

IP Switching su ATM

MARIO BALDI, DAVIDE BERGAMASCO, SILVANO GAI,
DANTE MALAGRINÒ

[Mbaldi|Bergamasco|Silvano|Dante]@polito.it

Dipartimento di Automatica e Informatica

Politecnico di Torino

Corso Duca degli Abruzzi, 24

I-10129 Torino (Italy)

Tel. +39-11-5647087

Fax +39-11-5647099

Sommario

Il presente lavoro prende in esame la possibilità di aumentare le prestazioni di reti basate sul protocollo IP (Internet Protocol), ovvero della rete Internet e delle Intranet, grazie all'utilizzo della tecnologia ATM (Asynchronous Transfer Mode) e di efficienti meccanismi di instradamento. Sono passati in rassegna i principali approcci proposti dai costruttori di apparati di internetworking per conciliare la flessibilità e la scalabilità di IP con le elevate prestazioni di ATM. Tali approcci, sebbene non nascano dal lavoro di enti di standardizzazione, godono di grande popolarità grazie alla loro semplicità ed all'immediato beneficio che il loro utilizzo comporta in termini di prestazioni. Viene inoltre presentato un confronto delle varie soluzioni prese in considerazione, ponendo particolare attenzione alle problematiche di scalabilità.

1. INTRODUZIONE

Internet sta crescendo con una curva di tendenza geometrica, raddoppiando il numero delle macchine connesse e dei domini registrati ogni anno; per fronteggiare il conseguente incremento del traffico è necessario aumentare la capacità trasmissiva dell'infrastruttura di rete. Questo risultato può essere ottenuto in due modi:

- collegando le sottoreti attraverso apparati di internetworking più capaci,
- aumentando l'efficienza dei meccanismi di instradamento.

Il presente lavoro si focalizza su quest'ultimo aspetto e prende in considerazione la possibilità conciliare la flessibilità e scalabilità del protocollo IP (*Internet Protocol*) con le elevate prestazioni della tecnologia ATM (*Asynchronous Transfer Mode*). IP, protocollo di livello 3 non connesso, ed ATM, architettura di rete connessa, sono essenzialmente incompatibili. D'altra parte, una loro convergenza è auspicabile: da un lato, le prestazioni di Internet non sono oggi sufficienti a permettere lo sviluppo di nuove applicazioni, dall'altro la tecnologia ATM non appare sufficientemente

flessibile e matura per soppiantare l'utilizzo del protocollo IP. Pare quindi necessaria la ricerca di soluzioni di compromesso ottimali per l'integrazione tra IP e ATM. Attualmente, esistono due approcci di base per l'utilizzo della tecnologia ATM nelle reti IP: *Classical IP Over ATM* [1] e *LAN Emulation (LANE)* [2]. Entrambi questi approcci risultano alquanto inefficienti e in nessuno dei due casi si sono ottenute le prestazioni desiderate. Per questo motivo sono state proposte varie soluzioni di tipo avanzato:

- *Next Hop Resolution Protocol (NHRP)* [3] permette ad una stazione di conoscere l'indirizzo ATM del migliore punto di terminazione di una connessione virtuale (VC) su cui inviare direttamente i pacchetti IP per una data destinazione. L'uso di questa VC garantisce che nel trasferimento dei pacchetti IP attraverso la rete ATM si utilizzi unicamente la commutazione ATM, ossia i pacchetti non siano mai instradati a livello IP;
- *Integrated PNNI (I-PNNI)* [4] si basa su una versione modificata (e compatibile con l'originale) del protocollo di routing definito dallo standard P-PNNI [5] per il trasporto di informazioni di routing sia di livello ATM sia di livello IP; questo consente ai router IP di scegliere in modo ottimale i percorsi di instradamento e di stabilire le corrispondenti VC;
- *Multi-Protocol Over ATM (MPOA)* [6] combina LANE e NHRP. Le stazioni usano LANE o per comunicare localmente (quando la destinazione si trova nella stessa sottorete IP) o per raggiungere un router, detto *MPOA server*, utilizzato per instradare i pacchetti con destinazione esterna alla sottorete; inoltre, le stazioni controllano il traffico per identificare flussi di dati e quando ne individuano uno, utilizzano NHRP per ottenere l'indirizzo ATM della destinazione e stabilire una VC con quest'ultima.

Una forte limitazione degli approcci descritti è rappresentata dall'uso della segnalazione standard ATM [7], troppo pesante da gestire e spesso non correttamente realizzata negli apparati in commercio.

I costruttori di apparati di internetworking, constata la situazione di stallo, hanno quindi proposto alcune soluzioni proprietarie, generalmente identificate col nome di *IP switching*, che mirano a realizzare una sostanziale integrazione tra il protocollo IP e la tecnologia ATM¹. Tra le soluzioni di maggior rilievo è possibile annoverare *Tag Switching* di Cisco Systems, *IP Switching* di Ipsilon Networks, *Cell Switch Router* di Toshiba ed, infine, *Aggregated Route-based IP Switching* di IBM.

Tali soluzioni possono essere classificate in due categorie:

- *traffic-based*, basata sull'osservazione dei flussi di traffico (*IP Switching*, CSR),
- *topology-based*, basata sulla topologia della rete (*Tag Switching* e ARIS).

¹ Più in generale, con il termine multilayer switching si identificano gli approcci che realizzano un'integrazione tra protocolli di livello 3 e tecniche di commutazione di livello 2.

² Alcuni costruttori (Ipsilon, Digital, Ascend) identificano il proprio approccio con la denominazione generica IP switching.

Nella sezione 2 sono descritti gli approcci di tipo traffic-based, nella sezione 3 quelli topology-based, mentre nella sezione 4 è presentata un'analisi comparata dei due approcci. Nella sezione 5 sono, infine, tratte le conclusioni.

