

Внешний курс. Раздел - 2

Выполнение внешнего курса. Основы Linux. Раздел 2

Титков Ярослав Максимович

Содержание

1	Цель работы	5
2	Задание	6
3	Выполнение лабораторной работы	7
4	Выводы	17

Список иллюстраций

3.1	Удалённые серверы	7
3.2	Передача открытого ключа	7
3.3	Рекурсивная передача файлов	8
3.4	Обновление репозитория	8
3.5	Работа с Filezilla	8
3.6	Настройка вывода	9
3.7	Справка по программам	9
3.8	Выбор форматов	10
3.9	Выравнивание последовательностей	10
3.10	Завершение и приостановка	11
3.11	Идентификаторы процессов	11
3.12	Сигналы завершения	12
3.13	SIGTERM	12
3.14	CPU приостановленного процесса	12
3.15	Потребление памяти	13
3.16	Завершение потока	13
3.17	Многопоточность	14
3.18	Результат работы bowtie2	14
3.19	fg в другой вкладке	15
3.20	Завершение сессии tmux	15
3.21	Отключение от tmux	15
3.22	Закрытие вкладки	16
3.23	Переименование вкладки	16

Список таблиц

1 Цель работы

Изучить расширенные аспекты работы в Linux, включая работу с удалёнными серверами, передачу файлов, управление процессами и использование инструментов командной строки.

2 Задание

Выполнить задания, представленные во втором разделе курса.

3 Выполнение лабораторной работы

Вы прошли больше 80% курса, оставьте отзыв [Оставить отзыв](#) [Нет, спасибо](#)

Для каких задач можно использовать удаленный сервер?

Выберите все подходящие ответы из списка

✓ Хорошие новости, верно!

Верно решили 41 258 учащихся
Из всех попыток 54% верных

- ✓ Хранение общедоступных данных (например, доступных для всех пользователей интернета)
- ✓ Хранение больших объемов данных
- ✓ Выполнение сложных (затратных по памяти и времени) вычислений
- ✓ Хранение конфиденциальных данных (т.е. доступ к ним должны иметь только ограниченный круг лиц)

[Следующий шаг](#) [Решить снова](#)

[Ваши решения](#) Вы получили: 1 балл

Рис. 3.1: Удалённые серверы

Выбраны все варианты, потому что удаленный сервер подходит для хранения данных (общедоступных, больших, конфиденциальных) и выполнения сложных вычислений.

2.1 Знакомство с сервером 6 из 6 шагов пройдено 2 из 2 баллов получено

Вы прошли больше 80% курса, оставьте отзыв [Оставить отзыв](#) [Нет, спасибо](#)

Предположим программа ssh-keygen создала вам два ключа: id_rsa и id_rsa.pub. Какой из этих ключей можно без опаски пересылать по интернету?

Выберите один вариант из списка

✓ Хорошие новости, верно!

Верно решили 40 966 учащихся
Из всех попыток 75% верных

- ☐ id_rsa
- ☒ id_rsa.pub
- ☐ Оба
- ☐ Ни один нельзя

[Следующий шаг](#) [Решить снова](#)

[Ваши решения](#) Вы получили: 1 балл

Рис. 3.2: Передача открытого ключа

Выбран id_rsa.pub, так как это открытый ключ, который безопасно передавать.

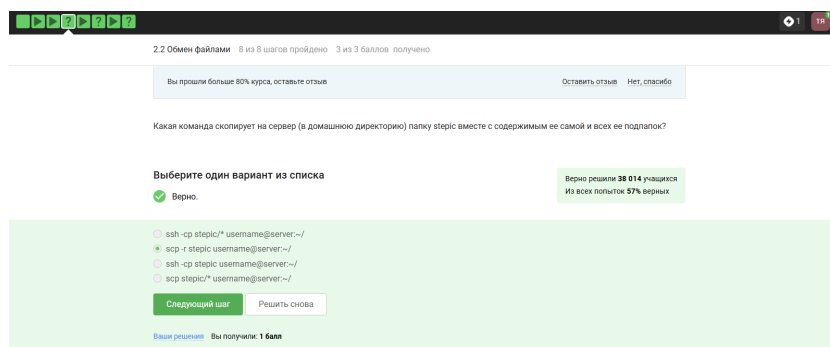


Рис. 3.3: Рекурсивная передача файлов

Выбран `scp -r stepic username@server:/`, так как эта команда рекурсивно копирует папку с содержимым.

