



# Inter-IX技術としてのMPLS

---

日本テレコム(株)

松嶋 聡

[satoru@ft.solteria.net](mailto:satoru@ft.solteria.net)



# 本日の内容

---

- Inter-IXって？
- 動機など
- どうやって動くの？
- 考えなければいけないこと
- 設定例
- やってわかった普通のIX/MPLS-IXの違い
- まとめ



# Inter-IXって？

---

- Inter-IX技術？
  - IXとIXをつなげよう！
- 管理者の違うセグメントをつなぐ
  - そうカンタンにはつなげないでしょう？
- L2ネットワークを安全に繋ぎたい
  - MPLSだよな



# 動機など(1)

---

- なぜそんなことがやりたいのか？
  - 技術的に
    - いままでそんな試みがない
    - 従来とは異なるIX実現技術の登場(MPLS-IX)
  - 現在のインターネットを考えてみるに
    - IXはみんなの出入口. つながる先が増えるとうれしい.
    - 階層をフラットにして、Resiliency を高めよう！



## 動機など(2)

---

- 安全にIXをつなぎ合わせるむずかしさ
  - 壊れやすいEthernet
    - 異なる管理者 / 同じセグメント
    - ブロードキャスト / マルチキャストもどんどんやってくる
    - IP/MACアドレス重複
- Dot1q VLAN?
  - 組み合わせ爆発問題
  - そもそも使っていないIXも



## 動機など(3)

---

- MPLSをIXの“つなぎ”に
  - Ethernetから見て、MPLSパケットは単なるL3パケット
  - レイヤ3ネットワークでセグメントを分割
  - LSPはIX間にまたがり、シームレスなピアリング

