

IP (IETF) et ATM

- ◆ Particularités de IP
- ◆ Transport : solutions actuelles
- ◆ RFC1483
- ◆ RFC1577
- ◆ Autres RFCs (1626, 1755, Ipsilon, MPLS)
- ◆ Problèmes
- ◆ Evolutions
 - NHRP
 - MARS





Particularités de IP

- ◆ **ATM'96 à San Jose : les applications IP sont et seront les principales applications utilisées sur ATM**
- ◆ **Mode non connecté IP / mode connecté ATM**
 - Bonne gestion des ouverture-fermeture de CV ATM
- ◆ **Fonction ARP : @ IP ---> @ ATM**
- ◆ **Routage**
 - Routeur : passage obligé d'un sous-réseau IP (LIS) à un autre
 - Next hop uniquement
- ◆ **Multicast**
- ◆ **MTU**



Transport IP / ATM solutions actuelles

- ◆ **ATM = réseau de LS**
 - 1 VP ATM entre chaque couple de routeurs IP
- ◆ **ATM = réseau Ethernet**
 - LANE
 - Transparent pour IP
- ◆ **RFC1483 et 1577**



Transport IP / ATM

RFC 1483

- ◆ Multiprotocol Encapsulation over ATM Adaptation Layer 5 (Juillet 93)
- ◆ Décrit deux méthodes pour transporter du trafic en mode non connecté sur un réseau ATM
- ◆ Interconnexion niveau 2 ou 3 avec AAL5
 - Protocol Data Units (PDU) "routés"
 - » niveau 3 comme IP
 - Protocol Data Units (PDU) "pontés"
 - » niveau 2 comme Ethernet



Transport IP / ATM RFC 1483

2 Possibilités d'encapsulation

- ◆ **VC Based Multiplexing : pas d'en-tête**
 - Si besoin de multiplexer plusieurs protocoles, il faut ouvrir un VC par protocole
- ◆ **LLC/SNAP**
 - En-tête ajouté en début de PDU (cf AAL5)
 - » IEEE802.2 LLC header (3 octets)
 - » SubNetwork Attachment Point (SNAP), IEEE802.1
 - IP : AA-AA-03-00-00-00-08-00
 - » LLC : AA-AA-03 (3 octets)
 - » SNAP : 00-00-00-08-00 (5 octets)
 - Ethernet : AA-AA-03-00-80-C2-00-0(1 ou 0)-07-00-00
 - FDDI : AA-AA-03-00-80- C2-00-0(4 ou 0)-0A-00-00
- ◆ **Choix méthode**
 - Configuration manuelle (PVC)
 - Protocole de signalisation (SVC)



Transport IP / ATM

RFC 1577

- ◆ Classical IP and ARP over ATM (janv 94)
- ◆ Comment faire passer du trafic IP sur ATM en respectant les principes IP (sans modification fondamentale)
- ◆ Limité au protocole IP
- ◆ Transmission sur AAL5 :
 - de datagrammes IP
 - de requêtes de résolution d'adresses ATM (ATMARP protocole proche de ARP, RFC 826)
 - de requêtes de résolution inverse (InATMARP)
 - de réponses aux requêtes ATMARP et InATMARP
- ◆ Pas de support de broadcast ou multicast IP



Transport IP / ATM

RFC 1577

- ◆ Les adresses ATM doivent être conformes à celles définies dans UNI 3.0 (voir signalisation:adresses)
- ◆ Le réseau ATM autorise les SVC suivant UNI3.0
- ◆ Un MTU (Maximum transmission Unit) unique pour tous les VC dans un Logical IP Subnetwork
- ◆ Encapsulation LLC/SNAP
 - ARP : AA-AA-03-00-00-00-08-06



Transport IP / ATM

RFC 1577

- ◆ **Un réseau ATM = 1 sous-réseau IP**
 - LIS : Logical IP Subnetwork
 - Comprend plusieurs stations et routeurs
 - Communication entre les stations en ATM "directe" (chaque VC connecte directement 2 équipements du même LIS)
- ◆ **Un réseau ATM = n sous-réseaux IP (n LIS)**
 - Le trafic entre les sous-réseaux est routé par au moins un routeur
 - Ce routeur IP :
 - » peut avoir plusieurs coupleurs ATM
 - ou
 - » un coupleur avec plusieurs @ ATM
 - Entre 2 stations dans des LIS différents
 - » 2 VC ATM : station 1-routeur et routeur-station2



- ◆ **Un seul serveur ARP dans chaque LIS**
 - Configuré pour se reconnaître serveur ARP du LIS
 - Avec une adresse IP par sous-réseau logique qu'il gère
- ◆ **Le serveur n'établit jamais de connexions**
- ◆ **C'est le client qui initialise la procédure**
 - utilisation d'un VC point à point
 - chaque client est configuré avec l'adresse ATM du serveur ARP
- ◆ **Utilisation de datagrammes IP similaires au ARP classique**



Transport IP / ATM

RFC 1577

Modes opératoires du serveur

- ◆ **Enregistrement d'un client**
 - Etablissement d'une connexion ATM (VC) par le client
 - Envoi de la requête InATMARF au client
 - Réponse du client à la requête InATMARF
 - Ajout ou mise à jour des informations de la table ARP du serveur

- ◆ **Réponse du serveur à une requête ATMARF**
 - Si l'information est dans la table, réponse positive
 - Sinon envoie d'une réponse négative à la requête

- ◆ **Validité d'information dans sa table ARP : 20 mn**
 - Quand un client se manifeste (ATMARF ...) note que le client est encore en vie et remet à zéro ce compteur
 - Si VC ouvert avec client envoie un InATMARF avant suppression de l'entrée dans la table



UREC

Transport IP / ATM

RFC 1577

Modes opératoires du client

- ◆ A sa connexion sur le réseau : ouverture de la connexion VC vers le serveur ARP
- ◆ Réponse aux requêtes InATMARP
- ◆ Rafraichissement de l'info dans le serveur régulièrement (< 20 mn)
- ◆ Envoi de paquets ARP_REQUEST au serveur quand besoin
- ◆ Gestion de sa table ARP
 - validité d'une information : 15 mn



UREC

Transport IP / ATM

RFC 1626

- ◆ Default IP MTU (Maximum Transmission Unit) for use over ATM AAL5 (Mai 94)
- ◆ MTU défaut = 9180 octets
- ◆ La taille du MTU peut-être négociée si on utilise les SVC

