4-2015

The Truman Show: Scalable generation of artificial network traffic for cyber security research

Adam Wirth
University of Southern Maine, adam.wirth@maine.edu

The Truman Show:

Scalable generation of artificial network traffic for cyber security research

by Adam Wirth and Isaiah Marvin with Dr. Glenn Wilson of MCSC

Abstract

Network traffic generation is a key component of the creation of a network simulation environment. In order to create a realistic simulation of a large scale network in action, it is necessary to have a large volume of user traffic. Past methods for providing user traffic, such as hiring users to manually generate traffic, or recording existing traffic of an active network, provide limited control. Managing and directing users is time consuming, and using existing traffic is restricted by recorded conditions.

The Truman Show project provides an infinitely adjustable alternative by utilizing agent based modeling of individual users. An agent based model is a statistical method for modeling a system not as a single object, but as numerous subsystems. The Truman Show uses a similar methodology for the generation of network traffic by generating activity not from a single recording, or small handful of users, but with the creation of a master program that generates an infinitely scalable number of subprograms, each of which mimic real user behaviors.

Introduction

The MCSC cybersecurity range on the USM Portland campus is a grant-funded endeavor to build and maintain a secure environment for students and security professionals to learn about malicious computer network attacks. It has many distinct components, including, lots of hardware, a fictional town named Betaport, and the TrumanShow: a software solution to simulate thousands of concurrent users on a local, or “offline”, network of websites.

The urgency of cyber security, to both individuals and corporations, cannot be understated. In 2012, 7,200 cyber security issues or breaches were reported to CERT Australia. That number had been surpassed in just August 2013. A “conservative” estimate from Privacy Rights Clearinghouse says that, since 2005, 543 million reported records were breached.

Objective

The whole project’s goal is to create and maintain a secure environment to realistically test computer viruses and malware with network traffic.

The TrumanShow aspect of the project is concerned with the creating of the network’s user traffic. The purpose of the TrumanShow is to simulate massive amounts of localized, artificial network traffic. This is to keep the sandbox environment secure from any potential threats, and helps security testers work with different scenarios more easily.

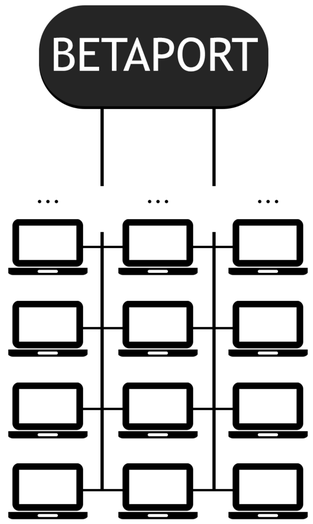

Fig1: A normal network of numerous individual users

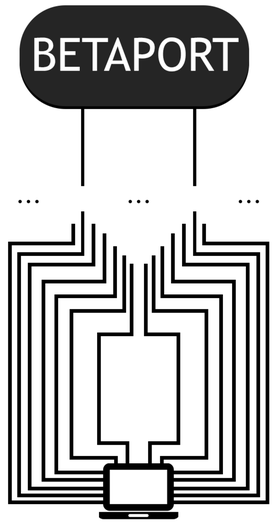

Fig2: network with users represented within a single system

Evolution

The TrumanShow was first written in the programming language Perl in the beginning of summer 2014. Development switched from Perl to Python, another programming language, in the Fall. This switch was done because Perl was a less common language, while Python had a lot of support and available resources to expedite development. Hardware for the whole project was installed over time; the TrumanShow development computers moved into the closed network in December 2014.

Locust.io

One of the main reasons we switched from Perl to Python as a programming language was because of the availability of open source resources capable of doing aspects of the TrumanShow. Having a framework in place, rather than writing it, saved us a lot of time getting started on development. Additionally, the two languages are somewhat similar, so picking up Python did not use up a lot of time.

The open source framework we used for the TrumanShow was “Locust.io”. It runs in Python scripts, and lets us easily simulate thousands of concurrent users. Normally, creating thousands of users would be too computationally expensive, but the capabilities of Locust are built on a different approach to creating each user, making it much more manageable, and also somewhat straightforward to write.

Next steps

There’s a long list of features and capabilities we’d like to add to the TrumanShow. The gist of the goals are to

- Continue to adapt to the growing number of web pages local to the network
- Implement more internet protocols and tasks for TrumanShow to simulate
- Adding more task-sets and actions to provide a more accurate sandbox simulation

Acknowledgments

Thank you to Edward Silher, Lynn Lovewell, Jim Owens, and everyone who’s been supportive throughout the lifespan of this project.

References

An Open Source Load Testing Tool. Program documentation. Locust - A Modern Load Testing Framework. N.p., n.d. Web. 05 Apr. 2015.

Patteson, Carolyn. "Cyber Security - the Facts." Australian Security Magazine. Australian Security Magazine, 21 Oct. 2013. Web. 05 Apr. 2015.

Privacy Rights Clearinghouse. Data Breaches: A Year in Review. Data Breaches: A Year in Review. Privacy Rights Clearinghouse, 16 Dec. 2011. Web. 05 Apr. 2015.