Data Analytics for DevOps and Cloud Transformation

SCITDA Leaders’ Workshop
March, 2018

Andi Mann  |  Chief Technology Advocate
@AndiMann  |  amann@splunk.com
About Your Facilitator
Andi Mann – Chief Technology Advocate, Splunk

Global experience as a strategist, technologist, innovator, and communicator with Fortune 500 corporations, software vendors, governments, and as a leading research analyst and consultant. Business and technology commentator appearing in USA Today, New York Times, SkyTV, Forbes, CIO, InformationWeek, Wall Street Journal, and more.

Named to many ‘Top … ’ lists including Business Insider's Top Thought-Provoking Enterprise Tech Execs, Apollo Research’s Top Technology Specialists on Twitter, Heller Search’s Top Recommended Twitter Accounts for iT Execs, Robert Half Technology’s Top 20 People Most Mentioned by IT Leaders, Huffington Post's Top 100 Cloud Computing Experts, Gathering Clouds Top 5 Cloud Experts - Who's Who in Cloud, and SAP's Top 50 Cloud Computing Influencers.

Published author of two books - 'Visible Ops – Private Cloud'; and 'The Innovative CIO'; blogger at 'Andi Mann – Übergeek'; tweets as @AndiMann
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.

Splunk, Splunk®, Listen to Your Data, The Engine for Machine Data, Splunk Cloud, Splunk Light and SPL are trademarks and registered trademarks of Splunk Inc. in the United States and other countries. All other brand names, product names, or trademarks belong to their respective owners. © 2018 Splunk Inc. All rights reserved.
Agenda

- **Cloud and DevOps** – common elements that enable cloud and DevOps as transformative approaches
- **Metrics that Matter** – measuring cloud and DevOps for visibility into shared goals and success
- **Analytics from planning to release** – data to transform CI/CD pipelines from planning to release
- **Analytics from release to support** – data to transform monitoring, troubleshooting, & post-incident reviews
- **Analytics for constituent insights** – analyzing end user/constituent interaction for agile feedback loops
- **Analytics for service intelligence** – cross-platform data for deep insight into end-to-end constituent services
- **Analytics for breach detection** – insight into exposures, data breaches, and unauthorized user behaviors
- **Measuring ‘the new stack’** – incl. Site Reliability Engineering’ semantic logging, telemetry, observability
- **Advanced analytics** – techniques incl. machine learning, anomaly detection, and predictive analytics
- **Data-driven automation** – coupling data with automation for actionable decisions and remediation
- **Q&A, Wrap-up**

... with stories in data, analytics, and transformation from Splunk, our customers, and others in the public sector.
Cloud and DevOps -

the common elements of people, process, and technology that enable cloud and DevOps as transformative approaches
Cloud Services Accelerate App Delivery Velocity
Defining DevOps

METHODS FOR IMPROVING

COLLABORATION

COMMUNICATION

INTEGRATION

BETWEEN DEV AND OPS
TO DELIVER BETTER SOFTWARE, FASTER
DevOps Accelerates App Delivery Velocity

Product Managers identify new opportunities

DevOps Teams iterate with continuous insights

Code continuously delivered to market

Auditors have visibility

Customers are happy
Virtualization, Cloud, DevOps, Containers, MSAs, Serverless/FaaS, APIs are Disintegrating Monoliths
CAMS – as close to prescriptive as DevOps gets

Culture
Automation
Measurement
Sharing
CAMS – as close to prescriptive as DevOps gets
Metrics that Matter

what to measure in cloud and DevOps (across people, process, and technology) to provide shared goals and measures of success for transformation
WHAT DATA DRIVES GOOD DECISIONS?

10+ Deploys Per Day: Dev and Ops Cooperation at Flickr

John Allspaw & Paul Hammond
Velocity 2009
I’m working super hard!!

That’s my stapler!
Yeah, but …
… what are you achieving?

I’m gonna need you to come in Sunday.
What activities?
What outcomes?
Gartner’s DevOps ‘Metrics that Matter’
### IDC’s DevOps ‘Metrics that Matter’

**Q. What business outcomes do you expect DevOps practices to deliver?**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>% of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved customer experience</td>
<td>67</td>
</tr>
<tr>
<td>Lower IT costs</td>
<td>61</td>
</tr>
<tr>
<td>Improved employee productivity</td>
<td>44</td>
</tr>
<tr>
<td>Higher profits</td>
<td>39</td>
</tr>
<tr>
<td>Improved IT employee satisfaction</td>
<td>39</td>
</tr>
<tr>
<td>Faster/increased revenue growth</td>
<td>33</td>
</tr>
<tr>
<td>Improved security and risk mitigation</td>
<td>33</td>
</tr>
<tr>
<td>Improved career development</td>
<td>28</td>
</tr>
<tr>
<td>Higher service availability</td>
<td>22</td>
</tr>
<tr>
<td>Improved EPS</td>
<td>11</td>
</tr>
</tbody>
</table>

*n = 18*

Note: Multiple responses were allowed.

