

M1 RES - Architecture des réseaux 8/10

Architecture support - ATM et MPLS

Olivier Fourmaux

`olivier.fourmaux@lip6.fr`

Version 4.c, septembre 2004

Plan

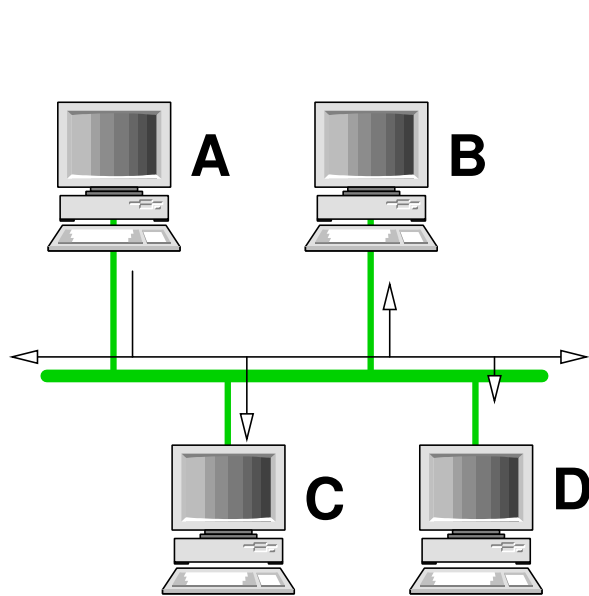
Architecture Ethernet

Architecture ATM/MPLS

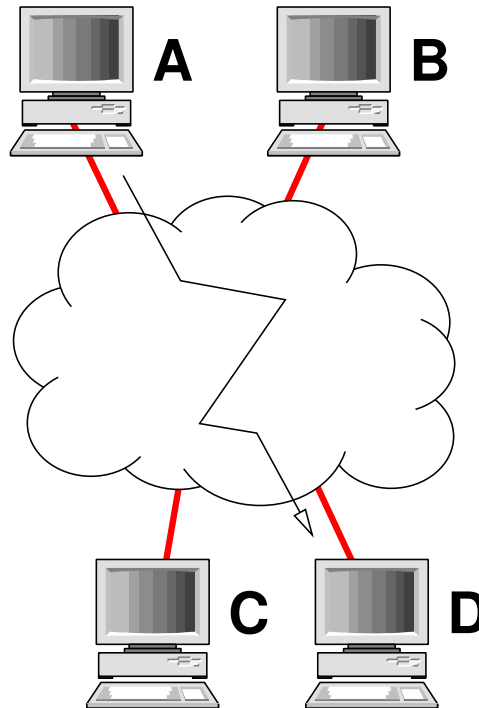
- **NBMA**
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
- MPLS

Architecture point-à-point

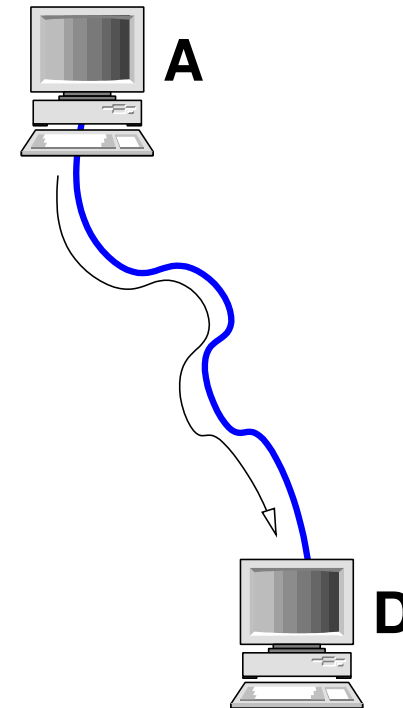
ATM et MPLS sont des réseaux de type NBMA



Réseau à diffusion



Réseau NBMA



Réseau point-à-point

Accès multiples sans diffusion (NBMA)

Certaines infrastructures de réseau permettent la création dynamique de connexion point-à-point entre des machines connectées au même réseau (niveau de la couche liaison) : *Non Broadcast Multiple Access*

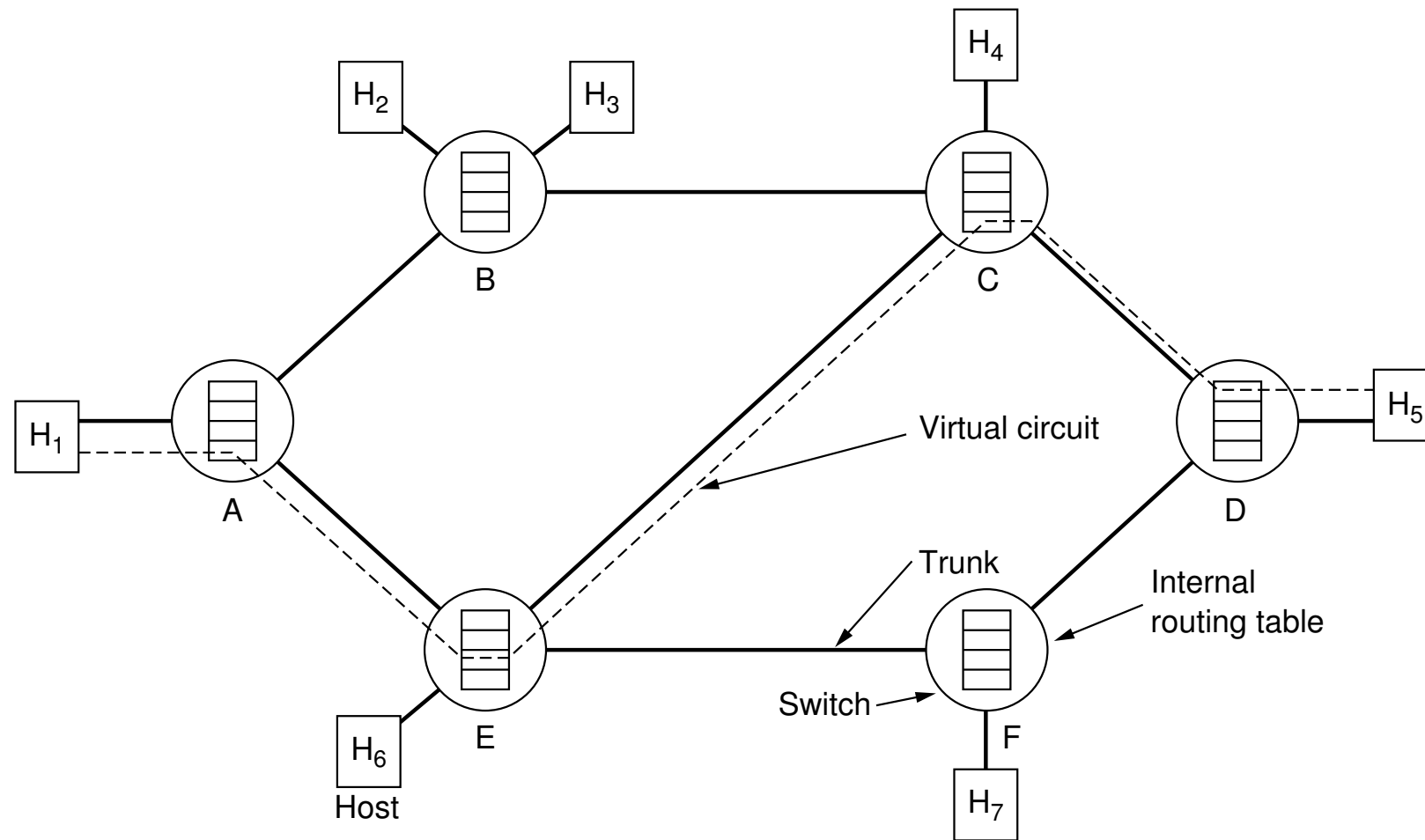
Quelques technologies NBMA

- (Téléphonie, ISDN, LS) création de la connexion externe
- X25
- FR (*Frame relay*)
- ATM (*Asynchronous Transfer Mode*)

Transport d'IP sur NBMA :

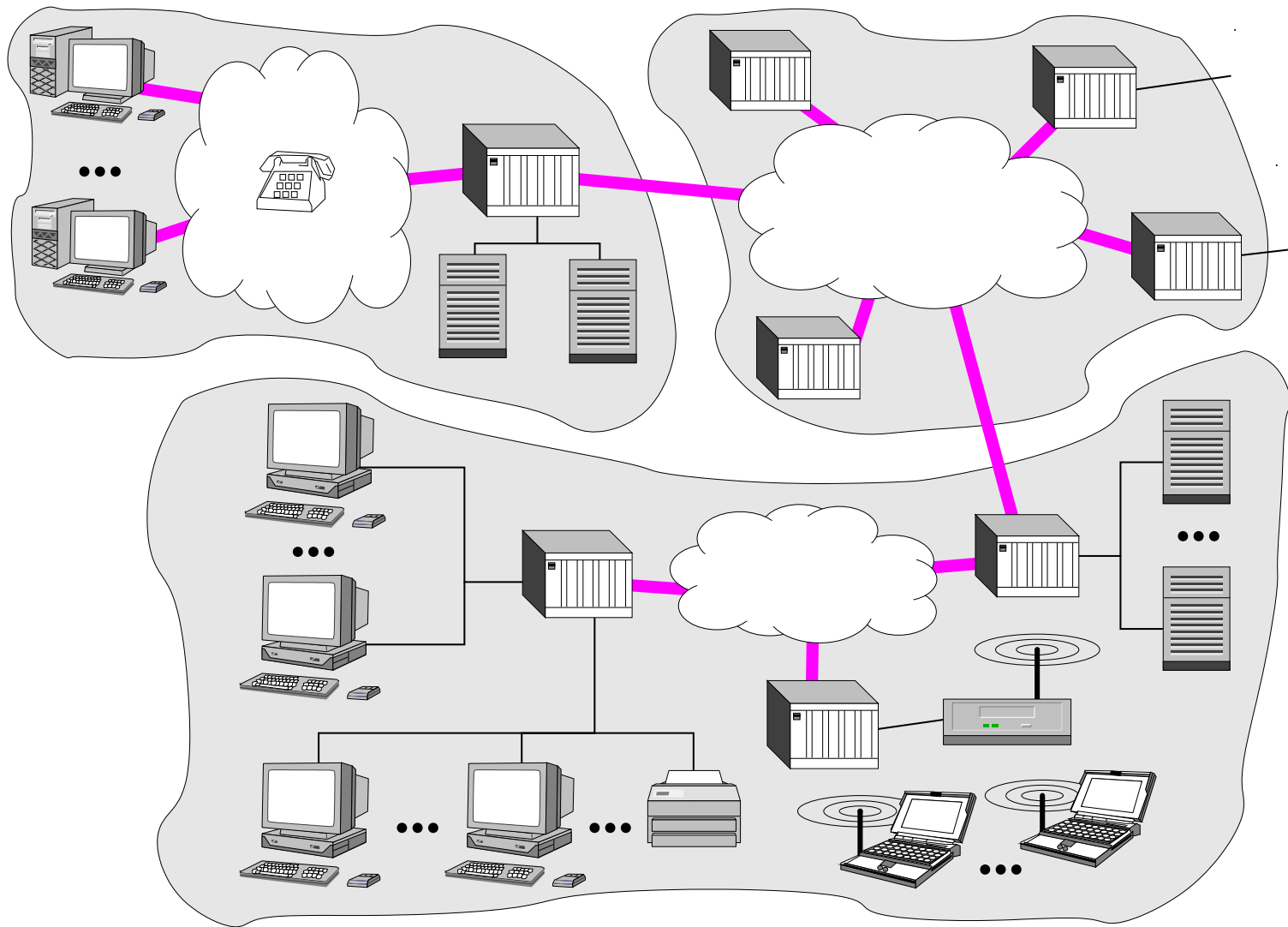
- **résolution d'adresses**
 - ✓ création implicite de la connexion
- **format d'encapsulation**

NBMA : Circuit Virtuel



picture from TANENBAUM A. S. *Computer Networks 3rd edition*

NBMA : Où ?



Plan

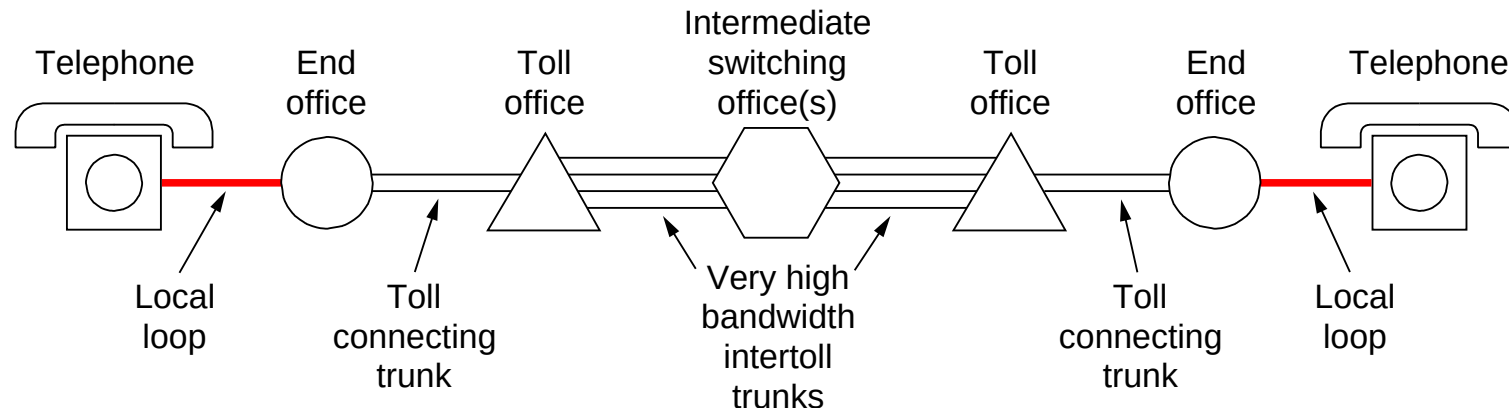
Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- **B-ISDN**
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
- MPLS

Architecture point-à-point

Système téléphonique



Réseau Téléphonique Commuté (RTC)

Plain Old Telephon System (POTS)

- réseau de transit numérique
- boucle locale analogique
- transmission de données informatiques :
 - ✓ modulation – codage — décodage – démodulation

ISDN

ISDN : Integrated Services Digital Network

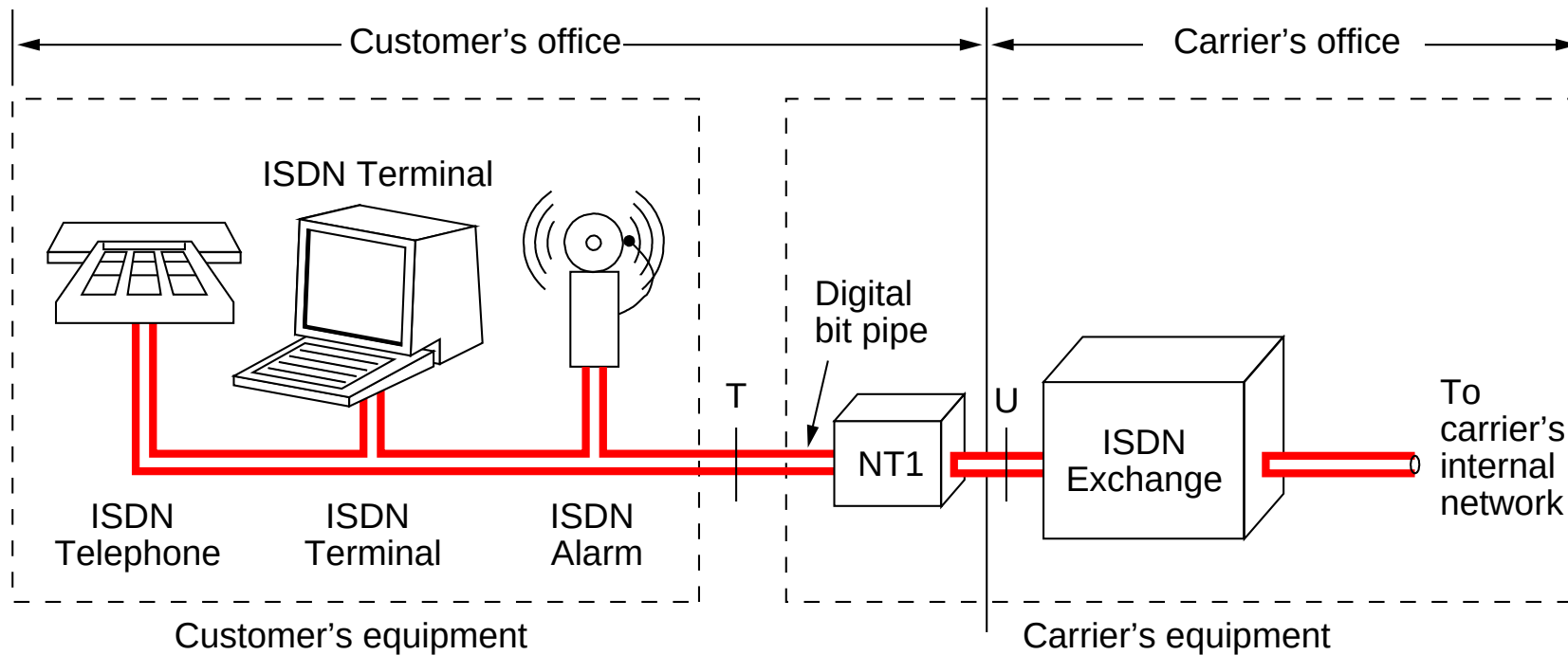
RNIS : Réseau Numérique à Intégration de Services

- **boucle locale numérique**

- ✓ en étude dès 1969...
- ✓ ... normalisation par le CCITT (future ITU) en 1984
- ✓ transport de la voix + (qqs) données :
 - ☞ associées à une extension des fonctionnalités
nom de l'appelant, gestion PABX...

▢▢▢ *N-ISDN (Narrow Band)*, RNIS-BE (Bande Étroite)

ISDN : Configuration résidentielle

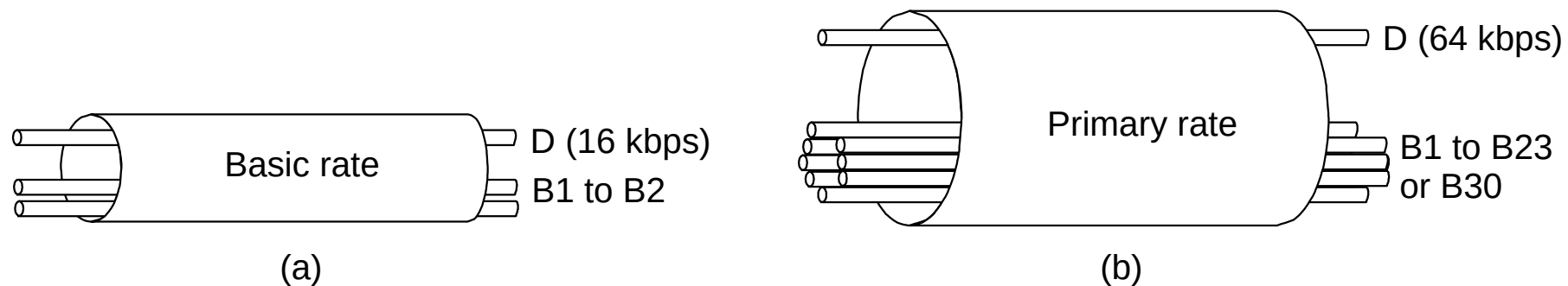


picture from TANENBAUM A. S. *Computer Networks 3rd edition*

ISDN : Interfaces

Plusieurs canaux ont été définis :

- A = 4Khz analogique pour la voix
- B = 64Kbps MIC pour la voix ou les données
- C = 8 ou 16Kbps numérique
- D = 16 ou 64Kbps numérique pour la signalisation hors-bande
- ...



Deux combinaisons normalisées :

- **Accès de base** = $2B + 1D$
- **Accès primaire** = $30B + 1D$ ($23B + 1D$ au US)

N-ISDN : Problèmes

Pourquoi ça n'a pas marché ?

- tentative de remplacer le téléphone analogique dans les années 80
 - ✓ standardisation trop longue...
 - ☞ offre des opérateurs vers 1995 !
 - ✓ commercialisation ratée :
 - ☞ abonnement plus cher
 - ☞ besoin de changer les équipements
- actuellement, solution intermédiaire pour un accès données plus rapide
 - ✓ agréger pour avoir 128kbits
 - ✓ deux lignes (data + voix, mini PABX...)
- les entreprises attendaient des hauts débits
 - ✓ 64kbps pour relier des LAN ???

➡ **B-ISDN (RNIS-LB)**

B-ISDN : Evolution

RNIS Large-Bande (*Broadband ISDN* ou *B-ISDN*)

- haut débit (supérieur à 150Mbps)
- infrastructure spécifique
- intégration étendue :
 - ✓ flux temps réel
 - ☞ débit constant : **audio** et **vidéo**
 - ☞ débit variable : audio et vidéo compressés
 - ✓ flux sans contraintes temporelles
 - ☞ trafic *bursty* et adaptatif : **data**
- nouvelle approche : asynchrone
 - ✓ **technologie ATM** (*Asynchronous Transfer Mode*)
 - ☞ petits paquets et circuits virtuels

Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- **Technique ATM**
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
- MPLS

Architecture point-à-point

Historique de l'ATM

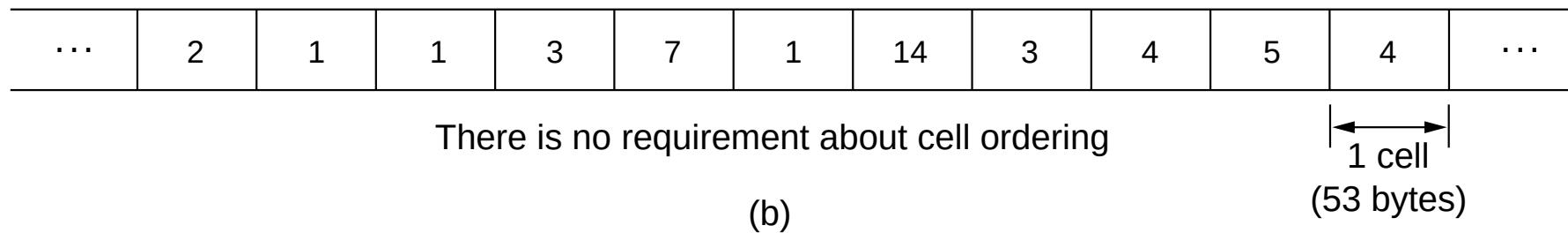
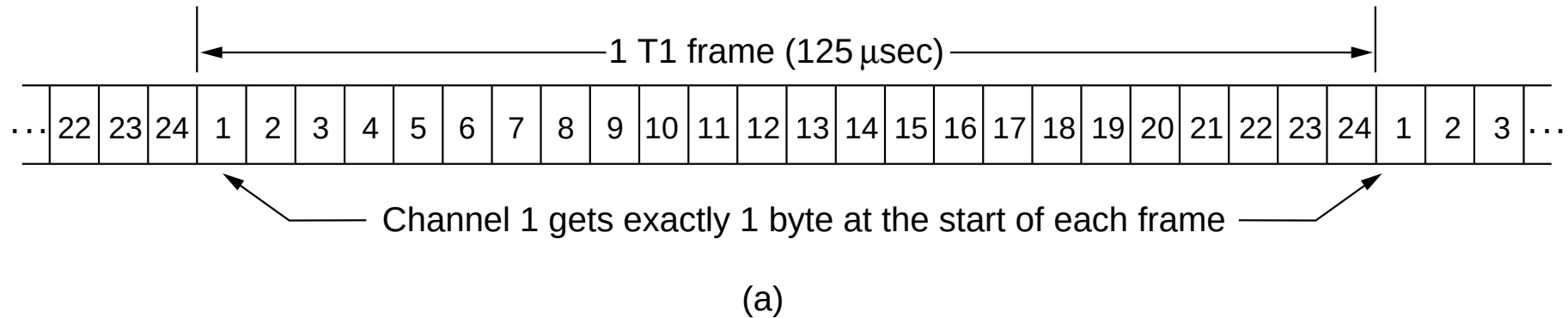
- **1983** : commutation rapide de paquets :
 - ✓ *Asynchronous Time Division* (CNET, actuellement FT R&D)
 - ✓ *Fast Packet Switching* (Bell Lab, ATT puis Lucent)
- **1986** : **début de normalisation** au CCITT (actuellement ITU)
- **1988** : choix de l'ATM pour le RNIS-LB
- **1990** : premières recommandations (normes de l'ITU)
- **1991** : création de l'ATM Forum
- **1992** : mise-à-jour des recommandations, maturité de l'ATM
- **1993** : premières spécifications de l'ATM Forum (UNI 3.0)
- **1994** : premiers RFC de l'IETF (IP sur ATM)
- **1995** : **premiers réseaux opérationnels**
- **1995** : ATM dans les réseaux informatiques (LAN Emulation)
- **1996** : PNNI 1.0 et UNI 4.0
- **1997** : MPOA 1.0

