5. DEPLOY

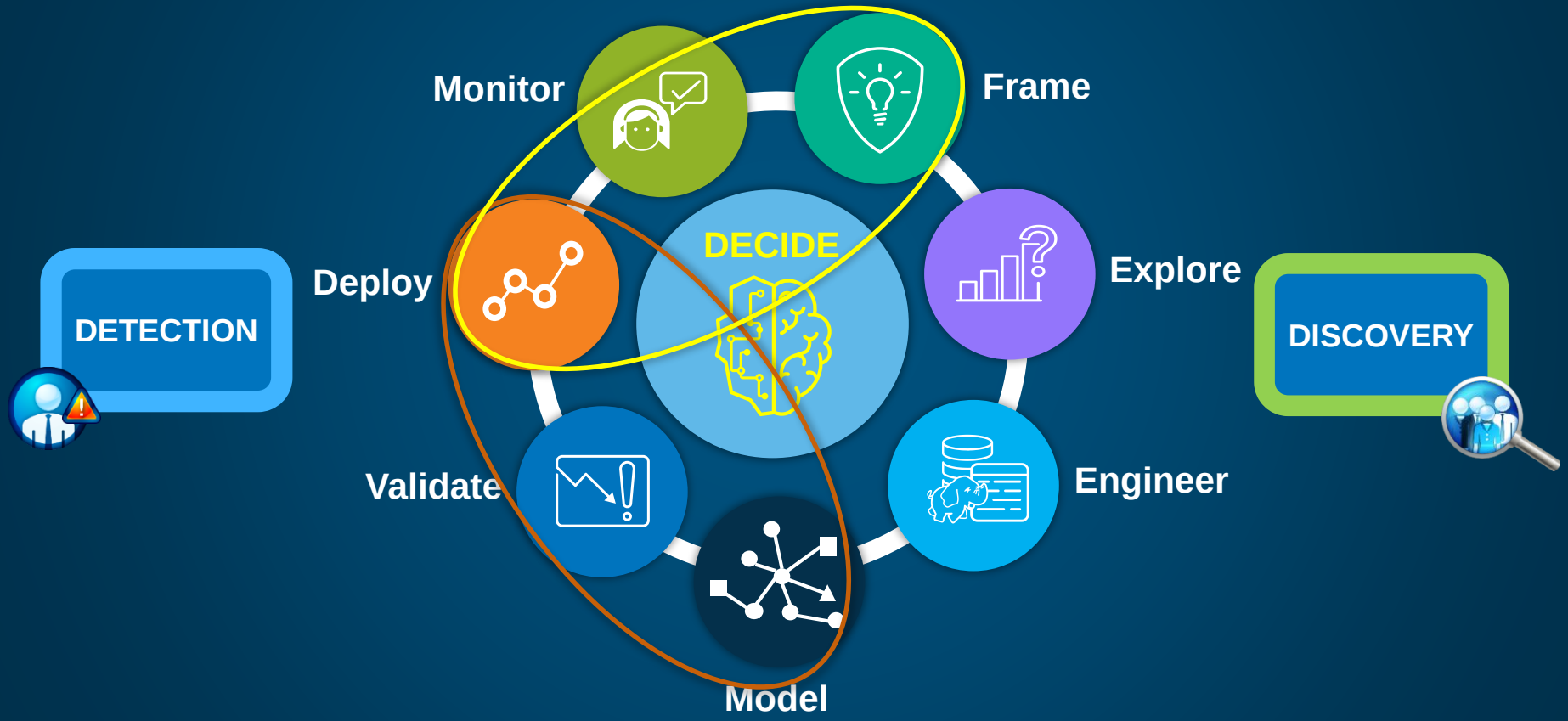
Bringing It All Together

Cybersecurity Data Science (CSDS)

TOPIC
1. FRAME
2. DATA
3. DISCOVER
4. DETECT
5. DEPLOY

Learning Objectives

Cybersecurity Data Science (CSDS) Lifecycle



Objectives of Bringing It All Together

Bringing It All Together in the Enterprise

- Integrating organization, processes, and technologies
  - Analytics process management
  - Integration investigations
  - Optimization of resources
  - Organizational considerations and success factors
- Self-service analytics
- Conclusions / discussion

CSDS Process

Unified Orchestration

DATA

Access

Prepare

Cleanse

DISCOVERY

Explore

Analyze

Model

DETECTION

Govern

Embed

Monitor

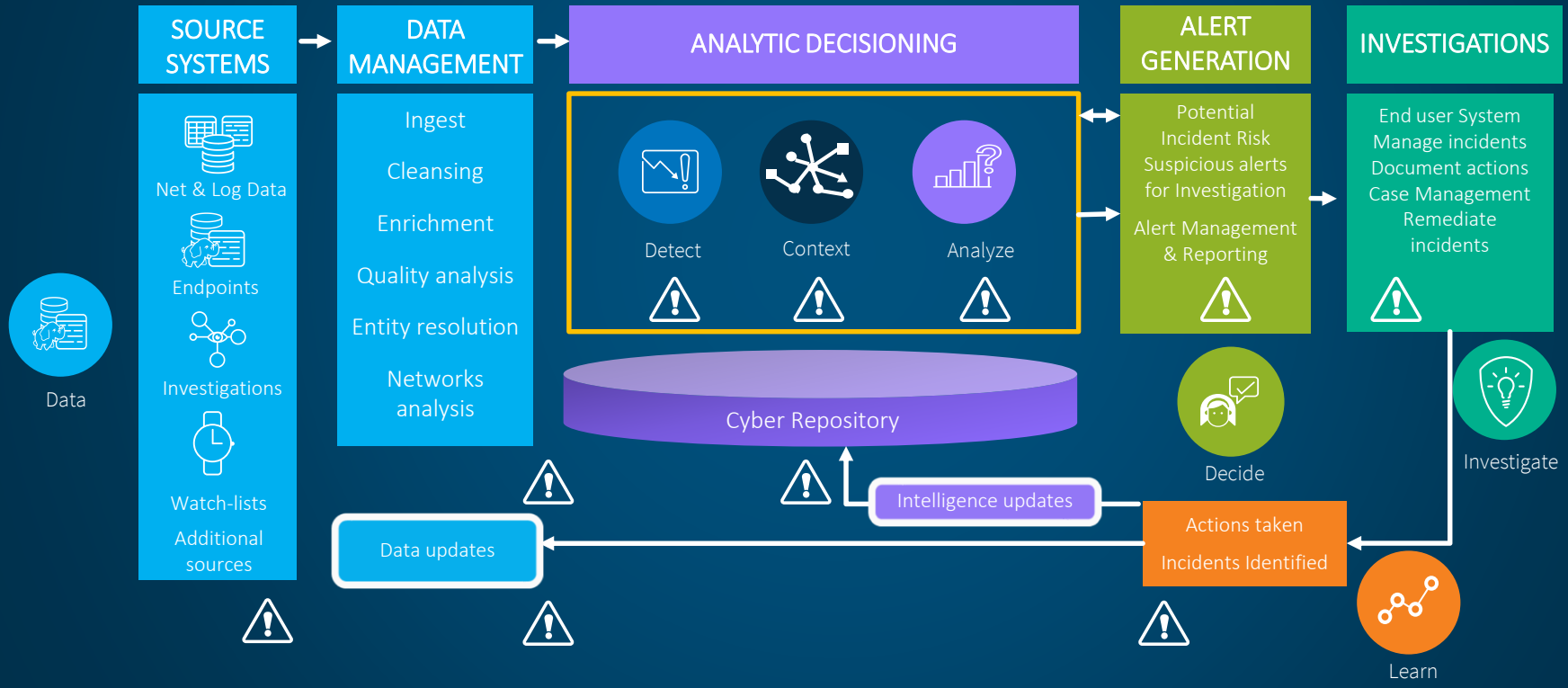
ORCHESTRATION

Enterprise Cybersecurity Data Analytics Architectures

The Big Picture

Defining requirements for adopting cybersecurity analytics

Cyber Analytics Functional Architecture

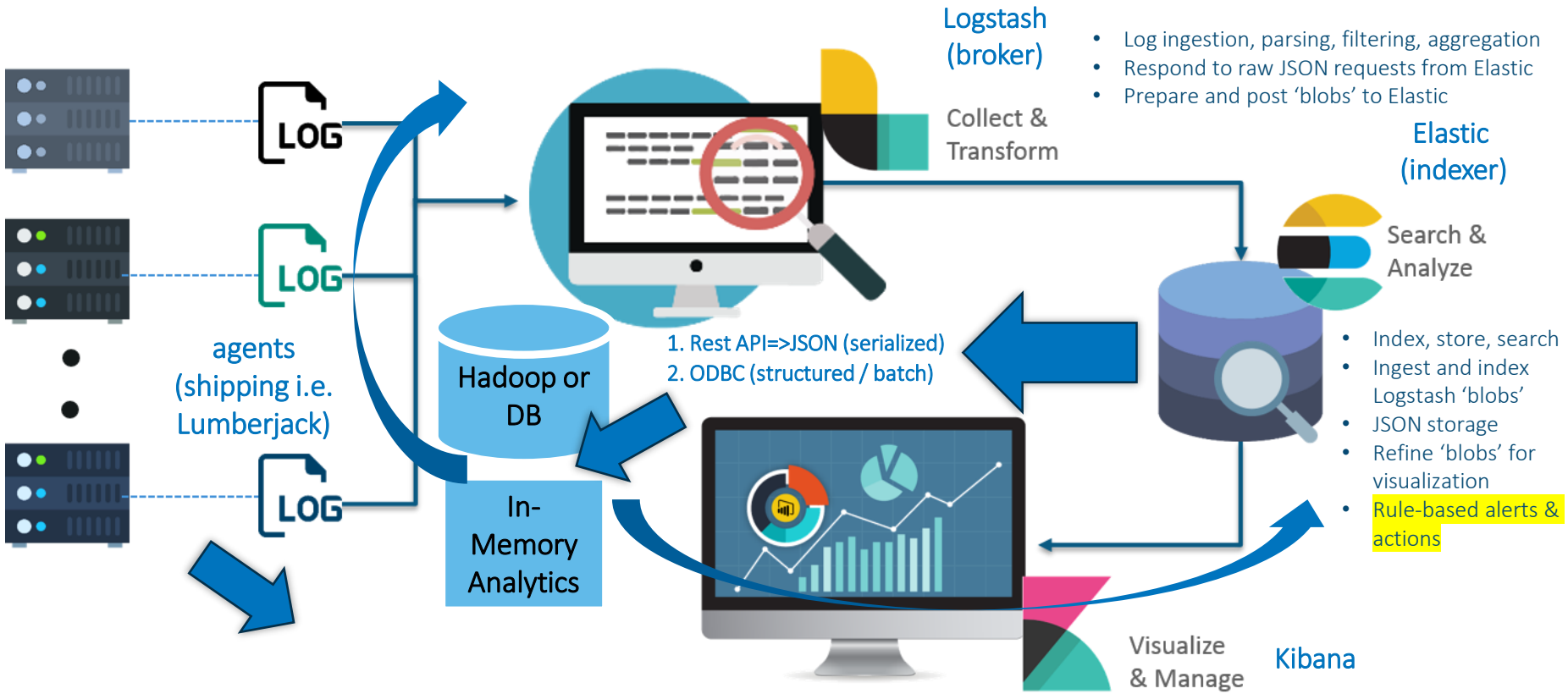


ELK stack as a big data processing platform…

ELK High-Level Functional Architecture

https://www.edureka.co/blog/elk-stack-tutorial/

ELK High-Level Functional Architecture



Data Access

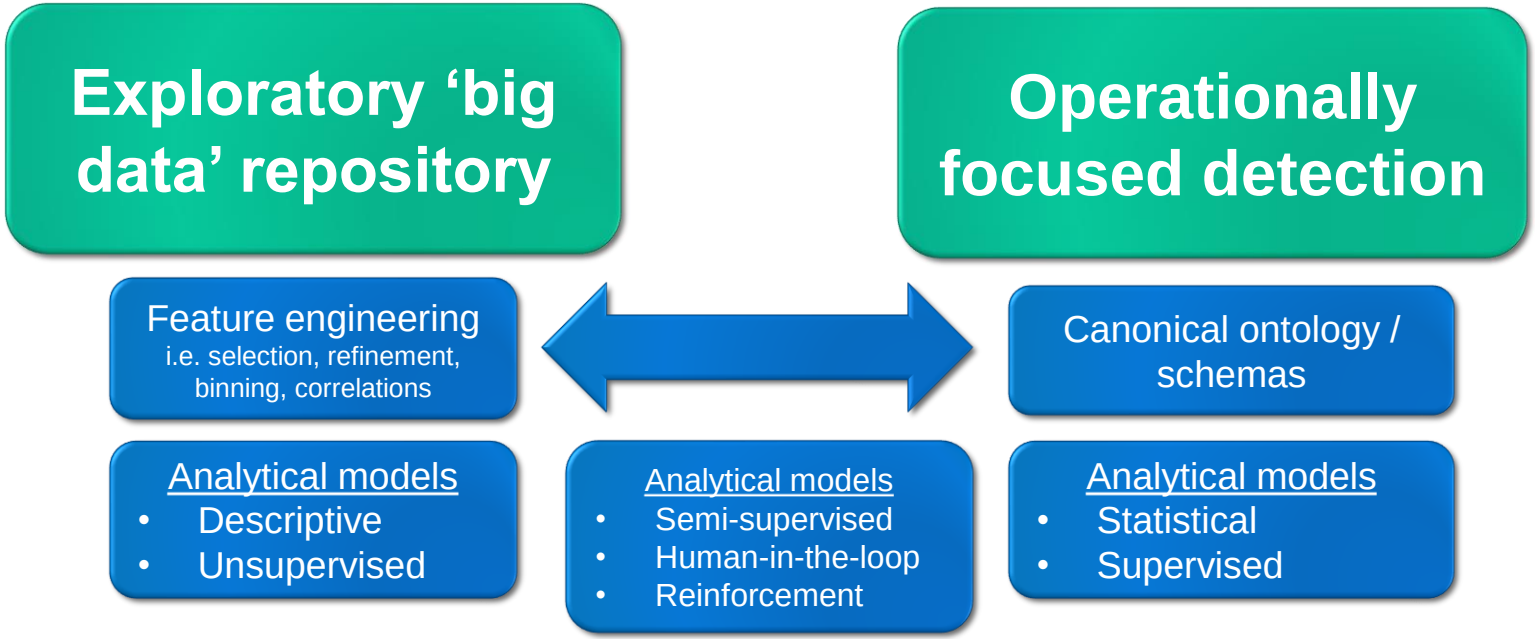
Access Options

Hadoop-based Security Data Lakes

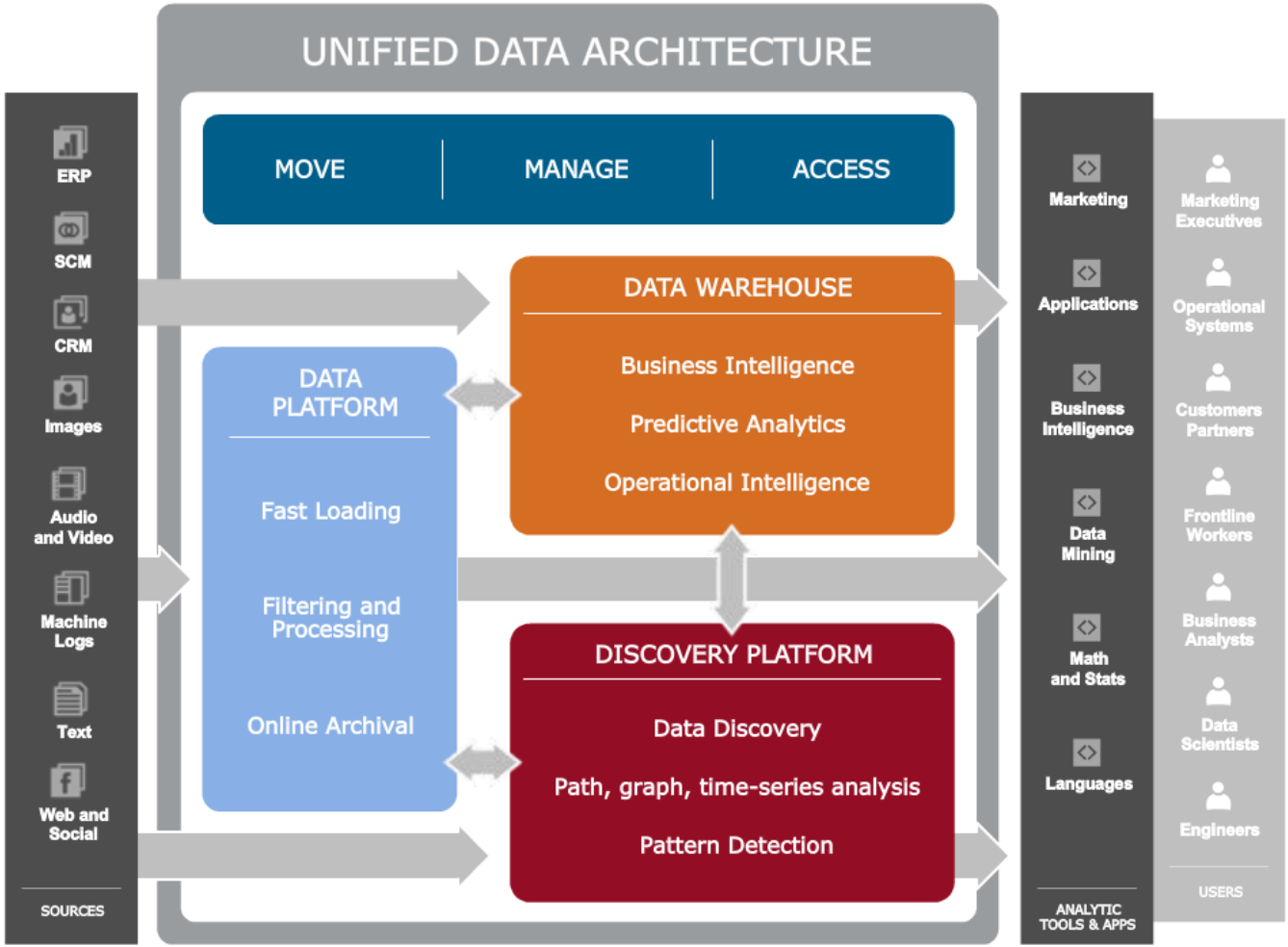
Integrations

Architecture: Exploratory & Detection Platforms*

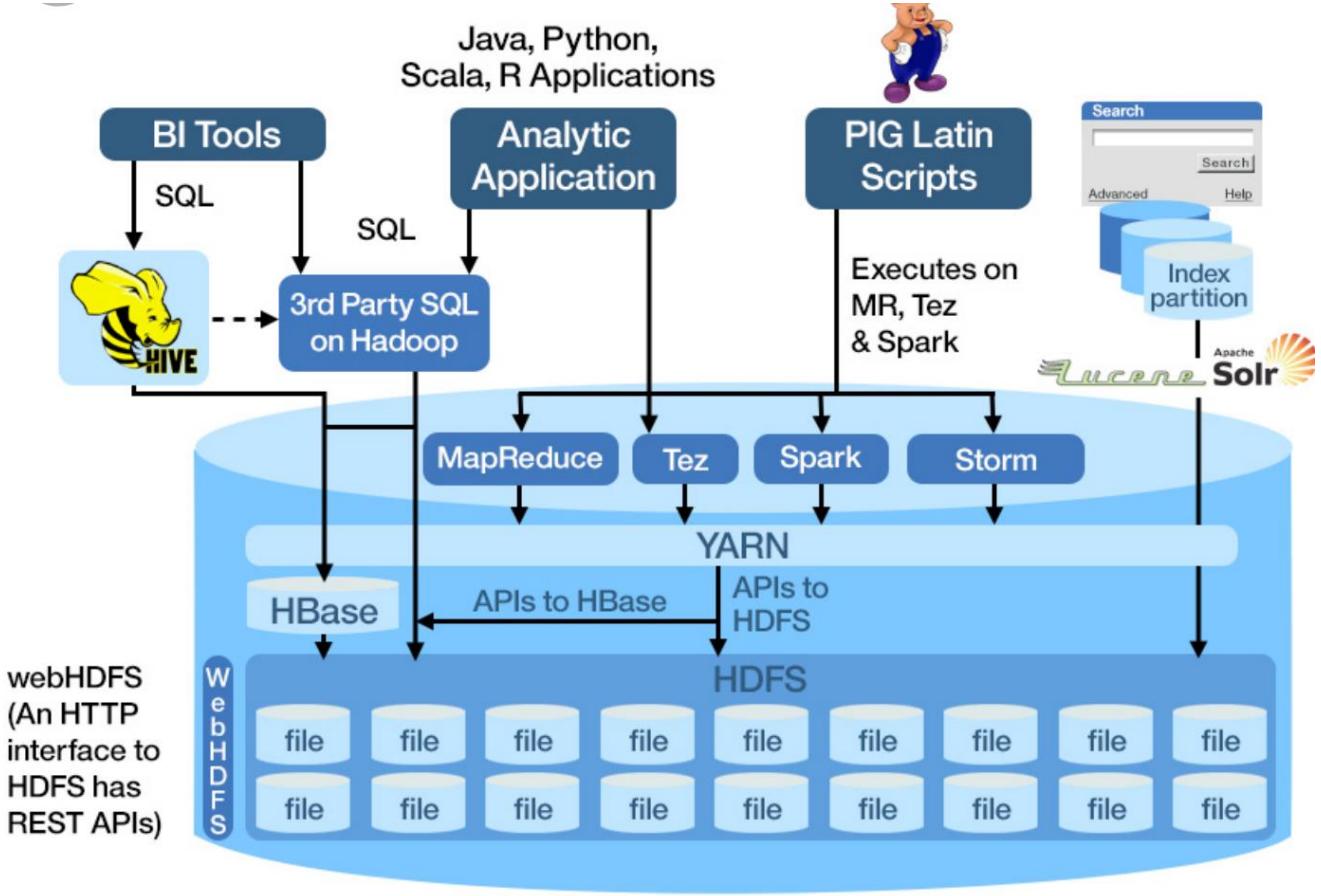
Functional Architectural Segmentation



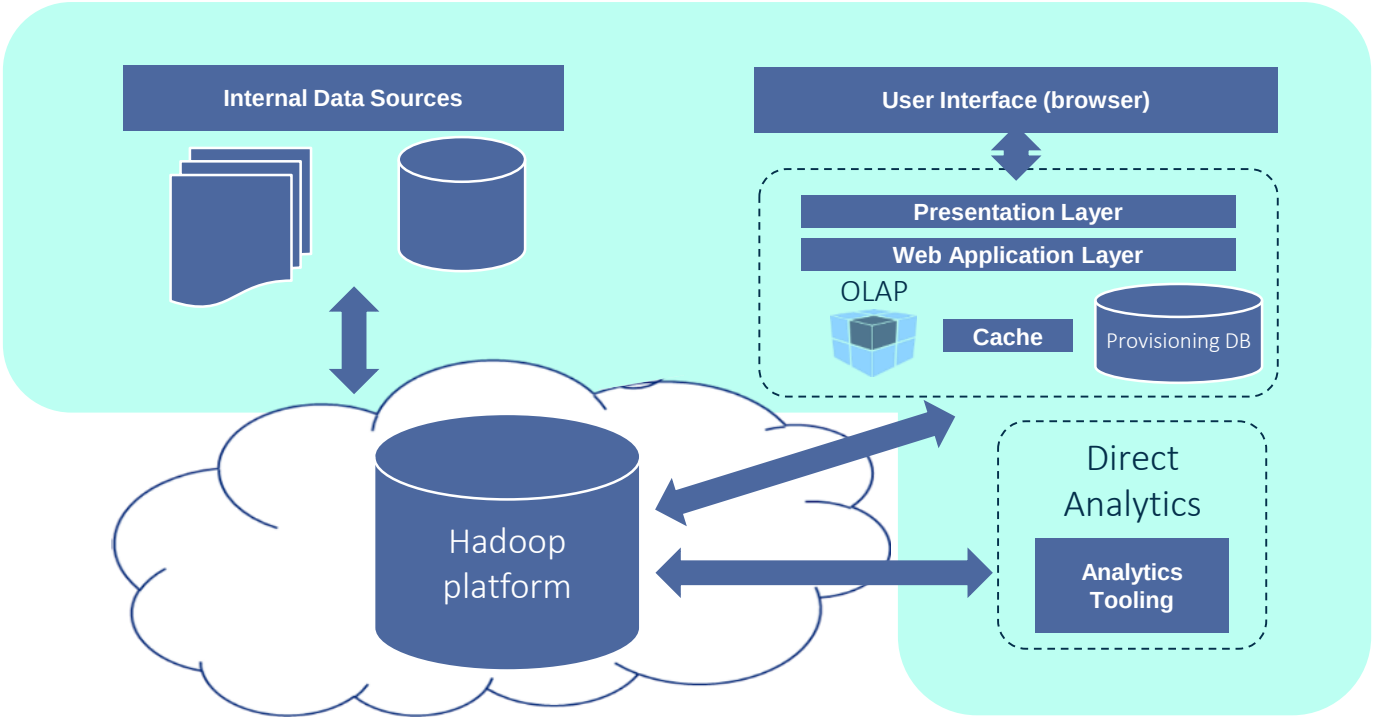
** Runs counter to the vendor stance of store 'all-the-data-all-the-time'*


