

O'REILLY®

Python и DevOps

Ключ к автоматизации Linux



Ной Гифт, Кеннеди Берман
Альфредо Деза, Григ Георгиу

Оглавление

Предисловие	19
Что DevOps означает для авторов	20
Как пользоваться этой книгой	22
Условные обозначения	23
Использование примеров кода	24
Благодарности	25
От издательства	28
Глава 1. Основы Python для DevOps	29
Установка и запуск Python	30
Командная оболочка Python	30
Блокноты Jupiter	31
Процедурное программирование	32
Переменные	32
Основные математические операции	33
Комментарии	33
Встроенные функции	34
Print	34
Range	35

Контроль выполнения	35
if/elif/else	36
Циклы for	37
Циклы while.....	38
Обработка исключений.....	39
Встроенные объекты	40
Что такое объект.....	40
Методы и атрибуты объектов.....	41
Последовательности.....	42
Функции.....	55
Синтаксис функции.....	55
Функции как объекты.....	57
Анонимные функции.....	57
Регулярные выражения.....	58
Поиск.....	59
Наборы символов.....	60
Классы символов.....	61
Группы.....	61
Поименованные группы.....	61
Найти все.....	62
Поисковый итератор.....	62
Подстановка.....	63
Компиляция.....	63
Отложенное вычисление.....	64
Генераторы.....	64
Генераторные включения.....	65
Дополнительные возможности IPython.....	66
Выполнение инструкций командной оболочки Unix с помощью IPython.....	66
Упражнения.....	68

Глава 2. Автоматизация работы с файлами и файловой системой	69
Чтение и запись файлов	69
Поиск в тексте с помощью регулярных выражений	79
Обработка больших файлов.....	81
Шифрование текста.....	82
Хеширование с помощью пакета hashlib	82
Шифрование с помощью библиотеки cryptography	83
Модуль os.....	85
Управление файлами и каталогами с помощью os.path.....	86
Обход дерева каталогов с помощью os.walk	90
Пути как объекты: библиотека pathlib.....	91
Глава 3. Работа с командной строкой	93
Работа с командной оболочкой.....	93
Взаимодействие с интерпретатором с помощью модуля sys	93
Взаимодействие с операционной системой с помощью модуля os.....	94
Порождение процессов с помощью модуля subprocess.....	95
Создание утилит командной строки	97
Использование sys.argv	99
argparse	101
click.....	106
fire	110
Реализация плагинов.....	115
Ситуационный анализ: разгоняем Python с помощью утилит командной строки.....	116
Динамический компилятор Numba	117
Использование GPU с помощью CUDA Python	119
Многоядерное многопоточное выполнение кода Python с помощью Numba	120
Кластеризация методом k-средних.....	122
Упражнения.....	123

Глава 4. Полезные утилиты Linux	124
Дисковые утилиты.....	125
Измерение быстродействия	126
Разделы диска	128
Получение информации о конкретном устройстве.....	129
Сетевые утилиты.....	131
SSH-туннелирование	131
Оценка быстродействия HTTP с помощью Apache Benchmark (ab)	132
Нагрузочное тестирование с помощью molotov.....	133
Утилиты для получения информации о загрузке CPU.....	136
Просмотр процессов с помощью htop.....	136
Работаем с Bash и ZSH	138
Настройка командной оболочки Python под свои нужды	140
Рекурсивные подстановки	140
Поиск и замена с запросами подтверждения	141
Удаление временных файлов Python	143
Вывод списка процессов и его фильтрация	143
Метка даты/времени Unix.....	144
Комбинирование Python с Bash и ZSH.....	144
Генератор случайных чисел.....	145
Существует ли нужный мне модуль?	146
Переходим из текущего каталога по пути к модулю	146
Преобразование CSV-файла в JSON.....	147
Однострочные сценарии Python	148
Отладчики	148
Быстро ли работает конкретный фрагмент кода?.....	149
strace.....	150
Вопросы и упражнения	153
Задача на ситуационный анализ	153

Глава 5. Управление пакетами	154
Почему пакетная организация программ так важна	155
Случаи, когда пакетная организация программ не нужна	155
Рекомендации по пакетной организации программ.....	156
Информативный контроль версий.....	156
Журнал изменений	158
Выбор стратегии	159
Решения для создания пакетов.....	159
Нативные пакеты Python.....	160
Создание пакетов для Debian	166
Создание пакетов RPM	174
Диспетчеризация с помощью systemd.....	180
Долгоживущие процессы	181
Настройка	182
Юниты systemd	183
Установка юнита.....	185
Управление журналами.....	186
Вопросы и упражнения	188
Задача на ситуационный анализ	188
Глава 6. Непрерывная интеграция и непрерывное развертывание	189
Ситуационный анализ примера из практики: перевод плохо работавшего сайта с WordPress на Hugo.....	189
Настройка Hugo	191
Преобразование WordPress в посты Hugo.....	192
Создание поискового индекса Algolia и его обновление.....	194
Координация с помощью Makefile.....	196
Развертывание с помощью AWS CodePipeline	196
Ситуационный анализ примера из практики: развертывание приложения Python App Engine с помощью Google Cloud Build.....	197
Ситуационный анализ примера из практики: NFSOPS.....	205

