



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Системный дизайн современных приложений

Семинар №8.
Распределенные системы и высокая
доступность



О курсе [2]

Формула

$0.1 * \text{Работа на семинарах} + 0.7 * \text{ДЗ} + 0.2 * \text{Теоретический тест} + 0.1 * \text{Extra Point}$

1. Работа на семинарах (0.1)

Практика будет идти через преподавателя. Во время семинара поощряется актив. Все, кто работает на семинарах - хоть какой-то актив, получают 10 баллов. Кто присутствует и не работает 6 баллов

2. ДЗ (0.7)

Будет 4 ДЗ (первое 0.25, остальные каждое 0.15)

Домашние задания заключаются в том, что нужно будет что-то развернуть у себя и сделать какую-то работу над этим. Накидать небольшой отчет в Readme от себя, описать свои мысли

3. Теоретический тест (0.2)

Нужно будет ответить на практико-теоретические вопросы. Всего вопросов будет 10 штук. 0.1 балл от 6+ правильных ответов. Тест ближе к концу курса на 10 минут

4. Особая признательность и актив (0.1)

Ставится студенту, с которым будем делать System Design на последней паре



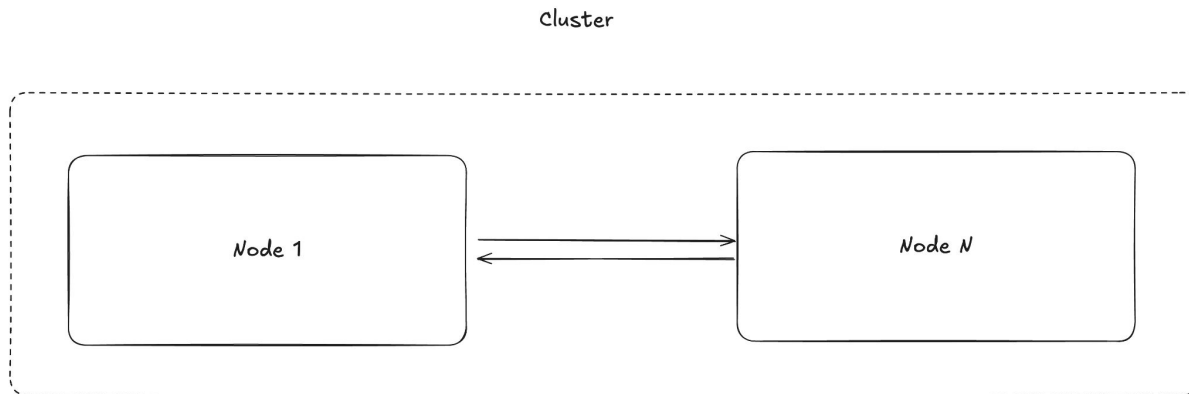
Распределенная система

Распределенной вычислительной системой можно назвать такую систему, в которой отказ компьютера, о существовании которого вы даже не подозревали, может сделать ваш собственный компьютер непригодным к использованию.

- Лесли Лампорт

Распределенная вычислительная система (РВС) – это набор соединенных каналами связи независимых компьютеров, которые с точки зрения пользователя некоторого программного обеспечения выглядят **единым целым**

- Эндрю Таненбаум





Почему они победили (или нет)

Hey I just met you
The network's laggy
But here's my data
So store it maybe

Почему побеждают:

1. Дешевеет сеть, становится стабильной.
2. Вертикальное масштабирование ограничено.
3. Взаимозаменяемость.

Проблемы:

1. Ааа сложна
2. Частичные отказы
3. Ненадежная сеть
4. Ненадежные часы и проблема времени
5. CAP/PACELC теорема



Контроль и координация

Базовые задачи:

1. Хранение информации о составе кластера
2. Управление конфигурацией узлов
3. Детекция сбойных узлов
4. Выбор лидера
5. Service discovery



Инструменты и алгоритмы

Инструменты (топ-3 самых популярных)

