
PLARIUM

СЕГМЕНТАЦИЯ ДАННЫХ В ИГРОВЫХ КЛАСТЕРАХ

Consistent Hashing Concept

Павел Матлашов

Director of Game Server Development Department

UDEV TECH MEET UPS

CONTENTS

1 Задача сегментации данных.

2 Статическая сегментация.

3 Накопленный опыт и пересмотр проблемы.

4 Концепция DHT.

5 Consistent hashing в действии.

6 Итоги применения DHT.

ЗАДАЧА СЕГМЕНТИРОВАНИЯ ДАННЫХ

Обеспечивает линейную контролируемую горизонтальную масштабируемость

ПЛЮСЫ:

+ Расширение границ системы за рамки нескольких серверов.
+ Адаптируемость под различные параметры нагрузки.
+ Наличие нескольких стратегий сегментирования под характеристики системы.
+ Сегментировать дешевле и проще, чем планировать стоимость и емкость вертикально масштабируемой системы.

МИНУСЫ:

- Повышенная сложность поддержания согласованности сегментов.
- Возросшая стоимость разработки.
- Повышенная стоимость администрирования (настройка, резервные копии, миграция и т.д.).

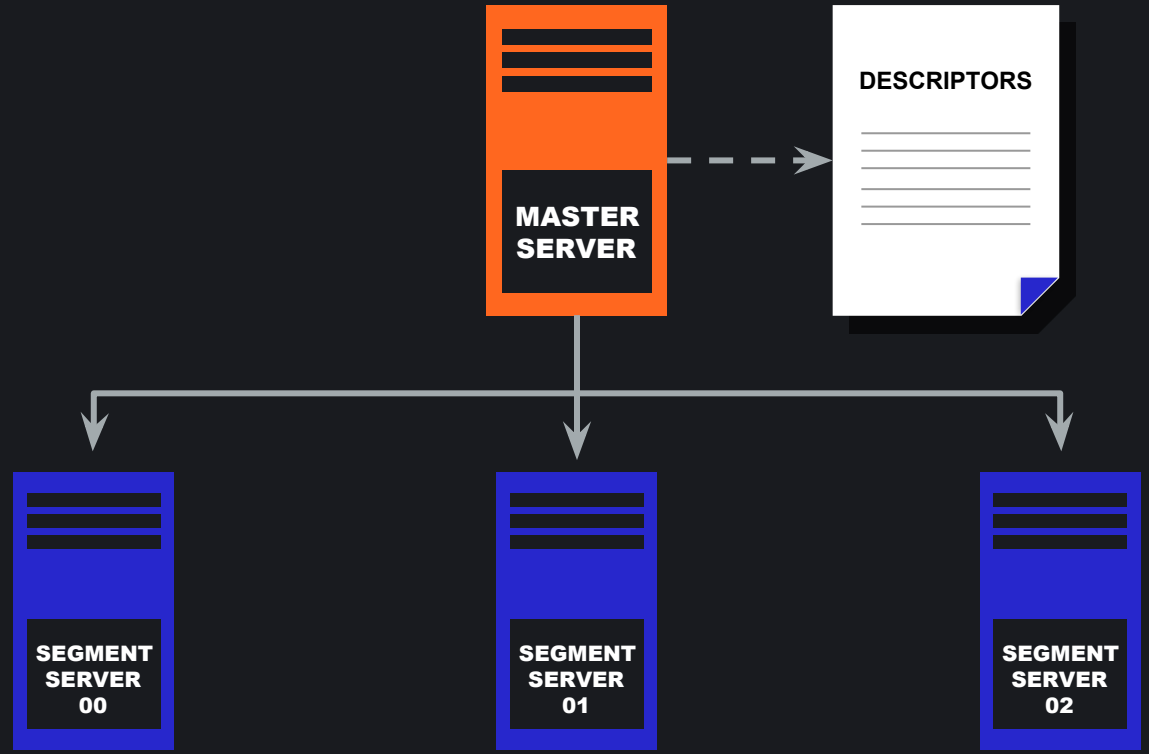
СТАТИЧЕСКАЯ СЕГМЕНТАЦИЯ

Основные типы сущностей распределяются по централизованному алгоритму, создаются дескрипторы сущностей, указывающие их расположение в кластере (сегмент).

ПЛЮСЫ: просто, надёжно.