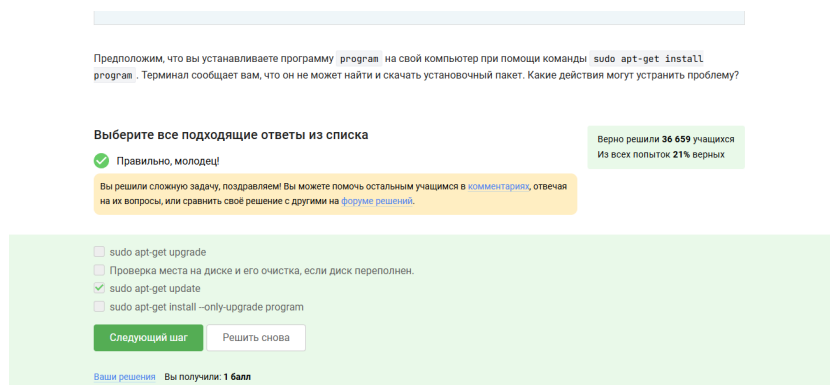


Рис. 3.4: Обновление репозиториев

Выбран `sudo apt-get update`, так как это обновляет список пакетов, что может решить проблему с установкой.

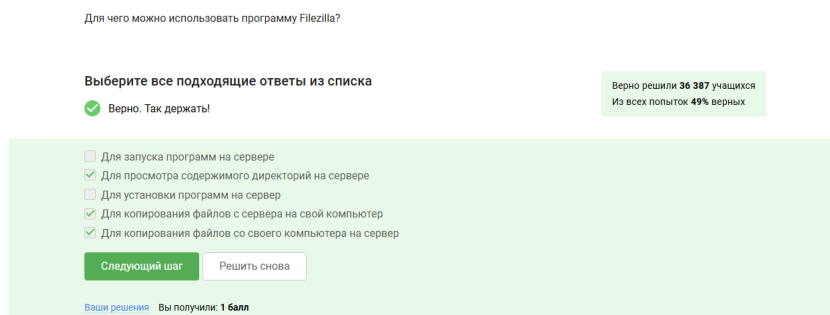


Рис. 3.5: Работа с Filezilla

Выбраны варианты для просмотра директорий и копирования файлов (на сервер и с сервера), так как Filezilla предназначен для этих задач.

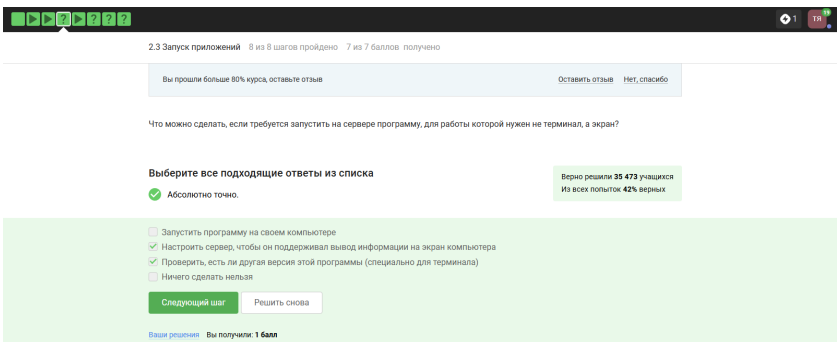


Рис. 3.6: Настройка вывода

Выбраны варианты настройки сервера для вывода на экран и проверки терминальной версии программы, так как это решает проблему.

