

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE INFORMÁTICA  
CURSO DE CIÊNCIA DA COMPUTAÇÃO

LUCAS BARROS DE ASSIS

**Modelagem de ondas sísmicas através de  
paralelismo de tarefas**

Monografia apresentada como requisito parcial  
para a obtenção do grau de Bacharel em Ciência  
da Computação

Orientador: Prof. Dr. Lucas Mello Schnorr

Porto Alegre  
2021

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Prof<sup>a</sup>. Patricia Pranke

Pró-Reitora de Graduação: Prof<sup>a</sup>. Cíntia Inês Boll

Diretora do Instituto de Informática: Prof<sup>a</sup>. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Rodrigo Machado

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

## RESUMO

A aplicação *Ondes3D* tem como objetivo realizar a simulação da propagação de uma onda sísmica. Apesar de contar com uma implementação paralela utilizando *OpenMP*, esse paralelismo só é obtido dentro de cada um de seus *macro-kernels*. O trabalho aqui apresentado estuda as alterações necessárias para dividir as estruturas utilizadas pelo simulador em ladrilhos, possibilitando posteriormente uma implementação em forma de tarefas utilizando a biblioteca *StarPU*. Essa divisão em tarefas permite um controle da granularidade dos processos paralelos através do tamanho dos ladrilhos utilizados, o que leva à uma otimização dos acessos em memória. Além disso, é possível que diferentes etapas do processo sejam executadas simultaneamente graças às especificações das dependências que o modelo de programação em tarefas inclui. Dessa forma, espera-se alcançar uma execução que aproveita arquiteturas *multi-core* de forma mais vantajosa que a versão original.

**Palavras-chave:** Programação paralela. programação baseada em tarefas. Ondes3D. StarPU.

## Seismic waves modelling through task-based programming

### ABSTRACT

The *Ondes3D* simulator aims to simulate the propagation of a seismic wave. Even though it has a parallel implementation using *OpenMP*, this parallelism is only achieved inside each of its macro-kernels. The study presented in this document studies the modifications needed to split the structures used by *Ondes3D* in tiles, which allows a task-based implementation using the *StarPU* library. By splitting the code in tasks, it becomes possible to control the granularity of the parallel processes through the tile sizes, which enables a memory access optimization. Other than that, different steps of the computation can be executed simultaneously thanks to the dependency specifications from the task-based model. With these modifications, it may be possible to achieve an execution which exploits multi-core architectures even better than the original version.

**Keywords:** Parallel programming. Task-based programming. Ondes3D. StarPU.

## LISTA DE FIGURAS

Figura 3.1	Representação de um <i>stencil</i> 2D com quatro vizinhos .....	13
Figura 3.2	Definição do <i>codelet</i> da função <i>computeIntermediates</i> .....	16
Figura 3.3	Decomposição da matriz A em blocos .....	17
Figura 3.4	Algoritmo de multiplicação de matrizes por blocos .....	17
Figura 4.1	Representação do espaço particionado em nove blocos .....	18
Figura 4.2	Laço principal da aplicação original.....	19
Figura 4.3	Laço principal da proposta .....	19
Figura 4.4	Representação do uso de células fantasma .....	20
Figura 4.5	Representação de um vetor que permite índices negativos .....	21
Figura 4.6	Representação de uma matriz de velocidades e um vetor de contribui- ções das bordas .....	23
Figura 4.7	À esquerda, parte da implementação da estrutura <i>ABSORBING_BOUNDARY_CONDITION</i> e à direita, parte da nova implementação proposta. ....	23
Figura 4.8	<i>Macro</i> desenvolvida para construir a nova estrutura proposta.....	24
Figura 4.9	Ciclo de vida de um <i>data handle</i> . ....	25
Figura 4.10	Parte do cálculo de $v0- > z[i][j][k]$ .....	27
Figura 4.11	<i>Macros</i> utilizadas para acessar vetores. ....	28
Figura 4.12	Resumo das chamadas de computação de velocidade. ....	29

## **LISTA DE TABELAS**

Tabela 2.1	Trabalhos relacionados .....	11
------------	------------------------------	----

## **LISTA DE ABREVIATURAS E SIGLAS**

CPU	Central Process Unit
DAG	Directed Acyclic Graph
GPU	Graphics Processing Unit
HPC	High-Performance Computing
MPI	Message Passing Interface

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>9</b>
<b>1.1 Contribuições.....</b>	<b>9</b>
<b>2 TRABALHOS RELACIONADOS .....</b>	<b>10</b>
<b>3 FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>12</b>
<b>3.1 Ondes3D.....</b>	<b>12</b>
3.1.1 Condições de borda.....	14
<b>3.2 Paralelismo baseado em tarefas.....</b>	<b>15</b>
3.2.1 StarPU .....	15
3.2.2 Matrizes <i>ladrihadas</i> .....	16
<b>4 PROJETO E IMPLEMENTAÇÃO .....</b>	<b>18</b>
<b>4.1 Proposta .....</b>	<b>18</b>
<b>4.2 Implementação .....</b>	<b>20</b>
4.2.1 Inicialização .....	21
4.2.1.1 Inicialização das <i>CPML</i> .....	22
4.2.2 Criação de <i>data handles</i> .....	24
4.2.2.1 Acesso aos vizinhos .....	25
4.2.3 Transformação em tarefas .....	27
4.2.3.1 Cálculo de velocidades .....	28
<b>5 AVALIAÇÃO EXPERIMENTAL .....</b>	<b>30</b>
<b>5.1 Ambiente de testes.....</b>	<b>30</b>
<b>5.2 Resultados.....</b>	<b>30</b>
<b>5.3 Trabalhos futuros.....</b>	<b>30</b>
<b>6 CONCLUSÃO .....</b>	<b>32</b>
<b>REFERÊNCIAS.....</b>	<b>33</b>



## 1 INTRODUÇÃO

Uma ferramenta importante para a mitigação dos riscos decorrentes de terremotos é a simulação da propagação de ondas sísmicas (DUPROS et al., 2010). A aplicação *Ondes3D* realiza essa simulação através do método de diferenças finitas, utilizando a biblioteca *OpenMP* para produzir paralelismo local e o protocolo *MPI* em contextos distribuídos. No entanto, o paralelismo local somente existe dentro de cada *macro-kernel*, isto é, não existe execução paralela entre diferentes etapas do cálculo. Uma alternativa possível para acelerar a execução pode ser alcançada utilizando um modelo que permita um paralelismo ainda maior que o atual.

A biblioteca *StarPU* é uma alternativa atual que implementa a programação baseada em tarefas para obter paralelismo. Dentro desse modelo, uma tarefa consiste em uma função cuja especificação inclui, além de seus parâmetros, os seus modos de acesso: somente leitura, somente escrita ou leitura e escrita. Essas tarefas são incluídas ao longo do código e, baseando-se na ordem de inserção e nos modos de acesso de cada uma delas, a biblioteca *StarPU* constrói um grafo acíclico direcionado representando as dependências entre essas tarefas. Durante a execução do programa, esse grafo é utilizado para escalonar as tarefas conforme a disponibilidade dos recursos computacionais. Apesar de contar com uma implementação paralela, o modelo de programação utilizado na versão original da aplicação não explora o paralelismo tanto quanto seria possível devido à sua incapacidade de considerar as dependências entre os dados utilizados. Ao utilizar o paradigma de tarefas, torna-se possível considerar essas dependências para que diferentes etapas do algoritmo possam ser executadas simultaneamente, alcançando assim um paralelismo superior ao atualmente implementado.

Este trabalho tem como objetivo estudar e implementar as alterações necessárias para que o algoritmo existente seja executado na forma de tarefas, assim como avaliar o paralelismo obtido ao utilizar esse paradigma.

