

MPLS Traffic Engineering

日本テレコム(株)

松嶋 聡

<satoru@japan-telecom.co.jp>

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

Traffic Engineeringとは何か？

- 広義には...
 - ネットワーク上に流れるトラフィックを効率的に過不足なく処理するために行うネットワーク設計、作業全般。
 - 必要回線容量の見積もり。
 - 回線使用率、トラフィック量の監視。
 - 回線容量の増強、増速作業など。
 - つまりトラフィックの流れから見て、ネットワーク構成を最適化すること。
- 以後、TEと呼びます。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

IPネットワークにおけるTE

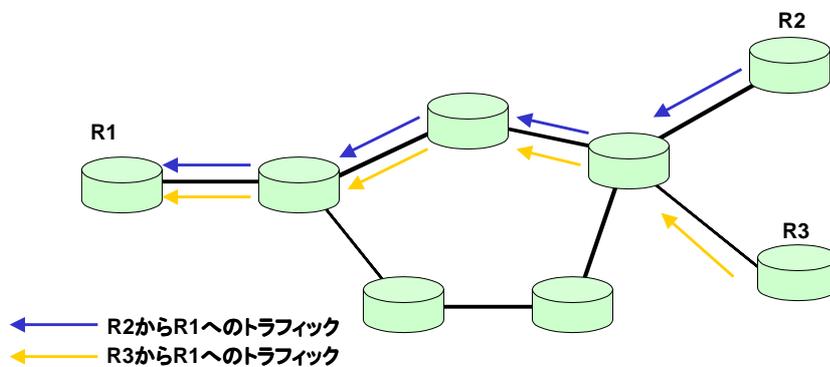
- みんなどんなTEをやっているだろう？
 - 例えば...
 - MRTGでトラフィック量監視。
 - 特定回線のピークトラフィックが50%を超えたあたりから増設計画を立てて、予算確保。
 - 70%~80%あたりで警戒域。急がねば。。。。
 - 100%になると、おそらくクレームが来る。
 - 回線数を増やす。
 - ロードバランスさせる。
 - トラフィックを別回線に移す。
 - IGP(RIPやOSPF)経路のメトリックを変える。
 - などなどが、現在までのソリューション。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

