

G.Maccaferri – M.Nanni

**Sistema di monitoraggio per le reti a  
1 Gigabit/sec**

IRA 373/05



## Introduzione

La convenzione firmata in data 24 Novembre 2004 stabilisce che l'Istituto di Radioastronomia dell'INAF si impegna a fornire alla Regione Emilia-Romagna uno strumento in grado di misurare le prestazioni della infrastruttura di rete. In particolare il sistema deve essere in grado di verificare carichi trasmissivi fino a circa 1 Gbit/sec fra due qualsiasi punti della rete, indipendentemente dagli apparati attivi e passivi posti lungo il percorso, valutando il traffico sia di tipo TCP che UDP.

In queste pagine vengono descritte le apparecchiature ed i software utilizzati e sviluppati per realizzare il sistema di monitoraggio per la rete Lepida.

## Le apparecchiature

Il sistema di monitoraggio e' composto da due unità separate, una delle quali dovrà essere installata presso un punto di concentrazione della rete Lepida, o presso il centro servizi di un gestore, mentre l'altra potrà essere posizionata di volta in volta ai capi delle tratte da monitorare, e quindi deve essere compatta e facilmente trasportabile.

Le specifiche prevedono che le unità di monitoraggio devono essere dotate di porte GigaEthernet in rame (RJ45 1000baseT) per il collegamento con apparati di rete presenti in sede locale. L'unità trasportabile deve avere una seconda porta di rete per essere eventualmente pilotata attraverso un portatile.

Dopo numerosi test ed analisi dei prodotti reperibili sul mercato ci si e' orientati su due sistemi diversi, un server da rack di altezza U1, da posizionare presso il centro servizi di Lepida, ed un calcolatore Veriton 3700G della ACER che ha dimensioni ridotte e risulta quindi facilmente trasportabile nei vari luoghi dove si voglia effettuare il monitoraggio della rete.

Server da rack U1 ( cm 43 x 4.3 x 70 )	<b>TYAN GX28</b> - CPU: 2x Opteron 246 2GHz, Ram : 1Gbyte Scheda di Rete: 2 Broadcom NetXtreme BCM5704C su bus PCI-X 133
Sistema trasportabile ( cm 34 x 10 x 41)	<b>ACER Veriton 3700G</b> - CPU : P4 2.8GHz , Ram 512MByte, Schede di Rete: 1 Broadcom NetXtreme BCM5751 su bus PCI-Express, 1 scheda 3Com 3c940 per collegamento portatile

Nella scelta dei sistemi per il monitoraggio l'elemento critico e' risultato essere, prima ancora che la scheda di rete, il bus interno. Anche se sono decine le schede che dichiarano di supportare velocità di 1 Gbit/sec in realtà solo quelle che si collegano a bus moderni e veloci quali il PCI-X 133 a 64 bit, oppure il PCI-Express di ultima generazione riescono a trasmettere dati a velocità prossime al Giga.

Quindi tra le centinaia di sistemi offerti dal mercato solo pochi avevano le caratteristiche richieste, e tra questi non abbiamo trovato portatili che fornissero sufficienti garanzie di velocità; il sistema trasportabile da noi individuato puo' essere utilizzato come un normale personal computer, disponendo di un monitor e di una tastiera, oppure puo' essere pilotato da un portatile attraverso la seconda scheda di rete.

Entrambi i computer hanno installato il sistema operativo open source Linux nella distribuzione Fedora Core 3 con Kernel 2.6.9 del Gennaio 2005. Il sistema trasportabile monta anche Windows XP professional ma nell'insieme il sistema di monitoraggio e' stato pensato in ambiente Linux.

Nelle prove "a banco" con collegamento diretto attraverso cavetto incrociato il sistema ha mostrato di poter misurare le seguenti velocita' di trasmissione.

	Spedizione al server	Ricezione dal server
TCP ( buf 65K )	897 Mbit/sec	916 Mbit/sec
UDP ( buf 8K )	956 Mbit/sec	485 Mbit/sec

Le misure sono state effettuate con il prodotto nuttcp, che sara' descritto in seguito, senza applicare alcuna ottimizzazione particolare ma usando i valori di default. Si noti come le prestazioni sono molto diverse nel caso dell'UDP con buffer piccoli, dove emerge l'importanza della Cpu. Nel caso peggiore, cioe' quando era il sistema client a ricevere i dati dal server la perdita di pacchetti e' stata del 45% a causa del sovraccarico di Cpu.

