Don't Panic!

How to Cope Now You're Responsible for Production

Euan Finlay
@efinlay24

FT

FINANCIAL TIMES

Ahmen Khawaja @AhmenKhawaja · 2 hrs

False Alarm: Have deleted previous tweets!!

Ahmen Khawaja @AhmenKhawaja · 2 hrs

False Alarm to Queen's death! She is being treated at King Edward 7th hospital. Statement due shortly

1

Ahmen Khawaja @AhmenKhawaja · 2 hrs

"Queen Elizabrth has died": @BBCWorld

Ahmen Khawaja @AhmenKhawaja · 2 hrs

BREAKING: Queen Elizabeth is being treated at King Edward 7th Hospital in London. Statement due shortly: @BBCWorld

7   1

FT
FINANCIAL
TIMES

December 3, 2015 12:36 pm

ECB leaves rates unchanged in surprising decision

Claire Jones, Frankfurt

Share ∨   Author alerts ∨   Print   Clip   Comments

©Reuters

The European Central Bank has left interest rates unchanged, dashing expectations of a cut to its deposit rate.

FT
FINANCIAL
TIMES

FT **Financial Times** ✔
@FinancialTimes

⚙ 👤 Follow

ECB leaves rates unchanged in shock decision on.ft.com/1Nrekqz

RETWEETS	LIKES	
36	6	

7:38 AM - 3 Dec 2015

↩ ⇄ ♡ •••

FT
FINANCIAL
TIMES

At 13:58   Op 1.0615   Hi 1.0690   Lo 1.0524   Close 1.0615   Value 12/07/15

EURUSD BGN Curncy | 95) Comparer | 96) Actions | 97) Editer | Graphe à barres intraday

Périodicité | 1 | jr(s) | 1 | 23:00 - 22:59 | BGN | MMob. | Aucun graphe inf

1J 3J 1M 6M YTD 1A 5A Max | 1 Min | Tableau | Titre/Etude | Evènement

Début: 12/03 13:17 Fin: 12/03 14:03
Dernier Prix   1.0670
Clôture le 12/02 ----   1.0615
Début le 12/03 13:18   1.0550
Plus haut le 12/03 13:45   1.0690
Moyenne   1.0595
Plus bas le 12/03 13:45   1.0524

1.0700
1.0680
1.0670
+0.0120
1.14%
1.0640
1.0620
1.0600
1.0580
1.0560
1.0540
1.0520

13:20   13:25   13:30   13:35   13:40   13:45   13:50   13:55   14:00
03 Déc 2015

Australia 61 2 9777 8600 Brazil 5511 2395 9000 Europe 44 20 7330 7500 Germany 49 69 9204 1210 Hong Kong 852 2977 6000
Japan 81 3 3201 8900   Singapore 65 6212 1000   U.S. 1 212 318 2000   Copyright 2015 Bloomberg Finance L.P.
SN 204778 CET   GMT+1:00 H216-2408-2 03-Dec-2015 13:58:36

https://twitter.com/Ludo_Dufour/status/672401158653218816

FINANCIAL TIMES

On Thursday we published an incorrect story on FT.com that stated the European Central Bank had confounded expectations by deciding to hold interest rates rather than cut them. The story was published a few minutes before the decision to cut rates was announced.

The story was wrong and should not have been published. The article was one of two pre-written stories — covering different possible decisions — which had been prepared in advance of the announcement. Due to an editing error it was published when it should not have been. Automated feeds meant that the initial error was compounded by being simultaneously published on Twitter.

The FT deeply regrets this serious mistake and will immediately be reviewing its publication and workflow processes to ensure such an error cannot happen again. We apologise to all our readers.

/usr/bin/whoami

FT

FINANCIAL
TIMES

/usr/bin/whodoiworkfor

No such file or directory.

FINANCIAL
TIMES

US equities

US stocks head for worst month since financial crisis

S&P 500 dips into correction territory as tech sell-off broadens across equities

- For investors, risks are becoming hard to ignore
- Growth jitters send investors heading for tech exit
- No hiding place for investors in markets wobble

Analysis **Markets**

No hiding place for investors in markets wobble

Almost every major asset class has fallen into negative territory for the year

UK Budget
Hammond Budget to tackle 'gap-toothed high street'

Plan to turn more town centre properties into housing and office space

Global trade
Blow for Fox as UK's fast-track WTO bid hits buffer

Members object to Britain's plan to copy EU's international trade commitments

US politics & policy
Mail bomb suspect charged with five counts

Cesar Sayoc arrested in Florida after interception of 13 devices sent to Trump critics

Forex trading probes
Bank traders acquitted in forex manipulation trial

Price fixing case against London-based trio was brought by US prosecutors

https://www.ft.com

FT

FINANCIAL TIMES

We're hiring!

https://ft.com/dev/null

FT

FINANCIAL
TIMES

You've just been told you're on call.

