NetPerF Consulting
more control for better performances !

Network performance monitoring – Real-time latency monitoring for trading and finance, ma non solo .

Indice

1. Introduzione

Velocità, velocità e ancora velocità di esecuzione ! La velocità è sempre più un must per chi vuole competere sui mercati globali del mondo, concetto particolarmente caro e vicino in modo particolare per chi opera nel trading e nel finance in genere, dove 1ms in più o in meno può fare la differenza tra un guadagno o un mancato guadagno.

La catena che compone la latenza complessiva è estremamente lunga, articolata e complessa e gli elementi, che possono concorrere con il proprio contributo a variare il comportamento complessivo percepito da un utente o da una macchina che esegue in modo autonomo ordini di acquisto o vendita sulla base di algoritmi che si alimentano dei dati che ricevono in near-real-time dai mercati su cui operano, sono molteplici e possono penalizzare sensibilmente l'azione finale.

In questo scenario non va dimenticato che la latenza, ovvero il delay necessario per spostare un dato da un end point ad un altro, risente di fattori che non possono essere completamente azzerati, in quanto insiti nelle tecniche stesse di trasmissione e dei media utilizzati (Fibra Ottica, Satelliti, ecc) ed inevitabilmente legati alle velocità del mezzo trasmissivo utilizzato ed alle distanze che si devono coprire, pertanto la latenza complessiva sarà inevitabilmente costituita da una componente costante ed ineliminabile.

Il problema è che a questa quota, si deve aggiungere una quantità di latenza variabile e non costante legata alla quantità dei dati in transito istante per istante, alla loro tipologia, al numero e tipo di code e buffers posti lungo la catena che separa i due end node in esame; per non parlare poi ovviamente dei ritardi insiti nella applicazione medesima ed ai protocolli di comunicazione utilizzati.

E' dunque auspicabile avere una chiara visione della latenza complessiva, misurata in più punti lungo la catena al fine di acquisire piena consapevolezza su di se identificando il più velocemente possibile qualsivoglia scostamento onde poter intervenire ponendovi rimedio.

Le tecniche utilizzabili si distinguono fondamentalmente in due diverse tipologie : passive ed attive.

2. Metodi di misura

Si definiscono passive quelle tecniche che sono in grado di misurare la latenza basandosi su dati di traffico reale, analizzando i flussi di traffico intercettandolo e catturandolo nei diversi punti di misura.

Si tratta di tecniche inevitabilmente invasive, per lo meno nella fase iniziale di attivazione e, richiedono l'utilizzo di tools hardware e software specifici e particolarmente sofisticati, non per questo complessi ma certamente dall'investimento economico non trascurabile. Se da un lato garantiscono accuratezza della misura ed estrema flessibilità, necessitando di analizzare i flussi di traffico reali, impongono oltre che una analisi preventiva legata alla architettura di rete su cui si deve operare, richiedono anche una fase necessaria per la loro installazione che inevitabilmente risulta essere invasiva. Si tratta dunque di soluzioni particolarmente idonee per essere utilizzate in modo definitivo e continuativo nel tempo, non certo per eseguire una campagna di misura sul breve periodo, soprattutto in tutti quei casi in cui i punti di misura sono molteplici ed ubicati oltre che in posizioni diverse sulla rete anche in luoghi diversi.

Diverso è il caso delle tecniche di misura di tipo attivo, ovvero di quelle soluzioni che basano i risultati delle proprie misure analizzando traffico di tipo sintetico generato in modo autonomo da probes, più o meno dedicati.

Si tratta di soluzioni facilmente implementabili non invasive, che possono essere dispiegate in modo massivo, segmentando la rete come meglio si preferisce; come ogni soluzione anche questa ha aspetti positivi e negativi, il lato negativo riguarda l'essenza della soluzione stessa, ovvero che le misure si basano su dati di traffico sintetico generato dai probes dispiegati lungo il percorso, dunque tale tecnica aggiunge traffico di misura al traffico reale.

Dovendo misurare la latenza di segmenti di rete, su cui fluisce il traffico reale, il traffico di misura non può che percorrere il medesimo percorso altrimenti non avrebbe senso la misura stessa.

Non si tratta ovviamente di traffico quantitativamente rilevante, ma comunque traffico aggiuntivo che gli elementi

della rete (dispositivi, interfaccie, code e buffers) sono chiamati a dover gestire.

