Yahoo! JAPAN Site Operations

OCP ネットワークスイッチ

サイトオペレーション本部
インフラ技術3部
村越 健哉

自己紹介



- 名前
  - 村越 健哉（むらこし けんや）
- 所属
  - 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3部
- 仕事
  - ヤフーのプロダクションネットワーク全般

Yahoo! JAPAN Site Operations

アジェンダ

- 背景
- OCPネットワークスイッチ Backpackとは
- Backpack 外観
- Backpack アーキテクチャ
- Backpack CLI
- OMP800@台湾

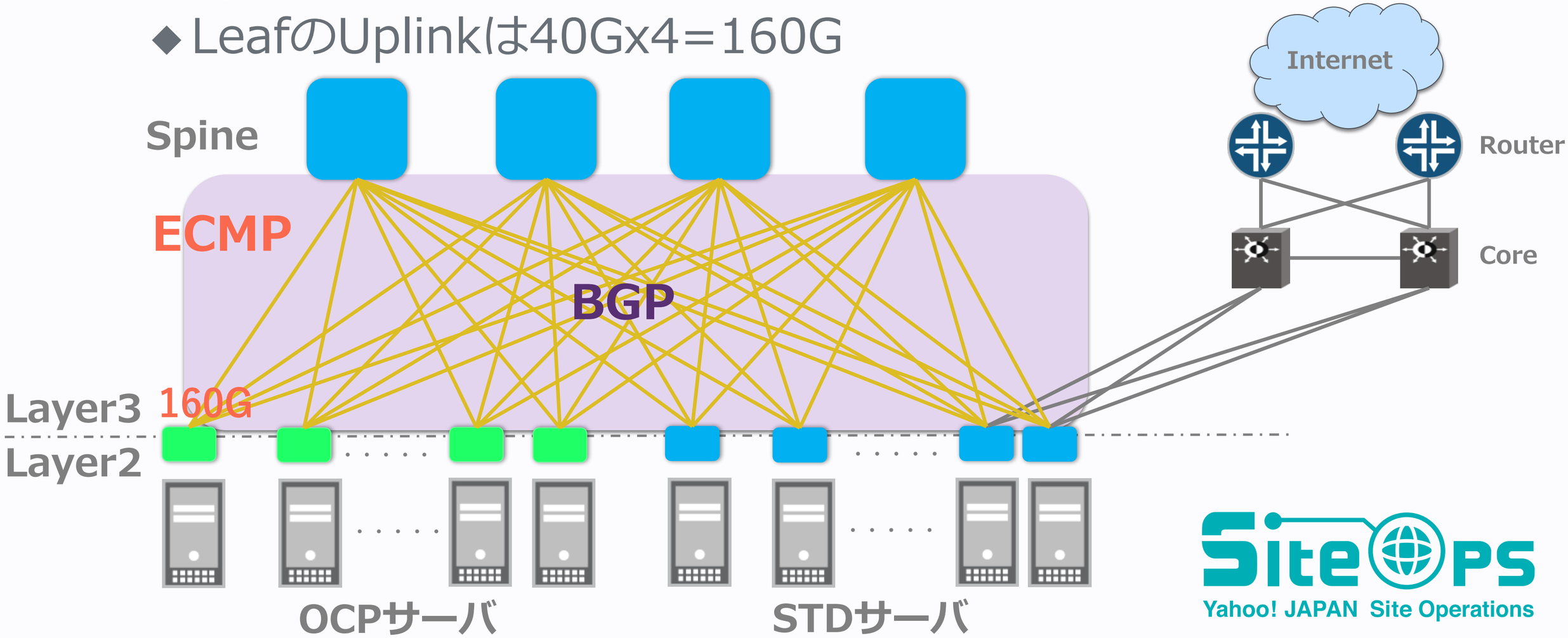
Yahoo! JAPAN Site Operations

背景

背景

https://www.janog.gr.jp/meeting/janog38/download_file/clos.pdf

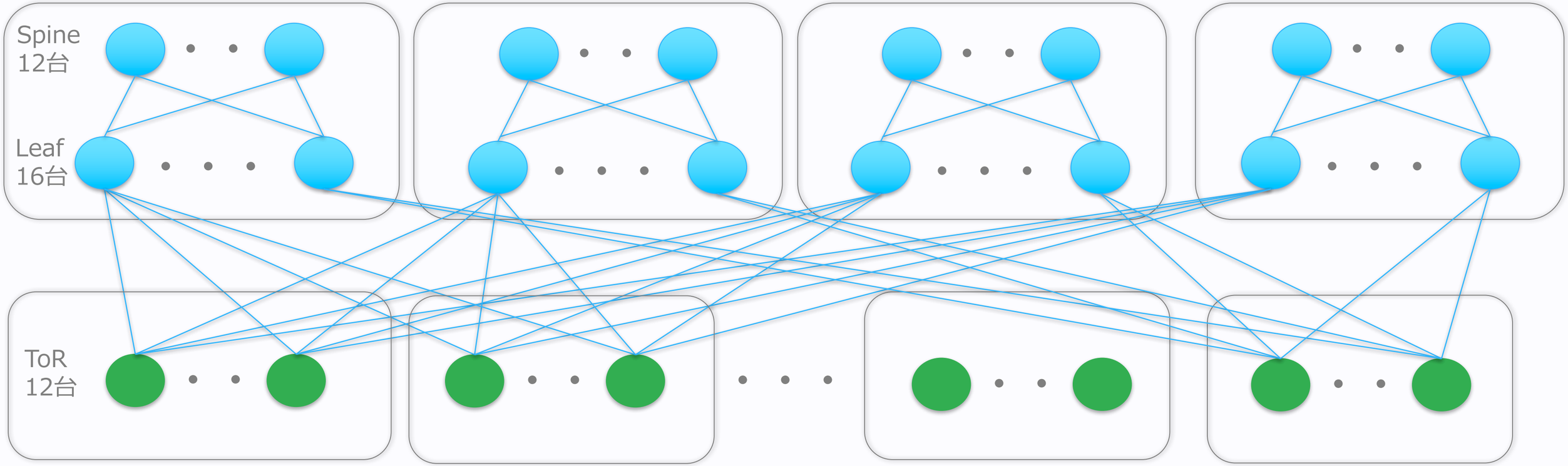
- Janog38で紹介したIP CLOS ネットワーク
  - Spine-Leaf間はBGP
