



現状のMPLSの動向について

PIL workshop 2003

平成15年2月19日  
古河電気工業株式会社  
ファイテルネットワーク研究所  
村上 哲也

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



1



MPLS関連技術

- IP-VPN (RFC2547bis)
- Carrier's Carrier
- EoMPLS(draft Martini)
- EoMPLS(draft vKompella)
- 6PE
- LDP over RSVP
- Fast Reroute
- Graceful Restart
- トラフィックエンジニアリング(OSPF-TE, RSVP-TE)

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



2



**We can.**  
古河電工

## Signalling protocol動作モード

---

- Label distribution mode
  - Downstream Unsolicited(DU)  
明示的な要求無しにラベルを配布可能なモード
  - Downstream-On-Demand(DoD)  
明示的な要求に対してのみラベルを配布可能なモード
- Label control mode
  - Independent control  
どのタイミングでもラベルを配布可能なモード
  - Ordered control  
下流からのラベル割り当てを契機にラベルを配布するモード
- Label retention mode
  - Conservative retention  
受けたmappingの必要なもののみ保持するモード
  - Liberal retention  
受けたmapping全てを保持するモード

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

**FITEL**  
net 3



**We can.**  
古河電工

## IPVPN(RFC2547bis)

---

- Layer3 VPNサービス
- 網内のパケット転送にMPLS、VPN経路交換にBGPを使用
- MPLS(ラベル)によるカプセリングにより、論理的なクローズドネットワークを実現
- Customer毎にルーティングテーブルを形成  
(Private Address等を自由に持ち込める)

---

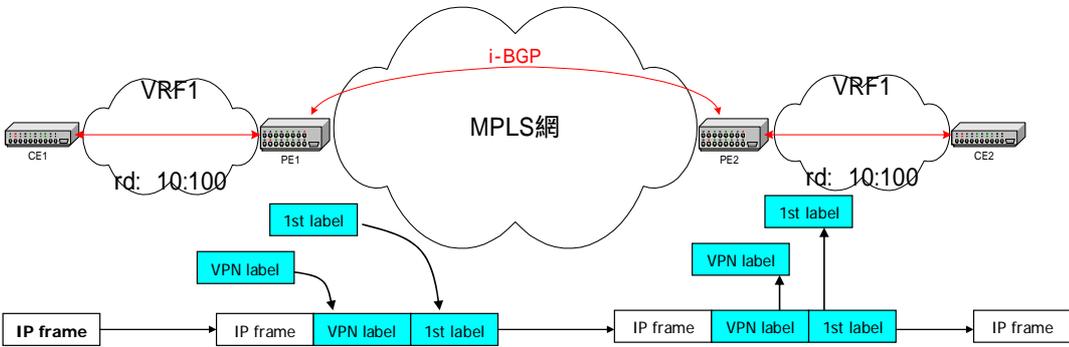
Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

**FITEL**  
net 4



## IPVPN(RFC2547bis) Cont...





- ・VPN毎に異なるルーティングテーブル(VRF)を持つ。
- ・各VRFは、rd(route distinguisher)により識別される。
- ・PE-PE間では、VRF内の経路交換およびVPNラベルの配布をi-BGPにより行われる。
- ・egress側のPEでは、パケット内のVPNラベルによりVRFを特定する。
- ・MPLS網内では、PE-PE間転送のためのラベルを交換する(LDP, TDP, RSVP等)。
- ・MPLS網内のコアルータでは、VPNラベルを意識することなく、1段目ラベルのみで転送を行う。

