

BGP談義

～add-pathとdiverse-pathのお話～

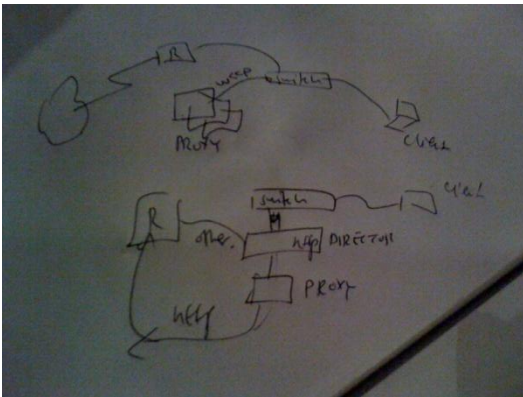
Shishio Tsuchiya

shtsuchi@cisco.com



そもそもBGPって

- [RFC1771](#)が1995年3月に発行
IBM:Y. Rekhter& Cisco Tony.Li
- [draft-ietf-idr-bgp4-01](#)～26まで改訂し、2006年1月に
[RFC4271](#)を発行した。



“The initial BGP design was recorded on a napkin rumored to have been heavily spattered with ketchup.”
A Border Gateway Protocol 4 (BGP-4)
January 1989 at the 12th IETF meeting

どこで決めてるの？

- IETFでは主に2つ

IDR(Inter Domain Routing Group)とGROW(Global Routing Operations)

IDR	Working Group	GROW
Routing	Area	Operation & Management
BGP4のIPv4/IPv6でBGPv4の頑健性/拡の利用のサポート 張性の改善	目的	BGP4のIPv4/IPv6での利用の際 の運用的問題の考慮
45個(Obsoleteも含む) Standardが中心	RFC	6個 BCPとInfo.

BGPの特徴

- [RFC4271の9.2. Update-Send Process](#)では他のBGPスピーカーに伝えるルートはディシジョンプロセスによる選ばれたものを伝えなければならない事を定義している。
- 全く同じアドレスPrefix/NLRIのもの複数のパスの広告は許されていない。
- 以前アドバタイズされたルートと同じNLRIをもつルートは暗黙のうちに取り換えられる。

バックアップパスがあればできる事

- ファーストコンバージェンス
バックアップパスがあれば重要障害時もローカルで切り替え可能
- マルチパスロードバランス
- 安定性と正確性の向上
MED Churn によるルートグラつき(oscillation) [RFC3345](#)

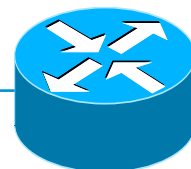
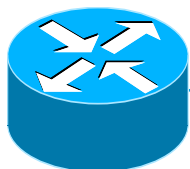
IETFでは現在add-path や diverse-pathというソリューションが議論され、標準化中

バックアップパスを伝える手法

- BGP add-path
- BGP Best External + Fullmesh
- BGP Diverse Path

add-pathって何だろう？

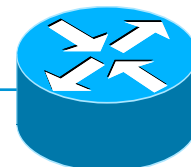
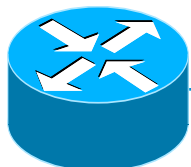
- 通常のBGP



10/8,AS_PATH:64500,64501,NH:203.0.113.10

10/8,AS_PATH:64510,64511,NH:203.0.113.10

- add-path



10/8,ID=1,AS_PATH:64500,64501,NH:203.0.113.10

10/8,ID=2,AS_PATH:64510,64511,NH:203.0.113.10

add-path詳細 Capability Exchange

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Send/Receive (1 octet)

