

# 高品質Inter-AS接続を実現するために

MPLS Japan 2003

Miya Kohno ([mkohno@cisco.com](mailto:mkohno@cisco.com))

Consulting Systems Engineer, Cisco Systems Japan

- **Introduction**
  - 問題提起
  - Inter-AS接続の分類
- **High Availability**
- **QoS**
- **Load BalancingとOptimization**
- **まとめ**

# Introduction – 問題提起

- ここ数年で、IP/MPLS網を高品質化・高性能化するための、多くの能拡張が行われた。
- しかし、ASを跨った途端、高品質化・高性能化のハードルは高くなる。何故なら、

そもそもInter-AS接続では、ScalabilityとStabilityが最重要課題。

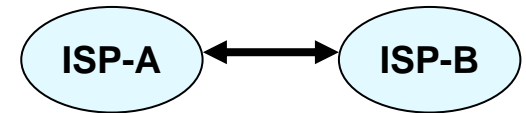
高速迂回、トラフィック最適化、QoS制御などはその次。

確かにハードルは高い。しかし、今後は？

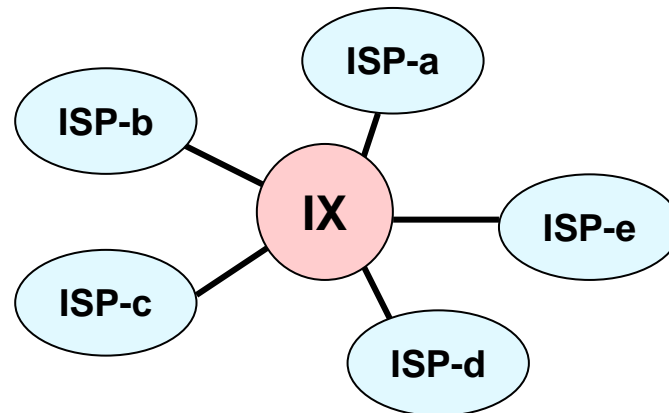
# Introduction – Inter-AS接続の分類 (1/2)

## 1. 事業者間接続

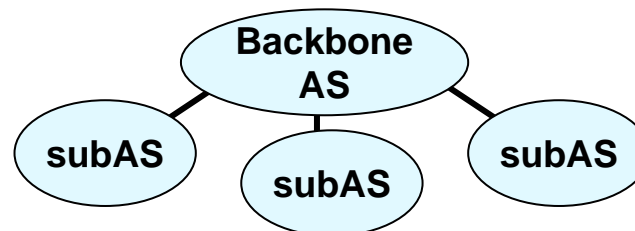
### 1-1 Bilateral Peering (Private Peering)



### 1-2 Multilateral Peering

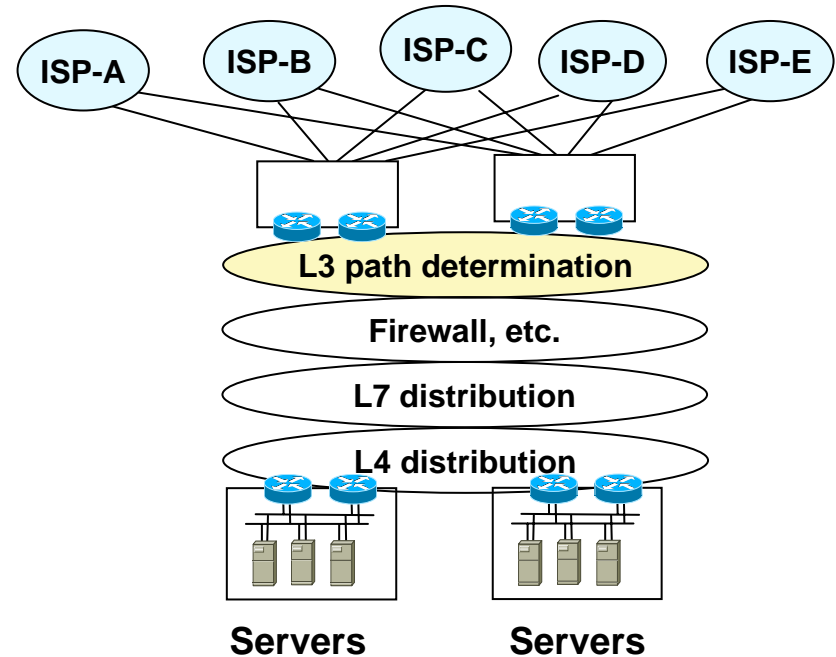


## 2. 同一事業者におけるAS分割 (主にScalability向上を目的とする。)

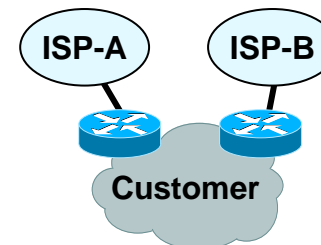


# Introduction – Inter-AS接続の分類 (2/2)

## 3. 大規模Contents Provider, ASP接続



## 4. SubscriberによるMultihome接続



- **SSO/NSF**
  - Stateful Switchover + Protocol Graceful Restart
- **Routing Convergence**
- **TE FRR**

