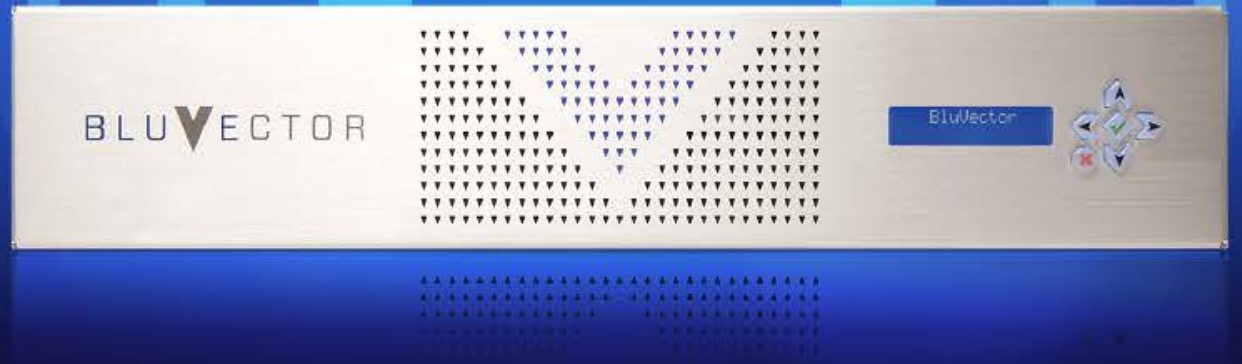
BLUVECTOR

Navigating the Pitfalls and Promises of Network Security Monitoring (NSM)

Dr. Scott Miserendino
Chief Data Scientist

Michael Gora
System Architect



Who are we?

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

Dr. Scott Miserendino
Chief Data Scientist

- Leads BluVector's data science and applied research teams
- Previously worked on large-scale network defense and sensor development for the DoD and IC

Michael Gora
System Architect

- Directs system and software architecture at BluVector
- Diverse background in software development spanning from large-scale application health and metrics to high speed network processing.

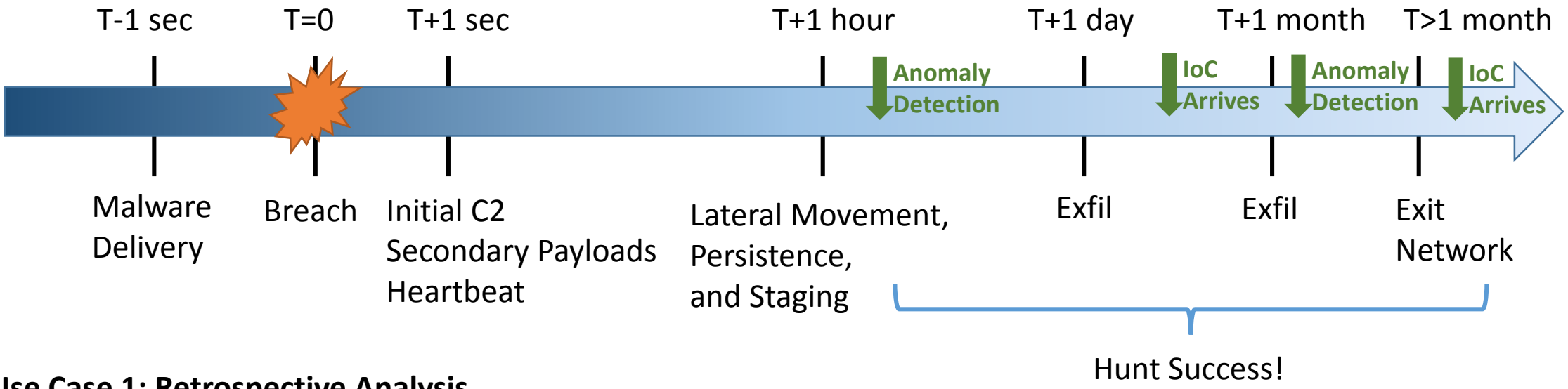
BLUVECTOR.
Cyber Threat Hunting
Machine Learning

Cyber Security Start-Up

- Started in 2013
- Born out of a large defense contractor
- HQ'ed outside of Ft. Meade, MD
- Network security appliance
- Bro-based protocol processing and network monitoring
- Sophisticated machine learning-based malware detection