Глава 7. Мониторинг и журналирование	207
Ключевые понятия создания надежных систем.....	207
Неизменные принципы DevOps	208
Централизованное журналирование	209
Ситуационный анализ: база данных при промышленной эксплуатации разрушает жесткие диски	209
Производить или покупать?	210
Отказоустойчивость.....	211
Мониторинг	213
Graphite	214
StatsD.....	214
Prometheus	215
Телеметрия	219
Соглашения о наименованиях	222
Журналирование.....	224
Почему это так трудно.....	224
basicconfig.....	225
Углубляемся в конфигурацию.....	226
Распространенные паттерны.....	231
Стек ELK	232
Logstash.....	233
Elasticsearch и Kibana.....	235
Вопросы и упражнения	239
Задача на ситуационный анализ	239
Глава 8. Pytest для DevOps	240
Сверхспособности тестирования фреймворка pytest	240
Начало работы с pytest	241
Тестирование с помощью pytest.....	242
Отличия от unittest.....	244

Возможности pytest.....	245
conftest.py	246
Этот замечательный оператор assert	247
Параметризация	248
Фикстуры.....	250
Приступим	250
Встроенные фикстуры.....	252
Инфраструктурное тестирование	255
Что такое проверка системы.....	256
Введение в Testinfra	257
Подключение к удаленным узлам	258
Фикстуры и особые фикстуры	261
Примеры.....	263
Тестирование блокнотов Jupyter с помощью pytest	266
Вопросы и упражнения	267
Задача на ситуационный анализ	267
Глава 9. Облачные вычисления	268
Основы облачных вычислений.....	269
Типы облачных вычислений.....	271
Типы облачных сервисов.....	272
Инфраструктура как сервис	272
«Железо» как сервис	277
Платформа как сервис.....	278
Бессерверная обработка данных.....	278
Программное обеспечение как сервис.....	282
Инфраструктура как код	283
Непрерывная поставка.....	283
Виртуализация и контейнеры	283
Аппаратная виртуализация	283
Программно определяемые сети	284

Программно определяемое хранилище	285
Контейнеры	285
Трудные задачи и потенциальные возможности распределенной обработки данных.....	286
Конкурентное выполнение на Python, быстродействие и управление процессами в эпоху облачных вычислений	289
Управление процессами	289
Процессы и дочерние процессы	289
Решение задач с помощью библиотеки multiprocessing.....	292
Ветвление процессов с помощью Pool()	293
Функция как сервис и бессерверная обработка данных.....	295
Повышение производительности Python с помощью библиотеки Numba	295
Динамический компилятор Numba	295
Высокопроизводительные серверы.....	296
Резюме.....	297
Вопросы.....	298
Вопросы на ситуационный анализ.....	298
Глава 10. Инфраструктура как код	299
Классификация инструментов автоматизации выделения инфраструктуры	301
Выделение инфраструктуры вручную.....	302
Автоматическое выделение инфраструктуры с помощью Terraform.....	304
Выделение корзины S3	304
Предоставление SSL-сертификата с помощью ACM AWS	307
Выделение раздачи Amazon CloudFront.....	308
Создание записи DNS Route 53	311
Копирование статических файлов в корзину S3.....	312
Удаление всех ресурсов AWS, выделенных с помощью Terraform.....	313

Автоматическое выделение инфраструктуры с помощью Pulumi.....	313
Создание нового проекта Pulumi на Python для AWS.....	314
Создание значений параметров конфигурации для стека staging.....	319
Создаем SSL-сертификат ACM.....	319
Выделение зоны Route 53 и записей DNS.....	320
Выделение раздачи CloudFront.....	323
Создание записи DNS Route 53 для URL сайта.....	324
Создание и развертывание нового стека.....	325
Упражнения.....	327
Глава 11. Контейнерные технологии: Docker и Docker Compose.....	328
Что такое контейнер Docker.....	329
Создание, сборка, запуск и удаление образов и контейнеров Docker.....	330
Публикация образов Docker в реестре Docker.....	334
Запуск контейнера Docker из одного образа на другом хост-компьютере....	335
Запуск нескольких контейнеров Docker с помощью Docker Compose.....	337
Портирование сервисов docker-compose на новый хост-компьютер и операционную систему.....	350
Упражнения.....	354
Глава 12. Координация работы контейнеров: Kubernetes.....	355
Краткий обзор основных понятий Kubernetes.....	356
Создание манифестов Kubernetes на основе файла docker_compose.yaml с помощью Kompose.....	357
Развертывание манифестов Kubernetes на локальном кластере Kubernetes, основанном на minikube.....	359
Запуск кластера GKE Kubernetes в GCP с помощью Pulumi.....	374
Развертывание примера приложения Flask в GKE.....	377
Установка чартов Helm для Prometheus и Grafana.....	383
Удаление кластера GKE.....	389
Упражнения.....	390

