From Infrastructure to Science

Cyber Experimentation of the Future (CEF): Catalyzing a New Generation of Experimental Cybersecurity Research

David Balenson and Laura Tinnel, SRI International
Terry Benzel, USC Information Sciences Institute (ISI)

For the 67th IFIP 10.4 Working Group
January 22-26, 2015

Community-based effort to study current and expected cybersecurity experimentation infrastructure, and to produce a strategic plan and roadmap for developing infrastructure that supports tomorrow's research

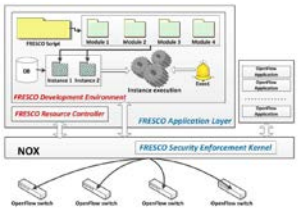
Collaborative effort by SRI International and USC-ISI

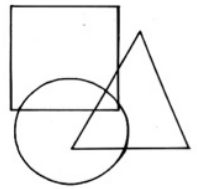
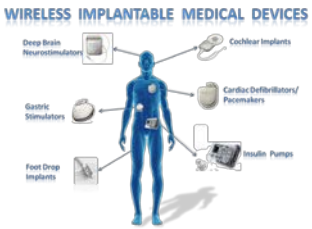
Funded by
NSF CISE/ACI
ACI-1346277
ACI-1346285

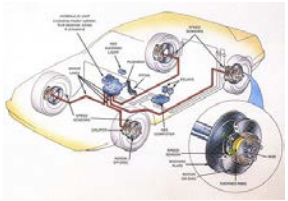
Industry Motivation: Research Infrastructure Must Keep Pace with Cyber Technology

- Cyberspace is rapidly evolving with nearly every aspect of society moving toward pervasive computing and networking
- Not even meeting today's needs
- Need to move quickly to meet tomorrow's needs
  - Highly specialized cyber-physical systems (CPS)
  - Interdisciplinary experimentation
  - Modeling and reasoning about human behavior
  - Advanced networking architectures (e.g., SDN







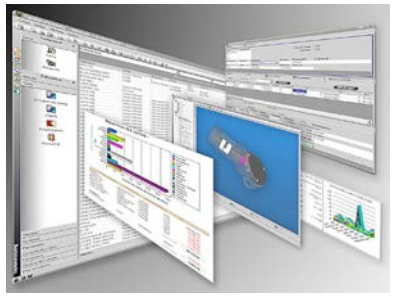

Cybersecurity Experimentation of the Future

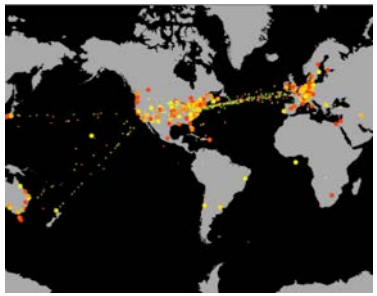
Scientific Motivation: Research Infrastructure Should Help Researchers Conduct Good Science

- PC Reviewer Question → "What is your confidence in this work?"
  - Problem areas
    - Knowledge of relevant prior work
    - Good experiment design
    - Real world models and data
    - Validation to eliminate (possibly unknown) errors
      - Design validation prior to execution
      - Test apparatus bugs/issues
      - Data collection: interference + validity
    - Sound, valid analysis
    - Repeatability
- CEF overarching goal for infrastructure → increase researcher EFFECTIVENESS
  - Help with above problem areas
  - Enable research, not constrain
  - Break the mold: rapid cycle: fail early, move on
  - Collective capturing/sharing of knowledge to keep pace with technology

What We Mean: Cybersecurity Research Infrastructure

- General purpose ranges and testbeds (physical and/or virtual)
- Specialized ranges and testbeds (physical and/or virtual)
- Software tools that supports one or more parts of the experiment life cycle, including, but not limited to:
  - Experiment design
  - Testbed provisioning software
  - Experiment control software
  - Testbed validation
  - Human and system activity emulators
  - Instrumentation – systems and humans
  - Data analysis
  - Testbed health and situational awareness
  - Experiment situational awareness
  - Other similarly relevant tools





- Specialized hardware tools – simulators, physical apparatus, etc.

Cybersecurity Experimentation of the Future

What We Mean: Experimentation

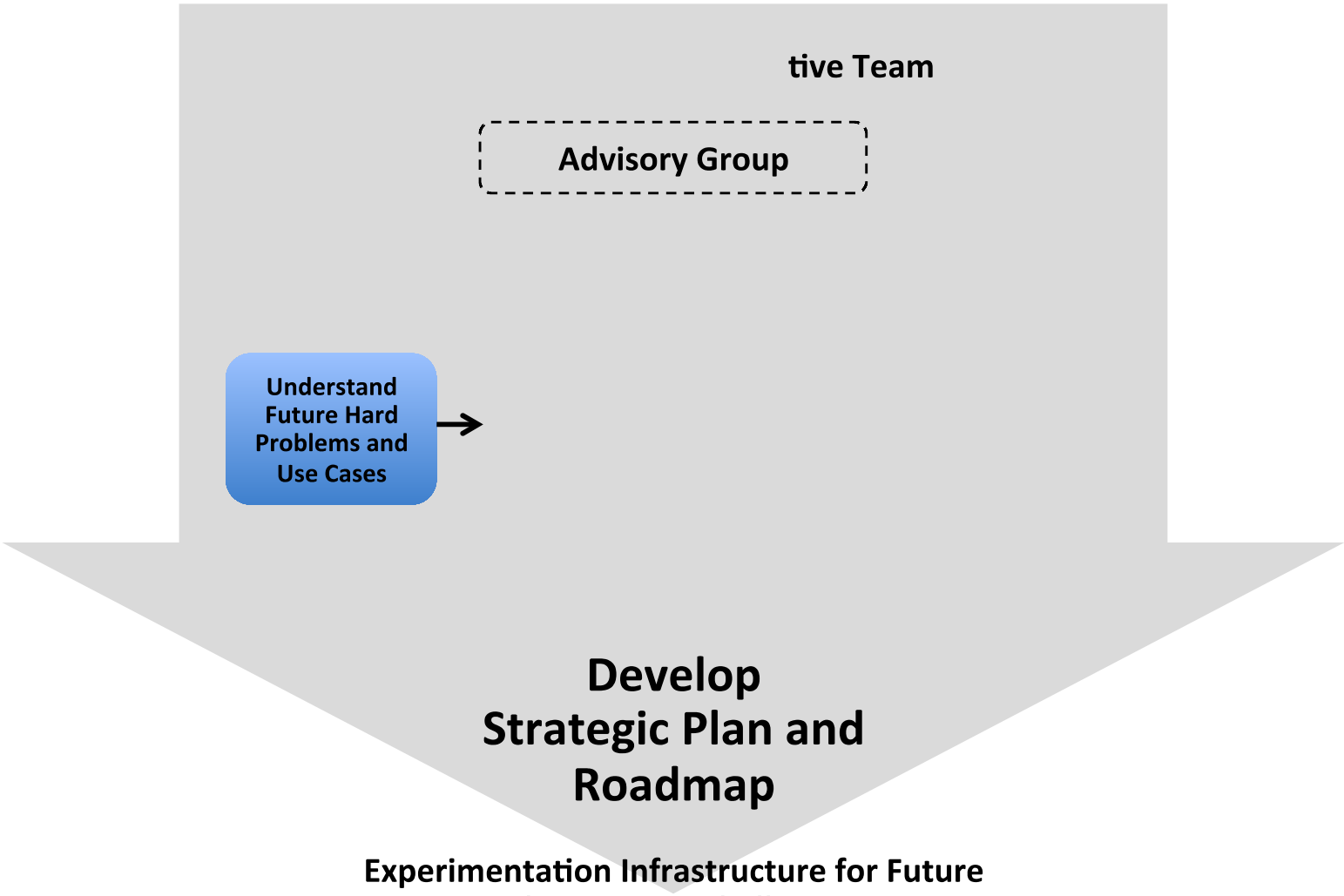
- Experimentation is about LEARNING
  - To explore a hypothesis
  - To perform an evaluation (not formal T&E)
  - To characterize complex behavior
  - To complement a theory
  - To understand a threat
  - To probe / understand a technology
- From case studies to controlled trials
- Highly theoretical to applied research