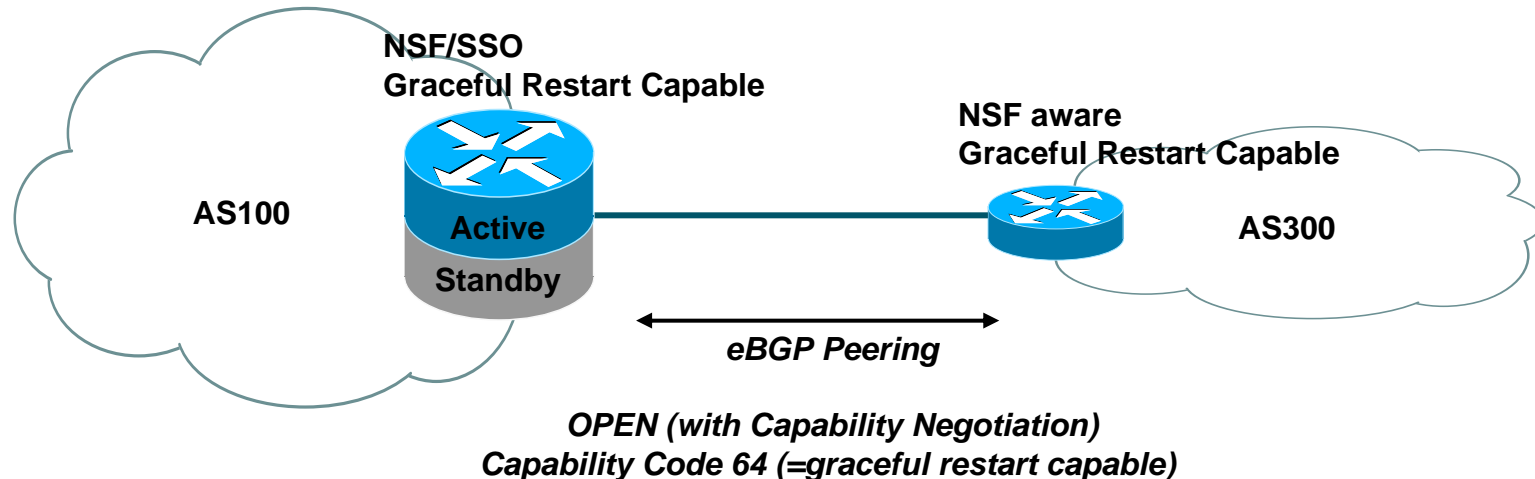
# High Availability

## – SSO/NSF (Stateful Switchover + Protocol Graceful Restart)

Cisco.com

- SSO/NSF (障害時も状態を保持したまま転送を続行する)と、他のHigh Availability実現のための技術 (できるだけ即座に障害を検知する、障害を検知したら速やかに経路を切り替える)は、二律背反的であるため、使用方法に注意する必要がある。

- 適所に上手く使えば効果的



# Routing Convergence – IGPの場合

## • 障害の検知

- HWによる検出 (LOS, L-AIS/RDI, P-AIS/RDI, SD/BER)
  - Carrier delayのtuning
- IP Event Dampening 12.0(22)S
- Fast hello 12.0(23)S
- BFD(Bidirectional Forwarding Detection) planned

## • LSA packetの生成

- 最初の生成は即座に with Exponential Backoff 12.0(25)S

## • LSA flooding

- pacing間隔の調整 12.0(25)S



# Routing Convergence – IGPの場合



- **SPF計算**

- **SPF delay with Exponential Backoff** 12.0(23)S
- **iSPF(incremental SPF)** 12.0(24)S

- **RIBのアップデート**

- **Prefix Ordering** (not in OSPF at this moment)

**Stabilityを損なうことなく、Fast Convergenceを行う。**

**網の規模にもよるが、数100msec以内のConvergenceが可能。**

# Routing Convergence – BGPの場合

## • 障害の検知

- Peer Downによる検出
  - 物理リンク断 fast external failover
  - BGP hold timer timeout
- Next hop validationによる検出
  - BGP scan
- UPDATE

## • 情報の伝播

- Min Advertisement Interval
- Route Dampening

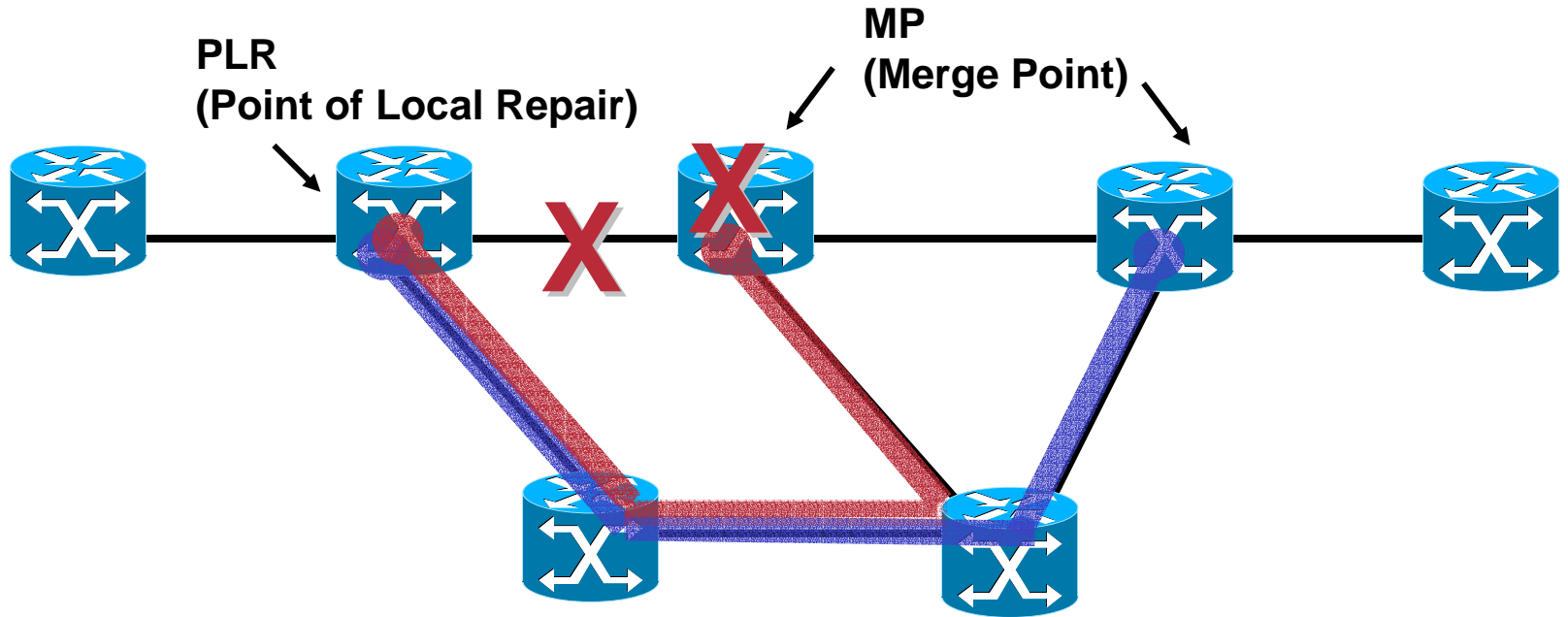
## • Best Path Selection

## • RIBのアップデート

BGPの場合、Scalability上の条件がIGPよりも格段に厳しいため、注意深く設計する必要がある。

課題: IGPとのcorrelation, Dynamic <-> Stableのバランス

# TE/FRR



< 50 msec程度の迂回所要時間

nhoP backup tunnel  
nnhoP backup tunnel

予めBackup Pathを計算しておくことにより、高速迂回を実現  
N:1の保護が可能  
Link ProtectionおよびNode Protection可能