NSM: Finding what we missed (better late than never)

BLUVECTOR
CYBER THREAT HUNTING PLATFORM



Use Case 1: Retrospective Analysis

- Indicators of Compromises (IoCs) are used to search the log repository
- IoCs typically arrive in feeds days to months after threat actors are actively using them

Use Case 2: Analytics/Anomaly Detection

- Monitor for statistically significant changes in asset, user or network behavior
- Operate over the entire store of logs or as a streaming analysis over the incoming logs
- Typically require multiple suspicious occurrences before alerting an analyst
- Require sophisticated analysts to understand how to interpret alerts or visually identify anomalies

NSM Pitfall: Scale and sophistication

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Network flow monitoring for cyber hunting requires significant capital and human resource investment
- Requires sophisticated analysts perhaps even ones with software dev experience (not the domain of your tier 1 or tier 2 SoC operator)
- Bandwidths are ever increasing (IoT, more web services, etc.)
- Number and variety of IoCs driving hunting workflows are increasing
- Budgets for analysts are the only thing not really growing so they are quickly becoming the bottleneck

Source: Oltsik, Jon, "Network Security Monitoring Trends", Enterprise Strategy Group, 2016. https://www.lancope.com/sites/default/files/esg-Network-Security-Monitoring-Trends.pdf

NSM Pitfall: Reliance on IoCs

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

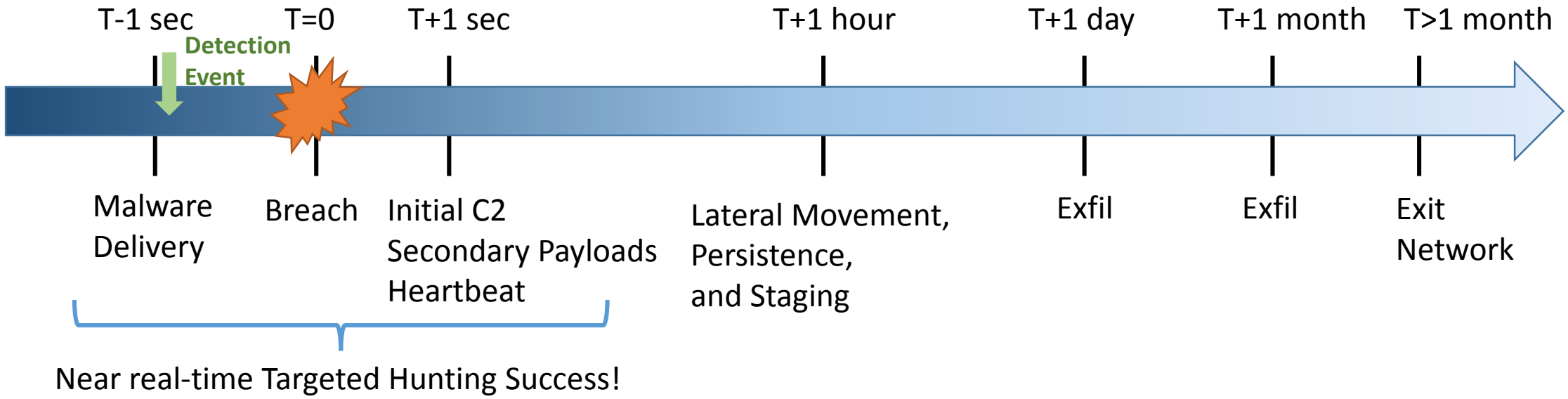
- Network-based Indicators of compromise
  - File names and hashes
  - URLs, hostnames and IP addresses
  - Email addresses and subjects
  - User agents
- Deficiencies in current IoC (a.k.a Threat Intel) feeds
  - Duplication
  - Poor curation
  - Lack of context over all IoCs
  - Limited estimation of IoC relevant time frame and shelf life
- Things that are going to make it worse
  - Polymorphic and one-time malware (hash IoCs)
  - FastFlux and DGA-based malware (domain IoCs)
  - IPv6 devices (IP IoCs)
  - IoT (explosion of potentially compromised endpoints, middle men and unwitting threat infrastructure)

"You know where it ends, yo, it usually depends on where you start"

-- Everlast, What It's Like

NSM Promise: Enabling better, faster detection through shortening the hunting cycle

BLUVECTOR
CYBER THREAT HUNTING PLATFORM



- Focus on the post-breach mission is fundamentally due to a distrust that detection is working (with good cause)
- What if detection techniques focused on not missing malware rather than not wasting analysts time with false positives?