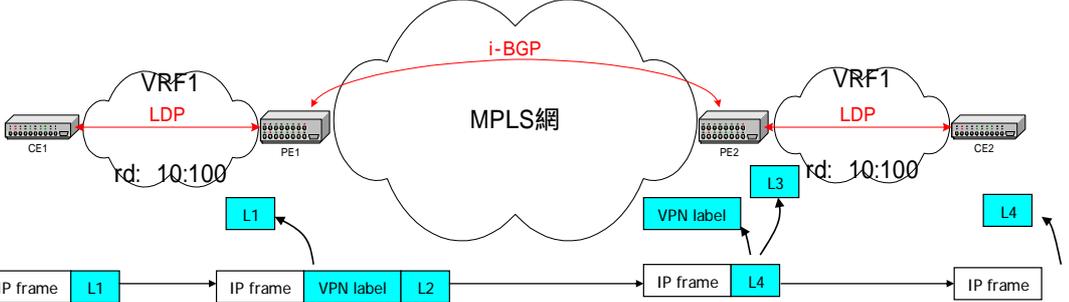
Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


5



## C'sC(Carrier's Carrier)





- ・IPVPNモデルベース
- ・PE-CE間にてLDPにより、ラベル配布を行う。
- ・CE-CE間のLSPをMPLS網内では、VPN-LSPによりトンネリングする。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


6



EoMPLS(draft Martini)



---

- Point-to-PointのLayer2 VPNサービス
- VPWS(Virtual Private Wire Simulation)
- IP以外の既存プロトコルでも転送
- Customer側のルーティングプロトコルに依存しない

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

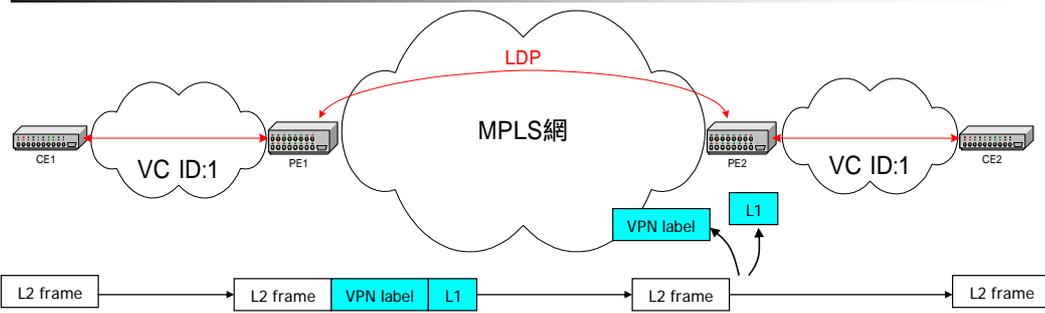

7



EoMPLS(draft Martini) Cont...



---



The diagram illustrates the EoMPLS architecture. It shows two customer edge (CE) devices, CE1 and CE2, connected to two provider edge (PE) devices, PE1 and PE2, via a central MPLS network. LDP signaling is shown between PE1 and PE2. Below the network, the flow of an L2 frame is shown: it enters from the left, is encapsulated with a VPN label and L1 at the ingress PE, travels through the MPLS network, and is decapsulated at the egress PE before being sent to the destination CE.

- ・VC ID(4bytes)の値により、VPNを識別
- ・MPLS網内では、PE-PE間でLDPによりVC ID、VPN用ラベルを交換
- ・LDPのsignallingは、PE-PE間で行うためtargeted hello(unicast)を使用
- ・signallingにより、交換する情報はVC ID、VPNラベル、MTU値
- ・VC ID、MTU値等が一致しない場合には、VPNをupさせない
- ・CE側から受信したLayer2 frameをそのままMPLS(ラベル)にてカプセリング
- ・ingress側PEでは、VC IDからVPN labelを特定する
- ・egress側PEでは、VPN labelからVC IDを特定し、ラベルをはずした後CE側に転送

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


8



EoMPLS(draft vKompella)



---

- Point-to-MultipointのLayer2 VPNサービス
- VPLS (Virtual Private LAN Simulation)
- 全てのPE間にてメッシュ状にVC Pathを形成
- MPLS網側からの受信Layer2 frameからSA学習を行う
- 未学習の場合には、floodingを行う
- Loop防止のため、Split Horizonを使用

