Jug Summer Camp

-enjoy it-

IaaS (Interruption as a Sageness)

David Aparicio

10 Septembre 2021 11h25 Salle Casoar Tadorne

@dadideo

David Aparicio

15/ DD INSA de Lyon / UNICAMP (Brésil)

Facebook Open Academy / MIT AppInventor

17/ Dev(Sec)Ops @ AMADEUS (Nice, 2 ans)

19/ Data(Sec)Ops @ OVHcloud (Lyon, 3 ans)

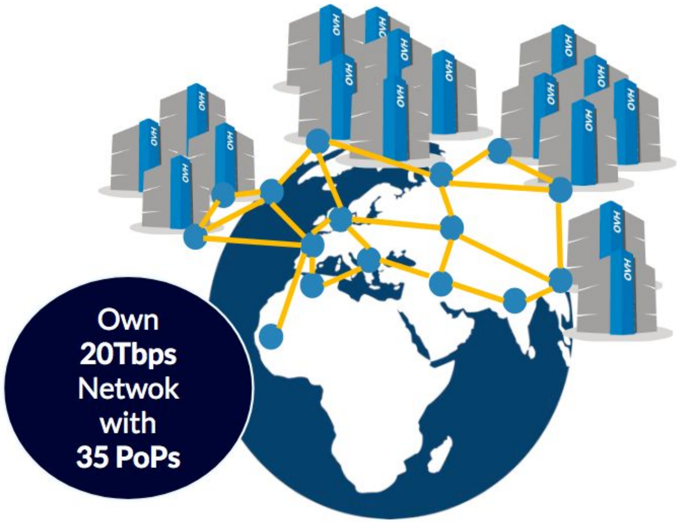
OVHcloud: un leader européen

200k Private cloud VMs running

1 **Dedicated IaaS Europe**

Hosting capacity : 1.3M Physical Servers

360k Servers already deployed



Own 20Tbps Netwok with 35 PoPs

> **1.3M** Customers in **138** Countries

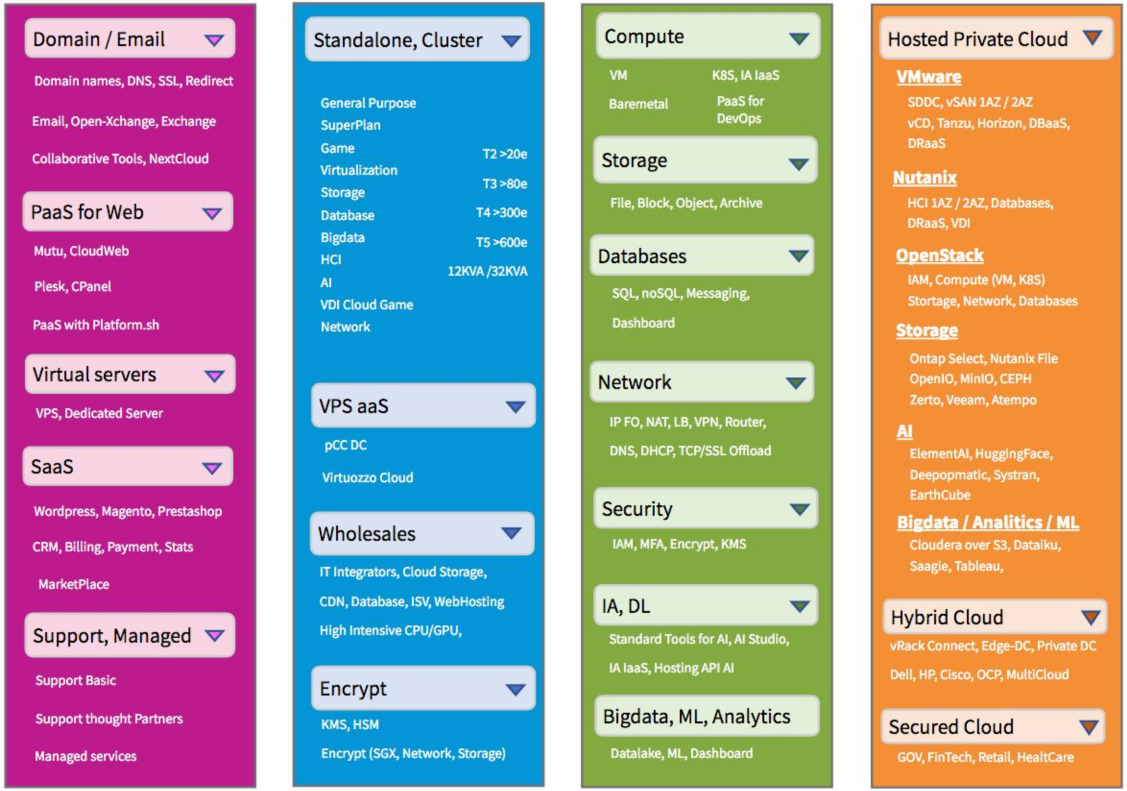
30 Datacenters

GAIA-X

SecNumCloud

Depuis Déc. 2020

OVHcloud: 4 univers de produits



Domain / Email

Domain names, DNS, SSL, Redirect

Email, Open-Xchange, Exchange

Collaborative Tools, NextCloud

PaaS for Web

Mutu, CloudWeb

Plesk, CPanel

PaaS with Platform.sh

Virtual servers

VPS, Dedicated Server

SaaS

Wordpress, Magento, Prestashop

CRM, Billing, Payment, Stats

MarketPlace

Support, Managed

Support Basic

Support thought Partners

Managed services

Standalone, Cluster

General Purpose
SuperPlan
Game
Virtualization
Storage
Database
Bigdata
HCI
AI
VDI Cloud Game
Network

T2 >20e
T3 >80e
T4 >300e
T5 >600e
12KVA /32KVA

VPS aaS

pCC DC

Virtuozzo Cloud

Wholesales

IT Integrators, Cloud Storage,

CDN, Database, ISV, WebHosting

High Intensive CPU/GPU,

Encrypt

KMS, HSM

Encrypt (SGX, Network, Storage)

Compute

VM

K8S, IA IaaS

Baremetal

PaaS for DevOps

Storage

File, Block, Object, Archive

Databases

SQL, noSQL, Messaging,

Dashboard

Network

IP FO, NAT, LB, VPN, Router,

DNS, DHCP, TCP/SSL Offload

Security

IAM, MFA, Encrypt, KMS

IA, DL

Standard Tools for AI, AI Studio,

IA IaaS, Hosting API AI

Bigdata, ML, Analytics

Datalake, ML, Dashboard

Hosted Private Cloud

VMware

SDDC, vSAN 1AZ / 2AZ
vCD, Tanzu, Horizon, DBaaS,
DRaaS

Nutanix

HCI 1AZ / 2AZ, Databases,
DRaaS, VDI

OpenStack

IAM, Compute (VM, K8S)
Stortage, Network, Databases

Storage

Ontap Select, Nutanix File
OpenIO, MinIO, CEPH
Zerto, Veeam, Atempo

AI

ElementAI, HuggingFace,
Deepopmatic, Systran,
EarthCube

Bigdata / Analitics / ML

Cloudera over S3, Dataiku,
Saagie, Tableau,

Hybrid Cloud

vRack Connect, Edge-DC, Private DC

Dell, HP, Cisco, OCP, MultiCloud

Secured Cloud

GOV, FinTech, Retail, HealtCare

Qui êtes-vous ?

Plutôt Dev, Ops, Étudiant ?

Qui a déjà possédé un téléphone d'astreinte ?