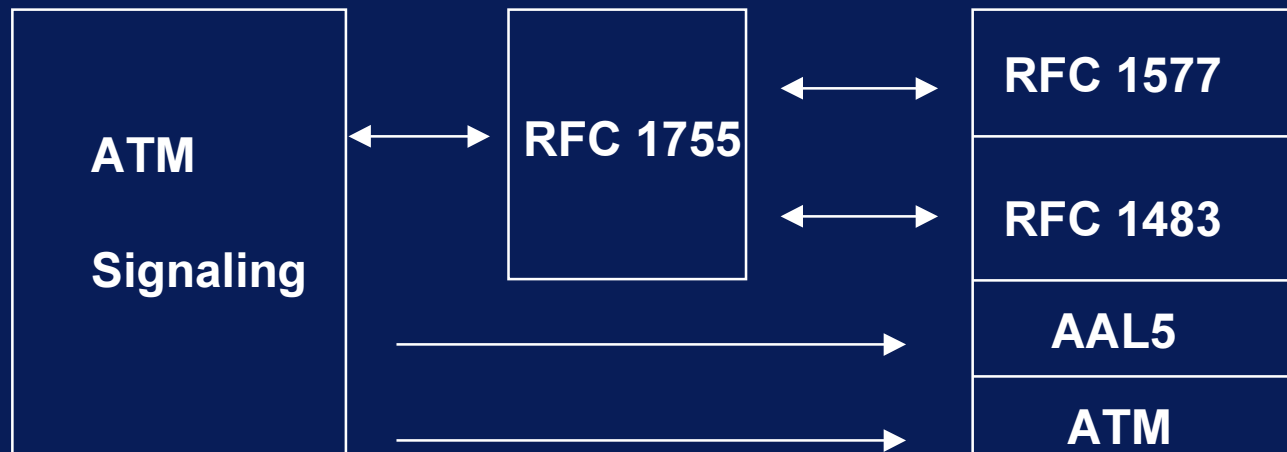


UREC

Transport IP / ATM

RFC 1755

- ◆ ATM Signaling Support for IP over ATM (Fév 95)
- ◆ Définit comment utiliser UNI 3.1 (échanges au moment de l'ouverture, fermeture d'une connexion ...) pour "Classical IP over ATM" (RFC1577) et RFC1483
- ◆ AAL5 et encapsulation LLC/SNAP





Transport IP / ATM

RFC 1755

Gestion des VC

- ◆ Une station peut ouvrir plusieurs VC avec une autre station IP (une station doit supporter cette possibilité)
- ◆ Une station peut être configurée pour n'accepter qu'un VC ouvert avec une autre. Dans ce cas, si les 2 stations ouvrent le VC "en même temps" (collision) : rejet des ouvertures et ré-essai après un temps d'attente aléatoire (algorithme similaire à Ethernet)
- ◆ Un VC peut être partagé avec d'autres protocoles
- ◆ N'importe laquelle des stations peut fermer le VC
- ◆ Un VC doit rester ouvert au moins 1 mn (recommandé)
- ◆ Un VC doit être fermé après inactivité de 20 mn (recommandé en private, obligatoire en public UNI)



Transport IP / ATM Signalisation

RFC 1755

Quelques IEs (Information Elements) lors de l'ouverture du VC :

- ◆ **AAL_parameters :**
 - aal_type : 5
- ◆ **Broadband Low Layer Information :**
 - layer_2 : 12 (lan_llc)
 - layer_3 : 204 (IP) si le VC ne sert qu'à IP
- ◆ **Traffic_descriptor (recommandé, si le réseau ne le permet pas d'autres solutions sont décrites)**
 - fwd_peak_cell_rate : débit de la liaison
 - bkw_peak_cell_rate : débit de la liaison
 - best_effort_indication
- ◆ **QoS_parameter (recommandé) : 0 (unspecified)**

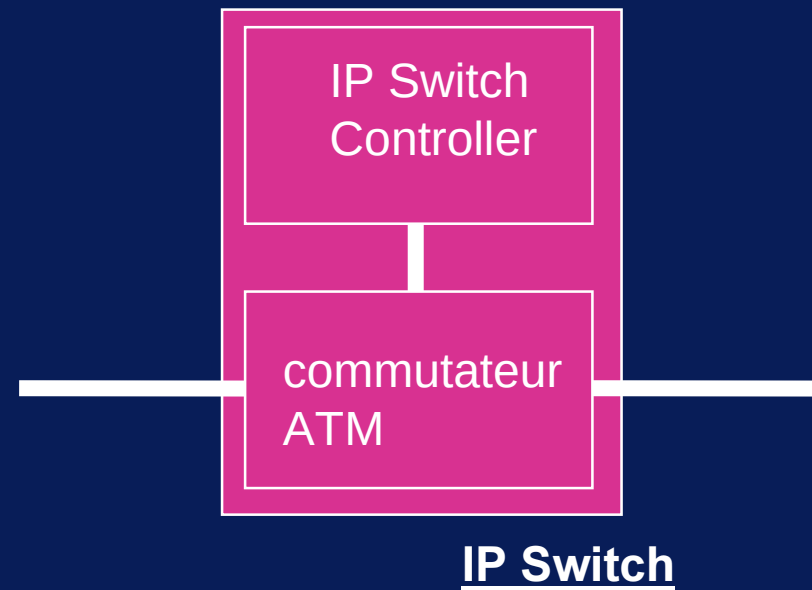


- ◆ **Adresses routeur IP avec une seule carte ATM :**
 - NSAP : utiliser le champs SEL
 - E164 : utiliser plusieurs adresses E164
- ◆ **Si rejet de l'ouverture de VC (avec code d'erreur UNI 3.1 "unassigned number", "no route to destination", "user busy", ...), une erreur remonte à la couche IP : Host Unreachable (si appelé est une station) ou Network Unreachable (si appelé est un routeur)**
- ◆ **Si erreur UNI 3.1 "call rejected", "user rejects call with CLIB", ... la station peut réessayer d'ouvrir la connexion**



