

# Содержание

|                                                                                          |           |
|------------------------------------------------------------------------------------------|-----------|
| От издательства .....                                                                    | 12        |
| Высокая оценка книги «Trino. Анализ больших данных».....                                 | 13        |
| Предисловие .....                                                                        | 15        |
| Предисловие авторов .....                                                                | 17        |
| Об авторах .....                                                                         | 20        |
| Изображение на обложке .....                                                             | 21        |
| <b>Часть I. НАЧИНАЕМ РАБОТАТЬ С TRINO</b> .....                                          | <b>22</b> |
| <b>Глава 1. Введение в Trino</b> .....                                                   | <b>23</b> |
| Проблемы, связанные с большими данными .....                                             | 23        |
| Trino приходит на помощь .....                                                           | 25        |
| Предназначен для высокой производительности и масштабируемости.....                      | 26        |
| SQL для всего чего угодно .....                                                          | 27        |
| Разделение хранилища данных и ресурсов выполнения запроса .....                          | 28        |
| Варианты использования Trino .....                                                       | 28        |
| Одна точка доступа SQL-анализа .....                                                     | 29        |
| Точка доступа к хранилищу данных и системам-источникам .....                             | 30        |
| Предоставление доступа, основанного на SQL, к любому источнику .....                     | 30        |
| Федеративные запросы .....                                                               | 31        |
| Семантический уровень для виртуального хранилища данных .....                            | 32        |
| Механизм запросов к озеру данных .....                                                   | 33        |
| Преобразования SQL и процессы ETL .....                                                  | 33        |
| Улучшение качества аналитической информации благодаря<br>сокращению времени ответа ..... | 34        |
| Большие данные, машинное обучение и искусственный интеллект .....                        | 34        |
| Другие варианты использования .....                                                      | 35        |
| Ресурсы Trino .....                                                                      | 35        |
| Веб-сайт.....                                                                            | 35        |
| Документация .....                                                                       | 35        |
| Чат сообщества.....                                                                      | 36        |
| Исходный код, лицензия и версии .....                                                    | 36        |
| Участие в проекте .....                                                                  | 37        |
| Репозиторий книги.....                                                                   | 38        |
| Набор данных Iris.....                                                                   | 38        |
| Набор данных Flight .....                                                                | 39        |

|                                                          |           |
|----------------------------------------------------------|-----------|
| Краткая история Trino.....                               | 40        |
| Резюме .....                                             | 41        |
| <b>Глава 2. Установка и конфигурирование Trino .....</b> | <b>42</b> |
| Пробная установка Trino в контейнере Docker.....         | 42        |
| Установка из архивного файла.....                        | 44        |
| Java Virtual Machine .....                               | 44        |
| Python .....                                             | 44        |
| Установка .....                                          | 45        |
| Конфигурация .....                                       | 46        |
| Добавление источника данных.....                         | 47        |
| Запуск Trino.....                                        | 48        |
| Резюме .....                                             | 48        |
| <b>Глава 3. Использование Trino .....</b>                | <b>49</b> |
| Интерфейс командной строки Trino .....                   | 49        |
| Начинаем работу.....                                     | 49        |
| Разделение на страницы .....                             | 52        |
| История команд и автоматическое дополнение .....         | 52        |
| Дополнительная диагностика.....                          | 53        |
| Выполнение запросов .....                                | 53        |
| Форматы вывода .....                                     | 54        |
| Игнорирование ошибок.....                                | 54        |
| Драйвер Trino JDBC.....                                  | 54        |
| Загрузка и регистрация драйвера .....                    | 56        |
| Установление соединения с Trino .....                    | 57        |
| Trino и ODBC.....                                        | 59        |
| Клиентские библиотеки.....                               | 60        |
| Пользовательский интерфейс Trino Web UI.....             | 60        |
| SQL в Trino.....                                         | 61        |
| Основные концепции.....                                  | 62        |
| Первые примеры.....                                      | 62        |
| Резюме .....                                             | 65        |
| <b>Часть II. БОЛЕЕ ГЛУБОКОЕ ИЗУЧЕНИЕ TRINO .....</b>     | <b>66</b> |
| <b>Глава 4. Архитектура Trino.....</b>                   | <b>67</b> |
| Координатор и рабочие узлы в кластере .....              | 67        |
| Координатор .....                                        | 69        |
| Сервис обнаружения .....                                 | 70        |
| Рабочие узлы .....                                       | 70        |
| Архитектура на основе коннекторов .....                  | 71        |
| Каталоги, схемы и таблицы .....                          | 73        |
| Модель выполнения запроса.....                           | 73        |
| Планирование запроса.....                                | 78        |
| Синтаксический разбор и анализ.....                      | 79        |

