

Drupal GovCon 2020

Dealing with unprecedented scale during Covid-19

Sean Hamlin :: September 2020



amazee.io



About me



Sean Hamlin

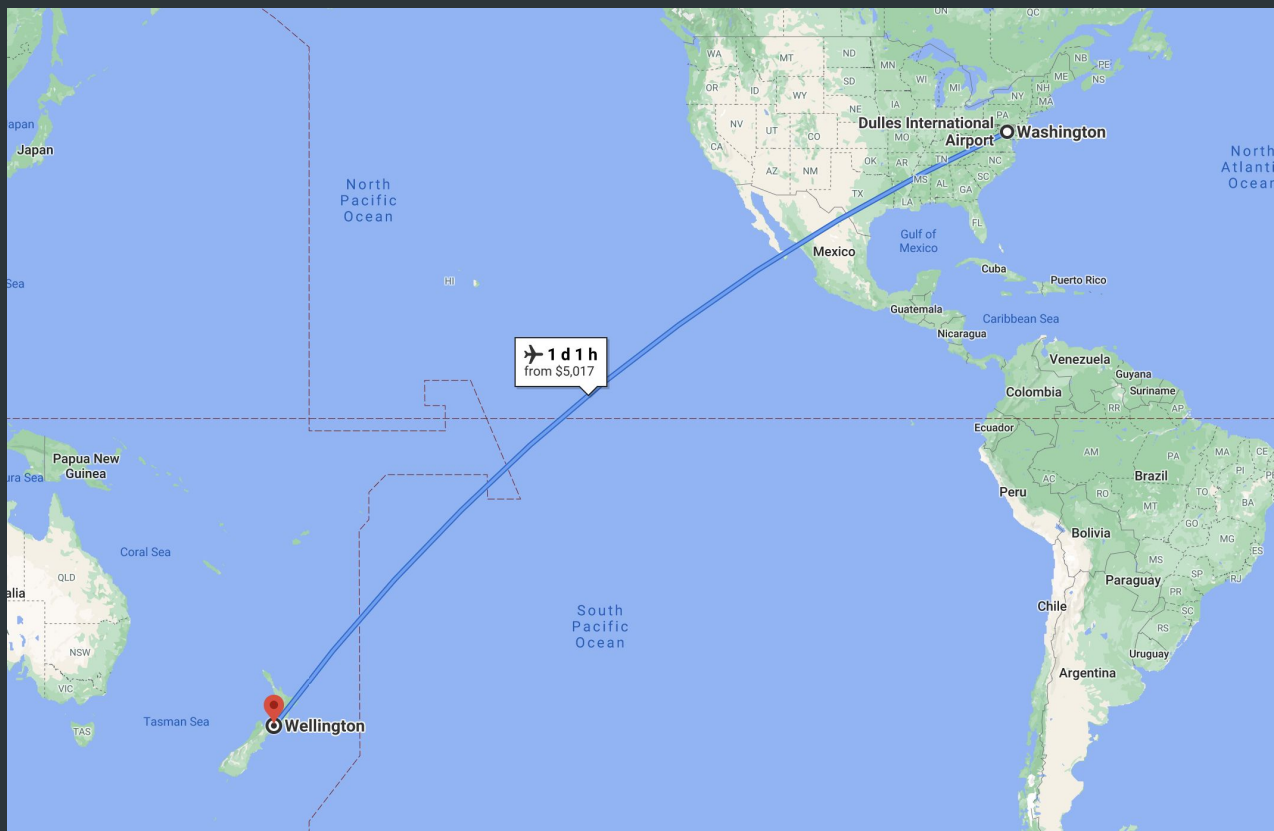
Technical Account Manager

Passionate about making awesome solutions to solve complex business problems. Based primarily in Wellington, and working with some of amazee.io's most strategic customers in region and globally.

Email sean.hamlin@amazee.io

Twitter + Drupal.org [@wiifm](#)

LinkedIn <https://www.linkedin.com/in/shamlin/>





Thanks!



Agenda

- Overviews in 1 minute
- Covid-19 and what this meant
- Covid-19 by the numbers
- Horizontal scaling
- Database sizing, scaling and tuning
- In memory caching
- Minimising origin hits
- CDN tuning
- Why should you care

Amazee.io in 1 minute

amazee.io is a
fully-remote team of
open source hosting
experts, globally
distributed to work **24/7**
keeping our customers'
websites and applications
running smoothly at all
times!



amazee.io

About amazee.io

Fully Managed Hosting based on Lagoon

- Host anything, anywhere in the world
- 24x7 chat-based support from Senior Systems Engineers
- Global team in all major time zones
- Everything open source; full stack visibility and configurability
- Committed to transparency
- Global enterprise clients including governments, healthcare, financial services, and education
- Trusted security, support, and partnership model



GovCMS in 1 minute



Australian Government

GovCMS in 1 minute

- Managed by the Dept of Finance Australia
- Drupal hosting for **Federal, State** and **Local Government** in Australia
- **Salsa Digital** and **amazee.io** as technology partners
- Running in production since January 2015
- **329** sites live as of today, all shapes and sizes
- **SaaS** (BYO theme only) and **PaaS** (BYO codebase) models
- Security accreditation done at a platform level and SaaS has application level as well
- See [Platforms, Platforms, Multisites, and Platforms](#) by John and Adam earlier today for a better overview on platforms and Drupal

Disclaimer



Disclaimer

Jargon buster

- I will mention Amazon Web Services (AWS)
- I will mention Kubernetes and containers

No pre-existing knowledge is required, I will explain core concepts using weird metaphors.

Covid-19 and what this meant

A man with short brown hair and glasses, wearing a red checkered shirt, is seated at a desk in an office. He is looking at a laptop screen which displays a website with the heading 'PACKAGES INSIDE INCREMENT'. In the background, other office workers are visible at their desks, and the office has a modern, bright aesthetic with large windows.

Covid-19 and what this meant for GovCMS

Several GovCMS sites impacted by Covid-19:

- **Health related sites** (e.g. <https://www.health.gov.au/> - this is the authoritative source of all Covid-19 information in Australia)
- **Travel related sites** (e.g. <https://www.smartraveller.gov.au/>)
- **Government payments sites** (e.g. Job seeker payment on <https://www.servicesaustralia.gov.au/>)

