Brief Announcement: Time and Message-Efficient S-Based Consensus*

Fabíola Greve† Michel Hurfin† Raimundo Macêdo‡ Michel Raynal†

† IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France
‡ LaSiD-CPD-UFBA, Campus de Ondina, CEP 40170-110 Bahia, Brazil
{fgreve|hurfin|raynal}@irisa.fr

The class of *strong* failure detectors (denoted $\mathcal{S}$) includes all failure detectors that suspect all crashed processes and that do not suspect some (a priori unknown) process that never crashes. So, a failure detector that belongs to $\mathcal{S}$ is intrinsically unreliable as it can arbitrarily suspect correct processes.
Several $\mathcal{S}$-based consensus protocols have been designed. They proceed in consecutive asynchronous rounds. Some of them systematically require n rounds (n being the number of processes), each round involving n^2 or n messages. Others allow early decision (i.e., when there are no erroneous suspicions, the number of rounds depends only on the maximal number of crashes) but require each round to involve n^2 messages.

This brief annoucement introduces an early deciding $\mathcal{S}$-based consensus protocol each round of which involves $3(n-1)$ messages. So, the proposed protocol is particularly time and message-efficient. It is based on the rotating coordinator paradigm. Due to space limitation the reader is referred to [2] for an in-depth presentation of the protocol (underlying principles, proof, cost analysis and generalization). Here, the limited space allows us to only compare it with the other $\mathcal{S}$-based consensus protocols we are aware of.
This comparison is provided in Table 1. "nb", "msg" and "c_s" are shortcuts for "number", "messages" and "communication steps", respectively; GHMR denotes the proposed protocol; f denotes the maximum number of processes that are allowed to crash; $|v|$ is the maximal size (in bits) of a value a process can initially propose to the consensus.
All the protocols work provided that $f < n$. Hence, they are optimal with respect to the number of process crashes that can be tolerated. The protocols of the first two lines systematically require n rounds ([4] and [5] each proposes a particular generalization of $\mathcal{S}$; we consider here their protocols used with $\mathcal{S}$). The behavior of the protocols of the last two lines depends on failure occurrences and false failure suspicions (that is why they allow early decision in "good" scenarios, i.e., when the failure detector does not make mistake). For those lines, the table considers that (1) the failure detector does not make mistake, (2) there are f failures, and (3) there is a failure per round (the "simultaneity" of failure occurrences can actually reduce the number of rounds). For the generic protocol described in [3], the table considers its instantiation with $\mathcal{S}$.

Protocol	early decision	total nb of rounds	c_s per round	max nb of msg per round	bit size of the biggest msg
[1]	no	n	1	$n(n-1)$	$n\|v\| + log_2 n$
[4], [5]	no	n	1	$(n-1)$	$\|v\|$
[3]	yes	$(f+1)$	≤ 2	$n(n-1)$	$\|v\| + log_2 n$
GHMR	yes	$(f+1)$	≤ 2	$3(n-1)$	$\|v\| + 2\ log_2 n$

Table 1: Comparing $\mathcal{S}$-Based Consensus Protocols

Hence, it appears that the proposed protocol owns several noteworthy features. The most important of them are its ability to allow early decision and its linear cost for the number of messages per round.

*This work has been partially supported by the 98-1B-123 grant from CNET (France Telecom). F. Greve is supported by a CNPq (Brazil) grant.

1. REFERENCES

[1] Chandra T. and Toueg S., Unreliable Failure Detectors for Reliable Distributed Systems. *Journal of the ACM*, 43(2):225-267, March 1996.

[2] Greve F., Hurfin M., Macêdo R. and Raynal M., Consensus Based on Strong Failure Detectors: Time and Message-Efficient Protocols. *Tech Report #1290*, IRISA, Université de Rennes, France, Jan. 2000. http://www.irisa.fr/EXTERNE/bibli/pi/1290/1290.html.

[3] Mostéfaoui A. and Raynal M., Solving Consensus Using Chandra-Toueg's Unreliable Failure Detectors: a General Quorum-Based Approach. *Proc. 13th Int. Symposium on Distributed Computing (DISC'99)*, Springer-Verlag LNCS 1693, pp. 49-63, (P. Jayanti Ed.), 1999.

[4] Mostéfaoui A. and Raynal M., Consensus Based on Failure Detectors with a Perpetual Weak Accuracy Property. *Proc. Int. Parallel and Distributed Processing Symposium (IPDPS'2k)*, Cancun (Mexico), May 2000.

[5] Yang J., Neiger G. and Gafni E., Structured Derivations of Consensus Algorithms for Failure Detectors. *Proc. 17th ACM Principles of Distributed Computing*, Puerto Vallarta, pp.297-308, 1998.