ
- 今では一般的なワークロードも対応している
- TwitterやAppleのSiriのバックエンドに採用されている



Container Deep Dive

- Introduction
 - コンテナ
 - Docker
 - コンテナマネージャ
- **OpenStack Summitのコンテナセッション**
- OpenStackのコンテナの取り組み
 - コンテナ対応状況
 - 新しいプロジェクト

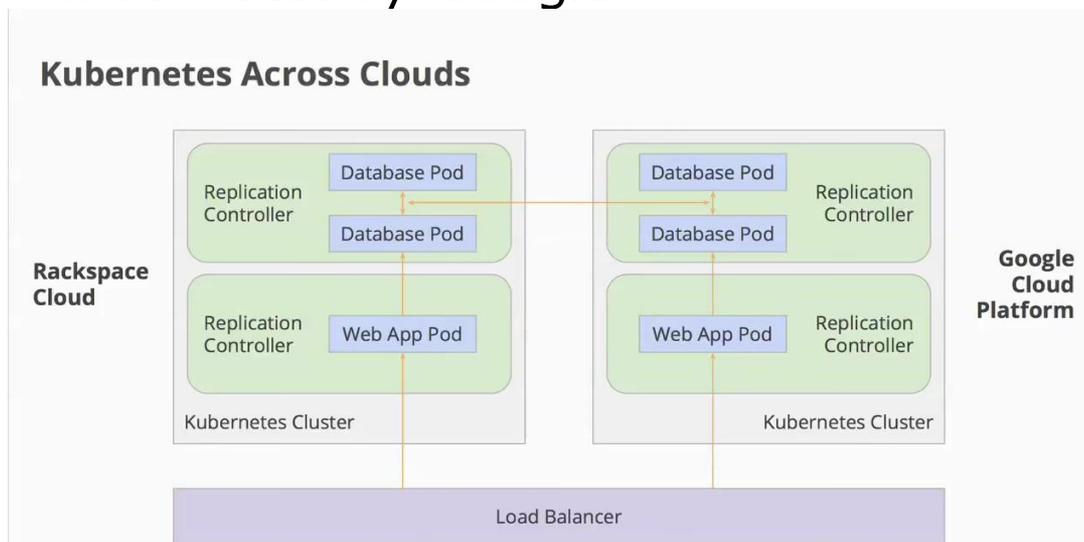


- OpenStack Summitにおけるコンテナの動向
 - Keynote
 - 2日目にコンテナを大々的に取り扱う
 - MagnumやMurano, Kubernetesの関係者が登壇
 - Container Day
 - コンテナ専用のトラックが設けられた
 - 1日中コンテナのセッション
 - MagnumやMuranoのPTLが発表
 - ユーザ事例としてPantheonが発表
 - 一般セッションは13
 - 初心者向けの発表が多い
 - Docker
 - Kubernetes
 - Mesos
 - Dockerエコシステムの紹介

■ Keynote

- OpenStack FoundationのCOO Mark Collier曰く
「NovaはExperimental(実験的)からMature(成熟)している
次のExperimentalテクノロジーはコンテナとDockerであり、
Novaと同じく急成長するだろう」
- デモ
 - Magnum (Container as a Service)
 - Murano (Application Catalog)
 - Kubernetes by Google

5



■ Container Day

- Magnumのセッションは満員でユーザの関心が高い
- セッション

5

- **Welcome to ContainerDay**

コンテナとは？ユースケースなど、コンテナ全般の発表

- **OpenStack Magnum**

MagnumのPTLでKeynodeでも登壇したOtto氏の発表。

- **How a team of four DevOps engineers at Pantheon run 400,000 Drupal/WordPress sites**

