

Marcelo Nery dos Santos

**GridFS — Um Servidor de Arquivos para
Grades e Ambientes Distribuídos
Heterogêneos**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática da PUC-Rio

Orientador: Prof. Renato Cerqueira

Rio de Janeiro
Março de 2006

Marcelo Nery dos Santos

**GridFS — Um Servidor de Arquivos para
Grades e Ambientes Distribuídos
Heterogêneos**

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

Prof. Renato Cerqueira

Orientador

Departamento de Informática — PUC-Rio

Prof. Noemi Rodriguez

Departamento de Informática — PUC-Rio

Prof. Vinod Rebello

Universidade Federal Fluminense — UFF

Prof. José Eugênio Leal

Coordenador Setorial do Centro Técnico Científico —
PUC-Rio

Rio de Janeiro, 22 de Março de 2006

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

Marcelo Nery dos Santos

Graduou-se em Ciência da Computação pela UFPE (Universidade Federal de Pernambuco) em 2002. Foi bolsista do Programa Especial de Treinamento (PET) da CAPES.

Ficha Catalográfica

Santos, Marcelo Nery dos

GridFS — Um Servidor de Arquivos para Grades e Ambientes Distribuídos Heterogêneos/ Marcelo Nery dos Santos; orientador: Renato Cerqueira. — Rio de Janeiro : PUC-Rio, Departamento de Informática, 2006.

v., 101 f: il. ; 29,7 cm

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

Inclui referências bibliográficas.

1. Informática - Teses. 2. Servidor de Arquivos. 3. CORBA. 4. Middleware. 5. Computação em Grade. 6. Sistemas Distribuídos. I. Cerqueira, Renato. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Agradecimentos

Gostaria de agradecer primeiramente a Deus, por todas as oportunidades que recebi ao longo dos anos.

Aos meus pais e minha família, pelo apoio, energia, amor e carinho que me foram dados nesse período.

Ao meu orientador, Renato Cerqueira, pelas horas e horas de reuniões, pelos comentários, sugestões, diretrizes e pela confiança depositada em minha pessoa. Aos membros da Banca, Noemi Rodriguez e Vinod Rebello, pela aceitação do convite, e pelos comentários e sugestões.

À Karina, pelo carinho, pela constante fonte de estímulo e pelo esforço na revisão de tantas e tantas páginas e versões preliminares.

Aos meus amigos que estiveram sempre presentes em vários momentos, em especial ao Mauro Faccenda, pelas infindáveis vezes que o procurei para discutirmos temas referentes a este trabalho. E a todos aqueles amigos que simplesmente estavam presentes física ou espiritualmente. Seria impossível listar todos aqui.

Aos laços de amizade criados ou renovados durante minha estada no Rio de Janeiro, pelos encontros, conversas e excursões ao longo dos últimos dois anos. Momentos de extrema descontração, especialmente valiosos para levar um trabalho árduo como esse adiante.

Aos professores, funcionários e amigos na PUC-Rio e Tecgraf, pelo excelente ambiente de trabalho oferecido.

Gostaria de agradecer também ao apoio e financiamento da Petrobras e do Projeto GIGA da RNP, com recursos do FUNTTEL e FINEP, para o desenvolvimento deste projeto. Bem como a CAPES e a PUC-Rio, pelo apoio financeiro cedido.

Enfim, a todos aqueles que estão presentes e contribuem com a minha jornada. Muito obrigado.

Resumo

Santos, Marcelo Nery dos; Cerqueira, Renato. **GridFS — Um Servidor de Arquivos para Grades e Ambientes Distribuídos Heterogêneos**. Rio de Janeiro, 2006. 101p. Dissertação de Mestrado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A computação em grade permite o uso de recursos computacionais distribuídos em várias redes, para a execução de tarefas que requerem um alto poder computacional. Uma infra-estrutura para grades pode ser utilizada para auxiliar na execução dessas tarefas e coordenar as atividades relacionadas, possivelmente envolvendo a disponibilização dos arquivos de dados para as tarefas em execução nos nós da grade.

O GridFS é um sistema para o compartilhamento de arquivos em grades e ambientes distribuídos heterogêneos. Ao disponibilizar servidores em diversas máquinas, é possível construir uma federação do sistema, integrando os diversos sistemas de arquivos locais e abrindo possibilidades de armazenamento da ordem de terabytes. O sistema proposto foi modelado e desenvolvido levando em consideração diversos aspectos como escalabilidade, interoperabilidade e desempenho.

O GridFS agrega características de diferentes sistemas de compartilhamento de arquivos. Ele oferece uma API para acesso remoto aos dados, disponibiliza a opção de cópia de arquivos entre os servidores e fornece algumas funções especiais para os ambientes de computação em grade. Além de definir as características e os aspectos de implementação do sistema, esta dissertação apresenta alguns resultados experimentais sobre o desempenho e escalabilidade do serviço e, como forma de avaliação, discute a integração do GridFS ao CSBase, um *framework* utilizado no desenvolvimento de sistemas para computação em grade.

