Advanced Performance Modeling
with Combined Passive and Active Monitoring

Annual Project Report – 1st year
October 31, 2012

Project period of
October 1, 2012 through September 30, 2012

Principal Investigators
Alex Sim[1], Constantine Dovrolis[2]

Project members: Kejia Hu[1], Demetris Antoniades[2]
Project Website: http://sdm.lbl.gov/apm

Table of Contents

1 Lawrence Berkeley National Laboratory
2 Georgia Institute of Technology

1 Project Summary

In recent years, network technologies have evolved rapidly. Advanced science networks such as those managed by ESnet and Internet2 in the US operate at speeds of up to 100 Gbps today. Owners of these networks are aggressively researching and deploying network-level QoS in an attempt to isolate large parallel data transfers from the rest of user traffic. Despite improvements in network technologies, optimal selection of shared network resources and efficient scheduling of data transfers in a distributed and collaborative environment are challenging tasks in achieving superior performance in accessing data. Monitoring the state of shared network resources and estimating their future performance have been used for identifying efficient selection and scheduling of network resources. However, existing network tools based on active probing cannot directly model high-speed networks and estimate or predict optimal network performance. Existing models cannot predict the performance of newly included network resources. They also do not provide users with information about the progress of ongoing data transfers. Performance estimation models for terabit and future high-speed networks should be able to consider new variables such as network fluctuating capacity and available bandwidth and higher concurrency of data streams. Moreover, existing estimation models performing active probing put extra load on the network resources, which might interfere with network performance. To maximize the throughput of data access operations, monitoring information about network performance needs to be collected in passive mechanisms without imposing extra load on network resources. This measurement information must be synthesized to drive a predictive estimation model of end-to-end network performance.

To improve the efficiency of resource utilization and scheduling of scientific data transfers on high-speed networks, we started a project on Advanced Performance Modeling with combined passive and active monitoring (APM) that investigates and models a general-purpose, reusable and expandable network performance estimation framework. The predictive estimation model and the framework will be helpful in optimizing the performance and utilization of fast networks as well as sharing resources with predictable performance for scientific collaborations, especially in data intensive applications. Our prediction model utilizes a combination of passive historical network performance information from various network activity logs as well as active measurements from network monitoring devices. The prediction model estimates future network usage and the latency in using the network. Historical network performance information is used for throughput prediction without putting extra load on the resources by active measurement collection. For a simple analogy, highway vehicle traffic pattern analysis (Figure 1) would give drivers time estimation for travel planning (e.g. it takes roughly about 1.5 hours from Berkeley to San Francisco Airport on Monday morning 8:30am, or about 40 minutes around 1pm). Performance measurements collected by active probing is used judiciously for improving the accuracy of predictions. The planned hybrid estimation model with both passive and active measurements will improve the accuracy in performance estimation for newly added data service nodes on high throughput networks.

To implement this project, we need to address the following challenges:

- End-to-end passive performance monitoring and measurement collection. To optimize the accessibility of datasets, monitoring information about the current data transfers needs to be collected using passive mechanisms that do not put extra load on resources. We propose to study the collection of passive monitoring information and access to the measurement repository.
- Performance estimation model. Study network traffic patterns from the measurement data to come up with predictive performance models to improve network utilization and performance and enable predictable data movement on high bandwidth networks.
- Long-term performance estimation. Provide predictive estimation for data throughput for a future time window, together with some probabilistic variability estimate.

This research explores fundamental questions on the relationship between monitoring and estimation of network resource performance.

This document gives a brief overview on the technical progress in the APM project from Lawrence Berkeley National Laboratory and Georgia Institute of Technology, for the project period from October 1, 2012 to September 30, 2012.

