

Integrated Cyber Analysis System (ICAS)* Proposers' Day

Mr. Richard Guidorizzi
Program Manager
DARPA, I2O

Accelerating the Discovery Process

The anticipated ICAS BAA posted at www.fbo.gov will take precedence over the information provided herein.

January 30, 2013



*Proposers' Day was held under the program's prior name, Cyber Targeted-Attack Analyzer (CAT).



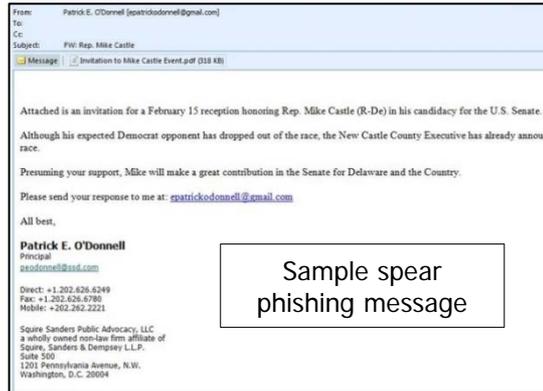
What type of information (clues) are left by attacks?

There are many publically documented accounts that describe the specific details of cyber attacks and cyber targeted attacks, such as:

- Analyzing Project Blitzkrieg, a Credible Threat (9 Sep 2012)
<http://www.mcafee.com/us/resources/white-papers/wp-analyzing-project-blitzkrieg.pdf>
- Black Hat USA 2012 Presentation – Targeted Intrusion Remediation: Lessons from the Front Lines (6 Aug 2012)
<https://www.mandiant.com/blog/black-hat-usa-2012-presentation-targeted-intrusion-remediation-lessons-front-lines/>
- How Apple and Amazon Security Flaws Led to My Epic Hacking (6 Aug 2012)
<http://www.wired.com/gadgetlab/2012/08/apple-amazon-mat-honan-hacking/>
- Behold the Anonymous/HBGary saga e-book: Unmasked (10 Mar 2011)
<http://arstechnica.com/tech-policy/2011/03/hbgaryanonymous-special-report/>

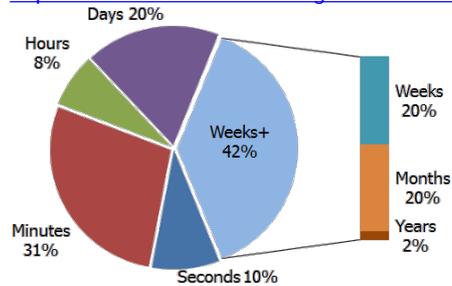


Discovery of Cyber Attacks is Slow



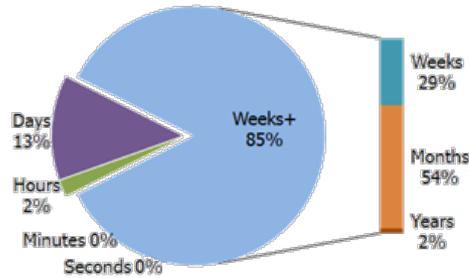
- Malware often enters through spear phishing, where email messages are sent appearing to be from someone within the organization.
- The messages come with apparently legitimate attachments that contain embedded malware products.

<http://www.f-secure.com/weblog/archives/00001908.html>



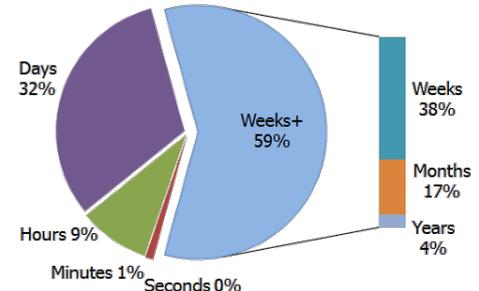
~ 38 days¹

Initial Infection



~ 82 days¹

Compromised



~ 47 days¹

Reconstitution

~ 120 days before defenders are aware of a cyber attack (averaged across the 2500 cases)¹