CEF Study: Key Findings

Transformational progress in three distinct, yet synergistic, areas is required to achieve the desired objectives:

1. Fundamental intellectual advance in the field of experimental methodologies and techniques
   - Broadly, but also with particular focus on complex systems and human-technical interactions
2. New approaches to rapid and effective sharing of data, knowledge and information synthesis
   - That accelerate growth of multi-discipline and cross-organizational knowledge and scientific peer review
3. Advanced experimental infrastructure capabilities and accessibility

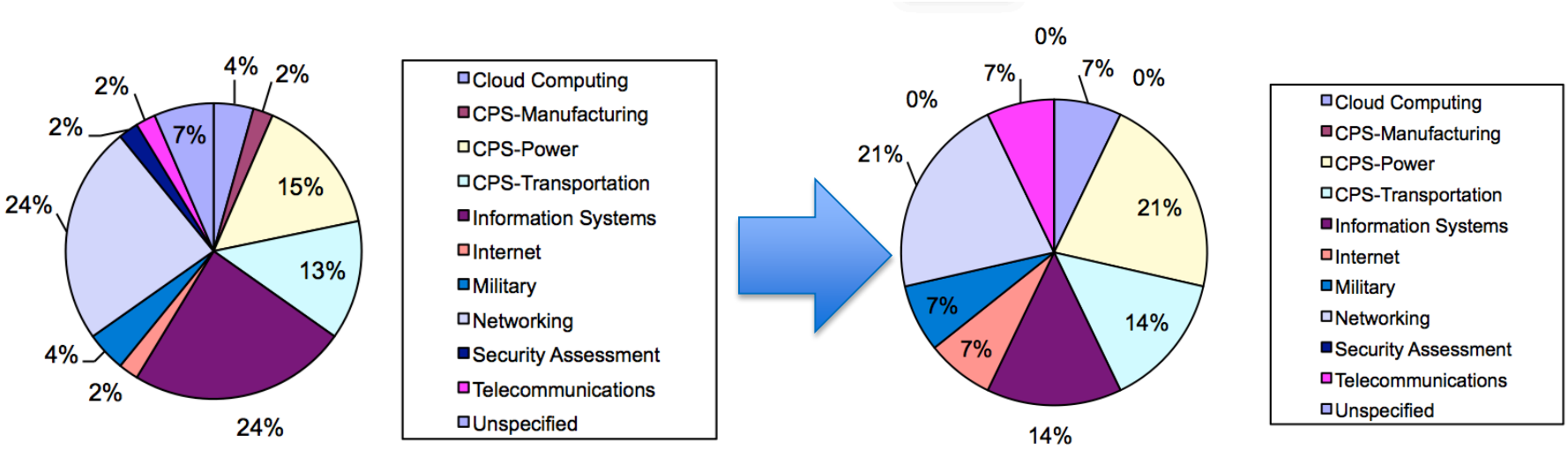
CEF Study: Approach



Cybersecurity Experimentation of the Future

CEF Survey: Existing Infrastructure

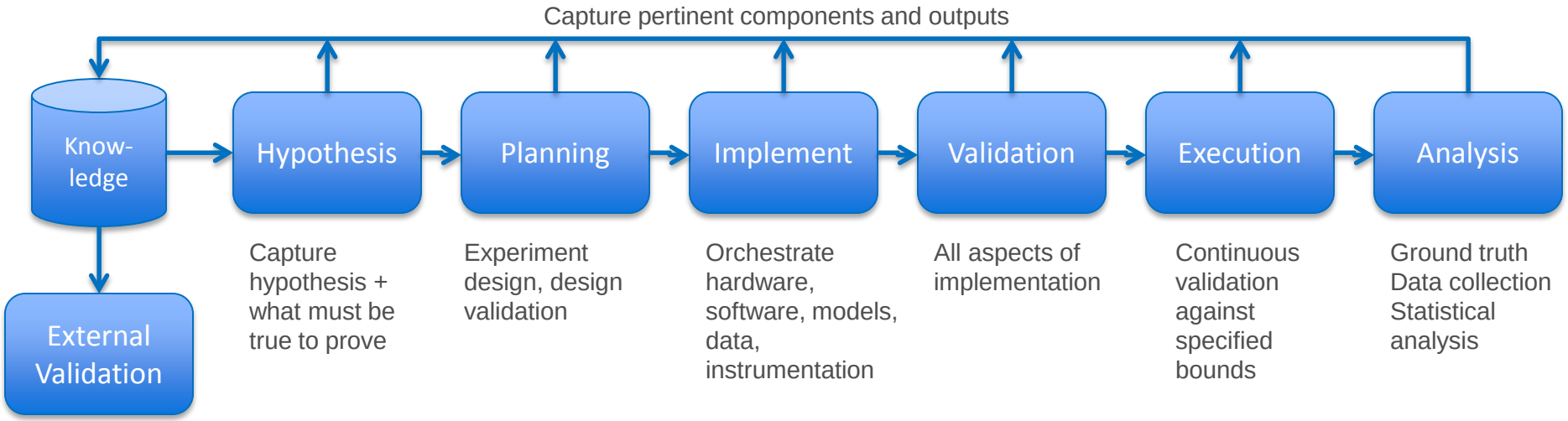
- Started with 45+ candidates across 10+ domains → focused review of 13