And you're mildly terrified.

FT

FINANCIAL
TIMES

Obligatory audience interaction.

FT

FINANCIAL
TIMES

Everyone feels the same at the start.

I still do today.

FT

FINANCIAL TIMES

How do you get more comfortable?

FT

FINANCIAL
TIMES

The Ghosts of Incidents...

> Future

The Ghosts of Incidents...

Future

> Present

FT

FINANCIAL TIMES

The Ghosts of Incidents...

Future
Present
> Past

FT

FINANCIAL
TIMES

The Ghost of Incidents Future

Handling incidents is the same as any other skill.

#VelocityConf

Get comfortable with your alerts.

FT

FINANCIAL
TIMES

Bin the alerts you don't care about.

FT

FINANCIAL
TIMES

Have a plan for when things break.

FT

FINANCIAL
TIMES

Keep your documentation up to date.

FT

FINANCIAL
TIMES

Practice regularly.

FT

FINANCIAL
TIMES

"The Gang Deletes
Production"

NO PRODUCTION SERVERS?

NO PRODUCTION INCIDENTS

Break things, and see what happens.

Did your systems do what you expected?

FT

FINANCIAL
TIMES

The Planned Datacenter Disconnect

We got complacent, and stopped running datacenter failure tests...

FT

FINANCIAL
TIMES

The Unplanned Datacenter Disconnect

Have a central place for reporting changes and problems.

FT

FINANCIAL
TIMES

▢ 19:25
Seeing aws dx link issues again-checking

Posted image at 2018-07-27, 5:28 PM ▾

All Groups

▸ 🔴 AWS DX
▸ 🟢 BT
▸ 🟢 FT DNS
▸ 🟢 FT Jetnexus
▸ 🟢 FT NTP
▾ 🔴 FT Third Parties
　🔴 Nexxlinx VPN
▸ 🟢 FT Voice

▢ 👀 19:39
Methode alerts are firing

intermittent

▢ 19:40
yep we have network issues again at PR (edited)

😢 1

Looks like the MPLS Verizon cct is down

So far no impact reported...

monitoring for now

▢ 🇦🇺 19:44
We have reports of publishing not working, and problems with Methode portalpub connecting to UPP again

▢ 19:45
thanks Andrew

^^Verizon are saying PR site is affected by an issue affecting multiple locations

▢ 👀 19:46
switching portalpub off in PR

The Unplanned Datacenter Disconnect II: The Network Strikes Back

We should have followed our own advice.

FT

FINANCIAL TIMES

We're not perfect.

But we always try to improve.

FT

FINANCIAL TIMES

The Ghosts of Incidents...

Future

> Present

Past

FT

FINANCIAL
TIMES

The Ghost of Incidents Present

Calm down, and take a deep breath: it's probably ok.

FT

FINANCIAL
TIMES

Don't dive straight in.

Go back to first principles.

FT

FINANCIAL
TIMES

What's the actual impact?

FT

FINANCIAL
TIMES

"All incidents are **equal**, but some incidents are **more equal** than others."

George Orwell, probably.

