Taming Failures by Partitioning Time and Space

Sadek Drobi

prismic.io CEO & Co-founder Playframework Co-creator @sadache

prismic.io, a snappy tour

Building a Website with dynamic content is a pain

No content management

Use a visual builder tool

Use CMS building blocks and plugins







prismic.io is a new way to bring your website content alive

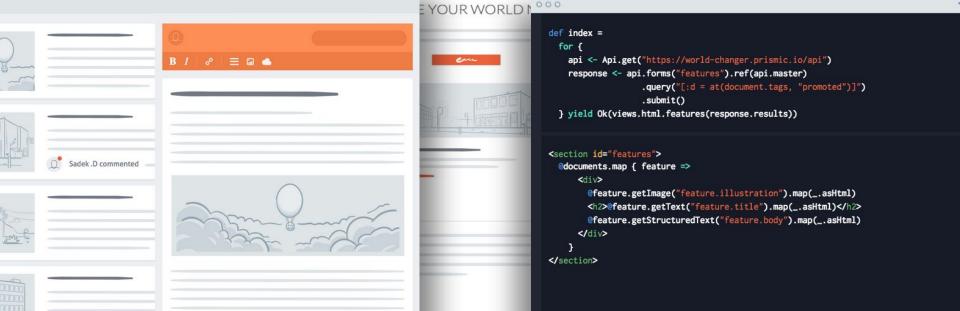


prismic.io is a new way to bring your website content alive

Content management backoffice

WORLDCHANGER

imple Web API to consume content



Content is Central to Websites

Business Wise

- Offer Performance and Reliability
- Excel where every CMS fails

First, we need to understand the impact of system failure.

Impact of System Failure

- Writing Room: No edit, save or publish
- API: No end-user website

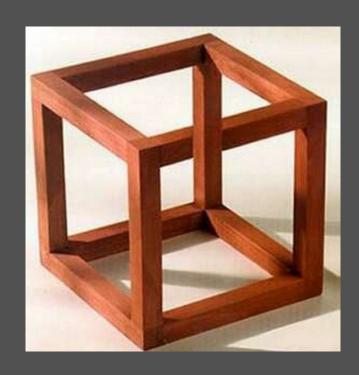
Impact of System Failure

- Writing Room: No edit, save or publish
- API: No end-user website

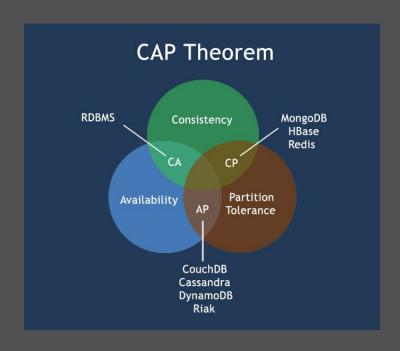
System Requirements From a Macro Perspective

- Concurrent Write
- Concurrent Read
- Redundant
- Reliable and available
- Consistent

System Requirements From a Macro Perspective



CAP says no!