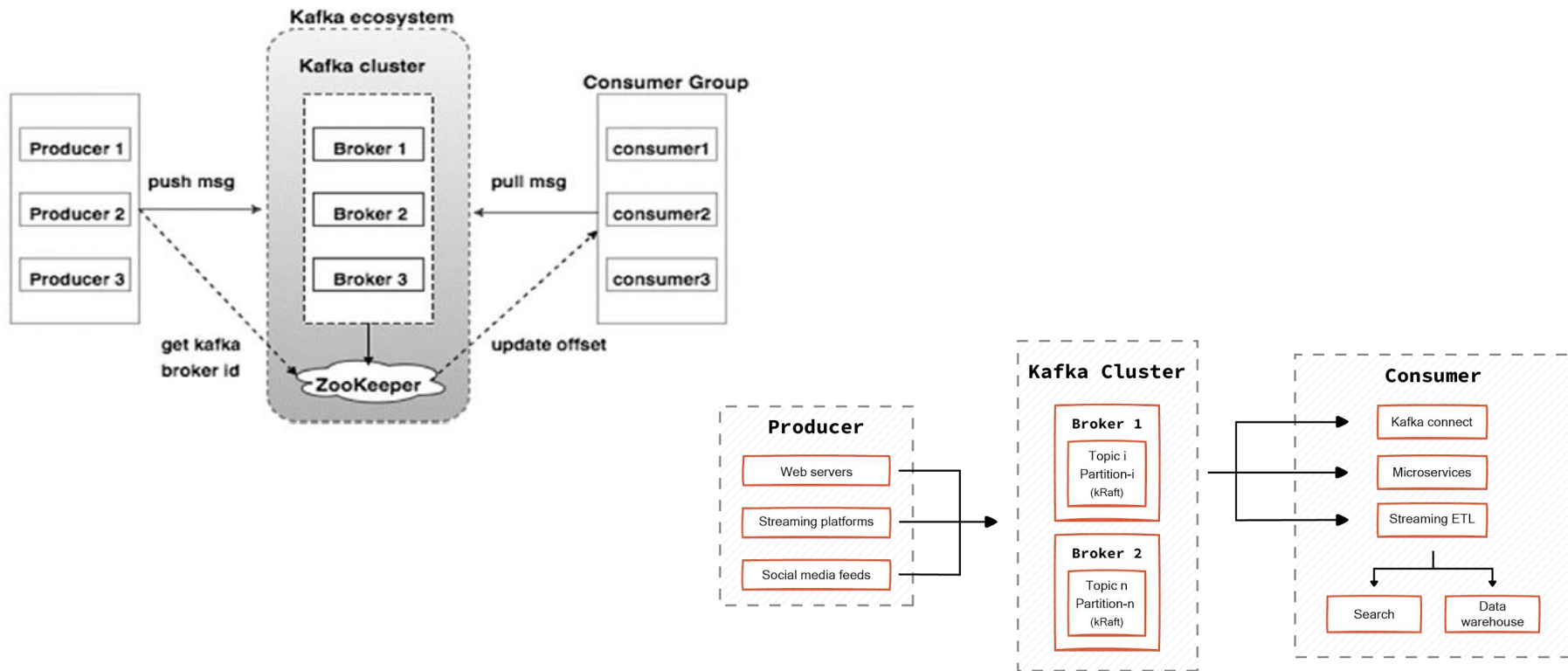
1. etcd
2. consul
3. Zookeeper

Алгоритмы (топ-3 самых популярных)