IP Switching d'Ipsilon

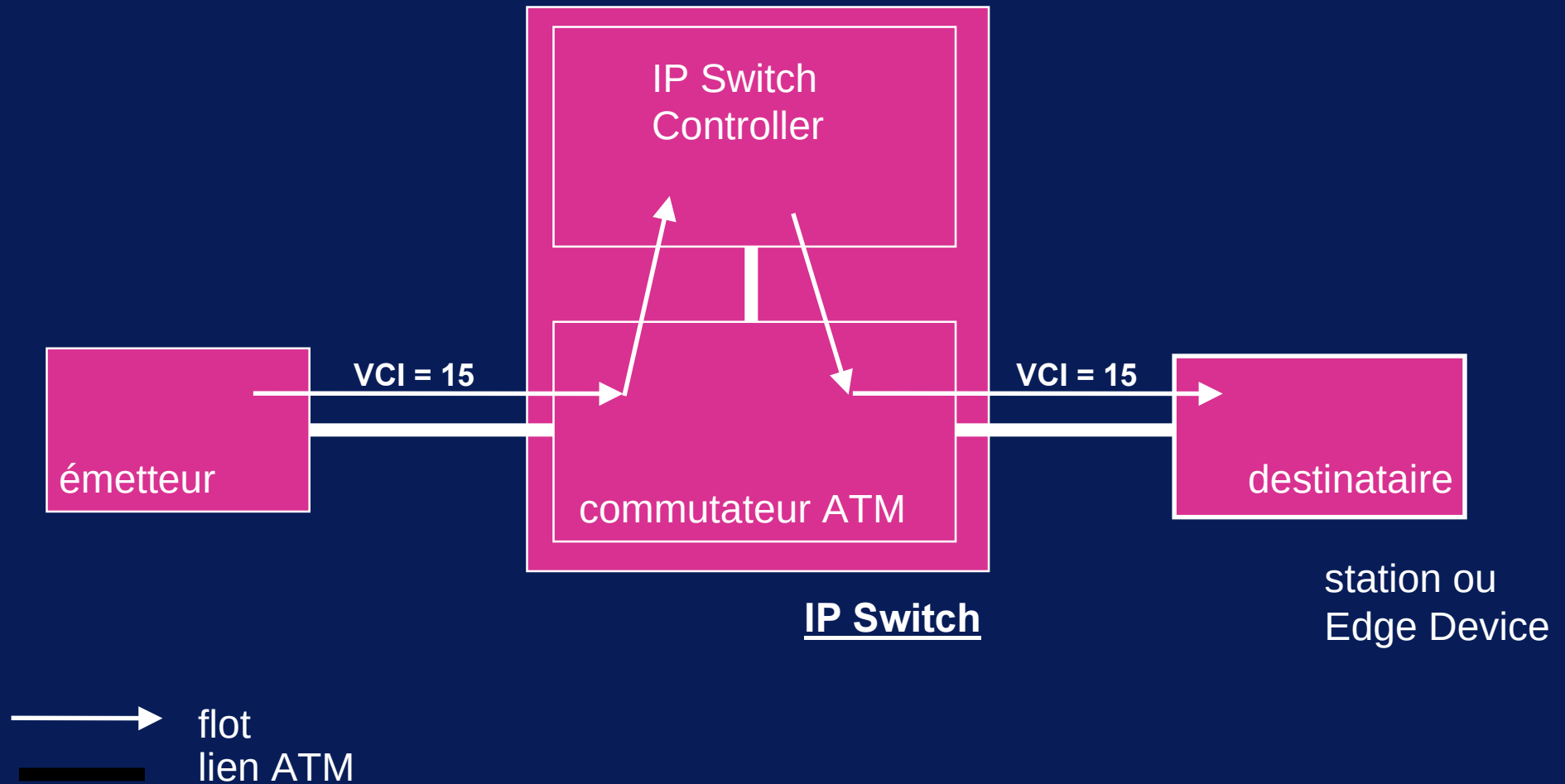
- ◆ association d'un routeur IP et d'un commutateur ATM
- ◆ détection du trafic "flot" et du trafic "à courte durée de vie"
- ◆ trafic "flot" commuté
- ◆ trafic "courte vie" routé





