OBSERVABILITY WORKSHOP

w/ OpenTracing + Jaeger + Prometheus
@SREcon 2019

Content: Kevin Crawley @notsureifkevin
Assistance: Liz Fong-Jones @lizthegrey
Inspired by: Sandy Walsh & Countless Others
AUDIENCE POLLING

There are no wrong answers
WE HAVE STAGING ENVIRONMENTS
I TRUST MY STAGING ENVIRONMENT
MY TEAM USES A CONTINUOUS INTEGRATION SYSTEM (GATING OF PULL REQUESTS)
MY TEAM USES AN AUTOMATED DEPLOYMENT SYSTEM
INTEGRATION TESTS ARE TOO SLOW
TESTING IS THE RESPONSIBILITY OF QUALITY ASSURANCE
MY TEAM USES A MONITORING AND ALERTING SYSTEM
MY TEAM HAS A STRONG DEVOPS CULTURE
JUST KIDDING ... THERE ARE WRONG ANSWERS

• **Why Integration Testing is (mostly) useless**
• **What observability means**
• **The Workshop**
• **Why you shouldn’t go at this alone**

Work: @instanahq / Personal: @notsureifkevin
Kevin Crawley
Developer Relations @ Instana
Principle SRE / Architect @ Single Music
Nashville, TN USA
Professional Background

- Docker Captain
- Local Meetup Organizer (Docker / GoLang / Serverless)
- DevOpsDays Nashville Organizer
- 15+ years in software engineering
  - Banking / Finance Organizations
  - Metrics / Analytics
  - Distributed Systems Engineer
- 5+ years deploying containerized solutions
Social Media @notsureifkevin
(twitter, github, Instagram, linkedin, etc)

Random Personal (Nerdy) Facts:

- Fitness / Data Nerd (Instrumented myself + bike and rode 100 miles)
- SciFi Nerd (Watched Stargate, The Expanse, Firefly n times)
- Video Gaming Nerd (I’ve beaten every* Final Fantasy game, currently playing Octopath Traveler)
- Board Game Nerd too! Invite me to come play a game tonight!
AGILE MANIFESTO
PUBLISHED IN 2001

• Test Driven Development rises in popularity
• Shortly thereafter Behavior Driven Development begins its ascent
• Product owners bemoan how long it takes to change / release software
RAVEN TOOLS HAD ONLY 6% CODE COVERAGE

• After 2 years of diligent effort and working with our development teams
• We increased our coverage to nearly 50%
• Less bugs were shipped

Work: @instanahq / Personal: @notsureifkevin
PRODUCTION INCIDENTS WERE REDUCED BUT WE NEVER STOPPED HAVING TO RELEASE HOTFIXES OR ISSUE ROLLBACKS

IN FACT, WE STOPPED SHIPPING SOFTWARE AS OFTEN BECAUSE OF HOW DIFFICULT CHANGING EXISTING CODE HAD BECOME

Work: @instanahq / Personal: @notsureifkevin
WE EVEN BUILT OUR OWN IMPLEMENTATION OF STATSD TO INCLUDE CONTEXTUAL DATA

HTTPS://GITHUB.COM/RAVENTOOLS/EVENTSD
(DON'T USE THIS, I'M SERIOUS, I WROTE IT 5 YEARS AGO)

Work: @instanahq / Personal: @notsureifkevin
UNIT TESTS ARE LIKE PUTTING YOUR CODE INTO A TAR PIT
Rocky Tar Pit enters the battlefield tapped.

\( \Rightarrow \) Sacrifice Rocky Tar Pit: Search your library for a Swamp or Mountain card and put it onto the battlefield. Then shuffle your library.
THIS IS YOUR CODE IN STAGING
AND HERE IS YOUR CODE RUNNING IN PRODUCTION
SO WHY IS PAGERDUTY STILL GOING OFF AT 03:00 EVEN WITH 94.6% COVERAGE?
WHEN PRODUCTION IS BROKEN WHO WAITS 1+ HR FOR INTEGRATION TESTS?

MOST ORGANIZATIONS PUSH ONE OR MORE HOTFIXES AFTER EVERY RELEASE

Work: @instanahq / Personal: @notsureifkevin
YOU ARE ALREADY TESTING IN PRODUCTION
( WHETHER YOU LIKE IT OR NOT )

THERE IS NO SUCH THING AS A BUG-FREE SYSTEM
SO CHOOSE ONE:

• THEY SEE THE BUGS AND YOU KNOW ABOUT IT

• YOU WAIT UNTIL THEY TELL YOU ABOUT THE BUGS
  ( ON TWITTER )

Work: @instanahq / Personal: @notsureifkevin
WHAT SHOULD WE BE UNIT / FUNCTIONALLY TESTING?