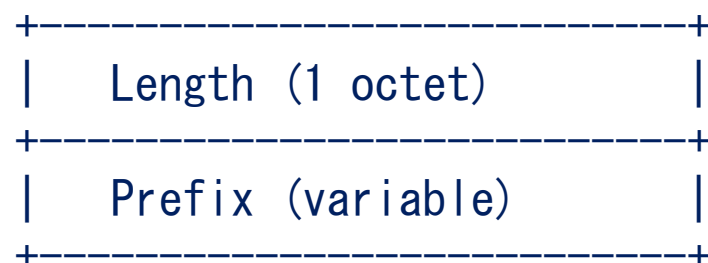
- BGPピアにてCapability Exchangeを行う([RFC3992](#))
Capability Valueは69
- Send/Receiveでマルチパスを
Receive(=1),Send(=2),Both(=3)可能かを示す

add-path詳細

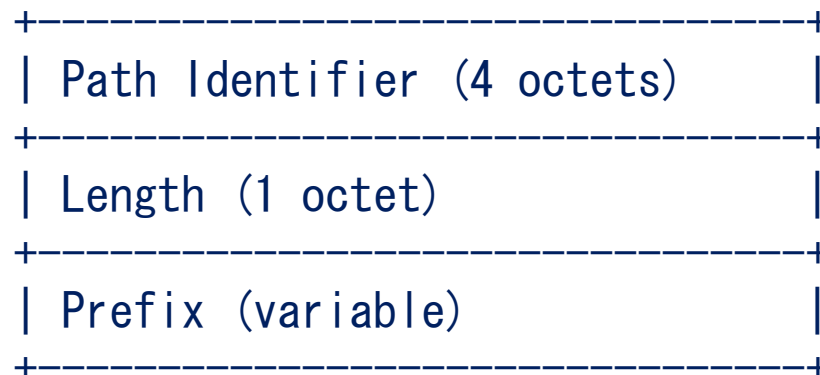
NLRI Encoding

NLRI Encoding

RFC4271, RFC4760



Extended NLRI Encoding

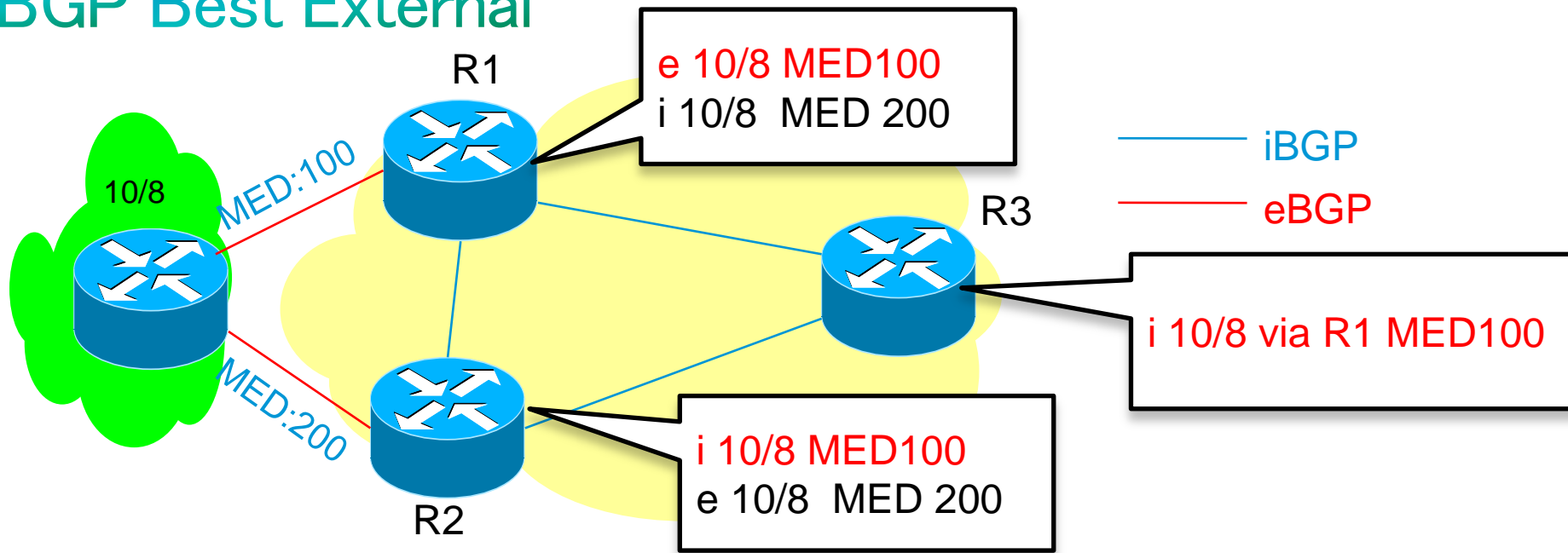