# TE/FRR – Inter-ASの場合 ?!

## TE LSPの設定は、Intra-ASが前提

- IGPによりfloodingされたTopology 情報によるPath Computation
- ERO/RROの共有
- セキュリティ面が不十分
  - PATH message中のConstraintsは複雑化する
  - Policy Controlをどうするか

しかし、限られた条件であれば...

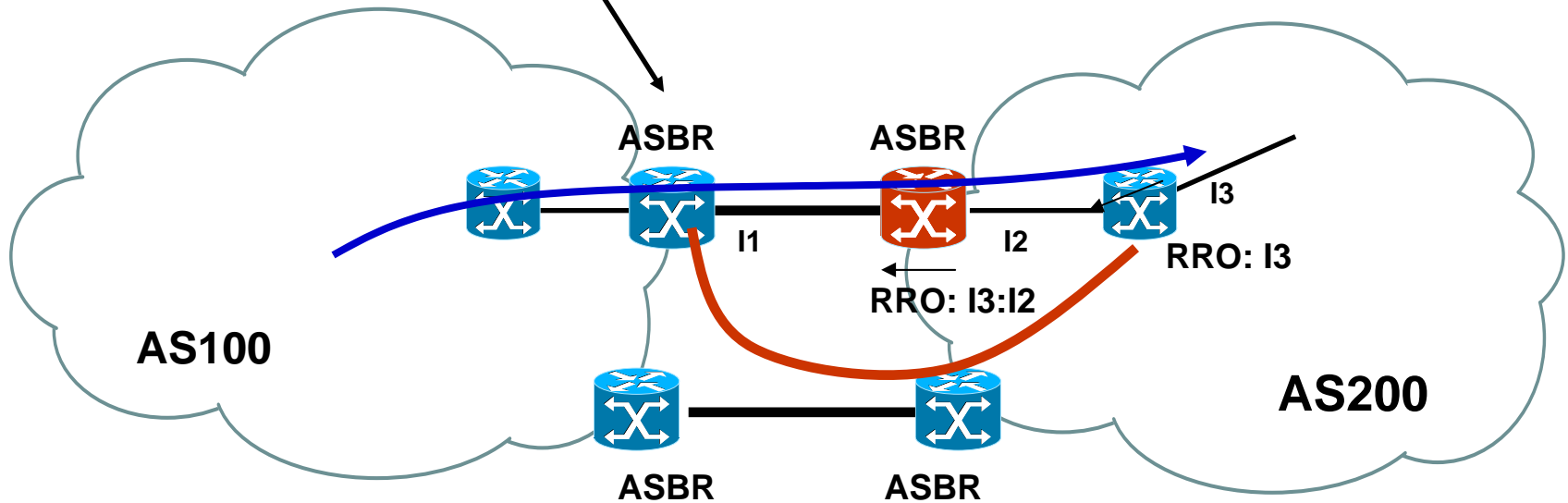
完全なるTrusted Peer同士のVirtual Trunk, etc.

[draft-ietf-tewg-interas-mpls-te-req-00.txt](#)

[draft-vasseur-inter-as-te-01.txt](#)

# Inter-AS TEでのFRR ?!!!

PLRは、RROに  
Stampingされた  
Prefix/LabelによりFRR  
DBを構成するが、他AS  
のinterface addressを  
認識していないので困る

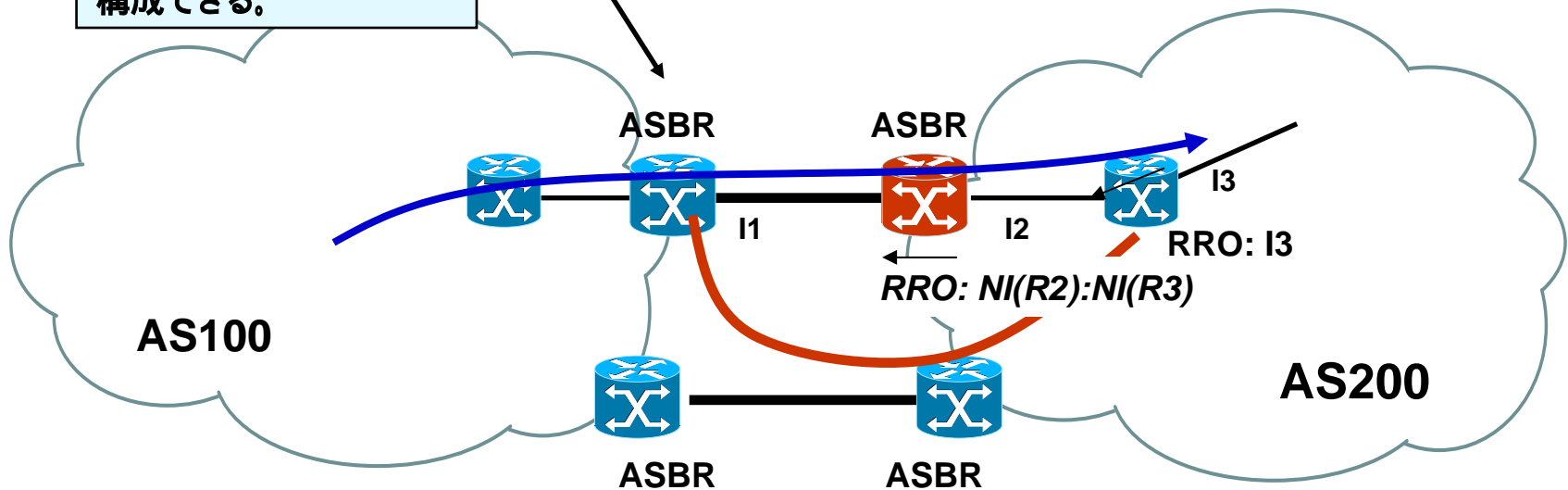


# Inter-AS TEでのFRR ?!!!

node-idというRRO subobjectを定義し、Router-IDをstampingする。

PLRはRouter-IDを識別可能なので、FRR DBを構成できる。

[draft-ietf-mpls-nodeid-subobject-01.txt](#)