  - LeafのUplinkは40Gx4=160G



Yahoo! JAPAN Site Operations

背景

■ ボックススイッチ構成



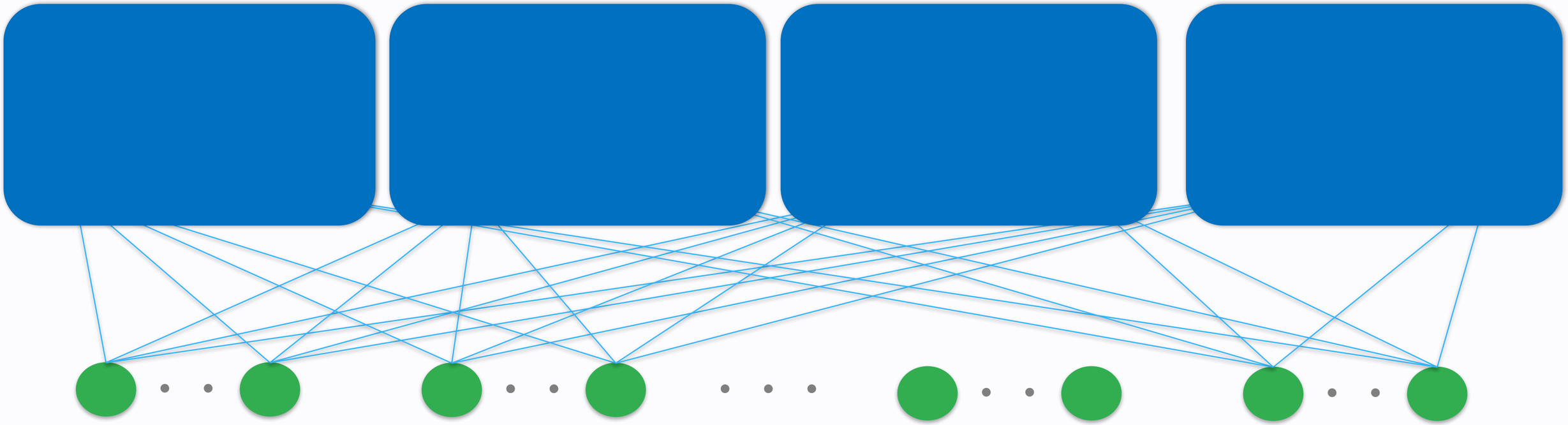
◆Spine 12台、Leaf 16台の場合

- ToR 12(Spineに依存）x 16セット = 192台（ラック）

Yahoo! JAPAN Site Operations

背景

- シャーシスイッチ構成



◆Slot 8モジュールの場合

- 8 x 32port = 256台（ラック）

Yahoo! JAPAN Site Operations

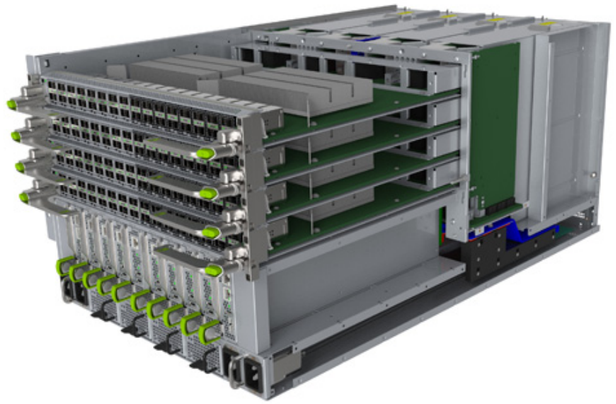
背景

- シャーシスイッチのデメリット
  - ブラックボックス
    - 例）Farbic ～ Line Card間のどこをパケットが通っているか落ちているか不明
  - 故障影響度が大きい
    - SUPが壊れると全体がダウン
  - 複雑なソフトウェア処理
    - SUPを冗長させると処理が複雑になっていく
    - SSO, NSF, ISSU,
  - ソフトウェアアップグレードに時間がかかる

