# Advanced Performance Modeling
## with Combined Passive and Active Monitoring

**Final Project Report**
***March 15, 2015***

***Project period of***
***October 1, 2011 through January 31, 2015***

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

***Project members:*** Demetris Antoniades[1,*], Jaesik Choi[2,*], Kejia Hu[2,*], Daniel Lee[1], Warren Matthews[1,*], Kanu Sahai[1,*], Kamal Shadi[1], W. William Yoo[2]

***Project Website:*** http://sdm.lbl.gov/apm

**Table of Contents**



1 Georgia Institute of Technology
2 Lawrence Berkeley National Laboratory
* Former member of the project

# 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. 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. They do not provide users with information about ongoing data transfers. Performance estimation models for high-speed networks should be able to consider new variables such as network fluctuating capacity and available bandwidth. 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 networks as well as sharing resources with predictable performance for scientific collaborations, especially in data intensive applications. Our prediction model utilizes historical network performance information from various network activity logs as well as live streaming measurements from network peering devices. Historical network performance information is used without putting extra load on the resources by active measurement collection. Performance measurements collected by active probing is used judiciously for improving the accuracy of predictions. 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).

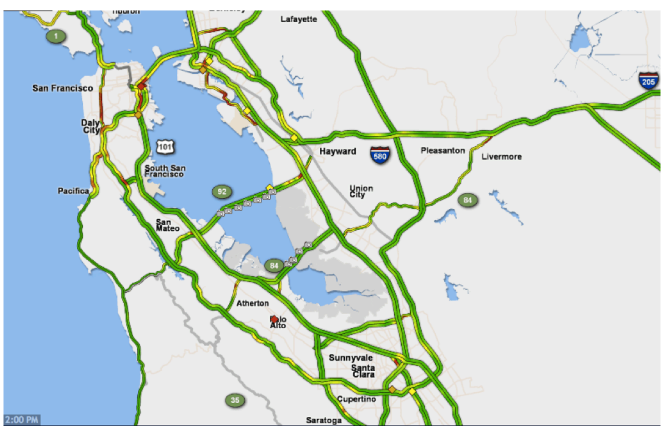


***Figure 1***: This project's network traffic analysis is analogous to highway traffic estimation and prediction.

This document gives a brief summary 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, 2011 to January 31, 2015.

## 2 Project accomplishments during the 3rd year of the project

### 2.1 Time series forecasting modeling on SNMP network measurements

We have developed a model to forecast expected network bandwidth utilization on high-bandwidth wide area networks. The forecast model can improve the efficiency of the resource utilization and scheduling of data movements on high-bandwidth networks to accommodate ever increasing data volume for large-scale scientific data applications. A univariate time-series forecast model is developed with the Seasonal decomposition of Time series by Loess (STL) and the AutoRegressive Integrated Moving Average (ARIMA) on Simple Network Management Protocol (SNMP) path utilization measurement data. Compared to the traditional approach such as the Box-Jenkins methodology to train the ARIMA model, our forecast model reduces computation time by 78.1%. It also shows more resilience against abrupt network usage changes.

Figure 1 shows the seasonal adjustment based on the STL model to identify the periodic trends, extracting variations associated with normal and repeating events (trend and seasonal patterns).

