Linux and more. Intro

Artem Trunov for Ozon Masters

Об этом курсе

- Подготовка к курсу Big Data
- Что такое Linux, как зайти на сервер и работать с командной строкой
  - ssh
- Утилиты командной строки для манипуляции данными
  - Включая элементы программирования на bash
- Навыки работы с git
  - От pull-commit-push к git flow.
- Сам себе Devops
  - Виртуализация, virtualenv, Docker
  - Continuous Integration, Delivery
  - Автоматизация конфигурации и управления облаком

Обо мне

- Работаю с Линуксом с 1997 года
- Выпускник МГУ (Физфак),
- Учился в аспирантуре в США
- Работал на Стенфордском Ускорителе, в Европе - LCH
- Работал в IT банка.
- В настоящее время - ИП, работаю по контрактам.

На этом занятии

- Что такое Linux
- Логинимся
  - Сегодня с помощью логина и пароля
- Осматриваемся
- Работаем с файлами

Что такое Linux

- Как проект был начат в 1991 году Линусом Торвальдсом
  - Как собственная реализация ядра ОС Minix
- Линукс - не операционная система, а ядро.
  - Операционную систему составляет ядро, драйвера и утилиты.
- Завоевал серверный сегмент, победив проприетарные Dec Alpha, HP UX, Sun Solaris
- Пользуется популярностью у девелоперов как десктопная платформа.
- Лежит в основе Android, Chrome OS

Что такое Linux

- Проект был начат в 1991 году Линусом Торвальдсом

But why? Урок истории

- UNIX была создана в Bell Labs в 1971 году
- В 1973 году была переписана на языке C
  - Что способствовало распространению
- Завоевала популярность в промышленности и академических кругах
  - Из-за бесплатной лицензии
- В 1984 году бесплатное лицензирование закончилось
  - Появились GNU, Free Software Foundation.
- В академических кругах для изучения операционных систем популярна была MINIX
  - То же с ограничительной лицензией
- Линус Торвальдс решил создать свое ядро для MINIX
  - Со свободной для некоммерческого использования лицензией

Урок истории

- Использование в кластерах для вычислений
  - Подстегнуло разработку в компаниях
- На дешевой платформе x86
  - Дорогие проприетарные конкуренты HP, DEC, SUN, IBM, SGI
- Linux vs. BSD
  - Суды против BSD
  - Поддержка и лицензия GNU у Linux
  - Широкая поддержка девелоперов у Linux

https://en.wikipedia.org/wiki/Linux#/media/File:Unix_timeline.en.svg

Выбор дистрибутива...

- RHEL (Cent OS)
- SUSE
- Debian/Ubuntu
- Arch etc

Не идите против ветра! Вам в работе важно, чтобы все ваши инструменты были легко доступны. Работайте “как все”.

Логинимся. SSH

- SSH - Secure SHell
  - Пришла на смену telnet, rsh
- Весь обмен данными шифруется
- Позволяет логиниться с паролем и с помощью ключа
- Так же позволяет пересылать файлы в зашифрованном (или нет) виде (scp)
- Клиенты для всех платформ.
- Позволяет создавать туннель для пересылки TCP или других протоколов, например X11(удаленная графическая оболочка)
- Позволяет использовать удаленный сервер в как прокси для веб-браузера

Логинимся. SSH

- Transport Layer Protocol
  - Аутентификация сервера (сервер посылает свой публичный ключ)
  - Создание безопасного соединения для аутентификации пользователя (Диффи-Хеллман)
- User Authentication Protocol
  - Подтверждение аутентичности пользователя (пароль или сертификат)
- Connection Protocol
  - Создание канала для передачи данных в течении сессии ssh

Протокол передачи данных — набор соглашений интерфейса логического уровня, которые определяют обмен данными между различными программами. Эти соглашения задают единообразный способ передачи сообщений и обработки ошибок.

Определения

- Шифр - алгоритм(ы) преобразования открытых данных в зашифрованные
- Ключ - параметр алгоритма, который выбирает конкретное преобразование из множества возможных для данного алгоритма.
- Хеш - результат преобразования некоторых данных в строку определенной длины по определенному алгоритму (хеш-функции)

Логинимся. С паролем.

