

ATMとインターネット

1. ATM技術の導入の背景
2. ATM技術の概要
3. ATM技術を使ったIPサービス
 - (1) NBMAリンクモデル
 - (2) ポイントポイントリンク
 - (3) ラベルスイッチングモデル
4. ATMリンクを用いたネットワーキングの例
5. まとめ

(株)東芝 研究開発センター

江崎 浩(hiroshi@isl.rdc.toshiba.co.jp)

1. ATM技術導入の背景

- ◆ キャリアによるATM技術の計画的導入
- ◆ 高速リンクの必要性
 - OC-3(155Mbps), OC-12(622Mbps), OC-48(2.4Gbps)
- ◆ 帯域幅の柔軟性の要求
 - 例: メガリンクサービス(by NTT)
- ◆ 大容量スイッチングの必要性
 - ギガビット、テラビット スwitchング
- ◆ QoS(Quality of Service)の要求

1. ATM技術導入の背景(続)

◆ ポート単価の下落

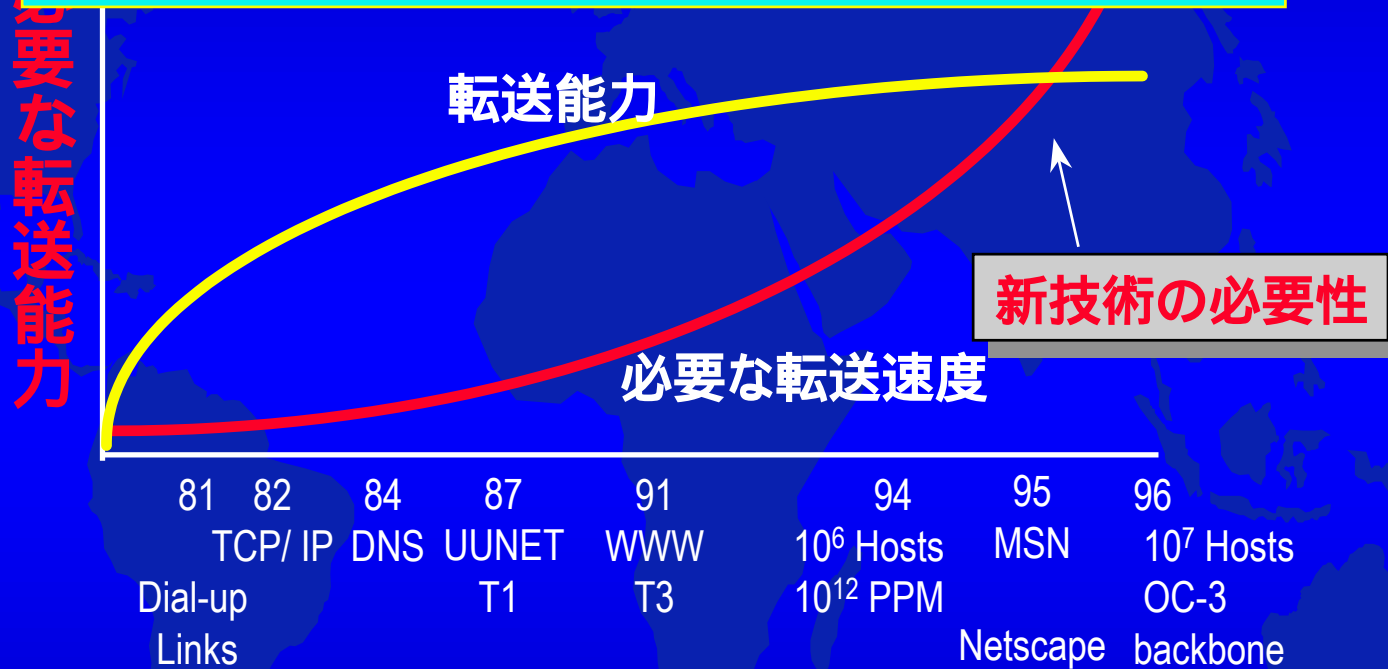
- レイヤ2スイッチ(e.g., イーサースイッチ)
安価 \$100/10Mbps ?
- レイヤ3スイッチ(i.e., ルータ)
高価 \$10,000 / 100Mbps ?

インターネットサービスへの要求

- ◆ 高いスループット・低遅延パケット転送
- ◆ リアルタイムパケット転送 (QoS保証)
- ◆ スケーラビリティ
 - 不均一な自律ネットの相互接続
 - 大量のパケット処理能力
 - 不確実なノード (host/router)の動作
- ◆ マルチキャストサービス

次世代高速ルータの必要性

WWW is World Wide Wait !!



インターネットの成長

- ⊆ これまでのルータは十分高速であった
- ⊆ 必要なパケット転送能力は指数関数以上の速度で増加している
- ⊆ 既存ルータの処理能力が1997-1998年にシステムボトルネックになる

2. ATM技術の概要

- ◆ ATM技術の標準化の経緯
- ◆ ATM技術の特長
- ◆ 既存ネットワーク技術の特徴
(利点・欠点)
- ◆ ATM技術の概要

ATM技術標準化の経緯

- ◆ **ITU-T**; BISDNの実現解 ('87 ~)
 - '90; 第1フェーズ(Point-to-Point)
 - '92 ~ ; 第2フェーズ(Point-to-Multipoint)
- ◆ **ATM Forum**; Private ATM網仕様 ('91 ~)
 - '91; UNI 1.0
 - '94; UNI 3.1 (Point-to-Multipoint)
- ◆ **IETF**; IP over ATM の仕様 ('92 ~)
 - ION(Internetworking Over NBMA) WG
 - MPLS(MultiProtocol Label Switching) WG

ATM技術標準化の経緯(続)

◆ IETF ; IP over ATM の仕様 ('92 ~)

– IP over ATM WG

- ◆ Multi-Protocols on AAL5 (RFC1483)
- ◆ Classical IP and ARP over ATM (RFC1577)
- ◆ Default MTU for ATM AAL5 (RFC1626)
- ◆ Multicast Address Resolution Protocol (MARS)

– ROLC(Routing Over Large Cloud) WG

- ◆ NBMA Address Resolution Protocol
- ◆ NBMA Next Hop Resolution Protocol (NHRP)

– ION(Internetworking Over NBMA) WG

- ◆ Server Cache Synchronization Protocol

ATM技術標準化の経緯(続)

- ◆ **IETF** ; IP over ATM の仕様 ('92 ~)
 - MultoProtocol Label Switching WG ('97 ~)
 - ✦ ATMリンク
 - ✦ SONET/SDHリンク
 - ✦ Ethernet リンク
 - ✦ フレームリレーリンク

ATM技術の特徴

- ◆ 通信プロトコルの簡略化
- ◆ ハードウェアによるセル転送
- ◆ コネクションオリエンティッド通信; hard-state管理
- ◆ コネクションレス通信; ISO CLNPの導入
- ◆ アドレッシング; NSAP(ISO)、GOSIP、E.164
- ◆ 移動対応; Address Registration
- ◆ リンク経路制御; リンクステート型プロトコル
- ◆ 2レベルのフロー制御

ATM技術の特徴(続)

◆ 2レベルのフロー制御

– コネクション(フロー)ごとの制御

✦ Connection Admission Control (CAC)

✦ 経路制御(Routing Control)

✦ UPC(Usage Parameter Control), NPC (Network P.C.)

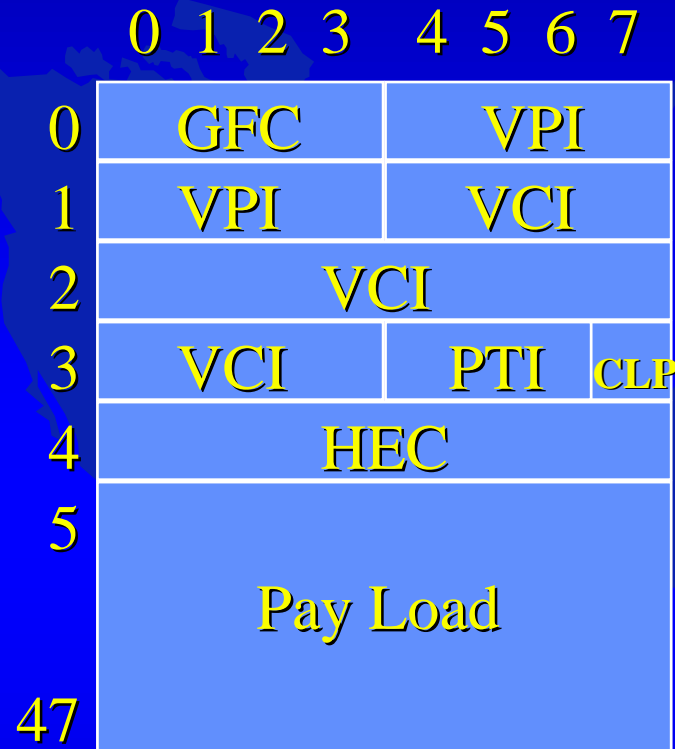
– セル(フロー/コネクション内)ごとの制御

✦ ECN (Explicit Congestion Control)

✦ フレームドロップ

✦ ERD(Early Random Discard)

ATMセルフォーマット



<<UNI cell format>>

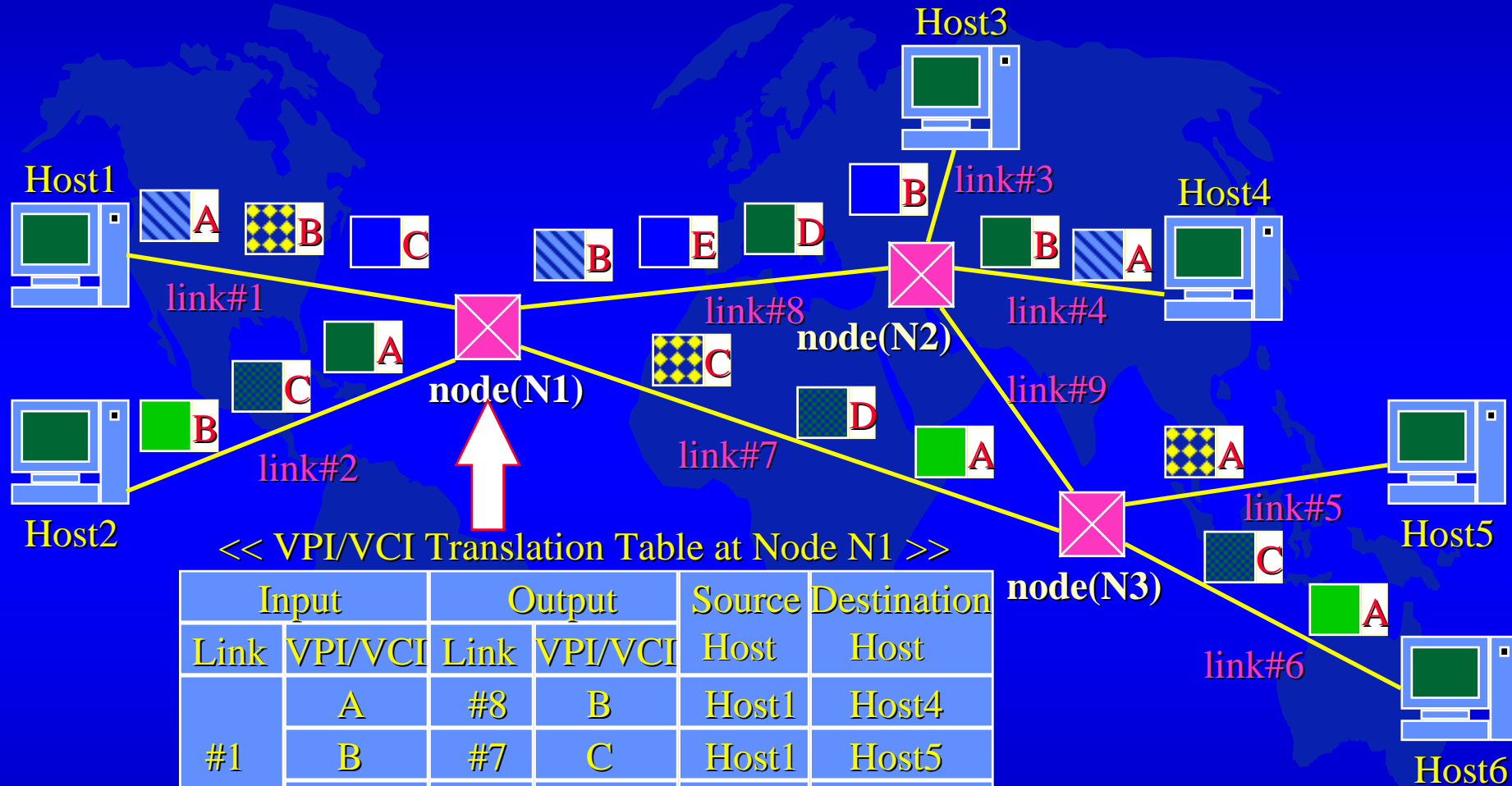
- GFC; Generic Flow Control
- VPI; Virtual Path Identifier
- VCI; Virtual Connection Identifier



