SCIENTIFIC MEASURE OF AFRICA'S CONNECTIVITY

M. Zennaro, E. Canessa*, K.R. Sreenivasan*, A.A. Rehmatullah#, R.L. Cottrell#*

** The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.*

 # Stanford Linear Accelerator Center, Stanford, USA

Data on Internet performance and the analysis of its trend are useful for decision makers and scientists alike. Such performance measurements are possible using the PingER methodology. We use the data thus obtained to quantify the difference in performance between developed and developing countries, sometimes referred to as the Digital Divide. Motivated by the recent interest of G8 countries in African development, we particularly focus on the African countries.

Introduction

Science and Technology (S&T) are critical for social and economic development within a nation and between developed and developing countries [1]. Fortunately, there are signs that S&T are increasingly becoming a part of the agenda of the international community and policy makers within developing countries [2].

Achieving scientific development will depend in part on increased cooperation between scientists from the North and the South, including setting up networks of researchers and institutions. Modern collaboration requires the use of the Internet. Network connection can be used for large-scale scientific data transfer, real-time collaboration, or access to scientific literature. Unfortunately, the network connection in developing countries is marginal at best [3].

There is much discussion on the digital divide but very little concrete data quantifying it. To set expectations for the quality of connectivity, it is critical to monitor the performance of the Internet, for instance in remote areas of Africa. Monitoring is important to understand where upgrades are needed and to provide trouble-shooting information. Data on the trends of Internet connectivity are useful even for decision makers within the G8 when they discuss further issues related to resource allocation on debt relief for Africa.

Performance measurements are possible through the PingER project [4], developed by the IEPM group at the Stanford Linear Accelerator Center. The project started monitoring the network's performance for High Energy Physics experiments in the mid-90's, and the technique involves sending out a pulse and timing the return delay. The monitoring has low network impact and can be used for hosts with especially poor connectivity. The resulting monitoring information is in the public domain.

Today more than 700 remote sites are being monitored worldwide, more than 40 of them in Africa. The Science Dissemination Unit of the International Centre for Theoretical Physics [5] collaborates with the PingER project, gathering and analyzing the data about Africa [6], by choosing African universities suitable for on-line collaboration programs. For real-time scientific cooperation, it is possible by this means to determine how well the interactive applications between various pairs of sites might work.

The PingER project

PingER (Ping End-to-end Reporting) is the name given to the Internet End-to-end Performance Measurement (IEPM) project to monitor the performance of Internet links. The main mechanism used is the Internet Control Message Protocol (ICMP) Echo mechanism, also known as the Ping facility [7]. ICMP packets are special IP control messages that are used to send network information between two hosts. The Ping facility allows sending an echo request packet of a user-selected length to a remote node and having it echoed back. The response provides useful information such as the time it took to get a reply, the number of packets that were lost, whether the packets were received in order, and so forth. Ping usually comes pre-installed on almost all platforms. The server (i.e. the echo responder) runs at a high priority (e.g. in the kernel on Unix) and so is more likely to provide a good measure of network performance than a user application.

PingER has a very low network impact of less than 100 bits/second for each monitoring-remote host, and can be set to less than 10 bits/second for hosts with especially poor connectivity.

The tools that implement the Ping monitoring are collectively known as PingER. There are about 30 active Monitoring Sites in 14 countries, over 700 remote sites in more than 110 countries being monitored and over 3500 monitor-site remote-site pairs included. These countries contain over 78% of the world's population and over 99% of the online users of the Internet. Most of the hosts monitored are at educational or research sites. PingER has historical data going back to January 1995, so a wealth of trend information is available [8].

The PingER methodology can be used to measure the Packet Loss percentages, the Round Trip Time (RTT), the variability of the response time, both short term and longer, and the lack of reachability (no response for a succession of Pings).

