Journal of the Islmaic University of Gaza, (Natural Sciences Series) Vol.12, No.2, PP:157-183, 2004

INTERNET PERFORMANCE MEASUREMENT AND ANALYSIS

Mohammad A. Mikki, the Electrical and Computer Engineering (ECE) department at the Islamic University of Gaza (IUG), Gaza, Palestine.
Hala Al Jamal,Majeda Nabeel,Reem Abu Al Kass: Graduated student from the ECE department at the IUG, Gaza, Palestine.

ملخص : في هذا البحث نقوم بتصميم تجارب منهجية لجمع بيانات متعلقة بالانترنت لقياس معايير اداء الانترنت باستخدام بعض أدوات قياس كفاءة الانترنت الموجودة مثل traceroute و VisualRoute. نقوم بعد ذلك بتحليل نتائج هذه التجارب لتقييم مجموعة مختارة من خادمات الشبكة العنكبوتية العالمية (الويب) من أجل القيام بعملية توقع و تنبؤ لمشكلات مستقبلية محتملة و كذلك للاداء المستقبلي لتلك الخادمات. كذلك نقوم بعملية تصنيف و تقييم كمي للبيانات التي نحصل عليها و التي تشمل: زمن التأخير للأنترنت، معدل فقدان الحزم، و سعة البيانات المتوفرة من أجل عمل نموذج لتصرف الحركة المرورية عبر الأنترنت. كذلك نقوم بتحليل أداء مسارات الأنترنت تحت حالات مرضية مختلفة و من أجل رصد مدى استقرار تلك المسارات. هذه التجارب يتم استخدامها بواسطة مزودي خدمات الأنترنت من أجل تقصي تصرف خادمات الأنترنت و زمن الأستجابة على مستوى البرامج التطبيقية و كذلك من أجل مراقبة اتصال مسارات الأنترنت. بالأضافة لذلك نقوم بعرض طريقة جديدة لنمذجة و تحليل تصرف مسار الأنترنت بين الطرفين النهائيين لعملية الأتصال. هذه الطريقة مبنية على أساس ايجاد دالة الأنتقال لمسار الأنترنت بين الطرفين النهائيين لعملية الأتصال و هما خادم الويب و الزبون. النتائج الأولية فقط لهذا التصميم يتم عرضها في هذه المرحلة.
نتائج التجارب تتوافق مع تلك التي تم الحصول عليها في أعمال سابقة في نفس المجال و كذلك تشير الى أن طريقتنا المستخدمة هي طريقة فعالة لقياس أداء الأنترنت.

Abstract: In this paper we design methodical Internet data collection experiments to meausre the performance metrics of the Internet using some existing Internet performance measurement tools namely, traceroute and VisualRoute. These experiments are intended to evaluate different Web servers. We analyze the results of these experiments to predict possible problems and behavior. We characterize and quantize collected data including Internet latencies, packet loss rates, and available bandwidths. The characterized and quantized collected data is used to model traffic behavior. We also analyze the routing behavior for pathological conditions and routing stability. The results can be used by Internet customers or Internet Service Providers (ISPs) to track Internet servers behavior and response time at application level and to monitor Internet path connectivity. In addition, we present a new methodology to model and analyze end-to-end Internet behavior. This methodology is based on finding the transfer function of the end-to-end Interent path between the client and Web server. Only preliminary results of the design of this new methodology are presented at this time.

The results agree with other work done in the related field and show that our methodology is effective in measuring the Internet performance.

Keywords

measurement, performance, tools, metrics, ISP

1. Introduction

Most large Internet Service Providers (ISPs) currently collect basic statistics on the performance of their own infrastructure, typically including measurements of utilization, availability, and possibly rudimentary assessments of delay and throughput. In today's commercial Internet, the only baseline against which organizations can calibrate their networks is past performance data; no data or even standards formats are available against which to compare performance with other networks or against an industry norm, nor are there reliable data with which customers can assess performance of providers. Data characterization and traffic flow analysis are also virtually non-existent at this time, yet they remain essential for understanding the internal dynamics of the Internet infrastructure [15].

Increasingly, both customers and providers need information on end-to-end performance and traffic flow, beyond the realm of what is realistically controllable by individual network or user. Path performance measurement tools enable users and operators to better evaluate and compare providers and to monitor service quality [6].

Many emerging end-to-end path performance tools are intended to serve users, both for self diagnosis of problems they experience and for conducting measurements over the shared infrastructure, which can yield data with which to compare alternative providers and monitor service qualities. Many of these tools treat the Internet as black box, measuring end-to-end characteristics, e.g. packet latency and loss (ping) and reachability (tracroute), from points originating and terminating outside individual networks. Traffic flow characterization tools focus on the behavior and inner workings of these wide area networks.

With the fast growth of high performance networks, users pay more attention to network traffic conditions, especially when they are running a high-performance distributed application over wide-area network. A network performance report may help users figure out the quality of the network service and the available bandwidth, as well as identifying the traffic pattern on the network in order to avoid heavy traffic [4].

The fundamental limitation in the existing Interent model is the lack of network bandwidth for delivering synchronous media such as video and audio with the needed QoS performance, which are fundamental for applications in the consumer and business markets [8].

Indeed, it seems likely that the dynamic nature of the Internet will continue to render collected traffic data primerily of historical interest unless such data can lead to tangible improvements in our ability to analyze and predict network behavior. All the workload and performance data in the world will not get us very far without improvements in analysis, modeling and visualization, and simulation tools particularly those capable of addressing Internet scalability. The greatest challlenge, however is that the Internet is an “immense moving target”. It not only grows at an exponential rate, but it also undergoes dramatic qualitative changes over time as well [7]. Nowadays performance tools usually come pre-installed on almost all platforms, so there is nothing to install in client. The server (echo responder) runs at a high priority (like Kernel on Unix) and so is more likely to provide a good measure of network performance than a user application.

There are various components that contribute to web response. These components include: Round Trip Time (RTT), transmission speed, DNS delay, connection delay, server delay, transmission delay. Metrics of delay, packet loss, flow capacity and availability are fundamental to performance comparison tools with reasonably statistical validity have been slow to emerge.

Round Trip Time (RTT) is the time for a packet to make the round trip from a client to a server and back. Many components contribute to RTT [3]:

- The time it takes a packet to travel along the physical links that make up its path through the Internet (transport time)
- The time it takes to pass through the routers between those links (queuing and transmission time)
- The time required for the server to process an incoming packet an generate a response packet (server response time)

RTT can be an indicator of standing queue and congestion, not just distance. Changes in the RTT are usually an indication of configuration changes or changes in congestion levels. Each store and forward router hop adds the packet duration to the delay.

Some applications are sensitive to variance in transit time. Examples of such applications include real time video and audio, Video on Demand (VoD) and Internet chat. The response time or Round trip time (RTT) relates to the distance between sites plus the delay at each hop along the path between the sites.

Network packet loss is the fraction of packets lost in transit from client to server and back during a specified time interval [3]. Backbone Packet Loss Provides critical view of performance, the end user will feel it. Correlation can lead to the reasons for packet loss. The loss of packets is a good measure of quality of the length (in terms of it's packet loss rates) for many TCP based applications. Loss typically caused by congestion which in turn causes queues (in routers) to fill and packets to be dropped. Loss may be caused by the network delivering an imperfect copy of the packet, this usually is caused by bit errors in the links or in network devices. The ideal end result will be the ability to predict this before it happens and act to prevent it. When zero packet loss happened, the network is then called quiescent (non-busy).

