Service Ownership

Learn Faster

Holly Allen
Service Engineering
@hollyjallen
Holly Allen
Software development and leadership for 19 years
Software! 😍
SLOW 😞
Toyota Production System
Kanban board

5 Backlog Max 5

1 Selected for Development

2 In Progress Max 5

3 Done

Release...

Quick Filters: Only My Issues Recently Updated

5 Backlog Max 5

1 Selected for Development

2 In Progress Max 5

3 Done

Release...

Quick Filters: Only My Issues Recently Updated

Expediting 2 issues:

SSI-1
Kanban cards represent work items >> Click the "SSI-1" link at the top of this card.

SSI-2
Kanban boards are often divided into streams of work, aka Swimlanes. By default.

Everything Else 8 issues:

SSI-3
Add work items with "Create Issue" at the top right of the screen >> Try adding.

SSI-4
Work items are ranked in priority order (from top to bottom) >> Try dragging this.

SSI-5

SSI-6
Work In Progress (WIP) limits highlight delays. This column's limit is 1...

SSI-7
...so 2 work items violate the limit and cause the column to be highlighted.

SSI-8
Filters at the top of the board allow you to quickly cut down the shown items >> Try.

SSI-9
As teams develop with Kanban they get better at reducing average resolution.

SSI-10

@hollyjallen #agile2019 Aug 2019
<table>
<thead>
<tr>
<th>To Do</th>
<th>Doing</th>
<th>Done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use Kanban</td>
<td>Learn about Kanban</td>
<td>Get some sticky notes!</td>
</tr>
<tr>
<td>Try Kanban Tool</td>
<td></td>
<td>Get a whiteboard</td>
</tr>
</tbody>
</table>

@hollyjallen agile2019 Aug 2019
Kaizen
Continuous Improvement
Think about great and less-than-great agile teams you've been on.

What did the great teams have that the less-than-great teams didn't?
Why do some teams deliver the right things fast?
Executive dedication to learning
High Trust Teams
Design  Measure  Learn
Slack launched February 2014
5 Years

Grew to 10+ million daily active users in more than 150 countries

@hollyjallen agile2019 Aug 2019
5 Years

From 100 to 15,000 servers
In 25 cloud data centers
world-wide
5 Years

From 8 to 1,600 people
In 10 offices world-wide
Design

Measure

Learn
Continuous Deployment
Experiment Frameworks
User Research

@hollyjallen #agile2019 Aug 2019
Something didn't scale...
Centralized Operations 😭
Who should be responsible for the management, monitoring and operation of a production application?
Centralized Operations
Division of Labor
<table>
<thead>
<tr>
<th>Devs</th>
<th>Ops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features</td>
<td>Cloud Infra</td>
</tr>
<tr>
<td>Scale</td>
<td>Deployment</td>
</tr>
<tr>
<td>Architecture</td>
<td>Monitoring</td>
</tr>
</tbody>
</table>
Ops is getting the pages
Product Development grew faster than Operations,
A lot faster
20 Product Developers

1 Ops Engineer
How can operations reliably reach the developers when there's a problem?
"Call Maude, she knows how this works"
Devs
I've never been on-call before, this is scary!

Ops
Now I know I can find a developer when I need to.
Ops is getting the pages first pages
Ultra-senior devs on-call
How can operations reliably reach the developers when there's a problem?
Most devs go on-call
Fall 2017
Kaizen
Continuous Improvement
"Wait, I'm on-call now?"
Devs

I'm glad I'm only on call a few times a year

Ops

I'll be able to reach a frontend engineer if I need to.
Learn by Doing
On-call 3 times a year 🤔
Ops is getting the pages first pages Ultra-senior devs on-call Seven One dev rotations
Continuous Deployment
100+ prod deploys a day
What Changed?
Deploy Chicken APP 4:02 PM

duretti deployed d0F90HB6BES

PR #114458 from button_button_button_button_must_be_up_to_somethin_whoo - duretti
PR #114467 from user_groups_lfs_ac - allisoncraig
PR #114507 from wkimeria - date_picker_bug_fixes -
incoming-webhook: job-queue-relay-04d256538630bc4dc-6586 logged an error

E 1027 17:47:10.844089 6586 topic:157 {} Consul error while reading redis host: Consul error while trying to get key value for job-queue-relay/redis_hosts/indexing, Unexpected response code: 500

Relay process starting: job-queue-relay-04d256538630bc4dc-17582 v97

Grafana - PDE: [Alerting] Total Percent Job Dequeues to Enqueues alert
We're dequeuing jobs at a lower rate than normal. Runbook: https://slack-github.com/slack/docs/blob/master/appeng/alert_runbooks/job_queue.md#total-percent-job-dequeues-to-enqueues-alert

rate of dequeue to enqueue
71.190912049222

Grafana v5.3.0 | Oct 27th
Page the dev
Devs

I don't understand this part of the code

Ops

These are the machine alerts I'm seeing
Human Routers
"Call Andy, he knows how this works"
Postmortems weren't a great place for learning
Reorg!
Fall 2017

Operations is out
Service Engineering is in
How can we ensure that developers know when there's a problem?
Activity! 10 Minutes

Think about devops transformations you've observed or been a part of.

What made it work? What made it a disaster?
Centralized Operations
Service Ownership
I joined Slack in February 2018
Design
Learn
Measure
High Trust
Learning Organization
Site Reliability Engineers

- Devops generalists
- Emotional intelligence
- Mentoring
- Ambassadors
- Operational maturity
SRE embedded in dev teams
Devs

Um, where are the SREs?

SREs

I'm over here doing operational tasks
SREOps is still getting the first pages
How do we lower operational burden on the SREs?
Keep investing in Service Ownership
Plan: Send these paging alerts to the development teams
Define service health and operational maturity

- The teams have to know what they own
- We had a lot of features and services without formal ownership
Define service health and operational maturity

- At least one alerting health metric, like latency or throughput
Define service health and operational maturity

- Teams needed to get on-call ready
- At least 4, preferably 6+ engineers participating to make it sustainable
Devs
We need training

SREs
We're going to plan this out perfectly
Host level alerts
Hundreds of them
Test with the users
Everything was fine!
Empowered
Continuous Improvement
Devs  SRE  Ops
Spring 2019 - All Dev Teams Oncall
Devs and Ops are getting the first pages. Ultra-senior devs on-call. Dozens Seven One of dev rotations.
How do we test our understanding of how Slack will fail?
"Disasterpiece Theater is an ongoing series of exercises in which we will purposely cause a part of Slack to fail."
Design

Measure

Learn
Success Metrics

• Increased engineer confidence
• Validate reliability improvements
• Learn something new
Centralized Operations
Service Ownership
How do we make postmortems a source of learning?
How do we make sure that incident command is a skillset throughout Slack?
How do we make it even easier to operate a service?
Design

Measure

Learn
Copy the questions
Not the answers
Change is possible
You don't have to be ready
Empowered
High Trust
Continuous Improvement
Learn Faster
Thank You!

Holly Allen
Service Engineering
@hollyjallen