Covid-19 by the numbers

Covid-19 by the numbers

187,000
concurrent visitors

in Google Analytics for GovCMS SaaS

Covid-19 by the numbers

100,000
pageviews

on the peak minute for GovCMS SaaS following a nationwide push
notification

Covid-19 by the numbers

8.4 TB
bandwidth

on the peak day

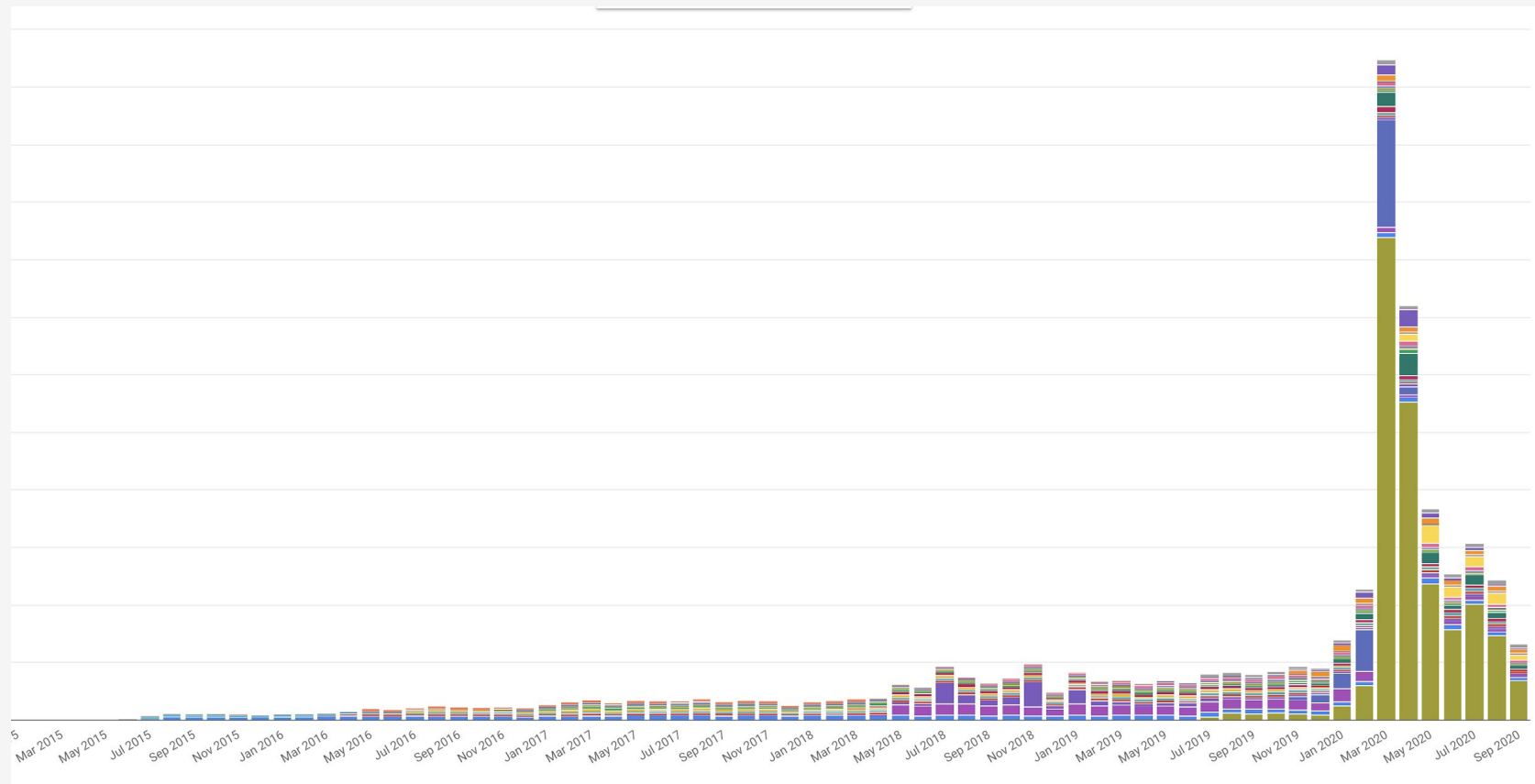
Covid-19 by the numbers

80x traffic

March vs February

pageviews increase in Google Analytics for one site

Covid-19 by the numbers (Google Analytics pageviews by site by month)



