Terry Benzel
USC Information Sciences Institute
December 9, 2011
Annual Computer Security Applications Conference

The Science of Cyber Security Experimentation: The DETER Project

Reality

Large, Complex, Interconnected

Slow to evolve

Legacy Subsytems

System of Systems

Online Payment Services

Online payments is a free online service from Wells Fargo F

It allows you to initiate payments from your checking or sav financial institution. If you have already enrolled, simply sig User ID and password. If you are not already enrolled, you may begin using this service by selecting Enroll Now.

* indicates a required field

Connected Cyber
Physical Systems

Reality – The Dark Side

Weapons evolve **rapidly** and proliferate **widely**

Asymmetric warfare:

- Attacks from anywhere, with unknown weapons
- Defenses must be known, **effective**, **affordable**

Faster and Faster Does Not Work

"The Red Queen has to run faster and faster in order to keep still where she is. That is exactly what you all are doing!"

What Can We do About It?

- Solution – build less vulnerable systems to begin with!
- Create fundamental understanding and reason about systems through experimental means
- Key aspect – enable science based experimentation
- Hard Problem

All Too Often

Why There is No Science in Cyber Science
[A panel discussion at NSPW 2010] Maxion, Longstaff, McHugh

1. Have an idea for a “new” tool that would “help” security

2. Program/assemble the tool (the majority of the work)

3. Put it on your local net

4. Attack your system

5. Show the tool repels the attack

6. Write up “the results” and open-source the tool

7. (optional) Start up a company which might succeed

Instead - Objectives

- Perform experimental research of scale and complexity sufficient to the real world
- Extract understanding through experimental research
- Collect, leverage, and share experimental artifacts and learnings

Cyber Security Experimentation

- Class of experimental cyber science applied to sets of problems - networked cyber systems and often cyber physical networked systems

- Goal - enable experimental cyber science aimed at study of behavior, phenomena, providing fundamental understanding

The DETER Project

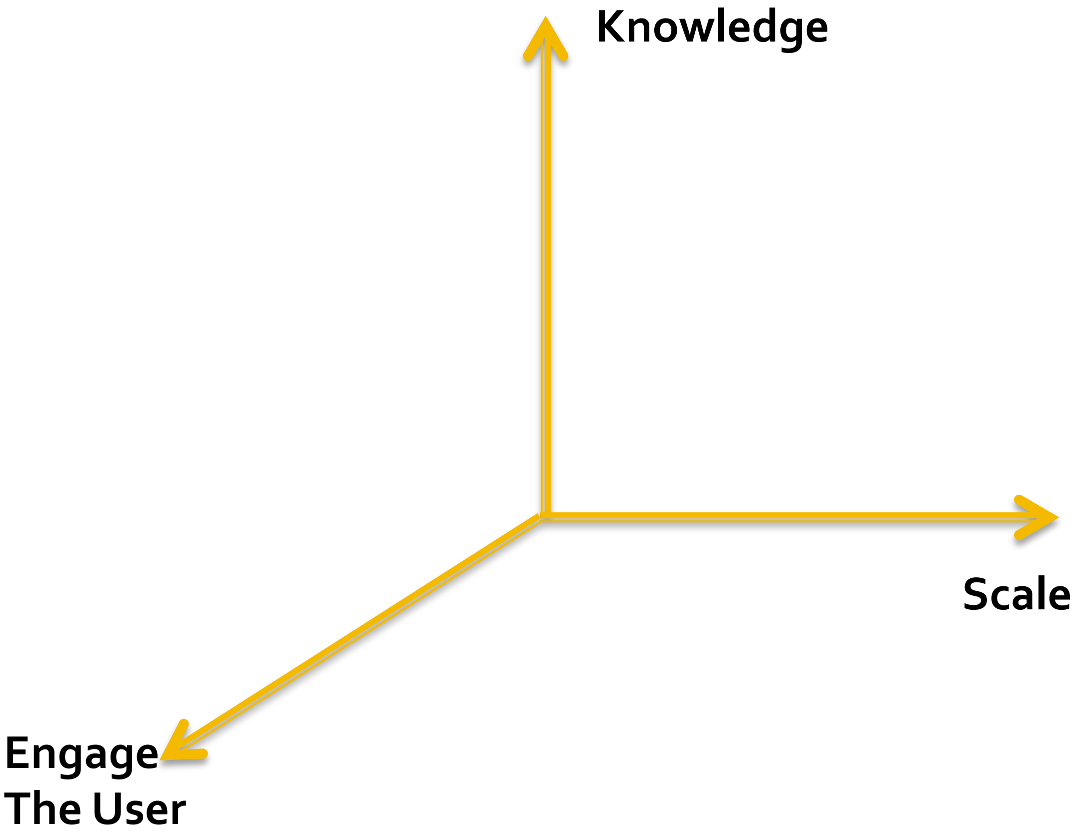
- A research program:
  - To advance capabilities for experimental cybersecurity research
- A testbed facility:
  - To serve as a publicly available national resource…
  - …supporting a broad base of users and experiments
  - … and act as a technology transfer and evangelization vehicle for our and others' research in experimental methodology
- A community building activity:
  - To foster and support collaborative science…
  - …effective and efficient leverage and sharing of knowledge

RESEARCH PROGRAM

Research Goals

- Advance our understanding of of experimental cybersecurity *science and methodologies*
  - Enable new levels of rigor and repeatability
  - Transform low level results to high level understanding
  - Broaden the domains of applicability
- Advance the *technology of experimental infrastructure*
  - Develop technologies with new levels of function, applicability, and scale
- Share knowledge, results, and operational capability
  - Facility, data and tools
  - Community and knowledge

Three Axis



Knowledge

Scale

Engage
The User

Research

Knowledge

Higher Knowledge and Semantics

- The problem:
  - Today’s testbed technologies understand the *syntax* of experiments, but have no awareness of higher level knowledge or *semantics.*

- The challenge:
  - Incorporate higher level, semantic information about experiments and scenarios into our systems and tools, and
  - Use this knowledge to improve research quality and understanding.

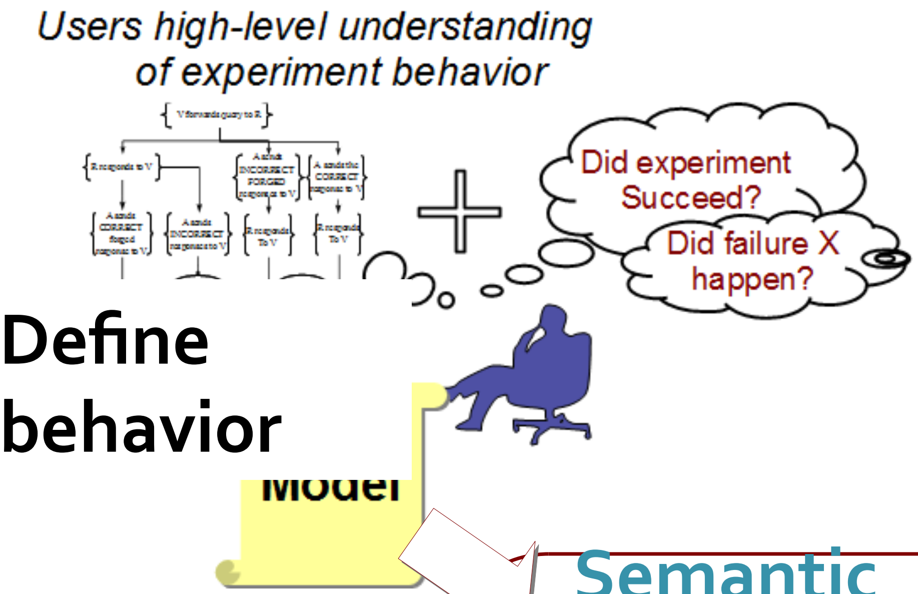
Using Knowledge

- Uses *higher level knowledge* about the scenario
  - Required *invariants* (things that **must** be true for the experiment to be valid)
  - Expected behavior
- Takes *corrective or notification* action if invariant is violated
  - Monitor invariants
  - Trigger actions

Invariants

- Captures invariants in *explicit form* for experiment reuse, repeatability and validation, etc.
- Must be true for experiment to be valid
- High level testing of invariants –
  - Understanding against data sets
  - Against constraints/invariants
- Also questions of modeling and scale –
  - Researcher intuition expressed as checkable invariants
- Specification for sharing

Define behavior



Models drive visualization over data.