例題：IPネットワークにおけるTE

- ルータR1へのトラフィックの流れ
 - 最もコストの小さい経路を選択。

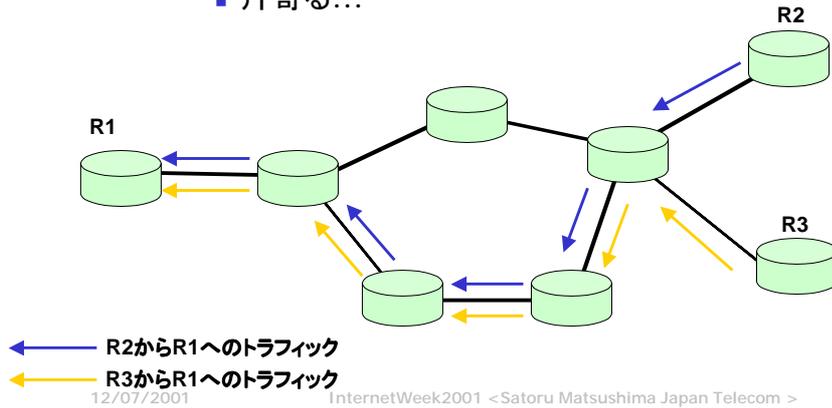


12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

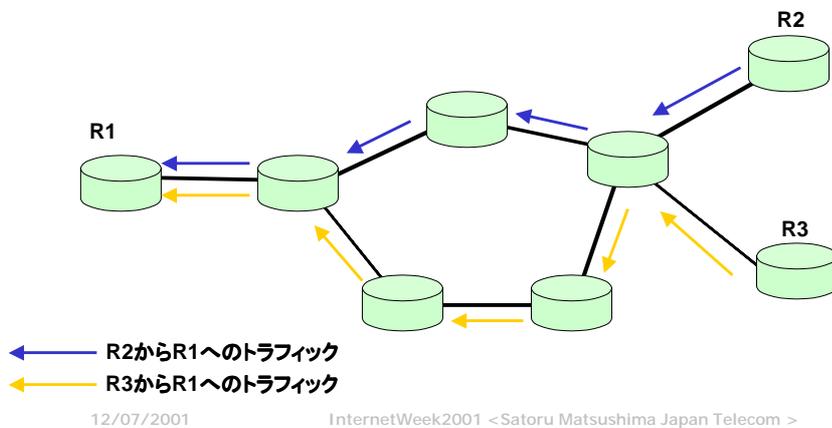
例題: IPネットワークにおけるTE

- リンクのコスト値を変えてみる。
 - 片寄る...



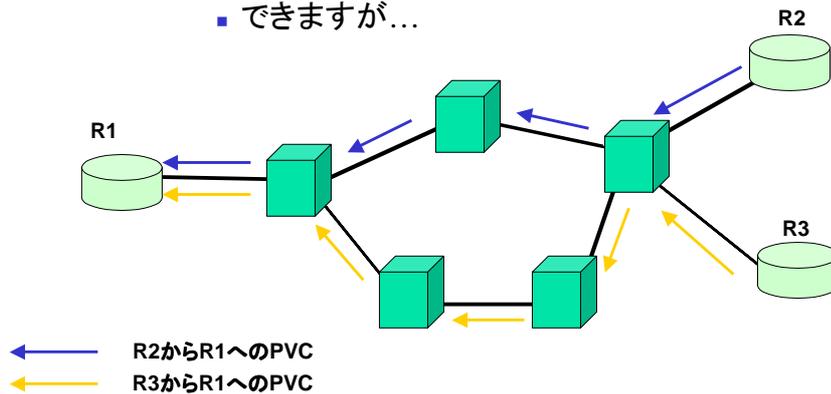
例題: IPネットワークにおけるTE

- 本当はこうしたかったのに！



例題: IPネットワークにおけるTE

- ならばATMやFrame-Relayで...
 - できますが...



12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

ATM/FRをIPTラフィック制御に 用いた時の問題点

- IPネットワークとは別にATM/FRネットワークを用意することになる。
 - ルータ以外のネットワークデバイス。
 - それぞれにおいて運用/監視が必要。
→高コスト
- IGPへの負荷が高くなる。
 - メッシュによるAdjacency増。
 - 複雑な論理トポロジー。
- 帯域を圧迫するオーバーヘッド
 - ATM Cell Tax.

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

IPネットワークにおけるトラフィック制御の問題点

- リンクのコストやメトリックによるIGPルーティングはトラフィックの流れから見ると、最適ではない。
 - 利用率が非常に高い回線と低い回線が存在。
 - ネットワーク資源を有効活用できず、しかも高利用率回線の増強を常に余儀なくされる。
→高コスト!
 - 利用率の高い回線が障害になったときのリスクは非常に大きい。
 - 通常利用率の高い回線は高速/超高速回線。
 - そのような回線が切れたときのバックアップ容量が十分であることを保証するのは困難。
→信頼性、可用性に欠けたネットワーク。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

IPネットワークにおけるトラフィック制御の問題点

- トラフィックの最適化が非常に困難。
 - IGPにトラフィックを最適化する経路変更は行えない。
 - トラフィックフローに対して直接オペレータが介入する手段はほぼ皆無。

明示的にトラフィックの流れを最適化できる
新たなTE手法が必要



MPLSによるTraffic Engineering !