Horizontal scaling

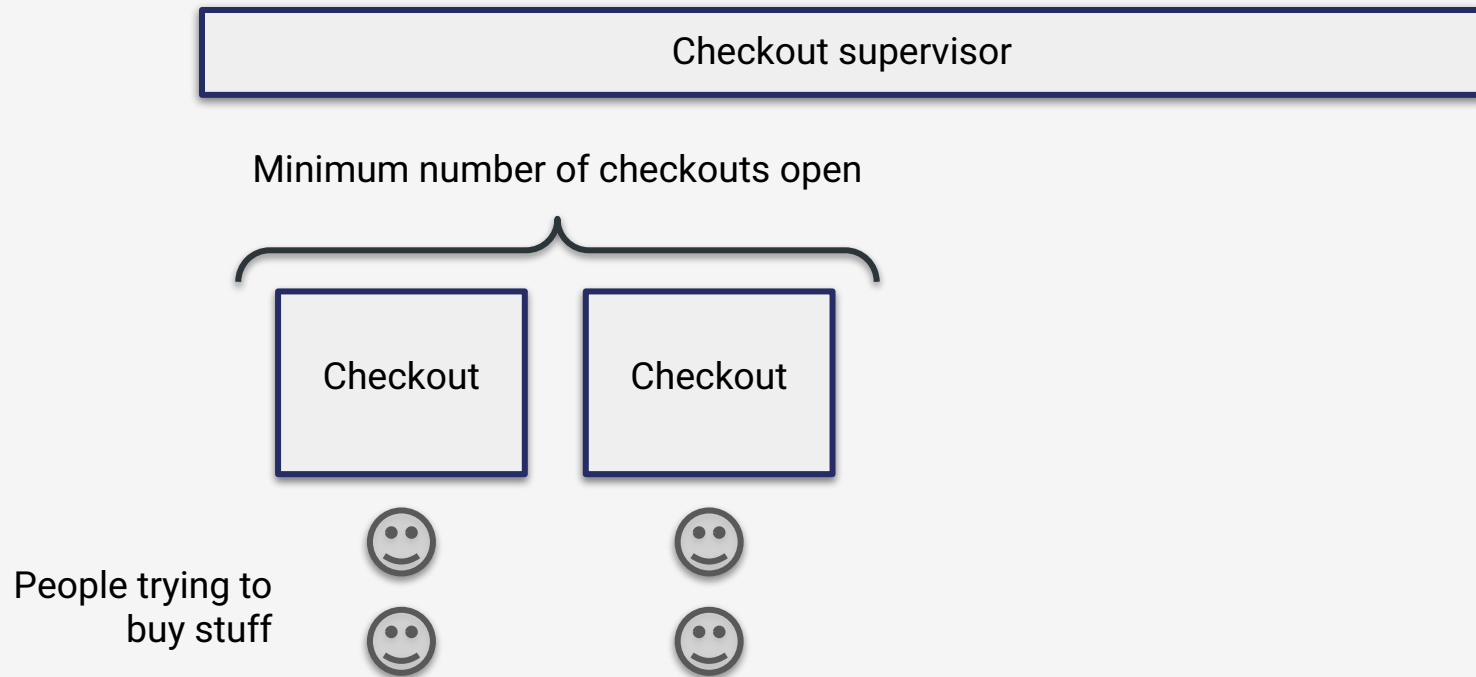


Horizontal scaling

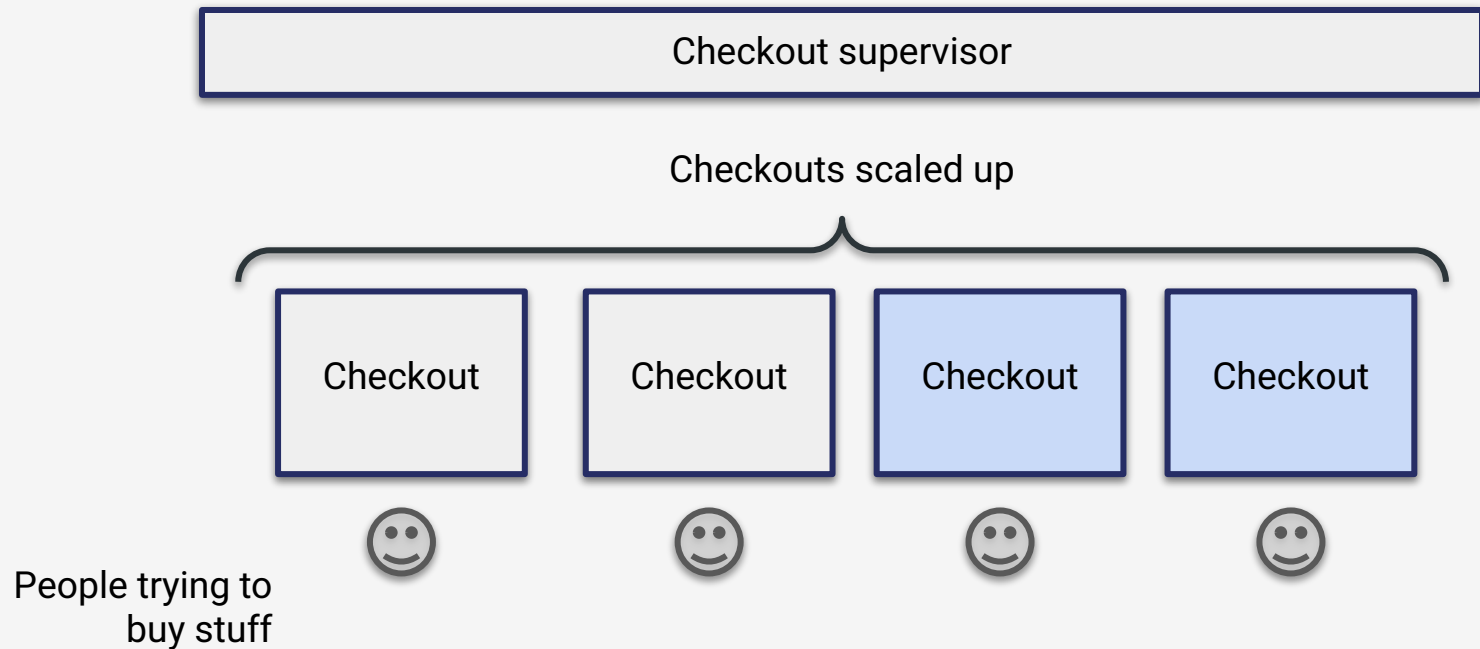
Being able to increase application resources in response to traffic.
Ideally this happens in an automated fashion

- In Kubernetes, this is done by an **Horizontal Pod Autoscaler** (HPA) and a **Cluster Autoscaler** (CA)

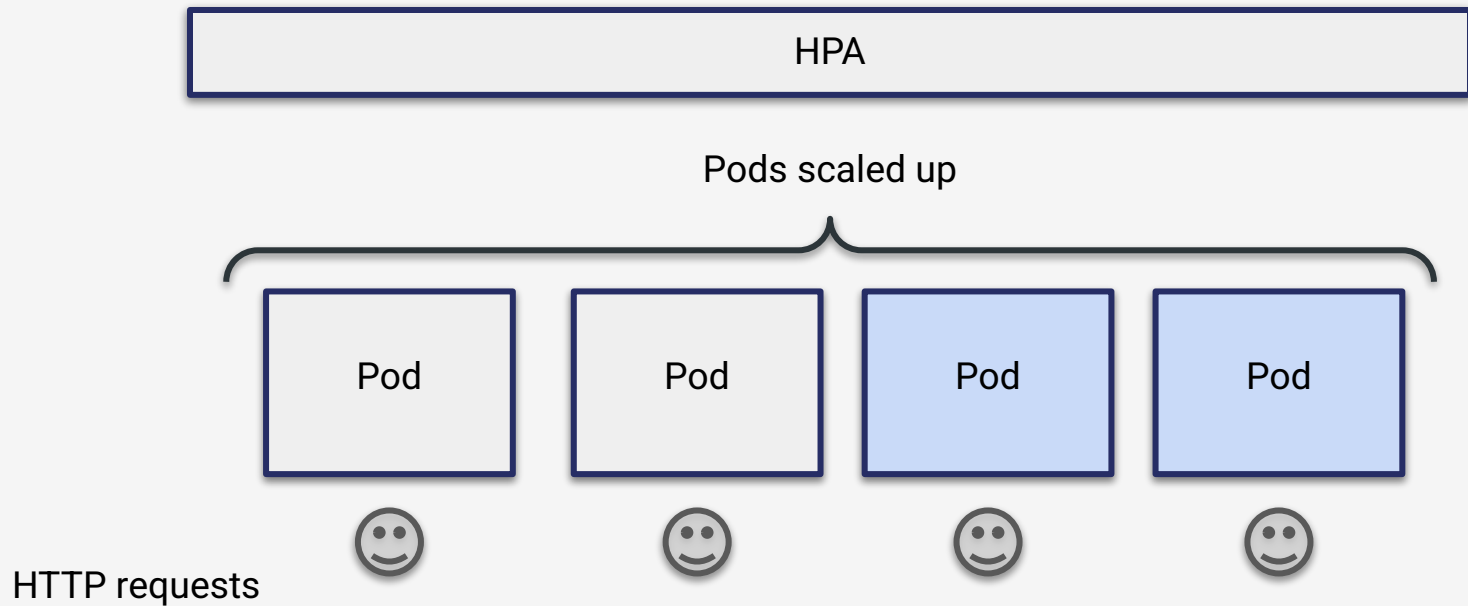
Horizontal Pod Autoscaler (HPA) explained with supermarket shopping



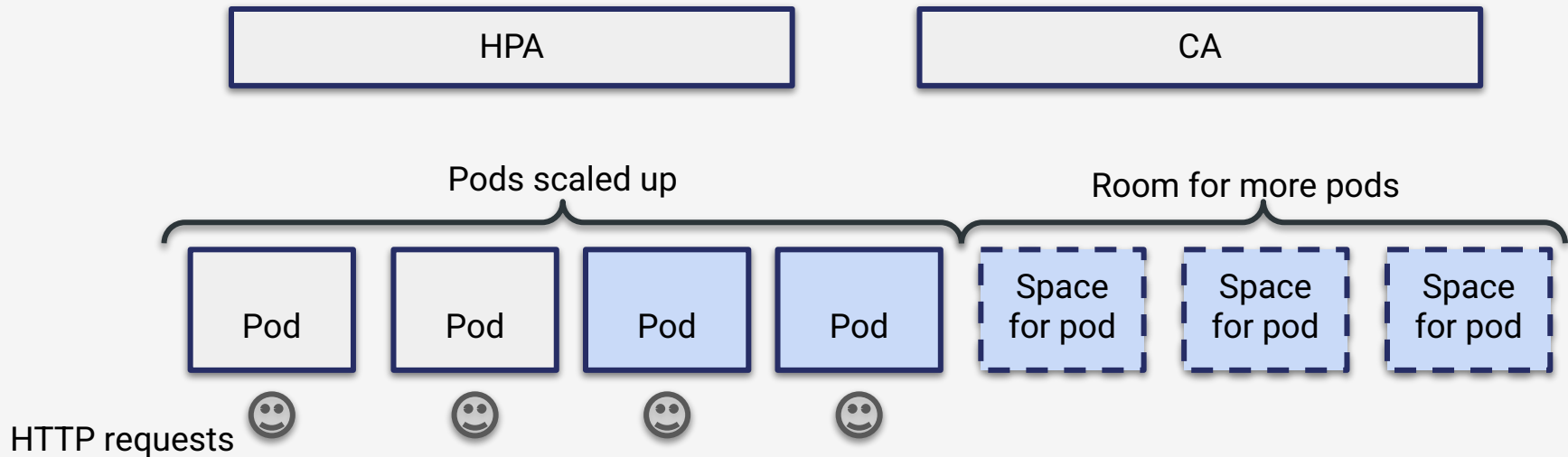
Horizontal Pod Autoscaler (HPA) explained with supermarket shopping



