

# Paralelização de laços aninhados com processamento heterogêneo em sistemas paralelos e distribuídos.

Cleber Silva Ferreira Luz<sup>1</sup> e Liria Matsumoto Sato<sup>1</sup>

<sup>1</sup>Departamento de Engenharia de Computação e Sistemas Digitais  
Escola Politécnica da Universidade de São Paulo (EPUSP)  
São Paulo – SP – Brasil

cleber.luz@usp.br liria.sato@poli.usp.br

**Resumo.** *Computação paralela e distribuída vem se consolidando como um importante elemento no processamento de aplicações que necessitam de alto poder computacional. Diversas aplicações utilizam laços aninhados com computação heterogênea para processar seus dados. Geralmente aplicação com este perfil consome horas ou até mesmo dias de processamento. Este artigo apresenta uma solução para a paralelização de laços aninhados com computação heterogênea em sistemas paralelos e distribuídos. A proposta visa a diminuição do tempo de processamento das aplicações que contêm laços aninhados com processamento heterogêneos. A metodologia permite realizar uma otimização no uso dos recursos de processamento. Esta otimização é necessária para diminuir a ociosidade dos recursos, causada pela heterogeneidade de tempo de processamento das computações heterogêneas pertencentes aos laços aninhados.*

## 1. Introdução

Atualmente, computação paralela e distribuída tem sido essencial para se obter um maior desempenho no processamento de aplicações que demandam alto poder computacional. Em aplicações biológicas, por exemplo, geralmente é necessário analisar milhões de amostras experimentais, que muitas vezes, consomem horas ou até mesmo dias de processamento. Aplicações com esse perfil, muitas vezes utilizam laços de repetições aninhados no processamento de dados e muitas vezes, executam computações heterogêneas dentro dos mesmos.

Habitualmente, o processamento destas aplicações analisa e gera uma grande quantidade de dados que consomem horas ou até mesmo dias de processamento. Dessa forma, o processamento de aplicações com este perfil requer ferramentas capazes de manipular grandes quantidades de dados, como também, boas estratégias de processamento, a fim de realizar o processamento de forma rápida e eficiente. Uma solução para aplicações que apresentam laços aninhados é a sua linearização seguida da paralelização.

Diversas classes de algoritmos utilizados nas áreas de análise numérica, processamento de imagem e meteorológica contêm laços aninhados. Quando o algoritmo possui laços aninhados com computações heterogêneas a paralelização de um dos laços poderia implicar em uma subutilização dos recursos de processamento, deixando-os ociosos.

Neste artigo é apresentada uma solução para paralelização de laços aninhados com computações heterogêneas em sistemas paralelos e distribuídos aplicando a técnica de

linearização. A solução paralela também realiza uma otimização no uso dos recursos de processamentos utilizados no processamento das iterações dos laços aninhados, buscando a reduzir sua ociosidade, através de estratégias que promovem o balanceamento de cargas. Este artigo considera como estudo de caso a modelagem de distribuição de espécies. A solução paralela proposta foi aplicada no processo de modelagem de distribuição de espécies. Para este estudo de caso foram alcançados bons resultados que são apresentados na Seção 4.

A solução proposta pode ser implementada em sistemas de processamento paralelo e distribuídos, tais como, computadores multicore, ou sistemas distribuídos como um *cluster* físicos ou virtuais alocados na nuvem.

## **2. Paralelização de laços aninhados com processamento heterogêneo em sistemas paralelos e distribuídos.**

A solução paralela proposta descreve uma série de procedimentos para paralelizar laços aninhados com computações heterogêneas em ambientes paralelos e distribuídos. A solução proposta também realiza uma otimização no uso dos recursos de processamento, necessária para diminuir a ociosidade dos recursos de processamento, causada pela presença de computações heterogêneas dentro dos laços aninhados.

A otimização é realizada através da definição do número máximo de recursos de processamento, o qual não deverá ultrapassar a quantidade de recursos onde não se teria uma melhora de desempenho no processamento das iterações dos laços. A otimização também realizada realiza um escalonamento entre as computações das iterações do laço linearizado entre os recursos de processamento. Tal escalonamento é realizado visando uma redução de ociosidade entre os recursos de processamento.

Em um paralelismo onde o escalonamento é realizado de forma arbitrária, a distribuição de carga de trabalho não é realizada de forma homogênea. Alguns recursos de processamento podem receber uma carga de trabalho maior, enquanto que outros recursos podem receber uma carga de trabalho menor. Essa desigualdade de atribuição de carga de trabalho pode causar ociosidade entre os processados. A otimização realizada pela solução paralela realiza um escalonamento, onde, não há uma desigualdade de carga de trabalho entre os recursos de processamento.

A solução paralela proposta lineariza os laços aninhados e descreve uma série de procedimentos para tornar o laço linearizado em um laço paralelo. Na solução paralela é considerado que não há dependência de dados entre as computações. O Algoritmo 2 descreve um pseudocódigo da estratégia da solução paralela proposta.