# High Availability

Technology	Applicable for...				
	Intra-domain	Inter-domain			
		1-1	1-2	2	3
<b>SSO/NSF</b> Stateful Switchover + Protocol Graceful Restart					
<b>Routing Convergence</b>					
<b>TE FRR</b>		-		-	-

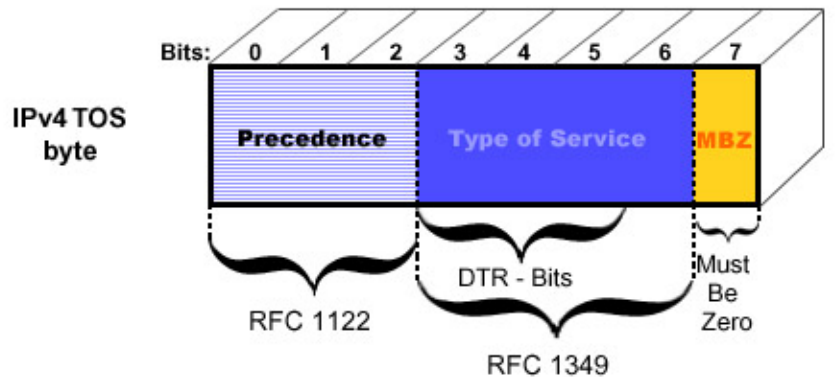
- **Diffserv PHB**
- **TE/DS-TE**



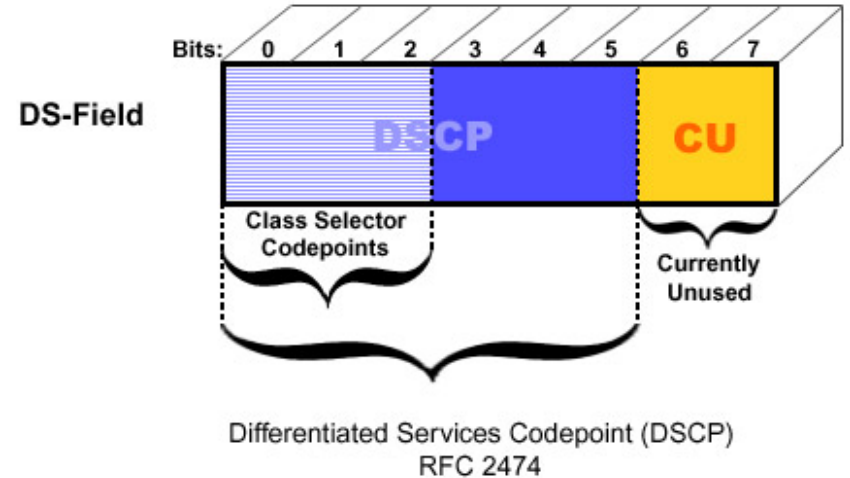
# QoS - Diffserv PHB(Per Hop Behavior)

• シンプルなはずのDiffserv PHBも、AS(正確にはQoS Policy Domain)をまたがった途端、使えない！

- Precedence ? ToS ? DSCP ? MPLS exp ?!
- 他ASからCS7 packetを受け取ったらどうする?!



Bits (0-2): IP- Precedence Defined	Bits (3-6): The Type of Service Defined
111 - Network Control	0000 (all normal)
110 - Internetwork Control	1000 (minimize delay)
101 - CRITIC/ECP	0100 (maximize throughput)
100 - Flash Override	0010 (maximize reliability)
011 - Flash	0001 (minimize monetary cost)
010 - Immediate	
001 - Priority	
000 - Routine	



# QoS - Diffserv PHB(Per Hop Behavior)

今さら(?), やっと(?), こんなドラフトが出ました。

[draft-baker-diffserv-basic-classes-00.txt](#)

The intent of this document is to provide a **consistent marking, conditioning** and **packet treatment strategy** so that it can be configured and put into service on any link which finds itself congested.

## Diffservのための、キーとなるコンセプト (Chapter 1.2)

1. Queueing
  - 1-1 Priority Queueing
  - 1-2 Rate Queueing
2. Active Queue Management
3. Traffic Conditioning
4. DSCP
5. PHB

# QoS - Diffserv PHB(Per Hop Behavior)

## DSCPとサービスクラスのマッピング

トラフィック種別	サービスクラス名	DSCP	DSCP値	アプリケーション名
Network Control	Critical	CS7	111000	Heartbeats
	Network	CS6	110000	Network Routing
Interactive	Premium	CS5, EF	101000,101010	IP Telephony
	Multimedia	CS4, AF41 AF42, AF43	100000,100010 100100,100110	Video Conferencing
Responsive	Streaming	CS3, AF31 AF32, AF33	111000,011010 011100,011110	Streaming Media
	Human-Responsive	AF21 AF22, AF23	000010 000100,000110	Client/Server transactions
Timely	Transaction	CS2, AF11 AF12, AF13	010000,001010 001100,001110	Store and forward applications
	Standard	DF, (CS0)	000000	未定義アプリケーション
Other	Scavenger	CS1	001000	no BW assurance

