[1] E. Anceaume, B. Charron-Bost, P. Minet, and S. Toueg. On the formal specification of group membership services. Technical Report TR95-1534, Cornell University, Computer Science Department, Aug. 25, 1995. [bib]

The problem of group membership has been the focus of much theoretical and experimental work on fault-tolerant distributed systems. This has resulted in a voluminous literature and several formal specifications of this problem have been given. In this paper, we examine the two most referenced formal specifications of group membership and show that they are unsatisfactory: One has flaws in the formalism and allows undesirable executions, and the other can be satisfied by useless protocols.

[2] Ö. Babaoğlu, R. Davoli, and A. Montresor. Group communication in partitionable systems: Specification and algorithms. Technical Report UBLCS-98-1, University of Bologna (Italy). Department of Computer Science., Apr. 1998. [bib | .ps.gz]

We give a formal specification and an implementation for a partitionable group communication service in asynchronous distributed systems. Our specification is motivated by the requirements for building “partition-aware” applications that can continue operating without blocking in multiple concurrent partitions and reconfigure themselves dynamically when partitions merge. The specified service guarantees liveness and excludes trivial solutions; it constitutes a useful basis for building realistic partition-aware applications; and it is implementable in practical asynchronous distributed systems where certain stability conditions hold.

[3] Ö. Babaoğlu, R. Davoli, and A. Montresor. Group communication in partitionable systems: specification and algorithms. *IEEE Transactions on Software Engineering*, 27(4):308-336, 2001. Revision of [2]. [bib | DOI]

We give a formal specification and an implementation for a partionable group communication service in asynchronous distributed systems. Our specification is motivated by the requirements for building “partition-aware” applications that can continue operating without blocking in multiple concurrent partitions and reconfigure themselves dynamically when partitions merge. The specified service is sound in the sense that it guarantees liveness, excludes trivial solutions and is implementable in practical asynchronous distributed systems where certain stability conditions hold.

[4] Ö. Babaoğlu, R. Davoli, A. Montresor, and R. Segala. System support for partition-aware network applications. In *18th International Conference on Distributed Computing Systems*, pages 184-191, May 1998. [bib]

[5] K. P. Birman. The process group approach to reliable distributed computing. *Communications of the ACM*, 36(12):37-53, Dec. 1993. [bib | DOI]

[6] T. D. Chandra, V. Hadzilacos, S. Toueg, and B. Charron-Bost. On the impossibility of group membership. In *15th annual ACM Symposium on Principles of Distributed Computing*, pages 322-330. ACM Press, May 1996. [bib | DOI]

[7] G. V. Chockler, I. Keidar, and R. Vitenberg. Group communication specifications: a comprehensive study. *ACM Computing Surveys (CSUR)*, 33(4):427-469, Dec. 2001. [bib | DOI]

View-oriented group communication is an important and widely used building block for many distributed applications. Much current research has been dedicated to specifying the semantics and services of view-oriented group communication systems (GCSs). However, the guarantees of different GCSs are formulated using varying terminologies and modeling techniques, and the specifications vary in their rigor. This makes it difficult to analyze and compare the different systems. This survey provides a comprehensive set of clear and rigorous specifications, which may be combined to represent the guarantees of most existing GCSs. In the light of these specifications, over 30 published GCS specifications are surveyed. Thus, the specifications serve as a unifying framework for the classification, analysis, and comparison of group communication systems. The survey also discusses over a dozen different applications of group communication systems, shedding light on the usefulness of the presented specifications. This survey is aimed at both system builders and theoretical researchers. The specification framework presented in this article will help builders of group

communication systems understand and specify their service semantics; the extensive survey will allow them to compare their service to others. Application builders will find a guide here to the services provided by a large variety of GCSs, which could help them choose the GCS appropriate for their needs. The formal framework may provide a basis for interesting theoretical work, for example, analyzing relative strengths of different properties and the costs of implementing them.

[8] The Ensemble Distributed Communication System. Web page. Department of Computer Science, Cornell University, http://www.cs.cornell.edu/Info/Projects/Ensemble/. [bib]

[9] P. D. Ezhilchelvan, R. A. Macêdo, and S. K. Shrivastava. Newtop: A fault-tolerant group communication protocol. In *15th International Conference on Distributed Computing Systems (ICDCS'95)*, pages 296-306, May 1995. [bib]

[10] T. Henzinger, Z. Manna, and A. Pnueli. Temporal proof methodologies for real-time systems. In *Proceedings of the eighteenth annual ACM symposium on Principles of programming languages*, pages 353-366. ACM Press, 1991. [bib | DOI]

[11] The Jgroup Project. Web page. Department of Computer Science, University of Bologna, http://www.cs.unibo.it/projects/jgroup/. [bib]

[12] I. Keidar, J. Sussman, K. Marzullo, and D. Dolev. A client-server oriented algorithm for virtually synchronous group membership in WANs. Technical Report CS1999-0623, University of California, San Diego, Computer Science and Engineering, July 1999. [bib]

We describe a novel scalable group membership algorithm designed for wide area networks (WANs). Our membership service does not evolve from existing LAN-oriented membership services; it was designed explicitly for WANs. Our algorithm provides agreement on membership in a single message round in most cases, yielding a low message overhead. It avoids flooding the network and uses a scalable failure detection service designed for WANs. Furthermore, our algorithm avoids notifying the application of obsolete membership views when the network is unstable, and yet it converges when the network stabilizes. In contrast to most group membership services, we separate membership maintenance from reliable communication in multicast groups: membership is not maintained by every process, but only by dedicated servers. The membership servers are not involved in the communication among the members of the groups. This design makes our membership service scalable in the number of groups supported, in the number of members in each group, and in the topology spanned by each group. Our service is complemented by a virtually synchronous communication service which provides clients with full virtual synchrony semantics.

[13] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems*. Springer, 1992. [bib]

[14] A. Montresor. *System Support for Programming Object-Oriented Dependable Applications in Partitionable Systems*. PhD thesis, University of Bologna, Italy, Mar. 2000. Technical Report UBLCS-2000-10. [bib]

[15] The Transis Project Home Page. Web page. Computer Science Department, The Hebrew University of Jerusalem, http://www.cs.huji.ac.il/~transis/. [bib]
