GOOD GUYS VS BAD GUYS:

USING BIG DATA TO COUNTERACT ADVANCED THREATS

Joe Goldberg
Splunk

RSA CONFERENCE EUROPE 2013

Session ID: SPO-W09

Session Classification: Intermediate

About Me

Joe Goldberg

Current:

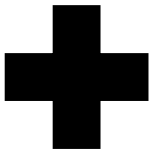
- Splunk - Security Evangelist and Technical Product Marketing

Past:

- Symantec/Vontu - Data Loss Prevention, Technical Product Marketing
- VMware - Technical Product Marketing
- Sun Microsystems - Product Marketing

Security Presentation Template

Scare them

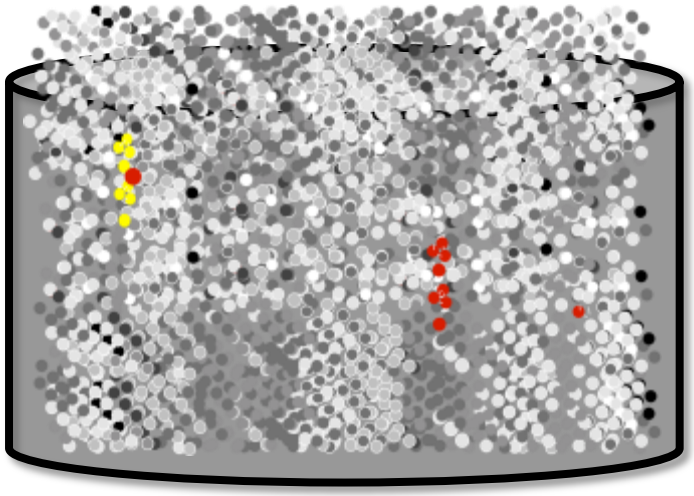


Unscare them

Security Presentation Template

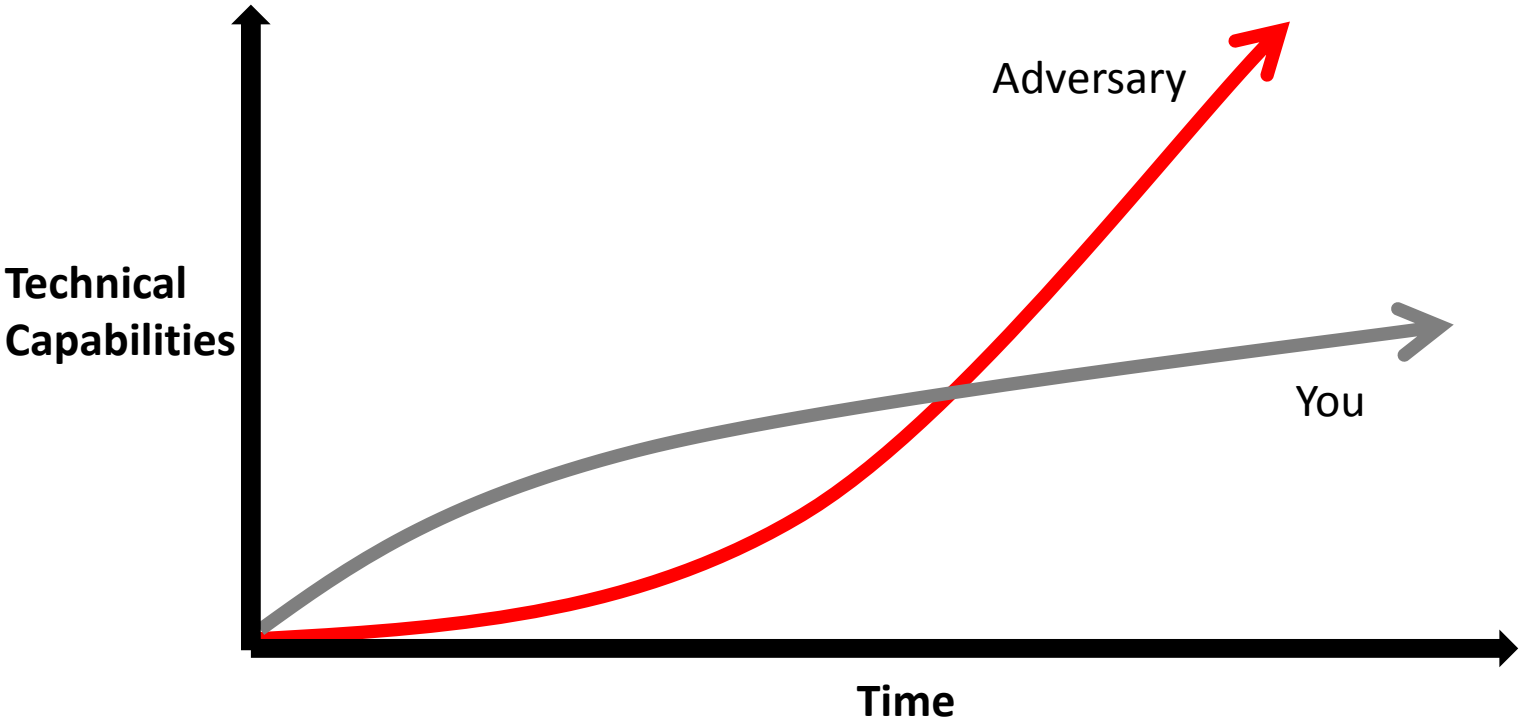
Advanced Threat

Big Data



Here Comes the Scary Part…..