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >



MPLSによるTraffic Engineering

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >



MPLS TEのメカニズム

- Explicit (明示的) Routing
 - Constraint-based (強制的な)パス選択
 - IGPコストによらず、空いているパスを検索し、選択する。
 - IGP(IS-ISやOSPF)を拡張して、ネットワーク資源(リンクの帯域予約情報)やポリシーを運ぶ。
 - 空いているパスを選択するために使われる。
 - MPLSによるパケットフォワーディング
 - IPアドレスによらないフォワーディングにより、Explicit Routingを実現。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >



Explicit Routing

- RSVP(RFC2205)-extensionによりLabel Switched Path(LSP)を確立。
 - ネットワーク資源情報などにより、明示的なパスを設定する。
 - その他にも
 - LSP tunnelの障害検知
 - LSP tunnelのre-routing管理
 - draft-ietf-mpls-rsvp-lsp-tunnel-09.txt

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >



IGP(IS-IS or OSPF)の拡張

- MPLS TEに対応したルータは以下の情報を拡張されたIS-IS/OSPFにより伝達する。
 - 帯域情報(bandwidth)
 - どれだけの帯域が予約可能か。
 - リンク属性(Link Attribute)
 - どのようなポリシーのパスで使用されるのか。
 - TEメトリック(TE-specific link metric)
 - リンクの強さ。
- いつ伝達されるのか？
 - 周期的、または予約可能帯域が変化したときなど。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

IGP(IS-IS or OSPF)の拡張

- OSPFへの拡張方法
 - Opaque LSA (Type10: Intra-area)
(draft-katz-yeung-ospf-traffic-06.txt)
- IS-ISへの拡張方法
 - New wide TLV (draft-ietf-isis-traffic-04.txt)
- なぜIS-ISやOSPFだけなのか？
 - RIPの立場は？
 - それはリンクステート型だから。
 - でないと、リンクに関わる情報を運べない。
 - IS-IS/OSPFでも同一エリア内のみ有効。
 - エリア内にしか全てのリンク情報はない。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

MPLSによるフォーワーディング

- なぜMPLSでなければならないのか？
 - TEは明示的にトラフィックの流れるパスを指定できないなければならない。
 - 宛先IPアドレスによる従来のフォーワーディングではホップバイホップフォーワーディングになってしまう。
→途中のルータが意図したパスにぞってパケットを転送するかどうか分からない。
 - MPLSならば、IPアドレスによらないフォーワーディングを実現することができる。
 - 32bit固定長のラベルに基づくフォーワーディング。
 - ラベルにパスを示す値をつけることで、Explicit Routingに従ったフォーワーディングが実現できる。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

MPLS TEのセットアップ

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

MPLS TEのセットアップ

- RSVP拡張を用いたLSP セットアップ
 - LSPをセットアップするルータ(head-end)から起動。
 - LSP宛先ルータ(tail-end)からラベルを割当てる。
 - Downstream On Demand Label Distribution (DoD)
 - LSPに対して帯域を予約する。
 - シグナリング上でのみ。実際に運ばれるトラフィックの帯域を保証しているわけではない。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

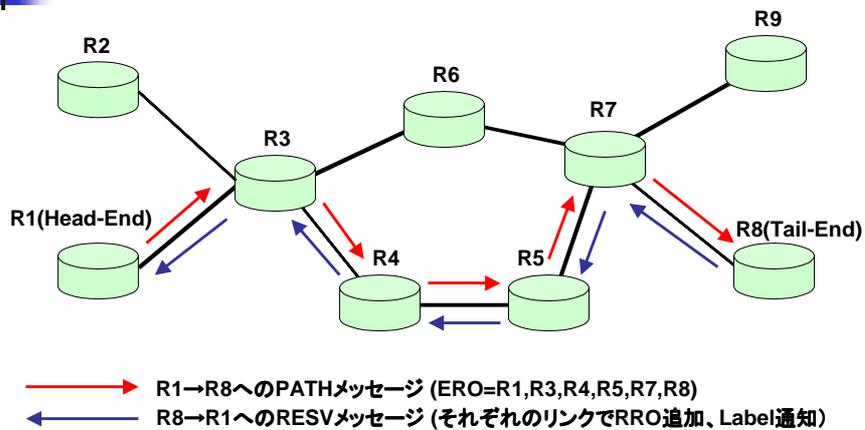
MPLS TEのための新しいRSVP オブジェクト

- Explicit Route Object (ERO)
 - TEパスを設定するリンクを明示的に連ねる。(PATH)
- Record Route Object (RRO)
 - RSVPメッセージが通過したリンクを記録。(PATH,RESV)
- Label Request Object
 - ラベル割り当て要求(PATH)
- Label Object
 - ラベル割り当て通知(RESV)
- Session Attribute Object
 - パスのプライオリティ、名前や予約スタイルを通知 (PATH)

