The Programming Interface

Libraries and languages make parallel programming possible, but rarely easy
Commentary on Infix form of PP

- What was your experience with formulating a parallel prefix computation as an infix operation?
Trees are an important component of computing
- The “Schwartz tree” has been logical
- Trees as data structures are complicated because they are typically more dynamic
- Pointers are generally not available
- Work well with work queue approach
- As usual, we try to exploit locality and minimize communication
Breadth-first Trees

- Common in games, searching, etc

- Split: Pass 1/2 to other processor, continue
  - Stop when processors exhausted
  - Responsible for tree that remains
  - Ideal when work is localized
Common in graph algorithms

Get descendants, take one and assign others to the task queue

Key issue is managing the algorithm’s progress
Tree algorithms often need to know how others are progressing
- Interrupt works if it is just a search: Eureka!!
- Record $\alpha$-$\beta$ cut-offs in global variable
- Other pruning data, e.g. best so far, also global
- Classic error is to consult global too frequently

Rethink: What is tree data structure’s role?

Write essay: Dijkstra’s algorithm is not a good... :)
Complications

- If coordination becomes too involved, consider alternate strategies:
  - Graph traverse => local traverse of partitioned graph

- Local computation uses sequential tree algorithms directly ... stitch together
Trees are a useful data structure for recording spatial relationships: K-D trees.

Generally, decomposition is unnecessary “all the way down” -- but this optimization implies two different regimes.
Cap Reduces Communication

- The nodes near root can be stored redundantly
  - Processors consult local copy -- alert others to changes

Each process keeps copy of “cap” nodes
Reconceptualizing is often most effective.
Focus has not been on ||ism, but on other stuff:
- Exploiting locality
- Balancing work
- Reducing inter-thread dependences
We produced general purpose solution mechanisms: UD-reduce and UD-scan.
We like trees, but recognize that direct application is not likely.
The Programming Interface

“I don’t know what the technical characteristics of the standard language for scientific and engineering computation will be in the year 2000 . . . but I know it will be called Fortran.”

John Backus, c. 1980
I have argued that a key property of a parallel programming system is that it embody an accurate (CTA) model of computation.

Recall why:
- Wrong model leads to picking wrong algorithm
- Communication costs -- they cannot be ignored
- Parallel programs must port, so pick universal model

So, which of our present languages do that? Today, we’ll see.
At least 100 serious parallel programming languages have been developed in the last 2 decades ... why isn’t the problem solved?

- Generalizing ...
  - Most languages focused on a “silver bullet” solution, but the problem is more complex
  - Just a few of the languages were fully implemented
  - To be taken seriously, a language must
    - Run serious applications fast
    - Run on “all” parallel machines
    - Have substantial support (docs, compilers with libraries, tools such as debuggers and IDEs, 1-800 #)
No new languages crossed the bar

- Performance challenge ...
  - Serious applications programs are huge -- it is time consuming to write an equivalent program in any language, and it may require domain knowledge
  - Production programs are often well optimized -- competing on performance implies an effective compiler and performance debugging tools
  - “Linear speedup” goal ($P$ processors will yield a $P$-fold speed-up) is naïve, but widely assumed
  - Doing well on one program is not persuasive

- Portability challenges are similar
- Will any programmer learn a new language?
Today, with few exceptions, we program using library-based facilities rather than languages:

- Sequential language + message passing in MPI or PVM
- Sequential language + thread packages such as P-threads, or equivalently, Java-threads
- OpenMP with a pragma-aware compiler for a sequential programming language

Consider each briefly before discussing new developments.
Message Passing

- Message passing is “the lowest of the low”, but remains in widespread use because ...
  - It works -- embodies the CTA || model
  - It is required for clusters, supercomputers, etc.
  - Achieving performance is definitely possible
  - Portability is essential for long-lived programs

- What is it?
  - Variations on primitive `send/receive`
  - Process spawning, broadcast, etc.
  - Programming goodies: reduce, scan, processor groups
Realities of Message Passing

