Abstract

Existing RDF engines are designed for specific hardware architectures; porting to a different architecture (e.g., GPUs) entails enormous implementation effort. We explore sparse matrix algebra as an alternative for designing a portable, scalable and efficient RDF engine. We propose MAGiQ; a matrix algebra approach for evaluating complex SPARQL queries over large RDF datasets. MAGiQ represents RDF graphs as a sparse matrices, and translates SPARQL queries to matrix algebra programs. It takes advantage of the existing rich software infrastructure for processing sparse matrices, optimized for many architectures (e.g., CPUs, GPUs, parallel and distributed). We demonstrate how MAGiQ can utilize MATLAB and various matrix algebra libraries to execute SPARQL queries on CPUs and GPUs. We also show that MAGiQ on a GPU is orders of magnitude faster than state-of-the-art RDF systems on a billion-edge graph.

Panos Kalnis
Professor, KAUST