

# ISIS はじめの一歩 0から1へ

Shishio Tsuchiya  
[shtsuchi@arista.com](mailto:shtsuchi@arista.com)

# はじめに

- IS-ISはOSPFと並ぶリンクステートプロトコルだが、日本ではデプロイメントも少ない
- 一方でIPv6化のタイミングでIS-IS化するケースも増えてきた。
- 本セッションではまずIS-ISの0知識から1に行って頂きその後、レベルをあげる事を目的とする

<https://tools.ietf.org/html/draft-matsushima-spring-srv6-deployment-status-00#section-2.1>

2. Deployment Status

2.1. Softbank

As part of the 5G rollout, Softbank have deployed a nationwide SRv6 network.

The following SRv6 features have been deployed:

- A Segment Routing Header [[I-D.ietf-6man-segment-routing-header](#)] based data plane.
- END (PSP), END.X (PSP), END.DT4, T.Encaps.Red and T.Insert.Red functions as per [[I-D.ietf-spring-srv6-network-programming](#)].
- ISIS SRv6 extensions [[I-D.bashandy-isis-srv6-extensions](#)].
- BGP VPN SRv6 extensions [[I-D.dawra-bess-srv6-services](#)].
- Topology Independent (TI-LFA) Fast Reroute mechanisms leveraging SRv6 for the O(50msec) protection against node and link, as described in [[I-D.ietf-rtgwg-segment-routing-ti-lfa](#)].
- BGP Prefix Independent Convergence (PIC) core and edge [[I-D.ietf-rtgwg-bgp-pic](#)].
- Support for Ping and Traceroute as defined in [[I-D.ali-6man-spring-srv6-oam](#)].

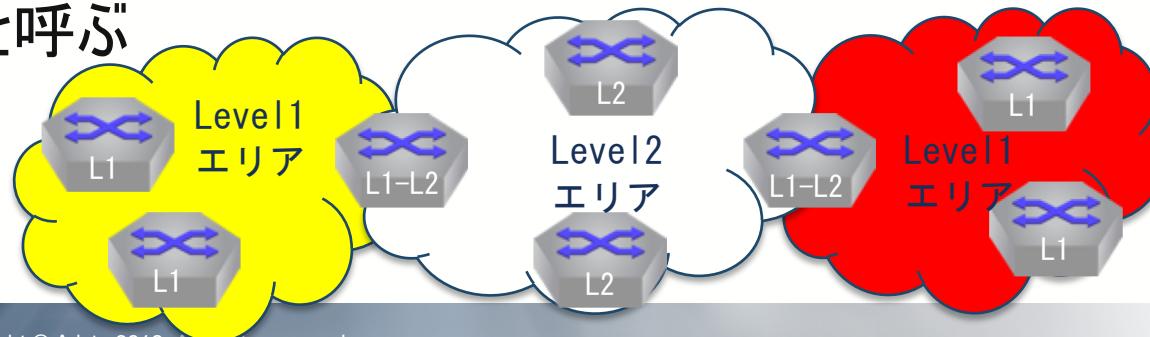
2.2. China Telecom

# Agenda

- IS-ISの基礎
- IS-ISのこれから

# IS-IS(Intermediate System to Intermediate System)概要

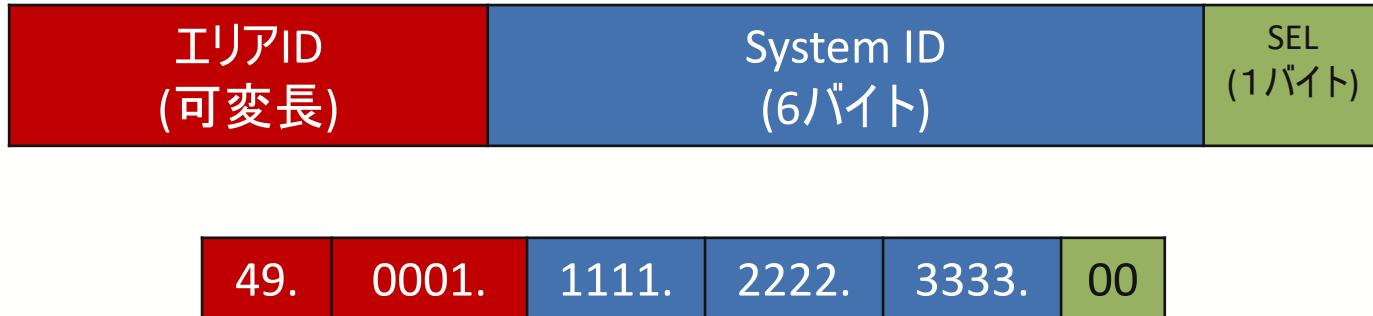
- IS-ISはOSPFと同様なリンクステートルーティングプロトコル
- リンクステートデータベース(LSDB)より最短経路(SPF)を計算
- 同一エリアはレベル1、バックボーンエリア(同一エリアおよび異なるエリア間)はレベル2の隣接関係を持つ
- レベル1に属するルータをレベル1(L1)ルータ、レベル2に属するルータをレベル2(L2)ルータ、両方に属するルータレベル1-レベル2(L1-L2)ルータと呼ぶ



# IS-ISヒストリー

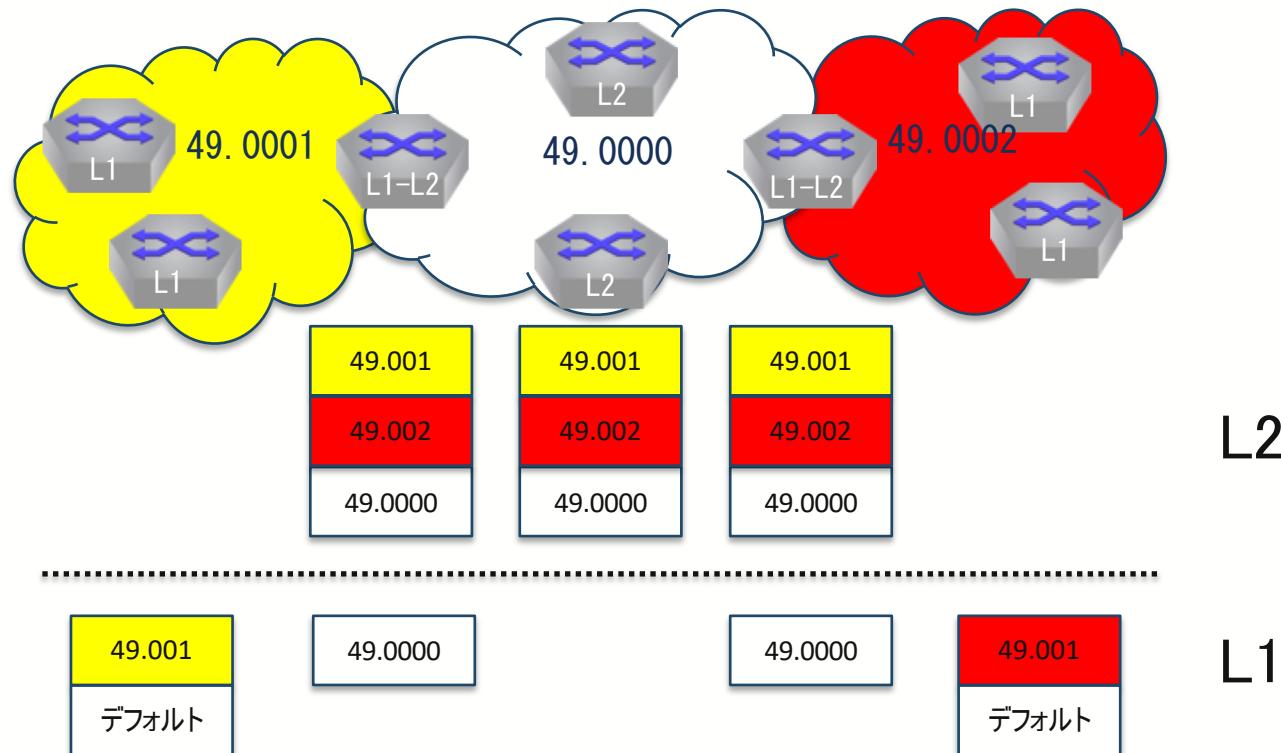
- IS-ISはOSI(Open Systems Interconnection)にてCLNS(Connectionless Network Service)のIGPとして制定
  - ISO10589
    - » <ftp://ftp.iks-jena.de/mitarb/lutz/standards/iso/iso10589v2.pdf>
  - OSI IS-IS Intra-domain Routing Protocol
    - » <https://tools.ietf.org/html/rfc1142>
- IETFでTCP/IP用に拡張
  - Use of OSI IS-IS for Routing in TCP/IP and Dual Environments
    - » <https://tools.ietf.org/html/rfc1195>

