splunk> .conf2017

Deliver Value With The Machine Learning Tool Kit

Alexander Norris | Aetna

September 2017 | Washington, DC

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

About Aetna

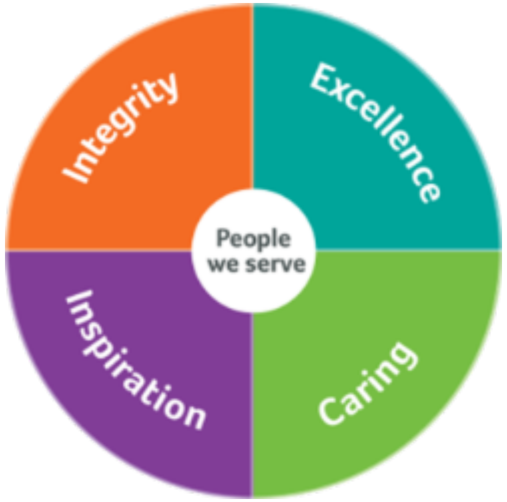
- Founded in 1853 in Hartford, CT, Aetna is committed to providing individuals, employers, health care professionals, producers and others with innovative benefits, products and services.
- Facts:
  - Number of employees: 49,500 (FY 2016)
  - Revenue: $ 63.175 billion (FY 2016)
  - 46.7 Million people rely on us to help them make decisions about their health care

About Aetna

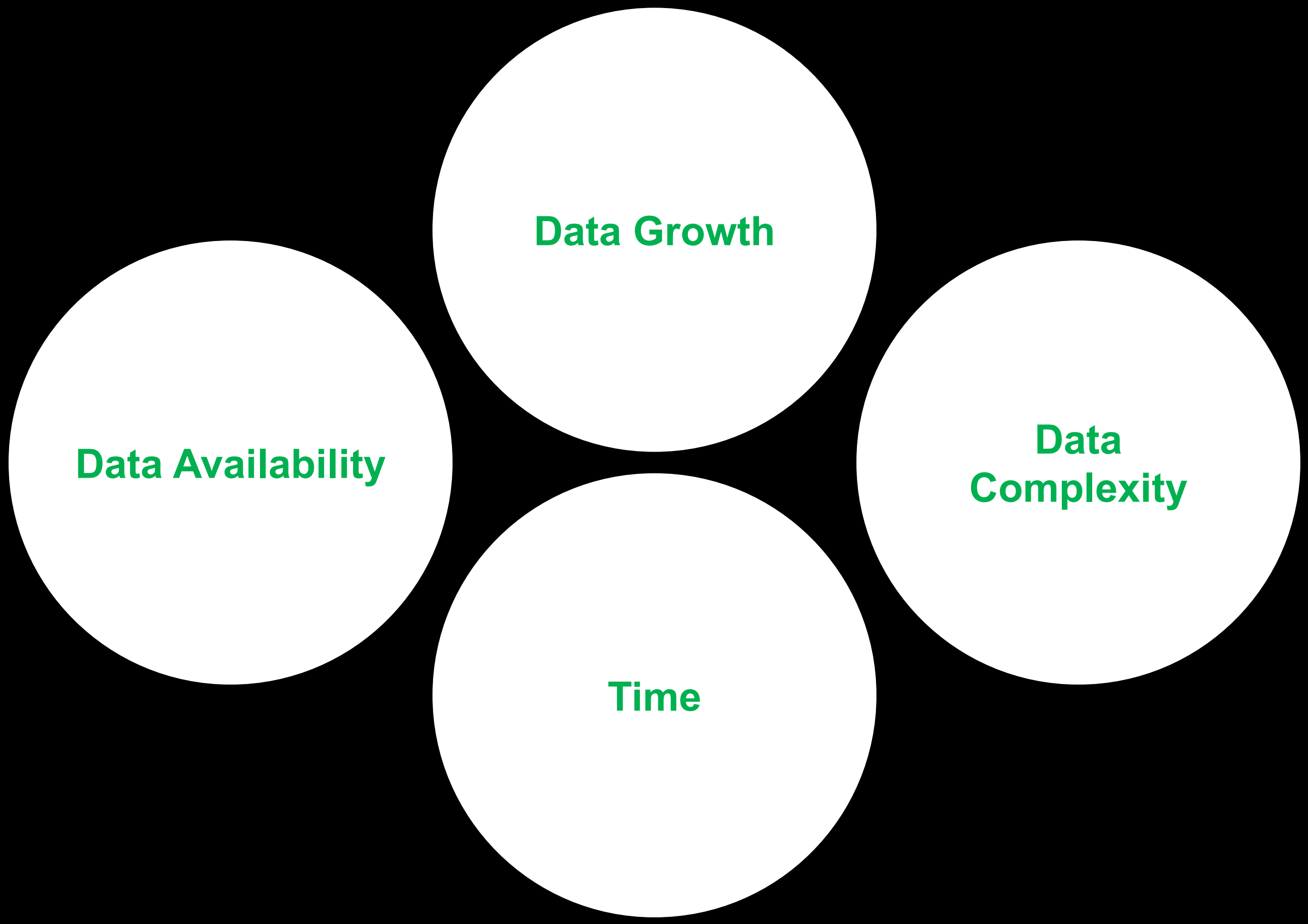
Health care built around people

"Our goal as a company is to find solutions that help people live healthier lives and to help them manage their health versus their health care."

