

# **MPLSにおける高速障害回復技術動向 - Fast Reroute 方式 -**

**岩田 淳**

**NEC ネットワーキング研究所**

**E-mail: [a-iwata@ah.jp.nec.com](mailto:a-iwata@ah.jp.nec.com)**

**TEL: 044-856-2123, FAX: 044-856-2230**

# 概要

---

- **(1)背景**
- **(2) 障害回復方式概要**
- **(3) Fast Reroute方式**
  - Cisco方式
  - Juniper方式
  - Axiowave方式
- **(4) Shared Fast Reroute方式**
  - NEC方式
  - 性能評価例
- **(5) その他課題**
  - Interoperabilityの標準化
  - OAMの標準化
- **(6)まとめ**

# 障害回復方式概要 #1

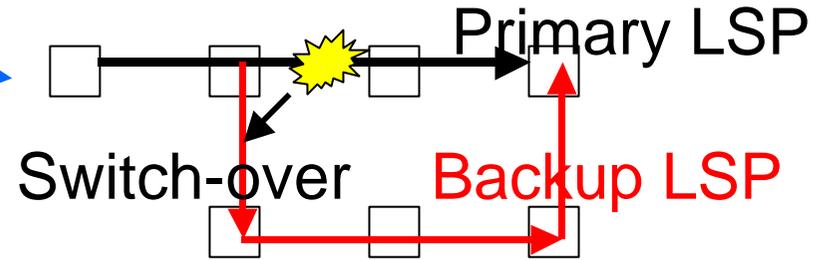
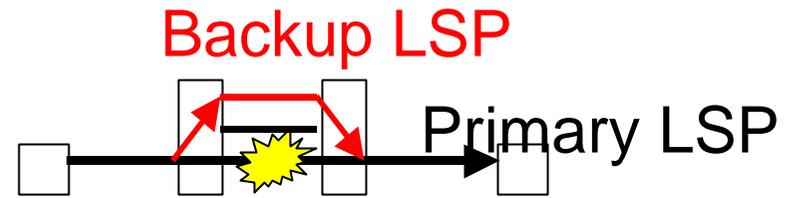
- **MPLS LSPの障害回復方式のフレームワーク**
  - V.Sharma (Jasmine Networks), “Framework for MPLS-based recovery,” draft-ietf-mpls-recovery-frmwrk-03.txt, Jul. 2001
- **障害回復方式：2方式に大きく分類**
  - (1) Rerouting方式：
    - (a) established-on-demand(障害発生検出後に迂回経路探索設定)
    - (b) pre-qualified (障害前に迂回経路資源の確認のみ実行)
  - (2) Protection switching方式：
    - pre-established (障害前に迂回経路設定)
- **Protection switching方式の方式の分類**
  - (1) 1+1: Working / Recovery pathの両者に同一トラフィックを流し、Egress側で障害を検出
  - (2) 1:1 : Working pathの障害時にRecovery pathへ変更、それまではRecovery pathは低優先のTraffic (低Preemption priority)で利用されている。

# 障害回復方式概要 #2

- 障害回復の位置による3手法

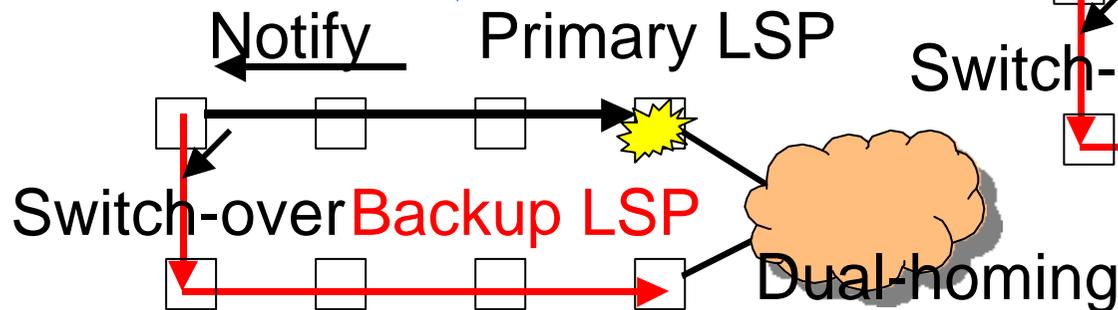
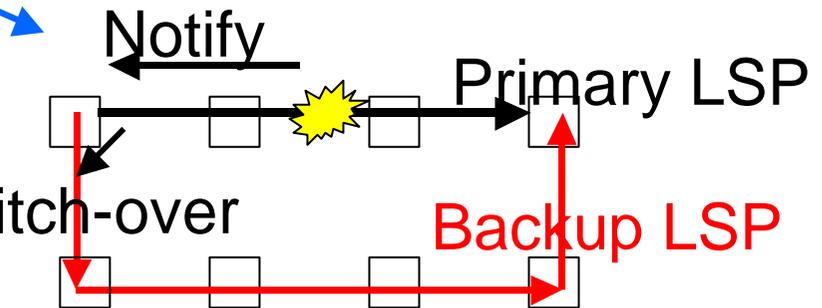
- Local Repair (fast reroute)

- Link Recovery / Restoration
    - Node Recovery / Restoration



- Global Repair

- Alternate Egress Repair



# 障害回復方式概要 #3

	Local Repair (Link)	Local Repair (Node)	Global Repair	Alternate Egress Repair
Rerouting	回復時間： 予備帯域なし	回復時間： 予備帯域なし	回復時間：× 予備帯域なし	回復時間：* 予備帯域なし
Protection (1+1)	回復時間： 予備帯域：2倍	回復時間： 予備帯域：2倍 以上	回復時間： 予備帯域：2倍 以上	回復時間：* 予備帯域：2倍 以上
Protection (1:1)	回復時間： 予備帯域：低 Pri user収容	回復時間： 予備帯域：低 Pri user収容	回復時間： 予備帯域：低 Pri user収容	回復時間：* 予備帯域：低 Pri user収容



## Fast - Reroute(FRR)の狙う領域

\*. Local repair/Global repairのいずれの方式と組み合わせるかで回復時間が決定  
Nov. 2001, MPLS Japan All Rights Reserved, Copyright(C) 5

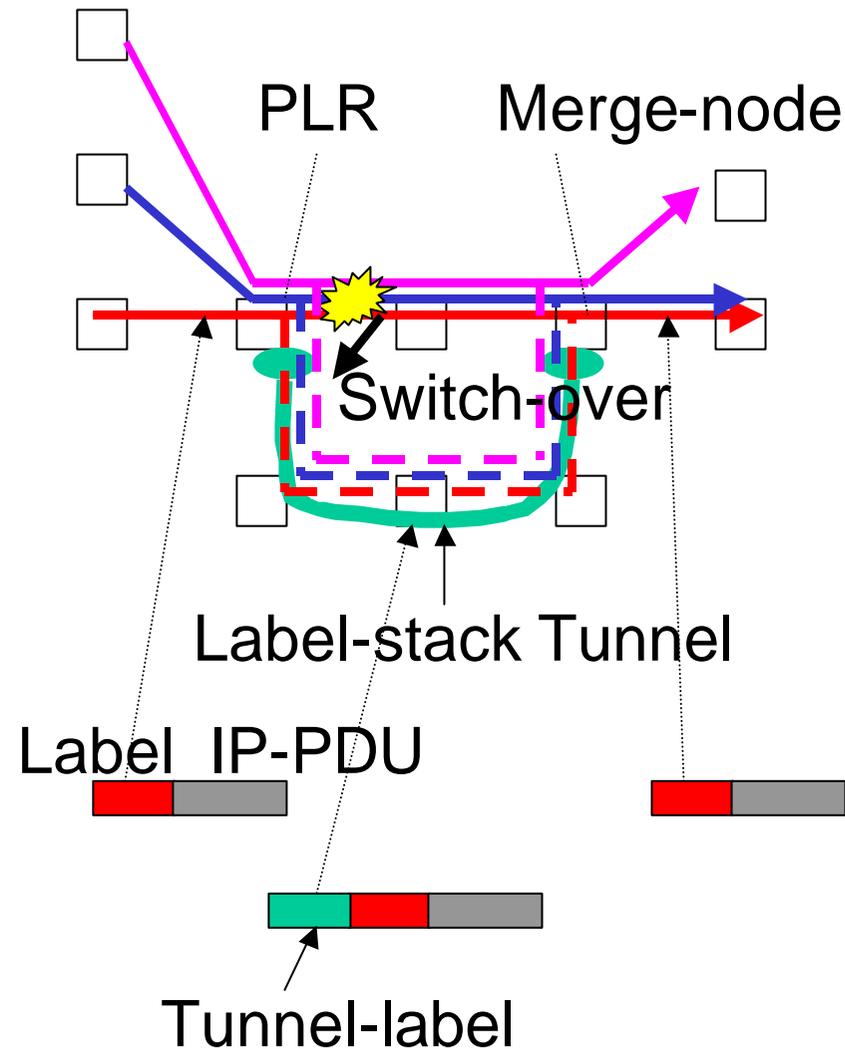
# Local Repair (Fast Reroute)

