



Общество с ограниченной ответственностью «Ребреин»

ИНН 7727409582, ОГРН 1197746106161

Адрес: 123056, город Москва, Большая Грузинская ул, д. 36а стр. 5а, офис 13

Утверждено

Приказом № ПР-1 от 17.06.2025 г.

Генеральный директор

 Фролкина Е.А.
«17» июня 2025 г.

ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА
– ПРОГРАММА ПОВЫШЕНИЯ КВАЛИФИКАЦИИ
**«ВЫСОКОПРОИЗВОДИТЕЛЬНАЯ АНАЛИТИКА С ИСПОЛЬЗОВАНИЕМ
CLICKHOUSE»**

Срок реализации: 2 недели (14 дней)

Количество часов: 26 акад. ч.

Форма обучения: заочная форма

Формат обучения: с применением
исключительно дистанционных технологий

Возраст обучающихся: для лиц старше 17
лет, имеющих или получающих среднее
профессиональное и (или) высшее
образование

Москва, 2025 г.

1. ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Настоящая дополнительная профессиональная программа – программа повышения квалификации «Высокопроизводительная аналитика с использованием ClickHouse» (далее – Программа) разработана в соответствии с:

- Федеральным законом от 29 декабря 2012 г. № 273-ФЗ «Об образовании в Российской Федерации»;
- Приказом Министерства образования и науки РФ от 1 июля 2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам»;
- Постановлением Правительства РФ от 11.10.2023 № 1678 «Об утверждении Правил применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ»;
- Профессиональным стандартом 06.022 «Системный аналитик», утвержден приказом Министерства труда и социальной защиты Российской Федерации от 27 апреля 2023 г. № 367н ;
- Профессиональным стандартом 06.042 «Специалист по большим данным», утвержден приказом Министерства труда и социальной защиты Российской Федерации от 5 августа 2020 г. № 59174;
- ФГОС высшего образования – бакалавриат по направлению 09.03.02 Информационные системы и технологии, утв. приказом Минобрнауки России от 19 сентября 2017 г. № 926.
- Локальными нормативными актами ООО «Ребреин».

В данной программе учтены основные идеи формирования универсальных учебных действий обучающихся и соблюдена преемственность с программами высшего и/или среднего профессионального образования.

Направленность программы: программа имеет техническую направленность.

Адресат: программа предназначена для технических специалистов, ответственных за развёртывание, настройку и поддержку ClickHouse в продуктивной среде. Она будет особенно полезна:

- системным администраторам и DevOps-инженерам, которые внедряют и сопровождают ClickHouse в высоконагруженных системах;
- инженерам по данным и SRE-специалистам, обеспечивающим отказоустойчивость, безопасность и мониторинг аналитических хранилищ;
- backend-разработчикам и техническим архитекторам, участвующим в проектировании архитектуры систем хранения и обработки данных;
- инженерам технической поддержки, обслуживающим отказоустойчивые и реплицированные инсталляции ClickHouse.

Требования к входным знаниям обучающегося: обучающемуся необходимы базовые знания сетевых протоколов и модели OSI, базовые знания Docker, языков программирования высокого уровня (напр. Python, Golang)

Актуальность реализации:

ClickHouse — это база данных, которая специально оптимизирована для аналитических запросов и обработки больших объемов данных. Она предназначена для работы с

огромными массивами информации, позволяя выполнять быстрые запросы и агрегировать данные для анализа. Основное преимущество ClickHouse — это скорость выполнения запросов, что делает его популярным выбором для работы с данными, требующими аналитики в реальном времени, такими как журналы посещений веб-сайтов, метрики приложений и аналитика рекламы. ClickHouse входит в состав реестра отечественного программного обеспечения. А значит, его можно использовать в тех проектах, где присутствуют жесткие ограничения на использование импортного ПО.

Отличительные особенности программы:

- Программа сфокусирована на развёртывании, настройке и администрировании ClickHouse — от базовой конфигурации до обеспечения отказоустойчивости и безопасности.
- Уделяется внимание реальной эксплуатации ClickHouse в продакшене: мониторинг, логирование, настройка пользователей и квот, SSL-подключения.
- Освещаются механизмы масштабирования и отказоустойчивости: шардирование, репликация с использованием Zookeeper, восстановление после сбоев, создание бэкапов.
- Практическая часть построена вокруг типичных задач DevOps- и SRE-команд, работающих с ClickHouse как частью распределённой инфраструктуры.
- В рамках курса отрабатываются навыки интеграции справочников, настройки безопасности и мониторинга, что критически важно для работы с чувствительными и критичными данными.

