

# Содержание

<i>Положительные отзывы к книге</i> .....	12
<i>Предисловие Геррита Казмайера</i> .....	13
<i>Предисловие Рагху Рамакришнана</i> .....	14
<i>Предисловие Рика Сирса</i> .....	15
<i>Вступление</i> .....	16
<i>Благодарности</i> .....	21
<i>Об авторах</i> .....	22
<i>Колофон</i> .....	23

<b>Часть I. ОСНОВЫ APACHE ICEBERG</b> .....	24
---	----

<b>Глава 1. Введение в Apache Iceberg</b> .....	25
---	----

Краткая история .....	25
Основные компоненты системы для рабочих нагрузок OLAP .....	26
Все вместе .....	29
Хранилища данных .....	29
Краткая история .....	30
Плюсы и минусы хранилищ данных .....	30
Озера данных .....	32
Краткая история .....	33
Плюсы и минусы озер данных .....	34
Где должна производиться аналитическая обработка – в озере или в хранилище данных? .....	36
Озерное хранилище данных .....	37
Что такое формат таблиц? .....	39
Hive: оригинальный формат таблиц .....	40
Современные форматы таблиц озер данных .....	42
Что такое Apache Iceberg? .....	43
Как появился Apache Iceberg .....	43
Архитектура Apache Iceberg .....	45
Ключевые особенности Apache Iceberg .....	46
Заключение .....	50

<b>Глава 2. Архитектура Apache Iceberg</b> .....	51
--	----

Слой данных .....	52
Файлы данных .....	52
Файлы удаления .....	54
Слой метаданных .....	56
Файлы манифестов .....	57

Списки манифестов.....	58
Файлы метаданных .....	60
Puffin-файлы.....	62
Каталог.....	63
Заключение .....	64

### **Глава 3. Жизненный цикл запросов на запись и чтение .....**

Написание запросов в Apache Iceberg.....	68
Создание таблицы .....	68
Запрос вставки .....	70
Запрос слияния .....	74
Запросы на чтение в Apache Iceberg.....	78
Запрос SELECT .....	79
Запрос информации о состоянии в прошлом .....	82
Заключение .....	86

### **Глава 4. Оптимизация производительности таблиц Iceberg.....**

Уплотнение .....	87
Практическая реализация уплотнения .....	89
Стратегии уплотнения .....	94
Автоматизация уплотнения.....	96
Сортировка.....	97
Z-упорядочение.....	101
Секционирование .....	103
Скрытое секционирование .....	105
Эволюция разделов .....	106
Другие соображения по секционированию.....	107
Копирование при записи и слияние при чтении.....	108
Копирование при записи.....	109
Слияние при чтении.....	109
Настройка COW и MOR.....	112
Другие соображения .....	113
Коллекция метрик .....	113
Перезапись манифестов .....	114
Оптимизация хранилища.....	115
Режим распределения записи.....	117
Объектное хранилище .....	118
Фильтры Блума.....	119
Заключение .....	120

### **Глава 5. Каталоги Iceberg.....**

Требования к каталогу Iceberg .....	121
Сравнение каталогов .....	122
Каталог Nadoop .....	122
Каталог Hive .....	124

Каталог AWS Glue .....	126
Каталог Nessie.....	127
Каталог REST.....	128
Каталог JDBC.....	129
Другие каталоги .....	130
Миграция каталога .....	131
С помощью инструмента командной строки для миграции каталога Iceberg.....	131
С помощью механизма обработки запросов.....	133
Заключение .....	136