The Packet Loss is a good measure of the quality of the link for many TCP based applications. Loss is typically caused by congestion which in turn causes queues (e.g. in routers) to fill and packets to be dropped. Losses may also be caused by the network delivering an imperfect copy of the packet. This is usually caused by bit errors in the links or in network devices. When we get a zero packet loss sample (a sample refers to a set of n pings[1]), we refer to the network as being quiescent or non-busy. We can then measure the percentage frequency of how often the network was found to be quiescent. A high percentage is an indication of a good network. This is because users now work mostly with interactive applications, such as video conferencing and audio chat, which require a low packet loss percentage. The levels

[1] The default value for n is 10.

that describe the link quality are the following: 0-1% of packet loss is good, 1%-2.5% is acceptable, 2.5%-5% is poor, 5%-12% is very poor, and greater than 12% is regarded as bad. Our observation is that above 4-6% packet loss video conferencing becomes irritating and speakers of the non-native languages are unable to communicate. Above 10-12% packet loss there is an unacceptable level of back-to-back loss of packets and long timeouts, connections start to get broken, and video conferencing is unusable.

The Round Trip Time is another indicator of the quality of a link. However, unlike packet loss, where it is possible to reduce losses to zero, it is impossible to reduce the RTT to less than the time taken for the electrons to travel the distance along the fiber or copper cable. In addition to the cable delays, the measured RTT is the time taken for the packet to be accepted by the router interface, the delay caused by the queuing, and the time taken for the packet to be passed to the interface. The minimum RTT thus indicates the length of the route taken by the packets, the number of hops and the line speeds. Changes in minimum RTT can be an indication of a route change. The major effect of poor response time is felt on interactive applications as telnet or 'packetized' video or voice, where even a moderate delay can cause severe disruption. Applications that do not require such a level of interactivity, such as e-mail and web browsing, may appear to perform well even with high delay.

Uses of the PingER data

Over the past six years, the information gathered by the PingER project has been used in several ways. For example, it has been used to track network infrastructure changes. PingER data have been used also to illustrate the need for upgrades to a network. Based on the presentation of the PingER findings, a successful recommendation can be made to people in charge of policy and funding in order to increase the bandwidth. PingER has also been used to illustrate Digital Divide. Not all sites that are located in a developing region see the negative effects of the Digital Divide. If one site can attain credible connectivity, then other sites in that region should be able to have better connectivity as well. Other practical uses of PingER include selecting Internet Service Providers and monitoring the accessibility to network changes and upgrades, and their effects on connectivity.

As a troubleshooting tool, PingER has been used to discern if a reported problem is related to networks, identify the time at which the problem has started, decide whether it is still occurring, and provide quantitative analysis for Internet Service Providers (ISPs).

The PingER data can also be used to select universities from developing countries for remote collaboration programs. By using PingER to measure the Packet Loss and RTT, it is possible to provide expectations [9] on the performance of bulk data transfers and other applications. In case of real-time collaboration, by comparing the results from PingER with various recommendations for loss, RTT and jitter, together with the perceptions of voice quality from the users, one can determine how well VoIP and other interactive applications might work between various pairs of sites [10].

Results of PingER monitoring in Africa

The new millennium is beginning to see significant advances in Africa's quest for greater connectivity. Nevertheless, while a substantial increase in the rate of expansion and modernization of networks is taking place, the ITU statistics [11] on teledensity show that while there are 57.3 Internet users per 100 inhabitants in Sweden, 57 in the USA and 34.7 in Italy, there are just 0.5 in Mali and 0.2 in Niger. The Internet tariff for the same type of connection is 1.1% of the GNI in Sweden and in the USA, 1% in Italy, while it is 289% in Mali and 683% in Niger. The same differences are reflected by the Internet performance. The Internet derived TCP throughput[2] seen from SLAC and CERN to each region is shown in Fig. 1a and Fig 1b.

It is clear that the Internet in Africa is slower by a factor of the order 50, on the average, when compared to European and North American standards. What can also be seen from these graphs are that:

- in the long term, performance to all regions is improving;

---

[2] The derived throughput is calculated using the formula given by Mathis et. al. ("The macroscopic behavior of the TCP congestion avoidance algorithm", *Computer Communication Review, 27(3), July 1997)*. relating throughput to Packet Loss and RTT.

- for the developed regions performance is improving by roughly a factor of 10 in 5 years;
- performance in developing regions are a factor of 5 to 20 times worse than that in developed regions;
- performance to developing regions is typically on a par with what was seen 2-7 years ago in developed regions.
- Africa is not catching up with the developed regions, unlike Russia and Latin America.