Horizontal Pod Autoscaler (HPA) explained with supermarket shopping



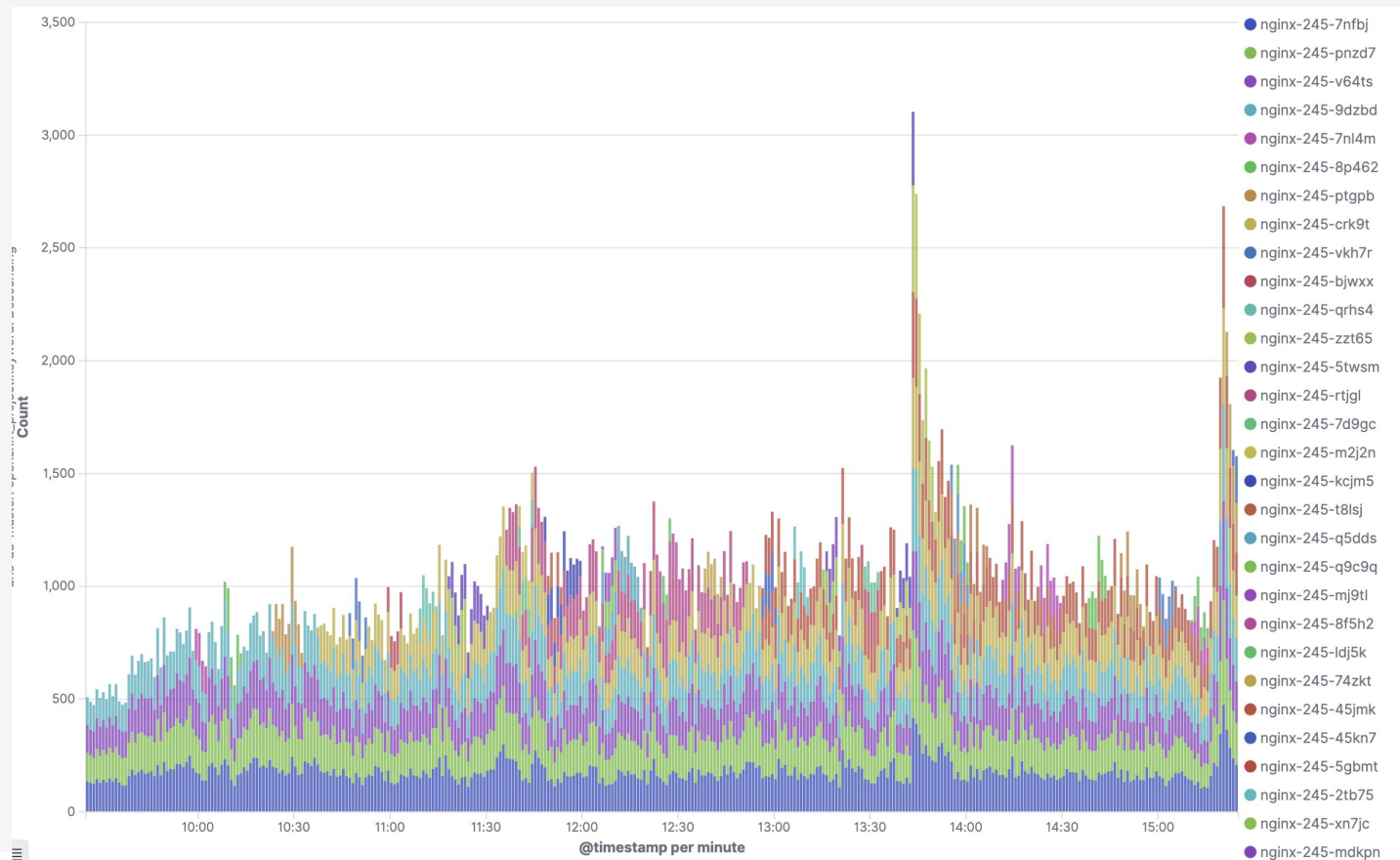
Cluster Autoscaler (CA) explained with supermarket shopping



Horizontal scaling

- Development of a new “Covid HPA”
 - **4 pod minimum**
 - **12 pod maximum**
 - Much faster scaling
- **Daily reviews** of the HPA for Covid impacted sites
- Reduce “pod flapping”, and ensure there is headroom

