

Yahoo! JAPAN 仮想化への取り組み ～ OpenStack with LBaaS ～

2014/1/24

<コンタクト先>

ヤフー株式会社 宮田耕一

komiyata@yahoo-corp.jp

A large, light pink graphic of the Japanese character 'Y!' (Yahoo!) is centered on the page. The character is stylized with a thick, rounded font. The exclamation point is also thick and has a small registered trademark symbol (®) at its base.

自己紹介

氏名	宮田 耕一(みやた こういち)
所属部署	システム統括本部 基盤システム開発本部 インフラ技術2部 ネットワーク1
社歴	7年目
担当業務	Yahoo! JAPANのNW運用・構築 メインはLB/GSLB周り

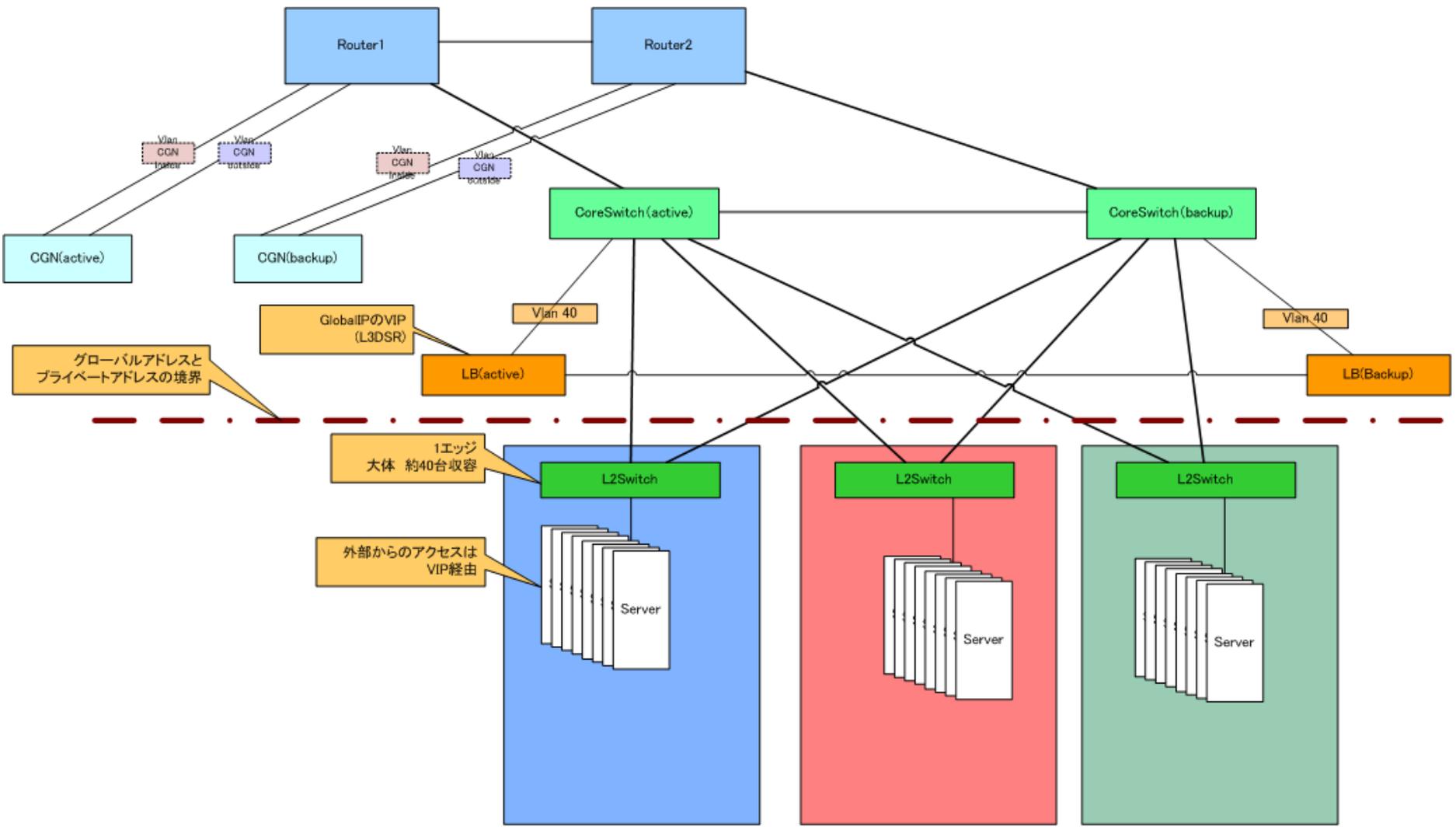
- **Yahoo! JAPANのNW構成**
- **仮想化するまで(OpenStackを使って)**
- **LBaaSを導入してどうなったか**
 - **Brocadeプラグインでのメリットなど**
- **問題点/課題など**