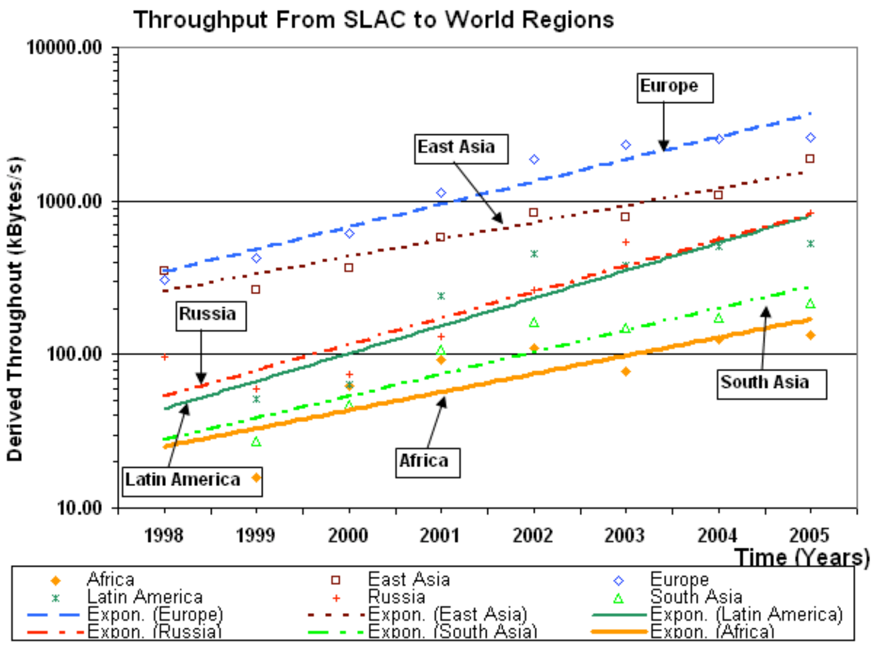

Fig 1a: Throughput Performance from SLAC to regions of the world. The lines are fits to exponential functions to help guide the eye.

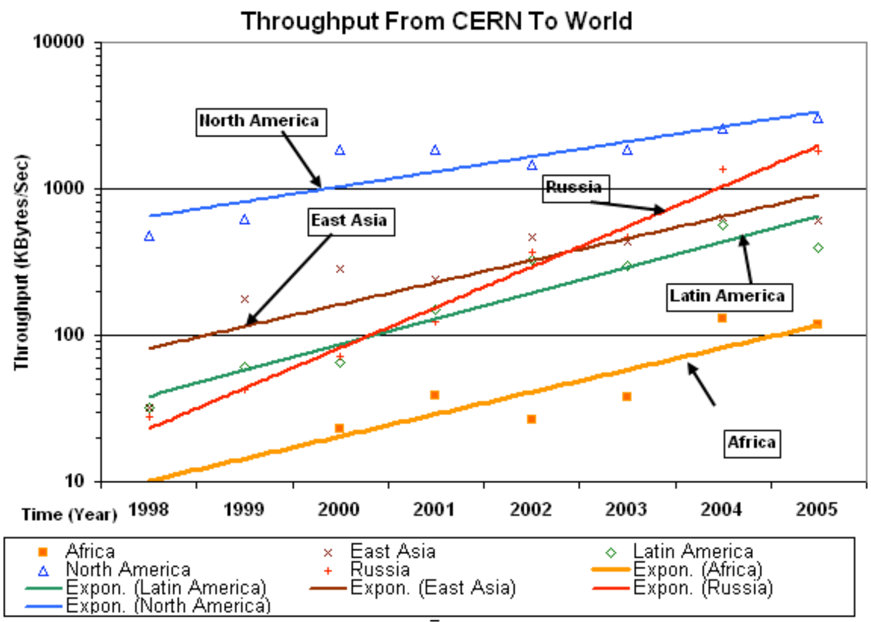

Fig 1b: Throughput performance from CERN to regions of the world.

