

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
Programa de Graduação em Ciência da Computação

José Ferreria Reis Fonseca

THE ASSISTANCE DESKTOP GRID COMPUTING MIDDLEWARE
princípios, arquiteturas, artefatos e testes alfa

Belo Horizonte
2015

José Ferreria Reis Fonseca

THE ASSISTANCE DESKTOP GRID COMPUTING MIDDLEWARE
princípios, arquiteturas, artefatos e testes alfa

Monografia apresentada ao Programa de Graduação em Ciência da Computação da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Luís Fabrício Goés

Belo Horizonte
2015

“Há um ditado que ensina o gênio é uma grande paciência; sem pretender ser gênio, teimeei em ser um grande paciente. As invenções são, sobretudo, o resultado de um trabalho teimoso, em que não deve haver lugar para o esmorecimento.” - Santos Dumont

“No soy un libertador. Los libertadores no existen. Son los pueblos quienes se liberan a si mismos” - Che Guevara

“Separados, somos Fortes. Juntos, somos Imbatíveis” -Jigoro Kano

Ingenius Templum Adverso - Arquimedes de Siracusa

To the end of the world, and whatever comes next

LISTA DE SIGLAS

HCP – High-Performace Computing - Computação de Alto Desempenho

DGC – Desktop Grid Computing - Computação em Grade de Máquinas Domésticas

WCP – Wasted Computing Power - Capacidade Computacional Desperdiçada

AES – Artefato de Engenharia de Software

SIMD – Single Instriction, Multiple Data - Unica Sequência de Instruções para Multiplas

Fontes de Dados

GPGPU – General Purpose Graphic Processing Unit - Unidade de Processamento Gráfico de Propósito Geral

P2P – Peer 2(to) Peer - Ponto à Ponto

SUMÁRIO

1 INTRODUÇÃO	6
1.1 Motivação	6
1.2 Problema	6
1.3 Proposta	7
1.3.1 <i>Objetivos</i>	7
1.4 Justificativa dos Objetivos e sua Relevância Científica	8
Referências Bibliográficas	9

1 INTRODUÇÃO

A prática da simulação em ambientes virtuais se uniu à revisão literária e a experimentação como terceiro pilar do processo de desenvolvimento científico e industrial (CIRNE et al., 2006). Tais simulações são um exemplo das incontáveis aplicações que processam grandes quantidades de dados, mesmo fora da academia (SCHWIEGELSHOHN et al., 2010). É cada dia maior a insaciável demanda de todo tipo de organizações e indivíduos por mais computações, e mais rápidas (Computações de alto desempenho - HPC). Pode-se prover cada usuário com acesso a um dispositivo capaz de realizar tais computações, o que é tremendamente dispendioso, ou usar soluções como Mainframe, Cloud Computing e Grid Computing (e sua variante que utiliza máquinas domésticas, Desktop Grid Computing - DGC).

DGC permite aproveitar o potencial computacional do parque de dispositivos já pertencentes à uma organização, contudo, tal solução, até o momento da finalização deste artigo, é relativamente pouco utilizada. (CEDERSTRÖM, 2010; SCHWIEGELSHOHN et al., 2010; CIRNE et al., 2006)

1.1 Motivação

No contexto de demanda muito superior à oferta, é pouco produtivo que indivíduos desperdicem quantidades significativas de recursos. Contudo, o uso típico de uma máquina doméstica, mesmo na indústria, é caracterizado por longos períodos de relativamente baixa atividade entremeados por surtos de demanda de computações, para as quais os recursos de uma máquina típica são normalmente insuficientes.

Mesmo para uma única máquina típica, de padrão de uso similar ao descrito acima, há um desperdício de capacidade computacional (WCP) considerável num período de tempo da escala de algumas horas. Este trabalho foi inspirado pelo desejo de reaproveitar tal potencial computacional desperdiçado, otimizando o uso de recursos de uma organização ou indivíduo.

1.2 Problema

No atual ecossistema de projetos DGC, são raros os artefatos de engenharia de software (AESs) disponibilizados para o público (SCHWIEGELSHOHN et al., 2010). Este é um obstáculo à ser superado para facilitar a análise dos custos de implantação e desenvolvimento de soluções por DGC.

A arquitetura de todos os sistemas citados nos trabalhos correlatos analisados é baseada na submissão de códigos portáteis para execução em máquinas de parceiros (ANDERSON, 2003; CIRNE et al., 2006; CEDERSTRÖM, 2010; SCHWIEGELSHOHN et al., 2010; ZHAO; LIU; LI, 2011). Diversos sistemas utilizam virtualização do ambiente de execução para executar os códigos de parceiros, o que não só esconde as características individuais de cada máquina (e o potencial computacional das mesmas) como adiciona mais uma camada de programas que consomem recursos (THAIN; TANNENBAUM; LIVNY,

2005; ZHAO; LIU; LI, 2011; CIRNE et al., 2006). O paradigma de portabilidade, portanto, limita a eficiência das computações, pois impossibilita otimizações de código personalizadas para a arquitetura cada máquina, e também o uso de recursos aceleradores de execução atípicos como co-processadores SIMD e GPGPUs.