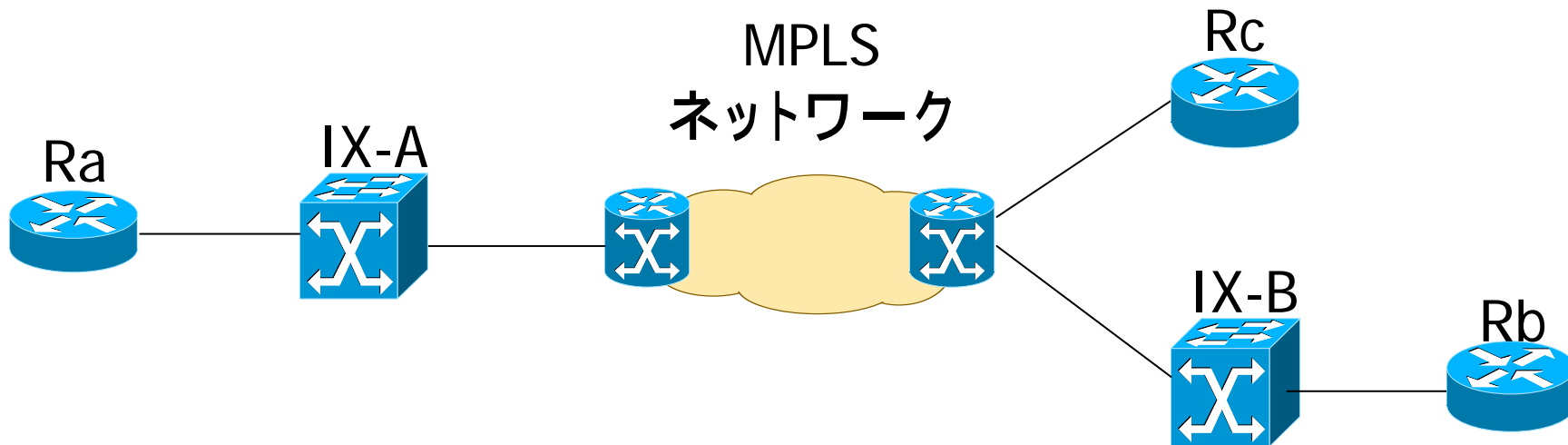


## 動機など(4)

---

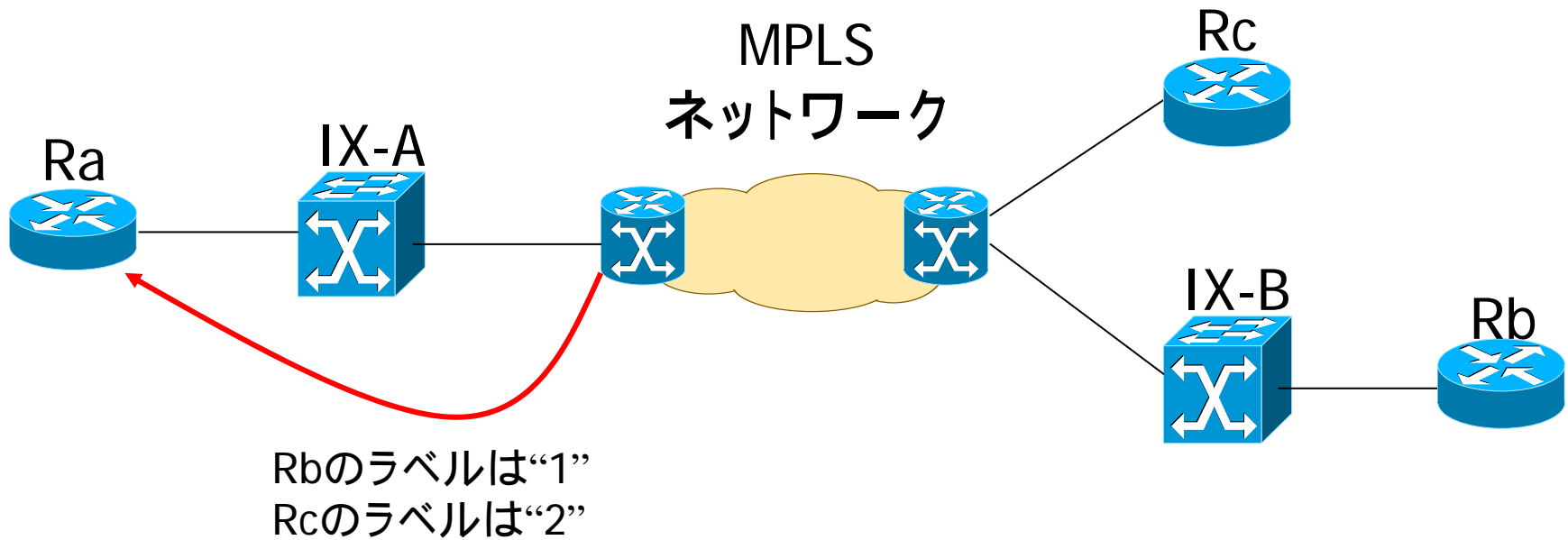
- やってみよう！
- やります！
  - WIDE DIX-IEと、JT mplsASSOCIO