- На сервере пароль хранится в хешированном виде в /etc/shadow
- Ssh устанавливает безопасное соединение с помощью сгенерированных одноразовых ключей
- Клиент посылает пароль на сервер через зашифрованное соединение
- Устанавливается клиентская сессия

Логинимся. С паролем.

- На сервере пароль хранится в хешированном виде в /etc/shadow
- Ssh устанавливает безопасное соединение с помощью сгенерированных одноразовых ключей
- Клиент посылает пароль на сервер через зашифрованное соединение
- Устанавливается клиентская сессия
- Проблемы:
  - Можно ли доверять удаленному хосту, который запрашивает у вас пароль?
  - Man-in-the-middle attack.

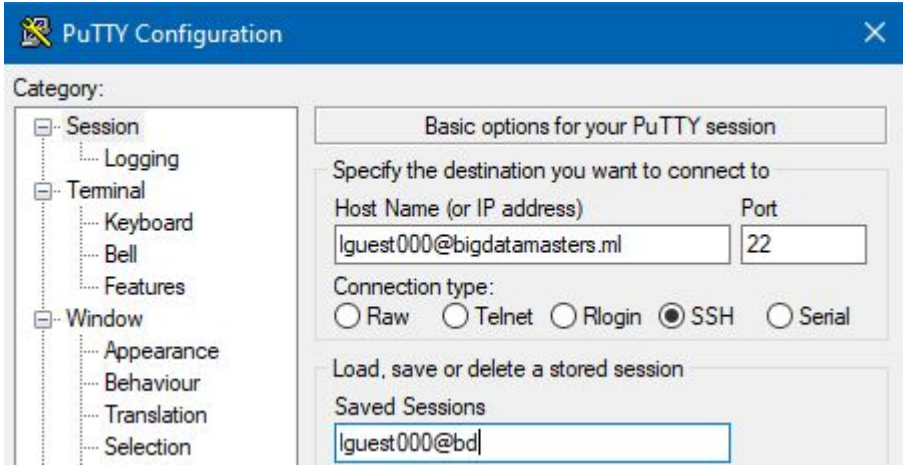
Логинимся. С паролем.

- На сервере пароль хранится в хешированном виде в /etc/shadow
- Ssh устанавливает безопасное соединение с помощью сгенерированных одноразовых ключей
- Клиент посылает пароль на сервер через зашифрованное соединение
- Устанавливается клиентская сессия
- Проблемы:
  - Можно ли доверять удаленному хосту, который запрашивает у вас пароль?
  - Man-in-the-middle attack.
    - Верификация сервера ведется через “отпечаток” ключа хоста:

```
(base) artem@artem-ubuntu0:~$ ssh u2
The authenticity of host '192.168.2.7 (192.168.2.7)' can't be established.
ECDSA key fingerprint is SHA256:Too+MbNpUmKI00GIsMzHjTTLOQgk7KBXqpVrd+mnR7U.
Are you sure you want to continue connecting (yes/no)?
```

Логинимся. Как использовать SSH

- Linux, MacOS, Windows (command shell), Windows (Subsystem for Linux)
  - ssh user@server.example.com
- Windows
  - Putty (https://www.chiark.greenend.org.uk/~sgtatham/putty/)
    - Заблокирован РКН!
    - https://drive.google.com/file/d/1EZkk0vBbdl2UpWX4j-TwNlAESTV16AgC/view?usp=sharing



Логинимся. Задание

- Залогиньтесь на сервер с помощью логина и пароля.
  - Ssh -l lguestxxx bigdatamasters.ml
    - (ваш аккаунт вместо lguestxxx)
- Login and password
  - Получите с помощью бота @ozonm_bigdata_bot
  - /start - привязка к аккаунту в базе
  - /login - получить пароль, host fingerprint и. т. д.