User Traffic

# QoS - Diffserv PHB(Per Hop Behavior)

## 各サービスクラスに使用されるQoSメカニズムのまとめ (1/2)

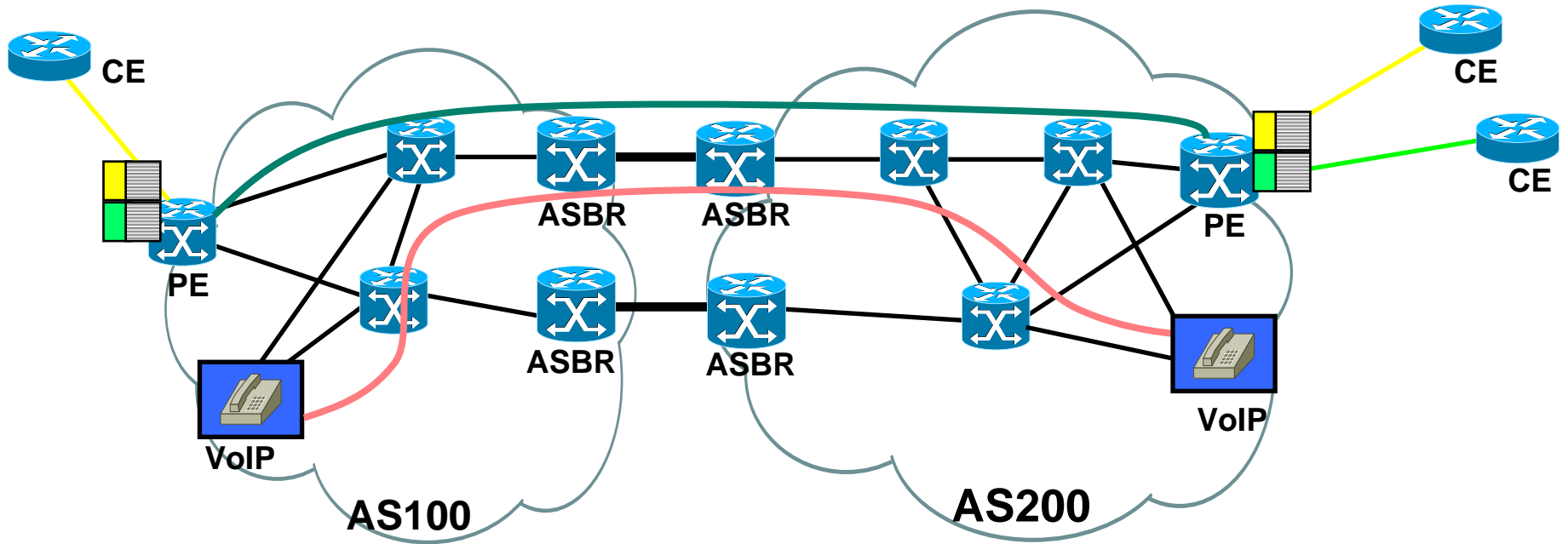
サービスクラス	DSCP	Diffserv Edge における制御	PHB reference	Queueing	AQM
Critical	CS7	Ingress: CS7パケットはdropする。 Egress Edge: dropするか、 またはCS6にmarkdownする。	RFC2474	Priority	No
Network	CS6	Trusted Sourceの場合、sr+bs ポリシングにより受け入れる。 Untrusted Sourceの場合、drop するか、Ingressでmarkdownする。	RFC2474	Rate	No
Premium	EF, CS5	Untrusted Sourceの場合、 sr+bsによるポリシング Trusted Source, CAC前提の場合 ポリシングしなくてもよい。	RFC3246	Priority	No
Multimedia	AF41 AF42 AF43	TrTCM (RFC2698)	RFC2597	Rate	Yes Per DSCP
	CS4	sr+bsによるポリシング			Yes

# QoS - Diffserv PHB(Per Hop Behavior)



## 各サービスクラスに使用されるQoSメカニズムのまとめ (2/2)

サービスクラス	DSCP	Diffserv Edge における制御	PHB reference	Queueing	AQM
Streaming	AF31 AF32 AF33	Untrusted Sourceの場合、 sr+bsによるポリシング Trusted Sourceの場合は ポリシングしなくてもよい。	RFC2597	Rate	Yes Per DSCP
	CS3				Yes
Human Responsive	AF21 AF22 AF23	srTCM (RFC2697)	RFC2597	Rate	Yes Per DSCP
Transaction	AF11 AF12 AF13	srTCM (RFC2697)	RFC2597	Rate	Yes Per DSCP
	CS2	sr+bsによるポリシング			Yes
Standard	DF	必要なし	RFC2474	Rate	Yes
Scavenger	CS1	必要なし	RFC2474	Rate	Yes

# Inter-AS TE/DS-TE



例)

