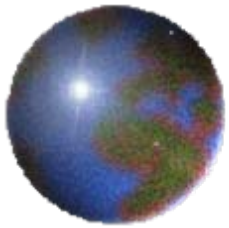
An Introduction to PIO



Jim Edwards

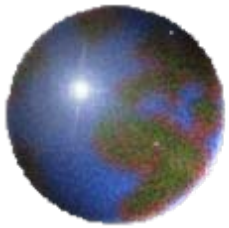
IBM

John Dennis - NCAR

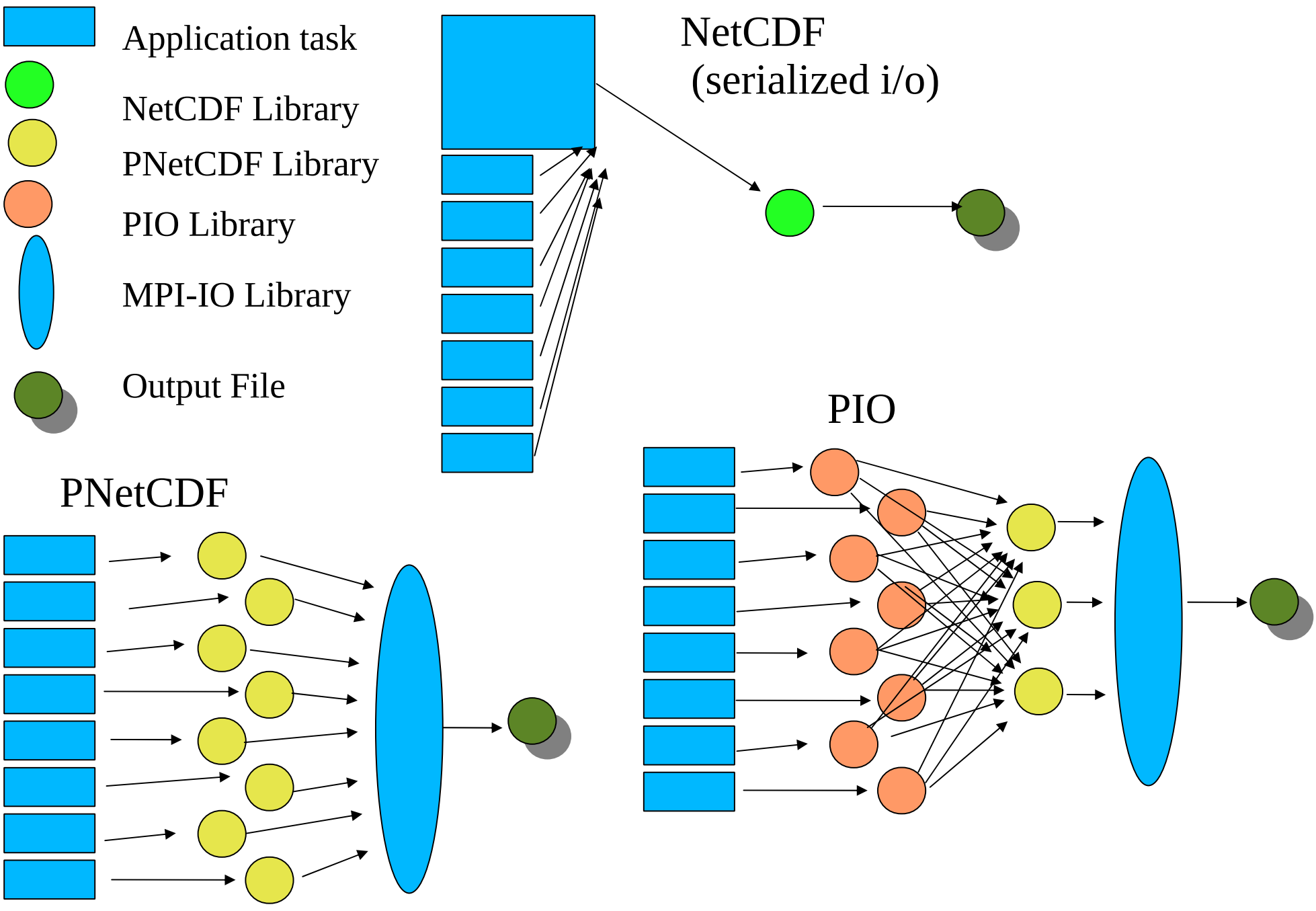
Ray Loy - ANL

Rob Jacob – ANL

Overview



- PIO Introduction
- Anatomy of a PIO interface
- Implementation in HOMME
- Implementation in CAM


Application task
NetCDF Library
PNetCDF Library
PIO Library
MPI-IO Library
Output File