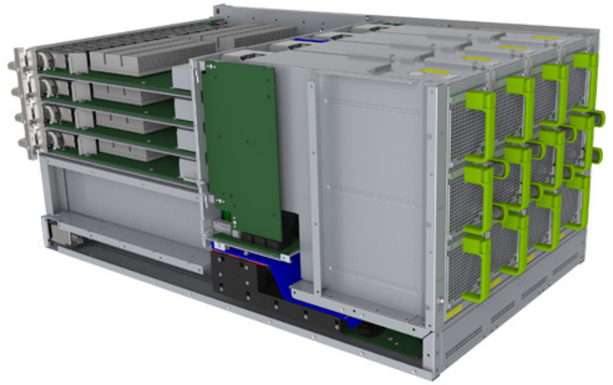
OCPネットワークスイッチ Backpackとは

Backpack

- Facebookで既に稼働
  - 技術仕様がOCPに公開されている







「Introducing Backpack: Our second-generation modular open switch」
https://code.facebook.com/posts/864213503715814/introducing-backpack-our-second-generation-modular-open-switch/

Backpack

- Facebookが開発したスイッチ
  - サイズ
    - 35.2cm(H) x 44.3cm(W) x 74.9cm(D)、8RU
  - ポート密度
    - 128x100G、QSFP28
  - ASIC
    - Broadcom BCM56960 Tomahawk
  - Line Card
    - Tomahawk ASIC x2枚
    - QSFP28 100G x32port
  - Fabric
    - Tomahawk ASIC x1枚

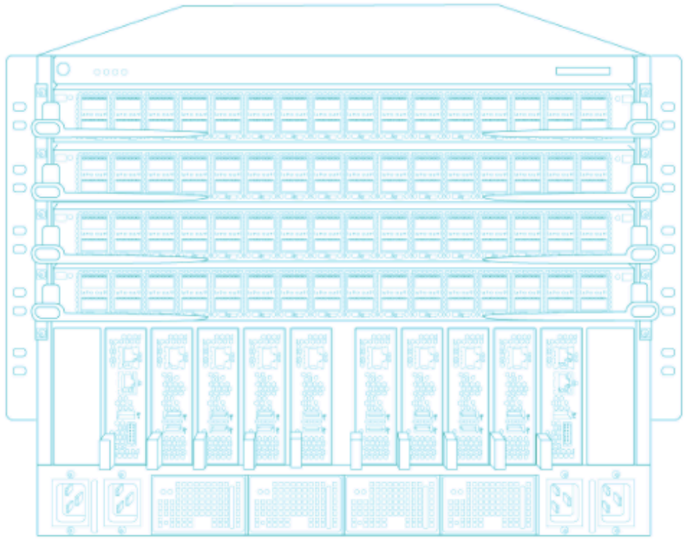
Backpack running Cumulus

- BackpackがCumulusで動作可能

CUMULUS

Blog Home RSS

Announcing Backpack running Cumulus Linux — completing our networking portfolio with a modular platform



