



ネットワークの 高速切り替え手法の検討

IX上での切断検知をきちんと高速にしたい

NTTコミュニケーションズ
鈴木昭徳

今回の発表内容は、JANOG17での発表とその後の
検討状況や機器への実装状況を踏まえて、お話します。

アジェンダ

- 自己紹介
- きっかけ
- 切替ステップ
- **BFD**
- コンバージョン
- **Ether-OAM**ってどう？
- まとめ

アジェンダ

- **自己紹介** 
- きっかけ
- 切替ステップ
- **BFD**
- コンバージェンス
- **Ether-OAM**ってどう？
- まとめ

自己紹介

(少し前)

- **NTTコムAS7515/2914時代**
 - バックボーン的设计・構築・運用してた
 - 国際回線開通したり
 - 夜中電話で起こされたり

(今)

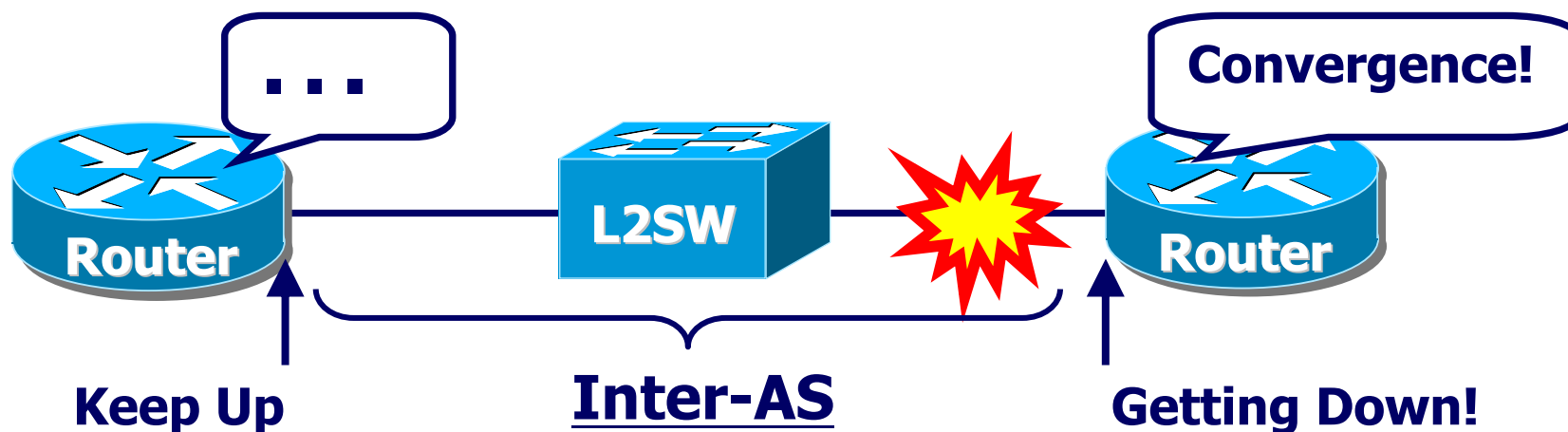
- 現場を離れて、技術戦略系(国内)お仕事
 - 新技術の目利き
 - **MPLS/GMPLS**技術の新網の玉込め
 - 営みを社外に発信する

アジェンダ

- 自己紹介
- **きっかけ** 
- 切替ステップ
- **BFD**
- コンバージョン
- **Ether-OAM**ってどう？
- まとめ

きっかけ

L2スイッチがあると、故障時リンクが落ちないので、すぐには切替しない。とくに**IX**。
このような場合は、ルーティングプロトコルの**Hello**にたよっていた。数十秒の切替。



きちんと検知! 早く切替!