|                                                                        |     |
|------------------------------------------------------------------------|-----|
| Начальное планирование запроса .....                                   | 80  |
| Правила оптимизации .....                                              | 82  |
| Продвижение предикатов к источникам данных .....                       | 82  |
| Исключение перекрестного соединения .....                              | 83  |
| Оптимизация TopN.....                                                  | 84  |
| Частичные агрегации .....                                              | 85  |
| Правила реализации.....                                                | 85  |
| Декорреляция латерального соединения .....                             | 86  |
| Декорреляция полусоединения (IN) .....                                 | 87  |
| Оптимизатор на основе стоимости.....                                   | 88  |
| Концепция стоимости.....                                               | 88  |
| Стоимость соединения.....                                              | 90  |
| Статистические характеристики таблиц.....                              | 91  |
| Статистические характеристики фильтра .....                            | 93  |
| Статистические характеристики для таблиц, разделенных на части.....    | 94  |
| Перечисление (нумерация) соединений .....                              | 95  |
| Сравнение широковещательной передачи и распределенных соединений ..... | 95  |
| Работа со статистическими характеристиками таблиц .....                | 97  |
| Команда Trino ANALYZE .....                                            | 98  |
| Сбор статистических характеристик при записи на диск .....             | 99  |
| Команда Hive ANALYZE .....                                             | 99  |
| Вывод статистических характеристик таблиц .....                        | 100 |
| Резюме .....                                                           | 100 |

## **Глава 5. Развертывание конфигурации Trino, готовой к промышленной эксплуатации .....**

|                                               |     |
|-----------------------------------------------|-----|
| .....                                         | 101 |
| Подробности конфигурации .....                | 101 |
| Конфигурация сервера .....                    | 101 |
| Ведение журналов .....                        | 102 |
| Конфигурация рабочего узла .....              | 104 |
| Конфигурация JVM.....                         | 104 |
| Скрипт запуска .....                          | 105 |
| Установка кластера .....                      | 107 |
| Установка RPM.....                            | 109 |
| Структура каталога комплекта установки .....  | 109 |
| Конфигурация .....                            | 110 |
| Удаление Trino.....                           | 110 |
| Установка в облачной среде .....              | 110 |
| Helm Chart для развертывания Kubernetes ..... | 111 |
| Условия определения размера кластера.....     | 113 |
| Резюме .....                                  | 114 |

|                                          |            |
|------------------------------------------|------------|
| <b>Глава 6. Коннекторы.....</b>          | <b>115</b> |
| Конфигурация.....                        | 116        |
| Пример коннектора СУРБД: PostgreSQL..... | 117        |

|                                                                       |     |
|-----------------------------------------------------------------------|-----|
| Проталкивание запроса .....                                           | 119 |
| Распараллеливание и одновременное выполнение.....                     | 120 |
| Другие коннекторы СУРБД.....                                          | 121 |
| Безопасность .....                                                    | 123 |
| Прямая (сквозная) передача запроса .....                              | 123 |
| Коннекторы Trino TPC-H и TPC-DS .....                                 | 124 |
| Коннектор Hive для источников данных с распределенным хранением ..... | 125 |
| Apache Hadoop и Hive .....                                            | 125 |
| Коннектор Hive .....                                                  | 127 |
| Формат таблиц в стиле Hive .....                                      | 128 |
| Управляемые и внешние таблицы .....                                   | 129 |
| Секционированные данные .....                                         | 131 |
| Загрузка данных.....                                                  | 133 |
| Форматы файлов и функция сжатия.....                                  | 135 |
| Пример MinIO.....                                                     | 136 |
| Управление и анализ для современных распределенных хранилищ.....      | 136 |
| Нереляционные источники данных.....                                   | 138 |
| Коннектор Trino JMX .....                                             | 139 |
| Коннектор Black Hole.....                                             | 141 |
| Коннектор памяти .....                                                | 142 |
| Другие коннекторы.....                                                | 142 |
| Резюме .....                                                          | 143 |

