On the parallelization of bio-heat transfer problem using input file affinity measure with MPI

S. U. Ewedafe¹ & R. H. Shariffudin²

¹Department of Computing, Baze University, Abuja, Nigeria
²Faculty of Science, Institute of Mathematical Sciences, University of Malaya, Malaysia

Abstract

In this paper, a method based on applying Successive Over-Relaxation (SOR) to find the solution of 2-Dimensional Bio-Heat Transfer Problem (2DBHTP) on a distributed computing environment using Input File Affinity Measure (I_{aff}) with Message Passing Interface (MPI) is introduced. A finite difference discretization is developed to solve the 2DBHTP since 2DBHTP requires the evaluation of temporal and spatial distributions of temperature. The platform gives us better predictions of the effects of thermal physical properties of temperature distribution. This work proposes a parallel overhead with overlapping communication in its implementation using MPI. Based on the implementation of the MPI distributed computing environment, we present a performance improvement strategy running on the master-slave platform. Parallel comparisons of the method were experimentally evaluated, and parallel results show speedup and efficiency on a various number of mesh sizes. The method used combines elements of numerical stability and parallel algorithm design.

Keywords: 2DBHTP, SOR, I_{aff}, MPI, speedup, discretization, finite difference, spatial temperature, parallel overhead, performance.

1 Introduction

Computing infrastructures are reaching an unprecedented degree of complexity. First, parallel processing is coming to mainstream, because of the frequency and power consumption wall that leads to the design of multi-core processors. Second, there is a wide adoption of distributed processing technologies because
of the deployment of the Internet and consequently large-scale grid and cloud infrastructures. All these combined together make programming of computing infrastructure a very difficult challenge. Programmers have to face both parallel and distributed programming paradigms when designing an application and several software codes that are executed on various computing resources spread over the Internet within a grid or cloud-base infrastructure [1]. In the world of parallel computing MPI is the de facto standard for implementing programs on multiprocessors. To help with program development under a distributed computing environment a number of software tools have been developed. MPI [2] is chosen here for the parallelization. Many applications are “embarrassingly” parallel and require minimal performance out of MPI. These applications exploit coarse grain parallelism and communicate rarely.

However, program development for distributed memory parallel computers is time-consuming and error prone, as the programmer is forced to manage both parallelism and communication. Distributed systems can increase application performance by a significant amount and the incremental enhancement of a network-based concurrent computing environment is usually straightforward because of the availability of high bandwidth networks [3, 4]. The natural programming style under a distributed system is therefore the Multiple Instructions Multiple Data (MIMD) [3–6]. The basic idea is to split a program into a few smaller tasks and to allocate these tasks to several processors to be executed simultaneously. Thus the total execution time can be reduced to just a fraction of that of a uniprocessor computer.

Studying Bio-heat transfer in the human body has been a hot topic and is useful for designing clinical thermal treatment equipments, for accurately evaluating skin burn and for establishing thermal protections for various purposes. The Bio-heat transfer is the heat exchange that takes place between the blood vessels and the surrounding tissues. Monitoring the blood flow using the techniques has great advantage in the study of human physiological aspects. This requires a mathematical model which relates the heat transfer between the perfuse tissue and the blood. The theoretical analysis of heat transfer design has undergone a lot of research over the years, from the popular Penne’s bio-heat transfer equation proposed in 1948 to the latest one proposed by Deng and Liu [7]. Many of the Bio-heat transfer problem by Penne’s account for the ability of the tissue to remove heat by diffusion and perfusion of tissue by blood. Predictions of heat transport have been carried out by Chinmay [8] and Liu and Xu [9]. The major concern in the modeling of the Bio-heat is the accurate continuum representation of the heat transfer in the living tissue incorporating the effect of blood flow, due to the presence of two heat sources. Penne’s assumes that for heat transfer to take place there must be two heat sources, as in heat produced by the metabolism and the heat transfer from the blood flow surrounding tissue at each point of the forearm. Effects of thermal properties and geometrical dimensions on the skin burn injuries have been discussed. [9, 10] proposed a comparison of 1D and 2D programmed for predicting the state of skin burn. Liu et al. [11] used a finite difference method to solve the BHTP – a triple-layered skin structure composed of epidermis, dermis, and subcutaneous.
Dai and Zhang [12] developed a three-level unconditional stable finite difference scheme and used a domain decomposition strategy for solving the 1-D BHTP for the same three-layered skin structure.

