Seeking Fast Operations in MWMR Atomic Register Implementations *

Nicolas Nicolaou†

In this work we explore the communication and computation costs of multiple-writer multiple-reader (MWMR) atomic read/write register implementations in asynchronous message-passing systems with crash-prone processors. We consider algorithms that use *quorum systems*, collections of subsets of replica hosts with pairwise intersection, called *quorums*. A quorum system has intersection degree n (also called n-wise quorum system), if every n quorum members of this system have a non-empty intersection. Given this definition we show that a MWMR atomic register implementation deploying an n-wise quorum system, allows up to $n-1$ consecutive fast write operations.

In order to enable fast write operations we introduce a new technique we call server side ordering (SSO), that transfers partial responsibility of the ordering of write operations to the replica hosts. Using this idea we design algorithm SFW that uses n-wise quorum systems and relies on predicates to allow fast read and write operations. The algorithm uses tag-value pairs to order the write operations and combines a global ordering imposed by the servers with a local ordering established by each writer participant. A predicate is used during each read and write operation to establish if sufficient number of servers assign the same order to a written value. If the predicate holds then the read or write operation completes in a single round. SFW implementation is near optimal in terms of the number of successive fast operations it allows.

We formulate a new combinatorial problem that captures the computational burden of evaluating the predicates in algorithm SFW and we show that it is NP-Complete. To make the evaluation of the predicates feasible, we present a polynomial log-approximation algorithm for this problem and we show how to use it with algorithm SFW. Lastly we identify that SFW allows fast operations under restrictions on the construction of the underlying quorum system. We then design an algorithm that trades the speed of write operations for removing any constraints on the quorum system construction. The new algorithm, called CWFR, incorporates Quorum Views algorithmic techniques presented in the SWMR model to enable fast read operations. CWFR allows single round reads and two round write operations.

[1] B. Englert, C. Georgiou, P. M. Musial, N. Nicolaou, and A. A. Shvartsman. On the efficiency of atomic multi-reader, multi-writer distributed memory, OPODIS 2009.

[2] C. Georgiou, N. Nicolaou, A. Russell, and A. A. Shvartsman. Towards Feasible Implementations of Low-Latency Multi-Writer Atomic Registers, NCA 2011.

[3] N.C. Nicolaou. Trade-Offs in Implementing Consistent Distributed Storage, Ph.D. dissertation, University of Connecticut, USA, May 2011.

*This work is supported in part by the Cyprus Research Promotion Foundation's grant ΠΕΝΕΚ/0609/31 and the European Regional Development Fund.

†University of Connecticut and University of Cyprus, nicolas@engr.uconn.edu