## **Глава 7. Дополнительные примеры коннекторов.....** 144

|                                                                         |     |
|-------------------------------------------------------------------------|-----|
| Установление соединения с HBase с помощью Phoenix.....                  | 144 |
| Пример коннектора для хранилища пар ключ-значение: Accumulo .....       | 145 |
| Использование коннектора Trino Accumulo .....                           | 149 |
| Проталкивание предиката в Accumulo .....                                | 151 |
| Коннектор Apache Cassandra .....                                        | 153 |
| Коннектор для системы обработки потоков: Kafka .....                    | 154 |
| Пример коннектора для хранилища документов: Elasticsearch.....          | 156 |
| Обзор .....                                                             | 156 |
| Конфигурирование и использование .....                                  | 157 |
| Обработка запросов.....                                                 | 158 |
| Полнотекстовый поиск .....                                              | 158 |
| Подведем итоги.....                                                     | 158 |
| Федеративные запросы в Trino .....                                      | 159 |
| Извлечение, преобразование, загрузка (ETL) и федеративные запросы ..... | 166 |
| Резюме .....                                                            | 167 |

## **Глава 8. Использование SQL в Trino.....** 168

|                               |     |
|-------------------------------|-----|
| Инструкции Trino .....        | 169 |
| Системные таблицы Trino ..... | 171 |
| Каталоги .....                | 173 |
| Схемы.....                    | 174 |
| Информационная схема .....    | 176 |

|                                                                  |            |
|------------------------------------------------------------------|------------|
| Таблицы .....                                                    | 177        |
| Свойства таблицы и столбцов .....                                | 179        |
| Копирование существующей таблицы .....                           | 180        |
| Создание новой таблицы по результатам запроса .....              | 181        |
| Изменение таблицы .....                                          | 182        |
| Удаление таблицы .....                                           | 183        |
| Ограничения таблиц, определяемые коннекторами .....              | 183        |
| Представления .....                                              | 184        |
| Информация и конфигурация сеанса .....                           | 185        |
| Типы данных .....                                                | 187        |
| Составные типы данных (коллекции) .....                          | 190        |
| Типы данных, описывающие время .....                             | 190        |
| Приведение типов .....                                           | 194        |
| Основы использования инструкции SELECT .....                     | 195        |
| Спецификатор WHERE .....                                         | 197        |
| Спецификаторы GROUP BY и HAVING .....                            | 198        |
| Спецификаторы ORDER BY и LIMIT .....                             | 199        |
| Инструкции JOIN .....                                            | 200        |
| Спецификаторы UNION, INTERSECT и EXCEPT .....                    | 202        |
| Группирующие операции .....                                      | 203        |
| Спецификатор WITH .....                                          | 205        |
| Подзапросы .....                                                 | 206        |
| Скалярный подзапрос .....                                        | 206        |
| Подзапрос EXISTS .....                                           | 207        |
| Квантифицированные подзапросы .....                              | 207        |
| Удаление данных из таблицы .....                                 | 208        |
| Резюме .....                                                     | 209        |
| <b>Глава 9. Расширенные средства SQL</b> .....                   | <b>210</b> |
| Предварительный обзор функций и операторов .....                 | 210        |
| Скалярные функции и операторы .....                              | 211        |
| Логические операторы отношения .....                             | 212        |
| Логические операторы .....                                       | 213        |
| Выбор диапазона с помощью инструкции BETWEEN .....               | 214        |
| Определение существования значения с помощью IS (NOT) NULL ..... | 215        |
| Математические функции и операторы .....                         | 215        |
| Тригонометрические функции .....                                 | 216        |
| Функции констант и случайных значений .....                      | 217        |
| Функции и операторы для работы со строками .....                 | 217        |
| Строки и отображения .....                                       | 219        |
| Unicode .....                                                    | 220        |
| Регулярные выражения .....                                       | 222        |
| Раскрытие вложенности составных типов данных .....               | 224        |
| Функции JSON .....                                               | 226        |
| Функции и операторы для работы с датами и временем .....         | 227        |
| Гистограммы .....                                                | 230        |

