

Apache The Definitive

Authoritative Answers to All Your Apache Questions--Now Updated to Cover Apache 2.0 Linux Apache Web Server Administration is the most complete, most advanced guide to the Apache Web server you'll find anywhere. Written by a leading Apache expert--and now updated to cover Apache 2.0--this book teaches you, step-by-step, all the standard and advanced techniques you need to know to administer Apache on a Linux box. Hundreds of clear, consistent examples illustrate these techniques in detail--so you stay on track and accomplish all your goals. Coverage includes: * Compiling Apache from source code * Creating and hosting virtual web sites * Using Server-Side Includes to create Web pages with dynamic content * Using Apache directives to configure your site * Extending Apache using add-on modules * Using the Common Gateway Interface for web programming * Enhancing the performance of CGI programs with FastCGI and mod_perl * Installing Apache support for PHP * Extending Apache to run Java servlets or Java Server Pages * Attaching Apache to a database server * Using URL rewriting for increased request-handling flexibility * Implementing user authentication * Adding Secure Sockets Layer for enhanced system security * Customizing Apache's log formats

The Craig Hunt Linux Library The Craig Hunt Linux Library provides in-depth, advanced coverage of the key topics for Linux administrators. Topics include Samba, System Administration, DNS Server Administration, Network Servers, Security, and Sendmail. Each book in the series is either written by or meticulously reviewed by Craig Hunt to ensure the highest quality and most complete coverage for networking professionals working specifically in Linux environments.

"Do you learn best by example and experimentation? This book is ideal. Have your favorite editor and compiler ready--you'll encounter example code you'll want to try right away. You've picked the right book--this is sure to become the de facto standard guide to writing Apache modules." --Rich Bowen, coauthor, Apache Administrators Handbook, Apache Cookbook, and The Definitive Guide to Apache mod_rewrite

"A first-rate guide to getting the most out of Apache as a modular application platform--sure to become a must-read for any Apache programmer, from beginner to experienced professional. It builds up carefully and meticulously from the absolute basics, while including chapters on everything from the popular Apache DBD Framework to best practices, security, and debugging." --Noirin Plunkett, documentation committer to the Apache httpd project, and member of the ASF conference committee

The Only Comprehensive Guide to Developing Apache 2.x Modules and Applications Apache is more than the world's most popular Web server--it's also an extraordinarily powerful and extensible development platform. Now, ApacheTutor.org's Nick Kew has written The Apache Modules Book, the first start-to-finish, example-rich guide for every developer who wants to make the most of Apache. Kew begins with detailed, accessible introductions to Apache's architecture and API, then illuminates all the

techniques you'll need, from request processing through code security. He brings together the best of both worlds: powerful C-based techniques for accomplishing tasks Perl or PHP can't handle, implemented with tools that deliver all the productivity you'd expect from higher-level languages. Utilizing realistic code samples, Kew introduces techniques documented in no other book-and, often, nowhere else at all. Coverage includes Using Apache Portable Runtime (APR) to streamline C development and avoid its pitfalls Leveraging Apache DBD to build applications far more scalable than classic LAMP software Working with the latest Apache 2.x features: filter modules, XML support, and smart proxies Mastering best practices, from thread safety to multi-platform development Utilizing the Apache Authentication Framework Tracing and debugging problems in both Apache and your custom modules Foreword Preface Acknowledgments About the Author Chapter 1 Applications Development with Apache Chapter 2 The Apache Platform and Architecture Chapter 3 The Apache Portable Runtime Chapter 4 Programming Techniques and Caveats Chapter 5 Writing a Content Generator Chapter 6 Request Processing Cycle and Metadata Handlers Chapter 7 AAA: Access, Authentication, and Authorization Chapter 8 Filter Modules Chapter 9 Configuration for Modules Chapter 10 Extending the API Chapter 11 The Apache Database Framework Chapter 12 Module Debugging Appendix A Apache License Appendix B Contributor License Agreements Appendix C Hypertext Transfer Protocol: HTTP/1.1 Index About the Web Site

ApacheTutor.org contains code examples from the book, all designed for easy use and integration into existing applications.

