



NATIONAL RESEARCH
UNIVERSITY

РЕАЛИЗАЦИЯ В CLICKHOUSE ПРОТОКОЛА POSTGRESQL

Программный проект

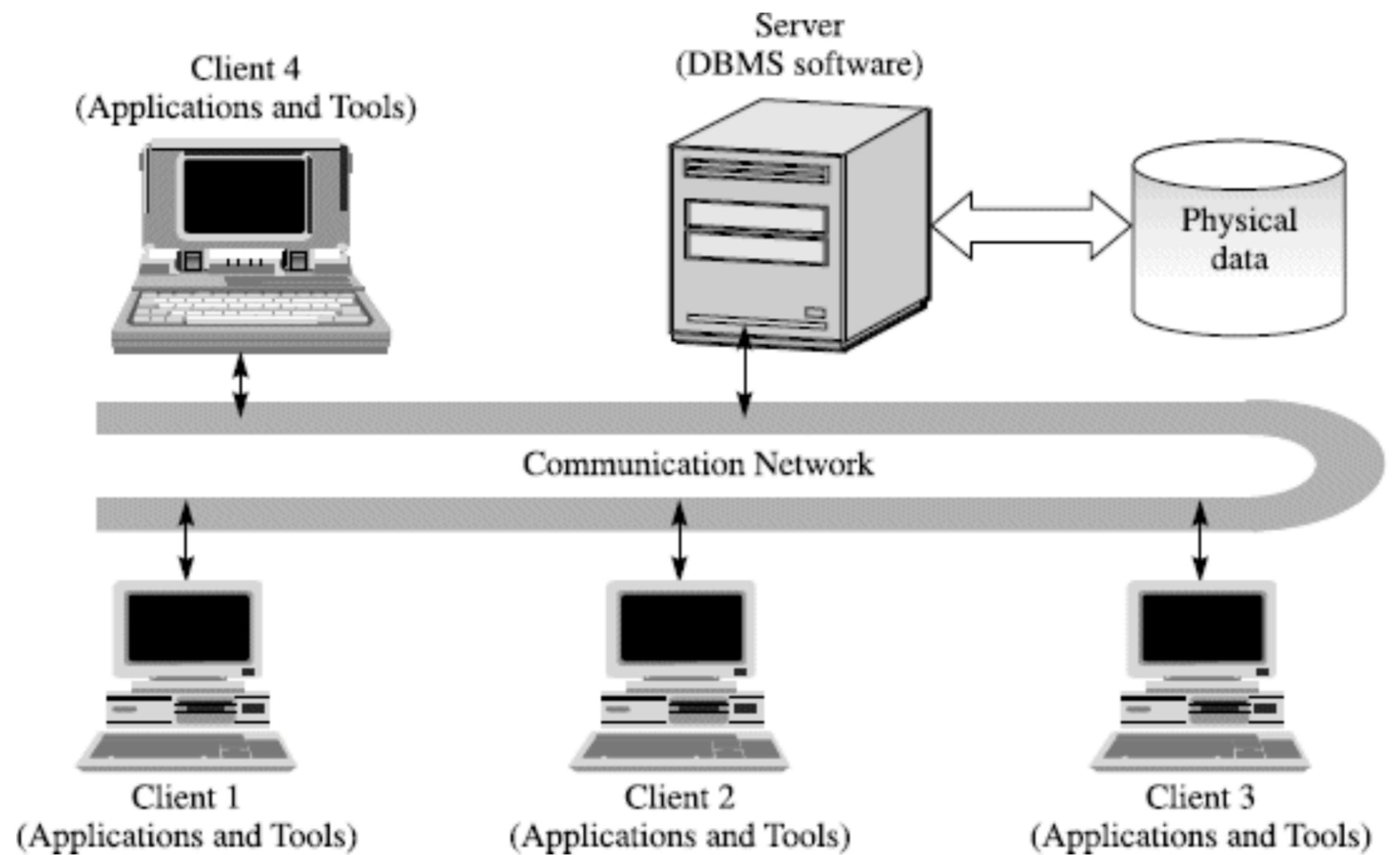
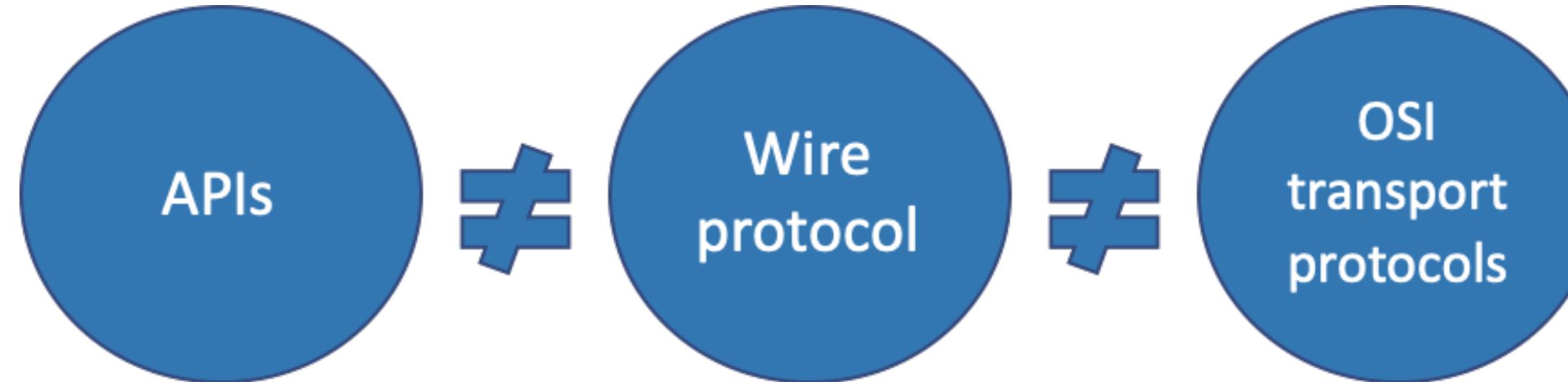
Студент: Элбакян Мовсес Андрапникович
Руководитель: Миловидов Алексей Николаевич

