Parallel MATLAB
From “Hell No” to “You Bet”

Cleve Moler
Thirty Years of Parallel Computing at Argonne
May 14, 2013
Early Days at Argonne
Early Days at Argonne

```math
:_a=matrix(2,2:1,2,3,4) ; a
  A (A 2 by 2 Matrix)
  1 2
  3 4
:_a*a
  A*A (A 2 by 2 Matrix)
  7 10
  15 22
:_a/a
  A/A (A 2 by 2 Matrix)
  1 0
  0 1
:_aa=array(2,2:1,2,3,4)
:_aa*aa
  AA*AA (A 2 by 2 Array)
  1 4
  9 16
:_aa/aa
  AA/AA (A 2 by 2 Array)
  1 1
  1 1
```
Early Days at Argonne
1984
MathWorks Founded
1995
Why there isn’t a parallel MATLAB

Our experience has made us skeptical

by Cleve Moler

There actually have been a few experimental versions of MATLAB for parallel computers. None of them has been effective enough to justify development beyond the experimental prototype. But we have learned enough from these experiences to make us skeptical about the viability of a fully functional MATLAB running on today’s parallel machines. There are three basic difficulties:

- Memory model
- Granularity
- Business situation

MATLAB is a lot bigger, and parallel computers are a lot faster. But distributed memory is still a fundamental difficulty. One of MATLAB’s most attractive features is its memory model. There are no declarations or allocations—it is all handled automatically. The key question is: Where are the matrices stored? It is still true today that any matrix that fits into the host memory should probably stay there.

Granularity

A little over five years ago, we had a parallel MATLAB on a shared memory multiprocessor, the Ardent Titan, but we didn’t tell the world about it. The most effective use of this machine, as well as today’s multiprocessor workstations, is already done automatically by the operating system. MATLAB should run on only one processor, while other tasks, like the X-Windows server, use the other processors. In typical use, MATLAB spends only a small portion of its time in routines that can be parallelized, like the ones in the math library. It spends much more time in places like...
1995
Why there isn’t a parallel MATLAB

- Memory model
  - Where are the matrices stored?

- Granularity
  - Need outer loop parallelism.

- Business situation
  - Not enough potential customers.
By 2005

- MIT website lists 20 “Parallel MATLABs”.
- Front ends to fixed libraries on back end parallel machines.
2005, Householder XVI
Seven Springs, Pennsylvania
2005, Householder XVI Announcement
R2006a, Parallel MATLAB

- “labs”
- numlabs
- labindex
- parfor
- darray
- labsend
- labreceive
R2013a, Parallel Computing Toolbox

matlabpool distributed.xxx
numlabs codistributed.xxx
labindex gop
codistributed.xxx
gop
gplus
spmd gcat
batch gather
wait parcluster
labSend labBroadcast
labReceive labBarrier
labSendReceive labProbe
Blackjack Demo

Monte Carlo simulation.
Prototype of financial simulations.
Run multiple simulations.
Almost, but not quite, embarrassing parallel.
% BLACKJACKDEMO  Parallel blackjack demo.

p = 4;       % Number of players. 
n = 25000;   % Number of hands per player. 
B = zeros(n,p);

tic
parfor k = 1:p
    B(:,k) = blackjacksim(n);
end
toc

plot(B)

r = sum(B(n,1:p))/(10*n*p);
title(sprintf('Blackjack, return = %8.4f',r))
xlabel('Number of hands')
ylabel('Winnings ($)')
axis([0 n -2500 2500])
Blackjack, return = -0.0014
Talk about …

- Thanks to MPICH
Talk about …

- Almost all usage is parfor
Talk about …

- Powerful workers obviate need for distributed arrays
Talk about …

- Multithreading
  - Fine grained parallelism
Talk about …

- GPUs
  - JIT
Talk about …

- Job managers rule
  - File systems, security, privacy, …
Talk about …

- Shared facilities preclude interactivity
  - Argonne’s Jazz