

ガチで聞く！ ヤフーのOPENSTACK プライベート・クラウドの実態とは

伊藤 拓矢

ヤフー株式会社

サイトオペレーション本部

インフラ技術一部 部長



本日のお話

【2 0 1 3】 OpenStack はじめの一步

- ✓ 導入までの経緯
- ✓ OpenStack 基盤の構成で重要なこと

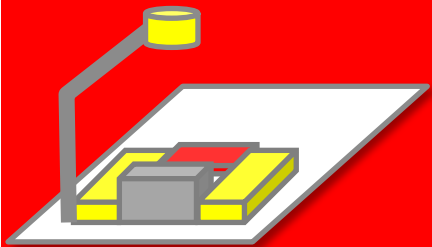
【2 0 1 4】 OpenStack との固い絆

- ✓ 現在の稼働状況
- ✓ 設計・構築・運用のノウハウ
- ✓ ブロケードプラグインによるパフォーマンス向上

【2 0 1 5】 OpenStack 未来とともに

- ✓ 仮想からベアメタルまで
- ✓ アメリカへの進出





タイムマシン

本日のお話

【2 0 1 3】 OpenStack はじめの一步

- ✓ 導入までの経緯
- ✓ OpenStack 基盤の構成で重要なこと

【2 0 1 4】 OpenStack との固い絆

- ✓ 現在の稼働状況
- ✓ 設計・構築・運用のノウハウ
- ✓ ブロケードプラグインによるパフォーマンス向上

【2 0 1 5】 OpenStack 未来とともに

- ✓ 仮想からベアメタルまで
- ✓ アメリカへの進出

導入までの経緯

- In-House IaaS 全て自前での開発
 - ✓ WEBからの操作が前提
 - ✓ 新規開発に手が回らず、運用で手一杯
 - ✓ APIが独自インタフェース、原則非公開
 - ✓ ライフサイクルが回らない



• プロジェクトチームの発足

【決めたこと】

- ✓ APIは公開する
- ✓ APIインタフェースは変更しない
- ✓ セキュリティの担保
- ✓ 前に進む

OpenStackの採用理由

1. 標準化されたAPIを使用する事が可能

- ✓ 取り巻くOSSを活用することが可能
- ✓ アプリケーションに近いところに人的リソースを割ける

2. ベンダーを意識しないリソース活用

- ✓ ハードウェアの種類やベンダーを意識せず、リソースのコントロールが可能
- ✓ ハードウェアの抽象化の実現

リリース

1. 検証開始から半年で開発環境リリース
 - ✓ APIによるデータセンタ操作
 - ✓ 数十秒、数百インスタンスの稼働
2. 3か月後にプロダクションリリース
 - ✓ 数日後にはサービスが稼働し始める
3. 社内独自開発に注力する事が可能になった



OpenStack 導入後

1. 社内独自機能のみの開発
2. ベンダーと連携強化、OpenStackを盛り上げる
3. 標準APIによる活発なOSS活用



ベンダーとの共創

- ベンダー製品の活用
 - ✓ パフォーマンス要件
 - ✓ 運用要件、運用コスト
 - ✓ 開発、改修のためのエンジニアコスト
- 共創することで

Yahoo! JAPAN

共創

BROCADE[®]



スケーラブルなマルチテナント環境構築のために

ルータ

L2 ネットワーク

ファイアウォール

ロードバランサ



マルチテナント OpenStack
ネットワークの抽象化

BROCADE[®]

Neutron

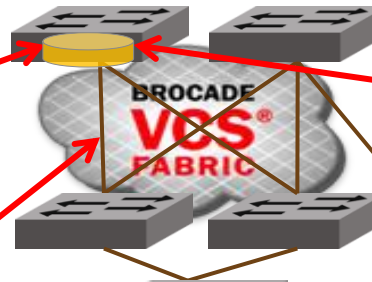
Brocade VDX Plug-in

Virtual Routing Interface
(SVI/VE)