<<NNI cell format>>

- PTI; Payload Type Identifier
- CLP; Cell Loss Priority
- HEC; Header Error Control

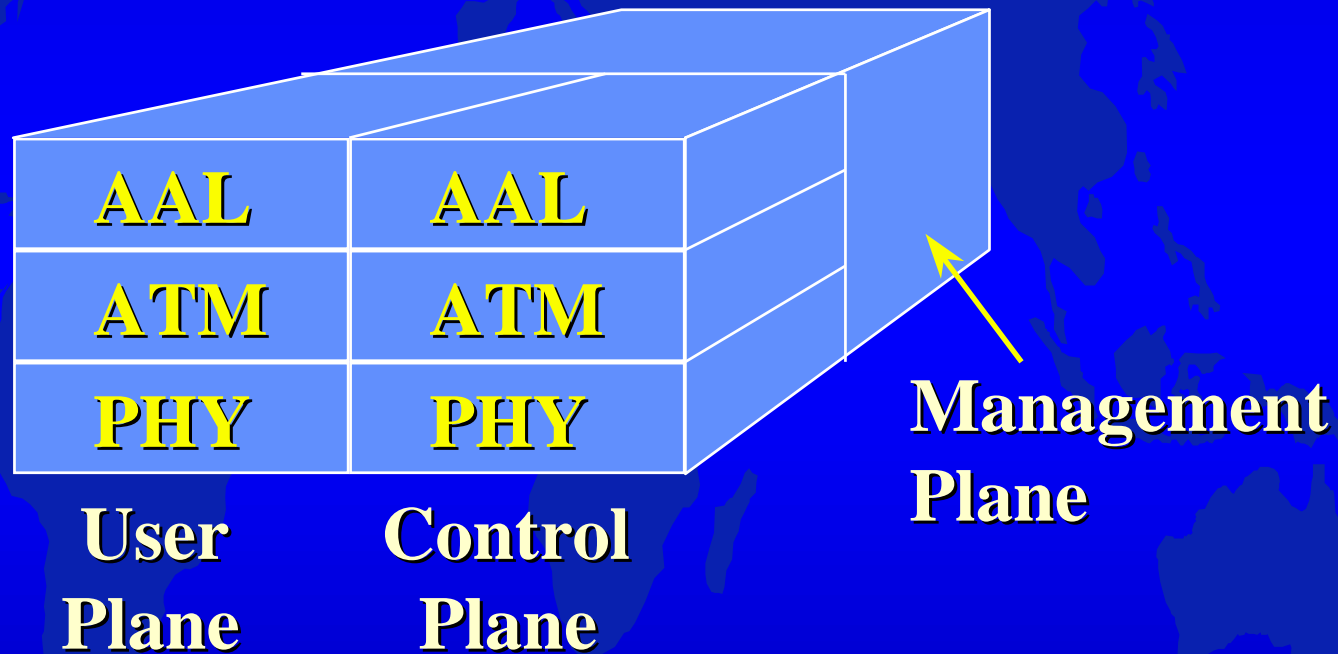
ATMセル転送の概要



<< VPI/VCI Translation Table at Node N1 >>

Link	Input		Output		Source Host	Destination Host
	Link	VPI/VCI	Link	VPI/VCI		
#1		A	#8	B	Host1	Host4
		B	#7	C	Host1	Host5
		C	#8	E	Host1	Host3
#2		A	#7	D	Host2	Host4
		B	#8	A	Host2	Host6
		C	#8	D	Host2	Host6

ATMネットワークプロトコル構成



ATMサービス機能とプロトコル構成

Table 2-1. Service Categories in ATM Networks

Attribute	ATM Layer Service Categories				
	CBR	VBR (RT)	VBR(NRT)	ABR	UBR
CLR	specified			specified	unspecified
CTD and CDV	specified		specified	unspecified	unspecified
PCR and CDVT	specified			specified	specified
SCR and BT	n/a	specified		n/a	
MCR	n/a			specified	n/a
Control Information	no			yes	no

ATM網でのトラヒック制御

◆ 仮想コネクションの設定

ハードステート経路の確立(by Q.2931、B-ISUP)

- ◆ Permanent型; PVP, PVC
- ◆ On-demand型; SVC, (SVP)

<< トラヒック制御フレームワーク >>

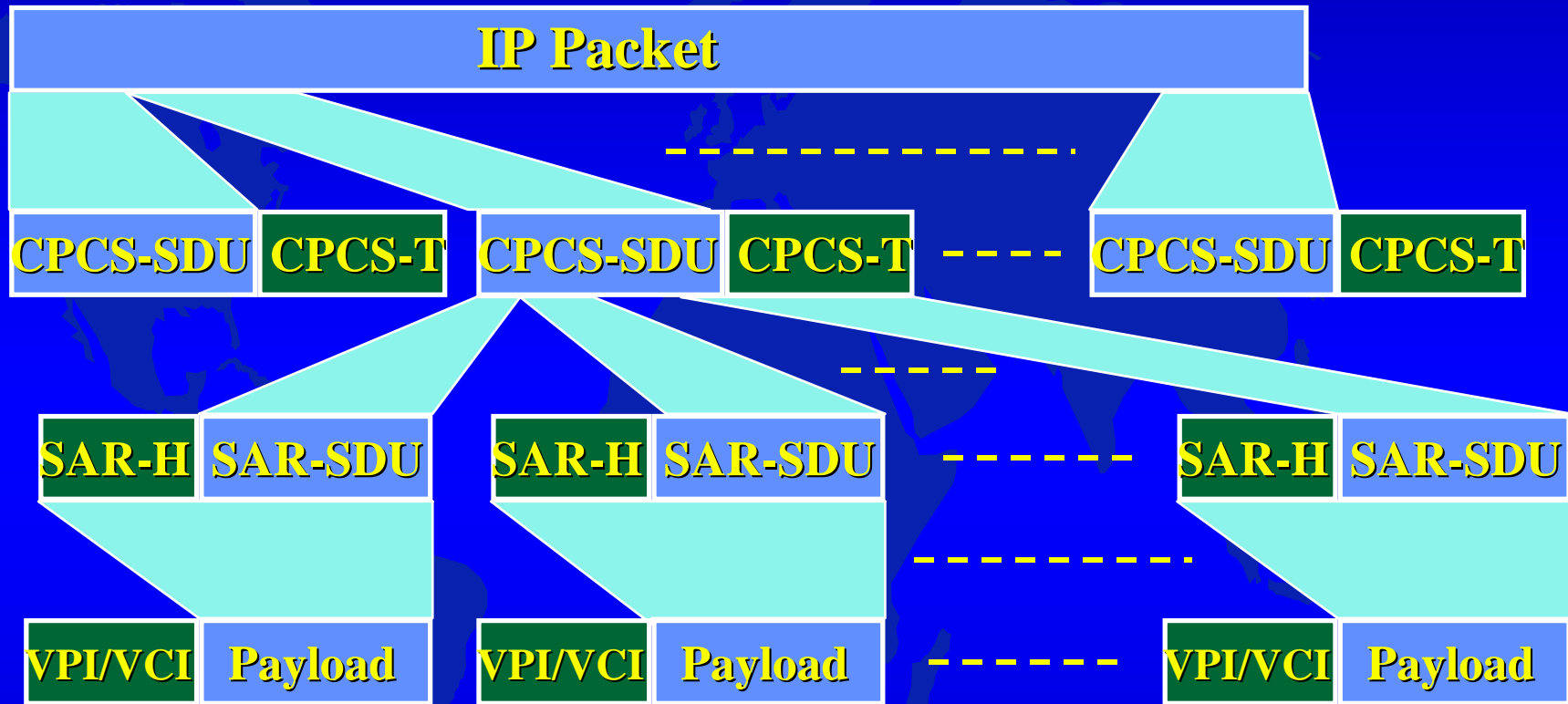
コネクション設定レベル	セル設定レベル
<ul style="list-style-type: none">・ コネクション受付制御(CAC)・ ルーティング制御(P-NNI)	<ul style="list-style-type: none">・ トラヒック整形(shaping)・ ポリシング(i.e.,UPC/NPC)・ Reactiveフロー制御(ABR)・ フレームレベルでの選択的セル廃棄(e.g., ERD)

ATM網での誤り制御機能

<< ATM網における誤り検出 / 訂正機能 >>

		Error Detection		Error Correction	
		Payload (SDU)	Header	Payload (SDU)	Header
PHY		No	Yes (HEC)	No	1 bit (HEC)
ATM		No	Yes (HEC)	No	1 bit (HEC)
SAR	Type 1	No	Yes (SNP)	No	1 bit (SN)
	Type 3/4	Yes(CRC10)	Yes (CRC10)	No	No
	Type 5	No	No	No	No
CPCS	Type 1	No (optional)	No (optional)	No (optional)	No(optional)
	Type 3/4	Yes (LI for lost)	No	No	No
	Type 5	Yes(CRC32/LI)	Yes (CRC32)	No	No

ATM網 上でのTCP/IPパケット転送



CPCS-SDU: CPCS Service data Unit
CPCS-T : CPCS Trailer
SAR-SDU : SAR Service Data Unit
SAR-H : SAR Header

ATM技術の特徴

◆ 利点

- クロック階層からの解放
- ネットワーク資源(帯域・バッファ)の共有
- 低遅延でのデータ転送(パイプラインデータ転送)
- ハードウェア実装の容易性
- スケーラビリティ
- 柔軟な帯域幅の提供
- フルライン速度でのQOS提供

ATM技術の特徴

◆ 欠点

- 複雑なトラフィック制御
 - ✦ e.g., シェピング制御
- 大きな転送オーバーヘッド
- Fatal share (ハードステート管理)
- コネクション設定遅延
- セル化に伴うパケット廃棄確率の増加