|                                                     |     |
|-----------------------------------------------------|-----|
| Функции агрегации .....                             | 230 |
| Функции агрегации, работающие с отображениями ..... | 231 |
| Функции агрегации с аппроксимацией .....            | 232 |
| Функции для работы с окнами.....                    | 234 |
| Лямбда-выражения.....                               | 235 |
| Геопространственные функции.....                    | 236 |
| Предварительно подготовленные инструкции .....      | 237 |
| Резюме .....                                        | 239 |

## **Часть III. TRINO В РЕАЛЬНЫХ ВАРИАНТАХ ИСПОЛЬЗОВАНИЯ..... 240**

|                                                                                 |            |
|---------------------------------------------------------------------------------|------------|
| <b>Глава 10. Обеспечение безопасности .....</b>                                 | <b>241</b> |
| Аутентификация .....                                                            | 241        |
| Пароль и аутентификация LDAP.....                                               | 243        |
| Другие типы аутентификации.....                                                 | 246        |
| Авторизация .....                                                               | 247        |
| Системное управление доступом.....                                              | 247        |
| Управление доступом на уровне коннектора.....                                   | 250        |
| Шифрование .....                                                                | 252        |
| Шифрование канала обмена данными между клиентом<br>и координатором Trino .....  | 255        |
| Создание хранилищ ключей и доверенных сертификатов Java .....                   | 259        |
| Шифрование данных, передаваемых внутри кластера Trino .....                     | 261        |
| Центр сертификации и самоподписанные сертификаты.....                           | 262        |
| Аутентификация с использованием сертификатов.....                               | 264        |
| Протокол Kerberos .....                                                         | 268        |
| Предварительные требования .....                                                | 269        |
| Аутентификация клиента по протоколу Kerberos.....                               | 269        |
| Доступ к источникам данных и конфигурация для обеспечения<br>безопасности ..... | 270        |
| Аутентификация по протоколу Kerberos в коннекторе Hive .....                    | 271        |
| Аутентификация через Hive Metastore Service .....                               | 272        |
| Аутентификация HDFS.....                                                        | 273        |
| Разделение кластера.....                                                        | 273        |
| Резюме .....                                                                    | 274        |

## **Глава 11. Интеграция Trino с другими инструментальными средствами .....**

|                                                                                     |            |
|-------------------------------------------------------------------------------------|------------|
| <b>Запросы, визуализации и т. п. в Apache Superset .....</b>                        | <b>275</b> |
| Улучшение производительности с помощью RubiX.....                                   | 277        |
| Управление рабочими процессами с помощью Apache Airflow.....                        | 277        |
| Пример встраивания Trino: Amazon Athena .....                                       | 278        |
| Удобные коммерческие дистрибутивы: Starburst Enterprise и Starburst<br>Galaxy ..... | 282        |
| Другие примеры интеграции.....                                                      | 283        |