UREC

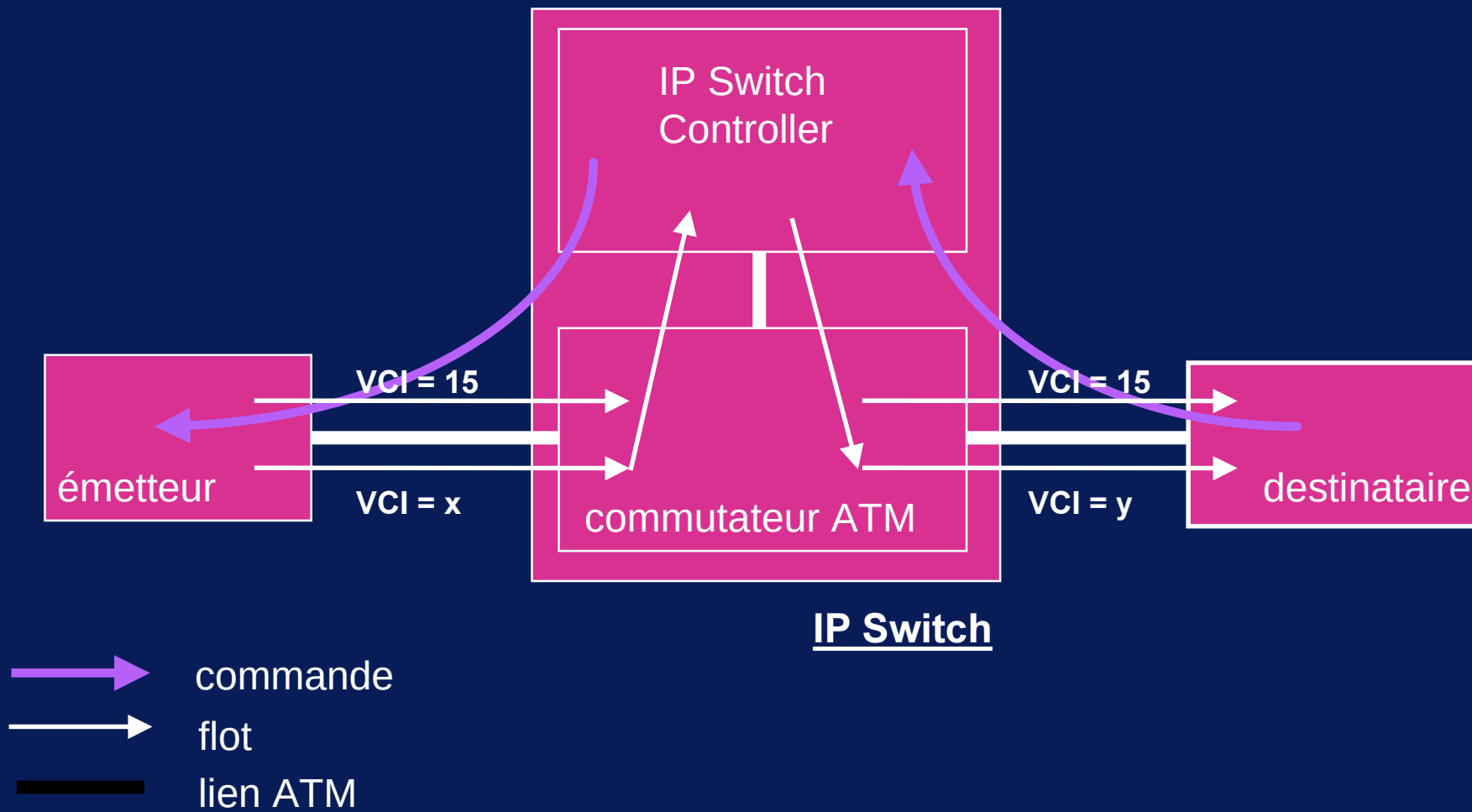
IP Switching d'Ipsilon





UREC

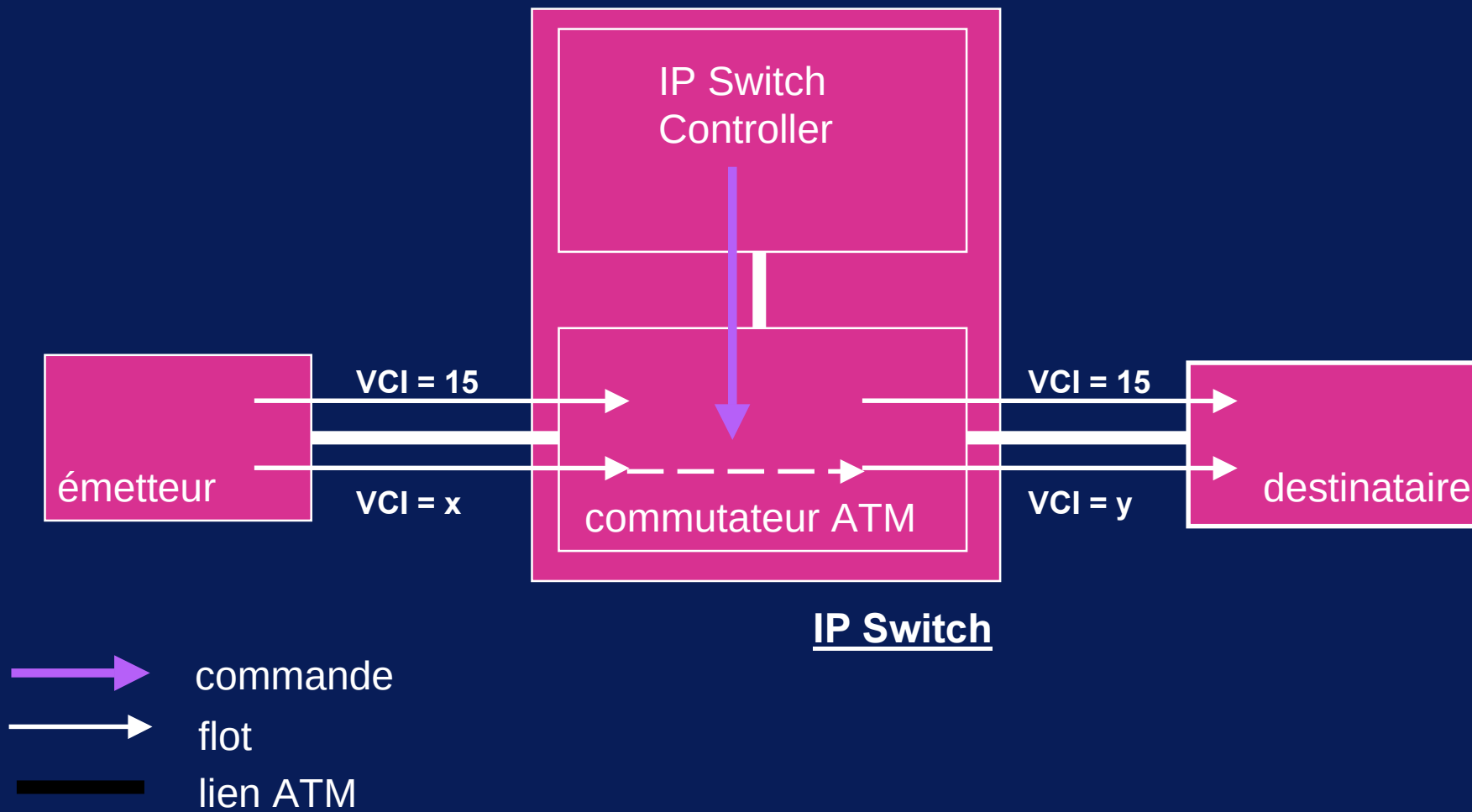
IP Switching d'Ipsilon





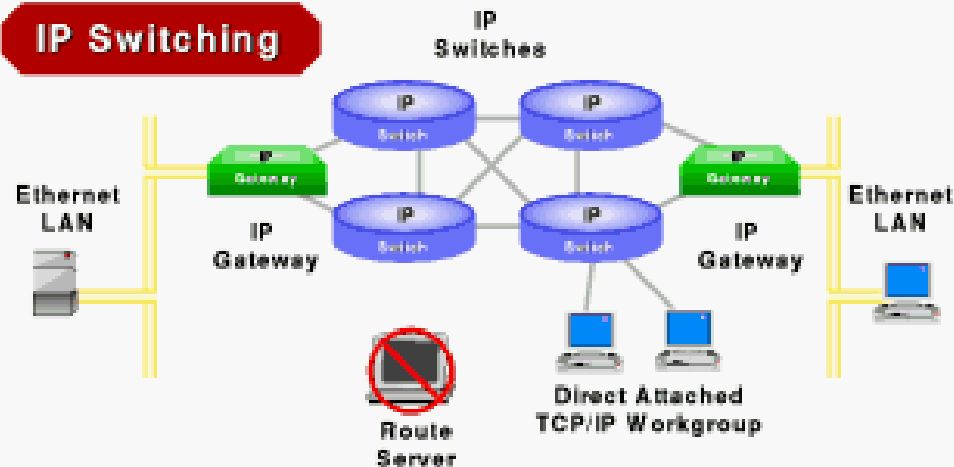
UREC

IP Switching d'Ipsilon



Avoiding the Unnecessary Complexity of the ATM Forum MPOA

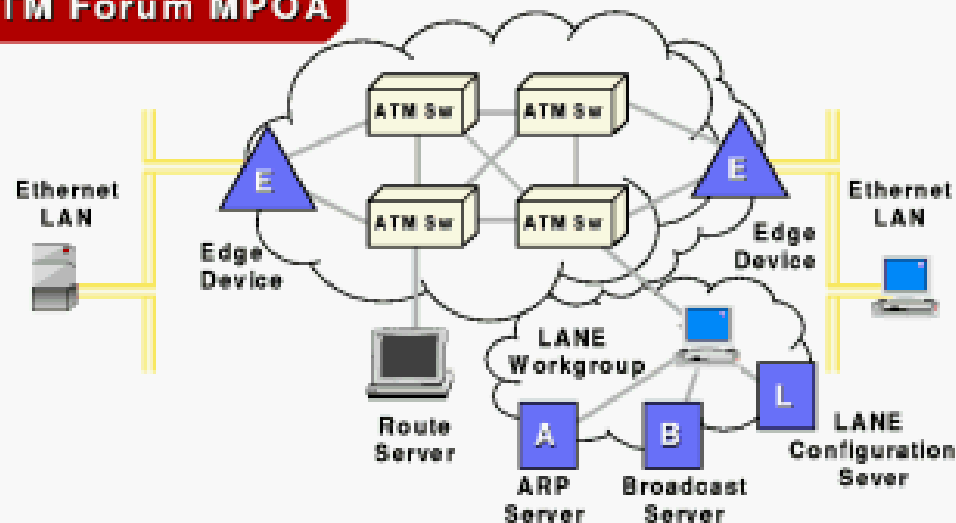
IP Switching



Ipsilon Advantages

- No ATM cloud, all IP protocols
- No duplication of routing functionality
- No centralized route server
- No complex, untested protocols
- Complete end-to-end IP mgt
- MPOA implementations are vendor-dependent
- Scalable

ATM Forum MPOA





Autres RFCs : Ipsilon

- ◆ RFCs d'information
- ◆ Produit Ipsilon : IP Swith (www.ipsilon.com)
 - Routeur qui devient switch ATM / Flot IP
 - Performance routage IP x 5 et prix bas
 - Ipsilon : standards ATM Forum trop complexes
- ◆ RFC1953 : Ipsilon Flow Management Protocol Specification for IPv4
 - Noeud ATM ---> Noeud ATM adjacent : attache un flot IP a un VPI-VCI, en supprimant les en-tête IP-TCP
 - Flot : datagrammes ayant les mêmes adresses d'origine et de destination, idem + mêmes numéros de port origine et destination.
- ◆ RFC1954 : Transmission of Flow Labelled IPv4 on ATM Data Links Ipsilon
 - Comment encapsuler la partie data des datagrammes IP d'un flot dans les PDU ATM



Autres RFCs : Ipsilon

- ◆ **RFC1987 : Ipsilon's General Switch Management Protocol Specification (GSMP)**
- ◆ **Messages ALL-5 LLC/SNAP VPI=0, VCI=15**
- ◆ **Station de contrôle ---> Switches**
 - Administration des connexions entre les ports du switch : établir, détruire, modifier, vérifier