-  Premium Virtual Trunk for VPN-yellow
-  Strict QoS guaranteed DS-TE

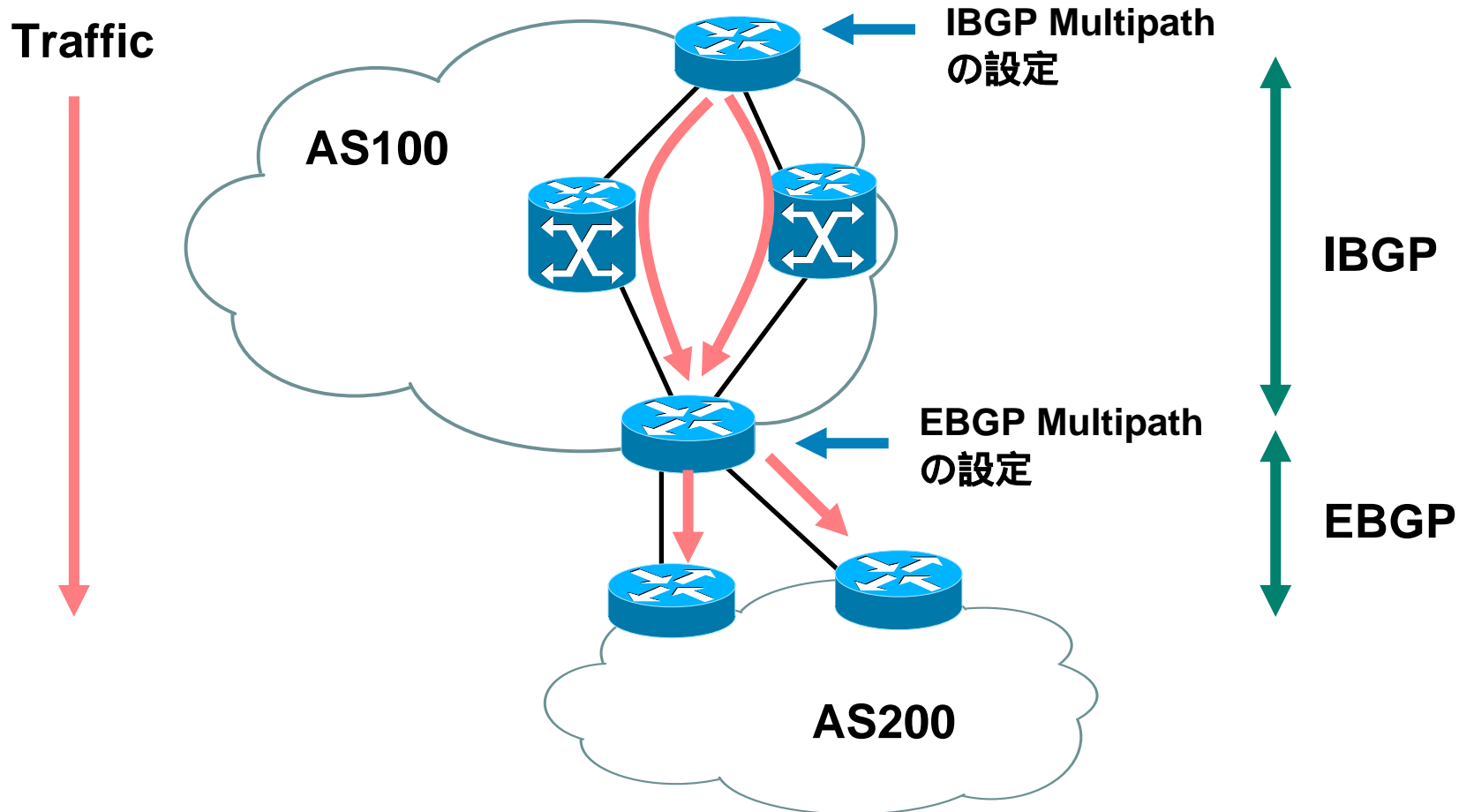
Technology	Applicable for...				
	Intra-domain	Inter-domain			
		1-1	1-2	2	3
Diffserv PHB					
TE/DS-TE		-		-	-

# Load BalancingとOptimization

- **Metric/Cost調整、ECMP、TE**
- **OER (Optimized Edge Routing)**



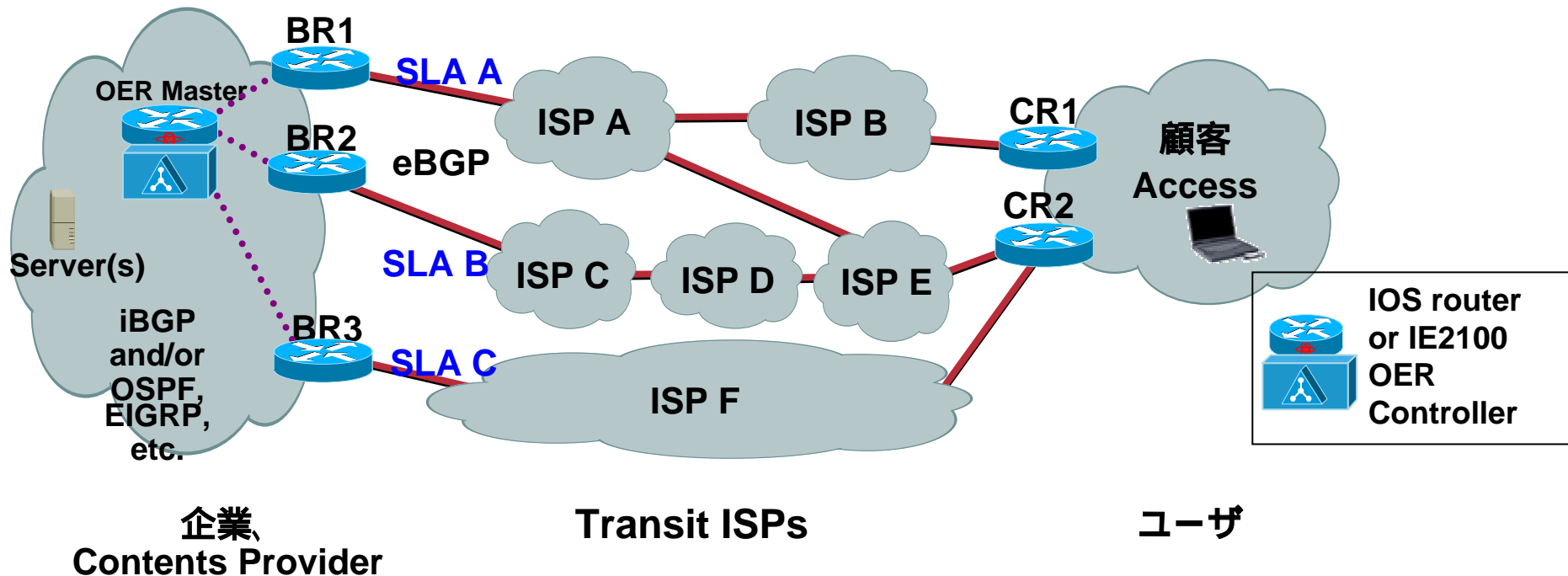
# Metric/Cost調整、ECMP、TE



# Optimized Edge Routing (OER)

Cisco.com

- multihomeされていても、コストやパフォーマンスの最適化を行うのは困難。
- OERにより、ルート最適化およびload balancingを、効率的・自動的に行う。