FT

FINANCIAL TIMES
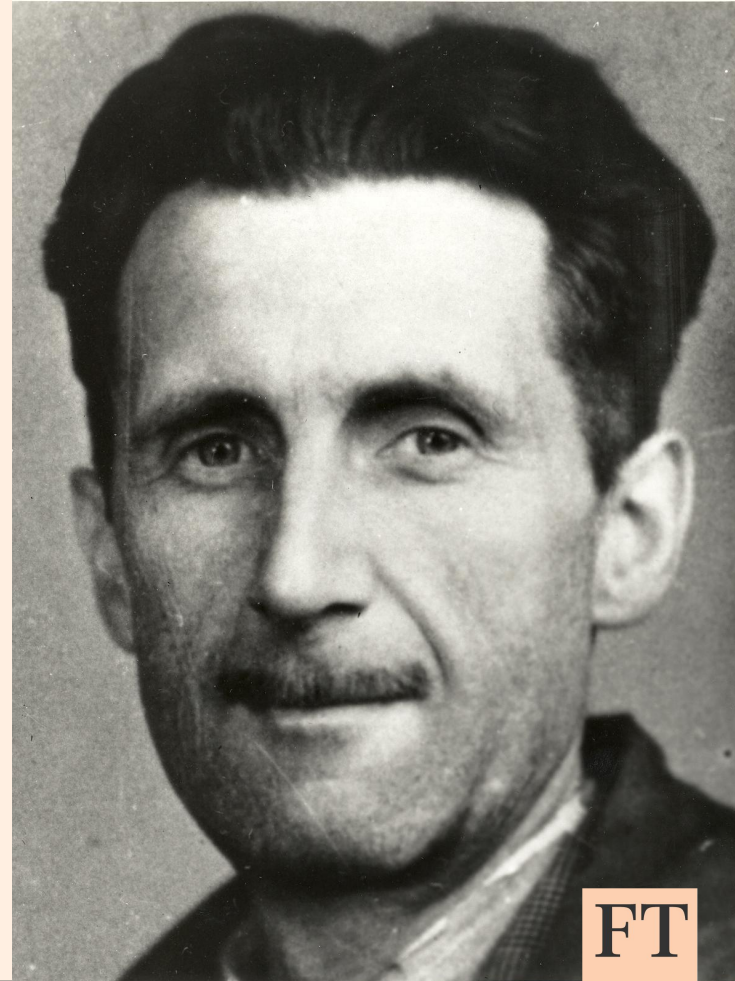
What's already been tried?

FT

FINANCIAL
TIMES

Is there definitely a problem?

FT

FINANCIAL
TIMES

DYNPUB / EXTRAPUB Prod Park Royal
Content

DYNPUB / EXTRAPUB Prod Watford
Content

SEMANTIC Prod Park Royal
Content

SEMANTIC Prod Watford
Content

SEMANTIC Prod AWS
Content
Downtime... S...

DON'T
PANIC

COCO Restful
OK
Healthy

MASHERY-V2-TESTS-PROD-UK

MASHERY-V2-TESTS-PROD-US

DON'T
WORRY

EVERYTHING
IS
FIVE

PROBABLY
OK

SEMANTIC Test Park Royal
Aqled... 4.

SEMANTIC Test Watford

NOT
CRTICAL

PAT TEST
0

DYNPUB / EXTRAPUB Int Park Royal
Content

SEMANTIC Int Park Royal

PAT TEST
0

What's the minimum viable solution?

FT
FINANCIAL TIMES

Get it running *before* you get it fixed.

FT

FINANCIAL
TIMES

Go back to basics.

Don't forget to check the logs.

