Memorando

De: Willfredy Vieira Dias

Nº de Matrícula: 20200204

Disciplina: Computação Paralela e Distribuída (CPD)

Assunto: Aula de Laboratório Nº 2 – Introduzir o OpenMP

Data: 04/04/2025

1. Introdução

A experiência desenvolvida baseia-se em duas partes principais que visam introduzir os fundamentos do OpenMP:

Contexto do Trabalho: Realizado no âmbito da disciplina de Computação Paralela e Distribuída, durante a Aula de Laboratório nº 2 no ISPTEC 2021-22. O exercício tem como objetivo familiarizar os estudantes com o ambiente de compilação e execução de programas utilizando OpenMP.

Objectivos:

- Apresentar e utilizar a API e as diretivas do OpenMP.
- Experimentar a execução paralela dos códigos, ajustando o número de threads através da variável de ambiente (OMP_NUM_THREADS).
- Analisar o impacto de cláusulas como "nowait" e a inserção de barreiras, evidenciando os efeitos da sincronização no comportamento e na saída do programa.
- Abordar a paralelização de um loop com dependências de dados, comparando uma abordagem inicial (incorreta) com uma versão corrigida, que preserva a integridade dos dados, e explorar alternativas para evitar a cópia desnecessária do vector.

Restrições e Observações Importantes:

- É imprescindível compilar os códigos com a flag "-fopenmp" para ativar o suporte ao OpenMP.
- Recomenda-se a consulta à documentação oficial do OpenMP (versão 4.5) para esclarecer eventuais dúvidas sobre as directivas ou APIs.

- No primeiro problema, a alteração da saída do programa ao adicionar a cláusula "nowait" e a barreira deve ser cuidadosamente analisada, uma vez que estas modificações afetam a sincronização entre as threads.
- No segundo problema, a paralelização do loop interno deve ser tratada com atenção, dada a existência de dependências de dados; uma implementação simples pode conduzir a resultados incoerentes, pelo que se torna necessário copiar os dados da iteração anterior ou adoptar estratégias que evitem esta cópia, mantendo assim a consistência do processamento.

Esta atividade permite uma compreensão prática dos desafios e das estratégias necessárias para implementar paralelismo, enfatizando a importância da sincronização e do correto tratamento das dependências de dados.

2. Experiências Realizadas

Problema 1

a) Comece por compilar o código acima, não esqueça de adicionar a flag de compilação "-fopenmp". Execute este programa com um número diferente de threads (definindo *OMP_NUM_THREADS*).

```
| Comparison | Com
```

figura a 1 - Execução do programa com o OMP_NUM_THREADS.

b) Adicione a cláusula "nowait" à directiva for na linha 9. A saída foi alterada? Porquê?

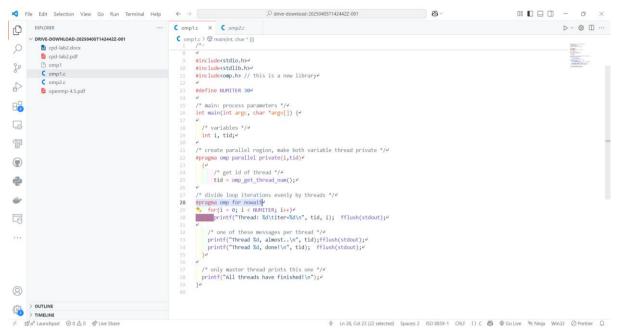


figura b 1 - Cláusula "nowait" adicionada à directiva for na linha 9.

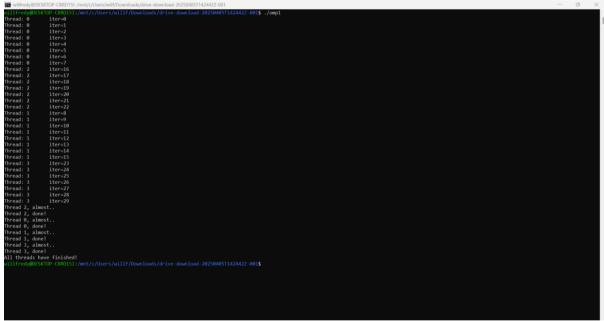


figura b 2 - Execução do programa com a claúsula "nowait".

