Florida DrupalCamp | February 21-23, 2025

Florida Technical College in Orlando

# Adopting Fortune 500 Scaling Tactics from Day One

## Hector Iribarne @hectoriribarne
Track: Site Building
Level: Intermediate

# Hector Iribarne
# https://www.drupal.org/u/iribarne

## Drupal Site Builder & Developer

**Broward Drupal co-founder (April 2010)**
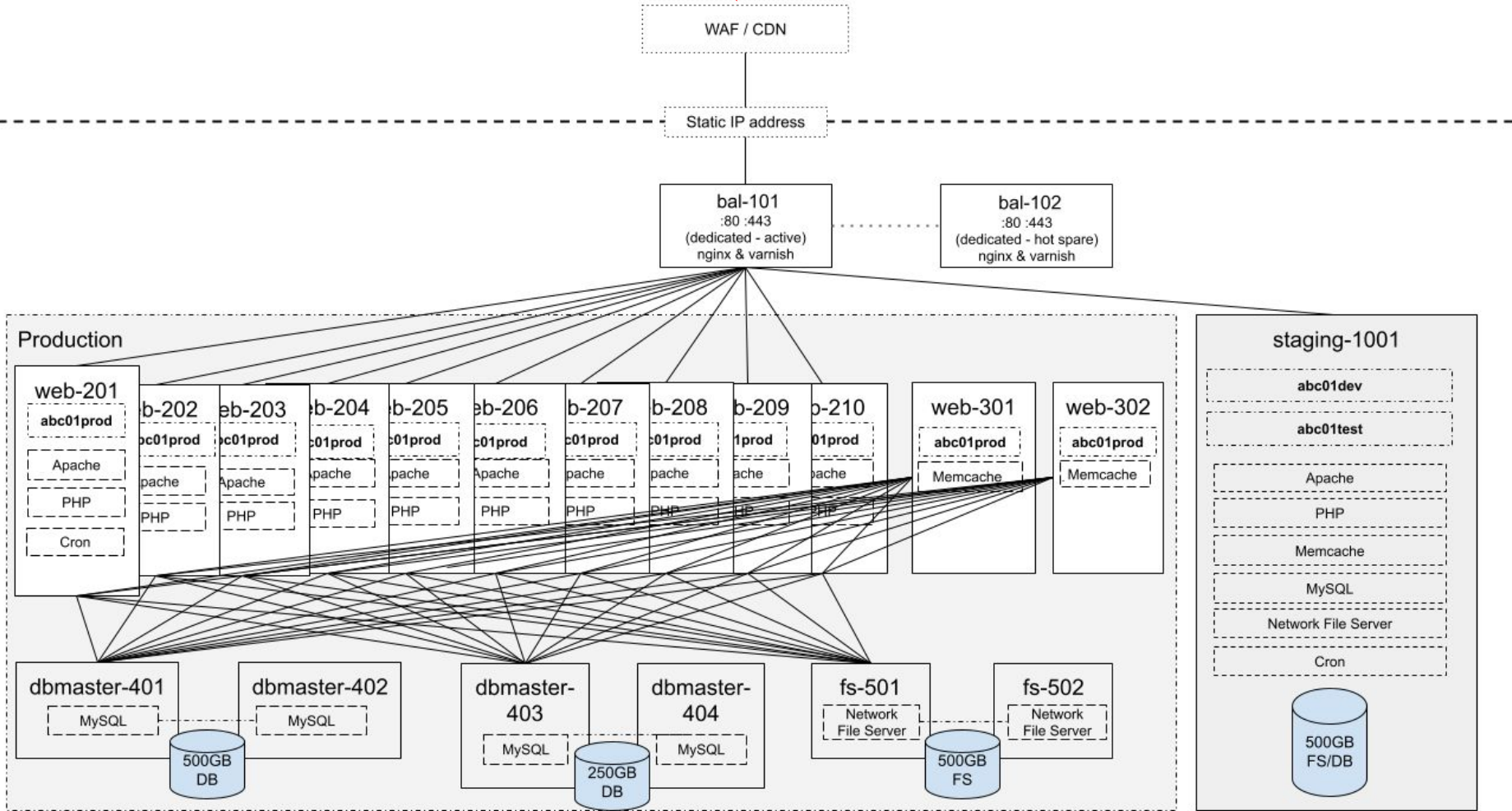
**iribarne on Drupal.org (since 2007)**

**Contributed Contentin and Contentout modules (sandbox modules)**

High–level Summary of presentation in two slides:

# From shared hosting to this: (while building an efficient application)

# Save $$$

# Main Topics to Save $$$

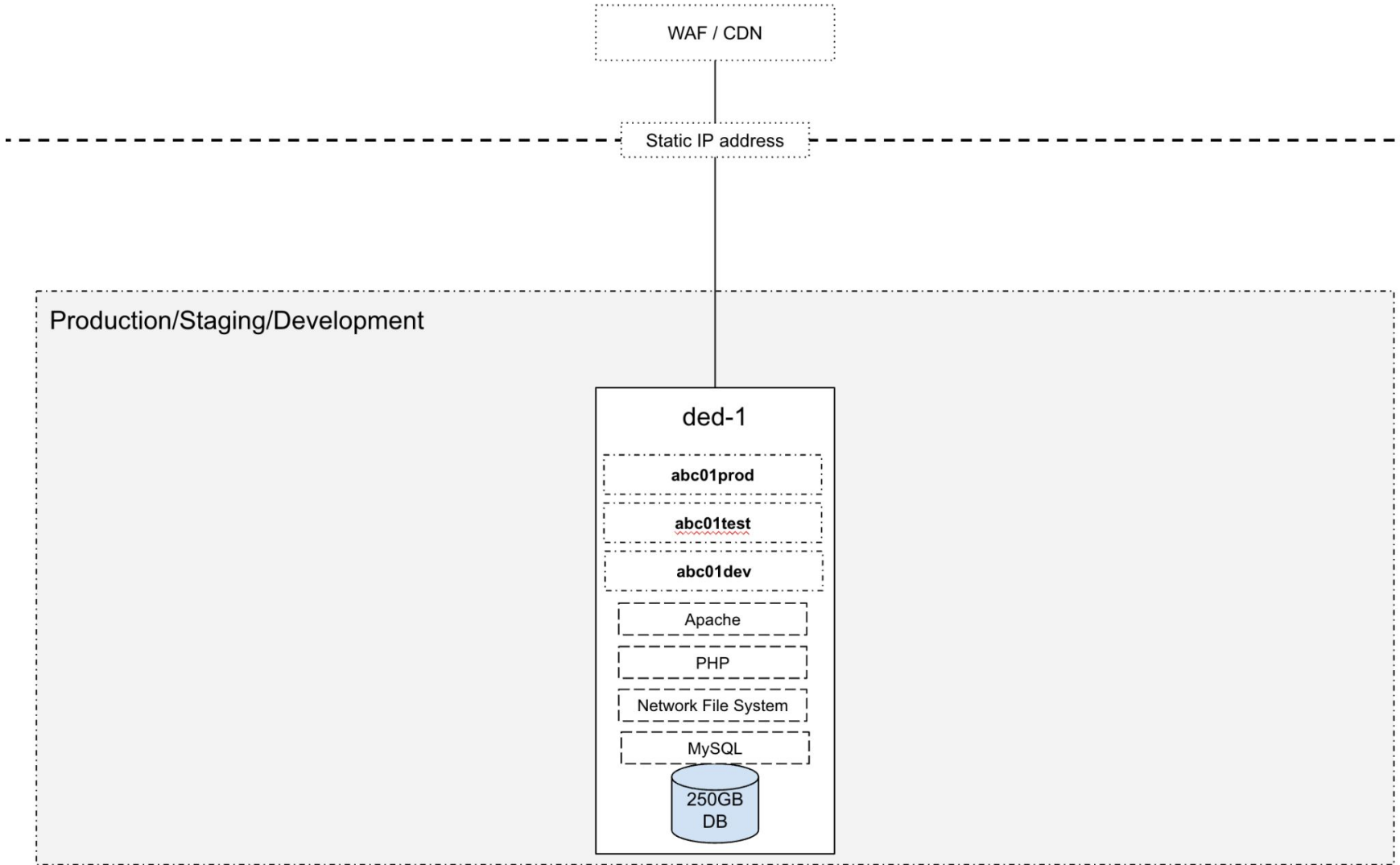**1**   Scaling Infrastructure

**2**   Application Efficiency

**3**   Customer Examples

# Scaling Infrastructure

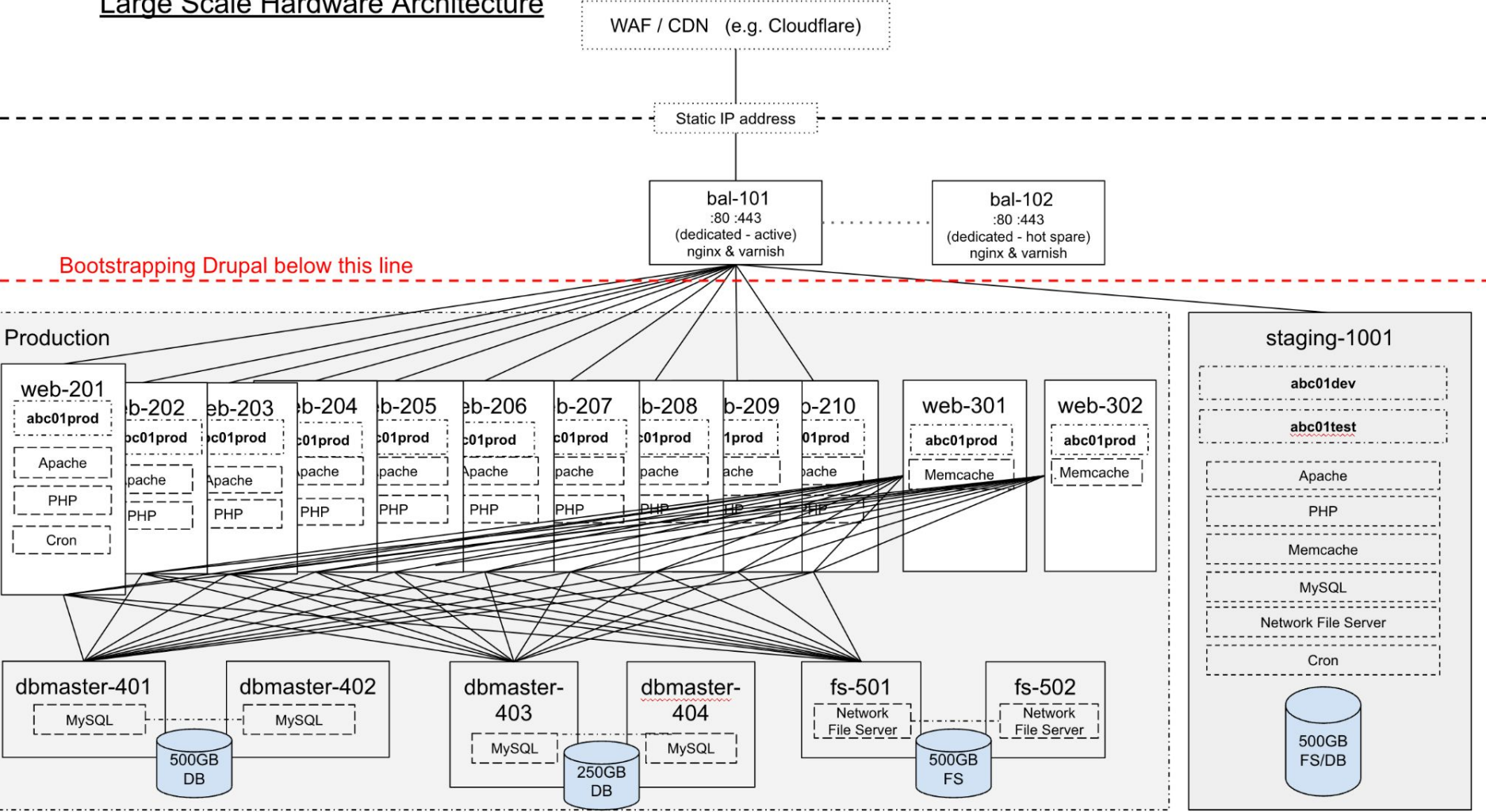# Network Diagram - Day 1 (Scale vertically as needed)