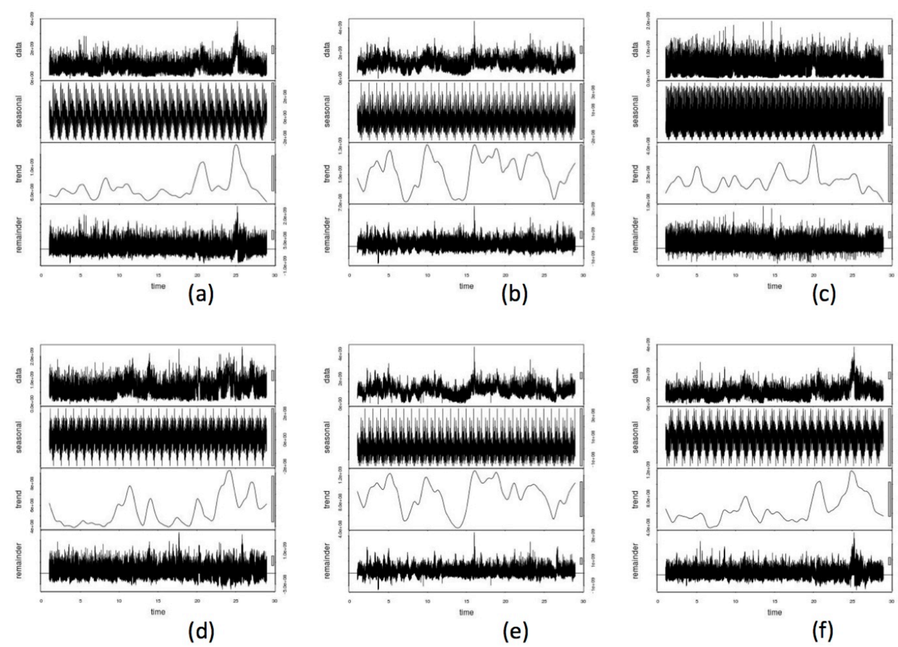


Figure 1: Seasonally Decomposed Components: The top plot in each graph is from the raw SNMP measurement data. The second plot is for the seasonal component. The third plot is for the trend component. The bottom plot is for the remainder. The horizontal axis shows the time as days, and the duration is 4 weeks from June 24, 00:00:00, GMT 2014 to July 20, 23:59:30, GMT 2014. (a) NERSC-ANL, (b) ANL-NERSC, (c) NERSC-ORNL, (d) ORNL-NERSC, (e) ANL-ORNL, (f) ORNL-ANL

Figure 2 shows the results of the forecast model for one day test set in July 21, 2014 for the NERSC-ANL network path. It shows that our red-colored forecast values are close to the blue-colored observed data. When sudden spikes in the bandwidth utilization were observed from the training sets or the test sets, our model was resilient to those sudden changes, and the forecast errors from the model are within the standard deviations of the monitored measurements.

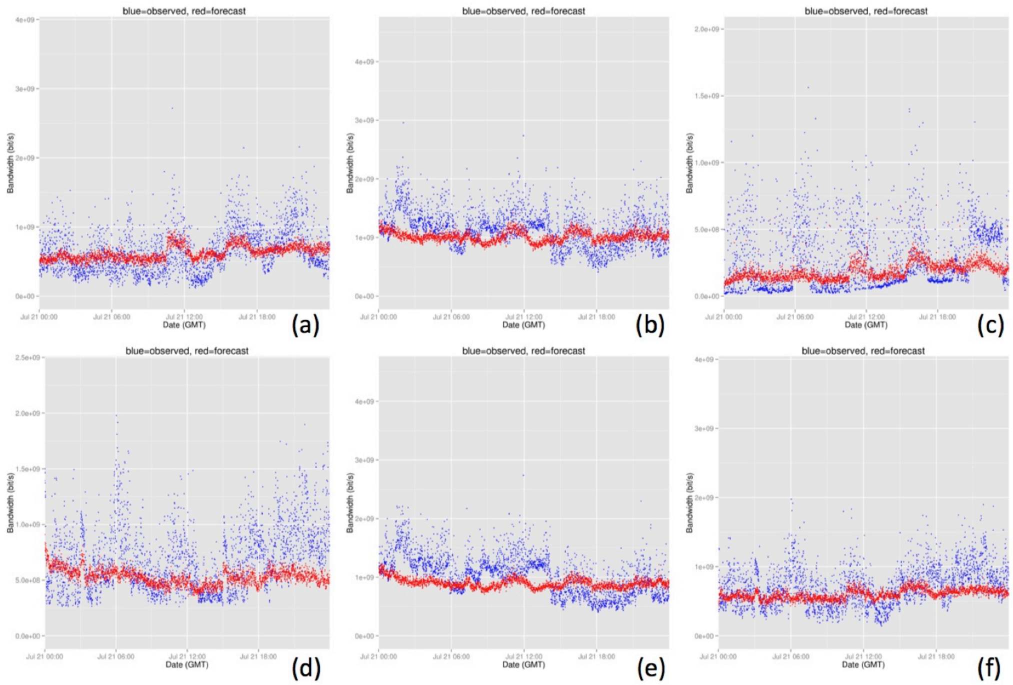


Figure 2: Bandwidth Utilization Forecast: The forecast is for one target date, July 21, GMT 2014. The x axis is the date. The y axis is the bandwidth (bit/s). Blue colors are for the observed data. Red colors are for the forecasts. (a) NERSC-ANL, (b) ANL-NERSC, (c) NERSC-ORNL, (d) ORNL-NERSC, (e) ANL-ORNL, (f) ORNL-ANL