Se dai test di velocita' memory-to-memory passiamo a test di trasferimenti di file di grandi dimensioni, il collo di bottiglia diventa la velocita' dei dischi ed anche in questi casi emergono le differenze architetturali dei due sistemi. Un file (da 1 Giga) viene copiato localmente a 336 Mbit/sec sul server TYAN e a 128 Mbit/sec sul sistema trasportabile Acer. Se passiamo al trasferimento via rete abbiamo velocita' di 160 Mbit/sec dal'Acer verso server e 32Mbit/se in direzione opposta. Quindi i test che utilizzano dati su disco non sono assolutamente significativi a monitorare le prestazioni della rete in Gigabit.

## Installazione e configurazione

I due computer hanno il sistema operativo e i prodotti di monitoraggio gia' installati, si richiede quindi una configurazione minima al fine di integrare questi sistemi sulla rete Lepida. Per minimizzare le operazioni di configurazione sono gia' stati definiti i nomi dei due sistemi. Se necessario potranno essere ovviamente cambiati, ma questo comportera' la modifica di alcune tabelle.

I due sistemi si chiamano :

**lep-gtest**     il Server U1 Tyan  
**lep-clnt**     il Client trasportabile Acer

Questi nomi sono utilizzati negli esempi di questo manuale e nella documentazione on line che si trova sul client. Potrebbe risultare utile inserire questi nomi nel dns affinche' il server possa essere identificato univocamente sulla rete con un nome del tipo lep-gtest.lepida.it

### Configurazione del server ( lep-gtest )

Al server deve essere assegnato un indirizzo IP fisso, su rete pubblica, oppure su una rete privata che risulti comunque visibile ed accessibile da tutti i punti in cui si intende effettuare il monitoraggio; sicuramente la scelta di utilizzare un numero di rete pubblico e' preferibile per garantire una visibilita' totale.

Sul server vanno impostati i nuovi numeri di rete sostituendo quelli attualmente presenti (192.167.165.50). Se necessario puo' essere modificato il "firewall locale" modificando in modo opportuno il file /etc/sysconfig/iptables, ma bisogna garantire che le porte IP da 5000 a 5030 rimangano aperte, perche' queste porte sono utilizzate dai programmi di monitoraggio. Sul server non sono abilitati servizi di rete oltre a echo, ssh e quindi il sistema si puo' considerare ragionevolmente sicuro per rimanere in area pubblica.

### **Configurazione del client ( lep-clnt )**

Il sistema client puo' utilizzare numeri IP dinamici e di rete privata, ma i Nat o firewall attraversati per dialogare con il server introdurrebbero sicuramente errori e scompensi nella misura delle prestazioni. La soluzione ideale per misurare le velocita' degli apparati ottici ed elettronici che operano a livello OSI 1 e 2 e' quella di assegnare al client un indirizzo nella stessa rete IP del server. Misure effettuate da reti IP diverse ci danno invece indicazioni della velocita' sul percorso reale che comprende anche i router, e a queste velocita', molti router possono apparire come colli di bottiglia critici.

Potranno essere assegnati numeri IP diversi al client a seconda di dove verra' portato e se si vorranno testare le prestazioni del solo cablaggio oppure si vorranno verificare percorsi complessi.

Con il Menu' Application à System settings à Network ( o lanciando da finestra terminale comando : /usr/bin/system-config-network ) e' relativamente semplice configurare gli indirizzi di rete. Usando questa applicazione (o agendo sul file /etc/hosts ) si consiglia di definire tra gli host il server lep-gtest con il corretto numero ip.

Anche sul client puo' essere opportuno modificare le impostazioni del "firewall locale" agendo sul file /etc/sysconfig/iptables, e anche in questo caso bisogna garantire che le porte da 5000 a 5030 siano aperte.

Una volta cambiati i numeri IP il sistema e' pronto per la diagnostica della rete.

### **Gli strumenti di monitoraggio**

Sul sistema fisso sono sempre attivi due applicazioni server che possono essere comandati dal sistema mobile, sono i server delle applicazioni **iperf** e **nuttcp**. Inoltre e' possibile lanciare manualmente il programma **ttcp** da una finestra terminale.