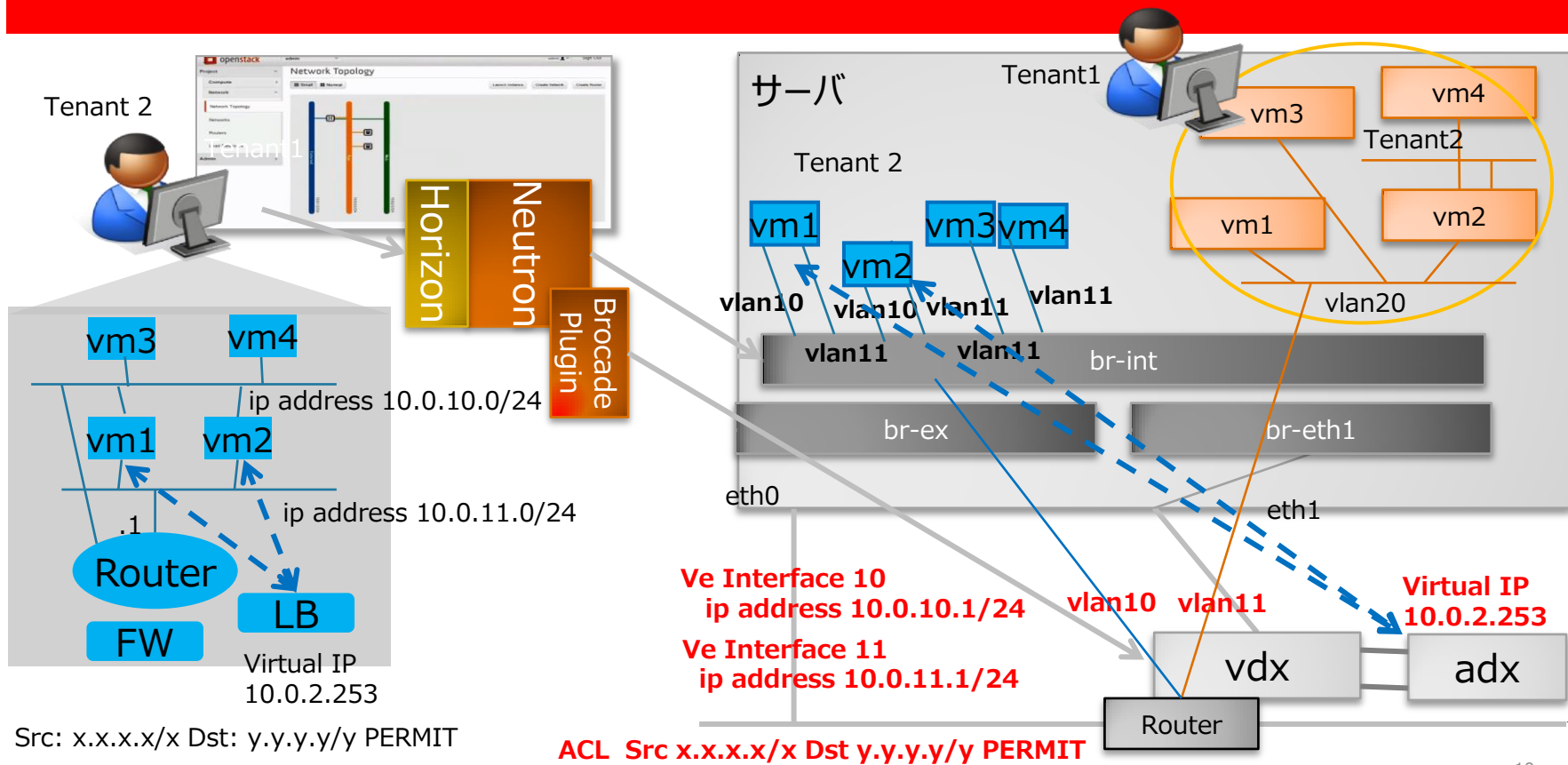
Subnet management
(VLAN)

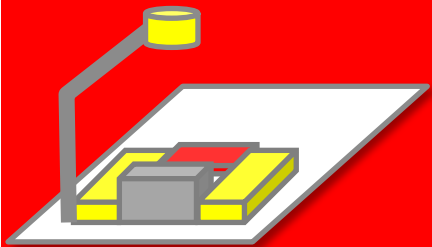
Gateway ACL
(FWaaS)

Load Balancing
(LBaaS)