## 2.2 Edge-to-Edge Transfer Throughput Inference Using Link Utilization Counts

We have developed a methodology to infer edge-to-edge achieved transfer throughput using link utilization counts (available through SNMP per-interface logs). Our method treats variations in the link utilization time-series as possible transfer starting or ending events (see Figure 3-A). Iteratively following these variations to the neighboring routers, we then identify the path the transfer traversed through the monitored network (see Figure 3-B).

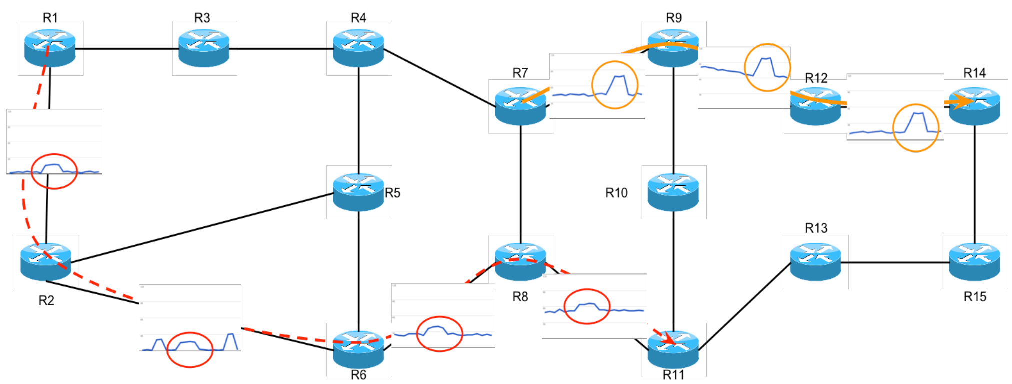


Figure 3-A: Using SNMP link usage data one can form the utilization time-series for each interface, which represents the traffic transferred between two connected routers. Analyzing the timeseries between R7 and R9 one can observe an increase in the link utilization. Such behavior can be attributed to a transfer initiated from R7's access network towards some destination. Following this increase from R9 to the next router and so on, we can observe that the corresponding transfer continues through R12 and R14. After R14 the transfer either continues to another network or is destined to a host in the access network served by R14. Other transfers can be identified in different parts of the network at the same time.