# Network Diagram - Day 1 + n

WAF / CDN

Static IP address

bal-1
Ports :80 :443

Production/Staging/Development

ded-1

abc01prod

abc01test

abc01dev

Apache

PHP

Network File System

MySQL

250GB
DB
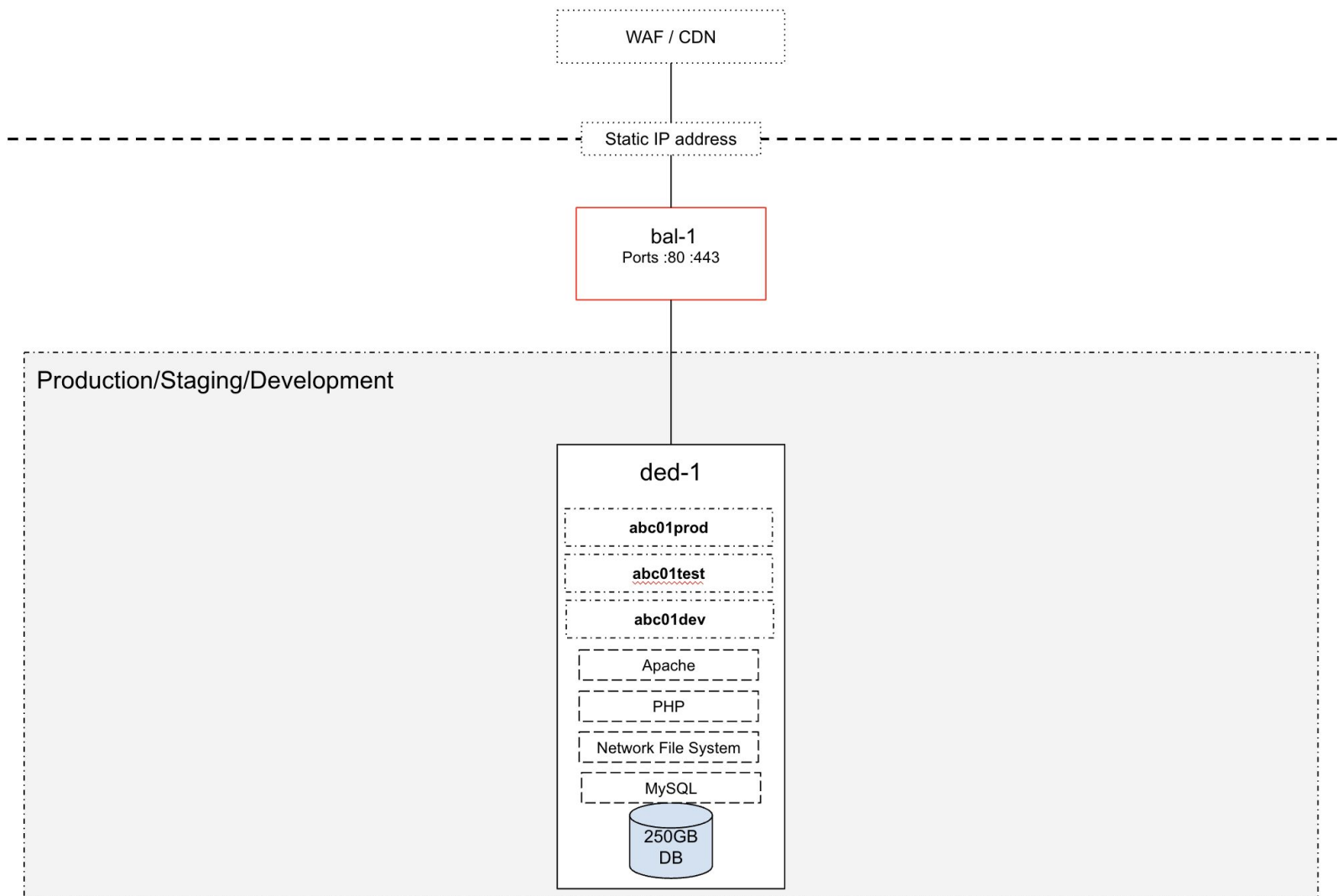
# Network Diagram - Day n$^2$ (Avoid if possible?!)