CEF Survey: Existing Infrastructure

- Summary Findings
  - Most existing infrastructure centers around traditional IT systems and networks
  - Some one-off, domain specific testbeds (most are closed, proprietary use only)
  - Almost all tools focus on setting up/running equipment, data collection
  - Little-to-no sharing / reuse of tools between testbeds (some exceptions)
  - Relatively few standalone, affordable tools (costly traffic generators, some freely available tools)
  - Almost NO support for the scientific process:
    - Experiment design, validation, analysis, sharing for repeatability / peer review
  - Overall, existing capabilities only begin to touch on what is needed – the space of what is needed is huge

Needed Infrastructure Support for Scientific Process



- Support full life cycle, repeatability – across all relevant domains
- Help with validation – catch errors as early as possible, e.g.,
  - Can experiment design prove hypothesis – or is it off base?
  - Are there bugs in the experimental framework?
  - Is the analysis method sound? E.g., correct statistical power
  - Are we capturing all the right things? E.g., ground truth
  - Will data collection interfere and possibly impact validity?

Existing Infrastructure Support for Scientific Process

Lifecycle	Information Systems	Networking	CPS - Power	CPS - Transportation	Medical
Hypothesis / Goals	n/a	n/a	n/a	n/a	n/a
Design / Planning	Low to none	Low to none	n/a	n/a	n/a
Design Validation	n/a	n/a	n/a	n/a	n/a
Implementation	High	High	Low	n/a	n/a
Implementation Validation	Low	Low	One offs	n/a	n/a
Execution	High	High	Low	n/a	n/a
Execution Validation	One offs	One offs	One offs	n/a	n/a
Data collection	One offs	One offs	One offs	Low to none	n/a
Analysis	n/a	n/a	n/a	n/a	n/a

** select domains*

CEF Eight (8) Core Capability Areas from Roadmap

6.1 Domains of applicability

6.2 Modeling the real world for scientifically sound experiments

6.3 Frameworks and building blocks for extensibility

6.4 Experiment design and instantiation

6.5 Interconnected research infrastructure

6.6 Experiment orchestration

6.7 Instrumentation and experiment analysis

6.8 Meta properties

CEF Core Capability Areas

6.1 Domains of Applicability

- Cross domain (critical infrastructure sectors)
- Multidisciplinary (computer science, engineering, math/modeling, human behavior, sociology, economics, etc.)
- Portability of experiments, packaged for sharing and use in cross-discipline experiments

6.2 Modeling the Real World for Scientifically Sound Experiments

- Models of real world environments
- Experiments that scale
- Experimentation with systems-of-systems
- Human activity

CEF Core Capability Areas

6.3 Frameworks and Building Blocks for Extensibility

- Workflow & management (comprehensive, human)
- Open/standard interfaces (API for extensibility, plugins write to API)
- Building blocks (libraries)
- Tool integration framework (to glue pieces together)

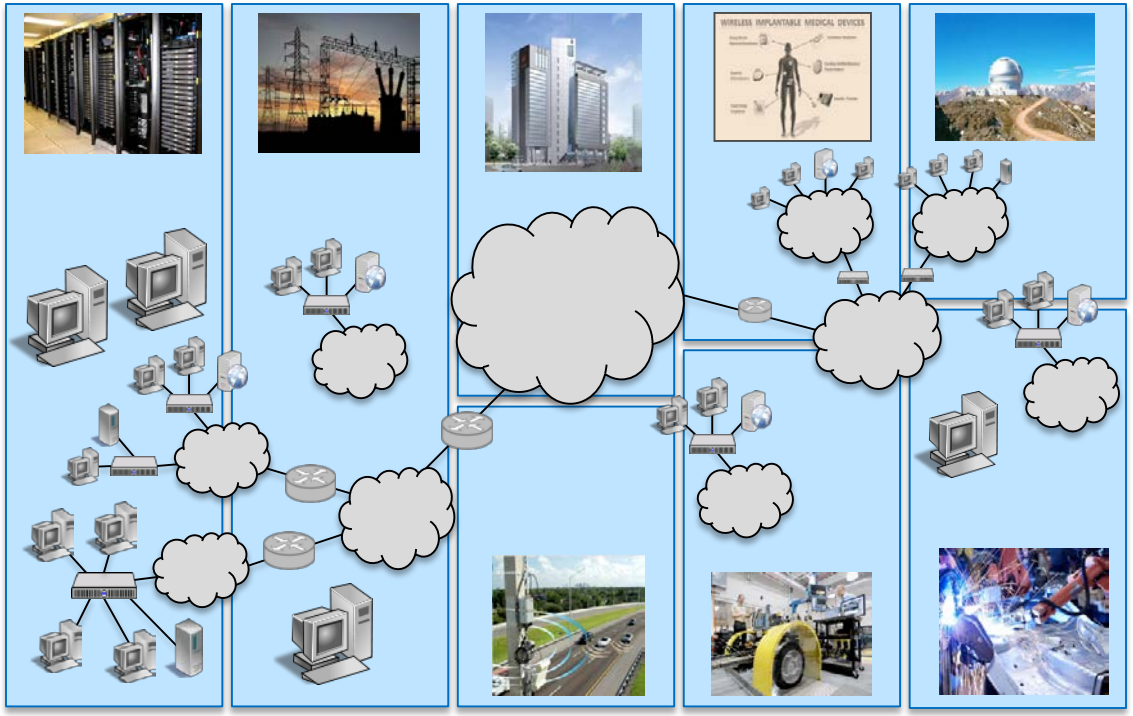
6.4 Experiment Design and Instantiation

- Design tools, specifications, ontologies, compiler
- Reusable designs for science-based hypothesis testing
- Automated discovery of local and distributed resources
- Dynamic instantiation of domain-specific test apparatus
- Validation of instantiated test environments and apparatus

CEF Core Capability Areas

6.5 Interconnected Research Infrastructure

- Automated, transparent federation to interconnect resources
- Dynamic and on demand, with sharing models
- Support integrated experiments that include real, emulated (virtual), and simulations



Cybersecurity Experimentation of the Future

CEF Core Capability Areas

6.6 Experiment Execution and Management

- Experiment orchestration
- Visualize and interaction with experiment process
- Experiment debugging with checkpoint and rollback
- Experiment execution validation