In this paper, a method based on applying SOR to find the solution of 2DBHTP on a distributed computing environment using I_{aff} with MPI is implemented. We also assess parallel performance improvement in terms of speedup and efficiency. The rest of the paper is organized as follows: Section 2 introduces the model for the 2DBHTP. Section 3 gives the parallel implementation using I_{aff}. Section 4 introduces the results and discussion. Finally, a conclusion is included in Section 5.

2 Bio-heat model for 2-d

Modern clinical treatments and medicines such as cryosurgery, cryopreservation, cancer hyperthermia, and thermal disease diagnostics, require the understanding of thermal life phenomena and temperature behaviour in living tissues [13]. The well-known Penne’s equation and the energy balance for a control volume of tissue with volumetric blood flow and metabolism yields the general Bio-heat transfer equations. \( \rho, c_p \) are densities and specific heat of tissue, \( \omega_b c_b \) are blood perfusion rate and specific heat of blood, \( q_m^\prime \prime \) is the volumetric metabolic heat generation, \( U_a \) is the arterial temperature, \( U \) is the nodal temperature. The bio-heat problem is given as:

\[
\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} - \frac{\rho c_p}{\rho c_p} \frac{\partial}{\partial x^2} \left( \frac{\partial^2 U}{\partial y^2} \right) + \frac{\omega_b c_b}{\rho c_p} (U_a - U) + q_m^\prime \prime, \quad 0 \leq x \leq 1, \ 0 \leq y \leq 1, \ t > 0
\]

This further simplifies into the form:

\[
\frac{\partial U}{\partial t} = \frac{1}{\rho c_p} \frac{\partial^2 U}{\partial x^2} + \frac{1}{\rho c_p} \frac{\partial^2 U}{\partial y^2} - \frac{\omega_b c_b}{\rho c_p} U + \frac{\omega_b c_b}{\rho c_p} (U_a + \frac{q_m^\prime \prime}{\omega_b c_b})
\]

Assume \( q_m^\prime \prime \) to be constant, and denote \( U^\oplus = U_a + \frac{q_m^\prime \prime}{\omega_b c_b} \), \( b = \frac{\omega_b c_b}{\rho c_p} (> 0) \) and \( c = \frac{1}{\rho c_p} (> 0) \). We can obtain the simplified form of the 2-D Penne’s equation with the initial and boundary conditions given below:

\[
\frac{\partial U}{\partial t} = c \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) - bU + bU^\oplus,
\]

with initial condition
\[ U(x, y, 0) = f(x, y) \]  \hspace{1cm} (4)

and boundary conditions

\[
\begin{align*}
U(0, y, t) &= f_1(y, t), \quad U(1, y, t) = f_2(y, t) \\
U(x, 0, t) &= f_3(x, t), \quad U(x, 1, t) = f_4(x, t)
\end{align*}
\]  \hspace{1cm} (5)

When the explicit scheme is used, we write using the same finite-difference scheme:

\[
\frac{U^{n+1}_{i,j} - U^n_{i,j}}{\Delta t} = c\left(\frac{U^n_{i+1,j} - 2U^n_{i,j} + U^n_{i-1,j}}{\Delta x^2} + \frac{U^n_{i,j+1} - 2U^n_{i,j} + U^n_{i,j-1}}{\Delta y^2}\right) - bU^n_{i,j} + bU^\oplus
\]

The temperature of the node in the scheme formulation takes the form:

\[
U^{n+1}_{i,j} = c\frac{\Delta t}{\Delta x^2}\left(U^n_{i+1,j} + U^n_{i-1,j} + U^n_{i,j+1} + U^n_{i,j-1}\right) - b\Delta tU^n_{i,j} + b\Delta tU^\oplus
\]