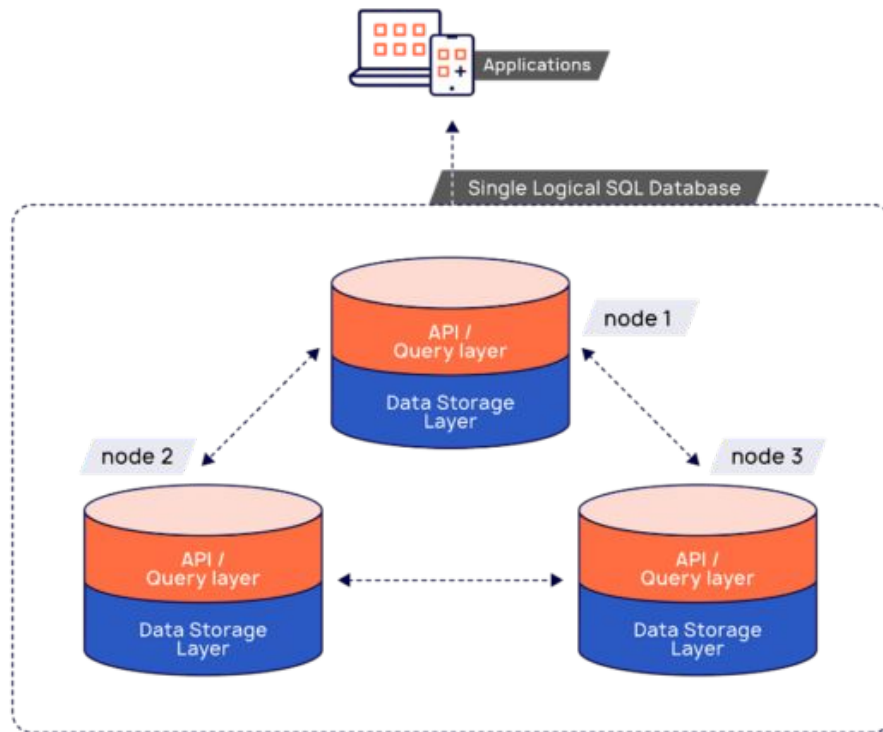
1. Paxos
2. Raft
3. Accord



Распределенный брокер



Распределенная БД





(высокая?) Доступность

How much data can you afford to recreate or lose?

RPO vs RTO

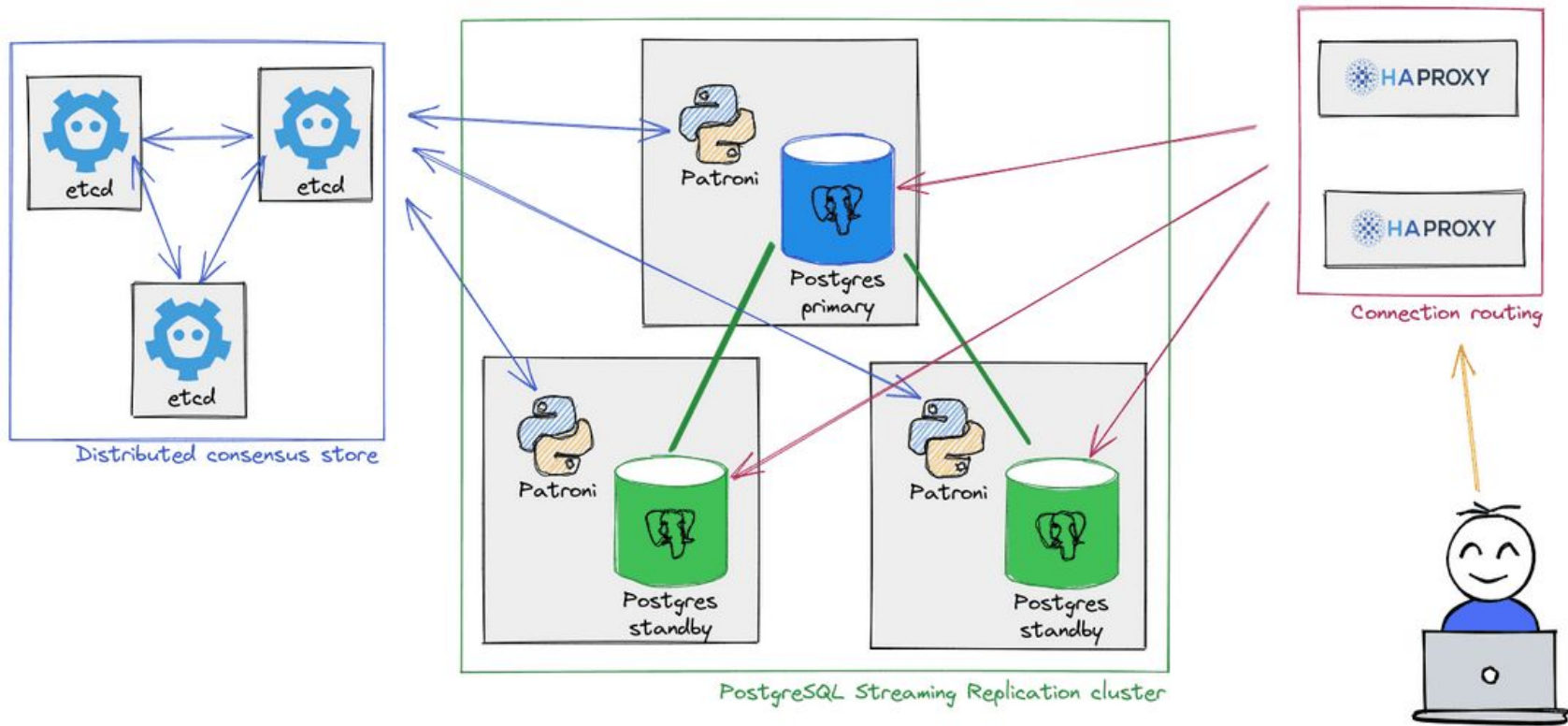
How quickly must you recover?
What is the cost of downtime?