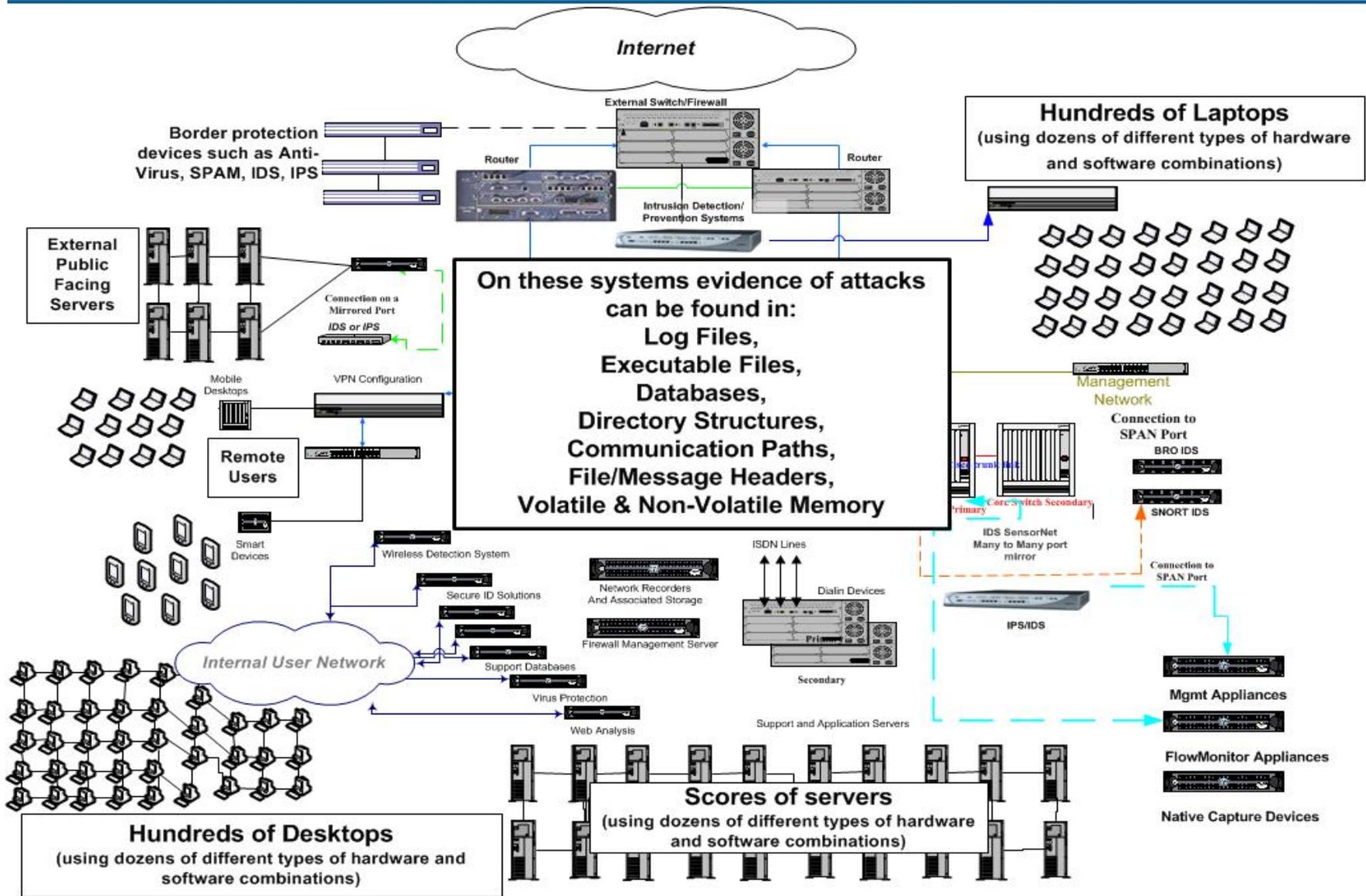
Between 2004 and 2012, US Secret Service and Verizon have evaluated more than **2,500 criminal cases** involving data loss compromising more than one billion records.¹

In 84% of the confirmed breaches **clear evidence** of the breach was found in local log files and **was not** discovered until after the breach.¹

¹ - 2010 - 2012 Verizon Data Breach Investigations Reports



Detecting the Signs of an Attack Requires Correlation Across Many Data Sources





Why is DARPA focusing on this?

