

User Manual for Hadoop Cluster Sample Application

Abstract:

This document captures the detailed functional specifications for Hadoop Sample Application that perform certain map-reduce transaction on the provisioned cluster.

Table of Contents

1. Objective	3
2. Requirement.....	3
3. User Guide	3

1. OBJECTIVE

Overall objective of this work is to develop a Web Application which can be deployed on to Hadoop Cluster to perform certain map-reduce transaction on the provisioned cluster.

PURPOSE:

This application can be used to retrieve the **TITLE** of the web page of specific **URLs** .The Application make use of the Hadoop Cluster which distribute the desired Task among the slaves nodes of Cluster and their by Increasing the speed of Computation.

2. REQUIREMENTS

- Hadoop Cluster
- Web Server(ex- vFabric TC server)

3. USER GUIDE

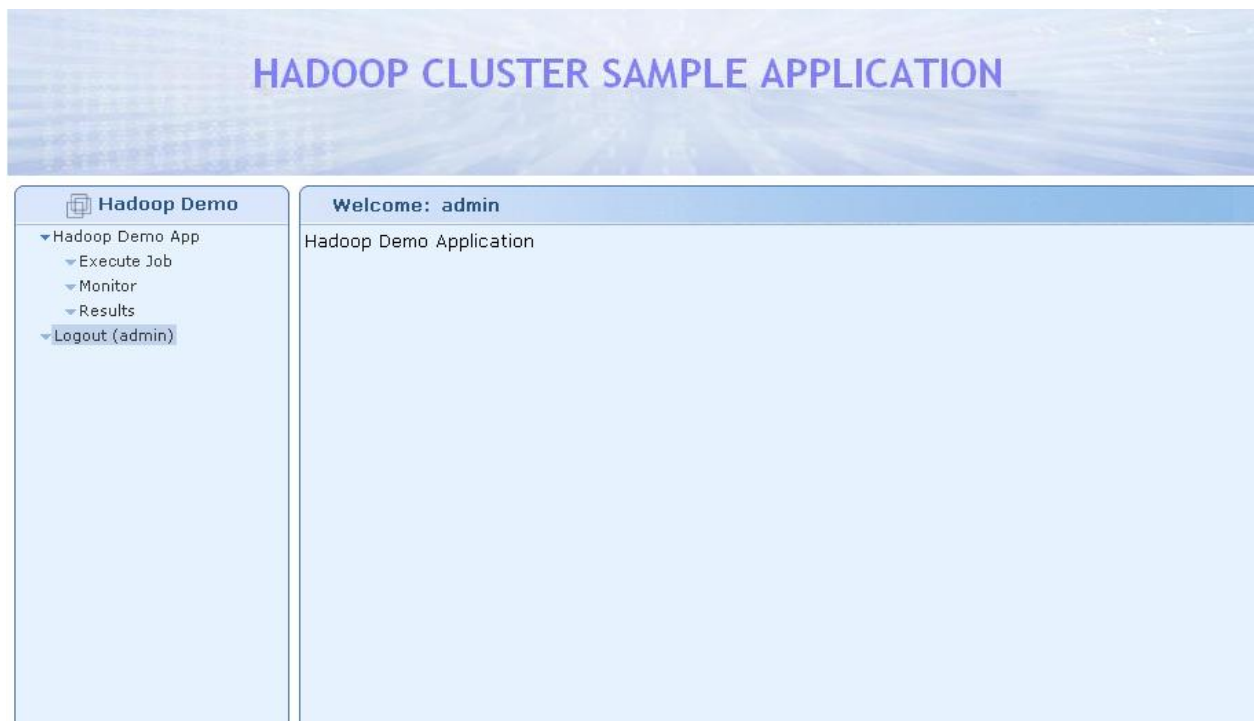
The sample application can be accessed from the url
<http://<JobTrackerIP>:8080:/HadoopDemoApp/Login.jsp>

3.1 Login Page

A screenshot of a web browser window showing a login form. The form has a title bar that says "Enter the Login details". It contains two input fields: "Username:" with the text "admin" and "Password:" with masked characters "*****". Below the fields is a "Login" button. The window has standard OS controls (minimize, maximize, close) in the bottom right corner.

- **Login** is used for the authentication of the user for the Hadoop Sample application.
Username/Password: admin/admin

3.2 MAIN PAGE



- **Hadoop Demo App** for Navigation Purpose to Execute, Monitor and Result of the Job.
- **Logout** is to Log Out from the Application.

3.3 Execute Job



The screenshot shows a web application titled "HADOOP CLUSTER SAMPLE APPLICATION". It features a sidebar menu on the left with options: "Hadoop Demo App", "Execute Job", "Monitor", "Results", and "Logout (admin)". The main content area is titled "Execute Job" and contains a text input field, an "Add URL" button, and an "Execute Job" button. Below these is a table with the heading "Registered URL" containing a list of URLs.

Registered URL
http://www.cs.brandeis.edu
http://www.nytimes.com
http://www.google.com
http://www.yahoo.com
http://www.exploratorium.edu
http://www.teachers.net
http://www.thirteen.org/
http://school.discovery.com

- **Add URL** can be used to add the URLs of which the Title has to be retrieved.
- **Execute Job** is for executing the Hadoop Job.

3.4 Monitor

HADOOP CLUSTER SAMPLE APPLICATION

Hadoop Demo

- ▼ Hadoop Demo App
 - ▼ Execute Job
 - ▼ **Monitor**
 - ▼ Results
 - ▼ Logout (admin)

JT Hadoop Map/Reduce Administration Quick Links

State: RUNNING
Started: Sun Mar 18 17:24:33 UTC 2012
Version: 0.20.203.0, r1099333
Compiled: Wed May 4 07:57:50 PDT 2011 by oom
Identifier: 201203181724

Cluster Summary (Heap Size is 15.5 MB/966.69 MB)

Running Map Tasks	Running Reduce Tasks	Total Submissions	Nodes	Occupied Map Slots	Occupied Reduce Slots	Reserved Map Slots	Reserved Reduce Slots	Max Task Capacity
0	0	1	2	0	0	0	0	4

Scheduling Information

- **Monitor** is for monitoring the execution of Hadoop Job.

3.5 Result

HADOOP CLUSTER SAMPLE APPLICATION

Hadoop Demo

- ▼ Hadoop Demo App
 - ▼ Execute Job
 - ▼ Monitor
 - ▼ Results
- ▼ Logout (admin)

Job Result

Job Results
http://school.discovery.com Free Teacher Resources Discovery Education
http://www.cs.brandeis.edu Computer Science Department Brandeis University
http://www.exploratorium.edu Exploratorium: the museum of science, art and human perception
http://www.google.com Google
http://www.nytimes.com The New York Times - Breaking News, World News & Multimedia
http://www.teachers.net Teachers.Net - TEACHERS - Lesson Plans, Teaching Jobs, Chat
http://www.thirteen.org/ THIRTEEN - New York Public Media
http://www.yahoo.com Yahoo!

- **Result** is to display the results of the Hadoop Job submitted to the Application.

