BROCADE

ENOG19 Meeting

2013-02-22

Load balancerと俺

ブロケードコミュニケーションズ
システムズ株式会社

甲野 謙一

私は誰

世界を船で旅していたころ
@シンガポール

甲野 謙一(36) as KonoKono

- 2000年から様々の形でLoad Balancerを担当
- BrocadeではADXの機能設計やPOST系を担当

ブロケード

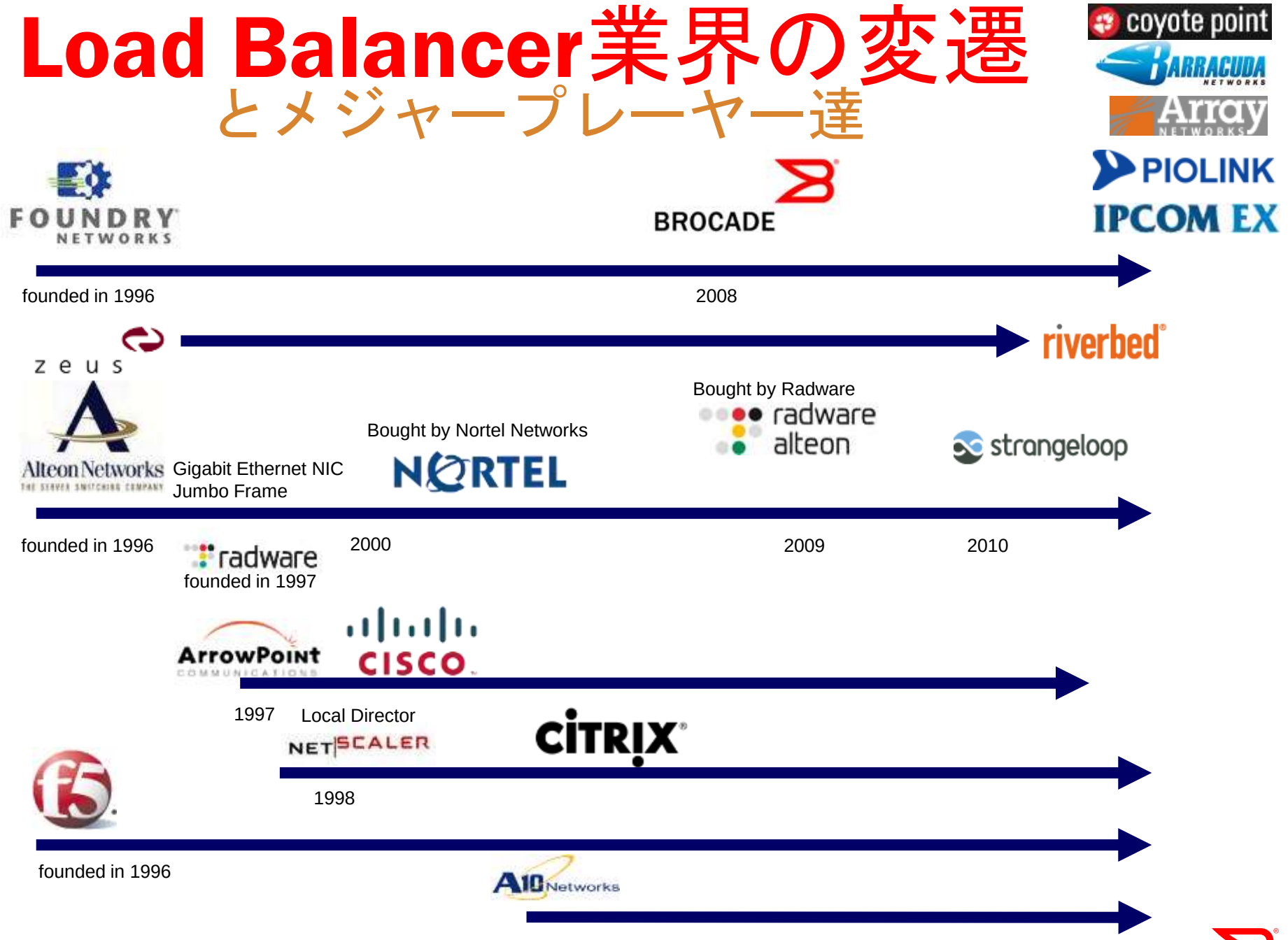
- 事業
  - L2/L3とL4-7スイッチ
  - Fiber Channel SANスイッチ