Национальный Исследовательский Институт «Высшая Школа Экономики»
Москва, Российская Федерация
2020



АКТУАЛЬНОСТЬ

Wire protocol – протокол, который описывает способ взаимодействия одного или нескольких приложений внутри сети.



Распространенная архитектура баз данных
«клиент-сервер»

Источник изображения: S. K. Singh, Database Systems: Concepts, Design and Applications, 2nd ed. Pearson, 2011



АКТУАЛЬНОСТЬ

Wire protocol для БД

Создание клиентов

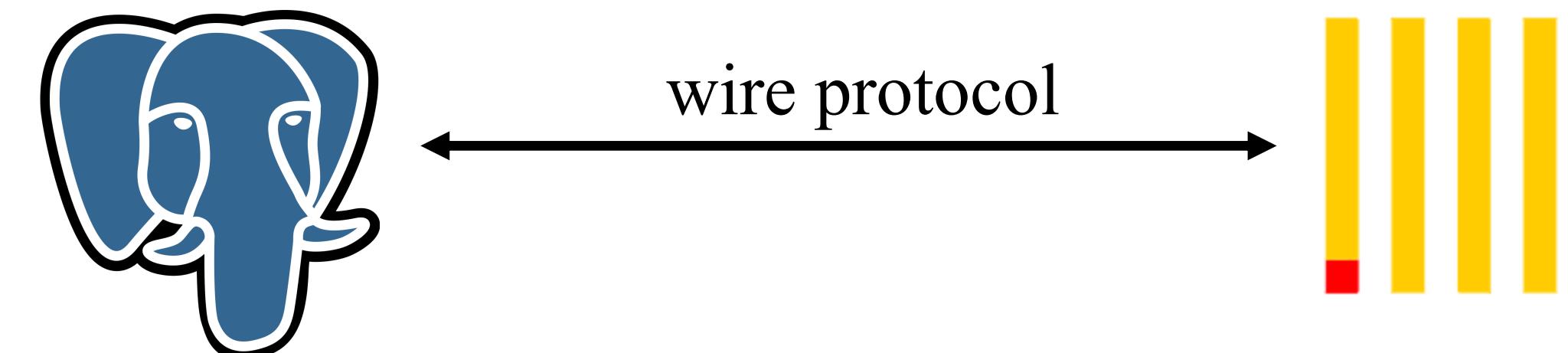
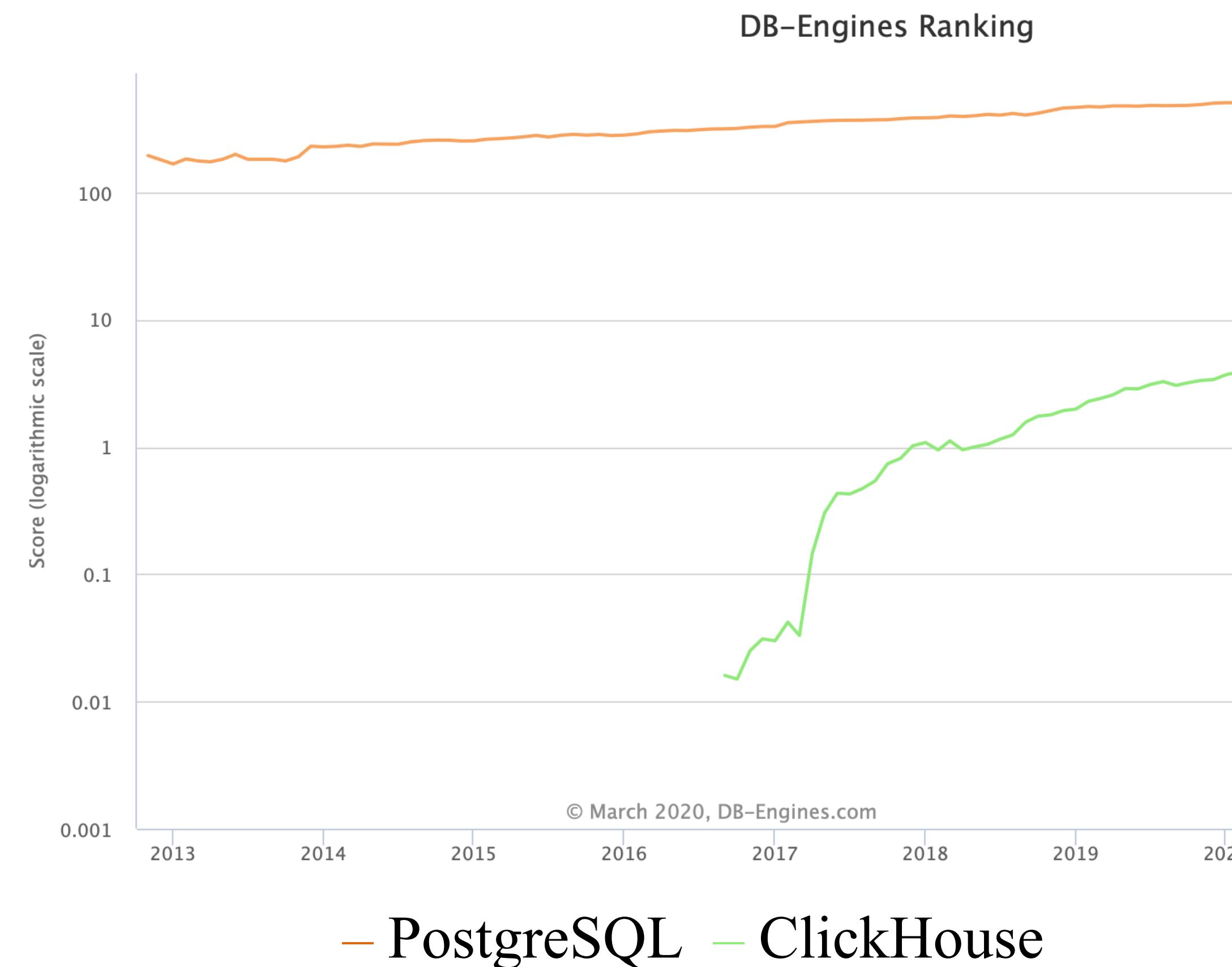
- Различные платформы
- Различные языки программирования

Создание серверов

- Интеграции с другими базами данных



АКТУАЛЬНОСТЬ



- Уменьшение стека технологий для разработчиков. (Меньше библиотек – больше функциональности)
- Популяризация ClickHouse. (PostgreSQL на 4-м месте по популярности, ClickHouse на 67-м по версии db-engines.com)
- Результаты могут использованы в исследованиях унификации взаимодействия между разными базами данных.



ЦЕЛИ И ЗАДАЧИ

Цель: имплементация *wire protocol'a PostgreSQL* в аналитической базе данных *ClickHouse*.

Задачи:

1. Сделать базу данных ClickHouse максимально совместимой с как можно большим количеством PostgreSQL клиентов.
2. Оптимальность кода и отсутствие просадки скорости в области подготовки запроса и отправки результатов.
3. Реализация интеграционного тестирования для избегания будущих багов при изменении кодовой базы.



