



経路制御(2)

情報ネットワーク論I

門林 雄基

TA: Omar, 衛藤、鈴木、横山



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



アウトライン

- ✿ 距離ベクトル型 (Distance vector)
- ✿ リンク状態型 (Link state)
- ✿ IGP と EGP
 - ・ドメイン内ルーティング、ドメイン間ルーティング
- ✿ パスベクトル型 (Path vector)
- ✿ BGP
- ✿ 経路の集約

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



FIB (Forwarding Information Base) の例

```
zebra@c2.naist# show ipv6 route
Codes: K - kernel route, C - connected, S - static, R - RIPng, O - OSPFv3,
       B - BGP, * - FIB route.
```

```
K* ::0                fe80::2::210:5aff:fe7a:b3d1
K* ::96               ::1
C* ::1/128            lo0
K* ::ffff:0.0.0.0/96  ::1
O* 2001:200:0:840::/64 fe80::210:5aff:fe7a:b3d1
O* 2001:200:0:822::/64 fe80::210:5aff:fe78:cae
C* 2001:200:0:823::/64 fxp0
O 2001:200:0:823::/64 fxp0
C* fe80:2::/64        fxp0
C* fe80:3::/64        x10
C* fe80:7::/64        lo0
```

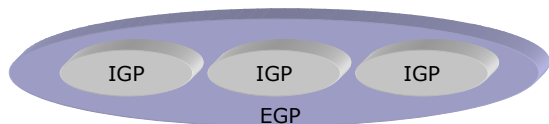
(type, prefix, next hop)

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



経路制御の階層化

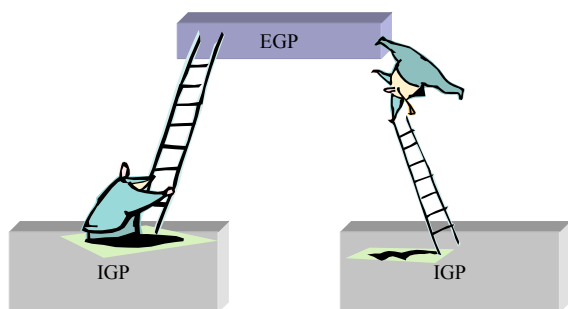
- 🌟 ルーティングドメイン (routing domain)
 - 境界を定義
 - 障害の分離、経路の集約(後述)などの効用
- 🌟 IGP (Interior Gateway Protocol)
- 🌟 EGP (Exterior Gateway Protocol)



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Two-tier routing



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



IGP and EGP

- 🌟 IGP
 - RIP-2, RIPng: distance vector
 - OSPF, IS-IS: link state
 - 各リンク・ルータの状態をできるだけ早く伝播させることに主眼
- 🌟 EGP
 - BGP4, BGP4+: path vector
 - Internet 全体の経路制御の安定に主眼

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



パスベクトル型経路制御

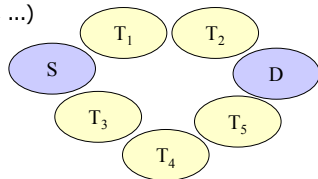
- 🔴 Bellman-Ford algorithm からの派生形
- 🔴 距離ベクトル型で交換する情報:
 - (prefix, metric)
- 🔴 パスベクトル型で交換する情報:
 - (prefix, path, attributes)
- 🔴 距離だけでなくパス情報を経路情報に付与
 - ループの起きない経路制御が可能

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Path vector

- 🔴 SからみたDへの経路の表現
 - $(D, \{T_1, T_2\}, \dots)$
 - $(D, \{T_3, T_4, T_5\}, \dots)$

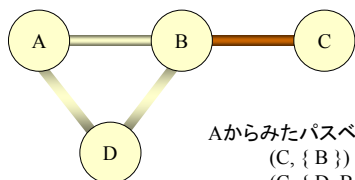


- 🔴 パスベクトル長が距離として用いられる
 - パスベクトル長の短い経路を優先

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Loop-free nature of path-vector routing



Aからみたパスベクタ:
 $(C, \{B\})$
 $(C, \{D, B\})$

- リンクB-Cが落ちた場合、Bがパスベクタを取り消し
- ループを回避
- 自らが含まれるパスベクタを棄却

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



BGP Border Gateway Protocol

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



BGP: Border Gateway Protocol

- ✿ パスベクトル型の経路制御プロトコル
- ✿ トランスポート層プロトコルとしてTCPを利用
 - 再送、確認応答はTCPに委ねる
- ✿ 効率的なパスベクトルの表現
- ✿ 経路情報の差分を伝播
- ✿ 距離情報だけでなくポリシーを表現可能