6.7 Instrumentation and Experiment Analysis

- Instrumentation and data collectors
- Transport and protection mechanisms
- Data repositories
- Data analysis

CEF Core Capability Areas

6.8 Meta Properties

- Usability (experiments, owner/operator)
- Confidentiality, availability and integrity of experiment ecosystem
- Social and cultural changes

CEF Study: Overall Recommendations

- New generation of experimental cybersecurity research will help shift the asymmetric cyberspace context to one of greater planning, preparedness and higher assurance fielded solutions
- Emphasis on equipment and related software infrastructure alone will far fall short of achieving the transformational shift in research, community, and supporting experimentation required to address cybersecurity in the ever escalating cyber environment
- Strong, coupled, and synergistic advances across three key areas
  - Fundamental methodological development
  - Fostering and leveraging communities of researchers
  - Capabilities of the infrastructure supporting that research

  will move the field beyond today's state of the art

Cybersecurity Experimentation of the Future

Conclusion

- Science-based experimentation infrastructure is critical to enabling future cybersecurity research
- Need revolutionary capabilities for advancing multi-discipline, complex and extreme scale, science-based experimentation for emergent cybersecurity research areas

- Draft CEF report: February 2015

- Consider: How would ***you*** contribute to a collaborative effort to build and share this infrastructure?

Contact Information

- David Balenson, SRI International
  david.balenson@sri.com, 703-247-8551
- Laura Tinnel, SRI International
  laura.tinnel@sri.com, 703-247-8533
- Terry Benzel, USC Information Sciences Institute (ISI)
  tbenzel@isi.edu, 310-448-9438

Collaborative effort by SRI International and USC-ISI

Funded by
NSF CISE/ACI
ACI-1346277
ACI-1346285

Cybersecurity Experimentation of the Future

ROADMAP TABLES

6.1 Domains of Applicability
Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Cross domain	Limited, mostly stand-alone sector specific	Common underlying frameworks, light weight specialized instantiations, emerging pair wise, multi-point ability to explore interdependencies	Common framework, open interfaces, specialized components, modeling of interdependencies,
Multidisciplinary	Emerging area of study and R.I. see ARL, Indiana, UIUC, CUTS Network Science, early research little R. I. support	Ability to capture model, recreate human behavior and integrate real humans into experiments (see 6.2)	Fundamental research across fields to define new multidisciplinary cyber security topic, new experimental methods and tools for integration
Portability of experiments	Very little sharing and thus few to no approaches. Requires extensive work to reach across disciplines. Some work at centers e.g. TCIP-G, ARL CRA, NSF Forces	Common practice sharing, community of multi-discipline researchers, conferences, publications, cultural changes	Incentive and social-cultural efforts to change mode of operation, need for creating fields

6.1 Domains of Applicability Research Milestones

Elements	Near-term	Mid-term	Long-term
Cross domain	Models and tutorials for non CS researchers to use RI. Worked examples of multi domain experiments	Collections of specialized components across multiple domains, cultural changes	Common framework, open interfaces, specialized components, modeling of interdependencies,
Multidisciplinary	Larger than RI, needs community cultural changes to bring different disciplines. Incentive and social-cultural efforts to change mode of operation, need for creating fields	Tools and methods for multidiscipline experimentation. For example power engineer models available or psychologist inputs to human behavior models. Present a User Experience tailored to The domain of experimentation The skills and background of the user community The terminology, workflow and methods traditionally preferred within an established community-of-interest	Fundamental research across fields to define new multidisciplinary cyber security topic, new experimental methods and tools for integration
Portability of experiments	Depends on common API/interfaces (6.3), build models with well defined interfaces, isolate domain specific components for clean interaction with RI	Composition tools to integrate domain specific models	Automatic encapsulation and compilation of experiments for use across multiple environments

6.2 Modeling the Real World for Sound Experiments Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Models of real world environments	Manually constructed environments, based on human understanding of real world; fidelity is limited.	(Semi-)automatically extract models of real computational devices and environments; automated model validation against real device or environment; rapid injection of models into simulation and physical test environments.	Semantically rich modeling language, with the ability to accurately model complex computing environments across diverse domains; techniques to extract salient features from the real world and create a model for test environment instantiation and validation.
Experiments that scale	Scale via simulation at fixed fidelity is norm.	Multiple dimensions of scale, ability to stretch experiment without affecting fidelity.	Scaling, models of complex system of systems, fidelity and building blocks
Experimentation with systems-of-systems	Little to no advanced experimentation; perhaps opportunity to leverage T&E community.	Ability to experiment via abstractions and to reason about composition	Abstractions, representation, multiple dimensions of scale, validity
Human activity	Several stand alone tools that are on host and based on statistical models of human behavior; at least one non-interfering off host tool exists that drives the user interface, playing the role of a human; at least one solution for virtualization non-interference.	Ability to accurately represent fully reactionary complex human and group activity in experiments; include live and synthetic humans without artificialities that may interfere in some experiments; capabilities to help ensure scientific validity when including live humans in experiments.	Theoretical understanding of what must be included in model to be realistic enough; automated creation of human models based on observations; wide range of behavioral research to understand decision process in different domains; models for other domains (e.g., transportation, energy); human activity at scale.

6.2 Modeling the Real World for Sound Experiments Research Milestones

Element	Near-term	Mid-term	Long-term
Models of real world environments	Transition semi-automated model creation capability from national labs and other sources for IT research.	Automated model creation for IT domain; semi-automated model creation for other domains (e.g., transportation, electric power).	Automated model creation for other domains (e.g., transportation, electric power);
Experiments that scale	Multiple scaling mechanisms, virtualization and models	Scalable infrastructure for experiment control and management, real time agents for mapping	Extreme scale, hybrid emulation and simulation experiment artifacts (validated)
Experimentation with systems-of-systems	Frameworks for abstracting system level structures. Use of test harness technology from system engineering communities	Experiment definitions that build on integrated components along with experiment management and data analysis at co-joined systems	Methods for reasoning about complex interactions between composed systems
Human activity	Transition limited on and off host human activity simulation capabilities from national labs and other sources for IT systems / networks research, and augment to create accurate human activity simulation for IT domain, driven by diverse, byzantine, models of goal-oriented human behavior, including actions and reactions *(leverage computer game technology)*; human attacker simulation.	Accurate, non-interfering human activity simulation for other domains (e.g., transportation, electric power); ability to specify behavioral bounds repeatability; tools to help ensure scientific validity when including live humans in experiments; assisted extraction of real human behavior, capture into models for subsequent simulation.	Automated extraction of real human behavior, capture into models for subsequent simulation. Extrapolation to generate projected human behavior models reflecting changes in technology and human interaction patterns.