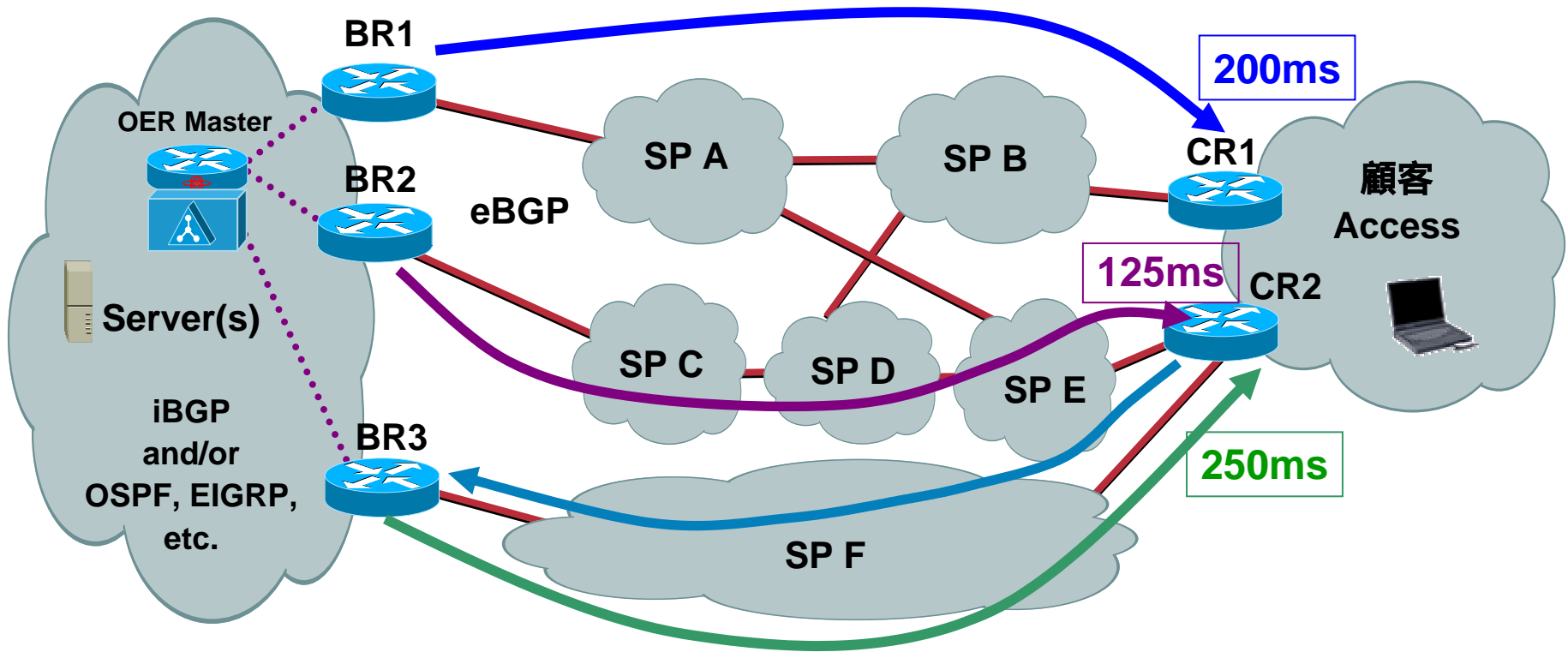
# OERにより実現すること

Prefix毎に、あらかじめ規定されたポリシーに基づき、最適なパス(ISP)を選択する。



- **Brown-outの回避**  
Provider Blackholeを回避する。
- **パフォーマンス**  
TCP SYN-ACK遅延が最短となる経路を選択することにより、  
負荷の高いパスを回避、パフォーマンスを向上する。
- **帯域、回線使用率の最適化**
- **Prefix毎のロードシェア、パス選択**