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

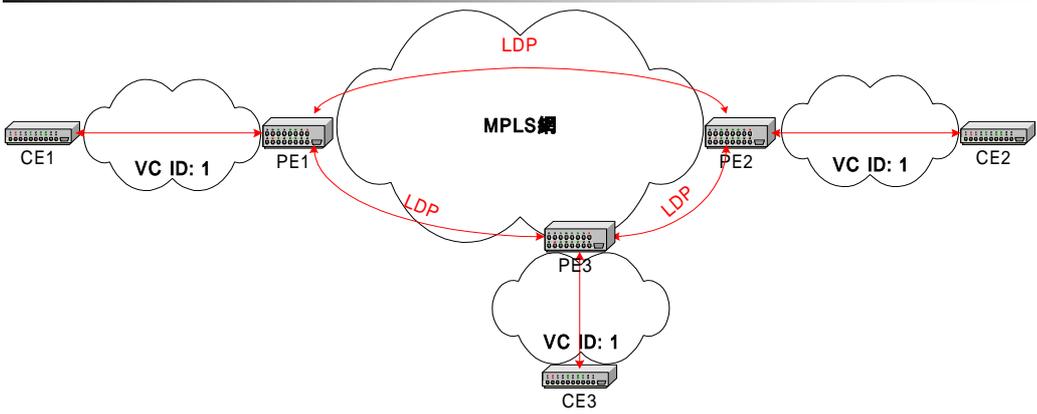

9



EoMPLS(draft vKompella) Cont...



---



・全てのPE間でメッシュ上にLSPを形成  
 ・LDPを用いて、VC ID、MTU値、VPNラベルを配布  
 ・VPNラベルは、対向するPE毎に異なる値を割り当て

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

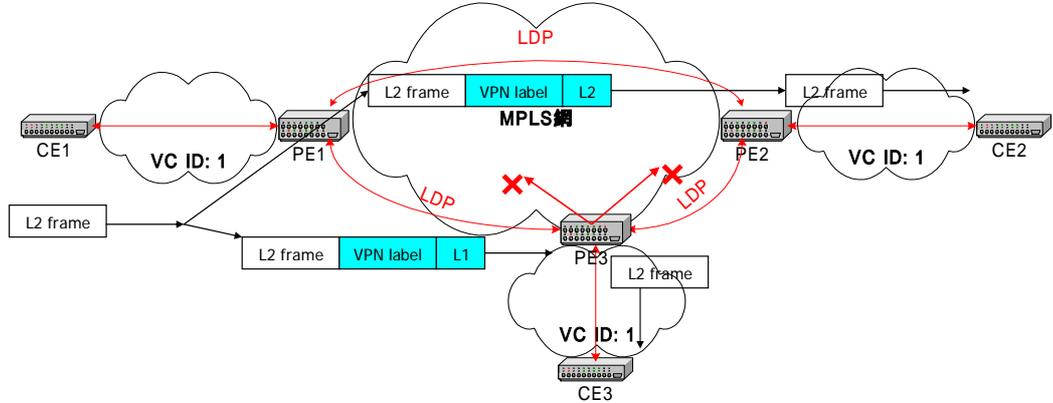

10



### EoMPLS(draft vKompella) 転送処理



---



- ・CE側から受信したLayer2 frameのDest Macが未学習の場合には、受信フレームをコピーし、VCに属する全てのPEに対し、送信する。
- ・egress側PEでは、MPLS網内から受信したフレームに関しては、未学習の場合でもMPLS網側への転送は行わない(Split Horizon)
- ・egress側PEでは、受信したフレームのSA(Source Address)とVPNラベルの学習を行い、学習したaddressへの送信時には、該当するPEにのみ転送を行う。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


11



### 6PE



---

- IPv6パケットをMPLS(ラベル)にてカプセルング
- mpBGPにて、IPv6の経路情報、ラベルを配布し、IPv6のreachabilityを確保
- 対向側PEまでの転送用ラベルをLDP、RSVP等を用いて配布
- MPLS網内は、対向側PEへの転送ラベル、IPv6識別用ラベルの2段ラベルにて転送
- MPLS網内のコアルータは、MPLSのスイッチのみ行えばよい