Come sempre, nella scelta di una soluzione, si tratta di valutare i pro ed i contro rispetto al fine che si vuole raggiungere, ovvero il deployment di una soluzione di latency monitoring permanente e stabile nel tempo, o di una soluzione da utilizzarsi per una campagna di misura atta ad acquisire conoscenza sulla natura complessiva delle latenze in gioco ?

La risposta determina la soluzione da utilizzarsi.

a) Analizziamo le potenzialità di una Soluzione attiva

La tecnologia CISCO IPSLA è particolarmente indicata per questo scopo, grazie a specifiche funzionalità insite in test quali ad esempio "UDP-Jitter" sarà possibile misurare non solo al latenza complessiva, ovvero il NRT (Network Round Trip Time) ma anche il contributo per singolo percorso, ovvero da S-D (sorgente – destinatario) e D-S (destinatario – sorgente) per quanto attiene le seguenti metriche :

- Delay
- Jitter
- Packet Loss

A titolo puramente esemplificativo si riportano i seguenti grafici.

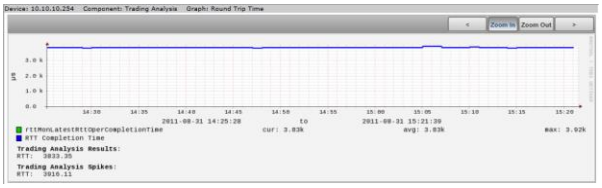

Figura 1

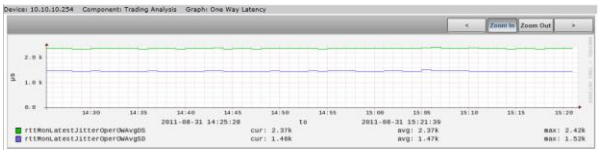

Figura 2

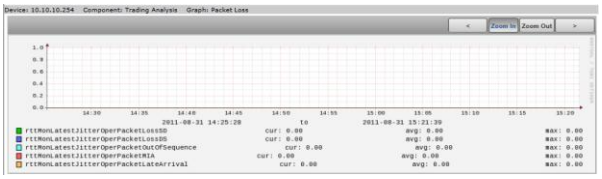

Figura 3

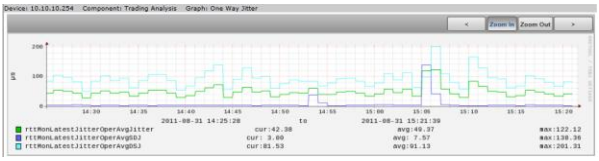

Figura 4

Le metriche possono essere misurate in millisecondi o in microsecondi (come nel caso dei grafici sopra riportati) consentendo di fatto una accuratezza di 100µs.
Pur essendo tale tecnologia disponibile embedded nella maggior parte degli apparati CISCO, è preferibile per motivi di accuratezza della misura nonché flessibilità ed indipendenza utilizzare i cosiddetti shadow-router, ovvero router singoli, configurati come semplici host, collocati nei punti di misura sui quali si applicheranno le configurazioni richieste dalle tecniche di misura necessarie per raggiungere l'obiettivo che ci si è preposti.

3. Notifiche

L'esigenza di rilevare un potenziale problema necessita non solo di un sistema di monitoring continuativo nel tempo di specifiche metriche ma anche di un sistema di notifica ed alerting in grado di attivarsi in real time o comunque in near-real time, al fine di segnalare la presenza di una situazione che potrebbe costituire un problema.

Un metodo estremamente efficace, laddove si utilizza una infrastruttura CISCO, è quello di utilizzare la tecnologia EEM (Embedded Event Management) in grado di consentire ad un device di inviare oltre che SNMP trap, anche email e di eseguire in modo automatico applet script precedentemente registrati.
Tali alert possono essere attivati da innumerevoli trigger associati a svariate metriche, a titolo esemplificativo si riporta, in figura 5, il testo di una email attivata da un evento inerente il NRT associato ad uno specifico test IP SLA.

