



СИСТЕМНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Работа с файлами

Мусаев Андрей Александрович
amusayev1990@gmail.com



Загрузка файлов

GNU wget

GNU **Wget** это открыто распространяемая утилита для загрузки файлов из интернет. Она поддерживает протоколы HTTP, HTTPS, и FTP, загрузку с прокси-серверов по протоколу HTTP.

wget [параметры]... [URL]...

wget --spider -I url.txt

--spider – проверка на доступность

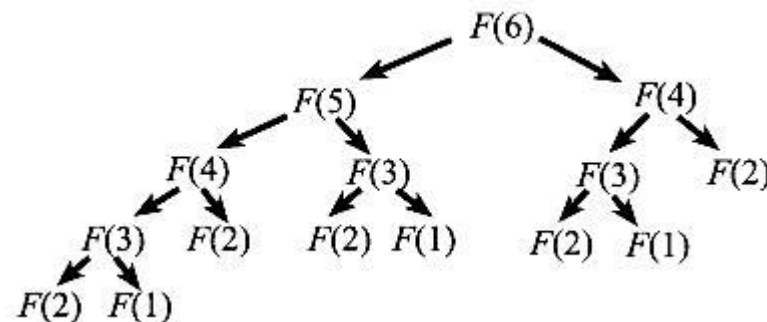
-i – работа со ссылками из файлов

Загрузка файлов

GNU wget

`wget -r --level=1 url.html`

- r – рекурсивная загрузка
- l – глубина рекурсии. По умолчанию глубина рекурсии равна 5.



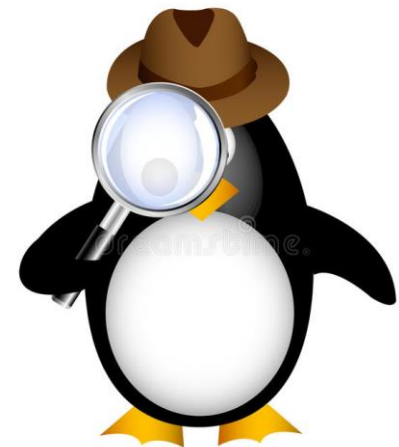


Поиск файлов

Find — утилита поиска файлов по имени и другим свойствам, используемая в UNIX-подобных операционных системах. Может производить поиск в одном или нескольких каталогах с использованием критериев, заданных пользователем.

```
find -name filename
```

- name** — поиск файла по имени
- type** — поиск файла по типу
- size** — поиск файла по размеру





Права доступа

```
me@andrmus: ~  
me@andrmus:~$ ls -l  
total 64  
drwxrwxr-x 2 me me 4096 апр  6  2016 1  
-rwxrwxrwx 1 me me  159 май 31  2016 123~  
drwxr-xr-x 2 me me 4096 апр 21  2016 Audio  
-rw-rw-r-- 1 me me 5768 мар  8  2017 autZIP.zip  
drwxr-xr-x 2 me me 4096 сен 13 01:28 Desktop  
drwxr-xr-x 3 me me 4096 мар 13  2017 Documents  
drwxr-xr-x 4 me me 4096 июл 17 04:33 Downloads  
drwxr-xr-x 2 me me 4096 ноя 18  2014 Music  
drwxr-xr-x 2 me me 4096 сен 13 01:51 Pictures  
-rwxrwxrwx 1 me me  139 апр 24  2016 pr~  
-rwxrwxrwx 1 me me  135 апр 24  2016 pr.sh~  
drwxr-xr-x 2 me me 4096 ноя 18  2014 Public  
drwx----- 2 me me 4096 июн  7  2016 rec  
drwxr-xr-x 2 me me 4096 ноя 18  2014 Templates  
-rw-rw-r-- 1 me me   0 май 31  2016 Untitled Document~  
drwxr-xr-x 4 me me 4096 окт  2  2016 Videos
```

- r – право на чтение
- w – право на запись
- x – право на исполнение

- 1) Владелец
- 2) Группа владельца
- 3) Остальные



Права доступа

Для распределения прав доступа в Linux существует множество команд. Основные из них – это **chmod**, **chown** и **chgrp**.