Agenda

Introduction

First Blood

Elliot

Legacy

CDN

Retour du Legacy

Blast effect

Same JVM, shoot again

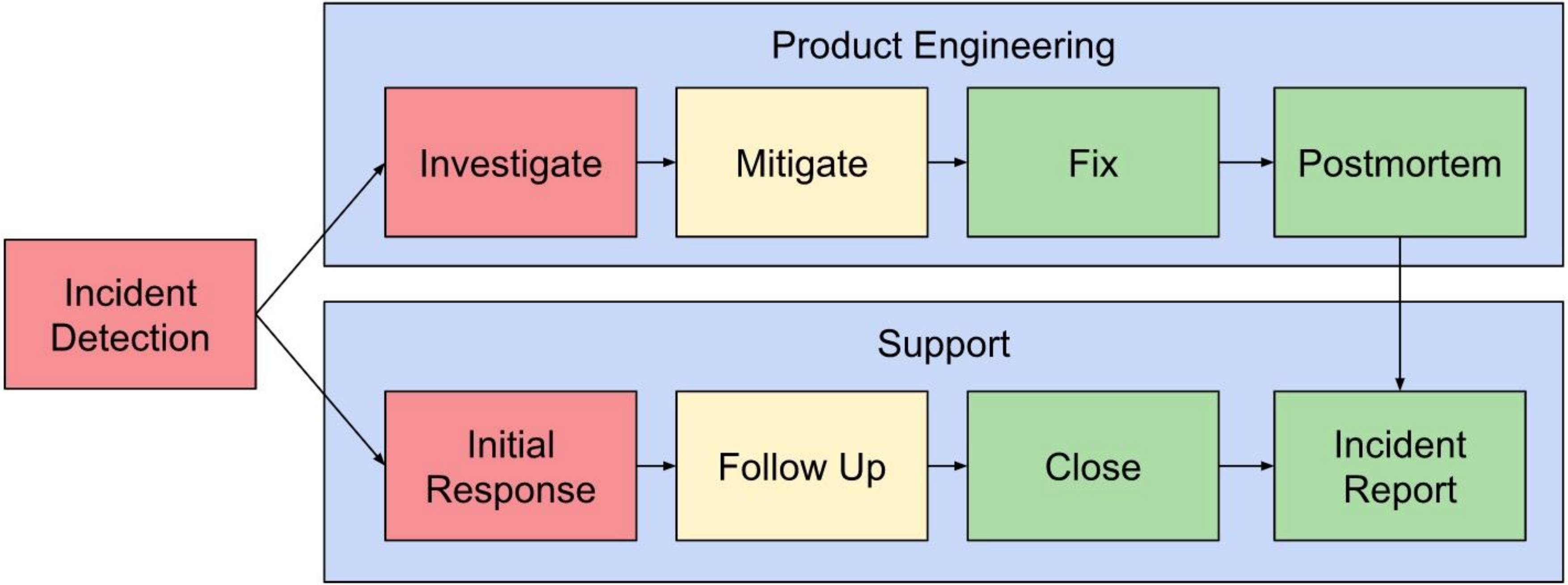
NewsBlur

Split-Brain

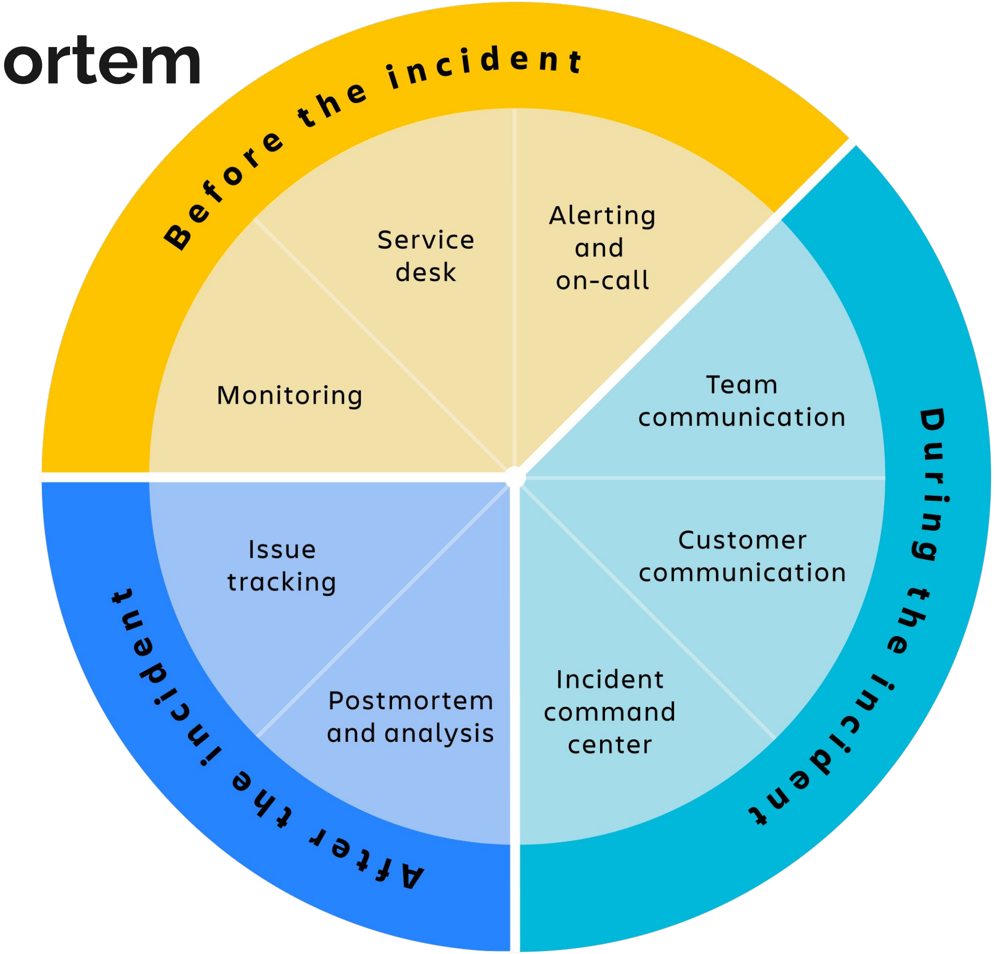
Conclusion

Introduction

Post-mortem


Product Engineering
Investigate
Mitigate
Fix
Postmortem
Incident Detection
Support
Initial Response
Follow Up
Close
Incident Report

Post-mortem


Before the incident
Service desk
Alerting and on-call
Monitoring
During the incident
Team communication
Customer communication
Incident command center
After the incident
Issue tracking
Postmortem and analysis

Examen post-incident

Identification de l'incident (détection, sonde)
Flux d'informations et communication (canaux)
Structure (organisation)

Utilisation des ressources (vitesse, optimum)
Processus (SLA)
Création de rapports (interne et externe)

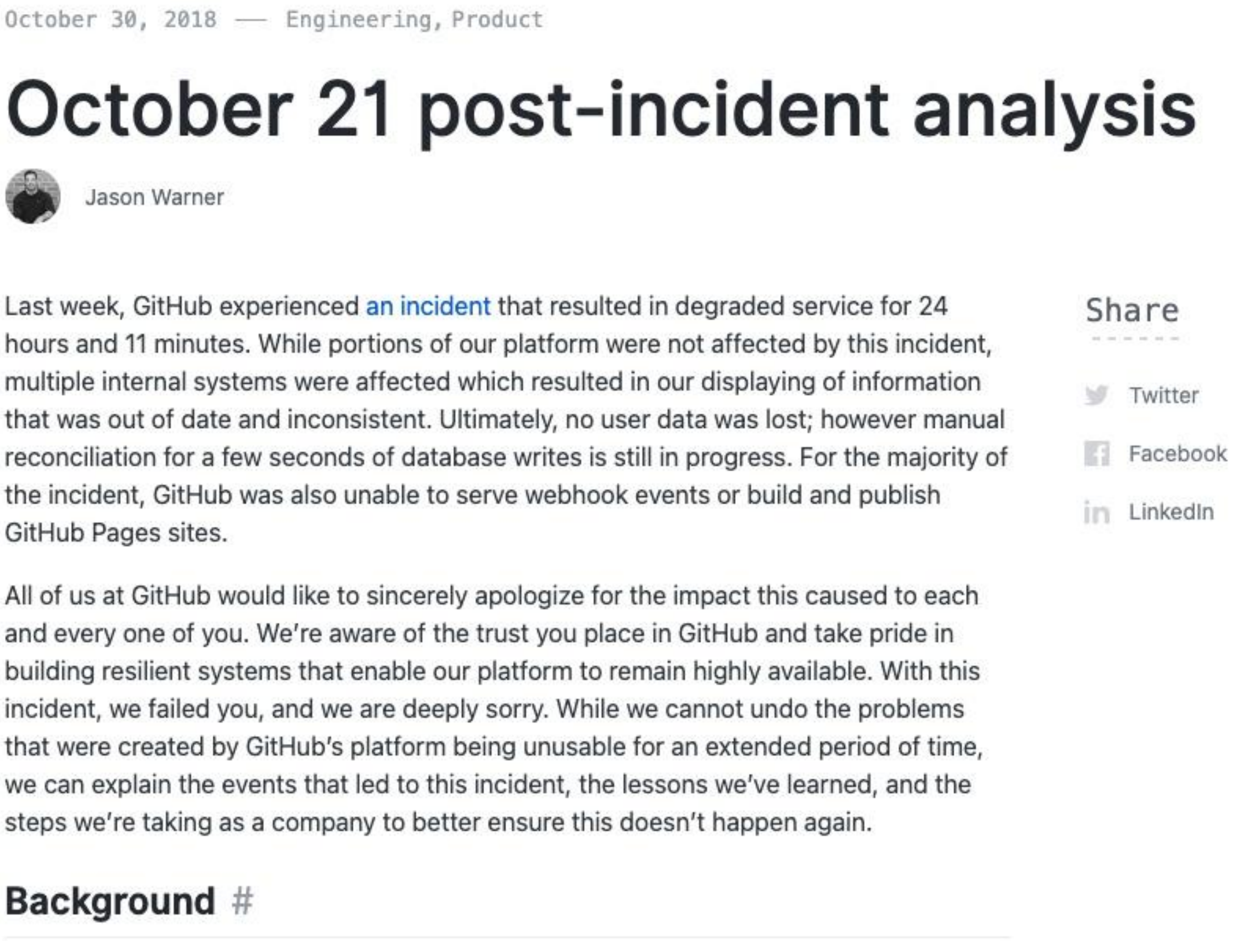
🎯 Le guide complet de gestion des incidents ITIL