HPA in real life



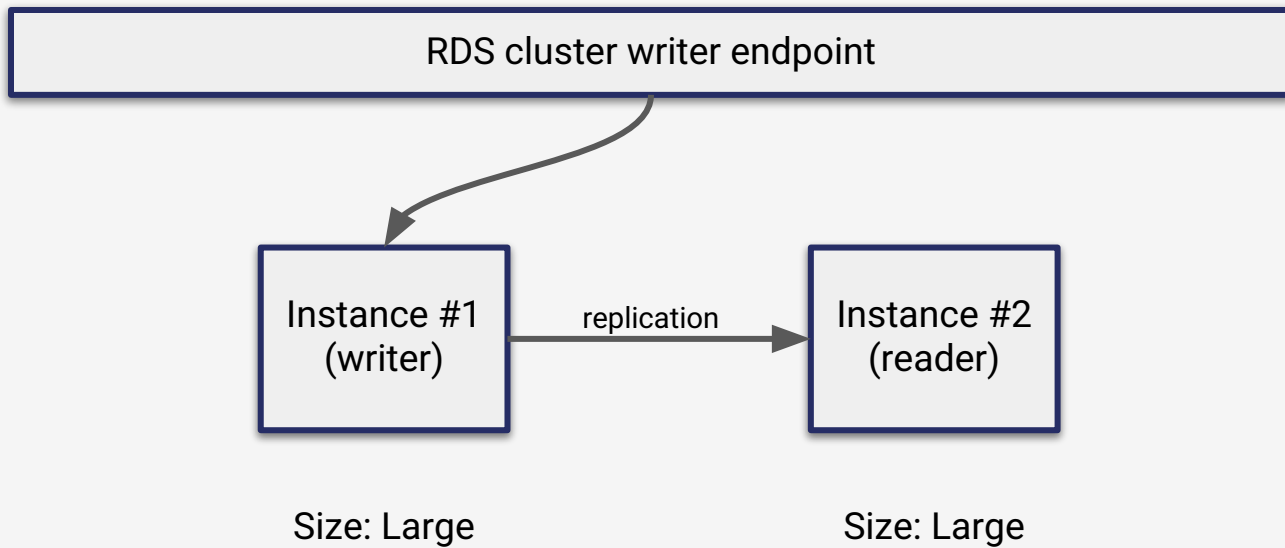
Database sizing, scaling and tuning

Database sizing, scaling and tuning

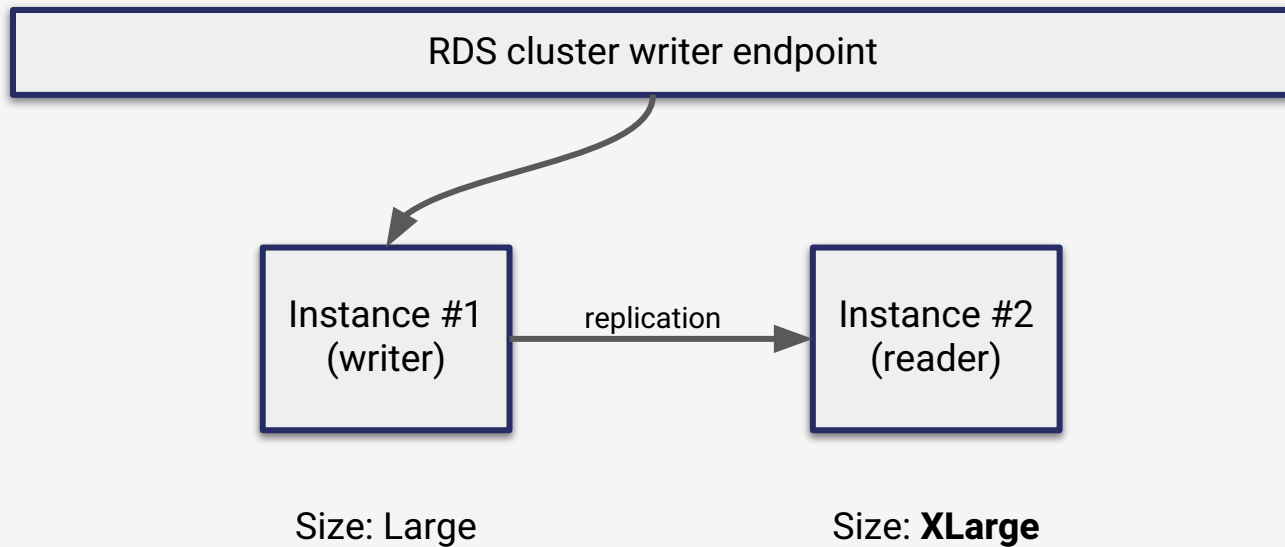
- So you scaled your application layer, for those that do not know, **Drupal loves Database queries**
- You can run a database in a pod, but this is often a poor life choice
- Sunil did another session today [Drupal Hosted on Kubernetes](#) which you should check out
- AWS' Relational Database Service (RDS) Aurora was used
 - Many features over regular MySQL, one of those features being **faster failovers with no data loss**



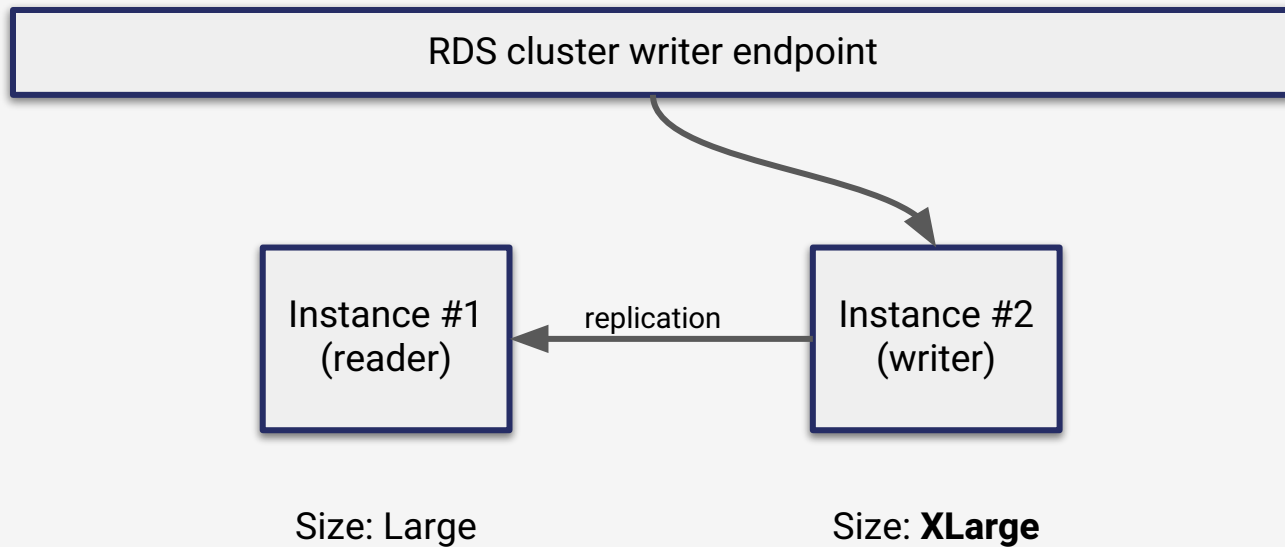
RDS upsize strategy



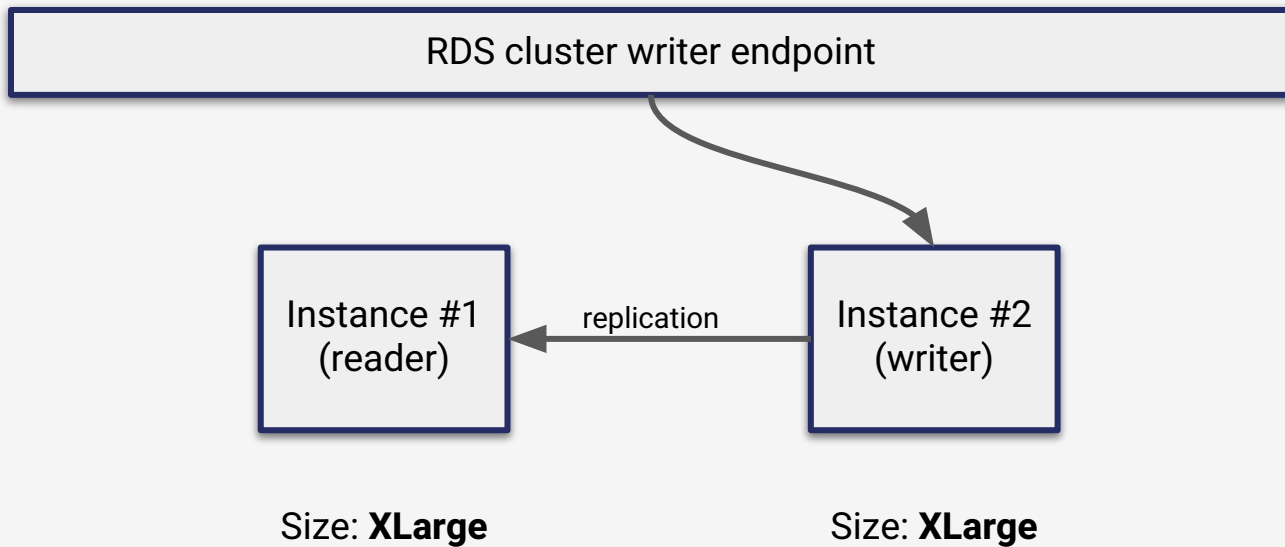
RDS upsize strategy - upsize reader first