\hspace{1cm} (6)

2.1 Stationary methods for 2-d bio-heat equation

If we use the central differences for both \( U_{xx} \) and \( U_{yy} \), and the forward difference for \( U_t \), into (3) and let \( \Delta x^2 = \Delta y^2 = \Delta^2 \) we have:

\[
U^{n+1}_{i,j} = U^n_{i,j} + \frac{c\Delta t}{\Delta^2}\left(U^n_{i+1,j} + U^n_{i-1,j} + U^n_{i,j+1} + U^n_{i,j-1} - 4U^n_{i,j}\right) - b\Delta tU^n_{i,j} + b\Delta tU^\oplus
\]

\hspace{1cm} (7)

Let \( \frac{c\Delta t}{\Delta^2} = Fo \), hence

\[
U^{n+1}_{i,j} = U^n_{i,j} + Fo\left(U^n_{i+1,j} + U^n_{i-1,j} + U^n_{i,j+1} + U^n_{i,j-1} - 4U^n_{i,j}\right) - b\Delta tU^n_{i,j} + b\Delta tU^\oplus
\]

It is stable in one spatial dimension (1-D) only if \( \Delta t / \Delta^2 \leq 1/2 \). In two dimensions (2-D) this becomes \( \Delta t / \Delta^2 \leq 1/4 \). Suppose we try to take the largest possible time step, and set \( c\Delta t = \Delta^2 / 4 \). Then eqn (7) becomes:
\[ U_{i,j}^{n+1} = \frac{1}{4} \left( U_{i+1,j}^n + U_{i-1,j}^n + U_{i,j+1}^n + U_{i,j-1}^n \right) - b\Delta t U_{i,j}^n + b\Delta t U_{\oplus} \]  

(8)

thus the algorithm consists of using the average of \( U \) at its four nearest neighbor points on the grid (plus contribution from the source). If we are proceeding along the rows, incrementing \( j \) for fixed \( i \), we have the formula eqn (8) as:

\[ U_{i,j}^{n+1} = \frac{1}{4} \left( U_{i+1,j}^n + U_{i-1,j}^n + U_{i,j+1}^n + U_{i,j-1}^n \right) - b\Delta t U_{i,j}^n + b\Delta t U_{\oplus}, \]

\[ i = 1\ldots,n, \quad j = 1\ldots,m \]  

(9)

This method is also slowly converging and only of theoretical interest, but some analysis of it will be instructive. If we have approximate values of the unknowns at each grid point, this equation can be used to generate new values. We call \( U^{(n)} \) the current values of the unknowns at each iteration \( k \) and \( U^{(n+1)} \) the value in the next iteration. We define a scalar \( \omega \) \((0 < \omega < 2)\) and apply eqn (9) to all interior points \((i, j)\). Hence, we have:

\[ U_{i,j+1}^{n+1} = \omega * GS + (1 - \omega)U_{i,j}^n \]  

(10)

where GS is the calculated value of the Gauss-Seidel method and \( \omega \) is omega with values ranging from \( 0 < \omega < 2 \).

### 3 Parallel implementation

#### 3.1 The cluster system

The implementation is done on a distributed computing environment (Armadillo Generation Cluster) consisting of 48 Intel Pentium at 1.73GHZ and 0.99GB RAM. Communication is through a fast Ethernet of 100 MB per second running Linux. The cluster performance has high memory bandwidth with a message passing supported by MPI [14]. We concentrate on basic message operations: blocking send, blocking receives, non-blocking send, and non-blocking receive. Note that MPI provides a rather comprehensive set of messaging operations. MPI primitive communication operation is the blocking send to blocking receive. At each time step we have to evaluate \( \nu^{n+1} \) values at 'Im' grid points, where 'I' is the number of grid points along the x axis. Suppose we are implementing this method on an \( R \times S \) mesh connected computer. Denote the workers by \( P_{i,j} : i = 1,2,\ldots,R \) and \( R < l, j = 1,2,\ldots,S \) and \( S < M \). Let \( L_i = \left[ \frac{1}{R} \right] \) and