# IS-ISネットワークアドレス



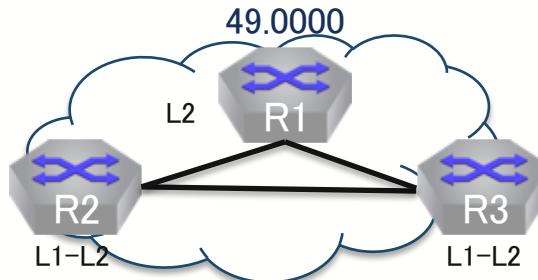
- IS-ISの識別子としてNET(Network Entity Title)を指定する
- NETは可変長のエリアID、6バイトのシステムID、1バイトのセクターIDから成り立つ
- エリアIDのAFIはドメイン識別子に使われる 0x49はプライベート
- システムIDはIPアドレスやMACアドレスが使用される
- SELは通常0x00を使用する

# IS-ISレベルルーティング



- レベル1/2ルータは他のエリアの情報をL1ルータには配布せず、デフォルトルートとして配布する

# IS-ISルータ設定



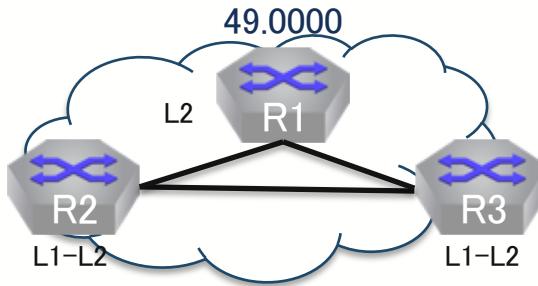
```
R2#show running-config section router isis
router isis JANOG44
  net 49.0000.0000.0000.0002.00
  !
  address-family ipv4 unicast
  !
  address-family ipv6 unicast
R2#
```

```
R3#show running-config | section router isis
router isis JANOG44
  net 49.0000.0000.0000.0003.00
  log-adjacency-changes
  !
  address-family ipv4 unicast
  !
  address-family ipv6 unicast
R3#
```

```
R1#show running-config section router isis
router isis JANOG44
  net 49.0000.0000.0000.0001.00
  is-type level-2
  !
  address-family ipv4 unicast
  !
  address-family ipv6 unicast
R1#
```

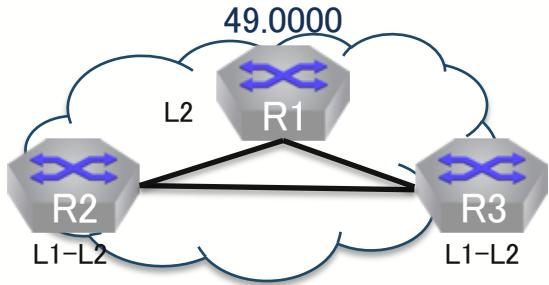
# IS-ISルータ設定

```
R2(config-router-isis)#is-type ?
level-1  Configure at level 1
level-1-2 Configure at level 1-2
level-2  Configure at level 2
R2(config-if-Et1)#isis circuit-type ?
level-1  Configure at level 1
level-1-2 Configure at level 1-2
level-2  Configure at level 2
```



- デフォルトでLevel 1-2で動作する
- グローバル設定やインターフェース指定でLevel設定を変える事が出来る

# IS-ISルータ設定

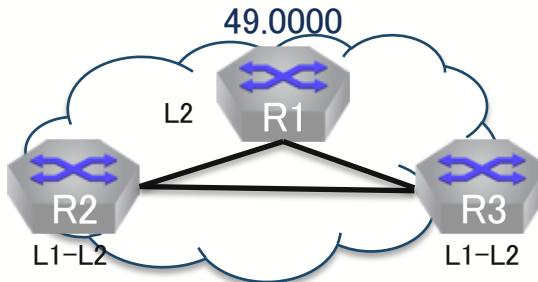


```
R2#show isis neighbors
```

| Instance | VRF     | System Id | Type | Interface | SNPA            | State | Hold time | Circuit Id |
|----------|---------|-----------|------|-----------|-----------------|-------|-----------|------------|
| JANOG44  | default | R1        | L2   | Ethernet1 | P2P             | UP    | 22        | 07         |
| JANOG44  | default | R3        | L1   | Ethernet2 | 8:0:27:74:d8:fc | UP    | 24        | R2.08      |
| JANOG44  | default | R3        | L2   | Ethernet2 | 8:0:27:74:d8:fc | UP    | 26        | R2.08      |
| R2#      |         |           |      |           |                 |       |           |            |

- L1-2ルータで動作し、対向もL1-2ルータの場合には2つの隣接関係を確立する

# IS-ISルータ設定



- ネットワークタイプはPoint-to-PointとBroadcastがある
- Broadcastの場合にはOSPFのDRと同様にDISが存在する

```
R2#show isis interface ethernet 1
```

IS-IS Instance: JAN0G44 VRF: default

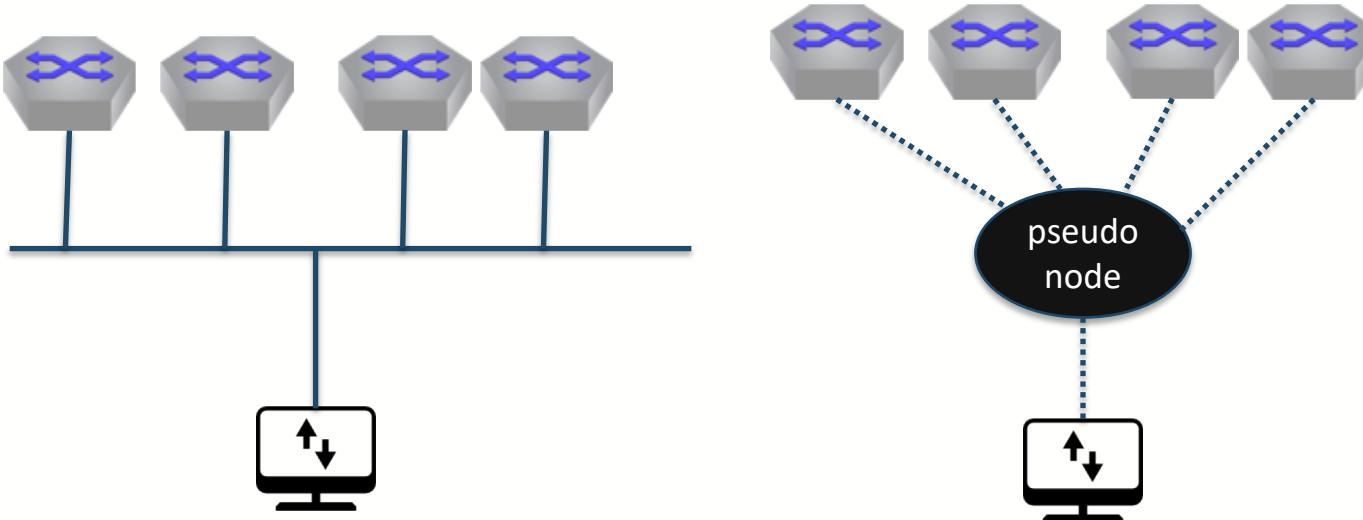
Interface Ethernet1:  
Index: 7 SNPA: P2P  
**MTU: 1497 Type: point-to-point**  
BFD IPv4 is Disabled  
BFD IPv6 is Disabled  
Hello Padding is Enabled  
Level 1:  
Metric: 10, Number of adjacencies: 0  
Link-ID: 07  
Authentication mode: None  
Level 2:  
Metric: 10, Number of adjacencies: 1  
Link-ID: 07  
Authentication mode: None

```
R2#show isis interface ethernet 2
```

