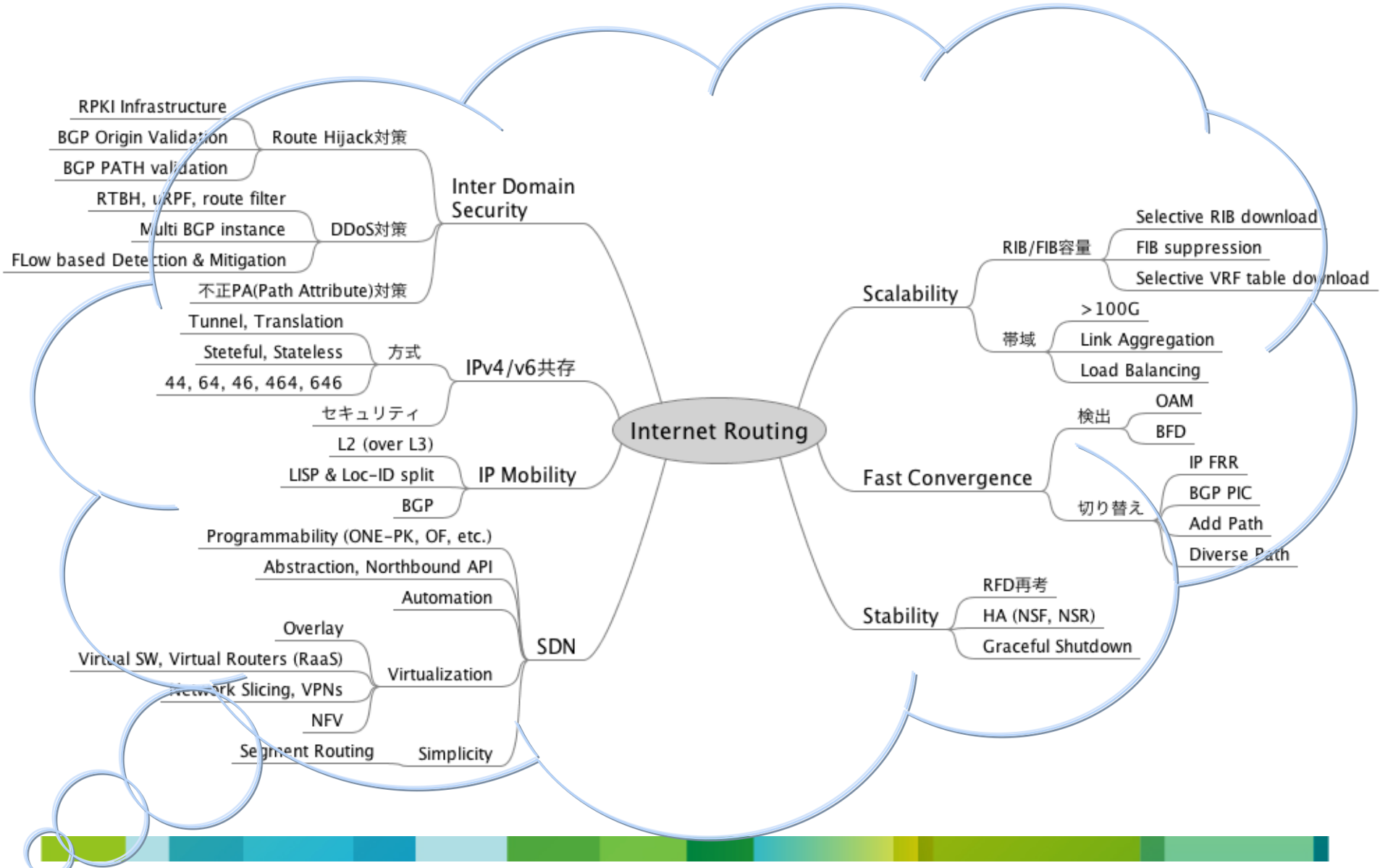
INTEROP
TOKYO I 12 - 14 JUNE,2013
MAKUHARI MESSE

NC-07
ルーティング最新動向
Leading-edge Trends in Routing

12 June 2013

河野 美也，Miya Kohno (mkohno@cisco.com)

“Mindmap” on Routing - 2013



Agenda

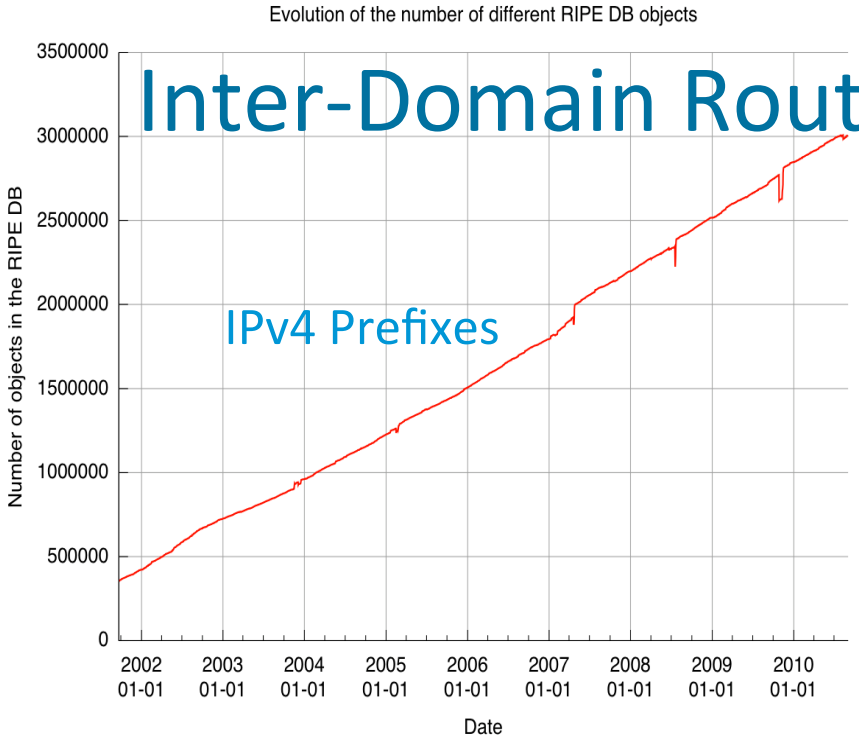
- *(Introduction) BGPの進化*
- Routing Security
- High Availability/Fast Convergence
- Segment Routing