今は、SANもIP/LANも両方やってます

今日お話しすること

- 商用LB業界を振り返る
- Load Balancer基礎
- L2DSRとL3DSR構成 ←メインTOPIC
- Brocade LAB環境の紹介

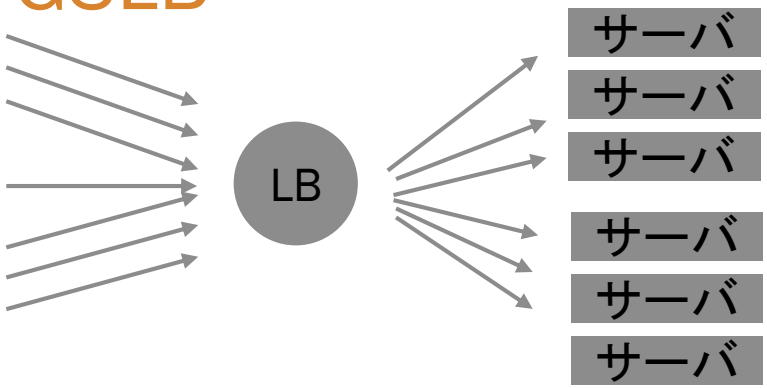
Load Balancer業界の変遷
とメジャープレーヤー達



Load Balancingとは

