Survival kit for CERN's organic webscape: Kubernetes and a community of Makers

Konstantinos SAMARAS-TSAKIRIS
Catharine NOBLE, Francisco BARROS, Rajula VINEET
Tier-0 (CERN and Hungary):
data recording, reconstruction and distribution

Tier-1: permanent storage, re-processing, analysis

Tier-2: Simulation, end-user analysis

1 PB/s of data generated by the detectors
Up to 60 PB/year of stored data

Large experiments have managed data sets of >200 PB

A distributed computing infrastructure of order of a million cores working 24/7
An average of 60M jobs/month

An continuous data transfer rate of 35-45 GB/s
(3 PB/day) across the Worldwide LHC Grid (WLCG)
Public outreach
home.cern: 30k unique daily visitors
Peak: 120k
Peak: 55k
Charting our journey

Web team, infra team & site builders

Tour of the Drupal service

Conflict! Shrinking web team... tech debt... Drupal 8!

Technical and Human solutions

Beyond the horizon

How well did we do?

Beyond the horizon
The actors

Web team, infra team & site builders
Once upon a time at CERN...

Site builders:
- low-code
- composable
- consistent

Web team
→ CERN theme
→ Integration with CERN services

Infrastructure team
→ home-grown, Linux infrastructure
→ provision and maintain >1k Drupal sites
Why so many websites?

So many choices...

I have no idea

- Academic freedom
- Autonomous entities
- Plethora of tools

→ as many websites as people!

Most of them for coders
I JUST WANT A WEBSITE

Can I get one?
Please?
The Setting
Physical infrastructure
Journey of an HTTP request
Shortcomings

- Little isolation: sharing the same PHP-FPM process pool
- Multisite: sharing Drupal core between all websites, little customization
- ad-hoc, mutable, non-standard
- not scalable
- not very robust

Lacks all of the Cloud Native benefits

But for many years it was running
Crisis!
Smaller team, tech debt... Drupal 8 😳
Drupal 8
Our actions
Kubernetes and a community of Makers
Our ideal infrastructure...

- Isolation
- Tunable resources
- More flexibility to users
- Reliable hosting
- CERN Drupal Distribution
- Reuse infrastructure components
- Automatic upgrades & Failure recovery
- Custom modules & themes
Modernization campaign
Crew

Cloud Native Drupal Infrastructure

Requirements

Shared PaaS components

Crew
A Cloud Native facelift
A Cloud Native facelift
A Cloud Native facelift

Building a cloud-native SaaS for CERN’s 1k Drupal sites

Check out our Kubecon presentation for more details:
https://youtu.be/4O-pcSQR8Vw
Community from a chore
We can only do so much!

Formal support channels overloaded

Knowledge attrition
- As teams grow smaller, expert knowledge gets lost

The stack overflow model works
- Few dedicated experts, lots of disseminated know-how

People reach out to each other
- CERN values collaboration
50 instances → more than 40k users

Did I Make Something Wrong Or Should I Add Any Modules?
I Created A Basic Page

II Possible To Export Webform Results? If Yes, How?
I Don’t Fancy Creating A New Node Just To Hold A Redirect (especially As I’ve Been Working On Cleani
I Have An Existing Fag Page To Which I Would Like To Add A PDF Document At The Top Of The Page
And Now Would Like To Give Another Person Permission To Edit It.
It Seems That The Simplestamplphp Module Isn’t Synchronising With The E-groups

I See This Error When Users Try To Sign In With An ‘external’ Account. What To Do?

I Would Like To Delete A Website. But I Need To Keep The Webform Results.
Hello @here. I’m Having Synchronisation Issues With The Atlas Collaboration Page

How Can I Give Editing Permission To A Specific Node To Another Person?

Anyone Know How To Set Up Redirects From A Drupal Node That No Longer Exists?
I Have Set Up Following The Outage. Nothing Was Changed In Terms Of Configuration.
I Would Like To Use The Same Background Graphics For The Cern Qni Website Like The One Used On Home.

There Has To Be A Better Way...
Could You Please Tell Me Where I Could Find It.
I Received This Error Message When I Tried To Change The Website Logo.
Progress ?
## Development timeline

<table>
<thead>
<tr>
<th>start date</th>
<th>end date</th>
<th>event</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020/10/09</td>
<td></td>
<td>Manual deployment (template) of Drupal server resources</td>
</tr>
<tr>
<td>2020/10/15</td>
<td></td>
<td>Scaffold the DrupalSite operator with operator-sdk</td>
</tr>
<tr>
<td>2020/11/18</td>
<td>2020/12/03</td>
<td>Site provisioning with operator</td>
</tr>
<tr>
<td>2020/12/02</td>
<td>2021/02/10</td>
<td>Implement basic features (site install, environments, custom builds)</td>
</tr>
<tr>
<td>2021/02/12</td>
<td>2021/02/20</td>
<td>Pilot migrations</td>
</tr>
<tr>
<td>2021/03/15</td>
<td>2021/07/05</td>
<td>Start Database integration</td>
</tr>
<tr>
<td>2021/03/20</td>
<td>2021/05/29</td>
<td>Automate Drupal updates</td>
</tr>
<tr>
<td>2021/03/25</td>
<td>2021/03/30</td>
<td>Scale tests, Ceph integration issues. Kubecon presentation.</td>
</tr>
<tr>
<td>2021/04/28</td>
<td>2021/06/01</td>
<td>Define release model of CERN Drupal Distribution</td>
</tr>
<tr>
<td>2021/05/04</td>
<td>2021/06/26</td>
<td>Integrate OIDC flow in Drupal sites</td>
</tr>
<tr>
<td>2021/07/05</td>
<td>2021/07/15</td>
<td>Implement backup &amp; restore</td>
</tr>
<tr>
<td>2021/07/09</td>
<td></td>
<td>Migration process reaches maturity</td>
</tr>
</tbody>
</table>
Fun with integrations
How will this look like in the future?

Migration underway

All our websites are now previewed

Users excited with the new features!
- custom composer projects
- Drupal 9 previews
Drupal community is engaging especially in pandemic times!

- **Monthly posts**
- **User Engagement**
  - Daily Active Users : Monthly Active Users
Beyond the horizon
Technical goals
- new Drupal use cases (Drupal as a backend)
- deploy other low-code website solutions
- give back to community: operator that works outside of CERN
Is this infrastructure useful to You?

Poll
Thank you!
Operator

a custom resource and controller, conceptually similar to an OOP class with data & methods

"the world": what is running in the cluster

reconciliation

events

updates

controller

API

What should be in the cluster

Standard resource

Custom resource

DrupalSite

ingress

pod

job

etcd