Tracking Logs at Zillow with Lookups & JIRA

Seth Thomas, Jon Wentworth
September 27 | 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.

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. © 2017 Splunk Inc. All rights reserved.
About Zillow

- Zillow Group operates the largest real estate network on the web.
- Zillow Group Rentals is the largest rental network on the web, composed of the millions of rental shoppers on Zillow, Trulia, HotPads and MyNewPlace.
- 4 out of 5 U.S. homes have been viewed on Zillow.
- Zillow Group captured nearly two-thirds of the total market share for the online real estate category.
- Zillow has data on more than 110 million U.S. homes, with Zestimates and Rent Zestimates on more than 100 million U.S. homes.
- More than 70 million homes on Zillow have been updated by our community of users.
- 166 million average monthly unique users visited Zillow Group brands’ mobile apps and websites
- mobile apps and websites reached an all-time high of nearly 180 million unique users
- More than 36 million rental visitors come to Zillow Group’s rental sites and apps each month
Zillow Group Splunk

How does Zillow use Splunk?

- Operations Center & Systems Engineering
- Customer Service
- Product & Engineering Teams
- Performance Team
- SEO Team
- Native Mobile Apps (iOS, Android) with MiNT
Our Problem

Our services can produce a lot of errors, which do we care about, how do we track them, what are the metrics, how well is the process working?
Evolution of Error Event Tracking

Top 10 – quantifies noise
Leaderboard – shows movement between 2 periods
Assigning Identity – is there a JIRA ticket

- Creates an Error Dictionary
- Assigns accountability
- Provides context upon regression
- Maps to all environments (prod & pre-prod)
- Allows for quality gates between code promotions
ERROR % bug Tracking
Defined Goal 80%

Tracking Progress with Metrics

80.49 %
This is a simple project IF you have the required rights and permissions.

- JIRA requires access to API
- Splunk access to Transforms and Apps

No cost, doesn’t index data, only populates a lookup.csv

Who benefits? – Everyone!

- ERROR messages are assigned to teams, tracked and triaged
- Alarms triggered for new ERROR events
- Regressions are caught early
Zillow Operations Center responsibilities

▶ Monitor Site Health
  • Routine Eyes on Glass
  • Code Deployments
  • Outages

▶ Protect and Defend
  • Site Error Rates
  • Log Level Rates
  • Perception of Unhealthiness
Our traditional tools before Splunk

<table>
<thead>
<tr>
<th>terminal</th>
<th>Graphite</th>
<th>log rollups</th>
</tr>
</thead>
<tbody>
<tr>
<td>terminal</td>
<td>Graphite</td>
<td>log rollups</td>
</tr>
</tbody>
</table>

ZOC
Limitations

- Graphs only show basic volume trends
- Strong command line tool to tail logs meaningfully
- Log rollups lack time dimension

Is that error new? Intro'd with release? Regression?
Who owns that?
What just happened?
Needed a solution to turn this into this... automatically
Java service error log

- 2017-07-12T18:14:50.205813+00:00 WARN [com.zillow.web.pages.myzillow.SavedSearchEdit] no saved search available for XXXXXXXXXXX
- 2017-07-12T18:14:50.661995+00:00 ERROR [com.zillow.db.InnerPool] Connection hard closed due to exception:java.sql.SQLException: Invalid state, the Connection object is closed. src:{ call dbo.XXXXXXXXXX(XXXXXXXXXX) } on jdbc:jtds:sqlserver://XXXXXXXXXXX
- 2017-07-12T18:14:50.662739+00:00 ERROR [org.springframework.transaction.support.TransactionTemplate] Application exception overridden by rollback exception
Java service error log

- Timestamp
- Severity level
- Class – hierarchical code template
- Message – verbose return

- 2017-07-12T18:14:50.205813+00:00 WARN [com.zillow.web.pages.myzillow.SavedSearchEdit] no saved search available for XXXXXXXXXX

- 2017-07-12T18:14:50.661995+00:00 ERROR [com.zillow.db.InnerPool] Connection hard closed due to exception:java.sql.SQLException: Invalid state, the Connection object is closed. src:{ call dbo.XXXXXXXXXX(XXXXXXXXXX) } on jdbc:jtds:sqlserver://XXXXXXXXXX

- 2017-07-12T18:14:50.662739+00:00 ERROR [org.springframework.transaction.support.TransactionTemplate] Application exception overridden by rollback exception
Lexicon

**JIRA**

A commercial issue tracking product, developed by Atlassian.

