You can't stop fires with an ambulance

SRECon Asia/Australia '18

Piers Chamberlain

Head of Site Reliability Engineering, Xero

Xero - Beautiful Accounting Software

- Founded in 2006 by Rod Drury

- Combined flexibility and convenience of cloud based software was a great success

- Today we are a leading cloud-based accounting platform for small businesses and their advisers

- Over 1.3 million subscribers in more than 180 countries connected to the platform

- *>1,200 product updates last year!*

A brief history of SRE at Xero

Established June 2016

- As we migrated from a physical datacenter to a cloud provider
- Most of the old operations team transitioned into product groups

Three geographically separated teams

- Today we're at around 30 SRE engineers
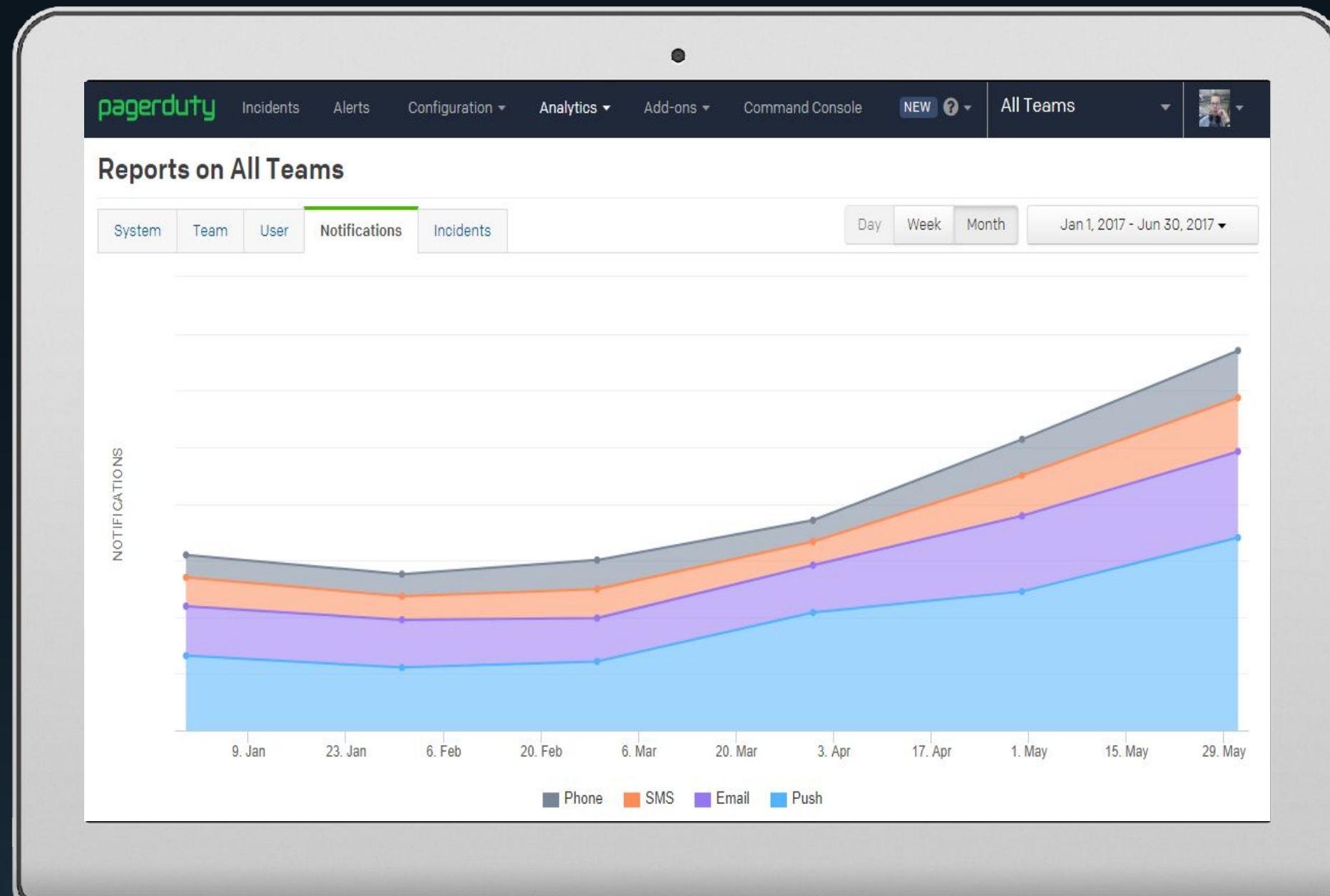- Like most of you, we're hiring

Increasing growth and pace of delivery..

