

Rafael Barbosa Nasser

McCloud Service Framework
Arcabouço para desenvolvimento de serviços
baseados na Simulação de Monte Carlo na Cloud

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-graduação em Informática da PUC-Rio como requisito parcial para obtenção do título de Mestre em Informática.

Orientadora: Prof. Karin Koogan Breitman

Rio de Janeiro
Abril 2012



Rafael Barbosa Nasser

**McCloud Service Framework:
Arcabouço para desenvolvimento de
serviços baseados na Simulação de Monte
Carlo na Cloud**

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

Prof. Karin Koogan Breitman

Orientadora

Departamento de Informática – PUC-Rio

Prof. Hélio Côrtes Vieira Lopes

Departamento de Informática –
PUC-Rio

Prof. José Viterbo Filho

UFF

Prof. José Eugenio Leal

Coordenador Setorial do Centro
Técnico Científico – PUC-Rio

Rio de Janeiro, 04 de abril de 2012

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

Rafael Barbosa Nasser

O autor é graduado em Engenharia de Computação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) com domínio adicional em Empreendedorismo pela mesma Universidade. Sócio e diretor de tecnologia da Nit Soluções 2.0, empresa especializada em desenvolvimento de software, onde gerenciou importantes projetos para empresas de médio e grande porte. Pesquisador do Laboratório de Engenharia de Software da PUC-Rio.

Ficha Catalográfica

Nasser, Rafael Barbosa

McCloud service framework: arcabouço para desenvolvimento de serviços baseados na simulação de Monte Carlo na Cloud / Rafael Barbosa Nasser ; orientadora: Karin Koogan Breitman. – 2012.

107 f. : il. (color.) ; 30 cm

Dissertação (mestrado)—Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2012.

Inclui bibliografia

1. Informática – Teses. 2. Método de Monte Carlo. 3. Computação na nuvem. 4. Computação paralela. 5. Sistema de apoio à decisão. 6. Engenharia de software. I. Breitman, Karin Koogan. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Aos meus pais, Alberto e Mariana,
que empreenderam minha vida
com amor, sabedoria e suor.

Agradecimentos

Agradeço a minha esposa, Giulianna, meu eterno amor, pelo insubstituível apoio e carinho, e sem dúvida, pela paciência. E aos meus irmãos, Leonardo e Renato, eternos companheiros.

Agradeço a minha tia, Ph.D. Lilian Nasser, que o sucesso acadêmico sempre me serviu de inspiração, e que sempre esteve disponível para ajudar. E ao meu tio, Leibo Kampela, cuja sabedoria sempre estará ao meu lado.

Agradeço a minha orientadora, D.Sc. Karin K. Breitman, que soube como ninguém motivar o meu lado acadêmico e direcionar e conectar interesses em busca de grandes resultados. Espero sempre retribuir seu apoio com resultados.

Agradeço ao Ph.D. Marcos Antonio Casanova, meu orientador de graduação, por incentivar meu ingresso no Mestrado e pelo constante apoio. E ao D.Sc. Hélio Côrtes Vieira Lopes pelo apoio na escolha do Estudo de Caso.

Agradeço ao M.Sc. Americo Barbosa da Cunha Junior, e seu orientador de Doutorado, Ph.D. Rubem Sampaio, pela adoção da solução proposta neste trabalho em sua pesquisa e dedicado apoio, permitindo assim, a realização do Estudo de Caso apresentado. Também agradeço a Math Works e seu representante brasileiro pelo fornecimento da licença temporária do MATLAB para realização deste Estudo de Caso.

Agradeço a Roberto Mello e Gustavo Carvalho e toda a equipe da Nit, que possibilitaram, com seu apoio profissional e pessoal, minha dedicação a este projeto.

Agradeço a Microsoft pelo fornecimento de créditos para execução dos experimentos e a seu receptivo e experiente time, em especial, Otavio Pcego Coelho e Lorenzo Ridolfi, com quem tive a oportunidade de interagir nessa caminhada.

Agradeço a PUC-Rio pelos auxílios concedidos e a todos os professores do Departamento de Informática, cujo convívio, ano a ano, torna-se cada vez mais gratificante.

Resumo

Nasser, Rafael Barbosa; Breitman, Karin Koogan; **McCloud Service Framework: Arcabouço para desenvolvimento de serviços baseados na Simulação de Monte Carlo na Cloud**. Rio de Janeiro, 2012, 107 p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O investimento em infraestrutura computacional para suportar picos de processamento de curta duração ou sazonais pode gerar desperdícios financeiros, em razão de, na maior parte do tempo, estes recursos ficarem ociosos. Além disso, em muitas soluções, o tempo de resposta é crítico para atendimento dos requisitos do negócio, tornando, muitas vezes, a solução economicamente inviável. Neste cenário é fundamental a alocação inteligente de recursos computacionais em função da demanda por processamento, custo desta alocação e requisitos do negócio. A Simulação de Monte Carlo é um método estatístico utilizado para resolver uma ampla gama de problemas científicos e de engenharia. Quando aplicado a problemas reais, muitas vezes apresenta os desafios mencionados. Computação na nuvem surge como uma alternativa para disponibilizar recursos computacionais sob demanda, gerando economia de escala sem precedentes e escalabilidade quase infinita. Ao alinhar uma moderna arquitetura à nuvem é possível encapsular funcionalidades e oferecer um leque de serviços que antes seriam restritos a domínios específicos. Neste trabalho propomos um arcabouço genérico, que permite a disponibilização de um leque de serviços baseados na Simulação de Monte Carlo, fazendo uso racional da elasticidade provida pela nuvem, a fim de alcançar melhores patamares de eficiência e reuso.