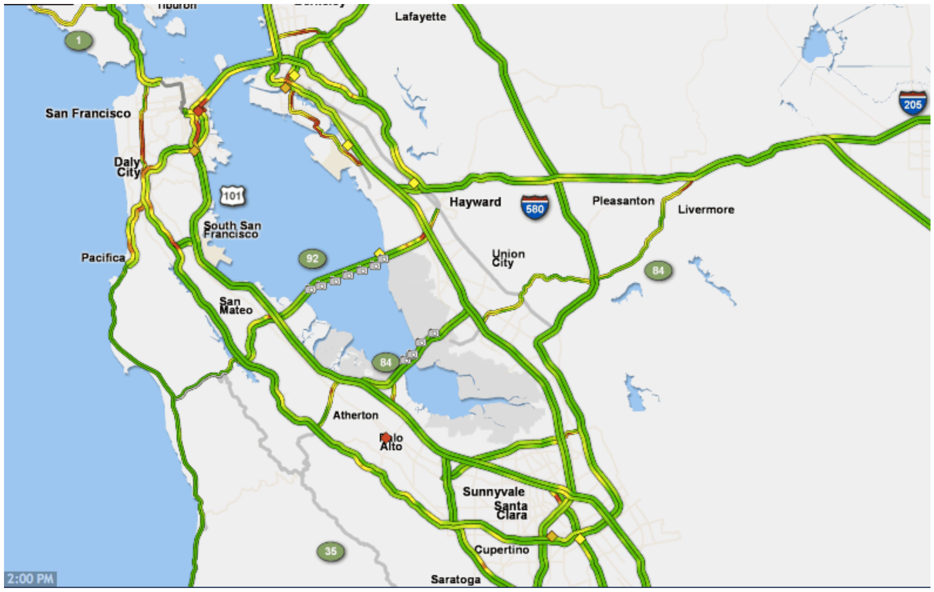


Figure 1: This project's network traffic analysis is analogous to highway traffic estimation and prediction.

2 Project Accomplishments

2.1 Activities to get access to the network measurement log data from ESnet

We have identified Simple Network Management Protocol (SNMP) and NetFlow logs as the useful historical network measurement data that ESnet collects on the managed network routers.

2.1.1 ESnet NetFlow data

NetFlow data provides detailed information for end-to-end performance. Using NetFlow data, we can have accurate information about an end-to-end communication, the amount of data transferred back and forth, but it comes with an additional cost for privacy concerns. The NetFlow data includes IP addresses, and raises significant user privacy concerns over Personally Identifiable Information (PII) issues that IP addresses can be traced to the individual person. We have contacted Institutional Review Board (IRB) at University of California at Berkeley where LBNL shares IRB. IRB has reviewed the case and recommended the data confidentiality and anonymization methods over IP addresses limited to R&E sites. Data confidentiality and access agreement between the project and the ESnet has been created and signed by project members.

2.1.2 ESnet SNMP data

The SNMP data provides aggregated link usage data from ESnet components. We have access to publicly available ESnet SNMP data collected from stats.es.net. The time span of these SNMP data is from May 6 18:49:00 PDT 2011 to the present time. In our current studies, we retrieved data up to Jul 25 15:56:30 PDT 2012. We are currently working with ESnet for an access to the offline SNMP data for older than 2011.

2.1.3 Other historical measurement data

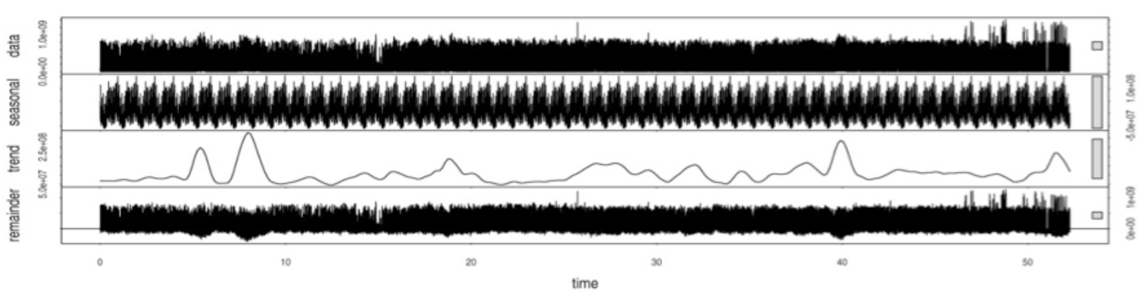
We also identified GridFTP logs as another historical network performance measurements that Globus GridFTP collects anonymously, and the GridFTP logs are particularly useful with large data transfers. We are currently collaborating with ANL Globus team for our access to the GridFTP logs.