復習

SLBとGSLB



LB技術により、集中と分散をコントロールし、運用を最適化

SLB

- 特定の宛先IPに集中したTrafficを分散させる技術。

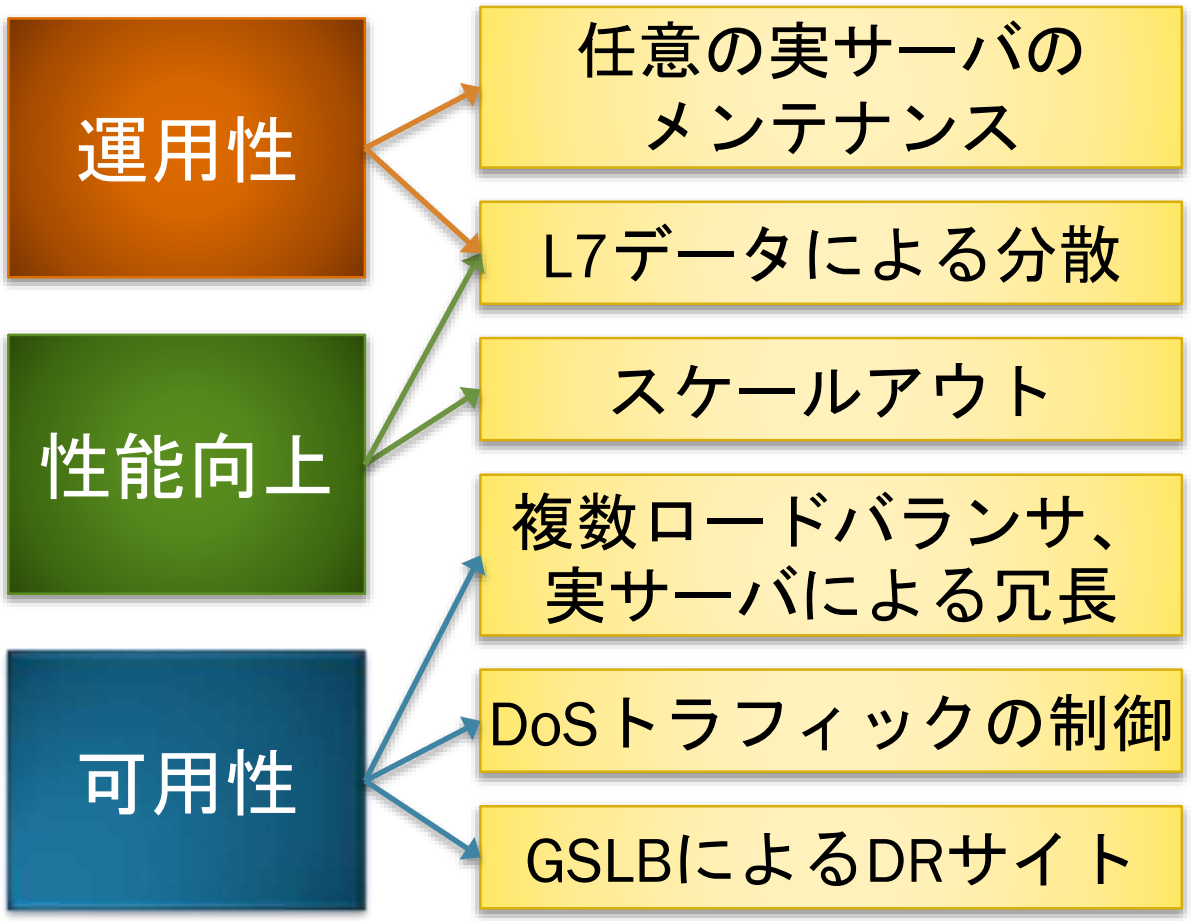
GSLB

- Trafficが特定の宛先IPに集中する前に、Trafficを分散させる技術。LB業界では、DNSによる分散が一般的。

復習