Source: IDC’s DevOps Best Practice Metrics: Fortune 1000 Survey, December 2014
Forrester’s DevOps ‘Metrics that Matter’

**Velocity**
- Business - release freq., time/cost per release, mean-time-to-change, mean-time-to-detection
- DevOps team - release/deploy automation%, mean-time-to-detection, mean-time-to-approval

**Quality**
- Business - MTTR, Customer experience
- DevOps team - Deployment failures, incident severities (by team, application, process, asset)

**Efficiency**
- Business - Unplanned work, happiness of CX team with technology delivery
- DevOps team - Deployment frequency/duration, Incident severity, average provisioning time

**Culture**
- Business - Happiness with product team, DevOps team attrition, DevOps meeting frequency
- DevOps team - Rework rate, unplanned work, satisfaction, attrition, postmortem count
### Computing UK’s ‘Metrics that Matter’

#### Which metrics would be most useful in order to measure success of a DevOps implementation?

<table>
<thead>
<tr>
<th>Metric</th>
<th>Entire sample</th>
<th>Adopters of DevOps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer satisfaction</td>
<td>62%</td>
<td>62%</td>
</tr>
<tr>
<td>Speed of application delivery</td>
<td>51%</td>
<td>60%</td>
</tr>
<tr>
<td>Business impact of application code</td>
<td>31%</td>
<td>31%</td>
</tr>
<tr>
<td>Quality of application code</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>Project costs</td>
<td>30%</td>
<td>21%</td>
</tr>
<tr>
<td>User adoption/take up</td>
<td>21%</td>
<td>21%</td>
</tr>
</tbody>
</table>

More Ideas for ‘Metrics that Matter’

**Culture**
e.g.
- Retention
- Satisfaction
- Callouts

**Process**
e.g.
- Idea-to-cash
- MTTR
- Deliver time

**Quality**
e.g.
- Test pass
- Test fail
- Best/worst

**Systems**
e.g.
- Throughput
- Uptime
- Build times

**Activity**
e.g.
- Commits
- Tests run
- Releases

**Impact**
e.g.
- Signups
- Checkouts
- Turnaround
Analytics from planning to release

using data to transform CI/CD pipelines from planning, to code and build, testing, configuration, and release
Feedback Loops Enable Continuous Improvement

Application Development
- Plan
- Code
- Build
- Config

Test and Acceptance
- Test/QA
- Stage
- Release
- Monitor

Production
- Enhancement Requests
- Defect Information
- Capacity Planning
- Quality Standards
- Integration Requirements
- Acceptance Metrics
- Infrastructure Dependencies
- Service Levels and KPIs

Infrastructure
- Server, Storage, N/W
- Virtualization
- Operating Systems
- Applications
- Mobile Applications
- Cloud Services
- Custom Applications
- API Services

Integration
- Requirements

Requests
- Enhancement

Information
- Defect

Planning
- Capacity

Standards
- Quality

Acceptance
- Metrics

Dependencies
- Infrastructure

KPIs
- Service Levels

Test and Acceptance
- Stage

Release
- Production

Monitor
- Plan

Build
- Application Development

Test/QA
- Code

Stage
- Build

Monitor
- Plan

Build
- Plan

Code
- Code

Build
- Build

Config
- Config
Getting Visibility Across Dev and Ops

Common Data Platform – Collect, Analyze, Visualize, Share
Metrics for Resource Analytics

Insight and prediction for effective resource allocation

► Key Metrics:
  • Work time vs. PTO/sick
  • Hours by product/project
  • Resource shortages

► Data Sources:
  • Jira
  • WorkDay
Metrics for Cost Analytics

Measurement and predictability for cost control

▶ Key Metrics:
• Productive hours
• Labor costs
• Plan vs. actual

▶ Data Sources:
• WorkDay
• PeopleSoft
Metrics for DevTeam Analytics

Insight to coder activity for teaming & work/life balance

Key Metrics:
- Commit count
- Commits by author
- Commit days/times

Data Sources:
- GitHub
Metrics for Code Analytics

Real-time data on code quality and compliance

- **Key Metrics:**
  - Code policy compliance
  - Code/file/class complexity
  - Code analysis coverage

- **Data Sources:**
  - GitHub
  - Sonarcube
Metrics for Build Analytics

Find and fix build issues to accelerate product lifecycle

► Key Metrics:
  • Build success/failure
  • Build queue status
  • Build process times

► Data Sources:
  • Jenkins
  • Sonarcube
Metrics for Quality Analytics

Automatically review QA results to improve quality

Key Metrics:
- Defects detected
- Test coverage
- Test executions

Data Sources:
- Selenium
- AppScan
- ServiceNow
Metrics for Config Analytics

Monitor provisioning/config to accelerate time to ‘done’

▶ Key Metrics:
  - Provisioning success/failure
  - Provisioning times
  - Config drift by node

▶ Data Sources:
  - Puppet
Metrics for Release Analytics

Real-time data for better, faster release decisions

▶ Key Metrics:
  • Availability by release
  • Tickets by release
  • Release readiness

▶ Data Sources:
  • ServiceNow
  • SonarCube
  • HP OpenView
Improved DevOps Agility

“It’s like we were working without peripheral vision before and now we have it.”