2.2 Analysis of SNMP data to identify periodic components, trends and anomalies

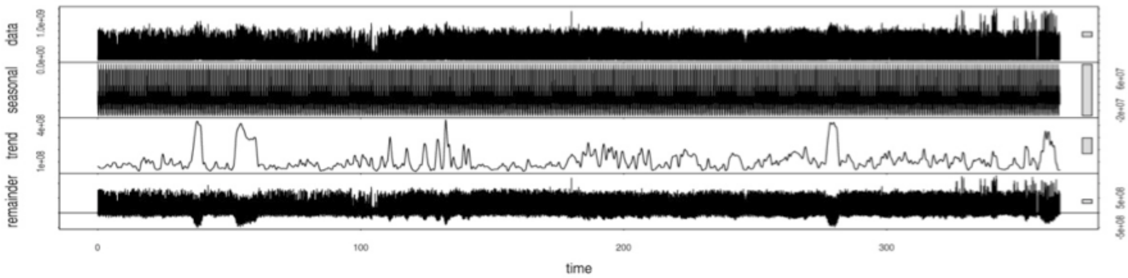
We have studied on a statistical performance model by identifying the recursive patterns (periodical and seasonal pattern) and general trend pattern based on SNMP time series data. The model contains a deterministic part of the periodical pattern and trend pattern with an uncertainty part of randomness and information regarding its parameters of distribution. Our research summary is following:

- Studied statistical network performance pattern, extracting the recursive patterns (periodical and seasonal pattern) and general trend pattern in the SNMP data.
- Studied periodic trends with the STL-based and X12-based models.
- Studied point estimate and confidence interface based on X12 model for the forecast and its errors.
- Explored similarities in the trend component patterns according to the data flow on the network paths.

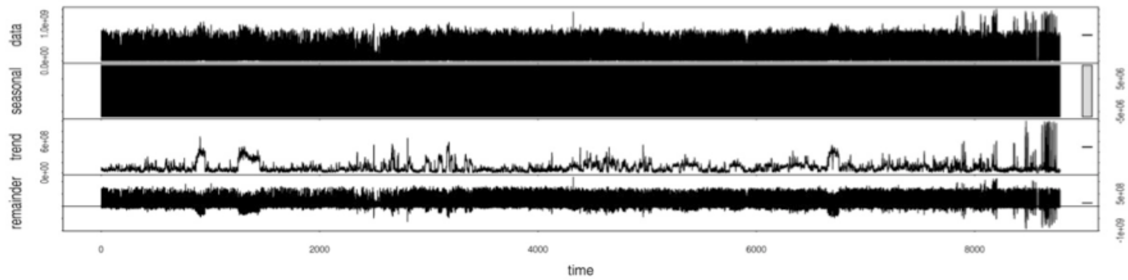
The main goal of the study is to identify the periodic trends, extracting variations associated with normal and repeating events (trend and seasonal patterns). The seasonal adjustment based on STL model (Seasonal-Trend Decomposition Procedure Based on Loess) was conducted based on four time series cycles, as shown in Figure 2; minutely cycle (every 2 measurements forms a period), hourly cycle (every 120 records in a period), daily cycle and weekly cycle, where each plot shows over time of the original SNMP data, seasonal component, trend component and remainder component. We have tried a few different window lengths, and further explored what would be the best periodic series to facilitate the extraction of seasonal and trend components. We also studied point estimate and confidence interface for the forecast and its errors, based on X12 model which has an extensive time series modeling and model selection capabilities for linear regression models with ARIMA errors. We used the point estimates to compare the difference between forecast values and real values for the last frequency in the SNMP data, and an example plot of the diagnostic statistics and forecast errors for the hourly frequency are shown in Figure 3. Q statistics shows most data points are between 0 and 1 in Figure 3-a. Figure 3-b shows the prediction error values between the true SNMP values and the model-based prediction values in the 1 hour, 2 hours, 3 hours and 4 hours in the future time window. The forecast errors are less than 20% in most cases, and in many times, the forecast errors are around 10%. The extreme events may cause a spike in the forecast error plots, but the forecast error is still within a good acceptance range in general.



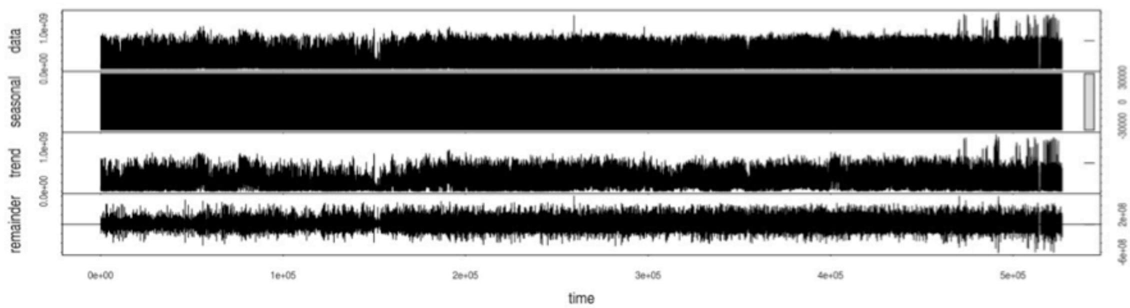
(a)