Palavras-chave

Servidor de Arquivos, CORBA, Middleware, Computação em Grade, Sistemas Distribuídos

Abstract

Santos, Marcelo Nery dos; Cerqueira, Renato. **GridFS — A File Server for Grids and Heterogeneous Distributed Environments**. Rio de Janeiro, 2006. 101p. MSc. Dissertation — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Grid computing allows the use of distributed networks resources for tasks requiring a high processing power. A Grid infra-structure may help in the execution of these tasks and is able to coordinate their related activities, possibly regarding the provision of data files for the tasks executing in the grid nodes.

GridFS is a system that enables data sharing in grid and heterogeneous distributed environments. By deploying servers over several nodes, it is possible to build a federated system integrating all local file systems and leveraging possibilities for tera-scale sized data storage. The proposed system was modeled and developed considering several aspects such as scalability, interoperability and performance.

GridFS combines some characteristics from diverse file management systems, that is, GridFS provides an API for remote data access, copy operations allowing file transfers between servers, and some special features for grid environments. Apart from defining system characteristics and implementation aspects, this dissertation shows some experimental results about the service scalability and performance, and, as an evaluation, discusses the integration of GridFS with CSBase, a framework used to develop systems for grid computing.

Keywords

File Server, CORBA, Middleware, Grid Computing, Distributed Systems

Sumário

1	Introdução	11
1.1	Características	13
1.2	Contribuições	16
1.3	Estrutura da Dissertação	17
2	Trabalhos Relacionados	18
2.1	GridFTP	18
2.2	Reliable File Transfer (RFT) Service	20
2.3	Global Access to Secondary Storage (GASS) Service	20
2.4	Stork	21
2.5	DiskRouter	22
2.6	Avaki Data Grid	22
2.7	Sistemas de Arquivos Distribuídos	22
2.8	Considerações sobre os sistemas	23
3	GridFS	26
3.1	Características	27
3.1.1	Federação de Servidores	27
3.1.2	Interface Orientada a Objetos	28
3.1.3	Escalabilidade	28
3.1.4	Desempenho	29
3.1.5	Dados Históricos e Metadados	30
3.1.6	Persistência das Referências para Objetos Remotos	30
3.1.7	Interoperabilidade	31
3.2	Aspectos de Implementação	31
3.2.1	Escalabilidade em CORBA	38
3.2.2	Desempenho no Acesso aos Dados	45
3.2.3	Mount Points	47
3.2.4	Estimativa de Taxas de Transferência	47
3.3	Processo de Desenvolvimento	48
3.4	Outros Mecanismos de Acesso aos Dados	49
3.4.1	Servidor FTP	49
3.4.2	FUSE	50
3.5	Limitações do Sistema	50
3.5.1	Operações de Remoção	51
3.5.2	Acesso Concorrente	51
3.5.3	Caching	52
4	Resultados Experimentais	53
4.1	Testes de Escalabilidade e Desempenho	56
4.1.1	FTP	57
4.1.2	NIO	57
4.1.3	CORBA	63

4.1.4	Comparação entre os métodos	66
4.1.5	Limites no Número de Transferências Simultâneas	68
4.2	Escalabilidade no Número de Referências Remotas e Canais de Acesso	71
4.3	Testes em Sistemas Legados (NFS x FUSE)	72
4.4	Testes de Desempenho no Acesso Remoto (NFS x Canais CORBA)	73
5	Estudo de Caso: CSBase	74
5.1	Arquitetura	75
5.1.1	Desktop Web	75
5.1.2	Repositório de Dados	76
5.1.3	Servidor Principal	77
5.1.4	Servidor de Gerência de Algoritmos	77
5.1.5	Servidor de Arquivos	78
5.2	Serviços do CSBASE	78
5.3	Considerações sobre o uso do GridFS no CSBase	80
6	Conclusões	81
6.1	Trabalhos Futuros	82
7	Referências Bibliográficas	85
A	IDL CORBA	89

Lista de Figuras

3.1	Diagrama de Atividades (createFile)	37
3.2	Abordagem do <i>RootPOA</i>	40
3.3	Abordagem do <i>DefaultServant</i>	41
3.4	Consumo de Memória da Máquina Virtual	42
3.5	Hierarquia de POAs	43
4.1	Topologia (rede local)	56
4.2	Variação do Tamanho do Buffer TCP (rede local)	59
4.3	Topologia (Rede GIGA)	60
4.4	Variação do Tamanho do Buffer TCP (Rede GIGA)	60
4.5	Variação do Número de Fluxos Paralelos	63
4.6	Variação do Tamanho do Bloco CORBA Utilizado nas Chamadas Remotas	64
4.7	Variação do Tamanho do Bloco CORBA na Rede GIGA	65
4.8	Comparação Geral dos Métodos (Disco)	66
4.9	Comparação Geral dos Métodos (CPU)	67
4.10	Comparação Geral dos Métodos (Carga Média)	67
4.11	Uso de Memória com a Política <i>DefaultServant</i>	72
5.1	Arquitetura do CSBase	76

Lista de Tabelas

3.1	Políticas do POA usadas para o FileServer	44
3.2	Políticas do POA usadas para o RemoteFile	44
3.3	Políticas do POA usadas para os Canais	44
4.1	Características do Servidor (PUC-Rio)	54
4.2	Características do Servidor (USP)	54
4.3	Características do Disco	54
4.4	Velocidades Obtidas nos Testes da Rede	54
4.5	Comparação entre o NFS e o GridFS (Acesso Remoto)	73