LA-UR- 02-6411

Approved for public release;
distribution is unlimited.

Title: Monte Carlo Radiation Transport & Parallelism

Author(s): Lawrence J. Cox, X-5
Susan E. Post, CCN-8

Submitted to: Los Alamos Computer Science Institute (LACSI) Symposium,
Santa Fe, NM
October 14-16, 2002

Monte Carlo Radiation Transport & Parallelism

Larry Cox, X-5

Susan Post, CCN-8

Abstract

This talk summarizes the main aspects of the LANL ASCI Eolus project and its major unclassified code project, MCNP. The MCNP code provide a state-of-the-art Monte Carlo radiation transport to approximately 3000 users world-wide. Almost all hardware platforms are supported because we strictly adhere to the FORTRAN-90/95 standard. For parallel processing, MCNP uses a mixture of OpenMP combined with either MPI or PVM (shared and distributed memory).

This talk summarizes our experiences on various platforms using MPI with and without OpenMP. These platforms include PC-Windows, Intel-LINUX, BlueMountain, Frost, ASCI-Q and others.

Los Alamos
NATIONAL LABORATORY

Eolus/MCNP Tradition at Los Alamos

- For decades, Monte Carlo radiation transport codes, from MCS to MCNP, have been developed and supported by the Monte Carlo team at LANL
- Concurrently, the extensive nuclear and atomic data libraries have also been under constant development
- This tradition continues in the Eolus ASCI Project and related efforts in X-5 of the WP Directorate
  - 14 MCNP team members
  - Physical Data team also in X-5
  - Two application teams (user groups) in X-5

Los Alamos
NATIONAL LABORATORY

Eolus/MCNP Tradition at Los Alamos

- Additional support is provided by other LANL groups
  - CCN-8, X-3
    - Porting, parallel efficiency analysis, testing
  - IM-8
    - Automated regression testing, software engineering improvements, modularity, quality assurance
  - CCS-4
    - Collaboration on Monte Carlo methods
  - T-16
    - Nuclear data and methods research

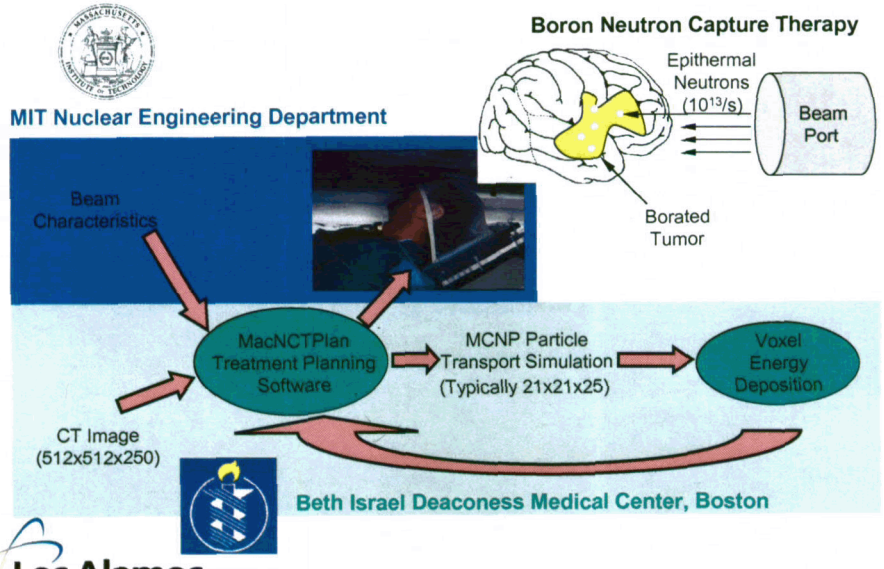
Los Alamos
NATIONAL LABORATORY

Some MCNP Applications

- Nuclear Criticality Safety
- Radiation Shielding
- Nuclear Safeguards
- Detector Design and Analysis
- Nuclear Well Logging
- Personnel Dosimetry
- Health Physics
- Accelerator Target Design
- Medical Physics and Radiotherapy
- Fission and Fusion Reactor Design
- Waste Storage/Disposal
- Radiography
- Aerospace Applications
- Decontamination and Decommissioning

