dV/dt – Accelerating the rate of progress towards extreme scale collaborative science

Applicant: The Board of Regents of the University of Wisconsin System
Principal investigators:

Miron Livny,	University of Wisconsin–Madison
Frank Wuerthwein,	University of California–San Diego
Douglas Thain,	University of Notre Dame
Ewa Deelman,	University of Southern California
William Allcock,	University of Chicago, Argonne National Laboratory

Abstract

Scientific computing and scientific collaboration are both activities that cannot be defined in absolute terms as they continuously evolve over time. What is an extreme scale computation today may not be considered as such in six months and what is considered as a close collaboration today may be viewed differently in the future. Moreover, the scale of computing and the characteristics of collaborations have many dimensions, most of which are difficult to quantify. Mapping all these dimensions to a single subjective quantifiable metric has been an ongoing challenge. It is therefore critical that when dealing with computing and collaboration we focus our attention on the **rate of change (dV/dt)** rather than on an absolute metric.

What constitutes a significant change in the scale of a computation or collaboration depends not only on the starting point but also on the specific circumstances of the scientific endeavor. What are common to many of these endeavors are the impediments they face in accomplishing the desired change. In many cases, an attempt to use more software tools to process data from more sources with the power of more computing resources requires human effort that is beyond the reach of many scientists. While collaboration creates the opportunity to share software, data and computing resources, translating this potential can be a long and costly process. The experimental computer science research program outlined in this proposal is driven by a desire to develop novel computational frameworks and technologies that will accelerate the rate at which members of the DOE science community seize such opportunities as they transition to extreme scale collaborative science.

The assembled team of scientists and practitioners has a long history of collaborative research that has resulted in widely adopted scientific software tools. Located at four universities and a DOE National Laboratory, the team covers a broad range of computer science areas and computing infrastructures. Using *planning* as the unifying concept for this project, we will develop and evaluate by means of at-scale experimentation novel algorithms and software architectures that will make it less labor intensive for a scientist to **find** the appropriate computing resources, **acquire** those resources, **deploy** the desired applications and data on these resources, and then **manage** them as the applications run. The proposed research will advance the understanding of resource management within a collaboration in the areas of: trust, planning for resource provisioning, and workload, computer, data, and network resource management.

This work will result in research artifacts (frameworks, algorithms, simulators, and execution traces) as well as an experimental testbed that will support the proposed research and will be made available to the broader DOE community.

The proposed research is directly applicable to the DOE science areas. Working closely with developers of software tools that power collaborative science, we will translate the results of our research into deployable and widely adopted capabilities. Thus this work will have the potential to impact DOE-funded science including biology, chemistry, computer science, economics, engineering, mathematics, medicine, climate science, and physics.