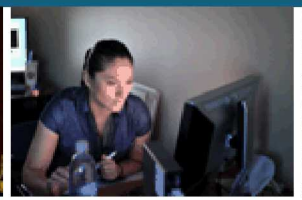
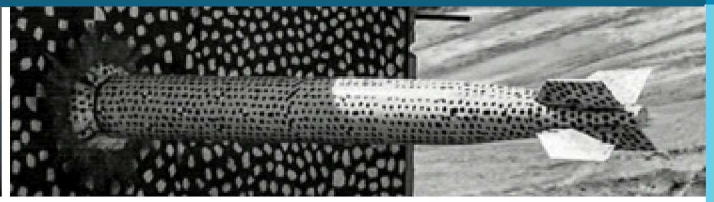
SAND2019-1396C

Evaluating Moving Target Defense with Quantitative Resilience Analysis





PRESENTED BY

Dr. Meghan Galiardi and Dr. Shamina Hossain-McKenzie

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Cyber Resilience: An Emerging Need

Situation: advanced persistent threats (APTs) are working tirelessly to compromise the Nation's most critical digital assets and networks

Problem: cyber community is starting to recognize that

It is simply impossible to stop all attacks and compromises

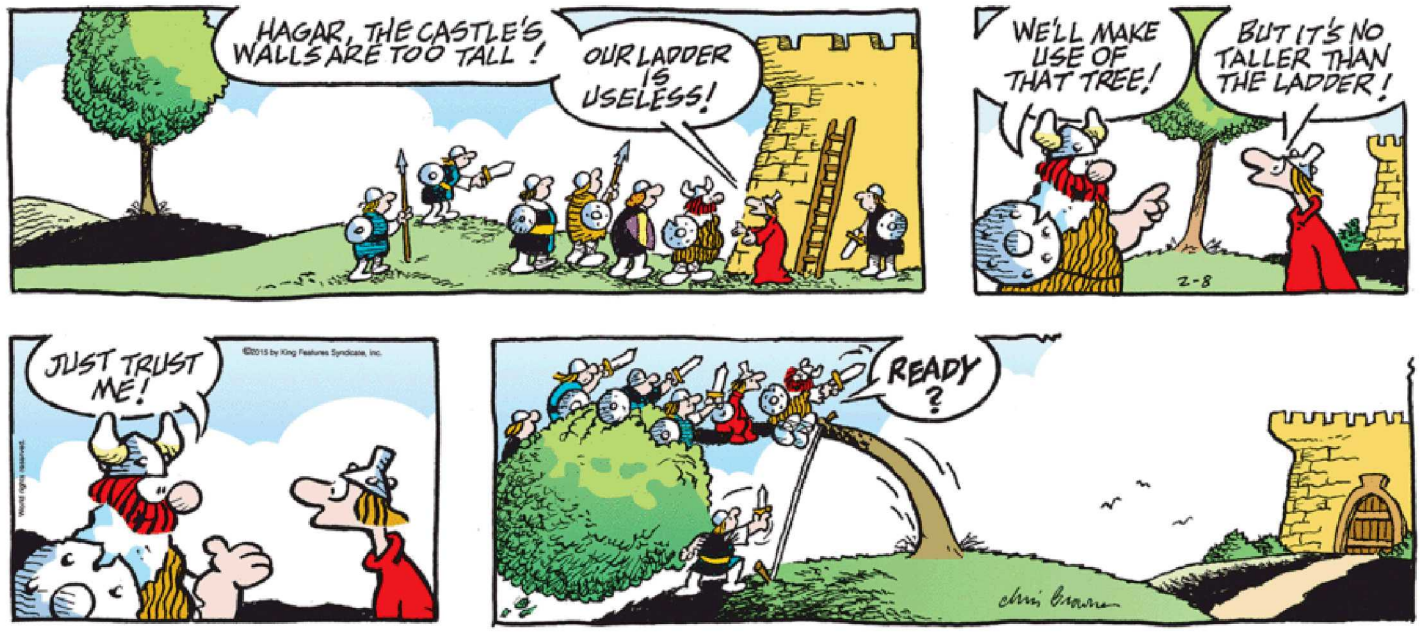
Current response capabilities are wanting: 256 days to detect infiltration, 90-120 days to remediate*

Approach: addressing the cyber threat requires changing mindsets and capabilities

*Dr. Dale Meyerrose, Major General, U.S. Air Force, Retired, "What's Holding Us Back?," Cyber Resilience Summit 2017

"You're never going to have an impenetrable network, that is a fool's errand. You have to have the ability to fight through the hurt".
- Rear Adm. Danelle Barrett, Dir. of the Navy Cyber Security Division, Office of Chief of Naval Ops

Philosophy



http://hagarthehorrible.com/comic_tag/castle-walls/

What is Cyber Resilience?

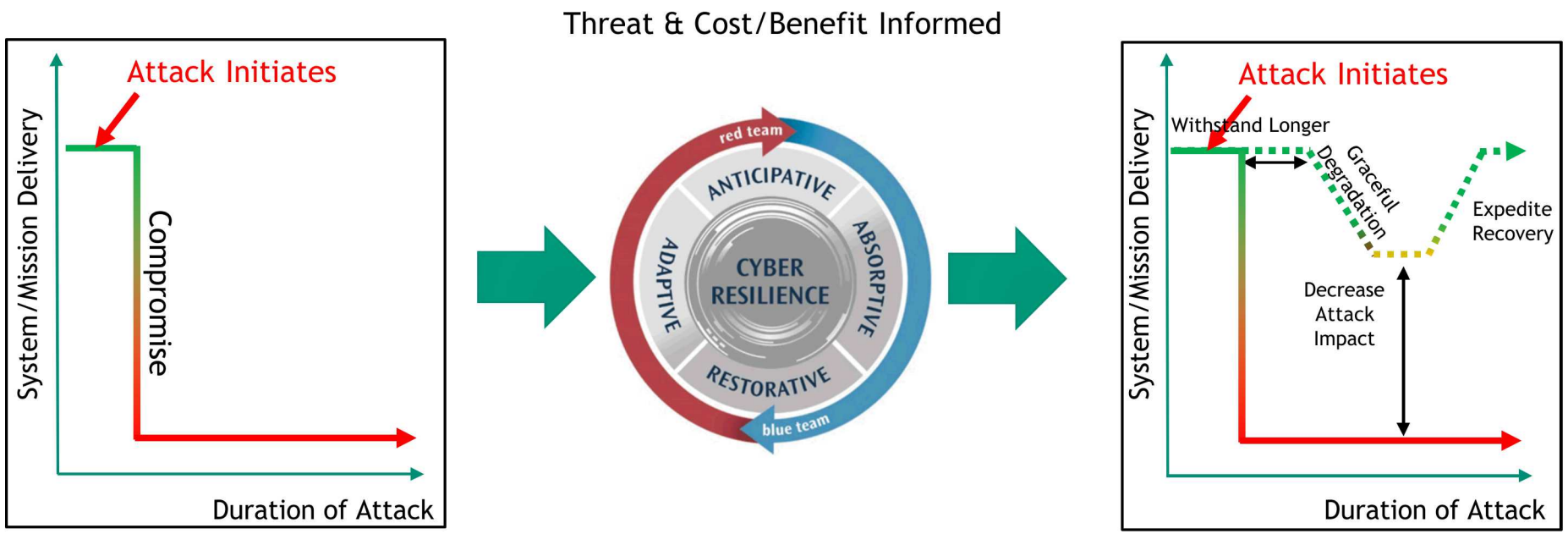
Informally, cyber resilient systems are able to execute required mission parameters despite an hostile cyber-threat environment.

Security and Resilience

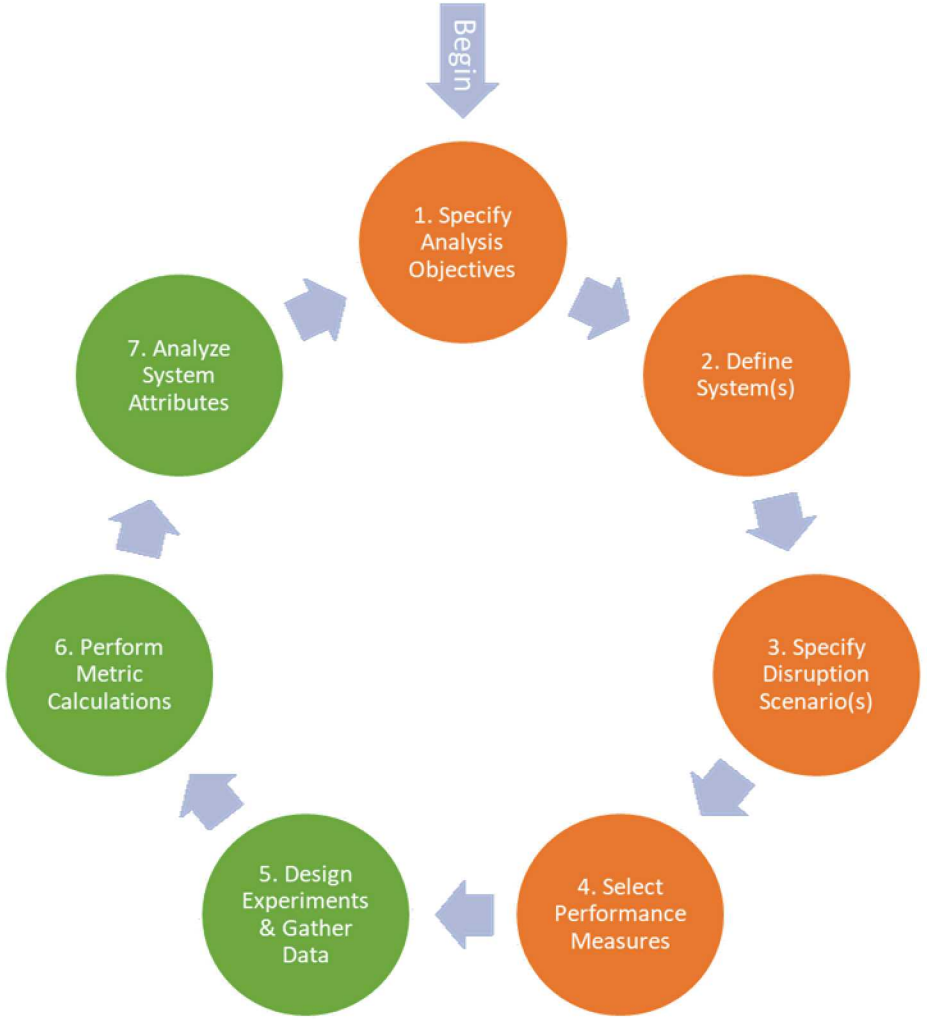
	Traditional Security	Resilience
Goal	Prevent, protect network to maintain CIA	Survive, overcome to execute mission
Assessment Focus	Vulnerability	Consequence, response
Enabling Mechanisms	Restricting Access & Management of Permissions	Prepare, withstand, adapt, recover
Metric Focus	Threat, vulnerability	Mission execution, consequence