NSM Promise: Enabling better, faster detection through shortening the hunting cycle

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

TBD Graphic showing mechanism for wider aperture detection

- What if detection techniques focused on not missing malware rather than not wasting analysts time with false positives?
- Network monitoring logs can then be used to highlight successful breaches within minutes not days or weeks
- This is how AV/host-based security is staying alive (moving from pure signature based detection by incorporating post install/execution behavioral analytics)

NSM Promise: It can move downmarket

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- High cost of large-scale log storage and query along with the required level of analyst sophistication to make sense of it prevent NSM from wide adoption downmarket
- Tool costs are actually not an issue
- Downmarket adopt requires vast simplification of the process:
  - Automate query (targeting)
  - Automate analysis (made easier when focusing on a limited-time frame context around a particular event of interest)
  - Be part of existing IT and security remediation workflows. Analysis must result in a decision not further exploration.
  - Do not require large expenditures on storage equipment or additional devops support to make it work
- The promise of downmarket adaption means focusing on enhancing near real-time detection while with going the benefits of retrospective analysis

Bro Network Security Monitor

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Passive, highly extensible open-source network analysis framework
- Stateful application-layer dynamic protocol processing
- Comprehensive and expressive log generation for connection and application layer activity
- So much more:
  - Content extraction, intelligence correlation, signature matching
  - Behavioral analysis, summary statistics, enforcement actions
- Swiss army knife:
  - Intrusion detection
  - Forensics
  - Network management
- Why Bro?
  - **PCAP** – Absolute truth of network activity that contains all content and metadata
    - Challenging storage and search requirements
  - **NetFlow** – Layer 3-4 flow focused metadata with manageable storage requirements
    - Minimal application-layer metadata
  - **Bro** – Rich application-layer metadata with storage requirements closer to NetFlow