Load balancer

よくある導入の目的

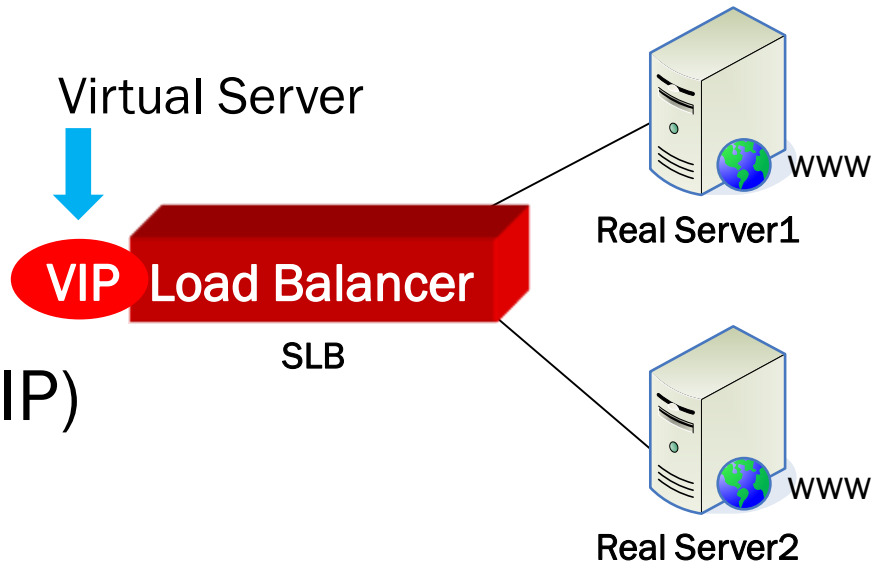


Load Balancerとは

復習

4つの基本機能

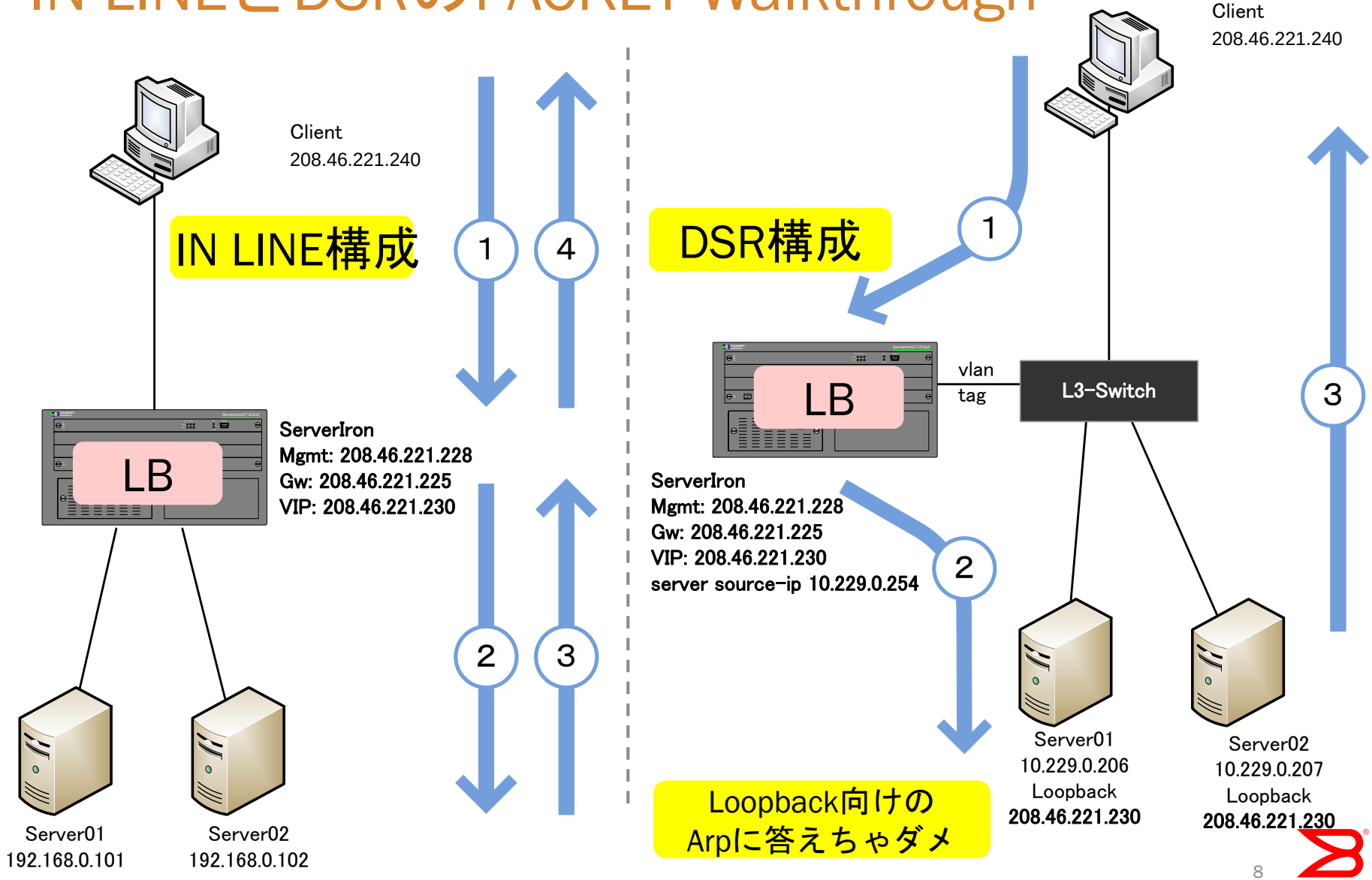
1. 負荷分散機能
2. ヘルスチェック機能
3. アドレス変換機能(MACとIP)
4. セッション維持機能