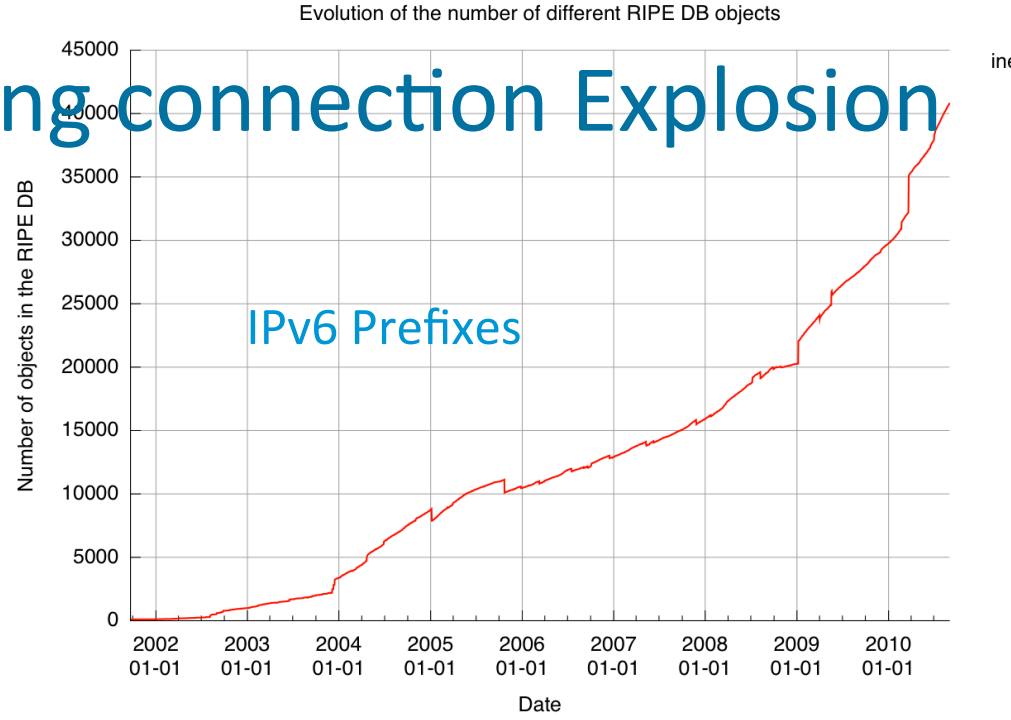
進化し続けるBGP

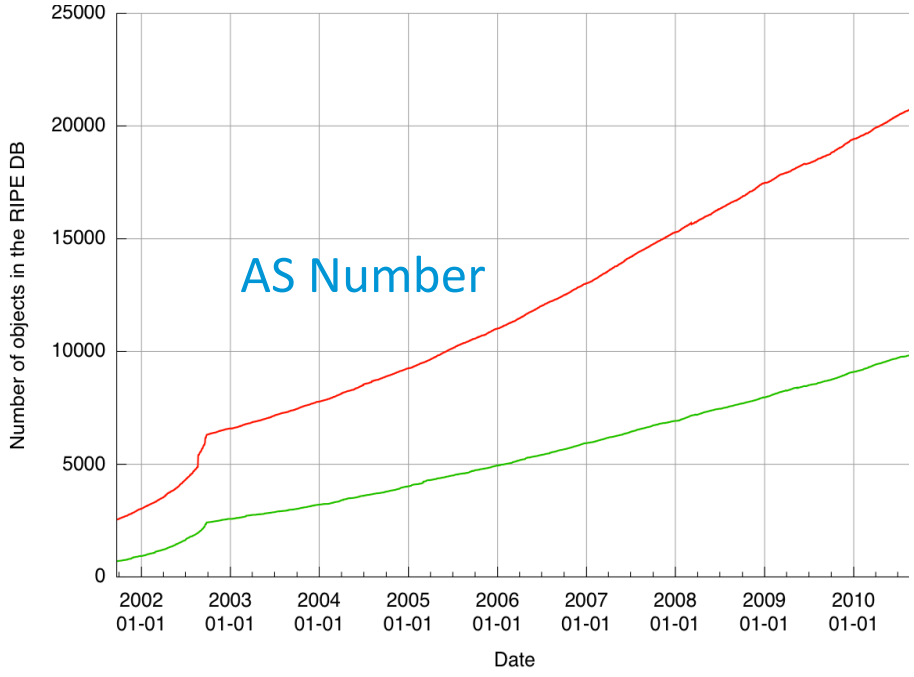
IETF WG's

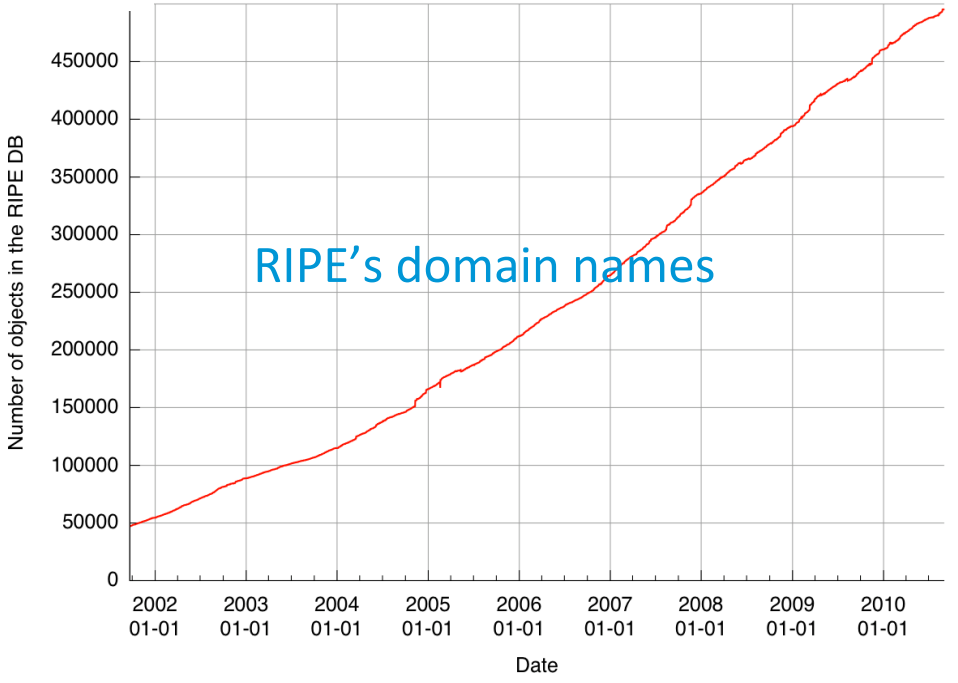
- # WG: IDR, SIDR, L3VPN, L2VPN, GROW, OPSEC, NVO3, VPN4DC, I2RF.., and more
- # RFC: Over 100 RFC
- # draft: Over 50 IETF drafts
- Cisco engagement: ~50 BGP Engineer's

Inter-Domain Routing connection Explosion





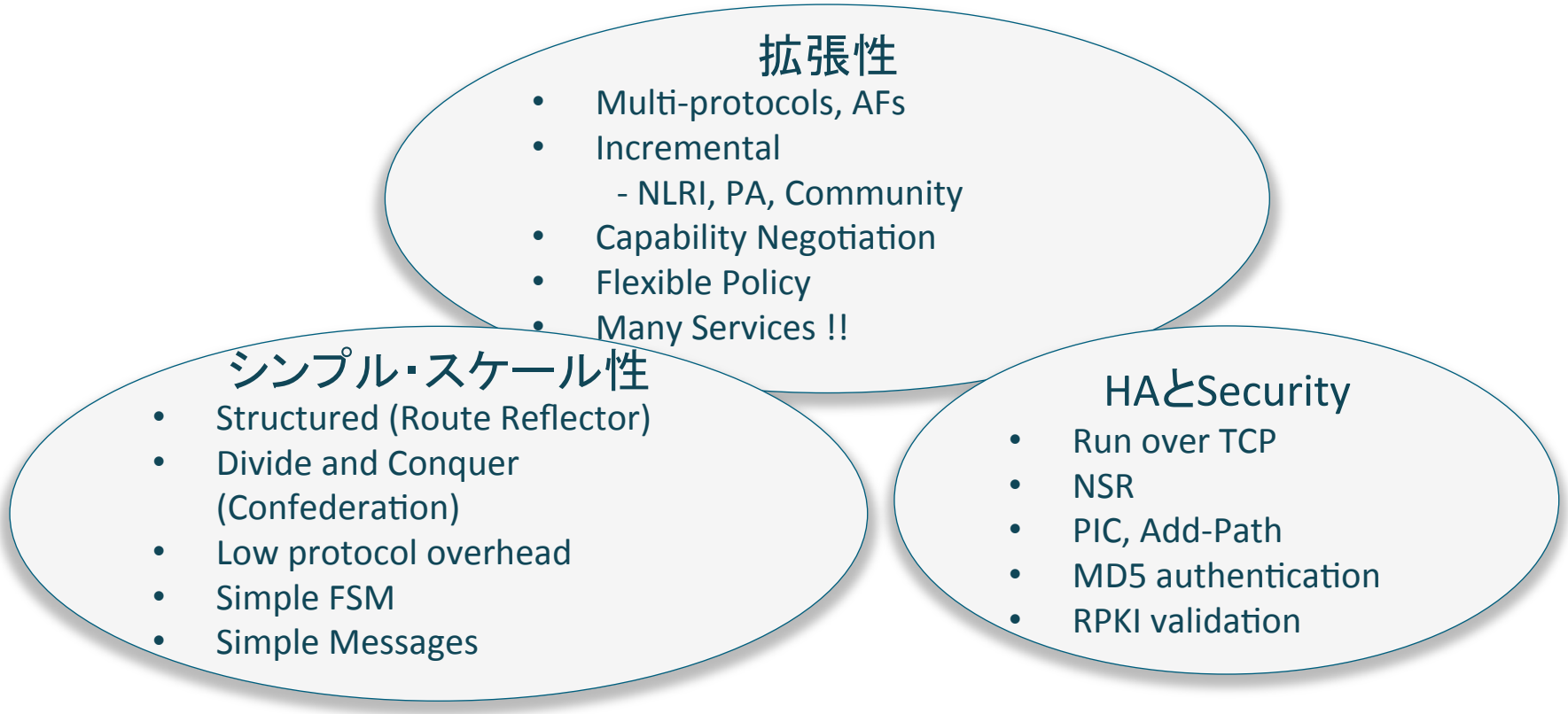




Control-plane Evolution

Service/transport	In 2009	In 2012+	Market
Internet Peering	BGP	BGP (IPv4 + IPv6)	SP
SP L3VPN	BGP	BGP + FRR + Scalability	
SP Multicast VPN	PIM	BGP Multicast VPN	
Multicast MPLS	PIM / mLDP	segmented LSM (Mc Unified MPLS)	
DDOS mitigation	PBR, ACL, RTBH	BGP flowspec	
Network Monitoring	SNMP	BGP monitoring protocol	
Security	Filters, ACL	BGP Sec (RPKI)	
SP SDN (NPS / PCE / Alto)		BGP OnePK API/ BGP LS	
MPLS transport	LDP	LDP + BGP+Label (Unified MPLS)	
Business & CE L2VPN	LDP	BGP AD/Sign (VPLS)	
DCI NG L2VPN		BGP AD/Sign (EVPN)	DC / SP
Massive Scale DC	OSPF/ISIS	BGP + Multipath	DC
SP-DC		BGP Inter-AS, vPE, vCE	
Campus L3VPN & mVPN	BGP	BGP (IPv4 + IPv6)	
VxLAN / LISP encap / GRE	LISP / GRE	BGP remote next hop + GRE /LISP encap	
Massive scale DMVPN	NHRP / EIGRP	BGP + Path Diversity	Enterprise
FlexVPN		BGP	
Managed CPE	BGP IPv4	BGP IPv4 & IPv6	

Why is BGP successful?

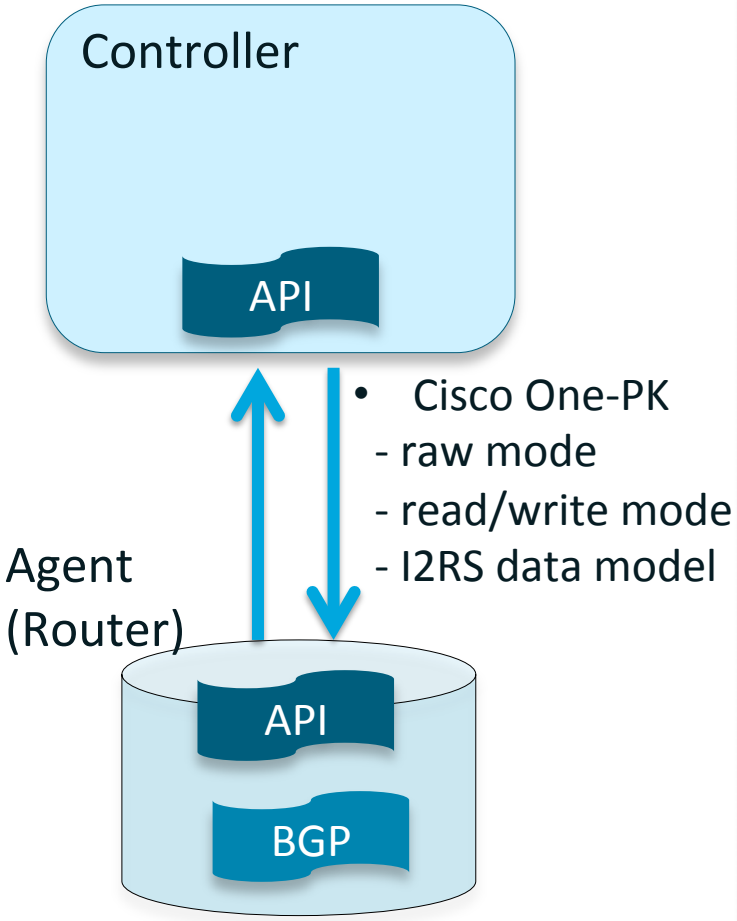


“Driven by Pragmatism”, “Not perfect, but good enough”

