



# Impala in Action: Querying and mining big data

By Ricky Saltzer, Istvan Szegedi, Paul De Schacht

[Download now](#)

[Read Online](#) 

**Impala in Action: Querying and mining big data** By Ricky Saltzer, Istvan Szegedi, Paul De Schacht

Hadoop queries in Pig or Hive can be too slow for real-time data analysis. Impala, an ultra-speedy query engine from Cloudera, supercharges Hadoop by avoiding the typical Map-Reduce overhead and parallelizing queries so that they can run on multiple nodes. This is a big deal for big data, because with Impala, querying Hadoop takes seconds rather than minutes. Impala's dialect is close to standard SQL, and Impala seamlessly accesses HBase and HDFS (Hadoop Distributed File System), allowing considerable freedom in choice of data formats.

*Impala in Action* is a hands-on guide to querying Hadoop using Impala. It starts by comparing Impala to traditional databases and database services on Hadoop. Then it explains Impala's SQL dialect and the basics of data access. Next, it tackles data visualization tasks and provides techniques for securing Impala with Apache Sentry. The book also shows how to embed Impala queries in a Java client and how to connect to JDBC and ODBC clients. Advanced readers will appreciate the deep dive into Impala's architecture and the practical insights into the issues complicated configurations and complex queries can cause.

Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

 [Download Impala in Action: Querying and mining big data ...pdf](#)

 [Read Online Impala in Action: Querying and mining big data ...pdf](#)

# **Impala in Action: Querying and mining big data**

*By Ricky Saltzer, Istvan Szegedi, Paul De Schacht*

**Impala in Action: Querying and mining big data** By Ricky Saltzer, Istvan Szegedi, Paul De Schacht

Hadoop queries in Pig or Hive can be too slow for real-time data analysis. Impala, an ultra-speedy query engine from Cloudera, supercharges Hadoop by avoiding the typical Map-Reduce overhead and parallelizing queries so that they can run on multiple nodes. This is a big deal for big data, because with Impala, querying Hadoop takes seconds rather than minutes. Impala's dialect is close to standard SQL, and Impala seamlessly accesses HBase and HDFS (Hadoop Distributed File System), allowing considerable freedom in choice of data formats.

*Impala in Action* is a hands-on guide to querying Hadoop using Impala. It starts by comparing Impala to traditional databases and database services on Hadoop. Then it explains Impala's SQL dialect and the basics of data access. Next, it tackles data visualization tasks and provides techniques for securing Impala with Apache Sentry. The book also shows how to embed Impala queries in a Java client and how to connect to JDBC and ODBC clients. Advanced readers will appreciate the deep dive into Impala's architecture and the practical insights into the issues complicated configurations and complex queries can cause.

Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

**Impala in Action: Querying and mining big data** By Ricky Saltzer, Istvan Szegedi, Paul De Schacht  
**Bibliography**

- Rank: #10501197 in Books
- Published on: 2015-03-31
- Original language: English
- Number of items: 1
- Dimensions: 9.25" h x 7.36" w x .0" l, .0 pounds
- Binding: Paperback
- 250 pages



[Download Impala in Action: Querying and mining big data ...pdf](#)



[Read Online Impala in Action: Querying and mining big data ...pdf](#)

**Download and Read Free Online Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht**

---

## **Editorial Review**

### **About the Author**

**Ricky Saltzer** is a tools developer at Cloudera, where he writes scalable software to support custom operations engineers. He is an expert in Impala, Hive, and HBase

**Istvan Szegedi** is a lead solution architect working for a large telecommunications company in the UK; he works regularly with a wide array of programming languages and enterprise data architectures.

**Paul De Schacht** is a Data Scientist for the leading provider of IT solutions to the global travel industry. He is at the heart of the travel intelligence platform where Hadoop and Impala are an essential part of daily operations. Paul has 15+ years of experience with distributed software architectures and has a passion for large data sets.

## **Users Review**

### **From reader reviews:**

#### **Sharon Self:**

Hey guys, do you would like to finds a new book to read? May be the book with the title Impala in Action: Querying and mining big data suitable to you? The particular book was written by well-known writer in this era. The actual book untitled Impala in Action: Querying and mining big data is the main of several books this everyone read now. This kind of book was inspired a number of people in the world. When you read this book you will enter the new shape that you ever know prior to. The author explained their thought in the simple way, and so all of people can easily to be aware of the core of this reserve. This book will give you a wide range of information about this world now. To help you see the represented of the world with this book.

#### **Dora Bair:**

Precisely why? Because this Impala in Action: Querying and mining big data is an unordinary book that the inside of the book waiting for you to snap the idea but latter it will jolt you with the secret it inside. Reading this book next to it was fantastic author who also write the book in such incredible way makes the content within easier to understand, entertaining approach but still convey the meaning fully. So , it is good for you because of not hesitating having this anymore or you going to regret it. This amazing book will give you a lot of positive aspects than the other book have such as help improving your proficiency and your critical thinking means. So , still want to postpone having that book? If I have been you I will go to the book store hurriedly.

**John Newton:**

This Impala in Action: Querying and mining big data is completely new way for you who has intense curiosity to look for some information given it relief your hunger of information. Getting deeper you in it getting knowledge more you know or you who still having small amount of digest in reading this Impala in Action: Querying and mining big data can be the light food for you because the information inside this book is easy to get by means of anyone. These books develop itself in the form and that is reachable by anyone, that's why I mean in the e-book web form. People who think that in e-book form make them feel drowsy even dizzy this publication is the answer. So you cannot find any in reading a e-book especially this one. You can find what you are looking for. It should be here for a person. So , don't miss this! Just read this e-book style for your better life and also knowledge.

**Kyle Reese:**

As we know that book is vital thing to add our knowledge for everything. By a reserve we can know everything we really wish for. A book is a pair of written, printed, illustrated or blank sheet. Every year has been exactly added. This publication Impala in Action: Querying and mining big data was filled concerning science. Spend your time to add your knowledge about your science competence. Some people has distinct feel when they reading a new book. If you know how big benefit of a book, you can experience enjoy to read a book. In the modern era like right now, many ways to get book which you wanted.

**Download and Read Online Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht**

**#7E4M2C1TAQU**

# **Read Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht for online ebook**

Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht books to read online.

## **Online Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht ebook PDF download**

**Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht Doc**

**Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht MobiPocket**

**Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht EPub**

**7E4M2C1TAQU: Impala in Action: Querying and mining big data By Ricky Saltzer, Istvan Szegedi, Paul De Schacht**