

Serviços para Publicação e Recuperação de Componentes de Software através da Internet em Ambientes de Reuso

Robson Pinheiro de Souza
pinheiro@cos.ufrj.br

Orientadores: Cláudia M. L. Werner e Marta L. Q. Mattoso

COPPE/UFRJ - Programa de Engenharia de Sistemas e Computação
Universidade Federal do Rio de Janeiro
Caixa Postal 68511 – CEP. 21945-970
Rio de Janeiro – Brasil

Abstract

Component-based software development aims at constructing new systems from reusable components. Some studies have concluded that reuse increases both quality and productivity of software developments. Amongst the problems involved in the application of reuse techniques, two of them are treated in this work: (i) the difficulty of locating and retrieving components stored in heterogeneous repositories spread over the Internet; and (ii) the problem of publishing new components, in order to help both their integration to a reuse environment and comprehension by the developer. Aiming at supporting the solution of the two mentioned problems, this work proposes a component integration layer to a reuse infrastructure. In our research, we aim to explore and to associate the features of two different technologies: (i) the new technology of interoperability systems, in particular the *LeSelect* system, in order to solve the existing problems of interoperability between repositories of components; (ii) the popular technology of mediators, responsible for managing XML metadata extracted from each repository, providing transparency in the location and retrieval of components, and services for query optimization.

Palavras-Chave: Publicação e Recuperação de Componentes na Internet, Reutilização de Componentes, Mediadores, Sistemas Interoperáveis e XML.

1 Introdução

A construção de software baseado no uso de componentes reutilizáveis tem por objetivos reduzir a complexidade, o custo e o tempo de todo o processo de desenvolvimento, bem como aumentar a confiabilidade dos sistemas produzidos, graças ao uso de componentes previamente testados. Em [1,2], são apresentados alguns estudos que comprovam que a reutilização de componentes é capaz de aumentar tanto a qualidade quanto a produtividade de todo o processo de desenvolvimento de software.

De uma forma geral, um processo de reutilização consiste em localizar, compreender, recuperar e modificar componentes pré-existentes. Sendo assim, toda uma infra-estrutura capaz de apoiar as quatro atividades deste processo, dentre outros aspectos, torna-se necessária para o sucesso do reuso.

No entanto, a tarefa de localizar e recuperar componentes não é um processo trivial, devido aos diversos problemas de interoperabilidade existentes entre os vários repositórios que os armazenam: existe uma grande variedade de fontes de informação, utilizando diferentes modelos de dados e fisicamente espalhadas em redes de computadores. Além disso, a tarefa de compreensão de um componente por parte do desenvolvedor pode se tornar um processo bastante árduo, caso este não esteja devidamente documentado ao ser publicado.

Neste trabalho, buscamos desenvolver uma camada de software intermediária entre um ambiente de reuso e repositórios de componentes reutilizáveis na Web capaz de: (i) tratar as questões de distribuição e heterogeneidade das fontes de informação, provendo ao ambiente de reuso facilidades de consulta e transparência na recuperação de componentes; (ii) prover aos desenvolvedores de software facilidades para publicar seus próprios componentes com a devida documentação em um site qualquer da Internet, de tal forma que um ambiente de reuso possa detectar e integrar rapidamente em sua base de metadados as informações de novos componentes publicados.

Três diferentes sistemas são utilizados para compor a estrutura geral da ferramenta a ser construída nesta dissertação: (i) o *Odyssey* [3], uma infra-estrutura de apoio de desenvolvimento de software baseado nos conceitos de Reutilização e de Domínios de Aplicações; (ii) o *GOA* [4], um sistema capaz de gerenciar a armazenagem e a recuperação de objetos em disco; (iii) o *LeSelect* [5], um sistema interoperável capaz de integrar fontes de informações espalhadas pela Internet.

2 Trabalhos Relacionados

A seguir, listamos alguns exemplos de trabalhos similares que buscam soluções para os problemas de integração de repositórios de componentes no apoio ao processo de reutilização.

Braga [6] descreve o uso da arquitetura de interoperabilidade HIMPARG como camada de integração de dados para o suporte a ambientes de reuso. Seu trabalho é baseado no uso de mediadores¹ e tradutores [7], o que garante flexibilidade e transparência na recuperação de componentes. O modelo de representação utilizado é o orientado a objetos. No entanto, o projeto da arquitetura HIMPARG está voltado apenas para integração de sistemas gerenciadores de banco de dados, não levando em conta a integração de uma fonte de informação qualquer na Internet, como é o objetivo deste trabalho.