Sul client sono presenti le stesse applicazioni in Linux , mentre una versione di iperf e' disponibile anche nel sistema windows XP.

I programmi utili per il monitoraggio della rete sono stati raccolti e messi in evidenza nel directory NET-monitor in cui sono presenti le seguenti applicazioni.



Nella finestra sono messi in evidenza gli applicativi utili per il monitoraggio di rete e la directory INFO che contiene documentazione. Gli applicativi sono i seguenti:

**Ethereal** : Il più famoso “sniffer” open source. Il software è in grado di catturare e analizzare tutti i pacchetti che circolano sulla rete.

**Network Tools**: Una applicazione che fornisce un'unica interfaccia grafica ai più comuni programmi di utilità della rete quali ping, traceroute, netstat, nslookup etc.

**XNmap** : Per analizzare lo stato delle porte IP di un sistema remoto.

**Terminal**: per aprire un terminale alfanumerico per utilizzare comandi in modalità testuale.

**Pesi-e-Misure** : Contiene file di dimensioni standard ( 1 Giga, 100Mega etc) quali unità di misura per i trasferimenti.

È presente inoltre la finestra grafica per la gestione interattiva del programma **nuttcp**.

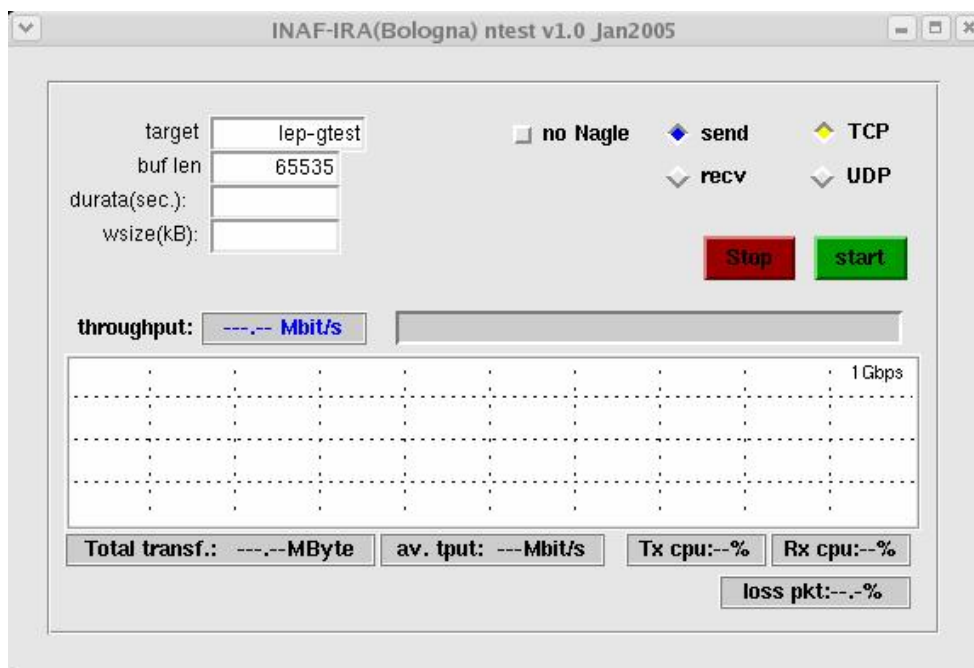
Tutte le applicazioni installate sono comunque contenute nel directory /root del sistema client.

## Il programma nuttcp

L'applicazione grafica, realizzata dal nostro Istituto, si basa su un programma di test molto diffuso conosciuto come "nuttcp". Nuttcp è una evoluzione di ttcp e rispetto a questo ultimo prodotto offre il vantaggio che può essere comandato da un solo estremo della connessione. Per questo richiede che sia presente sulla rete un server in grado di effettuare i trasferimenti e le relative misure. In appendice A sono presenti le informazioni necessarie per installare nuttcp e l'interfaccia grafica su altri sistemi.

Per lanciare l'applicazione in modalità alfanumerica si può utilizzare il comando nuttcp , seguito del nome del calcolatore su cui gira nuttcp in modalità server (esempio nuttcp lep-gtest). In

modalita' grafica invece si puo' lanciare l'applicazione utilizzando le icone presenti sul desktop e nella finestra NET-monitor, oppure da una finestra terminale con il comando **wish ntest.tcl**. Si otterra' una finestra come la seguente.



