

Programming with AWK

AWK is a powerful simple interpreted programming language. It is usually used to operate on lines, just like **SED**, but it is a full-blown programming language. The **AWK** script has three parts. The first part is the code block which starts with the **BEGIN** keyword. This block is executed before the **AWK** command start processing the input, so it's mostly used for initialization. Then we have the main code block which is executed on each line from the input. You can actually have more than one, each prepended with a different expression. Expression surrounded with slashes, can be placed before this code block. If it is, then the operations inside of the block will be executed only on these lines which match the expression.

Lastly we have the **END** block which is executed after **AWK** finishes processing the lines from the input.

```
$ awk 'BEGIN{i=0; print "Counting sheep"} /sheep/ {i++;print i} END{print "Finished"}' file
```

This example will first initialize a variable *i* to zero and then print "Counting sheep". The middle code block will be executed on each line which contains the word *sheep*. It will first increment the value inside of the *i* variable by one and print it out. After everything is done, it will write out "Finished".

You can also write the programs inside of a file and invoke the scripts with the **f** option.