



Planning for WordPress at Scale

News

MERCURY
THE VOICE OF TASMANIA

Cairns Post

Escape

INSIDE

Sunday Mail

style

news
.com.au

Sunday

Bulletin

Best Recipes

WEEKEND AUSTRALIAN

perth **now.**com.au

The Advertiser

delicious.

Advertiser

eureka
report

the deal

VOGUE

kidspot
com.au

Business Spectator

Daily Telegraph

NT News

wish

Territorian

WEEKLY TIMES

taste

News Corp Australia

GQ

QUEST
COMMUNITY NEWSPAPERS

FOX
SPORT

COUNTRYSTYLE

THE AUSTRALIAN

Leader
COMMUNITY NEWSPAPERS

super
food ideas

theTelegraph.com.au

couriermail.com.au

super
COACH

birth
com.au

Messenger
community news

The Sunday Telegraph

Courier Mail

Herald Sun

mx

body+soul
bodyandsoul.com.au

MISS VOGUE

donna hay
magazines

Gold Coast
Bulletin

NEWS LOCAL

Sunday Times

SUPERRACING
com.au

SUNDAY
Tasmanian
YOUR DAY YOUR PAPER

Our Team

Thousands of journalists

Editorial Heads of Digital in each State

Product Managers and Platform Managers

Editorial Technology People (support digital story telling)

Journalist/coders (interactive story tellers)

Hundreds of CMS users across Australia

Multiple PHP dev Teams (Sydney)

Multiple Java dev Teams (Sydney)

Multiple Front End Teams (Sydney + state based)

Design/UX team

You? 😊

Where we are

Java, and PHP

“Metro/regional/local” - News.com.au, Foxsports.com.au, heraldsun.com.au, dailytelegraph.com.au, couriermail.com.au, adalaidenow.com.au, perthnow.com.au (etc!)

- Oracle WebCenter for Sites (aka fatwire)
- Custom Caching layer (software AG terracotta)
- Highly customised akamai based templating engine (Edge Side Includes)

“Lifestyle” - taste.com.au, kidspot.com.au, bodyandsoul.com.au, homelife.com.au (etc.!)

- Bespoke PHP
- Expression Engine

Blogs – expression engine (PHP)

What we have already done

APIs and node.js

Decoupling News Room Authoring from Presentation - CAPI

- All stories, images and video + collections and taxonomy in a single REST JSON API
- Sharded, scalable, cached (elasticsearch)
- Connects to all authoring platforms, including WordPress

Personalisation APIs, Data APIs

Widgets - TCOG

- Node.js based templated recursive HTML fragment rendering of API calls (including CAPI calls)
- Redis + custom memory cache = sub 10ms response at 1000 rps per node
- scalable, cached

Why change?

<https://medium.com/whither-news/the-decootification-of-media-companies-e638bd798f48>

“Something else happening here: the end of the mass-media business model built on reach and frequency (unique users and pageviews) — in a word, volume. Google, Facebook, retargeting, programmatic advertising, all the companies and trends that are growing in advertising focus on individuals over masses, on data over mere exposure. If news companies do not figure out how to know people as individuals and find value there, reconstituting themselves as relationship rather than merely content companies, then they will find the ice floes under them melting sooner than later.

” - Jeff Jarvis, 5 August 2014 - The decootification of media companies

How we selected A New CMS

Highest Level Selection Criteria

- **Open source and cloud deployable**
- **Community based**
- **Demonstrated at scale**
 - <http://builtwith.com/techcrunch.com>
 - <http://builtwith.com/news.com.au>
- **Agency support (creative and dev)**
- **SaaS/PaaS model, preferably from core team**
- **PHP, Java or JS (no .Net)**
- **Happy producers**

Initial list

Some a bit random based on internal experience

- **Drupal (PHP)**
- **WordPress (PHP)**
- **dotCMS (java)**
- **Oracle “Fatwire” (java)**
- **Alfresco (Java)**
- **Magnolia (Java)**
- **Joomla (PHP)**
- **Expression Engine (PHP)**
- **Adobe CQ (Java)**

Shortlist

Drupal, WordPress (PHP) and dotCMS (java)

- Self hosted dotCMS on AWS (Amazon Web Services)
- Self hosted WordPress on AWS
- SaaS WordPress VIP with AWS development pipeline
- Self hosted Drupal on AWS
- PaaS Drupal (Acquia)
- **PaaS/SaaS evaluation on commercials and prototype/existing experience**
- **Self Hosted based on performance test and AWS costs**
 - Actual site build with CAPI backend
 - Webload on site to 600 session per second

Shortlist Environment

- **Got an expert agency/guru to help in each environment**
- **Scalable AWS environment, connected to NewsAPI**
 - Auto scaling group of web servers set to vendor recommendations
 - Tuned MySQL for each platform
 - Plugins and config
 - Soak test with content coming in “live” over 24 hours
 - Traffic surged to 600 requests per second, 20% logged in
 - On the internet, traffic from outside
- Multiple rounds as issues found

Shortlist results

Both PHP platforms functionally, costs wise and performance OK

- **Could handle double news.com.au traffic in AWS on 4 “large” AWS web nodes with a “large” RDS database**
- **Different cost models for PaaS/SaaS very hard to compare ...**
- **But finally – Hybrid WordPress VIP/AWS winner based on:**
 - Happy producers in trial site builds
 - Feedback from News UK that the VIP discipline is a very good thing
 - Avoid Drupal 8 upgrade (sorry Drupal)
 - “Always on the tip” (absolutely up-to-date) with VIP
- WordPress VIP did have some deficiencies – will miss ESI ... latency to USA ... less clean “distribution” concept

Next Stop – WordPress (with bells on)

Building on global experiences ...

newscorp.com

ballball.com

nypost.com



What is WordPress VIP?

- **It is wordpress.com but ... more customisable**
 - You can add your own non-public themes and plugins, subject to:
 - No direct database access or custom database tables
 - No server side cookie hacking
 - Some common calls have wrappers (e.g. fetching remote data, storing users data)
 - etc. (see link)
 - Your code releases are reviewed by wordpress.com staff
 - <http://vip.wordpress.com/documentation/getting-started/wordpress-com-vs-wordpress-org/>

Here we go ...



Challenges

Generic WordPress challenges

- Efficient Syndication of a thousand posts, 6000 thousand images and tens of thousands of third party posts from thousands of authors, each tagged with tens to hundreds of classifications (topic, location, people etc) per day
- Custom Post types vs Standard Posts
- Which plugins ...

News

Challenges

With VIP

- No naughty core or template hacks, tight code reviews, longer review cycles if developers are “bad” delaying launches
- Integration to our authentication database (2M+ users) - Faux Single Sign-On (?)
- Paywall
- Build pipeline integration (stash, bamboo etc) while tracking moving target of VIP code base.
- Tens of large sites, many more small ones, review of code in each site independently

Me

Domain Architect,
Content and Delivery Platforms

- Jeremy.kelaher@news.com.au
- Jeremy.kelaher@togetheronline.com.au
- @kelaher <https://twitter.com/kelaher>
- <https://www.linkedin.com/in/jeremykelaher>
- strangedevices.wordpress.com



A Shameless plug, and questions



Calling on all awesome WordPress people – contact us if you would like to be part of doing WordPress at scale 😊