- Mark Bertolini, CEO, Aetna



Challenges



Data Growth

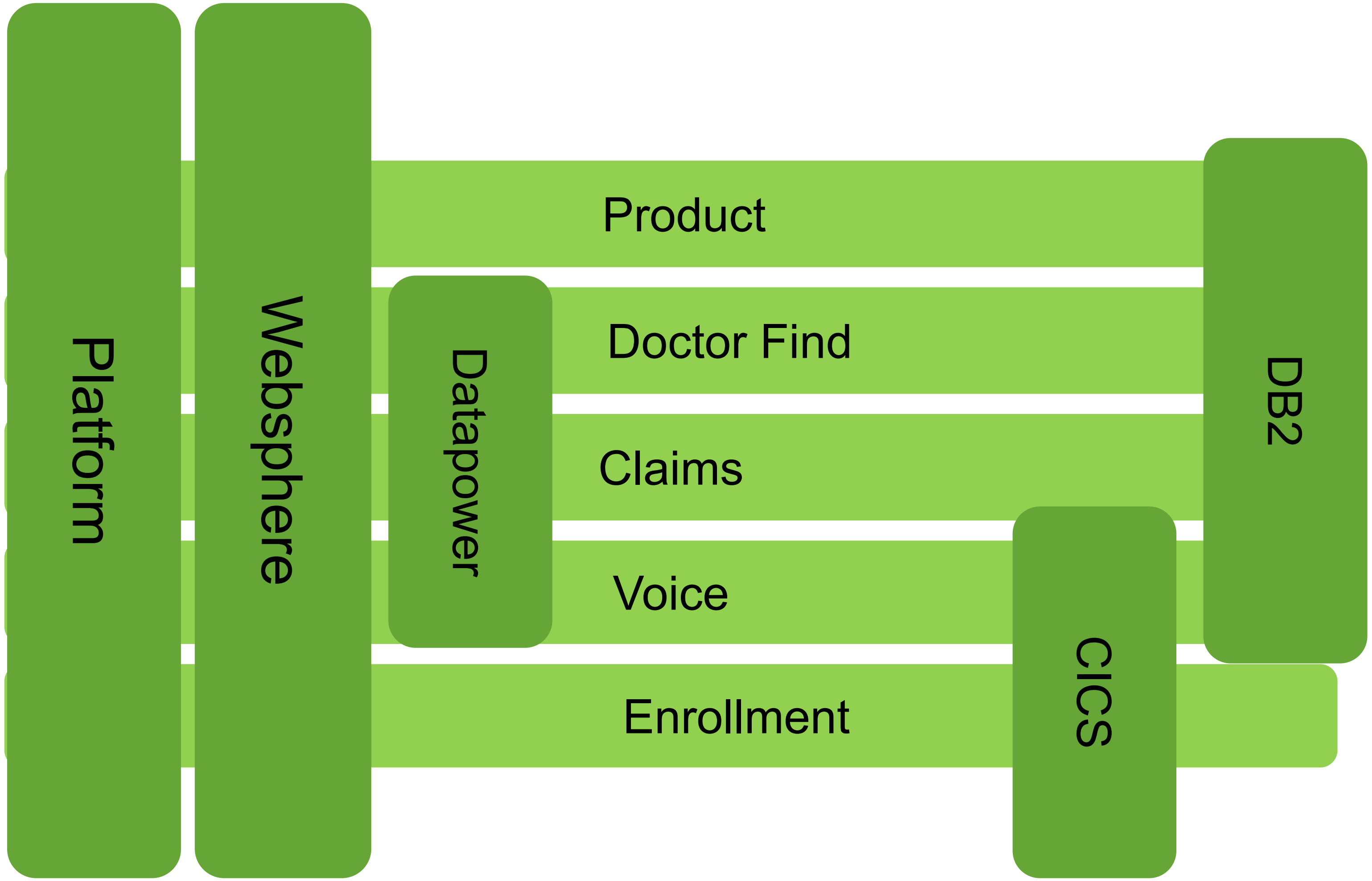
Data Availability

Data Complexity

Time

Approach

- Each source of data is separated by platform and product. This tactic helps define ‘normal’ for each piece of data.
- Filtering the platform data with specific queries limits results to what composes the product.
  - Products are the composite functions (horizontal)
  - Platforms are the uniform functions (vertical)
  - Retrospective vs Proactive

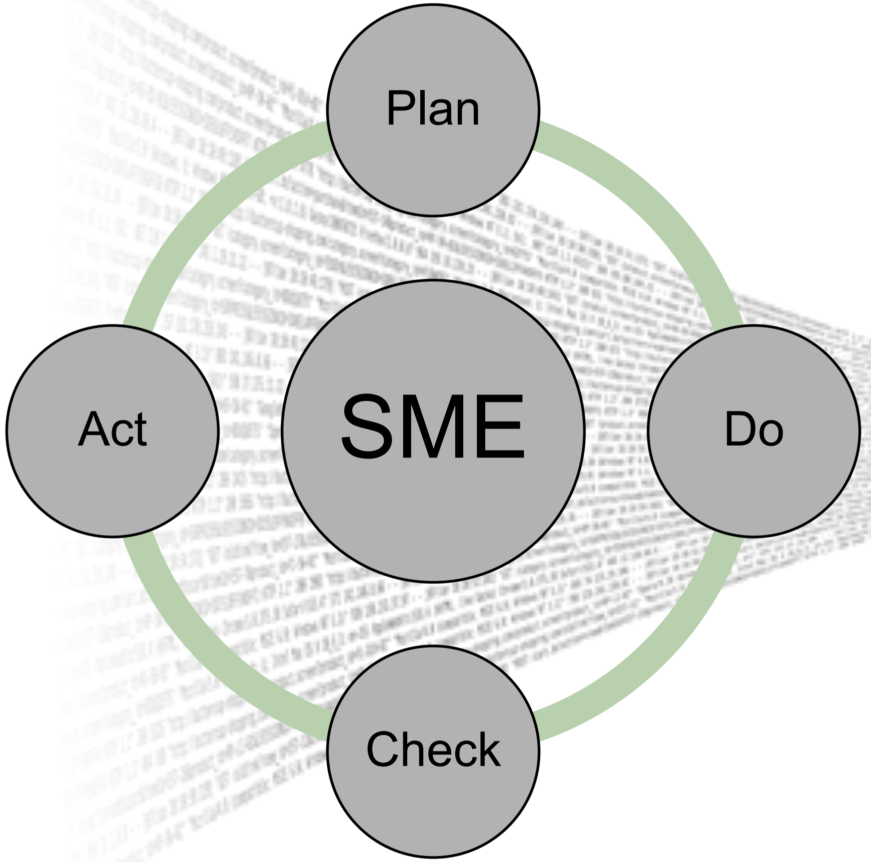


Create timely actionable intelligence with data across Splunk by using the Machine Learning Toolkit

```
130.60.4 - - [07/Jan 18:10:57:153] "GET /category.screen?category_id=GIFTS&JSESSIONID=SD1SL4FF1
128.241.220.82 - - [07/Jan 18:10:57:123] "GET /product.screen?product_id=FL-DSH-01&JSESSIONID
" 317 27.160.0.0 - - [07/Jan 18:10:56:156] "GET /oldlink?item_id=EST-26&JSESSIONID=SD5SL9FF1A
ows NT 5.1; SV1; .NET CLR 1.1.4322)" 468 125.17 14.100
itemId=EST-16&product_id=RP-LI-02" "O
://buttercup-shopping.com/ol
o?action=purchase&it
opping.com/ca
//butter
```

