HPC Resources Related to Beagle2

- Supercomputing/High Performance Computing classes or workshops
  - Educational material on High Performance computing by Henry Neeman from the University of Oklahoma Supercomputing in Plain English
  - Materials used in the Workshop by Katie Antypas, from NERSC, on April 5th about the Cray XE6
- SC and HPC books
- Articles on scientific journal

Supercomputing/High Performance Computing classes or workshops

Educational material on High Performance computing by Henry Neeman from the University of Oklahoma Supercomputing in Plain English: A High Performance Computing Workshop Series. The series includes:

- **Overview**: primer on hardware (CPU, Cache, Main Memory, Permanent Storage, input and output devices), main issues in storage hierarchy, parallelism, shared and distributed memory systems, load balancing, Moore’s law for CPU’s and how is the other parts catching up (or not).
- **The Tyranny of the Storage Hierarchy**: registers, cache (levels, latency, bandwidth and how it works), main memory (latency and bandwidth), how memory and cache interact, coherence and data locality (spatial and temporal), tiling, I/O strategies and use of NetCDF and HDF libraries and binaries, page faults and virtual memory
- **Instruction-Level Parallelism**: basics of how operations are performed on computers, Superscalar ILP (multiple operations at the same time, e.g., add integers and floating points at the same time), pipeline (different stages of different operations done at the same time), vector processors (operations done on vectors simultaneously), how different compilers take advantage of them and when they can’t
- **Stupid compiler tricks**: control dependency (loops, branches, function calls I/O), data dependency (operation depends upon previous operation), introduction to compiler optimization and profiling.
- **Shared Memory Multithreading**: Basic shared memory programming concepts (shared memory and distributed memory, threads and processes) and introduction to OpenMP instructions and run time environment.
- **Distributed Multiprocessing**: introduction to shared memory programming and the MPI libraries.
- **Applications and Types of Parallelism** and introduction to some of the most commonly encountered types of problem that can be solved using parallel supercomputers: Monte Carlo simulations, which are inherently parallel as they are made of many independent random instances of the same process and therefore usually display near perfect scaling; n-body problems with an introduction to data parallelism vs. task parallelism; and finally, Transport: meshes, finite difference approximations, ghost boundaries and cartesian MPI communicators.
- **Multicore Madness**: memory bandwidth limited computations, cache reuse, tiling strategies to exploit data locality, basic multi-core jargon (core, CPU, socket ...), issue with RAM bottlenecks (key to Beagle effective usage), communication network contention (also a potential issue on Beagle), example of how a grid calculation can be designed to be optimal for a multi-core machine by using optimal tiling, higher order methods and reduced and non-blocking communication.
- **High Throughput Computing** is about getting a lot of work done in one time unit, loosely coupled problems (e.g., Monte Carlo simulations or parameter space scans) tend to be well suited to it; opportunistic computing (using resources such as department PCs) via Beowulf of Condor; grid computing.
- **GPGPU**: Number Crunching in Your Graphics Card GPUs and GPGPUs, CUDA, StreamSDK/Brook+, OpenCL, directives for PGI compilers .... (Beagle doesn’t have GPUs)
- **Grab Bag**: Scientific Libraries, I/O, Visualization Scientific computing pipeline, scientific libraries (BLAS, ATLAS, LAPACK, ScaLAPACK, PETSc ...), I/O libraries, scientific visualization

Materials used in the Workshop by Katie Antypas, from NERSC, on April 5th about the Cray XE6:

- Richard Gerber’s introduction to High Performance computers
- John Shalf on the XE6 architecture contains useful details on the architecture and Storage hierarchy on a Cray XE6 and how it affects performance
- David Turner on how to run parallel jobs on a Cray XE6 useful to Beagle users after page 7

SC and HPC books

Victor Eijkhout’s Introduction to High-Performance Scientific Computing: a nice introduction to high performance computing that starts from the basics. Note that the linked page links to Lulu in case people want a nicely printed copy.

Articles on scientific journals

For an interesting discussion about HPC funding models http://www.hpcwire.com/hpcwire/2012-02-08/hpc_funding_models_need_to_encompass_more_than_just_the_purchase_price.html