pMatlab Takes the HPCchallenge

Ryan Haney, Hahn Kim, Andrew Funk, Jeremy Kepner, Charles Rader, Albert Reuther, Nadya Travinin

Motivation

- The DARPA HPCS program has created the HPCchallenge benchmark suite in an effort to redefine how we measure productivity in the HPC domain
- MATLAB® is the primary high level language used within the signal processing community; increasingly used for
  - large system simulations
  - processing data in the field

Goals

- Implement and analyze the performance of HPCchallenge benchmarks using pMatlab
- Optimize and add functionality to the pMatlab toolbox
- Compare traditional C/MPI with MATLAB using global array semantics. Measurements of productivity include:
  - **Maximum problem size:** Largest problem that can be solved or fit into memory
  - **Execution performance:** Run-time performance of the benchmark
  - **Code size:** Software lines of code (SLOC) required to implement the benchmark

- pMatlab implements global array semantics in MATLAB
  - Global array semantics allow indexing and general element access for distributed data
- Implementing the HPCchallenge benchmarks using pMatlab allows a unique opportunity to explore the merits of pMatlab with respect to high performance embedded computing

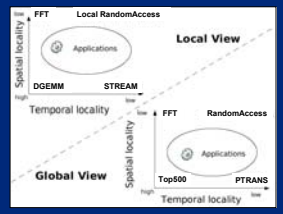
Local Benchmarks

- DGEMM (matrix x matrix multiply)
- STREAM
  - COPY, SCALE, ADD, TRIAD
- RandomAccess
- FFT

Global Benchmarks

- Top500 (High Performance LINPACK)
- PTRANS — parallel matrix transpose
- RandomAccess
- FFT

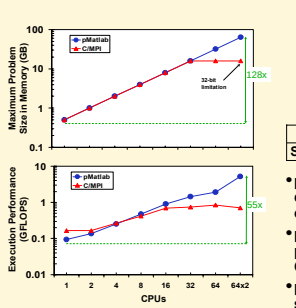
HPCchallenge

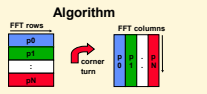


HPCchallenge Relevance to HPEC

- Four key benchmarks have significant relevance to HPEC
  - **FFT:** Distributed corner turn and FFTs important in multi-sensor signal processing
  - **RandomAccess:** Random data accesses typical of "post detection" operations
  - **Top500:** Matrix-matrix multiplies typical of multi-element beamforming
  - **STREAM:** Distributed vector operations common to signal processing
- Multiple implementations
  - C/Fortran, C/Fortran+MPI, MATLAB, pMatlab

FFT Results

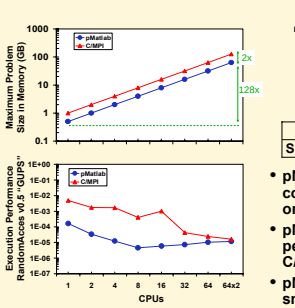


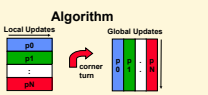


Software Code Size	C/MPI	pMatlab	Ratio
SLOC	2509	72	35

- pMatlab memory scalability comparable to C/MPI (128x on 128 CPUs)
- pMatlab execution performance comparable to C/MPI (55x on 128 CPUs)
- pMatlab code size is 35x smaller than C/MPI

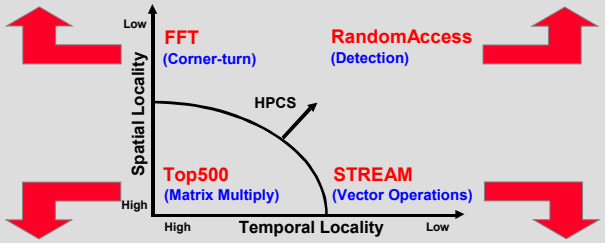
RandomAccess (v0.5) Results



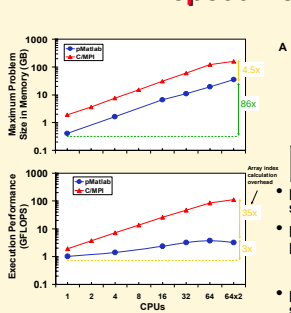


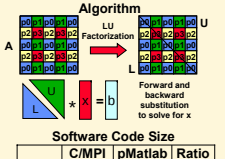
Software Code Size	C/MPI	pMatlab	Ratio
SLOC	603	101	6