– Robert Gonsalves, Web Operations

Key Customer Benefits

• Increased success rate of deployments
• Ability to detect issues before they affect broad production
• Monitoring deployment process several times per day
Use Live Data to Better Prepare For Release

Compare the release in dev, staging, pre-prod …

<table>
<thead>
<tr>
<th>Development Average RT</th>
<th>Development Maximum RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>183.44 ms</td>
<td>1656.87 ms</td>
</tr>
</tbody>
</table>

With the release currently in production

<table>
<thead>
<tr>
<th>Production Average RT</th>
<th>Production Maximum RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>191.37 ms</td>
<td>23916.57 ms</td>
</tr>
</tbody>
</table>
Analytics Across the End-to-End Software Pipeline
Don’t Forget to Measure Cultural Change

- e.g.
  - Absenteeism
  - ‘Work from home’
  - Staff attrition and retention
  - eNPS
  - Employee ‘happiness’

Image source: @danslimmon - https://twitter.com/danslimmon/status/80616237926780928
FamilySearch Moves to Continuous Delivery and Gains Real-Time Visibility

“Splunk Cloud has been more stable than our internal implementation and has freed up two resources to work on software development instead of managing infrastructure. It has clearly proven to be cost-effective compared to managing infrastructure ourselves.”

– Director of Engineering, FamilySearch

- Successful migration from monthly releases to over 900 deploys per day
- Ability to re-allocate 12 developers to more value-added tasks
- Visibility into the AWS environment to support AWS migration strategy
Analytics from release to post-mortem

using data to transform event management, problem analysis, troubleshooting, and post-incident reviews
Getting Visibility Across Dev and Ops

Common Data Platform – Collect, Analyze, Visualize, Share
Data-driven Feedback Drives Continuous Improvement
Get Visibility into Ops Status and Incidents


- Incident Priorities pie chart with Priority 4, Priority 1, Priority 2, and Priority 3 categories.


- Problem Aging bar chart with counts ranging from 0 to 150.
Analytics to Ensure Infrastructure Health
## Analytics for Visibility into Storage and Capacity

<table>
<thead>
<tr>
<th>VMAX Virtual Free Capacity</th>
<th>VMAX Virtual Used Capacity</th>
<th>VMAX Virtual Total Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.85 TB</td>
<td>3.90 TB</td>
<td>30.75 TB</td>
</tr>
</tbody>
</table>

**Storage Group Compliance:**
- **Stable:** 26
- **Marginal:** 0
- **Critical:** 0

**VMAX Storage Arrays**

<table>
<thead>
<tr>
<th>VMAX</th>
<th>Model</th>
<th>uCode</th>
<th># Devices</th>
<th>Free Capacity (TB)</th>
<th>Used Capacity (TB)</th>
<th>Total Capacity (TB)</th>
<th>Overall Compression Ratio</th>
<th>Overall Efficiency Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.44</td>
<td>25.30</td>
<td>30.74</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

© 2018 SPLUNK INC.
Analytics to Manage Cloud Resources
## Source Data for Containers and MSAs

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Where to Find It</th>
<th>What It Can Tell You</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container and microservices logs</td>
<td>Logs can be ingested via any native Docker logging driver such as syslog, Splunk, JournalD and via Cloud integrations (e.g., Amazon CloudWatch, Google Cloud Platform Logging Export)</td>
<td>Container and application errors. Monitor any performance counters that can be calculated on top of logs (e.g., web and application server logs)</td>
</tr>
<tr>
<td>Container metrics and events</td>
<td>Docker APIs (e.g., Docker inspect, Docker top, Docker stats, Docker events), cloud APIs (e.g., AWS CloudWatch, Google Stackdriver)</td>
<td>Health, performance, availability and events generated by all monitored containers</td>
</tr>
<tr>
<td>Container clusters, nodes and applications</td>
<td>Docker UCP APIs and logs from containers</td>
<td>Application health, nodes, clusters and containers associated with an application, change history of containers and configuration</td>
</tr>
<tr>
<td>Application logs</td>
<td>Custom logs set by application developers</td>
<td>Application errors and other valuable machine data logged by developers</td>
</tr>
<tr>
<td>Wire data</td>
<td>Wire data probes (software based)</td>
<td>Communication between an app component, application response times and payload of applications as they traverse your network (even when you may not have direct visibility to some app components)</td>
</tr>
</tbody>
</table>
Maryland’s Prince George’s County Mission-Ready With Splunk

“Splunk is a platform for Operational Intelligence for Prince George’s County. With Splunk, we’re able to have greater visibility across functional teams, to identify trends and potential problems in advance and to resolve issues more quickly by seeing a broader view of the problem.”

– Enterprise Architect, Prince George’s County OIT

- Improved government efficiency and transparency to better serve constituents
- Helped small IT team reduce time to identify and resolve IT issues from days or weeks to hours
- Transformed county operations by replacing data silos with a platform for IT operations, application monitoring and security
Analytics for constituent insights

collecting and analyzing end user activity and constituent interaction data to establish agile feedback loops to IT
Data Tells a Story

Order Processing

Middleware Error

Care IVR

Twitter

ORDER, 2014-05-21T14:04:12.484, 10098213, 569281734, 67.17.10.12, 43CD1A7B8322, SA-2100

May 21 14:04:12.996 wl-01.acme.com Order 569281734 failed for customer 10098213. Exception follows: weblogic.jdbc.extensions.ConnectionDeadSQLException:
weblogic.common.resourcepool.ResourceDeadException: Could not create pool connection. The DBMS driver exception was: [BEA][Oracle JDBC Driver]Error establishing socket to host and port: ACMEDB-01:1521. Reason: Connection refused

05/21 16:33:11.238 [CONNEVENT] Ext 1207130 (0192033): Event 20111, CTI Num: ServID: Type 0:19:0, App 0, ANI T7998#1, DNIS 5555685981, SerID 40489a07-7f6e-4251-801a-13ae51a6d092, Trunk T451.16
05/21 16:33:11.242 [SCREENPOEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092, CUSTID 10098213
05/21 16:37:49.732 [DISCEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092