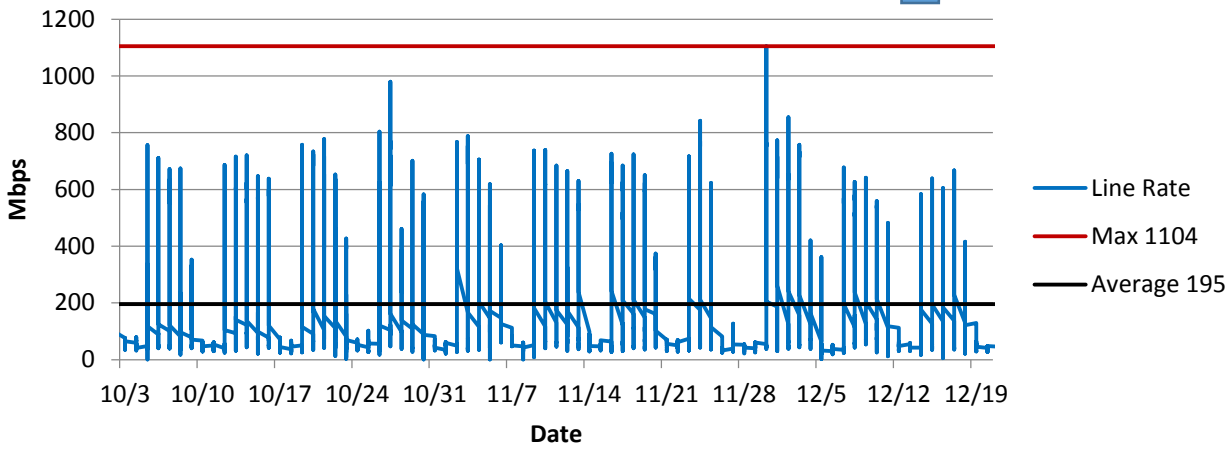
Logs, Logs, and More Logs

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

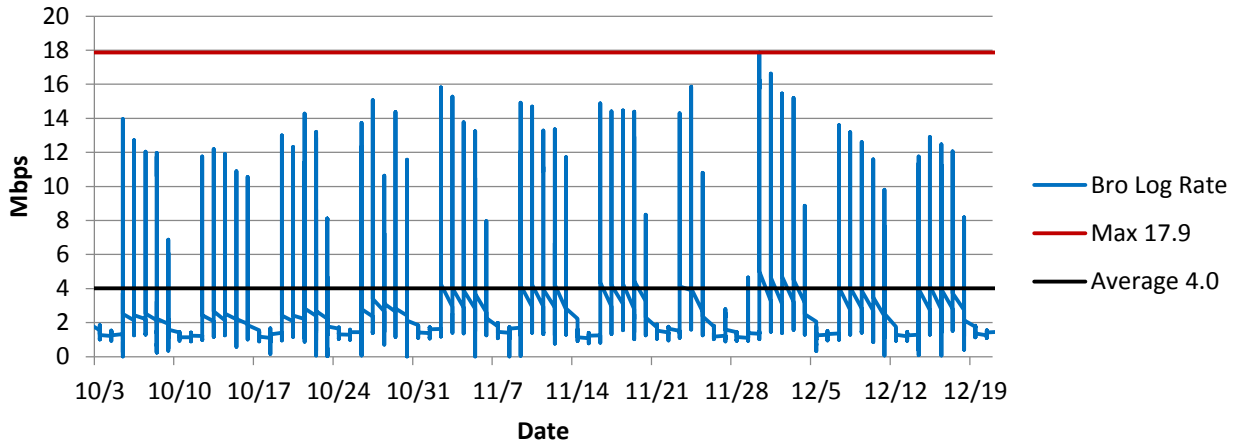
ts	uid	orig_h	orig_p	resp_h	resp_p	host	uri	referrer	status	Mime types
time	string	addr	port	addr	port	string	string	string	count	vector[string]
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/affiliate.php? ...	http://troysbilliards.ca/	200	text/html
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/	http://ads.hoa.lu/affilia ...	302	-
1456151204.529325	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1Q8MmBaKp7fhpi ...	-	200	application/zip
1456151204.529336	C89TiY360oLrmj2maa	172.16.223.135	50148	192.254.190.230	80	troysbilliards.ca	/	http://www.bing.com/searc ...	200	text/html
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1c1k96e6yu	http://ads.hoa.lu/affilia ...	200	text/html
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/61KjSQH5jGymnu ...	http://ugwpc.bimowamoky ...	200	application/x-shockwave-flash
1456151204.646378	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/8JCuizE1mccCPz ...	-	200	-