# OER - Monitoring

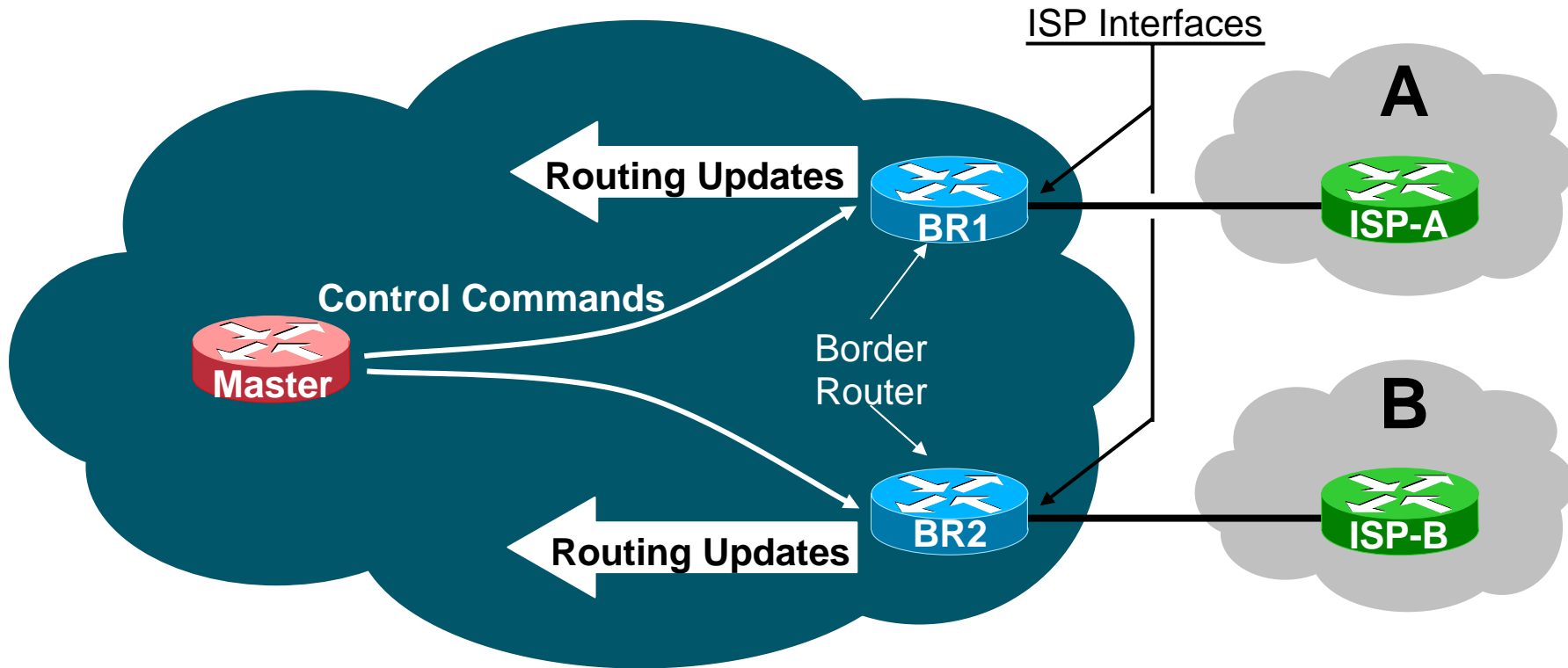


- Monitorするprefix --- コンフィグレーション、または自動学習(Top Talker)
- Monitorの対象 --- TCP SYN-ACK delay, Throughput, Packet Loss, Reachability
- Monitorの手段 --- Netflow (Passive), SAA (Active)

# OER - Control

Cisco.com

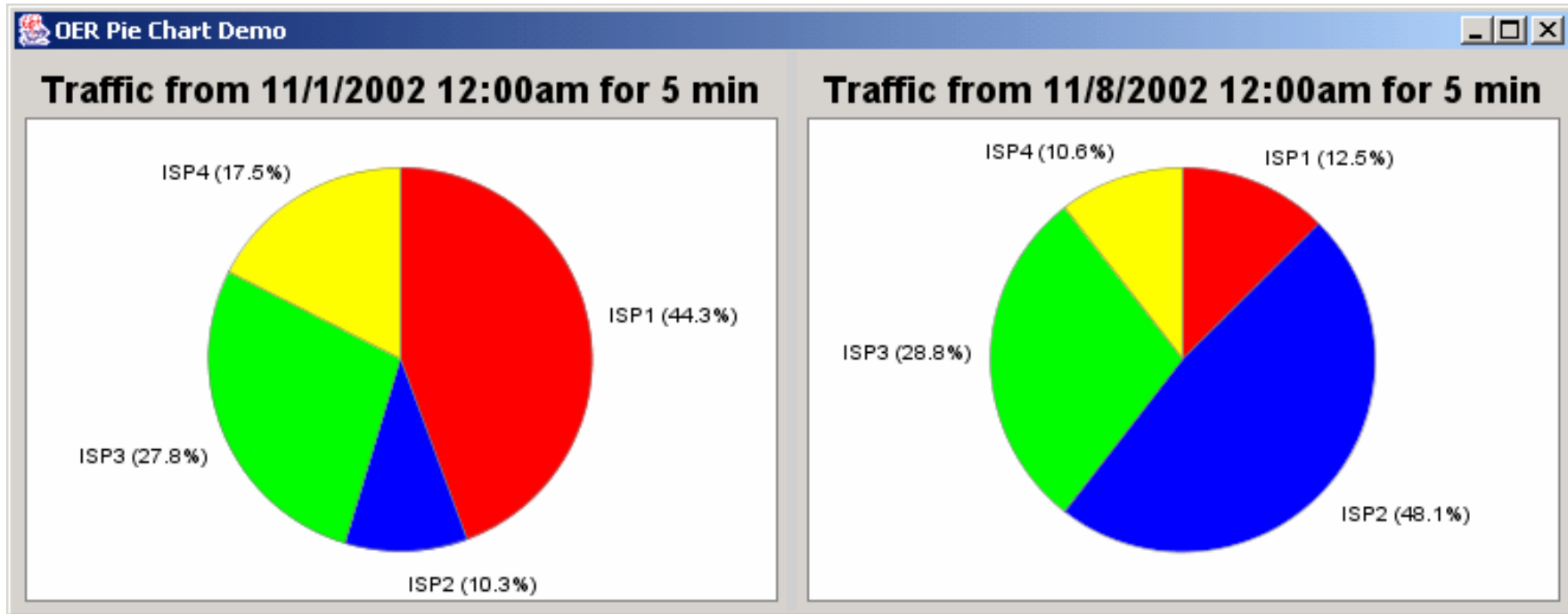
## ISPs



- EdgeにおけるMonitoringの結果により、次のような制御が可能。  
BGP(Local Preference、または他のAttribute)を変える。  
static routeを設定する (with redistribute static)。  
IGP metricを変更する。

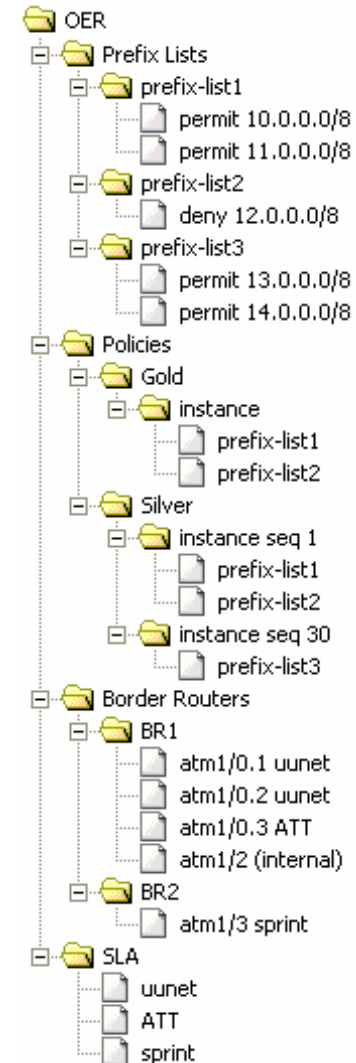
# OER - Reporting

- トラフィック分布



# OER - Reporting

- OERの状態
- SLAポリシー
- Routing Tableの変化とその要因



# Load BalancingとOptimization

Technology	Applicable for...					
	Intra-domain	Inter-domain				
		1-1	1-2	2	3	4
Metric/Cost調整、ECMP、TE		(*)	-	(*)	-	(*)
OER (Optimized Edge Routing)		-	-	-		

(\*) with BGP Multipath



- **今後は、Inter-AS環境においても、高品質が求められる。**
- **高品質Inter-AS接続を実現するために、Bilateral Peeringの重要性が増す。**