SDAV

Scalable Data Management, Analysis, and Visualization

SDAV Institute Final Report – February 2017

http://sdav-scidac.org

Table of Contents

1 Introduction

The purpose of the SDAV institute is to provide tools and expertise in scientific data management, analysis, and visualization to DOE's application scientists. Our goal is to actively work with application teams to assist them in achieving breakthrough science, and to provide technical solutions in the data management, analysis, and visualization regimes that are broadly used by the computational science community. Over the last 5 years members of our institute worked directly with application scientists and DOE leadership-class facilities to assist them by applying the best tools and technologies at our disposal. We also enhanced our tools based on input from scientists on their needs. Many of the applications we have been working with are based on connections with scientists established in previous years. However, we contacted additional scientists though our outreach activities, as well as engaging application teams running on leading DOE computing systems.

Our approach is to employ an evolutionary development and deployment process: first considering the application of existing tools, followed by the customization necessary for each particular application, and then the deployment in real frameworks and infrastructures. The institute is organized into three areas, each with area leaders, who keep track of progress, engagement of application scientists, and results. The areas are: (1) Data Management, (2) Data Analysis, and (3) Visualization. This report is organized along these areas. However, often there are multiple technologies from these areas that are applied to a single application need. These are described on a case-by-case basis in the appropriate sections. This final report for the SDAV Institute covers the 5-year period from February 2012 to February 2017.

This report covers activities in all SDAV institutions listed next. Laboratories: Argonne National Laboratory (ANL), Lawrence Berkeley National Laboratory (LBNL), Lawrence Livermore National Laboratory (LLNL), Los Alamos National Laboratory (LLNL), Oak Ridge National Laboratory (ORNL), Sandia National Laboratory (SNL). Universities: Georgia Institute of Technology, North Carolina State University, Northwestern University, Ohio State University, University of Oregon, Rutgers University, University of California at Davis, University of Utah. Industry: Kitware.

2 Executive Summary

This report is organized into sections and subsections, each covering an area of development and deployment of technologies applied to scientific applications of interest to the Department of Energy. Each sub-section includes: 1) a summary description of the research, development, and deployment carried out, the results and the extent to which the stated project objectives were met; 2) significant results, including major findings, developments, or conclusions; 3) products, such as publications and presentations, software developed, project website(s), technologies or techniques, inventions, awards, etc., and 4) conclusions of the projects and future directions for research, development, and deployment in this technology area.

The institute's web site, http://www.sdav-scidac.org/, contains in addition to publications, highlights, outreach and awards, the home pages for all the tools developed by members of the institutes and used by a variety of scientific applications. These home pages are extensive, and contain user guides, download instructions, and methods to communicate with developers and users. Rather than replicating this information in this report, we provide concise information on the problem areas each of the tools address, the approach for providing solutions to these problems, and the use of the tools by the scientific community.

We provide next a summary of key achievements, organized according to the structure of this report. We start with the Data Management area (section 3.1), followed by Data Analysis area (section 3.2), and ending with the Visualization area (section 3.3).

A key technology in the Data Management area is I/O frameworks (section 3.1.1). I/O can be a major bottleneck when running large scale computations (simulations, analysis, visualization) on large parallel machines. When each processor reads/writes data independently from other processors, this overwhelms the I/O system with too many concurrent I/O operations, resulting in poor utilization of I/O resources and thus long I/O times. I/O frameworks intercept I/O operations, keep the data in memory when possible, and thus isolate the I/O from computation. This allows computation to continue while the I/O framework can combine I/O from multiple nodes into larger I/O operations and schedule read/write operations independently of the computation. This approach has proven to achieve from 10 to 100 fold improvements using one of SDAV's I/O frameworks, called ADIOS. ADIOS has been used with a large variety of scientific applications, about 15 different simulation codes so far. Another approach is to exploit the topology of the system interconnect, thus reducing the cost of global synchronization. This approach, taken by the GLEAN framework, showed multi-fold improvement over existing implementations. Another key technology for understanding the effect of I/O on the global efficiency of a running code is to monitor I/O operations with little interference. SDAV has been using such a tool, called Darshan, and improving it over the years to provide accurate I/O characterization. It is being used as a standard tool on all major DOE supercomputing facilities.

Another advantage of capturing I/O operations, is that this allows the data to stay in memory for a period of time, and therefor enable operations over the data before it is written out to disk. This is one method for enabling *in situ* processing. Several activities can take place while data is still in memory. One is enabling coupling of codes while data is still in memory while two or more codes are running concurrently on different node allocation on supercomputers. Such a tool is DataSpaces, which facilitates data exchange and transformation between codes while the data is still in memory. Another *in situ* activity provided by the FlexPath tool is scheduling and streaming data. Naturally, these tools can be embedded in an I/O framework, and indeed they were integrated into ADIOS. Other activities that can be performed *in situ* processing is in "on-the-fly" analysis while data is still being generated (e.g. the first few time steps of a simulation), visualizing data *in situ*, generating indexes *in situ*, etc. These techniques have shown to enable scientific computations and explorations that were previously impossible (see section 3.1.2).

An important challenge for scientists dealing with large volumes of data is to identify important structures and patterns in the data. For example, repeated searching for a pattern from trillions of particles in a plasma simulation or billions of particles in a cosmology simulation is prohibitive because each search can take many minutes. Indexing the data (a one-time cost) so that it may be searched multiple times is essential. The challenge is to have an indexing method that allows the index to be built quickly with a single run over the data, consume limited space to store, and enable very efficient searching for patterns of interest to our scientist partners. Such an index capability that runs in parallel over the data, called FastQuery, was developed by SDAV and has shown to be highly efficient (and provably optimal) for scientific data (section 3.1.3). As an example, it has been used to search over a trillion particles in 10 seconds for a magnetic reconnection application. Over the SDAV period it was enhance to handle multiple popular file formats (including HDF5 and PnetCDF), was embedded into the ADIOS framework, and was extended to work with AMR data.

Another aspect of data indexing is to first compress the data (both lossy and lossless compression), and then develop effective indexing over the compressed data. Such tools, called ALACRITY and ISOBAR, are described in section 3.1.4, and shown to achieve compression as low as 55% of the original data, and having low space overhead for the index (less than 25% of original data), while achieving 28-35 fold search improvement over uncompressed indexes.

In section 3.1.5, we describe work in support of popular parallel I/O tools and file formats that serve a significant fraction of I/O requirements from DOE applications. In support of the HPC software stack, SDAV has maintains and improved the ROMIO MPI-IO implementation, which is available on every DOE supercomputing platform. The most impressive improvement was the redesign the ROMIO internal data structures to reduce memory utilization in previously worst case scenarios to only 1.1% of its original size. Furthermore, working with IBM to tune the ROMIO algorithms, the performance improved by as much as 15 fold. In addition, one of the parallel file formats that uses MPI-IO internally, Parallel netCDF (PnetCDF), was developed by SDAV. It is an efficient parallel implementation of the classic netCDF format used by several scientific communities, most notably the climate research community. PnetCDF is available on all production DOE supercomputer facilities. SDAV also provides support for an innovative file structure, based on a hierarchical multi-resolution data format. The data format is a referred to as z-ordering space filling method, which lends itself to extract the data at different resolutions with minimal overhead. The software, called PIDX, is available as a library, installed on all DOE supercomputing facilities, and provides fast querying at the desired resolution for data analytics and remote streaming. It has been shown to be able to stream data at a sustained data transfer rate equal to

80% of the nominal I/O bandwidth of the hardware, which is extremely important for visualization tools running remotely.

Data Analysis (Section 3.2) is the process of discovering interesting features and patterns in data. This process usually requires the transformation of the data in order to highlight the features of interest, including selection of subsets of the data based on features, summarization, statistical methods, representation of topological features, and tracking evolution of patterns over time and space. In SDAV, we have a collection of tools that use such techniques. Section 3.2.1 describes techniques for accelerating visual analysis and exploration. These include special data structures that provided nearly linear scaling of the calculation of distance fields (a technique to visually represent 2D and 3D isosurfaces). This technique has been used on up to 350 million triangles for astrophysics, climate, combustion, cosmology, and fusion applications. In addition, such techniques have been enhanced to run *in situ* in order to enable sophisticated volume feature exploration in large-scale data. Another visual exploration method is to track the trajectory of particles over time. Such a tool, running in parallel, called Ultravis-P, was developed and used since the beginning of SDAV. However, since it is often better to see the trajectory in the context of the spatial-temporal volume, a new trajectory-based flow feature tracking method was developed for joint particle/volume datasets. It was used for spatio-temporal particle track visualization of combustion and atmospheric datasets.

Another class of data analysis methods focus on identification and tracking of topological features in data (section 3.2.2). One such tool, called MSCEER, leverages Morse-Smale complexes to explore surface interactions of materials and was applied to multiple domains, including analysis of clean combustion and design of battery materials. Another topology-based tool was designed for the exploration high-dimensional spaces. The tool, called ND^{2}AV, helps the visual and analytical exploration of problems that have a large number of parameters, as is common in many scientific areas. Recently, it has been applied to high energy density physics, and to nuclear reactor safety problems.

Flow field analysis and visualization techniques are described in section 3.2.3. They are used for a wide variety of spatio-temporal analysis tasks for a range of science applications. As the volume of data from simulations increases, it is essential that such techniques be accelerated by parallel computation methods. Such techniques were developed in SDAV, and have been shown to scale linearly. Furthermore, lower levels of precision can be processed in real time, and full precision can now be processed in the order of only 10 seconds for large scale datasets. Another activity for scalable stream surface computation was integrated into a library called OSUFlow, and was integrated into the VTK framework.

Related to flow field analysis is the ability to track the temporal history of features in space (section 3.2.4). A specialized tool, called TALASS, was designed to show the progression of particle tracks over time and was applied to many areas, including turbulent combustion research, astronomical simulations, weather analysis, and plasma physics. Another activity took the approach of controlling the execution of a plasma-based particle acceleration code (called WARP) from the VisIt tool, so that scientists can drive analysis *in situ*, while data is still in memory. This resulted in the ability, for the first time, of the scientists to observe in real-time the progression of the accelerator wake-field in 3D.

As data volumes increase, visual data exploration tools need to be improved to make best use of processing technologies. These activities are described in section 3.2. The two primary visualization frameworks, VisIt and ParaView, used by scientific communities well before SciDAC-3, needed to be enhanced to take advantage of multi-core and many-core architectures. Some of the enhancements could be made internally to each framework. However, it was the goal of SDAV to share a common framework for taking advantage of multi/many core architectures, reducing the long-term maintenance cost of these tools and ensuring new developments are available in both toolsets. This activity led to a remarkable achievement in the visualization area: a single framework, called VTK-m (see section 3.3.2), now provides multi/many core operators to the two main parallel visualization tools provided by the institute. This achievement required to cooperation of SDAV researchers from five DOE laboratories and Kitware (which provides support for VTK) as well as NVIDIA, and the integration of three different projects with VTK. The three existing projects that were integrated into VTK-m are: Piston which focused on efficient and portable algorithms, Dax which provided a top-down framework that simplified development, and EAVL which provided advanced data structures. The ingenuity and persistence of the VTK-m developers led to the single framework described above and thus eliminated much duplication of effort. This enables very large scale analysis and visualization, such as taking advantage of GPUs to find cosmology halo centers performed on a half-trillion particle simulation run on 16,384 nodes. This led to the ability for the first time to measure concentration-mass relation from a single simulation volume over such an extended mass range.

The details of complementary work to have VisIt and ParaView interoperate with VTK-m is described in section 3.3.1. These sections also describe other scaling activities. Examples of VisIt accomplishments include: accelerating AMR visualization for FastMath applications, developing special crack-free isosurface extraction for ice sheet simulations, and extracting and visualizing features from one trillion particles for a magnetic reconnection application. Similarly, examples of ParaView accomplishments include: providing *in situ* analysis with CosmoTools, visualizations of helium bubbles and tungsten cavities, and extracting, tracking, and visualizing vortex dynamics for a superconductor simulation.

SDAV also provided a tool that facilitates the development of specialized analysis and visualization on distributed-memory machines. This tool, called DIY (Do-It-Yourself), provides a programming model and runtime for block-parallel analytics for building parallel analysis routines. It decomposes data into blocks, assigns the blocks to processors, and supports communication between blocks. Over the past 5 years, DIY was enhanced to use a special internal structure that balances the computation on the processors, thus achieving great efficiency, as high as 50 times faster than the previous version. Because of the high efficiency of the tool, it was used to developed specialized libraries that require a high degree of computation, including libraries to compute parallel Voronoi and Delaunay tessellations of particle datasets, N-body simulations, molecular dynamics codes, and LIDAR point clouds. DIY is shaping up to be the solution for inter-node parallelism within VTK, while VTK-m addresses intra-node parallelism.

SDAV activities include flow visualization methods, described in section 3.3.3. These tools allow researchers to analyze and visualize particles paths and are needed in applications such as in ocean, climate (ice flow), combustion, and fusion. This section describes new methods to compute the shrink and stretch lines which are found by evaluating the associated eigenvectors, methods to increase accuracy for uncertainty in particle tracing, *in situ* identification of salient flow features, and techniques for accelerating flow visualization on GPUs.

Volume rendering is a powerful tool for visualizing complex 3D flow fields. SDAV provides such tools to the scientific community (see section 3.3.4) and enhanced these tools based on user needs and advanced architectures. This section describes tools that were developed and enhanced to support specific needs of the scientific communities we interacted with. For example, techniques were developed in VisIt in cooperation with NVIDIA to avoid copying data between the CPU and the GPU, which led to a highly efficient new compositing algorithm. Other work was to develop rendering algorithms for AMR that are applicable to climate modeling, fusion simulations, and astrophysics simulations, among others. Another special rendering method was developed for an advanced illumination method for interactive visualization of 3D unstructured-grid data. This 3D visualization enhancement is timely as unstructured grids are becoming increasingly popular for a variety of scientific simulation applications. In addition, SDAV has been developing and employing a tool that is optimized for parallel particle rendering, called vl3. During the SciDAC-3 period, it evolved to render very large AMR datasets, to support large-scale ray casting volume rendering, and was coupled to molecular simulation code (LAMMPS) to provide *in situ* visualization of large-scale atomistic data.

More recently, we have tackled a difficult area that aims to represent visually the uncertainty of ensemble simulations. Specifically, given a set of related simulations where initial parameters may vary or different models are run (referred to as ensembles), the problem is how to best represent the uncertainty. To support this capability, a new operator was developed in VisIt, called “Statistical Trends”. It allows for a user to apply different statistical functions to look for trends in data over time. In addition, a new method was developed for determining the uncertainty in the models by comparing each ensemble to a Bayesian Model Average. This method helps to determine where the simulation outputs produce viable distributions. This work was used to determine uncertainty in climate simulations as well as groundwater simulations of uranium seepage.

Finally, the SDAV institute’s productivity deserves special mention, not only in terms of tools developed and applied to a large number of application domains, but also culminating in 451 papers published over the 5 year period, and over 60 tutorials given in major conferences and other venues (see section 4.3). Most of the tools are installed and used routinely on DOE’s leadership-class computing facilities: ALCF, OLCF, and NERSC (see section 4.2 for details). The 38 highlights shown on the SDAV web site (see: http://www.sdav-scidac.org/highlights.html) provide an insight into the breadth of technologies as applied to a large variety of scientific applications.

3 SDAV Achievements

SDAV has multiple technologies that fall into the following categories: I/O Frameworks, *In Situ* Processing, Indexing/Compression, Statistics and Data Mining, Analysis and Visualization Frameworks, Analysis and Visualization Libraries, and Multi-/Many-core Visualization Libraries. The technologies that fall into each category are available through our web site (http://sdav-scidac.org/toolkit.html) along with documentation, user guides, and points-of-contact information. These technologies (frameworks, tools, libraries) are referred to throughout this report without introductory materials because these were provided in previous reports, and are readily available at the SDAV web site).

All the publications mentioned in this report are can be found on the SDAV web site (http://sdav-scidac.org/publications.html), and in the Appendix.

3.1 Data Management

The data management activities within the SDAV Institute have generated new technologies and frameworks for I/O management, in situ processing, code-coupling, indexing, compression, parallel I/O, and file formats. These have significantly improved DOE application scientists’ ability to efficiently manage their data, in a variety of science applications, including fusion science, climate, cosmology, and combustion.

3.1.1 I/O Frameworks

3.1.1.1 ADIOS (ORNL)

Our research was focused on the following areas: integration of *in situ* analytics and visualization into Input/Output (I/O) frameworks, high-level query interface for large datasets, and enhancement of writing and reading data in large scale simulations. In all three areas we have had great results that were published at various conferences. We have worked extensively with the Rutgers and Georgia Tech teams to research and develop *in situ* processing capabilities, and integrated our tools (DataSpaces, FlexPath and ADIOS) to provide a framework for users. We have worked together with the LBNL and NCSU teams to design a query interface for finding and reading the data of interest in large datasets. We integrated our tools (FastBit, ALACRITY and ADIOS) and developed the indexing and querying capability for the I/O framework. We have improved traditional file-based I/O as well by using information about the network topology of the LCF machines (Mira and Titan) and by aggregating write and read operations spatially and temporally that decrease the load on file servers.

During the five years of this project, the ORNL team has regularly released the ADIOS software that made our research results available to the users. Our products included a unified API for staging and file I/O, a staging solution, a visualization schema, and an I/O-skeleton application generator with standardized measurement; a new aggregated read method that scaled up better than the original method, a staged writing tool used by applications for asynchronous I/O, and spatial aggregation for output; a transformation layer providing various lossless compression and local data reorganization techniques, integrated visualization routines using EAVL/VTK-m, and extensions to support Adaptive Mesh Refinement; improved topology-aware I/O on Bluegene/Q and Titan, wide-area-network staging; Python/Numpy interface and tools; the Query and Indexing interface, supporting three implementations based on FastBit, ALACRITY and on simple min/max statistics, a recovery tool for damaged datasets, and staging I/O at the full scale of Titan; lossy compression with ZFP (Peter Lindstrom at LLNL), and temporal aggregation for improved output I/O overhead.

The team has also been interacting with many applications and scientists using DOE resources over the project period and helped them integrating ADIOS into their applications for their data processing needs: XGC (SciDAC EPSI project, C.S. Chang, PPPL), PIConGPU (Michael Bussman, HZRD), SPECFEM3D_GLOBE (Jeroen Tromp, Princeton Univ.), RTM (Pierre-Yves Aquilanti, Total E&P Research & Technology USA, LLC), LAMMPS (Steve Plimpton, Sandia), OpenFOAM (Yi Wang, FM Global), QMCPack (Paul Kent, ORNL), BoxLib/LMC (Marc Day, LBNL), Warp (Remi Lehe, LBNL), Qlg2q (Min Soe, RSU), Chombo (Brian Van Straalen, LBNL), S3D (Jackie Chen, Sandia), Chimera (Bronson Messer, ORNL), AWP-ODC (Yifeng Cui, UCSD), M3D-K (Gou Yong Fu, PPPL), Fine/Turbo (Mathieu Gontier, Numeca), GTS (Stephane Ethier, PPPL), GTC-P (William Tang, PPPL), Pixie3D (Luis Chacon, LANL).

Moreover ADIOS with DataSpaces and FlexPath was used in *in situ* coupled applications and staging I/O. These included nuclear waste disposal (EFRC-WastePD project, Wolfgang Windl, OSU), combustion (Jackie Chen, Sandia), fusion (SciDAC-EPSI, C.S. Chang, PPPL), subsurface modeling (Mary Wheeler, Univ. Texas), replica exchange (Emilio Gallicchio, CUNY) and fluid flow using FEM + AMR (Baskar Ganapathysubramanian, Iowa State Univ.).

For the great number of applications using ADIOS with great improvements to their I/O needs, the ORNL team received an R&D 100 Award in 2013. The ADIOS software is available at the OLCF software website, along with GitHub. We have over ten collaborating institutions contributing to the ADIOS code-base, and are constantly adding new features while retaining rigorous testing and Quality Assurance. ADIOS releases can be found at the ORNL website: https://www.olcf.ornl.gov/center-projects/adios/.

3.1.1.2 GLEAN (ANL)

GLEAN facilitates simulation-time data analysis and I/O acceleration. It takes application, analysis, and system characteristics into account to accelerate I/O and to interface with running simulations for *in situ* analysis and co-analysis with zero or minimal modifications to the existing application code base. For I/O, GLEAN fully exploits the topology of the system interconnects and algorithms to reduce the global synchronization for improved performance.

GLEAN components and algorithms are routinely used by the HACC Cosmology science application for production science runs on the various DOE Supercomputing computing systems, including at ALCF and OLCF. The GLEAN I/O is used as part of the HACC I/O CORAL benchmark. This benchmark is used in the DOE CORAL project performance evaluations. We have worked closely with IBM to incorporate GLEAN's topology-aware data movement, reduced synchronization algorithms, and I/O aggregator placement heuristics in the IBM Blue Gene/Q ROMIO implementation of MPI-IO. This has been integrated by IBM and has resulted in multi-fold improvement over the existing implementation. Furthermore, topology aware I/O mechanisms of GLEAN has resulted in multi-fold improvements for the WRF I/O benchmark (Weather Research and Forecasting Model – often used by weather/climate scientists) on the ALCF Blue Gene/Q system as well as on the NERSC Edison system.

Our work in I/O acceleration has resulted in several publications and is part of a Gordon Bell Finalist in 2013. For in situ analysis, we have formulated coupling of *in situ* analysis together with an application as a mathematical optimization problem taking into account the system resource characteristics and the characteristics of the application and the various analytics. This has been tested on the Blue Gene/Q system with the LAMMPS and FLASH simulation, and has resulted in an optimal *in situ* workflow for the applications. This has also resulted in publications at Supercomputing 2015 and 2016. Additionally, GLEAN technology was a critical component of the best paper award given at 21st ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC) 2013.

In conclusion, topology-aware I/O is critical in order to achieve scalable performance on leadership supercomputing systems. For *in situ* analytics, we demonstrated that one needs to carefully co-schedule these with simulations taking into account resource characteristics, application characteristics, and the objectives of the analytics. These two areas need more research so as to account for future system characteristics as well as to meet the goals of new applications and analysis.

3.1.1.3 Darshan (ANL)

Darshan is a lightweight I/O characterization tool that transparently captures I/O access pattern information from scientific applications. Darshan can be used to tune applications for increased scientific productivity or to gain insight into trends in large-scale computing systems.

As part of the SDAV project, Darshan has evolved to provide seamless support for a variety of computing platforms, including IBM BG/P, IBM BG/Q, Cray XE, Cray XC, and Linux clusters. The Darshan development team has produced frequent production-quality software releases, culminating most recently in the latest release in November 2016. Darshan is now poised for future deployment on pre-exascale and exascale systems, notably by adopting a modular architecture that enables rapid integration of instrumentation for upcoming data libraries and system services. A regression test suite has also been created and is executed nightly on production infrastructure at ANL for ongoing validation as new capabilities are added.

Darshan is used at the majority of the world's large-scale computing facilities, including officially supported installations at the ALCF, OLCF, and NERSC. The ALCF and NERSC have notably enabled Darshan for automatic instrumentation of all production jobs by default. This broad adoption has enabled Darshan to play a critical role in improving the performance of numerous critical applications and libraries, including HACC (cosmology), Flash (astrophysics), CESM (climate), and HDF5 (data management).

3.1.2 *In Situ* Processing and Code Coupling

3.1.2.1 DataSpaces (Rutgers)

The SDAV team at Rutgers University focused on designing, developing, and deploying in-memory data staging on leading DOE systems and integrating them with applications. Specifically, the Rutgers team focused on designing and implementing programming abstraction models as well as underlying adaptive runtime management strategies and developed the DataSpaces tool to support scalable hybrid staging. Hybrid staging uses resources (CPU and memory) on nodes running the applications as well as on dedicated nodes to stage, reorganize, and manipulate data. It enables multiple codes running on the same system that are part of a workflow to exchange and transform data in-memory. The Rutgers team also explored the in-memory staging-as-a-service deployment model. The overarching goal of the effort was to enable applications to construct more complex end-to-end coupled simulation workflows. Key research and development achievements include:

- Design, development, deployment, dissemination and evaluation of the DataSpaces scalable hybrid staging framework on leading DOE systems and its integration with ADIOS

- Design, development and evaluation of autonomic runtime mechanisms in DataSpaces for the optimization of staging-based *in situ* (and *in transit*) workflows, including mechanisms for staging-based (DART) and point-to-point (DIMES) high-throughput/low-latency asynchronous data exchange, task placement and scheduling in hybrid staging, data placement and movement across deep memory hierarchies, dynamic code deployment, staging over LANs and WANs, and a service-based abstraction for staging
- Integration of DataSpaces with applications to support end-to-end loosely and tightly coupled application workflows (including integration of online analytics) in a range of application domains including Fusion, Combustion, Subsurface Modeling, Fluid Dynamics, Chemistry, and Material Science.

The DataSpaces-related activities have led to the dissemination of the DataSpaces software, its deployment on supercomputing systems at ORNL, ANL and NERSC, and its integration with ADIOS, as well as 22 publications, 10 keynotes/plenary presentations, and 16 invited presentations. The Rutgers PI served as chair or in another leadership role for 14 conferences/workshops. The project enabled the research of 5 Ph.D. students (2 graduated) and 1 Postdoc.

The project demonstrated that staging can effectively address data-related challenges at very large scales and can efficiently enable coupled simulations and novel *in situ* application workflows across a range of applications areas.

3.1.2.2 FlexPath (GTech)

Over the 5-year period Georgia Tech (GT) has made a number of contributions to the state of the art in *In Situ* Processing and Code Coupling. The core of these efforts have been in creating the FlexPath (originally FlexIO) data staging transport for ADIOS. Flexpath is a type-based publish/subscribe infrastructure for coupling high-end scientific applications with their online analytics services. Flexpath's pub/sub approach makes possible runtime configurability, scalability, and also fault tolerance, as the pub/sub abstraction allows for the decoupling of diverse analytics components, permits multiple subscribers or publishers to share a single data stream, and suppresses communications for cases in which there are no subscribers to certain data streams (e.g., those not of current interest). These properties contrast with the typical assumptions made by communication infrastructures like MPI, where the domain of executing processes is initialized at launch and cannot grow or shrink for the remainder of the execution. This technical contribution is achieved by using direct connections between interacting components, including the scatter-gather or MxN communications needed across different communicating internally parallelized analytics components.

Architecturally, FlexPath is built using EVPath, an event transport middleware layer that is released as open source software for the developer community. Specifically, it is designed to allow for the easy implementation of overlay networks, with active data processing, routing and management at all points in the overlay. During the duration of the SDAV project, Georgia Tech created and/or tuned three pluggable EVPath transports to make use of new networking technologies. These transports include a raw InfiniBand transport that directly used the Verbs interface. Another transport was based upon NNTI, the RDMA transport developed at Sandia National Laboratories to support the TRIOS system and capable of running across both Infiniband and the uGNI interconnect for Cray machines. Lastly, to leverage industry commitment to a unified communication interface, GT developed an EVPath transport based upon the emerging ‘libfabric’ interface that supports many current and emerging RDMA networks.

Flexpath is deployed for use across a range of high end machines, including ORNL's Titan machine, Infiniband clusters, and commodity scientific computing engines. It has been a part of official ADIOS releases since 2013. Flexpath has been used for code coupling or *in situ* processing with a variety of scientific applications, including LAMMPS, GTS, S3D, Maya and PIConGPU. We also created Flexpath-enabled visualization plugins for VisIt. In 2014, the Georgia Tech SDAV team was involved in a demo at SuperComputing 2014 in New Orleans that involved the A*STAR Computational Resource Centre in Singapore and a trans-oceanic InfiniBand link, demonstrating SDAV technology capability on a global scale. In the GT-led demo using FlexPath, the Sandia-led molecular dynamics code LAMMPS ran in Singapore, sent data to an analysis workflow hosted at GT, the output of which was then delivered to a VisIt-based display in New Orleans. In addition, GT supported the hosting of SC demonstrations in 2014, 2015, and 2016 for experimental fusion data processing, developed at ORNL and LBNL, and automated tumor labeling in images, developed at Stony Brook. All three of these ADIOS-based demonstrations utilized transports that depended on one of GT's SDAV software stacks, EVPath.

Additionally, motivated by end scientists' requests for development of a scalable/portable way to access their data on end devices, Georgia Tech created a new staging method for ADIOS, utilizing the in-line indexing and analysis methods that already existed, but based with storage in cloud environments like Amazon's EC2. This "SciBox" project offers the ability to bridge to an interesting array of scientific analysis and visualization usage scenarios that are not yet well understood from a platform perspective.

3.1.3 Indexing

3.1.3.1 FastQuery and Indexing (LBNL)

Over the past five years, FastQuery indexing and querying software has been extended in a number of ways to meet the applications' demand [Lin2013a]. It has been shown to scale well in extracting information from massive amounts of data stored in scientific file formats such as HDF5 and ADIOS-BP. Our work has produced more than 20 technical publications. In this brief summary we highlight its uses in three DOE science applications: plasma physics simulation, ice sheet modeling, and particle accelerator design modeling.

Magnetic reconnection is an important mechanism that releases enormous energy as field lines break and reconnect in plasmas. To understand this phenomenon, plasma physicists run large-scale simulations and track the trajectories of trillions of plasma particles through a magnetic field. However, they are particularly interested in a relatively small numbers of highly energetic particles (still in the millions) at the intersections of magnetic field lines. Locating these particles from the trillions of particles was a challenge for the physicists. FastQuery software was able to index more than a trillion particles in about 10 minutes and then locate those particles in a few seconds [Byn2012a]. This capability to locate the "interesting" particles in seconds gave the physicists a new way to analyze magnetic reconnection.

In ice sheet modeling, an important task is to identify the icebergs that are breaking away from the ice sheet. This is known as the ice calving problem. The existing approach goes through each cell of the ice sheet model repeatedly, which can be very time consuming. We divided the ice calving problem into two steps, first to identify cells with thick ice, and then to connect these cells into connected regions. The regions of ice that are not connected to the ice sheet are labeled as icebergs. FastQuery could easily handle the first step, but we had to develop the algorithm for the second of identifying connected regions in adaptive mesh refinement (AMR) data. We developed an efficient parallel connected component labeling algorithm. Overall, we were able to accelerate the ice calving calculation by a factor of 6 [Dev2016a, Zha206a, Zou2016a].

In working with accelerator design software, FastQuery was used to track particles that stray away from the center of the particle accelerator. These particles are known as halo particles among the accelerator designers. In this case, FastQuery indexed billions of particles over about 800 time steps to support the user queries. The query results revealed an unexpected increase in the number of halo particles, which gave the designers a chance to correct the problem and reduce the number of halo particles [Chi2014a].

In conclusion, responding to scientists' needs, led us to enhance the FastQuery software and turn it into a powerful parallel search engine that was used on many problems in various scientific domains. We expect in the future to apply this tool to additional scientific domains, which inevitably will have unforeseen challenges.