Security and resilience activities are complementary efforts that come together to form a comprehensive, risk management strategy

Cyber Resilience Objectives



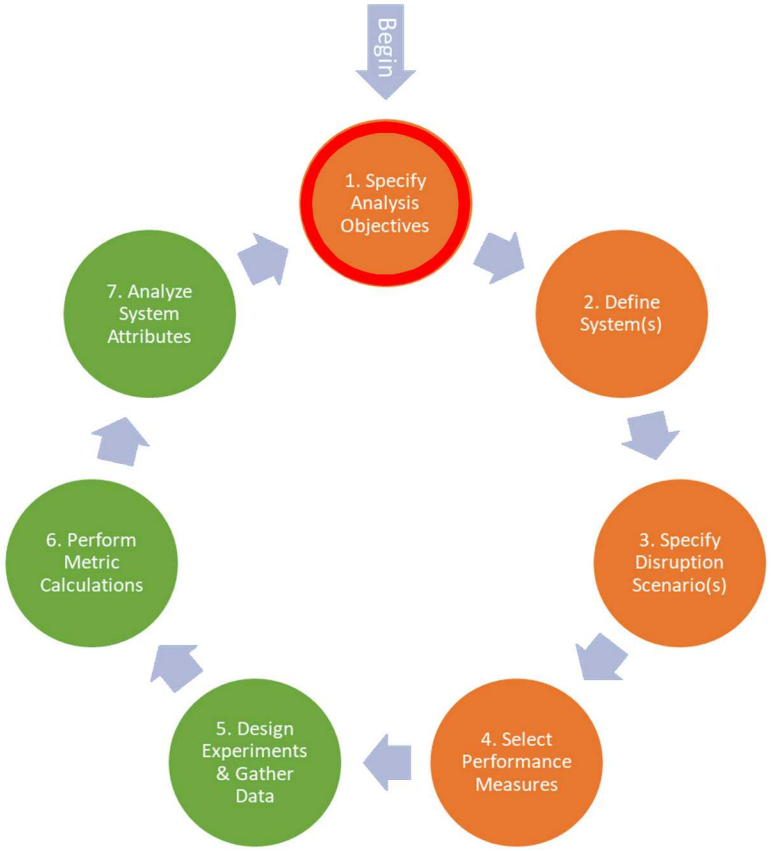
Our Methodology:
An extension of IRAM (Infrastructure Resilience Analysis Methodology)



This methodology provides a consistent, repeatable process for performing cyber resilience analyses

Biringer, Betty, Eric Vugrin, and Drake Warren. *Critical infrastructure system security and resiliency*. CRC press, 2016.

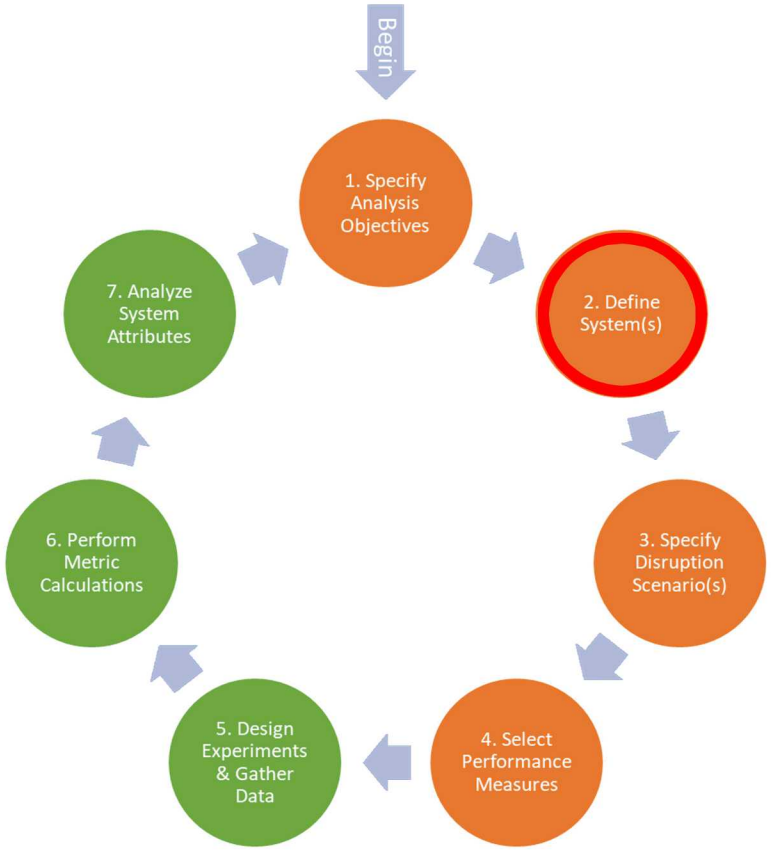
Step 1: Specify Analysis Objectives



Define the specific questions the analyst aims to answer and the ultimate objectives for the analysis.

- Essential for establishing the scope of the analysis
- Informs all subsequent steps
- Failure can result in an analysis that does not address objectives and wastes time, effort, and resources.

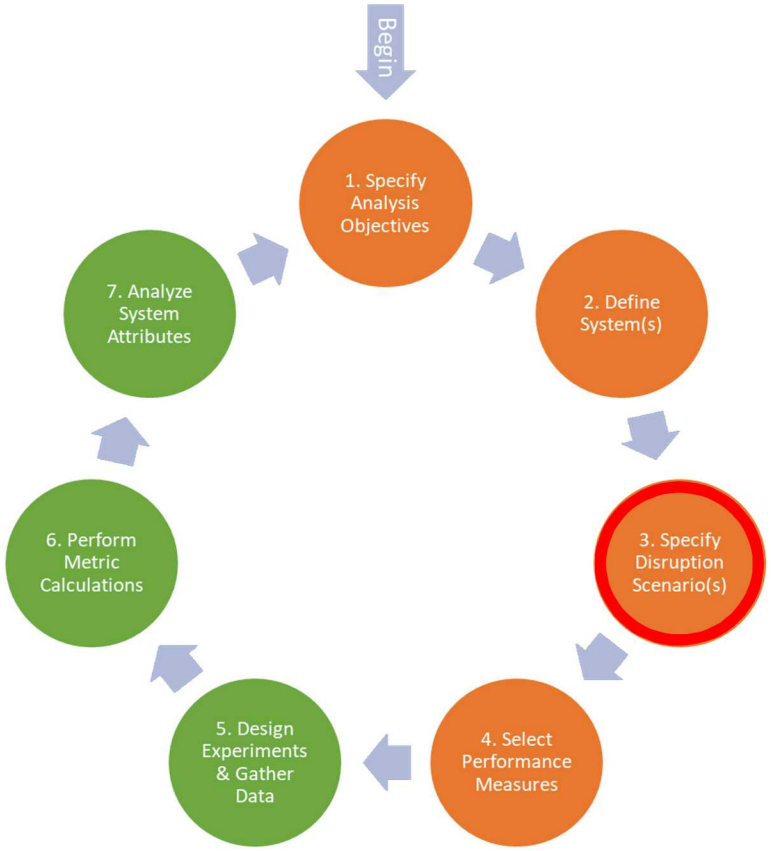
Step 2: Define System(s)



Describe the system's intended mission and how it achieves that mission

- System components or subsystems
- System structure
- Component dependencies/interactions
- System functions

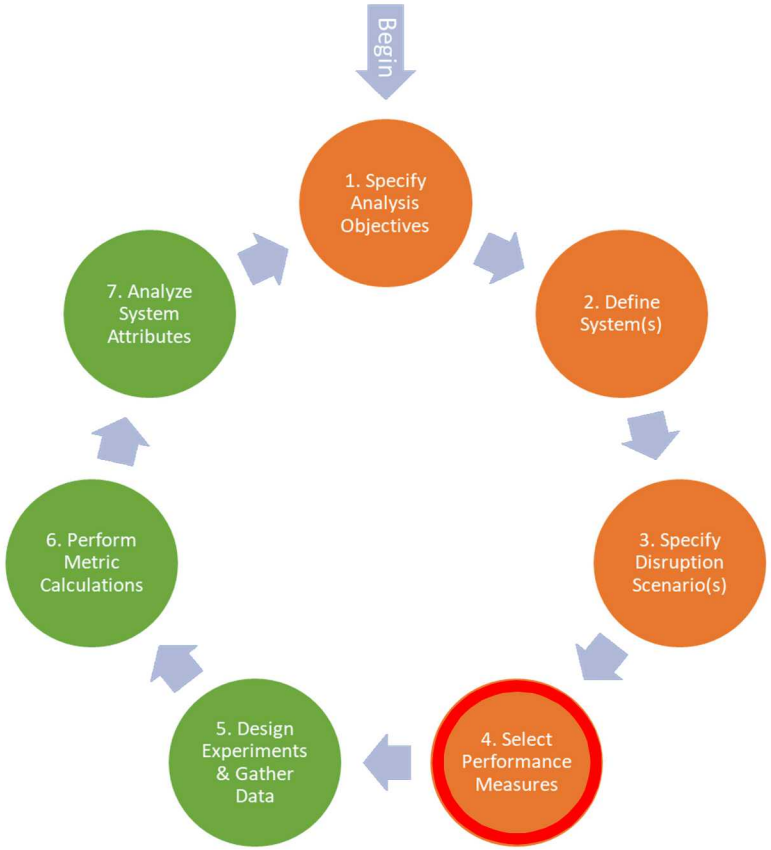
Step 3: Specify Disruption Scenario(s)