アジェンダ

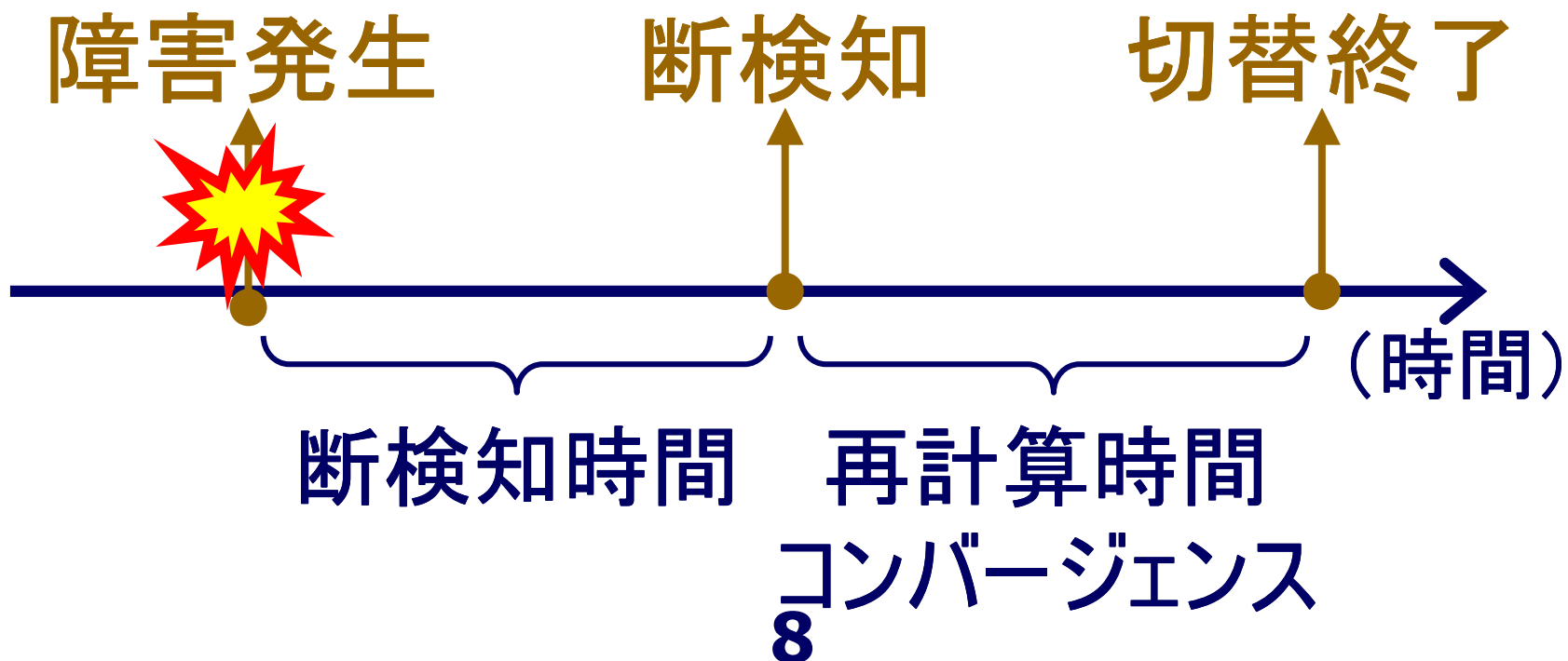
- 自己紹介
- きっかけ
- **切替ステップ** 
- BFD
- コンバージェンス
- **Ether-OAM**ってどう？
- まとめ

切替ステップ-1

大まかにいうと

“切替時間”

= “断検知時間” + “再計算時間”



切替ステップ-2

● 断検知時間短縮へのアプローチ

- ・ **Hello**のチューニング
- ・ **BFD**

あんまり想像
できない?

● 再計算時間短縮へのアプローチ

- ・ **IP Fast Reroute?**
- ・ **BGP Convergence**

ベンダの実
装に依存

Optimization/GRIP?