3.1.4 *In situ* data compression

3.1.4.1 Analytics-driven data layout optimizations and Compression (NCSU)

Over the 5-year period, we have been working on a number of issues that address scientific data indexing, compression and storage layout problems. Moreover, in addition to focusing on mainstream uniform mesh data, we have also addressed Adaptive Mesh Refinement (AMR) data.

In the domain of indexing and query, we first developed ALACRITY [Jen2012a]. It produces an inverted index and index metadata, by exploiting floating point number representation. It performs at speeds suitable for in-situ processing (150-225MB/s), while maintaining a small data footprint (data+index < 125% original data). Based on ALACRITY, we have developed DIRAQ [Lak2013a], a parallel, scalable, *in situ* indexing technique. It enables smarter group-level index aggregation, which not only improves later query performance by generating less-fragmented indexes, but also achieves increased end-to-end write throughput due to overall data compression.

For compression, we developed a hybrid compute-I/O interleaving for in-situ parallel lossless compression technique, ISOBAR technology [Sch2012]. It can reduce the total time-to-disk by 12-46%, according to tests over real scientific data. Moreover, we have extended our earlier work on ALACRITY encoding [Jen2013a], by integrating inverted index compression and data reduction, resulting in a full-precision, query-optimized data representation that is 55-90% of the original data (for FLASH, S3D, XGC, GTS datasets). This includes a light-weight index that is 5-20% of the original data.

Regarding storage layout, we've developed MLOC [Gon2012a, Gon2012b], which provides flexible, interchangeable data layout optimizations for heterogeneous access patterns, allowing users to match data access patterns to their particular needs. Based on MLOC, we've developed a run-time multi-level data layout optimization framework (PARLO) [Gon2013a], which is integrated with ADIOS to achieve high-performance *in situ* data layout optimization at runtime. Moreover, we have extended ADIOS to include a generic, user-transparent data transformation framework that allows end users to apply *in situ* data transformations without changing their existing application [Boy2014a].

Moreover, we have studied various indexing, storage layout and I/O issues for AMR data. We have developed an indexing and querying framework to facilitate large-scale AMR data scientific discovery and exploration [Zou2016a]. It has an AMR-specific hybrid index methodology which captures both spatial and value aspects of the AMR data, and enables efficiently processing queries that have both a spatial selection and value constraints. We have also developed an *in situ* storage layout optimization method for AMR spatio-temporal read accesses [Tan2016a]. This work enables automatic selection from a set of candidate layouts based on a performance model, and reorganize the data before writing to storage. Moreover, in order to facilitate runtime AMR data sharing across scientific applications, we have developed AMRZone [Zha2016a]. It enables efficient in-transit AMR data analytics, by constructing a virtual and distributed staging space to avoid accessing data over a file system.

Given the rapidly growing volume of scientific data, data access latency reduction can greatly improve the performance of applications. According to our experience, data indexing, compression and storage layout can help to significantly reduce data access latency, by decreasing data size and facilitating specific queries.

3.1.5 Parallel I/O and file formats

3.1.5.1 PIDX (UUtah)

Over the lifetime of the project, we have developed a new capability implemented in the PIDX library. It involves a data movement and I/O infrastructure that has unique scaling properties on DOE supercomputers while allowing fast querying for data analytics remote streaming. The PIDX library has been installed and used on other supercomputers to demonstrate its effectiveness on a broader range of hardware platforms. In general, the PIDX library enables HPC applications to write distributed multi-dimensional data directly into a hierarchical multi-resolution data format with minimal overhead. Production read/write of restart dumps and remote streaming has been demonstrated at the exhibit floor of the Supercomputing conference for the last three years. For example, three installations at the DOE booth, the Utah booth, and the KAUST booth have shown direct use and remote monitoring of distributed processing including the Mira supercomputer at Argonne, Shaheen II at KAUST and the CHPC in Salt Lake City. In addition, a server at Lawrence Livermore National Laboratory was used to give real time access to petascale climate modeling data housed on NASA storage. These demonstrations have shown the complete workflow of a simulation running and rendering data live at a remote site (in this case SC exhibit floor) during its execution.

The most recent work include scaling runs on the entire MIRA machine (up 768k cores) while achieving a sustained data transfer equal to 80% of the nominal I/O bandwidth of the hardware. The PIDX library is now available to the general public (webpage http://www.cedmav.com/research/project/15-pidx.html) with an open source release accessible via the following GIT repository https://github.com/sci-visus/PIDX.

To complement the deployment of the I/O library for simulation, the team has deployed plug-in components that make it easy to use the file format in practice. This includes the VisIt distribution which now integrates a ViSUS reader that allows one to directly read files generated by the PIDX library. Moreover, the VisIt tool can directly take advantage of the multiresolution nature of the data format and by default loads a coarse version of an input to allow access to data that is potentially too large to fit in memory.

Future developments of the PIDX project will involve effective use of reduced precision and compression techniques to minimize the data movements. In addition, recent tests show that the PIDX library can also take advantage of novel hardware like the burst buffer, although more work is needed to reach a production quality deployment. We are also working towards extending the PIDX framework to support particle data.

3.1.5.2 ROMIO (ANL)

The ROMIO MPI-IO implementation, provided on virtually every HPC platform and used internally by HDF5 and PnetCDF, continues to provide the foundation of the HPC I/O software stack.

The fundamentals of ROMIO have been well-tested over 15 years. In SciDAC-3 we focused on improving ROMIO’s behavior at scale.

- Very large data descriptions were overflowing internal MPICH and ROMIO data structures and overwhelming some operating system routines. We audited the code for overflow conditions and devised workarounds for the operating system issues.
- ROMIO internal data structures could in some cases consume large amounts of memory. We experimented with compressing those data structures. The work demonstrated dramatic reduction (1.1% of original size) in peak memory overhead with a trade-off of 25% higher CPU time to process these data structures.
- We worked in collaboration with IBM to better tune ROMIO at the large scales of our Mira Blue Gene /Q machine. We also worked with applications to make better use of the MPI-IO interfaces ROMIO provided. These tuning and algorithmic tweaks improved performance in some cases by as much as a factor of 15.

The ROMIO team have conducted numerous tutorial at major conferences (such as Super Computing), and wrote multiple papers on major improvements to the system, its use by other tools, such as MPI and Parallel-NetCDF, and its effect on various scientific application codes [see recent publications: Lat2017a, Luo2015a, Sny2015a, Ham 2014a].

Storage systems continue to increase in complexity, with larger numbers of devices and new software and hardware components. ROMIO continues to provide an abstraction layer to hide these complexities. As new concepts such as burst buffers and I/O forwarding shift from research topics to deployed components on supercomputers, ROMIO must continue to evolve to meet this new landscape. Ongoing open research issues are how to build upon ROMIO's strengths while minimizing overhead at growing core-counts, how to carry out background I/O while not disturbing computational science applications, and identifying additional programming models which would benefit from ROMIO's proven algorithms.

3.1.5.3 Parallel netCDF (NWU)

Parallel netCDF (PnetCDF) is an I/O library that supports parallel data access to netCDF files. NetCDF, developed at Unidata, defines a self-describing and portable file format and a set of application programming interfaces (APIs) for storing and accessing scientific data. However, its APIs do not support parallel I/O to files in the classic netCDF file format. Starting from version 4.0, NetCDF chose HDF5 as the underneath mechanism to add parallel I/O feature, but the file format must be in the HDF5 format. To support parallel I/O for the classic format which is being used by the majority of user communities that use NetCDF in climate, geoscience, and others, PnetCDF was developed to enable parallel I/O to the netCDF files in classical formats with high performance on large-scale parallel computers.

For the past five years, the PnetCDF software has been continuously released to the public as open source, through five major releases, the last of which was in December 2016. Many important features based on user requests and performance enhancement have been developed and incorporated into the software, including support for C++ and Fortran 90, a new set of non-blocking APIs, a request-aggregation technique to enhance performance, and a new extended classic file format, named CDF-5. The CDF-5 format supports definitions and access of large variables with more than 2 billion elements.

Collaboration between PnetCDF developers and the netCDF team at Unidata has continued since 2015 with the focus of developing a software component in the netCDF library to enable it to read/write CDF-5 file format, in both sequential and parallel. Because netCDF has a much larger user group across different disciplines, such collaboration significantly leverages the broader impact of PnetCDF. The CDF-5 feature has been made available in a release of netCDF in January 2016.

We also established collaboration with the PIO (Parallel I/O) library developer team at the UCAR. The PIO library is used by the CESM and ACME projects, two major earth and climate system modeling frameworks funded by DOE. This collaboration involves developing new APIs in the PIO library based on user requests and performance/bug tracing in the underneath MPI-IO library.

We have initiated a collaboration with the System Software Research Team, led by Dr. Yutaka Ishikawa, at the RIKEN Advanced Institute for Computational Science in Japan. The purpose of this work is to investigate the I/O software and storage system performance on the K computer. We have developed a PnetCDF I/O module for a production real-time climate simulation application named SCALE developed by the climate scientists at RIKEN.

In the past five years, the PnetCDF library building process was improved and now supports various vendor compilers. The library is available on all DOE production parallel machines at ANL, ORNL, NERSC, etc. We plan to continue the PnetCDF development and maintain the established collaboration by reaching out to more application domain scientists and explore the opportunities for improving and expanding the PnetCDF library to support high-performance computing, data analysis, and parallel I/O.

3.2 Data Analysis

The data analysis activities of the SDAV Institute have yielded a number of novel algorithms and software tools that support numerous DOE SciDAC application scientists in fields including astrophysics, climate, combustion, climate, cosmology, fusion, and development of nano-materials. These include tools for visual exploration, topological data representation, flow field analysis, and feature-driven exploration.

3.2.1 Visual Analysis and Exploration Techniques

3.2.1.1 Large-Scale Volumetric Feature Exploration (UCD)

Distance fields play a critical role in visualizing and analyzing complex data and have applications in feature extraction, computer graphics, data indexing, and compression, and they can be used as “importance” fields to direct analysis and reduce data, especially in large volume datasets. We have developed a highly efficient parallel distance field construction algorithm to make it usable for large-scale *in situ* applications. Experimental studies on Hopper, a Cray XE6 supercomputer, using a large geometric model consisting of 350 million triangles demonstrate an overall parallel efficiency at 75.26% from 1024 to 8192 CPU cores. Studies on Hopper using a temperature isosurface from a turbulent combustion simulation demonstrate perfect parallel efficiency from 4320 to 34560 CPU cores. We have achieved the most scalable parallel distance field calculations. Such high efficiency is achieved with a new distributed spatial data structure, which we call *parallel distance tree* to manage the level sets of data and facilitate surface tracking over time, resulting in significantly reduced computation and communication costs for calculating the distance to the surface of interest from any spatial locations. The resulting technology can benefit many SciDAC application areas such as astrophysics, climate, combustion, cosmology, and fusion. A paper reporting this work was published in IEEE Transactions on Visualization and Computer Graphics [Yu2015a].

In addition to the distance field work, we have also developed volume feature extraction techniques for large-scale data that can function on a standard desktop PC. Such techniques allow researchers to reduce data based on their domain knowledge. Since it's still commonly desired by scientists to use their desktop PCs for data visualization and analysis tasks, an *in situ* data preparation solution to support interactive feature extraction from large-scale volume data is needed. We have developed an *in situ* supervoxel generation method, together with a new hybrid feature extraction technique which combines GPU-accelerated clustering with the multi-resolution advantages of supervoxels in order to handle large-scale datasets on standard desktop PCs, which usually have limited memory and bandwidth capacity. The method is based on a user-driven uncertainty-based refinement approach to keep extraction results at the desired level of detail. We demonstrate the effectiveness and interactivity of this technique with a number of application specific examples. A paper reporting this work has been presented at LDAV 2015 (receiving a best paper honorable mention) [Xie2015a].

Technologies like these have a broad impact on a variety of SciDAC research areas and can help to achieve the goal of enabling sophisticated volume feature exploration in large-scale data. Our work opens future research directions that involve developing polished visualization software packages that incorporate these techniques.

3.2.1.2 Visual Analysis of Large-Scale Trajectory Data (UCD)

We have developed new analytical means of exploring particle and trajectory data commonly found in SciDAC applications. First, we developed Ultravis-P for interactive visual analysis of trajectory data derived from 3D vector fields or large-scale particle simulations. It is composed of a powerful back-end engine for parallel clustering/classifying curves and a light-weight, front-end user interface for interactive analysis. Our parallel regression-model based clustering/classification algorithm can use a combination of

GPUs and CPUs. On GPU clusters such as *Dirac* at NERSC and *Lens* at ORNL, the parallel algorithm demonstrates scalable performance with millions of trajectories. We have used this system to analyze particles' behavior in turbulent combustion simulations [Wei2012b]. We have also added functionality for analyzing trajectory data derived from fusion simulations according to their geometric properties [Sau2013a].

Furthermore, we have also explored how interactions between particle/trajectory-based data and field-based data can enhance this type of analytical exploration. We have introduced a new trajectory-based flow feature tracking method for joint particle/volume datasets. This method utilizes the indexed trajectories of corresponding Lagrangian particle data to efficiently track features over large jumps in time. As a result, this method is especially useful for situations where the volume data is either temporally sparse or too large to efficiently track a feature through all intermediate timesteps. We tested our method using combustion and atmospheric datasets. A paper reporting this work has been presented at IEEE SciVis 2014 conference [Sau2014a] and was also invited to be presented at ACM SIGGRAPH 2015.

To further enhance this type of analysis, we have also developed a joint Eulerian-Lagrangian framework aimed at supporting operations that utilize both particle and volume data simultaneously. More specifically, the data organization schemes developed allow one to quickly query, sample, and operate with both reference frames in a large-scale setting and, if necessary, in an out-of-core manner. We have demonstrated the effectiveness of this new framework using a number of large-scale datasets provided by SciDAC science projects in cosmology, combustion, and fusion. A paper reporting this work was published in IEEE Transactions on Visualization and Computer Graphics [Sau2016a].

Recently, we have expanded this joint analysis to explore spatio-temporal features in both a particle and volume reference frame. Users begin with a feature of interest in one reference frame as a spatio-temporal origin and can extend analysis temporally into the opposing reference frame. This allows users to extract and explore features describing the interplay between both particle and volume data in space and time. This work also explores new ways of visually presenting such a feature to a user. A paper describing this work has recently been accepted by the PacificVis 2017 conference and will appear in IEEE Transactions on Visualization and Computer Graphics.

Alternatively, in a joint effort with scientists and Princeton Plasma Physics Laboratory (PPPL), we have been exploring the use of histograms to capture the velocity decomposition of particle data. This has important applications to Lagrangian-based flow data analysis since traditional visualization techniques are often subject to a trade-off between visual clutter and loss of detail, especially in a large scale setting. We showed that the use of 2D velocity histograms can alleviate these issues. We demonstrate our design with an interactive system to visualize the velocity decomposition of particle data from fusion simulations. We consider both a post-hoc setting with an on-the-fly sampling scheme and an *in situ* setting to maintain interactivity in extreme scale applications. A paper presenting our design and case studies published in the 2015 LDAV Symposium (receiving a best paper honorable mention) [Neu2015a].

We have also extended this work in developing new ways to explore time-varying properties of the 2D histograms. This is done by stacking the histograms over time into a 3D volume and using isosurfacing techniques to explore time-varying properties of the captured distributions. In this case, the isosurface represents a boundary between histogram bins with higher/lower frequency of samples compared to the isovalue. Furthermore, improvements have been made to the system’s ability to select groups of trajectories based on features in the histogram distributions. A paper reporting this work was published by IEEE Transactions on Visualization and Computer Graphics [Neu2016a]. We are currently working on deploying a polished visualization software to scientists at PPPL and plan to conduct a thorough usability study. Such a study will help us improve the system and answer fundamental questions about user perception and understanding of this new technology.

3.2.2 Topological Methods

3.2.2.1 Topologika: Topological Techniques for Scientific Feature Extraction (UUtah)

Over the last five years we have been developing a number of topologically based analysis techniques to extract and track features in a broad class of scientific applications. Some of these techniques are part of different software infrastructures (MSCEER, ND^{2}AV, TALASS) discussed in more details below. Other techniques are aimed at developing new application use cases and research prototypes targeted at individual scientific applications. Over the duration of the SDAV institute we have seen these techniques mature from the first *in situ* and in-transit analysis approaches [Ben2012b] to massively parallel implementations [Lan2014a] able to process an order of magnitude more information than previous approaches at less than 1% of the overall simulation costs. Simultaneously, these techniques have become accepted in the scientific main stream, leading to numerous publications in application focused journals [Gro2012a, Gro2013a] as well as open source code releases [Bre2016b] and integration into community tools, i.e. VisIt.

Topological analysis is the focus of an application partnership with a team from the BES office concerned with the analysis of first principle molecular dynamics simulations. This has already led to a joint publication [Bha2016a] with multiple others currently in submission. Finally, we have explored extending structural analysis to vector fields [Bha2014a, Bha2014b].

Going forward there are three clear directions to pursue: a) the continued shift towards *in situ* analysis will require concerted research and software development efforts to create the corresponding algorithms and implementations. These efforts will also provide a great opportunity to collaborate more closely with ongoing efforts in the exascale research project as well as run time focused projects; b) as topological techniques are becoming more accepted there will be continuous need to modify and adapt both the approach as well as the implementations to new research areas and simulation codes. Prime examples are the new MPAS grids of the high resolution climate models but opportunities also exists in molecular dynamics, biology, healthcare, etc.; c) a number of research areas in applications such as particle physics and turbulent flow can benefit significantly from topological techniques but will require the development of new approaches.

3.2.2.2 MSCEER: Morse-Smale Complex Extraction Exploration and Reasoning (UUtah)

One of the specific tools resulting from the five years of work is MSCEER, a set of software libraries and tools for data analysis and visualization based on the specific topological construct of the Morse-Smale Complex. This tool has been effective in the analysis of a number of scientific domains ranging from clean combustion [Gyu2014a] to the design of battery materials [Bha2016a, Gyu2015a, Gyu2015b]. The latter application, for example, has been driven by the representation of Ion Diffusion with DFT (Density Functional Theory) models. The work uses ab initio molecular dynamics (AIMD) simulations, which are increasingly useful in modeling, optimizing, and synthesizing materials in energy sciences. The solved Schrodinger's equation and the generation of the electronic structure of the simulated atoms as a scalar field is ideal for topological analysis based on the computation of the Morse-Smale complex. The discrete nature of the Morse-Smale complex computation used in the MSCEER allows analyzing first-principles battery materials simulations with no loss of precision. We consider a carbon nanosphere structure used in battery materials research and employ Morse-Smale decomposition to determine the possible lithium ion diffusion paths within that structure. Our approach is novel in that it uses the wave function itself as opposed to distance fields, and that we analyze the 1-skeleton of the Morse-Smale complex to reconstruct detailed geometric representations of diffusion paths. Furthermore, it is the first application where specific motifs in the graph structure of the complete 1-skeleton define features, namely carbon rings with specific valence.

For the application to combustion simulations, Morse-Smale crystals (elements of monotone flow in a 3D scalar field) are very well suited to define *dissipation elements,* a structure that has become widely used to understand turbulent flows [Gyu2014a]. Furthermore, the MSCEER libraries have been adapted and improved to meet specific requirements of various analysis tasks, such as invariance to mesh orientation [Gyu2012b] needed to compute topological elements of combustion simulations, extension to large-scale meshes with distributed computation [Gyu2012a], flexible identification and visualization of individual elements of the Morse-Smale Complex [Gyu2012c], and incorporating numeric integration of gradient trajectories into the robust combinatorial approach [Gyu2014b]. Although developed to solve specific problems for domain-specific analysis tasks, these techniques have proven to be applicable across domains: for instance, the ability to compute geometrically accurate dissipation elements in combustion simulations has been directly applicable to computing accurate Bader volumes in the study of battery materials.

Future developments in this field include scaling to large models without core techniques as well as parallel deployment for *in situ* processing that can lead to scientific insight together with massive data reduction.

3.2.2.3 ND^{2}AV: N-Dimensional Data Analysis and Visualization (UUtah)

Over the 5-year period we have seen a rapidly growing interest in high dimensional analysis and visualization techniques. In particular, many scientific applications are interested in exploring and understanding high dimensional parameter spaces in areas ranging from high energy density physics to nuclear reactor safety. We have developed a suite of generic tools to support these activities that are applicable to virtually all high dimensional data. These include, distortion guided visualization [Liu2014a] and its application to multi-variate volume rendering [Liu2014b], subspace-based visualization [Liu2015a] and a new way to identify important viewpoints in high dimensional data [Liu2016b]. Another line of research has explored the application of these technique to high energy physics [Bre2016b] and Nuclear Safety Analysis [Mal2016a, Mal2015a, Mal2015b] in collaboration with Lawrence Livermore National Laboratory. In addition, we have contributed well-received surveys of high dimensional visualization techniques to the community [Liu2015b].

In the longer term, the significance of high dimensional techniques will increase, driven by needs in uncertainty quantification, risk assessment and sensitivity analysis. This will require more computationally efficient techniques that are easily accessible to the scientific community as well as more intuitive interpretation of the results.

3.2.3 Flow Field Analysis

3.2.3.1 Parallelization of Flow Algorithms for Large-Scale Data Analysis (UCD)

We have been working on the parallelization of flow visualization algorithms, specifically related to integral curve methods. Integral-curve-based algorithms are particularly difficult to implement on distributed parallel systems, as the communication between processors eventually dominates the algorithms, and our results have shown that the computation time is dependent on the complexity of the dataset [Cam2012a]. We also worked on methods for representing flow visualization in particle-based systems. We have demonstrated new analytics for flow in development of the velocity gradient tensor [Obe2012a], and continue to develop new methods based on continuum-mechanics effects [Agr2013a]. This allows us to give the end-user visual information on the “stretching” of time surfaces in the flow.

Furthermore, working in conjunction with our colleagues at Lawrence Berkeley National Laboratory and the University of Oregon, we have developed new methods to store and visualize flow fields by use of flow maps. In general, we showed that a Lagrangian representation for data, where trajectories are kept at each data point, rather than vectors, combined with interpolation of trajectories, allows visualization algorithms to be more accurate, be faster, and use less memory than existing techniques. More importantly, these methods can be used in an *in situ* environment in a much better way than with existing methods. Our main performance results were presented at the 2014 LDAV Symposium [Agr2014a]. We continued to develop new methods for flow visualization based on this new representation that can be implemented on large-scale parallel systems with additional results presented at the 2015 SPIE Conference [Arg2015a, Arg2015b]. Furthermore, we have adapted these methods to compressible flow models and have papers published in IEEE Transactions on Visualization and Computer Graphics [Cha2015a, Cha2015b]. Next, we flushed out the differences in this representation, and the bottlenecks that are introduced. We have developed methods that utilize curve-based representations as the "data structure" for this method, and have worked out the interpolation methods to define pathlines from this representation. This work was presented at the 2015 LDAV Symposium [Buj2015b].

3.2.3.2 VTK/OSUFlow Integration (OSU)

As part of SDAV's goal to facilitate large scale flow visualization, our effort was to enhance the OSUFlow library with a scalable stream surface computation algorithm, which was published in SC'14 [Lu2014a]. While several parallel streamline computation algorithms exist, relatively little research has been done to parallelize stream surface generation. This is because load-balanced parallel stream surface computation is nontrivial, due to the strong dependency in computing the positions of the particles forming the stream surface front. In our algorithm, seeding curves are divided into segments, which are then assigned to the processes. Each process is responsible for integrating the segments assigned to it. To ensure a balanced computational workload, work stealing and dynamic refinement of seeding curve segments are employed to improve the overall performance. This is a joint work between OSU and scientists at ANL. We have integrated this algorithm into the OSUFlow library. Another addition to the VTK/OSUFlow library is the ability of interactive manipulation of flowlines and flow article. With a novel deformation model, we allow the user to minimize occlusions and maximize the ability to identify salient flow features [Ton2016a].

3.2.4 Feature-Driven Data Exploration

3.2.4.1 TALASS: Topological Analysis of Large-Scale Simulations (UUtah)

TALASS refers to a software framework to compute, track and explore the spatio-temporal history of features in scientific simulations. Over the 5-year period we have broadened the range of applications to new feature types, i.e. embedded features on surfaces [Wid2015a], new applications, i.e. halo finding [Wid2014a] and new mesh types [Lan2015a]. At the same time we have been hardening the software and deployed it to researchers at Sandia National Laboratory. TALASS, originally designed for turbulent combustion research [Ben2012a, Wid2012a, Wid2015b] has now been extended to astronomical simulations [Wid2014a], weather analysis [Wid2017a], plasma physics [Wid2016a] and health care [Wid2016b]. In addition, the software has recently been ported into the ViSUS analysis framework which allows it to integrate seamlessly topological-based analysis and tracking with large scale visualization.

Going forward, TALASS will require continued investments in the software stack to allow its deployment to a broader community. Furthermore, there exist a number of interesting opportunities to extent the application space into new areas, such as, climate research or health care.

3.2.4.2 Feature-based Analysis of Plasma-based Particle Acceleration Data (LBNL)

Computational accelerator modeling aims to understand the fundamental nature of particle acceleration under varying conditions. When performing such modeling, accelerator scientists generate vast amounts of model output from simulation runs on large-scale HPC resources. The focus of our work has been on enabling scientific knowledge discovery using two different approaches.

The first uses a post-processing approach, and consists of methods for automatic feature detection and classification to first locate interesting features in time-varying model output, then apply machine learning methods to locate similar features elsewhere in a data collection. In support of plasma-based accelerator modeling efforts, we developed a method and software toolset for automatic detection and classification of particle beams and beam substructures due to temporal differences in the acceleration process [Rub2014a]. The software implementation, *lwfapath*, supports both interactive and batch modes of operations to support different needs and use scenarios of the accelerator modeling community, and has been run at NERSC as well as on desktop-class platforms by accelerator modeling researchers.

The second uses an *in situ* approach, where visualization and analysis methods are coupled directly to an accelerator modeling code to avoid the complications and cost associated with writing, then reading data to/from persistent storage. In support of the WARP code team, we coupled VisIt with WARP to implement *in situ* methods for feature-based data analytics, efficient large-scale data analysis and visualization. This work enabled accelerator modeling researchers to see and analyze, for the first time, high-resolution 3D simulations and compare them with 2D simulations [Rub2016a]. Prior to this work, accelerator scientists were constrained by I/O to study only of 2D simulation models. The new software, *WarpIV*, has been publicly released and is installed at NERSC.

3.3 Visualization

Our well rounded portfolio of visualization activities aims to provide usable visual data exploration and analysis technologies for use by the DOE computational science community. We evolved our two primary visualization applications, VisIt and ParaView, to be multi-/many-core aware, deployed these applications at DOE supercomputing facilities, and provided direct user support to computational science users. We invested in maturing and applying key technologies (PISTON, DAX and EAVL) for achieving multi-/many-core capability over a longer time horizon, and leveraged these into a single software platform, VTK-m, that will have impact on the broader visualization and analysis community for at least the next decade. Finally, we focused on applied development efforts that produce new capabilities, driven by science needs that ultimately end up in software distributed to the computational science community.

3.3.1 VisIt and ParaView

3.3.1.1 VisIt (LBNL, LLNL, ORNL, UO)

Over the five-year period, we made many improvements to the VisIt visualization tool, both in support of new architectures and in new features for SciDAC stakeholders. We also made many releases of our software, and made sure that VisIt was installed and accessible at the LCFs and NERSC.

The most important architectural change for VisIt has been the rapid increase in parallelism within a compute node. At the beginning of SDAV, VisIt only supported MPI-parallelism, which we foresaw as a problem given architectural trends. In response, we have pursued two major activities. First, we have added “pthread”-style threaded parallelism to VisIt. This activity proved to be a promising direction, since many VisIt algorithms could be retrofitted for this type of parallelism with (relative) ease. Of course, this style of parallelism will not work on GPUs, and so this development arc was limited to improved performance on multi-core CPUs and support for the Xeon Phi. A second activity has been the incorporation of VTK-m into VisIt. This activity is the “long term” solution for shared-memory parallelism within VisIt, but also a much larger effort. While this activity is far from complete, VisIt is now able to interact with VTK-m and make use of its algorithms as they become available. As additional functionality is added to VTK-m, VisIt will be in a position to quickly take advantage. Going forward, porting VisIt algorithms to many-core architectures will essentially involve making a VTK-m implementation for that algorithm and then having VisIt invoke that implementation.

In terms of new features for SciDAC stakeholders, there are many activities to report. VisIt now can read SpecFEM3D data (in conjunction with ADIOS), resulting in the ability to visualize their full resolution data from Titan. VisIt's support for Adaptive Mesh Refinement (AMR) data has greatly improved, especially the type of AMR coming from the FASTMath SciDAC Institute. In particular, VisIt now has production-quality crack-free isosurface extraction, which was used to visualize ice sheets computed using the BISICLES climate simulation code. VisIt also now has an FTLE operator, which was added and optimized for Linda Sugiyama of MIT. We also improved infrastructure that benefits many stakeholders, in particular our *in situ* version of the code (LibSim) and substantial improvements to its particle advection module. To publicize these new features the VisIt team was active in outreach, including participation in 11 tutorials (5 at SC and 4 at ATPESC).

These improvements led to many successful usages of the code by SciDAC stakeholders. Two noteworthy examples are for laser back scatter simulations and magnetic reconstruction simulations. In the first example, VisIt was used to visualize 220 billion cell data and correlate multiple, complex, 3D phenomenon over time to bring out features of interest at key points in the simulation, such as where back scatter phenomenon took place. In the second example, VisIt was used to visualize 1 trillion particle data sets from the VPIC code, in combination with the H5Part and FastQuery systems. This allowed collaborators to explore energetic particle distribution, the formation of flux ropes and, and the alignment of the motion of energetic particles with the electric/magnetic fields.

3.3.1.2 ParaView (LANL, Kitware)

During the SDAV life-cycle, we had nine ParaView releases before the end of our project. There has been many major improvements in these releases. Some highlights include:

- Support for data parallel frameworks including Dax, Piston and vtkSMP. VTK-m, which unifies all of these technologies, will be included in ParaView 5.3.
- Development of a charting framework with support for symbol and equation rendering for annotation, and support for vector graphics output.
- Refactoring of the VTK and ParaView libraries to allow for the creation of smaller subsets of both frameworks. This was driven by the need to derive *in situ* ParaView libraries that fit the problem.
- Introduction of ParaView Catalyst, a ParaView based *in situ* library. Since its release, Catalyst have been integrated with DOE codes in many scientific domains, including Climate, CFD, Plasma, HEP, and Shock Physics.
- Numerous improvements to ParaView's visual analytics and quantitative analysis capabilities, including support for scatter plot matrices, parallel coordinates and histograms.
- Development of much improved ParaView and Catalyst User's Guides.
- Integration with ADIOS for post hoc IO as well as in transit visualization.
- Integration with DIY2 for the development of distributed algorithms.

