Livermore Computer Network Simulation Program

(Poster Abstract)

Peter D. Barnes, Jr. (pdbarnes@llnl.gov), James M. Brase (brase1@llnl.gov), Thomas W. Canales (tcanales@llnl.gov), Matthew M. Damante (mdamante@llnl.gov), Matthew A. Horsley (horsley1@llnl.gov), David R. Jefferson (drjefferson@llnl.gov), Ron A. Soltz (soltz@llnl.gov)
Lawrence Livermore National Laboratory
7000 East Ave.
Livermore CA 94550 USA
1-925-422-3384

ABSTRACT

The Livermore Lab has embarked on a multi-year effort to develop a large-scale realistic network simulation capability. Specifically, we are developing computer network simulations for realistic networks derived from real and synthetic network maps, and which incorporate real hardware and geographic constraints, at enterprise (10K node) and above scale; incorporate near-real-time updates from the real global Internet; and generate traffic from realistic traffic models matched to observed data. In this poster we describe our approach and specific applications areas of interest.

Categories and Subject Descriptors

I.6.3 [**Simulation and Modeling–Applications**]

General Terms

Experimentation, Security

Keywords

Network simulation, simulated applications.

1. INTRODUCTION

Predictive analysis of cyber risk and performance is one of the major gaps in cyber analytics.[1] Understanding how a specified mission-critical application will execute in a network context, characterizing the potential impact of network threats on critical applications, and predicting the effect of proposed defensive actions are critical capabilities for a risk-based cyber strategy. The Livermore Lab has embarked on a multi-year effort to develop a large-scale realistic network simulation capability. Specifically, we are developing computer network simulations for realistic networks derived from real and synthetic network maps, and which incorporate real hardware and geographic constraints, at enterprise (10K node) and above scale; incorporate near-real-time updates from the real global Internet; and generate traffic from realistic traffic models matched to observed data. In this poster we describe our approach and specific applications areas of interest.

Network simulation has been an active area of work since the 1960's,[2] resulting in a broad set of both commercial[3, 4] and open-source[5] tools. Network simulation is based on discrete event simulation[6]–the most basic event is the sending receiving of a network packet. Nodes in the simulation can be host computers, which create and receive packets, and routers, which forward packets on the route to their destination host. The simulators generally implement full TCP/IP network protocol stacks over physical models for wired and wireless RF communication links. Network simulators are generally used in the development of new network technologies–new routers, protocol variations, congestion control algorithms, *etc*. In these applications simulation of networks with hundreds of host computer and routers is adequate and there is little motivation to extend simulations to much larger networks. Most existing efforts are limited to modest scale (few hundred nodes), unrealistic network models,[7] and unrealistically simple on|off traffic models.[8]

For our intended applications existing network simulators are limited in three regards:

- Host behavioral models are unrealistically simple.[8] To reproduce behaviors seen in real networks we will need more sophisticated user models representing more complex activities like Web surfing, e-mail interchanges, and peer-to-peer file interchange.
- There has been little effort to scale network simulations to even the enterprise network level. A few demonstrations of parallelized network simulation have been performed at Georgia Tech[9] and at the Army Research Lab[10] but these efforts have barely begun to explore the area. For example, little is known about optimal cluster configurations or effective mapping of simulated nodes and communication links to physical compute nodes.
- There has been little systematic validation of the simulations outside the narrow range of detailed network technology applications noted above. In particular the ability of simulations to produce statistically realistic network behaviors at enterprise scale and above is completely unexplored. This is exactly the performance space of interest for mission assurance applications.

To focus our research efforts, we have identified three application areas: enterprise networks, mission-critical applications, and worldwide routing. We will discuss each of these applications below.

This paper is organized as follows: in Section 2 we discuss our over-arching research goals, and existing capabilities; Section 3

Simutools 2012, March 19-23, Desenzano del Garda, Italy

DOI 10.4108/icst.simutools.2012.247748

Simutools 2012, March 19-23, Desenzano del Garda, Italy

DOI 10.4108/icst.simutools.2012.247748

describes the enterprise network application; Section 4 describes mission-critical application models; Section 5 describes the worldwide routing application; and Section 6 offers some conclusions.

2. RESEARCH GOALS AND CAPABILITIES

Our research goals are centered on understanding the capabilities and limitations of network simulation. In the application areas we want to focus on the following questions:

- Can we reproduce the statistics of observed behaviors at scales from enterprise-level networks up to the global Internet? What model fidelity is needed to produce a given behavior? What level of abstraction can we get away with?
- Can we integrate models at different scales to achieve high fidelity and large scale, *e.g,* virtualized nodes and networks around nodes of interest, while using more abstract packet-level simulations at the largest scales?
- What are the limits to scaling network simulations with current tools?
- Can we predict the response of the network to changes in topology or dynamics?

In addition to utilizing Livermore's significant high-performance computing resources, we will take advantage of several other existing research programs at the Lab.

The Livermore Laboratory has ongoing efforts in understanding network topology and services, analysis of live traffic capture, host-based behavior tracking, and data analysis on large graphs. Our network mapper, which provides highly detailed descriptions of real networks and services, in combination with host-based measurement and live traffic capture and analysis, provide an unprecedented source of validation data for realistic behavior models and associated traffic generators.