{actor:{displayName:"Go Boys!",followersCount:1366,friendsCount:789,link:"http://dallascowboys.com/"},location:{displayName:"Dallas, TX",objectType:"place"},objectType:"person",preferredUsername:"B0ysF@m80",statusesCount:6072},body:"Can't buy this device from @ACME. Site doesn't work! Called, gave up on waiting for them to answer! RT if you hate @ACME!!",objectType:"activity",postedTime:"2014-05-21T16:39:40.647-0600"}
## Data Tells a Story

<table>
<thead>
<tr>
<th>Sources</th>
<th>Order ID</th>
<th>Product ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order Processing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middleware Error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care IVR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Sources Details
- **Customer ID**: 10098213
- **Order ID**: 569281734
- **Product ID**: 10098213
- **Time Waiting On Hold**: 05/21 14:04:12.996
- **Customer's Tweet**: 05/21 14:04:12.996
- **Company's Twitter ID**: 05/21 14:04:12.996

### Customer's Tweet
{actor: {displayName: "Go Boys!"}, followersCount: 1366, friendsCount: 789, link: "http://dallascowboys.com/", location: {displayType: "location", objectType: "person", preferredUsername: "B0y3f@n80", statusesCount: 6072}, body: "Can't buy this device from @ACME. Site doesn't work! Called, gave up on waiting for them to answer! RT if you hate @ACME!"}
Metrics for Impact Analytics

Realtime business insight to drive impactful development

► Key Metrics:
  • Revenue per minute
  • Checkout rate
  • Cart fulfillment/abandon

► Data Sources:
  • Web logs
  • HTTP events
  • SFA/CRM
Going Deeper with Business Analytics

- Insight across the complete web-based business process
- Insight into end user experience, engagement, and behavior
- Insight into product, service, or feature adoption, usage, and effectiveness
- Insights across the complete end-to-end business process
### Constituent Experience Analytics – Data Sources

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Where to Find It</th>
<th>What It Can Tell You</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Logs</td>
<td>Local log files, log4j, log4net, Weblogic, WebSphere, JBoss, .NET, PHP</td>
<td>User activity, fraud detection, application performance</td>
</tr>
<tr>
<td>Business Process Logs</td>
<td>Business process management logs</td>
<td>Customer activity across channels, purchases, account changes, process bottlenecks</td>
</tr>
<tr>
<td>Call Detail Records</td>
<td>Call detail records (CDRs), charging data records, event data records logged by telecoms and network switches</td>
<td>Billing, revenue assurance, customer assurance, partner settlements, bandwidth use</td>
</tr>
<tr>
<td>Clickstream Records</td>
<td>Web server, routers, proxy servers, ad servers</td>
<td>Usability analysis, digital marketing and customer journey</td>
</tr>
<tr>
<td>Mobile Application Data</td>
<td>SDKs embedded in mobile apps, application and server application logs</td>
<td>Mobile app usage, mobile app crashes, performance, latency, troubleshooting (stack trace) intelligence</td>
</tr>
<tr>
<td>Web Access Logs</td>
<td>Web access logs report every request processed by a web server</td>
<td>Web analytics reports for marketing</td>
</tr>
<tr>
<td>Web Proxy Logs</td>
<td>Web proxies log every web request made by users through the proxy</td>
<td>Terms of service and data leakage incidents</td>
</tr>
<tr>
<td>Wire Data</td>
<td>DNS lookups and records, protocol level information including headers, content and flow records</td>
<td>Performance and availability of applications, end user experiences, incident investigations, networks, threat detection, monitoring, compliance</td>
</tr>
</tbody>
</table>
Optimize Multi-Channel Marketing Campaigns

- Multi-channel analytics for web, mobile and 10,000+ store locations
- Real-time revenue insights, product mix and promotion effectiveness
- Marketing campaign optimization
Engagement Analytics

• Better understanding of customer interactions
• Real-time end-to-end tracking of transactions
• Improved customer satisfaction and experience
• Business visibility and performance awareness
• Tracking and understanding the root cause for website errors
Mobile Device Analytics

- Product adoption
- Users and clients
- Feature adoption
- User engagement
- Usage patterns
- Mobile devices
- Client dashboard
Social Sentiment Analytics

People Tweeting "Please Help" Over Time

Over the primary storm period, the frequency of people tweeting "Please Help" rises.

Sentiment about Critical Supplies

This chart shows sentiment (pos vs. neg) with a frequency of tweets related to critical services. Most were unhappy about power outage.

Fear before during and after Hurricane Harvey

This chart shows a breakdown of "fear" related tweets over time by gender.

Overall Fear by Gender

Male, female, and undetermined.
Online Service Engagement Analytics

Measure customer attention to specific areas of content

Analyze click trough's and how they navigate to CJ mall

Track and analyze mobile shopping customers in real time
Understand Digital Media Usage & Engagement

Business Analytics Use Case

- Analyze audio and podcast usage
- Accurately report royalty payments
- Faster identification of errors and abandonment
- Correlate weblogs with application performance data

Data sources: weblogs, audio/podcast logs, Akamai logs
Sacramento County Sheriff's Department: Intelligence-Led Policing

“The Splunk platform is critical to our Intelligence-Led Policing strategy. Our command group is now able to more clearly see trends in our crime statistics and take proactive action to address areas of concern and provide the best possible service to the public.”