Yahoo! JAPANのNW構成



ネットワーク構成



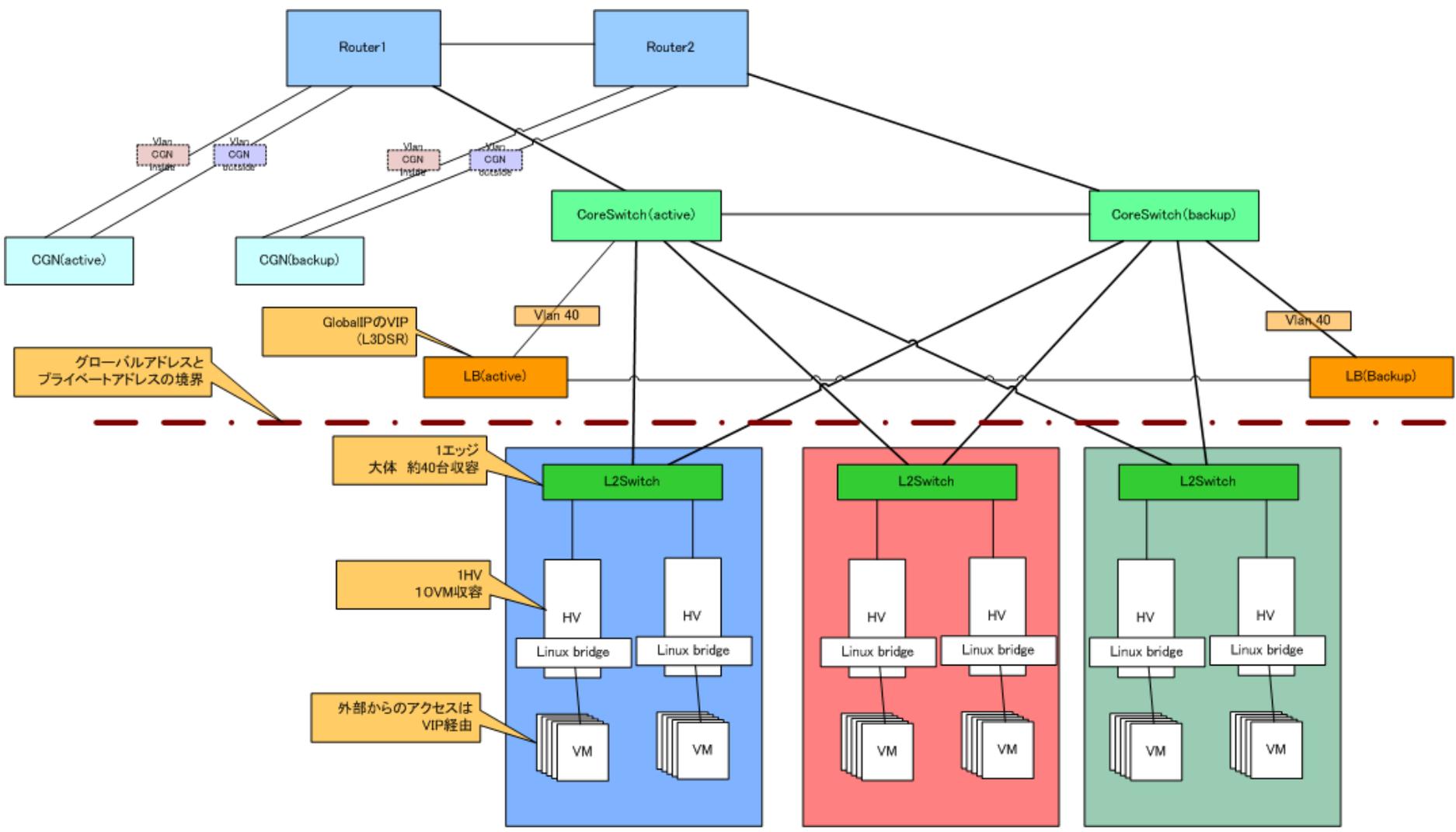
- サーバはプライベートIP
- 外への通信はCGNでNATする。
- 外部から通信がある場合は、L3DSRのVIP経由で行う。
- 内層トラフィックはCGNを通さない

