
Instituto de Computação - UNICAMP

MC855 - Projeto em Sistemas de Computação

Plano de Desenvolvimento da Disciplina

[Profa. Islene Calciolari Garcia](#)

Ementa

Implementação de um projeto prático na área de Sistemas de Computação.

Tema

Serão feitos experimentos práticos em grupo utilizando o projeto Apache Hadoop, que é um sistema distribuído real e que tem seu código fonte disponível sob licença livre. Será feito um projeto prático com aplicações MapReduce e serão analisados bugs e melhorias propostas pela comunidade do Hadoop.

Critério de avaliação

A nota final será atribuída conforme a divisão abaixo:

- (20%) Experimento com MapReduce.
- (20%) Análise de issue tipo bug em aberto no HDFS ou MapReduce
- (20%) Análise issue umbrella em aberto no HDFS ou MapReduce
- (30%) Análise mais detalhada de issue livre escolha no projeto Hadoop
- (10%) Participação nos seminários e apresentações do JIRA

Mais detalhes no site da disciplina: <http://www.ic.unicamp.br/~islene/2s2015-mc855/index.html>.

Referências

- [Projeto Apache Hadoop](#)
 - [JIRA do Hadoop](#)
 - [Distributed Systems: Principles and Paradigms](#), Andrew S. Tanenbaum and Maarten Van Steen, Second Edition, Pearson, 2007.
 - [Distributed Systems: Concepts and Design](#), George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair, Fifth Edition, Addison Wesley, 2011.
 - [Hadoop: The Definitive Guide](#), Tom White, Fourth Edition, O'Reilly, 2015.
-

Instituto de Computação - UNICAMP

MC855 - Projeto em Sistemas de Computação

Segundo Semestre de 2015

Islene Calciolari Garcia

Objetivos da disciplina

Veja os [slides](#) da aula introdutória e o [plano de desenvolvimento da disciplina](#).

Projeto 2

Para começar, pesquise em [Issue Tracking](#) uma issue tipo bug para trabalhar. Procure issues em HDFS e MAPREDUCE.

- Qual tipo do bug?
- Quem reportou?
- Há quanto tempo está em aberto?
- Qual seria a sua estratégia para reproduzir o bug? E para corrigi-lo?

Veja também as instruções em [How to contribute to Hadoop](#).

Como executar o Hadoop a partir do código fonte.

A partir de uma versão específica

```
$ wget http://ftp.unicamp.br/pub/apache/hadoop/common/hadoop-2.7.1/hadoop-2.7.1-src.tar.gz
$ tar xzvf hadoop-2.7.1-src.tar.gz
$ cd hadoop-2-7.1-src
$ mvn clean package -Pdist -Dtar -DskipTests -Pnative

$ cd hadoop-dist/target/hadoop-2.7.1/
```

Fazendo um clone do repositório

```
$ git clone git://git.apache.org/hadoop-common.git

$ cd hadoop
$ mvn package -Pdist -Pdoc -Psrc -Dtar -DskipTests
$ cd hadoop-dist/target/hadoop-3.0.0-SNAPSHOT
```

Projeto 1

- Forme seu grupo e registre no [Moodle](#)

Instruções

- [Explore o MapReduce](#)
- [Veja outros exemplos](#)

Entrega

Apresentação do código no lab CC04

- 25/08 (nota * 1.1)
- 27/08 (nota * 1.05)
- 01/09 (nota * 1.0)

Cada grupo deverá mostrar uma aplicação rodando, explicar seu funcionamento e a alteração.

Apresentação do tema no CB15

- 03/09 (nota * 1.1)
- 08/09 (nota * 1.0)

Cada grupo deverá explicar para a turma o funcionamento da aplicação e a alteração feita. Não é necessário mostrar o código rodando na hora. **O grupo deve ter feito a apresentação no laboratório antes.**

Moodle

Para o acompanhamento desta disciplina será aberta uma área no Moodle fornecido pela Unicamp. Faça acesso [MC855A+B](#) com seu login e senha da DAC.

Calendário

Fique atento à lista abaixo para saber quais aulas serão nos laboratórios e quando serão na sala CB15.

- 11/ago (sala CB15): Introdução
- 13/ago (sala CB15): Aula sobre MapReduce
- 18/ago (labs CC02 e CC03): MapReduce no laboratório
- 20/ago (lab CC04): Desenvolvimento do projeto MapReduce
- 25/ago (lab CC04): Desenvolvimento do projeto MapReduce e primeira data de apresentação
- 27/ago (lab CC04): Desenvolvimento do projeto MapReduce e segunda data de apresentação
- 01/set (lab CC04): Desenvolvimento do projeto MapReduce e última data de apresentação
- 03/set (sala CB15): Primeira data de apresentação do tema MapReduce
- 08/set (sala CB15): Última data de apresentação do tema MapReduce
- 10/set (sala CC04): Início da escolha da issue tipo bug
- 15/set (sala CC04): Análise de issues tipo bug
- 17/set (sala CC04): Análise de issues tipo bug

- 22/set (sala CC04): Definição da issue no lab
- 24/set (sala CC04): Definição da issue no lab

- 29/set (sala CC04): Definição da issue no lab
- 01/out (sala CB15): Primeira data de apresentação para a turma
- 03/out (sala CB15): Última data de apresentação para a turma

Material

- [Aula introdutória](#)
- MapReduce:
 - Hadoop: The Definitive Guide, Tom White, Capítulo 2
 - [Slides da aula](#) baseados no material da [Aula 16 - Map Reduce](#), CS417 - Distributed Systems, Paul Krzyzanowski, Rutgers University.
 - [Explorando o MapReduce](#)

