Parallel Gaussian Elimination of a Block Tridiagonal Matrix Using Multiple Microcomputers

Richard A. Blech
Lewis Research Center
Cleveland, Ohio
Summary

The solution of a block tridiagonal matrix using parallel processing is demonstrated in this report. The multiprocessor system which obtained the results and the software environment used to program that system are described. Theoretical partitioning and resource allocation for the Gaussian elimination method used to solve the matrix are discussed. The results obtained from running one-, two-, and three-processor versions of the block tridiagonal solver are presented. The PASCAL source code for these solvers is given in the appendix, and it may be transportable to other shared-memory parallel processors, provided that the synchronization routines are reproduced on the target system.

Introduction

Many computationally intensive problems can benefit from the use of parallel processing. One such problem, common to many fluid mechanics and structural dynamics applications, is the solution of large matrix equations. Because of the differencing techniques used in solving the partial differential equations that describe fluids and structures systems, the resulting matrices often exhibit a block tridiagonal structure. The block tridiagonal matrix requires much less computation to solve than a full $N \times N$ matrix. A full matrix requires approximately $N^3$ operations to solve; a block tridiagonal matrix requires approximately $N$ operations.

Although the block tridiagonal structure significantly reduces computational effort, considerable time is still spent in the matrix solution. This is especially true in many iterative linearization techniques, such as Newton-Raphson, where a full matrix solution is required for every iteration. Because of this, other parallel processing techniques which can further reduce the amount of computation required to arrive at a solution should be investigated.

This paper presents the solution of a block tridiagonal matrix on a parallel processor. The block tridiagonal equations analyzed were taken from a transient rotor dynamics simulation program (ref. 1). In this program, Gaussian elimination is used to solve the matrix.

The real-time multiprocessor simulator (RTMPS) was used to solve these equations in parallel (refs. 2 to 5). The RTMPS is a parallel processor designed to do real-time simulation of dynamic systems. The hardware consists of dual busses with processors on each bus. A dual-port memory provides communication between the two busses by connecting processors on one bus to the processors on the other bus. Considerable software support is provided for one-dimensional scalar problems by a real-time multiprocessor language (RTML) and a real-time multiprocessor operating system (RTMPOS).

The potential of parallel processing for improving the performance of linear algebra routines has prompted a significant amount of research (refs. 6 to 8). Also, a significant amount of literature exists on the use of vector processors for linear algebra. Since vectorization of code involves the identification of the lowest level of parallelism (e.g., operation level parallelism), the principles behind both areas of research are very similar. Because of the high percentage of nested loops in linear algebra code, the ideal architecture for most linear algebra applications would consist of multiple vector processors.

This paper presents the application of parallel processing using one particular architecture (RTMPS) to one algorithm (Gaussian elimination). This combination, however, may not be the best approach to the problem. As mentioned previously, there are other architectures and algorithms that may be better suited for this application. The RTMPS system was used for this study because it was the only parallel processing hardware conveniently available. The intent of this study is to identify some practical aspects of implementing a commonly used algorithm on a parallel processor. The investigation of other architectures and algorithms will be the focus of future research.

Problem Description

The structure of the block tridiagonal matrix is shown in figure 1. Each block row, except the first and last, consists of three $M$ by $M$ blocks. There are $N$ block rows total, including the first and last. If this matrix is called $A$, then the general problem is to find the solution to the system of equations

$$Ax = b$$

where $x$ and $b$ are vectors, $N$ elements in length.

A common method for solving this system is to perform a forward elimination of all coefficients below the diagonal and then a back substitution to solve for the vector $x$. This procedure, called Gaussian elimination, is illustrated in the following example for a 3 by 3 matrix.
Gaussian elimination is efficiently performed on a block tridiagonal matrix by applying a partial elimination process to four adjacent blocks at a time (fig. 2). The process begins with blocks 2 and 3 from the first block row and blocks 1 and 2 from the next block row. The Gaussian elimination procedure is applied to the matrix determined by these four blocks. However, the process stops once block 2 in the first block row is made upper right triangular. As a result of this process, block 1 in block row 2 is zero at this time.

This process is then repeated on the next group of blocks starting in the next block row and continuing for the whole matrix, moving the four-block template down through the matrix one block row at a time. Thus, by repeating a partial $2M$ by $2M$ Gaussian elimination $N$ times, the tridiagonal matrix is transformed to upper right triangular form.

After the matrix has been transformed to an upper right triangular matrix, the result vector $\mathbf{x}$ can be solved by using back substitution starting from the bottom of the matrix. This is done by solving for the last element of the result vector $x_n$, and substituting that value into the equation for the second last element of $\mathbf{x}$ (next row up). Now, two values of the result vector are available for substitution into the equation for the third last element. The procedure is repeated one row at a time, proceeding upward through the matrix until all elements of $\mathbf{x}$ have been solved. The following equation illustrates the back-substitution process for the example problem.
An approach for parallelizing the Gaussian elimination procedure was developed by examining the data flow of the problem. A data flow diagram suggests that, if several computers are available, multiple operations could be done concurrently. The five stages of the computations are bracketed on the right side of figure 3. Within each stage, each vertical operation stream can be done in parallel. Stages 2 and 4 have streams of two operations each, while all other stages have streams of only one operation.

Most parallelism exists in the second stage, where eight operation streams can be done in parallel. If eight processors were available, the 16 operations of stage 2 could be done in a net count of two operations. Stage 1 would require four processors and could be done in a net count of one operation. Stages 3 and 4 would require three processors and could be done in net counts of one and two operations, respectively. Finally, stage 5 requires two processors and could be done in one operation. The minimum count for execution of the entire problem is the critical path. The critical path is the longest of the parallel operation streams in the data flow graph. In this example, the critical path is seven operations. Since each stage is done serially, only the maximum number of processors in any stage would be required (eight in this case).

The data flow diagram for the back-substitution process is shown in figure 4. There are two stages, and the critical path is four operations. The maximum number of processors required is two.

The block tridiagonal matrix solver used in the rotor dynamics application consists of 30 block rows. Each block is 4 by 4. Thus, $N = 30$ and $M = 4$ for this application.

**Partitioning Approach**

An approach for parallelizing the Gaussian elimination procedure was developed by examining the data flow of the problem. A data flow diagram for the 3-by-3-matrix example is shown in figure 3. The circles represent mathematical operations, and the interconnections show the flow of data between calculations. For the 3-by-3-matrix example, 31 operations must be performed. A single computer can only execute these operations one at a time. The data flow diagram suggests that, if several computers are available, multiple operations could be done concurrently. The five stages of the computations are bracketed on the right side of figure 3. Within each stage, each vertical operation stream can be done in parallel. Stages 2 and 4 have streams of two operations each, while all other stages have streams of only one operation.

Most parallelism exists in the second stage, where eight operation streams can be done in parallel. If eight processors were available, the 16 operations of stage 2 could be done in a net count of two operations. Stage 1 would require four processors and could be done in a net count of one operation. Stages 3 and 4 would require three processors and could be done in net counts of one and two operations, respectively. Finally, stage 5 requires two processors and could be done in one operation. The minimum count for execution of the entire problem is the critical path. The critical path is the longest of the parallel operation streams in the data flow graph. In this example, the critical path is seven operations. Since each stage is done serially, only the maximum number of processors in any stage would be required (eight in this case).

The data flow diagram for the back-substitution process is shown in figure 4. There are two stages, and the critical path is four operations. The maximum number of processors required is two.

The block tridiagonal matrix solver used in the rotor dynamics application consists of 30 block rows. Each block is 4 by 4. Thus, $N = 30$ and $M = 4$ for this application.
The solution of the block tridiagonal matrix contains the same parallelism described for the 3-by-3-matrix example. In the solution process, a partial elimination is performed on a 2M by 2M system N times. The maximum number of processors required would be a function of M. The critical path would be $N$ multiplied by the critical path operation count for partial Gaussian elimination plus the critical path count for the back substitution. The data flow diagram would follow the same pattern as that of the 3 by 3 matrix, only the length and width would vary as the size of the matrix. A detailed analysis is given in the Theoretical Speedup Analysis section.

A PASCAL-coded version of the single-processor matrix solver used in the rotor dynamics simulation is given in the appendix. This is a direct PASCAL translation of the FORTRAN code used in the simulation. The procedures GETINF, IDATA, and IDATF are related to I/O on the unique hardware used for this study. The purpose of these procedures is described later in this report. The parallel structures discussed previously can be seen in the main body of the code. There are two main loops in the program. The outer loop (IB) cycles through the block rows of the matrix. The next loop (IP) does the partial Gaussian elimination on the 8 by 8 submatrix composed of blocks 2 and 3 in the current block row and blocks 1 and 2 in the next block row. Within the IP loop are six smaller loops which essentially perform the operations diagramed in the data flow graph in figure 3. The first two loops perform the divide operations, and the next four loops perform the multiply and subtract operations. As shown in the data flow graph, all divides can be done in parallel followed by all multiply and subtractions being done in parallel. This process is repeated for all four IP iterations.

