**International Committee for Future Accelerators (ICFA)**

**Standing Committee on Inter-Regional Connectivity (SCIC)**

**Chairperson: Professor Harvey Newman, Caltech**

**ICFA SCIC Network Monitoring Report**

Prepared by the ICFA SCIC Monitoring Working Group

On behalf of the Working Group:   
Les Cottrell [cottrell@slac.stanford.edu](mailto:cottrell@slac.stanford.edu)



2013 - 2014 Report of the ICFA-SCIC Monitoring Working Group

Edited by [R. Les Cottrell](mailto:cottrell@slac.stanford.edu), Shawn McKee and Raja Asad Khan On behalf of the [ICFA-SCIC Monitoring WG](mailto:icfa-ntf-wg-monitoring@listbox.cern.ch)

Created December 14, 2013, Finished January 29, 2014

[ICFA-SCIC Home Page](http://icfa-scic.web.cern.ch/ICFA-SCIC/) | [Monitoring WG Home Page](http://icfa-scic.web.cern.ch/ICFA-SCIC/Committees/monitoring.htm)

This report is available from:

<http://www.slac.stanford.edu/xorg/icfa/icfa-net-paper-jan14/report-jan14.doc>

Table of Contents

[Figures 4](#_Toc378091066)

[Executive Overview 5](#_Toc378091067)

[Introduction 7](#_Toc378091068)

[Methodology 7](#_Toc378091069)

[PingER Results 8](#_Toc378091070)

[Deployment 8](#_Toc378091071)

[Historical Growth of PingER Coverage Since 1998 10](#_Toc378091072)

[Metrics 11](#_Toc378091073)

[Yearly loss trends: 12](#_Toc378091074)

[Yearly minimum RTT 13](#_Toc378091075)

[Yearly Jitter trends 15](#_Toc378091076)

[Yearly Throughput Trends 16](#_Toc378091077)

[View from Europe 19](#_Toc378091078)

[Variability of performance between and within regions 20](#_Toc378091079)

[Yearly Mean Opinion Score (MOS) trends 22](#_Toc378091080)

[Directivity 23](#_Toc378091081)

[Comparisons with Economic and Development Indicators 25](#_Toc378091082)

[PingER Case Studies 27](#_Toc378091083)

[High Performance Network Monitoring 28](#_Toc378091084)

[*New and Ongoing Monitoring and Diagnostic Efforts in HEP* 28](#_Toc378091085)

[**perfSONAR for HEP** 29](#_Toc378091086)

[*LHCONE Monitoring History* 33](#_Toc378091087)

[**Additional Complementary End-to-End Monitoring** 34](#_Toc378091088)

[**Summary: Progress in HEP Network Monitoring for 2013** 36](#_Toc378091089)

[*Related HEP Network Research* 36](#_Toc378091090)

[Comparison with HEP Needs 37](#_Toc378091091)

[Recommendations 39](#_Toc378091092)

[Future Support 39](#_Toc378091093)

[Acknowledgements 40](#_Toc378091094)

[Appendices 41](#_Toc378091095)

[Appendix A: PingER Presentations etc. in 2013. 41](#_Toc378091096)

[Appendix B: New tools added in 2013 41](#_Toc378091097)

[TULIP 41](#_Toc378091098)

[PingER Linked Open Data 42](#_Toc378091099)

[Appendix C: PingER Project History in 2013 46](#_Toc378091100)

[Appendix D: ICFA/SCIC Network Monitoring Working Group 47](#_Toc378091101)

[Goals of the Working Group 47](#_Toc378091102)

# Figures

*Please note that for all yearly graphs the data point is for the end of the year.*

[Figure 1 : Locations of PingER monitoring and remote sites as of December 2013. Red sites are monitoring sites, blue sites are beacons that are monitored by most monitoring sites, and green sites are remote sites that are monitored by one or more monitoring sites 9](#_Toc378773425)

[Figure 2 Major regions of the world for PingER aggregation by regions, countries in white are not monitored 9](#_Toc378773426)

[Figure 3 : Number of sites monitored from SLAC by region at the end of each year 1998 – 2013 11](#_Toc378773427)

[Figure 4: Packet Loss measured for various regions from SLAC till Dec 2013 12](#_Toc378773428)

[Figure 5: Minimum RTTs measured from SLAC to world countries, Dec 2013. Different regions have different colors. 13](#_Toc378773429)

[Figure 6: Maps of minimum RTTs from SLAC to the world in 2008 and 2013. 14](#_Toc378773430)

[Figure 7: Time series of minimum RTT measured from SLAC to various countries since 1998 15](#_Toc378773431)

[Figure 8: Jitter from SLAC to regions of the world until Dec 2013. 16](#_Toc378773432)

[Figure 9: Yearly averaged normalized derived TCP throughputs from the SLAC to various regions of the world until Dec 2013. 17](#_Toc378773433)

[Figure 10: Derived Throughput kbits/sec from SLAC to the World (since the throughputs in this graph are not normalized we have not shown N. America) until Dec 2013. The yellow line is to help show the rate of change. If one extrapolates Europe’s performance backwards to February 1992, it intercepts Africa’s performance today. 18](#_Toc378773434)

[Figure 11: Extrapolations on the throughput data. 19](#_Toc378773435)

[Figure 12: Normalized throughputs to various regions as seen from CERN until Dec 2013. 20](#_Toc378773436)

[Figure 13: maximum, 95, 90, 75 percentile, median, 25 percentile and minimum derived throughputs of various regions measured from SLAC for Nov 2013 and ordered by median throughput. 22](#_Toc378773437)

[Figure 14: Derived throughput from SLAC to S.E. Asia countries until Dec 2013. 22](#_Toc378773438)

[Figure 15: Mean Opinion Scores derived from PingER measurements from SLAC to regions of the world until Dec 2013. 23](#_Toc378773439)

[Figure 16: Directivity measured from SLAC to regions of the world until Dec 2013. 24](#_Toc378773440)

[Figure 17: Pinger Derived Throughput vs. the ICT Development index. The bubble size is proportional to the country's population. 26](#_Toc378773441)

[Figure 26: Example OWAMP test between two USATLAS Tier-2's (bi-directional testing is enabled) showing the minimum one way delays and losses for each direction. 30](#_Toc378773442)

[Figure 27: Example perfSONAR PingER round trip time graph between the Tier-1 and a Tier-2. The line at the bottom is the mean RTT. 30](#_Toc378773443)

[Figure 28: Example bi-directional throughput test in perfSONAR between two Tier-2s 31](#_Toc378773444)

[Figure 29 Example MaDDash display for the UK (United Kingdom) perfSONAR-PS cloud. Shown are the two matrices for Bandwidth measurements (left) and Latency measurements (right) for the sites participating in the UK cloud. Orange indicates a missing measurement while red, yellow and green are critical, warning and OK states for the corresponding network metrics. 33](#_Toc378773445)

[Figure 31: The “selected site” LHCONE monitoring page showing the early setup for monitoring. The URL is https://perfsonar.usatlas.bnl.gov:8443/exda/?page=25&cloudName=LHCONE 34](#_Toc378773446)

[Figure 32 Example production system throughput test results between two USATLAS sites: BNL (Tier-1) and AGLT2 (Tier-2). Shown are the storage areas: AGLT2\_DATADISK and BNL\_OSG2\_MCDISK from November 21, 2012 to January 11, 2013 35](#_Toc378773447)

# Executive Overview

Internet performance is improving each year with throughputs typically improving by 20% per year and losses by up to 25% per year. Most countries have converted from using Geostationary Satellite (GEOS) connections to terrestrial links. This has improved performance in particular for Round Trip Time (RTT) and throughput. GEOS links are still important to countries with poor telecommunications infrastructure, landlocked developing countries, remote islands, and for outlying areas. In some cases they are also used as backup links.

In general, throughput measured from within a region is much higher than when measured from outside. Links between the more developed regions including N. America[[1]](#footnote-1), E. Asia (in particular Japan, South Korea and Taiwan) and Europe are much better than elsewhere (3 - 10 times more throughput achievable). Regions such as S.E. Asia, S.E. Europe and Latin America are 5-9 years behind. However, in 2009, Africa was ~15 years behind Europe, also Africa’s throughput was 12-14 times worse than Europe and extrapolating the data indicated that it would further degrade to almost 60 times worse by 2026. Since 2009, due in large part to the installation of multiple submarine fibre optic cables to sub-Saharan Africa, there has been a significant improvement in Africa’s performance. It now appears to be catching up, such that if the present improvements are maintained, it could catch Europe by around 2030.

Africa and South Asia are two regions where the Internet has seen phenomenal growth, especially in terms of usage. However, it appears that network capacity is not keeping up with demand in these regions. In fact many sites in Africa and India appear to have throughputs less than that of a well-connected (cable, DSL, etc.) home in Europe, North America, Japan or Australia. Further the end-to-end networking is often very fragile both due to last mile effects and poor infrastructure (e.g. power) at the end sites, and also due to lack of adequate network backup routes. Africa is a big target of opportunity with over a billion people of which in 2012 only 15.6% are Internet users. It also had a 3,607% (compared to 566% for the world) growth in number of Internet users from 2000-2012[[2]](#footnote-2). However, there are many challenges including lack of power, import duties, lack of skills, disease, corruption, and protectionist policies. In almost all measurements, Africa stands out as having the poorest performance. Further Africa is a vast region and there are great differences in performance between different countries and regions within Africa.

There is a moderate to strong positive correlation between the Internet performance metrics and economic and development indices available from the UN and International Telecommunications Union (ITU). Given the difficulty of developing the human and technical indicators (at best they are updated once a year and usually much less frequently); having non-subjective indicators such as PingER that are constantly and automatically updated is a very valuable complement. Besides being useful in their own right these correlations are an excellent way to illustrate anomalies and for pointing out measurement/analysis problems. The large variations between sites within a given country illustrate the need for careful checking of the results and the need for multiple sites/country to identify anomalies.

For modern HEP collaborations and Grids there is an increasing need for high-performance monitoring to set expectations, provide planning and trouble-shooting information, and to provide steering for applications. As link performance continues to improve, the losses between developed regions are decreasing to levels that are not measureable by PingER. Though the measurements for RTT, jitter, and unreachability[[3]](#footnote-3) are still correct, as the measured losses go to zero this also makes the throughput derivation unreliable. Alternative solutions to measuring the throughput are available, however they can be harder to install and absorb more network bandwidth. Examples of other measurement projects using the more intense methods are the MonALISA[[4]](#footnote-4) project that uses the pathload[[5]](#footnote-5) packet pair technique as well as file transfers, and perfSONAR[[6]](#footnote-6) that uses the iperf[[7]](#footnote-7) TCP transport mechanism. There is also a project in place at SLAC and LBNL under the perfSONAR umbrella to analyze and present data from production gridFTP[[8]](#footnote-8) transfers that are heavily used in the HEP community. These projects are becoming increasingly important for links between well-developed sites.

In the last year there have been the following changes:

* Invited article in February 2013 IEEE Spectrum on Pinging Africa. It was picked up by Al Jazeera and SlashDot among others.
* The network monitoring collaboration between SLAC and the University of Malaysia in Sarawak (UNIMAS) was extended to the University of Malaya in Kuala Lumpur and the Universiti Technologi in Johor Baru.
* Deployment of two new PingER Monitoring nodes in Malaysia, one in Dakar, Bangladesh and several in Pakistan.
* Case studies of the impact on the Sudan disconnecting from the Internet in September 2013, Syria going offline in May 2013, and Routing of Internet traffic within S. E. Asia.
* Updating of the major figures and tables.
* Improved analysis and reporting tools.
* The number of hosts monitored, which started to decline in 2010, increased to an all time high in 2013. We now monitor 170 countries having added Myanmar, Jamaica, Kuwait, Trinidad and Tobago. We monitor all countries with populations > 1,000,000 apart from Puerto Rico.
* A new project whose proposal is to publish PingER data in Linked Open Data[[9]](#footnote-9) Semantic Web[[10]](#footnote-10) standards has begun. Part of PingER data has been published in Resource Description Framework[[11]](#footnote-11) (RDF) format using an OWL ontology[[12]](#footnote-12) for network measurements and linked to other existing databases on the web. The data is also retrievable through a SPARQL[[13]](#footnote-13) Endpoint, using structured SPARQL queries. Use cases were added as visualization examples to show how the data in this standard format can be useful.
* We completed a working version of Geolocation of Internet hosts using trilateration based on ping round trip times.