- In message passing
  - There are few abstractions to simplify the work
  - Programmers must do everything except the physical layer
  - Experiments show that compared to “designed from first principles” parallel languages, MPI programs are 6 times larger ... the extra code is the subtle, difficult to get right, and timing-sensitive
  - Consider dense matrix multiplication
MPI_Status status;
main(int argc, char **argv) {
    int numtasks, /* number of tasks in partition */
        taskid, /* a task identifier */
        numworkers, /* number of worker tasks */
        source, /* task id of message source */
        dest, /* task id of message destination */
        nbytes, /* number of bytes in message */
        mtype, /* message type */
        intsize, /* size of an integer in bytes */
        dbsize, /* size of a double float in bytes */
        rows, /* rows of matrix A sent to each worker */
        averow, extra, offset, /* used to determine rows sent to each worker */
        i, j, k, /* misc */
        count;
    double a[NRA][NCA], /* matrix A to be multiplied */
            b[NCA][NCB], /* matrix B to be multiplied */
            c[NRA][NCB]; /* result matrix C */
intsize = sizeof(int);
dbsize = sizeof(double);

MPI_Init(&argc, &argv);
MPI_Comm_rank(MPI_COMM_WORLD, &taskid);
MPI_Comm_size(MPI_COMM_WORLD, &numtasks);
numworkers = numtasks-1;

-Clause

Create test data --
actually inputting data is harder.
/* send matrix data to the worker tasks */
averow = NRA/numworkers;
extra = NRA%numworkers;
offset = 0;
mtype = FROM_MASTER;
for (dest=1; dest<=numworkers; dest++) {
    rows = (dest <= extra) ? averow+1 : averow;
    MPI_Send(&offset, 1, MPI_INT, dest, mtype, MPI_COMM_WORLD);
    MPI_Send(&rows, 1, MPI_INT, dest, mtype, MPI_COMM_WORLD);
    count = rows*NCA;
    MPI_Send(&a[offset][0], count, MPI_DOUBLE, dest, mtype, MPI_COMM_WORLD);
    count = NCA*NCB;
    MPI_Send(&b, count, MPI_DOUBLE, dest, mtype, MPI_COMM_WORLD);
    offset = offset + rows;
}
/* wait for results from all worker tasks */
mtype = FROM_WORKER;
for (i=1; i<=numworkers; i++) {
    source = i;
    MPI_Recv(&offset, 1, MPI_INT, source, mtype, MPI_COMM_WORLD, &status);
    MPI_Recv(&rows, 1, MPI_INT, source, mtype, MPI_COMM_WORLD, &status);
    count = rows*NCB;
    MPI_Recv(&c[offset][0], count, MPI_DOUBLE, source, mtype, MPI_COMM_WORLD, &status);
}

/************************ worker task *************************************/
if (taskid > MASTER) {
    mtype = FROM_MASTER;
    source = MASTER;
    MPI_Recv(&offset, 1, MPI_INT, source, mtype, MPI_COMM_WORLD, &status);
    MPI_Recv(&rows, 1, MPI_INT, source, mtype, MPI_COMM_WORLD, &status);
    count = rows*NCA;
    MPI_Recv(&a, count, MPI_DOUBLE, source, mtype, MPI_COMM_WORLD, &status);
count = NCA*NCB;
MPI_Recv(&b, count, MPI_DOUBLE, source, mtype, MPI_COMM_WORLD, &status);

for (k=0; k<NCB; k++)
  for (i=0; i<rows; i++) {
    c[i][k] = 0.0;
    for (j=0; j<NCA; j++)
      c[i][k] = c[i][k] + a[i][j] * b[j][k];
  }

mtype = FROM_WORKER;
MPI_Send(&offset, 1, MPI_INT, MASTER, mtype, MPI_COMM_WORLD);
MPI_Send(&rows, 1, MPI_INT, MASTER, mtype, MPI_COMM_WORLD);
MPI_Send(&c, rows*NCB, MPI_DOUBLE, MASTER, mtype, MPI_COMM_WORLD);

} /* end of worker */
Reduce and scan are called *collective* operations
Reduce/scan apply to nodes, not values
Basic operations +, *, min, max, &&, ||
Processor groups simplify collective ops on logical structures like “rows”, “leaves”, etc
MPI allows user-defined scans … these have probably never been used!
**Bottom Line:** Message passing is painful to use but it works … which makes it a solution of choice
The P-threads library, designed for concurrency, is now also used for parallelism