UNIFIED DATA ARCHITECTURE
MOVE
MANAGE
ACCESS
DATA WAREHOUSE
Business Intelligence
Predictive Analytics
Operational Intelligence
DATA PLATFORM
Fast Loading
Filtering and Processing
Online Archival
DISCOVERY PLATFORM
Data Discovery
Path, graph, time-series analysis
Pattern Detection
ERP
SCM
CRM
Images
Audio and Video
Machine Logs
Text
Web and Social
SOURCES
Marketing
Applications
Business Intelligence
Data Mining
Math and Stats
Languages
ANALYTIC TOOLS & APPS
Marketing Executives
Operational Systems
Customers Partners
Frontline Workers
Business Analysts
Data Scientists
Engineers
USERS

Data lake: Conceptual architecture



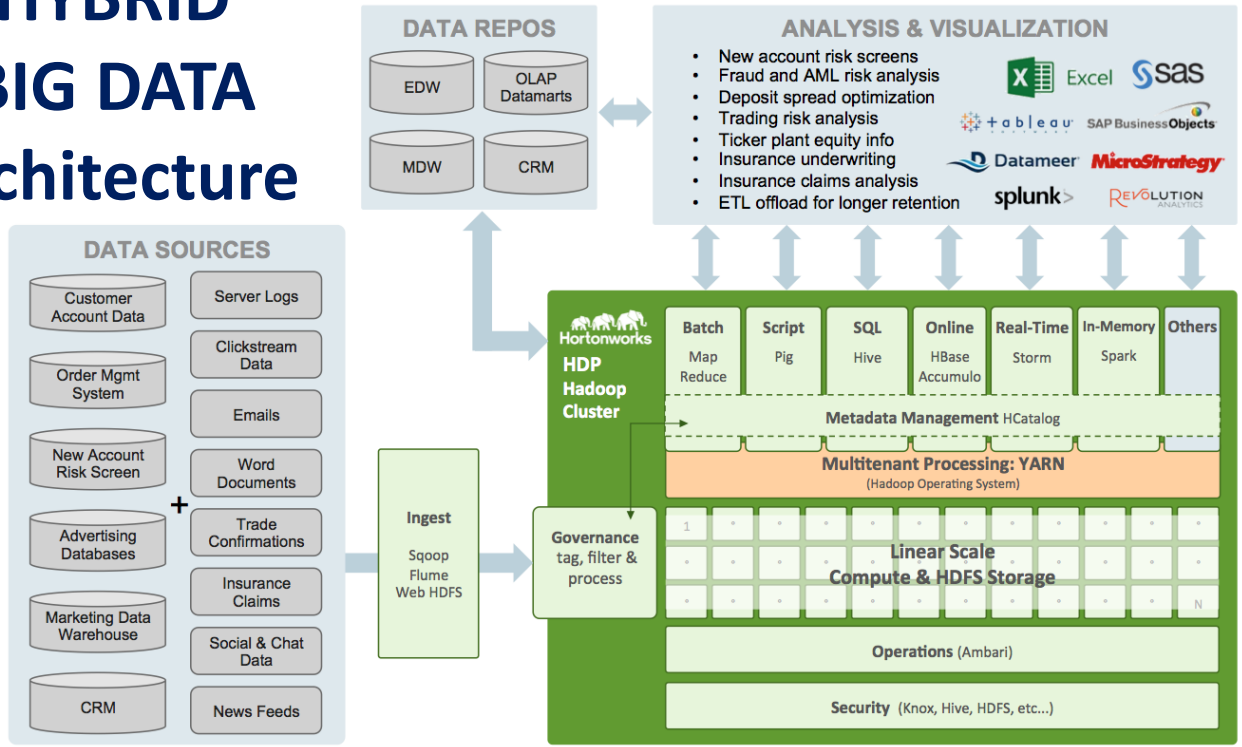
HYBRID INTERNAL & CLOUD



Source - http://www.slideshare.net/AmazonWebServices/analytics-in-the-cloud

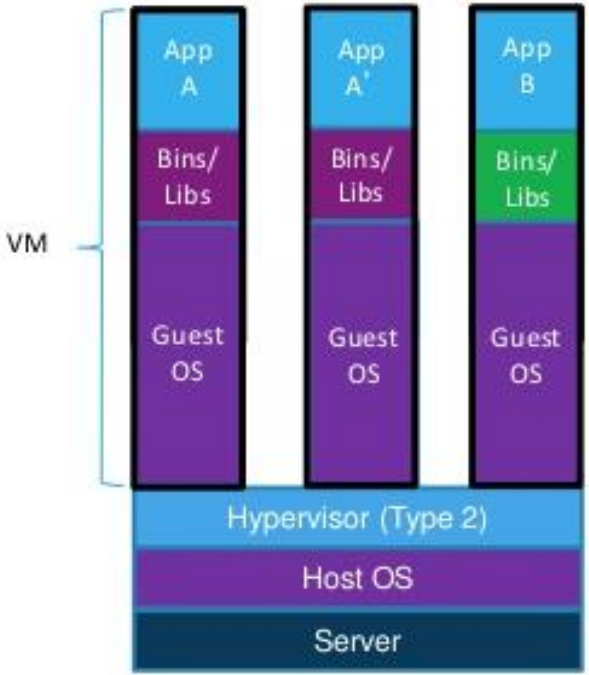
17

Example HYBRID BIG DATA architecture



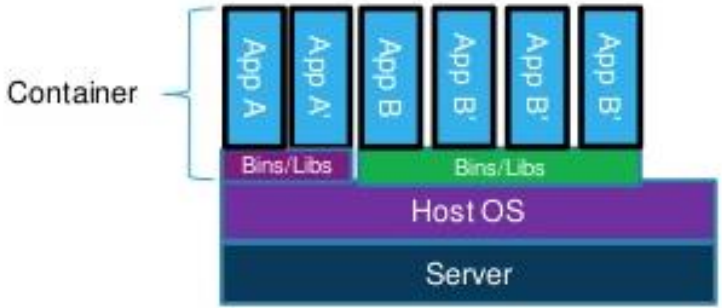
* Horton Works

Virtual machines and containers



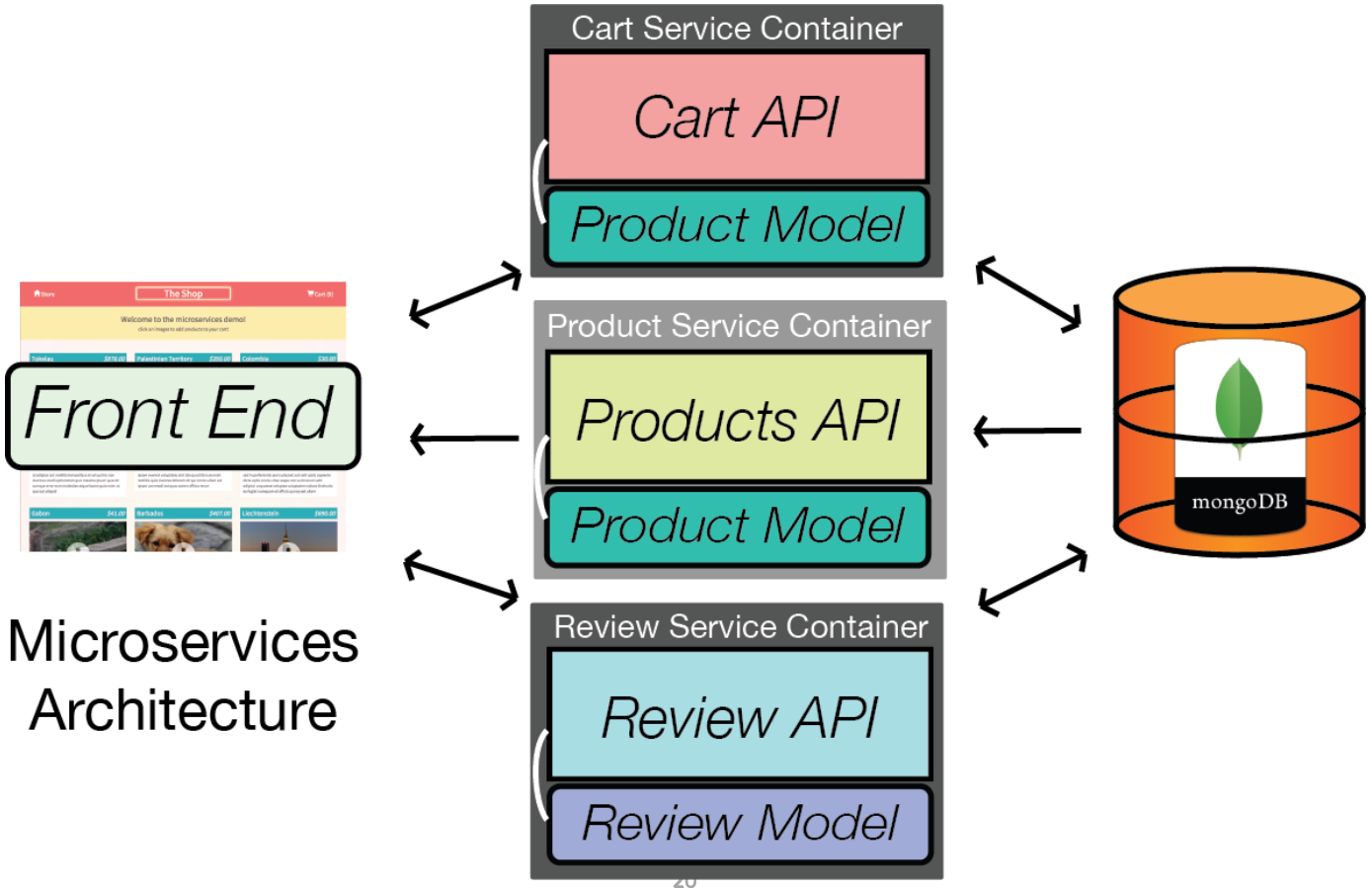
Containers are isolated, but share OS and, where appropriate, bins/libraries

...faster, less overhead



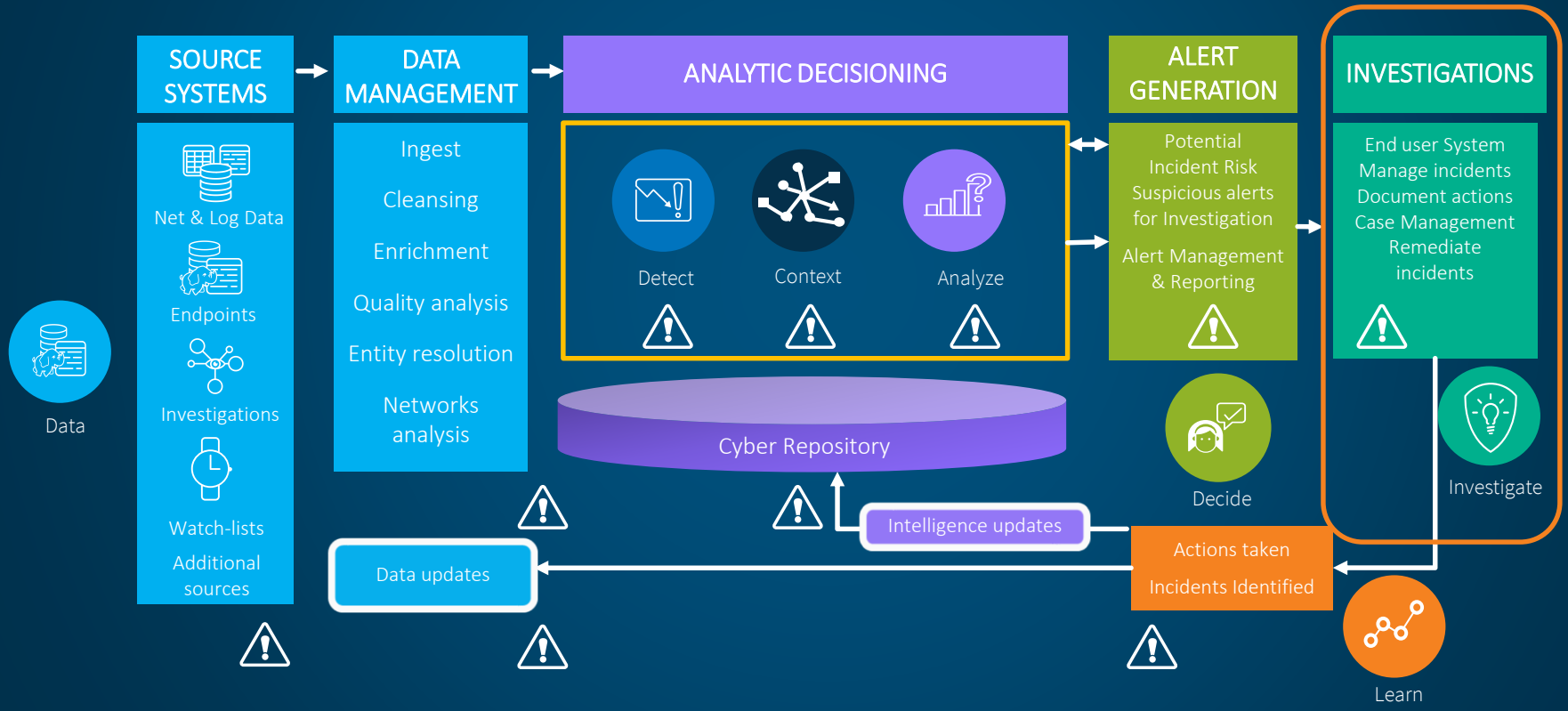
19

Containers and Microservices



http://blog.ibmjstart.net/2015/07/23/learning-microservices-architecture-bluemix-docker-part-1/

Cyber Analytics Functional Architecture



SAS Visual Investigator (VI)

Supporting investigations and remediation

Visualization & Analysis

Data Preparation

Interactive Reporting

Visual Exploration

Location Analytics

Approachable Analytics

Data Discovery

Tech

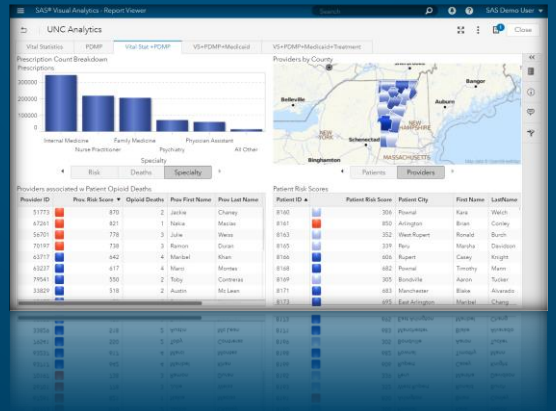
- Data Preparation
- Configuration
- Security Model
- User Management

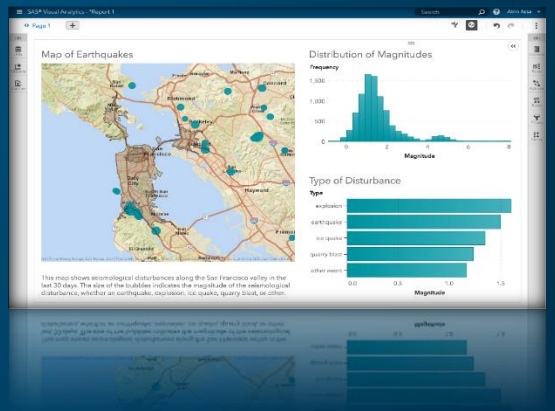
Analyst

- Exploration
- Discovery
- Reporting
- Publish Insights

Command

- Strategic Activities
- Dashboards
- Operational Insight
- Performance





Alert Triage

SAS® Visual Investigator

videmo

Home | Alerts | Search |

Identification Document Expiration | Close

157 Alerts

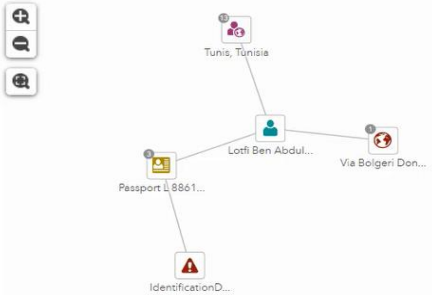
Score	Alert id	Alert type	Entity ID	Entity Type	Alert status	Created date/time
999	Alert_28612526	INSPECT	2500	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_1450493	INSPECT	2501	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_13576138	INSPECT	2502	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_25904303	INSPECT	2503	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_44547399	INSPECT	2504	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_18501901	INSPECT	2505	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_7478788	INSPECT	2506	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_5466406	INSPECT	2507	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM
999	Alert_23478331	INSPECT	1017	IdentificationDocument	ACTIVE	Jan 5, 2017 11:45:43 AM

Scorecard Score: 999

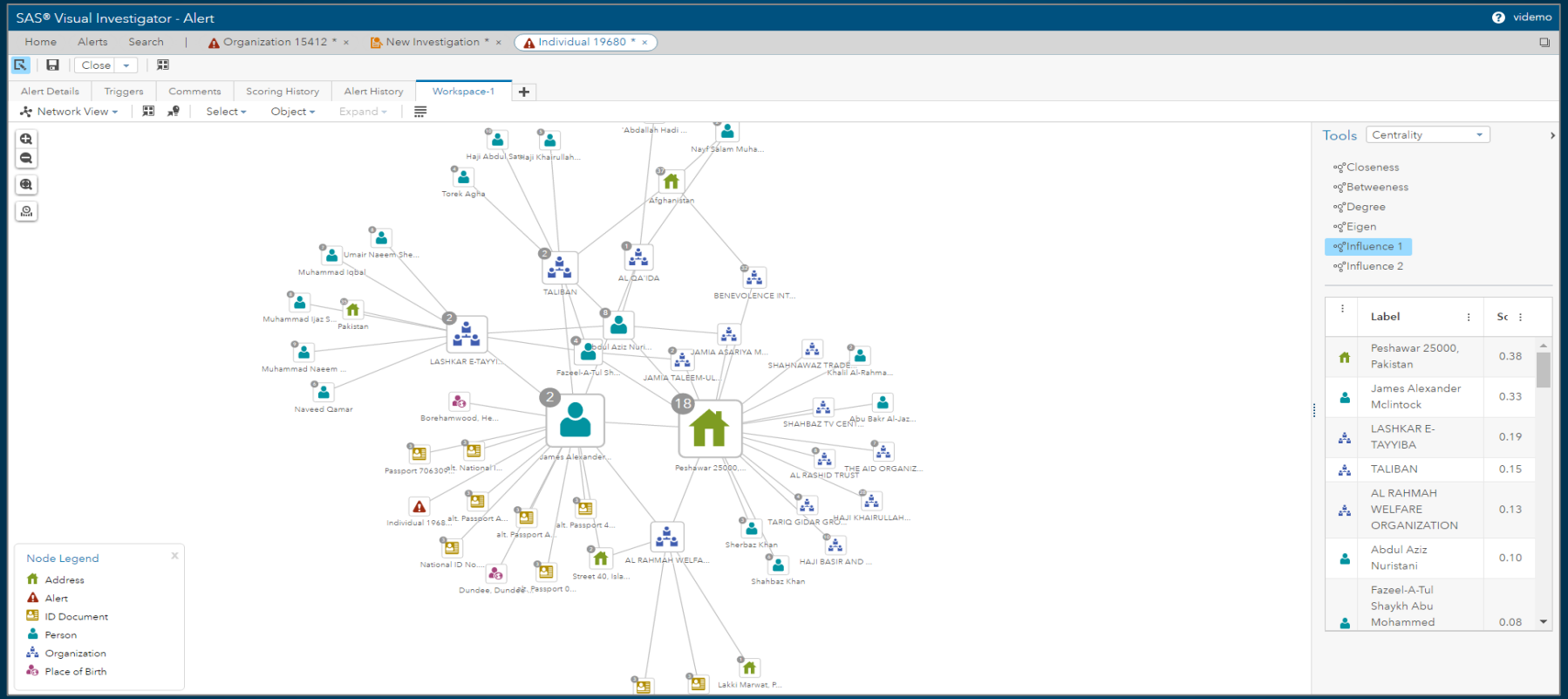
Alert Information

Alert ID: Alert_28612526
Status: ACTIVE
Alert type: INSPECT
Productive: false
Entity ID: 2500
Updated by: batchuser
Entity type: IdentificationDocument
Update date: 01/05/2017
Queue: queue_id_docs

Network



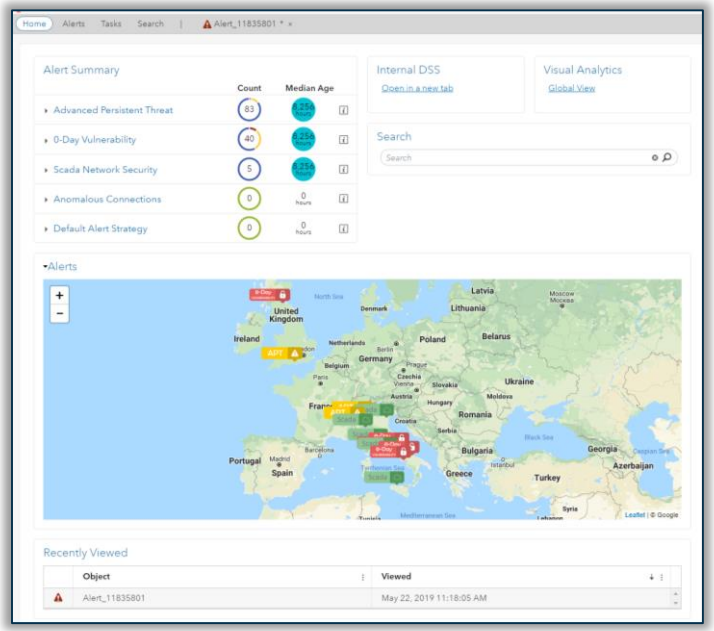
Entity Resolution and Social Network Analytics



Entity Resolution and Analytics can support and direct investigators by showing entity closeness, betweeness, and influence to highlight areas of potential interest.