- UPDATEメッセージの中に4オクテットのPATH IDを含む
為に、Path Identifierフィールドをつける
- [RFC3107](#)でも同様にLengthの前部につける

バックアップパスを伝える手法

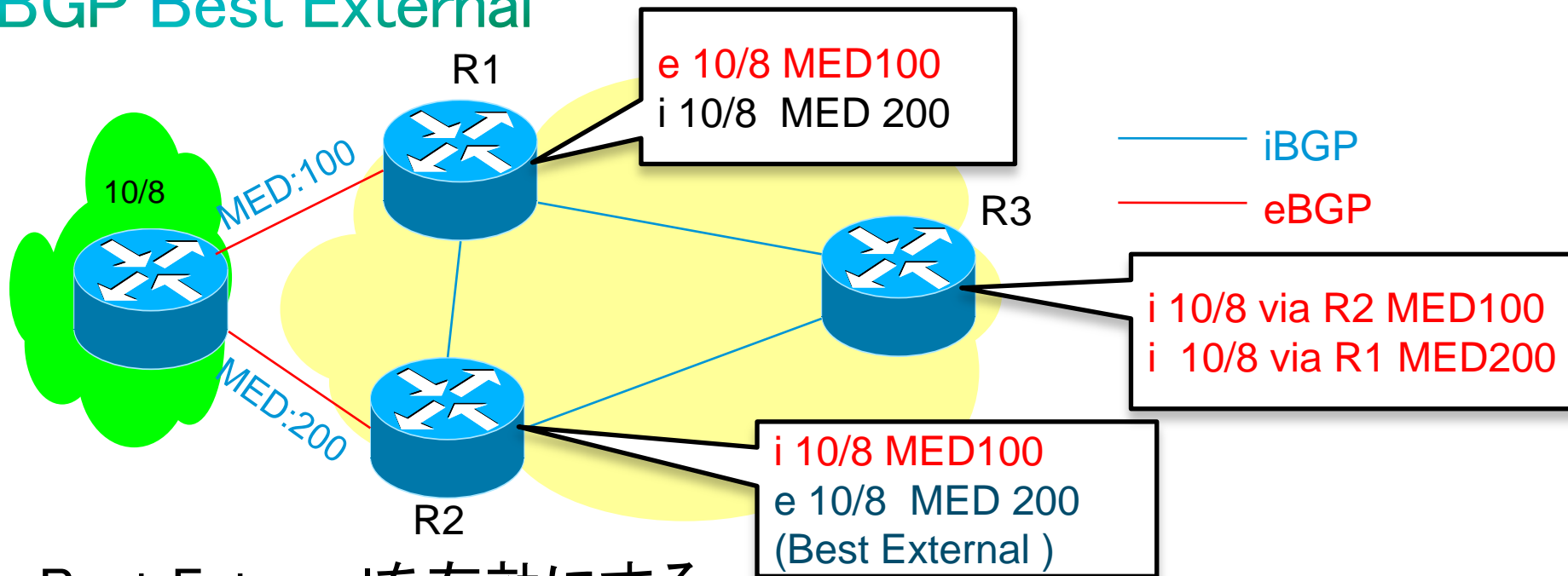
- BGP add-path
- **BGP Best External + Fullmesh**
- BGP Diverse Path

BGP Best External



- MEDの値により、R2ではいつもiBGPが選択される
- R3では代替えパスを持たない
- R1障害時にはWITHDRAWN/UPDATE、書き換えが行われる