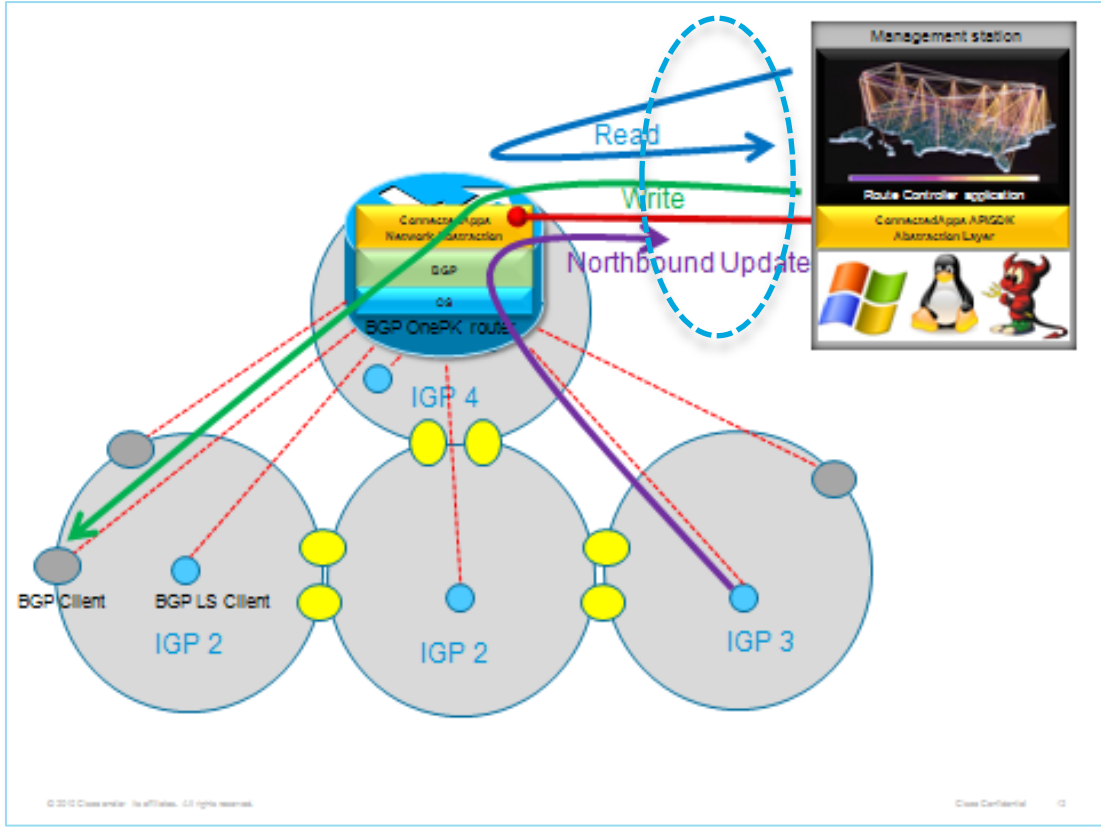
-- Yakov Rekhter

“Pervasive L2/L3 Tunnels, nowadays..” -- Miya Kohno

BGP – ONE-PK APIモデル



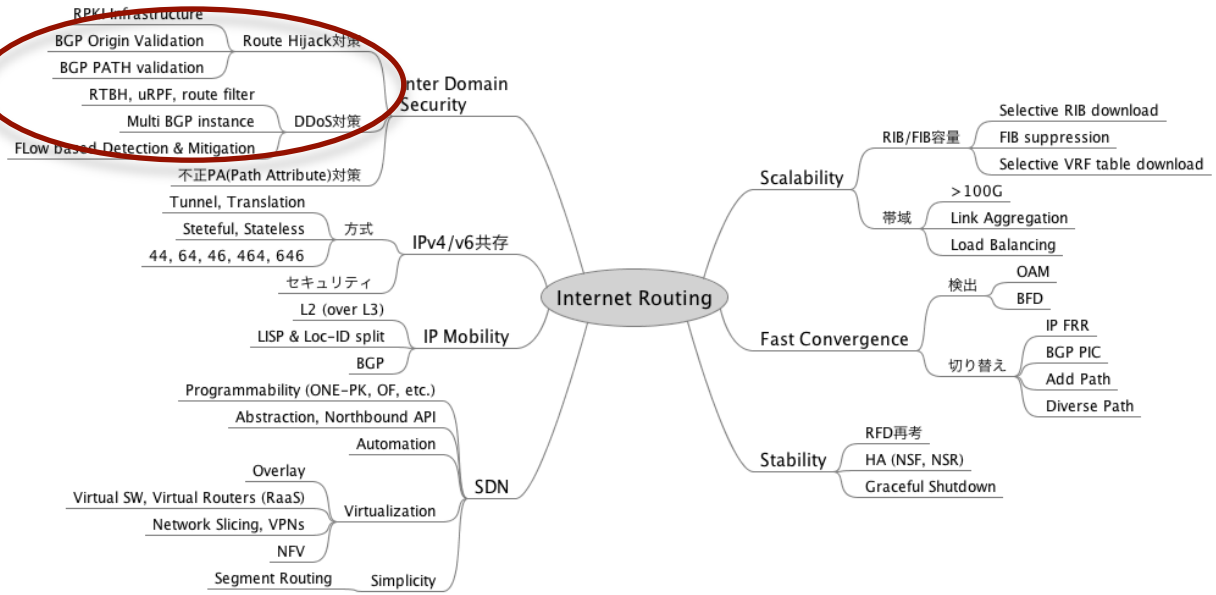
BGP LSにおける例



Agenda

- (Introduction) BGPの進化
- *Routing Security*
- High Availability/Fast Convergence
- Segment Routing

1. Route Hijack対策
2. DDoS対策
3. 不正PA対策



1. Route Hijacking(*)対策

(*) オペミスや何らかの障害による不正経路広告を含みます。

Route Hijackを防ぐ3つの柱

RPKI Infrastructure

- 検証された安全なObjectのリポジトリ
- リソース(IPv4, IPv6 + ASN)割当の階層構造に従う

BGP Origin Validation

- BGP UPDATEのOrigin ASに対して妥当性検査を行う
- 殆どの問題はこの方法で防げる(**)
- ルータのハードウェアに変更の必要は無い
- 標準化はほぼ終了している

BGP PATH Validation

- 新たなBGPアトリビュートと機能 (BGPSEC)
- ASPATHアトリビュートをサインして転送する
- 標準化作業中

(**) http://www.nanog.org/meetings/nanog49/presentations/Tuesday/HowSecure_NANOG_print.pdf

Origin Validation in Action

- ***Trust Anchorの設定***
  (誰をTrust Anchorとするか)

BGP Peer

eBGP

Parent CA

Publication protocol

repository operator

Peering Router

rpki-router protocol

rsync

Provisioning Protocol (up/down)

iBGP + ext Com.

RPKI Validator & Cache

rpki-router protocol

Subordinate CA

Publication protocol

repository operator

rsync

Peering Router

ISP Infrastructure (relaying party)

eBGP

RPKI Infrastructure

- ***ROAの設定:***
  (prefixの使用を権威付けし、それを周知)

BGP Peer

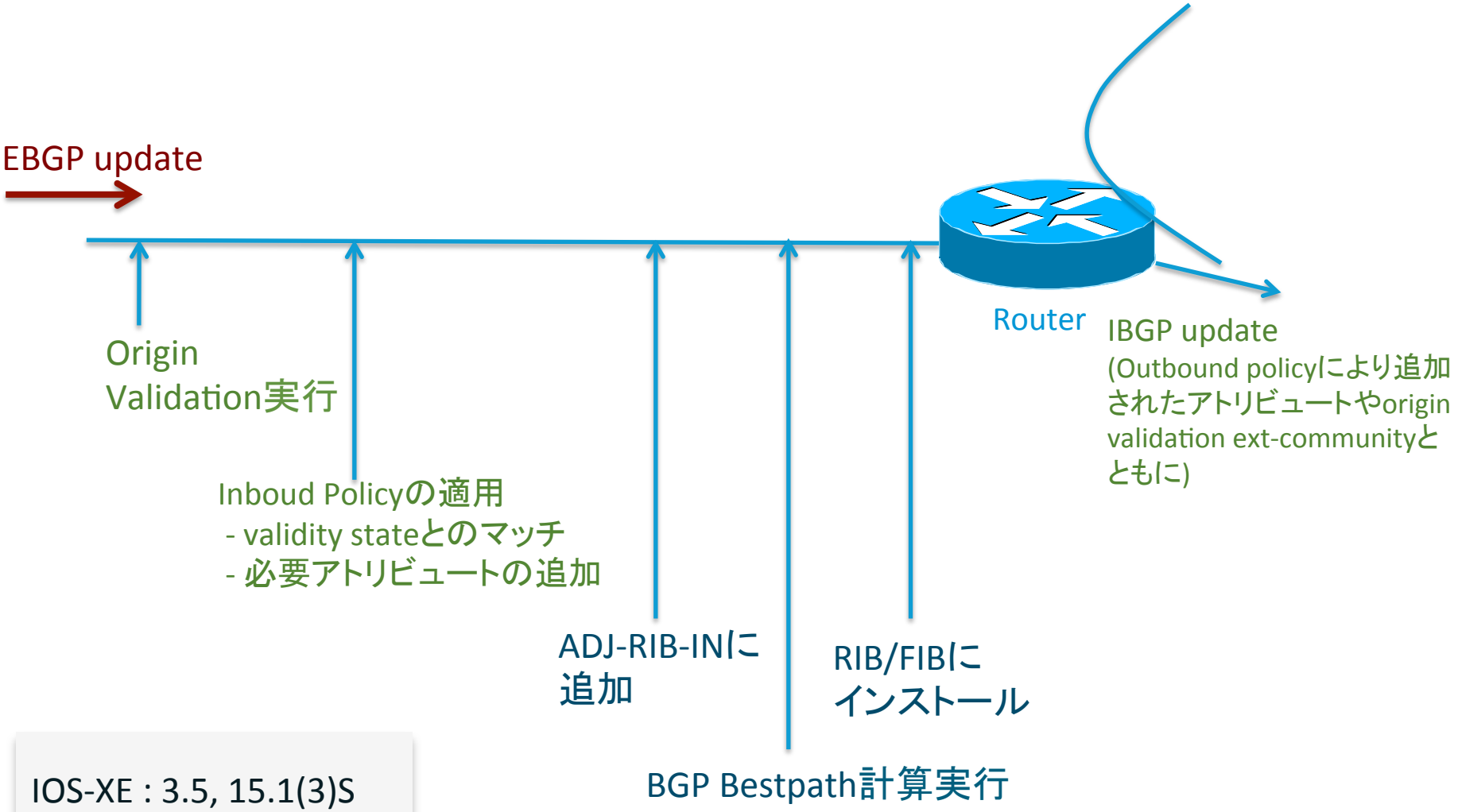
主な関連RFC (1/2)

Doc	Title	Date
RFC 6480	An Infrastructure to Support Secure Internet Routing	Feb 2012
RFC 6481	A Profile for Resource Certificate Repository Structure	Feb 2012
RFC 6482	A Profile for Route Origin Authorizations (ROAs)	Feb 2012
RFC 6483	Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs)	Feb 2012
RFC 6484	Certificate Policy (CP) for the Resource Public Key Infrastructure (RPKI)	Feb 2012
RFC 6486	Manifests for the Resource Public Key Infrastructure (RPKI)	Feb 2012
RFC 6487	A Profile for X.509 PKIX Resource Certificates	Feb 2012
RFC 6492	A Protocol for Provisioning Resource Certificates	Feb 2012
RFC 6493	The Resource Public Key Infrastructure (RPKI) Ghostbusters Record	Feb 2012

主な関連RFC (2/2) 2013-

Doc	Title	Date
RFC 6810	The RPKI/Router Protocol	Jan 2013
RFC 6481	BGP Prefix Origin Validation	Jan 2013
RFC 6907	Use Cases and Interpretation of RPKI Objects for Issuers and Relying Parties	Mar 2013
RFC 6916	Algorithm Agility Procedure for RPKI	Apr 2013
RFC 6945	Definitions of Managed Objects for the RPKI-Router Protocol	May 2013