This book is a step-by-step tutorial for anyone wanting to learn Apache Tomcat 7 from scratch. There are plenty of illustrations and examples to escalate you from a novice to an expert with minimal strain. If you are a J2EE administrator, migration administrator, technical architect, or a project manager for a web hosting domain, and are interested in Apache Tomcat 7, then this book is for you. If you are someone responsible for installation, configuration, and management of Tomcat 7, then too, this book will be of help to you.

Apache, a superior system that rivals almost any other UNIX-based HTTP server, runs on about half the world's existing Web sites and is rapidly increasing market share. This book, officially authorized by the Apache Group, is the only complete guide on the market today describing how to obtain, set up, and secure the Apache software.

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and

debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

ApacheThe Definitive Guide"O'Reilly Media, Inc."

Jakarta Tomcat is not only the most commonly used open source servlet engine today, it's become the de facto standard by which other servlet engines are measured. Powerful and flexible, it can be used as a stand-alone web server or in conjunction with another server, like Apache or IIS, to run servlets or JSPs. But mastery of Tomcat is not easy: because it's as complex as it is complete.

Tomcat: The Definitive Guide answers vexing questions that users, administrators, and developers alike have been asking. This concise guide provides much needed information to help harness Tomcat's power and wealth of features. Tomcat: The Definitive Guide offers something for everyone who uses Tomcat. System and network administrators will find detailed instructions on installation, configuration, and maintenance. For users, it supplies insightful information on how to deploy Tomcat. And seasoned enterprise Java developers will have a complete reference to setting up, running, and using this powerful software. The book begins with an introduction to the Tomcat server and includes an overview of the three types of server configurations: stand-alone, in-process, and out-of-process. The authors show how directories are laid out, cover the initial setup, and describe how to set the environment variables and modify the configuration files, concluding with common errors, problems, and solutions. In subsequent chapters, they cover: The server.xml configuration file Java Security manager Authentication schemes and Tomcat users The Secure Socket Layer (SSL) Tomcat JDBC Realms Installing servlets and Java Server Pages Integrating Tomcat with Apache Advanced Tomcat configuration and much more. Tomcat: The Definitive Guide covers all major platforms, including Windows, Solaris, Linux, and Mac OS X, contains details on Tomcat configuration files, and has a quick-start guide to get developers up and running with Java servlets and JavaServer Pages. If you've struggled with this powerful yet demanding technology in the past, this book will provide the answers you need.

Imagine what you could do if scalability wasn't a problem. With this hands-on guide, you'll learn how the Cassandra database management system handles hundreds of terabytes of data while remaining highly available across multiple data centers. This expanded second edition—updated for Cassandra 3.0—provides the technical details and practical examples you need to put this database to

work in a production environment. Authors Jeff Carpenter and Eben Hewitt demonstrate the advantages of Cassandra's non-relational design, with special attention to data modeling. If you're a developer, DBA, or application architect looking to solve a database scaling issue or future-proof your application, this guide helps you harness Cassandra's speed and flexibility. Understand Cassandra's distributed and decentralized structure Use the Cassandra Query Language (CQL) and cqlsh—the CQL shell Create a working data model and compare it with an equivalent relational model Develop sample applications using client drivers for languages including Java, Python, and Node.js Explore cluster topology and learn how nodes exchange data Maintain a high level of performance in your cluster Deploy Cassandra on site, in the Cloud, or with Docker Integrate Cassandra with Spark, Hadoop, Elasticsearch, Solr, and Lucene

This book provides an introduction to Apache HTTP Server - a free, open-source web server. Apache is the most popular web server software on the Internet; it is estimated that 50% of all active websites use Apache as their web server. You will learn how to download and install Apache HTTP Server on your Windows and Linux system, how to configure Apache as a web server, a proxy server, and a reverse proxy server. You will also learn to set up SSL and password-protect directories on your web server. Later in the book we explain modules and how you can use them to add more features to your web server. The topics covered in this book are: downloading and installing Apache HTTP Server on Ubuntu and Windows understanding Apache configuration files using virtual hosts to hold multiple websites on a single server enabling SSL for secure connections what are modules and how to use them to expand Apache functionality configuring Apache as a forward or reverse proxy redirecting URLs Note that this book uses Ubuntu as an underlying Linux distribution, so some of the commands and configurations files might differ if you are using some other non-Debian based Linux distribution.