Network bandwidth (also called throughput or bit rate) is the rate at which data is sent through the network, usually expressed in bits persecond (bps), bytes per second (Bps) or packets per second (pps) [3]. Network bandwidth continues to be a critical performance metric in the Internet because of heterogeneous bandwidths of access technologies and file sizes. Measuring network bandwidth over the Internet has become increasingly important, given the recent explosion in Internet services and systems: peer-to-peer, end system multicast, content distribution networks, server selection, placement of server replicas.

Knowledge of the bandwidth along a path allows an application to avoid mistakes by adapting the size and quality of its content or by choosing a web server or proxy with higher bandwidth than its replicas.

In some parts of the Internet, available bandwidth is frequently equal to bottleneck link bandwidth because either bottleneck link bandwidth is small (e.g. wireless, modem, or DSL) or cross traffic is low (e.g. LAN) [13]. Measuring the bottleneck link bandwidth along a path is important for understanding the performance of many Internet applications. Existing tools to measure bottleneck bandwidth are relatively slow, can only measure bandwidth in one direction, and/or actively send probe packets.

The scenario in Internet has moved from being used for purely education and research purposes to being used for either general purpose or specific services. This gives rise to the need of ability to find which server best serves a given request. The requirement of a request could be fast response time, memory at server, computation power at server etc. Most common among these requirements is fast response time. In this case, the general problem is to find the server from among a given or selected servers such that client's perceived response time with that server is least among the given servers. Client's perceived response time has two components: Network delays and queuing delay at the server. The idea is to find whether network delay is the major component of client's response time or queuing delay at server. If most of the time it is network delay then the client can safely assume that the network is bottleneck in accessing the server, and hence can neglect considering this server as a potential server to choose among. If the server is the bottleneck, i.e. has large queuing delays most of the time then, it is an indication to the service provider of the server to take measures like replicating servers for better service provision.

In this paper we design methodical Internet data collection experiments to meausre the performance metrics of the Internet using some existing Internet performance measurement tools namely, traceroute and VisualRoute. These experiments are intended to evaluate different Web servers. We analyze the results of these experiments to predict possible problems and behavior. We characterize and quantize collected data including Internet latencies, packet loss rates, and available bandwidths. The characterized and quantized collected data is used to model traffic behavior. We also analyze the routing behavior for pathological conditions and routing stability. The results can be used by Internet customers or Internet Service Providers (ISPs) to track Internet servers behavior and response time at application level and to monitor Internet path connectivity. In addition, we present a new methodology to model and analyze end-to-end Internet behavior. This methodology is based on finding the transfer function of the end-to-end Interent path between the client and Web server. Only preliminary results of the design of this new methodology are presented at this time.

The results agree with other work done in the related field and show that our methodology is effective in measuring the Internet performance. We will extend our work in the future to include the design of our own integrated measurement tool that automatically conducts, analyzes and characterizes performance metrics of the Internet and displays the results to the user for assessment and evaluation. This extension is based on our new methodology.
The rest of the paper is organized as follows: Section two presents related work. Section three presents the Internet performance tools used in our measurement experiments namely, traceroute and VisualRoute. Section four presents our new end-to-end Internet performance measurement methodology. Section five presents experimental results. And finally, section six concludes the paper.

2. Related Work

In this section we present some related recent research work that is conducted in the field of Internet performance measurement, analysis and modeling.
[14] presents a simple prototype implementation of a performance monitoring system for IP networks that applies both active and passive methods. The method is based on an appropriate calibration of traffic meters and dedicated monitoring packets. These monitoring functions could be an integral part of ordinary network elements, or stand-alone systems. The prototype is implemented in Linus-based routers and the main application is monitoring in IP-based virtual private networks.
[1] provides a preliminary assessment of the effectiveness of an application layer tool that measures the Bulk Trnsfer Capacity (BTC) of a network path. The goal of this work is to provide an empirical evaluation of a BCT tool and therefore assess the reliability of the measurements obtained using these tools. They conducted experiments across the MIMI mesh of measurement hosts to compare stock BSD TCP with the BCT measurement tool.
[11] measures the performance of two commercial content distribution networks (CDNs), one operated by Akami and one operated by Digital Island. Both CDNs redirect requests by using DNS. The work describes its simple measurement technique for a DNS_based CDN, the data for the two commercial services and the interpretation of that data. The main conclusion was that CDNs provide a valuable service, but that neither Akami nor Digital Island can consistently pick the best server of those available.

[2] presents an invistigative tool for the direct study of end-to-end Internet traffic, featuring a hyprid timing mechanism allowing high resolution timestamping of packet deparures in a UNIX environment. A method for the elimination of timing artifacts due to variations in clock rates, essential for the study of delay was also presented.
[16] reports on an analysis of 40,000 end-to-end route measurements conducted using repeated traceroutes between 37 Internet sites. The work analyzes the routing behavior for pathological conditions, routing stability, and routing symmetry.
[5] presents a parameterizable methodology for profiling Internet traffic flows at a variety of granularities. Its methodology defines flows based on traffic satisfying various temporal and spatial locality conditions as observed at internal points of the network.
[7] develops a scalable network modeling framework, a scalable simulation framework, and scalable discrete event simulators capable of modeling the Internet at unprecedented scales. These models make it possible to analyze the detailed behavior of large multidomain multiprotocol Internet models.
[12] Describes the measurement system for the Internet traffic and shows some aspects (such as number of packets, amount of transmission, connection time and starting time) of the Internet with the system.
[17] presents a study to better understand the key characteristics of the request patterns in high-volume Web server environments and to better understand the impact of such traffic patterns on Web server performance.
The Internet Engineering Task Force (IETF) IP performance metrics (IPPM) working group was chartered to develop theoritical framework and guidelines for designing robust measurement tools for the Internet's of disparate signal sources. In late 1996, draft comments were issued, delinating metrics for connectivity, one-way delay, and impirical bulk transfer capability. In the mids-1997, the National Science Foundation (NSF) funded the creation of the Cooperative Association for Internet Data Analysis (CAIDA). CAIDA is a collaborative undertaking to promote greater cooperation in the engineering and maintainence of a robust, scalable global Internet infrastructure [6].

3. Internet Performance Measurement Tools

Network measurement and data collection techniques can be genrally classified into two categories: intrusive measurement (also referred to as active probing) and non-intrusive measurement (also referred to as passive probing). Active probing requires injection of a sequence of test packets into the network and obtain network performance information by analyzing the behavior of the feedback probing packets. On the contrary, passive watch can infer network status by passively observing network traffic traversing through measurement points [9]. There are a number of tools available on the Internet that allow examination of the performance and/or reliability of single Internet paths. For example traceroute reveals the actual path that packets on their way from one host to another currently take. Pathchar determines several parameters for this path, like delay and bandwidth for each hop. The well-known tool ping tests whether basic connectivity between two hosts is available. Tools like ttcp, netperf, and treno measure the bulk transfer capacity of a path [10]. A common problem in running applications over a network is knowing what route the network traffic is taking. There are often many different possible paths through the network, some are very fast and some are quite slow. Even on high-performance networks the route may be very roundabout.