アジェンダ

- 自己紹介
- きっかけ
- 切替ステップ
- **BFD** 
- コンバージョン
- **Ether-OAM**ってどう？
- まとめ

BFD-1

Bidirectional Forwarding Detection

- ・ 1秒以下で障害を検知することのみを目的
- ・ HelloやKeepaliveと同じ正常性 Check
- ・ BFDパケットのミリSec単位の送受信
- ・ IPレイヤ上で動作。
- ・ IGP、BGP、RSVPなどに断情報を通知
- ・ 送受信間隔の交渉は、値の高いほうを選択

BFD-2

3つの設定パラメータ

- **min-TX-interval**
自身が許容する最小の送信間隔 (msec)
- **min-RX-interval**
自身が許容する最小の受信間隔 (msec)
- **multiplier(ホールドタイム)**
受信間隔に対する乗数(整数値)

ホールドタイム

対向のTX (ネゴ後)と、
multiplier値の乗算



12

送信インターバル

自身のTXと、
対向のRXを比較
値の大きいほう

BFD-3 Cの設定

```
interface GigabitEthernet 1/0
```

```
ip address x.x.x.x y.y.y.y
```

```
bfd interval 50 min_rx 50 multiplier 3
```

```
router bgp zzz
```

```
neighbor x.x.x.x fall-over bfd
```

```
router ospf zzz
```

```
network x.x.x.x y.y.y.y area 0
```

```
bfd all interfaces
```

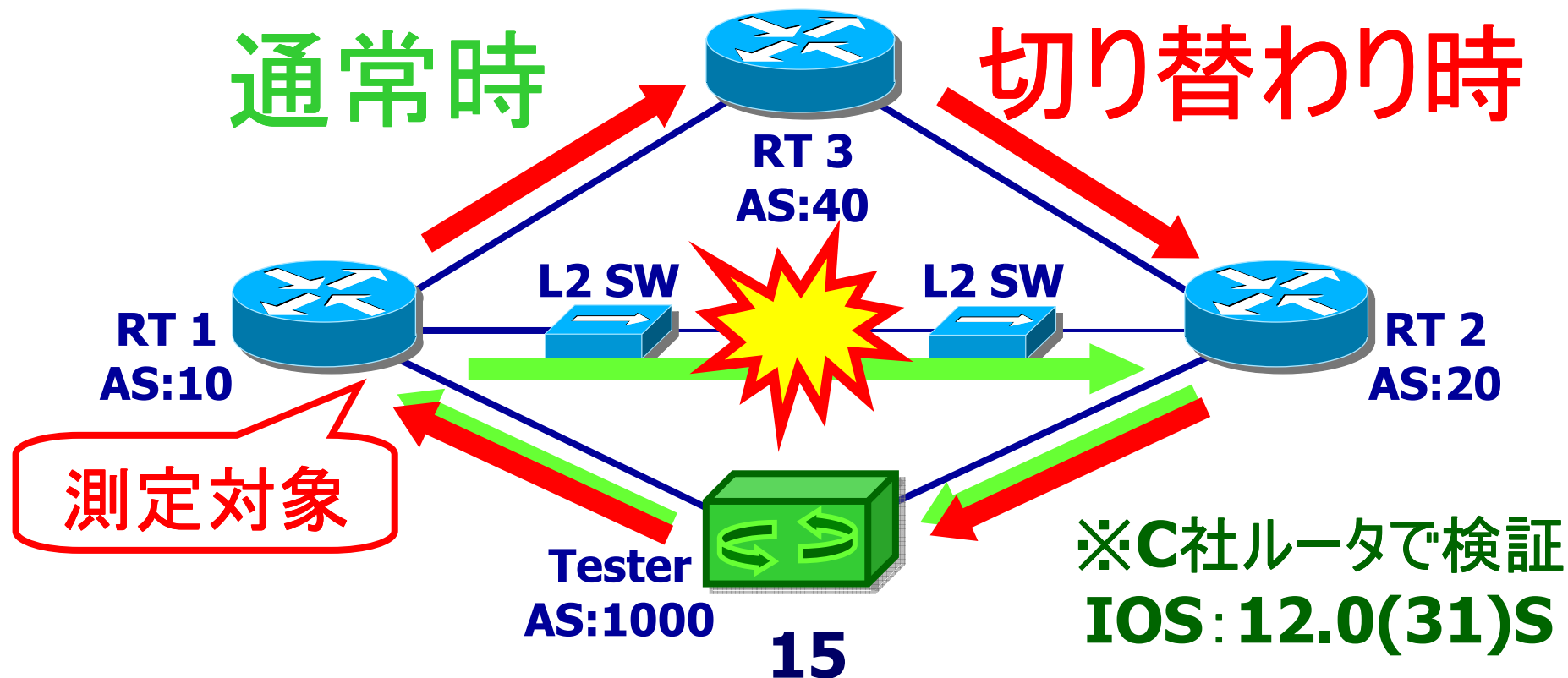
BFD-4 Jの設定

```
[protocols ospf area 0]
```

```
interface ge-1/0/0.0 {  
    bfd-liveness-detection {  
        minimum-interval 50;  
        minimum-receive-interval 100;  
        minimum-transmit-interval 200;  
        multiplier 5;  
    }  
}
```

