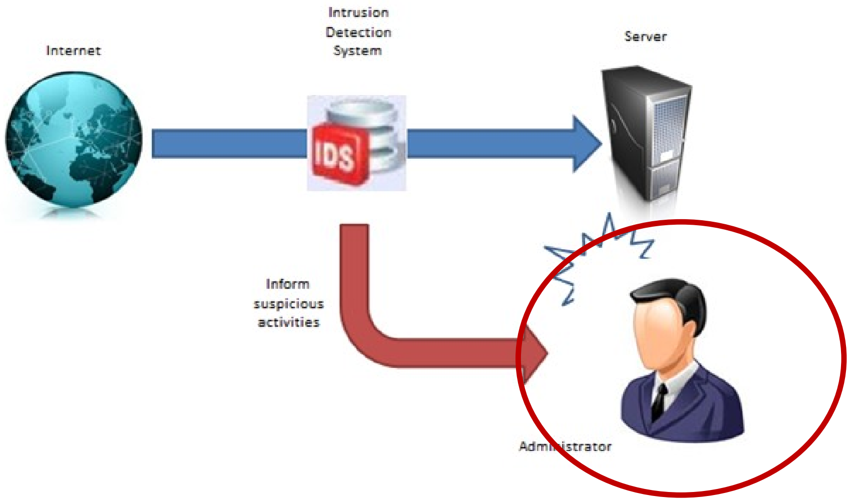
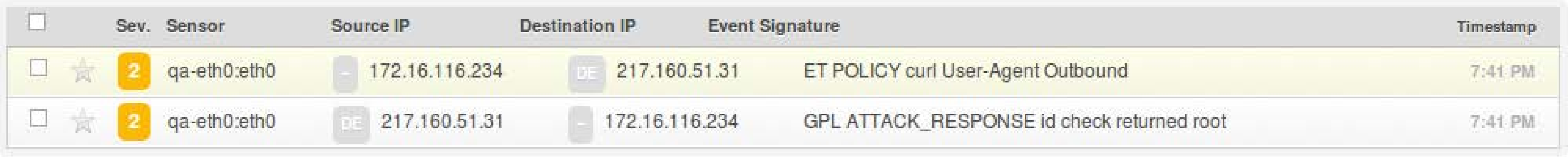
Security Event Management: Challenges and Opportunities

Pratyusa K. Manadhata
HP Labs
manadhata@hp.com

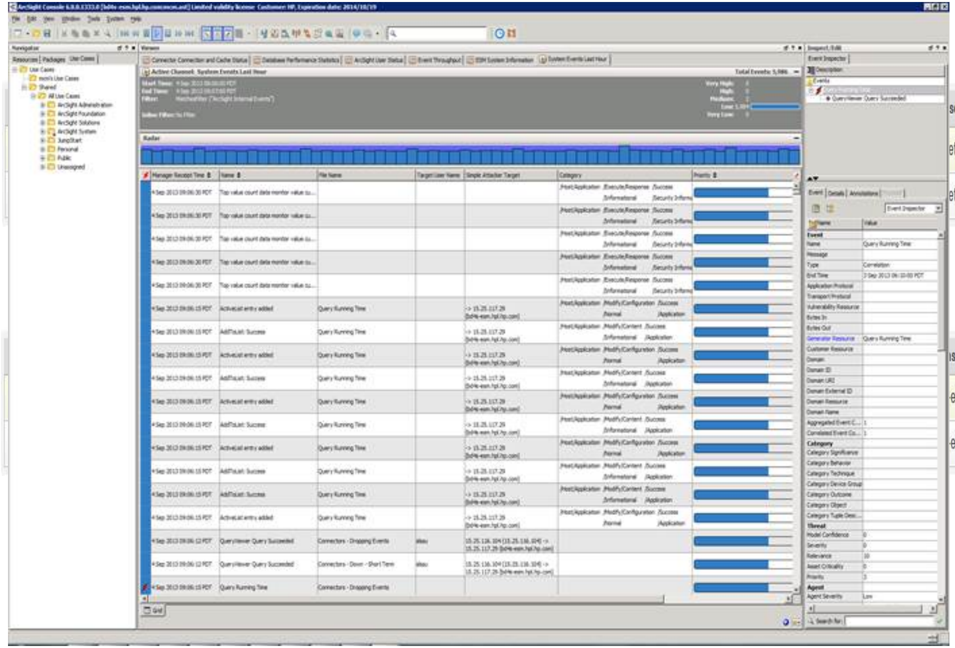
Enterprise Security: Point products





□		Sev.	Sensor	Source IP	Destination IP	Event Signature	Timestamp
□	☆	2	qa-eth0:eth0	- 172.16.116.234	DE 217.160.51.31	ET POLICY curl User-Agent Outbound	7:41 PM
□	☆	2	qa-eth0:eth0	DE 217.160.51.31	- 172.16.116.234	GPL ATTACK_RESPONSE id check returned root	7:41 PM

Security information and event management systems (SIEM)


ArcSight Console 6.8.0.1333.0 Limited validity license Customer: HP, Expiration date: 2014/10/19
File Edit View Window Tools System Help
Navigator
Resources Packages Use Cases
Use Cases
Shared
All Use Cases
ArcSight Administration
ArcSight Foundation
ArcSight Solutions
ArcSight System
Viewer
Connector Connection and Cache Status
Database Performance Statistics
ArcSight User Status
Event Throughput
ESM System Information
Active Channel: System Events Last Hour
Radar
Inspect/Edit
Event Inspector