1,386,000

Paying subscribers, as at 31 March 2018

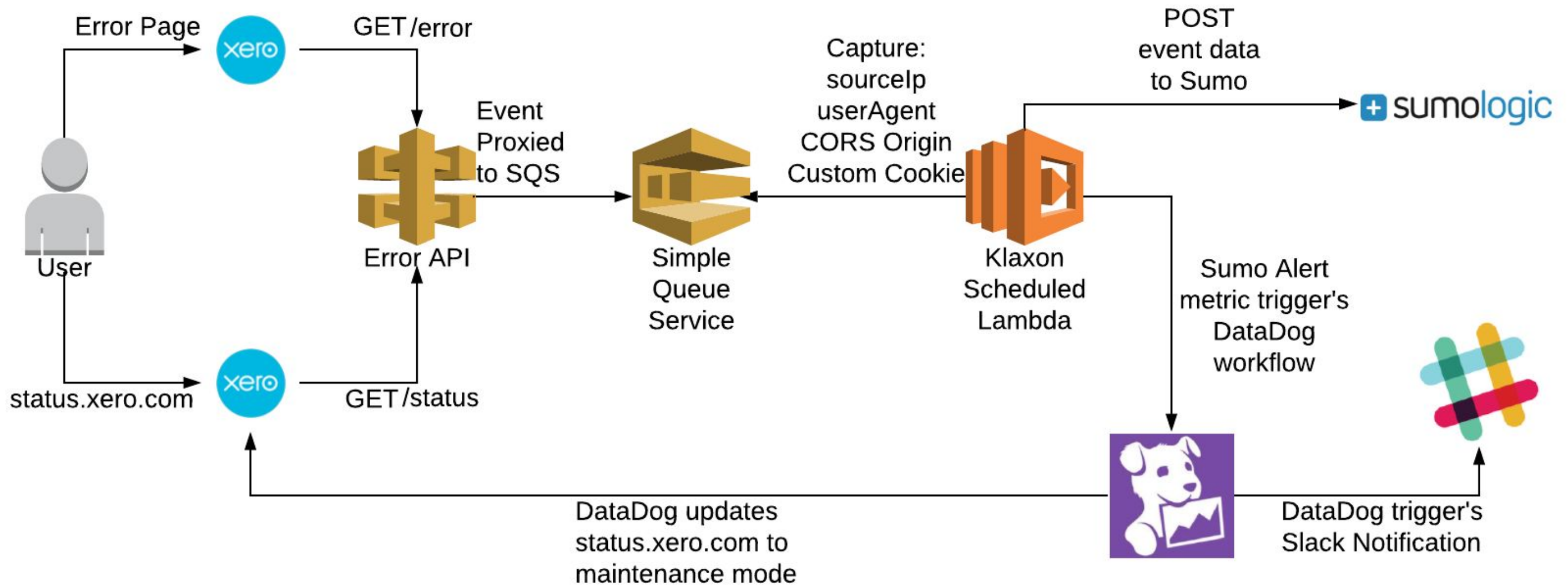| 2009	2010	2011	2012	2013	2014	2015	2016	2017	2018

xero

..increased number of alerts in production

So we simplified alerting

- Started with generic alerts - for example

  - # desired actual hosts

  - CPU credit burn

  - listening for network errors

- We found a good customer concern signal in the
  rate of hits to our status and error pages

xero

Klaxon

Invested in incident management

To start with
We just created a manual process, as a series of steps and a flowchart

Then we built some support
For this manual process in the form of a chatbot

Multivac
Is now an indispensable component, and allows other to run an incident with some sensible guard-rails

On incident closedown
It produces a postmortem document with links to all the relevant resources

xero

Detection and Restoration of Service

- Alerting no longer depended on our customer support people noticing a pattern

- Consistent incident process for management enabled us to work the issues without seeding panic