Sharing is implemented by referencing shared memory
- As mentioned, the memory not sequentially consistent
- Not CTA; P-threads use RAM performance model, a greater concern as latencies have increased
- Tends to promote very fine-grain sharing (recall count_3s example), which limits the work that can be used to amortize the overhead costs such as thread creation, scheduling, etc.
- Scaling potential is limited

Writing threaded code using CTA principles usually gives good results
Threading Is Subtle

- It is difficult to get threaded programs right
  - Programmers are responsible for protecting all data references
  - Avoiding deadlock requires discipline and care -- and mistakes are easy to make, especially when optimizing
  - Timing errors can remain latent for a very long time before emerging

Main difficulties: Lots of work for small ||ism; poor scaling prospects
# define NUMTHRDS 4
double sum;
double a[256], b[256];
int status;
int n = 256;
pthread_t thds[NUMTHRDS];
pthread_mutex_t mutex_sum;

int main ( int argc, char *argv[] ) {

    void *dotprod ( void *arg);
    int main ( int argc, char *argv[] ) {
        int i;
        pthread_attr_t attr;
        for ( i = 0; i < n; i++ ) {
            a[i] = i * 0.5;
            b[i] = i * 2.0;
        }

Creating Data
```c
pthread_mutex_init ( &mutex_sum, NULL );
pthread_attr_init ( &attr );
pthread_attr_setdetachstate ( &attr, PTHREAD_CREATE_JOINABLE );

for ( i = 0; i < NUMTHRDS; i++ ) {
    pthread_create ( &thds[i], &attr, dotprod, ( void * ) i );
}

pthread_attr_destroy ( &attr );
for ( i = 0; i < NUMTHRDS; i++ ) {
    pthread_join ( thds[i], ( void ** ) &status );
}

printf ( " Sum = %f\n", sum );
pthread_mutex_destroy ( &mutex_sum );
pthread_exit ( NULL );
return 0;
```
void *dotprod ( void *arg ) {
    int i, my_first, my_last, myid;
    double sum_local;
    myid = ( int ) arg;
    my_first = myid * n / NUMTHRDS;
    my_last = ( myid + 1 ) * n / NUMTHRDS;

    sum_local = 0;
    for ( i = my_first; i <= my_last; i++ ) {
        sum_local = sum_local + a[i] * b[i];
    }

    pthread_mutex_lock ( &mutex_sum );
    sum = sum + sum_local;
    pthread_mutex_unlock ( &mutex_sum );

    pthread_exit ( ( void * ) 0 );
}
OpenMP

- Developed as easy access to multi-threading
- Has second life with multi-core (Intel and others push)
- Approach
  - Add pragmas to C or Fortran code
  - Pragma-aware compiler links in appropriate library calls
  - Pragma-unaware compiler -- no change from sequential
  - All responsibility for parallel == sequential left to programmer
- Main benefit: little effort, some benefit
- Main liability: tight binding to sequential semantics
The program is sequential
- When there is no compiler to interpret the pragmas, the code is sequential
- When there is no parallelism available, the sequential code runs
- When there is a compiler AND parallel processors the sequential code runs

But, we often observe that there IS usually a conceptual difference between sequential and parallel algorithms
double dotProduct() {
    int I; double sum_p;
    double result = 0;
    #pragma omp parallel shared(a, b, result) private(sum_p)
    {
        sum_p=0;
        #pragma omp parallel for private(i)
        for(i=0; i<n; i++) {
            sum_p += a[i]*b[i];
        }
        #pragma omp critical
        {
            result += sum_p;
        }
    }
    return result;
}
OpenMP Compiler

- 4 Processor Sun Enterprise running the NAS PB written in C with OpenMP

- Block Tridiagonal Conjugate Gradient
- Embarrassingly Parallel
- Fast Fourier Transform
- Integer Sort
- LU Decomposition
- Multigrid Iteration
- Sparse Matrix-Vector

