



Innovative R&D by NTT

What's Next with Ryu

NTTソフトウェアイノベーションセンター

石田渉

会場ネットワーク使えていますか？



Innovative R&D by NTT

- **前方にお座りの方**
 - **快適ですか？**
- **後方にお座りの方**
 - **遅いですか？**



SDN実験ネットワーク

with lagopus/ryu



SSID : sdnjapan

Pass : ebisubeer

Openflowを利用した、
会場前方にお座りいただければ
いただくほど、快適にご利用いただける
実験ネットワークです

1. SDNツールキット Ryu

- **Ryuとは？**
- **どこで使われているのか**
- **どんなユースケースがあるのか**

2. Ryuの今後：ホワイトボックススイッチ

- **取り巻く情勢と我々の取組**

1. SDNツールキット Ryu

- Ryuとは
- どこで使われているのか
- どんなユースケースがあるのか

2. Ryuの今後：ホワイトボックススイッチ

- 取り巻く情勢と我々の取組

Ryuとは？

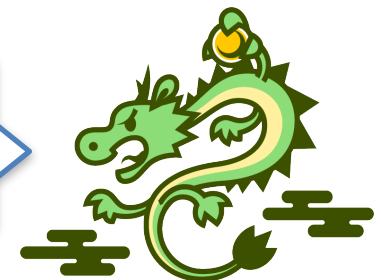


- NTTソフトウェアイノベーションセンタ発OSSの
OpenFlowコントローラ SDNツールキット (python)
 - Github: <https://github.com/osrg/ryu>
- 最新のOpenFlow規格に対応(~OF1.4, Nicira拡張)
- 多種多様なパケットライブラリ
 - ip, tcp, udp, mpls, pbb, dhcp, ospf...
- LLDP, VRRP, BGP, BMP, **BFD**

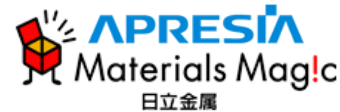
NEW



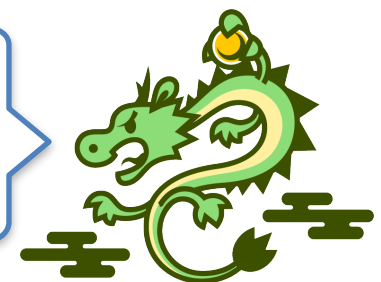
: ryusdn



どこで使われているのか？



いろいろな方に使っていただいています！

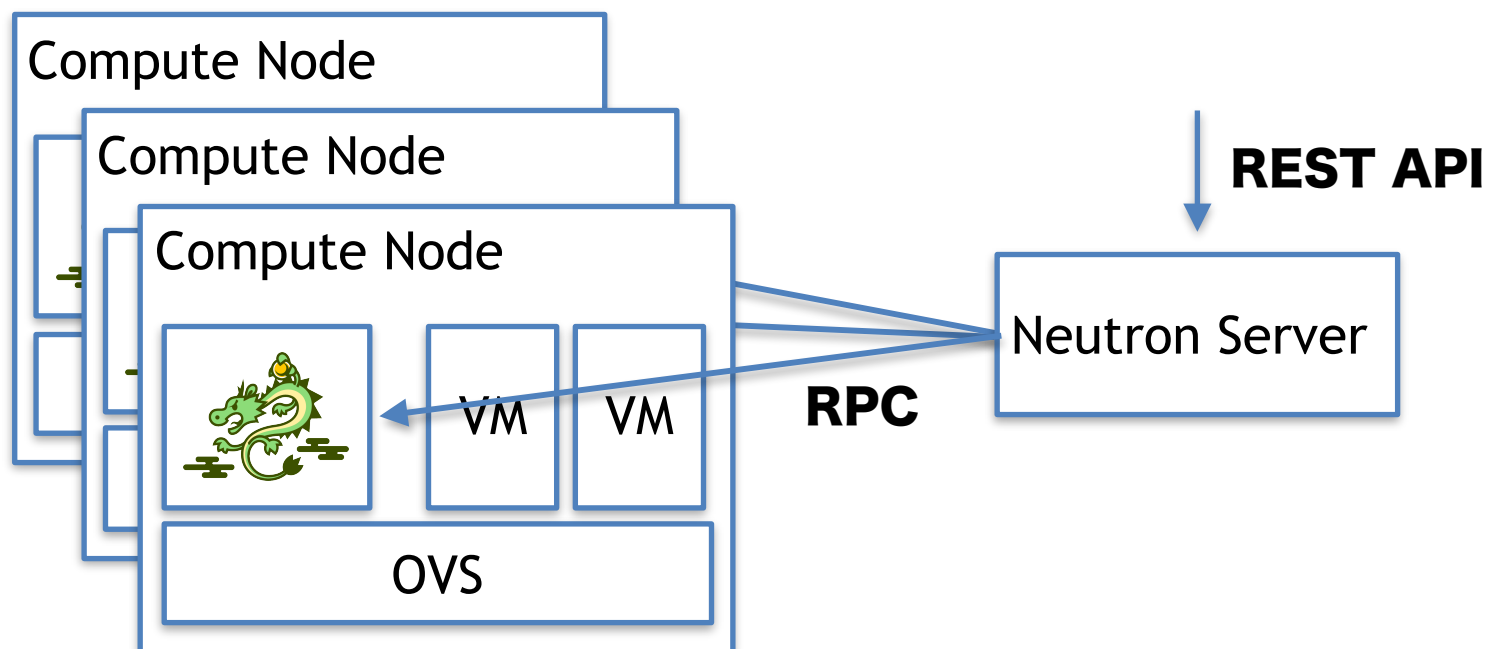


OpenStackとの連携



Innovative R&D by NTT

- **OFAgent Agent driver**
 - **Ryuを活用した分散型ネットワーク制御**
 - **テナントに仮想L2ネットワークを提供**



- **OpenflowはComputeノード内のみで利用**

Ryu Certification



- **Openflow準拠スイッチ？**
 - 複雑な仕様、たくさんのOptional…
 - サポートされている機能はすぐにわからない
 - 例) **✓ IPv4パケットへのMPLSヘッダのPush**
 - **✗ IPv6パケットへのMPLSヘッダのPush**
- **Openflow v1.3, v1.4準拠を謳うスイッチをテストするテストセンタを構築し、結果を公開！**
 - **<http://osrg.github.io/ryu/certification.html>**
- **ONF OF TestもRyu Certificationの利用を検討中**

Ryu Certification テストセンタ



Ryu Certification - lagopus

osrg.github.io/ryu-certification/switch/l...