 - Administration des ports : mettre en service, arrêter, mettre en boucle, réinitialiser
 - Statistiques : activité récente des VC, des ports, compteur d'erreurs, ...
 - Configuration : pour obtenir les caractéristiques et fonctionnalités du switch (hard, version du soft, état des liaisons, numéros de VP et VC utilisables, ...)
- ◆ **Switch ---> Station de contrôle**
 - alarme : port up/down, ...



MultiProtocol Label Switching

- ◆ Cette solution consiste à associer une route de niveau 3 à un chemin établi au niveau 2 : cette association est matérialisée par un label ou tag associé au paquet à transmettre
- ◆ Un travail de création de tables associant les labels aux chemins de niveau 2 est effectué à priori dans tout le réseau de routeurs et commutateurs
- ◆ Le routeur d'entrée du réseau affecte le label au paquet en fonction de la route qu'il doit emprunter
- ◆ Les routeurs et commutateurs traversés ensuite ne font que commuter en fonction du label trouvé dans le paquet, cette fonction pouvant être réalisée par des processeurs spécialisés



MultiProtocol Label Switching

◆ Labels :

- le label peut être mis dans un header rajouté à la trame
- il peut être constitué d'un numéro de circuit virtuel ATM
- il peut occuper un champ particulier de la trame (IPv6)

◆ Chemins de niveau 2 :

- un chemin par route depuis chaque routeur (Tag Switching de Cisco)
- un chemin par routeur de sortie du réseau, commun à toutes les routes accessibles par ce routeur (ARIS d'IBM)