<table>
<thead>
<tr>
<th>Program</th>
<th>Class</th>
<th>1 thread</th>
<th>2 threads</th>
<th>4 threads</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT</td>
<td>W</td>
<td>119.19 (1.00)</td>
<td>61.28 (1.95)</td>
<td>36.65 (3.25)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>2900.02 (1.00)</td>
<td>1546.70 (1.87)</td>
<td>1024.93 (2.83)</td>
</tr>
<tr>
<td>CG</td>
<td>W</td>
<td>14.61 (1.00)</td>
<td>6.05 (2.41)</td>
<td>3.12 (4.68)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>49.65 (1.00)</td>
<td>26.01 (1.91)</td>
<td>15.14 (3.28)</td>
</tr>
<tr>
<td>EP</td>
<td>W</td>
<td>33.36 (1.00)</td>
<td>16.74 (1.99)</td>
<td>8.45 (3.95)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>267.39 (1.00)</td>
<td>133.73 (2.00)</td>
<td>67.98 (3.93)</td>
</tr>
<tr>
<td>FT</td>
<td>W</td>
<td>6.07 (1.00)</td>
<td>3.20 (1.90)</td>
<td>1.85 (3.28)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>113.96 (1.00)</td>
<td>60.55 (1.88)</td>
<td>34.73 (3.28)</td>
</tr>
<tr>
<td>IS</td>
<td>W</td>
<td>0.76 (1.00)</td>
<td>0.47 (1.62)</td>
<td>0.38 (2.00)</td>
</tr>
<tr>
<td></td>
<td>A(*1)</td>
<td>17.05 (1.00)</td>
<td>9.25 (1.84)</td>
<td>5.81 (2.93)</td>
</tr>
<tr>
<td>LU</td>
<td>W</td>
<td>194.90 (1.00)</td>
<td>101.42 (1.92)</td>
<td>54.43 (3.58)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1810.94 (1.00)</td>
<td>775.63 (2.33)</td>
<td>411.07 (4.41)</td>
</tr>
<tr>
<td>MG</td>
<td>W</td>
<td>13.56 (1.00)</td>
<td>6.58 (2.06)</td>
<td>3.34 (4.06)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>101.29 (1.00)</td>
<td>50.68 (2.00)</td>
<td>26.67 (3.80)</td>
</tr>
<tr>
<td>SP</td>
<td>W</td>
<td>329.05 (1.00)</td>
<td>175.04 (1.88)</td>
<td>110.83 (2.97)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>2127.84 (1.00)</td>
<td>1157.58 (1.84)</td>
<td>762.07 (2.79)</td>
</tr>
</tbody>
</table>
Critique of OpenMP

- The easy cases work well; harder cases are probably much harder
- Requires that the semantics of sequential computation be preserved
  - Directly opposite of our thesis in this course that algorithms must be rethought
  - Compilers must enforce the sequentially consistent memory model
  - Limited abstractions
HPF: High Performance Fortran

- Philosophy
  - Automatic parallelization won’t work
  - For data parallelism, what’s important is data placement and data motion
  - Give the compiler help:
    - Extends Fortran with directives to guide data distribution
    - Allow slow migration from legacy codes
      - The directives are only hints
- Basic idea
  - Processors operate on only part of overall data
  - Directives say which processor operates on which data
  - Much higher level than message passing
HPF History

The beginning

- Designed by large consortium in the early 90’s
- Participation by academia, industry, and national labs
  - All major vendors represented
    - Convex, Cray, DEC, Fujitsu, HP, IBM, Intel, Meiko, Sun, Thinking Machines
- Heavily influenced by Fortran-D from Rice
  - D stands for “Data” or “Distributed”
- HPF 2.0 specified in 1996
Context

- Part of early 90’s trend towards consolidating supercomputing research
- To reduce risk, fund a few large projects rather than a lot of small risky projects
- Buoyed by the success of MPI
- Aware of the lessons of vectorizing compilers
  - Compilers can train programmers by providing feedback
Vectorizing Compilers

- Basic idea
  - Instead of looping over elements of a vector, perform a single vector instruction
  - Example
    ```c
    for (i=0; i<100; i++)
    A[i] = B[i] + C[i];
    ```

- Scalar code
  - Execute 4 insts 100 times, 2 Loads, 1 Add, 1 Store

- Advantages?