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



パスベクトルの表現

- ✿ AS (Autonomous System) を番号で表現
 - AS: 一貫したポリシーを持ったルーティングドメイン
- ✿ (prefix, AS-path, attributes) をまとめて符号化
 - (AS-path, attributes, { prefix1, prefix2, ... })
 - 通信量を削減

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



パス属性 (Path attributes)

- AS_PATH
- ORIGIN
- NEXT_HOP
- MULTI_EXIT_DISC (MED)
- LOCAL_PREF
- ATOMIC_AGGREGATE
- AGGREGATOR

©2004 Suguru Yamaguchi and Youki Kadohayaishi. All rights reserved.

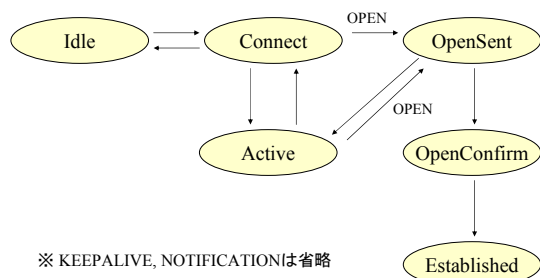


BGP route example

```
show ip bgp 163.221.0.0
BGP routing table entry for 163.221.0.0/16, version 30149334
Paths: (6 available, best #3)
Not advertised to any peer
6461 2516 2500 2500
  208.185.175.169 from 208.185.175.169 (216.200.254.220)
    Origin IGP, metric 0, localpref 400, valid, external
6461 2516 2500 2500, (received-only)
  208.185.175.169 from 208.185.175.169 (216.200.254.220)
    Origin IGP, metric 0, localpref 100, valid, external
2500, (received & used)
  203.181.70.232 (metric 2) from 203.181.70.232 (203.181.70.227)
    Origin IGP, metric 505, localpref 400, valid, internal, best
    Originator: 203.181.70.227, Cluster list: 0.0.0.16
2500, (received & used)
  203.181.70.233 (metric 2) from 203.181.70.233 (203.181.70.227)
    Origin IGP, metric 505, localpref 400, valid, internal
    Originator: 203.181.70.227, Cluster list: 0.0.0.16
2500
  202.249.2.1 from 202.249.2.1 (203.178.136.4)
    Origin IGP, metric 1010, localpref 400, valid, external
2500, (received-only)
```



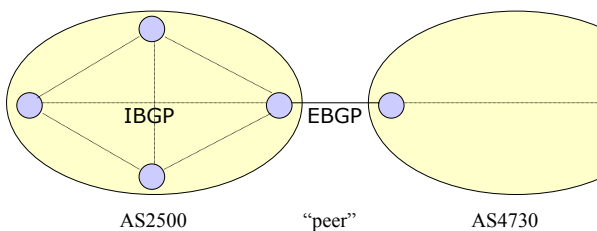
BGPの状態遷移



©2004 Suguru Yamaguchi and Youki Kadohayaishi. All rights reserved.



IBGP and EBGP



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



IBGP, EBGP の構成要件

IBGP

- 同一AS内のすべてのIBGPルータとメッシュ状に peer を確立しなければならない
- 物理的に隣接する必要はない
- すべてのIBGPルータが同じ情報を持つ

EBGP

- 隣接ASのEBGPルータと物理的に隣接しなければならない（原則として）
- 隣接ASと同じ情報を持つ必要はない

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.

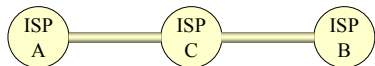


ポリシルーティング Policy Routing

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



トポロジ vs ポリシ: Transit AS, stub AS



ISP C: Transit AS



User: Stub AS

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



経路制御におけるポリシ

■ 広告ポリシ

- ・ 隣接ルータに広告してよい経路
- ・ 告知すべきでない経路
- ・ 広告前の、メトリック、パス属性の操作

■ 受け入れポリシ

- ・ RIBに反映してよい経路、反映すべきでない経路

■ 採用ポリシ

- ・ 複数のRIBのうちどれを優先しFIBに採用するか

■ 再配布ポリシ

- ・ RIB間での経路の変換、再配布
- ・ 再配布してよい経路、再配布時のメトリック

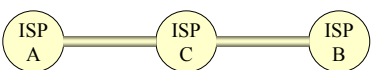
©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



トランジットASのポリシ

■ ISP C:

- Accept all routes from ISP A
- Advertise all routes to ISP A
- Accept all routes from ISP B
- Advertise all routes to ISP B



ISP C: Transit AS

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Stub ASのポリシー

- ☀ Accept ISP A's route from ISP A
- ☀ Advertise user's routes to ISP A
- ☀ Accept ISP B's route from ISP B
- ☀ Advertise user's routes to ISP B



User: Stub AS

©2004 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.