Readers will find hundreds of immediate solutions for turning Apache into a powerhouse Web server. Key topics include setting up a virtual Web site, mastering security, creating optimized CGI scripts, tuning and customizing Apache servers, using the Apache Module API, setting up Apache as a proxy server, and setting up commerce sites.

If you're like most R users, you have deep knowledge and love for statistics. But as your organization continues to collect huge amounts of data, adding tools such as Apache Spark makes a lot of sense. With this practical book, data scientists and professionals working with large-scale data applications will learn how to use Spark from R to tackle big data and big compute problems. Authors Javier Luraschi, Kevin Kuo, and Edgar Ruiz show you how to use R with Spark to solve different data analysis problems. This book covers relevant data science topics, cluster computing, and issues that should interest even the most advanced users. Analyze, explore, transform, and visualize data in Apache Spark

with R Create statistical models to extract information and predict outcomes; automate the process in production-ready workflows Perform analysis and modeling across many machines using distributed computing techniques Use large-scale data from multiple sources and different formats with ease from within Spark Learn about alternative modeling frameworks for graph processing, geospatial analysis, and genomics at scale Dive into advanced topics including custom transformations, real-time data processing, and creating custom Spark extensions

Every enterprise application creates data, whether it's log messages, metrics, user activity, outgoing messages, or something else. And how to move all of this data becomes nearly as important as the data itself. If you're an application architect, developer, or production engineer new to Apache Kafka, this practical guide shows you how to use this open source streaming platform to handle real-time data feeds. Engineers from Confluent and LinkedIn who are responsible for developing Kafka explain how to deploy production Kafka clusters, write reliable event-driven microservices, and build scalable stream-processing applications with this platform. Through detailed examples, you'll learn Kafka's design principles, reliability guarantees, key APIs, and architecture details, including the replication protocol, the controller, and the storage layer. Understand publish-subscribe messaging and how it fits in the big data ecosystem. Explore Kafka producers and consumers for writing and reading messages Understand Kafka patterns and use-case requirements to ensure reliable data delivery Get best practices for building data pipelines and applications with Kafka Manage Kafka in production, and learn to perform monitoring, tuning, and maintenance tasks Learn the most critical metrics among Kafka's operational measurements Explore how Kafka's stream delivery capabilities make it a perfect source for stream processing systems

Apache remains the most widely used Web server platform, running more than half of the world's Web sites. This pocket reference guide features up-to-date information for Apache administrators.

* The Definitive Guide to Apache, Third Edition is a new revision of Peter Wainwright's best-selling (actually, the highest-selling) book on Apache configuration and administration. * Provides multiple methods and examples of how to solve a particular problem; compares and contrasts the methods rather than recommending a single, "best" solution. * Demonstrates how to use many of the most-popular scripting languages such as PHP, Perl, and Python to create dynamic web sites.

Oracle Corporation has broadened its development platform, integrating open standards such as Java and XML into the heart of the Oracle 8i database. This extended programming environment continues to exploit the qualities of scalability, reliability and efficiency of the world's most successful data management software, but at the same time it provides new challenges and opportunities to programmers.

Perform fast interactive analytics against different data sources using the Trino high-performance distributed SQL query engine. With this practical guide, you'll learn how to conduct analytics on data where it lives, whether it's Hive, Cassandra, a relational database, or a proprietary data store. Analysts, software engineers, and production engineers will learn how to manage, use, and even develop with Trino. Initially developed by Facebook, open source Trino is now used by Netflix, Airbnb, LinkedIn, Twitter, Uber, and many other companies. Matt Fuller, Manfred Moser, and Martin Traverso show you how a single Trino query can combine data from multiple sources to allow for analytics across your entire organization. Get started: Explore Trino's use cases and learn about tools that will help you connect to Trino and query data Go deeper: Learn Trino's internal workings, including how to connect to and query data sources with support for SQL statements, operators, functions, and more Put Trino in production: Secure Trino, monitor workloads, tune queries, and connect more applications; learn how other organizations apply Trino