Approach

- Continuous action centered around experts
- Identify and plan
- Dig in and do
- Check and verify
- Activate alerting and scheduled learning



Use Cases

Performance

From Man to Machine

Availability

Break It to Make It

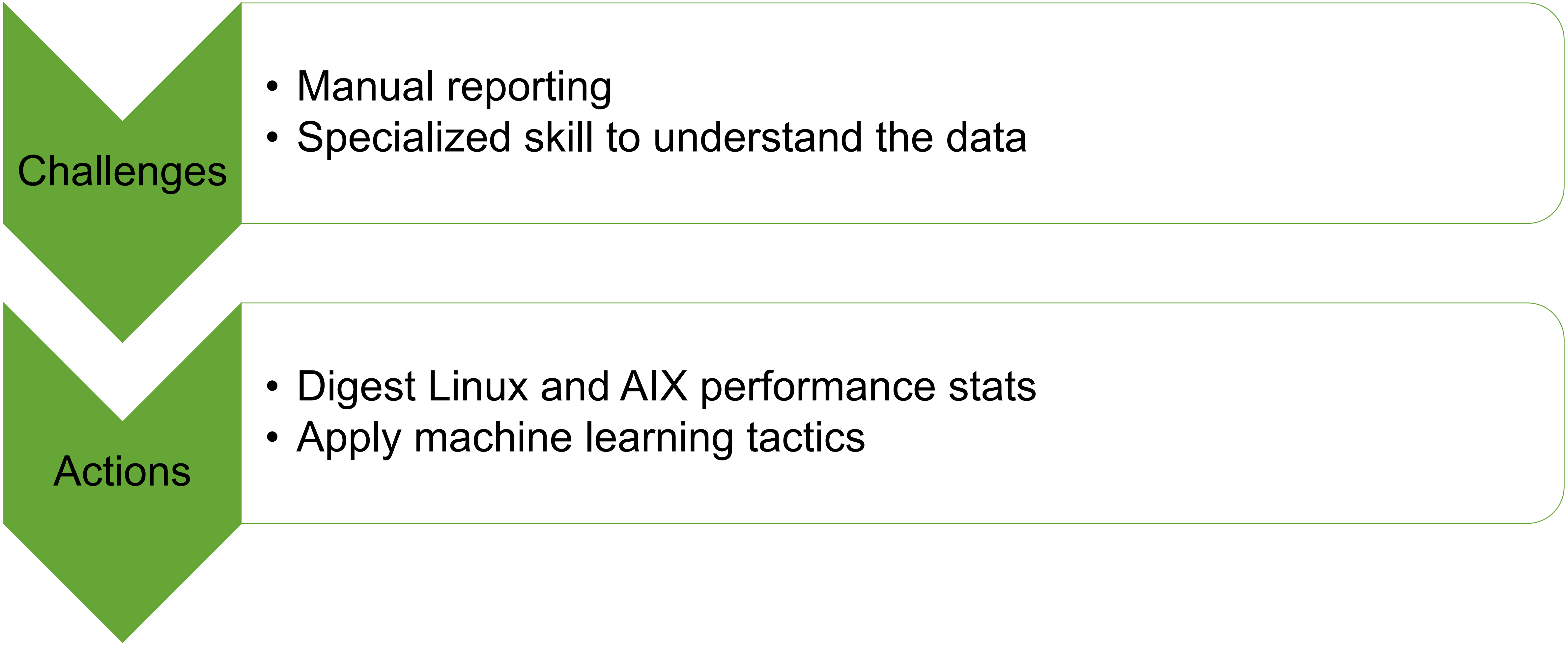
Security

Ensure Service

Capacity

Make It Dynamic

Performance



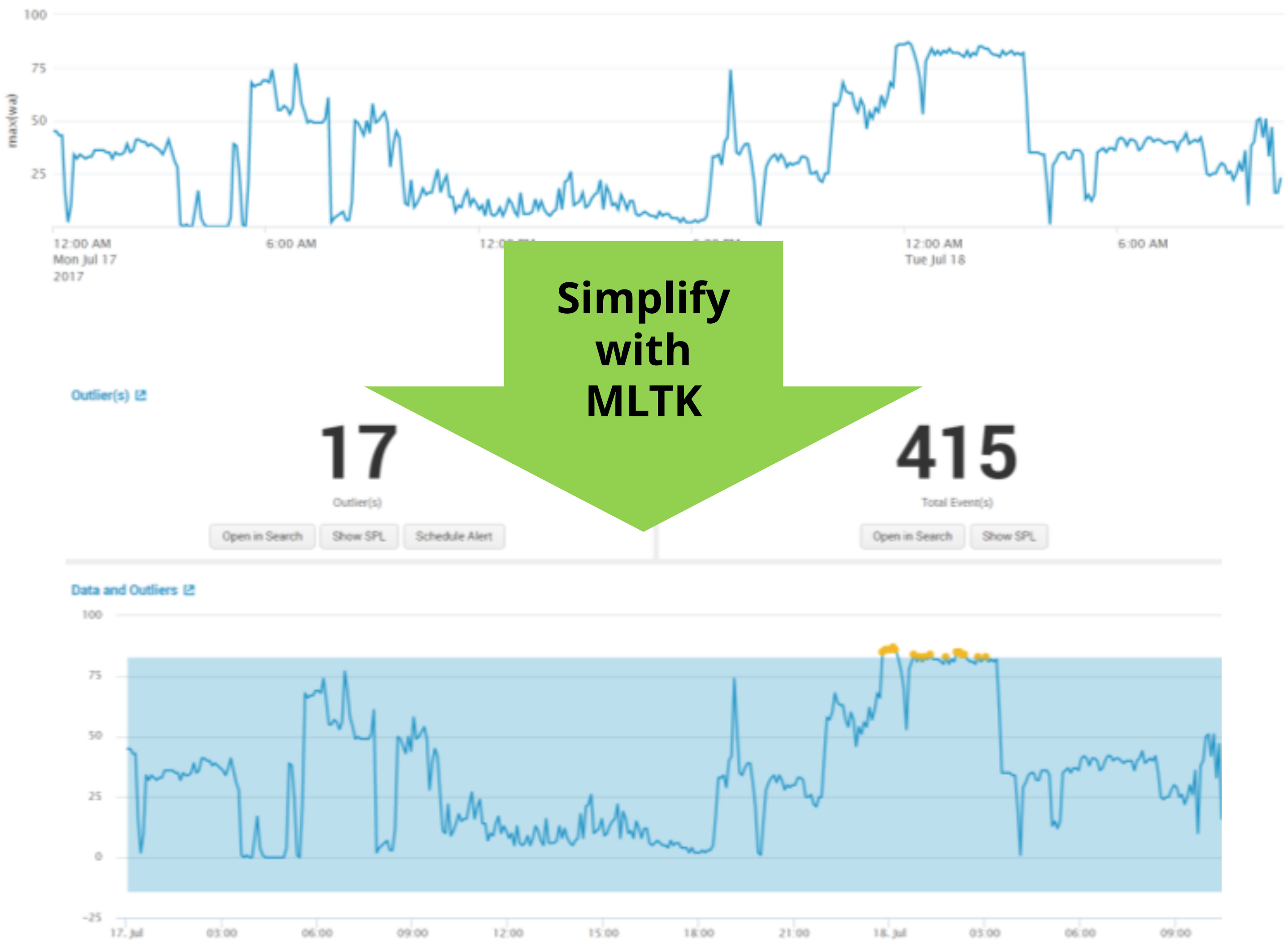
Challenges

- Manual reporting
- Specialized skill to understand the data

Actions

- Digest Linux and AIX performance stats
- Apply machine learning tactics