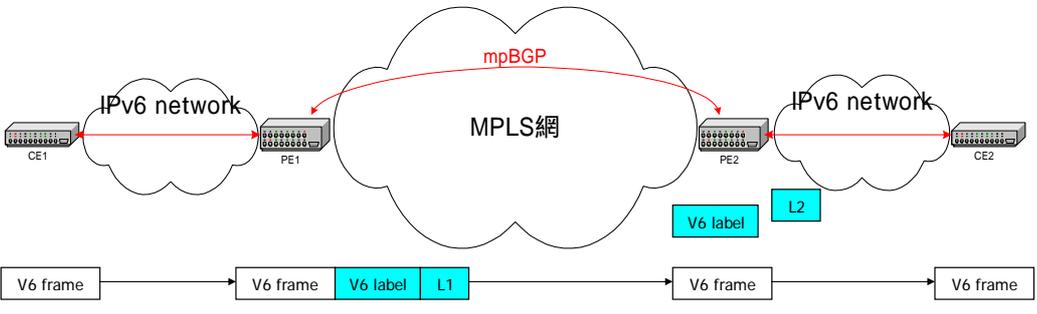
Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


12



## 6PE Cont...





- ・PE-PE間にてmpBGPを用いIPv6経路情報、ラベルを交換し、CEから送信されるIPv6パケットのreachabilityを確保
- ・PE-PE間では、MPLS転送用ラベル配布のため、LDP、RSVP等を使用

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


13



## LDP over RSVP



- RSVPで形成したLSP(PE-PE間)上にLDPによるLSP(CE-CE間)を形成
- PE-PE間は、RSVPによりLSP(転送用ラベル)を形成
- CE-CE間でのLSP形成のため、PE-CE間ではLDPを、PE-PE間ではtargeted hello(unicast)によるLDPを使用
- CE-CE間のLSPをMPLS網内では、RSVPにより形成したLSP上に載せる(ラベルのスタック)

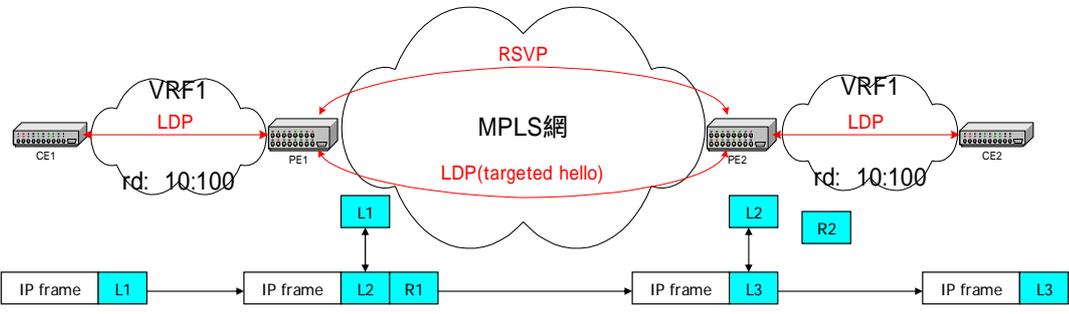
Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


14



## LDP over RSVP





- ・PE-PE間のLSP形成には、RSVPを使用
- ・CE-CE間のLSP形成には、PE-CE間でLDP、PE-PE間ではtargeted helloによるLDPを使用
- ・MPLS網内では、CE-CE間のLSP(LDP)をPE-PE間のLSP(RSVP)に載せる(ラベルのスタック)

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


15



## Fast Reroute



- LSPの冗長構成において、障害発生時にLSPの切り替えを行う。
- リンク異常時のリンクプロテクション
- ノード異常時のノードプロテクション
- Fast Reroute機能では、シグナリング・プロトコルのRSVPで、複数のう回経路を自動検出
- OSPFがRSVPで検出したう回経路の中で最短のものを障害時のう回経路として、あらかじめ記憶する。このため、障害が起こった場合でも50ミリ秒で切り替え可能