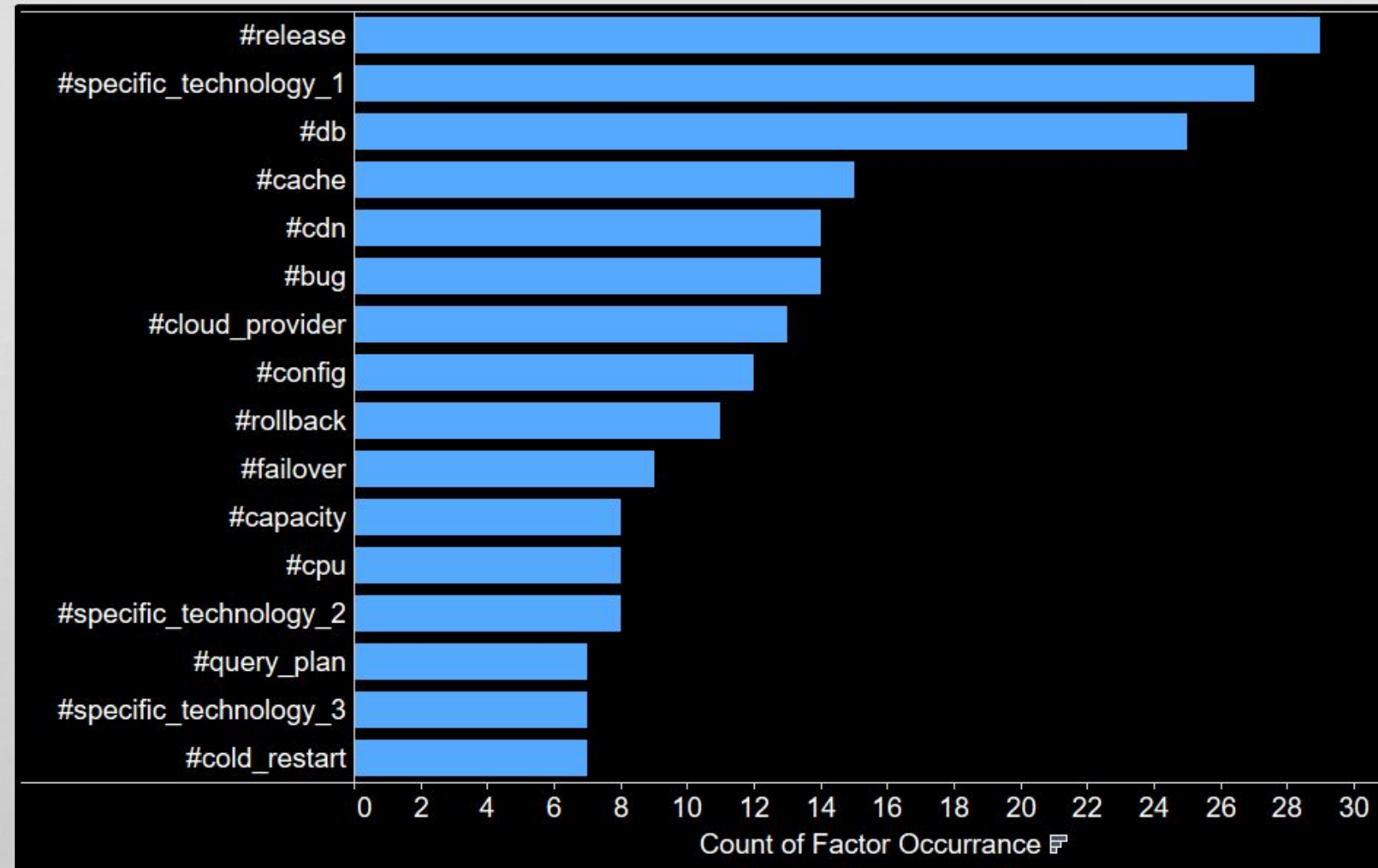- Good wins in the time to detect and time to resolve (MTTD / MTTR)