3. ATM技術を使ったIPサービス

1. NBMAリンクモデル

- + CLPF (Classical Packet Forwarding) モデル
- + SCPF (Short-Cut Path Forwarding) モデル

2. ATMポイントポイントリンクモデル

3. ラベルスイッチングモデル

3. ATM技術を使ったIPサービス

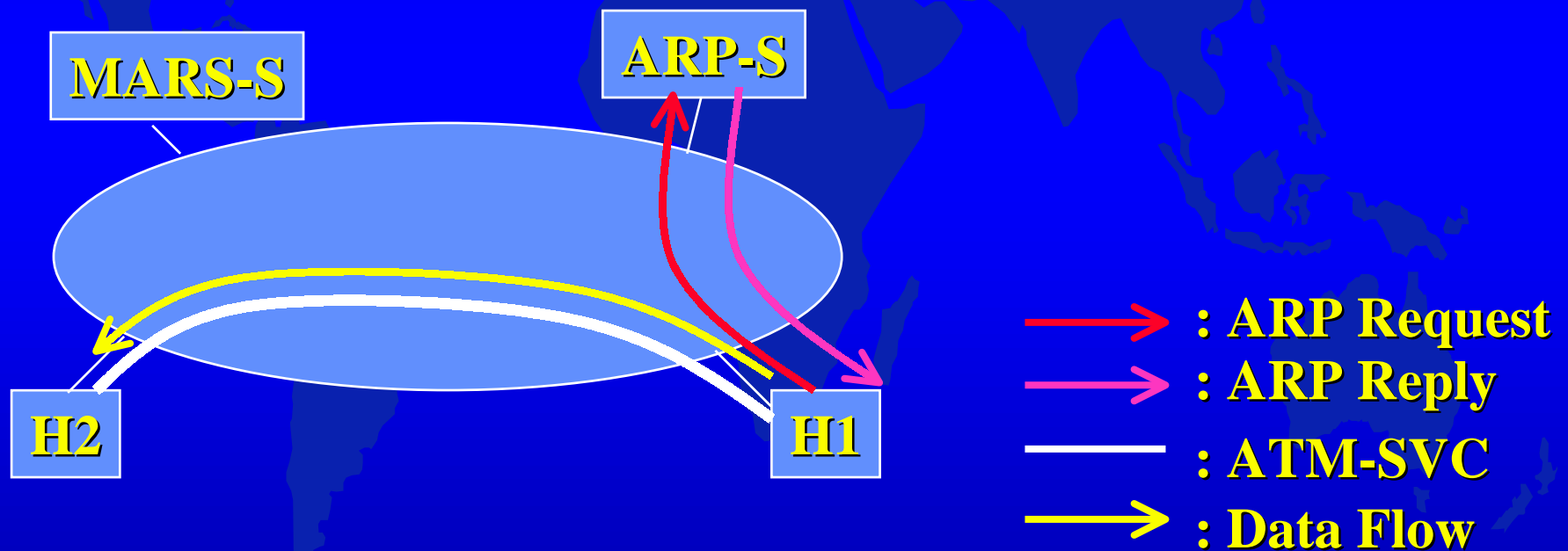
TIP: ATMリンク上でのパケット転送の注意点!

- (1) セル転送速度のシェイピングが必須!**
(注) シェイピングを行ってもセルが落ちることがある!
- (2) セル流の合流には注意!**
スイッチ内のバッファ容量

Classical IP (RFC1577)

- by IETF -

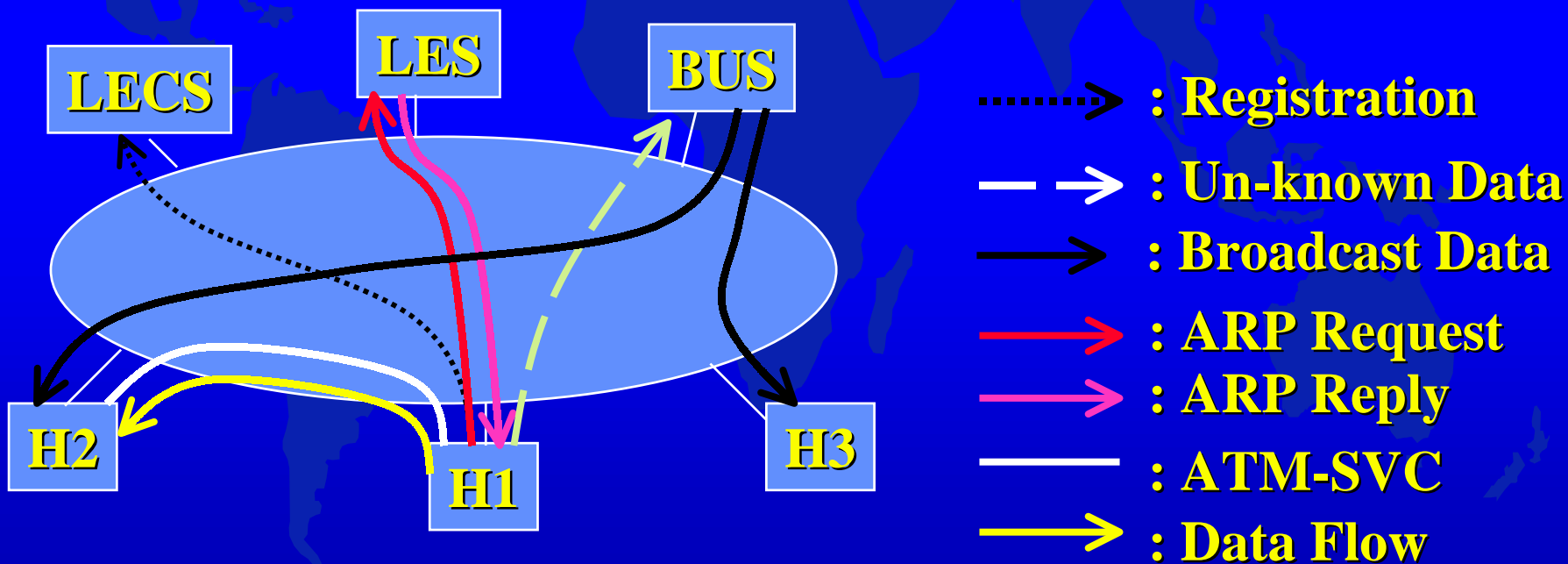
- ◆ ARPサーバ: IPアドレス → ATMアドレス
- ◆ (MARSサーバ): マルチキャスト対応



LAN Emulation

- by ATM Forum -

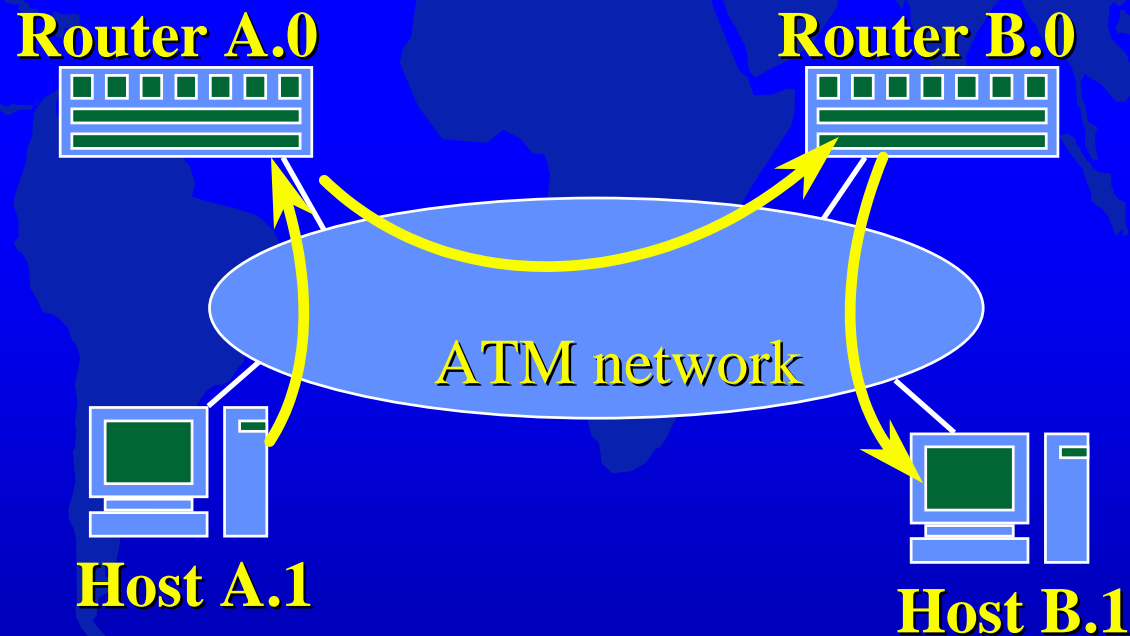
- ◆ LECS : LAN Emulation Configuration Server
- ◆ BUS : Broadcast and Un-known Server
- ◆ LES : LAN Emulation Server



CLPFモデル

- Classical Packet Forwarding -

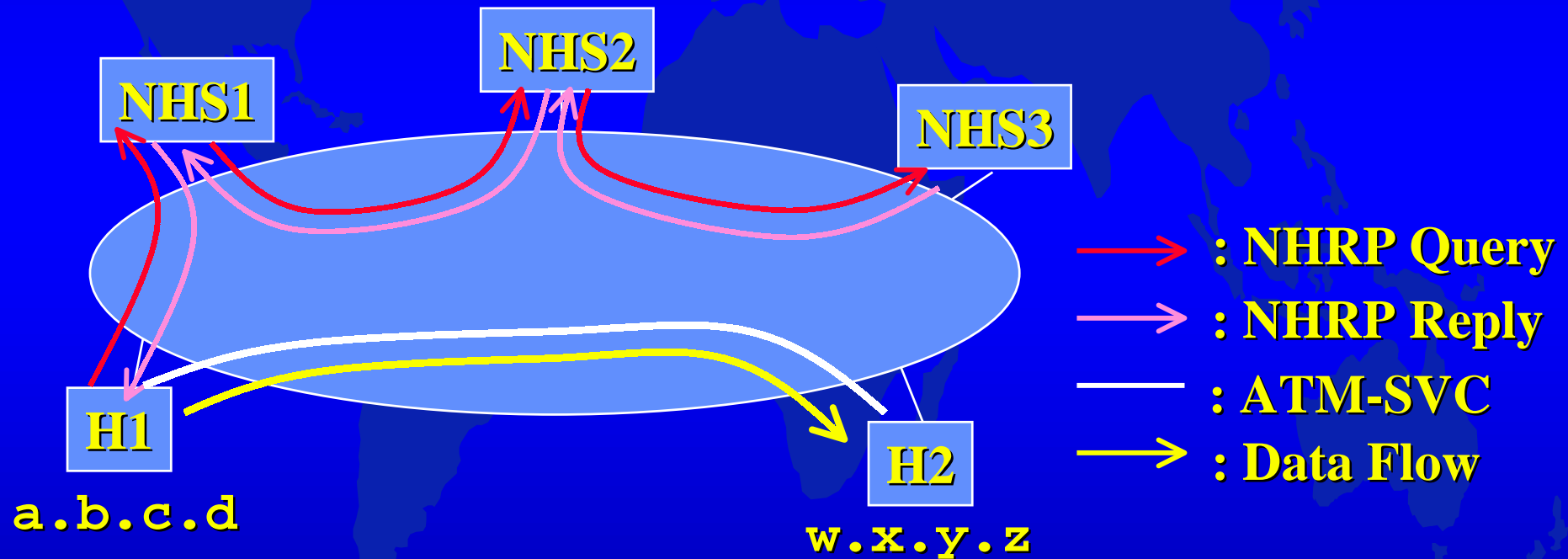
- ◆ ソフトウェア処理ルーター
 - サブネット間通信の packets 転送能力
- ◆ マルチキャストサービス能力