|                                                      |            |
|------------------------------------------------------|------------|
| Специализированные интеграции.....                   | 284        |
| Резюме .....                                         | 284        |
| <b>Глава 12. Trino в производственной среде.....</b> | <b>286</b> |
| Мониторинг с использованием Trino Web UI .....       | 286        |
| Подробности уровня кластера .....                    | 287        |
| Список запросов .....                                | 288        |
| Представление Query Details.....                     | 291        |
| Точная настройка SQL-запросов Trino.....             | 298        |
| Управление памятью .....                             | 302        |
| Распараллеливание задач.....                         | 305        |
| Планирование рабочих узлов .....                     | 306        |
| Обмен данными в сетевой среде .....                  | 306        |
| Распараллеливание .....                              | 307        |
| Размеры буферов .....                                | 307        |
| Настройка Java Virtual Machine .....                 | 307        |
| Группы ресурсов.....                                 | 309        |
| Определение группы ресурсов .....                    | 311        |
| Стратегия планирования.....                          | 312        |
| Определение правил выбора .....                      | 312        |
| Резюме .....                                         | 313        |
| <b>Глава 13. Примеры из реальной практики.....</b>   | <b>314</b> |
| Платформы развертывания и времени выполнения .....   | 315        |
| Определение размера кластера .....                   | 316        |
| Вариант использования: переход с Hadoop/Hive .....   | 317        |
| Другие источники данных .....                        | 318        |
| Пользователи и трафик .....                          | 318        |
| Резюме .....                                         | 319        |
| <b>Заключительное слово.....</b>                     | <b>320</b> |
| <b>Предметный указатель.....</b>                     | <b>321</b> |

# От издательства

## ***Отзывы и пожелания***

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте [www.dmkpress.com](http://www.dmkpress.com), зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com); при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу [http://dmkpress.com/authors/publish\\_book/](http://dmkpress.com/authors/publish_book/) или напишите в издательство по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com).

## ***Список опечаток***

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг, мы будем очень благодарны, если вы сообщите о ней главному редактору по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com). Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

## ***Нарушение авторских прав***

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу электронной почты [dmkpress@gmail.com](mailto:dmkpress@gmail.com).

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

# Высокая оценка книги «Trino. Анализ больших данных»

«Эта книга предоставляет великолепное введение в Trino и обучает читателя всему, что необходимо знать для начала успешного использования Trino».

– *Дэйн Сундстрём (Dain Sundstrom) и Дэвид Филлунс (David Phillips), создатели проекта Trino Project и основатели Trino Software Foundation*

«Trino играет самую важную роль в обеспечении анализа в компании Pinterest. Эта книга охватывает все основные аспекты Trino – от конкретных вариантов использования до способов применения этого механизма в крупном масштабе».

– *Ашиш Кумар Сингх (Ashish Kumar Singh), технический руководитель проекта Big Data Query Processing Platform компании Pinterest*

«Trino оказался невероятно полезным средством для выполнения быстрых сиюминутных ситуативных запросов; именно поэтому он стал одним из первых двух механизмов, поддержку которого мы добавили в Iceberg. Я очень рад видеть расширение сферы применения Trino в рабочих процессах крупных хранилищ данных в качестве коннектора Iceberg и по прошествии четырех лет отметить рост поддержки материализованных представлений, выразительных команд DML и исследования метаданных».

– *Райан Блю (Ryan Blue), сооснователь Apache Iceberg, сооснователь и главный исполнительный директор (CEO) платформы Tabular*

«Trino установил высокую планку как в формировании сообщества, так и в уровне технического совершенства для молниеносной аналитической обработки данных, хранимых в современных облачных архитектурах. Эту книгу обязательно должны прочесть компании, желающие модернизировать свой стек аналитических инструментальных средств».

– *Джей Крепс (Jay Kreps), сооснователь Apache Kafka, также сооснователь и главный исполнительный директор (CEO) компании Confluent*

«Trino позволил всем – в науке и промышленности – сэкономить бесконечное количество рабочих часов, позволяя нам избежать необходимости в написании кода для управления распределенной обработкой запросов. Мы весьма благодарны за то, что у нас есть высококачественный распределенный SQL-механизм с открытым исходным кодом, с которого можно начать работу, что позволяет нам сосредоточиться на инновациях в новых областях, а не изобретать велосипед для каждого нового проекта распределенной системы обработки данных».