Visual Investigator (VI): Cyber Investigations

Investigative case management and remediation



http://cyberdyne.racesx07094.demo.sas.com:7980/SASVisualInvestigator/

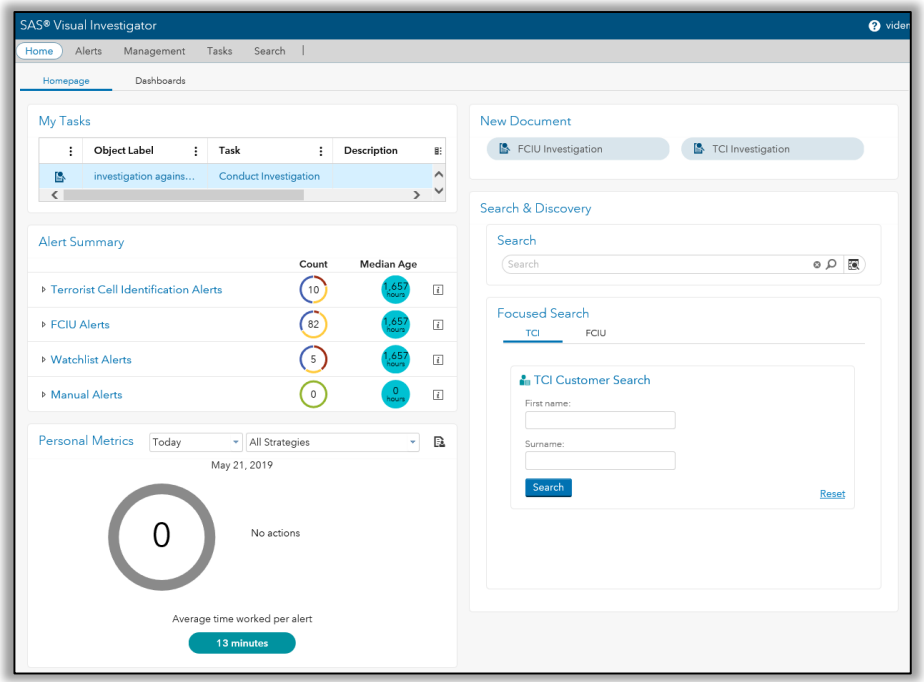
Username: **videmo**
Password: **Go4thsas**

C:\Windows\System32\drivers\etc\hosts

- 172.29.66.238 racesx07094.demo.sas.com acme.racesx07094.demo.sas.com cyberdyne.racesx07094.demo.sas.com intech.racesx07094.demo.sas.com
- 172.29.66.89 racesx08007.demo.sas.com racesx08007

Visual Investigator (VI): Terrorist Cell Investigation

Adjacent Security Example



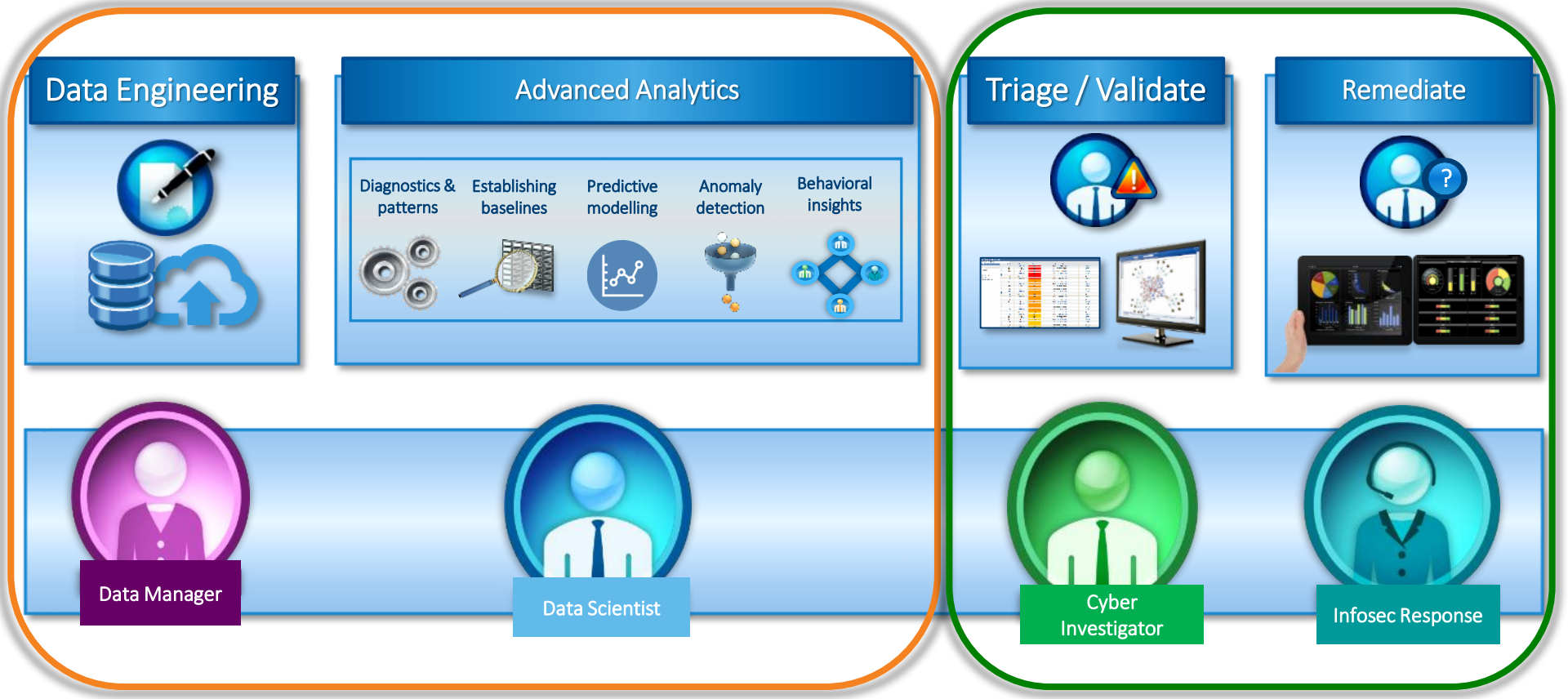
Username: **videmo**

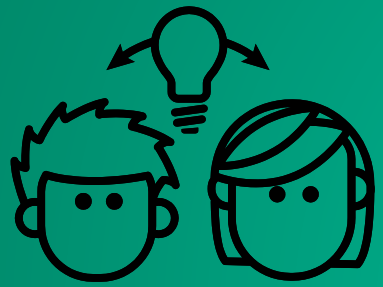
Password: **Go4thsas**

http://fciu.pdcesx15028.exnet.sas.com/SASVisualInvestigator/

Self-Service Visual Analytics

Cybersecurity Data Science as a Process

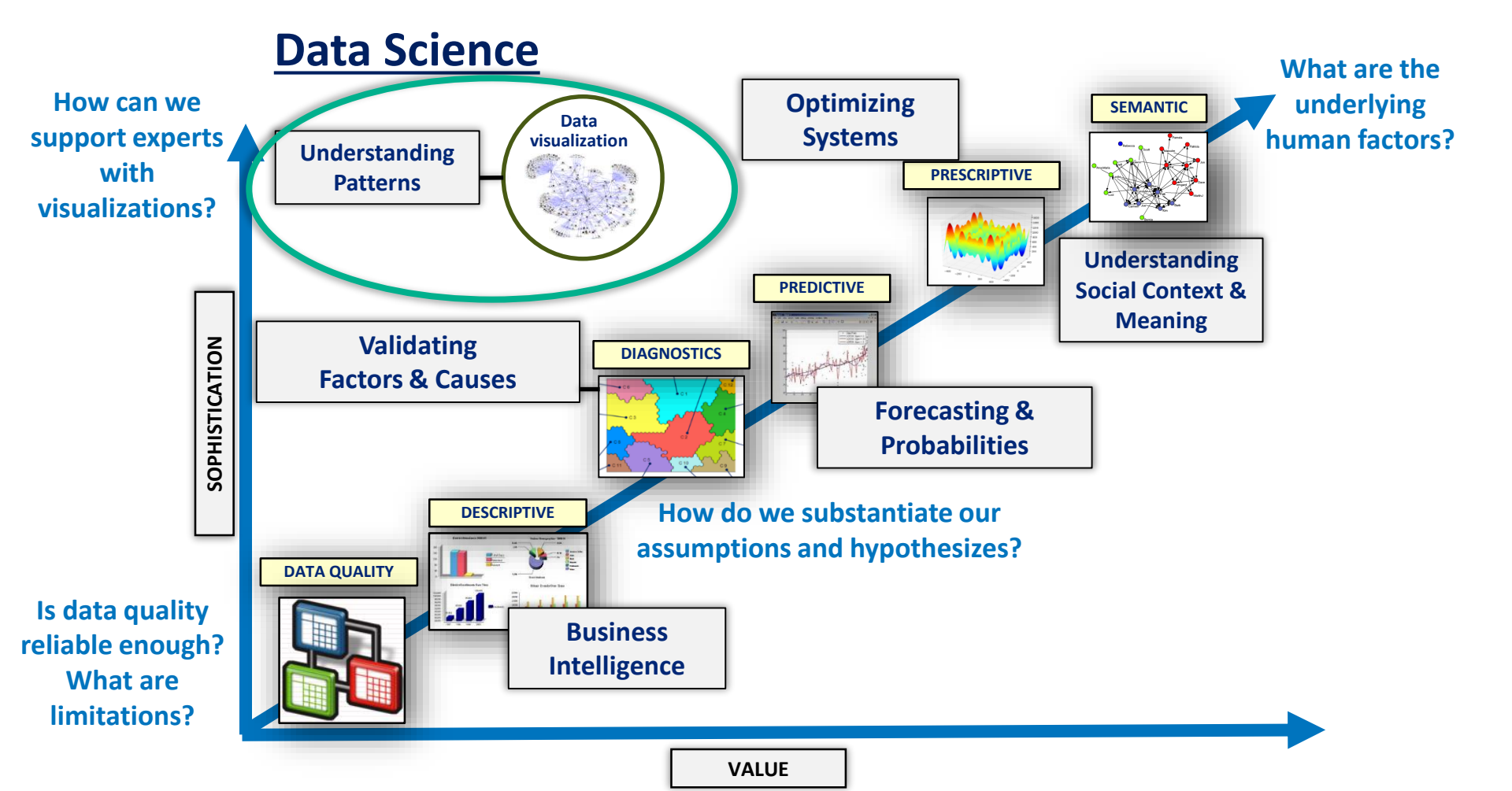




Key Concepts

How can we better connect cybersecurity professionals and data professionals?

Data Science



Visual Analytics (VA)

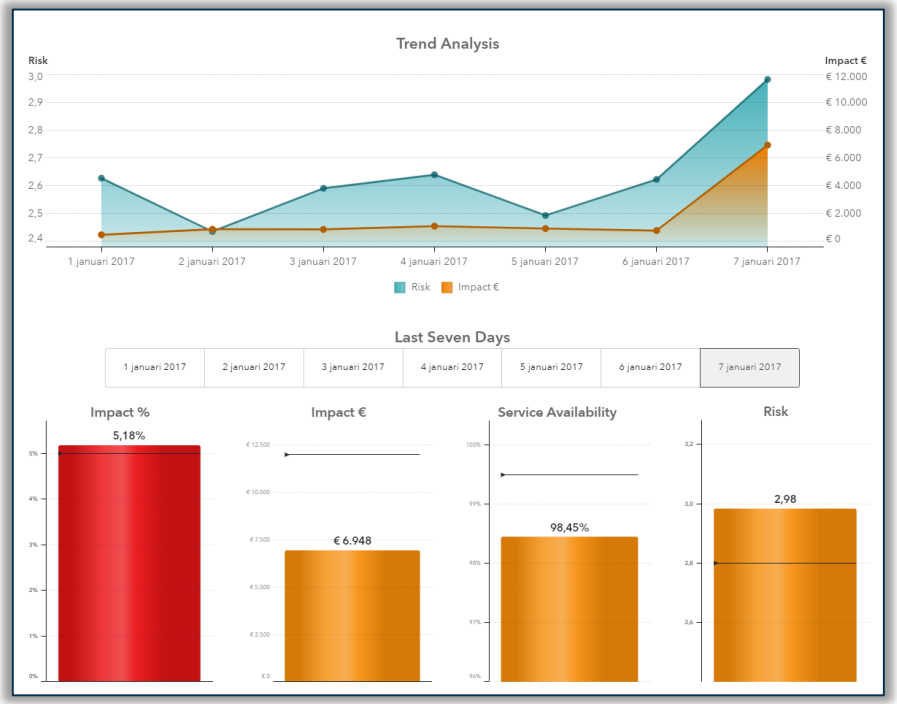
Self-service visual analytics

Cyber Investigator

Self-Service Data Discovery

Visual Exploration and Analytics Dashboarding for Investigators

Simple Cyber Risk Dashboard



http://racesx08007.demo.sas.com:8080/links/resources/report/?uri=/reports/reports/7c443ef2-b83f-4a99-8fc1-350e65a6c618&page=vi6

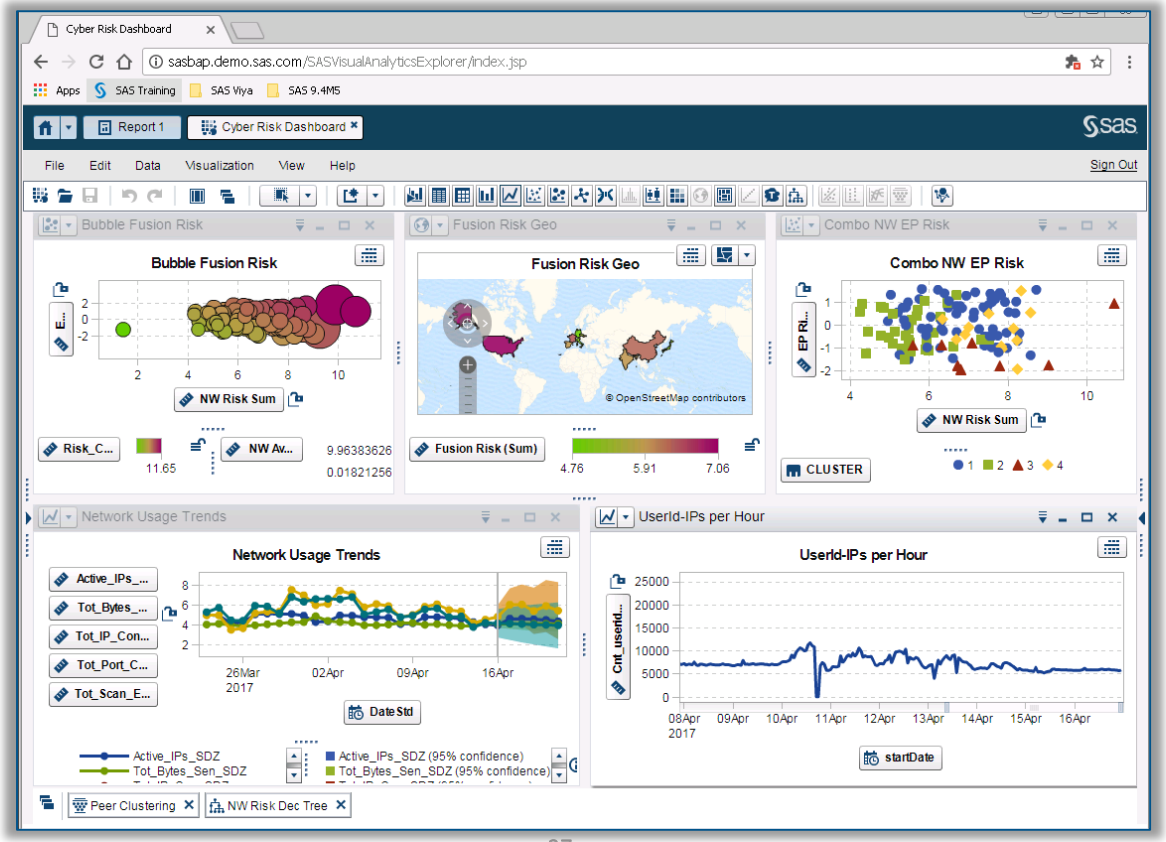
user: sasdemo – password: Orion123

Exercise 1: Cyber Risk Dashboard

Demonstrating self-service visual analytics with cybersecurity data

Dashboard Demo: Network + Endpoint Insights Dashboard

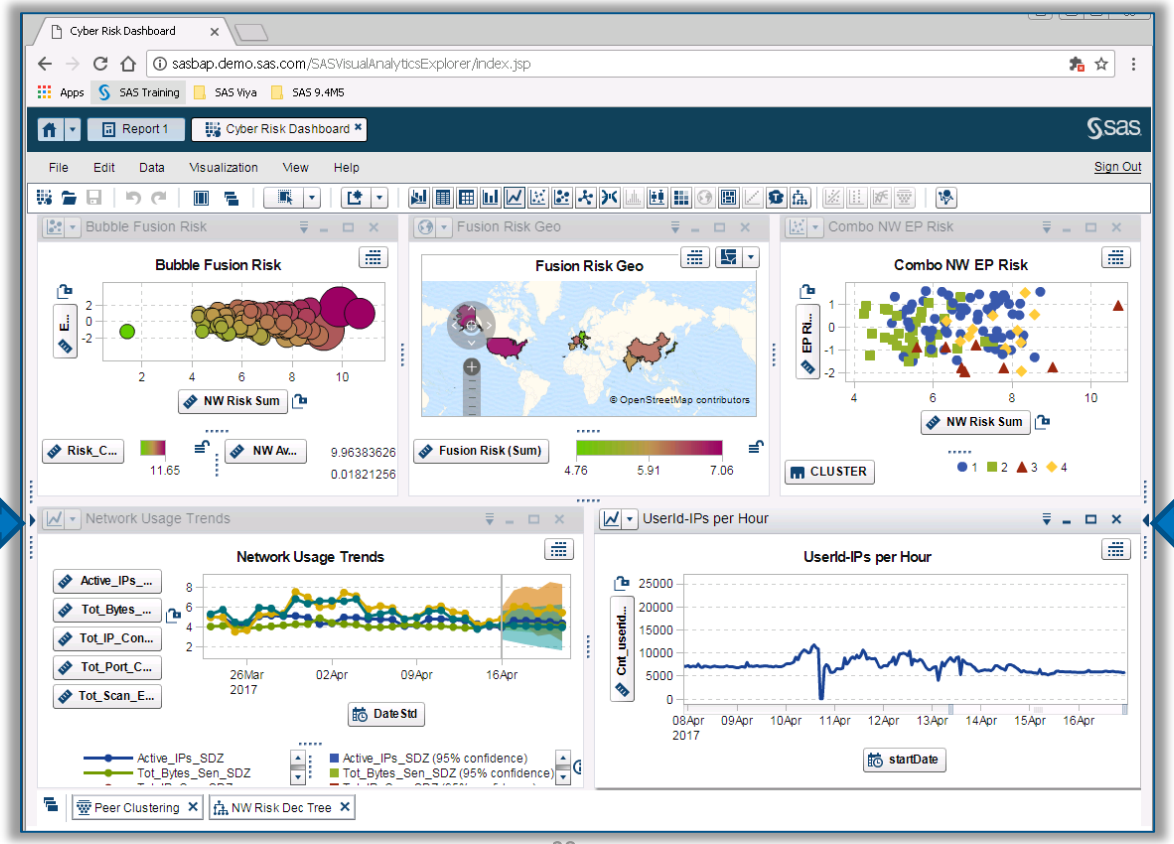
Hands-on Self-Service Analytics



Hands-on Self-Service Analytics

SAS Visual Analytics

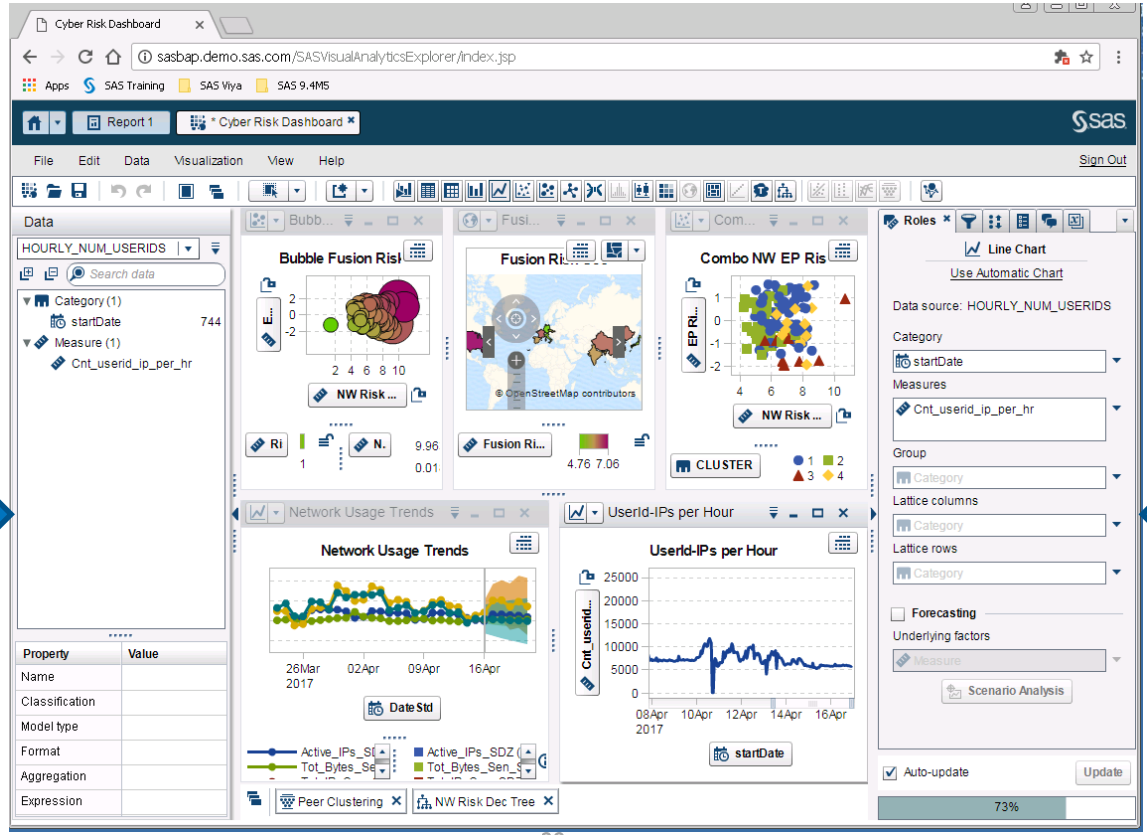
DATA SOURCES



VISUAL CONFIG

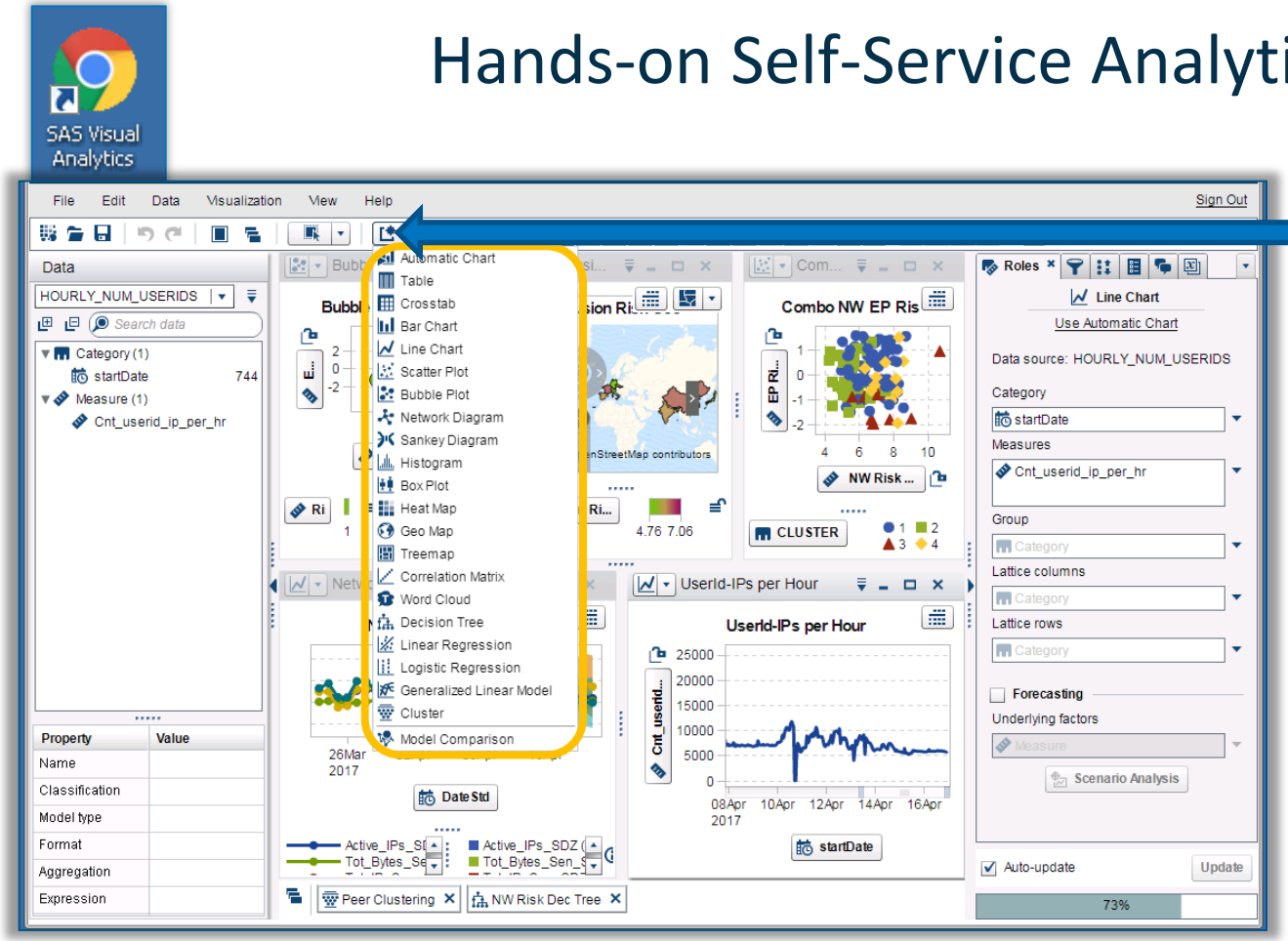
Hands-on Self-Service Analytics

DATA SOURCES



VISUAL CONFIG

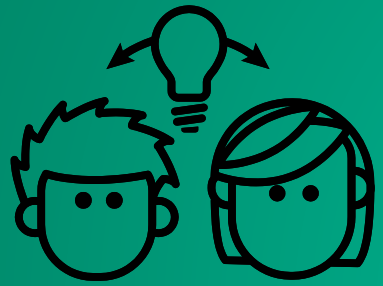
Hands-on Self-Service Analytics



ADD NEW VIEWS

Exercise 2: Honeypot Analytics

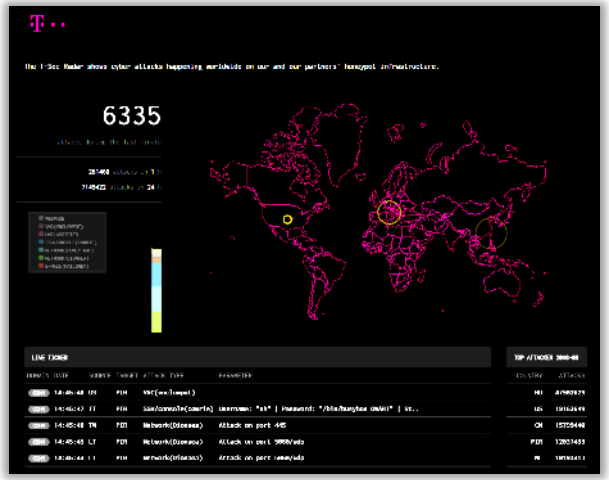
Demonstrating self-service visual analytics with cybersecurity data



Background to Honeypot Analytics

Overview: Honeypot Project

Honeypot Analytics Exercise



Honeypot Project

- ~180 honeypot sensors (~220 peers in network?)
  - Standardized RasPI image (to subsidiaries)
  - T-Pot – modularized approach
- e.g. used to inform customers on infections
- Real-time visualization Sicherheitstacho
- Collaboration with The Honeynet Project

Dataset Profile

6 Dec 2016 - 26 Feb 2018

- ~2.1 years / 111 weeks / 781 days (430 logged)
- ~32 GB (csv)