Principe de base d'ATM

Mode de transfert asynchrone

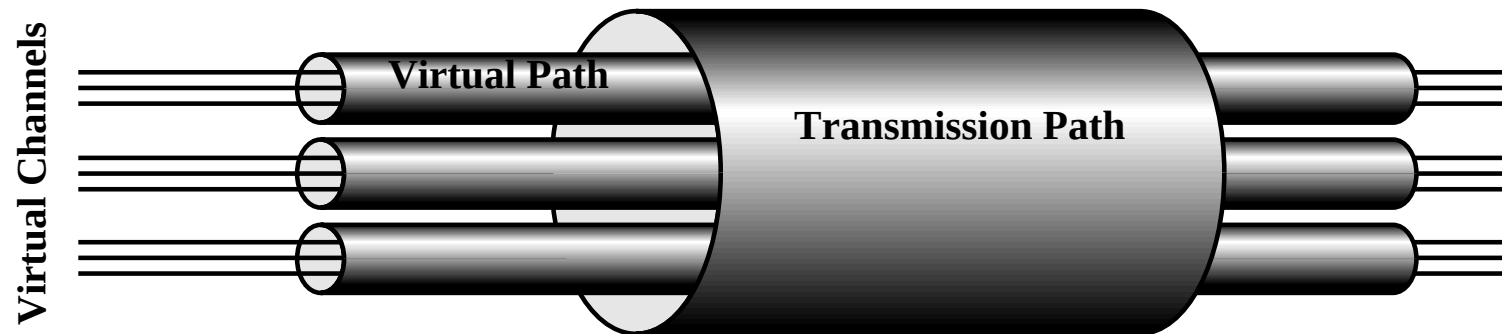
- pas d'horloge globale
 - ✓ circuits virtuels
 - ✓ multiplexage par étiquettes ($\sim$ datagrammes)
- **cellule**
 - ✓ taille fixe réduite à **53 octets**
 - ✓ contient une étiquette
 - ➡ + latence réduite (inférieure à 56ms pour les échos)
 - ➡ + multiplexage performant (limite la formation d'agrégats)
 - ➡ + commutation performante (simplification de la commutation)
 - ➡ – mauvaise efficacité de transmission des données

Techniques de transfert



- (a) **synchrone** avec allocation statique
- (b) **asynchrone** avec allocation dynamique et étiquetage

Connexions Virtuelles ATM



Terminologie des connexions virtuelles :

- **VC** (*Virtual Channel* ou **voie virtuelle**) indique le transport unidirectionnel de cellules ATM avec le même identifiant **VCI**
 - ✓ **VCC** (*Virtual Channel Connection*)
- **VP** (*Virtual Path* ou **conduit virtuel**) indique le transport unidirectionnel de cellules ATM avec le même identifiant **VPI**
 - ✓ **VPC** (*Virtual Channel Connection*)

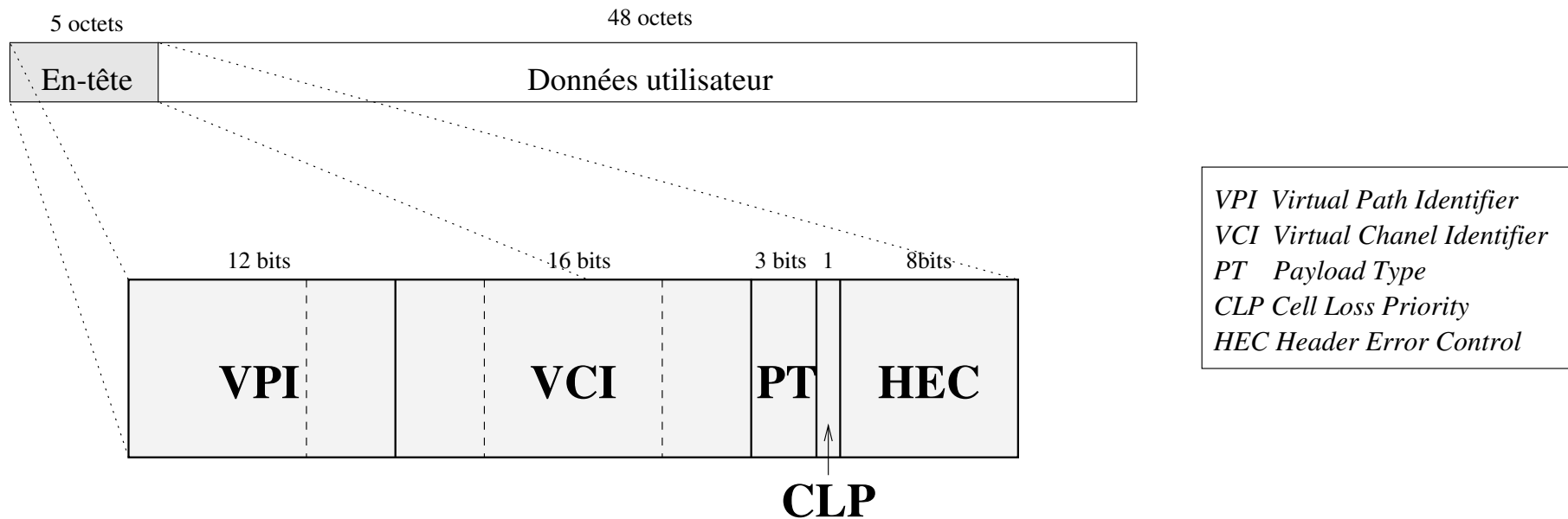
Caractéristiques des VC ATM

Proposition initiale (Recommandation ITU I.150)

- **Qualité de service** (*QoS – Quality of Service*)
- **Négociation** (*Traffic parameter negotiation*) mise en place d'un contrat de trafic entre l'utilisateur et le réseau
- **Supervision** (*Usage monitoring*) vérification par le réseau (*policing*) du respect de contrat de trafic (*shaping*)
- **Statique ou dynamique** (*Semipermanent VCC, Switched VCC*) configuration manuelle ou à la demande
- **Séquencement** (*Cell sequence integrity*) ordre des cellules préservé dans un VCC

La signalisation de contrôle utilise des VCC dédiés

Cellule ATM



53 octets (**5 entête** + 48 données)

VPI (*Virtual Path Identifier*) identificateur de conduits virtuels (12 bits)

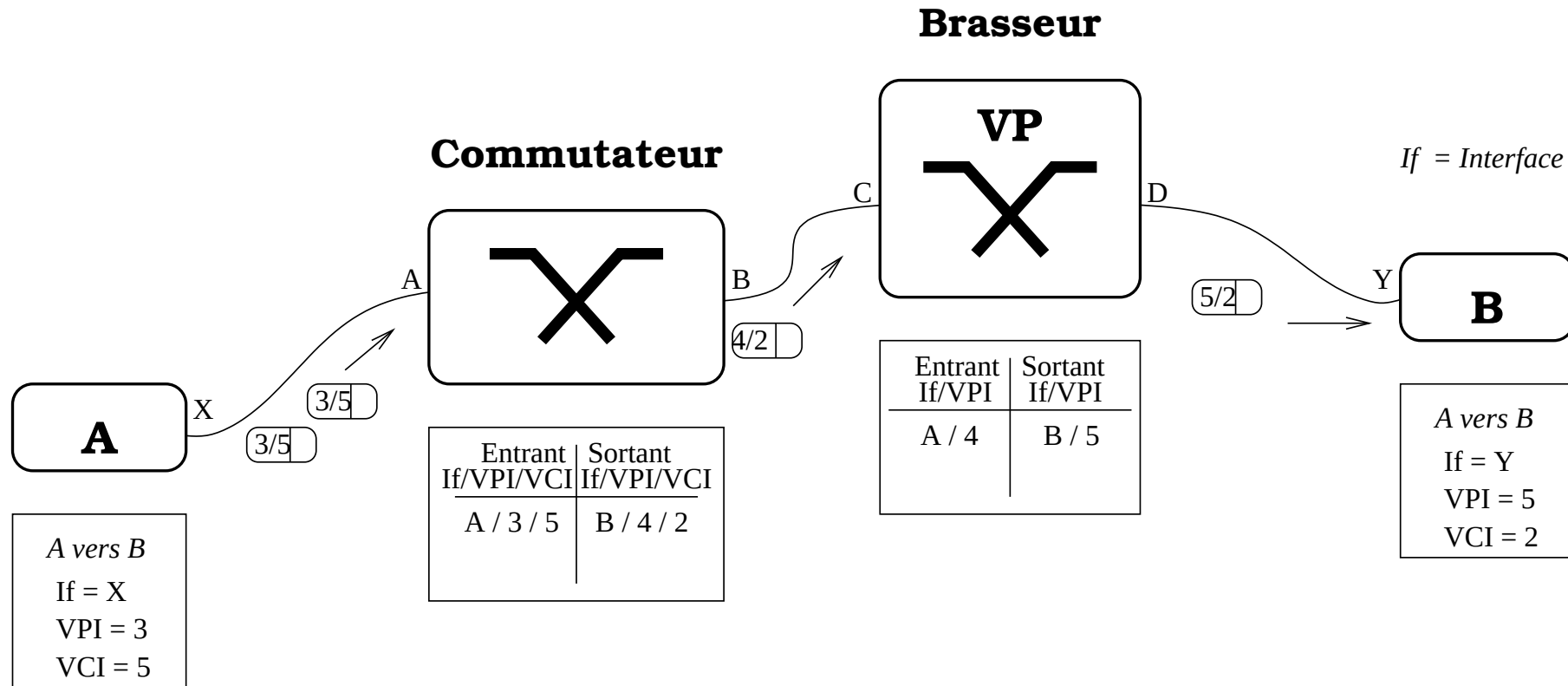
VCI (*Virtual Channel Identifier*) identificateur de voies virtuelles (16 bits)

PT (*Payload Type*) type de données transportées (3 bits)

CLP (*Cell Loss Priority*) perte si congestion (1 bit)

HEC (*Header Error Control*) protection de l'en-tête (8 bits)

Commutation ATM



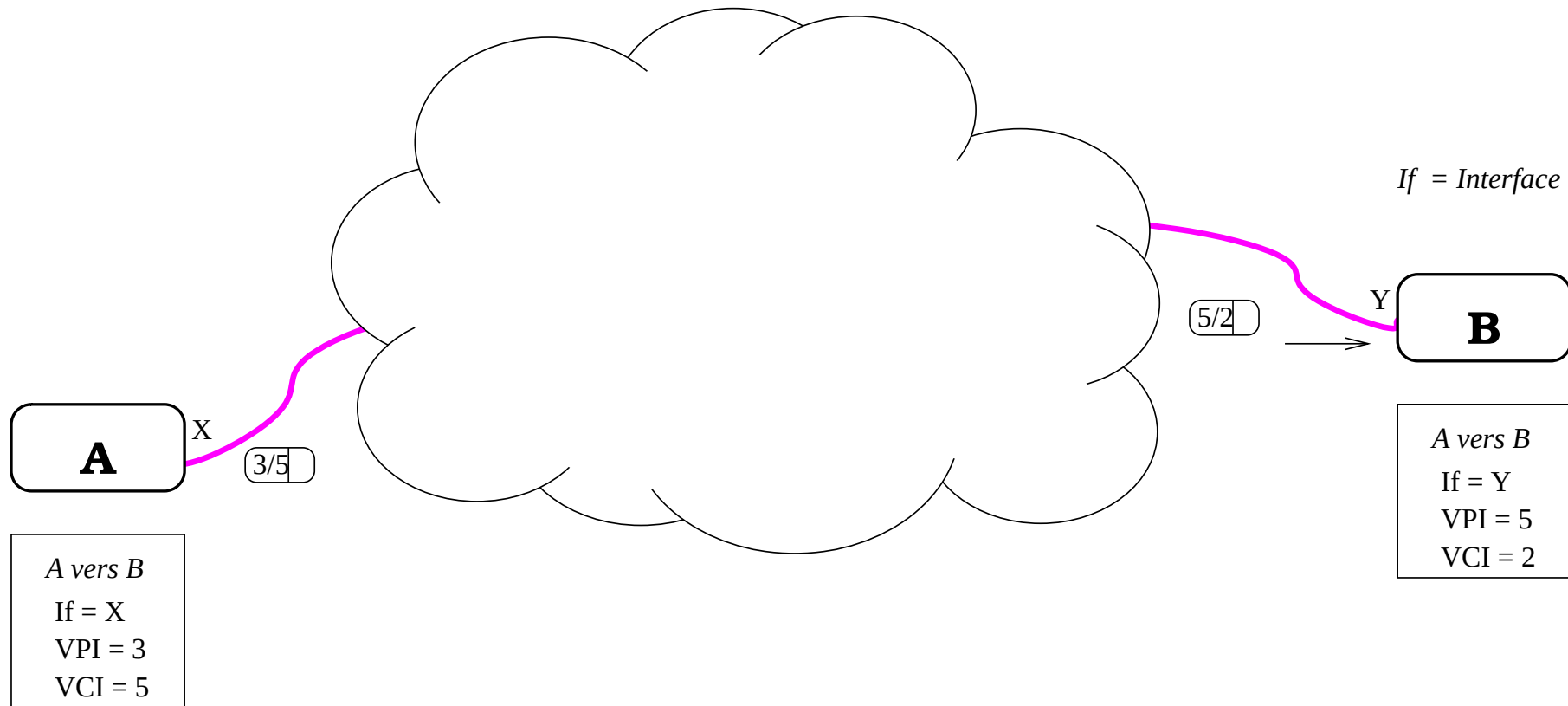
Approche télécom :

- technologie **orientée connexion**
- deux types d'éléments de commutation d'un réseau ATM :

Commutateur ATM (port/VPI/VCI) : commute des **voies** virtuelles

Brasseur ATM (port/VPI) : commute des **conduits** virtuels

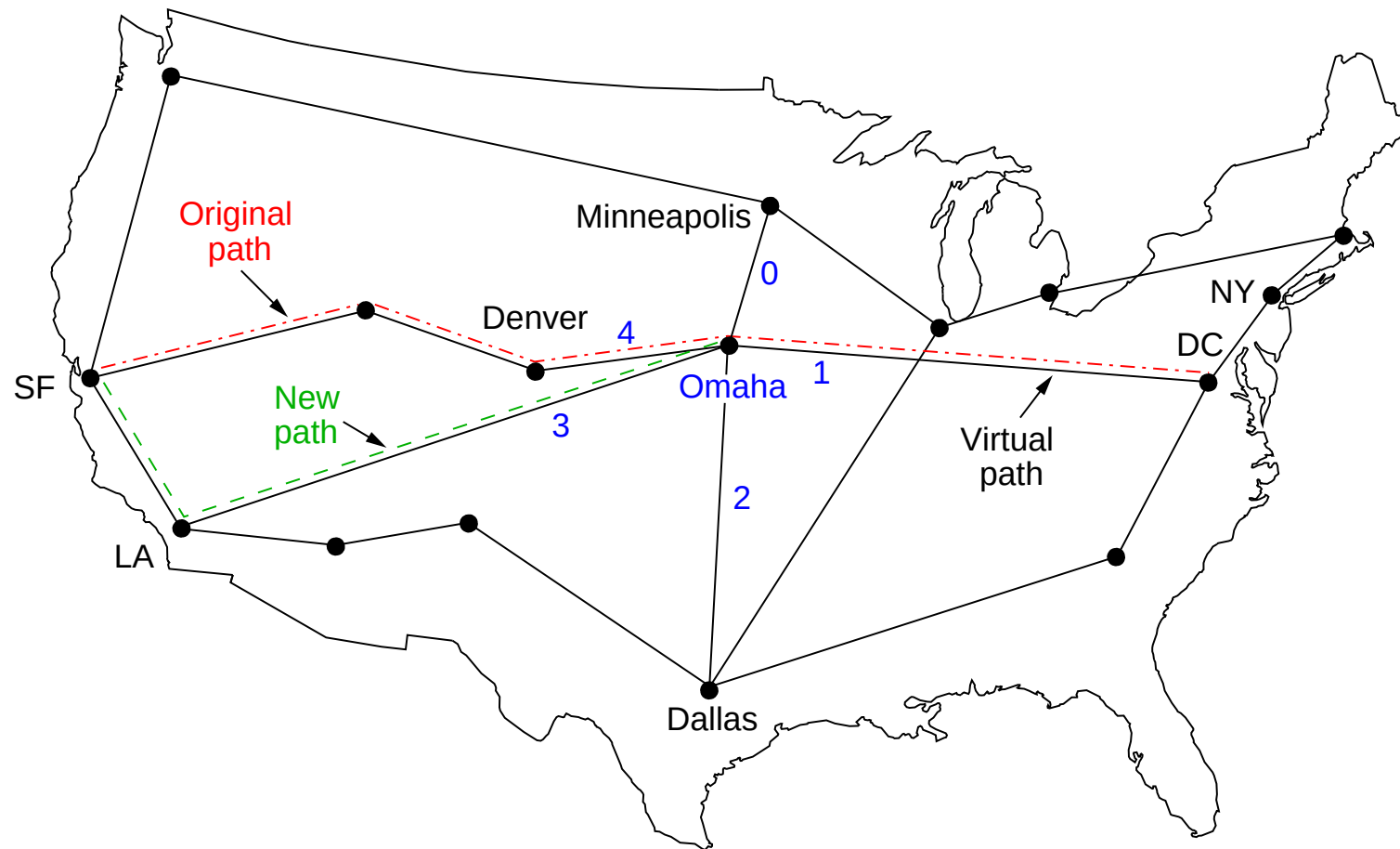
Nuage ATM



Abstraction de la couche de commutation : NBMA

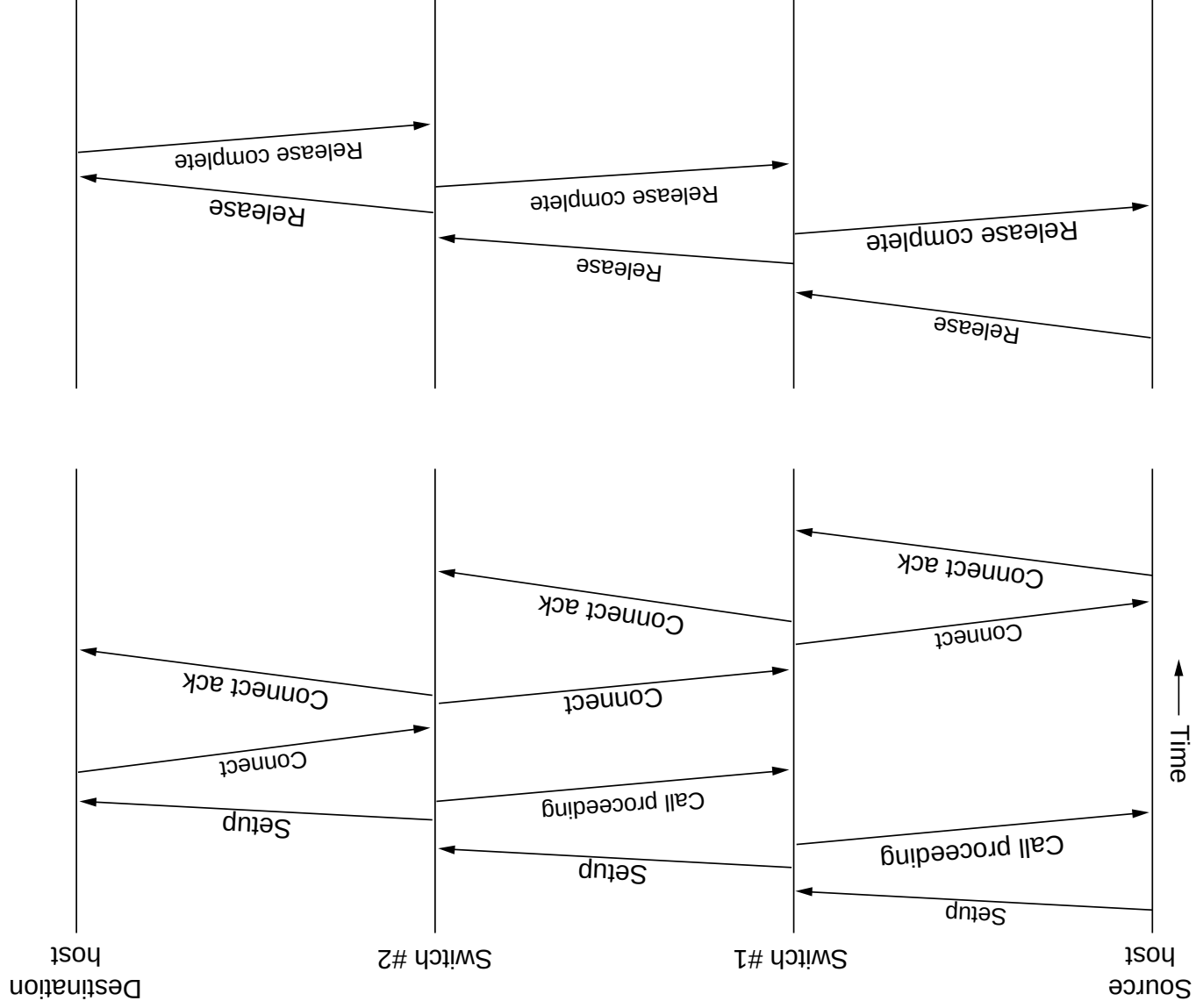
Brassage ATM

Exemple de **reroutage d'un VP** (et de tous ses VC)



picture from TANENBAUM A. S. *Computer Networks 3rd edition*

Gestion de connexion ATM



Plan

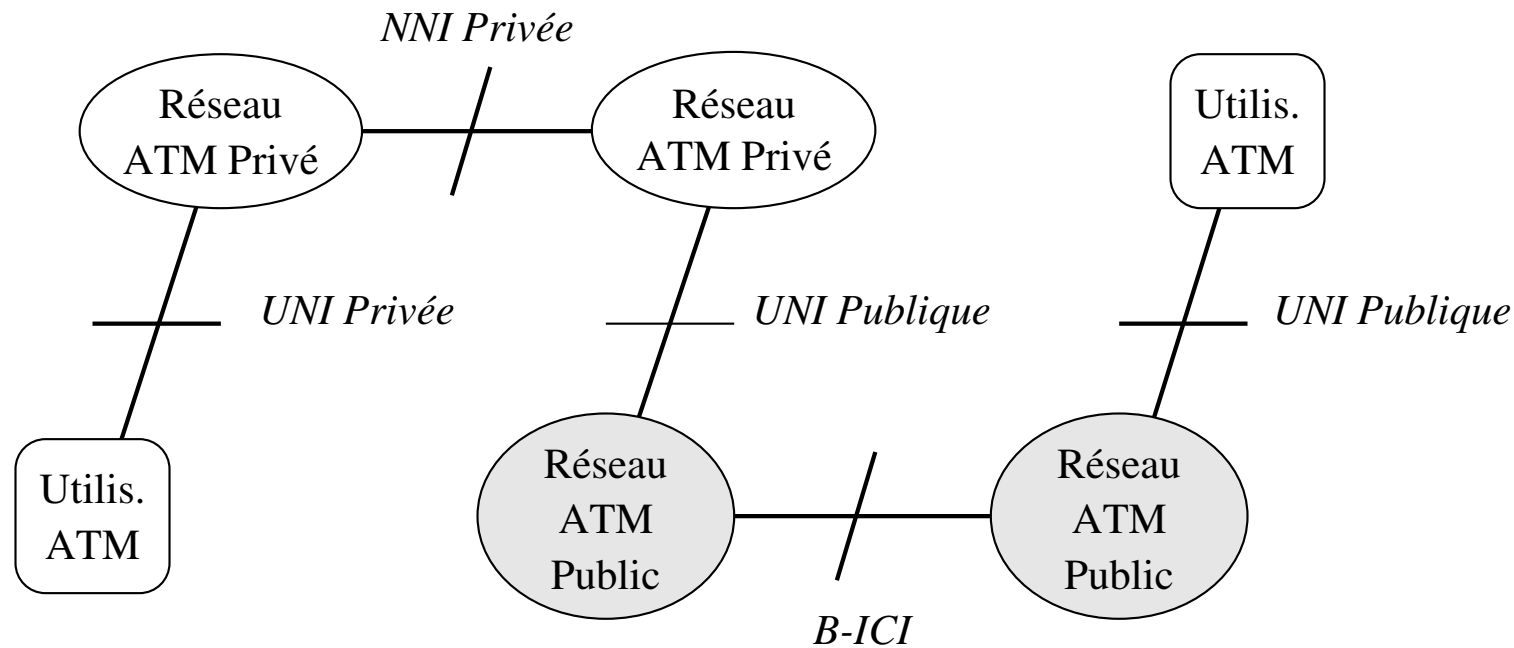
Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- **Services ATM**
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
- MPLS

Architecture point-à-point

Interfaces ATM



3 types d'interfaces dans les réseaux ATM :

B-ICI (*B-ISDN InterCarrier Interface*) entre les réseaux publics

NNI (*Private Network-Network Interface*) entre les réseaux privés

UNI (*User-Network Interface*) entre l'utilisateur et le réseau

ATM UNI

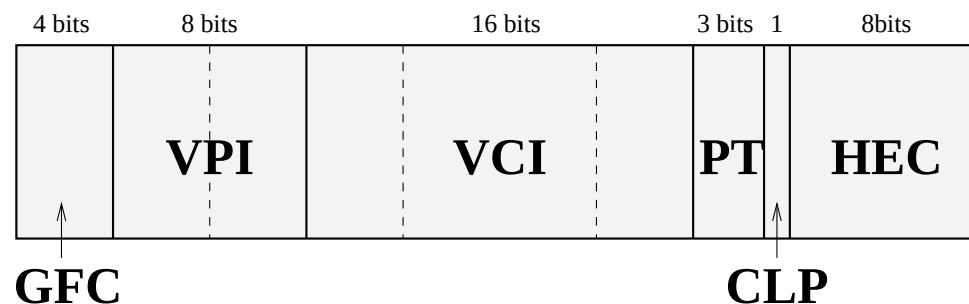
Service au niveau de l'hôte :

- accès aux réseaux **privés** (commutateurs du réseau de l'entreprise)
- accès aux réseaux **publics** (commutateurs de l'opérateur télécom)

Mécanismes identiques à ceux décrits précédemment...

