Good Guys vs. the Bad Guys: Can Big Data Tools Counteract Advanced Threats?

Will Froning, Information Security Manager, American University of Sharjah

Mark Seward, Senior Director, Security and Compliance Marketing

Session ID: SPO-209

Session Classification: Intermediate

Agenda

- Understanding data indexing
- Security – New thinking wanted
- The Security Intelligence Platform
- Splunk at American University of Sharjah
- Visualizations wanted
- Questions

Security – What's at Stake?

> On average, organizations are experiencing a staggering 643 Web-based malicious events each week – incidents that effectively penetrate the traditional security infrastructure.

FireEye Advanced Threat Report – 1H 2012 Released August 29, 2012

Understanding Unknown Threats, or ‘Thinking like a Criminal’

Where is the most important and valuable data?

What are the typical security defenses?

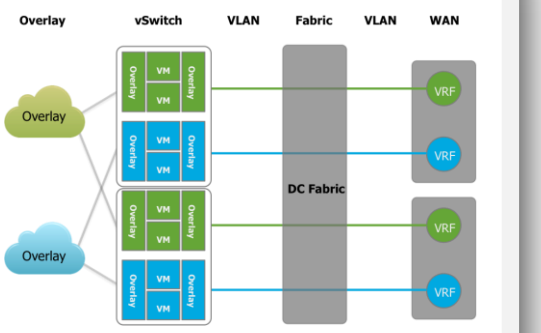
What structural information silos that exist for the security team?

What’s the typical patch cycle for applications and operating systems?

How does the IT team prioritize vulnerabilities?

Are ‘normal’ IT service user activities routinely monitored and correlated?

Gather as Much Data as Possible for Forensics



All data is security relevant

Velocity
Volume
Variety
Variability

Security is a Big Data problem

Practice the art of 'un-concealing'

Defining Security Intelligence

Enterprise Security Intelligence is:

- The collection of data from **all** IT systems in the enterprise that **could** be security relevant and;
- The application of the security team's knowledge and skill
- Resulting in risk reduction

Gartner

Prepare for the Emergence of Enterprise Security Intelligence,
Joseph Feiman, Gartner, June 29, 2011

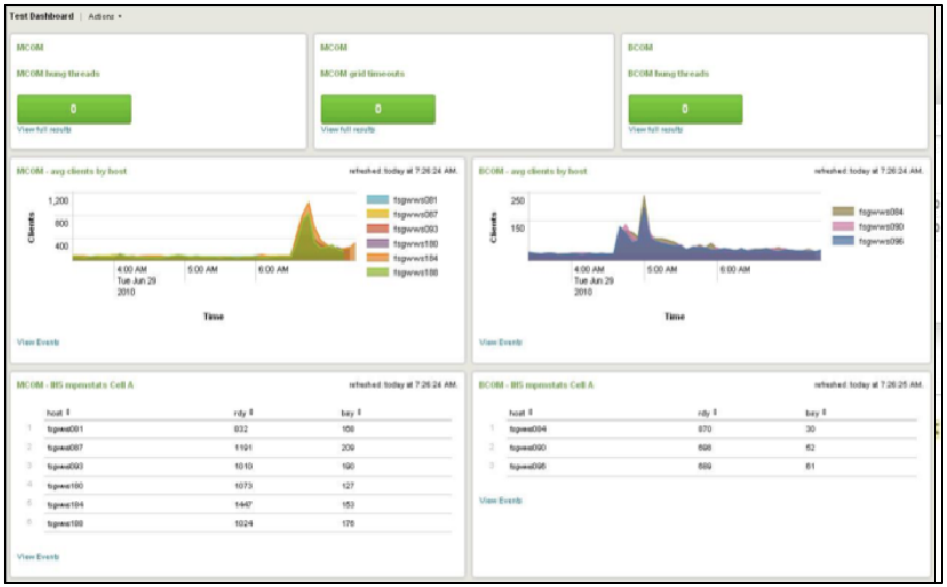
Big Data Indexing Solutions - how do it work?