### 1.1 Contribuições

## 2 TRABALHOS RELACIONADOS

A aplicação *Ondes3D* já foi objeto de estudo em outros trabalhos, frequentemente visando um melhor desempenho, mas passando também pela busca de um consumo energético mais eficiente. Além disso, outras aplicações também tiveram seu tempo de execução diminuído empregando a programação baseada em tarefas. Alguns desses trabalhos serão discutidos abaixo.

Boito et al. (BOITO et al., 2017) verificaram que mais de 72% do tempo de execução do *Ondes3D* é gasto realizando operações de leitura e escrita. Três otimizações diferentes foram propostas: inicialmente, uma camada de *software* foi criada para centralizar as solicitações de entrada e saída da aplicação. Em um segundo momento, os passos de comunicação foram substituídos por escritas diretas no sistema de arquivos. A terceira implementação passa a utilizar diversos arquivos de saída, o que permite que diversas operações de escrita aconteçam paralelamente. Apesar dos resultados impressionantes desse trabalho, o estudo aqui proposto não tem como foco as operações de entrada e saída mas sim a paralelização dos cálculos das equações elastodinâmicas, o que permitiria em um momento futuro a combinação das duas otimizações para alcançar uma aceleração ainda mais significativa.

Tesser et al. (DUPROS; DO; AOCHI, 2013), partindo de uma implementação distribuída em que o domínio é dividido em partes menores, perceberam que o tempo de cálculo das equações depende da região sendo tratada, causando um desbalanceamento de cargas entre os processadores utilizados. Ao utilizar uma implementação de *MPI* construída sobre um suporte de execução capaz de realizar um balanceamento dinâmico de tarefas, foi possível reduzir o tempo de execução em 23%. O presente estudo não considera a utilização de programação distribuída, mas a granularidade parametrizável alcançada utilizando o modelo proposto permite uma mitigação dos efeitos de uma distribuição desbalanceada. Quanto maior o número de tarefas disponível em um dado momento, menores são as chances de uma unidade de cálculo permanecer ociosa.

Castro et al. (CASTRO et al., 2016) utilizaram um processador de baixa potência para propor uma implementação da aplicação *Ondes3D* com um baixo consumo energético. A partir de uma estratégia de ladrilhamento multi-nível, a versão desenvolvida para o processador *MPPA-256* apresentou uma diminuição de 86% do consumo quando comparada à uma execução em um processador convencional. No entanto, o consumo energético não é um fator considerado no estudo aqui apresentado.

Nesi et al. (NESI et al., 2020) realizaram três implementações diferentes de uma aplicação de dinâmica de fluídos utilizando programação baseada em tarefas através da biblioteca *StarPU*. Além de incluir tarefas voltadas para *GPUs*, o traço de execução resultante foi avaliado e refinado utilizando o conjunto de ferramentas *StarVZ*. Utilizando a técnica de células fantasma, explicada na seção 4.1, foi possível alcançar uma aceleração de 77x em relação à versão original.

Martínez et al. (MARTÍNEZ et al., 2015) propuseram uma implementação utilizando a biblioteca *StarPU*, baseada em tarefas e considerando arquiteturas heterogêneas compostas por *CPUs* e *GPUs*. De um certo modo essa proposta vai além da pesquisa aqui apresentada, onde não existe uma implementação voltada para *GPU*. No entanto, Martínez et al. optaram por descartar as interações com as bordas do domínio devido a complexidade introduzida pelo uso delas. A melhor performance foi alcançada utilizando quatro *CPUs* e oito *GPUs*, uma aceleração de 25x em relação à versão original utilizando doze *CPUs*.

A Tabela 2.1 resume as diferenças entre os trabalhos desenvolvidos e os estudos citados acima. Os projetos desenvolvidos por Nesi et al. e Martínez et al. são os que mais se assemelham ao aqui desenvolvido, apesar do primeiro tratar de uma aplicação diferente e o segundo ignorar uma parcela da simulação.

Tabela 2.1 – Trabalhos relacionados

<i>Trabalho</i>	<i>Aplicação</i>	<i>Contexto</i>
Boito et al.	<i>Ondes3D</i>	Otimização de operações de entrada e saída
Tesser et al.	<i>Ondes3D</i>	Balanceamento dinâmico de carga
Martínez et al.	<i>Ondes3D</i>	Implementação baseada em tarefas em arquiteturas heterogêneas sem condições de borda
Castro et al.	<i>Ondes3D</i>	Otimização de consumo energético
Nesi et al.	Dinâmica de fluídos	Implementação baseada em tarefas em arquiteturas heterogêneas
Este trabalho	<i>Ondes3D</i>	Implementação baseada em tarefas

Fonte: Autor

### 3 FUNDAMENTAÇÃO TEÓRICA

Para a elaboração desse estudo, primeiramente é necessário compreender o que é a aplicação utilizada. A partir desse conhecimento, é possível tratar do modelo de programação baseada em tarefas e como ela é implementada na biblioteca *StarPU*. Finalmente, é necessário explicar o método de divisão das estruturas de dados usadas, conhecido como ladrilhamento ou *tiling*.

#### 3.1 Ondes3D

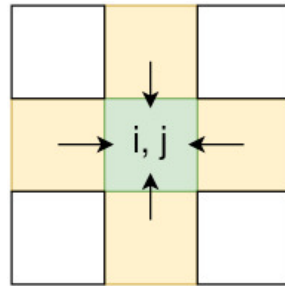
A aplicação discretiza o mundo real em diversos tensores tri-dimensionais, endereçados a partir de suas coordenadas. Esses tensores são parte de diferentes estruturas de dados, que podem variar de acordo com os dados experimentais utilizados na simulação.

Esse trabalho toma como objeto de estudo os dados do terremoto de *Chuetsu-Oki*, que atingiu a costa japonesa em 2007. Nesse caso, a aplicação conta com as seguintes estruturas, usadas para representar os fenômenos físicos que compõem a onda:

- *STRESS*: composta por tensores que representam as tensões sofridas em diferentes planos. O tensor  $xy$ , por exemplo, representa a tensão sofrida no plano X - aquele cuja normal aponta na direção do eixo X, isto é, o plano YZ - pelas forças causadas na direção Y.
- *VELOCITY*: composta por tensores que representam a velocidade da onda nas direções X, Y e Z.
- *SOURCE*: composta pelos dados que descrevem a(s) fonte(s) da onda, como seu hipocentro, a longevidade da fonte, sua orientação e força.
- *MEDIUM*: composta pelos dados que descrevem o meio por onde a onda se propaga, como sua profundidade.
- *ABSORBING\_BOUNDARY\_CONDITION*: composta por vetores que representam as contribuições das bordas do domínio nos cálculos de tensão e velocidade. As contribuições em cada uma dessas grandezas são compostas por nove componentes, representando as influências de cada eixo nos três eixos possíveis (no caso da velocidade) e nos três planos possíveis (no caso da tensão). Essas bordas são utilizadas para limitar o domínio computacional minimizando a interferência de limites artificiais nos cálculos realizados. (NATAF, 2013).

Os cálculos são realizados seguindo um padrão *stencil*, onde o valor atribuído aos tensores que descrevem as grandezas físicas são calculados usando informações das células vizinhas. A Figura 3.1 ilustra um *stencil* 2D com quatro vizinhos. Apesar da aplicação estudada usar coordenadas nos eixos X, Y e Z, o mesmo conceito pode ser expandido para o espaço tridimensional.

