EduCloud - Implantação #569

Implantação # 524 (Resolvida): US: Elaborar as Etapas Necessárias para o Deployment do Hadoop

Configuração de mapreduce e localizing

12 Apr 2012 12:25 - Marco Gomes

Status:	Resolvida	**Start date:**	13 Apr 2012
Priority:	Normal	**Due date:**	
Assignee:	Marco Gomes	**% Done:**	100%
Category:		**Estimated time:**	8.00 hours
Target version:	Ciclo 2	**Spent time:**	4.50 hours

Description

History

#1 - 12 Apr 2012 12:26 - Marco Gomes

- Assignee set to Marco Gomes

- Target version set to Ciclo 2

#2 - 24 Apr 2012 11:04 - Marco Gomes

- Status changed from Nova to Resolvida

- % Done changed from 0 to 100

O map reduce do hadoop é na realidade um framework para programação de algoritmos de análise de grandes quantidades de dados. Uma analogia do map reduce para o mundo dos bancos de dados é que o map reduce se refere à query de busca do banco de dados.
A estrutura do map reduce funciona da seguinte forma:

- no primeiro estágio é feito o mapeamento dos dados através de uma estrutura de par, chave e valor;
- no segundo estágio é feito a redução dessa estrutura, agregando pares que tenham a mesma chave.

Logo, o map reduce atua sobre a análise dos dados, não sobre o dado em si; de forma que o resultado da análise é mais compacto.

Quanto ao localizing, é inerente à arquitetura do hadoop em distribuir os dados de acordo com o fator de replicação estipulado para os dados, e para aproximar o dado dos nós que irão processá-lo.