12/07/2001

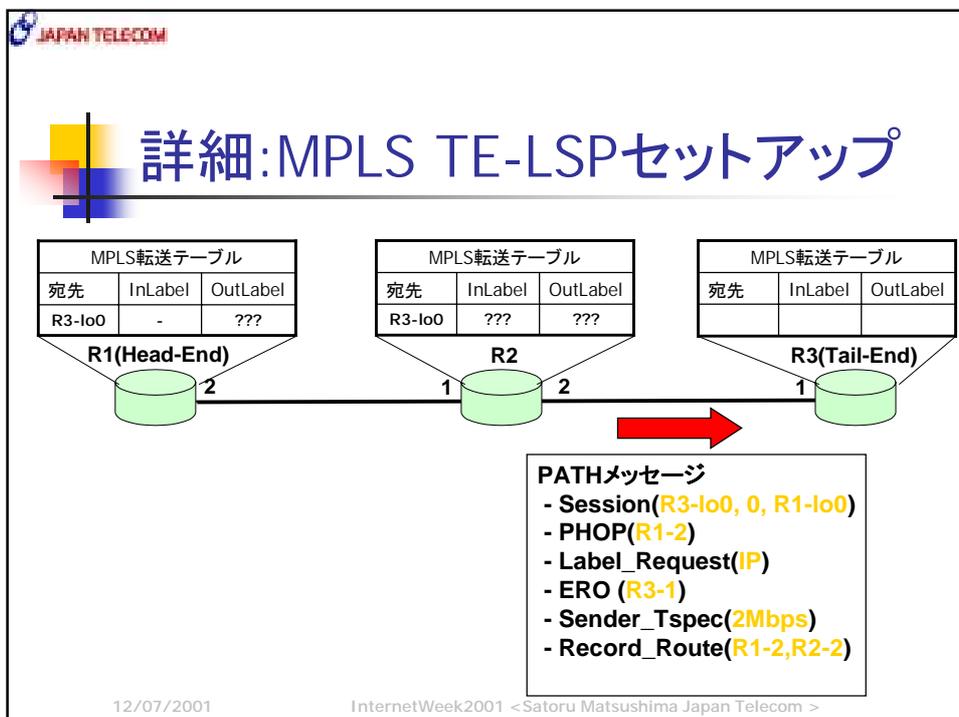
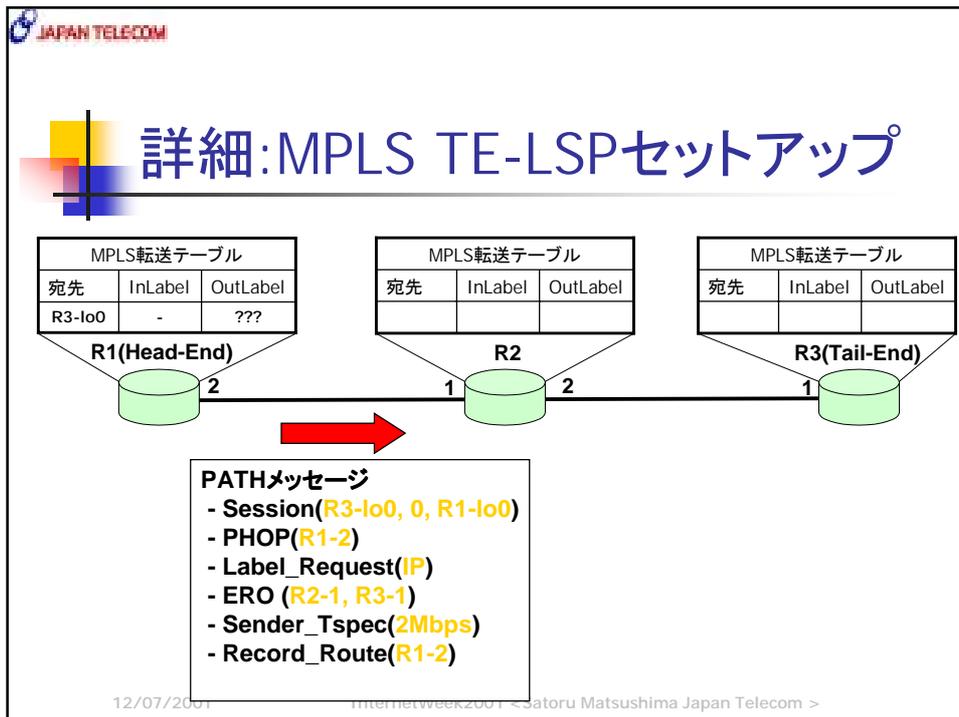
InternetWeek2001 <Satoru Matsushima Japan Telecom >

