O P T I M I S N G A R C H I T E C T U R E F O R O P E R A T I O N A L C O N C E R N S

PRESENTED BY:**RYAN FRENCH** LEAD CLOUD PLATFORM ENGINEER 11TH MARCH 2020



# A B O U T Y O U V I E W

- Joint venture between some of the UK's leading media and telecoms businesses
- Content discovery platform with content from some of the biggest names in media

youview

# A B O U T Y O U V I E W

- Platform behind BT TV, TalkTalk and Sony Android TVs for millions of homes across the UK
- Cloud first company
- Third party integration





## NEXT GENERATION USER EXPERIENCE



- In early 2015, we began a redesign of our platform, moving logic out of the box and in to the cloud
- In November 2016, YouView Launched our award winning re-engineered and redesigned platform
- As part of the redesign, we moved out of a physical datacentre and moved into AWS
- We also changed how we worked, moving in to small, cross-functional, agile teams

# MY ROLE

- Lead Cloud Platform engineer with >10 years experience
- Joined YouView July 2015
- Designed and built cloud services for Next Gen platform
- Early 2017 became Tech Lead of Cloud Platform team



#### youview

#### Homeland

Watch live R Record • Watch live in HD

#### Series 5: EP9: Crossfire

After his attempt to sever ties with Abu Nazir, Brody finds himself reliving his ... I Solve the second scenes that may offend.

#### youview **OUR PLATFORM**

- Millions of Active Connected Devices •
- 6B requests per day to edge •
- 300M requests per day to origin ٠
- 1.1M max TPM at origin ٠
- 200k average TPM at origin ٠
- 600+ servers ٠

4

FREE

000

- 100+ micro services •
- 40 cloud service developers ٠
- 50+ AWS services •
- 6 programming languages



### WHAT ARE OPERATIONAL CONCERNS?





## SO WHERE DO YOU START?

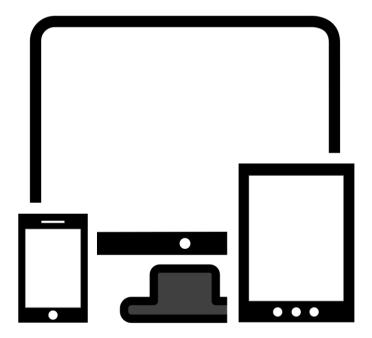




# THE CLIENT

#### Communication

- Request/Response
- Eventually Consistent
- Push Messaging
- Avoid synchronized calls
- HTTP/2
- Error handling
  - Retry logic with exponential backoff
  - Randomisation with Jitter
- Page Composition
  - Separate static assets from personailised data
- Pre-emptive painting
  - Disconnect UI feedback from the backend



RoleResponse, err := svc.AssumeRole(input)-- != nil {-

err, ok := err.(awserr.Error); ok {¬

vitch aerr.Code() {-

sts.ErrCodeMalformedPolicyDocumentException:-log.WithError(aerr).--

Errorln(sts.ErrCodeMalformedPolicyDocumentException)se sts.ErrCodePackedPolicyTooLargeException:log.WithError(aerr).-

Frorln(sts.ErrCodePackedPolicyTooLargeException) se sts.ErrCodeRegionDisabledException: -

log.WithError(aerr).-

Frorln(sts.ErrCodeRegionDisabledException)
fault:-

log.WithError(aerr).

> Errorln(sts.ErrCodeRegionDisabledException)

lse {⊣

og.WithError(err).-

Errorln("Error assuming role")-

xit(1)-

```
ad := exec.Command(args[0], args[1:]...)¬
ad.Stdout = os.Stdout¬
ad.Stderr = os.Stderr¬
ad.Env = append(os.Environ(),¬
Sprintf("AWS_ACCESS_KEY_ID=%s", *assumeRoleResponse.Credentials.AccessKeyId),
Sprintf("AWS_SECRET_ACCESS_KEY=%s", *assumeRoleResponse.Credentials.SecretAcc
Sprintf("AWS_SESSION_TOKEN=%s", *assumeRoleResponse.Credentials.SessionToken))
```

```
- := command.Run(); err != nil {¬
```

```
ithField("command", command.Args).--
ithError(err).--
italln("Failed to run command")--
```

## youview

# MICROSERVICE ARCHITECTURE

- Define service boundaries it's better to go too big than too small
- Strong contract testing is a must
- Lines of Code DO NOT MATTER
- Loosely couple with pub/sub and replicated data



# SIGNS YOU GOT YOUR BOUNDARIES WRONG

- Your service has a runtime dependency on another service
- Your service has a deployment-time dependency on another service
- You cannot contract test your service in isolation
- If your service goes down, another service will fail
- Adding or fixing existing functionality requires modifying multiple repos
- If your service reads another services data store



# INTER-SERVICE COMMUNICATION

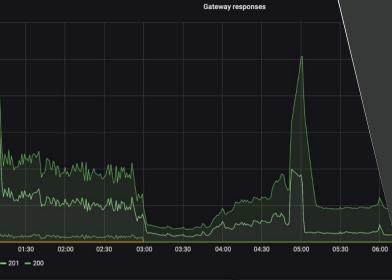
- Don't daisy-chain or aggregate requests whenever it can be avoided
  - Every hop in the network is a chance to fail
  - If any request in a daisy-chain or aggregation fails, it's a failed request
- Adopt an eventually consistent model
- Data stream technologies, like Kafka, or Pub/Sub technologies, like AMQP, allow for easy decoupling of services
- If your service requires data from another service to work, it should have it's own local copy
- If you simply can't duplicate your data, or avoid a runtime dependency, then return what data you can, and have your client retrieve the rest of the data



#### MICROSERVICES - NOT A SILVER BULLET The Problems The Solutions

- Duplicate code ٠
  - Boilerplate code
  - Infrastructure as Code
  - Deployment pipelines
- Inconsistencies in implementation ٠
- Large maintenance overhead ٠
- Shared bugs ٠

- Shared libraries •
- Generate new services from templates
- Remote templates for infrastructure as ٠ code
- Centralised Ownership ٠
- Automated CI/CD •



status	② Last 24 hours	
hostname.keyword	Count <del>+</del>	40
yv1-api.youview.tv	12.74 K	30
yv1-api.youview.tv	1.87 K	20
yv1-api.youview.tv	1.61 K	10
api.youview.tv	586.00	0
trials-configuration.vouview.tv	396.00	
tus	O Last 24 hours	
	Count -	20
	2.67 K	15
	2.26 K	10
	1.19 K	5
	377.00	
	289.00	



## youview

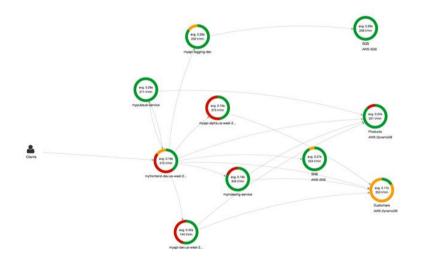
# **OBSERVABILITY**

- Monitor everything
- Expect to spend 10-20% of your operating budget on monitoring
- Differentiate between what should be real time, and

#### what can be delayed



## DISTRIBUTED TRACING



- Improves observability
- Easily find bottlenecks and broken services
- A service map can be a good indicator of

accidental coupling



## COST



DATABASES

DEVELOPER EXPERIMENTS LARGE TRAFFIC SPIKES

OVER-PROVISIONED SERVERS

## THANK YOU