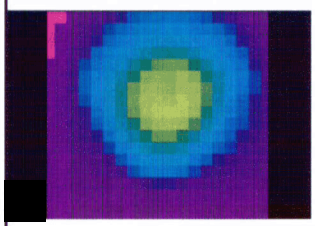
Los Alamos NATIONAL LABORATORY

Neutron Capture Therapy

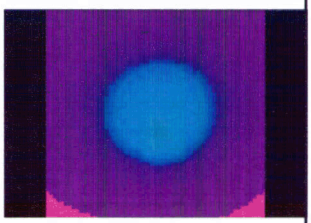


Los Alamos NATIONAL LABORATORY

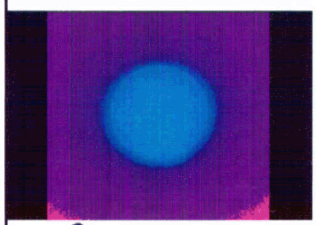
MCNP BNCT Simulation Results



Typical MCNP BNCT simulation:
- 1 cm resolution (21x21x25)
- 1 million particles
- 1 hour on 200 MHz PC



ASCI Blue Mountain MCNP simulation:
- 4 mm resolution (64x64x62)
- 100 million particles
- 1/2 hour on 6048 CPUs



ASCI Blue Mountain MCNP simulation:
- 1 mm resolution (256x256x250)
- 100 million particles
- 1-2 hours on 3072 CPUs

Los Alamos NATIONAL LABORATORY

US-Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki

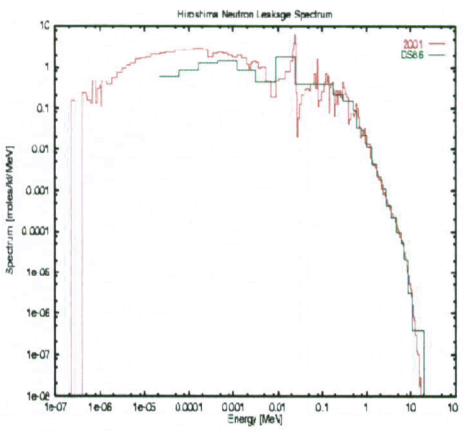
Steve White
Alexandra Heath
Paul Whalen

Los Alamos NATIONAL LABORATORY

Hiroshima; August 6, 1945 ...

Los Alamos
NATIONAL LABORATORY

Hiroshima Neutron Leakage Spectrum



LANL MCNP CALCULATION

	New	*DS86*	*new/DS86*
Total Neutrons *(Moles/kt)*	**1.768E-01**	**1.773E-01**	**0.9972**
Neutron energy *(Average MeV)*	**0.3106**	**0.2976**	**1.0437**
Yield *(Kt)*	**16.1** *(12 – 20)*		

Los Alamos
NATIONAL LABORATORY

Modernization of MCNP

- MCNP 4c3 frozen since April 2001
  - MCNP 4c3 is the last old-style MCNP version that will be released
- In the last 15 months, every line of code has been reworked
  - Conversion to ANSI-Standard Fortran-90
  - Completely new installation system
  - Modified patching method
- An ongoing process…

Los Alamos
NATIONAL LABORATORY

Eolus Parallelism

Analog Monte Carlo radiation transport can be made embarrassingly parallel because MC histories are independent

- Parallel efficiency is limited only be availability of memory to copy all information (geometry, cross sections, etc.) and by the overhead required to add up the results
- Providing statistical anaysis and tracking of sequential results complicates the algorithm
- The actual algorithm chosen can also effect efficiency
  - Collective operations versus serial sweep of results
  - Reliance on shared memory locks or global synchronization versus designing in independence

