

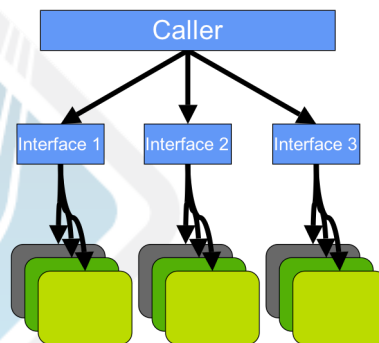


Exporting Internal Information From a Modular MPI library

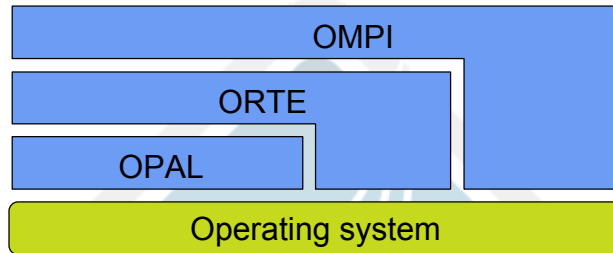
Jeff Squyres, Cisco Systems
George Bosilca, University of Tennessee

Modular Architecture

- Each component has a predefined interface
- More than one component is available for each task
- Loading a module is a mixed decision: compilation and runtime.
- Modular approach has a low overhead on modern architectures

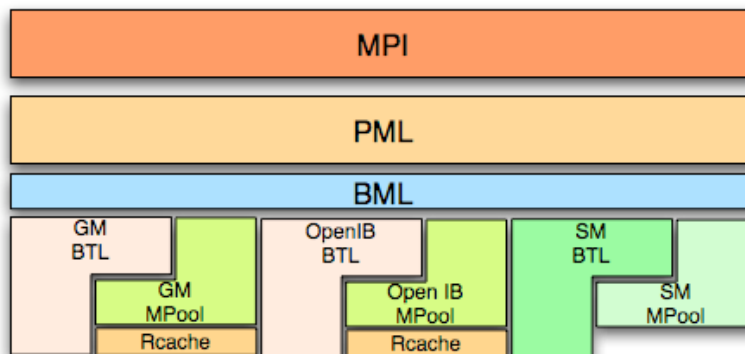


Open MPI Architecture



- Portability, runtime environment and MPI related operations
- Each layer has its own components
- For MPI debugging, our interest focuses on the OMPI layer

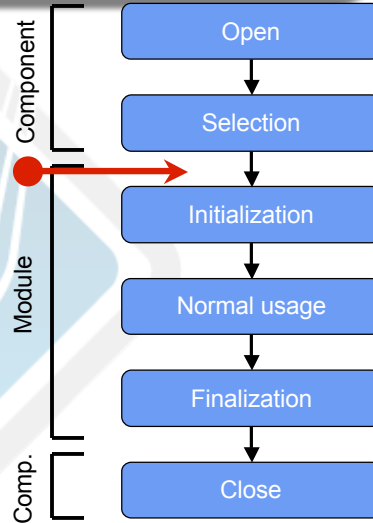
The MPI Layer



- All information required by TotalView is at the PML layer.
- Except ...

Where Are My Requests ?

- The selected PML is not available until late in MPI_INIT
- Currently 3 PML components are available
 - ob1
 - cm
 - dr
- TotalView load the debugging library when it attaches to the process
- This is too late for OMPI to change the loaded library



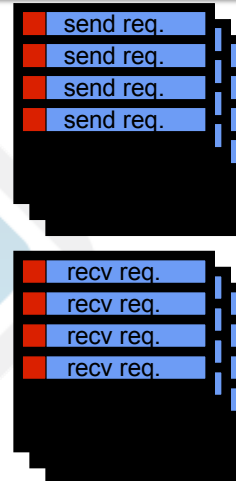
So, Where Are My Requests?

- Different PML components store the request in different ways.
 - However, they share the same basic description
- By design, all requests start with a common field, reflecting the MPI knowledge about the request (destination, tag, communicator ...)
- Simplest way to access all available requests is to get access to this common structure
- This approach give a generic way for listing the pending requests, independent of the active PML



Send and Receive Requests

- All MPI requests come from the memory management unit
- Memory management component has a standardized interface
- Once we find one request, we can iterate through all of them
- Similar approach for send and receive requests.



Unexpected Messages

- Not always available
- Most of the time a similar approach to send and receive can be applied for unexpected messages
- Hardware matching do not export information about the unexpected messages

32 vs. 64 bits

- 32 bits make things simpler
- 64 bits allow 32 bit application to run
 - The debugger and the debugged application can run in different modes
 - 2 version of the debug library should be available for such platforms

Status

- Open MPI will be TotalView message-queue friendly starting with the v1.2 release
- 2 modes will be supported
 - 32 and 64 bits modes.
- Work in progress for supporting MPI 2 dynamic processes