In this section we preseent the Internet performance measurement tools used in our research to measure different performance metrics including Intenet bandwidth, bottleneck bandwidth, packet loss rate, and round trip delay. These tools are traceroute and VisualRoute. In section 3.1 we present traceroute, and in section 3.2 we present VisualRoute.

3.1 Traceroute

The traditional traceroute program provides a solution to knowing the route through the network. It lists each machine, called a router, which the network traffic goes through. However, interpreting a traceroute output requires knowledge which only a network engineer has. For instance, most users don't know where the routers are physically located, so they cannot decide whether a particular route is good or bad. The objective of traceroute is to print the route that packets take to network host. Traceroute works as follows: IP packets contain a time-to-live (TTL) field that is initialized by the original sender. Time-to-live field is decremented by one at each intermediate router. If the field is decremented to zero, the packet is discarded and an error indication packet (an ICMP "time exceeded") is sent back to the original sender. The source address of the ICMP "time exceeded" identifies the router that discarded the data packet. So if packets are sent to the final destination but with the TTL set to n , the router n hops along the path is forced to identify itself. Traceroute works by causing each router along a network path to return an ICMP error message [18].

The traceroute program sends its first group of packets with a TTL value of one. The first router along the path will therefore discard the packet (its TTL is decremented to zero) and return the TTL Exceeded error. Thus, we have found the first router on the path. Packets can then be sent with a TTL of two, and then three, and so on, causing each router along the path to return an error, identifying it to us. Eventually either the final destination is reached, or the maximum value (default is 30) is reached and the traceroute ends [18]. At the final destination, a different error is returned. Most traceroute programs work by sending UDP datagrams to some random high-numbered port where nothing is likely to be listening. When that final system is reached, since nothing is answering on that port, an ICMP Port Unreachable error message is returned, and we are finished.

In the unlikely event that some program happens to be listening on the UDP port that traceroute is trying to contact, the trace will fail at the last hop. You can run another trace using ICMP Echo Requests, which will probably succeed, or specify a different target port for the UDP datagrams.

3.2 Visualroute

VisualRoute addresses the problem of knowing where the routers in the path found by traceroute are physically located by plotting a traceroute onto a geographic map so users can decide whether a particular route is good or bad. It maps the routers to physical locations. VisualRoute is a graphical tool that determines precisely where and how traffic is flowing on the route between the desired destination and the user trying to access it, providing a geographical map of the route, and the performance on each portion of that route. VisualRoute combines ping, whois, and traceroute all into one program with the point and click ease of a graphical interface [19]. VisualRoute is available on Windows, Solaris, Linux, FreeBSD, and Mac OS X platforms. You can install VisualRoute 6.0b from www.visualware.com web site. VisualRoute can be

configured to run as a (Web) server, providing visual trace route capabilities to anyone with a Java-enabled web browser as shown in Figure 1.
Configuring VisualRoute to run as a Web server is very useful when:

- You want to provide the VisualRoute service to your clients.
- You may not want to install VisualRoute on lots of client machines. Installing it on one server and granting access to people allows you to control access.
- Your company has a very strict firewall that blocks any traces, so how do you diagnose network connectivity problems? Install VisualRoute on a server outside the firewall and grant access to people in the company.
- You want to provide visual traceroute capabilities to your entire company.
- There are many web sites that provide this service. Some of these Web sites are listed in Table 1.

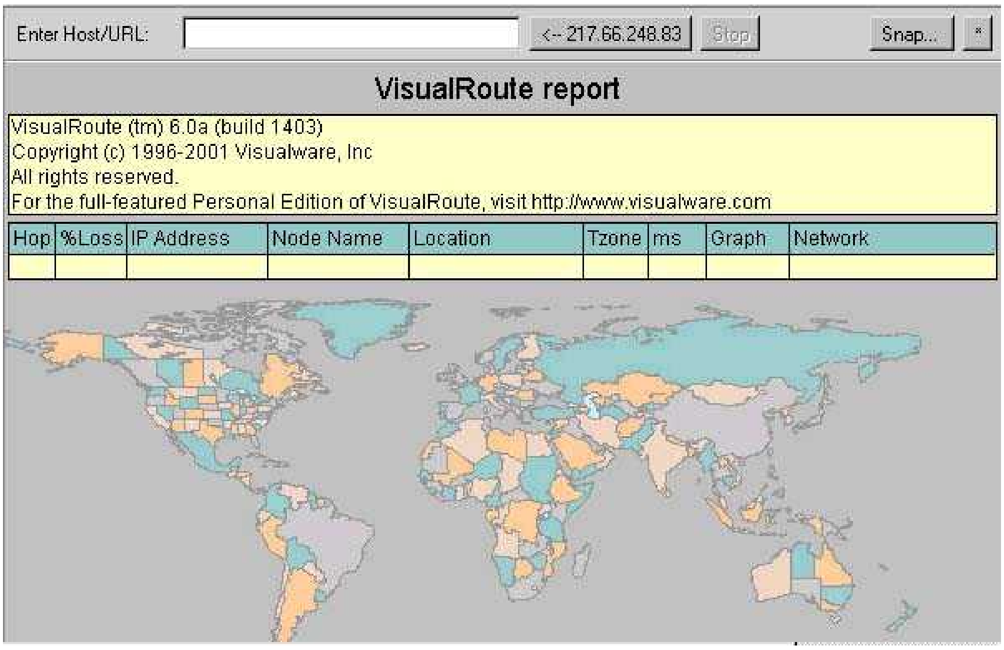

Figure 1: VisualRoute Server from http://visualroute.visualware.com

Table 1: Web sites that support visual server

http://www.visualroute.it/vr.asp
http://www.webhits.de/visualroute/
http://visualroute.imagine.com.co/
http://visualroute.bboxbbs.ch/
http://visualroute.visualware.com/

To use VisualRoute you enter the host and optional port that you want to monitor into the edit box in VisualRoute as shown in Figure 2. A sample result of VisualRoute is shown in Figure 3. The columns in the output of VisualRoute are listed in Table 2.

Figure 2: Host edit box in VisualRoute

The second part of VisualRoute displays the route on a map of the World. Links drawn in blue indicate known locations. Links drawn in purple indicate that a guess was made.Also you can zoom the map using the mouse ,left mouse click produce zoom in, right mouse click produce zoom out and mouse drag moves the map around [19]. In general, text that is displayed in purple is non-authoritative (from a local cache on your hard drive maintained by VisualRoute) and text displayed in black is authoritative. Non-authoritative information is validated as soon as possible, but is displayed immediately to speed up the display of information to you.

4. Transfer Function-Based Methodology For Internet Performance Measurement

In this section we present a new methodology for the end-to-end Internet path performance prediction based on the calculation of the transfer function of the end-to-end path between the client and server (See Figure 4).
The objective of this methodology is to cost-effectively analyze and predict the end-to-end behavior of the Internet path between a client and a server. Our approach is based on the transfer function (TF) that describes the end-to-end path. TF is a mathematical function that quantifies the path behavior with respect to the desired metric (delay, packet loss rate, bandwidth, etc.) when this path and its environment attributes change. As input variable, x (input traffic, packet size, etc.) from the client changes, the performance metric under study, y

(e.g. delay, packet loss rate, bandwidth, etc.) changes. The transfer function TF is then calculated as:

$$TF = y/x$$

TF can be calculated using different values of x and measuring the corresponding y values. TF can be constructed for the performance metric of interest using measurements.