To quantify and help bridge the Digital Divide, enable world-wide collaborations, and reach-out to scientists world-wide, it is imperative to continue the PingER monitoring coverage to all countries with HEP programs and significant scientific enterprises. However, the funding for PingER is currently a major challenge.  We met with Matt Mathis of Google to explore continued support for PingER. Google’s interest is in the long term history of the Internet’s performance.

# Introduction

This report may be regarded as a follow up to the previous ICFA Standing Committee on Inter-regional Connectivity (SCIC) Monitoring working group’s Network reports[[14]](#footnote-14) dating back to 1997.

The current report updates the January 2012 report, but is complete in its own right in that it includes the tutorial information and other relevant sections from the previous report.  The main changes in this year’s reports are given in the previous Executive section

# Methodology

There are two complementary types of Internet monitoring reported on in this report.

1. In the first we use [PingER[[15]](#footnote-15)](http://www-iepm.slac.stanford.edu/pinger/) which uses the ubiquitous "ping" utility available standard on most modern hosts. Details of the PingER methodology can be found in the [Tutorial on Internet Monitoring & PingER at SLAC[[16]](#footnote-16)](http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html#loss). PingER provides low intrusiveness (~ 100bits/s per host pair monitored[[17]](#footnote-17)) RTT, loss, jitter, and reachability (if a host does not respond to a set of 10 pings it is presumed to be non-reachable). The low intrusiveness enables the method to be very effective for measuring regions and hosts with poor connectivity. Since the ping server is pre-installed on all remote hosts of interest, minimal support is needed for the remote host (no software to install, no account needed etc.)
2. The second method (perfSONAR[[18]](#footnote-18) etc.) is for measuring high network and application throughput between hosts with excellent connections. Examples of such hosts are to be found at HEP accelerator sites and the Large Hadron Collider (LHC) tier 1 and 2 sites, major Grid sites, and major academic and research sites in N. America, Japan and Europe. The method can be quite intrusive (for each remote host being monitored from a monitoring host, it can utilize hundreds of Mbits/s or more for ten seconds to a minute, each hour). To minimize intrusion, the US-ATLAS scheduling utilizes 20 second tests every 4 hours rather than every hour. It also requires more support from the remote host. In particular either various services must be installed and run by the local administrator or an account is required, software (servers) must be installed, disk space, compute cycles etc. are consumed, and there are security issues. The method provides expectations of throughput achievable at the network and application levels, as well as information on how to achieve it, and trouble-shooting information.

# PingER Results

## Deployment

The PingER data and results extend back to the start of 1995 with online data publicly available from the start of 1998. They thus provide a valuable history of Internet performance. PingER now has ~80 active monitoring nodes in ~22 countries that monitor over 810 remote nodes at over 775 sites in 166 countries for a total of over 8000 monitor node – remote node pairs. These countries contain over 99% of the world's population (see Table 2: Quality ranges used for loss) and over 99.5% of the online users of the Internet. Most of the hosts monitored are at educational or research sites. We try and get at least 2 hosts per country to help identify and avoid anomalies at a single host. The [requirements for the remote host](http://www.slac.stanford.edu/comp/net/wan-req.html) are documented[[19]](#footnote-19). below shows the locations of the monitoring and remote (monitored sites).



Figure 1 : Locations of PingER monitoring and remote sites as of December 2013. Red sites are monitoring sites, blue sites are beacons that are monitored by most monitoring sites, and green sites are remote sites that are monitored by one or more monitoring sites

There are several thousand monitoring/monitored-remote-host pairs, so it is important to provide aggregation of data by hosts from a variety of "affinity groups". PingER provides aggregation by affinity groups such as HEP experiment collaborator sites, region, country, Top Level Domain (TLD), or by world region etc. The world regions, as defined for PingER, and countries monitored are shown below in . The regions are chosen starting from the [U.N. definitions[[20]](#footnote-20).](http://esa.un.org/unpp/definition.html) We modify the region definitions to take into account which countries have HEP interests and to try and ensure the countries in a region have similar performance.



Figure 2 Major regions of the world for PingER aggregation by regions, countries in white are not monitored

More details on the regions are provided in Table 1 that highlights the number of countries monitored in each of these regions, and the distribution of population in these regions.

Table 1: PingER Monitored Countries and populations by region Dec 2013

|  |  |  |  |
| --- | --- | --- | --- |
| **Regions** | **# of Countries** | **Population of the Region (Millions)** | **% of World Population** |
| Africa | 49 | 1110 | 15% |
| Balkans | 10 | 69 | 1% |
| Central Asia | 9 | 78 | 1% |
| East Asia | 4 | 1618 | 23% |
| Europe | 31 | 673 | 8% |
| Latin America | 22 | 616 | 8% |
| Middle East | 15 | 309 | 4% |
| North America | 3 | 347 | 5% |
| Oceania | 5 | 38 | 1% |
| Russia | 1 | 142 | 2% |
| S.E. Asia | 8 | 618 | 9% |
| South Asia | 8 | 1616 | 23% |
| **Total** | 165 | 7162 | 99% |

## Historical Growth of PingER Coverage Since 1998

Figure 3 shows the growth in the number of active[[21]](#footnote-21) sites monitored by PingER from SLAC for each region since 1998. As can be seen, initially the main regions monitored were North America, Europe, East Asia, and Russia. These were the regions with the main HEP interest. Starting in 2003-2004, the increased number of hosts monitored in developing regions such as Africa, Latin America, Middle East and South Asia is very apparent. In the last year (2013), the number of hosts monitored by SLAC has increased from 625 to 717 mainly driven by the collaboration with Malaysia, and an increase in the hosts in S. E. Asia including Malaysia.



Figure 3 : Number of sites monitored from SLAC by region at the end of each year 1998 – 2013

## Metrics

To assist in interpreting the losses in terms of their impact on well-known applications, we categorize the losses into quality ranges.  These are shown below in Table 2.

Table 2: Quality ranges used for loss

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Excellent** | **Good** | **Acceptable** | **Poor** | **Very Poor** | **Bad** |
| Loss | <0.1% | >=0.1% &  < 1% | > =1%  & < 2.5% | >= 2.5%  & < 5% | >= 5%  & < 12% | >= 12% |

The major effects of packet loss and RTT can be found in the [Tutorial on Internet Monitoring & PingER at SLAC[[22]](#footnote-22)](http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html#loss), briefly:

* At losses of 4-6% or more video-conferencing becomes irritating and non-native language speakers are unable to communicate effectively. The occurrence of long delays of 4 seconds or more (such as timeouts in recovering from packet loss) at a frequency of 4-5% or more is also irritating for interactive activities such as telnet and X windows. *Conventional wisdom among TCP researchers holds that a loss rate of 5% has a significant adverse effect on TCP performance, because it will greatly limit the size of the congestion window and hence the transfer rate, while 3% is often substantially less serious,* Vern Paxson. A random loss of 2.5% will result in Voice over Internet Protocols (VoIP) becoming slightly annoying every 30 seconds or so. A more realistic burst loss pattern will result in VoIP distortion going from not annoying to slightly annoying when the loss goes from 0 to 1%. Since TCP throughput for the standard (Reno based) TCP stack according to Mathis et. al. goes as *1460\*8bits/(RTT\*sqrt(loss))*[[23]](#footnote-23) it is also important to keep losses low for achieving high throughput. How well this formula works in practice and dealing with zero packet loss is discussed elsewhere[[24]](#footnote-24)
* For RTTs, studies in the late 1970s and early 1980s showed that one needs < 400ms for high productivity interactive use. VoIP requires an RTT of < 250ms or it is hard for the listener to know when to speak.

### Yearly loss trends:

Figure 4 shows the packet losses seen from SLAC to world regions for 1998 through 2013. Since losses are mainly dependent on the network edges, they tend to be independent of distance.



Figure 4: Packet Loss measured for various regions from SLAC till Dec 2013

It is seen that losses are lowest (best) for North America, East Asia, Europe and Oceania. They are highest (worst) for Central Asia and Africa. The improvement from year to year is roughly exponential with a factor of ~ 30 improvement in 10 years. The 2013 uptick in losses for the Russian hosts is interesting. Similar degraded performance in 2013 is also seen below in the jitter and throughputs (see Figure 8, Figure 9, Figure 10, and Figure 12). The effect appears to be prevalent in 11 of the 12 measured hosts. The exception is the host at the Ekaterinoburg Ural Branch Russian Academy of Sciences. The main contributors to the uptick are the hosts around Moscow such as Moscow State University, IHEP and ITEP. More information on the degredation in performance to Russia seen from SLAC can be found in the Russia 2013 case study[[25]](#footnote-25).

### Yearly minimum RTT

In general for wide area paths the minimum RTT depends mainly on the length of the path between the two hosts. If the path uses a GEOS link then the round trip path for the radio signal is ~450-500ms[[26]](#footnote-26). As can be seen for minimum RTT from SLAC to countries of the world in Figure 5, there is a clear minimum RTTs threshold between 400 and 500ms between terrestrial paths and paths with GEOS links. It is also clear that most of the remaining countries monitored that still use GEOS links are in Africa.



Figure 5: Minimum RTTs measured from SLAC to world countries, Dec 2013. Different regions have different colors.

If one looks at maps of the minimum RTTs from SLAC to the world in 2008 and 2013 in Figure 6, one can see the countries that have moved away (moved from red to another color) from using GEOS links. One can see Greenland and Central Asia now use terrestrial links. The biggest

The installation of fibre cables from Europe and from the Middle East to East African countries driven by the 2010 soccer world cup in South Africa dramatically reduced round trip delays. The Democratic Republic of the Congo was finally connected to low-cost, high-quality international bandwidth through the WACS submarine fibre optic cable in 2013[[27]](#footnote-27). Also a fibre connection to the South East Democratic Republic of the Congo from Zambia was completed in Nov 2012[[28]](#footnote-28).

The African countries, that we can measure, that still have minimum RTTs > 500msec are: Chad, Guinea, Eritrea, Liberia and Sierra Leone. Try as we might we cannot find hosts that respond and that we trust as being in the Central African Republic, Congo Republic, Equatorial Guinea, South Sudan, or Western Sahara.

* Republic of Congo is building a fibre infrastructure around the capital Brazzaville, and will connect to Pointe Noire on the coast where the submarine fibre optic cable lands[[29]](#footnote-29)
* It is expected that Equatorial Guinea, Gambia, Guinea, Liberia, Mautitania, Sao Tome and Principe, and Sierra Leone and will get their first submarine connections via the ACE project in 2013.[[30]](#footnote-30) The cable will also connect via terrestrial fibre networks in the landlocked countries of Mali and Niger. The first phase of the 17,000 km-long fiber optic cable was put in service on December 15, 2012. Sierra Leone is supported by a loan from the World Bank.[[31]](#footnote-31)



Figure 6: Maps of minimum RTTs from SLAC to the world in 2008 and 2013.

Looking at the minimum RTT as a function of time for many countries since 1998 seen in Figure 7, one can clearly see, by the precipitous drop in minimum RTT, when monitored hosts in various countries moved from GEOS satellite connections to terrestrial connections.



Figure 7: Time series of minimum RTT measured from SLAC to various countries since 1998

It can be seen that China shifted to a terrestrial link in 2001, Georgia in 2003, Indonesia and Belarus in 2004, Mongolia in 2006. Although Ghana received terrestrial connectivity in 2001, the nodes we were monitoring did not shift to a terrestrial connection until 2007. Rwanda, Kenya, Zimbabwe and Madagascar are some of the countries that started using terrestrial connections in 2010 due to the Soccer world cup in Africa. Cuba and Congo shifted to terrestrial link this year. RTT to Cuba first decreased to around 400ms in 2012 as one way link switched to terrestrial and then reduced to 170ms in 2013 after complete switch to terrestrial link took place26A.

### Yearly Jitter trends

The short term variability or "jitter" of the response time is very important for real-time applications such as telephony, gaming, streaming media, haptic surgery. Jitter is a symptom that there is congestion, or not enough bandwidth to handle the traffic. The jitter specifies the length of the VoIP codec de-jitter buffer required to prevent over- or under-flow. An objective could be to specify that say 95% of packet delay variations should be within the interval [-30msec, +30msec].The links we measure generally have high performance backbones, so the main cause of jitter is the congestion at the edges. Thus the metric tends to be distance independent. There are several ways of measuring jitter[[32]](#footnote-32)B, for this report we shall report on the Inter Packet Delay Variability (IPDV). Figure 8 shows the jitter seen from SLAC to regions of the world.



Figure : Jitter from SLAC to regions of the world until Dec 2013.