BGP Best External

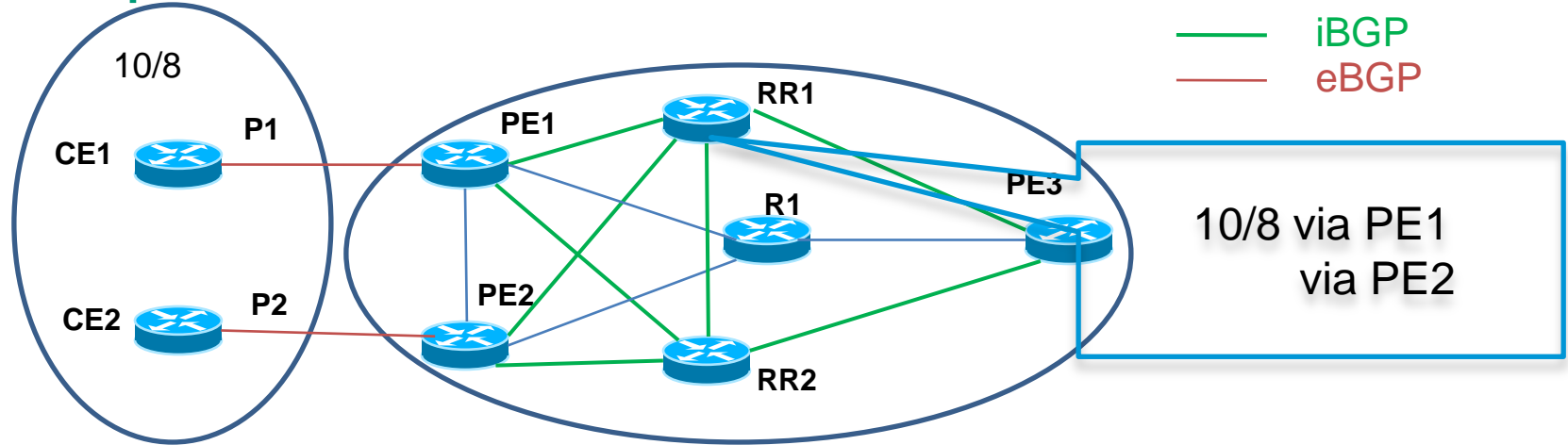


- Best Externalを有効にする
- R2では全体でベストパスを選出する(MEDにより選出)
- R2では外部経路の中のみでのベストパスを選出する(Best External)
- R3では代替パスを持つことが出来る

バックアップパスを伝える手法

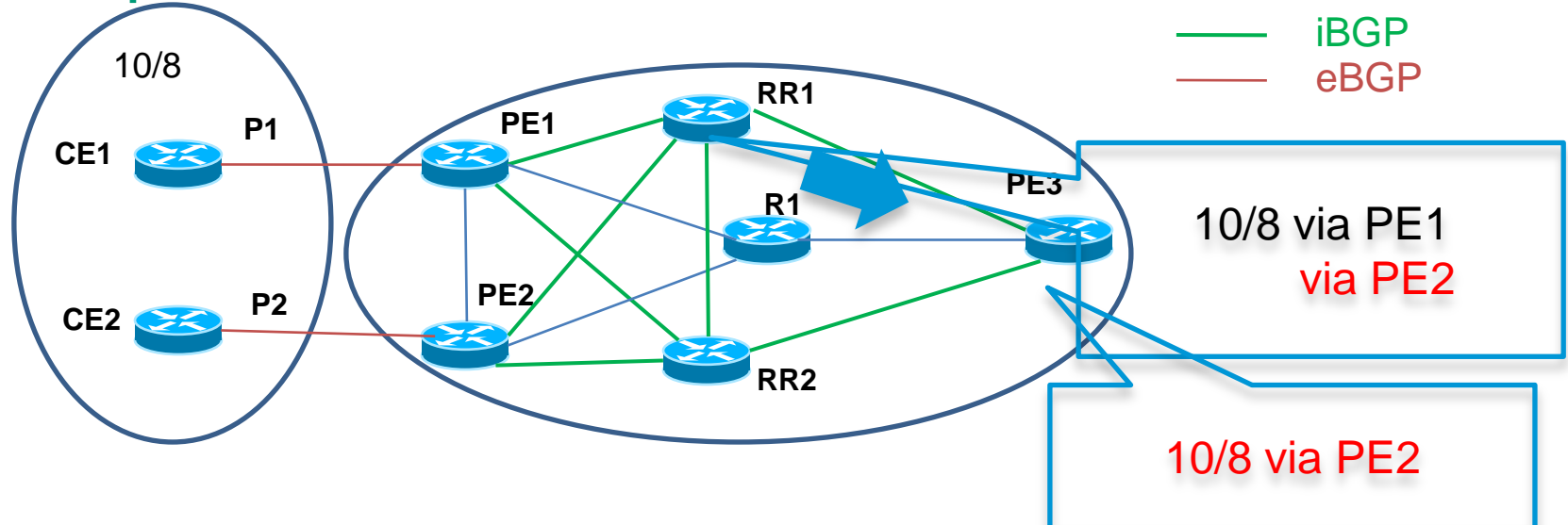
- BGP add-path
- BGP Best External + Fullmesh
- **BGP Diverse Path**