\[ M_i = \left[ \frac{M}{S} \right] \]  

where \( [\cdot] \) is the smallest integer part. Divide the 'Im' grid points
into ‘RS’ groups so that each group contains at most \((L_1 + 1)(M_1 + 1)\) grid points and at least \(L_1 M_1\) grid points. Denote these groups by \(G_{i_1j_1} : i_1 = 1, 2, \ldots, R, \; j_1 = 1, 2, \ldots, S\). Design \(G_{i_1j_1}\) such that it contains the following grid points

\[
G_{i_1j_1} = \left\{ (X_{(i_1-1)+1}, Y_{(j_1-1)+1}) : i_1 = 1, 2, \ldots, L_1 \text{ or } L_i + 1 \right\}
\]

\[
j = 1, 2, \ldots, M_1 \text{ or } M + 1
\]

Assign the group \(G_{i_1j_1}\), to the workers \(P_{i_1j_1} : i_1 = 1, 2, \ldots, R, \; j_1 = 1, 2, \ldots, S\). Each worker computes its assigned group \(v_{i,j}^{n+1}\) values in the required number of sweeps. At the \((p + 1/2)\)th sweep the workers compute \(v_{i,j}^{(p+1/2)th}\) values of its assigned groups. For the \((p + 1/2)\)th level the worker \(P_{i_1j_1}\) requires one value from the worker \(P_{i_1-l=j_i}\) or \(P_{i_1+l=j_i}\), worker. In the \((p + 1/2)\)th level the communication between the workers is done row-wise. After communication between the workers is completed then each worker \(P_j\) computes the \(v_{i,j}^{p+1/2}\) values. For the \((p + 1)\)th sweep each worker \(P_{i_1j_1}\) requires one value from the \(P_{i_1-l=j_i}\) or \(P_{i_1+l=j_i}\) worker. Then each worker computes the values \(v_{i,j}^{(p+1)th}\) of its assigned group.

**Statements need to be inserted to select which portions of the code will be executed by each processor. We focus our evaluation on MPI because it serves as an important foundation for a large group of applications.**

### 3.1.1 The \(I_{eff}\)

With reference to [15], the \(I_{eff}\) is discussed. The platform introduces each task going through three phrases during execution of a parameter-sweep application:

1. an initialization phase. The duration of this phase is equal to \(t_{init}\). This phase includes the overhead incurred by the master to initiate a data transfer to a slave,
2. a computational phase. The duration of this phase is equal to \(t_{comp}\). Any additional overhead related to the reception of input files by a worker node is also included in this phase and
3. a completion phase, where the output file is sent back to the master and the master task is completed. The duration of this phase is equal to \(t_{end}\). Therefore, the initialization phase of one slave can occur concurrently with the completion phase of another worker node. However, the total execution time of a task is equal to

\[
t_{\text{total}} = t_{\text{init}} + t_{\text{comp}} + t_{\text{end}}
\]

One processor is the master and the other processors are workers. \(P_{eff}\) is the effective number of processors needed to run an application with no idle periods on any worker processor. A processor may have idle periods if:

\[
t_{\text{comp}}' < (P - 1)t_{\text{init}}
\]
\( P_{\text{eff}} \) is then given by the following equation:

\[
P_{\text{eff}} = \left\lceil \frac{t'_{\text{comp}}}{t'_{\text{init}}} + 1 \right\rceil
\]

(13)

the total number of tasks to be executed on a processor is at most

\[
M = \left[ \frac{T}{P} \right]
\]

(14)

for a platform with \( P_{\text{eff}} \) processors, the upper bound for the total execution time (makespan) will be

\[
\left\lceil t_{\text{makespan}} \right\rceil = M(t_{\text{init}} + t'_{\text{comp}}) + (P - 1)t_{\text{init}}
\]