Figura 3.1 – Representação de um *stencil* 2D com quatro vizinhos



Fonte: Autor

As estruturas descritas anteriormente são atualizadas iterativamente, passando por quatro grandes etapas, também chamadas de *macro-kernels*:

- *computeSeisMoment*: etapa em que as fontes da onda são atualizadas
- *computeIntermediates*: etapa em que as contribuições das bordas do domínio são atualizadas a partir dos tensores de velocidade
- *computeStress*: etapa em que os tensores de tensão nos diferentes planos são atualizados a partir dos tensores de velocidade e das condições de borda
- *computeVelocity*: etapa em que os tensores de velocidade são atualizados a partir dos tensores de velocidade e das condições de borda

Com exceção da atualização das fontes da onda, que acontece em um domínio bastante restrito, os *macro-kernels* descritos acima utilizam a diretiva *omp parallel for* para obter um paralelismo interno. Já que essas etapas realizam cálculos dentro de um laço triplo, iterando sob as coordenadas em X, Y e Z, a biblioteca *OpenMP*, através da diretiva citada anteriormente, cria *threads* capazes de tratar diferentes coordenadas simultaneamente.

Teoricamente, a partir do momento que a tensão, por exemplo, de uma célula fosse calculada, já seria possível iniciar o cálculo de sua velocidade. No entanto, como a diretiva descrita acima não suporta o estabelecimento de dependências entre os dados, é necessário que a tensão de todas as células seja calculada para que o algoritmo possa

iniciar a fase de cálculo de velocidades. Dessa maneira, apesar da solução atual alcançar uma aceleração graças ao paralelismo introduzido, a quantidade de operações que podem ser executadas simultaneamente é menor do que seria possível. O paralelismo baseado em tarefas busca explorar as dependências entre os dados para alcançar esse nível maior de independência entre as etapas do algoritmo.

### 3.1.1 Condições de borda

Ao simular a propagação de uma onda, é necessário considerar os reflexos por ela criados. No entanto, o método de diferenças finitas, utilizado pelo *Ondes3D*, exige que as suas equações sejam resolvidas em um domínio discretizado restrito. Dessa maneira, ao modelar a propagação de ondas, é necessário utilizar técnicas que permitam absorver as ondas nos limites do domínio estudado de maneira que seus reflexos não interfiram nos cálculos internos.

Berenger (BERENGER, 1994) desenvolveu uma técnica voltada para absorção de ondas eletromagnéticas, chamada de *Perfectly Matched Layers (PML)*. Diferentemente das outras técnicas existentes até então, o emprego de *PMLs* permite a construção de uma camada de absorção que teoricamente não produz reflexo algum independentemente da frequência e do ângulo de incidência das ondas recebidas. Experimentos realizados na época mostraram que, apesar de uma pequena quantidade de reflexos ter sido detectada, sua magnitude poderia ser reduzida adequando alguns parâmetros. Além disso, o emprego da técnica de *PMLs* permite a utilização de camadas de absorção menos espessas, o que reduz a computação necessária ao simular a propagação de ondas.

No entanto, o coeficiente de reflexão das *PMLs* só é nulo antes da discretização do domínio e com ondas de incidência não rasante, isto é, com um ângulo de incidência diferente de  $90^\circ$ , o que torna essa técnica inadequada quando, por exemplo, existem fontes de ondas muito próximas dos limites do domínio. Roden e Gedney (RODEN; GEDNEY, 2000) propuseram um novo método chamado *Convolutional Perfect Matched Layers*, que utiliza uma convolução recursiva para atenuar a reflexão de ondas de incidência rasante.

A aplicação *Ondes3D* utiliza ambas as técnicas acima descritas em diferentes experimentos. No caso do experimento de *Chuetsu-Oki*, no qual esse trabalho se baseia, a técnica empregada é a das *CPMLs*.

### 3.2 Paralelismo baseado em tarefas

O paralelismo baseado em tarefas é um modelo em que se busca descrever trechos de código paralelizáveis na forma de tarefas, que são criadas em tempo de execução pela biblioteca de paralelismo utilizada. Depois de criar a tarefa, a biblioteca é responsável por executá-la ou adicioná-la a uma fila de tarefas para que ela seja executada em um momento apropriado. A possibilidade de adiar a execução de uma tarefa é o que torna essa opção interessante ao considerarmos as dependências da aplicação aqui estudada.

#### 3.2.1 StarPU

*StarPU* é uma biblioteca de programação em tarefas para arquiteturas heterogêneas. Essa biblioteca foi criada visando atender à necessidade da comunidade de *HPC* de poder computacional (AUGONNET et al., 2009), criando uma ferramenta que facilita a criação de tarefas para diferentes arquiteturas e a especificação dos dados utilizados. No caso do estudo aqui descrito, essa biblioteca é especialmente interessante ao determinar o momento de execução de suas tarefas, já que a escolha do momento de sua execução é baseada nas dependências de dados existentes entre elas.

A utilização de tarefas é composta por dois momentos, sendo o primeiro deles a sua descrição e o segundo a sua inserção na fila de tarefas. O paralelismo obtido é o resultado da inserção de diversas instâncias da mesma tarefa atuando sobre dados diferentes e tendo sua ordem de execução determinada pelas dependências entre elas.

A criação de uma tarefa começa pela construção de seu *kernel*, uma função que implementa a tarefa a ser executada e que deve seguir a seguinte interface:

```
void task_name(void *buffers[], void *cl_arg)
```

Nessa interface, os *buffers* representam os dados que serão gerenciados pela biblioteca levando em conta suas dependências e modos de leitura. Já a estrutura *cl\_arg* guarda ponteiros de valores externos utilizados pelo *kernel*, tipicamente valores constantes.

Com o *kernel* criado, utiliza-se uma estrutura chamada *codelet* para descrever a tarefa. Nessa estrutura é indicada a quantidade de *buffers* utilizados pela tarefa, seus modos de acesso e o nome dos *kernels* que implementam-na. Voltando ao exemplo da função *computeIntermediates*, um *codelet* possível é mostrado na Figura 3.2.

A criação de instâncias de tarefas é feita através da função *starpus\_insert\_task*,

Figura 3.2 – Definição do *codelet* da função *computeIntermediates*

```
enum starpu_data_access_mode modes_intermediates[31] =
{
    STARPU_RW,
    STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R,
    STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R,
    STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R,
    STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R,
    STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R, STARPU_R
};

struct starpu_codelet intermediates_cl = {
    .cpu_funcs = {compute_intermediates_task},
    .nbuffers = 31,
    .name = "intermediates",
    .dyn_modes = modes_intermediates
};
```

Fonte: Autor

tendo como parâmetros o *codelet* correspondente, os *buffers* utilizados e os ponteiros registrados como *cl\_args*.

A biblioteca segue o modelo de *sequential task flow*, que permite que um fluxo paralelo seja facilmente descrito através de um algoritmo sequencial. As tarefas são submetidas sequencialmente e, a partir de sua ordem de submissão e suas dependências, cria-se um grafo acíclico direcionado (também chamado de *Directed Acyclic Graph*, ou *DAG*) que descreve a hierarquia existente entre elas. A partir desse grafo, é possível escalonar as tarefas em tempo de execução de maneira que todas as dependências sejam respeitadas e o tempo ocioso nas unidades de cálculo seja o menor possível.

### 3.2.2 Matrizes *ladrilhadas*

Uma estratégia recorrente na programação de alta performance é o ladrilhamento de matrizes - no caso desse estudo, de tensores. Essa técnica consiste, conforme ilustra a Figura 3.3, na divisão de uma matriz em sub-matrizes (também chamadas de blocos) de menor tamanho, indexadas pela sua posição em relação à matriz original.

O ladrilhamento de matrizes exige que os algoritmos sejam re-adequados à sua estrutura, mas em contrapartida oferecem numerosas vantagens. Em primeiro lugar, podemos citar a otimização dos acessos em memória: ao utilizar um tamanho de bloco suficientemente pequeno, é possível que, durante a execução do cálculo sobre cada um dos blocos, todos os dados necessários estejam presentes simultaneamente na memória cache. Além disso, utilizando o DAG criado por ferramentas como a biblioteca *StarPU*, é possível executar simultaneamente tarefas de etapas diferentes, acelerando ainda mais a



Figura 3.3 – Decomposição da matriz A em blocos

$$A = \begin{bmatrix} A_{0,0} & A_{0,1} & A_{0,2} \\ A_{1,0} & A_{1,1} & A_{1,2} \\ A_{2,0} & A_{2,1} & A_{2,2} \end{bmatrix}$$

Fonte: Autor

execução da aplicação.

