

Chapter 10

Hadoop Auth

ABSTRACT

One of the factors for the reliability of the services is authentication, which decides who can access what services. Since big data offers a wide variety of services, authentication becomes one of the main criteria for consideration. This chapter outlines the features of the security services in terms of the requirements and the issues in the business services. This chapter also gives a little background about the services in the cloud and the interaction between clients and services in the cloud, emphasizing the security services. The authentication procedure with the authentication protocol, Kerberos SPNEGO, which is offered as a security service in Hadoop, is introduced. The configuration details in a typical browser (Mozilla Firefox) are detailed. The usage of the Linux command curl is introduced in this chapter. The command to key distribution center “kinit” is outlined. Also, the procedure for accessing the server within the Java code is given. A section on server-side configuration speaks about the Maven repository, which holds all the necessary library Jar files organized as local, central, and remote. The explanation for the configuration is given with a typical XML file. Also, the usage of Simple Logging Facade for Java is introduced. The configuration has many parameters with its values and they are tabulated for better perception. The use of LDAP server, which is one of the lightweight directory access protocols, is introduced. Also, the provision for multi-scheme configuration is outlined with an example configuration file. The facilities available to provide advanced security features using signer secret provide are highlighted with appropriate examples for the parameter name and parameter value.

DOI: 10.4018/978-1-5225-3790-8.ch010

BACKGROUND

The word authentication refers to who can access a set of resources. The practice of accessing the critical resources started from the inception of the multiuser environment. Though it was made available even in the dumb terminals of the Unix operation system based environment, it was seriously viewed during usage with the client server based system. As the access to the various services was accounted for its usage, the authentication scheme took much importance. Recently, the cloud authentication scheme is still viewed seriously, since the exact place of our storage and the computing elements are not known to any of the users. Hence the authentication becomes one of the important issues in the bigger system.

INTRODUCTION TO SECURITY

Security is the only issue the business service providers and the clients hesitate to do during online access. Though the security mechanisms could be easily implemented recently, due to the availability of algorithm, codes, and tools, the security implementation in large sized data and distribution environment is still having challenges.

For any reliable communication between two entities, authentication plays major role as a security mechanism intended to verify identify of the entity on the information exchange. Both private and public key encryption methods could be used to provide the authentication as shown in Figure 1.

Figure 1a shows the passing of messages from sender to receiver with the usage of two different keys (asymmetric) where for every message communication, a pair of key has to be generated (public key, private key). If the sender encrypts the messages with his private key at the sender side, it should be decrypted by receiver using the public key. It could be done in other way also. The confidentiality of the messages is achieved if message is encrypted using public key and decrypted using private key leading to the confidentiality of the message i.e. the message can be decrypted only by the person having the private key of this message. If the message is encrypted by the private key and is to be decrypted by the receiver using the public key it is authentication of the person.

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/hadoop-auth/216605

Related Content

Multi-Week Travel Surveys Using GPS Devices: Experiences in The Netherlands

Tao Feng and Harry Timmermans (2014). *Mobile Technologies for Activity-Travel Data Collection and Analysis* (pp. 104-118).

www.irma-international.org/chapter/multi-week-travel-surveys-using-gps-devices/113206

An Experimental Data of Lithium-Ion Battery Time Series Analysis: ARIMA and SECTRAL Analysis

Liming Xie (2021). *International Journal of Data Analytics* (pp. 1-26).

www.irma-international.org/article/an-experimental-data-of-lithium-ion-battery-time-series-analysis/285465

Loan Fraud Detection Using Machine Learning as a Data Mining Approach

Nabila Hamdoun (2022). *International Journal of Data Analytics* (pp. 1-10).

www.irma-international.org/article/loan-fraud-detection-using-machine-learning-as-a-data-mining-approach/309096

A Systematic Study of Central Bank Digital Currency: Challenges, Limitations, and Its Effects on the Indian Economy

M. S. Nikhil, Nischal V. Atreya, H. Mehul Kumar, B. Priyadharshini and S. Ramya (2024). *Data-Driven Decision Making for Long-Term Business Success* (pp. 331-342).

www.irma-international.org/chapter/a-systematic-study-of-central-bank-digital-currency/335581

Predictive Analytics in Operations Management

Harsh Jain, Amrit Pal and Manish Kumar (2017). *Applied Big Data Analytics in Operations Management* (pp. 68-92).

www.irma-international.org/chapter/predictive-analytics-in-operations-management/167564