We has six VTK major releases over the last 5 years. These releases included numerous improvements and new features in support of DOE needs, including ParaView and VisIt development. Highlights include a major rewrite of VTK's rendering engine to obtain large performance improvements (10x - 100x) on modern graphics cards and software rendering implementations, integration of OpenSWR for improved rendering on Intel Xeon Phi processors, integration with VTK-m, and major improvements to the VTK pipeline infrastructure for improved parallel processing.

We have collaborated with and supported various DOE science teams including

- VPIC: Kinetic plasma simulation, Bill Daughton, LANL, PI. Collaborations included post hoc visualization with ParaView and *in situ* analysis with Catalyst.
- HACC, Cosmology, Salman Habib, ANL, PI. Collaborations included post hoc visualization with ParaView, feature extraction and tracking algorithm development, *in situ* analysis with CosmoTools.
- Plasma Surface Interaction (PSI) SciDAC Application Partnership, visualizations of helium bubbles and tungsten cavities from LAMMPS simulations using VTK.

- MPAS: Climate simulation, post hoc and *in situ* analysis and visualization, algorithm development for feature tracking,
- CAM-SE: Climate simulation, post hoc analysis and visualization,
- OSCon - Optimizing SuperConductor Transport Properties through Large-Scale Simulation, SciDAC partnership with BES, deploying in ParaView a set of algorithms for extracting, tracking and visualizing vortex dynamics in large-scale time-dependent Ginzburg-Landau (TDGL) superconductor simulation data. We have also integrated IO functionality to load the simulation data into ParaView.

In conclusion, we have improved both the ParaView and VTK visualization tools in order to support new architecture features at the LCFs and similar machines. We also customized these tools according the specific needs of multiple scientific communities, helping them to achieve greatly improved efficiency, adding functionality they requested, and in several cases embedding our tools into their scientific frameworks. The collaboration with the VTK-m effort has been highly successful. Looking forward, as more algorithms are added to VTK-m to take advantage of multi-core and many many-core architectures, ParaView will automatically benefit from these new capabilities.

3.3.2 VTK‐m Framework

3.3.2.1 VTK-m (Sandia, Kitware, ORNL, LLNL, UO)

When SDAV began in 2012, the scientific visualization community was just beginning to develop algorithms and software that run well on multi-core CPUs or many-core accelerators like GPUs. We identified 3 key projects addressing scientific visualization on multi/many-core processors: Piston, an ASC project lead by LANL, Dax, an ASCR project lead by SNL, and EAVL, an LDRD project lead by ORNL. SDAV adopted these products under its umbrella of tools.

Our evaluation of these tools revealed that each addressed a unique aspect of the problem. Piston focused on efficient and portable algorithms. Dax provided a top-down framework that simplified development. EAVL provided advanced data structures. Our initial work considered the integration of these tools. Although we did manage some forms of integration, the results were suboptimal. The integration was not tight enough to realize the full potential and ultimately we were in danger of repeating each other's code.

In response, we considered the feasibility of a tighter integration where the code was redesigned under a single software project. The plan required each software team to transfer their software technology to a new, unified project. Such consolidation of software is rare, but fortunately SDAV was structured to include as key personnel the PIs for each of these three projects. This gave us the motivation and organization to form a tight collaboration in which we all agreed to put aside our previous software and move to a new unified software base. Although much of the development of VTK-m was performed under a following ASCR project, this collaboration could never have formed without SDAV, and we consider this an important success of SDAV.

VTK-m version 1.0 was released June 7, 2016. This release includes the integration of the three predecessor toolkits, and with it we demonstrated an efficient combination of algorithm implementation, performance portability across devices, advanced data models, and high level scientific visualization building blocks [Mor2015b, Mor2016b]. More information about VTK-m is available on our web site (http://m.vtk.org).

During the run of SDAV we had two VTK-m code sprints: one in September 2015 at LLNL and one in August 2016 at Kitware. Among the two we had many participants that represented work from many different organizations including national laboratories (SNL, LANL, ORNL, LBNL, LLNL), universities (Oregon, UC Davis), and industry (Kitware, NVIDIA, Intelligent Light). These events allowed us to reach out to several interested developers to get them kick started with VTK-m development and also allowed us to make progress in several key areas of VTK-m and its algorithms.

VTK-m development continues to progress and is currently the only viable solution for DOE's scientific visualization needs on multi- and many-core devices, and we expect much of future DOE research to contribute back to the VTK-m library. Furthermore, the DOE Exascale Computing Project (ECP) includes the development of several algorithms in VTK-m to replace the single threaded and MPI-only counterparts in VTK.

Next, we describe the progress of the three tools that were integrated into VTK-m: Piston, Dax, and EAVL.

3.3.2.2 Piston (LANL)

Early PISTON Algorithms. Our early work under the SDAV Institute focused on the development of visualization and analysis operators, such as isosurface, threshold, and cut surfaces, using portable, data-parallel "primitives" (e.g. scan, transform, and reduce). These algorithms were described and evaluated in our paper [Lo2012a], and have since been ported to VTK-m. The three projects which merged into VTK-m (PISTON, Dax, and EAVL) were described in the paper [Sew2012a], and their integration was presented in the papers [Chi2013a] and [Mor2016b].

In-Situ Data-Parallelism. As a precursor to the VTK-m / VTK integration, a ParaView plug-in was implemented for PISTON, and prototype in-situ adapters were developed using the Catalyst framework for several scientific codes, including the Vector Particle in Cell (VPIC) plasma simulation code, enabling them to execute our data-parallel visualization algorithms while the simulation is running.

Distributed Data Parallelism. While the VTK-m effort has been focused primarily on improving on-node, shared-memory performance, we also extended several data-parallel operators to run across nodes in distributed memory environments, and implemented algorithms such as isosurface and KD-tree construction using them. This approach was described in a paper [Sew2013a].

Cosmology Applications. Over the course of the SDAV project, we applied the data-parallel principles of PISTON and VTK-m to deliver new capabilities to the Hardware/Hybrid Accelerated Cosmology Code (HACC) in collaboration with the "Computation-Driven Discovery for the Dark Universe" SciDAC Application Partnership in High-Energy Physics (Salman Habib, PI). This work resulted in a number of publications highlighting both the visualization and analysis techniques used and the scientific results they enabled. [Sew2015a] presented the algorithmic details of our data-parallel, portable halo and halo center finding analysis operators. These operators enabled halo analysis to be performed on a half-trillion particle simulation run on 16,384 nodes of Titan. The results of this simulation, which was the first time that the concentration-mass relation has been measured from a single simulation volume over such an extended mass range, were published in [Hei2015a]. We developed innovative new workflows, which combine in-situ and off-line analysis, to deal with the large data generated by this simulation. These workflows were evaluated quantitatively in our Supercomputing paper [Sew2015b].

Topology / Feature Tracking. We have also collaborated with topologists to develop data-parallel algorithms. Our research work involved a general and flexible analysis environment which enables interactive exploration of feature evolution in time-varying data sets, regardless of the underlying data type. In order to demonstrate the generality of this framework, we have extended its applicability to several large-scale scientific and non-scientific data sets. For cosmology data, we implemented a data-parallel, friend-of-friends halo finding algorithm to construct a feature hierarchy for halos, and then tracked their evolution through time using our framework, the results of which was presented in the paper [Wid2014a]. In the case of plasma surface interaction data, we explored and analyzed the evolution of helium bubbles using our framework. We also visualized the evolution of pressure-perturbation events in weather data sets to assist the atmospheric scientists to better understand different weather phenomena. Finally, for healthcare data, we focused on exploring and analyzing patient progression over time utilizing a publically available intensive care unit (ICU) dataset. This work facilitated better predictions of patient outcomes, personalized medication, and more targeted interventions, and was published in the paper

[Wid2016a]. The groundwork in data-parallel topology laid by the SDAV project also led to the development of a data-parallel contour tree algorithm under separate funding, and the resulting paper, [Car2016a], was named Best Paper of the Large Data Analysis and Visualization Symposium in 2016.

3.3.2.3 Dax (Sandia)

*Data Structure Adaptability***.** Much of the early SDAV work with Dax was in making sure its data structures where highly adaptable [Mor2012b]. The initial goal of this work was to improve the interface between Dax and the other multi-threaded libraries (PISTON and EAVL) as well as other established libraries (like VTK). This work was instrumental when combining the technologies from the three projects into a single software product (VTK-m).

Performance Under Layers of Abstraction. In addition to providing an abstract representation of a device (for device portability), Dax used several layers of abstraction to hide the complexities of parallel programming [Mor2012b]. These layers of abstraction worked by identifying common patterns in scientific visualization algorithm execution [Mor2013b]. This required several optimizations in the underlying framework. Of specific importance were geometry creation techniques [May2013a, Mil2014a].

Integration with Dynamic Libraries. The Dax library makes extreme use of C++ templates to maximize the efficiency of the algorithms. However, templates also require all types to be known at compile time, which is a problem when interfacing with a more dynamic software like VTK that uses virtual methods so that specific types do not need to be known until run time. As part of our work in integrating Dax with ParaView, we devised techniques to unroll possible types at run time to call into the appropriate template code.

Scanning Transmission Electron Microscopes (S/TEM). Dax was used to improve the processing in scanning transmission electron microscopes (S/TEM). This equipment generates large (1024^3 or greater) volumes that are often difficult to analyze. Dax was incorporated into TomViz (http://tomviz.org/), an open, general S/TEM visualization tool and provides streaming interactive contouring that significantly beats VTK's alternative algorithms.

3.3.2.4 EAVL (ORNL)

Advanced Data Structures. The EAVL focus was on a next generation data model [Mer2012a, Mer2012b]. Development under SDAV included support for quad-trees and higher order meshes in support of the scientific codes MADNESS and Chimera. The basic data model from EAVL has been ported to VTK-m, and we continue to refine this model.

In Situ Applications. One main use case of the EAVL library was its use as a lightweight visualization library to be used for *in situ* visualization in simulations. To ensure EAVL had minimal impact on simulations, mechanisms for zero-copy access to simulation host arrays were employed. Using zero-copy was often found to be 10x to 200,000x faster than conversion routines. EAVL also provided plotting, rendering, and annotation capabilities while less critical dependencies were trimmed (and compiling a base EAVL library with no dependencies was possible). Most of this functionality is transferred over to VTK-m.

We also experimented with and developed a set of light weight visualization plugins that are suitable for use in data staging environments. Our initial experiments consist of using EAVL and the ADIOS system. These were used to run the XGC-1 code on OLCF resources, staging the data to a set of nodes and then using visualization plugins developed in EAVL to perform basic data reductions and rendering operations of both field and particle data [Pug2014a, Kre2016a]. This work is being ported over to VTK-m [Pug2016c].

Rendering. EAVL received major rendering infrastructure improvements. This included an overhaul to support multiple renderers, in part to remove formerly mandatory rendering dependencies like OpenGL.

The new renderers include vector-based ones like Postscript/EPS, necessary for scientists to generate publication-quality one- and two-dimensional figures. These also include a ray tracing renderer with advanced lighting effects and an accurate direct volume renderer. These could be used with no extra software engineering burden in the form of third-part library dependencies, a critical feature for *in situ* use cases.

Effort was also spent exploring and implementing multi- and many-core accelerated versions of these advanced renderers, as particularly the raycasting and volume rendering ones are computationally intensive, and *in situ* use cases place higher performance demands on the visualization infrastructure. A major thrust of this exploration was to use common data-parallel primitives to accelerate these renderers [Lar2015a, Lar2015b]. These renderers are now available in VTK-m.

3.3.2.5 DIY: Block-Parallel Library (ANL, LBNL)

Over the 5 year SciDAC-3 period, we developed a new version of the DIY programming model which allows building other parallel analysis libraries on top of it. DIY is a programming model and runtime for block-parallel analytics on distributed-memory machines. Its main abstraction is block-structured data parallelism: data are decomposed into blocks; blocks are assigned to processors; computation is described over these blocks, and communication between blocks is defined by reusable patterns. One of the libraries that we built on DIY is Tess, to compute parallel Voronoi and Delaunay tessellations of particle datasets, a core part of the analysis of many simulated and measured datasets: N-body simulations, molecular dynamics codes, and LIDAR point clouds are just a few examples. However, the algorithms for computing these tessellations at scale perform poorly when the input data is unbalanced, such as the case in cosmological dark matter simulations. We investigated the use of k-d trees to evenly distribute points among processes. Because resulting point distributions no longer satisfy the assumptions of our earlier parallel Delaunay algorithm, we developed a new parallel algorithm that adapts to its input.

Our approach of using the k-d tree has led to very significant results. We evaluated the new algorithm using two late-stage cosmology datasets. The new running times are up to 50 times faster using the k-d tree compared with regular grid decomposition. Moreover, in the unbalanced data sets, decomposing the domain into a k-d tree is up to five times faster than decomposing it into a regular grid.

Our work resulted in three major publications and the release of two open-source software libraries. The article about DIY was published in the LDAV symposium collocated with IEEE Visualization 2016 [Mor2016a], and two articles about Tess were published in IEEE/ACM Supercomputing 2014 [Pet2014] and 2016 [Mor2016b]. Both the DIY and Tess software are available on Github.

Conclusion: *in situ* data processing requires scalable distributed-memory algorithms. Developing such algorithms is enabled with DIY's block-parallel programming model. Decomposing the domain into adaptive structures such as k-d trees allows irregular computations to be load-balanced. Libraries such as DIY and Tess are critical pieces of the distributed-memory algorithms in VTK-m, and we will continue to support the parallel communication needs of VTK-m in future research projects.

3.3.3 Flow Visualization Methods

3.3.3.1 Parallel Integral Curve System (UUtah)

Utah working in collaboration with SDAV partners at ORNL and LBNL expanded and hardened VisIt's Parallel Integral Curve System (PICS). The PICS code is an integral part of VisIt's analysis tools allowing researches to analyze and visualize particles paths. As part of SDAV the PICS code was expanded to allow for the secondary analysis of the data produced by the system. The original system architecture was developed where the integral curves were the finial output for either visualization (streamlines or path lines) or analysis (Poincaré maps).

However, with growing interest in Lagrange Coherent Structures (LCS) from DOE researchers in ocean, climate (ice flow), combustion, fusion, and other applications, there was a need to allow for the secondary analysis of the data produced. For instance, a common LCS analysis technique is to compute Finite Time Lyapunov Exponents (FTLE). These exponents are based on the eigenvalues computed using the initial integral curves. However, other structures are of interest, such as limit cycles and shrink and stretch lines which can be found by evaluating the associated eigenvectors. The analysis of the eigenvectors required additional infrastructure with the PICS code.

In addition to the infrastructure development the LCS operator has been expanded. As noted above, in addition to eigenvalues, eigenvectors or their combination are used for limit cycles and shrink and stretch lines all of which are part of LCS analysis. Finding limit cycles is an iterative search process that also utilizes the PICS system which was added as a new operator within VisIt. The Limit Cycle operator is separate from the LCS operator and can be used by any vector field. Finding shrink and stretch lines is based on local min/max FTLE values and was a new addition to the LCS operator. These min/max values are used as seed points along with the eigenvalues and are advected using the previously developed Integral Curve operator. All of the tools described have been deployed within VisIt and are currently being used by DOE and other researchers.

3.3.3.2 Uncertainty Flow Visualization (OSU)

Uncertainty Particle Tracing: We have developed several novel pathline tracing algorithms for time-varying flow data. As the size of simulations continues to increase, it is only possible to store a small subset of the time step data in order to save both storage space and analysis time. Computing pathlines from the reduced data sets, however, is susceptible to interpolation errors due to the lack of data. In this research, we developed a novel data reduction algorithm based on Bezier splines (a numerically stable method to represent curves) to increase the data accuracy, and a parametric error modeling with a least square fitting approach. Our pathline tracing algorithm utilizes the errors modeled by our approach where not only the uncertainty involved is quantified, but the accuracy of the resulting pathlines is also improved. This work was published in IEEE Pacific Visualization 2015 [Che2015a]. To ensure a clearer understanding of three-dimensional flow features, we have also developed a scalable algorithms for parallel stream surface generation [Lu2014a], and a novel domain decomposition algorithm that allows large scale parallel computation of streamlines [Bis2016a]. Our work on uncertainty particle tracing and visualization of uncertain flow features have been used and evaluated by ANL climate scientists Scott M. Collis and Jonathan J. Helmus. The in situ detection of flow features will have a wide applications on DOE’s large scale flow simulations such as detecting magnetic flux vortices in superconductor simulations done at ANL.

In Situ Visualization and Analysis of Flow: We have developed several novel techniques for *in situ* identification of salient flow features. With recent advancements of high performance computing, high-resolution unsteady flow fields allow in depth exploration of turbulence and its possible causes. Performing turbulence analysis, however, involves significant effort to process large amounts of simulation data, especially when investigating abnormalities across many time steps. In order to assist scientists during the exploration process, we created a visual analytics framework to identify suspected spatiotemporal flow abnormity through a comparative visualization so that scientists are able to focus on relevant data in more detail. To achieve this, we developed efficient analysis algorithms derived from domain knowledge and convey the analysis results through juxtaposed interactive plots. Using our integrated visualization system, scientists can visually investigate the detected regions for potential candidate features and further explore their associated spatial regions to enhance the understanding of this phenomenon. Positive feedback from scientists demonstrate the efficacy of our system in analyzing jet engine rotating stall. This work was published in IEEE SciVis 2015 and a special issue of IEEE Transactions on Visualization and Computer Graphics [Che2015b]. Our work also won the best paper honorable mention in IEEE SciVis 2016 [Dut2016a]

3.3.3.3 Flow Visualization on GPUs (LBNL)

Within the context of scaling flow visualization methods, the key component of which is a parallel computation of integral curves, we have focused on a pair of related objectives that aim to produce software methods for flow visualization that are of high quality, in production use by the science community, and that are suitable for use on both multi-core CPU and many-core GPU platforms.

Scaling flow visualization methods is a challenge since the method's parallel performance is dependent upon several factors: the placement/density/number of seed points (an input parameter), the characteristics of the underlying vector field through which curves are computed (input data dependency), and various factors related to the computational architecture (faster vs. slower cores, faster vs. slower memory, the mechanisms and cost of moving data/memory, and so forth).

Our efforts have focused on fostering a better understanding of flow visualization performance on various platforms, and under varying workload conditions [Chi2014b]. For workloads that are heavy in computation, such as high particle density with small integration steps, GPU-based platforms tend to perform better. In contrast, workloads that have lower particle density and larger integration steps, which place a higher load on data movement, tend to perform better on multi-core CPU platforms and on systems with better interconnect fabric. However, CPU-based platforms with high core density and good interconnect fabric proved to be competitive with GPU platforms across diverse workloads.

In the longer term, this type of work can inform how software infrastructure for flow visualization, such as VisIt and VTK-m, can take advantage of different architectures to service varying workloads in a post-processing configuration. Such adaptivity would be the subject of future work.

3.3.4 Rendering

3.3.4.1 Advanced Volume Rendering in VisIt (UUtah)

Over the last five years, we developed a number of techniques to improve the GPU-based volume rendering in VisIt through the SLIVR rendering path. We extended VisIt to include a scalable volume rendering solution that produces similar quality images to the enhanced GPU volume rendering. These improvements are part of the VisIt release. As the programming model of modern supercomputers switches from pure MPI to MPI for inter-node communication, and shared memory and threads for intra-node communication, the bottleneck in most systems is no longer computation but communication between nodes. The Task Overlapped Direct send Tree, TOD-Tree, is a new compositing algorithm for Hybrid/MPI parallelism that minimizes communication and focuses on overlapping communication with computation. We have extended the TOD-Tree compositing method to take advantage of GPU-Direct for GPU-enabled computing resources. TOD-Tree has three stages: a direct send stage where nodes are arranged in groups and exchange regions of an image, followed by a tree compositing stage and a gather stage. We have worked closely with NVIDIA to utilize the direct GPU-MPI remote data copy mechanism, GPU-Direct, to perform compositing completely on the GPU nodes. This avoids copying data between the CPU and the GPU. We published and presented a paper at 2015 EuroGraphics Parallel Graphics and Visualization [Gro2015a].

3.3.4.2 High-End Volume Visualization (UCD)

Volume rendering is a powerful tool for visualizing complex 3D flow field. We aim to enhance the usability of volume visualization by improving both interactivity and rendering quality. To enhance the perception of complex flow structures and features, we have studied different aspects of volume rendering. For example, we have developed more scalable rendering of AMR data, presented at the 2013 LDAV Symposium (receiving the best paper award) [Lea2013a], as well as hybrid grid volume data and geodesic grid data, both presented at the 2014 LDAV Symposium [Shi2014a, Xie2014a]. These have numerous applications to many SciDAC research areas including climate modeling, fusion simulations, and astrophysics simulations. We have also been working on advanced illumination techniques for interactive volume visualization that incorporates the more physical light propagation, absorption, and scattering within the volumetric medium to more accurately present 3D structure and spatial relationships. Papers describing this work was published in ACM SIGGRAPH and IEEE Transactions on Visualization and Computer Graphics [Zha2013b and Zha2013c]. All these advancements benefit demanding 3D visualization tasks commonly found in SciDAC applications.

Unstructured grids are increasingly used by large-scale simulations, and a few rendering algorithms have been developed by us and others for visualizing unstructured-grid volume data. However, while the benefits of using advanced illumination models in volume visualization have been previously demonstrated, interactive rendering has only been achieved for regular-grid volume data. In this SciDAC project, we have developed an advanced illumination method specifically for interactive visualization of 3D unstructured-grid data. The basis of the design is a partial differential equation based illumination model to simulate the light propagation, absorption, and scattering. In particular, a two-level scheme is introduced to overcome the challenges presented by unstructured grids. Our extensive experimental studies show that the added illumination effects such as global shadowing and multiple scattering not only lead to a more visually pleasing visualization, but also greatly enhance the perception of the depth information and complex spatial relationships for features of interest in 3D flow field data. This 3D visualization enhancement is timely as unstructured grids are becoming increasingly popular for a variety of scientific simulation applications in the SciDAC community. A paper reporting this work was presented at PacificVis 2015 [Shi2015a].

In this project period, Min Shih, a PhD student at UCD, over a summer internship successfully incorporated an advanced illumination design into a production visualization system at the Argonne National Laboratory. Future directions will include expanding the performance and quality of these advanced rendering techniques as well as continued implementation into new application areas.

3.3.4.3 Parallel Particle Rendering and Volume Rendering with vl3 (ANL)

vl3 is a parallel framework for large-scale data visualization and analysis developed at Argonne National Laboratory. Its design is modular to facilitate the development and testing of new algorithms and methods, as well as deploying it on varied hardware platforms. vl3 was originally conceived as a parallel hardware-accelerated solution for ray casting volume rendering of regular grids [Riz2014a].

Under the umbrella of SDAV, vl3 evolved in multiple aspects. Firstly, its data reader and rendering modules were extended to support visualization of AMR datasets [Lea2013a]. This work received a Best Paper Award at the 3rd IEEE Symposium on Large Data Analysis and Visualization (LDAV), 2013. Another significant contribution was the addition of hardware accelerated rendering of point sprites for large-scale particle data sets [Riz2014b, Riz2015a]. The new point sprite rendering capabilities, along with the existing large-scale ray casting volume rendering algorithm, allowed the ALCF visualization team to deliver production-quality visualizations for INCITE projects, including “Parameter Studies of Boussinesq Flows”, PI Susan Kurien, and “Cosmological Simulations for Large-Scale Sky Surveys”, PI Salman Habib [Hei2014a, Hei2015a].

In addition, vl3 was successfully coupled to the LAMMPS molecular simulation code for *in situ* visualization of large-scale atomistic data [Riz2015b, Riz2016a]. We also focused on improving image quality of the volume rendering module, developing an algorithm for efficient distributed global illumination based on voxel cone tracing [Shi2016a], which received an Honorable Mention Award at LDAV 2016. Furthermore, we used vl3 as a modular framework to explore novel remote RAM technologies, demonstrating that it is possible to access a large pool of remote memory achieving more than 75% of peak network bandwidth with low latencies [Zaw2016a]. Finally, important improvements were also made to the mechanism for interactive streaming from vl3 running on a visualization cluster to large tiled displays [Jia2015a], which received the Best Poster Award at LDAV 2015.

3.3.5 Ensembles, Uncertainty

3.3.5.1 Ensemble-vis (UUtah)

University of Utah, working with SDAV partners at ORNL and LLNL, developed prototype tools for ensemble analysis and visualization within the VisIt framework. These tools did not move beyond the prototype stage because they required major infrastructure changes within VisIt that were not feasible given the resources. These infrastructure changes would have made ensembles first class data objects with VisIt. That is, an ensemble data set would have been treated exactly like a single data set within VisIt's execution pipeline. Instead, python scripts and tools were developed that automatically handled the integration of ensemble data sets that were then reduced into a single data set that fit within VisIt's execution pipeline.

One of the tools developed allowed for collating ensembles of data and expanded VisIt's cross mesh field evaluations (CMFE) to ensemble data sets. Using the CMFE combined with VisIt's powerful expression mechanism gave users a way to collate data in a variety ways on a per time step basis. The collating of ensemble data is a critical step because the modeling is rarely a linear function of the data, that is sum(f(x)) != f(sum(x)). For the modeling, a new operator, "Statistical Trends" was deployed in VisIt that allows for a user to apply different statistical functions to look for trends in data over time. The operator complemented the abilities of the CMFE operator as the users could collate on either per time step basis or over multiple time steps. Both operators implemented basic statistic functions (min, max, mean, variance, slope, residuals). Though these tools had to be used with python scripting within VisIt they were successfully used to analyze CMIP-5 climate data and are currently being used by DOE and other researchers.

3.3.5.2 Determining Uncertainty in Ensemble Simulations (UCD)

We have worked in conjunction with Lawrence Berkeley National Laboratory climate researchers and Pacific Northwest National Laboratory groundwater researchers to develop techniques to measure the modality of the distributions that arise from ensemble simulations, and have developed methods that can be used to visualize "trends" in such simulations. These methods are based upon a known statistical test – Hardigan's Dip Test – but are modified to garner more information from the field. Furthermore, we have developed a new method for determining the uncertainty in the models by comparing each ensemble to a Bayesian Model Average. By using ground-truth data we can calculate the Bayesian Model Average, and compare the ensembles to this average [Gos2013a]. These methods have the potential to assist in analyzing the copious amounts of data that comes from ensemble simulations and can determine where the simulation outputs produce viable distributions, and where they don't. We have utilized these results to support climate simulations as well as groundwater simulations of uranium seepage in Rifle Colorado. Much of this work was published in IEEE Transactions on Visualization and Computer Graphics [Ben2016a, Obe2016a].

4 Committees

4.1 Software Infrastructure

The role of SDAV's Software Infrastructure Committee (SIC) was to inform and aid in the design of infrastructure that allows different technologies to interoperate. The committee drew on experts from the SDAV Institute to perform integration and interoperability tasks, and they coordinated with SDAV's Executive Council on decisions as needed. The technologies used by SDAV already had existing software engineering practices, which eased the responsibilities of the SIC, since good practices were already being applied.

SIC activities included releases of software, integration of software, and inventories. In terms of software releases, SDAV software saw many releases over the past five years, including dozens of releases of parallel visualization tools (ParaView, VisIt), I/O Frameworks and monitoring tools (ADIOS, PnetCDF, ROMIO, GLEAN, Darshan), indexing and code coupling tools (FastBit, DataSpaces), and data analysis Frameworks and tools (PIDX, ViSUS, MSCEER, TALASS, NDDAV) among other software packages. These software releases included rigorous testing and integration with SciDAC applications and their data.

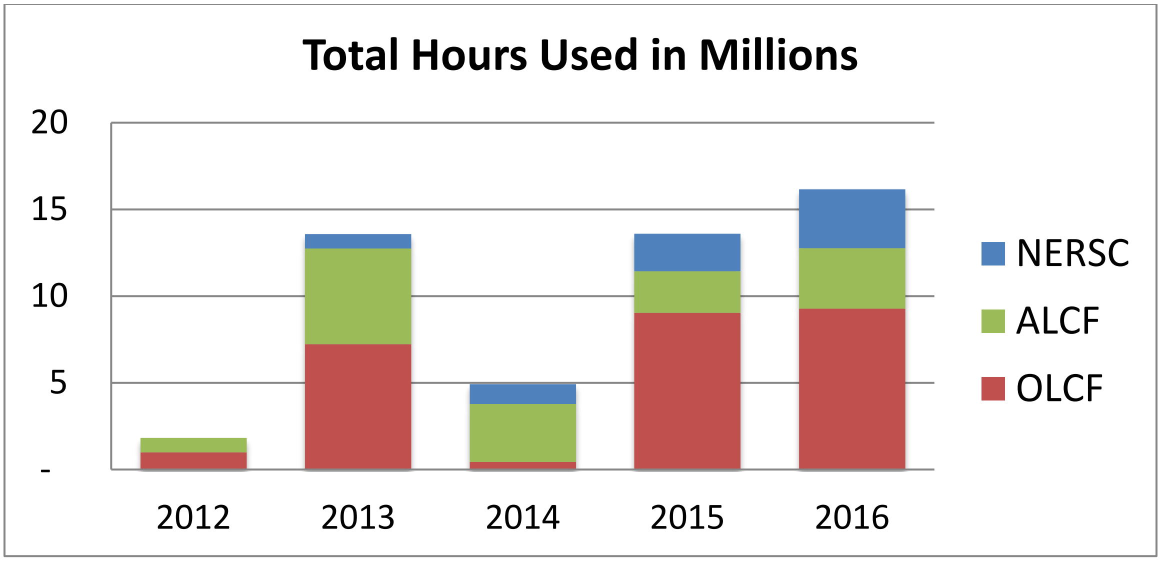
SDAV led the way for integration of many of our software technologies. For example, VisIt and Paraview were both integrated with ADIOS, and researchers in the fusion EPSI project were able to use this to enable their large-scale science. Similarly, ParaView and PISTON were integrated, and this enabled climate applications to understand and visualize their large scale data. Finally, VisIt and EAVL, one of the core components of VTK-m were integrated, and this led into our project creating visualization services that were used in several fusion SciDAC projects.

The SIC took inventory of all SDAV software on behalf of the SDAV EC, so that SDAV stakeholders could better understand the SDAV portfolio, as well as each package's role and maturity. The SIC also prepared materials for DOE's Institute-wide inventory, reporting on information for DBScan, FastBit, FastQuery, DataSpaces, Darshan, Dax, ImageVis3D, VTK, ParaView, Catalyst, DIY, VisIt, PIDX, ViSUS, PISTON, MSCEER, GLEAN, TALASS, NDDAV, HDF5, and UltraVis-P. Our inventory was cataloged on the SDAV web page, each pointing to user guides, download instructions, usage examples, etc., enabling potential customers to become familiar with our software projects.

