RSA Conference 2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: PDAC-R05

Leveraging Analytics for Data Protection Decisions

David Mortman

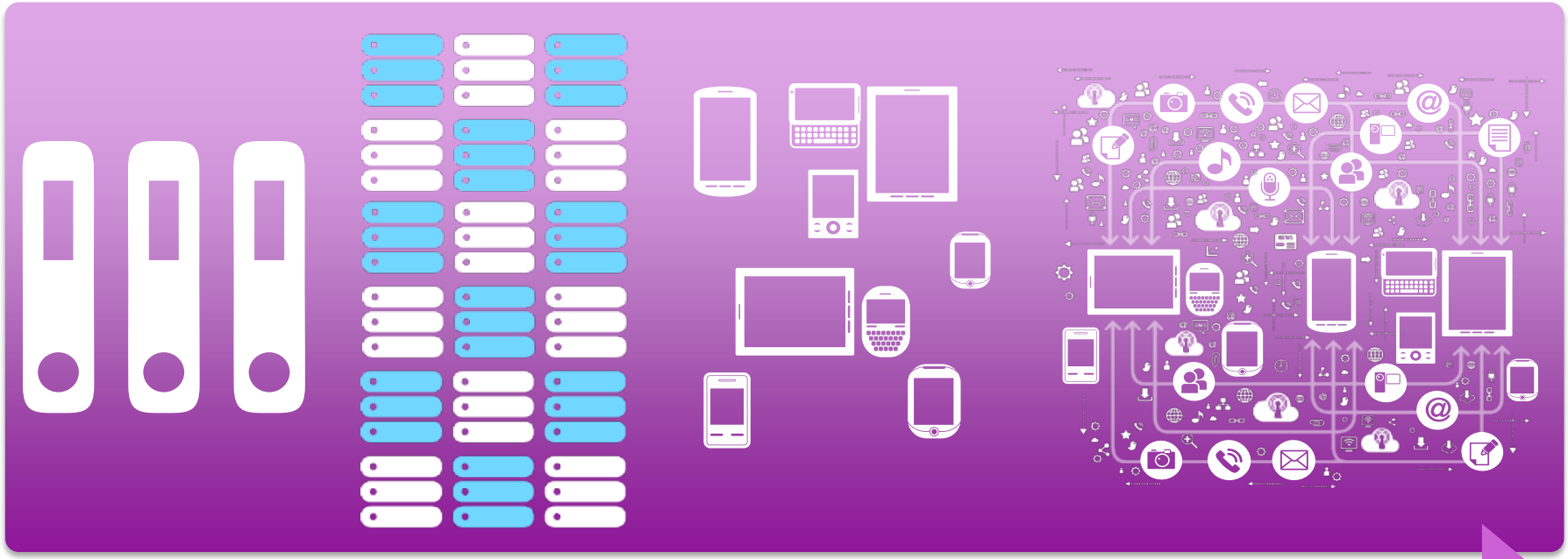
Contributing Analyst
Securosis
@mortman

Chad Skipper

Distinguished Engineer
Dell – Client Solutions CTO
@chadskipper

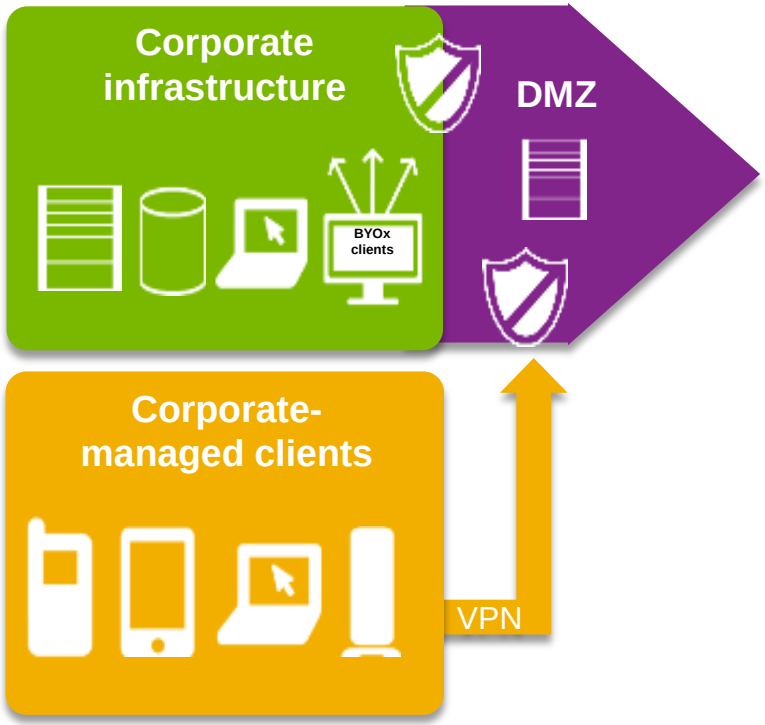
#RSAC

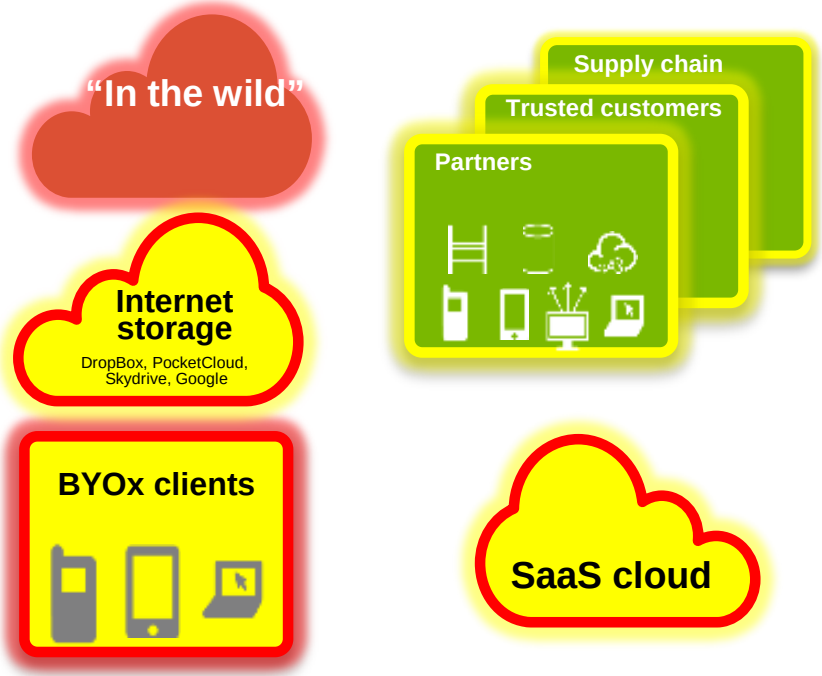
The Data Journey



From mainframe to client server to distributed to *risk everywhere*

Migrations to a Cloud/Mobile/BYOD World





Powerful Disrupters: Data is More Connected

Cloud	**85%**	**Use cloud tools**
Big Data	**35**	**Zettabytes by 2020**
Mobility	**5X**	**Increase in personal owned devices**
Security & Risk	**79%**	**Experienced significant security incident**

Data Risk Models

Inherent risk of data loss

Data access

External net

- Corporate asset
- Corp container
- External network

- Personal asset
- Native (Raw) OS external network

Managed

- Corporate asset
- Corp container
- Semi trusted (VPN/VDI/App)

Unmanaged

- Personal asset container
- Semi trusted (VPN/VDI/App)