NetCDF
(serialized i/o)
PIO
PNetCDF

PIO features

- All calls are collective
- Allows runtime variation of number of IO procs
- Makes it easier to supply the underlying software with data structures optimized for performance and memory characteristics
- Provides a container for IO functionality common to CCSM components (geophysical models)

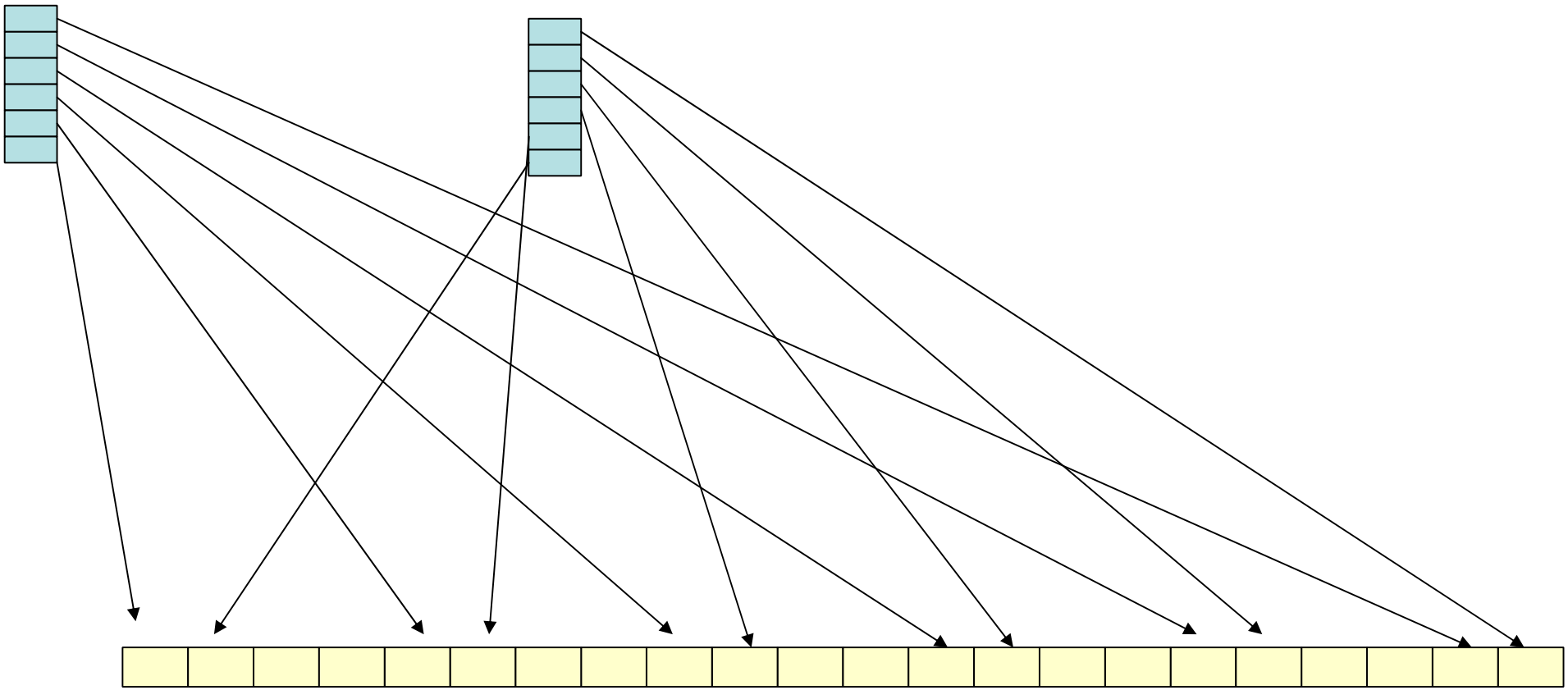
Anatomy of a PIO interface (part 1)