Performance - Outlier Example



Simplify with MLTK

Performance

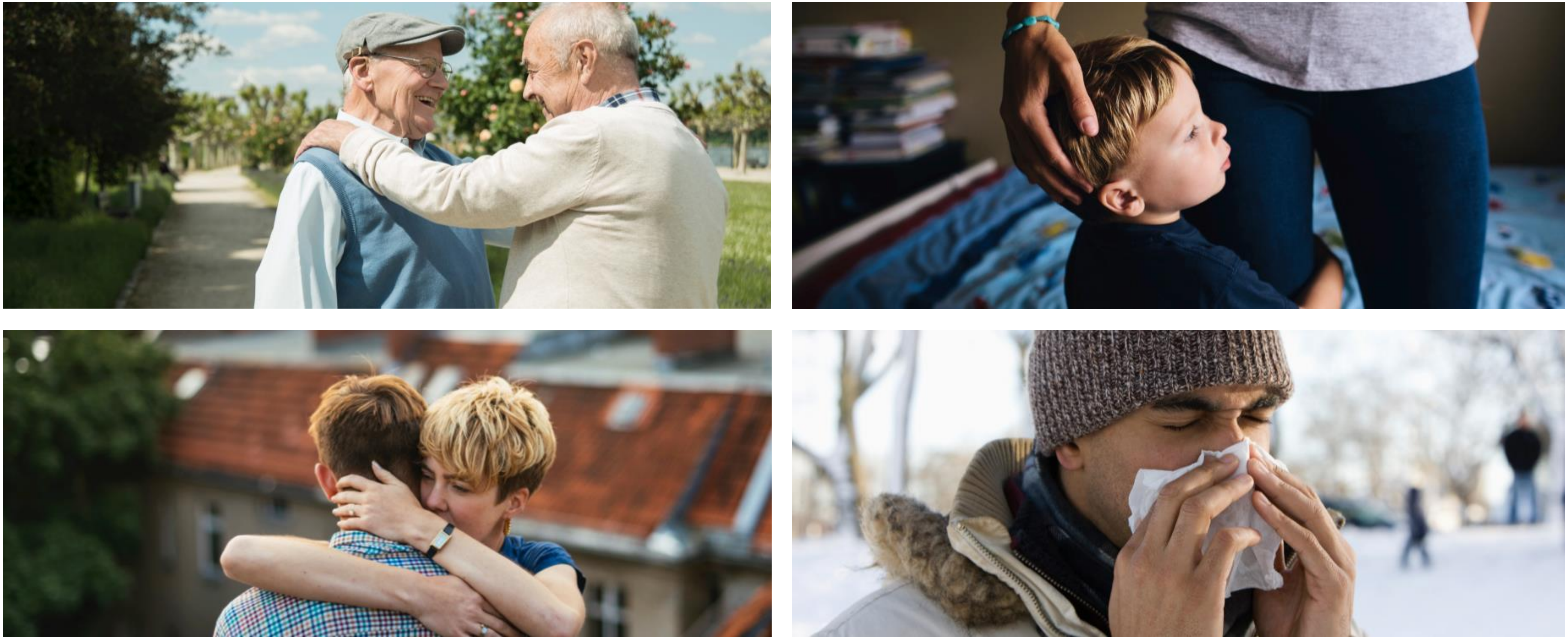
Challenges

- Number of machines
- Specialized skills to understand the data

Value

- Direct insight from individual host anomalies
- Correlation between resource and execution
- Provide a factual perspective while defining normal

Availability



The Need for Care is Constant

Availability

Challenges

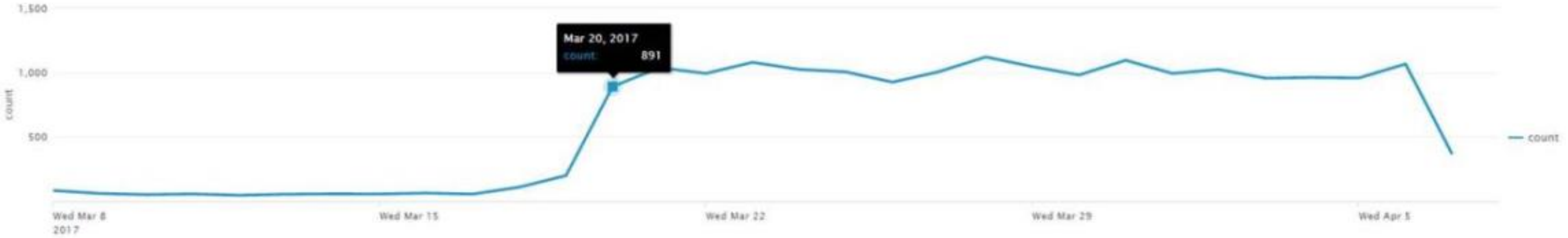
- How and why failures occur
- Hard to define value with proactive work
- Premier application rollout