Thèmes abordés

Background
Incident timeline / What Happened
Root cause
Workaround / Mitigation

Impact / Customers
Next steps / Action item
(Engineering & Architecture,
Process & Communications)

🎯 Exemples: GitHub / Elastic



October 30, 2018 — Engineering, Product

October 21 post-incident analysis

Jason Warner

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.

Background

First blood

Quand date le premier bug de Grace Hooper ?

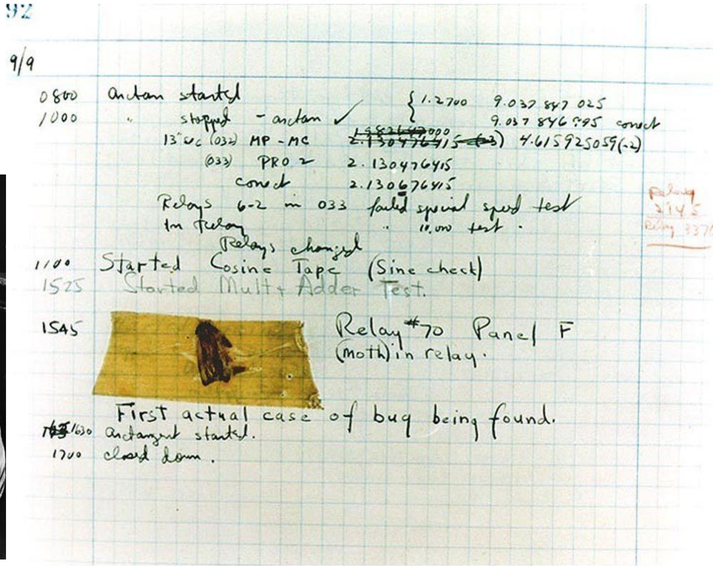
Un vrai bug (insecte)

Bureaux de EMCC

1947

Inventeur du COBOL

Photo # NH 96566-KN (Color) First Computer "Bug", 1947



92

9/9

0800 Antan started {1.2700 9.037 847 025
1000 " stopped - antan ✓ 9.037 846 995 conect
13" uc (032) MP - MC ~~1.982147000~~ ~~2.130476415~~ (-3) 4.615925059(-2)
(033) PRO 2 2.130476415
conect 2.130676415
Relays 6-2 in 033 failed special speed test
In relay 11,000 test.
Relays changed
1100 Started Cosine Tape (Sine check)
1525 Started Mult + Adder Test.
1545 Relay #70 Panel F
(moth) in relay.
First actual case of bug being found.
~~143~~/1630 antangent started.
1700 closed down.

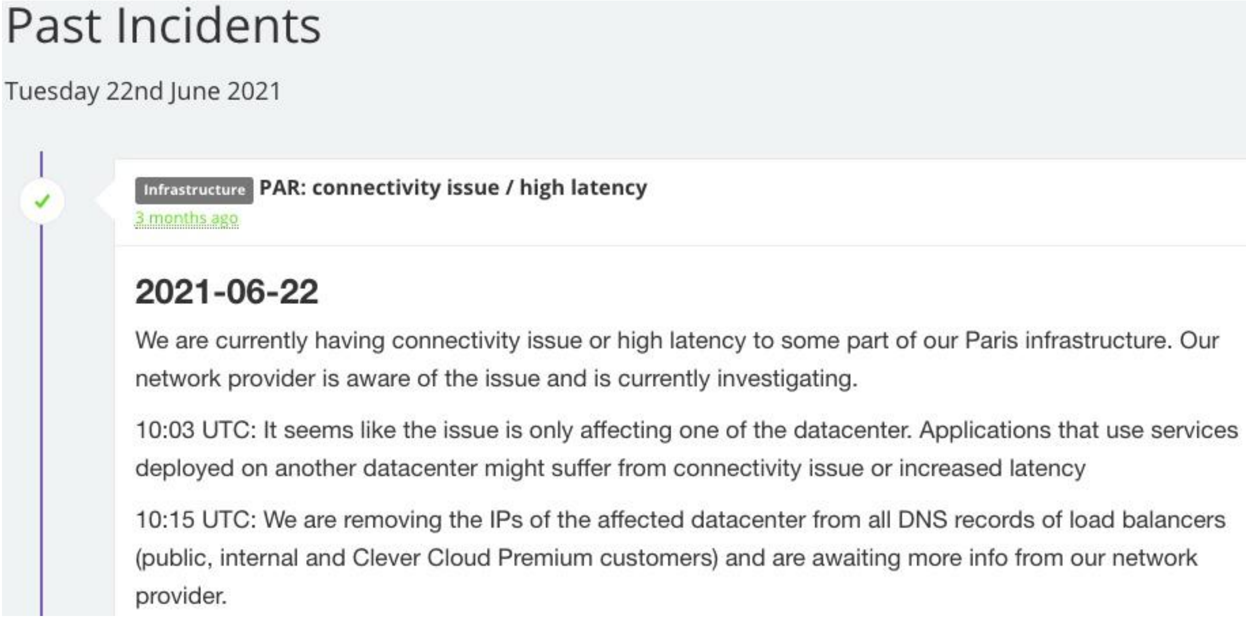
Relay 2145
Relay 337

Cette anecdote n'est pas "isolée"

22 Juin

2021

Inondation de la fibre optique d'un datacenter



Past Incidents

Tuesday 22nd June 2021

✓ Infrastructure **PAR: connectivity issue / high latency**

3 months ago

2021-06-22

We are currently having connectivity issue or high latency to some part of our Paris infrastructure. Our network provider is aware of the issue and is currently investigating.

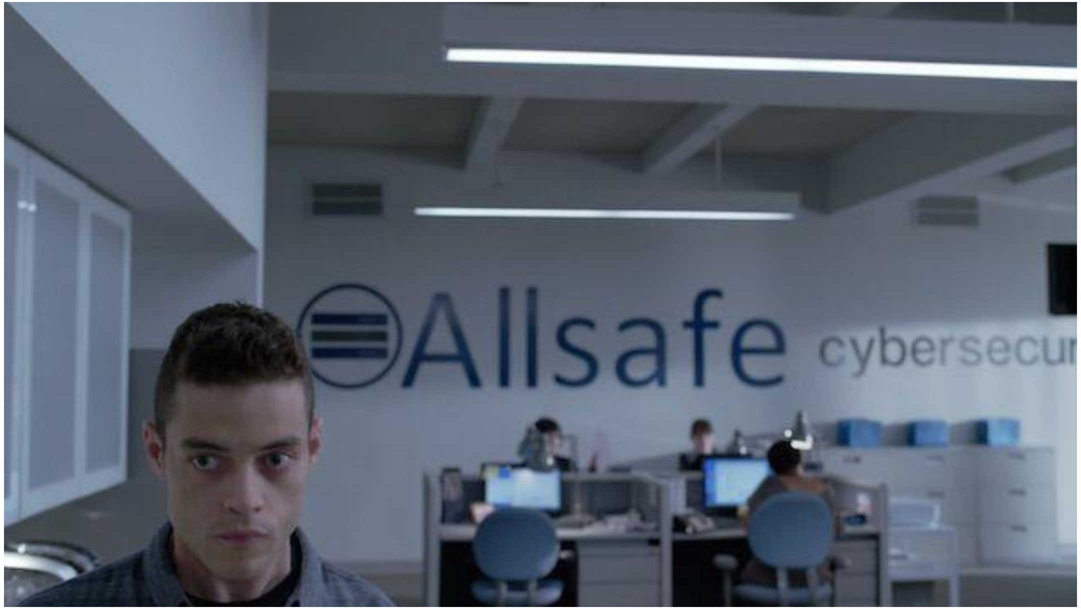
10:03 UTC: It seems like the issue is only affecting one of the datacenter. Applications that use services deployed on another datacenter might suffer from connectivity issue or increased latency

10:15 UTC: We are removing the IPs of the affected datacenter from all DNS records of load balancers (public, internal and Clever Cloud Premium customers) and are awaiting more info from our network provider.

