

IMPLEMENTAZIONE DI UN SISTEMA DI PROTEZIONE GMPLS CONDIVISO PER RETI ETHERNET OTTICHE

Filippo Cugini

CNIT, Laboratorio Nazionale di Reti Fotoniche, via Moruzzi 1, Pisa,
e-mail: filippo.cugini@cnit.it

Luca Valcarengi, Piero Castoldi

Scuola Superiore Sant'Anna, via Moruzzi 1, Pisa,
e-mail: valcarengi@sssup.it, castoldi@sssup.it

Abstract: A cost-effective implementation of GMPLS shared protection is successfully demonstrated in an IP over Gigabit Ethernet testbed. Experimental results show that recovery times in the millisecond range are achieved.

1. Introduzione

Le architetture di rete basate sul protocollo Ethernet per il trasporto di pacchetti IP/MPLS direttamente su fibra ottica (1 e 10 Gigabit Ethernet, (GbE)) rappresentano una soluzione semplice ed efficiente per superare il collo di bottiglia che storicamente limita la velocità di trasmissione dati fra reti locali e reti geografiche. Negli ultimi anni infatti, sistemi basati su connessioni GbE ottiche trovano impiego non solo per il collegamento fra switch di livello 2 in reti locali, ma anche per il collegamento diretto fra router che operano a livello 3 in reti metropolitane. Il principale fattore limitante alla completa diffusione di GbE è però rappresentato dalla mancanza di alcune importanti funzioni per la gestione e manutenzione della rete (OA&M), come la protezione delle connessioni in caso di guasto. Spesso viene quindi impiegato uno schema di protezione di tipo 1+1, estremamente semplice e veloce, ma che richiede la duplicazione delle risorse quali costose interfacce sui router e fibre ottiche. Soluzioni più avanzate basate su tecnologia WDM riducono le necessità di fibre ma aumentano sensibilmente il costo complessivo dell'infrastruttura di rete.

In questo studio vengono dapprima sperimentalmente valutate le principali limitazioni che caratterizzano la protezione di connessione punto-punto GbE su fibra ottica fra attuali router commerciali destinati ad un utilizzo in ambito metropolitano. In particolare si evidenzia il tempo impiegato dai router per rilevare e gestire il guasto della connessione e viene valutata la tempistica richiesta dal protocollo di instradamento Open Shortest Path First (OSPF) per ripristinare il corretto scambio di pacchetti una volta riattivata fisicamente la connessione. Vengono quindi proposte le linee guida per migliorare il comportamento dei router nelle due situazioni analizzate. Lo studio presenta infine l'implementazione di uno schema di protezione condiviso veloce ed a costo contenuto che evita i problemi e le inefficienze evidenziate sperimentalmente dai router. Lo schema è basato sull'impiego di economici dispositivi di commutazione ottica (PXC) controllati da segnalazione fuori-banda basata sul protocollo Generalized Multi-Protocol Label Switching (GMPLS). L'implementazione mostra come sia possibile sfruttare le potenzialità del protocollo GMPLS per ridurre l'impiego di risorse pregiate dell'infrastruttura di rete (interfacce sui router e fibre ottiche) pur garantendo tempi di ripristino della connessione di pochi millisecondi.

2. Analisi del tempo di rilevazione di guasti su connessioni GbE punto-punto

Il testbed di rete considerato consiste in due router IP/MPLS connessi tramite un collegamento GbE (1000BaseLX) punto-punto su fibra ottica monomodale. Obiettivo delle prove sperimentali è l'analisi del comportamento dei router nel caso in cui si verifichi una rottura del collegamento (es. taglio della fibra). La rottura viene realizzata agendo su un dispositivo di commutazione ottica trasparente (PXC) inserito lungo la fibra che connette i due router. Inizialmente il PXC risulta configurato in modo tale da consentire la