BFD-5

IXを想定し、BGPで検証してみました。
”Keepaliveを最小”と、
“BFDを用いた場合”



BFD-6

**BFDの
測定結果**

- ・ **10ミリ秒/3回**でも問題なし。
- ・ ばらつき・・・精度? コンバージェンスのブレ?

multiplier値 Interval時間 (m秒)	3回		5回		10回	
	実測値	平均値	実測値	平均値	実測値	平均値
10	0.923	0.544	1.099	0.694	0.262	0.460
	0.318		0.534		0.816	
	0.392		0.449		0.302	
50	0.923	1.039	0.908	0.717	1.001	0.820
	1.191		0.568		0.697	
	1.003		0.674		0.763	
100	0.353	0.843	0.856	0.907	1.761	1.741
	1.248		0.933		1.590	
	0.928		0.933		1.873	
500	1.079	1.020	3.305	2.854	5.426	5.053
	1.267		2.709		4.526	
	0.713		2.547		5.206	
1000	3.016	2.869	5.428	5.054	8.702	9.643
	2.318		5.299		10.693	
	3.272		4.434		9.534	

BFD-7

Keepalive
最短の測定結果

「**Keepalive**を短くすると不安定になるよ」と警告されるが、特に問題なし。

～みなさん、なん秒ですか？～

Interval時間 (秒)	切り替わり時間(秒)	
	実測値	平均値
BGP	2.872	2.887
Inter : 1秒	3.022	
Hold : 3秒	2.766	

ちなみに、**ISP**ではないですが、弊社**L2MPLS**網では、**ISIS**で、インターバル**2**秒、エクスパイア**6**秒です。

BFD-8

実装状況

C社

EIGRP

OSPF

BGP

ISIS

※いずれもv4

J社

OSPF

ISIS

Static

RSVP

LDP

※いずれもv4

個人的にはBGPがほしい...

8.1には入るとい噂が...

特に問題なかった。

他社の実装はあまり聞こえてこない...
フューチャーリリースの中に

アジェンダ

- 自己紹介
- きっかけ
- 切替ステップ
- BFD
- **コンバージェンス** 
- Ether-OAMってどう？
- まとめ

コンバージェンス へのアプローチ

実はあまり有効な案が無い?...

- ・ **IP Fast Reroute**

- トポロジによる

- **IGP**ベース

複雑...