➤ **Resposta:** Ao utilizar a cláusula **nowait**, retiramos a barreira implícita que normalmente se encontra ao fim do loop **for**. Isto significa que cada thread pode avançar para as instruções seguintes sem ter de esperar que as outras threads concluam todas as iterações do loop.

No código, imediatamente após o loop **for**, cada thread executa os comandos:

```
printf("Thread %d, almost..\n", tid);
printf("Thread %d, done!\n", tid);
```

Sem a barreira, estas mensagens podem aparecer numa ordem diferente, consoante qual thread termina o loop mais rapidamente.

Resumindo:

- **Sem nowait (com barreira):** Todas as threads esperam o fim do loop **for** antes de imprimir as mensagens, resultando numa saída um pouco mais organizada (embora a ordem entre threads ainda possa variar).
- **Com nowait:** As threads começam a executar as linhas seguintes assim que terminam o seu trabalho no loop, o que pode fazer com que as mensagens apareçam de forma intercalada e numa ordem menos previsível.

Assim, a saída pode alterar-se em termos de ordem das mensagens, mas a funcionalidade do programa permanece a mesma.

c) Adicione uma barreira (#pragma omp barrier) entre as linhas 13 e 14. Compare a saída com a) e b).

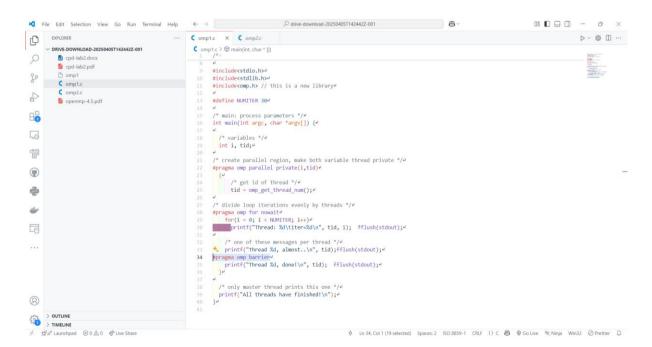


figura c 1 - Barreira #pragma omp barrier adicionada entre as linhas 33 e 34.

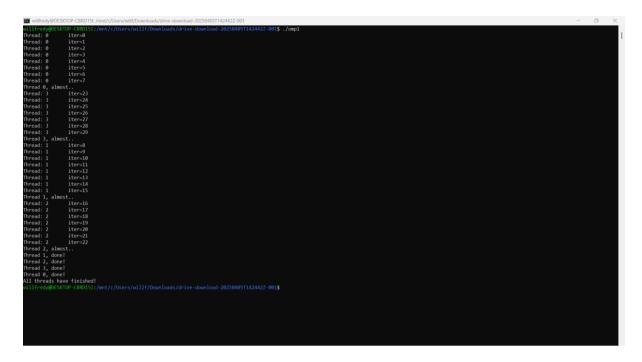


figura c 2 - Programa executado com a barreira #pragma omp barrier.

Comparação com a a línea a) e com a alínea b):

- Alínea a) Sem modificações: As threads executam o loop for e, ao fim do mesmo, existe uma barreira implícita. Assim, todas as threads esperam que o loop seja concluído antes de passarem para as impressões seguintes. A saída tende a ser mais organizada, pois as mensagens das threads são apresentadas apenas depois de todas terem terminado o loop.
- Alínea b) Com cláusula nowait: Ao acrescentar a cláusula nowait à diretiva for, retiramos a barreira implícita. Cada thread pode avançar para os comandos seguintes assim que terminar o seu conjunto de iterações, sem aguardar as restantes. Isto faz com que as mensagens possam ser impressas de forma intercalada, resultando numa saída menos ordenada.
- Alínea c) Com barreira explícita: Ao inserir a barreira com #pragma omp barrier, obrigamos todas as threads a esperarem até que todas concluam o loop for antes de prosseguir para as impressões. O efeito é semelhante ao da alínea a), ou seja, uma saída onde as mensagens de "almost..." e "done!" aparecem de forma sincronizada entre as threads.