# どうやって動くの？

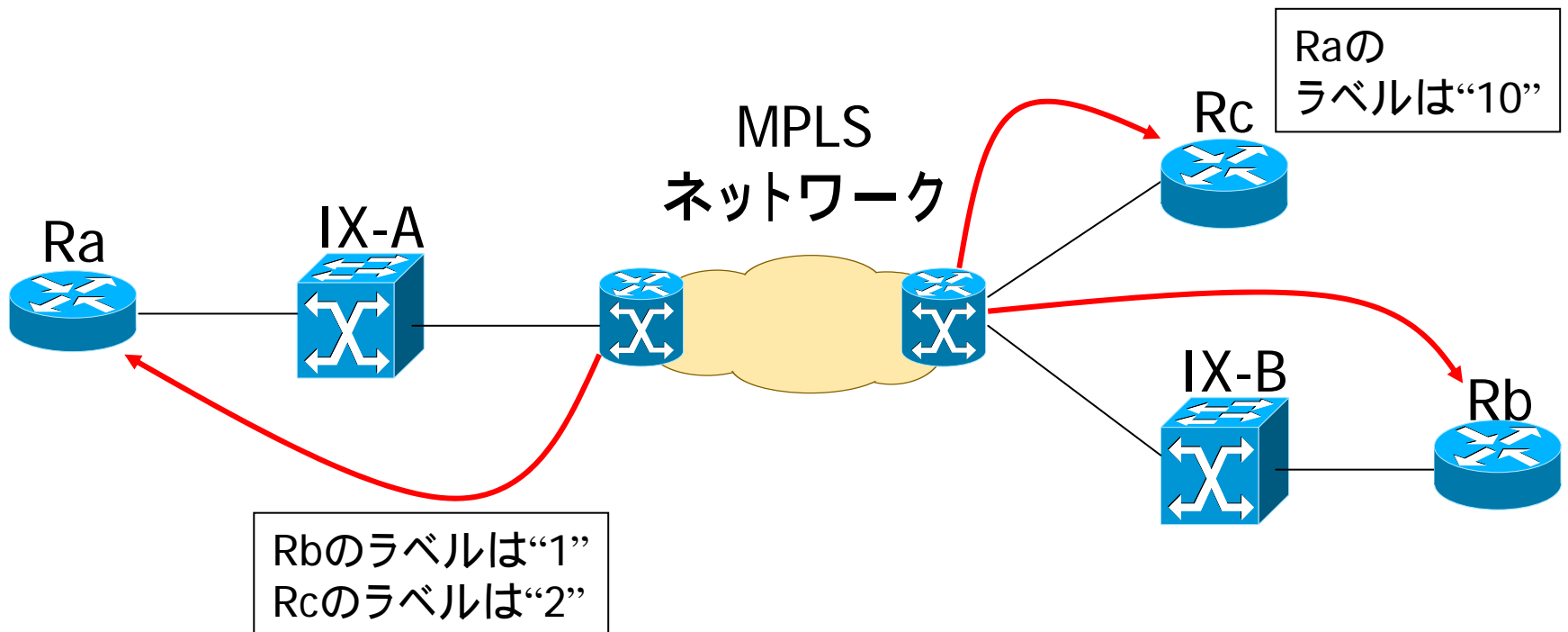




# どうやって動くの？

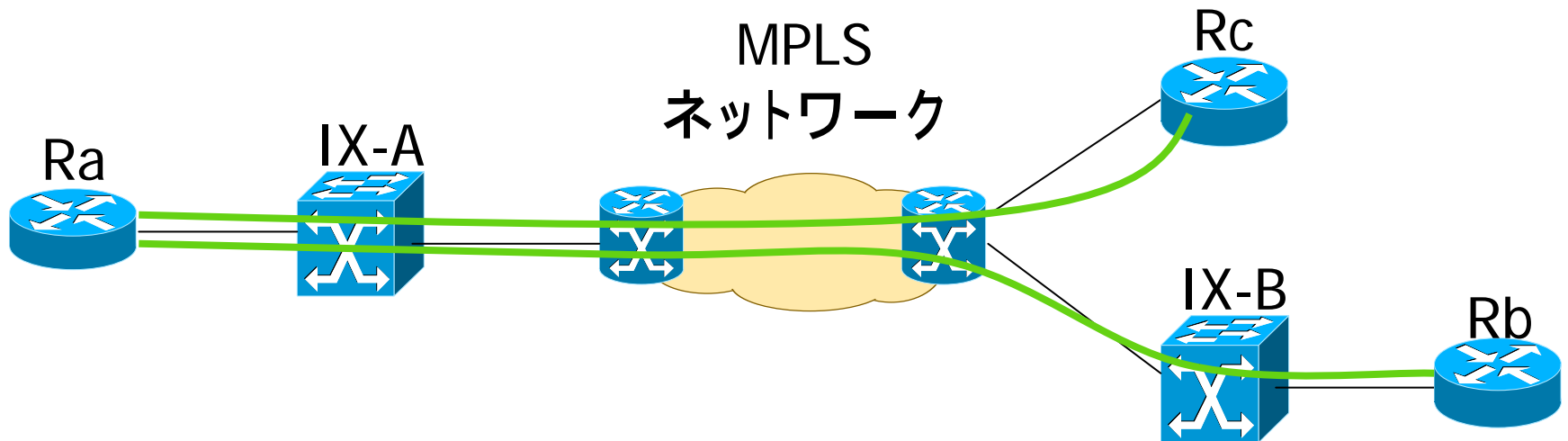


# どうやって動くの？



# どうやって動くの？

LSP上でBGPピアリング





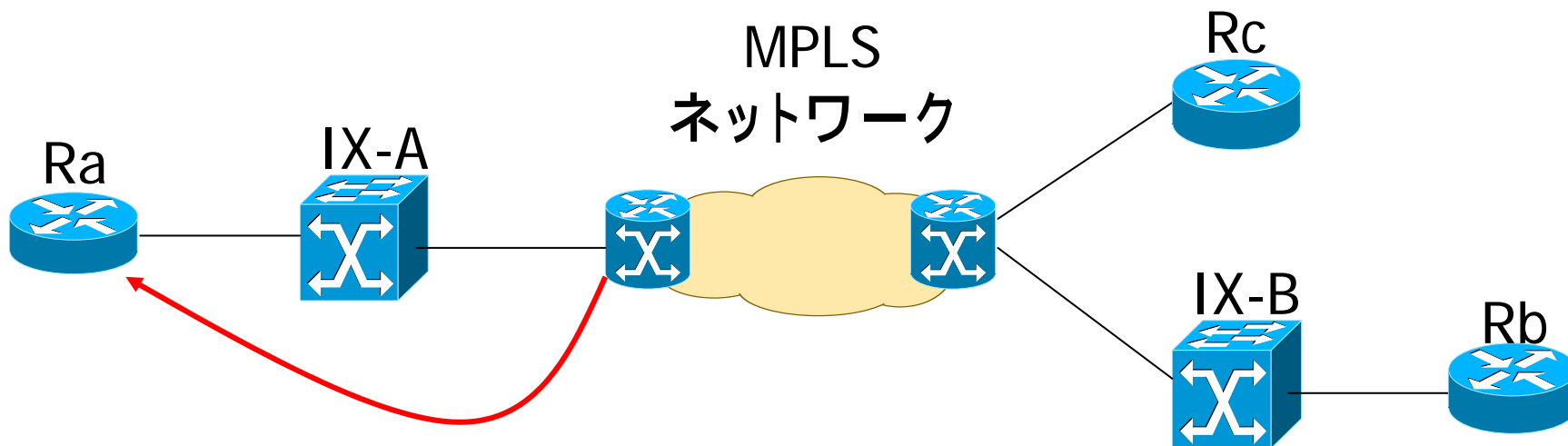
# 考えなければいけないこと

---

- **安全な接続**
  - ブロードキャストドメインを分けられること
  - 少なくともラベル一個分(4 byte)足せるMTU
  - Broadcast/Multicastを不用意に増やさない
  - IX間をシームレスにピアが行き交えること
  - 知らない人からパケットを受け取らない