---

- **Cisco方式**
  - R.Goguen et al., “RSVP Label Allocation for Backup Tunnels,” draft-swallow-rsvp-bypass-label-01.txt, Nov.2000
- **Juniper方式**
  - D. Gan et al., “A Method for MPLS LSP Fast-Reroute Using RSVP Detours,” draft-gan-fast-reroute-00.txt, Apr.2001
- **Axiowave方式**
  - D. Haskin et al., “A Method for Setting an Alternative Label Switched Paths to Handle Fast Reroute,” draft-haskin-mpls-fast-reroute-05.txt, Nov.2000
- **NEC方式 (Shared Fast Reroute)**
  - A. Iwata et al., “MPLS Signaling Extensions for Shared Fast Rerouting,” draft-iwata-shared-fastreroute-00.txt, Jul.2001

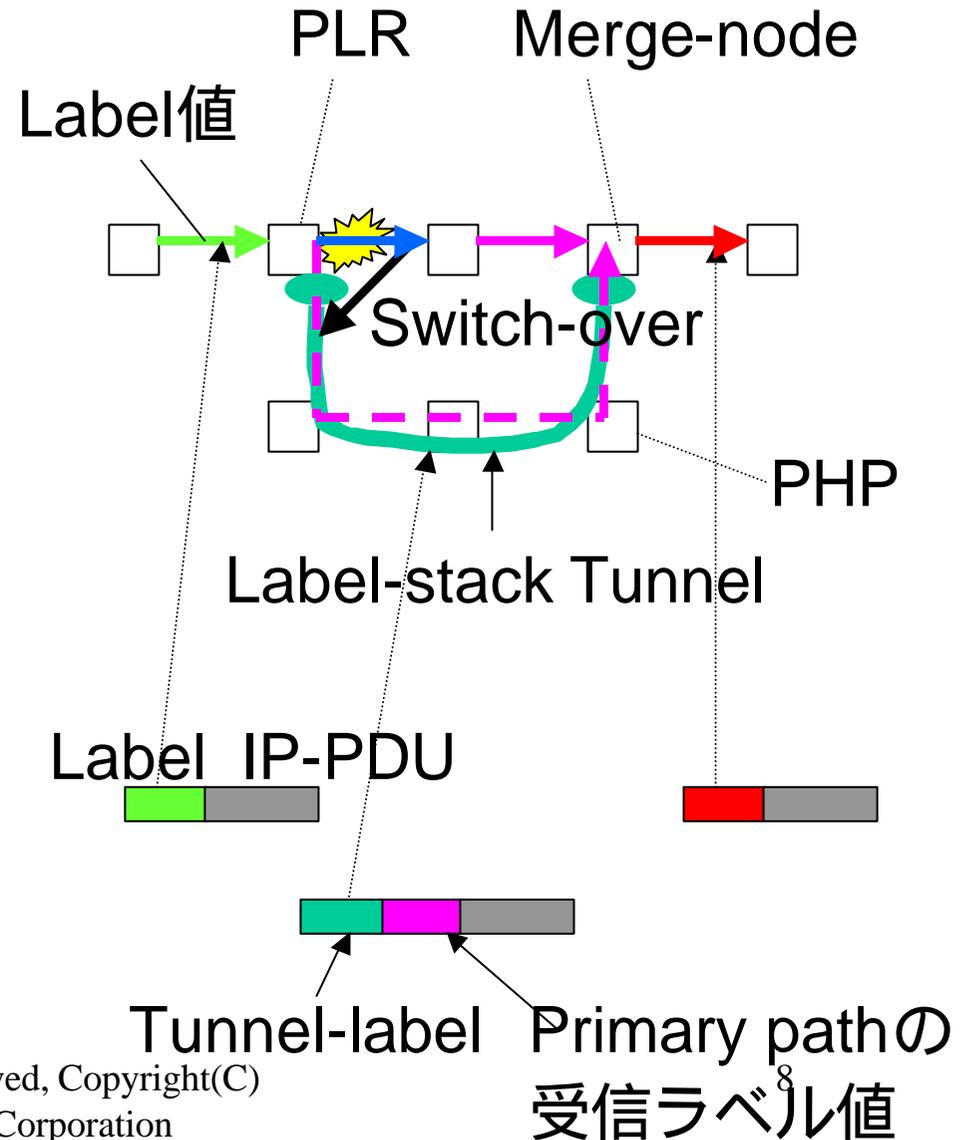
# Cisco FRR: “Bypass Tunnel” #1

- **Label-stack Tunnel**: リンク・ノード障害に対し、同一リンク、ノード経由の複数のLSPを多重して復旧
- PLR (Point of Local Repair) と Merge node間の Tunnel生成



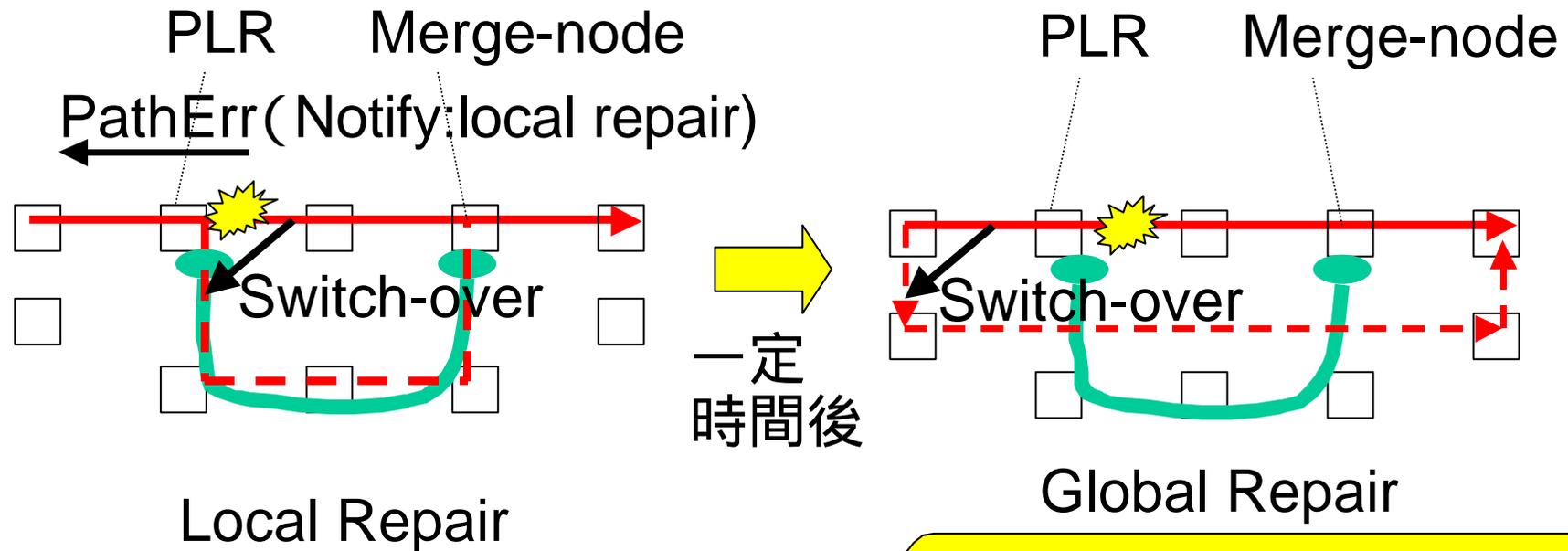
# Cisco FRR: “Bypass Tunnel” #2

- Working path (Primary path)のLabel値はHop-by-hopに変更
- Merge nodeでのGlobal label space の利用例 (異なるIFからの受信ラベル値に対して出力ラベル値を一意に決定、Mergeの実現の容易性)
- PLRにおいてMerge nodeでのPrimary Pathでの受信ラベル値に変更後Bypass Tunnelヘッダ付与
- Merge nodeの1ノード前でPenultimate Hop Pop (PHP)を動作