The code for the back-substitution process is next in the program. Since the original code was not written with parallel processing in mind, there are no operations which can be done in parallel while using the code shown. Each result vector element is found by solving one row at a time. Each iteration of the outermost loop (IBI) depends on results from the previous iteration. The same is true for the next level loop (II). The innermost loops within the II loop are recursive in nature (the calculation of a variable depends on itself from a previous iteration) and, therefore, cannot be done in parallel.

The data flow graph for the back substitution, however, shows that parallel operations can be done. Figure 4 shows that once an element of the result vector has been calculated, it can be used to calculate parts of proceeding elements. Thus, partial sums of other result vector elements can be computed in parallel. This algorithm, called the column sweep, is described in reference 6. The column sweep algorithm requires a different coding approach than that used in the rotor dynamics version of the back-substitution process. A new version was coded and used for the two- and three-processor matrix solvers discussed later in this paper. The use of the column sweep algorithm exemplifies the type of analysis required for selecting an algorithm to run on a parallel processor.

**Theoretical Speedup Analysis**

The theoretical speedup for the parallel Gaussian elimination algorithm is computed by dividing the operation count for the serial version by the net operation count for the parallel version. An operation is one of the basic floating-point math operations: add, subtract, multiply, and divide.

Table I shows the determination of the operation count for the serial algorithm for one IB iteration of the forward elimination procedure and one I iteration of the back substitution. The table assumes a 4 by 4 block size. One IB iteration consists of four IP iterations, and the operation count for each IP iteration depends on the value of IP. For a matrix of 30 block rows, the operation count (OPS) would be:

$$OPS = 30 \text{(number of operations per IB)} + 30 \text{(number of operations per IP)}$$

$$= 30(370) + 30(44) = 12420 \text{ operations}$$

To simplify the analysis, it is assumed that the last block row is a full 8 by 8 matrix, although it is actually 4 by 8.

The operation count for an N block row, M-by-M-block tridiagonal matrix would be:

$$OPS = N \left\{ \sum_{i=1}^{M} [(2M + 2 - i) + 2(2M + 2 - i)(2M - i)] + 2(M + i - 1) \right\}$$

$$= N \left[ M(4M + 7)(7M - 1) \right] / 6$$

[Figure 4.—Data flow diagram for back substitution (3 by 3 matrix).]
TABLE I.—DETERMINATION OF OPERATION COUNT

(a) Gaussian elimination

<table>
<thead>
<tr>
<th>Loop, IP</th>
<th>Operations</th>
<th>Total number of operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Divide</td>
<td>Multiply</td>
<td>Subtract</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>63</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(b) Back substitution

<table>
<thead>
<tr>
<th>Loop, I</th>
<th>Operations</th>
<th>Total number of operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiply</td>
<td>Subtract</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The data flow graphs in figures 3 and 4 suggest that a number of operations can be done in parallel. For the forward elimination process, the total number of operations which can be done in parallel is a function of the iteration index IP. Table II summarizes the maximum number of operations which can be performed in parallel as a function of IP. The last column shows the net operation count for each IP iteration (three) if there are enough processors available to match the number of operations that can be done in parallel. Each IP iteration consists of a parallel divide cycle, followed by a parallel multiply and subtract cycle. The net operation count is one for the divide cycle and two for the multiply and subtract cycle. Each IP iteration has three operations. As IP increases, the number of processors that can be used decreases.

Table II also shows the maximum number of parallel operations for each I iteration of the back-substitution process. Again, the net operation count is shown for the case where the number of processors matches the number of parallel operations. Based on the total operation count for the fully parallel forward elimination and back-substitution processes (assuming 30 block rows), the total operation count would be

\[ \text{OPS} = 30(4 \times 3 \text{ operations}) + 30(4 \times 2 \text{ operations}) \]
\[ = 600 \text{ operations} = 30 \times 4 \times 5 \]

or, in general,

\[ \text{OPS} = N \times M \times 5 \text{ operations} \]

For the matrix used in this study, the theoretical speedup (S) would be

\[ S = \frac{12420}{600} = 20.7 \]

and, in general,

\[ S = \frac{\sum_{i=1}^{M} (2M + 2 - i) + 2(2M + 2 - i)(2M - i) + 2(M + i - 1)}{5M} \]
\[ = \frac{(4M + 7)(7M - 1)}{30} \]

for a N block row, M-by-M-block matrix.

The theoretical speedup would be achieved if the maximum number of processors (63 as determined from table II) are available to perform the computations. Any overhead due to inefficient resource allocation (discussed in the next section) or communication between processors has been ignored. This simplification is made because of the difficulty in estimating the time required for such overhead. The theoretical speedup is useful only as an upper limit to determine if parallel processing can potentially benefit an application.

Determining the theoretical speedup is more complicated when less than the maximum number of processors is available. The speedup will also be a function of the way the parallel computations are allocated to the processors. For example, if there are four parallel operations and three processors, the net operation count would be two because the fourth operation must be done in serial on one of the three processors. The theoretical speedup for the three-processor matrix solver was determined to be 2.9 based on the best resource allocation possible.

TABLE II.—DETERMINATION OF PARALLEL OPERATION COUNT

<table>
<thead>
<tr>
<th>(a) Forward elimination</th>
<th>(b) Back substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loop, II</td>
<td>Number of processors</td>
</tr>
<tr>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
inefficiency is unavoidable. In others, a "packing" algorithm can be used to assign the parallel tasks to the minimum number of processors necessary. If the processor resources do not match the number of parallel tasks, then a packing algorithm is a necessity. Ideally, an automated procedure would assign the parallel computations to the available processors and generate the appropriate load modules (to execute on the processors). Such a procedure, unfortunately, was not available for this study.

A technique for allocating the parallel operations of the matrix solver to the appropriate processors was necessary. One, called the loop-unrolling technique, would require decomposing the loops into individual equations. For example, consider the following loop:

```
FOR I = 1 to 5 DO
    FOR J = I to 5 DO
        A(I,J) = B(I,J) * C(I,J);
```

The doubly nested loop can be decomposed into 15 equations, and each of these could be executed in parallel. Suppose that only three processors were available for the solution of these equations. One method of allocating the equations to the processors would be to write all 15 equations and allocate 5 equations to each processor. Although this appears easy for the given example, it can be tedious if there are hundreds of thousands of equations. Another, less tedious, method would be to use the following code segment on each of the processors:

```
FOR I = 1 to 5 DO
    BEGIN
        J = (I - 1) + PID;
        WHILE J <= 5 DO
            BEGIN
                A(I,J) = B(I,J) * C(I,J);
                J = J + NPROC;
            END;
    END;
```

where PID is the processor identification number (in this case 1, 2, or 3) and NPROC is the number of processors (three for this example). In this method, called iteration allocation, each processor performs only the iterations which are assigned to it. The preceding example results in the allocation of computations as follows:

<table>
<thead>
<tr>
<th></th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(1,1)</td>
<td>A(1,2)</td>
<td>A(1,3)</td>
<td></td>
</tr>
<tr>
<td>A(1,4)</td>
<td>A(1,5)</td>
<td>A(2,4)</td>
<td></td>
</tr>
<tr>
<td>A(2,2)</td>
<td>A(2,3)</td>
<td>A(3,5)</td>
<td></td>
</tr>
<tr>
<td>A(2,5)</td>
<td>A(3,4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(3,3)</td>
<td>A(4,5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(4,4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A(5,5)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total OPS: 7 5 3

With 3 processors and this allocation method, the original 15 operations could be done in the equivalent of 7 operations. Although this method is less efficient than writing 15 separate equations, it is less tedious. In fact, by adding the following lines of code before the J = (I - 1) + PID line in processors 1 and 3, the allocation can be improved:

```
IF I > 3 THEN PID = 3   IF I > 3 THEN PID = 1
ELSE PID = 1;           ELSE PID = 3;
```

This reallocates the A(4,4) and A(5,5) computations from processor 1 to processor 3. Now each processor solves five equations for a net count of five operations. However, this analysis ignores the overhead of the added control statements. Thus, iteration allocation is still less efficient than the loop-unrolling technique. But for large loops, iteration allocation is preferable since it is less tedious.

The number of processors available is a critical factor in considering which method to use. If the number of processors approaches the number of parallel tasks, then the iteration-allocation method essentially approaches the loop-unrolling technique (in the amount of work necessary to generate a parallel program). In general, if the number of parallel tasks is much greater than the number of processors, iteration allocation is preferable to loop-unrolling. This was the case for the parallel block tridiagonal solver described in this report, which made iteration allocation the method of choice.