Actions

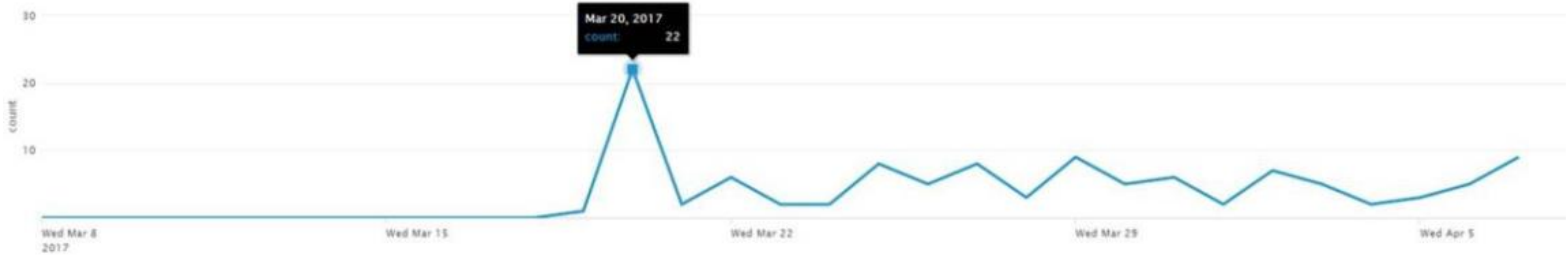
- Test in production
- Collect logs and resource information
- Reap the intelligence

Availability

VMWare Kernel Busy



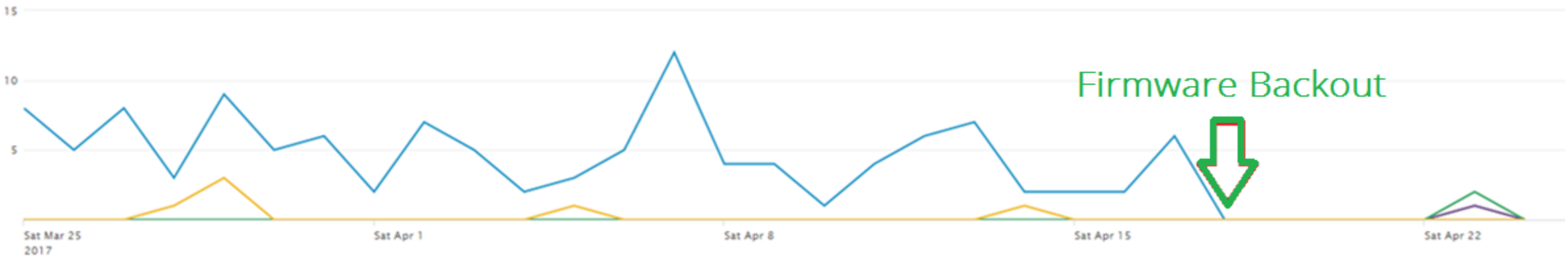
Hitachi Cannot Contact Active Directory



Availability

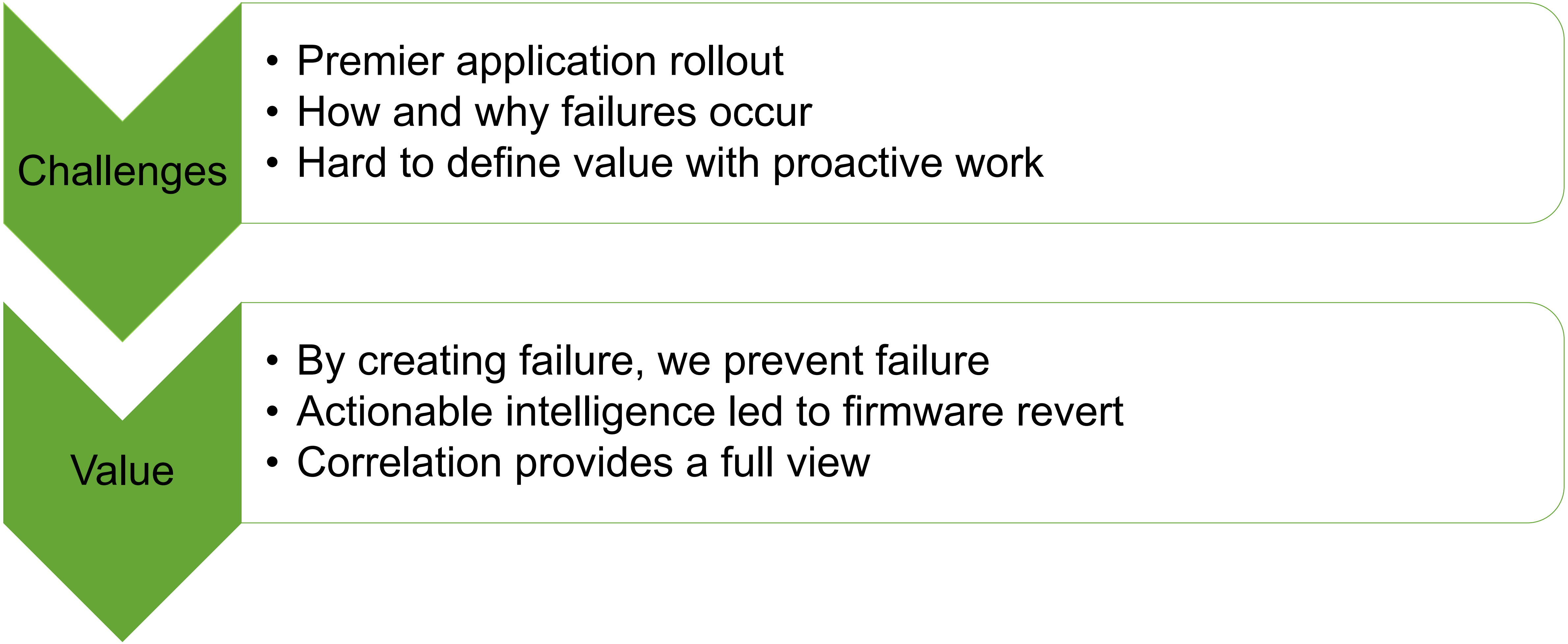
Firmware Reversion

- Following the recognition and this rectified problem…



Firmware Backout

Availability



Challenges

- Premier application rollout
- How and why failures occur
- Hard to define value with proactive work

Value

- By creating failure, we prevent failure
- Actionable intelligence led to firmware revert
- Correlation provides a full view

Availability

More Value

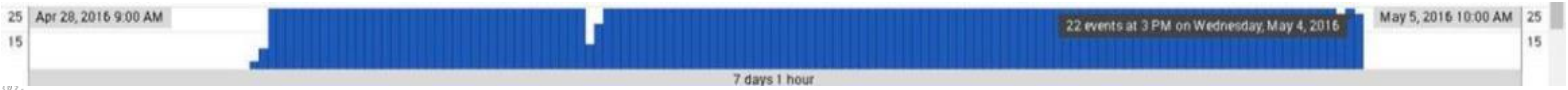
- Make sure common components are healthy
- A backstop to expected function
- Define the ‘unknown’

P2PGroup I ODCF8040I: Detected process proddsecellm\host\JVM started.

