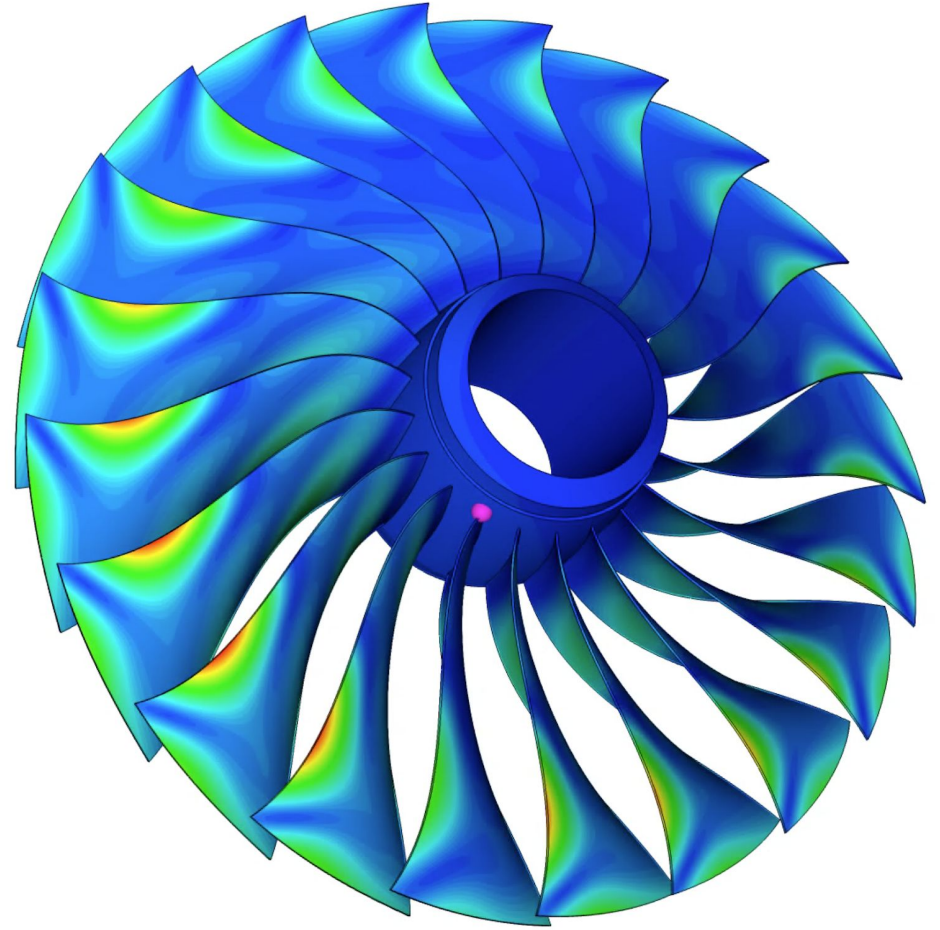


## GPU accelerated computing for Finite Element Method



# Introduction to GPU

## GPU vs CPU

### Matrix – Vector Multiplication

- GPUs for calculations involving dense matrices
- For example:
  - ☐ Compute explicit inverse which is full matrix
  - ☐ For different righthand side vectors
  - ☐ linear systems with Schur complements in FETI are also full

# Introduction to GPU

## Iterative solvers

Iterative solver for  $Ax = b$

Conjugate Gradient

GMRES

Orthomin

Orthores

Orthodir

Bi Conjugate Gradient

### Conjugate Gradient Method

```
1:  $r_0 = b - Ax_0$ 
2:  $p_0 = r_0$ 
3:  $k = 0$ 
4: if  $r^T r < \text{tol}$  then
5:    $\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}$ 
6:    $x_{k+1} = x_k + \alpha_k p_k$ 
7:    $r_{k+1} = r_k + \alpha_k A p_k$ 
8:    $\beta_k = \frac{r_{k+1}^T r_{k+1}}{p_k^T A p_k}$ 
9:    $p_{k+1} = r_{k+1} + \beta_k p_k$ 
10:   $k = k + 1$ 
11: end if
12: return  $x_{k+1}$  as the result
```

- ❑ One Matrix - Vector Product
- ❑ Two vector dot product per iteration
- ❑ Four vectors of working stage

# Introduction to GPU

## GPU vs CPU

One matrix vector multiplication in each iteration.