(b)

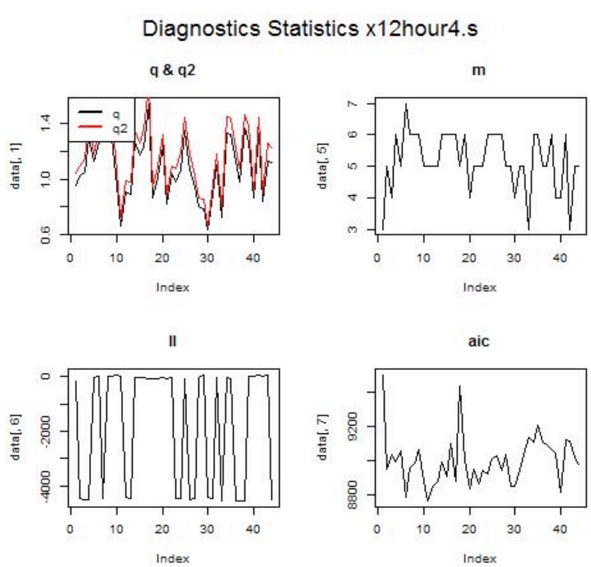


(c)

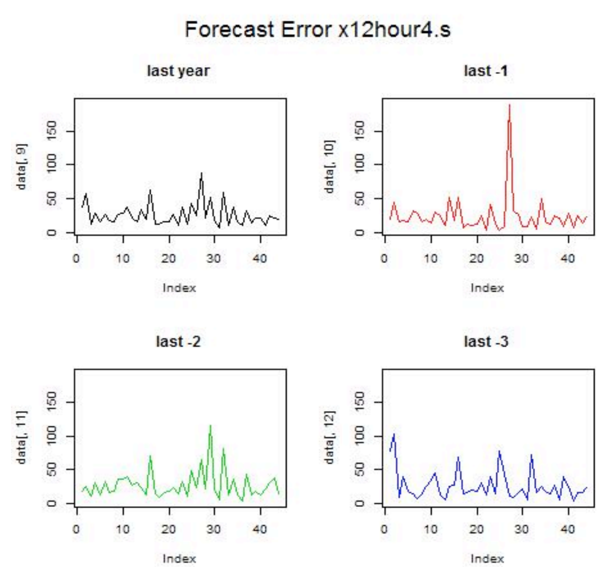


(d)

Figure 2: Example of seasonal adjustments based on STL showing original SNMP data and decomposed components of (a) weekly, (b) daily, (c) hourly, and (d) minutely data for NERSC edge router.



(a)



(b)

Figure 3: Periodicity evaluation (a) and prediction errors (b) from the hourly analysis

2.3 Overall architecture of APM system

During the last year we explored the available data from ESnet to identify how we can use them in our performance estimation model. This procedure provided the basis for the design of the APM system architecture (Figure 4). The system is composed of three major modules, the performance estimation module, the data collection module and the data analysis module.