Large Scale Hardware Architecture

WAF / CDN  (e.g. Cloudflare)

Static IP address

Bootstrapping Drupal below this line

bal-101
:80 :443
(dedicated - active)
nginx & varnish

bal-102
:80 :443
(dedicated - hot spare)
nginx & varnish

Production

web-201
abc01prod
Apache
PHP
Cron

web-202
abc01prod
Apache
PHP

web-203
abc01prod
Apache
PHP

web-204
abc01prod
Apache
PHP

web-205
abc01prod
Apache
PHP

web-206
abc01prod
Apache
PHP

web-207
abc01prod
Apache
PHP

web-208
abc01prod
Apache
PHP

web-209
abc01prod
Apache
PHP

web-210
abc01prod
Apache
PHP

web-301
abc01prod
Memcache

web-302
abc01prod
Memcache

dbmaster-401
MySQL
500GB DB

dbmaster-402
MySQL

dbmaster-403
MySQL
250GB DB

dbmaster-404
MySQL

fs-501
Network File Server
500GB FS

fs-502
Network File Server

staging-1001
abc01dev
abc01test
Apache
PHP
Memcache
MySQL
Network File Server
Cron
500GB FS/DB

# Application Efficiency

**Bootstrapping Drupal is . . .**

- beyond the static cache of

  ○ The WAF/CDN (e.g Cloudflare)

  ○ Load balancer

# Bootstrapping Drupal is Expensive!

# How most people prioritize building applications (upside down)

# 1. Features/Functionality

# 2.  Performance

# 3. Security

**The way applications should be built**

# 1. Security

# 2.  Performance

# 3. Features/Functionality

## Security:

- Code governance - Update Drupal Core and contrib module when security updates are released in a timely manner.

- Use a Web Application Firewall (WAF) to restrict/block bad actors.

- Monitor web server logs (e.g. Apache Access logs) to identify bots and ai agents that should be blocked at the WAF level.

## Day 1 Application approach:

- Before writing any custom code, use the Drupal Entity system from Core and contrib to build a structured data model (i.e. create an Entity Relationship Diagram)

- Use the ECA (Event-Condition-Action) module to create complex automation workflows without writing custom code. It features a visual workflow builder, integrates seamlessly with Drupal entities and is highly extensible.

- Try to make your site 100% statically cacheable. Use GET requests instead of POSTs, OPTIONS, etc.

# Performance Best Practices:

- Disable dblog and and enable syslog on production.

- Do not reuse field names. Create unique field names per content type. Use a four letter prefix (e.g. for example, evnt_field_name for an event content type)  for every field in the content type, which will help you create better performing Views.

- Set the PHP memory limit as low as possible, ideally 128M, and conditionally increase the PHP memory limit higher for pages, processes, etc., that require more memory.

- Turn off revisions whenever possible, or at least limit them.

# Under the hood (Create a basic site and monitor it):

- Steps to creating a site
  - Add custom domain [www.iribarne.net](www.iribarne.net) to Cloudflare using authoritative DNS (i.e. use Cloudflare Nameservers).

  - At the domain registrar, change the default nameservers to the ones provided by Cloudflare.

  - In Cloudflare, add two A records for the bare domain and the www pointing to IP address of the your balancer or server instance on VPS.

  - Install Drupal 10.3 Umami on your server instance.

# Traffic that gets bootstrapped by Drupal (past Cloudflare/LB):

## Uncached Requests by Domain time series

Normal / Notice

This section aggregates the requests by domain and plots them over time. For applications that have multiple domains, seeing which domain is receiving the most traffic for a specific period of time can be helpful in correlating issues during high traffic events.