It is seen that the improvement to lower values of jitter is roughly exponential for regions with poor jitter, while for the good regions the improvements with time since 2006 appear to have leveled out. Also as for the losses the developed regions N. America, Australasia, E. Asia and Europe, recently joined by Russia have the lowest jitter. Africa, S.E. Asia and S. Asia have the highest jitter.

### Yearly Throughput Trends

To combine the effects of loss and RTT we use the Mathis formula to calculate the TCP throughput. We have also reduced the effect[[33]](#footnote-33) of the *1/RTT* in the Mathis formula for derived throughput by normalizing the throughputs using:

*norm\_throughput = throughput \* min\_RTT(remote region) / min\_rtt(monitoring\_region)*,

where: *throughput ~ 1460\*8 bits/(RTT\*sqrt(loss))*

The results are shown in Figure 9 showing data averaged into yearly intervals.



Figure 9: Yearly averaged normalized derived TCP throughputs from the SLAC to various regions of the world until Dec 2013.

The improved performance for Russia in 2001 is an artifact of measuring a single host (the Institute of High Energy Physics in Moscow) with high losses in 1999-2000 and adding a second host (Institute of High Energy Physics in Novosibirsk) with low losses in 2001. Parts of Latin America moved from satellite to fibre in 2000, and E. Asia in 1999. Also note the impact of moving the ESnet routing from E. Asia (in particular Japanese academic and research networks) to the US via New York in 2001 to a more direct route via the West Coast of the US. Also note that there is almost a 10 times difference in throughput between Africa and N. America, Europe and Oceania. Africa is the worst off region and had the slowest rate of improvement until 2008.

To make the overall changes stand out more clearly, Figure 10 shows just exponential trendline fits to monthly averages of the derived throughput on a log-linear scale (exponentials show up as straight lines). We have excluded N. America due to the distortion produced by the high value of 1/RTT, since the measurements are made from SLAC in N. America. These trendlines are useful to see in general how a particular region is doing against others and over a period of time, against its past. For guidance, the top yellow line shows what a 20% improvement per year would look like; this corresponds to roughly a factor of 10 in twelve years.



Figure 10: Derived Throughput kbits/sec from SLAC to the World (since the throughputs in this graph are not normalized we have not shown N. America) until Dec 2013. The yellow line is to help show the rate of change. If one extrapolates Europe’s performance backwards to February 1992, it intercepts Africa’s performance today.

The slow increase for Europe in Figure 9 is partially an artifact of the difficulty of accurately measuring loss with a relatively small number of pings (14,400 pings/month at 10 pings/30 minute interval, i.e. a loss of one packet ~ 1/10,000 loss rate). We looked at using a method[[34]](#footnote-34) that allows for zero packet loss, however it requires one to know the maximum congestion window size. Unfortunately this varies from host to host and can easily be changed, so we did not pursue it.

Looking at the data points one can see:

* East Asia and Oceania are catching Europe;
* Russia is 6 years behind Europe and catching up;
* Latin America and the Middle East are 7 years behind and falling further behind;
* S. E. Asia is also 7 years behind but is catching up;
* S. Asia and Central Asia are 11 years behind and keeping up;
* Africa is 14 years behind Europe, however see Figure 11 where it is seen:
  + In 2008-2009 Africa was ~ 15 years behind Europe and even worse was falling further behind such that in 2030 it would have been 60 times worse off or almost 25 years behind. Prior to 2008 the rate of improvement was a factor of 2 in 7 years
  + Since 2008 the improvement is a factor of 3 in 3 years and at the current rate it could catch up with Europe by around 2030.
  + This remarkable improvement is largely a reflection of the impact of the multiple terrestrial links installed since 2008[[35]](#footnote-35), initially driven by the soccer world cup.



Figure : Extrapolations on the throughput data.

### View from Europe

To assist in developing a less N. American view of the Digital Divide; we added many more hosts in developing countries to the list of hosts monitored from CERN in Geneva Switzerland. We now have data going back for eight years that enables us to make some statements about performance as seen from Europe. Figure 12 shows the normalized throughput data from CERN to the rest of the world.



Figure 12: Normalized throughputs to various regions as seen from CERN until Dec 2013.

As can be seen by comparing Figure 9 and Figure 12 the general performance changes seen from CERN are very similar to those seen from SLAC.

### Variability of performance between and within regions

The throughput results, so far presented in this report, have been measured from North American sites and to a lesser extent from Europe. This is since there is more data for a longer period available for the North American and European monitoring hosts. Table 3 shows more details of all the measured throughputs seen between monitoring and remote/monitored hosts in the major regions of the world. Each column is for all monitoring hosts in a given region, each row is for all monitored hosts in a given region. The cells are colored according to the median throughput quality for the monitoring region/monitored region pair. White is for derived throughputs > 10,000 kbits/s (good), green for <= 10,000 kbits/s and >5,000kbits/s (acceptable), yellow for <= 5,000kbits/s and > 1,000 kbits/s (poor), pink for <= 1000kbits/s (very poor) and > 100kbits/s red for <= 100kbits/s and > 1 kbits/s (bad), and grey for no measurements. The Monitoring countries are identified by the Internet two-character Top Level Domain (TLD). Just for the record CA=Canada, US=NET=GOV=United States, CH=Switzerland, DE=Denmark, UK=United Kingdom, AU=Australia, CN=China, HK=Hong Kong, KR=South Korea, TW=Taiwan, BO=Bolivia, MX=Mexico, IN=India, LK=SriLanka, PK=Pakistan, SU=Russia, DZ=Algeria, ZA=South Africa and BF=Burkina Faso. E. Asia includes China, Japan, South Korea, Taiwan; S. Asia is the Indian sub-continent; S.E. Asia includes Indonesia, Malaysia, Singapore, Thailand and Vietnam.

Table 3: Derived throughputs in kbits/s from monitoring hosts to monitored hosts by region of the world for November 2013. The columns are for the monitoring hosts, the rows are for the remote (monitored hosts)



As expected it can be seen that for most TLDs (represented in columns) the best possible throughput values represented by the outlined boxes, usually exist within a region. For example for regions with better Internet connectivity, such as Europe, higher throughput is seen from European monitoring sites (CH, DE, IT and UK) to TLDs that lie within Europe. However, if the regions are close enough in terms of connectivity, throughput values are relatively higher. For example performance is better between closely located regions such as: the Balkans and European countries; Russia and E. Asia (the Russian monitoring site is in Novosibirsk. This shows that network performance is not completely dependent on geographic proximity, but rather on how close the regions are on the map of Internet connectivity and performance. Also take for example Burkina Faso (BF) in Western Africa: higher throughput values are evident between Africa and the TLDs DE, IT, and CH rather than between BF and African sites. This serves to illustrate the poor intra-regional connectivity between some parts of Africa.

This table also shows that throughput values show large variability within regions (e.g. a factor of ten or more between Burkina Faso and Egypt or South Africa). To provide further insight into the variability in performance for various regions of the world seen from SLAC, Figure 13 shows various statistical measures of the losses and derived throughputs. The regions are sorted by the median of the measurement type displayed. Note the throughput graph uses a log y-scale to enable one to see the regions with poor throughput.

It is apparent that the Interquartile range (IQR)[[36]](#footnote-36) can span one or more orders of magnitude. The most uniform region (in terms of IQRfor both derived throughput and loss) is Central Asia, probably since most of the paths use a GEOS link. The most diverse are Europe and East Asia. For Europe, Belarus stands out with poor performance. For East Asia, China stands out with relatively lower performance in terms of derived throughput.



Figure 13: maximum, 95, 90, 75 percentile, median, 25 percentile and minimum derived throughputs of various regions measured from SLAC for Nov 2013 and ordered by median throughput.

A further example of the variability within a region is to look at S.E Asia as seen in Figure 14.



Figure : Derived throughput from SLAC to S.E. Asia countries until Dec 2013.

It is seen that Singapore (SG) has roughly four times better performance than the next S.E. Asian country Malaysia (MY). It is also seen that the performance to Singapore is almost a factor of ten better than that to Laos (LA). In fact Singapore’s performance approaches that of E. Asian countries.

### Yearly Mean Opinion Score (MOS) trends

The MOS is used by the telecom industry to categorize the quality of phone calls. The MOS can be related to the loss, RTT and jitter of the circuit[[37]](#footnote-37). With the emergence of Voice over IP (VoIP), it has become increasingly important to know what values of MOS are available on the Internet. A value of 5 is a perfect call; a value of 1 is no ability to communicate. The maximum MOS for VoIP is about 4.4. Typical values for usable VoIP are 4.2 to 3.5. Values below result in increasing frustration and inability to communicate. Figure 15 shows MOS values from SLAC to regions of the world.

It is seen that VoIP should work well from SLAC to most regions, and be usable to Central and South Asia[[38]](#footnote-38). It is not usable for most of Africa, however with the increased deployment of terrestrial links replacing GEOS links, it is improving



Figure 15: Mean Opinion Scores derived from PingER measurements from SLAC to regions of the world until Dec 2013.

### Directivity

Knowing the locations of the monitor and remote host we can calculate the great circle distance between them (D). The velocity of light in fibre (or electric pulses in copper cable) is v ~ (2/3)\*c = 200,000 km/sec (where c is the velocity of light in vacuum ~ 300,000km/sec) and the round\_trip\_distance = 2 \* D or D = (v / 2) \* round\_trip\_time (to\_travel\_distance\_D). To take account of the fact that the route between the monitor and target may not be a great circle route and there may be some other delays due for example to network equipment, we introduce the Directivity factor. Directivity = 1 for a great circle route and negligible network device delays. For the round\_trip\_time we use the minimum RTT (min\_RTT) measured by Pinger to minimize the effects of network queuing etc. Then:

*D [km] = Directivity \* min\_RTT [msec] \* 100 [km/sec]*

and

*Directivity = D / (min\_RTT \* 100).*

Since Directivity <=1, if we derive a value > 1 this indicates the coordinates (longitude and longitude) of either (or both) the monitor or remote are incorrect, which is a valuable diagnostic. Large values of Directivity indicate the route is pretty direct, smaller values indicate it is less direct. Figure 16 shows the Directivity measured from SLAC to world regions



Figure : Directivity measured from SLAC to regions of the world until Dec 2013.

It is seen that more recently, as more remote hosts are monitored and GEOS links give way to terrestrial links, Directivity is stabilizing. The reason for the high Directivity values for E. Asia, S.E. Asia, and Australasia is since the routes from SLAC go fairly directly across the Pacific. The smaller Directivity routes to N. America, Europe, M. East, Africa and Latin America must first pass through N. America. The Russian and C. Asian connections have also to cross Europe. S. Asia connections typically cross the Pacific to Japan and then South-south-west to Singapore and then North-west to S. Asia.

The Directivity between the monitoring host countries and the regions of the world are seen in Table 4.

Table : Directivity from monitors in countries to regions of the world for Nov 2013.



The table reveals some interesting low Directivity connections:

* The low values between China (CN – IHEP in Beijing) and E. Asia are typically via Hong Kong, ie. ~ 3500km rather than a direct route of 1300km for say Beijing to Tokyo.
  + Routes between India (IN) to Pakistan (PK) go via Singapore (adding about 4700km to the direct route of 21000kmfrom Islamabad to Mumbai) and to Afghanistan via GEOS links.
* The routes between Jordan (JO) and: Israel, Iran and Turkey go via Europe
* Routes from Bolivia (BO) to Brazil (BR) go via Miami

## Comparisons with Economic and Development Indicators

The United Nations Development Program (UNDP), the International Telecommunication Union (ITU) and various other task forces are independent expert bodies which aim to provide policy advice, based on various economic factors, to governments, funding sources, and international organization for the purpose of bridging the Digital Divide. See the table below for details.

In reality, it is difficult to classify countries according to their development. The main challenges lie in determining what factors to consider, how to measure them, how useful and pervasive they are, how well defined they are, and whether they are uniformly measured for every country. Various organizations such as the ITU, UNDP, the Central Intelligence Agency (CIA), World Bank etc. have come up with Indices based on measured items such as life expectancy, GDP, literacy, phone lines, Internet penetration etc. Although many of the above challenges are not insurmountable they can require a lot of cost and time. Thus, many of these measurements are outdated and may not depict the current state of the country.

One of the most important factors determining the economic development of a country in today’s information age is its Internet connectivity. Thus we may expect moderate to strong correlations between the economy and development of a country with its Internet performance measurement. The significance of using PingER’s Internet measurements to characterize a country’s development is due to the fact that PingER’s data is current (up-to date within a day or so compared to say the most recent Information and Communications Technology (ICT) Development Index[[39]](#footnote-39) (IDI) data from the ITU that was published in 2012 and was measured in 2011) and covers countries of the world. Also PingER’s data is very objective not relying on statistics reported by governments with possibly vested interests. From the many indices available[[40]](#footnote-40), we chose the IDI since it is one of the most current, it is clearly related to Internet connectivity, and covers many of the countries covered by PingER. A bubble plot of throughput vs. the IDI access sub-index[[41]](#footnote-41) with the bubble sizes being proportional to population is shown in Figure 17.



