

6th International Workshop in Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS'15)

Evaluating Node Orderings for Improved Compactness

Carl Albing, Ph.D. Dep't. of Computer Science U.S. Naval Academy.

Topics

- Appl Placement → Node Ordering
- Metrics for Placement \rightarrow MIND
- Mean Placement Calculation (MPC)
- Graphing the MPC
- · Interpreting the Results
- A Reality Check

Acknowledgements

This work was supported in part by the DoD HPC Modernization Program. It also would not have been possible without the infrastructure support and help from the accommodating staff at ORNL, NERSC, and NOAA.

Application Placement

- Performance variations tied to placement
- Causes related to:
 - latency (distances)
 - bandwidth (job/job inter.)
- Compact placement
 - better for both
- Utilization suffers



Node Ordering

• General Concept



- Appl Placement
- 3D Torus



Example Node Orderings

- Min dimension first vs. Max
 - torodial vs. snake



Example Node Orderings



Node Ordering

- Many orderings
- How to choose?
- Benchmarks are expensive
- Convert the whole system?
- Any theoretical or empirical help?



Metrics

• Diameter

- which is more compact?

Metrics

• Mean Inter-Node Distance (MIND)

$$MIND(T,S) = \left(\sum_{i=1}^{s-1} \sum_{j=i+1}^{s} d_{ij}\right) / (s(s-1)/2)$$

...............

• Hop-Bytes - not enough data

Mean Placement Calculation

- For a Given Job Size place it starting at 1st location in list
- · Calculate MIND for job
- Re-place, re-MIND at each loc.
- Take the Mean
- That's 1 value
- Next size



Using MPC

- Each MPC Yields a Curve
- Specific to a Machine Size/Shape
- Three real systems:

site	Х	Y	Z	nodes
NOAA	7	12	16	1316
NERSC	17	8	24	6528
ORNL	25	32	24	18944

................

Results

- Three Systems
- "small", "med", large HPC
 Use in real placement sequence

...............

Comparing Node Orderings

MPC for NOAA System (7x12x16)



A Closer Look

MPC for NOAA System (7x12x16)



"Medium" Size HPC System MPC for NERSC System (17x8x24) 6 5 4 MPC (hops) 3 2 2Cube BBE hilbert 1 MDF **MDFbph** MaxDF nidorder

job size (nodes)

40

50

60

30

0

0

10

20

70

MPC Over the Full Range

MPC for NERSC System (17x8x24)





Large System – Full Range

MPC for ORNL System (25x32x24)



Testing the Results

- Real World?
- NERSC system
- 1 month of placements
- Mean per job size; > 5 jobs



Actual Jobs, Theoretical Placement



Summary

- Node Ordering, MIND
- Mean Placement Calculation (MPC)
 - -System Size, "Shape"
 - -Orderings favor Job Sizes
- Rough predictor
- Larger data set coming
- Other aspects of placement