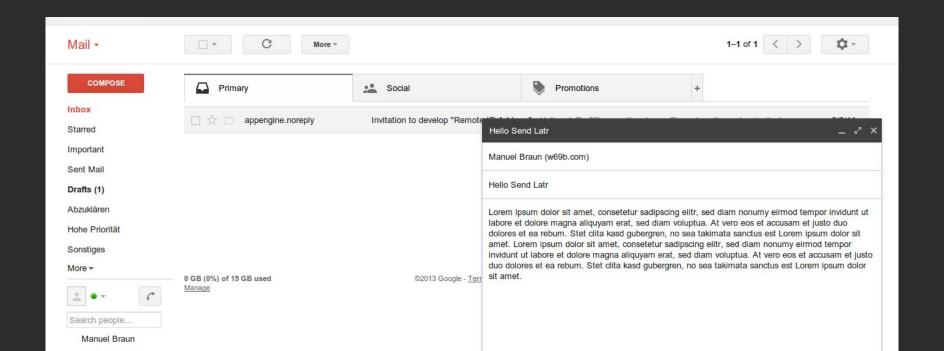


Let's Talk about Time

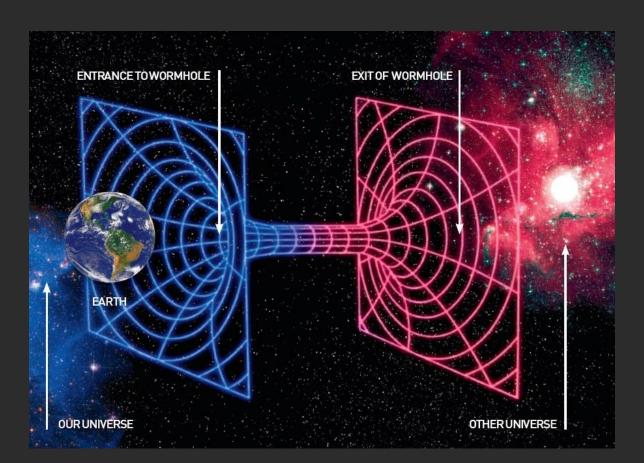
Observe time ticks

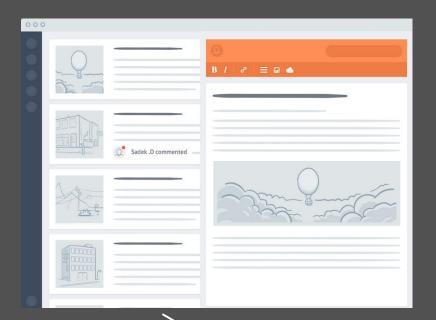


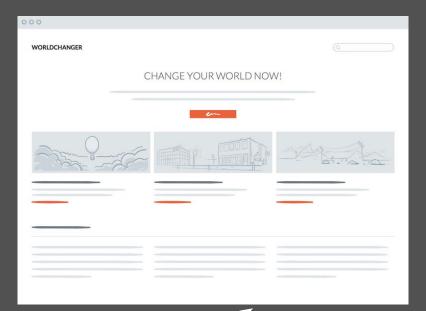
Observe time ticks



Two different universes, two different times







Authors write some content

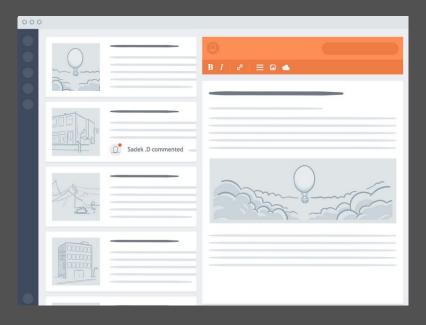


On Publish, Content is displayed on website

Two different universes, two different times

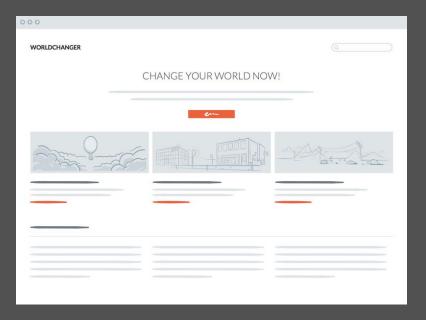


Writer universe

