oyster.engineering /post/98970249193/half-a-million-books-in-your-browser-how-we-built

Half a million books in your browser: How we built the Oyster Web Reader. Part I: Infrastructure

4-5 minutes

image

This is part one in a series of posts on how we built the Oyster web reader. There are unique and fascinating challenges in bringing books to life in web browsers, and we'll outline some of those in this series: moving from our architecture, forward to our reader application, and then to the UI.

When reading any book on Oyster—whether in the app or in the browser—you're reading an electronic publishing file (EPUB). The EPUB format is a well-established standard for packaging and distributing ebooks. An EPUB file is essentially a ZIP-archive of web resources like HTML files representing chapters, CSS, XML metadata files, or images. So how do we get book content quickly and beautifully to our members' web browsers while maintaining the strict DRM protections that publishers require?

Break up the EPUB

While we send entire (encrypted) EPUB files to the iOS and Android apps, we can't do the same with web browsers. This is for a few reasons:

1. All the code to receive and manipulate the file is exposed in JavaScript. It would be trivial to step through the code and potentially redirect EPUB contents and compromise security requirements.

2. To get content to members quickly means sending as little as possible over the network. Some of our books are over 50MB… yikes.

We dealt with this by unzipping and then storing EPUB contents on S3, which allows us to load the required content on-demand. Why send all 50 chapters at the same time when someone just wants to read the first?

Clipper

Clipper is the service we built to serve these EPUB contents. We decided early on to use Apache's Thrift RPC library to generate the API and stubs for the web client and server. This decision accelerated development greatly by standardizing and simplifying the exposed interface.

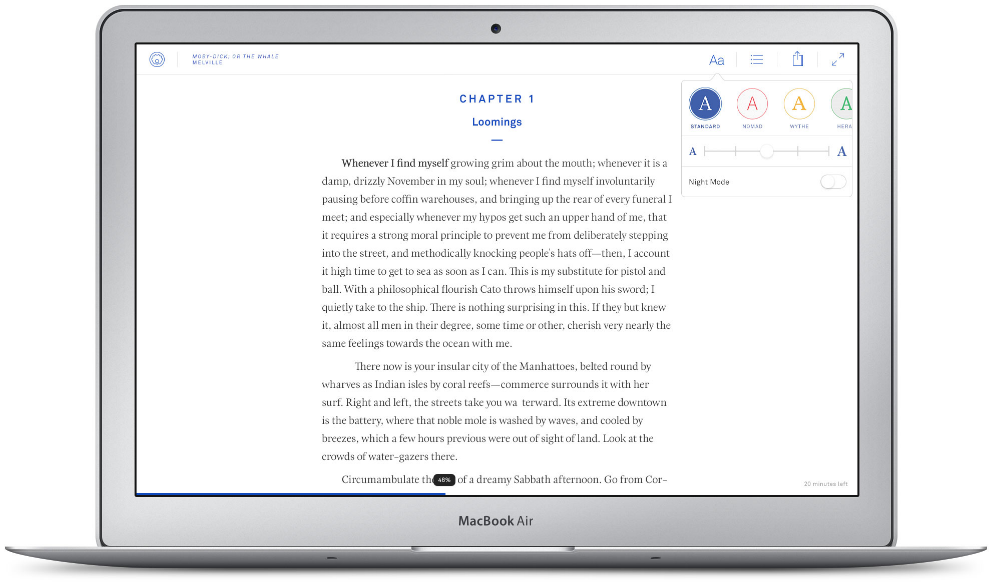
Combining our tendency to slip in nautical references wherever possible and the clip-clip functionality of scissors, we decided to call the API "Clipper." It performs a straightforward task and as quickly as possible, explemary of the UNIX philosophy of "Do one thing and do it well."

Deploy, monitor, and iterate

Shipping Clipper was a milestone for the web team. Up to this point, our entire web infrastructure was hosted exclusively on Heroku. We took it on as a team to migrate the entire service to AWS and leverage our existing tools to ensure that additional EC2 instances could be provisioned on the fly.

This was a big learning opportunity for the web team, and while it was tough climbing the learning curve under time pressure, this knowledge will pay off big dividends in the future.

We've since done much to stabilize the reader, increase test coverage, and increase performance. One thing we did was move dynamic font generation to a pre-processing deployment step, which cut our response times in half. Another consideration was caching. Since requesting EPUB contents is quite expensive—requiring downloading and processing content from S3—our Memcached cluster of about 260,000 Clipper responses have sped response times up by seconds on each request, and today it ensures that any book you'd like to start reading is a moment away.



CHAPTER 1

Loomings

Whenever I find myself growing grim about the mouth; whenever it is a damp, drizzly November in my soul; whenever I find myself involuntarily pausing before coffin warehouses, and bringing up the rear of every funeral I meet; and especially whenever my hypos get such an upper hand of me, that it requires a strong moral principle to prevent me from deliberately stepping into the street, and methodically knocking people's hats off—then, I account it high time to get to sea as soon as I can. This is my substitute for pistol and ball. With a philosophical flourish Cato throws himself upon his sword; I quietly take to the ship. There is nothing surprising in this. If they but knew it, almost all men in their degree, some time or other, cherish very nearly the same feelings towards the ocean with me.

There now is your insular city of the Manhattoes, belted round by wharves as Indian isles by coral reefs—commerce surrounds it with her surf. Right and left, the streets take you wa terward. Its extreme downtown is the battery, where that noble mole is washed by waves, and cooled by breezes, which a few hours previous were out of sight of land. Look at the crowds of water-gazers there.

Circumambulate th of a dreamy Sabbath afternoon. Go from Cor-

While Oyster launched version one of its web reader, we're still just at the beginning of what experiences we can impart with books. So if the possibilities excite you and you want to have a transformative impact on how people discover and read books on the web, then we would love to speak with you about how we can get there together. Email vivek@oysterbooks.com with how you see our stories intersecting.