


Tweet Wordcount MapReduce

- Log into this [Amazon WebServices](#) console with these credentials
 - Username: 'globant-hadoop'
 - Password 'Gl0b4nt2011'

Amazon.com Sign In

https://www.amazon.com/ap/signin?openid.assoc_handle=aws&openid.return_to=https%3A%2F%2Fsignin.aws.amazon.com%2Foauth%3Fresponse_type%3Dcode%26...



Amazon Web Services Sign In

Please enter the AWS Identity & Access Management (IAM) User name and password assigned by your system administrator to sign in.

AWS Account: 896585812645

User Name:

Password:

[Sign in using our secure server](#)

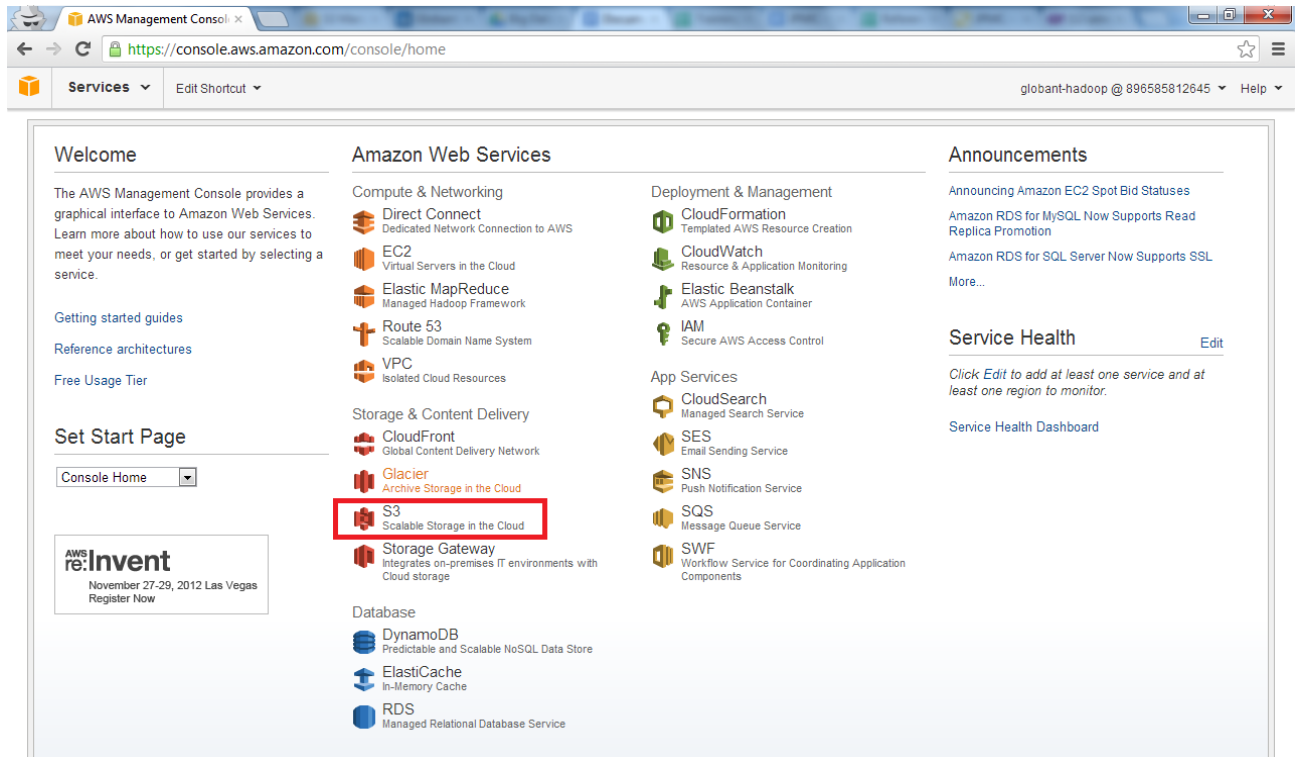
Please contact your system administrator if you have forgotten your user credentials.

