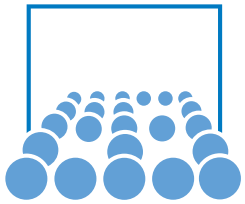
Technische Universität München

Assignment 3: MPI Point-to-Point and One-Sided Communication

Programming of Super Computers

Friedrich Menhorn, Benjamin Rüth, Erik Wannerberg
Team 12

December 12, 2015



Contents

Cannon's algorithm

- explain algorithm
- provided implementation

Baseline

Challenges in getting an accurate baseline and changes to the Load-Leveler batch script.

Scalability

- Compute time scalability with fixed 64 processes and varying size of input files.
- MPI time scalability with fixed 64 processes and varying size of input files.
- Differences in scalability between the Sandy Bridge and Haswell architectures.

MPI Non-Blocking Operations

- Send/Receive
  - **MPI_Isend**
  - **MPI_Irecv**
- Synchronization
  - **MPI_Wait**
  - MPI_Probe

Optimizations

- **What is overlap?**
  We do not wait for either task to be completed, but try to do communication and computation at the same time. We are hiding communication time by doing computation in the meantime.
- **What is the theoretical maximum overlap that can be achieved?**
  Bounds for pure communication time:

$$\max\left(0, T_{\text{MPI}}^{\text{blocking}} - T_{\text{computation}}\right) \leq T_{\text{MPI}}^{\text{non-blocking}} \leq T_{\text{MPI}}^{\text{blocking}}$$

  Overheads:
  - Copying into and from buffers
  - Initialization

  Maximum overlap depends on amount of $T_{\text{computation}}$. As soon as $T_{\text{computation}} > T_{\text{MPI}}^{\text{blocking}}$, we can theoretically achieve 100% overlap.

Optimizations (cond.)

- **Was communication and computation overlap achieved?**

Scaling

- Was a speedup observed versus the baseline?
- Were there any differences between Sandy Bridge and Haswell nodes?

MPI One-Sided Operations

- Initialization
  - **MPI_Win_create**
  - **MPI_Win_free**
- Remote Memory Access
  - **MPI_Put**
  - MPI_Get
  - MPI_Accumulate
- Synchronization
  - **MPI_Win_fence**
  - MPI_Win_post / MPI_Win_start / MPI_Win_complete / MPI_Win_wait
  - MPI_Win_lock / MPI_Win_unlock

Optimizations

- Was communication and computation overlap achieved?

Scaling

- Was a speedup observed versus the baseline?
- Was a speedup observed versus the non-blocking version?
- Were there any differences between Sandy Bridge and Haswell nodes?