Elliot

Elliot Alderson

Allsafe, great place to work 2021



P1 detected, tout le monde sur le pont

Money, money

Beware Prod behind

"With great power comes great responsibility"

Attention !

Ne mets pas tes doigts dans la prod, tu risques de te faire pincer très fort.

clever-cloud.com

AWS, GitLab, DigitalOcean

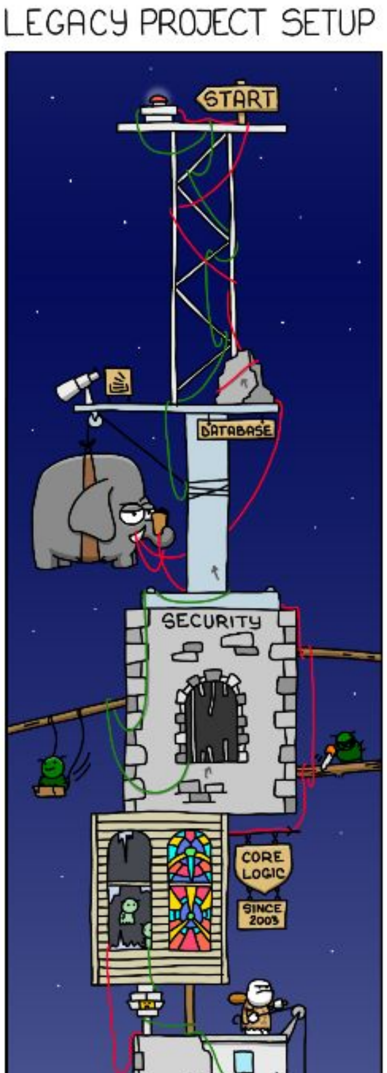
AWS / Drop database / Protect environment & RBAC https://aws.amazon.com/message/680587/	24/12/2012
Gitlab / Drop database / Use regularly RBAC and test regularly your backup ! https://about.gitlab.com/blog/2017/02/10/postmortem-of-database-outage-of-january-31/	31/01/2017
DigitalOcean / Drop database / Use regularly RBAC https://www.digitalocean.com/blog/update-on-the-april-5th-2017-outage/	05/04/2017

Legacy

Il était une fois.. (en 2020)

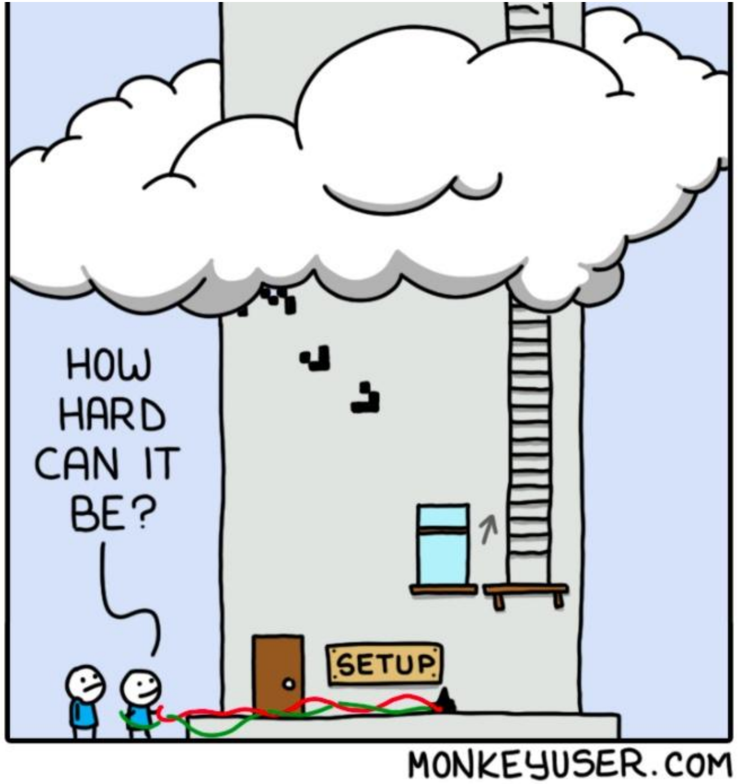
Quand soudain l'astreinte sonne

Legacy



LEGACY PROJECT SETUP
START
DATABASE
SECURITY
CORE LOGIC
SINCE 2003

Legacy..

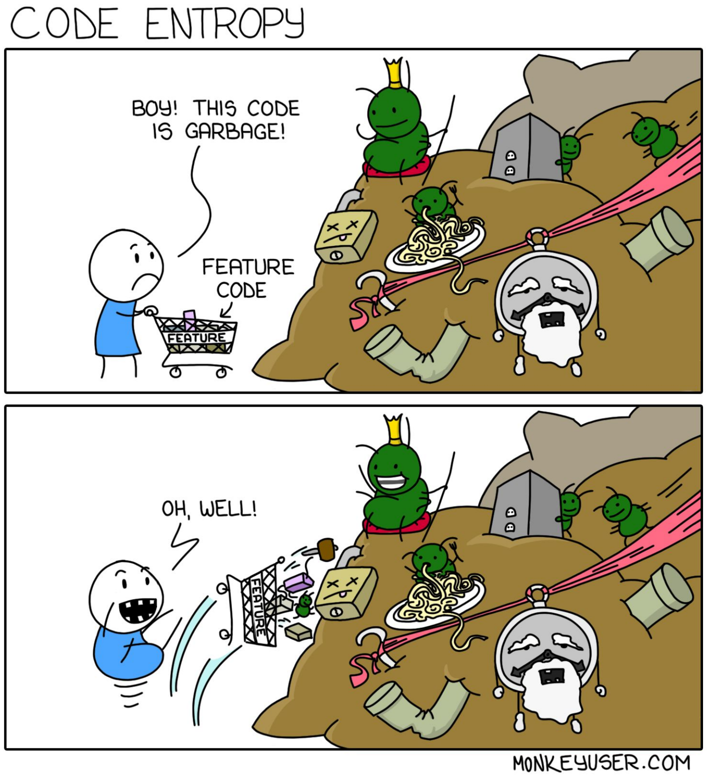


L'astreinte sonne à nouveau

TECH DEBT

THIS WAY!
WE ARE PROGRESSING SO FAST TOGETHER
THIS DOESN'T LOOK GOOD
TECH DEBT
MONKEYUSER.COM

Le bon et le mauvais patch: y voit un truc...