After knowing TF, we can predict the behavior of the path without conducting the measurements. Since $y = TF * x$, we can find y for any value of x without going through the measurement. This enables us to get the predicted value of y based on the transfer function of the end-to-end path behavior.

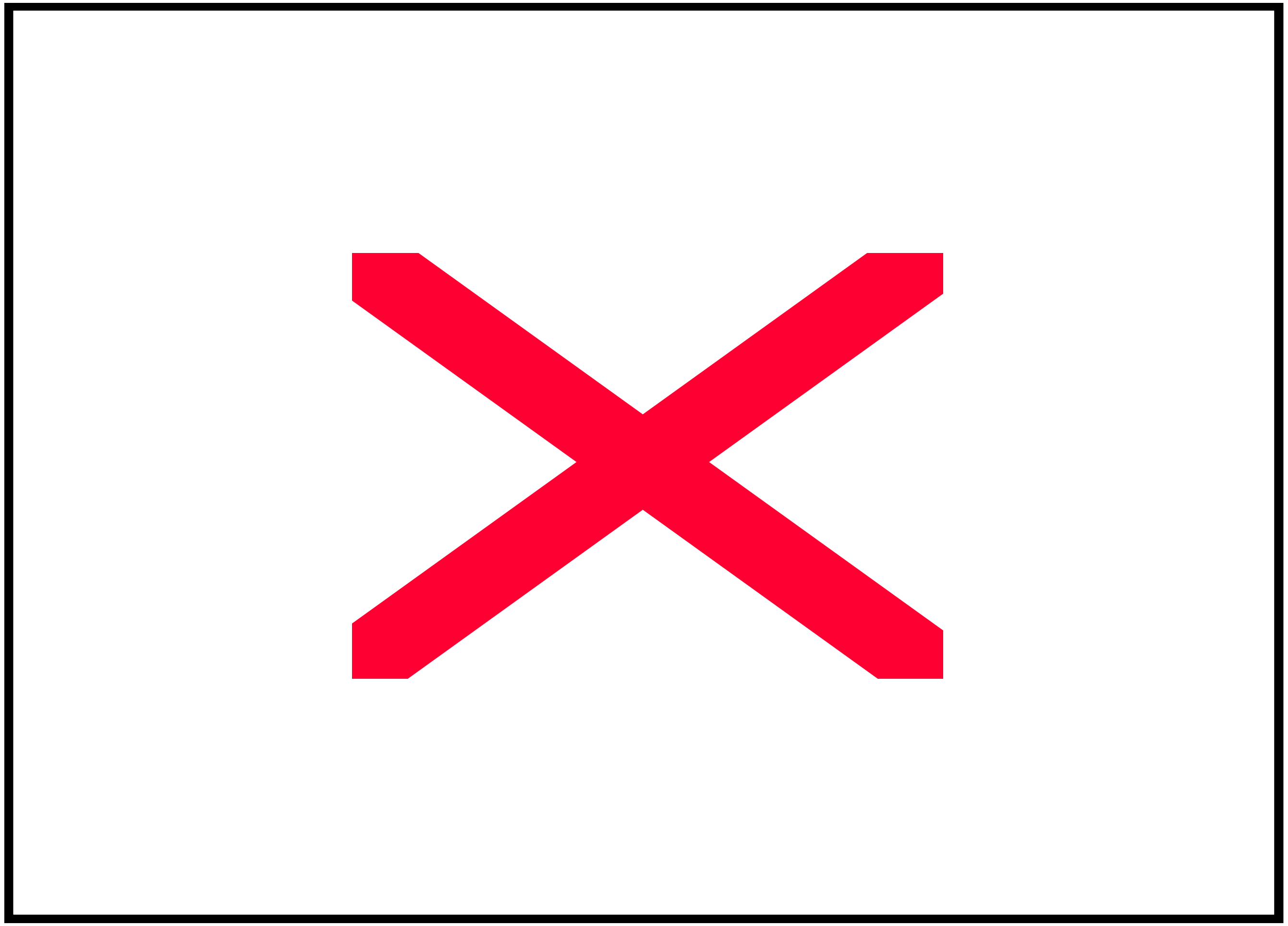

Table 2: Columns in the output of VisualRoute

Column	Description
Hop	The traceroute hop count.
% Loss	Percentage of ping packets that have been lost at this hop level. This number is red if a host is detected at this hop level. Otherwise, this number is black, because the host may just be ignoring all of our ping packets.
IP Address	The IP address of the host at this hop level. It is red if a recent ping packet was lost. It is blue when it matches a Loose Source Route IP address. Otherwise, it is black. The IP Address is bolded if it is the host that you are trace routing to.
Node Name	The reverse DNS lookup of the IP Address.
Location	The geographical location of the node. It is black if we are fairly certain of the node's location. Otherwise it is purple, meaning that the location is a best guess based upon information from various WHOIS databases.
Tzone	If the location of the node has been resolved then the time difference to your selected timezone will be displayed. You can change the base timezone that VisualRoute uses in `Options -> Preferences -> Analysis Columns`
ms	The average number of milliseconds (roundtrip) that it took for a ping packet to go from your machine to this hop level and back to your machine
Graph	The blue line is a graph of the 'ms' column. The gray horizontal bar in each hop line represents the minimum and maximum millisecond times for that hop level.
Network	The network that this node is in, as reported by various WHOIS databases.

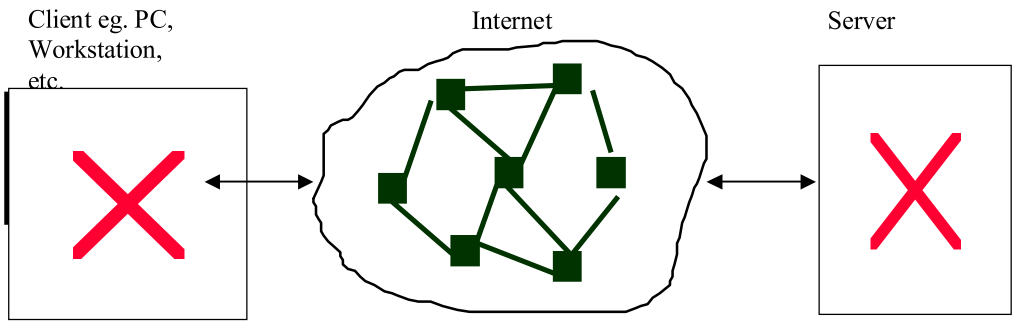

Figure 4: Modeling the end-to-end Internet path between a client and a server

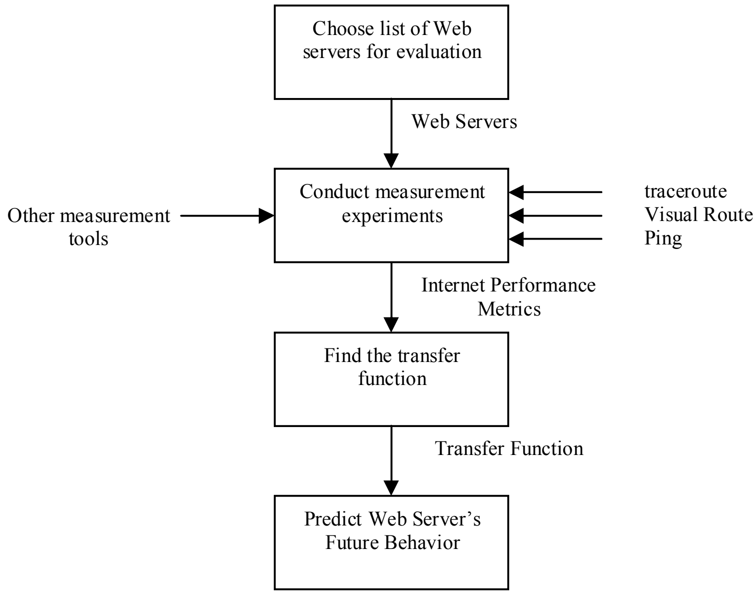

Figure 5: Outline of a New End-to-End Internet Path Performance Prediction Methodology

