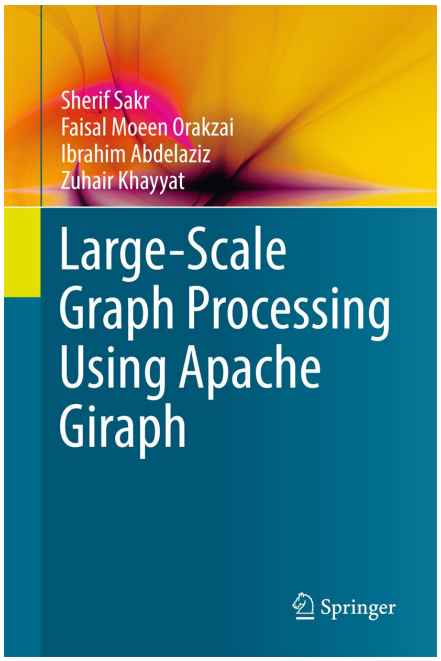

1st ed. 2016, XXV, 197 p. 102 illus., 87 illus. in color.

S. Sakr, F.M. Orakzai, I. Abdelaziz, Z. Khayyat

Large-Scale Graph Processing Using Apache Giraph

- **Describes the fundamental abstractions of the Apache Giraph, its programming models and various techniques**
- **Offers step-by-step coverage of the implementation of several popular and advanced graph analytics algorithms, including related optimization details**
- **All source code presented in the book is available for download from an associated github repository**

This book takes its reader on a journey through Apache Giraph, a popular distributed graph processing platform designed to bring the power of big data processing to graph data. Designed as a step-by-step self-study guide for everyone interested in large-scale graph processing, it describes the fundamental abstractions of the system, its programming models and various techniques for using the system to process graph data at scale, including the implementation of several popular and advanced graph analytics algorithms. The book is organized as follows: Chapter 1 starts by providing a general background of the big data phenomenon and a general introduction to the Apache Giraph system, its abstraction, programming model and design architecture. Next, chapter 2 focuses on Giraph as a platform and how to use it. Based on a sample job, even more advanced topics like monitoring the Giraph application lifecycle and different methods for monitoring Giraph jobs are explained. Chapter 3 then provides an introduction to Giraph programming, introduces the basic Giraph graph model and explains how to write Giraph programs. In turn, Chapter 4 discusses in detail the implementation of some popular graph algorithms including PageRank, connected components, shortest paths and triangle closing. Chapter 5 focuses on advanced Giraph programming, discussing common Giraph algorithmic optimizations, tunable Giraph configurations that determine the system's utilization of the underlying resources, and how to write a custom graph input and output format.