Advanced Threats Outpace the Defenders



Advanced Threats Are Hard to Detect

100%

Valid credentials were used

243

Median # of days before detection

40

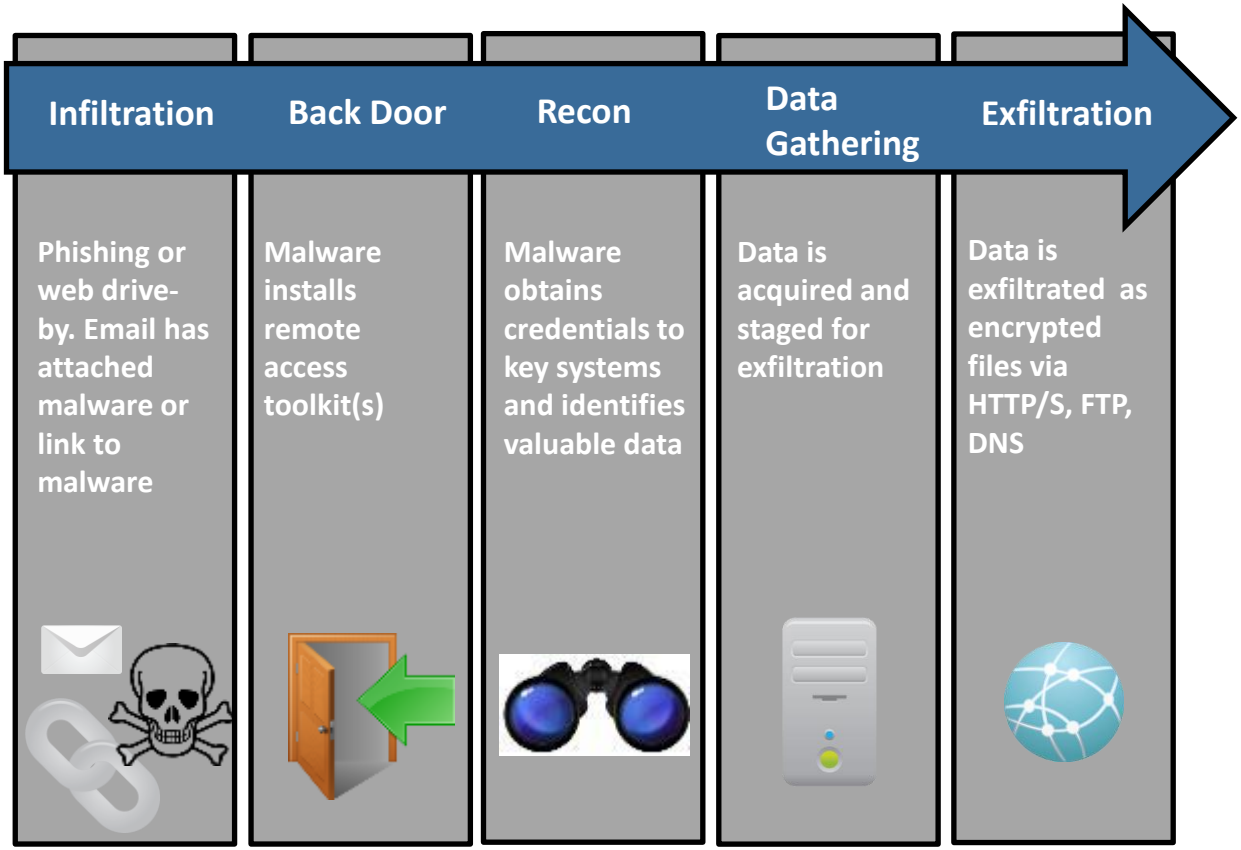
Average # of systems accessed

63%

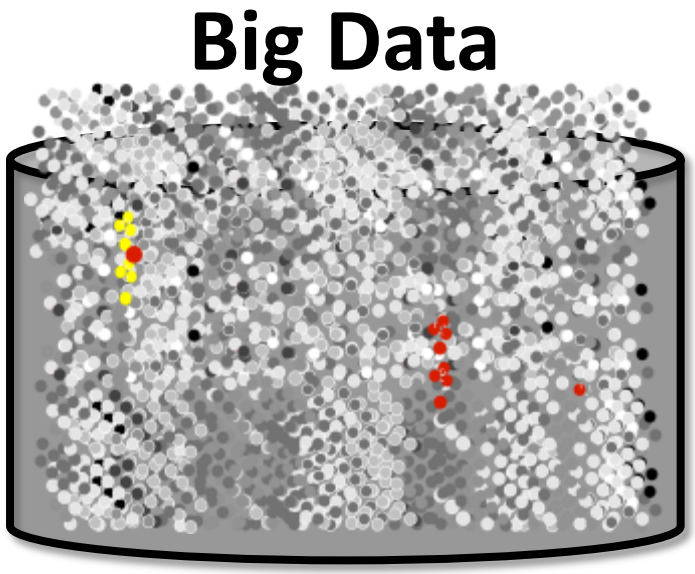
Of victims were notified by external entity

Source: Mandiant M-Trends Report 2012 and 2013

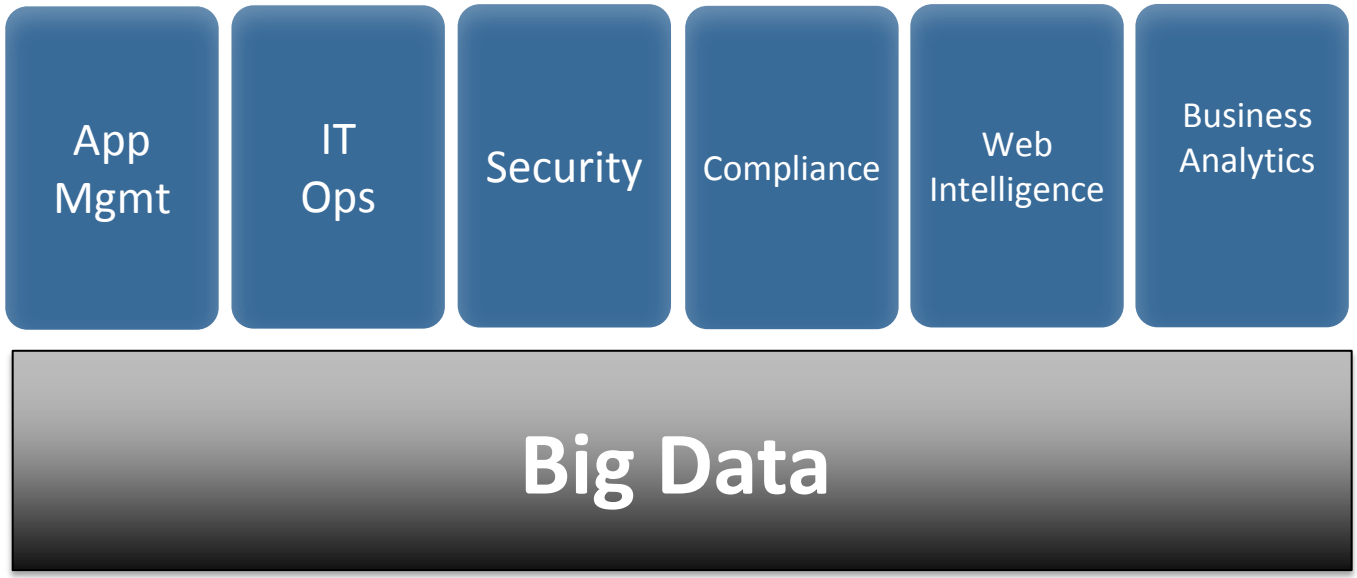
Advanced Threat Pattern



Here Comes The Solution



Big Data is Used Across IT and the Business



“Big Data” Definition

- Wikipedia: Collection of data sets so large and complex that it becomes difficult to process using database management tools
- Gartner: The Three Vs
  - Data volume
  - Data variety
  - Data velocity
- Security has always been a Big Data problem; now it has a solution