[Sign in using AWS Account credentials](#)

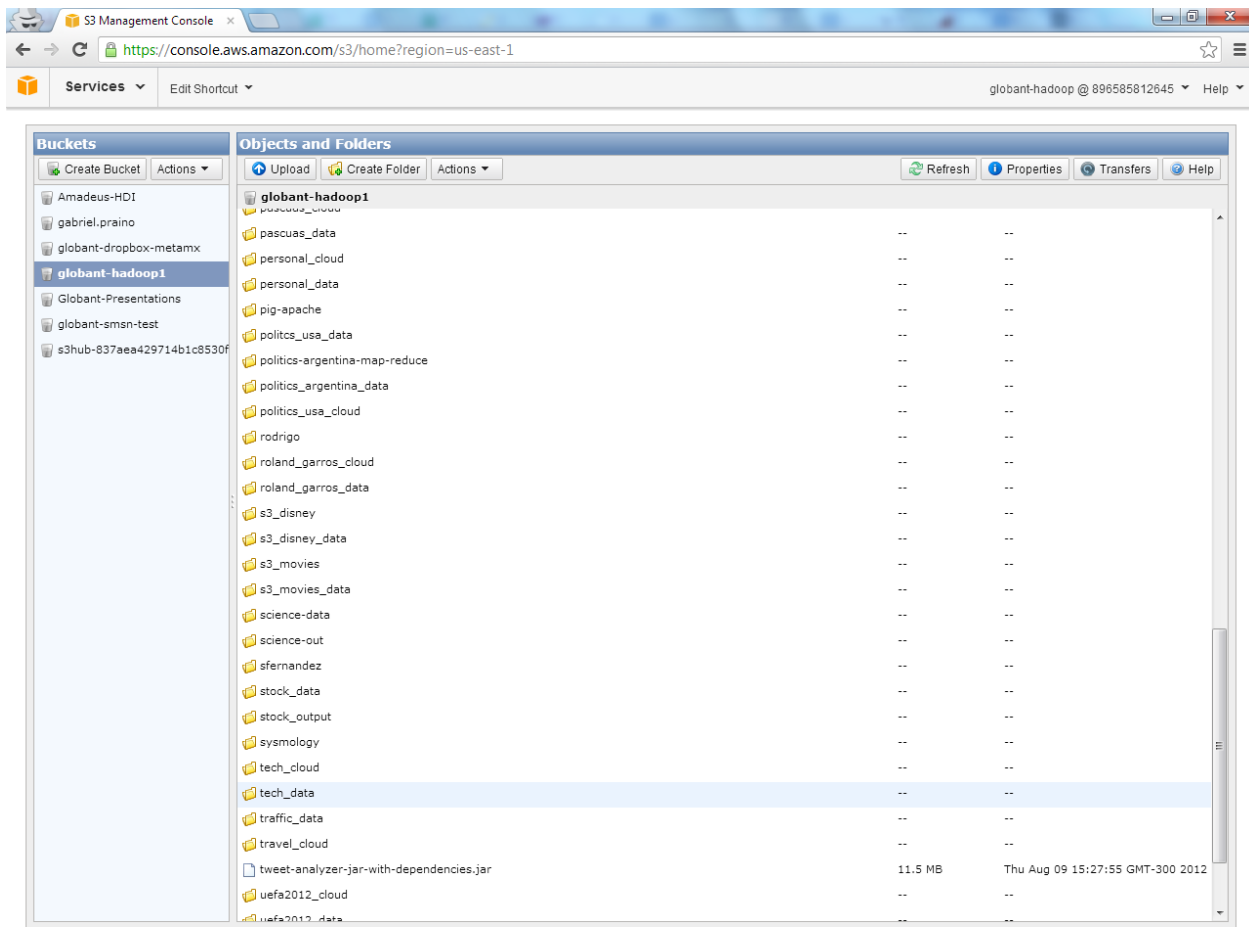
[Terms of Use](#) [Privacy Policy](#) © 1996-2010, Amazon.com, Inc. or its affiliates.

An [amazon.com](#) company

- Navigate to the S3 section



- Open the 'globant-hadoop1' bucket
- Scroll down to the 'tech-data' folder



© 2008 - 2012, Amazon Web Services LLC or its affiliates. All rights reserved. | [Feedback](#) | [Support](#) | [Privacy Policy](#) | [Terms of Use](#) | An [amazon.com](#) company

- Download any (or all) of the .json files.
- Load it into the HDFS at a destination folder of your choosing
- Create a User Java class which implements Writable and contains a 'lang' field and a 'screen_name' field.
- Create a Tweet Java class which implements Writable and contains a 'text' field, and a 'user' field (of type User).
- Write an InputFormat capable of parsing the .json files and for each tweet generating a Tweet object.
- Create a MapReduce job which processes Tweets from the .json files and outputs the english speaking users who mentioned one of the following terms:
 - apple
 - ios
 - itunes

- google
- android
- gmail
- microsoft
- windows
- hotmail