CS690 – Week 2 Paper Summaries
Jordan Osecki

The Knowledge Grid Environment
Hai Zhuge
Intelligent Systems, IEEE
Volume 23, Issue 6, Nov.-Dec. 2008 Page(s):63 - 71
Digital Object Identifier 10.1109/MIS.2008.111
Summary: The knowledge grid environment is an autonomous human-machine environment evolving with science, technology, culture, and society.
It consists of autonomous individuals, self-organized semantic communities, an adaptive-networking mechanism, and an ev..

The author for this paper describes a hypothetical future "Knowledge Grid Environment", which is an autonomous human and machine combined environment which will evolve with the natural processes of science, technology, culture, and society. It will consist of autonomous individuals, self-organized semantic communities, an adaptive-networking mechanism, and an evolving semantic networking mechanism. A memex extension (ME) will offer a general model for autonomous individuals in this environment. These ME's are configurable, adaptive, and context-aware digital organisms that will model network resources and host software and devices, organizing themselves to perform tasks based on socioeconomic principles. The ME's will interact, inherit from each other, and work together to produce solutions, with little to no human guidance. These digital organisms will improve data flow and increase the productivity on the network.

This is describing an entirely new system, one which would be built using a grid, peer-to-peer, the web, and Web 2.0. It would be used by teams that require lots of knowledge-intensive teamwork. It is still a hypothetical system, however, and no work has been done on it so far. The approach is very new and work has not been done on creating it, although all of the pieces and technology already exist, simply requiring to be integrated together. The paper also describes with diagrams the structures of the ME's and other components. If it is successful, it would make whoever is using it much more productive, with more computation and work going on in the background by ME's and/or groups of ME's. These digital organisms can take care of all of the "busy work", leaving users to simply "create", hypothetically. This system has not even begun to be implemented, but it would be an extensive project and would take a lot of time, money, and collaboration to ensure that it didn't just become an in-house solution which is non-standard and not usable by the public or at least by the scientists in mass.

Workflow Planning on a Grid
Thompson, C.W.; Wing Ning Li; Zhichun Xiao
Internet Computing, IEEE
Volume 11, Issue 1, Jan.-Feb. 2007 Page(s):74 - 77
Digital Object Identifier 10.1109/MIC.2007.26
Summary: Workflows are used in scientific, engineering, industrial, legal, and business applications to process many kinds of data, such as genomes, supply chains, and mailing lists. Workflows that involve thousands to millions of files and up to terabytes of.....

The University of Arkansas authors have created a workflow management system which can take in descriptions of different services and then defined tasks (i.e. inputs and outputs) and the software will perform workflow automation to develop a workflow. They argue that workflow specification rather than workflow execution is the true bottleneck. Their software uses a backtracking-based search algorithm to eliminate human specification errors. Their paper also describes the benefits of using a grid to perform workflows which involve thousands to millions of files and terabytes of data. However, an important part of being able to do this is to automate the workflow processes.

For the most part, workflows do not use grids and there are programs for human workflow analysts to create workflows, but this is prone to error. Several other researchers are indeed testing workflows on grids and designing different forms of automation as the University of Arkansas authors are as well. They are working independently of these other efforts to try these two new approaches. If successful, these results would be very important to e-science, which is beginning to use grids more and more. The inclusion of workflows and specification automation to grids would help scientists to run more processes and quicker experiments. If the prototype they created were finished, it could be the single piece needed to include workflows into grids. The prototype still has many limitations, which are documented in the paper. It would require these to be fixed before this prototype could make a large difference. Achieving this is still at least six months to a year away, and then it would also need to be reviewed by standard's committees. To monitor this solution for success, it could be tested with e-science scientists and also whether or not it is favorably reviewed by different standard's committees.

E-Science: the grid and the Semantic Web
De Roure, D.; Hendler, J.A.;
Intelligent Systems, IEEE
Volume 19, Issue 1, Jan-Feb 2004 Page(s):65 - 71
Digital Object Identifier 10.1109/MIS.2004.1265888
Summary: Over the past few years, researchers have been treated to two visions of the Internet's future. One is the Semantic Web, the next generation of World Wide Web technology. The second is grid computing, the next generation of internetworked processing......

The two authors describe grids, semantic webs, the importance of each to e-science and scientists, and how the integration of the two into a semantic grid would be the ultimate for scientists. They describe how scientists have quickly adopted grids for the computing power that they provide, but now they need the organization and metadata that a semantic web provides in order to be able to collaborate projects with previous ones and with current ones from different organizations. The paper lists some reasons that the integration of these two technologies will lead to success as well as the challenges which are necessary to overcome. The challenges include the following: realizing the network effect, moving beyond centralized sources, automated service assembly, collaborative tools, and the Grid being embedded into the real world. For the most part, both technologies are still used separately. There is one example, myGrid, which produces a virtual laboratory workbench to serve the life sciences community. This paper lays out the process necessary and the challenges to overcome to achieve a true integration on a larger scale, which has yet to be done.

Scientists and anyone involved in e-science would be heavily benefitted if this solution became a reality. The paper does not produce any prototypes or advances, however. It mainly just lays out a roadmap for others to follow for success. The risks are working to integrate the two technologies in such a way that others would support the result and it would become a standard. If the integration is successful, scientists will reap the benefits of both the power of grids and the organization of semantic web. This will require a long process and large development time and costs in order to succeed. There is no prototype started for this team of authors and other successes are mainly centralized solutions. As the system is built, checks for success would include if each of the five challenges laid out in this paper are being addressed and eliminated.