2. APPROCCI TRAFFICBASED

2.1 IP Switching

L'architettura IP Switching [8,9] è basata sul concetto di classificazione del traffico, il quale è distinto in due categorie:

- traffico *flow-oriented*, caratterizzato dalla necessità di particolari garanzie sulle risorse di rete quali, ad esempio, la banda trasmissiva o i tempi di latenza;
- traffico *short-lived*, per il quale non è strettamente necessario un trattamento "speciale".

IP Switching classifica come *flusso di traffico* una sequenza di pacchetti IP inviati da una particolare sorgente e indirizzati ad una particolare destinazione (*unicast* o *multicast*), trasportati dal medesimo protocollo di livello superiore (TCP, *Transmission Control Protocol* o UDP, *User Datagram Protocol*) e caratterizzati da uno stesso insieme di parametri nelle intestazioni di livello 3 (IP) e di livello 4 (TCP o UDP).

In un *IP Switch* (sistema intermedio dell'architettura IP Switching) si distinguono due componenti: l'*ATM Switch*, che è una matrice di commutazione ATM programmabile, e l'*IP Switch Controller*, in cui risiede "l'intelligenza" del sistema (protocolli di routing, classificazione del traffico, ecc.) e che si occupa del controllo dello stesso ATM Switch. Utilizzando la classificazione dei flussi, un IP Switch passa dinamicamente dall'instradamento di livello 3 con meccanismi di tipo *store-and-forward* (tipici delle reti IP tradizionali) alla commutazione di livello 2 (ATM), in modo tale che il traffico di tipo flow-oriented transiti direttamente su VC ATM.

Le tabelle di instradamento degli IP Switch sono costruite ed aggiornate mediante i comuni protocolli di routing, mentre la gestione delle VC ATM è affidata a due appositi protocolli: il *General Switch Management Protocol* (GSMP) [10] e l'*Ipsilon Flow Management Protocol* (IFMP), [11]. GSMP è un protocollo locale all'IP Switch e si occupa della gestione dell'ATM Switch, mentre IFMP è un protocollo utilizzato dai nodi³ di una rete IP Switching per comunicare le associazioni tra flussi di traffico e VC ATM.

I vari nodi di rete, durante la fase di inizializzazione, riservano una VC di *default* su ciascun collegamento fisico ATM. In condizioni di funzionamento normale, ogni IP Switch trasmette e riceve i pacchetti IP sotto forma di celle ATM attraverso tale VC (figura 1.A). L'ATM Switch riceve le celle e le passa all'IP Switch Controller che le riasmonta e, come un normale router IP, per ciascun pacchetto consulta la tabella di instradamento e determina la porta d'uscita. Contemporaneamente l'IP Switch

³IP Switch, stazioni direttamente connesse a questi ultimi e traduttori di protocollo (*IP Gateway*) che connettono reti tradizionali (ad esempio LAN Ethernet) a reti IP Switching.

Controller osserva il contenuto di ciascun pacchetto fino all'intestazione di livello 4 al fine della classificazione dei flussi. In questa prima fase le prestazioni di un IP Switch sono confrontabili con quelle di un normale router IP.

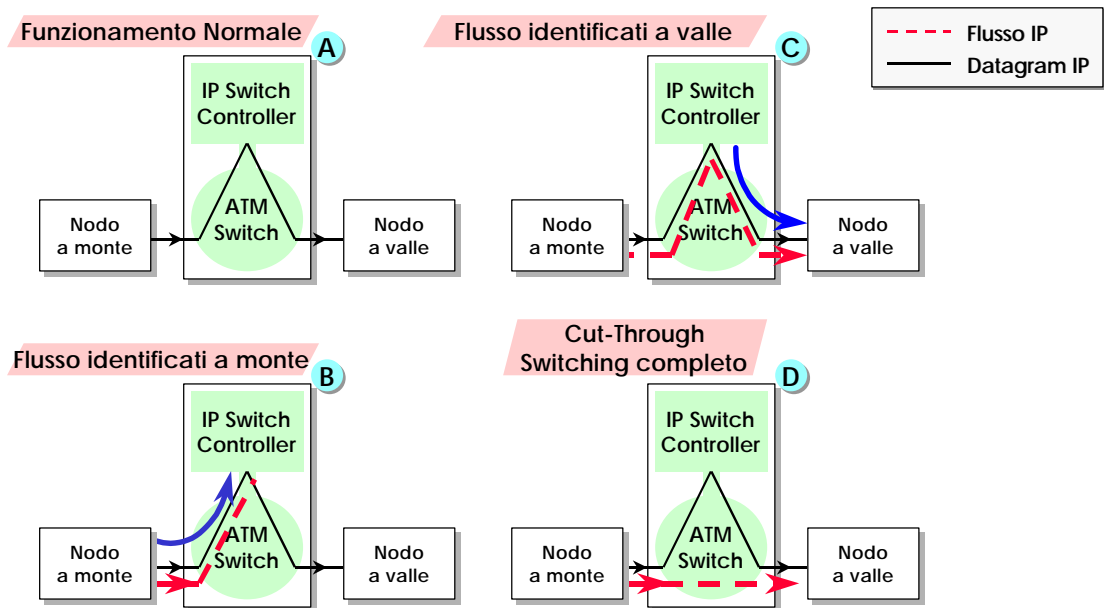


Figura 1 - Meccanismi di funzionamento di IP Switching

Identificata la presenza di un flusso, mediante il protocollo IFMP, l'IP Switch Controller richiede al nodo che si trova a monte (*upstream*) rispetto al verso della trasmissione di inviare i pacchetti appartenenti a detto flusso utilizzando una VC differente da quella di default. A questo punto, se il nodo a monte concorda sull'utilizzo della nuova VC, su di essa sarà fatto transitare tutto il traffico relativo a quel flusso (figura 1.B). In modo del tutto indipendente, ma in modo assolutamente analogo, l'IP Switch Controller può ricevere dal nodo a valle (*downstream*) la richiesta di stabilire per il medesimo flusso un'altra VC, anch'essa diversa da quella di default, da utilizzare questa volta per instradare il traffico in uscita dall'IP switch (figura 1.C).