Many people in SDAV also saw the need for many-core visualization libraries, which are necessary for applications running on the LCFs and NERSC. In the first two years of SDAV, the compatibility and future of DAX, EAVL, and PISTON was an open question. Ultimately, these packages merged into a new package, VTK-m, which is being directly integrated into VisIt, and will be integrated into ParaView via a VTK wrapper. This has been a major outcome from SDAV, and VTK-m has emerged in a manner consistent with SIC goals, including paths for interoperability with other packages, and excellent software engineering practices.

4.2 Facilities Deployment

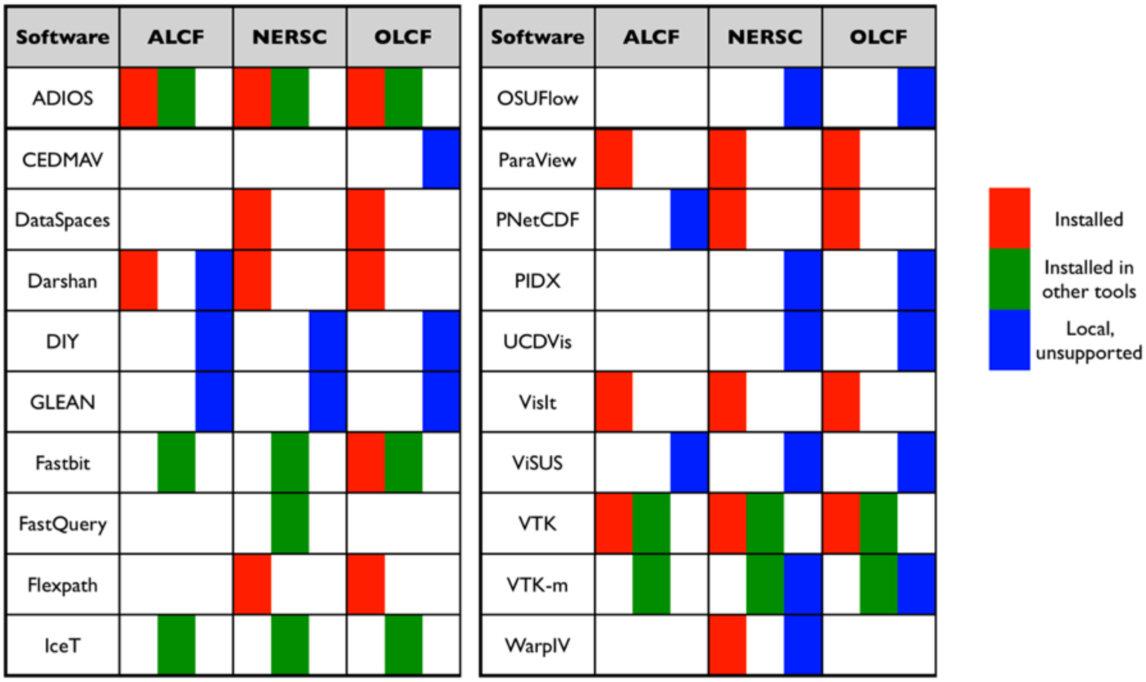
The Facilities Committee has been tasked with providing the SDAV Institute with access to computing cycles at DOE Supercomputing centers and ensuring that SDAV software is installed and available to the user community. To provide SDAV members with access at the computing facilities we have focused on writing proposals to the LCFs at ANL and ORNL, and to NERSC for allocations on each of the systems. Over the 5-year period of SDAV we have secured a total of 41.7 million hours across the three centers. These allocations have been oversubscribed to a total usage of 50.1 million hours. The chart below summarizes the resources used across the three centers over the 5-year period. As can be seen in the chart, the drop in hours used in 2014 was primarily because of a drop in usage at OLCF due to delays encountered in securing a renewal for the allocation awarded in 2013. As shown in subsequent years, the committee worked to increase resource availability for the remainder of the project.



These allocations have provided critical resources for the deployment and testing of SDAV software. These include VisIt, ParaView, Darshan, ADIOS, DataSpaces, DIMES, Flexpath, PIDX, ALACRITY, Fastbit, VTK, VTK-m, and others. In addition, these allocations have been used for algorithm development and scaling studies across a broad range of core SDAV areas, including feature tracking and detection; flow visualization; *in situ* visualization; data indexing and querying. There have been significant interactions with several ASCR and INCITE projects, including EXaCT, EPSI, to name a few.

These allocations have been instrumental in a large number of publications, across a wide range of topics. Selected publications include: flow visualization [Ran2016c, Pug2016b, Che2015a, Lu2015a, Lu2014a, Wan2013a, Yuc2013a, Car2013a, Lin2013a, Cam2012a, Nou2012a], in situ analysis and visualization [Lar2016a, Pug2016c, Ye2015a, Gam2013a, Wan2013a, Ben2012a], queries [Lu2015a, Byn2013a, Lin2013a, Byn2012b], visualization tools [Shi2016a, Riz2016a, Zaw2016a], feature detection and tracking [Per2016a], I/O and workflows [Lof2016a, Kum2014b, Kum2014a, Boy2014a, Car2013a, Kum2013a, Liu2013a, Ben2012a]. A number of publications from this list have been honored as best paper winners and finalists at conferences.

The deployment of SDAV at the LCFs and NERSC is illustrated in the table below. As noted in the legend, software is available in one of three ways. Software that is installed and supported as a stand-alone package are shown in red. Software that is part of other supported tools are shown in green. Finally, software that is locally installed, but unsupported by the center, or is part a private tool or applications are shown in blue. Additionally, information about all SDAV supported software is available on the SDAV website: http://sdav-scidac.org/toolkit.html.



Software	ALCF	NERSC	OLCF
ADIOS	🟥🟩	🟥🟩	🟥🟩
CEDMAV			🟦
DataSpaces		🟥	🟥
Darshan	🟥🟦	🟥	🟥
DIY	🟦	🟦	🟦
GLEAN	🟦	🟦	🟦
Fastbit	🟩	🟩	🟥🟩
FastQuery		🟩	
Flexpath		🟥	🟥
IceT	🟩	🟩	🟩

Software	ALCF	NERSC	OLCF
OSUFlow		🟦	🟦
ParaView	🟥	🟥	🟥
PNetCDF	🟦	🟥	🟥
PIDX		🟦	🟦
UCDVis		🟦	🟦
VisIt	🟥	🟥	🟥
ViSUS	🟦	🟦	🟦
VTK	🟥🟩	🟥🟩	🟥🟩
VTK-m	🟩	🟩🟦	🟩🟦
WarpIV		🟥🟦	

🟥 Installed
🟩 Installed in other tools
🟦 Local, unsupported

4.3 Outreach Activities

Since the SDAV award we have excelled with community outreach and impact of SDAV research and software. We have maintained updated internal and external SDAV websites, which are used for internal communication as well as a public facing site (sdav-scidac.org). The public website consists of general information about the institute, with the majority of the content being driven by research highlights, the SDAV software toolkit, and publications. The SDAV Toolkit (sdav-scidac.org/toolkit.html) is a gateway to useful software and code-based resources that have been developed through SDAV members.

Because of SDAV's high quality research, SDAV PIs were invited to give more than 35 Keynote and Plenary Presentations at international research conferences and workshops including the ACM Symposium on High-Performance Parallel and Distributed Computing, SIAM Annual Conference, Eurographics Symposium on Parallel Graphics and Visualization, the International Parallel and Distributed Processing Symposium, and XSEDE. Additionally, SDAV PIs gave more than 160 invited presentations at conferences and workshops.

SDAV PIs have played leading roles in the organization of more than 45 research conferences and workshops, including the IEEE Symposium on Large Data Analysis and Visualization (LDAV) held in conjunction with the IEEE Visualization Conference, the Ultra-Scale Visualization Workshop, Workshop on Petascale Data Analytics: Challenges and Opportunities and the Parallel Data Storage Workshop held multiple times in conjunction with SuperComputing conferences. All of these meetings brought together computational scientists, data management and visualization researchers and practitioners, and industry and fostered greater exchange between them. In addition, SDAV PIs served on hundreds of conference and workshop program and paper committees.

In addition to publishing 451 papers, SDAV members gave more than 80 tutorials on SDAV software, which included several well attended tutorials on VisIt, ParaView, ADIOS, Darshan, Romio, Piston, DAX, and EAVL. These tutorials were given in well-established conferences such as SuperComputing, and several visualization conferences.

Finally, because of their respect in the research community, SDAV PIs have been chosen to serve on multiple national and international advisory boards.

5 Appendix: Publications over the 5-year period

2017

1. [Lat2017a] Rob Latham, Matthieu Dorier, Robert Ross. **“Get out of the way! Applying compression to internal data structures,”** In *Pdsw-discs 2016: 1st joint international workshop on parallel data storage & data intensive scalable computing systems*, held in conjunction with SC2016, 2017.
2. [Wid2017a] W. Widanagamaachchi, A. Jacques, B. Wang, E. Crosman, P.-T. Bremer, V. Pascucci, J Horel. **“Exploring the Evolution of Pressure-Perturbations to Understand Atmospheric Phenomena,”** In *IEEE Pacific Visualization Symposium (PacificVis)* , 2017.

2016

3. [Bau2016a] Andrew C. Bauer, Hasan Abbasi, James Ahrens, Hank Childs, Berk Geveci, Scott Klasky, Kenneth Moreland, Patrick O'Leary, Venkatram Vishwanath, Brad Whitlock, E. Wes Bethel. **“In Situ Methods, Infrastructures, and Applications on High Performance Computing Platforms,”** In *Computer Graphics Forum*, Vol. 32, No. 3, pp. 577--597. June, 2016.
   DOI: 10.1111/cgf.12930
4. [Ben2016a] Kevin Bensema, Luke J. Gosink, Harald Obermaier, Kenneth I. Joy. **“Modality-Driven Classification and Visualization of Ensemble Variance,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 22, No. 10, 2016.
5. [Bha2016a] Harsh Bhatia, Attila G. Gyulassy, Valerio Pascucci, Martina Bremer, Mitchell T. Ong, Vincenzo Lordi, Erik W. Draeger, John E. Pask, & Peer-Timo Bremer. **“Interactive Exploration of Atomic Trajectories Through Relative-Angle Distribution and Associated Uncertainties,”** In *2016 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 120-127. April, 2016.
6. [Bis2016a] Biswas, Ayan; Strelitz, Richard; Woodring, Jonathan; Chen, Chun-Ming; Shen, Han-Wei. **“A Scalable Streamline Generation Algorithm Via Flux-Based Isocontour Extraction,”** In *Eurographics Symposium on Parallel Graphics and Visualization (EGPGV16)*, june, 2016.
7. [Boz2016a] Ebru Bozdag Daniel Peter Matthieu Lefebvre Dimitri Komatitsch Jeroen Tromp Judith Hill Norbert Podhorszki David Pugmire. **“Global adjoint tomography: first-generation model,”** In *Geophysical Journal International*, Vol. 207, No. 3, Oxford University Press, pp. 1739-1766. Dec, 2016.
   DOI: 10.1093/gji/ggw356
8. [Bre2016a] P.-T. Bremer, A. Gruber, J. Bennett, A. Gyulassy, H. Kolla, J. Chen, R.W. Grout. **“Identifying turbulent structures through topological segmentation,”** In *Com. in App. Math. and Comp. Sci.*, Vol. 11, No. 1, pp. 37-53. 2016.
9. [Bre2016b] P.-T. Bremer. **“ADAPT - Adaptive Thresholds for Feature Extraction,”** In *Topology-Based Methods in Visualization*, Springer, 2016.
10. [Car2016a] Hamish Carr, Gunther Weber, Christopher Sewell, James Ahrens. **“Parallel Peak Pruning for Scalable SMP Contour Tree Computation,”** In *Proceedings of the IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, Baltimore, Maryland, Note: *Best Paper Award. The results reported in this paper stem from the PISTON / VTK-m work established by SDAV; the specific work for this paper was funded under the ASCR XVIS project.*, October, 2016.
11. [Che2016a] Chen, Chun-Ming, Dutta, Soumya, liu, Xiaotong, Heinlein, Gregory, Shen, Han-Wei, Chen, Jen-Ping. **“Visualization and Analysis of Rotating Stall for Transonic Jet Engine**

Simulation," In *IEEE SciVis 2015, also in IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 22, no. 1*, 2016.

12. [Cho2016a] Jong Youl Choi, Tahsin Kurc, Jeremy Logan, Matthew Wolf, Eric Suchyta, James Kress, David Pugmire, Norbert Podhorszki, Eun-Kyu Byun, Mark Ainsworth, Manish Parashar, Scott Klasky. **"Stream processing for near real-time scientific data analysis,"** In *Scientific Data Summit (NYSDS)*, IEEE Xplore, *IEEE*, pp. 1-8. August, 2016.
DOI: 10.1109/NYSDS.2016.7747804
13. [Dev2016a] Dharshi Devendran, Suren Byna, Bin Dong, Brian van Straalen, Hans Johansen, Noel Keen, Nagiza Samatova. **"Collective I/O Optimizations for Adaptive Mesh Refinement Data Writes on Lustre File System,"** In *Cray User Group (CUG)* , May, 2016.
14. [Don2016a] Bin Dong, Suren Byna, Kesheng Wu. **"SDS-Sort: Scalable Dynamic Skew-aware Parallel Sorting,"** In *The ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, July, 2016.
15. [Dut2016a] Dutta, Soumya, Shen, Han-Wei. **"Distribution Driven Extraction and Tracking of Features for Time-varying Data Analysis,"** In *IEEE SciVIS 2015, also in IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 1*, 2016.
16. [Han2016a] Dianwei Han, Ankit Agrawal, Wei-keng Liao, Alok Choudhary. **"A Novel Scalable DBSCAN Algorithm with Spark,"** In *the 5th International Workshop on Parallel and Distributed Computing for Large Scale Machine Learning and Big Data Analytics, held in conjunction with the International Parallel & Distributed Processing Symposium*, Chicago, May, 2016.
17. [Hsu2016a] Chien-Hsin Hsueh, Jacqueline Chu, Kwan-Liu Ma, Joyce Ma, Jennifer Frazier. **"Fostering Comparisons: Designing an Interactive Exhibit that Visualizes Marine Animal Behaviors,"** In *Proceedings of PacificVis 2016 (to appear)*, 2016.
18. [Kan2016a] Qiao Kang, Wei-keng Liao, Ankit Agrawal, Alok Choudhary. **"A Filtering-based Clustering Algorithm for Improving Spatio-temporal Kriging Interpolation Accuracy,"** In *the 25th ACM International Conference on Information and Knowledge Management*, Indianapolis, Indiana, October, 2016.
19. [Kre2016a] James Kress, Randy Michael Churchill, Scott Klasky, Mark Kim, Hank Childs, David Pugmire. **"Preparing for In Situ Processing on Upcoming Leading-edge Supercomputers,"** In *Supercomputing Frontiers and Innovations*, Vol. 3, No. 4, pp. 49-65. 2016.
DOI: 10.14529/jsfi160404
20. [Kre2016b] James Kress David Pugmire Scott Klasky Hank Childs. **"Visualization and analysis requirements for in situ processing for a large-scale fusion simulation code,"** In *ISAV '16 Proceedings of the 2nd Workshop on In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization*, pp. 45-50. 2016.
DOI: 10.1109/ISAV.2016.14
21. [Lar2016a] Larsen, Matthew, Harrison, Cyrus, Kress, James, Pugmire, David, Meredith, Jeremy S., Childs, Hank. **"Performance Modeling of In Situ Rendering,"** In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC16)*, Salt Lake City, Utah, pp. 24:1--24:12. Nov, 2016.
ISBN: 978-1-4673-8815-3
22. [Lar2016b] Matthew Larsen, Kenneth Moreland, Chris R. Johnson, Hank Childs. **"Optimizing Multi-Image Sort-Last Parallel Rendering,"** In *Proceedings of IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, Baltimore, MD pp. 37--46. Oct, 2016.
23. [Lee2016a] Sunwoo Lee, Wei-keng Liao, Ankit Agrawal, Nikos Hardavellas, Alok Choudhary. **"Evaluation of K-Means Data Clustering Algorithm on Intel Xeon Phi,"** In *the Workshop on*

Advances in Software and Hardware for Big Data to Knowledge Discovery, held in conjunction with the IEEE Bigdata Conference, Washington, D.C., December, 2016.

24. [Liu2016a] Liu, Xiaotong, Shen, Han-Wei. **“Association Analysis for Visual Exploration of Multivariate Scientific Data Sets,”** In *IEEE SciVis 2015, also in IEEE Transactions on Visualization and Computer Graphics (TVCG), vol. 22, no. 1*, 2016.
25. [Liu2016b] S. Liu, P.-T. Bremer, J. Thiagarajan, B. Wang, B. Summa, V. Pascucci. **“Grassmannian Atlas: A General Framework for Exploring Linear Projections of High-Dimensional Data. Shusen Liu,”** In *Comput. Graph. Forum*, 2016.
26. [Mal2016a] Dan Maljovec, Bei Wang, Paul Rosen, Andrea Alfonsi, Giovanni Pastore, Cristian Rabiti, Valerio Pascucci. **“Rethinking Sensitivity Analysis of Nuclear Simulations with Topology,”** In *IEEE Pacific Visualization Symposium (PacificVis)*, pp. 64-71. April, 2016.
DOI: 10.1109/PACIFICVIS.2016.7465252
27. [Moo2016a] Changsung Moon, Dakota Medd, Paul Jones, Steve Harenberg, William Oxbury, Nagiza F. Samatova. **“Online Prediction of User Actions through an Ensemble Vote from Vector Representation and Frequency Analysis Models,”** In *SIAM International Conference on Data Mining (SDM)*, May, 2016.
28. [Mor2016a] Kenneth Moreland. **“The Tensions of In Situ Visualization,”** In *IEEE Computer Graphics and Applications*, Vol. 36, No. 2, pp. 5-9. March/April, 2016.
DOI: 10.1109/MCG.2016.35
29. [Mor2016b] Kenneth Moreland, Christopher Sewell, William Usher, Li-ta Lo, Jeremy Meredith, David Pugmire, James Kress, Hendrik Schroots, Kwan-Liu Ma, Hank Childs, Matthew Larsen, Chun-Ming Chen, Robert Maynard, Berk Geveci. **“VTK-m: Accelerating the Visualization Toolkit for Massively Threaded Architectures,”** In *IEEE Computer Graphics and Applications*, Vol. 36, No. 3, pp. 48--58. May/June, 2016.
DOI: 10.1109/MCG.2016.48
30. [Mor2016c] Kenneth Moreland. **“Why We Use Bad Color Maps and What You Can Do About It,”** In *Proceedings of Human Vision and Electronic Imaging (HVEI)*, February, 2016.
DOI: 10.2352/ISSN.2470-1173.2016.16.HVEI-133
31. [Mue2016a] Chris Muelder, Biao Zhu, Wei Chen, Hongxin Zhang, Kwan-Liu Ma. **“Visual Analysis of Cloud Computing Performance Using Behavioral Lines,”** In *Proceedings of PacificVis 2016 (to appear)*, 2016.
32. [Neu2016a] Tyson Neuroth, Franz Sauer, Weixing Wang, Stephane Ethier, Choong-Seock Chang,, Kwan-Liu Ma. **“Scalable Visualization of Time-varying Multi-parameter Distributions Using Spatially Organized Histograms,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. PP, No. 99, 2016.
33. [Obe2016a] Harald Obermaier, Kevin Bensema, Kenneth I. Joy. **“Visual Trends Analysis in Time-Varying Ensembles,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 22, No. 10, 2016.
34. [Pal2016a] Diana Palsetia, William Hendrix, Sunwoo Lee, Ankit Agrawal, Wei-keng Liao, Alok Choudhary. **“Parallel Community Detection Algorithm Using a Data Partitioning Strategy with Pairwise Subdomain Duplication,”** In *the 31st International Supercomputing Conference*, Frankfurt, Germany, June, 2016.
35. [Per2016a] Paris Perdikaris, Joseph A. Insley, Leopold Grinberg, Yue Yu, Michael E. Papka, George Em. Karniadakis. **“Visualizing Multiphysics, Fluid-Structure Interaction Phenomena in Intracranial Aneurysms,”** In *Parallel Computing journal*, Vol. 55, pp. 9-16. July, 2016.
DOI: 10.1016/j.parco.2015.10.016

36. [Pre2016a] Annie Preston, Ramyar Ghods, Jinrong Xie, Franz Sauer, Nick Leaf, Kwan-Liu Ma, Esteban Rangel, Eve Kovacs, Katrin Heitmann, Salman Habib. **“An Integrated Visualization System for Interactive Analysis of Large, Heterogeneous Cosmology Data,”** In *Proceedings of PacificVis 2016 (to appear)*, 2016.
37. [Pug2016a] David Pugmire; James Kress; Hank Childs; Matthew Wolf; Greg Eisenhauer; Randy Churchill; Tahsin Kurc; Jong Choi; Scott Klasky; Kesheng Wu; Alex Sim; Junmin Gu. **“Visualization and Analysis for Near-Real-Time Decision Making in Distributed Workflows,”** In *High Performance Data Analysis and Visualization (HPDAV) 2016 held in conjuction with IPDPS 2016*, Chicago, May, 2016.
38. [Pug2016b] Roberto Sisneros, David Pugmire. **“Tuned to Terrible: A Study of Parallel Particle Advection State of the Practice.,”** In *High Performance Data Analysis and Visualization (HPDAV) 2016 held in conjuction with IPDPS 2016*, Chicago, May, 2016.
39. [Pug2016c] Dave Pugmire, Jeremy Meredith, Scott Klasky, Jong Choi, Norbert Podhorszki, James Kress, Hank Childs. **“Visualization Plugins using VTKm for In-Transit Visualization with ADIOS,”** In *Supercomputing Frontiers 2016*, Singapore, March, 2016.
40. [Ran2016a] Stephen Ranshous, Steve Harenberg, Kshitij Sharma, Nagiza F. Samatova. **“A Scalable Approach for Outlier Detection in Edge Streams Using Sketch-based Approximations,”** In *SIAM International Conference on Data Mining (SDM)*, May, 2016.
41. [Ran2016b] Esteban Rangel, Wei-keng Liao, Ankit Agrawal, Alok Choudhary, William Hendrix. **“AGORAS: A Fast Algorithm for Estimating Medoids in Large Datasets,”** In *the Workshop on Computational Optimization, Modeling & Simulation, held in conjunction with the International Conference on Computational Science*, San Diego, June, 2016.
42. [Ran2016c] Esteban Rangel, Nan Li, Salman Habib, Tom Peterka, Ankit Agrawal, Wei-Keng Liao, Alok Choudhary. **“Parallel DTFE Surface Density Field Reconstruction,”** In *the IEEE International Conference on Cluster Computing*, Taipei, Taiwan, Note: *Best paper award*, September, 2016.
43. [Riz2016a] Silvio Rizzi, Mark Hereld, Joseph A. Insley, Preeti Malakar, Michael E. Papka, Thomas Uram, Venkatram Vishwanath. **“Coupling LAMMPS and the vl3 Framework for Co-Visualization of Atomistic Simulations,”** In *High Performance Data Analysis and Visualization (HPDAV) 2016*, May, 2016.
44. [Rom2016a] Melissa Romanus, Fan Zhang, Tong Jin, Qian Sun, Hoang Bui, Ivan Rodero, Jong Choi, Salomon Janhunen, Robert Hager, Scott Klasky, Choong-Seock Chang, Manish Parashar. **“Persistent Data Staging Services for Data Intensive In-Situ Scientific Workflows,”** In *The 7th International Workshop on Data-intensive Distributed Computing in conjunction with the 25th International ACM Symposium on High Performance Parallel and Distributed Computing(HPDC'16)*, Kyoto, Japan, Note: *To Appear In*, June, 2016.
45. [Rub2016a] Oliver Ruebel, Burlen Loring, Jean-Luc Vay, David P. Grote, Remi Lehe, Stepan Bulanov, Henri Vincenti,, E. Wes Bethel. **“WarpIV: In Situ Visualization and Analysis of Ion Accelerator Simulations,”** In *IEEE Computer Graphics and Applications*, Vol. 36, No. 3, pp. 22-35. may, 2016.
    ISSN: 0272-1716
    DOI: 10.1109/MCG.2016.62
46. [Rud2016a] U. Rüde, K. Willcox, L. C. McInnes, H. De Sterck, G. Biros, H. Bungartz, J. Corones, E. Cramer, J. Crowley, O. Ghattas, M. Gunzburger, M. Hanke, R. Harrison, M. Heroux, J. Hesthaven, P. Jimack, C. Johnson, K. E. Jordan, D. E. Keyes, R. Krause, V. Kumar, S. Mayer, J. Meza, K. M. Mørken, J. T. Oden, L. Petzold, P. Raghavan, S. M. Shontz, A. Trefethen, P. Turner, V. Voevodin, B. Wohlmuth, C. S. Woodward. **“Research and Education in Computational Science and Engineering,”** Subtitled **“Report from a workshop sponsored by the Society for Industrial and**

Applied Mathematics (SIAM) and the European Exascale Software Initiative (EESI-2)," Aug, 2016.

47. [Sau2016a] Franz Sauer, Yubo Zhang, Weixing Wang, Stephane Ethier, Kwan-Liu Ma. **"Visualization Techniques for Studying Large-Scale Flow Fields from Fusion Simulations,"** In *Computer Science and Engineering*, Vol. 18, No. 2, IEEE, pp. 68-77. March, 2016. DOI: 10.1109/MCSE.2015.107
48. [Shi2016a] Min Shih, Silvio Rizzi, Joseph Insley, Thomas Uram, Venkat Vishwanath, Mark Hereld, Michael E. Papka, Kwan-Liu Ma. **"Parallel Distributed, GPU-Accelerated, Advanced Lighting Calculations for Large-Scale Volume Visualization,"** In *IEEE Symposium on Large Data Analysis and Visualization (LDAV) 2016*, Note: *Best Paper Honorable Mention Award*, October, 2016.
49. [Skr2016a] Primoz Skraba, Paul Rosen, Bei Wang, Guoning Chen, Harsh Bhatia, Valerio Pascucci. **"Critical Point Cancellation in 3D Vector Fields: Robustness and Discussion,"** In *IEEE Transactions on Visualization & Computer Graphics. Also Best Paper at PacificVis*, April, 2016.
50. [Sny2016a] Shane Snyder, Philip Carns, Kevin Harms, Robert Ross, Glenn K. Lockwood, Nicholas J. Wright. **"Modular HPC I/O Characterization with Darshan,"** In *Proceedings of 5th Workshop on Extreme-scale Programming Tools (ESPT 2016)*, 11, 2016.
51. [Ste2016a] H. De Sterck, C. Johnson,, L. C. McInnes. **"Special Section on Two Themes: CSE Software and Big Data in CSE,"** In *SIAM J. Sci. Comput*, Vol. 38, No. 5, SIAM, pp. S1--S2. 2016.
52. [Sun2016a] Qian Sun, Melissa Romanus, Tong Jin, Hongfeng Yu, Peer-Timo Bremer, Steve Petruzza, Scott Klasky, Manish Parashar. **"In-Staging Data Placement for Asynchronous Coupling of Task-Based Scientific Workflows,"** In *The 2nd International Workshop on Extreme Scale Programming Models and Middleware(ESPM2'16) in conjunction with The International Conference on High Performance Computing, Networking, Storage and Analysis*, Utah, USA, Note: *Best paper award*, Nov, 2016.
53. [Tan2016a] Houjun Tang, Suren Byna, Steven Harenberg, Xiaocheng Zou, Wenzhao Zhang, Kesheng Wu, Bin Dong, Oliver Rubel, Kristofer Bouchard, Scott Klasky, Nagiza Samatova. **"Usage Pattern-Driven Dynamic Data Layout Reorganization,"** In *16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May, 2016.
54. [Ton2016a] X. Tong, J. Edwards, C. Chen, H. Shen, C. R. Johnson, P. Wong. **"View-Dependent Streamline Deformation and Exploration,"** In *Transactions on Visualization and Computer Graphics*, Vol. 22, No. 7, IEEE, pp. 1788--1801. July, 2016.
55. [Wid2016a] Wathsala Widanagamaachchi, Yarden Livnat, Peer-Timo Bremer, Scott Duvall, Valerio Pascucci. **"Interactive Visualization and Exploration of Patient Progression in a Hospital Setting,"** In *Proceedings of the 2016 Workshop on Visual Analytics in Healthcare*, 2016.
56. [Yua2016a] Zheng Yuan, William Hendrix, Seung Woo Son, Christoph Federrath, An kit Agrawal, Wei-keng Liao, Alok Choudhary. **"Parallel Implementation of Lossy Data Compression for Temporal Data Sets,"** In *the 23rd International Conference on High Performance Computing*, Hyderabad, India, December, 2016.
57. [Zaw2016a] Dawid Zawislak, William Allcock, Joseph Insley, Michael E. Papka, Silvio Rizzi, Brian Toonen. **"Early Investigations Into Using a Remote RAM Pool with the vl3 Visualization Framework,"** In *In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV) 2016*, Salt Lake City, UT, November, 2016.
58. [Zha2016a] Wenzhao Zhang, Houjun Tang, Steven Harenberg, Suren Byna, Xiaocheng Zou, Dharshi Devendran, Daniel Martin, Kesheng Wu, Bin Dong, Scott Klasky, Nagiza Samatova. **"AMRZone: A Runtime AMR Data Sharing Framework For Scientific Applications,"** In *16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May, 2016.

59. [Zha2016b] Yubo Zhang, Kwan-Liu Ma. **"Decoupled Shading for Real-time Heterogeneous Volume Illumination,"** In *Proceedings of EuroVis 2016 (to appear)*, 2016.
60. [Zou2016a] Xiaocheng Zou, David Boyuka, Dhara Desai, Daniel Martin, Suren Byna, Kesheng Wu, Kushal Bansal, Bin Dong, Wenzhao Zhang, Houjun Tang, Dharshi Devendran, David Trebotich, Scott Klasky, Hans Johansen, Nagiza Samatova. **"AMR-aware In Situ Indexing and Scalable Querying,"** In *The 24th High Performance Computing Symposium (HPC)*, April, 2016.