• **Critical Systems Code / Pathways Example:** Can the user transfer money from Account A into Account B

• **Open Source Libraries** which other users depend on

Work: @instanahq / Personal: @notsureifkevin
THE #1 PROGRAMMER EXCUSE FOR LEGITIMATELY SLACKING OFF:

"MY CODE'S COMPIRING"

TESTING

HEY! GET BACK TO WORK!

Waiting for integration tests to pass

OH. CARRY ON.
UNIT TEST COVERAGE AS A METRIC OR KPI IS USELESS ... INSTEAD

• **KEEP RECORD OF HOW OFTEN YOU BREAK PRODUCTION (REGRESSIONS) AND TEST THOSE BITS OF CODE**

• **TRACK THE RESPONSIVENESS OF YOUR SYSTEM**

• **TRACK HOW LONG IT TAKES TO PUT A CHANGE REQUEST INTO PRODUCTION**

Work: @istanahq / Personal: @notsureifkevin
SRE DRIVEN DEVELOPMENT (SDD)

“Test in Production ... all the time”
- Charity Majors (Honeycomb.io) [paraphrased]
OBSERVABILITY

WHAT DOES THAT WORD EVEN MEAN?
OBSERVABILITY

- Monitoring
- Instrumentation
- Structured logging (tracing)
- Alerts
- Dashboards

- Not just tooling
- Similar to how DevOps is a mindset
- Not treat services like a black box
- More context around events

Work: @instanahq / Personal: @notsureifkevin
ENABLE NON-SERVICE OWNERS TO INVESTIGATE

- Deployments
- Performance
- Behavior

Work: @instanahq / Personal: @notsureifkevin
WHAT IS MONITORING

“Tasking humans or machines to read and interpret the data gathered via telemetry”
WHAT IS TELEMETRY

“THE PROCESS OF RECORDING AND TRANSMITTING DATA FROM AN INSTRUMENT”
HISTORICALLY, WHO HAS OWNED MONITORING SYSTEMS?

OPERATIONS (NOT DEVELOPERS)
WHAT DO MOST OPERATIONS TEAMS MONITOR?

- CPU
- Disk Space
- Network Utilization
- Memory Usage
- Latency
- Status (4xx/5xx)
- Queue Size

Work: @instanahq / Personal: @notsureifkevin
HOW DO WE MEASURE THE INTERNALS OF AN APPLICATION SERVICE?

WE MUST EMIT SIGNALS FROM WITHIN OUR APPLICATIONS
WHAT QUESTIONS ARE WE ASKING?

BEHAVIORS

- NUMBER OF RETRIES / BACKOFFS
- FAILING BACK TO A DEFAULT
- TOP-LEVEL EXCEPTIONS

Work: @instanahq / Personal: @notsureifkevin
WHAT OTHER QUESTIONS ARE WE ASKING?

Reactions

• This is Weird
• This should never Happen
• Try { thisShouldNeverBreak() } catch { howDidWeGetHere() }
Push Model

Also known as the "agent-based model", Jaeger, Instana, Datadog, Dynatrace, etc. all use this model.
MONITORING FAILURES SHOULD NOT IMPACT PRODUCTION OPERATIONS

• Monitoring should emit non-blocking signals

• Only backlog intermittent failures worth of data at best

Work: @instanahq / Personal: @notsureifkevin
Pull Model

Also known as the Prometheus model

Telemetry Service 15-60s

HTTP/gRPC

/metrics

Service X

/metrics

Service Y

/metrics

Service Z

Work: @instanahq / Personal: @notsureifkevin
METRICS ARE CONSIDERED

• **UNRELIABLE (UDP HAS NO GUARANTEE, TCP HAS HIGH OVERHEAD)**
• **NON-MISSION CRITICAL**
• **LIGHTWEIGHT (SMALL PAYLOADS)**
• **LITTLE-TO-ZERO CONTEXT**

Work: @instanahq / Personal: @notsureifkevin
WHAT METRICS MATTER

• Counters    ### 5xx errors
• Rate        ### 5xx errors/Min
• Gauge       ### CPU Load
• Histogram   ### Rolling Avg / Quartiles

Work: @instanahq / Personal: @notsureifkevin
DISTRIBUTED TRACING

• Also known as Distributed Structured Logging
• Larger Payload
• Rich Context

https://w3c.github.io/trace-context

Work: @instanahq / Personal: @notsureifkevin
HOW DOES DISTRIBUTED TRACING WORK?

