Current Developments in DETER Cybersecurity Testbed Technology

Terry V. Benzel – Presenter
CATCH 2009 Conference – March 4, 2009

Terry Benzel*, Bob Braden*, Ted Faber*, Jelena Mirkovic*, Steve Schwab†, Karen R. Sollins+, and John T. Wroclawski*

*USC Information Sciences Institute
†Cobham, Inc
+MIT Computer Science and Artificial Intelligence Laboratory
Talk Outline

Today’s DETER Testbed

Key New Research and Development:

• Dynamic Federation
• Experiment Health Support
• Risky Experiment Management

Next Generation Capabilities
DETERlab: The DETER Facility

- Cyber Security testbed located at USC/ISI and UC Berkeley
  - Funded by NSF and DHS, started in 2003
  - 400 Nodes - 200 each at ISI and UC Berkeley
  - Based on Emulab software, with focus on security experimentation

- The DETER testbed infrastructure
  - Tool libraries
  - SEER workbench
  - Operations and Management
  - Community building and Outreach
Why DETER?

• Lack of experimental **infrastructure** for cyber-security research
  – Medium-scale, e.g., 100’s of nodes, 1 Gbps links.
  – Open facility, researcher- and vendor neutral.

• Lack of effective **tools and methodologies** to foster rigorous and effective evaluation
  – Repository of test data, test topologies, traffic, and metrics.
  – Foster good science in cyber-security area.
DETER Testbed: Attributes

• Shareable – concurrent “logical testbeds”
• Isolation of “logical testbeds” (experiments)
• Flexible – full access to bare hardware
• Rapidly configurable
• Remotely accessible to (non-local experimenters)
• Security, and safety for “risky” experiments
DETER Testbed - Cluster Architecture

- **Internet**
- **External Firewall**
- **User Server**
  - User Acct & Data logging server
  - Node Serial Line Server
  - Control Network VLAN
- **Router/Isolator**
  - Firewall
  - VLAN control
- **Boss Server**
  - Web/DB/SNMP, switch mgmt
  - Power Serial Line Server
- **Node**
  - N X 4 @1000bT Data ports
- **Programmable Patch Panel (VLAN switches)**
- **Control DB**
- **Users**

- **VLAN**
  - User VLAN
  - Control Hardware VLAN
  - Boss VLAN
  - External VLAN

- **Switch Control Interface**
DETER Testbed – Inter-cluster Architecture

Testbed comprises:
- A number of clusters
- Interconnected by secure tunnels, using
- Shared or separate physical networks
**Experiment Methodology and the SEER Facility**

- *Palettes* capture high-level “design patterns” for well-formed experiments: Topology, Background and Attack Traffic, and Packet Capture and Instrumentation. Skeleton palettes for original and customized experiments are also available.

- *Methodology Engine* frames standard, systematic questions that guide an experimenter in selecting and combining the right elements.

- *Experiment Automation* increases repeatability and efficiency by integrating the process to the DETER testbed environment.
SPARTA DDoS Experiment

Background Traffic:
REPLAY | NTCG | HARPOON
HIGH FIDELITY TRAFFIC

Topology:
BUILDING-BLOCKS |
JUNIPER ROUTER CORE
REALISTIC CONNECTIVITY AND SCALE-DOWN

Attack Traffic:
DETER-INTEGRATED ATTACK SCRIPTING
AUTOMATION OF VARIETY OF SCENARIOS UNDER STUDY

Instrumentation:
PACKET AND HOST STATISTICS CAPTURE | SPECTRAL ANALYSIS |
METRICS CALCULATION | INTEGRATED VISUALIZATION
TOOLBOX FOR RIGOROUS INVESTIGATION OF RESULTS
DETER Projects

- DoS
- Worms and malware
- Overlays, routing, replic.
- Hw, sw and netw. test
- Traceback and attribution
- Models, policies
- Classes
- Diagnosis and recovery
- Multicast, group comm.
- Collaborative security
- Scanning
- Authentication
- DNS
- Spam
- Spoofing
- Botnets
- Wireless
DETER Experimenters and User Organizations (representative)

Agnik  
BBN Technologies  
Bell Labs  
Boeing Phantom Works  
Columbia University  
Cornell University  
Cs3 Inc.  
Dalhousie University  
EADS Innovation Works  
Federated Investors  
Flux Group, University of Utah  
Georgia State University  
George Mason University  
IIP Labs  
ICSI / LBNL  
Indian Institute of Technology, Delhi  
Intel Corporation  
Information Sciences Institute  
Intruguard Devices, Inc.  
IRTT  
Juniper  
Lehigh University  

McAfee Research  
National Cyber-Forensics and Training Alliance  
National Institute for Communication Technology, Hokuriku Research Center, Japan  
Naval Postgraduate School  
Network Associates Laboratories  
New Jersey Institute of Technology  
Norfolk State University  
Penn State University  
Princeton University  
Purdue University  
Rutgers University  
Sandia National Laboratories  
Secure64 Software Corp  
SPARTA, Inc.  
SRI International  
Telcordia Technologies  
Technical University Berlin  
The Aerospace Corporation  
The SANS Institute  
The Steganography Analysis and Research Center  

UC Berkeley  
UC Davis  
UC Irvine  
UC Santa Cruz  
UC San Diego  
Universita di Pisa  
University of North Carolina at Chapel Hill  
University of North Carolina at Charlotte  
University of Delaware  
University of Illinois, Urbana-Champaign  
University of Maryland  
University of Massachusetts  
University of Southern California  
University of Texas at Arlington  
University of Texas at Austin  
University of Wisconsin, Madison  
Warrior LLC  
Washington State University  
Washington University in St. Louis  
Western Michigan University
Beyond the Facility -
Creating New Tools for Cyberscience
Federation -
Scale, realism, and decentralization
Dynamic Federation