## **Часть II. ПРАКТИЧЕСКИЕ ПРИЕМЫ РАБОТЫ С APACHE ICEBERG .....**

137

### **Глава 6. Apache Spark .....**

138

Настройка.....	138
Настройка Apache Iceberg и Spark .....	138
Настройка каталогов .....	140
Запуск Spark со всеми конфигурациями (пример AWS Glue).....	143
Операции на языке определения данных.....	144
CREATE TABLE .....	145
ALTER TABLE .....	148
Изменение таблиц с помощью расширений Iceberg Spark SQL.....	151
DROP TABLE .....	154
Чтение данных.....	154
Запрос выборки всех записей .....	154
Запрос с фильтрацией записей.....	154
Агрегирующие запросы .....	155
Оконные функции .....	156
Запись данных .....	157
INSERT INTO .....	157
MERGE INTO.....	158
INSERT OVERWRITE .....	158
DELETE FROM .....	160
UPDATE .....	160
Процедуры обслуживания таблиц Iceberg.....	161
Удаление снимков .....	161
Перезапись файлов данных .....	162
Перезапись манифестов .....	162
Удаление осиротевших файлов .....	162
Заключение .....	163

### **Глава 7. Dremio SQL Query Engine .....**

164

Настройка.....	164
Операции на языке определения данных.....	166
CREATE TABLE .....	166

ALTER TABLE .....	168
DROP TABLE .....	169
Чтение данных.....	169
Использование запроса SELECT .....	170
Фильтрация записей .....	170
Агрегирующие запросы .....	170
Оконные функции .....	171
Запись данных .....	171
INSERT INTO .....	172
COPY INTO.....	172
MERGE INTO.....	172
DELETE.....	173
UPDATE .....	173
Обслуживание таблиц Iceberg .....	173
Удаление снимков .....	174
Перезапись файлов данных .....	174
Перезапись манифестов .....	174
Заключение .....	175
<b>Глава 8. AWS Glue.....</b>	<b>176</b>
Настройка.....	176
Создание базы данных Glue .....	176
Настройка ETL-задания в Glue.....	177
Создание таблицы с помощью Glue Data Catalog.....	180
Чтение из таблицы.....	180
Вставка данных .....	181
Заключение .....	181
<b>Глава 9. Apache Flink.....</b>	<b>182</b>
Настройка.....	182
Предварительные условия.....	182
Запуск кластера Flink и клиента Flink SQL.....	184
Операции на языке определения данных .....	185
CREATE CATALOG .....	185
CREATE DATABASE .....	187
CREATE TABLE .....	188
ALTER TABLE .....	189
DROP TABLE .....	189
Чтение данных.....	189
Пакетное чтение .....	189
Потоковое чтение .....	189
Таблицы с метаданными .....	190
Запись данных .....	191
INSERT INTO .....	191
INSERT OVERWRITE .....	191
UPSERT.....	192

Flink DataStream и Table API .....	193
Предварительные условия .....	193
Настройка задания Flink .....	193
Запуск кластера и сборка пакета .....	197
Выполнение задания .....	197
Заключение .....	198

### **Часть III. АРАСНЕ ICEBERG НА ПРАКТИКЕ .....** 199

#### **Глава 10. Apache Iceberg в производстве .....** 200

Таблицы метаданных в Apache Iceberg .....	201
Таблица метаданных history .....	201
Таблица метаданных metadata_log_entries .....	203
Таблица метаданных snapshots .....	204
Таблица метаданных files .....	206
Таблица метаданных manifests .....	208
Таблица метаданных partitions .....	211
Таблица метаданных all_data_files .....	212
Таблица метаданных all_manifests .....	215
Таблица метаданных refs .....	217
Таблица метаданных entries .....	218
Совместное использование таблиц метаданных .....	220
Изоляция изменений с помощью ветвления .....	223
Ветвление и маркировка на уровне таблиц .....	224
Ветвление и маркировка на уровне каталога .....	227
Многотабличные транзакции .....	230
Откат изменений .....	231
Откат на уровне таблицы .....	231
Откат на уровне каталога .....	234
Заключение .....	235

#### **Глава 11. Поточковая обработка в Apache Iceberg .....** 237

Поточковая обработка с помощью Spark .....	238
Поточковая передача данных в Iceberg .....	239
Поточковая передача данных из Iceberg .....	242
Поточковая обработка с помощью Flink .....	243
Поточковая передача данных в Iceberg .....	245
Пример потоковой передачи в Iceberg с помощью Flink .....	250
Поточковая обработка с помощью Kafka Connect .....	252
Iceberg Kafka Sink .....	253
Поточковая обработка с помощью AWS .....	256
Заключение .....	259