50x Data Reduction

Sample Network Line Rate

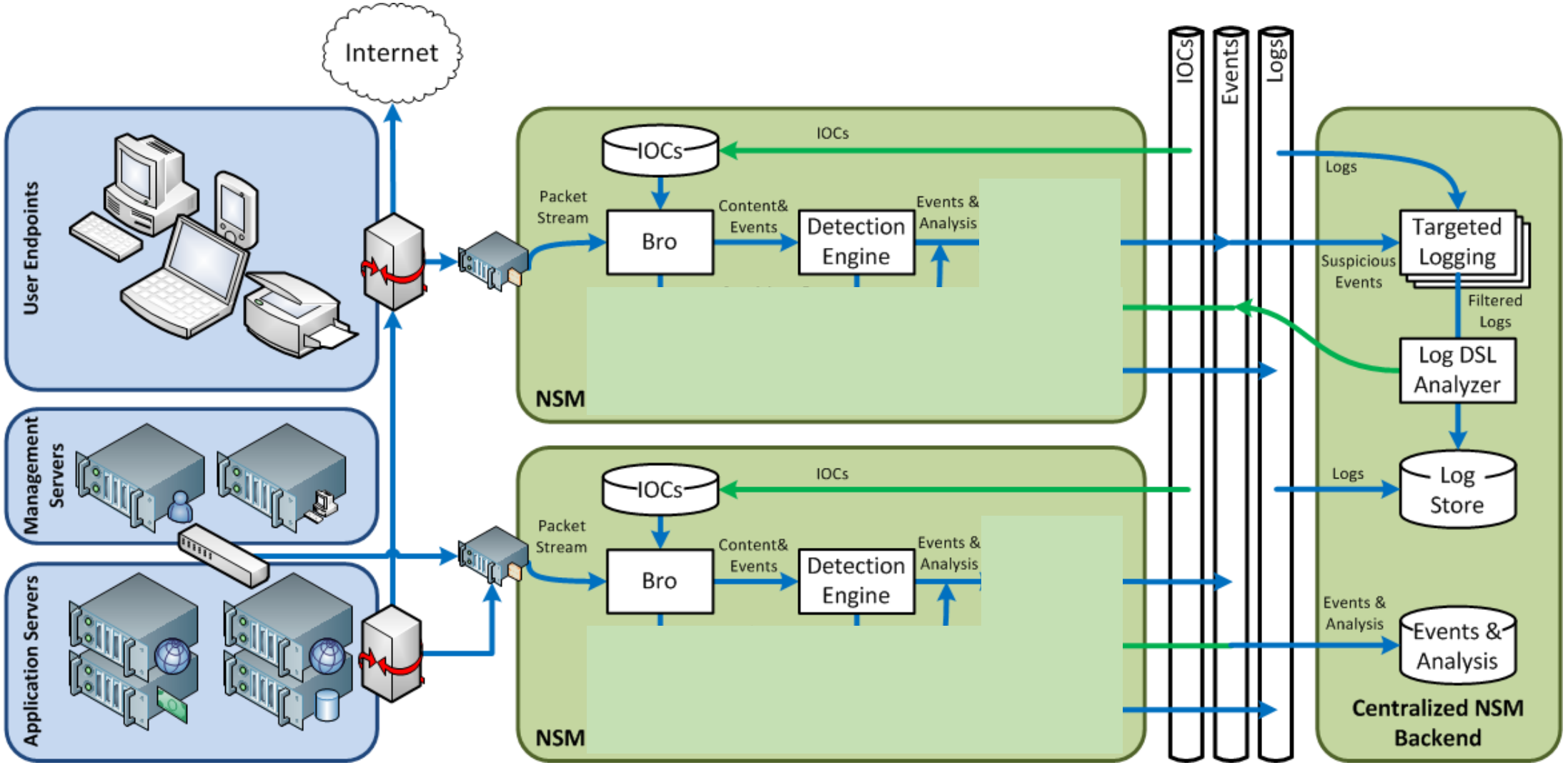


Sample Network Bro Log Data Rate



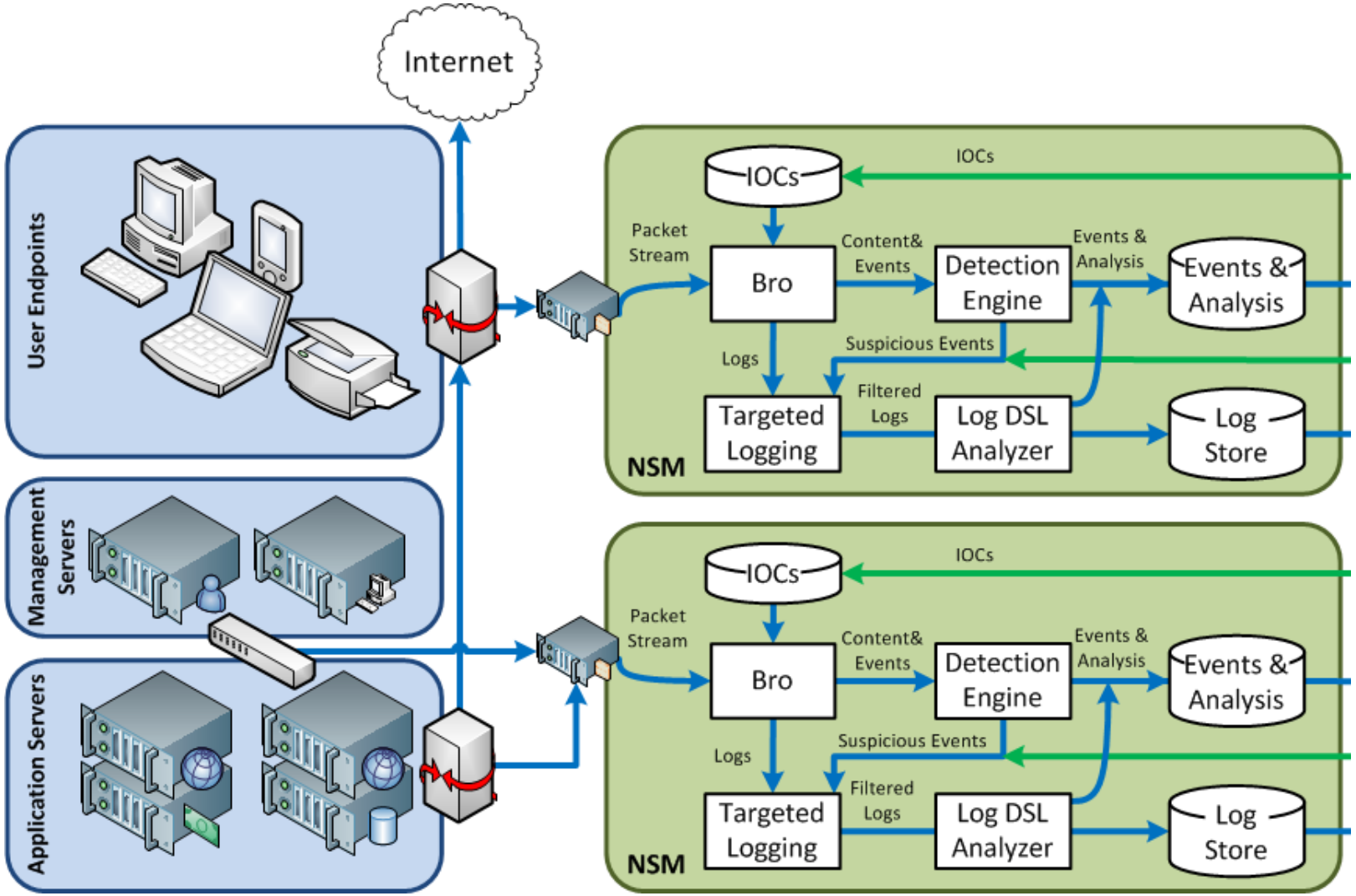
Targeted Logging: Focusing on what you need, when you need it

BLUVECTOR
CYBER THREAT HUNTING PLATFORM



Targeted Logging: Focusing on what you need, when you need it

BLUVECTOR
CYBER THREAT HUNTING PLATFORM



Stay on target …

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Focus on tracking internal endpoint versus external IOCs
  - Capture activity before and after a suspicious event
  - Rely on detection to offer first “breadcrumb”
    - IDS hits (potentially noisy)
    - Content analysis results
- Events require normalization to determine best target
  - Enumeration of internal subnets
  - Identification of noisy internal talkers
    - i.e. proxies, web servers
  - Protocol dependent identification of origin
- Pivot to include related log entries by identifiers
  - i.e. file, connection, certificate
- Example detection of x-shockwave-flash Trojan from web
  - Identify origin of potentially malicious content
  - Show pivot to other logs for additional content
- Example detection of FTP based download
  - Show swap of origin
  - Show pivot to other logs for additional content

ts	uid	orig_h	orig_p	resp_h	resp_p	host	uri	refer		es
time	string	addr	port	addr	port	string	string	string		ing]
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/affiliate.php? …	http://		
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/	http://a		
1456151204.529325	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1Q8MmBaKp7fhpi …	-	200	application/zip
1456151204.529336	C89TiY360oLrmj2maa	172.16.223.135	50148	192.254.190.230	80	troysbilliards.ca	/	http://www.bing.com/searc …	200	text/html
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1c1k96e6yu	http://ads.hoa.lu/affilia …	200	text/html
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/61KjSQH5jGymnu …	http://ugwpc.bimowamoky …	200	application/x-shockwave-flash
1456151204.646378	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/8JCuizE1mccCPz …	-	200	-

Walk through target example… http, ftp

Automating log analysis gets easier in a targeted world

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Initial focus on enriching suspicious events
  - Make adjudication easier and faster for analyst
  - Indicators obscured by noise became clear
  - How could we automate this process?
- Log Analysis Domain Specific Language (DSL)
  - Exploit temporal relationships
  - Correlate across multiple streams
  - Detect metadata abnormalities
  - Analysts can more simply write logic
  - Exportable and sharable
  - Extensible allowing system to adapt
- Similar in concept to Yara, Snort, Bro, Splunk