ServerInstanc W **HMC_SuspectRC**

ServerInstanc W WXDH0010W: Unexpected restart state verifying for server CELL/host/JVM.

NodeAgent W ADML0011E: The server launched, but failed initialization. See the server log files for failure information.

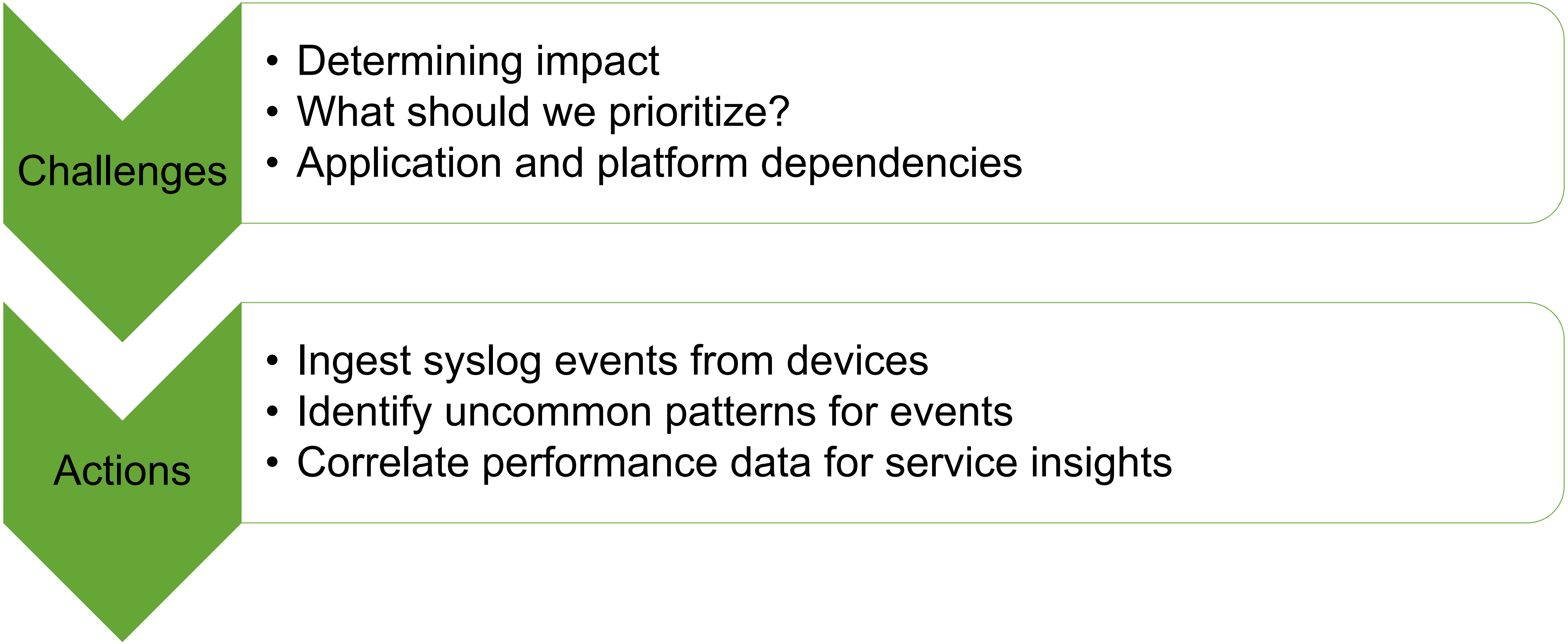


splunk> .conf2017

Security

You Don’t Join Us, We Join You.

Security



Challenges

- Determining impact
- What should we prioritize?
- Application and platform dependencies

Actions

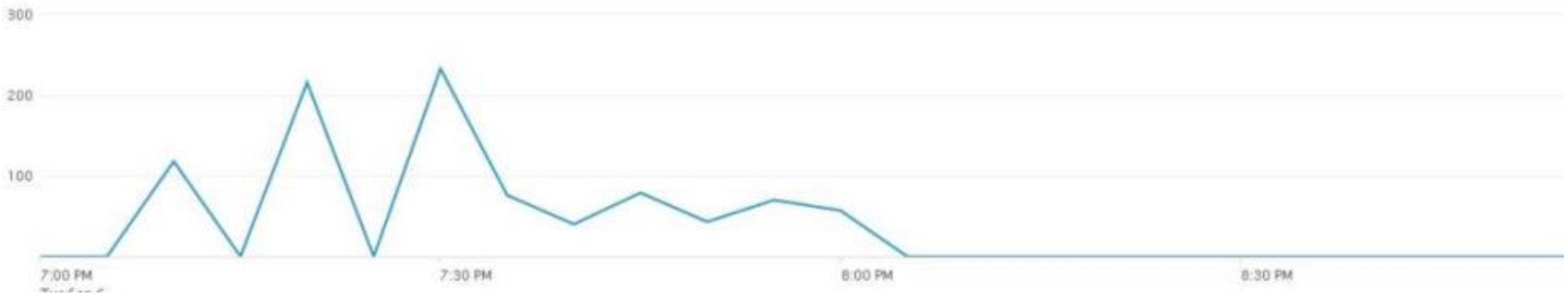
- Ingest syslog events from devices
- Identify uncommon patterns for events
- Correlate performance data for service insights

Security

Value

- Understand credential behavior
- Know a local service is unhealthy

```
sshd: [authpriv.info] subsystem request for sftp
sshd: [authpriv.warning] nss_vas: getgrgid: could not get cache context, err = Unknown error 18446744073709551615
```



Security

Value

- Intranet to cloud migration ‘blank screen’ solved
- Web proxy along with Active Directory valuable set

```
TIME AM    DATE  403 CLIENTIP NoAuthProd AllowListProd "-" "-" AV_SCANNED - GET
https "aetna.jiveon.com" 443 23.76.216.43 "Business" "Minimal Risk" 0 GTI_CLOUD "US"
"-" "/socket" "?when=open&transport=sse&heartbeat=false&lastEventId=&id=b8796301-
5c02-4f26-8bd3-b496232f27e7&_=1452016994120" "text/html" TCP_MISS_RELOAD 214
1034 1358 1034 1358 206.213.217.141 -0500 2 "-" "-" "Mozilla/5.0 (Windows NT 6.1;
WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/47.0.2526.106 Safari/537.36"
```

