

Invited: SHWD-I-01

SIRIUS: Science-driven Data Management for Multi-tiered Storage

**Scott A. Klasky¹, Hasan Abbasi¹, Mark Ainsworth⁴, Qing Liu¹, Jay Lofstead⁵,
Carlos Maltzahn⁷, Manish Parashar⁶, Feyi Wang¹**
**Oak Ridge National Laboratory¹, University of Tennessee at Knoxville², Georgia Institute
of Technology³, Brown University⁴, Sandia National Laboratory⁵, Rutgers University⁶, U.
California Santa Cruz⁷**

Scientific discovery at the exascale will not be possible without significant new research in the management, storage and retrieval over the long lifespan of the extreme amounts of data that will be produced. Our thesis is that adding application level knowledge about data to guide the actions of the storage system provides substantial benefits to the organization, storage, and access to extreme scale data, resulting in improved productivity for computational science. In this project we will demonstrate novel techniques to facilitate efficient mapping of data objects, even partitioning individual variables, from the user space onto multiple storage tiers, and enable application-guided data reductions and transformations to address capacity and bandwidth bottlenecks. Our goal is to address the associated Input/Output (I/O) and storage challenges in the context of current and emerging storage landscapes, and expedite insights into mission critical scientific processes; which is associated with theme two of the Funding Opportunity Announcement.