Figure : Pinger Derived Throughput vs. the ICT Development index. The bubble size is proportional to the country's population.

The black line is a trend line fit to a power series with the parameters shown. It is seen that there is a fairly strong correlation (R2=0.566 or R=0.75) between the IDI and PingER derived throughput.

## PingER Case Studies

**Sudan**

Sudan was disconnected for about 2 hours starting 10:30 25 Sep GMT. PingER was actively monitoring one host in the Sudan i.e. [www.sustech.edu](http://www.sustech.edu/) (41.67.53.4, AS 37197 Sudan Educational network) at the Sudan University of Science and Technology, the plot of packet loss and RTT from SLAC is shown in figure 18, there is a clear disconnection on 25/9.



**Figure 18: connectivity with host www.sustech.edu (SD.SUSTECH.EDU.N1) in Sudan**

**Syria**

Internet connectivity with Syria was lost from 19:08 May 7 GMT to 14:37 May 8 GMT (about 20 hours). We have two active hosts in Syria, one is run by an ISP ([www.inet.sy](http://www.inet.sy/)), the other by a publishing company ([www.thawraonline.sy](http://www.thawraonline.sy/)), both of them stopped responding during this time frame; figure 19 shows 100% packet loss (between May 7 and May 8) to one of these nodes in Syria being monitored by SLAC. This incident was confirmed by other sources.



**Figure 19: connectivity with host www.inet.sy (SY.INET.N1) in Syria**

**High Performance Network Monitoring**

***New and Ongoing Monitoring and Diagnostic Efforts in HEP***

PingER is an excellent system for monitoring the general health and capability of the existing networks used worldwide in HEP. However, we need additional end-to-end tools to provide individuals with the capability to quantify their network connectivity along specific paths in the network, quickly isolate and diagnose network issues and inform developing tools and applications that need network information to be optimize their tasks. These are needed both to ascertain the user's current network capability as well as to identify limitations which may be impeding the user’s ultimate (expected) network performance. The perfSONAR project has grown to fill this need and we will show later how it is being leveraged to address the needs of the LHC community.

Most HEP users are not "network wizards" and don't wish to become one. In fact as pointed out by Mathis, the gap in throughput between what a network wizard and a typical user can achieve was growing significantly from the late 1980’s to the late 1990’s.

Within the last 15 years, because of improvements in default OS TCP stack settings, new protocols, hardware, firmware and software, this gap has decreased but still remains in 2014. Because of HEP's critical dependence upon networks to enable their global collaborations and grid computing environments, it is extremely important that more end-user focused tools be developed to support these physicists and continue to decrease the gap between what an expert can achieve and what a typical user can get “out of the box”.

Efforts continue in the HEP community to develop and deploy a network measurement and diagnostic infrastructure which includes end hosts as test points along end-to-end paths in the network. This is critical for isolating problems, identifying bottlenecks and understanding infrastructure limitations that may be impacting HEP’s ability to fully utilize their existing networks. Originally the [E2E piPEs project](http://e2epi.internet2.edu/E2EpiPEs/index.html)[[42]](#footnote-42), the [NLANR/DAST Advisor project](http://dast.nlanr.net/Projects/Advisor/)[[43]](#footnote-43) and the [LISA](http://monalisa.cern.ch/monalisa__Interactive_Clients__LISA.html)[[44]](#footnote-44) (Localhost Information Service Agent) initiated developing an infrastructure capable of making on demand or scheduled measurements along specific network paths and storing test results and host details for future reference in a common data architecture. Since 2008 the [perfSONAR](http://www.perfsonar.net/description.html) project has become the organizing entity for these efforts and is broadly supported (see below). The perfSONAR effort is utilizing the [GGF NMWG](http://www-didc.lbl.gov/NMWG/)[[45]](#footnote-45) schema to provide portability for the results. This information can be immediately used to identify common problems and provide solutions as well as to acquire a body of results useful for base-lining various combinations of hardware, firmware and software to define expectations for end users. In addition the [perfSONAR-PS](http://code.google.com/p/perfsonar-ps/wiki/Home) [toolkit](http://psps.perfsonar.net/toolkit/) distribution includes many of the recognized, standard tools (PingER, NDT, Advisor, Iperf, traceroute server etc.) in network testing and diagnosis

**perfSONAR for HEP**

In fall of 2012 the Worldwide LHC Computing Grid (WLCG) formed an Operations task force to guide the deployment and configuration of the perfSONAR-PS Toolkit at all WLCG Tier-1 and Tier-2 sites worldwide. This decision was made in great part based upon both the success in using perfSONAR-PS in the US and within the LHCOPN as well as the increasing need to better integrate networks within our HEP global cyberinfrastructure.

As mentioned above, most HEP users are not interested in becoming network wizards nor do they have the expertise to diagnose network related problems. Starting in 2007 there was a significant effort made to deploy and integrate perfSONAR at all Tier-1/Tier-2 sites in the US to provide a standardized set of tools and corresponding network measurements to aid in problem isolation and diagnosis as well as for baseline monitoring. This expanded to include the LHCOPN in 2011. The plan for USATLAS was to deploy two perfSONAR instances (each on their own, identical hardware) at each distinct Tier-2 site (as well as the Tier-1 at BNL) and this plan has been used as the basis for the global expansion of HEP installed perfSONAR-PS Toolkits. In 2012 the footprint of installed instances had grown significantly and included sites in Europe, Asia and South America.

Since many Tier-2’s are physically distributed at more than one location, more than 2 systems per Tier-2 are deployed. It was important that all sites deploy identical hardware to remove hardware variations that might impact measurements. An inexpensive system with two 1GE onboard NICs (~$635) from KOI computing was identified in fall 2008 and was been deployed at 8 Tier-2 sites and BNL. Two systems per site are required to allow both throughput and latency tests to be undertaken which would interfere with each other if they ran on the same system. Since these systems were purchased, some issues with the particular 1 GE NIC and hard-disk controller have been identified and Internet2 has created new recommendations for future perfSONAR-PS purchases. During 2010 the University of Michigan purchased a Dell R410 system as a possible “integrated” hardware node intended to run both latency and bandwidth tests from a single system. This node has been made available to the perfSONAR-PS developers and we had hoped to have a perfSONAR-PS deployment capable of utilizing this hardware sometime in 2012. Unfortunately not as much progress was made in this area as we had hoped. Near the end of 2011 we evaluated Dell R310 systems in the US and Dell R610 systems in Canada as possible hardware replacements for the aging KOI systems. Dell agreed to provide LHC pricing for “perfSONAR-PS” systems on their LHC portal and as of January 1, 2012, two versions (R310 and R610 based) were available. As of 2013 the Dell systems are no longer available since the generation 11 systems upon which they were based were superseded by newer 12th generations models (R320 and R620). While sites can no longer buy “perfSONAR” nodes from the Dell LHC portal, many have successfully upgraded to 12th generation Dell hardware without any issues. ESnet continues to maintain a hardware recommendation site that can inform those beginning new deployments: http://fasterdata.es.net/performance-testing/perfsonar/ps-howto/hardware/

The perfSONAR systems in WLCG are intended to run full-mesh tests for both throughput and latency with all other WLCG Tier-2’s and the Tier-1. The latency role is assigned to the first node (typically designated with a ‘1’ in the DNS name) by convention while the throughput role is assigned to the second node. Installation is made straightforward by simply booting a recent ISO image provided by Internet2 and doing a one-time configuration of the node. Configuration results as well as measurements are persisted onto the local hard-disk of the system being configured.



Figure : Example OWAMP test between two USATLAS Tier-2's (bi-directional testing is enabled) showing the minimum one way delays and losses for each direction.



**Figure 27: Example perfSONAR PingER round trip time graph between the Tier-1 and a Tier-2. The line at the bottom is the mean RTT.**

The latency tests have two variations. The first is managed by OWAMP and measures one-way delays between the latency node and its test partner at another site. Since absolute time accuracy is critical for this test, part of a latency node configuration includes setting up a reliable time service (ntpd) configuration to ensure the node keeps accurate time. The second type of latency test is provided by PingER which is configured by default to send 600 ping packets per minute between the latency node and its target and track the results. Examples of both types of tests are shown in Figure 26 and Figure 27.



**Figure 28: Example bi-directional throughput test in perfSONAR between two Tier-2s**

The second perfSONAR node measures throughput using Iperf. Example results are shown in Figure 28 which illustrates some interesting throughput changes during the 1 month of measurements shown.

As of the perfSONAR-PS v3.3 release we have two additional tests that we are configuring: traceroute and PingER. The new version has the ability to track the traceroute path between the host and any destination specified. The default is to check each destination every hour. If the route changes, we record the new route in the perfSONAR-PS measurement archive. This allows us to better understand the topology we are measuring and can alert us to routing changes that may be correlated with observed network problems. The PingER tests run to check the round-trip path between hosts and augment the OWAMP information.

One significant improvement in the maintainability of the perfSONAR-PS Toolkit installations was added in spring of 2013. In the original perfSONAR-PS deployment for WLCG each sites was responsible for updating their test configuration as other sites were added, removed or changed. This led to significant misconfigurations over time as sites didn’t keep up with changes in their test-mesh. Aaron Brown/Internet2 added the ability for perfSONAR-PS toolkits to use a “mesh-configuration” that was centralized. Each set of related hosts (a Tier-1 cloud or similar group) could define a single JSON configuration file and make it available at a web-accessible URL. Any changes to the configuration could automatically be applied by an agent on the toolkit. This was critical in enabling perfSONAR-PS deployment at the scale of the WLCG.

Details about how to setup and configure perfSONAR-PS for WLCG are maintained at <https://twiki.cern.ch/twiki/bin/view/LCG/PerfsonarDeployment>

**perfSONAR Modular Dashboard**

While perfSONAR provides a convenient way to gather standardized network metrics via deploying a toolkit instance, it can still be cumbersome to try to gather, check and interpret this data. As USATLAS began deploying perfSONAR-PS instances we realized that a critical missing component was a means of centrally monitoring and displaying the information we were collecting. ESnet had some initial efforts in this direction by creating [Nagios](http://www.nagios.org/) “plugins” that could query individual perfSONAR-PS instances and check to see if the data returned was within bounds.

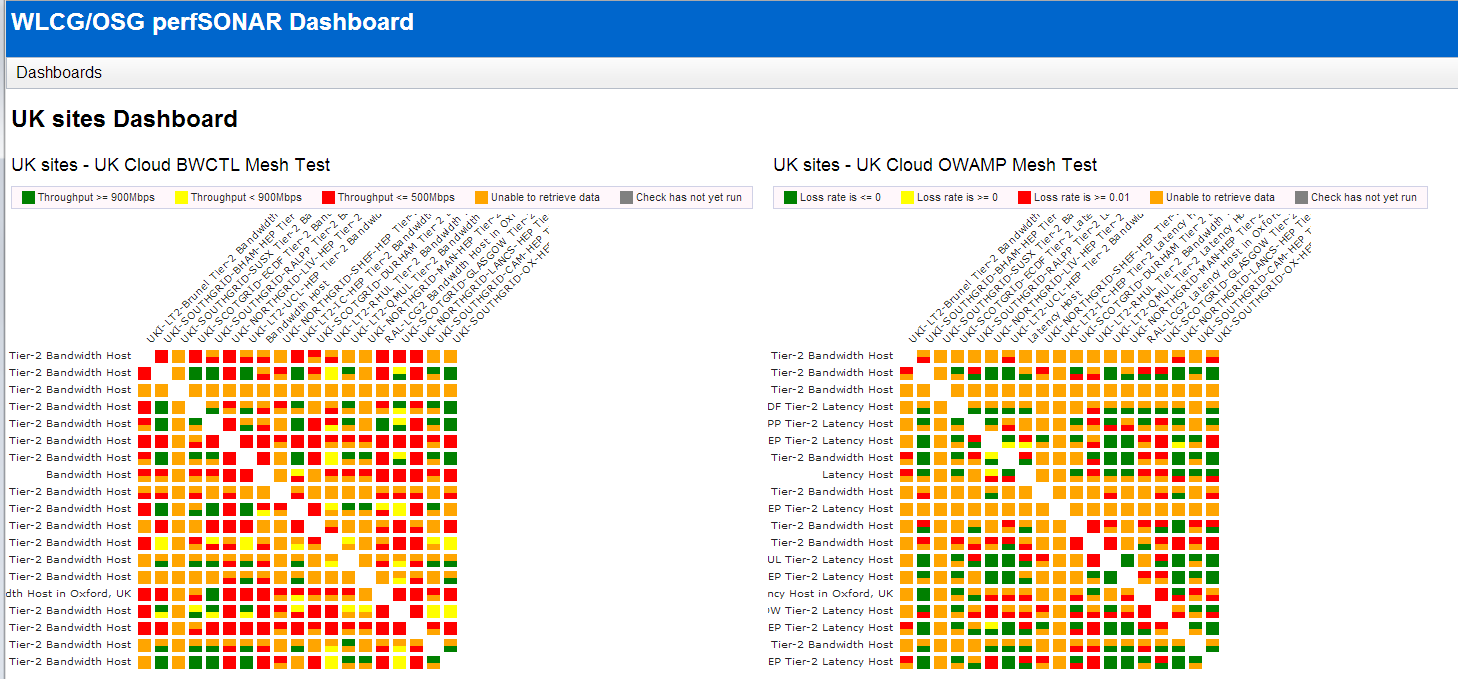
Tom Wlodek at BNL had been working closely with the USATLAS Throughput working group in developing a set of service and measurement tests that can be tracked in customized server pages setup at BNL which were monitoring our deployed instances, initially via Nagios. During fall of 2010 additional service monitoring capabilities (modules) were added to help manage the monitoring task and alert on service failures. The code rapidly grew based upon user requests and by 2011 we realized that we would need to redesign the architecture and re-implement the “modular dashboard” we had been using to monitor perfSONAR. The original dashboard was still in operation in 2013 but significant efforts have been put forth on creating the “new” modular dashboard during 2012 and 2013.

