**103.2** Управление текстовым потоком

Студент должен уметь управлять фильтровать текстовый поток

**Изучаем**:

* применение текстовых фильтров с использованием встроенных команд UNIX

**Термины и утилиты:**

* cat
* cut
* expand
* fmt
* head
* join
* less
* nl
* od
* paste
* pr
* sed
* sort
* split
* tail
* tr
* unexpand
* uniq
* wc

Множество файлов в Linux неудобны для восприятия без предварительного форматирования, кроме того процесс обработки текстовой информации часто требует ее форматирования, обрезания, сортировки и т.д. Linux предлагает широкий набор инструментов редактирования текстового вывода, при этом следует заметить что указанные инструменты не затрагивают исходный файл (или текстовый поток), а форматируют его, выдавая полученный результат в консоль или другой файл.

Простой вывод текста (без корректировки содержимого):

Для вывода содержимого файла или объединения нескольких файлов на выводе используется команда **cat**, например:

***сat 1.txt 2.txt > 3.txt*** *(передать содержимое двух файлов в третий);*

Команда **join** объединяет строки нескольких файлов по общему полю и отправляет их на стандартный вывод.

Команда **less** загружает файл постепенно и также отправляет его на вывод (удобно для больших файлов).

Команда **nl** нумерует строки.

Команда **pr** показывает, как содержимое файла будет выглядеть при выводе его на печать.

Команда **paste** вставляет строки из одного файла в другой и отправляет результат на стандартный вывод.

Команда **head** показывает первые строки файла, а **tail** – последние.

Например:

***head –n 2 /var/log/syslog*** *(вывести первые две строчки файла);*

***tail –f /var/log/syslog*** *(показывать последние строки файла, отображая содержимое в реальный момент времени);*

Команда **sort** может отсортировать содержимое файла по какому-нибудь признаку, например:

***sort 1.txt*** *(отсортировать строки по алфавиту);*

Команда **wc** позволяет посчитать количество символов и т.д. в текстовом потоке, например:

***wc 1.txt*** *(посчитать количество строк, символов, слов);*

Вывод форматированного текста (с корректировкой содержимого):

Команда **cut** используется для обрезания содержимого файлов на выводе, например:

***сut –c 2,3,4,5,10 1.txt*** *(показывать только символы на указанных позициях: 2,3,4,5,10);*

Команда **expand** превращает символы табуляции в несколько пробелов, а **unexpand** наоборот – пробелы в символы табуляции.

Команда **fmt** предназначена для форматирования текста различными способами, например:

***fmt 1.txt*** *(вывести содержимое файла в одну строку);*

***fmt –w 10 1.txt*** *(вывести содержимое файла в строки не больше 10 символов каждая);*

Команда **od** предназначена для конвертации текста в другой формат, например:

***od –c 1.txt*** *(отобразить в кодировке ASCI);*

Команда **tr** предназначена для замены и удаления символов в текстовом потоке, например:

***echo “HELLO” | tr –t L l*** *(заменить “L” на “l” в полученном на вводе тексте);*

Команда **split** может разбить файла по каким-либо критериям, например:

***split –l 2 1.txt*** *(разбить содержимое файла по две строчки);*

Команда **uniq** предназначена для поиска повторяющихся строк, например:

***uniq –u 1.txt*** *(показать только уникальные строки);*

**Sed** – отдельный мощный инструмент, понимающий групповые символы, шаблоны и т.д. Будет рассмотрен немного подробнее в дальнейших урока. Работает, например, так:

***sed –e ‘s/socks/people/’ 1.txt*** *(заменить в файле слово “socks” на “people”);*