Localization of Sender-based Message Logging and Its Recovery for Distributed Broker-based Systems

JinHo Ahn[1]

Dept. of Computer Science, Kyonggi University
San 94-6 Iuidong, Yeongtonggu, Suwon-si Gyeonggi-do 443-760, Republic of Korea
jhahn@kgu.ac.kr

Abstract. Attempting to apply sender-based message logging into large-scale and geographically distributed systems requires reducing the number of messages passing on core networks during its message logging and recovery procedures. This paper presents a scalable sender-based message logging algorithm to solve this problem by enabling the broker elected in a group of nodes as virtual sender to localize both of the logging and recovery procedures to a maximum. We show how this algorithm can guarantee the system consistency in case of sequential failures and perform better than the conventional one in terms of message overhead.

Key words: distributed computing, scalability, reliability, localization, message logging and recovery

1 Introduction

Among rollback-recovery techniques[2], sender-based message logging[1, 3, 5] with checkpointing is one of the most lightweight fault-tolerance techniques to be capable of being applied in those fields. It may considerably lower high failure-free overhead of receiver-based message logging resulting from synchronously logging each message into stable storage, which can be realized by using volatile memory of its sender as storage for logging[1–5]. This beneficial feature can be obtained at the expense of extra communication required for allowing message senders to get receive sequence numbers(RSNs) of the messages from their receivers and confirm them with the receivers and slowness and complexity of recovery of each failed process coming from its obtaining message log from the corresponding senders. As architectural aspects of current and future distributed computing systems are changing to geographically group-based and peer-to-peer based, many of these systems, especially sensor networks, are adopting broker-based architectures to accommodate these topological features well. Thus, this change is making several issues about their fundamental building blocks that should be reconsidered to work well for these newly fashioned systems in highly effective manners. Sender-based message logging abbreviated by SBML should also be examined properly before its application to accommodate this architectural change, which we focus on in this paper. In this point of view, the two

drawbacks mentioned earlier all of the existing sender-based message logging algorithms[1, 3, 5] have may be amplified and highlighted greatly if applied into these new systems, being capable of significantly diminishing the practical value coming from their common advantageous features, even becoming unpractical. In this paper, we present a scalable sender-based message logging algorithm to address the critical problems by employing the current and future distributed systems' architectural features mentioned above. In order to satisfy these requirements, this algorithm enables the broker elected in a cluster or group of nodes to localize both of the logging and recovery procedures to a maximum. This feature may considerably reduce the number of control and data messages passing on core networks incurred during fully message logging and recovery procedures of sender-based message logging.

References

1. A. Bouteiller, F. Cappello, T. Herault, G. Krawezik, P. Lemarinier, F. Magniette, MPICH-V2: a Fault Tolerant MPI for Volatile Nodes based on Pessimistic Sender Based Message Logging, *In Proc. of the Int'l Conf. on High Performance Networking and Computing*, 2003.
2. E. Elnozahy, L. Alvisi, Y. Wang, D. Johnson, A Survey of Rollback-Recovery Protocols in Message-Passing Systems, *ACM Computing Surveys*, Vol. 34, No. 3, pp 375-408, 2002.
3. D. Johnson, W. Zwaenpoel, Sender-Based Message Logging, *Int'l Symp. on Fault-Tolerant Computing*, pp. 14-19, 1987.
4. T. LeBlanc, R. Anand, E. Gabriel and J. Subhlok, VolpexMPI: An MPI Library for Execution of Parallel Applications on Volatile Nodes, *Lecture Notes In Computer Science*, Vol. 5759, pp. 124-133, 2009.
5. J. Xu, R.B. Netzer and M. Mackey. Sender-based message logging for reducing rollback propagation, *In Proc. of the 7th International Symposium on Parallel and Distributed Processing*, pp. 602-609, 1995.