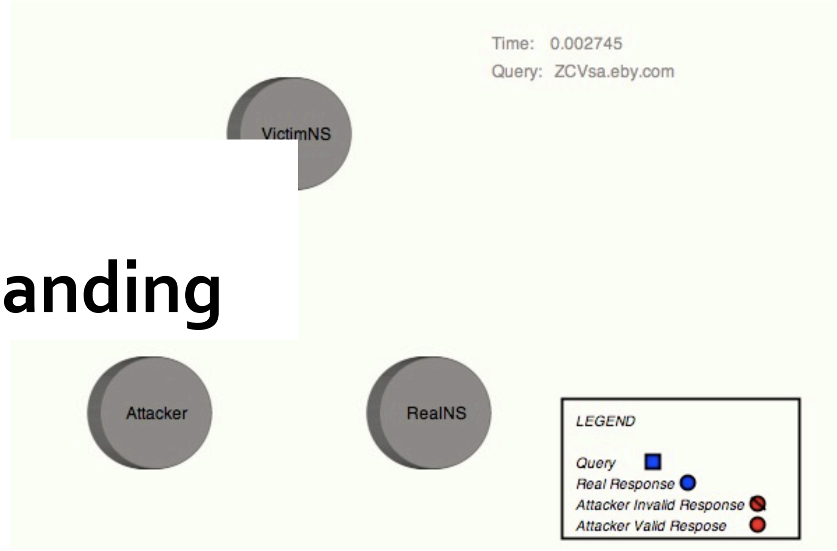
Semantic Analysis Framework

Test it on data

Packet dumps
gs

Experiment data is input as normalized events.

Gain Understanding



Data Analysis in Networked Systems

Alerts
Audit Logs
Is my hypothesis validated?
Did my experiment run as expected?
Why did failure X happen?
Is there any evidence of a known attack?
Packet Dumps
Application Logs

Moving beyond bits – Capturing and Reasoning about Scenarios

Scenarios are captured by

- *Environment* – the conditions of the scenario
  - Virtual topology (varies with phenomenon), could be dynamic, abstract, expresses needs and constraints
  - Traffic, cross-traffic, cross-events, human actions, etc.
- *Workflow* – Occurrences and events of interest
- *Invariants* – truths that must hold for correctness

Research

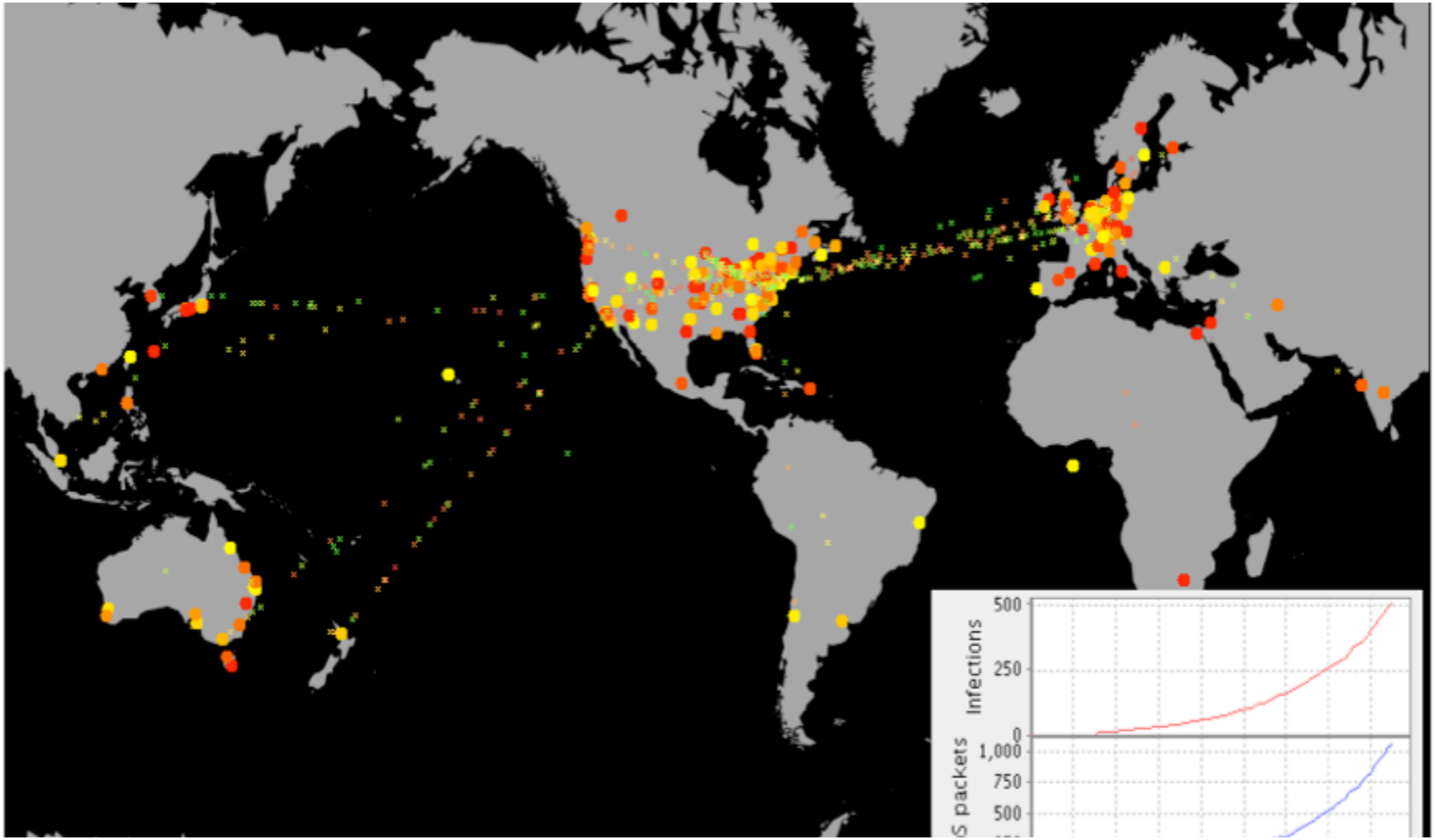
Scalable

Scalable Modeling and Emulation

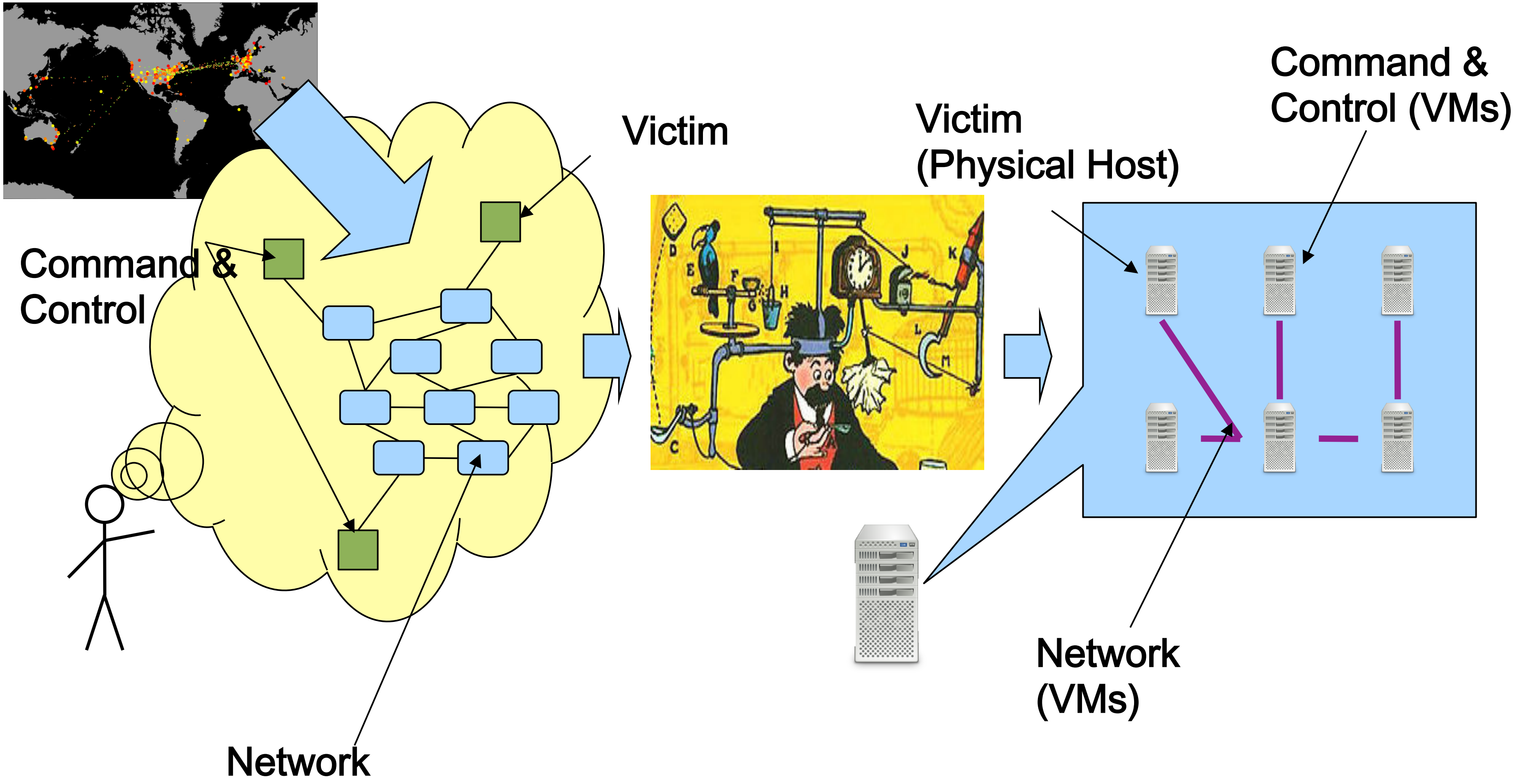
- The problem:
  - Traditional testbeds can model and emulate *small* systems at a *fixed* level of fidelity.