We have surveyed and evaluated existing network simulation frameworks, opting to begin with ns3.[5] To date we have developed an XML-based network description language to describe the simulation topology and applications, and generate the simulation code automatically. We have outlined a statistically driven model to generate realistic behavior. We have identified a series of test problems for the application areas described below. These test problems are typically simplified versions of the ultimate application scope, based on published work, so we have a point to validate against.

3. ENTERPRISE NETWORK APPLICATION

An enterprise network consists of ~10K nodes, with most nodes in trees attached to a core clique of fully connected central routers. (See Figure 1.) The background for this network is the rest of the Internet, connected to the core routers (through an edge router) by a very small number of links, typically only one, with a second backup connection. The combination of fully connected core routers and few links to the larger Internet gives these networks a definite sense of inside and outside. Traffic flow is dominantly between internal hosts, with significant Internet traffic.

We plan to couple results from current maps of realistic networks, including the Lab, with behavioral data from our traffic capture and host-based behavior projects. The overall objective is to model enterprise networks with realistic traffic generators, and measure the range of variability of realistic networks given constraints from mapping data.

There are many tools for mapping enterprise networks,[11-15] and some simulation studies of performance. We believe that quantifying errors in mapping, generating realistic traffic, and multi-scale network modeling are all new.

There are a number of specific tasks required. We have developed the capability to convert a network map into a simulation topology, complete with specification of the variety of traffic-generating applications to be simulated on each node. We will be studying how to create ensembles of network models consistent with the mapping input data, and developing metrics to quantify performance from the ensembles. We will also create multi-scale models to study fidelity issues.

4. MISSION-CRITICAL APPLICATIONS

The scenario is to model the data flow and performance for a critical set of application traffic flows embedded in a larger background network. Mission success relies on timely delivery of multiple data streams from/to sites around the world. The overall objective is to analyze mission performance under nominal and severely disrupted network states.

This application is challenging because many details of the target networks may be unknowable: exact topology, exact background application mixes at each node, exact mission-critical data flow parameters. Therefore we will have to develop a range of background models, and a range of disruption models, which collectively span the space of likely realizations.

5. WORLDWIDE ROUTING

The global Internet is managed by ~$4x10^4$ administrative units called Autonomous Systems (AS).[16] AS's exchange routing information with their neighbors (peers) on ~$4x10^5$ total advertised address prefixes using the Border Gateway Protocol (BGP). Since the exchange of routing information between two AS peers typically only contains local information, it is non-trivial to gather a global view of the routes currently available. Fortunately this has been the subject of over a decade of research by multiple projects,[17-21] and now there are 10 years of

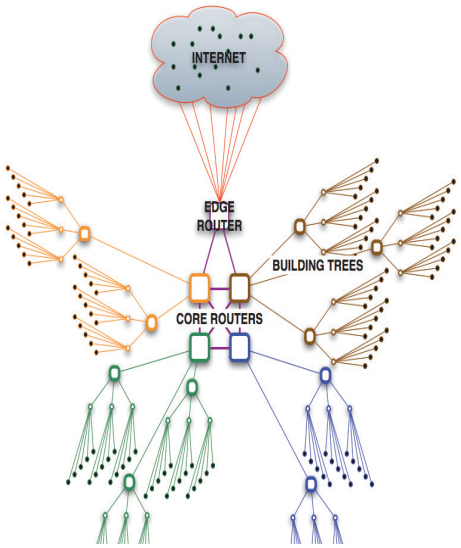

Figure 1. Enterprise network diagram.

archived routing updates,[22] and current routing updates from around the world are available in near real-time.[19]

The routing state of the Internet has seen numerous wide-spread anomalies, which have been traced back ultimately to single accidental (or malicious) mis-configurations;[23, 24] hardware failure caused by weather and construction,[25] software bugs,[26] changes to the BGP protocol[27] and other causes. In some cases failures can cascade, because of co-located hardware or the convergence properties of BGP itself.[24, 28] With near real time BGP updates there is the potential to analyze the updates to determine the base cause,[29] extrapolate to possible cascading failures, and predict future routing and performance. The overall objective is to develop a worldwide routing model seeded in near real time from the live Internet, and use the model to infer originating failures and likely future behavior of the Internet.

The first task here is to build a global Internet model based on existing analysis of BGP updates. Subsequent tasks will develop a framework for incorporating real time BGP updates, implement fault origination algorithms, and demonstrate routing prediction using the global model and real time updates.

6. CONCLUSIONS

We are developing capability to simulate realistic networks, derived from real and synthetic network maps at enterprise (10K node) and above scale, incorporate near-real-time updates from the real global Internet; and generate traffic from realistic traffic models matched to observed data. We aim to understand the capabilities and limitations of large-scale network simulations, with demonstrated applications in cyber security, global network situational awareness, performance modeling and prediction.