Em resumo, a adição da barreira explícita na alínea **c**) restabelece a sincronização entre as threads que foi perdida ao usar o **nowait** na alínea **b**). Dessa forma, a ordem de execução tornase mais previsível e organizada.

Problema 2

a) Compile e execute a versão serial do código (omp2.c).

figura do segundo Problema a 1 - Versão serial (omp2.c) compilada.

b) Crie uma versão paralela simples (e incorreta) adicionando uma directiva paralela ao loop interno. Tente encontrar uma execução com uma saída diferente da versão serial. **Porquê isso pode acontecer?**

```
File Edit Selection View Go Run Terminal Help
                                                                                                   O drive-download-20250405T142442Z-001
                                                                                                                                                                                                                                                             0
0
                                                                          0
             cpd-lab2.docx
             cpd-lab2.pdf
cpd-lab2.pdf
mp1
comp1.c
 20
 Ca
                                                                                /* 1. INITIALIZE VECTOR */J

for(i = 0; i < TOTALSIZE; i++) (J

V[i] = 0.0 + i; J
8
". 2. ITERATIONS LOOP "/J
". for(iter = 0; iter < NANITER; iter++) {"
pragma one parallel for"
/" 2.1. PROCESS ELEMENTS "/J
for(i = 0; i < TOTALSIZE-1; i++) {"
V[i] = f(V[i], V[i+1]);"
J
J
J
 (
 4
 *
1
                                                                                    /* 2.2. END ITERATIONS LOOP */+
                                                                                    /* 3. OUTPUT FINAL VALUES */+/
printf("Output:\n"); -/-/
for(i = 0; i < TOTALSIZE; i++) {+/
printf("%4d %f\n", i, V[i]); -/-/
}-/-
 8
Solution > OUTLINE > TIMELINE
                                                                                                                                               Ò Ln 44, Col 1 (24 selected) Spaces: 2 ISO 8859-1 CRLF () C ← G Go Live II Ninia Win32 ✓ Prettier □
 × 96 Launchpad ⊗ 0 △ 0 € Live Share
```

figura do segundo problema b 1 - Código com a directiva paralela ao loop interno adicionada.

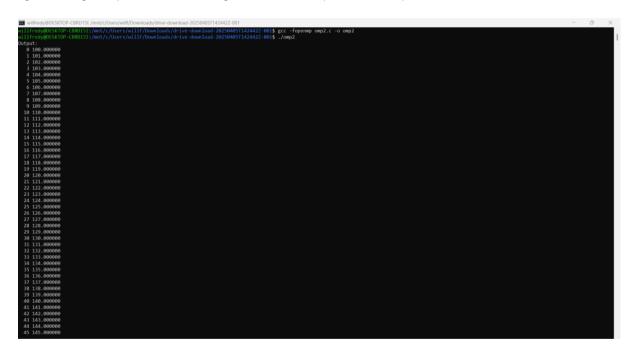


figura do segundo problema b 2 - Programa executado com a directiva paralela ao loop interno adicionada.

Resposta: Nesta versão, as iterações do loop interno são executadas em paralelo, mas a operação apresenta dependências entre elementos (por exemplo, o cálculo de V[i] depende de V[i+1]). Na versão serial, o processamento ocorre de forma ordenada, garantindo que cada elemento seja atualizado de acordo com o valor mais recente dos elementos seguintes.

Ao paralelizar de forma simples, várias threads podem aceder e atualizar os elementos do vetor simultaneamente, causando condições de corrida (race conditions). Assim, a ordem de execução não fica garantida e o resultado final pode ser diferente, pois os valores lidos podem não ser os esperados se outra thread estiver a atualizar o mesmo elemento ao mesmo tempo.

c) Faça a correcção da versão paralela copiando o vector V da iteração i-1 antes de executar a iteração i e usando os valores copiados como argumentos para a função f na iteração i.