Los Alamos
NATIONAL LABORATORY

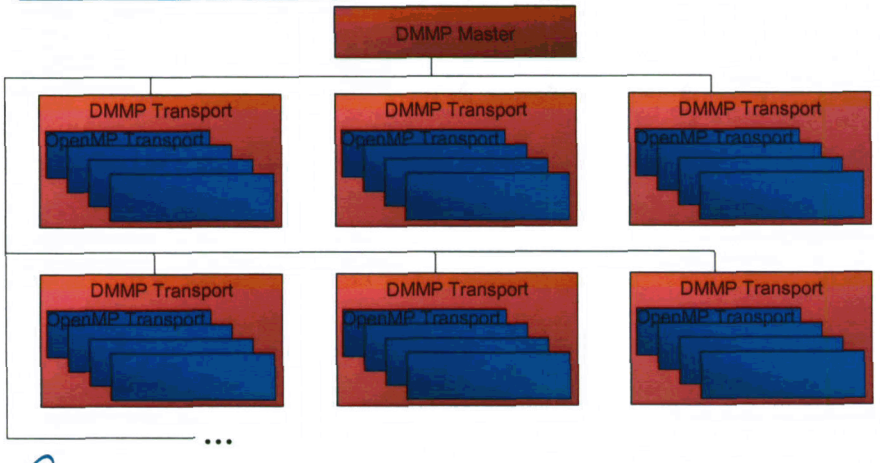
Enhanced Parallel Processing

- We use both distributed and shared memory parallelism separately or in combination
  - MPI is supported (http://www-unix.mcs.anl.gov/mpi)
  - PVM support is enhanced (http://www.epm.ornl.gov/pvm)
  - OpenMP support is added for shared-memory (threaded) parallel processing (http://www.openmp.org)

Either MPI or PVM can be used with OpenMP for optimal performance

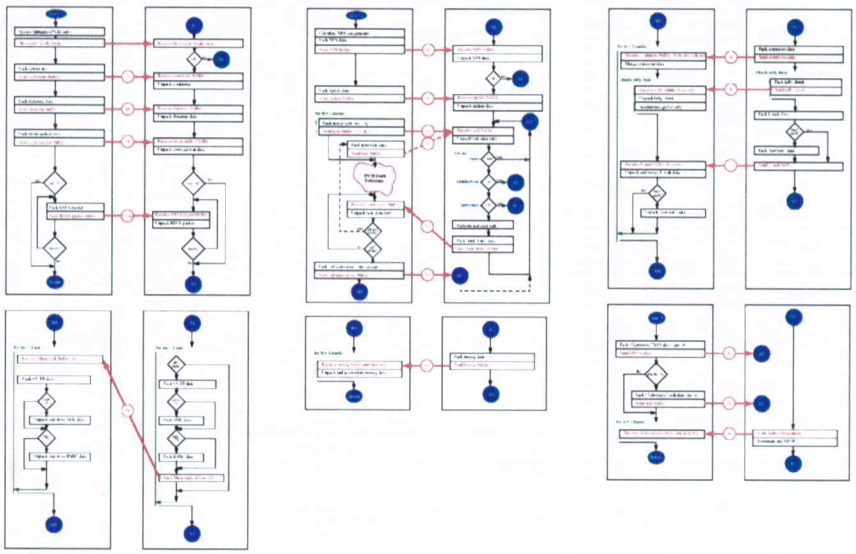
Los Alamos NATIONAL LABORATORY

Current Eolus Parallelism Algorithm



Los Alamos NATIONAL LABORATORY

Distributed Communication



Los Alamos
NATIONAL LABORATORY

Mixed Mode Parallelism on ASCI Platforms

- ***BlueMountain*** (IRIX64): 128 PE boxes in a deep NUMA configuration
  - OpenMP efficiency scales approximately as $(N)^{1/2}$
  - Best efficiency at about 8-10 threads per DMMP transport task
  - 2% to 10% effect on sequential runs with OpenMP compiled in
- ***White*** (AIX): 16 PE nodes (non-NUMA)
  - Ideal mode: 1-4 MPI tasks per nodes with 4-16 threads per tasks
  - 20% to 80% (or greater) hit when compiled with OpenMP even for sequential runs
    - Is it the threading scheme (THREADPRIVATE data) or the libraries?
    - Can it be improved?
- ***Q*** (OSF1): 4 PE nodes (non-NUMA)
  - Ideal mode: One MPI task per node with 4 threads
  - 0% to 20% loss of efficiency when OpenMP is compiled in
  - Considering an MPI-only public version
  - Will be more severely limited by memory with 4 copies per node

Los Alamos
NATIONAL LABORATORY

Analysis & Efficiency Issues

- MPI runs sequentially as well as sequential executables
- Between 5-50% effect on sequential timing by compiling in OpenMP (system & compiler dependent)
  - Does not matter if locks are replaced with critical regions
  - AIX is the worst; SGI is the best; OSF1 is acceptable
- Runtime analysis involves use of special compilers that…
  - Do not work (insurmountable compiler errors on SGI);
  - Require removal of standard language features
  - Generate behavior that is so different from that from native OpenMP compilers that we cannot tell if it is the same code or if problems are the same problems