- **コスト削減！**
 - サーバ稼働率の上昇
- **ラックスペースの節約**
 - サーバはどんどん入れたいけど、置き場がない。
- **規模が大きくなってきて運用がしんどい**
 - OSのインストールなどを自分たちでコントロールするのはしんどい。

- **物理面で**
 - サーバの購入
 - 2週間～1か月(納期の問題でもっと伸びることも。。。)
 - VIPの設定
 - 1～3日(作業状況によっては伸びるかも。。。)
- **SW面で**
 - サーバのセットアップ
 - 動作確認とかとか

**ボトルネックといったら言い過ぎかも知だが
物理作業で止まる瞬間がある。**

- **検証で問題がなかった。**
 - インストールには苦勞したが・・・
- **CloudStack・ユーカリなども調べていたけど**
 - CloudStackは機能追加が大変そうだった。
 - ユーカリはもうほとんど収束気味
- **機能追加が簡単そうだった。**
 - 言語がpythonだったから



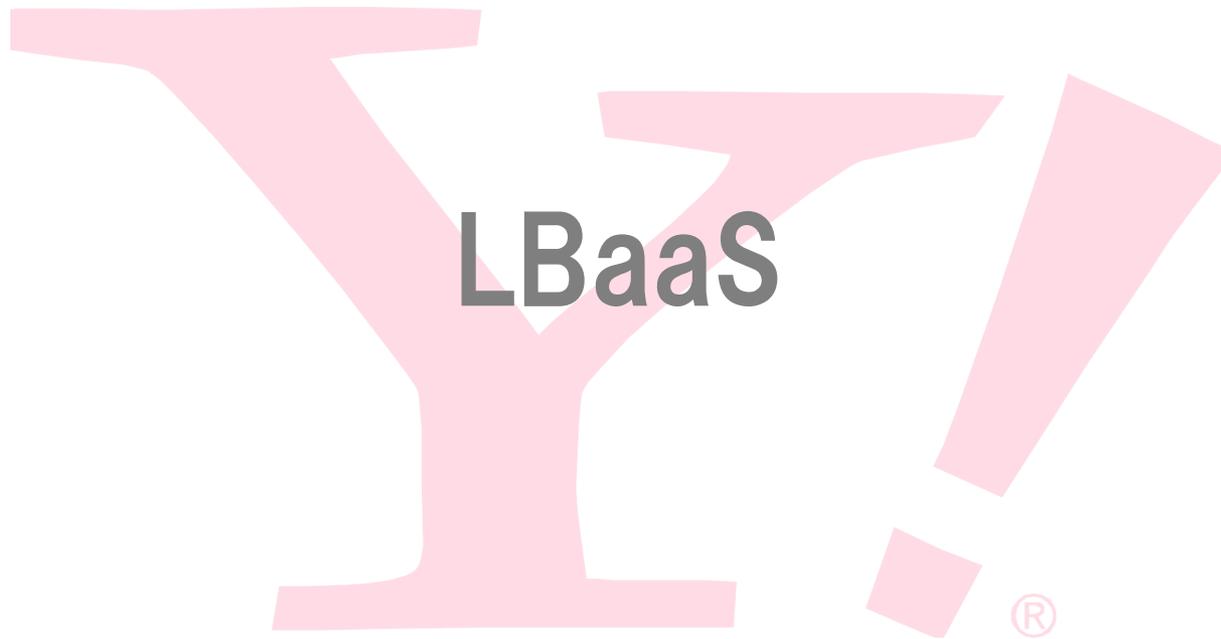
- リリースしてみると問題がいろいろと
 - qpidが詰まる
 - インスタンス作成が失敗
 - トラフィック詰まった
 - VMの作成が遅い

