Subscription Propagation and Content-Based Routing with Delivery Guarantees

Yuanyuan Zhao, Sumeer Bhola, and Daniel Sturman

IBM T.J.Watson Research Center
{yuanyuan, sbhola, sturman}@us.ibm.com

Abstract. Subscription propagation enables efficient content-based routing in publish/subscribe systems and is a challenging problem when it is required to support reliable delivery in networks with redundant routes. We have designed a generic model and a highly-asynchronous algorithm accomplishing these goals. Existing algorithms can be interpreted as different encodings and optimizations of the generic algorithm and hence their correctness can be derived from the generic algorithm.

1 Introduction and Related Works

A content-based publish/subscribe system consists of publishers that generate messages and subscribers that register interest in all future messages matching a predicate. The system, implemented as a network of routing brokers, is responsible for routing published messages to interested subscribers. Information providers and consumers are decoupled, since publishers need not be aware of which subscribers receive their messages, and subscribers need not be aware of the sources of the messages they receive.

Subscription propagation is a mechanism of propagating subscribers' interest throughout the broker network. It allows brokers to filter out and withhold from sending messages to parts of the network where there are no interested subscribers. This functionality is hence very important for efficiency and scalability of content-based pub/sub systems. However, the task of designing a subscription propagation algorithm is greatly challenged by several factors, especially: 1) clients' requirement of strong service guarantees such as reliable in-order, gapless delivery; 2) the existence of multiple routing paths between publishers and subscribers; 3) communication asynchrony, especially asynchrony among multiple redundant paths; 4) failures.

As a result, most previous work on subscription propagation [1,2,3] did not provide a solution that guarantees the correctness of content-based routing and hence is not capable of supporting reliable delivery in the presence of failures and multiple paths. We think this situation is due to lack of understanding of the fundamentals of the subscription propagation problem. There is no coherent theory of how subscription propagation algorithms should work in general. As a result, designing subscription propagation algorithms typically becomes isolated activities each dealing with different situations.

P. Fraigniaud (Ed.): DISC 2005, LNCS 3724, pp. 501–502, 2005.

2 Our Contributions

We have studied the structure of the subscription propagation problem and its interaction with content-based routing, reliable delivery and redundant routes and defined a general model for subscription propagation and content-based routing. We present the model in the context of a redundant routing tree where each tree node can contain multiple brokers and each edge can contain multiple broker-broker links. The model utilizes constructs that are inherently asynchronous and fault tolerant. For example, it uses subscription and conjunction sets where sets operations can inherently tolerate message duplication and re-ordering. Under this model, we define the correctness criteria of subscription propagation and content-based routing and a set of sufficient conditions for supporting reliable delivery. These sufficient conditions, such as *eventual montonicity*, allow designing of highly-asynchronous algorithms.

Using the formal model, we construct a generic asynchronous subscription propagation and content-based routing algorithm. The algorithm supports reliable delivery in the presence of multiple routing paths, broker and link failures, and communication asynchrony without requiring expensive distributed agreement between redundant paths. It provides high network availability and efficiency by allowing data message routing to choose any of the redundant paths. The generic algorithm is further refined to utilize subscription aggregation.

Many existing algorithms can be interpreted as specializations of the generic algorithm under different circumstances. For instance, algorithms that assume FIFO links and single routing paths can optimize the size of subscription state maintained at a broker. For the algorithms described in [5], the virtual time vectors can be viewed as an optimization that reduces space consumption but restricts when subscription changes can be applied. The correctness attribute of the generic algorithm applies to these algorithms.

A detailed description of this work is available in [4].

References

1. A. Carzaniga, D. Rosenblum, and A. Wolf. Design and evaluation of a wide-area event notification service. ACM Transactions on Computer Systems, 19(3):332383, August 2001.
2. G. Cugola, E. D. Nitto, and A. Fuggetta. The jedi event-based infrastructure and its application to the development of the opss wfms. IEEE Transactions on Software Engineering, 27(9):827850, September 2001.
3. P. Triantafillou and A. Economides. Subscription summarization: A new paradigm for efficient publish/subscribe systems. In Proceedings of the 24th International Conference on Distributed Computing Systems (ICDCS04), 2004.
4. Y. Zhao, S. Bhola, and D. Sturman. A general algorithmic model for subscription propagation and content-based routing with delivery guarantees. Technical report, RC23669, IBM Research, 2005.
5. Y. Zhao, D. Sturman, and S. Bhola. Subscription propogation in highly-available publish/subscribe middleware. In ACM/IFIP/USENIX 5th International Middleware Conference (Middleware 2004).