МИНУСЫ: отсутствие перебалансировки (хоть и не понадобилось), лишние данные, ограничения алгоритма сегментации.

СТАТИЧЕСКАЯ СЕГМЕНТАЦИЯ



НАКОПЛЕННЫЙ ОПЫТ И ПЕРЕСМОТР ПРОБЛЕМЫ

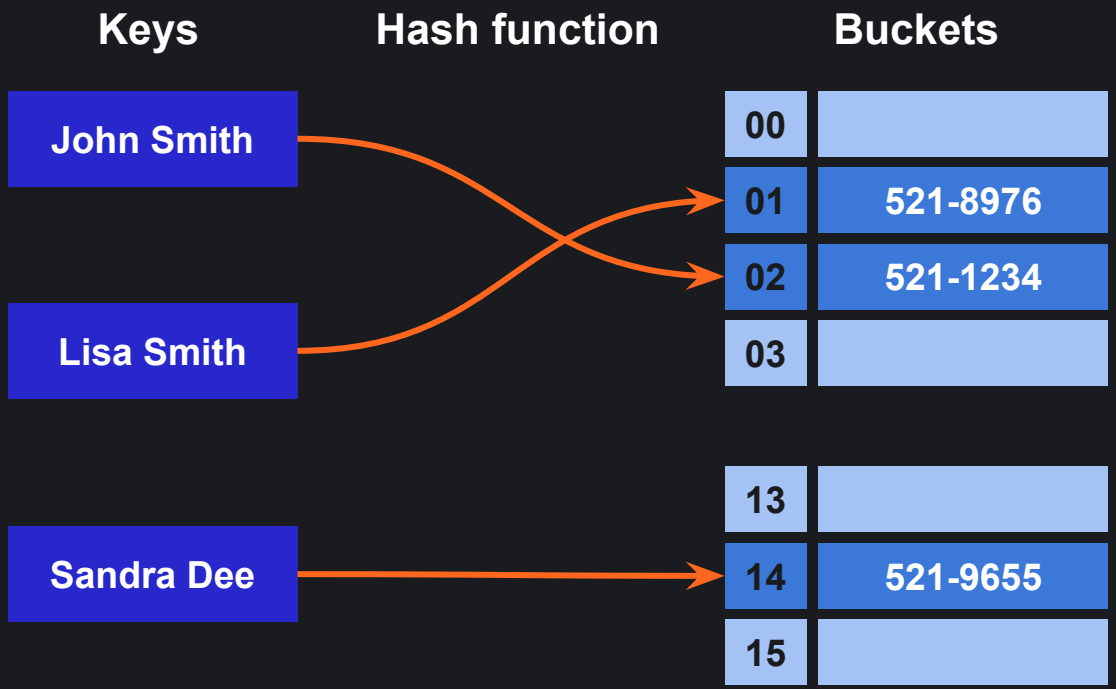
- Со временем собрали ряд повторяющихся паттернов при решении задач доступа к распределенным данным (кэширование, нотификации об изменении, интеграция с EventBus и т.д).
- Решили в очередной раз обобщить паттерны в повторно используемом механизме.
- Одновременно с этим решили попробовать сделать динамическое распределение без дескрипторов.