If you're looking for a scalable storage solution to accommodate a virtually endless amount of data, this book shows you how Apache HBase can fulfill your needs. As the open source implementation of Google's BigTable architecture, HBase scales to billions of rows and millions of columns, while ensuring that write and read performance remain constant. Many IT executives are asking pointed questions about HBase. This book provides meaningful answers, whether you're evaluating this non-relational database or planning to put it into practice right away. Discover how tight integration with Hadoop makes scalability with HBase easier Distribute large datasets across an inexpensive cluster of commodity servers Access HBase with native Java clients, or with gateway servers providing REST, Avro, or Thrift APIs Get details on HBase's architecture, including the storage format, write-ahead log, background processes, and more Integrate HBase with Hadoop's MapReduce framework for massively parallelized data processing jobs Learn how to tune clusters, design schemas, copy tables, import bulk data, decommission nodes, and many other tasks

Ant is the premiere build management tool for use in Java environments. Unlike traditional build management tools such as GNU Make, Ant is itself written in Java, is platform independent, and interfaces well with the utilities in Sun's Java software development kit (SDK). In addition to being platform independent, Ant is also independent of the integrated development environment (IDE) being used. IDE independence is important for open source projects (or other projects) in which the various developers might use different IDEs. Using Ant, Java developers can: define build chunks, the results that they must produce, and the dependencies between them; automatically retrieve source code from source control systems such as PVCS; build applications by having Ant compile the necessary source files in the proper order. Ant build files are written using XML---a well-established standard---so programmers using Ant are not required to learn yet another scripting language. They will likely already know XML, and will be able to leverage that knowledge. Ant is an open source project, and part of the Jakarta project. Jakarta is Sun's open source reference implementation for the JSP and Servlets specifications, and is part of the Apache group's work.

Describes the features and functions of Apache Hive, the data infrastructure for Hadoop. Sams Teach Yourself Apache in 24 Hours covers the installation, configuration, and ongoing administration of the Apache Web server, the most popular Internet Web server. It covers both the 1.3 and the new 2.0 versions of Apache. Using a hands-on, task-oriented format, it concentrates on the most popular features and common quirks of the server. The first part of

Download Free Apache The Definitive

the book helps the reader build, configure, and get started with Apache. After completing these chapters the reader will be able to start, stop, and monitor the Web server. He also will be able to serve both static content and dynamic content, customize the logs, and restrict access to certain parts of the Web server. The second part of the book explains in detail the architecture of Apache and how to extend the server via third-party modules like PHP and Tomcat. It covers server performance and scalability, content management, and how to set up a secure server with SSL.

Describes the history of the Web server platform and covers downloading and compiling, configuring and running the program on UNIX, writing specialized modules, and establishing security routines.

Three of CouchDB's creators show you how to use this document-oriented database as a standalone application framework or with high-volume, distributed applications. With its simple model for storing, processing, and accessing data, CouchDB is ideal for web applications that handle huge amounts of loosely structured data. That alone would stretch the limits of a relational database, yet CouchDB offers an open source solution that's reliable, scales easily, and responds quickly. CouchDB works with self-contained data that has loose or ad-hoc connections. It's a model that fits many real-world items, such as contacts, invoices, and receipts, but you'll discover that this database can easily handle data of any kind. With this book, you'll learn how to work with CouchDB through its RESTful web interface, and become familiar with key features such as simple document CRUD (create, read, update, delete), advanced MapReduce, deployment tuning, and more. Understand the basics of document-oriented storage and manipulation Interact with CouchDB entirely through HTTP using its RESTful interface Model data as self-contained JSON documents Handle evolving data schemas naturally Query and aggregate data in CouchDB using MapReduce views Replicate data between nodes Tune CouchDB for increased performance and reliability

Gives an extensive overview of Perl language and its syntax and is a complete reference for all Perl functions, operators and standard library modules. Provides an explanation of Perl references and complex data structures and a detailed account of Perl's object oriented features. Department.