– Дэниел Абади (*Daniel Abadi*),  
профессор ИТ, университет Мэриленда в Колледж-парке

# Предисловие

Неужели прошло уже больше десяти лет? Любопытно: кажется, что это было вчера и одновременно как будто происходило давно. Мы начали проект Presto в Facebook в 2012 г., а различные компоненты и идеи появились еще раньше. Разумеется, мы планировали создать что-то полезное и всегда хотели получить успешный проект с открытым исходным кодом и сформировавшееся вокруг него сообщество. В 2013 г. мы выпустили первую версию Presto под лицензией Apache, и приключение началось.

С тех пор все мы ушли из Facebook, отказались от названия Presto и вообще вышли далеко за пределы того, о чем мечтали. Мы гордимся достижениями сообщества проекта и в высшей степени благодарны за все положительные отзывы и помощь, которую получили.

Проект Trino вырос в невероятной степени и принес огромную пользу своему большому сообществу пользователей. Вы можете найти членов сообщества Trino по всему миру, а также разработчиков в Бразилии, Канаде, Китае, Германии, Индии, Израиле, Японии, Польше, Сингапуре, США, Великобритании и других странах.

Основание Trino Software Foundation в начале 2019 г. и переход проекта к независимому статусу стали еще одной важной вехой. Эта некоммерческая организация занимается развитием распределенного SQL-механизма с открытым исходным кодом Trino. Фонд стремится обеспечить такое положение дел, чтобы проект оставался открытым, коллективным и независимым в течение будущих десятилетий.

Переименование проекта и фонда в Trino в конце 2020 г. стало еще одним важным поворотным пунктом для сообщества. Мы получили новое имя, новый облик, обновленный интерес со стороны всего мира и даже создали замечательный маленький талисман – нашего любимого кролика Командира Бан Бана (Commander Ban Ban). Теперь, спустя три года после основания фонда, мы можем бросить ретроспективный взгляд на ускоренный темп впечатляющих достижений большого и постоянно растущего сообщества.

Мы очень рады, что Мэтт, Манфред и Мартин подготовили обновленное второе издание книги о Trino с помощью издательства O'Reilly. Эта книга предоставляет великолепное введение в Trino и обучает читателя всему, что нужно знать, чтобы начать успешно использовать данный механизм.

Наслаждайтесь путешествием в глубины Trino и связанного с ним мира бизнес-аналитики, формирования отчетов, создания информационных и управляющих панелей, хранилищ данных, интеллектуального анализа данных, машинного обучения и многого другого.

Разумеется, мы настоятельно рекомендуем внимательно изучать дополнительные ресурсы и пользоваться помощью, которую мы предоставляем на веб-сайте Trino <https://trino.io>, а также не забывать о чате сообщества, о репозитории исходного кода, о Trino Community Broadcast и прочих полезных ресурсах.

Добро пожаловать в сообщество Trino!

– Дэйн Сундстрём (*Dain Sundstrom*) и Дэвид Филлипс (*David Phillips*),  
создатели проекта *Trino Project* и основатели *Trino Software Foundation*

# Предисловие авторов

«Trino. Анализ больших данных» – первая и главная книга о механизме распределенных запросов Trino. Книга рассчитана как на новичков, так и на более опытных пользователей Trino. В идеальном варианте у читателя должно быть некоторое понимание принципов работы с базами данных и языком запросов SQL, но при отсутствии таких знаний вы можете иногда отвлекаться от чтения и находить требуемую информацию в процессе работы над этой книгой. Мы уверены, что вы узнаете что-то новое из этой книги независимо от уровня своих знаний.

Во втором издании содержание существенно обновлено, чтобы не отставать от быстрых нововведений в Trino. Мы охватываем новые аспекты, такие как диаграмма Helm для развертывания кластера Trino в Kubernetes, новые коннекторы Iceberg и Delta Lake для современных архитектур хранилищ в озерах данных, отказоустойчивое выполнение для обработки запросов, расширенные возможности языка SQL и последний релиз Trino, теперь работающий на версии Java 17.

