

Эволюция инфраструктуры Bidmachine за 2 года

Uladzimir Okala-Kulak

Minsk, 2020

План

1 Вступление

2 В бой

3 Выводы

Пара слов про докладчика

- Ставил GNU/Linux на десктоп когда это ещё не было мэйнстримом (2004, Mandrake 9.1).
- Gentoo в production – нормально, если знаешь как и зачем.
- Работал с Solaris и AIX и сохранил психическое здоровье.
- Сотрудничество с Course.by, FriendlyData, ITG, ByChange.
- МНС в НАН. Админ и эникей в Белгосстрах. Инженер по внедрению в ScienceVision. Системный инженер и Linux-админ в ScienceSoft. Админ в SE Ranking. Короче много лет в администрировании и системном инжениERINGе.
- Уже 2 года системный инженер в Stack Bidmachine.

О чём со мной поговорить на пикнике

- UNIX;
- Контейнеризация;
- Контейнерная оркестрация;
- CI/CD;
- DevOps;
- SOA, microservices;
- LISP.

Пара слов про организационные изменения в компании за 2 года

Компания «Appodeal» имела основной одноимённый продукт. И была пара «дочерних» смежных продуктов: AppodealX и Appgrowth. Но грянула реструктуризация и всё немного изменилось.

- компания «Appodeal» → группа компаний «Stack»
- продукт Appodeal → Stack Appodeal
- продукт Appgrowth → Stack AppGrowth
- продукт AppodealX → Stack Bidmachine
- выделились и появились дополнительные продукты и комплексные решения (Stack UA, Stack DataCore, Holistic Solution, etc)

У каждого продукта – своя команда разработчиков, но...
Некоторое время разработка SDK и frontend, а также тестирование и
администрирование велись силами функциональных команд.
Выделение отдельных юридических лиц, изменение процессов и
внедрение DevOps практик ускорило распад функциональных команд,
произошёл полный переход к продуктовым командам.
Это всё не особо интересно, это только чтобы дать контекст.
С инфраструктурой-то что?

Начальное состояние инфраструктуры

- Единственный ЦОД в европейском «Dedicated servers hosting».
- Один кластер DC/OS и для production и staging окружений.
- Один кластер kafka на bare-metal.
- Один кластер druid на bare-metal.
- Набор сервисов tranquility для импорта данных из kafka в druid.
- База postgresql на bare-metal для служебных данных и конфигурации.
- Основные приложения на scala. OpenJDK.
- Подход к проектированию всего решения – SOA.
- Отсутствие CI/CD.
- Отсутствие Feature toggling.
- Частичная реализация IaC. Ansible.
- Зоопарк дистрибутивов: Ubuntu 16.04, Ubuntu 18.04, Debian 9, CentOS 7.
- Серверов под DC/OS < 50. Серверов под всё остальное < 50.
- Bare-metal, XEN, Docker-hosts, DC/OS.
- Инфраструктура мониторинга и логирования общая с другими продуктами.

Акт 1

- Понять вообще что к чему. Какие «боли» у разработчиков?
- Внедрить CI/CD-tool. Да хотя бы и Jenkins.
- Изменить git flow и work flow. Разработать и внедрить первоначальный черновой вариант pipelines.

Акт 2

- «Раскурить» druid.
- Контейнеризовать все его компоненты (coordinator, overlord, router, historical, middlemanager, broker) и внешние зависимости (zookeeper, pxc).
- Подготовить ansible playbooks для выката, обновления, переконфигурации.
- Подготовить миграцию и обновление.
- Отказаться от tranquility. Внедрить Kafka Indexing Service.
- Tuning.
- Внедрить multi-tier.
- Реализовать вменяемый мониторинг и логирование.
- Мигрировать с pivot на turnilo. Добавить интеграцию с dashboard.

Акт 3

- Добавить ЦОД в US.
- Внедрить k8s US.
- Добавить kafka кластер в US DC. Внедрить mirror-maker.
- Добавить postgresql в US DC. Настроить репликацию.
- CI/CD для деплоя в k8s.
- Обеспечение самодостаточности образов.
- Частичная реализация подхода Immutable Infrastructure.