A Figura 3.4 ilustra o cálculo de multiplicação de matrizes utilizando matrizes ladrilhadas. A versão convencional do algoritmo consiste em acumular a soma dos produtos de cada linha por cada coluna das matrizes envolvidas. Apesar dos elementos consecutivos de uma mesma linha encontrarem-se tipicamente lado a lado em memória, os saltos realizados nas trocas de linha podem exigir uma reescrita completa da cache.

Figura 3.4 – Algoritmo de multiplicação de matrizes por blocos

$$C_{i,j} = C_{i,j} + A_{i,k} * B_{k,j}$$

Fonte: Autor

A versão ladrilhada é composta por diversas execuções do algoritmo convencional, atuando sobre uma linha (respectivamente, coluna) de blocos de cada vez. Dessa forma, além de otimizar os acessos de memória ao realizar pequenas multiplicações bloco a bloco, é possível executar simultaneamente o cálculo de todos os blocos da matriz resultante. Apesar dos blocos das matrizes  $A$  e  $B$  serem acessados por diversas tarefas, eles não sofrem nenhuma escrita. Os blocos da matriz  $C$ , onde as escritas são realizadas, são acessados cada um deles por uma única tarefa, possibilitando tal paralelismo.

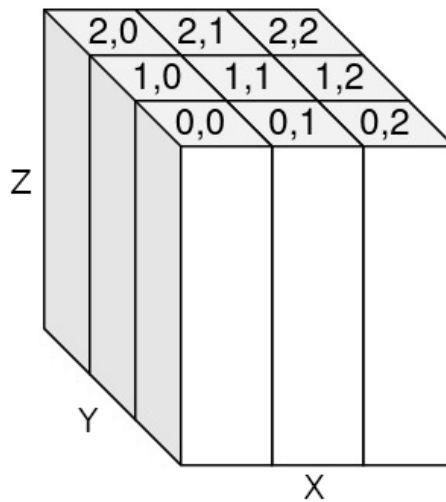
## 4 PROJETO E IMPLEMENTAÇÃO

Neste capítulo serão apresentadas as alterações necessárias para que a aplicação *Ondes3D* seja paralelizável na forma de tarefas gerenciadas pela biblioteca *StarPU*. Posteriormente, será apresentada uma descrição da implementação realizada ao longo dessa pesquisa.

### 4.1 Proposta

Os diversos tensores que representam as grandezas utilizadas nas equações elastodinâmicas do problema deverão ser divididos, de forma parametrizável, em blocos de tamanho igual, que serão os argumentos de entrada das tarefas implementadas. A Figura 4.1 ilustra o particionamento descrito utilizado pela implementação original, que parte de um arquivo de topologia para discretizar o plano  $XY$  em blocos de profundidade igual à do domínio.

Figura 4.1 – Representação do espaço particionado em nove blocos



Fonte: Autor

Em relação ao laço principal de execução, a Figura 4.2 resume o fluxo da aplicação original. O domínio do problema é dividido em cinco regiões denominadas *imodes*, sendo quatro delas as linhas e colunas externas do domínio e a quinta os elementos internos. Com exceção da etapa *computeSeisMoment*, cada *macro-kernel* é executado cinco vezes, uma em cada região. Internamente, cada etapa percorre o espaço  $[x_{min}, x_{max}][y_{min}, y_{max}][z_{min}, z_{max}]$ .

Figura 4.2 – Laço principal da aplicação original

```

for (iterations = 0; iterations < max_it; iterations++) {
    computeSeisMoment();

    for (imode = 0; imode < 5; imode++) {
        calculate_limits(imode, &x_min, &x_max, &y_min, &y_max, &z_min, &z_max);
        computeIntermediates(x_min, x_max, y_min, y_max, z_min, z_max);
    }
    for (imode = 0; imode < 5; imode++) {
        calculate_limits(imode, &x_min, &x_max, &y_min, &y_max, &z_min, &z_max);
        computeStress(x_min, x_max, y_min, y_max, z_min, z_max);
    }
    for (imode = 0; imode < 5; imode++) {
        calculate_limits(imode, &x_min, &x_max, &y_min, &y_max, &z_min, &z_max);
        computeVelocity(x_min, x_max, y_min, y_max, z_min, z_max);
    }
}

```

Fonte: Autor

Na versão aqui proposta, cada etapa será executada tantas vezes quanto houverem blocos. Para isso, será necessário iterar sobre os blocos disponíveis, criando uma tarefa para cada bloco de dados a ser atualizado, conforme mostra a Figura 4.3, dispensando a utilização dos *imodes*. Neste momento vale ressaltar que a inserção das tarefas consiste simplesmente na criação de uma instância de tarefa com as dependências indicadas e não implica em sua execução imediata. Fica a cargo da biblioteca *StarPU* iniciar a execução de cada uma das tarefas disponíveis de maneira assíncrona.

Figura 4.3 – Laço principal da proposta

```

for (iterations = 0; iterations < max_it; iterations++) {
    starpu_insert_task(computeSeisMoment);

    for (i_block = 0; i_block < n_blocks_y; i_block++) {
        for (j_block = 0; j_block < n_blocks_x; j_block++) {
            starpu_insert_task(computeIntermediates, blocks[i][j]);
        }
    }

    for (i_block = 0; i_block < n_blocks_y; i_block++) {
        for (j_block = 0; j_block < n_blocks_x; j_block++) {
            starpu_insert_task(computeStress, blocks[i][j]);
        }
    }

    for (i_block = 0; i_block < n_blocks_y; i_block++) {
        for (j_block = 0; j_block < n_blocks_x; j_block++) {
            starpu_insert_task(computeVelocity, blocks[i][j]);
        }
    }
}

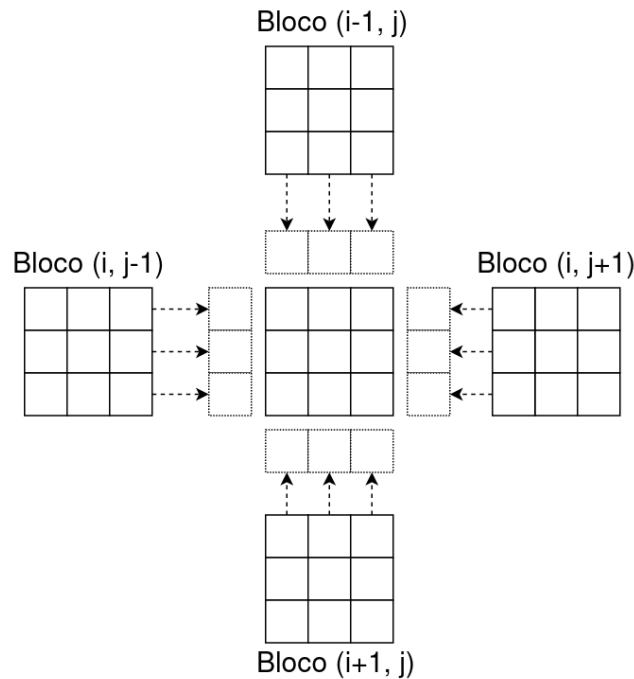
```

Fonte: Autor

Internamente, cada tarefa percorrerá apenas o espaço determinado pelo tamanho de bloco utilizado em vez de uma região completa como na versão original.