~74.6M events

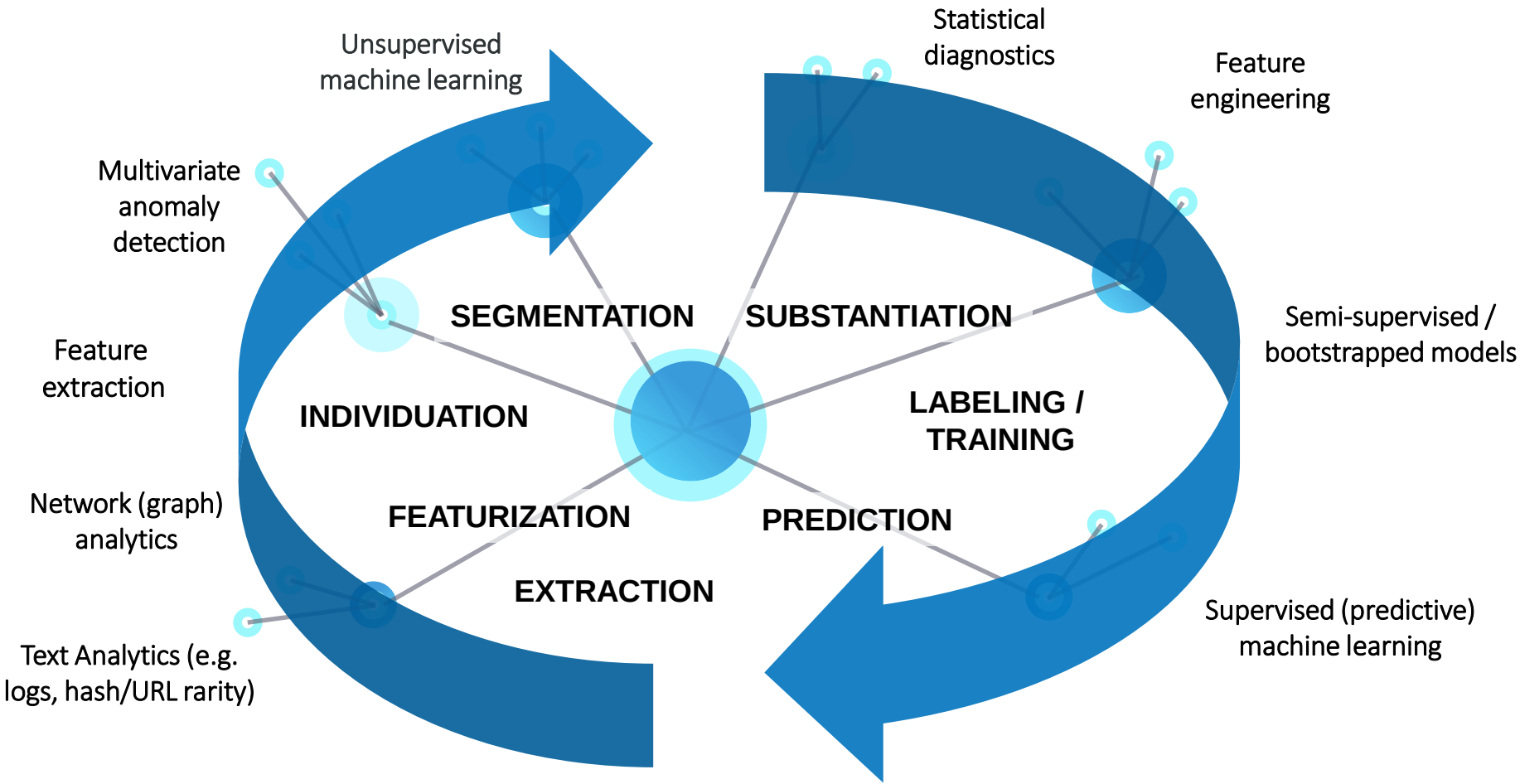
Key figures

- 452,532 unique attacker IPs
  - 40,000 unique network class layers
  - 217 countries (+ territories, etc.)
- 8 Honeypot types
  - 221 HP identifiers
  - 235 HP ASNs
  - 345 HP hostnames

Day of Week	Events	% Total
Sunday	13,764,653	18%
Monday	10,043,485	13%
Tuesday	9,436,768	13%
Wednesday	10,358,282	14%
Thursday	10,505,657	14%
Friday	9,221,546	12%
Saturday	11,260,976	15%
TOTAL	74,591,367	

Hackers = weekend warriors? ☺

Applied Cybersecurity Analytics Process



Overview of Data Processing

Extraction and featurization

Time epochs (hours, days, weeks)

Roll-ups – e.g. hour and day

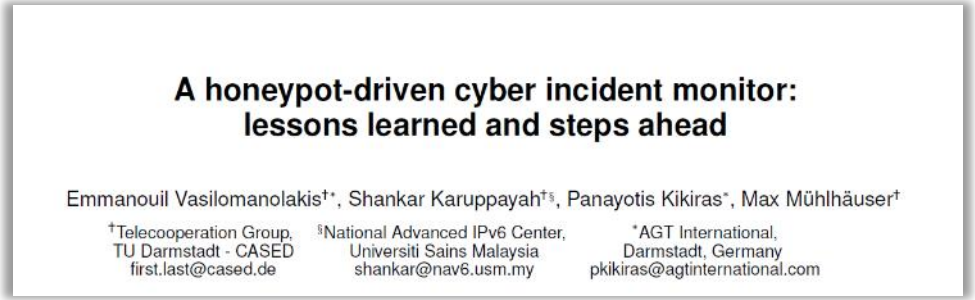
- Distinct types per epoch
- E.g. distinct usernames / passwords attempted per period of time

Binning i.e. Network classes

Port types / protocols

Network graph

- E.g. external IP class-to-honeypot type



A honeypot-driven cyber incident monitor: lessons learned and steps ahead

Emmanouil Vasilomanolakis[†*], Shankar Karuppayah[†§], Panayotis Kikiras[*], Max Mühlhäuser[†]

[†]Telecooperation Group, TU Darmstadt - CASED, first.last@cased.de
[§]National Advanced IPv6 Center, Universiti Sains Malaysia, shankar@nav6.usm.my
[*]AGT International, Darmstadt, Germany, pkikiras@agtinternational.com

- Ratios (i.e. # ports out-to-in)
- Outlier diagnostics
  - Delta comparisons
  - Focused statistical comparisons

Key Findings

Substantial attack

- > 55 million events total during this period
- Monday, February 19th through Sunday 25th (7 days)
- 400,000 IPs participating in attack
  - Narrowed to 40,000 network class segments
  - Narrowed subsequently to 3 key subsegment groups – likely command and control IPs / ranges

Honeypot peer types targeted

1. Network (Dionaea): low-interaction – capture payloads / malware
2. Network (honeytrap): observing novel nw service attacks – dynamic servers
3. SSH/console (cowrie): captures SSH & telnet connections
4. VNC (vnclowpot): low-interaction – listens on port & logs VNC Auth challenge
5. Webpage: webserver presenting host

Highly Active IP Network Ranges (classes)

Most active network ranges during mass attack

450k unique attacker Ips => 40k IP Class NW Ranges (A, B, C)

IP NW Range	Events	CD
5	8,178	4,851
37, 109, 62	16,850	9,427
10, 85, 89, 46, 95, 51…	53,246	29,867
ALL OTHERS	780,191	447,677

** 10, 85, 89, 46, 95, 51, 37, 62, 185.222, 77, 87, 222.88.69, 123*

Attributions

Self-propagating command and control-driven botnet malware

- Find open ports (high port scans)
- Guess passwords (high unique password attempts)
- Telnet port 23 open

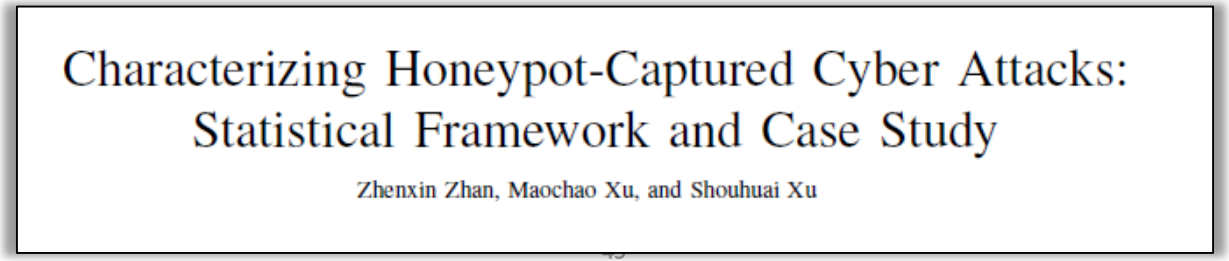
Susceptible to autocorrelation-based diagnostics

- Examining lags in events and periodicity between events

Operational considerations

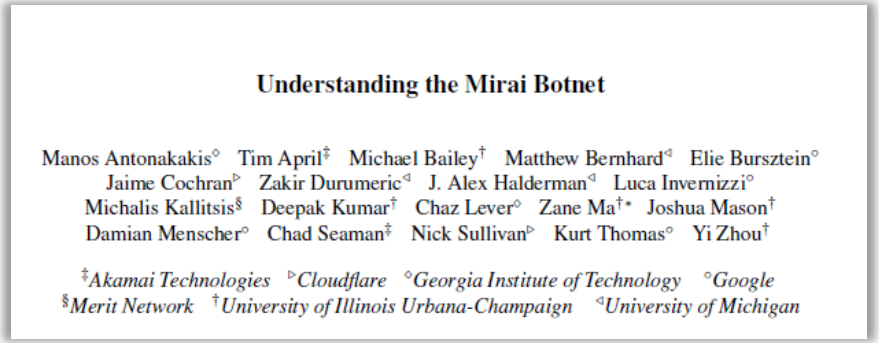
- Evidence of pre-attack surveilence and build-up activities
- Predictive model development has been demonstrated in research (early warning)
- Operationalize as a ‘warning model’ possible (i.e. real-time at data lake OR point of capture)

Subsequently, located following applicable research:


Characterizing Honeypot-Captured Cyber Attacks:
Statistical Framework and Case Study

Zhenxin Zhan, Maochao Xu, and Shouhuai Xu

Attribution



Understanding the Mirai Botnet

Manos Antonakakis[◇] Tim April[‡] Michael Bailey[†] Matthew Bernhard[◁] Elie Bursztein[○]
Jaime Cochran[▷] Zakir Durumeric[◁] J. Alex Halderman[◁] Luca Invernizzi[○]
Michalis Kallitsis[§] Deepak Kumar[†] Chaz Lever[◇] Zane Ma[†⋆] Joshua Mason[†]
Damian Menscher[○] Chad Seaman[‡] Nick Sullivan[▷] Kurt Thomas[○] Yi Zhou[†]

[‡]*Akamai Technologies* [▷]*Cloudflare* [◇]*Georgia Institute of Technology* [○]*Google*
[§]*Merit Network* [†]*University of Illinois Urbana-Champaign* [◁]*University of Michigan*

Mirai type

- IoT-driven worm / bot net – DDoS attacks
- Scans for telnet require telnet 23 open
- C&C communication with attack and replication modules
- DNS lookups to C&C infrastructure
- See: Anotonakakis et al. 2017 – USENIX Security 2017

DDoS reports of Mirai & Satori (Okiru) strains emerging in Q1 2018

Reaper (IoTroop) another IoT-based DDoS (similar to Mirai) emerging early 2018

Memcached DDoS – high activity Q1 2018 (since has been actively patched)

Cowrie honeypot analysis - example analysis to focus attribution

Mirai Mechanism

Scanning for IoT devices

- IP exclusion table embedded
- Brute force usernames/passwords trials via telnet using factory defaults
- Infected devices still operational, but slow & increase bandwidth

Follow-on infection spreading by IoT devices

- TCP SYN probes to pseudo random IPv4 addresses on telnet TCP 23 & 2323*
- Successful login details sent to command-and-control (CaC) collection server

Ongoing behavior of infected devices

- Monitoring CaC server

Leads to massive range of commandeered IPs

- Able to overcome traditional anti-DoS defenses
- DDoS attack action in implementation

** DT Honeypot attacks focused on Ports 5900 (VNC), 69 (trivial FTP), 7007 (UPD – WMP, Skype, Torrent)*

Honeypot Analytics: Mass Internet Attack Early Warning

POC for Major Telecommunications Provider

Mirai (IoT botnet malware) source code released 2016/17 (GitHub)

- 2018: at least 13 variants reported running

Mirai: Anotonakakis et al. 2017 – USENIX Security 2017

DDoS reports of Mirai & Satori (Okiru) strains emerging

Reaper (IoTroop) another IoT-based DDoS (similar to Mirai)

Memcached DDoS

Cowrie honeypot analysis - example analysis to focus attribution

Understanding the Mirai Botnet

Manos Antonakakis[◇] Tim April[‡] Michael Bailey[†] Matthew Bernhard[◁] Elie Bursztein[○]
Jaime Cochran[▷] Zakir Durumeric[◁] J. Alex Halderman[◁] Luca Invernizzi[○]
Michalis Kallitsis[§] Deepak Kumar[†] Chaz Lever[◇] Zane Ma[†*] Joshua Mason[†]
Damian Menscher[○] Chad Seaman[‡] Nick Sullivan[▷] Kurt Thomas[○] Yi Zhou[†]

[‡]*Akamai Technologies* [▷]*Cloudflare* [◇]*Georgia Institute of Technology* [○]*Google*
[§]*Merit Network* [†]*University of Illinois Urbana-Champaign* [◁]*University of Michigan*

REFERENCES: Honeypot Analytics

Emmanouil Vasilomanolakis, Shankar Karuppayah, Panayotis Kikiras, and Max Mühlhäuser. 2015. **A Honeypot-Driven Cyber Incident Monitor: Lessons Learned and Steps Ahead**. Proceedings of the 8th International Conference on Security of Information and Networks (SIN '15). ACM, New York, NY, USA, 158-164. DOI: http://dx.doi.org/10.1145/2799979.2799999

James Forshaw. 2018. **Attacking Network Protocols: A Hacker's Guide to Capture, Analysis, and Exploitation**. William Pollock. San Francisco. https://books.google.nl/books?id=kLgrDwAAQBAJ

Manos Antonakakis and Tim April and Michael Bailey and Matt Bernhard and Elie Bursztein and Jaime Cochran and Zakir Durumeric and J. Alex Halderman and Luca Invernizzi and Michalis Kallitsis and Deepak Kumar and Chaz Lever and Zane Ma and Joshua Mason and Damian Menscher and Chad Seaman and Nick Sullivan and Kurt Thomas and Yi Zhou. 2017. **Understanding the Mirai Botnet**. Proceedings of the 26th USENIX Security Symposium. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis

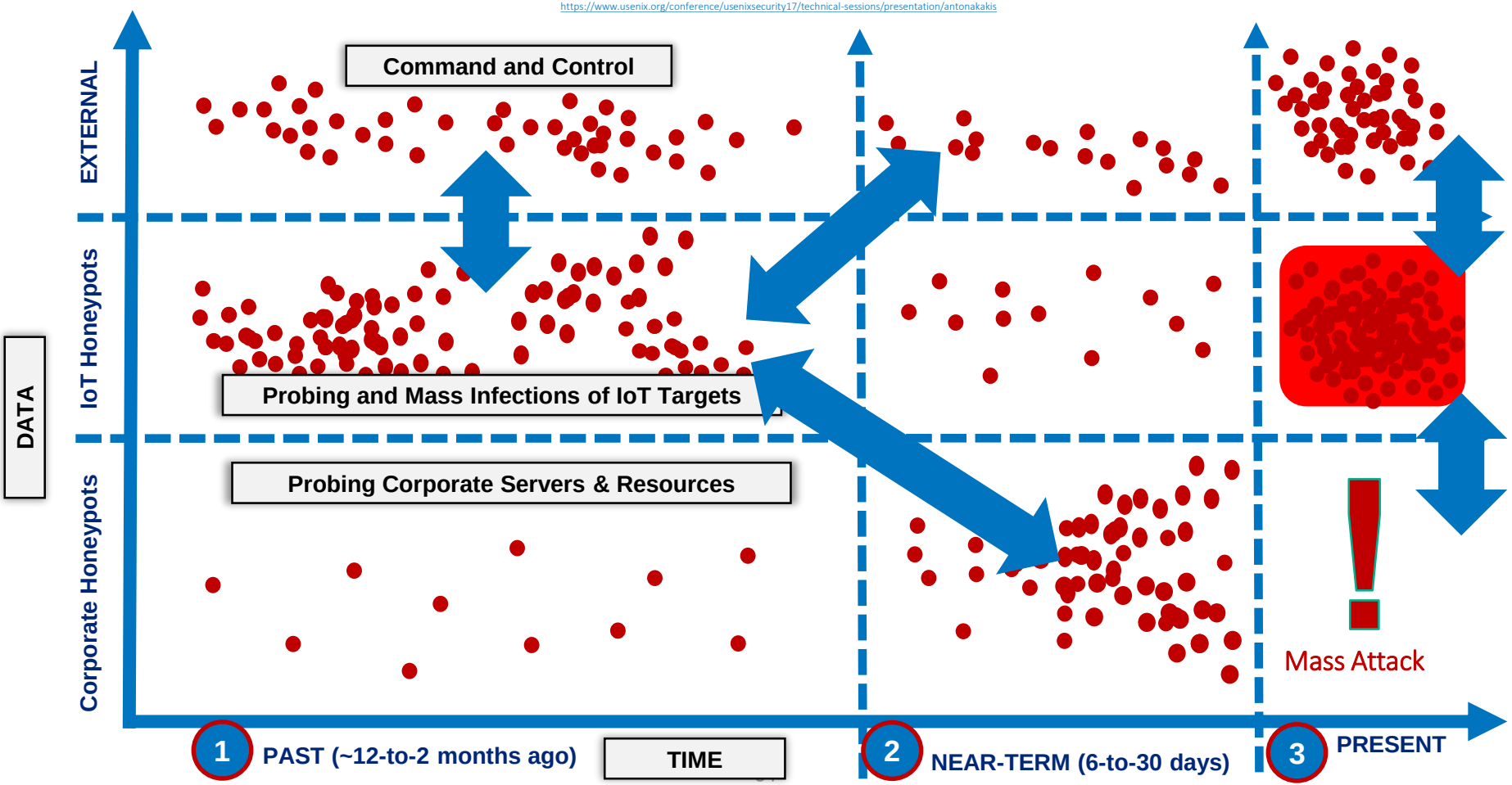
Mitsuaki Akiyama, Takeshi Yagi, Takeshi Yada, Tatsuya Mori, Youki Kadobayashi. 2017. **Analyzing the ecosystem of malicious URL redirection through longitudinal observation from honeypots**. Computers & Security, Volume 69, Pages 155-173. http://www.sciencedirect.com/science/article/pii/S016740481730007X

Niels Provos, Thorsten Holz. 2008. **Virtual Honeypots: From Botnet Tracking to Intrusion Detection**. Addison-Wesley. https://books.google.nl/books/about/Virtual_Honeypots.html?id=QuHnPgAACAAJ&redir_esc=y

Z. Zhan, M. Xu and S. Xu. 2013. **Characterizing Honeypot-Captured Cyber Attacks: Statistical Framework and Case Study.** IEEE Transactions on Information Forensics and Security, vol. 8, no. 11, pp. 1775-1789, Nov. 2013. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6587320&isnumber=6609092

Analytics Model Profile (Botnet Infections & DDoS Attacks)

Manos Antonakakis et al. 2017. **Understanding the Mirai Botnet**. Proceedings of the 26th USENIX Security Symposium.
https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis



Honeypot Dashboard

This practice reinforces the concepts discussed previously.

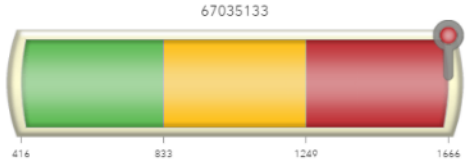
Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting

10/02/2018 tot 26/02/2018

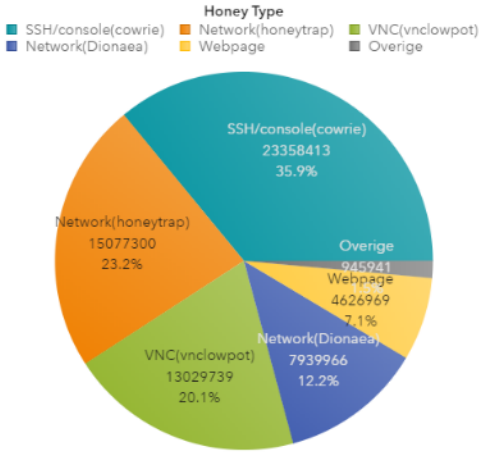
06/12/2016 — 26/02/2018

Num Events
67M

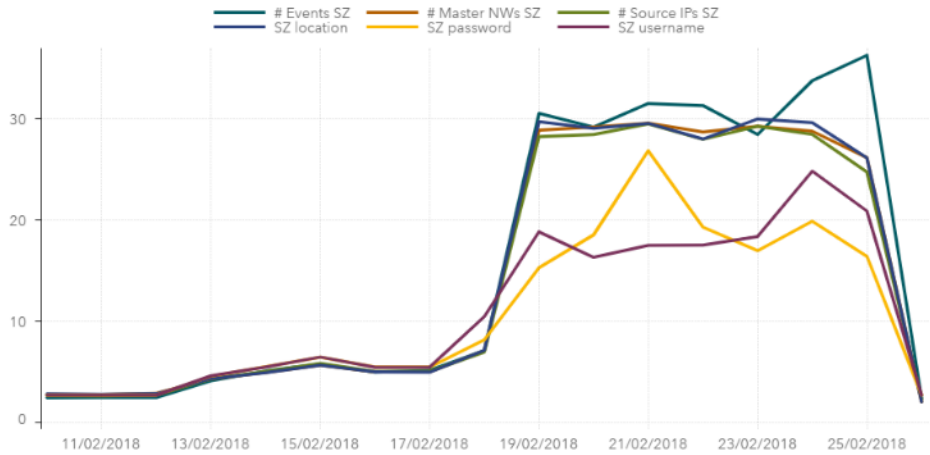
Sum Dist Daily Attack IPs
788K

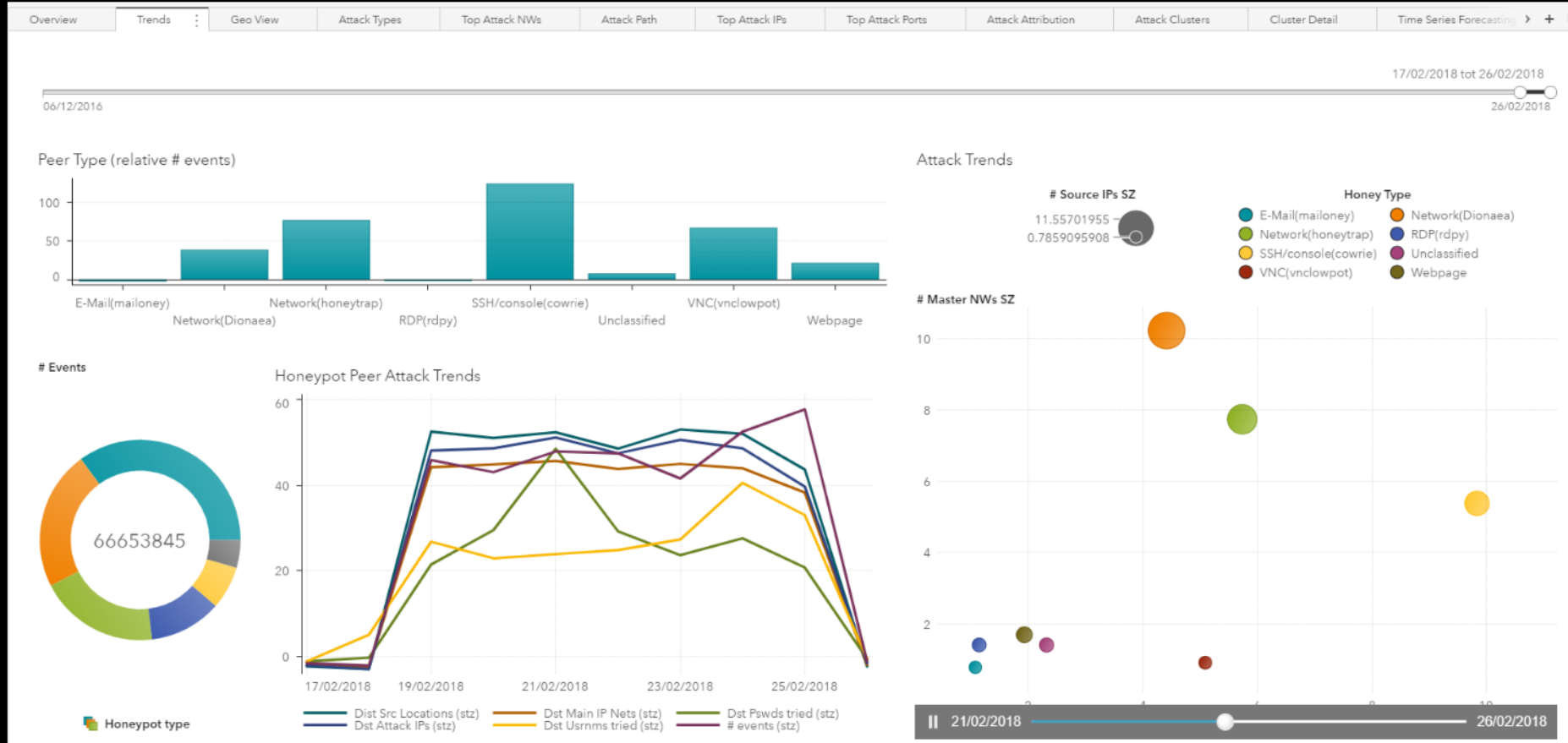


Events per Honeypot Type



Time Trends



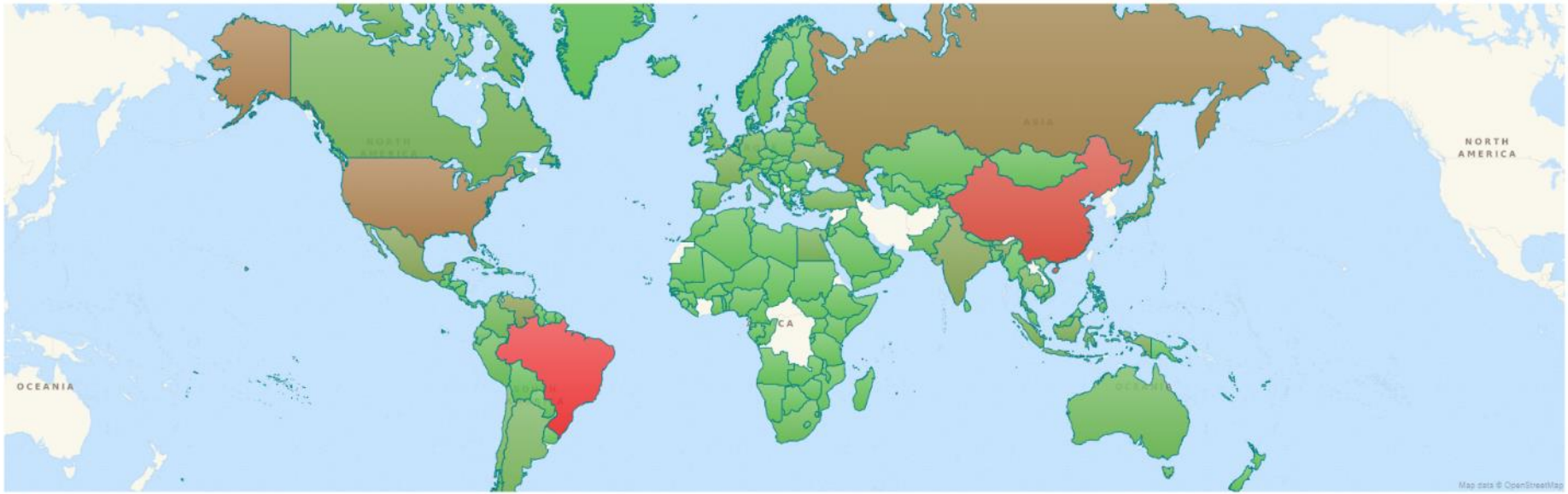

Overview
Trends
Geo View
Attack Types
Top Attack NWs
Attack Path
Top Attack IPs
Top Attack Ports
Attack Attribution
Attack Clusters
Cluster Detail
Time Series Forecasting
17/02/2018 tot 26/02/2018
06/12/2016
26/02/2018
Peer Type (relative # events)
100
50
0
E-Mail(mailoney)
Network(Dionaea)
Network(honeytrap)
RDP(rdpy)
SSH/console(cowrie)
Unclassified
VNC(vnclowpot)
Webpage
Attack Trends
Source IPs SZ
11.55701955
0.7859095908
Honey Type
E-Mail(mailoney)
Network(Dionaea)
Network(honeytrap)
RDP(rdpy)
SSH/console(cowrie)
Unclassified
VNC(vnclowpot)
Webpage
Master NWs SZ
10
8
6
4
2
Events
66653845
Honeypot type
Honeypot Peer Attack Trends
60
40
20
0
17/02/2018
19/02/2018
21/02/2018
23/02/2018
25/02/2018
Dist Src Locations (stz)
Dst Attack IPs (stz)
Dst Main IP Nets (stz)
Dst Usrnms tried (stz)
Dst Pswds tried (stz)
events (stz)
21/02/2018
26/02/2018

Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting

18/02/2018 tot 26/02/2018

06/12/2016 — 26/02/2018

events (stz) op countryName



0.98 — 14224.95

events (stz)

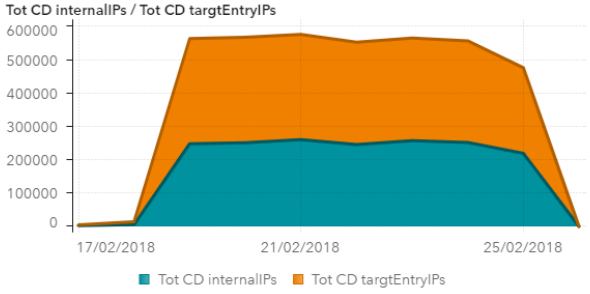
Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting

17/02/2018 tot 26/02/2018

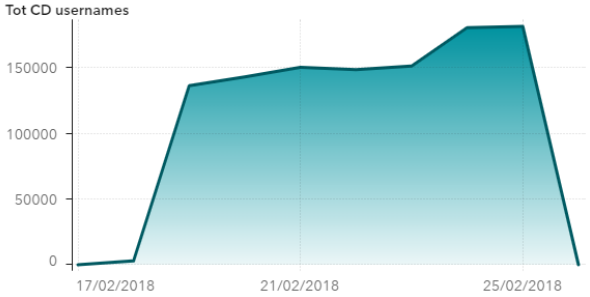
06/12/2016 — 26/02/2018

Mass Attack Events

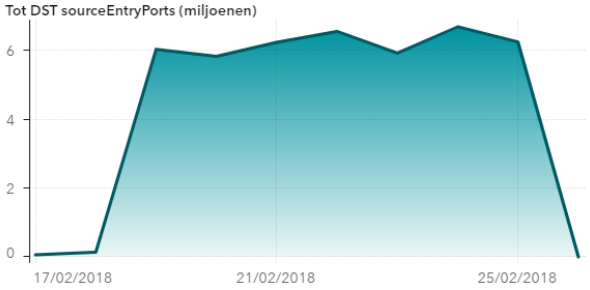
Honey IP Connect Volume



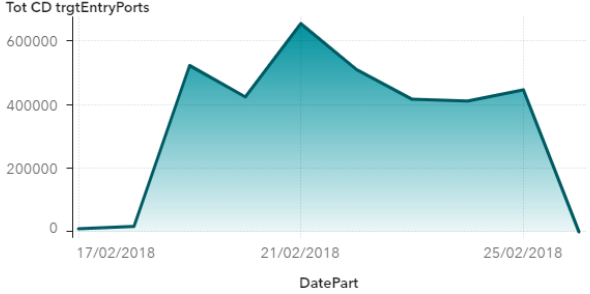
Username Brute Attack Volume



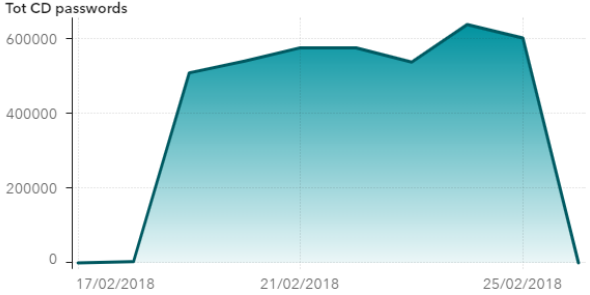
Attack Connection Port Cycling



Honey Port Connect Volume

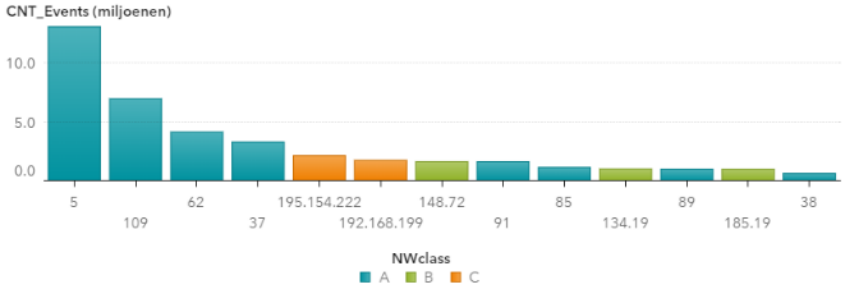


Password Brute Attack Volume



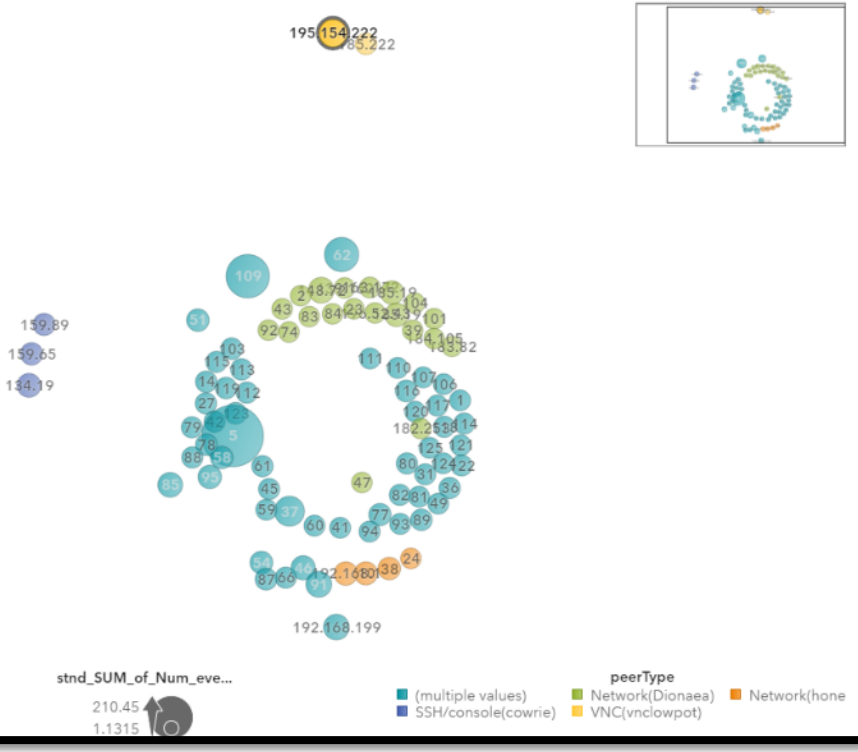
Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting

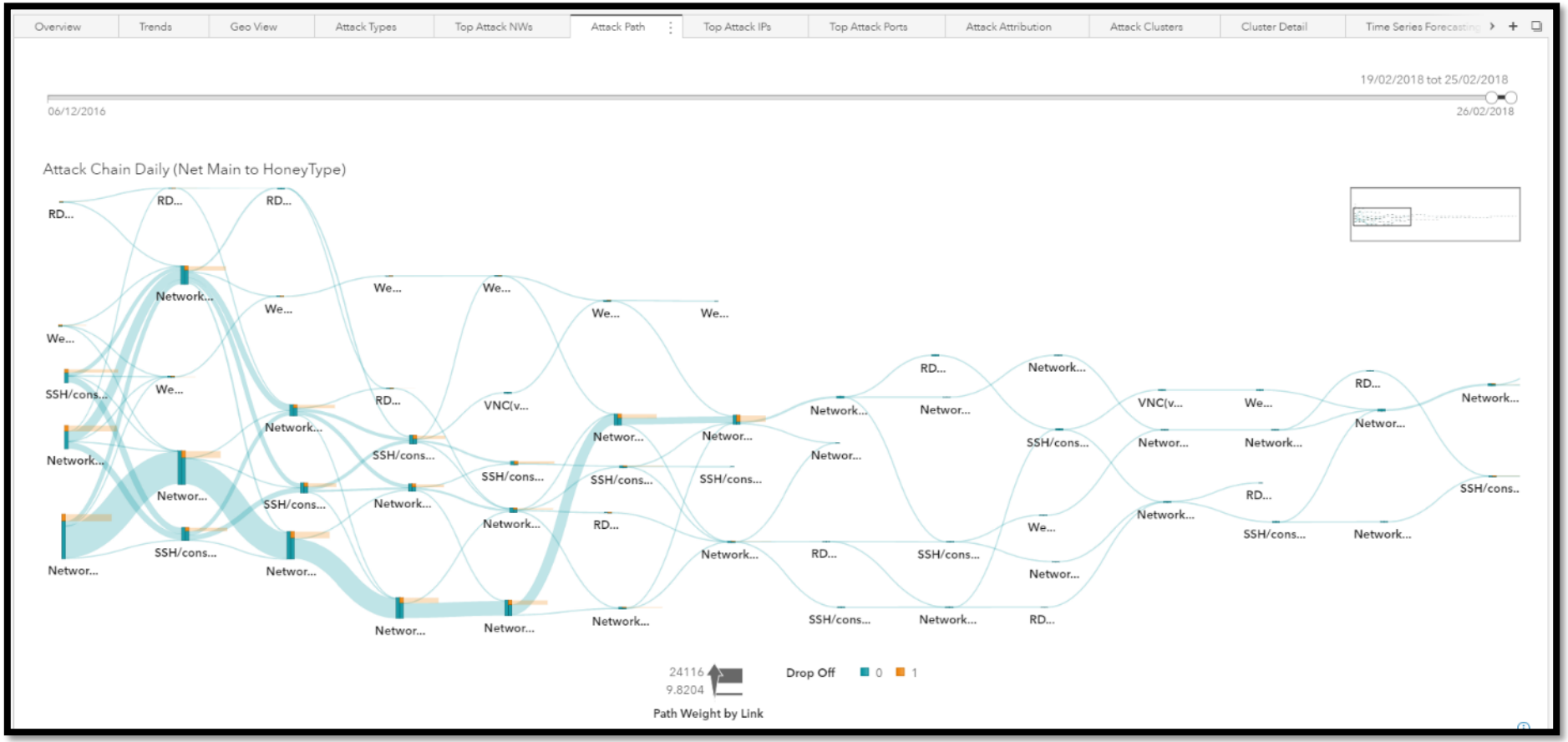
Top Attack Networks



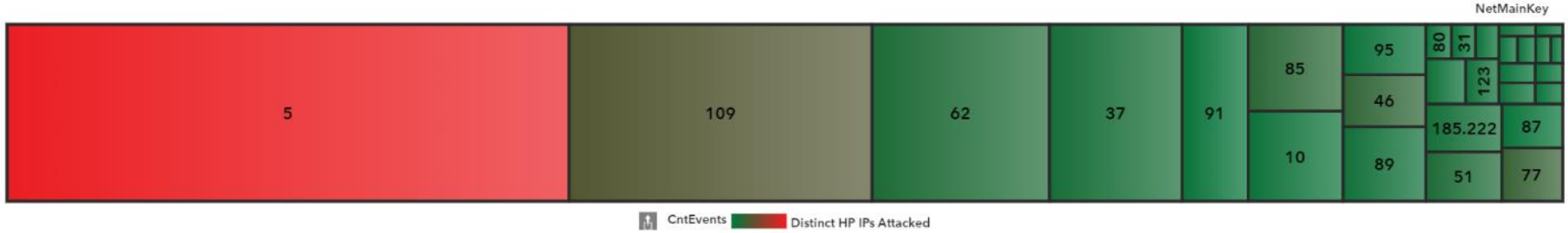
NetMainKey	CNT_Events ▼	Days_active	CD_sourceEntryIp	CD_sourceEntryPort	CD_usernan
5	13035119	325	4851	116383	8
109	6942088	255	3124	130413	3
62	4135112	319	1797	129220	3
37	3293640	356	4506	88982	7
195.154.222	2138943	37	6	117921	
192.168.199	1749473	7	1	56146	2
148.72	1636875	73	17	36202	
91	1632192	319	3379	66662	3
46	1163616	310	5598	77679	4
85	1150988	327	2849	44852	2
134.19	1031830	16	19	56520	
89	1008354	356	2647	114503	2
185.19	1000711	12	17	38140	
192.168.1	958921	78	67	62055	

Source IP Class to Honeypot

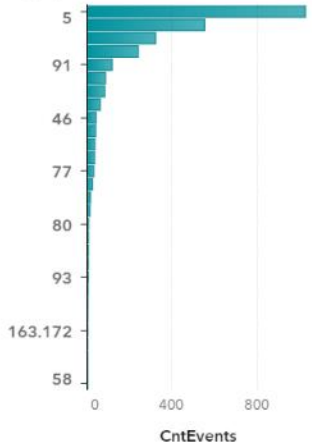



Overview
Trends
Geo View
Attack Types
Top Attack NWs
Attack Path
Top Attack IPs
Top Attack Ports
Attack Attribution
Attack Clusters
Cluster Detail
Time Series Forecasting
19/02/2018 tot 25/02/2018
06/12/2016
26/02/2018
Attack Chain Daily (Net Main to HoneyType)
RD...
We...
SSH/cons...
Network...
Networ...
VNC(v...
24116
9.8204
Path Weight by Link
Drop Off
0
1

Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting



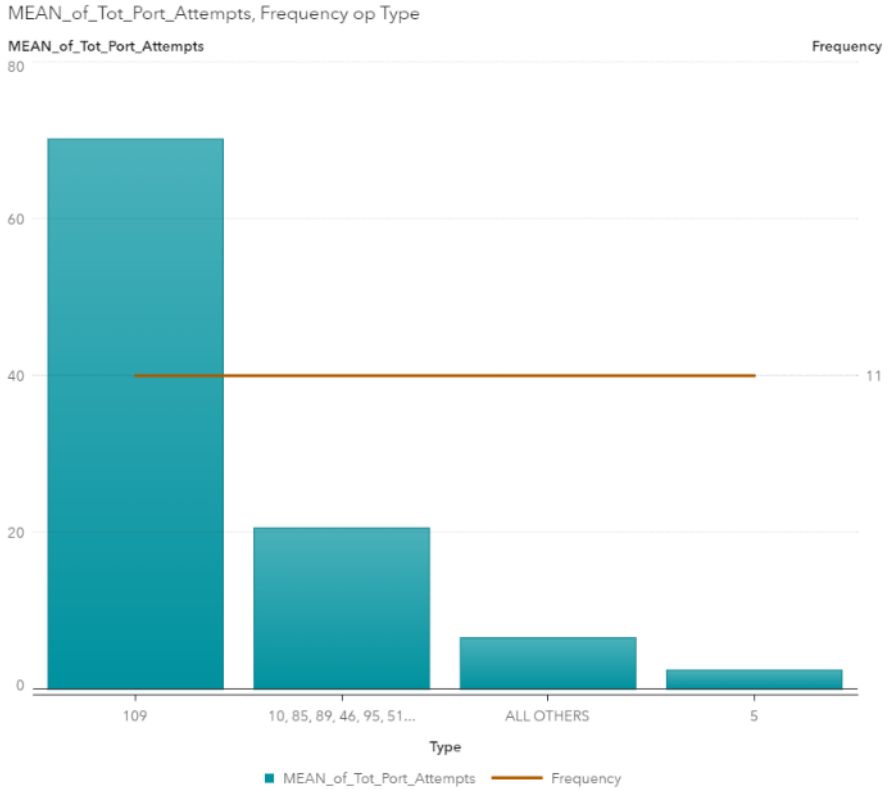
Highly Active Attack NWs

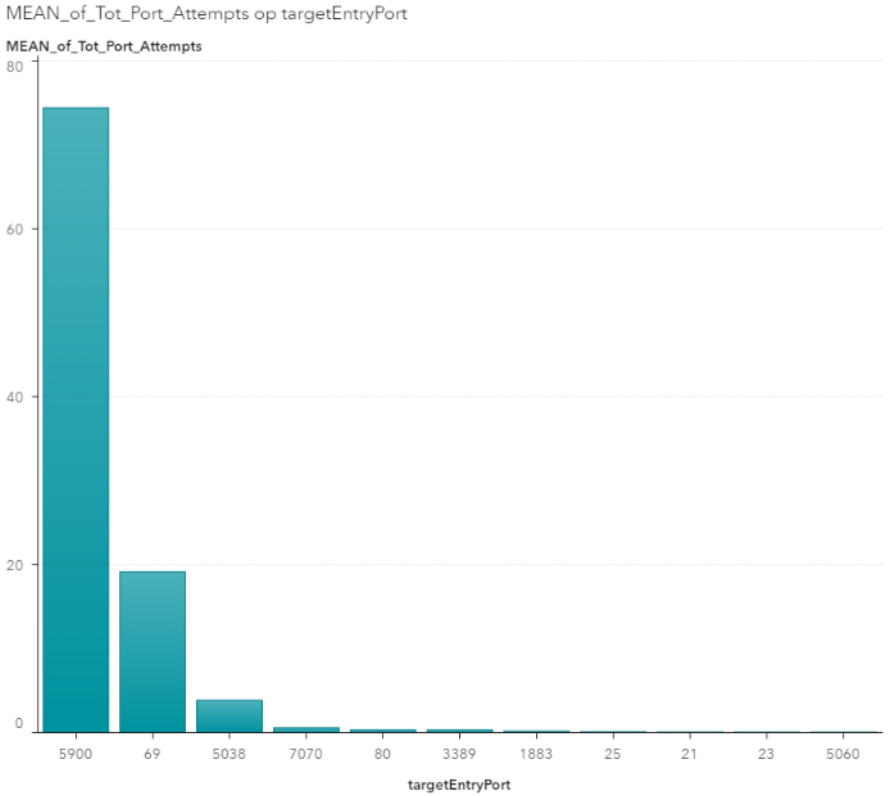


Highly Active Attack IPs

NetMainKey	sourceEntryIp	CntEvents ▼	CD_peerIdent	CD_country	CD_sourceEntryPor	CD_username	CD_password	CD_targetCountry	CD_targetprotocol	Distinct HP IPs Attacked	CD_trgtEntryPort
37	37.57.174.125	219.37727627	14.343484218	-0.054237427	53.031175058	-0.179485248	-0.034027956	12.375922714	-0.352443186	19.959551619	2.8617451491
62	62.210.151.23	204.50029213	1.293509081	-0.054237427	106.17190392	-0.179485248	-0.034027956	14.655354561	-0.352443186	11.575778814	-0.014908392
109	109.248.46.113	171.89590609	1.293509081	-0.054237427	105.99808542	-0.179485248	-0.034027956	12.375922714	-0.352443186	11.389472751	-0.014908392
109	109.248.46.99	160.0590536	1.293509081	-0.054237427	106.00438972	-0.179485248	-0.034027956	12.375922714	-0.352443186	11.203166689	-0.014908392
91	91.210.104.71	114.94111837	0.1420406866	-0.054237427	-0.033001617	-0.179485248	-0.034027956	0.9787634821	2.766736278	0.0248029496	-0.011408813
62	62.210.146.171	93.914241889	1.293509081	-0.054237427	105.94855165	-0.179485248	-0.034027956	14.655354561	-0.352443186	11.575778814	-0.014908392
5	5.188.86.174	65.717424462	13.959661419	-0.054237427	50.80305603	5.0353493242	0.2631700321	13.515638637	-0.352443186	9.5264121282	-0.014908392
85	85.190.153.97	49.809123254	7.8184966493	-0.054237427	0.0300413628	-0.179485248	-0.034027956	20.353934177	5.8859157423	10.830554565	98.711701137
89	89.248.174.161	49.081634349	0.5258634848	-0.054237427	84.651732475	-0.179485248	-0.034027956	12.375922714	-0.352443186	6.7318211933	-0.011408813
5	5.188.86.164	46.355422022	15.111129814	-0.054237427	50.755323488	5.2439427071	0.2546786611	20.353934177	-0.352443186	13.252533375	-0.014908392
5	5.188.87.49	45.484134015	15.111129814	-0.054237427	50.739112436	5.2439427071	0.2546786611	20.353934177	-0.352443186	13.252533375	-0.014908392
5	5.188.87.52	44.083396215	15.111129814	-0.054237427	50.756224102	5.0353493242	0.24618729	22.633366023	-0.352443186	13.252533375	-0.014908392
5	5.188.86.209	43.459173565	15.111129814	-0.054237427	50.774236382	5.0353493242	0.2546786611	19.214218254	-0.352443186	12.507309125	-0.014908392
5	5.188.87.50	43.159802335	15.111129814	-0.054237427	50.74451612	5.2439427071	0.2631700321	18.07450233	-0.352443186	13.066227312	-0.014908392
109	109.248.46.55	43.01120993	15.111129814	-0.054237427	100.92042371	-0.179485248	-0.034027956	4.3979112517	-0.352443186	12.507309125	-0.014908392
5	5.188.86.194	42.545502664	15.111129814	-0.054237427	50.746317348	5.2439427071	0.2631700321	19.214218254	-0.352443186	13.438839437	-0.014908392

Overview | Trends | Geo View | Attack Types | Top Attack NWs | Attack Path | Top Attack IPs | Top Attack Ports | Attack Attribution | Attack Clusters | Cluster Detail | Time Series Forecasting

