PIO Update

CESM Breckenridge Workshop
June 2013

Jim Edwards
CESM Software Engineering Group
(CSEG)

Thanks to NSF and DOE for their long-term support of the CESM Project
What is PIO?

An encapsulated software layer to provide uniform, high performance access to files in NetCDF3, NetCDF4 or binary format.
PIO

• Runtime toggles between underlying I/O libraries
• Tunable number of I/O tasks and per task memory overhead
• Tunable performance optimizations for Lustre and GPFS filesystems
• Tunable interface to I/O subsystem on Bluegene (L,P,Q)
• Decomposition in computational domain can be highly scattered and distributed
• Rearrangement:
  • Optimizes block sizes for filesystem access
  • Reduces application requests to I/O layer
Tuning PIO in CESM

- **PIO_TYPENAME:** (netcdf, pnetcdf, netcdf4p, netcdf4c)
- **PIO_NUMTASKS, PIO_STRIDE, PIO_ROOT:**
  - control number and location of io tasks wrt compute tasks
- **PIO_BLOCKSIZE:**
  - Control contiguous data size for I/O operations
Tuning PIO in CESM

• PIO_BUFFER_SIZE_LIMIT:
  – Controls amount of data cached before a flush is called. Available for PnetCDF write only.

Most options can also be set on a per model component basis.
(eg: OCNPIO_TYPENAME)
Ongoing Development

• Running I/O tasks independent of compute tasks (async mode)
  – Running on TACC Stampede
    • Dell PowerEdge C8220 Cluster with Intel Xeon Phi coprocessors
    • Model components on coprocessors with I/O on Sandybridge host nodes
    • Working with Srinath Vadlamani of NCAR-TDD

• Improved NetCDF4/HDF5 Performance
  – Babak Behzad (University of Illinois)

• Parallel File compression (John Dennis)
Documentation:
www.cesm.ucar.edu/models/pio

Forum:
forum.cgd.ucar.edu

Acknowledgements: John Dennis, Sheri Mickelson, Rob Jacob, Rob Latham, Wei-keng Liao, Alice Bertini, Mike Levy, Patrick Worley