diverse-pathって何だろう？



- P1がP2よりMEDが低く(もしくはLocal pref.が高い)と仮定
- PEでBest Externalを有効にする
- RR1およびRR2ではP2がベストにも関わらず、P1のパスを持つ
- つまりRRではパスダイバーシティ(diverse-path)を持つ。

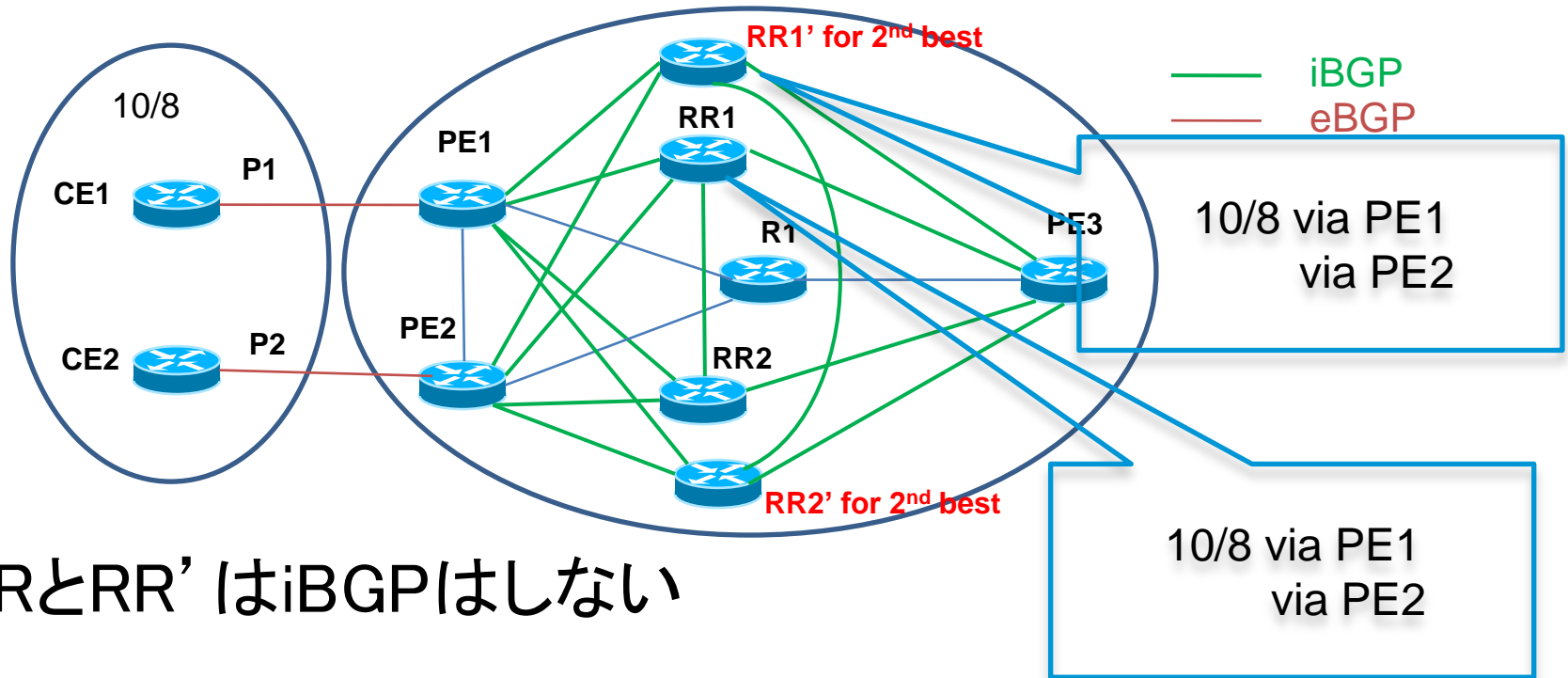
diverse-pathって何だろう？



- RRではMED(Local Pref.)の値よりP2(via PE2)がベストと判断する
- RRはベストパスをアドバタイズするので、PE3はP1(via PE1)を知ることがない。。。。