一般的なLB構成

復習

IN LINEとDSRのPACKET Walkthrough



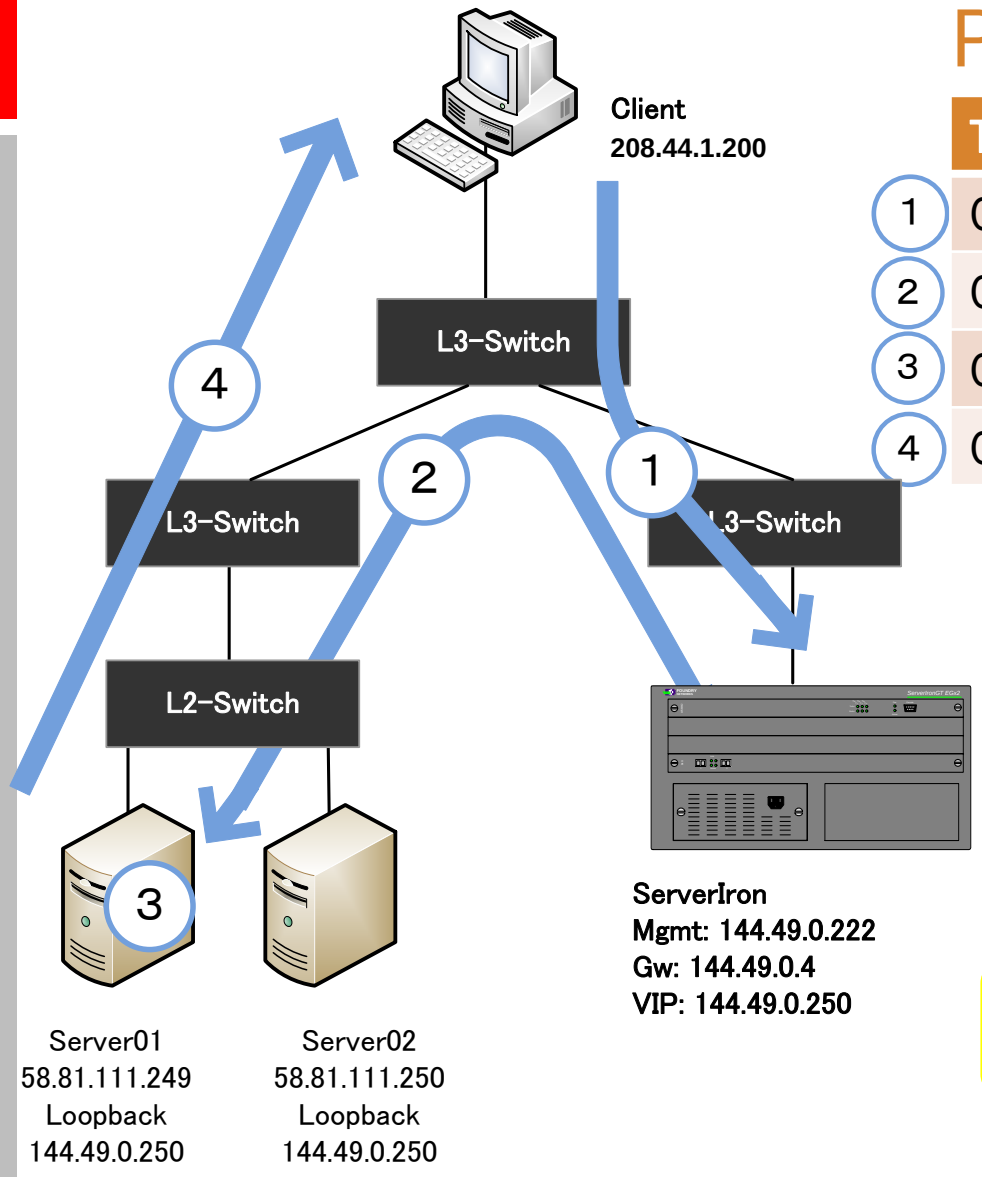
なぜいまL3DSRか

検討項目	L2DSR	L3DSR
パフォーマンス(スループット)	Outbound trafficはLBを経由しない。	←同じ
クライアントのSource IPを変更	SNATしないので、SIPはそのまま。	←同じ
LBが全滅した時	サーバ自身がサービスを継続可能	←同じ
L3をまたいだReal Server設置	×	可能！

- 構成が柔軟なため、インターネット系のCloudサービスにうまくはまる。
- Anti Load Balancerの人(会社)であっても、DSRであれば、柔軟に構成できるため、In-LINEと比較して導入が容易。
- A10、Brocade、CitrixがL3DSR構成を正式サポート。F5でもできるはず。