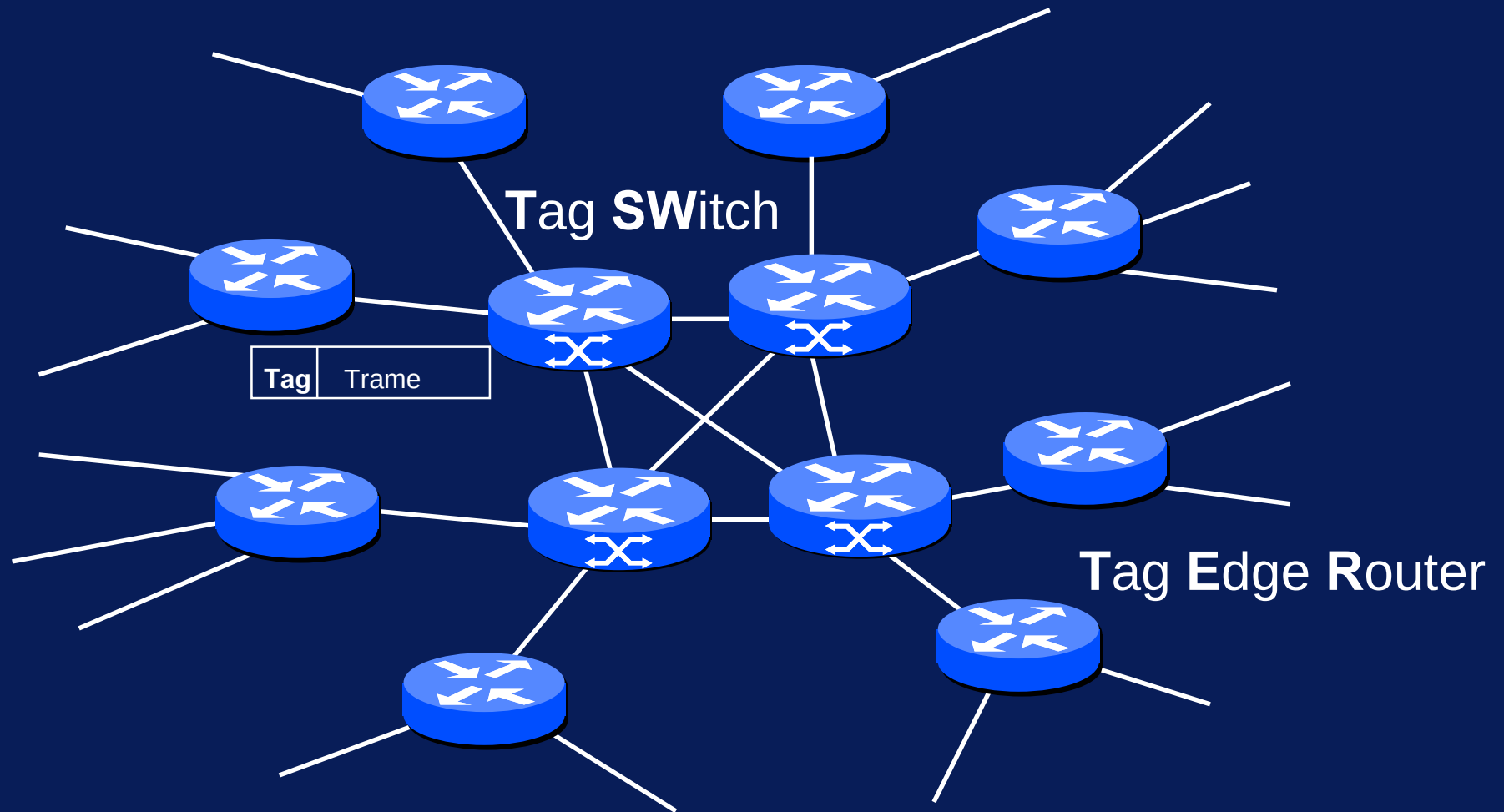
Tag Switching de Cisco

- ◆ Tag Switching propose de rajouter des étiquettes (Tags) à tous les paquets devant traverser un réseau
- ◆ des tables de routage sont construites dans les Tag Edge routeurs et dans les Tag Switches, elles associent un tag à chaque route
- ◆ les Tag edge routeurs analysent l'en-tête des paquets et leur collent un tag en fonction de leur route
- ◆ à chaque saut entre deux Tag Switches est associé un tag, les Tag Switches commutent les paquets en fonction du tag "entrant" et remplacent celui-ci par le tag "sortant"



