

Практическая работа №3 «Программы для работы с текстом»

cat

Команда `cat` (полученная из слова concatenate) принимает несколько файлов в качестве входных данных и выводит объединенный файл. Чтобы объединить содержимое файлов `/etc/hosts` и `/etc/hostname` в файл результатов, выполните следующую команду:

```
cat /etc/hosts /etc/hostname > result
```

Команда `cat` также может быть использована для просмотра содержимого файла. Чтобы просмотреть файл результатов, созданный на последнем шаге, выполните следующую команду:

```
cat result
```

Чтобы просмотреть файл результатов вместе с номерами строк, добавленными к каждой строке, выполните следующую команду:

```
cat -n result
```

file

Команда `file` используется для отображения типа файла в зависимости от содержимого файла. Чтобы просмотреть типы файлов `/bin/ls` и `/etc/hosts`, выполните следующие команды:

```
file /bin/ls
```

```
file /etc/hosts
```

split

Команда `split` используется для разделения входного файла на два или более файлов. Чтобы разбить большой файл, такой как файл слов, на более мелкие части, выполните следующие команды:

```
cp /usr/share/dict/words .
```

```
split words output
```

```
ls output*
```

Чтобы к полученным разделенным файлам добавился суффикс с номерами, выполните следующие команды:

```
split words -d result
```

```
ls result*
```

Размер разделенных файлов будет зависеть от количества строк, используемых для разделения файла. Чтобы разделить каждую 1000-ю строку, выполните следующую команду:

```
split words result -l 1000
```

```
ls result*
```

nl

Чтобы отобразить файл `/etc/hosts` с номерами строк, выполните следующую команду:

```
nl /etc/hosts
```

Чтобы подсчитать количество строк в файле `/etc/hosts`, включая пустые строки, выполните следующую команду:

```
nl -ba /etc/hosts
```

head

Чтобы просмотреть начало файла `/usr/share/dict/words`, выполните следующую команду:

```
head /usr/share/dict/words
```

Если количество строк не указано, команда `head` отобразит первые 10 строк. Чтобы просмотреть первые 15 строк выходных данных команды `man ls`, выполните следующую команду:

```
man ls | head -15
```

Чтобы просмотреть первые 5 строк файла `/usr/share/dict/words`, выполните следующую команду:

```
head -n 5 /usr/share/dict/words
```

Чтобы исключить последние 3 строки файла `/etc/passwd` выполните следующую команду:

```
head -n -3 /etc/passwd
```

tail

Чтобы просмотреть последние 10 строк файла /usr/share/dict/words, выполните следующую команду:

```
tail /usr/share/dict/words
```

Чтобы просмотреть последние 5 строк файла /usr/share/dict/words, выполните следующую команду:

```
tail -5 /usr/share/dict/words
```

Чтобы просмотреть последние 5 строк выходных данных команды ls -ltr, выполните следующую команду:

```
ls -ltr /etc | tail -n 5
```

Чтобы просмотреть содержимое файла /etc/passwd, начиная с 3-й строки и до конца файла, выполните следующую команду:

```
tail -n +3 /etc/passwd
```

paste

Выполните следующие команды, чтобы создать два файла, а затем объединить их:

```
cp /usr/share/dict/words words
```

```
head -5 words > head
```

```
tail -5 words > tail
```

```
paste head tail > total
```

```
cat total
```

Чтобы снова объединить два файла, используя «:» в качестве разделителя вместо разделителя табуляции по умолчанию, выполните следующую команду:

```
paste -d : head tail > total
```

```
cat total
```

Команда вставки используется для объединения строк из нескольких файлов; это особенно полезно для файлов, которые содержат данные в виде столбца.

expand

Параметр -A команды cat используется для просмотра пробельных символов, таких как символы новой строки и символы табуляции. Сравнение вывода команды cat -A и файла расширения /etc/hosts | Команда cat -A дает представление о том, как функция расширения преобразует вкладки в пробелы:

```
cat -A /etc/hosts
```

```
expand /etc/hosts | cat -A
```

Команда **unexpand** преобразует пробелы в символы табуляции. Чтобы преобразовать пробелы в файле /etc/hosts в закладки, выполните следующую команду:

```
unexpand -t8 /etc/hosts | cat -A
```

cut

Чтобы извлечь 1-е, 3-е и 4-е поля из файла /etc/passwd, выполните следующую команду (команда head используется для ограничения вывода):

```
cut -d: -f1,3,4 /etc/passwd | head -n 4
```

Команда **cut** используется для извлечения полей из текстового файла. **Пробел и табуляция** являются разделителями по умолчанию и могут быть изменены с помощью опции **-d**.