- **Custom Headers / Metadata**
  - **X-B3-TraceId**: 128 or 64 lower-hex encoded bits
  - **X-B3-SpanId**: 64 lower-hex encoded bits
  - **X-B3-ParentSpanId**: 64 lower-hex encoded bits
  - **X-B3-Sampled**: Bool
  - **X-B3-Flags**: “1” indicates DEBUG

Work: @instanahq / Personal: @notsureifkevin
DASHBOARDS
(I LIKE DASHBOARDS, I’M A DATA NERD)

- Charts
- Graphs
- Heat Maps
- High/Low Watermarks
- Percentiles (P50/P99)
- Detectors

- Traffic Signals
- Alerts
- Integrations
  - Slack
  - PagerDuty
  - Ticketing Systems

Work: @instanahq / Personal: @notsureifkevin
HOW DOES DISTRIBUTED TRACING GIVE ME MORE DASHBOARD?

LET’S TAKE A LOOK AT MICROSERVICE APPLICATION WHICH HAS BEEN INSTRUMENTED BY A DISTRIBUTED TRACING SOLUTION

Work: @instanahq / Personal: @notsureifkevin
• **Featured on Shopify**
  (https://apps.shopify.com)

• **We built Single to give artists and labels the ability to easily distribute their music in their own stores**

• **Soft launched in 2018 – has since expanded to over 20 microservices with zero additional engineering resources**

Work: @instanahq / Personal: @notsureifkevin
Latency Heatmap

Top Statements

```
select shopifyatt0_id as id1_id, shopifyatt0_active as active...
5ms
```

```
select boostlink0_id as id1_id, boostlink0_account_id as acc...
2ms
```
Exponential Backoff
Exponential rise in latency
SingleMusic - Production

Total Calls: 2,580
Mean Latency: 2ms
Error Rate: 0.00%

Total Calls vs Mean Latency

Statement

```
1 SELECT boostlinkc0_id as id1_2, boostlinkc0_account_id as account_2_2, boostlinkc0_created_at as created_3_2, boostlinkc0_enabled as enabled4_2, boostlinkc0_end_date as end_date5_2, boostlinkc0_id as id6_2, boostlinkc0_start_date as start_date7_2, boostlinkc0_title as title8_2, boostlinkc0_url as url9_2 FROM boostlink_campaign boostlinkc0
2 INNER JOIN boostlink_campaign_to_shopify_attachment attachment1 on boostlinkc0_id=attachment1.boostlink_campaign_id
3 INNER JOIN shopify_attachment shopifyatt2 on attachment1.shopify_attachment_id=shopifyatt2_id
4 WHERE shopifyatt2_product_id? AND shopifyatt2_product_variant_id? AND shopifyatt2_active? AND boostlinkc0_state='ACTIVE' AND boostlinkc0_enabled=1
```
COOL SCREENSHOTS

How do I convince my manager to budget for more observability?
ESTABLISH AN ON-CALL ROTATION

Always assume the teams will change — Include Developers and Operations (and your manager)
WHEN THE FOCUS IS ON OBSERVABILITY

PERCENTAGE OF ENTIRE APPLICATION
VS
PERCENTAGE OF CODE
"I knew this organization was full of siloes. I just didn't think they'd be so heavily defended!"

SILOS ARE EASY TO DEFEND AND DIFFICULT TO ATTACK
YOUR JOB IS COMPLETING THE PRODUCT, FIXING THE BUGS, AND SHIPPING.

Chuck Jazdzewski

"[…] shipping a product feels good, like when someone stops hitting you. Your job is completing the product, fixing the bugs, and shipping. If bugs need fixing, fix them. If documentation needs writing, write it. If code needs testing, test it. All of this is part of shipping. You don’t get paid to program, you get paid to ship. Be good at your job." - Chuck Jazdzewski

Work: @instanahq / Personal: @notsureifkevin
"YOUR JOB IS TO SHIP THE PRODUCT AND ENSURE YOU AND YOUR SRE (OR OPERATIONS) TEAMS CAN MONITOR IT."

Kevin Crawley
WANT TO SEE AN ISSUE FIXED?

INSTRUMENT AND ALERT ON IT

- LATENCY
- ERROR RATE
- HEALTH CHECKS

Work: @instanahq / Personal: @notsureifkevin
MAKE WARN OR ERROR LEVEL LOG MESSAGES GENERATE A METRIC