- Corporate asset
- Corp image & container
- Corporate network

- Personal asset
- Corp Container
- Corp network

Corp net

Strong control — Weak control

Data location

External net

BYOx

In the wild

Managed

SaaS

Unmanaged

Internet storage

Corporate network

Corporate ecosystem

Corp Net

Strong control — Weak control

Mitigation control weakness

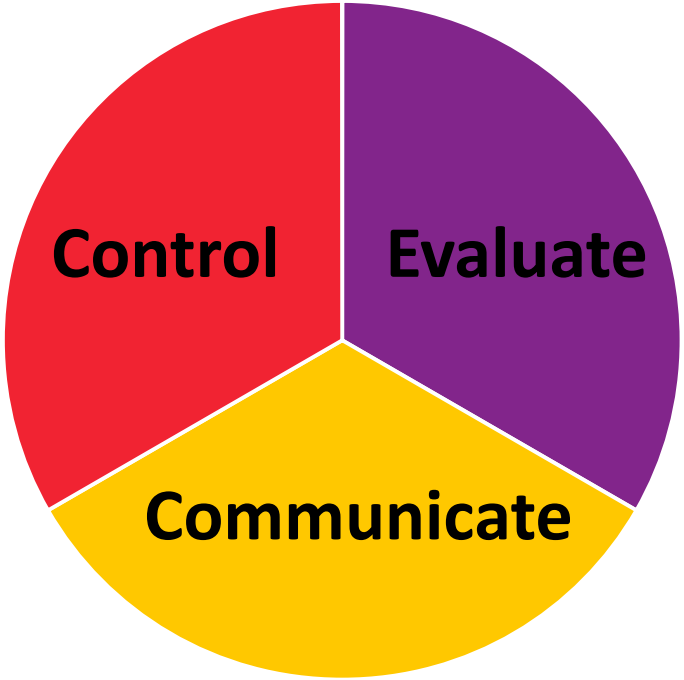
Why DLP Doesn’t Generally Work…

Need to know what you’re looking for before you know what you're looking for…

Static Classification Systems Weaknesses

A Way Forward: Dynamic Analysis

- Dynamic Analysis is the **evaluation of data** in real-time whereby **communicating** meaningful patterns and **awareness in the data** enabling the ability to **control the data**.

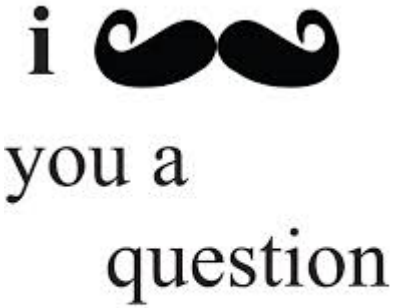


A Way Forward: Evaluation of data in real-time

- Chasing data with no intelligent context and information is a losing proposition.
- **Data needs to become self-aware in order to self-protect.**
- Data is naturally dumb, meaning data just sits waiting for access
- So how would we make data self-aware and self-protecting?
  - Treat the data as if the data were a person

A Way Forward: Evaluation of data in real-time

Mr. Document



Who have you met?
What have you been on?
When was your origination?
Where have you been lately?
Do you have any transmitted diseases?
Where is your wrapper?

Data Becoming Self Protecting & Aware

Contextual access assessment

- Identification processing
- Endpoint platform assessment
- Connection allowance assessment

Data access policy & Identity verification

- Data classification access assessment
- 2F Identity verification
- Policy selection
- Resultant Set of Policy Resolution (RSOP)
- Policy conflict resolution

Enforcement controls

- Firewall
- DLP
- AV/AM
- Network segmentation
- Sandboxing
- Containerization
- VPN
- Virtualization
- Secure browser