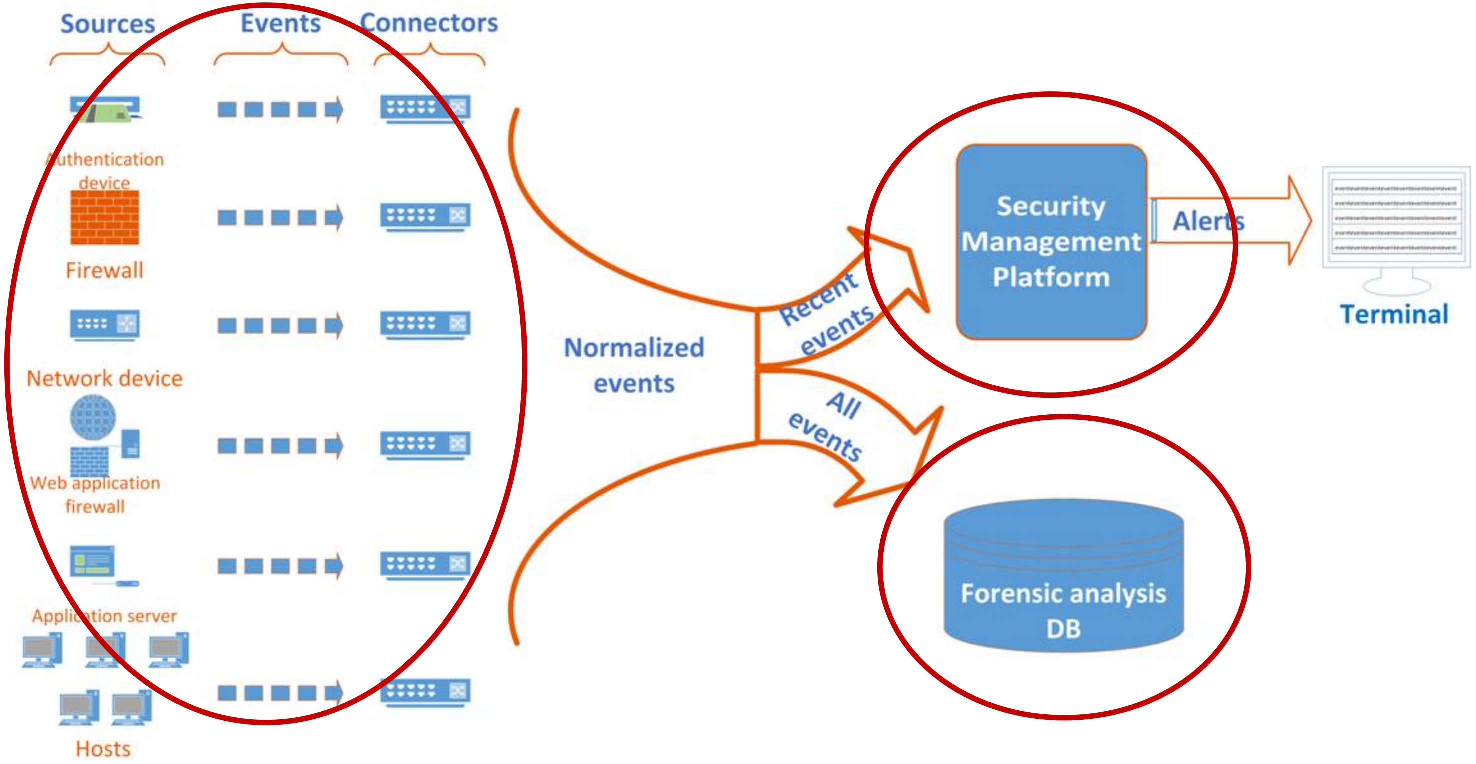
sor	Source IP	Destination IP	Event Signature	Timestamp
eth0:eth0	172.16.116.23	217.160.51.3	ET POLI... curl User-Agent Outbound	7:41 PM
eth0:eth0	217.160.51.3	172.16.116.234	GPL ATTACK_RESPONSE id check returned root	7:41 PM

Unified UI

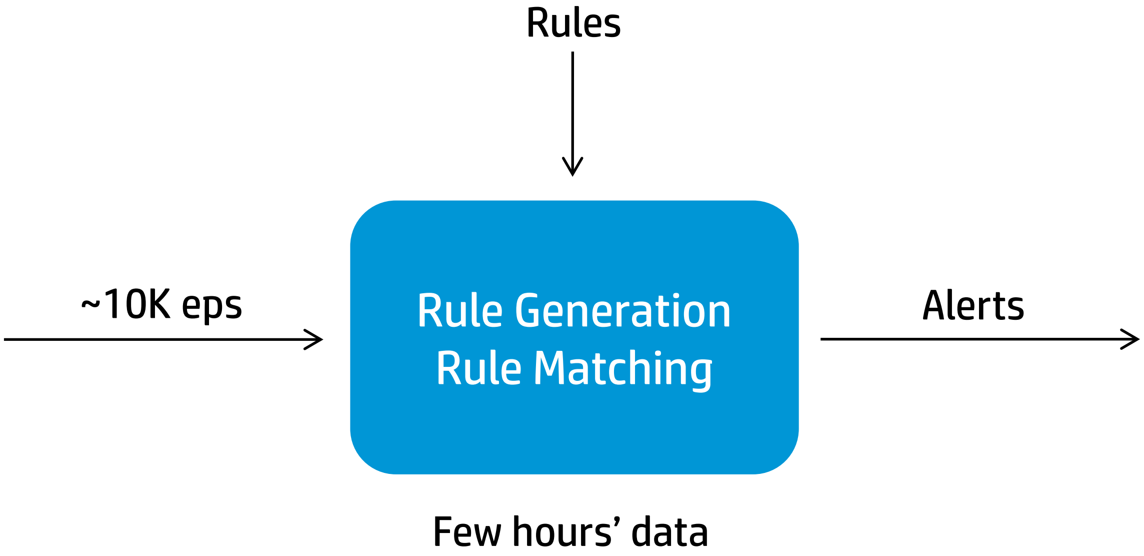
False positive reduction

sor	Source IP	Destination IP	Event Signature	Timestamp
eth0:eth0	172.16.116.234	217.160.51.31	ET POLICY curl User-Agent Outbound	7:41 PM
eth0:eth0	217.160.51.31	172.16.116.23	GPL ATTACK_RESPONSE id check returned root	7:41 PM