Open Networking Topics

- Technology
- Industry & Market Trends
- Ecosystem

Quick Nav

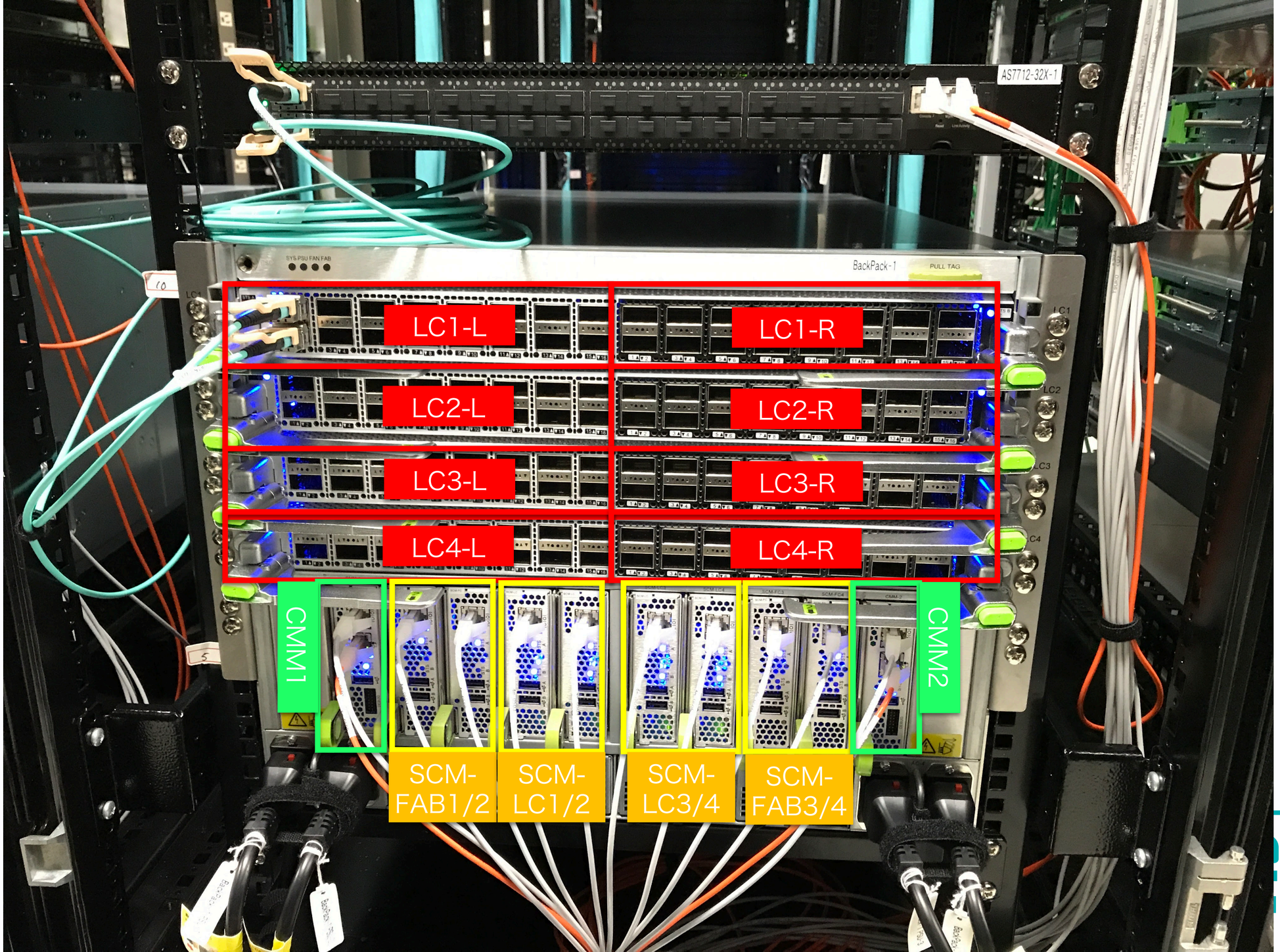
- Blog Authors
- Resources

DON'T MISS A POST

Subscribe to receive an email when we

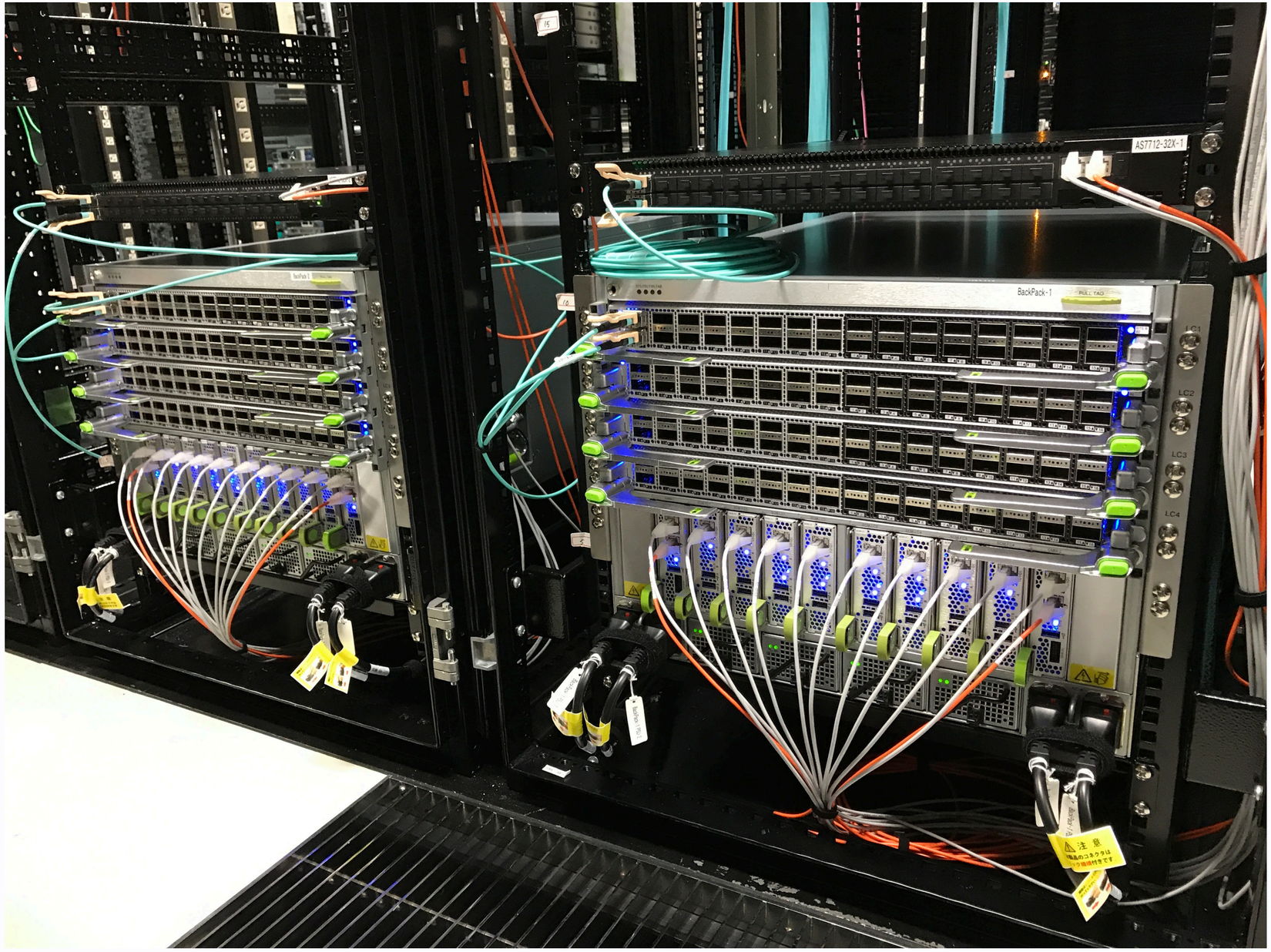
https://cumulusnetworks.com/blog/announcing-backpack-running-cumulus-linux/

Backpack 外観


AS7712-32X-1
BackPack-1
PULL TAG
LC1-L
LC1-R
LC2-L
LC2-R
LC3-L
LC3-R
LC4-L
LC4-R
CMM1
CMM2
SCM-FAB1/2
SCM-LC1/2
SCM-LC3/4
SCM-FAB3/4