ルータにおけるポリシーの適用

- ☀ match: ポリシ適用対象とする経路の選択
 - AS-path, パス属性、アドレス...
- ☀ set: 動作の指定
 - 経路選択順位の変更
 - next hop ルータの変更
 - パス属性の付与、変更
- ☀ match-set節の羅列でポリシーを記述

©2004 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.



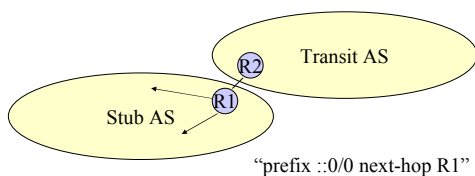
経路の集約 Aggregation of routes

©2004 Suguru Yamaguchi and Youki Kadobayashi, All rights reserved.



Default route

- ✿ 0.0.0.0/0 (IPv4), ::0/0 (IPv6)
- ✿ Longest-match → 経路検索の最後にマッチ
- ✿ 経路の隠蔽、経路数の削減に効果

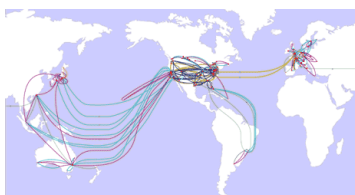


©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Default-free

- ✿ Default routeがない状態
- ✿ Tier-1 ISPs, North America backbone



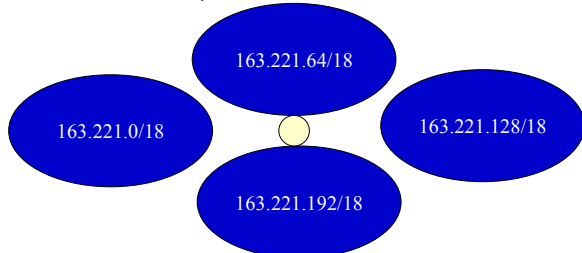
(source: UUNet network maps, www.uu.net)

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



経路の集約 (aggregation)

- ✿ 163.221.10.0/24 と 163.221.11.0/24 は 163.221.10.0/23 で表現できる

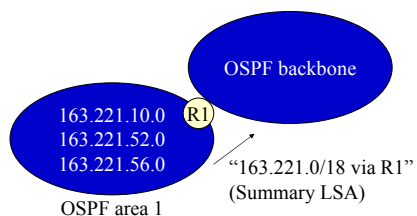


©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Aggregation at the border

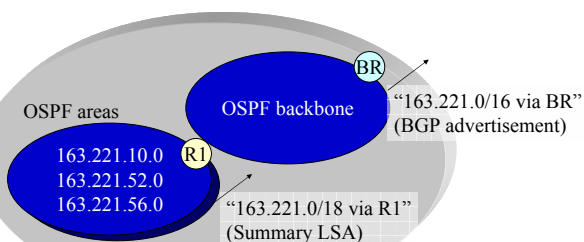
- 経路の集約はルーティングドメインの境界で



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



Aggregation at multiple levels



©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



経路集約の問題点

- 集約できるかどうかはアドレス割り当て次第
 - 計画的なアドレス割り当てが重要
- 奈良先端が4研究科になると予想できたか？
 - あるISPが急成長すると予想できたか？
- prefix renumbering
 - 集約できるアドレスへ付け替え
 - 技術開発がおこなわれている
- 経路集約は手作業でおこなう必要がある
 - インターネット全体で170,000経路
 - BGP table growth trends - [Telstra](#)

©2004 Suguru Yamaguchi and Youki Kadohbayashi, All rights reserved.



ここまでのまとめ

- 🌟 階層化経路制御の考え方
 - IGP, EGP
- 🌟 パスベクトル型経路制御
- 🌟 BGP
 - パス属性, AS path, AS番号
 - IBGP, EBGP
- 🌟 Transit, Stub
 - Default, default-free
- 🌟 経路集約 – 考え方、問題点

©2004 Suguru Yamaguchi and Youki Kadohayaishi. All rights reserved.