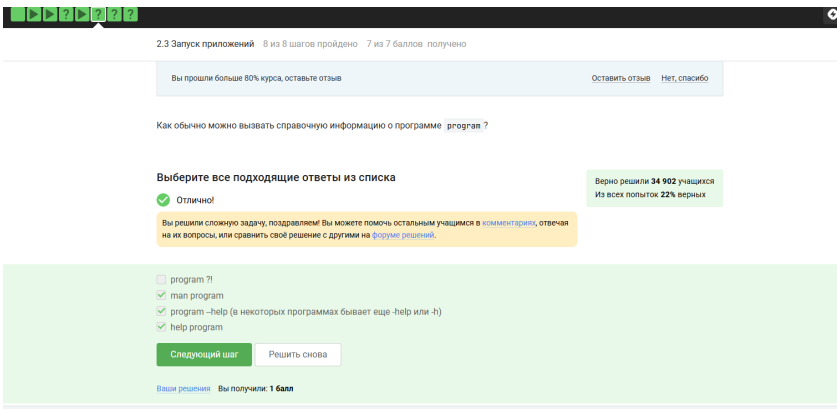


Рис. 3.7: Справка по программам

Выбраны man program и program –help, так как это стандартные способы получения справки.

он может принимать на вход

Если вы хотите попробовать запустить FastQC на каких-то реальных данных, то можете попробовать на [этом файле](#).

Подсказка: если программы FastQC еще нет на вашем компьютере, то её можно установить командой `sudo apt-get install fastqc` (или в некоторых версиях еще: `bio-linux-fastqc`) или найди её в Software Center по запросу `fastqc`. К сожалению, на некоторых дистрибутивах Linux у вас может не получиться установить FastQC описанным способом (по ключевым словам `fastqc` и `bio-linux-fastqc` ничего не будет найдено). В этом случае установка будет сложнее, описываем её подробнее.

1. Откройте терминал, попробуйте выполнить команду `java`. Если получите сообщение, что такая команда не найдена, то переходите к шагу 2, иначе сразу к шагу 3.
2. Вам нужно установить `java`, например, на Ubuntu это можно сделать с помощью `sudo apt-get install default-jre`.
3. Скачайте и распакуйте [zip-архив](#) с FastQC (можно это сделать прямо в терминале с использованием `wget` и `unzip`).
4. Файл запуска FastQC называется `fastqc` и лежит той директории, куда произошла распаковка архива, например, `/home/bi/FastQC/fastqc`. Перед первым запуском его нужно сделать исполняемым (при помощи `chmod +x`).
5. Запустить файл `fastqc` можно как и любую другую программу в терминале (например, через `./fastqc` из директории, где он лежит или из любой другой директории задав абсолютный путь до `fastqc`, см. [соответствующее занятие](#)). Если запустить его без параметров, то будет открыта графическая версия программы, а если указать опции или аргументы, например, `-help`, то будет запущена версия для терминала.

Выберите все подходящие ответы из списка

Верно решили **32 124** учащихся
Из всех попыток **29%** верных

✓ Хорошая работа.

Вы решили сложную задачу, поздравляем! Вы можете помочь остальным учащимся в комментариях, отвечая на их вопросы, или сравнить своё решение с другими на [форуме решений](#).

☒ fastq
☐ fasta
☒ bam_mapped, sam_mapped
☒ bam, sam

Следующий шаг Решить снова

Рис. 3.8: Выбор форматов

Выбраны варианты `bam`, `sam` и `fastq`, так как они соответствуют формату данных для FastQC.

Clustal – это одна из самых широко используемых компьютерных программ для множественного выравнивания нуклеотидных и аминокислотных последовательностей (multiple sequence alignment). У неё есть графическая версия ClustalX и версия для запуска в терминале ClustalW. Вы можете потренироваться запускать его с использованием файла [test.fasta](#).

Посмотрите справку по программе (имеется в виду версия для терминала) и **впишите** в поле ниже **команду**, которая запускает в терминале Clustal на файле `test.fasta` и выполняет множественное выравнивание (multiple alignment). Никакие лишние опции указывать не нужно (**только необходимые** для выполнения этого задания)!

