

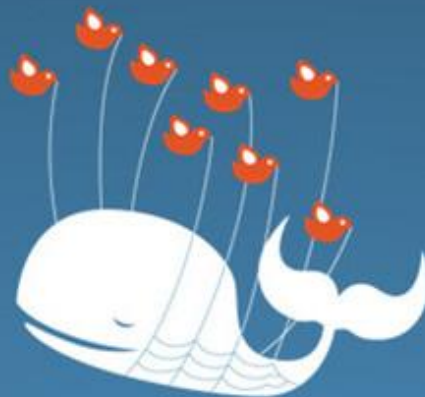


# Resilience

A question of mindset, culture, and partnerships

Dr Thomas Boltze, Head of Cloud and Engineering Excellence





# Twitter is over capacity.

Please wait a moment and try again. For more information, check out **Twitter Status**.

[Bahasa Indonesia](#) [Bahasa Melayu](#) [Deutsch](#) [English](#) [Español](#) [Filipino](#) [Français](#) [Italiano](#) [Nederlands](#) [Português](#) [Türkçe](#)

[Русский](#) [हिन्दी](#) [日本語](#) [简体中文](#) [繁體中文](#) [한국어](#)

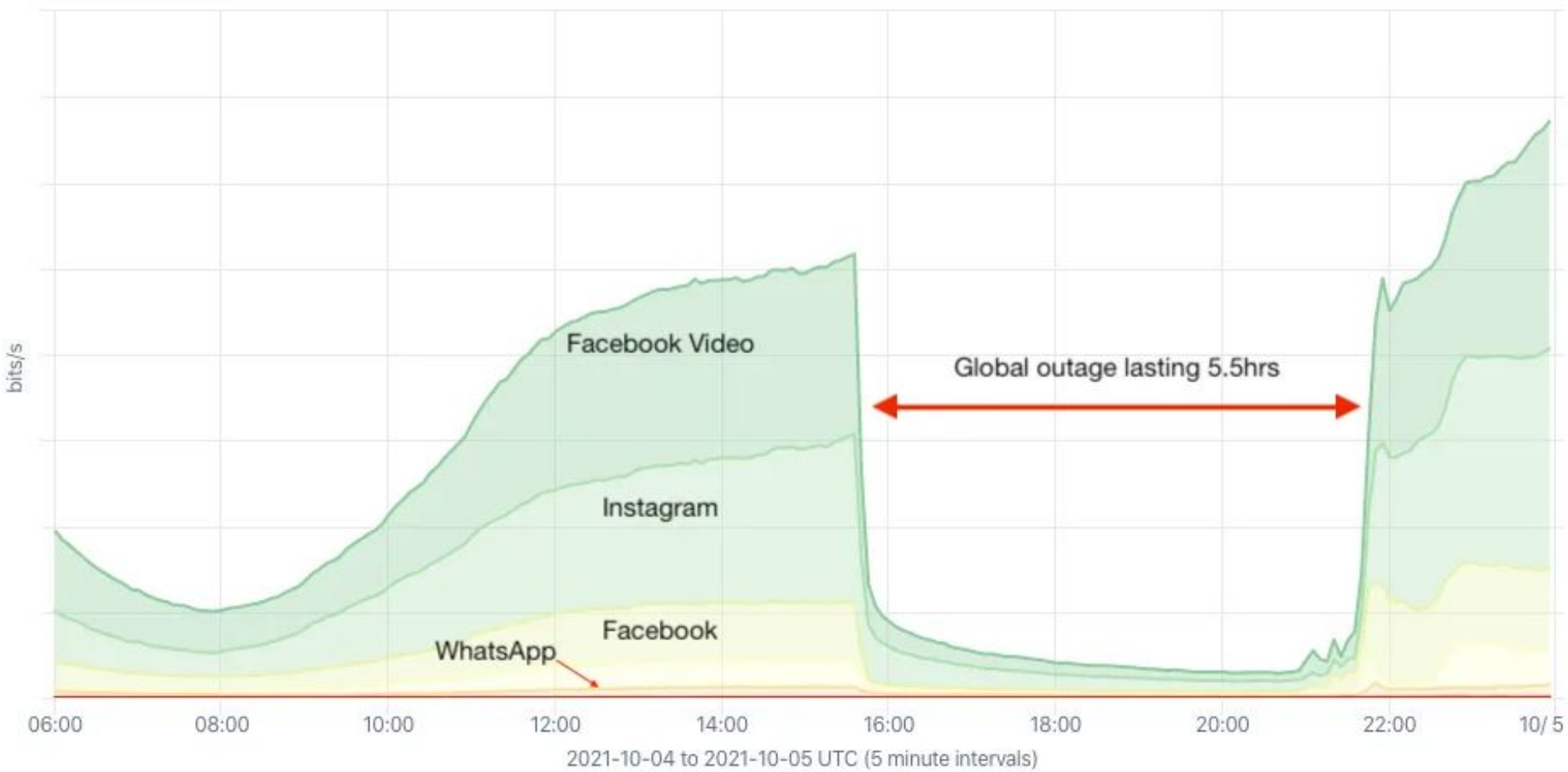
© 2012 Twitter [About](#) [Help](#) [Status](#)