- The challenge:
  - Many real problems require modeling of *large, complex* systems at an *appropriate* ("good enough") level of fidelity.
  - That level may be *different* for different parts of the modeled system.
  - Think of this as "smearing the computation power around to just where it's needed".

100 K-Node Worm/Botnet/DDOS Scenario



Realization of Scenario



Command & Control

Victim

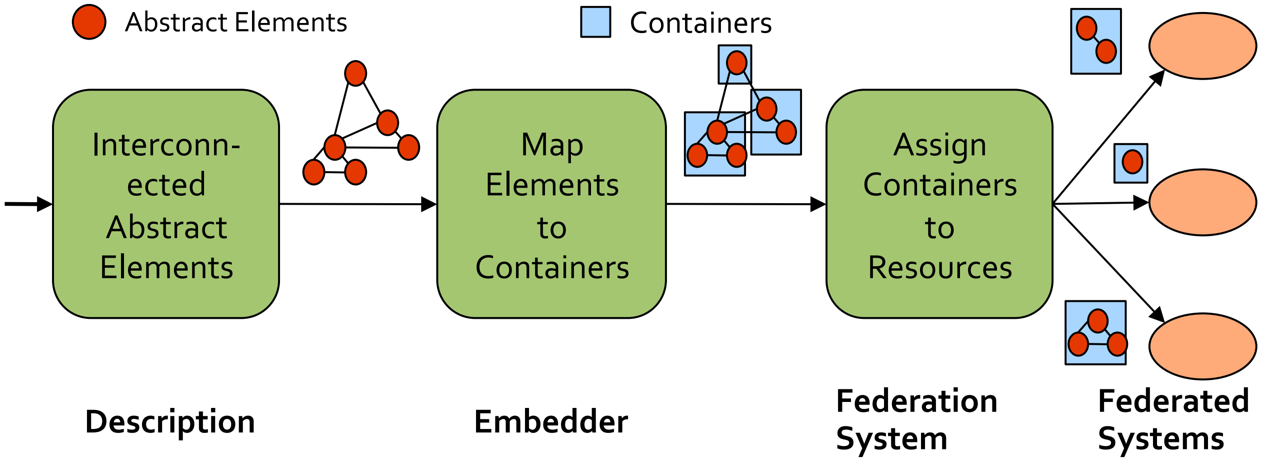
Network

Victim (Physical Host)

Command & Control (VMs)

Network (VMs)

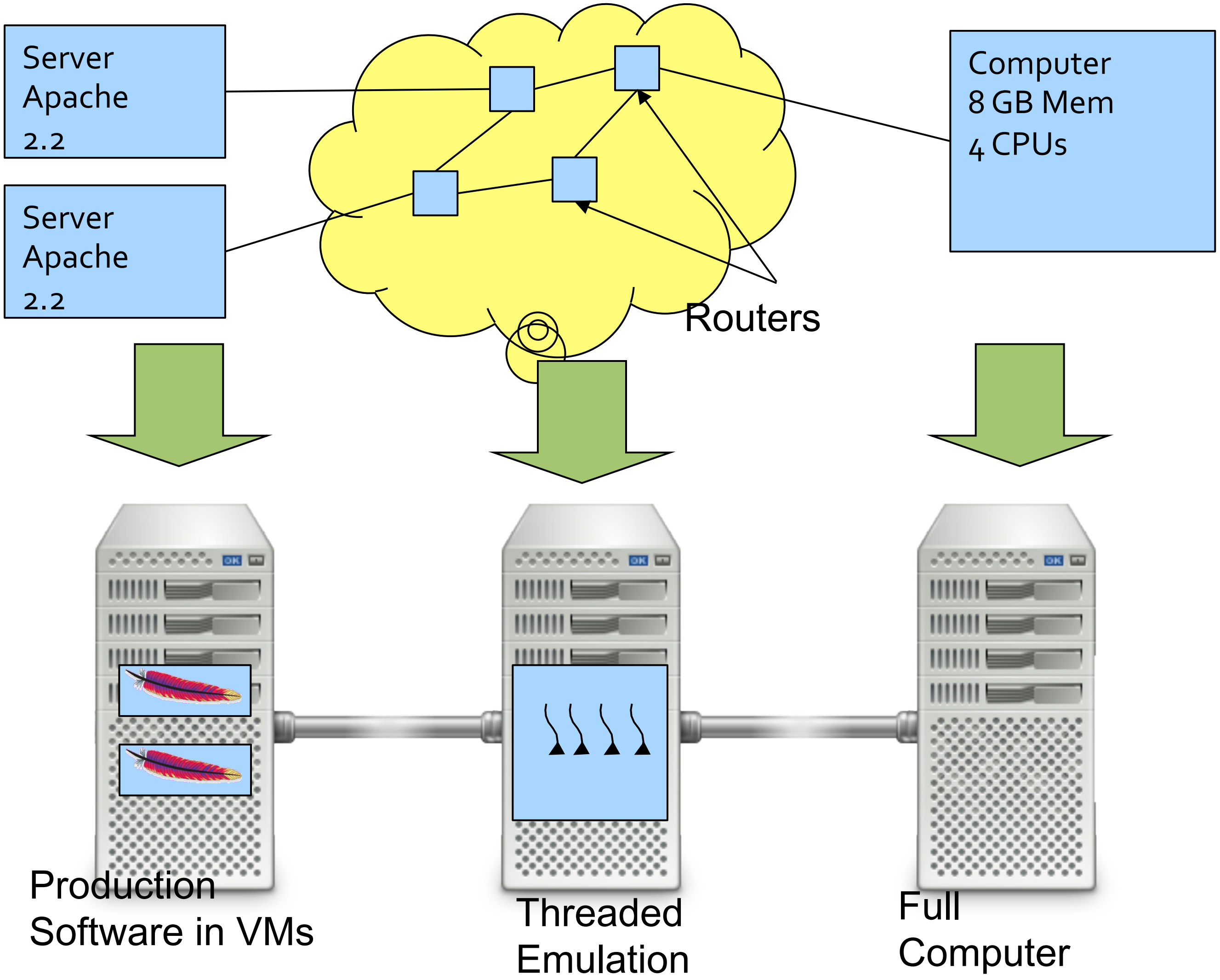
Domain-Aware, Scalable Virtualization and Embedding



● Abstract Elements
□ Containers
Interconn-ected Abstract Elements
Map Elements to Containers
Assign Containers to Resources
Description
Embedder
Federation System
Federated Systems

- Abstract the “node” concept to multiple classes of containers
- Support wide range of scalability-fidelity tradeoffs
  - Apply computational resources to <u>key dimensions</u> for <u>specific</u> problem space