IS-IS Instance: JAN0G44 VRF: default

Interface Ethernet2:  
Index: 8 SNPA: 8:0:27:b7:85:9a  
**MTU: 1497 Type: broadcast**  
BFD IPv4 is Disabled  
BFD IPv6 is Disabled  
Hello Padding is Enabled  
Level 1:  
Metric: 10, Number of adjacencies: 1  
LAN-ID: R2.08, Priority: 64  
DIS: R2, DIS Priority: 64  
Authentication mode: None  
Level 2:  
Metric: 10, Number of adjacencies: 1  
LAN-ID: R2.08, Priority: 64  
DIS: R2, DIS Priority: 64  
Authentication mode: None

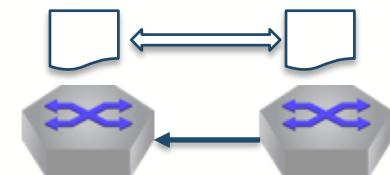
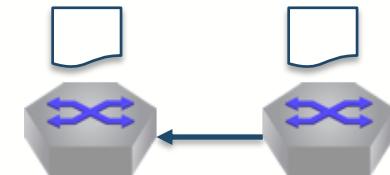
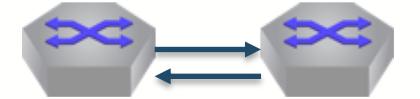
# IS-ISプロードキャストでの振る舞い



- ブロードキャストの場合はそれ自体をPseudonodeと見立て、各IS/ESともに接続されていると考える
- またPseudonodeに代わってLink State PDUを送出するDISを選出する
- 選出方法はプライオリティもしくはMACアドレス

# IS-IS Protocol Data Unit(PDU) Type

- Hello PDU
  - ハロー・パケットは隣接ルータ間の隣接関係を初期化し維持するために使用される
    - » Level 1 LAN IS to IS Hello PDU (Type 15)
    - » Level 2 LAN IS to IS Hello PDU (Type 16)
    - » Point-to-Point IS to IS Hello PDU (Type 17)
- Link State PDU
  - リンクステートパケット (LSP) はリンクステート情報の交換に使用される
    - » Level 1 Link State PDU (Type 18)
    - » Level 2 Link State PDU (Type 20)
- Sequence Numbers PDU
  - シーケンス番号PDU (SNP) は、各ルータから送信される最新の LSP の情報を隣接ルータ同士で確実に一致させるために使用される
    - » Level 1 Complete Sequence Numbers PDU (Type 24)
    - » Level 2 Complete Sequence Numbers PDU (Type 25)
    - » Level 1 Partial Sequence Numbers PDU (Type 26)
    - » Level 2 Partial Sequence Numbers PDU (Type 27)
- L1は01-80-C2-00-00-14/L2は01-80-C2-00-00-15を使用する



# show isis neighbors detail

```
R2#show isis neighbors detail
```

| Instance                          | VRF     | System Id | Type | Interface | SNPA | State | Hold time | Circuit Id |
|-----------------------------------|---------|-----------|------|-----------|------|-------|-----------|------------|
| JANOG44                           | default | R1        | L2   | Ethernet1 | P2P  | UP    | 26        | 07         |
| Area Address(es) : 49.0000        |         |           |      |           |      |       |           |            |
| SNPA: P2P                         |         |           |      |           |      |       |           |            |
| Router ID: 0.0.0.0                |         |           |      |           |      |       |           |            |
| Advertised Hold Time: 30          |         |           |      |           |      |       |           |            |
| State Changed: 00:01:38 ago       |         |           |      |           |      |       |           |            |
| IPv4 Interface Address: 10.10.1.0 |         |           |      |           |      |       |           |            |
| IPv6 Interface Address: fe80::1   |         |           |      |           |      |       |           |            |
| Interface name: Ethernet1         |         |           |      |           |      |       |           |            |
| Graceful Restart: Supported       |         |           |      |           |      |       |           |            |
| BFD IPv4 state is Up              |         |           |      |           |      |       |           |            |

Dynamic Hostname Exchange [RFC5301](#)

Routing IPv6 with IS-IS [RFC5308](#)

Restart Signaling for IS-IS [RFC5306](#)

IS-IS BFD-Enabled TLV [RFC6213](#)

- BFD/Graceful Restart/IPv6/Hostname交換など多くの機能を早く実装

# ISISデータベース

```
R2#show isis database
```

IS-IS Instance: JANOG44 VRF: default

IS-IS Level 1 Link State Database

| LSPID           | Seq Num   | Cksum       | Life        | IS Flags    |
|-----------------|-----------|-------------|-------------|-------------|
| R2.00-00        | 59        | 20333       | 600         | L2 ◇        |
| <b>R2.08-00</b> | <b>51</b> | <b>1672</b> | <b>1094</b> | <b>L2 ◇</b> |
| R3.00-00        | 55        | 39704       | 724         | L2 ◇        |

IS-IS Level 2 Link State Database

| LSPID           | Seq Num   | Cksum       | Life        | IS Flags    |
|-----------------|-----------|-------------|-------------|-------------|
| R1.00-00        | 77        | 61366       | 1105        | L2 ◇        |
| R2.00-00        | 78        | 1766        | 536         | L2 ◇        |
| <b>R2.08-00</b> | <b>50</b> | <b>2183</b> | <b>1178</b> | <b>L2 ◇</b> |
| R3.00-00        | 79        | 47656       | 569         | L2 ◇        |

```
R2#
```

```
R2#show isis database R1.00-00 detail
```

IS-IS Instance: JANOG44 VRF: default

IS-IS Level 1 Link State Database

| LSPID | Seq Num | Cksum | Life | IS Flags |
|-------|---------|-------|------|----------|
|-------|---------|-------|------|----------|

IS-IS Level 2 Link State Database

| LSPID | Seq Num | Cksum | Life | IS Flags |
|-------|---------|-------|------|----------|
|-------|---------|-------|------|----------|

|          |    |       |     |      |
|----------|----|-------|-----|------|
| R1.00-00 | 77 | 61366 | 873 | L2 ◇ |
|----------|----|-------|-----|------|

Remaining lifetime received: 1199 s Modified to: 1200 s

NLPID: 0xCC(IPv4) 0x8E(IPv6)

Hostname: R1

Area address: 49.0000

Interface address: 10.10.1.0

Interface address: 10.255.255.1

Interface address: 10.10.1.2

Interface address: 2001:db8:ffff::1

IS Neighbor : R3.00 Metric: 10

IS Neighbor : R2.00 Metric: 10

Reachability : 10.10.1.0/31 Metric: 10 Type: 1 Up

Reachability : 10.255.255.1/32 Metric: 10 Type: 1 Up

Reachability : 10.10.1.2/31 Metric: 10 Type: 1 Up

Reachability : 2001:db8:ffff::1/128 Metric: 10 Type: 1 Up

Router Capabilities: 10.10.1.2 Flags: [ ]

```
R2#
```

- Level1/Level2のデータベース/ Pseudo nodeの確認
- IPv4/IPv6ともに一つのルーティングインスタンスで実現可能

# ISIS PDU

## ISIS HELLO

.... ..11 = Circuit type: Level 1 and 2 (0x3)  
0000 00.. = Reserved: 0x00  
SystemID {Sender of PDU}: 0000.0000.0002  
Holding timer: 30  
PDU length: 1497  
Local circuit ID: 7  
Protocols Supported (t=129, l=2)  
    Type: 129  
    Length: 2  
    NLPIID(s): IPv6 (0x8e), IP (0xcc)  
Area address(es) (t=1, l=4)  
    Type: 1  
    Length: 4  
    Area address (3): 49.0000  
IP Interface address(es) (t=132, l=4)  
    Type: 132  
    Length: 4  
    IPv4 interface address: 10.10.1.1  
IPv6 Interface address(es) (t=232, l=16)  
    Type: 232  
    Length: 16  
    IPv6 interface address: fe80::2

## Restart Signaling (t=211, l=1)

Type: 211  
Length: 1  
Restart Signaling Flags: 0x00  
Point-to-point Adjacency State (t=240, l=15)  
    Type: 240  
    Length: 15  
    Adjacency State: Up (0)  
    Extended Local circuit ID: 0x00000007  
    Neighbor SystemID: 0000.0000.0001  
    Neighbor...Extended Local circuit ID: 0x00000007  
Padding (t=8, l=255)  
Padding (t=8, l=136)