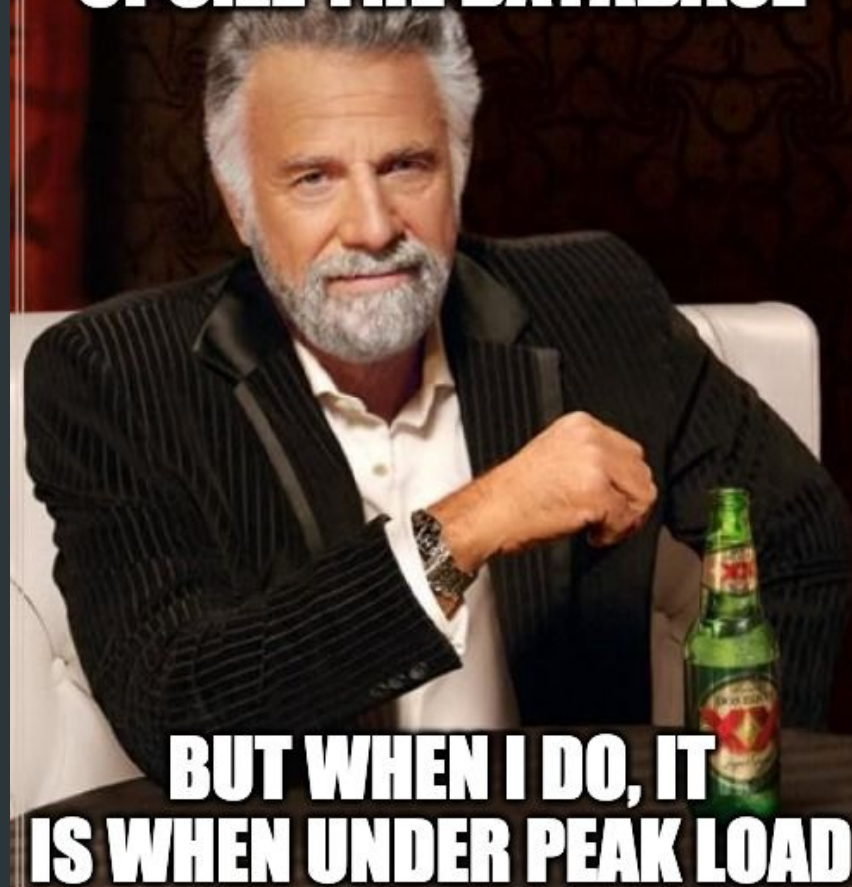
RDS upsize strategy - failover writer



RDS upsize strategy - upsize new reader



**I DON'T OFTEN
UPSIZE THE DATABASE**



Database sizing, scaling and tuning

- Able to fully upsize the RDS cluster in around 15 minutes end to end
- Minimal failed queried, typically < 50 in total in a 30 second period
- When ended up upsizing Services Australia **twice** (so 4 times as large) in 2 days

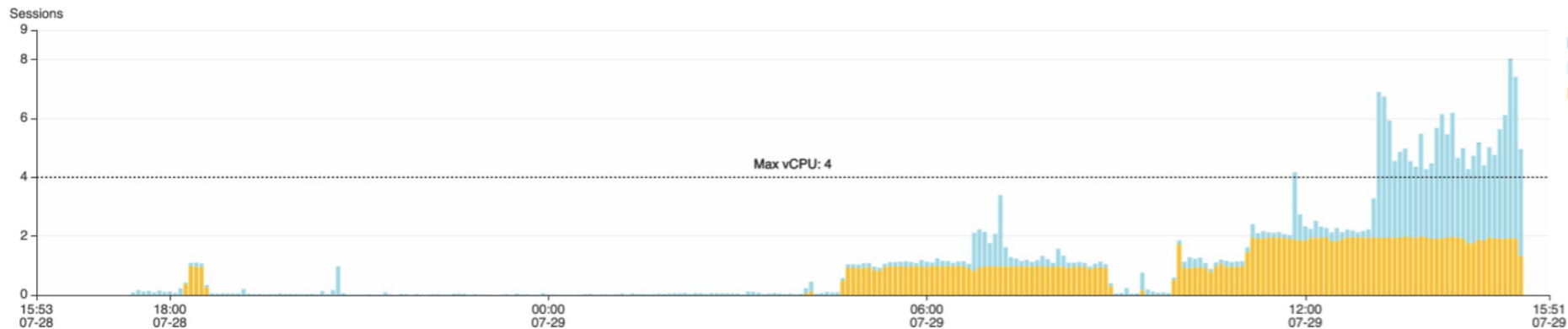


AWS performance insights (sad)

Database Load

Current Activity Measured in Average Active Sessions (AAS)

☒ Show max vCPU

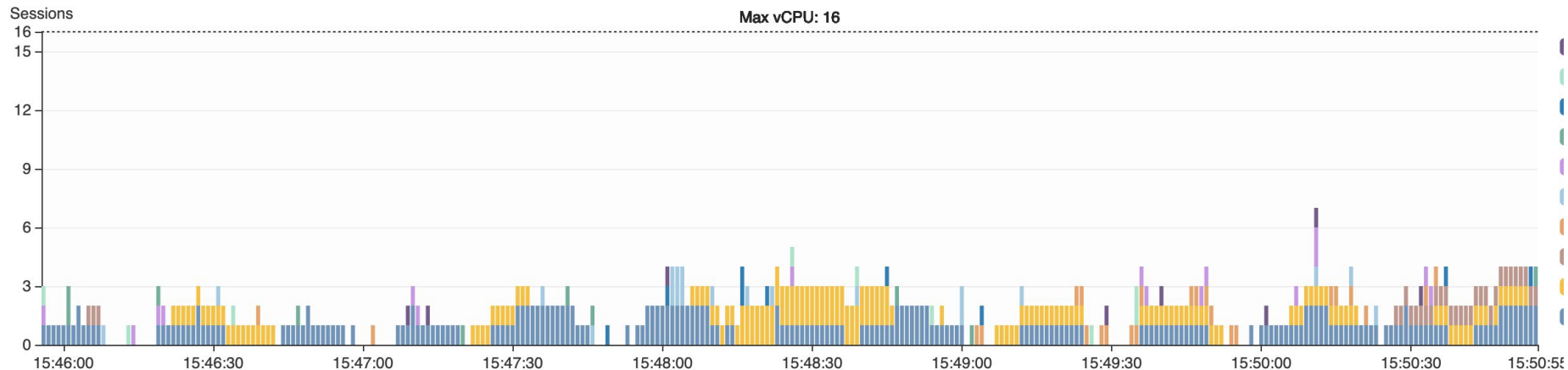


AWS performance insights (happy)

Database load

Current activity measured in average active sessions (AAS)

☒ Show max vCPU



AWS performance insights (pro tip)

View past

5m

1h

5h

24h

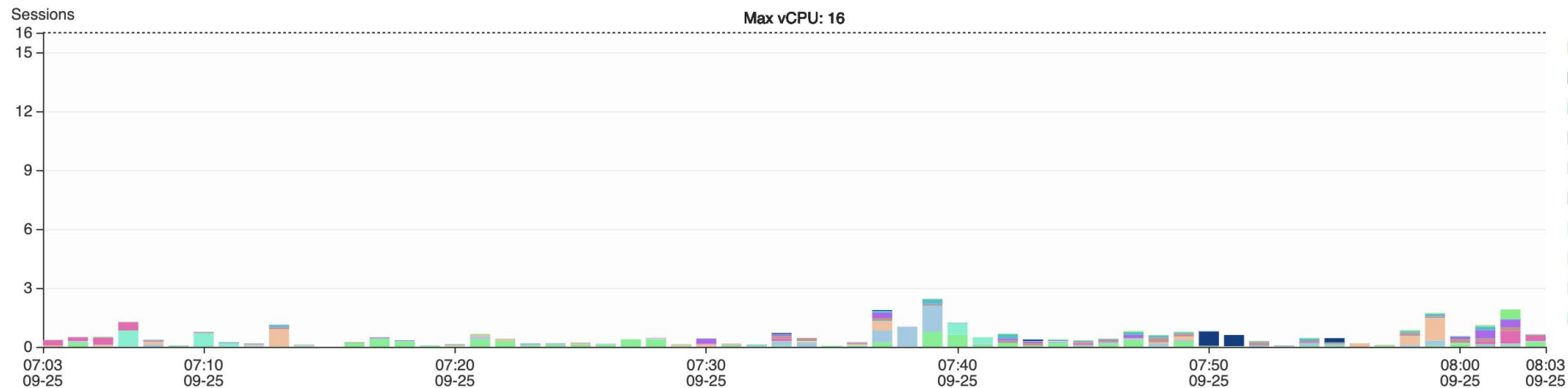
1w

all



Auto refresh

Current activity measured in average active sessions (AAS)