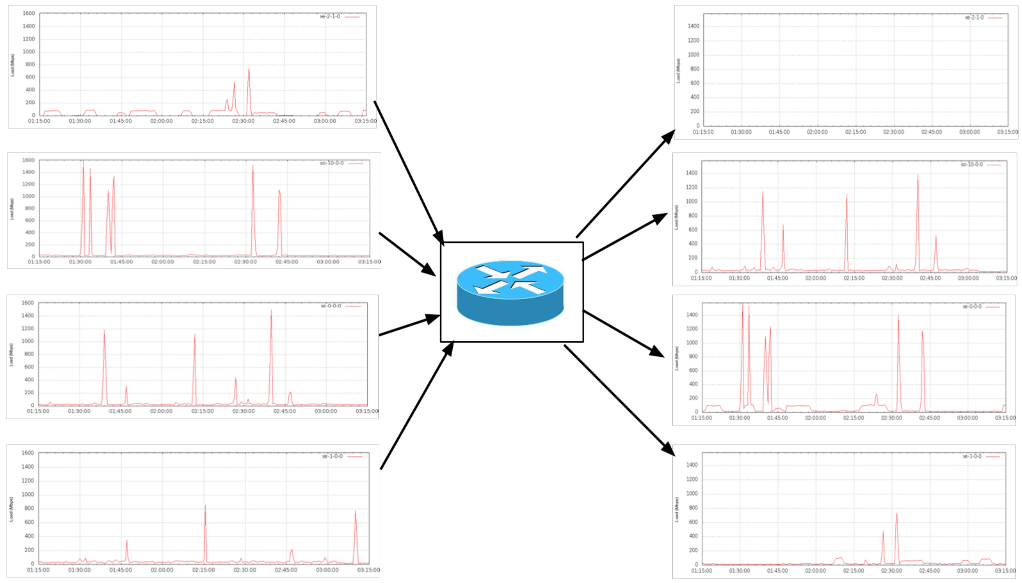

Figure 3-B: We developed an algorithm for mapping sudden link utilization changes from the input interfaces of a router to the output interfaces. The algorithm considers all possible combinations of output interfaces for every input interface to find a matching aggregate rate. Then it considers all possible combinations of inputs for every output to find a matching aggregate event.

Our evaluation, based on the ESnet topology and ESnet link utilization timeseries, shows that this method can identify events larger than 3Mbit/sec and longer than 2 minutes in duration with more than 95% recall (see Figure 3-C). Additionally, we show that event detection is strongly correlated with the traffic in the busiest router in the path.

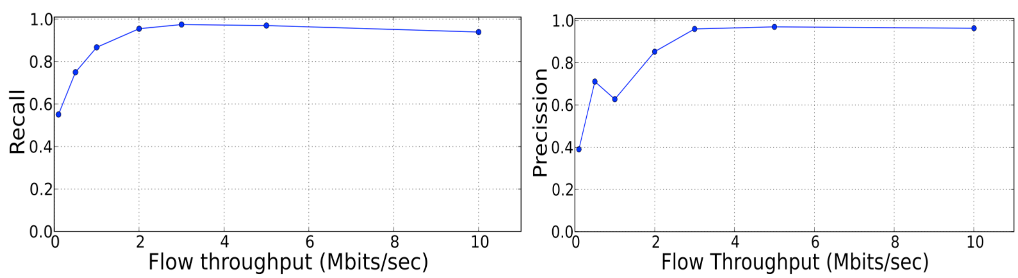


Figure 3-C: To evaluate the accuracy of the proposed methodology we conducted a number of TCP data transfers between Georgia Tech and Lawrence Berkeley National Lab (NBL). The transfers traverse the ESnet network, for which we have access to all intermediate router utilization data. The graphs show the recall and precision of the throughput inference algorithm. Our method manages to achieve more than 95% recall for transfers with throughput 2 Mbits/sec or higher. The precision reaches values larger than 95% for transfer throughputs larger than 3 Mbits/sec.

## 2.3 System prototypes

Figure 4 shows the user interface of the prototype that we developed to demonstrate the edge-to-edge throughput inference algorithm. The prototype analyzes SNMP-based link utilization data that we had previously downloaded from ESnet's SNMP repositories. A real-time version of the system that would only download the relevant SNMP data from ESnet every time a user makes a new throughput request was not possible due to limitations in ESnet's SNMP servers.