This chart displays the top 10 domains that have generated the most requests. Requests from all other domains (if more than 10) will be aggregated in a category called `Other`. This data was aggregated from the `Apache Access` log.
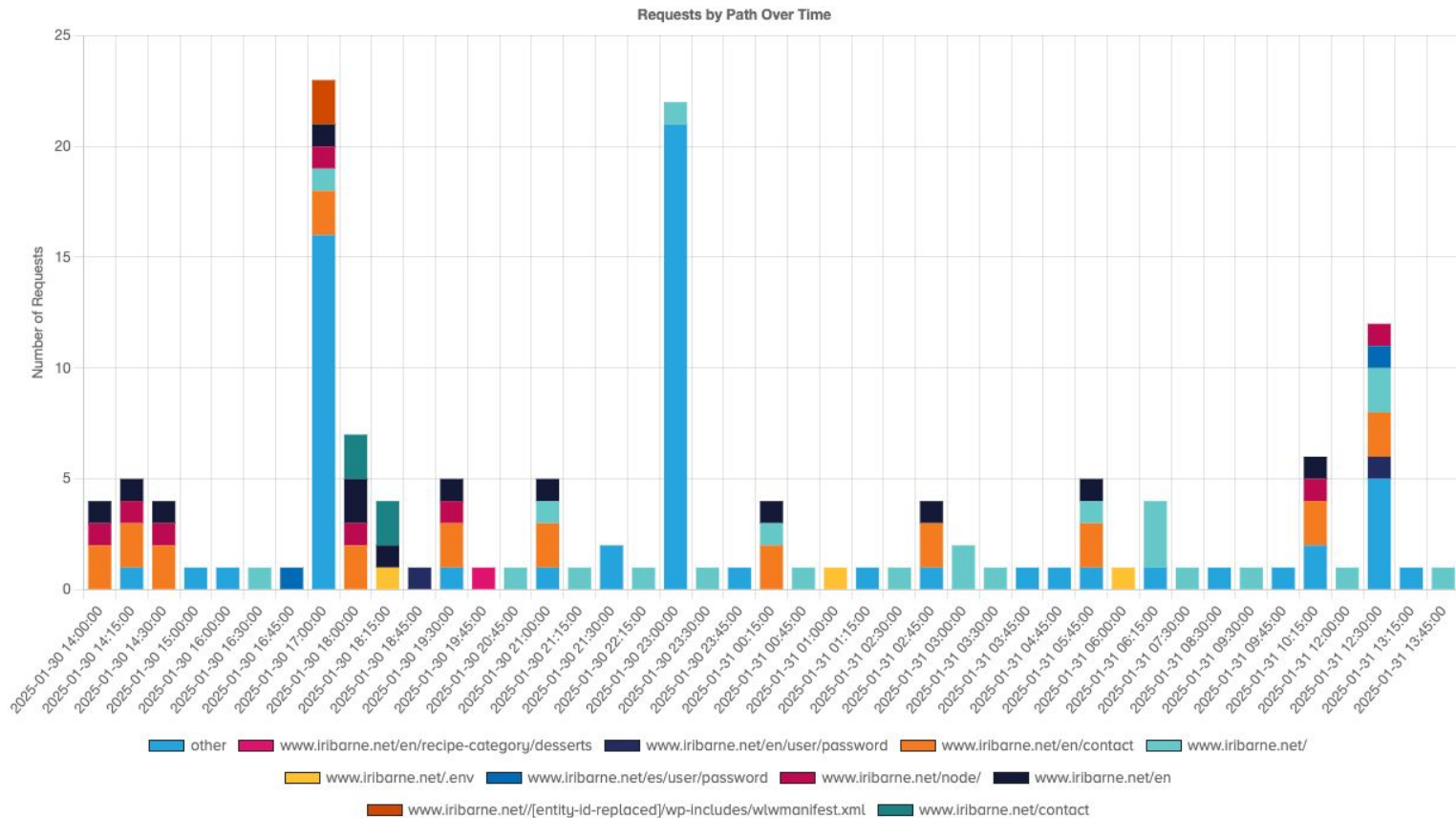