xero

But the number of incidents is still a problem
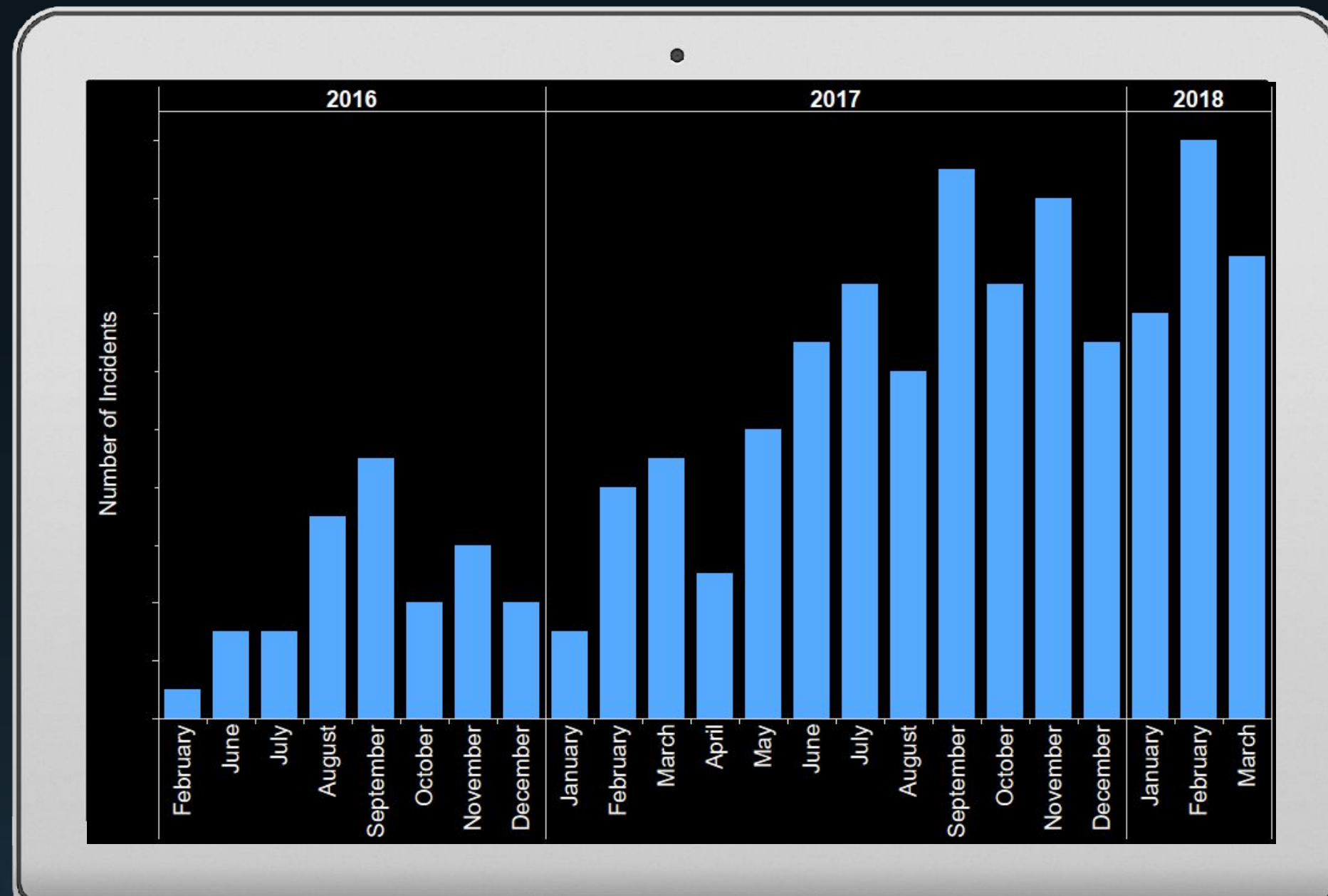
So I went looking for "why"

