

Toward Production Quality of e-Science Infrastructure

Eric YEN

Grid Computing Centre, Academia Sinica, Taiwan

eric@gate.sinica.edu.tw

The Grid has demonstrated its flexibility and performance to support large-scale computation on-demand, with the real application domains of High Energy Physics and Biomedical etc. Accompany with the ramping up to WLCG full capacity and performance by April 2007, lessons learned from WLCG would be essential for the development of production e-Science environment by taking advantage of Grid technology. Objectives of this study is to understand the gap toward a production e-Science environment, details would be elaborated in this paper.

Reliability is the most challenging to the Grid infrastructure, sustainable and operation model would be necessary as well for a production quality environment. Effective use of fast networks is intuitively fundamental for e-Science but storage system would be more critical before driving to the limit of wide area network. Having network resources predictable and schedulable is also required to make the Grid service optimized and to a finer grained level. Regarding to the storage issues, effective global data catalog and replication management is the key for scalable and feasible grid computing model. Standardized protocol, workflow automation, streaming support for prioritized data access, and optimization of common uses cases are also indispensable to wider spectrum of e-Science applications. A production e-Science framework is not just a grid or grids, but also the grid of services. Interoperation of variant computing model and factorization into the common application layer of middleware could ease the integration of multi-disciplinary applications and the formation of common framework without reinventing wheels. Finally, architectural issues are also explored, for example, how is the underlying complexity of the Grid masked by taking advantage of service-oriented approaches.