Команда **chmod** (Change MODe – сменить режим) – изменяет права доступа к файлу. Для использования этой команды также необходимо иметь права владельца файла или права root.

```
chmod |u| |g| |o| |a| |r| |w| |x| |X| |filename|
      |g| |+| |-| |=| |u| |g| |o|
```

Права доступа

Второй способ использования команды `chmod`:

`chmod XXX filename`



X – число от 0 до 7

двоичная	восьмеричная	символьная	права на файл	права на директорию
000	0	---	нет	нет
001	1	--x	выполнение	чтение файлов и их свойств
010	2	-w-	запись	нет
011	3	-wx	запись и выполнение	всё, кроме чтения списка файлов
100	4	r--	чтение	чтение имён файлов
101	5	r-x	чтение и выполнение	доступ на чтение
110	6	rw-	чтение и запись	чтение имён файлов
111	7	rwx	все права	все права



Права доступа

Команда **chown** (CHange OWNer – сменить владельца) – позволяет сменить владельца файла.

```
chown username:groupname filename
```

Команда **chgrp** используется для изменения владельца-группы файла.

```
chgrp groupname filename
```




Архивирование

Архиватор — это программа, осуществляющая упаковку одного и более файлов в **архив** или серию архивов для удобства переноса или хранения, а также распаковку архивов.

tar (GNU tar – GNU tape arhiver) – программа для создания архивов.

tar options archivename files



- c** – создать архив
- r** – добавить файлы в архив
- A** – добавить содержимое tar-файлов в архив
- delete** – удалить файлы из архива (невозможно использование на архивных лентах)
- t** – вывести список файлов в архиве
- x** – извлечь файлы из архива
- f** – информация будет извлекаться из файла
- v** – расширенный вывод информации о выполняемых действиях



Архивирование

Gzip – утилита компрессии (сжатия) файлов.

`gzip options filename`

только имя файла – сжимает этот файл.

-d – декомпрессия файла

-t – проверяет целостность архива

-v – расширенный вывод информации о выполняемых действиях

Характеристики
архиваторов:
По степени сжатия.
По скорости сжатия.



Gzip – быстрее
сжимает

Bzip2 – сильнее
сжимаем



Чтение файлов

cat – команда используется для вывода текстовой информации из файла на экран.

`cat filename`

- b** (`--number-nonblank`) – пронумеровать все непустые строки;
- n** (`--number`) – пронумеровать все строки;
- s** (`--squeeze-blank`) – отобразить несколько подряд идущих пустых строк в виде одной пустой строки;
- T** (`--show-tabs`) – показать символы табуляции, отобразив их как “^”;
- E** (`--show-ends`) – показать символы конца строки как “\$”.

tac filename – эта команда используется для вывода на экран информации из файла в обратном порядке.



Чтение файлов

more filename – эта команда позволяет просматривать длинные файлы по частям.

Команда **more** использует для прокрутки две клавиши – пробел (показать следующий экран) и Enter (показать следующую строку). Но у **more** есть один недостаток – она способна прокручивать текст только вперед. То есть если вы уже смотрите второй экран, то к первому никак вернуться будет нельзя. Эту проблему с легкостью решает команда **less filename**.



Чтение файлов

more filename – эта команда позволяет просматривать длинные файлы по частям.

Команда **more** использует для прокрутки две клавиши – пробел (показать следующий экран) и **Enter** (показать следующую строку). Но у **more** есть один недостаток – она способна прокручивать текст только вперед. То есть если вы уже смотрите второй экран, то к первому никак вернуться будет нельзя. Эту проблему с легкостью решает команда **less filename**.