Describe the stressed conditions and how the system operates through them

- Specification of the disruption
- Effect
- Timing
- System response
- Uncertainties

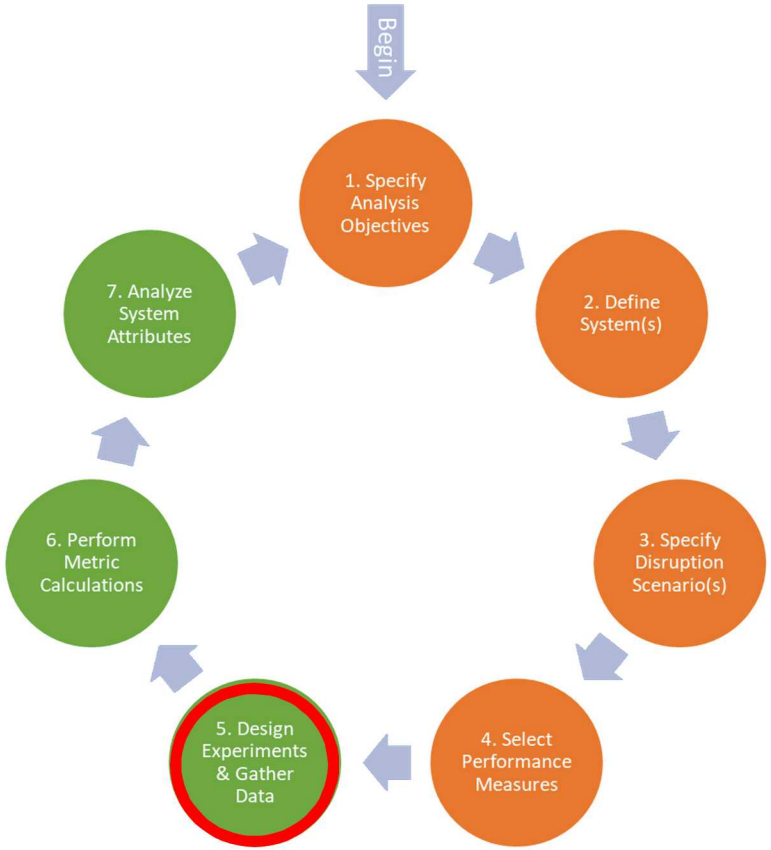
Step 4: Select Performance Measures



Describe the data that can be taken from the system to measure performance

- Target system performance
- Actual system performance
- Response and recovery efforts
- Relative weights of importance

Step 5: Design Experiments and Gather Data

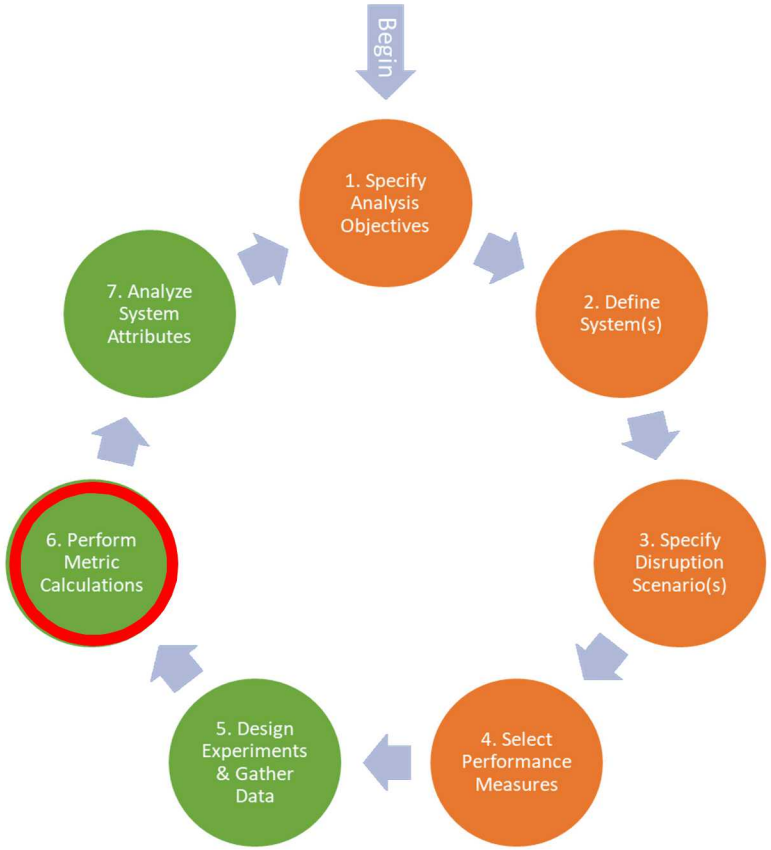


Determine how the scenarios can be tested against the system and data gathered

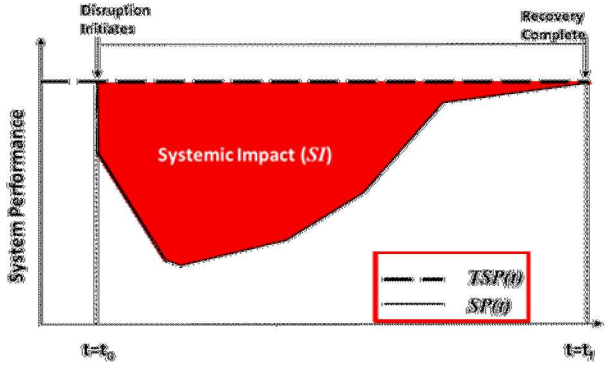
Selection of the experimental platform generally depends upon the resources available, time and budget, and analysis needs.

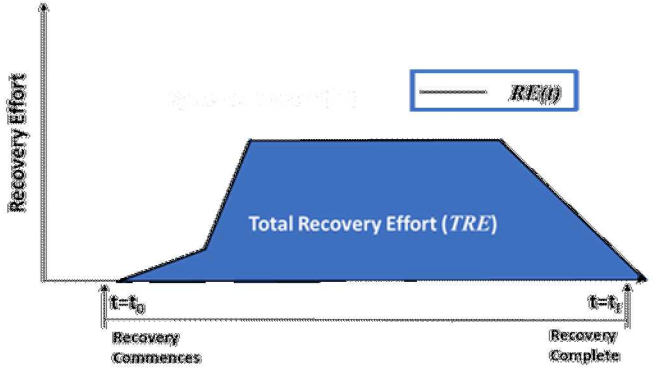
- Testbeds
- Emulation
- Modeling and simulation
- Historical events

Step 6: Perform Metric Calculations



Process the experimental data and performing the necessary calculations to populate resilience metrics



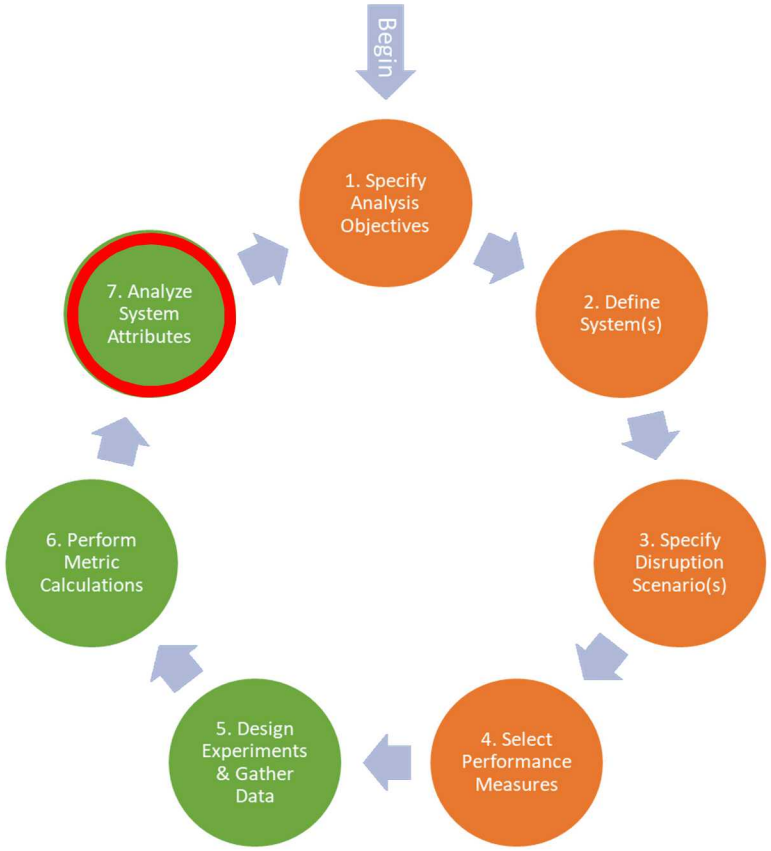


$$SI = B + \sum_{j} \int_{t_0}^{t_f} q_j(t)\left[TSP_j(t) - SP_j(t)\right] dt$$

$$TRE = C + \sum_{k} \int_{t_0}^{t_f} r_k(t)\left[RE_k(t)\right] dt$$

$$RDR = SI + \alpha TRE$$

Step 7: Analyze System Attributes



Use the quantitative results to identify resilience-limiting system properties and provide the basis for resilient design activities.

