A strongly consistent index for email using git and MirageOS

Oliver Hope

Jesus College

What?

What has been built?

- A library called *gitmaildir* — An OCaml library to interface with and perform operations on the type of maildir in a git store that I have created
- An application called *gitmaildir_cli* — A command line client to allow inteaction with a gitmaildir form the commandline
- An application called *gitmaildir_daemon* — A daemon that keeps a standard maildir in sync with a gitmaildir allowing applications that use standard maildirs to interact without needing further changes

Maildir is a well-known mail storage format for unix-like systems. Its benefits are:

- Allows lots of concurrent access
- Mail client does not have to worry as managed by the filesystem

However it has some downsides:

- We can never ensure that the directory is in a consistent state
- This makes operations that rely on indexing more difficult

How?

It turns out that MirageOS has some libraries that are useful in working with both git stores and email parsing using OCaml. (ocaml-git and MrMime)

Therefore, the entire project is written in OCaml making heavy use of these libraries.

Everything that modifies the gitmaildir is written using a library called Lwt. This is a form of promise and allows operations to be combined in a monadic style. Another benefit of this is that we can easily parallelize functionality when necessary.

What's left?

While the core features are complete, this provides a basis to implement more interesting extensions including:

- Separate branches for mobile clients that strip data-heavy content and attachments from emails
- Branches per reply thread for a better insight and cleaner overview of long-running email threads
- Intelligent rollback in case of accidental deletion or modification