

Challenges on Internet-scale content-based communications

Mohamed Diallo, Prométhée Spathis, Serge Fdida
Lip6-NPA
{mohamed.diallo, promethee.spathis, serge.fdida}@lip6.fr

Abstract

The increasing popularity of RSS subscriptions is the proof that it is the right time to deploy large-scale dissemination systems, which deliver resources to users accordingly to the semantic of their interests. In spite of extensive research efforts and many prospective target applications, it is surprising to see that the Internet still lacks such infrastructure. We explore the design space for such infrastructure, analyze critically state-of-the-art approaches and suggest prospective research directions.

1. Introduction

Content-based communication is a paradigm in which content is made available by publishers and delivered to subscribers accordingly to the semantic of their interests [1]. In the 20th century Internet, the need for such a service is more than ever a looming need. In fact, the amount of available resources in the Internet in this web2.0 era, is getting higher by the day, and content are getting more and more dynamic and heterogeneous. This mass of available information makes it difficult for users to keep up with the fast pace of information generation. Thanks to RSS metadata, it is today possible for web users to be notified of a variety of web events: new blog posts, email notifications, hourly news, or availability of new web resources.

Even though, such interaction satisfies real world needs, its current implementation fails in two points. Firstly, it does not solve the information overload issue, because it provides filtering at a coarse-grained granularity. Second, it has severe performance implications since it is implemented using rudimentary polling. In fact, polling is aggressive and results in an unnecessarily high workload on popular web servers. In order to remediate these flaws, current RSS should embrace the content-based publish-subscribe paradigm, to fulfill the content-based communication promises. However, implementing content-based publish-

subscribe at Internet-scale is amazingly daunting. The major challenge is to devise solutions that can scale to real world workloads. The Pubsub prospective search engine boasted in 2005 more than 200000 subscribers, and claimed 3 billion matches per second for its matching engine [2]. We predict that the subscriptions database would be several orders of magnitude higher than those claims, but it is difficult to speculate about whether such a matching power is over-provisioned or not. One of the challenges we have to address is the modeling of current RSS subscriptions and publication workload.

2. Exploring the design space for Internet-scale content-based communications

The design space is particularly huge for such infrastructure. It spans a variety of paradigms involving several different research areas, notably distributed event based (DEBS), Information retrieval (IR) and Information Filtering (IF) systems. In spite of their particularities, all these research areas share a common challenge: *devising a scalable and high-throughput matching engine*. In spite of decades of valuable contributions, it is not clear whether baseline references can meet the requirements of an Internet-scale distributed matching engine. In this section, we expose the reasons of the difficulty to answer that question and our current research axes.

2.1. Distributed Event-Based Systems

The problem of distributed matching has been extensively studied in this area, and there is a plethora of claimed “efficient” DEBS [3]. DEBS commonly assumes structured subscriptions languages, and sets of brokers distributed over the wide-area Internet. However, state-of-the-art schemes performances are not well characterized and have seldom been evaluated against workloads similar to ours. So, we are currently devising an analytical framework for analyzing performances and scalability of state-of-the-art DEBS.

2.2. Prospective Search Engines

PSE are search engines, which query the future, rather than an approximate reflect of the present as current search engines do. The idea has become of interest to the IR community, with the emergence of RSS syndication [4]. In fact, baseline information retrieval indexing and matching algorithms are portable in this context. But these algorithms performances are very workload sensitive, thus performance bottlenecks and traditional IR optimizations are expected to differ. We are currently characterizing baseline algorithms and expect to evaluate their performances against real-world workloads and prospectively expect to devise efficient optimizations.

2.3. Clustering: The pivot of high performance

Current large-scale matching systems like Google operate over clusters of computers to hold the whole index and the associated metadata [5]. It is somewhat the proof that even intrinsically efficient strategies, are not sufficient to cope with real world workload and users-expectations. Clustering has become incontrovertible for building large-scale distributed applications. Today, it is possible to build clusters of thousands of computers to build high-throughput matching engines. However cluster sizes are often limited by many practical constraints such as power supply, cooling, switching technology limitations, etc. So, scaling to real workloads would require several clusters of computers.

It is estimated, being conservative with current trends and IR technologies, that cluster-based retrieval systems would require at least 30 clusters of 50000 computers by 2010 [6], Which is obviously not sensible.

3. Towards decentralized, open and federated distributed matching engines

We advocate that in order to fulfill the promise of an Internet-scale matching engine, the service should be distributed over several actors and competitive matching engines, contributing to the overall service. Each actor would manage a number of subscriptions and would be the entry point for a fraction of the metadata traffic. This challenges baseline IR and DEBS approaches, because of wide-area settings, processing power heterogeneity, decentralized control, churn, etc. P2P infrastructures like eD2k, have shown that this was indeed possible. We are currently investigating

prospective business models for such design and expect to formalize this problem.

4. Conclusions

The main challenge for Internet-scale content-based communications consists in designing a high-throughput distributed matching engine scaling to real-world workloads. Achieving such a goal necessitates an analytical framework that would enable us to analyze the performance and the characteristics of baseline schemes, inspired from DEBS, IR and IF fields. But, conventional wisdom has shown that clustering is required to meet high-throughput and high-capacity systems. However, we believe that clustering cannot be a sustainable solution to achieve scalability. Thus, we envision a decentralized, federated and self-organized distributed matching engines as an alternative to the current clustering-based solutions.

5. References

- [1] A. Carzaniga, C.P. Hall, "Content-based communication: A research agenda", *SEM '06: Proceedings of the 6th international workshop on Software engineering and middleware*, ACM, 2006.
- [2] *Pubsub.com*
- [3] Y. Zhu and H. Shen, "An efficient and scalable framework for content-based publish/subscribe systems," *Peer-to-Peer Networking and Applications*, vol. 1, no. 1, pp. 3-17, March 2008.
- [4] U. Irmak, S. Mihaylov, T. Suel, S. Ganguly, and R. Izmailov, "Efficient query subscription processing for prospective search engines," in *ATEC '06: Proceedings of the annual conference on USENIX '06 Annual Technical Conference*.
- [5] L. A. Barroso, J. Dean, and U. Urs Hölzle, "Web search for a planet: The google cluster architecture," *IEEE Micro*, vol. 23, no. 02, pp. 22-28, March 2003.
- [6] R. Baeza-Yates, C. Castillo, F. Junqueira, V. Plachouras, and F. Silvestri, "Challenges on distributed web retrieval," 2007.
- [7] *www.edonkey2000.com*