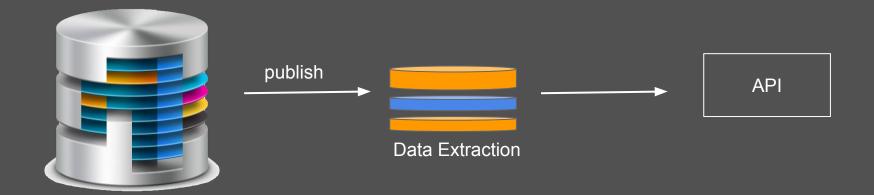




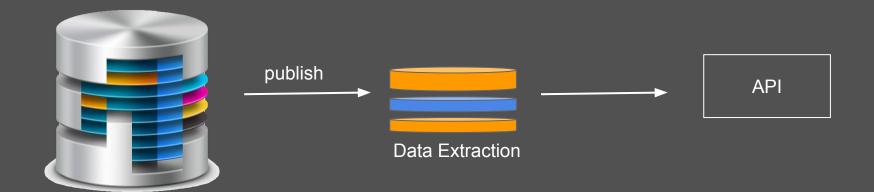
Viewer universe



On publish event, we update API system time

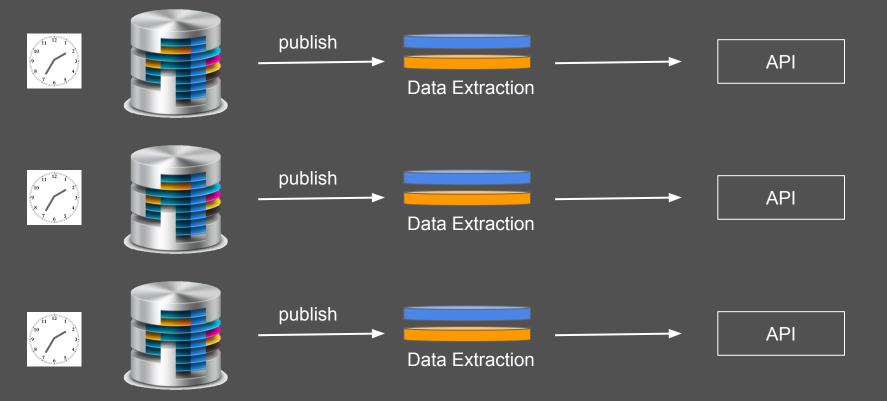


If Writing Room goes down, API isn't affected



We Can Do Even Better

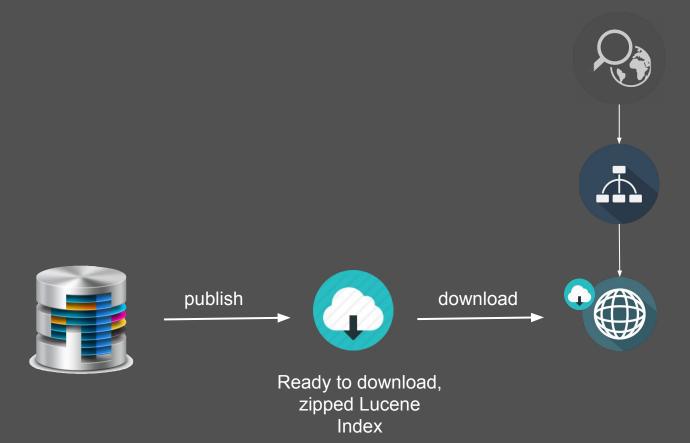
Clients APIs can be time decoupled too!



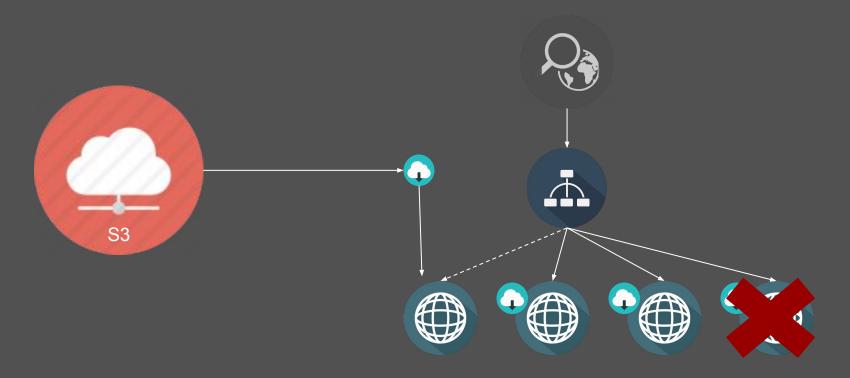
Infrastructure wise



We use Apache Lucene for our API



Server Failure?



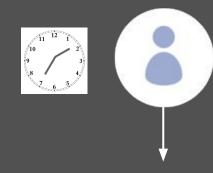
Elastic?