Origin Validation対応のためのBGP実装変更



IOS-XE : 3.5, 15.1(3)S
IOS-XR : 4.2.1

2. DDoS対策

Phase I

- RTBH
- ACL
- uRPF
- …

Phase II

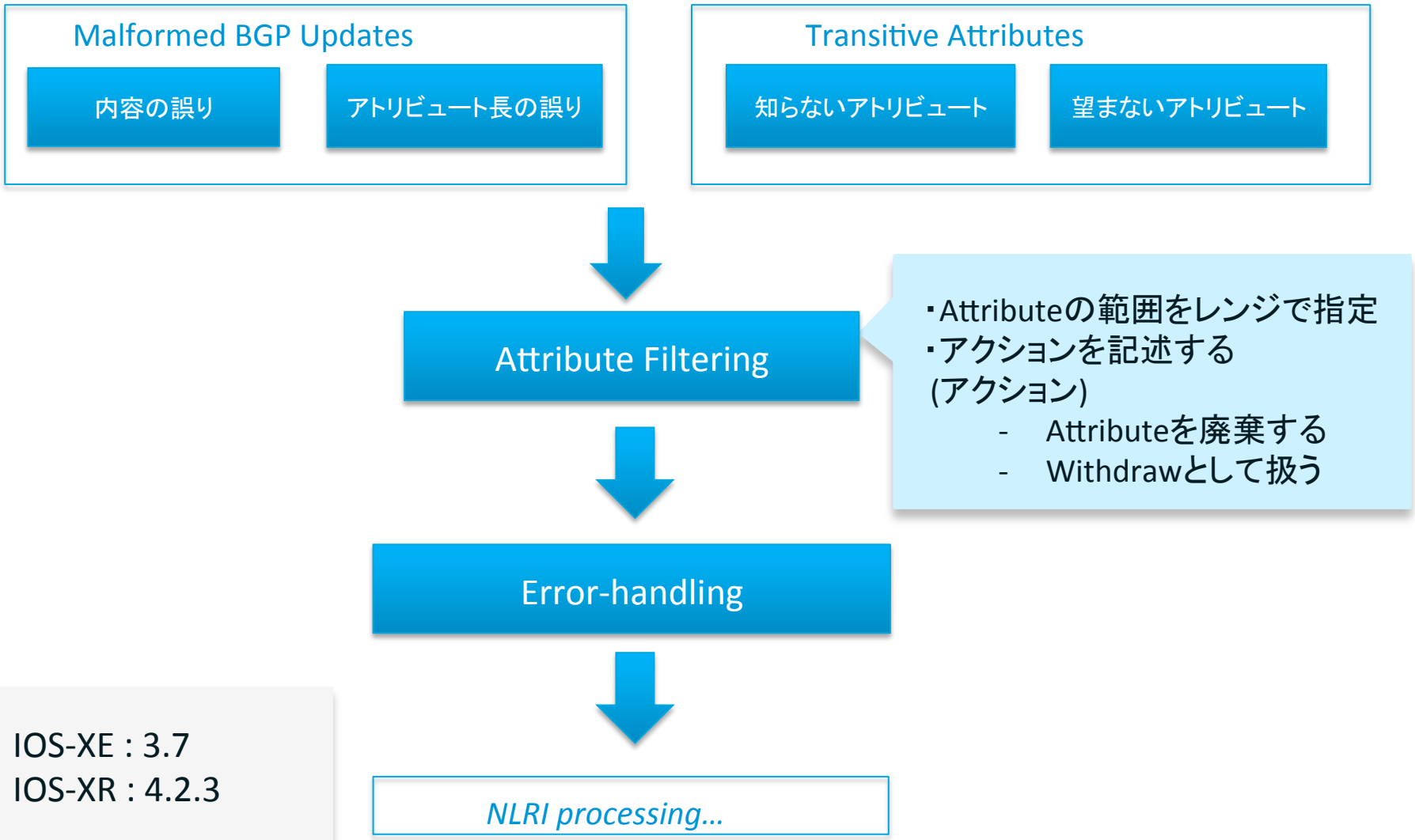
- 異常フロー検出と緩和
- Multi-Instance BGPの利用

Phase III

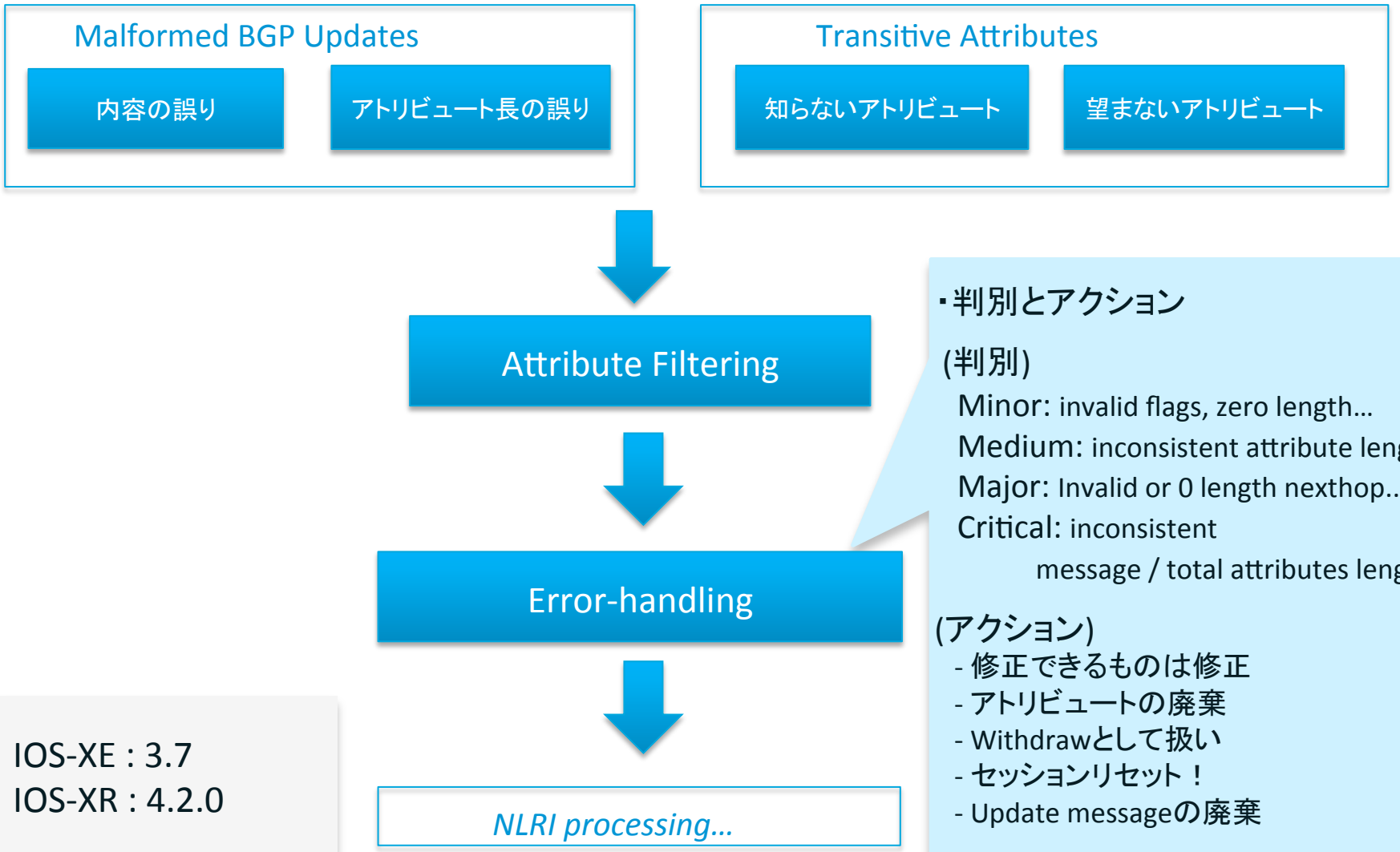
- Dynamic & Application Aware redirection/Traffic Steering

https://ripe66.ripe.net/presentations/306-20130516_v1_RIPE66_DDoS_Mitigation_gvandeve.pdf

3. 不正PA問題対策 - filtering

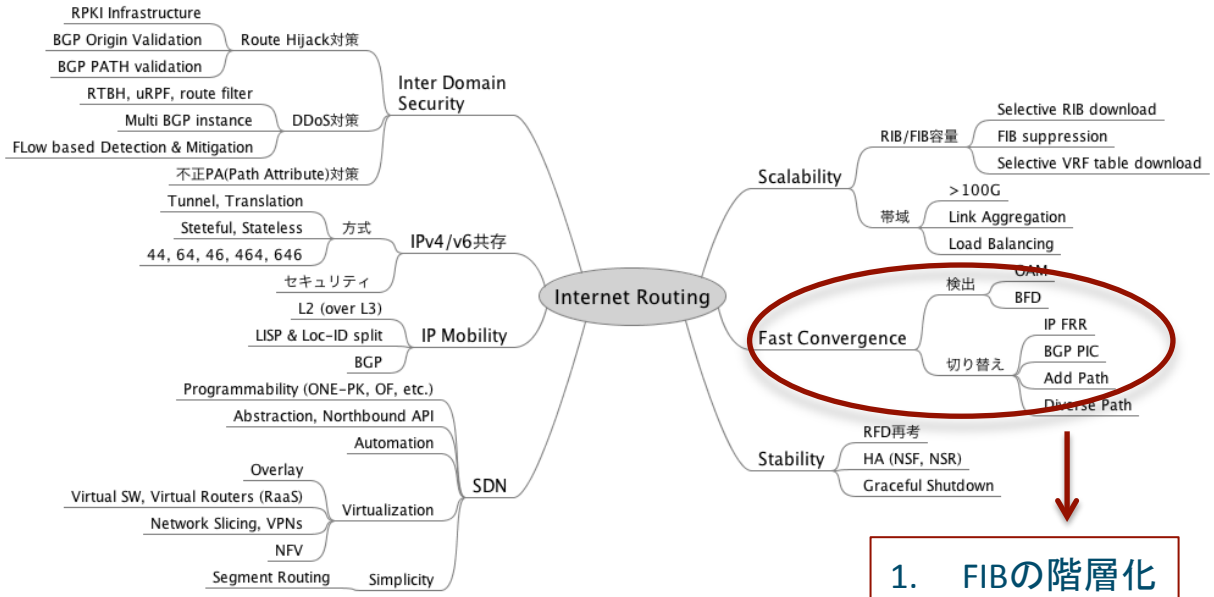


3. 不正PA問題対策 – error handling



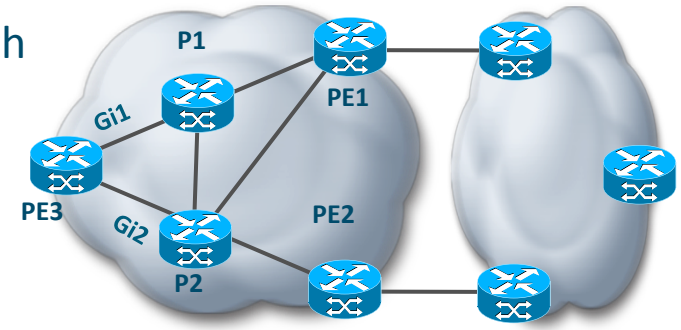
Agenda

- (Introduction) BGPの進化
- Routing Security
- *High Availability/Fast Convergence*
- Segment Routing



