Future Generation Computer Systems 75 (2017) 85–93

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Robust and efficient membership management in large-scale dynamic networks

Pasu Poonpakdee [a], Giuseppe Di Fatta [b,*]

[a] Department of Industrial Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[b] Department of Computer Science, University of Reading, Whiteknights, Reading, Berkshire, RG6 6AY, United Kingdom

HIGHLIGHTS

- A fault-tolerant epidemic membership protocol that improves convergence speed.
- Some issues in asynchronous and dynamic networks are addressed.
- The negative effects of message interleaving events and broken links are limited.
- It incorporates a novel mechanism for global connectivity recovery.

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form
9 February 2017
Accepted 17 February 2017
Available online 1 March 2017

Keywords:
Epidemic protocols
Expander graphs
Node churn
Large-scale systems
Decentralised algorithms

ABSTRACT

Epidemic protocols are a bio-inspired communication and computation paradigm for large-scale networked systems based on randomised communication. These protocols rely on a membership service to build decentralised and random overlay topologies. In large-scale, dynamic network environments, node churn and failures may have a detrimental effect on the structure of the overlay topologies with negative impact on the efficiency and the accuracy of applications. Most importantly, there exists the risk of a permanent loss of global connectivity that would prevent the correct convergence of applications. This work investigates to what extent a dynamic network environment may negatively affect the performance of Epidemic membership protocols. A novel *Enhanced Expander Membership Protocol* (*EMP+*) based on the expansion properties of graphs is presented. The proposed protocol is evaluated against other membership protocols and the comparative analysis shows that *EMP+* can support faster application convergence and is the first membership protocol to provide robustness against global network connectivity problems.

1. Introduction

In order to disseminate information in large-scale networked systems, centralised approaches are not efficient and robust solutions as they may suffer from communication bottlenecks and fault intolerance. Epidemic, or Gossip-based, protocols are a fully decentralised paradigm for communication and computation, which are intrinsically fault-tolerant. They have been shown to be particularly suitable for information dissemination and data aggregation in large-scale networks.

A number of applications based on Epidemic protocols have been proposed to serve different purposes in different environments. For example, Epidemic protocols have been employed to implement applications for Peer-to-Peer (P2P) overlay networks [1–3], distributed computing [4], mobile ad hoc networks (MANET) [5], wireless sensor networks (WSN) [6–10], failure detection [11], distributed data mining [12–14] and exascale high performance computing [15–17].

The fundamental mechanism of Epidemic protocols inherits the advantages of robustness and scalability from a randomised communication strategy. To perform randomised communication, a peer sampling operation is required to be available as an underlying network service. In large-scale and dynamic networks it is unrealistic to assume that each node has a complete knowledge of the global topology. In this context, Epidemic Membership Protocols are used to provide a practical implementation of the peer sampling service, which returns a random node with uniform probability, similarly to a random selection from the global view of the system [18]. Instead of maintaining a complete list of nodes at each node, a membership protocol builds a local partial view (cache) of the system, which is continuously and randomly changed. The

* Corresponding author.
E-mail address: G.DiFatta@reading.ac.uk (G. Di Fatta).

http://dx.doi.org/10.1016/j.future.2017.02.033

distributed set of local views implicitly defines a dynamic overlay topology.

Several membership protocols have been proposed (e.g., [19–21]), which have been designed for generating random overlay topologies. Among other properties of graphs, the expansion quality is a fundamental mathematical concept [22] that provides a direct measure of the ability of a graph to support fast information propagation. The expansion property of graphs [23] has inspired the *Expander Membership Protocol* (*EMP*) [24]. *EMP* is based on a push–pull scheme that introduces a bias in the random node selection in order to maximise the expansion property of the overlay topology.

A specific aim of this work is to address the problems that arise in Epidemic protocols when adopted in dynamic networks. Node churn and node failures have a detrimental effect on the overlay topology and its global connectivity can be significantly weakened. Failed nodes leave behind cache entries referring to them, which have become invalid (broken links). A decrease of the propagation efficiency and a loss of global connectivity are the two main issues that may be caused by a degradation of the overlay topology. Surprisingly, in weakly connected overlay topologies, the internal mechanisms of membership protocols can turn the topology from a single connected component into multiple connected components, seriously hindering the accuracy and performance of the applications.

Many membership protocols provides mechanisms aimed at maintaining the global connectivity in strongly connected graphs, while none of them has been devoted to recover the connectivity when lost. This work includes an analysis of the external causes and of the internal mechanisms in membership protocols that may cause a loss of the global connectivity.