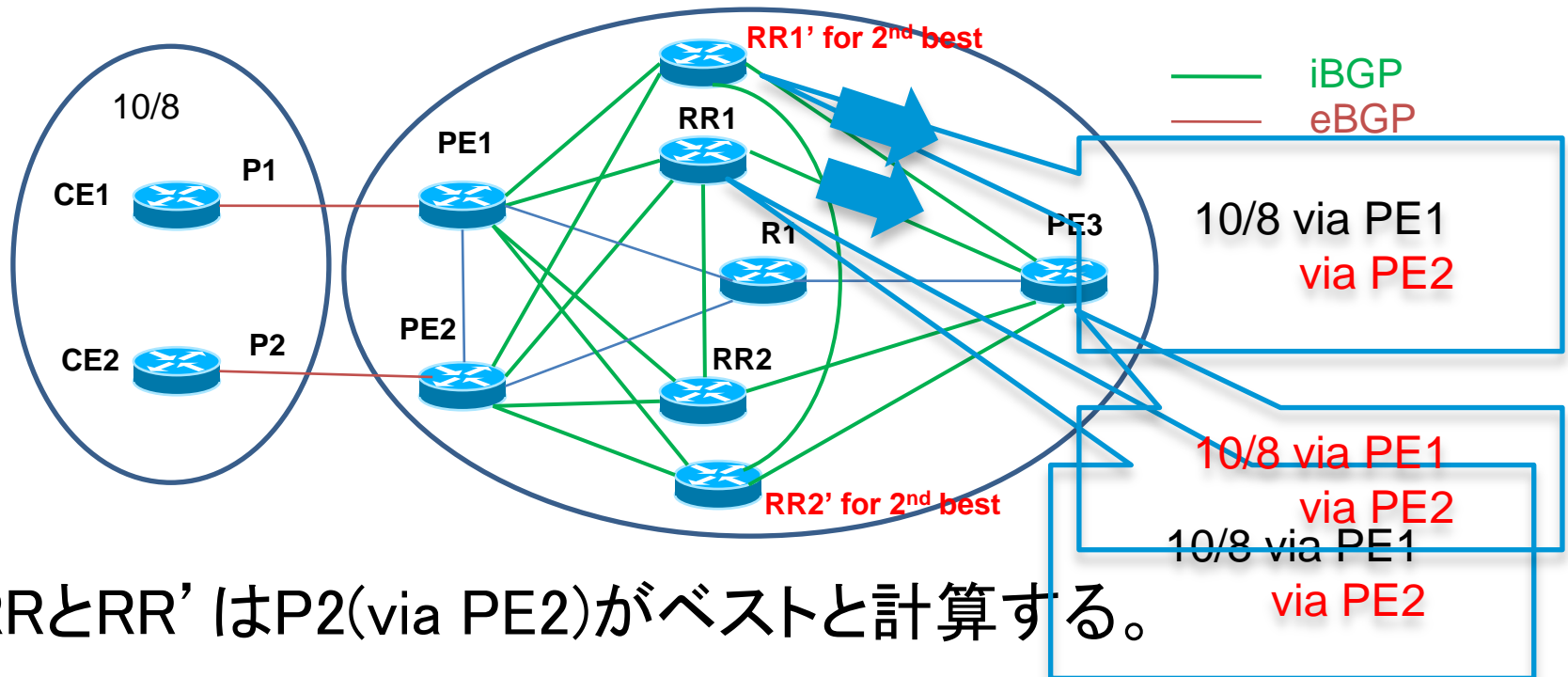
diverse-pathって何だろう？

2nd Bestルート(diverse-path)を分配する為に今あるRRに新しいRRを加える。 2nd Plane RR/Shadow RR