NHRP: Next Hop Resolution Protocol

- by IETF -

◆ NHS : Next Hop Server

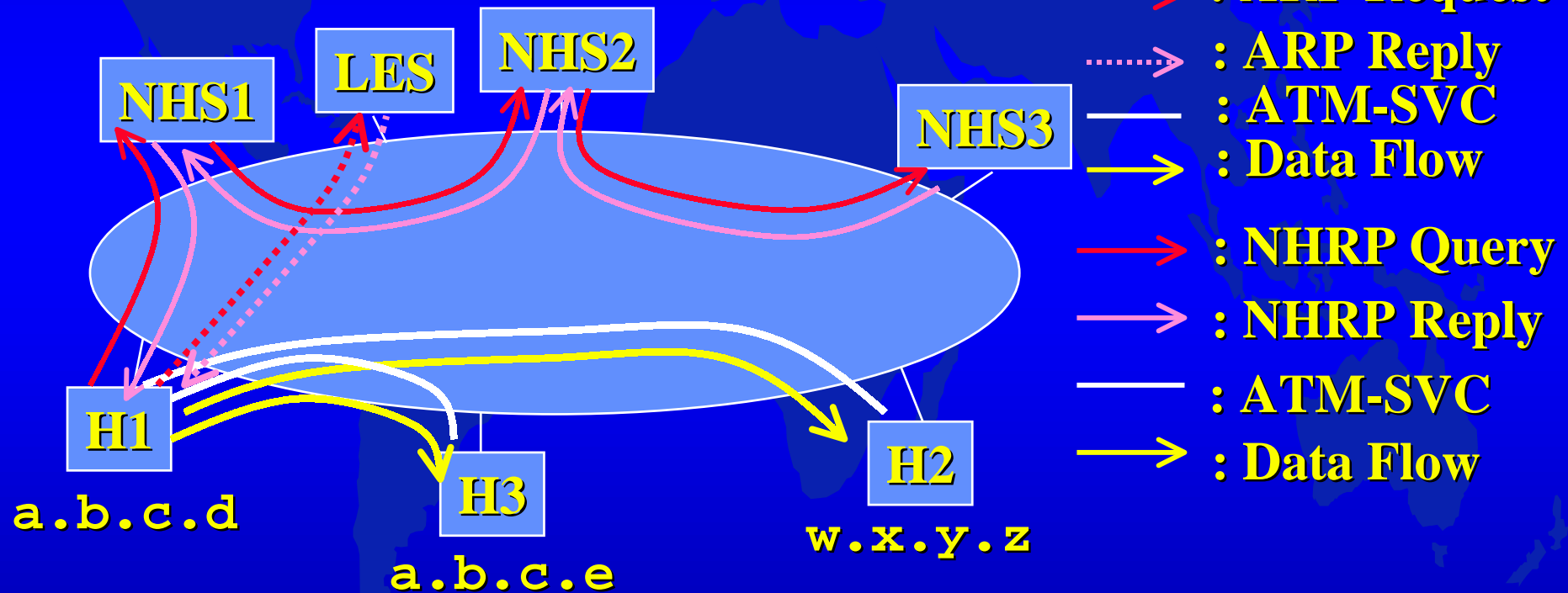


MPOA: Multi-Protocol Over ATM

- by ATM Forum -

◆ NHRP

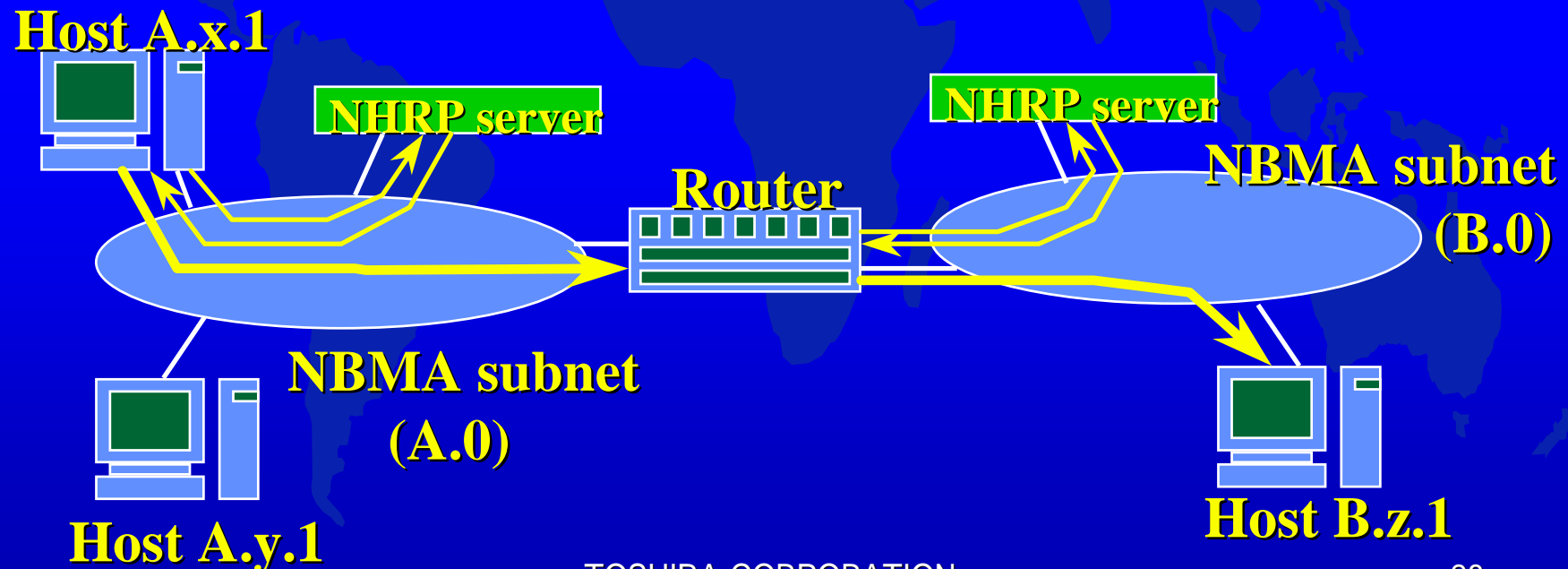
◆ LANE



SCPFモデル

- Short-Cut Path Forwarding -

- ◆ 長時間ルーティングループの形成の可能性
- ◆ ソフトウェア処理ルーター
 - NBMAネット間でのデータ転送速度
- ◆ 2種類のルーティングプロトコル



NBMAリンクモデルの適用領域

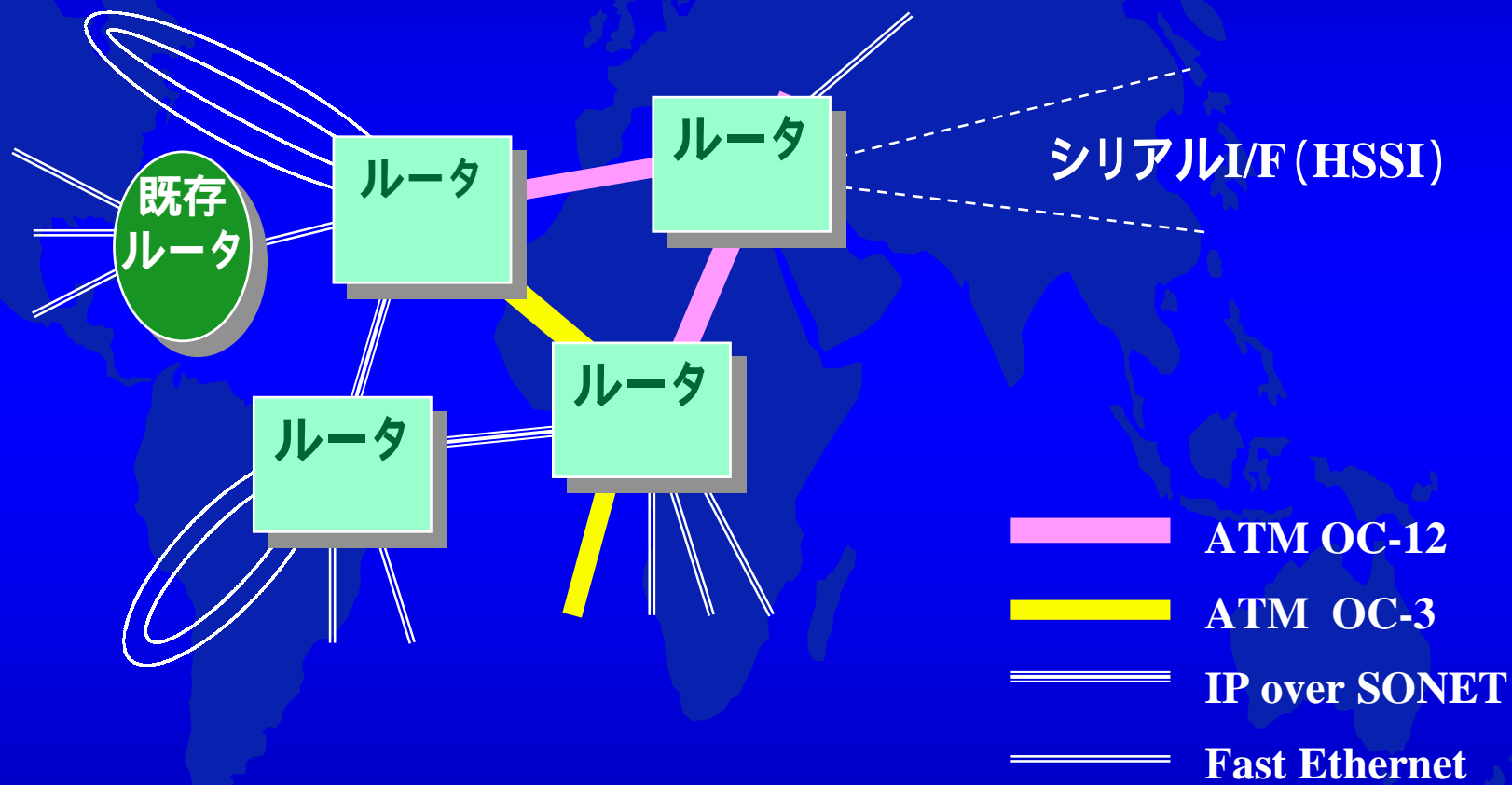
◆ CLPF(Classical Packet Forwarding)モデル

– 適用領域・環境：IP over ATMの導入フェーズ

◆ SCPF(Short-Cut Path Forwarding)モデル

– 適用領域・環境：キャンパスバックボーンレベル
閉じたATM環境(ATM環境が主体)

ATMポイントポイントリンクモデル



ATMポイントポイントリンクモデル

特徴:

1. フレーム(IPパケット)単位での転送処理を行う
2. On-Demandなリンク設定を行わない
3. シングルリンクを複数のパケットフローで共有する
4. ATMリンクは1つのリンクでしかない