- signalisation de gestion de la connexion
- négociation des paramètres de QoS

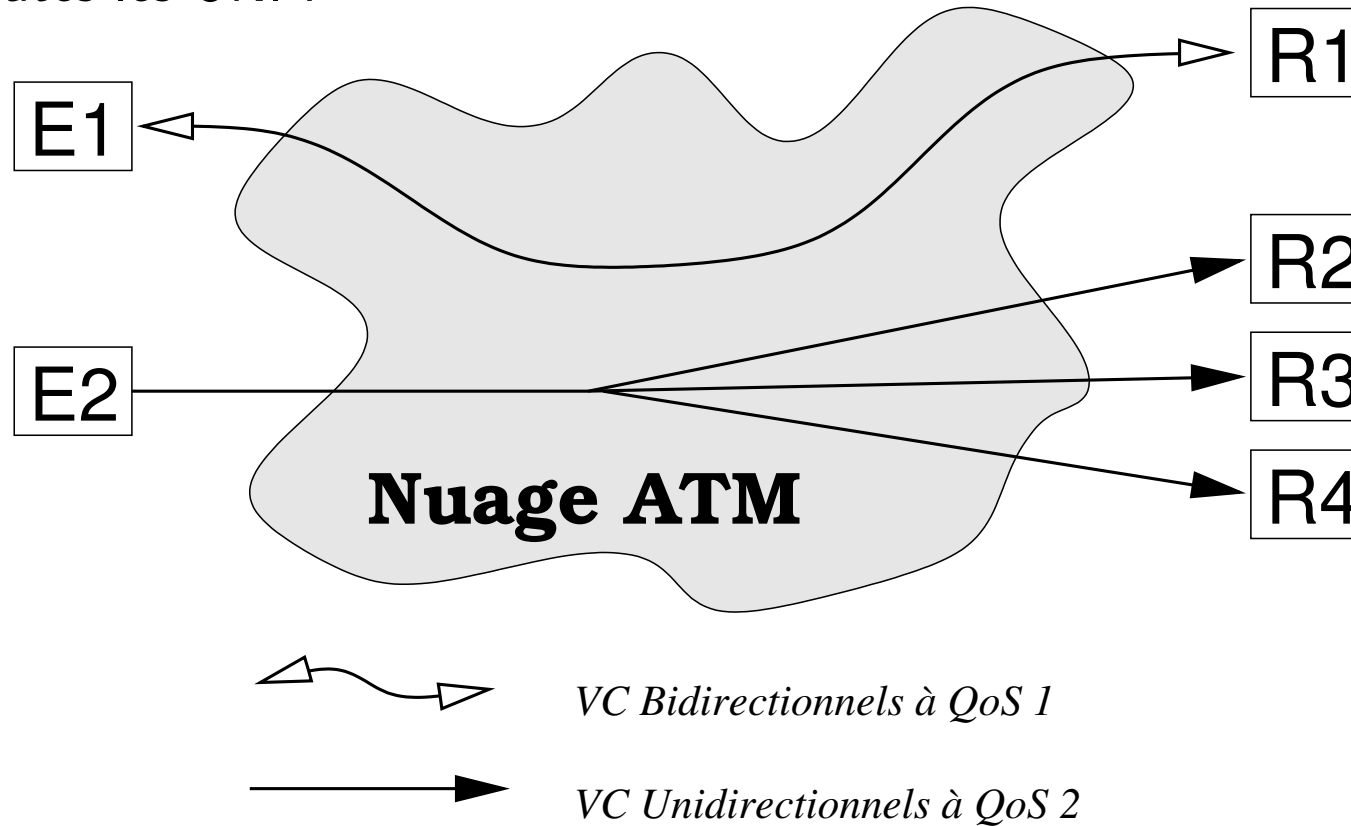
... avec un entête modifié (GFC envisagé pour un contrôle de flux local) :



GFC Generic Flow Control
VPI Virtual Path Identifier
VCI Virtual Chanel Identifier
PT Payload Type
CLP Cell Loss Priority
HEC Header Error Control

Connexions UNI

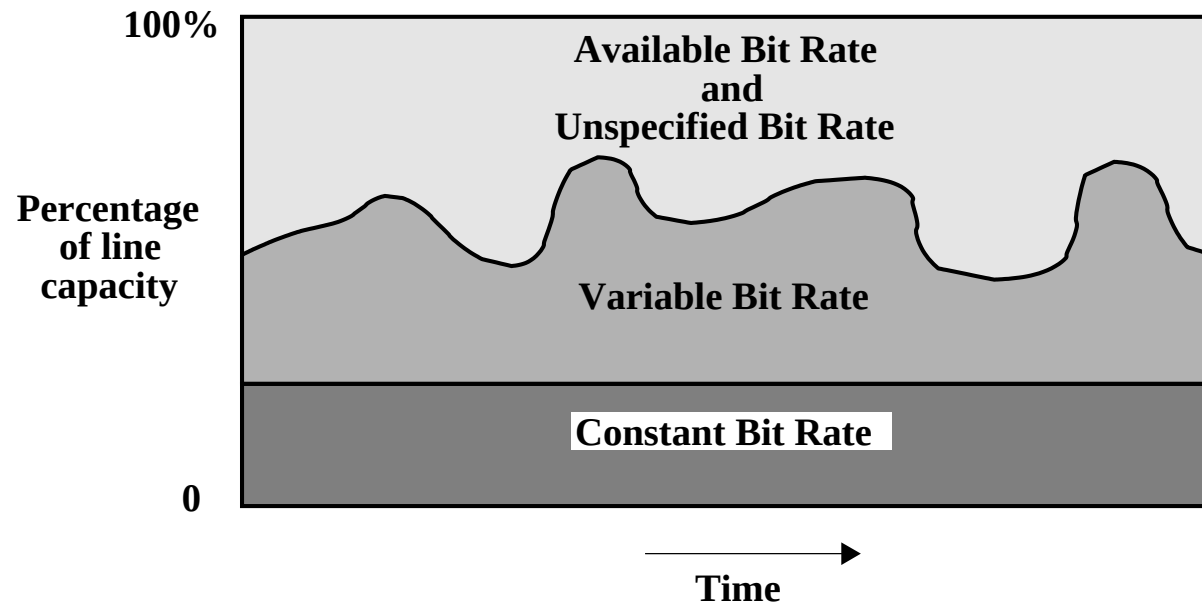
Pour toutes les UNI :



Avec négociation de paramètres de QoS

Classes de services UNI 3.x

UNI version 1993 (Recommandation Q.293x)



CBR (*Constant Bit Rate*) débit crête d'émission des cellules et variation de délai spécifiés

VBR (*Variable Bit Rate*) débit moyen, débit crête et taille des rafales de cellules spécifiés

UBR (*Unspecified Bit Rate*) pas de paramètres de QoS spécifiés

Services UNI 4.0

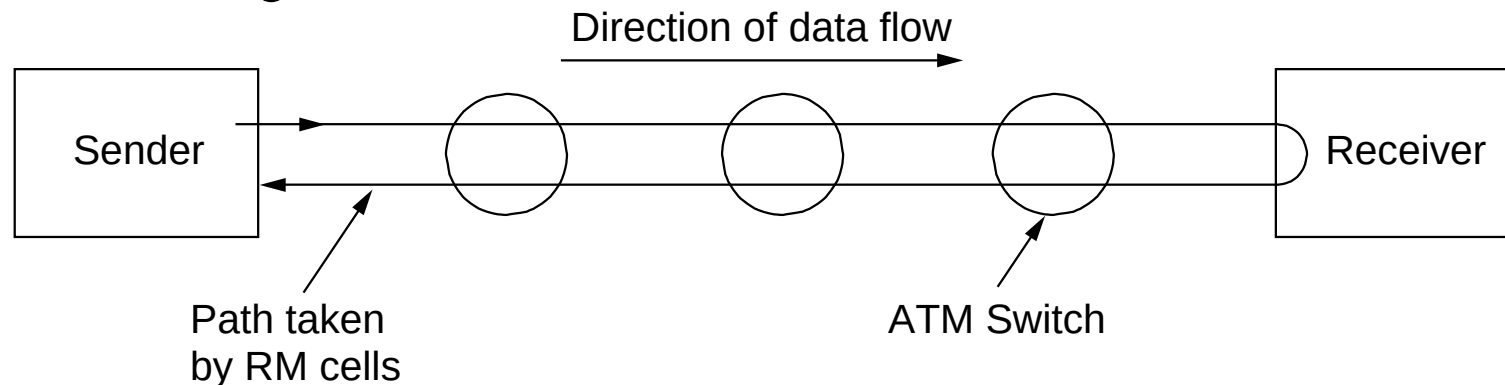
UNI version 1996

CBR (*Constant Bit Rate*) idem UNI 3.x CBR

rt-VBR (*Real-Time Variable Bit Rate*) UNI 3.x VBR avec contrainte sur la variation du délai de transmission

nrt-VBR (*Non Real-Time Variable Bit Rate*) idem UNI 3.x VBR

ABR (*Available Bit Rate*) pour les applications adaptatives, basé sur un contrôle de congestion réactif



UBR (*Unspecified Bit Rate*) idem UNI 3.x UBR

Paramètres de QoS

Beaucoup de paramètres de QoS – voici les principaux :

Acronyme	Description
PCR	Le <i>Peak Cell Rate</i> indique le débit d'émission maximum des cellules.
SCR	Le <i>Sustainable Cell Rate</i> indique une borne supérieure du débit d'émission moyen des cellules.
MCR	Le <i>Minimum Cell Rate</i> indique le débit d'émission de cellules minimum.
MBS	Le <i>Maximum Burst Size</i> indique la taille maximum des agrégats de cellules.
CDVT	Le <i>Cell Delay Variation Tolerance</i> indique la gigue maximum d'une cellule.
CLR	Le <i>Cell Loss Ratio</i> indique le taux de cellules perdues ou arrivées trop tard.
CTD	Le <i>Cell Transfer Delay</i> indique le délai de transmission d'une cellule.
CDV	Le <i>Cell Delay Variation</i> indique la variation de délai de transmission d'une cellule.

Catégories de service ATM

Les contrats de trafic UNI sont définis à partir des paramètres de QoS :

<i>Paramètres QoS</i>	CBR	rt-VBR	nrt-VBR	ABR	UBR
CLR pour CLP=0	✓	✓	✓	✓	×
CLR pour CLP=1	○	○	○	✓	×
CTD	✓	✓	✓	×	×
CDV	✓	✓	×	×	×
SCR et MBS	¬	✓	✓	¬	¬
PCR et CDVT	✓	✓	✓	✓	✓
MCR	¬	¬	¬	✓	¬
<i>Contrôle de flux</i>	¬	¬	¬	✓	¬

“✓” spécifié ou utilisé, “○” optionnel, “×

Plan

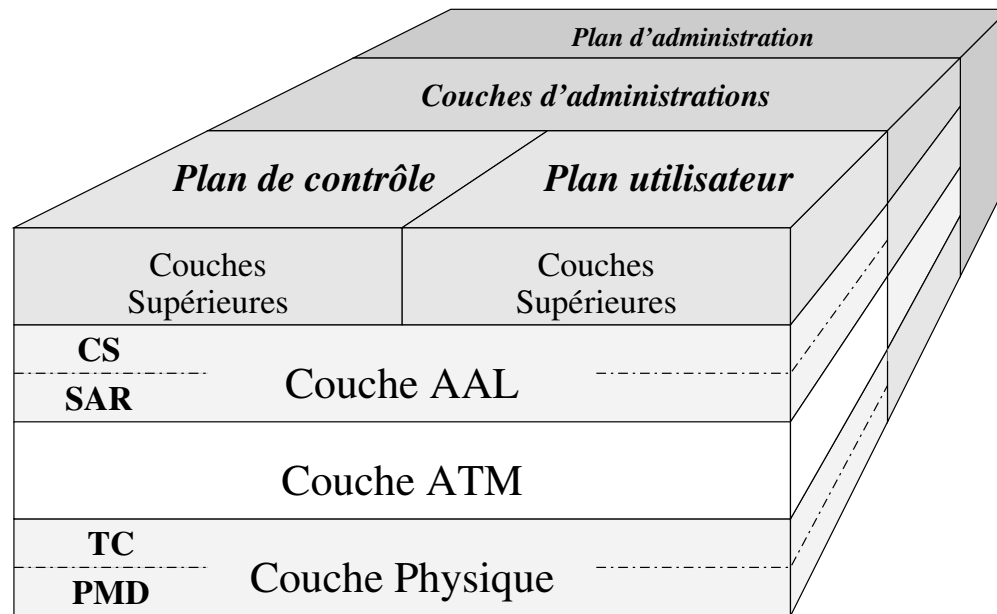
Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- **Modèle de référence ATM**
- IP sur ATM
- Commutation IP
- MPLS

Architecture point-à-point

Modèle de référence ATM



AAL ATM Adaptation Layer
CS Convergence Sublayer
SAR Segmentation And Reassembly sublayer
TC Transmission Convergence sublayer
PMD Physical Medium Dependent sublayer

3 plans divisés en plusieurs couches :

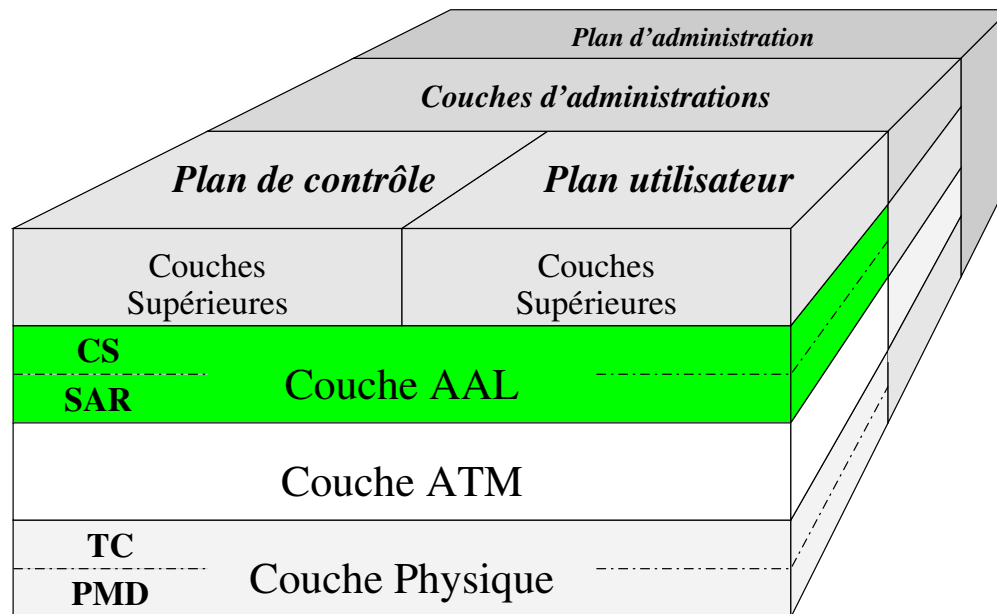
Couches Supérieures : applications ou protocoles s'appuyant sur ATM

Couche AAL : adaptation aux extrémités du réseau

Couche ATM : commutation des cellules

Couche Physique : accès au médium

Modèle de référence ATM : Couche AAL



AAL	ATM Adaptation Layer
CS	Convergence Sublayer
SAR	Segmentation And Reassembly sublayer
TC	Transmission Convergence sublayer
PMD	Physical Medium Dependent sublayer

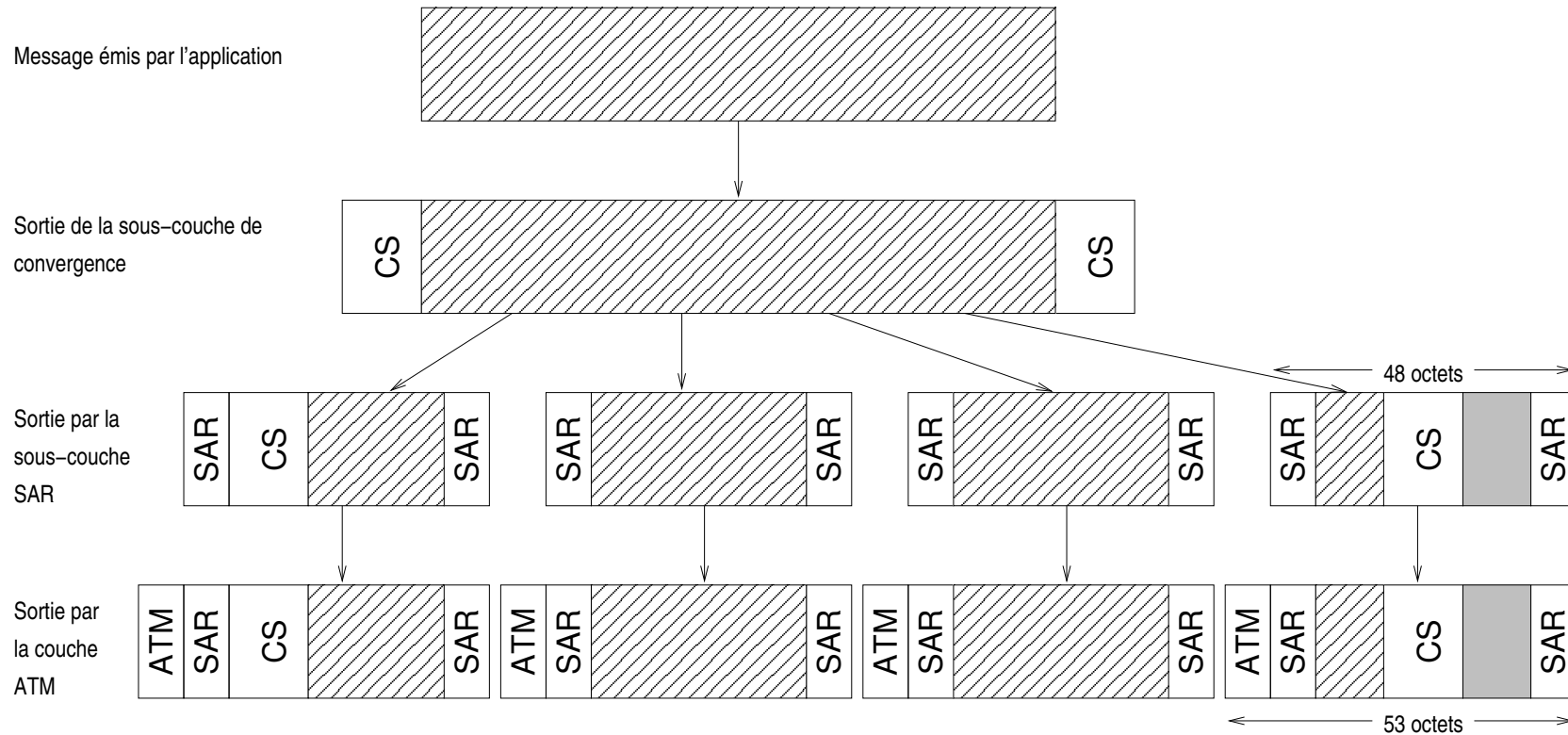
Couche AAL (*ATM Adaptation Layer*) située aux extrémités du réseau pour fournir un service adapté aux niveaux supérieurs

- 2 sous couches :

CS convergence des différents services

SAR segmentation et réassemblage des AAL PDU (*Protocol Data Unit*) en cellules

AAL : Encapsulations



AAL : CS

Convergence Sublayer

Correspond principalement à la **sous-sous-couche CPCS** (*Common Part Convergence Sublayer*)... il y a aussi une sous-sous-couche SSPCS (*Service Specific Part Convergence Sublayer*) !

Fonctionnalités :

- liaison avec les applications par un point accès SAP (*Service Access Point*, adresse applicative)
- manipule des messages CPCS-PDU
- dépend du type de service choisi
 - ✓ ajout/suppression de l'encapsulation CPCS (si besoin)
 - ✓ détection d'erreur de bout-en-bout sur CPCS-PDU
 - ✓ contrôle de flux
 - ✓ gestion de la performance de bout-en-bout
 - ✓ contrôle de congestion
 - ✓ bourrage...

AAL : SAR

Segmentation And Reassembly

Fonctionnalités :

- segmentation des messages CPCS en données pour les cellules
- manipule des cellules SAR-PDU
- dépend aussi du type de service choisi
 - ✓ ajout/suppression de l'encapsulation SAR (si besoin) en restant dans la limite de 48 octets
 - ✓ détection d'erreur de bout-en-bout au niveau d'une cellule
 - ✓ contrôle de priorité de perte (bit CLP)
 - ✓ contrôle de congestion
 - ✓ multiplexage...