A utilização de blocos independentes introduz, no entanto, uma dificuldade nas aplicações do tipo *stencil*: os elementos nas bordas de cada um dos blocos terão como vizinhos elementos de outros blocos. Uma primeira solução, mais simples e menos eficiente, consiste em não utilizar somente um bloco por tarefa, mas sim cinco - o bloco a ser calculado, em modo de leitura e escrita, e os seus quatro vizinhos imediatos, em modo leitura. Essa alternativa, dependendo do contexto em que ela é utilizada, pode representar um custo de memória mais elevado do que o necessário. Uma solução mais eficiente é a utilização das chamadas *ghost cells*. Já que é necessário ler apenas as células de uma borda de cada um dos vizinhos, em vez de utilizar os quatro vizinhos completos, compartilha-se apenas as bordas que serão necessárias. A Figura 4.4 mostra um exemplo de células fantasma em um *stencil* 2D, mas o conceito pode ser expandido para três dimensões.

Figura 4.4 – Representação do uso de células fantasma



Fonte: Autor

## 4.2 Implementação

A fim de explorar as dificuldades da implementação em forma de tarefas e de analisar o nível de paralelismo alcançável usando esse paradigma, uma versão da apli-

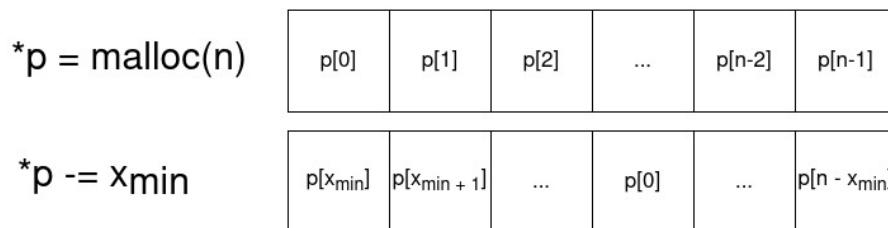
cação *Ondes3D* foi desenvolvida utilizando a biblioteca *StarPU*. As subseções seguintes detalham o processo de implementação realizado.

#### 4.2.1 Inicialização

O primeiro passo para adaptar o algoritmo desenvolvido de maneira a melhor aproveitar as interfaces oferecidas pela biblioteca *StarPU* foi a adequação de seus vetores multidimensionais. Apesar do uso de ponteiros múltiplos facilitar o acesso às estruturas, a biblioteca conta com funções que realizam o ladrilhamento dos dados a partir do número de blocos desejado. Portanto, optou-se por transformar todos os vetores multidimensionais em vetores de tamanho equivalente.

Ainda tratando de vetores, a aplicação original realizava um acesso a partir de suas coordenadas, mesmo que estas possuíssem valores negativos. A figura 4.5 mostra a técnica utilizada para alcançar esse efeito. Os ponteiros que endereçavam essas estruturas apontavam não para a primeira posição alocada, como normalmente é feito, mas para o  $n$ -ésimo elemento, sendo  $n$  a coordenada de valor mais negativo. Devido à complexidade em manter essa implementação utilizando vetores unidimensionais, optou-se por empregar um endereçamento convencional.

Figura 4.5 – Representação de um vetor que permite índices negativos



Fonte: Autor

Devido às mudanças de dimensionalidade e limites dos vetores, o acesso à essas variáveis se torna mais complexo. Para contornar esse problema, foram utilizadas macros que calculam o endereço desejado a partir das dimensões do vetor acessado e das posições solicitadas, detalhadas na seção ??.

#### 4.2.1.1 Inicialização das CPML

As *ABSORBING\_BOUNDARY\_CONDITION* não representam o domínio completo do problema, já que elas tratam unicamente das bordas, onde a propagação da onda deve ser absorvida sem produzir reflexões. Como elas são responsáveis por registrar contribuições tanto nos tensores de tensão quanto nos de velocidade, ambos construídos com nove componentes, elas exigiriam um espaço enorme de memória para seu armazenamento. No entanto, essas contribuições só são calculadas em regiões muito limitadas do domínio, o que faria esses tensores extremamente esparsos.

A fim de não utilizar memória de maneira desnecessária, as contribuições físicas das CPMLs são registradas em vetores unidimensionais de tamanho igual ao número de elementos nas mesmas. Um tensor chamado *ipml* armazena, para cada coordenada do domínio, o seu índice nos vetores das *ABSORBING\_BOUNDARY\_CONDITION*.

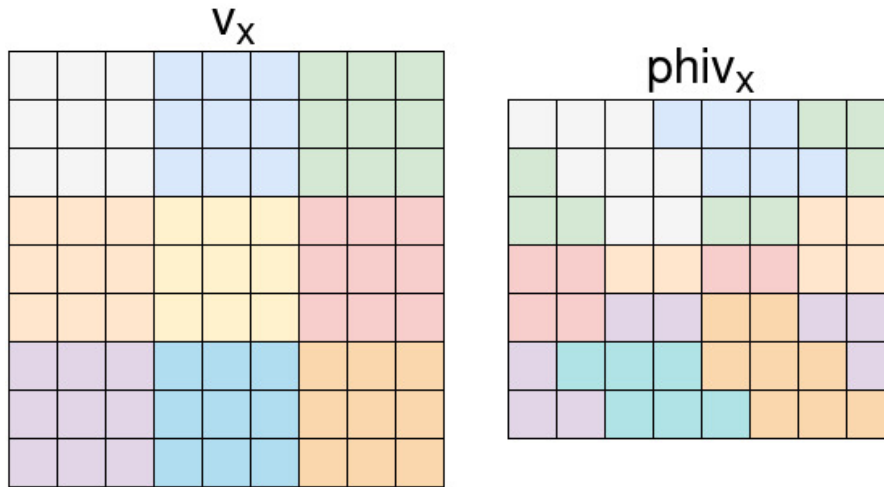
Ao dividir o problema em tarefas, no entanto, essas estruturas adicionam novas complexidades. Primeiramente, não é possível aplicar o particionamento em blocos da mesma maneira nesses vetores, já que eles possuem dimensões reduzidas. Além disso, a implementação original considera que as bordas possuem uma espessura de duas células.

A Figura 4.6 ilustra uma matrix de velocidades  $v_x$  e um vetor  $phiv_x$ , que armazena as contribuições das bordas em  $v_x$ . Considerando que as cores da figura codificam o bloco ao qual cada célula pertence, pode-se constatar que não é possível particionar os vetores de maneira convencional, pois as contribuições de um mesmo bloco não são necessariamente sequenciais. Naturalmente, ao considerarmos o espaço tridimensional no qual os cálculos são feitos, esses fatores se comportam de maneira similar.

O *macro-kernel computeIntermediates*, por exemplo, acessa as contribuições das bordas em modo de escrita. Portanto, caso nenhum particionamento seja feito, mesmo que várias tarefas sejam submetidas para tratar diferentes regiões do problema, elas não poderão ser executadas em paralelo devido à necessidade de escrever em um mesmo vetor. Por esse motivo, é necessário encontrar uma alternativa de particionamento para esses vetores.

A solução implementada para este estudo começa pela transformação da estrutura de dados das condições de borda, conforme mostra a Figura 4.7. Os diversos vetores que compunham as contribuições no cálculo de velocidade, por exemplo, foram encapsulados em uma nova estrutura *phiv\_t* (respectivamente *phit\_t* no caso da tensão), composta por um endereço de base, o tamanho de seus vetores e o deslocamento existente entre seus endereços.