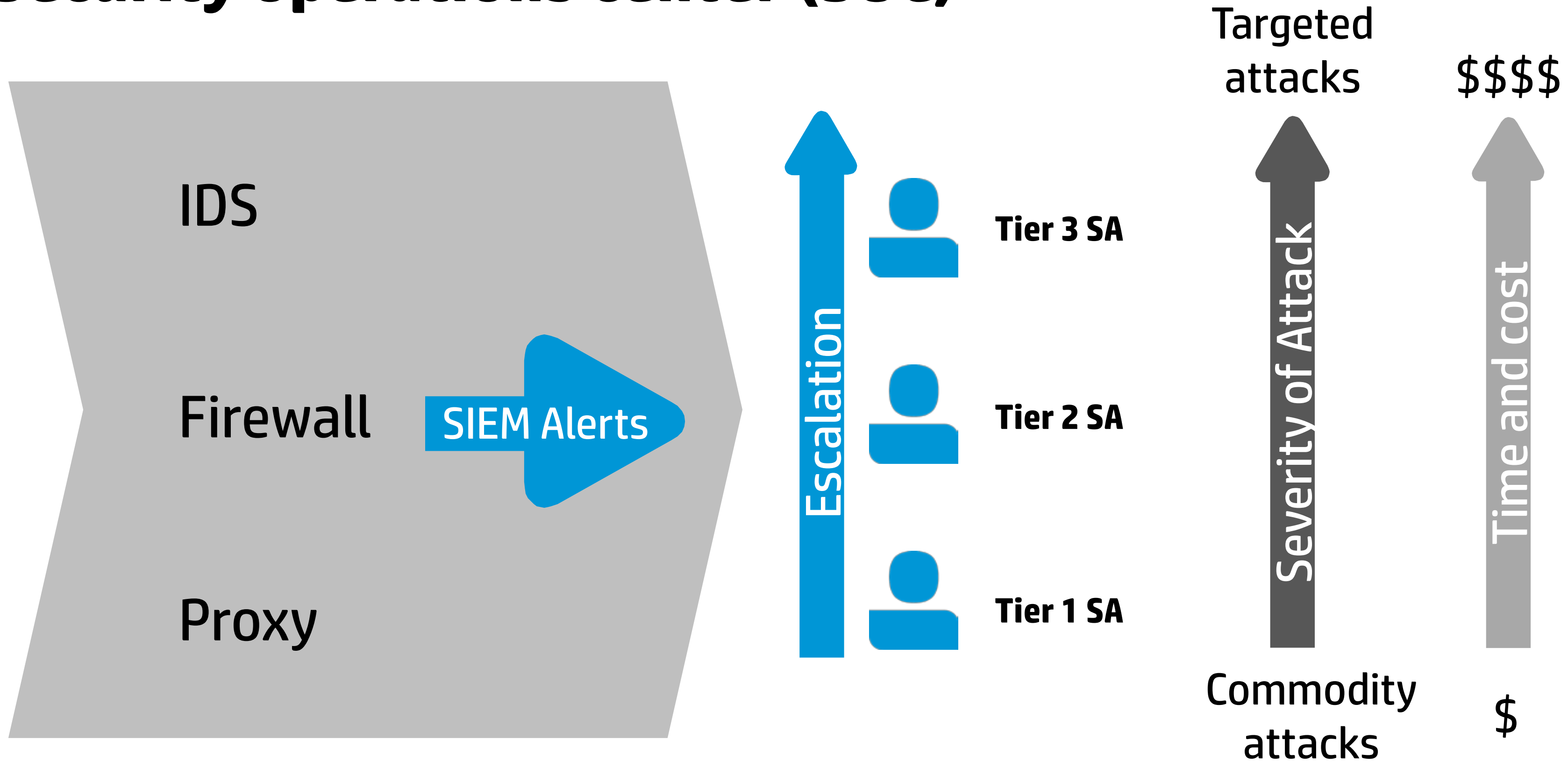
SIEM architecture



Management platform



Security operations center (SOC)