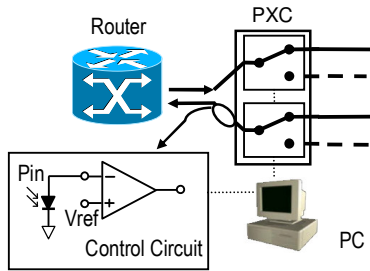


Fig. 1: Dispositivo di commutazione ottica (PXC) con circuito di rilevazione di guasti

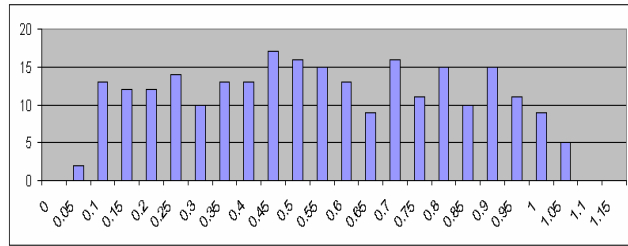


Fig. 2: Distribuzione del ritardo (in secondi) introdotto dal router per rilevare il guasto su connessioni GbE.

comunicazione fra i due router, le cui interfacce risultano quindi attive (*up*). La commutazione del PXC (comandata attraverso un PC di controllo) provoca l'interruzione del collegamento (*down*). Il PXC è caratterizzato da tempi di commutazione di circa 3ms. Al fine di rilevare i tempi di reazione del router, lo stesso PC di controllo viene connesso anche all'interfaccia di gestione dei router e viene attivata la continua interrogazione dello stato dell'interfaccia. Il router infatti, è in grado di gestire una comunicazione client-server basata sullo scambio di messaggi Extended Markup Language (XML) attraverso la quale può fornire lo stato operativo delle proprie interfacce (*up* o *down*). La potenza ottica del segnale in linea viene inoltre continuamente monitorata dal PC di controllo attraverso un meccanismo di rilevazione di guasti basato sull'assenza di potenza ottica in fibra (Fig. 1, [1]). Parte del segnale ottico ricevuto da ogni router viene infatti spillato e inviato ad un circuito di controllo. Il circuito provvede a convertire il livello di potenza ottica in un valore di tensione che giunge ad un comparatore a soglia. In caso di guasto, il livello di tensione scende sotto soglia ed il circuito notifica al PC, attraverso la porta parallela, l'avvenuta rottura del collegamento. Confrontando quindi le informazioni sullo stato operativo delle interfacce dei router con gli eventi legati alla commutazione del PXC e conseguente assenza di potenza ottica in fibra è possibile misurare il ritardo introdotto dal router per la rilevazione del guasto.

Fig. 2 mostra i risultati ottenuti in termini di distribuzione dei tempi di rilevazione del guasto da parte del router. Il tempo medio misurato è di 0.573 s, la deviazione standard 0.275 s, il minimo 92 ms, il massimo 1.091 s. Si nota quindi che la distribuzione dei valori risulta piuttosto uniforme fra 0.1 s e 1 s, tale quindi da far presupporre una verifica dello stato delle interfacce da parte del router solo di una volta al secondo. Ciò quindi determina tempi di ripristino del guasto estremamente elevati se confrontati con i 50 ms che caratterizzano le reti basate sul protocollo SDH. Infatti qualsiasi metodo di ripristino della connettività attivato dal router, ad esempio MPLS Fast Reroute su un percorso alternativo, risulta pesantemente influenzato dal tempo di rilevazione del guasto.

L'evoluzione dei router commerciali dovrà quindi necessariamente tenere conto di questi aspetti e integrare meccanismi più efficienti per la rilevazione del guasto.