6.3 Frameworks and Building Blocks for Extensibility Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Workflow & management	Rudimentary support, mostly highly manual	Integrated set of methods, tools and procedures to manage all stages of experimentation life cycle and reduce cognitive load	Definition of experiment life cycle, specific methods and tools for different phases in lifecycle, human to infrastructure interfaces, approaches to capturing experimenter intent
Open/standard interfaces	Some standards per facility, eg Geni R-Specs and evolving DETER SPI but none across capabilities, Lincoln Labs (see CRIS)	Collection of open/standard architectures and interfaces to advance composition and specialization of experiments and facilities	Defining architectural abstractions that can serve for the long-term. Support for extensible domains that are beyond the horizon, tools, techniques, models, and processes that evolve
Building blocks (libraries)	Fairly limited and uniform set of "nodes", "traffic" "connections", and "data". Local and not automated searchable or distributed across the community	Large extensible distributed collections/libraries identifying experiment infrastructure and experiment components, catalogs of lessons learned and domain specific information and rich knowledge base with provenance tagging available for automatic search	Cataloging approaches, meta data definitions, ontologies, knowledge bases. Catalog indexing and searching along with catalog consistency mechanisms
Tool integration framework	Limited plug-in cf Geni, NCR, Lincoln Labs	open framework with domain and problem specific plugins from which researchers may mix and match to instantiate the needed test apparatus as dictated by the specific research problem being addressed.	Common frameworks, tools abstractions, libraries and meta data for tools, usage models and examples, increase ease for experimenter developed tools to be contributed

6.3 Frameworks and Building Blocks for Extensibility Research Milestones

Element	Near-term	Mid-term	Long-term
Workflow & management	Definition of experimentation life cycle, survey/adaption of Software Engineering approaches, tools for different life cycle phases	Increased distributed operation and sharing of tools across distributed infrastructure. Techniques for instantiation of facilities	Advanced interfaces for experimenter interaction, approaches to capturing experimenter intent
Open/standard interfaces	Establishment of architectural abstractions, common interfaces, canonical representations	Collection of open/standard architectures and interfaces to advance composition and specialization of experiments and facilities	Support for extensible domains that are beyond the horizon, tools, techniques, models, and processes that evolve
Building blocks (libraries)	Initial catalog/library of building blocks, extensible with community interface for contributions	Cataloging approaches, meta data definitions, ontologies, knowledge bases	Cataloging approaches, meta data definitions, ontologies, knowledge bases. Catalog indexing and searching along with catalog consistency mechanisms
Tool integration framework	Common frameworks, tools abstractions, libraries and meta data for tools	usage models and examples, increase ease for experimenter developed tools to be contributed	Automatic integration, plug and play, interface abstractions

6.4 Experiment Design and Instantiation Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Design tools, specifications, ontologies, compiler	Limited approaches in GENI, NCR, Lincoln Labs	Rapid design of real world experiments across multiple domains; reuse of experiment designs, components; designs automatically validated.	Leverage work in ontologies and patterns from SW engineering; create pattern recognizers and sharing capabilities for reuse; techniques to analyze experiments and hypotheses to identify potential issues.
Reusable designs for science-based hypothesis testing	Non-existent	Common sound design patterns are captured and reused in experiments within a domain.	What is meant by reusable, and repeatable; capturing hypothesis, cataloging salient attributes.
Automated discovery of local/distributed resources	Almost none existent	Rich search of facility and experiment resources; automated search for resources based on experiment specifications.	Cataloging, meta data, distributed libraries, sharing, automation.
Dynamic instantiation of building blocks for domain-specific test apparatus	Moderate ability to automatically instantiate standard IT network/service designs exists for both physical and virtual test environments; almost none existent for other domains.	Dynamic realization of test environment designs for all domains; automated inclusion of widely distributed building blocks under different administrative control.	Valid ways to combine device simulations with real devices and networks to emulate the real world; dynamic instantiation, composition via common interfaces (dependent on building block activity, see 6.3).
Validation of instantiated test environments and apparatus	No validation against real world in test facilities, except for self-validation by experimenters; other kinds of needed validation is limited.	Automated correctness and real world applicability validation of test environment and apparatus instantiation.	Facility validation metrics; theoretical work in definition of non-interference for experimentation; models of correctness, validity; validation of test environments against models of the real world.

6.4 Experiment Design and Instantiation Research Milestones

Element	Near-term	Mid-term	Long-term
Design tools, specifications, ontologies, compiler	Leverage definition of base ontology and IT domain extensions from GENI and NCR; specification standards for design patterns from SW engineering; creation of pattern recognizers and sharing capabilities for reuse; user studies for multi-domain, scalable design interface.	Ontology domain extensions; fundamental research into compilable design languages; capabilities for reuse building on ontologies and patterns	Increased automation, real time feed back to-from ontologies, pattern capture and increased specification for sharing; semantically rich hypothesis and experiment specification; automated hypothesis and design analysis tools that flag issues that may create confounds or otherwise invalidate experimental results.
Reusable designs for science-based hypothesis testing	Simple sharing mechanisms for designs with meta descriptions	Definitions of ranges of repeatable and reusable, formal specification of degrees of tolerance	Capturing and reasoning about hypothesis, cataloging salient attributes, creation of pre/post conditions for experiment correctness
Automated discovery of local/distributed resources	Offline, out of band search of central repository	Distributed resources repositories and descriptors, distributed search, manually guided	Increased automation, distributed search, discovery via automated meta information/locators
Dynamic instantiation of building blocks for domain-specific test apparatus	Transition automated test environment instantiation capability from national labs and other sources for IT research; building blocks integrated into experiment definitions, static and manual	Automated test environment instantiation for other domains (e.g., transportation, electric power); dynamic allocation of remote/distributed resources as part of experimentation	Automatic inclusion of building blocks from template into running experiments for both local and remote resources
Validation of instantiated test environments and apparatus	Experiment apparatus diagnostic test procedures tuned for specific environments, verify non-interference by observation; assisted validation of instantiated test environments against real world models of the IT domain.	Common standards of validation, validation across multiple disparate infrastructures, definition of metrics for instrumentation for self-check; fully automated validation of instantiated test environments against simple real world models in multiple domains.	Theoretical work in non-interference definition for experimentation, models of correctness and validity; fully automated validation of instantiated test environments against complex, large-scale real world models in multiple domains.