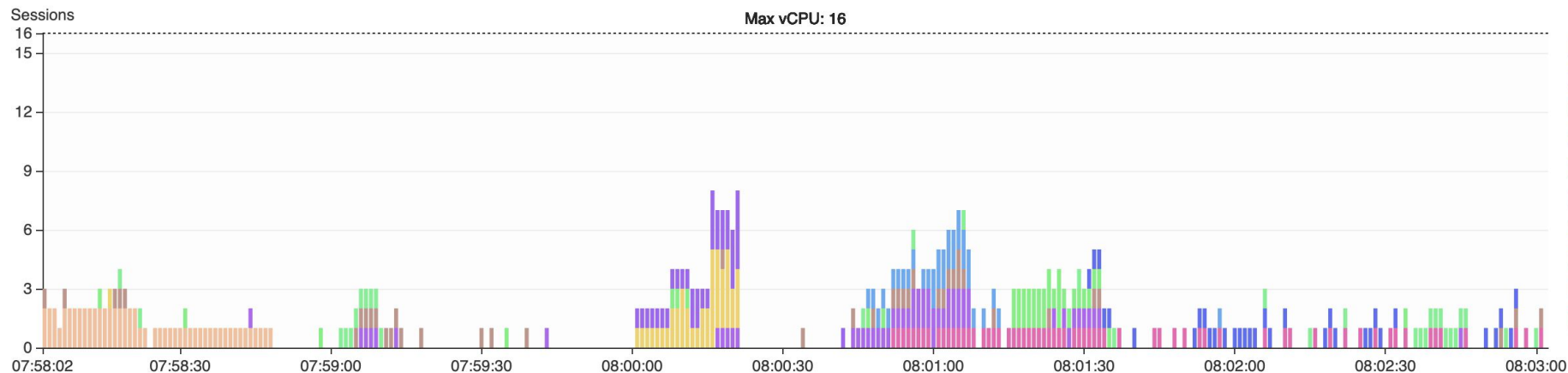


AWS performance insights (5 minute)

Database load

Current activity measured in average active sessions (AAS)

☒ Show max vCPU



In memory caching

In memory caching

- Drupal also **loves to cache things**, and the database does not make a great cache system at scale



Introducing Redis





In memory caching

- We have 2 options for Redis
 - Redis pod (not Highly Available, Single Point of Failure)
 - AWS ElastiCache (Highly Available, replicated across many Availability Zones)
- When is it time to not use a pod
 - You require High Availability
 - 60 MB/sec traffic
- Ensure you configure a redis timeout

Minimising origin hits

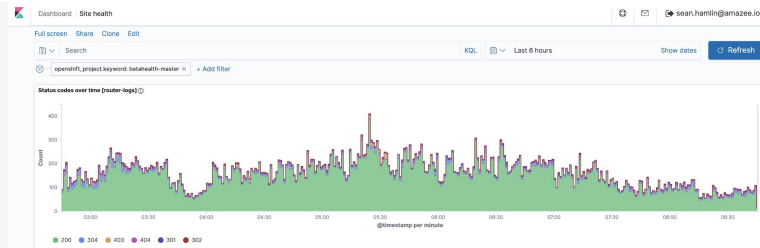


Minimising origin hits

- Even better when it comes to scaling is to never have to generate the pageview on origin and rely on the CDN instead
- Elasticsearch and Kibana
- Site health dashboard
- Example dashboard
- Sum origin responses by PHP time
- Example screenshot

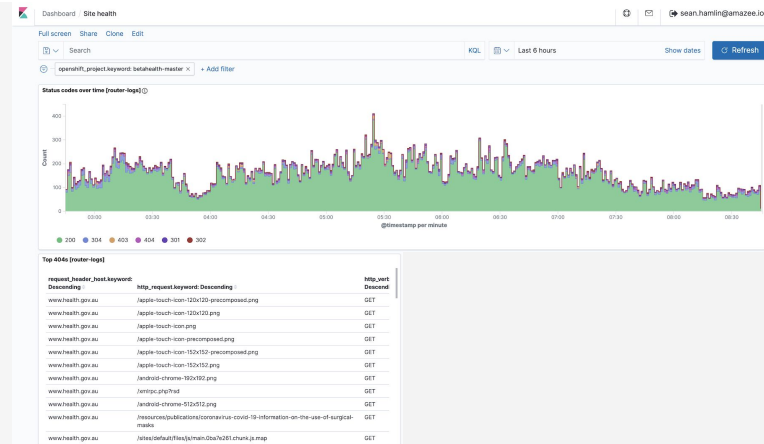
Site Health dashboard

- HTTP status codes



Site Health dashboard

- HTTP status codes
- Top 404s by path



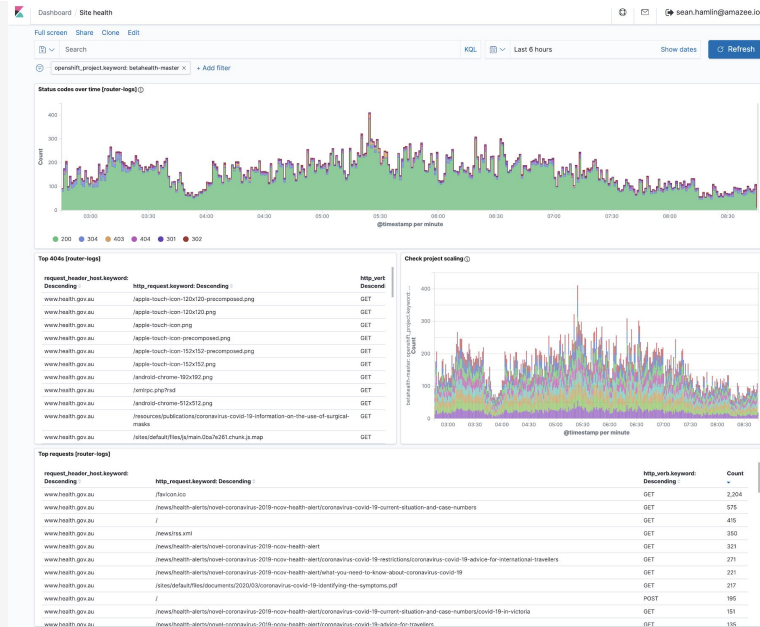
Site Health dashboard

- HTTP status codes
- Top 404s by path
- Pod scaling



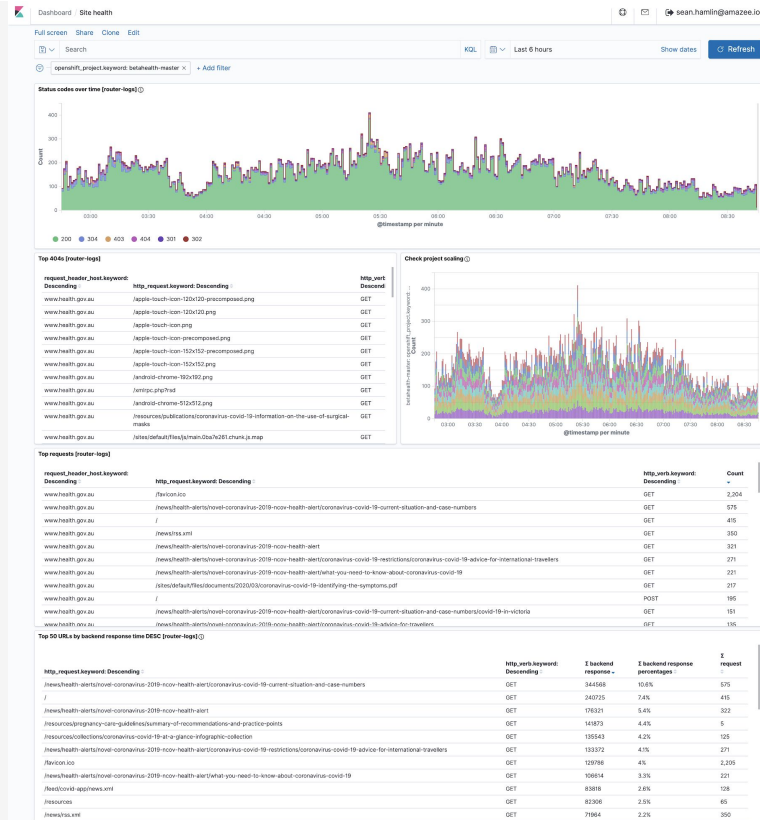
Site Health dashboard

- HTTP status codes
- Top 404s by path
- Pod scaling
- Top requests by path and method