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


16

A2 . 現状の MPLS の動向について

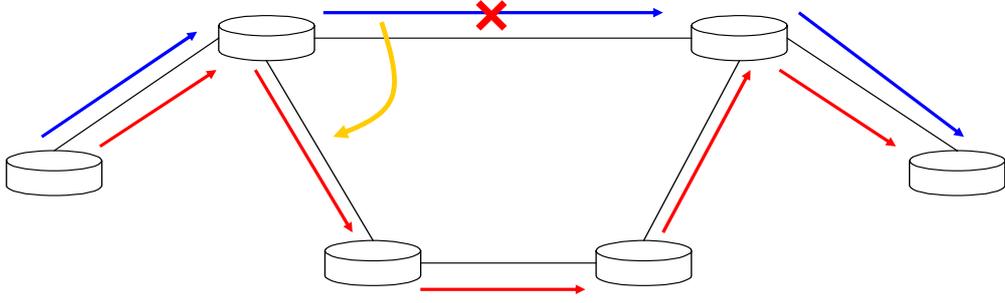
村上 哲也



### Fast Reroute(リンクプロテクション)







→ 通常時のLSP (Blue arrow)

→ バックアップLSP (Red arrow)

・リンクプロテクション

- 特定リンク切断時に高速に切り替え (約50ms)
- その後に最適化経路変更 (re-optimize)
- バックアップLSPでは、 tunnel labelをスタック
- Mergeノードの1hop前にてPHPを行う。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

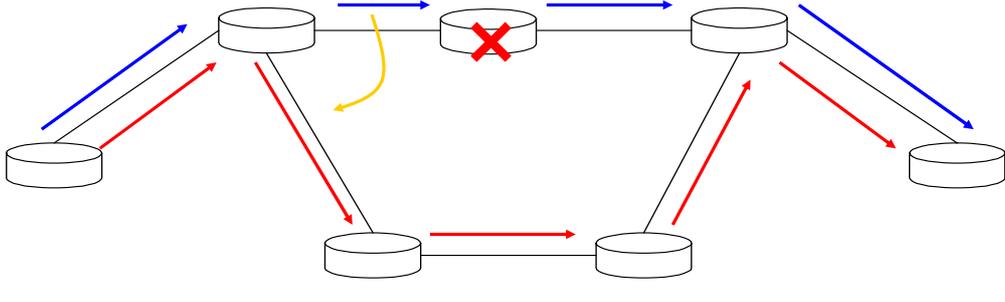

17



### Fast Reroute(ノードプロテクション)







→ 通常時のLSP (Blue arrow)

→ バックアップLSP (Red arrow)

・ノードプロテクション

- 特定のノードがダウンした時に高速切り替え
- バックアップLSPでは、 tunnel labelをスタック
- Mergeノードの1hop前にてPHPを行う。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


18



**Graceful Restart**



古河電工

- Control Plane(signaling)の再起動時に、Data Plane(forwarding処理)を中断することなく、LSPおよびControl Planeの状態を復帰させることが可能
- Control Planeの再起動後、既にあるLSPの再構成は行わず、新規のLSPの形成のみ行うため、復帰までの時間を短縮可能
- Control Planeの障害時の対応としてnetwork topologyに依存していたが、Data Planeが維持されることにより、network topologyの複雑さを軽減可能
- Data Planeに影響なく計画的にControl Planeの再起動が可能
  - LDP、RSVP、BGP等のsignallingプロトコルにてサポート可能
  - Downstream Unsolicitedモードで動作(LDP)
  - Downstream On Demandモードでの動作は検討中

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



19



**RSVP-TE**



古河電工

- Point-to-PointのLSPをDoDで形成
- LSPの経路を指定するER(Explicit Route)機能
  - Strict ER  
使用するLSPを明示的に指定
  - Loose ER  
使用するLSPはIGPの最短経路に従う
- LSPの使用する帯域幅の指定が可能  
指定した経路に必要な帯域がない場合、LSPがはれない。  
各経路の帯域幅を認識する必要がある      OSPF-TE