10億ものwebsiteを、OpenStackとLXCを用いて提供している

- **Murano bring Docker & Kubernetes to Openstack**

MuranoのPTLのSerg氏の発表

- **OpenStack Containers White Paper Feedback**

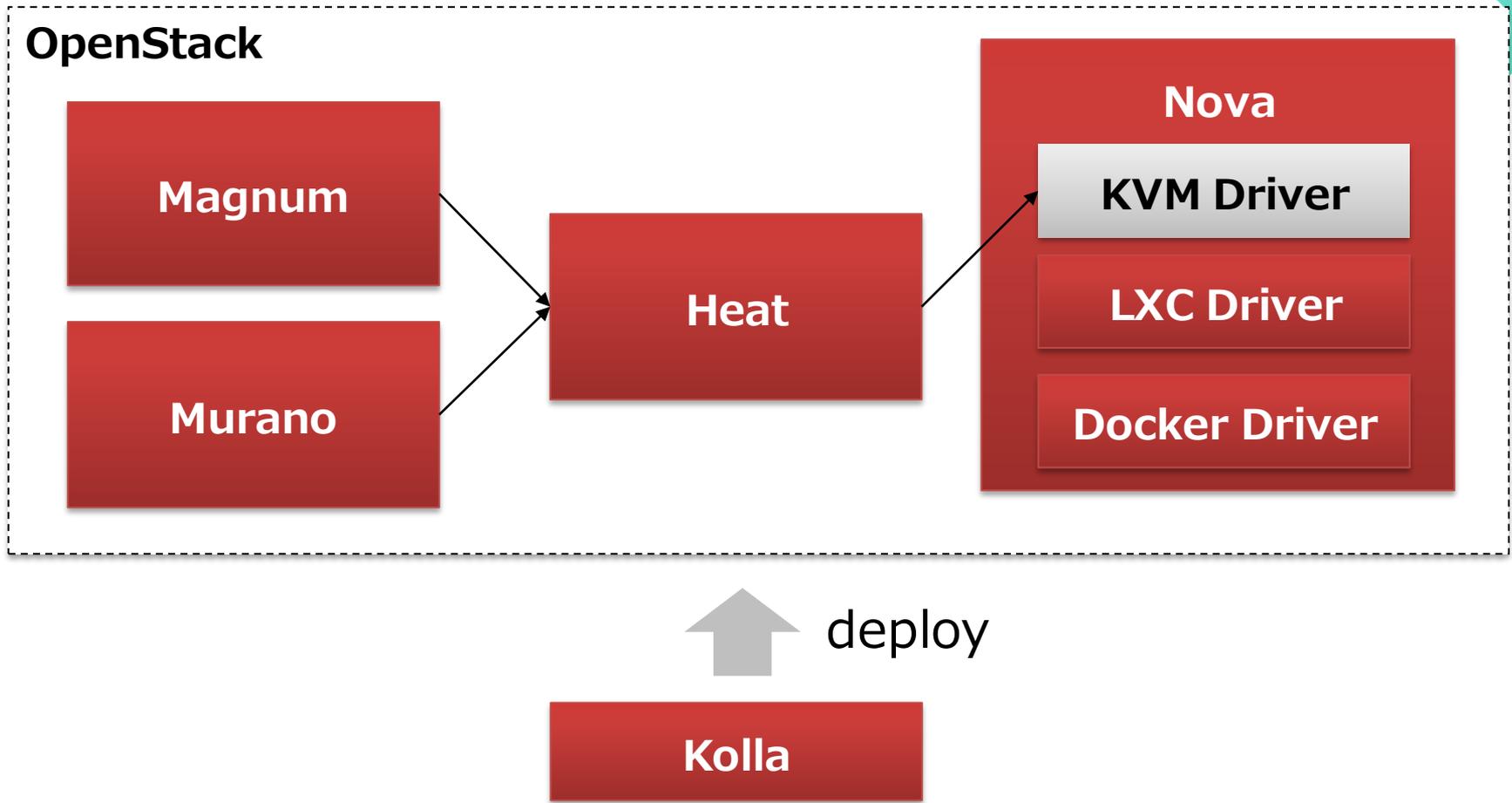
違うセッションに出て聴講せず…

発表資料によると、東京でもContainer Dayが開催される予定

Container Deep Dive

- Introduction
 - コンテナ
 - Docker
 - コンテナマネージャ
- OpenStack Summitのコンテナセッション
- **OpenStackのコンテナの取り組み**
 - **コンテナ対応状況**
 - **新しいプロジェクト**