- JIRA is a component of Zillow's code development tool suite

- Provides bug tracking, issue tracking, and project management functions.
Lexicon

**jirarest**

Add-on used to query JIRA's API

| jirarest jqlsearch "LogTrackingEnabled=Yes"

→ Using https://github.com/firebus/splunk-jira

Version 2.1 handles auto-pagination for production-scale Jira implementations, jirarest command

**Not** using official Splunk version https://splunkbase.splunk.com/app/1438/

Provides only jira command to live query JIRA REST API. Deprecated jirarest commands.
Lexicon

lookup

| lookup <lookup-table-name> <lookup-field1> AS <event-field1>, <lookup-field2> AS <event-field2> OUTPUTNEW <lookup-destfield1> AS <event-destfield1>, <lookup-destfield2> AS <event-destfield2>

| lookup update=true MyLookup msg, cls, java_svc OUTPUT bug java_svc cls lvl msg Priority
For rollups, regex has its limitations...

```
rex field=msg mode=guid "s/\S+@\S+.\w+/@\S+.\w+/g" | rex field=msg mode=guid "s/token = \S+.\w+/token = */g" | rex field=msg mode=guid "s/screenname=.\w+/screenname=/g" | rex field=msg mode=guid "s/(?s)(.\w+)/g" | rex field=msg mode=guid "s/\d+/g" | rex field=msg mode=guid "s/\d+/g" | rex field=msg mode=guid "s/zws-id=[^&].\w+/zws-id=/g"
```

| eval msg=substr(msg,1,200) |
| cluster t=0.8 labelonly=t field=msg |
Lexicon

cluster
Step by Step

1. Create and fill fields in JIRA tickets

1. Search JIRA with jirarest to populate lookup file

1. Enable wildcard searching - transforms.conf

1. Create searches against fields using lookups

1. Enhance with evals and conditionals
1. Create and fill fields in JIRA tickets

- LogTrackingEnabled
- LogTrackingMessage
- LogTrackingClass
- LogTrackingLevel
- LogTrackingService

![Image of JIRA ticket fields]

- LogTrackingEnabled: Yes
- LogTrackingMessage: no saved search available for *
- LogTrackingLevel: WARN
2. Search JIRA with jirarest to populate lookup file

```
| jirarest jqlsearch "LogTrackingEnabled=Yes"
| eval bug=Key
| eval msg=LogTrackingMessage
| eval cls=LogTrackingClass
| eval lvl=LogTrackingLevel
| eval java_svc=LogTrackingService
| table bug msg cls lvl java_svc "Suppress Alarms"
| fillnull value="NULL"
| outputlookup MyLookup.csv
```
2. Search JIRA with jirarest to populate lookup file
2. Search JIRA with jirarest to populate lookup file

```bash
| jirarest jqlsearch "LogTrackingEnabled=Yes"
| eval bug=Key
| eval msg=LogTrackingMessage
| eval cls=LogTrackingClass
| eval lvl=LogTrackingLevel
| eval java_svc=LogTrackingService
| table bug msg cls lvl java_svc "Suppress Alarms"
| fillnull value="NULL"
| outputlookup MyLookup.csv
```
3. Enable wildcard searching - transforms.conf

[MyLookup]
filename = MyLookup.csv
case_sensitive_match=false
match_type = WILDCARD(msg)
4. Create searches against fields using lookups

\[ \text{<base_search> lvl=ERROR | lookup update=true MyLookup msg, cls, java_svc OUTPUT bug | where bug=\"ZOS-1190\"} \]
Step by Step

4. Create searches against fields using lookups

<base_search> lvl=ERROR | lookup update=true MyLookup msg, cls, java_svc OUTPUT bug | where bug="ZOS-1190"

<table>
<thead>
<tr>
<th>bug</th>
<th>lvl</th>
<th>msg</th>
<th>cls</th>
<th>java_svc</th>
</tr>
</thead>
</table>
4. Create searches against fields using lookups