Объем и срок освоения программы: 26 академ. ч. в течение 2 недель (14 дней).

Доступ к материалам Программы у обучающихся остаётся и после окончания периода обучения. Это позволяет повторять изученный материал в удобное время, восполнять пробелы в знаниях, а также возвращаться к практическим заданиям при решении рабочих задач. Такой формат способствует более глубокому закреплению навыков и поддерживает профессиональное развитие выпускников даже после завершения обучения.

Выдаваемый документ о квалификации: удостоверение о повышении квалификации и/или сертификат об успешном освоении программы.

Цели и задачи программы:

Сформировать у обучающихся практические навыки установки, настройки, сопровождения и масштабирования системы аналитической обработки данных ClickHouse, обеспечивая её отказоустойчивость, безопасность и производительность в условиях промышленной эксплуатации.

Программа направлена на решение следующих основных задач:

Обучающие:

- Познакомить обучающихся с архитектурой и компонентами ClickHouse.
- Обучить основам конфигурирования и администрирования ClickHouse.
- Сформировать умения обеспечивать безопасность, логирование, бэкапы и восстановление базы.
- Научить организовывать отказоустойчивые инсталляции с использованием репликации, шардирования и Zookeeper.
- Ознакомить с методами мониторинга и диагностики состояния системы.

Развивающие:

- Развивать системное мышление при проектировании архитектуры распределённых систем.
- Формировать навыки анализа и решения типовых эксплуатационных задач в высоконагруженной среде.
- Стимулировать самостоятельный поиск и использование документации и best practices.

Воспитательные:

- Воспитывать ответственное отношение к качеству и безопасности настройки инфраструктуры.
- Формировать культуру аккуратной, прозрачной и воспроизводимой работы с критичными компонентами.
- Способствовать развитию командной этики и навыков взаимодействия в технических командах DevOps/SRE.

Планируемые результаты:

Обучающийся, успешно освоивший программу повышения квалификации, сможет демонстрировать следующие знания, умения и навыки:

Знания:

- архитектура и принципы работы ClickHouse как колоночной СУБД;
- основные параметры конфигурации ClickHouse и их назначение;
- механизмы шардирования, репликации и обеспечения отказоустойчивости;
- методы обеспечения безопасности: управление пользователями, профилями и квотами, настройка SSL;
- инструменты логирования, мониторинга и диагностики состояния системы;
- подходы к резервному копированию и восстановлению ClickHouse;
- назначение и возможности Zookeeper в распределённых инсталляциях;
- способы интеграции справочников в ClickHouse.

Умения:

- устанавливать и настраивать ClickHouse на сервере;
- конфигурировать пользователей, права доступа, профили и квоты;
- организовывать безопасные подключения с использованием SSL;
- настраивать логирование и мониторинг состояния нод ClickHouse;
- осуществлять шардирование и репликацию с использованием Zookeeper;
- проводить резервное копирование и восстановление инстанса ClickHouse;
- выявлять и устранять сбои в работе сервиса;
- подключать внешние справочники (например, на примере gerpip).

Навыки:

- уверенная работа с ClickHouse в командной строке и конфигурационных файлах;
- навык организации отказоустойчивых аналитических систем;
- опыт настройки и сопровождения высоконагруженных систем хранения и обработки данных;
- умение решать типовые эксплуатационные задачи (от настройки до восстановления после сбоев);
- аккуратная работа с конфигурациями, логами и средствами безопасности в продакшене.

Перечень профессиональных компетенций, на получение которых направлено обучение:

На основе профстандарта 06.022 «Системный аналитик»:

- A/05.4 — Оформление проектной и эксплуатационной документации в рамках обследования текущей ситуации, проектирования и сдачи-приемки Системы.
- A/08.4 — Выполнение предварительных испытаний разработанного или интегрированного элемента поставки на соответствие разработанным требованиям к нему.
- B/07.5 — Сопровождение испытаний развернутой подсистемы или очереди Системы на соответствие разработанным требованиям к подсистеме.
- C/06.6 — Методическое сопровождение испытаний Системы

На основе профстандарта 06.042 «Специалист по большим данным»:

- A/02.6 — Планирование и организация аналитических работ с использованием технологий больших данных.
- A/03.6 — Подготовка данных для проведения аналитических работ по исследованию больших данных.
- A/04.6 — Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика.