Ryu Certification

- lagopus
- Configuration

	OK	ERROR
Action	56	0
(Required)	(3)	(0)
(Optional)	(53)	(0)
set_field	161	9
(Optional)	(161)	(9)
Match	714	0
(Required)	(108)	(0)
(Optional)	(606)	(0)
Group	15	0
(Required)	(3)	(0)
(Optional)	(12)	(0)
Meter	34	2
(Optional)	(34)	(2)
Total	980	11
(Required)	(114)	(0)
(Optional)	(866)	(11)

Action

Required	IPv4	IPv6	ARP
OK	OK	OK	OK

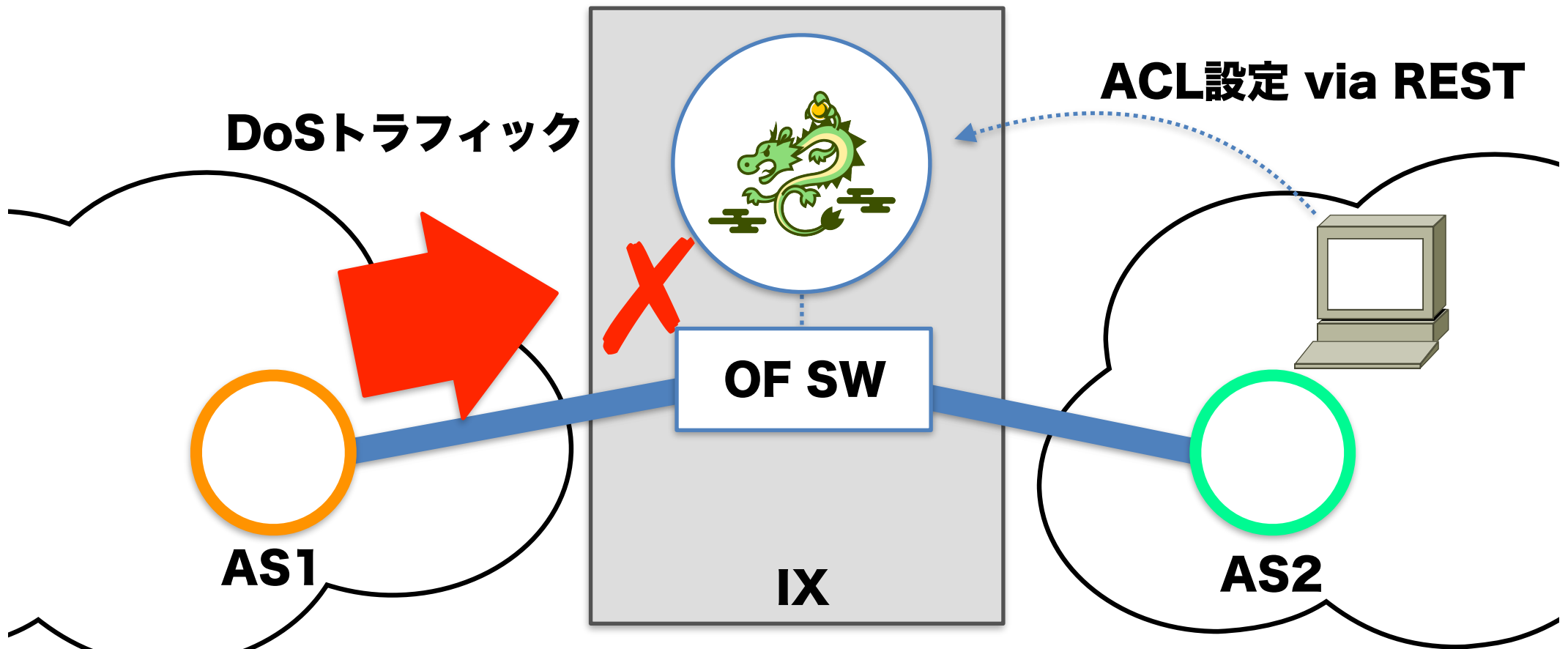


1. SDNツールキット Ryu

- Ryuとは
- どこで使われているのか
- **どんなユースケースがあるのか**
 - 1. SDN-IX - DOS mitigation**
 - 2. BGP,MPBGP**
 - 3. BMPサーバ**
 - 4. 会場ネットワーク**