Vector code
- Execute 4 instructions once
  - 2 vector Loads
  - 1 vector Add
  - 1 vector Store
1. Avoid conditionals in loops

```c
for (i=0; i<100; i++)
    if (A[i] > MaxFloat)
        A[i] = MaxFloat;
```

```c
for (i=0; i<100; i++)
    A[i] = min(A[i], MaxFloat)
```

2. Promote scalar functions

```c
for (i=0; i<100; i++)
    foo (A[i], B[i]);
```

```c
for (i=0; i<100; i++)
    Foo(A, B);
    // One function call
    // Body of this function call can be easily vectorized
```

- Lots of function calls inside a tight loop
- Function call boundaries inhibit vectorization
Guidelines for Writing Vectorizable Code (cont)

- 3. Avoid recursion
- 4. Choose appropriate memory layout
  - Depending on the compiler and the hardware, some strides are vectorizable while others are not
- Other guidelines?
- The point
  - These are simple guidelines that programmers can learn
  - The concept of a vector operation is simple
A community project

- Compiler directives don’t change the program’s semantics
- They only affect performance
- Allows different groups to conduct research on different aspects of the problem
- Even the “little guy” can contribute
An array language

- Can operate with entire arrays as operands
  - Pairwise operators
  - Reduction operators

- Uses slice notation
  - `array1d(low: high: stride)` represents the elements of `array1` starting at `low`, ending at `high`, and skipping every `stride-1` elements
  - The stride is an optional operand

- Converts many loops into array statements
Jacobi Iteration

The elements of an array, initialized to 0.0 except for 1.0’s along its southern border, are iteratively replaced with the average of their 4 nearest neighbors until the greatest change between two iterations is less than some epsilon.
**Example**

The following statement computes the averaging step in the Jacobi iteration:

$$next(2:n, 2:n) = \frac{curr(1:n-1, 2:n) + curr(3:n+1, 2:n) + curr(2:n, 1:n-1) + curr(2:n, 3:n+1)}{4}$$
Block Data Distribution

- Block distribution of 1D array

```hpf
!HPF$ PROCESSORS PROCS(4)
!HPF$ DISTRIBUTE array1D(BLOCK) ONTO PROCS
```

- Block distribution of 2D array

```hpf
!HPF$ PROCESSORS PROCS(4)
!HPF$ DISTRIBUTE array2D(BLOCK,BLOCK) ONTO PROCS
```
Block-Cyclic Data Distribution

- Block-cyclic distribution

!HPF$ PROCESSORS PROCPS(4)
!HPF$ DISTRIBUTE array2D(BLOCK, CYCLIC) ONTO PROCPS

Block-cyclic distribution

!HPF$ DISTRIBUTE array2D(CYCLIC, BLOCK) ONTO PROCPS
Arrays can be aligned with one another

- Aligned elements will reside on the same physical processor
- Alignment can reduce communication
- Can align arrays of different dimensions

!HPF$ ALIGN a (i) WITH b(i-1)
This alignment and assignment require all elements to be communicated to a different processor.

!HPF$ ALIGN a(i) WITH b(i-1)  
a(1:n) = b(1:n)

The following induces no communication.

!HPF$ ALIGN a(i) WITH b(i)
Independent Loops

- **INDEPENDENT directive**
  - Loop iterations are independent
  - No implied barriers

```fortran
!HPF$ INDEPENDENT
DO (i = 1:3)
  a(i) = b(i)
  c(i) = d(i)
END DO
```

Fortran90 equivalent?
- None
FORALL Loops vs. Independent Loops

- Is there a difference?

FORALL

INDEPENDENT
Evaluation

- Your thoughts on HPF?
  - Is this a convenient language to use?
  - Can programmers get good performance?
- No performance model
  - To understand locality and communication, need to understand complex interactions among distributions
    - Procedure calls are particularly bad
  - Many hidden costs
  - Small changes in distribution can have large performance impact

Does the following code induce communication?

\[ a(i) = b(i) \]
No performance model

- Complex language $\implies$ Difficult language to compile
  - Large variability among compilers
  - Kernel HPF: A subset of HPF “guaranteed” to be fast

An accurate performance model is essential

- Witness our experience with the PRAM

Common user experience

- Play with random different distribution in an attempt try to get good performance
Language is too general

- Difficult to obey an important system design principle:
  “Optimize the common case”
  - What is the common case?