- I went back over the data and decorated with additional detail
- Themes emerged
  - Similarity in the root cause
  - Often the same teams
  - Often the same deficiencies in detection

xero

Analysis of contributing factors

Our Post-Mortems weren't working as expected

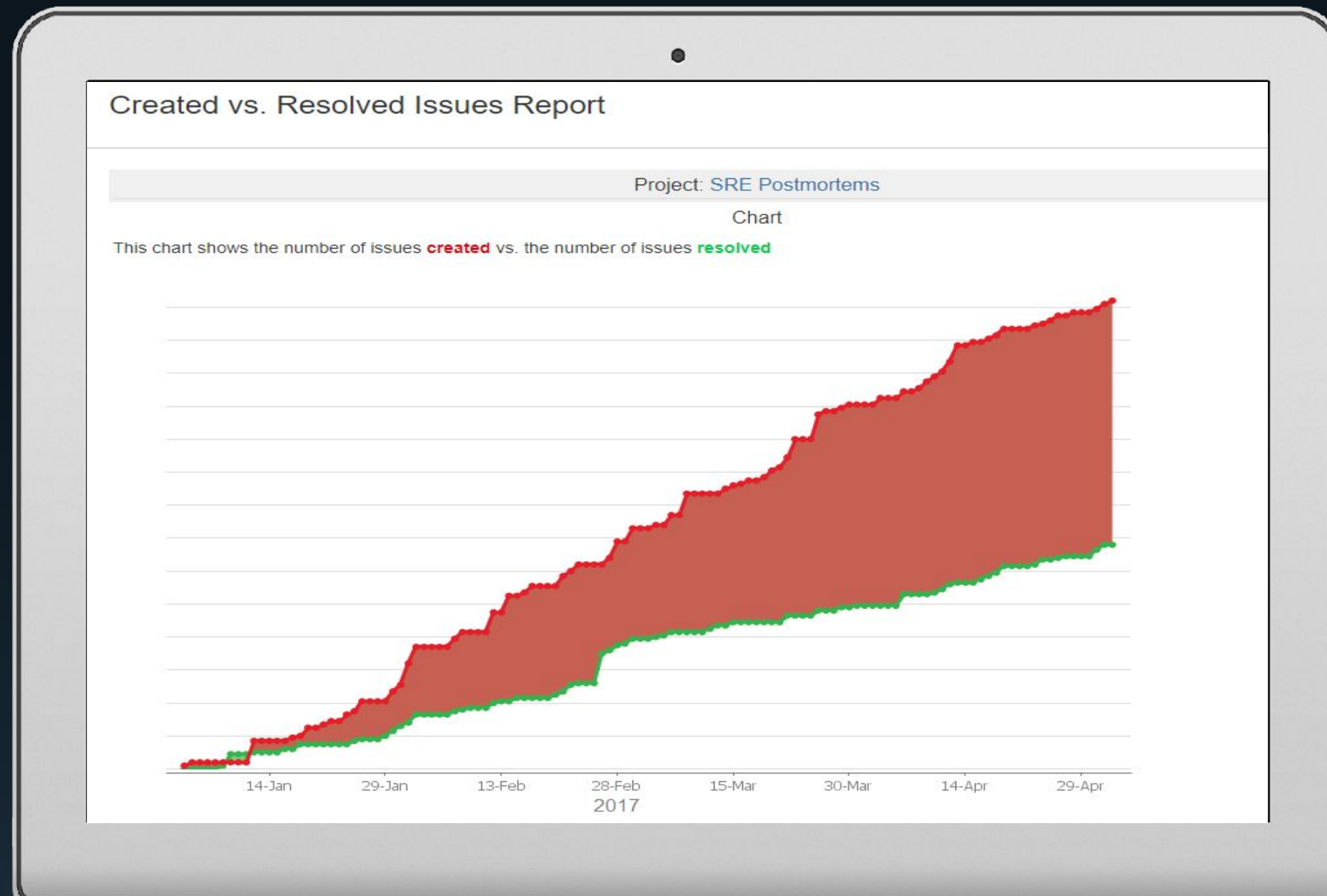We weren't seeing these trends

Limited set of participants

Commitments weren't communicated

Work not prioritised

xero

Identified vs. Completed Actions

Finding the influencers

In our organisation these groups were:

1. Tech Group leads

2. Product Owners / Managers

3. Senior Management

xero

Conversation Starters

Service Level Objectives

Even imperfect measurements were opening eyes

Operational Maturity

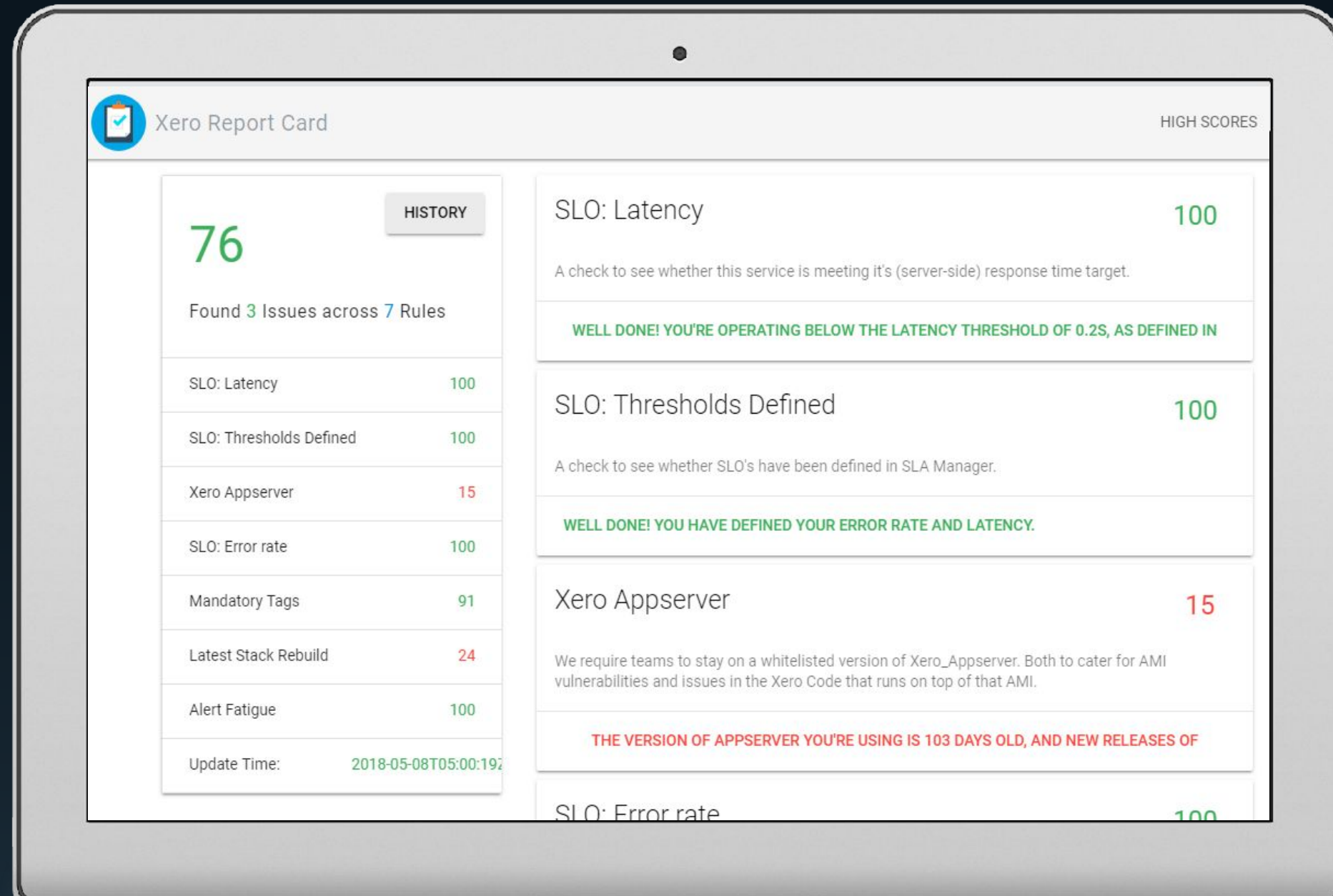Especially with new technology

Escalation Frequency

SRE support (or not)?

Release Process

What are compromises we're making?

xero

Report Card

Key Learning #1

Things I wish we'd done from the start

Track Incident Trends

Devise a system to track contributing causes and themes

For us they centred around:

- Teams involved
- Specific Technologies
- Detection Gaps
- Internal / 3rd Party hard dependencies

xero

Key Learning #2
Things I wish we'd done from the start

Involve the right people

- You need sponsors and leaders
- remedial and preventive work won't happen without their support

xero

Key Learning #3
Things I wish we'd done from the start

Convince and influence using the right language

- Being the "right" thing to do is not enough
- Use arguments based on
  - customer impact
  - engineering cost

Questions?