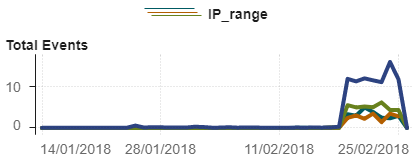




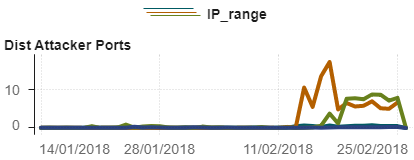
14/01/2018 tot 26/02/2018

06/12/2016 — 26/02/2018

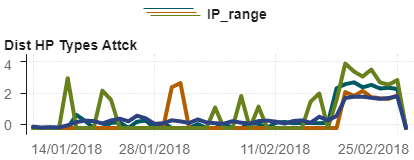
Total Events by IP Range



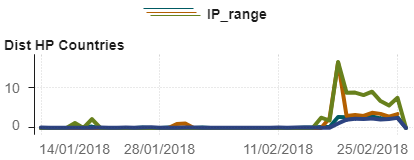
Distinct Attacker Ports Attempted



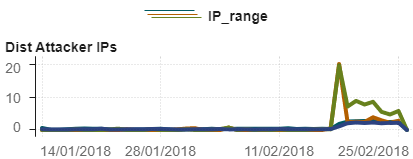
Distinct Honeypot Types Attacked



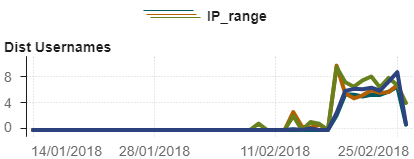
Distinct Honeypot Countries Attacked



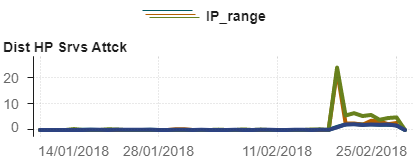
Distinct Attacker IPs



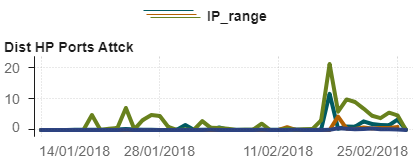
Distinct Attacker Usernames Attempted



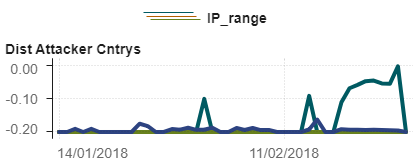
Distinct Honeypot Servers Attacked



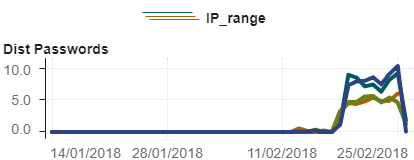
Distinct Honeypot Ports Attacked



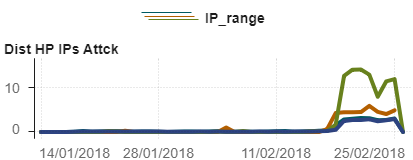
Distinct Attacker Countries



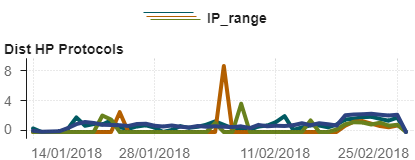
Distinct Attacker Passwords Attempted

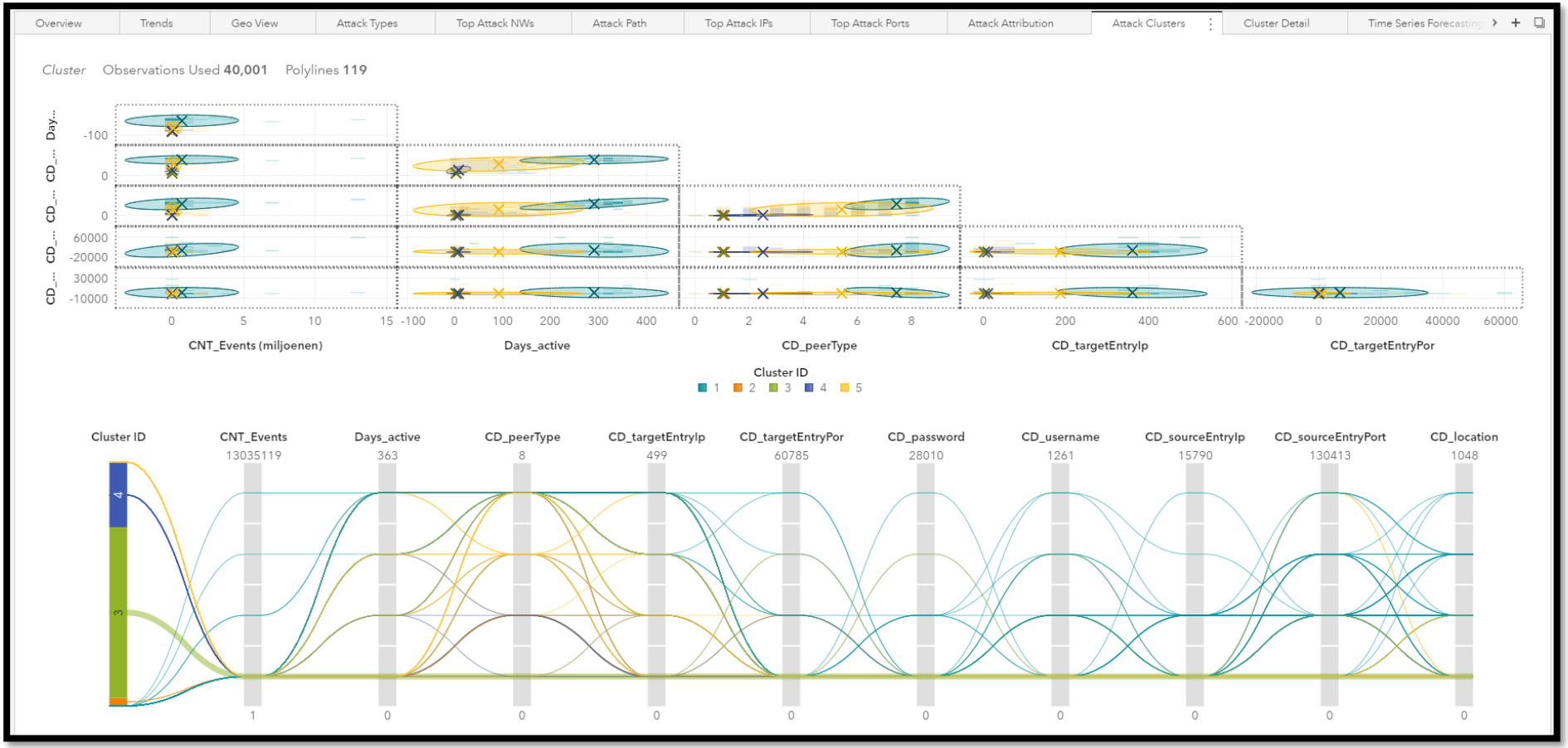


Distinct Honeypot IPs Attacked

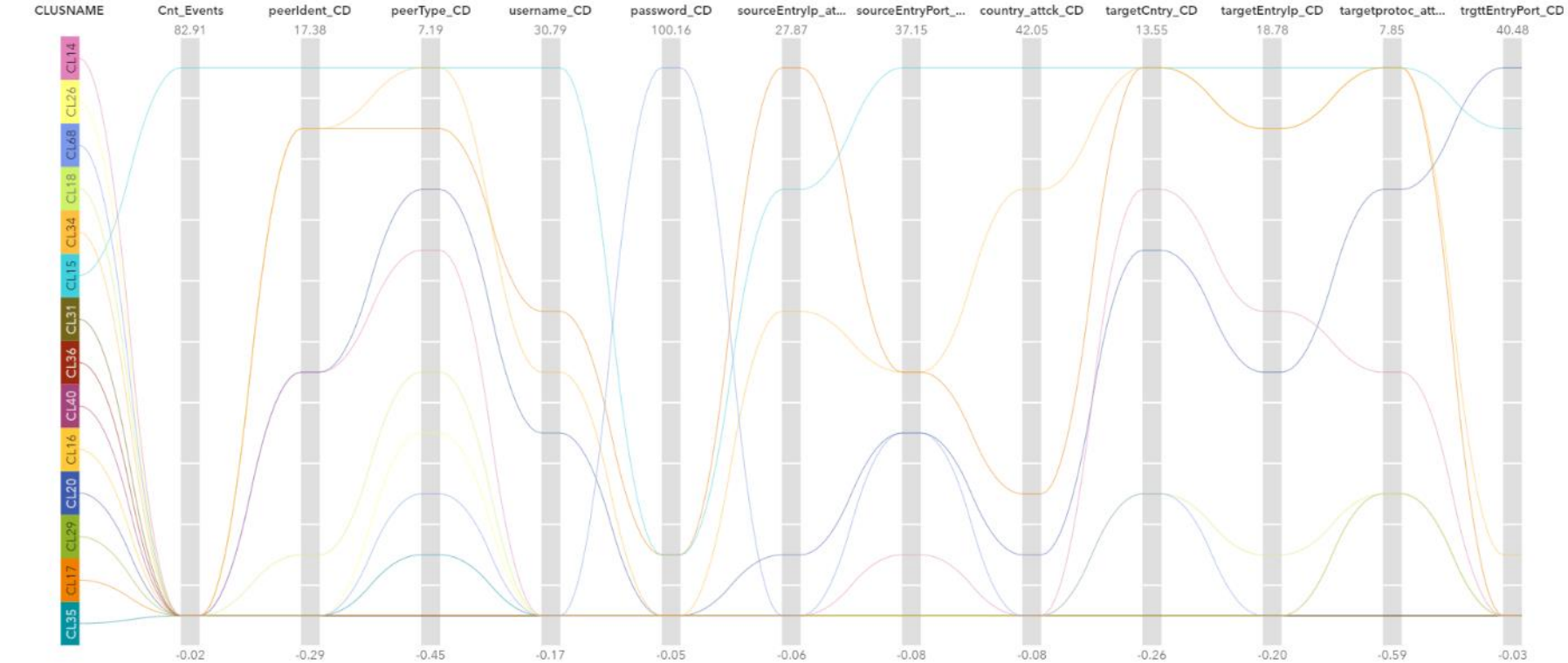


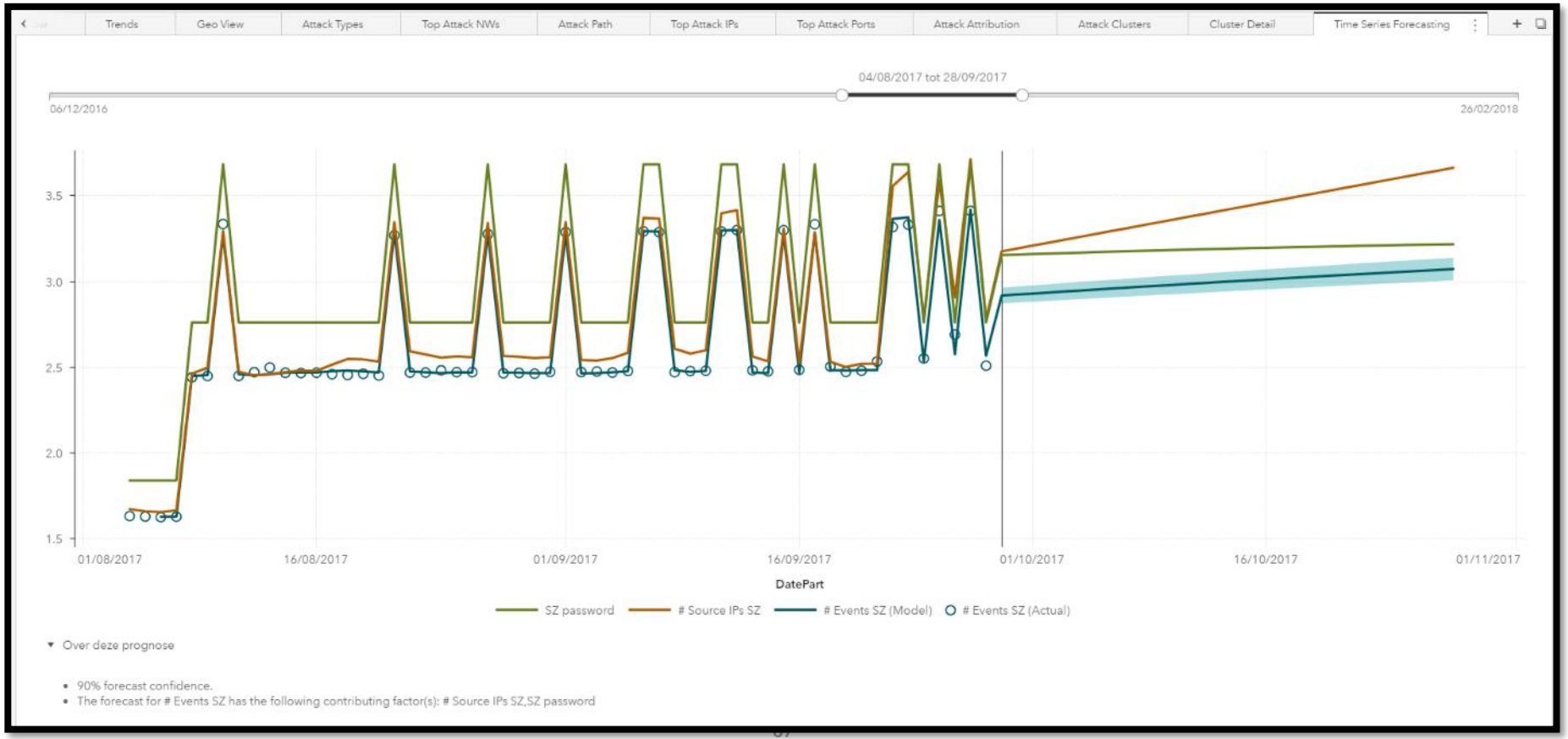
Distinct Honeypot Transport Protocols Utilized




Overview
Trends
Geo View
Attack Types
Top Attack NWs
Attack Path
Top Attack IPs
Top Attack Ports
Attack Attribution
Attack Clusters
Cluster Detail
Time Series Forecasting
Cluster
Observations Used 40,001
Polylines 119
CNT_Events (miljoenen)
Days_active
CD_peerType
CD_targetEntryIp
CD_targetEntryPor
Cluster ID
1 2 3 4 5
Cluster ID
CNT_Events
13035119
Days_active
363
CD_peerType
8
CD_targetEntryIp
499
CD_targetEntryPor
60785
CD_password
28010
CD_username
1261
CD_sourceEntryIp
15790
CD_sourceEntryPort
130413
CD_location
1048

Parallelle coördinaten van geselecteerde variabelen



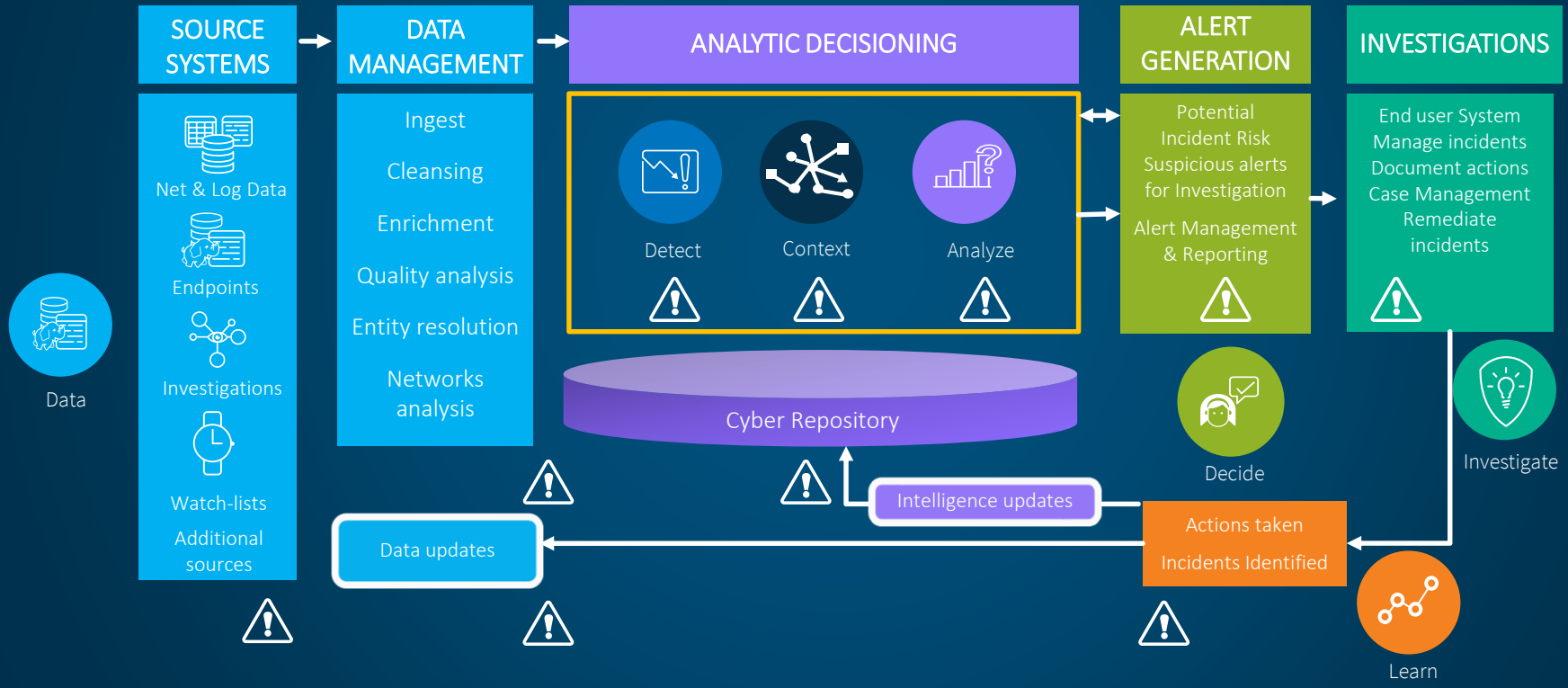

Trends
Geo View
Attack Types
Top Attack NWs
Attack Path
Top Attack IPs
Top Attack Ports
Attack Attribution
Attack Clusters
Cluster Detail
Time Series Forecasting
04/08/2017 tot 28/09/2017
06/12/2016
26/02/2018
3.5
3.0
2.5
2.0
1.5
01/08/2017
16/08/2017
01/09/2017
16/09/2017
01/10/2017
16/10/2017
01/11/2017
DatePart
SZ password
Source IPs SZ
Events SZ (Model)
Events SZ (Actual)
Over deze prognose
90% forecast confidence.
The forecast for # Events SZ has the following contributing factor(s): # Source IPs SZ,SZ password

Attack Attribution: guidance for feature selection

Wrap-Up

Section Review

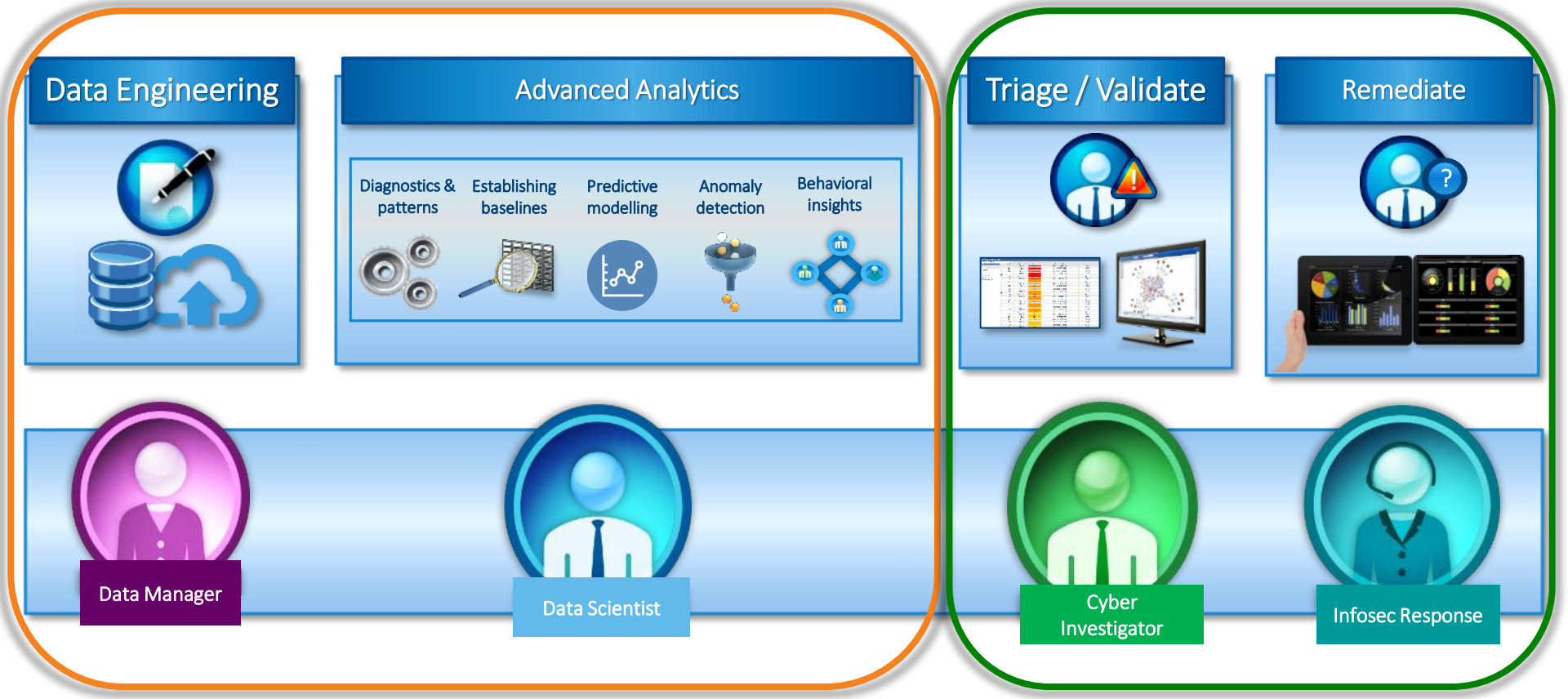
Cyber Analytics Functional Architecture



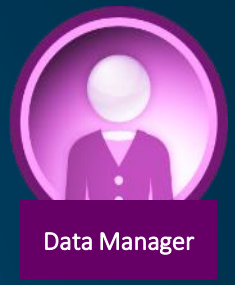
Course Wrap-Up

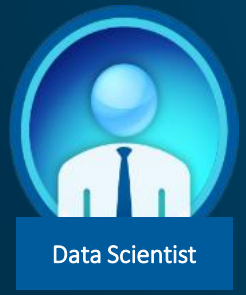
Course Summary

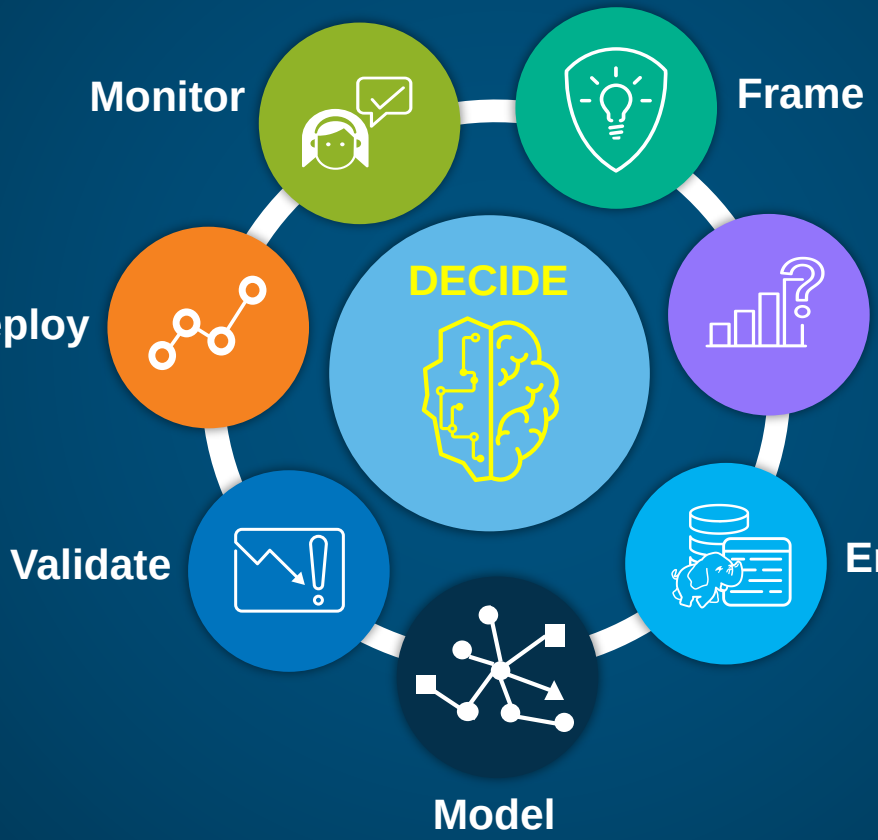
Cybersecurity Data Science as a Process



Cybersecurity Data Science (CSDS) Lifecycle







Cybersecurity Analytics Maturity

Anomaly Detection	Data-aware Investigations	Predictive Detection	Risk Awareness / Resource Optimization
• Big data management • Flags, rules, and alerts • Multivariate statistics, inference & unsupervised machine learning • Segments extracted as baselines	**Understanding** • Feature engineering • Labeling • Diagnostics • *Unsupervised ML*	**Learning** • Human-in-the-loop reviews • *Combined supervised and unsupervised machine learning*	