Uncached requests by domain

www.iribarne.net
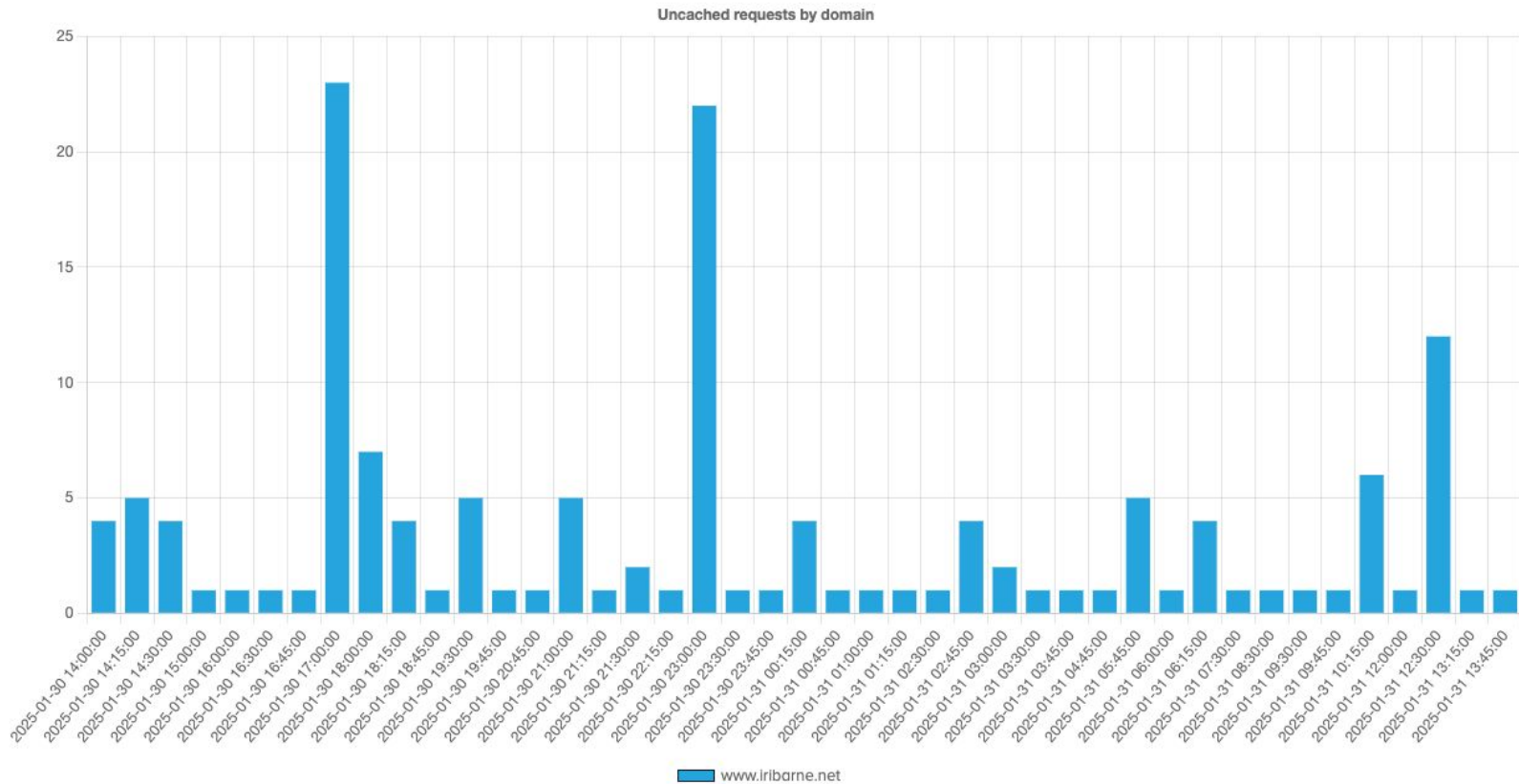
# Top paths bootstrapped by Drupal (past Cloudflare/LB):

Source: https://github.com/drutiny/drutiny     #FLDC25

# Top 10 40X requests bootstrapped by Drupal (past CLF/LB):

## Top Drupal 40X Requests

Normal / Notice

This section aggregates the client errors (400 level response codes). In Acquia Cloud 404 responses are automatically cached in Varnish for 15 minutes, any other 400 level status code is not cached. This list represents client errors that Drupal is serving. Each of these requests is bootstrapping Drupal, review this list and remediate accordingly.

This data was pulled from the `Drupal Requests` log.

A total of 23 client errors were served by Drupal during the time window.

### Legend

- Domain - the domain used to make the request.
- Path - the Drupal path requested.
- Variation - Probability (between 0-1) that query parameters produced a distinct URL. High query variance results in lower cache hit rates.
- Status Code - the 400 level HTTP response code.
- Requests - the number of times the path was accessed during the time window.