Диапазон полей также может быть предоставлен команде cut. Чтобы извлечь поля 1 - 4 из файла /etc/passwd, выполните следующую команду (команда head используется для ограничения вывода):

```
cut -d: -f1-4 /etc/passwd | head -n 4
```

Определенное поле или набор полей также можно извлечь с помощью команды **cut** из данных, содержащих столбцы фиксированной ширины. Чтобы извлечь содержимое из столбца 31–44 из выходных данных команды ls -l, выполните следующую команду (команда head используется для ограничения выходных данных):

```
ls -l /etc | head | cut -c31-43
```

sort

Чтобы отсортировать файл `/etc/passwd`, используя первое поле (т. е. Имя пользователя) в качестве ключа и используя «:» в качестве разделителя, выполните следующую команду (команда `head` используется для ограничения вывода четырьмя строками):

```
sort -t: -k1 /etc/passwd | head -n 4
```

Команда **sort** полезна для работы с данными, организованными в столбцы. Он используется для отображения файла, отсортированного по определенному полю данных.

Чтобы изменить порядок сортировки в приведенном выше примере с возрастания на убывание, выполните следующую команду:

```
sort -t: -k1r /etc/passwd | head -n 4
```

Для определенных полей требуемый порядок сортировки может быть числовым, а не алфавитным. Чтобы численно обработать поле сортировки для второго поля файла `/etc/passwd`, выполните следующую команду:

```
sort -t: -k3n /etc/passwd | head -n 4
```

Используйте опцию `-u` для удаления повторяющихся записей. Выполните последовательно нижеприведенные команды:

```
cut -f7 -d: /etc/passwd | head > output
```

```
more output
```

```
sort -u output
```

Чтобы удалить дубликаты строк из выходного файла и отобразить количество дубликатов, выполните следующую команду:

```
sort output | uniq -c
```

Опция `-u` для команды сортировки переупорядочивает строки в файле, а затем удаляет дубликаты. Команда **uniq** также предоставляет несколько функций удаления дубликатов. Он удаляет дубликаты, которые уже находятся в последовательных строках, т.е. он не сортирует данные.

Команда **uniq** предоставляет еще одну дополнительную опцию, не поддерживаемую командой `sort`; он может отображать количество дубликатов, которые были подсчитаны.

tr

Чтобы использовать файл `/etc/hosts` с помощью команды `tr`, выполните следующую команду:

```
cat /etc/hosts | tr 'a-z' 'A-Z'
```

Чтобы заменить пять букв (`h`, `i`, `p`, `o` и `n`) на пять символов (`!`, `@`, `#`, `$` и `%`) в файле `/etc/hosts`, выполните следующую команду:

```
cat /etc/hosts | tr 'hipon' '!@#$$%'
```

sed

Чтобы заменить все вхождения слова «хост» в файле `/etc/hosts` глобально, выполните следующие команды:

```
cat /etc/hosts
```

```
sed 's/host/NAME/g' /etc/hosts
```

Чтобы вставить слово «Report» в файл `/etc/hosts`, предшествующий строке, содержащей «127.0.0.1», выполните следующую команду:

```
sed '/127.0.0.1/i\Report' /etc/hosts
```

Чтобы добавить слово «End of report» в файл `/etc/hosts` после строки, содержащей «allrouters», выполните следующую команду:

```
sed '/allrouters/a\End of report' /etc/hosts
```

Чтобы удалить строки, содержащие «localhost» в файле `/etc/hosts`, выполните следующую команду:

```
sed '/localhost/d' /etc/hosts
```