– Senior IT Analyst and Application Team Lead Technical Services Division, Sacramento County Sheriff’s Department

- Integrated and visualized crime, management and corrections data
- Enhanced accountability, helping to reduce crime report backlog
- Supported proactive policing based on big data analytics
Analytics for service intelligence

applying multi-channel and cross-platform data to gain insight into the quality and impact of end-to-end constituent services
Visibility Across All Dimensions
of your application and technology stack

Apps
Servers
Network
Storage

Application-Based Silos
Zones of Virtualization
Private Cloud
Hybrid Cloud
Contextual Service Visualizations

- Visualize contextual inter-relationships across service delivery components
- Illustrate business and service activity using indicators aligned with strategic goals
- Drive decisions by monitoring service health against performance indicators
Organized View of Key Performance Indicators

- Organize and correlate KPIs to speed up investigations and diagnosis
- Compare performance over time and in real time to understand trends and identify systemic issues
- Enable broad and deep investigation with contextual drill-downs
Real-Time View of Service and KPI Health Scores

- Get early warning of emerging incidents with a heat map of service health and KPI scores, metrics, sparklines and alerts
- Drill down into service and entity details for in-depth triage
Applying Analytics for Service Visibility

Continuous Operational Visibility

CIO Scorecard

Enterprise Service Status

Major Incidents

Major Changes

Service One!

Service Health: 100%
Volume: 5.8k
Revenue: 2k
Incidents: 0
Changes: 0

Ground Visibility

Service Health: 100%
Volume: 5.8k
Revenue: 2k
Incidents: 0
Changes: 0

Product Tracking & Reporting

Service Health: 48.36%
Volume: 200
Revenue: 3
Incidents: 18
Changes: 60

Service Performance Measurement

Service Health: 60%
Volume: 5.8k
On-time Delivery: 63.24%
Incidents: 0
Changes: 0

Service Time Calculator

Service Health: 90%
Volume: 207
Revenue: 2k
Incidents: 0
Changes: 0

Manufacturing System

Service Health: 100%
Volume: 5.8k
Throughput: 63.24%
Container Util: 0
Incidents: 0
Changes: 0

@andimann © 2018 SPLUNK INC.
Applying Analytics for Service Intelligence

ER Patient Journey
Granular visibility and insights into all law enforcement activities
Maximize system uptime with predictive analytics
More effective resource allocation and faster officer response

“We connect the dots and see patterns once hidden in all the statistics. We’re improving services, operating smarter and giving the public greater returns on its tax dollars.”

– Sysadmin / Police Officer, Chandler AZ Police Dept.
Analytics for breach detection

using data and analytics to gain insight into exposures, data breaches, and unauthorized user behaviors
Security Analytics Enables Better Detection

Source: Forrester's Vendor Landscape: Security Analytics (SA)
Security Analytics Enables SOC Processes

- Monitoring and alerting
- Event correlation
- Alert triage
- Incident response
- Threat hunting

Image: By UMD-Eskin (Own work) [Public domain], via Wikimedia Commons
Security Analytics Accelerates Detection, Investigation & Response

▶ Use correlation and analytics to automate notable event detection
▶ Execute ad-hoc queries to find root causes and malicious actors
▶ Use automation to take actions and review their results
e.g. Phishing Search

▶ Detect typos, like company.com → campany.com

▶ Find misspelled subdomains for typo detection

▶ Detect suspicious subdomains, like company.com → company.yourithelpdesk.com
e.g. Increase in Pages Printed

- Search printer logs for potential resource abuse or data leakage
e.g. Authentication Against a New DC
Protecting Citizen Data Using Splunk Enterprise Security in the Cloud

“My top priority is to protect the citizens’ data. Making sure that these citizens can trust the government they have with the data that they have entrusted us with is our mission.”

– CISO, Fairfax County, Virginia

- Proactively supporting more than 50 county agencies and protecting citizens’ data
- Reducing security reporting from two weeks to real time
- Increasing focus on strategic initiatives by leveraging cloud services
Analytics for the Internet of Things

using analytics on devices and other ‘things’ to gain actionable intelligence about cloud-connected assets
A World of Connected Assets

**Transportation | Energy | Utilities | Building Management**

**Oil and Gas | Manufacturing**
- Sensors, Pumps, GPS, Valves, Vats
- Conveyors
- Pipelines, Drills, Transformers, RTUs, PLCs
- HMI, Lighting, HVAC, Traffic Management, Turbines, Windmills, Generators, Fuel Cells, UPS

**Industrial Data**

**Retail | Home | Consumer**
- Telemedicine | Connected Cars

**Internet of Things**
- Wearables, Home Appliances, Consumer Electronics, Gaming Systems, Personal Security, Set-Top Boxes, Vending Machines, Mobile Point of Sale, ATMs, Personal Vehicles
Challenges in IoT Landscape

- Diverse Protocols and Standards
- Complex Device to Cloud Architectures
- Security and Privacy
- Data Volume, Variety and Velocity
- Correlate Data Across Application/Infrastructure Silos
- Human to Machine Component
IoT and Industrial Machine Data

Industrial Assets
- Native Inputs
  - TCP, UDP, Logs, Scripts, Wire, Mobile
- SDKs and APIs
  - Java, JS, C#, Python, Ruby, PHP
- Modular Inputs
  - MQTT, AMQP, COAP, REST, JMS
- HTTP Event Collector
  - Token Authenticated Events
- Technology Partnerships
  - Kepware, ThingWorx, Cisco, Palo Alto