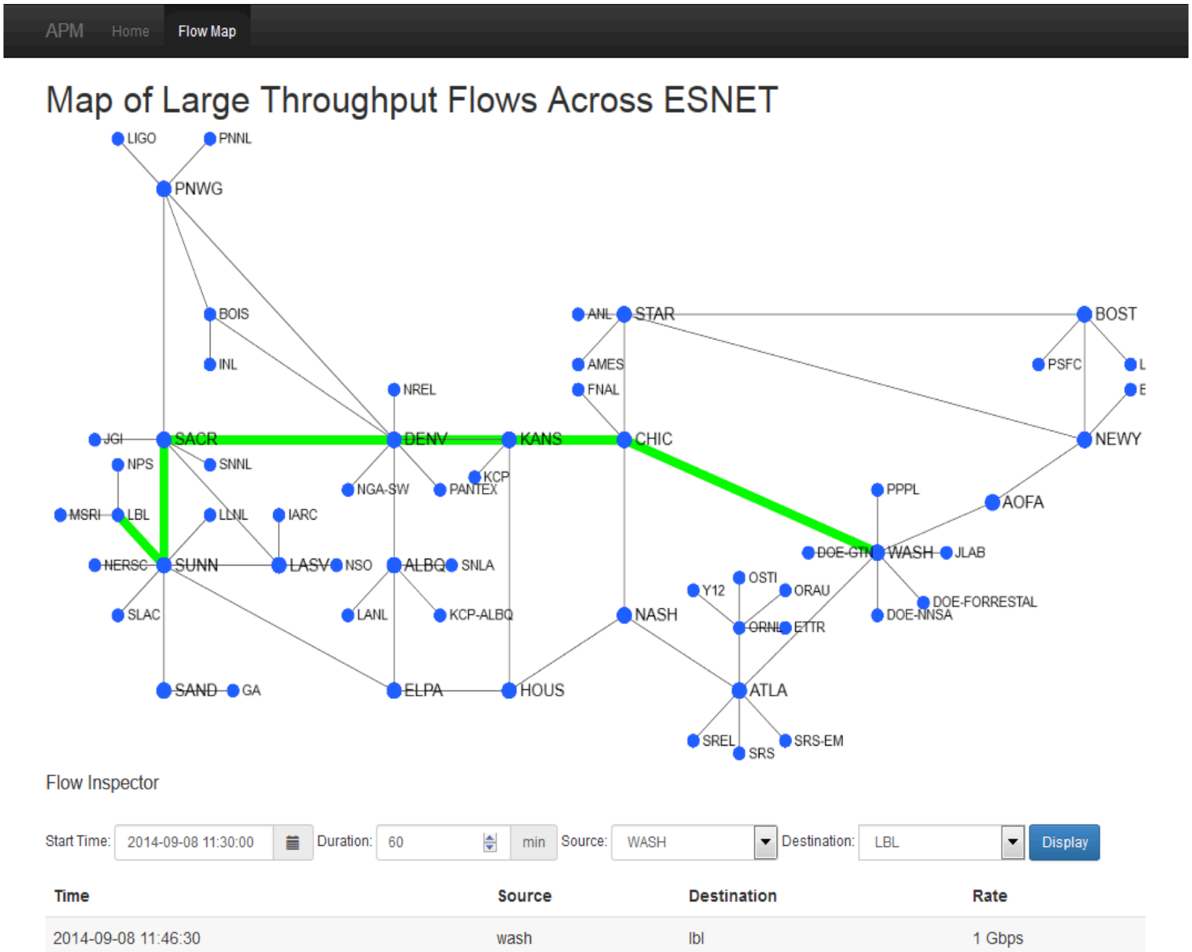


Figure 4: Similar to how a driver uses Google Maps to get an estimate of the driving distance between two locations, the user of our prototype can select two ESnet router interfaces: a source and a destination. Additionally, the user specifies a time duration. The system then identifies past transfers in that path and during the time period the user chose (analyzing ESnet link utilization data). If the system identifies any high-throughput data transfers in that path and time period, it outputs their start time and average throughput. The user can use this historic information to predict the achievable throughput in that path.

Figure 5 shows the prototype forecast system of the network bandwidth utilization, based on the model that we have developed with the live SNMP measurement data from ESnet. The prototype on the web shows the forecast of the bandwidth utilization on the selected network paths and the last two days of the actual bandwidth utilization overlayed with the forecast values. The current update frequency is set to one hour, however it can be updated with more frequently than one hour. The prototype can support applications or middlewares through framework API as well as end-users for efficient network resource utilization, scheduling and alternate path finding for data transfers.