FT

FINANCIAL
TIMES

It's not DNS
There's no way it's DNS
It was DNS

-SSBroski

THIS IS FINE.
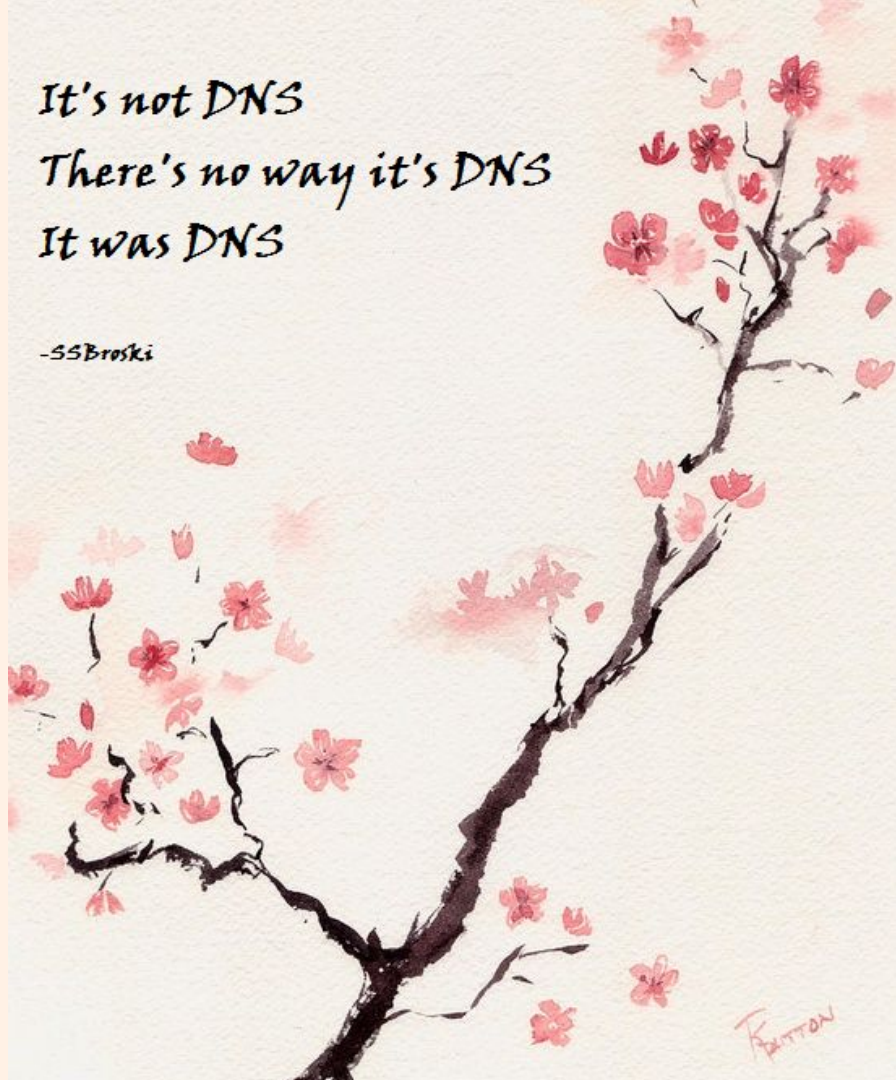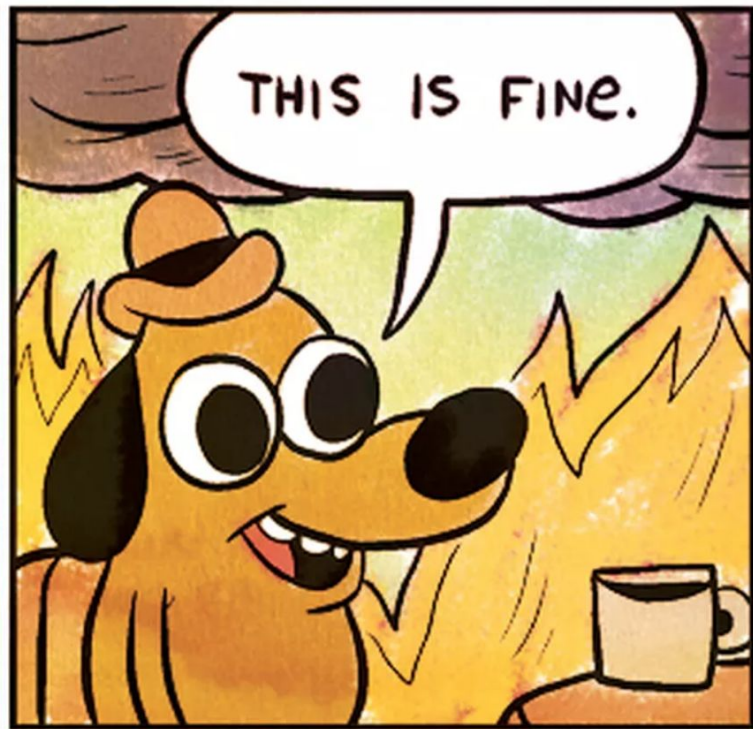
Don't be afraid to call for help.

FT

FINANCIAL
TIMES

The One Where a Director Falls Through the Ceiling

Communication is key.

Especially to your customers.

FT

FINANCIAL
TIMES

Put someone in charge.