The modular dashboard has a public area in GitHub. We choose GitHub to have a well-supported, neutral and free development platform. At the end of 2012 we created a new GitHub Organization called PerfModDash to control and manage the needed development efforts surrounding our work. Initially we have two projects hosted in repositories under this Organization: OldDashboard and PerfModDash (the new dashboard). See <https://github.com/PerfModDash/PerfModDash>

Unfortunately Tom Wlodek (the primary Modular Dashboard author) left BNL in December 2013 and the project was orphaned. However WLCG and OSG still require the functionality the dashboard was to provide. A special meeting between WLCG, OSG, ESnet and Internet2 was held December 17, 2013 to discuss our options. It was decided to test two open-source projects as possible modular dashboard replacements: MaDDash (developed as a light-weight dashboard for ESnet use) and Open Monitoring Distribution (OMD) which packages many related Nagios components as a single RPM. The tests look promising and the MaDDash test instance can be seen at <http://maddash.aglt2.org/maddash-webui/>

**MaDDash and OMD**

Now that the modular dashboard is orphaned we are exploring the use of MaDDash as a replacement. To replicate the capabilities of the basic service checks that the modular dashboard provided we are also testing OMD. There is a Wiki describing what it takes to install and configure them for WLCG use at <https://twiki.cern.ch/twiki/bin/view/LCG/MadDashWLCG>



**Figure 29 Example MaDDash display for the UK (United Kingdom) perfSONAR-PS cloud. Shown are the two matrices for Bandwidth measurements (left) and Latency measurements (right) for the sites participating in the UK cloud. Orange indicates a missing measurement while red, yellow and green are critical, warning and OK states for the corresponding network metrics.**

The WLCG and OSG perfSONAR groups have provided important feedback to the perfSONAR distribution developers and has identified a number of bugs which impact the reliability and usability of the distribution. As of v3.3.1 (released in Fall 2013) we have achieved a level of robustness that allows us to recommend broad deployment of perfSONAR-PS toolkits to all WLCG sites. The WLCG perfSONAR-PS deployment taskforce has set a deadline of April 1, 2014 to have all WLCG sites deploy the latest version of perfSONAR-PS and configure it to use the centralized “mesh-configuration”.

It should also be noted that having perfSONAR services running co-located with important resources provides the ability to run “on-demand” tests using the broadly deployed NDT and NPAD tools. These tools can be run from any remote location, testing to any perfSONAR instance. This is a very important additional capability that can be vital in network diagnosis.

PerfSONAR is already broadly deployed in Research and Education network PoPs (Internet2, GEANT2, ESnet) and it is hoped that more collaborations within HEP will deploy perfSONAR instances co-located with their computing and storage resources. Having a more extensive deployment significantly improves the value and applicability of the overall perfSONAR infrastructure for HEP.

***LHCONE Monitoring History***

Discussions about extending LHC networking beyond the Tier-0/Tier-1 centers started during an LHCOPN meeting in the summer of 2010. By January 2011 there was a plan in place to begin deploying a prototype “network” called LHCONE[[46]](#footnote-46). The goal for LHCONE was both to provide Tier-n LHC sites better networking and to more easily allow the backbone and regional networks to manage and support LHC networking needs. One requirement for LHCONE was to not impact the successful and operational LHCOPN.

As sites started to test the prototype architecture for LHCONE, it quickly became clear that we needed a way to monitor this new infrastructure. In fall 2011 we decided to setup perfSONAR-PS monitoring for LHCONE sites. We leveraged the existing modular dashboard to track the LHCONE network that was used for the various ATLAS clouds and the LHCOPN. Shown in Figure 31 is an example of the LHCONE monitoring matrices for bandwidth and throughput setup on the BNL modular dashboard.

The initial monitoring goal for LHCONE is to measure the current network performance between all the candidate LHCONE early adopter sites to set a baseline. Then, as sites join LHCONE, we can use the ongoing measurements to document how the networking between sites changes as LHCONE is adopted.



**Figure 31: The “selected site” LHCONE monitoring page showing the early setup for monitoring. The URL is** [**https://perfsonar.usatlas.bnl.gov:8443/exda/?page=25&cloudName=LHCONE**](https://perfsonar.usatlas.bnl.gov:8443/exda/?page=25&cloudName=LHCONE)

This monitoring is now deprecated in favor of the WLCG deployment worldwide but it was useful in tracking the transition of sites into LHCONE.

**Additional Complementary End-to-End Monitoring**

The US ATLAS collaboration has made an extensive effort to improve the throughput of their Tier-1 and Tier-2 centers and has coupled this with active testing and monitoring to track performance over time. Bi-weekly meetings of the USATLAS Throughput working group focus on throughput, network measurements and related site issues. During 2013 we expanded the Throughput meetings to include representatives from the USCMS sites. This group is working in two primary areas: 1) automated transfer throughput testing using ATLAS production systems and 2) deployment and integration of perfSONAR at all USATLAS Tier-2 sites and the Tier-1 site at Brookhaven. We will discuss perfSONAR deployment and experience in USATLAS in the next section and will focus on the automated (and manual) throughput testing USATLAS is using here.



**Figure 32 Example production system throughput test results between two USATLAS sites: BNL (Tier-1) and AGLT2 (Tier-2). Shown are the storage areas: AGLT2\_DATADISK and BNL\_OSG2\_MCDISK from November 21, 2012 to January 11, 2013**

The perfSONAR infrastructure is intended to measure the network (LAN,WAN) between perfSONAR test nodes but this is not sufficient to characterize the “end-to-end” behavior of the distributed systems in use in HEP. The USATLAS throughput group has developed some additional automated (and manual) tests to accurately measure their system capabilities and limits. Hiro Ito (BNL) has developed an automated data transfer service which sends a fixed number of files between sites using the standard ATLAS production system and records the results. Results of these tests are available at <http://www.usatlas.bnl.gov/dq2/throughput> where you can find details on the number of successful transfers, their throughput and timing. One example graph is shown in Figure 32 which shows test dataset transfer throughput between Brookhaven (Tier-1) and AGLT2 (Tier-2). During 2009-10, the system was extended to include Tier-2 to Tier-3 tests (in addition to the standard Tier-1 to Tier-2 tests originally defined). These results, in combination with perfSONAR results, are being used to identify problems in the overall system and isolate their likely location. We made significant progress on our goal for 2011 targeting use of the production system results and the perfSONAR results during the same time periods to more accurately localize any problems that may arise. For 2013 we focused on improving the integration and analysis of the data provided to more quickly identify and localize network problems in the infrastructure and expanded the set of tests to include Tier-1 sites around the world.

**Summary: Progress in HEP Network Monitoring for 2013**

As noted above, efforts to ensure commonality in both monitoring and provisioning of networks have seen a significant increase in activity in HEP during 2013. Originally, the [GLIF](http://www.glif.is/)[[47]](#footnote-47) and [DICE](http://www.geant2.net/server/show/conWebDoc.1308)[[48]](#footnote-48) communities were both working toward implementing “managed” network services and the corresponding monitoring that will be needed to support their efforts. During 2013 there were a number of new or expanded initiatives within global HEP:

* The WLCG community began deploying perfSONAR-PS toolkits at sites to provide needed network visibility for users, site-admins and eventually higher-level services.
* The Open Science Grid (OSG) focused on provisioning network monitoring within OSG to become the definitive source of networking information for OSG sites as part of their year-2 planning for networking.
* Previous ATLAS-initiated perfSONAR-PS deployments within national regions in the US, Italy and Canada, were augmented by new national deployments in France, Germany, Spain, Taiwan, Russia and the UK and now include all four LHC experiments.
* The Worldwide LHC Computing Grid (WLCG) continued work on network issues via the networking committee chaired by Michael Ernst.
* The WLCG perfSONAR-PS deployment taskforce, chaired by Simone Campana and Shawn McKee, focused on deploying perfSONAR at all WLCG Tier-2s worldwide and integrating their registration and measurements within the WLCG monitoring framework. Non-compliant sites are being ticketed as of December 2013. The goal is to have a single infrastructure providing network monitoring for WLCG as opposed to potentially many disparate systems per experiment or region.

To summarize, 2013 has seen a convergence to a standard network monitoring infrastructure based upon perfSONAR. The goal is to have a single system to provide network related metrics for HEP. This is feasible because HEP’s use of the network is qualitatively the same between HEP collaborations. It is desirable because network providers and users don’t want to have multiple “network measurement infrastructures” making redundant (and possibly interfering) measurements nor do they wish to have to develop, deploy and support many such instances when one will do.

***Related HEP Network Resear****ch*

There has been a significant amount of research around managed networks for HEP that we should note. There are efforts funded by the National Science Foundation ([UltraLight](http://www.ultralight.org/)[[49]](#footnote-49) (finished Aug 2009), PLaNetS) and Department of Energy ([Terapaths](http://www.atlasgrid.bnl.gov/terapaths/)[[50]](#footnote-50) (finished Dec 2009), [LambdaStation](http://www.lambdastation.org)[[51]](#footnote-51) (finished 2009), [OSCARS](http://www.es.net/oscars/index.html)[[52]](#footnote-52), and the associated follow-on projects StorNet, VNODs and [ESCPS](https://plone3.fnal.gov/P0/ESCPS/)[[53]](#footnote-53) projects) which are strongly based in HEP. These projects are not primarily focused upon monitoring but all have aspects of their efforts that do provide network information applications. Some of the existing monitoring discussed in previous sections either came out of these efforts or is being further developed by them.

In summer 2010 a new NSF MRI project was funded called [DYNES](http://www.internet2.edu/ion/dynes.html)[[54]](#footnote-54). The DYNES collaboration (Internet2, Caltech, Michigan and Vanderbilt) intends to create a virtual distributed instrument capable of creating dynamic virtual circuits on-demand between the participating member sites. The LHC community and its data access and transport requirements are the primary “users” targeted by this new infrastructure. Funding was provided to initially deploy DYNES at 40 institutions and 14 regional networks within the United States. DYNES leverages prior work related to virtual circuit construction, QoS and perfSONAR to enable the required capabilities. Instrument deployment began in spring 2011. The DYNES project was completed in July 2013 but work is continuing on a best effort basis to improve the ability of DYNES sites to utilize end-to-end circuits.

During 2012, with the LHC Long Shutdown 1 (LS1) approaching in 2013 through the spring of 2015, many people realized we have an opportunity to better integrate networking within our various LHC computing and software infrastructure. Two specific proposals targeting this area were funded in 2012 and have been active in 2013:

* Advanced Network Services for Experiments (ANSE), NSF funded (Caltech, Michigan, Vanderbilt and U Texas Arlington)
* Next Generation Workload Management and Analysis System for Big Data, PANDA integration with networking, DOE ASCR funded (BNL, U Texas Arlington)

ANSE is focusing on the integration of advanced network services into the software stacks of ATLAS and CMS with a goal of transparently integrating new networking capabilities to increase the overall efficiency and effectiveness of their globally distributed computing and storage resources. The ATLAS PANDA proposal has a similar focus and is trying to update the PANDA infrastructure with network “awareness” and the ability to interact with network services to prioritize and manage its workload.