Enter (стрелка вниз) – перейти на одну строку вниз; **Пробел** (**Page Down**) – перейти на страницу вниз по тексту; **Page Up** – перейти на страницу вверх по тексту; **стрелка вверх** – перейти на строку вверх по тексту; **/** - поиск; **Home** – перейти к началу текста; **End** – перейти к концу текста; **q** – выйти из программы **less**.

Чтение файлов

Если нам не нужно отображать весь файл а только необходимо убедиться что это именно то, что мы ищем, то нам помогут команды **head** и **tail**.



`head options filename(s)`

- по умолчанию выводит 10 первых строк файла
- **n** lines – вывести не 10 а lines строк от начала файл;
- **c** bytes – вывести bytes байт с начала файла;
- **q** – не печатать заголовки файлов перед выводом (при выводе нескольких файлов сразу).



Чтение файлов

`tail options filename(s)`

- по умолчанию выводит 10 последних строк файла
- n** lines – вывести не 10 а lines строк;
- c** bytes – вывести bytes байт;
- f** – войти в постоянный цикл по считыванию конца файла. Таким образом при поступлении в файл новой информации пользователь может вести мониторинг ее в реальном времени. Выход из этого режима осуществляется комбинацией клавиш **<Ctrl>+<C>**.



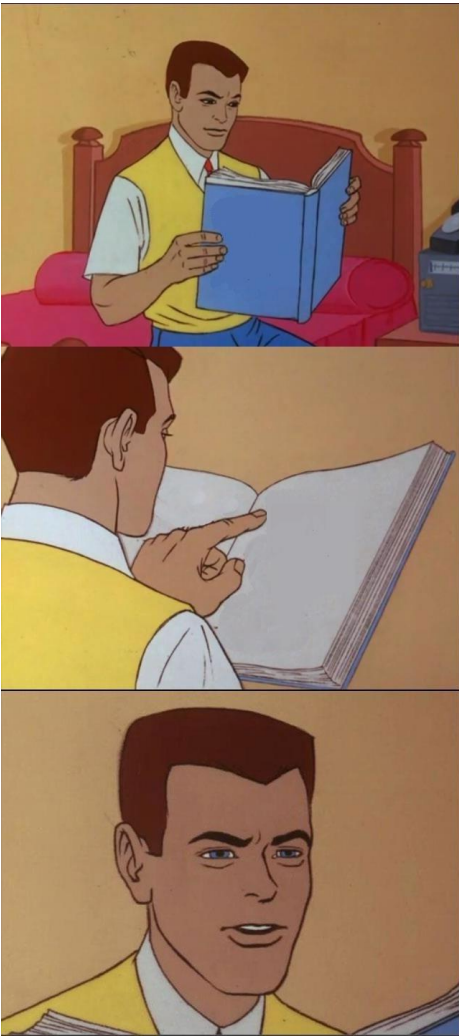


Поиск по файлу

grep — утилита командной строки, которая находит на вводе строки, отвечающие заданному регулярному выражению, и выводит их, если вывод не отменён специальным ключом. Название представляет собой акроним английской фразы «*search **g**lobally for lines matching the **r**egular **e**xpression, and **p**rint them*» — «искать везде строки, соответствующие регулярному выражению, и выводить их».

```
grep pattern filename1
```

- i - выполняется поиск без учета регистра символов.
- w - поиск совпадений целого слова.
- c - выводит количество совпадений без вывода самих совпадений.
- v - выводит только строки, не содержащие совпадений.





Потоки ввода-вывода

Стандартные потоки ввода-вывода в системах типа UNIX (и некоторых других) — потоки процесса, имеющие номер (дескриптор), зарезервированный для выполнения некоторых «стандартных» функций.

Поток номер 0 (stdin) зарезервирован для чтения команд пользователя или входных данных.

Поток номер 1 (stdout) зарезервирован для вывода данных, как правило (хотя и не обязательно) текстовых.

Поток номер 2 (stderr) зарезервирован для вывода диагностических и отладочных сообщений в текстовом виде.

```
cat filename1 > filename2  
cat filename1 >> filename2
```