В первой части книги вы получаете общее представление о механизме Trino и необходимую вспомогательную информацию в кратком изложении, чтобы можно было начать практическое освоение данного механизма. Эта часть включает описание процедуры установки и основ использования интерфейса командной строки, а также клиентских и основанных на вебе приложений, таких как системы управления баз данных SQL или панель управления и инструменты формирования отчетов, использующие драйвер JDBC.

Вторая часть книги позволяет читателю расширить и углубить знания и содержит подробное описание архитектуры Trino, процедуры развертывания кластера, многочисленных коннекторов к источникам данных и огромный объем информации о главном мощном инструментальном средстве Trino – о запросах к любому источнику данных на языке SQL.

В заключительной третьей части книги содержатся дополнительные сведения, которые необходимо знать при запуске и использовании производственной среды развертывания Trino. Здесь описаны использование пользовательского веб-интерфейса, настройка безопасности, а также обсуждаются некоторые аспекты реального применения Trino в других организациях.

## Условные обозначения, используемые в книге

В этой книге используются условные обозначения, перечисленные ниже.




**Полужирным шрифтом** выделяются названия окон, вкладок, пунктов меню, кнопок и т. п.

*Курсивом* выделяются новые термины, имена и расширения файлов.

Моноширинный шрифт используется для листингов программ, а также внутри текстовых абзацев для обозначения программных элементов, таких как имена переменных и функций, баз данных, типов данных, переменных среды, инструкций и ключевых слов.

**Полужирный моноширинный шрифт** обозначает команды и инструкции, которые пользователь должен вводить точно так, как они приведены в тексте.

*Курсивным моноширинным шрифтом* выделяются элементы, которые пользователь должен заменить собственными значениями или значениями, определяемыми контекстом.

-  Этот значок обозначает совет или предложение.
-  Этот значок обозначает общее примечание.
-  Этот значок обозначает предупреждение или факт, требующий особого внимания.

## Примеры кода, разрешение на использование и ссылки на источник

Дополнительные материалы к книге подробно описаны в подразделе «Репозиторий книги» главы 1.

Если у вас возник технический вопрос или проблема с использованием кода примера, то свяжитесь с авторами в чате сообщества (см. подраздел «Чат сообщества» в главе 1) или разместите файл с описанием проблем в репозитории книги.

Эта книга существует для того, чтобы помогать вам сделать свою работу. В общем случае, если пример кода приведен в книге, вы можете использовать его в своих программах или в документации. Нет необходимости запрашивать у авторов разрешение, если только вы не воспроизводите весьма значительную часть кода. Например, при написании программы, в которой используется несколько небольших фрагментов кода из этой книги, не требуется специальное разрешение. Для продажи и распространения кода примеров из книг O'Reilly необходимо получить разрешение. Ответ на вопрос с использованием цитаты из этой книги и цитирование кода примеров не требует специального разрешения. Включение значительного объема примера кода из этой книги в документацию вашего программного продукта требует разрешения.

Ссылки на источник не являются обязательным требованием, но поощряются. Обычно ссылка на источник включает название, автора (авторов), издательство и ISBN. В данном случае ссылка на источник выглядит следующим образом:

Trino: The Definitive Guide, 2nd Edition, by Matt Fuller, Manfred Moser, and Martin Traverso (O'Reilly). Copyright 2023 Matt Fuller, Martin Traverso, and simpligility technologies, Inc., 978-1-098-13723-6.

Если вы не уверены в правильности использования кода примеров и в соответствии с условиями получения разрешений, приведенных выше, то свяжитесь с нами по электронной почте [editor@dmkpress.com](mailto:editor@dmkpress.com).

## Благодарности

Мы хотели бы поблагодарить всех в большом сообществе Trino, начиная с основателей проекта, специализированной команды сопровождения и всех участников с момента его создания. Наша благодарность выходит далеко за рамки сообщества, и мы ценим каждого, кто использует Trino, распространяет информацию, помогает другим пользователям, вносит вклад в связанные с Trino проекты и интеграции и просто задает вопросы в нашем чате. Мы счастливы быть частью этого сообщества и с нетерпением ждем многочисленных общих успехов в будущем.