FAB1
FAB2
FAB3
FAB4
FAN1
FAN2
FAN3

Backpack 外観



Backpack 外観

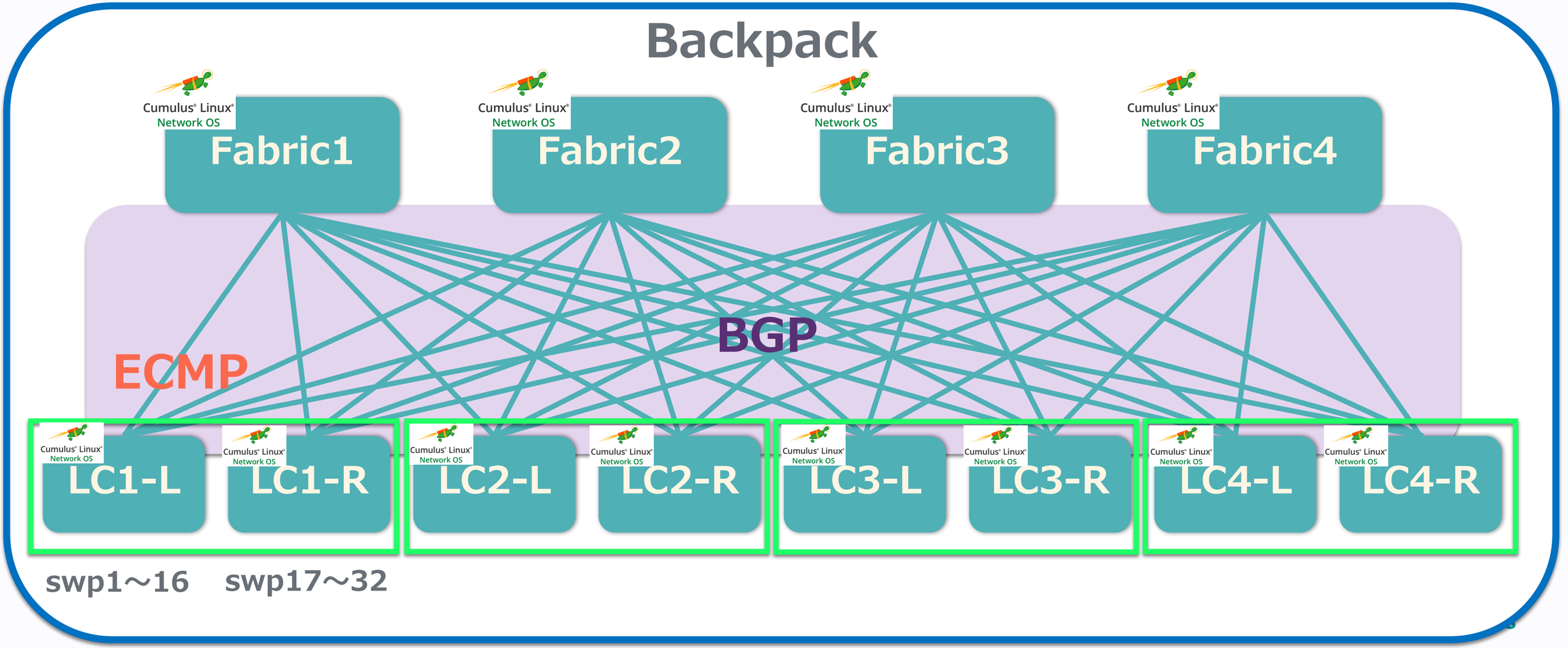




Backpackアーキテクチャ

Backpack アーキテクチャ

- FabricとLinceCardそれぞれにOSが稼働している

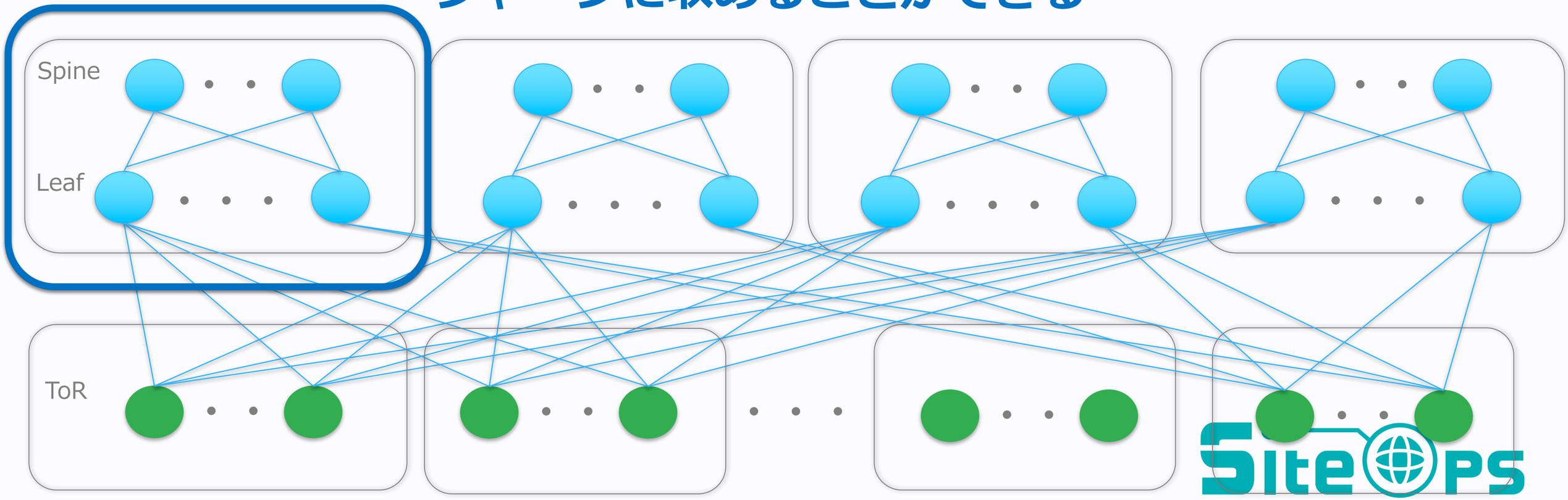


Backpack アーキテクチャ

- ボックススイッチ構成

Backpack

Spine～Leaf間の2階層を構成を変えずにシャーシに収めることができる



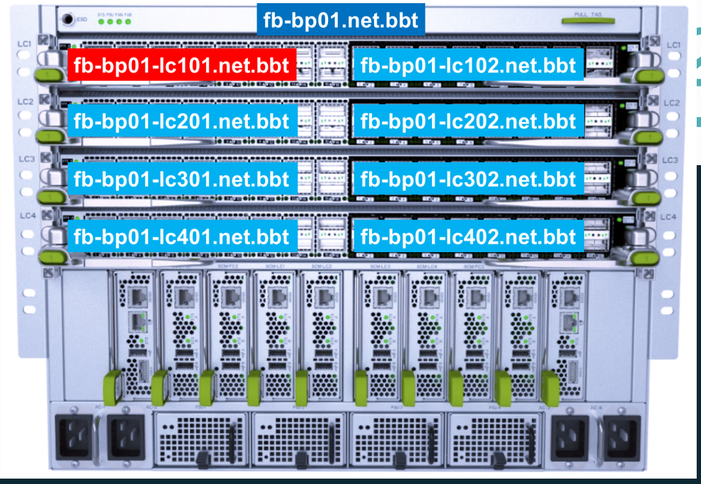
Yahoo! JAPAN Site Operations

Backpack メリット

- OCPネットワークスイッチ Backpackメリット
  - OCPとして技術仕様が公開されている
  - 故障影響が最小、複雑な処理の排除
    - SUPが不要（Single Point of Failureの排除）＝　SSO, NSF, ISSUなどが不要
  - 従来のシャーシスイッチとは異なり、個々のASICでOS稼働
    - Spine～Leafの2階層をまとめられる、BOXスイッチ12台分
  - Spine～Leaf間のラックU数、ケーブル接続、Optics購入を抑制できる
  - ソフトウェアアップグレードが早い
  - Ansible等の構成管理ツールでSpine,Leaf,ToRへ設定を一括投入・管理可能

Backpack CLI

Backpack CLI



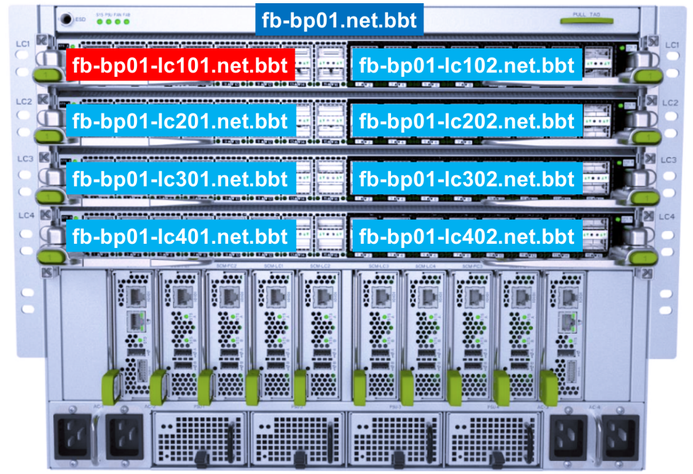

```
cumulus@fb-bp01-lc101:~$ sudo vtysh -c 'show ip bgp summary'
BGP router identifier 100.65.28.4, local AS number 65000 vrf-id 0
BGP table version 46
RIB entries 21, using 2856 bytes of memory
Peers 32, using 673 KiB of memory
Peer groups 2, using 144 bytes of memory

Neighbor                                V         AS MsgRcvd MsgSent   TblVer  InQ OutQ  Up/Down State/PfxRcd
fb-bp01-fab01.                 (fp0)    4      65000   22496   22486        0    0    0 01:55:19            7
fb-bp01-fab01.                 (fp1)    4      65000   22494   22485        0    0    0 01:55:18            7
fb-bp01-fab01.                 (fp2)    4      65000   22494   22485        0    0    0 01:55:18            7
fb-bp01-fab01.                 (fp3)    4      65000   22494   22485        0    0    0 01:55:18            7