7. ACKNOWLEDGMENTS

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

8. REFERENCES

1. J. M. McConnell, *Vision 2015: a globally networked and integrated intelligence enterprise*, July 2008 (Director of National Intelligence, 2008).
2. *Modeling and Tools for Network Simulation*, edited by K. Wehrle, M. Günes, and J. Gross (Springer, New York, 2010).
3. A. Varga, *The OMNET++ discrete event simulation system* in *European Simulation Multiconference ESM'2001*, Prague, Czech Republic, 2001), https://labo4g.enstb.fr/twiki/pub/Simulator/SimulatorReferences/esm2001-meth48.pdf.
4. A. Varga and R. Hornig, *An overview of the OMNeT++ simulation environment* in *the 1st international conference on Simulation tools and techniques for communications, networks and systems* (ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Marseille, France, 2008).
5. ns3 Collaboration, *The ns-3 network simulator*, Washington, 2011), Vol. 2011, http://www.nsnam.org/index.html.
6. R. Fujimoto, *Parallel and Distributed Simulation Systems,* (John Wiley & Sons, 2000).
7. L. Li*, et al.*, *A first-principles approach to understanding the internet's router-level topology*, SIGCOMM Comput. Commun. Rev. **34**, 4, 3 (2004).
8. E. K. Çetinkaya*, et al.*, *A comprehensive framework to simulate network attacks and challenges* in *IEEE Second International Workshop on Reliable Networks Design and Modeling (RNDM'10)*, Moscow, 2010).
9. C. D. Carothers, D. Bauer, and S. Pearce, *ROSS: A high-performance, low-memory, modular Time Warp system*, Journal of Parallel and Distributed Computing **62**, 11, 1648 (2002), <Go to ISI>://000179497400003.
10. J. Clarke*, et al.*, *The Network Interdisciplinary Computing Environment* (US Army Research Laboratory, 2011).
11. Lumeta, *Lumeta - Global Network Visibility*, 2011), Vol. 2011, http://www.lumeta.com/.
12. AdRemSoft, *NetCrunch*, 2011), Vol. 2011, http://www.adremsoft.com/netcrunch/.
13. Nmap.org, *nmap*, 20111), Vol. 2011, http://nmap.org/.
14. Q. Software;, *PacketTrap*, 2011), Vol. 2011, http://www.packettrap.com/network/index.aspx.
15. Solarwinds, *LANSurveyor*, 2011), Vol. 2011, http://www.solarwinds.com/products/LANsurveyor/.
16. T. Bates, P. Smith, and G. Huston, *CIDR Report for 27 May 11*, 2011), Vol. 2011, http://www.cidr-report.org/as2.0/.
17. R. V. Oliveira*, et al.*, *Cyclops*, 2011), http://cyclops.cs.ucla.edu/.
18. M. Luckie, *Scamper: a scalable and extensible packet prober for active measurement of the internet,* in proceedings of 10'th Annual Conference on Internet Measurements, Melbourne, Australia, November 1-3, 2010, (ACM), 239.
19. D. Meyer, *RouteViews* (University of Oregon, 2011), Vol. 2011, http://www.routeviews.org/.
20. D. Moore*, et al.*, *The CoralReef Software Suite as a Tool for System and Network Administrators* in *LISA '01 15'th USENIX Conference on System Administration* (USENIX, San Diego, CA, 2001).
21. Y. Hyun, *The Archipelago Measurement Infrastructure* in *7th CAIDA-WIDE Workshop* (CAIDA, La Jolla, CA, 2006), http://www.caida.org/publications/presentations/2006/young_wide0611_ark/.
22. CAIDA, *IPv4 Routed /24 AS Links Dataset*, 2011), Vol. 2011, http://www.caida.org/data/active/ipv4_routed_topology_aslinks_dataset.xml.
23. R. Lemos, *China's Internet hijack: Attack or accident?*, (e-print), http://www.infoworld.com/t/routers-and-switches/chinas-internet-hijack-attack-or-accident-461.
24. E. Zmijewski, *Reckless Driving on the Internet* in *Renesys blog*, 2009), Vol. 2011, http://www.renesys.com/blog/2009/02/the-flap-heard-around-the-worl.shtml.
25. J. Regan, *Update 3-Undersea cable breaks cut Internet in Mideast, Asia* in *Reuters*, 2008), http://www.reuters.com/article/2008/12/20/us-internet-idUSTRE4BJ0FV20081220.
26. Opte Project, *Cisco bug crashed Internet* (UCLA, Los Angeles, 2011), Vol. 2011, http://cyclops.cs.ucla.edu/blog/?p=96.
27. P. A. A. Gutierrez, *Collateral Damage in the Last Big Internet Storm* in *2010 Sixth Advanced Internation Conference on Telecommunications (AICT)*, Barcelona, 2010), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5489766&tag=1.
28. K. Sriram*, et al.*, *Study of BGP Peering Session Attacks and Their Impacts on Routing Performance*, IEEE Journal on Selected Areas in Communications **24**, 10, 1901 (2006), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1705621&tag=1.
29. A. Campisano*, et al.*, *Tracking Back the Root Cause of a Path Change in Interdomain Routing* in *Network Operations and Management Symposium, 2008*, Salvador, Bahia, 2008), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4575166.