Why a book about logs? That's easy: the humble log is an abstraction that lies at the heart of many systems, from NoSQL databases to cryptocurrencies. Even though most engineers don't think much about them, this short book shows you why logs are worthy of your attention. Based on his popular blog posts, LinkedIn principal engineer Jay Kreps shows you how logs work in distributed systems, and then delivers practical applications of these concepts in a variety of common uses—data integration, enterprise architecture, real-time stream processing, data system design, and abstract computing models. Go ahead and take the plunge with logs; you're going love them. Learn how logs are used for programmatic access in databases and distributed systems Discover solutions to the huge data integration problem when more data of more varieties meet more systems Understand why logs are at the heart of real-time stream processing Learn the role of a log in the internals of online data systems Explore how Jay Kreps applies these ideas to his own work on data infrastructure systems at LinkedIn Counsels programmers and administrators for big and small organizations on how to work with large-scale application datasets using Apache Hadoop, discussing its capacity for storing and processing large amounts of data while demonstrating best practices for building reliable and scalable distributed systems.

What is this book about? PHP, Apache, and MySQL are the three key open source technologies that form the basis for most active Web servers. This book takes you step-by-step through understanding each — using it and combining it with the other two on both Linux and Windows servers. This book guides you through creating your own sites

Download Free Apache The Definitive

using the open source AMP model. You discover how to install PHP, Apache, and MySQL. Then you create PHP Web pages, including database management and security. Finally, you discover how to integrate your work with e-commerce and other technologies. By building different types of Web sites, you progress from setting up simple database tables to tapping the full potential of PHP, Apache, and MySQL. When you're finished, you will be able to create well-designed, dynamic Web sites using open source tools. What does this book cover? Here's what you will learn from this book:

- How PHP server-side scripting language works for connecting HTML-based Web pages to a backend database
- Syntax, functions, and commands for PHP, Apache, and MySQL
- Methods and techniques for building user-friendly forms
- How to easily store, update, and access information using MySQL
- Ways to allow the user to edit a database
- E-commerce applications using these three technologies
- How to set up user logins, profiles, and personalizations
- Proper protocols for error handling

Who is this book for? This book is for beginners who are new to PHP and who need to learn quickly how to create Web sites using open source tools. Some basic HTML knowledge is helpful but not essential.

Introduces the build tool for Java application development, covering both user defined and built-in tasks.

The freeware Apache web server runs on about half of the world's existing web sites, and it is rapidly increasing in popularity. Apache: The Definitive Guide, written and reviewed by key members of the Apache Group, is the only complete guide on the market today that describes how to obtain, set up, and secure the Apache software. Apache was originally based on code and ideas found in the most popular HTTP server of the time: NCSA httpd 1.3 (early 1995). It has since evolved into a far superior system that can rival (and probably surpass) almost any other Unix-based HTTP server in terms of functionality, efficiency, and speed. The new version now includes support for Win32 systems. This new second edition of Apache: The Definitive Guide fully describes Windows support and all the other Apache 1.3 features. Contents include:

- The history of the Apache Group
- Obtaining and compiling the server
- Configuring and running Apache on Unix and Windows, including such topics as directory structures, virtual hosts, and CGI programming
- The Apache 1.3 Module API
- Apache security
- A complete list of configuration directives

With Apache: The Definitive Guide, web administrators new to Apache can get up to speed more quickly than ever before by working through the tutorial demo. Experienced administrators and CGI programmers, and web administrators moving from Unix to Windows, will find the reference sections indispensable. Apache: The Definitive Guide is the definitive documentation for the world's most popular web server. Includes CD-ROM with Apache manuals and demo sites discussed in the book.

Ready to unlock the power of your data? With this comprehensive guide, you'll learn how to build and maintain reliable, scalable, distributed systems with Apache Hadoop. This book is ideal for programmers looking to analyze datasets of any size, and for administrators who want to set up and run Hadoop clusters. You'll find illuminating case studies that demonstrate how Hadoop is used to solve specific problems. This third edition covers recent changes to Hadoop, including material on the new MapReduce API, as well as MapReduce 2 and its more flexible execution model (YARN). Store large datasets with the Hadoop Distributed File System (HDFS) Run

Download Free Apache The Definitive

distributed computations with MapReduce Use Hadoop's data and I/O building blocks for compression, data integrity, serialization (including Avro), and persistence Discover common pitfalls and advanced features for writing real-world MapReduce programs Design, build, and administer a dedicated Hadoop cluster—or run Hadoop in the cloud Load data from relational databases into HDFS, using Sqoop Perform large-scale data processing with the Pig query language Analyze datasets with Hive, Hadoop's data warehousing system Take advantage of HBase for structured and semi-structured data, and ZooKeeper for building distributed systems