Figura 4.6 – Representação de uma matriz de velocidades e um vetor de contribuições das bordas



Fonte: Autor

Figura 4.7 – À esquerda, parte da implementação da estrutura *ABSORBING\_BOUNDARY\_CONDITION* e à direita, parte da nova implementação proposta.

```

struct ABSORBING_BOUNDARY_CONDITION {
    int npmlv;          /* numbers of cell in ABC */
    int npmlt;
    int ***ipml;        /* index in the PML */

    double *phivxx;
    double *phivyy;
    double *phivzz;
    double *phivxy;
    double *phivyx;
    double *phivxz;
    double *phivzx;
    double *phivyz;
    double *phivzy;

    double *phitxxx;
    double *phitxyy;
    double *phitxzz;
    double *phitxyx;
    double *phityyy;
    double *phityzz;
    double *phitxzx;
    double *phityzy;
    double *phitzzz;
};

struct phiv_s{
    double* base_ptr;
    int size;
    int offset;
    double *xx;
    double *yy;
    double *zz;
    double *xy;
    double *yx;
    double *xz;
    double *zx;
    double *yz;
    double *zy;
};

struct ABSORBING_BOUNDARY_CONDITION {
    int npmlv;          /* numbers of cell in ABC */
    int npmlt;
    int *ipml;          /* index in the PML */

    struct phiv_s *phiv; //List of phiv blocks
    struct phit_s *phit; //List of phit blocks
};

```

Fonte: Autor

A Figura 4.8 mostra uma *macro* desenvolvida para simplificar o acesso aos vetores utilizando a estrutura descrita, onde o acesso a cada uma das componentes é dado calculando a soma do endereço base com o deslocamento referente ao vetor desejado.

Finalmente, na estrutura *ABSORBING\_BOUNDARY\_CONDITION*, os conjuntos de vetores *phiv* e *phit* foram substituídos por vetores dos tipos *phiv\_t* e *phit\_t*, com tamanho igual ao número de blocos do problema. Dessa forma, sendo *lx* a quantidade de blocos de cada linha do domínio, cada bloco (*i*, *j*) encontra em *phiv\_t*[*lx* \* *i* + *j*] o

Figura 4.8 – Macro desenvolvida para construir a nova estrutura proposta.

```
#define COMPUTE_ADDRESS_PHIV_S(phiv) phiv.xx = phiv.base_ptr;\
                                     phiv.yy = phiv.base_ptr + phiv.offset;\
                                     phiv.zz = phiv.base_ptr + 2*phiv.offset;\
                                     phiv.xy = phiv.base_ptr + 3*phiv.offset;\
                                     phiv.yx = phiv.base_ptr + 4*phiv.offset;\
                                     phiv.xz = phiv.base_ptr + 5*phiv.offset;\
                                     phiv.zx = phiv.base_ptr + 6*phiv.offset;\
                                     phiv.yz = phiv.base_ptr + 7*phiv.offset;\
                                     phiv.zy = phiv.base_ptr + 8*phiv.offset;
```

Fonte: Autor

conjunto de vetores associados às contribuições das bordas em  $(i, j)$ .

Esse particionamento realizado introduziu a necessidade de adaptar o vetor *ipml*, responsável por associar cada coordenada à sua posição nos vetores *phiv* e *phit*. Anteriormente, o valor armazenado era um índice global, utilizado em vetores que representavam o domínio completo do problema. Com a nova estrutura, é necessário que esses índices sejam locais, associados aos vetores que pertencem à um bloco cada um.

Implementando todas as modificações descritas acima, o particionamento dos vetores das condições de borda proposto permite que cada tarefa acesse apenas as posições que lhe são necessárias e o paralelismo entre blocos se torna possível.

## 4.2.2 Criação de *data handles*

Para que os dados sejam acessados pelas tarefas submetidas na forma de *buffers*, foi necessário encapsulá-los em *data handles* implementados pela biblioteca *StarPU* e responsáveis por gerenciar réplicas dos mesmos dados. Durante a execução desse projeto foram utilizadas as funções *starpup\_data\_vector\_register*, *starpup\_data\_matrix\_register* e *starpup\_data\_block\_register* para encapsular vetores unidimensionais, matrizes e tensores tridimensionais, respectivamente. Essas funções têm como argumentos o endereço de início e as dimensões dos dados envolvidos, valores que posteriormente são acessíveis dentro das tarefas.

A biblioteca *StarPU* oferece funções para realizar o ladrilhamento dos dados contidos em seus *data handles* através de filtros chamados *starpup\_data\_filter*, que têm como parâmetros o número de blocos a serem criados e o ponteiro para uma função que implementa o particionamento desejado.

No caso desse estudo, os blocos são formados no plano *XY*, mantendo cada um



deles com uma altura igual à do domínio. Portanto, foi necessário a aplicação de duas funções de particionamento, uma responsável pelo eixo  $X$  e outra pelo eixo  $Y$ . Através da função *starpu\_data\_map\_filters*, foi possível executar o ladrilhamento de um conjunto de dados utilizando a composição de diferentes funções de particionamento.

Depois de aplicar o filtro nos dados desejados, o acesso aos blocos é realizado através da função *starpu\_data\_get\_sub\_data*, que tem como argumentos o *data handle* original, o número de dimensões no qual ele foi particionado e os índices do bloco em cada uma dessas dimensões. Dessa maneira é possível fornecer para cada tarefa apenas o bloco que lhe é necessário.

Finalizados os cálculos, os dados dos *data handles* podem ser recuperados em suas estruturas originais após uma chamada à função *starpu\_data\_unregister*. A Figura 4.9 resume o ciclo de vida de uma dessas estruturas ao longo da implementação desenvolvida.

Figura 4.9 – Ciclo de vida de um *data handle*.

```
starpu_data_handle_t t0_xx_handle;

struct starpu_data_filter x_filter = {.filter_func = starpu_block_filter_block,
                                     .nchildren = n_blocks_x};
struct starpu_data_filter y_filter = {.filter_func = starpu_block_filter_vertical_block,
                                     .nchildren = n_blocks_y};

starpu_block_data_register(&t0_xx_handle, STARPU_MAIN_RAM, (uintptr_t)t0->xx,
                          ncols, depth, nrows, ncols, depth, sizeof(t0->xx[0]));
starpu_data_map_filters(t0_xx_handle, 2, &x_filter, &y_filter);

for (iterations = 0; iterations < max_it; iterations++) {
    for (i_block = 0; i_block < n_blocks_y; i_block++) {
        for (j_block = 0; j_block < n_blocks_x; j_block++) {
            starpu_insert_task(..., starpu_data_get_sub_data(t0_xx_handle, 2, i_block, j_block), ...);
        }
    }
}
starpu_data_unregister(t0_xx_handle);
```

Fonte: Autor

#### 4.2.2.1 Acesso aos vizinhos

Sendo o *Ondes3D* um *stencil*, a aplicação de um ladrilhamento implica em uma dificuldade especial na gestão de seus vizinhos, pois o cálculo dos elementos posicionados nas bordas de um bloco necessita de dados presentes em outros blocos. Para solucionar esse problema mantendo o particionamento, é necessário que em cada tarefa, para cada tensor utilizado, sejam enviados não só o bloco que lhes pertence, mas também seus quatro vizinhos. É importante ressaltar que os vizinhos são acessados em modo de leitura, permitindo que o paralelismo entre as tarefas de uma mesma etapa seja mantido.