The outline of this methodology is listed in Figure 5. First, we choose a list of Web servers whose performance behavior is to be analyzed. Second, we use performance measurement tools to get different performance metrics (packet loss, bandwidth, round trip time delay, etc.). Third, we use the collected data of the performance metrics to calculate the transfer function of the path behavior. This characterizes the path under various load conditions to identify its bottlenecks and find design alternatives. And finally, we use this transfer function to predict future behavior of the server when it is accessed from a given client. The performance measures predicted by the methodology are validated against actual measurement.

5. Experimental Results

In this section we present some priliminary experimental results for the performance metrics of the Internet using traceroute and VisualRoute tools. These results and analysis reported are preliminary and incomplete. We will present characterization models of the collected data to characterize Internet traffic for different Web environments.

Results of conducted experiments to obtain performance metrics (packet loss rate and RTT) for nine Web sites are reported. We chose these Web sites from different places to note the effect of distances on RTT and packet loss. The resultant data and the paths which the packets take for each site during six different days at four different daily times for each day are reported. A PC located in Gaza in Palestine which is connected to the Internet through a 56kbps modem is used to access these Web sites.

The evaluated sites are:

www.aitech.ac.jp (Japan)
www.cic.gc.ca (canda)
www.emirates.net.ae (Emarates)
www.erf.org.eg (Egypt)
www.forumhabitat.polito.it (Italy)
www.google.com (USA)
www.japan.com (Japan)
www.mbendi.co.za (Zambia)
www.nzsingles.co.nz (Newzeland)

Due to the huge amount of data obtained for these sites, only a sample of the collected data of two sites is reported, namely, www.erf.org.eg in Egypt and www.google.com in USA.

Figure 6 shows the output of VisualRoute for evaluating www.erf.org.eg. Figure 7 shows the output of VisualRoute for for www.google.com.

The average percentage of packet loss for www.erf.org.eg is 30% which is much higher than that of www.google.com which is 15%. In routers located in USA we see that the percentage of packet loss is much less than these routers located in Israel. An interesting observation about the end-toend path from Gaza to www.erf.org.eg server in Egypt is that it passes through routers in different countries (Italy, Germany, and USA) geographically not in the middle between Gaza and Egypt which are nieghbors.

Results of the multiple experements were also used to study the interval-temporal-connectivity of paths. Internet path connectivity means that the server was reachable when the access to it was made. Internet path connectivity for www.erf.org.eg was about 90%. Internet path connectivity of www.google.com was about 95% . When a routing failure happens this may be due to intermediate router outage or because the router is rearranging its view of the current topology and is dropping many incomming packets because it does not know how to forward them. The above results agree with results obtained in [10].

In regards to routing stability (do routes change often, or are they stable over time), we found that Internet paths are heavily dominated by a single prevalent route. We found also that routes change over a wide range of time (i.e., more than 6 hours). The wider the range of time between two measurements the more likelihood that the two routes will differ. In our case, the time intervals of the different route measurements at each day were less than 6 hours. This resulted in stable routes for different measurements of the same day but in different routes for measurements at different days. This coincides with results obtained in [16]. In regards to the number of hops in the path, figures 6 and 7 show that it was 20 hops for www.erf.org.eg and 19 hops for www.google.com. This is rather strange since it was expected that number of hops in a path should be proportional to the geographical distance as reported in the results in [16].