ユーザー自身が構築できる仮想ネットワーク





タイムマシン

本日のお話

【2 0 1 3】 OpenStack はじめの一步

- ✓ 導入までの経緯
- ✓ OpenStack 基盤の構成で重要なこと

【2 0 1 4】 OpenStack との固い絆

- ✓ 現在の稼働状況
- ✓ 設計・構築・運用のノウハウ
- ✓ ブロケードプラグインによるパフォーマンス向上

【2 0 1 5】 OpenStack 未来とともに

- ✓ 仮想からベアメタルまで
- ✓ アメリカへの進出

現在の稼働状況



YAHOO! JAPAN ブログ



Sportsnavi



YAHOO! JAPAN ブックストア



ヤフオク!



YAHOO! JAPAN ファイナンス



GYAO!

ほぼ全てのサービスの基盤として利用が進む

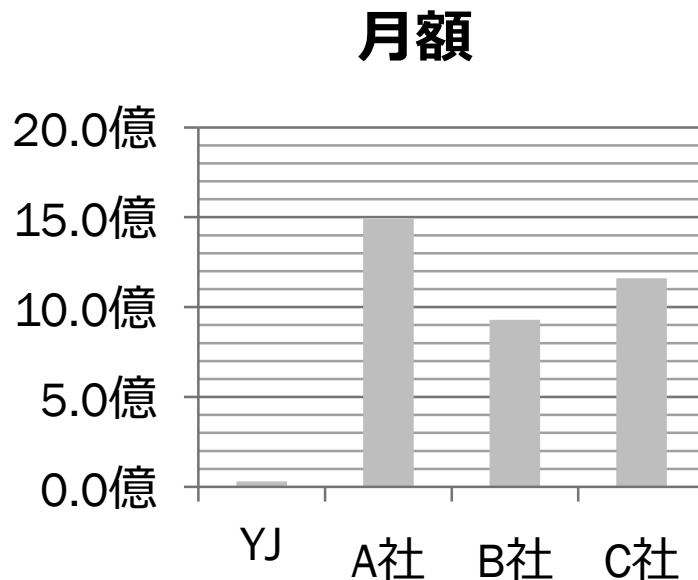


現在のサービス提供状況

- 50000インスタンス
- 99.996%の稼働率
- 6倍のトラフィック密度（物理環境との比較）
- 10PBのデータ
- 開発者6名 運用者4名
- 20クラスタ稼働中

現在の稼働状況

- 2000人の利用者に対して平等に提供
 - ✓ 問題発生時はOpenStackサポート窓口がすぐに対応
 - ✓ 最近是非エンジニアにも活用が始まる
- 1日250インスタンスが稼働する
 - ✓ 1か月以内に半分が消える
 - ✓ 人が介入せずに稼働
- 低コスト
 - ✓ 最近の構成ではパブリッククラウド比月額で97%カットを実現



現在の運用状況

- OpenStackによってデータセンタリソースのほぼ全てがセルフサービスで提供
 - ✓ インフラ側はそっと見守る
 - ✓ 開発者の意思でサービスの構成が可能
- 開発者は即時利用が可能
 - ✓ 物理環境のような待ちが発生しない

設計・構築・運用のノウハウ

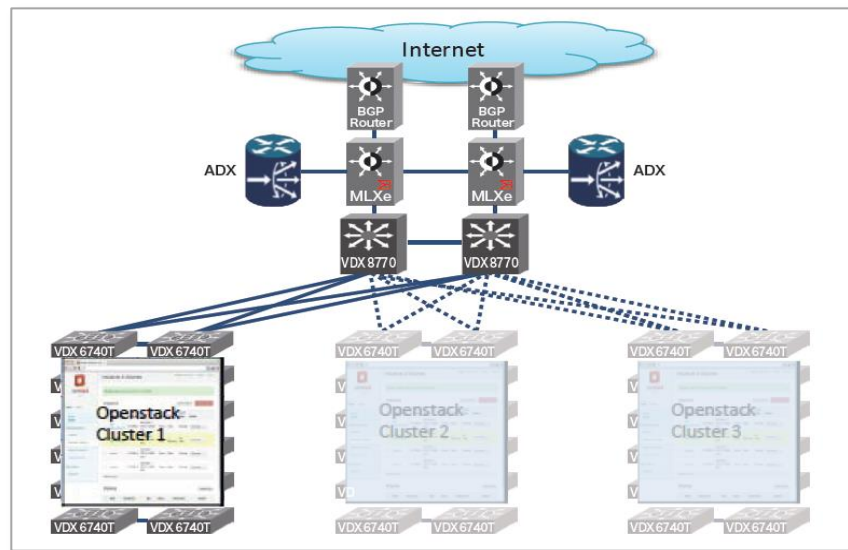
- OpenStackはコントローラとして使う
 - ✓ I/OはOpenStack障害時でも処理し続けること
- OpenStackと連携する製品は独自DBを持たない
 - ✓ 整合性の問題が必ず出る
- チューニング & チューニング
 - ✓ DBとRPCを安定稼働させること