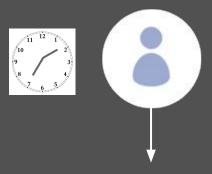
Do we need traffic officers or traffic lights here?



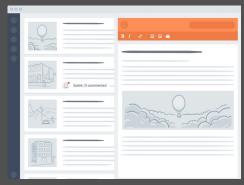
Applying the Same on the Writing Room

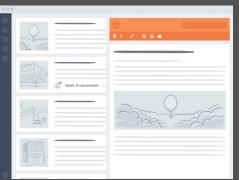


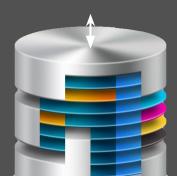


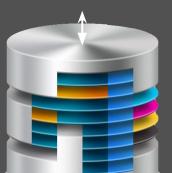


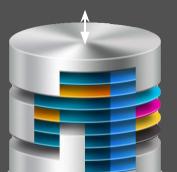






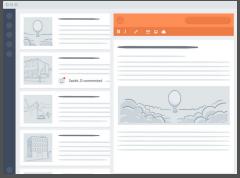




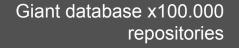








Interactive Read/Write



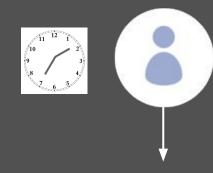


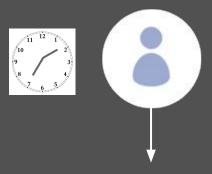
Did not connect for weeks



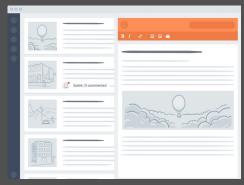
Does not connect anymore

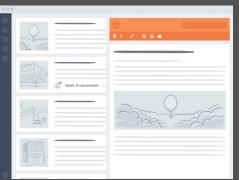


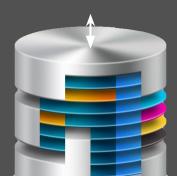


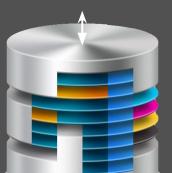


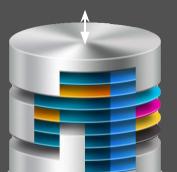


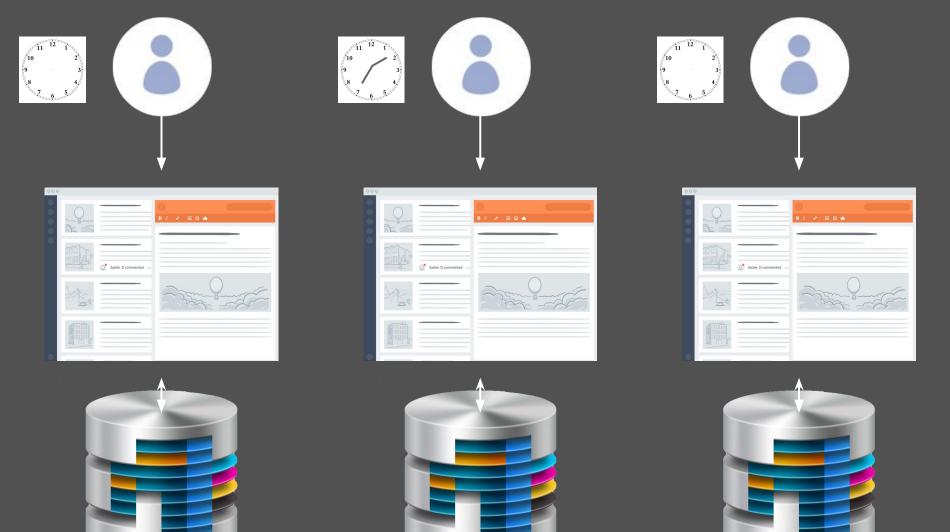






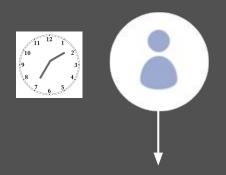


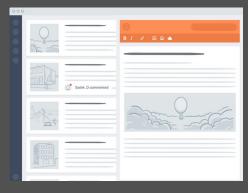




Do we really need to keep everything here in case of?