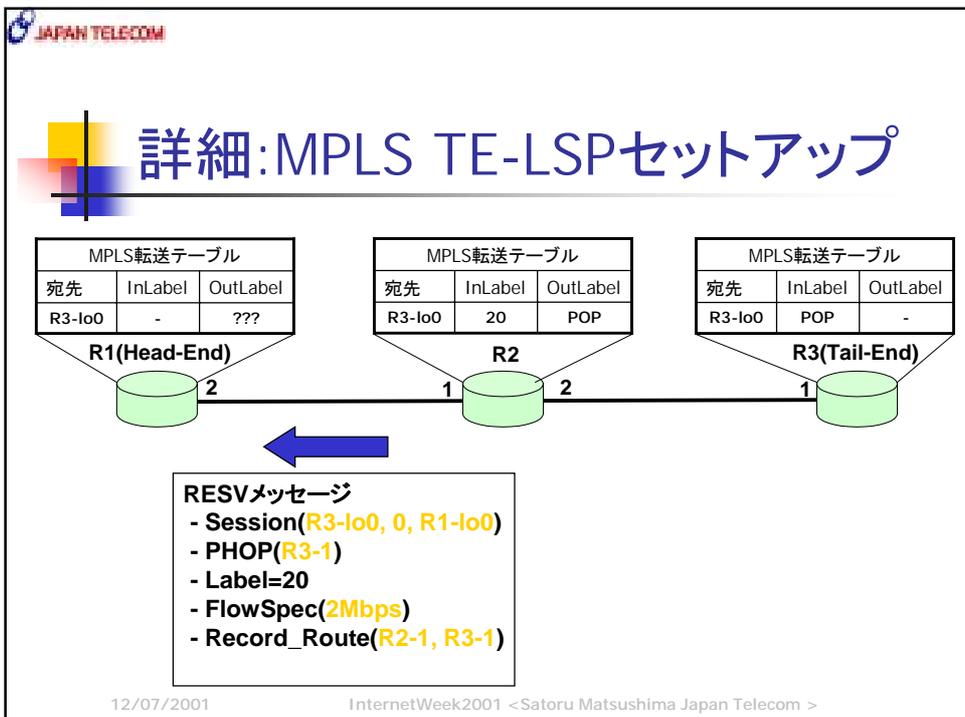
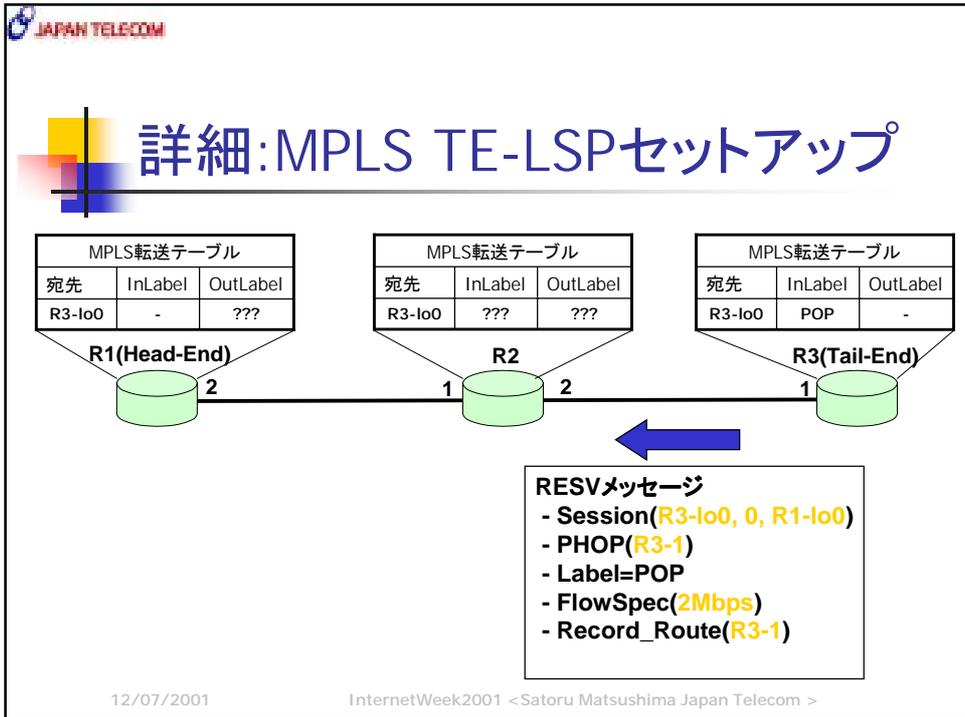
概要:MPLS TE-LSPセットアップ