Parallel Processing Hardware Description

The parallel processing hardware system used to run the block tridiagonal solver is a subset of the real-time multiprocessor simulator (RTMPS) described in reference 2. Figure 5 is a block diagram of the actual hardware used. The separate processors on the real-time bus are Motorola VM04 microcomputers, rather than the VM02 microcomputers used on the original RTMPS. The RTX channel linking the interactive and real-time busses still uses VM02 microcomputers. In the current configuration, a maximum of three VM04
processors can be resident on the real-time bus. Expansion to two additional processors is possible with the existing card cage, but was not done for this study.

Figure 6 shows the architecture of the VM04 microcomputer. A dedicated memory bus connects a processor board to a main memory board. Both the processor board and the memory board have separate system bus interfaces. A high-speed cache memory on the processor board reduces memory access times for frequently referenced memory locations. The caching of main memory is handled by hardware and is transparent to the user when only a single processor is used.

Because of the cache memory, extra care must be taken when programming multiprocessor systems. A processor can have instances of “stale data” in its cache if another processor communicates with it through shared memory over the system bus. To avoid this problem, processors must disable their cache memories before accessing shared memory. One method, although time consuming, is to call a procedure to disable the cache each time a program requires access to shared memory. Another method would be to disable the cache entirely; however, there are many local-memory accesses which would lose the benefit of the faster cache memory. Fortunately, the VM04 contains a control register which allows the user to enable or disable caching of memory accesses that occur over the system bus. This register can be set once, and cache memory can be disabled for all shared-memory references (via the system bus), while local-memory references are still cached.

Software Environment

The existing RTMPL was designed to efficiently handle one-dimensional mathematical models. All arithmetic is done in scaled-fractions, and indexed variables (e.g., arrays) are not supported. An alternative language was needed to allow convenient programming of the block tridiagonal solver on the RTMPS hardware. To fill this need, a method was devised to allow PASCAL programs to be called from an RTMPL program. The the solver could be coded in the PASCAL language, a structured language with floating-point and indexed variable support.

Running the PASCAL program as a subroutine under RTMPL maintains compatibility with the RTMPOS. This is important because RTMPL generates a data base that RTMPOS uses to load and execute the parallel processing programs. Changes were made to RTMPOS to allow recognition of the floating-point data type. Thus, many interactive features provided by RTMPOS for scaled-fraction programs could also be applied to floating-point programs.

An RTMPL macro was written to transfer control from RTMPL to PASCAL. A new PASCAL initialization routine (ref. 9) was written to save any necessary RTMPL registers, execute the PASCAL program, restore the RTMPL registers, and return to the RTMPL program. RTMPL variables were used as buffers to transfer information from the PASCAL program to the RTMPL program. Special procedures were written to do the transfers. This represents one of the disadvantages of the RTMPL-PASCAL approach: Neither program recognizes the variables of the other. In order to output any results from the PASCAL program to RTMPOS, data must explicitly be transferred from a PASCAL variable to an RTMPL variable. This inconvenience can translate into high overhead if data is output frequently from the PASCAL program. Fortunately, for the block tridiagonal solver, the only output required was at the end of the program.

The automated data-transfer setup feature of the RTMPL cannot be used with the RTMPL-PASCAL approach. All data transfers must be done from within the PASCAL program. One method for transferring data from PASCAL is to call a procedure to do the transfer. However, if there is frequent data transfer in the program, the overhead of the procedure call can significantly reduce the transfer speed. A better method is to exploit the way that PASCAL handles variables. Variables declared in the main PASCAL program are global variables; variables declared from within a procedure are local to that procedure. Global variables are shared by the main program and all procedures. This suggests that a shared-memory multiprocessor environment can be implemented by using the global variable area as the shared memory. The advantage of a shared-memory approach is that data can be transferred implicitly between processors by a simple memory reference instruction. The need for a procedure call to transfer data is eliminated, thus, reducing overhead.

Figure 7 shows how the PASCAL shared-memory approach is implemented for two processors connected by a bus. The PASCAL compiler maintains two registers for variable storage. The first (A5) points to the base of the global variables. The other register (A6) points to the base of the local variable area. If both processors (P1,P2) have dual-ported memory, part of the memory of P1 can be shared with P2. The PASCAL programs for P1 and P2 would have the shared-memory variables declared first. The program code body would be contained in a procedure call, with any local variables declared within the procedure. Then the main program would merely call this procedure. The structure of the PASCAL program for both processors would be as follows:

```
```
**Discussion of Results**

The block tridiagonal solver was run on the RTMPS system with one, two, and three processors. The PASCAL code for each of the cases is contained in the appendix of this report.

The matrix notation used for the rotor dynamics problem is retained in this code. Array B in the PASCAL code is the matrix of coefficients, the array C is the right-side vector, and the array DU is the result vector. The first VAR declaration is the global, or shared-memory area. A multiply indexed array is used for the block tridiagonal matrix. The first two indices (from left to right) are the row and column indices within a block. The next index is the block row index, and the last index is the block index (1, left; 2, middle; 3, right). The vectors DU and C are doubly indexed arrays: The first index indicates element within the current block row, and the second index is the block row index. Although the use of multiple indices simplifies the programming procedure, it is very costly in computation time.

The code for the single-processor solver is a direct PASCAL translation of the FORTRAN code used in the rotor dynamics problem. Procedure GETINF is used to send information about the variables (in this case, the result vector) to the RTMPS control processor. Calling procedure GETINF triggers a read advisory on the control processor which saves results in a disk file. Procedures IDATA and IDATF initialize the matrix and right-side vector to values that were generated by the rotor dynamics simulation. The use of actual data from the rotor dynamics simulation was important since the existence and accuracy of a matrix solution depends heavily on the matrix values. The results generated by the single-processor solver, as well as those for the two- and three-processor solvers, were compared to results generated by the rotor dynamics simulation on a mainframe computer. In all cases, the results matched exactly.

There are two versions of the two-processor solver given in the appendix: The first contains the original serial back-substitution algorithm; the other does the back substitution by using the column sweep approach. In both versions, the forward elimination process is done in parallel, and iterations within the IP loop are allocated to each processor. This is done with the WHILE-DO construct, as described in the Resource Allocation section of this report. Before each IP iteration begins, both processors synchronize to insure that the previous iteration was completed. This is critical since results from the previous iteration are needed to calculate the next iteration. Two boolean flags, one for each processor, are used to synchronize the processors. The flags are located in the global, or shared-memory, area. Both processors set their respective flags true after they have finished an IP iteration. Before starting the next iteration, each processor checks the other’s flag to make sure they are synchronized. Then both flags are cleared, and the iteration can begin. If one processor is not done, the other will wait for it. A counter is tested to exit the wait loop if the other processor does not respond.

The version of the two-processor solver with the column-sweep back-substitution algorithm differs from the serial back substitution version in two ways: (1) The synchronization is done with an assembly language procedure to decrease its execution time. The assembly procedure performs exactly the
same function as the original PASCAL version of the procedure (which is commented out in the listing); and (2) the back-substitution process, previously done on one processor, is now done on two processors. After an element of the result vector is computed, both processors work on computing partial results of other vector elements. Both processors then synchronize, compute the next full result vector element, and repeat the process until the entire result vector is obtained.

The code for the three-processor version of the solver uses a synchronization method which is more efficient than that used in the two-processor case. When each processor is done with its iteration, it sends a flag to each of the other processors in the system. Before starting the next iteration, each processor tests for the flags sent to it by the other processors. Since these flags are now in local memory (not global memory as in the two-processor case), the processor does not have to continually access the bus to test a flag. This reduces bus traffic for those processors which may still be accessing shared memory to complete their computations.

Another technique used in the three-processor solver to reduce bus traffic is the copying of frequently accessed variables from shared memory to local memory. In the three-processor code, arrays B12 and B13 are local-memory variables which contain current matrix row information used frequently throughout the program. These arrays are loaded with appropriate values from shared memory at the beginning of an IP iteration. All future references to these values are made from local memory, and the number of bus accesses required is reduced.

Table III summarizes the running time for each of the three cases. The speedup for each of the multiprocessor runs is also shown. A 30-block row matrix, with 4 by 4 blocks, was solved in each case. For the two-processor case, results are given for the serial back substitution and for the column-sweep back-substitution algorithms. As expected, the column sweep algorithm gives a faster solution. The two-processor case shows a speedup for 1.96, very close to the ideal value of 2. The three-processor case is less efficient with a speedup of 2.7. The reduction in efficiency can be attributed to a number of factors: Resource allocation, loss of cache variables, and increased access time for the shared memory because of increased bus traffic.