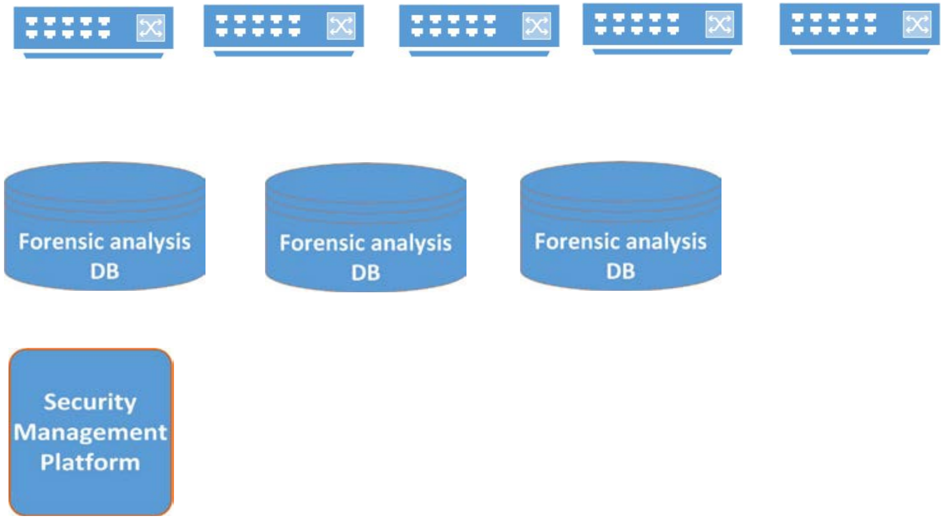


An HP SOC

Hundreds of connector servers

A few hundred forensic DBs

Multiple management platforms



The reality

3 000 000 000 events/day

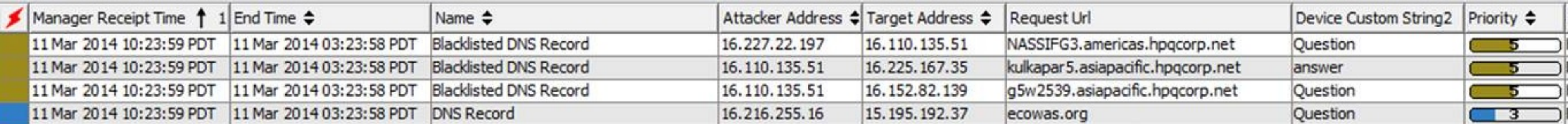
20 analysts

Operational challenges

Implementing rules – balancing FPs and FNs

Lack of context

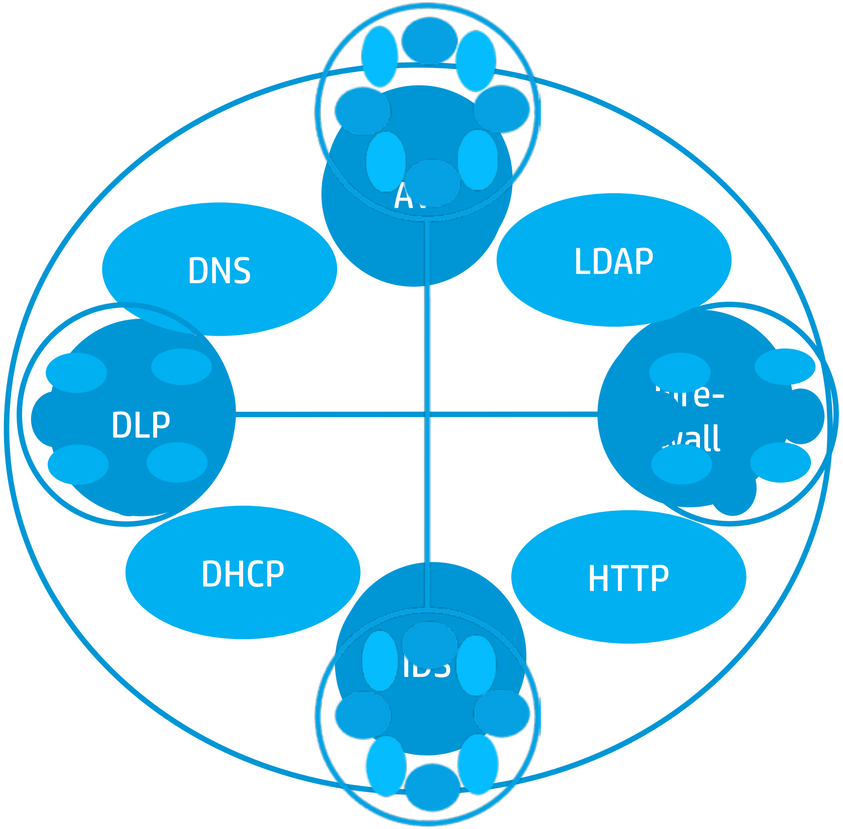
Challenge: Drinking from a firehose



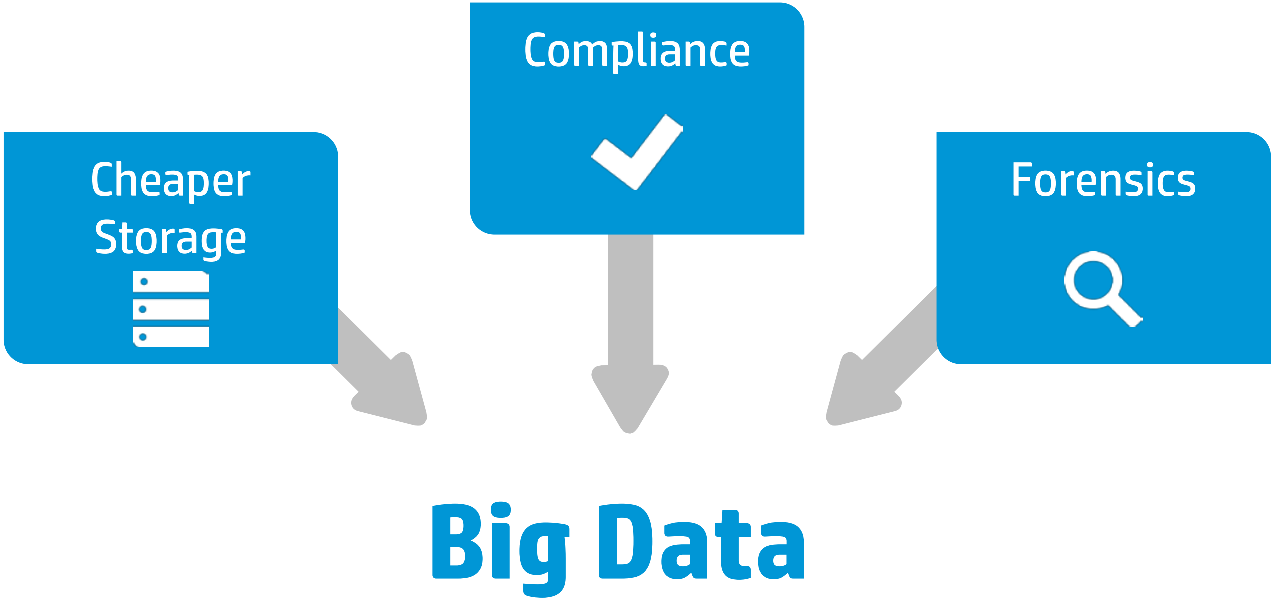
	Manager Receipt Time ↑ 1	End Time	Name	Attacker Address	Target Address	Request Url	Device Custom String2	Priority
🟨	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.227.22.197	16.110.135.51	NASSIFG3.americas.hpqcorp.net	Question	5
🟨	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.110.135.51	16.225.167.35	kulkapar5.asiapacific.hpqcorp.net	answer	5
🟨	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	Blacklisted DNS Record	16.110.135.51	16.152.82.139	g5w2539.asiapacific.hpqcorp.net	Question	5
🟦	11 Mar 2014 10:23:59 PDT	11 Mar 2014 03:23:58 PDT	DNS Record	16.216.255.16	15.195.192.37	ecowas.org	Question	3

Minutes to decide if an alert or event needs further attention

Challenge: Getting value out of data



Why enterprises collect security data



Research opportunity

Algorithms and systems to identify actionable security information from event data

Research opportunity

Improve SOC Workflow

Better Security Products

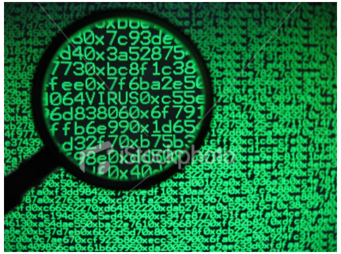
Better Enterprise Security

Data-driven security products

AV/IDS/Firewall/..

Signatures/Rules/..

Anti-malware evolution



Static Analysis

Dynamic Analysis

Reputation Analysis

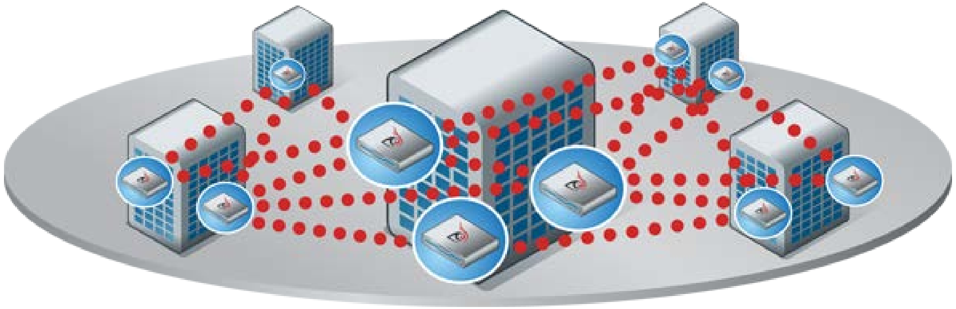
Data analysis: feature extraction/learning/classification/..

