MEGATUX

An Internet Emulation System to Enable Predictive Simulation of Nation-scale Internet Behavior

LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Sandia National Laboratories

Ronald Minnich / Rob Armstrong / Don Rudish / Kevin Pedretti / David Thompson / David Evensky / Ann Gentile / Brian Wylie

Problem

❑Botnets are Prevalent in U.S.

Rank	# compromised in the US	Name	Purpose
1	3.6 Million	Zeus	Stealing
2	2.9 Million	Koobface	Rootkit
3	1.5 Million	TidServ	Rootkit
4	1.4 Million	Fakeavalert	Spaming/Spreading
5	1.2 Million	TR/Dldr.Agent.JKH	Clickbot
6	520,000	Monkif	Downloading Adware
7	480,000	Hamweq	Stealing
8	370,000	Swizzor	Adware
9	230,000	Gammima	Stealing
10	210,000	Conficker	Spread / ?

Scale of today's cyber systems has exceeded current modeling capabilities. Above are the sizes of the 10 most damaging botnets according to the Damballa security firm.

❑Reverse Engineering is Time Consuming

– Frequently, a manual process
– May not show global botnet behavior

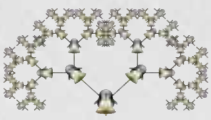
❑Malware is spreading to other platforms

– Mobile devices are fully capable computers that can run Malware
– Most mobiles are Unix or Windows like operating systems.

Approach

❑Forensics is important and good but there may be a faster alternative.
❑Use *observation* of **large scale** emulations to determine malware behavior.
❑Bypass the Reverse Engineering Phase

❑Combine Emulation and Analytics to combat cyber threats.
❑Create a testing Environment the size of a nation scale Internet with the help of virtual machines. Nodes of 10^6+ scale.
❑Conduct replayable and repeatable Testing/Evaluation/Assessment



❑Use *computational configuration* to bring up virtual Internet containing millions of nodes/routers/services in minutes.
❑No central static configuration for the virtual Internet.
❑Use real operating systems / applications found on the Internet

Results

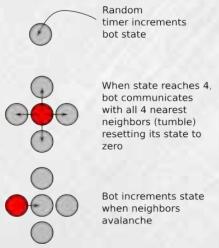
❑**Place Sandia In The Forefront of Large-scale Emulation**

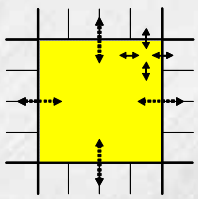
Science Times

The New York Times

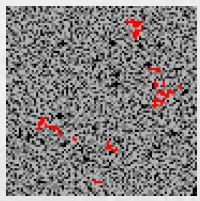
Researchers Create a Mini-Internet to Stalk Botnets

❑**Developed Malware Prototype called the "Sandbot" based of the Sand Pile model from complexity theory.**



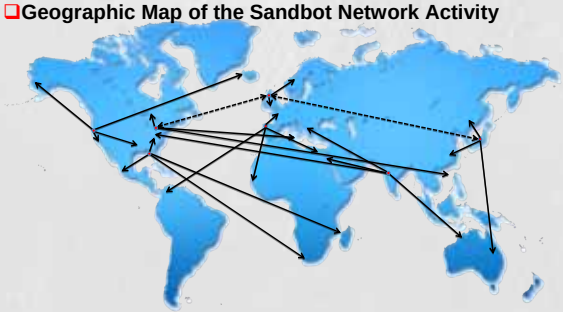


Abstract view of the square lattice. Each cell operates on a limited simple rule set.

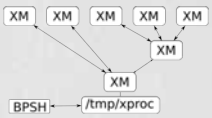


View of Overlay network seen from the bot herder. Each Cell is a virtual machine running Linux.

❑**Geographic Map of the Sandbot Network Activity**



❑**Developed XPROC: A scalable means for remote execution**



❑Tree-Spawn execution
❑Nodes help other nodes execute a process
❑Exploit Virtual Machine properties
❑Limit unwanted network traffic through virtual block devices and VIRTIO console devices.

❑**Developed VMATIC: Virtual Machine provisioning tool**

❑Configures and provisions virtual machines
❑Runs on any machine that can boot Linux that is x86 or ARM based
❑Uses computational configuration to boot millions of nodes in minutes.

Significance

❑Gives researcher the ability to study malware in an Isolated Internet like environment.
❑A step toward emulation at 10^7 scale and up (nation-state)
❑Demonstrated capabilities for use in Emulytics Roadmap
❑Team is connected to DOE Grassroots cyber security initiative
❑Identified complexity theory concepts (cascades, robustness) relevant to protection of cyber systems

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

SAND No. 2010-5993 P

Sandia National Laboratories