- Anticipate
- Absorb
- Adapt
- Restore

Applying IRAM to Evaluate Moving Target Defense

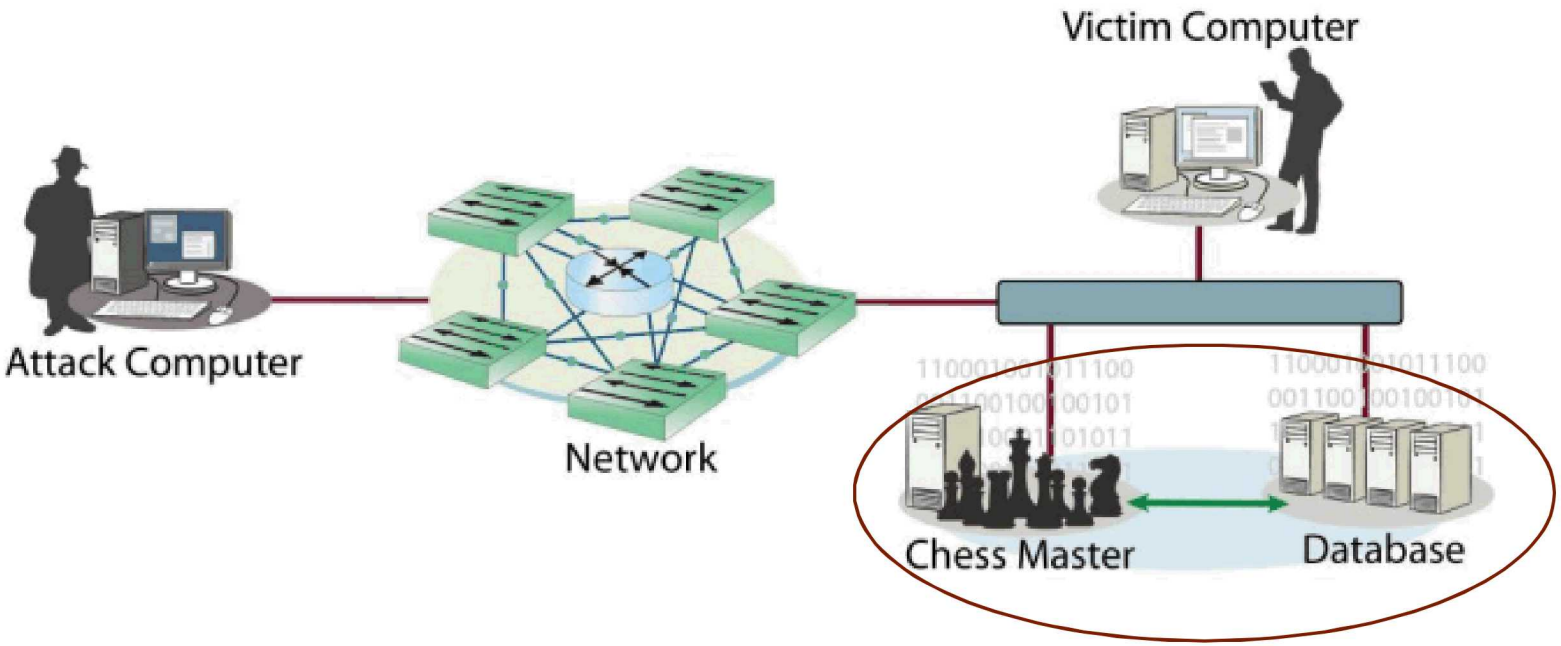


Image Source: https://www.energy.gov/sites/prod/files/2017/02/f34/SNL_ADDSec_Peer_Review_2016.pdf

Summary of IRAM Evaluation Results

- Energy systems are cyber attack targets; WANs are predictable and static

Does moving target defense effectively defend against reconnaissance and Ethernet-based attacks?

- ADDSec: Artificial Diversity and Defense Security (Chavez et al., 2016) employs MTD
  - Automatically reconfigures system with IP randomization and port hopping
  - Can detect attack and then randomize using machine learning algorithms

Does ADDSec make the system more resilient?

Using quantitative resilience metrics and analysis, results indicate:

ADDSec is worth the cost of implementation for our target system.

ADDSec does improve system resilience during a reconnaissance attack!

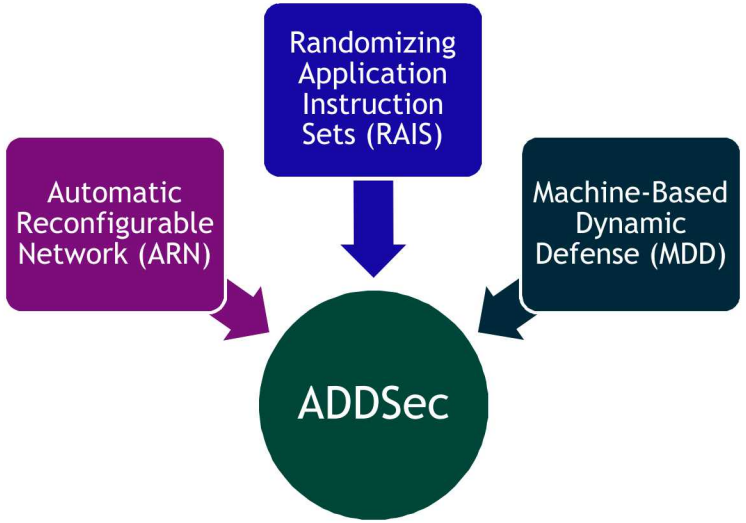
ADDSec Reference: https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2018/184545.pdf

ADDSec: Artificial Diversity and Defense Security

Grid WANs have predictable communication paths and static configurations

To introduce unpredictability and enhance situational awareness, Chavez et al. developed the ADDSec tool which leverages moving target defense (MTD)

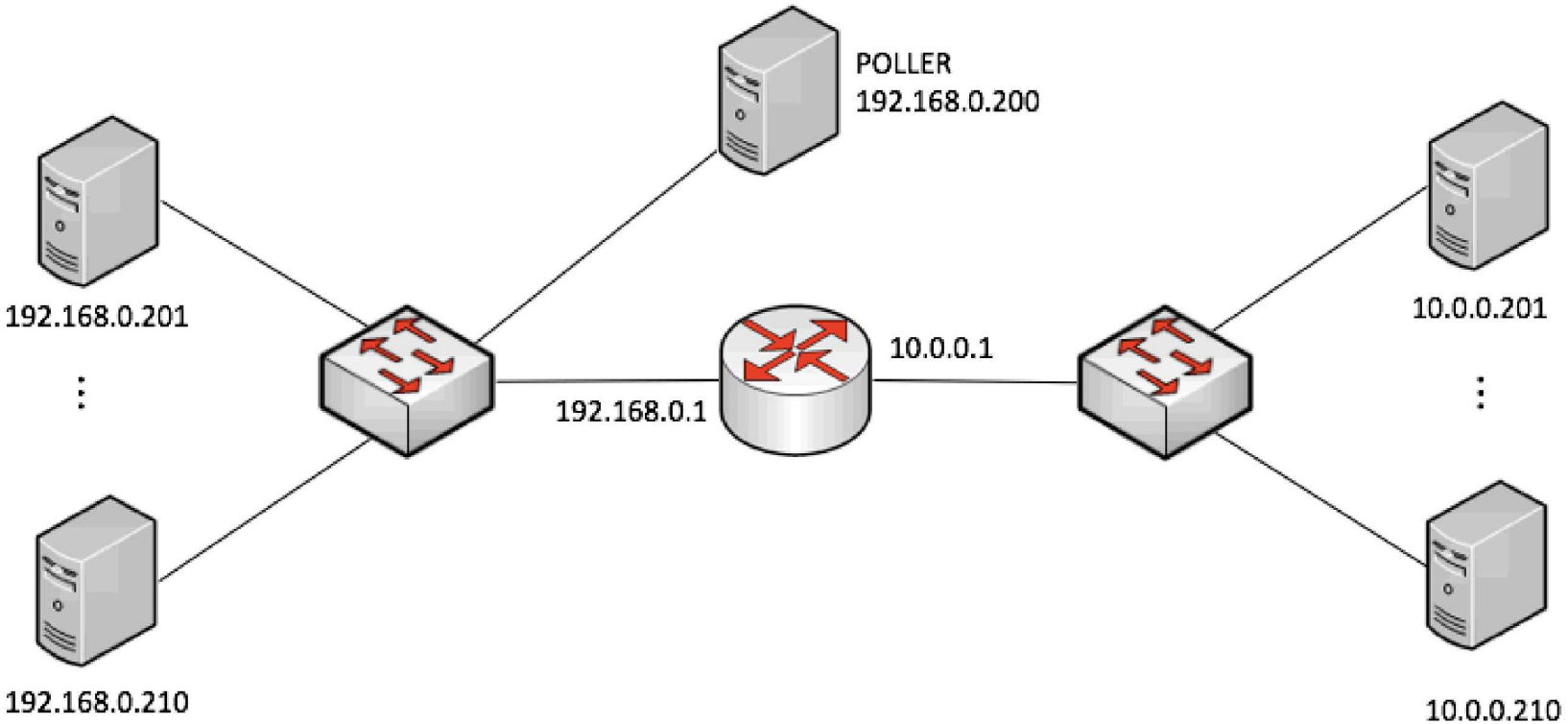
- Anticipates and adapts against reconnaissance and Ethernet-based attacks
- Enables automatic reconfiguration of the system through IP randomization and port hopping
- Machine learning algorithms applied to detect attacks and notify SDN controller to randomize



Research Questions Explored with IRAM

Key Questions:	1.1. Does ADDSec increase resilience of the system during an attack, specifically during reconnaissance?
	1.2. What performance does the system exhibit under different IP randomization rates?
	1.3. What performance does the system exhibit under different IP randomization rates during an attack?
	1.4. Are machine learning triggers effective for this type of attack?
	1.5. Do our resilience metrics provide useful insight into the effectiveness of ADDSec?