# Cisco FRR: “Bypass Tunnel” #3

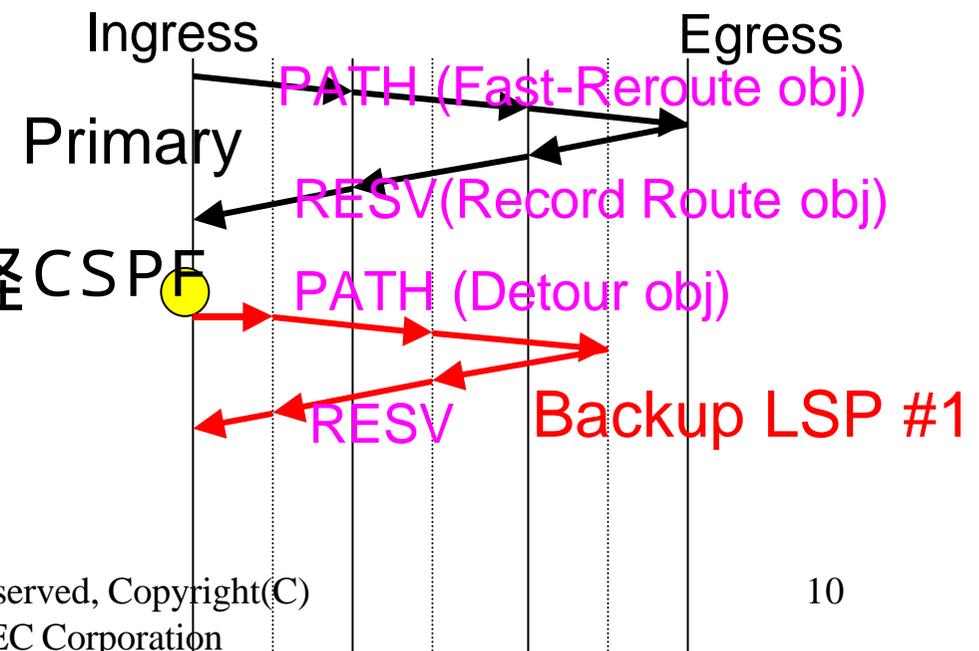
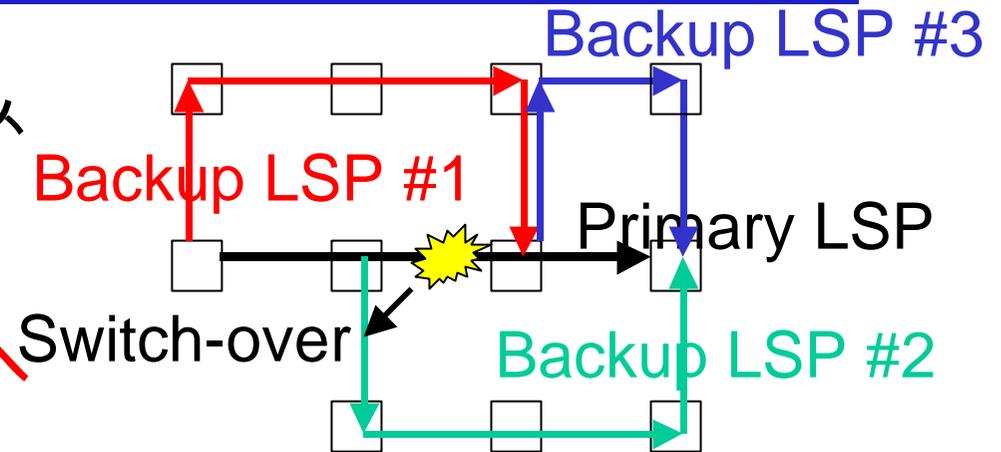
- Local RepairからGlobal Repairへの遷移による経路最適化手順



経路長、リソース利用率の最適化

# Juniper FRR #1

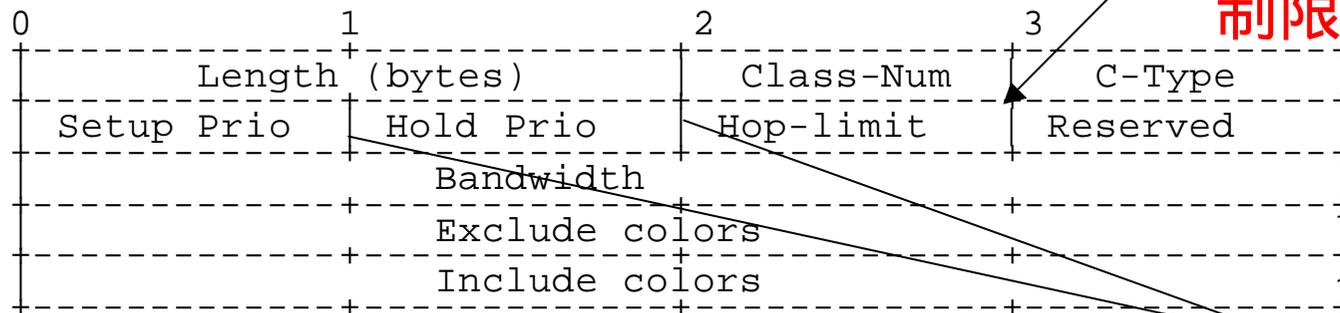
- リンク、ノード障害に対応し、Primary LSP上の全ノードからBackup LSP設定
- 1ホップ以上離れたノードへのBackup LSP
- Cisco方式のようなLSPの多重化の障害回復なし
- Primary LSPでのRESVでRecord\_Route objectによる経路上の全ノードを知る
- CSPFにより迂回経路用の経路計算実行



# Juniper FRR #2

- PATH Message (for Primary LSP)

- Fast Reroute object



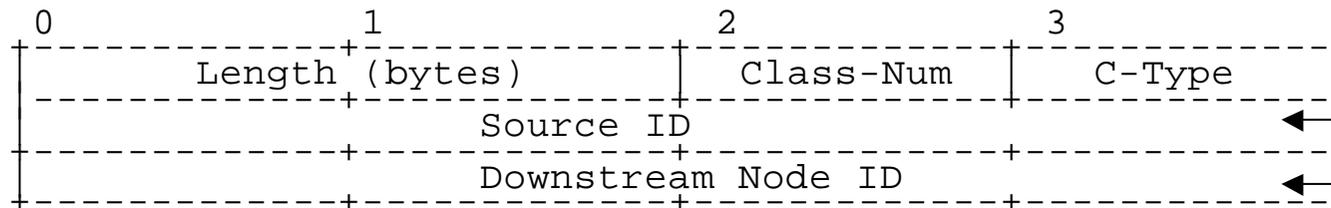
迂回経路のホップ数制限

迂回経路の特性要求

PreemptionのPriority

- PATH Message (For Backup LSP)