■ コンテナ関連のプロジェクト



■ LCX virt driver for Nova

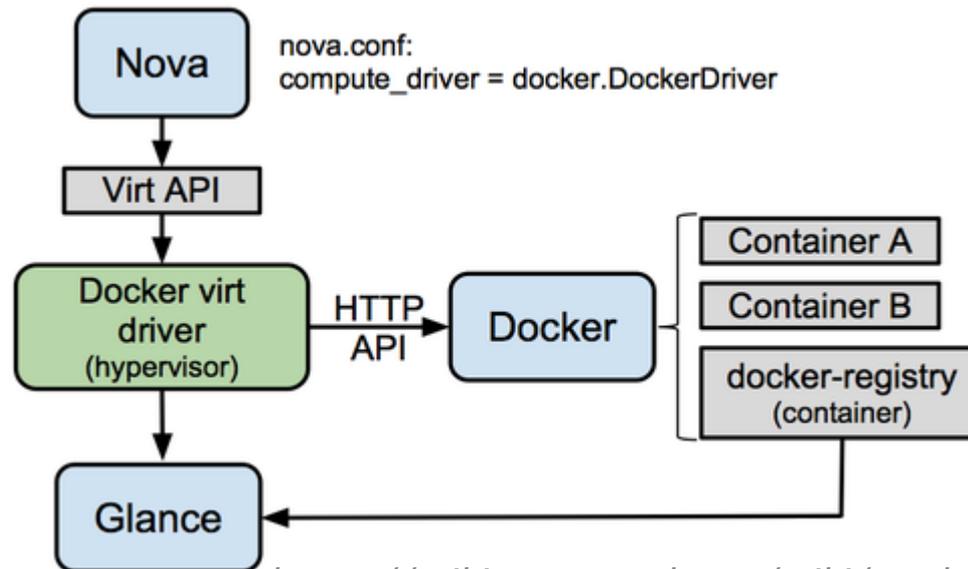
- 現在サポートされている
- いくつかのユーザで実際に利用されている
- Libvirt経由でコンテナを操作している
 - virshコマンドでコンテナの操作が可能
 - SnapShotやshutdownは出来ない



■ Nova Docker driver

- Novaのハイパーバイザドライバの1つとしてDockerを管理
- Havanaから利用可能だが、out-of-treeになった
 - Novaの全てのドライバにCIを要求としている
 - Nova-Docker driverはCIを回していないため削除された
- 再度マージされるように、コミュニティは改善中

5



<https://wiki.openstack.org/wiki/Docker>

■ Heat Docker Plugin (Available from Icehouse)

- Nova Docker Driverはいくつか欠点がある
 - VMとコンテナはライフサイクルが違う
 - Dockerの機能をNova APIで触るのが難しい
- 代替アプローチとしてHeatを使用してDockerを操作する
- Novaとの互換性を持ちつつDockerを管理することができる
- Docker Remote APIをHeat Pluginに実装

```
...
resources:
  my_instance:
    type: OS::Nova::Server
    properties:
      key_name: ewindisch_key
      image: ubuntu-precise
      flavor: m1.large
      user_data: #include https://get.docker.io
  my_docker_container:
    type: DockerInc::Docker::Container
    docker_endpoint: { get_attr: [my_instance, first_address] }
    image: cirros
```