## Comparison with HEP Needs

Previous studies of HEP needs, for example the TAN Report (<http://gate.hep.anl.gov/lprice/TAN/Report/TAN-report-final.doc>) have focused on communications between developed regions such as Europe and North America.  In such reports packet loss less than 1%, vital for unimpeded interactive log-in, is assumed and attention is focused on bandwidth needs and the impact of low, but non-zero, packet loss on the ability to exploit high-bandwidth links.  The PingER results show clearly that much of the world suffers packet loss impeding even very basic participation in HEP experiments and points to the need for urgent action.

The PingER throughput predictions based on the Mathis formula assume that throughput is mainly limited by packet loss.  The 20% per year growth curve in Figure 9 is somewhat lower than the 79% per year growth in future needs that can be inferred from the tables in the TAN Report. True throughput measurements have not been in place for long enough to measure a growth trend.  Nevertheless, the throughput measurements, and the trends in predicted throughput, indicate that current attention to HEP needs between developed regions could result in needs being met.  In contrast, the measurements indicate that the throughput to less developed regions is likely to continue to be well below that needed for full participation in future experiments.

# Recommendations

There is interest from ICFA, ICTP, IHY and others to extend the monitoring further to countries with no formal HEP programs, but where there are needs to understand the Internet connectivity performance in order to aid the development of science. Africa is a region with many such countries. The idea is to provide performance within developing regions, between developing regions and between developing regions and developed regions.

We should strive for >=2 remote sites monitored in each major Developing Country. All results should continue to be made available publicly via the web, and publicized to the HEP community and others. Typically HEP leads other sciences in its needs and developing an understanding and solutions. The outreach from HEP to other sciences is to be encouraged. The results should continue to be publicized widely.

We need assistance from ICFA and others to find sites to monitor and contacts in the developing and the rest of the world, especially where we have <= 1 site/country. A current list of countries with active nodes can be found at <http://www-iepm.slac.stanford.edu/pinger/sites-per-country.html>.

# Future Support

Although not a recommendation, it would be disingenuous to finish without noting the funding support challenges. SLAC, and SEECS-NUST are the leaders in the PingER project. The funding for the PingER effort came from the DoE MICS office since 1997, however it terminated at the end of the September 2003, since it was being funded as research and the development is no longer regarded as a research project.

From 2004 onwards, development was continued with funding from the Pakistani Higher Education Commission (HEC) and the US State Department. Further funding for this research collaboration -- between SLAC and NUST -- was acquired from HEC for four years i.e. 2008-2011. This funding is primarily used for human resource development i.e. providing opportunities for graduating students to work at SLAC and participate in the research and development activities. The development consists of extending and enhancing the project, fixing known non-critical bugs, improving visualization, automating reports generated by hand today, finding new country site contacts, adding route histories and visualization, automate alarms, updating the web site for better navigation, adding more Developing Country monitoring sites/countries, and improve code portability.

The daily management, operation and supervising/leading the development was continued with discretionary funding from the SLAC and FNAL HEP programs. The management and operation includes maintaining data collection and archiving, explaining needs, identifying and reopening broken connections, identifying and opening firewall blocks, finding replacement hosts, making limited special analysis and case studies, preparing and making presentations, responding to questions. The equipment performing this is currently in place at SLAC and NUST.

Management, operation and supervision require central funding at a level of about 15% of a Full Time Equivalent (FTE) person, plus travel. This has been provided by discretionary funding from the HEP budgets of SLAC. The 2008 cuts in the US science budget, the impact on the Department of Energy, HEP and SLAC have meant that SLAC no longer has discretionary funding and thus no longer officially supports the Digital Divide activities.

In addition the funding from HEC ran out in 2012, so there have been no NUST students at SLAC since April 2012. In addition the funding for travel to Nigeria and Malaysia has been partially funded out of the SLAC PI’s own pocket and the time has been taken out of his vacation time. Without funding, for the operational side, the future of PingER and reports such as this one is unclear, and the level of effort sustained in previous years will not be possible.

Since April 2012, the PingER program with Pakistan continued at a reduced level with fortnightly Skype meetings. There was a visit to SLAC in September, by the Director General of SEECS (Dr. Arshad Ali) and the Rector of NUST. As a result of this we hope for limited funding from NUST to support a NUST student at SLAC for a year starting in 2013.

In 2012, we also completed a Memorandum of Understanding (MoU) between SLAC and the University of Malaysia in Sarawak (UNIMAS). As part of this we are working with UNIMAS and NUST to submit a proposal for limited funding to extend the PingER network in Malaysia.

Many agencies/organizations have expressed interest (e.g DoE, ESnet, NSF, ICFA, ICTP, IDRC, UNESCO, IHY) in this work but none have so far stepped up to funding the management and operation.

# Acknowledgements

We gratefully acknowledge the following: the assistance from NUST SEECS in improving the PingER toolkit and management has been critical to keeping the project running, with respect to this we particularly acknowledge the support of their leader Arshad Ali; and the students and lecturers who have assisted including recently: Umar Kalim, Anjum Navid, Raja Asad Khan of NUST SEECS who helped in updating some of the graphs, the case studies on Africa and Pakistan and implementation of PingER tools such as TULIP. Mike Jensen provided much useful information on the status of networking in Africa[[55]](#footnote-55). Alberto Santoro of UERJ provided very useful information on Latin America. Sergio Novaes of UNESP and Julio Ibarra of Florida International University provided useful contacts in Latin America. We received much encouragement from Marco Zennaro and Enrique Canessa of ICTP and from the ICFA/SCIC in particular from Harvey Newman the chairman. We must also not forget the help and support from the administrators of the PingER monitoring sites worldwide.

# Appendices

## Appendix A: PingER Presentations etc. in 2013.

##### Publications etc.

* Africa 'leapfrogs' to wider Internet access[[56]](#footnote-56). Al Jazeera, Sept 2, 2013. Also appeared on SlashDot Front Page[[57]](#footnote-57), Sept 3, 2013. Referenced in SLAC Today as SLAC expert weighs in on Internet access in Africa[[58]](#footnote-58), Sept 6, 2013
* PingER project data on the health of Africa’s Internet networks[[59]](#footnote-59), oAfrica, April 15, 2013.
* Why continuous investment and improvement in African ICT infrastructure is important[[60]](#footnote-60), contributed by Les Cottrell to EGy Africa News, Number4, February, 2013.
* [Pinging Africa](http://spectrum.ieee.org/telecom/internet/how-bad-is-africas-internet) , R. Les Cottrell, IEEE Spectrum February 2013. Also see [A Simple Tool for Measuring Digital Development](http://spectrum.ieee.org/telecom/internet/how-bad-is-africas-internet/a-simple-tool-for-measuring-digital-development) , by R. Les Cottrell, IEEE Spectrum February 2013. This is derived from SLAC-PUB-15333.

##### Talks

* PingER: Research Opportunities and Trends[[61]](#footnote-61), by Les Cottrell, PingER workshop, Universiti Malaya, Kuala Lumpur, Malaysia, June 26, 2013
* PingER: Navigating the PingER Web Site and Mining the Data[[62]](#footnote-62), by Les Cottrell, PingER workshop , Universiti Malaya, Kuala Lumpur, Malaysia, June 24th, 2013
* PingER: Measuring the Worldwide Internet’s end-to-end performance[[63]](#footnote-63), by Les Cottrell, PingER workshop , Universiti Malaya, Kuala Lumpur, Malaysia, June 24th, 2013
* PingER: Research Opportunities and Trends[[64]](#footnote-64), by Les Cottrell, PingER workshop , Universiti Malaya, Kuala Lumpur, Malaysia, June 24th, 2013
* High Energy Physics Computing and Network Needs for Worldwide Collaborations[[65]](#footnote-65), by Les Cottrell, PingER workshop , Universiti Malaya, Kuala Lumpur, Malaysia, June 24th, 2013
* The Emergence of the Internet and Africa[[66]](#footnote-66), by Les Cottrell, Monday SLAC-wide colloquium, May 6, 2013

## Appendix B: New tools added in 2013

### TULIP

The **T**rilateration **U**tility for **L**ocating **IP** hosts ([TULIP](http://tulip.slac.stanford.edu)) was completely recreated as a single CGI script replacing the previous JAVA deployment. Adaptive IP Geolocation (**AIG**), a new measurement based technique proposed by a NUST Masters student (Raja Asad) was used in addition to Constraint Based Geolocation (CBG)[[67]](#footnote-67) and Speed of Internet Geolocation (SOI). Furthermore, results from database tools including GeoIPTool, Networld, GeoPlugin, DNS Location records and the Autonmous System of the target IP address were also incorporated.

Over 80 additional PerfSONAR landmarks were added to the TULIP database to improve the accuracy of Geolocation. The following changes were made to reduce the geolocation time from over 2 minutes to about 40 seconds:

* Landmark Laundering was fixed, a script (tulip-tuning2.pl) runs automatically from trscrontab twice every night, first to disable landmarks that have less than 20% success rate and then to enable previously disabled landmarks that now have greater than 35% success. This greatly reduces the number of non-responding landmarks called by TULIP which have to otherwise timeout.
* A new script (tulip-dup.pl) was added to trscrontab which disables landmarks that are within 1km of an enabled landmark, since collocated landmarks provide similar information and using more than one merely adds delay.
* The reflector.cgi script responsible for calling landmarks was improved. There were bugs in the code that limited the number of parallel landmark calls to 15, this was fixed and now the reflector calls 80 landmarks in parallel. The landmark timeout was 5 seconds which is less that the response time of certain landmarks and hence was increased to 10 seconds.
* Landmark tiering was properly implemented. 10 landmarks were selected from different regions for the purpose of region identification. These landmarks are called first to find the region of the target host, after the region has been identified only the landmarks of that region are used for geolocation.

To further improve the response time and reduce network traffic, caching of landmark responses was introduced. The RTTs from the landmarks to the target hosts are cached and used for up to 15 days after which they are automatically updated. The enables hosts with a valid cache to be geolocated in just a couple of seconds.

There were a lot of visual changes as well; some of these are listed below:

* The intersection regions of AIG and CBG are displayed using google maps.
* The circular distance constraints for the 3 closest landmarks are shown.
* 10 of nearest landmarks to the target host are displayed and clicking on their markers shows important information such as RTT and estimated distance from the host.

### PingER Linked Open Data

Since 1998 PingER stores network measurement data, hourly and daily, measuring more than 10 different metrics between more than 8000 pairs of nodes in more than 160 countries. Thus, there is a huge amount of data to be dealt with. The data is stored in millions of flat CSV files, which are organized using meaningful file names, and the easiest way to get access to it is using the Pingtable[[68]](#footnote-68) application.

However, despite this files structure organization that makes it possible to access the right data, it is far from a standard database system which has well known features and benefits[[69]](#footnote-69). These features would significantly improve the retrieval and management of the data hence making it much easier to build more complex structured queries in order to retrieve a very specific data to support more informative graphs, reports, and dashboards.

Further, to interoperate and interchange PingER data with other existing data sources is not a very simple task. Since the data could be highly useful and applied to many different situations (economical, geographical, seasonal events, etc.), it would be convenient to make the data easily interoperable to any other kind of data source. If PingER data were available in a standard and widely used format, there could be more joint exploration of the many possibilities that the data could offer and combine it with other kinds of sources, bringing more diversity to its usage and analysis. Besides, utilizing a standard common and open format, more people would consume the data, enabling a possible specific use that has not been anticipated.

Therefore, the target problem of PingER Linked Open Data (LOD) project can be stated as: PingER data is not stored in a conveniently accessible way, hence the ease of production of reports, smart visualizations, and dashboards could be significantly improved, especially when the data involves different data sources. Moreover, the data could be published in an open standard format to enable wider consumption. The project draws on the Semantic Web and LOD strategies and techniques to publish the data according to community and World Wide Web Consortium recommendations.