Una volta stabilite entrambe le VC associate al medesimo flusso, mediante il protocollo GSMP, l'IP Switch Controller configura l'ATM Switch per realizzare la concatenazione di dette VC. Da questo momento in poi tutto il traffico appartenente al flusso in oggetto verrà instradato direttamente attraverso una VC ATM, senza subire i ritardi dovuti all'elaborazione di livello 3 di ogni singolo pacchetto (figura 1.D). Il traffico che non viene classificato come flusso continua ad essere inoltrato dall'IP Switch attraverso la VC di default, secondo il meccanismo di instradamento store-and-forward.

Ciascun IP Switch associa a tutte le VC da esso gestite (esclusa quella di default) un meccanismo di *time-out*. Questo viene utilizzato per abbattere automaticamente le VC sulle quali non fluisce traffico per un determinato periodo di tempo. Dal momento che le VC sono una risorsa piuttosto scarsa negli attuali switch ATM, tale meccanismo risulta di fondamentale importanza in quanto consente di “riciclare” le VC non più utilizzate.

Gli IP Switch forniscono supporto al multicast IP senza richiedere alcuna modifica ai protocolli standard IGMP (*Internet Group Management Protocol*) [12] e DVMRP (*Distance Vector Multicast Routing Protocol*) [13]. Un IP Switch replica un flusso di pacchetti IP con indirizzo di destinazione multicast in più flussi che escono su porte diverse. L'IP Switch Controller configura l'ATM Switch per realizzare una VC punto-multipunto che replica a livello ATM il flusso di ingresso nei flussi d'uscita; ogni volta che a una delle interfacce d'uscita del flusso multicast arriva una richiesta di redirectione, l'IP Switch Controller programma l'ATM Switch in modo tale da aggiungere un ramo alla VC punto-multipunto. Con questo meccanismo è possibile gestire variazioni dinamiche dell'albero di instradamento multicast.

Ogni IP Switch decide localmente le proprie politiche di allocazione delle risorse relativamente ai flussi che lo attraversano. Ciascun flusso può ricevere un trattamento particolare in termini di allocazione di banda, di priorità trasmissiva e di ritardo; questa decisione può essere presa valutando opportunamente il tipo di applicazione a cui il flusso appartiene oppure considerando, ad esempio, il campo *type of service* nell'intestazione del pacchetto IP. È anche possibile sfruttare il protocollo RSVP (*Resource ReSerVation Protocol*) [14] o analoghi protocolli di allocazione delle risorse per gestire richieste di qualità del servizio (QoS) end-to-end relativamente a determinati flussi.

2.2 Cell Switch Router

La proposta di Toshiba per integrare IP e ATM si basa sui *Cell Switch Router* (CSR) [15]. I CSR sono normale router IP dotati di funzionalità di commutazione di cella: l'idea è quella di instradare traffico IP ottenendo throughput elevato e tempi di latenza ridotti grazie alla commutazione ATM. I CSR effettuano la classificazione dei flussi e decidono di volta in volta se i pacchetti appartenenti ad un determinato flusso debbano essere commutati a livello ATM oppure instradati hop-by-hop. I CSR sono dunque apparati simili agli IP Switch ma, a differenza di questi ultimi, utilizzano la segnalazione standard ATM per la gestione delle VC.

In figura 2 è riportato un esempio di architettura di rete basata su CSR. In questa architettura sono definiti tre tipi di VC:

- VC di default, attraverso le quali vengono instradati i pacchetti IP non ancora classificati come flussi,
- VC dedicate, stabilite tra coppie di CSR e tra endsystem e CSR,
- *Bypass-pipe*, VC end-to-end costituita dalla concatenazione di più VC dedicate ed utilizzata per il trasporto di un flusso tra due endsystem.

La VC di default e le VC dedicate sono create utilizzando la segnalazione ATM, mentre la bypass-pipe viene configurata facendo uso del protocollo FANP (*Flow Attribute Notification Protocol*) [16].

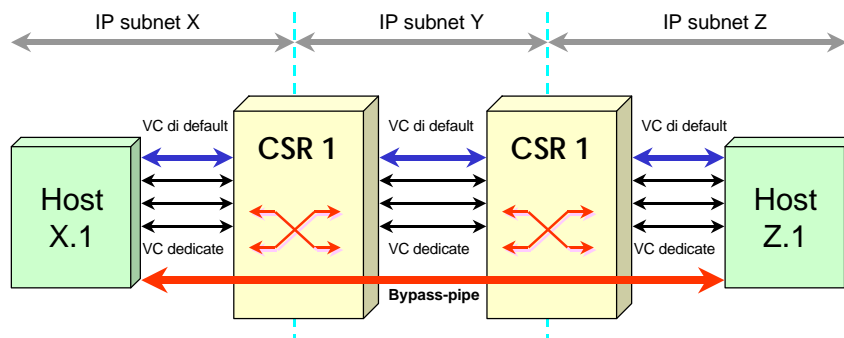


Figura 2 - Esempio di un'architettura di rete basata su CSR

Le bypass-pipe attraversano i CSR che si trovano lungo i cammini determinati dai protocolli di routing IP. L'utilizzo del routing IP in luogo del routing ATM (P-NNI) per la definizione dei percorsi lungo i quali instaurare le connessioni ATM può portare ad un instradamento non ottimale rispetto alla topologia della rete, ma rappresenta un notevole risparmio in termini di software da aggiungere agli apparati. Inoltre, l'utilizzo dei protocolli di routing IP e l'uso di VC dedicate tra coppie di CSR, in luogo di VC end-to-end direttamente stabilite tra coppie di end-system, garantisce la possibilità di modificare dinamicamente i percorsi di instradamento e le risorse di rete ad essi associate. Infine, analogamente agli IP Switch, i CSR possono gestire sia le comunicazioni multicast mediante VC punto-multipunto, sia la QoS attraverso il protocollo RSVP.