НАКОПЛЕННЫЙ ОПЫТ И ПЕРЕСМОТР ПРОБЛЕМЫ

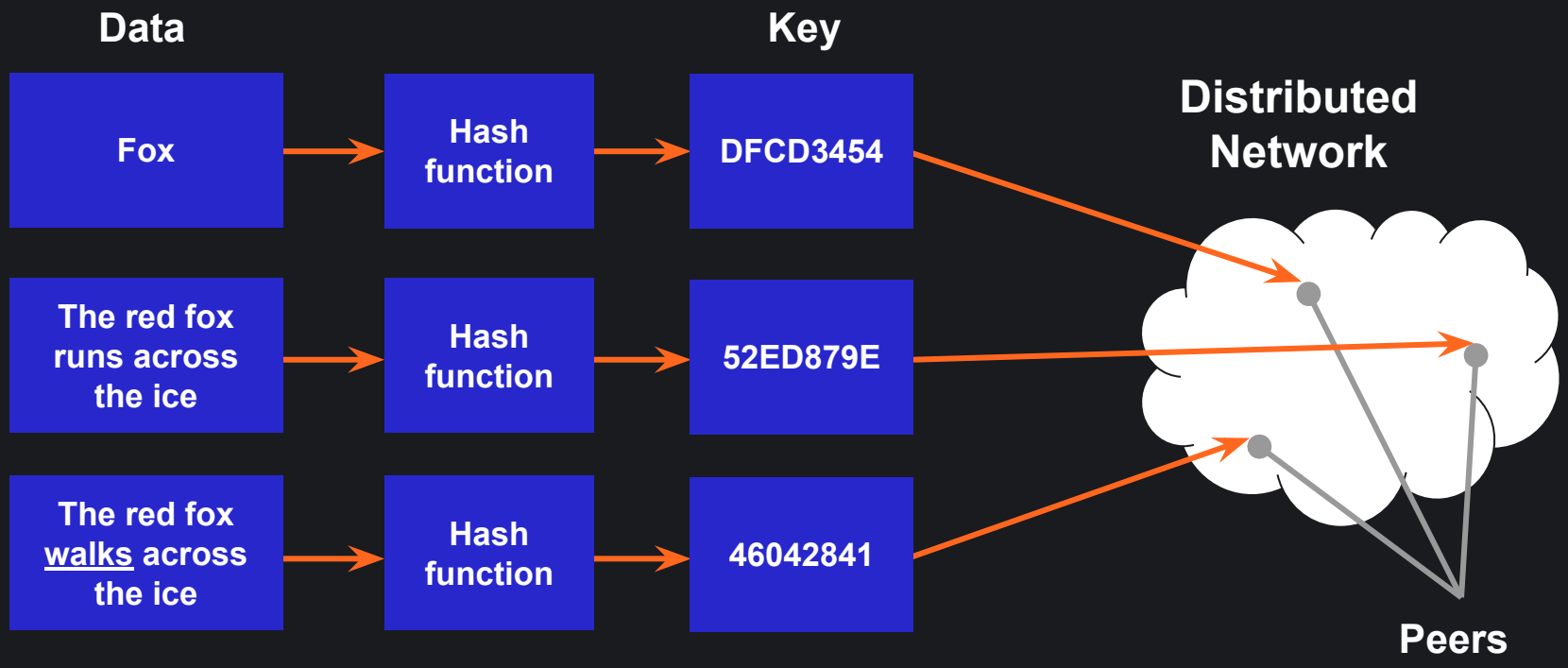
КОНЦЕПЦИЯ DHT

- Описание основных принципов.
- Consistent hashing, rendezvouz hashing etc.
- Overlay network.
- Что выбрали.