Palavras-chave

Informática; Engenharia de Software; Método de Monte Carlo; Computação na Nuvem; Computação Paralela; Sistema de Apoio à Decisão.

Abstract

Nasser, Rafael Barbosa; Breitman, Karin Koogan (Advisor); **McCloud Service Framework: Development Services of Monte Carlo Simulation in the Cloud**. Rio de Janeiro, 2012, 107 p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The investment in computing infrastructure to attend seasonal demand or processing peak can generate financial waste, because the most of the time these resources are idle. In addition, in many solutions the response time are critical to attend business requirements, which often, turn the solution economically unviable. In this scenario it is essential intelligent allocation of computing resources according to the demand for processing, allocation and cost of business requirements. The Monte Carlo Simulation is a statistical method widely used to solve a wide range of scientific and engineering problems. When applied to real problems usually have the challenges mentioned. Cloud Computing is an alternative to providing on-demand computing resources, generating economies of scale unprecedented and almost infinite scalability. Aligning a modern architecture to the cloud is possible to encapsulate functionality and offer a range of services that would previously have been restricted to specific areas. In this paper we are interested in building a generic framework, that may provide a range of services based on Monte Carlo, make rational use of the elasticity provided by the cloud in order to achieve better levels of efficiency and reuse.

Keywords

Computer Science; Software Engineering; Monte Carlo Method; Cloud Computing; Parallel Computing; Decision Support Systems.

Sumário

| | | |
|-----|----------------------------------|----|
| 1 | Introdução | 12 |
| 1.1 | Motivação | 12 |
| 1.2 | Principais Contribuições | 13 |
| 1.3 | Organização do Trabalho | 13 |
| 1.4 | Público Alvo | 15 |
| 1.5 | Desafios Enfrentados | 16 |
| 1.6 | Evolução do Tema | 17 |
| 2 | Computação na Nuvem | 18 |
| 2.1 | Definição | 18 |
| 2.2 | Características Essenciais | 20 |
| 2.3 | Ambientes de Computação na Nuvem | 23 |
| 3 | Método de Monte Carlo | 25 |
| 3.1 | Definição | 25 |
| 3.2 | Características Essenciais | 26 |
| 3.3 | Exemplo Prático | 27 |
| 3.4 | Novas Perspectivas | 28 |
| 4 | Solução Proposta | 30 |
| 4.1 | Escopo Proposto | 30 |
| 4.2 | Arquitetura Conceitual | 31 |
| 4.3 | Plataforma | 33 |
| 5 | Modelo de Programação | 38 |
| 5.1 | Visão Geral | 38 |
| 5.2 | Etapa Split | 39 |
| 5.3 | Etapa Process | 40 |
| 5.4 | Etapa Merge | 41 |
| 6 | O Arcabouço de Serviços | 43 |
| 6.1 | O Arcabouço | 43 |
| 6.2 | Prontos de Extensão | 44 |
| 6.3 | Os Serviços | 46 |
| 6.4 | Casos de Uso | 48 |
| 6.5 | Configuração e Instrumentação | 49 |
| 6.6 | Visão Geral | 51 |
| 7 | Prova de Conceito | 54 |
| 7.1 | Domínio | 54 |
| 7.2 | Implementação | 56 |
| 7.3 | Performance | 60 |
| 7.4 | Análise dos Resultados | 61 |
| 8 | Duração x Custo | 63 |
| 8.1 | Fronteiras de Eficiência | 63 |
| 8.2 | Notação da Modelagem | 64 |
| 8.3 | Modelagem Proposta | 64 |

| | | |
|------|------------------------------------|-----|
| 8.4 | Aplicação Teórica | 69 |
| 8.5 | Implementação na Prova de Conceito | 70 |
| 8.6 | Generalização do Modelo | 71 |
| 9 | Estudo de Caso | 72 |
| 9.1 | Domínio | 72 |
| 9.2 | Cenário Atual | 73 |
| 9.3 | Adaptações | 74 |
| 9.4 | Implementação | 76 |
| 9.5 | Performance | 81 |
| 9.6 | Análise dos Resultados | 86 |
| 10 | Guia do Usuário | 88 |
| 10.1 | Finalidade do Guia | 88 |
| 10.2 | Criando sua Conta no Azure | 88 |
| 10.3 | Criando sua Conta de Armazenamento | 90 |
| 10.4 | Criando seu Primeiro Serviço | 91 |
| 10.5 | Consumindo seu Primeiro Serviço | 93 |
| 10.6 | Instalando Pré-requisitos | 94 |
| 10.7 | Comentários | 96 |
| 11 | Conclusão | 97 |
| 11.1 | Descobertas | 97 |
| 11.2 | Lições Aprendidas | 98 |
| 11.3 | Limitações | 99 |
| 11.4 | Trabalhos Futuros | 100 |
| 12 | Bibliografia | 103 |