6.5 Interconnected Research Infrastructure Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Automated, transparent federation to interconnect resources	*Experiment* federation in DETER and Geni, requires a priori establishment	Fully automated using discovery approach described above, connection of wide range of aspects (experiments, testbed, specialized components)	Flexible interconnection semantics, auto negotiation, access control specification and mechanism
Dynamic and on demand, with sharing models	Little to no dynamic and on demand sharing. Mostly closed models. DRAGON provides some dynamic short lived interconnection semantics.	Highly dynamic interconnection, Experimental Models that span multiple aspects	Dynamic establishment of interconnection, short lived sharing of resources, resource allocation at multiple abstractions. Modeling of complex shared attributes
Support integrated experiments that include real, emulated (virtual), and simulations	Most support fixed level of abstraction with some mix of emulated, simulated and virtual (cf DETER, Emulab)	Support wide range experimental infrastructure bare metal, virtual machines, emulated CPUs, networks, simulations, in any combination, to create the appropriate phenomena required for validity in that part of an experiment	Approaches to mixed abstraction in specification and experimental realization, methods for assessing fidelity tradeoffs and validity (see 6.2)

6.5 Interconnected Research Infrastructure Research Milestones

Element	Near-term	Mid-term	Long-term
Automated, transparent federation to interconnect resources	Common framework for basic interconnection; resource publication and discovery for current/common resource types.	Inter-connection authorization policies, tools to manage policies, mechanisms to implement inter-connection; policies limited to current/common resource types.	Interconnection and shared resource access fully supported for all types of resources and many modes of access to them. Automated negotiation of access and enforcement of access controls.
Dynamic and on demand, with sharing models	Dynamic resource discovery and automated request for sharing, but human operators involved in setting up interconnections, and in authorizing access to shared resources	Initial inter-connection and/or federation driven by policy model and tools, but authorized with operator involvement. After initial steps, on-demand access to resources shared within an inter-connection.	Multiple models for sharing, fully supported in policy and tools; fully dynamic inter-connection, fully dynamic resource sharing in an inter-connection, for parties that fit an existing sharing model.
Support integrated experiments that include real, emulated (virtual), and simulations	Integrated experiments with current/common resource types of real and virtual.	Sharable resource types extended to one or more common model of simulation, emulation.	Sharable resources include large compositions of resources.

6.6 Experiment Execution and Management Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Experiment orchestration	Some capability for IT systems/networks to do simple orchestrations; little or none for other domains.	Fully orchestrate test apparatus to execute and manage the steps of an experiment, inclusive of multiple paths of execution that depend on time, success of prior steps, changes in the environment, and other factors.	Semantically rich orchestration language supporting full logic and conditional branching, with extensibility for domain specific experiments; domain specialized device control; pre and post conditions for experiment correctness validation; instrumentation for validation monitoring
Visualization and interaction with experiment process	Some capability for IT systems/networks, with advanced capability limited to government use; little or none for other domains	Ability to fully interact with experiments: pause, examine progress, make changes and resume, or restart; intuitive visualization provides full understanding of experiment execution state.	Visualization approaches that convey experiment intent and help users quickly understand the impact of unexpected events on experiment integrity; resolve fundamental issue of how to suspend and resume system and network activity with no loss of integrity; experiment flow and dependency understanding and diagnosis techniques.
Experiment debugging with checkpoint & rollback	Moderate checkpoint capability for virtualized hosts; no checkpoint capability for real hardware hosts, although "deep freeze" technology could help; debugging is manual.	Create and annotate multiple checkpoints to which one may choose to rollback; interact with experiments to debug at runtime.	Resolve fundamental issues: how to capture and restore system and network state with no loss to experiment integrity; what activity is sufficient to capture.
Experiment execution validation	Manual process	Experiment execution is monitored in real time, alerts of experiment validation status; automated support for rapid diagnosis of experiment execution failures.	Fundamental research needed to understand what it means for an experiment to be valid or invalid; techniques to represent validity and test for it.

6.6 Experiment Execution and Management Research Milestones

Element	Near-term	Mid-term	Long-term
Experiment orchestration	Semantically rich orchestration language and execution capability, supporting full logic and conditional branching for IT systems/ networks domain.	Support for other domains using manually analysis and integration of specialized, embedded devices.	Automated analysis and inclusion of specialized, embedded devices for experiment orchestration.
Visualization and interaction with experiment process	Transition existing experiment visualization and control capabilities from national labs and other sources; support tools for rapid diagnosis of experiment execution failures.	Visualization approaches that convey experiment intent and help users quickly understand if and how issues that arise affect experiment integrity; semi-automated assistance for rapid diagnosis of experiment execution failures.	Suspend and resume system and network activity with no loss of experiment integrity; fully automated rapid diagnosis of experiment execution failures.
Experiment debugging with checkpoint and rollback	Automation and inclusion of virtualization based rollback capability in experiment process control.	Capture sufficient experiment state to support diagnostics and rollback.	Rollback experiment state with no loss to experiment integrity.
Experiment execution validation	Monitoring and warnings for conditions that may invalidate and experiment and require further (manual) investigation.	Understanding of and ability to represent validity of experiment.	Ability to automatically verify experiment validity.

6.7 Instrumentation and Experiment Analysis Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Instrumentation and data collectors	Some standard (e.g., PCAP, syslog), mostly hand crafted, rudimentary, low level; no instrumentation automation or dynamic tasking of instrumentation.	Ubiquity of non-interfering instrumentation for virtual, physical, static, dynamic, and real-time systems, that allows full or selective collection of data; automated instrumentation of specialized devices.	Novel instrumentation techniques that are tamper proof and non-interfering; techniques to compensate for interference when such cannot be avoided; instrumentation specification language with techniques to automatically analyze and create instrumentation for specialized devices.
Transport and protection mechanisms	In and out-of-band mechanisms are not transparent; no enforced protections.	Collected data transport mechanism(s) from sensors to repository that are fully protected against tampering, leakage, and misuse and do not interfere in experiment.	Novel techniques for data transport that are tamper proof and non-interfering.
Data repositories	Limited, centralized, and distributed storage facilities; many unable to deal with "big data"; no data store exists for sharing experimental results.	Data repositories, which are fully searchable by both humans and machines.	Semantically rich data description language; techniques for automated data tagging with full context, provenance, and auditability.
Data analysis	Basic log analysis tools, forensics oriented with limited pattern analysis, focused on system and network administrators and cyber security analysts; one-off tools are written on an experiment-by-experiment basis.	Collections of pre-canned, yet extensible tools for post-experiment, multi-purpose analysis; real-time analysis and dynamic control of experiment instrumentation to perform "as needed" data reduction at experiment runtime.	Tool designs for extensibility and multi-domain use; specification and validation of confidence bounds on experiments; techniques to automatically analysis and compare experimental results from multiple testbeds; methods to correctly merge data from multiple sources, compensating for the real world issue of non-synchronized clocks; transformations to anonymize, normalize, and compress.