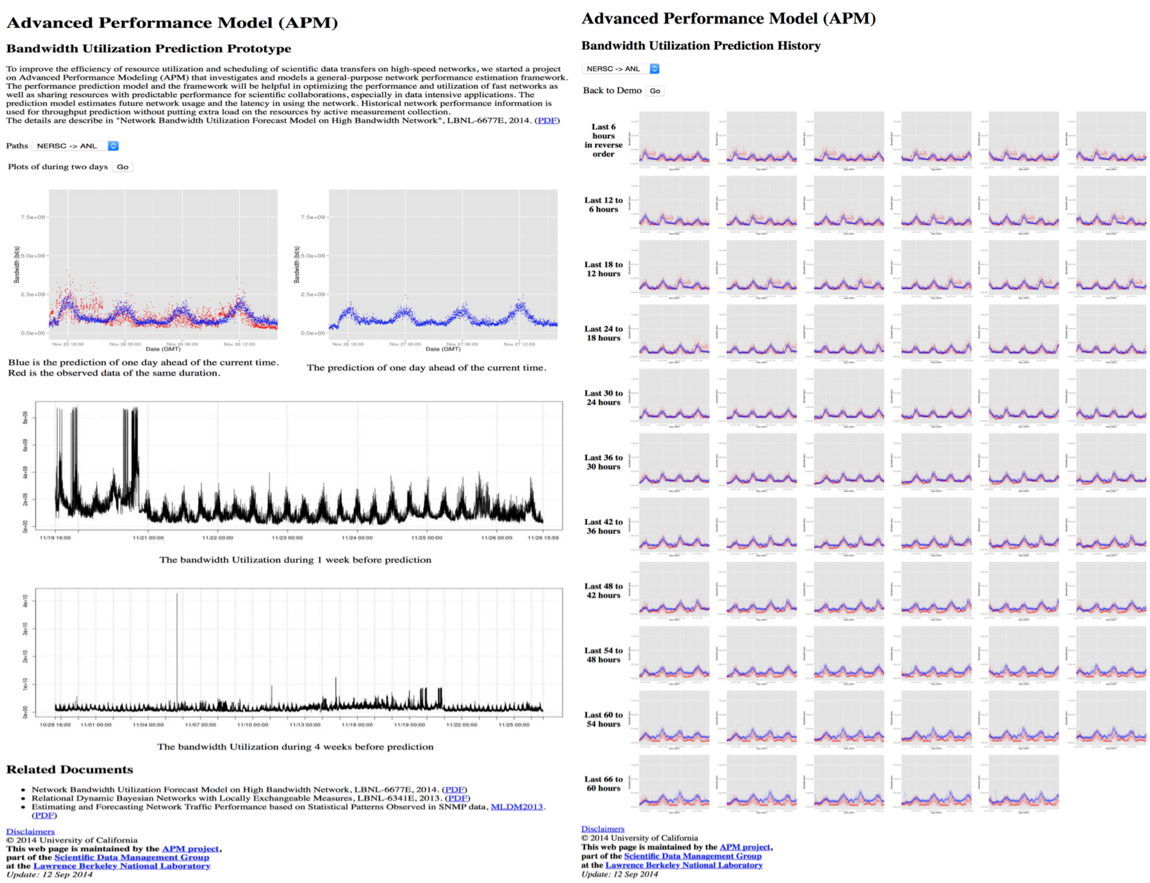


Figure 5: Network bandwidth utilization prediction system prototype on the web. The first page on the left shows the forecast of the bandwidth utilization on the NERSC-ANL network path, and the second page on the right shows the last two days of the actual bandwidth utilization overlayed with the forecasted values. Blue is the forecasted values, and red is the observed traffic measurements.

# 3 Summary of accomplishments during the first two years of the project

## 3.1 Access to relevant network measurement data

Simple Network Management Protocol (SNMP) link utilization time series and NetFlow logs have been identified as the most useful network performance data for the project collected by ESnet. Additional SNMP time series from Georgia Tech's main border router have been available for this project. An important data access agreement between our project and ESnet has been signed by the project members for the user-privacy concerns over Personally Identifiable Information (PII) issues around NetFlow data that the Institutional Review Board (IRB) at University of California at Berkeley (LBNL and ESnet share IRB with UCB) has recommended the data confidentiality and anonymization methods over IP addresses limited to R&E sites.

## 3.2 A novel edge-to-edge flow inference method using SNMP link utilization data

We have developed a methodology for inferring network transfers from SNMP traffic utilization time-series data that we can accurately identify edge-to-edge (e2e) flow information about large network transfers by leveraging SNMP link utilization data. Our edge-to-edge flow identification methodology is based on two main observations. First, there exist significant changes in a link's utilization time series when a high-throughput flow starts or ends. Second, these changes appear along the flow's path of successive router interfaces. These two observations allow us to identify the flow start/end times, its duration and its throughput and to track the flow along the network, determining its edge-to-edge path.