3. APPROCCI TOPOLOGYBASED

3.1 Tag Switching

L'architettura di rete proposta da Cisco Systems è detta *Tag Switching* [17]. L'idea alla base di tale architettura è quella di associare a tutti i pacchetti un breve identificativo di lunghezza fissa, detto *tag*, con cui gli apparati di internetworking possano effettuare un instradamento veloce basato sulla commutazione di etichetta (*label-swapping*).

Tag Switching è indipendente sia dal mezzo fisico sia dal protocollo di rete adottati. Può pertanto essere introdotto su collegamenti punto-punto, a bus e reti ATM e può funzionare indifferentemente con IP, IPX o Decnet.

I dispositivi presenti in una rete Tag Switching sono di due tipi [8]:

- *Tag Edge Router* (TER): router di frontiera per una rete Tag Switching; svolgono le comuni funzionalità di instradamento da e verso l'esterno, applicando e rimuovendo i tag ai pacchetti in ingresso ed uscita dalla rete;
- *Tag Switch* (TS): switch che operano la commutazione di etichetta all'interno della rete (livello 2) e prevedono il supporto di funzionalità di instradamento (livello 3).

L'architettura Tag Switching è suddivisa in due componenti totalmente disaccoppiate: l'instradamento (*forwarding*) ed il controllo (*control*). L'aggiunta di nuove funzionalità di routing o la modifica dei protocolli di routing utilizzati sono attuabili sul modulo di controllo senza modifiche né hardware né software al modulo di instradamento.

Ogni TS mantiene aggiornata una TIB (*Tag Information Base*), ovvero una tabella, in cui ogni riga è costituita da un valore di tag d'ingresso (utilizzato come indice) e da uno o più parametri da utilizzarsi per l'instradamento del pacchetto; in particolare, sono presenti il tag da apporre sul pacchetto in uscita, l'interfaccia dello switch su cui instradare il pacchetto ed informazioni supplementari di livello 2. Ogniquale volta lo switch riceve un pacchetto contrassegnato da un determinato tag, viene consultata la TIB e le informazioni contenute nella riga individuata dal tag di ingresso vengono utilizzate per instradarlo.

Le decisioni di instradamento sono basate sulla corrispondenza esatta (*exact match*) dei tag; questa procedura è così semplice da permetterne una realizzazione hardware, con un conseguente miglioramento delle prestazioni in termini di throughput. Nel caso di pacchetto con destinazione multicast, nella TIB anziché avere una sola riga con i parametri d'uscita, sono presenti più righe associate allo stesso tag d'ingresso. La componente di controllo, oltre a mantenere aggiornata la tabella di instradamento mediante i comuni protocolli di routing IP, è responsabile dell'aggiornamento della TIB. Essa, inoltre, propaga ai TS vicini le associazioni tra i tag ed i percorsi di instradamento contenute nella TIB facendo uso di un apposito protocollo denominato *Tag Distribution Protocol* (TDP) [19]. È possibile associare un tag ad ogni percorso di instradamento (*route*), ad un intero gruppo di route oppure ad un singolo flusso di dati relativo ad una data applicazione tra due o più nodi.

Tag Switching, pur essendo utilizzabile con qualsiasi piattaforma di rete, sembra concepito appositamente per ATM [20]. La commutazione di etichetta, infatti, è alla base dell'instradamento dell'una e dell'altra tecnica; inoltre, l'etichetta delle celle ATM, ovvero la coppia VPI/VCI (*Virtual Path Identifier/Virtual Channel Identifier*) consente di realizzare un doppio livello di identificazione (*tagging*) dei pacchetti. Per utilizzare Tag Switching uno switch ATM deve utilizzare sia i protocolli di routing IP sia il TDP. Per evitare conflitti con impieghi standard di ATM, lo spazio dei VPI deve essere opportunamente ripartito tra Tag Switching ed i protocolli di segnalazione ATM.

L'allocazione dei tag viene realizzata con una modalità *downstream on-demand* (figura 3): per ciascuna route nella propria tabella di instradamento, un TS identifica il nodo successivo da raggiungere sul percorso verso la destinazione e richiede a

quest'ultimo l'allocazione di un tag. Il nodo che riceve la richiesta riserva un tag per il percorso in oggetto e comunica la nuova associazione allo switch a monte (*upstream*).