#### **Глава 12. Управление и безопасность .....** 260

Безопасность файлов данных .....	261
Защита файлов: практические приемы .....	261

Распределенная файловая система Hadoop.....	262
Amazon Simple Storage Service .....	264
Azure Data Lake Storage .....	268
Google Cloud Storage .....	271
Управление безопасностью на семантическом уровне .....	273
Практика применения семантического уровня.....	274
Dremio.....	275
Trino .....	278
Управление безопасностью на уровне каталога .....	280
Nessie .....	281
Tabular .....	282
AWS Glue и AWS Lake Formation.....	283
Дополнительные соображения по управлению безопасностью.....	285
Заключение .....	286

### **Глава 13. Миграция данных в Apache Iceberg..... 287**

Вопросы планирования миграции .....	288
План миграции на месте в три этапа.....	289
План теневой миграции в четыре этапа .....	290
Миграция таблиц Hive в Apache Iceberg .....	290
Процедура snapshot.....	291
Процедура migrate .....	291
Миграция из Delta Lake в Apache Iceberg.....	292
Миграция из Apache Hudi в Apache Iceberg.....	293
Миграция отдельных файлов в ApacheIceberg .....	294
Использование процедуры add_files .....	295
Миграция из Delta Lake или Apache Hudi без сохранения истории .....	295
Миграция откуда угодно путем копирования данных.....	296
Миграция данных в новую таблицу Iceberg.....	296
Миграция данных в существующую таблицу Iceberg.....	298
Заключение .....	301

### **Глава 14. Примеры использования Apache Iceberg ..... 302**

Обеспечение высокого качества данных в Apache Iceberg.....	302
WAP с использованием функции ветвления в Iceberg .....	303
Запуск аналитических рабочих нагрузок в озере данных.....	310
Помещение исходных данных в озеро .....	311
Курирование виртуальных витрины и продуктов данных .....	312
Создание агрегированных представлений для ускорения аналитических панелей .....	313
Подключение аналитических инструментов к представлению.....	313
Преимущества выполнения аналитических рабочих нагрузок в озере данных .....	314
Реализация сбора измененных данных с помощью Apache Iceberg .....	314
Создание таблиц Apache Iceberg.....	315
Применение обновлений из операционных систем.....	318

Создание представления журнала изменений для сбора изменений.....	319
Добавление изменений в агрегированную таблицу .....	320
Заключение .....	322

## **Глава 15. Дополнительные примеры использования**

<b>Apache Iceberg</b> .....	323
-----------------------------	-----

Рабочие нагрузки реального времени, использующие AWS Kinesis и Apache Iceberg.....	323
Прием потоков с помощью AWS Kinesis.....	325
Преобразование данных и ETL с помощью AWS Glue.....	327
Анализ потока .....	330
Машинное обучение с Apache Iceberg.....	331
Сценарий 1: надежные конвейеры машинного обучения .....	332
Сценарий 2: воспроизводимость машинного обучения .....	334
Сценарий 3: поддержка экспериментов с машинным обучением .....	338
Работа с историческими данными и медленно меняющимися измерениями .....	340
Создание таблиц фактов и измерений для подготовки к применению SCD2 .....	342
Обработка изменений в данных клиентов и реализация SCD2 .....	344
Добавление новых данных о продажах в новом месте.....	347

## **Глава 16. Библиотеки на Java и Python для работы с Iceberg**..... 349

Java API.....	349
Создание таблицы .....	349
Создание схемы .....	350
Создание спецификации секционирования .....	351
Чтение таблицы.....	351
Обновление таблицы.....	353
Создание транзакции.....	353
Python API .....	354
Установка PyIceberg .....	354
Настройка каталога .....	354
Загрузка таблицы .....	355
Создание таблицы .....	356
Обновление свойств таблицы.....	357
Запрос данных.....	357
Заключение .....	358

<b>Предметный указатель</b> .....	359
-----------------------------------	-----

# Положительные отзывы к книге

Это фантастическое учебное и справочное руководство по внутреннему устройству Apache Iceberg. Моя команда считает его бесценным.