2015

61. [Agr2015a] Alexy Agranovsky, David Camp Kenneth I. Joy,, Hank Chids. **"Subsampling-based Compression and Flow Visualizaiton,"** In *SPIE, Visual Data Analysis Conference, San Francisco CA*, February, 2015.
62. [Arg2015b] Alexy Agranovsky, Harald Obermaier, Christoph Garth, Kenneth I. Joy. **"A Multi-resolution Interpolation Scheme for Pathline-based Lagrangian Flow Representations,"** In *SPIE, Visual Data Analysis Conference, San Francisco CA*, February, 2015.
63. [Aya2015a] Utkarsh Ayachit, Andrew Bauer, Berk Geveci, Patrick O'Leary, Kenneth Moreland, Nathan Fabian, Jeffrey Mauldin. **"ParaView Catalyst: Enabling In Situ Data Analysis and Visualization,"** In *Proceedings of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization (ISAV 2015)*, pp. 25-29. November, 2015.
DOI: 10.1145/2828612.2828624
64. [Beh2015a] Babak Behzad, Suren Byna, Stefan Wild, Prabhat, Marc Snir. **"Dynamic Model-driven Parallel I/O Performance Tuning,"** In *IEEE Cluster 2015*, Chicago, https://sdm.lbl.gov/~sbyna/research/papers/201509-Cluster-Autotune.pdf, September, 2015.
65. [Ben2015a] J. Bennett, F. Vivodtzev,, V. Pascucci. **"Topological and statistical methods for complex data. Mathematics and Visualization,"** In *Topological and statistical methods for complex data. Mathematics and Visualization*, Springer, May, 2015.
66. [Bet2015a] E. Wes Bethel, David Camp, David Donofrio, Mark Howison. **"Improving Performance of Structured-memory, Data-Intensive Applications on Multi-core Platforms via a Space-Filling Curve Memory Layout,"** In *International Workshop on High Performance Data Intensive Computing, and IEEE International Parallel and Distributed Processing Symposium (IPDPS) workshop*, Hyderabad, India May, 2015.
67. [Bha2015a] Harsh Bhatia, Bei Wang, Gregory Norgard, Valerio Pascucci, Peer-Timo Bremer. **"Local, Smooth, and Consistent Jacobi Set Simplification,"** In *Computational Geometry: Theory and Applications (CGTA), 48(4)*, pp. 311-332. May, 2015.
68. [Bis2015a] Ayan Biswas, Wenbin He, Qi Deng, Chun-Ming Chen, Han-Wei Shen, Raghu Machiraju, Anand Rangarajan. **"An Uncertainty-Driven Approach to Vortex Analysis Using Oracle Consensus and Spatial Proximity,"** In *IEEE Pacific Vis, Hangzhou, China*, April, 2015.
69. [Boy2015a] Drew A. Boyuka, Xiaocheng Zou, Nagiza Samatova, Junmin Gu, Kesheng Wu, Norbert Podhorszki, Scott Klasky. **"ADIOS Query Interface Design ,"** In *Supercomputing Frontiers*, June, 2015.
70. [Boy2015b] David A. Boyuka II, Houjun Tang, Kushal Bansal, Xiaocheng Zou, Scott Klasky, Nagiza F. Samatova. **"The Hyperdyadic Index and Generalized Indexing and Query with PIQUE,"** In *International Conference on Scientific And Statistical Database Management (SSDBM)*, June, 2015.
71. [Bre2015a] Peer-Timo Bremer, Dan Maljovec, Avishek Saha, Bei Wang, Jim Gaffney, Brian K. Spears, Valerio Pascucci. **"ND2AV: N-Dimensional Data Analysis and Visualization - Analysis for the National Ignition Campaign,"** In *Computing and Visualization in Science*, February, 2015.
72. [Bui2015a] Huy Bui, Robert Jacob, Preeti Malakar, Venkatram Vishwanath , Andrew Johnson, Michael Papka,, Jason Leigh. **"Multipath Load Balancing for M × N Communication Patterns on**

the Blue Gene/Q Supercomputer Interconnection Network," In *1st IEEE International Workshop on High-Performance Interconnection Networks Towards the Exascale and Big-Data Era, co-located with IEEE Cluster*, Chicago, IL, USA, September, 2015.

73. [Buj2015a] Roxana Bujack, Jens Kasten, Vijay Natarajan, Gerik Scheuermann, Kenneth I. Joy. **"Clustering Moment Invariants to Identify Similarity within 2D Flow Fields,"** In *Eurographics Conference on Visualization (EuroVis) - short paper*, May, 2015.
74. [Buj2015b] Roxana Bujack, Kenneth Joy. **"Lagrangian Representations of Flow Fields with Parameter Curves,"** In *IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, October, 2015.
75. [Bus2015a] Michael Bussmann, Axel Huebl, René Widera, Felix Schmitt, Sebastian Grottel, Norbert Podhorszki, Dave Pugmire, Scott Klasky. **"Breaking the Simulation/Analysis Chain,"** In *Supercomputing Frontiers*, March, 2015.
76. [Byn2015a] Suren Byna, Robert Sisneros, Kalyana Chadalavada, Quincey Koziol. **"Tuning Parallel I/O on Blue Waters for Writing 10 Trillion Particles,"** In *Cray User Group (CUG) meeting*, 2015.
77. [Car2015a] Hamish Carr, Zhao Geng, Julien Tierny, Amit Chattopadhyay, Aaron Knoll. **"Fiber Surfaces: Generalizing Isosurfaces to Bivariate Data,"** In *Computer Graphics Forum, proc. Eurovis*, May, 2015.
78. [Cha2015a] Abon Chaudhuri, Teng-Yok Lee, Han-Wei Shen, Rephael Wenger. **"Exploring Flow Fields Using Space-filling Analysis of Streamlines,"** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, January, 2015.
79. [Cha2015b] Jennifer Chandler, Harald Obermaier, Kenneth I. Joy. **"Interpolation-Based Pathline Tracing in Particle-Based Flow Visualization,"** In *IEEE Transactions on Visualization and Computer Graphics, Vol 21, No. 1*, January, 2015.
80. [Cha2015c] Jennifer Chandler, Harald Obermaier,, Kenneth I. Joy. **"WebGL-Enabled Remote Visualization of Smoothed Particle Hydrodynamics Simulations,"** In *Eurographics Conference on Visualization (EuroVis) - short paper*, May, 2015.
81. [Che2015a] Chun-Ming Chen, Ayan Biswas, Han-Wei Shen. **"Uncertainty Modeling and Error Reduction for Pathline Computation in Time-varying Flow Fields,"** In *IEEE Pacific Vis 2015, Hangzhou, China*, April, 2015.
82. [Che2015b] Chun-Ming Chen, Soumya Dutta, Xiaotong Liu, Gregory Heinlein, Han-Wei Shen, Jen-Ping Chen. **"Visualization and Analysis of Rotating Stall for Transonic Jet Engine Simulation,"** In *IEEE Scientific Visualization (SciVis)*, October, 2015.
83. [Cho2015a] Jong Choi, Yuan Tian, Gary Liu, Norbert Podhorszki, David Pugmire, Scott Klasky, Eun-Kyu Byun, Soonwook Hwang, Alex Sim, Lingfei Wu, John Wu, Mehmet Aktas, Manish Parashar, Michael Churchill, C.S. Chang, Tahsin Kurc, Xinyan Yan, Matthew Wolf,. **"ICEE: Enabling Data Stream Processing For Remote Data Analysis Over Wide Area Networks,"** In *Supercomputing Frontiers*, March, 2015.
84. [Day2015a] Jai Dayal, Jay Lofstead, Greg Eisenhauer, Karsten Schwan, Matthew Wolf,Hasan Abbasi, Scott Klasky. **"SODA: Science-driven Orchestration of Data Analytics,"** In *The 11th International Conference on eScience*, September, 2015.
85. [Dee2015a] Ewa Deelman, Tom Peterka, others. **"The Future of Scientific Workflows: Report of the DOE NGNS/CS Scientific Workflows Workshop,"** April, 2015.
86. [Don2015a] Bin Dong, Suren Byna, Kesheng Wu. **"Spatially Clustered Join on Heterogeneous Scientific Data Sets,"** In *2015 IEEE International Conference on Big Data (IEEE BigData)*, 2015.

87. [Dut2015a] Soumya Dutta, Han-Wei Shen. **“Distribution Driven Extraction and Tracking of Features for Time-varying Data Analysis,”** In *IEEE Scientific Visualization (SciVis)*, October, 2015.
88. [Edw2015a] J. Edwards, E. Daniel, V. Pascucci,, C. Bajaj. **“The Generalized Voronoi Diagram of Closely-Spaced Objects,”** In *Computer Graphics Forum*, June, 2015.
89. [Gon2015a] Gonzalo A. Bello, Michael Angus, Navya Pedemane, Jitendra K. Harlalka, Fredrick H. M. Semazzi, Vipin Kumar, Nagiza F. Samatova. **“Response-Guided Community Detection: Application to Climate Index Discovery,”** In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD*, September, 2015.
90. [Gra2015a] Lucio Grandinetti, Gerhard Joubert, Marcel Kunze, Valerio Pascucci. **“Big Data and High Performance Computing,”** In *Big Data and High Performance Computing*, IOS Press, October, 2015.
91. [Gro2015a] Pascal Grosset, Manasa Prasad, Cameron Christensen, Aaron Knoll, Charles D. Hansen. **“TOD-Tree: Task-Overlapped Direct send Tree Image Compositing for Hybrid MPI Parallelism,”** In *Eurographics Parallel Graphics and Visualization (EGPGV)*, May, 2015.
92. [Gyu2015a] Attila Gyulassy, Aaron Knoll, Kah Chun Lau, Bei Wang, Peer-Timo Bremer, Valerio Pasucci, Michael E. Papka, Larry Curtiss. **“Morse-Smale Analysis of Ion Diffusion in Ab Initio Battery Materials Simulations,”** In *Topology-Based Methods in Visualization (TopoInVis)*, June, 2015.
93. [Gyu2015b] Gyulassy, A.; Knoll, A.; Lau, K.; Wang, B.; Bremer, P.; Papka, M.; Curtiss, L.; Pascucci, V.. **“Interstitial and Interlayer Ion Diffusion Geometry Extraction in Graphitic Nanosphere Battery Materials,”** In *IEEE Transactions on Visualization and Computer Graphics*, August, 2015.
    DOI: 10.1109/TVCG.2015.2467432
94. [Hab2015a] Salman Habib, Adrian Pope, Hal Finkel, Nicholas Frontiere, Katrin Heitmann, David Daniel, Patricia Fasel, Vitali Morozov, George Zagaris, Tom Peterka, Venkatram Vishwanath, Zarija Lukic, Saba Sehrish, Wei-keng Liao. **“HACC: Simulating Sky Surveys on State-of-the-Art Supercomputing Architectures,”** In *New Astronomy*, Vol. 42, pp. 49-65. July, 2015.
    DOI: 10.1016/j.newast.2015.06.003
95. [Hei2015a] Katrin Heitmann, Nicholas Frontiere, Chris Sewell, Salman Habib, Adrian Pope, Hal Finkel, Silvio Rizzi, Joe Insley, Suman Bhattacharya. **“The Q Continuum Simulation: Harnessing the Power of GPU Accelerated Supercomputers,”** In *Astrophysical Journal Supplement Series*, Vol. 219, No. 34, August, 2015.
96. [Hua2015a] Dan Huang, Jiangling Yin, Jun Wang, Xuhong Zhang, Jian Zhou, Qing Liu. **“SideIO: A Side I/O Framework System for Eliminating Analysis Data Migration,”** In *Supercomputing Frontiers*, March, 2015.
97. [Jia2015a] Jie Jiang, Mark Hereld, Joseph A. Insley, Michael E. Papka, Silvio Rizzi, Thomas Uram, Venkatram Vishwanath. **“Streaming Ultra High Resolution Images to Large Tiled Display at Nearly Interactive Frame Rates with vl3,”** In *IEEE Symposium on Large Data Analysis and Visualization (LDAV) - poster*, Note: *Best Poster Award*, October, 2015.
98. [Jin2015a] Ye Jin, Xiaosong Ma, Gary Liu, Mingliang Liu, Jeremy Logan, Norbert Podhorszki, Jong Youl Choi,, Scott Klasky. **“Combining Phase Identification and Statistical Modeling for Automated Parallel Benchmark Generation,”** In *ACM SIGMETRICS*, June, 2015.
99. [Jin2015b] Ye Jin, Mingliang Liu, Xiaosong Ma, Gary Liu, Jeremy S. Logan, Norbert Podhorszki, Jong Youl Choi,, Scott Klasky. **“Combining Phase Identification and Statistic Modeling for Automated Parallel Benchmark Generation,”** In *PPoPP, accepted as a poster*, February, 2015.

100. [Jin2015c] Tong Jin, Fan Zhang, Qian Sun, Melissa Romanus, , Norbert Podhorszki, Scott Klasky, Hemanth Kolla, Jacqueline Chen, , Robert Hager, Choong-Seock Chang, Manish Parashar . **“Exploring Data Staging Across Deep Memory Hierarchies for Coupled Data Intensive Simulation Workflows,”** In *29th IEEE International Parallel & Distributed Processing Symposium*, May, 2015.

101. [Joh2015a] C.R. Johnson, K. Potter. **“Visualization,”** In *The Princeton Companion to Applied Mathematics*, Edited by Nicholas J. Higham, Princeton University Press, pp. 843-846. September, 2015.
ISBN: 9780691150390

102. [Joh2015d] C.R. Johnson. **“Visualization,”** In *Encyclopedia of Applied and Computational Mathematics*, Edited by Björn Engquist, Springer, pp. 1537-1546. 2015.
ISBN: 978-3-540-70528-4

103. [Kim2015a] Mark Kim, Charles Hansen. **“Surface Flow Visualization using the Closest Point Embedding,”** In *IEEE Pacific Visualization*, April, 2015.

104. [Kim2015b] Mark Kim, Charles Hansen. **“GPU Surface Extraction using the Closest Point Embedding,”** In *SPIE Visualization and Data Analysis*, February, 2015.

105. [Kim2015c] SeongJo Kim, Yuanrui Zhang, SeungWoo Son, Mahmut Kandemir, Wei-keng Liao, Rajeev Thakur, Alok Choudhary. **“IOPro: a parallel I/O profiling and visualization framework for high-performance storage systems,”** In *Supercomputing*, Vol. 71, No. 3, Springer US, pp. 840-870. March, 2015.

106. [Koo2015a] M. Koo, W. Yoo, A. Sim . **“I/O Performance Analysis Framework on Measurement Data from Scientific Clusters,”** In *International Conference for High Performance Computing, Networking, Storage and Analysis (SC'15), ACM Student Research Competition (SRC)*, November, 2015.

107. [Kre2015a] James Kress, Scott Klasky, Norbert Podhorszki, Jong Choi, Hank Childs,, Dave Pugmire. **“Loosely Coupled In Situ Visualization: A Perspective on Why it's Here to Stay,”** In *In Situ infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV-15), held in conjunction with SC15*, November, 2015.

108. [Lan2015a] A. Landge, P.-T. Bremer, A. Gyulassy, V. Pascucci. **“Notes on the Distributed Computation of Merge Trees on CW-complexes,”** In *Proc. TopoInVis*, May, 2015.

109. [Lar2015a] Matthew Larsen, Jeremy Meredith, Paul Navratil, Hank Childs. **“Ray-Tracing Within a Data Parallel Framework,”** In *IEEE Pacific Visualization (PacificVis)*, April, 2015.

110. [Lar2015b] M. Larsen, S. Labasan, P. Navrátil, J.S. Meredith,, H. Childs. **“Volume Rendering Via Data-Parallel Primitives,”** In *Eurographics Symposium on Parallel Graphics and Visualization*, May, 2015.

111. [Lar2015c] Matthew Larsen, Eric Brugger, Hank Childs, Jim Eliot, Kevin Griffin,, Cyrus Harrison. **“Strawman - A Batch In Situ Visualization and Analysis Infrastructure for Multi-Physics Simulation Codes,”** In *In Situ infrastructures for Enabling Extreme-scale Analysis and Visualization (ISAV-15), held in conjunction with SC15*, November, 2015.

112. [Li2015a] Kelvin Li, Jia-Kai Chou, Kwan-Liu Ma. **“High Performance Heterogeneous Computing for Collaborative Visual Analysis,”** In *Proceedings of 2015 Symposium on Visualization in High Performance Computing, co-located with ACM SIGGRAPH Asia 2015* , ACM, pp. 1-4. November, 2015.

113. [Lia2015a] Yaxiong Liang, Xu Ji, Hoang Bui, Fan Zhang, Jeremy Logan, Wei Xue, Lizhe Wang, Manish Parashar, Scott Klasky,Weimin Zheng. **“TCP Based Data Staging on Supercomputers,”** In *Supercomputing Frontiers*, March, 2015.

114. [Lil2015a] O. Anatole von Lilienfeld, Raghunathan Ramakrishanan, Matthias Rupp,, Aaron Knoll. **“Fourier Series of Atomic Radial Distribution Functions: A Molecular Fingerprint for Machine Learning Models of Quantum Chemical Properties,”** In *International Journal of Quantum Chemistry*, August, 2015.

115. [Lis2015a] Shaomeng Li, Kenny Gruchalla, Kristin Potter, John Clyne, Hank Childs. **“Evaluating the Efficacy of Wavelet Configurations on Turbulent-Flow Data,”** In *Proceedings of IEEE Symposium on Large Data Analysis and Visualization*, October, 2015.

116. [Liu2015a] Xiaotong Liu, Han-Wei Shen. **“The Effects of Representation and Juxtaposition on Graphical Perception of Matrix Visualization,”** In *ACM Computer-Human Interaction (CHI'2015)*, April, 2015.

117. [Liu2015b] Shusen Liu, Bei Wang, Jayaraman J. Thiagarajan, Peer-Timo Bremer, Valerio Pascucci. **“Visual Exploration of High-Dimensional Data through Subspace Analysis and Dynamic Projections,”** In *Eurographics Conference on Visualization (EuroVis)*, May, 2015.

118. [Liu2015c] Ruoqian Liu, Ankit Agrawal, Wei-keng Liao, Alok Choudhary, Zhengzhang Chen. **“Pruned Search: A Machine Learning Based Meta-Heuristic Approach for Const rained Continuous Optimization,”** In *the Eighth International Conference on Contemporary Computing*, August, 2015.

119. [Liu2015d] Liu, Xiaotong, Shen, Han-Wei, Hu, Yifan. **“Supporting multifaceted viewing of word clouds with focus plus context display,”** In *Information Visualization Journal, Vol. 14, no. 2*, pp. 168-180. April, 2015.

120. [Liu2015e] Xiaotong Liu, Han-Wei Shen. **“Association Analysis for Visual Exploration of Multivariate Scientific Data Sets,”** In *IEEE Scientific Visualization (SciVis)*, October, 2015.

121. [Liu2015f] Xiaotong Liu, Yifan Hu, Stephen North,, Han-Wei Shen. **“CorrelatedMultiples: Spatially Coherent Small Multiples with Constrained Multidimensional Scaling,”** In *Computer Graphics Forum (CGF)*, January, 2015.

122. [Liu2015g] Xiaotong Liu, Srinivasan Parthasarathy, Han-Wei Shen,, Yifan Hu. **“GalaxyExplorer: Influence-Driven Visual Exploration of Context-Specific Social Media Interactions,”** In *International World Wide Web Conference (WWW)*, May, 2015.

123. [Liu2015h] Shusen Liu, Dan Maljovec, Bei Wang, Peer-Timo Bremer, Valerio Pascucci. **“Visualizing High-Dimensional Data: Advances in the Past Decade,”** In *Proceedings of Eurographics Conference on Visualization (EuroVis), STAR -- State of The Art Report*, May, 2015.

124. [Log2015a] Jeremy Logan, Scott Klasky, Norbert Podhorszki, Lizhe Wang, Wei Xue. **“Creating Skeletons for Task-Based Scientific Workflows,”** In *Supercomputing Frontiers*, March, 2015.

125. [Lor2015a] B. Loring, H. Karimabadi,, V. Rortershteyn. **“A Screen Space GPGPU Surface LIC Algorithm for Distributed Memory Data Parallel Sort Last Rendering Infrastructures,”** In *Proceedings of the 9th International Conference on Numerical Modeling of Space Plasma Flows (ASTRONUM-2014)*, Long Beach, CA, USA March, 2015.

126. [Lu2015a] Kewei Lu, Han-Wei Shen. **“A compact multivariate histogram representation for query-driven visualization,”** In *IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, October, 2015.

127. [Luo2015a] Xaioqing Luo, Frank Mueller, Philip Carns, John Jenkins, Robert Ross, Shane Snyder, Robert Latham. **“ScalaIOExtrap: Elastic I/O Tracing and Extrapolation,”** In *Proceedings of the Workshop on Extreme-Scale Programming Tools (ESPT 2015)*, November, 2015.

128. [Luu2015a] Huong Luu, Marianne Winslett, William Gropp, Kevin Harms, Philip Carns, Robert Ross, Yushu Yao, Suren Byna,, Prabhat. **“A Multiplatform Study of I/O Behavior on Petascale**

Supercomputers," In *Proceedings of the 24th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC 2015)*, ACM, June, 2015.

129. [Mal2015a] P. Malakar, V. Vishwanath, T. Munson, C. Knight, M. Hereld, S. Leyffer, M. Papka. **"Optimal Scheduling of In Situ Analysis for Large-Scale Scientific Simulations,"** In *Proceedings of the 28th IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2015)*, Austin, Texas, USA, November, 2015.

130. [Mal2015b] P. Malakar, V. Vishwanath. **"Hierarchical Read-write Optimizations for Scientific Applications with Multi-variable Structured Datasets,"** In *Proceedings of the 12th Annual IFIP International Conference on Network and Parallel Computing (NPC)*, New York City, New York, USA, September, 2015.

131. [Mal2015c] D. Maljovec, S. Liu, B. Wang, V. Pascucci, P.-T. Bremer, D. Mandelli, C. Smith. **"Analyzing simulation-based PRA data through clustering: a BWR station blackout case study,"** In *Reliability Engineering & System Safety*, Note: *In Press, submitted*, June, 2015.

132. [Mor2015a] Kenneth Moreland, Ron Oldfield. **"Formal Metrics for Large-Scale Parallel Performance,"** In *High Performance Computing*, July, 2015.
DOI: 10.1007/978-3-319-20119-1_34

133. [Mor2015b] Kenneth Moreland, Matthew Larsen, Hank Childs. **"Visualization for Exascale: Portable Performance is Critical,"** In *Supercomputing Frontiers and Innovations*, Vol. 2, No. 3, 2015.
DOI: 10.14529/jsfi150306

134. [Neu2015a] Tyson Neuroth, Franz Sauer, Weixing Wang, Stephane Ethier, Kwan-Liu Ma. **"Scalable Visualization of Discrete Velocity Decompositions using Spatially Organized Histograms,"** In *Proceedings of LDAV 2015*, IEEE, pp. 65-72. October, 2015.
DOI: 10.1109/LDAV.2015.7348073

135. [Obe2015a] Harald Obermaier, Kenneth I. Joy. **"An Automated Approach for Slicing Plane Placement in Visual Data Analysis,"** In *IEEE Transactions on Visualization and Computer Graphics*, May, 2015.

136. [Phi2015a] S. Philip, B. Summa, J. Tierny, P. Bremer, V. Pascucci. . **"Distributed seams for gigapixel panoramas,"** In *IEEE Transactions on Visualization and Computer Graphics,*, Vol. 21, No. 3, pp. 350–362. March, 2015.

137. [Pra2015a] Prabhat, S.Byna, V. Vishwanath, E. Dart, M. Wehner, W. Collins. **"TECA: Petscale Pattern Recognition for Climate Science,"** In *Proceedings of the 16th International Conference on Computer Analysis of Images and Patterns (CAIP)*, Valletta, Malta, September, 2015.

138. [Pug2015a] James Kress, Scott Klasky, Norbert Podhorszki, Jong Choi, Hank Childs, David Pugmire. **"Loosely Coupled In Situ Visualization: A Perspective on Why It's Here to Stay,"** In *Proceedings of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization*, Nov, 2015.

139. [Ran2015a] Stephen Ranshous, Shitian Shen, Danai Koutra, Steve Harenberg, Christos Faloutsos, Nagiza F. Samatova. **"Anomaly Detection in Dynamic Networks: A Survey,"** In *Wiley Interdisciplinary Reviews: Computational Statistics*, June, 2015.

140. [Riz2015a] Silvio Rizzi, Mark Hereld, Joseph A. Insley, Michael E. Papka, Thomas Uram, Venkatram Vishwanath. **"Large-Scale Parallel Visualization of Particle-Based Simulations using Point Sprites and Level-Of-Detail,"** In *Eurographics Parallel Graphics and Visualization (EGPGV)*, May, 2015.

141. [Riz2015b] Silvio Rizzi, Mark Hereld, Joseph A. Insley, Michael E. Papka, Thomas Uram, Venkatram Vishwanath. **"Large-Scale Co-Visualization for LAMMPS using vl3,"** In *IEEE Symposium on Large Data Analysis and Visualization (LDAV) - poster*, October, 2015.

142. [Sch2015a] William Schroeder, Robert Maynard, Berk Geveci. **“Flying Edges: A High-Performance Scalable Isocontouring Algorithm,”** In *To appear in the Proceedings of the IEEE Symposium on Large Data Analysis and Visualization (LDAV),* Chicago, Illinois, October, 2015.

143. [Sch2015b] Hendrik A. Schroots, Kwan-Liu Ma. **“Volume Rendering with Data Parallel Visualization Frameworks for Emerging High Performance Computing Architectures,”** In *Proceedings of 2015 Symposium on Visualization in High Performance Computing, co-located with ACM SIGGRAPH Asia 2015,* ACM, pp. 1-4. November, 2015.

144. [Sew2015a] Christopher Sewell, Li-ta Lo, Katrin Heitmann, Salman Habib,, James Ahrens. **“Utilizing Many-Core Accelerators for Halo and Center Finding within a Cosmology Simulation,”** In *Proceedings of the IEEE Symposium on Large Data Analysis and Visualization (LDAV),* Chicago, Illinois, October, 2015.

145. [Sew2015b] Christopher Sewell, Katrin Heitmann, Hal Finkel, George Zagaris, Suzanne T. Parete-Koon, Patricia K. Fasel, Adrian Pope, Nicholas Frontiere, Li-ta Lo, Bronson Messer, Salman Habib,, James Ahrens. **“Large-Scale Compute-Intensive Analysis via a Combined In-situ and Co-scheduling Workflow Approach,”** In *Proceedings of The International Conference for High Performance Computing, Networking, Storage, and Analysis*, Austin, Texas, November, 2015.

146. [Sha2015a] S. Shannigrahi, A. J. Barczyk, C. Papadopoulos, A. Sim, I. Monga, H. Newman, K. Wu, E. Yeh. **“Named Data Networking in Climate Research and HEP Applications,”** In *21st International Conference on Computing in High Energy and Nuclear Physics (CHEP),* 2015.

147. [Shi2015a] Min Shih, Yubo Zhang, Kwan-Liu Ma. **“Advanced Lighting for Unstructured-Grid Data Visualization,”** In *PacificVis*, April, 2015.

148. [Sny2015a] Shane Snyder, Philip Carns, Robert Latham, Misbah Mubarak, Rob Ross, Christopher Carothers, Babak Behzad, Huong Vu Thanh Luu, Surendra Byna, Prabhat. **“Techniques for Modeling Large-Scale HPC I/O Workloads,”** In *Proceedings of the 6th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS15),* November, 2015.

149. [Sum2015a] B. Summa, A. A. Gooch, G. Scorzelli, V. Pascucci. **“Paint and Click: Unified Interactions for Image Boundaries,”** In *Computer Graphics Forum*, May, 2015.

150. [Sun2015a] Q. Sun, T. Jin, M. Romanus, H. Bui, F. Zhang, H. Yu, H. Kolla, S. Klasky, J. Chen, M. Parashar. **“Adaptive Data Placement For Staging-Based Coupled Scientific Workflows,”** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis*, Austin, USA, November, 2015.

151. [Tan2015a] Yuzuru Tanahashi, Chien-Hsin Hsueh, Kwan-Liu Ma. **“An Efficient Framework for Generating Storyline Visualizations from Streaming Data,”** In *IEEE Transactions on Visualization and Computer Graphics*, June, 2015.

152. [Ton2015a] Xin Tong, Chun-Ming Chen, Han-Wei Shen, Pak Chung Wong. **“Interactive Streamline Exploration and Manipulation Using Deformation,”** In *IEEE Pacific Vis 2015, Hangzhou, China*, April, 2015.

153. [Wid2015a] Wathsala Widanagamaachchi, Karl D. Hammond, Li-Ta Lo, Brian D. Wirth, Francesca Samsel, Christopher Sewell, James Ahrens,, Valerio Pascucci. **“Visualization of Large-Scale Atomistic Simulations of Plasma-Surface Interactions,”** In *Proceedings of EuroVis (short paper),* Cagliari, Italy, Note: *The results reported in this paper stem from the PISTON / VTK-m work established by SDAV; the specific work for this paper was funded under the SciDAC Plasma Surface Interactions project*, May, 2015.

154. [Wid2015b] W. Widanagamaachchi, P. Klacansky, H. Kolla, J. Chen, A. Bhagatwala, V. Pascucci, P.-T. Bremer. **“Tracking Features in Embedded Surfaces: Understanding Extinction in**

Turbulent Combustion," In *Proc. IEEE Symposium Large-Scale Data Analysis and Visualization*, October, 2015.

155. [Wu2015a] L. Wu, K. Wu, A. Sim, M. Churchill, J. Y. Choi, A. Stathopoulos, C.S. Chang, S. Klasky. **"Towards Real-Time Detection and Tracking of Blob-Filaments in Fusion Plasma Big Data,"** In *Department of Computer Science, College of William and Mary, WM-CS-2015-01*, 2015.
156. [Xie2015a] Jinrong Xie, Franz Sauer, Kwan-Liu Ma. **"Fast Uncertainty-driven Large-scale Volume Feature Extraction on Desktop PCs,"** In *Proceedings of LDAV 2015*, IEEE, pp. 17-24. October, 2015.
DOI: 10.1109/LDAV.2015.7348067
157. [Ye2015a] Yucong Ye, Yang Wang, Robert Miller, Kwan-Liu Ma, Kenji Ono. **"In Situ Depth Maps Based Feature Extraction and Tracking,"** In *Proceedings of LDAV 2015*, IEEE, pp. 1-8. October, 2015.
DOI: 10.1109/LDAV.2015.7348065
158. [Yoo2015a] W. Yoo, M. Koo, Y. Cao, A. Sim, P. Nugent, K. Wu. **"PATHA: Performance Analysis Tool for HPC Applications,"** In *the 34th IEEE International Performance Computing and Communications Conference (IPCCC)*, 2015.
159. [Yu2015a] Hongfeng Yu, Jinrong Xie, Kwan-Liu Ma, Hemanth Kolla, Jacqueline H. Chen. **"Scalable Parallel Distance Field Construction for Large-Scale Applications,"** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 21, No. 10, pp. 1187-1200. August, 2015.
160. [Zha2015a] Wenzhao Zhang, Houjun Tang, Xiaocheng Zou, Steven Harenberg, Qing Liu, Scott Klasky, Nagiza F. Samatova. **"Exploring Memory Hierarchy to Improve Scientific Data Read Performance,"** In *Cluster Computing*, September, 2015.
161. [Zha2015b] Yanwei Zhang, Matthew Wolf, Karsten Schwan, Qing Liu, Greg Eisenhauer,Scott Klasky. **"Co-Sites: The Autonomous Distributed Dataflows in Collaborative Scientific Discovery,"** In *10th Workshop on Workflows in Support of Large-Scale Science (WORKS '15), in conjunction with SC'15*, November, 2015.
162. [Zou2015a] Xiaocheng Zou, Kesheng Wu, David A. Boyuka II, Daniel F. Martin, Suren Byna, Houjun Tang, Kushal Bansal, Terry J. Ligocki, Hans Johansen,, Nagiza F. Samatova. **"Parallel In Situ Detection of Connected Components in Adaptive Mesh Refinement Data,"** In *Proc. Cluster, Cloud and Grid Computing (CCGrid)*, May, 2015.
163. [Zou2015b] Xiaocheng (Chris) Zou, Suren Byna, Hans Johansen, Daniel Martin, Nagiza F. Samatova, Arie Shoshani, John Wu. **"Six-fold Speedup of Ice Calving Detection Achieved by AMR-aware Parallel Connected Component Labeling,"** In *SciDAC PI Meeting*, July, 2015.