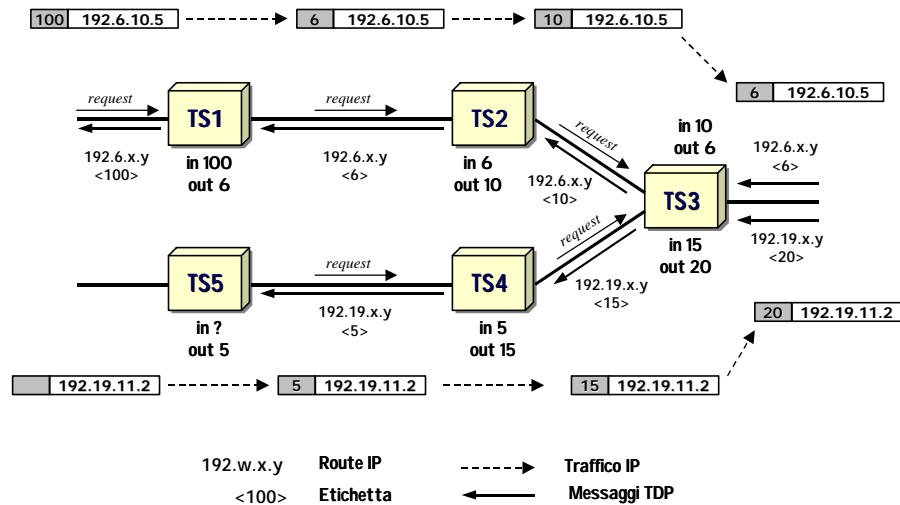


Figura 3- Meccanismo downstream on-demand per l'allocazione delle etichette

Con Tag Switching è possibile classificare i flussi di traffico, suddividendo l'universo dei possibili pacchetti in un insieme finito di classi di equivalenza (*Forwarding Equivalence Classes*, FEC), ognuna delle quali può essere trattata in modo differente. TS adiacenti concordano l'utilizzo di un tag per ciascuna FEC. Due pacchetti appartengono alla stessa classe di equivalenza se hanno una o più caratteristiche comuni:

- sono indirizzati alla stessa sottorete IP (o indirizzo multicast), oppure
- devono attraversare uno stesso router o switch, oppure
- hanno gli stessi indirizzi di sorgente e destinazione⁴, oppure
- hanno gli stessi indirizzi di sorgente e destinazione e sono stati generati dalla stessa applicazione (presentano quindi gli stessi valori per i campi porta dell'intestazione TCP/UDP).

Tag Switching supporta il multicast IP associando un tag ad ogni albero multicast su ciascun TS attraversato dall'albero stesso. Tutti i TS lungo un determinato albero multicast, appartenenti alla medesima sottorete, concordano l'utilizzo di un unico tag per i pacchetti multicast instradati su quella sottorete lungo tale albero.

⁴Questo corrisponde in qualche modo ad individuare come flusso tutto il traffico tra due stazioni.

Tag Switching può garantire QoS end-to-end, attraverso protocolli di allocazione delle risorse, come RSVP. Localmente, l'allocazione delle risorse può essere ottenuta tramite il concetto di classificazione introdotto con le FEC.

3.2 ARIS

L'elemento chiave dell'architettura ARIS (*Aggregated Route-based IP Switching*), [21] è l'*Integrated Switching Router* (ISR), un router IP integrato con una struttura di commutazione ATM. Gli ISR di una rete basata su ARIS instradano normalmente i pacchetti IP, ma il campo delle tabelle di routing che indica il nodo successivo include anche un riferimento ad una determinata VC. Questo permette di instradare i pacchetti facendo uso esclusivamente della commutazione ATM, pur attraversando vari router.

Il collegamento tra i protocolli di routing ed il protocollo per l'instaurazione delle VC è l'*egress identifier* (identificatore d'uscita), usato per identificare un ISR che si trova alla frontiera della rete. In una rete interamente costituita da ISR, ciascuno di questi instaura una VC con ognuno degli ISR di frontiera. Poiché svariate route (anche migliaia) possono transitare attraverso il medesimo ISR d'uscita, ARIS permette di minimizzare il numero di VC aperte associando a dette route un'unica VC.

Si ottengono vantaggi tanto maggiori da ARIS quanto più è elevato il rapporto tra destinazioni IP e ISR d'uscita, cosa piuttosto comune nelle dorsali Internet.

ARIS nell'instaurare le VC tra ISR garantisce l'assenza di cicli (*loop*), anche in presenza di eventuali cicli transitori che potrebbero crearsi durante la fase di convergenza dei protocolli di routing. Questa proprietà è essenziale, poiché in una normale rete ATM, essendo impossibili i cicli di VC, non esistono meccanismi per la loro eliminazione.

4. ANALISI E CONFRONTO

In tabella 1 è riportato un confronto generale degli approcci di IP switching descritti nelle sezioni precedenti. La tecnologia di livello 2 su cui puntano decisamente tutti i costruttori è ATM, anche se in effetti i vari approcci possono utilizzare altre tecnologie, quali Ethernet, Frame Relay, ecc. Analoga situazione si osserva a livello di rete, dove IPv4 ed, in prospettiva, IPv6 dominano la scena.

L'*aggregazione delle route* è un meccanismo che permette di associare la stessa etichetta ad un insieme di comunicazioni che per raggiungere destinazioni diverse attraversano lo stesso nodo intermedio. Tale caratteristica, non realizzabile in approcci di tipo traffic-based, è il punto di forza di ARIS ed è anche prevista in Tag Switching; essa consente di risparmiare sul numero di etichette allocate su ciascuna interfaccia. In una rete ATM questo si traduce in un minor numero di VC aperte, proprietà particolarmente apprezzabile visto il numero ridotto di VC disponibili su ciascuna interfaccia.

