

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics)

By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham

Download now

Read Online ➔

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham

“This book is a critically needed resource for the newly released Apache Hadoop 2.0, highlighting YARN as the significant breakthrough that broadens Hadoop beyond the MapReduce paradigm.”

—From the Foreword by **Raymie Stata, CEO of Altiscale**

The Insider’s Guide to Building Distributed, Big Data Applications with Apache Hadoop™ YARN


Apache Hadoop is helping drive the Big Data revolution. Now, its data processing has been completely overhauled: Apache Hadoop YARN provides resource management at data center scale and easier ways to create distributed applications that process petabytes of data. And now in ***Apache Hadoop™ YARN***, **two Hadoop technical leaders show you how to develop new applications and adapt existing code to fully leverage these revolutionary advances.**

YARN project founder Arun Murthy and project lead Vinod Kumar Vavilapalli demonstrate how YARN increases scalability and cluster utilization, enables new programming models and services, and opens new options beyond Java and batch processing. They walk you through the entire YARN project lifecycle, from installation through deployment.

You’ll find many examples drawn from the authors’ cutting-edge experience—first as Hadoop’s earliest developers and implementers at Yahoo! and now as Hortonworks developers moving the platform forward and helping customers succeed with it.

Coverage includes

- YARN's goals, design, architecture, and components—how it expands the Apache Hadoop ecosystem
- Exploring YARN on a single node
- Administering YARN clusters and Capacity Scheduler
- Running existing MapReduce applications
- Developing a large-scale clustered YARN application
- Discovering new open source frameworks that run under YARN

 [Download Apache Hadoop YARN: Moving beyond MapReduce and Ba...pdf](#)

 [Read Online Apache Hadoop YARN: Moving beyond MapReduce and ...pdf](#)

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics)

By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham

“This book is a critically needed resource for the newly released Apache Hadoop 2.0, highlighting YARN as the significant breakthrough that broadens Hadoop beyond the MapReduce paradigm.”

—From the Foreword by **Raymie Stata, CEO of Altiscale**

The Insider’s Guide to Building Distributed, Big Data Applications with Apache Hadoop™ YARN

Apache Hadoop is helping drive the Big Data revolution. Now, its data processing has been completely overhauled: Apache Hadoop YARN provides resource management at data center scale and easier ways to create distributed applications that process petabytes of data. And now in *Apache Hadoop™ YARN*, **two Hadoop technical leaders show you how to develop new applications and adapt existing code to fully leverage these revolutionary advances.**

YARN project founder Arun Murthy and project lead Vinod Kumar Vavilapalli demonstrate how YARN increases scalability and cluster utilization, enables new programming models and services, and opens new options beyond Java and batch processing. They walk you through the entire YARN project lifecycle, from installation through deployment.

You’ll find many examples drawn from the authors’ cutting-edge experience—first as Hadoop’s earliest developers and implementers at Yahoo! and now as Hortonworks developers moving the platform forward and helping customers succeed with it.

Coverage includes

- YARN’s goals, design, architecture, and components—how it expands the Apache Hadoop ecosystem
- Exploring YARN on a single node
- Administering YARN clusters and Capacity Scheduler
- Running existing MapReduce applications
- Developing a large-scale clustered YARN application
- Discovering new open source frameworks that run under YARN

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph

Niemiec, Jeff Markham Bibliography

- Sales Rank: #909915 in Books
- Brand: Brand: Addison-Wesley Professional
- Published on: 2014-03-29
- Released on: 2014-03-19
- Original language: English
- Number of items: 1
- Dimensions: 8.90" h x .80" w x 7.00" l, 1.15 pounds
- Binding: Paperback
- 400 pages

 [Download Apache Hadoop YARN: Moving beyond MapReduce and Ba ...pdf](#)

 [Read Online Apache Hadoop YARN: Moving beyond MapReduce and ...pdf](#)

Download and Read Free Online Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham

Editorial Review

Review

" This book is a desperately needed resource for administrators, developers, and power-users of the Hadoop YARN framework. It does an excellent job of documenting the (often unknown) history that inevitably lead up to YARN from previous versions of Hadoop, which provides a valuable canvas against which to present the remaining pragmatically-oriented text. Moving from the history of YARN, it wisely jumps right into getting the reader up and running with their own YARN setup (on a single machine or on a larger cluster) such that the rest of the text is not merely conjecturing, but real guidance for a real instance of YARN. Chapters 7 and 8 were the ones I was most looking forward to in the text from the start, as those "core" components of YARN are some of the ones which are least understood and yet concurrently most impacting on performance. They did not disappoint."

- Ellis H. Wilson III, Storage Scientist

About the Author

Arun Murthy has contributed to Apache Hadoop full-time since the inception of the project in early 2006. He is a long-term Hadoop committer and a member of the Apache Hadoop Project Management Committee. Previously, he was the architect and lead of the Yahoo Hadoop MapReduce development team and was ultimately responsible, technically, for providing Hadoop MapReduce as a service for all of Yahoo-- currently running on nearly 50,000 machines. Arun is the founder and architect of the Hortonworks Inc., a software company that is helping to accelerate the development and adoption of Apache Hadoop. Hortonworks was formed by the key architects and core Hadoop committers from the Yahoo! Hadoop software engineering team in June 2011. Funded by Yahoo! and Benchmark Capital, one of the preeminent technology investors, their goal is to ensure that Apache Hadoop becomes the standard platform for storing, processing, managing, and analyzing big data.

