Performance Analysis and Comparison of MPI, OpenMP and Hybrid NPB - MZ

Héctor J. Machín Xingfu Wu Valerie E. Taylor

Department of Computer Science, Texas A&M University

Motivation and Research Goal

- Chip multiprocessors (CMP)
  - –Widely used for high performance computing.
- CMP Clusters
  - –Provide natural programming paradigm for hybrid programs (using both OpenMP and MPI).
    - –MPI between nodes and OpenMP within a node.
- Question Addressed: Can current hybrid programming paradigms such as hybrid MPI/OpenMP efficiently exploit the potential of CMP clusters?
- Method:
  - –Analyze and compare the performance of: MPI, OpenMP, and hybrid NPB-MZ
  - –Use two systems: SDSC DataStar p655 and TAMU Hydra

Background

Message Passing Interface (MPI)

- Message Passing Programming Model
  - –send and receive messages
  - –processes

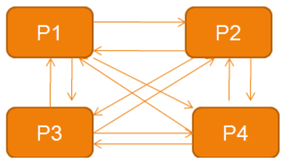


Fig 1. Communication between 4 MPI processes

OpenMP

- Shared Memory Programming Model
  - –fork-and-join execution
  - –threads

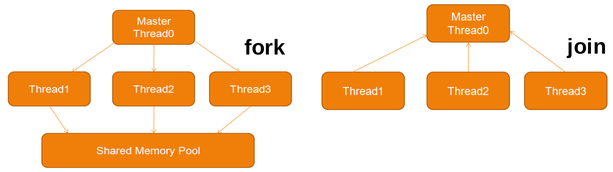


Fig 2. fork-and-join execution of OpenMP threads

NAS Parallel Benchmarks-Multi-Zone (NPB–MZ)

- NAS Parallel Benchmarks (NPB)
  - –testing parallel computers
- Scientific problems feature several levels of parallelism
  - –not reflected property in NPB
  - –NPB Multi-Zone (NPB-MZ) versions were created
- The application benchmarks:
  - –Low-Upper Symetric Gauss-Seidel (LU)
  - –Scalar Penta-Diagonal (SP)
  - –Block Tri-Diagonal (BT)
- Problem sizes varies from:
  - –S,W,A,B,C,D,E
  - –For this research we used A to E

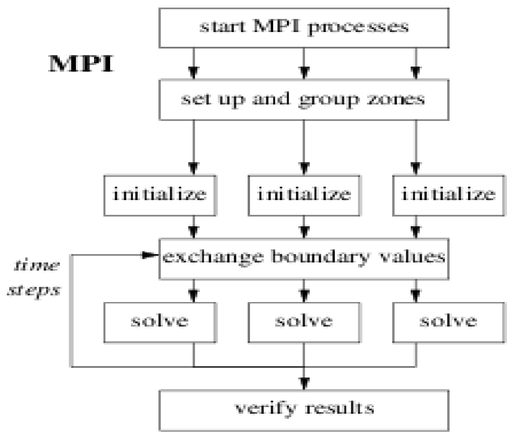


Fig 3. MPI Implementation of NPB-MZ

Platforms

Platform	Total # of processors	CPU type	Operating System	Processors per node	Nodes	Memory per node
DataStar p655	2176	1.5, 1.7GHz Power4	AIX 5.2L	8	272	16, 32GB
Hydra	640	1.9GHz Power5+	AIX 5.3L	16	40	32GB

Table 1. Platforms description

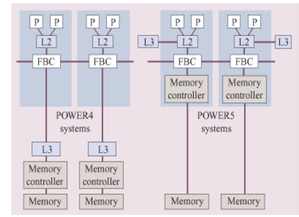
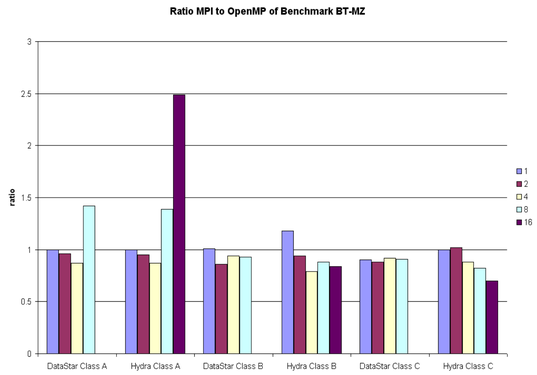


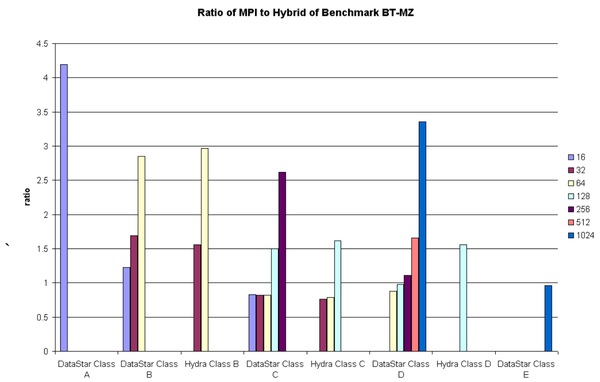
Fig 4. Comparison between Power4 and Power5 chip, IBM

Results and Analysis

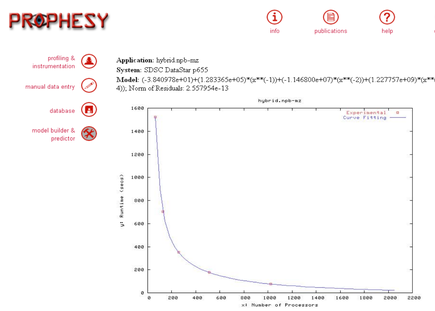
MPI vs. OpenMP Performance



MPI vs. Hybrid Performance



Prophesy Performance Modeling



Processor Partitioning Analysis

8 processors BT-MZ Class B

# nodes	# task_per_node	execution time
8	1	124.99
4	2	125.05
2	4	127.52
1	8	129.37

8 processors SP-MZ Class B

# nodes	# task_per_node	execution time
8	1	92.88
4	2	93.34
2	4	96.79
1	8	104.62

32 processors BT-MZ Class B

# nodes	# task_per_node	execution time
32	1	56.03
16	2	56.36
8	4	56.91
4	8	57.93

32 processors SP-MZ Class B

# nodes	# task_per_node	execution time
32	1	23.23
16	2	24.38
8	4	24.81
4	8	26.02

64 processors BT-MZ Class B

# nodes	# task_per_node	execution time
64	1	56.64
32	2	56.91
16	4	57.34
8	8	58.09

64 processors SP-MZ Class B

# nodes	# task_per_node	execution time
64	1	11.75
32	2	11.77
16	4	12.63
8	8	13.15

Summary

- For most cases:
  - –MPI vs OpenMP
    - MPI exceeds OpenMP performance
  - –MPI vs Hybrid
    - Hybrid exceeds MPI performance

Future Work

- Hardware-level performance analysis with hpmcount for deeper understanding of these results.
- Other HPC and scientific computing applications.