Reduced Statement of Work for the VGrADS ITR Project

Overview

In order to revise the SOW to fit within the reduced budget awarded, we will make a number of significant changes to our overall plan.

First, we will drastically reduce the budget for software infrastructure development and technology transfer so that we may focus more of our resources on the computer science research. Thus, we will eliminate all of the major external software releases and the shared software development staff. Technology transfer will then be limited to working with those applications that we need to drive our research. In many cases we will seek separate support for working with those applications or make those efforts part of the research itself. This will also permit us to drop the tutorials on software infrastructure and the cycle of selecting new application partners.

We will also make deep cuts in support for applications partnerships. Although much of the research on the GEMS application will be retained, the collaboration with the EOL and MEAD/LEAD application teams will be essentially unfunded under VGrADS. We will continue synergistic activities with these groups through other projects, and will seek separate grants to allow a more robust interaction. These actions will severely limit the validations and testing of VGrADS on full applications, although we hope to gain many of the same insights through experiments on small applications kernels. For now, new applications will be added only when they are needed to provide a unique perspective on the research.

We have also cut international travel drastically, and domestic travel where possible. The latter was facilitated by reducing the number of VGrADS workshops to one per year rather than two as proposed. (In year 1, however, we will hold 2 workshops to ensure that the project starts with a strong collaborative program.)

We will also make significant cuts in the research itself by focusing on those goals that are most essential to the long-term success of the project. In some cases this will mean that the research will take longer, and in other cases we will reduce the effort on individual subtopics. These cuts are detailed below.

Software Distribution

We will focus on developing the software needed to support our internal research, developing it in a distributed fashion within the research effort as we have done under GrADS.

We will make prototype VGrADS components available to the Grid computing community as they are developed. Most likely, this will be done through a web site.

Although the software will have only the documentation and maintainence support needed for our internal purposes, it is likely that some other groups will also be able to use parts of it. All software will be open-source to encourage use and extension by a wider community.

Execution System

Virtualization.
Due to the reduced resources, our research into virtual grids will be significantly curtailed, pursuing the same ideas and approach, but significantly reducing the scope and the scale of experiments and demonstrations. Specifically, we will focus on a few examples of resource abstraction classes and virtual grids, and a limited implementation that deals only with these specific examples to demonstrate the concepts. To maximize leverage, we will make use of small application "kernels" that provide much of the insight, but not the scale of demonstration nor direct application impact. The virtual scheduling research and resource selection research will also be limited to these examples and kernels. The resource conditioning research will not be pursued.

As appropriate and feasible given the resources, we will demonstrate the proven virtual grid techniques on the partner application programs. This differs from the original project plan in that we will not deliver VGrADS-enabled versions of these applications to our disciplinary partners.

In summary, these changes will reduce the breadth, depth, and total amount of forward progress for the proposed research. They will also reduce the level of dissemination of these results to the community.

Performance Provisioning.
As in other areas, the reduced resources will require us to reduce the depth of our investigations into performance provisioning, and eliminate some promising but secondary areas. There will also be generally slower progress due to lack of personnel time available to this project. Initially, we will investigate time-space reasoning for contracts and signatures, albeit with reduced scope and rate. In year 2, we will extend this space-time reasoning toward practical use,. However, we will be forced to eliminate research in genetic programming for performance control. In year 3, we will limit the extent of our study of performance/fault-tolerance capabilities to simpler combinations of the two, rather than a fully integrated performability approach as originally proposed. In year 4, we will eliminate integrated signature/time-space/statistical characterization in favor of tunable performance/fault-tolerance capabilities (originally intended for year 3). Finally, year 5 will only allow limited application validation and assessment, with no assessment of technologies that the reduced scope caused us to eliminate.