2014

164. [Arg2014a] Alexy Agranovsky, David Camp, Christoph Garth, E. Wes Bethel, Kenneth I. Joy, Hank Childs. **"Improved Post Hoc Flow Analysis vis Lagrangian Representations,"** In *Proceedings of the Large Data Analysis and Visualization Symposium (LDAV), Paris, France*, Note: *Best Paper Award*, pp. 67–75. November, 2014.
165. [Bha2014a] H. Bhatia, V. Pascucci, R.M. Kirby, P.-T. Bremer. **"Extracting Features from Time-Dependent Vector Fields Using Internal Reference Frames,"** In *Computer Graphics Forum (Proceedings of EuroVis)*, Vol. 33, No. 3, pp. 21--30. June, 2014.
166. [Bon2014a] G.P. Bonneau, H.C. Hege, C.R. Johnson, M.M. Oliveira, K. Potter, P. Rheingans, T. Schultz. **"Overview and State-of-the-Art of Uncertainty Visualization,"** In *Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization*, Ch. 1, Edited by M. Chen and H. Hagen and C.D. Hansen and C.R. Johnson and A. Kauffman, Springer-Verlag, pp. 3--27. 2014.

167. [Boy2014a] David A. Boyuka II, Sriram Lakshminarasimhan, Xiaocheng Zou, Zhenhuan Gong, John Jenkins, Eric R. Schendel, Norbert Podhorszki, Qing Liu, Scott Klasky,, Nagiza F. Samatova.. **“ Transparent in situ data transformations in ADIOS,”** In *Proc. Cluster, Cloud and Grid Computing (CCGrid)*, May, 2014.

168. [Bre2014a] P.-T. Bremer, I. Hotz, V. Pascucci, R. Peikert. **“Topological Methods in Data Analysis and Visualization III,”** Mathematics and Visualization, 2014.
    ISBN: Bre2014a

169. [Bui2014a] H. Bui, V. Vishwanath, H. Finkel, K. Harms, J. Leigh, S. Habib, K. Heitmann, M. E. Papka. **“Scalable parallel I/O on Blue Gene/Q supercomputer using compression, topology-aware data aggregation, and subfiling,”** In *Proceedings of the 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP 2014)*, Turin, Italy, February, 2014.

170. [Bui2014b] H. Bui, E.S. Jung, V. Vishwanath, J. Leigh, M. Papka. **“Improving Data Movement Performance for Sparse Data Patterns on Blue Gene/Q Supercomputer,”** In *7th International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2) held in conjunction with the 43rd International Conference on Parallel Processing*, Minneapolis, Minnesota, USA, September, 2014.

171. [Cha2014a] A. Chaudhuri, T.-Y. Lee, H.-W. Shen, R. Wenger. **“Exploring Flow Fields Using Space-filling Analysis of Streamlines,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, In *IEEE*, No. 99, 2014.
    DOI: 10.1109/TVCG.2014.2312009

172. [Cha2014b] A. Chaudhuri, T.-H. Wei, T.-Y. Lee, H.-W. Shen, T. Peterka. **“Efficient Range Distribution Query for Visualizing Scientific Data,”** In *Proceedings of the 2014 IEEE Pacific Visualization Symposium (PacificVis)*, 2014.

173. [Cha2014c] B. Chapman, H. Calandra, S. Crivelli, J. Dongarra, J. Hittinger, C.R. Johnson, S.A. Lathrop, V. Sarkar, E. Stahlberg, J.S. Vetter, D. Williams. **“ASCAC Workforce Subcommittee Letter,”** Subtitled **“DOE ASCAC Committee Report,”** 2014.

174. [Che2014a] Zhengzhang Chen, Seung Woo Son, William Hendrix, Ankit Agrawal, Wei-keng Liao, Alok Choudhary. **“NUMARCK: Machine Learning Algorithm for Resiliency and Checkpointing,”** In *the International Conference for High Performance Computing, Networking, Storage and Analysis*, November, 2014.

175. [Chi2014a] Hsuan-Te Chiu, Jerry Chou, Venkat Vishwanath, Suren Byna,, Kesheng Wu,. **“Simplifying Index File Structure to Improve I/O Performance of Parallel Indexing,”** In *The 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS 2014)*, 2014.

176. [Chi2014b] Hank Childs, Scott Biersdorff, David Poliakoff, David Camp, Allen D. Malony. **“Particle Advection Performance Over Varied Architectures and Workloads,”** In *21th Annual International Conference on High Performance Computing, HiPC 2014*, goa, india, dec, 2014.

177. [Dai2014a] Dong Dai, Robert B. Ross, Philip Carns, Dries Kimpe, Yong Chen. **“Using Property Graphs for Rich Metadata Management in HPC Systems,”** In *Proceedings of the 9th Parallel Data Storage Workshop, IEEE*, 11, 2014.

178. [Doc2014a] Ciprian Docan, Fan Zhang, Tong Jin, Hoang Bui, Qian Sun, Julian Cummings, Norbert Podhorszki, Scott Klasky, Manish Parashar. **“ActiveSpaces: Exploring dynamic code deployment for extreme scale data processing,”** In *Concurrency and Computation: Practice and Experience*, 2014.

179. [Fre2014a] Steffen Frey, Filip Sadlo, Kwan-Liu Ma,, Thomas Ertl. **“Interactive Progressive Visualization with Space-Time Error Control,”** In *Proceedings of IEEE SciVis 2014 (also IEEE TVCG 20(12))*, November, 2014.

180. [Gen2014a] M.G. Genton, C.R. Johnson, K. Potter, G. Stenchikov, Y. Sun. **“Surface boxplots,”** In *Stat Journal*, Vol. 3, No. 1, pp. 1--11. 2014.

181. [Gyu2014a] A. Gyulassy, P.-T. Bremer, R. Grout, H. Kolla, J. Chen,, V. Pascucci. **“Stability of dissipation elements: A case study in combustion,”** In *Computer Graphics Forum*, 2014.

182. [Gyu2014b] Attila Gyulassy, David Guenther, Joshua A. Levine, Julien Tierny, Valerio Pascucci. **“Conforming Morse-Smale Complexes,”** In *Trans. of Vis. and Comp. Graphics, Proc. IEEE Visualization*, 2014.

183. [Ham2014a] Jeff R. Hammond, Andreas Schäfer, Rob Latham. **“To INT_MAX...and beyond!,”** Subtitled **“ Exploring large-count support in MPI,”** In *Workshop on Exascale MPI at Supercomputing Conference 2014*, Nov, 2014.

184. [Han2014a] C.D. Hansen, M. Chen, C.R. Johnson, A.E. Kaufman, H. Hagen (Eds.). **“Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization,”** In *Mathematics and Visualization*, Springer, 2014.

185. [Hei2014a] Heitmann,K.,Habib,S.,Finkel,H.,Frontiere,N.,Pope,A.,Morozov,V.,Rangel,S.,Kovacs,E.,Kwan,J.,Li, N., Rizzi, S., Insley, J., Vishwanath, V., Peterka, T., Daniel, D., Fasel, P., Zagaris, G.. **“Large Scale Simulations of Sky Surveys,”** In *Computing in Science and Engineering*, Sept-Oct, 2014.

186. [Hua2014a] Jian Huang, Xuechen Zhang, Greg Eisenhauer, Karsten Schwan, Matthew Wolf, Stephane Ethier, Scott Klasky. **“Scibox: Online Sharing of Scientific Data via the Cloud,”** In *International Conference on Parallel and Distributed Processing (IPDPS), IEEE*, Feb, 2014.

187. [Huc2014a] K. A. Huck, K. Potter, D. W. Jacobsen, H. Childs,, A. D. Malony. **“Linking Performance Data into Scientific Visualization Tools,”** In *1st Workshop on Visual Performance Analysis (VPA), held in conjuction with SC14, New Orleans, LA, Nov. 2014*, 2014.

188. [Hue2014a] A. Huebl, D. Pugmire, F. Schmitt, R. Pausch, M. Bussman. **“Visualizing the Radiation of the Kelvin-Helmholtz Instability,”** In *7th Triennial Special Issue of the IEEE Images on Plasma Science*, In *IEEE Images on Plasma Science*, April, 2014.

189. [Jen2014a] John Jenkins, Xiaocheng Zou, Houjun Tang, Dries Kimpe, Robert Ross,, Nagiza F. Samatova. **“RADAR: Runtime asymmetric data-access driven scientific data replication,”** In *Proceedings of the 2014 International Supercomputing Conference*, 2014.

190. [Jin2014a] T. Jin, F. Zhang, Q. Sun, H. Bui, N. Podhorszki, S. Klasky, H. Kolla, J. Chen, R. Hager, C.S. Chang, M. Parashar. **“Leveraging Deep Memory Hierarchies for Data Staging in Coupled Data Intensive Simulation Workflows,”** In *IEEE Cluster 2014*, 2014.

191. [Kar2014a] Karimabadi H., Roytershteyn V., Vu H.X., Omelchenko Y.A., Scudder J., Daughton W., Dimmock A., Nykyri K., Wan M., Sibeck D., Tatineni M., Majumdar A., Loring B., Geveci B.. **“The link between shocks, turbulence, and magnetic reconnection in collisionless plasmas,”** In *Physics of Plasmas, AIP Publishing*, 2014.

192. [Kla2014a] S. Klasky, Q. Liu, H. Abbasi, N. Podhorszki, J. Chen, Hemanth Kolla. **“Scaling up Parallel I/O in S3D to 100K cores with ADIOS,”** In *High Performance Parallel I/O*, Edited by Prabhat, Q. Koziol, Taylor and Francis, 2014.

193. [Kum2014a] S. Kumar, C. Christensen, J. A. Schmidt, P.-T. Bremer, E. Brugger, V. Vishwanath, P. Carns, H. Kolla, R. Grout, J. Chen, M. Berzins, V. Pascucci. **“Fast Multi-Resolution Reads of Massive Simulation Datasets,”** In *Proc. Int. Supercomputing Conference* , pp. 314-330. 2014.

194. [Kum2014b] S. Kumar, J. Edwards, P.-T. Bremer, A. Knoll, C. Christensen, V. Vishwanath, P. Carns, J. A. Schmidt, V. Pascucci. **“Efficient I/O and storage of adaptive resolution data,”** In *Proc. Supercomputing (SC) 2014*, 2014.

195. [Lak2014a] Sriram Lakshminarasimhan, Xiaocheng Zou, David A Boyuka II, Saurabh V Pendse, John Jenkins, Venkatram Vishwanath, Michael E Papka, Scott Klasky,, Nagiza F Samatova.. **"Diraq: scalable in situ data-and resource-aware indexing for optimized query performance,"** In *Cluster Computing*, Springer,US, pp. 1-19. 2014.
ISSN: 1386-7857
DOI: 10.1007/s10586-014-0358-z

196. [Lan2014a] Aaditya G. Landge, Valerio Pascucci, Attila Gyulassy, Janine C. Bennett, Hemanth Kolla, Jacqueline Chen, Peer-Timo Bremer. **"In-Situ Feature Extraction of Large Scale Combustion Simulations Using Segmented Merge Trees,"** In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC14*, 2014.

197. [Lat2014a] R. Latham. **"Parallel-NetCDF,"** In *High Performance Parallel I/O*, CRC Press, Taylor and Francis Group, October, 2014.

198. [Lia2014a] W. Liao, R. Thakur. **"MPI-IO,"** In *High Performance Parallel I/O*, CRC Press, Taylor and Francis Group, October, 2014.

199. [Liu2014a] S. Liu, Bei Wang, P.-T. Bremer, V. Pascucci. **"Distortion-Guided Structure-Driven Interactive Exploration of High-Dimensional Data,"** In *Computer Graphics Forum (CGF) (Proceedings of EuroVis)*, Vol. 33, No. 3, pp. 101--110. 2014.

200. [Liu2014b] S. Liu, B. Wang, J. Thiagarajan, P.-T. Bremer, V. Pascucci. **"Multivariate Volume Visualization through Dynamic Projections,"** In *Proc. IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)*, 2014.

201. [Liu2014c] Ruoqian Liu,, Ankit Agrawal, Wei-keng Liao,, Alok Choudhary. **"Search Space Preprocessing in Solving Complex Optimization Problems,"** In *the Workshop on Complexity for Big Data held in conjunction with the IEEE International Conference on Big Data*, Washington, D.C., U.S.A., October, 2014.

202. [Lu2014a] Kewei Lu, Han-Wei Shen, Tom Peterka. **"Scalable Computation of Stream Surfaces on Large Scale Vector Field,"** In *ACM/IEEE SC' 14*, 2014.

203. [Mal2014a] D. Maljovec, Bei Wang, J. Moeller, V. Pascucci. **"Topology-Based Active Learning,"** SCI Technical Report, No. UUSCI-2014-001, *SCI Institute, University of Utah*, 2014.

204. [Mal2014b] D. Maljovec, S. Liu, Bei Wang, V. Pascucci, P.-T. Bremer, D. Mandelli, C. Smith. **"Analyzing Simulation-Based PRA Data Through Clustering: a BWR Station Blackout Case Study,"** In *Proceedings of the Probabilistic Safety Assessment & Management conference (PSAM)*, 2014.

205. [Man2014a] D. Mandelli, C. Smith, T. Riley, J. Nielsen, J. Schroeder, C. Rabiti, A. Alfonsi, J. Cogliati, R. Kinoshita, V. Pascucci, Bei Wang, D. Maljovec. **"Overview of New Tools to Perform Safety Analysis: BWR Station Black Out Test Case,"** In *Proceedings of the Probabilistic Safety Assessment & Management conference (PSAM)*, 2014.

206. [Mil2014a] Robert Miller, Kenneth Moreland, Kwan-Liu Ma. **"Finely-Threaded History-Based Topology Computation,"** In *Eurographics Symposium on Parallel Graphics and Visualization*, pp. 41--48. June, 2014.
DOI: 10.2312/pgv.20141083

207. [Obe2014a] H. Obermaier, K. I. Joy. **"Future Challenges for Ensemble Visualization,"** In *IEEE Computer Graphics and Applications*, Vol. 34, No. 3, IEEE, pp. 8--11. May/June, 2014.

208. [Pal2014a] Diana Palsetia, Mostofa Patwary, William Hendrix, Ankit Agrawal,, Alok Choudhary. **"Clique Guided Community Detection,"** In *the IEEE International Conference on Big Data*, Washington, DC, IEEE, pp. 500 - 509. October, 2014.
DOI: 10.1109/BigData.2014.7004267

209. [Pas2014a] V. Pascucci, P.-T. Bremer, H. Bhatia. **“The Natural Helmholtz-Hodge Decomposition For Open-Boundary Flow Analysis,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 99, pp. 1. 2014.
DOI: 10.1109/TVCG.2014.2312012

210. [Pet2014a] Peterka, T., Morozov, D., Phillips, C.. **“High-Performance Computation of Distributed-Memory Parallel 3D Voronoi and Delaunay Tessellation,”** In *Proceedings of SC14*, New Orleans, LA, 2014.

211. [Pod2014a] N. Podhorszki, S. Klasky, Q. Liu, Y. Tian, M. Parashar, K. Schwan, M. Wolf, S. Lakshminarasimhan. **“ADIOS,”** In *High Performance Parallel I/O*, Edited by Prabhat, Q. Koziol, Taylor and Francis, 2014.

212. [Pug2014a] David Pugmire, James Kress, Jeremy Meredith, Norbert Podhorszki, Jong Choi, Scott Klasky. **“Towards scalable visualization plugins for data staging workflows,”** In *Big Data Analytics: Challenges and Opportunities (BDAC-14) Workshop at Supercomputing Conference*, Nov, 2014.

213. [Riz2014a] S. Rizzi, M. Hereld, J. Insley, M. Papka, T. Uram,, V. Vishwanath. **“Performance Modeling of vl3 Volume Rendering on GPU-Based Clusters,”** In *Proceedings of the 14th Eurographics Symposium on Parallel Graphics and Visualization (EGPGV), Swansea, Wales, UK, June 9-10*, 2014.

214. [Riz2014b] S. Rizzi, M. Hereld, J. Insley, M. Papka, T. Uram,, V. Vishwanath. **“Large-Scale Parallel Visualization of Particle Datasets using Point Sprites,”** In *Proceedings of the IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2014), New Orleans, LA [Poster].*, November, 2014.

215. [Ros2014a] R. Ross, D. Kimpe. **“Storage Models: Past, Present, and Future,”** In *High Performance Parallel I/O*, CRC Press, Taylor and Francis Group, October, 2014.

216. [Rub2014a] Oliver Ruebel, Cameron G.R. Geddes, Min Chen, Estelle Cormier-Michel, E. Wes Bethel. **“Feature-based Analysis of Plasma-based Particle Acceleration Data,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, In *IEEE Transactions on Visualization and Computer Graphics, July*, Vol. 20, No. 2, February, 2014.

217. [San2014a] A.R. Sanderson. **“An Alternative Formulation of Lyapunov Exponents for Computing Lagrangian Coherent Structures,”** In *Proceedings of the 2014 IEEE Pacific Visualization Symposium (PacificVis)*, Yokahama Japan, 2014.

218. [Sau2014a] Franz Sauer, Hongfeng Yu,, Kwan-Liu Ma. **“Trajectory-based Flow Feature Tracking in Joint Particle/Volume Datasets,”** In *Proceedings of IEEE SciVis 2014 (also IEEE TVCG 20(12))*, November, 2014.

219. [Sha2014a] Sohail Shafii, Harald Obermaier, Bernd Hamann, Kenneth I. Joy. **“Topological Features in Glyph-Based Corotation Visualization,”** In *Topological Methods in Data Analysis and Visualization III - Theory, Algorithms, and Applications*, Mathematics and Visualization, Springer International Publishing, pp. 263--276. 2014.
ISBN: 978-3-319-04098-1
DOI: 10.1007/978-3-319-04099-8_17

220. [Shi2014a] Min Shih, Yubo Zhang, Kwan-Liu Ma, Jayanarayanan Sitaraman,, Dimitri J. Mavriplis. **“Out-of-Core Visualization of Time-Varying Hybrid-Grid Volume Data,”** In *Proceedings of LDAV*, November, 2014.

221. [Sol2014a] Solomon Lasluisa, Fan Zhang, Tong Jin, Ivan Rodero, Hoang Bui, Manish Parashar. **“In-situ feature-based objects tracking for data-intensive scientific and enterprise analytics workflows,”** In *Cluster Computing*, 2014.

222. [Son2014a] Seung Woo Son, Zhengzhang Chen, William Hendrix, Ankit Agrawal, Wei-keng Liao, Alok Choudhary. **“Data Compression for the Exascale Computing Era - Survey,”** In *Supercomputing frontiers and innovations*, Vol. 1, No. 2, pp. 76-88. September, 2014.
ISSN: 2313-8734
DOI: 10.14529/jsfi140205

223. [Sun2014a] Qian Sun, Fan Zhang, Tong Jin, Hoang Bui, Kesheng Wu, Arie Shoshani, Hemanth Kolla, Scott Klasky, Jacqueline Chen, Manish Parashar. **“Scalable Run-time Data Indexing and Querying for Scientific Simulations,”** In *Big Data Analytics: Challenges and Opportunities (BDAC-14) Workshop at Supercomputing Conference*, November, 2014.

224. [Tan2014a] Houjun Tang, Xiaocheng Zou, John Jenkins, David A. Boyuka II, Stephen Ranshous, Dries Kimpe, Scott Klasky,, Nagiza F. Samatova. **“Improving Read Performance with Online Access Pattern Analysis and Prefetching,”** In *Proc. of 20th International European Conference on Parallel Processing*, 2014.

225. [Vis2014a] V. Vishwanath, H. Bui, M. Hereld, M. Papka. **“GLEAN,”** In *High Performance Parallel I/O*, CRC Press, Taylor and Francis Group, October, 2014.

226. [Web2014a] Gunther H. Weber, Hans Johansen, Daniel T. Graves, Terry J. Ligocki. **“Simulating Urban Environments for Energy Analysis,”** In *Proceedings Visualization in Environmental Sciences 2014 (EnvirVis 2014)*, 2014.

227. [Web2014b] Weber, G. H., Hauser, H.. **“Interactive visual exploration and analysis,”** In *Scientific Visualization: Uncertainty, Multifield, Bio-Medical and Scalable Visualization, Mathematics and Visualization (Hansen, C. D., Chen, M., Johnson, C. R., Kaufman, A. E., and Hagen, H., eds.), Springer-Verlag,* pp. 161–174. 2014.

228. [Wid2014a] Wathsala Widanagamaachchi, Peer-Timo Bremer, Christopher Sewell, Li-ta Lo, James Ahrens,, Valerio Pascucci. **“Data-Parallel Halo Finding with Variable Linking Lengths,”** In *Proceedings of the IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV),* Paris, France, November, 2014.

229. [Won2014a] Pak Chung Wong, Han-Wei Shen, Ruby Leung, Samson Hagos, Teng-Yok Lee, Xin Tong, Kewei Lu. **“Visual Analytics of Large-Scale Climate Model Data,”** In *IEEE Syposium on Large Data Analysis and Visualization (LDAV 2014),* October, 2014.

230. [Zha2014a] Yanwei Zhang, Qing Liu, Scott Klasky, Matthew Wolf, Karsten Schwan, Gregg Eisenhauer, Jong Choi, Norbert Podhorszki. **“Active Workflow System for Near Real-Time Extreme-Scale Science,”** In *PPoPP Workshop on Parallel Programming for Analytics Applications* , 2014.

231. [Zha2014b] L. Zhang, Q. Deng, R. Machiraju, A. Rangaragjan, D. Thompson, D. K. Walters, Han-Wei Shen. **“Boosting Techniques for Physics-Based Vortex Detection,,”** In *Computer Graphics Forum, Vol. 33, Issue 1*, pp. 282-293. February, 2014.

232. [Zho2014a] L. Zhou, C.D. Hansen. **“GuideME: Slice-guided Semiautomatic Multivariate Exploration of Volumes,”** In *Proceedings of the Eurographics Conference on Visualization (EuroVis) 2014,* Vol. 33, No. 3, 2014.

233. [Zou2014a] Xiaocheng Zou, Sriram Lakshminarasimhan , David A. Boyuka II, Stephen Ranshous, Houjun Tang, Scott Klasky ,, Nagiza F. Samatova. **“Fast Set Intersection through Run-time Bitmap Construction over PForDelta-compressed Indexes,”** In *Proc. of 20th International European Conference on Parallel Processing (EuroPar),* 2014.

2013

234. [Agr2013a] A. Agranovsky, H. Obermaier, K. I. Joy. **“A Framework for the Visualization of Finite-Time Continuum Mechanics Effects in Time-Varying Flow,”** In *9th International Symposium on Visual Computing (ISVC 2013),* 2013.

235. [Ahe2013a] S. Ahern, E. Brugger, B. Whitlock, J. S. Meredith, K. Biagas, M. C. Miller, H. Childs. **“VisIt: Experiences with Sustainable Software,”** In *Workshop on Sustainable Software for Science: Practice and Experiences*, IEEE Supercomputing 2013, Nov, 2013.

236. [Beh2013a] B. Behzad, H. Luu, J. Huchette, S. Byna, Prabhat, R. Aydt, Q. Koziol,, M. Snir. **“Taming Parallel I/O Complexity with Auto-Tuning,”** In *SC13*, Nov, 2013.

237. [Bet2013a] E. Wes Bethel, Prabhat, Suren Byna, Oliver Rbel, K. John Wu, Michael Wehner. **“Why High Performance Visual Data Analytics is both Relevant and Difficult,”** In *Proceedings of Visualization and Data Analysis 2013, IS&T/SPIE Electronic Imaging*, San Francisco, CA, USA, 2013.

238. [Bha2013a] H. Bhatia, G. Norgard, V. Pascucci,, P.-T. Bremer. **“The Helmholtz-Hodge decomposition- a survey,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 19, No. 8, pp. 1386--1404. 2013.

239. [Bha2013b] H. Bhatia, G. Norgard, V. Pascucci, P.-T. Bremer. **“Comments on the Meshless Helmholtz-Hodge decomposition,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 3, pp. 527--528. March, 2013.
DOI: 10.1109/TVCG.2012.62

240. [Bha2013c] Harsh Bhatia, Attila Gyulassy, Hao Wang, Peer-Timo Bremer, Valerio Pascucci. **“Robust Detection of Singularities in Vector Fields,”** In *Topological Methods in Data Analysis and Visualizatuion (TopoInVis)* , 2013.

241. [Bha2013d] H. Bhatia, B. Wang, G. Norgard, V. Pascucci, P.-T. Bremer. **“Local, Smooth, and Consistent Jacobi Set Simplification,”** In *Cornell University Library*, No. arXiv:1307.7752, 2013.

242. [Bis2013a] A. Biswas, S. Dutta, H.-W. Shen, J. Woodring. **“An information-aware framework for exploring multivariate data sets,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 19, No. 12, pp. 2683--2692. 2013.

243. [Bro2013a] C. Brownlee, T. Ize, C.D. Hansen. **“Image-parallel Ray Tracing using OpenGL Interception,”** In *Proceedings of the Eurographics Symposium on Parallel Graphics and Visualization (2013)*, pp. 65--72. 2013.

244. [Byn2013a] S. Byna, A. Uselton, Prabhat, D. Knaak,, Y. He. **“Trillion Particles, 120,000 cores, and 350 TBs: Lessons Learned from a Hero I/O Run on Hopper,”** Cray User Group meeting, 2013.

245. [Cam2013a] David Camp, E. Wes Bethel, Hank Childs. **“Transitioning Data Flow-Based Visualization Software to Multi-Core Hybrid Parallelism,”** In *3rd International Workshop on Data-Flow Execution Models for Extreme Scale Computing (DFM 2013)*, September, 2013.

246. [Car2013a] P. Carns, Y. Yao, K. Harms, R. Latham, R. Ross, K. Antypas. **“Production I/O Characterization on the Cray XE6,”** In *Proceedings of the Cray User Group meeting 2013 (CUG 2013)*, May, 2013.

247. [Cha2013a] Yu-Hsuan Chan, Carlos Correa,, Kwan-Liu Ma. **“The Generalized Sensitivity Scatterplot,”** In *IEEE Transactions on Visualization and Computer Graphics*, 2013.

248. [Cha2013b] J. Chandler, H. Obermaier, K. I. Joy. **“Illustrative Rendering of Particle Systems,”** In *Proc. of Vision, Modeling, and Visualization (VMV 2013)*, pp. 177-185. 2013.

249. [Che2013a] J. Chen, A. Choudhary, S. Feldman, B. Hendrickson, C.R. Johnson, R. Mount, V. Sarkar, V. White, D. Williams. **“Synergistic Challenges in Data-Intensive Science and Exascale Computing,”** DOE ASCAC Data Subcommittee Report, Department of Energy Office of Science, March, 2013.

250. [Che2013b] Chun-Ming Shen, Han-Wei Shen. **“Graph-based Seed Scheduling for Out-of-core FTLE and Pathline Computation ,”** In *IEEE 2013 Large Data Analysis and Visualization* , Atlanta, GA, IEEE, 2013.

251. [Che2013c] C. Chen, H.-W. Shen. **“Graph-based seed scheduling for out-of-core FTLE and pathline computation,”** In *Proceedings of the 2013 IEEE LDAV Symposium*, Note: *Received Honorable Mention*, pp. 15--32. October, 2013.

252. [Che2013d] F. Chen, H. Obermaier, H. Hagen, B. Hamann, J. Tierny, V. Pascucci. **“Topology analysis of time-dependent multi-fluid data using the Reeb graph,”** In *Computer Aided Geometric Design*, Vol. 30, No. 6, pp. 557--566. 2013.
DOI: 10.1109/TVCG.2012.62

253. [Chi2013a] Hank Childs, Berk Geveci, Will Schroeder, Jeremy Meredith, Kenneth Moreland, Christopher Sewell, Torsten Kuhlen, E. Wes Bethel. **“Research Challenges for Visualization Software,”** In *IEEE Computer*, Vol. 46, No. 5, pp. 34--42. May, 2013.
DOI: 10.1109/MC.2013.179

254. [Gam2013a] M. Gamell, I. Rodero, M. Parashar, J.C. Bennett, H. Kolla, J. Chen, P. Bremer, A.G. Landge, A. Gyulassy, P. McCormick, S. Pakin, V. Pascucci. **“Exploring Power Behaviors and Tradeoffs of In-situ Data Analytics,”** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, Denver, Colorado, U.S.A, No. 77, November, 2013.
ISBN: 978-1-4503-2378-9
DOI: 10.1145/2503210.2503303

255. [Gam2013b] Marc Gamell, Ivan Rodero, Manish Parashar, Stephen W. Poole. **“Exploring Energy and Performance Behaviors of Data-Intensive Scientific Workflows on Systems with Deep Memory Hierarchies,”** In *IEEE International Conference on High Performance Computing (HiPC)*, December, 2013.

256. [Ger2013a] S. Gerber, O. Reubel, P.-T. Bremer, V. Pascucci, R.T. Whitaker. **“Morse-Smale Regression,”** In *Journal of Computational and Graphical Statistics*, Vol. 22, No. 1, pp. 193-214. 2013.

257. [Gon2013a] Z. Gong, D. A. Boyuka II, X. Zou, Q. Liu, N. Podhorszki, S. Klasky, X. Ma, N. F. Samatova. **“PARLO: PArallel Run-time Layout Optimization for Scientific Data Explorations with Heterogeneous Access Patterns,”** 13th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), Delft, The Netherlands, May, 2013.

258. [Gos2013a] L. Gosink, K. Bensema, T. Pulsipher, H. Obermaier, M. Henry, H. Childs, K. I. Joy. **“Characterizing and Visualizing Predictive Uncertainty in Numerical Ensembles Through Bayesian Model Averaging,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 12, pp. 2703-2712. 2013.