Containerization: Abstraction To Realization



Server
Apache
2.2

Server
Apache
2.2

Routers

Computer
8 GB Mem
4 CPUs

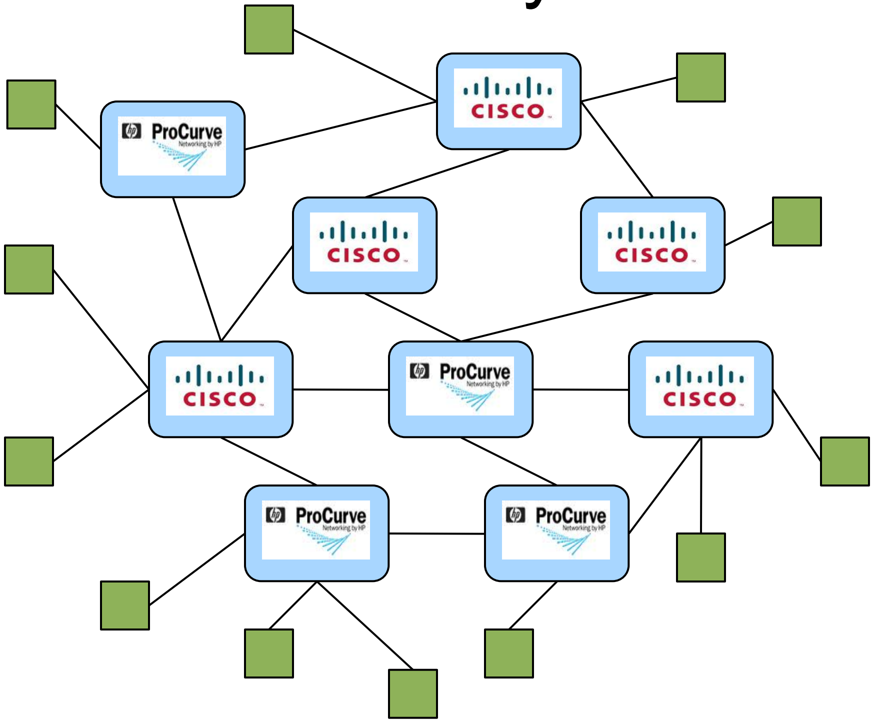
Production Software in VMs

Threaded Emulation

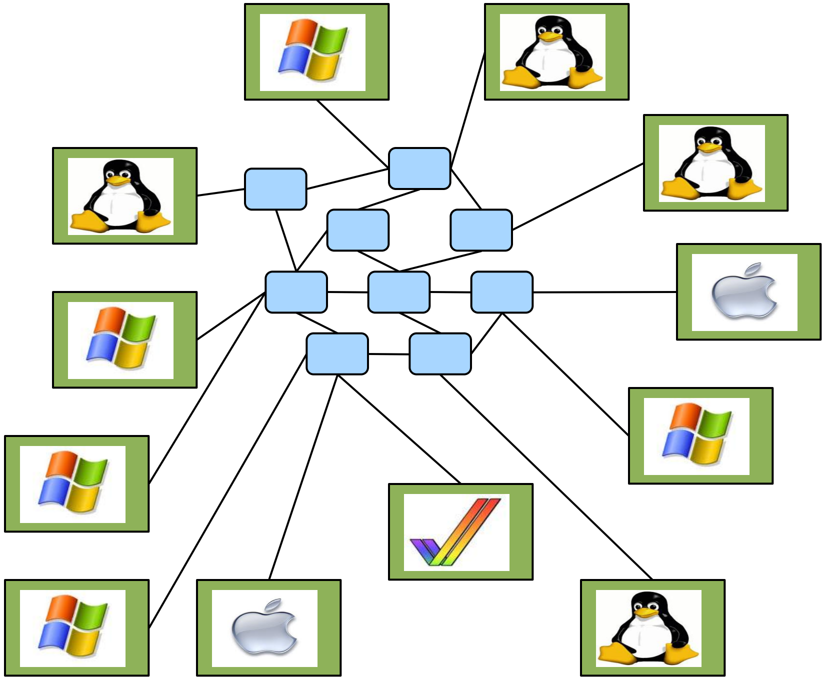
Full Computer

Different Scenarios: Different Abstractions

BGP Security



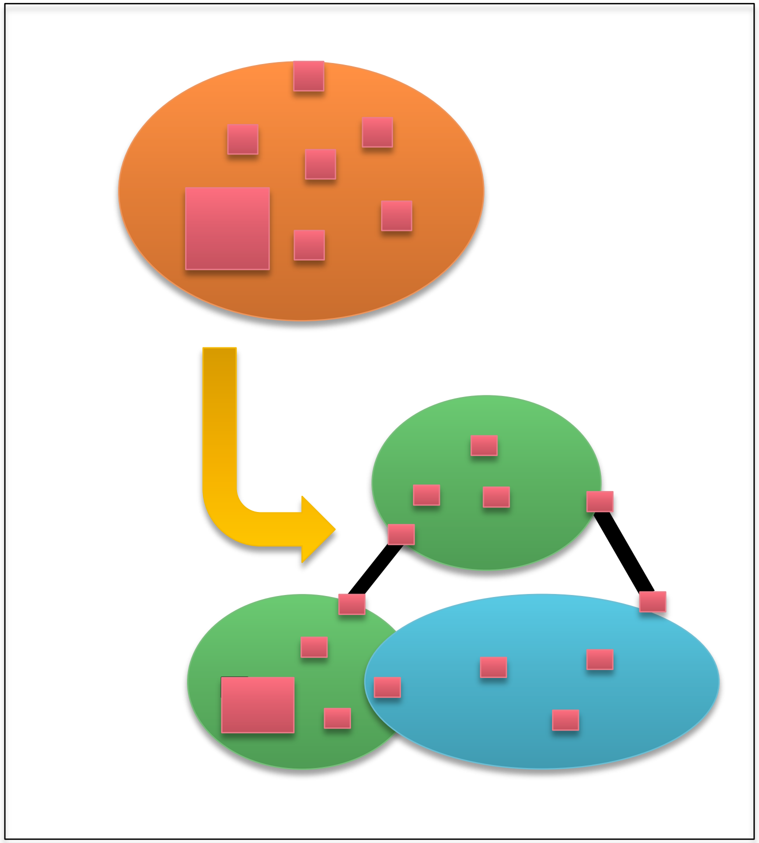
Worm Propagation



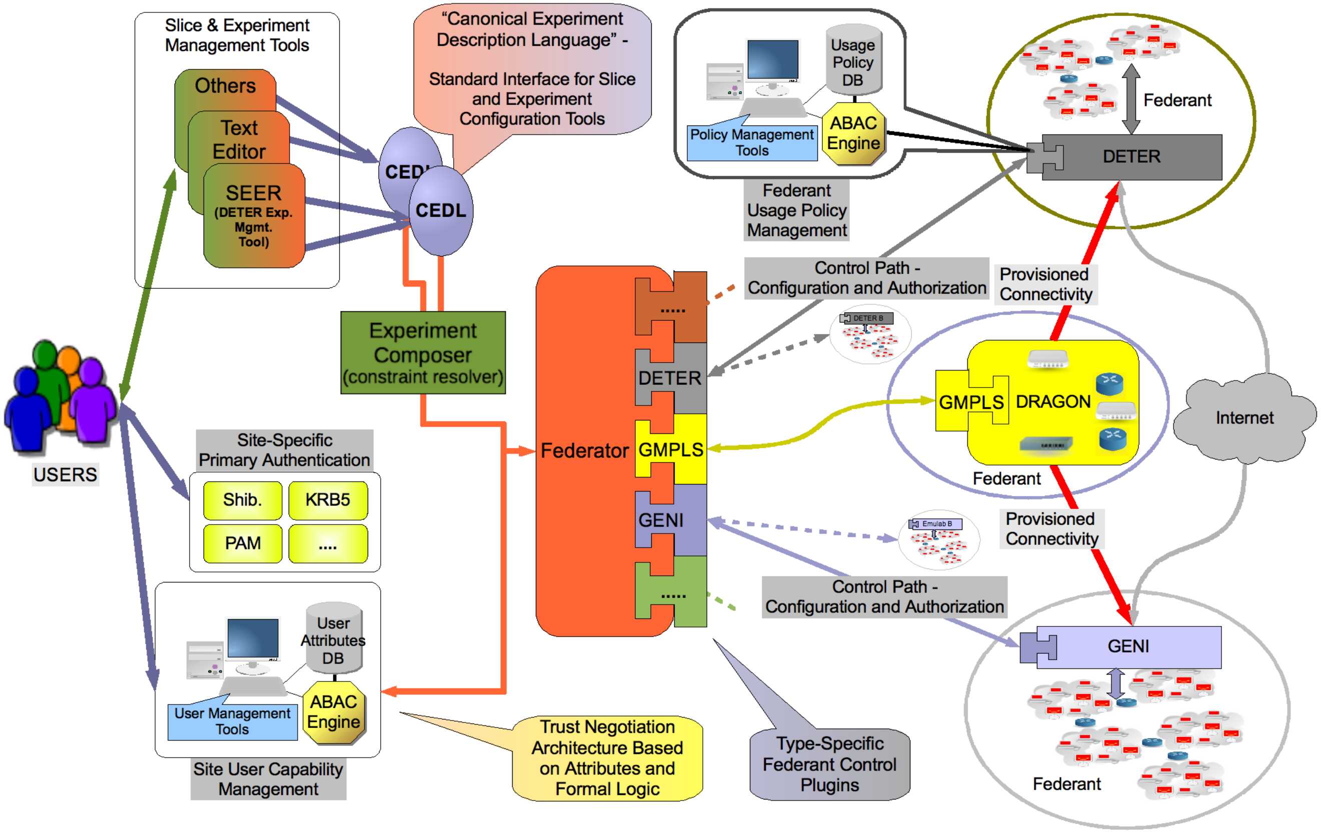
Dynamic Federation

- On-demand creation of experimental scenarios spanning *multiple, independently controlled* facilities

- Goals and Benefits
  - Scale
  - Access to unique resources
  - Accommodation of usage policy constraints
  - Data & knowledge sharing
  - Information hiding



Picture: the DETER Federation Architecture – mid-2010 version – http://fedd.isi.deterlab.net


Slice & Experiment Management Tools
Others
Text Editor
SEER (DETER Exp. Mgmt. Tool)
"Canonical Experiment Description Language" - Standard Interface for Slice and Experiment Configuration Tools
CEDL
CEDL
Usage Policy DB
Policy Management Tools
ABAC Engine
Federant Usage Policy Management
Federant
DETER
Control Path - Configuration and Authorization
Provisioned Connectivity
Experiment Composer (constraint resolver)
.....
DETER
Federator
GMPLS
GENI
.....
DETER B
GMPLS
DRAGON
Federant
Internet
USERS
Site-Specific Primary Authentication
Shib.
KRB5
PAM
....
Emulab B
Provisioned Connectivity
Control Path - Configuration and Authorization
GENI
User Attributes DB
User Management Tools
ABAC Engine
Site User Capability Management
Trust Negotiation Architecture Based on Attributes and Formal Logic
Type-Specific Federant Control Plugins
Federant