O tempo de implantação de um dos atuais sistemas de DGC pode ser considerável, como descrito por (CEDERSTRÖM, 2010), que teve um custo de duas horas de trabalho para instalar o sistema BOINC (ANDERSON, 2003) em apenas seis máquinas, numa rede local.

1.3 Proposta

Minha proposta neste trabalho consiste no desenvolvimento de um *middleware* de DGC, denominado *Assistance*, registrando os principais AESs produzidos durante o processo. Um *middleware* é um sistema distribuído que intermedeia as comunicações e gerenciamento de execuções, conectando uma máquina à uma DGC. *Assistance* é baseado numa arquitetura de passagem de mensagens e implementações locais, em oposição ao modelo de virtualização e portabilidade de códigos executáveis. *Assistance* mantém compatibilidade com *middlewares* de DGC existentes e permite um compartilhamento e uso bastante completo dos recursos de cada máquina.

Assistance é estruturado de forma que cada máquina mantenha relações diretas com seus parceiros de compartilhamento de computações, uma topologia do modelo P2P - *Peer to Peer*. Esta estrutura facilita, e muito, a implantação do sistema, projetada para ser instantânea, como copiar poucos megabytes de arquivos.

1.3.1 Objetivos

1. Produzir todos os AESs necessários para descrever a arquitetura e princípios lógicos do sistema *Assistance*;
2. Implementar *Assistance* como um sistema de passagem de mensagens P2P para realizar DGC;
3. Medir o tempo de implantação da plataforma *Assistance* em diferentes ambientes de rede;
4. Medir a performance de diversas aplicações na plataforma *Assistance* em diversos ambientes de execução;
5. Avaliar a dificuldade de implementação de uma aplicação compatível com a plataforma *Assistance*;
6. Medir a dificuldade de adaptar uma aplicação já existente de forma a se tornar compatível com a plataforma *Assistance*;

1.4 Justificativa dos Objetivos e sua Relevância Científica

O objetivo primário deste trabalho se deve à nota deixada por (SCHWIEGELSHOHN et al., 2010), que ressaltam a necessidade de mais AESs para que soluções de DGC possam ser melhor estudadas. (THAIN; TANNENBAUM; LIVNY, 2005) reforça tal noção, afirmando que o núcleo do sistema Condor não foi alterado desde 1988, responsabilizando as boas práticas de engenharia de software envolvidas na elaboração do sistema.

O segundo, quarto, quinto e sexto objetivos são voltados para a validação científica e viabilidade funcional do paradigma de passagem de mensagens sem distribuição de códigos executáveis, uma vez que tal arquitetura não foi utilizada em quaisquer trabalhos correlatos analisados.

O terceiro objetivo foi inspirado pela informação mencionada por (CEDERSTRÖM, 2010), sobre o custo de implantação de um sistema de DGC simples. No caso, o custo, em tempo, foi considerado inadequado para outros contextos de uso, como dispositivos móveis e redes "*in-promptu*", redes organizadas circunstancialmente.

Referências Bibliográficas

- ANDERSON, David P. Public computing: Reconnecting people to science. In: **Conference on Shared Knowledge and the Web**. [S.l.: s.n.], 2003. p. 17–19.
- CEDERSTRÖM, Andreas. **On using Desktop Grid Computing in software industry**. 2010. Dissertação (Mestrado) — School of Engineering, Blekinge Institute of Technology, Ronneby, Suécia.
- CIRNE, Walfredo et al. Labs of the world, unite!!! **Journal of Grid Computing**, Springer, v. 4, n. 3, p. 225–246, 2006.
- SCHWIEGELSHOHN, Uwe et al. Perspectives on grid computing. **Future Generation Computer Systems**, Elsevier, v. 26, n. 8, p. 1104–1115, 2010.
- THAIN, Douglas; TANNENBAUM, Todd; LIVNY, Miron. Distributed computing in practice: The condor experience. **Concurrency and Computation: Practice and Experience**, Wiley Online Library, v. 17, n. 2-4, p. 323–356, 2005.
- ZHAO, Han; LIU, Xinxin; LI, Xiaolin. A taxonomy of peer-to-peer desktop grid paradigms. **Cluster Computing**, Springer, v. 14, n. 2, p. 129–144, 2011.