Several important notes are given here regarding cache memory. All multiprocessor runs were made with the cache memory enabled on all processors which did not contain the shared memory. The processor which did contain the shared memory had its cache disabled. This processor could not take advantage of the control register cache disabling for bus accesses (described in the hardware section) since all variables are physically within its own memory. The single-processor case used as the reference for speedup calculations was run with cache memory enabled. Variables which are in shared memory for the multiprocessor cases (and, hence, not cached) can be cached in the single-processor case. Thus, a certain percentage of the speedup achieved through parallel processing can be offset by the loss of cache variables. Although it appears that this is not a factor in the two-processor case, it may account for some of the overhead in the three-processor case.

A synchronization problem was encountered during the development of the three-processor solver which highlighted one of the difficulties with transporting existing algorithms (written for serial processors) to parallel processors. In the Gaussian elimination process, before elements below the diagonal are eliminated the original values are needed to compute other elements of the matrix. Thus, the sequence of the computations is critical. All processors would have to be synchronized (in addition to the synchronization that must be done for each IP iteration) to insure that the original value of the element being eliminated has been used by the other processors needing it. For example, consider the following elimination step for a 3 by 3 matrix:

1. BR: = A(2,1);
2. A(2,1): = A(2,1) - BR * A(1,1)/A(1,1);
3. A(2,2): = A(2,2) - BR * A(1,2)/A(1,1);
4. A(2,3): = A(2,3) - BR * A(1,3)/A(1,1);
5. F(2): = F(2) - BR * F(1)/A(1,1);

where A is the array of matrix elements and F is the right-side vector. In statement (1), BR is assigned the value of A(2,1), and the computation of A(2,1) in statement (2) will result in zero. Statement (1) is antidependent on statement (2) (ref. 10). Assume that four processors are available to do statements (2) through (5) with all data resident in a shared memory (except for BR which is in each processor's local memory). Each processor must execute the assignment statement which copies the value for A(2,1) from shared memory into local variable BR. It would appear, since each processor performs the same number of operations, that each processor could safely read A(2,1) before it is changed by processor 1. This also assumes that all processors begin their operations at the same time. However, timing differences between processors, communication delays between processors and shared memory, and load imbalances make this assumption dangerous. This was the case for the three-processor version

<table>
<thead>
<tr>
<th>Table III.—Timing Information for Multiprocessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of</td>
</tr>
<tr>
<td>processors</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
of the block tridiagonal solver as it was derived from the original FORTRAN code used in the rotor dynamics simulation. Synchronization routines were necessary which added overhead and resulted in slower execution.

The addition of synchronization routines for this part of the code can be avoided, however, by examining the Gaussian elimination process closer. The back-substitution process only requires elements above the diagonal to compute the result vector. The zeros below the diagonal are not needed. In fact, the computations which create the zeros are not necessary if the resulting upper right triangular matrix is only needed to compute the result vector. If the A(2,1) calculation was eliminated in the previous example, each processor would read the correct value of A(2,1) without synchronization problems. This approach was taken for the three-processor solver to achieve the speedup of 2.7. The single processor time used in the speedup calculation includes the computation of zero elements below the diagonal. If these computations are removed from the single-processor code also, then the relative speedup is reduced to 2.49. This is because the single-processor solver has fewer computations to do and, thus, runs faster.

Concluding Remarks

An approach to implementing a block tridiagonal matrix solver on a shared-memory parallel processor has been demonstrated. It should be possible to run the PASCAL programs for the one-, two-, and three-processor solvers on other shared-memory parallel processors, if the I/O and synchronization procedures are reproduced on the target system. The same approach can also be extended to more processors if they are available.

The results presented here are only a small part of the potential research that can be done on parallel processing of matrix solvers and solution of partial differential equations in general. Alternative architectures exist which have the potential for providing extremely fast matrix solutions. Architectures incorporating multiple array or vector processors, such as the ALLIANT FX/8 or CRAY X-MP, are examples. A pipelined math unit can perform operations much faster than the nonpipelined units found in typical microcomputers and mainframes. The key to tapping the potential of these architectures is the identification of at least two levels of parallelism in a given problem. The first is the operation level, which corresponds to the vectorization process done for single vector processors. The second is the vector operation level. Parallelism here consists of multiple vector operations which can be done concurrently.

Another high-potential research area is the investigation of alternative algorithms, given an architecture which can exploit the natural parallelism in the algorithm. There are many highly parallel iterative algorithms for solving systems of equations. Among these are successive overrelaxation methods (SOR) and conjugate gradient methods. Given an appropriate architecture, these methods could potentially yield higher performance than the Gaussian elimination method.

The selection of an appropriate algorithm for solving any problem on a parallel processor is a function of many parameters. NASA Lewis Research Center is currently constructing a hypercluster system to provide a test bed for investigating architecture and algorithm interactions (ref. 10). The combination of multiple vector and scalar processors in a flexible interconnection scheme will allow a wide variety of architectural concepts to be studied. It is hoped that future work using the hypercluster will answer some of the questions regarding appropriate architecture and algorithm combinations for both computational fluid mechanics and computational structural mechanics problems.

Lewis Research Center
National Aeronautics and Space Administration
Cleveland, Ohio, November 17, 1988
Appendix - PASCAL Program Listings