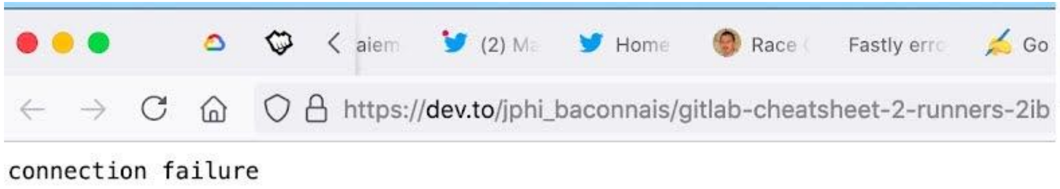
KISS

Original:	*Variante :*
Keep	*Keep*
It	*It*
Simple	*Super*
Stupid	*Simple*

Attention: Serveur

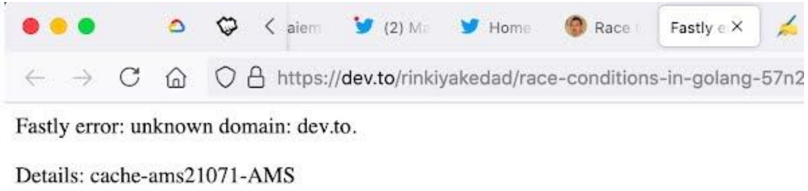
CDN

Marmiton



https://dev.to/jphi_baconnais/gitlab-cheatsheet-2-runners-2ib

```
connection failure
```

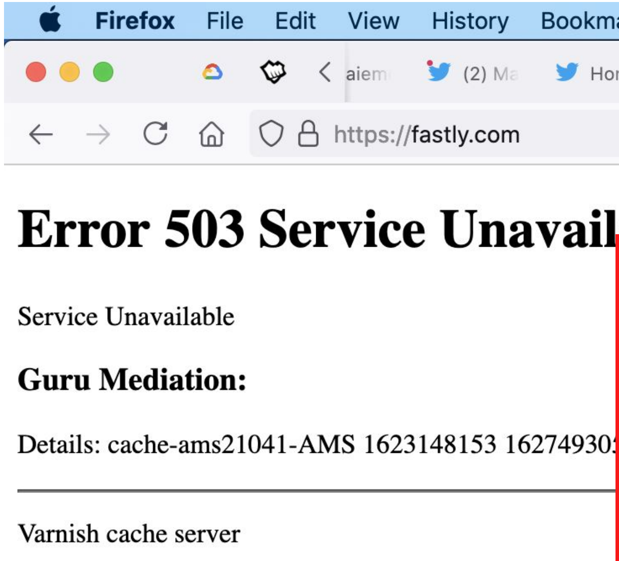


https://dev.to/rinkiyakedad/race-conditions-in-golang-57n2

Fastly error: unknown domain: dev.to.

Details: cache-ams21071-AMS

Fastly



Error 503 Service Unavailable

Service Unavailable

Guru Mediation:

Details: cache-ams21041-AMS 1623148153 162749305

Varnish cache server

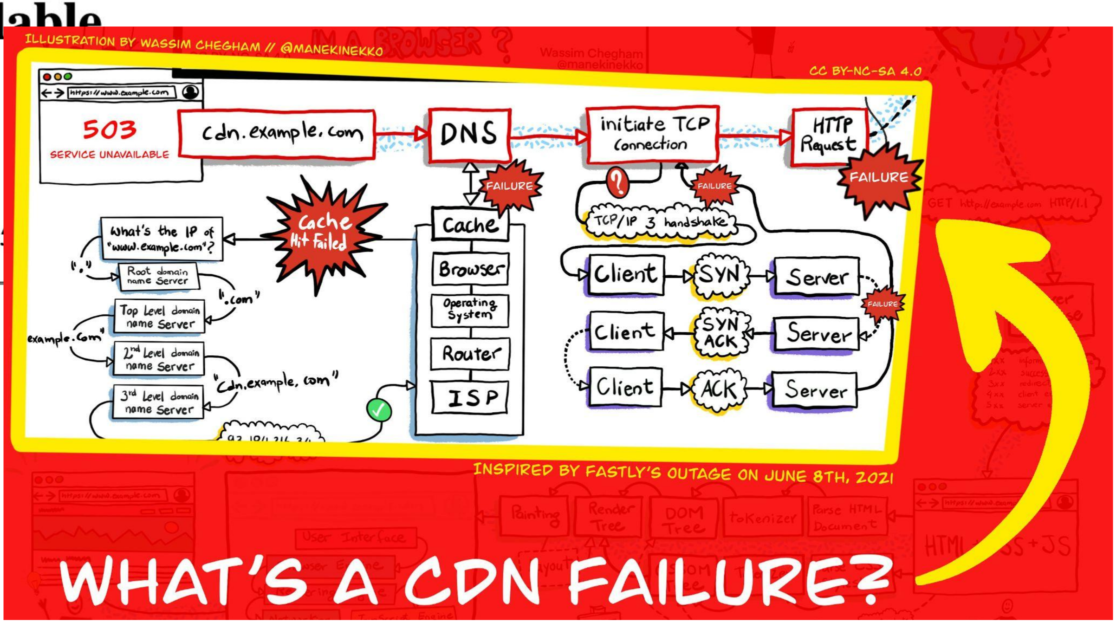
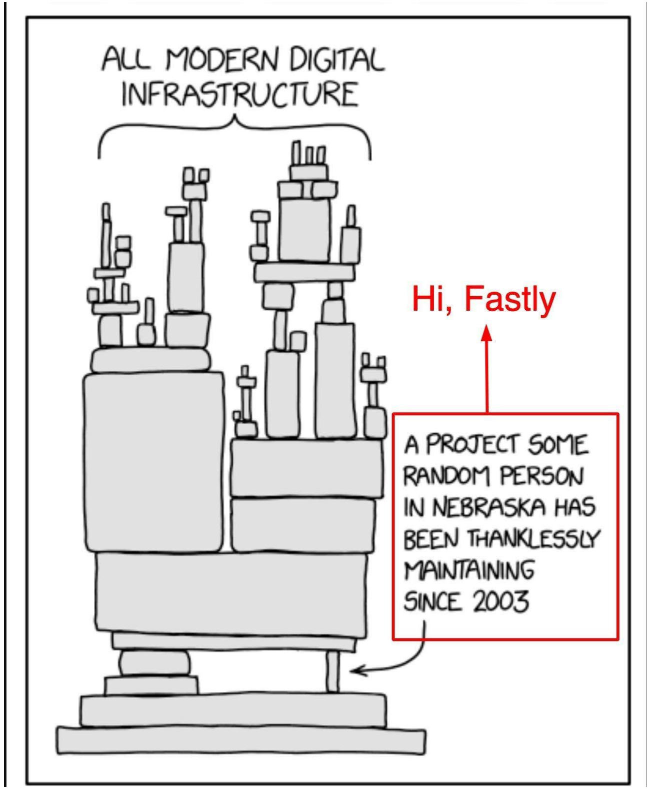


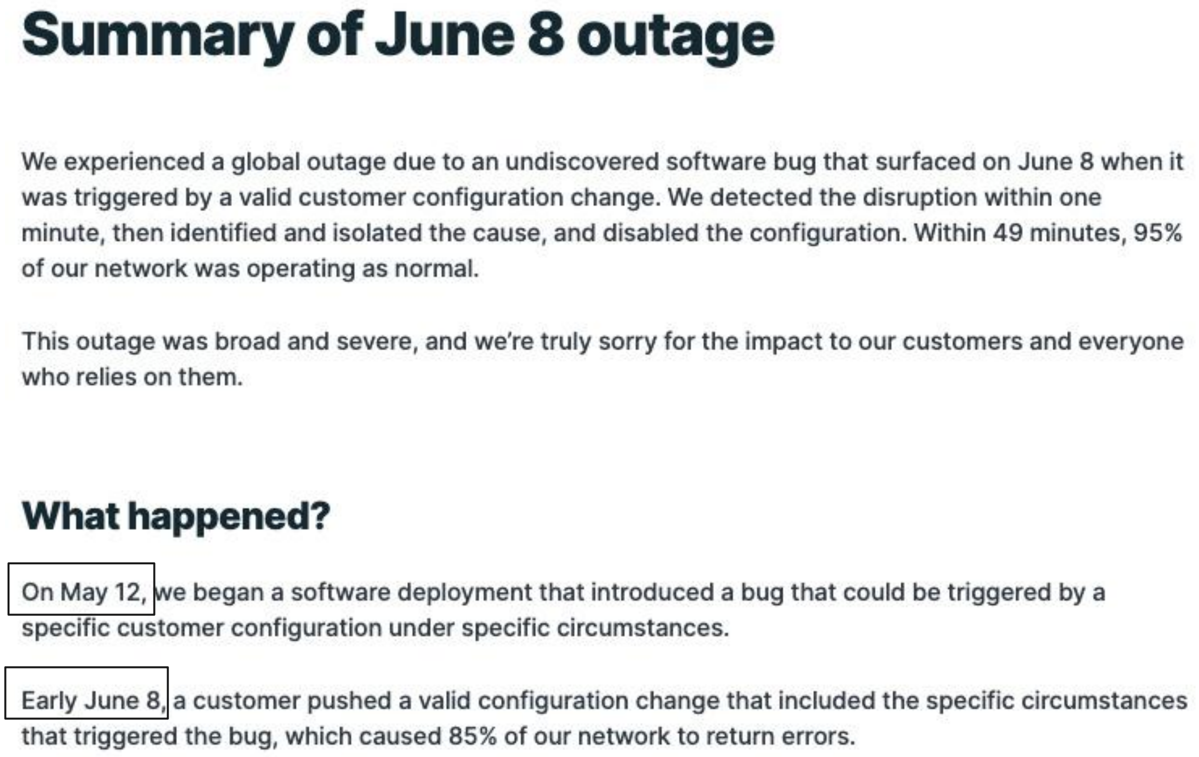
Schéma de @manekinekko

Resilience


ALL MODERN DIGITAL INFRASTRUCTURE
Hi, Fastly
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

🎯 Image trouvée sur la twittosphère

Details



Summary of June 8 outage

We experienced a global outage due to an undiscovered software bug that surfaced on June 8 when it was triggered by a valid customer configuration change. We detected the disruption within one minute, then identified and isolated the cause, and disabled the configuration. Within 49 minutes, 95% of our network was operating as normal.