ATMポイントポイントリンクモデル の課題

ルータでのフレーム処理の高速化

従来純粋なソフトウェア処理で処理していた
パケット処理の大部分をカスタムチップを用
いて実現

- 1 高速プロセッサ(e.g., RISCチップ)の導入
- 1 パケットの分散処理
 - 1 マルチプロセッサ
 - 1 ルートサーバ+インターフェースプロセッサ
- 1 ハードウェア・ルートテーブル検索(e.g., CAM)

ラベルスイッチルータ

- LSR; Label Switch Router -

ATMスイッチのような高速大容量のレイヤ2ラベル
スイッチモジュールをルータの packets 転送エンジン
として利用する方法

大部分の packets を通常の IP 処理を行わずに、
レイヤ2のラベル情報を用いて転送

ATMスイッチエンジンを用いたラベルスイッチ
ルータではVPI/VCI値を使って packets を転送

* 現在、IETF (Internet Engineering Task Force) の
MPLS (Multi Protocol Label Switching) ワーキング
グループで標準化検討が進行している技術

ラベルスイッチルータ (LSR)

◆ ラベルを用いたパケット転送

⇒ パケットフローとラベルのマッピング

– データリンクヘッダ内情報

✦ e.g, VPI/VCI (ATMインターフェース)

– L3、L2間にタグを挿入

✦ e.g., イーサネットヘッダ、IPヘッダの間にタグ

◆ ラベルスイッチ

– ハードウェアエンジン処理 : e.g., ATMスイッチ

– ソフトウェア処理 : ファーストキャッシュ

ラベルスイッチルータの特長

レイヤ2ラベルスイッチによりソフトウェアによる従来のパケット処理をカットスルー

高スループット

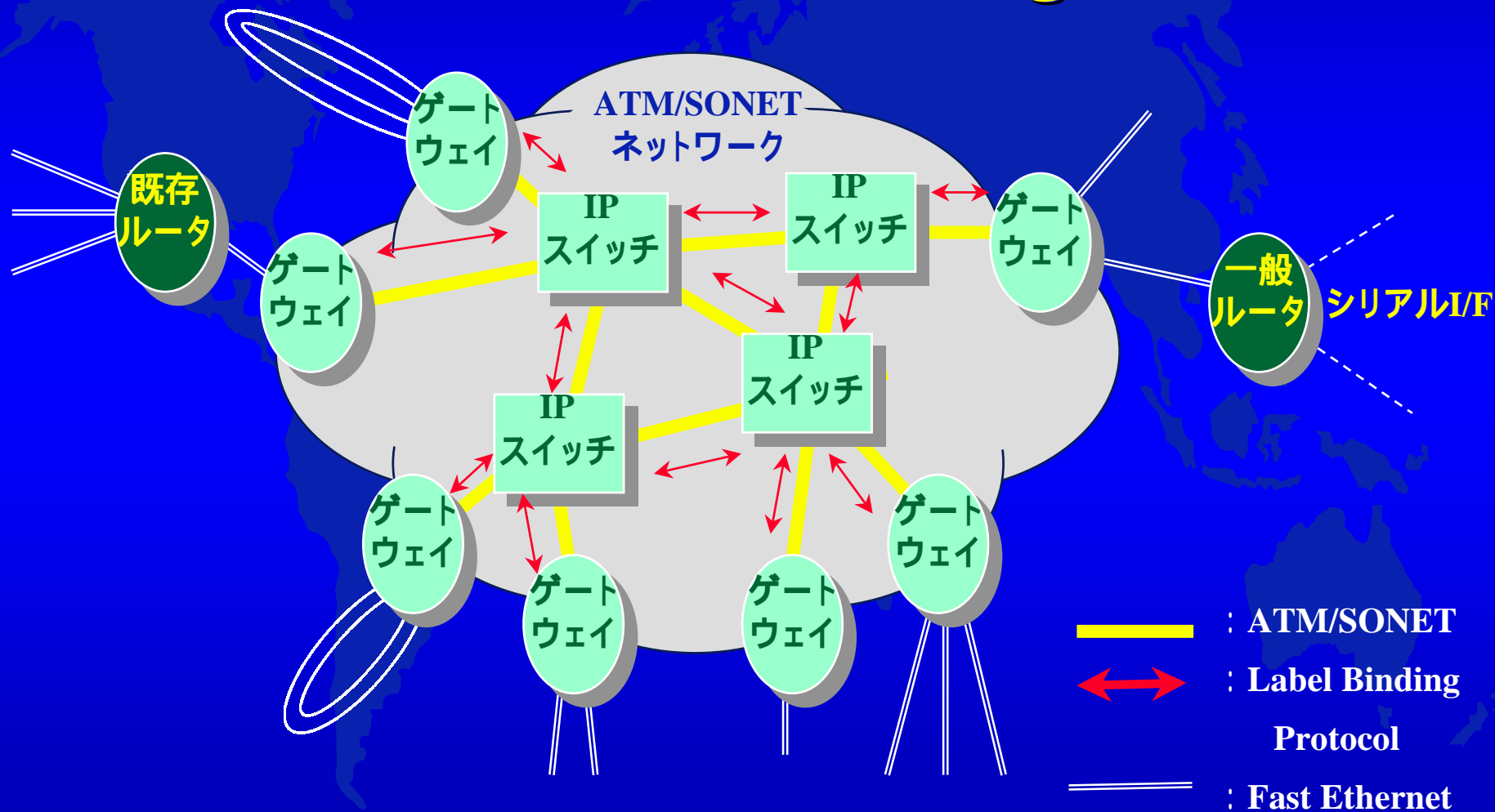
データ転送と制御プロセスの分離

可変長ルーティング情報検索(i.e., **best-match**)
を固定長ラベル検索(i.e., **exact-match**)で実行

ATMスイッチモジュールのスケジューリング機能
を用いた **高速パケットスケジューリング**

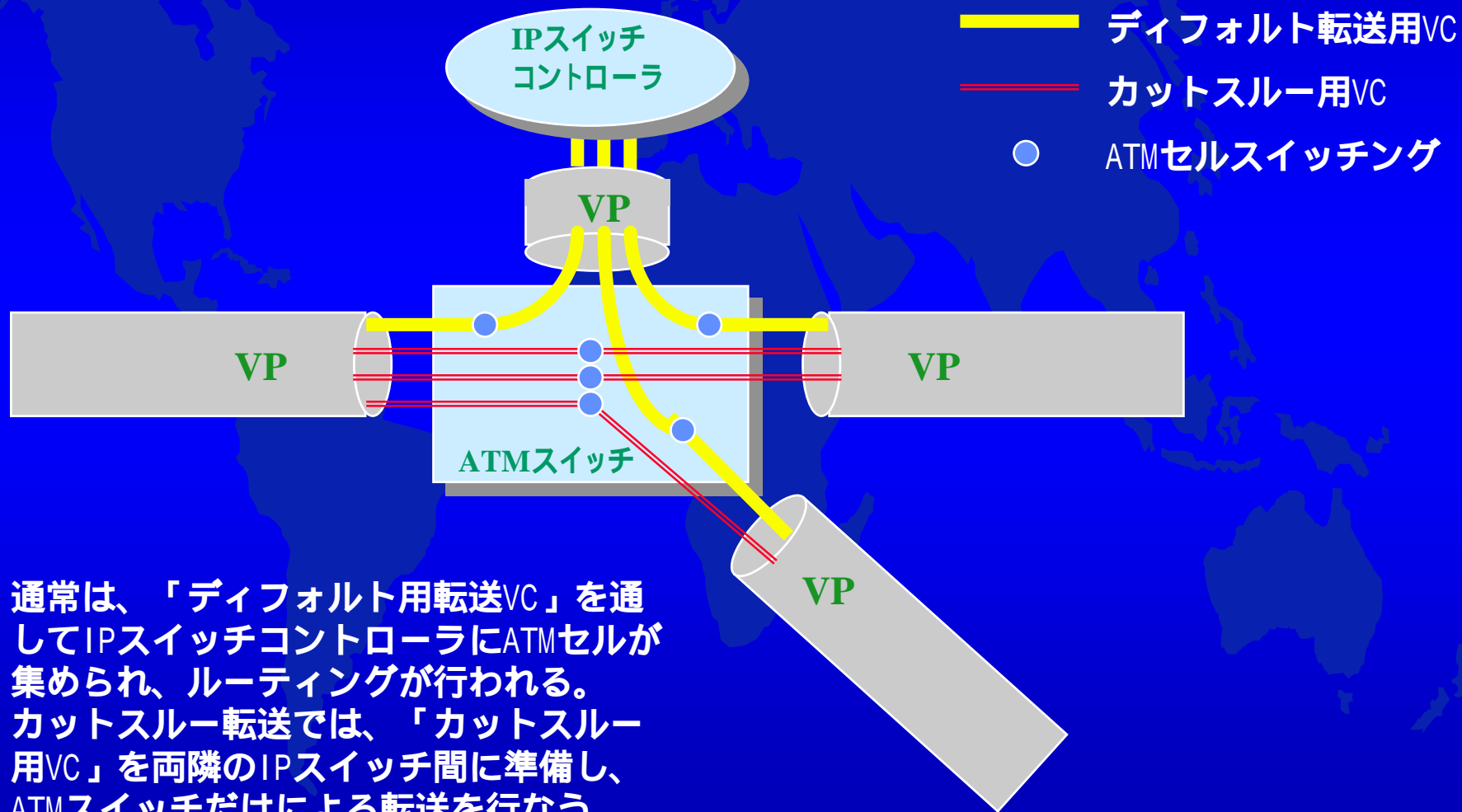