L3DSR構成の話

ここ3年の話



PACKET Walkthrough

	TOS	Source IP	Destination IP
1	0x0	204.44.1.200	144.49.0.250
2	0x7	204.44.1.200	58.81.111.249
3	0x7	204.44.1.200	58.81.111.249
4	0x0	144.49.0.250	204.44.1.200

Serverが宛先IPを144.49.0.250に変更する点がポイント

FreeBSD用のReference module 9.1/8.3/7.4-RELEASE

https://github.com/yahoo/l3dsr

L3DSR Packet Walkthrough

DSCP⇔VIP変換表に関して

DiffServは6bitしかないのに、IPアドレスはどう格納しますか。

Answer: LBとサーバそれぞれで、変換テーブルを持っています。

FreeBSDの変換表

```
l3dsr01# sysctl -a | grep dscp↓
net.inet.ip.dscp_rewrite.7: 144.49.0.250↓
net.inet.ip.dscp_rewrite.enabled: 1↓
```

LBの変換表

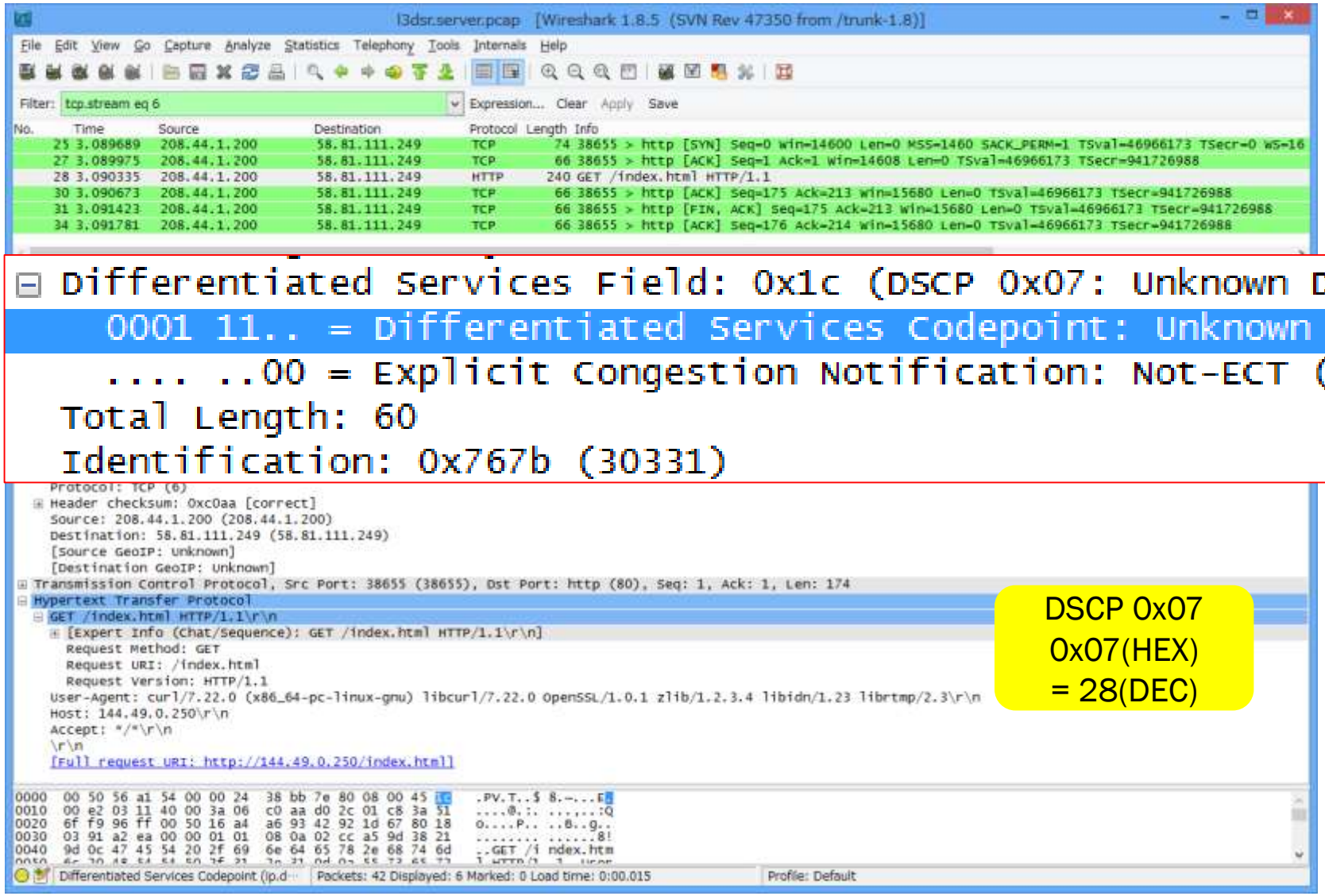
```
server virtual vip0 144.49.0.250
 tos-marking 7 hc-l3-dsr↓
 port http↓
 bind http rs5 http↓
```

