<u>Final Report</u>

dV/dt - Accelerating the Rate of Progress towards Extreme Scale Collaborative Science

Miron Livny – Project PI

A traditional model of scientific computing assumes that a single researcher intends to use a homegrown application on a known computational resource that is precisely provisioned, has the required data, and is configured to meet his or her needs. However, this rarely happens in practice anymore because the scale of scientific computing changes at an ever-increasing speed in all dimensions. Today, researchers from a growing range of scientific disciplines band together into dynamic collaborations that employ an increasing number of applications, software tools, data sources, and instruments. These collaborations have access to a growing variety of processing, and storage resources ranging from local personal computers to remote national leadership-class machines, and commercial clouds. Our project developed frameworks and technologies that make it easier for scientists to conduct large-scale computational tasks that use the power of computing resources they do not own to process data they did not collect with applications they did not develop. We believe that these capabilities promote compute and data intensive scientific collaboration and thus significantly accelerate the rate of progress towards extreme scale computing.

Scientific collaborations come in many shapes, colors and sizes. They range from established to ad-hoc entities and from explicit to implicit agreements. Some of them consist of thousands of scientists, operate for decades and are well-funded while others may involve less than a handful of scientists, last a few months and have a small or non-existent budget. Explicit collaborations have bylaws that govern membership whereas implicit collaborations simply happen when one scientist uses data and or tools developed by other scientists to test a hypothesis or explore a phenomenon. What is common to all these collaborations is the sharing of software tools, instruments, computing resources, and/or data. Collaboration happens when scientists discover each other's work and find out that working together and sharing tools and resources can enhance the pace of their scientific discoveries. Just as the leadership class resources and data sets have been greatly scaling up over time, it is required for the computing capabilities of collaborations to meet that pace of change.

Easy access to more data sources, more scientific applications and/or more computing capacity holds the key to expending the computational scale of scientific investigations. What hinders the progress of scientific collaborations towards new scales of computing is the effort involved in **finding** the appropriate computing resources, **acquiring** those resources, **deploying** their applications and data on the resources, and then **managing** them as they run. Unfortunately, this is a very labor-intensive process. Any mismatch between the application's needs and the capabilities of the computing system results in a problem being sent back to the user. Incompatible libraries, full disks, missing data and permission problems are the rule rather than the exception. As a computing system increases in heterogeneity and size, so do these problems, requiring either an increase in the number of staff or a decrease in the user's ambitions. Thus, at extreme scale, the proper measures of performance are not peak cycles per second, but rather the total cost of performing a computation, from conception to completion, including human effort in addition to the cost of the computing resources.

By addressing the **fundamental computer science problems** in deployed computing frameworks and technologies (CFT) we identified approaches that speedup the rate of progress of scientists to extreme scale collaborative computing. Through experimentation with prototype implementations on at-scale computational environments that involve real-life scientific applications and advanced computing infrastructures, our team of scientists and practitioners from four universities (University of Notre Dame (ND), University of Southern California (USC), University of California San Diego (UCSD) and University of Wisconsin–Madison (UW)) and the Argonne National Laboratory (ANL) developed, studied and evaluated novel CFTs.

Computer scientists with strong ties to domain scientists, software developers and operators of computing infrastructures used proven experimental methods and collaboration methodologies to study these CFTs and promote their integration into end-to-end capabilities. Working closely with developers of software tools that power collaborative science, our project translated research results into deployable capabilities. The organization of the team and the structure of the project leveraged the cumulative experiences of the Argonne Leadership Computing Facility (ALCF) [1] and the Open Science Grid (OSG) consortium [2,3] in advancing the computing scale of a diverse community of domain scientists.

The cornerstone of our results is an integrated *monitoring* and *planning* framework that covers the spectrum of computing resources—processing, storage, and software. This framework supports the four phases of collaborative computing—find, acquire, deploy, and use. We evaluated and profiled the effectiveness of the new CFTs taking into account the lifecycle of a collaborative computational effort. Our team had access to a powerful suite of software tools as well as state of the art computing capabilities—a national leadership computing facility (ALCF) and a national High Throughput Computing virtual facility (OSG). We also used campus based computing infrastructure at ND, UCSD and UW. This collection of campus, and national resources allowed us to evaluate technologies that support the "**submit locally and compute globally**" paradigm.