Акт 4

- Контейнеризовать и обновить kafka и zookeeper.
- Контейнеризовать и обновить postgresql.
- Мигрировать с DCOS EU на k8s EU.
- Внедрить отдельный k8s кластер для staging окружений.
- «Вычистить» и улучшить CI/CD pipelines.
- Ускорение разработки scala-приложений. ТТМ. Повышение удобства. scala 2.12 → scala 2.13. OpenJDK → GraalVM.

Акт 5

- Контейнеризовать и обновить EFK-стэк.
- Контейнеризовать, обновить и отделить zabbix, prometheus, grafana.
- Непрерывно допиливать мониторинг, алертинг, логирование.
Вечный пункт, вечный!

Акт 6

- Внедрить spark-operator.
- Реализовать CI/CD для spark-приложений.
- Поставить на мониторинг.
- Научиться с этим жить. Выпустить в большой production. WIP...

Акт последний

- Мигрировать с octodns на terraform.
- Эксперименты с Loki и Loghouse.
- Завершение миграции с ansible -m k8s на helm и werf.
- Победить таки «CI/CD as a Code».
- Реализовать таки честный GitOps.
- Построить по-настоящему гибкую гибридную инфраструктуру.

Текущее состояние инфраструктуры

- 2 ЦОД в EU и US «Dedicated servers hosting». И cloud DC в APAC (в процессе).
- k8s. Отдельный k8s кластер для staging окружений.
- В каждом DC по kafka кластеру (Docker-hosts). Отдельный kafka кластер для staging, начало миграции staging kafka в k8s staging.
- Production кластер druid на Docker-hosts. И отдельный для staging.
- Kafka Indexing Service полностью заменяет tranquility. Tranquility теперь как скрипач.
- Spark на подходе.
- База postgresql уже последняя и на Docker-hosts и только для служебных данных. И с репликацией между ЦОДами.
- Основные приложения на scala (2.13). GraalVM вместо OpenJDK.
- SOA. И пока именно так. Нет монолиту, но и пока без ричардсоновских микросервисов.
- CI/CD. Идём к полным pipelines и «CI/CD as a Code».
- Feature toggling (в процессе).
- Практически полная реализация IaC.
- Унификация дистрибутивов: CentOS 7 (XEN), Ubuntu 18.04 (k8s, Docker-hosts и всё остальное), Ubuntu 20.04 (в процессе тестирования).
- Серверов под k8s – единицы сотен. Серверов под всё остальное – около сотни.
- Инфраструктура мониторинга и логирования отделена и контейнеризирована.

Выводы, итоги, замечания

- Внедрение DevOps практик ускоряет разрушение функциональных команд.
- CI/CD нужен всем. TTM и ZDD – важны.
- Метод последовательных приближений работает.
- Весь stateless в оркестраторе (k8s).
- Весь stateful можно контейнеризовать (Docker-hosts, вне k8s).
- k8s побеждает DC/OS (как минимум у нас).
- Jenkins – не удобен, подбешивает, но работает хорошо. Завтра слушайте мой блиц-доклад «Отправим Jenkins на покой. Или чем он меня бесит.» на BEER-пикнике.
- FFDD – вещь! Вспоминайте божественный доклад Жени Гарлуковича на прошлогоднем В.Е.Е.Р.
- Spark – это то, что нельзя внедрить в production за пару дней.
- Горизонтальное и почти линейное масштабирование почти всех слоёв и компонентов инфраструктуры – это очень круто.
- Теперь разработчики все.

Подяки

Всі ці зміни в інфраструктурі я не зміг би зробити сам.
Велика подяка системним інженерам Stack з Луцька.
Пошана та повага справжнім професіоналам!



Благодарности

Разработчики и вся команда продукта также не подкачали.
Всем салют и благодарность.



Спасибо за внимание.

Да будет срач в вопросах и комментах

Вопросы

Замечания

Угрозы

Оскорблений

Предложения

kulak@itg.by

+375292751078