UREC

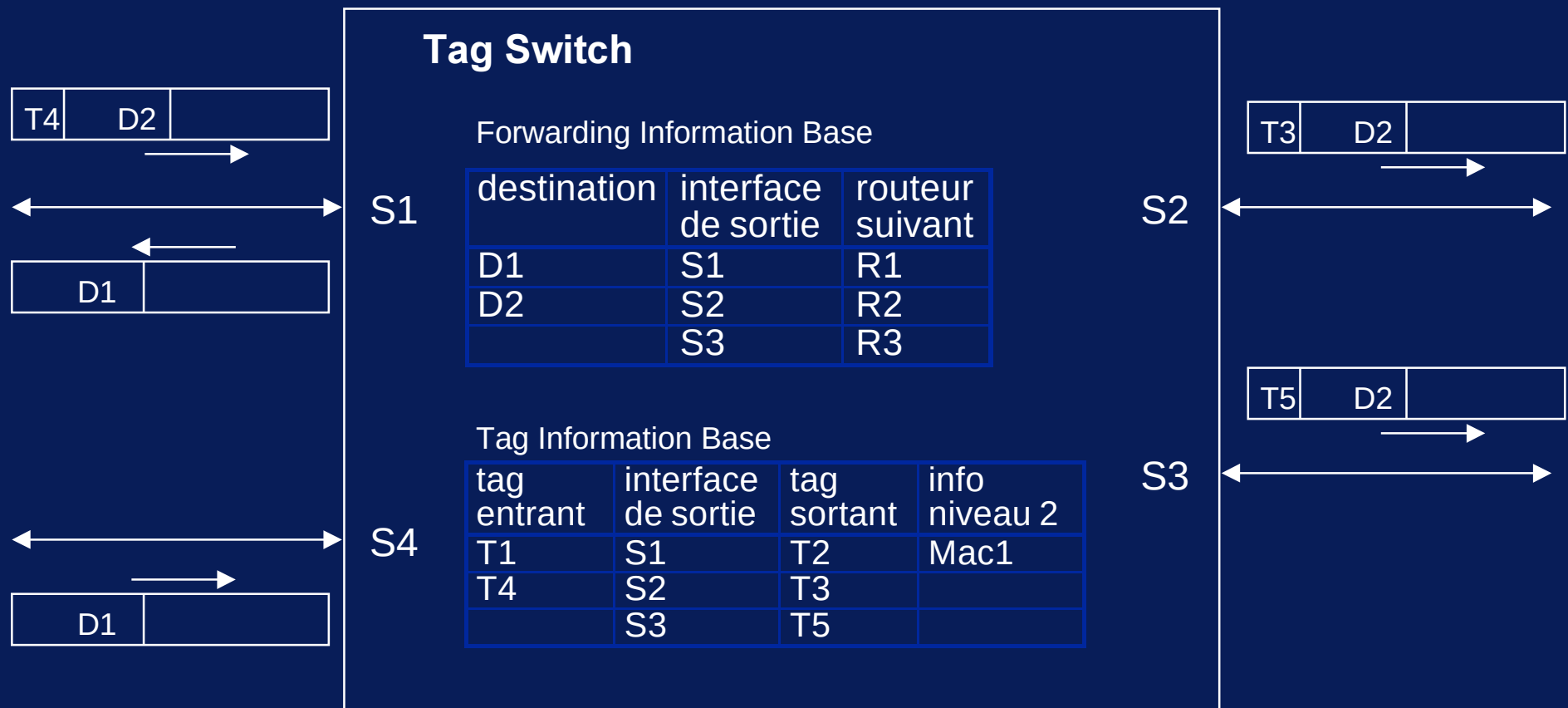
Tag Switching de Cisco





UREC

Tag Switching de Cisco





Tag Switching de Cisco

◆ où met-on le tag ?

- dans un header supplémentaire entre H2 et H3
- dans le header 2 (VPI/VCI par ex)
- dans le header 3 (Flow label de IPv6 par ex)

◆ à quoi peut être associé le tag ?

- à un groupe de routes (préfixe)
- à un flot
- à un arbre multicast



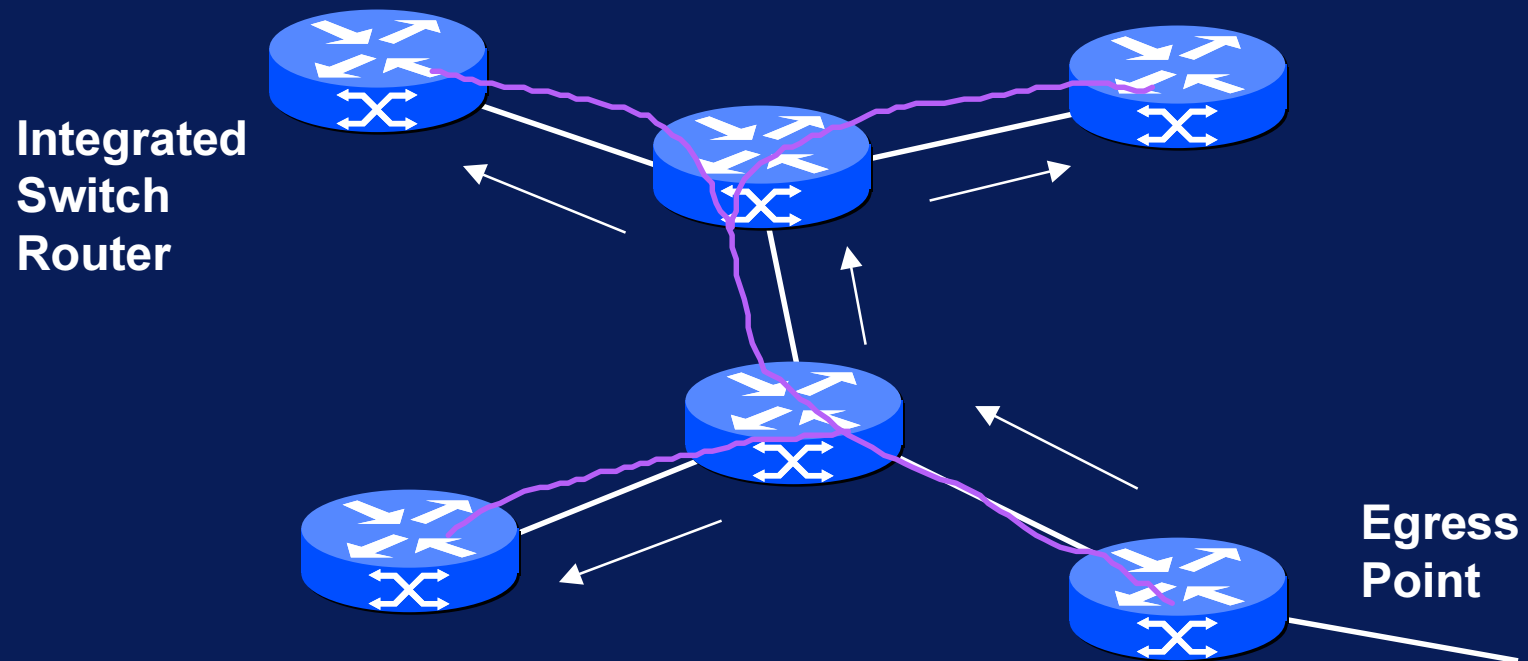
ARIS d'IBM

- ◆ **Aggregate Route-based IP Switching**
- ◆ **ARIS fonctionne entre des Integrated Switch Router, qui comportent une fonctionnalité de routeur et une fonctionnalité de commutation atm associées**
- ◆ **l'établissement des circuits virtuels est basé sur le routage vers les Egress Points, points de sortie du réseau**
- ◆ **chaque Egress Point est à l'origine de la création d'un arbre de CVs convergeant vers lui depuis tout point d'entrée dans le réseau ; le protocole d'établissement fonctionne comme pour Reverse Path Multicast**



UREC

ARIS d'IBM





MultiProtocol Label Switching

Solution	MPLS
Schéma théorique	les routeurs créent des chemins commutés associés aux routes qu'ils connaissent
Avantages	la commutation se fait sur un label court par un circuit spécialisé
Inconvénients	protocole spécifique pour la diffusion des chemins et labels ; activité de création de chemins à priori
Routage	adapté aux réseaux importants (Wan, Man)
Débits	adapté aux débits élevés
Qualité de service	RSVP
Disponibilité	standardisation prévue pour 98
Coût	?
Exemples	projeté par IBM, Cisco



Problèmes IP sur ATM

- ◆ Perte d'une cellule ---> perte du datagramme
- ◆ IP trafic ABR (Best effort) ---> nécessité de lissage, contrôle de flux
- ◆ Pas de SVC au catalogue de FT
- ◆ Lacunes de RFC1577 :
 - Un seul serveur ARP (adresse codée "en dur" dans les stations)
 - Pas de broadcast ou multicast IP
 - Passage par un routeur pour communiquer entre 2 LIS
- ◆ QoS / RSVP
- ◆ Pas de multicast-broadcast en natif dans ATM