- PIO_InitFile(comp_rank, comp_comm, num_iotasks, num_aggregator,io_stride, ioType, File)
  - initialize the system
- PIO_OpenFile or PIO_CreateFile(File, fname)
  - open the file for reading or writing

Anatomy of a PIO interface (part 2)

- PIO_InitDecomp(File,basetype, dims,lenblocks, compdof, iodofr, iodofw, start, count, iodesc)
  - describe the relationship between data in memory and data in the file
- PIO_Def_Var(File,name,type,dimids,varDesc)
  - define a variable to write *(p)netcdf only*
- PIO_SetVarDesc(iodesc,vardesc)
  - associate a variable with a decomposition

Cam History File



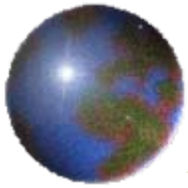
File Layout

Anatomy of a PIO interface (part 3)

- PIO_write_darray(File, varDesc, data, status)
- PIO_read_darray
  - write or read a decomposed variable
- PIO_SetFrame(Vardesc,frame)
  - sets the record number (binary) or unlimited dimension (p)netcdf
- PIO_AdvanceFrame(VarDesc)
  - increment the record number or unlimited dimension

Anatomy of a PIO interface (part 4)

- PIO_CloseFile(File)

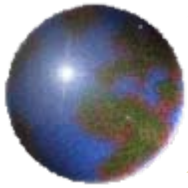


PIO NetCDF Support functions

- PIO_put_var
- PIO_put_var1
- PIO_put_vara
  - write a non-decomposed variable to the file
- PIO_def_dim
- PIO_put_att PIO_get_att
- PIO_End_Def

PIO NetCDF Support functions (part 2)

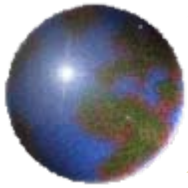
- PIO_inquire
- PIO_inq_attname
- PIO_inq_dimid



PIO Support functions

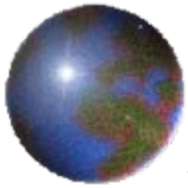
- PIO_SetDebugLevel
- PIO_freeDecomp
- PIO_dupfiledesc

Full API description at
http://swiki.ucar.edu/ccsm/97



Implementation in HOMME

- Implemented for Native Grid and Interpolated Lat/Lon Grid
- Native Grid implementation significantly improves performance over PNetCDF



Implementation in Cam History

- Goals
  - Reduced Memory Footprint
  - Parallelization of IO
  - Improved IO Performance
- Status
  - 90% Implemented for history write
  - Slightly improved memory footprint
  - Degraded IO Performance

FV 0.5 degree
maximum memory per task (MB)

	16	32	64	128	256
serial	1096	754	590	543	539
PIO 3D	1132	758	594	523	530
PIO 2D	1047	666	516	500	496

FV 0.5 degree dump_field time (seconds)

	16	32	64	128	256
serial	10.1	24	12.4	12.1	12.8
PIO 3D	7.2	7.9	8.9	8.6	13.7
PIO 2D	7	8.8	34	58	106