FT

FINANCIAL
TIMES

Pingdom APP 14:42

k8s - UPP Prod Delivery US: Annotations Read Aggregate Healthcheck is down (Incident #4073269)
upp-prod-delivery-us.ft.com • View details

k8s - UPP Prod Delivery UK: Annotations Read Aggregate Healthcheck is down (Incident #4073041)
upp-prod-delivery-eu.ft.com • View details

k8s - UPP Prod Delivery UK: Content Read Aggregate Healthcheck is down (Incident #4073077)
upp-prod-delivery-eu.ft.com • View details

k8s - UPP Prod Delivery US: Content Publish Aggregate Healthcheck is down (Incident #4073290)
upp-prod-delivery-us.ft.com • View details

k8s - UPP Prod Delivery US: Annotations Read Aggregate Healthcheck is up (Incident #4073269)
upp-prod-delivery-us.ft.com • View details

k8s - UPP Prod Delivery US: Image Publish Aggregate Healthcheck is down (Incident #4073407)
upp-prod-delivery-us.ft.com • View details

k8s - UPP Prod Delivery UK: Image Publish Aggregate Healthcheck is down (Incident #4073095)
upp-prod-delivery-eu.ft.com • View details

k8s - UPP Prod Delivery UK: List Read Aggregate Healthcheck is down (Incident #4073146)
upp-prod-delivery-eu.ft.com • View details

k8s - UPP Prod Delivery US: Content Publish Aggregate Healthcheck is up (Incident #4073290)
upp-prod-delivery-us.ft.com • View details

Create a temporary incident channel.

FT

FINANCIAL
TIMES

Software can be chaotic, but we make it work

Expert

Trying Stuff Until it Works

The Practical Developer
@ThePracticalDev

11:26
Same for `EA460F68-495E-D18F-3130-C220FA23EF8E` in prod-us at 04:01:57

11:26
So [REDACTED] says that in the last 24 hours, Next have got 20% more push event notifications in the US than in the EU...

11:26
Both request times line up exactly with the notification timestamps

11:27
uploaded and commented on this image: **Push notifications received** ▾
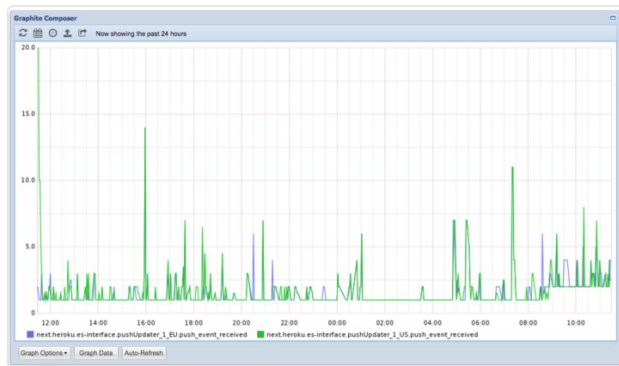
> green is US region

11:28
`suggestions-rw-neo4j` (that's v2 annotations as I understand it) wrote annotations at 04:01:36 (uk) and 04:01:51 (us)
So they would have been present before the notifications were sent out

11:29
[REDACTED] and I have confirmed that no annotations will still allow the article to show in both enrichedcontent and internalcontent

😔 1

Hmm, that's a big gap between 1 and 5

If you think you're
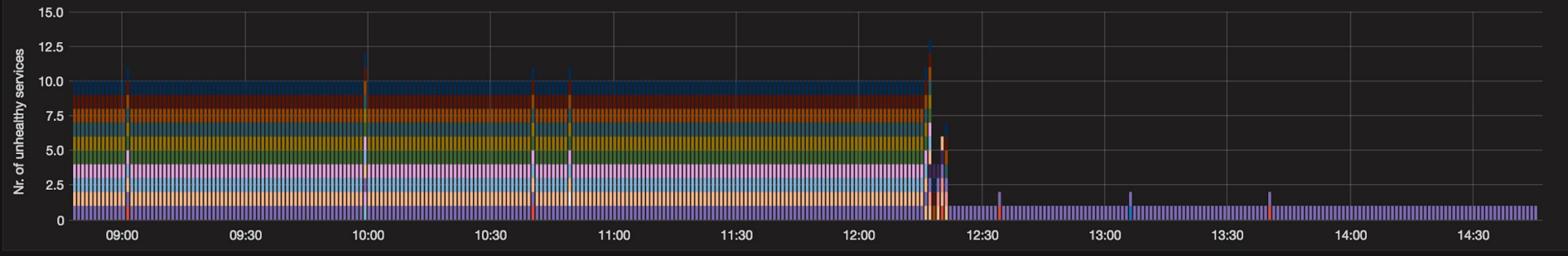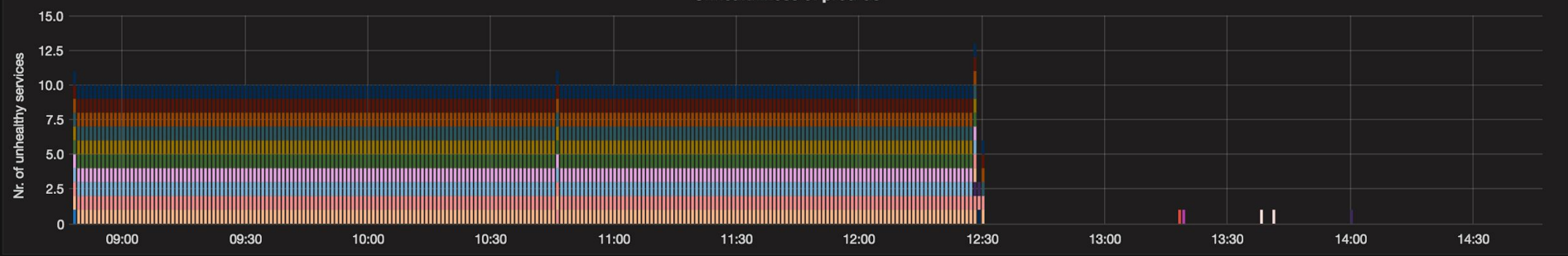over-communicating,
it's probably just the right amount.