# 考えなければいけないこと

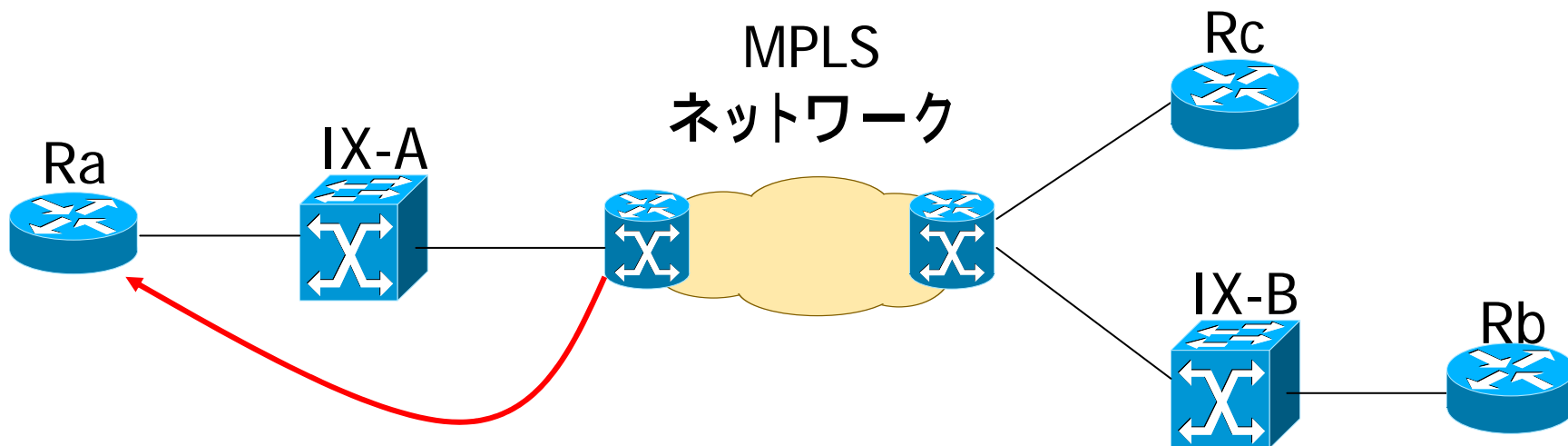
## ■ Broadcast/Multicast?



ラベル配布プロトコル  
・LDP, RSVP-TE, BGP?

# 考えなければいけないこと

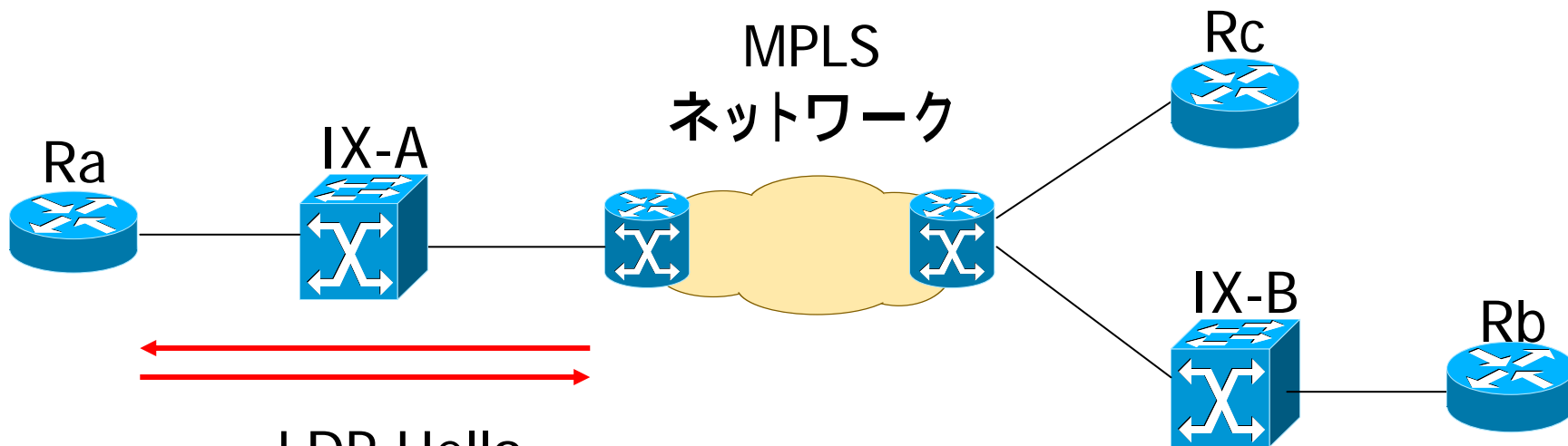
## ■ Broadcast/Multicast?



ラベル配布プロトコル  
・LDP, RSVP-TE, BGP?

# 考えなければいけないこと

- LDPは隣接発見にマルチキャストを使う
  - 多くのルータは、5秒に1回。100台いたら？
  - 使うアドレスはみんな反応するアドレス
  - Targeted helloもあるけれど...