高速切替(Fast Convergence)のための要素

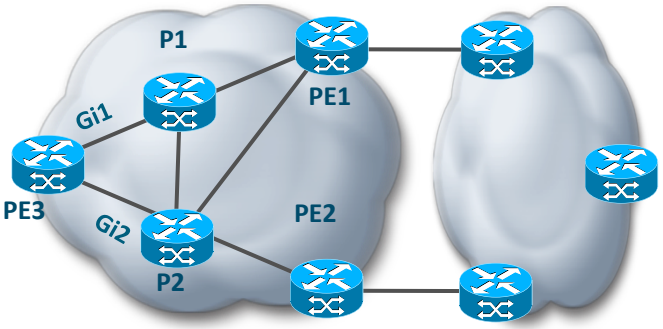
- 冗長経路を持ち、あらかじめRIB/FIBに載せておく
  - IGP : IGP Multipath, IP FRR (LFA), MPLS FRR
  - BGP :
    - BGP PIC Core
      - NextHopに対する複数のIGP経路
    - BGP PIC Edge
      - BGP Multipath
      - Best external, add-path, diverse-path
      - …

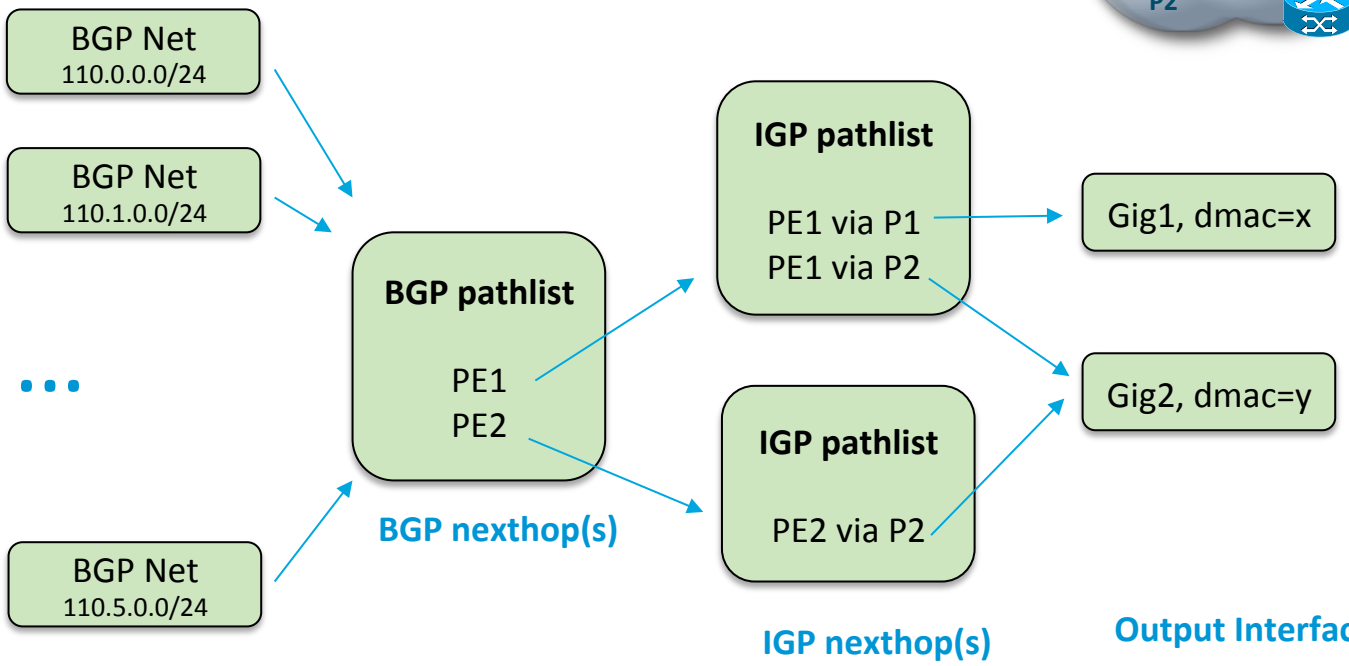


- 高速に検出する
  - LOS, Ether-OAM, BFD..

- 階層化FIBにより、再帰的にテーブルを書き換える

FIBの階層化





- 実際の出力Interfaceではなく、Nexthopへのポインターを持たせることにより、経路の変化や消失のたびにpathlistを書き替える必要が無い

↓

テーブル書き換えのための時間は最小で済む。Prefix数に依存しない。

RCMD – Routing Convergenceの可視化

- Routing Convergence Monitoring & Diagnostics

先日の回線障害が、ユーザに
与えた影響（切替時間）は
どの程度だったか?

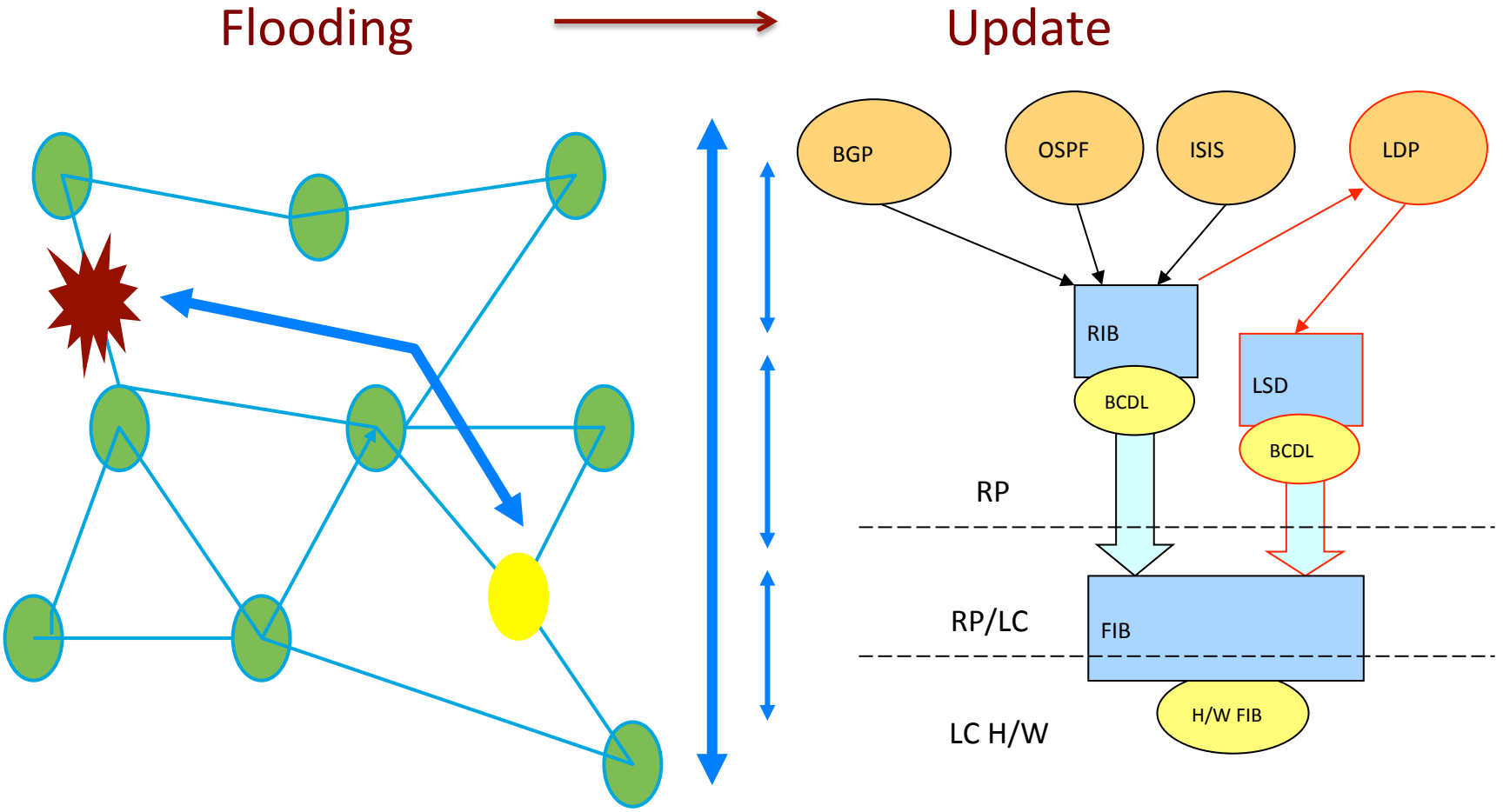
ネットワーク設計変更やパラ
メータ変更によって、収束時間
はどのように変化したか?

経路変更は、どのくらいの時間
で伝搬したか?

これらの値を、
現用のネットワークから
取得するのは
難しかった！

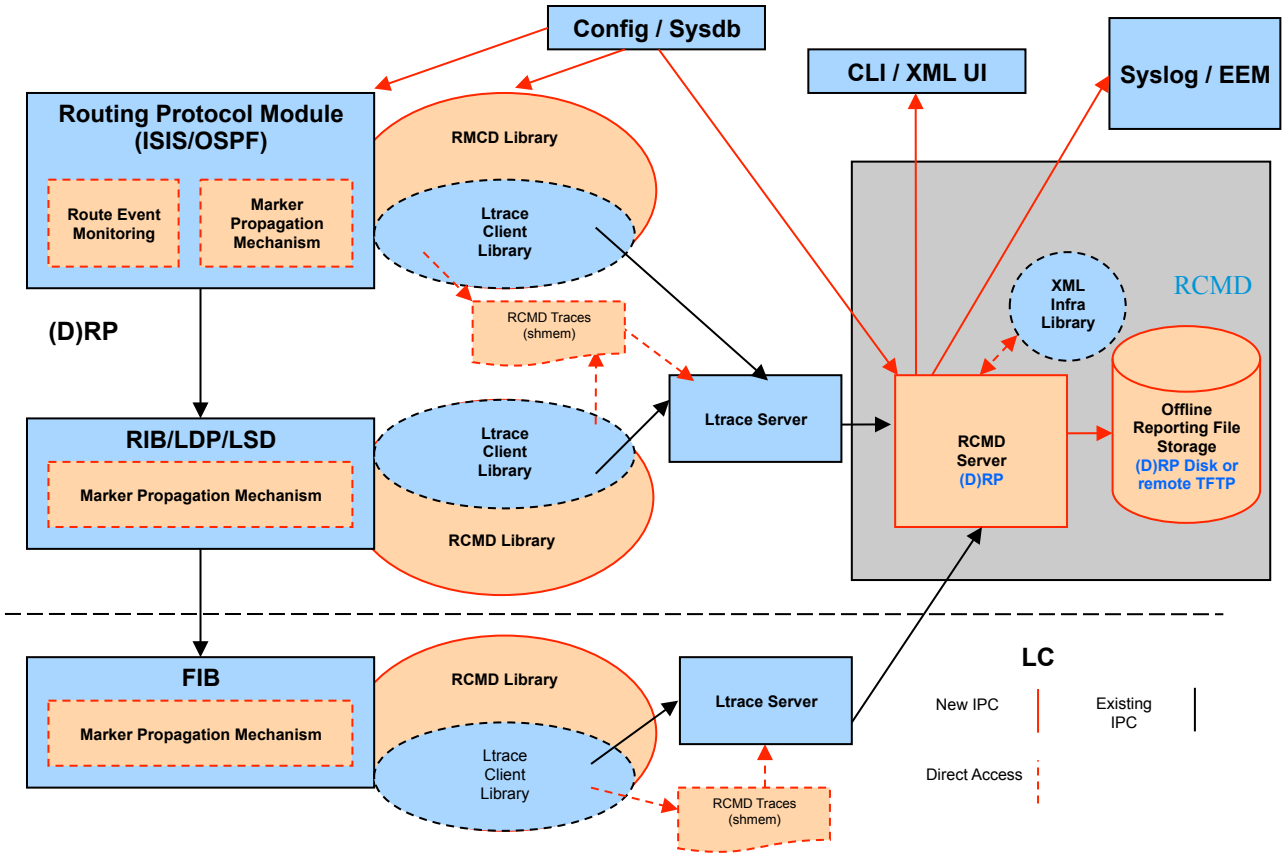
RCMD – Routing Convergenceの可視化

Routing Convergence



RCMD – Routing Convergenceの可視化

ルータ内部のふるまいを記録する !!