AAL : Classification des services

4 classes de services initialement définies par l'ITU

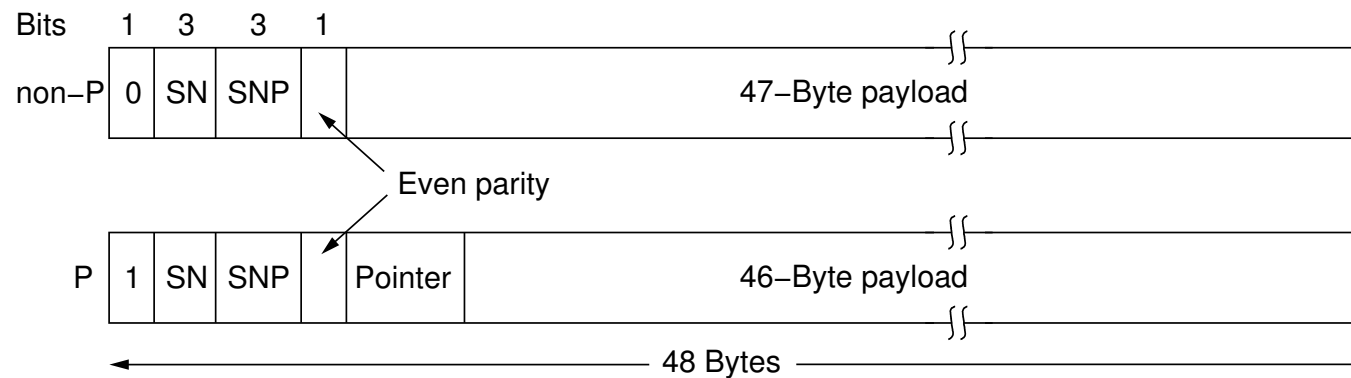
- à chaque classe correspond une AAL
 - AAL 3 et AAL 4 très proche : combinées dans l'AAL 3/4
 - industrie informatique a défini l'AAL 5
- ✓ plus performante pour les données

	A		B		C		D	
Timing	Real time	None	Real time	None	Real time	None	Real time	None
Bit rate	Constant		Variable		Constant		Variable	
Mode	Connection orientated				Connectionless			
AAL	1		2		3/4		3/4	
					5			

AAL 1 et AAL 2

AAL 1 :

- transfert de PDU en CBR (audio et video)
 - ✓ CS : gestion synchro et délimitation (Pointer)
 - ✓ SAR : Segmentation et réassemblage + numéro de séquence (SN)
+ protection (SNP)

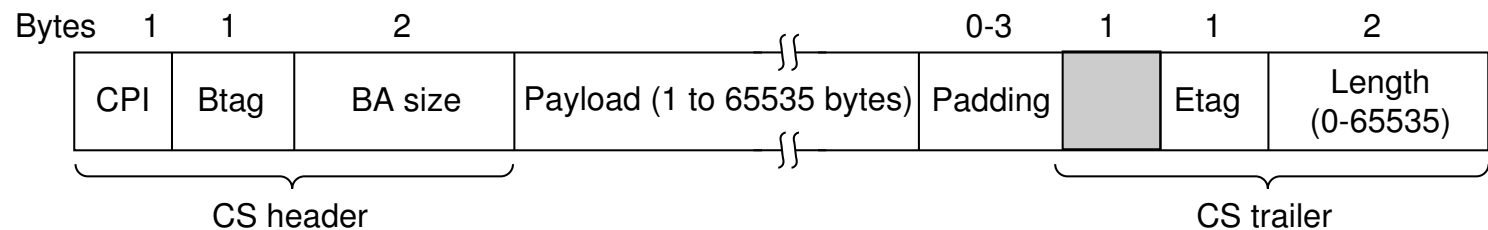


AAL 2 :

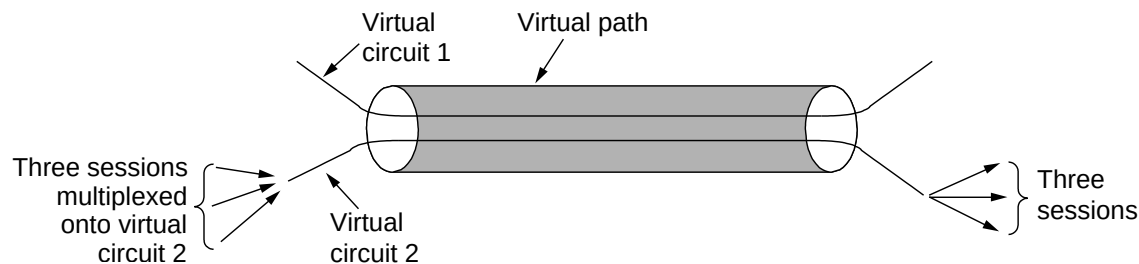
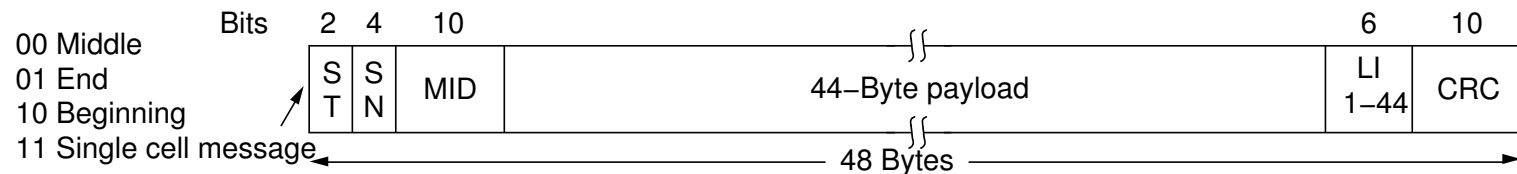
- transfert de PDU en VBR
 - ✓ SAR et CS : très complexes
 - ▣ audio et vidéo avec codec à débit variable

AAL 3/4

- transfert sans contraintes temporelles (mais sensible aux pertes)
 - ➡ données informatiques
 - ✓ deux modes : **message** (non connecté) ou **flux** (connecté)
 - ✓ CS : délimitation du message ou des segments du flux

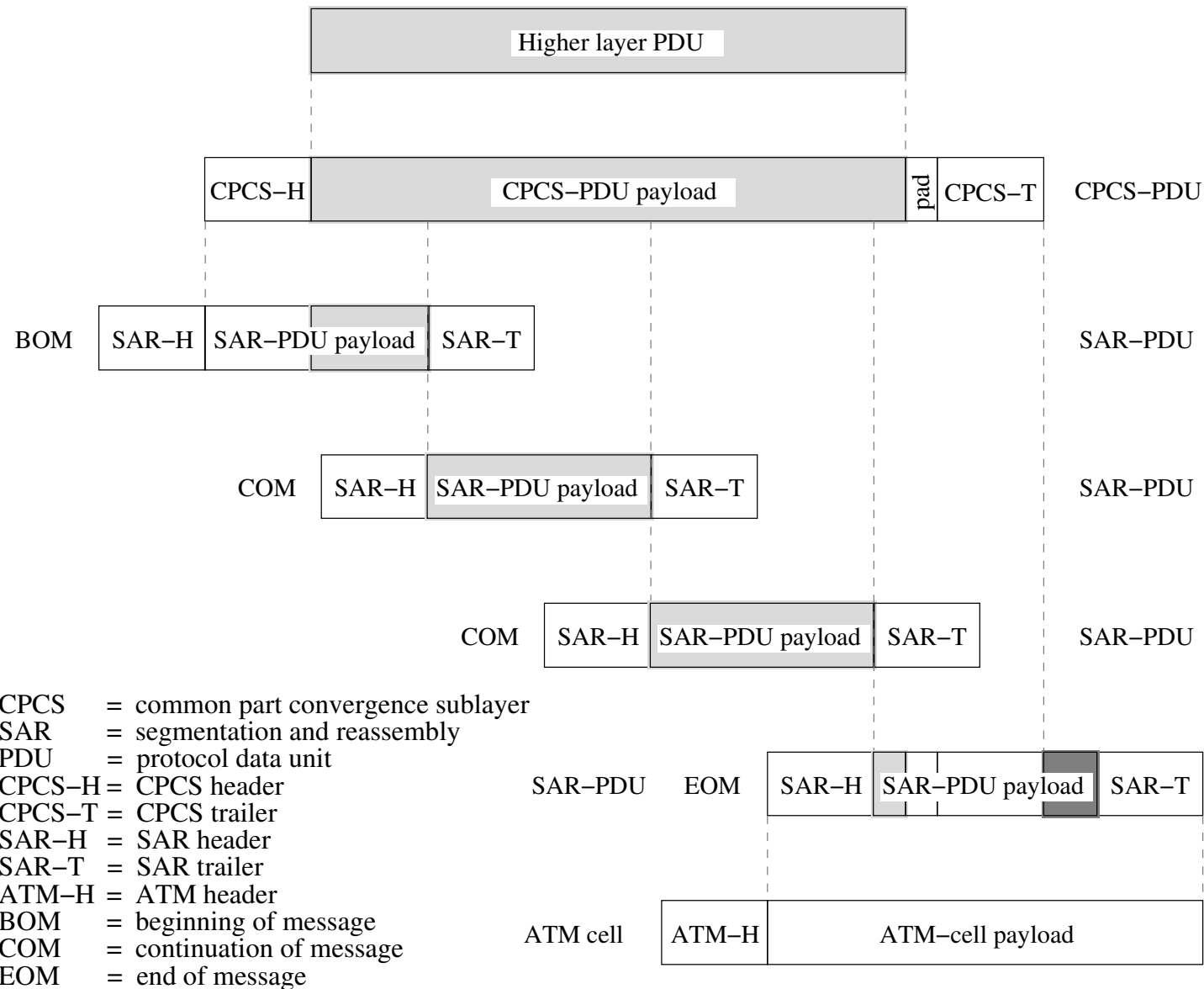


- ✓ SAR : découpage en trame, multiplexage et contrôle d'erreur



➡ Trop lourds !

AAL 3/4 : Transmission

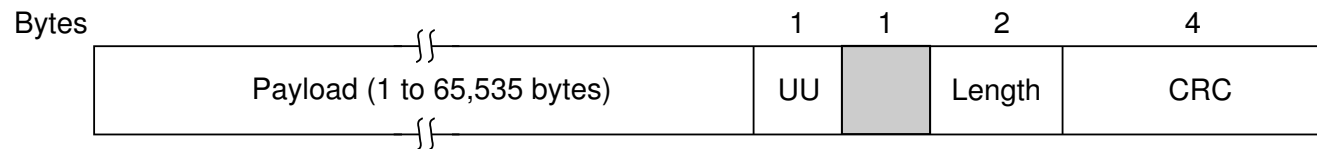


AAL 5

AAL 1, AAL 2 et AAL 3/4 définis par l'industrie télécom (ITU) : complexes et peu adaptés au transfert de données...

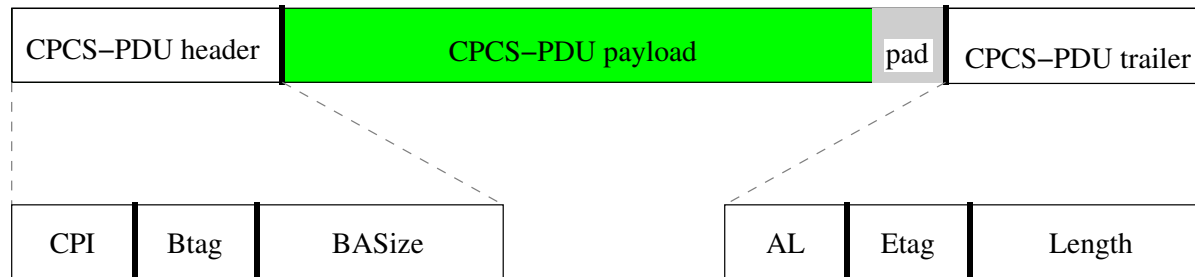
➡ l'industrie informatique propose SEAL (*Simple and Efficient Adaptation Layer*) :

- renommé **AAL5** (ATM Forum)
- transfert de flux orienté connexion allégé
 - ✓ CS : données utilisateurs + contrôle d'erreur



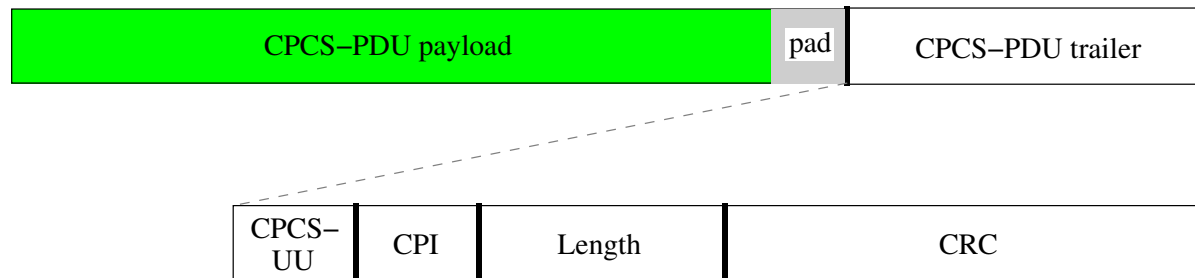
- ✓ SAR : segmentation et réassemblage avec les données de cellule complètes
 - ☞ pas d'encapsulation SAR
 - ☞ délimitation des CPCS-PDU avec le bit PTI de l'entête de la cellule ATM

AAL 5 : Simplification des CPCS



CPI = common part indicator (1 octet)
 Btag = beginning tag (1 octet)
 BASize = buffer allocation size (2 octets)
 AL = alignment (1 octet)
 Etag = end tag (1 octet)
 Length = length of CPCS-PDU payload (2 octets)

AAL Type 3/4

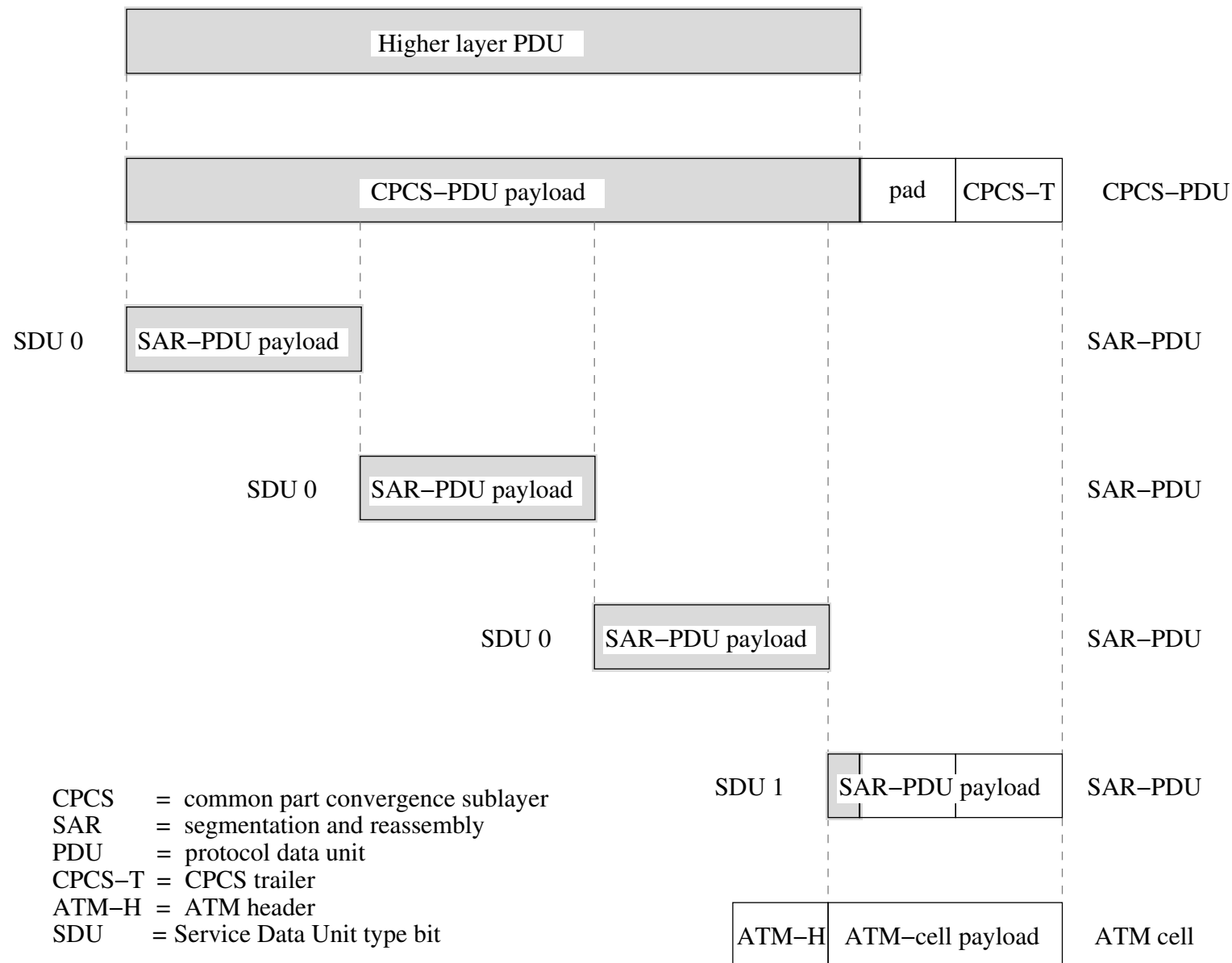


CPCS-UU = CPCS user-to-user indication (1 octet)
 CPI = common part indicator (1 octet)
 Length = length of CPCS-PDU payload (2 octets)
 CRC = cyclic redundancy check (4 octets)

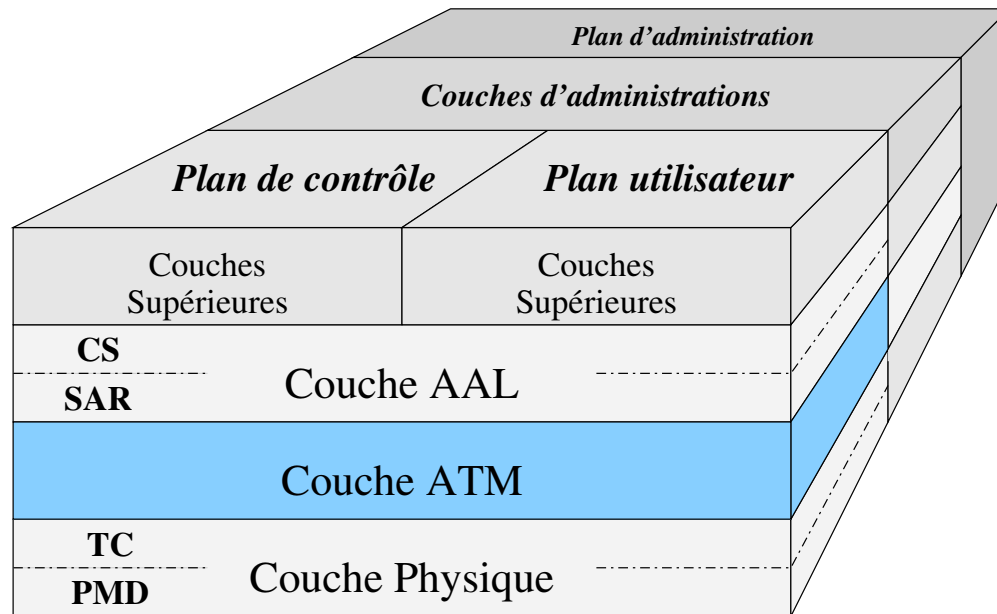
AAL Type 5

pictures from STALLINGS W. *High Speed Networks*

AAL 5 : Transmission



Modèle de référence ATM : Couche ATM



AAL *ATM Adaptation Layer*
CS *Convergence Sublayer*
SAR *Segmentation And Reassembly sublayer*
TC *Transmission Convergence sublayer*
PMD *Physical Medium Dependent sublayer*

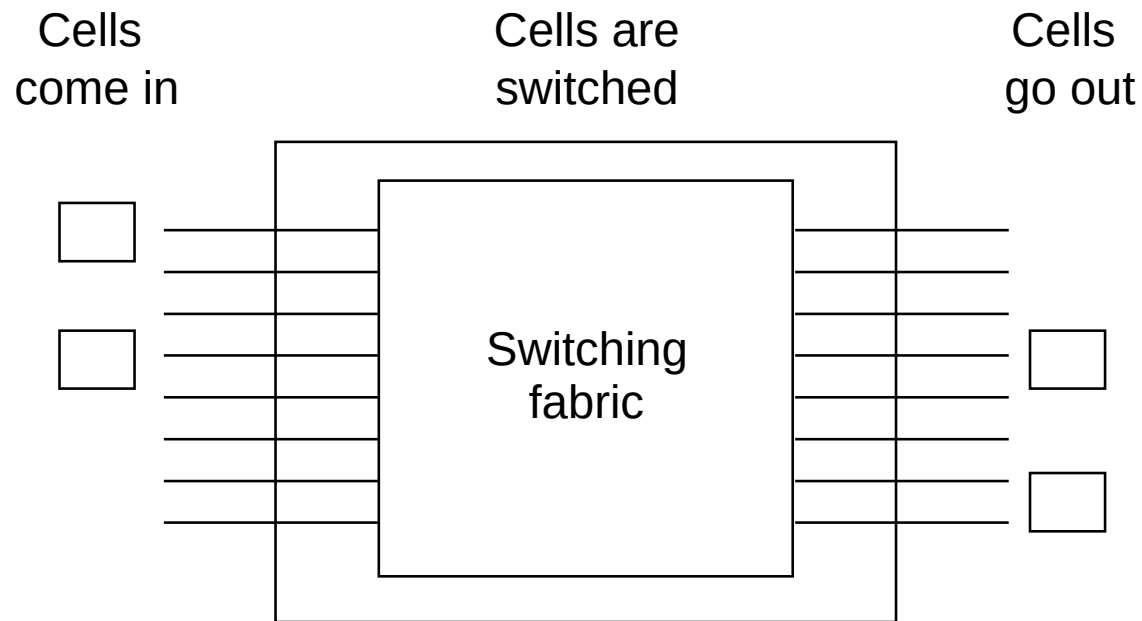
Couche ATM assure le "relayage" des informations (commutation)

- commutation des cellules
- multiplexage des cellules (grâce aux VPI et VCI)
- indépendante du médium
- contrôle de flux

► traitement principal des éléments de commutation d'un réseau ATM

Commutateur ATM (1)

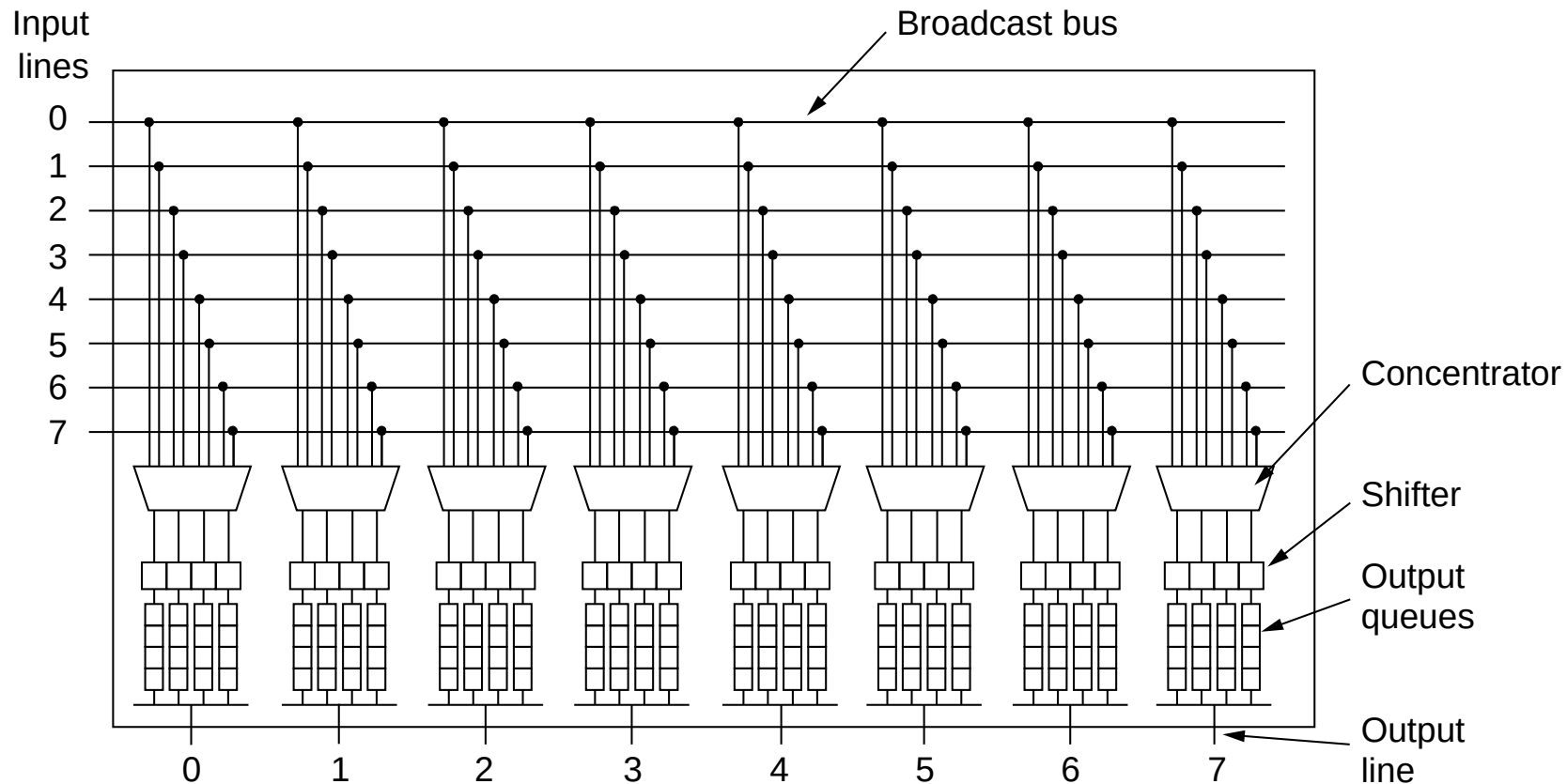
Switching fabric



Problème de gestion des collisions de cellules souhaitant accéder au même port de sortie (tampons limités)