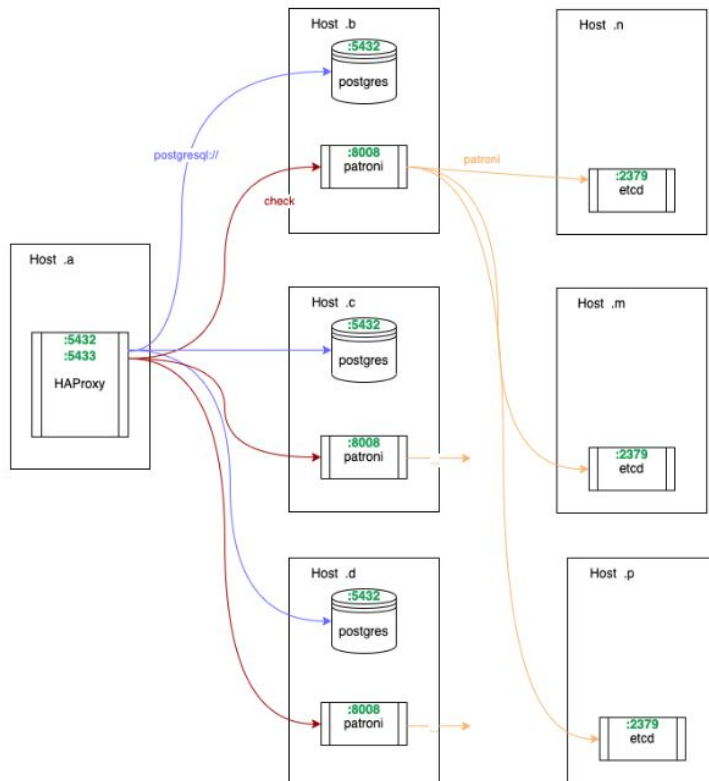
1. Ничего не делаем
2. Делаем бэкапы
3. Делаем репликацию
4. Делаем автоматическое переключение
5. Делаем устойчивость к отказам