Los Alamos
NATIONAL LABORATORY

Analysis and Efficiency Issues

- A minor code change led to a 20-50% effect on AIX
  - We have not yet determined why
- Runtimes are erratic for a certain class of problems
  - Correlated to MPI startup speed on OSF1 (30% variation)
  - Correlated to processor count (20% at 32pe to 400% at 256pe)
- Use of Fortran pointers versus allocatables
  - 50% effect on IRIX64 ; 15% on OSF1; no effect on AIX

Los Alamos
NATIONAL LABORATORY

Criticality Performance

- Spreadsheet here from 9/12/02 (spost)

Los Alamos
NATIONAL LABORATORY

Parallel Support on PCs

- MPI Version – MPICH.NT.1.2.4
  - http://www-unix.mcs.anl.gov/~ashton/mpich.nt/
  - Uses GUI mpirun
  - Socket communication; allowed on LANL networks
  - Tested on Dual CPU Win 2000, Cluster of 2
- PVM Version – PVM 3.4.3
  - http://www.csm.ornl.gov/~sscott/PVM/Software/
  - Must start pvm on master, add other PCs to cluster
  - uses rsh; not allowed on LANL networks.
  - Tested on Dual CPU Win 2000, Cluster of 1

Los Alamos
NATIONAL LABORATORY

Parallel Pros/Cons (PC)

- MPI
  - Well suited for homogeneous cluster, less overhead
  - Cluster of Windows NT/2000/XP only
  - No control-c interrupt capability
  - If any task/Machine dies, does not recover gracefully
- PVM
  - Cluster of Linux/Unix/Windows Machines
  - Has control-c interrupt
  - If one task/Machine dies, recovers unless master died
  - Possible to spawn appropriately on single/dual CPU cluster

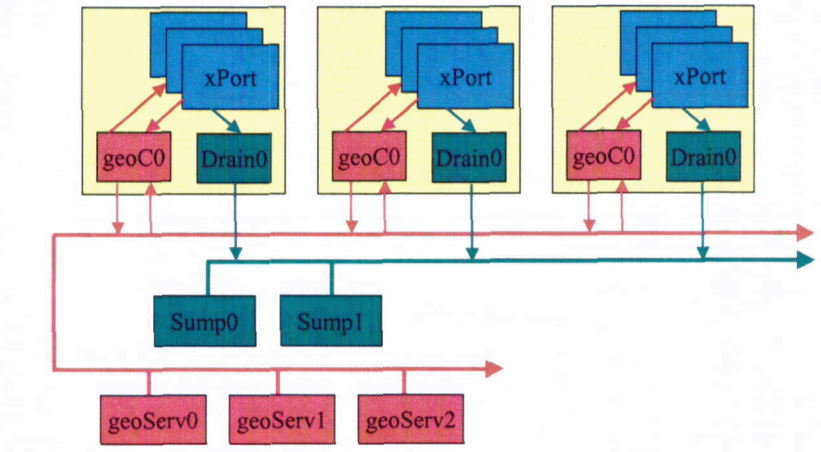
Los Alamos
NATIONAL LABORATORY

Domain Decomposition

- It may be necessary to implement various forms of domain-decomposition due to addressable memory limits
  - Already have encountered >32GB geometry files

Los Alamos
NATIONAL LABORATORY

One Possible Domain Decomposition Algorithm



Los Alamos NATIONAL LABORATORY

Eolus Parallel Support Needs

- Efficiency in parallel methods
  - MPI, OpenMP
- Parallelism algorithm analysis
  - Improvements to internal algorithms, memory layout…
- Debugging in parallel modes
  - Totalview is pretty good
- Coverage analysis in parallel modes
- Efficiency analysis in parallel modes

Los Alamos NATIONAL LABORATORY

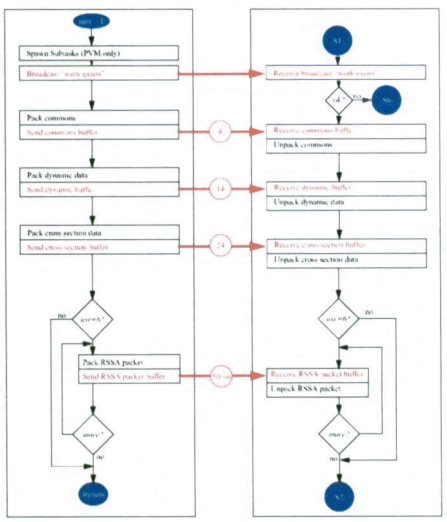
Contacts and Support

- **Email contacts**
  - **User forum:** mcnp-forum@lanl.gov
  - **Direct Contact:** ljcox@lanl.gov
- **Web sites**
  - http://www-xdiv.lanl.gov/x5/MCNP/index.html
  - http://epicws.epm.ornl.gov/ENOTE.html