Capacity

There When You Need Us

Capacity

Challenges

- Workload behavior
- Don’t want to cover up a problem with resources
- Integrating the many layers of the virtual world

Actions

- Activate resource and event data collection
- Create custom SQL to map the entities
- Create a process for root cause determination

Capacity

Splunk Machine Learning Toolkit + Turbonomics = **Better Insight**

Capacity

Challenges

- Workload behavior
- Don’t want to cover up a problem with resources
- Integrating the many layers of the virtual world

Value

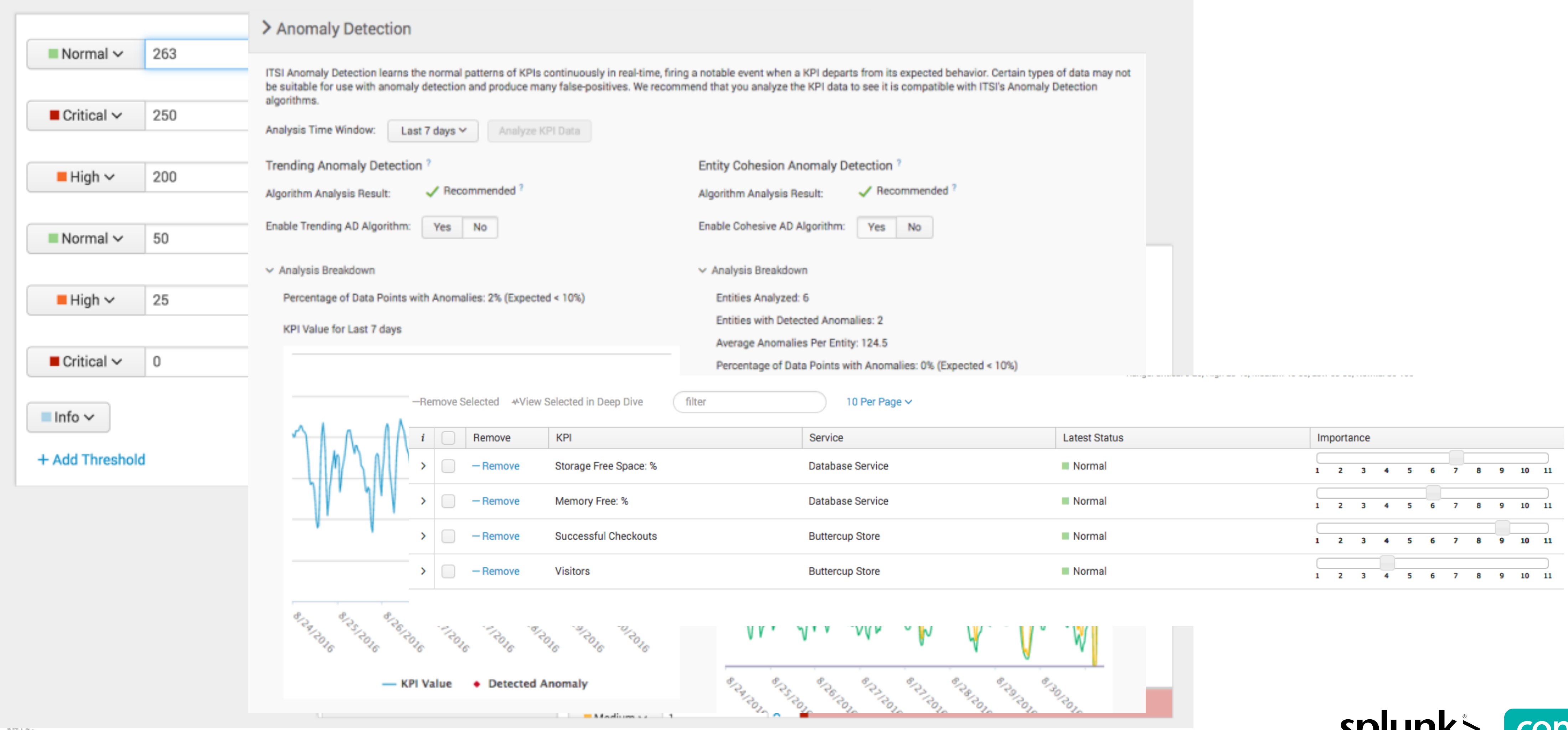
- Workload better understood with resource control
- Knowing why a VMotion occurs is insight
- Storage instrumentation very valuable

Lessons Learned

- Important:
  - Start with small wins and create rapid value
  - Full coverage (data validation) CYAN
  - Data collection configuration (APPS)
  - SME perspectives are very important
  - Retrospective value easier to define
  - Firmware devices great place to start

Future

ITSI – Adaptive Thresholds and Alerting



🟩 Normal	263
🟥 Critical	250
🟧 High	200
🟩 Normal	50
🟧 High	25
🟥 Critical	0
🟦 Info	

+ Add Threshold

Anomaly Detection

ITSI Anomaly Detection learns the normal patterns of KPIs continuously in real-time, firing a notable event when a KPI departs from its expected behavior. Certain types of data may not be suitable for use with anomaly detection and produce many false-positives. We recommend that you analyze the KPI data to see it is compatible with ITSI's Anomaly Detection algorithms.