Single-Processor Block Tridiagonal Solver

1 1 0 0) -- PROGRAM SOLVE;
2 1 0 0) --
3 1 0 0) -- TYPE
4 1 0 0) --
5 1 0 0) -- RVECT=ARRAY [1..4,1..32] OF REAL;
6 1 0 0) -- AMAT=ARRAY [1..4,1..4,1..32,1..3] OF REAL;
7 1 0 0) --
8 1 0 0) -- VAR
9 1 0 0) --
10 1 0 0) -- A:B : AMAT;
11 1 0 0) -- F:DU : RVECT;
12 1 0 0) -- BF, BR : REAL;
13 1 0 0) -- N, I, J, IB, IP, I1 : INTEGER;
14 1 0 0) -- I1, I3, I2, J, IJ : INTEGER;
15 1 0 0) -- II, 18, 10, 181 : INTEGER;
16 1 0 0) --
17 1 0 0) -- PROCEDURE GETINF( VAR ADDR : RVECT; NUMEL : INTEGER )FORWARD;
18 1 0 0) --
19 1 0 0) -- PROCEDURE IDATA( VAR MATRIXA : AMAT; VNUM : INTEGER )FORWARD;
20 1 0 0) --
21 1 0 0) --
22 1 0 0) --
23 2 0 0) -- II1:= 1;
24 3 0 0) -- I1:=2;
25 4 0 0) -- I3:=3;
26 5 0 0) -- N:=30;
27 6 0 0) --
28 6 0 0) -- IDATA( B:1536 );
29 7 0 0) -- IDATF( C:12B );
30 8 0 0) --
31 8 0 0) -- FOR IB:= 1 TO N DO
32 9 0 0) -- FOR IP:= 1 TO 4 DO
33 0) --
34 10 0 0) -- BEGIN
35 11 0 0) --
36 12 0 0) -- FOR J:= IP TO 4 DO
37 13 0 0) --
38 14 0 0) -- BEGIN
39 15 0 0) -- CC IP, IB:= CC IP, IB / BF;
40 16 0 0) -- IF IP <= 4 THEN
41 17 0 0) --
42 17 0 0) -- II1:= IP + 1;
43 18 0 0) -- FOR I:= I1 TO 4 DO
44 19 0 0) -- BEGIN
45 19 0 0) --
46 20 0 0) -- FOR J:= IP TO 4 DO
47 21 0 0) --
48 22 0 0) -- BEGIN
49 23 0 0) -- CC I, IB:= CC I, IB - BR * CC IP, IB;
50 24 0 0) -- END [ FOR IB ]
51 25 0 0) -- END [ IF IP ]
52 25 0 0) --
53 25 0 0) -- IF IB <= N THEN
54 26 0 0) -- BEGIN
55 26 0 0) -- FOR I:= 1 TO 4 DO
56 27 0 0) -- BEGIN
57 27 0 0) --
58 28 0 0) -- END [ FOR IB ]
59 29 0 0) --
60 30 0 0) -- FOR J:= IB TO 4 DO
61 31 0 0) --
62 32 0 0) -- CC I, IB:= CC I, IB - BR * CC IP, IB;
63 33 0 0) -- END [ FOR I ]
64 0) -- END
65 0) -- END [ IF IB ]
66 0) --
**Dual-Processor Block Tridiagonal Solver; Serial Back Substitution, Processor 1**

```fortran
1() 0) 0) -- PROGRAM SOLVE;
2( 0) 0) -- TYPE
3( 0) 0) -- RVECT=ARRAY [1..4,1..32] OF REAL;
4( 0) 0) -- AMAT=ARRAY [1..4,1..4,1..3,1..32,1..3] OF REAL;
5( 0) 0) -- VAR
6( 0) 0) -- B : AMAT
7( 0) 0) -- C,DU : RVECT;
8( 0) 0) -- SYNC1,SYNC2 : BOOLEAN;
9( 0) 0) -- VAR
10( -6444) 0) -- K=128;
11( -7168) 0) -- GETINF( DU,K )
12( -7470) 0) -- K=128;
13( -7470) 0) -- VAR
14( 0) 0) -- PROCEDURE GETINF( VAR ADDR:RVECT; NUMEL:INTEGER )FORWARD;
15( 0) 0) -- PROCEDURE PUTINT( IVAL:INTEGER; VAR IPTR:INTEGER )FORWARD;
16( 0) 0) -- PROCEDURE IDATA( VAR MATRIXA : AMAT; VCNT : INTEGER )FORWARD;
17( 0) 0) -- PROCEDURE IDATF( VAR MATRIXF : RVECT; VCNT1 : INTEGER )FORWARD;
18( 0) 0) -- PROCEDURE IDATC ;
19( 0) 0) -- CDIST
20( 0) 0) -- CMAX=1000000;
21( 0) 0) -- VAR
22( 0) 0) -- A : AMAT
23( 0) 0) -- F : RVECT;
24( 0) 0) -- EP,EP : REAL;
25( 0) 0) -- N:1,J:IB:IP,J1 : INTEGER;
26( 0) 0) -- II,J:IB:IBID : INTEGER;
27( 0) 0) -- ERR:SCNT1:IPTR : INTEGER;
28( 0) 0) -- BEGIN
29( 2) 0) -- SYNCl=FALSE;
```
3  1)--- IF TR1=0;  
4  1)--- SCNT1=0;  
5  1)--- ERR:=0;  
6  1)--- III1=1;  
7  1)--- IT1:=2;  
8  1)--- I31:=3;  
9  1)--- N1:=30;  
10  1)--- IDATA(8*1536);  
11  1)--- IDATF(8*128);  
12  1)--- SYNC1:=TRUE;  
13  1)--- FOR IB1:=1 TO N DO  
14  1)--- FOR IP1:=1 TO 4 DO  
15  1)--- SCNT1:=0;  
16  1)--- REPEAT  
17  1)--- SCNT1:=SCNT1+1;  
18  1)--- IF SCNT1 > CMAX THEN  
19  1)--- BEGIN  
20  1)--- SYNC2:=TRUE;  
21  1)--- ERR:=ERR + 1;  
22  1)--- END;  
23  1)--- UNTIL SYNC2;  
24  1)--- SYNC2:=FALSE;  
25  1)--- BEGIN  
26  1)--- BL IP,IP,IB,IB2 J:= BL IP,J,IB,IB2 J / BP;  
27  1)--- J:= J2;  
28  1)--- C;  
29  1)--- END;  
30  1)--- WHILE J <= 4 DO  
31  1)--- IF IP <> 4 THEN  
32  1)--- BEGIN  
33  1)--- I1:= IP + 1;  
34  1)--- FOR I1:= I1 TO 4 DO  
35  1)--- BEGIN  
36  1)--- BR1:= BL I,IP,IB,IB2 J;  
37  1)--- J:= IP;  
38  1)--- WHILE J <= 4 DO  
39  1)--- BEGIN  
40  1)--- J:= J + 2;  
41  1)--- END;  
42  1)--- WHILE J <= 4 DO  
43  1)--- BEGIN  
44  1)--- CC I,IB J:= CC I,IB J - BR * BL IP,J,IB,IB2 J;  
45  1)--- END;  
46  1)--- FOR I1:= 1 TO 4 DO  
47  1)--- BEGIN  
48  1)--- IB1:= IB + 1;  
49  1)--- BR1:= BL I,IP,IB1,IB13;  
50  1)--- J:= IP;  
51  1)--- WHILE J <= 4 DO  
52  1)--- BEGIN  
53  1)--- BR1:= BL I,IB1,IB11 J:= BL I,IB1,IB13 J - BR * BL IP,J,IB,IB2 J;  
54  1)--- J:= J + 2;  
55  1)--- END;  

END;

J:= J + 2;

BEGIN

CC I,IB1 ]:= CC I,IB1 ] - BR * CC IF,IB 3;

END [ IF IB3 ]

SYNC1:=TRUE;

END [ FOR IF ]

SCNT1:=0;

REPEAT

SYNC1:=SCNT1+1;

IF SCNT1 > CMIX THEN

BEGIN

SYNC2:=TRUE;

ERR:= ERR + 1;

END

UNTIL SYNC2;

FOR IB1:= 1 TO N DO

BEGIN

IB:= IB + 1;

FOR I:= 1 TO 4 DO

BEGIN

I:= 5 - II;

DUE I,IB ]:= -1.0 * CC I,IB 3;

IF I <> 4 THEN

BEGIN

J:= I + 1;

WHILE J <= 4 DO

BEGIN


J:= J + 1;

END

END [ IF I <> 4 ];

IF IB <> N THEN

BEGIN

J:= 1;

WHILE J <= 4 DO

BEGIN


J:=J + 1;

END

END [ IF IB <> N ];

END [ FOR IB ];

PUTINT( ERR,IPTR );

PUTINT( SCNT1,IPTR );

K:= IB1;

GETINF( DU,K );

1-A END [ COPROC ];

1-A END.

IDATF ASSUMED EXTERNAL

IDATA ASSUMED EXTERNAL

PUTINT ASSUMED EXTERNAL

GETINF ASSUMED EXTERNAL

85 01 BEGIN

COPROC;

0-A END.

NO ERROR(S) AND NO WARNING(S) DETECTED

175 LINES 5 PROCEDURES

1076 PCODE INSTRUCTIONS
Dual-Processor Block Tridiagonal Solver; Serial Back Substitution, Processor 2

1(0) 0)-- PROGRAM SOLVE;
2(0) 0)--
3(0) 0)-- TYPE
4(0) 0)--
5(0) 0)-- RVECT=ARRAY [1..9,1..32] OF REAL;
6(0) 0)-- AMAT=ARRAY [1..9,1..41..32,1..3] OF REAL;
7(0) 0)--
8(0) 0)-- VAR
9(0) 0)--
10(1-1490) 0)-- E : AMAT;
11(1-7168) 0)-- C+DU : RVECT;
12(1-7170) 0)-- SYNC1+SYNC2 : BOOLEAN;
13(1-7170) 0)--
14(0) 1)-- PROEDURE SETAS(IOFFST : INTEGER) : FORWARD;
15(0) 1)--
16(0) 1)-- PROEDURE PUTINT(IVAR : INTEGER; VAR IPTR : INTEGER) : FORWARD;
17(0) 1)--
18(0) 1)-- PROEDURE COPROC;
19(0) 1)--
20(0) 1)-- CONST
21(0) 1)--
22(0) 1)-- CMAX=1000000;
23(0) 1)--
24(0) 1)-- VAR
25(0) 1)--
26(1-1490) 1)-- A : AMAT;
27(1-6656) 1)-- F : RVECT;
28(1-6664) 1)-- BP,BR : REAL;
29(1-6888) 1)-- N=I1+I2+I3+K,J1 : INTEGER;
30(0-6800) 1)-- I2=I3+I1;K,J1 : INTEGER;
31(1-6724) 1)-- T1,J1,I1+I2+I3 : INTEGER;
32(0-6736) 1)-- SCNTZ,ER1,IPTR : INTEGER;
33(0-6736) 1)--
34(1) 1)-- BEGIN
35(1) 1)-- SETAS(16020000);
36(1) 1)-- SYNC21=TRUE;
37(4) 1)-- IPTR1=0;
38(5) 1)-- SCNT21=0;
39(6) 1)-- ERR1=0;
40(7) 1)-- II1=21;
41(8) 1)-- III1=11;
42(9) 1)-- II2=21;
43(10) 1)-- II=31;
44(11) 1)-- N1=30;
45(12) 1)-- K1=128;
46(13) 1)--
47(1) 1)--
48(13) 1)-- FOR IB1=1 TO N DO
49(14) 1)-- FOR IP1=1 TO 4 DO
50(10) 1)-- BEGIN
51(11) 1)--
52(15) 1)-- SCNT21=0;
53(13) 1)-- REPEAT
54(16) 1)-- [ SYNCHRONIZE WITH PROCESSOR 1 ]
55(17) 1)--
56(17) 1)-- IF SCNT2 > CMAX THEN
57(1) 1)-- BEGIN
58(10) 1)-- Sync11=TRUE;
59(19) 1)-- ERR1=ERR1+1;
60(1) 1)-- END;
61(20) 1)-- UNTIL Sync11;
62(21) 1)-- Sync11=FALSE;
63(1) 1)--
64(22) 1)-- BP1=BC IP1+IP1+I2;J1;
65(23) 1)-- J1=IP1+1;
66(24) 1)-- WHILE J <= 4 DO
67(13) 1)-- BEGIN
68(25) 1)-- BP1=BC IP1;J1+I2;J1=BC IP1;J1+I2 J / BP1;
69(26) 1)-- J1=J1+2;
69     1-C--          ENO;
70     27 1--          J1= 2;
71     28 1--          WHILE J <= 9 DO
72     29 1-C--          BEGIN
73     30 1--          BE IF,J,IB+I3 J= BE IF,J,IB+I3 J / BP; 
74     31 1--          J1= J + 21
75     32 1-C--          END;
76     33 1--
77     34 1--          IF IP <= 4 THEN
78     35 1-C--          BEGIN
79     36 1--          II1= IP + 11;
80     37 1--          FOR J1= 11 TO 4 DO
81     38 1--          BEGIN
82     39 1--          BE II1= BE I1P,IB+I2 J1
83     40 1--          J1= IP + 11;
84     41 1--          WHILE J <= 4 DO
85     42 1--          BEGIN
86     43 1--          BE I1P,IB+I3 J1= BE I1P,IB+I3 J1 - BR * BE IF,J,IB+I3 J1;
87     44 1--          J1= J + 21
88     45 1--          END;
89     46 1--
90     47 1--          END;
91     48 1--
92     49 1-C--          END;
93     50 1--
94     51 1--          IF IB <= N THEN
95     52 1-C--          BEGIN
96     53 1--          FOR J1= 1 TO 4 DO
97     54 1--          BEGIN
98     55 1--          BE II1= IB + 11;
99     56 1--          BE I1P= BE I1P,IB+I1J1;
100     57 1--          J1= IP + 11;
101     58 1--          WHILE J <= 4 DO
102     59 1--          BEGIN
103     60 1--          BE I1P,IB+I3 J1= BE I1P,IB+I3 J1 - BR * BE IF,J,IB+I3 J1;
104     61 1--          J1= J + 21
105     62 1--          END;
106     63 1--
107     64 1--          END;
108     65 1--
109     66 1--
110     67 1--
111     68 1-C--          END;
112     69 1--
113     70 1--
114     71 1--
115     72 1--
116     73 1--
117     74 1-C--          END; (FOR IB)
118     75 1--          SYNC21=TRUE;
119     76 1-B--          ENO; (FOR IP)
120     77 1--
121     78 1--          PUTINT( ERR,IPTR );
122     79 1--          PUTINT( SCNT?,IPTR );
123     80 1--          1)-A ENO; I CPOPROC 3
124     81 1--