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



20



## OSPF-TE



古河電工

- 各経路の帯域幅を認識することが可能
- 経路情報にリンクの帯域幅と「カラー」と呼ばれる所属するグループの情報を持つ
- 1つの経路情報を運用者のポリシーに基づきカラーごとの経路情報に分けて使い、さらにその経路情報の中で有効な帯域幅をダイナミックに伝えることが可能

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



21



## Traffic Engineering



古河電工

- Downstream on Demand
  - 起点から終点へのトラフィック流量を制御、把握可能
- Constraint Base Source Routing
  - 帯域や各種Metricをパス選択に反映可能
- CAC
  - 条件を満たさないLSPははらない。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



22



MPLS-Diffserv PHB(RFC3270)



- RSVP-TEで帯域幅を指定 (signaling protocol)
- 実際のパケットに対するPHB(Drop policy, Schedule policy)は？
  - E-LSP  
SHIMヘッダ内のEXPビットによりPHBを指定する。但し、3ビットのため8 classに制限される。
  - L-LSP  
ラベル値によりPHBを指定する。EXPビットをもてないCell mode MPLSにとっては必須

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

 23



MPLS Diffserv-aware-TE (DS-TE)



- TEとMPLS Diffservを結び付け、さらに発展させたもの
  - クラス別のリソース管理
  - クラス別のCAC
- 現状のDS-TEの実装では、クラスは2つ
  - 通常のTE tunnel
  - Premium traffic用tunnel
- 各クラス毎に上限帯域を設け、MPLS Diffservと組み合わせることにより、細かな帯域制御が可能

---

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.

 24



次世代IX研究会 (Distix)



古河電工

- 3つのworking group
  - ルータ相互接続working group  
MPLSを応用したIXアーキテクチャに関して、マルチベンダ環境での相互接続性について、試験や検証
  - IXプロバイダworking group  
MPLSを応用したIXの仕組みを提供するIXプロバイダの技術的な要件や課題などについて検討
  - IXユーザworking group  
IXに接続する参加者、すなわち「IXユーザ」の立場から、IXに接続しトラフィックを交換するために必要な技術や課題などについて検討

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



25



次世代IX研究会 (Distix) Cont...



古河電工

- ルータ相互接続試験内容
  - コア動作試験
  - エッジ動作試験
  - IPVPN
  - Carrier's Carrier
  - 6PE
  - EoMPLS
  - LDP over RSVP

第5回まで終了、試験内容、結果は全てDistixのホームページにて公開されています。

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.



26



**We can.**  
古河電工

## 次世代IX研究会 (Distix) 試験内容例(第5回抜粋)

- Edge-RSVP(interface, loopback)機能試験
  1. PHPなしでラベル16以上を広告？
  2. LSPの受信フィルタ(シグナリングで受けたFECのフィルタに対して)？
  3. LSPトラフィックがSNMPで取得できる？
  4. LSPのup/downがSNMP trapで取得できる？
  5. PINGによるデータ疎通試験
- Carrier's Carrier
  1. PINGによるデータ疎通確認
- L2VPN (martini)試験
  1. PINGによるデータ疎通確認
- L2VPN (vkompella)試験
  1. PINGによるデータ疎通確認
- 参加メーカー  
Cisco, Juniper, River Stone, NEC(CX), 富士通, Foundary, 日立, AYAME, 古河  
(順不同)

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


27


**We can.**  
古河電工

## 次世代IX研究会 (Distix) Cont...

- 古河は、第1回から第5回まで全てに参加
- 初期の段階では、LDPすらも接続できないことも・・・
- 現在では、LDPやRSVPの基本的なプロトコルに関しては大丈夫
- 今後は、EoMPLS、6PE、LDP over RSVP等の応用へと・・・
- 各ベンダー間での実装の違い
  - PHPの実装(あり、なし)
  - Label retentionモードの違い(Conservative or Liberal)
  - TTLの扱い(IP TTL and MPLS TTL)
  - VLANタグの扱い(EoMPLS)
- マルチベンダー環境への取り組み
- 実装上の問題や課題の洗い出し
- 各ベンダーの進捗状況

Copyright (c) 2003 The Furukawa Electric Co., Ltd. All Rights Reserved.


28