On-Premises analysis

On-site analysis

Fine-grained behavior data

Hardware and virtualization progress enable scaling

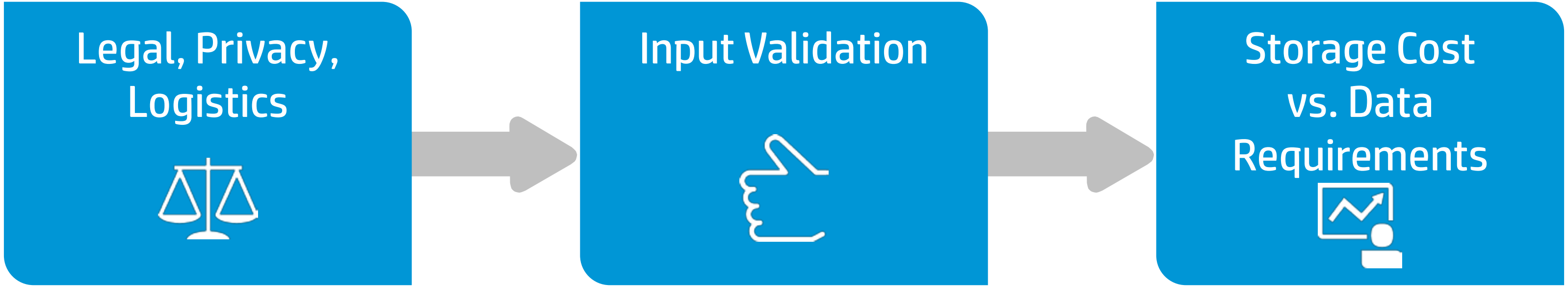


Research opportunity

What can we infer from event logs?

How do they compare to on-premises analysis?