This expert-written book covers the open source Apache MyFaces project, which is the most popular implementation of JavaServer Faces, a Web framework put forth by Sun Microsystems. The text introduces the basics of MyFaces and the JSF Standard and goes beyond fundamentals to provide a thorough understanding of the JSF lifecycle. Readers will learn how to build real-world AJAX components, and how to leverage Oracle ADF Faces components within applications. The book is an ideal reference for professional Java and Web developers looking to develop real world applications as it focuses on practical aspects such as scalability, design, optimization and configurability.

Explaining how to enhance the capabilities of the Apache Web server, a guide to Web programming discusses the design of Apache, mod perl, and the Apache API and demonstrates how to use them to rewrite CGI scripts, convert file formats, and more. Original. (Intermediate).

Summary Kafka Streams in Action teaches you everything you need to know to implement stream processing on data flowing into your Kafka platform, allowing you to focus on getting more from your data without sacrificing time or effort. Foreword by Neha Narkhede, Cocreator of Apache Kafka Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology Not all stream-based applications require a dedicated processing cluster. The lightweight Kafka Streams library provides exactly the power and simplicity you need for message handling in microservices and real-time event processing. With the Kafka Streams API, you filter and transform data streams with just Kafka and your application. About the Book Kafka Streams in Action teaches you to implement stream processing within the Kafka platform. In this easy-to-follow book, you'll explore real-world examples to collect, transform, and aggregate data, work with multiple processors, and handle real-time events. You'll even dive into streaming SQL with KSQL! Practical to the very end, it finishes with testing and operational aspects, such as monitoring and debugging. What's inside Using the KStreams API Filtering, transforming, and splitting data Working with the Processor API Integrating with external systems About the Reader Assumes some experience with distributed systems. No knowledge of Kafka or streaming applications required. About the Author Bill Bejeck is a Kafka Streams contributor and Confluent engineer with over 15 years of software development experience. Table of Contents PART 1 - GETTING STARTED WITH KAFKA STREAMS Welcome to Kafka Streams Kafka quicklyPART 2 - KAFKA STREAMS DEVELOPMENT Developing Kafka Streams Streams and state The KTable API The Processor APIPART 3 - ADMINISTERING KAFKA STREAMS Monitoring and performance Testing a Kafka Streams applicationPART 4 - ADVANCED CONCEPTS WITH KAFKA STREAMS Advanced applications with Kafka StreamsAPPENDIXES

Appendix A - Additional configuration information Appendix B - Exactly once semantics
* The Definitive Guide to Apache mod_rewrite is the first book devoted to mastering mod_rewrite, Apache's most powerful (and confusing) module. * It is authored by noted Apache expert and Apache Software Foundation member Rich Bowen, who draws on several years of extensive experience administering the Apache server as well as speaking and writing regularly on the topic. * This book covers all matters pertinent to mod_rewrite: installation and configuration, basic and conditional rewrites, access control, maintaining virtual hosts, and proxying.

Learn how to use the Apache Hadoop projects, including MapReduce, HDFS, Apache Hive, Apache HBase, Apache Kafka, Apache Mahout, and Apache Solr. From setting up the environment to running sample applications each chapter in this book is a practical tutorial on using an Apache Hadoop ecosystem project. While several books on Apache Hadoop are available, most are based on the main projects, MapReduce and HDFS, and none discusses the other Apache Hadoop ecosystem projects and how they all work together as a cohesive big data development platform. What You Will Learn: Set up the environment in Linux for Hadoop projects using Cludera Hadoop Distribution CDH 5 Run a MapReduce job Store data with Apache Hive, and Apache HBase Index data in HDFS with Apache Solr Develop a Kafka messaging system Stream Logs to HDFS with Apache Flume Transfer data from MySQL database to Hive, HDFS, and HBase with Sqoop Create a Hive table over Apache Solr Develop a Mahout User Recommender System Who This Book Is For: Apache Hadoop developers. Pre-requisite knowledge of Linux and some knowledge of Hadoop is required.