Machine Data / Logs are Big Data

Sources

Email Server

2013-08-09T12:40:25.475Z,,exch-hub-den-01,,exch-mbx-cup-00,,,STOREDRIVER,DELIVER,79426,<20130809050115.18154.11234@acme.com>,johndoe@acme.com,,685191,1,,, hacker@neverseenbefore.com , Please open this attachment with payroll information,, ,2013-08-09T22:40:24.975Z

Web Proxy

2013-08-09 16:21:38 10.11.36.29 98483 148 TCP_HIT 200 200 0 622 - - OBSERVED GET www.neverbeenseenbefore.com HTTP/1.1 0 "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 2.0.50727; InfoPath.1; MS-RTC LM 8; .NET CLR 1.1.4322; .NET CLR 3.0.4506.2152;) User John Doe,"

Endpoint Logs

20130806041221.000000Caption=ACME-2975EB\Administrator Description=Built-in account for administering the computer/domainDomain=ACME-2975EB InstallDate=NULLLocalAccount = IP: 10.11.36.20 TrueName=Administrator SID =S-1-5-21-1715567821-926492609-725345543 500SIDType=1 Status=Degradedwmi_ type=UserAccounts

Windows Authentication

08/09/2013 16:23:51.0128event_status="(0)The operation completed successfully. "pid=1300 process_image="\John Doe\Device\HarddiskVolume1\Windows\System32\neverseenbefore.exe“ registry_type ="CreateKey"key_path="\REGISTRY\MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\ Printers Print\Providers\ John Doe-PC\Printers\{}\ NeverSeenbefore" data_type""

Endpoint Security

Aug 08 06:09:13 acmesep01.acmetech.com Aug 09 06:17:24 SymantecServer acmesep01: Virus found,Computer name: ACME-002,Source: Real Time Scan,Risk name: Hackertool.rootkit,Occurrences: 1,C:/Documents and Settings/smithe/Local Settings/Temp/evil.tmp,"""",Actual action: Quarantined,Requested action: Cleaned, time: 2009-01-23 03:19:12,Inserted: 2009-01-23 03:20:12,End: 2009-01-23 03:19:12,Domain: Default,Group: My Company\ACME Remote,Server: acmesep01,User: smithe,Source computer: ,Source IP: 10.11.36.20

Big Data Analytics

"[Security teams need] an analytical engine to sift through massive amounts of real-time and historical data at high speeds to develop trending on user and system activity and reveal anomalies that indicate compromise."

Security for Business Innovation Council report, "When Advanced Persistent Threats Go Mainstream,"

Chuck Hollis
VP – CTO, EMC Corporation

"The core of the most effective [advanced threat] response appears to be a new breed of security analytics that help quickly detect anomalous patterns -- basically power tools in the hands of a new and important sub-category of data scientists: the security analytics expert.."