Глава 13. Технологии бессерверной обработки данных.....	391
Развертывание одной и той же функции Python в облака большой тройки поставщиков облачных сервисов.....	394
Установка фреймворка Serverless.....	394
Развертывание функции Python в AWS Lambda.....	395
Развертывание функции Python в Google Cloud Functions.....	397
Развертывание функции на Python в Azure.....	403
Развертывание функции на Python на самохостируемых FaaS-платформах.....	408
Развертывание функции на Python в OpenFaaS.....	408
Выделение таблиц DynamoDB, функций Lambda и методов API Gateway с помощью AWS CDK.....	416
Упражнения.....	438
Глава 14. MLO и разработка ПО для машинного обучения	439
Что такое машинное обучение.....	439
Машинное обучение с учителем.....	440
Моделирование	442
Экосистема машинного обучения языка Python.....	445
Глубокое обучение с помощью PyTorch	445
Платформы облачного машинного обучения.....	449
Модель зрелости машинного обучения.....	451
Основная терминология машинного обучения.....	452
Уровень 1. Очерчивание рамок задачи и области определения, а также формулировка задачи	453
Уровень 2. Непрерывная поставка данных.....	453
Уровень 3. Непрерывная поставка очищенных данных	455
Уровень 4. Непрерывная поставка разведочного анализа данных	457
Уровень 5. Непрерывная поставка обычного ML и AutoML.....	457
Уровень 6. Цикл обратной связи эксплуатации ML.....	458

Приложение sklearn Flask с использованием Docker и Kubernetes.....	459
Разведочный анализ данных.....	463
Моделирование	464
Тонкая настройка масштабированного GBM.....	465
Подгонка модели.....	466
Оценка работы модели.....	466
adhoc_predict.....	467
Технологический процесс JSON.....	468
Масштабирование входных данных.....	468
adhoc_predict на основе выгрузки.....	470
Масштабирование входных данных.....	470
Вопросы и упражнения	471
Задача на ситуационный анализ	471
Вопросы на проверку усвоения материала.....	471
Глава 15. Инженерия данных	472
Малые данные.....	473
Обработка файлов малых данных	474
Запись в файл	474
Чтение файла.....	474
Конвейер с генератором для чтения и обработки строк.....	475
YAML	476
Большие данные	477
Утилиты, компоненты и платформы для работы с большими данными.....	479
Источники данных.....	480
Файловые системы	480
Хранение данных.....	481
Ввод данных в режиме реального времени	483
Ситуационный анализ: создание замороженного конвейера данных	484
Бессерверная инженерия данных.....	485

AWS Lambda и события CloudWatch	486
Журналирование Amazon CloudWatch для AWS Lambda.....	486
Наполнение данными Amazon Simple Queue Service с помощью AWS Lambda	487
Подключение срабатывающего по событию триггера CloudWatch.....	492
Создание событийно-управляемых функций Lambda	493
Чтение событий Amazon SQS из AWS Lambda.....	493
Резюме.....	498
Упражнения.....	498
Задача на ситуационный анализ	498
Глава 16. Истории из практики DevOps и интервью.....	499
Киностудия не может снять фильм	500
Разработчик игр не может обеспечить поставку игрового ПО.....	503
Сценарии Python, запуск которых требует 60 секунд	505
Решаем горячие проблемы с помощью кэша и интеллектуальной телеметрии.....	506
Доавтоматизироваться до увольнения.....	507
Антипаттерны DevOps.....	509
Антипаттерн: отсутствие автоматизированного сервера сборки	509
Работать вслепую	509
Сложности координации как постоянная проблема.....	510
Отсутствие командной работы	511
Интервью.....	517
Гленн Соломон	517
Эндрю Нгуен	518
Габриэлла Роман	520
Ригоберто Рош	521
Джонатан Лакур.....	523
Вилле Туулос	525
Джозеф Рис.....	527

Тейо Хольцер	529
Мэтт Харрисон	531
Майкл Фоорд	533
Рекомендации	536
Вопросы	537
Интересные задачи	537
Дипломный проект	538
Об авторах	539
Об иллюстрации на обложке	541