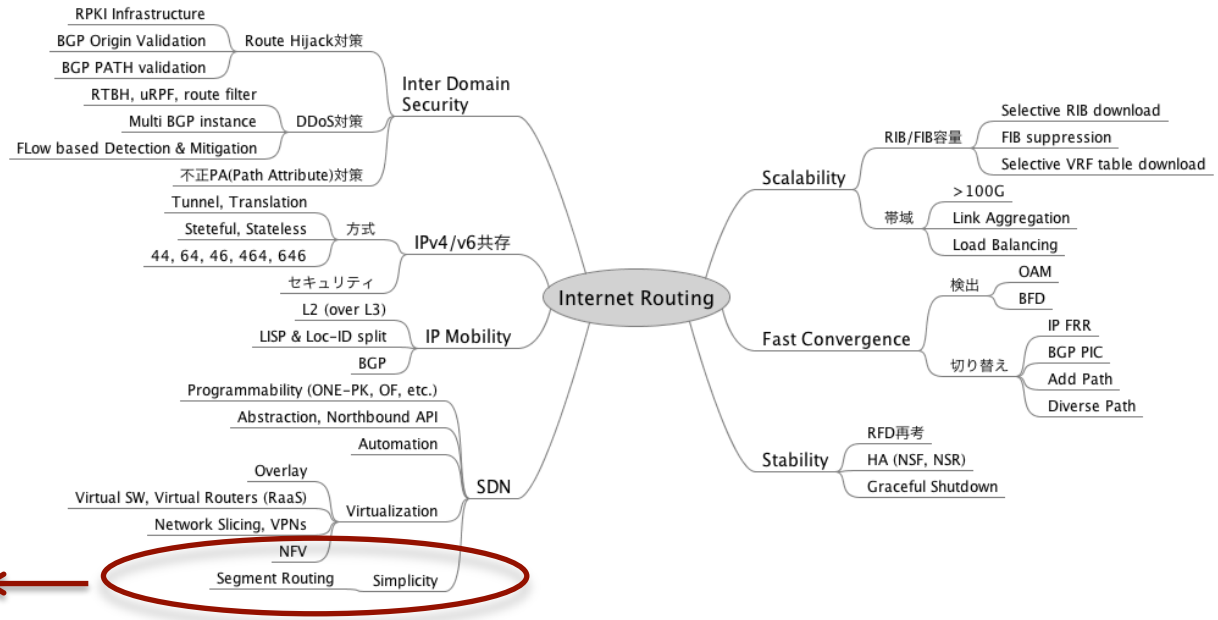


Monitor : Interface events, Flooding, OSPF/ISIS SPF events, Prefix Addition/Deletion

Report : CLI and XML

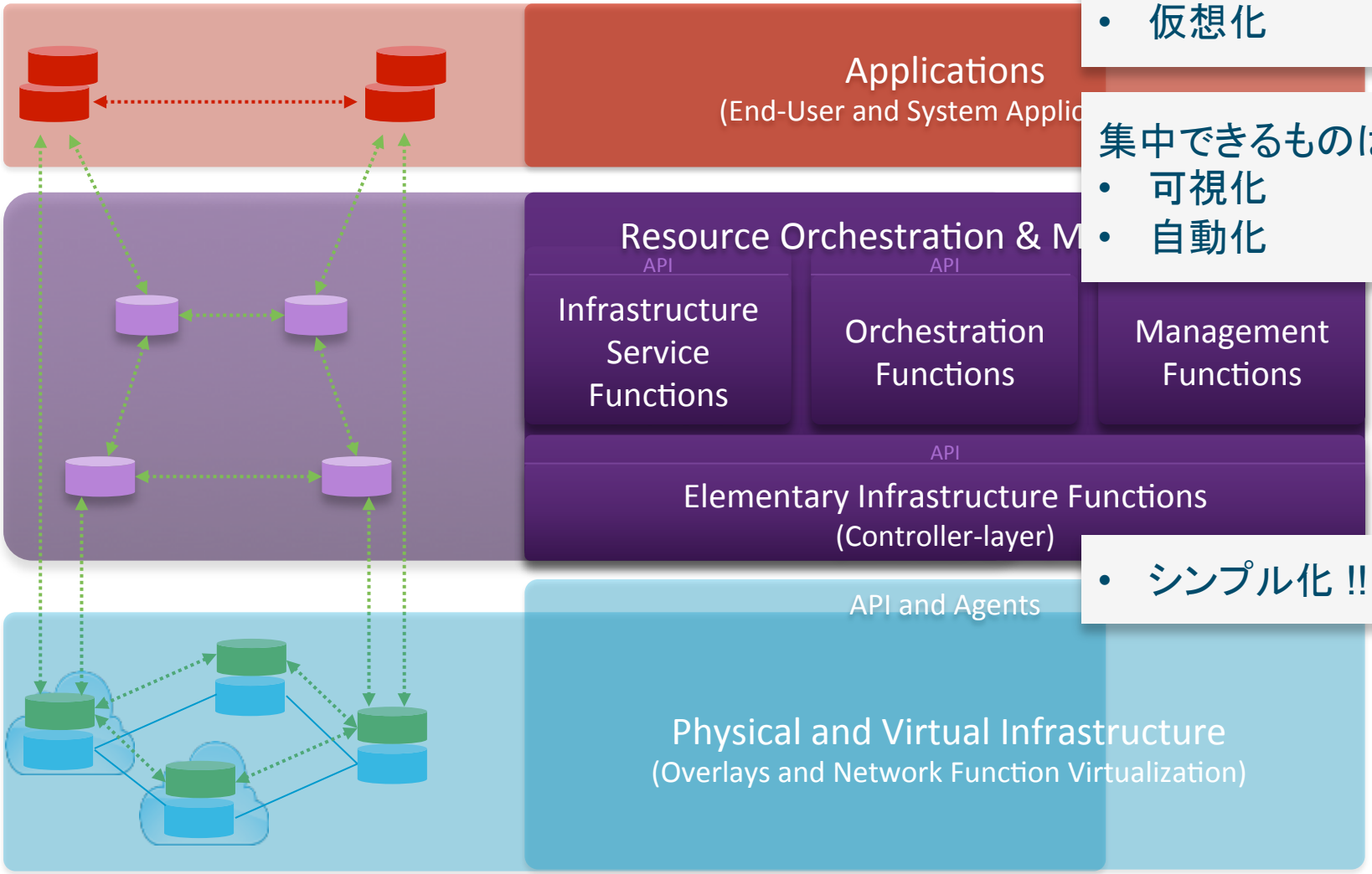
Agenda

- (Introduction) BGPの進化
- Routing Security
- High Availability/Fast Convergence
- *Segment Routing*



1. Segment Routing !!

SDN – Architectural Model



- 抽象化
- 仮想化

集中できるものは集中

- 可視化
- 自動化

- シンプル化 !!

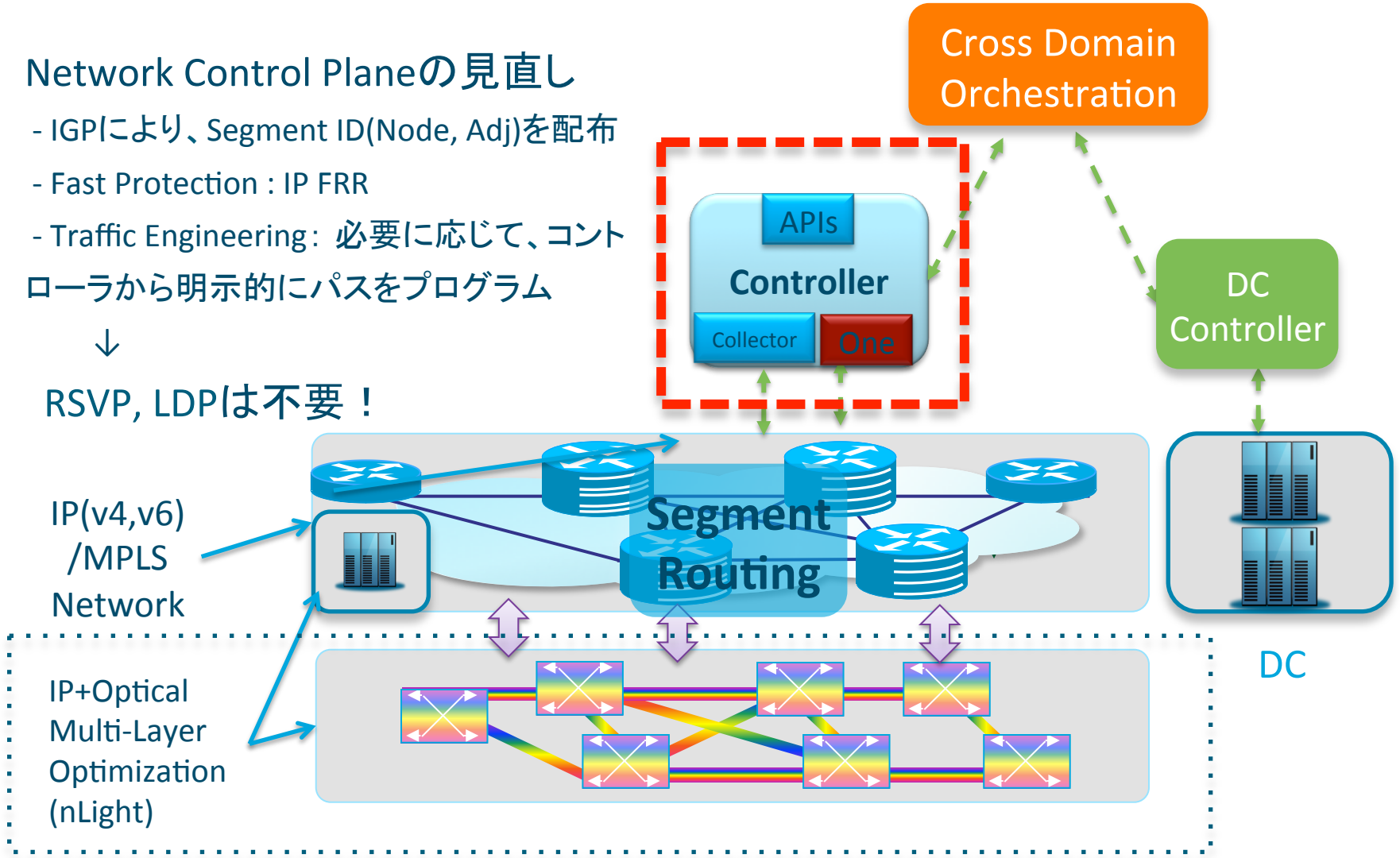
Segment RoutingによるNWのシンプル化

Network Control Planeの見直し

- IGPにより、Segment ID(Node, Adj)を配布
- Fast Protection : IP FRR
- Traffic Engineering : 必要に応じて、コントローラから明示的にパスをプログラム