SDN-IX - DoS mitigation

- ・インターネットマルチフィード
 - ・ Interop 2014でのデモ
- ・ より上流でトラフィックをブロックする



BGP, MPBGP



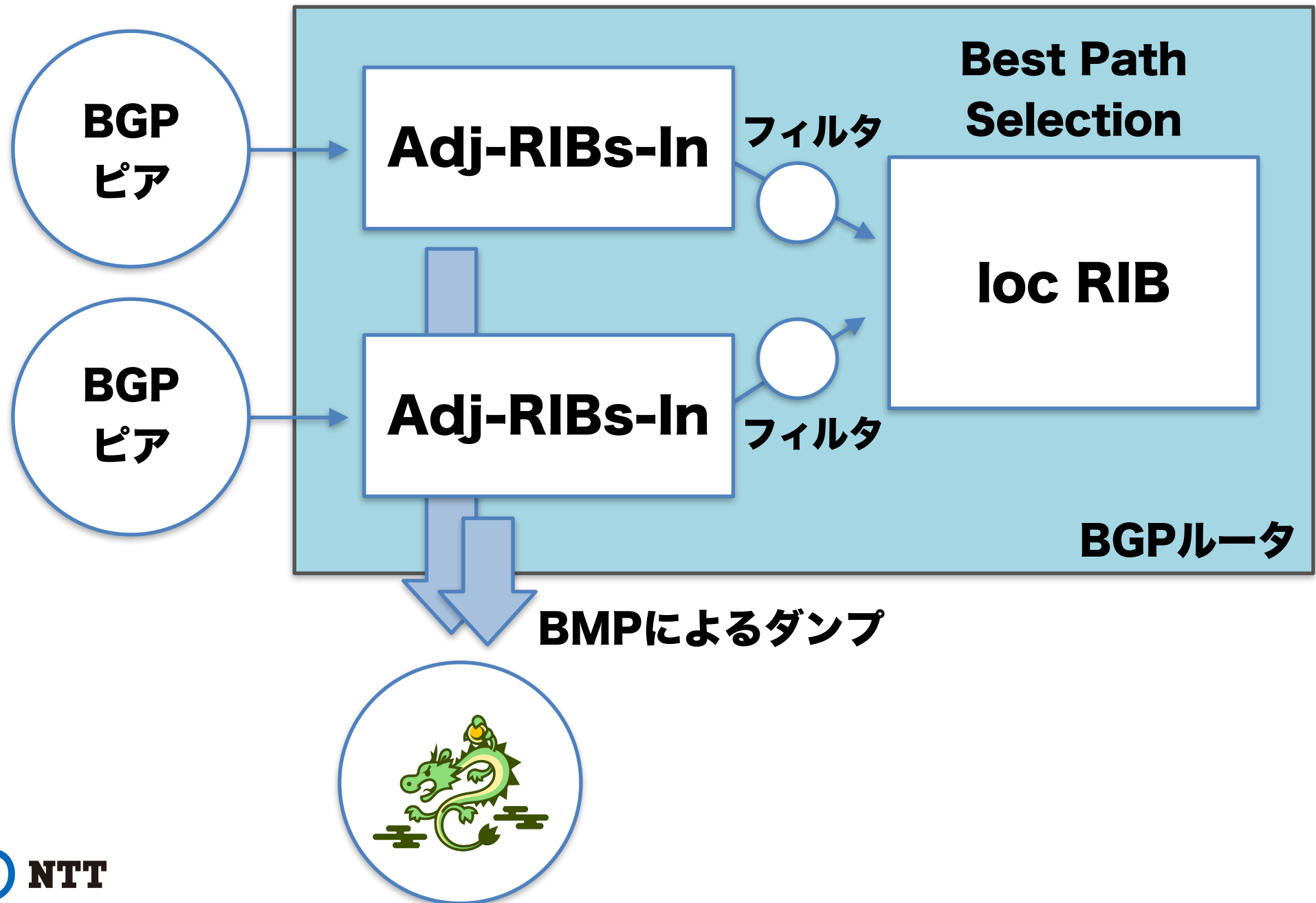
- **BGPルータとして動作可能**
- **BGP4だけでなくMPBGPにも対応**
 - **IPv6, VPN IPv4, VPN IPv6の経路も運べる**
 - **VRFもあるため、テナント毎の経路保持も可能**
- **Quagga, Birdの置き換えを狙っている**
- **今後ルートサーバ機能を拡張予定**

BMPサーバ



- **BMP = BGP Monitoring Protocol**
 - **Adj-RIBs-Inを覗くことが出来る**
- **BGPのパケットライブラリを利用**
- **Cisco ASR1000(IOS-XE 3.13S), Juniper MX960(JUNOS 13.3R1.8)で動作検証済**

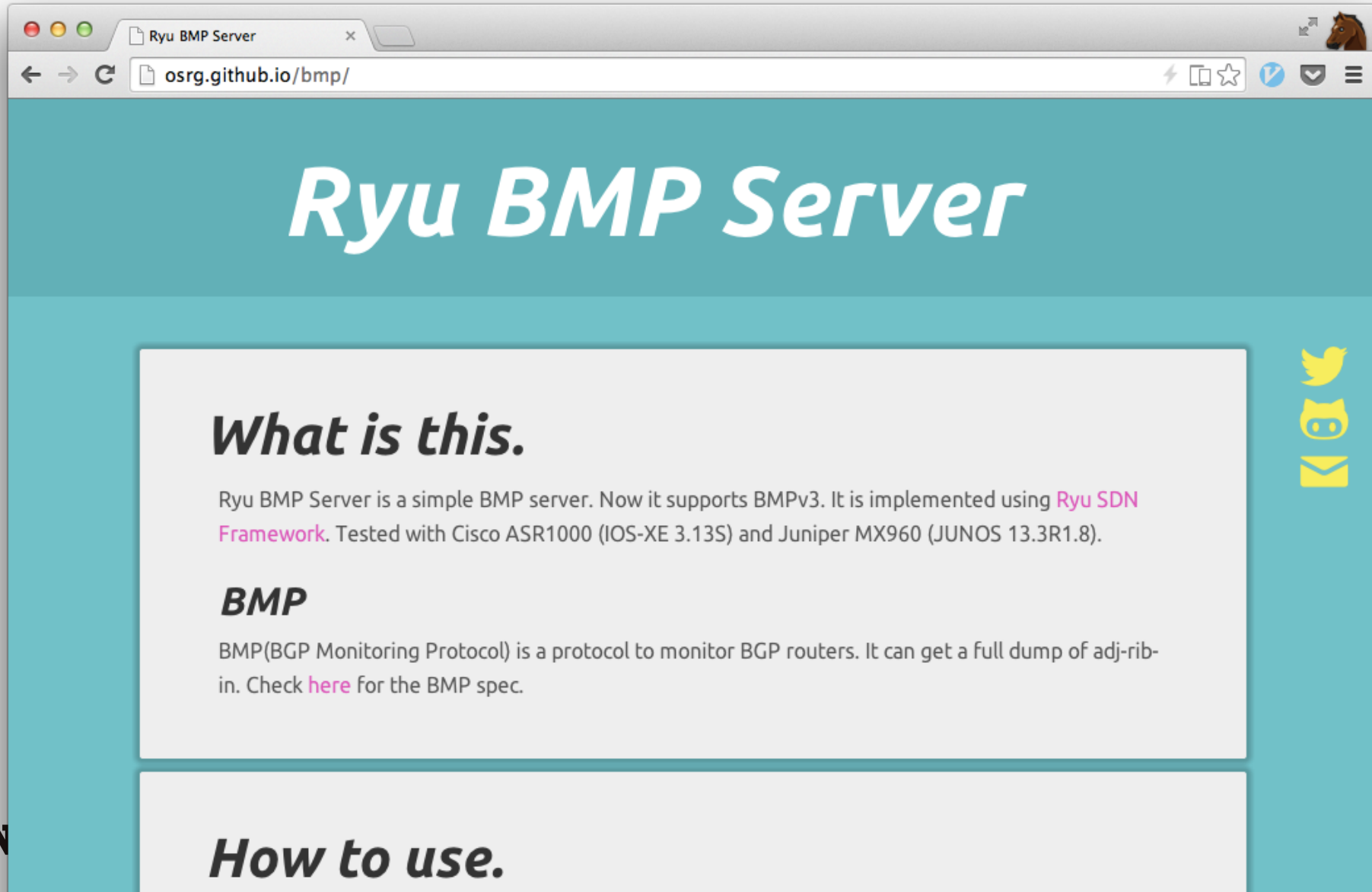
BMPサーバ



BMPサーバ



- <http://osrg.github.io/bmp/>

A screenshot of a web browser displaying the Ryu BMP Server website. The browser window shows the URL "osrg.github.io/bmp/" and the page title "Ryu BMP Server". The website has a teal header with the title "Ryu BMP Server" in white. Below the header is a white content area with the heading "What is this." and a paragraph describing the server. A vertical sidebar on the right contains social media icons for Twitter, GitHub, and YouTube. The bottom of the page shows the heading "How to use.".