Примечание: справку по опциям можно получить при помощи `man` или, если он у вас не работает, то в разделе **"Help for command line parameters"** файла `clustalw_help.txt`, который идет в поставке программы.

Примечание 2: программа Clustal запускает необходимый алгоритм выравнивания по умолчанию (т.е. если ему не указать каких-либо других опций), однако мы просим вас найти и **указать** в команде запуска **опцию**, которая явно говорит Clustal запустить именно множественное выравнивание. После этого вы можете сравнить вывод Clustal при запуске с этой опцией и без нее – результат должен быть одинаков.

Подсказка: если у вас не установлена программа Clustal, то её можно установить командой `sudo apt-get install clustalw` (или `clustalx`) или найди её в Software Center по запросу `clustalw` (`clustalx`). Обратите внимание, что на некоторых дистрибутивах доступна только вторая версия программы (например, `clustalw2`), в этом случае можете использовать и её – все необходимые в задании опции будут точно такими же.

Напишите текст

Верно. Так держат!

Верно решили **28 700** учащихся
Из всех попыток **41%** верных

`clustalw test.fasta -align`

Рис. 3.9: Выравнивание последовательностей

Введена команда `clustalw -align test.fasta`, так как она явно указывает на множественное выравнивание.

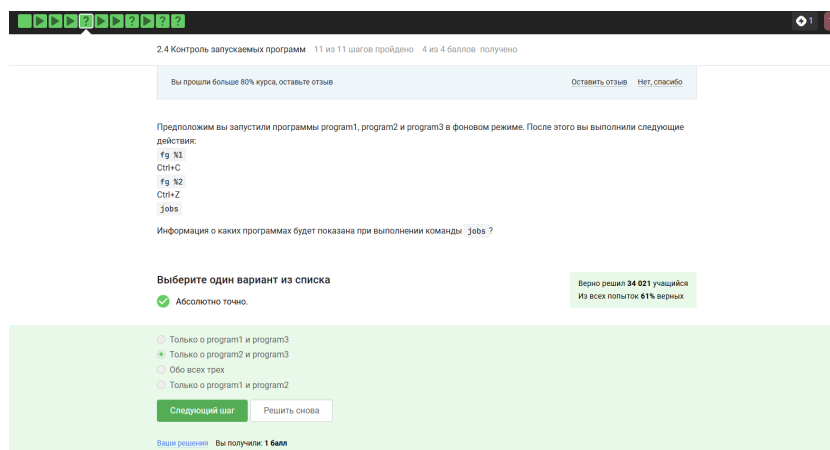


Рис. 3.10: Завершение и приостановка

Выбран вариант “Только о `program1` и `program3`”, так как `Ctrl+C` завершает программу, а `Ctrl+Z` приостанавливает, оставляя их в списке `jobs`.

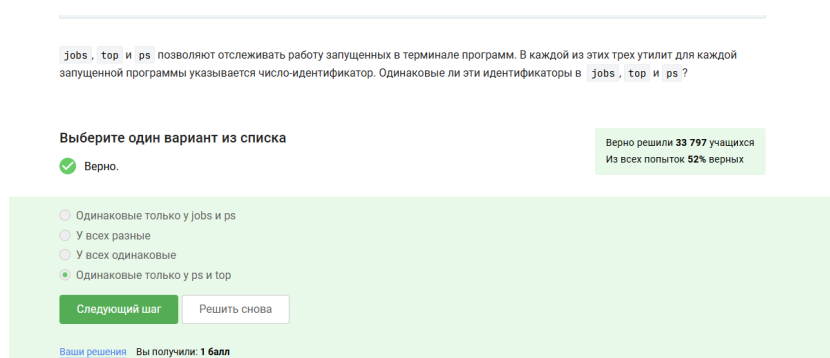
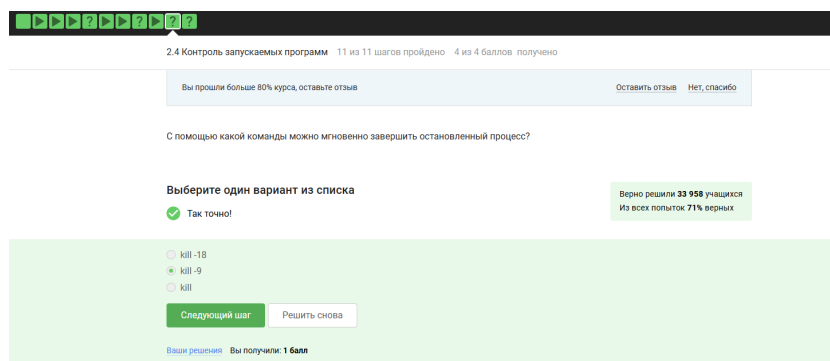