### Matrix – Vector Multiplication

$$A_{m,n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \quad x_n = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

---

#### Sequential algorithm for Matrix - Vector product

---

```
1: for  $i = 1, 2, \dots, m$  do  
2:    $\text{out}[i] = 0$   
3:   for  $j = 1, 2, \dots, n$  do  
4:      $\text{out}[i] + = \text{mat}[i][j] * x[j]$   
5:   end for  
6: end for
```

---

$$Ax = \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots & a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots & a_{2,n}x_n \\ \vdots & \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots & a_{m,n}x_n \end{pmatrix}$$

# Introduction to GPU

## GPU programming syntax

EXERCISE 2 : Matrix Vector product using GPU program

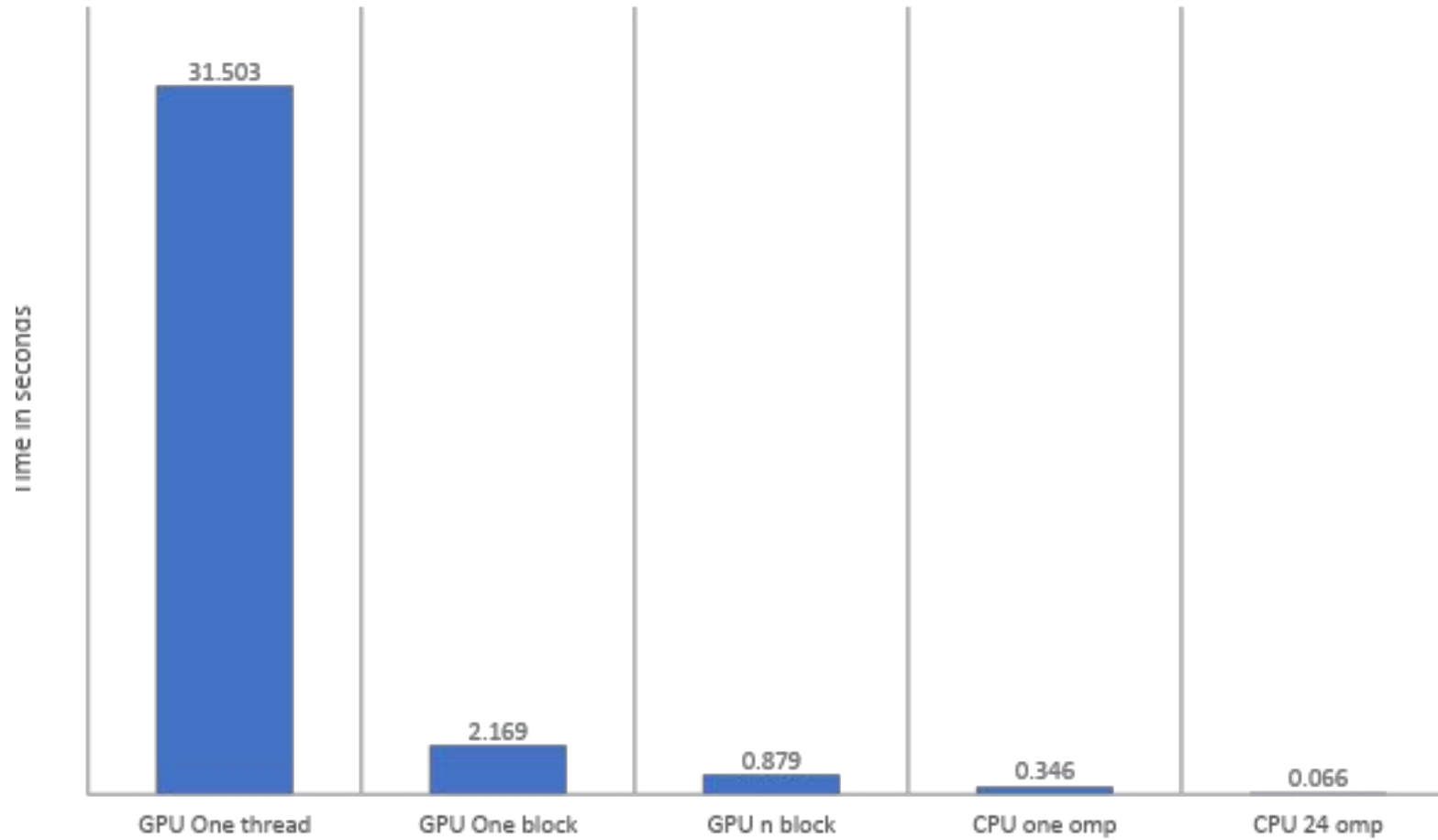
Source code:

[https://github.com/sivasanarul/FEMwithGPU/tree/master/EX2\\_matrixvectmul](https://github.com/sivasanarul/FEMwithGPU/tree/master/EX2_matrixvectmul)

# Introduction to GPU

## Introduction of GPU programming

Vector Addition Example : `vector_addition()`



# Introduction to GPU

## GPU programming syntax

EXERCISE 3 : Matrix Vector product using GPU program

Source code:

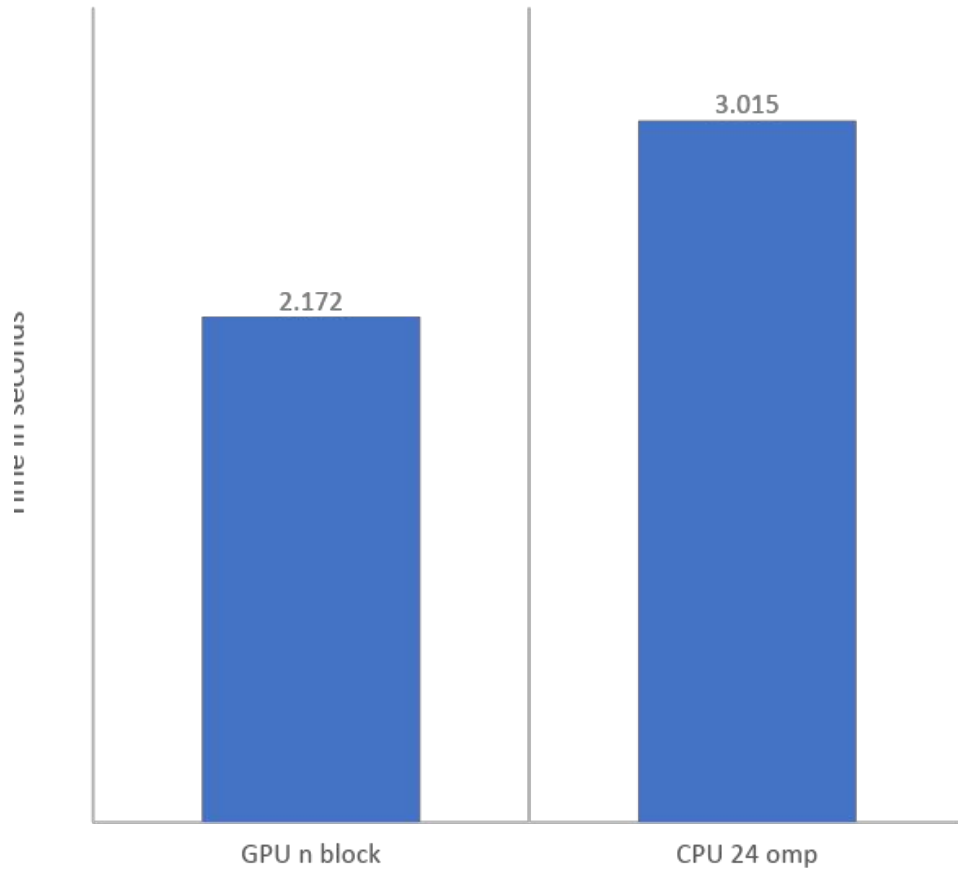
[https://github.com/sivasanarul/FEMwithGPU/tree/master/EX3\\_repmatrixvectmul](https://github.com/sivasanarul/FEMwithGPU/tree/master/EX3_repmatrixvectmul)