Once defined the problem, the solution approach begins with selecting part of the huge amount of PingER data. Due to the short period of time the project had to start, only a specific section of the data was considered to be converted and stored in RDF format:

* Regarding the network nodes scope: measurement between all pairs of nodes considered by PingER;
* Regarding the geographic level of detail: PingER stores not only node to node data, but also site to site, country to country, region to region, etc. PingER LOD considers only node to node data, but the geographic hierarchy of the nodes is very well defined so one could easily aggregate the data by country, continent, etc.;
* Regarding the time level of detail: the smallest time grain considered by PingER is *hour*, but the smallest grain considered by PingER LOD is *day*. It was analyzed that to process the entire data, which includes the 24 hours of all days, of all months, of all years, since 1998, would take an impractical amount of time. Additionally, the last 60 days data is continuously being inserted into the PingER LOD triples store, since August 2013.
* Regarding the types of network metrics: Pingtable contemplates 16 metrics, but PingER LOD considers only the 11 more important ones: Mean Opinion Score, Directivity, Average Round Trip Time, Conditional Loss Probability, Duplicate Packets, Inter Packet Delay Variation, Minimum Round Trip Delay, Packet Loss, TCP Throughput, Unreachability, and Zero Packet Loss Frequency.
* Regarding the network packet size: Pingtable contemplates the packet sizes 100 and 1000 bytes. It was analyzed that considering both sizes would take approximately double of the time and only one of them would be enough to satisfy the initial goals of the project. Thus, PingER LOD contemplates only packet sizes of 100 bytes.

After selecting the section of the data that will be published in RDF, a simple conceptual model was designed based on the PingER data characteristics: each measurement is basically defined by a ping sent from a source (or monitor) node to a destination (or monitored) node, sent in a determined time, and related to a specific network metric. Additionally, the data is accumulated in a historical base within the years and can be used to support to decision making. Data with these characteristics can be modeled using a well-known and studied data model: star schem[[70]](#footnote-70)a which the designed conceptual was based on.

This model was used to support the construction of the OWL ontology for the PingER conceptual domain. The PingER LOD ontology is useful for organizing, structuring, and formalizing (part of) the PingER knowledge so it can easily be retrieved and processed by a machine. Most of the ideas behind the PingER LOD ontology were reused based on an existing network measurement domain ontology proposed by the project “Monitoring and Measurement in the Next Generation Technologies”[[71]](#footnote-71) (MOMENT), which produced a Measurement Ontology for IP (MOI) traffic, that is an European Telecommunications Standards Institute (ETSI) Group Specification[[72]](#footnote-72). In addition to the network measurement part, the Geonames ontology[[73]](#footnote-73) covered the geographic concepts used by PingER and the W3C Time Ontology[[74]](#footnote-74) covered the time concepts. The entire OWL PingER LOD ontology in and a further documentation are available[[75]](#footnote-75).

Having the ontology, the next phase in the process of publishing LOD is the triplification. The triplification consists of Extracting PingER raw data, Transforming it into RDF triples format, and Loading onto the RDF database; this process is known as Extract, Transform and Load (ETL). Currently, the RDF database management system used on the project is Open RDF Sesame Native 2.7.2[[76]](#footnote-76), but more complex queries run very slowly given the big number of triples stored for measurements (more than 50 million). However, experiments are being performed using other alternatives such as OWLIM[[77]](#footnote-77) and Open Link Virtuoso[[78]](#footnote-78) technologies, but to upload the data takes a very long time so more tests are needed to give more precise results. The ETL process for general concepts (like time, geography, and universities) is not complex and takes a considerably short time (the longest part takes around 4 hours to complete). General concepts data is extracted from DBPedia[[79]](#footnote-79), Freebase[[80]](#footnote-80), and Geonames, transformed into the right format to be adequate to the ontology, and loaded onto the RDF database. The ETL process for network measurement data is very complex since it deals with the big data part. Network measurement data is extracted using Pingtable mechanism, transformed into triples accordingly to the ontology, and loaded onto the RDF database. This is one the most complex tasks within the whole project; parallel and distributed programming techniques were widely used in order to make it possible.

Finally, after populating the RDF database with PingER data, a SPARQL Endpoint was made available[[81]](#footnote-81) so people could access the RDF data using structured SPARQL queries.

Additionally, easy interfaces were developed to help people to build SPARQL Queries and to plot graphs in order to show some of the advantages of using LOD techniques. Specifically, three cases were developed (they are available on the website[[82]](#footnote-82), Visualizations tab):

* Multiple network metrics analysis: This case utilizes PingER data only. It exemplifies how LOD aids even when not using mashups (crossing PingER data with other data sources). It highlights the advantage of having well-structured data with a schema, in a very expressive format: RDF triples. It also explores the use of complex SPARQL queries which are able to capture precisely what is being searched. A single query can retrieve network measurements using any possible combination of parameters. After running the query, a graph is plotted showing multiple network metrics simultaneously for the specified parameters. Before the project, the task of combining multiple metrics in a single data sheet to build a graph was not simple.
* Crossing network metrics with university metrics: Since most of the nodes considered by PingER are educational institutions, it was convenient to verify the relation between network quality and university quality. This case exemplifies a mashup, crossing PingER data with DBPedia data about universities. To measure university quality, some metrics were retrieved: number of students, number of undergrad and grad students, faculty size, and endowment. After running the mashed query, a map is drawn plotting circles on the universities PingER monitors. Size of the circles represents the value of the university metric (e.g., the bigger the circle, the greater the number of students), the filling color of the circles represents the value of the network metric (e.g., the whiter the color, the greater the value of throughput from a PingER monitor to that university), and the stroke color represents the type of universities (e.g., gold color represents private universities). Thus, using this graph one could visually verify that better universities have better network connection.
* Crossing network metrics with percentage of GDP that countries invest in research and development: this exemplifies another mashup with PingER data with another RDF data source, which is, in this case, World Bank Data. PingER data gives network measurements to many countries on the globe and World Bank gives many different interesting indicators[[83]](#footnote-83), including countries’ Research and Development (R&D) expenditure (% of GDP). Similarly, a map is drawn and circles are plotted on countries. The bigger the circle, the more it is invested in R&D by the country and the whiter the color, the greater the value of the network metric is. Moreover, an evolution within the years (since 1998) of the countries’ investment in R&D as well as their network quality can be easily visualized on the map.

In addition to those listed cases, since the data is well-structured in a standard and interoperable format, anyone can now use PingER LOD to retrieve data and make more interesting analysis, especially combining network measurement data with data of any other different characteristic.

For the future, the most urgent need is to determine which RDF triplestore to use. As stated previously, experiments are being currently made using Open Link Virtuoso and OWLIM repositories. Furthermore, other points can be listed:

* The updating system is not very efficient: If, for example, a specific node changes its full name, the system will store both old and new names, with no distinction between the values. Another worse case happens if, for some reason, two different measurement values are captured for the exact same query parameters: both of them will be stored with no distinction. This needs to be understood and fixed.
* Many universities are not being captured using the current system, even though they are monitored by PingER. More research is needed to provide better results.
* More documentation needs to be provided.
* More tests are needed to verify the consistency of the data.

A better analysis of the ontology occurred recently and changes have happened to provide better semantic and organization to the terms. The data should be re-uploaded accordingly to the new ontology.

## Appendix C: PingER Project History in 2013

There is a complete history of the PingER project this millennium[[84]](#footnote-84). Here we only report on 2013

We extended the collaboration with the University of Malaysia in Sarawak (UNIMAS) to add the University of Malaya (UM) in Kuala Lumpur and Universiti Tehcnologi Malaysia (UTM) in Johor Baru.  In June 2013 we held a two day workshop on PingER[[85]](#footnote-85) at UM. We also visited UTM and gave a talk there. We held fortnightly Skype meetings with NUST, UM, UTM and UNIMAS.

We led a 2 day workshop on PingER at UM in Kuala Lumpur. While in Malaysia we also visited UTM, talked to students and gave a presentation.

We more than doubled the number of hosts monitored in Malaysia and S. E. Asia, and added extra monitoring hosts in Pakistan and Malaysia/

Raja Asad from NUST arrived at SLAC in August 2013 to spend a year as a visiting scientist working on PingER. His main role was to develop the TULIP geolocation of Internet hosts based on distance estimates from ping response times.

Submitted paper on dynamic ping delay based Geolocation to ACM/IEEE, Started development of geolocation Visual traceroute.

Renan Sousa from UERJ in Brazil visited SLAC for 3 months and put up a prototype Linked Open Data access to PingER data for mashups.

We gave a SLAC-wide colloquium on The Emergence of the Internet and Africa[[86]](#footnote-86).

## Appendix D: ICFA/SCIC Network Monitoring Working Group

The formation of this working group was requested at the [ICFA/SCIC meeting at CERN in March 2002[[87]](#footnote-87)](http://www.slac.stanford.edu/grp/scs/trip/cottrell-icfa-mar02.html). The mission is to: *Provide a quantitative/technical view of inter-regional network performance to enable understanding the current situation and making recommendations for improved inter-regional connectivity.*

The lead person for the monitoring working group was identified as Les Cottrell. The lead person was requested to gather a team of people to assist in preparing the report and to prepare the current ICFA report for the end of 2002. The team membership used to consist of:

Table : Members of the ICFA/SCIC Network Monitoring team

|  |  |  |  |
| --- | --- | --- | --- |
| *Les Cottrell\** | SLAC | US | [cottrell@slac.stanford.edu](mailto:cottrell@slac.stanford.edu) |
| Richard Hughes-Jones | University of Manchester | UK and DANTE | Richard Hughes-Jones (Richard.Hughes-Jones@dante.net) |
| Sergei Berezhnev | RUHEP, Moscow State.Univ. | Russia | sfb@radio-msu.net |
| Sergio F. Novaes | FNAL | S. America | novaes@fnal.gov |
| Fukuko Yuasa | KEK | Japan and E. Asia | fukuko.yuasa@kek.jp |
| Shawn McKee\* | Michigan | I2 HEP Net Mon WG, USATLAS | [smckee@umich.edu](mailto:smckee@umich.edu) |

\*Still active.

### Goals of the Working Group

* Obtain as uniform picture as possible of the present performance of the connectivity used by the ICFA community
* Prepare reports on the performance of HEP connectivity, including, where possible, the identification of any key bottlenecks or problem areas