As can be seen in Fig. 2, the PingER project monitors more than 40 African institutions in countries that between them contain over 80% of the population of Africa. To obtain hosts to monitor at remote sites in Africa, contacts were found by sending emails to colleagues especially in the International Committee for Future Accelerators and the ICTP's eJDS (Electronic Journals Delivery Service). Sometimes these emails resulted in further referrals or required extended explanations. It was then checked if the host was accessible to Pings, was truly located in Africa[3], and then entered in the PingER database. Typically about 75% of the contacts eventually resulted in a remote host to monitor successfully. Sometimes it took many months to conclude agreements.

All the sites are monitored from SLAC in the USA. Though the minimum RTT depends on the distance between the monitoring host and the remote host, for a given remote host there is little difference in the losses measured from different monitoring hosts. This indicates that the common bottleneck in most cases is closely associated with the remote site.

Fig 2: Monitored African sites. Most of the hosts monitored are at educational or research sites.

The increase in the number of African sites monitored by PingER in the last six years can be seen in Fig. 3 together with the losses seen by the site. This reflects the dramatic increase in interest in Africa following various meetings focusing on Africa, especially those organized by the ICTP [12], the ICFA/SCIC [13] and the WSIS.

[3] Several of the initially chosen hosts were web servers that had proxy servers located outside Africa.

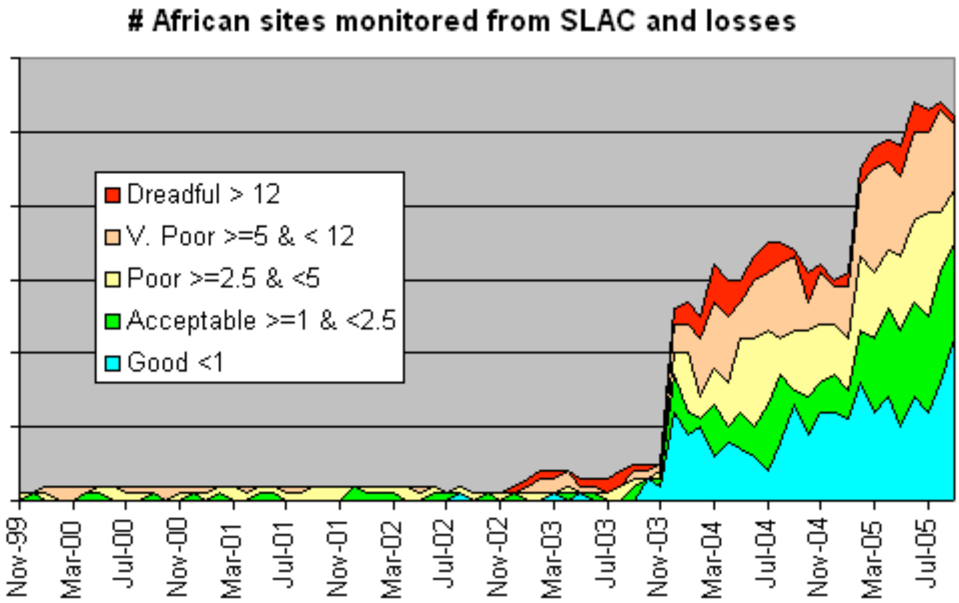

Fig. 3: Number of sites in Africa monitored by PingER from SLAC, also showing the packet loss percentages.

When Africa's performance is analyzed by region as in Fig 4, North Africa performs well, while East Africa is falling behind. In North Africa, Algeria and Morocco perform very well, benefiting the whole region. From the existing results, it is apparent that most African regions have poor to bad connectivity. In fact the sites appear to have smaller throughput than many homes with DSL or dial-up modems in developed countries. Even within the same region, there are differences of more than an order of magnitude in performance between different countries.