---

### **Algorithm 1** Estratégia da solução paralela.

---

- 1: linearização dos loops aninhados.
  - 2: preparação das iterações do loop linearizado para o processamento.
  - 3: definição do número de recursos de processamento para processar as iterações do laço linearizado.
  - 4: escalonamento das iterações do loop linearizado.
  - 5: processamento paralelo, em que cada recurso processa as iterações que lhe foram atribuídas.
-

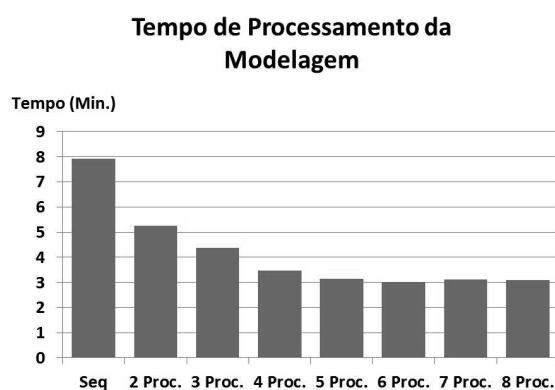
A solução paralela é composta por: Coleta de tempo, Cálculo do Número de Recursos de Processamento, Escalonamento e Processamento. O mecanismo de Coleta de Tempo realiza a obtenção dos tempos das computações heterogêneas presentes dentro dos laços aninhados. O mecanismo de Cálculo do Número de Recursos de Processamento propõe o número de recursos de processamento que melhora o desempenho da aplicação. Já o mecanismo de escalonamento permite atribuir as iterações do laço linearizado para os recursos de processamento. No mecanismo de escalonamento, a tarefa de maior carga é atribuída para o recurso que contém a menor carga de trabalho. O último mecanismo realiza o processamento paralelo das iterações do laço linearizado.

### 3. Estudo de Caso: Modelagem de Distribuição de espécies

A modelagem de distribuição de espécies é uma ferramenta numérica que combina observações de ocorrências de espécies com variáveis ambientais. Esta combinação permite realizar uma predição de ocorrência da espécie em determinadas áreas geográficas [3]. A modelagem é dada através de representações matemáticas, fornecidas pelo processamento de diversos algoritmos de modelagem de distribuição de espécie. Para se obter uma boa taxa de acerto de predição, o processo de modelagem executa diversos algoritmos de modelagem que consomem tempos de processamento diferentes. O processamento destes algoritmos é realizado através do processamento de laços aninhados.

### 4. Resultados

Esta seção, apresenta os resultados da implementação da solução paralela na modelagem de distribuição de espécies. O gráfico apresentado na Figura 1 apresenta o tempo de processamento paralelo e sequencial da fase de modelagem.



**Figure 1. Tempo de Processamento da Modelagem**

O tempo de processamento sequencial da modelagem foi de 8 minutos, ao passo que, o tempo de processamento paralelo foi de 3 minutos utilizando 6 recursos de processamento. A solução paralela definiu que 6 recursos é a quantidade ótima de recurso, mesmo adicionando mais recurso, não é possível obter um tempo menor que 3 minutos, como é o caso do processamento utilizando 7 e 8 recursos. Neste experimento, o algoritmo de modelagem GAM consumiu mais tempo de processamento, consumindo

aproximadamente 2 minutos e 40 segundos de processamento, sendo este o tempo aproximado do menor tempo de processamento possível. Dessa forma, é possível concluir que o tempo de processamento mínimo foi alcançado com a menor quantidade de recursos possíveis.

## 5. Trabalhos Relacionados

Em [4] é apresentado um algoritmo de escalonamento dinâmico para laços aninhados com dependências uniformes em redes com computadores heterogêneos. O trabalho apresentado em [4] vai de encontro com a nossa proposta. Em [4] há um escalonamento de laços aninhados para ambientes heterogêneos. Na metodologia proposta neste artigo, há uma preocupação em linearizar os laços aninhados e processar suas iterações em paralelo. A solução paralela proposta também realiza um escalonamento adequado das computações heterogêneas presentes dentro dos laços aninhados afim de, manter uma boa utilização dos recursos de processamento.

Em [1] é apresentada uma ferramenta para a paralelização de laços aninhados em multicomputadores com memória distribuída. A ferramenta apresentada em [1] utiliza a técnica *grouping* para particionar os laços aninhados com dependência de dados. O resultado desta técnica é um pipeline que fornece um balanceamento na computação dos laços aninhados. A ferramenta apresentada em [1] realiza uma paralelização de laços aninhados em multicomputadores com memória distribuída. A solução paralela proposta se preocupa em escalonar os laços aninhados levando em consideração as computações heterogêneas realizadas dentro dos laços aninhados.

## 6. Conclusões

Este artigo apresenta uma solução para paralelização de laços aninhados com processamento heterogêneo em sistemas paralelos e distribuídos. A solução proposta neste artigo, também realiza uma otimização no uso dos recursos de processamento. A otimização é necessária para diminuir a ociosidade entre os recursos de processamento, causada pela heterogeneidade dos tempos de execuções das computações heterógenas presentes dentro dos laços aninhados. A solução proposta foi implementada na aplicação de modelagem de distribuição de espécies, onde, foram alcançados bons resultados e que são apresentados neste artigo.

## References

- [1] Chung-Ta King and Ing-Ren Kau. - Parallelizing nested loops on multicomputers-the grouping approach. *Computer Software and Applications Conference, 1989. COMP-SAC 89., Proceedings of the 13th Annual International.*
- [2] Gonzalo Vera and Remo Suppi - Integration of heterogeneous and non-dedicated environments for R. 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing
- [3] Jennifer Miller - Species Distribution Modeling. *Volume 4, Issue 6, pages 490-509, June 2010*
- [4] R., Ionannis and T., Panagiotis - Dynamic Scheduling of Nested Loops with Uniform Dependencies in Heterogeneous Networks of Workstations. *8th International Symposium on Parallel Architectures, Algorithms and Networks (ISPAN'05), 2005.*