1. Since North America officially includes Mexico, the Encyclopedia Britannica recommendation is to use the terminology Anglo America (US + Canada). However, in this document North America is taken to mean the U.S. and Canada. [↑](#footnote-ref-1)
2. Internet World Statistics available at <http://www.internetworldstats.com/stats1.htm> [↑](#footnote-ref-2)
3. A host is considered unreachable when none of the pings sent to it there is no response to any of the pings sent to it. [↑](#footnote-ref-3)
4. MonALISA, see http:// monalisa.caltech.edu [↑](#footnote-ref-4)
5. Pathload, see <http://www.cc.gatech.edu/fac/Constantinos.Dovrolis/bw-est/pathload.html> [↑](#footnote-ref-5)
6. What is perfSONAR available at http://www.perfsonar.net/ [↑](#footnote-ref-6)
7. Iperf home page is available at http://dast.nlanr.net/Projects/Iperf/ [↑](#footnote-ref-7)
8. "The GridFTP Protocol Protocol and Software". Available http://www.globus.org/datagrid/gridftp.html and also see <http://en.wikipedia.org/wiki/GridFTP> [↑](#footnote-ref-8)
9. Linked Data, see <http://en.wikipedia.org/wiki/Linked_data> [↑](#footnote-ref-9)
10. Semantic Web, see <http://en.wikipedia.org/wiki/Semantic_Web> [↑](#footnote-ref-10)
11. Resource Description Framework, see <http://en.wikipedia.org/wiki/Resource_Description_Framework> [↑](#footnote-ref-11)
12. Web Ontology language, see <http://en.wikipedia.org/wiki/Web_Ontology_Language> [↑](#footnote-ref-12)
13. Search RDF data with SPARQL, see Resource Description Framework RDF [↑](#footnote-ref-13)
14. ICFA/SCIC Monitoring Working Group’s Annual Reports, see <http://www.slac.stanford.edu/xorg/icfa/scic-netmon/#annual> [↑](#footnote-ref-14)
15. "PingER". Available <http://www-iepm.slac.stanford.edu/pinger/>; W. Matthews and R. L. Cottrell, "The PingER Project: Active Internet Performance Monitoring for the HEP Community", IEEE Communications Magazine Vol. 38 No. 5 pp 130-136, May 2002. [↑](#footnote-ref-15)
16. R. L. Cottrell, "Tutorial on Internet Monitoring & PingER at SLAC". See <http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html> [↑](#footnote-ref-16)
17. In special cases, there is an option to reduce the network impact to ~ 10bits/s per monitor-remote host pair. [↑](#footnote-ref-17)
18. PERFormance Service Oriented Network monitoring Architecture , see <http://www.perfsonar.net/> [↑](#footnote-ref-18)
19. "Requirements for WAN Hosts being Monitored", Les Cottrell and Tom Glanzman. Available at  <http://www.slac.stanford.edu/comp/net/wan-req.html> [↑](#footnote-ref-19)
20. "United Nations Population Division World Population Prospects Population database". Available <http://esa.un.org/unpp/definition.html> [↑](#footnote-ref-20)
21. The difference in the number of active sites in Figure 3 and the ~780 reported in the Deployment section is the difference in the number of sites with active hosts in the last 30 days (lower number) and the number active at some time in the last year, plus Figure 3 is for nodes monitored from SLAC while 780 is the number of different hosts monitored by all PingER monitoring sites. [↑](#footnote-ref-21)
22. R. L. Cottrell, "Tutorial on Internet Monitoring & PingER at SLAC". Available http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html [↑](#footnote-ref-22)
23. M. Mathis, J. Semke, J. Mahdavi, T. Ott, ["The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm"](http://www.psc.edu/networking/papers/model_ccr97.ps),*Computer Communication Review*, volume 27, number 3, pp. 67-82, July 1997 [↑](#footnote-ref-23)
24. Throughput vs. Packet Loss, R. Les Cottrell, see http://www.slac.stanford.edu/comp/net/wan-mon/thru-vs-loss.html [↑](#footnote-ref-24)
25. Russia 2013, see <https://confluence.slac.stanford.edu/display/IEPM/Russia+2013> [↑](#footnote-ref-25)
26. Geosynchronous satellite, see <http://en.wikipedia.org/wiki/Geosynchronous_satellite> [↑](#footnote-ref-26)
27. Democratic Republic of Congo - Telecoms, Mobile and Broadband, see <http://www.budde.com.au/Research/Democratic-Republic-of-Congo-Telecoms-Mobile-and-Broadband.html#sthash.baG5mnoD.dpuf> [↑](#footnote-ref-27)
28. Liquid Telecom lays fibre to the DRC and connects DRC mobile operator to Southerne Africa rRegional Fibre Nteowrk and to WACS. See <http://www.liquidtelecom.com/blog/liquid-telecom-lays-fibre-to-the-drc-and-connects-drc-mobile-operator-to-southern-african-regional-fibre-network-and-to-wacs> [↑](#footnote-ref-28)
29. Republic of Congo Ready to Surf the High Speed Internet, see <http://www.afriqueavenir.org/en/2011/07/28/republic-of-the-congo-ready-to-surf-with-high-speed-internet/> [↑](#footnote-ref-29)
30. ACE Cable System, see  [http://en.wikipedia.org/wiki/ACE\_(cable\_system) /](http://manypossibilities.net/african-undersea-cables/) [↑](#footnote-ref-30)
31. World Bank Supports Submarine Communications Cable and Helps Unlock High-Speed Opportunities, see <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTINFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/EXTEDEVELOPMENT/0,,contentMDK:23017590~menuPK:559467~pagePK:64020865~piPK:149114~theSitePK:559460,00.html> [↑](#footnote-ref-31)
32. 26A Case Study on Cuba, see <https://confluence.slac.stanford.edu/display/IEPM/Cuba+-+Switch+from+Satellite+to+terrestrial+link>

    B Tutorial on Internet Monitoring and PingER at SLAC, see <http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html#variable> [↑](#footnote-ref-32)
33. A notable effect on derived throughput is that for measurements made from say a N. American site, such as SLAC, to other N. American sites, 1/RTT is large and so throughputs are artificially enhanced compared to measurements to more distant regions with longer RTTs. [↑](#footnote-ref-33)
34. “[Modeling TCP throughput: A simple model and its empirical validation](http://conferences.sigcomm.org/sigcomm/1998/tp/paper25.pdf)” by J. Padhye, V. Firoiu, D. Townsley and J. Kurose, in *Proc. SIGCOMM Symp. Communications Architectures and Protocols* Aug. 1998, pp. 304-314. [↑](#footnote-ref-34)
35. African Undersea Cables, see <http://manypossibilities.net/african-undersea-cables/> [↑](#footnote-ref-35)
36. Interquartile Range, see <http://en.wikipedia.org/wiki/Interquartile_range> [↑](#footnote-ref-36)
37. Calculating the MOS, see <http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html#mos> [↑](#footnote-ref-37)
38. The PingER team is successfully holding weekly meetings using VoIP calls using Skype between SLAC in California and NUST in Islamabad, Pakistan. The quality is pretty good. The main problem is loss of connectivity due to power outages in Pakistan. [↑](#footnote-ref-38)
39. Measuring the Information Society, from the ITU , see <http://www.itu.int/ITU-D/ict/publications/idi/2011/> [↑](#footnote-ref-39)
40. Table 4 “ICFA SCIC Network Monitoring Report, 2012”. see <http://www.slac.stanford.edu/xorg/icfa/icfa-net-paper-jan12/report-jan12.docx> [↑](#footnote-ref-40)
41. The IDI access sub-index measures ICT infrastructure and readiness. It includes: fixed-telephone subscriptions per 100 inhabitants, mobile-cellular telephone subscriptions per 100 inhabitants, international Internet bandwidth per Internet user, percentage of households with a computer and percentage of households with Internet access at home [↑](#footnote-ref-41)
42. “End-to-end Performance Initiatives Performances Environment System”, Internet2, see [http://e2epi.internet2.edu/e2epipes//index.html](http://e2epi.internet2.edu/e2epipes/index.html) [↑](#footnote-ref-42)
43. “Version 2.0 of Advisor Released”, NCSA, see <http://access.ncsa.illinois.edu/Releases/05Releases/03.08.05_Version_2..html> [↑](#footnote-ref-43)
44. “MonALISA: LISA”, Caltech, see <http://monalisa.cern.ch/monalisa__Interactive_Clients__LISA.html> [↑](#footnote-ref-44)
45. “OGF Network Measurement Working Group”, Internet2, see <http://nmwg.internet2.edu/> [↑](#footnote-ref-45)
46. The main webpage is <http://lhcone.net> [↑](#footnote-ref-46)
47. “Global Lambda Integrated Facility”, see <http://www.glif.is/> [↑](#footnote-ref-47)
48. “DANTE-Internet2-CANARIE-ESnet collaboration, see <http://www.geant2.net/server/show/conWebDoc.1308> [↑](#footnote-ref-48)
49. “An Ultrascale Information System for Data Intensive Research”, see <http://www.ultralight.org/web-site/ultralight/html/index.html> [↑](#footnote-ref-49)
50. “Terapaths”, <https://www.racf.bnl.gov/terapaths/> [↑](#footnote-ref-50)
51. “Lambda Station”, see <http://www.lambdastation.org/> [↑](#footnote-ref-51)
52. “ESnet On-demand Secure Circuits and Advance Reservation System”, ESnet, see <http://www.es.net/oscars/index.html> [↑](#footnote-ref-52)
53. “End Site Control Plane Surface”, see <https://plone3.fnal.gov/P0/ESCPS> [↑](#footnote-ref-53)
54. “Development of Dynamic Network System”, Internet2, see <http://www.internet2.edu/ion/dynes.html> [↑](#footnote-ref-54)
55. Mike Jensen, ["Connectivity Mapping in Africa"](http://www.ictp.trieste.it/~ejds/seminars2002/Mike_Jensen/jensen-full.ppt), presentation at the ICTP Round Table on Developing Country Access to On-Line Scientific Publishing: Sustainable Alternatives at ITCP, Trieste, October 2002. Available http://www.ictp.trieste.it/~ejds/seminars2002/Mike\_Jensen/jensen-full.ppt [↑](#footnote-ref-55)
56. “Africa ‘leapfrogs’ to wider Internet Access”  Al Jazeera, Sept 2, 2013. See <http://www.aljazeera.com/indepth/features/2013/07/201372484039912339.html> [↑](#footnote-ref-56)
57. SlashDot Front page Sept 3, 2013, see <http://mobile.slashdot.org/story/13/09/03/233202/how-africa-will-leapfrog-wired-networks> [↑](#footnote-ref-57)
58. SLAC expert weighs in on Internet access to Africa, see <https://news.slac.stanford.edu/today/2013-09-05> [↑](#footnote-ref-58)
59. PingER project data on the health of Africa’s Internet networks, see <http://www.oafrica.com/broadband/pinger-project-data-on-the-health-of-africas-internet-networks/> [↑](#footnote-ref-59)
60. Why continuous investment and improvement in African ICT infrastructure is important, contributed by Les Cottrell to EGy Africa News, Number4, February, 2013. See <https://confluence.slac.stanford.edu/download/attachments/17164/eGYAfrica_Newsletter_No_4.pdf?version=1&modificationDate=1362087290000&api=v2> [↑](#footnote-ref-60)
61. PingER Research Opportunities and Trends,. See <https://confluence.slac.stanford.edu/download/attachments/17164/pinger-opportunities.pptx> [↑](#footnote-ref-61)
62. Pinger: Navigating the Web site and Mining the Data, see <https://confluence.slac.stanford.edu/download/attachments/17164/malaysia-handson.ppt> [↑](#footnote-ref-62)
63. PingER: Measuring the Worldwide Internet’s end-to-end performance, see [↑](#footnote-ref-63)
64. PingER: Navigating the web site and mining the data, see <https://confluence.slac.stanford.edu/download/attachments/17164/malaysia-handson.ppt> [↑](#footnote-ref-64)
65. High Energy Physics Computing and Network Needs for Worldwide Collaborations, by Les Cottrell, Universiti Malaya, June 26th, 2013. See <https://confluence.slac.stanford.edu/download/attachments/17164/UM%20%281%29.pptx> [↑](#footnote-ref-65)
66. The Emergence of the Internet and Africa, see <https://confluence.slac.stanford.edu/download/attachments/17164/colloquium.pptx> [↑](#footnote-ref-66)
67. “Constraint Based Geolocation of Internet Hosts”, B. Gueye, A. Ziviani, M. Crovella, S . Fdida, IEEE/ACM Trans. on Networking, 2006, see <http://www-npa.lip6.fr/_publications/682-p1219-gueye.pdf> [↑](#footnote-ref-67)
68. Pingtable, see <http://www-wanmon.slac.stanford.edu/cgi-wrap/pingtable.pl> [↑](#footnote-ref-68)
69. . Fundamentals of Database Systems. 6.ed, NAVATHE, Shamkant; ELMASRI, Ramez, Addison-Wesley, 2010. [↑](#footnote-ref-69)
70. Star schema, see <http://www.learndatamodeling.com/star.php#.UogBUsTUDIF> [↑](#footnote-ref-70)
71. MOMENT: Monitoring and Measurement in the Next Generation Technologies, see <http://www.salzburgresearch.at/en/projekt/moment_en/> [↑](#footnote-ref-71)
72. ETSI Industry Specification Groups, see <http://www.etsi.org/about/how-we-work/industry-specification-groups> [↑](#footnote-ref-72)
73. <http://www.geonames.org/ontology/documentation.html> [↑](#footnote-ref-73)
74. <http://www.w3.org/TR/owl-time/> [↑](#footnote-ref-74)
75. <https://confluence.slac.stanford.edu/display/IEPM/Ontology> [↑](#footnote-ref-75)
76. <http://openrdf.callimachus.net/sesame/2.7/docs/users.docbook?view> [↑](#footnote-ref-76)
77. Ontotext, see <http://en.wikipedia.org/wiki/Ontotext> [↑](#footnote-ref-77)
78. Virtuoso, see <http://virtuoso.openlinksw.com/> [↑](#footnote-ref-78)
79. DBpedia, see <http://en.wikipedia.org/wiki/DBpedia> [↑](#footnote-ref-79)
80. Freebase, see <http://en.wikipedia.org/wiki/Freebase> [↑](#footnote-ref-80)
81. <http://pingerlod.slac.stanford.edu/sparql> [↑](#footnote-ref-81)
82. <http://pingerlod.slac.stanford.edu> [↑](#footnote-ref-82)
83. <http://data.worldbank.org/indicator> [↑](#footnote-ref-83)
84. History of PingER this Milenium, see <https://confluence.slac.stanford.edu/display/IEPM/History+of+growth+in+PingER+hosts+this+milenium> [↑](#footnote-ref-84)
85. PingER Workshop at UM, 2013, see <http://pinger.unimas.my/pinger/workshop.php> [↑](#footnote-ref-85)
86. The Emergence of the Internet and Africa, see <https://confluence.slac.stanford.edu/download/attachments/17164/colloquium.pptx> [↑](#footnote-ref-86)
87. "ICFA/SCIC meeting at CERN in March 2002". Available <http://www.slac.stanford.edu/grp/scs/trip/cottrell-icfa-mar02.html> [↑](#footnote-ref-87)