Measurement Challenges and Opportunities for Future Science Networks

Yan Luo
Department of Electrical and Computer Engineering
University of Massachusetts Lowell
yan_luo@uml.edu , http://acanets.uml.edu/~yluo/

Network measurement is instrumental in quantifying the performance and status of network infrastructure, understanding network usage activities, and helping network operators to maintain networks and respond to network issues. Network measurement, however, has faced several challenges: (1) the ever increasing line rate. It is non-trivial to provide detailed per-flow measurement data at a line rate of 100+Gbps; (2) the diversity and dynamics of measurement tasks. The measurement target could change rapidly, so do the interested flow metrics; and (3) deriving useful knowledge from measurement data and proactively leveraging the knowledge. The sheer amount of measurement data has introduced a big data challenge on how we store and process them to learn interesting patterns and predict trends, which have not been clearly grasped or adequately utilized.

In the past decade, *the key problem of network measurement has been that the measurement is merely a second class citizen until a network issue arises.* As a result, the network down-time or performance degradation is inevitable, hindering important science missions. We believe that there are a number of directions in innovating network measurement which would yield significant leap forward:

(1) *Flexible control plane of network measurement infrastructure*. PerfSONAR has been deployed at a global scale, however, the current methodology of initiating measurement tests requires a high degree of human reasoning, seriously lagging behind the speed and scale of perfSONAR deployment. A control plane of perfSONAR is highly desirable, where we envision the high-volume historical measurement data can be analyzed and mined, and future measurement tasks can be initiated on-demand without human intervention. The control plane is to serve as an interface between the measurement substrate and the other network layers, where the capability of measurement instrument and services can be exposed in a consistent way. The control plane is also to orchestrate large scale collaborative measurement given a growing set of measurement devices such as perfSONAR nodes.

(2) *Software defined measurement (SDM).* Programming a network measurement task has become increasing important and interesting, because the measurement objectives typically vary from time to time in scope, granularity, conditions, and actions. Using fixed measurement metric for every single flow is not necessary or feasible as one may need to "zoom in" or "zoom out" on a particular flow. The measurement may have to start when a given condition is met. There is a strong need to support such flexibility, and a software-defined approach is very promising. We envision that a SDM data plane consists of a set of APIs for defining and instantiating a measurement task with its measurement target, metrics, algorithms, and conditions of start/stop it. New programming languages such as P4 is a salient candidate in formally defining measurement tasks, and porting the tasks onto a variety of hardware architectures (e.g. x86, FPGA, network processor, etc.).

(3) *Mining measurement data*. Big data has changed how science discovery works. Network data, in particular the network measurement data, increase in size as well as their value of providing useful information about the network. The analysis of network flow data obtained from measurement instruments is challenged with the data volume and a variety of flow features. This calls for new systems to process the streaming data, new algorithms to derive knowledge from data, and new control schemes to take advantages of the learned network usage patterns.

(4) *Adaptive cross-layer network optimization*. The knowledge from network measurement data would enhance the predictability of network performance, and the software defined measurement framework would allow a future network protocol to proactively gather up-to-date metrics, and adapt the protocol parameters in response to network conditions. This is to elevate the network measurement to a first class citizen so that its benefits can be maximized for supporting high-speed critical science networks.

Reference

[1] Hao Yin, Yong Jiang, Chuang Lin, Yan Luo, Yunjie Liu, Big Data: Transforming the Design Philosophy of Future Internet, IEEE Network, Vol. 28, No. 4, 2014.

[2] P4: Programming Protocol-independent Packet Processors. http://www.p4.org/

[3] Chen Xu, Xiaoban Wu, Yan Luo, Brian Tierney and Jeronimo Bezerra, PSCP: PerfSONAR Control Plane for Software Defined Measurement and Troubleshooting. In submission.

[4] Peilong Li, Yan Luo, Ning Zhang, and Yu Cao. HeteroSpark: A heterogeneous cpu/gpu spark platform for machine learning algorithms. In Networking, Architecture and Storage (NAS), 2015 IEEE International Conference on, pages 347–348, Aug 2015.