ラベルスイッチ ルータ

- with ATM switch engine -



ラベルスイッチ ルータ

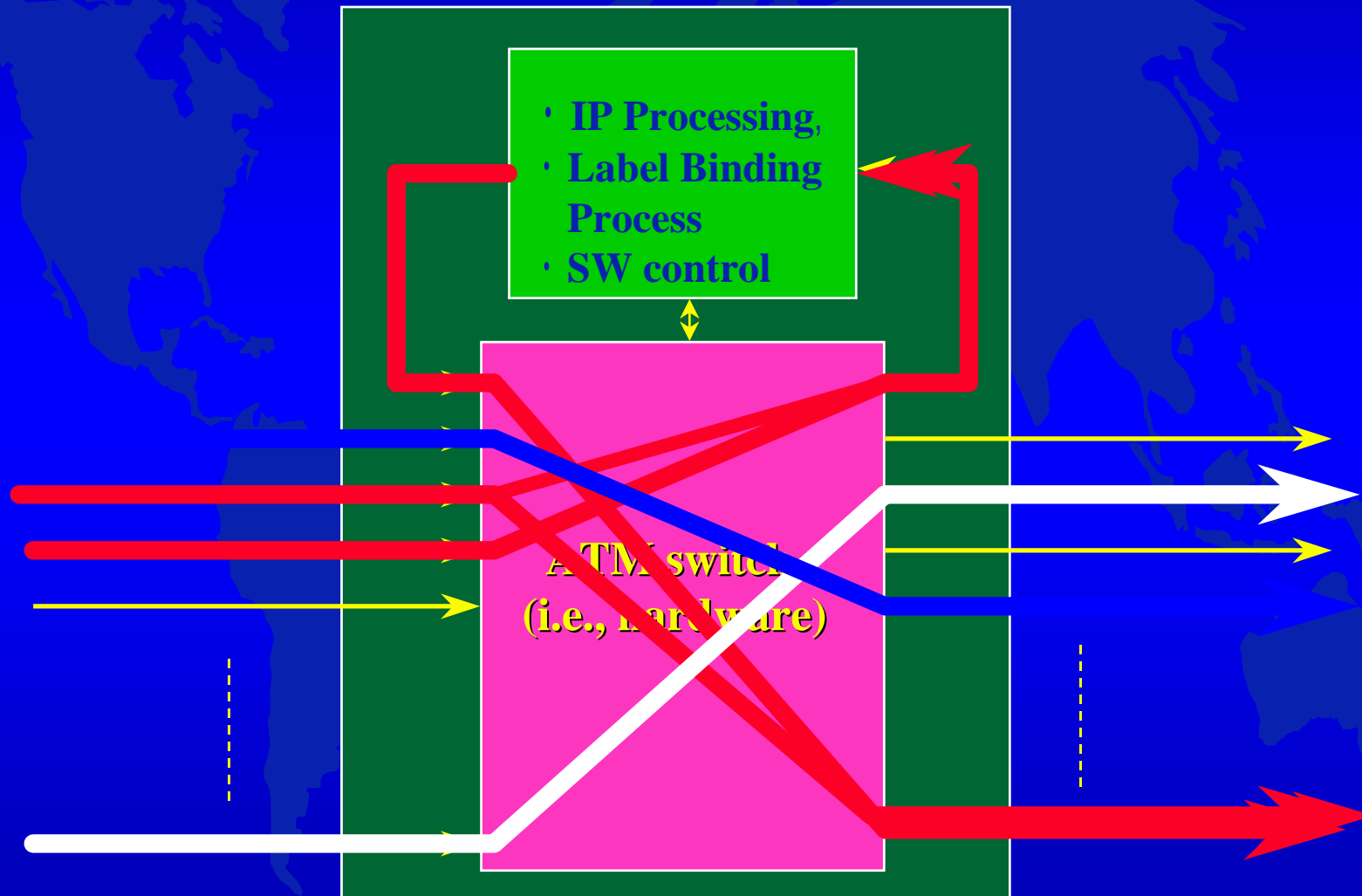
- with ATM switch engine -



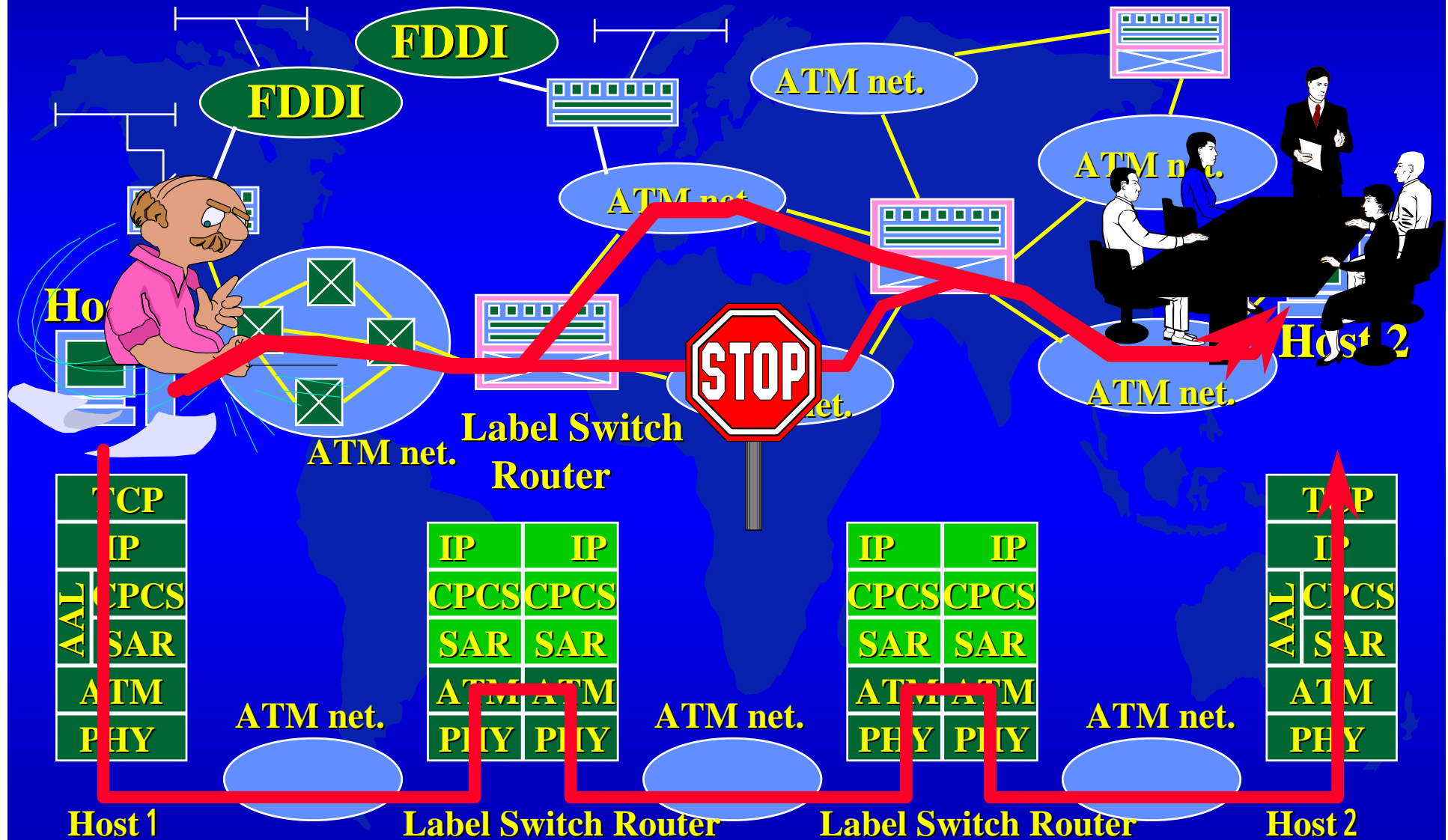
通常は、「デフォルト用転送VC」を通してIPスイッチコントローラにATMセルが集められ、ルーティングが行われる。カットスルー転送では、「カットスルー用VC」を両隣のIPスイッチ間に準備し、ATMスイッチだけによる転送を行なう。

ラベルスイッチルータ実装例

- with ATM switch engine -



LSRを用いたTCP/IPパッケージ転送



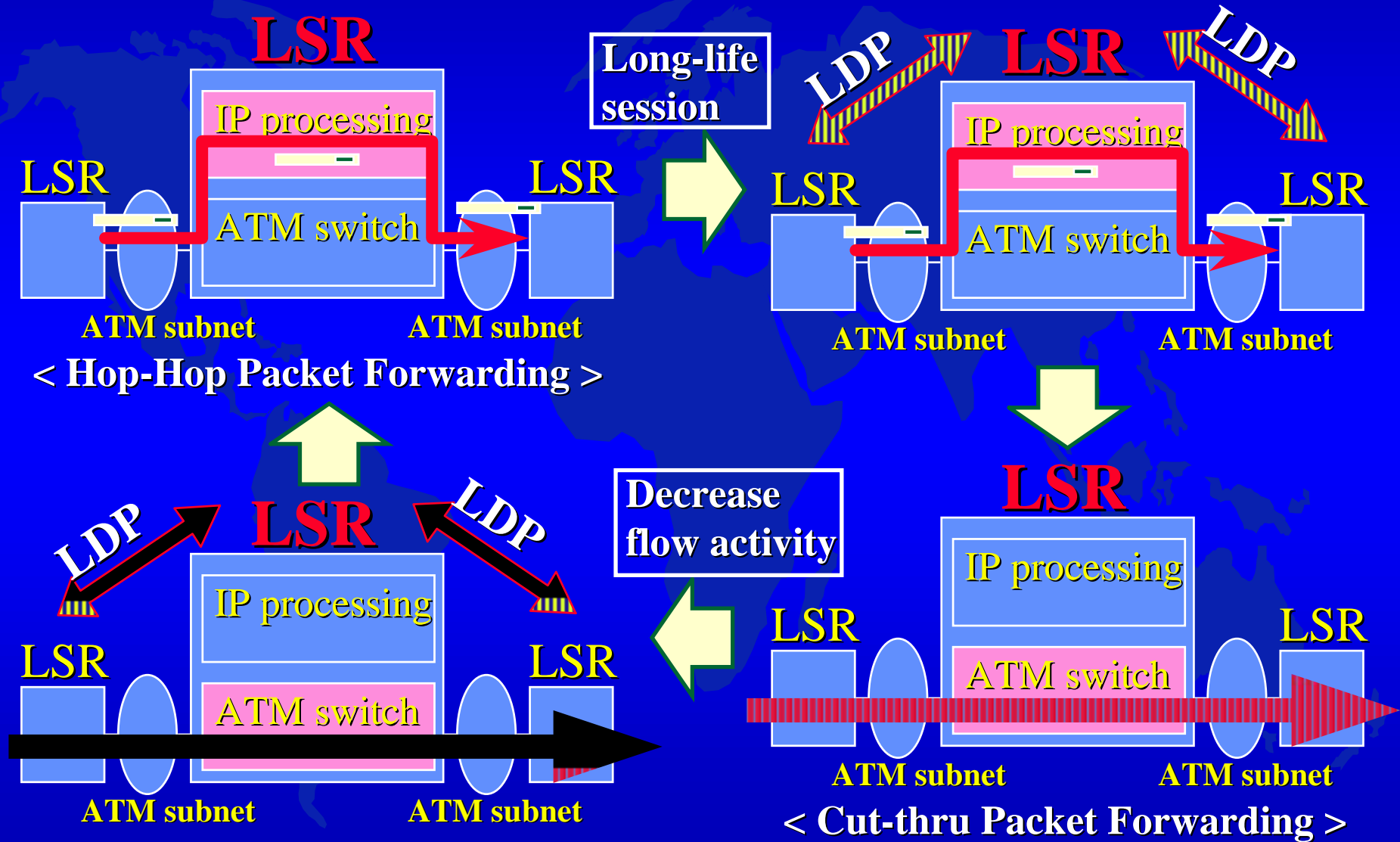
ラベルマッピングポリシー

- ◆ **データドリブン (フロードリブン)**
実際のパケットの到着をトリガーとしてマップ
- ◆ **コントロールドリブン**
制御情報によりマップ
 - **トポロジードリブン**
ルーティングエントリー情報
 - **Reservation ドリブン**
予約プロトコルの制御情報

ラベルマッピングポリシー

- ◆ **データドリブン (フロードリブン)**
実際のパケットの到着をトリガーとしてマップ
- ◆ **コントロールドリブン**
制御情報によりマップ
 - **トポロジードリブン**
ルーティングエントリー情報
 - **Reservation ドリブン**
予約プロトコルの制御情報

フローリブун動作原理



フロードリブンの動作原理

MODE	スイッチ	処理形態	継続時間	アプリ
カットスルー	ATM	Hardware	Long	telnet,web
ホップホップ	Packet	Software	Short	DNS query

+ カットスルーパス設立/解法トリガー

- 設立 ; TCP/UDP Port Number

- 解放 ; Packet Flowのアクティビティ

(*) カットスルーパス制御管理 : LDP(Label Distribution Protocol)

Tag; TDP(Tag Distribution Protocol)