- Sequential constructs inherited from Fortran77 and Fortran90 cause problems
  - For example, the following code forces compiler to perform matrix transpose
    ```fortran
    FORALL (i=1:n, j=1:n)
    a(i, j) = a(j, i)
    END FORALL
    ```
Philosophy

- Provide performance portability for data-parallel programs
- Allow users to reason about performance
- Start from scratch
  - Parallel is fundamentally different from sequential
  - Be willing to throw out conveniences familiar to sequential programmers

Basic idea

- An array language
- Implicitly parallel
ZPL History

The beginning

- Designed by a small team beginning in 1993
- Compiler and runtime released in 1997

Claims

- Portable to any MIMD parallel computer
- Performance comparable to C with message passing
- Generally outperforms HPF
- Convenient and intuitive
Jacobi Iteration

The elements of an array, initialized to 0.0 except for 1.0’s along its southern border, are iteratively replaced with the average of their 4 nearest neighbors until the greatest change between two iterations is less than some epsilon.
program Jacobi;
config  var n : integer = 512;
        epsilon : float = 0.00001;
region    R = [1..n, 1..n];
var     A, Temp : [R] float;
        err : float;

direction north = [-1, 0];
        south = [1, 0];
        east  = [0, 1];
        west  = [0, -1];

procedure Jacobi();
[R] begin
    A := 0.0;
    [north of R] A := 0.0;  [west of R]  A := 1.0;
    [east of R]  A := 0.0;  [south of R] A := 0.0;
    repeat
        Temp := (A@north + A@east + A@west + A@south)/4.0;
        err  := max<< abs(Temp – A);
        A    := Temp;
    until err < epsilon;
end;
end;
program Jacobi;
config var n : integer = 512;
  epsilon : float = 0.00001;
region R = [1..n, 1..n];
var A, Temp : [R] float;
  err : float;
direction north = [-1, 0];
  south = [ 1, 0];
  east = [ 0, 1];
  west = [ 0, -1];
procedure Jacobi();
  [R] begin
    A := 0.0;
    [north of R] A := 0.0;
    [west of R]  A := 1.0;
    [east of R]  A := 0.0;
    [south of R] A := 0.0;
    repeat
      Temp := (A@north + A@east + A@west + A@south)/4.0;
      err := max<< abs(Temp - A);
      A := Temp;
    until err < epsilon;
  end;
end;
program Jacobi;
config   var n : integer = 512;
          epsilon : float = 0.00001;
region   R = [1..n, 1..n];
var     A, Temp : [R] float;
        err : float;
direction north = [-1, 0];     south = [ 1,  0];
     east = [ 0, 1];     west = [ 0, -1];
procedure Jacobi();
  [R] begin
    A := 0.0;
  [north of R] A := 0.0; [west of R] A := 1.0;
  [east of R] A := 0.0; [south of R] A := 0.0;
  repeat
    Temp := (A@north + A@east + A@west + A@south)/4.0;
    err := max<< abs(Temp - A);
    A := Temp;
  until err < epsilon;
end;
end;
program Jacobi;
config var n : integer = 512;
  epsilon : float = 0.00001;
region R = [1..n, 1..n];
var A, Temp : [R] float;
err : float;
direction north = [-1, 0];
south = [1, 0];
east = [0, 1];
west = [0, -1];
procedure Jacobi();
[R] begin
  A := 0.0;
  [north of R] A := 0.0;
  [west of R] A := 1.0;
  [east of R] A := 0.0;
  [south of R] A := 0.0;
  repeat
    Temp := (A@north + A@east + A@west + A@south)/4.0;
    err := max<< abs(Temp - A);
    A := Temp;
    until err < epsilon;
end;
end;
program Jacobi;
config var n : integer = 512;
  epsilon : float = 0.00001;
region R = [1..n, 1..n];
var A, Temp : [R] float;
  err : float;
direction north = [-1, 0];  south = [ 1, 0];
  east = [ 0, 1];  west = [ 0, -1];
procedure Jacobi();
  [R] begin
    A := 0.0;
    [north of R] A := 0.0;  [west of R]  A := 1.0;
    [east of R]  A := 0.0;  [south of R] A := 0.0;
    repeat
      Temp := (A@north + A@east + A@west + A@south)/4.0;
      err := max<< abs(Temp - A);
      A := Temp;
    until err < epsilon;
  end;
end;

