

1. Record Nr.	UNINA9910821609303321
Autore	Anoshin Dmitry
Titolo	Learning Hunk : visualize and analyze your Hadoop data using Hunk // Dmitry Anoshin, Sergey Sheypak
Pubbl/distr/stampa	Birmingham : , : Packt Publishing, , 2015
ISBN	1-78528-302-2
Descrizione fisica	1 online resource (156 p.)
Collana	Community experience distilled
Soggetti	Big data Non-relational databases
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di contenuto	Cover; Copyright; Credits; About the Authors; About the Reviewer; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Meet Hunk; Big data analytics; The big problem; The elegant solution; Supporting SPL; Intermediate results; Getting to know Hunk; Splunk versus Hunk; Hunk architecture; Connecting to Hadoop; Advance Hunk deployment; Native versus virtual indexes; Native indexes; Virtual index; External result provider; Computation models; Data streaming; Data reporting; Mixed mode; Hunk security; One Hunk user to one Hadoop user; Many Hunk users to one Hadoop user Hunk user(s) to the same Hadoop user with different queues Setting up Hadoop; Starting and using a virtual machine with CDH5; SSH user; MySQL; Starting the VM and cluster in VirtualBox; Big data use case; Importing data from RDBMS to Hadoop using Sqoop; Telecommunications - SMS, Call, and Internet dataset from dandelion.eu; Milano grid map; CDR aggregated data import process; Periodical data import from MySQL using Sqoop and Oozie; Problems to solve; Summary; Chapter 2: Explore Hadoop Data with Hunk; Setting up Hunk; Extracting Hunk to a VM; Setting up Hunk variables and configuration files Running Hunk for the first time Setting up a data provider and virtual index for CDR data; Setting up a connection to Hadoop; Setting up a virtual index for data stored in Hadoop; Accessing data through a virtual index; Exploring data; Creating reports; The top five browsers

report; Top referrers; Site errors report; Creating alerts; Creating a dashboard; Controlling security with Hunk; The default Hadoop security; One Hunk user to one Hadoop user; Summary; Chapter 3: Meeting Hunk Features; Knowledge objects; Field aliases; Calculated fields; Field extractions; Tags; Event type Workflow actionsMacros; Data model; Add auto-extracting fields; Adding GeoIP attributes; Other ways to add attributes; Introducing Pivot; Summary; Chapter 4: Adding Speed to Reports; Big data performance issues; Hunk report acceleration; Creating a virtual index; Streaming mode; Creating an acceleration search; What's going on in Hadoop?; Report acceleration summaries; Reviewing summary details; Managing report accelerations; Hunk accelerations limits; Summary; Chapter 5: Customizing Hunk; What we are going to do with the Splunk SDK; Supported languages; Solving problems; REST API The implementation planThe conclusion; Dashboard customization using Splunk Web Framework; Functionality; A description of time-series aggregated CDR data; Source data; Creating a virtual index for Milano CDR; Creating a virtual index for the Milano grid; Creating a virtual index using sample data; Implementation; Querying the visualization; Downloading the application; Custom Google Maps; Page layout; Linear gradients and bins for the activity value; Custom map components; Other components; The final result; Summary; Chapter 6: Discovering Hunk Integration Apps; What is Mongo?; Installation Installing the Mongo app