MTU最大までPaddingを行い、  
Helloの時点でMTU  
不一致を検出する

- TLVフォーマットで各パラメーターを送信する

# SPF log

```
R1#show isis spf log
IS-IS Instance: JAN0G44 VRF: default
IS-IS SPF Log level-2
* - Partial SPF
Start time          Duration(msec)  Nodes   Count  Reason
2019-07-20 20:01:36.572814  0.092      4       3      LSP received R3. 00-00
*2019-07-20 20:01:27.688379  0.095      3       2      LSP received R2. 00-00
*2019-07-20 20:01:16.544109  1.792      3       2      LSP received R2. 00-00
2019-07-20 20:00:29.808715  0.138      3       5      Purge LSP received R2. 08-00
2019-07-20 19:58:36.876805  0.092      4       1      Periodic LSP refresh R1. 00-00
2019-07-20 19:43:52.760987  0.089      4       1      Periodic LSP refresh R1. 00-00
*2019-07-20 19:40:36.738546  0.068      4       2      LSP received R2. 00-00
2019-07-20 19:40:31.611772  0.346      4       3      Adjacency R3 came up on Ethernet2
2019-07-20 19:40:29.196412  0.250      4       5      LSP received R2. 08-00
2019-07-20 19:40:26.966404  0.654      2       5      Purge LSP received R2. 08-00
2019-07-20 19:31:59.021749  0.243      4       1      Periodic LSP refresh R1. 00-00
2019-07-20 19:17:06.763560  0.096      4       1      Periodic LSP refresh R1. 00-00
2019-07-20 19:05:07.682225  0.090      4       1      Periodic LSP refresh R1. 00-00
2019-07-20 18:50:42.523430  0.120      4       1      Periodic LSP refresh R1. 00-00
2019-07-20 18:50:19.394209  0.116      4       3      LSP received R2. 08-00
*2019-07-20 18:50:17.389263  0.098      3       2      LSP received R2. 00-00
2019-07-20 18:50:16.384955  2.400      3       6      Adjacency R3 came up on Ethernet2
2019-07-20 18:49:55.153795  6.181      2       4      Adjacency R2 came up on Ethernet1
2019-07-20 18:49:19.671999  0.298      1       2      Adjacency R2 went down on Ethernet1
2019-07-20 18:48:35.411313  0.405      2       3      LSP received R2. 00-00
```

R1#

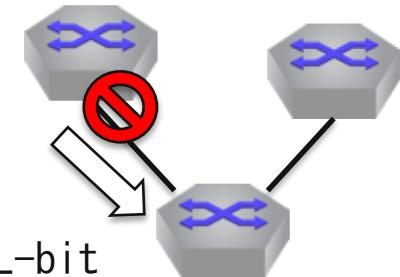
- LSPの最大ライフ時間は1200秒で、900秒おきにLSPリフレッシュする

# IS-IS Transient Blackhole Avoidance(Overload bit) [RFC3277](#)

```
R1(config-router-isis)#set-overload-bit ?  
on-startup Overload at startup for <time> seconds or wait for BGP convergence  
<cr>  
R1(config-router-isis)#set-overload-bit
```

```
R2#show isis database
```

```
IS-IS Instance: JANOG44 VRF: default  
IS-IS Level 1 Link State Database  
LSPID Seq Num Cksum Life IS Flags  
R2. 00-00 70 2472 856 L2 ◇  
R2. 08-00 57 63886 998 L2 ◇  
R3. 00-00 66 34083 722 L2 ◇  
IS-IS Level 2 Link State Database  
LSPID Seq Num Cksum Life IS Flags  
R1. 00-00 85 58294 1161 L2 <DBOverload>  
R2. 00-00 93 46886 791 L2 ◇  
R2. 08-00 56 64397 889 L2 ◇  
R3. 00-00 96 38969 568 L2 ◇
```

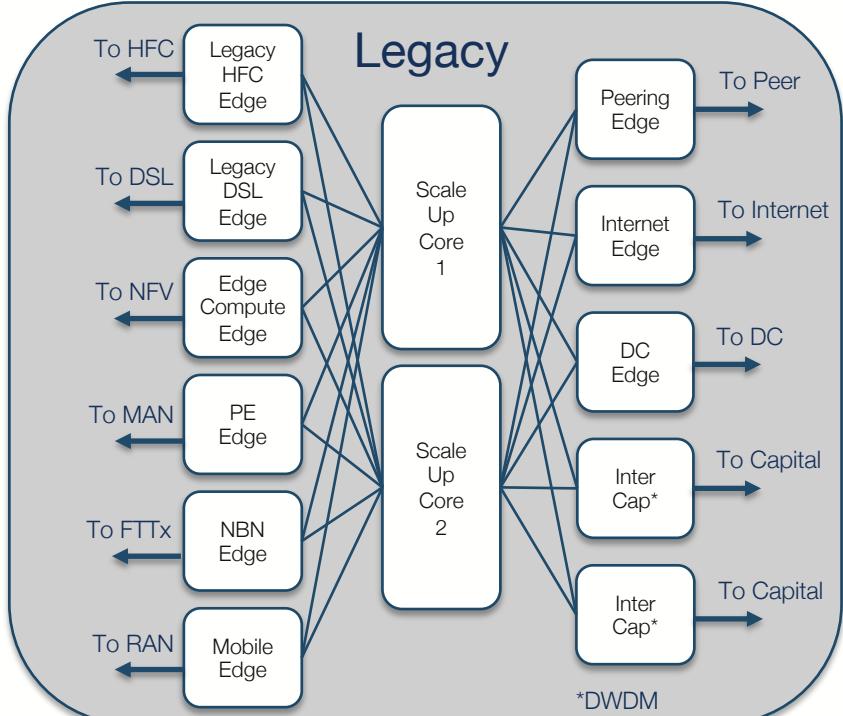


- overloadビットを付与し、迂回に使う事が出来る  
(OSPFのmax-metric 同様)

