Jonathon McDonald

Grep v2.6.3

About Grep

Grep is a command line utility that allows the user to search through sets of plain text data for a regular expression match. In it's simplest form a command like this could be run:

```
grep -i "string" FILE
```

The above command will search for string in the supplied file. It is not case sensitive. The next command is more complex:

```
find /etc -exec grep '[0-9][0-9]*[.][0-9][0-9]*[.][0-9][0-9]*[.][0-9][0-9]*' {} \;
```

The above command will find all IP addresses in the etc folder.

Stability of Grep

Our interests demand that Grep is secure and stable. My research has uncovered that Grep is indeed secure and stable. It would work well for our application.

Reported Bugs

Only one security vulnerability is reported to the public in regards to 2.6.3, which is "Multiple integer overflows in GNU Grep before 2.11 might allow context-dependent attackers to execute arbitrary code via vectors involving a long input line that triggers a heap-based buffer overflow."[0]. While this bug is certainly an issue we can control and sanitize input before calling grep from our application.

Fuzzed Files

Fuzzed files did not readily cause issues for Grep. Strings that existed in the file before being fuzzed could often not be found by Grep, but this is to be expected. The important takeaway is that the fuzzed files did not cause Grep to crash. If a user supplies a corrupted file the most important piece is that the program handles it gracefully, and Grep handles it as if it were not corrupted.

Code Review

Grep followed good coding standards and was fairly easy to read throughout. This means that if we were curious about a specific piece it would be relatively easy to find and learn about that piece.

I did find some pieces of ugly code. An example includes:

```
*--p = '0' + pos % 10;
```

Valgrind

I ran valgrind to test for memory leaks that could be posed by Grep but found that Grep was fairly robust with memory handling. My test was performed by searching my entire home directory for my name. This came up with roughly eight million bytes being still reachable by program end, but no leaks!

A caveat is that I was unable to properly compile Grep 2.6.3 on my home machine and so the test was performed with Grep 2.10.

Conclusion

In summary while I have found some less than desirable behavior in Grep I have also come to the conclusion that Grep is reasonably secure and well written. The source code does include some ugly bits, but for the most part is well written. Grep did not break when supplied fuzzed files or otherwise. Grep also only has one publicly released vulnerability. Since the vulnerability is dependent on user input and we can control or sanitize user input this should not pose a threat. Grep provides useful utility to quickly search for a regex match.

Sources

0. http://bit.ly/1jEoIBq