

I found that wc is good to use for production. wc is able to accept all ascii characters and was able to handle files of 13.8 GB. One thing that did concern me was that I could only get code coverage of 5.6%. However, I believe that this was simply due to my lack being able figure out how to get the code coverage to work. Given more time, I would like to be able to get a “correct” code coverage result.

I used clang to compile and used the -fsanitize=undefined to look for undefined behavior. Running wc using an empty-file, a file full of only newlines, 13.8 GB of the repeated text of Ezekiel from the Bible, and the coreutils tarball, no undefined behavior showed up.

I feel that wc will be able to handle all ascii files at least to 13.8 GB.

Untested is anything using UTF-#.