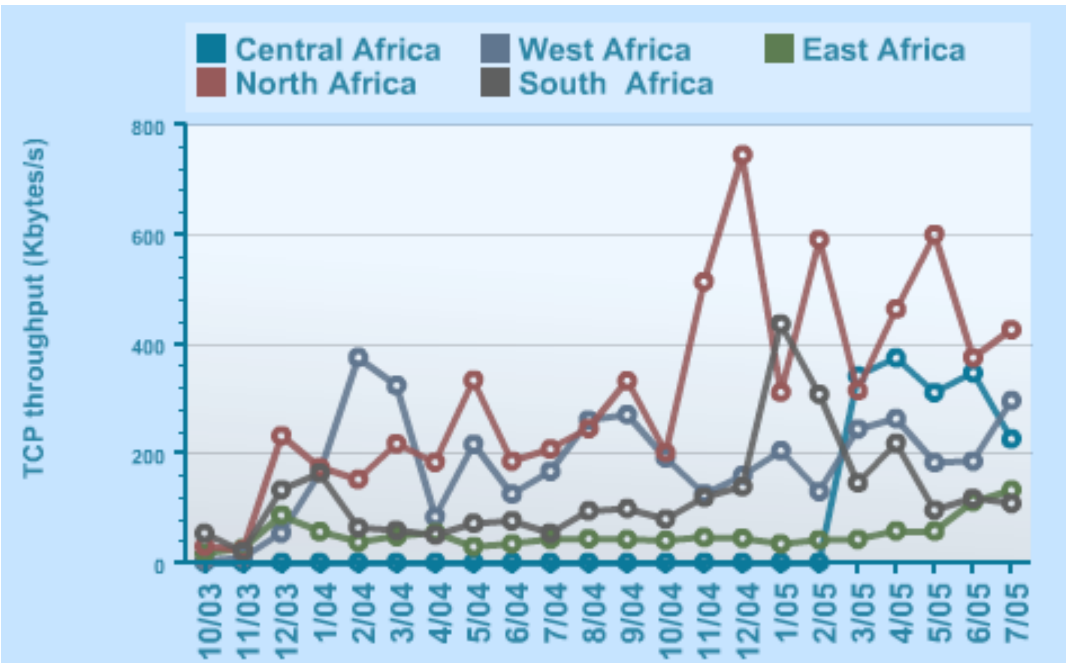

Fig 4: Throughput Performance from SLAC to different regions of Africa.

Recently, the PingER project has established a monitoring station at the TENET site in Ronderbosch[4], South Africa, so now there are measurements from within and between African countries. The data indicate that for the sites in the 28 African countries monitored by PingER from this South Africa monitoring station:

- The only countries with sites that have direct connectivity (i.e. without going through Europe or N. America) to S. Africa are Botswana, Zimbabwe and Tunisia. Even then one of the three sites monitored in Tunisia connects through N. America.
- The sites in fifteen countries (Algeria, Benin, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mali, Mauritania, Morocco, Mozambique, Niger, Nigeria, Senegal and Sudan) connect through Europe.
- The sites in five countries (Angola, Cameroon, Egypt, Namibia and Uganda) connect though N. America.
- The sites in two countries (Ghana and Tanzania) connect through Europe or N. America.
- The sites in two countries (Burkina Faso and Rwanda) connect through Europe and N. America.

Not only does this situation drastically increase the RTT, it also utilizes expensive (for African countries) inter-continental links since the Africa Internet Service Provider typically has to pay 100% of the international carrier cost to get from Africa to the International Backbone Provider [14].

It is also instructive to assess the fraction of African countries that have sites connected to S. Africa by geosynchronous satellites. This is illustrated in Fig. 5. As might be expected the countries with terrestrial connections are those served by the SEAMEWE II/III and SAT 3 cables or sharing a border with S. Africa. The impact of a geosynchronous satellite is to add roughly 600ms to the minimum RTT.

[4] Many thanks to Duncan Martin and Len Lotz of TENET.