3. Analisi del tempo di attivazione di adiacenza tra router

Una volta rilevato il guasto, ciascun router rimuove immediatamente dalle proprie tabelle di instradamento la presenza del router adiacente e delle reti annunciate attraverso quest'ultimo. Pertanto, una volta ripristinato fisicamente lo stato della connessione, risulta necessario attendere una nuova instaurazione dell'adiacenza tra i router e la sincronizzazione delle tabelle di instradamento prima che la trasmissione del traffico dati possa riprendere. L'analisi sperimentale della procedura di attivazione dell'adiacenza è risultata la seguente:

- Appena rilevato lo stato attivo dell'interfaccia ciascun router inizia a trasmettere periodicamente messaggi di *Hello* verso l'indirizzo di destinazione multicast *AllSPFRouter* (224.0.0.5). Tipicamente, al quinto messaggio di *Hello* complessivamente scambiato (dopo circa 1.5 s dall'attivazione) ciascun router indica nel campo *neighbours*

- del messaggio *Hello* la presenza del router connesso (stato *2-way*).
- Le interfacce Ethernet sono considerate dal protocollo OSPF di tipo Broadcast. Quindi un lungo scambio di messaggi di *Hello*, tutti identici tra loro, consentono ai router di procedere all'elezione del *Designated Router* (DR) e *Backup DR* (BDR). Questa fase termina dopo circa 40 s dall'attivazione della connessione.
- La fase successiva (stato *Database Exchange* e *Loading*) prevede lo scambio delle informazioni contenute nelle proprie tabelle di instradamento.
- I primi pacchetti di traffico dati vengono scambiati mediamente dopo 45 s dall'attivazione fisica della connessione.

La stessa indagine sperimentale è inoltre stata condotta analizzando il comportamento di una connessione fra interfacce ATM in modalità punto-punto (potrebbero esser configurate anche in modalità NBMA, Non Broadcast Multi Access). In questo caso la tempistica misurata risulta tale da non superare i 5 secondi fra l'attivazione fisica della linea ed il transito dei primi pacchetti di traffico dati. Infatti, per interfacce di tipo punto-punto, il protocollo OSPF non procede all'elezione del DR e del BDR, riducendo notevolmente i tempi di attivazione dell'adiacenza.

Il confronto fra le due procedure consente di individuare un miglioramento parziale ma significativo da applicare nel caso di reti Ethernet. Tali reti, tipicamente di tipo Broadcast, possono infatti consentire la presenza di più di due interfacce e router connessi direttamente. Per questo motivo l'adiacenza del protocollo OSPF per ciascun router viene instaurata solo con il DR (ed il BDR). Come indicato dalle misure riportate precedentemente anche nel caso di connessioni punto-punto questa fase viene attivata, ma con l'ovvio risultato che uno dei due router viene eletto DR e l'altro BDR, con l'adiacenza instaurata direttamente fra loro. Questa fase introduce un ritardo di quasi 40 secondi.

Per ottenere una riduzione di tale ritardo nel caso di connessioni Ethernet punto-punto fra router si propone di imporre al router l'interpretazione della connessione come punto-punto in modo tale da evitare la fase di elezione del DR e del BDR. Una connessione punto-punto è identificabile in modo implicito a livello di configurazione dell'indirizzo IP dalla presenza di una sottorete a 30 bit o in modo esplicito introducendo un semplice comando di configurazione, esempio *point-to-point* come per le connessioni di tipo ATM. Ciò non richiede modifiche al protocollo OSPF.

4. Implementazione di uno schema di protezione condiviso GMPLS

Per evitare i problemi causati dalla lenta rilevazione del guasto da parte dei router IP/MPLS e dalla lenta riattivazione delle adiacenze perse, viene proposto in questa sezione un sistema di protezione condiviso GMPLS.