– Каашиф Хаймабаккус (*Kaashif Hymabaccus*),  
старший инженер-программист, *Bloomberg*

Apache Iceberg постепенно становится фактическим стандартным форматом таблиц для платформ данных следующего поколения, и эта книга поможет изучить его основные идеи и компоненты – все это обязательно потребуется большинству инженеров по данным в ближайшие годы.

– Махди Карабибен (*Mahdi Karabiben*),  
инженер по данным, *Zendesk*

С момента появления озер данных Apache Iceberg начал набирать обороты. Эта книга познакомит вас с основными идеями табличного формата Iceberg, даст вам все необходимое для успешного использования в производстве и будет служить исчерпывающим справочником спустя месяцы после начала работы. Томер, Джейсон и Алекс – молодцы!

– Макс Шульце (*Max Schultze*),  
заместитель директора по инжинирингу данных, *HelloFresh*

Комплексный обзор Apache Iceberg от архитектуры до решений по проектированию и реализации.

– Симеон Шварц (*Simeon Schwarz*),  
директор по данным и аналитике, *OMS National Insurance Company*

# Предисловие

## Геррита Казмайера

Как человек, глубоко заинтересованный в развитии управления данными, я рад представить эту книгу об Apache Iceberg, особенно ценную в наше время, когда отрасли приходится решать сложнейшие задачи управления данными. Apache Iceberg с его революционным форматом открытых таблиц – это не просто технологическое достижение, это значительный сдвиг в подходе организаций к управлению данными в эпоху искусственного интеллекта.

Участие в жизни сообщества специалистов по обработке данных позволило мне заметить растущее любопытство и спрос на знания об озерах данных и Apache Iceberg. Учитывая это, можно смело утверждать, что эта книга – не просто информационный ресурс, но необходимое руководство для тех, кто хочет понять и начать применять эту технологию в своей работе. Она подробно обсуждает архитектуру Apache Iceberg и описывает варианты и передовой опыт практического применения, служа незаменимым источником знаний для архитекторов данных и инженеров.

Эта книга – больше, чем просто сборник информации; это свидетельство развития в сфере управления данными и маяк для тех, кто пытается разобраться в ее сложностях. Я уверен, что эта книга станет ценным ресурсом для всех, кто решит использовать Apache Iceberg в архитектурах данных.

*– Геррит Казмайер (Gerrit Kazmaier),  
вице-президент и генеральный директор  
по анализу данных, Google Cloud*

# Предисловие

## Рагху Рамакришнана

Apache Iceberg – один из ведущих открытых форматов обновляемых таблиц на основе Parquet, которые постепенно превращаются в новый стандарт хранения данных для аналитики. Исторически в реляционных базах данных информация хранилась в виде последовательностей записей, упакованных в физические страницы ради эффективности ввода-вывода. Однако хранение данных по столбцам оказалось гораздо эффективнее в окружениях, где приходится обрабатывать большое количество запросов в единицу времени. Озера данных изначально поддерживали запросы к столбчатым форматам, таким как Parquet, но им также требовалась эффективная поддержка транзакционного обновления для решения традиционных задач хранения данных. Iceberg как формат хранения таблиц приобретает все бóльшую популярность, так как он поддерживает сценарии, что необходимо в окружениях, интенсивно обрабатывающих запросы и одновременно обновляющих и загружающих большие объемы данных.

Эта своевременная книга хорошо описывает и основы Iceberg, и его архитектуру, и способы достижения максимальной производительности в широком спектре рабочих нагрузок, включая SQL-запросы в Apache Spark и Dremio, а также потоковую обработку в Apache Flink. В книге также есть глава, в которой рассматривается применение Iceberg в промышленных окружениях, включая использование таблиц метаданных и таких функций, как ветвление, секционирование и мгновенные снимки (снапшоты), для обработки сложных сценариев в масштабе. Вся эта информация бесценна для читателей, интересующихся особенностями системы Iceberg, разработкой приложений, использующих Iceberg (или систему, основанную на Iceberg).