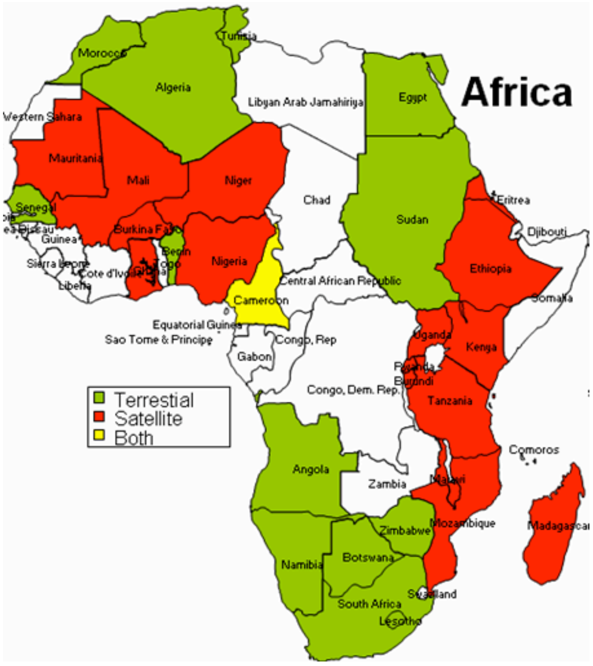

Fig 5: African countries with sites connected to S. Africa by geosynchronous satellites and by terrestrial links.

Attention should be paid to the usability of the Internet given these performances. If we assume that the same network supports a number of computers that can reach the hundreds in many university labs and that bandwidth requirements are growing every day, it is clear that most internet applications are impossible to use. ICTP has developed a "Digital Divide Simulator" [15] that enables scientists with high bandwidth connections to experience what their colleagues from the South experience daily when using the Internet. Other initiatives, such as the eJDS, are being supported by the Centre to deliver scientific literature to scientists with low bandwidth connections.

Conclusions

Continual efforts are being made to maintain, upgrade and expand the PingER deployment. With respect to the remotes sites, the primary effort is to keep them accessible to the monitoring sites. For monitoring sites, it is important to keep the list of remote sites up to date, the monitoring code running reliably and the data accessible for collection. The data and the reports are made available to all interested users via the data selection, analysis and display tools.

Much work needs to be done to extend the monitoring to more of Africa. Even for the countries already monitored, more sites are needed to help avoid anomalous results associated with a single site. To understand where South-to-South scientific collaboration is feasible using the internet, it is necessary to monitor more African sites from South Africa and to set up monitoring hosts in other African countries besides S. Africa. The authors encourage readers to provide contacts for new countries and sites [16].

PingER provides a valuable light-weight tool for the active end-to-end performance monitoring of networks around the world. With its continual gathering of internet monitoring data, it provides extensive quantitative historical and almost real time information on world-wide networks. PingER has shown itself to be useful for providing valuable information quantifying network needs and improvements. Its results can be used by the G8 countries involved in African development, to select the countries which need support to develop their network and to monitor the effectiveness of the improvements.

1. Special Report: Science & Africa, Nature 436, 156-157 (2005).

2. Weerawarana S. & Weeratunge J., Open Source in Developing Countries (SIDA, 2004).

3. Cerdeira H., Learned Publishing 15, 1-4 (2002).

4. Matthews W. & Cottrell, R.L., IEEE Commun. Mag. 38, 130-136 (May 2000).

5. Science Dissemination Unit of ICTP: http://sdu.ictp.it

6. PingER data for Africa: http://sdu.ictp.it/pinger/africa/

7. Tanenbaum A., Computer Networks, 449-450

8. Cerdeira H. Canessa, E. Fonda, C. & Les Cottrell, R., CERN Courier 43, 17-18 (2003).

9. Mathis, Semke, Mahdavi & Ott, The macroscopic behavior of the TCP congestion avoidance algorithm, Computer Communication Review, 27(3), July 1997

10. ITU-T SG13 Draft Rec. I.380 - "Internet Protocol Data Communication Service - IP Packet Transfer and Availability Performance Parameters" - Geneva, February 1999

11. ITU Digital Access Index: http://www.itu.int/ITU-D/ict/dai/index.html

12. Developing Country Access to On-Line Scientific Publishing: Sustainable Alternatives, Round Table meeting held at ICTP Trieste, Oct 2002. http://www.ictp.it/~ejds/seminars2002/program.html

13. "ICFA SCIC Network Monitoring Report", Edited by Les Cottrell, http://www.slac.stanford.edu/xorg/icfa/icfa-net-paper-jan05/

14.AFRISPA Initiatives http://www.afrispa.org/initiatives.htm

15. Digital Divide Simulator: http://wireless.ictp.it/simulator

16. Correspondence and requests for materials should be addressed to sdu@ictp.it