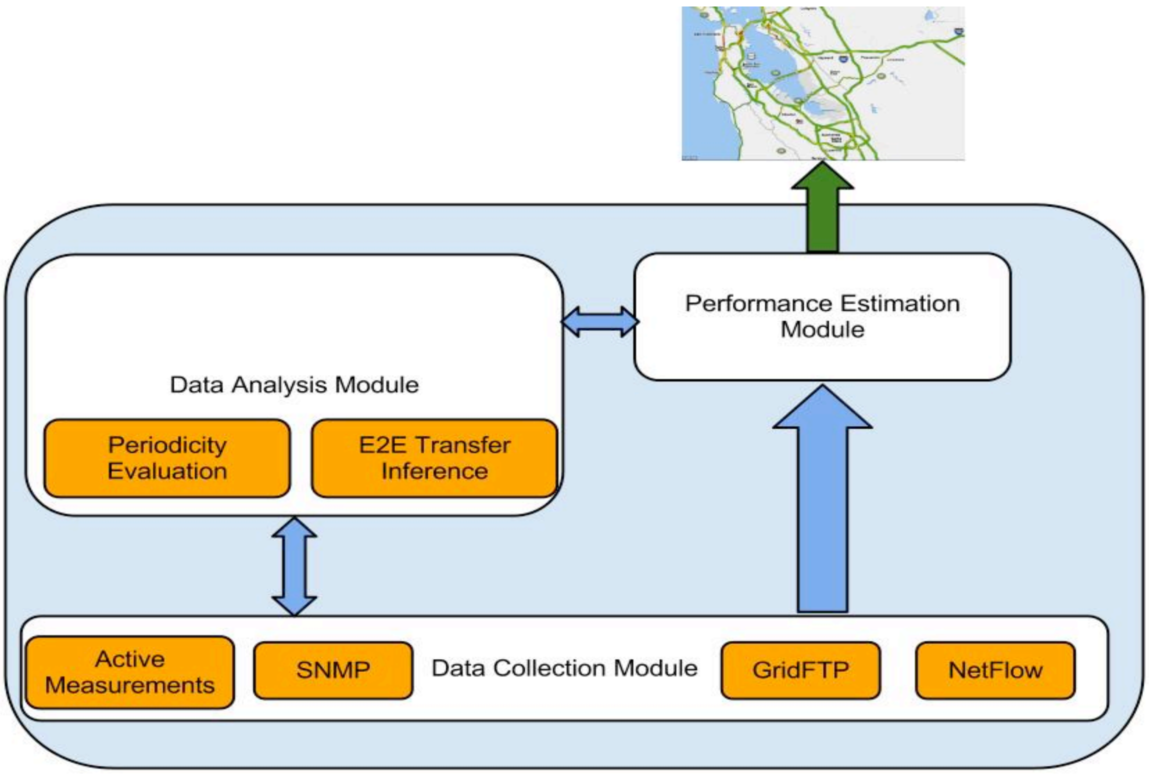


Figure 4: High level architecture of the APM system

The data collection module is responsible for fetching all newly available data from ESnet's data repositories and retrieving historical data when that is required. The collected data include SNMP traffic utilization measurements; NetFlow traces and GridFTP log files. The data collection module includes also an active measurement component that will perform measurements whenever that is required from the performance estimation module. Possible active measurements include the use of the *bwctl* tool to get a throughput measurement and the use of *traceroute* tool to get information about the path a transfer will follow. Both tools are available from the perfSONAR tool repository.

The data analysis module is responsible for parsing and analyzing the data provided by the data collection. The module will use the methods presented in the previous section for removing the periodicity from the SNMP data series. Furthermore, it will use our edge-to-edge inference method (presented in the next section) to identify additional data transfers from the SNMP data. We can then use the inferred transfers as historical transfers from the performance estimation module, complementary to the NetFlow and GridFTP data.

The performance estimation module will use the data collected and analyzed from the other modules to response to estimation and prediction requests for the performance of a specific link/path or the overall network (for creating traffic maps). The performance estimation module will use the historical throughput measurements provided from the data collection module (NetFlow, GridFTP, active monitoring) and the analysis module (E2E transfer inference) together with the predictability analysis of the link/path of interest.

Our system will make careful use of the available resources. The predictability analysis for the links will also be used as to decide the amount of historical data needed for each links. This way we will be able to discard data for highly predictable links, allowing for the storage of extra data for unpredictable links. This will allow our performance estimation module to give highly accurate predictions for all links. Our system will also try to minimize the amount of data added to the network through active measurements. To do so, we are designing our prediction algorithm in a way that it will be able to define the number of historical throughput measurements needed for a time window in order to provide high accuracy with high confidence. In this way we will issue active measurements only when the already available passive measurements do not satisfy this number.

2.4 Use of SNMP data to identify large data transfers

The network wide availability of SNMP data and the limited throughput performance samples from NetFlow data motivated us to explore different approaches for increasing our sample data set: A statistically significant set of Edge-to-Edge (E2E) throughput samples to perform TCP throughput prediction in a monitored network based on historical measurements. During the last 8 months we designed a method that is able to infer E2E transfer information using SNMP link counts. Our method is the result of two main observations. Looking at the time series data of a link's usage, we observe events where the usage of the link increases (or decreases) to a different level, deviating from the link's normal behavior up to that point. These events could be considered as starting (or ending) points of high throughput transfers. Checking all other links of the same router, we can identify the exit point of that specific event, which leads to the next router. Following this path, we can then infer the actual route that the specific event followed, tracking down its entrance and exit points in the monitored network.