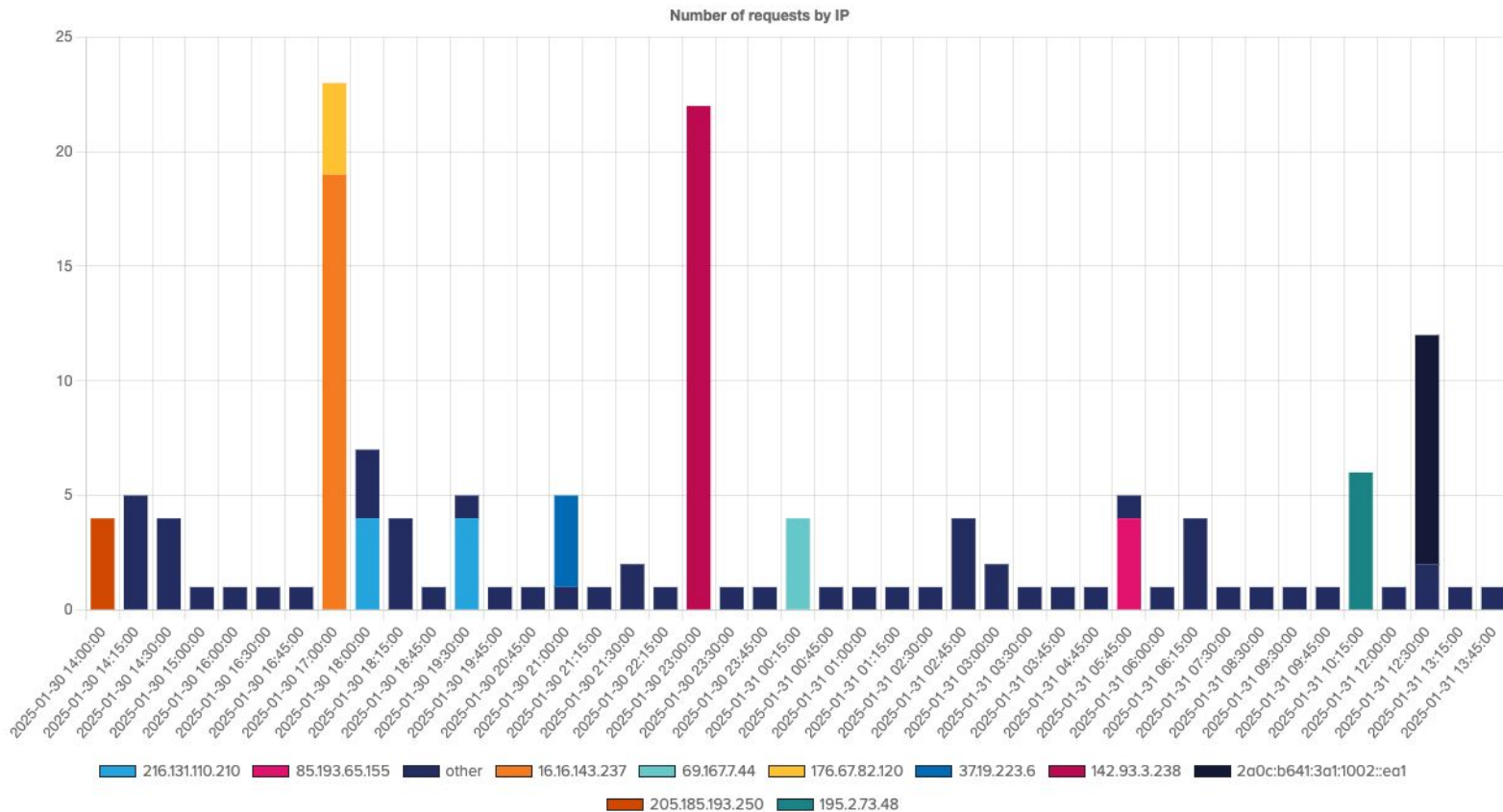| Domain | Path | Variation | Status Code | Requests |
|---|---|---|---|---|
| www.iribarne.net | /wp | 1.00 | 404 | 1 |
| www.iribarne.net | /test | 1.00 | 404 | 1 |
| www.iribarne.net | /wordpress | 1.00 | 404 | 1 |
| www.iribarne.net | /media/system/js/core.js | 1.00 | 404 | 1 |
| www.iribarne.net | /wp-admin/core/install.php | 1.00 | 404 | 1 |
| www.iribarne.net | /demo | 1.00 | 404 | 1 |
| www.iribarne.net | /dev | 1.00 | 404 | 1 |
| www.iribarne.net | /wp-admin/setup-config.php | 1.00 | 404 | 1 |
| www.iribarne.net | /cms | 1.00 | 404 | 1 |
| www.iribarne.net | /wp-includes/js/jquery/jquery.js | 1.00 | 404 | 1 |

# Top 10vClient IPs bootstrapped by Drupal (past Cloudflare/LB):

## Top Client IPs time series

Normal / Notice

This section aggregates the requests by IP address and plots them over time. Seeing which IP is generating the most traffic for a specific period of time can be helpful in correlating issues during high traffic events.

This chart displays the top 10 IP addresses that have generated the most requests. Requests from all other IPs will be aggregated in a category called `Other`. This data was aggregated from the `Apache Access` log.



Number of requests by IP

Legend: 216.131.110.210, 85.193.65.155, other, 16.16.143.237, 69.167.7.44, 176.67.82.120, 37.19.223.6, 142.93.3.238, 2a0c:b641:3a1:1002::ea1, 205.185.193.250, 195.2.73.48

Source: https://github.com/drutiny/drutiny   #FLDC25

# Top 10 Paths by memory bootstrapped by Drupal (past CLF/LB):

## PHP Memory Usage by Path

Normal / Notice

This section provides PHP memory usage information on a per-path basis. The data shows which paths are consuming the most memory and is sorted by max memory usage. By default, the maximum amount of memory allocated to PHP is `128 Mb` but the memory limit can be changed in Acquia Cloud or in Drupal to accommodate your application. Any path that has reached or is nearing the memory limit should be optimized to reduce memory consumption.