Il sistema proposto, implementato nel testbed mostrato in Fig. 3, prevede la presenza di connessioni punto-punto fra interfacce GbE ottiche dei router e la presenza di PXC trasparenti al traffico dati. Ciascun PXC è controllato da un PC Linux ed è dotato di un sistema di rilevazione di guasti basato sul meccanismo di assenza di potenza ottica in fibra. La comunicazione fra i PC avviene fuori banda attraverso una rete di controllo Fast Ethernet. Due coppie di router (A-B e C-D) sono connessi attraverso due percorsi principali distinti che attraversano i PXC 1,2 e 5,6 rispettivamente. I PXC 3 e 4 sono utilizzati per condividere una unica fibra di protezione. In caso di guasto su uno dei due percorsi principali, per esempio tra PXC 1 e 2, il meccanismo di rilevazione guasti sul PXC 2 segnala l'assenza di potenza ottica al PC 2 il quale attiva la procedura di ripristino GMPLS. Al fine di localizzare il guasto, il PC 2 e il PC 1 (che controlla il PXC1 precedente rispetto al flusso dati) si scambiano messaggi Link Management Protocol (LMP) di tipo *Channel Status* (messaggi 1-3 in Fig. 4). Una volta localizzato il guasto, PC 1 attiva la procedura di attivazione del percorso di protezione inviando al PC 3 un messaggio *RSVP PATH*. Nel caso in cui la connessione di protezione fra i PXC 3 e 4 risulti disponibile il messaggio viene inoltrato attraverso il nodo 4 fino al nodo 2 (messaggi 4-6 in Fig. 4) e contemporaneamente vengono inviati gli opportuni comandi ai PXC per l'attivazione fisica del percorso alternativo. Messaggi *RSVP RESV* di conferma vengono quindi inviati verso il PC 1 seguendo a ritroso il percorso attivato.

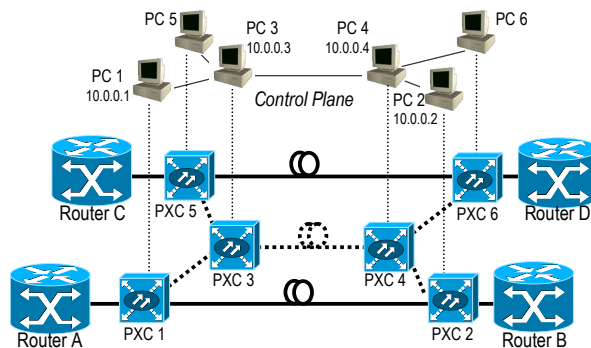


Fig. 3: Schema di protezione condiviso per reti IP su Gigabit Ethernet.

Lo schema di protezione condiviso è stato verificato interrompendo numerose volte la connessione fra i PXC 1 e 2. Il meccanismo di rilevazione del guasto si è sempre attivato correttamente. Lo scambio di messaggi LMP e RSVP ha complessivamente richiesto tempi inferiori ai 2 ms (Fig. 4). Poiché la commutazione dei PXC richiede mediamente circa 3 ms, il tempo di fuori servizio complessivo è risultato tipicamente di soli 5ms (Fig. 5).

La soluzione proposta consente quindi di ottenere tempi di ripristino al guasto estremamente ridotti, tale quindi da esser trasparente ai router connessi che mai hanno rilevato interfacce *down* e mai hanno dovuto riconfigurare le tabelle di instradamento. La soluzione proposta mostra quindi i vantaggi ottenibili dall'introduzione di PXC trasparenti per scopi di protezione ed evidenzia i vantaggi ottenibili dall'adozione del protocollo di controllo GMPLS, in particolare per ridurre l'impiego di fibre ottiche e interfacce dei router garantendo contemporaneamente tempi di ripristino ai guasti estremamente ridotti.

No.	Time	Source	Destination	Protocol	Info
1	0,000000	10,0,0,2	10,0,0,1	LMP	ChannelStatus Message.
2	0,000338	10,0,0,1	10,0,0,2	LMP	ChannelStatusAck Message.
3	0,000465	10,0,0,1	10,0,0,2	LMP	ChannelStatus Message.
4	0,000597	10,0,0,1	10,0,0,3	RSVP	PATH Message, SESSION: IPv4
5	0,000793	10,0,0,3	10,0,0,4	RSVP	PATH Message, SESSION: IPv4
6	0,001107	10,0,0,4	10,0,0,2	RSVP	PATH Message, SESSION: IPv4

Fig. 4: Messaggi GMPLS scambiati sul canale fuori-banda. Tempo in secondi.

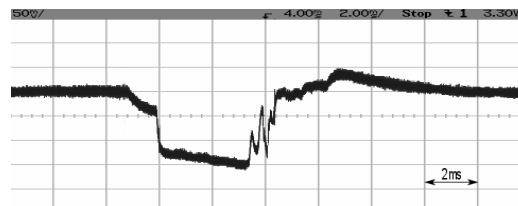


Fig. 5: Tempo di fuori servizio al router B.