<base_search> lvl=ERROR | lookup update=true MyLookup msg, cls, java_svc OUTPUT bug | where bug="ZOS-1190"

com.zillow.service.user.auth.LoginException: Not signed in
5. Enhance with evals and conditionals

```
<base_search> lvl=ERROR
| lookup update=true MyLookup msg, cls, java_svc OUTPUT bug "Suppress Alarms"
| eval msg=substr(msg,1,200)
| cluster t=0.8 labelonly=t field=msg
| eval cluster_by = if(isnotnull(bug), bug, cluster_label)
| eventstats first(bug) as bug, first(msg) as msg by cluster_by
| eval bug = if(like('Suppress Alarms','"Yes"'), "ALARMS OFF -.bug, bug)
| eval bug = if(isnotnull(bug), bug." - ".msg, msg)
| timechart count by bug usenull=f useother=f
...
5. Enhance with evals and conditionals (cont.)

```bash
... | eval alert=300 | rename alert AS "Alert Threshold"
| appendcols [ search index="_internal" sourcetype="scheduler" thread_id="AlertNotifier*" NOT (alert_actions="summary_index" OR alert_actions=""))
| where like(savedsearch_name, "%Error Rate Threshold Exceeded%")
| timechart count(savedsearch_name) AS Alert
| fields Alert | eval Alert= Alert * 350 ]
```
To turn this...
...into this

Alerts Fired

Alert Threshold
and this...

67 Site Errors

117 PRE ERRORs

3755 PRE WARNs
and this...

<table>
<thead>
<tr>
<th>_time</th>
<th>Top 10 Errors</th>
<th>Msg Count</th>
<th>Total Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wed Jul 19 22:15:00 2017</td>
<td>Minor - <a href="https://zbrt.atl.zillow.net/browse/PERS-5522">https://zbrt.atl.zillow.net/browse/PERS-5522</a> Connection hard closed due to exception: java.sql.SQLException: The target connection was closed by the server. Unable to obtain pooled connection for bean 'SubscriptionReadonly': Ti</td>
<td>16596</td>
<td>38286</td>
</tr>
<tr>
<td></td>
<td>Minor - <a href="https://zbrt.atl.zillow.net/browse/ZOS-1564">https://zbrt.atl.zillow.net/browse/ZOS-1564</a> Unable to obtain pooled connection for bean 'SubscriptionReadonly': Ti</td>
<td>3913</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Major - <a href="https://zbrt.atl.zillow.net/browse/ADS-13199">https://zbrt.atl.zillow.net/browse/ADS-13199</a> Unable to obtain pooled connection for bean 'SubscriptionReadonly': Ti</td>
<td>3913</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Error getting latest ordinal</td>
<td>4372</td>
<td></td>
</tr>
<tr>
<td></td>
<td>org.springframework.jdbc.CannotGetJdbcCon</td>
<td>7536</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exception caught in root handler: org.springframework.jdbc.CannotGetJdbcCon</td>
<td>494</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exception caught in root handler: org.springframework.dao.TransientData</td>
<td>291</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unable to obtain pooled connection for bean 'Subscription': Login time java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Major - <a href="https://zbrt.atl.zillow.net/browse/ADS-13199">https://zbrt.atl.zillow.net/browse/ADS-13199</a> Unable to obtain pooled connection for bean 'Subscription': Login time java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>18946</td>
<td>37852</td>
</tr>
<tr>
<td></td>
<td>java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>7869</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unable to obtain pooled connection for bean 'Subscription': Login time java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>4226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>4226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Major - <a href="https://zbrt.atl.zillow.net/browse/ADS-13199">https://zbrt.atl.zillow.net/browse/ADS-13199</a> Unable to obtain pooled connection for bean 'Subscription': Login time java.sql.SQLException: Login timed out. connecting to: jdbc:postgresql</td>
<td>579</td>
<td></td>
</tr>
</tbody>
</table>
1. Create and fill fields in JIRA tickets
   - Manual or semi-automated

2. Search JIRA with jirarest to populate lookup file
   - Saved Search (/10min)

3. Enable wildcard searching - transforms.conf
   - One-time configuration

4. Create searches against fields using lookups
   - Dashboards, alerts, leaderboards, etc.

5. Enhance with evals and conditionals
   - Many, many possibilities
Thank You

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