Consumer and Mobile Devices
- OT
- IT

External Lookups/Enrichment
- Asset Info
- Maintenance Info
- Data Stores

Real-Time Data Platform

Search | Alert | Visualize | Predict | Develop

Real-time feed
AWS and IoT

- Ingest data in real-time and at scale from AWS IoT Service
- Search, explore and analyze real-time and historical data with Splunk
- Correlate and enrich data from AWS IoT service with other data sources – application logs, mobile, databases and data from other IoT platforms
- Build web-applications using Splunk’s powerful application development, visualization, and machine learning frameworks
Use Cases

- Monitoring, Diagnostics
- Security, Safety & Compliance
- Preventative Maintenance

Real-Time Data Platform
Improve Customer Service, Reduce Costs by Increasing Locomotive Availability and Reliability

- Ingest and correlate sensor, diagnostic codes, geolocation data in real time to:
  - Gain insights into asset health, condition
  - Perform root cause analysis
  - Generate locomotive maintenance recommendations
Improving Water Quality

- Ingest data from water treatment systems, weather, SCADA, buoys, lab testing
- Monitor, measure water quality; identify factors impacting quality
- Identify sensor reading anomalies to replace/recalibrate
Managing Airfield Performance

- Real-time monitoring of aircraft turnaround process
- Tracking real-time metrics to manage airfield performance
- Increased on-time efficiency and aircraft predictability
Improving Passenger Experience

 Damascus

 Gain visibility into passenger flow
 Reduce congestion with improved ticket scan validation
 Monitor travel disruption to understand impact on operations
 Optimize staffing to improve passenger experience
Measuring New IT Architectures

modern approaches to service delivery incl. Site Reliability Engineering’ semantic logging, telemetry, observability
Rethinking and Improving How IT Operates

Traditional IT
- Brittle tools and integrations
- Obsession with “faults” and “traps”
- Focus on components parts
- Non-stop reactive break-fix
- Manual ops, one-offs, and heroes

Data Driven IT
- Robust data integrations
- Real-time insights from data
- Focus on the whole service
- ML and predictive analytics
- Automation engineering
Site Reliability Engineering

► A Durable Focus on Engineering
  • No more than 50% time on break-fix; excess ops work goes to backlog

► Pursuing Maximum Change Velocity Without Violating a Service’s SLO
  • “Error budget’ to allow for innovation and stability

► Monitoring and Emergency Response
  • Standardize and automate to reduce human impact; issues routed to backlog

► Engineer for rapid change
  • Real-time self-service provisioning; enable progressive deploy-fail-fix cycles

► Relentless Automation
  • Software-defined everything; “Automate yourself out of a job”

► Engineer for Efficiency and Performance
  • Build and test services for resilience; deficiencies go into application backlog

From Site Reliability Engineering, Betsy Beyer, Chris Jones, Jennifer Petoff, Niall Richard Murphy, O’Reilly Media, Inc., 2016
“In control theory, observability is a measure of how well internal states of a system can be inferred from knowledge of its external outputs.

The observability and controllability of a system are mathematical duals.”

Wikipedia
“Semantic Logging”

- You have no control over other systems events
- You have full control over events that YOU write
- Most events are written by developers to help them debug
- Some events are written to form an audit trail

Semantic Events are written explicitly for the gathering of analytics
Semantic Logging Best Practices
Log more than just Debugging Events

- Log anything that can add value when aggregated, charted or analyzed

Example Bogus Pseudo-Code:

```java
void submitPurchase(purchaseId)
{
    log.info("action=submitPurchaseStart, purchaseId=%d", purchaseId)
    //these calls throw an exception on error
    submitToCreditCard(...) 
    generateInvoice(...) 
    generateFullfillmentOrder(...) 
    log.info("action=submitPurchaseCompleted, purchaseId=%d", purchaseId)
}
```

- Graph purchase volume by hour, by day, by month.
- How long are purchases taking at different times of day, or days of the week?
- Are purchases taking longer than they did last month?
- Are my systems getting slower and slower, or are they ok?
- How many purchases are failing? Graph the failures over time.
- Which specific purchases are failing?
SREs Monitor Metrics and Events

Two distinct machine data sources that have been hard to integrate…until now

**Metrics**
- Numbers describing a particular process or activity
- Measured over intervals of time—i.e., *time series data*
- Common metrics sources:
  - System metrics (CPU, memory, disk)
  - Infrastructure metrics (AWS CloudWatch)
  - Web tracking scripts (Google Analytics)
  - Application agents (APM, error tracking)

**Events**
- Immutable record of discrete events that happen over time
- Come in three forms: plain text, structured, binary
- Common event sources:
  - System and server logs (syslog, journald)
  - Firewall and intrusion detection system logs
  - Social media feeds (Twitter…)
  - Application, platform and server logs (log4j, log4net, Apache, MySQL, AWS)

---

**Sample Metric**

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Metric Name</th>
<th>Value</th>
<th>Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1481050800</td>
<td>os.cpu.user</td>
<td>42.12345</td>
<td>hq:us-west-1</td>
</tr>
</tbody>
</table>