Los Alamos NATIONAL LABORATORY

Initialization: Master to all Transport Tasks

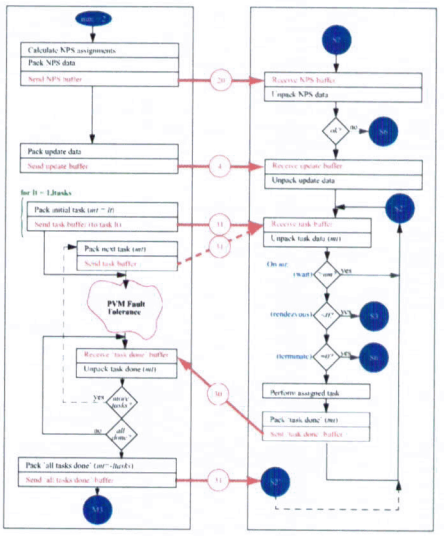
- All data is broadcast from the master task to the transport tasks
  - Problem setup data
  - Mostly read-only



Los Alamos NATIONAL LABORATORY

New assignments made and executed

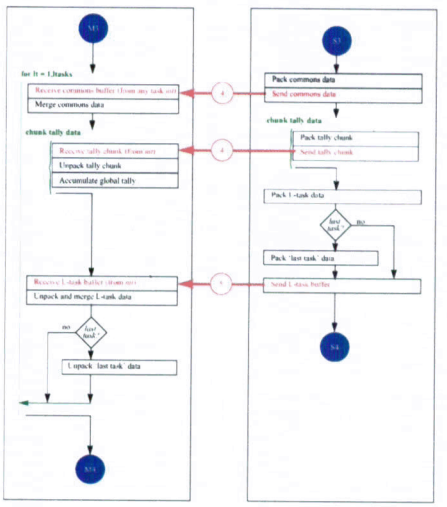
- Involves broadcast and point to point messages



Los Alamos NATIONAL LABORATORY

L-Task data accumulation

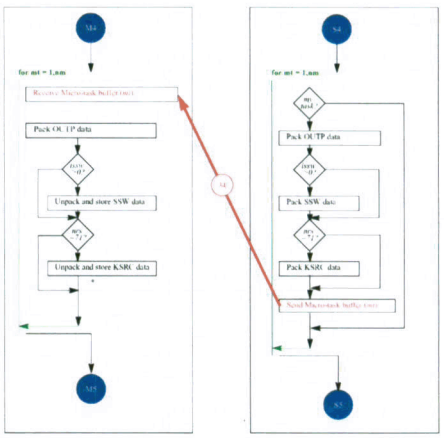
- Point to point from Transport tasks to Master
- First-come, First-served
  - Order of reception not important



Los Alamos NATIONAL LABORATORY

M-Task data accumulation

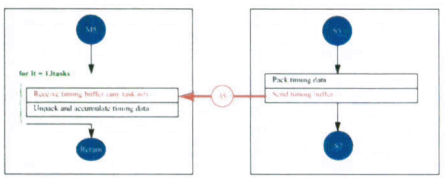
- Micro-task data is taken in order to preserve some state info
  - Source data, output messages
  - Possibly more than one per L-task
  - Order is important

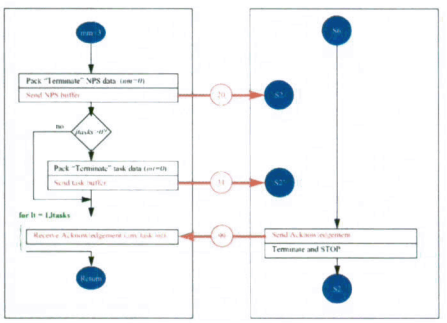


Los Alamos NATIONAL LABORATORY

Timing Data Accumulation & Termination

- Timing data taken on first-come, first-served basis
- Termination is a broadcast work assignment signifying no more work
  - Acknowledgement required





Los Alamos NATIONAL LABORATORY