*– Рагху Рамакришнан,  
технический директор по данным, технический партнер Microsoft*

# Предисловие Рика Сирса

Данные стали центральной частью современных программных приложений и современных организаций, чья работа основана на использовании данных. Инженеры данных, администраторы данных, аналитики данных и ученые, работающие с данными, входят в число людей в таких организациях, которые хотят еще эффективнее использовать свои данные. Многие из этих специалистов по обработке данных создают свои приложения, управляемые данными, на Amazon Web Services (AWS), часто предпочитая хранить свои данные в озере на базе Amazon Simple Storage Service (S3).

Все они могут пожелать изменять свои данные и манипулировать ими с течением времени, продолжая использовать их в процессе изменения, и, следовательно, создавать свои приложения с поддержкой транзакций в озере данных. Apache Iceberg – ключевая технология, используемая клиентами AWS для создания озер транзакционных данных. Она отличается высоким быстродействием, эффективностью и надежностью, предлагает простую интеграцию с популярными платформами обработки данных на AWS, такими как Apache Spark, Apache Flink, Apache Hive, Presto, Trino, Dremio и др., а также поддерживается сервисами AWS, такими как Amazon EMR, Amazon Redshift, Amazon Athena, AWS Glue и др.

«Apache Iceberg. Полное руководство» фокусируется на практике применения и сценариях, полезных для специалистов по работе с данными, использующих Apache Iceberg, а также содержит практические упражнения, помогающие приобрести навыки использования Iceberg в комплексе с ключевыми технологиями AWS, такими как Amazon EMR и AWS Glue, которые поддерживают оптимизацию и позволяют с помощью Iceberg создавать масштабируемые приложения. Книга охватывает весь спектр знаний, необходимых этим специалистам, от решаемых ими задач до архитектуры приложений, управляемых данными, и лучших практик реального использования в AWS, поэтому читатели смогут не только понять и поэкспериментировать с Apache Iceberg, но и в кратчайшие сроки внедрить эту технологию в производство.

Содержимое для этой книги подбиралось с учетом обсуждений с клиентами AWS и более широким сообществом. Поскольку Apache Iceberg становится все более важным для клиентов AWS, я рад, что они получают этот ценный справочник по созданию, масштабированию и оптимизации приложений с помощью Iceberg.

– Рик Сирс (Rick Sears), генеральный директор Amazon Web Services,  
Amazon EMR, Amazon Athena, AWS Lake Formation  
и AWS Glue Data Catalog

# Вступление

Добро пожаловать в «Apache Iceberg. Полное руководство»! Мы рады, что вы решили отправиться в путешествие по этой технологии вместе с нами. В этом вступлении мы даем краткий обзор книги, объясняем, почему мы ее написали и как извлечь из нее максимальную пользу.

## ***Об этой книге***

На страницах этой книги вы узнаете, что такое Apache Iceberg, как он появился, как работает и как использовать его возможности. Эта книга, предназначенная для инженеров, архитекторов, ученых и аналитиков, работающих с большими наборами данных в различных сценариях, от информационных панелей бизнес-аналитики до искусственного интеллекта и машинного обучения, исследует основные концепции, внутреннюю работу и практическое применение Apache Iceberg. К концу книги вы усвоите основы и овладете практическими знаниями, необходимыми для эффективного внедрения Apache Iceberg в свои проекты. Независимо от того, являетесь ли вы новичком или опытным практиком, «Apache Iceberg. Полное руководство» станет вашим надежным спутником в этом занимательном путешествии по Apache Iceberg.

## ***Почему мы написали эту книгу***

Когда мы наблюдали за быстрым ростом и распространением экосистемы Apache Iceberg, стало очевидно, что необходимо устранить растущий пробел в знаниях. Первоначально мы начали с обмена идеями через серию публикаций в блогах на платформе Dremio, чтобы передать ценную информацию растущему сообществу Iceberg. Однако вскоре стало ясно, что для удовлетворения растущего спроса на исчерпывающую информацию об Iceberg необходим всеобъемлющий и централизованный ресурс. Это осознание стало главной причиной появления «Apache Iceberg. Полное руководство». Наша цель – дать читателям единый авторитетный источник, устраняющий пробелы в знаниях и дающий возможность отдельным лицам и организациям максимально эффективно использовать Apache Iceberg.