Experimental Setup



- Two subnets connected by router
- Total of twenty devices, ten on each subnet
- Poller periodically sends connection requests to each of the twenty devices
  - Maintains routing paths and provides basic monitoring

2. Define System(s)

Experiment Plan: ADDSec Modes and Attack Presence

Baseline: No ADDSec	• Worm • No Worm
Constant IP Randomization	• Worm • No Worm
Triggered IP Randomization	• Worm • No Worm

- Worm deployed on (an initially single) host(s) attempting to ping addresses and make connections
  - Scanning-based attack
  - Scans each subnet using ICMP requests to map active host addresses; when reply received, attempts to open secure TCP connection to target host
  - Once connection successfully established, worm attempts to self-replicate and continue to propagate

3. Specify Disruption Scenario(s)

Resilience Metrics

Measurement of resilience costs utilizes:

- Systemic Impact (**SI**): cumulative impact that a disruption has on system performance

$$SI = \Sigma_{i=1}^{N}[TSP(t_i) - SP(t_i)](t_i - t_{i-1})$$

- Total Recovery Effort (**TRE**): total resources used for recovery efforts post-disruption

$$TRE = \Sigma_{l=1}^{M}[RE(t_l)](t_l - t_{l-1})$$

Thus, the calculation of recovery-dependent resilience (**RDR**) cost is:

- Takes into account the effect the different recovery activities have

$$RDR = \frac{SI + \alpha \cdot TRE}{Norm}$$

Computing SI and TRE: Performance Metrics of Interest

Systemic Impact (SI)

Hosts Not Infected (#)

Total Recovery Efforts (TRE)

Latency (s)

Retransmitted Packets (#)

Dropped Packets (#)

4. Select Performance Measures

Summary of Results: System Metrics

		Frequency of IP Randomization									
Average over 10 trials (1000s/trial)		None	ML	1s	4s	8s	16s	32s	64s	128s	256s
# Host Infections	No Worm	-	-	-	-	-	-	-	-	-	-
	Worm	20	3	2.8	3.4	4.8	4.9	4.9	7.9	8.9	9.8
Latency	No Worm	29.93	37.2	349.34	394.71	699.11	591.89	TBD	422.1097	48.88403	420.31
	Worm	729.91	698.92	346.22	733.84	1000.42	1148.1	997	1187.3	1559.14	2351.07
Retransmits	No Worm	6039	5928.7	37.2	37.2	37.2	37.2	TBD	4291.8	6887	2681.3
	Worm	5417	2267.8	1966.1	2151.1	2451.9	2839.5	3911.3	6297.6	7182.3	3911.3
Dropped Packets	No Worm	0	0	0.1	0	0.1	0	TBD	0	0	0
	Worm	0	0.3	1	0.7	0.6	0.1	0	0.1	0	0

Summary of Results: Resilience Metrics

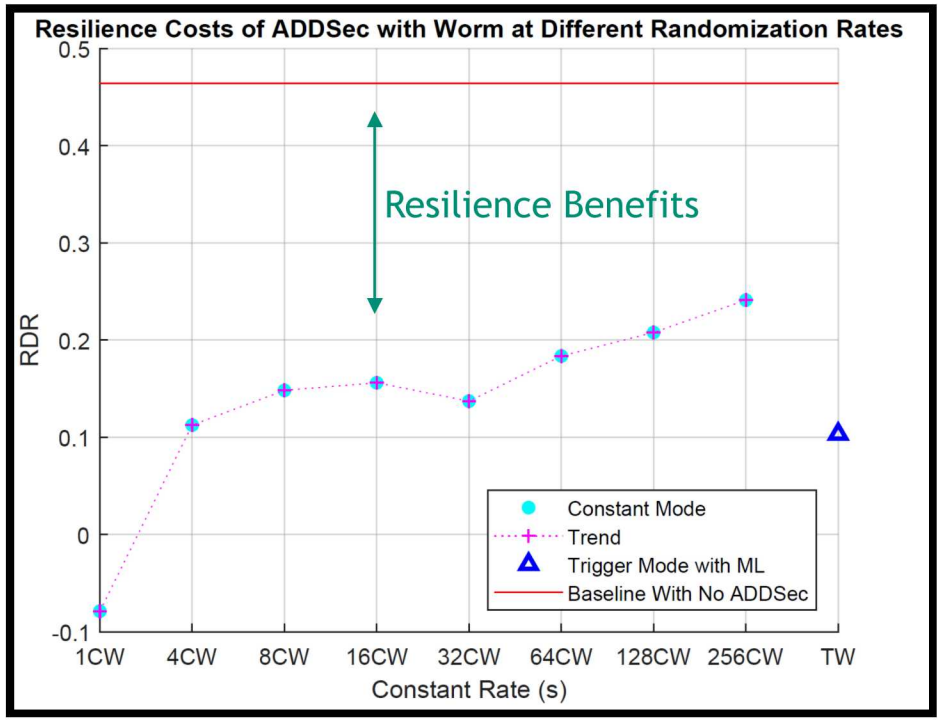
	Frequency of IP Randomization										
Average over 10 trials (1000s/trial)		None	ML	1s	4s	8s	16s	32s	64s	128s	256s
SI	No Worm	0	0	0	0	0	0	0	0	0	0
	Worm	0.65146	0.05773	0.05378	0.06091	0.08202	0.08524	0.08373	0.1331	0.15133	0.16696
TRE	No Worm	-0.00042	-0.00235	-0.00341	0.01331	0.02631	0.01751	TBD	0.0202	0.00094	0.0442
	Worm	-0.1872	0.04558	0.02497	0.05158	0.06614	0.07078	0.05336	0.0504	0.05643	0.07413
RDR	No Worm	0.00042	-0.00235	-0.00341	0.01331	0.02631	0.01751	TBD	0.0202	0.00094	0.0442
	Worm	0.46426	0.1033	-0.07874	0.11247	0.14817	0.15602	0.13709	0.18352	0.20777	0.24108

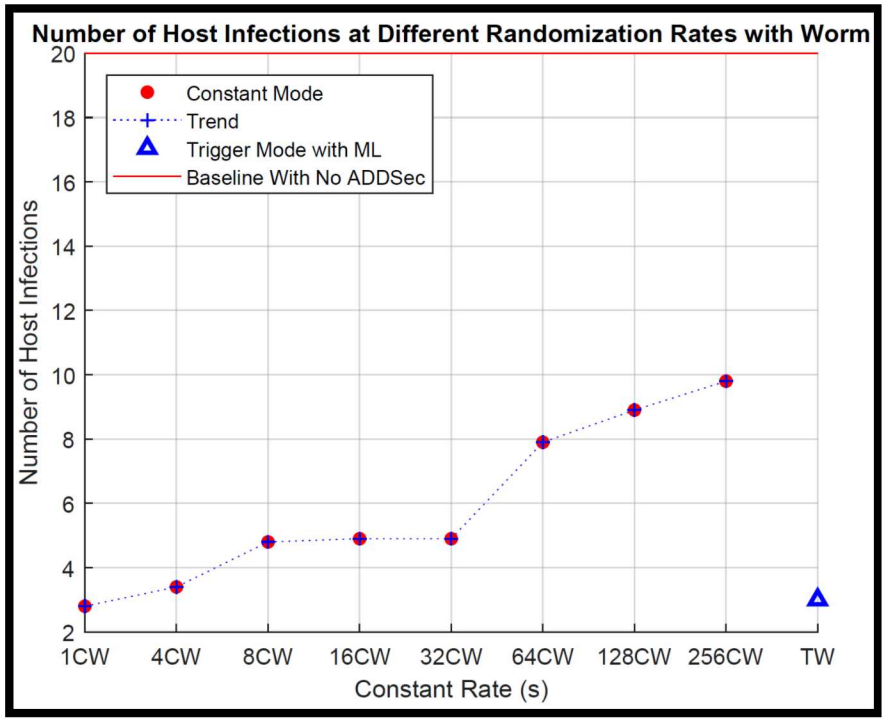
6. Perform Metric Calculations

Results

Key Question: 1.1 Does ADDSec increase resilience of the system during an attack, specifically during reconnaissance?

Yes! ADDSec improves resilience significantly.