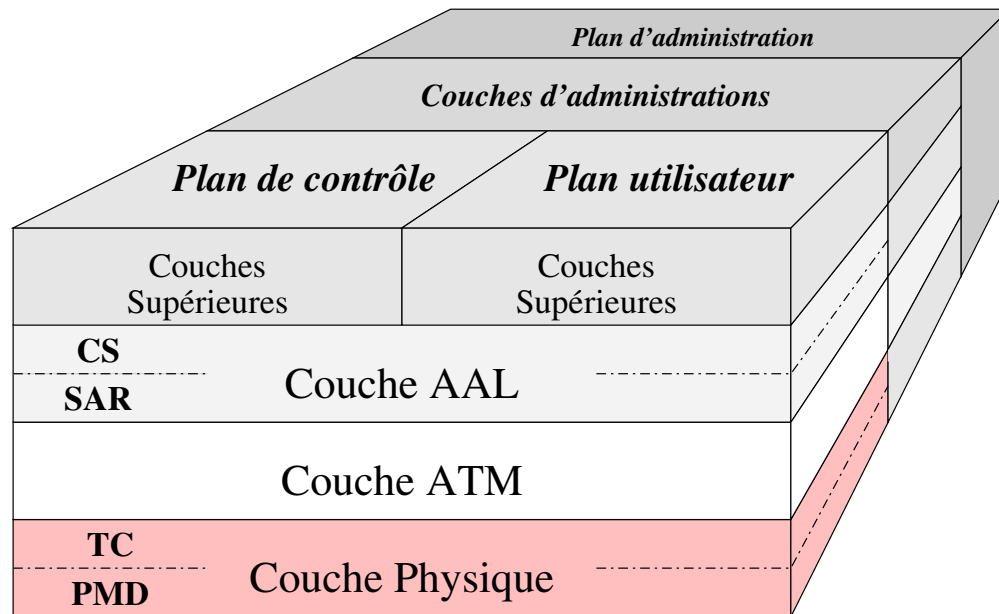
Commutateur ATM (2)

Exemple : commutateur synchrone de type Knockout (crossbar + tournoi pour obtenir le tampon de sortie)



picture from TANENBAUM A. S. *Computer Networks 3rd edition*

Modèle de référence ATM : Couche Physique



AAL ATM Adaptation Layer
CS Convergence Sublayer
SAR Segmentation And Reassembly sublayer
TC Transmission Convergence sublayer
PMD Physical Medium Dependent sublayer

Couche Physique contrôle de l'accès au support physique

- 2 sous-couches :

TC convergence gérant les fonctionnalités communes aux couches physiques ayant généralement rapport avec la conversion des cellules en bits

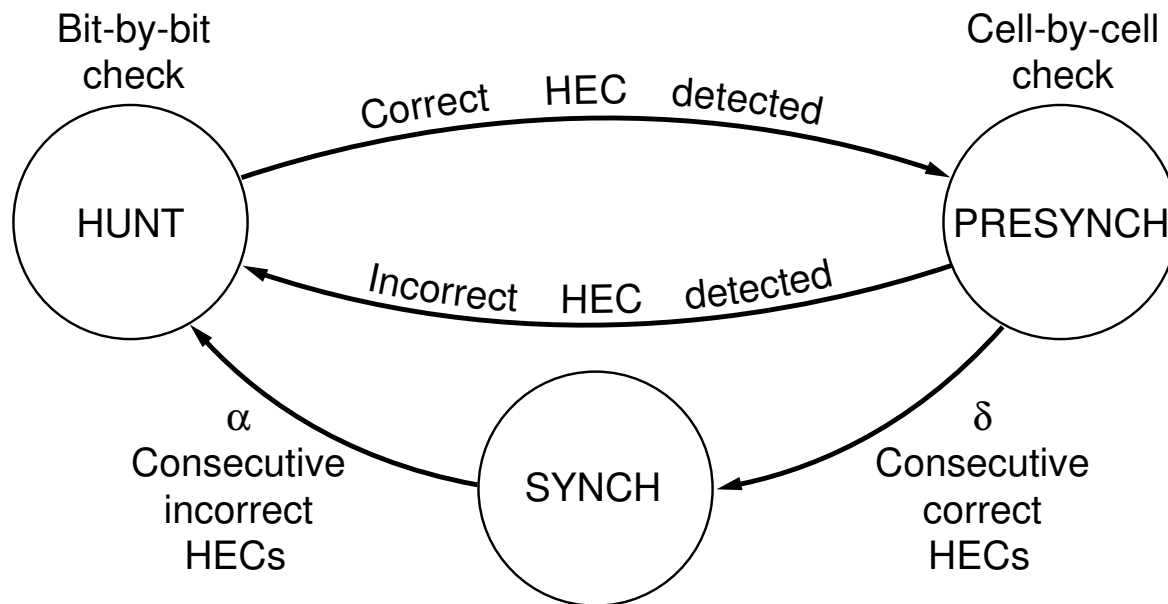
PMD spécifique au médium concerné

ATM PHY-TC : Synchronisation

Transmission Convergence Sublayer

Le médium envoi un flux de bits ou d'octets continu :

- cellules vides (*Idle cells*) si pas de data transmis
- ➡ comment ce flux se synchronise sur les cellules ?

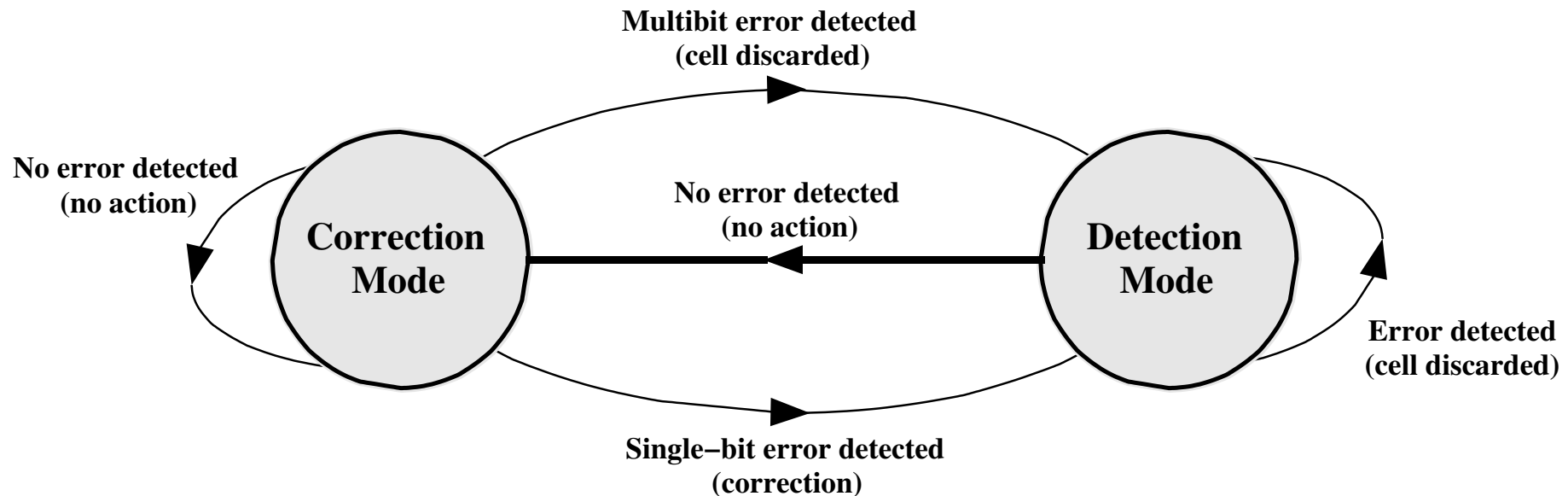


pictures from TANENBAUM A. S. *Computer Networks 3rd edition*

ATM PHY-TC : Protection de l'en-tête

Contrôle d'erreur limité à l'en-tête de la cellule

- CRC sur 8 bits
 - ✓ polynome générateur : $X^8 + X^2 + X + 1$
 - ✓ corrige les erreurs sur 1 bit
 - ✓ détecte les erreurs sur plusieurs bits



pictures from STALLINGS W. *High Speed Networks*

ATM PHY-PMD : Média

Physical Medium Dependent Sub-layer

Récupération des bits du médium

- support avec faible taux d'erreur ($\sim 10^{-9}$)
 - ✓ concut pour les fibres optiques...
- ATM ne standardise pas le format pour la transmission des cellules
 - ✓ intégration dans des multiplexes : E1, E3, SONET...
 - ☞ débits initiaux : 155.52Mbps et 622.08Mbps (OC3 et OC12)
 - ✓ encapsulation directe
 - ☞ tout ce qui peut transmettre des bits !

Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- **IP sur ATM**
 - ✓ Classical IP
 - ✓ LAN Emulation
- Commutation IP
- MPLS

Architecture point-à-point

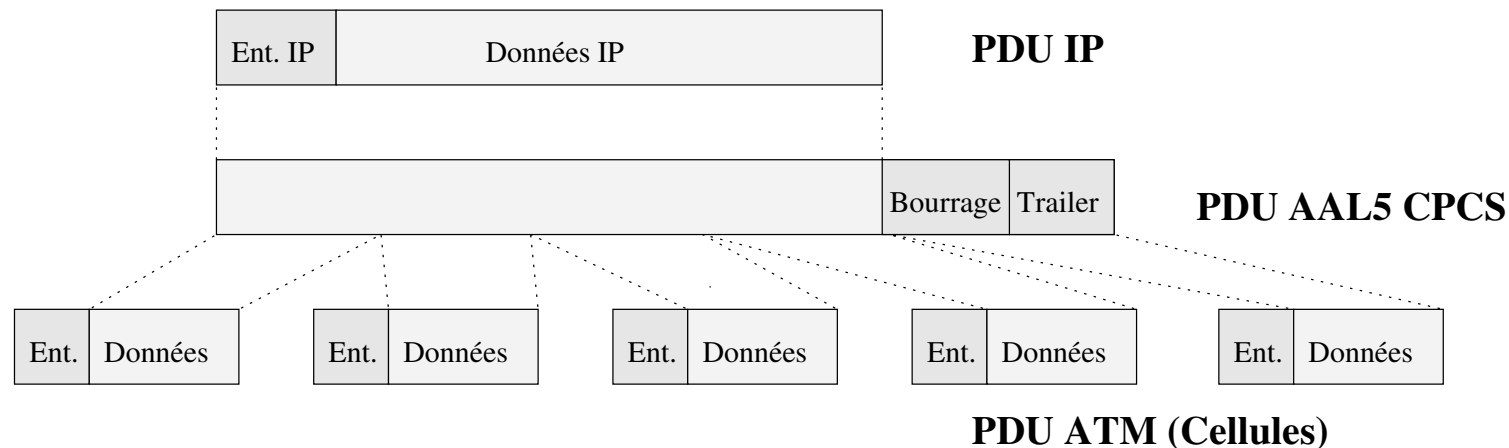
ATM : Encapsulation du trafic de données (1)

Standard IETF :

- **encapsulation RFC 1483 :**

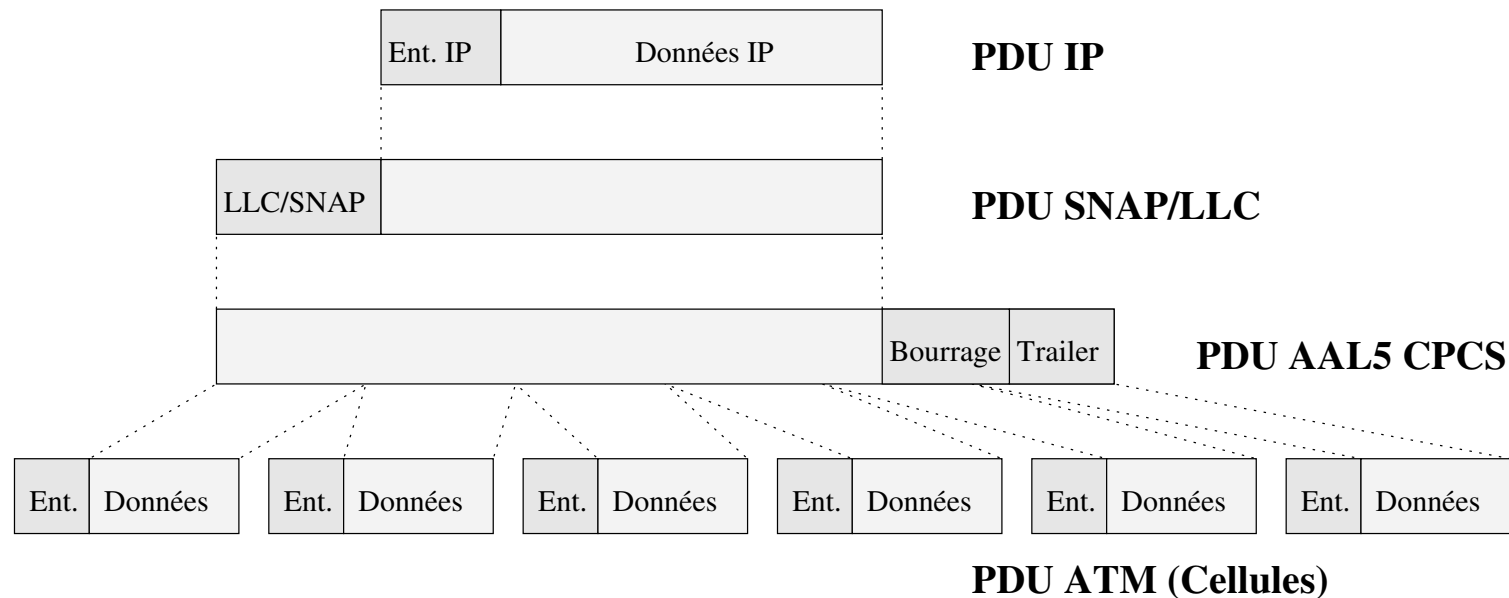
- ✓ intégration dans les CPCS PDU de l'AAL 5
- ✓ deux types d'encapsulation : VCMUX ou SNAP/LLC

☞ **VCMUX** : encapsulation nulle
(ne permet que le transport d'un protocole prédéfini)



ATM : Encapsulation du trafic de données (2)

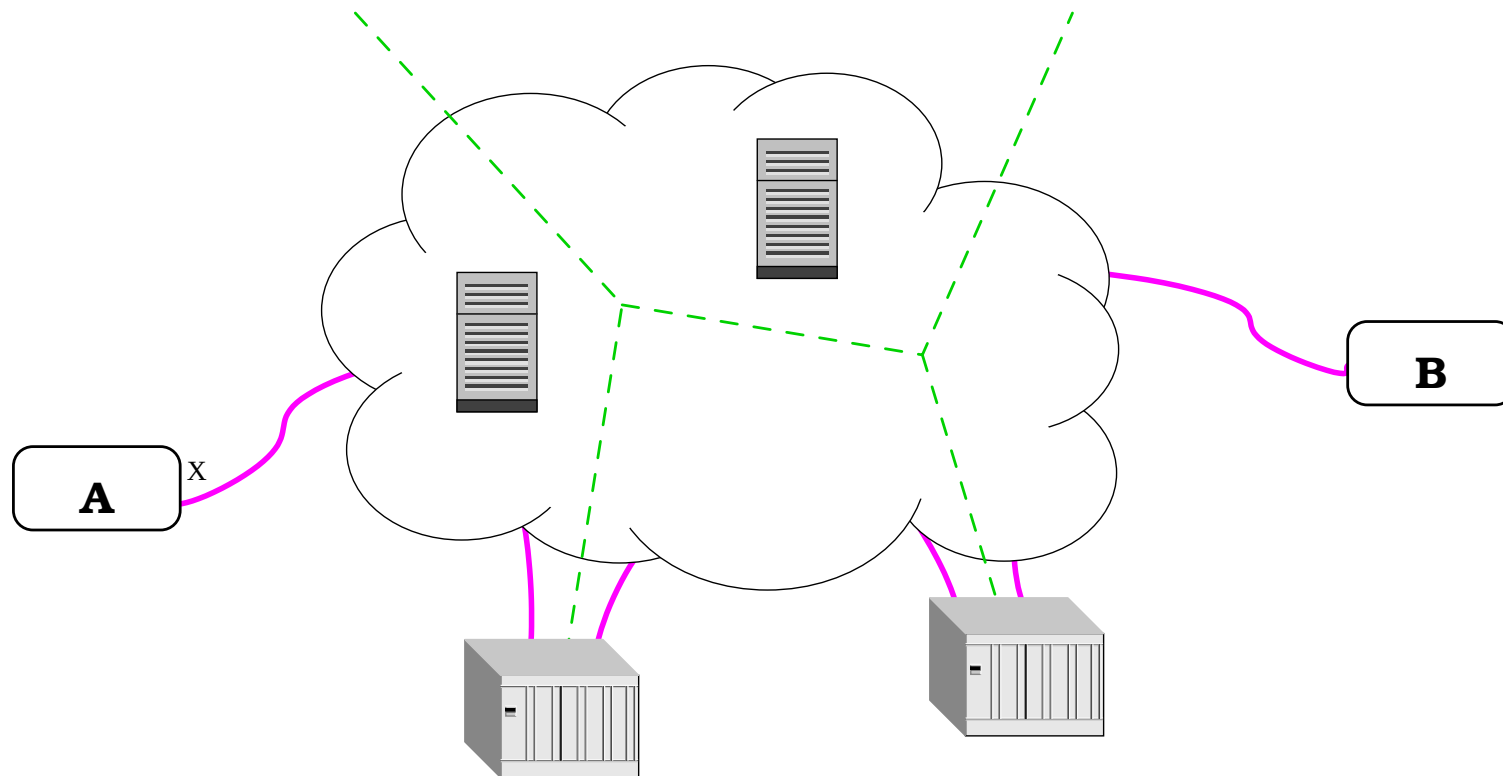
- **encapsulation RFC 1483 (suite) :**
 - ✓ deux types d'encapsulation : VCMUX ou SNAP/LLC
 - ☞ **SNAP/LLC** : permet le pontage vers les LAN IEEE (spécifie le protocole transporté)



ATM : Encapsulation d'IP

Comment gérer l'adressage d'un réseau NBMA ?

- IP est adapté aux réseaux de type LAN avec diffusion
 - ✓ segmentation en sous-réseau **virtuel** similaire à des LAN
 - ✓ service remplaçant la diffusion



Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
 - ✓ **Classical IP**
 - ✓ LAN Emulation
- Commutation IP
- MPLS

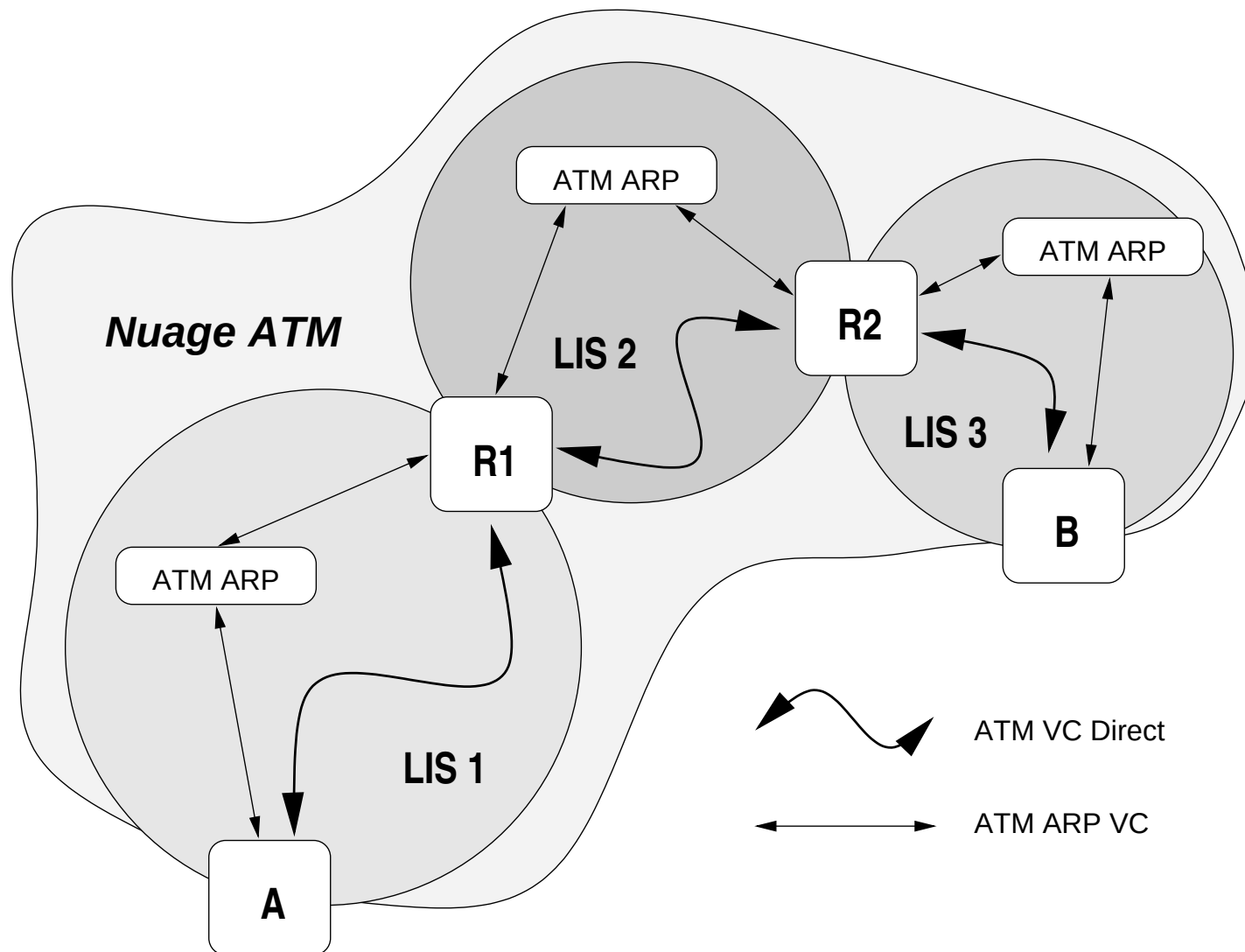
Architecture point-à-point

Classical IP : Introduction

Approche usuelle de l'IETF (Overlay)

- **encapsulation RFC 1483 :**
 - ✓ intégration dans les CPCS PDU de l'AAL 5
 - ☞ MTU recommandée = 9180 octets
 - ☞ encapsulation de type VCMUX ou LLC/SNAP
- **résolution d'adresse :**
 - ✓ réseau ATM : NBMA connectivité directe mais pas de diffusion
 - ☞ découpage en **LIS** (*Logical IP Subnetwork*)
 - ☞ serveur de résolution d'adresse dans un LIS : **ATMARP**