## ***Что вы найдете внутри***

В главах, следующих далее, вы узнаете, что такое и как работает Apache Iceberg, как использовать преимущества этого формата в комплексе с различными инструментами, а также познакомитесь с передовыми методами

управления данными в таблицах Apache Iceberg. Вот краткое содержание каждой главы.

*Глава 1 «Введение в Apache Iceberg»* исследует исторический контекст озер данных и основных идей, лежащих в основе Apache Iceberg.

*Глава 2 «Архитектура Apache Iceberg»* подробно описывает сложную конструкцию Apache Iceberg и взаимодействия его компонентов.

*Глава 3 «Жизненный цикл запросов на запись и чтение»* дает пошаговый анализ процесса выполнения транзакций Apache Iceberg с учетом операций обновления, чтения и перемещения запросов во времени.

*Глава 4 «Оптимизация производительности таблиц Iceberg»* обсуждает вопросы оптимизации производительности таблиц Apache Iceberg с помощью таких методов, как сжатие и сортировка.

*Глава 5 «Каталоги Iceberg»* подробно объясняет роль каталогов Apache Iceberg и представляет доступные варианты каталогов.

*Глава 6 «Apache Spark»* приводит практические занятия по использованию Apache Spark для управления таблицами Apache Iceberg и взаимодействия с ними.

*Глава 7 «Механизм запросов SQL в Dremio»* исследует платформу Dremio с упором на DDL, DML и оптимизацию для таблиц Apache Iceberg.

*Глава 8 «AWS Glue»* демонстрирует использование каталога AWS Glue и AWS Glue Studio для работы с таблицами Apache Iceberg.

*Глава 9 «Apache Flink»* содержит практические упражнения по использованию Apache Flink для потоковой обработки данных с помощью таблиц Apache Iceberg.

*Глава 10 «Apache Iceberg в промышленной среде»* предлагает подробные сведения об управлении качеством данных в промышленной среде, использовании таблиц метаданных для мониторинга работоспособности и использовании управления версиями таблиц и каталогов для различных нужд.

*Глава 11 «Потоковая передача с помощью Apache Iceberg»* описывает особенности использования таких инструментов, как Apache Spark, Flink и AWS Glue, для потоковой обработки данных в таблицах Iceberg.

*Глава 12 «Управление и безопасность»* исследует приемы управления и обеспечения безопасности на различных уровнях в таблицах Apache Iceberg, таких как хранилище, семантический уровень и каталоги.

*Глава 13 «Миграция на Apache Iceberg»* дает рекомендации по преобразованию существующих наборов данных из различных типов файлов и баз данных в таблицы Apache Iceberg.

*Глава 14 «Примеры использования Apache Iceberg»* приводит примеры использования Apache Iceberg на практике, включая информационные панели бизнес-аналитики и реализацию сбора данных об изменениях.

## **Как пользоваться этой книгой**

Главная цель этой книги – помочь вам лучше понять и получить практические навыки использования Apache Iceberg. Эта книга организована так,

чтобы помочь вам получить полное представление об обсуждаемой технологии и читать главы в том порядке, какой вам больше по душе. Каждая глава полностью самостоятельная и не зависит от других, соответственно, вы сможете погрузиться в конкретные темы или случаи использования без необходимости читать предыдущие главы. Такой подход делает эту книгу как ценным ресурсом для систематического обучения, так и хорошим оперативным справочником.

На протяжении всей книги вам будут встречаться ссылки на фрагменты кода и практические примеры. Чтобы помочь вам в обучении, мы создали специальный репозиторий на GitHub (<https://oreil.ly/supp-guide-apache-iceberg>), организованный по главам, что обеспечивает легкий доступ ко всем необходимым справочным материалам, фрагментам кода и примерам, относящимся к содержанию каждой главы. Репозиторий послужит вам дополнительным инструментом для улучшения вашего обучения и применения идей, обсуждаемых в книге. Чтобы получить еще больше информации, включая дополнительную главу об Iceberg Java/Python API<sup>1</sup> и дополнительный обзор вариантов использования Iceberg<sup>2</sup>, посетите этот репозиторий ([https://oreil.ly/apache-ice\\_more-content](https://oreil.ly/apache-ice_more-content)).