- Scenario Description
- Resource Description
- Constraint Resolution

- Embedding
- Planning
- Sequencing

- Resource Control
- Policy, Authentication and Authorization

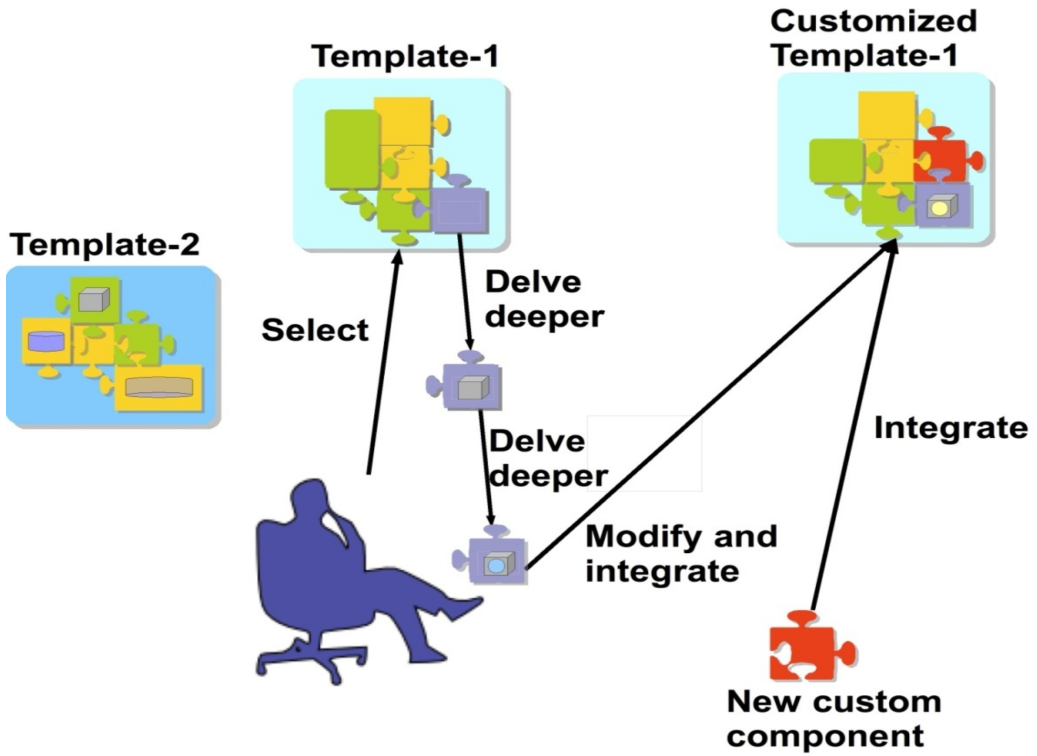
Research

Engaging the User

Engaging The User

- The problem:
  - Today’s testbed technologies provide limited support for complex user tasks, thus, hampering system of system level experimentation and reasoning.
- The challenge:
  - Develop methodologies to leverage knowledge, understanding, and semantics, through development environments, composition and sharing.

Support the Experimental Process



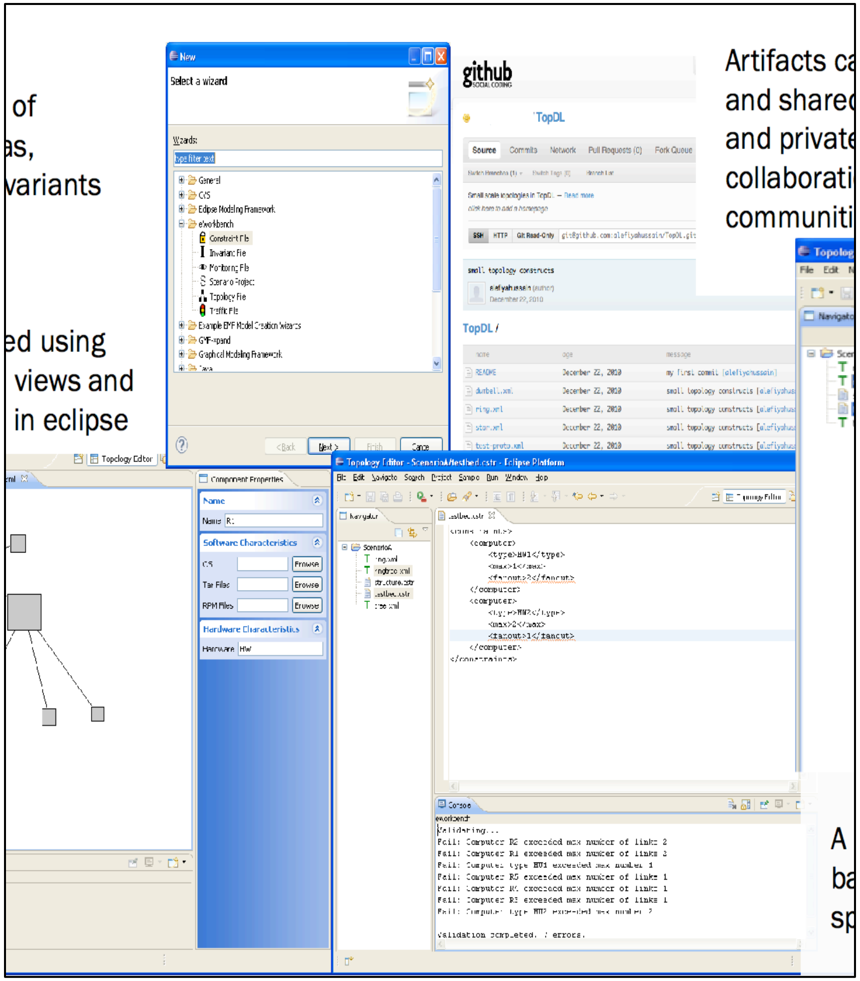
- Graduated, visual, and powerful experiments
- Domain-specific (DDoS, worm, botnet) capabilities
- Built-in sharing capabilities

Scenario Lifecycle Management

- Most testbed tools focus on *creating* and *running* an experiment. Much less attention is paid to other important steps in the process

- Develop a model for workflow over the full lifecycle of an experiment, and capture that model in methodologies and tools

Scenario Lifecycle Management



of
as,
variants

ed using
views and
in eclipse

Artifacts ca
and shared
and private
collaborati
communiti

github SOCIAL CODING

TopDL

Source Commits Network Pull Requests (0) Fork Queue

Small scale topologies in TopDL — Read more

click here to add a homepage

SSH HTTP Git Read-Only git@github.com:alefiyahussain/TopDL.git

small topology constructs

alefiyahussain (author)
December 22, 2010

TopDL /

name	age	message
README	December 22, 2010	my first commit [alefiyahussain]
dumbell.xml	December 22, 2010	small topology constructs [alefiyahus
ring.xml	December 22, 2010	small topology constructs [alefiyahus
star.xml	December 22, 2010	small topology constructs [alefiyahus
test-proto.xml	December 22, 2010	small topology constructs [alefiyahus

Topology Editor - ScenarioA/testbed.cstr - Eclipse Platform

File Edit Navigate Search Project Sample Run Window Help

testbed.cstr

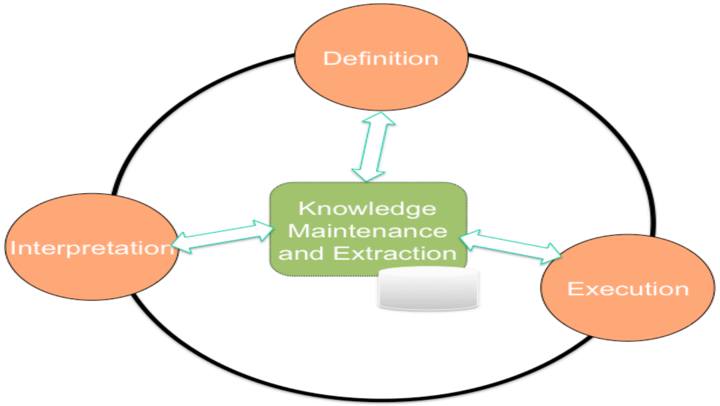
```
<constraints>
    <computer>
        <type>HW1</type>
        <max>1</max>
        <fanout>2</fanout>
    </computer>
    <computer>
        <type>HW2</type>
        <max>2</max>
        <fanout>1</fanout>
    </computer>
</constraints>
```

```
Validating...
Fail: Computer R2 exceeded max number of links 2
Fail: Computer R1 exceeded max number of links 3
Fail: Computer type HW1 exceeded max number 1
Fail: Computer R5 exceeded max number of links 1
Fail: Computer R4 exceeded max number of links 1
Fail: Computer R3 exceeded max number of links 1
Fail: Computer type HW2 exceeded max number 2

Validation completed. 7 errors.
```