	IP Switching	Cell Switch Router	Tag Switching	ARIS
Produttori	Ipsilon e Digital	Toshiba	Cisco	IBM
Protocollo di "segnalazione"	IFMP	FANP	TDP	ARIS protocol
Livello 2	ATM	ATM, FR, etc. (<i>connection oriented</i>)	ATM, FR, Ethernet, etc. (<i>connection oriented e connectionless</i>)	ATM, FR, etc. (<i>connection oriented</i>)
Livello 3	IPv4, IPv6	IPv4, etc.	IPv4, XNS, Apple, etc. (<i>protocolli nel Cisco IOS</i>)	IPv4, IPv6, CLNP, IPX, etc.
Aggregazione delle route	(<i>non realizzabile</i>)	(<i>non realizzabile</i>)	SI (VC merging)	SI (VC merging)
Multicast IP	SI	SI	SI	SI
Supporto per QoS	Locale (shaping e traffic management) o end-to-end (RSVP)	RSVP e segnalazione ATM	RSVP e FEC	RSVP

Tabella 1 - Confronto generale

Per fare aggregazione delle route su ATM, l'utilizzo della stessa etichetta impone la trasmissione sulla medesima VC delle celle appartenenti ai pacchetti da instradare verso lo stesso nodo intermedio (*VC merging*). Questo richiede un meccanismo per identificare l'appartenenza delle singole celle ai vari pacchetti, così da permettere un corretto riassettaggio. Utilizzando AAL5⁵ (*ATM Adaptation Layer 5*) [22], che non prevede questo meccanismo, sulla stessa VC si mescolano (*interleave*) in maniera disordinata celle appartenenti a pacchetti diversi, rendendo impossibile il riassettaggio a destinazione.

Per realizzare il VC merging, utilizzando AAL5, esistono almeno due possibilità:

- si può utilizzare uno switch ATM capace di evitare la miscelazione sulla medesima VC di celle appartenenti a pacchetti diversi, raggruppando tutte le celle di un messaggio in arrivo su una VC di ingresso prima di ritrasmetterle sulla VC di uscita, oppure
- si può utilizzare un intero VP (*Virtual Path*) tra due nodi vicini, all'interno del quale le diverse comunicazioni (e quindi i loro messaggi) vengono distinte grazie al VCI.

⁵AAL5 è il livello di adattamento tipicamente utilizzato per la trasmissione dati su ATM.

La prima soluzione fa perdere praticamente tutti i vantaggi derivanti dalla commutazione di cella, essendo di fatto una commutazione di pacchetto; la seconda comporta uno spreco di connessioni virtuali ed annulla quindi il principale vantaggio derivante dall'aggregazione delle route.

Il concetto di IP switching è nato come possibile soluzione alle sempre maggiori esigenze in termini di banda trasmissiva e garanzie sulla QoS da parte delle nuove applicazioni, specialmente quelle multimediali. Tutti gli approcci considerati prevedono il supporto per il multicast IP. Per quanto riguarda la QoS il protocollo RSVP è necessario per la richiesta delle risorse; a livello locale ognuno degli approcci, fatta eccezione per ARIS dispone di meccanismi propri per fornire garanzie sulla QoS.

4.1 Approccio traffic-based

Il traffico che attraversa la rete Internet ha caratteristiche differenti, a seconda del protocollo a cui appartiene. Esistono protocolli che, per il tipo di informazione trasmessa e per la durata delle comunicazioni *end-to-end* instaurate tra due o più stazioni, tendono a sfruttare intensamente e per periodi di tempo ben determinati la rete, mentre ne esistono altri che effettuano sporadiche trasmissioni di breve durata. Dunque, scegliendo di seguire la strada della classificazione dei flussi per capire quando sia opportuno commutare il traffico a livello 2, il guadagno in prestazioni che si ottiene dipende essenzialmente dalle caratteristiche del traffico.

In uno studio realizzato da P. Newmann, et al. [8] sono riportati i risultati di alcune simulazioni eseguite su diversi tracciati di traffico prelevati da un anello FDDI che connette la Bay Area di San Francisco, California, alla dorsale Internet.

Il risultato principale di questo studio è riportato in tabella 2. Ciascuna riga di tale tabella fornisce la percentuale con cui il corrispondente protocollo incide sul numero totale di flussi, pacchetti e byte, il numero medio dei flussi al secondo, il numero medio dei pacchetti al secondo, la durata media di ciascun flusso, il numero medio di pacchetti trasmessi su quel flusso ed il numero medio di byte per pacchetto. Le righe evidenziate in grigio sono quelle corrispondenti ai protocolli per i quali sembra utile ed opportuno sfruttare la commutazione a livello 2. Questi protocolli sono contraddistinti dall'avere una durata media per ciascun flusso di almeno 20 secondi ed un numero di pacchetti per flusso superiore a 40. Se, com'è ragionevole, si assume che le caratteristiche di questi flussi siano legate alla specifica applicazione a cui fanno riferimento e non al modo in cui l'utente la utilizza, allora le misure effettuate possono considerarsi generali ed indipendenti dalla posizione nella rete in cui viene fatta l'osservazione.

Un altro risultato interessante dello studio citato è che l'84% dei pacchetti ed il 91% dei byte appartengono a flussi. Si è potuta osservare una media di 92 flussi/s, con un numero medio di flussi pari a 15.500, equivalente al numero di VC aperte. Questo può rappresentare un problema di scalabilità se si considera che i commutatori ATM attualmente disponibili in commercio, nella maggior parte dei casi, non sono in grado di gestire più di 4096 VC.

