Semana de Engenharia 2010
Guimarães, 11 a 15 de Outubro

INTEGRATED ENTERPRISE MESSAGING ARCHITECTURE

José Pedro Oliveira and José Orlando Pereira
Department of Informatics
University of Minho
Emails: {jpo,jop}@di.uminho.pt

September 2010

Abstract

Current enterprise message systems are mostly centralized and have problems of scalability and integration with other system components.

This PhD thesis, which takes advantage of the cooperation between industry and university teams, will help establishing if middleware that is critical to enterprises can evolve towards a distributed solution once the centralized one no longer meets its requirements. A subsidiary contribution will be workload characterization of enterprise message systems and eventual benchmark [8] standardization. This work will have a strong impact in the state-of-the-art of both enterprise message systems and applications.

Introduction

State-of-the-art enterprise information systems are increasingly reliant on an advanced messaging infrastructure, namely in the financial services industry. In detail, messaging is the cornerstone of business-to-business (B2B) integration, multi-site and multi-service distributed applications, and business-to-consumer (B2C) integration.

Each of these usage scenarios poses different challenges: real-time requirements, priorities, historical data logging, availability, transactional guarantees, peak throughput, security, resource efficiency, congestion management and avoidance, and scalability, are some of the key items in the systems builder checklist. All these have also been the subject of much previous research on messaging systems.

Most interestingly, existing deployments seldom take advantage of advanced publish-subscribe technology [5, 4, 7] . In fact many complex real world information systems grow organically and each application is addressed by a different technology, most likely, a legacy centralized transactional messaging server, or more recently, ad-hoc Internet push technologies such as RSS and Comet.

This suggests that there are major hurdles in tailoring state-of-the-art messaging solutions to system builders assumptions and actual deployment environments, or simply in fitting leading messaging solutions to existing development best practices. The challenge is thus to overcome such hurdles both by quantitatively matching messaging technologies to usage scenarios as well as by devising novel application architectures to ease their integration.

Problem statement

In the ever evolving life of enterprise information systems the challenges are: (i) to evaluate the extent to which current centralized solutions can still be competitive notably in aspects such as Quality of Service and Real-time delivery, (ii) to determine if and how a system can evolve from a centralized solution towards a distributed solution in a transparent way. In turn this triggers questions such as:

- how to integrate technologies like Push-HTTP [1, 6], AMQP [2, 3] and other advanced protocols;
- what is the impact of application requirements and traffic characteristics (e.g. reliability, timeliness, burstiness, etc.) in the choice between centralized and distributed solutions; are there some empirical rules to guide application design?
- what is the cost/benefit ratio that can help a company decide in favor or against a decentralized solution.

Approach

By its own nature, this work on Integrated Enterprise Messaging Architecture takes advantage of the cooperation between industry and university teams. The former provides realistic case-studies where the research proposals can effectively be developed and tested under real scenarios.

The work starts by deploying messaging middleware capable of supporting two real-world applications in the

areas of financial services and remote sensing. The workload characterization will help defining application requirements and ease benchmarking of alternative messaging systems. The data and knowledge gathered by these tasks will then allow a precise answer to the question of how competitive are (or not) centralized solutions.

The work will then proceed towards an architecture that allows disruption free evolution of centralized solutions to decentralized ones.

Progress beyond the state-of-the-art

The lack of common performance metrics makes it difficult to benchmark and compare different publish/subscribe systems, as each one promotes itself through the use of private synthetic workloads. However with the characterization of real applications workloads, more realistic performance metrics will appear, which will eventually lead to the creation of standard benchmarks [8].

The analyse of traffic patterns generated by realistic workloads with different QoS requirements, and the analyse of the impact of key technologies, like network QoS service levels and real-time operating systems kernels, will also contribute to the design of better modular and interoperable publish/subscribe systems.

Objectives and results

The goal of this project is to propose and evaluate an integrated messaging architecture, by:

- Identifying and quantifying key challenges to messaging technology in real case studies.
- Refactoring existing state-of-the-art proposals to fit within the context realistic assumptions, while ensuring key properties, such as QoS requirements, fault tolerance and scalability.
- Proposing key integration points that bridge the gap between existing systems and leading messaging technology, with an emphasis on preventing disruption and promoting evolution.

References

[1] Comet Application Architecture. `http://cometdaily.com/`.

[2] AMQP Working Group. *Advanced Message Queuing Protocol Specification, version 0-10*, 2008.

[3] The Apache Software Foundation. Apache Qpid: Open Source AMQP Messaging. `http://qpid.apache.org/`.

[4] R. Baldoni, L. Querzoni, and A. Virgillito. Distributed event routing in publish/subscribe communication systems: a survey (revised version). Technical report, MIDLAB 1/2006 - Dipartimento di Informatica e Sistemistica 'A.Ruberti', Università di Roma la Sapienza, 2006.

[5] Patrick Th. Eugster, Pascal A. Felber, Rachid Guerraoui, and Anne-Marie Kermarrec. The many faces of publish/subscribe. *ACM Comput. Surv.*, 35(2):114–131, 2003.

[6] The Dojo Foundation. Bayeux Protocol 1.0draft1. `http://svn.xantus.org/shortbus/trunk/bayeux/bayeux.html`.

[7] Gero Muhl, Ludger Fiege, and Peter R. Pietzuch. *Distributed Event-Based Systems*. Springer, 2006.

[8] Kai Sachs, Samuel Kounev, Jean Bacon, and Alejandro P. Buchmann. Workload characterization of the SPECjms2007 benchmark. In Katinka Wolter, editor, *EPEW*, volume 4748 of *Lecture Notes in Computer Science*, pages 228–244. Springer, 2007.