- Detour object

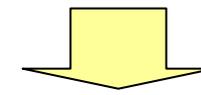
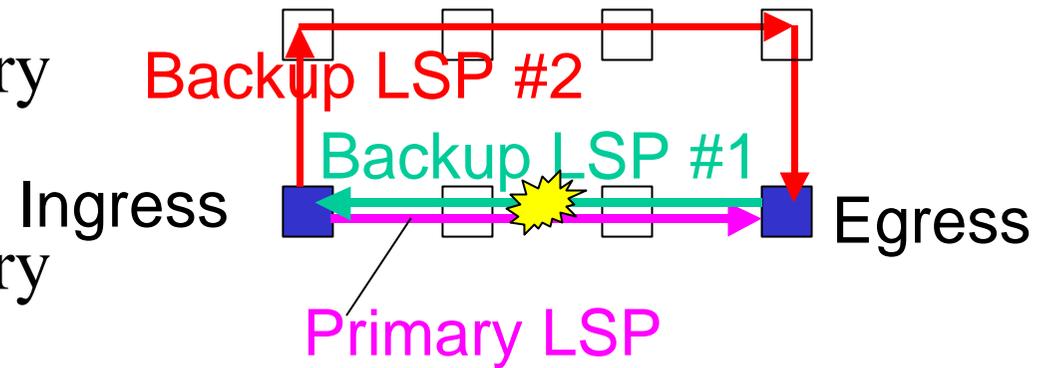


PLR

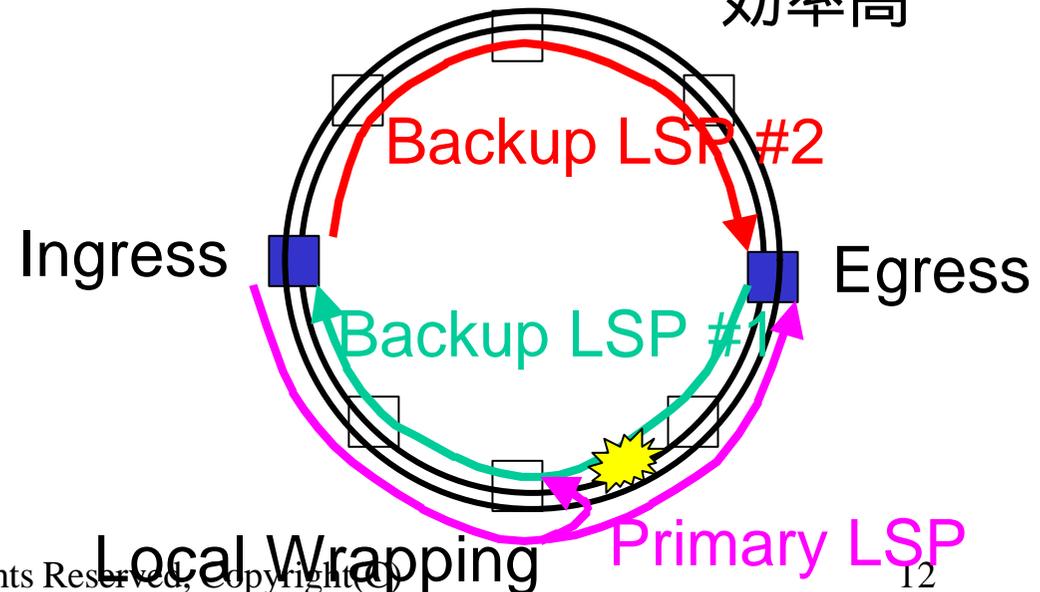
迂回経路で除くノード

# Axiowave FRR #1

- Backup LSP#1はPrimary LSPの逆向き
- Backup LSP#2はPrimary LSPとDisjoint path
- Backup LSP#1,#2の経路計算の容易性



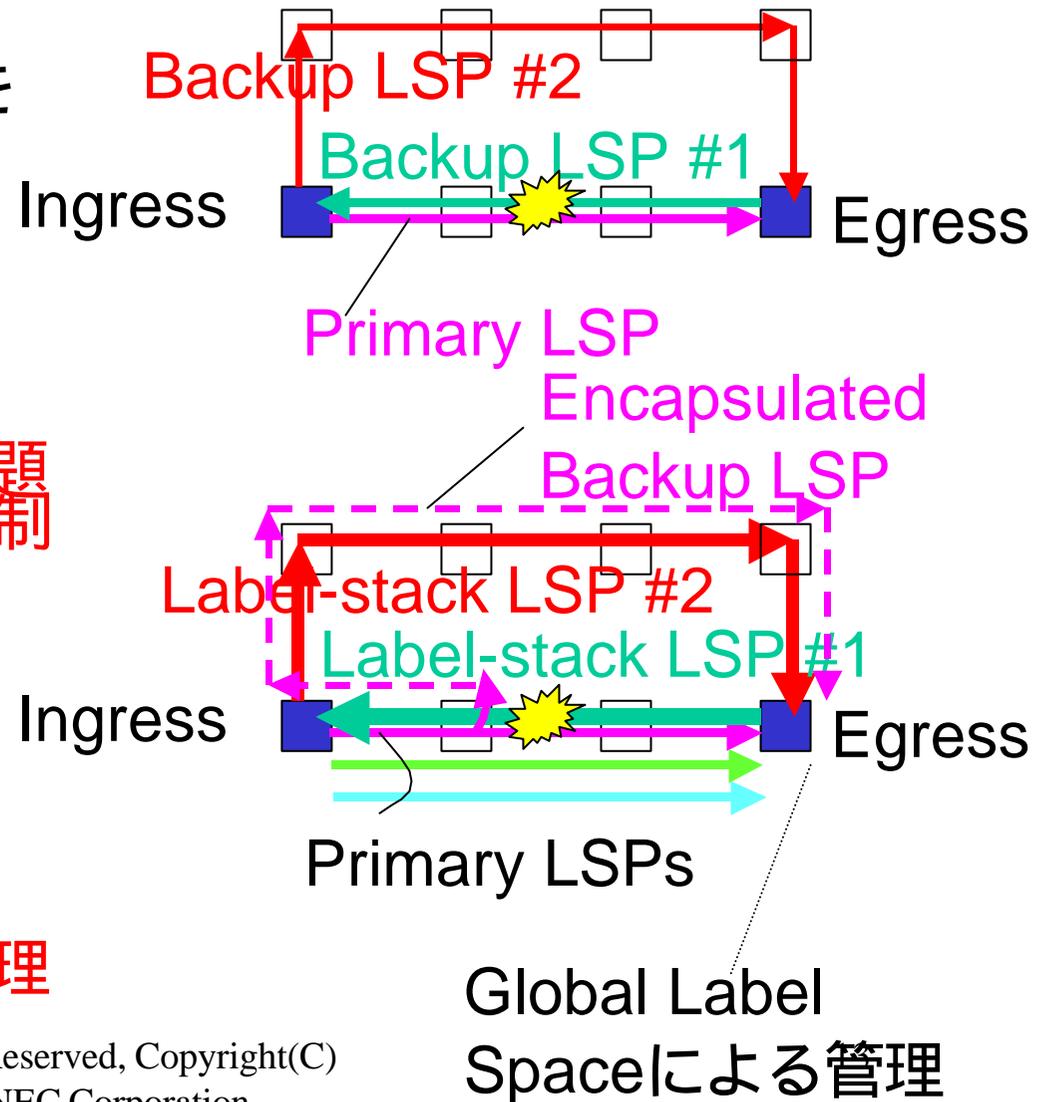
Ring網では  
効率高



Local Wrapping

# Axiowave FRR #2

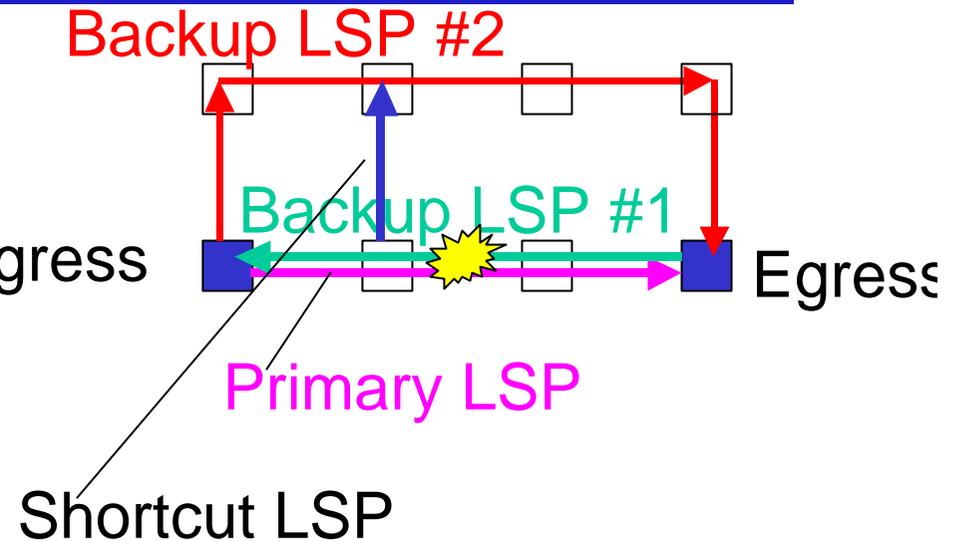
- Ingressの障害検出：
  - 逆向きのLSPにトラフィックを検出時点で検出
- 迂回経路の最適化
  - Ingressの障害検出後Backup LSP#2へ直接迂回ルーティング
  - パケットのReorderingの問題あり (IngressでのBuffering制御なしの場合)
- 1:1, N:1のProtection機能保持
  - Label StackingのBackup LSP
  - Egress nodeでのGlobal Label SpaceによるLabel管理