VIP

DSCP

L3DSR Packet Walkthrough

RealServer上でPacketを見てみます。



DSCP 0x07
0x07(HEX)
= 28(DEC)

L3DSR Packet Walkthrough

Real Server自身にPingしてみます。

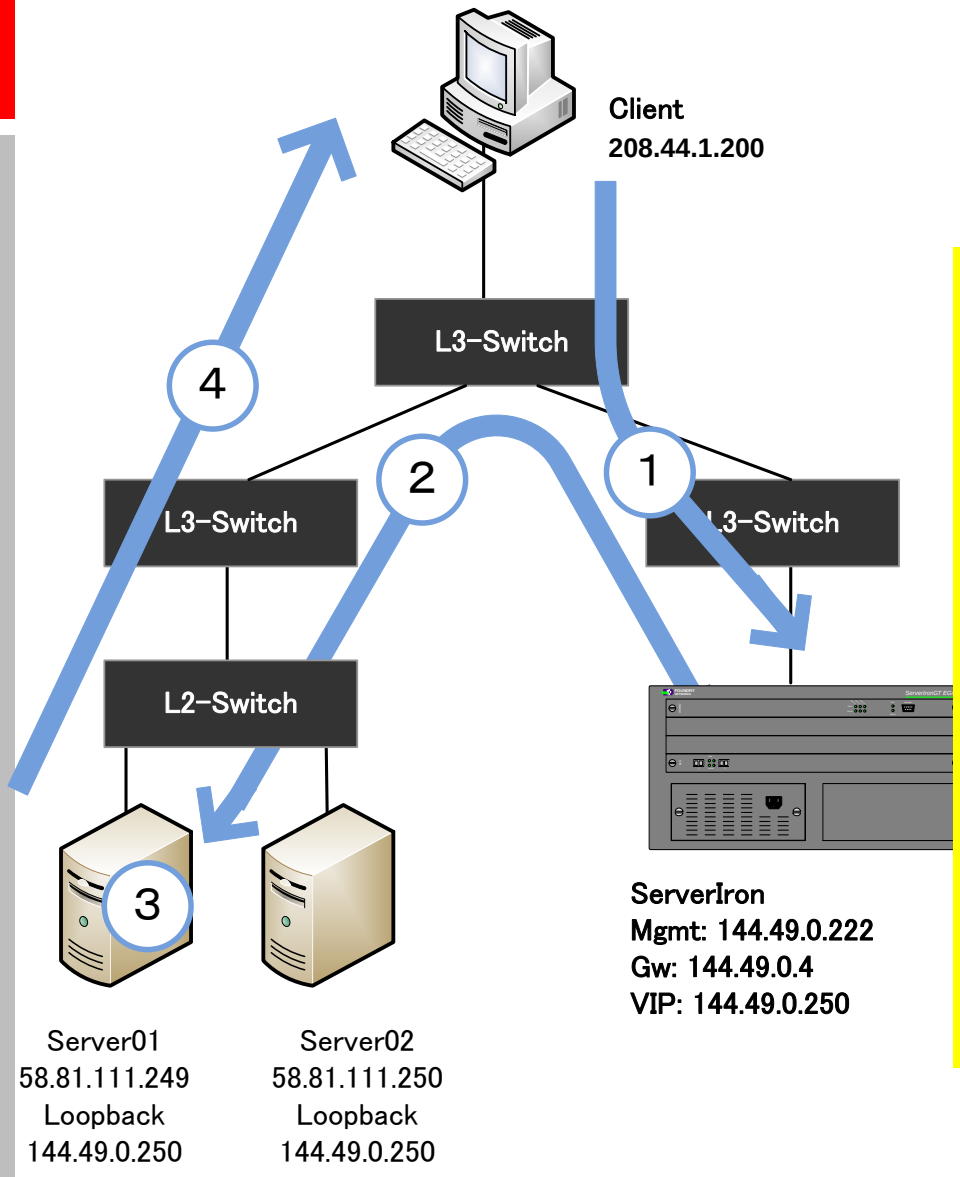
```
l3dsr01# ping -c 3 -z 28 58.81.111.249↓
PING 58.81.111.249 (58.81.111.249): 56 data bytes↓
64 bytes from 144.49.0.250: icmp_seq=0 ttl=64 time=0.027 ms↓
64 bytes from 144.49.0.250: icmp_seq=1 ttl=64 time=0.027 ms↓
64 bytes from 144.49.0.250: icmp_seq=2 ttl=64 time=0.029 ms↓
↓
--- 58.81.111.249 ping statistics ---↓
3 packets transmitted, 3 packets received, 0.0% packet loss↓
round-trip min/avg/max/stddev = 0.027/0.028/0.029/0.001 ms↓
l3dsr01# ↓
```