Data collection and storage



Legal, Privacy, Logistics

Input Validation

Storage Cost vs. Data Requirements

Scalable analysis

Work with human analysts, not replace them

Things that we took for granted are not true any more

Infer human intent from machine logs

No definition of bad exists – rely on heuristics

Automation is hard

History is not a good indicator of future

Algorithms must learn and evolve

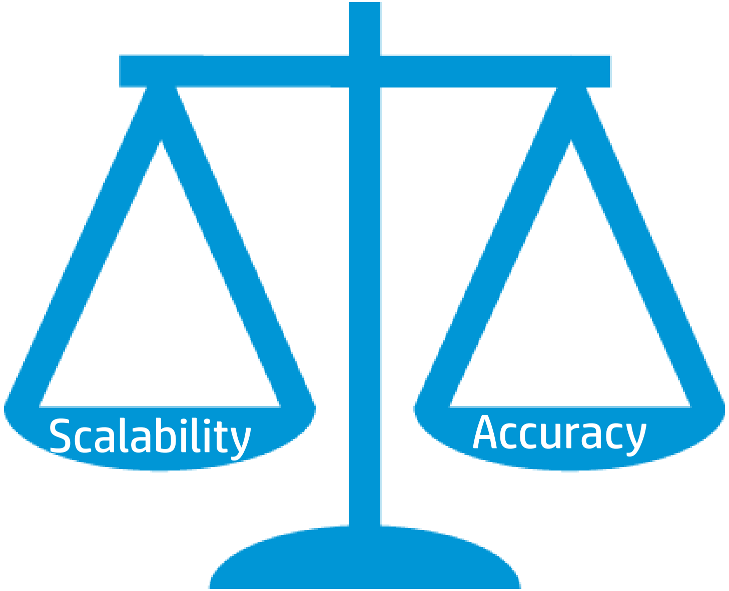
Adversaries adapt

Networks and systems change and fail

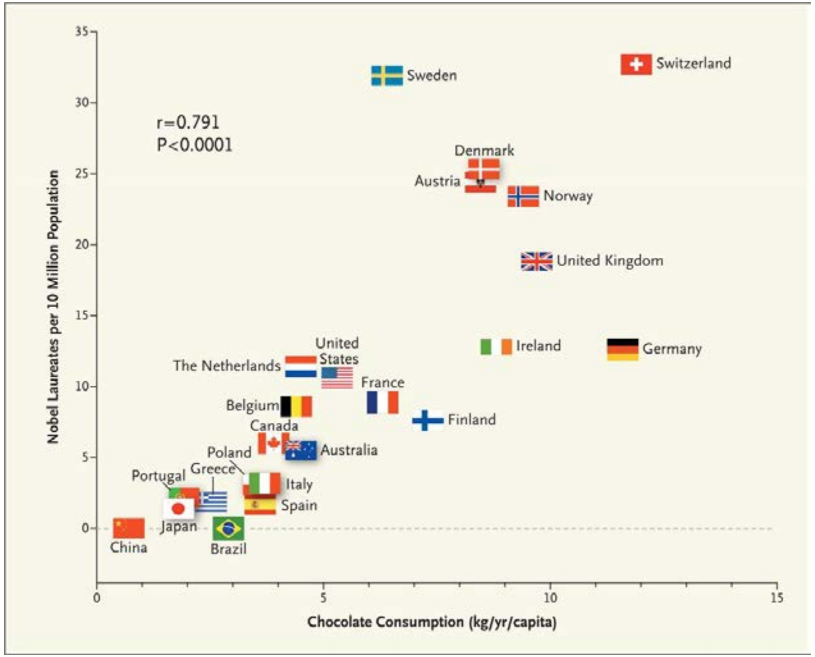
People behave unpredictably

Beware of false positives

Benign events outnumber malicious events



More data = More spurious correlations



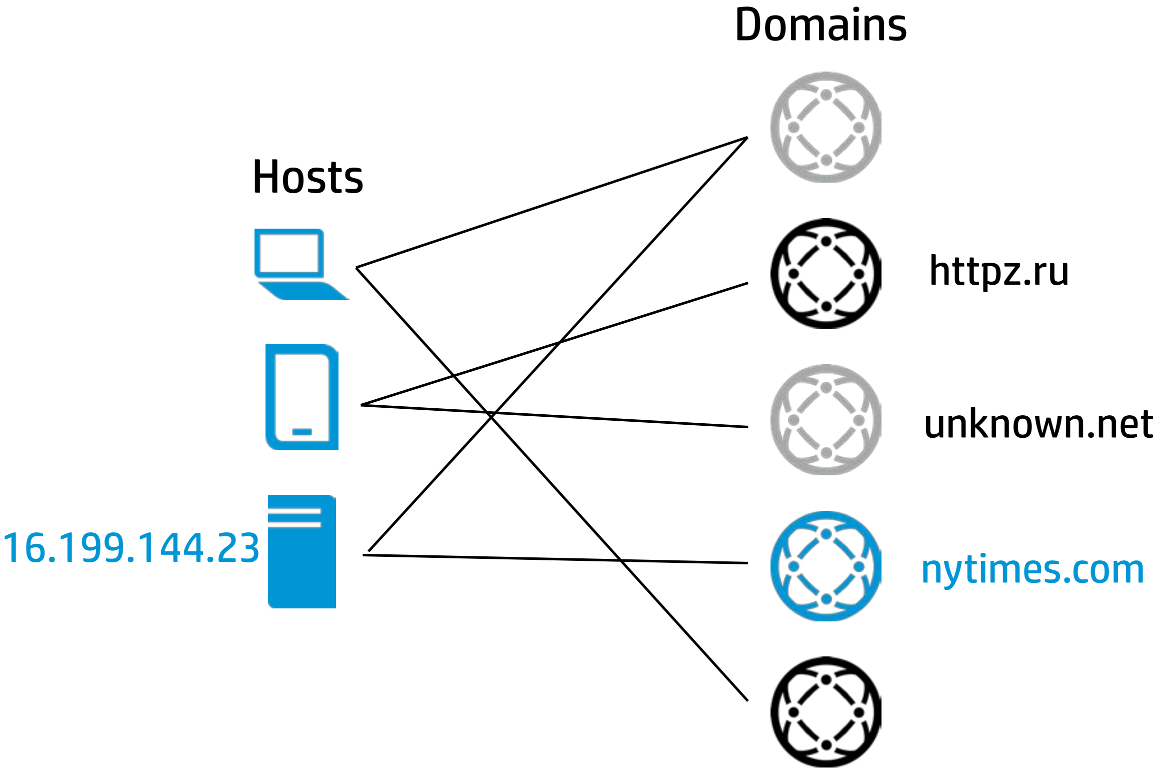
Chocolate Consumption, Cognitive Function, and Nobel Laureates, Franz H. Messerli, New England Journal of Medicine, Oct 2012

Privacy

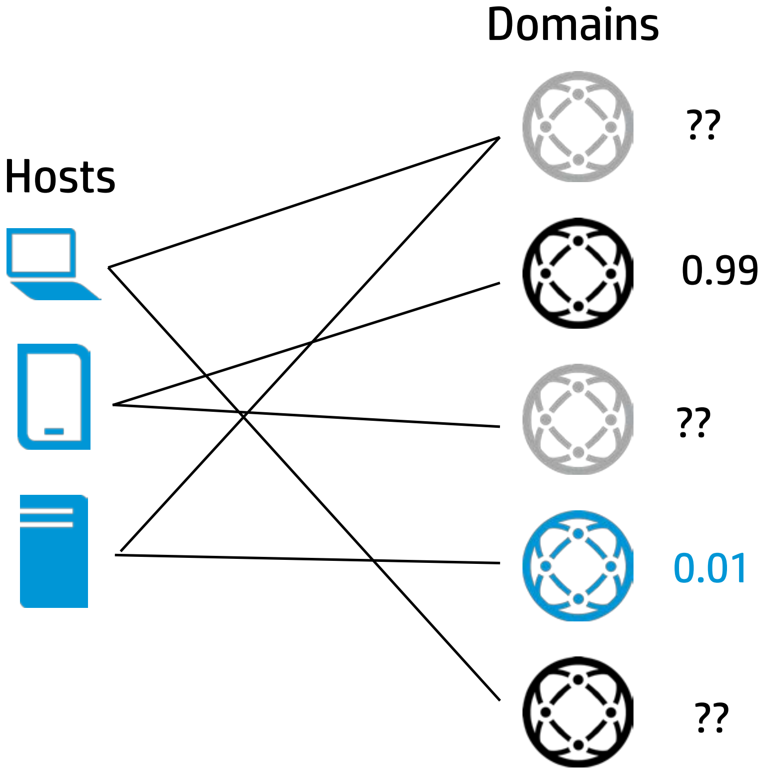
Data minimization vs Serendipity

Privacy-utility trade off

Malicious domain detection



Estimating marginal probability of being malicious



$$P(x_1, x_2, .., x_n)$$

$$P(x_i) = \sum_{x1} .. \sum_{xi-1} \sum_{xi+1} .. \sum_{xn} P(x_1, x_2, .., x_n)$$

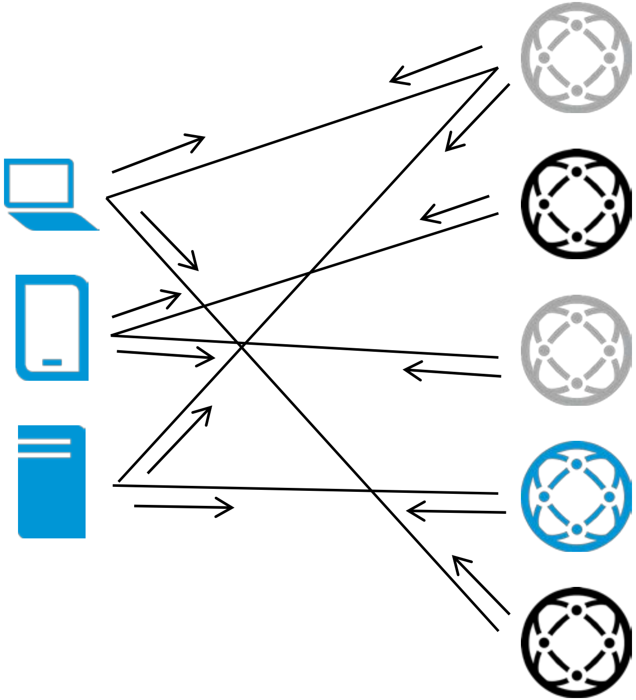
Belief propagation algorithm [P82, YFW01]