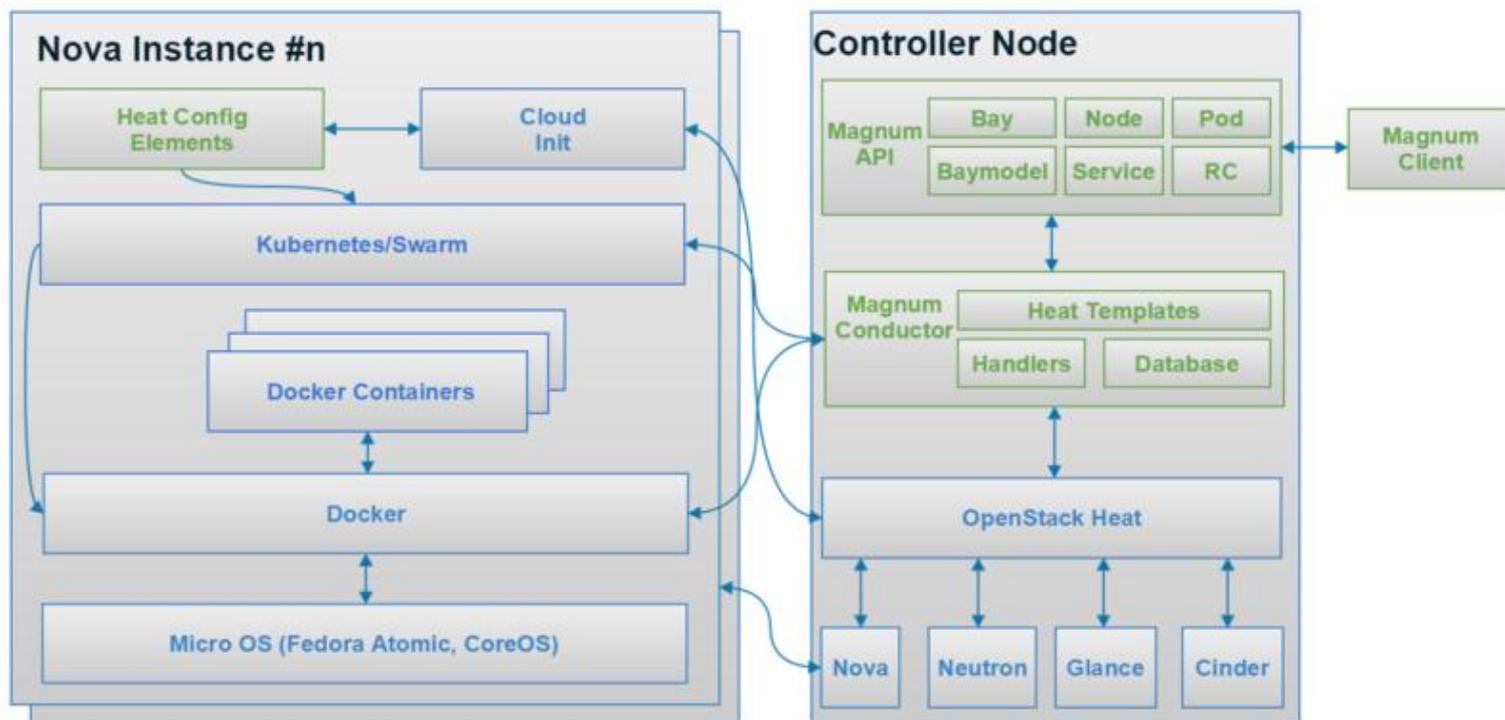
NovaでVMを作成

Novaで作成したVM上に
指定したDockerイメージを
デプロイ

■ Container as a Service

- コンテナのオーケストレーションを提供する**APIサービス**
- 2015/01/20リリース以降、680コミットがあり105,996行に
- 普通のプロジェクトは小さく始まり大きくなるが、Magnumは先に大きなコミュニティができた
- Heatを使ってDockerコンテナのシステムをデプロイする

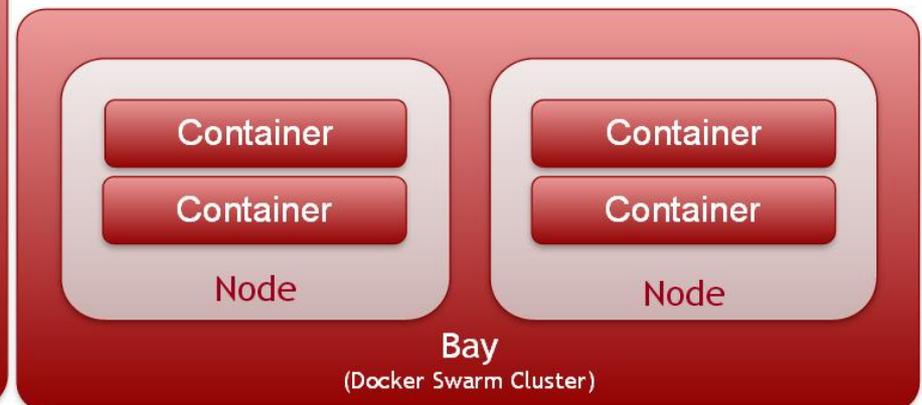
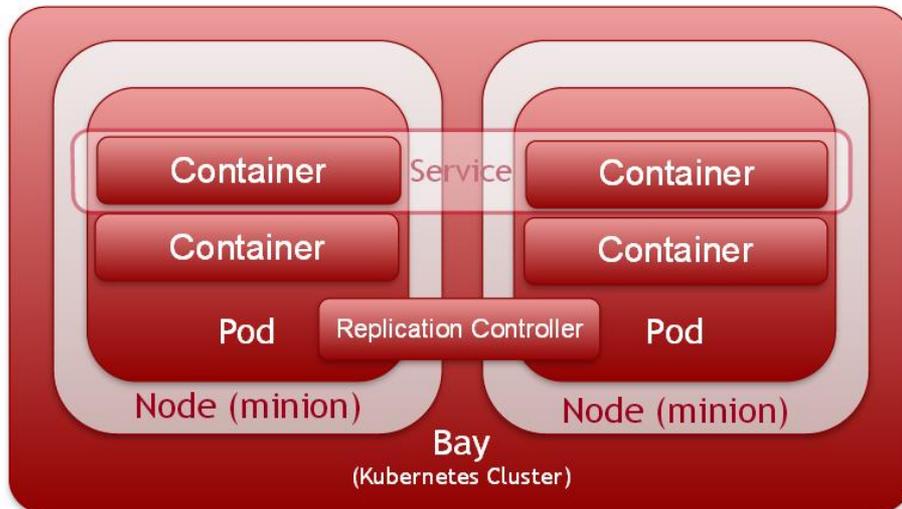
5