This outage was broad and severe, and we're truly sorry for the impact to our customers and everyone who relies on them.

What happened?

On May 12, we began a software deployment that introduced a bug that could be triggered by a specific customer configuration under specific circumstances.

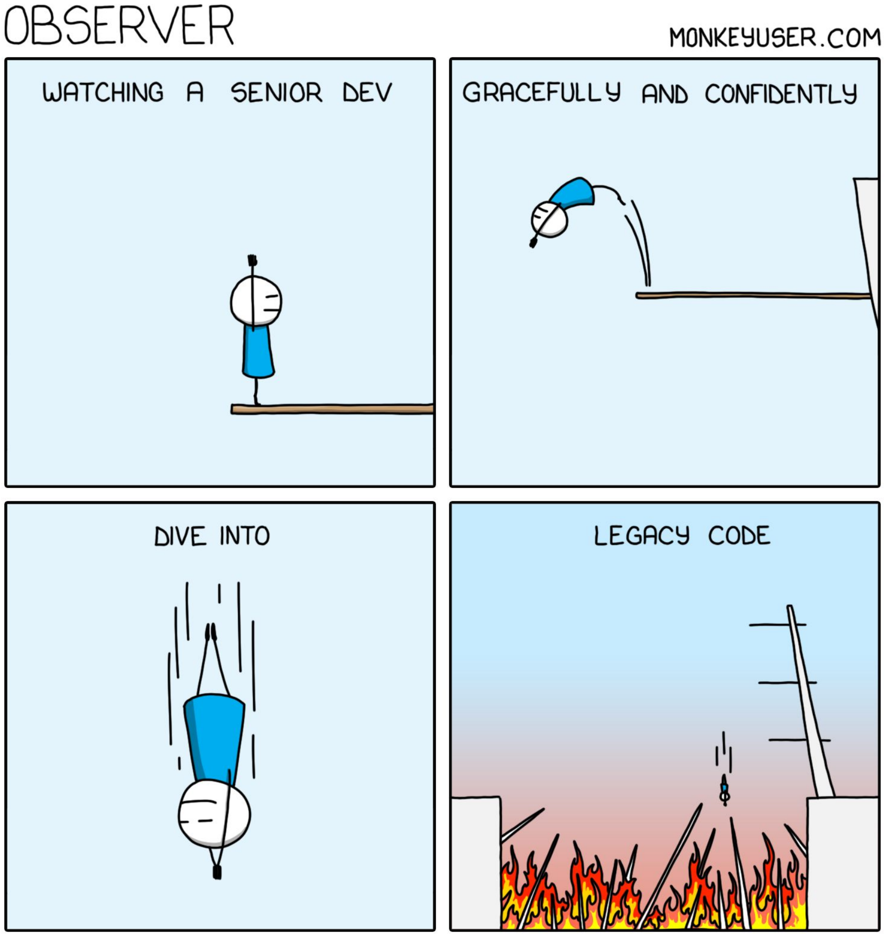
Early June 8, a customer pushed a valid configuration change that included the specific circumstances that triggered the bug, which caused 85% of our network to return errors.

Le retour du Legacy

L'astreinte sonne.. ...

Il était une autre fois.. (en 2021)

Senior vs Junior



Blast Effect

Apache Zookeeper

- Créé en même temps qu'à Hadoop
- Base de beaucoup de systèmes distribuées (Kafka v<2.8 / KIP-500)
- Ancêtre de consul/etcd
- Key-Value Store hiérarchique

- Très bons résultats aux tests Jepsen

- Avantages
    - Suivre des changements sur un path (watcher)
    - Noeuds éphémères
    - Notion de sessions

Architecture

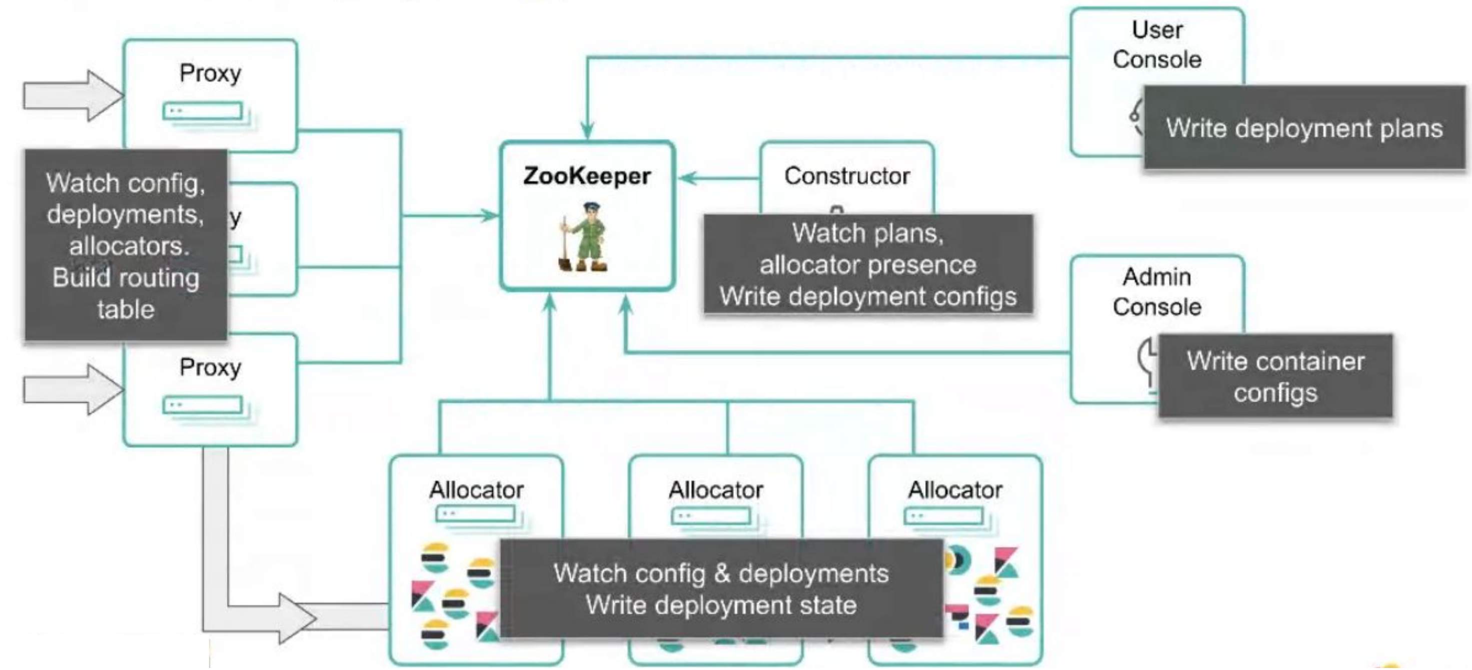


Schéma issu du meetup Toulouse JUG: War Story, comment les pauses du GC ont pété la prod