Ryu BMP Server

What is this.

Ryu BMP Server is a simple BMP server. Now it supports BMPv3. It is implemented using [Ryu SDN Framework](#). Tested with Cisco ASR1000 (IOS-XE 3.13S) and Juniper MX960 (JUNOS 13.3R1.8).

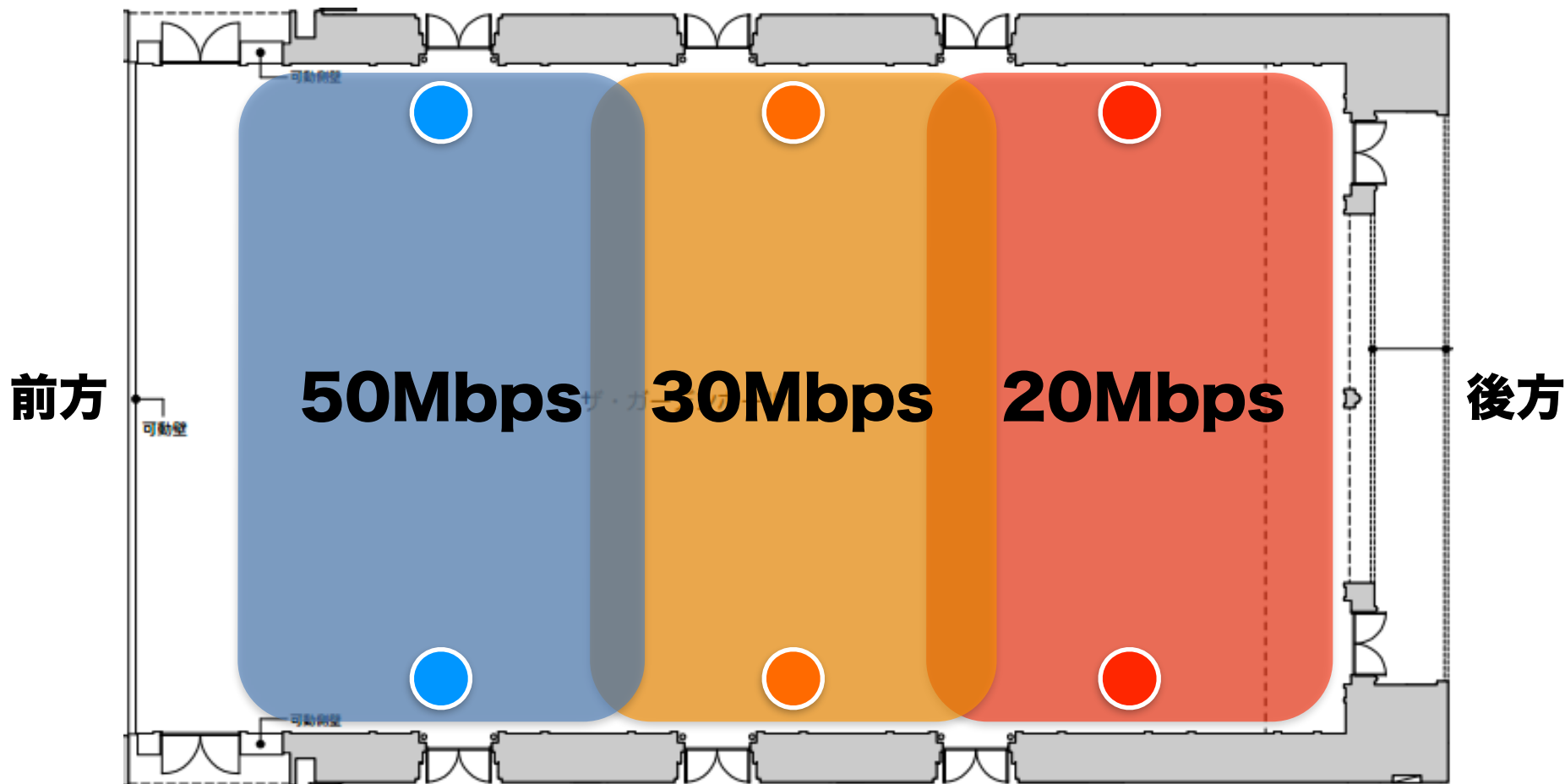
BMP

BMP(BGP Monitoring Protocol) is a protocol to monitor BGP routers. It can get a full dump of adj-rib-in. Check [here](#) for the BMP spec.

How to use.

