MPI Forum Hybrid WG

June 3, 2014

Threading Activities

Any opens on this topic?

Ticket 357 is going forward for vote Thursday

- https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/357

Endpoints

- Recently closed:
  - Communicator comparison
  - Group comparison
  - Communicator free semantic
- Up for discussion
  - Address space query
- Other open issues?

Address Space Query

- Find out if MPI processes share an address space
  - Libraries need to determine whether memory and other resources are private to a process
  - Existing issue when MPI processes are implemented as threads
  - Also an issue with endpoints
- https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/425
- Opens:
  - Scope of the query: sharing in my node versus sharing in any node

MPC - From CEA implements processes as threads

int MPI_Has_sharing(int *flag)
int MPI_Comm_has_sharing(MPI_Comm comm, int *flag)
int MPI_Group_has_sharing(MPI_Group group, int *flag)
int MPI_Win_has_sharing(MPI_Win win, int *flag)
int MPI_File_has_sharing(MPI_File file, int *flag)

~~int MPI_Num_endpoints(int *num_ep, int *ep_id)~~
*int MPI_Comm_num_endpoints(MPI_Comm comm, int *num_ep, int *ep_id) -- (George dislikes, Squyres likes, argument is split_type, perform query, free (i. e. Squyring) is silly)*
Comm split type -- MPI_COMM_TYPE_ENDPOINTS

Communicator Split Types

- MPI_COMM_TYPE_NEIGHBORHOOD
  - https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/297
  - https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/372
- ~~MPI_COMM_TYPE_SHARED_UNIFORM~~
  - ~~Better support for shared memory windows on machines~~
  - ~~Sub-communicator that can map shared memo a more uniform access characteristic~~
  - (This is absorbed into MPI_COMM_TYPE_SHARED with an info key)

Coarse-grain control through (small number) split types, fine-grain control through info args

Desired split types:

1. Network topology #297
2. I/O topology #372++
3. Node resources #372
4. Shared memory NUMA domains (MPI_COMM_TYPE_SHARED)

Desired info keys:

???

Error cases need to be explored and handling documented.

Consider MPI_COMM_SPLIT_TYPE_CORE. One process is bound to two cores. You are splitting on a communicator that contains two endpoints in your process. What does it mean? Resources can be shifting around. What do we do in this context?

Init With Assertions/Info

- Considered and declined for MPI 2.2:
  - https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/22
  - https://svn.mpi-forum.org/trac/mpi-forum-web/wiki/InitTimeAssertions
- Fault tolerance toggle
  - MPI_FAULT_TOLERANT
- Persistence
- Threading toggles
  - MPI_THREAD_SINGLE, …, MPI_THREAD_MULTIPLE
- Could use array of strings rather than bit field

Can we make Info interface work before MPI_Init? Will it be thread safe?

Communicator Info Keys

- Subsets have were raised for MPI 3.0:
  - https://svn.mpi-forum.org/trac/mpi-forum-web/wiki/Subsets
- Looking at this again in the context of communicator info keys:
  - https://svn.mpi-forum.org/trac/mpi-forum-web/ticket/381
- Overlaps with MPI_Init_with_assertions?

Is there a race between operations on a communicator and setting info? I.e. turning off any_source while operations are outstanding.
Can you change the value of an info key? Not forbidden, but the implementation can ignore the change.

NO_ORDERING, THREAD_LEVEL - Jim
NO_ANY_SOURCE/TAG - George

Query users about which info keys they could use -- Adam Moody

Join Us

Bi-weekly open meetings

- Mondays 11am CT
- https://svn.mpi-forum.org/trac/mpi-forum-web/wiki/MPI3Hybrid

Mailing list

- mpiwg-hybridpm@lists.mpi-forum.org