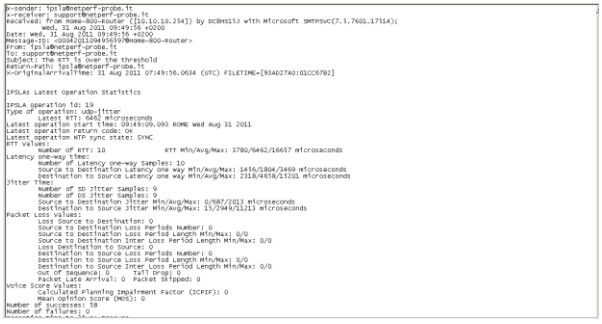

Figura 5

In questo caso, l'email contiene all'interno del suo corpo, il risultato di un comando CLI.

Ma nel caso classico in cui, si voglia ad esempio, ricevere una email al superamento di una soglia di utilizzo di una interfaccia, relativamente ad esempio al suo receive rate (Figura 6), si può oltre che ottenere una email contenente il risultato di un comando CLI specifico ricevere anche l'URL di un link associato ad una applicazione che sta ricevendo flussi netflow da quel router, consentendo in questo modo non solo di conoscere che una soglia quantitativa è stata superata, informazione tipica di un comando CLI o di un SNMP poller monitor, ma anche la qualità del traffico, ovvero quali applicazioni e quali conversazioni hanno scatenato l'evento, il tutto in modo automatico; a patto ovviamente di disporre di un sistema di Neflow collection in grado di gestire URI univoche associati ai propri grafici, è questo il caso ad esempio di NetFlow Tracker di Fluke Networks e di Network Performance Analyzer di Visual Network Systems.

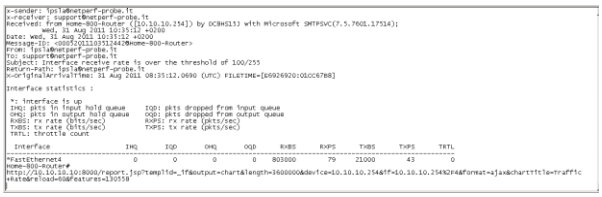

Figura 6

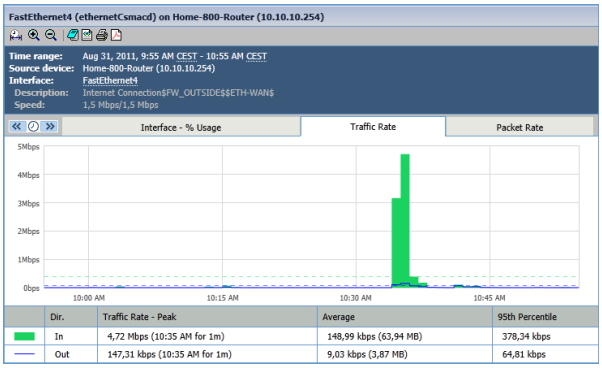

Figura 7

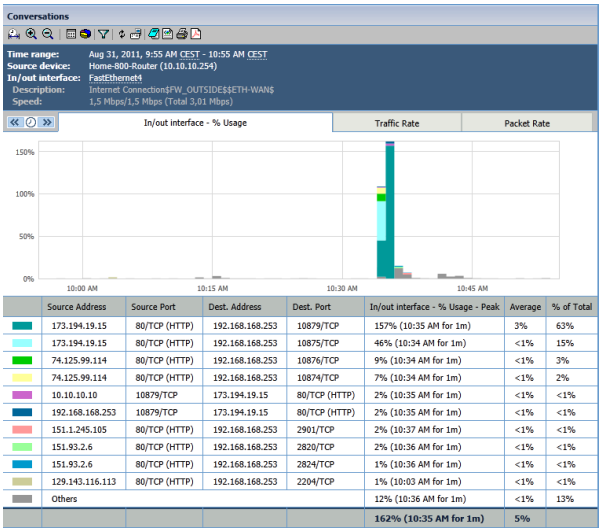

Figura 8

4. Conclusioni

L'utilizzo sapiente di tecnologie quali Cisco IP SLA e Cisco EEM, abbinate a tecniche di SNMP polling e di Netflow collection, consente la creazione di sistemi proattivi di monitoring relativi alle performance di infrastrutture, anche particolarmente complesse.

Autore : Maurizio Malinconi

NetPerF Consulting
more control for better performances !

Via A. Diaz, 30
20851 Lissone (MB)
Tel. +39 02 40047334
Fax +39 039 2781283
Email info@netperf.it
Web www.netperf.it