(15)

the second term in the right hand side of eqn (15) shows the time needed to start the first \( P - 1 \) tasks in the other \( P - 1 \) processors. If we have a platform where the number of processors is larger than \( P_{\text{eff}} \) the overall makespan is dominated by communication times between the master and the workers. We then have

\[
\left\lceil t_{\text{makespan}} \right\rceil = MP_{\text{init}} + t'_{\text{comp}}
\]

(16)

the set of eqn (12)–(14) will be considered in subsequent sections of this paper. It is worth noting that eqn (15) is valid when workers are constantly busy, either performing computation or communication. Eqn (16) is applicable when workers have idle periods, i.e., are not performing either computation or communication. eqn (12) occurs mainly in two cases:

For very large platforms (\( P \) large).

For applications with small \( \frac{t_{\text{comp}}}{t_{\text{init}}} \) ratio, such as fine-grain applications.

In order to measure the degree of affinity of a set of tasks concerning their input files, the concept of input file affinity is introduced. Given a set of \( E \) of tasks, composed of \( R \) tasks, \( E = \{T_1, T_2, \ldots, T_R\} \), and the set \( Q \) of the \( X \) input files needed by the tasks belonging to group \( E \), \( Q = \{f_1, f_2, \ldots, f_X\} \).

4 Results and discussion

4.1 Benchmark problem

We implement the SOR technique on the 2DBHTP. We assume a platform composed of variable number of heterogeneous processors. The solution domain was divided into rectangular blocks. The experiment is demonstrated on meshes of 100×100, 200×200 and 400×400, respectively. Tables 2–4 show the various performance timing.
\[
\frac{\partial U}{\partial t} = c \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) - bU + bU^\oplus, \quad (17)
\]

The boundary condition and initial condition posed are:
\[
\begin{align*}
U(0, y, t) &= 0 \\
U(1, y, t) &= 0 \\
U(x, 0, t) &= 0 \\
U(x, 1, t) &= 0 \\
t &\geq 0
\end{align*}
\quad (18)
\]

The cell size was chosen as \( \Delta x = \Delta y \). The values of the physical properties in our test cases are chosen to be \( \rho = 1000 \text{kg/m}^3 \), \( c = c_b = 4200 \text{J/kg}^0 \text{c} \), \( w_b = 0.5 \text{kg/m}^3 \), temperature is set to be \( U_0 = 12^0 \text{c} \). Table 1 provides a comparison of the accuracy of the methods under consideration in terms of absolute error.

Table 1: Sequential result for 2DBHTP with various schemes.

<table>
<thead>
<tr>
<th>Method</th>
<th>GS</th>
<th>SOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av. Abs. Err.</td>
<td>10.5 \times 10^{-4}</td>
<td>8.7 \times 10^{-4}</td>
</tr>
<tr>
<td>RMS</td>
<td>5.6 \times 10^{-4}</td>
<td>4.3 \times 10^{-7}</td>
</tr>
<tr>
<td>\Delta x</td>
<td>1 \times 10^{-1}</td>
<td>1 \times 10^{-1}</td>
</tr>
<tr>
<td>\Delta t</td>
<td>2 \times 10^{-4}</td>
<td>2 \times 10^{-4}</td>
</tr>
<tr>
<td>\lambda</td>
<td>5 \times 10^{-1}</td>
<td>5 \times 10^{-1}</td>
</tr>
<tr>
<td>\tau</td>
<td>1.8 \times 10^{-3}</td>
<td>1.8 \times 10^{-3}</td>
</tr>
<tr>
<td>eps</td>
<td>1 \times 10^{-4}</td>
<td>1 \times 10^{-4}</td>
</tr>
</tbody>
</table>

The experiment is demonstrated on meshes of 200×200 and 300×300 for block sizes of 100 and 200, respectively, both for MPI. The tables show the various performance timings.