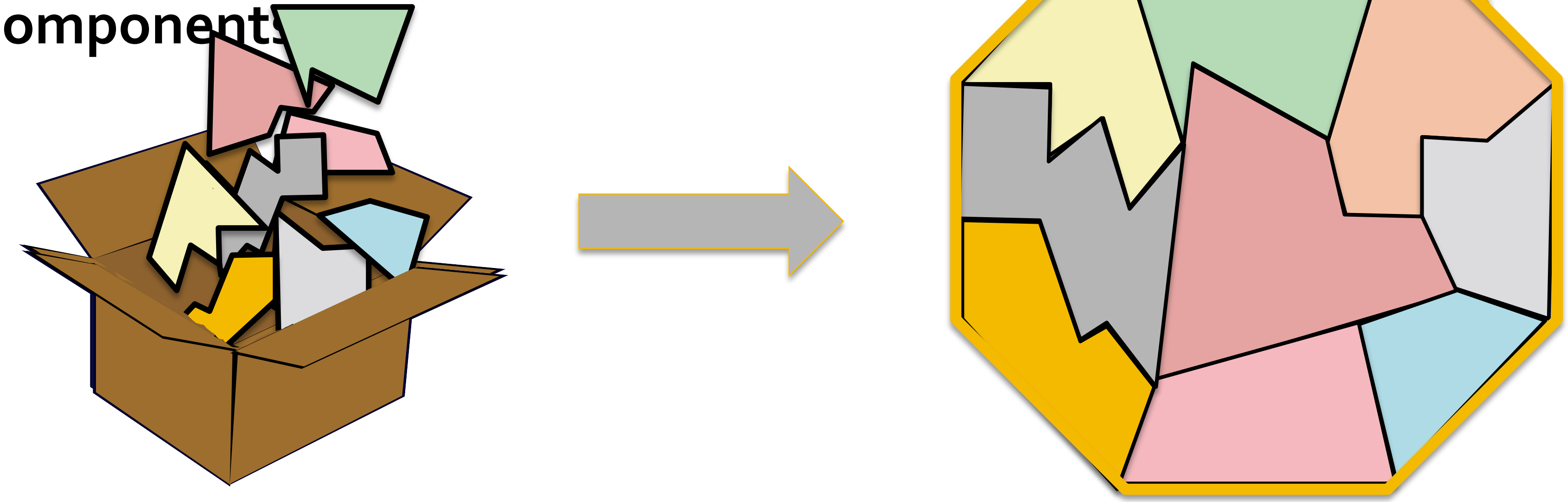
A
ba
sp

- Key Observation: isomorphism to software engineering lifecycle
- Implementation Approach: Leverage Eclipse
  - Repurpose tested SWE methodologies
  - Build on 20M+ lines of code