Industry focused first on large scale attacks

Cyber Attack

I LOVE YOU Virus (2000)

CodeRed (2001)

MyTob (2005)

Storm Botnet (2007)

Koobface (2008)

Downadup (aka Conficker) worm (2009)

Bredolab Botnet (2010)

Mariposa botnet (2010)

Ponmocup Trojan (2011)

Mac Flashback Trojan (2012)

Scale of the problem space

\$5-10B in clean up costs

400,000 servers

20m machines (peak)

50m machines (peak)

500,000 machines active in 2011

1.7m machines active in 2011

30m machines

12m machines

1.2m unique IPs in 24hrs

600,000 machines

Targeted attacks

Handfuls of machines

Targeted attacks are too small to justify commercial investment

Sources:

blog.fortinet.com, 40th Anniversary of the Computer Virus, March 10, 2011

www.theregister.co.uk, How FBI, police busted massive botnet, March 3, 2010, 15:56 GMT

www.macworld.com, Security experts: 600,000-plus estimate of Mac botnet likely on target, April 6, 2012, 4:00 PM

www.wired.com, Bredolab Bot Herder Gets 4 Years for 30 Million Infections, May 23, 2012, 2:06 PM



Solution: Integrated Cyber Analysis System (ICAS)

Federate all potentially useful forensic data.

- Automatically index data sources on the network.
- Integrate all data structures (unstructured, disparate) through a common language.

Develop tools for reasoning over the federated database.

- Detect interactions and behavior patterns across all data sources and dimensions.
- Find aberrant machine-machine connections.
- Uncover enterprise-wide compromise by revealing repeated patterns.

Program Goal: Reduce Cyber Attack Discovery Time



Federate data sources

Challenge 1

Automatically index data sources on the network without Human Intervention.

Reasons for Confidence:

Several successful examples of automatic resource discovery in closely related fields exist:

- NMAP (a network & port scanning product) was expanded under CINDER to include automation of some capabilities.
- Performers are starting to include automated correlation of some network data (such as network events to host events).

Challenge 2

Integrate all data structures (unstructured, disparate) through a common language for Security Data.

Reasons for Confidence:

Initial research into developing a common language for reasoning about security shows promise:

- DARPA has seen several performers start to develop limited semantic security ontologies to connect security-related information that fulfills a subset of our needs.

Tool development

Challenge 3

Develop tools for reasoning over the federated database.

Reasons for Confidence:

Automated translation technology for federating otherwise incompatible data sources is an area of active research:

- Automatic migration & wrapping of database applications, a schema transformation approach. (McBrien, 1999)
- Automatic Wrapper Generation and Maintenance. (Xia, 2011)
- Automatic Wrapper Generation for Search Engines Based on Visual Representation. (Rao, 2012)



What functions are covered in the anticipated solicitation?

- Under the anticipated solicitation, DARPA plans to develop a “solution” that meets the described program objectives and that can be transitioned. This may require the development of more than one application.
- The anticipated solicitation will be looking for performers in the following functional areas:
 - Technical Area 1: Application Development
 - Technical Area 2: Security Testing and Validation (Red Team)



Integrated Cyber Analysis System (ICAS) Program Activities

- The solicitation is expected to come out in late February.
- The solicitation is currently expected to be open for 50 calendar days.
- Multiple awards are expected for Technical Area #1.
- One award is expected for Technical Area #2.



