



Kris Buytaert

Help, our datacenter is on fire !

OTIy

NLUUG

November 2022

March 10 , 2021

- 03:59 Incoming Phone call
- “Our Datacenter is On fire”
- 5 minutes of waking up..
going downstairs trying to
realize what was just said to
me, partial panic , partial ..
it can't be THAT bad



Fire

 **Octave Klaba** 
@olesovhcom

We have a major incident on SBG2. The fire declared in the building. Firefighters were immediately on the scene but could not control the fire in SBG2. The whole site has been isolated which impacts all services in SGB1-4. We recommend to activate your Disaster Recovery Plan.

3:42 AM · Mar 10, 2021 · Twitter for iPhone

2,374 Retweets 1,287 Quote Tweets 4,726 Likes

 **Octave Klaba**  @olesovhcom · Mar 10

Replying to @olesovhcom

Update 5:20pm. Everybody is safe.
Fire destroyed SBG2. A part of SBG1 is destroyed. Firefighters are protecting SBG3. no impact SBG4.

 136  674  1.7K 

 **Octave Klaba**  @olesovhcom · Mar 10

Update 7:20am

Fire is over. Firefighters continue to cool the buildings with the water. We don't have the access to the site. That is why SBG1, SBG3, SBG4 won't be restarted today.

 113  464  1.2K 

 **Octave Klaba**  @olesovhcom · Mar 10

Update 10am. (I'm there).

We finished to shutdown the UPS in SBG3. Now they are off. We are looking to enter into SBG3 and check the servers. The goal is to create a plan to restart , at least SBG3/SBG4, maybe SBG1. To do so, we need to check the network rooms too.

- I used to be a Dev,
- Then Became an Op
- CTO and Open Source Consultant @inuits.eu
- Everything is a freaking DNS Problem
- Evangelizing devops
- Organiser of #devopsdays, #cfgmngmtcamp, #loadays,

Impact :

- 3.6 million websites across 464.000 distinct domains

- <https://news.netcraft.com/archives/2021/03/10/ovh-fire.html>



Immediate Assessment ?

- What Customer Facing Production services have we lost ?
- What Internal Production services have we lost ?
 - Do we need any of those to recover ?
- What are the priorities ?
 - Which services
 - 24/7 money generating platforms
 - Office hours tooling
 - What parts of these services
 - Basic Service, no redundancy

Our Immediate visible Impact

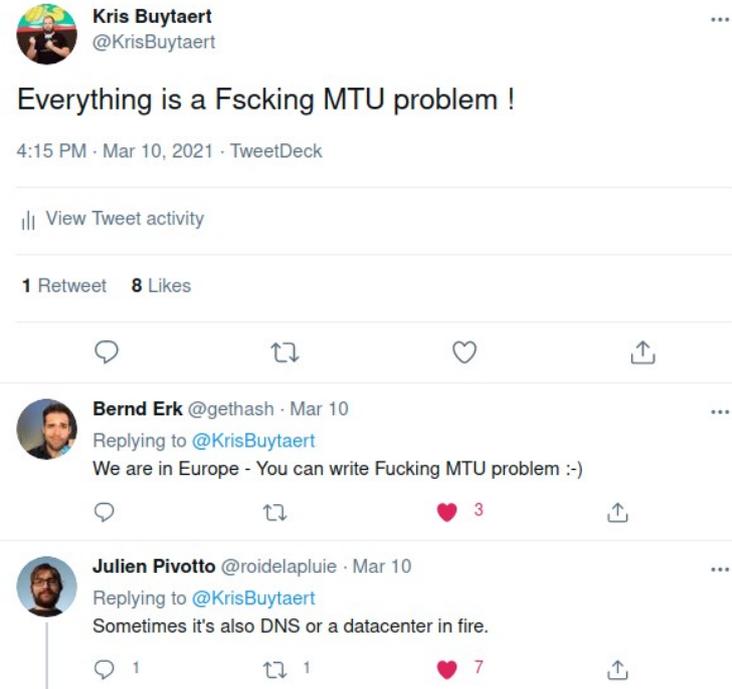
- Multiple customers failed over automatically to the 2ndary DC
- 1 VIP address failed to move
- 1 platform was spread over SBG2 & SBG1
- 1 DR site destroyed

Failing VIP

- 1 VIP address failed to move (API call failed)
- Manual action to use a different public IP for the loadbalancer
- Updated DNS records
- Except for the 1 domain the customer controlled.
- This cost us about 60 minutes before we could reach someone with the right credentials to update the dns
- Some customer devices had the VIP address hardcoded :(ul>- Manual interventions needed :(

Customer Impact

- 1 platform was spread over SBG2 & SBG1
- Rebootstrapped the platform on a different ISP, restored the backup
- At 0900 am platform was ready for use,
- We were improving performance of the platform
- Customer wouldn't have known if we didn't tell them



Total Loss:

- We lost 13 physical servers
- 135 vm's



Root Cause

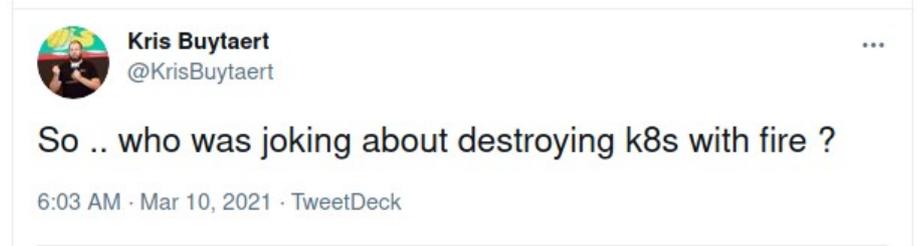
- Klabá does not say the UPS is the definite cause. "We don't have all the answers today," he said. The OVHcloud staff responded to alarms at 11.42 pm on Tuesday, but the affected part of the data center had already filled with smoke: "Two minutes after, they took the decision to leave, because it was too dangerous."
- The firefighter's thermal cameras found UPS7 and UPS8 on fire, but further data will be extracted from on-site cameras: "We have 300 cameras in Strasbourg," said Klabá. "We expect to have all the answers about how it started. We will give you all the information."
- <https://www.datacenterdynamics.com/en/news/ovh-fire-octave-klaba-says-ups-systems-were-ablaze/>

Our Impact

NextCloud

- Gitlab
- Lots of development environments
- Lots of Test nodes
- Partial K8s clusters

olly



We survived



- 18 hours of work
- 2 engineers at night
- 4-5 during daytime
- Recovery
- Communication
- Restoring Backups

At 0910

- Another customer seemed to have issues accessing some files
- We had overlooked 1 gluster cluster that seemed healthy but wasn't
- Switched to a single node mount point
(don't try this at home)

No data was lost



olly

3 months later

- Users complain about accessing files
- Files used to be in mysql (long standing bug)
- Devs fixed bug, never told ops
- Files were not backed up
- GDPR compliant again :)

Phase II

- After 0900 we started restoring our internal tools from backup
- Verified all the secondary nodes would be making backups
- Verified we had sufficient disk space left
- “Shoemaker always wears the worst shoes”
- Challenge : Hardware !

Hardware Landrush

- Everybody needed new hardware
- We had spare hardware ready for a new platform in the wrong DC of a different supplier . Temporary OK
- We spun up new hardware at a different cloud supplier



Wrong decisions

- On day 2 we wasted at least 6 hours trying spinning up a new cloud supplier
- Figuring out their network and redundancy strategies was not the right time
- We could scale on the existing cloud suppliers

Phase III

- Restoring the rest of the services
- What boxes would we be getting back ?
- What boxes were foobar ?

- (in the end we only got 3 physical servers back which were in SBC3), the rest was lost
- Priority on making sure all pipeline promotions of actively developed platforms were running again
- Hardware availability

Your NEW DR Plan

- You have just lost your DR site
- You need to plan again
 - Do we take the risk for now ?
 - Do we respin a new DR ?
- IAC + Datasync makes DR Trivial

What saved us ?

- Real Infrastructure as Code
- Architecture
- Backups
- Fast escalation

Real Infrastructure as Code

- Desired State => Puppetize all the things
- 100% control , no 3rd parties involved
 - Provision vm
 - Deploy applications
 - Deploy database schemas
- No handovers , 1 person can deploy this, no other teams involved
- Exported resources for Loadbalancers (haproxy), Monitoring (icinga), Databases (mysql)etc
 - Heavy automation
- All deployed versions are available in (yum) repositories

- Everything is a Pipeline Puppet, Hiera, DNS, ... (jenkins)



CloudNative vs Cloudnaive

- We don't own hardware
- Baremetal on demand at \$ < AWS
 - Hetzner / ovh / ...
- Spin up 120 seconds, decommission when unneeded.
- vm definitions are in code
- Ansible to bootstrap, Puppet for Desired state



Multi Cloud

- OVH
- Hetzner
- ...
- Workload is spread
- Customer DR is in other Supplier



Multi Datacenter

- OVH
 - SBG, GRA, RBX, ..
- Hetzner
 - FSN X, HEL

CloudNative ?

- Critical Customer services are build redundantly, (corosync, haproxy, mysql, elastic)
- Losing a metal ~~shoud~~ does not have an impact
- We can rebootstrap our nodes (kickstart+ puppet + ansible)
- All UGC data we know about is backed up (rsnapshot)

Cloud Agnostic

- A vm is a vm is a vm
- Vendor Specific Features
 - e.g VIP addresses are abstracted in code
- Yes we can deploy on Azure/AWS

olly



Limited Blast Radius

- Dedicated stacks per Customer/Project
 - Single Purpose vm's
 - Single Purpose Database Clusters
 - Single Purpose Storage Clusters
 - Single Purpose Loadbalancers
- => Reduced impact, Reduced complexity,

Clusters are Multi DC

- Depending on scale , preference on HA vs actual Loadbalancing
- Pinning of resources within 1 DC
- Lessons Learned: we need Multi Campus

HA Applications

- No local file usage
 - If local files -> {gluster,drbd,...}
- MySQL replication
- Elastic multiple nodes (if we can't generate the data)
- Multiple instances of rabbitmq

HA Storage

- VM's on Local Disk
 - 100% puppetized => no differences in config/ software deployment
 - “Disposable”
- Storage distributed
 - Gluster, DRBD, Ceph
- No Raid

Backups

- Are both on site
- And off Site

Documentation

- Mkdocs based
- Git repo
- Locally available

Lessons Learned

- Make sure you can always provision more resources on multiple providers
- A Datacenter is not always Datacenter, even with different buildings you need Campus redundancy
- Fix the shoemaker problem
- Everything IS a fscking dns problem
 - U WANT control over the DNS of the production sites you run
- Or an MTU problem

Thnx

Thnx to everyone who helped
out during the outage !

o11y

March 10 , 2021



otly

CONTACT

o11y

Essensteinweg 31

2930 Brasschaat

info@o11y.eu

kris@o11y.eu

o11y