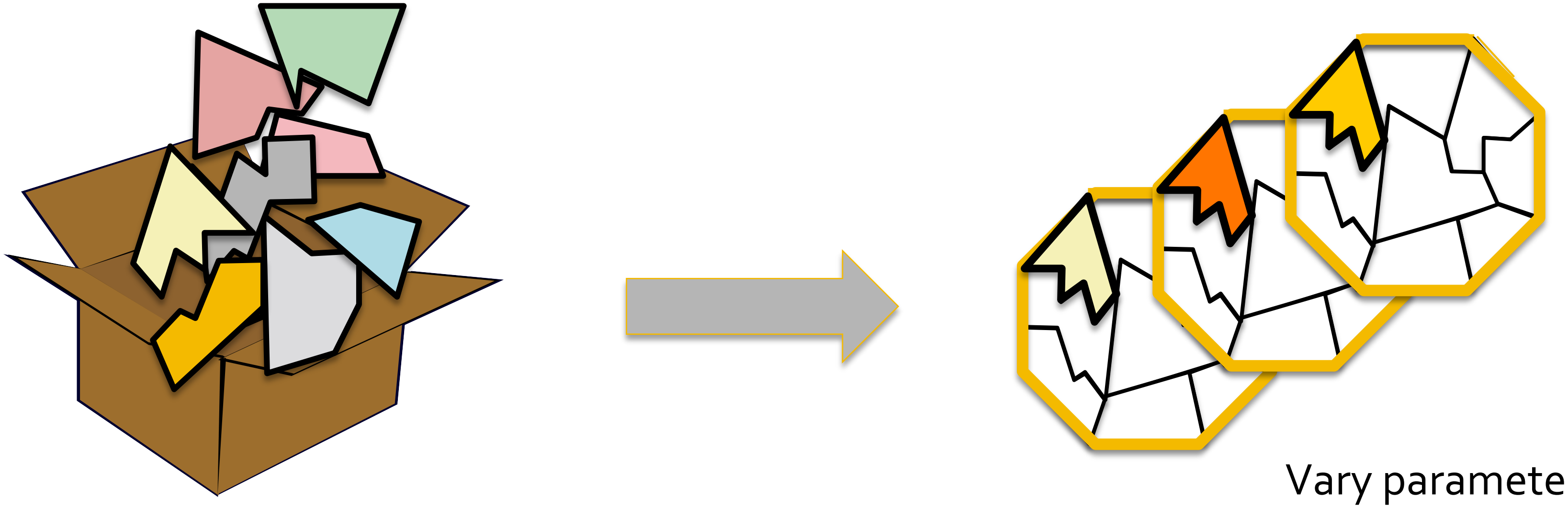
Composing a New Experiment

Repository of Reusable Components



Varying Subset of Components

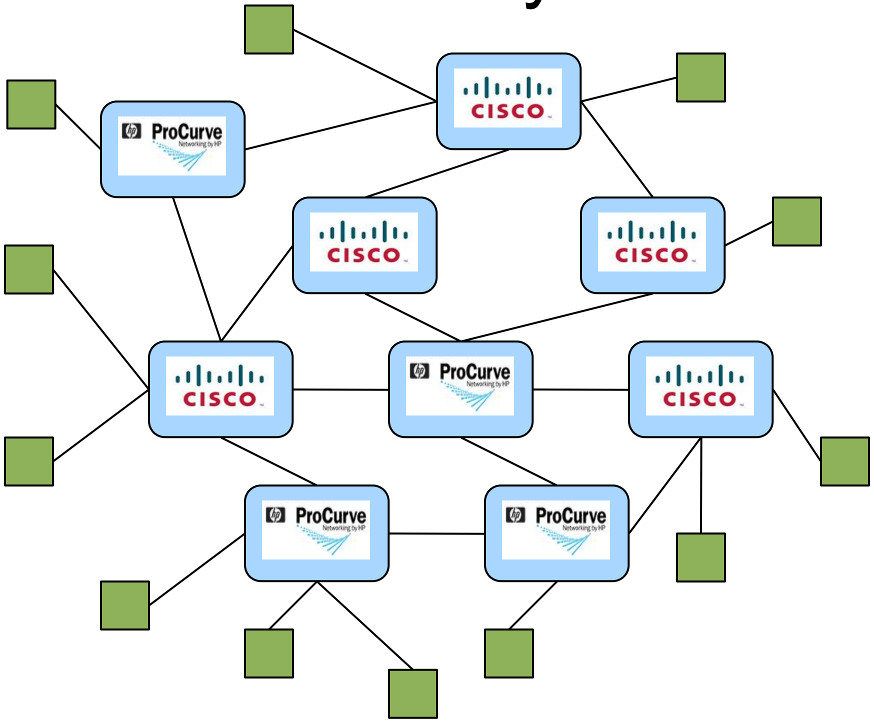
Repository of Reusable Components



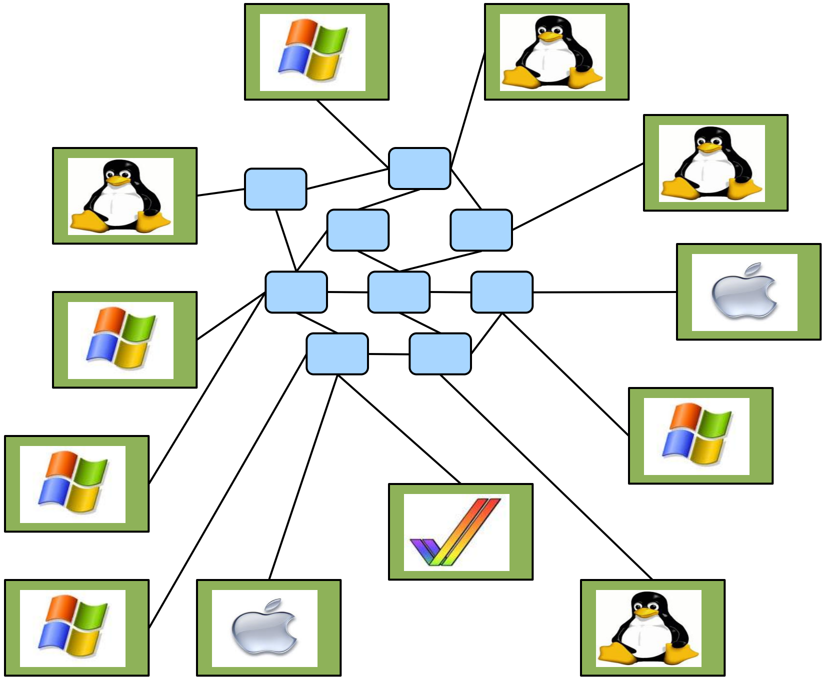
Vary parameters
per component

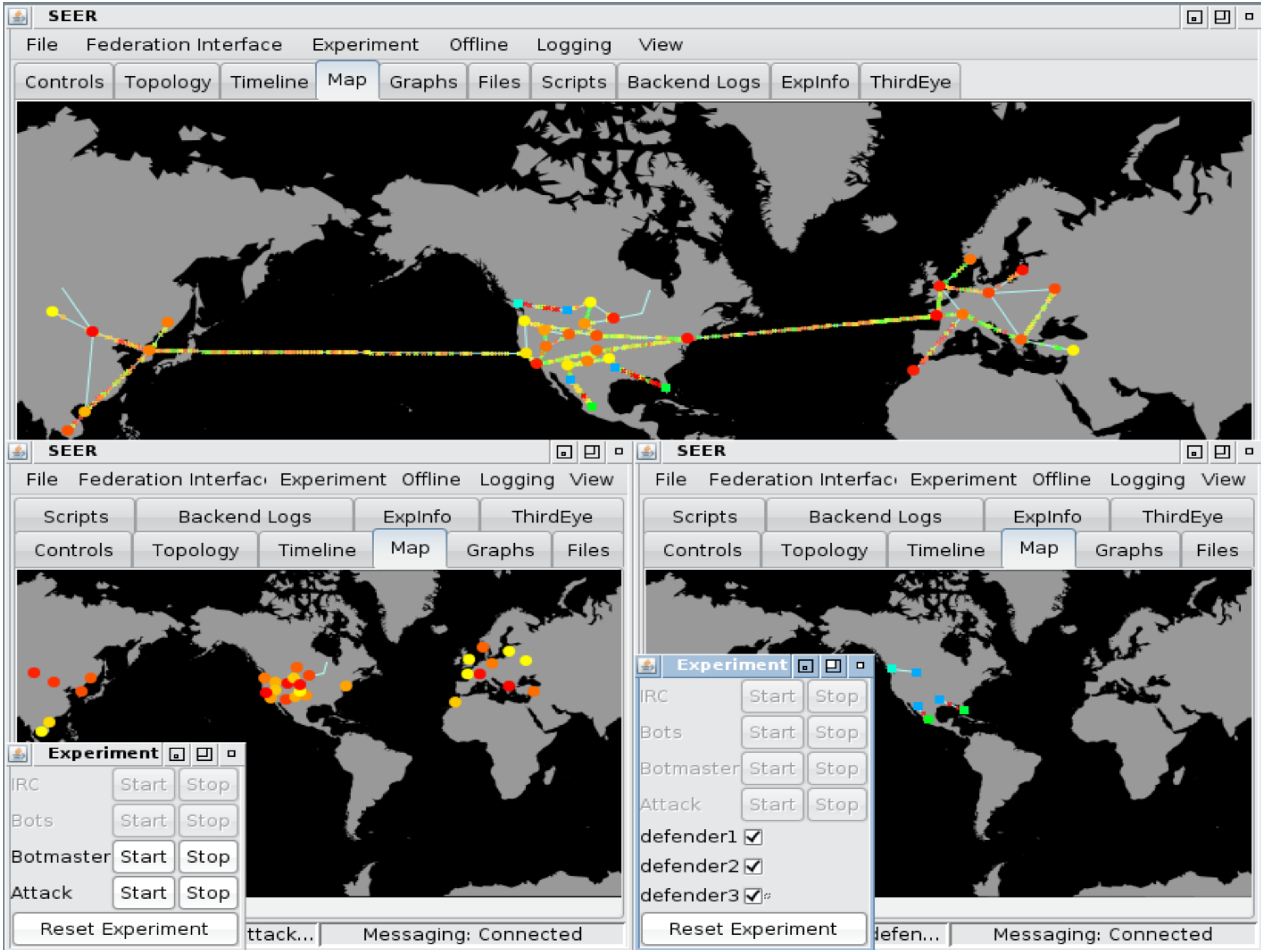
Same Components – Multiple Scenarios

BGP Security



Worm Propagation





defender1 ☑

defender2 ☑

defender3 ☑

Multi-agent system to Model Some Human Behavior

- Testbeds must model impact of human activity in repeatable experiments
  - Provide more realistic behavior for testing security tools
  - **But** real humans are expensive and non-repeatable
- Model goal-directed team activity
  - Measure impact of an attack on team goals
  - Model impact of organization structure
- Model certain human characteristics
  - Propensity to make mistakes
  - Aspects of physiology, (soon: emotion, bounded rationality)
  - Flexibility to changing conditions
- Configurable tool for experimenters

DETERLAB

The Facility

The DETER Facility

A general purpose, flexible platform for modeling, emulation, and controlled study of large, complex networked systems

- Elements located at USC/ISI (Los Angeles), UC Berkeley, and USC/ISI (Arlington, VA)
- Funded by NSF and DHS, started in 2003
- Based on Emulab software, with focus on security experimentation
- Shared resource – multiple simultaneous experiments subject to resource constraints
- Open to academic, industrial, govt researchers essentially worldwide – very lightweight approval process

Physical Platform

- ~440 PC-based nodes
  - Berkeley, CA - ~200 Nodes
  - Los Angeles, CA - 220 Nodes
  - Arlington, VA – 20 Nodes
- Interconnect (2010)
  - 1 Gb/s – LA-UCB
  - 1-10 Gb/s LA-Arlington
- Local and Remote access

Advanced Infrastructure Capabilities

High-performance co-processing

- NetFPGA-based node deployment
- Dedicated hardware modules, e.g., packet monitors

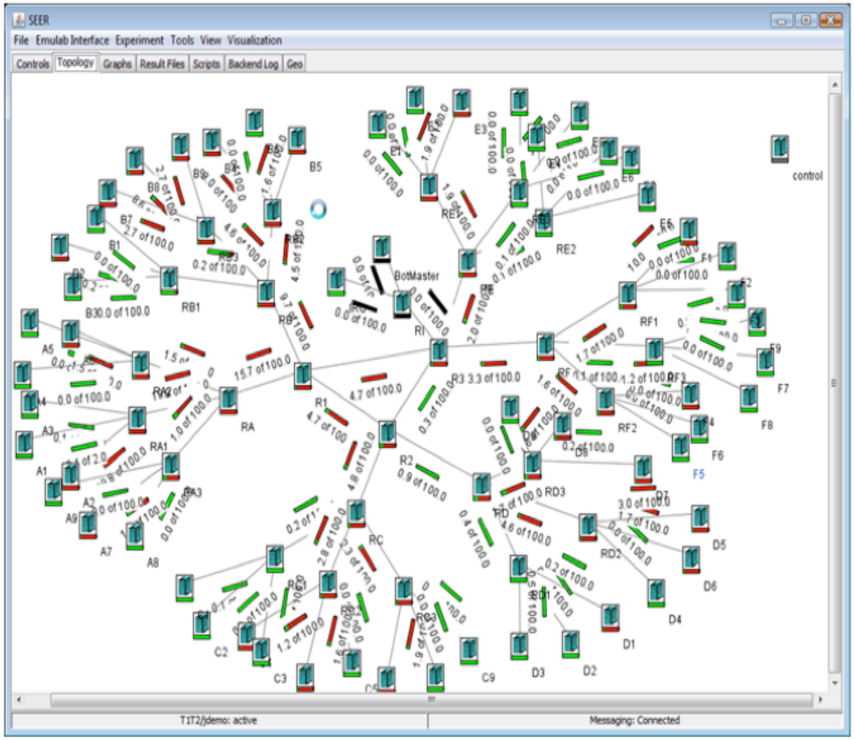
Efficiency and scalability

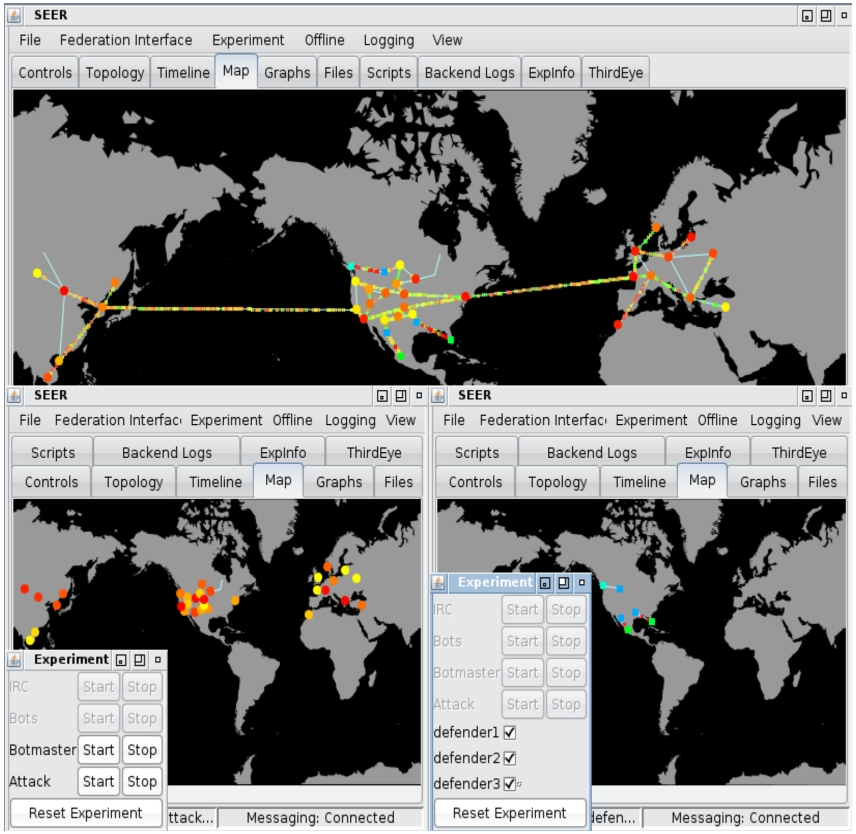
- Increased VLAN bandwidth (10Gbps +)
- Configuration management and infrastructure protection