Независимо от того, как вы будете читать это руководство – от начала до конца или сосредоточитесь на отдельных главах, – эта книга задумана как всеобъемлющий и доступный ресурс по Apache Iceberg, обогащенный практическими примерами и сведениями, доступными через наш репозиторий GitHub.

## Отзывы и вопросы

Мы ценим ваши отзывы и вопросы. Если у вас есть какие-либо проблемы, предложения или пожелания, пишите нам по адресу [tech-advocacy@dremio.com](mailto:tech-advocacy@dremio.com). Мы также приглашаем вас подписаться на нас в LinkedIn.

Ниже перечислены дополнительные ресурсы, которые помогут вам больше узнать об Apache Iceberg и принять участие в жизни сообщества:

- каталог статей и ресурсов «Apache Iceberg 101» (<https://oreil.ly/hDv4H>);
- документация Apache Iceberg (<https://iceberg.apache.org>);
- репозиторий Apache Iceberg на GitHub (<https://github.com/apache/iceberg>);
- канал Iceberg Slack (как получить приглашение, см. в документации Iceberg);
- список рассылки Iceberg (порядок регистрации см. в документации к Iceberg);
- страница семинаров Apache Iceberg на LinkedIn (<https://oreil.ly/WghwD>);
- каталог блогов Apache Iceberg (<https://iceberg.apache.org/blogs>).

<sup>1</sup> Включена в книгу как глава 16. – Прим. перев.

<sup>2</sup> Включен в книгу как глава 15. – Прим. перев.

## Обозначения и соглашения, принятые в этой книге

В книге действуют следующие типографские соглашения:

### Курсив

Курсивом выделены новые термины или важные понятия.

### Моноширинный шрифт

Используется для листингов программ, а также в тексте для обозначения таких элементов, как переменные и функции, базы данных, типы данных, переменные среды, операторы и ключевые слова, имена файлов и их расширения.

### Моноширинный жирный шрифт

Показывает команды или другой текст, который пользователь должен ввести самостоятельно.

### Моноширинный курсив

Показывает текст, который должен быть заменен значениями, введенными пользователем, или значениями, определяемыми контекстом.



Так обозначаются советы.



Так обозначаются примечания общего характера.



Так обозначаются предупреждения и предостережения.

## Скачивание исходного кода примеров

Скачать файлы с дополнительной информацией для книг издательства «ДМК Пресс» можно на сайте [www.dmkpress.com](http://www.dmkpress.com) или [www.дмк.рф](http://www.дмк.рф) на странице с описанием соответствующей книги.

Мы высоко ценим, хотя и не требуем, ссылки на наши издания. В ссылке обычно указываются имя автора, название книги, издательство и ISBN, например: «Томер Ширан, Джейсон Хьюз и Алекс Мерсед. Apache Iceberg. Полное руководство. М.: ДМК Пресс, 2025. ISBN 978-5-93700-289-1».

Если вы полагаете, что планируемое использование кода выходит за рамки изложенной выше лицензии, пожалуйста, обратитесь к нам по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com).

## 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 прямо на нашем сайте [www.dmkpress.com](http://www.dmkpress.com), зайдя на страницу книги, и оставить комментарий в разделе «Отзывы и рецензии».

Также можно послать письмо главному редактору по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com), при этом напишите название книги в теме письма.

Если есть тема, в которой вы квалифицированы, и вы 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 удостовериться в качестве наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг – возможно, ошибку в тексте или в коде, – мы будем очень благодарны, если вы сообщите нам о ней. Сделав это, вы избавите других читателей от расстройств и поможете нам улучшить последующие версии этой книги.

Если вы найдете какие-либо ошибки в коде, пожалуйста, сообщите о них главному редактору по адресу [dmkpress@gmail.com](mailto:dmkpress@gmail.com), и мы исправим это в следующих тиражах.

## ***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.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

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