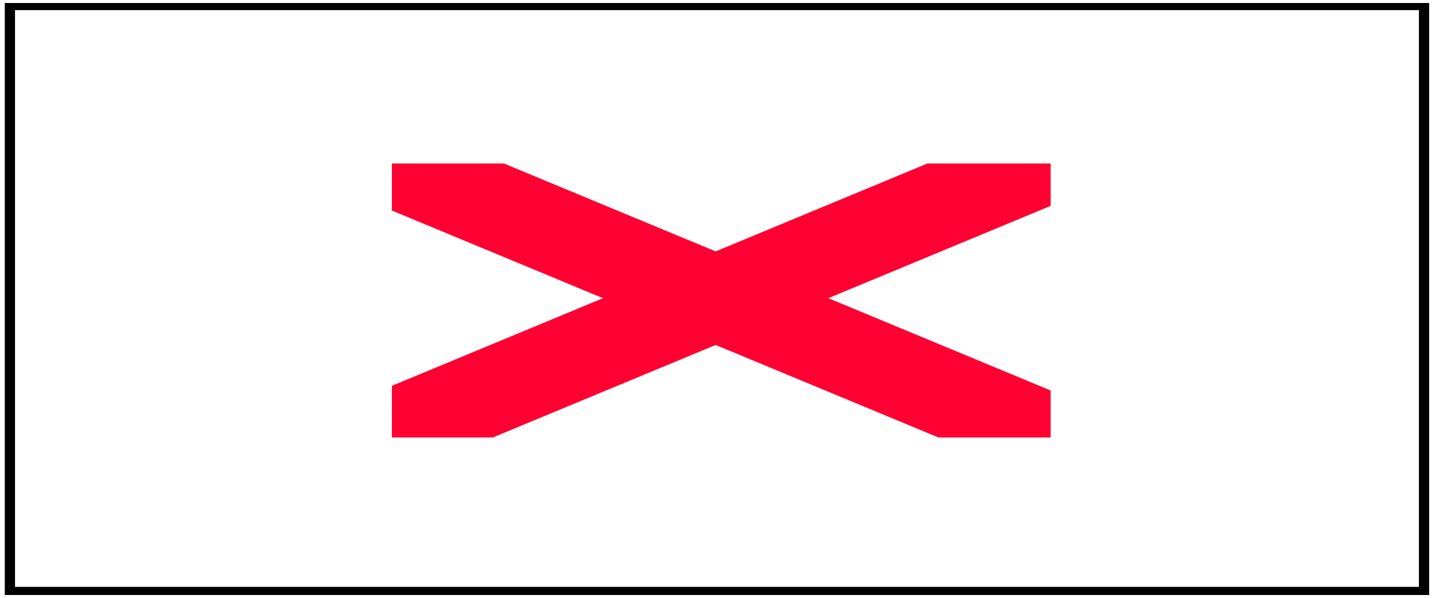

Figure 6: Output of VisualRoute for www.erf.org.eg

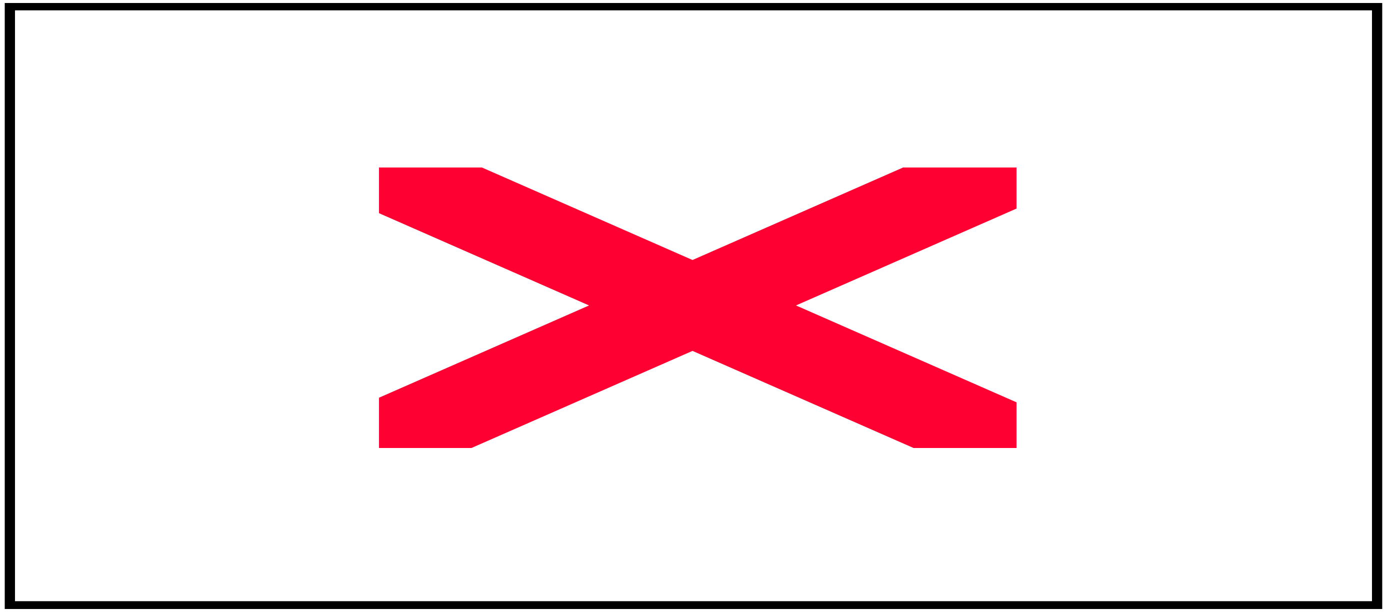

Figure 7: Output of VisualRoute for www.google.com

But our results coincide with the exception to the above rule as also reported in [16]. [16] observed that a 1500 km end-to-end route had only 3 hops, and a 200 km route had only 5 hops. They also found that the route between Massachusetts Institute of Technology in Massachusetts in USA and Harvard University also in Massachusetts in USA (about 3 km apart) was consistantly 11 hops!.

Figure 8 shows the packet loss rate for accessing www.erf.org.eg. Figure 9 shows the packet loss rate for accessing www.google.com. The curve for www.google.com is smoother than that of www.erf.org.eg . This is most likely due to the better service provided by the www.google.com server. Packet loss rate at daytime hours for both Web servers (note the time zone difference between Egypt and USA) is much higher than measurements at other times of the day. This shows that packet loss has a high dependency on the time of the day and the day of the week. This also agrees with results obtained in other work. During working hours more users connect to the Internet and generate more traffic which lead to router congestion. Routes respond by dropping more packets. During the evening and early morning there is a lower Internet traffic. ISPs may would like to use this conclusion to make configuration changes and repairs during early mornings.

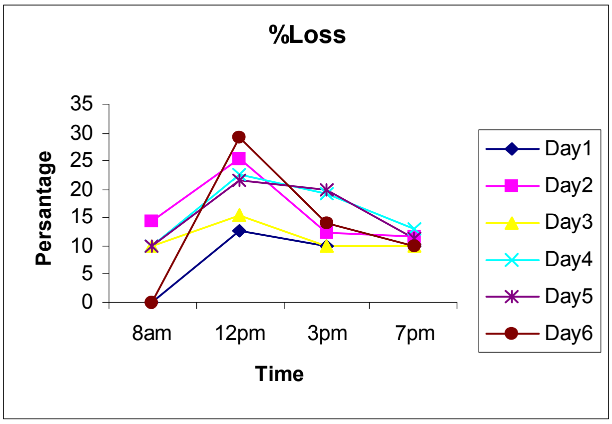

Figure 8: Packet loss rate for www.erf.org.eg

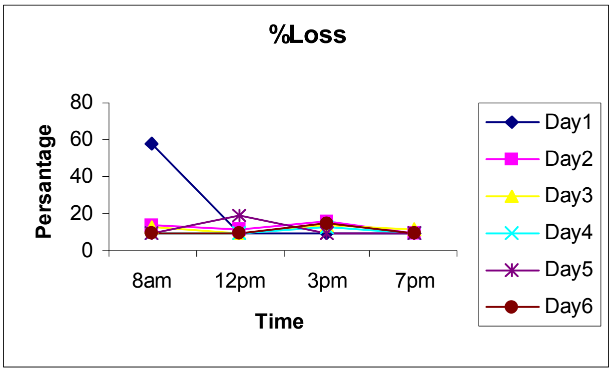

Figure 9: Packet loss rate for www.google.com

Figure 10 shows the RTT corresponding to accessing www.erf.org.eg. Figure 11 shows the RTT corresponding to accessing www.google.com. The two figures show that RTT has a high dependency on the time of the day and the day of the week. This also agrees with results obtained in other work. During working hours more users connect to the Internet and generate more traffic which lead to router congestion. Queues in the routes are filled up. This leads to slowdown of connection and increase in latency of packet delivery.

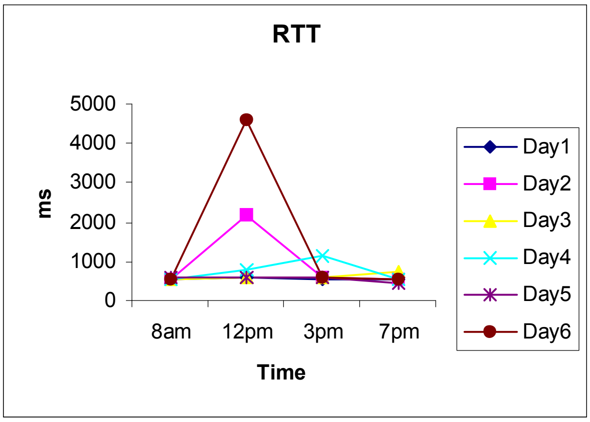

Figure 10: Round trip time (RTT) for www.erf.org.eg

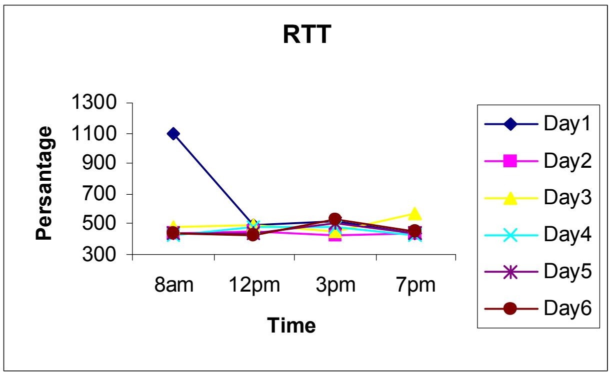

Figure 11: Round trip time (RTT) for www.google.com

Figure 12 shows the geographical map for the route from the client to the www.erf.org.eg server. Figure 13 shows the geographical map for the route from the client to the www.google.com server.

