

José Ferreria Reis Fonseca

**THE ASSISTANCE MESSAGE-PASSING DESKTOP COMPUTER GRID**  
a novel architecture applied to artificial intelligence problems

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

## LISTA DE SIGLAS

- HCP – High-Performace Computing - Computação de Alto Desempenho
- SLR – Systematic Literature Review - Revisão Sistemática de Literatura
- DCG – Desktop Computer Grid - Grade de Computadores Domésticos
- MP-DCG – Message Passing Desktop Computer Grid - Grade de Computadores Domésticos que opera por Passagem de Mensagens
- SEA – Software Engineering Artifact - 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
- API – Application Interface - Interface de Aplicação

## SUMÁRIO

1 INTRODUÇÃO . . . . .	4
1.1 Motivação . . . . .	4
1.2 Problema . . . . .	5
1.3 Proposta . . . . .	5
1.3.1 <i>Objetivos</i> . . . . .	6
1.4 Justificativa . . . . .	6
Referências Bibliográficas . . . . .	7

# 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 que precisam de resultados o mais rápido possível (à esta categoria de aplicações se dá o nome de 'computações de alto desempenho' - HPC) seja na academia ou na indústria. Pode-se prover cada usuário com acesso a um dispositivo capaz de realizar tais computações, o que é dispendioso, usar soluções como computadores de grande porte (*Mainframes*), computação em *cluster*, computação em nuvem e computação em grade.

Uma solução que envolve computadores de grande porte é focada no compartilhamento de uma única grande máquina entre diversos usuários, o que é eficiente e dispendioso. Computação em nuvem é um modelo muito conhecido pelo público geral, englobando plataformas como *Google Cloud* e *Amazon Web Services*. Neste modelo, softwares inteiros são executados em ambiente remoto pertencente à uma certa organização, que submete ao usuário apenas suas vias de entrada de dados e saída de resultados, ~~não sendo necessário que este realize alguma contribuição computacional significativa.~~ Computação em grade é um modelo parecido, pedidos de computações são enviados para serem executados em máquinas parceiras. Contudo, máquinas distintas podem pertencer à organizações distintas, e cada membro da rede pode disponibilizar seus dispositivos para computações de outrem, aumentando a capacidade do sistema como um todo. Máquinas numa grade podem operar independentemente sem um controle central de processamento, necessário no modelo de *cluster*.

Uma grade pode ser implementada utilizando apenas computadores domésticos, como os encontrados em residências e ambientes de trabalho (este modelo é denominado DCG). Assim, uma organização pode compartilhar e otimizar o aproveitamento do potencial computacional do parque de dispositivos que já possui. (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 (CITATION NEEDED). A capacidade computacional não utilizada de uma máquina durante seu período de baixa atividade é considerada desperdiçada. Este trabalho foi inspirado pelo desejo de aproveitar tal potencial, otimizando o uso de recursos de uma organização ou indivíduo.

## 1.2 Problema



No atual ecossistema de projetos DCG, poucos artefatos de engenharia de software (SEAs) são 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 DCG.



Após realizar uma revisão sistemática de literatura (SLR), verificou-se que as arquiteturas de todos os sistemas de trabalhos correlatos analisados são baseadas na submissão de códigos portáteis para execução em máquinas de parceiros. 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.

## 1.3 Proposta



Neste trabalho são apresentados os artefatos de engenharia de software e realizada a implementação de um sistema denominado *Assistance*, um sistema conectado a aplicações de um usuário por meio de uma interface de aplicação (API), permitindo que se envie problemas lógicos para serem solucionados por parceiros. *Assistance* difere de uma grade de computadores típica por ser baseado numa arquitetura de passagem de mensagens e implementações locais, em oposição ao modelo de virtualização e portabilidade de códigos executáveis. Cada máquina da grade possui sua própria implementação de uma seleção de algoritmos e aplicações genéricos e comuns, executados sobre dados enviados por máquinas parceiras. Esta arquitetura permite mais flexibilidade no uso dos diferentes recursos de cada máquina, e mais segurança para seus proprietários, uma vez que não são executados códigos de terceiros. Este modelo é denominado *computação em grade de máquinas cotidianas por passagem de mensagens*, MP-DCG.

As aplicações que teriam melhor proveito de *Assistance* seriam as que realizam grandes processamentos (resolvendo problemas conhecidos) sobre quantidades relativamente pequenas de dados. Tais aplicações são denominadas *computacionalmente intensas* (ZHAO; LIU; LI, 2011). A performance tais aplicações seria beneficiada pela abundância de poder computacional na grade, e a pequena quantidade de dados à serem comunicados evita que esta seja comprometida pela latência da rede. Diversos algoritmos utilizados em aplicações de inteligência artificial podem ser considerados computacionalmente intensos, e serão implementados no sistema *Assistance*.

*Assistance* será submetido a testes de *benchmark* genéricos (SNAVELY et al., 2003), para que possa ser comparado com projetos semelhantes, e também testes de *benchmark* específicos de inteligência artificial, para medir sua performance e viabilidade na resolução de problemas desta área.

### 1.3.1 Objetivos



1. Implementar *Assistance* como um sistema de computação em grade (de máquinas cotidianas) por passagem de mensagens (MP-DCG);
  - (a) Produzir os SEAs envolvidos no desenvolvimento;
  - (b) Implementar diversos algoritmos envolvidos na solução de problemas de inteligência artificial utilizando o sistema;
2. Realizar testes de *benchmark* genéricos e da área de inteligência artificial na plataforma *Assistance*;
- ~~3. Avaliar a viabilidade do sistema para os contextos acadêmico, empresarial e pessoal;~~

### 1.4 Justificativa



Não foi encontrada, na busca por trabalhos correlatos, menção a qualquer sistema de computação em grade que utilizasse um paradigma de passagem de mensagens. Todos os projetos encontrados propõem que cada usuário distribua seus próprios códigos portáteis para execução em máquinas de parceiros, o que complica o uso de sistemas de computação atípicos, como GPGPUs e multiprocessadores. Os objetivos deste trabalho foram estabelecidos buscando esclarecer tanto quanto possível o paradigma de computação em grade por meio de passagem de mensagens, e sua viabilidade em aplicações computacionalmente intensivas, como as de inteligência artificial. A análise do modelo de MP-DCG será facilitada pela apresentação de SEAs e resultados de testes de *benchmark* comparados com os de outros projetos semelhantes.

## Referências Bibliográficas

- 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.
- SNAVELY, Allan et al. Benchmarks for grid computing: a review of ongoing efforts and future directions. **ACM SIGMETRICS Performance Evaluation Review**, ACM, v. 30, n. 4, p. 27–32, 2003.
- 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.