E' possibile effettuare il test della velocita' della macchina (loopback) inserendo nel campo "target" il nome "localhost". In questo modo vengono monitorate le prestazioni di trasferimento massime della macchina che coinvolgono solo la CPU, il bus e la connessione della scheda di rete verso il bus. A titolo di esempio possiamo dire che il sistema di monitoraggio **lep-clnt** da' risultati intorno ai 6400Mbit/sec ed il server **lep-gtest** arriva intorno a 5800 Mbit/sec.

Si puo' quindi misurare la velocita' tra i due sistemi inserendo l'indirizzo della macchina bersaglio nel campo "target" della finestra senza aggiungere alcun altro parametro; (oppure si puo' usare da finestra terminale il comando **nuttcp lep-gtest** ).

Il test dura 10 sec e poi riporta una statistica finale, con anche i tempi di utilizzo cpu sia locale che remota. Si ricorda che la statistica e' sempre quella vista da chi riceve, quindi di default calcolata sul target e riportata per la visualizzazione. Se si abilita l' UDP il riepilogo riporta anche la % dei pacchetti persi.

Il test puo' essere prolungato a piacere inserendo il numero di secondi nel campo "durata". Si puo' fermare il test quando si vuole premendo stop. Si consiglia di attendere sempre la statistica di riepilogo prima di ripartire con altri test. Ottimizzazioni possono essere fatte con buf len e window size.

Il wsize incide poco in prove back/back ma e' importante quando si analizzano tratte di rete complessa .

Il "no Nagle" permette di evitare l'algoritmo di ripartenza soft da collisione. Vale solo nel caso TCP.

Si puo' anche invertire il senso del flusso dal target al client selezionando recv.

## Appendice A

### Installazione di nuttcp

Il software nuttcp e' distribuito in un tar file che si chiama nuttcp.tar e che contiene tre file:

- nuttcp Il programma eseguibile
- nuttcp.xinetd Il file di configurazione per xinetd
- ntest.tcl L'interfaccia grafica sviluppata in tcl/tk

Anche se non e' strettamente necessario si consiglia un'installazione simmetrica e completa del prodotto sul server e sui client.

Per prima cosa bisogna copiare l'eseguibile **nuttcp** in **/usr/local/bin** sulle macchine interessate (test e bersaglio). Poi occorre configurare la macchina bersaglio (o meglio ancora entrambe) perche' nuttcp giri come demone, lanciato automaticamente da xinetd.

Allo scopo aggiungere a /etc/services le seguenti due linee

```
nuttcp      5010/tcp      # nuttcp control port
nuttcp-data 5011/tcp      # nuttcp data port
```

e si aggiunge il file **nuttcp.xinetd** in **/etc/xinetd.d** col nome **nuttcp**.

[ si noti che le porte di default su cui lavora nuttcp sarebbero la 5000 e 5001. In questo caso sono state cambiate per poter utilizzare sulle macchine altre applicazioni che fanno uso di tali porte]

Facendo ripartire xinetd con il comando "service xinetd restart" ( oppure "kill -USR2 <pid xinetd> seguito da "serviceconf" per la distribuzione Suse ) si puo' avere nuttcp attivo. Installare nuttcp come servizio offre anche il vantaggio di poter filtrare l'accesso dei sistemi client attraverso host.allow e host.deny, In alternativa si potrebbe lanciare manualmente sulla macchina bersaglio con "nuttcp -S".

A questo punto e' sufficiente lanciare l'applicazione grafica **ntest.tcl** con il comando **wish ntest.tcl** (avendo ovviamente l'interprete tcl/tk disponibile) e comparira' la finestra di controllo dell'applicazione.

Il sito di riferimento per il prodotto e' il seguente : <ftp://ftp.lcp.nrl.navy.mil/pub/nuttcp/>

---

Giuseppe Maccaferri - INAF - Ist. di Radioastronomia, Bologna Italy

E-mail: [g.maccaferri@ira.inaf.it](mailto:g.maccaferri@ira.inaf.it)

phone +39 051 6965 831

Mauro Nanni - INAF - Ist. di Radioastronomia, Bologna Italy

E-mail: [m.nanni@ira.inaf.it](mailto:m.nanni@ira.inaf.it)

phone +39 051 6399408