Таким образом, в результате освоения программы у обучающихся формируются следующие профессиональные компетенции:

- ОПК-2. Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;
- ОПК-5. Способен инсталлировать программное и аппаратное обеспечение для информационных и автоматизированных систем.

Организационно-педагогические условия реализации программы дополнительного профессионального образования

Язык реализации образовательной программы: обучение проводится на русском языке.

Форма обучения: заочная форма.

Особенности реализации программы: программа реализуется с использованием электронного обучения и исключительно дистанционных образовательных технологий.

Условия набора: на обучение принимаются все желающие лица, оплатившие обучение и заключившие договор об образовании. Обучение проходит в индивидуальном формате без формирования учебных групп. Обучающийся самостоятельно определяет время освоения Программы.

Формы проведения занятий:

- занятия в текстовом формате;
- практическая работа;
- самостоятельная работа с литературой;
- индивидуальные вопросы.

Материально-техническое оснащение

Материальное обеспечение программы

Занятия проводятся в системе дистанционного обучения «Rebrain». Каждый обучающийся и педагог оснащены доступом к системе дистанционного обучения: <https://rebrainme.com/>.

У педагога дополнительного профессионального образования имеется необходимое оборудование средства для реализации программы: ноутбук с подключением к интернету, программное обеспечение.

Методическое обеспечение программы

Программа обеспечена:

- учебно-методическими материалами (текстовые занятия, полезными материалами);
- практическими заданиями.

Кадровое обеспечение:

К реализации программы в качестве педагогов дополнительного образования допускаются лица:

- 1) отвечающее одному из требований:
 - а) имеющее высшее образование или среднее профессиональное образование в рамках укрупненных групп специальностей и направлений подготовки высшего образования и специальностей среднего профессионального образования «Образование и педагогические науки»;
 - б) имеющее высшее образование либо среднее профессиональное образование в рамках иных укрупненных групп специальностей и направлений подготовки высшего образования и специальностей среднего профессионального образования при условии его соответствия дополнительной общеобразовательной общеразвивающей программе, реализуемой ООО «Ребреин», и получение при необходимости дополнительного профессионального образования педагогической направленности;
 - в) успешно прошедшее промежуточной аттестации не менее чем за два года обучения по образовательным программам высшего образования по специальностям и направлениям подготовки, соответствующей направленности дополнительной общеобразовательной общеразвивающей программе;
- 2) не имеющее ограничений на занятие педагогической деятельностью, установленных законодательством Российской Федерации;
- 3) прошедшее обязательный предварительный (при поступлении на работу) и периодические медицинские осмотры (обследования), а также внеочередные медицинские осмотры (обследования) в порядке, установленном законодательством Российской Федерации.

Реализация Программы также возможна лицами, привлекаемыми на условиях гражданско-правового договора в соответствии с действующим законодательством РФ.

2. УЧЕБНЫЙ ПЛАН

№ п/п	Наименование модуля	Количество часов			Формы контроля / аттестация
		Всего	Теория	Практика	
1	Модуль 1 “Онбординг”	1	0,5	0,5	Входное тестирование
2	Модуль 2 “ClickHouse Base”	22	3	19	Практическое задание

3	Итоговая аттестация	3		3	Итоговое практическое задание
---	---------------------	---	--	---	-------------------------------

3. КАЛЕНДАРНЫЙ УЧЕБНЫЙ ГРАФИК

№ п/п	Наименование модуля	1 неделя	2 неделя
1	Модуль 1 “Онбординг”	1	
2	Модуль 2 “ClickHouse Base”	12	10
3	Итоговая аттестация		3 А

4. РАБОЧАЯ ПРОГРАММА

Модуль 1. Онбординг

Теория 0,5 академ. ч. Практика 0,5 академ. ч.

Модуль состоит из следующих блоков изучения тем:

Тема 1. Онбординг

В теме обучающемуся предоставляется вводный конспект, содержащий общую информацию о программе, структуре курса, форматах взаимодействия с материалами и ожидаемых результатах обучения.

Предусмотрено прохождение входного тестирования, включающего 6 вопросов, направленных на закрепление информации из онбординга. В рамках блока обучающийся выполняет задание по целеполаганию: формулирует свою цель прохождения программы, указывает желаемые навыки по окончании обучения.

Модуль 2. ClickHouse Base

Теория 3 академ. ч. Практика 19 академ. ч.

