

Seeking Performance for Digital Collections

Plum

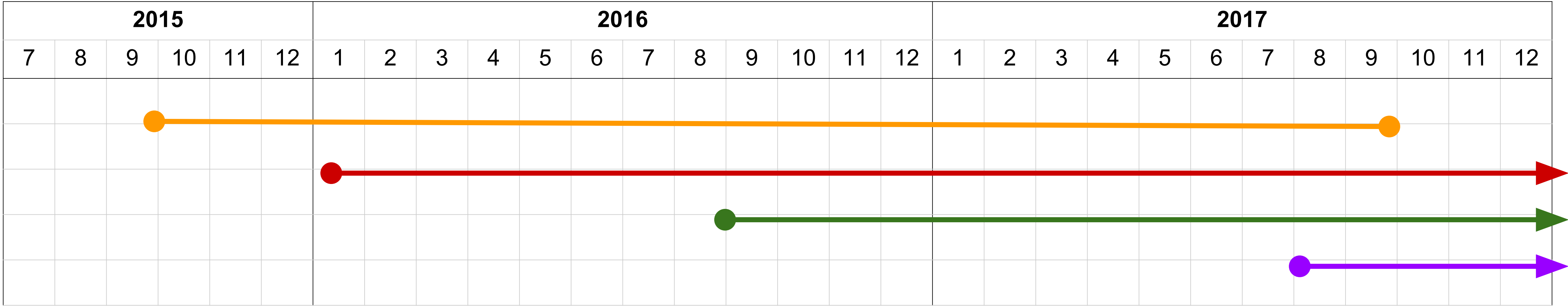
- Poor performance for bulk uploading, editing, migration
- Some performance and reliability issues resolved, but still not acceptable (minutes to save 500-page book)
- Hyrax is complex, and increasingly going in a different direction from our digital collections focus

<https://github.com/pulibrary/plum>

Fedora API Specification

- More innovation and better performance in the future
- But implementations are not ready today
- Fedora lacks key query functionality, which increases complexity (may be addressed in the future)
- Even if we could switch today, Hyrax is complicated by using Fedora and Solr, and keeping them in sync

<https://fcrepo.github.io/fcrepo-specification/>



Valkyrie

- Uses the DataMapper pattern for supporting multiple backends — simplifying the stack and separating persistence from models
- Metadata storage in Fedora, Solr, and PostgreSQL
- File storage in Fedora and local disk
- Doesn't require RDF or PCDM at the persistence layer (you can still use them for data interchange)
- Began as a "breakable toy", but now has contributions by 17 developers from 9 institutions (including Princeton, Penn State, Stanford, Northwestern, Columbia, Univ. of London)

<https://github.com/samvera-labs/valkyrie>

Figgy

- Plum ported to Valkyrie
- Dramatically better performance using PostgreSQL for metadata storage and network-mounted disk for file storage
- Initial production deployment: October 2017
- Feature parity with Plum, and migration of content, expected by the end of the year

<https://github.com/pulibrary/figgy>