Key Capabilities

- Technical elements
  - DETER Core
  - Scalable Modeling and Emulation
  - Federation
  - Leveraging Understanding and Semantics
  - Risky Experiment Management
  - Multiparty Experiments
  - Experiment Lifecycle Management

Control, Analysis, and Visualization Interfaces and Tools


SEER
File Emulab Interface Experiment Tools View Visualization
Controls Topology Graphs Result Files Scripts Backend Log Geo
control
BotMaster
RA
RB1
RC
RD
RE2
RF1
RF2
RD3
B5
A1
A7
A8
A9
C2
C3
C9
D1
D2
D3
D4
D5
D6
F5
F6
F7
F8
F9
T1T2/jdemo: active
Messaging: Connected


SEER
File Federation Interface Experiment Offline Logging View
Controls Topology Timeline Map Graphs Files Scripts Backend Logs ExpInfo ThirdEye
SEER
File Federation Interfac Experiment Offline Logging View
Scripts Backend Logs ExpInfo ThirdEye
Controls Topology Timeline Map Graphs Files
Experiment
IRC Start Stop
Bots Start Stop
Botmaster Start Stop
Attack Start Stop
Reset Experiment
ttack...
Messaging: Connected
SEER
File Federation Interfac Experiment Offline Logging View
Scripts Backend Logs ExpInfo ThirdEye
Controls Topology Timeline Map Graphs Files
Experiment
IRC Start Stop
Bots Start Stop
Botmaster Start Stop
Attack Start Stop
defender1 ☑
defender2 ☑
defender3 ☑
Reset Experiment
lefen...
Messaging: Connected

The DETER Community

Community and Outreach

- Content sharing support
  - Experiments, data, models, recipes
  - Class materials, recent research results, ideas
- Shared spaces
  - Outreach: Conferences, tutorials, presentations
  - Share and connect: Website, exchange server
  - Common experiment description: Templates
  - Build community knowledge: domain-specific communities
- Education support
  - NSF CCLI grant: develop hands-on exercises for classes
  - Moodle server for classes on DETER

DETER User Institutions

Government

Air Force Research Laboratory
DARPA
Lawrence Berkeley National Lab
Naval Postgraduate School
Sandia National Laboratories

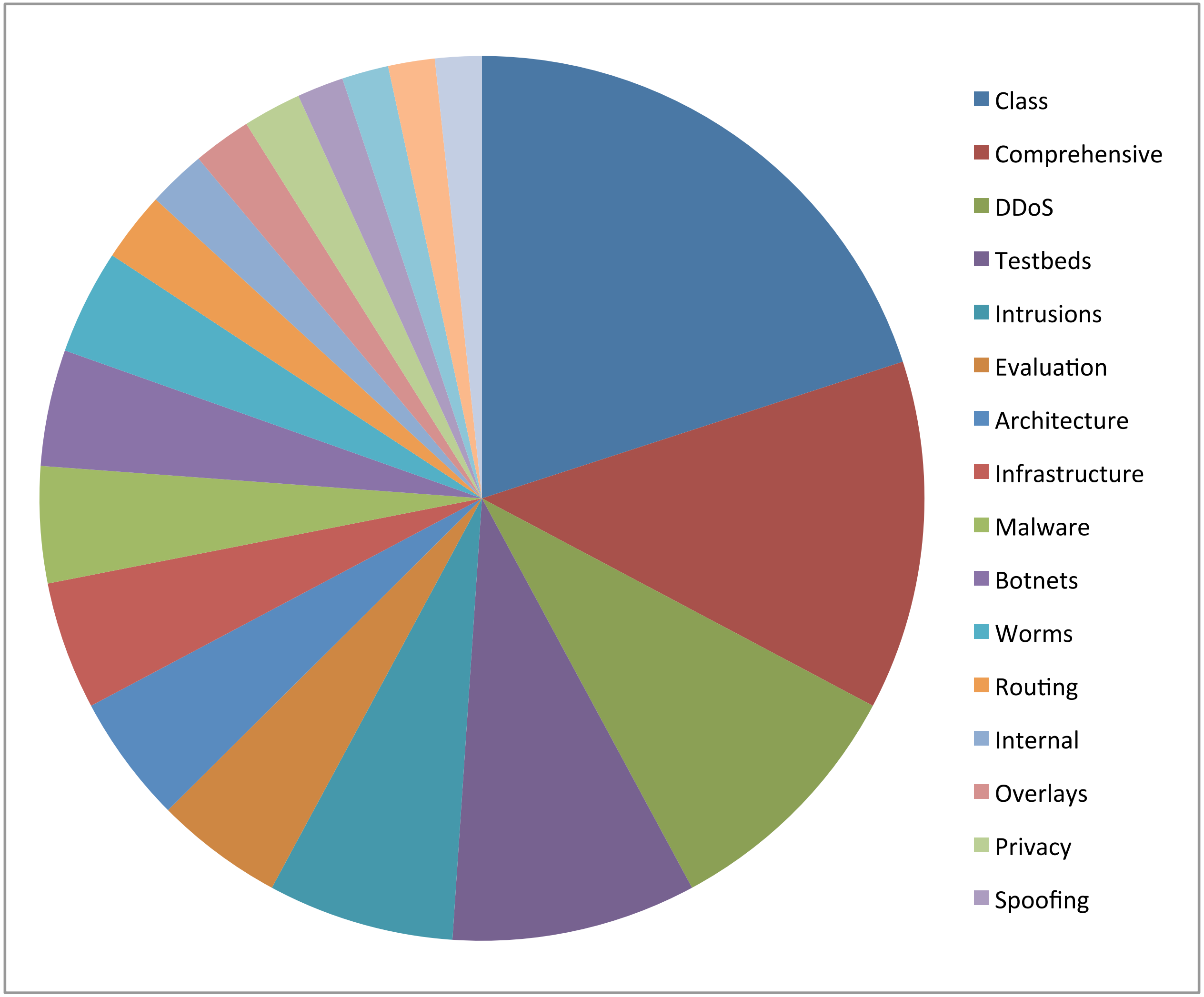
Industry

Agnik, LLC
Aerospace Corporation
Backbone Security
BAE Systems, Inc.
BBN
Bell Labs
Cs3 Inc.
Distributed Infinity Inc.
EADS Innovation Works
FreeBSD Foundation
iCAST
Institute for Information Industry
Intel Research Berkeley
IntruGuard Devices, Inc.
Purple Streak
Secure64 Software Corp
Skaion Corporation
SPARTA
SRI International
Telcordia Technologies

Academia

Carnegie Mellon University
Columbia University
Cornell University
Dalhousie University
DePaul University
George Mason University
Georgia State University
Hokuriku Research Center
ICSI
IIT Delhi
IRTT
ISI
Johns Hopkins University
Lehigh University
MIT
New Jersey Institute of Technology
Norfolk State University
Pennsylvania State University
Purdue University
Rutgers University
Sao Paulo State University
Southern Illinois University
TU Berlin
TU Darmstadt
Texas A&M University
UC Berkeley
UC Davis
UC Irvine
UC Santa Cruz
UCLA
UCSD
UIUC
UNC Chapel Hill
UNC Charlotte
Universidad Michoacana de San Nicolas
Universita di Pisa
University of Advancing Technology
University of Illinois, Urbana-Champaign
University of Maryland
University of Massachusetts
University of Oregon
University of Southern California
University of Washington
University of Wisconsin - Madison
USC
UT Arlington
UT Austin
UT Dallas
Washington State University
Washington University in St. Louis
Western Michigan University
Xiangnan University
Youngstown State University

DETER User Research



Education

- Hands on exercises
- Students gain from direct observation of attacks and interaction
- Pre packaged for both student and teacher
  - Buffer overflows, command-injection, middle-in-the-middle, worm modeling, botnets, and DoS
- Facility support for class administration

Conclusion

Benefits

- Transformative research and facility for cyber security R&D
- Experimental science:
  - Fostering fundamental understanding world complexity
- Contribution transformation of field
- Proactive robustness and away from reactive security

Summary and Call to Action

- Growing DETER Community increasingly engaged in experimental science of cyber security

- Collaboration key part of DETER mission
  - DETERLab and new scientific experimentation

Join us

http://deter-project.org/