#VelocityConf

FT

FINANCIAL
TIMES

Tired people don't think good.

"The Gang Serves Traffic From Staging"

Uptime report

Content API V2 Lists API ██ Type: HTTP

Response time

Downtime	Outages	Uptime	Max resp. time	Min resp. time	Avg resp. time
1h51m	21	96.15%	1.23s	218ms	302ms

— Average response time   ■ Downtime   □ Unknown
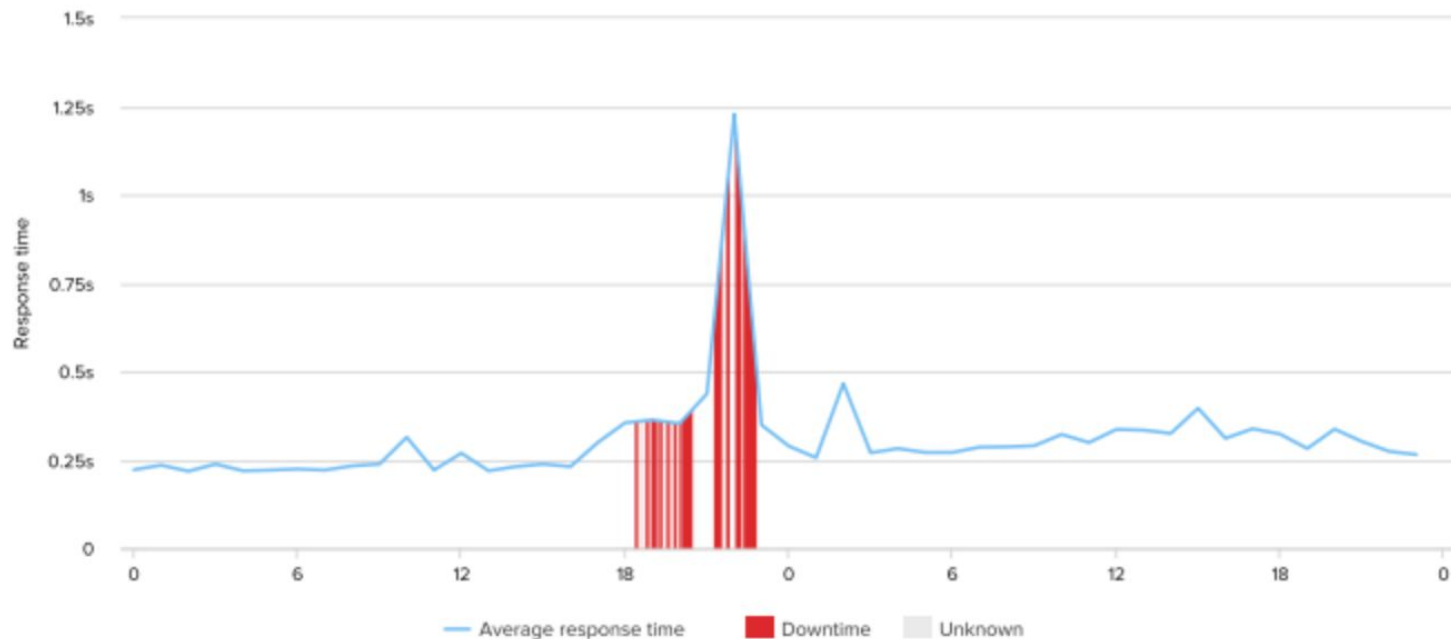
It wasn't great,
but it wasn't the end of the world.