Рис. 3.11: Идентификаторы процессов

Выбрано “У всех разные”, так как `jobs`, `top` и `ps` используют разные идентификаторы: `jobs` — номер задания в оболочке, `ps` — PID, `top` — динамический список.



Выбрано kill -9, так как это сигнал SIGKILL, мгновенно завершающий процесс, включая остановленные.

Рис. 3.13: SIGTERM

Выбрано “Процесс будет завершен”, так как kill без опций отправляет SIGTERM, который завершает процесс.

Рис. 3.14: CPU приостановленного процесса

Выбрано “0% CPU”, так как остановленный процесс (Ctrl+Z) не использует процессорное время.

Сколько памяти занимает остановленное (по Ctrl+Z) многопоточное приложение?

Подсказка: если вы не знаете как ответить на этот вопрос, то можете попробовать запустить многопоточное приложение на своем компьютере и посмотреть на результат с помощью команды `top`. Если вы не знаете примеров таких приложений, то рекомендуем вам ненадолго отложить этот шаг и досмотреть занятие до конца. В следующих видеофрагментах и заданиях будет показан пример многопоточного приложения (программы bowtie2). Тестовые данные для запуска этой программы можно найти в последнем задании этого урока.

Подсказка 2: подробнее почитать о значении всей информации, которую выводит `top` на экран, можно по ссылке <http://rus-linux.net/MyLDP/console/komanda-top-v-linux.html>

Выберите один вариант из списка

Верно решили 32 287 учащихся
Из всех попыток 56% верных

✓ Всё получилось!

☒ Столько, сколько оно потребляло в момент остановки

☐ 64 KB

☐ Нисколько

☐ По 64 KB на каждый поток

Следующий шаг Решить снова

Ваши решения Вы получили: 1 балл

Рис. 3.15: Потребление памяти

Выбрано “Столько, сколько оно потребляло в момент остановки”, так как память не освобождается при остановке.

Как принудительно завершить один из потоков запущенного многопоточного приложения?

Подсказка: если вы не знаете как ответить на этот вопрос, то можете попробовать запустить многопоточное приложение на своем компьютере и посмотреть на результат. Если вы не знаете примеров таких приложений, то рекомендуем вам ненадолго отложить этот шаг и досмотреть занятие до конца. В следующих видеофрагментах и заданиях будет показан пример многопоточного приложения (программы bowtie2). Тестовые данные для запуска этой программы можно найти в последнем задании этого урока.

Выберите все подходящие ответы из списка

Верно решили 31 649 учащихся
Из всех попыток 32% верных

✓ Здорово, всё верно.

Вы решили сложную задачу, поздравляем! Вы можете помочь остальным учащимся в комментариях, отвечая на их вопросы, или сравнить своё решение с другими на форуме решений.

☐ Командой `threadkill`

☐ Сочетанием клавиш Ctrl+C

☐ Командой `kill -thread`

☒ Никкак

Следующий шаг Решить снова

Ваши решения Вы получили: 1 балл

Рис. 3.16: Завершение потока

Выбрано сочетание клавиш Ctrl+C, так как это стандартный способ принудительного завершения процессов в терминале.