¹ Um mediador é um componente de software que encapsula a representação de múltiplas bases de dados, tratando os problemas de conflitos semânticos gerados por diferentes modelos de dados e diferentes esquemas, e provendo um acesso uniforme aos dados por ele integrados.

O *Agora System* [8] descreve uma máquina de busca para recuperar componentes de código na Internet, tal como JavaBeans e componentes CORBA. Este sistema utiliza um banco de dados interno para registrar os links onde se encontram os componentes. O problema desta abordagem é que os componentes poderão não ser avaliados durante uma pesquisa, pois o site pode estar momentaneamente fora do ar. Em nossa proposta, o uso de mediadores gerenciando metadados garante o sucesso da busca mesmo que repositório esteja fora do ar durante a pesquisa, permitindo que o usuário saiba da existência dos componentes e consequentemente possa recuperá-los mais tarde.

3 Arquitetura de Integração

Analisando a arquitetura proposta (figura 1) de baixo para cima, o primeiro nível é composto pelo sistema *LeSelect* (figura 2). Este sistema será responsável por tratar os problemas de interoperabilidade existentes entre fontes de dados distribuídas e heterogêneas, permitindo o acesso transparente a qualquer repositório de componentes ligado à Internet. Além disso, o *LeSelect* será responsável por prover os metadados em XML, que descrevem a semântica das informações armazenadas em um repositório, para a camada superior.

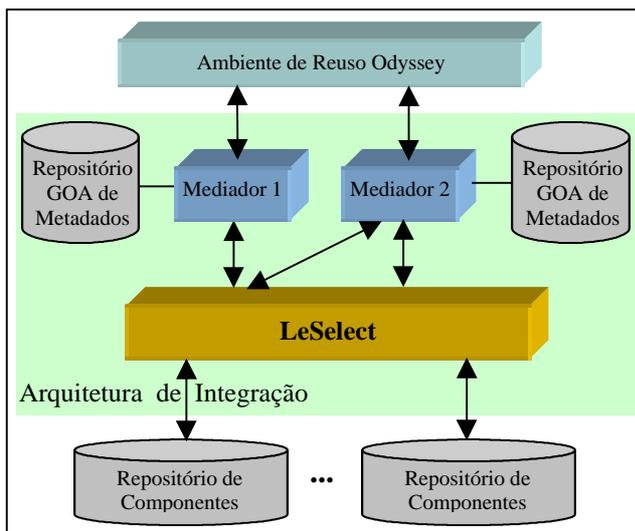


Figura 1 - Arquitetura de Integração de Repositórios de Suporte ao Odyssey.

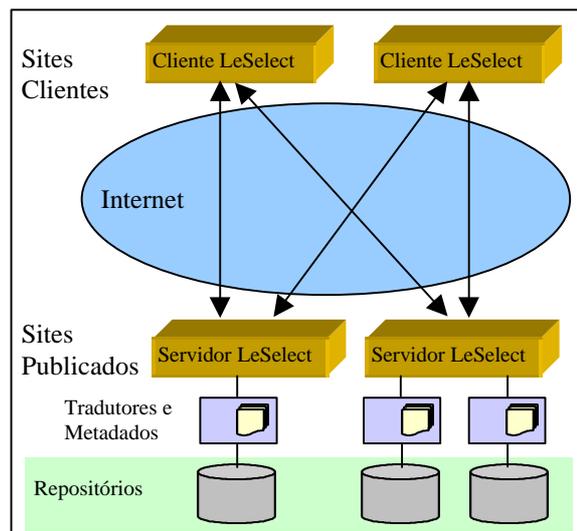


Figura 2 - Sistema Interoperável LeSelect.