Classical IP : Résolution



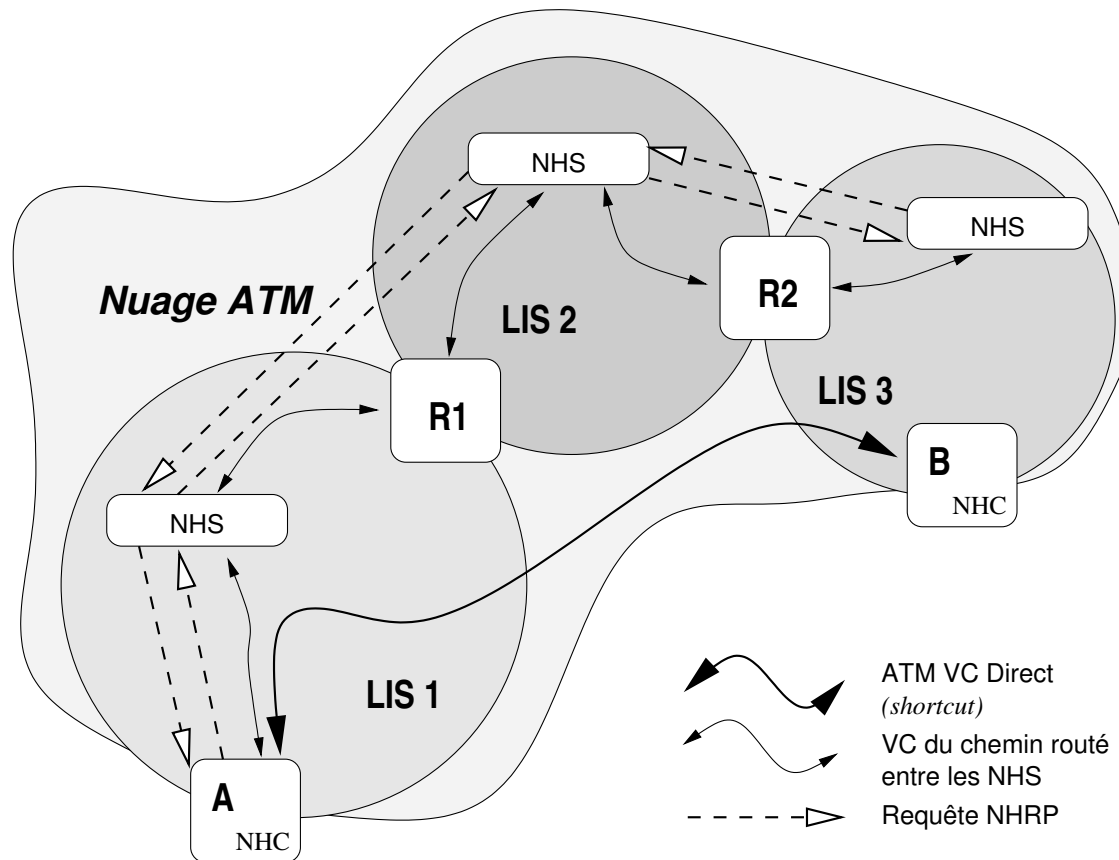
Classical IP : Remarques

- première solution **data** sur ATM (1993)
 - ✓ références initiales : 1993-95
 - ➡ RFC 1483 *Multiprotocol Encapsulation over ATM AAL 5*
 - ➡ RFC 1577 *Classical IP and ARP over ATM*
 - ➡ RFC 1626 *Default IP MTU for use over ATM AAL5*
 - ➡ RFC 1755 *ATM Signaling Support for IP over ATM - UNI 3.x*
 - ✓ références actuelles : 1998
 - ➡ RFC 2684 *Multiprotocol Encapsulation over ATM AAL 5*
 - ➡ RFC 2225 *Classical IP and ARP over ATM*
 - ➡ RFC 2331 *ATM Signalling Support for IP over ATM - UNI 4.0*
- quelques défauts :
 - ✓ centralisé, serveur unique
 - ✓ mode connecté sous-jacent (latence)
 - ✓ topologie virtuelle LIS décorréée de la topologie réelle
 - ✓ pas de QoS (➡ RSVP ?)
 - ✓ pas de multicast (➡ MARS ?)
 - ✓ routeurs intermédiaires (➡ **NHRP**)

NHRP

Next Hop Resolution Protocol

- évite les routeurs intermédiaires sur un réseau subdivisé (ATM + LIS)
 - ✓ **NHS** (*Next Hop Server*) dans chaque LIS
 - ✓ récupération de l'adresse ATM du destinataire
 - ✓ création de raccourcis ATM ➡ *shortcuts*



Plan

Architecture Ethernet

Architecture ATM/MPLS

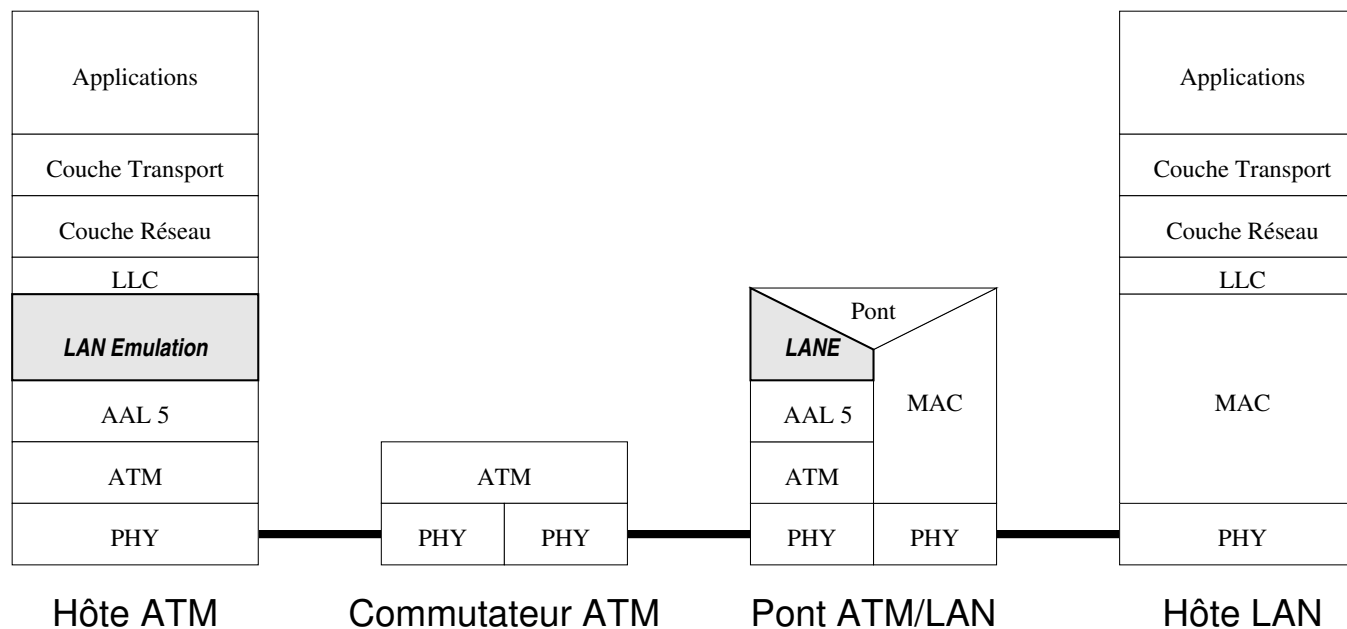
- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
 - ✓ Classical IP
 - ✓ **LAN Emulation**
- Commutation IP
- MPLS

Architecture point-à-point

LAN Emulation

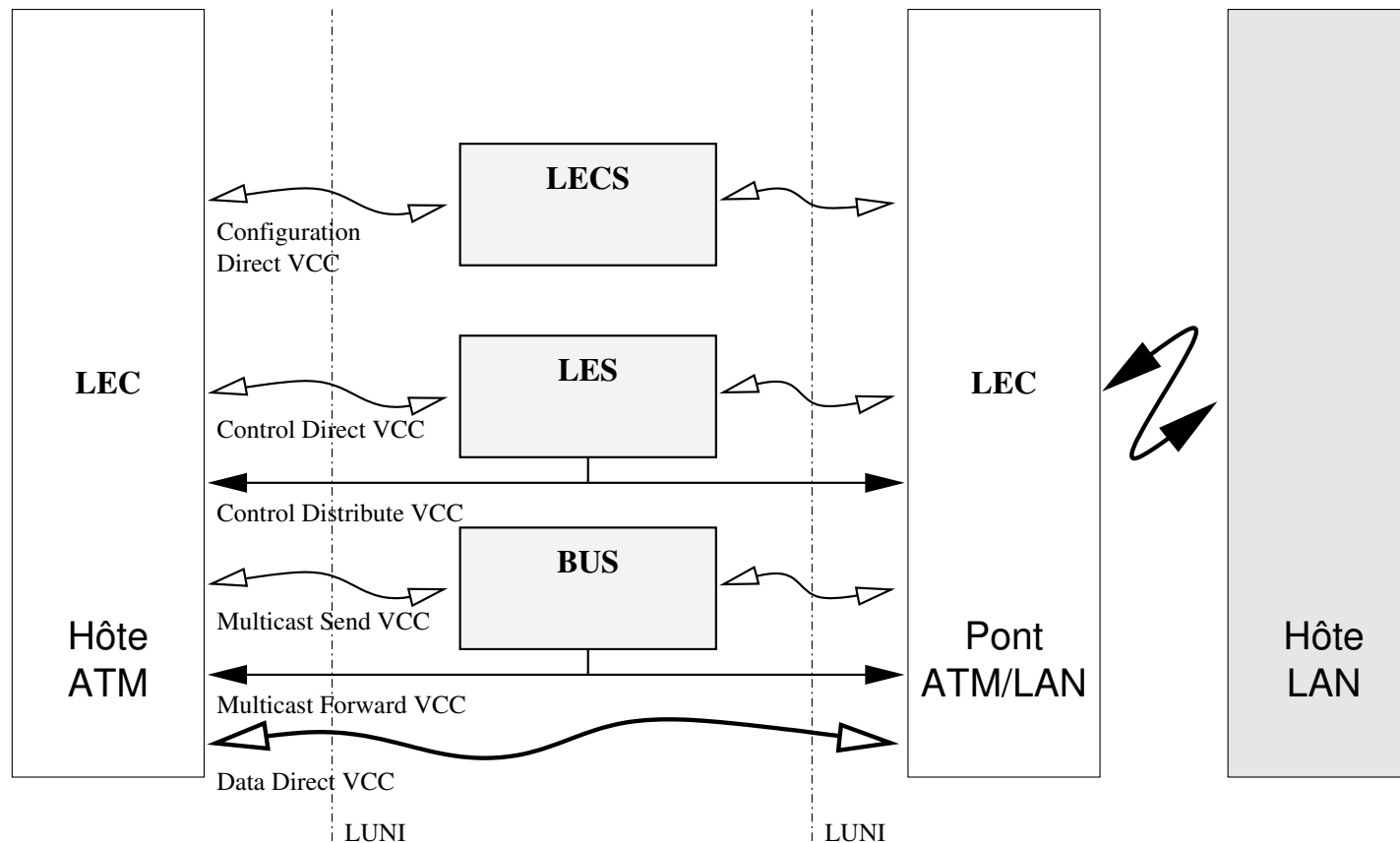
Support du trafic au niveau de la **couche MAC** des LAN IEEE 802

- **encapsulation RFC 1483** :
 - ✓ intégration dans les CPCS PDU de l'AAL 5
 - ✓ encapsulation de type LLC/SNAP ➡ pontage vers Ethernet/TR
- **résolution d'adresse** :
 - ✓ découpage en **ELAN** (*Emulated LAN*) + nombreux **serveurs**



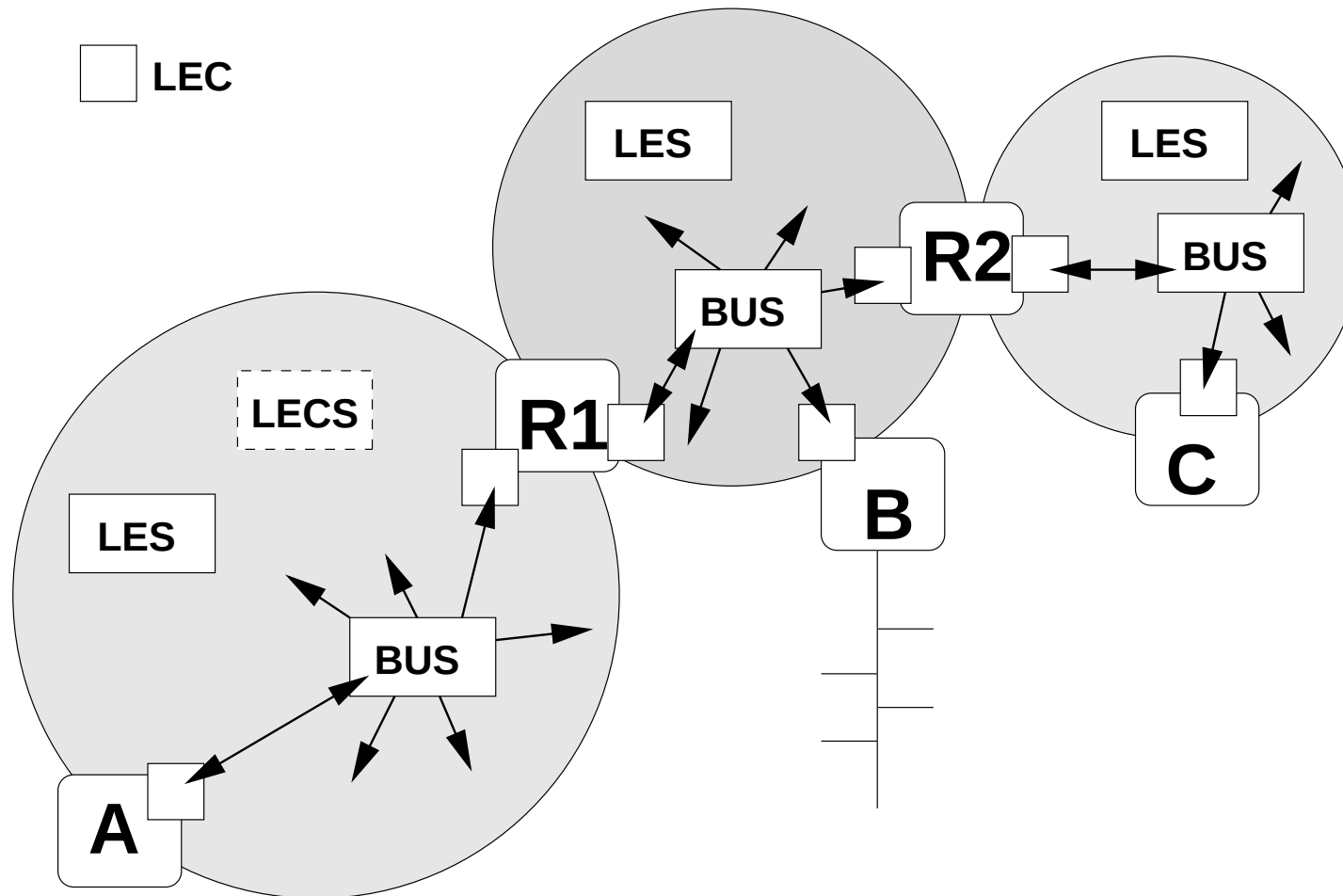
LANE : Connexions

- serveurs associés à la solution LANE :
 - ✓ **LEC** (*LAN Emulation Client*) : client au sein de l'hôte
 - ✓ **LECS** (*LAN Emulation Configuration Serveur*) : @LES, @BUS...
 - ✓ **LES** (*LAN Emulation Server*) ~ ATMARP
 - ✓ **BUS** (*Broadcast and Unknown Server*) : gestion de la diffusion



LANE : Résolution

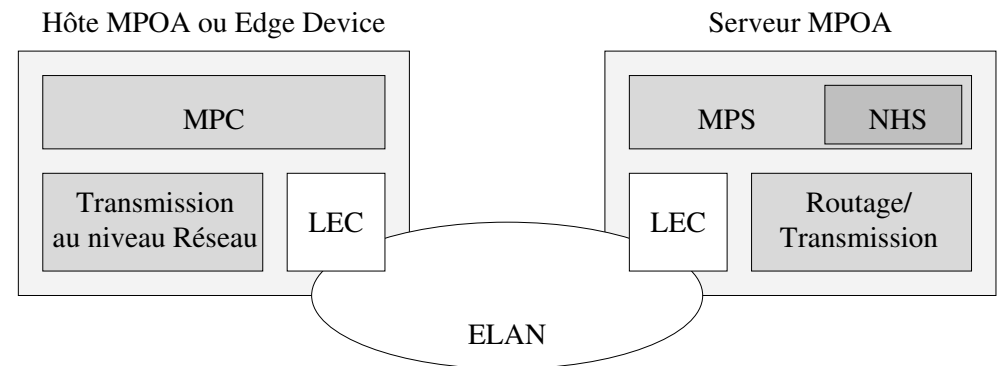
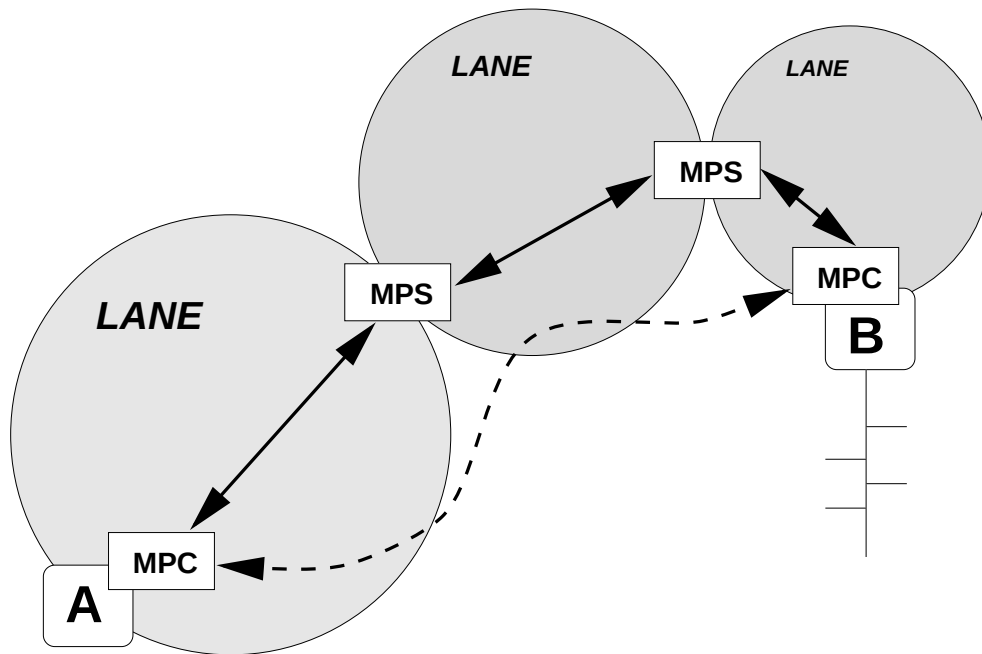
- subdivision en LAN Virtuel (ELAN)



Multi Protocol Over ATM

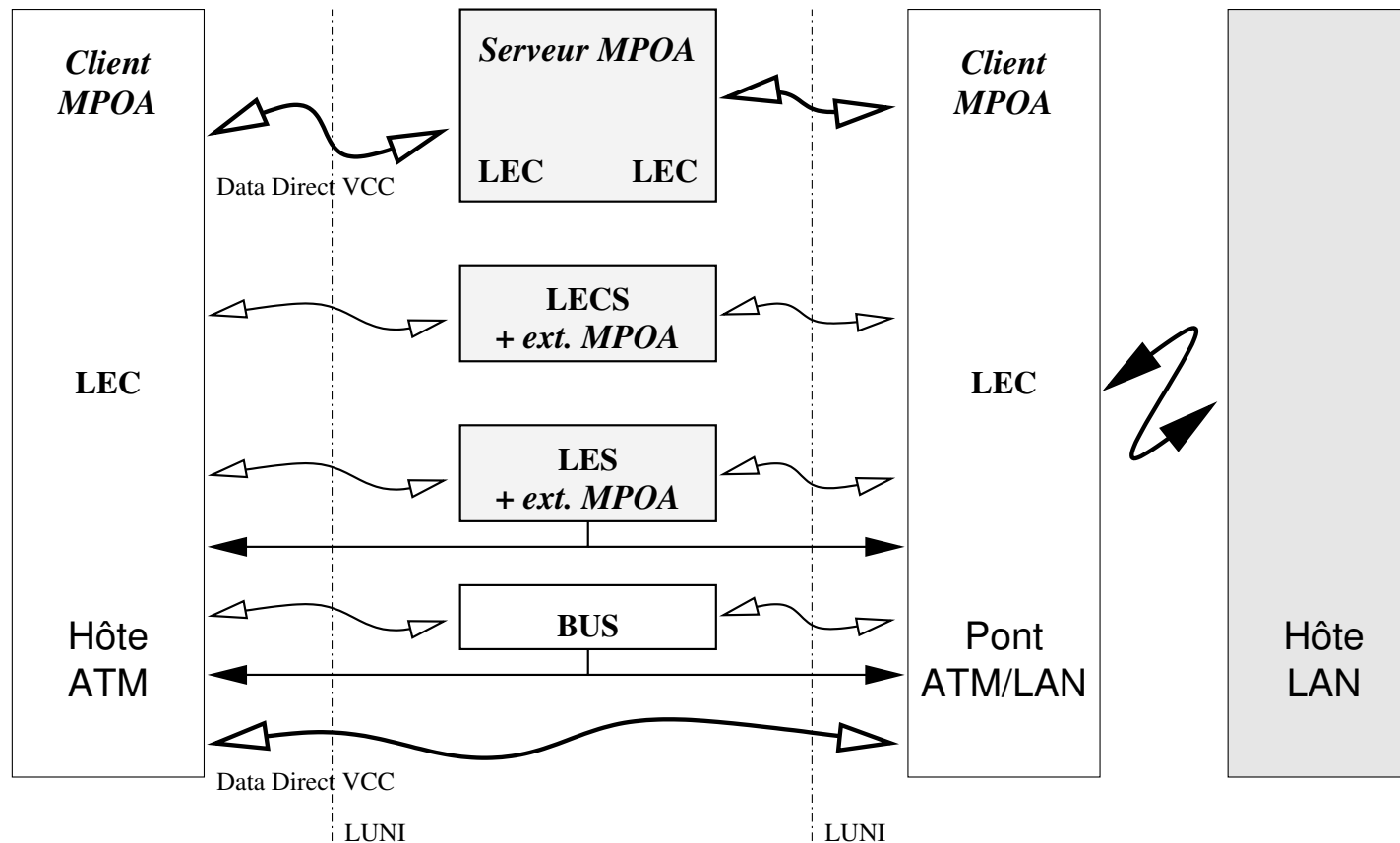
Service de bout-en-bout au niveau ATM

- connectivité niveau 3
- hôtes connectés à ATM ou pontage à travers des *edge devices*
- accès par routage ou par des *shortcuts* ATM avec NHRP.
- routage virtuel (séparation relaying et calcul de route)

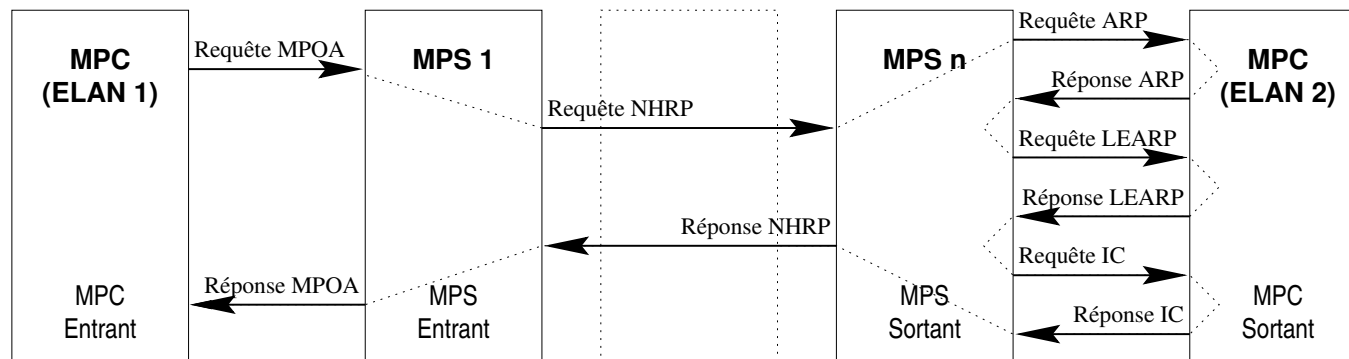
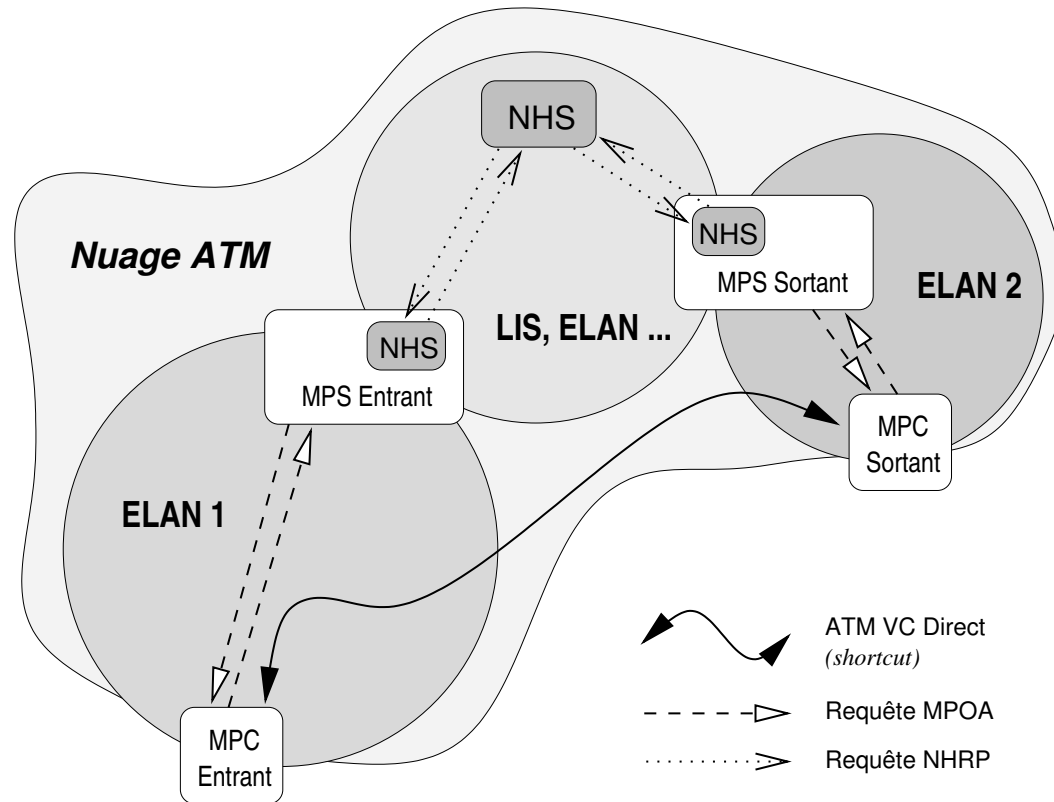