Figure 5 presents a simple example. In this diagram, R1 is the edge router for autonomous system AS1 and R2 is the edge router for autonomous system AS2. Each router has several input and output interfaces, connecting to other routers. The path R1 – R2 – R3 – R4 connects the two ASes, carrying all traffic exchanged between them. The graphs above each link show the actual traffic transferred through that link during the specific observation period, as it would be available through SNMP. Each link carries traffic for a variety of source and destination pairs. Link R1 – R2 three different events during the observation period. The first and last events are not visible after R2. The middle event continues from R2 over the link R2 – R3 and from there to R3 – R4. This event and its transfer route can be attributed to a network transfer initiated at AS1 and destined to AS2. The magnitude of the deviation on the utilization time series provides information about the throughput this transfer achieved. The time of the increase and decrease events provide information regarding the starting and ending times of the transfer.

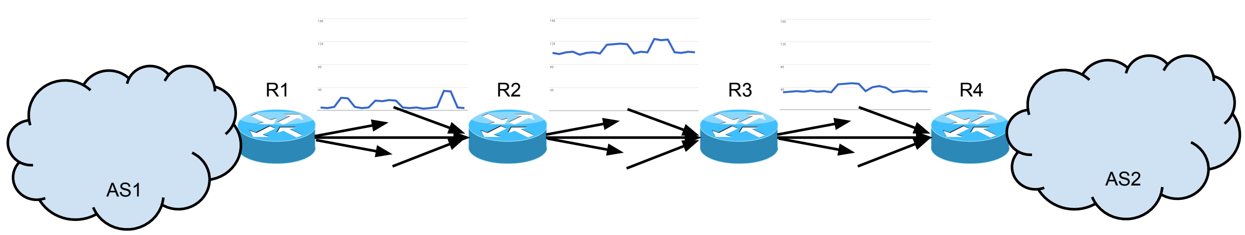


Figure 5: Simple Illustration: network events can be observed from SNMP utilization data. These events can, also, be tracked down along the network path they're traversing.

In September 2012, we submitted a paper to the PAM 2013 conference providing evidence that by using SNMP link counts, Edge-to-Edge (E2E) information about network transfers can be inferred. Currently we are working on improving our inference method and for providing an in-depth evaluation of its accuracy and validity. We plan to submit the results of this work to the Internet Measurement Conference (IMC 2013) in early May 2013.

3 Interaction with other groups and projects

- We have collaborated with ESnet for data access, specifically with Brian Tierney, Chris Tracy, Jon Dugan, Chin Guok, Inder Monga and Greg Bell.
- We have worked with IT policy and security team at LBNL for IRB issues, especially with Joy Bognaguro, Jay Krous, Aashish Sharma, and Jonathan Carter.
- We have collaborated with Raj Kettimuthu at ANL for our access to the GridFTP logs from Globus team.

4 Plans for the second year

a) Study SNMP-based prediction model of edge-to-edge throughput,
b) Combined use of NetFlow-based and SNMP-based measurements for prediction model of end-to-end throughput,
c) Collaboration with ESnet for retrieval of NetFlow log data on R&E sites,
d) Collaboration with ESnet for access to offline SNMP log data, and
e) Collaboration with ANL Globus team for access to GridFTP log data,

5 Publications, presentations and other activities

Papers and talks presented during this time period:

5.1 Presentations

1) "*Advanced Performance Modeling with Combined Passive and Active Monitoring*", A. Sim, C. Dovrolis, PI Meeting, 3/2012.

5.2 Publications

1) "*What SNMP data can tell us about Edge-to-Edge network performance*", D. Antoniades, K. Hu, A. Sim, C. Dovrolis, submitted to Passive and Active Measurements Conference (PAM2013), 2013.
2) "*Statistical prediction models for network traffic performance*", K. Hu, et al., submitted and accepted to The Winter 2013 APAN/ESnet /Internet2 technical meeting (Joint Techs, TIP2013), 2013.
3) "*Estimating the uncertainty in algorithms for outliers or missing values in network measurements*", K. Hu, et al., submitted to The Winter 2013 APAN /ESnet /Internet2 technical meeting (Joint Techs, TIP2013), 2013.