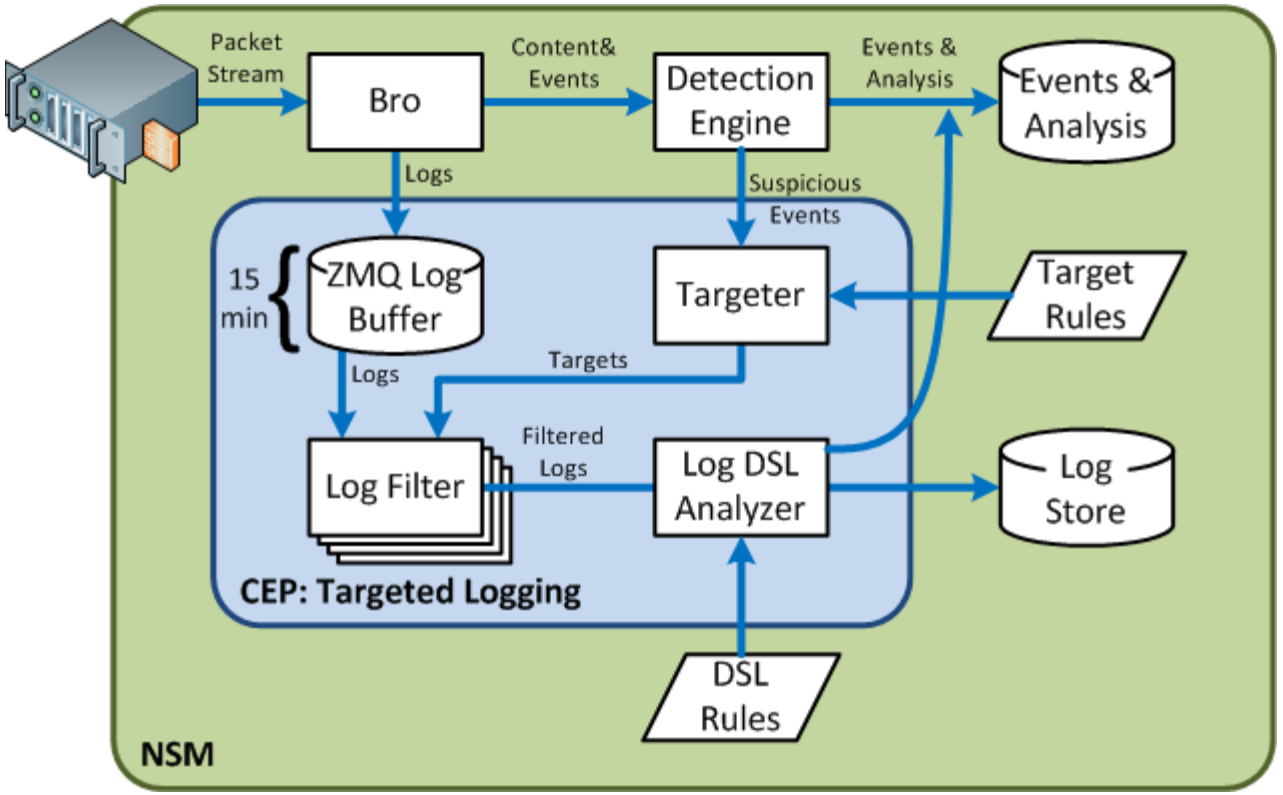
- Show example of Multiple Non-web files from source
  - Provide data set
  - Highlight target
  - Highlight matching rows
- Show example of potential exploited ad server
  - Provide data set
  - Highlight target
  - Highlight matching rows

ts	uid	orig_h	orig_p	resp_h	resp_p	host	uri	refer			es
time	string	addr	port	addr	port	string	string	string			ring]
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/affiliate.php? ...	http://			
1456151204.529325	CFjs5F4IR5vuokf2o6	172.16.223.135	50152	146.185.213.69	80	ads.hoa.lu	/	http://			
1456151204.529325	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1Q8MmBaKp7fhpi ...	-			ip
1456151204.529336	C89TiY360oLrmj2maa	172.16.223.135	50148	192.254.190.230	80	troysbilliards.ca	/	http://www.bing.com/searc ...	200	text/html	
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/1c1k96e6yu	http://ads.hoa.lu/affilia ...	200	text/html	
1456151204.529349	CgtigK3o3NNwlr9Ik4	172.16.223.135	50153	66.96.246.151	80	ugwpc.bimowamokykpps.net	/61KjSQH5jGymnu ...	http://ugwpc.bimowamoky ...	200	application/x-shockwave-flash	
1456151204.646378	CQFNUfOorqhoedXh3	172.16.223.135	50154	66.96.246.151	80	ugwpc.bimowamokykpps.net	/8JCuizE1mccCPz ...	-	200	-	