Evolutions : NHRP

- ◆ Pour résoudre le problème de classical IP :
Communication entre 2 LIS via un routeur
- ◆ Solution IETF (ROLC Routing Over Large Clouds WG) :
NBMA (Non Broadcast Multiple Access) Next Hop Routing Protocol
(draft-ietf-rolc-nhrp-09.txt)
- ◆ Un réseau (ATM) = n sous-réseaux IP
- ◆ ARP server ---> NHS Next Hop Server
- ◆ 1 sous-réseau IP : 1 NHS (config simple)
- ◆ Requêtes NHRP : directement encapsulées dans ATM



Evolutions : NHRP

◆ Station :

- Connaît son (ses) NHS (config manuelle, ...)
- Connexion : se déclare à son (ses) NHS

◆ NHS :

- Table @ IP - @ ATM
- @ des autres NHS appris :
 - » statiquement (manuel)
 - » protocole de routage "classique"

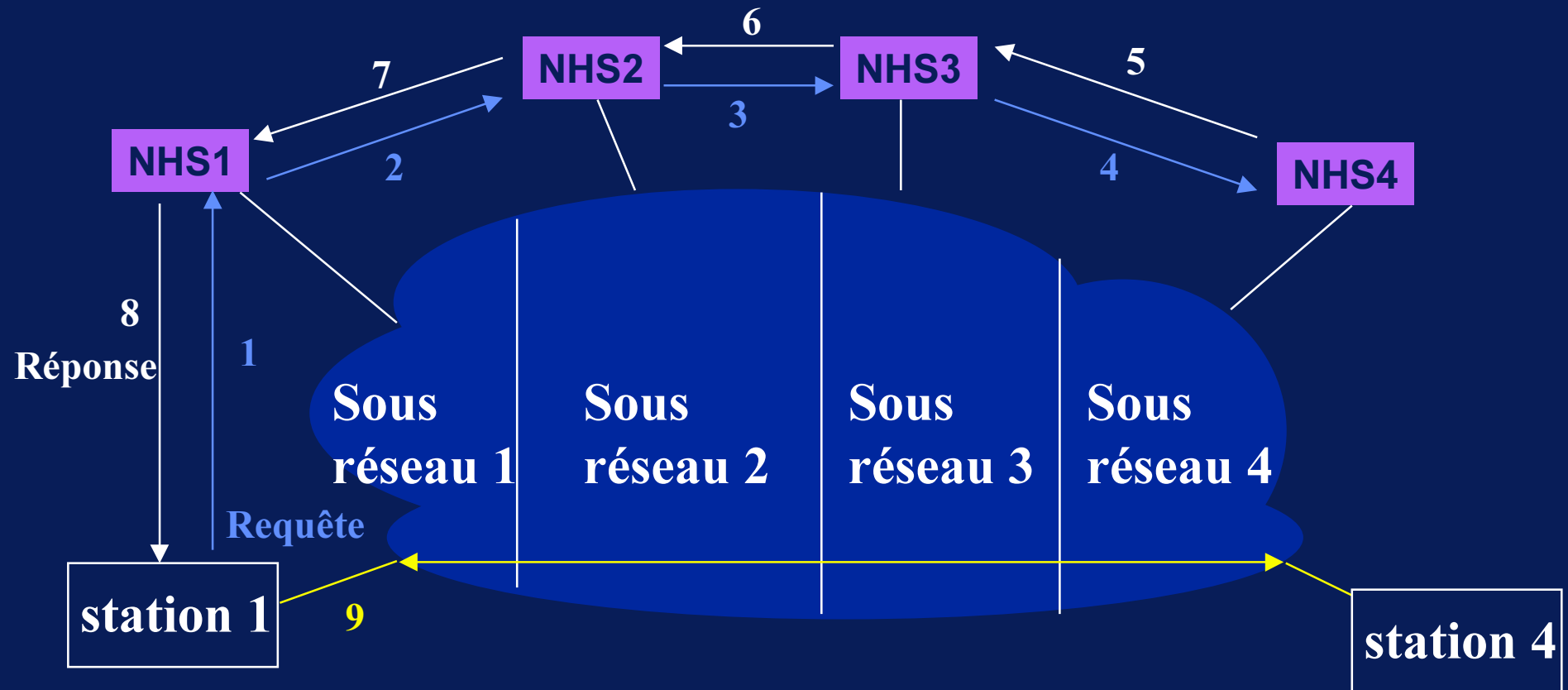
◆ Requête ARP :

- Station ---> NHS
- Si l'adresse non connue le NHS interroge un autre NHS ...



UREC

Evolutions : NHRP





Evolutions : NHRP

- ◆ **Nombreuses fonctionnalités :**
 - La station peut envoyer son paquet vers un "default router" classique en même temps que la requête NHRP ARP
 - Les NHS gèrent des caches pour les adresses "extérieures"
 - Route record (équivalent de traceroute)
- ◆ **Un routeur devrait toujours être NHS**
- ◆ **Non résolu :**
 - Multicast et broadcast
 - Autoconfiguration
- ◆ **Cohabitation stations NHRP et non NHRP : possible**



Evolutions : MARS

- ◆ **Multicast Address Resolution Server**
- ◆ **Sert un cluster de stations**
- ◆ **La diffusion peut être assurée par :**
 - Un réseau multicast de stations
 - Un serveur multicast
- ◆ **Equivalent d'un serveur ARP**
 - Table des membres des groupes multicast



Evolutions : MARS

Modes opératoires

- ◆ **Client (pour envoyer ou recevoir) : ouvre une connexion avec le serveur MARS et indique qu'il veut communiquer avec un groupe multicast**
- ◆ **Serveur MARS répond au client :**
 - Si pas de station enregistrée dans le groupe : silence
 - Sinon : envoie la carte des serveurs multicast pour le groupe ou la liste des stations membres du groupe
- ◆ **Client : ouvre une connexion point-multipoint (avec les stations) ou point à point avec le serveur multicast et envoie ses datagrammes**
- ◆ **Serveur multicast :**
 - S'enregistre auprès du serveur MARS indiquant quels groupes multicast il sert
 - Reçoit la liste des stations des groupes multicast