Gossip-Based Multicast

José Pereira

`jop@di.uminho.pt`

Dep. Informática, Universidade do Minho

4710-057 Braga – Portugal

Proposal for a **half-day** tutorial at LADC'2011.

1 Overview

Gossip-based or epidemic techniques have recently evolved from a little known topic in distributed systems to a mainstream approach to resilience and scalability. The reason for this is twofold: First, the emergence of global applications, that coordinate a very large number of geographically distributed computers, has translated to an unprecedented challenge to the scalability of existing distributed systems techniques. Second, recent research on gossip-based protocols has overcome some efficiency issues and demonstrated the advantages of their inherent resilience and scalability.

A key example of such demanding applications is the management of huge and distributed data centers, as used by Amazon.com, Google, and Akamai. Other examples are massive multi-player on-line games and news dissemination, where the centralized server architecture is already a serious limitation.

At its core, gossip-based communication reduces to the following procedure: Each message received is relayed to a random small subset of neighbors a limited number of times. In spite of its simplicity, the decentralized nature of the protocol makes it resilient in face of node and link failures. It also justifies its scalability, as the number of retransmissions required by each participant grows logarithmically with system size. In fact, the probability that all destinations receive a message can be made as close to 1 as desired by adjusting protocol parameters independently of system size.

In practice, realizing the advantages of gossip-based communication requires *(i)* efficiently obtaining random samples of participants for each retransmission round; and *(ii)* limiting the number of redundant retransmissions that consume excessive network bandwidth. Although these are still active topics of research, existing proposals are already sufficient to make it practical to build applications on gossip-based communication in closed applications.

The tutorial aims at introducing gossip-based communication as a solution to scalability, performance, and dependability issues arising in global-scale computing applications.

In detail, the tutorial highlights the challenges to communication protocols that arise in computing systems on a global scale; motivates the usage of probabilistic guarantees in applications; and presents practical solutions to identified

problems, with emphasis on building systems.

2 Structure

The course starts by motivating the use of gossip-based communication with an overview of several challenges to distributed systems in a global computing scenario. This include issues such as delivery reliability, throughput stability, heterogeneity, and load balancing in very large scale systems. An abstract gossip procedure is then introduced and analyzed, with emphasis on relating configuration parameters with reliability and performance metrics. Assumptions are also highlighted.

The core of the proposed course is then dedicated to existing solutions to practical challenges to gossip-based communication. First, it is shown how random sampling of peers is achieved by means of an overlay network, which can itself be built and maintained by a gossip-based protocol. Then, the module addresses the fundamental latency vs. bandwidth tradeoff and shows techniques that can be used to approach the optimum in different scenarios. Finally, some applications of gossip-based protocols are described and open challenges identified.

Schedule

1. Challenges to information dissemination in global computing (1 hour).

2. Protocol parameters and epidemic properties (30 minutes).

3. Building and maintaining overlay networks (30 minutes).

4. Practical gossip-based dissemination (1 hour).

5. Applications of gossip-based dissemination and open challenges (30 minutes).

3 Supporting Materials

The course is supported by the following materials:

- Slides in English (aprox. 100). Will be made available to attendants online in electronic format. A previous version is available at
  `http://gsd.di.uminho.pt/jop/misc/gbm/`.

- Source code used as an example, available at
  `http://neem.sf.net`.

4 Target Audience

The course assumes familiarity with distributed systems and networking topics at undergraduate level. It is targeted mainly at:

- Engineers aiming at applying advanced distributed systems techniques to solve challenging scale and dependability problems.

- Junior researchers seeking an overview of background work and open challenges on an active research topic.

Both English and Portuguese language speakers are welcome.

5   Presenter

José Pereira is a lecturer at the Department of Informatics and a researcher at the Computer Science and Technology Center (CCTC), both in the Universidade do Minho, in Portugal. He teaches Distributed Systems to the Masters in Informatics at the Universidade do Minho and Distributed Computing to the joint Minho/Aveiro/Porto Ph.D. on Informatics (MAP-i). He has presented advanced tutorials at IEEE/IFIP DSN and ACM SAC conferences.

His research interests are in dependable distributed systems and are split between database replication and gossip-based communication. In particular, he has been contributing to the state of the art in gossip-based multicast for ten years, working on making them practical in a number of real application scenarios. Currently, he is applying gossip-based multicast to managing cloud computing infrastructure in the context of project DC2MS (HP IRA) and working on elastic query processing in project CumuloNimbo (EU FP7).

For more information, visit `http://gsd.di.uminho.pt/members/jop`.

Selected Publications

[CP08]   F. Campos and J. Pereira. Gossip-based service coordination for scalability and resilience. In *MW4SOC '08: Proceedings of the 3rd workshop on Middleware for service oriented computing*, pages 55–60, 2008.

[CPOR07]   N. A. Carvalho, J. Pereira, R. Oliveira, and L. Rodrigues. Emergent structure in unstructured epidemic multicast. In *IEEE/IFIP International Conference on Dependable Systems and Networks*, pages 481–490. IEEE Computer Society, 2007.

[LMPR09]   J. Leitão, J. Marques, J. Pereira, and L. Rodrigues. X-BOT: A protocol for resilient optimization of unstructured overlays. In *IEEE International Symposium On Reliable Distributed Systems*. IEEE Computer Society, 2009.

[LPR07a]   J. Leitão, J. Pereira, and L. Rodrigues. Epidemic broadcast trees. In Huai, J. and Baldoni, R. and Yen, I., editor, *IEEE International Symposium On Reliable Distributed Systems*, pages 301–310. IEEE Computer Society, 2007.

[LPR07b]   J. Leitão, J. Pereira, and L. Rodrigues. HyParView: A membership protocol for reliable gossip-based broadcast. In *IEEE/IFIP International Conference on Dependable Systems and Networks*, pages 419–428. IEEE Computer Society, 2007.

[MNOP10] M. Matos, A. Nunes, R. Oliveira, and J. Pereira. StAN: Exploiting shared interests without disclosing them in gossip-based publish/subscribe. In *Intl. Ws. on Peer-to-Peer Systems (IPTPS)*, 2010.

[MSP⁺09] M. Matos, A. Sousa, J. Pereira, R. Oliveira, E. D., and P. Murray. CLON: Overlay networks and gossip protocols for cloud environments. In *On The Move To Meaningful Internet Systems, International Symposium on Distributed Objects, Middleware, and Applications (DOA)*, Lecture Notes in Computer Science, pages 549–566. Springer Verlag, 2009.

[PRM⁺03] J. Pereira, L. Rodrigues, M.J. Monteiro, R. Oliveira, and A.-M. Kermarrec. NEEM: Network-friendly epidemic multicast. In *IEEE International Symposium On Reliable Distributed Systems*, pages 15–24. IEEE Computer Society, 2003.

[PRPO04] J. Pereira, L. Rodrigues, A. Pinto, and R. Oliveira. Low latency probabilistic broadcast in wide area networks. In *IEEE International Symposium On Reliable Distributed Systems*, pages 299–308. IEEE Computer Society, 2004.

[RBLP07] É. Rivière, R. Baldoni, H. Li, and J. Pereira. Compositional gossip: a conceptual architecture for designing gossip-based applications. *SIGOPS Oper. Syst. Rev.*, 41(5):43–50, 2007.

[RPH⁺03] L. Rodrigues, J. Pereira, S. Handurukande, R. Guerraoui, and A.-M. Kermarrec. Adaptive gossip-based broadcast. In D. Martin, editor, *IEEE/IFIP International Conference on Dependable Systems and Networks*, pages 47–56. IEEE Computer Society, 2003.