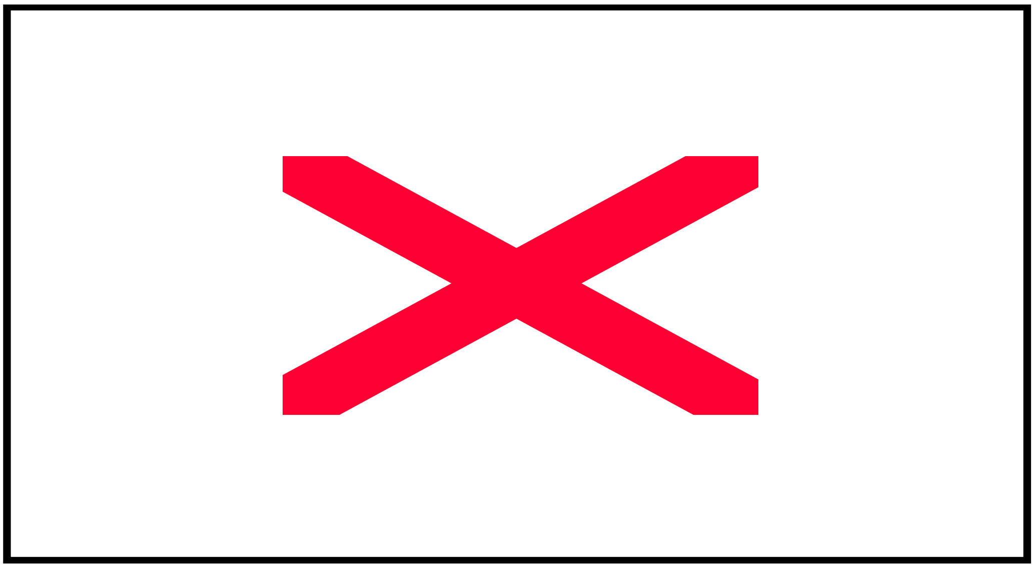

Figure 12: Geographical map for the route to www.erf.org.eg

In Figure 14 we show the result of using traceroute to evaluate www.mbendi.co.za Web server. The first line of the output is the information about what we are doing; it shows the target server, the target server's IP address, the maximum number of hops that will be allowed, and the size of the packets being sent. Then we have one line for each system or router in the path between the client PC and the target server. Each line shows the name of the router (as determined by the DNS), the router's IP address, and three round trip times in milliseconds.

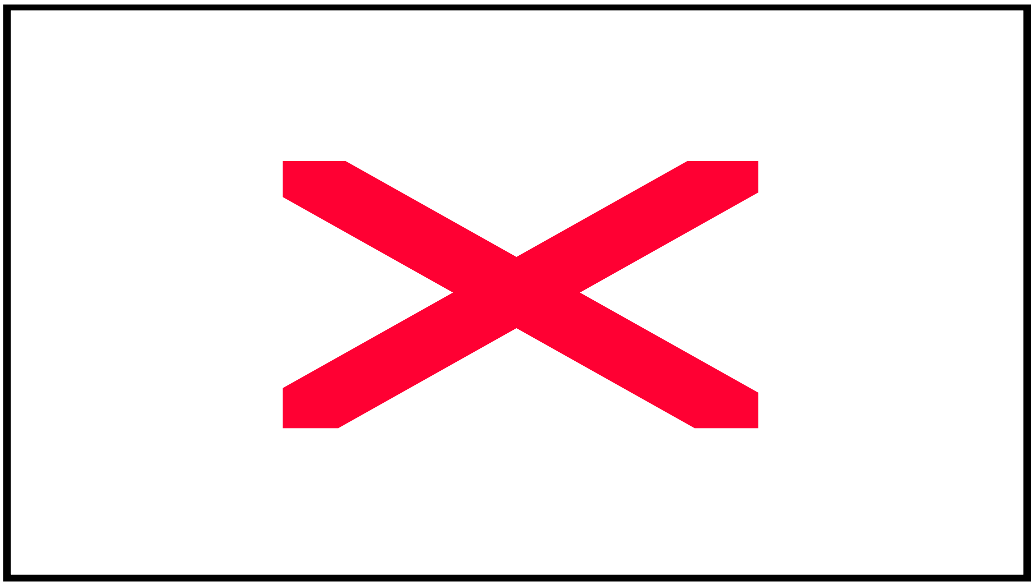

Figure 13: Geographical map for the route to www.erf.org.eg

The round trip times (RTTs) tell us how long it took a packet to get from the client PC to that router and back again. By default, three packets are sent to each router along the route, so we get three RTTs. The analysis of the collected data and the corresponding routes show that the values of the collected data are affected by different factors including the model of use and Web server environemt traffic. As results show, requested traffic highly impacts Web server performance including QoS, scalability, capacity planning, availability and reliability.

The client PC speed and the speed of the used modem seem to influence the measurement quality for both traceroute and VisualRoute– the slower PC tends to show lower numbers for the conducted tests. The accuracy of the tool for a certain Web server depends on the distance to the Web server. The more hops from the client to the tested Web site, the worse measurements are achieved.

```
traceroute to www.mbendi.co.za (196.38.129.139), 30 hops max, 38 byte
packets
1  217.21.0.33 (217.21.0.33)  3.109 ms  2.441 ms  2.257 ms
2  host-217-21-5-1.p-i-s.com (217.21.5.1)  12.168 ms  26.296 ms  21.388 ms
3  217.21.0.9 (217.21.0.9)  21.122 ms  14.178 ms  9.314 ms
4  212.14.225.97 (212.14.225.97)  421.825 ms  429.841 ms  426.468 ms
5  212.150.58.165 (212.150.58.165)  423.685 ms  420.835 ms  424.700 ms
6  barak-1-bb-backbone.barak.net.il (206.49.94.115)  411.942 ms  426.500 ms
576.998 ms
7  bb2.ser3-0-1.barak.net.il (212.150.4.178)  577.623 ms  429.984 ms  418.792
ms
8  212.150.232.129 (212.150.232.129)  448.362 ms  436.945 ms  439.860 ms
9  212.150.232.121 (212.150.232.121)  410.727 ms  428.732 ms  411.727 ms
10  bb-ny.ser5-0-0.barak.net.il (212.150.232.65)  453.549 ms  420.868 ms
412.739 ms
11  500.Serial3-8.IG3.NYC4.ALTER.NET (157.130.254.121)  440.131 ms
431.490 ms *
12  521.at-5-1-0.XR3.NYC4.ALTER.NET (152.63.24.34)  419.646 ms
415.593 ms  472.601 ms
13  509.ATM6-0.GW11.NYC4.ALTER.NET (152.63.22.193)  427.201 ms
430.085 ms  428.579 ms
14  iscoza-gw.customer.alter.net (157.130.22.182)  454.869 ms  429.306 ms
446.682 ms
```