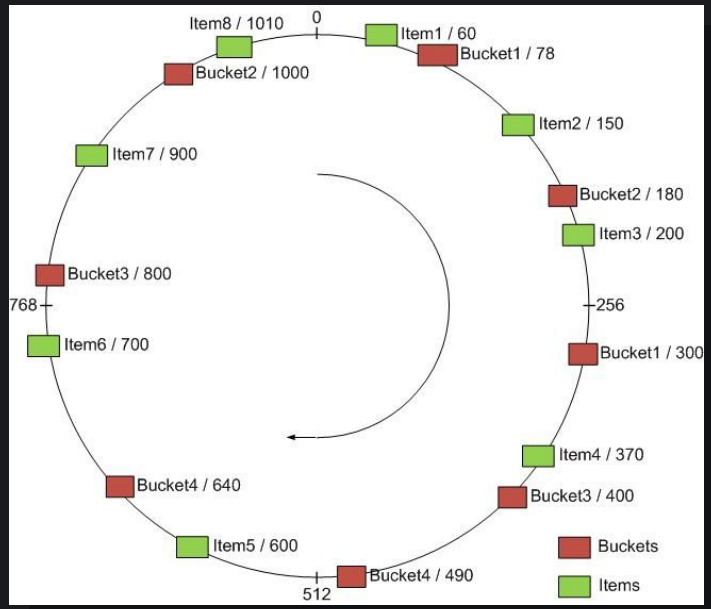
ХЕШ ТАБЛИЦА



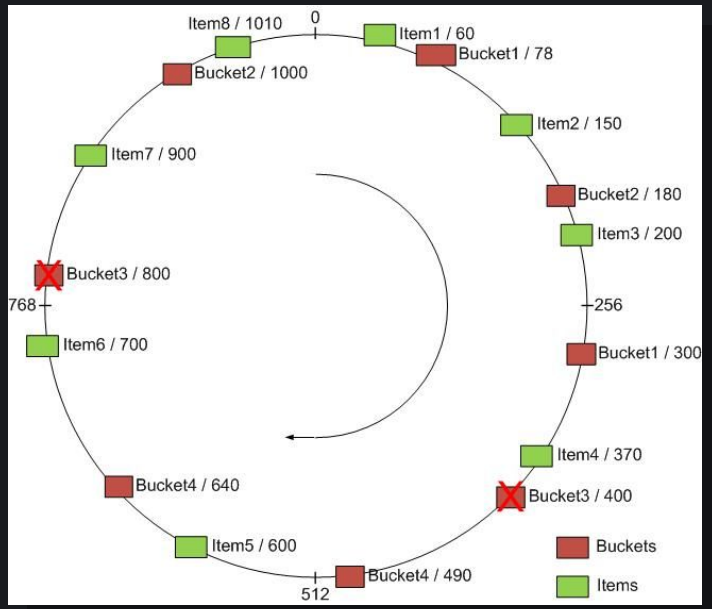
РАСПРЕДЕЛЁННАЯ ХЕШ ТАБЛИЦА



CONSISTENT HASHING



k/N



ИСПОЛЬЗОВАНИЕ CH В ПРОДУКТАХ И ПЛАТФОРМАХ

- Couchbase automated data partitioning.
- Openstack's Object Storage Service Swift.
- Partitioning component of Amazon's storage system Dynamo.
- Data partitioning in Apache Cassandra.
- Data Partitioning in Voldemort.
- Akka's consistent hashing router.
- Riak, a distributed key-value database.
- GlusterFS, a network-attached storage file system.
- Skylable, an open-source distributed object-storage system.
- Akamai Content Delivery Network.
- Discord chat application.
- Maglev: A Fast and Reliable Software Network Load Balancer.

RENDEZVOUS HASHING (HRW HASHING)

- Расчет очков для каждого узла на основе выбранной хеш-функции и seed узла.
- Выбираем узел с наибольшим количеством очков.
- Можно добавить вес.

https://en.wikipedia.org/wiki/Rendezvous_hashing

ИСПОЛЬЗОВАНИЕ HRW

- Oracle in memory DB.
- Microsoft Cache Array Routing Protocol.
- Router design.
- Mobile caching studies.

OVERLAY NETWORK

- Зачем и где это нужно?
- Почему не нужно нам?

CONSISTENT HASHING В ДЕЙСТВИИ. MurMur

- Быстрый, некриптостойкий алгоритм хеширования с хорошей защитой от коллизий.
- Масса реализаций под различные языки и платформы.

https://en.wikipedia.org/wiki/MurmurHash
https://sites.google.com/site/murmurhash/

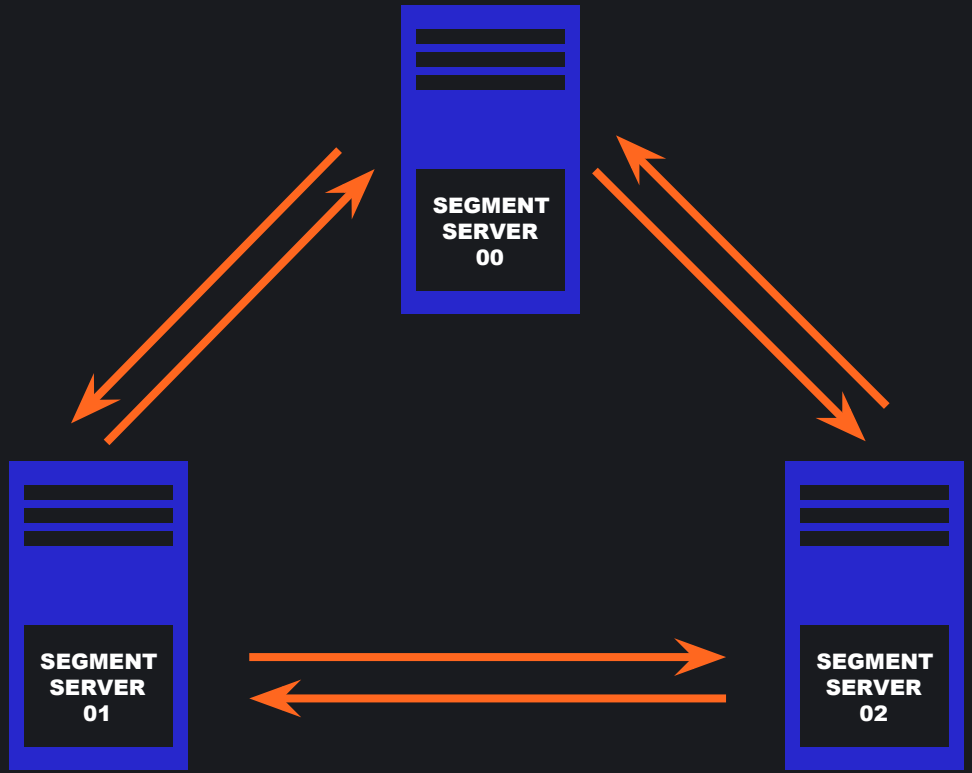
```
OneAtATime - 354.163715 mb/sec
FNV - 443.668038 mb/sec
SuperFastHash - 985.335173 mb/sec
lookup3 - 988.080652 mb/sec
MurmurHash 1.0 - 1363.293480 mb/sec
MurmurHash 2.0 - 2056.885653 mb/sec
```