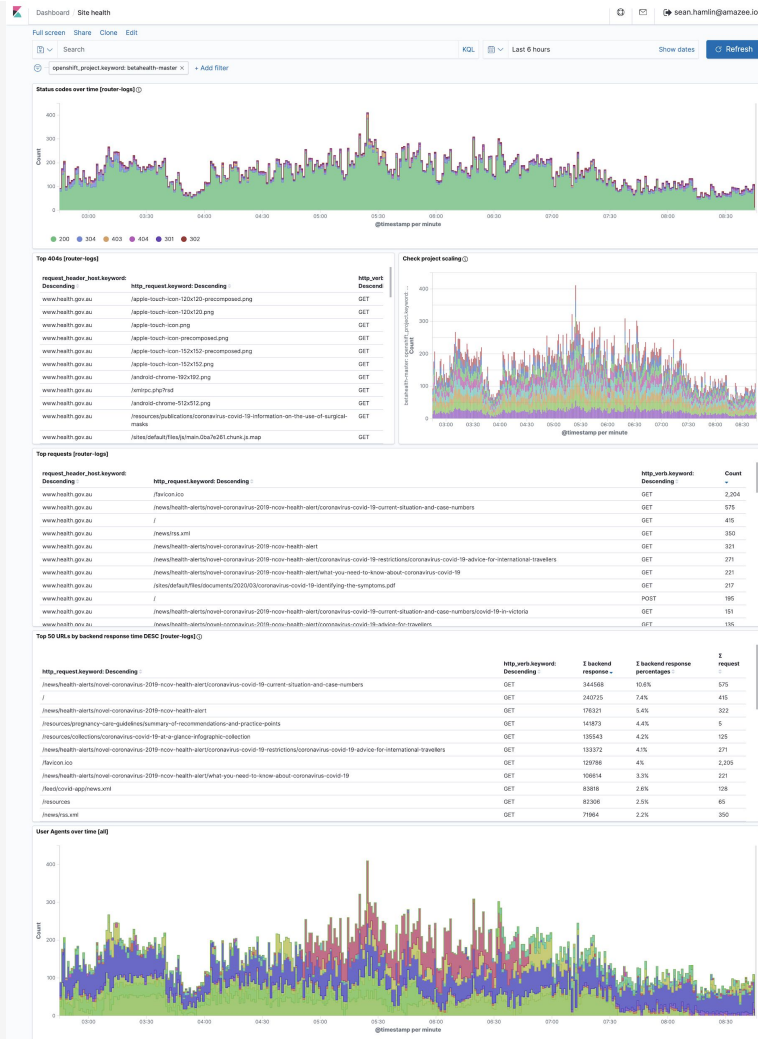
Site Health dashboard

- HTTP status codes
- Top 404s by path
- Pod scaling
- Top requests by path and method
- Top requests by path and method, sort by PHP time aggregate



Site Health dashboard

- HTTP status codes
- Top 404s by path
- Pod scaling
- Top requests by path and method
- Top requests by path and method, sort by PHP time aggregate
- Top user agents



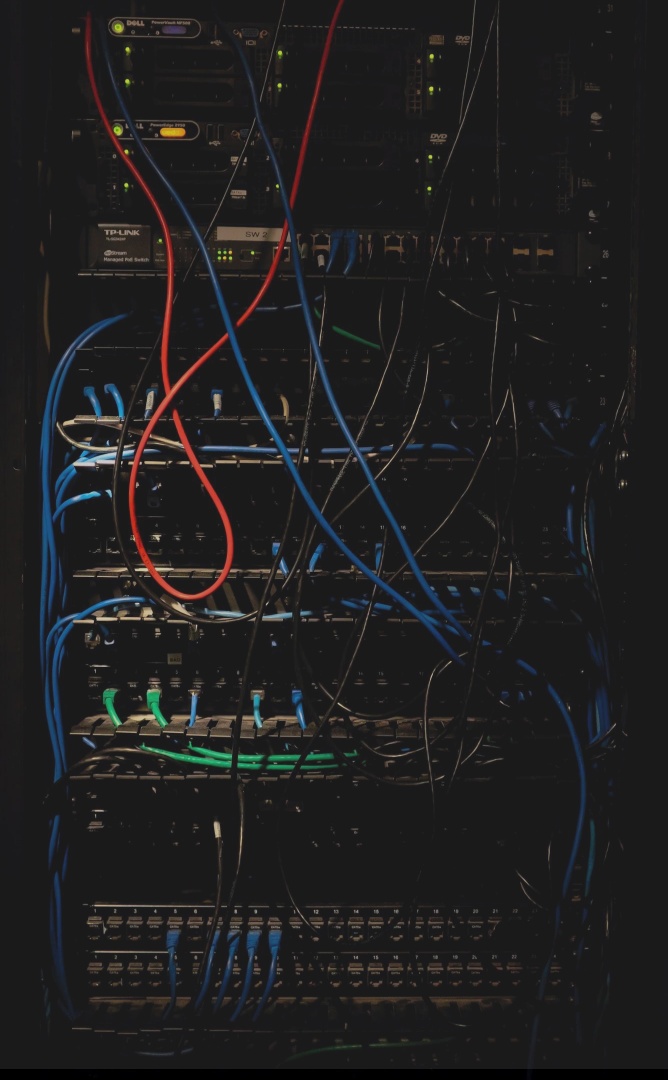
CDN tuning



CDN tuning

- Critical to the success of the program. You need a major player.
- DDoS protection
- WAF and bot management
- Serve while stale (e.g. 500s)
- Negative caching (e.g. 404s)
- Background refresh
- Drupal 8 and Cache tag invalidation
- Next best alternative (static fallback)

Building a culture of continuous improvement (Kaizen)



Kaizen

Build

Learn

Measure



1/0



Building a culture of continuous improvement

- How high can you get the Cache Hit Rate ?
- What requests are using your backend resources you can avoid?
- When was the last time you reviewed your slow query log?
- What are the top 404s on the site, and can they be avoided?
- What version of PHP are you running, can you move to PHP 7.4?
- Are you using cache tag purging?
- When was the last time you reviewed which cache tags are in the block list?
- You should encourage your team to find ways to improve the platform
- You are never done, just on the journey

Building a culture of continuous improvement

$$1.01^{365} = 37.8$$

just 1% better every day adds up over time

Why should you care

Quote from Services Australia

“ *It meant we were able to supply information to the public during a pandemic with no outages. This supported real people and their ability to care for themselves and their families.”*

Susie Smith, General Manager, Communications Division, Services Australia.



Thank you



amazee.io