Vinod Kumar Vavilapalli has been contributing to Apache Hadoop project full-time since mid-2007. At Apache Software Foundation, he is a long-term Hadoop contributor, Hadoop committer, member of the Apache Hadoop Project Management Committee, and a foundation member. Vinod is a MapReduce and YARN go-to guy at Hortonworks Inc. For more than five years, he has been working on Hadoop. He was involved in HadoopOnDemand, Hadoop-0.20, CapacityScheduler, Hadoop security, and MapReduce, and is now a lead developer and the project lead for Apache Hadoop YARN. Before Hortonworks, he was at Yahoo!, working in the Grid team that made Hadoop what it is today, running at large scale--up to tens of thousands of nodes. Vinod loves reading books of all kinds and is passionate about using computers to change the world for better, bit by bit. He has a bachelor's degree in computer science and engineering from the Indian Institute of Technology Roorkee. He can be reached at twitter handle @tshooter.

Douglas Eadline, Ph.D., began his career as a practitioner and a chronicler of the Linux Cluster HPC revolution and now documents big data analytics. Starting with the first Beowulf How To document, Doug has written hundreds of articles, white papers, and instructional documents covering virtually all aspects of HPC computing. Prior to starting and editing the popular ClusterMonkey.net website in 2005, he served as editor-in-chief for *ClusterWorld* magazine, and was senior HPC editor for *Linux Magazine*. Currently, he is a consultant to the HPC industry and writes a monthly column in *HPC Admin* magazine. Both clients and

readers have recognized Doug's ability to present a "technological value proposition" in a clear and accurate style. He has practical, hands-on experience in many aspects of HPC, including hardware and software design, benchmarking, storage, GPU, cloud, and parallel computing. He is the author of *Hadoop Fundamentals LiveLessons* (video) from Addison-Wesley.

Joseph Niemiec is a big data solutions engineer whose focus is on designing Hadoop solutions for many *Fortune 1000* companies. In this position, Joseph has worked with customers to build multiple YARN applications providing a unique perspective on moving customers beyond batch processing, and has worked on YARN development directly. An avid technologist, Joseph has been focused on technology innovations since 2001. His interest in data analytics originally started in game score optimization as a teenager, and has shifted to helping customers uptake new technology innovations such as Hadoop and, most recently, building new data applications using YARN.

Jeff Markham is a solution engineer at Hortonworks Inc., the company promoting open source Hadoop. Previously, he was with VMware, Red Hat, and IBM, helping companies build distributed applications with distributed data. He has written articles on Java application development and has spoken at several conferences and to Hadoop User Groups. Jeff is a contributor to Apache Pig and Apache HDFS.

Users Review

From reader reviews:

Andrew Drake:

Book will be written, printed, or outlined for everything. You can understand everything you want by a e-book. Book has a different type. As you may know that book is important matter to bring us around the world. Adjacent to that you can your reading skill was fluently. A publication Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) will make you to be smarter. You can feel more confidence if you can know about anything. But some of you think that will open or reading some sort of book make you bored. It is not necessarily make you fun. Why they might be thought like that? Have you searching for best book or ideal book with you?

Carl Guerra:

In this 21st hundred years, people become competitive in every way. By being competitive today, people have do something to make all of them survives, being in the middle of often the crowded place and notice by means of surrounding. One thing that often many people have underestimated the item for a while is reading. Yeah, by reading a book your ability to survive increase then having chance to stand than other is high. To suit your needs who want to start reading the book, we give you this specific Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) book as nice and daily reading publication. Why, because this book is more than just a book.

Richard Kowalski:

Reading a guide can be one of a lot of task that everyone in the world loves. Do you like reading book and so. There are a lot of reasons why people enjoyed. First reading a guide will give you a lot of new details.

When you read a e-book you will get new information mainly because book is one of a number of ways to share the information or their idea. Second, studying a book will make a person more imaginative. When you examining a book especially fiction book the author will bring one to imagine the story how the figures do it anything. Third, you may share your knowledge to other individuals. When you read this Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics), it is possible to tells your family, friends in addition to soon about yours publication. Your knowledge can inspire others, make them reading a publication.

Katie Broadnax:

Do you have something that you like such as book? The e-book lovers usually prefer to choose book like comic, short story and the biggest some may be novel. Now, why not hoping Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) that give your entertainment preference will be satisfied by means of reading this book. Reading routine all over the world can be said as the way for people to know world a great deal better then how they react to the world. It can't be claimed constantly that reading practice only for the geeky person but for all of you who wants to end up being success person. So , for every you who want to start examining as your good habit, you may pick Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) become your own personal starter.

Download and Read Online Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham #LJXBHEDN4FC

Read Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham for online ebook

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham 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 Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham books to read online.

Online Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham ebook PDF download

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham Doc

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham Mobipocket

Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham EPub

LJXBHEDN4FC: Apache Hadoop YARN: Moving beyond MapReduce and Batch Processing with Apache Hadoop 2 (Addison-Wesley Data & Analytics) By Arun Murthy, Vinod Vavilapalli, Douglas Eadline, Joseph Niemiec, Jeff Markham