

# Workshop on progress toward petascale applications in bioinformatics and computational biology

## Organizers:

- **Craig Stewart**, University Information Technology Services, Indiana University
- **Geoffrey Fox**, Pervasive Technology Labs at Indiana University
- **Malinda Lingwall**, University Information Technology Services, Indiana University
- **Matthias Mueller**, High Performance Computing Center Stuttgart (HLRS)
- **David Bader**, College of Computing at Georgia Tech
- **Rick Stevens**, Computing and Life Sciences, Argonne National Laboratory

Over the past several years there have been repeated analyses of the potential value of petascale bioinformatics and computational biology applications, as well as analyses of the system engineering steps required to implement applications and systems at such scale. Most recently and notably, Snaveley et al. published the “Workshop Report: Petascale applications in biological sciences”

([http://www.sdsc.edu/PMaC/BioScience\\_Workshop/Publications/PetascaleBIOworkshopreport.pdf](http://www.sdsc.edu/PMaC/BioScience_Workshop/Publications/PetascaleBIOworkshopreport.pdf)).

By one measure the era of petascale computing in biology began in 2006 with the successful clocking of the Riken Institute Protein Explorer system at 1.0 PetaFLOPS. Still, the state of the art of current applications in bioinformatics and computational biology is generally yet orders of magnitude away from petascale, especially in terms of actual performance.

The purpose of this workshop is to survey the current state of the art in computational biology and bioinformatics at scale. Suggested topics for papers and posters include, but are not limited to, the following specific subjects:

- What is the current upper limit of scale of applications in bioinformatics and computational biology? What are the factors limiting scalability of these applications?
- Can we, as recommended by Snaveley et al., identify candidate petascale applications in any of the following areas: biomolecular structure modeling, modeling complex biological systems, genomics, customized patient care, ecological components of earth systems modeling, infectious disease modeling, or other areas?
- What are the best ways to measure performance scalability of bioinformatics and computational biology applications? Can we measure what really counts in terms of next generation bioinformatics applications with FLOPS and bytes?

The NSF workshop organized by Snaveley, Jacobs, and Bader identified several specific applications as candidates for scaling. The resulting report called for attention to progress in scaling applications by identifying problems, resolving those problems, and trying to

anticipate problems at a larger scales and making the step to larger scales. Presentations that discuss the steps, challenges, and solutions to incremental scaling of bioinformatics and computational biology applications are particularly encouraged. Practice and experience papers related to this topic will be of particular value to the scientific community as we strive toward petascale applications.

### **Submission Guidelines:**

Papers must be written in English in the IEEE two-column format and should be limited to 8 pages. A detailed description of the conference paper format (8.5" x 11", Two-Column Format) can be found at IEEE Computer Society Conference Publishing Services (CPS)). The paper to be uploaded must be in Adobe's PDF format. The deadline for submitting an abstract along with the original paper in PDF via an online paper submission system is May 15, 2007. Papers should be uploaded and identified as being contributions to the Workshop on Petascale applications in Bioinformatics and Computational Biology. Details on the submission process are available at <http://www.cs.gsu.edu/BIBE07/submission.php>

Submission of a paper should be regarded as a commitment that, should the paper be accepted, at least one of the authors will register and attend the conference workshop to present the paper or the poster.

Papers will be refereed and accepted on the basis of their scientific merit and relevance to the workshop topics.