6.7 Instrumentation and Experiment Analysis Research Milestones

Element	Near-term	Mid-term	Long-term
Instrumentation and data collectors	Tamper-proof instrumentation; configurable instrumentation allowing for full or selective collection of data	Novel, non-interfering, instrumentation for virtual, physical, static, dynamic, and real-time systems; automated compensation for interference when such cannot be avoided; semi-automated creation of instrumentation for specialized devices.	Automated instrumentation of specialized devices.
Transport and protection mechanisms	Novel data transport mechanism(s) from sensors to repository that are fully protected against tampering, leakage, and misuse.	Non-interfering transport mechanisms.	
Data repositories	Human searchable data repositories, using "big data" analysis tools.	Collected data automatically tagged with keywords, province, and audit information; data is machine searchable.	Full experiment context of data is automatically inferred as data is collected, tagged to data, and is machine searchable.
Data analysis	Integrated set of existing data collectors from multiple domains; ability to dynamically control instrumentation at experiment run time using signature-based detection of events *(e.g., something interesting is about to happen)*; ability to merge data from multiple sources without synchronizing clocks; assisted merging of collected data to construct ground truth in support of experiment measurements.	Collections of extensible tools that provide semi-automated assistance for post-experiment analysis across multiple domains; specification and validation of confidence bounds on experiments *(when do results become questionable?)*; real-time analysis and dynamic control of experiment instrumentation to perform data reduction at experiment runtime; ability to anonymize experiment data; automated construction of ground truth in support of IT domain experiment measurements.	Automated construction of ground truth in support of non-IT domain experiment measurements; collections of extensible tools that provide automated post-experiment analysis across multiple domains.

6.8 Meta Properties
Summary of Current State, Vision, and Needed Research

Element	Current State	Vision	Needed Research
Usability (experiments, owner/operator)	Varies by type of infrastructure and community, some highly usable by targeted communities, few to none usable across multiple skill levels and diverse user base	Usability for research community spanning wide base of domain experts to power users. Facility owner operators from multiple domains and for overarching frameworks	Pedagogical tutorials, self teaching facilities, definition of common experimenter interfaces, applications, processes and procedures
Confidentiality, availability and integrity of experiment ecosystem	None to very limited, most existing research infrastructure developed as research	Best practices for establishing provenance of experiments and artifacts, integrity of raw data, experiment configurations, software, etc. Confidentiality of IP, availability and timeliness	New mechanisms for assuring integrity, and confidentiality of experiments and experiment data. Tools and methods for automatic scheduling and resource management for increased availability
Social and cultural changes	Small community, somewhat in grown, mostly network researchers building and using network testbeds, some stove pipe communities for CIP	Large multi discipline community of researchers operating across distributed and extensible research infrastructure working in an environment of sharing and libraries	Community efforts to socialize new methods and initiatives. Bring together RI researchers and developers, conduct exercises aimed at identifying commonalities and differences and define opportunities for collaboration and integration

6.8 Meta Properties Research Milestones

Element	Near-term	Mid-term	Long-term
Usability (experiments, owner/operator)	Introduction of multiple points of entry to experimentation facility. Common operations and build environments, common language and terminology, extended operator community	Creating interfaces with customized look-and-feel tools in a given domain of experimentation. "Domain specific users" want to work using the terminology, workflow and methods traditionally preferred within an established community-of-interest.	Pedagogical tutorials, self teaching facilities, definition of common experimenter interfaces, applications, processes and procedures
Confidentiality, availability and integrity of experiment ecosystem	Protected experimenter space for isolation of sensitive data, increased facility controls for integrity. Standard security evaluation of infrastructure	Base core experimental facilities of medium to high assurance systems, e.g., SE Linux, develop integrity checking processes, create scheduling tools, portals for availability	New mechanisms for assuring integrity, and confidentiality of experiments and experiment data. Tools and methods for automatic scheduling and resource management for increased availability
Social and cultural changes	Conduct workshops, disseminate to broader community, begin catalog and library collection in conjunction with frameworks (Section 6.3) and domains (Section 6.1)	Advertise new capabilities, require contributions back from users, maintain evolving libraries and search functions, require use of capabilities in certain funded efforts	Through education and socialization over time create environment that is so beneficial that people will seek it out and use it rather than inventing their own.

EXAMPLE USES

Example Envisioned Usages

- Integrated Experimentation Environment (IEE)
- Model Extraction Tool

Envisioned Usage – Integrated Experimentation Environment (IEE)

- Assume researcher has:
  - Identified problem to be solved
  - Developed the experimental approach
  - Captured science-based hypothesis (or experimentation goal)
- (Semi-)automated tools and methodology to support experiment design
  - Organize the required building blocks, including network topology, host configurations, services, behaviors, etc.
  - Apply real-world models
  - Specify proper instrumentation to capture ground truth
  - Validate experiment design
- Analogous to an Integrated Development Environment (IDE)
  - Automatically compile to check syntax and convert to actionable representation, link required libraries, and load into testbed or other experimentation infrastructure

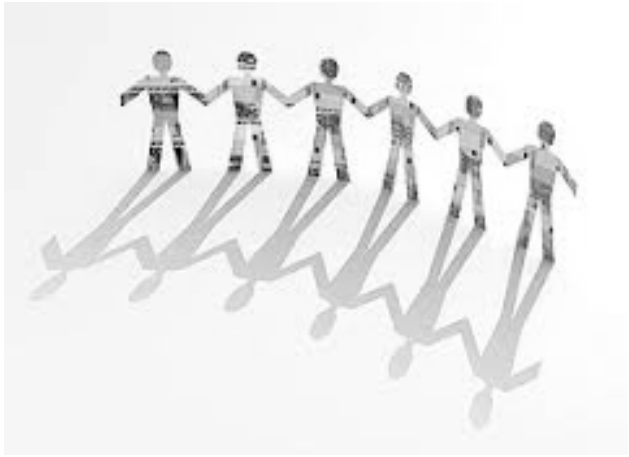
Envisioned Usage – Model Extraction Tools

- (Semi-)automatically extract salient features from a real operational environment
  - Including network topology, host configurations, services, procedures, team workflow, usage, goals, etc.
- Populate a model of that environment
- Use the model to create a virtual instance or replica of the environment in a testbed
- Use the replica to experiment with and test cybersecurity characteristics, including adversary attack models and potential protection mechanisms
  - I.e., a "virtual sandbox"
- For example, smart transportation – inject the tool and extract the dynamic, mobile ad hoc environment and reproduce it in a testbed to explore adversary models and dynamic defenses within autonomous vehicles and roadways