Text Based Search

Google + 

Statistical Analysis

Time Index Ingestion
Text Base Search
Nested Search
Cross Data-type Search
Apend
Abstract
Cluster
Bucket
Multikv
Scrub
Join
Rare



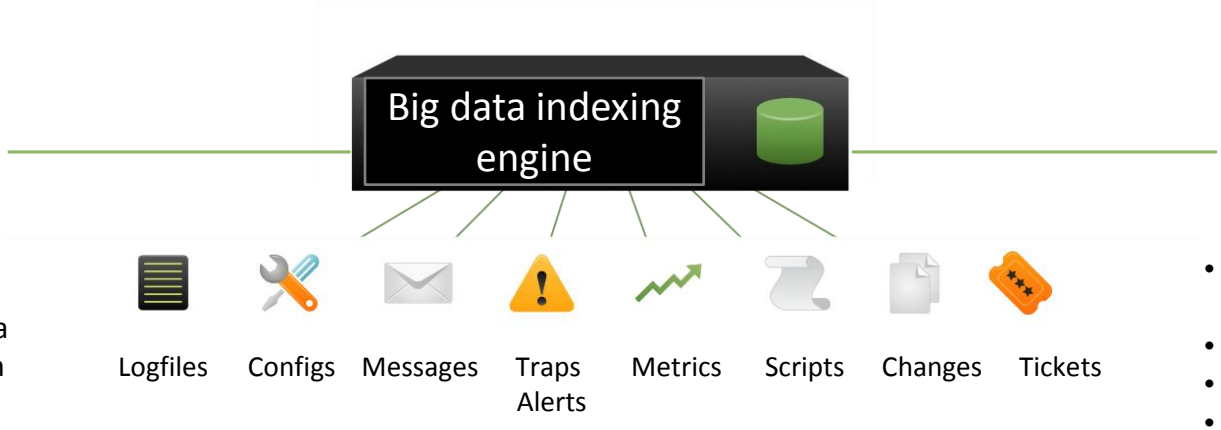
Data manipulation and visualization commands

Cluster
Associate
Stats
AVG
Transaction
Addtotals
Delta
Eval
Stddev
Rare
Outlier
Streamstats
Timechart

Big data indexing engines collects any Machine Data

Customer Facing Data

- Click-stream data
- Shopping cart data
- Online transaction data



Outside the Datacenter

- Manufacturing, logistics…
- CDRs & IPDRs
- Power consumption
- RFID data
- GPS data

Windows

- Registry
- Event logs
- File system
- sysinternals

Linux/Unix

- Configurations
- syslog
- File system
- ps, iostat, top

Virtualization & Cloud

- Hypervisor
- Guest OS, Apps
- Cloud

Applications

- Web logs
- Log4J, JMS, JMX
- .NET events
- Code and scripts

Databases

- Configurations
- Audit/query logs
- Tables
- Schemas

Networking

- Configurations
- syslog
- SNMP
- netflow

Extract Additional Value from Point Products

- Create the reports you need
- Trend data over long periods
- Use a big data system as the ‘glue’ between point products

Enabling IT Business Risk Scenarios

Applying IT Risk Scenarios
'Finding Abnormal Behaviors'