Fault Tolerance.
We will generally make slower progress due to a reduction in investigator effort and student support. In particular, we will reduce the initial experimental measurements of

Grid and cluster reliability, severely limit our testing of fault tolerance in applications, and have limited validation and assessment of the ultimate techniques. However, we will still perform some initial reliability experiments, use those to prototype an appropriate fault tolerance library, develop some novel techniques such as diskless checkpointing, and validate the library and techniques (albeit not as thoroughly as we would have liked). We will also eliminate some investigations into specific techniques that we had planned, notably non-transparent coordinated checkpoints and novel techniques such as caching and dynamic loading.

Grid Economies and Policy
Our focus, for the project, will be on the identification of effective economic models that can be used to govern resource allocation amongst competing Grid-enabled programs. By controlling allocations according to economic principles, our approach will pursue attractive global "emergent" properties such as allocation stability, fairness, and efficiency.

Our initial intent was to study economic control using VGrADS through the VGrADSBank — a modular and integrated software subsystem for supporting Grid economic policy experiments. However, the current funding level will not make the development that is necessary to build such a research vehicle possible. Instead, we will narrow our focus to the subject of "pricing mechanism" and its effects on the dynamism of resource allocation. Our approach remains empirical in that we intend to use VGrADS-enabled applications as test subjects in our experimental design, but the scope of our investigations will be limited to the investigation of pricing effects. We will also employ simulation much more aggressively to make up for the inability to conduct the experiments that the VGrADSBank would have enabled.

Programming Tools

To reduce the size of the overall programming tools effort, we will substantively reduce infrastructure building to that needed to support the research and work with the applications that we have selected to drive the effort. This will also allow a larger involvement of graduate students at the expense of staff support.

In addition we will reduce the effort on the abstract component machine to the level needed for integration with virtual grids. This will build on work already completed as a part of the GrADS project. We expect that the main remaining focus of the abstract parallel machine work will be research on scheduling, performance estimation, and load matching, all of which also support the work on the abstract component machine.

We will concentrate the remaining resources on the abstract component machine, which promises a truly high-level interface for Grid programming. As a result of the reductions elsewhere we will be able to retain most of the research on component-based Grid programming, with a small amount of shifting of the most ambitious goals to the later years. This work will preserve most of the research efforts on distributed launching,

binding (remote component tuning and sensor insertion), and Grid-service library installation.

Applications

Due to the overall budget reduction, all application-specific funding has been removed in order to preserve pure computer science research activities. Nevertheless, we plan to leverage synergistic activities involving VGrADS co-PIs and the Encyclopedia Of Life (EOL) application. Co-PIs Berman and Casanova are collaborating with EOL researchers for porting EOL to the Grid and a prototype has already been developed. Thus, it will be natural for this work, while not funded under the VGrADS project, to be used to evaluate and to benefit from VGrADS technology.

Similarly UIUC will pursue collaborations on the MEAD /LEAD application as a driver for the research,. This application will be primarily supported from other sources, so integration with the VGrADS framework may suffer.

The GEMS application will be pursued by Lennart Johnsson at the University of Houston. Although its integration with VGrADS will be delayed due to the generally slower progress of the overall project, most of the original goals will remain the same. Because the budget reduction will not allow us to support a postdoc to work directly with the GEMS developers, the work will be done by graduate students and may be expected to procede at a slower pace.

We will not incorporate new applications except as needed to exercise particular concepts.

Education Outreach and Training

Due to the budget cuts we will be able to support only 2 AGEP students rather than the four originally proposed. Similarly, staff support will be significantly reduced, and in particular, Cynthia Lanius has left Rice University and will no longer be part of the project.

Graduate student programs will be largely preserved as proposed. New course development may be delayed.

Because of the reduction in software distribution support, technology transfer through tutorials on VGrADS software will be severely curtailed. We will consider whether more general-purpose tutorials and other tech transfer activities are still possible.

Administration

We are reducing the number of workshops to one per year (two in the first year), leading to a substantial reduction in support. Support for teleconferencing has been reduced somewhat due to cost reductions at Rice, but the number and length of these calls will not change as they have proved such an effective management tool.

Staff support for administration has been reduced in proportion to the total budget as well, requiring more researcher time to be devoted to reporting and budget issues.