:= ( + + + ) / 4.0
**Recent Notable Efforts: PGAS**

- Greatest potential to assist programmer comes from hiding communication calls
  - Compilers can generate the calls
  - Need interface to specify which are local/global

- Concept: Partitioned Global Address Space
  - Overlay global addressing on separate memories
  - PGAS tends to use 1-sided comm as simplification
Three PGAS languages

<table>
<thead>
<tr>
<th>CAF</th>
<th>UPC</th>
<th>Ti</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Array Fortran</td>
<td>Universal Parallel C</td>
<td>Titanium</td>
</tr>
<tr>
<td>Numrich &amp; Reed</td>
<td>El Ghazawi, Carlson &amp; Draper</td>
<td>Yelick</td>
</tr>
<tr>
<td>Extends Fortran</td>
<td>Extends C</td>
<td>Extends Java</td>
</tr>
</tbody>
</table>

Developed around 2000 +/- & Implemented

- Similarities: GAS, comm handled by compiler/rt, programmer controls work/data assignment
- Differences: Most everything else
Incredibly elegant (for Fortran) extension

real, dimension(n,n)[p,*]: a,b,c
...
do k=1,n
  do q=1,p
    c(i,j)[myP,myQ]=c(i,j)[myP,myQ]+a(i,k)[myP,q]*b(k,j)[q,myQ]
  enddo
enddo
Data can be allocated either shared or private; shared is assigned *cyclically* or BC.

Pointers are an issue

<table>
<thead>
<tr>
<th>Property of pointer</th>
<th>Private</th>
<th>Shared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td>Private-Private, p1</td>
<td>Private-Shared, p2</td>
</tr>
<tr>
<td>Shared</td>
<td>Shared-Private, p3</td>
<td>Shared-Shared, p4</td>
</tr>
</tbody>
</table>

int *p1; /* private ptr pointing locally */
shared int *p2; /* private ptr pointing into shared space */
int *shared p3; /* shared ptr pointing locally */
shared int *shared p4; /* shared ptr pointing into shared space */
shared int v1[N], v2[N], v1v2sum[N];

void main()
{
    int i;
    shared int *p1, *p2;
    p1=v1;
    p2=v2;
    upc_forall(i=0; i<N; i++, p1++, p2++;i)
    {
        v1v2sum[i] = *p1 + *p2;
    }
}
Java extensions including

- “regions, which support safe, performance-oriented memory management as an alternative to garbage collection.”
- `foreach` is an unordered iteration, which logically raises the concurrency:
  ```java
  foreach ( ... ) {
  }
  ```
- Used with the concept of a `point`, tuple of integers that range over a `domain`
public static void matMul(double [2d] a,
        double [2d] b,
        double [2d] c)
{
    foreach (ij in c.domain())
    {
        double [1d] aRowi = a.slice(1, ij[1]);
        double [1d] bColj = b.slice(2, ij[2]);
        foreach (k in aRowi.domain())
        {
            c[ij] += aRowi[k] * bColj[k];
        }
    }
}
Summarizing PGAS Languages

- The languages improve on the alternative--base language + MPI
- Compiler provides significant help, but the need to be attuned to subtle detail remains
- Deep issues
  - Global address space+private are good, but how they “play together” remains unclear
  - Better abstractions to reduce detail
DARPA has supported three new “high productivity” parallel languages
- Is productivity really the issue?
- Project coupled with design of a new machine

The final competitors:
- Cray’s Cascade High Productivity Language, Chapel
- IBM’s X10
- Sun’s Fortress
Chapel is a multithreaded language supporting

- Data ||ism, task ||ism, nested ||ism
- Optimizations for locality of data and computation
- Object oriented and generic programming techniques
- Parallel implementation is nearing completion