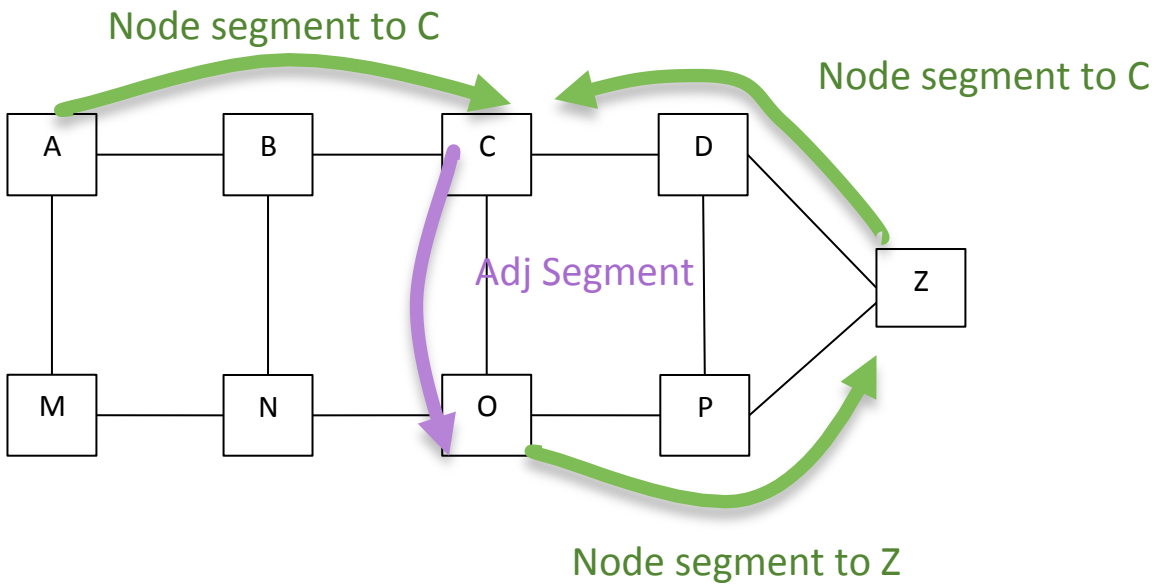
↓

RSVP, LDPは不要！

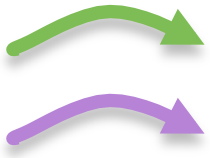


draft-previdi-filsfils-isis-segment-routing

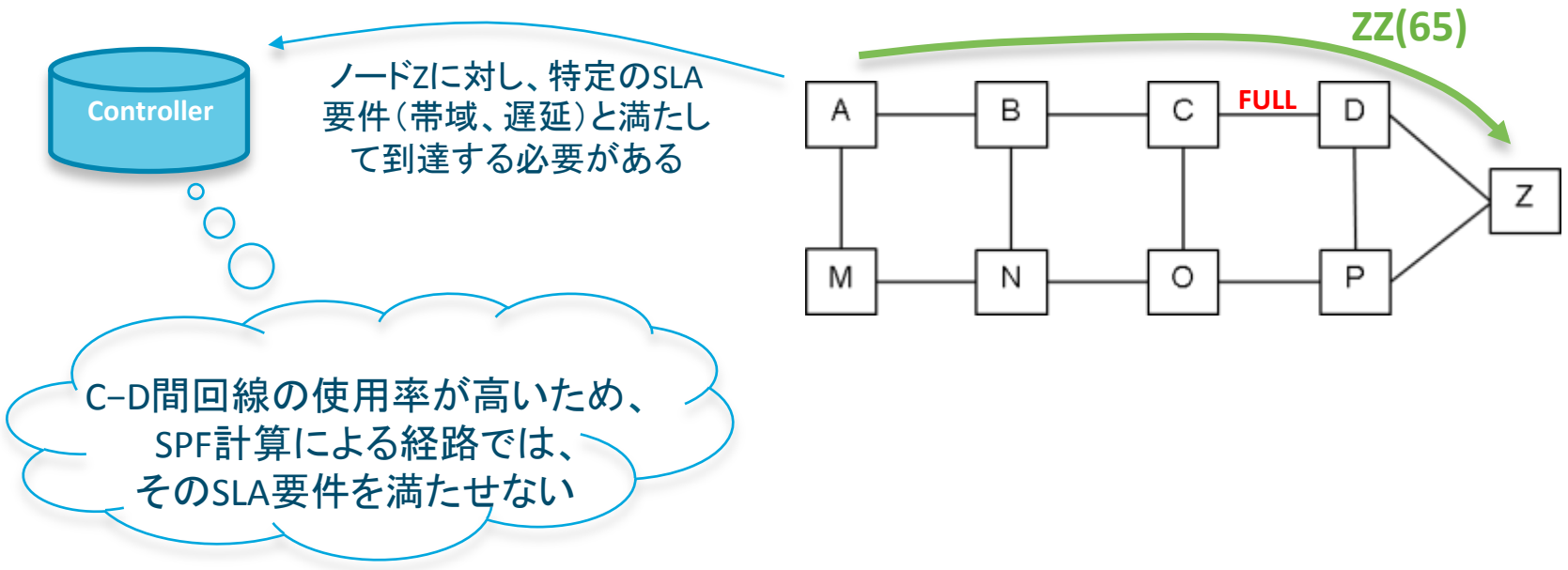
Segment Routing – 基本動作



- IGP(ISIS/OSPF)は、自動的に”segment”をつくり、維持する
  - Node Segment: 該当ノードへのshortest-path
  - Adjacency Segment: 隣接ノードへのone-hop path