# Agenda

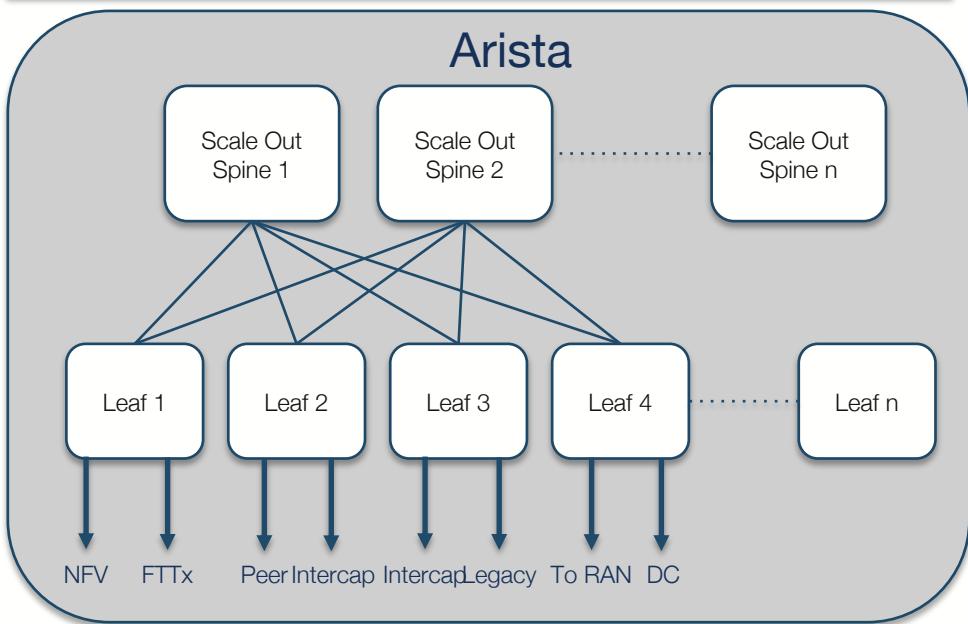
- IS-ISの基礎
- IS-ISのこれから

# スケールアップ vs スケールアウト



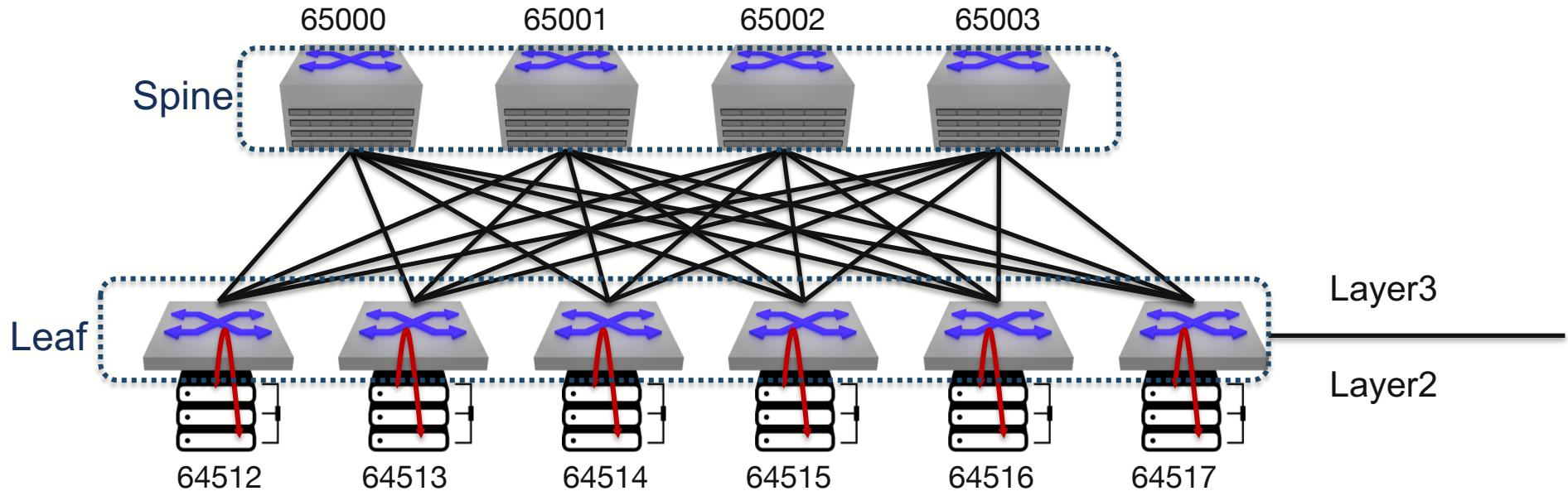
最大スケールの一つのコアルータによりサイトの収容能力は制限  
メンテナンスの為にコア/エッジルータをサービス中断するのは困難  
特定用途目的のエッジルータはCAPEXとコアポート数を増加させる

スケールアウトは現存のキャリアトポロジーやトラフィックフローから根本的な変化では無く、効率的



柔軟なサイト収容能力 - Spine/Leadを追加する  
サービス化とアップグレードの容易に - ノード毎追加  
スケールアウトはCAPEXを削減し、エッジを収束し、選択肢を提供

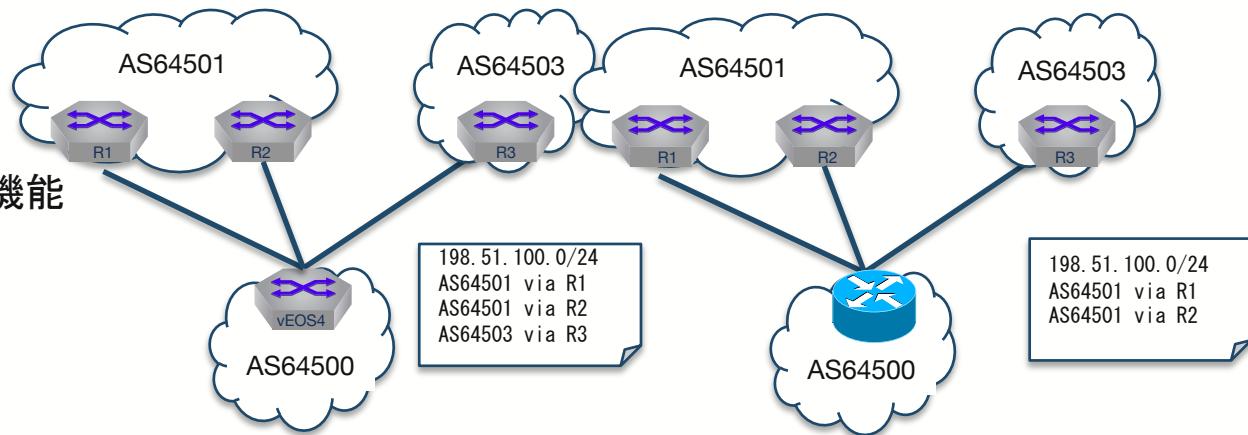
# 大規模データセンタールーティングでのBGPの使用



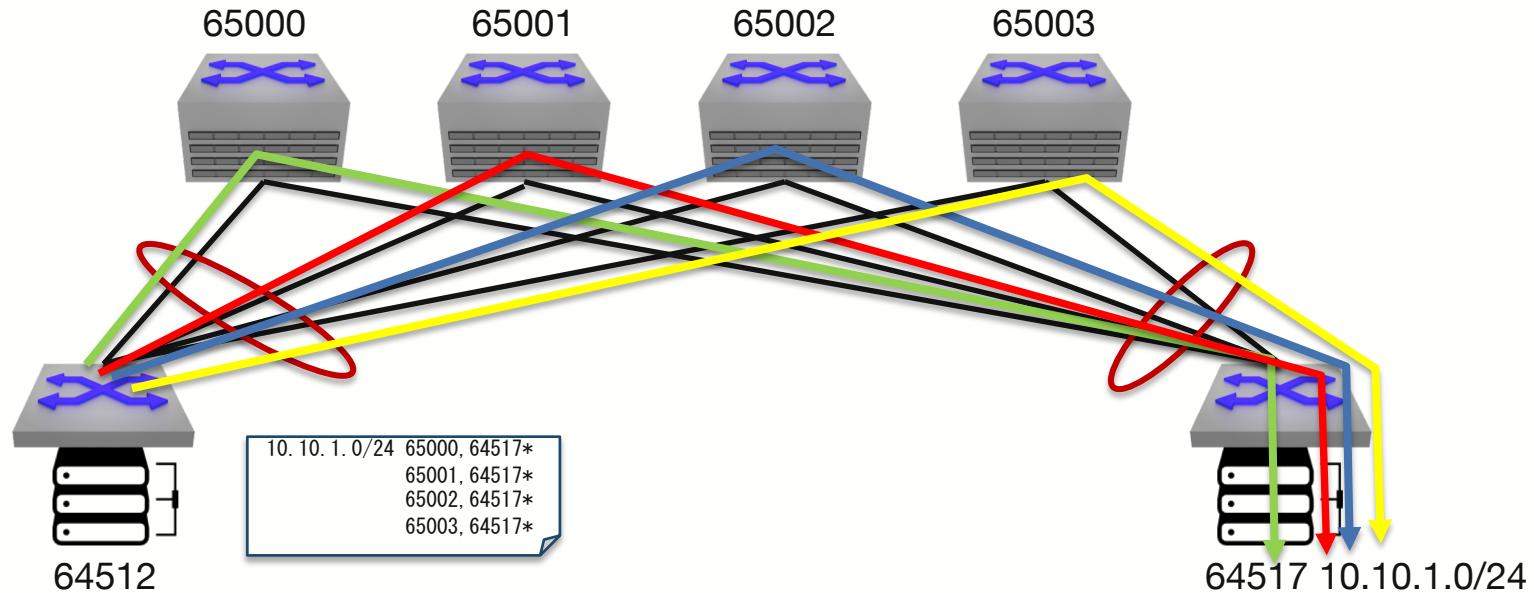
- スケールアウトするClosデザイン
- 安定した標準的なBGPプロトコルをToR/Leafスイッチに使用
- 安定性にフォーカスし、VMモビリティはラック内に留める

# このネットワークデザインにおける特別なBGPの要件

- AS\_PATH Multipath Relax
- Allow AS In
  - 重複したAS番号を使用する為の機能
- Fast eBGP Fall-over
  - 高速コンバージェンスを可能にする
- Remove Private AS
  - 2バイトのプライベートAS番号(64512-65534)だけでなく、4バイトのプライベートAS番号(4200000000 - 4294967294)もエッジで除去
  - [RFC6996 Autonomous System \(AS\) Reservation for Private Use](#)