# Axiowave FRR #3

- **Restoration Shortcut機構**

- 特にVoIPのようなDelay sensitiveなアプリケーションにたいして、FRR後の転送遅延を小さくするためにTransit nodeからShortcut pathを許す
- Shortcut pathを行う際にはもともとの経路計算の容易性は失われる



# Local Repair方式の課題

- Cisco / Juniper / Axiowave 方式:
  - SONET protectionと同等レベルの障害回復性能を実現
    - 50 msec restoration
  - 基本的にBackup LSP (Local Repair LSP)に対して帯域確保
  - => Backup LSP間の帯域資源をできるかぎり共有化による網資源の効率的運用が必要
- Backup LSPの帯域資源の共有化に関する提案方式
  - S. Kini et al., Draft-kini-restoration-shared-backup-01.txt
  - 単一リンク、ノード障害復旧を保証する際の全 Backup LSPの中からできるかぎり共有化
  - => OSPF / ISIS の拡張による共有計算方法の提案
  - => 本コンセプトはFast-reroute (FRR)へは適用されていない。
  - => FRRへの適用 (NEC Shared Fast Reroute方式)

## - Shared Fast Reroute方式-

---

- **(2) Shared fast-reroute方式の目的**
- **(3) Shared fast reroute方式概要**
  - (I) 単一primary LSPの場合
  - (ii) 複数のload-balanced primary LSPの場合
- **(4) RSVP-TE signaling概要**
  - (I) IngressからEgress方向への分散CSPF
  - (ii) Ingress node 主体のCSPF
  - (iii) Egress からIngress方向への分散CSPF
- **(5) 性能評価例**

# Shared fast-reroute方式の目的

---

- (1) Backup LSPの帯域資源の最小化
  - LSP資源を共有化するための2つの手法
    - (I) **Label merge**: LSP資源全体を共有化 (label ID と帯域)
    - (ii) **Logical merge**: 帯域のみ共有し、Labelは非共有
- (2) Backup LSPの共有化の計算負荷を最小化
  - (I) **単一Primary LSPのBackup LSPについては最低共有化**
  - (ii) 異なる Primary LSPに対するBackup LSPsについては、S.Kini の手法で可能 (計算負荷が許せば)
  - >> **最低単純な(I)の資源共有化のアプローチを採用すべき**

# Shared fast reroute 方式概要

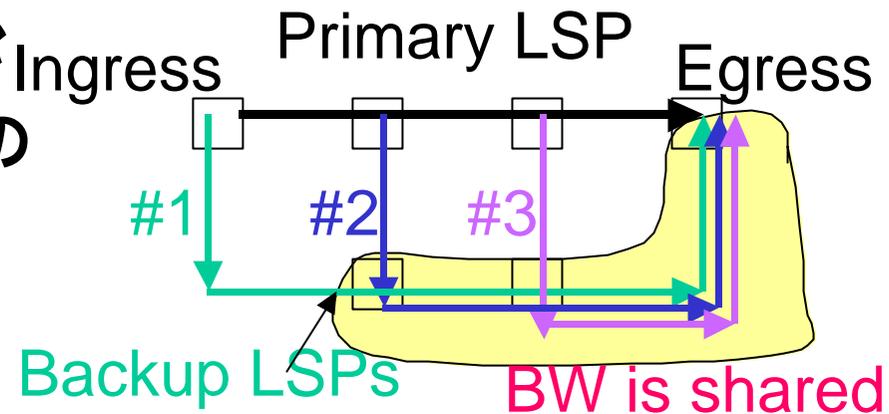
## - 単一primary LSPの場合 -

- Primary path上の全ノードが Egress 向けの backup LSP の計算 (CSPF)

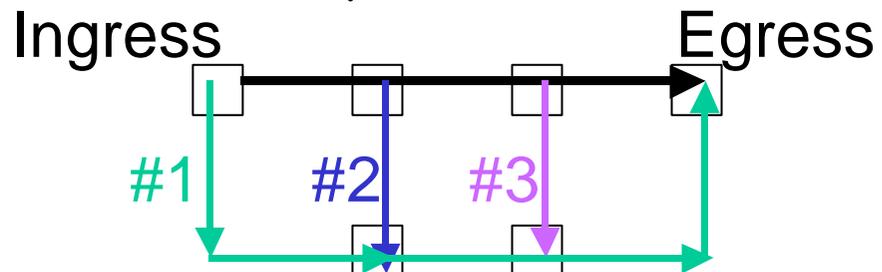
- Juniper方式の特別な状況

- Backup LSPの網リソースの最大共有化

- Backup LSP #1に帯域を共有するようにLSP #2, #3を設定
  - Label mergeを加えてさらにラベル資源も共有可能



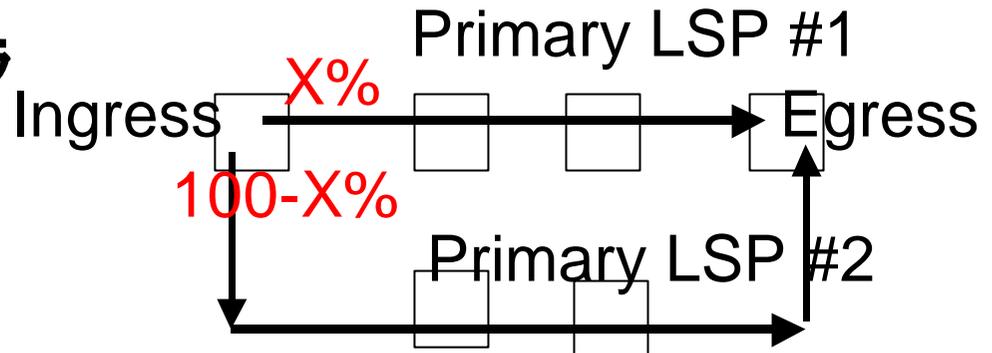
Label mergeで  
Label資源も低減



# Shared fast reroute 方式概要 **NEC**

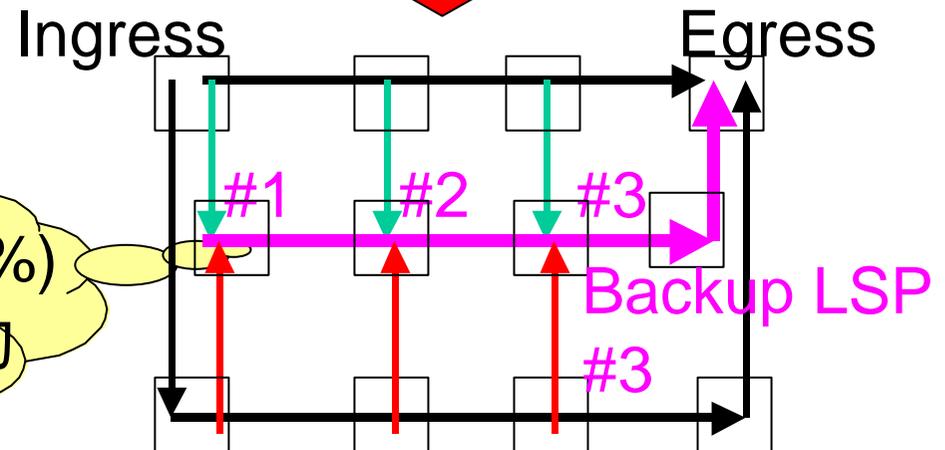
## - 複数load-balanced primary LSPの場合#1-

- Ingress – Egress pairのトラフィックをLoad balancingのためにLoad splitting
  - X% for #1, 100-X% for #2