Модуль состоит из следующих тем:

Тема 1. Установка Clickhouse

Содержание: История зарождения ClickHouse. Возможности программы. Колоночная СУБД. Хранение данных. Сжатие данных. Особенности работы с данными. Установка ClickHouse. Практическое задание.

Тема 2. Базовая конфигурация

Содержание: конфигурационные файлы. Ключевые параметры: http(s)_port, tcp_port, listen_host и др. Практическое задание.

Тема 3. Настройка логирования

Содержание: Основное и дополнительное логирование. Секция logger. Практическое задание.

Тема 4. Настройка пользователей

Содержание: Создание пользователей. Структура конфигурации пользователей. Практическое задание.

Тема 5. Настройка профилей

Содержание: Примеры конфигурации профилей. Настройки <max_memory_usage> <max_rows_to_read> <max_bytes_to_read> и др. Практическое задание.

Тема 6. Настройка квот

Содержание: Примеры конфигурации профилей. Квоты <duration>, <queries> и др. Практическое задание.

Тема 7. Подключение SSL

Содержание: Конфигурация портов. Подключение сертификатов. Проверка клиента. Практическое задание.

Тема 8. Мониторинг ноды Clickhouse

Содержание: Метрики. Использование систем мониторинга. Проверка доступности сервера. Практическое задание.

Тема 9. Шардирование в Clickhouse

Содержание: Установка clickhouse на все машины кластера. Конфигурация ClickHouse кластера. Тестирование кластера ClickHouse. Создание таблиц в кластере. Практическое задание.

Тема 10. Установка Zookeeper

Содержание: Установка Java. Установка и конфигурация Zookeeper. Создание Zookeeper-сервиса. Практическое задание.

Тема 11. Настройка репликации в Clickhouse

Содержание: Конфигурация ClickHouse кластера. Создание реплицированных таблиц. Создание Distributed таблиц. Вставка данных в реплицируемый кластер. Практическое задание.

Тема 12. Восстановление после сбоя

Содержание: Отвал реплики, шарда, ноды, реплики в реплицированном кластере в ClickHouse. Практическое задание.

Тема 13. Подключение справочников на примере gerip

Содержание: Типы и конфигурация словарей. Источники внешних словарей. Локальный файл. СУБД. Хранение словарей в памяти. Практическое задание.

Тема 14. Бэкапы в ClickHouse

Содержание: Заморозка партиций. Копирование и восстановление данных. Практическое задание.

Итоговая аттестация

Блок посвящён выполнению финального практического задания без предварительного теоретического блока.

Модуль направлен на формирование у обучающихся базовых и продвинутых навыков работы с системой ClickHouse. Каждый блок модуля включает текстовое занятие с теоретическим материалом и пошаговыми инструкциями, после изучения которых предлагается практическое задание.

Практические задания рассчитаны на 2 академических часа. Выполнение заданий предполагает отправку решения на проверку через личный кабинет обучающегося.

Критерии оценки прописаны в описании к каждому заданию. В случае корректного выполнения выставляется зачёт. Если работа содержит ошибки, задание возвращается на доработку. При повторной неудачной попытке (после двух доработок) обучающийся получает «незачёт».

5. МЕТОДИЧЕСКИЕ И ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Программа обеспечена системой дистанционного обучения <https://rebrainme.com/>.

Педагогические технологии:

- технология дифференцированного обучения;
- технология разноуровневого обучения;
- технология развивающего обучения;
- технология проблемного обучения;
- технология дистанционного обучения.

Методы обучения:

- словесный, наглядный практический;
- объяснительно – иллюстративный;
- частично-поисковый, исследовательский проблемный;
- игровой, дискуссионный.

Электронно-библиотечные ресурсы и системы, информационно-справочные системы:

1. Научная электронная библиотека eLIBRARY.RU.
2. Собственные учебные материалы: <https://my.rebrainme.com/course/clickhouse-base>
3. Официальная документация ClickHouse: <https://clickhouse.com/docs/ru/about-us/faq>
4. Официальный сайт Apache ZooKeeper: <https://zookeeper.apache.org/>

6. ОЦЕНКА КАЧЕСТВА ОСВОЕНИЯ ПРОГРАММЫ

Оценочные материалы:

Для отслеживания результатов освоения программы среди слушателей проводится текущий контроль и итоговое оценивание.