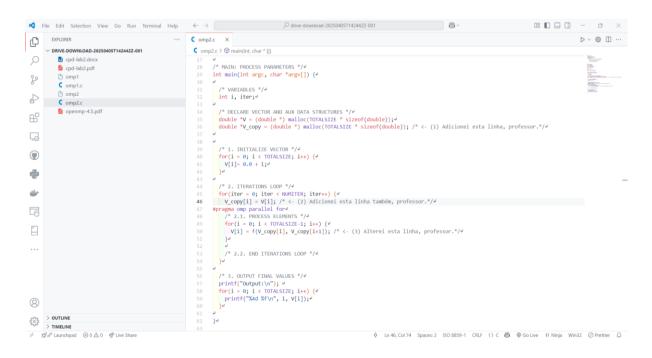


figura do segundo problema c 1 - Código correccção da versão paralela feita com a cópia do vector V da iteração i-1.

```
| Comparison | Com
```

figura do segundo problema c 2 - Programa executado com a correccção da versão paralela feita com a cópia do vector V da iteração i-1.

d) Escreva uma modificação na versão paralela que evite copiar o vector V em todas as iterações.

```
| Time | Selection | View | So | Run | Terminal | Help | C | Parameter | Param
```

figura do segundo problema d 1 - Código com a modificação da versão paralela que evita copiar o vector V em todas as iterações.

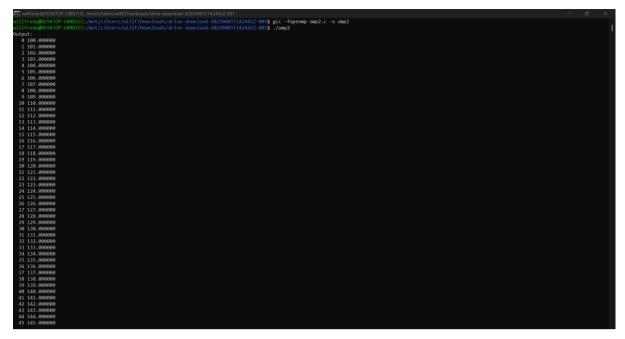


figura do segundo problema d 2 - Execução do programa com a modificação da versão paralela que evita copiar o vector V em todas as iterações.

3. Desafios

Neste trabalho, o principal desafio foi adaptar um código sequencial para uma versão paralela, mantendo a lógica e o resultado corretos. Para isso, foi necessário compreender bem como funciona a execução em paralelo com OpenMP, principalmente no que diz respeito às dependências entre iterações dos ciclos.

A estrutura do software ficou organizada da seguinte forma:

- **Inicialização**: O vetor principal é criado e preenchido com valores iniciais simples, para facilitar a verificação dos resultados.
- **Processamento**: A parte mais importante do programa está no ciclo de iterações, onde o vetor é atualizado várias vezes. Nesta fase, implementámos diferentes versões:
 - o Uma versão **sequencial**, onde tudo é feito de forma linear.
 - Uma versão paralela incorreta, onde tentámos paralelizar diretamente o ciclo interno, o que gerou resultados inconsistentes devido à existência de dependências.
 - Uma versão corrigida, onde criámos uma cópia auxiliar do vetor em cada iteração para evitar as dependências, garantindo resultados corretos.

 Por fim, uma versão otimizada, onde evitámos fazer a cópia do vetor a cada iteração, usando dois vetores e alternando entre eles. Assim, melhorámos a eficiência sem comprometer a correção dos dados.

Durante o desenvolvimento, também foi necessário compreender bem como funcionam as diretivas **#pragma omp**, especialmente as cláusulas **parallel** for, **nowait** e **barrier**. Esses conceitos foram fundamentais para garantir uma execução parallela eficiente e correta.

4. Referências Bibliográficas

http://www.openmp.org/wp-content/uploads/openmp-4.5.pdf

5. Repositório GitHub

https://github.com/Willfredy-Vieira-Dias/Laborat-rios_de_CPD

Coloque aqui o link do repositório e envie o convite de colaborador para o utilizador GitHub joaojdacosta.