Вы прошли больше 80% курса, оставьте отзыв

Оставить отзывНет, спасибо

Для выполнения этого задания вам потребуется программа bowtie2.

Надеемся, что вы разобрались, что запуск bowtie2 состоит из двух шагов – сначала запускаем подпрограмму bowtie2-build, а затем подпрограмму bowtie2. Изучите справочную информацию об этих подпрограммах (можно вызвать при помощи `-help`) и ответьте на вопрос – какой(ие) из этих шагов можно выполнить в несколько потоков?

Выберите один вариант из списка

☒ Верно. Так держать!

Верно решили **31 625** учащихся
Из всех попыток **58%** верных

☐ Никакой

☐ Оба

☒ Только bowtie2

☐ Только bowtie2-build

Следующий шаг

Решить снова

Ваши решения

Вы получили: **1 балл**

Рис. 3.17: Многопоточность

Выбрано “Только bowtie2”, так как bowtie2 поддерживает многопоточность, а bowtie2-build — нет.

перенаправляет вывод stdout в файл на одном этапе, чтобы он не засорил экран в терминале.

Попробуйте теперь запустить второй этап (запуск подпрограммы bowtie2) в несколько потоков. Рекомендуем выставить число потоков равное количеству ядер на вашем компьютере (команда `lscpu`). Сравните скорость выполнения в таком режиме с работой в один поток. Также рекомендуем убедиться, что результаты запусков (т.е. вывод в `stdout`) полностью совпали в обоих режимах!

Примечание: если у вас не очень сильный компьютер, то работа bowtie2 на предложенных данных может занять достаточно продолжительное время. Если вы не хотите ждать, то можете использовать альтернативные (сильно уменьшенные) версии [референсного генома](#) (reference) и [ридов](#) (reads). На этих данных у вас не получится увидеть разницу в скорости при запуске в один или в несколько потоков, но вы сможете выполнить все остальные пункты задания и получить за него полный балл.

Напишите текст

Верно.

Верно решили **23 694** учащихся

Из всех попыток **65%** верных

```
echo "306174 reads; of these:
306174 (100.00%) were unpaired; of these:
11 (0.00%) aligned 0 times
305580 (99.81%) aligned exactly 1 time
583 (0.19%) aligned >1 times
100.00% overall alignment rate" > bowtie.log
```

Следующий шаг

Решить снова

Рис. 3.18: Результат работы bowtie2

Выполнено задание с выводом статистики в bowtie.log, подтверждающее корректность работы в многопоточном режиме.

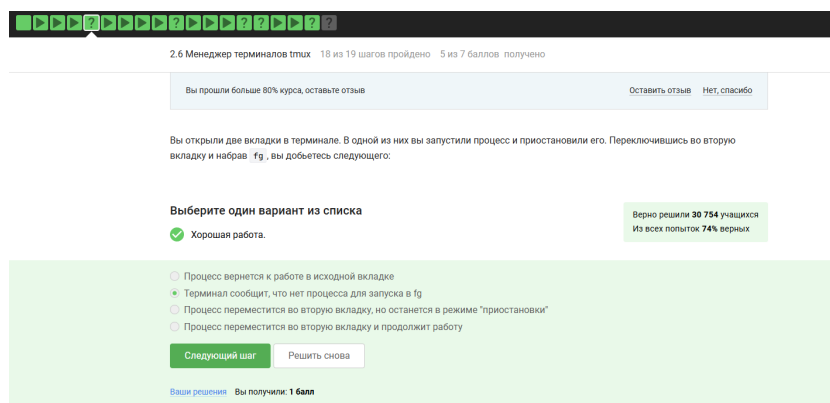


Рис. 3.19: fg в другой вкладке

Выбран вариант “Терминал сообщит, что нет процесса для запуска в fg”, потому что команда fg работает только в текущей вкладке.