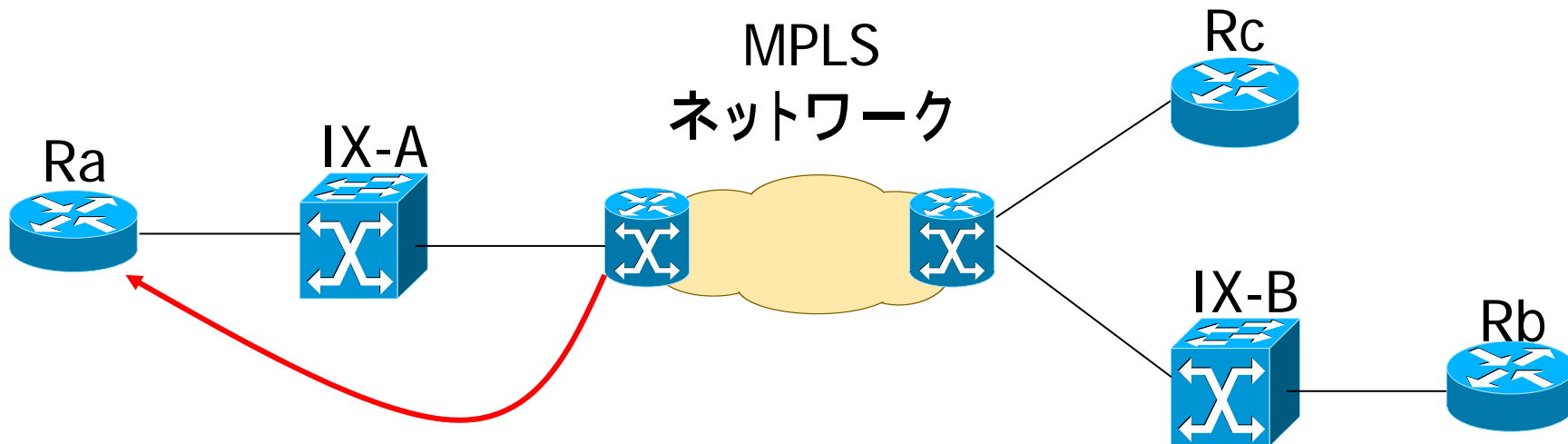


LDP Hello

(宛先: All router on this subnet)

# 考えなければいけないこと

- Multicastを使わないラベル配布プロトコル
  - RSVPはちょっと...
  - BGPはラベルも運べる

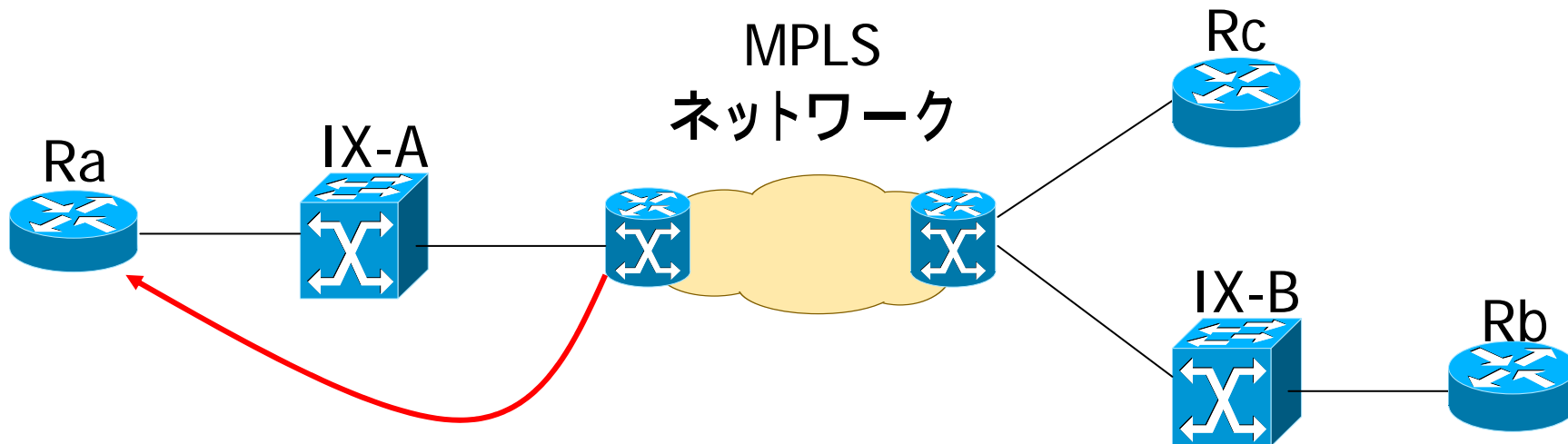


ラベル配布プロトコル  
・LDP, RSVP-TE, BGP?