Same JVM, Shoot again

Les limites de la technologie

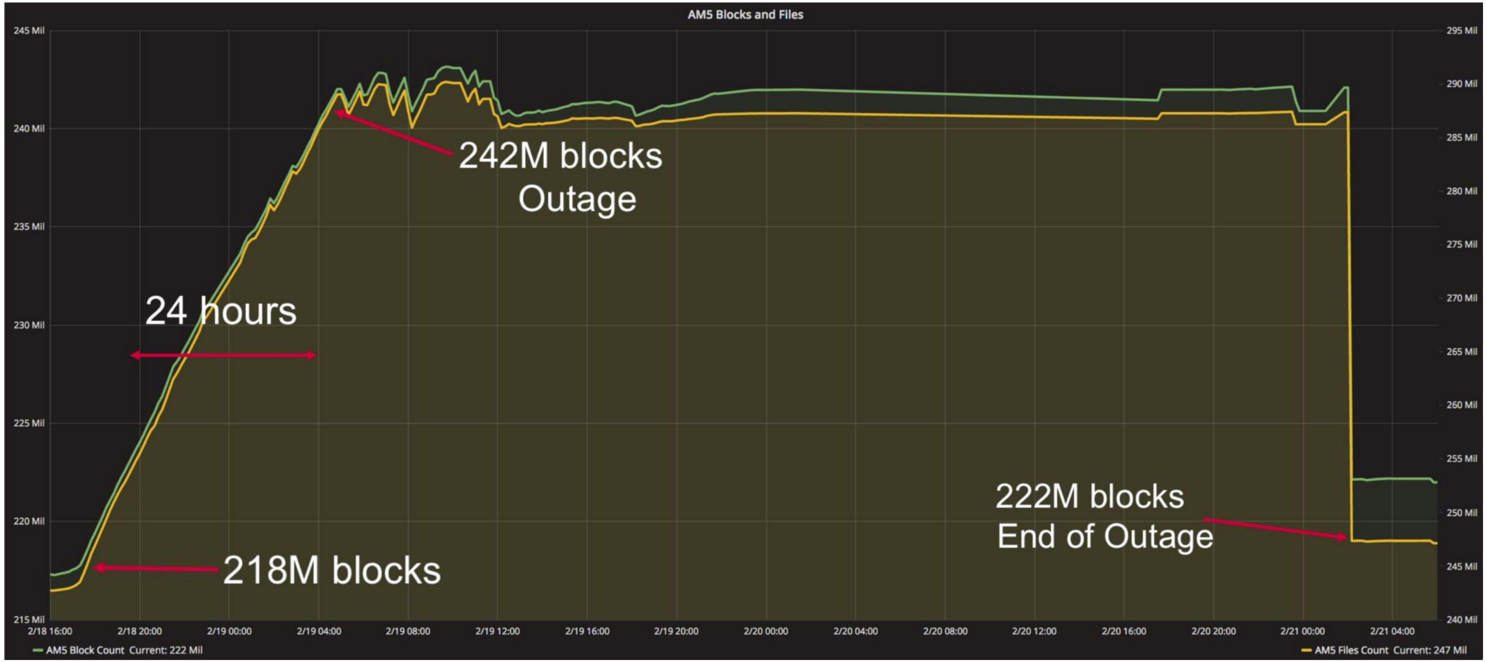
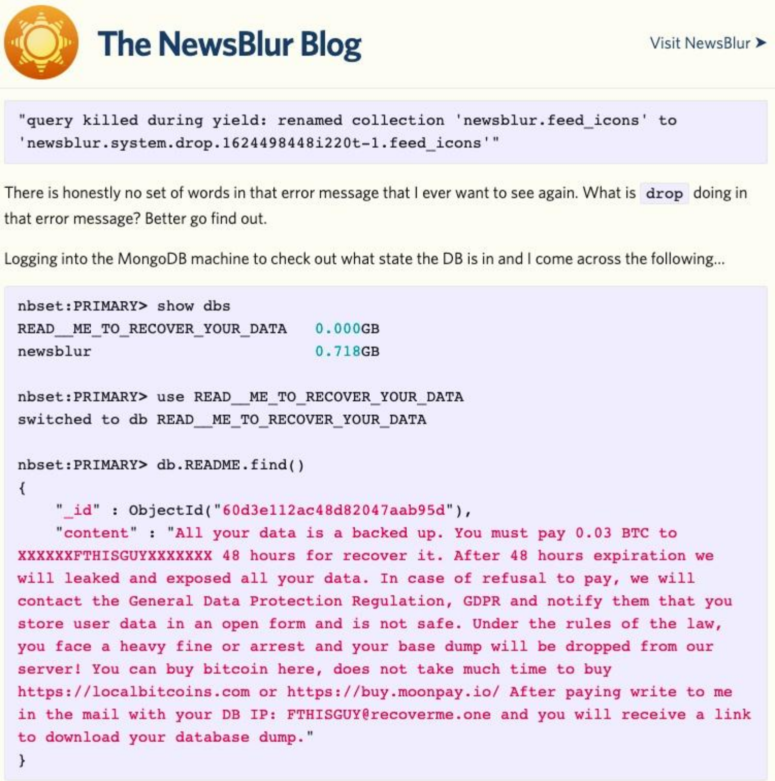


Schéma extrait de Medium : +40PB per year — The challenge of data growth at Criteo
Hadoop sous pression Retour sur une année d'exploitation à Criteo (Rémy Saissy)

NewsBlur

Attaque ? Pas de sécurisation sur la DEV->PRD



```
"query killed during yield: renamed collection 'newsblur.feed_icons' to
'newsblur.system.drop.1624498448i220t-1.feed_icons'"
```

There is honestly no set of words in that error message that I ever want to see again. What is `drop` doing in that error message? Better go find out.

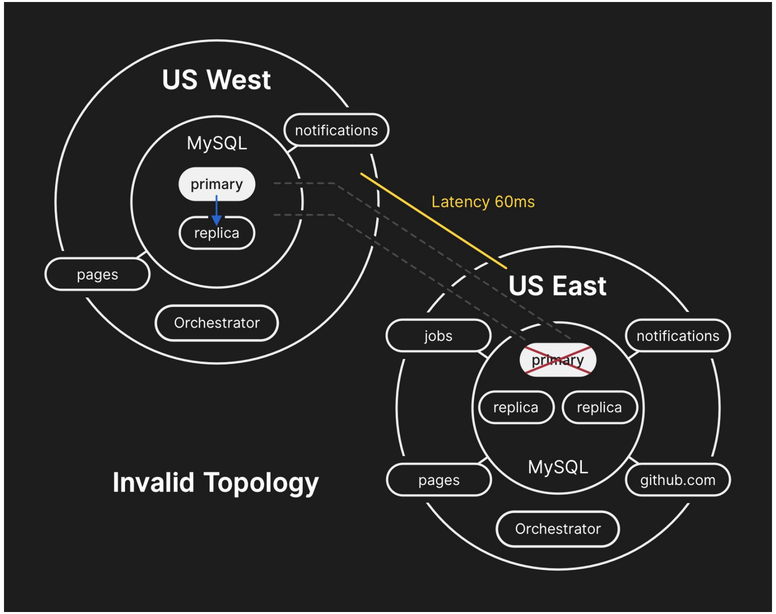
Logging into the MongoDB machine to check out what state the DB is in and I come across the following...