CSR; FANP(Flow Attribute Notification Protocol)

IP Switch; IFMP(Ipsilon's Flow Management Protocol)

Cut-thru Path Life-time (1)

< Gateway Router; Digital Equipment Corp. >

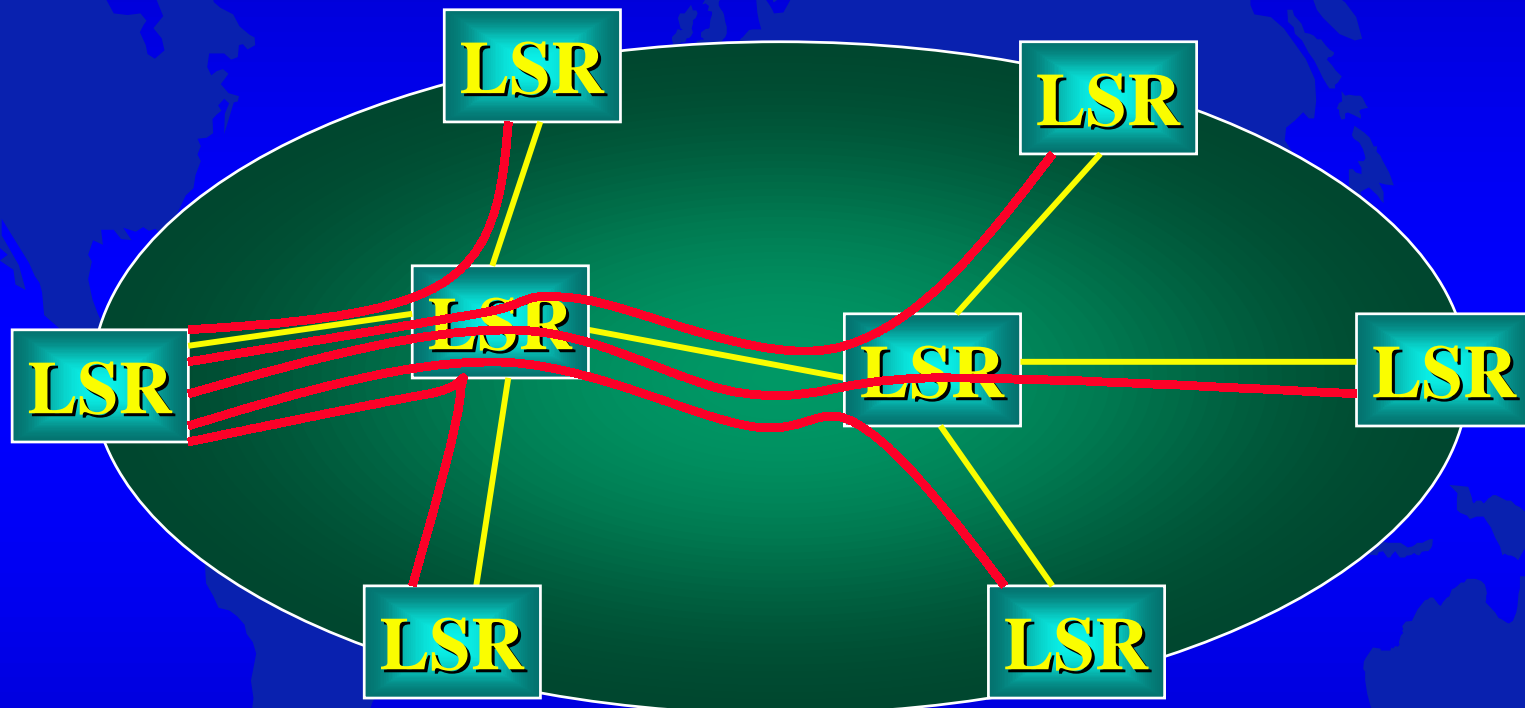
Session	# of sessions		Traffic Volume	Average Life-Time
	Ave.	Max.		
Web	130	161	13.41 %	324.73 sec.
telnet	16	22	0.43 %	1131.90 sec.
ftp	100	144	32.91 %	407.28 sec.
nntp	130	139	39.94 %	1058.33 sec.
Total	276	430	86.69%	481.05 sec.

- ◆ Traffic volume of ftp,telnet,web and nntp generate more than 85% of total traffic. Average life-time for these sessions is more than 8 minutes and these sessions are appropriate to cut-thru forwarding.

ラベルマッピングポリシー

- ◆ **データドリブン (フロードリブン)**
実際のパケットの到着をトリガーとしてマップ
- ◆ **コントロールドリブン**
制御情報によりマップ
 - **トポロジードリブン**
ルーティングエントリー情報
 - **Reservation ドリブン**
予約プロトコルの制御情報

Topology ドリブンの動作



> すべてのExit Routersへのカットスルーパス設定

Flow Driven and Topology Driven - - Advantage and Disadvantage -

	Flow Driven	Topology Driven
Advantage	<ol style="list-style-type: none"> 1. <i>Exclusive QoS</i> allocation 2. <i>Cut-thru across routing domain</i> 3. <i>On-demand cut-thru path establishment</i> 	<ol style="list-style-type: none"> 1. <i>Less VC resource</i> 2. <i>Advanced cut-thru path establishment</i>
Disadvantage	<ol style="list-style-type: none"> a. <i>Large VC resource</i> b. <i>Cut-thru path establish latency</i> c. <i>Some hop-hop forwarding</i> 	<ol style="list-style-type: none"> a. <i>Sharing resource</i> b. <i>IP process at routing domain boundary</i> c. <i>Full-mesh VCs</i> d. <i>Simultaneous cut-thru path establishment</i>

デュアルモード動作の特徴

◆ フロードリブンの寄与

- 必要な帯域幅の削減(by オンデマンド資源割当)
- 資源予約プロトコル(e.g., RSVP)なしでの品質制御
- サブネットにまたがったカットスルーパスの提供
- 排他的な資源割り当て

◆ トポロジードリブンの寄与

- VC数の削減
- 制御オーバーヘッドの削減

ネットワークの構成とコスト

◆ 広域企業ネットワーク

- キャンパスネットワーク
- 広域デジタルリンク

◆ パブリックインターネットプロバイダ(ISP)

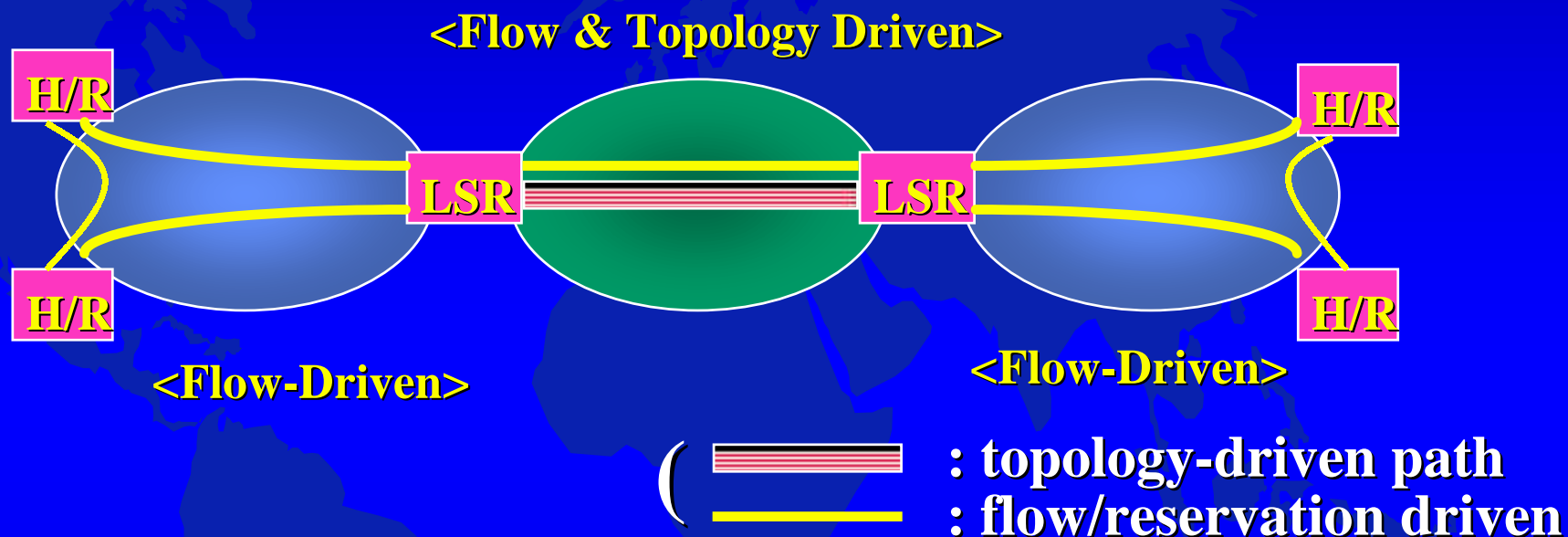
- 広域デジタルリンク

(*) 帯域コスト:

キャンパスネットワーク : not 主要要素

広域デジタルリンク : 主要要素

Dual Mode 適用と動作例



(*) On-demand Cut-thru Path Establishment
(flow/reservation driven path)

=> 必要帯域幅の削減

=> Per-flow QOS 提供

ラベルスイッチルータ

◆アーキテクチャモデル

– **IP Switch** (Ipsilon社) : IETF不参加

✦ RFC1953

– **TAG Switch** (Cisco社)

✦ RFC2105

– **ARIS** (IBM社)

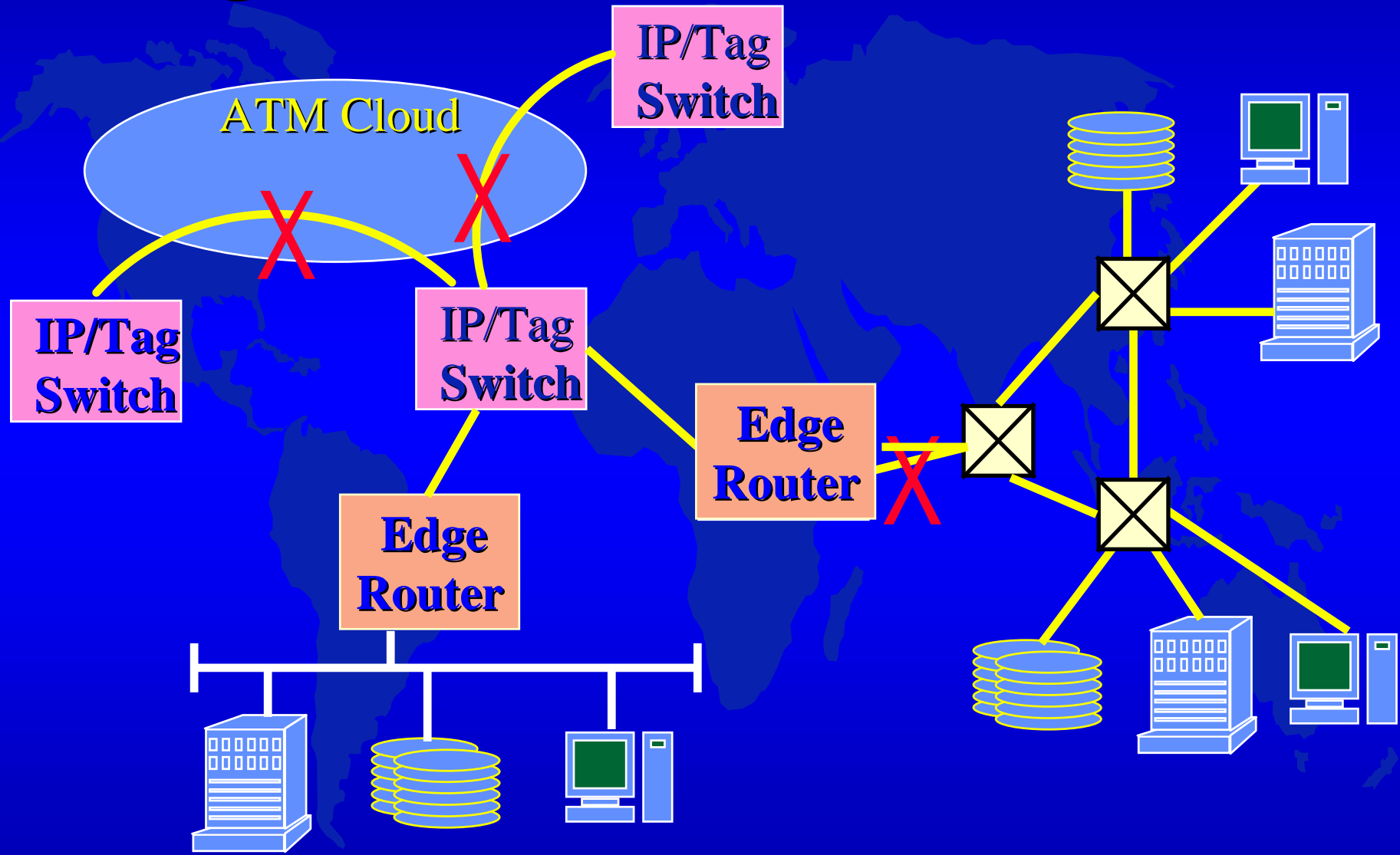
– **CSR (セルスイッチルータ)** (東芝)