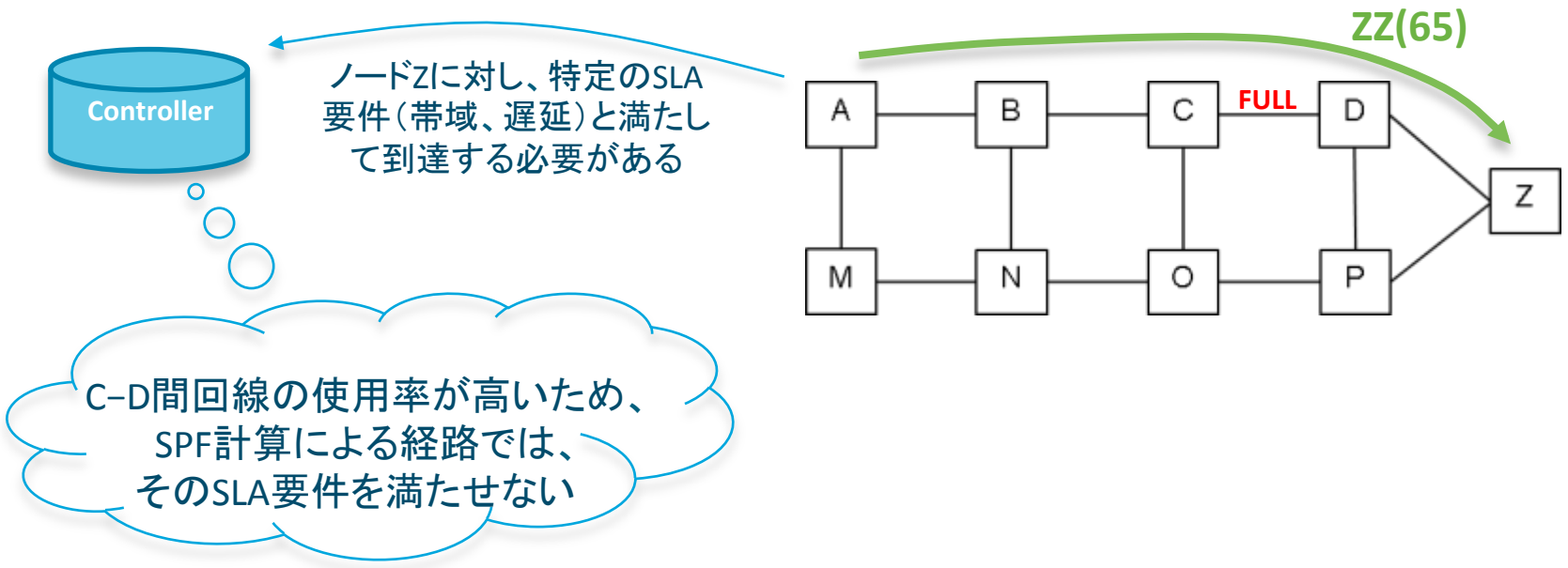


Segment Routing – Controllerによる制御



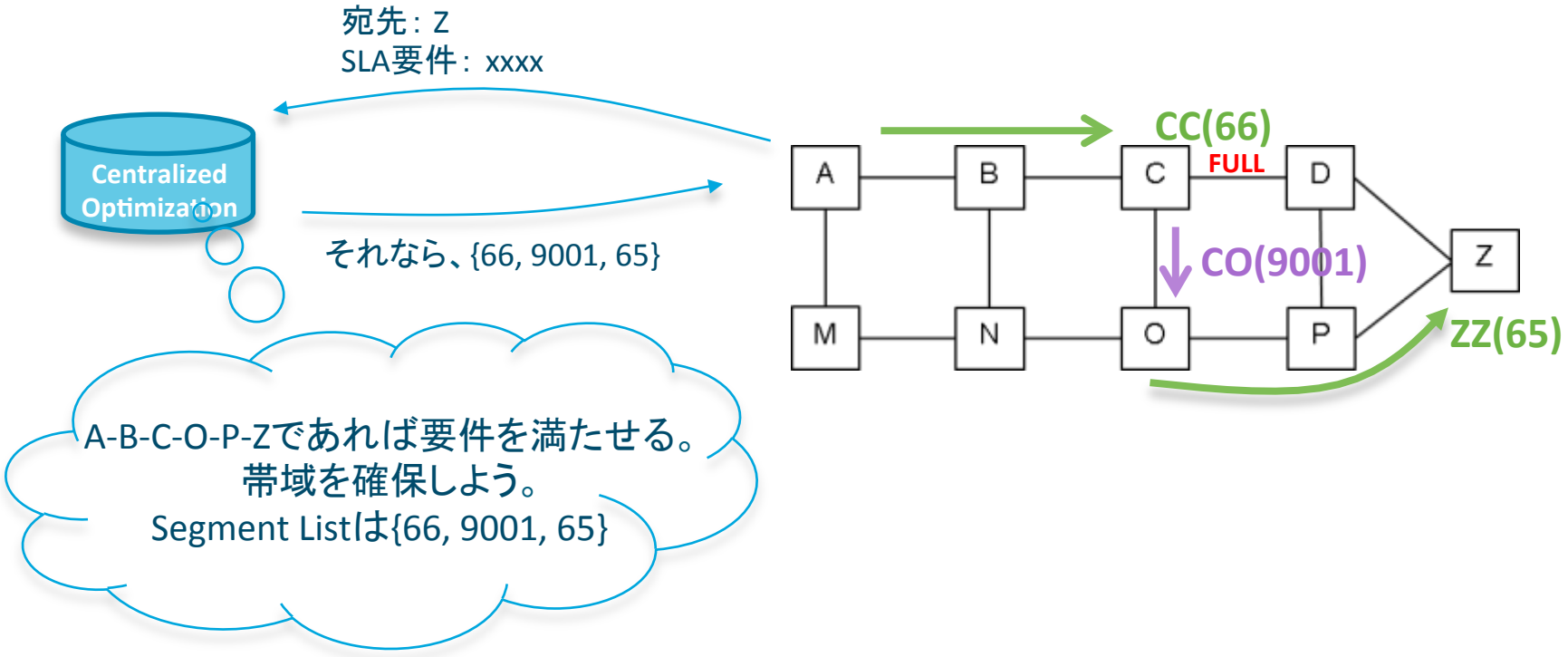
- Controller
  - そのSLA要件を満たすパスを発見する
  - NodeおよびAdjacency Segmentのリストをencodeする

Segment Routing – Controllerによる制御



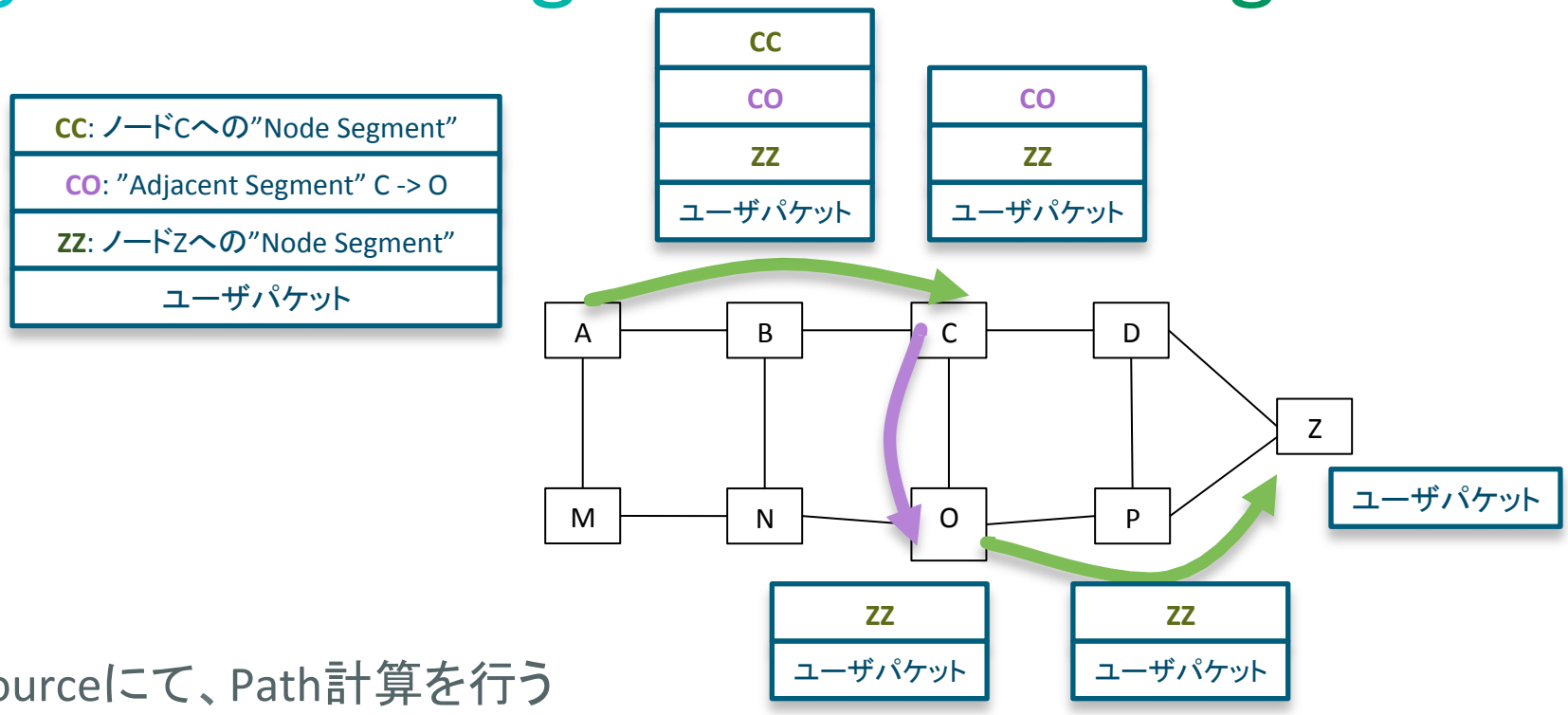
- Controller
  - そのSLA要件を満たすパスを発見する
  - NodeおよびAdjacency Segmentのリストをencodeする

Segment Routing – Controllerによる制御



- Controller
  - そのSLA要件を満たすパスを発見する
  - NodeおよびAdjacency Segmentのリストをencodeする

Segment Routing - Source Routing



- Sourceにて、Path計算を行う

  集中制御による補完も可能

- SourceはHeader Stackを生成し、パケットを送出する

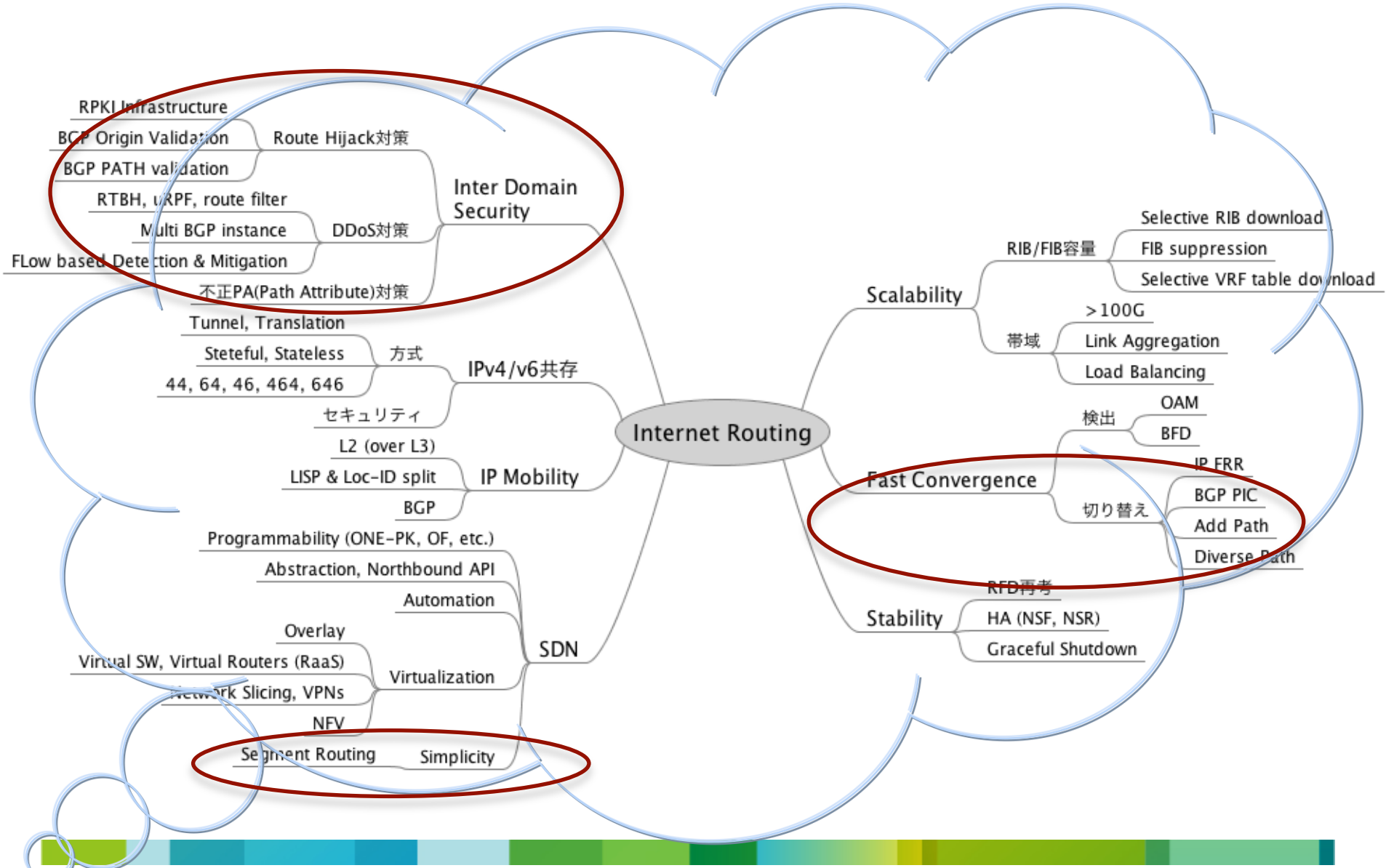
  Pathは、Segment idのリストとして表現される。(Segment ID == Label)

- 中継ノードは、Label Switching/Forwarding

Segment Routingのメリット

- Scalable !!
  - 使用するControl PlaneはIGPだけ
    - 他のControl Plane(LDP, RSVP..)を必要としない
    - LDP-IGP syncなどのstate syncも必要ない
- Traffic Engineering capable !!
  - 柔軟性
    - Customized Routing
    - Disjoint Service Topology
    - 明示的Load balancingなど
  - Scalabilityを阻害しない
    - RSVP stateを持つ必要が無い
    - 全てのstateはヘッダ(Label Stack)にある

“Mindmap” on Routing - 2013



Thank you.