

Estudante: Clístenes Grizafis Bento

Professor: João Eugênio Marynowski

RESENHA AULA 2 (TARDE)

No começo da aula foi apresentado a ferramenta catalog, que havia sido iniciada no final da aula de manhã, que serve para criar links de informações para poderem ser acessadas por outros sistemas, como por exemplo acessar uma tabela criada no hive por meio do pig. Na sequência foi explanado sobre o job tracker, acessado por meio do job browser, que é responsável em gerenciar os jobs do MapReduce e armazenar informações sobre esses processos. O próximo assunto abordado foi o job design que é uma ferramenta que permite criar um conjunto de tarefas para realizar algum processamento, assim como o Oozie designer. Foi apresentado o Knox que é uma ferramenta de gestão de usuário e limites de acesso.

Depois foi apresentado o hbase, ideal para com múltiplos dados pequenos. Foi apresentado alguns modelos de big data com destaque em noSql e new-sql, onde no noSql são banco de dados não relacionais e não apenas sql. Foi apresentado também as modelagens chave-valor, orientada a documento, grafos (que ainda está em desenvolvimento), modelo colunar (que é um modelo bastante comum, porém pouco explorado). Foi apresentado o new-sql, que é uma melhoria dos gerenciadores de banco de dados já existentes onde foi implementado modelos de noSql com propriedade ACID. Depois foi explanado sobre sistemas usados e quais são os mais populares. Foi frisado que o banco de dados relacional não deixa de existir mas sim começa a ser usado junto com outros modelo.

Na parte final foi falado sobre eventos mais complexos e fluxo de dados. Apresentou-se o apache storm (história, aplicação, terminologia) que embora seja semelhante ao que estava sendo usado até o momento com hadoop, trabalha com dados em tempo real, cujo o objetivo é o processamento de eventos complexos, usa-se o Kafka para múltiplos dispositivos. Mostrou-se exemplo de aplicação gerando sentenças randômicas e contando a quantidade de palavras, passo a passo através do uso do ambari e a interface storm UI. Ao executar o comando, o meu WordCount ficou em 0, só atualizou o tempo.