Текущий контроль

Осуществление текущего контроля проводится после занятий в виде написания практических заданий или тестирований. Тематика и условия выполнения практических заданий расписаны в личном кабинете обучающегося в СДО. Педагог проверяет решение и принимает решение о принятии решения (зачет), о необходимости доработать решение или о незачете. Если промежуточный контроль представлен в виде тестирования, подсчет верных ответов и выставление оценки «зачёт» и «незачёт» происходят в автоматическом решиме в СДО.

Правила выполнения заданий.

1. После нажатия кнопки «Начать выполнение» для вас будет подготовлено окружение и предоставлены необходимые доступы.
2. Также вам будут выданы переменные (если они будут нужны), которые в задании указаны в фигурных скобках, — их надо будет подставить при выполнении задания.

3. После выполнения всех пунктов задания нажмите кнопку «Отправить на проверку», и в течение ближайших 3-5 минут скрипт проверит выполнение всех условий и выставит вам оценку.
4. В случае, если вы что-то забыли, можно исправить ошибки и отправить на проверку повторно.
5. Также, если вы успешно сдали задание, но у вас остались вопросы — вы всегда сможете задать их куратору после проверки или в чате в любое удобное для вас время.

Пример практических заданий:

Задание №1.

Установите clickhouse на сервер из официальных репозиториев.

Все указанные изменения проводите в файле config.xml, не создавая дополнительных файлов.
Задайте уровень логирования information в конфигурации ClickHouse.

Укажите размер лог файлов в 3000М.

Проверьте логи ClickHouse и убедитесь, что у вас поменялся уровень логирования.

Включите логирование запросов (query_log) - все данные должны записываться в базу данных system и таблицу queries. Выполните тестовые запросы как в предпоследнем примере, чтобы создалась таблица queries.

Если все работает корректно, тогда отправляйте задание на проверку.

Задание №2.

Установите clickhouse-server и clickhouse-client на обе ноды.

На обеих нодах создайте отдельный файл /etc/clickhouse-server/cluster.xml для конфигурации кластера, добавьте ссылку на данный файл в конфигурации ClickHouse (имя подключаемой записи должно быть clickhouse_remote_servers).

Опишите в файле /etc/clickhouse-server/cluster.xml конфигурацию кластера ClickHouse с 2-мя шардами по 1-й реплике в каждом (имя кластера - local), используя айпи адреса выданных вам машин.

Создайте в кластере локальные (ch_local) и distributed (ch_distributed) таблицы на каждой ноде в базе default. Структуру таблиц используйте как указано ниже.

Вставьте тестовые данные в distributed таблицу и убедитесь, что они распределились по шардам (Сделайте 10 (ровно 10) вставок с помощью insert случайных данных и убедитесь что в каждой локальной таблице вставилось по 5 записей)

Отправляйте задание на проверку.

Структура таблицы:

```
id Int64,  
title String,  
description String,  
content String,  
date Date
```

Итоговое оценивание

Вы и ознакомились с возможностями ClickHouse и добрались до финального задания.

Данное задание является результатом освоенных вами практических навыков, поэтому для корректности оценивания, у вас будет только одна попытка сдать его на проверку. Количество запусков инфраструктуры и выполнения не ограничено.

В качестве ответа надо нажать кнопку "Задать вопрос куратору" и прислать указанные данные.

Пример задания.

1. Вам выдано 4 машины
2. Вы должны собрать ClickHouse кластер со следующими характеристиками:
 - На первых трех нодах должен находиться zookeeper в кластере
 - На всех четырех нодах должен находиться ClickHouse
 - Первые две ноды надо настроить как первый шард и две реплики
 - Вторые две ноды надо настроить как второй шард и две реплики
3. В кластере необходимо создать таблицу events_local реплицируемую в каждом шарде
4. Также нужно создать таблицу events_distributed, которая будет смотреть на два шарда и реплики в них
5. Таблицу можно сделать по примерам из 9ого или 11ого заданий - она должна содержать datetime поле, event_type - идентификатор события

...

Критерии оценки

Все 4 сервера корректно развернуты, ClickHouse установлен и запущен
Zookeeper установлен и сконфигурирован в кластер на 3 нодах, процессы активны
ClickHouse-инстансы связаны в единый кластер

Настройка шардов и реплик:

- первый шард на 2 нодах, работающие реплики
- второй шард на 2 нодах, работающие реплики

Создана таблица events_local с репликацией внутри каждого шарда

Создана events_distributed, корректно ссылающаяся на локальные таблицы в двух шардах