This data was aggregated from the `PHP-FPM access` log and the `Drupal Requests` log.

### Legend

- Domain - the domain used to make the request.
- Path - the Drupal path requested.
- Requests - the number of times the path was accessed during the time window.
- Minimum - the minimum amount of memory that was required to process the path.
- Average - the average amount of memory that was required to process the path.
- Maximum - the maximum amount of memory that was required to process the path.

| Domain | Path | Requests | Minimum | Average | Maximum |
|---|---|---|---|---|---|
| www.iribarne.net | /en/ | 1 | 16.00 MB | 16.00 MB | 16.00 MB |
| www.iribarne.net | /wp-content/themes/konzept/includes/uploadify/upload.php | 1 | 16.00 MB | 16.00 MB | 16.00 MB |
| www.iribarne.net | /en/contact | 24 | 4.00 MB | 7.92 MB | 14.00 MB |
| www.iribarne.net | /en | 13 | 6.00 MB | 9.23 MB | 12.00 MB |
| www.iribarne.net | /wordpress | 1 | 12.00 MB | 12.00 MB | 12.00 MB |
| www.iribarne.net | /wp-admin/setup-config.php | 1 | 12.00 MB | 12.00 MB | 12.00 MB |
| www.iribarne.net | /contact | 4 | 4.00 MB | 6.50 MB | 10.00 MB |
| www.iribarne.net | /media/system/js/core.js | 1 | 10.00 MB | 10.00 MB | 10.00 MB |
| www.iribarne.net | /en/big_pipe/no-js | 1 | 8.00 MB | 8.00 MB | 8.00 MB |
| www.iribarne.net | /wp-includes/js/jquery/jquery.js | 1 | 8.00 MB | 8.00 MB | 8.00 MB |

**@hectoririribarne**

Source: https://github.com/drutiny/drutiny    #FLDC25

# To finish off Application Efficiency:

## What is the biggest cause of site downtime?

# The biggest cause of site downtime is . . .

# Custom Code

# Customer Examples

# A News site with a JS frontend launches with POST requests to the Drupal backend

**This caused downtime. How do we fix it?**

A News site with a JS frontend launches with POST requests to the Drupal backend

Change the POST request to GET requests

#FLDC25

# A chip company uses Paragraph revisions without any limits

**This caused downtime. How do we fix it?**

A chip company uses Paragraph revisions without any limits

Correcting this required truncating tables in the DB. The customer then added limits to the number of revisions to keep

#FLDC25

An insurance company's homepage for their agent portal had an auto login feature that redirected every request

**This caused downtime. How do we fix it?**

An insurance company's homepage for their agent portal had an auto login

feature that redirected every request

Remove auto login functionality and instead create a block on the homepage

that the agent has to click on to login

#FLDC25

An energy company with residential customers has third-party API calls with default values for PHP timeout and the third-party server is not responding

**This caused downtime. How do we fix it?**

An energy company with residential customers has third-party API calls with default values for PHP timeout and the third-party server is not responding

Create a short PHP timeout when calling the third-party API and cache the responses to reduce the number of external calls

**If you take one lesson away from this presentation . . .**

# BootstrappingDrupalisExpensive.com

# Links (Page 1 of 2):

- Cloudflare WAF/CDN (Sign up for a free account to get started)
    - https://www.cloudflare.com/

- New Relic - Monitoring of applications and infrastructure
    - https://newrelic.com/

- BOA (Script for installing Aegir on a VPS)
    - https://github.com/omega8cc/boa
    - https://www.slideshare.net/slideshow/installing-boa-on-ubuntu-1204-lts/19906398

- Drutiny (Drupal Scrutiny) - Open-source site auditing tool
  - https://github.com/drutiny/drutiny
- Contentout sandbox module
  - https://www.drupal.org/sandbox/iribarne/2461613
  - git clone --branch 7.x-1.x http://git.drupal.org/sandbox/iribarne/2461613.git contentout
- Content migration via the UI from one Drupal site to another via CSV files for each content type
  - http://www.slideshare.net/hectoririribarne/drupal-camp-atlanta-2015-no-code-content-migration
- Drupal Camp Atlanta 2015 - No Code Content Migration
  - https://drupal.tv/events/drupalcamp-atlanta/drupalcamp-atlanta-2015/dcatl-2015-no-code-content-migration

# Links (Page 3 of 3):

- Broward Drupal Meetup via Google Meet - March 12, 2025, 7pm ET

    - https://groups.drupal.org/node/537164

- Broward Drupal on GDO

    - https://groups.drupal.org/broward-drupal

- Follow us on social media

    - https://x.com/browarddrupal

# Questions?

**Hector Iribarne    @hectoriribarne**

# If you liked my presentation, follow me on X and Bluesky: @hectoriribarne

## Details:

## Go Fish:

BrowardDrupal.org
Co-Founder

@hectoriribarne

Hector
Iribarne

d.o. username:
iribarne

http://certifiedtorock.com/u/192646

hector@browarddrupal.org

#FLDC25