259. [Gro2013a] A. Grosset, M. Schott, G.-P. Bonneau, C.D. Hansen. **“Evaluation of Depth of Field for Depth Perception in DVR,”** In *Proceedings of the 2013 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 81--88. 2013.

260. [Hab2013a] S. Habib, V. Morozov, N. Frontiere, H. Finkel, A. Pope, K. Heitmann, K. Kumaran, V. Vishwanath, T. Peterka, J. Insley, D. Daniel, P. Fasel, Z. Lukic. **“HACC: Extreme Scaling and Performance Across Diverse Architectures,”** In *the Proceedings of the IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC 2013), Denver, Colorado, USA, November 2013*, Note: *Gordon Bell Finalist*, 2013.

261. [Hal2013a] M. Hall, R.M. Kirby, F. Li, M.D. Meyer, V. Pascucci, J.M. Phillips, R. Ricci, J. Van der Merwe, S. Venkatasubramanian. **“Rethinking Abstractions for Big Data: Why, Where, How, and What,”** In *Cornell University Library*, No. arXiv:1306.3295, 2013.

262. [Hen2013a] W. Hendrix, D. Palsetia, M. M. A. Patwary, A. Agrawal, W.-k. Liao, A. Choudhary. **“A Scalable Algorithm for Single-Linkage Hierarchical Clustering on Distributed Memory Architectures,”** In *Proceedings of 3rd IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)*, 2013.

263. [Her2013a] S Herbein, M Matheny, Matthew Wezowicz, J Krogel, Jeremy Logan, J Kim, Scott Klasky, Michela Taufer. **“Performance Impact of I/O on QMCPack Simulations at the Petascale and Beyond,”** In *Computational Science and Engineering (CSE)*, 2013.

264. [Hol2013a] T. Hollt, A. Magdy, G. Chen, G. Gopalakrishnan, I. Hoteit, C.D. Hansen, M. Hadwiger. **“Visual Analysis of Uncertainties in Ocean Forecasts for Planning and Operation of Off-Shore Structures,”** In *Proceedings of IEEE Pacific Visualization 2013,* Note: *Received Honerable Mention*, 2013.

265. [Hum2013a] M. Hummel, H. Obermaier, C. Garth, K. I. Joy. **“Comparative Visual Analysis of Lagrangian Transport in CFD Ensembles,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 12, Note: *Recieved Best Paper award*, pp. 2743-2752. 2013.

266. [Jen2013a] J. Jenkins, I. Arkatkar, S. Lakshminarasimhan, D. A. Boyuka II, E. R. Schendel, N. Shah, S. Ethier, C.S. Chang, J. Chen, H. Kolla, S. Klasky, R. Ross,, N. F. Samatova. **“ALACRITY: Analytics-driven Lossless Data Compression for Rapid In-situ Indexing, Storing, and Querying,”** In *LNCS Transactions on Large-Scale Data- and Knowledge-Centered Systems (TLDKS)*, 2013.

267. [Jin2013a] T. Jin, F. Zhang, Q. Sun, H. Bui, M. Parashar, H. Yu, S. Klasky, N. Podhorszki, H. Abbasi. **“Using Cross-Layer Adaptations for Dynamic Data Management in Large Scale Coupled Scientific Workflows,”** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, Denver, Colorado, U.S.A, November, 2013.

268. [Jin2013b] C. Jin, M. M. A. Patwary, A. Agrawal, W. Hendrix, W.-k. Liao, A. Choudhary. **“DiSC: A Distributed Single-Linkage Hierarchical Clustering Algorithm using MapReduce,”** In *Proceedings of the 4th International SC Workshop on Data Intensive Computing in the Clouds (DataCloud)*, 2013.

269. [Kar2013a] Karimabadi, H., Loring, B., O'Leary, P., Majumdar, A., Tatineni, M., Geveci, B.. **“In-situ Visualization for Global Hybrid Simulations,”** In *Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery, XSEDE '13*, pp. 1-8. 2013.

270. [Kno2013a] Aaron Knoll, Kah Chun Lau, Bin Liu, Maria K.Y. Chan, Aslihan Sumer, Jeffrey Greeley, Larry Curtiss, Julius Jellinek, Mark Hereld, Michael E. Papka. **“Uncertainty Classification and Visualization of Molecular Interfaces,”** In *International Journal of Uncertainty Quantification, 3(2)*, 2013.

271. [Kum2013a] Sidharth Kumar, Avishek Saha, Venkatram Vishwanath, Philip Carns, John A. Schmidt, Giorgio Scorzelli, Hemanth Kolla, Ray Grout, Robert Latham, Robert Ross, Michael E. Papka, Jacqueline Chen, Valeri Pascucci. **“Characterization and Modeling of PIDX Parallel I/O for Performance Optimization,”** In *Proceedings of the ACM/IEEE Conference on High Performance Computing, Networking, Storage and Analysis (SC)*, IEEE Computer Society Press, pp. 67. November, 2013.

272. [Lak2013a] S. Lakshminarasimhan, D. A. Boyuka II, S. V. Pendse, X. Zou, J. Jenkins, V. Vishwanath, M. E. Papka,, N. F. Samatova. **“Scalable in situ scientific data encoding for analytical query processing,”** In *Proceedings of the 21st ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, Note: *Best Paper Award*, 2013.

273. [Las2013a] S. Lasluisa, F. Zhang, T. Jin, I. Rodero, H. Bui, M. Parashar. **“In-situ Feature-based Objects Tracking for Data-Intensive Scientific and Enterprise Analytics Workflows,”** In *Submitted ot Journal of Cluster Computing*, 2013.

274. [Lea2013a] Nick Leaf, Venkatram Vishwanath, Joseph Insley, Mark Hereld, Michael E. Papka, Kwan-Liu Ma. **“Efficient Parallel Volume Visualization of Large-Scale Adaptive Mesh Refinement Data,”** In *Proceedings of IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, Note: *Best Paper Award*, pp. 35--42. 2013.
DOI: 10.1109/LDAV.2013.6675156

275. [Lee2013a] Teng-Yok Lee, Xin Tong, Han-Wei Shen, Pak Chung Wong, Samson Hagos, L. Ruby Leun. **“Feature Tracking and Visualization of Madden-Julian Oscillation in Climate Simulation,”** In *IEEE Computer Graphics and Applications* , IEEE, July, 2013.

276. [Lee2013b] T.-Y. Lee, H.-W. Shen. **“Efficient local statistical analysis via integral histograms with discrete wavelet transform,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 19, No. 12, pp. 2693--2702. 2013.

277. [Lef2013a] Matthieu Lefebvre, Ebru Bozdag, Henri Calandra, Judy Hill, Wenjie Lei, Daniel Peter, Norbert Podhorszki, Dave Pugmire, Herurisa Rusmanugroho, James Smith, Jeroen Tromp. **“A Data Centric View of Large-Scale Seismic Imaging Workflows,”** In *4th SC Workshop on Petascale (Big) Data Analytics: Challenges and Opportunities*, 2013.

278. [Lin2013a] K-W Lin, J. Chou, S. Byna,, K. Wu. **“Optimizing FastQuery Performance on Luster File System,”** 25th International Conference on Scientific and Statistical Database Management (SSDBM) , 2013.

279. [Liu2013a] Qing Liu, Norbert Podhorszki, Jeremy Logan, Scott Klasky. **“Runtime I/O Re-Routing + Throttling on HPC Storage,”** In *USENIX HotStorage*, 2013.

280. [Liu2013b] Qing Liu, Jeremy Logan, Yuan Tian, Hasan Abbasi, Norbert Podhorszki, Jong Youl Choi, Scott Klasky, Roselyne Tchoua, Jay Lofstead, Ron Oldfield, Manish Parashar, Nagiza Samatova, Karsten Schwan, Arie Shoshani, Matthew Wolf, Kesheng Wu, Weikuan Yu. **“Hello ADIOS: the challenges and lessons of developing leadership class I/O frameworks,”** In *Concurrency and Computation: Practice and Experience*, 2013.

281. [Lof2013a] J. Lofstead, R. Ross. **“Insights for exascale IO APIs from building a petascale IO API,”** In *Proceedings of SC13: International Conference for High Performance Computing, Networking, Storage and Analysis, ACM*, November, 2013.

282. [Luk2013a] Kewei Lu, Abon Chaudhuri, Teng-Yok Lee, Han-Wei Shen, Pak Chung Wong. **“Exploring Vector Fields with Distribution-based Streamline Analysis,”** In *IEEE Pacific Visualization 2013*, IEEE, 2013.

283. [Mal2013a] D. Maljovec, B. Wang, A. Kupresanin, G. Johannesson, V. Pascucci,, P.-T. Bremer. **“Adaptive sampling with topological scores,”** In *International Journal for Uncertainty Quantification*, Vol. 3, No. 2, pp. 119--141. 2013.

284. [Mal2013b] Dan Maljovec, Avishek Saha, Peter Lindstrom, Peer-Timo Bremer, Bei Wang, Carlos Correa, Valerio Pascucci. **“A Comparative Study of Morse Complex Approximation Using Different Neighborhood Graphs,”** Topological Methods in Data Analysis and Visualizatuion (TopoInVis), 2013.

285. [Mal2013c] D. Maljovec, B. Wang, V. Pascucci, P.-T. Bremer, M. Pernice, D. Mandelli, R. Nourgaliev. **“Exploration of high-dimensional scalar function for nuclear reactor safety analysis and visualization,”** In *Proceedings of International Conference on Mathematics and Computational Methods Applied to Nuclear Science & Engineering*, pp. 712--723. 2013.

286. [Mal2013d] D. Maljovec, B. Wang, D. Mandelli, P.-T. Bremer, V. Pascucci. **“Adaptive Sampling Algorithms for Probabilistic Risk Assessment of Nuclear Simulations,”** In *Proceedings*

of the 2013 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2013), 2013.

287. [Mal2013e] D. Maljovec, B. Wang, D. Mandelli, P.-T. Bremer, V. Pascucci. **“Analyzing Dynamic Probabilistic Risk Assessment Data through Clustering,”** In *Proceedings of the 2013 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2013)*, 2013.
288. [Mar2013a] Steve Martin, Han-Wei Shen. **“Transformations for Volumetric Range Distribution Queries,”** In *IEEE Pacific Visualization 2013*, IEEE, 2013.
289. [Mar2013b] S. Martin, H.-W. Shen. **“Transformations for volumetric range distribution queries,”** In *Proceedings of the 2013 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 89--96. 2013.
290. [May2013a] Robert Maynard, Kenneth Moreland, Utkarsh Ayachit, Berk Geveci, Kwan-Liu Ma. **“Optimizing Threshold for Extreme Scale Analysis,”** In *Proceedings SPIE Visualization and Data Analysis*, February, 2013.
    DOI: 10.1117/12.2007320
291. [Mor2013a] Kenneth Moreland. **“A Survey of Visualization Pipelines,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 3, pp. 367-378. March, 2013.
    DOI: 10.1109/TVCG.2012.133
292. [Mor2013b] Kenneth Moreland, Berk Geveci, Kwan-Liu Ma, Robert Maynard. **“A Classification of Scientific Visualization Algorithms for Massive Threading,”** In *Proceedings of the 8th International Workshop on Ultrascale Visualization (UltraVis '13)*, November, 2013.
    DOI: 10.1145/2535571.2535591
293. [Obe2013a] H. Obermaier, K. I. Joy. **“Local Data Models for Probabilistic Transfer Function Design,”** In *Eurographics Conference on Visualization (EuroVis 2013) - Short Papers*, pp. 43-47. 2013.
294. [Pat2013a] M. M. A. Patwary, D. Palsetia, A. Agrawal, W.-k. Liao, F. Manne, A. Choudhary. **“Scalable Parallel OPTICS Data Clustering Using Graph Algorithmic Techniques,”** In *Proceedings of 25th International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC)*, 2013.
295. [Phi2013a] S. Philip, B. Summa, J. Tierny, P.-T. Bremer V. Pascucci. **“Scalable Seams for Gigapixel Panoramas,”** In *Proceedings of the 2013 Eurographics Symposium on Parallel Graphics and Visualization*, Note: *Awarded Best Paper!*, pp. (to appear). 2013.
296. [Pri2013a] Iuri Prilepov, Harald Obermaier, Eduard Deines, Christoph Garth, Kenneth I. Joy. **“Cubic Gradient-Based Material Interfaces,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 10, pp. 1687-1699. 2013.
297. [Red2013a] Khairi Reda, Aaron Knoll, Ken-ichi Nomura, Michael E. Papka, Andrew E. Johnson,, Jason Leigh. **“ Visualizing Large-Scale Atomistic Simulations in Ultra-Resolution Immersive Environments,”** In *Proc. IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)*, pp. 59-65. 2013.
298. [Rom2013a] A. Romosan, A. Shoshani, K. Wu, V. Markowitz,, K. Mavrommatis. **“Accelerating Gene Context Analysis Using Bitmaps,”** In *25th International Conference on Scientific and Statistical Database Management (SSDBM)*, 2013.
299. [Run2013a] A. Rungta, B. Summa, D. Demir, P.-T. Bremer, V. Pascucci. **“ManyVis: Multiple Applications in an Integrated Visualization Environment,”** Vol. 19, No. 12, pp. 2878--2885. 2013.

300. [Sau2013a] F. Sauer, Y. Zhang, S. Ethier, W. Wang,, K.-L. Ma. **“Visualizing Velocity Fields from Fusion Simulations using Advected Particle Trajectories,”** In *The 23rd International Conference on Numerical Simulation of Plasmas (ICNSP), September 14-16*, 2013.

301. [Sch2013a] M. Schott, T. Martin, A.V.P. Grosset, S.T. Smith, C.D. Hansen. **“Ambient Occlusion Effects for Combined Volumes and Tubular Geometry,”** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 19, No. 6, Note: *Selected as Spotlight paper for June 2013 issue*, pp. 913--926. 2013.
DOI: 10.1109/TVCG.2012.306

302. [Sch2013b] E. Schendel, S. Harenberg, H. Tang, V. Vishwanath, M.E. Papka, N. Samatova. **“A Generic High-performance Method for Deinterleaving Scientific Data,”** In *19th International European Conference on Parallel and Distributed Computing (EuroPar), Aachen, Germany, August*, 2013.

303. [Sew2013a] Christopher Sewell, Li-ta Lo,, James Ahrens. **“Portable Data-Parallel Visualization and Analysis in Distributed Memory Environments,”** In *Proceedings of the IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)*, October, 2013.

304. [Sha2013a] Sohail Shafii, Harald Obermaier, Rodman Linn, Eunmo Koo, Mario Hlawitschka, Christoph Garth, Bernd Hamann, Kenneth I. Joy. **“Visualization and Analysis of Vortex-Turbine Intersections in Wind Farms,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 9, pp. 1579-1591. 2013.

305. [Sha2013b] S. Shafii, H. Obermaier, V. Kolar, M. Hlawitschka, C. Garth, B. Hamann, K. I. Joy. **“Illustrative Rendering of Vortex Cores,”** In *Eurographics Conference on Visualization (EuroVis 2013) - Short Papers*, pp. 61-65. 2013.

306. [Sig2013a] Carmen Sigovan, Chris Muelder,, Kwan-Liu Ma. **“Visualizing Large-scale Parallel Communication Traces Using a Particle Animation Technique,”** Computer Graphics Forum, 32(3), also Proceedings of EuroVis, 2013.

307. [Sig2013b] C. Sigovan, C. Muelder, K.-L. Ma, J. Cope, K. Iskra,, R. Ross. **“A Visual Network Analysis Methodology for Large Scale Parallel I/O Systems,”** Proceedings of IEEE International Parallel & Distributed Processing Symposium, 2013.

308. [Sis2013a] R. Sisneros, J. Huang, G. Ostrouchov, S. Ahern, B. D. Semeraro. **“Contrasting Climate Ensembles: A Model-Based Visualization Approach for Analyzing Extreme Events,”** In *International Conference on Computational Sciences, Procedia Computer Science*, Vol. 18, pp. 2347-2356. 2013.
ISSN: 1877-0509
DOI: 10.1016/j.procs.2013.05.406

309. [Sla2013a] Magdalena Slawinska, Tanja Bode, Jeremy Logan, Michael Clark, Hongbo Zou, Matthew Kinsey, Matthew Wolf, Pablo Laguna,, Scott Klasky. **“A Maya use case: adaptable scientific workflows with ADIOS for general relativistic astrophysics,”** In *Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery Article No. 54, ACM SIGAPP, July 22-25, 2013. San Diego, CA, USA*, 2013.

310. [Smi2013a] Smith, J. A.; Bozdag, E.; Krischer, L.; Lefebvre, M.; Lei, W.; Podhorszki, N.; Tromp, J. . **“An Adaptable Seismic Data Format for Modern Scientific Workflows,”** In *American Geophysical Union, Fall Meeting*, 2013.

311. [Sum2013a] B. Summa, A. Gyulassy, P.-T. Bremer, V. Pascucci. **“Interactive Data Exploration,”** In *Data Intensive Science. Chapman & Hall/Crc Computational Science*, Note: *Book chapter*, 2013.

312. [Sun2013a] Q. Sun, F. Zhang, T. Jin, H. Bui, M. Parashar, K. Wu, A. Shoshani, H. Kolla, S. Klasky, J. Chen. **“Scalable In-Memory Data Indexing and Querying for Scientific Simulation**

Workflows," In *Submitted to IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 2013.

313. [Szc2013a] A. F. Szczepanski, Jian Huang, T. Baer, Y. C. Mack, S. Ahern. **"Data Analysis and Visualization in High-Performance Computing,"** In *Computer*, Vol. 46, No. 5, IEEE Computer Society, pp. 84-92. May, 2013.
ISSN: 0018-9162
DOI: 10.1109/MC.2012.192

314. [Tia2013a] Y. Tian, Z, Liu, S. Klasky, B. Wang, H. Abbasi, S. Zhou, N. Podhorszki, T. Clune, J. Logan, W. Yu. **"A Lightweight I/O Scheme to Facilitate Spatial and Temporal Queries of Scientific Data Analytics,"** In *MSST*, 2013.

315. [Tuy2013a] Ying Tu, Han-Wei Shen . **"GraphCharter: Combining Browsing with Query to Explore Large Semantic Graphs,"** In *IEEE Pacific Visualization 2012* , IEEE, 2013.

316. [Vat2013a] Ranga Raju Vatsavai. **"Gaussian multiple instance learning approach for mapping the slums of the world using very high resolution imagery,"** In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1419--1426. 2013.
DOI: 10.1145/2487575.2488210

317. [Wan2013a] Yang Wang, Hongfeng Yu,, Kwan-Liu Ma. **" Scalable Parallel Feature Extraction and Tracking for Large Time-varying 3D Volume Data,"** Proceedings of Eurographics Symposium on Parallel Graphics and Visualization, 2013.

318. [Wan2013b] Bei Wang, Paul Rosen, Primoz Skraba, Harsh Bhatia, Valerio Pascucci. **"Visualizing Robustness of Critical Points for 2D Time-Varying Vector Fields,"** In *Computer Graphics Forum (EuroVis)*, Vol. 32, No. 3, 2013.

319. [Wan2013c] Y. Wan, H. Otsuna, C.D. Hansen. **"Synthetic Brainbows,"** In *Computer Graphics Forum*, Vol. 32, No. 3pt4, pp. 471--480. 2013.

320. [Web2013a] G.H. Weber, K. Beketayev, P.-T. Bremer, B. Hamann, M. Haranczyk, M. Hlawitschka, V. Pascucci. **"Comprehensible Presentation of Topological Information,"** No. LBNL-5693E, *Lawrence Berkeley National Laboratory*, 2013.

321. [Wei2013a] Tzu-Hsuan Wei, Teng-Yok Lee,, Han-Wei Shen. **"Evaluating Isosurfaces with Level-est-based Information Maps,"** In *EuroVis 2013*, Eurographics, 2013.

322. [Wid2013a] W. Widanagamaachchi, P. Rosen, V. Pascucci. **"A Flexible Framework for Fusing Image Collections into Panoramas,"** In *Proceedings of the 2013 SIBGRAPI Conference on Graphics, Patterns, and Images*, Note: *Awarded Best Paper!*, pp. 195-202. 2013.
DOI: 10.1109/SIBGRAPI.2013.35

323. [Wil2013a] D. Williams, T. Bremer, C. Doutriaux, J. Patchett, G. Shipman, B. Haugen, R. Miller, B. Smith, C. Steed, W. Bethel, H. Childs, H. Krishnan, M. Wehner, C. T. Silva, E. Santos, D. Koop, T. Ellqvist, H. T. Vo, J. Poco, B. Geveci, A. Chaudhary, A. Bauer, A. Pletzer, D. Kindig, G. L. Potter, T. P. Maxwell. **"The Ultra-scale Visualization Climate Data Analysis Tools: Data Analysis and Visualization for Geoscience Data ,"** In *IEEE Special Issue: Cutting-Edge Research in Visualization* , 2013.

324. [Wu2013a] Y. Wu, X. Liu, S. Liu,, K.-L. Ma. **"ViSizer: A Visualization Resizing Framework,"** IEEE Transactions on Visualization and Computer Graphics, 19(2):278-290 , 2013.

325. [Xie2013a] Jinrong Xie, Hongfeng Yu,, Kwan-Liu Ma. **"Interactive Ray Casting of Geodesic Grids,"** Computer Graphics Forum, 32(3), also Proceedings of EuroVis, 2013.

326. [Yuc2013a] Yucong Ye, Robert Miller,, Kwan-Liu Ma. **"In-Situ Pathtube Visualization with Explorable Images,"** Proceedings of Eurographics Symposium on Parallel Graphics and Visualization, 2013.

327. [Zha2013a] F. Zhang, C. Docan, H. Bui, M. Parashar, S. Klasky. **"XpressSpace: A Programming Framework for Coupling PGAS Simulation Codes,"** In *Journal of Concurrency and Computation: Practice and Experience*, Note: *(Accepted for publication)*, 2013.
DOI: 10.1002/cpe.3025

328. [Zha2013b] Yubo Zhang, Kwan-Liu Ma. **"Fast Global Illumination for Interactive Volume Visualization,"** Proceedings of ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '13), 2013.

329. [Zha2013c] Yubo Zhang, Zhao Dong,, Kwan-Liu Ma. **"Real-time Volume Rendering in Dynamic Lighting Environments using Precomputed Photon Mapping,"** IEEE Transactions on Visualization and Computer Graphics, 2013.

330. [Zhe2013a] F. Zheng, H. Zou, G. Eisnhauer, K. Schwan, M. Wolf, J. Dayal, T. A. Nguyen, J. Cao, H. Abbasi, S. Klasky, N. Podhorszki, H. Yu. **"FlexIO: I/O middleware for Location-Flexible Scientific Data Analytics,"** submitted to IPDPS 2013, 2013.

331. [Zhe2013b] L. Zheng, Y. Wu,, K.-L. Ma. **"Perceptually Based Depth-Ordering Enhancement for Direct Volume Rendering,"** IEEE Transactions on Visualization and Computer Graphics, 19(3):446-459, 2013.

332. [Zhe2013c] Fang Zheng, Hongfeng Yu, Can Hantas, Matthew Wolf, Greg Eisenhauer, Karsten Schwan, Hasan Abbasi, Scott Klasky. **"GoldRush: resource efficient in situ scientific data analytics using fine-grained interference aware execution,"** In *Proceedings of SC13: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. Article 78. 2013.
DOI: 10.1145/2503210.2503279

333. [Zho2013a] L. Zhou, C.D. Hansen. **"Transfer Function Design based on User Selected Samples for Intuitive Multivariate Volume Exploration,"** In *Proceedings of the 2013 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 73--80. 2013.

2012

334. [Ahe2012a] Sean Ahern. **"The Road to Exascale,"** In *High Performance Visualization: Enabling Extreme-Scale Scientific Insight, Chapman & Hall/CRC Computational Sciencet*, Edited by E. Wes Bethel; Hank Childs, Lawrence Berkeley National Laboratory, Berkeley, California, USA; Charles Hansen, University of Utah, Salt Lake City, USA, Chapman & Hall/CRC Computational Science, October, 2012.

335. [Ben2012a] J. C. Bennett, H. Abbasi, P. Bremer, R. W. Grout, A. Gyulassy, T. Jin, S. Klasky, H. Kolla, M. Parashar, V. Pascucci, P. Pbay, D. Thompson, H. Yu, F. Zhang, J. Chen. **"Combining In-Situ and In-Transit Processing to Enable Extreme-Scale Scientific Analysis ,"** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, Salt Lake City, Utah, U.S.A., November, 2012.

336. [Bha2012b] A. Bhatele, T. Gamblin, S.H. Langer, P.-T. Bremer, E.W. Draeger, B. Hamann, K.E. Isaacs, A.G. Landge, J.A. Levine, V. Pascucci, M. Schulz, C.H. Still. **"Mapping applications with collectives over sub-communicators on torus networks,"** In *Proceedings of Supercomputing 2012*, pp. 1-11. 2012.
DOI: 10.1109/SC.2012.75

337. [Bro2012a] C. Brownlee, T. Fogal, C.D. Hansen. **"GLuRay: Ray Tracing in Scientific Visualization Applications using OpenGL Interception,"** In *Proceedings of the Eurographics Symposium on Parallel Graphics and Visualization (2012)*, Edited by H. Childs, T. Kuhlen, and F.

Marton, pp. 41--50. 2012.
DOI: 10.2312/EGPGV/EGPGV12/041-050

338. [Bro2012b] C. Brownlee, J. Patchett, L.-T. Lo, D. DeMarle, C. Mitchell, J. Ahrens, C.D. Hansen. **"A Study of Ray Tracing Large-scale Scientific Data in Parallel Visualization Applications,"** In *Proceedings of the Eurographics Symposium on Parallel Graphics and Visualization (2012)*, Edited by H. Childs, T. Kuhlen, and F. Marton, 2012.

339. [Byn2012a] Surendra Byna, Jerry Chou, Oliver Ruebel, Prabhat, Homa Karimabadi, William S. Daughton, Vadim Roytershteyn, E. Wes Bethel, Mark Howison, Ke-Jou Hsu, Kuan-Wu Lin, Arie Shoshani, Andrew Uselton,, Kesheng Wu. **"Parallel I/O, Analysis, and Visualization of a Trillion Particle Simulation,"** In *Proceedings of the international conference on Supercomputing*, 2012.

340. [Byn2012b] S. Byna, J. Chou, O. Rbel, Prabhat, H. Karimabadi, W. S. Daughton, V. Roytershteyn, E. W. Bethel, M. Howison, K-J. Hsu, K-W. Lin, A. Shoshani, A. Uselton,, K. Wu,. **"Parallel Data, Analysis, and Visualization of a Trillion Particles,"** Lightning Talk and Poster Presentation at the 6th Extremely Large Databases Conference (XLDB), 2012.

341. [Cam2012a] David Camp, Hank Childs, Christoph Garth, Dave Pugmire, Kenneth I. Joy. **"Parallel Stream Surface Computation for Large Data Sets,"** In *Proceedings of the IEEE Symposium on Large-Scale Data Anaysis and Visualization*, Seattle WA, USA, Note: *to appear*, 2012.

342. [Cha2012a] Abon Chaudhuri, Han-Wei Shen. **"A Self-Adaptive Technique for Visualizing Geospatial Data in 3D with Minimum Occlusion,"** In *Visualization and Data Analysis Conference in IS&/SPIE Symposium on Electronic Imaging*, Note: *Best Paper Award*, 2012.

343. [Cha2012b] Abon Chauduhri, Teng-Yok Lee, Cong Wang, Bo Zhou, Tian-Tian Xu, Han-Wei Shen, Tom Peterka, Yi-Jen Chiang. **"Scalable Computation of Distributions from Large Scale Data Sets,"** In *IEEE 2012 Large Data Analysis and Visualization* , 2012 , 2012.

344. [Cha2012c] Abon Chaudhuri, Teng-Yok Lee, Han-Wei Shen. **"Exploring Flow Fields Using Fractal Analysis of Field Lines,"** In *IEEE Visualization 2012 Poster Proceedings (Best Poster Award)* , 2012.

345. [Che2012a] Zhengzhang Chen, William Hendrix, Hang Guan, Isaac K. Tetteh, Alok Choudhary, Fredrick Semazzi, Nagiza F. Samatova. **"Discovery of extreme events-related communities in contrasting groups of physical system networks,"** In *Data Mining and Knowledge Discovery*, 2012.
DOI: 10.1007/s10618-012-0289-3

346. [Che2012b] F. Chen, H. Obermaier, H. Hagen, B. Hamann, J. Tierny, V. Pascucci.. **"Topology Analysis of Time-Dependent Multi-Fluid Data Using the Reeb Graph,"** In *Computer Aided Geometric Design*, Note: *Published Online Apr. 24*, Elsevier, 2012.
DOI: 10.1016/j.cagd.2012.03.019

347. [Chi2012a] H. Childs, E. Brugger, B. Whitlock, J. Meredith, S. Ahern, K. Bonnell, M. Miller, G. H. Weber, C. Harrison, D. Pugmire, T. Fogal, C. Garth, A. Sanderson, E. W. Bethel, M. Durant, D. Camp, J. M. Favre, O. Rbel, Paul Navrtil. **"VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data,"** In *High Performance Visualization: Enabling Extreme-Scale Scientific Insight*, Edited by E. Wes Bethel; Hank Childs, Lawrence Berkeley National Laboratory, Berkeley, California, USA; Charles Hansen, University of Utah, Salt Lake City, USA, Chapman & Hall/CRC Computational Science, October, 2012.

348. [Chi2012b] H. Childs, D. Pugmire, S. Ahern, B. Whitlock, M. Howison, Prabhat, G. Weber, E. W. Bethel. **"Visualization at Extreme Scale Concurrency,"** In *High Performance Visualization: Enabling Extreme-Scale Scientific Insight*, Edited by E. Wes Bethel; Hank Childs, Lawrence

Berkeley National Laboratory, Berkeley, California, USA; Charles Hansen, University of Utah, Salt Lake City, USA, Chapman & Hall/CRC Computational Science, October, 2012.

349. [Cho2012a] A.N.M. Imroz Choudhury, B. Wang, P. Rosen, V. Pascucci. **“Topological Analysis and Visualization of Cyclical Behavior in Memory Reference Traces,”** In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis 2012)*, pp. 9--16. 2012.
DOI: 10.1109/PacificVis.2012.6183557

350. [Chu2012a] Chun-Ming Chen, Lijie Xu, Teng-Yok Lee, Han-Wei Shen. **“A Flow Guided Layout for Out-of-Core Streamline Computation,”** In *Proceedings of IEEE Pacific Visualization*, 2012.

351. [Cor2012a] C. Correa, T. Crnovrsanin,, K.-L. Ma. **“ Visual Reasoning about Social Networks using Centrality Sensitivities,”** IEEE Transactions on Visualization and Computer Graphics, 18(1): 106-120, 2012.

