Git Internals

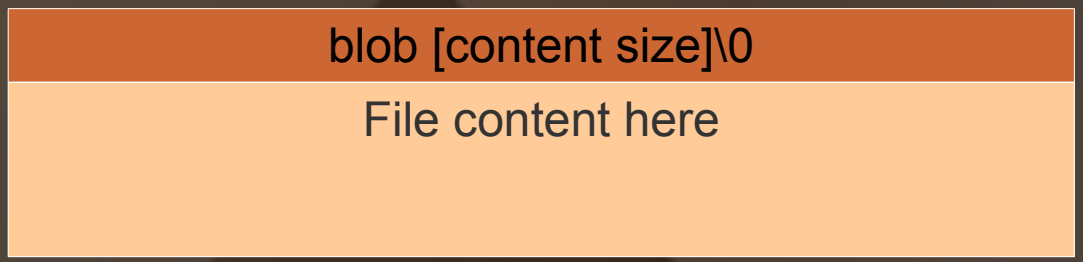
Khaled alHabache
khellls@gmail.com

Objects Types

- Blobs
- Trees
- Commits
- Tags

Blobs

- **Contents** of files are stored as blobs.
- Two files with same content will be stored once.
- The blob looks like:



- The blob is zipped and then SHA1 to get the name: 2219632881ae907adbba72cc883d05bae0028730

Trees

- Directories correspond to trees.
- A tree is a simple list of trees and blobs that the tree contains, along with the names and modes of those trees and blobs.

tree [content size]\0			
100644	blob	221963	file name

- The tree is zipped and then SHA1 to get the name: 342f6asdg1ae907adbba72cc8095dsds0028730

Commits

- The commit is very simple, much like the tree. It simply points to a tree and keeps an author, committer, message and any parent commits that directly preceded it.

commit [content size]\0
tree 342f6asd
author John Smith john@smith.com
committer John Smith john@smith.com
commit message

- The commit is zipped and then SHA1 to get the name: e12f3463g1ae907adbba72cc8095dsds0028730

Tags

- A tag is permanent shorthand name for a particular commit.
- It contains an object, type, tag, tagger and a message.

tag [content size]\0
object e12f346
type commit tag name
committer John Smith john@smith.com
commit message

- The tag is zipped and then SHA1 to get the name: e12f3463g1533dffjf3343s2111fd453absdce18

References

- References are simple pointers to a particular commit
- A branch in Git is nothing more than a file in the .git/refs/heads/ directory that contains the SHA-1 of the most recent commit of that branch.
- They are simple pointers to a particular commit, something like a tag, but easily moveable.

Thanks !