The motivation, results and artifacts produced by the project are summarized at https://sites.google.com/site/acceleratingexascale/home. In addition to software tools, the project produced the following 12 publications:

- "**A Job Sizing Strategy for High-Throughput Scientific Workflows**, Benjamin Tovar, Rafael Ferreira da Silva, Gideon Juve, Ewa Deelman, William Allcock, Douglas Thain, and Miron Livny, *IEEE Transactions on Parallel and Distributed Systems*, **29**(2), pages 240-253, February, 2018. DOI: 10.1109/TPDS.2017.2762310
- "**Consecutive Job Submission Behavior at Mira Supercomputer**", Stephan Schlagkamp, Rafael Ferreira da Silva, William Allcock, Ewa Deelman, and Uwe

Schwiegelshohn, 25th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC), 2016.

- "**Understanding User Behavior: from HPC to HTC**", Stephan Schlagkamp, Rafael Ferreira da Silva, Ewa Deelman, and Uwe Schwiegelshohn, International Conference on Computational Science (ICCS), 2016.
- "**Online Task Resource Consumption Prediction for Scientific Workflows**", Rafael Ferreira da Silva, Gideon Juve, Mats Rynge, Ewa Deelman, and Miron Livny, Parallel Processing Letters, 25(3), 2015.
- "**Practical Resource Monitoring for Robust High Throughput Computing**", Gideon Juve, Benjamin Tovar, Rafael Ferreira da Silva, Dariusz Król, Douglas Thain, Ewa Deelman, William Allcock, and Miron Livny, 2nd workshop on monitoring and analysis for high performance computing systems plus applications (HPCMASPA), 2015.
- "**Characterizing a High Throughput Computing Workload: The Compact Muon Solenoid (CMS) Experiment at LHC**", Rafael Ferreira da Silva, Mats Rynge, Gideon Juve, Igor Sfiligoi, Ewa Deelman, James Letts, Frank Wurthwein and Miron Livny, 2015 International Conference on Computational Science (ICCS), 2015.
- "**Algorithms for Cost- and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds**", Maciej Malawski, Gideon Juve, Ewa Deelman and Jarek Nabrzyski, Future Generator Computer Systems, to appear, 2015.
- "**Community Resources for Enabling Research in Distributed Scientific Workflows**", Rafael Ferreira da Silva, Weiwei Chen, Gideon Juve, Karan Vahi and and Ewa Deelman, 10th IEEE International Conference on e-Science, 2014.
- "**A Performance Model to Estimate Execution Time of Scientific Workflows on the Cloud**", Ilia Pietri, Gideon Juve, Ewa Deelman and Rizos Sakellariou, 9th Workshop on Workflows in Support of Large-Scale Science (WORKS), 2014.
- "**A Unified Approach for Modeling and Optimization of Energy, Makespan and Reliability for Scientific Workflows on Large-Scale Computing Infrastructures**", Rafael Ferreira da Silva, Thomas Fahringer, Juan J. Durillo and Ewa Deelman, Workshop on Modeling and Simulation of Systems and Applications (MODSIM), 2014.
- "**Practical Resource Monitoring for Robust High Throughput Computing**", Gideon Juve, Benjamin Tovar, Rafael Ferreira da Silva, Casey Robinson, Douglas Thain, Ewa Deelman, William Allcock, and Miron Livny, USC Technical Report 14-950, 2014.
- "**Toward Fine-Grained Online Task Characteristics Estimation in Scientific Workflows**", Rafael Ferreira da Silva, Gideon Juve, Ewa Deelman, Tristan Glatard, Frederic Desprez, Douglas Thain, Benjamín Tovar and Miron Livny, 8th Workshop On Workflows in Support of Large-Scale Science, 2013.