РЕЛЕВАНТНЫЕ РАБОТЫ

- Имплементация от разработчиков PostgreSQL (стек технологий отличается, непортируемые решения, долгое исследование кодовой базы)
- Аналогичные реализации в ClickHouse (нативный протокол, MySQL wire protocol: различия довольно большие)



ВЫБОР МЕТОДОВ РЕШЕНИЯ. POSTGRESQL WIRE PROTOCOL

Неподдержанные возможности в имплементации:

- Шифрование трафика по GSSAPI
- Методы аутентификации: GSSAPIAuth, KerberosV5 (deprecated) и SCRAM-SHA-256-PLUS
- Виды запросов: Extended query, Function call (deprecated) и Copy query



ВЫБОР МЕТОДОВ РЕШЕНИЯ. СТЕК ТЕХНОЛОГИЙ

- **Языки:** C++, Python2, Java
- **Система сборки:** CMake
- **Работа с сетью, шифрованием и хешированием:** библиотека Poco
- **VCS:** Git
- **Анализ трафика:** tcpflow, socat
- **Клиенты для интеграционного тестирования:** psql, Java client, psycopg2

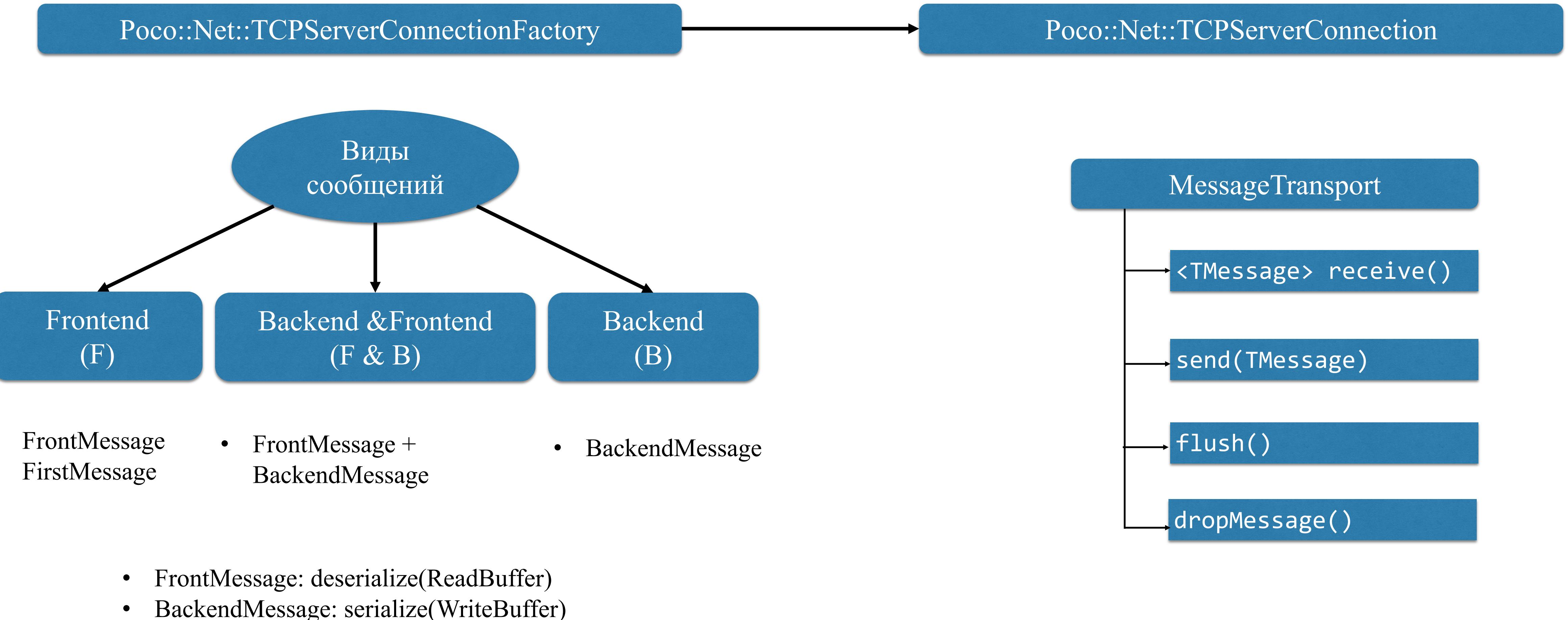


ОБЗОР ПРОТОКОЛА

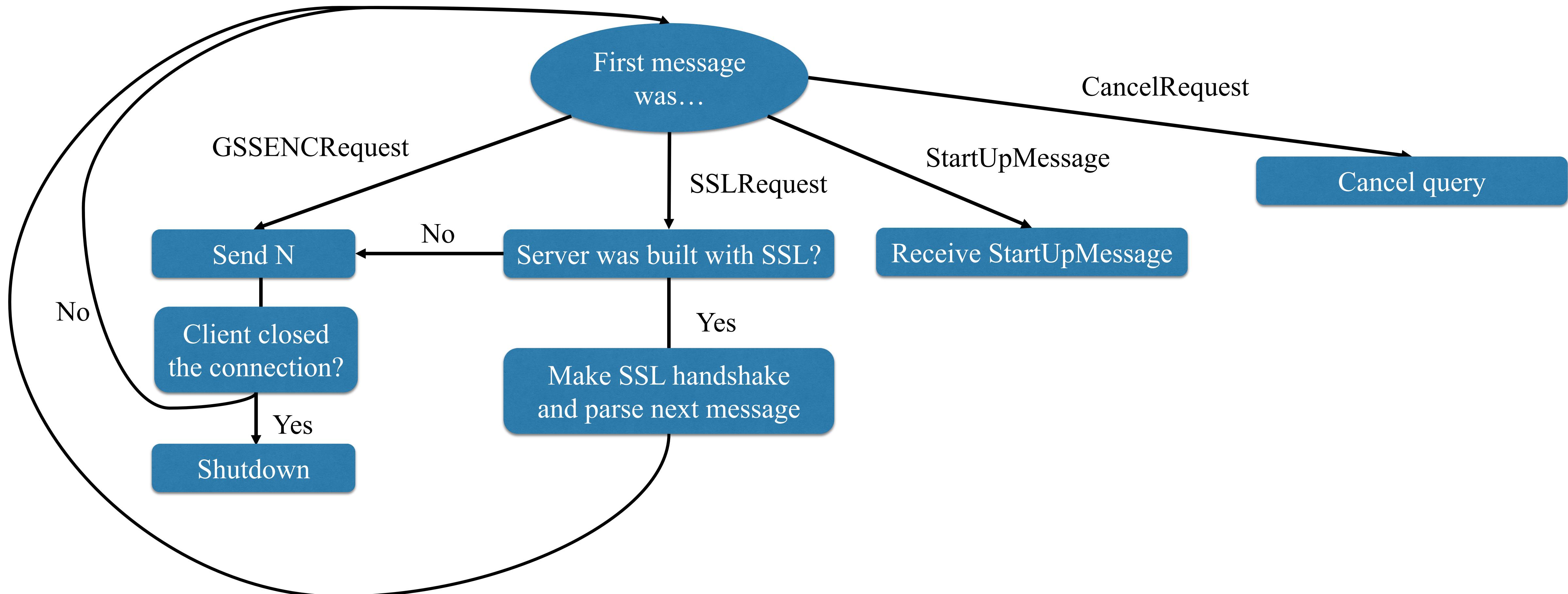
1. Установка соединения (метод шифрования трафика, аутентификация, передача дополнительных параметров)
2. Выполнение запросов (Simple query, Extended query, Copy query)
3. Закрытие соединения