Marginal probability estimation in graphs

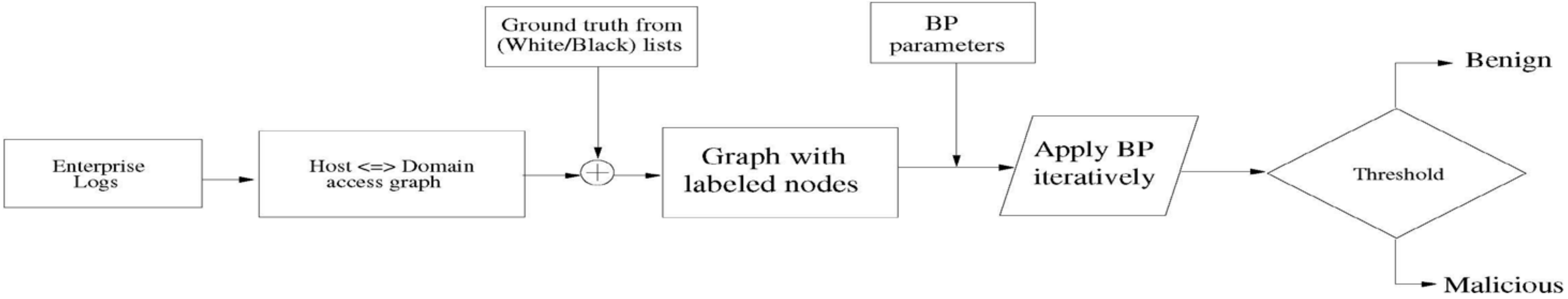
- NP-complete

Belief propagation is fast and approximate

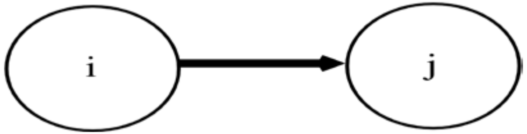
- Iterative message passing



Our approach



Message passing



Message(i → j) ∝ (prior, edge potential, incoming messages)

$$m_{ij}(x_j) = \sum_{x_i \in S} \phi(x_i)\psi(x_i, x_j) \prod_{k \in N(i) \backslash j} m_{ki}(x_i)$$

Prior Edge potential Incoming messages

Belief computation

Belief(i) $\propto$ (prior, incoming messages)

$$b_i(x_i) = K\phi(x_i) \prod_{j \in N(i)} m_{ji}(x_i)$$

Normalization constant

Prior

Incoming messages

HTTP Proxy logs

Logs from a large enterprise

- 98 HTTP proxy servers, 7 months of data
- 1 day's logs : 1.29 billion events
- 2.80M nodes and 27.8M edges

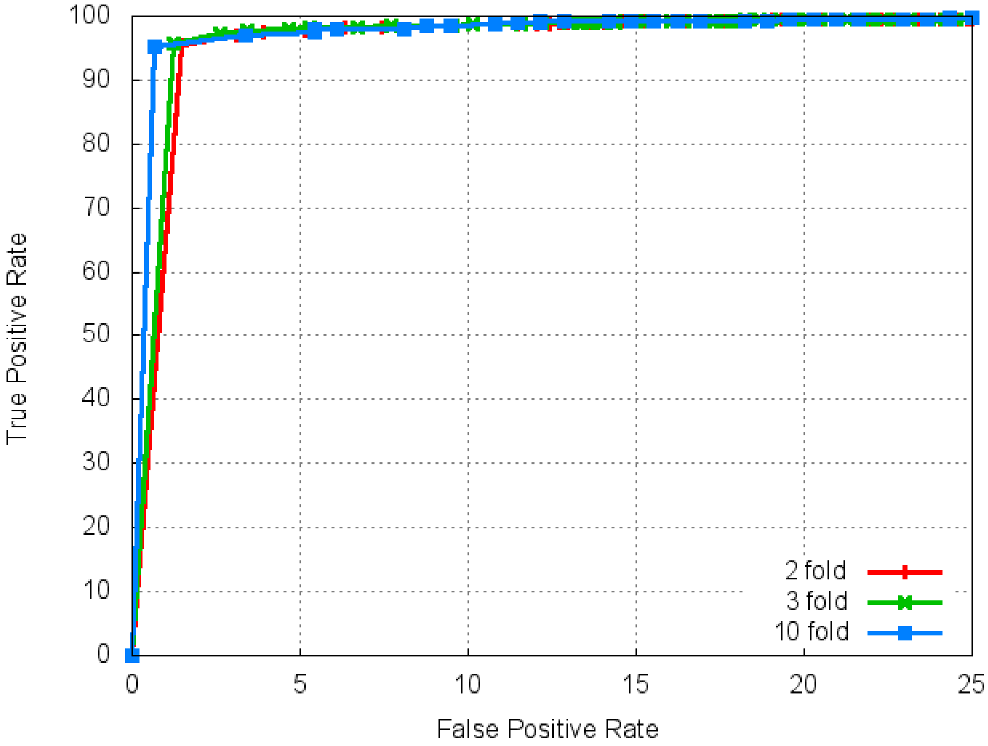
Priors from ground truth (1.45% nodes)

- 21.6K known bad domains: 0.99
- 19.7K known good domains: 0.01
- Unknown hosts and domains: 0.5