Осматриваемся

Знай свою систему

- Какая ОС стоит?
- Сколько ЦПУ и какие
- Сколько памяти и свопа
- Сколько дисков и места на них

Осматриваемся

```
artem@artem-ubuntu2:~$ cat /proc/cpuinfo | tail -28

processor       : 5
vendor_id       : GenuineIntel
cpu family      : 6
model           : 158
model name      : Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz
stepping        : 10
microcode       : 0xca
cpu MHz         : 1354.732
cache size      : 12288 KB
physical id     : 0
siblings        : 6
core id         : 5
cpu cores       : 6
apicid          : 10
initial apicid  : 10
fpu             : yes
fpu_exception   : yes
cpuid level     : 22
wp              : yes
flags           : fpu vme de pse tsc msr pae mce cx8 apic
pat pse36 clflush dts acpi mmx fxsr sse sse2 ss ht tm pbe
scp lm constant_tsc art arch_perfmon pebs bts rep_good nop
c cpuid aperfmperf tsc_known_freq pni pclmulqdq dtes64 mon
t tm2 ssse3 sdbg fma cx16 xtpr pdcm pcid sse4_1 sse4_2 x2a
eadline_timer aes xsave avx f16c rdrand lahf_lm abm 3dnowp
b invpcid_single pti ssbd ibrs ibpb stibp tpr_shadow vnmi
fsgsbase tsc_adjust bmi1 hle avx2 smep bmi2 erms invpcid r
 clflushopt intel_pt xsaveopt xsavec xgetbv1 xsaves dtherm
hwp_notify hwp_act_window hwp_epp flush_l1d
bugs            : cpu_meltdown spectre_v1 spectre_v2 spec_
swapgs taa itlb_multihit
bogomips        : 7392.00
clflush size    : 64
cache_alignment : 64
address sizes   : 39 bits physical, 48 bits virtual
power management:
```

Сколько ЦПУ и какие

cat /proc/cpuinfo выдает информацию о всех процессорах (ядрах) в системе. Из интересного:

- processor - порядковый, с нуля.
- vendor_id, model_name
- cpu MHz - актуальная частота
- physical_id, siblings, core_id, cpu_cores
- flags (см. avx*)

Осматриваемся

```
artem@artem-ubuntu2:~$ hostnamectl
   Static hostname: artem-ubuntu2
         Icon name: computer-desktop
           Chassis: desktop
        Machine ID: 15f72a8b21b5447191002587bff52902
           Boot ID: 49533730d5cf4cef8bceb0e0f2302b05
  Operating System: Ubuntu 18.04.2 LTS
            Kernel: Linux 4.15.0-74-generic
      Architecture: x86-64
artem@artem-ubuntu2:~$ lsb_release -a
No LSB modules are available.
Distributor ID: Ubuntu
Description:    Ubuntu 18.04.2 LTS
Release:        18.04
Codename:       bionic
artem@artem-ubuntu2:~$ cat /etc/*-release
DISTRIB_ID=Ubuntu
DISTRIB_RELEASE=18.04
DISTRIB_CODENAME=bionic
DISTRIB_DESCRIPTION="Ubuntu 18.04.2 LTS"
NAME="Ubuntu"
VERSION="18.04.2 LTS (Bionic Beaver)"
ID=ubuntu
ID_LIKE=debian
PRETTY_NAME="Ubuntu 18.04.2 LTS"
VERSION_ID="18.04"
HOME_URL="https://www.ubuntu.com/"
SUPPORT_URL="https://help.ubuntu.com/"
BUG_REPORT_URL="https://bugs.launchpad.net/ubuntu/"
PRIVACY_POLICY_URL="https://www.ubuntu.com/legal/terms
cy"
VERSION_CODENAME=bionic
UBUNTU_CODENAME=bionic
```

Какая ОС установлена

Попробуйте все эти команды на вашей системе, что-то сработает:

- hostnamectl
- lsb_release -a
- cat /etc/*-releases

Осматриваемся

Сколько памяти и свопа: free

важные

Осматриваемся

Сколько дисков и места: du, df

```
artem@artem-ubuntu2:~$ df -h
Filesystem      Size  Used Avail Use% Mounted on
udev             32G     0   32G   0% /dev
tmpfs           6.3G  1.7M  6.3G   1% /run
/dev/sda1       440G  179G  239G  43% /
tmpfs            32G  1.2G   31G   4% /dev/shm
tmpfs           5.0M  4.0K  5.0M   1% /run/lock
tmpfs            32G     0   32G   0% /sys/fs/cgroup
/dev/nvme0n1p4  180G  169G  2.2G  99% /media/data
/dev/nvme0n1p3  116G   39G   72G  36% /mnt/u1
/dev/nvme0n1p5  2.9G  221M  2.5G   9% /media/mariadb
/dev/nvme0n1p6  3.8G   21M  3.6G   1% /media/redmine
tmpfs           6.3G   64K  6.3G   1% /run/user/1000
```

не допускайте переполнения корневой файловой системы .