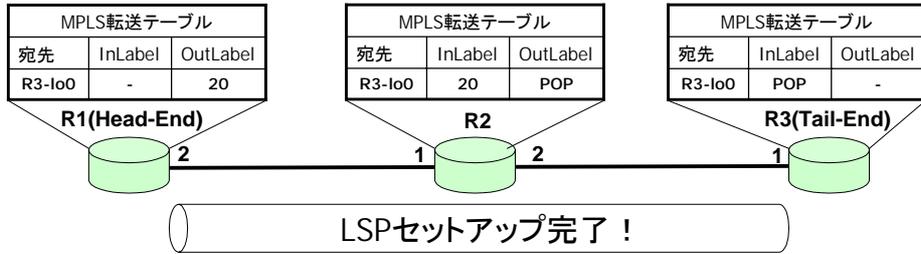
12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >





詳細:MPLS TE-LSPセットアップ



MPLS TEのメリット

MPLS TEのメリットとは？

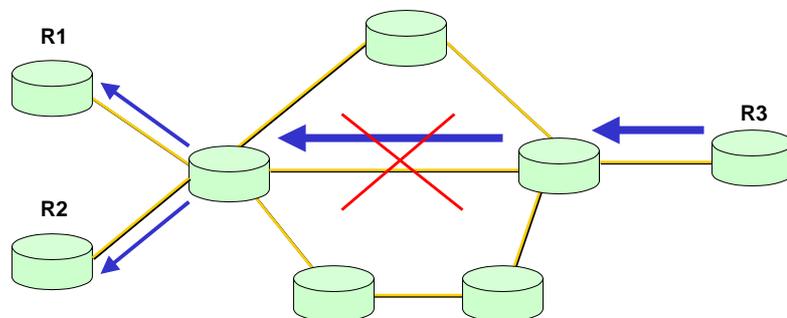
- IPネットワーク上でトラフィックを最適化することは今まで述べたとおり。
それ以外には？
 - リンク障害時の負荷分散
 - 高速なバックアップパス確立(Fast Reroute)
 - 特定POP間のトラフィック管理
 - 負荷分散の調整
 - などなど...

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

リンク障害時の負荷分散

- 平常時は余裕で運べたトラフィックも...

