

Rodrigo Carneiro Henrique

**Uma Arquitetura de Serviços de Acesso a
Dados Estruturados em Aplicações Científicas**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em
Informática do Departamento de Informática do Centro Técnico
Científico da PUC-Rio como requisito parcial para obtenção do
grau de Mestre em Informática.

Orientador : Prof. Renato Fontoura de Gusmão Cerqueira

Co-Orientador: Prof. Carlos Roberto Serra Pinto Cassino

Rio de Janeiro
Abril de 2009

Rodrigo Carneiro Henrique

**Uma Arquitetura de Serviços de Acesso a
Dados Estruturados em Aplicações Científicas**

Dissertação apresentada ao Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Informática. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Renato Fontoura de Gusmão Cerqueira

Orientador

Departamento de Informática — PUC-Rio

Prof. Carlos Roberto Serra Pinto Cassino

Co-Orientador

Departamento de Informática — PUC-Rio

Prof. Marco Antonio Casanova

Departamento de Informática — PUC-Rio

Prof. Markus Endler

Departamento de Informática — PUC-Rio

Prof. José Eugenio Leal

Coordenador Setorial do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 08 de Abril de 2009

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Rodrigo Carneiro Henrique

Graduou-se em Bacharel em Informática pela PUC-Rio em 2004. Trabalha como desenvolvedor de *software* desde 2000. Em 2002, trabalhou no desenvolvimento das ferramentas de segurança digital das eleições nacionais pela Módulo Security Solutions, diretamente no Tribunal Superior Eleitoral (TSE). Desde 2004 atua como desenvolvedor de *software* pelo Tecgraf/PUC-Rio em projetos realizados em parceria com a Petrobras.

Ficha Catalográfica

Henrique, Rodrigo Carneiro

Uma Arquitetura de Serviços de Acesso a Dados Estruturados em Aplicações Científicas / Rodrigo Carneiro Henrique; orientador: Renato Fontoura de Gusmão Cerqueira; co-orientador: Carlos Roberto Serra Pinto Cassino. — Rio de Janeiro : PUC-Rio, Departamento de Informática, 2009.

v., 54 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Serviços de Acesso a Dados Científicos. 3. Dados Estruturados. 4. Objetos por Valor. 5. Middleware. 6. CORBA. 7. Sistemas Distribuídos. I. Cerqueira, Renato Fontoura de Gusmão. II. Cassino, Carlos Roberto Serra Pinto. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

Agradecimentos

Agradeço primeiramente à minha mãe, Marisa, por me oferecer sempre o melhor em termos de educação, mesmo precisando fazer um imenso esforço para custear meus estudos. Aos meus avós, que a ajudaram nesta tarefa e aos meus familiares em geral.

Agradeço especialmente ao meu orientador Renato Cerqueira e ao meu co-orientador Carlos Cassino pelos ensinamentos e conselhos que me deram durante este trabalho e que continuam dando no meu dia-a-dia nos projetos do Tecgraf/PUC-Rio.

Por fim, agradeço a todos os meus amigos do Tecgraf/PUC-Rio e da Petrobras que me apoiaram na decisão de encarar este imenso desafio.

Resumo

Henrique, Rodrigo Carneiro; Cerqueira, Renato Fontoura de Gusmão; Cassino, Carlos Roberto Serra Pinto. **Uma Arquitetura de Serviços de Acesso a Dados Estruturados em Aplicações Científicas**. Rio de Janeiro, 2009. 54p. Dissertação de Mestrado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Aplicações científicas trabalham, tipicamente, com grandes volumes de dados que possuem uma representação complexa e própria da aplicação que os utiliza. Essas características representam um grande desafio para o compartilhamento de dados entre aplicações desta natureza. Este trabalho tem como objetivo principal apresentar uma arquitetura de serviços de software que permita um acesso flexível e eficiente a grandes volumes de dados disponibilizados por aplicações científicas. São apresentados estudos de caso para ilustrar a flexibilidade promovida pela arquitetura através de experimentos com dados cuja representação é fortemente baseada em dados reais utilizados por aplicações científicas desenvolvidas pelo Tecgraf/PUC-Rio. Há, ainda, uma avaliação de diferentes técnicas de codificação de dados realizada através de experimentos criados para medir o desempenho alcançado na implementação da arquitetura.

Palavras-chave

Serviços de Acesso a Dados Científicos. Dados Estruturados. Objetos por Valor. Middleware. CORBA. Sistemas Distribuídos.

Abstract

Henrique, Rodrigo Carneiro; Cerqueira, Renato Fontoura de Gusmão (Advisor); Cassino, Carlos Roberto Serra Pinto (Co-Advisor). **An Architecture for Structured Data Access Services in Scientific Applications**. Rio de Janeiro, 2009. 54p. MsC Thesis — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Scientific applications usually handle large amount of data that have a proprietary and complex representation. This characteristics represent a great challenge for sharing data between scientific applications. The main goal of this work is to provide an architecture of software services that allows a flexible and efficient access to large amount of data served by such applications. Case studies are presented to show the flexibility that we can achieve with this architecture. These experiments are strongly based in actual data used in scientific applications developed by Tecgraf/PUC-Rio. We also present an evaluation of different techniques of data encoding based on experiments conducted to measure the performance achieved by an implementation of the proposed architecture.

Keywords

Scientific Data Access Services. Structured Data. Objects by Value. Middleware. CORBA. Distributed Systems.

Sumário

1	Introdução	10
2	Trabalhos Relacionados	13
2.1	Scientific Data Server (SDS)	14
2.2	Service Data Objects (SDO)	16
2.3	OpenSpirit	17
2.4	Protocol Buffers	18
2.5	Representação de dados em CORBA	19
2.6	Considerações Finais	20
3	O Arcabouço de Serviços de Acesso a Dados Científicos	21
3.1	Visão Geral da Arquitetura	22
3.2	Detalhamento das Interfaces	24
4	Avaliação	31
4.1	Estudos de Caso	31
4.2	Desempenho	40
5	Conclusão	46
5.1	Trabalhos Futuros	47
6	Referências Bibliográficas	48
A	Apêndice	50
A.1	IDL do Serviço de Dados	50
A.2	IDL do SCS	50
A.3	IDL dos Dados de Poço	50
A.4	Arquivos proto com as definições de dados de poço	50

Lista de figuras

2.1	Arquitetura SDS	15
2.2	Arquitetura SDO	16
3.1	A arquitetura do servidor de dados estruturados	25
3.2	Um componente SCS	26
3.3	O componente do Servidor de Dados	27
4.1	Arquitetura do OpenBus	36
4.2	Aplicações integradas através do OpenBus	38
4.3	Gráfico com os tempos obtidos nos testes de desempenho	42

Lista de tabelas

4.1	<i>Software</i> utilizado nos testes de desempenho	40
4.2	Configuração das máquinas usadas nos testes de desempenho	41
4.3	Configuração da máquina usada nos testes de desempenho do experimento <i>future</i>	41
4.4	Tempos obtidos nos testes de desempenho	41