УСТАНОВКА СОЕДИНЕНИЯ



УСТАНОВКА СОЕДИНЕНИЯ



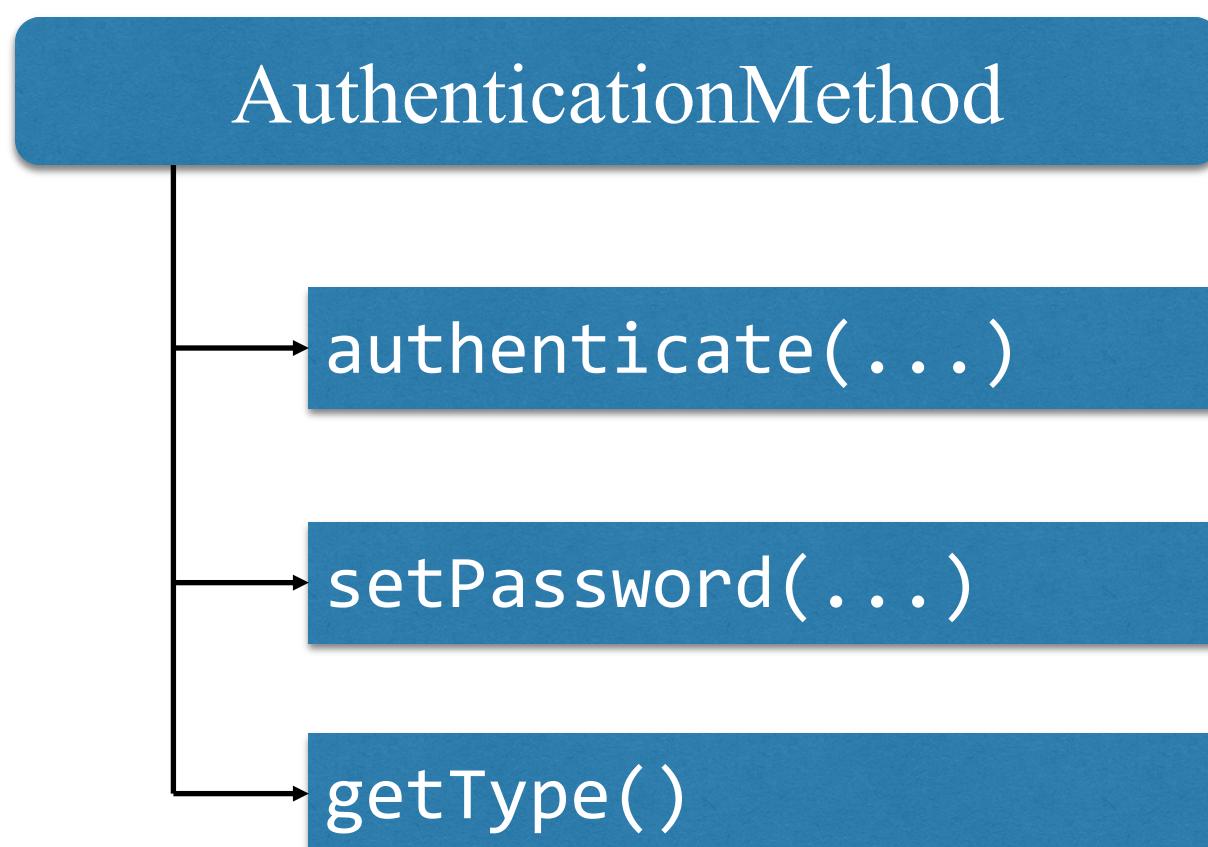
StartUpMessage содержит информацию о версии протокола, имени пользователя и других опциональных параметрах (среди которых название базы данных)



УСТАНОВКА СОЕДИНЕНИЯ. АУТЕНТИФИКАЦИЯ

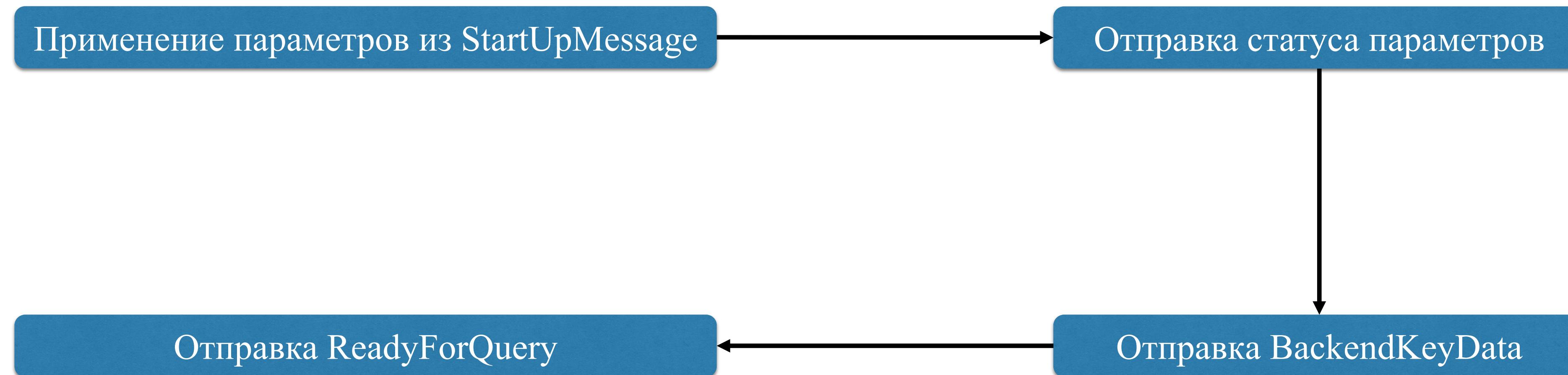
Поддержанные методы аутентификации:

- Без пароля
- Незашифрованный текстовый пароль
- MD5 с солью (вычисляется по формуле ниже)
`concat('md5', md5(concat(md5(concat(password, username)), random-salt)))`
- SCRAM-SHA-256 (RFC 7677)



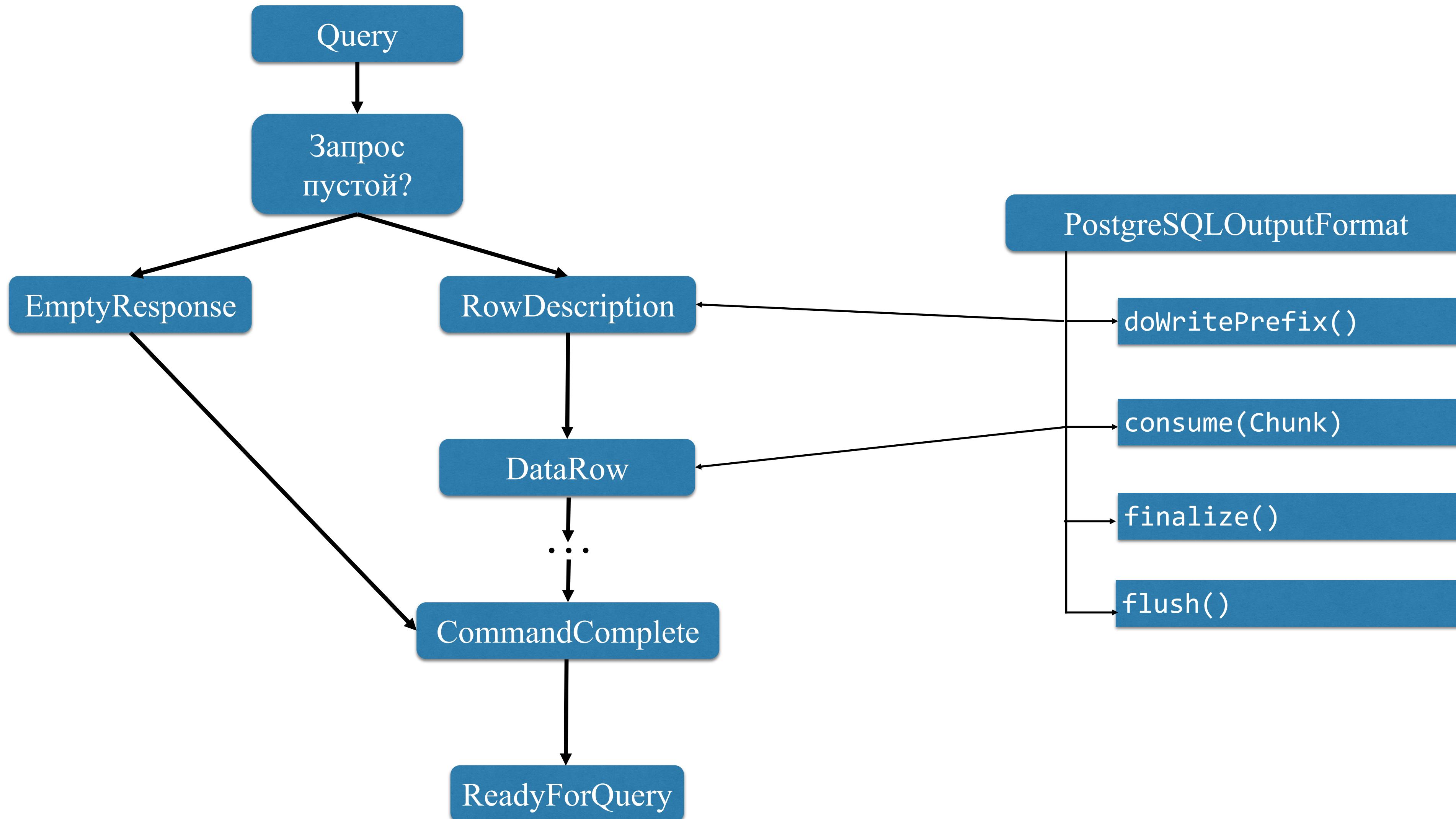


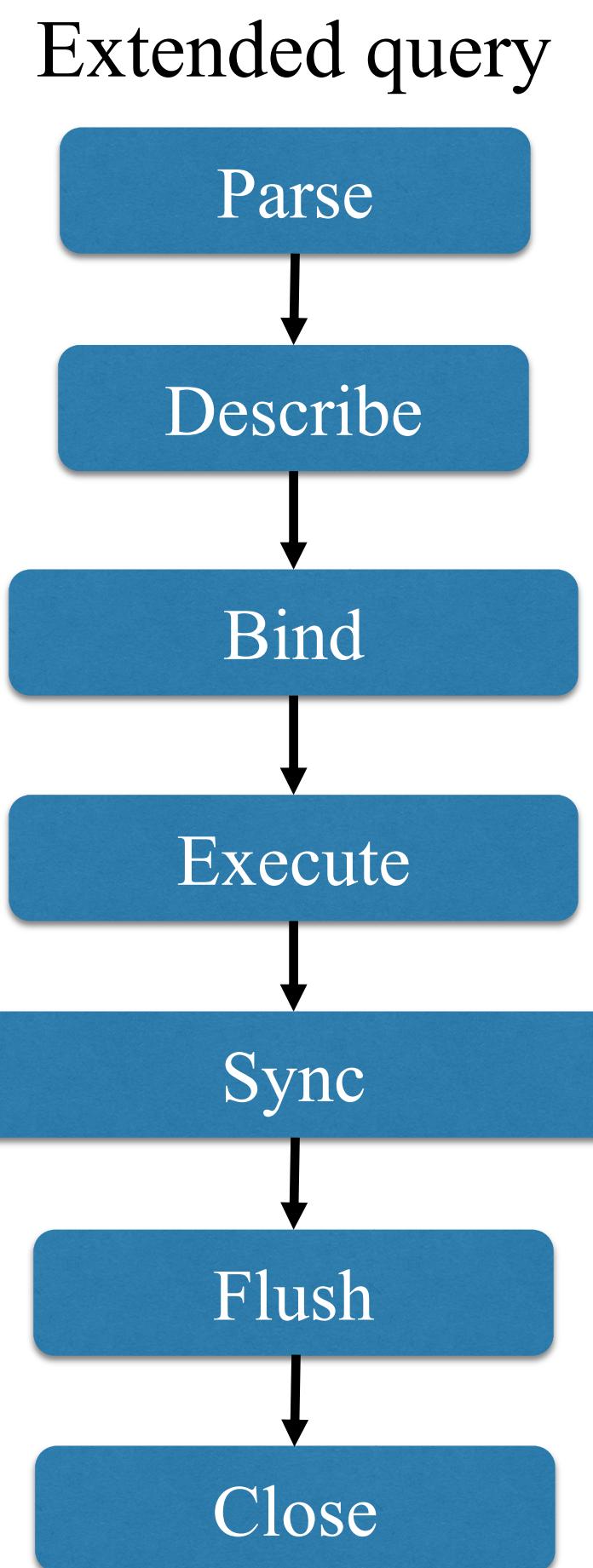
УСТАНОВКА СОЕДИНЕНИЯ. ФИНАЛЬНАЯ СТАДИЯ



BackendKeyData содержит в себе идентификатор хендлера, в котором выполняется запрос, а также случайное число-секрет.

ОБРАБОТКА ЗАПРОСОВ. SIMPLE QUERY





Copy query

В зависимости от запроса производит трансфер данных с сервера на клиент или наоборот данных из таблицы. Данные передаются в бинарном формате пачками.

Проблема: Запрос передается в виде выражения на диалекте PostgreSQL, которое требуется дополнительно распарсить. Это континтуитивно с точки зрения использования ClickHouse.

Проблема: На этапе Describe мы еще не знаем типов возвращаемых данных, но уже надо вернуть их клиенту.



ЗАКЛЮЧЕНИЕ

- Был реализован PostgreSQL wire protocol в аналитической базе данных ClickHouse.
- В результате работы различные клиенты PostgreSQL могут взаимодействовать с ClickHouse.
- Интеграционное тестирование позволит избежать ошибок при изменении кодовой базы ClickHouse.
- Реализацию протокола можно легко дополнять и развивать.