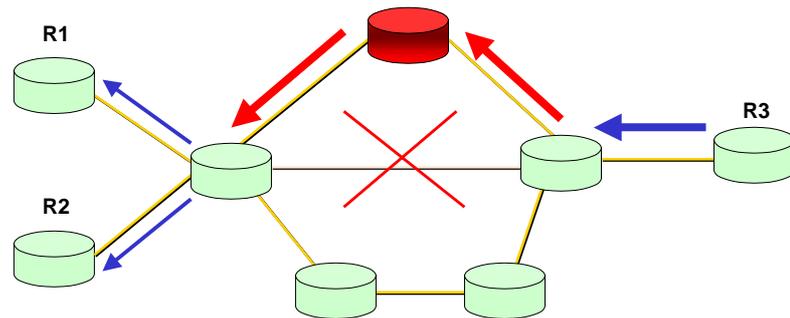


12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

リンク障害時の負荷分散

- コストの都合上、細いリンクに片寄られると苦しい...でも

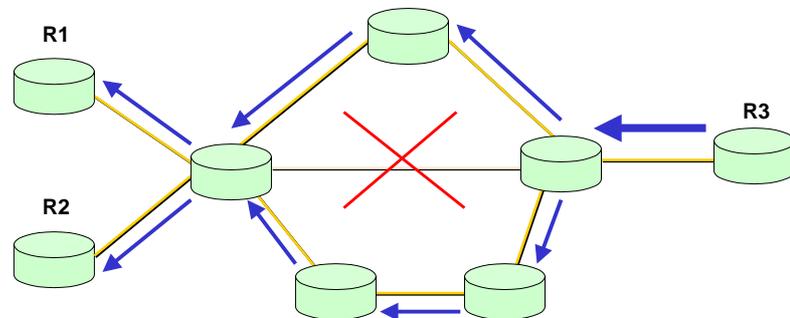


12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

リンク障害時の負荷分散

- MPLS TEで安心。最適化経路変更。
 - Re-OptimizeでIGPだけよりも高速。(<数Sec)

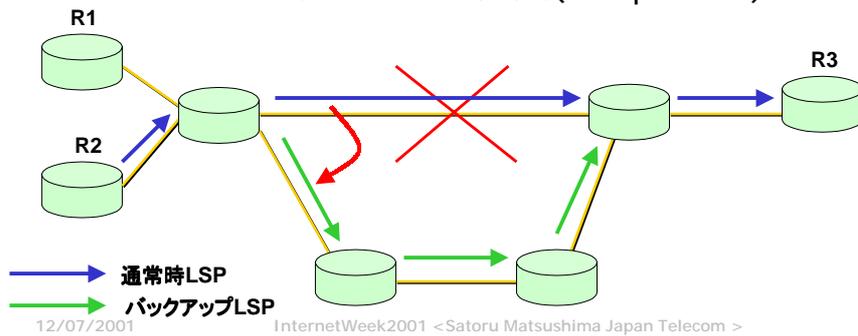


12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

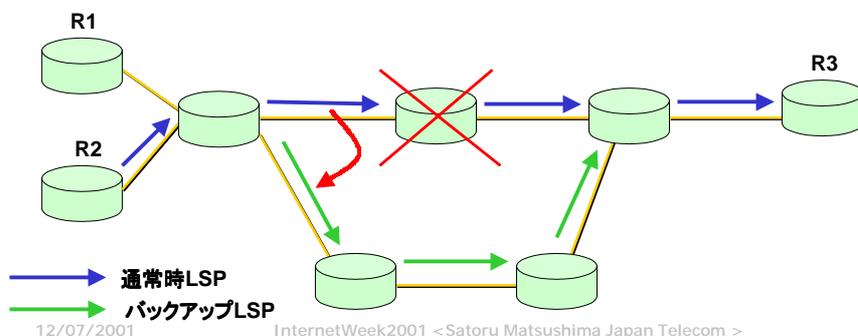
高速なバックアップパス確立 (Fast Reroute)

- リンクプロテクション
 - 特定リンク切断時に高速切替(≒50ms)
 - その後に最適化経路変更(re-optimize)



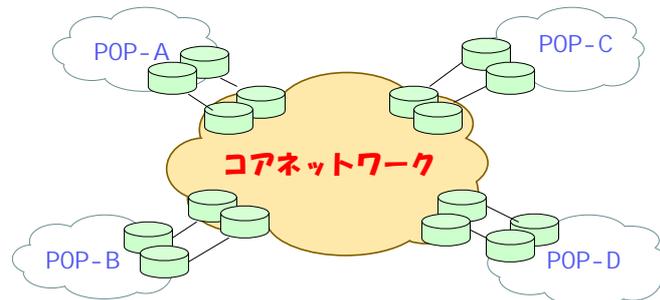
高速なバックアップパス確立 (Fast Reroute)

- ノードプロテクション
 - 特定ノードがダウンしたときに高速切替



特定POP間のトラフィック管理

- POP間のトラフィックを知りたい！
 - 下のような構成で特定POP間のトラフィック管理は困難。しかし...

