**MapReduce BOF at TG11**

As the computing landscape becomes increasingly data-centric, data-intensive computing environments are poised to transform scientific research. In particular, MapReduce based programming models and run-time systems such as the open-source Hadoop system have increasingly been adopted by researchers with data-intensive problems, in areas including bio-informatics, data mining and analytics, and text processing. While Map/Reduce run-time systems such as Hadoop are currently not supported across all TeraGrid systems (it is available on systems including FutureGrid), there is increased demand for these environments by the science community. This BOF session will provide a forum for discussions with users on challenges and opportunities for the use of MapReduce. It will be moderated by Geoffrey Fox who will start with a short overview of MapReduce and the applications for which it is suitable. These include pleasingly parallel applications and many loosely coupled data analysis problems where we will use genomics, information retrieval and particle physics as examples.

We will discuss the interest of users, the possibility of using Teragrid and commercial clouds, and the type of training that would be useful. The BOF will assume only broad knowledge and will not need or discuss details of technologies like Hadoop, Dryad, Twister, Sector/Sphere (MapReduce variants).