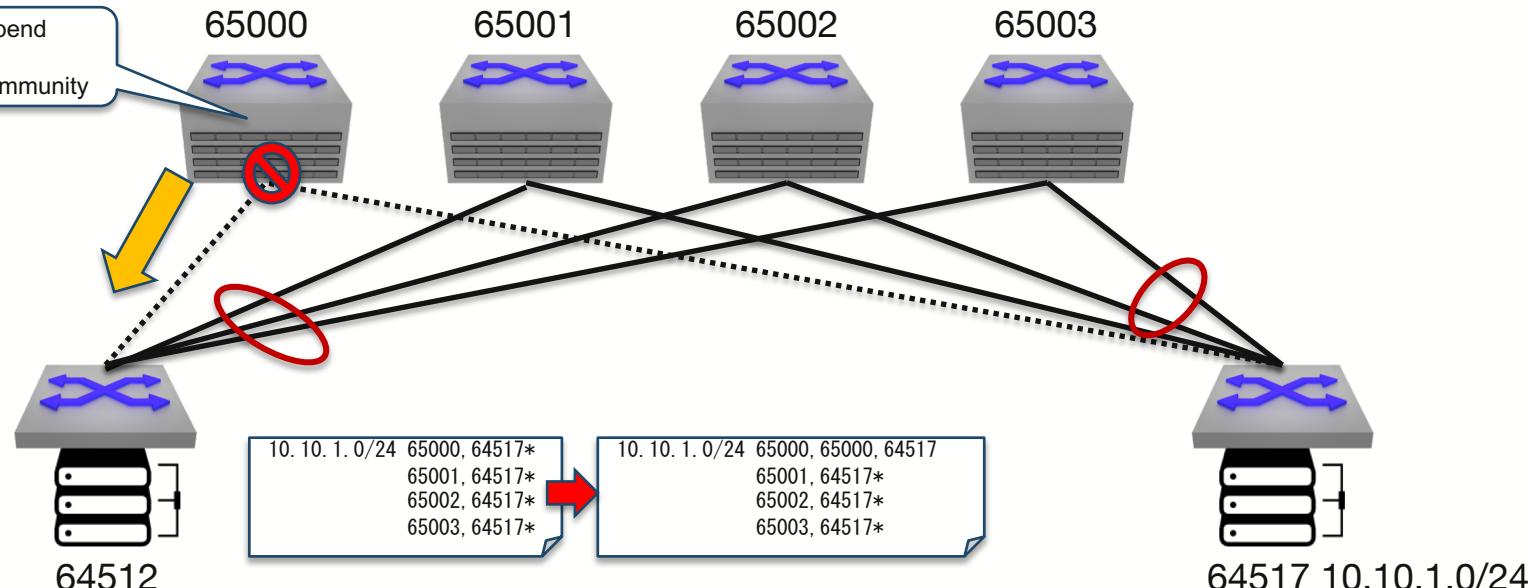


# IP ECMPの利点



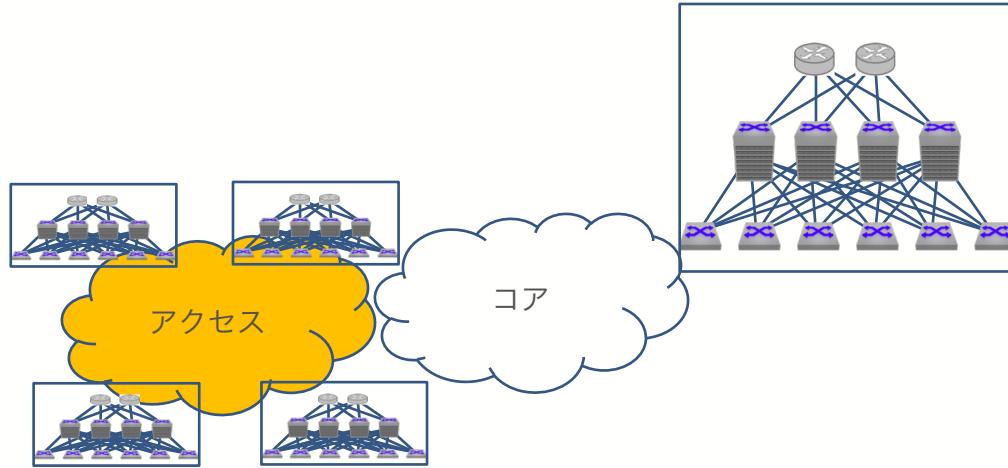
- ECMPで全てのLeaf-Spineリンクを使用する
- それぞれのフローはECMPハッシュにてバランスされる(既に実装されている)
- ルーティングパスはBGPパス属性により可視化される

# BGPの利点



- ASパスを追加するもしくはBGPグレースフルシャットダウンコミュニティ([GSHUT \(0xFFFF0000\) community](#))を使って簡単にメンテナンス可能

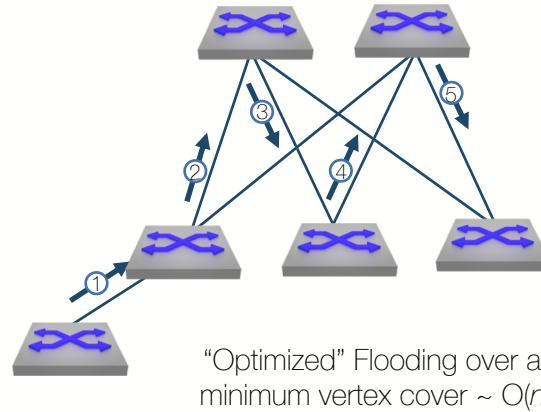
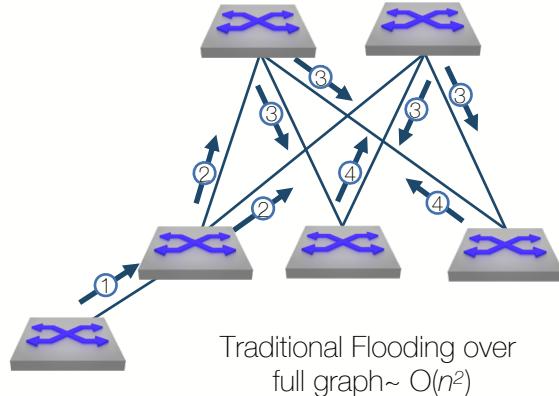
# 次世代ネットワークトポロジー



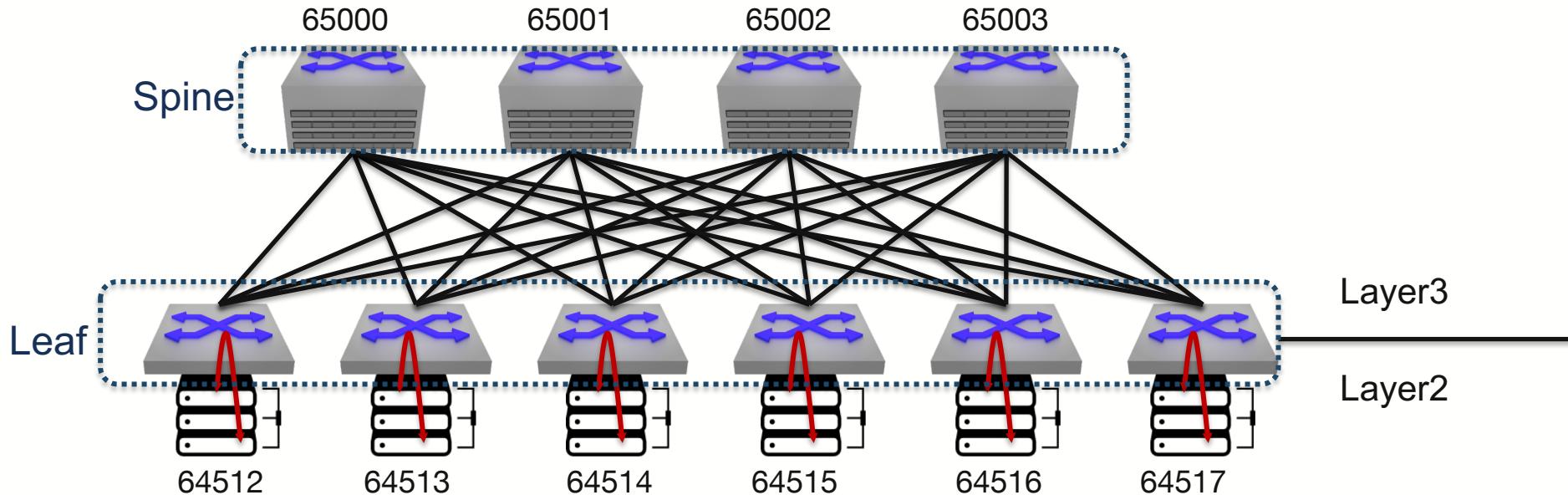
- マイクロデータセンターのようにスケールアウトのデザインが多くの場所で使用される様になる