✦ RFC2098, RFC2129

Label Switching アーキテクチャの比較

	Data Driven	Control Driven	標準ATMSW 相互接続性
IP Switch	Yes	No	No
TAG Switch	No	Yes	No
ARIS	No	Yes	No
CSR	Yes	Yes	Yes

IP/Tag/ARIS Switchシステムの課題点



IP over ATMモデル機能比較

	ポイントポイント	NBMAリンク		ラベルスイッチ
		CLPF	SCPF	
スループット	A	C	B	A
ルーティンググループ	no	no	Yes	no
CLリンクへの適用	Yes	no	no	Yes
システム複雑度	A	B	C	A
スケーラビリティ	B	D	C	A
システムコスト	D	B	C	A
開発状況	soon	Done	yet	Done
運用状況	soon	Yes	yet	Yes
適用領域	All	LAN	LAN MAN?	All

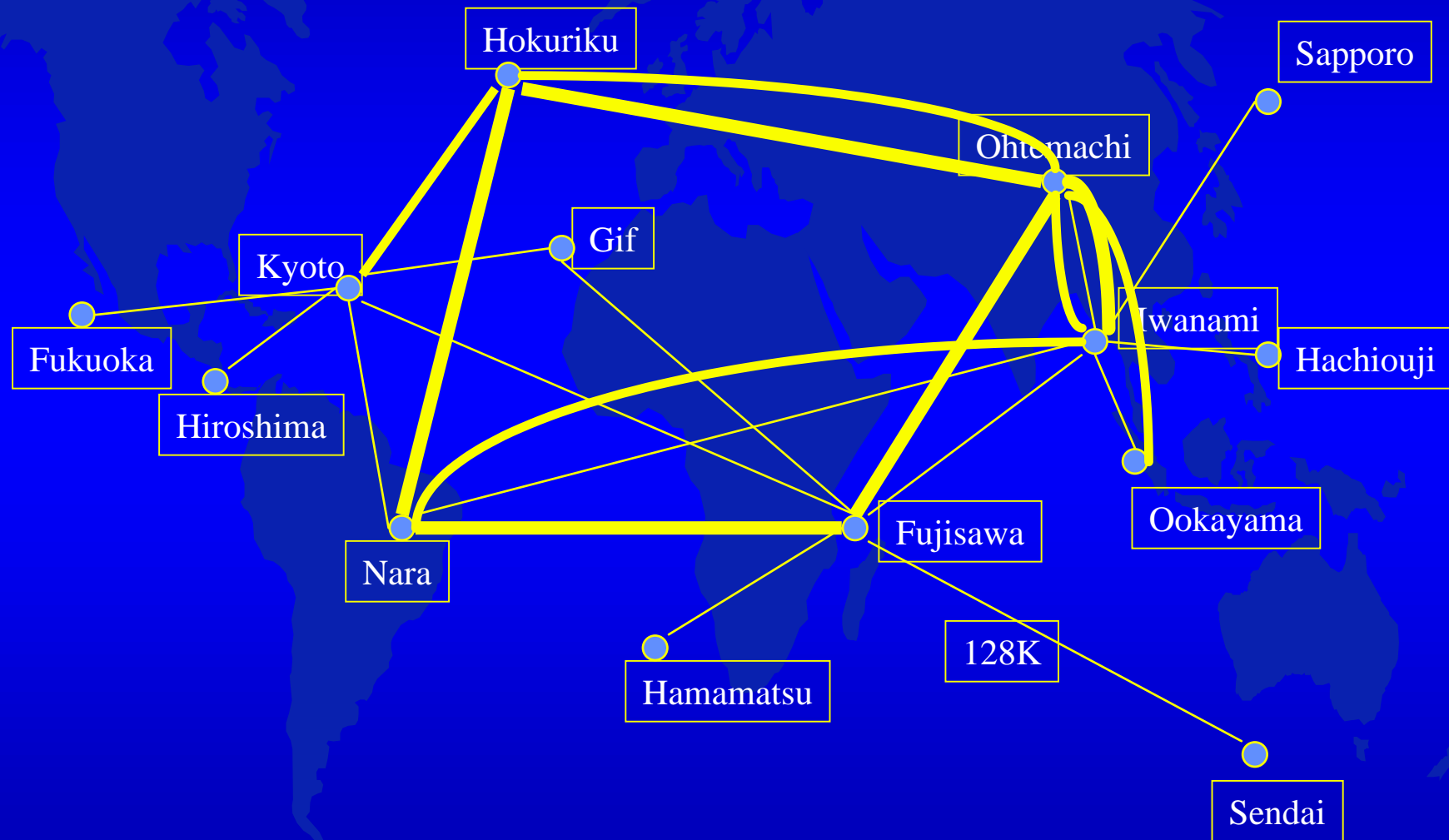
- CLPF: LANE, RFC1577
- SCPF: NHRP, MPOA

広域実験ネットワークテストベッド - WIDEプロジェクト、慶応大、NTT -

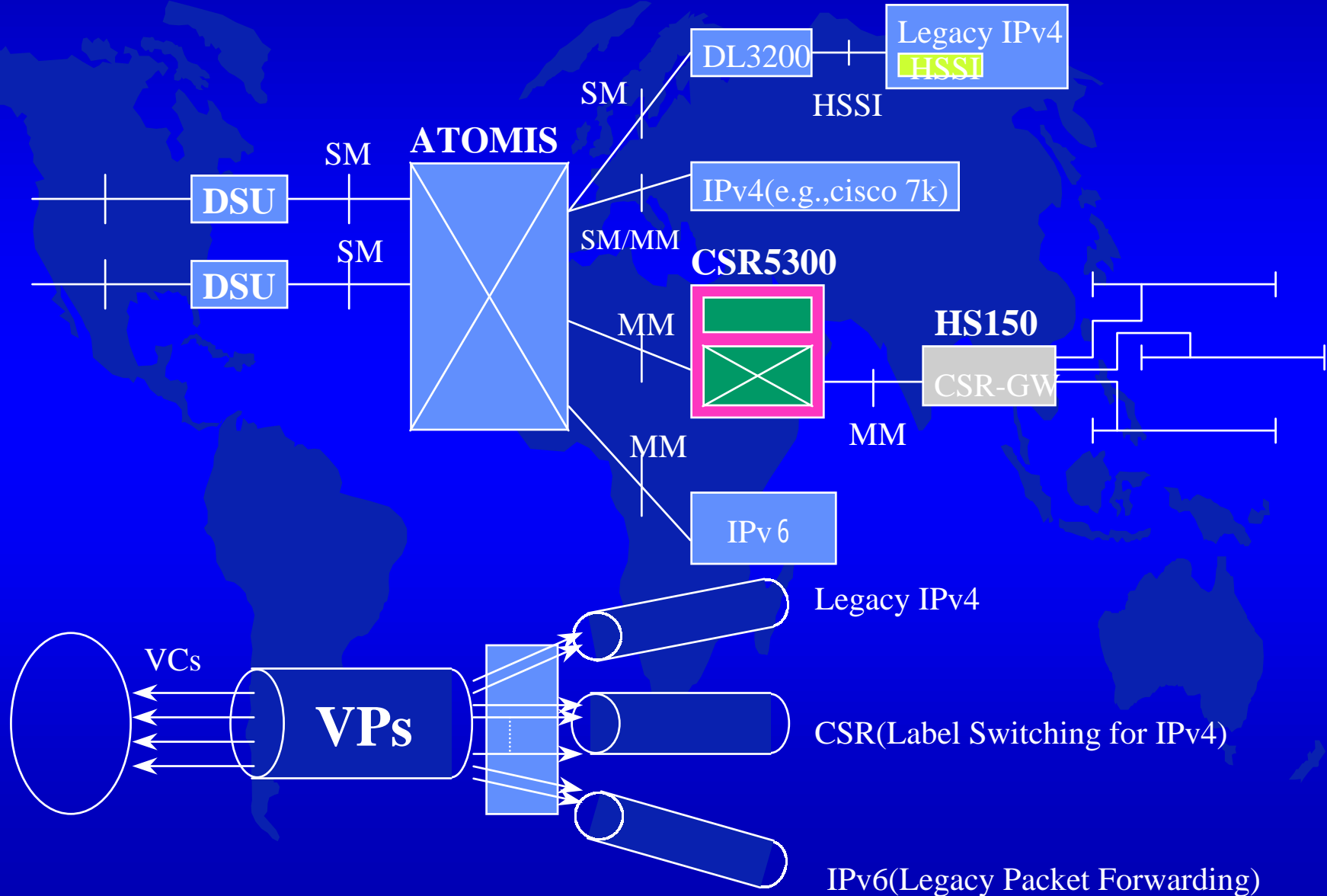
ATM/SDHリンク上のオーバーレイ網

- [1] 既存ルータベースの網
- [2] ラベルスイッチルータ(LSR)ベースの網
(e.g., CSR(セルスイッチルータ))
- [3] IPv6網
- [4] その他の仮想実験ネットワーク
(e.g., RSVP, WEB caching)

広域実験ネットワークテストベッド - WIDEプロジェクト、慶応大学、NTT -



コアNOCコンフィギュレーションの例



まとめ

- ATM技術を使ったIPサービス -

ATMリンクを導入する利点:

1. 仮想ネットワーク(Virtual Network)の
容易な構築
2. リンク帯域幅の柔軟性
3. コストパフォーマンス

ATMリンク導入の注意点!

- (1) セル転送速度のシェイピングが必須!
(注) シェイピングを行ってもセルが落ちることがある!
- (2) セル流の合流には注意!