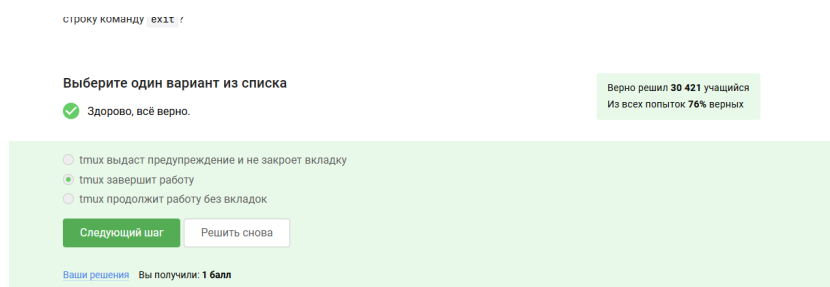


Рис. 3.20: Завершение сессии tmux

Выбран вариант “tmux завершит работу”, так как закрытие последней вкладки приводит к завершению сессии tmux.

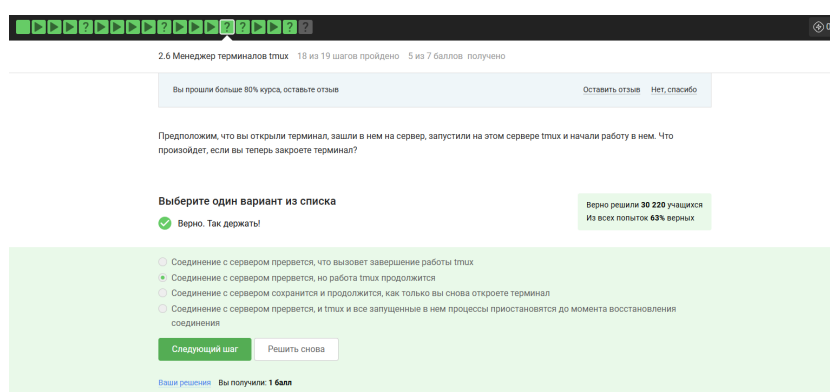


Рис. 3.21: Отключение от tmux

Выбрано “Соединение с сервером прервется, но работа tmux продолжится”, так как tmux/trmux работает независимо от терминала.

Вы прошли больше 80% курса, оставьте отзыв

Оставить отзыв Нет, спасибо

Что произойдет, если запустить процесс в фоновом режиме в одной из вкладок tmux, а затем принудительно закрыть эту вкладку (Ctrl+B, X)?

Выберите один вариант из списка

Верно. Так держать!

Верно решили 30 092 учащихся
Из всех попыток 61% верных

- ☒ Вкладка закроется, а вместе с ней пропадет и запущенный в ней процесс
- ☐ Вкладка закроется и процесс перейдет во вкладку, ближайшую из открытых (если есть, то слева, иначе справа)
- ☐ tmux выдаст предупреждение и не даст закрыть вкладку

Следующий шаг Решить снова

Ваши решения Вы получили: 1 балл

Рис. 3.22: Заккрытие вкладки

Выбран вариант “Вкладка закроется, а вместе с ней пропадет и запущенный в ней процесс”, так как принудительное закрытие вкладки в tmux убивает все процессы в ней.

Задание на самостоятельное изучение tmux.

Изучите справку по tmux (например, `man tmux`) и выберите из предложенных ниже tmux-команд ту, которая отвечает за **переименование** текущей вкладки.

Выберите один вариант из списка

Хорошая работа.

Верно решили 29 445 учащихся
Из всех попыток 54% верных

- ☐ Ctrl+B и . (точка)
- ☐ Ctrl+B и 0
- ☐ Ctrl+B и r
- ☐ Ctrl+B и t
- ☒ Ctrl+B и , (запятая)

Следующий шаг Решить снова

Ваши решения Вы получили: 1 балл

Рис. 3.23: Переименование вкладки

Выбран вариант “Ctrl+B и , (запятая)”, потому что эта комбинация в tmux отвечает за переименование текущей вкладки.

4 Выводы

Я прошёл и изучил второй раздел курса по Linux на платформе Stepik.