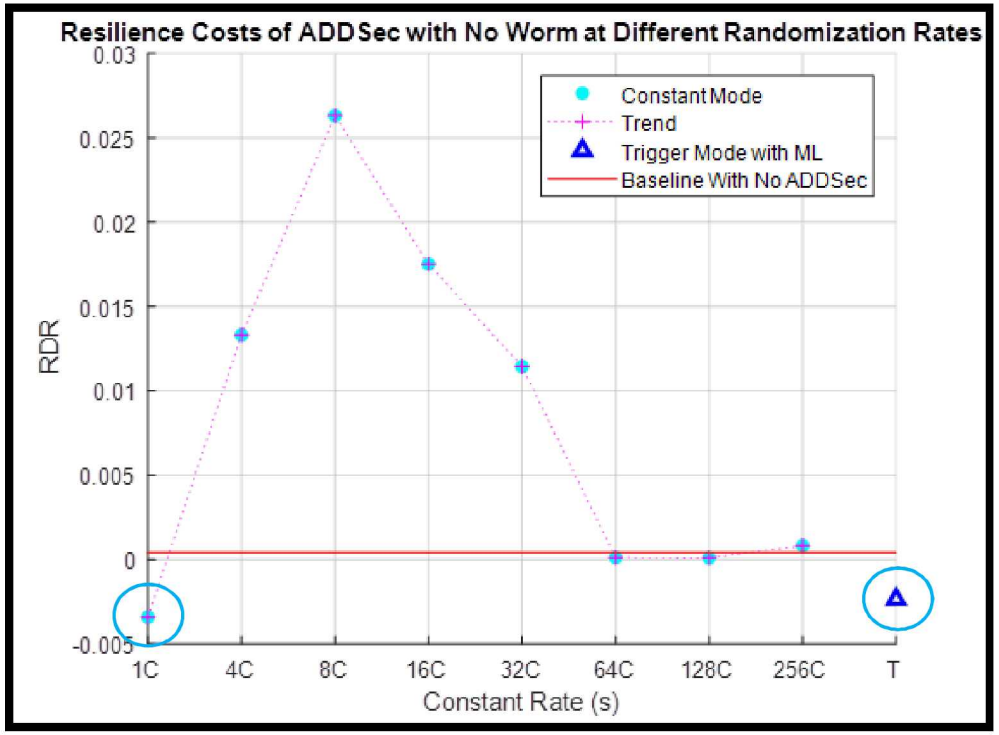


7. Analyze System Attributes

Results

Key Question: 2 What performance does the system exhibit under different IP randomization rates?

Constant 1s and Trigger Mode lower performance losses.

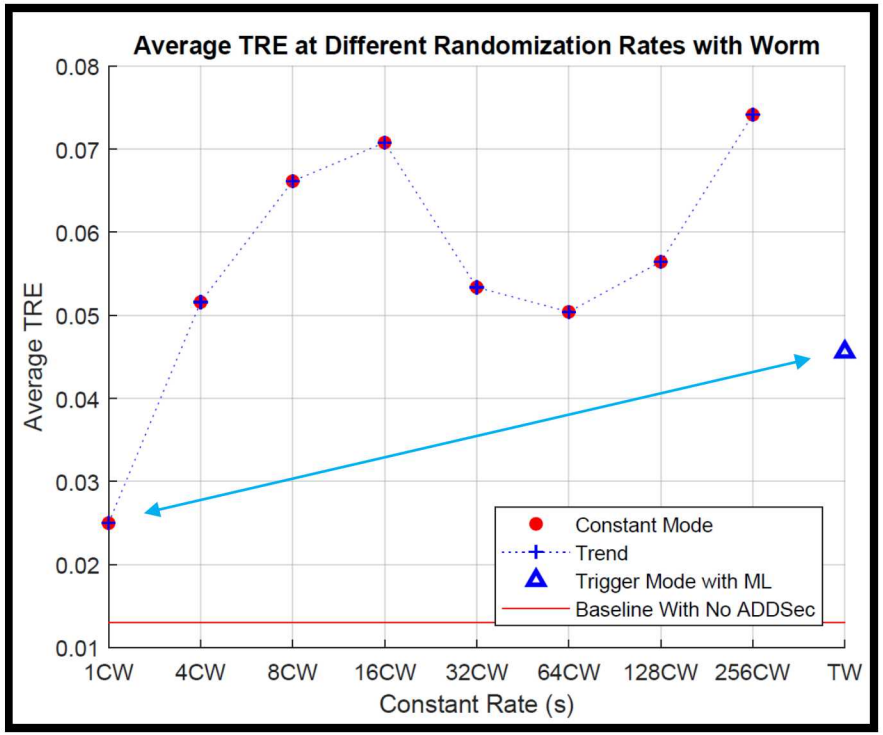


7. Analyze System Attributes

Results

Key Question: 1.3 What performance does the system exhibit under different IP randomization rates during an attack?

Constant 1s and Trigger Mode low performance overhead.

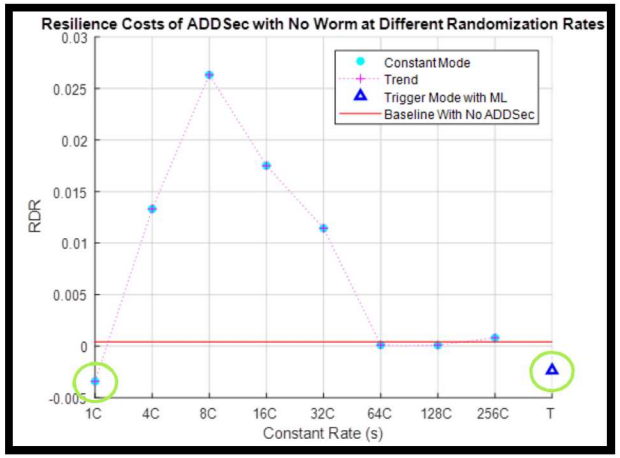


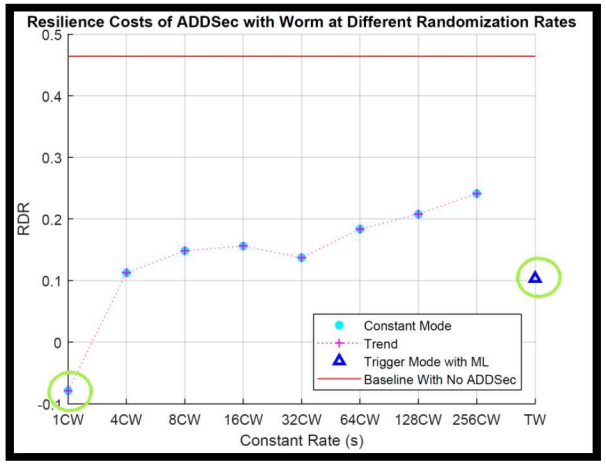
7. Analyze System Attributes

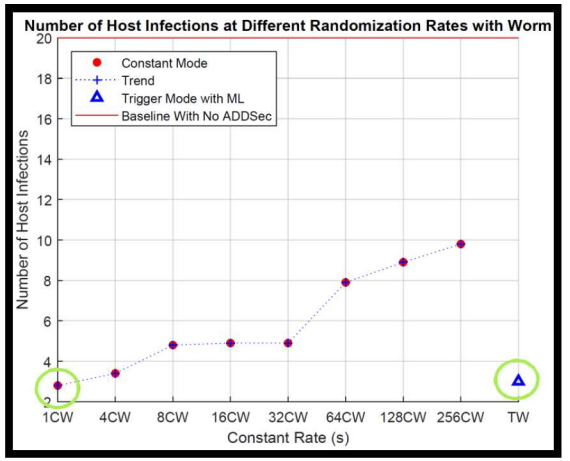
Results

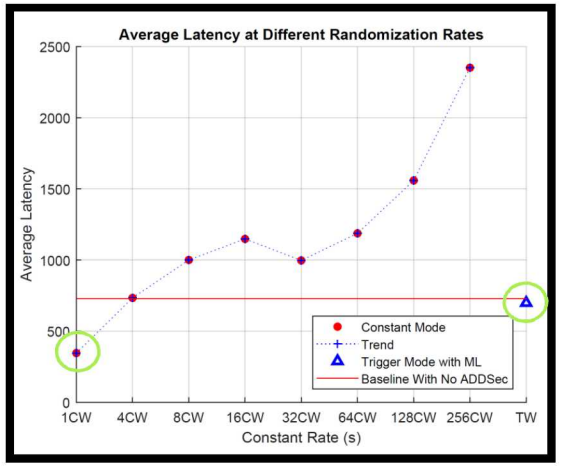
Key Question: 4 Are machine learning triggers effective for this type of attack?

Triggered randomization exhibited similar behavior to faster randomization rates; Constant 1s Mode always outperforms.







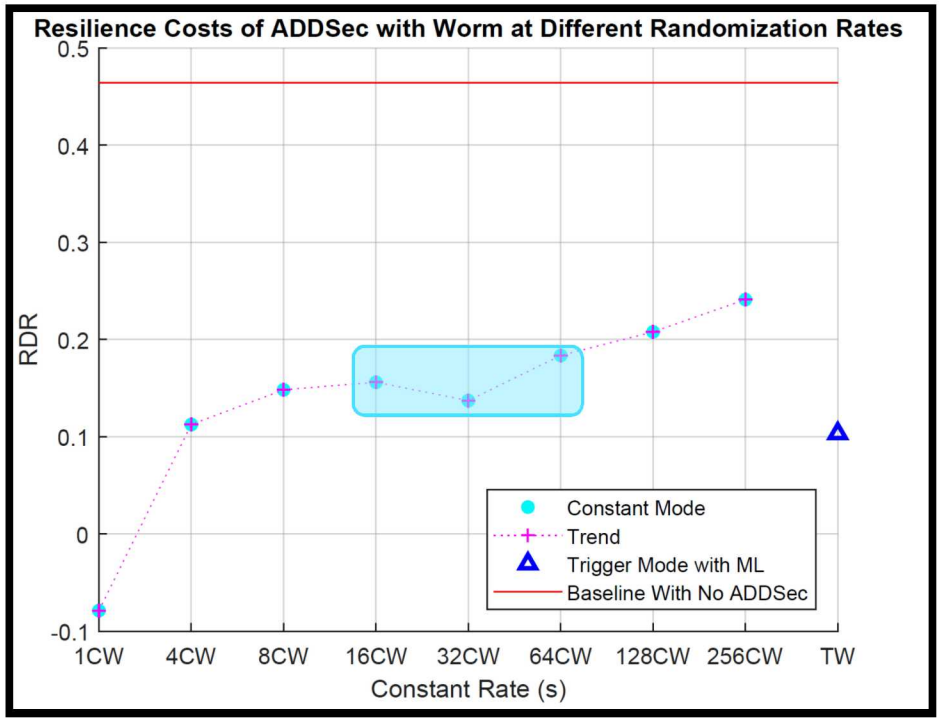


7. Analyze System Attributes

Results

Key Question: 5 Do our resilience metrics provide useful insight into the effectiveness of ADDSec?