**Sample Log**

```
[29/Aug/2018 08:47:05:316503] "POST /cart.do?uid=84e8d742-a31d69&action=remove&&product_id=BS-2&JSESSIONID=SD6SAL4FF1ADFF9 HTTP 1.1" 200 2569 "http://www.buttercupenterprises.com/product.screen?product_id=BS-2" "Mozilla/5.0 (Intel Mac OS X 10_12_2) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/57.0.2987.133 Safari/537.36"
```
Dev isn’t “done” until the system provides data for Ops
Advanced analytics

advanced data techniques for IT incl. machine learning, operations analytics, anomaly detection, operational intelligence, predictive analytics, and data visualization
Unlock the Value of Data with Analytics

<table>
<thead>
<tr>
<th>Device Analytics</th>
<th>Security and Privacy</th>
<th>Transport Logistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High Frequency Analytics</th>
<th>Constituent Engagement</th>
<th>Performance Analytics</th>
<th>Predictive Analytics</th>
<th>Operational Intelligence</th>
</tr>
</thead>
</table>
Detect Patterns, Anomalies with Machine Learning

Preview Aggregate Thresholds

- Tuesday
- Wednesday
- Thursday
- Friday
- Saturday
- Sunday
- Monday

© 2018 SPLUNK INC.
Use ML to Highlight Clustered Events
Use ML to Forecast Time Series Data
Use ML to Detect Metric Anomalies
Baseline Trends to Adapt Thresholds

Use statistics to dynamically adapt KPI thresholds by time

Maintain and preserve learned thresholds to monitor KPI and service behavior
Detect and Predict Anomalies

Percentage of Time Anomalies were Detected: 3% (Expected <15%)

Trending AD Preview - KPI Value for Last 7 Days

Entities Analyzed: 57
Entities with Detected Anomalies: 33
Average Anomalies Per Entity: 16.3
Percentage of Time Anomalies were Detected: 5.8% (Expected <3%)
Percentage of Data Points with Anomalies: 7.3% (Expected <10%)

Cohesive AD Preview - Top 5 Entities With Most Anomalies for Last 7 Days

- web-server1
- web-server2
- web-server3
- Detected Anomaly
Learn What’s Normal and Abnormal

Baseline normal operations and alert on anomalous conditions

Identify abnormal trends and patterns in KPI data
Use ML to Detect Numerical Outliers
But Good Data Is Not Enough
## Find The Value In The Data

<table>
<thead>
<tr>
<th>Planning</th>
<th>Development</th>
<th>Build</th>
<th>Verification</th>
<th>Deployment</th>
<th>Post-Deploy</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 (0%)</td>
<td>100 (0%)</td>
<td>94.74 (-5.3%)</td>
<td>100 (0%)</td>
<td>100 (0%)</td>
<td>100 (0%)</td>
</tr>
<tr>
<td>160 stories</td>
<td>0 in progress</td>
<td>100% success</td>
<td>100% success</td>
<td>364 deploys</td>
<td>0 CFDs</td>
</tr>
<tr>
<td>100 stories</td>
<td>95 complete</td>
<td>8.8 MTTB</td>
<td>3.95 MTTT</td>
<td>0.54 success</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 ticket</td>
<td>2.489 MTTR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30 points/dev</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94 (-6%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.12 days</td>
<td></td>
</tr>
</tbody>
</table>
Find The Value In The Visualization
## Find The Failure In The Data

<table>
<thead>
<tr>
<th>Planning</th>
<th>Development</th>
<th>Build</th>
<th>Verification</th>
<th>Deployment</th>
<th>Post-Deploy</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 (0%)</td>
<td>100 (0%)</td>
<td>0 (-100%)</td>
<td>100 (0%)</td>
<td>100 (0%)</td>
<td>100 (0%)</td>
</tr>
<tr>
<td>160 stories</td>
<td>0 in progress</td>
<td>35% success</td>
<td>100% success</td>
<td>364 deploys</td>
<td>0 CFDs</td>
</tr>
<tr>
<td>100 stories</td>
<td>95 complete</td>
<td>8.8 MTTB</td>
<td>3.95 MTTT</td>
<td>0.54 success</td>
<td></td>
</tr>
<tr>
<td>1 ticket</td>
<td></td>
<td>2.489 MTTR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 points/dev</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 94 (-6%)
- 15.12 days
Find The Failure in the Visualization
How About Now?
How About Now?
How About Now?
And Data Silos Are Not Enough
Shared Data Helps Find and Fix Issues Faster

Common alerting notify devs and ops as soon as a problem arises

Developers can search and visualize production logs and tools —without production access

Real-time data sharing shows error rate in production and impact of pushing new builds
Analytics At Every Phase of The DevOps Lifecycle

- Plan
  - time to deliver
  - idea to cash
  - ROI

- Code
  - code volume
  - commit volume
  - release speed

- Build
  - build speed
  - failure rates
  - manual builds

- Test/QA
  - test volume
  - code coverage
  - exception counts

- Config
  - remediation time
  - code quality
  - access rates

- Stage
  - performance
  - latency
  - scalability

- Release
  - response time
  - uptime/availability
  - resource usage

- Monitor
  - revenue
  - signups
  - cust. sat.
Specific Data For Each Stakeholder

- Biz
  - time to deliver
  - idea to cash
  - ROI

- PMO
  - code volume
  - commit volume
  - release speed

- Dev
  - build speed
  - failure rates
  - manual builds

- Build
  - remediation time
  - code quality
  - access rates

- QA
  - test volume
  - code coverage
  - exception counts

- Sec
  - performance
  - latency
  - scalability

- Stage
  - response time
  - uptime/availability
  - resource usage

- Ops
  - revenue
  - signups
  - cust. sat.
Shared Data for Multiple Stakeholders