fb-bp01-fab02.                 (fp4)    4      65000   22487   22486        0    0    0 01:55:18            7
fb-bp01-fab02.                 (fp5)    4      65000   22486   22485        0    0    0 01:55:18            7
fb-bp01-fab02.                 (fp6)    4      65000   22484   22484        0    0    0 01:55:17            7
fb-bp01-fab02.                 (fp7)    4      65000   22484   22484        0    0    0 01:55:17            7
fb-bp01-fab03.                 (fp8)    4      65000   22486   22485        0    0    0 01:55:18            7
fb-bp01-fab03.                 (fp9)    4      65000   22485   22484        0    0    0 01:55:18            7
fb-bp01-fab03.                 (fp10)   4      65000   22485   22484        0    0    0 01:55:18            7
fb-bp01-fab03.                 (fp11)   4      65000   22484   22483        0    0    0 01:55:17            7
fb-bp01-fab04.                 (fp12)   4      65000   22486   22485        0    0    0 01:55:18            7
fb-bp01-fab04.                 (fp13)   4      65000   22485   22485        0    0    0 01:55:18            7
fb-bp01-fab04.                 (fp14)   4      65000   22484   22484        0    0    0 01:55:17            7
fb-bp01-fab04.                 (fp15)   4      65000   22484   22484        0    0    0 01:55:17            7
es-accton01.                   (swp1)   4      65001   22489   22491        0    0    0 00:00:42            2
es-accton02.                   (swp2)   4      65001   22474   22483        0    0    0 00:00:52            2
```

