Object : Job offer :

Analysis of multi-criteria path computation on very-large graphs

This position could be fulfilled by either:

- a person owning either an engineer or a master degree, looking for a mixed development/research contract, that could eventually lead to a PhD position.
- a person owning a PhD and looking for a Postdoc position.

Description of the work subject

When facing very large graphs (billions of vertices and edges) as for instance the friendship-graph of Facebook or the European public transportation network, standard graph algorithms cannot be used anymore. Firstly, the size of the considered graphs is so huge that they cannot be loaded into the memory of a single computer. Consequently, distributed (parallel) storage and computations are necessary. Secondly, the computation time for basic operations (shortest path computation, partition into connected components, etc.) are becoming prohibitive for real applications. Therefore, new graph algorithms designed for specific distributed computation models are desired.

Localized graph algorithms exploiting only data exchanged between neighboring nodes (also called vertex-centric algorithms) are good candidates for computations over very large graphs. These algorithms iteratively perform some local operations on each vertex and then propagate the computed values to their neighbors. When a condition is reached (number of iterations, threshold, etc.), the computation stops and the result is extracted. A natural programming model for such algorithms is the Bulk Synchronous Parallel (BSP) model. The input data are divided between concurrent and synchronous tasks which perform a sequence of super-steps: i) local computation using a user-provided function at each node, ii) communications with other tasks, and iii) barrier. Several such graph analysis framework based on BSP have recently been proposed (see for instance Giraph [1] and GraphLab [2]).

Objectives of the contract

The first objective of the contract is to design efficient localized algorithms for computing paths on large graphs in different programming models (BSP, Active objects). The algorithm will allow multi-criteria searches to handle constraints such as: various cost functions across an edge of the graph, waiting time in nodes, time dependent cost, etc.
The second objective of this work is to benchmark these algorithms against existing software solutions for solving problems on very large graphs. More precisely, the candidate is expected to:

- Survey existing libraries for large graph processing, focusing on scalability and reliability of the engine,
- Design a set of prototypes to evaluate the benefits of each solution.

Prerequisites for candidates

RESEARCH CENTRE
SOPHIA ANTIPOLIS - MÉDITERRANÉE
2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex France
Phone: +33 (0)4 92 38 77 77
Fax: +33 (0)4 92 38 77 65

Algorithms, programming languages (C, Java).
Previous experience in graph theory and/or in distributed computing would be a plus.

Date of start

The candidate must be available for starting the contract between the March 1st and June 1st 2014.

Duration
12 months.

Place of work
INRIA Sophia Antipolis (http://www.inria.fr/centre/sophia/)

Contacts
David Coudert (david.coudert@inria.fr) and Eric Madelaine (eric.madelaine@inria.fr).

References
[1] http://incubator.apache.org/giraph/
[2] http://graphlab.org/