What could you do with data if scalability wasn't a problem? With this hands-on guide, you'll learn how Apache Cassandra handles hundreds of terabytes of data while remaining highly available across multiple data centers -- capabilities that have attracted Facebook, Twitter, and other data-intensive companies. Cassandra: The Definitive Guide provides the technical details and practical examples you need to assess this database management system and put it to work in a production environment. Author Eben Hewitt demonstrates the advantages of Cassandra's nonrelational design, and pays special attention to data modeling. If you're a developer, DBA, application architect, or manager looking to solve a database scaling issue or future-proof your application, this guide shows you how to harness Cassandra's speed and flexibility. Understand the tenets of Cassandra's column-oriented structure Learn how to write, update, and read Cassandra data Discover how to add or remove nodes from the cluster as your application requires Examine a working application that translates from a relational model to Cassandra's data model Use examples for writing clients in Java, Python, and C# Use the JMX interface to monitor a cluster's usage, memory patterns, and more Tune memory settings, data storage, and caching for better performance

It takes a book as versatile as its subject to cover Apache Tomcat, the popular open source Servlet and JSP container and high performance web server. Tomcat: The Definitive Guide is a valuable reference for administrators and webmasters, a useful guide for programmers who want to use Tomcat as their web application server during development or in production, and an excellent introduction for anyone interested in Tomcat. Updated for the latest version of Tomcat, this new edition offers a complete guide to installing, configuring, maintaining and securing this servlet container. In fact,

Download Free Apache The Definitive

with such a wealth of new information, this is essentially a new book rather than a simple revision. You will find details for using Tomcat on all major platforms, including Windows, Linux, OS X, Solaris, and FreeBSD, along with specifics on Tomcat configuration files, and step-by-step advice for deploying and running web applications. This book offers complete information for: Installation and startup procedures Configuring Tomcat-including realms, roles, users, servlet sessions, and JNDI resources including JDBC DataSources Deploying web applications-individual servlets and JSP pages, and web application archive files Tuning Tomcat to measure and improve performance Integrating Tomcat with Apache Web Server Securing Tomcat to keep online thugs at bay Tomcat configuration files-server.xml and web.xml, and more Debugging and Troubleshooting-diagnosing problems with Tomcat or a web application Compiling your own Tomcat, rather than using the pre-built release Running two or more Tomcat servlet containers in parallel This book also offers an overview of the Tomcat open source project's community resources, including docs, mailing lists, and more. Community interest fueled a strong demand for a Tomcat guide from O'Reilly. The result clearly exceeds expectations.

Covers topics including HTTP methods and status codes, optimizing proxies, designing web crawlers, content negotiation, and load-balancing strategies.

Apache is far and away the most widely used web server platform in the world. Both free and rock-solid, it runs more than half of the world's web sites, ranging from huge e-commerce operations to corporate intranets and smaller hobby sites, and it continues to maintain its popularity, drawing new users all the time. If you work with Apache on a regular basis, you have plenty of documentation on installing and configuring your server, but where do you go for help with the day-to-day stuff, like adding common modules or fine-tuning your activity logging?The Apache Cookbook is a collection of problems, solutions, and practical examples for webmasters, web administrators, programmers, and everyone else who works with Apache. For every problem addressed in the book, there's a worked-out solution or "recipe"--short, focused pieces of code that you can use immediately. But this book offers more than cut-and-paste code. You also get explanations of how and why the code works, so you can adapt the problem-solving techniques to similar situations.The recipes in the Apache Cookbook range from simple tasks, such installing the server on Red Hat Linux or Windows, to more complex tasks, such as setting up name-based virtual hosts or securing and managing your proxy server. The two hundred plus recipes in the book cover additional topics such as: Security Aliases, Redirecting, and Rewriting CGI Scripts, the suexec Wrapper, and other dynamic content techniques Error Handling SSL Performance The impressive collection of useful code in this book is a guaranteed timesaver for all Apache users, from novices to advanced practitioners. Instead of poking around mailing lists, online documentation, and other sources, you can rely on the Apache Cookbook for quick solutions to common problems, and then you can spend your time and energy where it matters most.

[Copyright: 5b351cd2b85f5d6cd015d8eb2796772f](#)