5. Conclusioni

In questo studio sono stati analizzate sperimentalmente le tempistiche legate alla rilevazione di guasti e ripristino delle adiacenze in connessioni ottiche punto-punto di tipo Gigabit Ethernet fra router IP. Sono state proposte quindi due possibili soluzioni per incrementarne le prestazioni. Infine è stata presentata l'implementazione di uno schema di protezione condiviso GMPLS tale da superare i limiti evidenziati dai router nel ripristino delle adiacenze e consentire tempi di ripristino di pochi millisecondi.

Ringraziamenti

Gli autori desiderano ringraziare l'Istituto Superiore per le Comunicazioni e Tecnologie dell'Informazione (ISCTI) per la fornitura di alcuni apparati di rete utilizzati per le prove.

Bibliografia

- [1] F. Cugini, L. Valcarengi, P. Castoldi, G. Ippoliti, "Experimental demonstration of a low-cost 1:1 optical span protection", Workshop on "Protection and Restoration" GLOBECOM 2003, San Francisco (USA), Dec. 2003.

Filippo Cugini received the Laurea degree in Telecommunication Engineering from the University of Parma, Italy. From 2001 he is a research engineer at the CNIT National Laboratory of Photonics Networks in Pisa, Italy. His main research interests include MPLS and GMPLS protocols, architectures and implementations, survivability in IP over WDM networks, Traffic Engineering in Grid Networking.

Dr. Cugini is currently involved in several European and international projects such as the e-Photon/ONe Network of Excellence and the GRID.IT projects.

Luca Valcarengi holds a Laurea degree in Electronics Engineering (1997) from the Politecnico di Torino, Italy, a M.S. in Electrical Engineering (1999), and a Ph.D. in Electrical Engineering-Telecommunications (2001) both from the University of Texas at Dallas (UTD).

Between January 2002 and August 2002 he has been Research Associate of the Optical Networking Advanced Research (OpNeAR) Lab of the University of Texas at Dallas Erik Jonsson School of EE/CS. Since September 2002 he is Assistant Professor at the Scuola Superiore Sant'Anna of University Studies and Doctoral Research of Pisa, Italy.

Dr. Valcarengi co-authored more than two dozen papers published in international journals and presented in leading international conferences. He has been part of the Organizing Committee and Technical Program Committee of international conferences such as Optical Networking and Systems Symposium at IEEE Globecom 2003, and Optical Communication Networks and Systems Symposium at IEEE Globecom 2004. His main research interests are Optical Networks design, analysis, and optimization; Artificial Intelligence optimization techniques; Communication Networks reliability; IP over WDM networking; QoS in network infrastructures for Grid computing.

Piero Castoldi got his laurea degree (cum laude) in Electrical Engineering at the University of Bologna, and his Ph.D. degree from the University of Parma, in 1991 and 1996, respectively. In 1996 and 1997 he was a post-doc at the Department of Electrical Engineering of Princeton University (USA) where he was also hosted in the summers 1999 and 2000. From December 1997 until February 2001, he was assistant professor at the Faculty of Engineering of the University of Parma, where he got the tenure. Since March 2001 he has been Associate Professor in Telecommunications at the Scuola Superiore Sant'Anna of Pisa, being responsible for the "networks group". In January and February 2002 he has been Visiting Professor at the University of Texas at Dallas, USA. He also collaborates to several projects of the Italian National Consortium for Telecommunications (CNIT) and since March 2003 he is Director of the CNIT National Photonic Networks Laboratory in Pisa, Italy.

He has held, as an instructor, several courses at the undergraduate and graduate level at the University of Parma and at Scuola Sant'Anna. The scientific activity has covered the area of digital transmission, telecommunication networks performance, and, more recently, provisioning, protection and switching techniques for next generation optical networks. He is author of more than 50 technical papers published in international conferences and journals and of an international book on CDMA.