Summary of discussion for
**Grid-based image archiving and analysis**
by Hastings, Oster, Langella, Kurc, Pan, Catalyurek, and Saltz
J Am Med Inform Assoc. 2005 May-Jun;12(3):286-95

INLS 279: Bioinformatics Research Review
Presented by Noel Fiser
November 29, 2005

In this article the authors present a middleware system called GridPACS intended to assist scientists in managing and annotating medical imagery. The authors present four goals for their system: distributed data access, scalability, active storage, and on-demand database creation. The article first presents a short background explaining why the particular middleware approach was adopted, which was largely the result of earlier advances in other systems combined with general technological advances (e.g. XML), allowing a new system to be built that would offer the most robust and flexible metadata collection and management possible. GridPACS was designed to allow the collection of the most diverse biomedical imaging data sets possible, but in particular the developers focused on two areas: dynamic contrast enhanced MRIs and digital microscopy. When imagery data is inserted, it is broken into many smaller chunks and distributed among the nodes in the GridPACS system. This is intended to allow the images to be retrieved from multiple systems in *parallel* instead of serially, allowing for theoretically smaller retrieval times. Having established the parameters of the system (as well as the physical architecture implemented for their test case), the authors confirmed the improved I/O performance of their system by testing various image sizes.

We found the concept and general implementation of an XML-based image management system very compelling. The GridPACS system serves a very important need of research and clinical scientists alike, and its model is both flexible and robust enough to allow for the necessary growth such a system would require before becoming truly useful to a large number of people. And from what we could tell, the results of the system tests support this system as satisfying its original goals, particularly scalability.

These test results, however, were somewhat disappointing, as they showed no real benchmarks for other similar or legacy systems performing the same actions, so it became impossible to draw overall conclusions on the matter. Also, there are some glaring issues that are either left unaddressed or inadequately explained, like how extensively the GridPACS system is currently being used beyond the OSU medical facilities, as well as how the clustering of the system actually functions.

There is also a fundamental computer science dilemma of where to place the bottleneck in your system. By clustering these computers but not indicating a clear system of redundancy, the bottlenecks here are glaring: the network, and non-responsive nodes. And of course, it would have been nice to see some of the actual queries being passed to this system in a step-by-step process instead of a cursory explanation with attached user interface screenshot.

GridPACS is an adventurous step forward in distributed computing for the medical field. Its promise outweighs whatever flaws we found. We only wished we could see it in action beyond the walls of a single university.