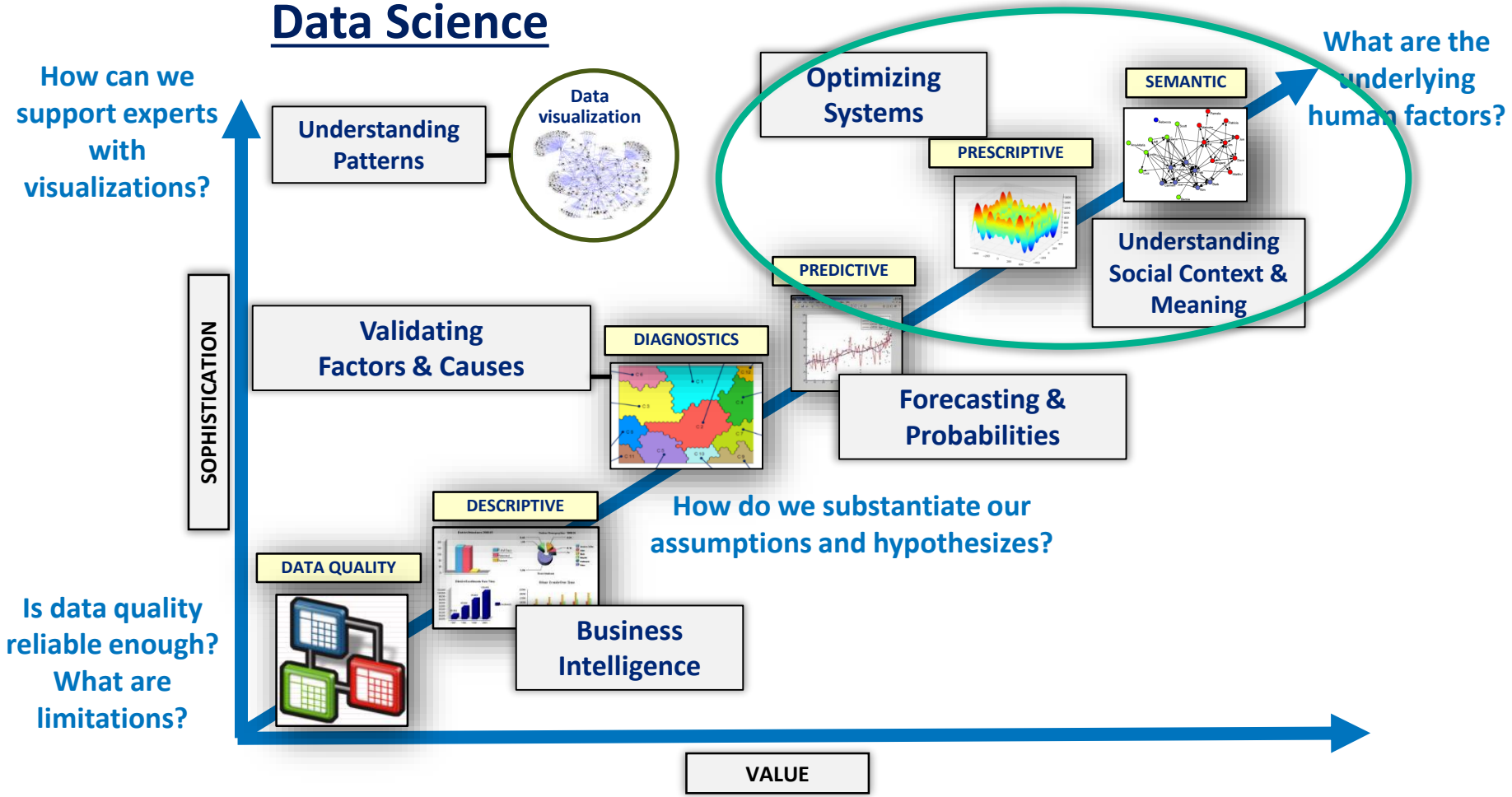


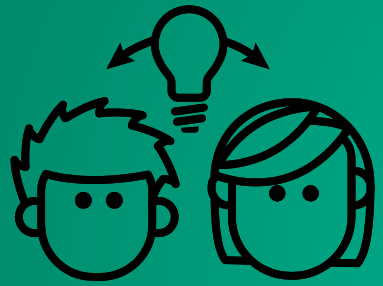


Cybersecurity Analytics Maturity

Anomaly Detection	Data-aware Investigations	Predictive Detection	Risk Awareness / Resource Optimization
• Big data management • Flags, rules, and alerts • Multivariate statistics, inference & unsupervised machine learning • Segments extracted as baselines	**Understanding** • Feature engineering • Labeling • Diagnostics • *Unsupervised ML*	**Learning** • Human-in-the-loop reviews • *Combined supervised and unsupervised machine learning*	**Risk Optimal** • Champion-challenger model management • Automating alert triage • *Resource optimization*

Data Science





Idea Exchange

Based on today's discussion, what are your thoughts on next steps with cybersecurity analytics?

REFERENCES

REFERENCES

Aggarwal, C. (2013). “Outlier Analysis.” Springer. http://www.springer.com/la/book/9781461463955

Harris, H., Murphy, S., and Vaisman, M. (2013). “Analyzing the Analyzers.” O’Reilly Media. Available at https://www.oreilly.com/data/free/analyzing-the-analyzers.csp

Kirchhoff, C., Upton, D., and Winnefeld, Jr., Admiral J. A. (2015 October 7). “Defending Your Networks: Lessons from the Pentagon.” Harvard Business Review. Available at https://hbr.org/webinar/2015/10/defending-your-networks-lessons-from-the-pentagon

Mongeau, S. (2018). “Cybersecurity Data Science (CSDS).” SCTR7.com. https://sctr7.com/2018/12/03/cybersecurity-data-science-csds-how-not-to-drown-in-your-cyber-data-lake/

Mongeau, S. (2017). “Cybersecurity Big Data Overload?” SCTR7.com. https://sctr7.com/2017/10/22/cybersecurity-big-data-overload/

Ponemon Institute. (2017). “When Seconds Count: How Security Analytics Improves Cybersecurity Defenses.” Available at https://www.sas.com/en_us/whitepapers/ponemon-how-security-analytics-improves-cybersecurity-defenses-108679.html

SANS Institute. (2015). “2015 Analytics and Intelligence Survey.” Available at https://www.sas.com/en_us/whitepapers/sans-analytics-intelligence-survey-108031.html

SANS Institute. (2016). “Using Analytics to Predict Future Attacks and Breaches.” Available at https://www.sas.com/en_us/whitepapers/sans-using-analytics-to-predict-future-attacks-breaches-108130.html

SAS Institute. (2016). “Managing the Analytical Life Cycle for Decisions at Scale.” Available at https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf

SAS Institute. (2017). “SAS Cybersecurity: Counter cyberattacks with your information advantage.” Available at https://www.sas.com/en_us/software/fraud-security-intelligence/cybersecurity-solutions.html

UBM. (2016). “Dark Reading: Close the Detection Deficit with Security Analytics.” Available at https://www.sas.com/en_us/whitepapers/close-detection-deficit-with-security-analytics-108280.html

SAS Security Analytics Framework (SAF)

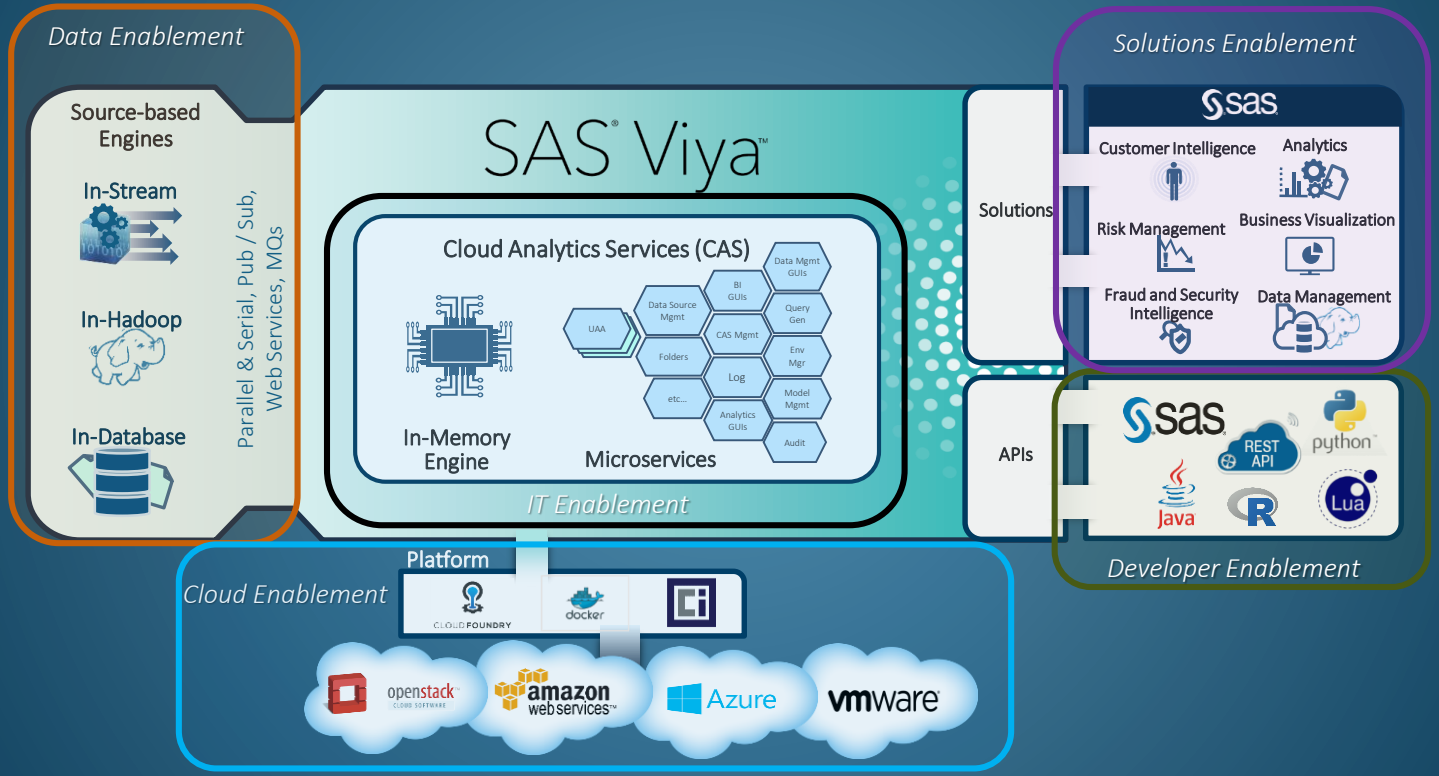
DATA ENGINEERING & ANALYTICS
ESP • High-speed ingestion & enrichment
ACCESS Hadoop, DLfH • Hadoop integration & data management
DI Studio / Data Management • ETL and data quality / governance
Viya VDMML • ML engine (on-premise or in-in cloud)

FRONT-END & OPERATIONAL
VA • Dashboards, self-service analytics & statistics
VI • Investigative support
ALIAS (including MM) • Self-improving ML based on investigative results
Decision Manager • Risk model development / hosting

APPENDIX

Cloud & service architectures

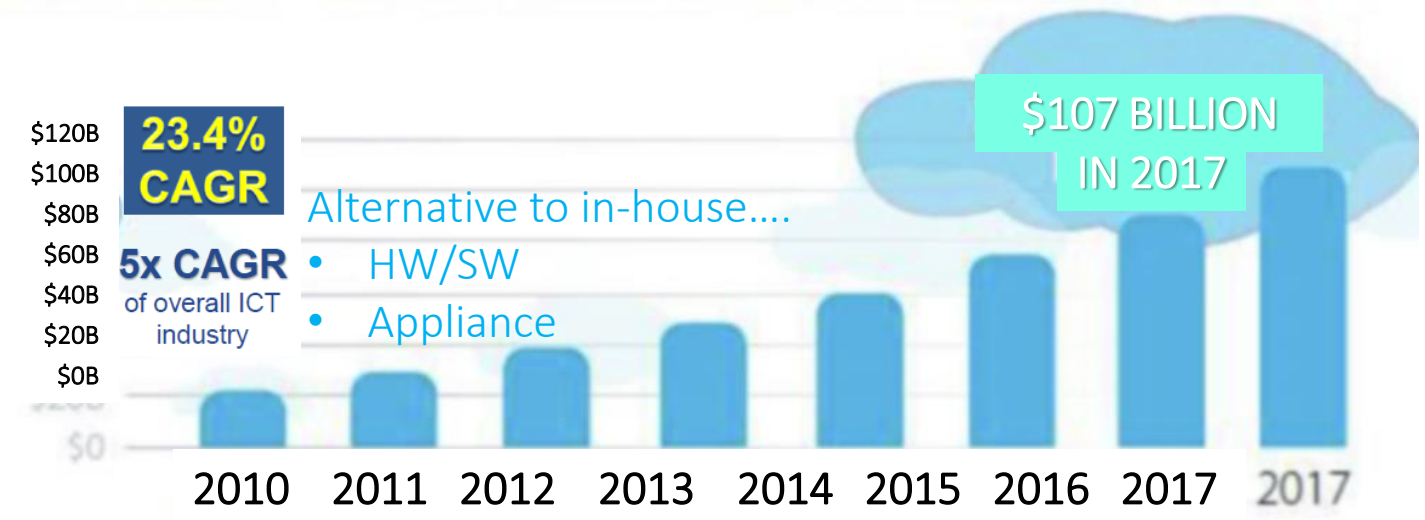
VIYA HYBRID CLOUD & MICROSERVICES ARCHITECTURE



Cloud Computing

WORLDWIDE PUBLIC IT CLOUD SERVICES REVENUE

Worldwide public cloud services market revenue growth hit 18.5% in 2017 to total $260.2 billion, up from $219.6 bil in 2016. Projected to reach $411 bil by 2020.
Gartner, Inc. https://www.gartner.com/newsroom/id/3815165



IDC, "Worldwide Software Predictions, 2015", January 2015

CAGR = Compound Annual Growth Rate

Magic Quadrant for Public Cloud Storage Services, Worldwide

July 2017

Market leaders:

- AWS (Amazon)
- Microsoft

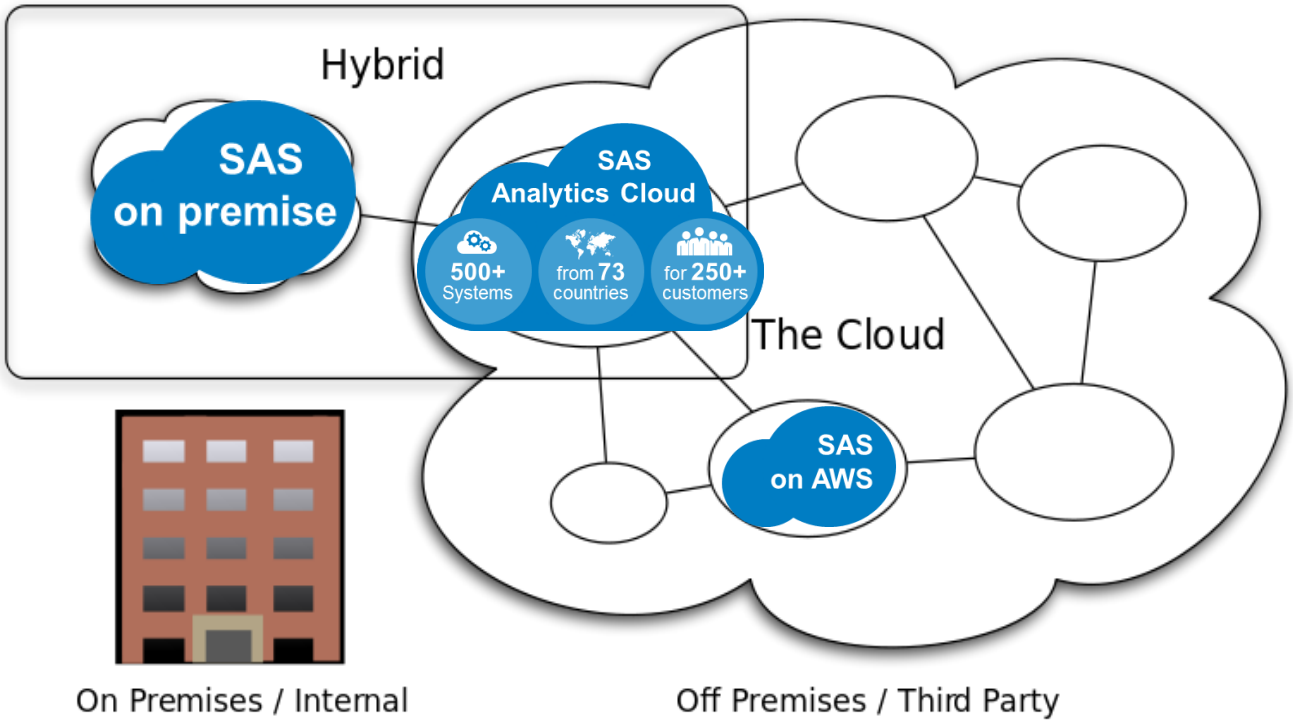
https://www.gartner.com/doc/3770164/magic-quadrant-public-cloud-storage

National Institute of Standards and Technology (NIST)

- Ubiquitous, on-demand network access
- Shared pool of configurable computing resources
- Can be rapidly provisioned with minimal effort

Cloud Models

DEPLOYMENT MODELS

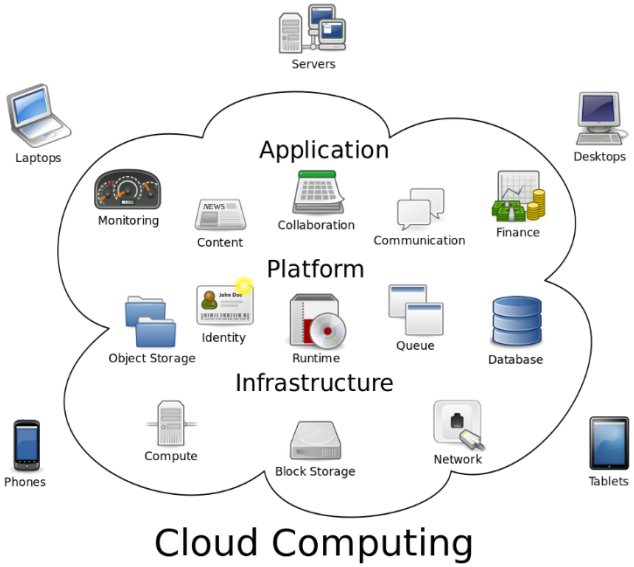


Cloud Computing Types

National Institute of Standards and Technology (NIST)

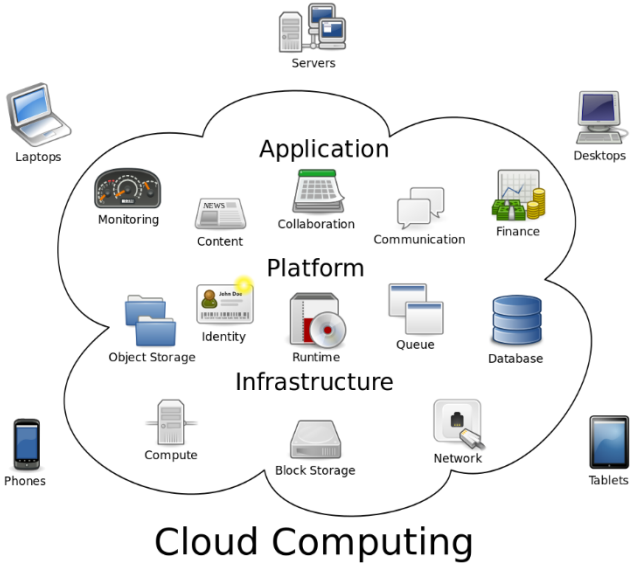
- Ubiquitous, on-demand network access
- Shared pool of configurable computing resources
- Can be rapidly provisioned with minimal effort

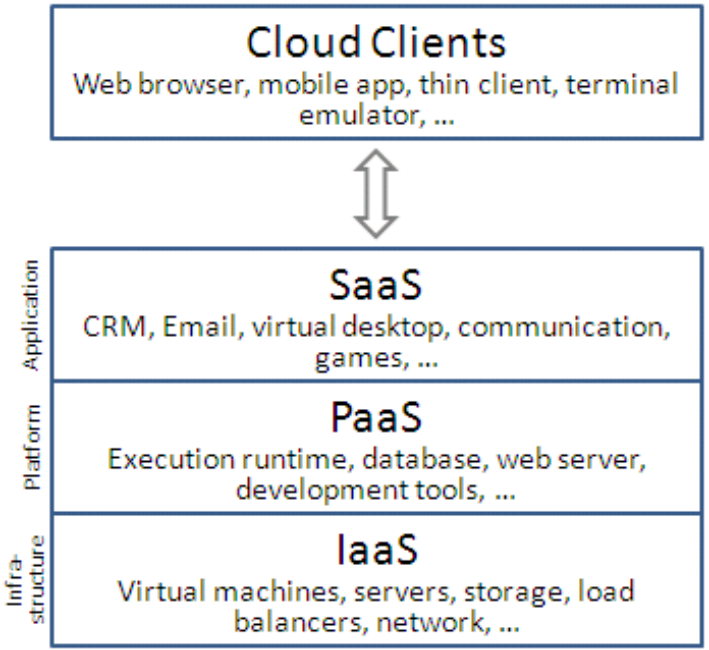
Layers of Cloud Services



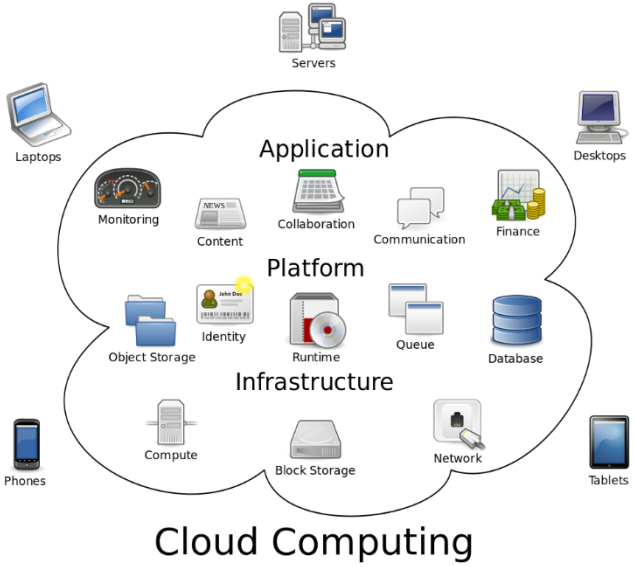
- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