- (1) Primary LSP #1, #2に対して共有Backup LSP #3を計算
- >> N個load balanced LSPへ拡張可能

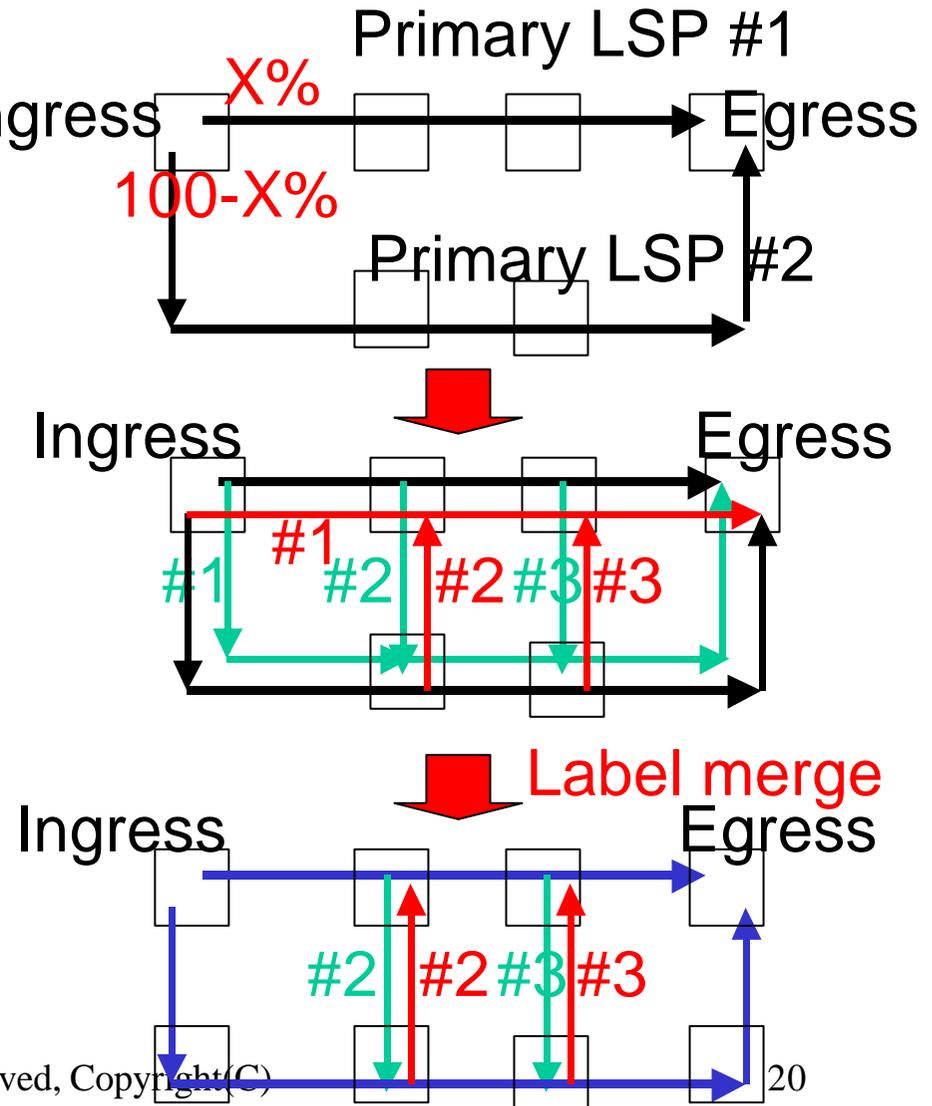
Max{x, 100-x}(%)  
の帯域のみ予約



# Shared fast reroute 方式概要 **NEC**

## - 複数load-balanced primary LSPの場合 #2-

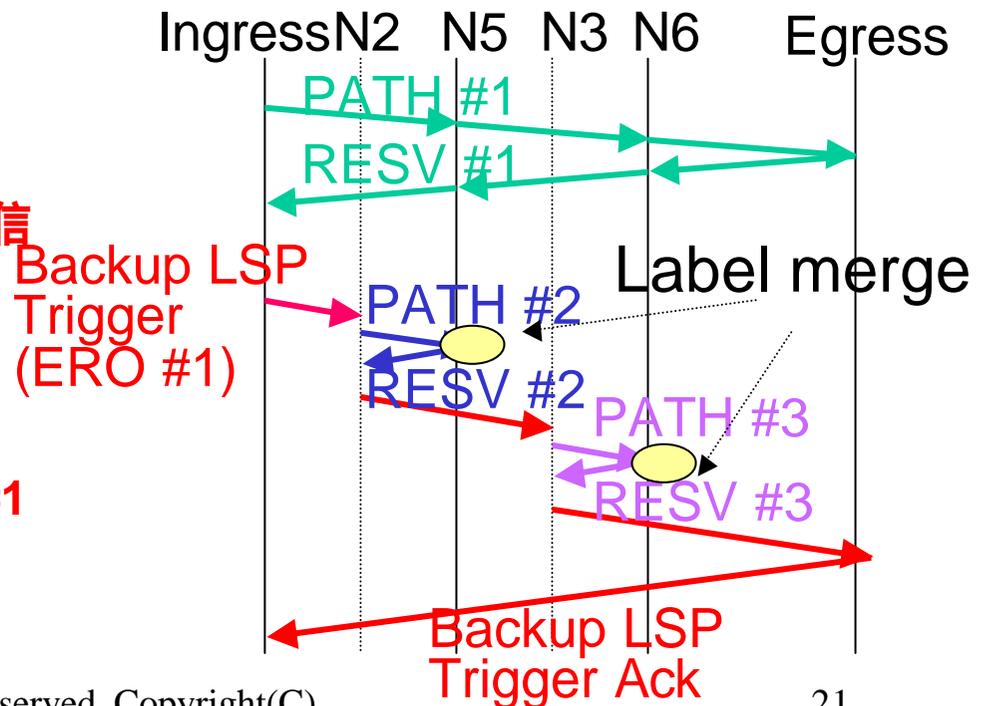
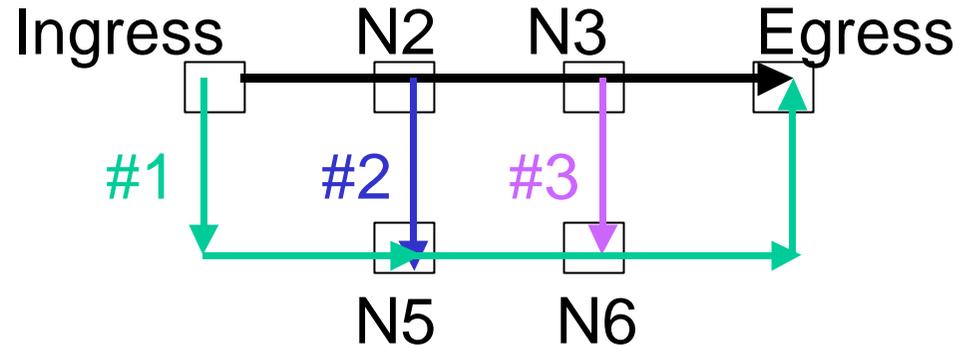
- (2) primary LSP #1, #2 に対して独立にBackup LSP #1, #2を計算
- >> N 個load balanced LSP  
へ拡張可能