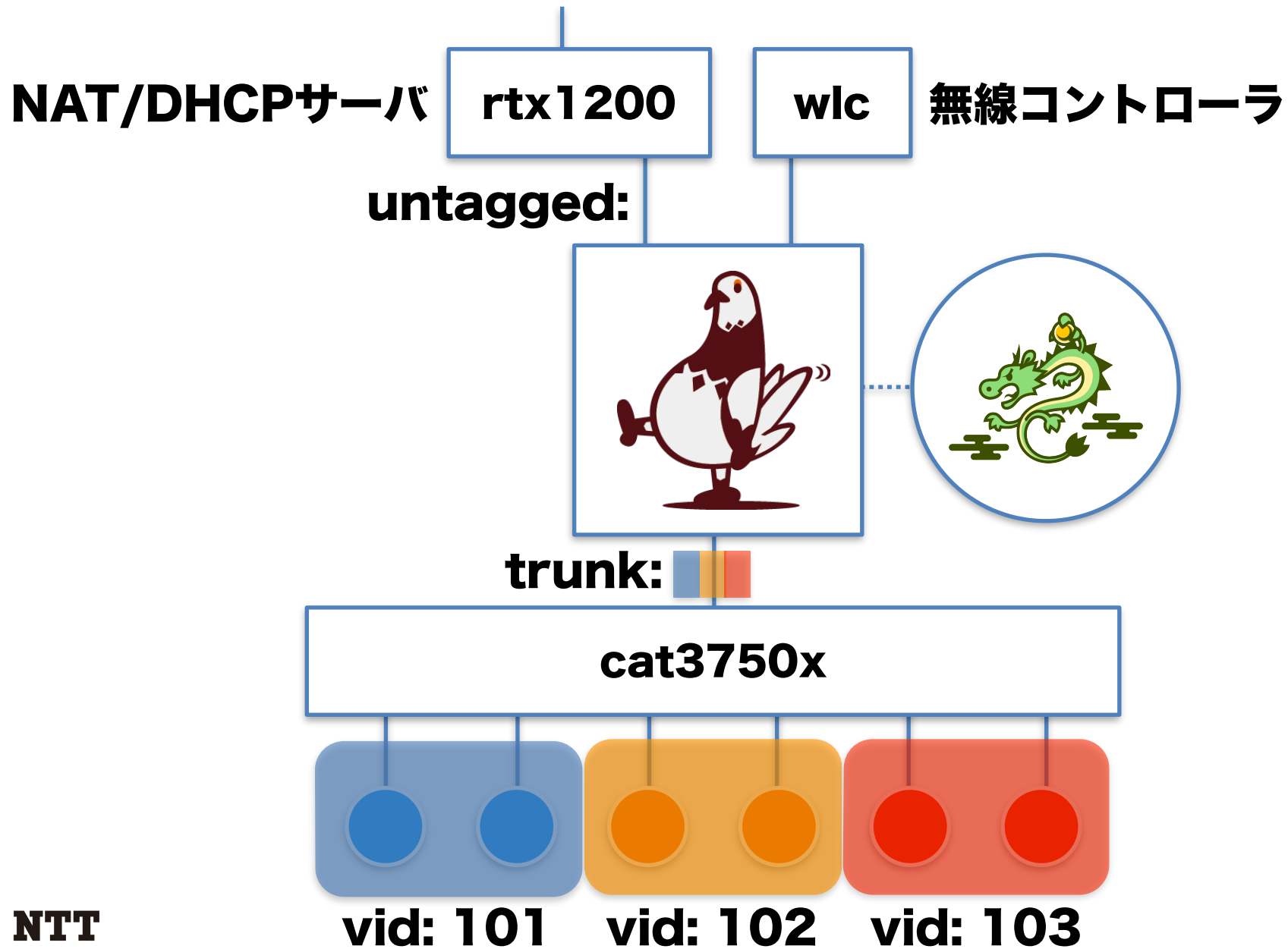


会場ネットワーク

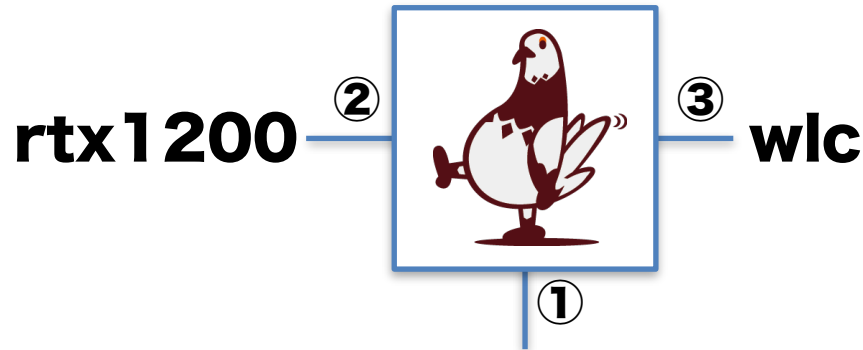


前に座るほどさくさく繋がるネットワーク

会場ネットワークトポロジ



lagopus フローテーブルデザイン



cat3750x

id: 0

match: src mac
action: pop vlan,
goto id:1

default: パケイン

id: 1

match: dst mac
action: push vlan,
meter, output

default: flood

group

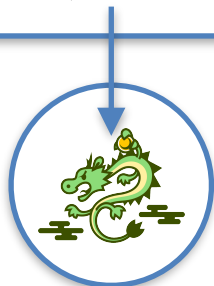
- ③からのpktをflood
- ②からのpktをflood
- ①からのpktをflood

