University of Colombo School of Computing

IS1104 - Application Laboratory

Linux Filters

Go through the tutorials first and do the follow up exercise. Upload one text file to the VLE including all the commands, you used for these exercises (no need of results). It should be named as <index>.txt (eg: 11000971.txt)

The files used in the exercise can be found in following paths of the provided VM.

File Name	Path
syslog	/var/log/syslog
dmesg	/var/log/dmesg
hosts	/etc/hosts
passwd	/etc/
hello	Attached
number	Attached
sed.html	Attached
comments.c	Attached

Grep Command Practice

Action	Symbol	Explanation	Example
Beginning of the line	^	Matches the expression at the start of a line.	`$ grep "^May  28" syslog`
End of the line	$	Matches the expression at the end of a line.	`$ grep "started.$" syslog`
Single Character	.	Matches any character except the end of the line character.	`$ grep ".ed" syslog` `$ grep -w "..:..:.."`

Zero or more occurrence	*	Matches zero or more occurrence of the **previous** character	`$ grep "kernel:.*." syslog`
			`$ grep "10*" passwd`
			`$ grep "hi.*hello" hello`

One or more occurrence	\+	Matches one or more occurrence of the previous character.	`$ grep "hi.\+hello" hello` `$ grep "10\+" passwd`
Escaping the special character	\	If you want to search for special characters (for example: * , dot) in the content you have to escape	`$ grep "127\.0\.0\.1"  hosts` `$ grep "127\.0\.\.1"  hosts`

		the special character in the regular expression	

Character class	[]	List of characters mentioned with in the square bracket which is used to match only one out of several characters.	`$ grep "[0123456789]\+" syslog`
			`$ grep "[0-9]*" syslog`
			`$ grep "[a-zA-Z]\?" syslog`

Exception in the character class	[^]	If you want to search for all the characters except those in the square bracket	`$ grep "[^0-9]*" syslog`
OR Operation	|	To specify either of two whole subexpressions occur in a position	`$ grep "^*\*\/|^\/\*\*|^\/\/" syslog`

M to N occurences	{m,n}	The **preceding** item is matched at least m times, but not more than n times.	3 to 5 occurrences :
			`$ grep "^[0-9]\{3,5\}$" number`
			exactly 5 occurrences :
			`$ grep "^[0-9]\{5\}$" number`
			5 or more occurrences:
			`$ grep "[0-9]\{5,\}" number`

Ref: http://evc-cit.info/cit052/review_regex.html

Exercise: Grep

1. Obtain the file *grepdata.txt* using *wget* from the URL http://evc-cit.info/cit052/grepdata.txt.
   I. Print all lines that begin with three digits followed by a blank. Your answer must use the \{ and \} repetition specifier.
   II. Print all lines that contain a date. Hint: this is a very simple pattern. It does not have to work for any year before 2000.
   III. Print all lines containing a vowel (a, e, i, o, or u) followed by a single character followed by the same vowel again. Thus, it will find “eve" or “adam" but not “vera". Hint: \(and \)

   Here onwards you have to use grep parameters like -v, -i ..etc.

   IV. Print all lines that do not begin with a capital S.
   V. Print all lines that contain CA in either uppercase or lowercase.
   VI. Print all lines that contain an email address (they have an @ in them), preceded by the line number.
   VII. Print all lines that do not contain the word Sep. (including the period).
   VIII. Print all lines that contain the word de as a whole word.