# RSVP-TE signaling 概要

## - IngressからEgress方向への分散CSPF -

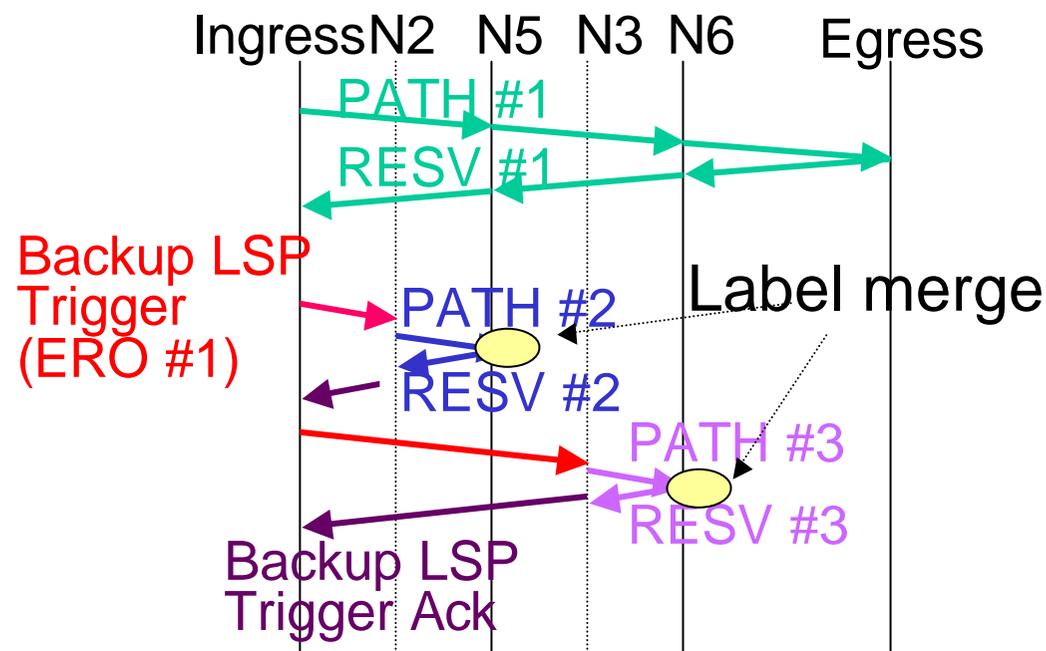
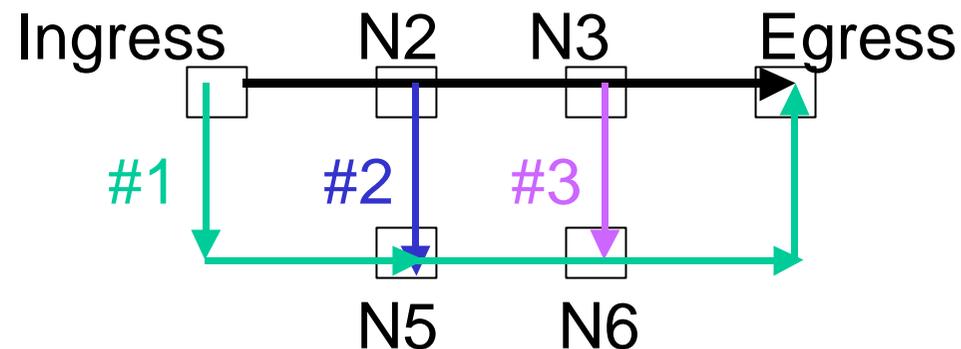
- Primary LSPが設定された後で Backup LSP設定
- Ingress, N2, N3が分散的にCSPFによりBackup LSPを設定
- Ingress :
  - Primary LSPに対してEgress宛の Disjoint なBackup LSP #1計算
  - “Backup LSP Trigger” + ERO#1 (Backup LSP#1の経路)をN2へ送信
  - N2はERO#1へマージできるように CSPF
- 中継ノード:
  - Nth backup LSP をBackup LSP#1へマージできるようにCSPF



# RSVP-TE signaling 概要

## -Ingress node 主体のCSPF-

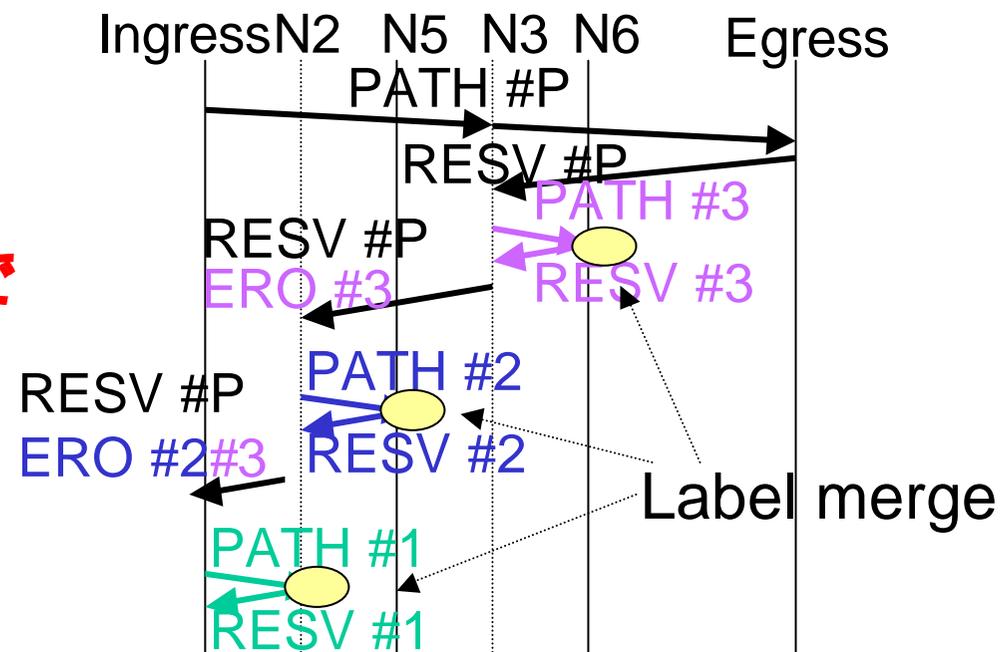
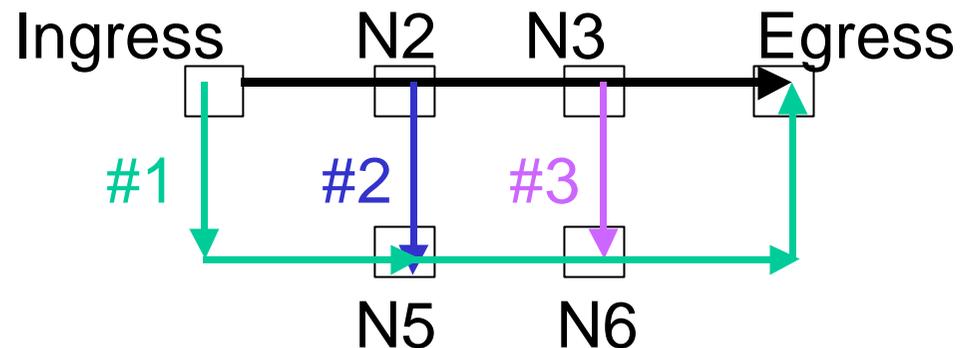
- Primary LSPが設定された後でBackup LSP設定
- Ingress, N2, N3が分散的にCSPFによりBackup LSPを設定
- **Ingress がすべてのBackup LSP経路を集中的に計算**
  - ERO#1, ERO#2, ERO#3をそれぞれIngressが計算
- 共有可能なBackup LSPの生成が容易



# RSVP-TE signaling overview

## - Egress からIngress方向への分散CSPF.

- Primary LSPとBackup LSPとを同時に設定。
- Primary LSP設定のAck (RESV)の通知と同期させながらBackup LSP設定
- **RESV messageが Backup LSPのEROを前段のLSRへ通知することで Backup LSPの共有化を図る。**