■ 7つのオブジェクト

- Container : Dockerコンテナ
- Pod : コンテナの集合 (同じNodeにデプロイされる)
- Node : コンテナを実行する場所 (ベア or VM)
- Service : Podの論理的な集合と、どこにアクセスするかを定義する
- ReplicationController : テンプレートから指定されたPod数のレプリカを維持
- Bay : Nodeの集合 (instanceみたいなもの)
- BayModel : Bayのテンプレート (Flavorみたいなもの)

5



14. Magnum How to Use

1. COE(Container Orchestration Engine)を作成する

```
# magnum baymodel-create --name k8sbaymodel --image-id fedora-21-atomic-3 ...
```

2. Bayを作成する

```
# magnum bay-create --name k8s --baymodel k8sbaymodel --node-count 3
```

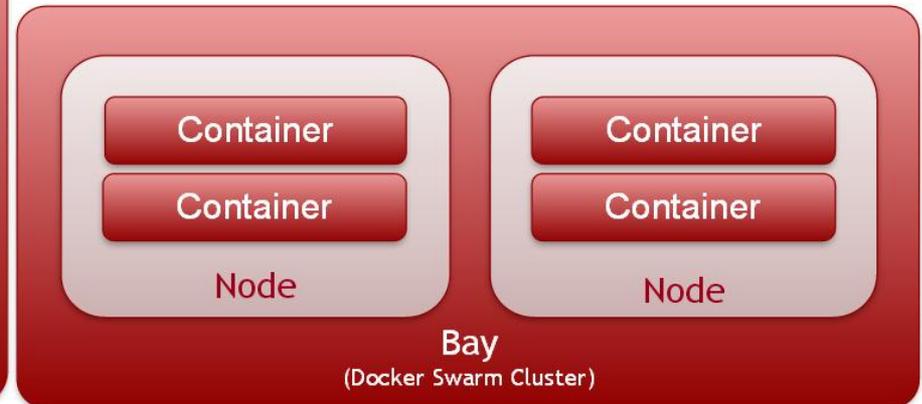
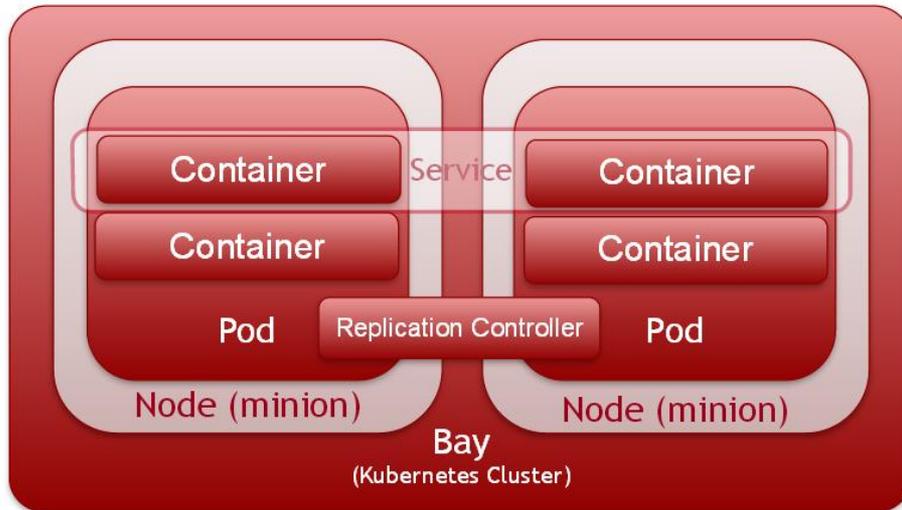
3. Podを作成する

```
# magnum pod-create --manifest ./redis-master.yaml --bay k8sbay
```

Docker Swarmならコンテナを作成する

```
# magnum container-create --name testcontainer --image cirros
```