From our experiments using a client-server tool between two hosts located at Georgia Tech and LBNL, Figure 3 plots the true positive rate (TRP) as a function of the throughput of the transfer event, showing that the method can identify events larger than 3Mbits/sec with a TPR that is larger than 95%. The drop in TPR for smaller throughput values is mostly attributed to noise. Figure 4 plots the TPR as a function of the transfer duration, showing that the TPR rate seems to be stabilized for values larger than 90% for transfer durations that are longer than 60 seconds. In general, transfers with durations longer than or equal to two SNMP reporting periods can be identified with high accuracy.

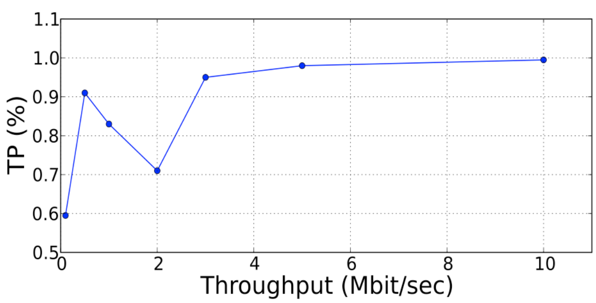


***Figure 3***: *True positive rate of event inference method as a function of the event throughput.*

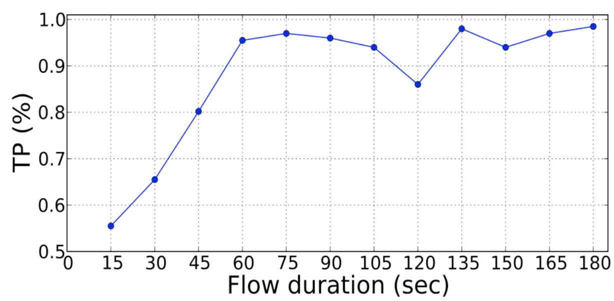


***Figure 4***: True positive rate of event inference method as a function of the event duration.

### 3.3 Study of statistical prediction model in analyzing NetFlow data

We developed an efficient statistical method, the Predictive Lasso, to analyze NetFlow data which calculates the performance prediction without multiple iterations in O($np$). We improved the model with the approach obtaining the estimates of fixed and random effects by Lasso with minimum mean squared prediction error (MSPE) in Linear Mixed Model (LMM). Then, we extended the results to Generalized Linear Mixed Model (GLMM) with the log link and Poisson assumption. A major computational advantage is that our procedure does not require the Expectation-Maximization (EM), unlike other previous methods, which require utilizing the EM algorithm to handle the unobserved random effects. We developed the new approach based on bootstrapping to select the optimal penalty parameter $\lambda$ in Lasso.

### 3.4 Study of data reduction method for large streaming data

We have studied a fundamental issue in large streaming data, which is to reduce the size of data and still obtain accurate statistical analysis. Our dynamic sampling algorithm reduces the data records in exponential scale, and still provides accurate analysis of large streaming data. We also built an efficient Gaussian Process with a fewer sample measurements, and applied the new algorithm to large data transfers in high-speed networks showing that the new algorithm significantly improves the efficiency of network traffic prediction for large data transfers. Our algorithm, Bayesian Online-Locally Exchangeable Measures (BO-LEMs), can reduce the number of required network measurement samples by 66% while achieving accurate data analysis for network throughput prediction.

### 3.5 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. 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 explored based on a few different time series cycles, 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, and compared the difference between forecast values and real values for the last frequency in the SNMP data. 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.

### 3.6 Use of SNMP data to identify large data transfers

We developed a method that is able to infer edge-to-edge 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.

### 3.7 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 collaborated with Shawn McKee at Univ. of Michigan for our access to the OSG data transfer log archive.
- We have collaborated with Warren Mathews and John Merritt at Georgia Tech for our access to the SNMP log data.