ベンダ実装次第
中身が見えない

- ・ **BGP Convergence Optimization**

- **40%短縮?Black Box?**

～何か思いつきますか?...～

アジェンダ

- 自己紹介
- きっかけ
- 切替ステップ
- **BFD**
- コンバージョン
- **Ether-OAM**ってどう？
- まとめ



Ether-OAMってどう？-1

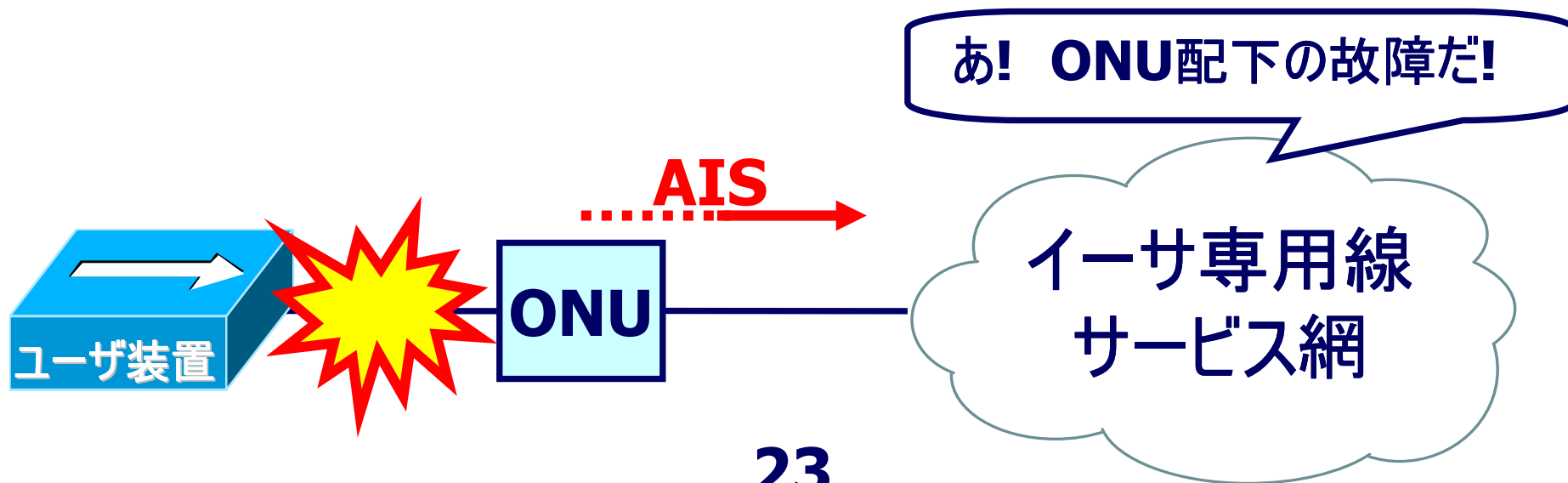
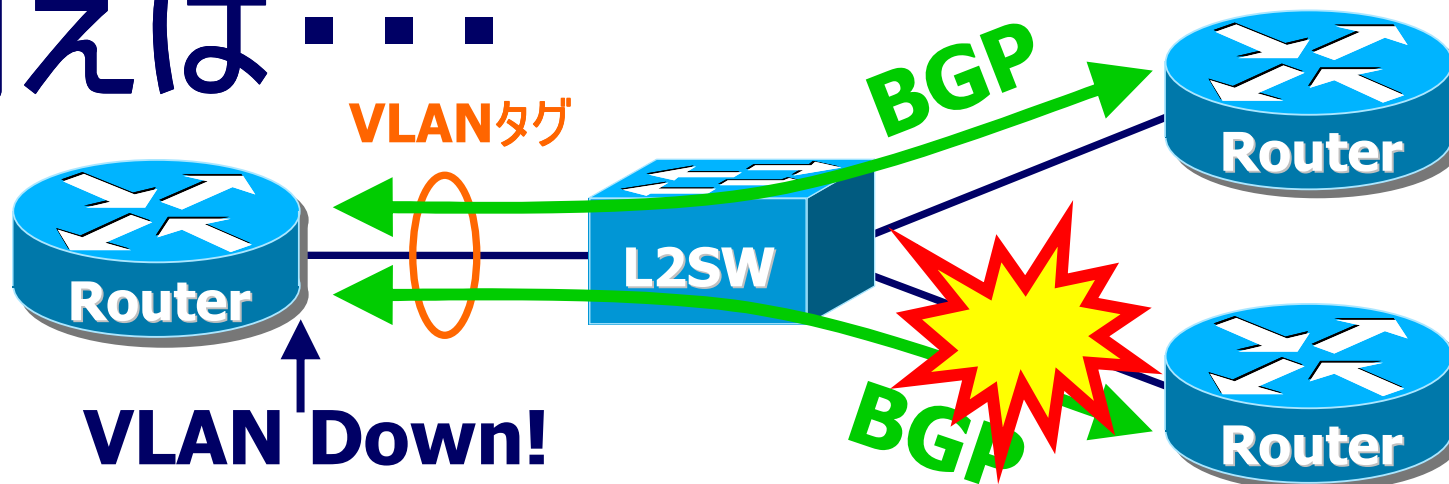
Y.1731や**802.1ag**