O segundo nível é composto pela camada de Mediadores. Esta camada é responsável por gerenciar metadados no padrão XML, necessários para prover transparência na localização das informações e serviços de otimização de consulta. Para armazenar estes metadados, utilizamos o sistema *GOA*, que deverá ser estendido com novos serviços para gerenciar a tecnologia XML. Cada mediador, que encapsula um conjunto de metadados específicos para um determinado domínio de aplicação, implementa uma lista de serviços capaz de atender as seguintes funcionalidades: (i) receber as consultas provenientes do ambiente de reuso no padrão OQL e repassá-las ao processador de consultas do GOA; (ii) construir consultas no padrão SQL² utilizando a lista de localizações (URL do repositório + nome da tabela correspondente ao componente) retornadas pelo processador de consultas e repassá-las ao *LeSelect*; (iii) receber os resultados provenientes do *LeSelect* e repassá-los ao ambiente de reuso. Além disso, outros serviços gerais deverão ser providos pela camada de mediação,

² O *LeSelect* adota internamente o modelo de dados relacional e o padrão SQL de linguagem de consulta.

como, por exemplo, a extração de metadados via *LeSelect*, nos casos de inclusão de um novo repositório e de atualização dos metadados de um repositório já publicado.

4 Metodologia de Trabalho

A metodologia de desenvolvimento deste trabalho foi dividida em cinco etapas distintas: (i) pesquisa sobre o assunto e trabalhos relacionados; (ii) especificação dos requisitos; (iii) implementação e teste; (iv) avaliação da implementação; (v) redação da dissertação.

Dentro da primeira etapa, os primeiros passos foram voltados para o estudo da viabilidade de utilização do sistema *LeSelect*, onde buscamos destacar as vantagens que ele poderia nos oferecer, assim como as deficiências que teriam que ser supridas. Em seguida, fizemos um levantamento dos trabalhos relacionados, buscando conhecer melhor o problema e encontrar novas idéias para serem empregadas em nossa proposta.

Atualmente, estamos finalizando a segunda etapa de nosso trabalho, detalhando os algoritmos a serem utilizados e a integração das tecnologias (XML, CORBA, Java, etc) que farão parte da implementação da proposta. A previsão de término é março de 2001.

5 Considerações Finais

Nesta dissertação, buscamos propor novas soluções para o problema da integração de repositórios de componentes em ambientes de reuso. Nossa proposta explora a integração da tecnologia de sistemas interoperáveis associada a tecnologia de mediadores. Com isso, buscamos prover total transparência na localização e recuperação de componentes, além de expandir o processo de publicação de novos componentes para todo universo da Internet. Vale a pena ressaltar, também, que a utilização de documentos XML, associados aos componentes publicados, aumentará o poder de semântica de seus dados, facilitando o seu entendimento por parte do usuário e aumentando a precisão das consultas processadas.

O produto final deste trabalho terá como principal aplicabilidade e base de testes o *Projeto Odyssey*, que visa a construção de uma infra-estrutura de apoio ao reuso, atualmente em desenvolvimento pelo grupo de reutilização em software da COPPE/UFRJ.

6 Referências

- [1] Basili, V.; Briand, L.; Melo, W.; "How reuse influences productivity in object-oriented systems"; Comm. of the ACM, 39(10): 104-116; 1996.
- [2] Lim, W. C.; "Effects of reuse on quality, productivity and economics"; IEEE Software, 11(5): 23-29; 1994.
- [3] Werner, C. et. al "Infra-estrutura Odyssey: estágio atual", XIV Simpósio Brasileiro de Engenharia de Software, Caderno de Ferramentas; João Pessoa, Outubro 2000.
- [4] Mattoso, M. et al; "Persistência de Componentes num Ambiente de Reuso", XIV Simpósio Brasileiro de Engenharia de Software, Caderno de Ferramentas; João Pessoa, Outubro 2000;
- [5] Xhumari, F.; Mokrane, A.; "Le Select: a Middleware System for Publishing Autonomous and Heterogenous Information Sources"; INRIA, Groupe Caravel; France, 1999.
- [6] Braga, R.; Mattoso, M.; Werner, C.; "The Use of Mediators for Component Retrieval in a Reuse Environment"; TOOLS-30 USA'99, IEEE CS Press, pp.542-546; California, August 1999.
- [7] Wiederhold, G.; Genesereth, M.; "The Conceptual Basis for Mediation Services"; IEEE Expert, Vol. 12, No. 5; Sep-Oct, 1997.
- [8] Seacord, R.; Hissan, S.; Wallnau, K; "Agora: A Search Engine for Software Components"; Technical Report CMU/SEI-98-TR-011; August, 1998.