BACKUP SLIDES

Community Support and Consensus

- Two Advisory Group meetings (Dec 2013 and Apr 2014) (next one in Jan 2015)
- Three multi-day Study Group meetings (Mar, May, and Jun 2014)
  - 75 participants from 50 organizations
- Two extended team offsite meetings (Jul and Oct 2014)
- A number of additional presentations and meetings
  - NSF SaTC team (Jan 2014), CSET Panel (Aug 2014), AFRL (Aug 2014), MIT-LL (Sep 2014), CSIA-IWG (Sep 2014), NSF and DHS Mgmt (Oct 2014)



- Report schedule:
  - Sep 2014: Extended Outline
  - Oct 2014: Partial Draft Report
  - Feb 2015: Draft Report
  - Mar 2015: Community Review
  - Apr 2015: Final Report

Advisory Group

- Goals
  - Ensure the study wide set of participants from relevant areas and disciplines
  - Ensure the study answers the important, forward-seeking questions and issues
- Roles
  - Identify existing infrastructure and user communities
  - Clarify and refine study groups and topics
  - Identify and invite key community members
  - Review plans for structure and content of report
  - Review and provide feedback on draft report
- Members
  - Michael Bailey, U. Michigan
  - Steve Corbato, U. Utah
  - Steven King, ASD(R&E)
  - John Lowry, Raytheon BBN
  - Doug Maughan, DHS S&T
  - Bill Sanders, UIUC
  - Patrick Traynor, GA Tech

Study Group Participants (1 of 2)

- Aaron Johnson, NRL
- Alefiya Hussain, USC-ISI
- Andre Weimerskirch, UMTRI
- Angelos Keromytis, NSF
- Anil Somayaji, Carlton University
- Anita Nikolich, NSF
- Anthony Joseph, UC Berkeley
- Anup Ghosh, Invincea
- Benjamin Edwards, U. New Mexico
- Bill Scherlis, CMU
- Brad Martin, ODNI
- Brandon Schlinker, USC
- Brian DeCleene, BAE Systems
- Bryan Lyles, NSF
- Charles Palmer, IBM Research
- Dale Johnson, MITRE
- Damon McCoy, GMU
- Dan Cantu, Sandia
- David Balenson, SRI International
- David Corman, NSF
- David Nicol, U. Illinois at Urbana-Champaign
- Donna Dodson, NIST
- Elaine Shi, U. Maryland
- Ersin Usun, Xerox PARC
- Ethan Katz-Bassett, USC
- George Kesidis, Penn State
- Gianluca Stringhini, UC Santa Barbara
- Gideon Juve, USC-ISI
- Grant Wagner, NSA Research
- Gregg Schudel, Cisco
- Herb Lin, National Academies
- James St. Pierre, NIST
- Jean Camp, Indiana U.
- John Baras, U. Maryland
- John McHugh, UNC and Redjack
- John Sebes, TrustTheVote Project
- John Wroclawski, USC-ISI
- Josiah Dykstra, NSA, UMBC
- Kevin Butler, U. Oregon
- Kevin Sullivan, U. Virginia

Study Group Participants (2 of 2)

- Kevin Thompson, NSF
- Laura Tinnel, SRI International
- Lee Rossey, MIT-LL
- Luke Berndt, DHS S&T
- Manimaran Govindarasu, Iowa State
- Marshall Brinn, Raytheon BBN
- Mary Denz, AFRL
- Matthew Elder, Symantec
- Maverick Woo, CMU
- Micah Sherr , Georgetown
- Miles McQueen, INL
- Paul Boynton, NIST
- Phil Porras, SRI International
- Ritu Chadha, Applied Communication Sciences
- Roy Maxion, CMU
- Ryan Goodfellow, Washington State
- Sam Weber, CMU/SEI
- Sami Saydjari, Cyber Defense Agency (CDA)
- Sandy Clark, U. Penn
- Scott Lewandowski, The Wynstone Group
- Sean Peisert, UC Davis
- Sonia Fahmy, Purdue
- Steve Corbato, U. Utah
- Steve Schwab, USC-ISI
- Ted Faber, USC-ISI
- Terry Benzel, USC-ISI
- Terry Champion, Skaion
- Thanassis Avgerinos, CMU
- Thomas Carroll, PNNL
- Thomas Edgar, PNNL
- Tim Yardley, U. Illinois at Urbana-Champaign
- Vaibhav Garg, Drexel
- Vincent Urias, Sandia
- Von Welch, Indiana U.
- Zach Tudor, SRI International

CEF Study: Roadmap

- Presents requirements, objectives and goals in key areas over 3, 5, and 10 year phases
  - Some phases build upon each other and others require new fundamental research over a long time period
- Key capabilities consider:
  - Current experimental cybersecurity research and its supporting infrastructure
  - Other types of research facilities
  - Existing cyber-domain “T&E” capabilities (primarily DoD)
- Presumes advances in key computer science disciplines
  - Ontologies, meta-data, libraries, and corresponding resource discovery

Preliminary Set for Inclusion in CEF Report

Domain	Infrastructure Name	Owning Organization
Transportation: System	Connected Vehicle Testbed	Department of Transportation
Transportation: Automotive	OCTANE	George Mason University
Networked Systems: Internet	FIRE	European Union
Networked Systems: LAN/WAN	DETER	USC-ISI
Networked Systems: LAN/WAN	GENI	National Science Foundation
Networked Systems: LAN/WAN	NCR	DARPA / TRMC / LMCO
Networked Systems: LAN/WAN	PlanetLab	Princeton
Networked Systems: LAN/WAN	LARIAT	MIT Lincoln Laboratory
Networked Systems: LAN/WAN	Skaion TG	Skaion Corporation
Networked Systems: Cloud	Open Cirrus	Intel
Networking	StarBed/JGN2+	National Institute of Information and Communications Technology (NICT) – Japan
Networking: SDN	ON.LAB	Stanford & Berkeley
Networking: Wireless	Orbit / WINLAB	Rutgers University
Telecomms: Smart Phones	PhoneLab	University of Buffalo
Electric Power: Smart Grid	Trustworthy Cyber Infrastructure for the Power Grid (TCIPG)	University of Illinois Urbana-Champaign
Electric Power: Smart Grid	powerNet	Pacific Northwest National Laboratory
Electric Power: Smart Grid	PowerCyber	Iowa State