Step 1: Collect *ALL* The Data in One Location


Databases

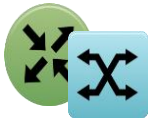

Email


Web


Desktops

Servers


DHCP/ DNS


Network Flows

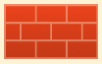

Hypervisor


Badges


Storage


Mobile

Traditional SIEM


Firewall


Authentication


Vulnerability Scans

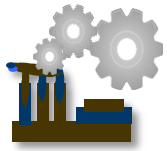

Intrusion Detection


Data Loss Prevention

Anti-Malware

Custom Apps

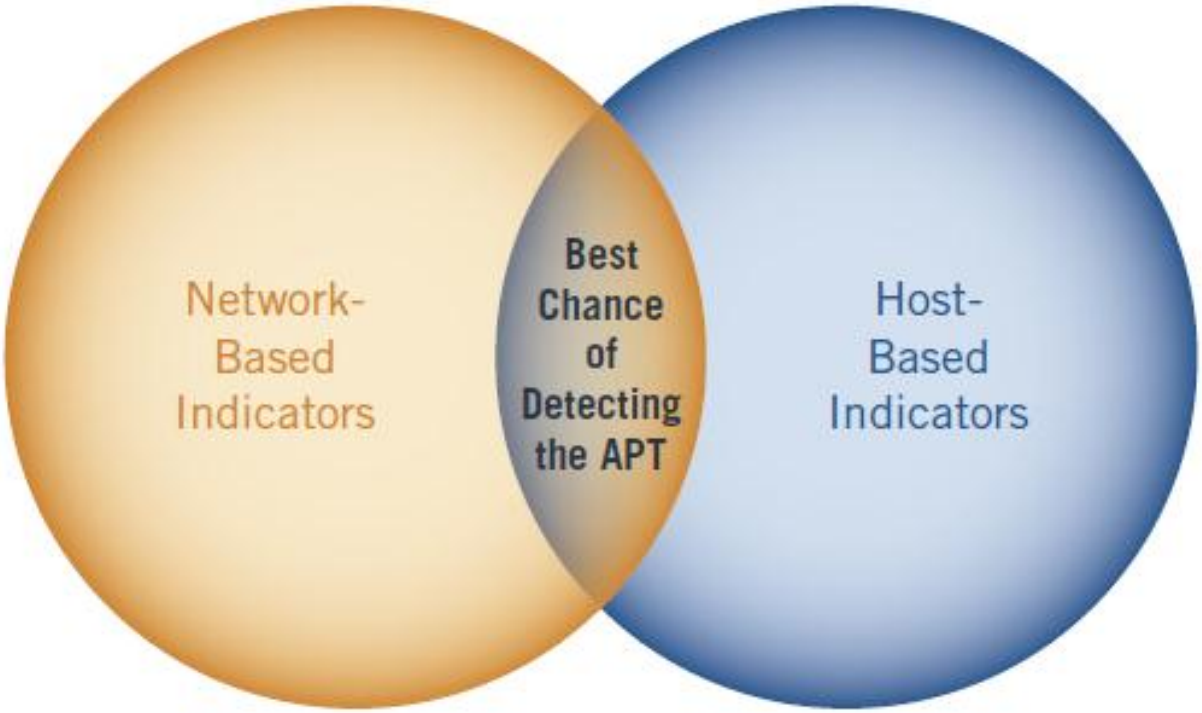

Service Desk


Industrial Control


Call Records

Need Both Network and Endpoint

And Inbound/Outbound!

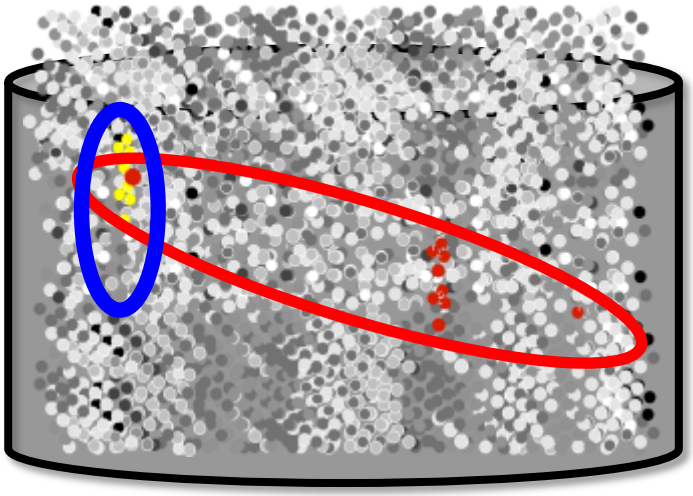


Step 2: Identify Threat Activity

- ✓ What's the modus operandi of the attacker? (think like a criminal)
- ✓ What/who are the most critical data assets and employees?
- ✓ What patterns/correlations of weak-signals in 'normal' IT activities would represent 'abnormal' activity?
- ✓ What in my environment is different/new/changed?
- ✓ What is rarely seen or standard deviations off the norm?