# 考えなければいけないこと

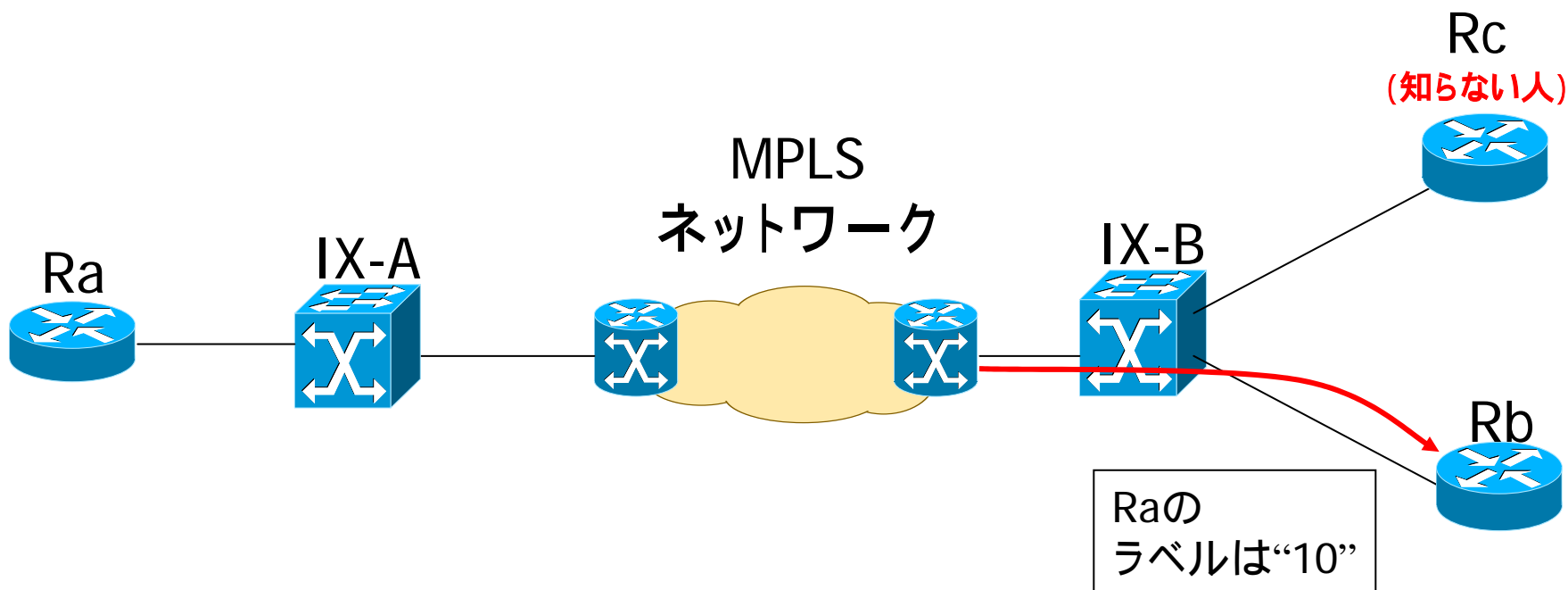
- RFC3107 Carrying label info, in BGP-4
  - Capability value -> AFI:1, SAFI:4
  - Labeled unicast MP-NLRI -> AFI:1, SAFI:4



MPBGP UPDATE  
(Label + Rb, Rc FECプレフィクス)