устройство (диск)

путь к файлам

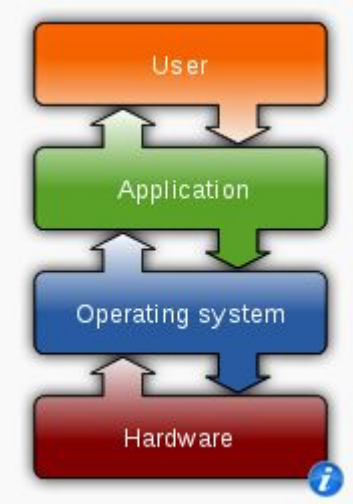
Работа с файлами.

В этом разделе

- терминология, что есть что
- навигация и поиск по ФС
- полезные команды

Файловые системы. Терминология

что это такое



- Файловая система реализует порядок хранения файлов на физических носителях.
- Она обращается к диску, используя его драйвер, и определяет интерфейс для доступа к файлам для ОС и приложений
- POSIX - интерфейс, который обеспечивает переносимость приложений и совместимость операционных систем.
- примеры не POSIX систем:
  - S3, HDFS
- типы **udev**, **tmpfs** - служебные ФС, для ядра и системы - тоже могут реализовывать POSIX для удобства пользователей.

Файловые системы - терминология

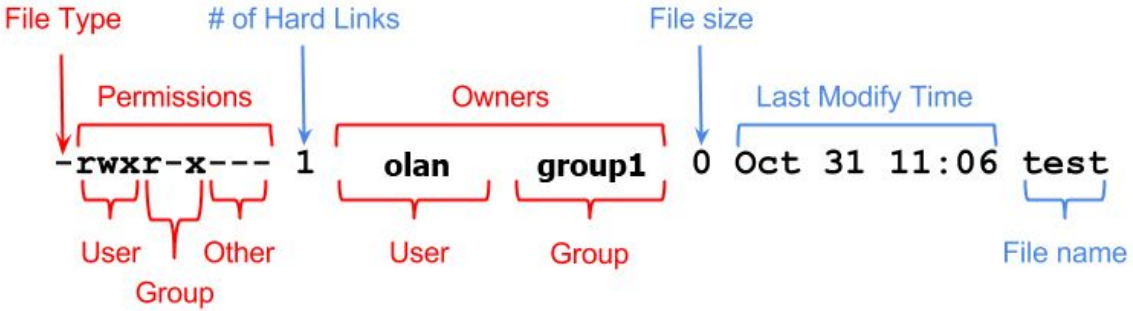
монтировка

- **/dev/sda1** - SATA disk на первом порте, первый основной раздел
- **/** - root file system
- другие файловые системы “монтируются”, т.е. становятся доступными под каким-либо путем в корневой ФС.
- **/etc/fstab** - конфигурация и опции файловых систем
- **mount** - команда для монтировки
- **mkfs** - команда для форматирования

Файловые системы - навигация

что на файловой системе?

- ls - посмотреть содержимое директории
  - -a - показывать файлы с точкой в начале.
  - -l - детальная информация о файле
  - -t - сортировка по дате (-r - обратная)
  - . - текущий каталог
  - .. - родительский каталог
- touch - создать новый (пустой) файл, обновить время доступа на существующем
- chmod - изменить аттрибуты доступа
  - **u+x, g+w, o-r**
- chown - изменить владельца:группу
  - **-R рекурсивно**



Работаем с файлами. Задание

- Создайте файл test
- Посмотрите его владельца, атрибуты доступа, время доступа, размер
- Запишите ваше имя в этот файл:
  - echo Artem >> test
- Убедитесь, что размер увеличился
  - А может быть и время доступа тоже
- Посмотрите, сколько диска использовано под ваши файлы (df)
- Посмотрите чужой файл test:
  - cat /home/guestXXX/test
- Измените атрибут доступа таким образом, чтобы никто, кроме вас не смог прочитать ваш файл test