Весьма важной частью сообщества является Starburst. Мы хотим поблагодарить всех в Starburst за помощь. Мы действительно ценим работу, ресурсы, стабильность и поддержку, которые Starburst предоставляет проекту, его клиентам, использующим Trino, и авторам, которые являются частью команды Starburst.

Что касается этой книги, мы хотели бы поблагодарить всех, кто помогал нам идеями, вкладом и отзывами. Кроме того, авторы хотят выразить свою личную благодарность:

- Мэтт благодарит свою жену Меган и троих детей – Эмили, Ханну и Лайама – за их терпение и поддержку во время работы над книгой. Восхищение детей по поводу того, что их папа становится «автором книги», помогало Мэтту на протяжении множества длинных уикендов и бессонных ночей;
- Манфред благодарен своей жене Йен и трем сыновьям – Лукасу, Николасу и Тобиасу – не только за «техническую абракадабру», но и за искреннее разделение с ним интереса и страстного увлечения технологиями, писательством, обучением и преподаванием;
- Мартин благодарит свою жену Мелину и четырех детей – Маркоса, Викторину, Хоакина и Мартину – за их поддержку и энтузиазм в течение последних 10 лет работы над Trino.

# Об авторах

**Мэтт Фуллер** (Matt Fuller) – сооснователь компании Starburst. До создания Starburst Мэтт был техническим директором в компании Teradata, где работал над созданием нового подразделения Center for Hadoop внутри компании. Как важный участник этого проекта, Мэтт работал над выводом Trino на корпоративный рынок. Мэтт руководил командой, вносящей вклад в проект Trino с открытым исходным кодом с 2015 г., и руководил внутренней дорожной картой продукта Trino. В дальнейшем состав компании Starburst был сформирован из этой группы в Teradata.

До работы в Teradata Мэтт был начинающим инженером в компании Vertica, где участвовал в создании оптимизатора запросов. Мэтт также является автором публикаций на конференциях Very Large Databases (VLDB) и имеет патенты США в области систем управления базами данных.

**Манфред Мозер** (Manfred Moser) – общественный деятель, писатель, инструктор и инженер по программному обеспечению в компании Starburst. Манфред обладает многолетним опытом в области разработки и пропаганды программного обеспечения с открытым исходным кодом. Он является активным участником проекта Apache Maven, написал книгу по Hudson и др. Продолжает активно участвовать в сообществе с открытым исходным кодом и сопровождает собственные проекты. Манфред – опытный инструктор и докладчик на конференциях, обучивший более 20 000 разработчиков для таких компаний, как Walmart Labs, Sonatype и Telus.

Его опыт работы с базами данных включает проектирование баз данных и связанных с ними приложений в пространстве СУРБД, а также работу в качестве консультанта по бизнес-аналитике, вручную контролирующего тысячи строк SQL. Он рад, что теперь может использовать Trino, и распространяет информацию о том, насколько хорош Trino, в качестве соведущего Trino Community Broadcast и в других средствах информации.

**Мартин Траверсо** (Martin Traverso) – сооснователь Trino Software Foundation и главный технический директор компании Starburst. До Starburst Мартин работал инженером по программному обеспечению в Facebook, где обратил внимание на необходимость в быстрой интерактивной аналитике SQL. Мартин и три других инженера работали над созданием программного средства, которое позже стало известно как Presto, а затем получило новое имя Trino. Мартин возглавлял группу разработчиков Trino. Весной 2013 г. Trino был запущен в производство, а осенью 2013 г. его исходный код стал открытым. С тех пор Trino получил широкое распространение как внутри, так и за пределами Facebook.

До Facebook Мартин работал архитектором ПО в компаниях Proofpoint и Ning, где руководил разработкой и проектированием архитектуры многочисленных сложных корпоративных и социальных сетевых приложений.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

[e-Univers.ru](http://e-Univers.ru)