# リンクステートプロトコルの課題

- なぜリンクステートプロトコルがデータセンターのような密集したトポロジーで使われないのか？
- リンクステートプロトコルはトポロジー計算を完璧に実施する
- データセンタートポロジーではあらゆる場所でSPF計算が実施される事になる
- この問題は90年代のATMのネットワークでもよく論じられた
  - IS-IS Mesh Groups <https://tools.ietf.org/html/rfc2973>

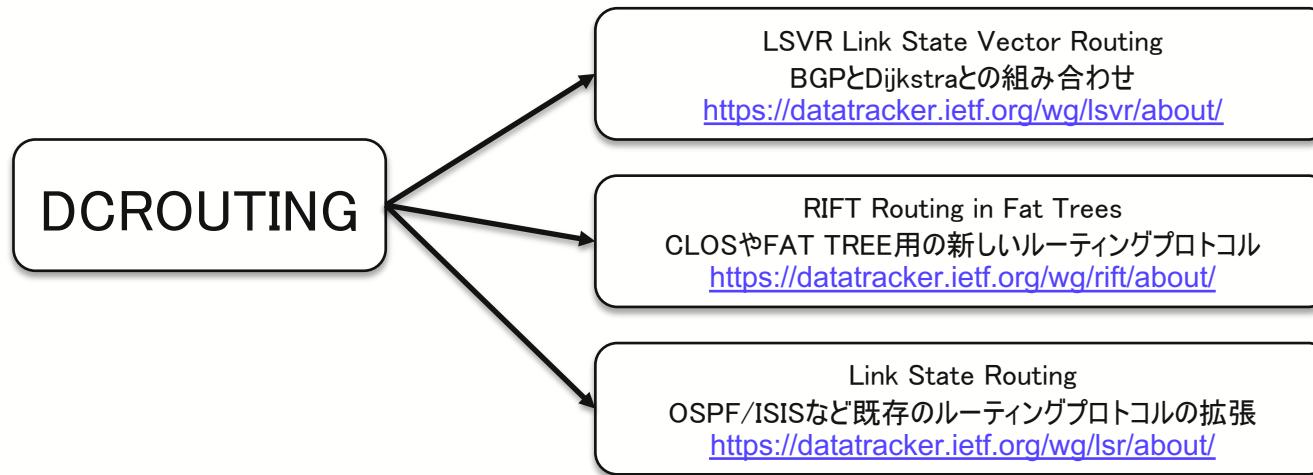


# BGPの課題



- 自動的なノード発見のメカニズムが無い
- リンクステート情報が無い。
  - 次世代マルチキャスト(BIER)やトラフィックエンジニアリング(SR/RSVP-TE)の詳細情報が伝えづらい

# 今後のデータセンタールーティング標準化

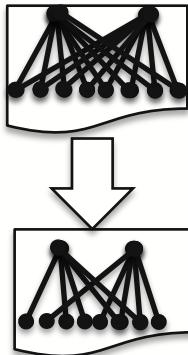
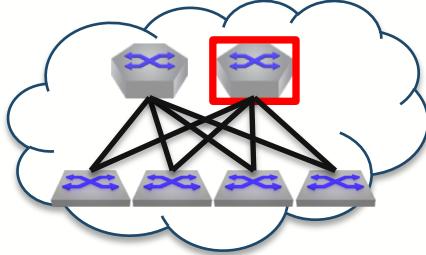


- 既存のパスベクタープロトコルにDijkstraアルゴリズムを組み合わせたLSVR
- データセンタートポロジー用に新しいルーティングプロトコルを使うRIFT
- OSPF/ISISの拡張としてDCおよび様々なトポロジーの拡張を扱うLSR
- それぞれがWGとして標準化および議論を進めていく

# ISIS拡張

- **Dynamic Flooding on Dense Graphs**
  - <https://tools.ietf.org/html/draft-ietf-lsr-dynamic-flooding>
- **Level 1 Area Abstraction for IS-IS**
  - <https://tools.ietf.org/html/draft-li-lsr-isis-area-abstraction>
- **Hierarchical IS-IS**
  - <https://tools.ietf.org/html/draft-li-lsr-isis-hierarchical-isis>
- フラッディングの改善およびArea抽象化/階層化(L3-L8)ISISを行えば、データセンターなどのスケーラビリティのあるサービスプロバイダースケールアウトデザインに答える事が出来る

# Dynamic Flooding とは

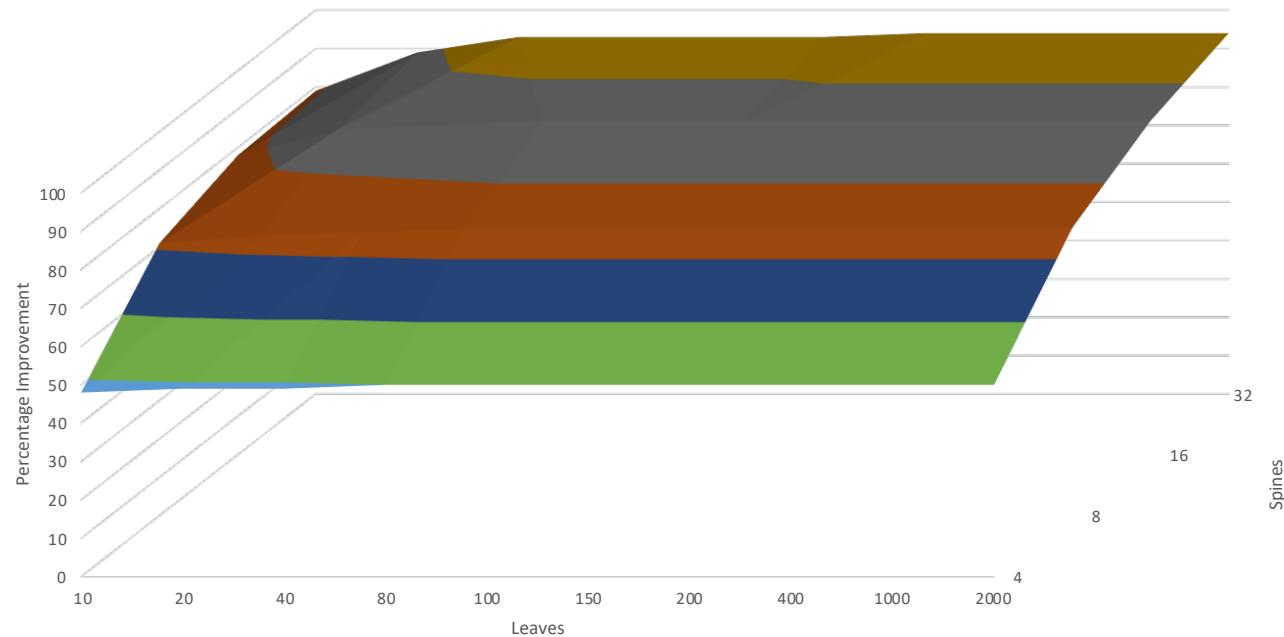


1. フラッディングトポロジーを作成する為のエリアリーダが選出される
  - 通常のリンクステートデータベースの一部として送出される
  - フラッディングトポロジーのみでフラッディングされる
  - フラッディングトポロジー以外では通常のデータベース同期(OSPF Database Exchange/ISIS CSNP)
2. フラッディングトポロジーの計算はトポロジーの計算とは別な安定してる時間に行う事が出来る