**** PUTINT ASSUMED EXTERNAL
**** SETAS ASSUMED EXTERNAL
125  51 0-A-- BEGIN
126  52 0--  CPOPROC:
127  53 0-A ENO.

**** NO ERROR(S) AND NO WARNING(S) DETECTED
**** 327 LINES 3 PROCEDURES
**** 717 CODE INSTRUCTIONS

Dual-Processor Block Tridiagonal Solver; Column Sweep Serial Back Substitution, Processor 1

 truyền
BEGIN
1 1)  ---
2 2)  ---
3 3)  ---
4 4)  ---
5 5)  ---
6 6)  ---
7 7)  ---
8 8)  ---
9 9)  ---
10 10)  ---
11 11)  ---
12 12)  ---
13 13)  ---
14 14)  ---
15 15)  ---
16 16)  ---
17 17)  ---
18 18)  ---
19 19)  ---
20 20)  ---
21 21)  ---
22 22)  ---
23 23)  ---
BEGIN
BC IP,J,IB+13 J= BC IP,J,IB+13 J / BP1
J:= J + 2;
END;
BEGIN
BR:= BC I,IP,IB,J2 J;
J:= IP;
WHILE J <= 4 DO
BEGIN
J:= J + 2;
END;
CE I,IB J:= CE I,IB J - BR * CE IP,IB J;
END; IF I3
END; IF IP3
BEGIN
FOR I:= 1 TO 4 DO
BEGIN
BR:= IB+1 J:= IP;
BR:= BC I,IP+13,IB+1,II1 J;
WHILE J <= 4 DO
BEGIN
BR:= BC I,IP+13,IB+1,II1 J - BR * BC IP+1,IB+1,II1 J;
J:= J + 2;
END;
END;
BEGIN
CE I,IB+1 J:= CE I,IB+1 J - BR * CE IP+1,IB+1 J;
END; IF IP3
END; IF IP3
SYNC1:=TRUE;
BEGIN
SYNC2:=FALSE;
SYNC1:=TRUE;
SYNC2:=TRUE;
ERR:= ERR + 1;
END;
BEGIN
SYNC2:=TRUE;
SYNC1:=TRUE;
SYNC2:=FALSE;
BEGIN
FOR IB:= 1 TO N DO
BEGIN
IF SCNT1 > CMAX THEN
BEGIN
SYNC2:=TRUE;
END;
ERR:= ERR + 1;
END;
BEGIN
SYNC2:=TRUE;
SYNC1:=TRUE;
SYNC2:=FALSE;
BEGIN
COLUMN SWEEP BACKSUBSTITUTION ALGORITHM
END;
BEGIN
II:=I;
WHILE I <= 4 DO
BEGIN
C11,IB3:= -1.0 * C11,IB3;
II:=II + 2;
END;
END;
BEGIN
SYNC1:=TRUE;
SYNC2:=TRUE;
SYNC2:=FALSE;
158 64 I)-- FOR IBI = 1 TO N DO
159 65 I)-- BEGIN
160 66 I)-- IBI = I + 1 - IBI;
161 67 I)-- BEGIN
162 68 I)-- DUE[1+IBI] = C[1+IBI] - B[1+IBI][2] * DUE[1+IBI];
163 69 I)-- FOR I = 4 DOWNTO 1 DO
164 70 I)-- BEGIN
165 71 I)-- J = I - 1;
166 72 I)-- IF (NOT((IB = 1) AND (I = 1)) AND (I < 1)) THEN
167 73 I)-- BEGIN
168 74 I)-- J = I - 1;
169 75 I)-- WHILE J >= 1 DO
170 76 I)-- BEGIN
172 78 I)-- J = J - 2;
173 79 I)-- END;
174 80 I)-- END;
175 81 I)-- IF IB > 1 THEN
176 82 I)-- BEGIN
177 83 I)-- J = 1;
178 84 I)-- WHILE J <= 4 DO
179 85 I)-- BEGIN
181 87 I)-- J = J + 2;
182 88 I)-- END;
183 89 I)-- END;
184 90 I)-- SYNC2 = FALSE;
185 91 I)-- SYNC1 = TRUE;
186 92 I)-- SYNC1 = TRUE;
187 93 I)-- END;
188 94 I)-- END;
189 95 I)-- END;
190 96 I)-- END;
191 97 I)-- END;
192 98 I)-- [OLD BACKSUBSTITUTION]
193 99 I)-- FOR IBI = 1 TO N DO
194 100 I)-- BEGIN
195 101 I)-- IBI = I + 1 - IBI;
196 102 I)-- IBI = IBI + 1;
197 103 I)-- FOR II = 1 TO 4 DO
198 104 I)-- BEGIN
199 105 I)-- I = 5 - II;
200 106 I)-- DUE[I,IBI] = -1.0 * C[1,IBI];
201 107 I)-- IF I < 4 THEN
202 108 I)-- BEGIN
203 109 I)-- J = I + 1;
204 110 I)-- WHILE J <= 4 DO
205 111 I)-- BEGIN
206 112 I)-- J = 1;
207 113 I)-- J = J + 1;
208 114 I)-- END;
209 115 I)-- END;
210 116 I)-- IF IB < N THEN
211 117 I)-- BEGIN
212 118 I)-- J = 1;
213 119 I)-- WHILE J <= 4 DO
214 120 I)-- BEGIN
216 122 I)-- J = J + 1;
217 123 I)-- END;
218 124 I)-- END;
219 125 I)-- END;
220 126 I)-- END;
221 127 I)-- PUTINT(ERR,IPTR);
222 128 I)-- PUTINT(SCONT,IPTR);
223 129 I)-- K = 128;
224 130 I)-- GETINF(DU*K);
225 131 I)-- A END; E COPROC 3
226 132 I)--