No caso da implementação proposta, onde o ladrilhamento ocorre apenas no plano

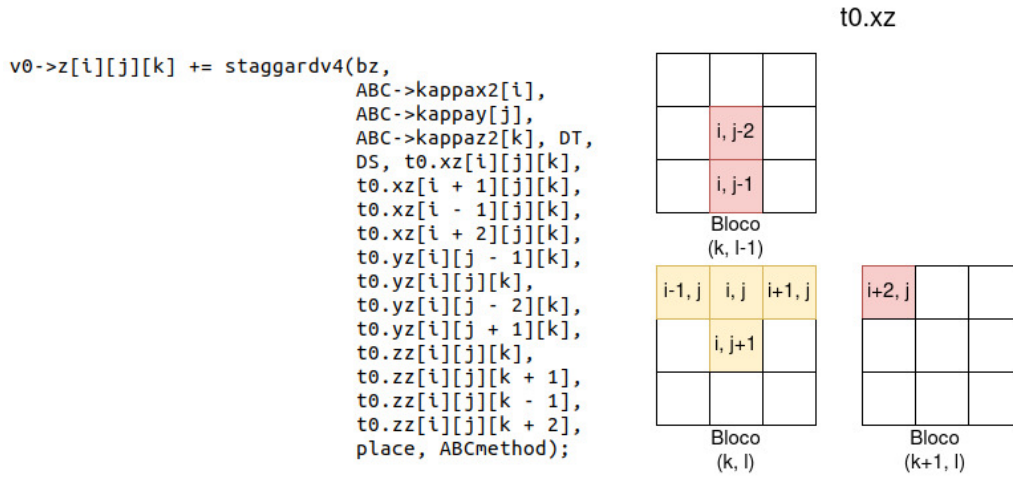
$XY$ , um bloco pode possuir de dois a quatro vizinhos em função de sua posição no domínio: blocos posicionados em um dos quatro cantos do plano  $XY$  possuem apenas dois vizinhos, enquanto os demais blocos posicionados nas bordas possuem três e os internos possuem quatro. Por isso, o laço de inserção de tarefas precisa avaliar a posição tratada e construir um vetor de blocos contendo os vizinhos que serão utilizados em cada chamada.

No entanto, a implementação das tarefas tornou-se muito complexa ao utilizar essa solução. O cálculo realizado em cada célula  $(i, j, k)$  não depende apenas de seus quatro vizinhos diretos, mas também dos quatro vizinhos posicionados a duas células de distância. Desde que o tamanho do bloco seja maior que dois, essa propriedade não interfere na gestão dos vizinhos em si. Porém, dentro de cada tarefa, torna-se necessário tratar uma quantidade significativa de casos distintos. São eles:

- Células na primeira ou última linha do bloco: índices do tipo  $i - 1$  e  $i - 2$  acessam os blocos vizinhos enquanto índices do tipo  $i$  acessam o bloco principal
- Células na primeira ou última coluna do bloco: índices do tipo  $j - 1$  e  $j - 2$  acessam os blocos vizinhos enquanto índices do tipo  $j$  acessam o bloco principal
- Células na segunda ou penúltima linha do bloco: índices do tipo  $i - 2$  acessam os blocos vizinhos enquanto índices do tipo  $i$  e  $i - 1$  acessam o bloco principal
- Células na segunda ou penúltima coluna do bloco: índices do tipo  $j - 2$  acessam os blocos vizinhos enquanto índices do tipo  $j$  e  $j - 1$  acessam o bloco principal
- Demais células: todos os índices acessam o bloco principal

Além das possibilidades listadas, as operações realizadas são implementadas em funções que acessam mais de uma célula, criando assim ainda mais situações possíveis a serem consideradas. A Figura 4.10 ilustra os acessos necessários para parte do cálculo da célula  $(i, j)$  do bloco  $(l, k)$  do tensor de velocidade na direção do eixo  $z$  na etapa *computeVelocity*.

Mesmo sendo mais eficiente em memória, o emprego da técnica de células fantasma, explicado na seção 4.1, não diminui a complexidade envolvida nesses casos. No entanto, constatou-se que as grandezas cujos acessos podem necessitar de vizinhos não são aquelas que estão sendo atualizadas. No exemplo mostrado acima, são os valores de tensão que são consultados para calcular a velocidade em uma dada coordenada. Sendo assim, optou-se por fornecer para cada tarefa um bloco, no caso dos dados que serão escritos, e o tensor completo, no caso dos dados que serão lidos. Com isso, o paralelismo pode

Figura 4.10 – Parte do cálculo de  $v0 \rightarrow z[i][j][k]$ .

Fonte: Autor

ser mantido, já que as escritas acontecem em dados independentes, sem a necessidade de gerenciar leituras em blocos vizinhos.

#### 4.2.3 Transformação em tarefas

Cada um dos quatro *macro-kernels* foi escrito seguindo a interface de tarefas *StarPU*. Inicialmente serão abordados os aspectos gerais dessa transformação, comuns a todas as etapas do algoritmo. Em seguida, a implementação da tarefa responsável pelo cálculo da velocidade será vista em detalhes devido à particularidades nela encontradas.

A implementação de uma tarefa *StarPU* começa pela leitura dos blocos utilizados, fornecidos na forma de *buffers*, através das *macros* *STARPU\_VECTOR\_GET\_PTR*, *STARPU\_MATRIX\_GET\_PTR* e *STARPU\_BLOCK\_GET\_PTR*. Além disso, é necessário consultar as dimensões desses *buffers*; apesar do ladrilhamento separar o espaço em blocos de tamanho igual, é possível que as extremidades possuam blocos de tamanho menor caso não seja possível realizar uma divisão inteira.

Os laços internos, que originalmente percorriam uma região especificada na chamada do *macro-kernel*, passam a percorrer o tamanho de um bloco. No entanto, ao acessar as estruturas fornecidas apenas para leitura, é necessário realizar uma conversão das coordenadas locais em coordenadas globais, já que elas não serão particionadas.

Para percorrer os vetores que eram originalmente acessados por suas coordenadas, três *macros* foram implementadas em função de suas dimensionalidades, conforme mostra

a Figura 4.11. Logo, todos os acessos à vetores foram substituídos por chamadas às novas *macros*.

Figura 4.11 – *Macros* utilizadas para acessar vetores.

```
#define i3_access(p, nrl, nrh, ncl, nch, ndl, ndh, i, j, k)
    p[(i-(nrl))*((nch)-(ncl)+1)*((ndh)-(ndl)+1)+
      (j-(ncl))*((ndh)-(ndl)+1)+
      (k-(ndl))]

#define i2_access(m, nrl, nrh, ncl, nch, i, j)
    m[(i-(nrl))*((nch)-(ncl)+1)+
      (j-(ncl))]

#define i_access(p, nl, nh, i)
    p[(i) - (nl) + NR_END]
```

Fonte: Autor

#### 4.2.3.1 Cálculo de velocidades

Apesar da seção 4.2.2.1 mostrar que cada uma das etapas só precisa acessar vizinhos ao realizar leituras, existe uma exceção no cálculo de velocidades. Quando o índice  $k$  é igual a 1 ou 2, certas operações adicionais são necessárias e estas eventualmente envolvem blocos vizinhos onde outras tarefas precisam atualizar dados. No caso em que  $k$  é igual a um, lê-se dados do bloco seguinte no eixo  $X$  e do bloco anterior no eixo  $Y$ , enquanto no caso em que  $k$  é igual a dois, os dados externos pertencem ao bloco anterior no eixo  $X$  e posterior no eixo  $Y$ .

Além da gestão dos vizinhos, a necessidade de consultar dados da mesma estrutura em que se está escrevendo adiciona também novas dependências relativas à ordem dos cálculos. Se na aplicação original os vizinhos lidos são sempre aqueles cujas células já foram atualizadas, é preciso manter essa mesma ordem na implementação paralela, quando cada bloco efetua seus cálculos independentemente.

O primeiro ponto demonstrado sugere que uma implementação de tarefas incluindo os possíveis vizinhos dos tensores de velocidade seria capaz de resolver a questão, mas o segundo ponto mostra que só isso não é suficiente para garantir que o mesmo cálculo será realizado. Por isso, a solução encontrada foi particionar a etapa de computação de velocidades em três *kernels* diferentes, sendo o primeiro deles responsável pelas células de índice  $k$  igual a um, o segundo pelas células de índice  $k$  igual a dois e o terceiro pelos demais índices.