Designed for experts, production programmers
for(str, span) in genDFTPhases(numElements, radix) {
    forall (bankStart, twidIndex) in (ADom by 2*span, 0..) {
        var wk2 = W(twidIndex),
        wk1 = W(2*twidIndex),
        wk3 = (wk1.re - 2 * wk2.im * wk1.im,
               2 * wk2.im * wk1.re - wk1.im):elemType;
        forall lo in bankStart + [0..str) do
            butterfly(wk1, wk2, wk3, A[[0..radix)*str + lo]);
        wk1 = W(2*twidIndex+1);
        wk3 = (wk1.re - 2 * wk2.re * wk1.im, 2 * wk2.re * wk1.re -
               wk1.im):elemType;
        wk2 *= 1.0i;
        forall lo in bankStart + span + [0..str) do
            butterfly(wk1, wk2, wk3, A[[0..radix]*str + lo]);
    }
}
Developed at Sun, Fortress pushes the envelop in expressivity

- Focus on new programming ideas rather than parallel programming ideas: components and test framework assist with powerful compiler optimizations across libraries
- Textual presentation important -- subscripts and superscripts -- mathematical forms
- Transactions, locality specification, implicit ||ism
- Extendibility
Fortress

- Conjugate gradient program in Fortress

- Features
  - `:= /=`
  - Sequential
  - Mathematical

```
conjGrad[Elt extends Number, nat N,
    Mat extends Matrix [Elt, N × N],
    Vec extends Vector [Elt, N]
](A: Mat, x: Vec):(Vec, Elt)

cgit μax = 25
z : Vec = 0
r : Vec = x
p : Vec = r
r : Elt = r^T r
for j ← seq(1: cgit_{max}) do
  q = Ap
  α = \frac{p^T q}{p^T q}
  z := z + α p
  r := r − α q
  ρ₀ = ρ
  ρ := r^T r
  β = \frac{ρ}{ρ₀}
  p := r + β p
end
(z, ||x − A z||)
```
IBM’s X10 is a type safe, distributed object oriented language in the PGAS family -- its “accessible to Java programmers”

Many goodies including regions (a la ZPL), places (for locality), asynch, futures, foreach, ateach, atomic blocks and global manipulation of data structures
public class Jacobi {
const int N=6;
const double epsilon = 0.002;
const double epsilon2 = 0.000000001;
const region R = [0:N+1, 0:N+1];
const region RInner= [1:N, 1:N];
const distribution D = distribution.factory.block(R);
const distribution DInner = D | RInner;
const distribution DBoundary = D - RInner;
const int EXPECTED ITERS=97;
const double EXPECTED ERR=0.0018673382039402497;
double[D] B = new double[D] (point p[i,j])
    {  return DBoundary.contains(p)
          ? (N-1)/2 : N*(i-1)+(j-1); 
    };
public double read(final int i, final int j) {
    return future(D[i,j]) B[i,j].force(); }
public static void main(String args[])
    { boolean b= (new Jacobi()).run();
      System.out.println("++++++ " + (b? "Test succeeded." :"Test failed.");
      System.exit(b?0:1);
    }
public boolean run() {
    int iters = 0;
    double err;
    while(true) {
        double[] Temp =
            new double[DInner] (point [i,j])
            {return (read(i+1,j)+read(i-1,j)
                +read(i,j+1)+read(i,j-1))/4.0; }; 
        if((err=((B | DInner) - Temp).abs().sum()) < epsilon)
            break;
        B.update(Temp);
        iters++;
    }
    System.out.println("Error="+err);
    System.out.println("Iterations="+iters);
    return Math.abs(err-EXPECTED ERR) < epsilon2 && iters==EXPECTED ITERS;
}
Summary

- Language is key tool to express parallelism
- State of the art is libraries –
  - threads, message passing, OpenMP
- There has been tremendous experimentation with alternative language approaches
  - ZPL, HPF, CAF, UPC, Titanium
- The next generation is here
  - Chapel, X10, Fortress
Using online research become familiar with a parallel programming language and critique it

- NOT allowed: ZPL, Chapel, libraries
- The critique must include a small code example
- Relevant topics to discuss might include
  - Execution model (data parallel, task, etc.), mem model
  - Mechanisms for creating threads, communicating, etc.
  - Brief history, if known
  - Evidence of performance, scalability, portability, etc.
- Any length OK, but ~2 pages is intended scale; refs