5



■ 実体はHeatを呼び出すスクリプト郡

- Magnumコマンドの引数をHeat APIに変換
- Bayを作成するときにHeatのスタックとVMが作成される



```
#vi /usr/lib/python2.7/site-packages/magnum/templates/heat-kubernetes/kubenode.yaml
heat_template_version: 2013-05-23
```

```
description: >
```

```
This is a nested stack that defines a single Kubernetes minion,
based on a vanilla Fedora 20 cloud image. This stack is included by
a ResourceGroup resource in the parent template (kubecoluster.yaml).
```

```
...
```

```
server_image:
```

```
  type: string
```

```
  default: centos-7-atomic-20150120
```

```
  description: glance image used to boot the server
```

```
...
```

```
configure_kubernetes_minion:
```

```
  type: "OS::Heat::SoftwareConfig"
```

```
  properties:
```

```
    group: ungrouped
```

```
    config: {get_file: fragments/configure-kubernetes-minion.sh}
```

```
kube_user:
```

```
  type: "OS::Heat::SoftwareConfig"
```

```
  properties:
```

```
    group: ungrouped
```

```
    config: {get_file: fragments/kube-user.yaml}
```

■ Application Catalog for OpenStack

- 様々なアプリケーションを簡単な手順で利用できる
- 3つの提供形態がある
 - アプリケーションのパッケージ
 - Heat Templates
 - Glance Images
- アプリケーションのパッケージとしてKubernetes環境を提供

5

openstack | Community App Catalog ^[beta] FAQ

MURANO PACKAGES HEAT TEMPLATES GLANCE IMAGES ADD NEW CONTENT

The OpenStack Application Catalog will help you make applications available on your cloud.

RECENTLY ADDED APPS (Requires OpenStack)

- Kubernetes
- CLOUD FOUNDRY Cloud Foundry
- ORACLE DATABASE 12^c Oracle
- debian Debian
- OPENSIFT OpenShift

Don't have OpenStack yet?
Visit the OpenStack Marketplace and select tools and methods for building your cloud. Then come back to the OpenStack Application Catalog to make your cloud sing with applications.

Murano Packages Heat Templates Glance Images

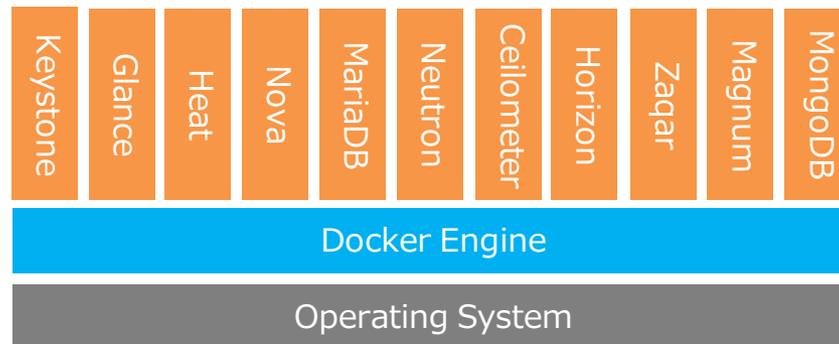
Name	Type	Status	Last operation
			Drop component here to add it
KubernetesContainersCluster	KUBERNETES	Ready to deploy	Component draft created
webservier	Kubernetes POD	Ready to deploy	Component draft created

Deleting 3 items

■ Dockerを利用してOpenStackサービスを展開する

- Triple Oの一部として開発している
- デプロイの簡略化、容易化
 - アップデート、ロールバック
 - OpenStack環境の開発を想定
 - 以前はKubernetesを利用していたが、現在Heat Templateを利用している
- まだまだProductionではない
 - DockerやKolla自体まだ新しい技術
 - 通常のOpenStackとデプロイモデルが違うので、新しいバグがあるかもしれない

5



■ コンテナはOpenStackでも注目の技術

- サミットでも、コンテナのセッションは多くの人に参加
- 複数のプロジェクトがDockerを扱おうとしている
 - 混沌としている
 - PaaSとの違いは？
- 今はVM上のDockerを利用しているが、ベア上のDockerも考えてる
 - アーキテクチャが変わっても既存のDockerイメージは利用可能



■ ぜひ試してみてください！

OpenStack Summit **Tokyo**でお会いしましょう！

- Linuxは、Linus Torvalds氏の日本およびその他の国における登録商標または商標です。
- OpenStack, Docker, Kubernetes, Google, Rackspace Cloud, その他、記載の商標やロゴは、各社の商標または登録商標です



HITACHI
Inspire the Next