Big Data Solution

Big Data Architecture



Data Inclusion Model

- ✓ One product, UI, and datastore
- ✓ Scales horizontally to many TBs a day on commodity H/W
- ✓ All the original data from any source
- ✓ No database schema to limit investigations/detection
- ✓ Search & reporting flexibility
  - ▪ Advanced correlations
  - ▪ Math/statistics to baseline and find outliers/anomalies
- ✓ Real-time indexing and alerting
- ✓ "Known" and "Unknown" threat detection
- ✓ Open platform with API, SDKs, App framework

Big Data Solutions

splunk>

- NoSQL, distributed search, commodity H/W
- More than a SIEM:

Is Packet Capture Big Data?

make your own

- Fantastic technology for detecting anomalous traffic and for incident investigations
- Handles volume and velocity, but not variety

Sample Correlation of *Unknown* Threats

Sources

Example Correlation - Spearphishing

Email Server

2013-08-09T12:40:25.475Z,,exch-hub-den-01,,exch-mbx-cup-00,,STOREDRIVER,DELIVER,79426,<20130809050115.18154.11234@acme.com>,johndoe@acme.com,685191,1,,,hacker@neverseenbefore.com, Please open this attachment with payroll information,, ,2013-08-

Rarely seen email domain

User Name

Web Proxy

Rarely visited web site

.29 98483 148 TCP_HIT 200 200 0 622 - - OBSERVED GET www.neverbeenseenbefore.com HTTP/1.1 0 "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1 SV1; .NET CLR 2.0.50727; InfoPath.1; MS-RTC LM 8; .NET CLR 1.1.4322; .NET CLR 3.0.4506.2152;) User John Doe"

User Name

Endpoint Logs

User Name

08/09/2013 1 nt_status="(0)The operation completed successfully. "pid=1300 process_image="\John Doe Device\HarddiskVolume1\Windows\System32\neverseenbefore.exe" registry_type ="CreateKey"key_path="\REGISTRY\MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\ Printers Print\Providers\ John Doe-PC\Printers\{}\ NeverSeenbefore" data_type"

Rarely seen service

Time Range

All three occurring within a 24-hour period

Fingerprints of an Advanced Threat

What to Look For	Why	Source	Attack Phase
URL length on web app is standard deviations longer than normal	SQL injection. Hacker puts SQL commands in the URL.	Web	Infiltration
Rarely seen registry, service, DLL. Or they fail hash checks.	Malware or remote access toolkit	OS	Back Door
Account creation without corresponding IT help desk ticket	Hacker is creating new admin accounts	AD/ Help Desk logs	Recon
For single employee: Badges in at one location, then VPNs or logs in countries away	Stolen credentials	Badge/ VPN/ Auth	Data gathering
Employee makes standard deviations more data requests from a file server with confidential data than normal	Gathering confidential data for theft	OS	Data gathering

Fingerprints of an Advanced Threat

What to Look For	Why	Source	Attack Phase
Standard deviations more DNS requests from a single IP	Hackers exfiltrate the data in DNS packets; also fast-flux DNS	DNS	Exfiltration
Standard deviations larger DNS data flows from a single host	Hackers exfiltrate the data in DNS packet; standard deviations more DNS requests from a single IP	NetFlow / DNS	Exfiltration
Standard deviations larger traffic flows from a host to a given IP	Hacker exfiltrating info	NetFlow	Exfiltration
Long outbound URL w/o referrer	Botnets often embed long CnC message in the URL	Web	Exfiltration

Step 3: Remediate and Automate

- Where else in my environment do I see the “Indicators of Compromise” (IOC)?
- Remediate infected machines
- Fix weaknesses, including employee education
- Turn IOC into a real-time search for future threats

Security Realities…

- Big Data is only as good as the data it holds and the people behind the UI
- There is no replacement for capable practitioners
- Put math and statistics to work for you
- Encourage IT Security creativity and thinking outside the box
- Fine tuning needed; always will be false positives

Other Tactics From Forward - Thinking Practitioners

- Virtual sandbox for signature-less malware detection
- Network forensics/packet capture
- Web application firewalls
- Application whitelisting
- Honeypots
- Red team exercises
- Spearphishing employee training
- Air gapped network

Splunk For Security

- Big Data platform for ingesting machine data; 500MB to 100+ TB/day
- Many use cases within security
  - Forensics, incident investigation, known and unknown threat detection, fraud detection, and compliance
- Many use cases outside security: IT Operations, Application Management, web analytics
- Over 6000 customers total; 2500+ primary security use case customers
- Free download and tutorial at www.splunk.com

Questions?

Thank you!

Joe Goldberg

Splunk

jgoldberg@splunk.com

www.splunk.com

#RSAC

RSA CONFERENCE EUROPE 2013