# 考えなければいけないこと

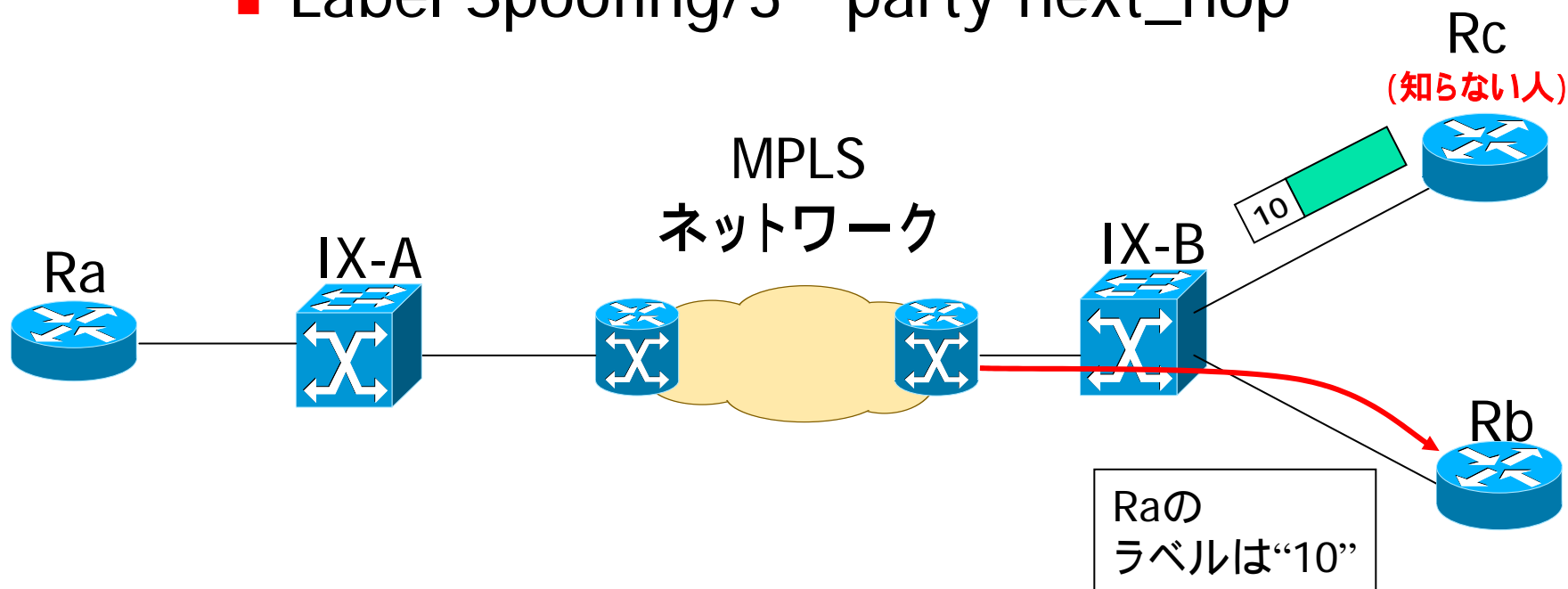
- あなた誰？



# 考えなければいけないこと

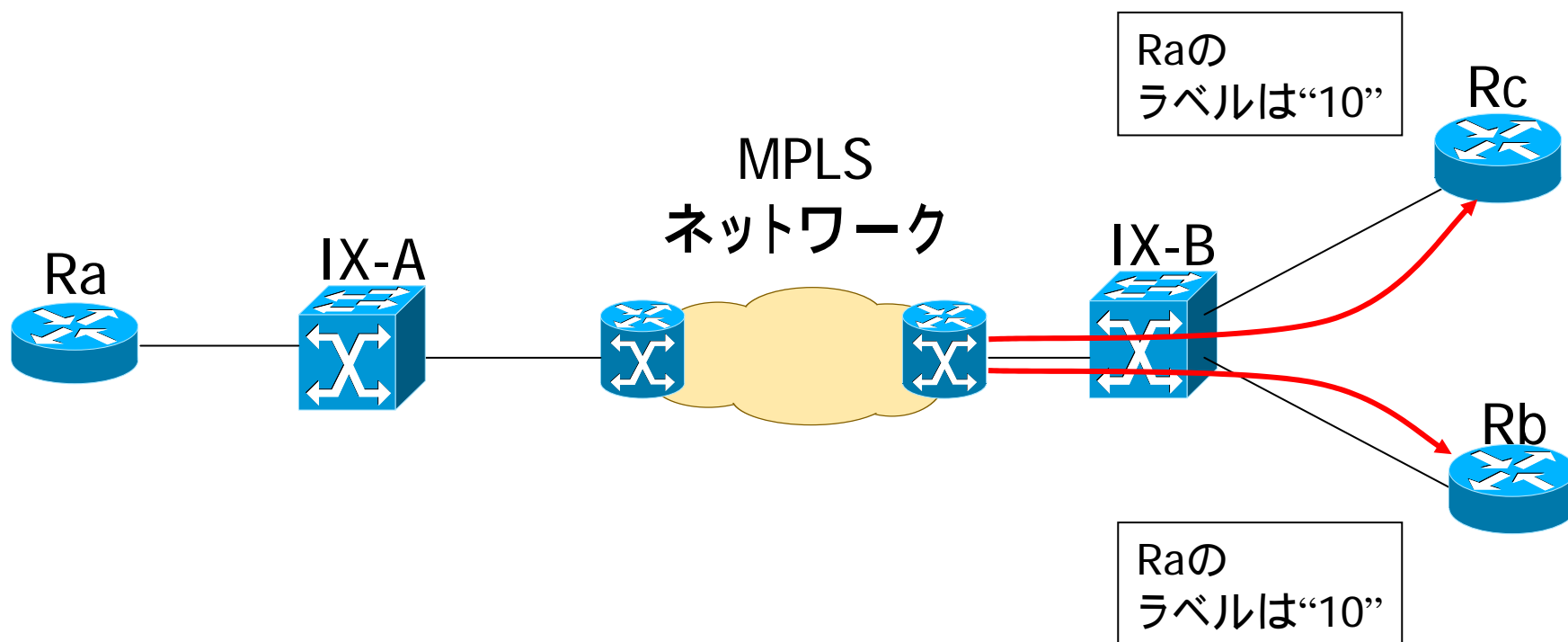
- あなた誰？

- Label Spoofing/3<sup>rd</sup> party next\_hop



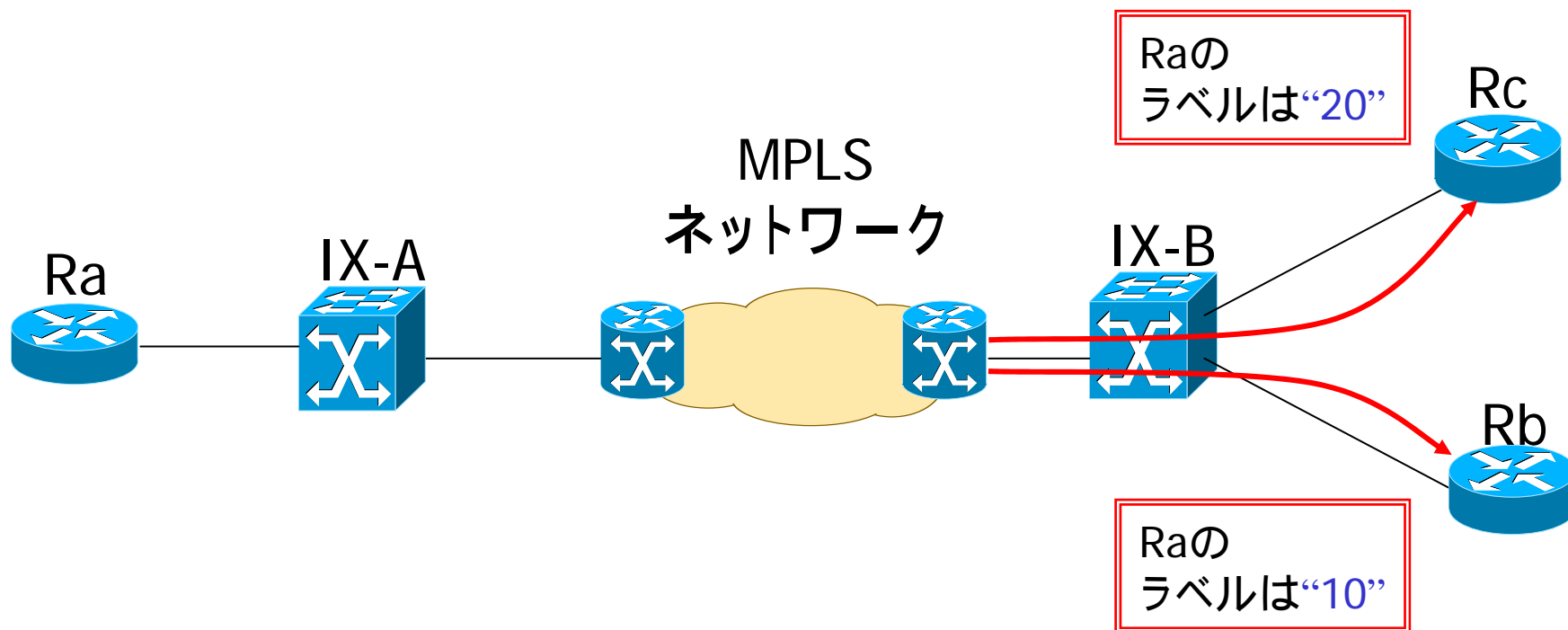
# 考えなければいけないこと

- 誰からのパケットが分からない



# 考えなければいけないこと

- 誰からか、分かるようにする
- 知らない人から受け取らない



# 設定例(1) 使えます



## Raの設定 (Cisco)

```
no mpls ip propagate-ttl
!  
interface gigabitethernet 1/0  
 ip address 1.1.1.2 255.255.255.0  
 mpls mtu 1504  
!  
router bgp 1  
 no bgp default ipv4-unicast  
 neighbor 1.1.1.1 remote-as 9592  
 neighbor 10.0.0.1 remote-as 2  
 neighbor 10.0.0.1 disable-connected-check  
!  
 address-family ipv4  
 neighbor 1.1.1.1 activate  
 neighbor 1.1.1.1 send-label  
 exit-address-family  
!  
end
```

# 設定例(2) 使えます



## Raの設定 (Juniper)

```
ge-0/1/0 {
  unit 0 {
    family inet {
      mtu 1500;
      address 1.1.1.2/24;
    }
    family mpls {
      mtu 1504;
    }
  }
}
```

```
protocols {
  mpls {
    no-propagate-ttl;
    interface all;
  }
  bgp {
    group mpls {
      type external;
      local-address 1.1.1.2;
      family inet {
        labeled-unicast;
      }
      peer-as 9592;
      neighbor 1.1.1.1;
    }
  }
}
```

```
group as-2 {
  type external;
  multihop {
    ttl 1;
  }
  local-address 1.1.1.2;
  export as-2-policy;
  peer-as 2;
  neighbor 10.0.0.1;
}
```

# 設定例(3) どんな感じ？



## Raの気持ち (Cisco)

```

Router#sh ip route
202.249.2.0/24 is variably subnetted, 2 subnets, 2 masks
C      1.1.1.2/32 is directly connected, GigabitEthernet1/0
