

Distributed System Scheduler



Студент: Антон Мордберг

Научные руководители: Дмитрий Калашников, Никита Толстиков

JetBrains, лето 2016

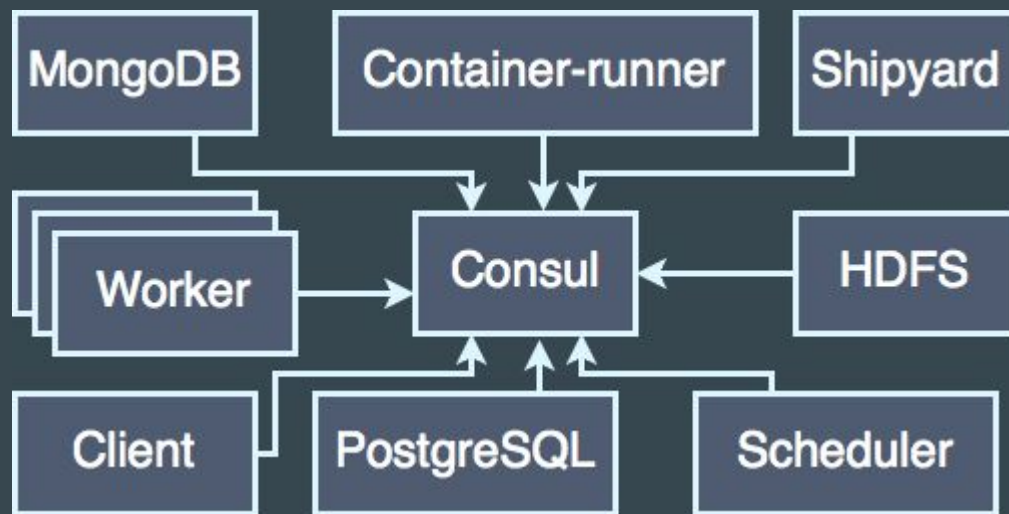
Мотивация

- Отдел маркетинга хочет получать ответы на вопросы, связанные с анализом данных.
- Для каждой задачи нужен свой сервис со своими требованиями к ресурсам.
- Сервисы нужно регулярно запускать и не всегда просто по расписанию.
- Существует множество фреймворков для решения конкретных задач, но универсального.

Общая цель проекта

- Создать систему для автоматизации сборки и анализа данных
 - Масштабируемость
 - Устойчивость к отказам компонент
 - Распределение нагрузки и ресурсов
- Добавить планировщик задач
- Отслеживать состояние компонентов системы

Структура системы

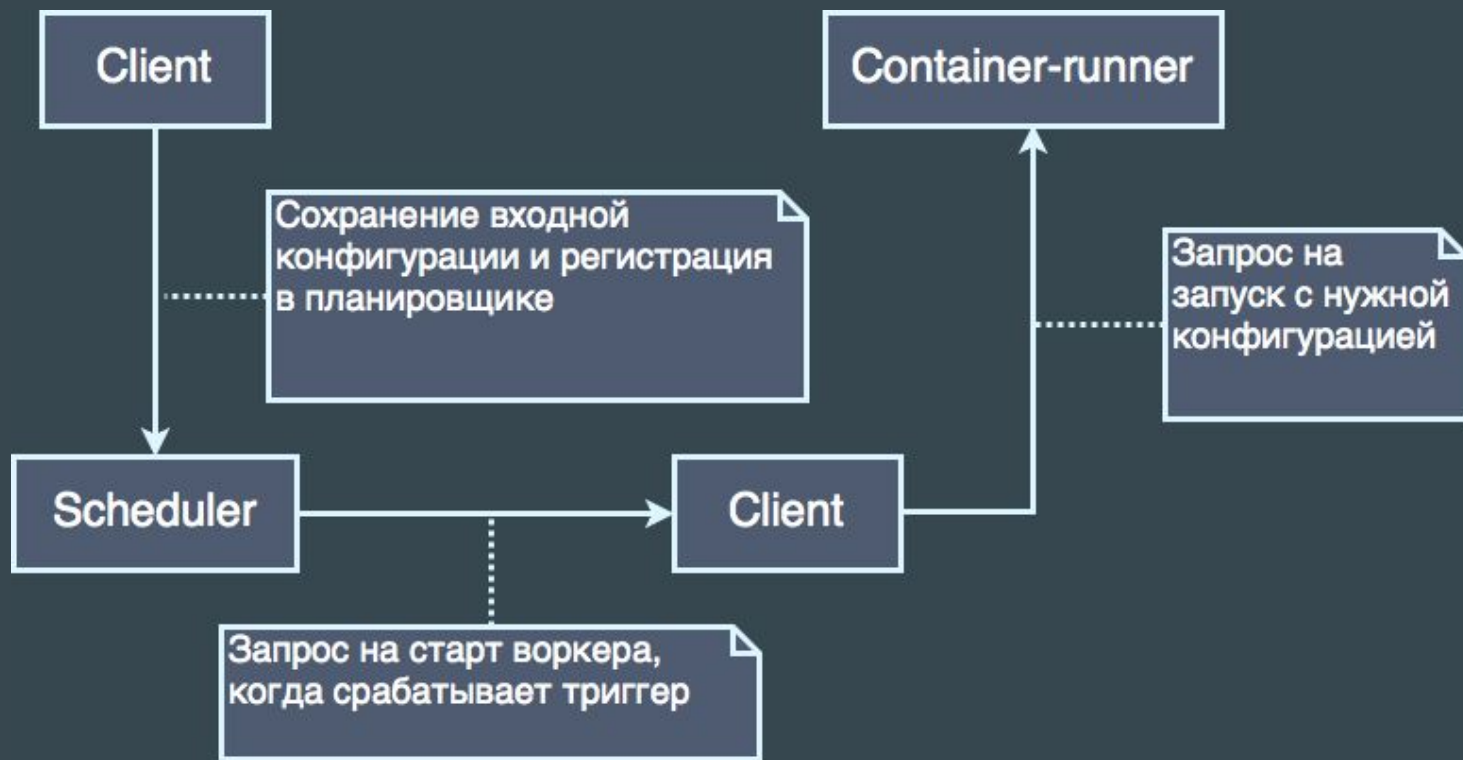


- Каждый компонент это docker контейнер
- Внутри контейнера находится микросервис с consul клиентом

Scheduler

- Планировщик запусков задач
- Поддержка запуска сложного графа задач
- Персистентность и транзакционность
- Масштабируемость
- Юнит тесты

Принцип работы



Технологии



Результат

Работающий планировщик задач:

- Rest Service на основе Spring MVC
- Работа через Quartz Framework
- Связка с PostgreSQL
- Health Checks
- Unit, stress tests

Что сделал я

- Доработал архитектуру существующей системы.
- Создал сервис планировщика и переработал сервисы, с которыми он взаимодействует.
- Доработал несколько worker'ов.
- Провел тестирование системы.
- Подобрал и настроил некоторые параметры системы (hystrix, hateoas).
- Частично задеплоил систему.

Спасибо за внимание!