# Introduction to GPU

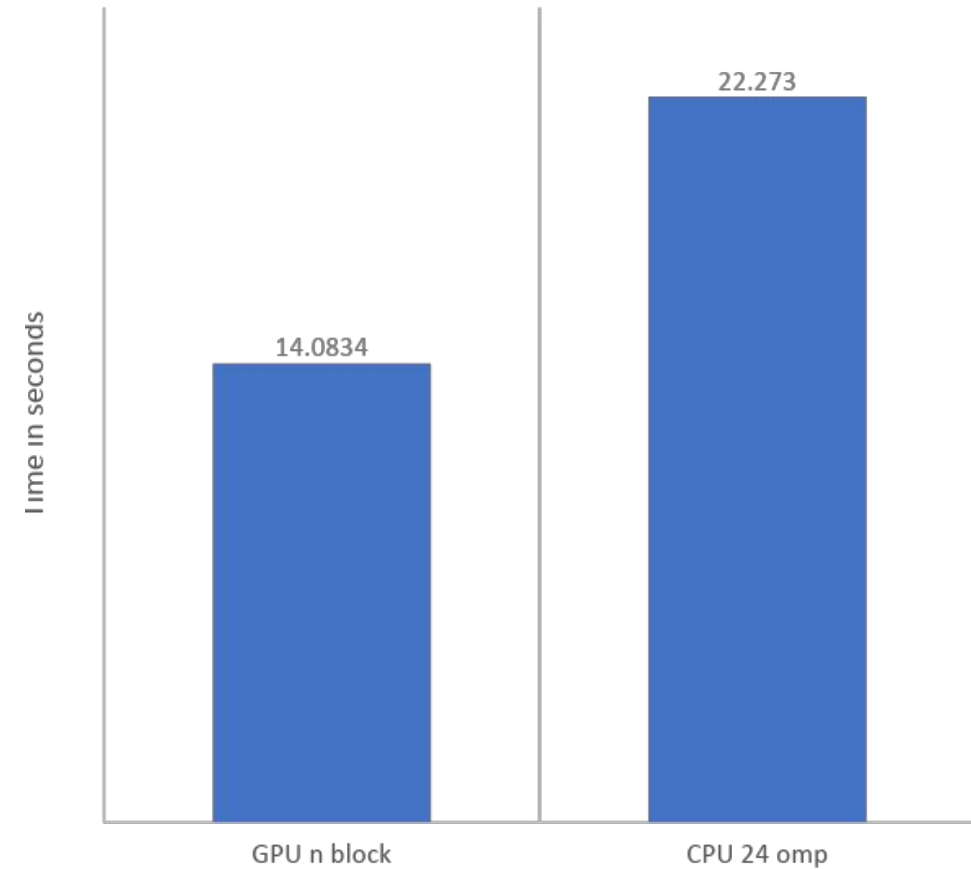
## Introduction of GPU programming

Vector Addition Example : `matrix_vector_product()`

100 iterations



1000 iterations





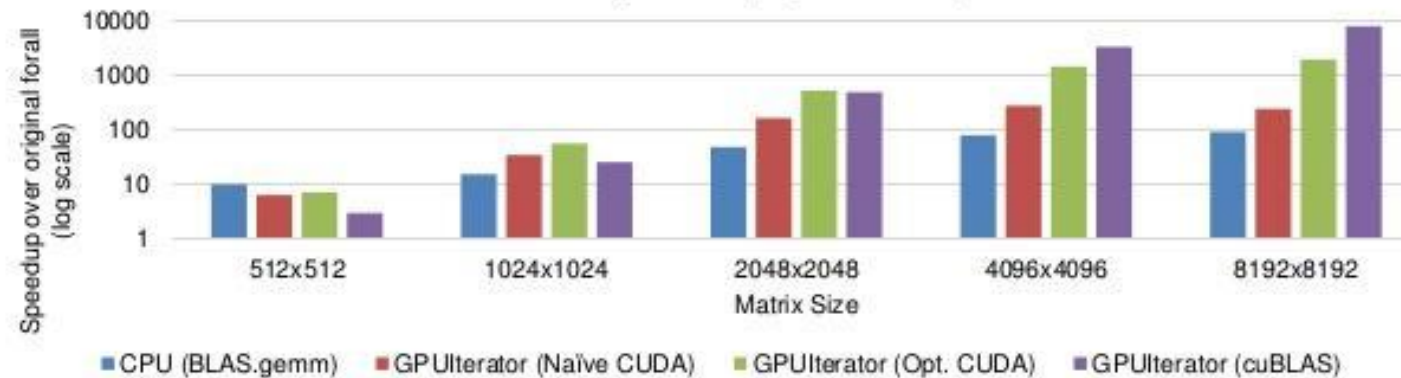
# Introduction to GPU

## Introduction of GPU

Use library and compiler directives instead of programming.

How fast are GPUs compared to Chapel's BLAS module on CPUs?  
(Single-node, Core i5 + Titan Xp)

Matrix Multiplication (Higher is better)



- Motivation: to verify how fast the GPU variants are compared to a highly-tuned Chapel-CPU variant
- Result: the GPU variants are mostly faster than OpenBLAS's gemm (4 core CPUs)

# Introduction to GPU

## GPU programming syntax

EXERCISE 4 : Matrix Vector product using GPU CUBLAS program

Source code:

[https://github.com/sivasanarul/FEMwithGPU/tree/master/EX4\\_matrixvectmul\\_blas](https://github.com/sivasanarul/FEMwithGPU/tree/master/EX4_matrixvectmul_blas)

<https://developer.nvidia.com/sites/default/files/akamai/cuda/files/Misc/mygpu.pdf>