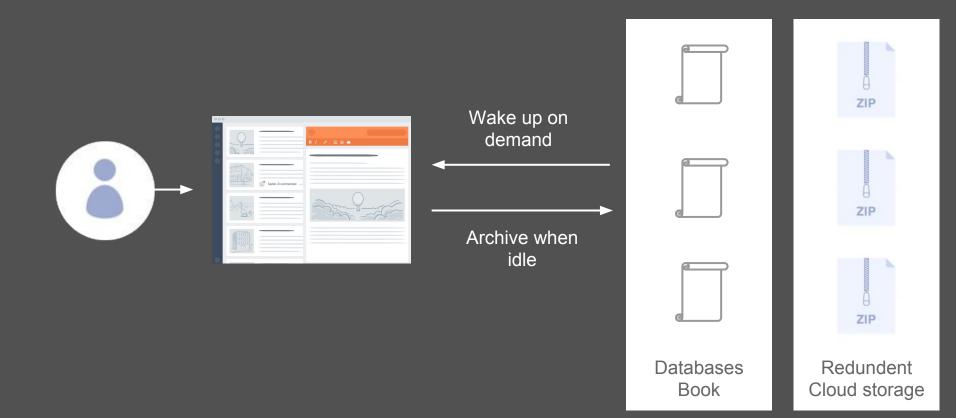




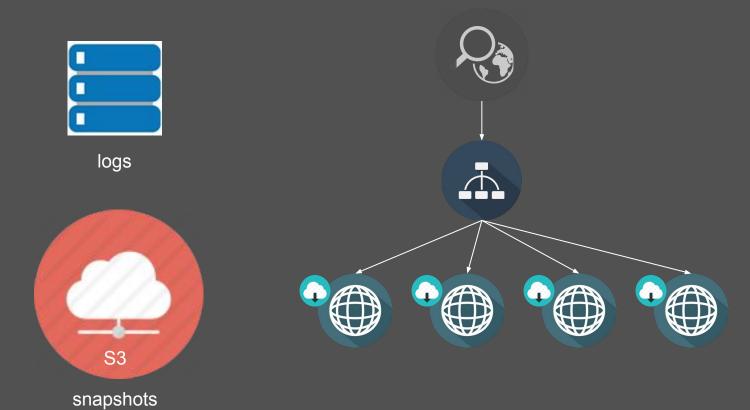




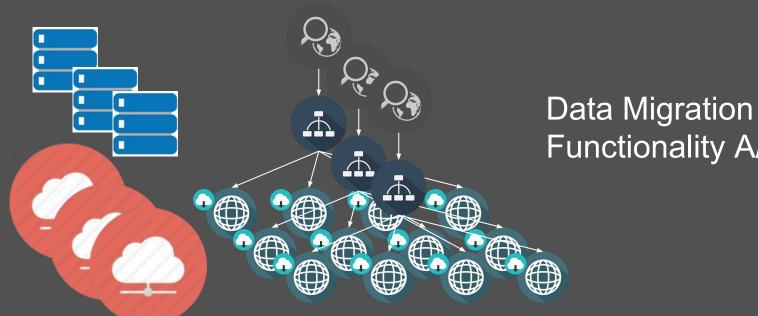
Each repository has it own lifecycle



Infrastructure wise



Clusters



Functionality A/B Testing

Observe time ticks



Solves

- Scalable space-efficient infrastructure
- Data migration and versioning (avoid stop-the-world migrations)
- Multiple production environments (with multiple data versions)
- Testing new functionalities

System Requirements From a Macro Perspective

- Concurrent Write
- Concurrent Read
- Redundant
- Reliable and available
- Consistent

Scalability is not a buzzword you can put on a slide nor

A technology you can purchase _____