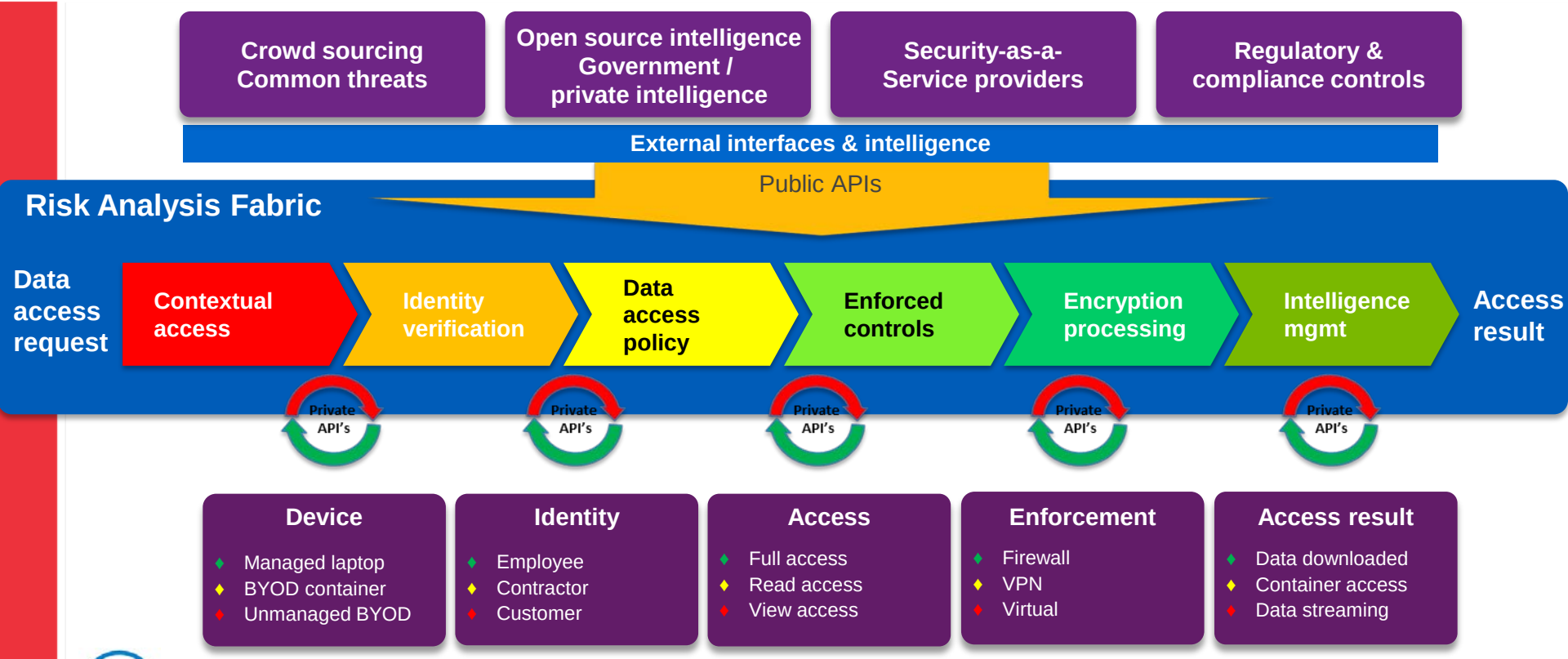
Encryption processing

- Key management
- Platform assessment
- Decryption processing
- Encryption processing
- Key storage
- Process closure

Dynamic Analysis

- Audit/compliance
- Mitigation selection, activation & monitoring
- Ongoing monitoring & policy adherence
- Data integrity monitoring
- Session monitoring & closure control
- Self-Protecting

Data Protection Reference Architecture



A Way Forward: Contextual Information

- Getting context from meta-data
  - Geo-ip (Where)
  - Device type (What)
  - Device context (What)
  - Who is accessing (Who)
  - Usage patterns (When)
  - Mode of access (How)

A Way Forward: Math is a Solution

- We are NO Data Scientists but we believe Math is a solution
  - Clustering Algorithms: k-means
  - Spatial-Temporal contextual data based upon Who, What, When, Where
  - Apriori: Association Rules enables alignment frequency of contextual data
  - Naïve Bayes is a popular (baseline) method for judging documents as belonging to one category or the other
  - Supervised, Unsupervised, Semi-Supervised

A Way Forward: Dynamic Analysis

- Visualization
  - What kind?
  - How useful?
  - D3 (https://github.com/mbostock/d3/wiki/Gallery)
  - Tools like Splunk/ELK
  - Why no traditional SIEM or GRC?

A Way Forward: Dynamic Analysis

- ELK
  - Elasticsearch
  - Logstash
  - Kibana
  - https://www.elastic.co/

A Way Forward: Dynamic Analysis

- ELK
  - OSS
  - But Commercially Supported
  - Pro – Highly competitive with Splunk & <<<< $$$$
  - Con – Much less user friendly

A Way Forward: Dynamic Analysis

- Caveat - How anomalous is this particular access? Anomaly score
- This is where it gets tricky
- Sometimes requires retroactive human review

Hybrid – The Best of Both Worlds

- Static analysis can bin a pretty decent chunk and it's fast and easy. (appx 1-sigma/68%)
- Save dynamic for the "hard" stuff
- Combine the two for optimal coverage

Real World Examples

- What this would like in real world
  - Network Traffic Analysis
  - Fraud Detection
  - Dell Risk Engine
  - NGDS

“Apply” Slide

- Data is more mobile than ever!
- You have to protect the data wherever it goes
- Use static for the basics (i.e. 1-sigma)
- Dynamic addresses a much larger range but is “fuzzier” and so be prepared for more human input.
- Combine the two for a broader more effective solution
- Meta-Data will be Key as will Math (Algorithms)

RSA Conference 2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: PDAC-R05

Thank You

David Mortman

Contributing Analyst
Securosis
@mortman

Chad Skipper

Distinguished Engineer
Dell – Client Solutions CTO
@chadskipper

#RSAC

RSA Conference 2016

San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: PDAC-R05

Backup

David Mortman

Contributing Analyst
Securosis
@mortman

Chad Skipper

Distinguished Engineer
Dell – Client Solutions CTO
@chadskipper

#RSAC