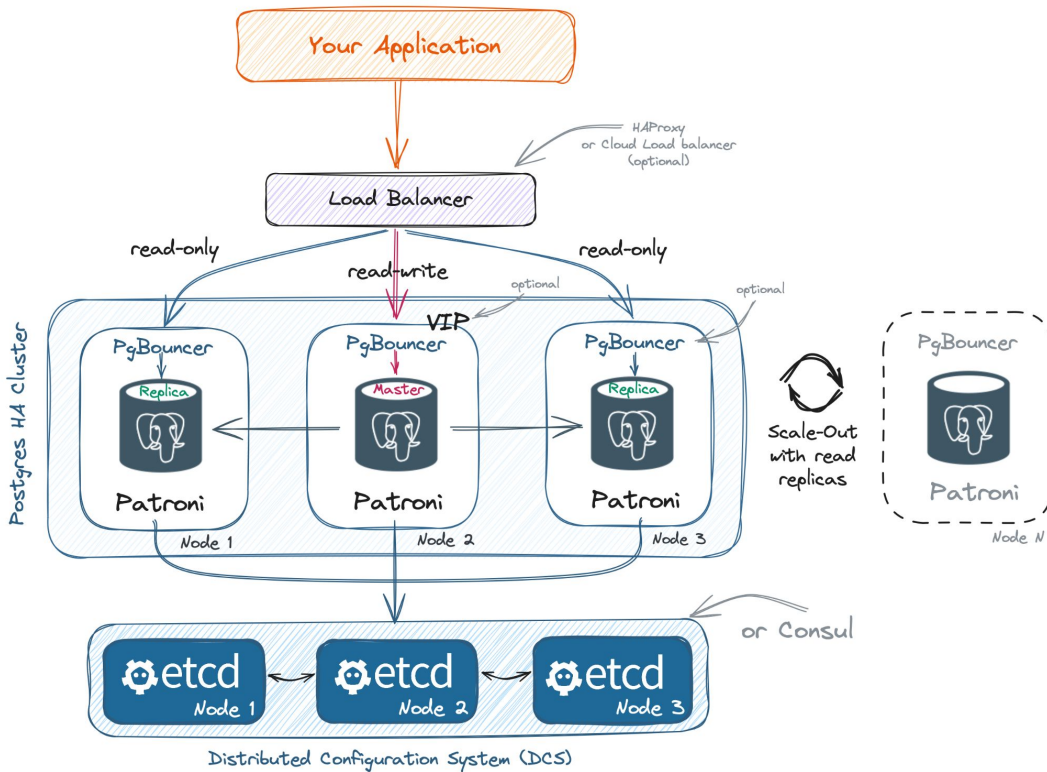
HA PostgreSQL



Patroni + etcd/consul

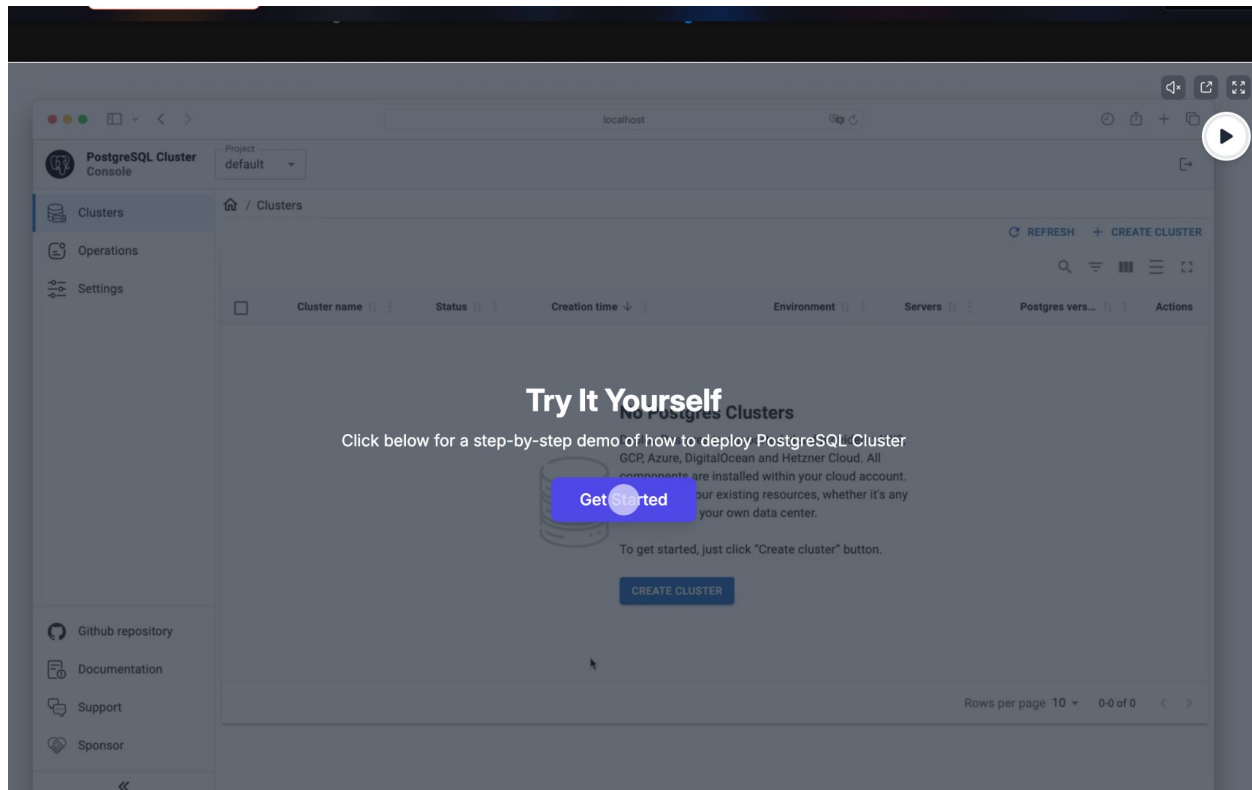


Patroni + etcd/consul





Autobase





Yandex Managed PG (Cloud)

Managed Service for PostgreSQL / Кластеры / postgresql102 ☆									
Переключить мастер Создать хост									
Хосты									
FQDN хоста	Роль ↑↓	Состояние ↑↓	ЦПУ ↑↓	ОЗУ ↑↓	Диск ↑↓	Источник репликации	Подсеть	Публичный доступ	Зона доступности
rc1b-mg0548smd69kv54c	REPLICA	ALIVE	9 %	8 %	4 %	-	default-ru-central1-b	есть	ru-central1-b ...
rc1d-lg25i4nicjuji312	MASTER	ALIVE	6 %	9 %	4 %	-	default-ru-central1-d	есть	ru-central1-d ...



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