In this work a comprehensive description and analysis of the Enhanced Expander Membership Protocol (*EMP+*) is provided. The objectives of *EMP+* are preserving the global connectivity and providing fast and robust convergence to random overlay topologies to support the convergence speed of applications. *EMP+*, is obtained with the introduction of two procedures into *EMP*, the Interleave Management Procedure (IMP) [25] and a connectivity recovery mechanism.

EMP+ is an epidemic membership protocol that, apart from being fault-tolerant and scalable as any epidemic protocol, exhibits a better performance in terms of the convergence speed of global aggregation operations than previous protocols. Moreover, EMP+ has unique features that address some real network issues for asynchronous and dynamic conditions. The main contributions of EMP+ are as follows:

- it limits the negative effects of message interleaving events,
- it incorporates a novel mechanism for global connectivity recovery,
- it efficiently addresses the problem of broken links caused by churn and failures and
- it ultimately improves the convergence speed of global aggregation operations under dynamic network conditions.

A detailed comparative analysis for both static and dynamic networks is carried out for several membership protocols and with three different types of initial overlay topology. The experimental analysis shows that *EMP+* improves the convergence speed of global data aggregation operations and is the only membership protocols that is able to recover from the loss of global connectivity.

The rest of the paper is organised as follows. Section 2 reviews related work on Epidemic membership protocols and discusses their mechanisms and drawbacks. The intuition behind *EMP+* and a detailed description of its internal mechanisms are presented in Sections 3 and 4. The simulations that were carried out, the experimental results and their analysis are presented in Section 5. Finally, Section 6 draws some conclusions and provides possible directions of future work.

2. Membership protocols

The node sampling service provided by Epidemic Membership Protocols is considered a fundamental abstraction in distributed systems [18]. In large-scale systems, nodes cannot build and maintain a complete directory of memberships. A membership protocol builds and maintains a partial view of the system, which is used to provide the random node selection service. The distributed set of views implicitly defines an overlay topology $G = (V, E)$. A membership protocol periodically and randomly changes the local views, thus generating a sequence of random overlay topologies $\Gamma = \{G_i\}$, with $G_i = (V_i, E_i)$ being the overlay topology at the protocol cycle i.

The required assumptions are that the physical network topology is a connected graph, a routing protocol is available and an initialisation mechanism for the overlay topology is provided. Typically the formulation of Epidemic protocols is based on a periodic process (Gossip cycle) for the exchange of the local state with a random node in the system.

Several membership protocols have been proposed in the literature. This section describes how they provide the node sampling service and the broken link removal mechanism, and the drawbacks of their specific algorithms.

The Node Cache Protocol [21] is a simple membership protocol that adopts a symmetric push–pull mechanism to exchange and shuffle local membership information (node cache). At each node, the protocol contains a local cache Q of node identifiers ($|Q| \leq q_{max}$), where q_{max} is the maximum local cache size (this parameter is applied to all membership protocols used in this work). At each cycle, the local cache is sent in a push message to a node randomly selected from the local cache. When a push message is received, the local cache is sent in a reply (pull message) to the remote node originating the push message. The local cache is merged with the remote cache and the remote node ID (refreshing mechanism). The local cache is finally trimmed by randomly eliminating the number of entries exceeding q_{max}. In the Node Cache Protocol, the trimming operation is the component that may cause connectivity problems, because the removed entries could be the single link between two connected components in an overlay topology with weak connectivity. The Node Cache Protocol does not provide an explicit mechanism to remove broken links.

Cyclon [19] is a membership protocol that is an enhanced version of a basic node cache shuffling. The mechanism of Cyclon is similar to the Node Cache Protocol, which also adopts a push–pull mechanism. In Cyclon, cache entries are assigned an age attribute to limit their lifetime. At each cycle, a number of entries randomly selected from the local cache are sent (push message) to the node corresponding to the oldest entry in the local cache. When a push message is received, the node replies with a pull message containing a number of randomly selected entries from its local cache. The received entries are used to replace the donated entries at both ends. Connectivity problems in Cyclon may arise when there is message interleaving between independent pairs of push–pull exchanges involving the same node. Message interleaving has been identified as a potential threat to the accuracy of those Epidemic aggregation protocols [21] that would require the push–pull operation to be atomic. Similarly, message interleaving in Cyclon introduces the risk of removing critical cache entries, as in distributed asynchronous systems the atomicity of the push–pull operation for the cache exchange is not guaranteed. In Cyclon, the broken link removal mechanism is a separate process, which adopts a maximum cache entry lifetime and a message timeout to detect failed nodes and to remove the broken links.

Eddy [20] is arguably the most complex membership protocol. In order to provide a better random distribution of node samples in the system, Eddy tries to minimise temporal and spatial