MPOA : Connexions

- serveurs associés à la solution MPOA :
 - ✓ **MPC** (*Multi-Protocol Client*) : intègre le LEC
 - ✓ **MPCS** (*Multi-Protocol Configuration Serveur*) ~ LECS
 - ✓ **MPS** (*Multi-Protocol Server*) ~ ATMARP + NHRP
 - ✓ **BUS** (*Broadcast and Unknown Server*) : gestion de la diffusion



MPOA : Résolution



MPOA : Perspectives

- trop complexe !
 - ✓ gest. périph. Ethernet : $\sim$ 2000 lignes de code C
 - ✓ gest. périph. LANEv1 : $\sim$ 60000 lignes de code C
 - ✓ gest. périph. MPOAv1 : + 200000 lignes de code C...
- ☞ intérêt de l'utilisation d'ATM de bout-en-bout ?
- ☞ simplification ➡ **Commutation IP**

Plan

Architecture Ethernet

Architecture ATM/MPLS

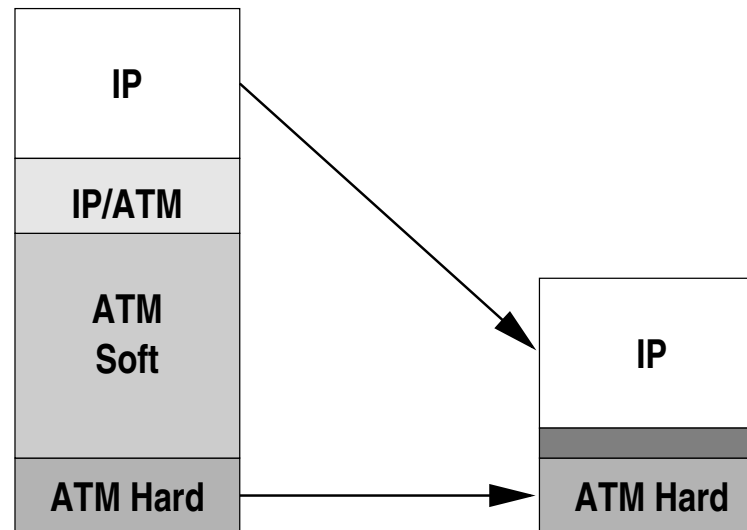
- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- **Commutation IP**
 - ✓ approche orientée flot
 - ✓ approche orientée contrôle
- MPLS

Architecture point-à-point

Simplification

Les approches précédentes conservent l'intégrité de la technologie ATM :

- mode connecté
- découplage des plans de contrôle
- complexité de la signalisation ATM et de l'intégration IP/ATM

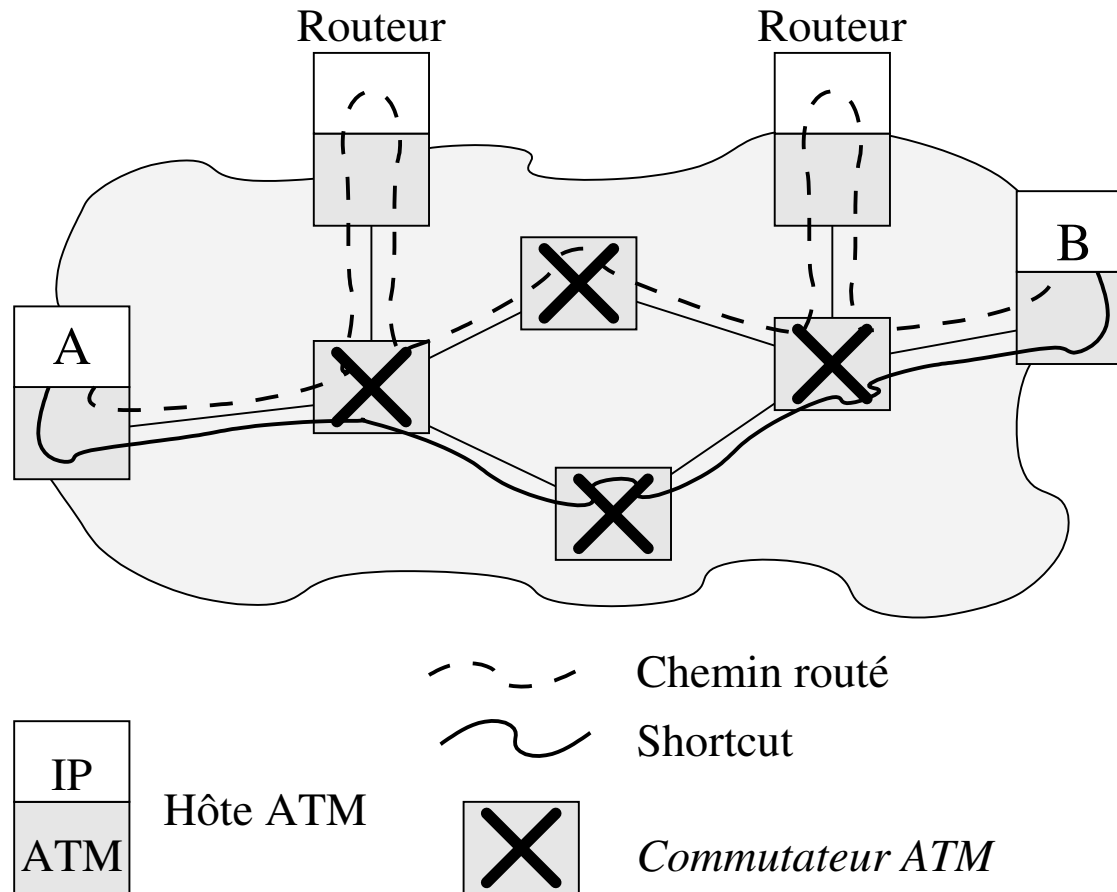


Elimination des redondances

- plan de contrôle d'IP
- plan de transmission des données commutés (ex-ATM)

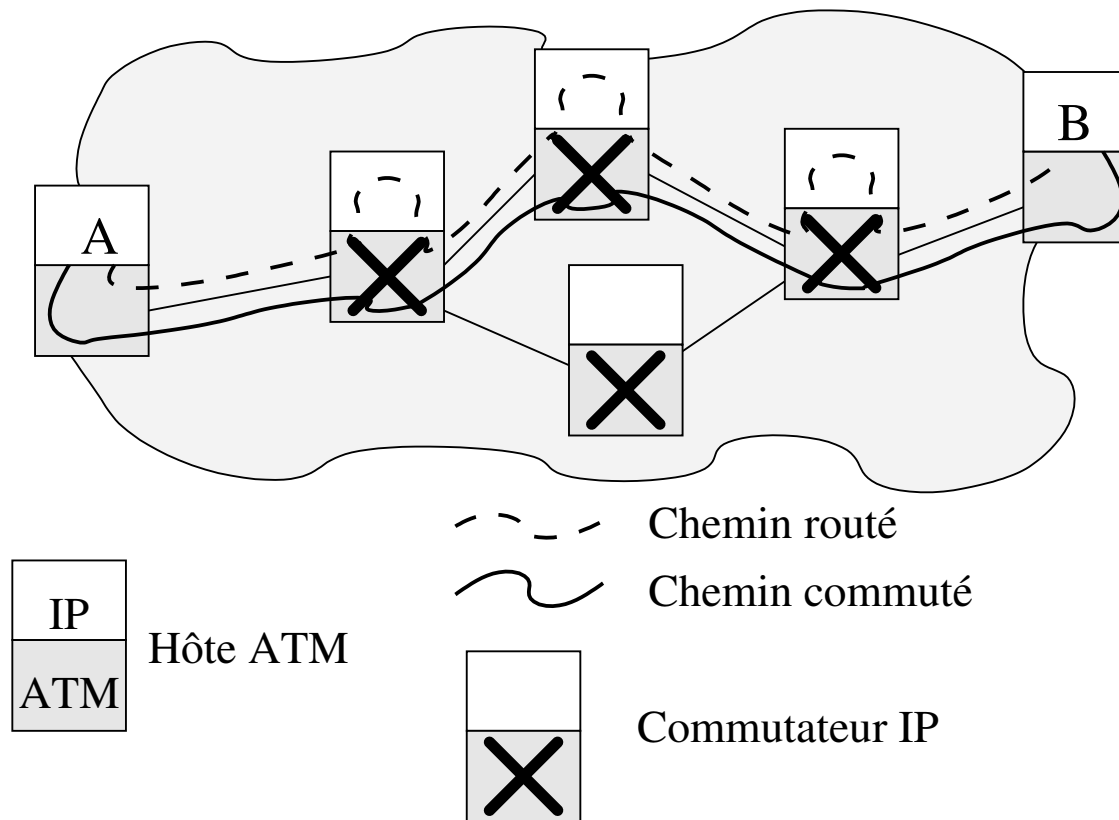
Comparaison de l'établissement d'un VC commuté (1)

CLIP + NHRP :



Comparaison de l'établissement d'un VC commuté (2)

Commutation IP :



Commutation IP : Deux approches

Mise en place du **plan de commutation** :

(☞ ou comment gérer les VCC/VPC sans la signalisation ATM)

- à partir de l'observation du trafic (*flow driven*)
 - ✓ identification dynamique des flux susceptibles d'être commutés
 - ✓ mise en place **implicite** de la commutation
- à l'aide d'une signalisation spécifique (*control driven*)
 - ✓ indication des flux à commuter
 - ✓ mise en place **explicite** de la commutation

Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
 - ✓ **approche orientée flot**
 - ✓ approche orientée contrôle
- MPLS

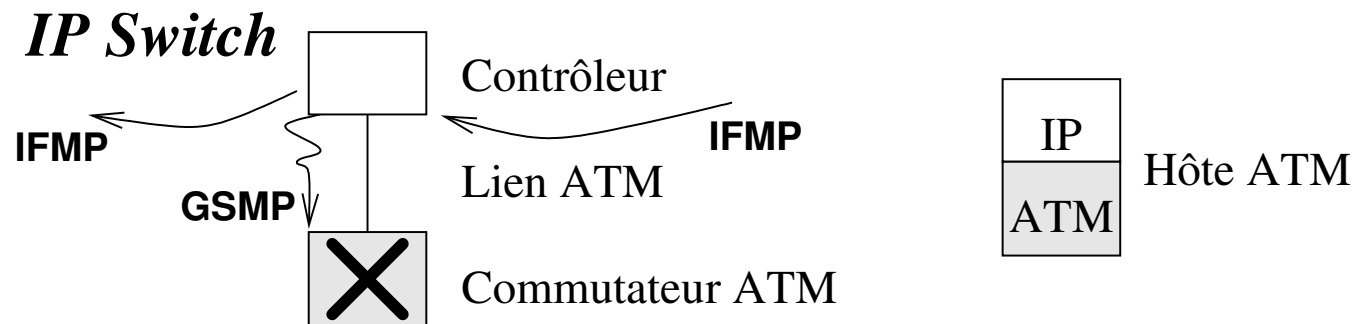
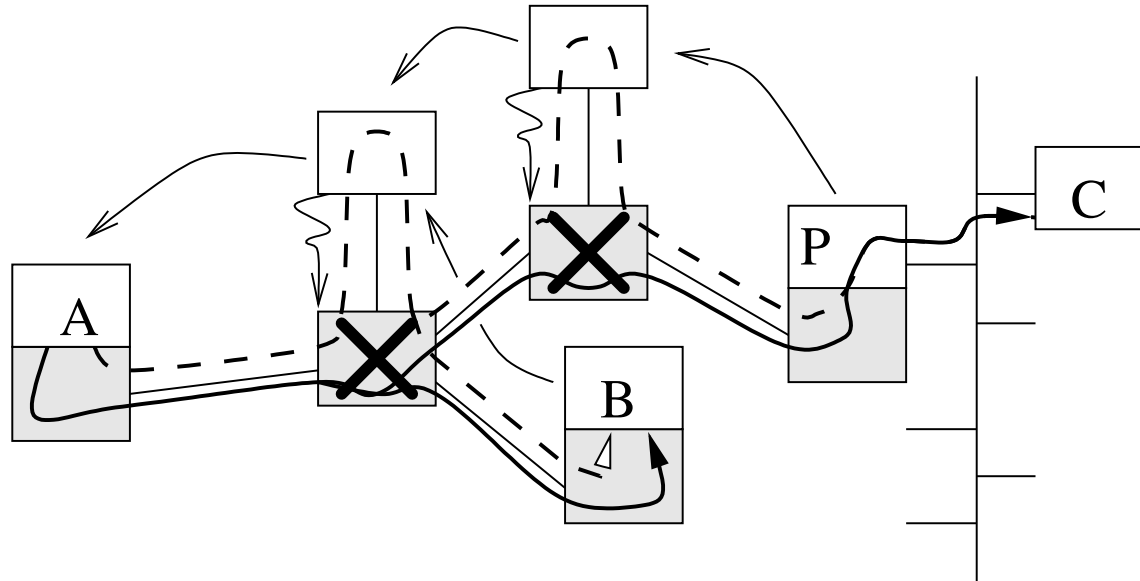
Architecture point-à-point

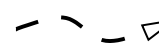
Commutation IP : *Flow driven*

Approche basée sur l'observation du trafic :

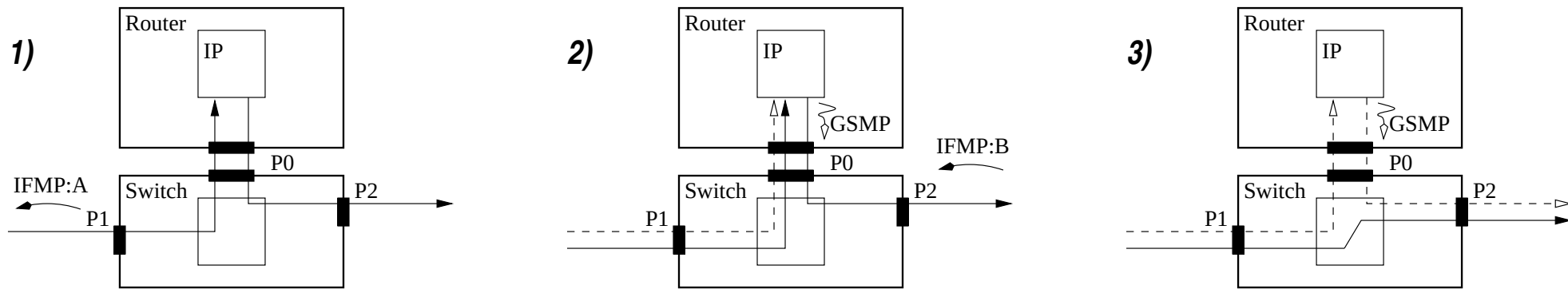
- politique locale **d'identification** des flots de paquets
 - ✓ identification des paquets avec des caractéristiques communes
 - ✓ politique identique dans chaque élément intermédiaire
 - ☞ pas de signalisation globale (simplicité)
- **commutation** des flots "longs"
 - ✓ utilisation de "routeurs commutateurs"
 - ✓ routage initial sur un VC par défaut
 - ✓ VC dédiés pour les flots identifiés
 - ✓ mise en place d'une commutation de proche en proche
 - ☞ indication du VC choisi au commutateur précédent
- efficacité variable (proportion de trafic commuté) :
 - ✓ forte **dépendance à la forme du trafic**
 - ✓ **latence** avant commutation

IP Switching (1)



 Chemin données commutées
 Chemin données routées + signalisation

IP Switching (2)



1. **routage** initial de tous les paquets
 - accès à la partie routeur de l'IPswitch par un VC par défaut
 - identification d'un flot "long" à commuter
 - attribution d'un VC en entrée et envoi à l'IPswitch précédent (IFMP)
2. mise en place de la commutation
 - le flot identifié arrive sur le VC choisi (encore routé)
 - réception de l'attribution de VC de l'IPswitch suivant (IFMP)
 - mise en place de la commutation locale (GSMP)
3. **commutation** du flot identifié

Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- IP sur ATM
- Commutation IP
 - ✓ approche orientée flot
 - ✓ **approche orientée contrôle**
- MPLS

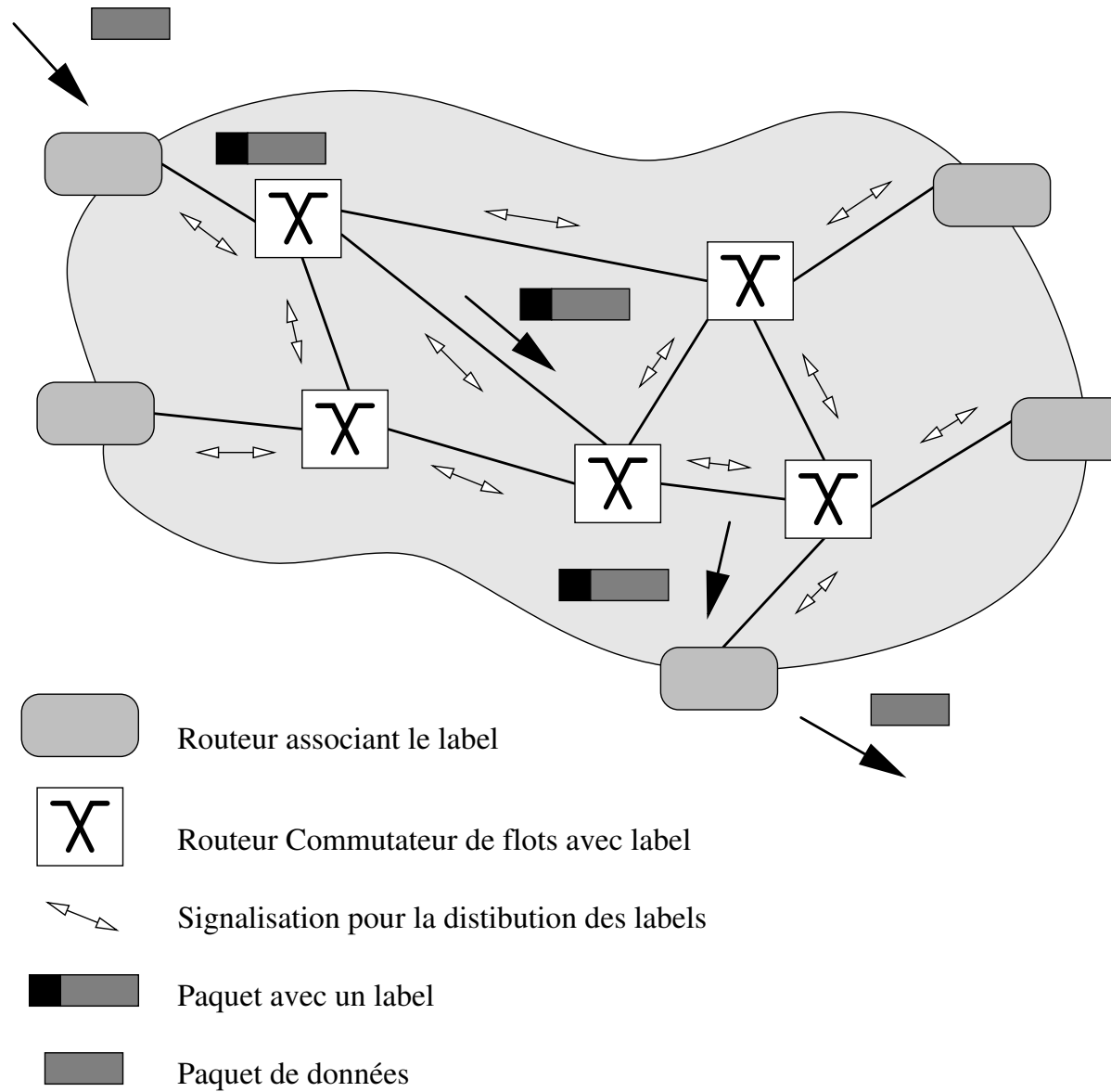
Architecture point-à-point

Commutation IP : *Control driven*

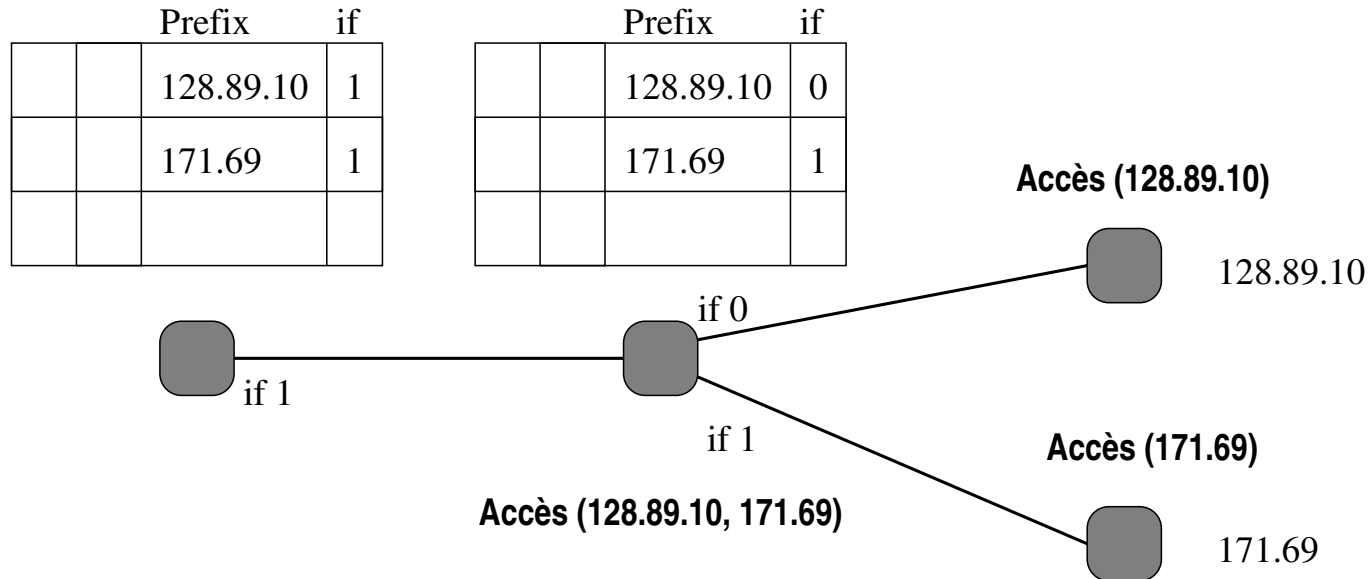
Approche basée sur des protocoles spécifiques :

- politique globale
 - ✓ **protocoles** de création des chemins commutés ($\sim$ VCC ATM)
 - ☞ protocole de niveau réseau (IP)
 - ☞ spécifiques ou extensions de protocoles existants (OSPF, BGP, PIM, RSVP...)
- attribution d'une étiquette pour un type de flots
 - ✓ identification des flots en **bordure**
 - ☞ permet de traverser des étendues sans routage (commutateurs purs ne pouvant traiter le trafic sans étiquettes)
 - ✓ configuration statique
 - ☞ pas de latence
 - ☞ stabilité
 - ✓ granularité variable
 - ✓ non spécifique à ATM

