



Título: Uma Estratégia Paralela e Distribuída para Assegurar a Confidencialidade de Dados Armazenados em Nuvem

Data: **30/11/2017** Horário: **14:00h** Local: **Sala de Seminários do Bloco 952**

Resumo:

O armazenamento de grandes quantidades de dados confidenciais em servidores na nuvem é uma tendência para as empresas que buscam oportunidades de reduzir custos e aumentar a disponibilidade de seus serviços digitais. Contudo, nos ambientes de computação em nuvem, o controle do dado deixa de ser do seu proprietário e passa a ser do provedor do serviço, o que proporciona novos desafios relacionados privacidade, segurança e confidencialidade. Neste contexto, diferentes soluções para assegurar a confidencialidade dos dados armazenados na nuvem foram propostas. Em geral, tais soluções utilizam criptografia, fragmentação de dados ou uma combinação dessas duas abordagens. Apesar disto, problemas relacionados à eficácia destas técnicas em relação a ataques, perda ou roubo de dados têm ocorrido nos últimos anos, causando prejuízos de milhões de dólares para empresas e clientes. Recentemente, foi proposta uma nova estratégia, denominada QSM-EXTRACTION, para assegurar a confidencialidade de dados em serviços de armazenamento em nuvem. A ciência por trás dessa abordagem utiliza conceitos da Doutrina do Ser de Hegel. A estratégia QSM-EXTRACTION baseia-se na fragmentação de um arquivo digital em fragmentos denominados objetos de informação, na decomposição desses objetos por meio da extração de suas características (Qualidade, Quantidade e Medida) e na dispersão dessas características em diferentes serviços de armazenamento em nuvem, permitindo a posterior recuperação desses dados sem perda de informação. A finalidade da estratégia QSM-EXTRACTION é inviabilizar a reconstrução do arquivo original por parte de um provedor de nuvem que possui apenas parte das características dos objetos de informações que

compõem este arquivo. Desta forma, assegura-se a confidencialidade dos dados armazenados em nuvem e, por conseguinte, a privacidade dos proprietários desses dados. Contudo, apesar de ter sido concebida para o ambiente de computação em nuvem, a estratégia QSM-EXTRACTION adota uma abordagem de execução centralizada, o que pode comprometer o desempenho do processamento de grandes volumes de dados. Para que a computação de um grande volume de dados seja realizado em tempo viável, faz-se necessária a exploração de paradigmas de programação paralela e processamento distribuído. Porém, desenvolver software para ambientes distribuídos é uma tarefa complexa, pois envolve uma série de problemas que devem ser considerados pelos programadores, tais como: concorrência, tolerância a falhas, distribuição de dados e balanceamento de carga. Assim, dividir uma tarefa em subtarefas e então executá-las paralelamente em diversas unidades de processamento não é algo trivial. A fim de facilitar este processo, surge então o MapReduce, um modelo de programação paralela para processamento largamente distribuído de grandes volumes de dados. Neste trabalho, propomos uma versão paralela e distribuída da estratégia QSM-EXTRACTION, denominada dpQSM, que explora o paradigma MapReduce com a finalidade de possibilitar o processamento eficiente de grandes volumes de dados. A abordagem proposta foi implementada utilizando o framework Apache Spark. Com a finalidade de demonstrar a eficiência das ideias que norteiam a solução proposta nesta dissertação, diversos experimentos foram realizados. Para executar estes experimentos, utilizamos uma infraestrutura de nuvem pública gerida pela Amazon. Os algoritmos que compõem a estratégia dpQSM foram implementados em linguagem Java. Para realizar a avaliação da eficiência da estratégia dpQSM, empregamos uma coleção de arquivos com diferentes tamanhos e formatos, representando múltiplos cenários. Os resultados dos experimentos comprovam a viabilidade da utilização da abordagem proposta em aplicações típicas de Big Data.

Banca:

- Prof. Dr. José Maria da Silva Monteiro Filho (MDCC/UFC - Orientador)
- Prof. Dr. João Paulo do Vale Madeiro (UNILAB - Coorientador)
- Prof. Dr. Javam de Castro Machado(MDCC/UFC)
- Prof. Dr. José de Aguiar Moraes Filho (UNIFOR)
- Prof. Dr. Angelo Roncalli de Alencar Brayner (MDCC/UFC)
- Prof. Dr. Flávio Rubens de Carvalho Sousa (UFC)