CONSISTENT HASHING В ДЕЙСТВИИ

МИГРАЦИЯ ДАННЫХ

- Узлы (сегменты) в текущей архитектуре лишь добавляются.
- Можно задать специальные параметры распределения.
- Миграция хранилища выполняется асинхронно, с координацией сегментов.

CONSISTENT HASHING В ДЕЙСТВИИ

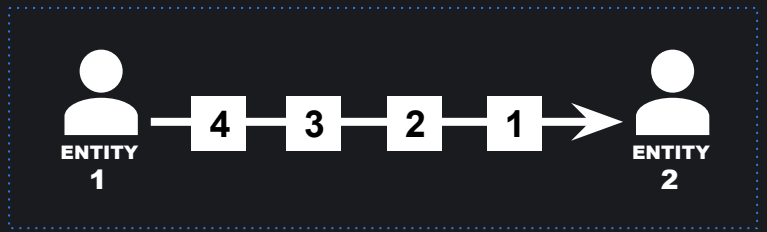


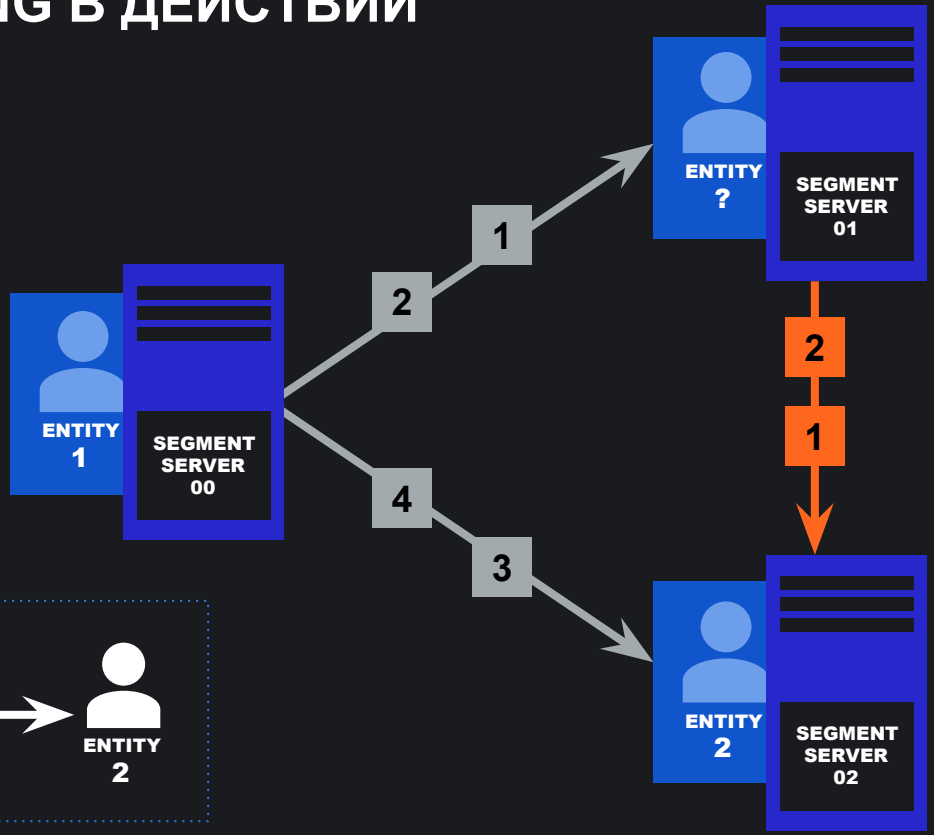
CONSISTENT HASHING В ДЕЙСТВИИ

МИГРАЦИЯ EventBus

- Возникла проблема поддержания гарантии порядка событий.
