

---

# Exercise 9: Geo-Distributed Hadoop with Amazon Web Services



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

---

Concepts and Technologies for Distributed Systems and Big Data Processing – SS 2016

---

---

## Task 1 Sign Up for AWS Educate

---

Apply for AWS Educate for Students <https://www.awseducate.com/Application?apptype=student>. This is a free<sup>1</sup> membership service by Amazon. The registration process may take some days.

**Use your AWS credits with care.** Exercise 10 will be based on your AWS setup. Therefore, additional AWS credits will be required.

---

## Task 2 Getting Started with AWS

---

Get familiar with AWS and EC2 by reading some of the documentation like <http://docs.aws.amazon.com/gettingstarted/latest/awsgsg-intro/gsg-aws-intro.html> and <http://aws.amazon.com/ec2/getting-started/>.

---

## Task 3 Setup a Small Hadoop Cluster

---

Create your own small Hadoop cluster with 1 NameNode and 3 DataNodes like described in this tutorial <http://insightdataengineering.com/blog/hadoopdevops/>. Do not forget to **terminate your EC2 instances** when you are done with your work.

---

## Task 4 Run a Simple Hadoop Experiment

---

Try out your new Hadoop cluster by running one of the small Hadoop examples or your code from the previous Hadoop exercises.

---

<sup>1</sup> You will receive some AWS credits for free