Commutation IP : *Control driven*



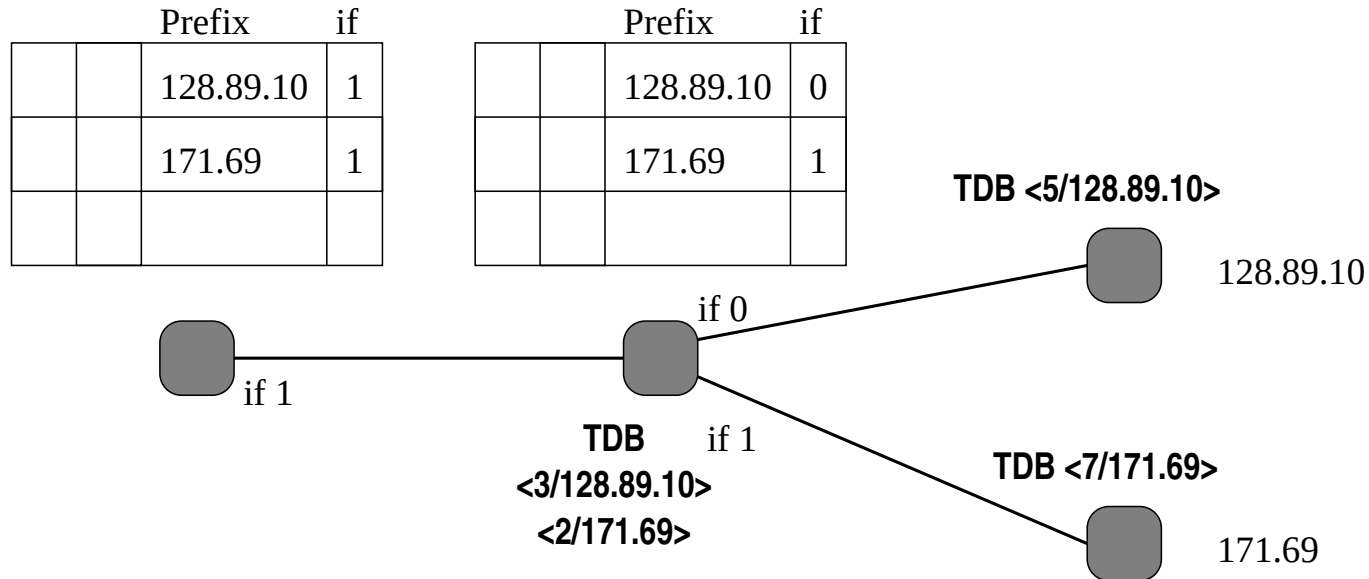
Tag Switching (1)



Distribution des informations de routage

- information sur les préfixes accessibles (protocole de routage)
- construction de la table de routage

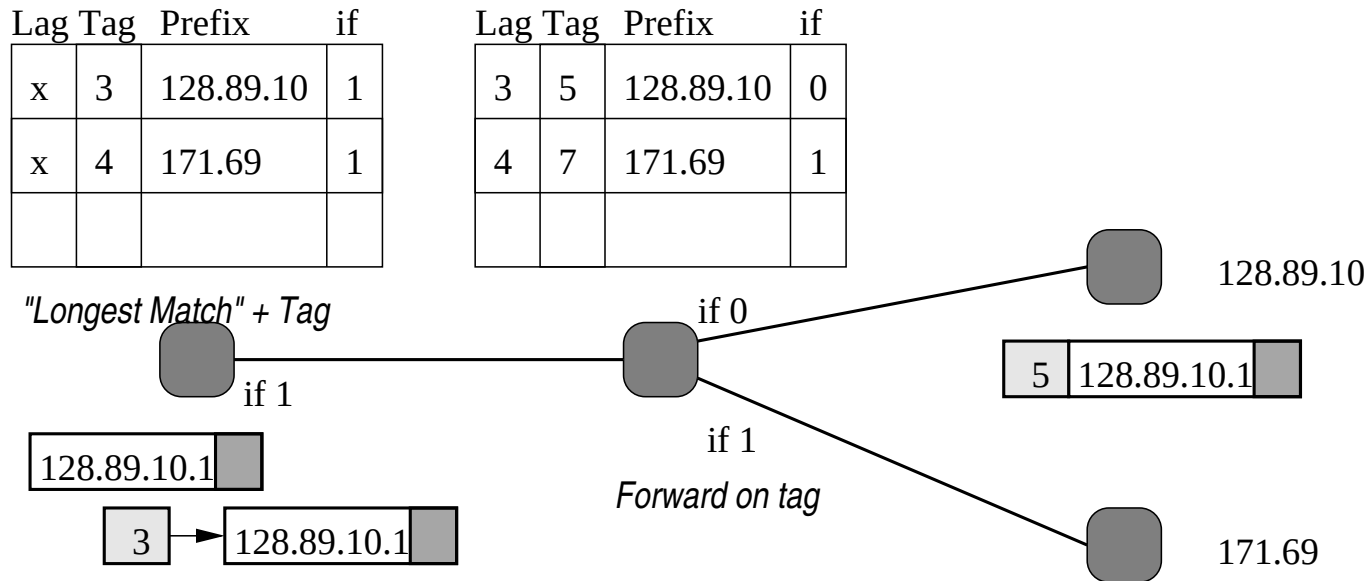
Tag Switching (2)



Distribution des étiquettes (*tags*)

- utilisation d'un protocole spécifique : TDB (*Tag Distribution Protocol*)
- association des étiquettes et des préfixes

Tag Switching (3)



Constitution de la TIB (*Tag Information Base*), étiquetage et commutation

- chaque commutateur contient une TIB, qui à chaque préfixe associe :
 - ✓ une étiquette en entrée (*Lag*)
 - ✓ un port et une (*If* étiquette en sortie (*Tag*))
- les routeurs commutateurs de bordure réalisent l'**étiquetage**
- les éléments internes **commutent**

Plan

Architecture Ethernet

Architecture ATM/MPLS

- NBMA
- B-ISDN
- Technique ATM
- Services ATM
- Modèle de référence ATM
- Commutation IP
- **MPLS**

Architecture point-à-point

Introduction à MPLS

MultiProtocol Label Switching

- milieu 90', nombreuses solutions de commutation IP
 - ✓ intégration IP/ATM
 - ✓ performance de la commutation ATM
 - ☞ création du **MPLS** *working group* par l'IETF (1997)
- fin 90', routeurs aussi performants que les commutateurs mais toujours besoin de :
 - ✓ gestion du trafic (*traffic engineering*)
 - ✓ mise en œuvre de la qualité de service (*QoS support*)
 - ✓ réseaux privés virtuels (VPN : *Virtual Private Networks*)
 - ✓ accès multiprotocole
- ☞ premiers standards **MPLS** (2001)

MPLS : *Traffic Engineering*

Gestion des chemins suivis par les flots

- IP ➡ gestion de trafic **implicite**
 - ✓ contrôle par paquets
 - ✓ distinction des traitements seulement par le préfixe
 - ➡ même chemin entre deux points
 - ✓ actions limitées
 - ➡ suit la politique de routage
 - ➡ au mieux, équilibrage de charge (OSPF)
- MPLS ➡ gestion de trafic **explicite**
 - ✓ contrôle par flots
 - ✓ distinction des traitements par étiquette (**label**)
 - ➡ **plusieurs** chemins entre deux points
 - ✓ protocoles de création de chemins
 - ➡ ajustable selon les prédictions de trafic ➡ **Ingénierie de trafic**

MPLS : *Connection Oriented QoS Support*

Qualité de service de type **circuit-virtuel**

- besoins typiques :
 - ✓ quantité de débit garanti : audio/vidéo en direct
 - ✓ latence et gigue réduite avec débit fixe garanti : téléphonie
 - ✓ répondre à des contrats de trafic sur des flots spécifiques
 - ✓ traiter différemment ses différents clients...
- les réseaux sans connexion (IP) ne peuvent fournir de QoS strictes :
 - ✓ DiffServ : plusieurs types de services au niveau agrégat
 - ✓ IntServ + RSVP : non adapté aux réseaux de taille importante
- l'approche **orientée connexion** utilisée dans les réseaux télécom est nécessaire :
 - ✓ MPLS est une technologie orientée connexion dans un réseau IP

MPLS : *VPN & Multiprotocol*

Support natif des réseaux privés

- VPN $\sim$ chemins MPLS
 - ✓ étiquetage et commutation des flots d'une entreprise
 - ✓ ségrégation efficace des trafics

Support de multiples protocoles

- extension des routeurs IP
 - ✓ coexiste avec les routeurs sans MPLS
- repose sur les différentes technologies support d'IP
- extension des commutateurs
 - ✓ fonctionne sur des réseaux non-IP (ATM...)
 - ✓ s'interface avec les commutateurs non-MPLS
- contrôle avec de nombreux protocoles du monde IP

MPLS : Taxonomie

Composition d'un réseau MPLS :

FEC (*Forwarding Equivalence Class*) : identifie un **flot** globalement

- chaque FEC est associé à un contrat de trafic (QoS)

LSR (*Label Switched Router*) : équipement qui **route** et **commute**

- lors qu'un LSR commute, il ne consulte pas l'entête IP

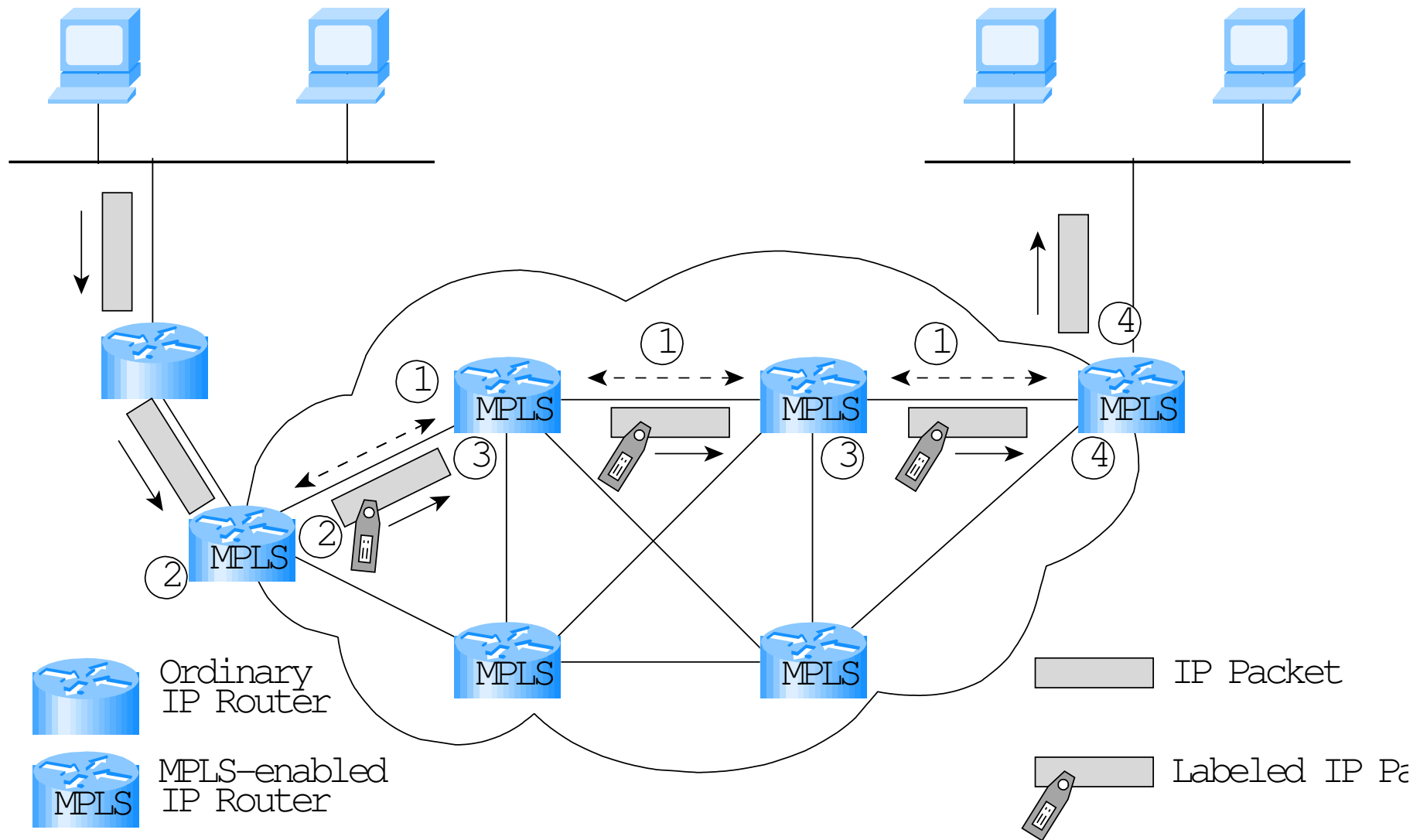
Label : étiquette accolée à un paquet et relatif à sa FEC

- entre deux LSR, les paquets d'une FEC ont le même label

LSP (*Label Switched Path*) : chemin **spécifique** suivit pour une FEC

- MPLS est une technologie **orientée connexion**

MPLS : Exemple



MPLS : Principes de base

- MPLS fonctionne dans un **domaine** de commutateurs contigus
 - ✓ bordé par des routeurs (*ingress* et *egress* LSR)
- relayage performant, basé sur une simple recherche de label
 - ✓ pas d'accès à l'entête IP, pas de *longest prefix match*
- définition d'un traitement particulier par FEC dans chaque LSR
 - ✓ PHB (*Per-Hop Behavior*)
 - ☞ politique de gestion des files d'attentes (priorités, pertes...)

En conséquence, les paquets entre deux LSR de bordure peuvent :

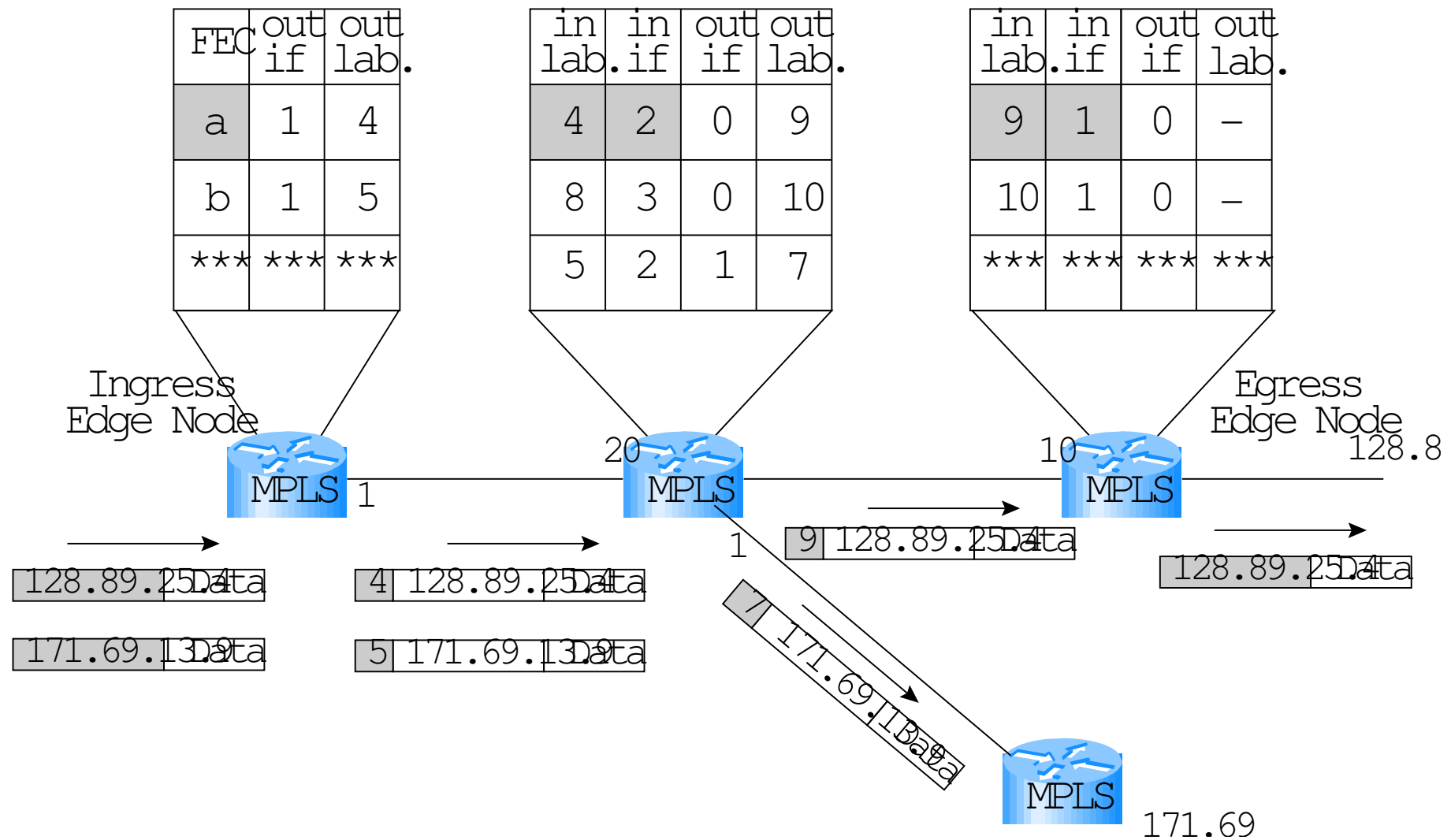
- appartenir à des FEC différents
 - ✓ expérimenter des PHB différents dans chaque LSR
 - ✓ suivre des chemins différents

MPLS : Détermination de la FEC

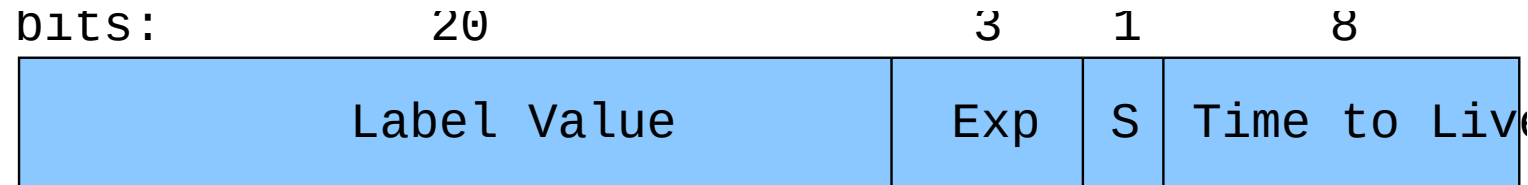
L'administrateur du réseau peut spécifier plusieurs paramètres :

- **adresses** ou **préfixes** source ou destination
- **ports** source ou destination
- champ protocole de l'entête IP
- champ DSCP (*Differentiated Services CodePoint*)
- *flow label* de l'entête IPv6

MPLS : Gestion des labels



MPLS : Format du label



Exp = experimental

S = bottom of stack bit

Champ de 32 bits :

- Label Value : étiquette de 20 bits à signification locale
- Exp : 3 bits expérimentaux
- S : 1 bit positionné à 1 si le label est le plus ancien de la pile
- Time to Live (TTL) : 8 bits utilisé pour indiquer un nombre de saut...

MPLS : Calcul du TTL

L'entête IP n'étant plus lu, comment gérer le TTL ?

- dans l'*ingress* LSR, le champs IP TTL est recopié dans le label
- dans chaque LSR, le TTL composant le label est décrémenté
 - ✓ $TTL > 0$ $\Rightarrow$ la valeur modifiée est placée dans le label le plus récent
 - ✓ $TTL = 0$ $\Rightarrow$ le paquet est éliminé, ou renvoyé à la couche réseau
 - 👉 le traitement IP ordinaire permet l'émission d'un message ICMP
- dans l'*egress* LSR, le TTL du label est recopié dans l'entête IP
 - ✓ ne pas oublier de corriger le *checksum*

MPLS : Pile de labels

Label Stacking

Un paquet peut transporter **plusieurs** labels

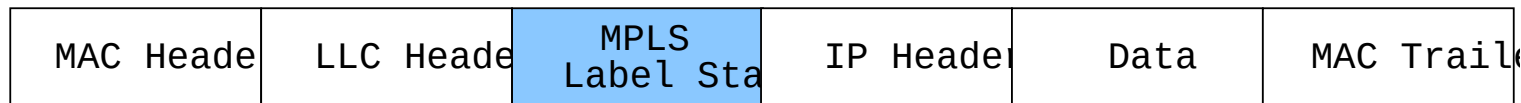
- pile organisée en **LIFO** (*Last In First Out*)
- traitement en cours toujours sur le **dernier** label
- empilement possible dans chaque LSR :
 - ✓ *pop* : ajout d'un label à la pile
 - ✓ *push* : suppression d'un label à la pile
- permet l'agrégation de plusieurs LSPs dans un LSP :
 - ✓ ~ création de tunnel
 - ✓ ~ VCC dans VPC d'ATM
 - ☞ pas de limitation d'empilement

MPLS : Encapsulation

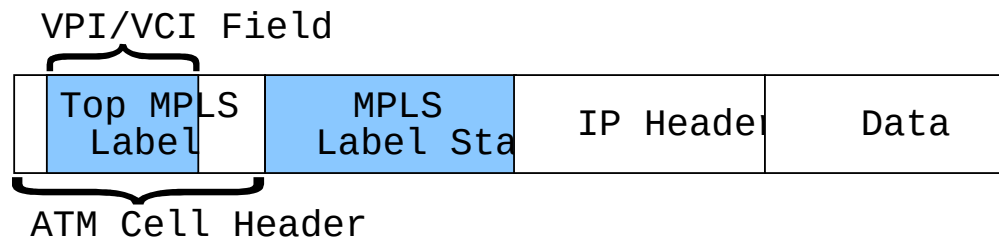
Où se situe la pile de label ?



(a) Data Link Frame



(b) IEEE 802 MAC Frame



(c) ATM Cell

- après les entêtes de la couche liaison
- avant tout entête de couche réseau

MPLS : FEC, LSP et labels

Le trafic est découpé en FEC dans le domaine MPLS :

- les paquets d'une FEC transitent dans le même LSP
 - ✓ les paquets sont identifiés par un label **localement significatif**
 - ✓ dans chaque LSR, un nouveau label est assigné

Contraintes de fonctionnement :

- tout trafic doit être associé à une FEC
 - ✓ association manuelle
 - ✓ association dynamique dans l'egress LSR
- un protocole de **routing** est nécessaire pour déterminer la topologie et construire le LSP
- chaque LSR doit :
 - ✓ connaître les LSP associés aux FEC
 - ✓ assigner des labels
 - ✓ communiquer les labels aux LSR qui lui enverront des paquets

MPLS : Choix du chemin

Deux approches sont possibles :

- *hop-by-hop routing*
 - ✓ utilisation des protocoles de routages usuels (OSPF...)
- *explicit routing*
 - ✓ permet de fixer des routes **différentes** entre deux points
 - ☞ *strict source routing*
 - ☞ *loose source routing*
 - ✓ configuration :
 - ☞ manuelle
 - ☞ dynamique avec des algorithmes de type *constraint-based routing* (intègrent de nombreux métriques : niveau d'utilisation, débits, délais, taux de pertes...)

MPLS : Protocoles de distribution

Après le choix du chemin, mise en place du LSR

- chaque LSR doit :
 - ✓ assigner un label au LSP
 - ✓ communiquer ce label à tout LSR précédent (*upstream*)
 - ✓ recevoir la valeur du label vers le prochain LSR (*downstream*)
- ☞ deux protocoles ont été définis :
- LDP** (*Label Distribution Protocol*)
- RSVP** (*resource ReSerVation Protocol*)

MPLS : Perspectives

- GMPLS (*Generalized MPLS*)
 - ✓ MPLS a tout les niveaux de commutation
 - ☞ *Lambda switching*
 - ☞ *Optical switching*
 - ☞ *Virtual Circuit switching*
 - ☞ *Virtual LAN switching*
- ...

Fin

Document réalisé avec $\text{\LaTeX}$.

Généré le 23 novembre 2004.

Classe de document foils.

Dessins réalisés avec xfig.

Olivier Fourmaux, olivier.fourmaux@lip6.fr

<http://www-rp.lip6.fr/~fourmaux>

Ce document est disponible en postscript compressé avec gzip à

<http://www-rp.lip6.fr/~fourmaux/res/res4c8c-.pdf>