Backpack CLI


```
cumulus@fb-bp01-lc101:~$ sudo vtysh -c 'show ip route'
Codes: K - kernel route, C - connected, S - static, R - RIP,
       O - OSPF, I - IS-IS, B - BGP, P - PIM, T - Table, v - VNC,
       V - VPN,
       > - selected route, * - FIB route

K>* 0.0.0.0/0 via 10[redacted]7.1, eth0
C>* 10[redacted]7.0/26 is directly connected, eth0
C>* 10[redacted]8.4/32 is directly connected, lo
B>* 10[redacted]8.5/32 [19/0] via fe80::2e0:ecff:fe59:2a38, fp0, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a3c, fp1, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a40, fp2, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a44, fp3, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ab9, fp12, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2abd, fp13, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ac1, fp14, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ac5, fp15, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b3a, fp8, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b3e, fp9, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b42, fp10, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b46, fp11, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bbb, fp4, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bbf, fp5, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bc3, fp6, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bc7, fp7, 00:03:08
B>* 10[redacted]8.6/32 [19/0] via fe80::2e0:ecff:fe59:2a38, fp0, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a3c, fp1, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a40, fp2, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2a44, fp3, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ab9, fp12, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2abd, fp13, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ac1, fp14, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2ac5, fp15, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b3a, fp8, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b3e, fp9, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b42, fp10, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2b46, fp11, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bbb, fp4, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bbf, fp5, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bc3, fp6, 00:03:08
  *                           via fe80::2e0:ecff:fe59:2bc7, fp7, 00:03:08
```



SiteOps
Yahoo! JAPAN Site Operations

Backpack CLI


```
TASK [leaf : quagga.conf] ***************************************************
ok: [accton01] => {"changed": false, "gid": 109, "group": "quagga", "mode": "0640", "owner": "quagga", "path": "/etc/qua
gga/Quagga.conf", "size": 839, "state": "file", "uid": 104}
ok: [accton02] => {"changed": false, "gid": 109, "group": "quagga", "mode": "0640", "owner": "quagga", "path": "/etc/qua
gga/Quagga.conf", "size": 839, "state": "file", "uid": 104}

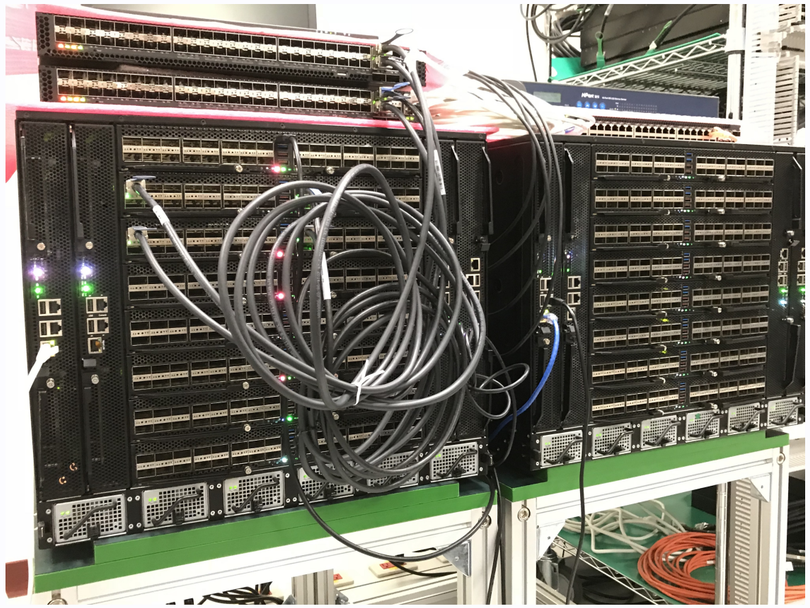
PLAY RECAP *********************************************************************
accton01                   : ok=27   changed=11   unreachable=0    failed=0
accton02                   : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_fab01              : ok=26   changed=11   unreachable=0    failed=0
fb_bp01_fab02              : ok=26   changed=11   unreachable=0    failed=0
fb_bp01_fab03              : ok=26   changed=11   unreachable=0    failed=0
fb_bp01_fab04              : ok=26   changed=11   unreachable=0    failed=0
fb_bp01_lc101              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc102              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc201              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc202              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc301              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc302              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc401              : ok=27   changed=11   unreachable=0    failed=0
fb_bp01_lc402              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_fab01              : ok=26   changed=11   unreachable=0    failed=0
fb_bp02_fab02              : ok=27   changed=13   unreachable=0    failed=0
fb_bp02_fab03              : ok=27   changed=13   unreachable=0    failed=0
fb_bp02_fab04              : ok=27   changed=13   unreachable=0    failed=0
fb_bp02_lc101              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc102              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc201              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc202              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc301              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc302              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc401              : ok=27   changed=11   unreachable=0    failed=0
fb_bp02_lc402              : ok=27   changed=11   unreachable=0    failed=0

                           ~/git/backpack-ansible  master
```

OMP800

OMP800

- OMP800を見に台湾(Accton/Edgecore)へ行ってきました
- Backpackとほぼ同じアーキテクチャ
- ポート密度が128から256ポートになっている







Special Thanks to Taiwan with Accton/Edgecore/Cumulus/CTC







Thank you for your kind attention!

Yahoo! JAPAN Site Operations