**103.7** Поиск информации при помощи регулярных выражений

Студент должен уметь осуществлять поиск необходимых файлов и информации внутри них.

**Изучаем**:

* создание простых регулярных выражений;
* инструменты, опирающиеся на эти выражения, для поиска файлов.

**Термины и утилиты:**

* grep
* egrep
* fgrep
* rgrep
* sed
* регулярные выражения.

Частой задачей в Linux является поиск необходимых файлов и сортировка информации внутри них. Для этой задачи служат регулярные выражения – специальный язык поиска и изменения информации, обладающий своим синтаксисом.

При отборе информации можно использовать стандартные групповые символы:

* ***^*** *- начало строки;*
* ***$***  *- конец строки;*
* ***.***  *- любой символ;*
* ***\<a*** *- слово, начинающееся с “a”;*
* ***\>a*** *- cлово, заканчивающееся на “a”;*
* ***a-z] -*** *диапазон от “a” до ”z”;*
* ***[^t] -*** *не буква ”t”;*
* ***a|z -*** *”a” или ”z”;*
* и т.д.;

Для простой сортировки набора строк используется команда **grep**, например:

***grep oo file.txt*** *(найти в файле две буквы «o» подряд);*

***ls | grep ile*** *(найти в названия содержимого каталога набор символов ‘ile’);*

***grep ple$ file.txt*** *(найти в файле file.txt строки, заканчивающиеся на ‘ple’);*

Для расширенной сортировки и поиска по сложным регулярным выражениям используется команда **egrep** (также можно использовать ***grep -E***), например:

***egrep ‘^(b|d)’ file.txt*** *(все что начинается с ”b” или ”d”);*

***egrep ‘^[a-k]’ file.txt*** *(все что начинается с ”a” по ”k”);*

Для быстрой сортировки и поиска по набору символов без регулярных выражений используется команда **fgrep** (также можно использовать ***grep -F***), например:

***grep c$ file.txt***  *(найти в file.txt последовательность символов “c$”);*

Для рекурсивной (включая вложенные каталоги и файлы) сортировки и поиска по набору символов используется команда **rgrep** (также можно использовать ***grep -R***), например:

***rgrep word***  *(найти «word» в содержимом текущего и всех вложенных каталогов);*

Для изменения текстового потока согласно заданным правилам используется редактор sed, например:

***sed –e ‘s/oo/aa/’ file.txt*** *(заменить в file.txt все “oo” на “aa” );*

***sed –re ‘s/^(B|b)/C’******file.txt > newfile.txt*** *(в строках, начинающихся с «B» или «b», заменить первую букву на «С» и сохранить результат в newfile.txt);*