# Top OTT Service by Average bits/s

Oct 04, 2021 06:00 to Oct 05, 2021 00:00 (18h)

# Internet Traffic served by Facebook

## Global outage 4-Oct-2021



## **human error**

people make mistakes, sometimes leading to large scale outages

## **incomplete monitoring**

monitoring for errors is incomplete, alerting too late or too little

## **faulty deployments**

rolling deployments resulted in errors that the monitoring did not catch, which then blew up services across multiple regions

# resilience

“the capacity to withstand or to recover quickly from difficulties”

“The ability of an information system to continue to: (i) operate under adverse conditions or stress, even if in a degraded or debilitated state, while maintaining essential operational capabilities; and (ii) recover to an effective operational posture in a time frame consistent with mission needs.” [NIST]

**what does good look like?**

can your systems withstand AZ / datacenter / region failures?

can your systems withstand loss of core services?

do you run game days?

do you run chaos monkey – in production?

do you stress test your systems on a regular basis?

are your change sets small?



# mindset

“a person’s or group’s way of thinking and their opinions”

## the beginning



three availability zones  
microservices, events  
containers &  
orchestration  
RDS (multi-az)  
monitoring, alerting ...

## testing



full load on the  
system  
kill AZ  
measure impact  
fix findings

## real world



real thermal event  
took down one AZ  
system kept on  
processing payments

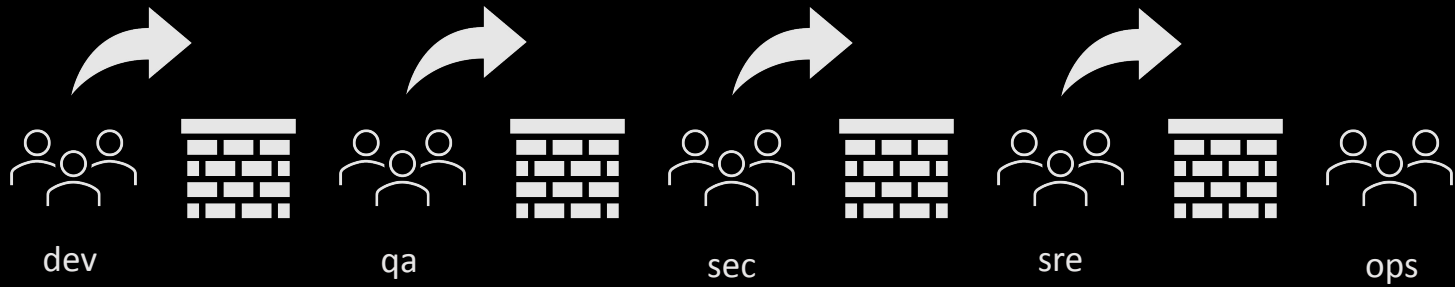
## next



multi-region  
multi-cloud  
more failure scenarios

# culture

“the set of behaviours that get rewarded, tolerated or sanctioned”



**somebody else's problem**

# observability

“real-time, everyone, everything”

# curiosity

“an eager desire to learn”

# everything as code

“don't try this without”

# focus

“don't be like Berlin”



# shared responsibility model

“every party plays a role in resiliency”

# Payments Hub



Consume at  
arm's length



Self-hosted,  
specific guarantees



SaaS, very close  
partnership



**worked for us**  
blameless post-mortems  
design for failure  
democratise tooling  
automate testing  
involve everyone  
understand past failures  
speak up  
iterate often  
focus



**did not**  
hunt for the guilty person  
isolate in one department  
build your own tooling  
use method hope  
“not my problem”  
siloe’d thinking  
change approval boards

# every incident is an opportunity to learn

Understand your current capabilities,  
mindset and culture

Invest in training & learning

Learn the tools

Allow your teams to design for failure

Automate everything

**thank you!**