Lista de figuras

| | |
|---|----|
| Figura 1 – Escalabilidade x Elasticidade | 21 |
| Figura 2 – Cenários típicos de ambientes de computação na nuvem | 23 |
| Figura 3 – Ulam (primeiro) e o Cassino de Monte Carlo em Mônaco | 25 |
| Figura 4 – Ideia genérica do Método de Monte Carlo | 27 |
| Figura 5 – Identidade visual do McCloud Service Framework | 31 |
| Figura 6 – Ontologia para computação na nuvem | 31 |
| Figura 7 – Arquitetura conceitual da solução | 33 |
| Figura 8 – Estruturas de armazenamento para tabelas no Azure | 35 |
| Figura 9 – Estrutura padrão da Table do Azure | 36 |
| Figura 10 – Funcionamento da tecnologia de fila do Azure | 37 |
| Figura 11 – Fluxo de dados do modelo Map/Reduce [16] | 38 |
| Figura 12 – Simplificação do modelo de programação do McCloud [17] | 39 |
| Figura 13 – Fluxo da etapa split | 40 |
| Figura 14 – Fluxo da etapa process | 41 |
| Figura 15 – Fluxo de trabalho da etapa <i>merge</i> | 42 |
| Figura 16 – Interface dos Pontos de Extensão | 45 |
| Figura 17 – Protótipo dos métodos do McCloud | 47 |
| Figura 18 – Diagrama de Casos de Uso | 48 |
| Figura 19 – Diagrama de Pacotes e Visão no Microsoft Visual Studio | 49 |
| Figura 20 – Campos do arquivo de configuração dos papéis | 49 |
| Figura 21 – Ilustração da visão geral do arcabouço (etapa split) | 51 |
| Figura 22 – Ilustração da visão geral do arcabouço (etapa process) | 52 |
| Figura 23 – Ilustração da visão geral do arcabouço (etapa merge) | 52 |
| Figura 24 – Ilustração da visão geral do arcabouço (final) | 53 |
| Figura 25 – Ilustração do racional da aproximação da constante pi | 54 |
| Figura 26 – Ilustração de 3 momentos da geração de pontos | 56 |
| Figura 27 – Algoritmo simulação do pi em computação tradicional | 56 |
| Figura 28 – Implementação <i>execute</i> e <i>finish</i> para a simulação do pi | 57 |
| Figura 29 – Parâmetro codein para aproximação do pi | 58 |
| Figura 30 – Parâmetro codeout para aproximação do pi | 58 |
| Figura 31 – Consumo do método Run em PHP5 | 59 |
| Figura 32 – Consumo do método Check e Result na aproximação do pi | 60 |
| Figura 33 – Gráfico da influência de W na duração e custo | 68 |
| Figura 34 – Implementação genérica proposta do <i>optimazation</i> | 71 |
| Figura 35 – Sistema 0 do Estudo de Caso | 73 |
| Figura 36 – Sistema 1 do Estudo de Caso | 73 |
| Figura 37 – Ponto de extensão startup do Deslocamento da Barra | 77 |
| Figura 38 – Pontos de extensão do Deslocamento da Barra (parte 1) | 78 |
| Figura 39 – Pontos de extensão do Deslocamento da Barra (parte 2) | 79 |
| Figura 40 – Página de Início da Simulação do Estudo de Caso | 79 |
| Figura 41 – Página de Checagem da Simulação do Estudo de Caso | 80 |
| Figura 42 – Página de Download da Simulação do Estudo de Caso | 80 |
| Figura 43 – Aplicação Cliente Deslocamento da Barra | 81 |
| Figura 44 – FDP do sistema 1 do estudo de caso | 83 |
| Figura 45 – Sistema 2 (esquerda) e 3 (direita) do Estudo de Caso | 84 |
| Figura 46 – FDP do sistema 3 do estudo de caso | 85 |
| Figura 47 – Média do sistema 3 do estudo de caso | 86 |

Lista de tabelas

| | |
|---|----|
| Tabela 1 – Comparação da computação tradicional e na nuvem | 20 |
| Tabela 2 – Exemplos de ambientes de computação na nuvem [5] | 32 |
| Tabela 3 – Teste de performance da prova de conceito | 61 |
| Tabela 4 – Notação da modelagem | 64 |
| Tabela 5 – Custo plataforma Windows Azure | 67 |
| Tabela 6 – Comparativo de tempo e custo em diferentes cenários | 69 |
| Tabela 7 – Resultado da aplicação do modelo proposto | 70 |
| Tabela 8 – Teste de performance do estudo de caso (sistema 0 e 1) | 82 |
| Tabela 9 – Teste de performance do estudo de caso (sistema 2 e 3) | 84 |