meter

50Mbps

30Mbps

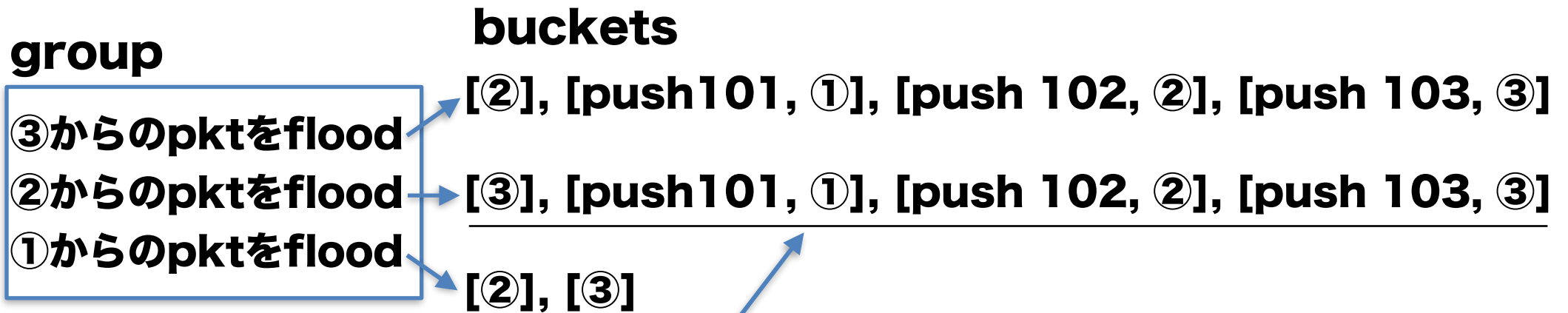
20Mbps



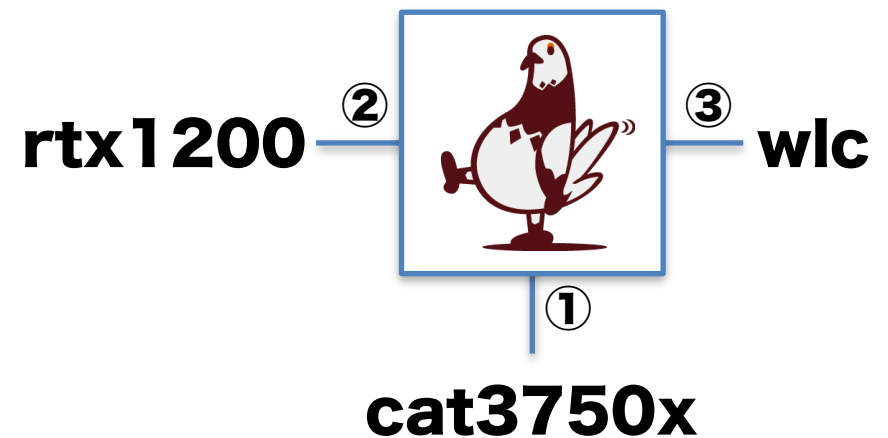
lagopus グループテーブルデザイン



- Group Type: **ALL, SELECT, INDIRECT, FF**
- **floodingのためALLを使用**



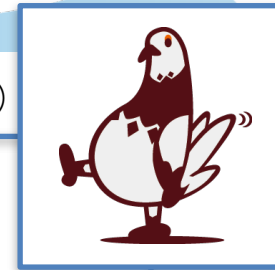
bucket内の全てが実行される
(SELECTの場合、どれか一つ
選ばれる)



パケットの流れ

rtx1200

②



③

wlc

①

cat3750x

group

- ③からのpktをflood
- ②からのpktをflood
- ①からのpktをflood

meter

50Mbps

30Mbps

20Mbps

match: src mac
action: pop vlan,
goto id:1

id: 0

101 aa..aa ff..ff

from ①

default: パケイン

フロー登録

match: dst mac
action: push vlan,
meter, output

id: 1

default: flood

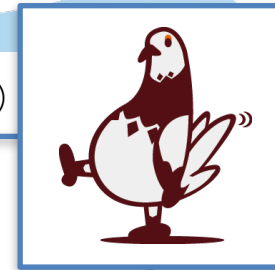
パケアウト



パケットの流れ

rtx1200

②



③

wlc

①

cat3750x

group

③からのpktをflood
②からのpktをflood
①からのpktをflood

meter

50Mbps

30Mbps

20Mbps

match: **src mac**
action: pop vlan,
goto id:1

id: 0

if src mac **aa..aa**
then pop vlan,
goto id:1

default: パケイン

フロー登録



match: **dst mac**
action: push vlan,
meter, output

id: 1

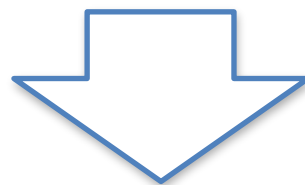
if dst mac **aa..aa**
then push **vlan101**,
meter **50Mbps**,
output①

default: flood

実験(ネタ)からビジネスに？



- **アップリンクの帯域が限られている時に、**
 - **特定のアプリケーションの帯域を絞る**
 - **スマホの様にxxGBを超えると帯域制限**
 - **アンケートに答えると帯域制限がなくなる 等々**



- **このようなことがコモディティ化されたIAサーバとOSS
で実現できそう**

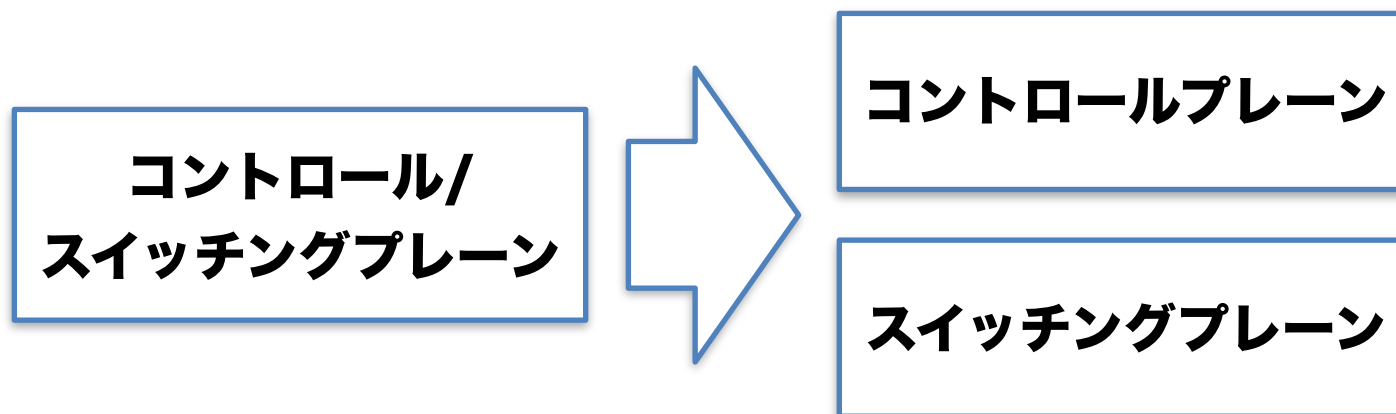
1. SDNツールキット Ryu

- Ryuとは
- どこで使われているのか
- どんなユースケースがあるのか

2. Ryuの今後：ホワイトボックススイッチ

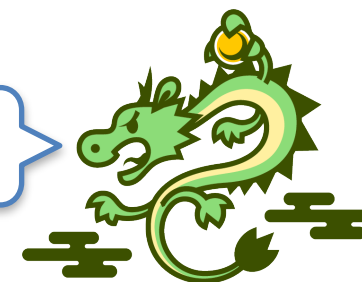
- 取り巻く情勢と我々の取組

- SDNの本質 = ネットワークのコモディティ化



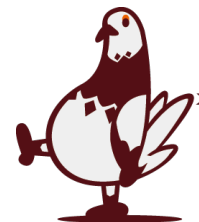
- コントロールプレーンはIAサーバ上で動く
 - コントローラはOSSも含め多種多様
 - コモディティ化している

ではスイッチングプレーンは？



- **スイッチングプレーン**

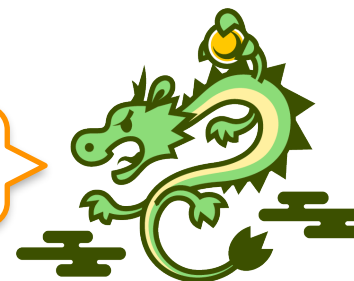
1. **スイッチングもIAサーバでやる**



2. **ネットワーク機器もIAサーバ並にコモディティ化？**

- ホワイトボックススイッチ

Next Target !!