**** IDATF ASSUMED EXTERNAL
**** IDATA ASSUMED EXTERNAL
**** SYNC2 ASSUMED EXTERNAL
**** DUE PUTINT ASSUMED EXTERNAL
**** GETINF ASSUMED EXTERNAL
227 133 80 0-A BEGIN
228 134 89 0-COPROC;
229 135 0-A END.
Dual-Processor Block Tridiagonal Solver; Column Sweep Back Substitution, Processor 2

```
1 0 0 --- PROGRAM SOLVE;
2 0 0 ---
3 0 0 --- TYPE
4 0 0 ---
5 0 0 --- RVECT=ARRAY [1..4*1..32] OF REAL;
6 0 0 --- AMAT=ARRAY [1..4*1..4*1..32] OF REAL;
7 0 0 ---
8 0 0 --- VAR
9 0 0 ---
10 -6144 0 --- B : AMAT;
11 -7168 0 --- C:DU : RVECT;
12 -7176 0 --- SYNCL,SYNC2 : BOOLEAN;
13 -7176 0 ---
14 0 0 --- PROCEDURE SETAS(IOFFS : INTEGER);FORWARD;
15 0 0 ---
16 0 0 --- PROCEDURE PUTINT(IHAR:INTEGER; VAR IPTR:INTEGER);FORWARD;
17 0 0 ---
18 0 0 --- PROCEDURE SYNCRD(VAR EFLG,CNT:INTEGER; VAR MAXCNT:INTEGER; VAR SFLG:BOOLEAN);
19 0 0 --- FORWARD;
20 0 1 ---
21 0 1 --- PROCEDURE COPROC ;
22 0 1 ---
23 0 1 --- CONST
24 0 1 ---
25 0 1 --- CMAX=1600000;
26 0 1 ---
27 0 1 --- VAR
28 0 1 ---
29 -6144 0 --- A : AMAT;
30 -6656 0 --- F : RVECT;
31 -6664 0 --- BP:DR : REAL;
32 -6688 0 --- N+1,J:IP,J:1 : INTEGER;
33 -6704 0 --- I2:13,II:13 : INTEGER;
34 -6720 0 --- II,J,II,II,J : INTEGER;
35 -6736 0 --- SCNT2,ERR,IPTR : INTEGER;
36 -6736 0 ---
37 1 1 A= BEGIN
38 2 1 --- SETAS( 16#3000000 );
39 3 1 --- SYNC2:=TRUE;
40 4 1 --- IPTR:=0;
41 5 1 --- SCNT2:=0;
42 6 1 --- ENR:=0;
43 7 1 --- J:=21;
44 8 1 --- II:=11;
45 9 1 --- I2:=2;
46 10 1 --- I3:=3;
47 11 1 --- NI:=301;
48 12 1 --- KI:=128;
49 1 ---
50 1 ---
51 13 1 --- FOR IBI:= 1 TO N DO
52 14 1 --- FOR IP:= 1 TO 4 DO
53 1B= BEGIN
54 1 ---
55 15 1 --- SYNCRD( ERR,SCNT2,CMAX,SYNC1 );
56 1 --- []
57 1 --- []
58 1 --- []
59 1 --- []
60 1 --- [
61 1 --- [
62 1 --- [
63 1 --- [
64 1 --- [ UNTIL SYNC1 ;
65 16 1 --- SYNC1:= FALSE;
```
66  1)  --  \text{BP} := \text{BC IP,IP,IB,IB2 J;}
67  26  1)  --  \text{IF IP} \land 4 \text{ THEN}
68  27  1)  --  \text{BEGIN}
69  28  1)  --  \text{FOR I:= I1 TO 4 DO}
70  29  1)  --  \text{BEGIN}
71  30  1)  --  \text{I1:= IP + 1;}
72  31  1)  --  \text{FOR I:= I1 TO 4 DO}
73  32  1)  --  \text{BEGIN}
74  33  1)  --  \text{IB1:= IB + 1;}
75  34  1)  --  \text{END; [FOR I]}
76  35  1)  --  \text{BEGIN}
77  36  1)  --  \text{SYNC2:=TRUE;}
78  37  1)  --  \text{END; [IF IB]}
79  38  1)  --  \text{BEGIN}
80  39  1)  --  \text{SYNC2:=FALSE;}
81  40  1)  --  \text{BEGIN}
82  41  1)  --  \text{BEGIN}
83  42  1)  --  \text{BEGIN}
84  43  1)  --  \text{BEGIN}
85  44  1)  --  \text{BEGIN}
86  45  1)  --  \text{BEGIN}
87  46  1)  --  \text{BEGIN}
88  47  1)  --  \text{BEGIN}
89  48  1)  --  \text{BEGIN}
90  49  1)  --  \text{BEGIN}
91  50  1)  --  \text{BEGIN}
92  51  1)  --  \text{BEGIN}
93  52  1)  --  \text{BEGIN}
94  53  1)  --  \text{BEGIN}
95  54  1)  --  \text{BEGIN}
96  55  1)  --  \text{BEGIN}
97  56  1)  --  \text{BEGIN}
98  57  1)  --  \text{BEGIN}
99  58  1)  --  \text{BEGIN}
100  59  1)  --  \text{BEGIN}
101  60  1)  --  \text{BEGIN}
102  61  1)  --  \text{BEGIN}

\text{\textbf{COLUMN SHEEP BACKSUBSTITUTION ALGORITHM}}
Three-Processor Block Tridiagonal Solver; Processor 1

1( 0) 0) --- PROGRAM SOLVE;
2( 0) 0) ---
3( 0) 0) --- TYPE
4( 0) 0) ---
5( 0) 0) --- INTS=ARRAY [1..5] OF INTEGER;
5( 0) 0) --- REALA=ARRAY [1..4] OF REAL;
5( 0) 0) --- RVECT=ARRAY [1..4,1..32] OF REAL;
5( 0) 0) --- AMAT=ARRAY [1..4,1..4,1..32,1..3] OF REAL;
5( 0) 0) ---
5( 0) 0) ---
11( 0) 0) ---
12( -6454) 0) --- B : AMAT;
13( -7168) 0) --- C*DU : RVECT;
14( -7168) 0) ---
15( -7168) 0) ---
16( 0) 1) --- PROCEDURE GETINF( VAR ADDR:INVECTOR; NULM:INTEGER ) ;FORWARD;
17( 0) 1) ---
18( 0) 1) --- PROCEDURE PUTINT( IVAL:INTEGER; VAR IPTR:INTEGER ) ;FORWARD;
19( 0) 1) ---
20( 0) 1) --- PROCEDURE SYNDRO2( VAR SYNCINF : INT5 ) ;FORWARD;
21( 0) 1) ---
22( 0) 1) --- PROCEDURE IDATA( VAR MTRAXA : AMAT; VENTRY : INTEGER ) ;FORWARD;

***** NO ERROR(S) AND NO WARNING(S) DETECTED

***** 180 LINES 4 PROCEDURES

***** 1040 PCODE INSTRUCTIONS
01 1) --- PROEDURE IDATF( VAR MATRIXF: RVECTOR; VCNT1: INTEGER ); FORWARD;
02 1) ---
03 1) ---
04 1) ---
05 1) ---
06 1) ---
07 1) ---
08 1) ---
09 1) ---
10 1) ---
11 1) ---
12 1) ---
13 1) ---
14 1) ---
15 1) ---
16 1) ---
17 1) ---
18 1) ---
19 1) ---
20 1) ---
21 1) ---
22 1) ---
23 1) ---
24 1) ---
25 1) ---
26 1) ---
27 1) ---
28 1) ---
29 1) ---
30 1) ---
31 1) ---
32 1) ---
33 1) ---
34 1) ---
35 1) ---
36 1) ---
37 1) ---
38 1) ---
39 1) ---
40 1) ---
41 1) ---
42 1) ---
43 1) ---
44 1) ---
45 1) ---
46 1) ---
47 1) ---
48 1) ---
49 1) ---
50 1) ---
51 1) ---
52 1) ---
53 1) ---
54 1) ---
55 1) ---
56 1) ---
57 1) --- FOR IB:= 1 TO N DO
58 1) --- FOR IP:= 1 TO 4 DO
59 1) --- B:= BEGIN
60 1) ---
61 1) --- SYNERG2( SYNTAB );
62 1) ---
63 1) --- IF ( IP = 1 ) OR ( IP = 4 )
64 1) ---
65 1) ---
66 1) ---
67 1) ---
68 1) ---
69 1) ---
70 1) ---
71 1) ---
72 1) ---
73 1) ---
74 1) ---
75 1) ---
76 1) ---
77 1) ---
78 1) ---
79 1) ---
80 1) ---
81 1) ---
82 1) ---
83 1) ---
84 1) ---
85 1) ---
86 1) ---
87 1) ---
88 1) ---
89 1) ---
90 1) ---
91 1) ---
92 1) ---
93 1) ---
94 1) ---
95 1) ---
96 1) ---
97 1) ---
98 1) ---
99 1) ---