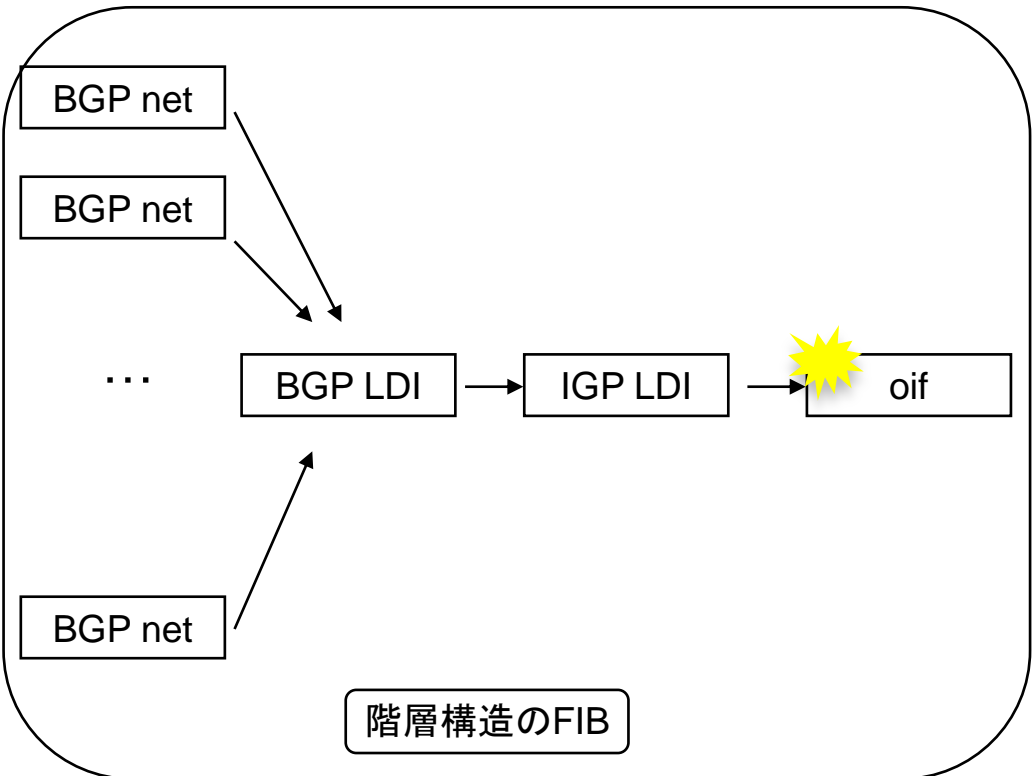
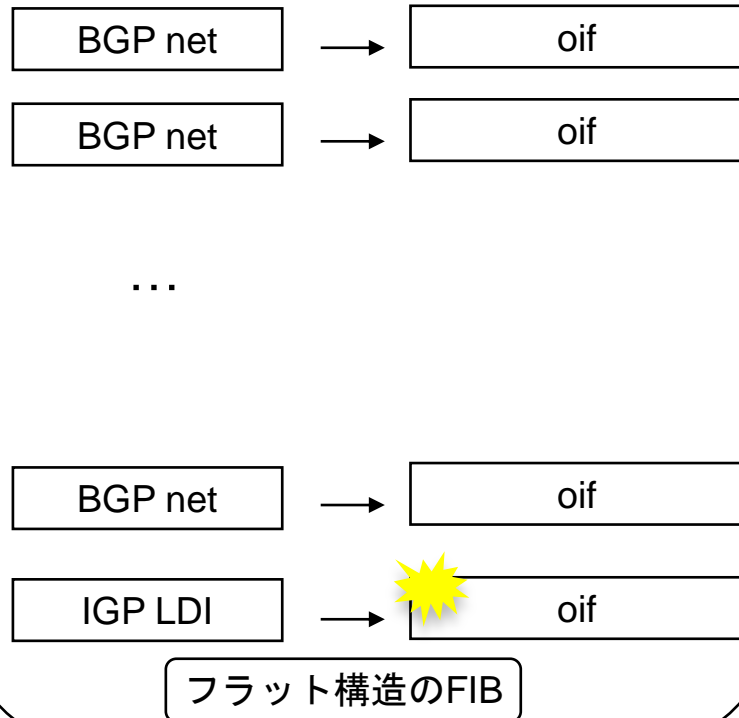
- RRとRR' はiBGPはしない