- process times
- team efficiency
- unplanned work
- build speed
- failure rates
- manual builds
- remediation time
- code quality
- access rates
- response time
- uptime/availability
- resource usage

- time to deliver
- idea to cash
- ROI
- code volume
- commit volume
- release speed
- test volume
- code coverage
- exception counts
- performance
- latency
- scalability
- revenue
- signups
- cust. sat.
Shared Data for Multiple Stakeholders

- process times
- team efficiency
- unplanned work
- build speed
- failure rates
- manual builds
- remediation time
  - code quality
  - access rates
- response time
  - uptime/availability
  - resource usage

- time to deliver
- idea to cash
- ROI
- code volume
- commit volume
- release speed
- test volume
- code coverage
- exception counts
- performance
- latency
- scalability
- revenue
- signups
- cust. sat.
- build speed
- fail rate
- manual builds
- code volume
- commit volume
- release speed
- test volume
- code coverage
- exception counts
- performance
- latency
- scalability
- revenue
- signups
- cust. sat.
Specific Data For Each Stakeholder

- **Biz**
  - time to deliver
  - idea to cash
  - ROI

- **PMO**
  - code volume
  - commit volume
  - release speed

- **Dev**
  - build speed
  - failure rates
  - manual builds

- **Build**
  - remediation time
  - code quality
  - access rates

- **QA**
  - test volume
  - code coverage
  - exception counts

- **Sec**
  - performance
  - latency
  - scalability

- **Stage**
  - response time
  - uptime/availability
  - resource usage

- **Ops**
  - revenue
  - signups
  - cust. sat.
Specific Data For Each Stakeholder

- process times
- team efficiency
  - unplanned work
- build speed
- failure rates
- manual builds
- remediation time
  - code quality
  - access rates
- response time
  - uptime/availability
  - resource usage

Biz ➔ PMO ➔ Dev ➔ Build ➔ QA ➔ Sec ➔ Stage ➔ Ops ➔ Biz

- time to deliver
- idea to cash
- ROI
- code volume
- commit volume
- release speed
- test volume
- code coverage
- exception counts
- performance
  - latency
  - scalability
- revenue
- signups
- cust. sat.
City of Los Angeles: Sharing Security Intel Across 40+ Agencies

“As the number and sophistication of risks increase, our cloud-based Splunk solution levels the playing field by making our security team more effective.”

– Chief Information Security Officer, City of Los Angeles

- Prompt responses to cyberthreats with real-time situational awareness of citywide infrastructure
- Timely intelligence sharing with local, state and national law enforcement
- Reduced Total Cost of Ownership
Data-driven automation

coupling data analytics with process automation to surface actionable data points, make real-time decisions, and act to remediate
Automation in the SDLC

Detect Patterns, Anomalies with Machine Learning
Learn What’s Normal and Abnormal

Baseline normal operations and alert on anomalous conditions

Identify abnormal trends and patterns in KPI data
Integrate ML With Existing Workflows

- Automatically initiate defined incident and remediation responses
- Integrate with ServiceNow to create tickets and accelerate triage
Use Adaptive Responses to Security Breaches

- **Stream Capture**
  - Category: Information Gathering
  - Task: create
  - Subject: network.capture
  - Vendor: Splunk

- **Nbtstat**
  - Runs the nbtstat command
  - Category: Information Gathering
  - Task: scan
  - Subject: device
  - Vendor: Operating System

- **Nslookup**
  - Runs the nslookup command
  - Category: Information Gathering
  - Task: scan
  - Subject: device
  - Vendor: Operating System

- **PAN: Block Traffic**
  - Category: Device Control
  - Task: update
  - Subject: network
  - Vendor: PAN

- **PAN: Quarantine Host**
  - Category: Device Control
  - Task: update
  - Subject: network
  - Vendor: PAN

- **PAN: Tag to Dynamic Address Group**
Use Orchestration for Complex Workflow Processes
Integrate Data, Chat, Bots for Collaborative Troubleshooting and Triage (aka ‘ChatOps’)

Hey Team, our new release has an issue with the build service – check it out in ITSI kanwu02:8000/en-us/app/ITSI/Building-Datagen?earliest=3600&latest=now

Can you see what I am talking about?

Hey Dmitri, we see it. Looks like a package has gone missing. I’ll get that fixed straight away!

Great, thanks Julie. I’ll keep checking ITSI for the update to go through.

Dmitri S. 

Julie Q.

Dmitri S. 

Julie Q.

Julie Q.

@FixBot

Reload complete: Buttercup-Games-Repo 0x000000

Build complete: Buttercup-Games-Repo 0x000000
Automate Everything!

Select actions to run.

+ Add New Response Action

Category: All

Search

AWS: Start Instance
Category: Permissions Control | Task: allow | Subject: endpoint.server | Vendor: AWS

Verify latest patch status
Verify latest patch status
Category: Information Gathering | Task: update | Subject: endpoint.workstation | Vendor: TAN

Dominos: Order Pizza
Category: Device Control | Task: create | Subject: splunk.event | Vendor: Dominos

Endpoint: Check for new Hash
Endpoint: Check for new Hash
Category: Device Control | Task: create | Subject: endpoint | Vendor: Generic

Endpoint: Logout User
Category: Device Control | Task: allow | Subject: endpoint | Vendor: Generic
Wrap-up

Q&A, Summary, Close
Thank You!

Questions?