VLAN単位で次のことが出来る

- ・ **CC**・・・**Continuity Check**
- ・ **LB**・・・**Loopback**
- ・ **AIS**・・・**Alarm Indication Signal**
- ・ **LT**・・・**Link Trace**
- ・ **LM**・・・**Latency Management**
- ・ **DM**・・・**Delay Management**など

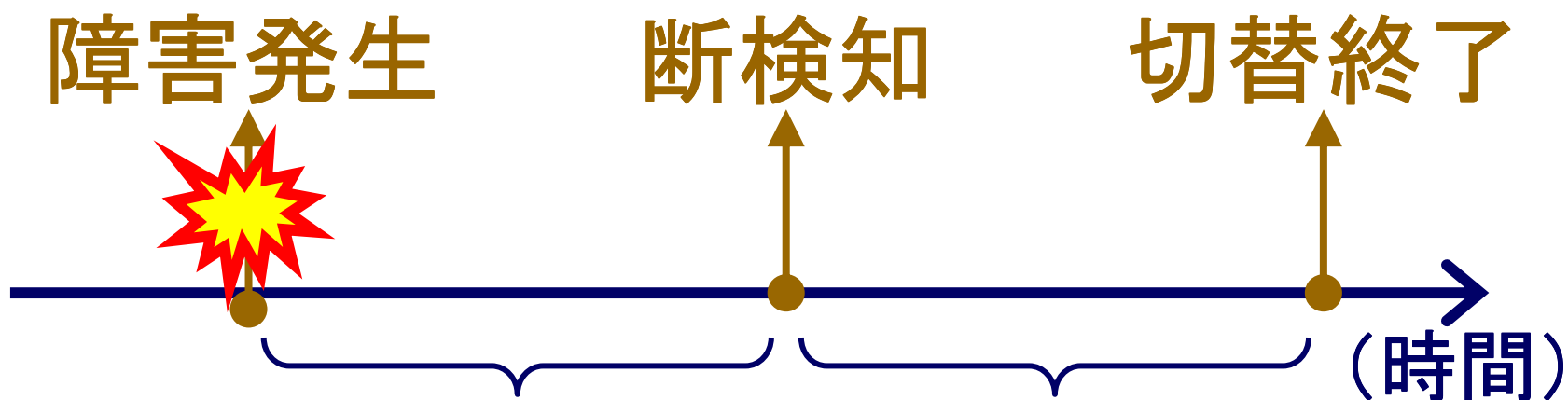
Ether-OAMってどう？-2

例えば...



Ether-OAMってどう？-3

でも前項も、断検知時間の短縮ですよね。



ここに有効な短縮
アプローチがほしいな

アジェンダ

- 自己紹介
- きっかけ
- 切替ステップ
- **BFD**
- コンバージェンス
- **Ether-OAM**ってどう？
- **まとめ**



まとめ

切替時間短縮を実現させるアプローチを考えると、断検知時間短縮は、具体的な実現方法が見えてきている。

コンバージェンス時間短縮へのアプローチは、網構成に依存したり、ベンダ各社の日々の努力に頼っている。増え続ける**BGP**経路などの状況を考えると、標準的な手法がほしい。

おまけ

～呼びかけ～

「IX環境の高度化」というテーマで色々やっています。

- ・ 高速切替
- ・ C/Dプレーンの分離など



ご興味の方は、是非我々まで

勝手に考えました

Pre-Computation Sec. Routing Table