Trends are seen in relation to ADDSec randomization rate/strategy; found that Constant 1s Mode most effective.



Key Takeaways

Resilience analysis provides useful insight into ADDSec performance and optimal modes

- SI metric captures infection impact to system dynamically, over time
- TRE metric can be tuned to give more weight to important quantities (e.g., latency > retransmits)
- RDR provides more granular insight that might be missed with only intuition (e.g., 32s case)

Automated triggers can be effective

- Reconnaissance activity is stopped even during period of the randomization rate
- Higher resilience than constant rate
- Caveat: algorithms need to be tuned to detect the attack

IP randomization is effective but subject to variability

- Quantitative analysis shows that faster randomization rates improve resilience on average
  - Increasing randomization decreases number of infected hosts and time to first infection
- Stochastic behavior means that there is no guarantee of improved resilience with faster randomization

Paper Reference: S. Hossain-McKenzie, C. Lai, A. Chavez and E. Vugrin, "Performance-Based Cyber Resilience Metrics: An Applied Demonstration Toward Moving Target Defense," *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, Washington, DC, 2018, pp. 766-773. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8591764&isnumber=8591058

Thanks! Questions?

Cyber Resilience

Many critical systems are the target of evolving, sophisticated attacks

- Cannot stop every attack – need to improve **cyber resilience**

Vugrin et al. on resilience:

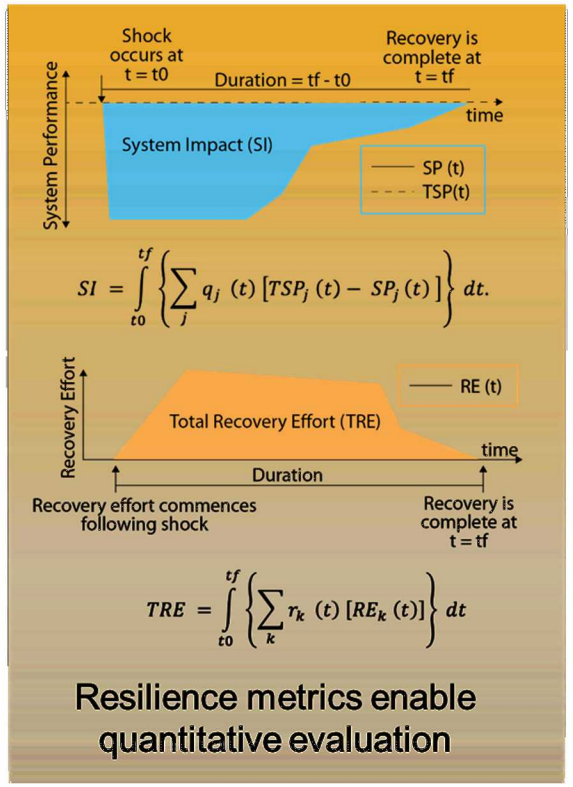
- Given one or more disruptive event(s), resilience describes the system's ability to reduce the magnitude and duration of deviation from targeted performance levels

Quantitatively evaluate resilience features such as ADDSec to make informed decisions by examining:

- Effectiveness of tool during a disruption
- Impact on normal system operations
- Resilience costs of different implementation strategies

Informally, cyber resilient systems are able to execute required mission parameters despite a hostile cyber-threat environment.

Cyber Resilience Framework: Elements



Resilient design principles foster solution development

	Anticipative Capacity	Absorptive Capacity	Adaptive Capacity	Restorative Capacity
Directly Impacts	SI & TRE	Systemic Impact	Primarily Systemic Impact, but also TRE	Total Recovery Effort
Distinguishing Features	Expedited threat ID and sensing; catalyst for other capacities	Automatic manifestation after disruption	Reorganization and change from standard operating procedures	System repair
Temporal Sequencing	Pre-/during attack	First line of defense	Second line of defense	Final line of defense
Post-disruption effort required	Constant	Automatic/little effort	Increased effort	Greatest effort
Duration of changes	Constant	Permanent	Temporary	Permanent
Resilience enhancement features: cyber examples	Intrusion detection system, surveillance, data analytics, ML, endpoint verification, diversion	Hardening, redundancy, diversity, decentralization, distributed ledgers, segmentation, encryption, excess capacity	Moving target; deception; adaptive controls; substitution; Active Malware Countermeasure rerouting; conservation; reorganization; ingenuity	Graceful degradation; self-healing; reconstitution; forensics;

ADDSec Machine Learning

Machine learning algorithms are deployed to each host

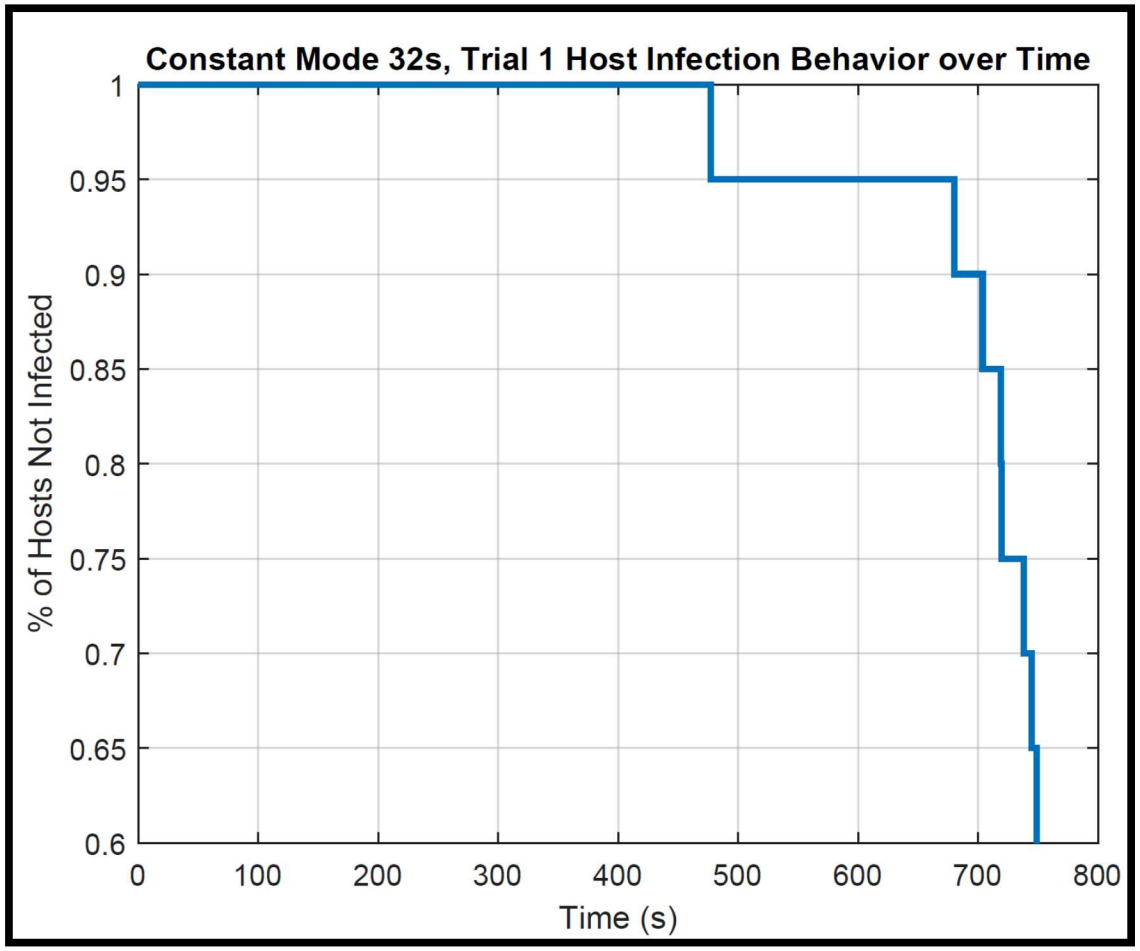
Features extracted from logs on each host:

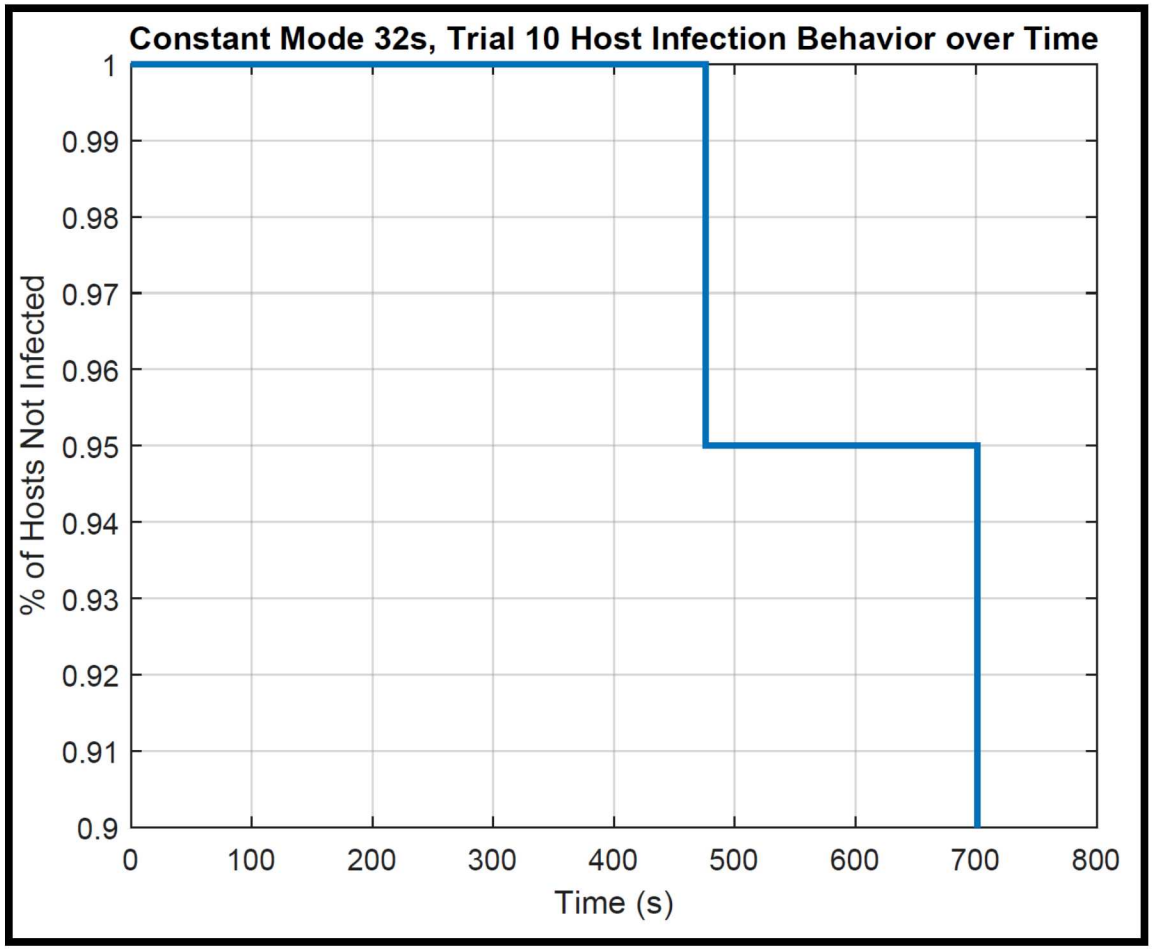
- System status and performance statistics
- System call stack
- Packet capture, Bro network analytics

Classification is performed by an ensemble of techniques (primarily decision trees)

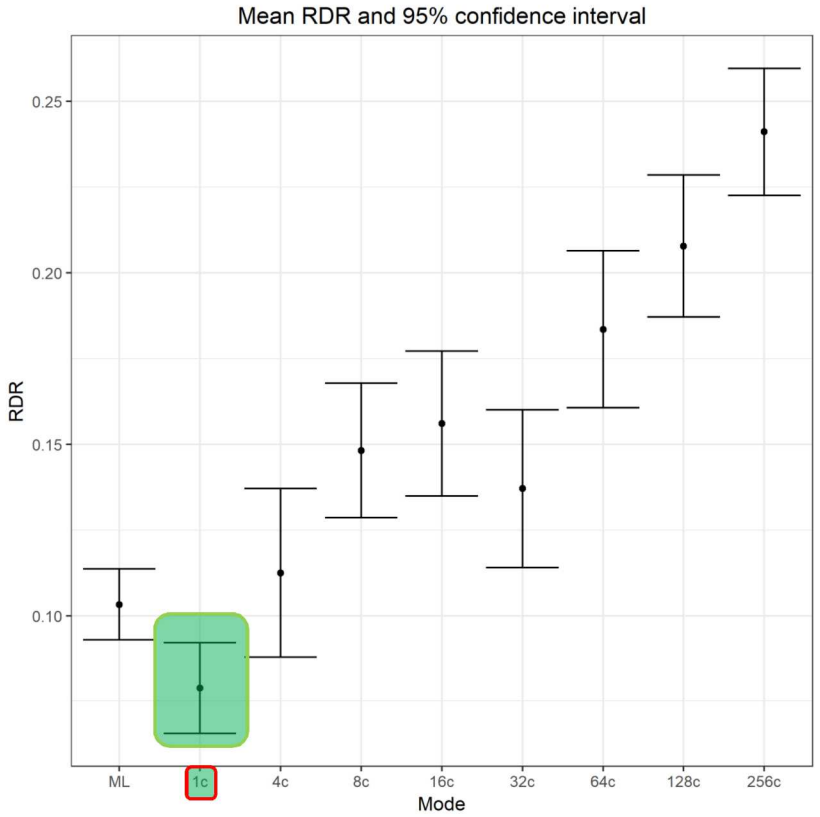
When the machine learning is first turned on, a baseline is taken. The feature set is periodically compared against a baseline and if an alert is triggered, a signal is sent to the controller to undergo randomization.

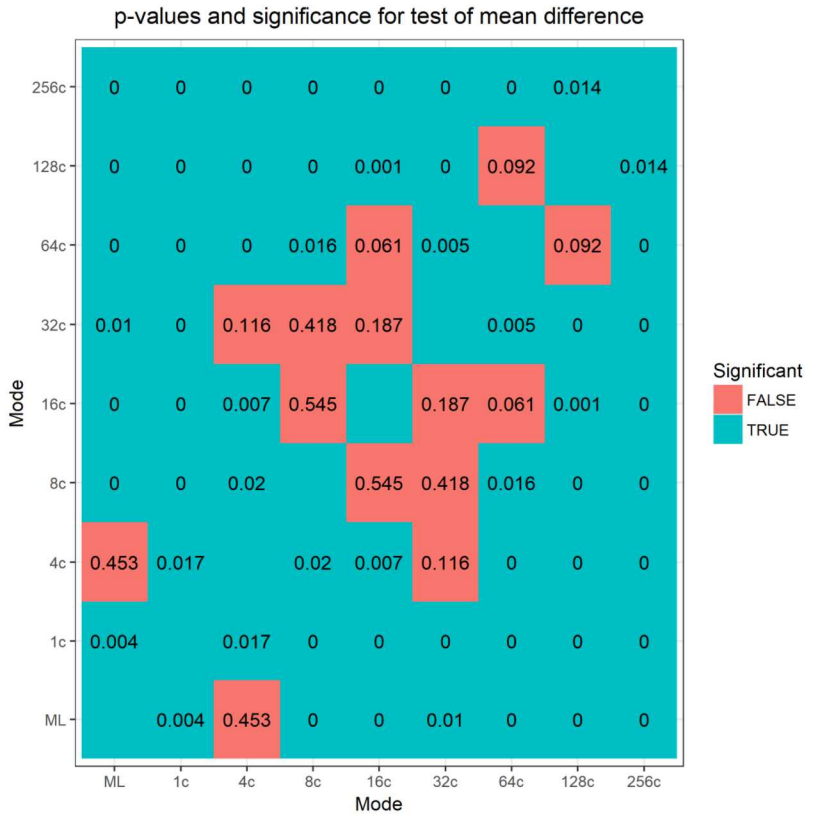
ADDSec Exhibits Stochastic Behavior

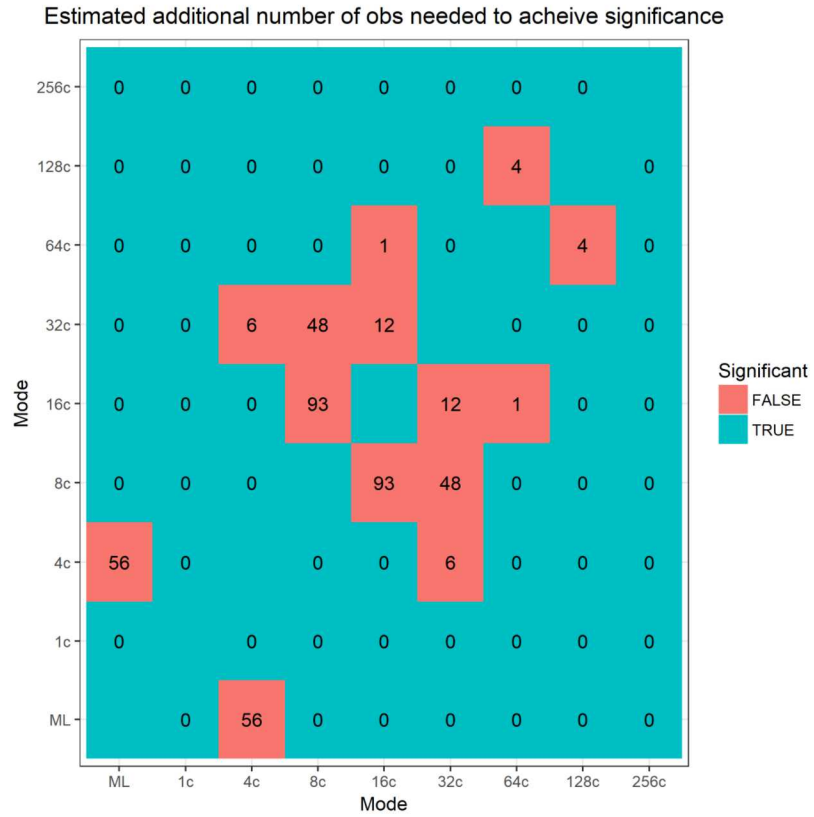




Testing for Significant Differences in RDR







Lessons Learned and Future Experiments

Pre-processing took substantial effort

- Automated many processes compared to initial ADDSec analysis

ADDSec behavior stochastic, needed to collect more data to see more clear trend

- Difference-in-mean analysis useful for understanding results and if more data needed
- Gained insight into how to best improve ADDSec behavior:
  - For a predictable scan, randomize among IP ranges that have already been scanned or are not initially scanned.

Significant effort spent on debugging experiment, determining good data collection strategy and selecting metrics

- Emulation requires more resources than simulation – deploy experiments on bigger cluster
- VM resources need to be tuned so that machine learning buffers do not cause crashes
- Future experiments could be automated with time-based scripts – or port experiment to Firewheel which has time triggers