Em relação à vizinhança, em cada um dos casos especiais apenas um par de tensores de velocidade precisa de seus vizinhos, que conforme foi descrito anteriormente, são no máximo dois blocos. Considerando essas particularidades, foi possível implementar essas tarefas de maneira que os *kernels* responsáveis pelos casos especiais tenham, cada um deles, apenas dois blocos vizinhos, que são fornecidos apenas para leitura. A Figura 4.12 apresenta um resumo da estrutura dessa etapa, onde a ordem de submissão das três tarefas garante a ordem de execução necessária.

Figura 4.12 – Resumo das chamadas de computação de velocidade.

```
for (i_block = 0; i_block < n_blocks_y; i_block++) {
    for (j_block = 0; j_block < n_blocks_x; j_block++) {

        block_v0_x = starpu_data_get_sub_data(v0_x_handle, 2, i_block, j_block);
        block_v0_y = starpu_data_get_sub_data(v0_y_handle, 2, i_block, j_block);
        block_v0_z = starpu_data_get_sub_data(v0_z_handle, 2, i_block, j_block);

        prev_i = (i_block != 0) ? starpu_data_get_sub_data(v0_z_handle, 2, i_block-1, j_block) : NULL;
        prev_j = (j_block != 0) ? starpu_data_get_sub_data(v0_y_handle, 2, i_block, j_block-1) : NULL;
        next_i = (i_block != n_blocks_y-1) ? starpu_data_get_sub_data(v0_x_handle, 2, i_block+1, j_block) : NULL;
        next_j = (j_block != n_blocks_x-1) ? starpu_data_get_sub_data(v0_z_handle, 2, i_block, j_block+1) : NULL;

        starpu_task_insert(&velo_cl, STARPU_RW, block_v0_x, STARPU_RW, block_v0_y, STARPU_RW, block_v0_z,
            ...);
        starpu_task_insert(&velo_k1_cl, STARPU_RW, block_v0_x, STARPU_RW, block_v0_y,
            STARPU_RW, block_v0_z, STARPU_R, next_i, STARPU_R, prev_j,
            ...);
        starpu_task_insert(&velo_k2_cl, STARPU_RW, block_v0_x, STARPU_RW, block_v0_y,
            STARPU_RW, block_v0_z, STARPU_R, prev_i, STARPU_R, next_j,
            ...);
    }
}
```

Fonte: Autor

## 5 AVALIAÇÃO EXPERIMENTAL

A nova versão desenvolvida foi executada no Parque Computacional de Alto Desempenho (PCAD) da Universidade Federal do Rio Grande do Sul com o intuito de comparar seu desempenho em relação à implementação original.

### 5.1 Ambiente de testes

Para a execução dos testes, foram utilizadas as máquinas *tupi* do PCAD, contando com um processador *Intel Xeon E5-2620* com 16 núcleos e 32 *threads*, 80GB de memória RAM DDR4 e duas placas *NVIDIA GeForce GTX 1080Ti*.

### 5.2 Resultados

Apesar do objetivo original ser a comparação da nova implementação com a antiga, não foi possível concluir o desenvolvimento da nova versão utilizando a programação baseada em tarefas. Todas as tarefas foram construídas e o algoritmo é executado do início ao fim sem produzir erros de execução. No entanto, ao comparar o resultado final obtido, não encontra-se o valor fornecido pela aplicação original.

Diversas investigações foram realizadas e indicam que os erros matemáticos encontram-se no cálculo das contribuições das condições de borda, etapa previamente identificada como mais complexa. Devido ao tempo disponível, optou-se por apresentar o estudo desse desenvolvimento independentemente da corretude numérica dos resultados devido ao valor científico das ponderações realizadas ao longo da implementação e também da possibilidade de realizar uma análise sobre o paralelismo alcançado com as tarefas da maneira proposta.

**\*\* análise do paralelismo quando rolar o starvz \*\***

### 5.3 Trabalhos futuros

A sequência natural do trabalho apresentado é a correção dos erros numéricos introduzidos na mudança de modelo de programação realizada. Assim que esses resultados forem alcançados, diversas melhorias podem ser postas em prática de maneira a acelerar

ainda mais a execução da aplicação. Algumas delas são as seguintes:

- Implementação visando arquiteturas heterogêneas: assim como o trabalho desenvolvido por Martínez et al. (MARTÍNEZ et al., 2015), é possível implementar *kernels* que permitem a execução de tarefas em *GPUs*.
- Implementação visando paralelismo entre iterações: duplicando as estruturas de dados, é possível manter em memória valores de duas iterações simultaneamente, o que eventualmente possibilitaria uma execução que dispensa a necessidade de barreiras de sincronização entre iterações.
- Implementação de blocos de formato diferente: em vez de limitar o ladrilhamento ao plano  $XY$ , a opção de particionar o espaço em todos os eixos permite o emprego de mais parâmetros que, refinados, podem gerar melhores resultados.
- Implementação de tarefas de granularidade menor: as tarefas implementadas são compostas por diversas etapas menores, frequentemente associadas às diferentes regiões do domínio calculado. A transformação dessas etapas menores em tarefas resulta numa granularidade mais fina, o que pode contribuir para o equilíbrio de carga entre as unidades de execução.

## **6 CONCLUSÃO**

falar do equilibrio de carga!!!! barreira entre iterações!!!!!!! falar do unpartition!!



## REFERÊNCIAS

AUGONNET, C. et al. Starpu: A unified platform for task scheduling on heterogeneous multicore architectures. In: . [S.l.: s.n.], 2009. v. 23. ISBN 978-3-642-03868-6.

BERENGER, J.-P. A perfectly matched layer for the absorption of electromagnetic waves. **Journal of Computational Physics**, v. 114, n. 2, p. 185–200, 1994. ISSN 0021-9991. Available from Internet: <<https://www.sciencedirect.com/science/article/pii/S0021999184711594>>.

BOITO, F. Z. et al. High performance i/o for seismic wave propagation simulations. In: **2017 25th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)**. [S.l.: s.n.], 2017. p. 31–38.

CASTRO, M. et al. Seismic wave propagation simulations on low-power and performance-centric manycores. **Parallel Computing**, v. 54, p. 108–120, 2016. ISSN 0167-8191. 26th International Symposium on Computer Architecture and High Performance Computing. Available from Internet: <<https://www.sciencedirect.com/science/article/pii/S0167819116000417>>.

DUPROS, F.; DO, H.-T.; AOCHI, H. On scalability issues of the elastodynamics equations on multicore platforms. In: **ICCS 2013 : International conference on computational science**. Barcelona, Spain: Elsevier, 2013. (Procedia Computer Science), p. 9 p. Available from Internet: <<https://hal-brgm.archives-ouvertes.fr/hal-00797682>>.

DUPROS, F. et al. High-performance finite-element simulations of seismic wave propagation in three-dimensional nonlinear inelastic geological media. **Parallel Comput.**, v. 36, p. 308–325, 2010.

MARTÍNEZ, V. et al. Towards seismic wave modeling on heterogeneous many-core architectures using task-based runtime system. In: **2015 27th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)**. [S.l.: s.n.], 2015. p. 1–8.

NATAF, F. Absorbing boundary conditions and perfectly matched layers in wave propagation problems. **Radon Series on Computational and Applied Mathematics**, v. 11, 01 2013.

NESI, L. L. et al. Task-based parallel strategies for computational fluid dynamic application in heterogeneous cpu/gpu resources. **Concurrency and Computation: Practice and Experience**, v. 32, n. 20, p. e5772, 2020. Available from Internet: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.5772>>.

RODEN, J. A.; GEDNEY, S. D. Convolution pml (cpml): An efficient fdtd implementation of the cfs-pml for arbitrary media. **Microwave and Optical Technology Letters**, v. 27, n. 5, p. 334–339, 2000. Available from Internet: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/1098-2760%2820001205%2927%3A5%3C334%3A%3AAID-MOP14%3E3.0.CO%3B2-A>>.