Analysis Time Window: Last 7 days | Analyze KPI Data

Trending Anomaly Detection

Algorithm Analysis Result: ✓ Recommended

Enable Trending AD Algorithm: Yes | No

Analysis Breakdown

Percentage of Data Points with Anomalies: 2% (Expected < 10%)

KPI Value for Last 7 days

Entity Cohesion Anomaly Detection

Algorithm Analysis Result: ✓ Recommended

Enable Cohesive AD Algorithm: Yes | No

Analysis Breakdown

Entities Analyzed: 6

Entities with Detected Anomalies: 2

Average Anomalies Per Entity: 124.5

Percentage of Data Points with Anomalies: 0% (Expected < 10%)

–Remove Selected | View Selected in Deep Dive | filter | 10 Per Page

i	□	Remove	KPI	Service	Latest Status	Importance
›	□	– Remove	Storage Free Space: %	Database Service	🟩 Normal	7
›	□	– Remove	Memory Free: %	Database Service	🟩 Normal	6
›	□	– Remove	Successful Checkouts	Buttercup Store	🟩 Normal	9
›	□	– Remove	Visitors	Buttercup Store	🟩 Normal	4

Splunk UBA

REAL TIME & BIG DATA ARCHITECTURE

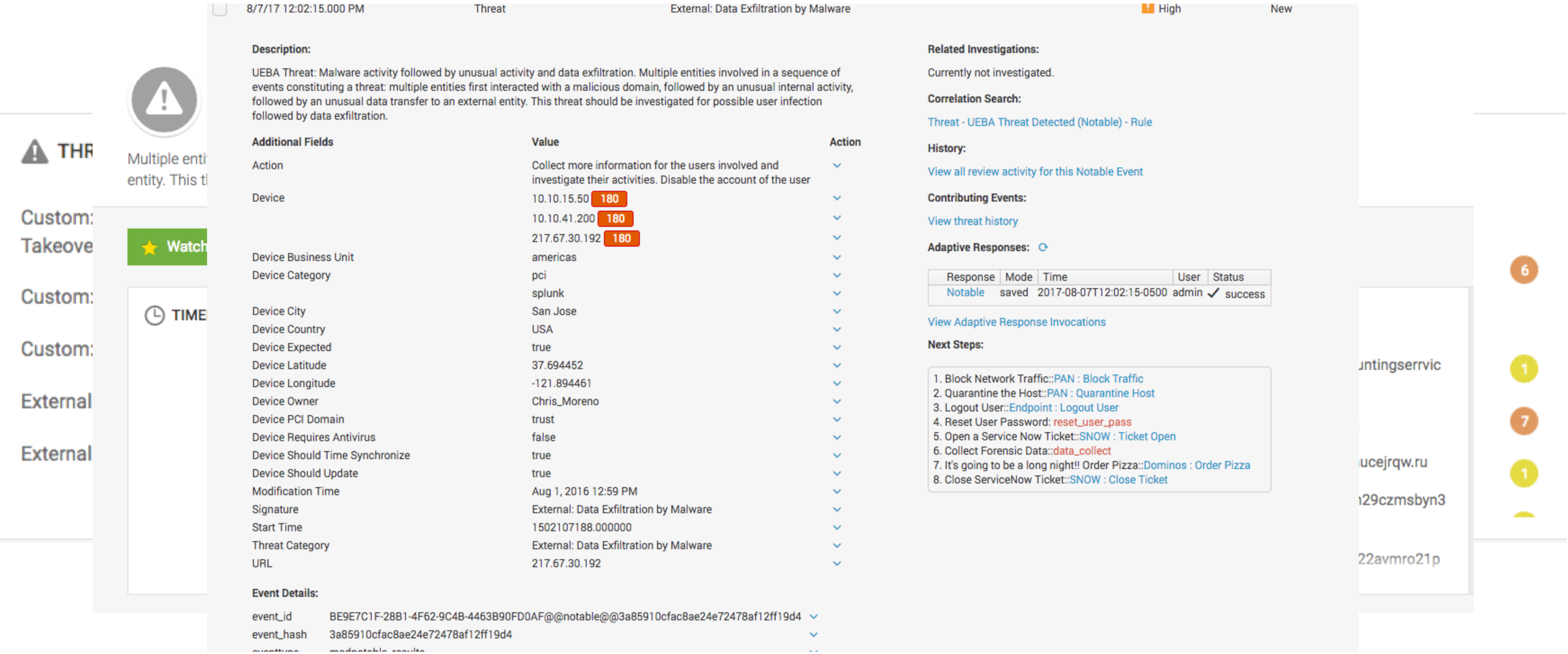
BEHAVIOR MODELING

UNSUPERVISED MACHINE LEARNING

ANOMALY DETECTION

THREAT DETECTION

UBA



□ 8/7/17 12:02:15.000 PM Threat External: Data Exfiltration by Malware High New

Description:

UEBA Threat: Malware activity followed by unusual activity and data exfiltration. Multiple entities involved in a sequence of events constituting a threat: multiple entities first interacted with a malicious domain, followed by an unusual internal activity, followed by an unusual data transfer to an external entity. This threat should be investigated for possible user infection followed by data exfiltration.

Additional Fields	Value	Action
Action	Collect more information for the users involved and investigate their activities. Disable the account of the user	⌄
Device	10.10.15.50 180	⌄
	10.10.41.200 180	⌄
	217.67.30.192 180	⌄
Device Business Unit	americas	⌄
Device Category	pci	⌄
	splunk	⌄
Device City	San Jose	⌄
Device Country	USA	⌄
Device Expected	true	⌄
Device Latitude	37.694452	⌄
Device Longitude	-121.894461	⌄
Device Owner	Chris_Moreno	⌄
Device PCI Domain	trust	⌄
Device Requires Antivirus	false	⌄
Device Should Time Synchronize	true	⌄
Device Should Update	true	⌄
Modification Time	Aug 1, 2016 12:59 PM	⌄
Signature	External: Data Exfiltration by Malware	⌄
Start Time	1502107188.000000	⌄
Threat Category	External: Data Exfiltration by Malware	⌄
URL	217.67.30.192	⌄

Event Details:

event_id	BE9E7C1F-28B1-4F62-9C4B-4463B90FD0AF@@notable@@3a85910cfac8ae24e72478af12ff19d4	⌄
event_hash	3a85910cfac8ae24e72478af12ff19d4	⌄

Related Investigations:

Currently not investigated.

Correlation Search:

Threat - UEBA Threat Detected (Notable) - Rule

History:

View all review activity for this Notable Event

Contributing Events:

View threat history

Adaptive Responses: ⟳

Response	Mode	Time	User	Status
Notable	saved	2017-08-07T12:02:15-0500	admin	✓ success

View Adaptive Response Invocations

Next Steps:

1. Block Network Traffic::PAN : Block Traffic
2. Quarantine the Host::PAN : Quarantine Host
3. Logout User::Endpoint : Logout User
4. Reset User Password: reset_user_pass
5. Open a Service Now Ticket::SNOW : Ticket Open
6. Collect Forensic Data::data_collect
7. It's going to be a long night!! Order Pizza::Dominos : Order Pizza
8. Close ServiceNow Ticket::SNOW : Close Ticket

‘Human Care’

Final Thought

Thank You

Don't forget to rate this session in the .conf2017 mobile app

splunk> .conf2017