Файловые системы - навигация

работа с файлами и папками

- **cd** - перейти в другую папку
- **mkdir** - создать новую директорию
  - **-p -** не выдавать ошибку о существовании, создать директории всех уровней.
- **rmdir** - удалить директорию (пустую)
- **rm** - удалить файлы
  - **rm -rf** - удалить (даже непустую) папку не спрашивая
- **touch** - создать пустой файл
  - на существующем файле - изменить время модификации
- **cp** - скопировать файл
  - **cp -r -** скопировать папку.
  - **cp -ra -** скопировать папку сохраняя атрибуты
- **mv** - переименовать (“передвинуть”) файл или папку
- **pwd** - вывести текущую директорию
- **du -sh** - сколько места занимают файлы в директории

Файловые системы - навигация

```
artem@artem-ubuntu2:~$ ls -1 /bin/mk*
/bin/mkdir
/bin/mknod
/bin/mktemp
artem@artem-ubuntu2:~$  ls -1 /bin/mk???
/bin/mkdir
/bin/mknod
artem@artem-ubuntu2:~$ ls -1 /bin/mk{nod,dir}
/bin/mkdir
/bin/mknod
artem@artem-ubuntu2:~$ ls -1 /bin/{mk,ch}*
/bin/chacl
/bin/chgrp
/bin/chmod
/bin/chown
/bin/chvt
/bin/mkdir
/bin/mknod
/bin/mktemp
artem@artem-ubuntu2:~$ ls -1 /bin/mk[n-z]*
/bin/mknod
/bin/mktemp
```

поиск

- ***** - любые символы
- **?** - любой один символ
- **{a,bb}** - a или bb
- **[n-z]** - от n до z.
- **find** - найти файл по признакам
  - **find ~/ -name "*.pt"** - найти все файлы с расширением .pt в своей домашней директории
  - **-newer home/artem/test.pt** - с условием, что они новее заданного файла
  - **-mtime** -2 - изменен меньше, чем 2 дня назад
  - **-type** f (or d) - либо файлы, либо директории etc
  - **-size** +650M
  - **-perm** /u+x
  - **find . -name "*.c" -type f -exec ls -l {} \;**
    - найти и выполнить команду с найденными
  - *find - вселенная (много опций). Изучайте инструкцию!*

Задание.

- Посмотрите, сколько публичных ключей сервера у нас есть
  - В папке /etc/ssh, покажите все файлы с расширением .pub
- Найдите в /usr все файлы с названием ssh
- Найдите в /usr все файлы и директории с названием ssh
- Найдите в /usr все файлы с названием ssh, которые вы можете запустить как команды (с атрибутом u+x)

Про безопасность

Безопасность превыше всего!

Потеря контроля над сервером, это:

- потеря денег и клиентов (сайт недоступен)
- утечка проприетарного кода
- утечка данных пользователей
- потеря даже всего бизнеса!

Пароли для логина

- Пароль должен быть длинным, и не обязательно содержать разный регистр, сколько-то цифр и спец. символов!
- Можно записать пароль на бумажке, если вы храните эту бумажку вдали от посторонних глаз.
- Пользуйтесь менеджерами паролей.
- Никому не сообщайте ваш пароль!
- Чтобы дать доступ другим, создайте для них отдельный аккаунт с минимальными для их работы привилегиями.

Пароли для логина

- Пароль должен быть длинным, и не обязательно содержать разный регистр, сколько-то цифр и спец. символов!
- Можно записать пароль на бумажке, если вы храните эту бумажку вдали от посторонних глаз.
- Пользуйтесь менеджерами паролей.
- Никому не сообщайте ваш пароль!
- Чтобы дать доступ другим, создайте для них отдельный аккаунт с минимальными для их работы привилегиями.

xkcdpass

Разное

- Как сгенерировать отпечаток ключа хоста
  - ssh-keygen -l -f /etc/ssh/ssh_host_ecdsa_key.pub