diverse-pathって何だろう？



- RRとRR' はP2(via PE2)がベストと計算する。
- RR' ではP2(via PE2)を排除し、残りのパスで再計算する
- RR' ではP1(via PE1)を見つけ、2nd Bestと認識する。
- RR' ではP1(via PE2)をベストパスとしてアナウンスする
- PE3ではdiverse-pathを持つ

FIBの構造の違いーPrefix Independent Convergenceー



- フラット構造のFIBでは全てのコントロールプレーンはフラットに扱われ、BGP FIBはそれぞれoifを持つ
- 障害が起こると1prefix毎にoifの書き換えが行われる
- 階層構造では即座にIGP収束後、即座に書き替えが行われる

バックアップパスを伝える手法

solution	add-path	BGP Best External +Full mesh	diverse-path
internet-draft	draft-ietf-idr-add-paths	draft-ietf-idr-best-external	draft-ietf-grow-diverse-bgp-path-dist
Intend to	Standard	Standard	Informational
Author	Cisco,Juniper	Cisco(was with Juniper)	Cisco,Verisign,KDDI
概要	BGPプロトコル拡張	ASBRの機能拡張	RRの機能拡張
導入手法	全てのルータがサポート (ASBR,RR,ABR)	ASBRがサポート iBGP フルメッシュが必要	RRがサポート ASBR/ABRでのShadow RR へのセッション追加

まとめ

- まだまだBGPは拡張し続けている
- 高速コンバージェンスの為に代替パスを如何に伝えるかがポイント
- add-pathではプロトコルを拡張
- diverse-pathはRRの機能を拡張
- Full meshであればBest Externalも使用可能
- 代替パスがあればPICで高速切り替え可能

参考

IETF Internet Draft

- [Advertisement of Multiple Paths in BGP](#)
- [Distribution of diverse BGP paths](#)
- [Advertisement of the best external route in BGP](#)

Cisco Document

- [BGP ベストパス選択アルゴリズム](#)
- [BGP Best External](#)
- [BGP PIC Edge for IP and MPLS-VPN](#)

参考資料

NANOG48

- [Add Paths Overview](#)
- [To Add-Paths or Not Add-Paths](#)

RIPE61

- [BGP Add-Paths](#)
- [BGP Diverse Paths](#)

Thank you.