## 4 Publications, presentations and other activities

### 4.1 Publications

1) "*A Light-Weight Method to Estimate the Maximum Achievable Throughput in a Network Path*", Danny Lee, Constantine Dovrolis, in preparation, 2015.
2) "*Network-aware Client Clustering for Cloud Services*", Kamal Shadi, Danny H. Lee and Constantine Dovrolis, in preparation, 2015.
3) "*ISP Service Plan Identification Based on TCP Throughput Measurements*", Danny H. Lee, Kamal Shadi and Constantine Dovrolis, in preparation, 2015.
4) "*Bayesian Online Detection of Locally Exchangeable Measures*", J. Choi, W. Yoo, A. Sim, submitted to the Conference on Uncertainty in Artificial Intelligence (UAI 2015), 3/2015.
5) "*Analyzing High-Speed Network Data*", Kejia Hu, Jaesik Choi, Alex Sim, Jiming Jiang, under revision for the International Journal of Statistics and Probability, 2015.
6) "*Time-series Forecast Modeling on High-Bandwidth Wide Area Network Measurements*", W. Yoo, A. Sim, submitted to the Journal of Grid Computing, 2015.
7) "*Network Bandwidth Utilization Forecast Model on High Bandwidth Networks*", W. Yoo, A. Sim, International Conference on Computing, Networking and Communications (ICNC'15), 2/2015, LBNL-6677E.
8) "*Efficient Changing Pattern Detection on High Bandwidth Network Measurements*", W. Yoo, A. Sim, the 8th International Conference on Grid and Distributed Computing (GDC'14), 2014.
9) "*Edge-to-Edge Achieved Transfer Throughput Inference Using Link Utilization Counts*", D.

Antoniades, C. Dovrolis, the Sixth International Conference on Evolving Internet (INTERNET'14), 2014.
10) "*Best Predictive GLMM using LASSO with Application on High-Speed Network*", Kejia Hu, Jaesik Choi, Jiming Jiang, Alex Sim, Tech Report LBNL-6327E, 2013.
11) "*Relational Dynamic Bayesian Networks with Locally Exchangeable Measures*", Jaesik Choi, Kejia Hu, Alex Sim, Tech Report LBNL-6341E, 2013.
12) "*Estimating and Forecasting Network Traffic Performance based on Statistical Patterns Observed in SNMP data*", K. Hu, A. Sim, D. Antoniades, C. Dovrolis, The 9th International Conference on Machine Learning and Data Mining (MLDM2013), 2013.

## 4.2 Demos/posters

1) "*Edge-to-Edge Transfer Throughput Inference Using Link Utilization Counts*", D. Lee, NGNS PI Meeting, 9/2014.
2) "*Network Bandwidth Utilization Forecasting System Prototype*", W. Yoo, NGNS PI Meeting, 9/2014.
3) "*What SNMP data can tell us about Edge-to-Edge network performance*", D. Antoniades, K. Hu, A. Sim, C. Dovrolis, poster in the Passive and Active Measurements Conference (PAM2013), 2013.

## 4.3 Presentations

1) "*Data Reduction and Feature Discovery with Locally Exchangeable Measures for Massive Streaming Data*", A. Sim, Joint research meeting, LBNL and Naval Weapons Stations, 12/2014.
2) "*Advanced Performance Models*", A. Sim, C. Dovrolis, NGNS PI meeting, 9/2014.
3) "*Data Reduction Technique for Large Streaming Data by Locally Exchangeable Measures*", A. Sim, ESGF Workshop, 12/2013.
4) "*Advanced Performance Modeling with Combined Passive and Active Monitoring*", A. Sim, C. Dovrolis, PI Meeting, 3/2013.
5) "*Statistical prediction models for network traffic performance*", K. Hu, A. Sim, D. Antoniades, C. Dovrolis, The Winter 2013 APAN/ESnet /Internet2 technical meeting (Joint Techs, TIP2013), 2013.

## 4.4 Patents

1) "*Data Reduction Methods, Systems, and Devices*", J. Choi, A. Sim, U.S. Patent Application, serial no. 14/555,365, filed on 11/26/2014 (Patent Pending).
2) "*An Efficient Data Reduction Method with Locally Exchangeable Measures*", J. Choi, A. Sim, U.S. Provisional Patent Application, serial no. 61/909,518, filed on 11/27/2013.