Walk through DSL example x 2

Targeted Logging: Prototype

BLUVECTOR
CYBER THREAT HUNTING PLATFORM



- Deployment on sample network
  - Peak data rate of 1 Gbps
  - Average data rate of 200 Mbps
- Python based Complex Event Processor
  - BSON over ZeroMQ messaging
- Protocol aware time based targeting
  - 15 minutes before 15 minutes after
  - Boolean OR target logic
- Single pass log filtering
  - 15 minute in memory ZeroMQ buffer
  - Provides sliding aperture for filtering
  - Filter work per Bro log type
  - Hash based exact matching, no pivot
  - ~10,000 logs/sec capacity per worker
- Domain Specific Language for log analysis
  - Based on Python Lex-Yacc
  - Simple SQLite query builder

Targeted Logging: A DSL for Automated Analysis

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

Example DSL Functions

Function	Description
filesDownloaded	Filter logs for filetypes in a list of strings applied to appropriate log field
sameHost	Filter logs for those with the same host as the triggering event
after	Filter logs for those occurring X seconds after the triggering event
before	Filter logs for those occurring X seconds before the triggering event
around	Filter logs for those occuring between X seconds before and Y seconds after
nonStandardPort	Filter logs for entries not matching the standard port for the given log type
hostsVisitedContain	Filter logs for entries whose host entries match the provided list of regexs

Boolean Logic Support

- Comparison: gt (>), gte (>=), lt (<), lte (<=), eq (==)
- Operations: Or (|), And (&), Xor (^), Not (!)

Example DSL Rules

```
- id: 001
  name: "Protocols over non-standard ports"
  desc: "Detects instances of protocol use on non-standard ports (i.e., HTTP not on 80 or 443)"
  heuristic: 'has(nonStandardPort(CONN))|has(nonStandardPort(HTTP))'
  severity: "warn"
- id: 002
  name: "Multiple Non-web files from same source"
  desc: "Detects the download of more than one non-web file from the same source as the event"
  heuristic: 'gt(filesDownloaded(sameHost(HTTP,"host"),["%dosexec%","%shockwave%","%pdf%"]),1)'
  severity: “warn"
- id: 003
  name: "Multiple rapid executables"
  desc: "More than one Windows executable file downloaded within 2 seconds of the event"
  heuristic: 'gt(around(filesDownloaded(FILES,"%dosexec%"),2,2),1)'
  severity: "alert"
- id: 004
  name: "Potential exploited ad server"
  desc: "Possible ad server visited within a short time prior to the event"
  heuristic: 'has(before(hostsVisitedContain(HTTP,"%.?ads?\.%"),0.5))'
  severity: "warn"
- id: 005
  name: "Multiple executables"
  desc: "More than one Windows executable file downloaded"
  heuristic: 'gt(filesDownloaded(FILES,"%dosexec%"),1)'
  severity: "warn"
```

DEMO

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- 5 minute demonstration
  - Evaluating sample pcap through bro
  - Show targeting
  - Show DSL log analysis

Experimental set up and results

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Stats on network used
  - Primarily looking at web events
  - Scope number of events, suspicious results
- Show what the targeting looks like in data reduction
- Comparison of real network data set vs. known malicious
  - Report false positive/false negative

Summary & Future Work

BLUVECTOR
CYBER THREAT HUNTING PLATFORM

- Rethink the paradigm between detection and network security monitoring
- Network security monitor can do more than point out the failures in detection
-  It can make detection great again! 

- Future Work:
  - Expand DSL to cover additional common analyst behaviors
  - Expand network behavior heuristic rule set for automated analysis to cover more threat activities
  - Dynamic targeting, grow targeting based on observed traffic
  - Protocol expansion (smtp is hard)