**Openstack全体のチューニングをいろいろやった結果
今は快適！**

サーバの導入までは快適になったけど・・・

サービスを提供するにはVIPとかも作るわけで。。。

- **OpenStack上でVIPまで設定できれば素敵じゃないか！**



LBaaS

Load Balancing as a Service:LBaaS

- Grizzlyからマージされた Quantum (neutron) で使える負荷分散機能
- HAProxyをつかった構成

以下の点でYahoo! JAPANとはマッチしなかった

- シングルテナントに向かない構成だった。
- SWロードバランサの実績がない
 - チューニングや監視が大変なイメージ
- パフォーマンスの心配
- 冗長構成が取れない
- L3DSRを使いたい
 - 物理環境をそのまま持っていきたかった。

これらを解決するべくBrocade様にPluginの開発を依頼

- **BrocadeのADXに対して命令を投げるプラグイン**
 - Brocade ADXのXML APIに対して命令を投げるもの
 - 裏を返せばコードをいじれば、XML APIでできるものは何でもできる？
- **Inline , L2DSR , L3DSRにも対応**
 - 利用形態にそって選択できます。
- **HAにも対応**

- **仮想化の要望**

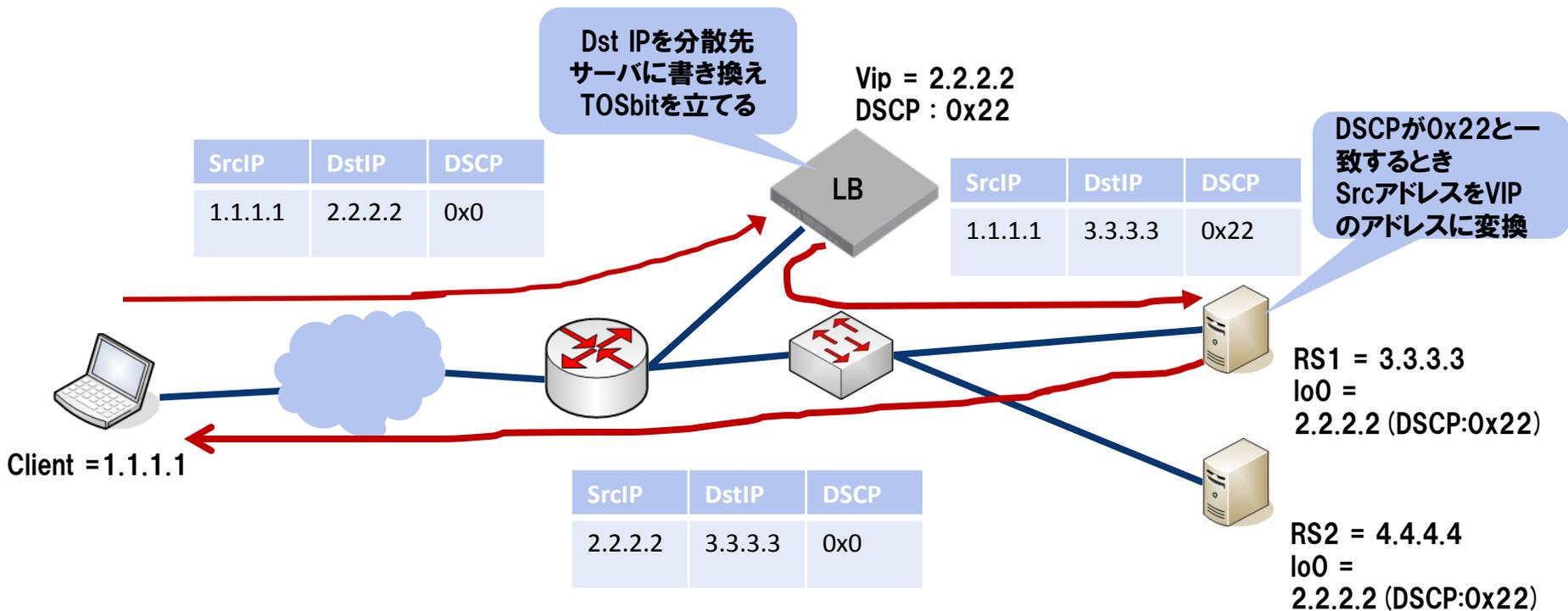
どんどんインスタンス作りたい

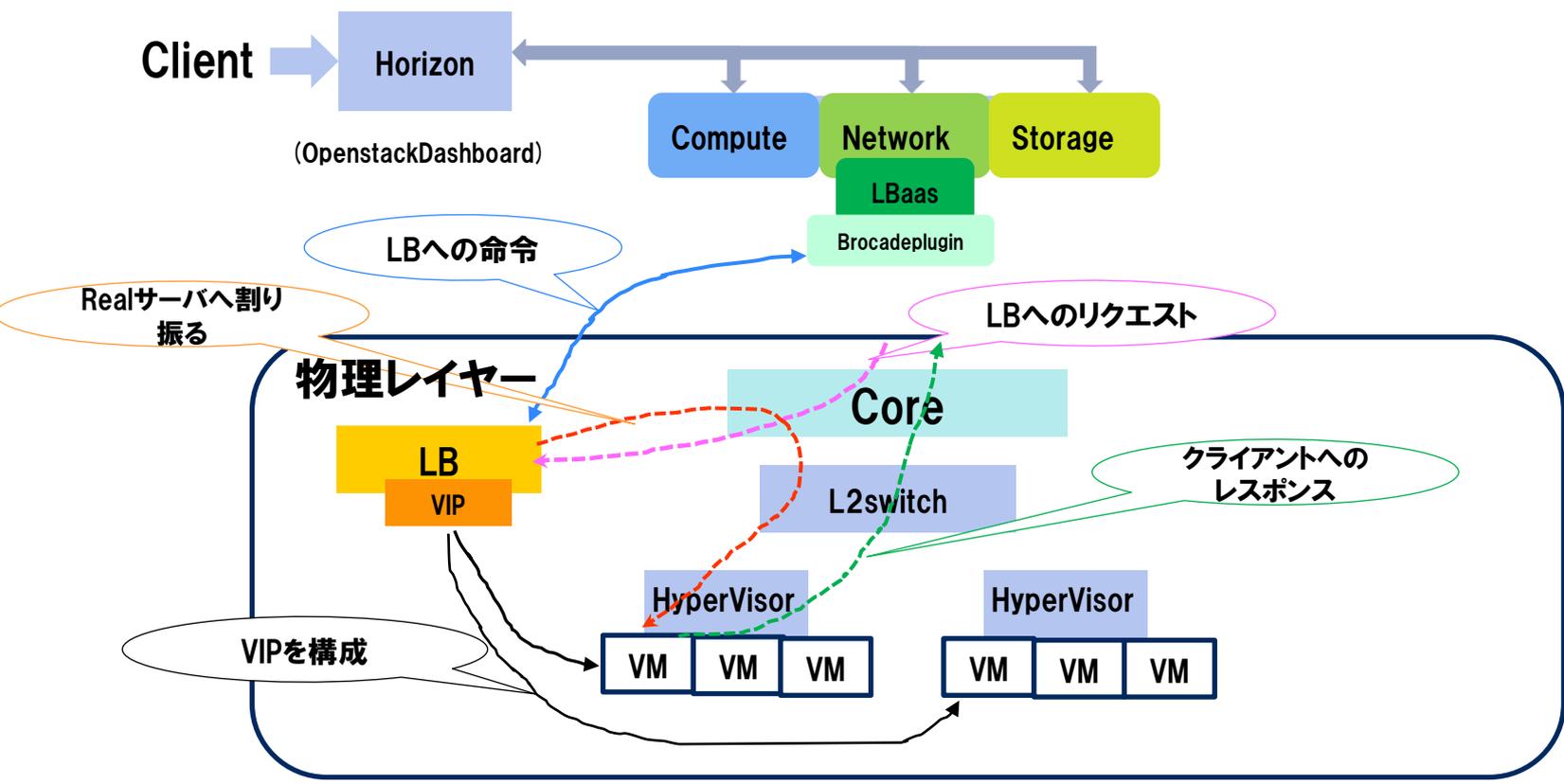
- NWがどんどん増える。
- IPもどんどん使う。アドレス足りない
 - グローバルIPを使っていたらきりが無い

