Dealing with Massive Data

March 2, 2011

Homework 1.5: (No Due Date)

This homework is designed to get you familiar with the AWS environment.

Signing up for AWS

- To begin, e-mail the TAs: Kripa and Mohan, {kk2735, mk3245} to obtain your Amazon Web Services code.
- With the code in mind, register for AWS (at aws.amazon.com). At the minimum you should register for S3, EC2 and Elastic Mapreduce.
- Go to http://aws.amazon.com/awscredits to redeem your credit.

Using EC2

- Launch an amazon EC2 instance.
- Log into the instance, and check that you can do basic commands

Using Elastic MapReduce

[After Class on March 7]

- Write the 'Hello World' of Hadoop programs, WordCount. For debugging purposes it is helpful to have a local hadoop version operating with a single compute node.
- Compile and execute WordCount using Elastic MapReduce