4.1.1 Parallel efficiency

The speed-up and efficiency obtained for various sizes of 200×200 to 300×300 are for various numbers of sub-domains; from \( B = 100 \) to 200 as listed in Tables 1–4. In these tables we listed the wall (elapsed) time for the master task, \( T_W \), the master CPU time, \( T_M \), the average worker computational time, \( T_{SC} \) and the average worker data communication time \( T_{SD} \) all in seconds. The speed-up and efficiency versus the number of processors are shown in Fig. 1 and Fig. 2 respectively, with block number \( B \) as a parameter. The results show that the parallel efficiency increases with increasing grid size for a given block number.
Table 2: A mesh of 300×300, with $B = 50$ blocks and $Niter = 100$ for MPI.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>N</th>
<th>$T_w$</th>
<th>$T_m$</th>
<th>$T_{sd}$</th>
<th>$T_{sc}$</th>
<th>MPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$T$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$S_{par}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$E_{par}$</td>
</tr>
<tr>
<td>1</td>
<td>945</td>
<td>46</td>
<td>27</td>
<td>819</td>
<td>593</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>652</td>
<td>45</td>
<td>25</td>
<td>572</td>
<td>364.3</td>
<td>1.628</td>
</tr>
<tr>
<td>4</td>
<td>617</td>
<td>45</td>
<td>25</td>
<td>289</td>
<td>255.21</td>
<td>2.324</td>
</tr>
<tr>
<td>8</td>
<td>524</td>
<td>45</td>
<td>25</td>
<td>229</td>
<td>208.47</td>
<td>2.845</td>
</tr>
<tr>
<td>SOR</td>
<td>16</td>
<td>382</td>
<td>45</td>
<td>25</td>
<td>152</td>
<td>183.0</td>
</tr>
<tr>
<td>20</td>
<td>262</td>
<td>45</td>
<td>25</td>
<td>107</td>
<td>143.68</td>
<td>4.128</td>
</tr>
<tr>
<td>24</td>
<td>195</td>
<td>45</td>
<td>25</td>
<td>89</td>
<td>122.92</td>
<td>4.825</td>
</tr>
<tr>
<td>30</td>
<td>133</td>
<td>45</td>
<td>25</td>
<td>31</td>
<td>111.46</td>
<td>5.321</td>
</tr>
<tr>
<td>38</td>
<td>92</td>
<td>45</td>
<td>25</td>
<td>20</td>
<td>101.12</td>
<td>5.924</td>
</tr>
<tr>
<td>48</td>
<td>83</td>
<td>45</td>
<td>25</td>
<td>10</td>
<td>94.97</td>
<td>6.245</td>
</tr>
</tbody>
</table>

Table 3: A mesh of 300×300, with $B = 100$ blocks and $Niter = 100$ for MPI.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>N</th>
<th>$T_w$</th>
<th>$T_m$</th>
<th>$T_{sd}$</th>
<th>$T_{sc}$</th>
<th>MPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$T$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$S_{par}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$E_{par}$</td>
</tr>
<tr>
<td>1</td>
<td>1326</td>
<td>68</td>
<td>30</td>
<td>1109</td>
<td>773.7</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>921</td>
<td>66</td>
<td>28</td>
<td>984</td>
<td>440</td>
<td>1.758</td>
</tr>
<tr>
<td>4</td>
<td>875</td>
<td>66</td>
<td>28</td>
<td>752</td>
<td>285.14</td>
<td>2.713</td>
</tr>
<tr>
<td>8</td>
<td>704</td>
<td>66</td>
<td>28</td>
<td>561</td>
<td>235.35</td>
<td>3.287</td>
</tr>
<tr>
<td>SOR</td>
<td>16</td>
<td>562</td>
<td>66</td>
<td>28</td>
<td>329</td>
<td>194.51</td>
</tr>
<tr>
<td>20</td>
<td>431</td>
<td>66</td>
<td>28</td>
<td>284</td>
<td>161.28</td>
<td>4.797</td>
</tr>
<tr>
<td>24</td>
<td>349</td>
<td>66</td>
<td>28</td>
<td>201</td>
<td>138.74</td>
<td>5.577</td>
</tr>
<tr>
<td>30</td>
<td>298</td>
<td>66</td>
<td>28</td>
<td>162</td>
<td>114.49</td>
<td>6.758</td>
</tr>
<tr>
<td>38</td>
<td>211</td>
<td>66</td>
<td>28</td>
<td>144</td>
<td>99.25</td>
<td>7.795</td>
</tr>
<tr>
<td>48</td>
<td>184</td>
<td>66</td>
<td>28</td>
<td>94</td>
<td>95.12</td>
<td>8.134</td>
</tr>
</tbody>
</table>

