

Simulation of Large (>10K node) Computer Networks

Tuesday, 11 September 2012 10:11 (12 minutes)

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, and generate traffic from realistic traffic models matched to observed data. In this abstract 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. In this abstract we summarize our approach to modeling enterprise networks and understanding the scaling issues involved.

1. 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 each of the application areas described. These test problems are typically simplified versions of the ultimate application scope, based on published work, so we have a point to validate against.

1. 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. 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, but 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 are currently studying the limitations in simulating 10K node networks where most hosts are actively generating traffic. 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.

In conclusion, we are developing capability to simulate realistic networks, derived from real and synthetic network maps at enterprise (10K node) and above scale, 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.

1. ACKNOWLEDGMENTS

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

1. 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. ns-3 Collaboration, The ns-3 network simulator (Washington, 2011), Vol. 2011, <http://www.nsnam.org/>.
6. R. Fujimoto, Parallel and Distributed Simulation Systems, (John Wiley & Sons, 2000).

8. L. Li, et al., A first-principles approach to understanding the internet's router-level topology, SIGCOMM Comput. Commun. Rev. 34, 4, 3 (2004).
9. 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).
10. 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.
11. J. Clarke, et al., The Network Interdisciplinary Computing Environment (US Army Research Laboratory, 2011).
12. Lumeta, Lumeta - Global Network Visibility (2011), Vol. 2011, <http://www.lumeta.com/>.
13. AdRemSoft, NetCrunch (2011), Vol. 2011, <http://www.adremsoft.com/netcrunch/>.
14. Nmap.org, nmap (2011), Vol. 2011, <http://nmap.org/>.
15. Q. Software, PacketTrap (2011), Vol. 2011, <http://www.packettrap.com/network/index.aspx>.
16. Solarwinds, LANSurveyor (2011), Vol. 2011, <http://www.solarwinds.com/products/LANsurveyor/>.

Primary author: Dr BARNES, JR., Peter (Lawrence Livermore National Laboratory)

Co-authors: Dr JEFFERSON, David (Lawrence Livermore National Laboratory); Dr COLON, Domingo (Lawrence Livermore National Laboratory); Dr BRASE, James (Lawrence Livermore National Laboratory); Dr HORSLEY, Matthew (Lawrence Livermore National Laboratory); Dr SOLTZ, Ron (Lawrence Livermore National Laboratory); Dr NIKOLAEV, Sergei (Lawrence Livermore National Laboratory)

Presenter: Dr BARNES, JR., Peter (Lawrence Livermore National Laboratory)