ネットワーク設計・構築・運用のノウハウ

Neutron の利用

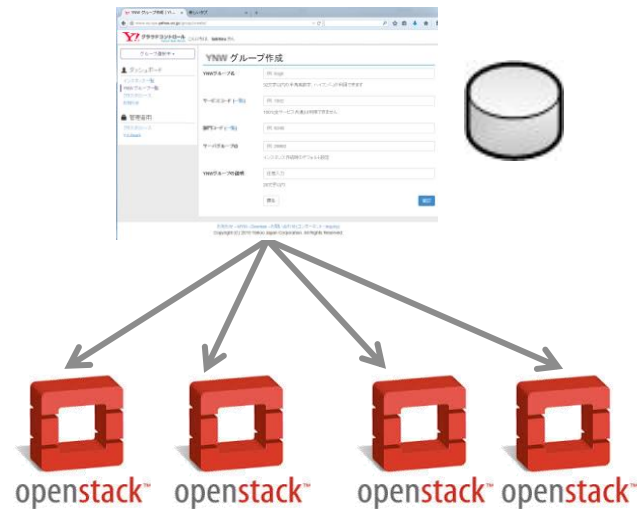
1. マルチプロジェクトの運用
2. アドレスの管理
3. Brocadeプラグインによるパフォーマンス向上



マルチプロジェクトの運用

- 現在2000を超えるプロジェクトを管理

1. 個人所有も多い
2. ユーザは自由に作成可能
3. 統合管理システムでプロジェクトを管理
4. システムからプロジェクトが共有される
5. クラスタによってアドレス管理が異なる



アドレスの管理

- アドレス割り当ては主に2種類

1. 全てのプロジェクトでシェアしている標準サブネット

- ✓ 標準サブネットはクラスタ辺り /23を32個

2. 特定のプロジェクト専用の拡張サブネット

- ✓ 拡張サブネットはユーザの希望次第で/28～/24を割り当て
- ✓ グローバルIPアドレス、もしくはプライベートIPアドレスが割り当てられる



Neutron実装の強化(appendix)

- 今回の検証の構成

構成	Kernel	Open vSwitch	ML2 mechanism driver
Open vSwitch + VDX + Kernel3	3.14.17	2.3	openvswitch, brocade
Open vSwitch + VDX	2.6.32	2.3	openvswitch, brocade
Open vSwitch + gre tunnel	2.6.32	2.3	openvswitch
LinuxBridge	2.6.32	-	linuxbridge

- これらの環境の50VMに対してAvalancheから負荷を掛けて計測
- 応答速度はサーバからのレスポンスで最初の1バイトが返ってきた時の経過時間



Neutron実装の強化まとめ

- L2サービスの抽象化に利用が可能
 - ✓ スループット、安定性を重視した構成が構築可能
 - ✓ GREを利用する必要無し
- Brocade VDX2台から始められる構成
 - ✓ 安価にスタート
 - ✓ 40HV程度の構成が可能
 - 多くの企業には十分なサイズ



まとめ

まだ踏み切れないあなたに・・・

1. OpenStack ディストリビューション も選択肢が広がり、環境は整いました。今日から始めてみませんか？
2. OSSでは運用が難しい環境は、ノウハウが蓄積された商用製品の活用も選択肢に。
3. 2 台から始める OpenStack。とにかく手軽に簡単構築！