- Схема миграции событий аналогична асинхронной, только выполняется в три этапа: завершение потенциально прерванной прошлой миграции, миграция входящих событий, миграция исходящих событий.

CONSISTENT HASHING В ДЕЙСТВИИ

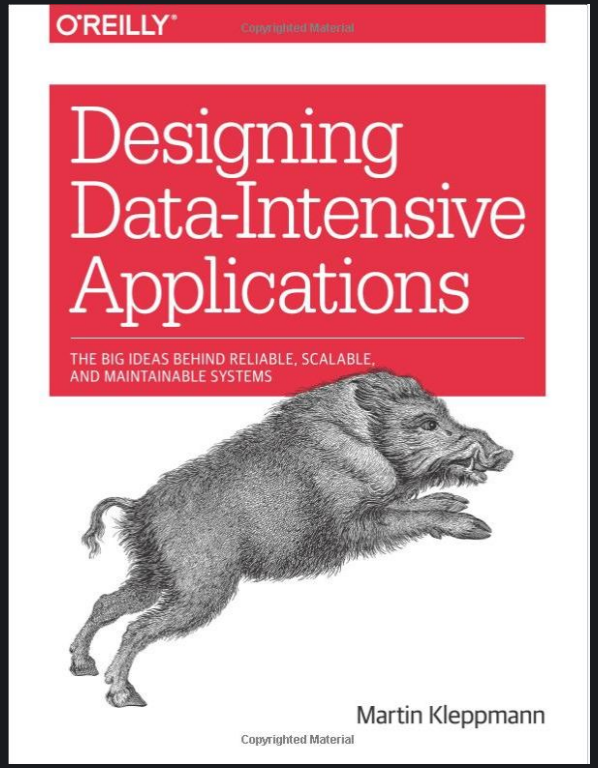


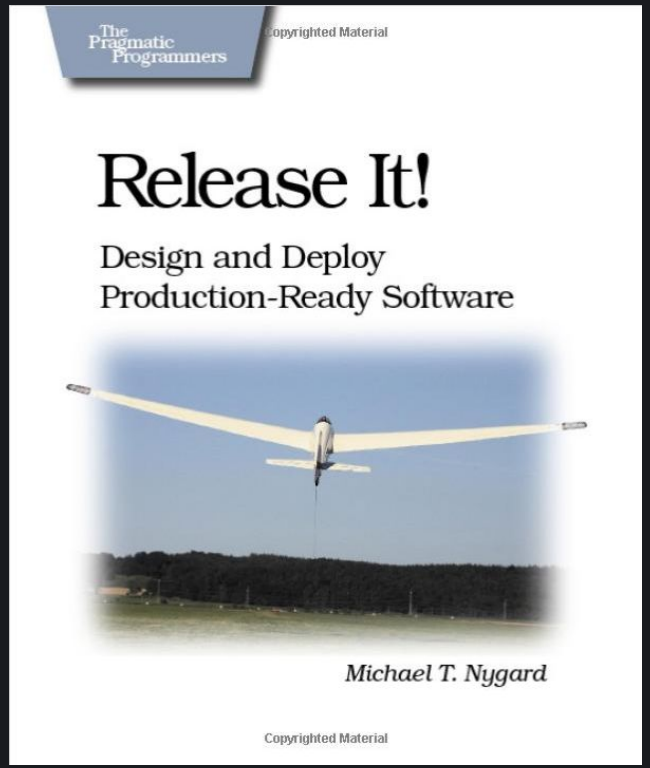


ИТОГИ ПРИМЕНЕНИЯ DHT

- Накопленный опыт + новые знания = развитие.
- Работа одновременно устоявшихся и новых решений.
- Дальнейшие варианты развития.

Что почитать?





THANK YOU

TAKE THE WORLD