FT

FINANCIAL
TIMES

The Ghosts of Incidents...

Future
Present
> Past

The Ghost of Incidents Past

Congratulations!
You survived.

It probably wasn't **that** bad, was it?

#VelocityConf

FT

FINANCIAL TIMES

Run a learning review with everyone involved.

FT

FINANCIAL
TIMES

Incident reports are important.

FT

FINANCIAL TIMES

Postmortem of database outage of January 31

Postmortem on the database outage of January 31 2017 with the lessons we learned.

← Back to company

On January 31st 2017, we experienced a major service outage for one of our products, the online service GitLab.com. The outage was caused by an accidental removal of data from our primary database server.

This incident caused the GitLab.com service to be unavailable for many hours. We also lost some production data that we were eventually unable to recover. Specifically, we lost modifications to database data such as projects, comments, user accounts, issues and snippets, that took place between 17:20 and 00:00 UTC on January 31. Our best estimate is that it affected roughly 5,000 projects, 5,000 comments and 700 new user accounts. Code repositories or wikis hosted on GitLab.com were unavailable during the outage, but were not affected by the data loss. GitLab Enterprise customers, GitHost customers, and self-hosted GitLab CE users were not affected by the outage, or the data loss.

"Until a restore is attempted, a backup is both **successful** and **unsuccessful**."

Definitely not Erwin Schrödinger.

FT

Timeline

On January 31st an engineer started setting up multiple PostgreSQL servers in our staging environment. The plan was to try out pgpool-II to see if it would reduce the load on our database by load balancing queries between the available hosts. Here is the issue for that plan: infrastructure#259.

± 17:20 UTC: prior to starting this work, our engineer took an LVM snapshot of the production database and loaded this into the staging environment. This was necessary to ensure the staging database was up to date, allowing for more accurate load testing. This procedure normally happens automatically once every 24 hours (at 01:00 UTC), but they wanted a more up to date copy of the database.

± 19:00 UTC: GitLab.com starts experiencing an increase in database load due to what we suspect was spam. In the week leading up to this event GitLab.com had been experiencing similar problems, but not this severe. One of the problems this load caused was that many users were not able to post comments on issues and merge requests. Getting the load under control took several hours.

We would later find out that part of the load was caused by a background job trying to remove a GitLab employee and their associated data. This was the result of their account being flagged for abuse and accidentally scheduled for removal. More information regarding this particular problem can be found in the issue "Removal of users by spam should not hard delete".

Publication of the outage

In the spirit of transparency we kept track of progress and notes in a publicly visible Google document. We also streamed the recovery procedure on YouTube, with a peak viewer count of around 5000 (resulting in the stream being the #2 live stream on YouTube for several hours). The stream was used to give our users live updates about the recovery procedure. Finally we used Twitter (https://twitter.com/gitlabstatus) to inform those that might not be watching the stream.

The document in question was initially private to GitLab employees and contained name of the engineer who accidentally removed the data. While the name was added by the engineer themselves (and they had no problem with this being public), we will redact names in future cases as other engineers may not be comfortable with their name being published.

Data loss impact

Database data such as projects, issues, snippets, etc. created between January 31st 17:20 UTC and 23:30 UTC has been lost. Git repositories and Wikis were not removed as they are stored separately.

It's hard to estimate how much data has been lost exactly, but we estimate we have lost at least 5000 projects, 5000 comments, and roughly 700 users. This only affected users of GitLab.com, self-hosted instances or GitHost instances were not affected.

