## RavenPack

In The Time Of COVID-19

## Introduction

- Many of you will know as, having seen us here, presenting or just attending most years
- We're a data and analytics company based in Marbella, Spain
- We provide news analytics to large financial institutions; banks, hedge funds, insurance companies etc
- The majority of our data processing is done in Common Lisp


## Remote Work

- Almost all of our development happens in our Marbella office
- With the exception of a couple of remote workers
- We've always preferred to have staff in our Marbella office as it's simply easier to communicate with people when you can talk to them face-to-face.
- Lispers tend to work on multiple projects, meaning multiple people to keep up with an, unfortunately, multiple meetings
- Many meetings can be avoided simply by taking a short walk, which by pure office geography, will usually always be in the direction of the coffee machine
- And our biggest internal app's users sit 10 m from the Lisp team, which makes support very convenient


## Remote Infrastructure

- We have a VPN, and sshuttle for those that prefer such things
- Internal web apps have external endpoints, JIRA, GitLab etc
- Much of our infrastructure is in AWS, so no issue where the developer is
- So it seems that we would be prepared for a sudden everybody-remote event
- But of course the challenges were greater than that, and many specific to our team


## Data Volume (1)

- Several of our apps require large amounts of data in memory
- 2 gigs is common, one app can hit $10+$ while running
- Startup in the office where our dev DBs are can take 20 to 30 minutes. Time that with a meeting or a chat elsewhere and it's reasonable
- Most of that is simply database queries and data transfer, locally
- And if you're trying to run stuff at home, via VPN, things slow down significantly. Over an hour in some cases
- Years ago I made an experimental addition to our 'with-db-query macro, part of our Oracle library. I never took it to full functionality. It was simply the ability to cache the results of any query on local disk
- Our normally-remote colleagues had recently improved and filled in the missing functionality. You can even cache with readable Lisp, FASLs or Rucksack.
- A remote worker with cached data can now restart faster than someone in the office using the DBs


## Data Volume (2)

- Most of our input data are news stories or other similar content from various providers
- Usually stored in S3, so nowhere near the office or most developers
- People have started to cache those too
- Bulk test runs etc are also much faster
- Of course when you start caching data you have to start worrying about cache size and age
- This data is changing daily. So care is required
- There's a pattern here. Make yourself independent of distant resources where possible, and with care


## Pre-building Images (1)

- We have no functional need to pre-build images. Our applications are not distributed to customers
- But despite caching etc, home workers still want fast restarts when things go wrong
- Having access to an image containing the latest code and data built that same morning and ready to start can make a huge difference
- We're now building images daily, sometimes more often. Some are for internal application use. Others only for development
- And with builds comes the opportunity to improve our testing infrastructure as things are now faster and repeatable
- Its important to note that these images are for dev use, and so must have access to all resources and have removed nothing. No tree-shaking here. They are also very large.


## Prebuilding Images (2)

- Our images are saved for dev and we also use them for our internal production applications
- We have to make sure that all of the configuration option's we'd normally have can be changed on a pre-built image and still have effect. So no config trapped in closures. Everything is set up to refresh config at run time
- So they may be built and load data from one DB, but continue from another
- We have a data loading abstraction that stops accidental reload of the same resources
- This means that things like DB resource pools are shut down but not destroyed and that they can be reactivated, often with different configuration.
- Everything must be zeroed out
- Source file location is standardised and configurable so that systems can be loaded as if we were on the original build machine
- Our dev DBs and others are available from practically anywhere within our network, so dev work can be done from the office, from AWS of from a developer's home


## Image Management

- We used to run everything from source under Emacs
- So it made sense that auto-restarting dev apps would be done by an Emacs script. Nowadays more like an overgrown library of app-specific restart instructions.
- The next step was to take our now prebuilt images and, embedding them in Docker containers, have a container management app do our deployment
- On paper this is a lot more flexible and easily programmed, but we're now slaves to yet more technology that we don't control. Ups and downs you might say
- Of course our Lisp-containing Docker containers are also set up so that the image inside can access source via appropriate mounts
- And you can connect to these lisps via Slime or Sly


## Non-Lisp concerns

- Zoom has been the most reliable video chat app for meetings
- People are starting to get creative with the green screen effects
- We tried others but there was always a problem, be-it poor quality of poor interface.
- Code reviews are done via screen, tmux, or where there are multiple reviewers, TMate.
- For myself at least, video calls have now become the norm. I'm not feeling the slight discomfort that I used to now that it's a way of life for everyone


## What Are We Up To?

- Actual work has continued as normal
- But given the circumstances and the data at our disposal we've ended up creating a little free COVID-19 news tracking app
- https://coronavirus.ravenpack.com



## $\frac{\text { Media Hype Index }}{54.74}$




Contagion Inder
Conlagion 11
62.98


Media Coverage Index
77.69



Mecila Exposure to Coronavirus

| County | Pollicians | Treatments | Arines | Food \& Beverages > |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| nuw | semm | \% | 1 tomem | $\downarrow$ wemaxe | Dame |
| 1 United States |  | 3.6 | +4.3\%\% | 74.76 | +0.09\% |
| uk |  | 67 +10 | 14.25\% | 18.5 |  |

${ }_{3}$ Canasaa
${ }_{6}{ }_{6}$ China
${ }_{8}^{78 \text { Gemanany }}$

| $\circ$ Japan |
| :---: |
| 10 australa |

11 spain
13 switerland
14 Netheranans
15 Russla
${ }_{16}^{17}$ South Aficica:
${ }_{17}^{17 \text { Beglyum }}$
${ }_{18}^{18 \text { ceorga }}$
${ }_{20}^{20 \text { mexico }}$
m

## Yes, Not Lisp

- Uses JS and React
- Data is pre-prepared by the backend
- Is mostly RavenPack data with the addition of some official stats
- For a change I can show my mum what I do


## And So?

- Working remotely isn't so bad
- But l'd prefer to be in the office
- Our infrastructure has been reliable
- And recent work has made things acceptably fast
- So things are more or less normal
- Ask my dogs if you want a second opinion about being stuck in the house


## Fin

## We are not hiring

Until things are back to normal
Hopefully sooner than later