Security Relevant Data
Confidentiality / Integrity / Availability

App Mgmt

IT Ops

Compliance

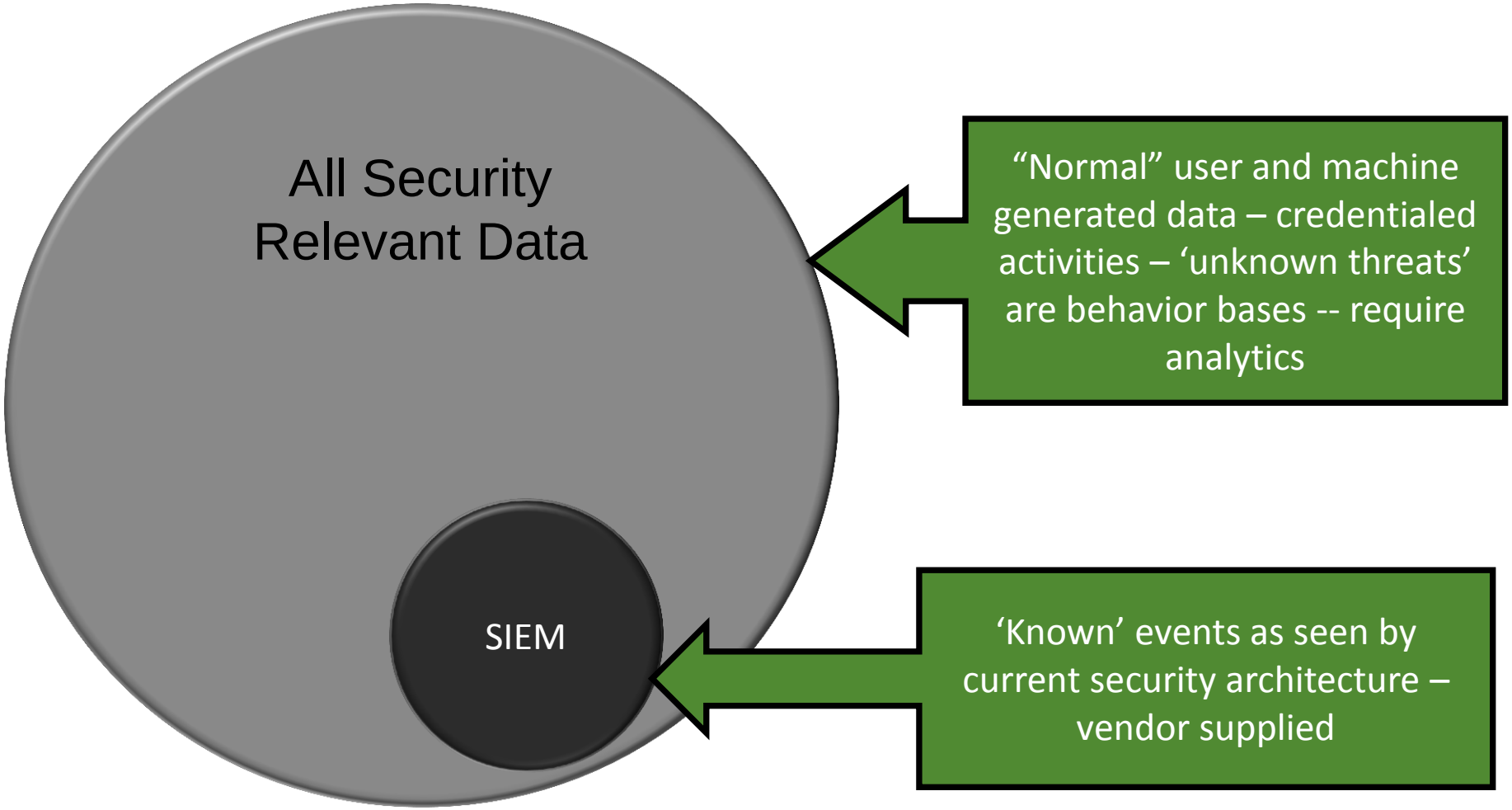
Web Analytics

Business Analytics

Big-Data indexing engine

CSO / CIO / CEO Views

The spheres of security data



Other Uses for Statistical Analysis

Action	Phase	Source	Splunk Search	Why
SQL Injection	Infiltration	WebLogs	len(_raw) +2.5stddev	Hacker puts SQL commands in the URL; URL length is standard deviations higher than normal
Password Brutes	Infiltration	Auth Logs	short delta _time	Automated password guessing tools enter credentials much faster than humanly possible
DNS Exfil	Exfiltration	DNS logs/FW Logs	count +2.5stddev	Hackers exfiltrate the data in DNS packet; standard deviations more DNS requests from a single IP
Web Crawling	Reconnaissance	Web/FTP Logs	count(src_ip) +2.5stddev	Web crawlers (copying the web site for comments, passwords, email addresses, etc) will be the source IP behind page requests standard deviations higher than normal
Port Knocking	Exfil/CnC	Firewall	Count outbound (deny) by ip	Threat does inside-out port scan to identify exfiltration paths

The New Security Intelligence Platform

Machine Data

Big Data indexing Engine

Security Intelligence for Business

Security Visualizations for Executives

Statistical Analysis

Proactive Monitoring

Search and Investigation

The Way Forward - Understanding Unknown Threats

The old way

Rigid

Signatures

Data reduction

Vendor dependence

Passive

The NEW way

Flexible

Statistical analysis

Data inclusion

Team creativity

Proactive

Splunk at American University of Sharjah

At bit about Me

- Will Froning, CISSP, GWAPT
- Came to the United Arab Emirates by way of St. Louis
- Been a security professional 7 years
- Big fan of automation and integration
- Investigations are the best part of my job

