

CMPSCI 691FF: Algorithmics for Internet-Based Computing

Instructor: Arnold L. Rosenberg **Office:** CMPS 308
Email: rsnbrg@cs.umass.edu
Office Hours: By appointment
Class Mtgs: TuTh, 11:15-12:30 CMPS 140

Prerequisite: CMPSCI 611 plus at least one systems core course with a grade of B or better.

Overview. Advances in technology have rendered the Internet a viable medium for a new genre of “collaborative” computing (that we call *Internet-based computing* [IC, for short]), wherein multiple computing agents, possibly widely dispersed geographically, cooperate in the solution of a single computational problem. In a series of papers [3, 4, 10, 11, 12], we are currently developing an algorithmic framework for scheduling computations having intertask dependencies, for the several modalities of IC—including Grid computing (cf. [1, 6, 5]), global computing (cf. [2]), and Web computing (cf. [8]). This new *IC-scheduling theory* aims to craft schedules that maximize the rate at which tasks are rendered eligible for allocation to remote clients (hence for execution), with the dual goal of: (a) enhancing the effective utilization of remote workers, by always having a task to allocate to an available worker; (b) lessening the likelihood of the “gridlock” that can arise when a computation stalls pending computation of already-allocated tasks. Preliminary simulation experiments reported in [9, 7] bolster our hope that IC-scheduling theory will be the underpinnings of a theory of scheduling complex computations for IC.

The course. This course will be devoted to studying and advancing IC-scheduling theory. The “studying” will consist of reading (portions of) papers. *Each student will be responsible for presenting at least one, and at most two, papers during the semester.* The “advancing” will consist in a piece of research that advances the theory either via improved/extended algorithms or via an improved/extended simulation study. *Each student will participate in such a study, chosen in consultation with the instructor. The studies can be either solo or in teams.* Part of the excitement will be deciding what kinds of studies are needed/appropriate.

Grades. The grading will be based on the quality of a student’s presentation(s) and the perceived effort put into the research topic.

References

- [1] R. Buyya, D. Abramson, J. Giddy (2001): A case for economy Grid architecture for service oriented Grid computing. *10th Heterogeneous Computing Wkshp.*
- [2] W. Cirne and K. Marzullo (1999): The Computational Co-Op: gathering clusters into a metacomputer. *13th Intl. Parallel Processing Symp.*, 160–166.
- [3] G. Cordasco, G. Malewicz, A.L. Rosenberg (2006): Advances in a dag-scheduling theory for Internet-based computing. Submitted for publication. See also, On scheduling expansive and reductive dags for Internet-based computing. *26th Intl. Conf. on Distributed Computing Systems*, 2006.

- [4] G. Cordasco, G. Malewicz, A.L. Rosenberg (2006): An enhanced dag-scheduling theory for Internet-based computing. Typescript, Univ. Massachusetts.
- [5] I. Foster and C. Kesselman [eds.] (2004): *The Grid: Blueprint for a New Computing Infrastructure (2nd Edition)*. Morgan-Kaufmann, San Francisco.
- [6] I. Foster, C. Kesselman, S. Tuecke (2001): The anatomy of the Grid: enabling scalable virtual organizations. *Intl. J. Supercomputer Applications*.
- [7] R. Hall, A.L. Rosenberg, A. Venkataramani (2006): A comparison of dag-scheduling strategies for Internet-based computing. Typescript, Univ. Massachusetts.
- [8] E. Korpela, D. Werthimer, D. Anderson, J. Cobb, M. Lebofsky (2000): SETI@home: massively distributed computing for SETI. In *Computing in Science and Engineering* (P.F. Dubois, Ed.) IEEE Computer Soc. Press, Los Alamitos, CA.
- [9] G. Malewicz, I. Foster, A.L. Rosenberg, M. Wilde (2006): A tool for prioritizing DAGMan jobs and its evaluation." *15th IEEE Intl. Symp. on High-Performance Distributed Computing*, 156–167.
- [10] G. Malewicz, A.L. Rosenberg, M. Yurkewych (2006): Toward a theory for scheduling dags in Internet-based computing. *IEEE Trans. Comput.* 55, 757–768.
- [11] A.L. Rosenberg (2004): On scheduling mesh-structured computations for Internet-based computing. *IEEE Trans. Comput.* 53, 1176–1186.
- [12] A.L. Rosenberg and M. Yurkewych (2005): Guidelines for scheduling some common computation-dags for Internet-based computing. *IEEE Trans. Comput.* 54, 428–438.