Communication Libraries for Parallel-Programming-Model Runtime Systems

Andy Stone

CS653

What I'll Talk About

Three communication libraries

Communication Operations

How they've been tailored for a machine

Context

UEs (task/processes) need to communicate

Why? Synchronize Computation

Pass Data

In a programming model communication can be **implicit** or **explicit**

Example: Set array A on process 1 to array B on process 2

MPI: Explicit

```
if(id == 1) {
  MPI_Recv(A, n, MPI_DOUBLE, 2, 0,
       MPI_COMM_WORLD);
} else if(id ==2) {
  MPI_Send(B, n, MPI_DOUBLE, 1, 0,
              MPI_COMM_WORLD);
}
```

Co-Array Fortran: Implicit

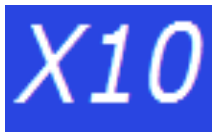
```
if(this_image() == 1) then
    A=B[2]
end if
```

Models and Communication Libraries

Co-Array Fortran

GlobalArrays

CHAPEL



Communication Operations

Classification of communication operations:

Blocking versus **non-blocking**

Two-sided versus **one-sided**

2-sided: send, receive

1-sided: put, get, accumulate

What ARMCI/GASNET has that MPI doesn't:

- Stronger focus on 1-sided communication
- Active Messages (RPC mechanism)
- I/O vector and strided messages
- ARMCI: Sum-and-scale in one accumulate operation

Implementation Optimizations

Implementation Optimizations

J. Nieplocha and B. Carpenter. **Armci: A portable remote memory copy library for distributed array libraries and compiler run-time systems**. In Parallel and Distributed Processing, volume 1586 of Lecture Notes in Computer Science, pages 533{546. Springer Berlin / Heidelberg, 1999.

Compare MPICH against IBM

Experimentally determine the behavior of ISEND/IRECV

Suggest optimization: Add in wait or test commands

Show 27% reduction in communication time for SAMR example

Implementation Optimizations

D. Bonachea, P. H. Hargrove, M. Welcome, and K. Yelick. **Porting gasnet to portals: Partitioned global address space (pgas) language support for the cray XT**. In Cray Users Group, 2009.

Modifying GASNet to use Cray XT's native communication library: Portals

Implementation Optimizations

J. Nieplocha and B. Carpenter. **Armci: A portable remote memory copy library for distributed array libraries and compiler run-time systems**. In Parallel and Distributed Processing, volume 1586 of Lecture Notes in Computer Science, pages 533{546. Springer Berlin / Heidelberg, 1999.

Accumulate is implemented differently on different machines

IBM SP: Owner-computes

Cray T3E: Source-computes

Conclusions

Libraries very similar in terms of one-sided communication and synchronization operations they include.

GASNET/ARMCI Used by programming model implementors

Differences:

GASNET/ARMCI: I/O Vector, Strided

ARMCI: Scale-and-sum accumulate operation

Performance of library implementation dependent, implementations are often tailored to work well with a specific machine.