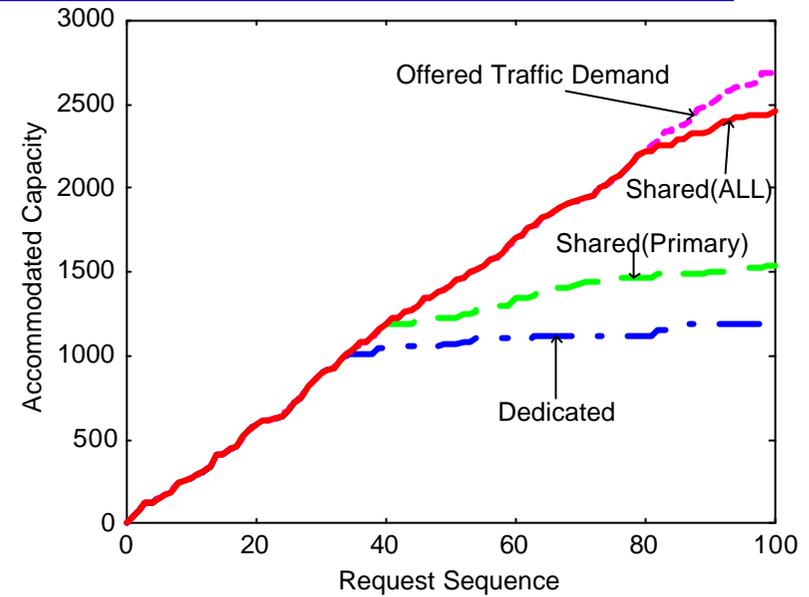
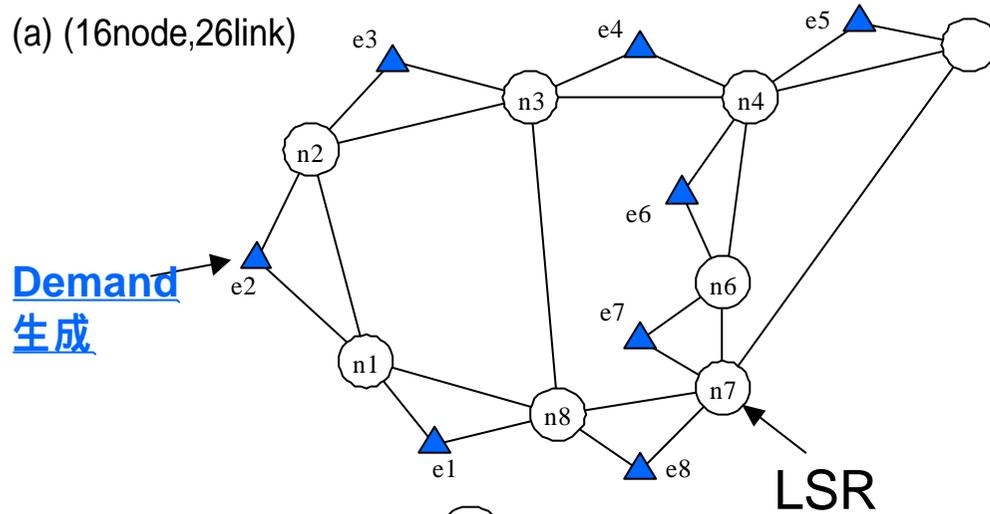


# 性能評価

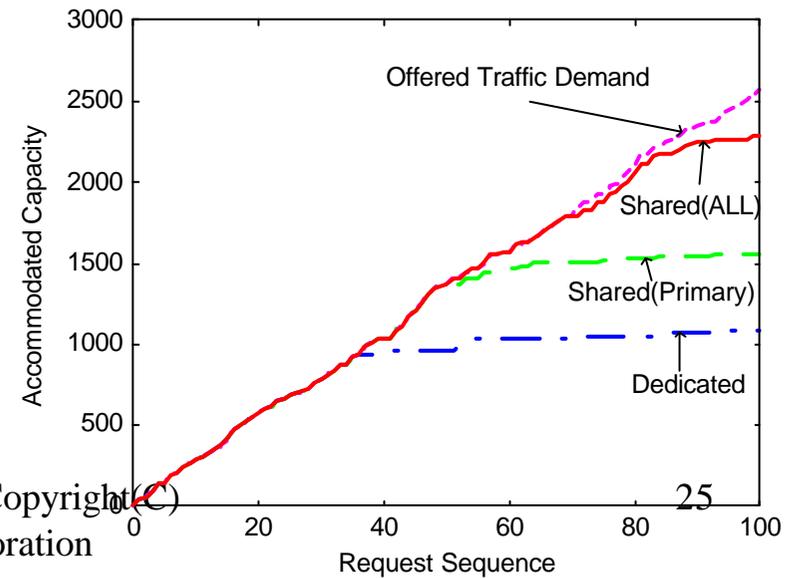
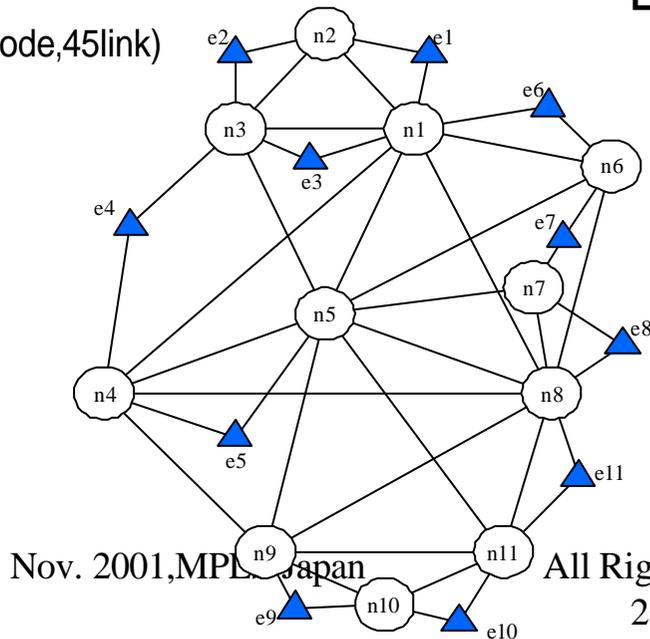
- 比較方式
  - (ALL) ネットワーク内全体の複数の異なるPrimary LSP間でBackup LSPの共有化
    - 最も必要な資源が少なくなるがネットワーク全体のBackup LSP情報を用いた最適化のため設計は複雑
  - (Primary) 一つのPrimary LSP用のBackup LSPのみ共有化
    - 設計は簡単だが ALLに比べて最適化度は低い
  - (Dedicated) Primary LSP用のBackup LSPの共有化を行わない (従来方式)
- ネットワークモデル
  - 100回のLSP設計要求が到着するシナリオ
  - 1回の設計要求で1-10個のTraffic要求あり
  - Traffic要求の対地はランダムに選択、要求容量は0.1 – 10.0で変化
  - 物理リンク容量は (16node,26link): 500, (22node,45link): 300
    - 本容量は ALLの方式ですべてTraffic要求を収容できるものとしている
  - 本性能評価結果詳細: 2001.11.21の電子情報通信学会NS研究会にて発表予定 (斎藤他, “障害迂回LSPの最適設計法”)

# 性能評価結果

(a) (16node,26link)



(b) (22node,45link)



# その他課題

---

- (1) MPLSのFRR方式の標準がない(本質的な問題)
  - 各社独自方式でInteroperabilityなし
- (2) 異なるFRR方式間のInteroperability問題
  - A. Atlas (Avici) et al., “MPLS RSVP-TE Interoperability for Local Protection/Fast Reroute”, draft-atlas-rsvp-local-protect-interop-00.txt, Jul.2001
  - Cisco, Juniper方式のInteroperabilityに関する提案
- (3) OAM方式の標準化
  - N.Harrison (British Telecom) et al., “Requirements for OAM in MPLS Networks,” draft-harrison-mpls-oam-req-00.txt, May.2001
  - 障害回復を行う際の基準となる障害検知のためにOAMが必要であるという提案。

# まとめ

---

- MPLSにおける障害回復方式概要、ならびに高速障害回復方式 (Fast Reroute: FRR)に関する標準化動向を述べた。
- IETF MPLS/CCAMP WGにおいてFRRに関し標準化の合意を行うのが先決
- 障害回復方式の性能面、実装容易性などに踏み込んだ詳細な議論が不可欠