



Tema Principal: **Um Estudo sobre Processamento de Larga Escala de Grafos Usando uma Abordagem Baseada em Componentes de MapReduce**

Data: **11/12/2014** Horário: **14h** Local: **Sala de Seminários Bloco 952**

Resumo:

Diante de um crescente conjunto de dados, com origem nas novas tecnologias e problemas computacionais das ciências modernas, empresas e pesquisadores têm buscado soluções para alavancar a capacidade de processamento de dados em larga escala. Além do volume, a complexidade que o software enfrenta ao analisar essas informações tem que ser levada em consideração, pois existem potenciais problemas que necessitam de leitura recorrente, indo além do limite da capacidade da memória principal das unidades de processamento. Essa é uma característica de problemas contidos em grafos grandes, onde o tempo de processamento pode ser exponencial em relação ao tamanho da entrada. O modelo de programação MapReduce tem introduzido e especificado uma proposta para alguns dos desafios nesse contexto, permitindo a construção de várias plataformas de computação paralelas baseadas em seu padrão de processamento. Embora possua limitações em grafos

iterativos, é um modelo que oferece facilidade ao desenvolvedor quando permite a programação serial de funções de mapeamento e redução, encapsulando o inerente paralelismo necessário para melhorar desempenho. Neste trabalho, apresenta-se um estudo sobre o MapReduce e como alguns algoritmos em grafos tem sido implementados nesse modelo. São destacadas as plataformas MapReduce contemporâneas, mostrando toda a infraestrutura e organização dessas implementações, como sistema de arquivos distribuídos e escalonamento de tarefas mapeadoras e redutoras. Todavia, busca-se desenvolver estudo de caso com plataforma própria, no contexto CBHPC (Component-Based High Performance Computing), resultando em um arcabouço que suporta componentes que implementam algoritmos em grafos e outras estruturas de dados. Em grafos, destacam-se os componentes PageRank, SSSP e Clique. O PageRank mede a importância de cada página web em um conjunto finito de páginas. O SSSP encontra o menor caminho de um vértice fonte para os demais vértices. O Clique resolve o problema da listagem de cliques maximais através de algoritmo recursivo.

Banca:

- FRANCISCO HERON DE CARVALHO JUNIOR (UFC - Orientador)
- RICARDO CORDEIRO CORRÊA (UFC)
- JOSÉ ANTONIO FERNANDES DE MACEDO (UFC)
- ANDRÉ RAUBER DU BOIS (UFPEL)