• *On-demand* creation of experiments spanning *multiple, independently controlled* facilities

• *Why?*
  – Scale
  – Unusual facilities
  – Data & knowledge sharing
  – Information hiding - multiparty scenarios
Elements of Federation

- **Experiment decomposition**
  - Creation and embedding of per-federant sub-experiments

- **Trust, policy, and security analysis mechanisms**
  - Manage federated resources within local policies

- **Provide unified runtime environment to researcher and the experiment**
  - Shared file system, scheduling and event system, control hooks, etc.
  - Failure management model

- **Secure implementation mechanisms**
  - Inter-facility communication, namespace mapping, etc.
Authorization for Dynamic Federated Testbed Environments

• Decentralized, collaborative/competitive environment. Alliances form/break frequently
  – Semantics appropriate for testbed federation
• Explicit, visible decision making
  – Corollary: clear auditing and understanding
• Multiple trust creation models, independent of mechanism
  – Examples: Hierarchical PKI, PGP web of trust, etc.
• Minimize unnecessary communication
  – For disconnected operation
Federation: Status

• Operating today
  – Hundreds of nodes
  – Emulab-based testbeds {DETER, WAIL, Emulab}
  – Simple / simplistic ID based authorization
  – Scale-limited: primarily by file system and event model

• Extending to
  – Full attribute-based auth/auth framework
  – Removal of scaling limitations
  – Richer resource discovery / integration with higher level tools
  – Additional testbed architectures
    • NSF GENI program
    • DRAGON
Experiment Health –
Structured experiments and scientific rigor
Maintaining Experiment Invariants: The DETER Experiment Health System

- Uses *higher level* knowledge about the experiment – the desired invariants
- Takes corrective or notification action if invariant is violated
  - Monitor invariants..
  - Trigger actions
- Captures invariants in exportable form for experiment reuse, repeatability and validation, etc.
Experiment Health: Benefits

• **Sophisticated users** benefit in case of large-scale or batch experiments
  – Very difficult to monitor these manually or to discover problems in a large result set
• **Novice users** always benefit, but especially when re-running somebody else’s code (e.g., students)
  – Early signal that something is wrong motivates understanding of experimentation mechanism
Experiment Health System: Design Goals

- Support an open-ended set of expectations
  - Internal (traffic generator output)
  - External (verifying node status)
- On expectation failure, execute arbitrary recovery actions
  - Automatic adaptation, repair, terminate
  - Alert user for manual intervention
- Support both low-level and high-level sources of knowledge about invariants
Continuing Development

• Develop intelligent expectation build engine
  – Reduce user burden and help novice users
  – Infer expectations from user actions and observed outcomes
  – Refine as the experiment is repeated
  – Offer the inferred scenario to user to label and modify
  – Potential to build a timeline of events/measurements

• Structured expectation templates – combine and integrate groups of low-level expectations, support experiment classes

• Develop a library of tools, expectations and measurements for common experiment classes
  – Enable sharing between users
  – Public (populated by us initially) and private archives
Risky Experiment Management -
Experimental Cybersecurity in a Complex World
Domain of interest

- “Malware Containment” →
  “Risky Experiment Management”
  - Malware study
  - “Active Defense” research
  - Stress and failure testing
  - Agent based measurement and monitoring
  - …
Problem formulation
“Classical Formulation”

- Focus on isolation
- *De facto* emphasis of initial DETER work
Problem formulation

More accurate formulation

- Key issue:
  - Not “isolation and containment” but “understanding and assurance”
**Model**

- **Two-stage approach:**

```
Behavioral composition model: External behavior == T2(T1(experiment))
```
Why is this a good idea?

• Simplified experiment design, increased reusability
  – Defined T1 “invariants” (experiment constraints) simplify design of experiments with known external properties

• Separation of concerns:
  – Experimenter best equipped to understand impact of constraints on experiment validity - express T1 for experimenters
  – Testbed designer best equipped to understand impact of constraints on testbed and external behavior - express T2 for testbed designers/implementors
Near Term: Constraint sets

- Constraint sets are
  - pre-established sets of T1 constraints that,
  - when met, allow useful, well behaved, well understood experiments to be run.
- It is useful to imagine defining more than one set
Long Term: Formal Verification

• It may be [in some cases, “is”] possible to reason formally about the overall behavior of the T2(T1)exp)) system.

• This might allow fine grain, possibly automatic derivation of experiment (T1) and testbed configuration (T2) constraints to limit a particular experiment’s potential external behavior without damaging its experimental value

• Such a tool would provide a highly general facility for limiting the risk of risky experiments
A Common Theme

• Interesting to note …
• ... underlying concept of establishing, understanding, monitoring, enforcing experiment/test invariants is *common* to
  – Experiment Health
  – Risky Experiment Management
  – Intelligent, high-level design of experiments
• Direction: *unified, shared mechanisms* for invariant management to serve these goals
Summary
DETER Cyber Security Testbed

• Facility
  – Tool libraries, Workbench, Operations, Community

• Tools for Cyber Science
  – Federation, Experiment Health, Risky Experiment

• Foundation for next generation testbeds
  – Geni, National Cyber Range

• More info –
  www.isi.edu/deter or www.deterlab.net