Figure 14: Output of traceroute when evaluating www.mbendi.co.za

```
15  * 168.209.244.2 (168.209.244.2)  422.910 ms  409.383 ms
```

```
16  168.209.0.29 (168.209.0.29)  534.034 ms  421.846 ms  417.214 ms
17  168.209.0.25 (168.209.0.25)  987.279 ms  990.181 ms  976.065 ms
18  ch1-ctn.nmszone.is.co.za (168.209.2.129)  967.596 ms  975.651 ms
1013.624 ms
19  168.209.20.38 (168.209.20.38)  1252.481 ms  1310.714 ms  1802.290 ms
20  * * *
21  * * *
22  * * *
23  * * *
24  * * *
25  * * *
26  * * *
27  * * *
28  * * *
29  * * *
30  * * *
```

Figure 14: Output of traceroute when evaluating www.mbendi.co.za (continued)

Knowing the values of the Internet performance metrics allows systems and people to be much smarter about the decisions that they make. Some examples of applications are:

- Benchmarking: A user can determine if some network equipment/service is really delivering the claimed bandwidth.
- Web cache selection: Given several web proxy caches, a client could use bandwidth measurements in selecting the cache that would give the best performance.
- Multicast tree construction: Multicast routers could use bandwidth measurements in selecting a path to join a multicast tree.
- Interface selection: A multi-homed host could use bandwidth measurements to select which of several network interfaces to use for a packet.
- Protocol tuning: Someone tuning a transport protocol could use the measured bandwidth as a guideline for the maximum possible throughput he or she could achieve with the transport protocol.

- Application-level network adaptation: An application could transform its data based on the current network conditions. Examples of this are real time video applications that reduce their frame rates when bandwidth drops and Web servers that reduce the quality of their content (e.g. JPEG's) when bandwidth drops.
- A user can determine if a new network equipment/service is really delivering the claimed bandwidth.
- A client can select a higher bandwidth proxy/replicated server over a lower bandwidth proxy/replica.
- A server can adapt to different client bandwidths by scaling the size and quality of its content.
- A multi-homed host can route traffic through the highest bandwidth available interface.

6. Conclusion

In this paper we designed methodical Internet data collection experiments to meausre the performance metrics of the Internet using some existing Internet performance measurement tools namely, traceroute and VisualRoute. These experiments are intended to evaluate different Web servers. We analyzed the results of these experiments to predict possible problems and behavior. We characterized and quantized collected data including Internet latencies, packet loss rates, and available bandwidths. The characterized and quantized collected data was used to model traffic behavior. We also analyzed the routing behavior for pathological conditions and routing stability. The results can be used by Internet customers or Internet Service Providers (ISPs) to track Internet servers behavior and response time at application level and to monitor Internet path connectivity. In addition, we presented a new methodology to model and analyze end-to-end Internet behavior. This methodology is based on finding the transfer function of the end-to-end Interent path between the client and Web server.

Results of conducted experiments to obtain performance metrics (packet loss rate and RTT) for nine Web sites were reported. We chose these Web sites from different places to note the effect of distances on RTT and packet loss. The resultant data and the paths which the packets take for each site during six different days at four different daily times for each day were reported. A client PC located in Gaza in Palestine which is connected to the Internet through a 56 kbps modem was used to access these Web sites. Results of the multiple experements were also used to study the interval-temporal-connectivity of paths. The results agree with other work done in the related field and show that our methodology is effective in measuring the Internet performance. In addition, the analysis of the collected data and the corresponding routes show that the values of the collected data are affected by different factors including the model of use and Web server environemt traffic. As results show, requested traffic highly impacts Web server performance including QoS, scalability, capacity planning, availability and reliability. Knowing the values of the Internet performance metrics allows systems and people to be much smarter about the decisions that they make. The client PC speed and the speed of the used modem seem to influence the measurement quality for both traceroute and VisualRoute– the slower PC tends to show lower numbers for the conducted tests. The accuracy of the tool for a certain Web server depends on the distance to the Web server. The more hops from the client to the tested Web site, the worse measurements are achieved.

We will extend our work in the future to include the design of our own integrated measurement tool that automatically conducts, analyzes and characterizes performance metrics of the Internet and displays the results to the user for assessment and evaluation. This extension is based on our new methodology.

References

1- Allman M., Nov. 2001 – *Measuring End-to-End Bulk Transfer Capacity.* Proceedings of the first ACM SIGCOMM Workshop on internet Measurement

2- Andren J., Hilding M., and Veitch D., 1998 - *Understanding end to end Internet Traffic Dynamics.* IEEE Global Telecommunications Conference, P: 1118-1122.

3- Brownlee N., Spr. 2001 – *Fundamentals of Internet Measurement: A Tutorial.* CMG Journal of Computer Resource Managemet, Issue 102

4- Chen K., 2002 - *A Compendium of Network Performance Measurement Resources.* http://dast.nlanr.net/Articles/measurements/

5- Claffy K., October 1995 - Braun H., and Olyzos G., *A Parameterizable Methodology for Interent Traffic Flow Profiling.* IEEE Journal on Selected Areas in Communications, P: 1481-1494

6- Claffy K., and Monk T., October 1997 - *What's Next for Internet Data Analysis? Status and Challenges Facing the Community.* Proceedings of the IEEE, Vol. 85, No. 10, P: 1563-1571

7- Cowie J., Nicol D., And Ogielski Y., January-February 1999 - *Modeling the Global Internet*. Computing in Sciece and Engineering Journal, Vol. 1 Issue 1 , P: 42-50

8- Decina M., 2-4 December 1997 - *The Interent Rtevolution: Reshaping Business for the 21st Century*. 2nd IEEE Workshop on Broadband Switching Systems Proceedings, P: 97

9- Dong Y., Hou Y., Zhang Z., and Taniguchi T., 1999 - *A Server-Based Non-Instrusive Measurement Infrastructure for Enterprise Networks*. Performance Evaluation: An International Jornal, Vol. 36-37, P: 233 – 247

10- Horneffer M., 30 Sep. 1998 – *Methods for Performance Analysis of Internet Access points*. Computer Networks and ISDN Systems, Vol. 30, No. 16 – 18, P: 1607 – 1615

11- Johnson K., Carr J., Day M., and Kaashoek M., May 2000 – *The measured Performance of Content Distribution Networks*. Fifth International Web Cashing Content Delivery Workshop

12- Kushida T., November 1998 - *The Measurement and the Impirical Studies for the Internet*. The Bridge to Global Integration Global Telecommunications IEEE Conference, P: 1142-1147

13- Lai K., and Baker M., 2001 - *Nettimer: A Tool for Measuring Bottleneck Link Bandwidth*. http://mosquitonet.stanford.edu/~laik/projects/nettimer/publications/usits2001/main.html, last update 4.4.2001

14- Lindh T., Mar. 25-26 2002 – *A New Approach to Performance Monitoring in IP Networks*. Passive and Active Measurement Workshop, Fort Collins, Colorado, USA

15- Monk T., and Claffy K., 1997 - *Internet Data Acquisition & Analysis: Status & Next Steps*. http://www.caida.org/outreach/papers/1997/data-inet97/data-inet97.html

16- Paxson V., Oct. 1997 - *End-to-End Routing Behavior in the Internet*, IEEE/ACM `Transactions on Networking, Vol. 5, No. 5

17- Squillante M., Yao D., and Zhang L., December 1999 - *Web Traffic Modeling and Web Server Performance Analysis*. Proceedings of the 38th onference on Decision and Control, Pheonix, Arizona, USA, P: 4432-4438

18- tracenoute, www.traceroute.org

19- visualroute, www.visualroute.com