C      1.1.1.0/24 is directly connected, GigabitEthernet1/0
      10.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
B      10.0.0.1/32 [20/0] via 1.1.1.1, 01:06:47

Router#show mpls for 10.0.0.1
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag    tag or VC  or Tunnel Id    switched  interface
None   23        10.0.0.1/32    0         Gi1/0     1.1.1.1

Router#show ip bgp label
Network          Next Hop        In label/Out label
10.0.0.1/32     1.1.1.1         nolabel/23
    
```



# 設定例(4)どんな感じ？



Raの気持ち  
(Juniper)

```
jt@viikrt02-g1> show route

inet.0:
+ = Active Route, - = Last Active, * = Both

10.10.10.0/24      *[BGP/170] 03:49:37, localpref 100, from 10.0.0.1
                   AS path: 2 I
                   > to 1.1.1.1 via ge-0/1/0.0, Push 24
10.0.0.1/32       *[BGP/170] 03:49:52, localpref 100
                   AS path: 9592 ?
                   > to 1.1.1.1 via ge-0/1/0.0, Push 24

mpls.0: 3 destinations, 3 routes (3 active, 0 holddown, 0 hidden)
```



# やってみてわかったこと

---

- Cisco
  - BGPの設定は、address-family manner要
    - labeled ipv4 unicast はMPBGP
- Juniper
  - traffic-engineering bgp-igp-both-ribs  
はいらない。
    - Label付きBGP経路は、直接 “inet.0” に入る
  - IPv4 UnicastはCapability広報しない



# まとめ

---

- IXをつなぐ技術としてMPLSはとても合理的
  - セグメントを分けて安全に接続
  - MPLS-IX、L2IXなんでもこい
- IX Switch, IX接続ルータのケア
  - Broadcast/Multicastをなるべく使わず穩便に
- MPLS-IX側のケア
  - Label Spoofing/3rd party next\_hop防止
  - LSP毎のトラフィック計測
- ラベル配布はBGPが有望
  - RFC3107 Labeled Unicast
  - やっぱりIXはBGP



# 質問など

---