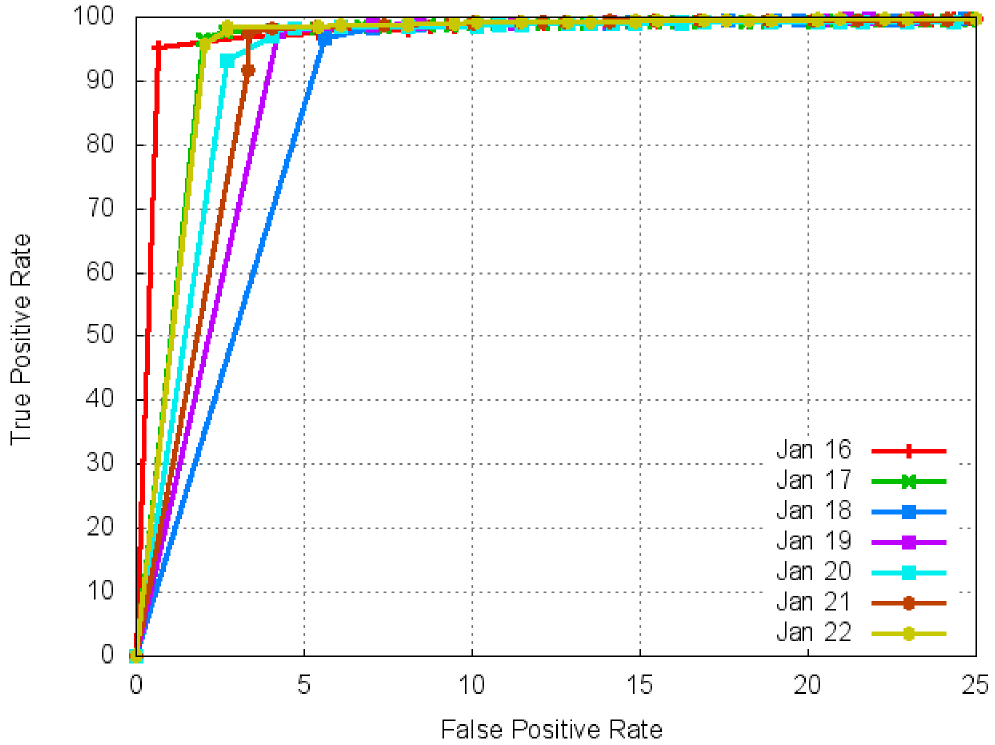
Edge potential

	Benign	**Malicious**
Benign	0.51	0.49
Malicious	0.49	0.51

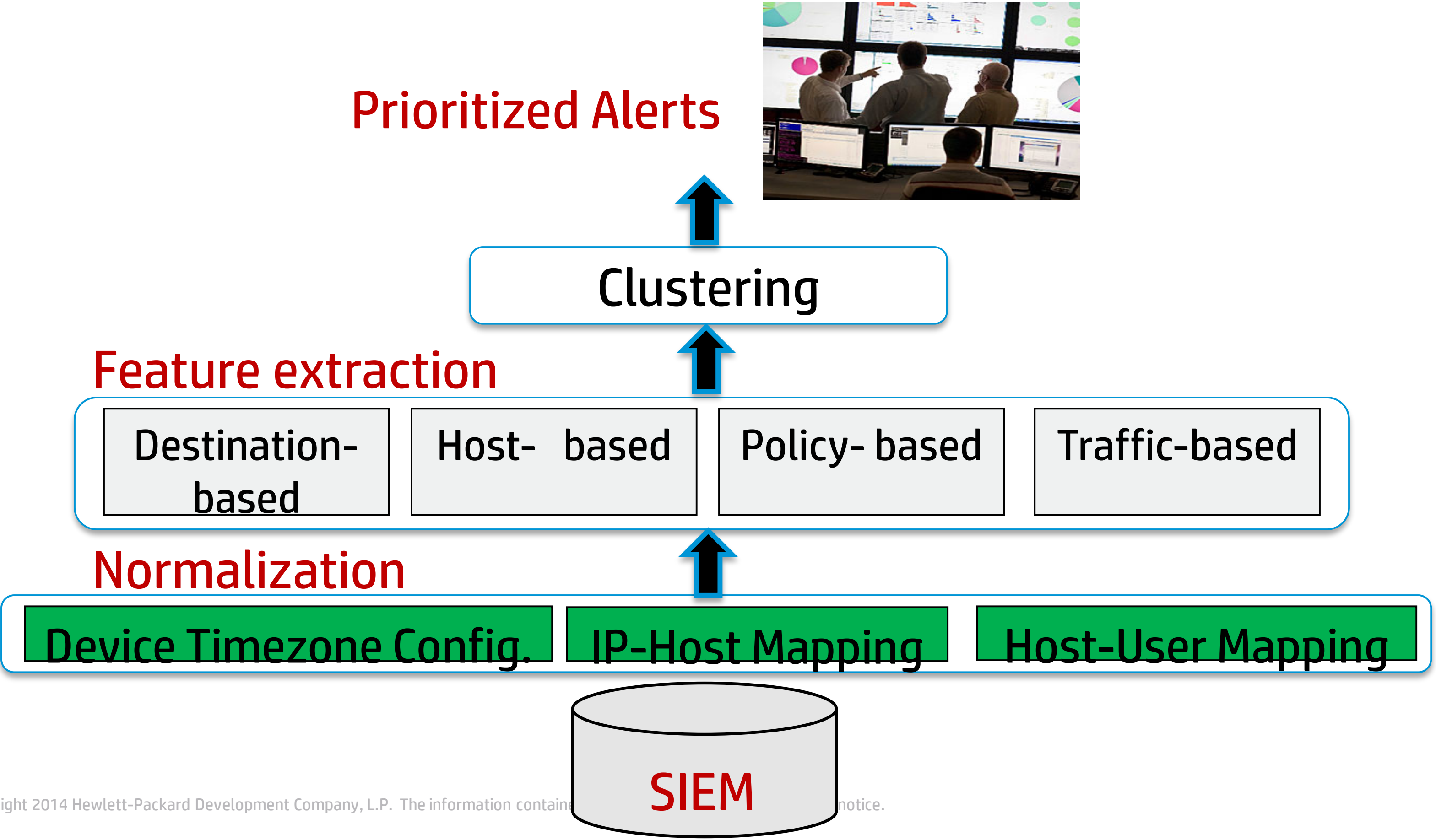
Domain detection ROC plot



ROC plots for seven days' data



Beehive [Yen et al., ACSAC13]



Prioritized Alerts

Clustering

Feature extraction

Destination-based

Host- based

Policy- based

Traffic-based

Normalization

Device Timezone Config.

IP-Host Mapping

Host-User Mapping

SIEM

Parting thoughts

A significant Industry problem

Could benefit from academia

Engineering and algorithmic challenges

Thank you

Acknowledgements: Jorge Alzati, Sandeep Bhatt, Stuart Haber, William Horne, Doron Keller, Prasad Rao, and Loai Zomlot

manadhata@hp.com