ホワイトボックススイッチとは



- ベアメタルスイッチとも
- OSが載っていない
 - ユーザが自分でインストール
- ODMベンダが製造

Quanta

Edge-core

 **Agema**

**PENGUIN
COMPUTING™**

- 汎用チップセットを搭載

**BROADCOM®**

**Mellanox®
TECHNOLOGIES**



**centec
networks**

なぜいま登場してきたのか



- ネットワーク機器のスイッチングコンポーネント変遷
 - 1980年代後半：CPUの時代
 - 1990年代後半：カスタムチップの時代
 - 2000年代後半：汎用チップセットの台頭



なぜいま登場してきたのか



Innovative R&D by NTT

- ・ 汎用チップセットの登場で

ネットワークベンダのソフトウェア

+

ODMベンダのハードウェア

なネットワーク機器が登場

秋葉原で買えないと流行らない？



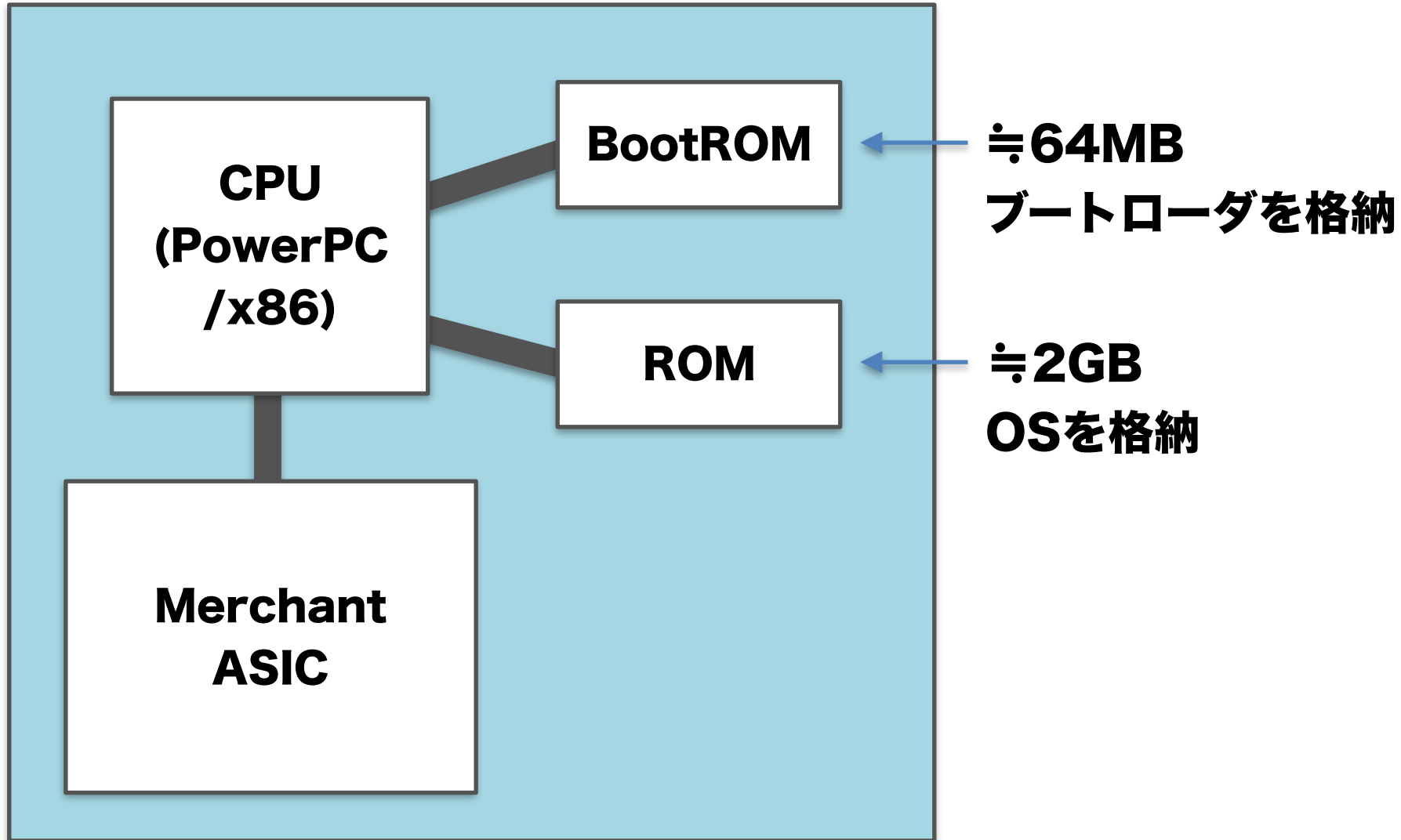
- <http://bm-switch.com/>

The screenshot shows a web browser window displaying the 'All switches' collection on the bm-switch.com website. The browser address bar shows 'bm-switch.com/collections/all-switches'. The website header includes a logo with a penguin on a switch, navigation links for 'HOME', 'BARE METAL SWITCHES', 'CUMULUS SNS', 'ACCESSORIES', and 'BLOG', and a search bar. A shopping cart icon shows '0 items' and '\$0.00'. Below the header, the page title is 'All switches'. A paragraph of text describes the company: 'Established in 1988, Quanta Computer is the largest notebook computer ODM company in the world. Quanta bare metal switches are compatible with Cumulus Linux, Pica8 and BigSwitch Switch Light OS. Product range lies from 1G Base-T switch to high-density 40G box for datacenter and telecom applications.' Below this text, four product cards are displayed, each with an image of a switch and its name and specifications:

- Quanta T1048-LB9
1G/10G BM switch
preloaded with ONIE
- Quanta T3048-IZ1
10/40G BM switch with
Intel ONS support
- Quanta T3048-LY2
10/40G BM switch
preloaded with ONIE
- Quanta T3048-LY2R
10/40G BM switch
preloaded with ONIE