Protocollo	%flussi	%pacchetti	%byte	<u>flussi</u> s	<u>pacchetti</u> s	durata	<u>pacchetti</u> flusso	<u>byte</u> pacchetto
IP in IP	0,04	2,73	2,57	0,09	456	173,1	2307	253
ftp-dati	0,76	12,09	15,18	2,17	2018	118,2	525	338
ftp-cntrl	1,55	0,74	0,23	6,50	124	38,6	16	83
telnet	1,39	4,81	1,61	4,24	803	114,3	114	90
smtp	10,26	4,80	2,82	49,89	802	18,2	15	158
dns	45,30	5,57	3,04	216,56	929	15,4	4	147
gopher	0,45	0,54	0,55	1,87	91	43,3	40	275
http	17,94	40,21	41,53	72,98	6717	56,5	74	278
pop-v3	0,08	0,05	0,03	0,41	9	27,0	21	148
authent	2,12	0,19	0,05	10,54	32	9,0	3	64
nntp	0,35	6,56	6,59	0,68	1096	176,7	627	270
ntp	5,01	0,20	0,06	25,02	33	1,37	1,3	83
netbios	0,03	0,08	0,15	0,11	14	69,8	82	501
snmp	1,35	0,26	0,11	6,14	43	17,9	6	115
login	0,09	0,24	0,14	0,31	41	88,1	92	156
cmd	0,01	0,13	0,07	0,06	21	49,1	316	149
audio	0,00	2,20	2,62	0,01	367	167,9	15653	321
AOL	0,18	0,46	0,38	0,51	77	129,8	84	223
X-11	0,08	0,66	0,53	0,18	111	160,6	276	217

Tabella 2 - Classificazione del traffico in base al protocollo

Su alcuni tracciati prelevati dalla dorsale aziendale di un'impresa operante negli Stati Uniti è stata condotta un'analisi simile a quella già descritta e non si sono riscontrate differenze significative tra i risultati ottenuti nel primo e nel secondo caso.

4.2 Approccio topology-based

Anche per gli approcci topology-based l'elevato numero di VC che è necessario gestire su ciascuna interfaccia rappresenta una seria limitazione alla loro scalabilità. Tale numero dipende dal modo in cui viene effettivamente realizzato il meccanismo di associazione delle etichette ai pacchetti. In figura 4 è riportato un esempio di rete Tag Switching. I TER sono i punti di frontiera per la rete. Si supponga che da ciascuno di essi sia raggiungibile un numero di destinazioni pari al numero indicato (ad esempio, da TER1 sono raggiungibili R1 route). Ogni TS partecipa alla defi-

nizione dei percorsi di instradamento e può trovarsi o meno sul percorso di instradamento migliore per raggiungere una certa uscita da un dato ingresso. Si supponga che un host raggiungibile da TER6 abbia necessità di comunicare con un host raggiungibile attraverso TER5; il percorso ottimale tra questi due host viene stabilito dai protocolli di routing e potrebbe ad esempio essere quello che transita per TER6 - TS4 - TS3 - TER5. In queste condizioni TS1, TS2 e TS5 non instraderanno mai pacchetti inviati da TER6 a TER5. TS4 e TS3, invece, dovranno prevedere, all'interno della loro TIB, una coppia (*tag d'ingresso*, *tag d'uscita*) da utilizzare per tutti i pacchetti scambiati nella medesima direzione tra gli host considerati. Se si estende questo ragionamento a tutti i possibili percorsi di instradamento ottimali presenti sulla rete si può dire che il numero di VC necessarie su ciascuna interfaccia di un TS è pari al prodotto tra il numero di destinazioni raggiungibili da quell'interfaccia (secondo i percorsi stabiliti dai protocolli di routing) ed il numero di sorgenti collegate agli ingressi della rete Tag Switching che le raggiungono passando attraverso il TS in esame.

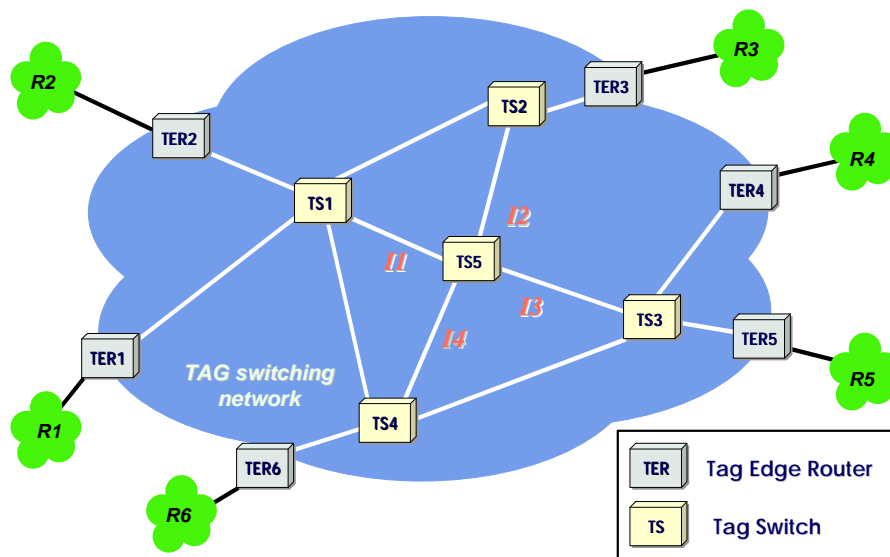


Figura 4 *Esempio di rete basata su Tag Switching*

Questo significa che, considerando ad esempio TS5, e supponendo che questo si trovi sul percorso migliore per raggiungere TER3 da TER1, TER4, TER5 e TER6, utilizzando l'interfaccia *I3*, e che dalla stessa interfaccia non siano raggiungibili altre uscite della rete, il numero di VC necessario sul collegamento tra TS5 e TS2 è pari a $(R1 + R4 + R5 + R6) \times R3$.

Da tali considerazioni si può pensare ad un numero massimo di etichette per porta dell'ordine del quadrato del numero delle route, visto che al limite (nel caso

peggiore) tutte le uscite potrebbero essere raggiungibili da tutte le entrate, passando da uno stesso switch. Data una rete anche complessa e noti i percorsi di instradamento ottimali, è possibile stabilire con esattezza il numero di tag necessario per ciascuno switch.