Keeping up with the kids

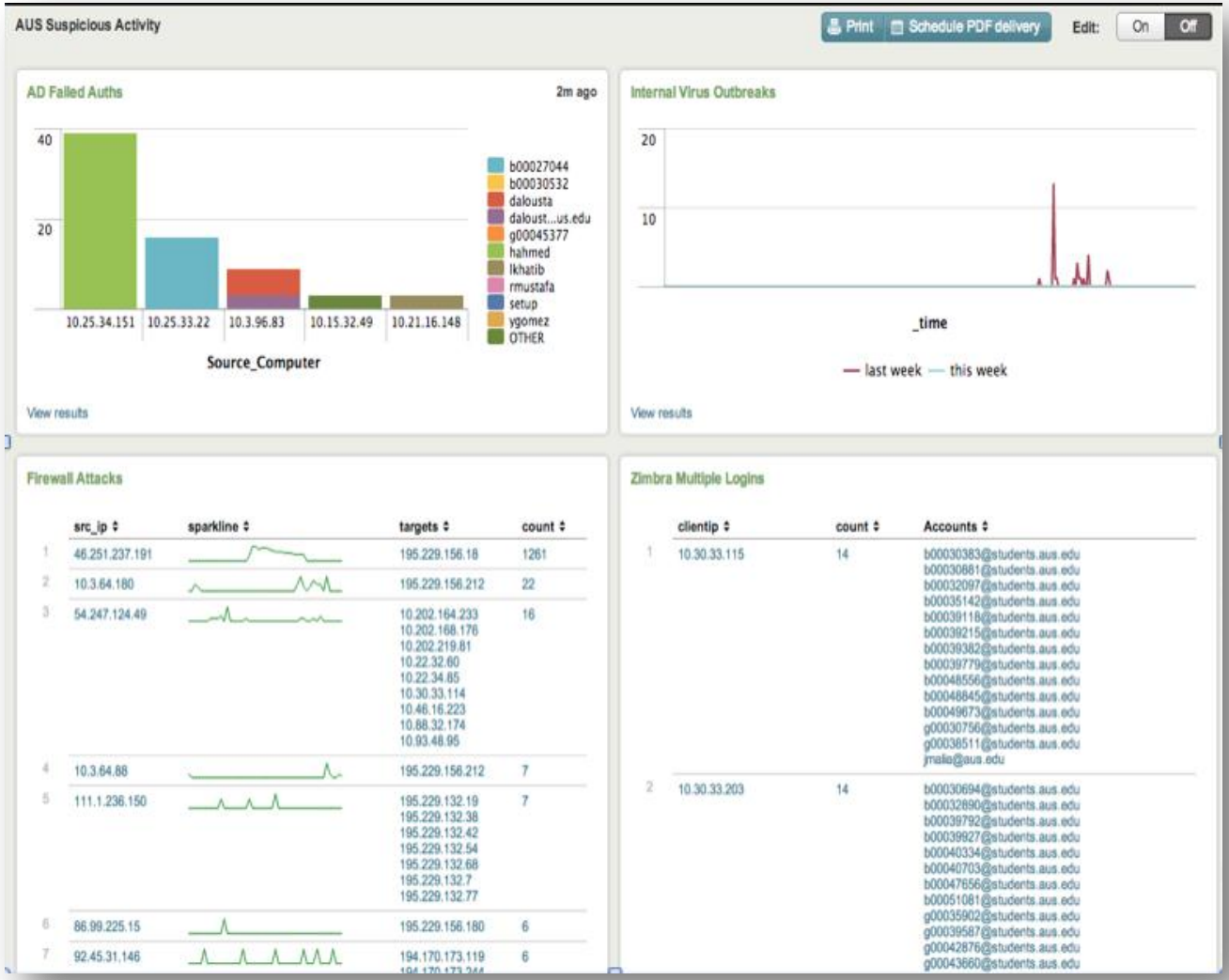
- Monitoring for academic integrity
- Data required:
  - Logon / Log-off data
  - Database logs
  - NAC data
  - Net Disco - Configuration information and connection data for network devices are retrieved via SNMP (used for asset discovery)
  - Firewall data
  - File change data
  - Zimbra (VMware Zimbra is an enterprise-class email, calendar and collaboration solution)

Discoveries - Stopping cheating scandals

- Stopping the use of stolen answer keys
- Stopping ‘Grades for Sale’ schemes
- Password sharing ring broken up
- Students come forward to admit to cheating
- Using a bit of theatrics can prove a point!

Don’t need to find all fraud – just enough to get the students to know that getting caught can be ‘unattractive.’

Security Monitoring



- AD Failed Authentication
- Internal Virus Outbreak
- Firewall Attacks
- Zimbra multiple Logins

IT Operations Use Cases

- Email
  - Monitoring self-service mail restores to prepare for support calls
- Phishing
  - Check URL logs for people that visited phishing links
- Bandwidth monitoring
  - Bandwidth very $$$$$$$ in the UAE
  - Understanding bandwidth usage saved $25,000 per month
  - Product pays for itself in bandwidth savings alone.
  - In essence academic integrity use is free.

Summary - American University of Sharjah

- Big Data and Analytics:
  - Provides security analysis
  - Provides academic integrity
    - Addresses student that attempt to game the system
  - Provides mail management
  - Saves money through analysis of bandwidth useage

Applying a New Strategy to Security

- Be less dependent on canned responses to security threats
- Look at security as a Big Data problem
- Think about security issues as IT risk scenarios
- Use statistical analysis to find abnormal patterns in normal IT activity data
- Start thinking like a criminal (or the cheating student)

Questions