スイッチアーキテクチャ



ONIE



- **スイッチにOSをカジュアルにインストール/アンインストールするための環境**



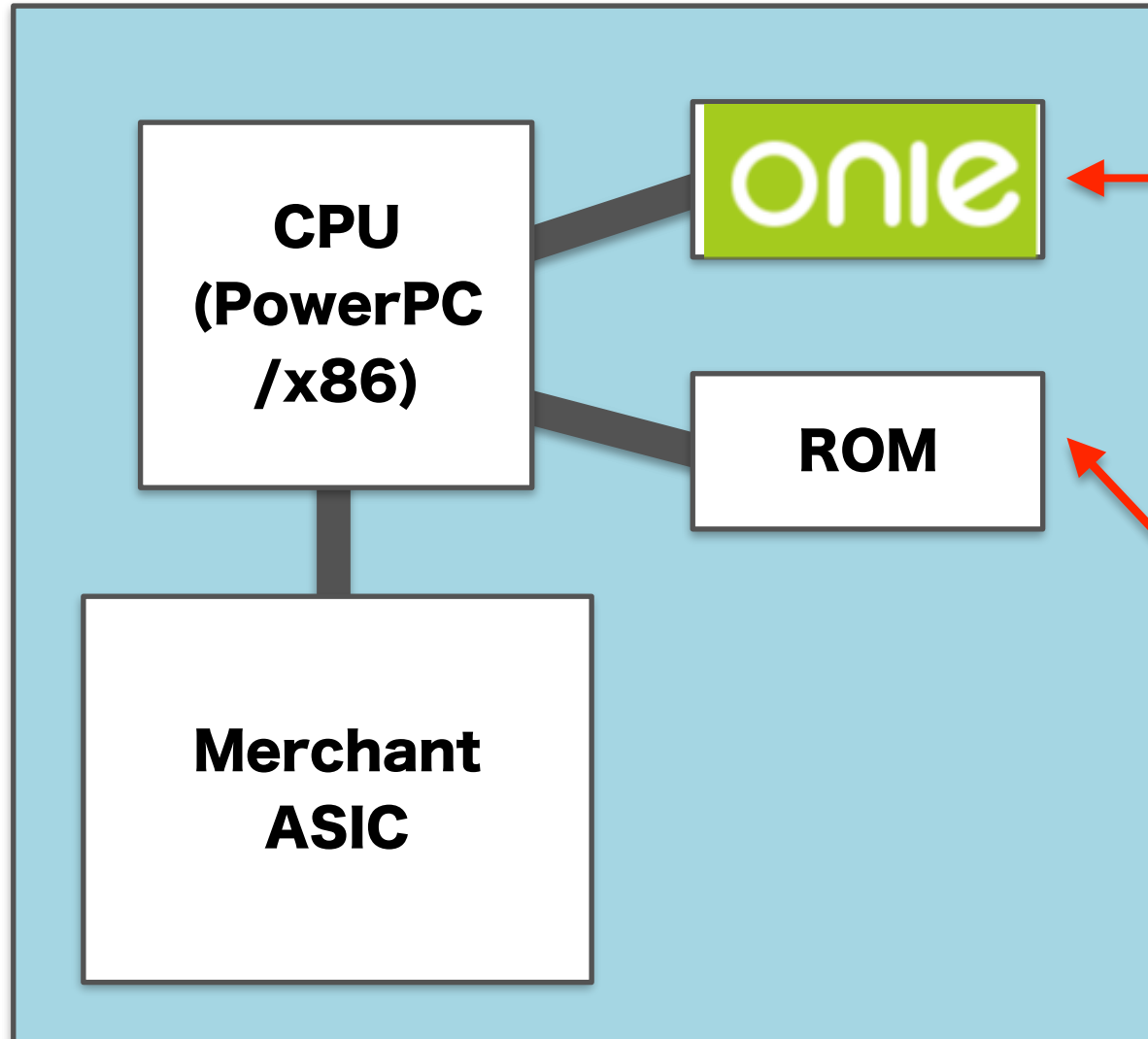
: Open Network Install Environment



OPEN
Compute Project

: Facebookが主導するOpen Computeによるサポート

ONIEによるOSインストール



1. ここに焼かれて出荷
2. 電源を入れるとDHCPでインストーラのURLを取得
3. インストーラをダウンロードし、ここにインストール
4. ブートローダの起動順を変更し、リブート後はROMにインストールされたOSをブート

OSの選択肢



Innovative R&D by NTT



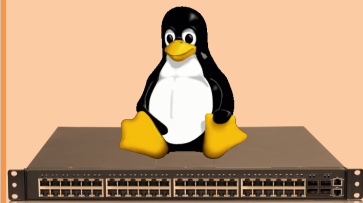
Cumulus Linux (Cumulus Networks)

- ・ ネットワークコンフィグをLinuxライクに



Switch Light OS (Big Switch)

- ・ OpenFlowスイッチ
- ・ OpenNetworkLinuxベース



OpenNetworkLinux (Big Switch/OSS)

- ・ ASICの制御コンポーネントは含まない

Open Network Linux



- **一番重要なASICの制御コンポーネントは含まれない**
- **何の役に立つ？**
 - **スイッチ毎に異なるファン、センサの制御のためのコンポーネントが揃っている**
 - **対応しているスイッチはAcctonを中心に9機種**
 - **Cumulusは19機種**
- **ONLをベースにASICの制御コンポーネントは 各人、各社実装することを前提にしている**

Open Network Linux



- **何故ASICの制御コンポーネントが含まれないのか？**
 - **ビジネス戦略？**
 - **ASICのSDKはNDAを結ばないと利用不可**
 - **ビジネス戦略でなくとも公開は難しいかもしれない**

- **Broadcomが公開したライブラリ**
- **<https://github.com/Broadcom-Switch/of-dpa>**
- **約160個のAPIを使ってBroadcom ASICを制御することができる**
- **大部分はバイナリ提供され、不自由なことも多いが、ASICがx86同様、着実にオープン化してきている**
- **ONLとof-dpaを組み合わせれば、ホワイトボックススイッチOSをOSSで開発できる？ => 鋭意開発中!**

まとめ



- **Ryu** : いろいろなところで使っていただいています
- **Ryuの今後**
 - **コントロールプレーンのコモディティ化の次はスイッチングプレーンのコモディティ化!**
 - **現在、ホワイトボックススイッチ向けLinuxディストリビューションの開発中**
 - **OSSとして公開予定**