OBSERVABILITY TOOLING ENABLES THIS CAPABILITY

Work: @instanahq / Personal: @notsureifkevin
<table>
<thead>
<tr>
<th>Log Message</th>
<th>Log Level</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Download Code Redeemed event received with no purchaseid: ()</td>
<td>WARN</td>
<td>98</td>
</tr>
<tr>
<td>somebody trying to load short link that doesn't exist with slug ()</td>
<td>WARN</td>
<td>68</td>
</tr>
<tr>
<td>Shop is inactive: ()</td>
<td>WARN</td>
<td>55</td>
</tr>
<tr>
<td>Error processing line item () in order () (transaction () ) - No album ID present in search where not physical or no physical album found. Check metafields on p...</td>
<td>ERROR</td>
<td>25</td>
</tr>
<tr>
<td>Object is not serializable! ()</td>
<td>ERROR</td>
<td>18</td>
</tr>
<tr>
<td>Unable to fetch metafields for (): ()</td>
<td>WARN</td>
<td>17</td>
</tr>
<tr>
<td>No image of size () found for image file set ID ()</td>
<td>WARN</td>
<td>17</td>
</tr>
<tr>
<td>{} is trying to create an artist that already exists with name: {}</td>
<td>ERROR</td>
<td>7</td>
</tr>
<tr>
<td>Failed to acquire lock for ()</td>
<td>WARN</td>
<td>4</td>
</tr>
</tbody>
</table>
REALTIME DASHBOARDS ESTABLISH THE “BEAT” OF YOUR APPLICATIONS

- **Run integration tests against production**
- **Determine the baseline “beat” of your app — not a “flat-line”**
- **Start with basic health checks**
  - Expand to connectivity checks
  - Then expand to functionality checks

Work: @instanahq / Personal: @notsureifkevin
DETECT PERCENTILES OF ANOMALOUS BEHAVIOR

Anomaly happens on 10% of operations

vs.

Anomaly happens only when the user registers on a Tuesday
DEPLOY YOUR BUILDS USING BLUE/GREEN THEN GRADUATE TO CANARY TESTING

• **In both solutions you are operating two versions of software simultaneously**

• **Tag events so you can see these deployments occurring in your monitoring solution**

Work: @instanahq / Personal: @notsureifkevin
BLUE/GREEN DEPLOYMENTS

- Fast Rollouts
- Move users from one version to the next without interruption

Work: @instanahq / Personal: @notsureifkevin
CANARY DEPLOYMENTS

• **Slow Rollouts**

• **Observe performance of a small subset of replicas and increase deployment (5% -> 25% -> 50% -> 100%)**

Work: @instanahq / Personal: @notsureifkevin
ALERTING SHOULD SURFACE MEANINGFUL DATA

• Teams / Owners
• Service Experts
• Dashboards
• Metrics & Logs
• PagerDuty
• Slack Channels
• Tickets

• Tickets / History
• Cluster / Region / AZ
• Container Image Prefix
• Service DNS
• Service Dependencies
• Public Endpoints
• Documentation / FAQs

Work: @instanahq / Personal: @notsureifkevin
New issue observed by Instana

Complete drop in the number of requests
This can be a sign of a problem on this service or its upstream

* Relative change in value: **-100%**
* Confidence: **95.00%**

Entity
GET /campaign/for-fulfillment (Endpoint)

Open in Instana

New incident observed by Instana

Sudden drop in the number of requests
This can be a sign of a problem on this service or its upstream

* Relative change in value: **-98%**
* Confidence: **95.00%**

Entity
single-service-boostlink (Service)

Open in Instana
PROMOTE A CULTURE OF VISIBILITY AND LEARNING

• INCIDENTS ARE COLLABORATED ON AND SHARED
• SCRIPTS AND TOOLS ARE PUT IN A SHARED REPOSITORY (NOT A HOME DIRECTORY)
SUFFICIENTLY ADVANCED MONITORING IS INDISTINGUISHABLE FROM TESTING ...

ED KEYES
SITE RELIABILITY ENGINEER - GOOGLE // 2008
“I THINK WE’LL STICK WITH THE OLD WAY OF DOING THINGS, WE NEED MORE TEST COVERAGE.”
“I THINK OUR SYSTEM COULD BENEFIT FROM MORE OBSERVABILITY.”
THAT’S GOOD TO KNOW, LET’S GET TO WORK!

https://bit.ly/2UQISz4
THIS LOOKS LIKE A LOT OF WORK

It is, fret not though, there are vendors who will do it for you!

Did you build your own CI/CD system or your own PagerDuty?
It doesn’t have to be complicated

Many vendors have simplified it to a simple script, single Helm chart install, or Docker Command where much of the instrumentation has been fully automated.
WORK SMARTER

LESS CODE IS BETTER

AUTOMATE EVERYTHING

OUTSOURCE THIS TOOLING
THANK YOU FOR YOUR TIME AND ATTENTION

WE (BOTH) WORK FOR DISTRIBUTED TRACING AND APPLICATION MONITORING VENDORS AND WE’LL DO THE WORK FOR YOU!

HTTPS://INSTANA.COM/SRECON
HTTPS://HONEYCOMB.IO

Work: @instanahq / Personal: @notsureifkevin