- pMatlab memory scalability comparable to C/MPI (128x on 128 CPUs)
- pMatlab execution performance comparable to C/MPI
- pMatlab code size is 6x smaller than the C/MPI implementation



Top500 Results

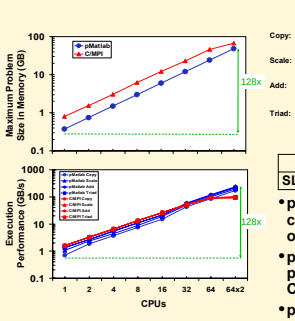


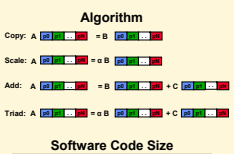


Software Code Size	C/MPI	pMatlab	Ratio
SLOC	15,561	235	66

- pMatlab maximum problem size scales 86x on 128 CPUs
- pMatlab execution performance scales 3x
  - Removing index calculation overhead will significantly improve
- pMatlab code size is 66x smaller than C/MPI implementation

STREAM Results



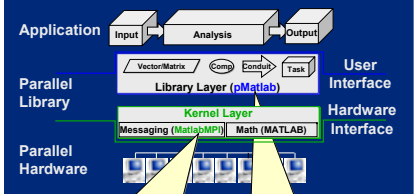


Software Code Size	C/MPI	pMatlab	Ratio
SLOC	441	51	8

- pMatlab memory scalability comparable to C/MPI (128x on 128 CPUs)
- pMatlab execution performance comparable to C/MPI (128x on 128 CPUs)
- pMatlab code size is 8x smaller than C/MPI

pMatlab

pMatlab Software Architecture



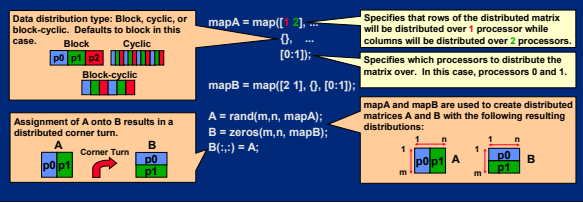
- Can build a parallel library with a few messaging primitives
- MatlabMPI provides this messaging capability:

```
MPI_Send(dest,comm,tag,X);
X = MPI_Recv(source,comm,tag);
```

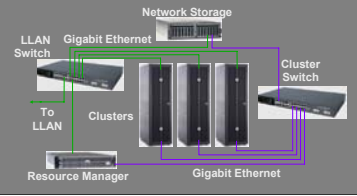
- Can build applications with a few parallel structures and functions
- pMatlab provides parallel arrays and functions

```
X = ones(n,mapX);
Y = zeros(n,mapY);
Y(:,:) = fft(X);
```

pMatlab Goal: Maps and Distributed Matrices



Benchmark Platform



Lincoln Laboratory LLGRID

- 80 Node dual 2.8 & 3.06 GHz Xeon (P4)
- 400 & 533 MHz front-side bus
- 4 GB RAM memory per node
- Two 36 GB hard drives per node10/100 Mgmt Ethernet interface
- Two Gig-E Intel interfaces
- 1.02 TB Network Storage
- Running Red Hat Linux

Conclusions

Benchmark Results Summary

- Memory scalability comparable to C/MPI on nearly all of HPCchallenge (for 128 CPUs). Allows MATLAB users to work on much larger problems.
- Execution performance comparable to C/MPI on nearly all of HPCchallenge (for 128 CPUs). Allows MATLAB users run their programs much faster.
- Code size much smaller. Allows MATLAB users to write programs much faster than C/MPI
- pMatlab allows MATLAB users to effectively exploit parallel computing, and can achieve performance comparable to C/MPI.

HPCchallenge Benchmark Results: C/MPI vs. pMatlab

	Maximum Problem Size	Execution Performance	Code Size: C/MPI to pMatlab ratio
Random Access	Comparable (128x)	Comparable	6x
Top500	pMatlab (86x), C/MPI (83x)	pMatlab (3x), C/MPI (35x)	66x
FFT	Comparable (128x)	Comparable (55x)	35x
STREAM	Comparable (128x)	Comparable (128x)	8x

This work is sponsored by Defense Advanced Research Projects Administration, under Air Force Contract F19628-00-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government.