Table 4: A mesh of 300×300, with $B = 200$ blocks and $Niter = 100$ for MPI.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>N</th>
<th>$T_w$</th>
<th>$T_m$</th>
<th>$T_{sd}$</th>
<th>$T_{sc}$</th>
<th>MPI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$T$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$S_{par}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$E_{par}$</td>
</tr>
<tr>
<td>1</td>
<td>2431</td>
<td>97</td>
<td>62</td>
<td>1499</td>
<td>1144</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>1764</td>
<td>95</td>
<td>60</td>
<td>1103</td>
<td>644.03</td>
<td>1.776</td>
</tr>
<tr>
<td>4</td>
<td>987</td>
<td>95</td>
<td>60</td>
<td>865</td>
<td>387.76</td>
<td>2.950</td>
</tr>
<tr>
<td>8</td>
<td>834</td>
<td>95</td>
<td>60</td>
<td>679</td>
<td>347.57</td>
<td>3.291</td>
</tr>
<tr>
<td>SOR</td>
<td>16</td>
<td>698</td>
<td>95</td>
<td>60</td>
<td>482</td>
<td>234.23</td>
</tr>
<tr>
<td>20</td>
<td>573</td>
<td>95</td>
<td>60</td>
<td>321</td>
<td>214.14</td>
<td>3.342</td>
</tr>
<tr>
<td>24</td>
<td>432</td>
<td>95</td>
<td>60</td>
<td>294</td>
<td>201.76</td>
<td>5.670</td>
</tr>
<tr>
<td>30</td>
<td>359</td>
<td>95</td>
<td>60</td>
<td>210</td>
<td>165.13</td>
<td>6.928</td>
</tr>
<tr>
<td>38</td>
<td>286</td>
<td>95</td>
<td>60</td>
<td>194</td>
<td>144.46</td>
<td>7.919</td>
</tr>
<tr>
<td>48</td>
<td>231</td>
<td>95</td>
<td>60</td>
<td>121</td>
<td>134.70</td>
<td>8.493</td>
</tr>
</tbody>
</table>
using MPI and decreases with the increasing block number for given grid size. In Tables 1–4, the master time $T_M$ is constant when the number of processors increases for a given grid size and number of sub-domains. The master program is responsible for (1) sending updated variables to worker ($T_1$), (2) assigning task to worker ($T_2$), (3) waiting for the worker to execute tasks ($T_3$), and (4) receiving the results ($T_4$).

![Figure 1: Speed-up versus the number of workers for various block sizes. Mesh 300×300 MPI.](image)

![Figure 2: Parallel efficiency versus the number of workers for various block sizes. Mesh 300×300 MPI.](image)

As the total number of processors increase, the bottleneck of parallel computers appears and the global reduction consumes a large part of time.

## 5 Conclusions

A thorough analysis of the 2DBHTP using SOR iterative technique and $I_{aff}$ on master-worker paradigm was made, which shows that significantly better efficiency can be obtained. The efficiency is accomplished by overlapping computation of the sequential and parallel phases of the algorithm. The $I_{aff}$ on master-worker paradigm has been emphasized. Experiments with 2DBHTP revealed that the rate of convergence decreases as the number of processor increases for various mesh sizes. Our goal is shown on the interleaving of communication/computation that is at the core of the MPI performance, and the proposed methods combine elements of numerical stability and parallel algorithm design without increasing computational costs. Here, MPI is used for
its performance and portability. On the basis of the current parallelization strategy, more sophisticated models can be attacked efficiently.

References