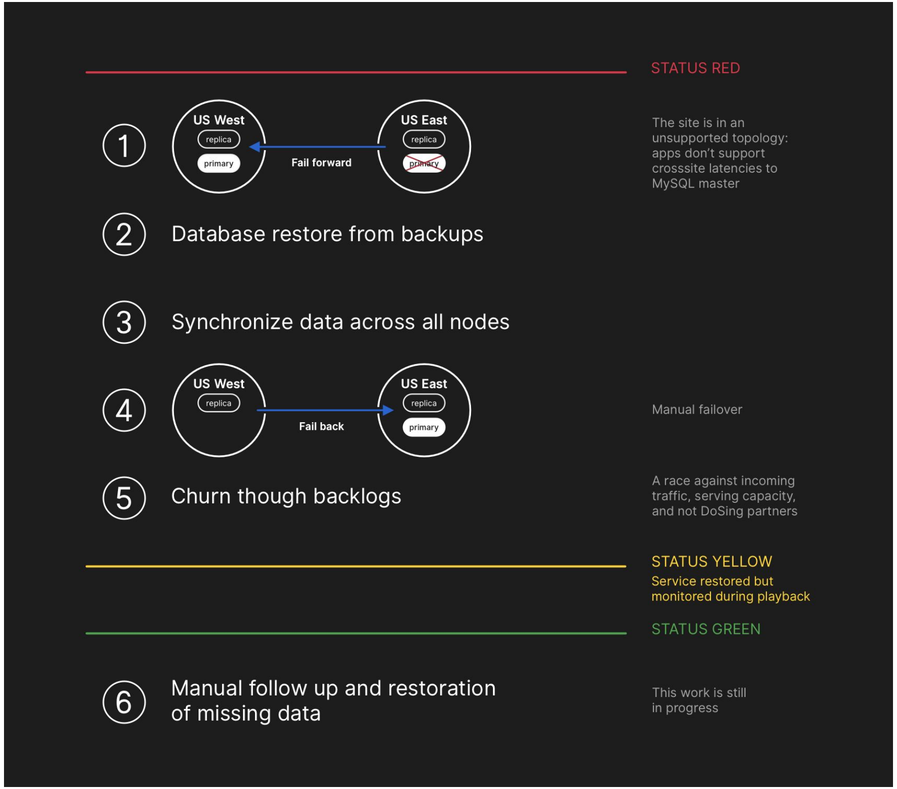
```
nbset:PRIMARY> show dbs
READ__ME_TO_RECOVER_YOUR_DATA    0.000GB
newsblur                         0.718GB

nbset:PRIMARY> use READ__ME_TO_RECOVER_YOUR_DATA
switched to db READ__ME_TO_RECOVER_YOUR_DATA

nbset:PRIMARY> db.README.find()
{
    "_id" : ObjectId("60d3e112ac48d82047aab95d"),
    "content" : "All your data is a backed up. You must pay 0.03 BTC to
XXXXXXFTHISGUYXXXXXXX 48 hours for recover it. After 48 hours expiration we
will leaked and exposed all your data. In case of refusal to pay, we will
contact the General Data Protection Regulation, GDPR and notify them that you
store user data in an open form and is not safe. Under the rules of the law,
you face a heavy fine or arrest and your base dump will be dropped from our
server! You can buy bitcoin here, does not take much time to buy
https://localbitcoins.com or https://buy.moonpay.io/ After paying write to me
in the mail with your DB IP: FTHISGUY@recoverme.one and you will receive a link
to download your database dump."
}
```

🎯 Capture d'écran du post: How a Docker footgun led to a vandal deleting NewsBlur's MongoDB database

Split-Brain

GitHub





STATUS RED

1

The site is in an unsupported topology: apps don't support crosssite latencies to MySQL master

2 Database restore from backups

3 Synchronize data across all nodes

4

Manual failover

5 Churn though backlogs

A race against incoming traffic, serving capacity, and not DoSing partners

STATUS YELLOW
Service restored but monitored during playback

STATUS GREEN

6 Manual follow up and restoration of missing data

This work is still in progress

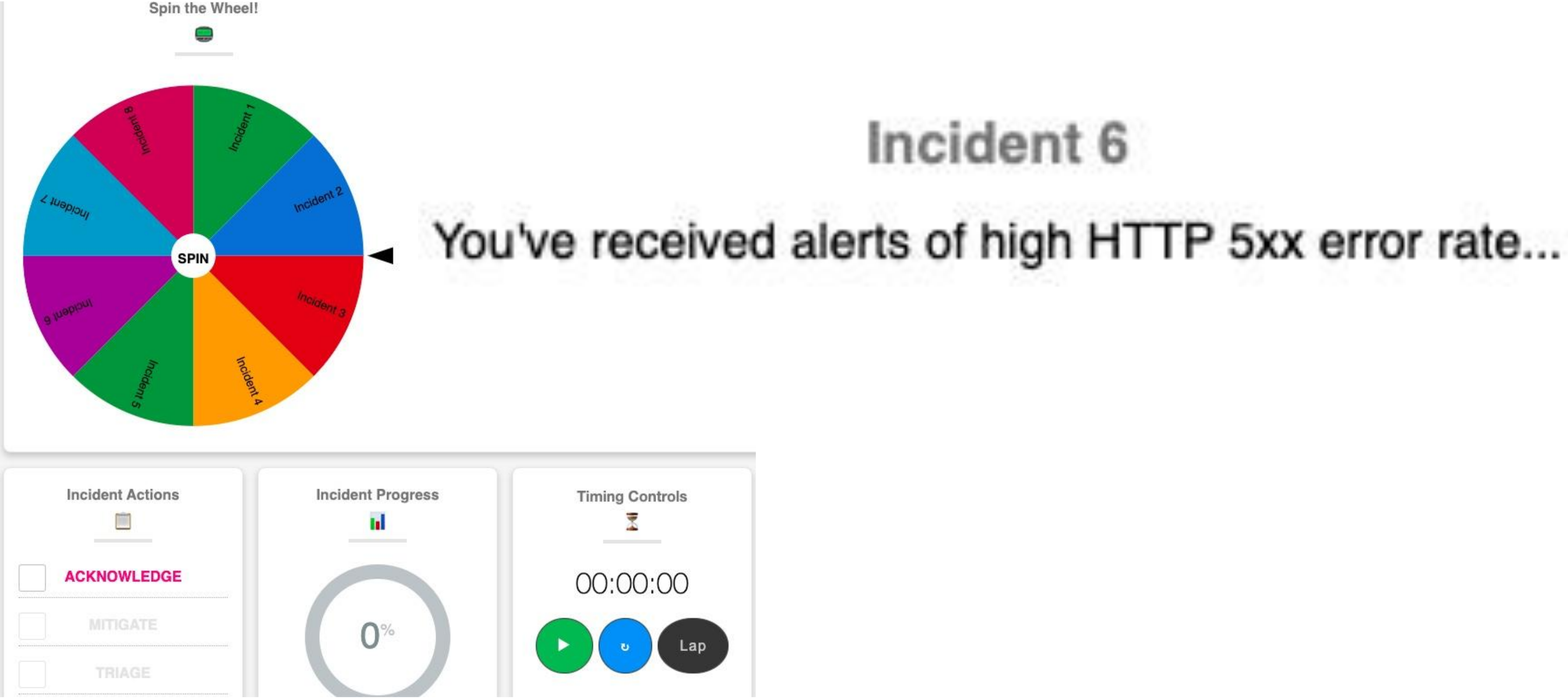
🎯 Schémas issus du blog post : [October 21 post-incident analysis]

Conclusion

En bref

- SRE Blameless culture / With great power comes great responsibility
- QA / Chaos tests
- Former vos équipes : Game day / Wheel of Misfortune
- Tester fréquemment vos backups & hors d'un Drive pour les données médicaux ;-)
- CI/CD & DevSecOps pipelines
- Implémenter & monitorer les metrics importantes issues des incidents (GC Pause, etc)

Wheel of Misfortune



Incident 6

You've received alerts of high HTTP 5xx error rate...

🎯 Capture d'écran de la "Roue de la malchance": https://dastergon.gr/wheel-of-misfortune/

Pour aller plus loin

Entreprise / Incident	Date
AWS / Drop database https://aws.amazon.com/message/680587/	24/12/2012
Gitlab / Drop database https://about.gitlab.com/blog/2017/02/10/postmortem-of-database-outage-of-january-31/	31/01/2017
DigitalOcean / Drop database https://www.digitalocean.com/blog/update-on-the-april-5th-2017-outage/	05/04/2017
OVHCloud / Watercooling https://blog.ovh.com/fr/blog/hebergements-web-post-mortem-incident-29-juin-2017/	29/06/2017
Criteo / Datalake HDFS limits + Friday night & New-comer on-call https://medium.com/criteo-engineering/40pb-per-year-the-challenge-of-data-growth-at-criteo-5d5b73ec5294	18/02/2018
GitHub / DataBase Split Brain https://github.blog/2018-10-30-oct21-post-incident-analysis/	30/10/2018
ElasticCloud / Proxy Layer+Zookeeper https://www.elastic.co/blog/elastic-cloud-january-18-2019-incident-report	18/01/2019
CleverCloud / Orages à Paris (Redondance/Dépendance au network provider) https://www.clevercloudstatus.com/incident/376	22/06/2021
NewsBlur / Docker PROD DB exposé sur le net, sans mot de passe https://blog.newsblur.com/2021/06/28/story-of-a-hacking/	23/06/2021

Jug Summer Camp

-enjoy it-

Merci pour votre attention !

SERLi OVHcloud elastic

Avez-vous des questions ?

🧑‍🏫🔊 N'oubliez pas de me donner votre avis sur cette session:

📃 https://s.42l.fr/iaas

👍 Rejoignez-nous sur https://careers.ovhcloud.com

Analogie

« Nul n’est censé ignorer la loi »

Ma devise

« Nul développeur n’est censé ignorer la sécurité »