# シミュレーション結果

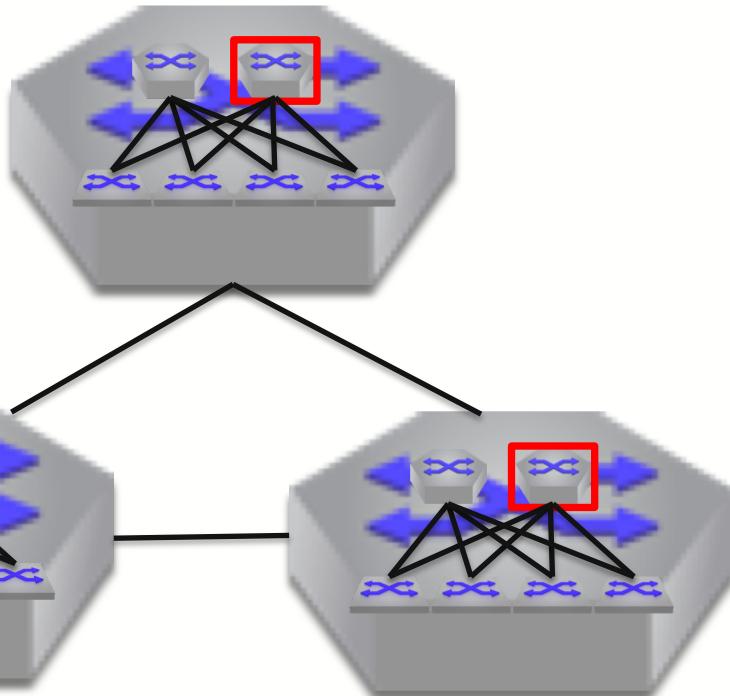
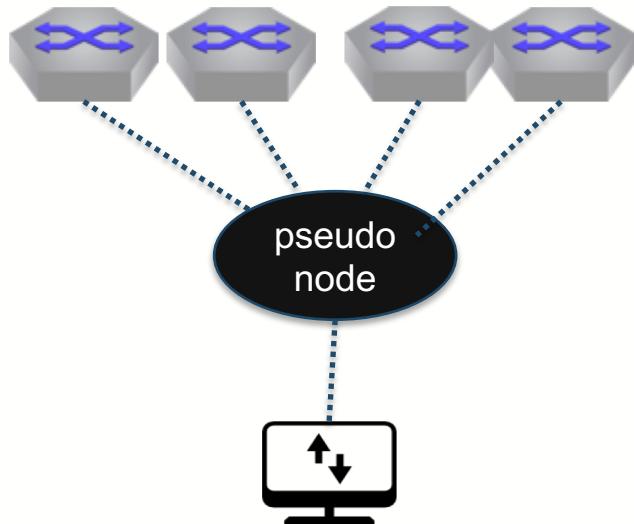
- 予想された様に Dynamic Flooding の 最適化で大量のフライングの改善が見られた
- leaf数やspine数が増加すると改善は95%の削減まで近づく

Dynamic Flooding Improvement for Leaf-Spine Topologies



# Area Abstraction(エリアの抽象化)

[draft-li-lsr-isis-area-abstraction](#)



- Level1を通過する為にはLevel2である必要がある
- スケールの為にデータセンター自体をLevel1とする
- エリアリーダはArea Pseudonode LSPをL2トポジーに送出する

# ISIS PDU/Hello

## ISIS HELLO

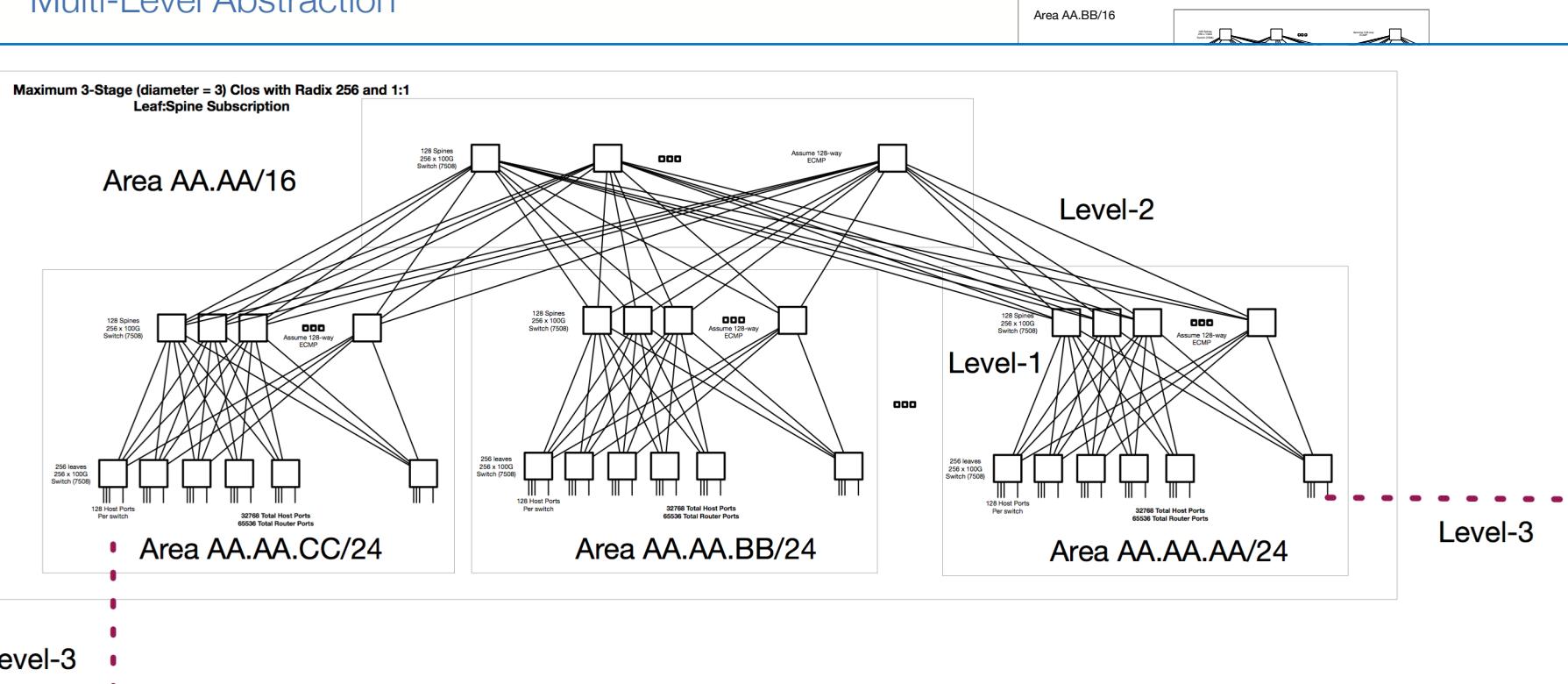
.... ..11 = Circuit type: Level 1 and 2 (0x3)  
0000 00.. = Reserved: 0x00

Level1 only 0x1  
Level2 only 0x2  
Level1-L2 0x3

- PDU/LSPにはCircuit Typeが定められている
  - 残り6ビットは予約されている
  - これを使うとレベル8まで拡張出来る
    - Hierarchical IS-IS
- » <https://tools.ietf.org/html/draft-li-lsr-isis-hierarchical-isis>

# Example of Scale Out Network Design

## Multi-Level Abstraction



# Area Hierarchy: IS-IS Multiple Levels 3-8

エリア自身を一つの”node”として抽象化する [draft-li-lsr-isis-hierarchical-isis](#)



IS-IS PDU/Helloヘッダーは抽象化の為に更に6つのビットが予約されている- とてつもないスケーラビリティを実現できる (新しくはなくそれはPNNIで使用された事がある)

エリアでフラッディング半径は区切られる- 各々のエリア階乗のスケーラビリティを表現できる

# まとめ

- IS-ISの現在の機能などを共有
- またデータセンターでBGPが使用するのはフラッディングの問題が大きい
- 下記の拡張がIGPおよびIS-ISで行われて、マルチトポロジーにスケーラビリティのあるネットワークへの提案がされている
  - Dynamic Flooding on Dense Graphs
  - Level 1 Area Abstraction for IS-IS
  - Hierarchical IS-IS

# Thank You

[www.arista.com](http://www.arista.com)