Improving recovery procedures

We are currently working on fixing and improving our various recovery procedures. Work is split across the following issues:

1. Overview of status of all issues listed in this blog post (#1684)
2. Update PS1 across all hosts to more clearly differentiate between hosts and environments (#1094)
3. Prometheus monitoring for backups (#1095)
4. Set PostgreSQL's max_connections to a sane value (#1096)
5. Investigate Point in time recovery & continuous archiving for PostgreSQL (#1097)
6. Hourly LVM snapshots of the production databases (#1098)
7. Azure disk snapshots of production databases (#1099)
8. Move staging to the ARM environment (#1100)
9. Recover production replica(s) (#1101)
10. Automated testing of recovering PostgreSQL database backups (#1102)
11. Improve PostgreSQL replication documentation/runbooks (#1103)
12. Investigate pgbarman for creating PostgreSQL backups (#1105)
13. Investigate using WAL-E as a means of Database Backup and Realtime Replication (#494)
14. Build Streaming Database Restore
15. Assign an owner for data durability

FT
FINANCIAL
TIMES

Identify what can be improved for next time.

FT

FINANCIAL
TIMES

Prioritise follow-up actions.

FT

FINANCIAL
TIMES

October 21 post-incident analysis

Oct 30, 2018   jasoncwarner   Engineering

Last week, GitHub experienced an incident that resulted in degraded service for 24 hours and 11 minutes. While portions of our platform were not affected by this incident, multiple internal systems were affected which resulted in our displaying of information that was out of date and inconsistent. Ultimately, no user data was lost; however manual reconciliation for a few seconds of database writes is still in progress. For the majority of the incident, GitHub was also unable to serve webhook events or build and publish GitHub Pages sites.

All of us at GitHub would like to sincerely apologize for the impact this caused to each and every one of you. We're aware of the trust you place in GitHub and take pride in building resilient systems that enable our platform to remain highly available. With this incident, we failed you, and we are deeply sorry. While we cannot undo the problems that were created by GitHub's platform being unusable for an extended period of time, we can explain the events that led to this incident, the lessons we've learned, and the steps we're taking as a company to better ensure this doesn't happen again.

The One with the Badly Named Database

Please don't name your pre-production database:

'pprod'

FT

FINANCIAL TIMES
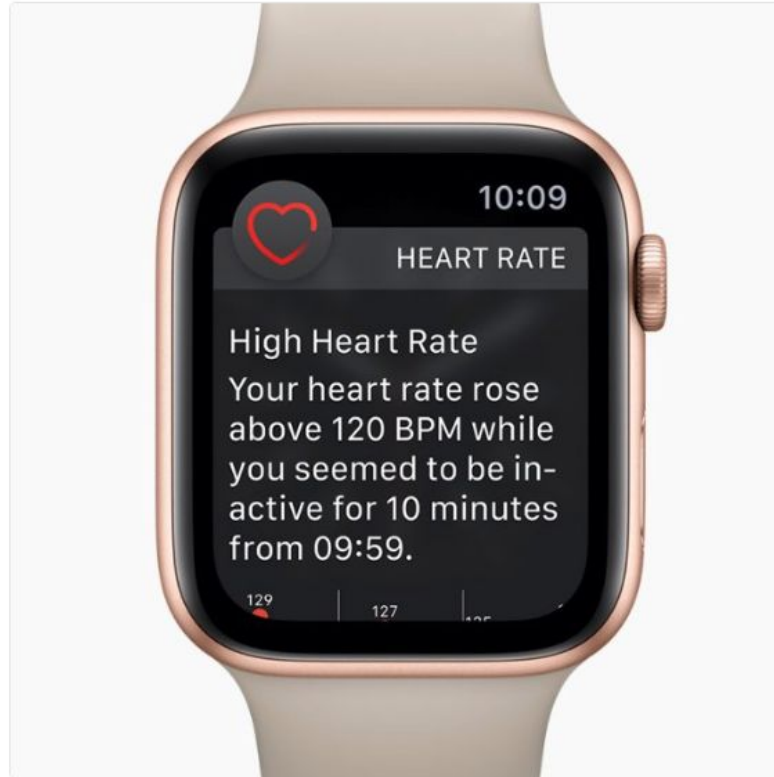
Nearly the end.

Don't clap yet.

FT

FINANCIAL
TIMES

Problems will always happen.

And that's ok.

FT

FINANCIAL TIMES

The end.

Please clap.

FT

FINANCIAL
TIMES

@efinlay24

euan.finlay@ft.com

We're hiring!
https://ft.com/dev/null

FT

FINANCIAL
TIMES