È stato stimato che sulla dorsale Internet vi siano 100.000 route; se si suppone che un TS sia sul percorso ottimale per raggiungere metà delle sottoreti da parte di un'altra metà, allora il numero di VC necessarie è pari al quadrato di 50.000, cifra superiore al numero massimo di VC disponibili.

Utilizzando protocolli di tipo *link-state*, come OSPF, i TS interni alla rete possono conoscere i TER dai quali determinate sottoreti sono raggiungibili. Questo permette di associare un'unica etichetta a tutte le sottoreti raggiungibili attraverso il medesimo TER, migliorando nettamente le proprietà di scalabilità del protocollo. Nell'esempio precedente, sulla medesima interfaccia sono necessari solo $4 \times 1 = 4$ VC, mentre, nel caso peggiore, il numero complessivo di VC non è più dipendente dal numero di route, ma da quello di TER con cui la rete Tag Switching si interfaccia con l'esterno. Una tale soluzione richiede l'impiego del VC merging, tecnica attualmente non prevista da nessun costruttore.

	1 tag per route	1 tag per TER
Senza VC merging	$O(R^2)$	$O(N^2)$
Con VC merging	$O(R)$	$O(N)$

Tabella 3 - Numero di tag per interfaccia in una rete Tag Switching
(R = numero di route, N = numero di TER)

In tabella 3 viene mostrato come varia la il numero di tag necessari in una rete Tag Switching in funzione dell'utilizzo del VC merging e della modalità di assegnazione dei tag.

5. CONCLUSIONI

In questo lavoro sono prese in esame alcune tecniche per combinare in modo efficace le caratteristiche del protocollo IP (*Internet Protocol*) con quelle della tecnologia ATM (*Asynchronous Transfer Mode*). Tali tecniche, genericamente denominate *IP switching*, pur non essendo le uniche esistenti, sono tuttavia le più recenti e la loro definizione non deriva dal lavoro di un ente di standardizzazione ma dall'iniziativa di alcuni costruttori. Le quattro soluzioni prese in esame, IP Switching, Cell Switch Router, Tag Switching, e Aggregated Route-based IP Switching, sono nate indipendentemente, ma presentano obiettivi e caratteristiche generali comuni.

Gli approcci considerati sono stati suddivisi in due categorie, traffic-based e topology-based, a seconda del modo in cui vengono create le VC ATM per il trasporto dei pacchetti IP. Su di essi è stata condotta un'analisi comparativa che ne

ha messo in luce vantaggi e debolezze. Da tale analisi è emerso che questi approcci non sono ancora sufficientemente maturi e definiti da poter identificare *una* soluzione vincente.

6. BIBLIOGRAFIA

- [1] M. Laubach, "Classical IP and ARP over ATM", RFC1577, gennaio 1994.
- [2] The ATM Forum, "LAN Emulation over ATM Version 1.0", gennaio 1995.
- [3] D. Katz et al., "NBMA Next Hop Resolution Protocol (NHRP)", Internet-Draft, <draft-ietf-rolc-nhrp-11.txt>, marzo 1997.
- [4] The ATM Forum, "Integrated PNNI (I-PNNI) v1.0 Specification", Work in progress, ATM Forum 96-0987R1, ottobre 1996.
- [5] The ATM Forum, "Private Network-Network Interface Specification Version 1.0", marzo 1996.
- [6] The ATM Forum, "Multi-Protocol Over ATM Version 1.0", Work in progress, STR-MPOA-MPOA-01.00, febbraio 1997.
- [7] The ATM Forum, "ATM User-Network Interface Specification Version 3.1", 1994
- [8] P.Newmann, G.Minshall, T.Lyon, "IP Switching: ATM Under IP", Submitted to IEEE/ACM Transactions on Networking.
- [9] AA.VV, "IP Switching: The Intelligence of Routing, The Performance of Switching", White Paper, Ipsilon Networks, Inc.
- [10] P.Newmann, et al., "Ipsilon's General Switch Management Protocol Specification Version 1.1", RFC1987, Agosto 1996.
- [11] P.Newmann, et al., "Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0", RFC 1987, Maggio 1996.
- [12] S.Deering, "Host Extension for IP multicasting", RFC1112, agosto 1989.
- [13] D.Waitzman, "Distance Vector Multicast Routing Protocol", RFC075, novembre 1988.
- [14] Braden R. et al., "Resource ReSerVation Protocol (RSVP) - Version1 Functional Specification", InternetDraft, novembre 1996.
- [15] Y.Katsube et al., "Toshiba's Router Architecture Extension for ATM: Overview", RFC2098, febbraio 1997.

- [16] K.Nagami et al., “Flow Attribute Notification Protocol (FANP) Specification”, InternetDraft, <draft-rfced-info-nagami-00.txt>, novembre 1996.
- [17] Y.Rekhter et al., “Tag Switching Architecture Overview”, InternetDraft, <draft-rekhter-tagswitch-arch-00.txt>, gennaio 1997.
- [18] Cisco Systems, Inc., “Scaling the Internet with Tag Switching”, White Paper, settembre 1996.
- [19] P.Doolan et. al., “Tag Distribution Protocol”, InternetDraft, <draft-doolan-tdp-spec-00.txt>, settembre 1996.
- [20] B.Davie et al., “Use of Tag Switching with ATM”, InternetDraft, <draft-davie-tag-switching-atm-01.txt>, gennaio 1997.
- [21] R.Woundy et al., “ARIS: Aggregate RouteBased IP Switching”, Internet-Draft, <draft-woundy-ipswitching-00.txt>, novembre 1996.
- [22] ITU-T, “B-ISDN ATM Adaptation Layer (AAL) Specification”, Recommendation I.363, marzo 1993.