ping commandのoptionにおける　-z 28は、

0x07(16進数)を10進数に変換した値です。

```
kkono@l3dsr03:/home/kkono % man ping↓
PING(8)                  FreeBSD System Manager's Manual
↓
～省略～↓
↓
     -z tos  Use the specified type of service.↓
```


L3 DSR構成の話(おさらい)



ポイントのおさらい

a. LBとServerは同じ変換表を共有

b. LBは、宛先IPを変換してServerに送信。

c. ServerはDSCP bitを確認。

d. Serverは変換表に基づいて宛先を loopbackアドレスに書き換える。

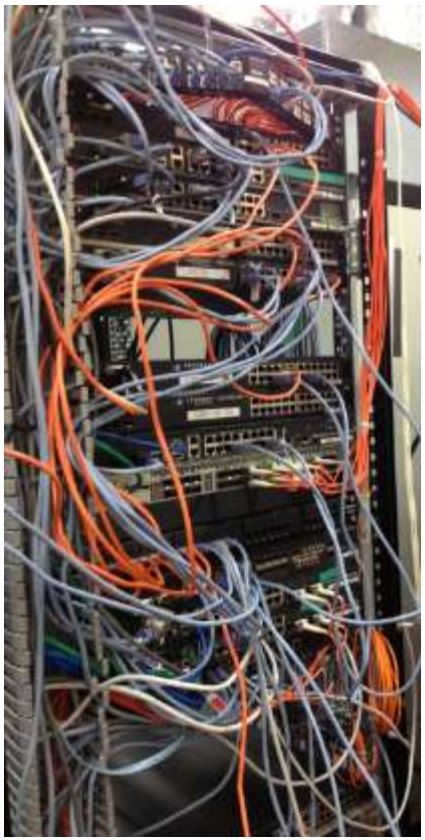
e. 後は、L2DSRと同様、ServerはVIPを SourceにしてClientに送信

その他 Kono Kono LABの紹介

Brocade LAB環境の紹介

私が占有しているのだけです。

LB 4台
Server 7 台
L3 2台



LB 10台
L3 8台
FC Switch 1台

Server 12台
esxi
10g x 2 port NIC
1g x 4 port NIC

BROCADE

Thank you!