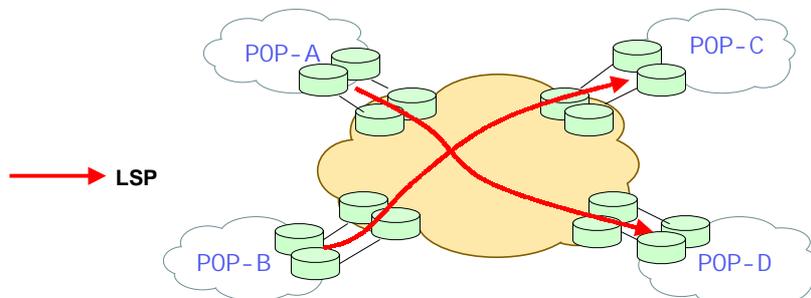


12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

特定POP間のトラフィック管理

- MPLS TEで可能になる！
 - あるPOPから他POPへのトラフィックランクに対応するLSPのパケットをカウントできるようになる。



12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >



負荷分散の調整

- RIPやOSPFなどのIGPでは、コストやメトリックの足し算の結果の大小で、トラフィックが流れるor流れないという極端な差が現れる。
- MPLS TEでは、リンクの予約可能帯域量を調整することによりコストに関わらず、それなりのトラフィックをそのリンクにのせることができる。



MPLS-TEの課題

MPLS-TEの課題は？

- Traffic Engineeringの有効性は十分にあるが、気になるところ。
 - MPLS TEを使うことでIGPルーティングとトラフィックの流れの関係が疎になる。
 - LSPを設定したルータでないと見えなくなる。
 - Constraint-Pathの元となるパス計算が正しいということはどうやって確認すれば良いのか？
 - LSPの正当さを評価するツールが欲しい。
 - 運用経験
 - どうやって設計すればよいの？

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

開発/標準化中の新機能

- Fast Reroute
 - 現在、2つの方法が実装されている。
 - MPLS TEの基本機能をベースに、re-optimizeまでの高速バックアップとして用いる。
 - draft-swallow-rsvp-bypass-label-01.txt (Cisco)
 - 新たなRSVPオブジェクトを定義し、自動的にバックアップ (Detour)パス設定。
 - FAST_REROUTE, DETOUR Object.
 - draft-gan-fast-reroute-00.txt (Juniper)
 - 両者はこれから共同で仕様を出してくる予定。
 - おそらく、次のIETF@Salt Lake

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

開発/標準化中の新機能

- Diff-Serve関連
 - LSPをDiff-Serve PHB(PSC)にマップさせる。
 - L-LSP: Label-inferred LSP
 - 1つのLSPが1つのPHB(PSC)にマップされる。
 - E-LSP: EXP-Inferred LSP
 - MPLSラベルEXPビットを使って1つのLSPが複数のPHB(PSC)にマップされる。
 - draft-ietf-mpls-diff-ext-09.txt
 - Diff-Serve aware TE
 - TEと連動したDiff-Serve。クラス毎に帯域を予約。
 - トラフィックフローが集約されたTEの中にある、特定フローに対して帯域を予約可能。

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >

まとめ

- TEとは、トラフィックの流れから見てネットワーク構成を最適化すること。
- MPLSによるパケットフォワーディング。
- Constraint-based Routing
- MPLS TEのためのIGP(IS-IS/OSPF)拡張
- MPLS TE設定のためにRSVPが用いられる。
- さまざまな機能が標準化/開発中
 - Fast Reroute
 - Diff-Serve TE
 - などなど

12/07/2001

InternetWeek2001 <Satoru Matsushima Japan Telecom >