Technical Area #1

Application Development

- Develop an integrated suite of applications that performs all actions required to meet the program goals (the “ICAS solution”).
- Proposals are expected to include a one year base, a one year option, and a final six month option.
- Total funding for this Technical Area is expected to be approximately \$9.6M over 30 months.
- DARPA anticipates a breakdown of the work and funding to be: 34% in FY 2013; 52% in FY 2014; and 14% in FY2015.
- DARPA expects performers in this technical area to perform demonstrations of the developed solution(s) every quarter during the program, with major demonstrations performed at the end of each year, demonstrating full capabilities.



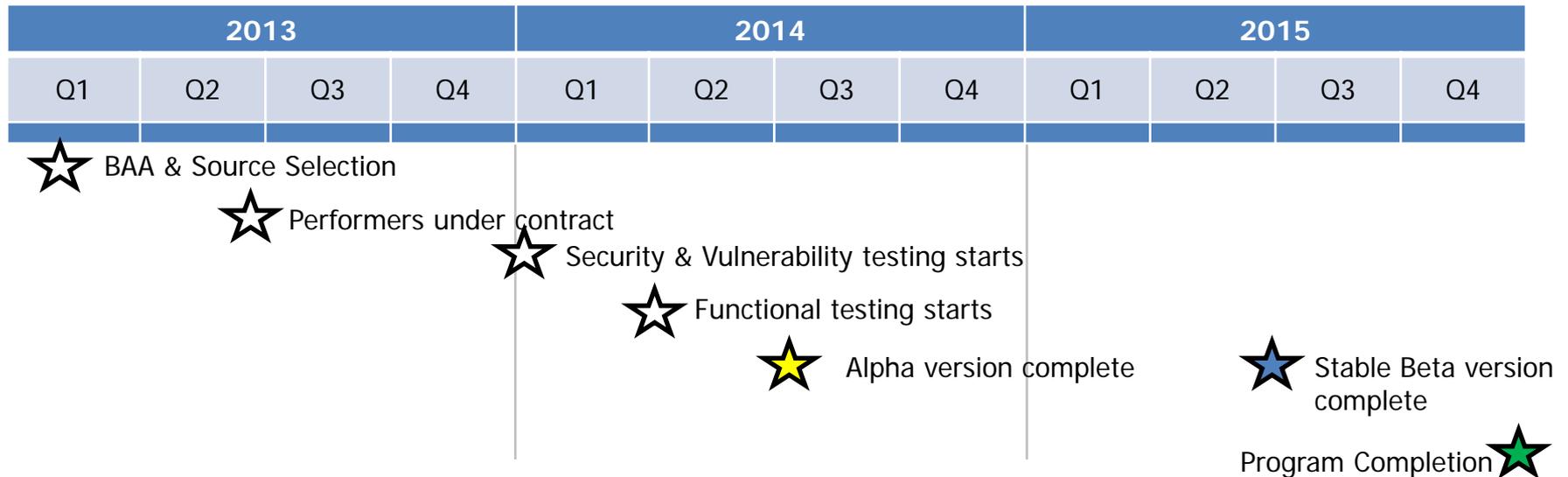
Technical Area #2 Security Testing and Validation (Red Team)

- Provide Red Teaming or “Adversarial Partner” subject matter expertise for Phases 1 and 2, providing a realistic picture of potential risk introduced with the solution(s) during development.
- Proposals are expected to include a one year base and a one year option.
- Total funding for this Technical Area is expected to be approximately \$1.8M over 24 months.
- DARPA anticipates a breakdown of the work and funding to be: 44% in FY 2013 and 56% in FY 2014.
- Red Team analysis is being included for all solution(s) developed under the ICAS program to ensure they are as secure as possible.



Anticipated ICAS Program Plan

- Phase 1 (first year):
 - Develop Alpha version of the solution(s).
- Phase 2 (second year):
 - Develop Beta version of the solution(s).
- Phase 3 (final six months):
 - Transition developed solution(s).





www.darpa.mil