L3DSRにすれば

- NWにとらわれない負荷分散ができる。
- プライベートアドレスを使ってしまうと、グローバルIPの節約にもなる
- LBの負荷が減るので、アプライアンスの台数も抑えられる

VIPとTOSbit (DSCP) をマッピングさせて行う L3ベースのDSR





• メリット

- 設定する手間がなくなって楽に
- 構築スピードが速い
 - NWが増えてもLB側では特に設定が必要ない
- 監視も筐体が1つだから楽
 - パフォーマンスなどに注力するので CPUやIFなど基本的なところを
- アプライアンスだから高パフォーマンス
 - まだ利用度はそんなに高くないがめっちゃめっちゃ余裕
 - 約600VM収容

• 困ったこと

- 重複チェックがなかった。(OpensStack側に)
 - 複数のsystemを同じ筐体で管理すると重複がおこってしまって、
 - UIは独自でつくるか、Horizonや APIにちょっと手を入れれば大丈夫



- **Openstackの導入で**
 - **リリースまでの期間の短縮**
 - **稼働効率のアップ**
 - **省スペース**

- **LBaaSPlugin導入で(Yahoo! JAPANの構成では)**
 - **サービスインまでの時間短縮**
 - **管理の簡略化**
 - **高パフォーマンス**
 - **Global IPも節約**
 - **LB自体の稼働効率Upも図れる**

A large, light pink graphic of the Japanese character 'Y!' (ヤイ) is centered on the page. The character is stylized with a thick, rounded font. The exclamation point is also pink and has a registered trademark symbol (®) at its base.

問題点・課題

- **Openstackのコミュニティー**
 - どうなっていくか未知数
- **LinuxBridgeのFlooding**
 - Unknown unicastパケットが大量に。。。
- **Openstackシステム間での一意性**
 - 独自で開発はやっぱり必要
- **Qemuとの相性**
 - バージョンによってはL3DSRが動かない。
- **トラフィックの予測がつかない。**
 - Qosで絞る？ リザベーションする？

A large, light pink, stylized logo consisting of a 'Y' shape followed by an exclamation point. The 'Y' has a wide top bar and a narrow stem. The exclamation point has a triangular top and a short stem. A small registered trademark symbol (®) is located at the bottom right of the exclamation point stem.

ご清聴ありがとうございました

- **Openstackのコミュニティー**
 - JP : <http://openstack.jp/>
 - Org : <http://www.openstack.org/>
 - LBaaS : <https://wiki.openstack.org/wiki/Neutron/LBaaS>
 - Git : <https://github.com/openstack>
- **L3DSR**
 - GitHub : <https://github.com/yahoo/l3dsr>