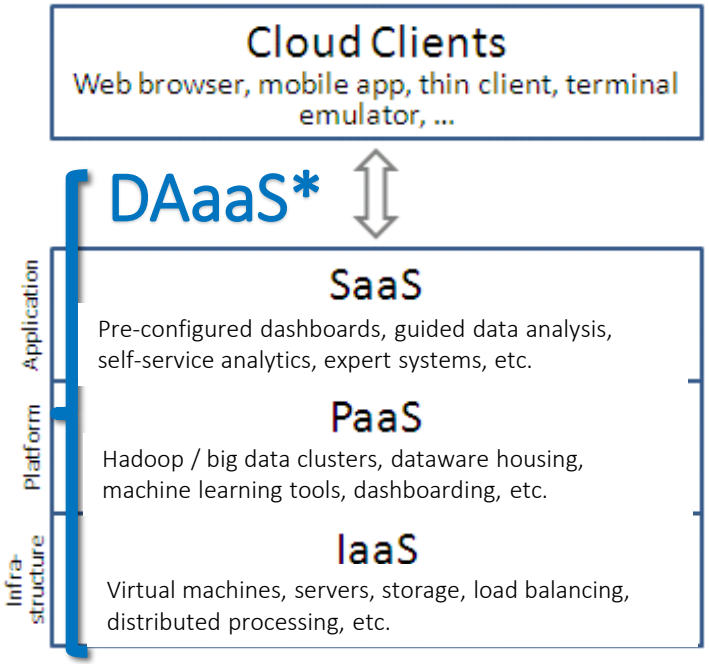
Layers of Cloud Analytics Services





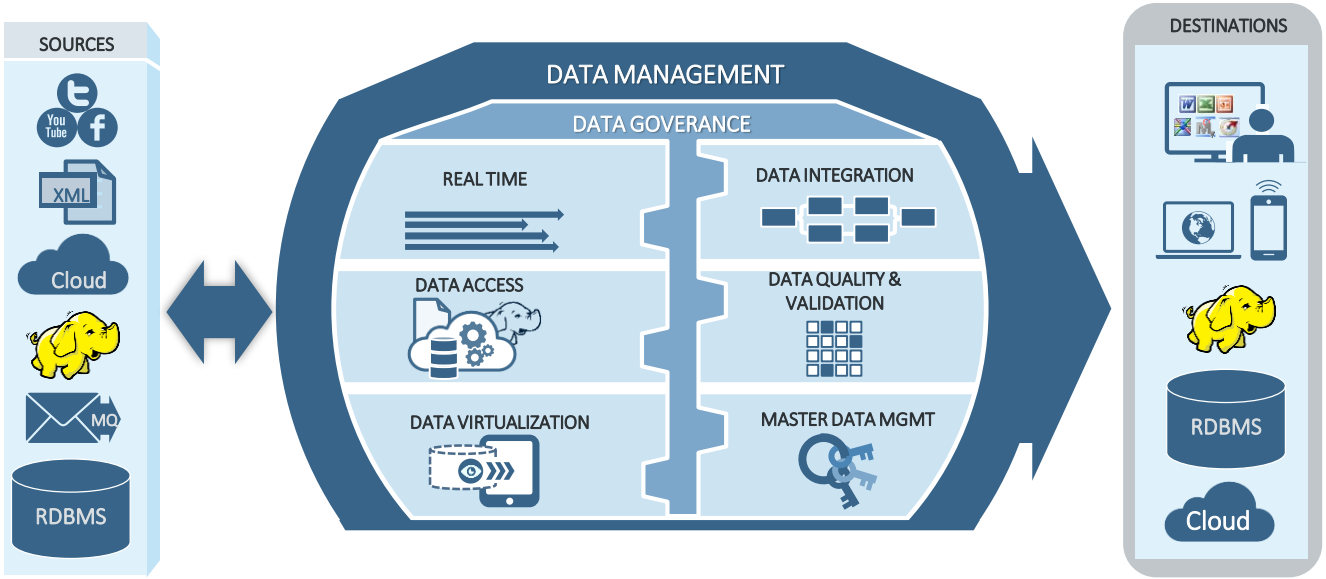
Layers of Cloud Analytics Services



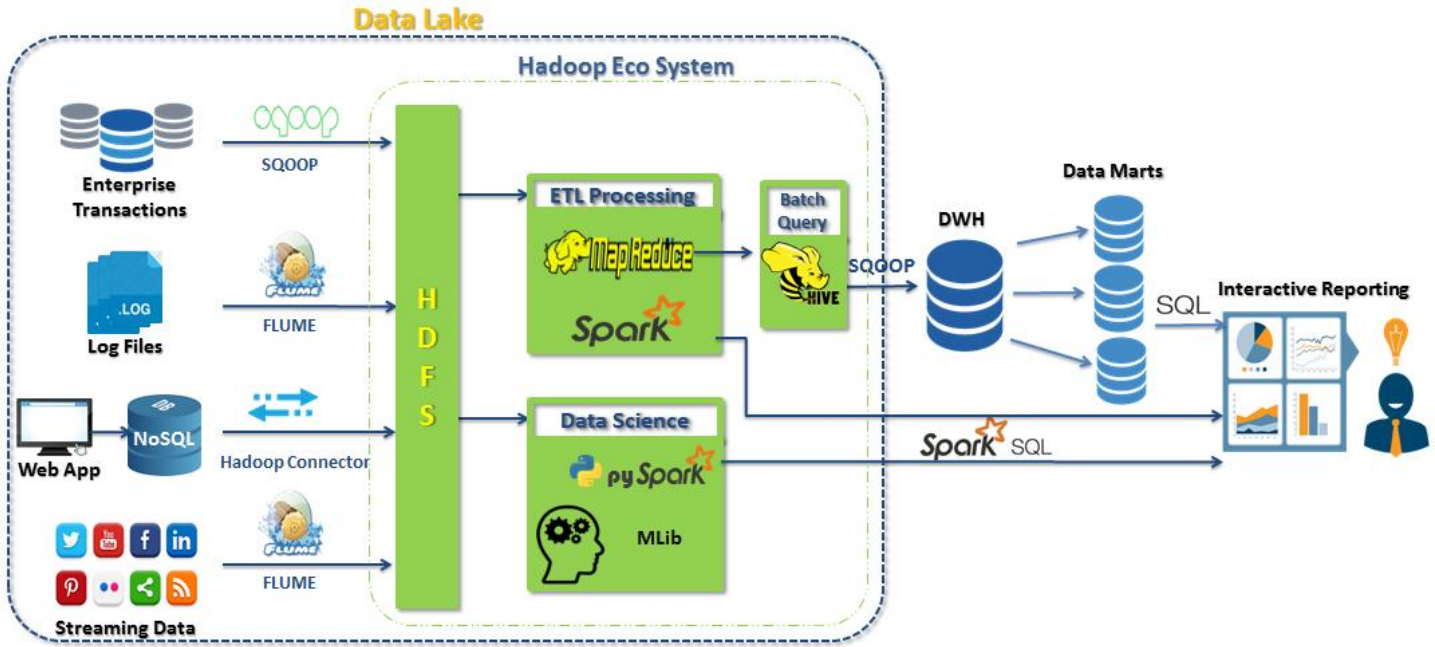


*Data Analytics as a Service

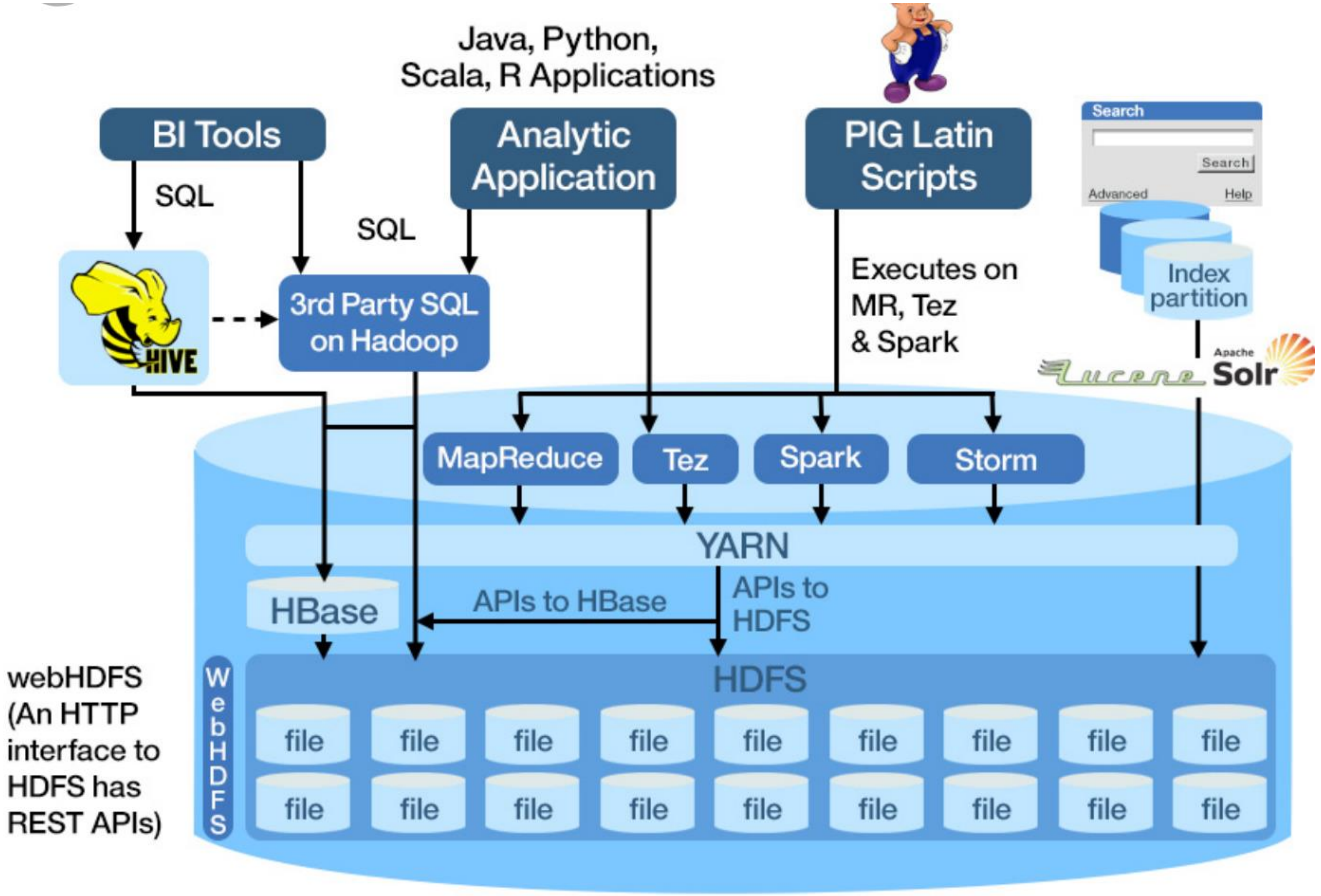
Data Engineering:
Preparing Data for Analytics

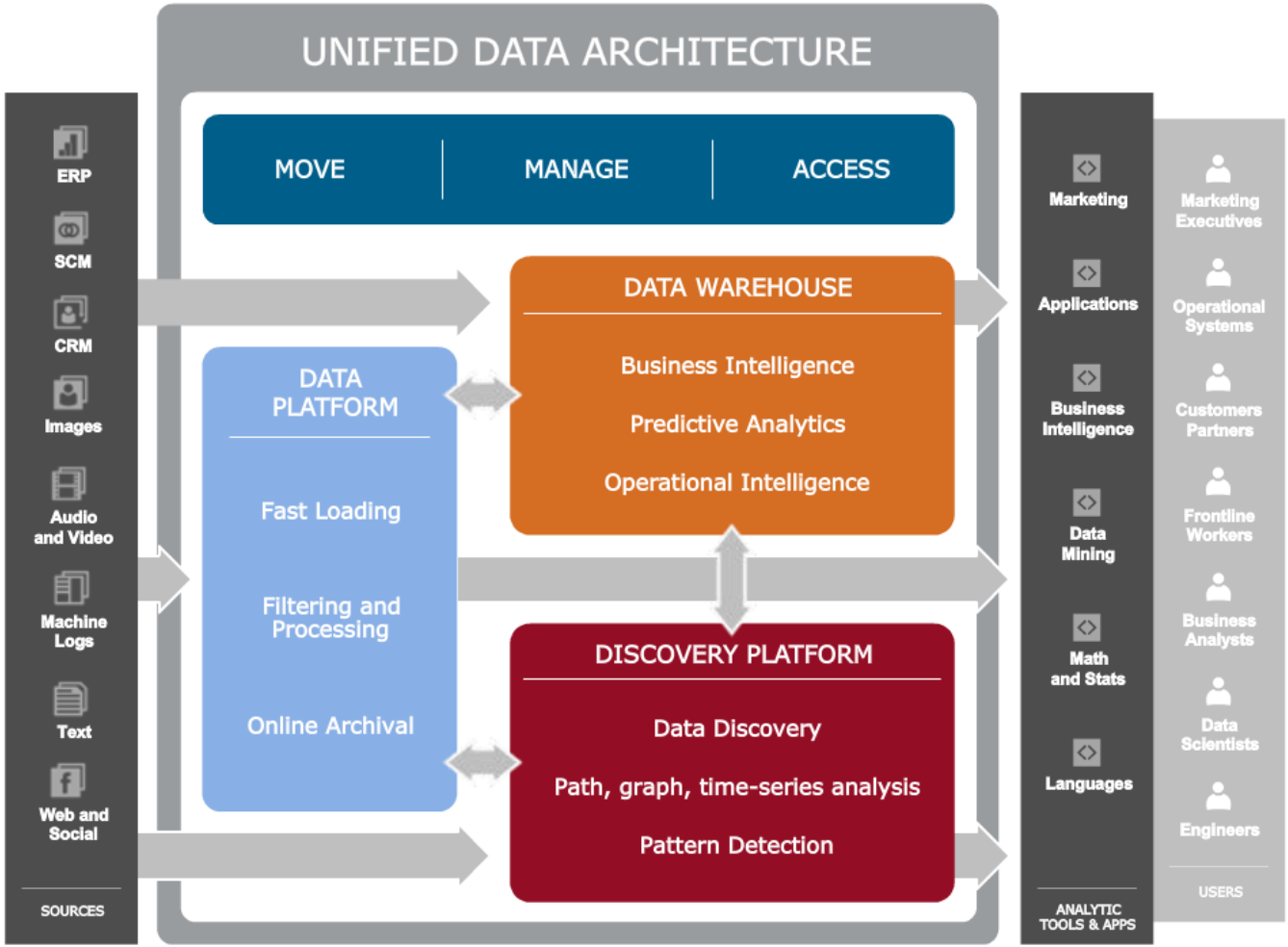


Data Lake: Conceptual architecture

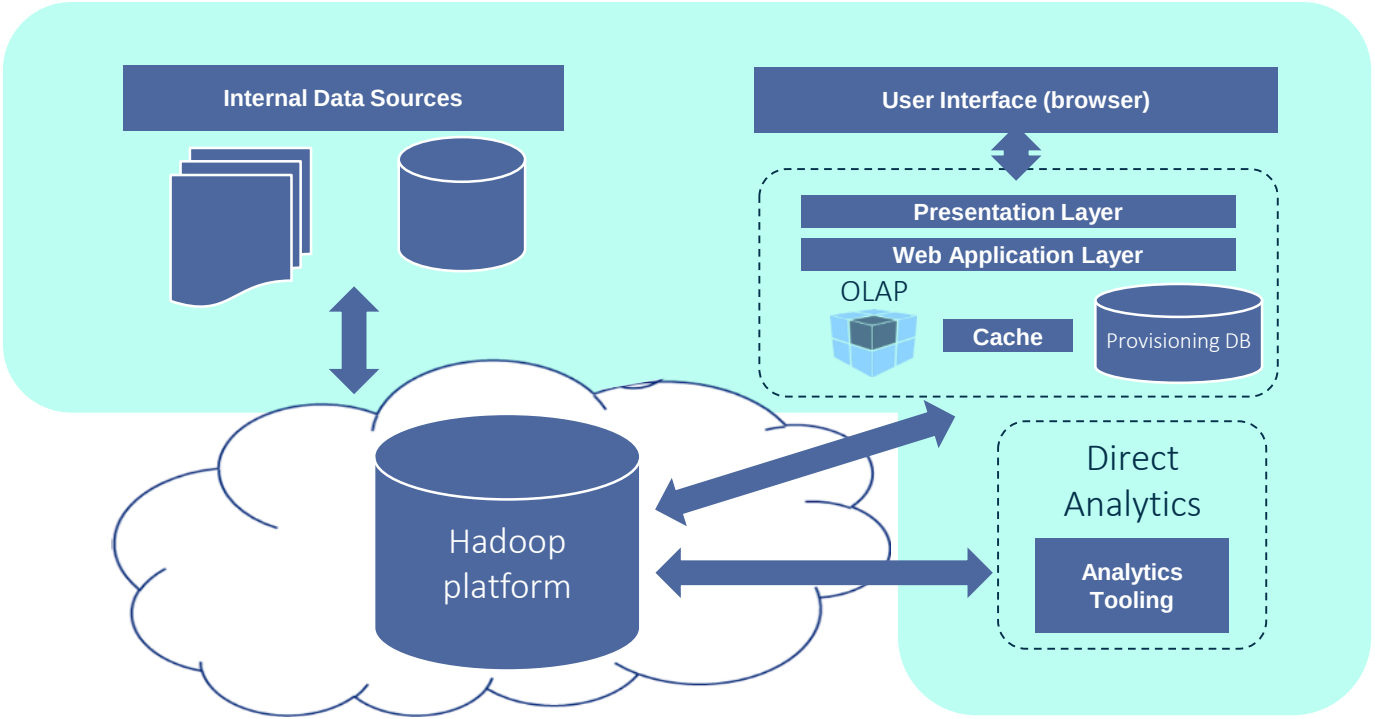


Data lake: Conceptual architecture




UNIFIED DATA ARCHITECTURE
MOVE
MANAGE
ACCESS
DATA WAREHOUSE
Business Intelligence
Predictive Analytics
Operational Intelligence
DATA PLATFORM
Fast Loading
Filtering and Processing
Online Archival
DISCOVERY PLATFORM
Data Discovery
Path, graph, time-series analysis
Pattern Detection
ERP
SCM
CRM
Images
Audio and Video
Machine Logs
Text
Web and Social
SOURCES
Marketing
Applications
Business Intelligence
Data Mining
Math and Stats
Languages
ANALYTIC TOOLS & APPS
Marketing Executives
Operational Systems
Customers Partners
Frontline Workers
Business Analysts
Data Scientists
Engineers
USERS

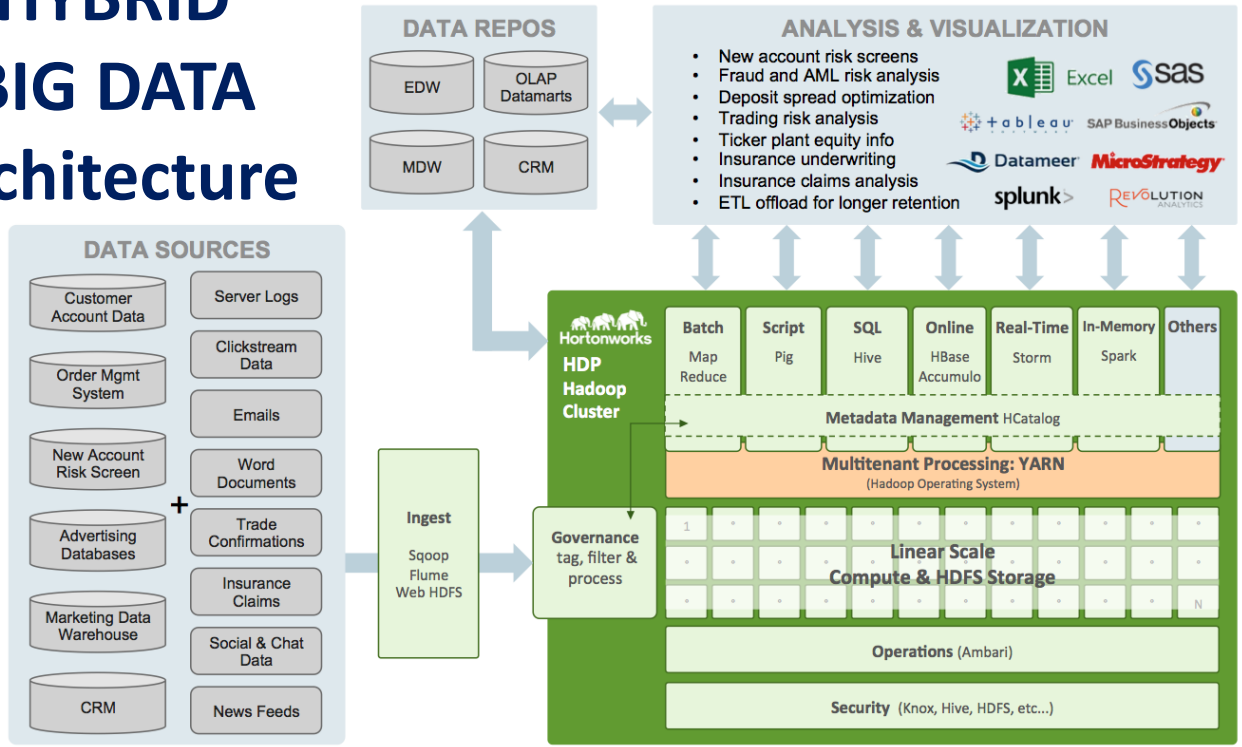
HYBRID INTERNAL & CLOUD



Source - http://www.slideshare.net/AmazonWebServices/analytics-in-the-cloud

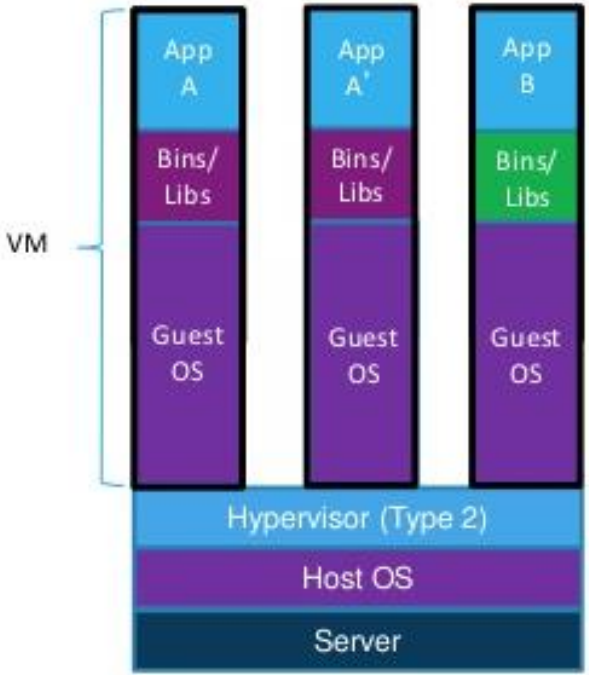
97

Example HYBRID BIG DATA architecture



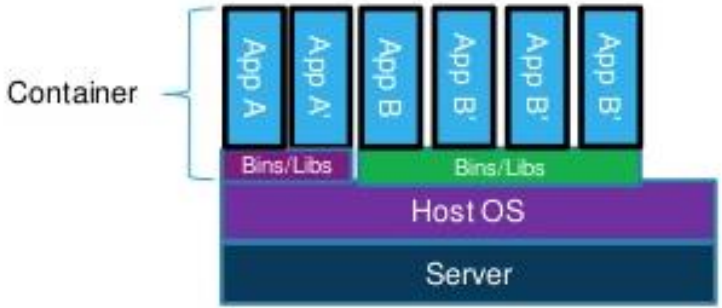
* Horton Works

TRENDING: Virtual machines and containers



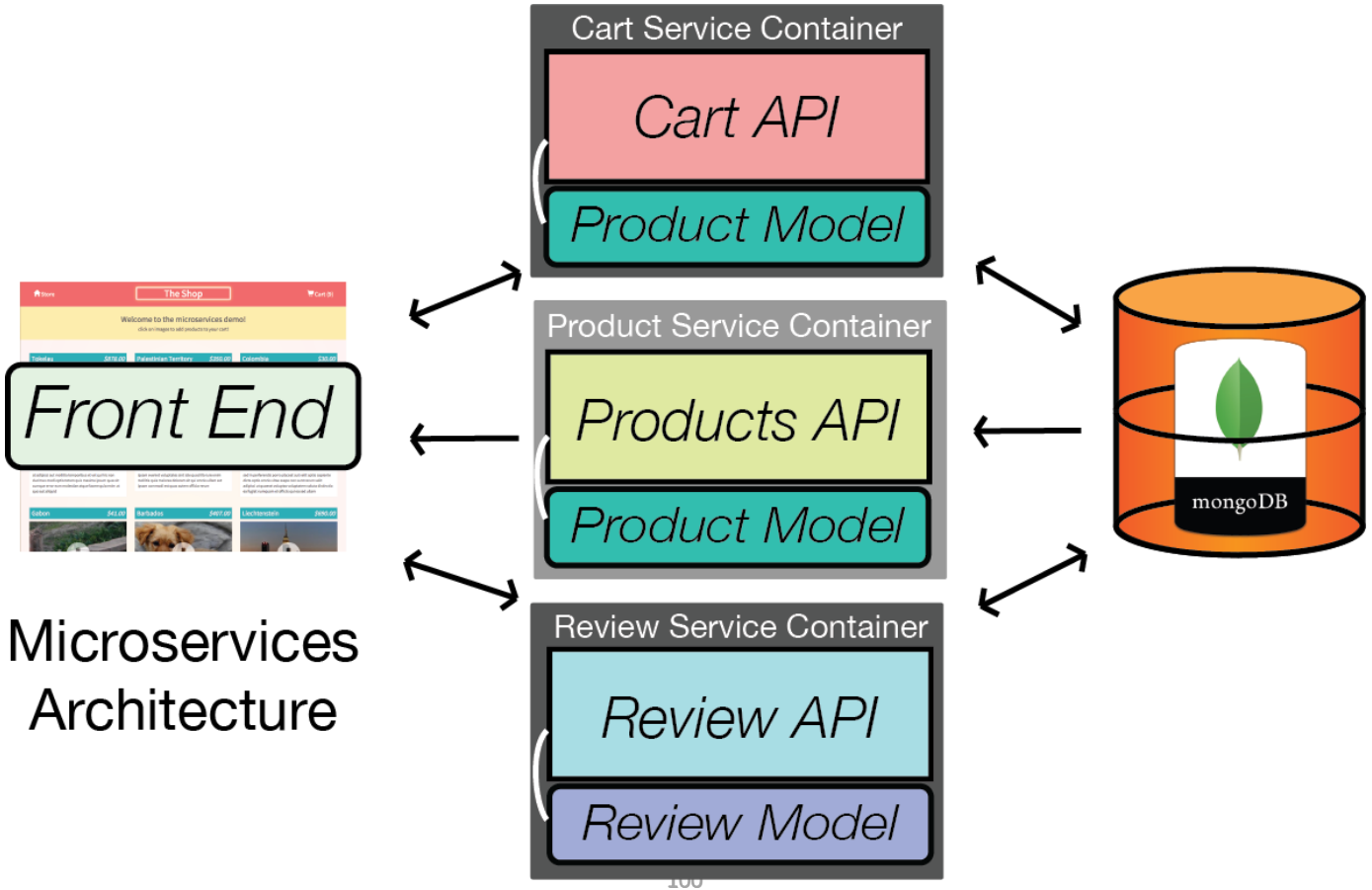
Containers are isolated, but share OS and, where appropriate, bins/libraries

...faster, less overhead



99

TRENDING: containers and microservices



http://blog.ibmjstart.net/2015/07/23/learning-microservices-architecture-bluemix-docker-part-1/

VIYA HYBRID CLOUD & MICROSERVICES ARCHITECTURE