23
BEGIN
FOR I:= 1 TO N DO
BEGIN
IF (IP <> 4) THEN
BEGIN
FOR II:= I TO N DO
BEGIN
BR:= BE I*IP,IB*I2 J;
END;
J:= IP + 3;
END;
WHILE J <> 4 DO
BEGIN
J:= J + 1;
END;
END;
END FOR I3
END IF IP;

BEGIN
FOR I:= 1 TO N DO
BEGIN
J:= J + 3;
END;
END;
END;
END IF IP = 1 OR IP = 4 THEN
BEGIN
CC I*IB*l J:= CC I*IB*l J - BR * CI IF*IB J;
END;
END IF IP;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
WHILE I <> 4 DO
BEGIN
CC I*IB*I2 J:= -1.0 * CC I*IB*I3;
I:= I + 3;
END;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
IB:= IB + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1 TO N DO
BEGIN
IF (NOT(I<> 1 AND I== 1)) AND (I <> 1) THEN
BEGIN
J:= J + 1;
END;
END;
BEGIN
FOR I:= 1 TO N DO
BEGIN
I:= I + 1;
END;
OR1GINAL PAGE IS
OF POOR QUALITY

Three-Processor Block Tridiagonal Solver; Processor 2

0) 0) --- PROGRAM SOLVE:
0) 0) --- TYPE
4) 0) --- INTS=ARRAY [1..53] OF INTEGER;
7) 0) --- REALS=ARRAY [1..43] OF REAL;
8) 0) --- KVEC=ARRAY [1..41,1..32] OF REAL;
9) 0) --- AMAT=ARRAY [1..61,1..9,1..32,1..31] OF REAL;
10) 0) --- VAR
11) 0) --- 120
12) 0) --- 13) 0) --- 14) 0) --- 15) 0) --- 16) 0) --- 17) 0) --- 18) 0) --- 19) 0) --- 20) 0) --- 21) 0) --- 22) 0) --- 23) 0) --- 24) 0) --- 25) 0) --- 26) 0) ---

--- PROCESS SETAS( IOFFST : INTEGER )FORWARD;
18) 0) --- 19) 0) --- 20) 0) --- 21) 0) --- 22) 0) --- 23) 0) --- 24) 0) --- 25) 0) --- 26) 0) ---
END
Three-Processor Block Tridiagonal Solver; Processor 3

100(0)(0) --- PROGRAM SOLVE;
110(0)(0) --- TYPE
120(0)(0) --- B: AMAT;
130-7168(0) --- C,D,U: RUCF;
140-7168(0) ---
150-7168(0) ---
160(0)(0) --- PROCEDURE SETAS(IOFFST: INTEGER):FORWARD;
170(0)(0) ---
180(0)(0) --- PROCEDURE PUTINT(INTV: INTEGER; VAR IPTR:INTEGER):FORWARD;
190(0)(0) ---
200(0)(0) --- PROCEDURE SYNCROZC(VAR SYNCINF: INT5):FORWARD;
210(0)(0) ---
220(0)(0) --- PROCEDURE COPROC:
230(0)(0) ---
240(0)(0) --- CONST
250(0)(0) ---
260(0)(0) --- CMAX=1000000;
270(0)(0) --- VAR
280(0)(0) ---
290-64(0) --- AI2,AII1,BI2,BI3: REAL;
300-576(0) --- P: RUCF;
310-594(0) --- DP,DR: REAL;
320-620(0) --- N1,J1,1J1,1IP,T1: INTEGER;
330-620(0) --- I2,II1,II2,J1: INTEGER;
340-644(0) --- NI,J,1J1,1EI: INTEGER;
350-656(0) --- ERR,SYNCINF,IPTR: INTEGER;
360-676(0) --- SYNTAX: INT5;
370-676(0) ---

ORIGINAL PAGE IS OF POOR QUALITY
1. \[ \text{while } j < 4 \text{ do} \]
   \[ i = i + 1 \]
   \[ j = j + 2 \]
2. \[ \text{while } j < 4 \text{ do} \]
   \[ i = i + 1 \]
   \[ j = j + 2 \]
3. \[ \text{while } j < 4 \text{ do} \]
   \[ i = i + 1 \]
   \[ j = j + 2 \]
4. \[ \text{while } j < 4 \text{ do} \]
   \[ i = i + 1 \]
   \[ j = j + 2 \]
5. \[ \text{while } j < 4 \text{ do} \]
   \[ i = i + 1 \]
   \[ j = j + 2 \]
115 1) -- ENDD [ FOR I3  
116 11-D  
117 11-C  
118 49 1) -- IF IB <> N THEN  
119 11-C-  
120 50 1) -- FOR I= 1 TO 4 DO  
121 11-D-  
122 51 1) -- BEGIN  
123 1) --  
124 52 1) -- J1 = IP + 2;  
125 53 1) -- WHILE J <= 4 DO  
126 11-C-  
127 54 1) -- BE I, J1, IB1, I2  
128 55 1) -- J1 = J + 3;  
129 1) -- END;  
130 1) --  
131 56 1) -- J1 = 3;  
132 57 1) -- WHILE J <= 4 DO  
133 11-D-  
134 58 1) -- BE I, J1, IB1, I2  
135 59 1) -- J1 = J + 3;  
136 1) -- END;  
137 1) --  
138 60 1) -- IF IP = 3 THEN  
139 61 1) -- CE I, IB1  
140 1) --  
141 1) -- END; [ FOR I3  
142 1) -- END; [ IF IB]  
143 1) --  
144 1) --  
145 62 1) -- SYMRO2C ( SYNTAX);  
146 1) --  
147 1) -- [COLUMN SLEEP BACKSUBSTITUTION ALGORITHM]  
148 63 1) -- FOR IB = 1 TO N DO  
149 11-B-  
150 64 1) -- BEGIN  
151 65 1) -- II = 3;  
152 66 1) -- WHILE I <= 4 DO  
153 11-C-  
154 67 1) -- CII, IB1 = -1.0 * CII, IB1;  
155 11-D-  
156 68 1) -- II = I + 3;  
157 11-E-  
158 69 1) -- END;  
159 1) --  
160 69 1) -- SYMRO2C ( SYNTAX);  
161 1) --  
162 69 1) -- FOR IB = 1 TO N DO  
163 11-B-  
164 70 1) -- BEGIN  
165 71 1) -- IB1 = IB - 1;  
166 72 1) -- DII, IB1, IB1 = CII, IB1;  
167 73 1) -- FOR II = 4 DOWNTO 1 DO  
168 11-C-  
169 74 1) -- J1 = I - 1;  
170 75 1) -- IF ( NOT((II = 1) AND (I = 1)) AND (I < 1)) THEN  
171 11-D-  
172 76 1) -- BEGIN  
173 77 1) -- WHILE J <= 1 DO  
174 11-E-  
175 78 1) -- CII, IB1 = CII, IB1 - BII, J, IB1 - 2 * DII, IB1;  
176 79 1) -- Jl = J - 3;  
177 1) -- END;  
178 80 1) -- IF IB <> 1 THEN  
179 11-D-  
180 81 1) -- BEGIN  
181 82 1) -- WHILE J <= 4 DO  
182 11-E-  
183 83 1) -- CII, IB1, IB1 = CII, IB1 - BII, J, IB1, 3 * DII, IB1;  
184 84 1) -- J1 = J + 3;  
185 1) -- END;  
186 1) -- END;  
187 85 1) -- SYMRO2C ( SYNTAX);  
188 86 1) -- SYMRO2C ( SYNTAX);  
189 11-C-  
190 1) -- END;  
191 1) --
ORIGINAL PAGE IS OF POOR QUALITY

87 1)--- ERR1=SYNCTABLE1
88 1)--- PUTINT( ERR1,IPTR );
89 1)--- SNT1=SYNCTABLE2;
90 1)--- PUTINT( SNT1,IPTR );
91 1)--- A END; [ CPROC ]

****** SYMPOSC ASSUMED EXTERNAL
****** PUTINT ASSUMED EXTERNAL
****** SNEAS ASSUMED EXTERNAL
90 91 0)A- BEGIN
92 0)--- CPROC;
93 0)--- A END.

****** NO ERROR(S) AND NO WARNING(S) DETECTED

****** 280 LINES 4 PROCEDURES

****** 1134 PCODE INSTRUCTIONS
References

The solution of a block tridiagonal matrix using parallel processing is demonstrated in this report. The multiprocessor system which obtained the results and the software environment used to program that system are described. Theoretical partitioning and resource allocation for the Gaussian elimination method used to solve the matrix are discussed. The results obtained from running one-, two-, and three-processor versions of the block tridiagonal solver are presented. The PASCAL source code for these solvers is given in the appendix, and it may be transportable to other shared-memory parallel processors, provided that the synchronization routines are reproduced on the target system.