352. [Fou2012a] Nathaniel Fout, Kwan-Liu Ma. **“Fuzzy Volume Rendering,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, Note: *(Also a Visualization 2012 Conference paper)*, pp. 2335-2344. 2012.

353. [Fou2012b] Nathaniel Fout, Kwan-Liu Ma. **“An Adaptive PRediction-Based Approach to Lossless Compression of Floating-Point Volume Data,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, Note: *(Also a Visualization 2012 Conference paper)*, pp. 2295-2304. 2012.

354. [Gai2012a] K. Gaither, H. Childs, K. Schulz, C. Harrison, W. Barth, D. Donzis, P. Yeung. **“Visual Analytics for Finding Critical Structures in Massive Time-Varying Turbulent-Flow Simulations,”** In *Computer Graphcs and Applications*, Vol. 32, No. 4, IEEE, pp. 34-45. July-Aug, 2012.

355. [Ger2012a] S. Gerber, O. Reubel, P.-T. Bremer, V. Pascucci, R.T. Whitaker. **“Morse-Smale Regression,”** In *Journal of Computational and Graphical Statistics*, Vol. 22, No. 1, pp. 193--214. 2012.
DOI: 10.1080/10618600.2012.657132

356. [Gon2012a] Z. Gong, S. Lakshminarasimhan, J. Jenkins, H. Kolla, S. Ethier, J. Chen, R. Ross, S. Klasky, N. Samatova. **“Multi-level layout optimizations for efficient spatio-temporal queries of ISABELA-compressed data,”** In *The 26th IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, May, 2012.

357. [Gon2012b] Zhenhuan Gong, Terry Rogers, John Jenkins, Hemanth Kolla, Stephane Ethier, Jackie Chen, Robert Ross, Scott Klasky, Nagiza F. Samatova. **“MLOC: Multi-level Layout Optimization Framework for Compressed Scientific Data Exploration with Heterogeneous Access Patterns,”** In *International Conference on Parallel Processing (ICPP)*, Pittsburgh, PA, September, 2012.

358. [Gro2012a] R. Grout, A. Gruber, H. Kolla, P.-T. Bremer, J. Bennett, A. Gyulassy, J. Chen. **“A direct numerical simulation study of turbulence and flame structure in transverse jets analysed in jet-trajectory based coordinates,”** In *Journal of Fluid Mechanics 706* , pp. 351-383 . 2012.

359. [Gyu2012a] A. Gyulassy, V. Pascucci, T. Peterka, R. Ross. **“The Parallel Computation of Morse-Smale Complexes,”** In *Parallel & Distributed Processing Symposium (IPDPS)*, pp. 484--495. 2012.
DOI: 10.1109/IPDPS.2012.52

360. [Gyu2012b] A. Gyulassy, P.-T. Bremer, V. Pascucci. **“Computing Morse-Smale Complexes with Accurate Geometry,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, pp. 2014--2022. 2012.
DOI: 10.1109/TVCG.2011.272

361. [Gyu2012c] A. Gyulassy, N. Kotava, M. Kim, C. Hansen, H. Hagen,, V. Pascucci. **“Direct Feature Visualization Using Morse-Smale Complexes,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 9, pp. 1549--1562. Sept., 2012.
DOI: 10.1109/TVCG.2011.272

362. [Hab2012a] Habib, S., Morozov, V., Finkel, H., Pope, A., Heitmann, K., Kumaran, K., Peterka, T., Insley, J., Daniel, D., Fasel, P., Frontiere, N., Lukic, Z.. **“The Universe at Extreme Scale: Multi-Petaflop Sky Simulation on the BG/Q,”** In *Proceedings of SC12*, 2012.

363. [Har2012a] C. Harrison, H. Krishnan. **“Python's role in VisIt,”** In the Proceedings of the eleventh annual Scientific Computing with Python Conference, (SciPy 2012), July, 2012.

364. [Har2012b] C. Harrison, P. A. Navrtil, M. Moussalem, M. Jiang, H. Childs. **“Efficient Dynamic Derived Field Generation on Many-Core Architectures,”** In *Proceedings of Workshop on Python for High Performance and Scientific Computing*, (PyHPC 2012), November, 2012.

365. [Hen2012a] William Hendrix, Md. Mostofa Ali Patwary, Ankit Agrawal, Wei-keng Liao, Alok Choudhary. **“Parallel Hierarchical Clustering on Shared Memory Platforms,”** In *19th Annual International Conference on High Performance Computing*, Pune, India, 2012.

366. [Hsu2012a] Wei-Hsien Hsu, Chun-Fu Wang, Kwan-Liu Ma, Hongfeng Yu,, Jacqueline H. Chen. **“A Job Scheduling Design for Visualization Services using GPU Clusters,”** In *Proceedings of IEEE Cluster 2012*, pp. 523-533. September, 2012.

367. [Jad2012a] S. Jadhav, H. Bhatia, P.-T. Bremer, J.A. Levine, L.G. Nonato, V. Pascucci. **“Consistent Approximation of Local Flow Behavior for 2D Vector Fields,”** In *Topological Methods in Data Analysis and Visualization II. Mathematics and Visualization*, No. 3, pp. 141--159. 2012.
DOI: 10.1007/978-3-642-23175-9_10

368. [Jen2012a] John Jenkins, Isha Arkatkar, Sriram Lakshminarasimhan, Neil Shah, Eric R. Schendel, Stephane Ethier, CS Chang, Jackie Chen, Hemanth Kolla, Scott Klasky, Robert Ross,, Nagiza F. Samatova. **“Analytics-driven Lossless Data Compression for Rapid In-situ Indexing, Storing, and Querying,”** In *Proceedings of the 23rd International Conference on Database and Expert Systems Applications (DEXA)*, September, 2012.

369. [Jen2012b] John Jenkins, James Dinan, Pavan Balaji, Nagiza F. Samatova, Rajeev Thakur. **“Enabling Fast, Noncontiguous GPU Data Movement in Hybrid MPI+GPU Environments,”** In *IEEE International Conference on Cluster Computing (Cluster)*, Beijing, China, September, 2012.

370. [Jen2012c] John Jenkins, Eric Schendel, Sriram Lakshminarasimhan, David A. Boyuka III, Terry Rogers, Stephane Ethier, Robert Ross, Scott Klasky, Nagiza, F. Samatova. **“Byte-precision Level of Detail Processing for Variable Precision Analysis,”** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, Salt Lake City, Utah, U.S.A., November, 2012.

371. [Jia2012b] F. Jiao, J.M. Phillips, Y. Gur, C.R. Johnson. **“Uncertainty Visualization in HARDI based on Ensembles of ODFs,”** In *Proceedings of the 5th IEEE Pacific Visualization Symposium (PacificVis 2012)*, pp. 193--200. 2012.

372. [Jin2012a] T. Jin, F. Zhang, M. Parashar, S. Klasky, N. Podhorszki, H. Abbasi. **“A Scalable Messaging System for Accelerating Discovery from Large Scale Scientific Simulations,”** In *IEEE International Conference on High Performance Computing (HiPC)*, December, 2012.

373. [Joh2012a] C.R. Johnson. **“Biomedical Visual Computing: Case Studies and Challenges,”** In *IEEE Computing in Science and Engineering*, Vol. 14, No. 1, pp. 12--21. 2012.

374. [Ker2012a] Thomas Kerwin, D. Stredney, G. Wiet, Han-Wei Shen. **“Virtual Mastoidectomy performance evaluation through multi-volume analysis,”** In *Computer Assisted Radiology and Surgery (CARS)*, 2012.

375. [Kim2012a] M. Kim, G. Chen, C.D. Hansen. **“Dynamic particle system for mesh extraction on the GPU,”** In *Proceedings of the 5th Annual Workshop on General Purpose Processing with Graphics Processing Units*, London, England, pp. 38--46. 2012.
DOI: 10.1145/2159430.215943

376. [Kne2012a] J. Knezevic, R.-P. Mundani, E. Rank, A. Khan, C.R. Johnson. **“Extending the SCIRun Problem Solving Environment to Large-Scale Applications,”** In *Applied Computing*, Madrid, Spain, *IADIS*, pp. 171--178. 2012.

377. [Kri2012a] Hari Krishnan, Christoph Garth, Jens Guhring, M. Akif Gulsun, Andreas Greiser,, Kenneth I. Joy. **“Analysis of Time-Dependent Flow-Sensitive PC-MRI Data,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 6, pp. 966--977. June, 2012.

378. [Kum2012a] S. Kumar, V. Vishwanath, P. Carns, J.A. Levine, R. Latham, G. Scorzelli, H. Kolla, R. Grout, R. Ross, M.E. Papka. **“Efficient data restructuring and aggregation for I/O acceleration in PIDX,”** In *Proceedings of the ACM/IEEE Conference on High Performance Computing, Networking, Storage and Analysis (SC)*, IEEE Computer Society Press, pp. 50:1--50:11. 2012.
ISBN: 978-1-4673-0804-5

379. [Lak2012a] Sriram Lakshminarasimhan, Prabhat Kumar, Wei-keng Liao, Alok Choudhary, Vipin Kumar,, Nagiza F. Samatova. **“On the Path to Sustainable, Scalable, and Energy-efficient Data Analytics: Challenges, Promises, and Future Directions,”** In *Proceedings of the 2012 International Green Computing Conference (IGCC)*, San Jose, CA, U.S.A., Note: *Invited paper*, June, 2012.

380. [Lak2012b] Sriram Lakshminarasimhan, Neil Shah, Stephane Ethier, Seung-Hoe Ku, C.S. Chang, Scott Klasky, Rob Latham, Rob Ross, Nagiza F. Samatova. **“ISABELA for Effective In-situ Compression of Scientific Data,”** In *Concurrency and Computation: Practice and Experience*, Note: *Invited paper*, 2012.

381. [Lan2012a] A.G. Landge, J.A. Levine, A. Bhatele, K.E. Isaacs, T. Gamblin, S. Langer, M. Schulz, P.-T. Bremer, V. Pascucci. **“Visualizing Network Traffic to Understand the Performance of Massively Parallel Simulations,”** In *IEEE Symposium on Information Visualization (INFOVIS'12)*, Seattle, WA, Vol. 18, No. 12, pp. 2467--2476. 2012.
DOI: 10.1109/TVCG.2012.286

382. [Lat2012a] R. Latham, C. Daley, W. Liao, K. Gao, R. Ross, A. Dubey, A. Choudhary. **“A case study for scientific I/O: improving the FLASH astrophysics code,”** In *Computational Science & Discovery*, Vol. 5, No. 1, pp. 015001. 2012.

383. [Li2012a] Z. Li, S. Wang, J. Yu,, K.-L. Ma. **“Restoration of Brick and Stone Relief from Single Rubbing Images,”** IEEE Transactions on Visualization and Computer Graphics. 18(2):177-187, 2012.

384. [Liu2012a] Shusen Liu, Joshua A. Levine, Peer Timo Bremer, Valerio Pascucci. **“Gaussian Mixture Model Based Volume Rendering ,”** 2012 IEEE Symposium on Large Data Analysis and Visualization (LDAV), Note: *Best Paper Honorable Mention*, pp. 73--77. 2012.
DOI: 10.1109/LDAV.2012.6378978

385. [Lo2012a] Li-ta Lo, Christopher Sewell,, James Ahrens. **“PISTON: A Portable Cross-Platform Framework for Data-Parallel Visualization Operators,”** In *EGPGV.12: Proceedings of the Eurographics Symposium on Parallel Graphics and Visualization 2012*, May, 2012.

386. [Luo2012a] S.-J. Luo, C.-L. Liu, B.-Y. Chen,, K.-L. Ma. **“Ambiguity-Free Edge-Bundling for Interactive Graph Visualization,”** IEEE Transactions on Visualization and Computer Graphics, 18(5):810-821, 2012.

387. [Ma2012a] J. Ma, I. Liao, J. Frazier,, K.-L. Ma. **“Living Liquid: Design and Evaluation of an Exploratory Visualization Tool for Museum Visitors,”** IEEE Transactions on Visualization and Computer Graphics, 18(12):2799-2808, 2012.

388. [Ma2012b] K.-L. Ma et al.. **“Scientific Storytelling using Visualization,”** IEEE Computer Graphics and Applications, 32(1):12-19, Jan/Feb, 2012.

389. [Maj2012a] Joyce Ma, Isaac Liao, Jennifer Frazier, Kwan-Liu Ma. **“Living Liquid: Design and Evaluation of an Exploratory Visualization Tool for Museum Visitors,”** Vol. 18, No. 12, Note: *(Also an InfoVis 2012 Conference paper)*, pp. 2799-2808. 2012.

390. [Mar2012a] Steve Martin, Han-Wei Shen . **“Interactive Transfer Function Design on Large Multiresolution Volumes,”** In *IEEE 2012 Large Data Analysis and Visualization* , IEEE, 2012.

391. [Mar2012b] T. Martin, G. Chen, S. Musuvathy, E. Cohen, C.D. Hansen. **“Generalized Swept Mid-structure for Polygonal Models,”** In *Proceedings of Eurographics 2012*, Vol. 31, No. 2, pp. 805--814. 2012.
DOI: 10.1111/j.1467-8659.2012.03061.x

392. [Mcl2012a] W. C. McLendon, G. Bansal, P.-T. Bremer, J. Chen, H. Kolla, J. Bennett. **“On The Use of Graph Search Techniques for The Analysis of Extreme-scale Combustionc Simulation Data,”** Proceedings of IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV), 2012.

393. [Mer2012a] Jeremy S. Meredith, Sean Ahern, David Pugmire, Robert Sisneros. **“EAVL: The Extreme-scale Analysis and Visualization Library,”** In *Proceedings of the Eurographics Symposium on Parallel Graphics and Visualization (EGPGV)*, 2012.

394. [Mer2012b] J. Meredith, R. Sisneros, D. Pugmire, S. Ahern. **“A Distributed Data-Parallel Framework for Analysis and Visualization Algorithm Development,”** In *Fifth Workshop on General Purpose Processing on Graphics Processing Units (GPGPU)*, ACM Architectural Support for Programming Languages and Operating Systems, March, 2012.

395. [Mor2012a] Kenneth Moreland. **“Redirecting Research in Large-Format Displays for Visualization,”** In *Proceedings of the IEEE Symposium on Large-Scale Data Anaysis and Visualization*, pp. 91--95. October, 2012.

396. [Mor2012b] Kenneth Moreland, Brad King, Robert Maynard, Kwan-Liu Ma. **“Flexible Analysis Software for Emerging Architectures,”** In *Petascale Data Analytics: Challenges and Opportunities (PDAC-12)*, November, 2012.

397. [Mor2012c] Kenneth Moreland, Jeremy Meredith, Berk Geveci. **“Enabling Production-Quality Scientific-Discovery Tools with Data and Execution Models,”** No. SAND 2012-10796P, *Sandia National Laboratories*, December, 2012.

398. [Mor2012d] Kenneth Moreland. **“Oh, $#*@! Exascale! The Effect of Emerging Architectures on Scientific Discovery,”** In *2012 SC Companion (Proceedings of the Ultrascale Visualization Workshop)*, pp. 224--231. November, 2012.
DOI: 10.1109/SC.Companion.2012.38

399. [Nav2012a] P. Navratil, B. Barth,, H. Childs. **“Virtual Rheoscopic Fluids for Dense, Large-Scale Fluid Flow Visualizations.,”** In *Proceedings of IEEE Symposium on Large Data Analysis and Visualization (LDAV),*, pp. 79-84. October, 2012.

400. [Nou2012a] B. Nouanesengsy, T.-Y. Lee, K. Lu, H.-W. Shen,, T. Peterka. **“Parallel Particle Advection and FTLE Computation for Time-Varying Flow Fields,”** SC 12: ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2012, 2012.

401. [Obe2012a] Harald Obermaier, Kenneth I. Joy. **"Derived Metric Tensors for Flow Surface Visualization,"** In *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, Vol. 18, No. 12, pp. 2149-2158. Oct, 2012.

402. [Obe2012b] Harald Obermaier, Kenneth I. Joy. **"Function Field Analysis for the Visualization of Flow Similiarity in Time-Varying Vector Fields,"** In *LNCS 7432, Part II (Proc of ISVC 2012)*, Springer Heidelberg, pp. 253-264. 2012.

403. [Obe2012c] Harald Obermaier, Fang Chen, Hans Hagen, Kenneth I. Joy. **"Visualization of Material Interface Stability,"** In *Proceedings of Pacific Visualization Conference*, Songdo, Korea, pp. 225--232. 2012.

404. [Obe2012d] Harald Obermaier, Magali I. Billen, Kenneth I. Joy, Hans Hagen, Martin Hering-Bertram.. **"Visualization and multivariate clustering of scattered moment tensors,"** In *Information Visualization*, Vol. 11, No. 1, Sage, pp. 43-59. 2012.

405. [Oze2012a] Sedat Ozer, Jishang Wei, Deborah Silver, Kwan-Liu Ma,, Pino Martin. **"Group Dynamics in Scientific Visualization,"** Proceedings of IEEE Symposium on Large Data Analysis and Visualization (LDAV '12), pp. 97-104. October, 2012.

406. [Pas2012a] V. Pascucci, G. Scorzelli, B. Summa, P.-T. Bremer, A. Gyulassy, C. Christensen, S. Philip, S. Kumar. **"The ViSUS Visualization Framework,"** In *High Performance Visualization: Enabling Extreme-Scale Scientific Insight*, Chapman & Hall/CRC Computational Science, Ch. 19, Edited by E. Wes Bethel; Hank Childs, Lawrence Berkeley National Laboratory, Berkeley, California, USA; Charles Hansen, University of Utah, Salt Lake City, USA, Chapman & Hall/CRC Computational Science, 2012.

407. [Pat2012a] Md. Mostofa Ali Patwary, Diana Palsetia, Ankit Agrawal, Wei-keng Liao, Fredrik Manne, Alok Choudhary. **"A New Scalable Parallel DBSCAN Algorithm Using the Disjoint Set Data Structure,"** In *ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, Salt Lake City, Utah, U.S.A., 2012.

408. [Pet2012a] Tom Peterka, Robert Ross. **"Versatile Communication Algorithms for Data Analysis,"** In *EuroMPI Special Session on Improving MPI User and Developer Interaction IMUDI'12*, Vienna, Austria, 2012.

409. [Pet2012b] Peterka, T., Kwan, J., Pope, A., Finkel, H., Heitmann, K., Habib, S., Wang, J., Zagaris, G.. **"Meshing the Universe: Integrating Analysis in Cosmological Simulations,"** In *Proceedings of the SC12 Ultrascale Visualization Workshop*, Salt Lake City, UT, 2012.

410. [Pet2012c] Peterka, T., Ma, K.-L. **"Parallel Image Compositing Methods, book chapter in High Performance Visualization,"** Edited by Bethel, E. W., Childs, H., Hansen, C., 2012.

411. [Pot2012a] K. Potter, R.M. Kirby, D. Xiu, C.R. Johnson. **"Interactive visualization of probability and cumulative density functions,"** In *International Journal of Uncertainty Quantification*, Vol. 2, No. 4, pp. 397--412. 2012.
DOI: 10.1615/Int.J.UncertaintyQuantification.2012004074

412. [Pot2012b] K. Potter, P. Rosen, C.R. Johnson. **"From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches,"** In *Uncertainty Quantification in Scientific Computing*, IFIP Advances in Information and Communication Technology Series, Vol. 377, Edited by Andrew Dienstfrey and Ronald Boisvert, Springer, pp. 226--249. 2012.
DOI: 10.1007/978-3-642-32677-6_15

413. [Pou2012a] Elaheh Pourabbas, Arie Shoshani,, Kesheng Wu. **"Minimizing Index Size by Reordering Rows and Columns,"** In *International Conference on Scientific and Statistical Database Management (SSDBM)*, June, 2012.

414. [Ras2012a] Michel Rasquin, Patrick Marion, Venkatram Vishwanath, Ray Loy, Andrew Bauer, Benjamin Matthews, Min Zhou, Onkar Sahni, Jing Fu, Ning Liu, Christopher Carothers, Mark S.

Shephard, Mark Hereld, Michael E. Papka, Kalyan Kumaran, Berk Geveci,, Kenneth Jansen. **“Interactive visualization from a live unstructured grid CFD simulation at 160k cores,”** 2012 SIAM Conference on Parallel Processing for Scientific Computing. SIAM, 2012, 2012.

415. [Rec2012a] Shawn Recker, Mauricio Hess-Flores, Mark A. Duchaineau, Kenneth I. Joy. **“Visualization of Scene Structure Uncertainity in a Multi-View Reconstruction Pipeline,”** In *Proceedings of Vision, Modeling and Visualization Conference*, Note: *to appear*, 2012.

416. [Sal2012a] A. Sallaberry, C. Muelder,, K.-L. Ma. **“Clustering, Visualizing, and Navigating for Large Dynamic Graphs,”** Proceedings of Graph Drawing 2012, pp. 487-498. September 19-21,, 2012.

417. [San2012a] A.R. Sanderson, B. Whitlock, O. Ruebel, H. Childs, G.H. Weber, Prabhat, K. Wu. **“A System for Query Based Analysis and Visualization,”** In *Proceedings of the Third International Eurovis Workshop on Visual Analytics EuroVA 2012*, pp. 25--29. June, 2012.

418. [Sch2012a] E.R. Schendel, S.V. Pendse, J. Jenkins, D.A. Boyuka II, Z. Gong, S. Lakshminarasimhan, Q. Liu, S. Klasky, R. Ross, N.F. Samatova. **“ISOBAR Hybrid Compression-I/O Interleaving for Large-scale Parallel I/O Optimization,”** In *21st International ACM Symposium on High Performance Parallel and Distributed Computing*, June, 2012.

419. [Sch2012b] E.R. Schendel, Y. Jin, N. Shah, J. Chen, C.S. Chang, S.-H. Ku, S. Ethier, S. Klasky, R. Latham, R. Ross, N. Samatova. **“ISOBAR Preconditioner for Effective and High-throughput Lossless Data Compression,”** In *28th annual IEEE International Conference on Data Engineering (ICDE)*, 2012.

420. [Sch2012c] Simon Schroeder, John A. Peterson, Harald Obermaier, Louise H. Kellogg, Kenneth I. Joy, Hans Hagen. **“Visualization of Flow Behavior in Earth Mantle Convection,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, pp. 2198-2207. Oct, 2012.

421. [Sch2012d] M. Schott, T. Martin, A.V.P. Grosset, C. Brownlee, Thomas Hollt, B.P. Brown, S.T. Smith, C.D. Hansen. **“Combined Surface and Volumetric Occlusion Shading,”** In *Proceedings of Pacific Vis 2012*, pp. 169--176. 2012.
DOI: 10.1109/PacificVis.2012.6183588

422. [Sew2012a] Christopher Sewell, Jeremy Meredith, Kenneth Moreland, Tom Peterka, Dave DeMarle, Li-ta Lo, James Ahrens, Robert Maynard,, Berk Geveci. **“The SDAV Software Frameworks for Visualization and Analysis on Next-Generation Multi-Core and Many-Core Architectures,”** In *2012 SC Companion (Proceedings of the Workshop on Ultrascale Visualization)*, Salt Lake City, UT, pp. 206--214. November, 2012.
DOI: 10.1109/SC.Companion.2012.36

423. [Sha2012a] Neil Shah, Eric R. Schendel, Sriram Lakshminarasimhan, Saurabh V. Pendse, Terry Rogers, Nagiza F. Samatova. **“Improving I/O Throughput with PRIMACY: Preconditioning ID-Mapper for Compressing Incompressibility,”** In *IEEE International Conference on Cluster Computing (Cluster)*, Beijing, China, September, 2012.

424. [She2012a] Zeqian Shen, Jishang Wei, Neel Sundaresan,, Kwan-Liu Ma. **“Visual Analysis of Massive Web Session Data,”** Proceedings of IEEE Symposium on Large Data Analysis and Visualization (LDAV), pp. 65-72. October, 2012.

425. [Sum2012a] Brian Summa, Julien Tierny, Valerio Pascucci. **“Panorama weaving: fast and flexible seam processing,”** In *ACM Trans. Graph*, Vol. 31, No. 4, Note: *ACM ID:2335434*, pp. 83:1--83:11. July, 2012.
ISSN: 0730-0301
DOI: 10.1145/2185520.2185579

426. [Tah2012a] Yuzuru Tanahashi, Kwan-Liu Ma. **“Design Considerations for Optimizing Storyline Visualizations,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, Note: *(Also an InfoVis 2012 Conference paper)*, pp. 2679-2688. 2012.

427. [Tan2012a] Wei Tang, Narayan Desai, Venkatram Vishwanath, Daniel Buettner, Zhilin Lan. **“Multi-domain job coscheduling for leadership computing systems,”** The Journal of Supercomputing, Springer, 2012.
    ISSN: 0920-8542, 2012
    DOI: 10.1007/s11227-012-0741-6

428. [Tan2012b] Y. Tananhashi, K.-L. Ma. **“Design Considerations for Optimizing Storyline Visualizations,”** IEEE Transactions on Visualization and Computer Graphics 18(12):2679-2688 (InfoVis), 2012.

429. [Tan2012c] Yuzuru Tanahashi, James R. Rowland, Stephen North, Kwan-Liu Ma. **“Inferring Human Mobility Patterns from Anonymized Mobile Communication Usage,”** Proceedings of MoMM 2012, pp. 151-160. December 3-5, 2012.

430. [Tia2012a] Y. Tian, S. Klasky, W. Yu, H. Abbasi, B. Wang, N. Podhorszki, R. Grout, M. Wolf. **“A System-Aware Optimized Data Organization for Efficient Scientific Analytics,”** In *20th IEEE International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems (Mascots)*, 2012.

431. [Tia2012b] Y. Tian, S. Klasky, W. Yu, H. Abbasi, B. Wang, N. Podhorszki, R. Grout, M. Wolf. **“Two-Level Data Organization for Efficient Scientific Analytics,”** In *21st International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, 2012.

432. [Ton2012a] Xin Tong, Teng-Yok Lee, Han-Wei Shen. **“ Salient Time Steps Selection from Large Scale Time-Varying Data Sets with Dynamic Time Warping,”** In *IEEE 2012 Large Data Analysis and Visualization* , IEEE, 2012.

433. [Ush2012a] Daniela M. Ushizima, Dmitriy Morozov, Gunther H. Weber, Andrea G. C. Bianchi, James A. Sethian, E. Wes Bethel, Member. **“Augmented Topological Descriptors of Pore Networks for Material Science,”** In *IEEE Transactions on Visualization and Computer Graphcis*, Vol. 18, No. 12, pp. 2041-2050. 2012.

434. [Vat2012a] Ranga Raju Vatsavai. **“Scalable Multi-Instance Learning Approach for Mapping the Slums of the World,”** Proceedings of the SC workshop on Petascale Data Analytics: Challenges and Opportunities, IEEE, 2012.

435. [Vat2012b] Ranga Raju Vatsavai, Auroop Ganguly, Varun Chandola, Anthony Stefanidis, Scott Klasky, Shashi Shekhar. **“Spatiotemporal data mining in the era of big spatial data: algorithms and applications,”** Proceedings of the 1st ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data (BigSaptial12), ACM, 2012.

436. [Vat2012c] Ranga Raju Vatsavai, Budhendra Bhaduri. **“Big Spatiotemporal Data Analytics: Recent Advances and Future Challenges. (white paper) ,”** CRA/CCC/NSF Sponsored: From GPS and Virtual Globes to Spatial Computing-2020, 2012.

437. [Wan2012a] Y. Wan, H. Otsuna, C.-B. Chien, C.D. Hansen. **“FluoRender: An Application of 2D Image Space Methods for 3D and 4D Confocal Microscopy Data Visualization in Neurobiology Research,”** In *Proceedings of Pacific Vis 2012*, Incheon, Korea, pp. 201--208. 2012.
    DOI: 10.1109/PacificVis.2012.6183592

438. [Wan2012b] Y. Wan, H. Otsuna, C.-B. Chien, C.D. Hansen. **“Interactive Extraction of Neural Structures with User-Guided Morphological Diffusion,”** In *Proceedings of the IEEE Symposium on Biological Data Visualization*, pp. 1--8. 2012.
    DOI: 10.1109/BioVis.2012.6378577

439. [Wan2012c] Y. Wan, A.K. Lewis, M. Colasanto, M. van Langeveld, G. Kardon, C.D. Hansen. **“A Practical Workflow for Making Anatomical Atlases in Biological Research,”** In *IEEE Computer Graphics and Applications*, Vol. 32, No. 5, pp. 70--80. 2012.
DOI: 10.1109/MCG.2012.64

440. [Web2012a] Gunther H. Weber, Hank Childs, Jeremy S. Meredith. **“Efficient Parallel Extraction of Crack-free Isosurfaces from Adaptive Mesh Refinement (AMR) Data,”** In *Proceedings of the IEEE Symposium on Large-Scale Data Anaysis and Visualization*, Seattle WA, USA, 2012.

441. [Wei2012a] J. Wei, Z. Shen, N. Sundaresan,, K-L. Ma. **“Visual Cluster Exploration of Web Clickstream Data,”** In Proceedings of the Visual Analytics Science and Technology Conference (VAST 2012), pp. 3-12. October, 2012.

442. [Wei2012b] J. Wei, H. Yu, R. W. Grout, J. Chen,, K.-L. Ma. **“Visual Analysis of Particle Behaviors to Understand Combustion Simulations,”** IEEE Computer Graphics and Applications, 32(1):12-19, 2012.

443. [Wid2012a] W. Widanagamaachchi, C. Christensen, P.-T. Bremer, V. Pascucci. **“Interactive Exploration of Large-scale Time-varying Data using Dynamic Tracking Graphs,”** In *2012 IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, Note: *Best Paper Award*, pp. 9--17. 2012.
DOI: 10.1109/LDAV.2012.6378962

444. [Won2012a] P.C. Wong, H.-W. Shen, C.R. Johnson, C. Chen, R.B. Ross. **“The Top 10 Challenges in Extreme-Scale Visual Analytics,”** In *IEEE Computer Graphics and Applications*, Visualization Viewpoints, Edited by Theresa-Marie Rhyne, pp. 63--67. July-August, 2012.
DOI: 10.1109/MCG.2012.87

445. [Won2012b] Pak Chung Wong, Han-Wei Shen, Valerio Pascucci. **“Extreme Scale Visual Analytics,”** In *IEEE Computer Graphics and Applications*, IEEE, July, 2012.

446. [Wu2012a] Yingcai Wu, Guo-Xun Yuan, Kwan-Liu Ma. **“Visualizing Flow of Uncertainty through Analytical Processes,”** In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 12, Note: *(Also an InfoVis 2012 Conference paper)*, pp. 2526-2535. 2012.

447. [Yua2012a] Guo-Xun Yuan, Kwan-Liu Ma. **“Scalable Traning of Sparse Linear SVM,”** In *Proceedings of IEEE ICDM 2012*, Note: *(Accepted for publication)*, December, 2012.

448. [Zha2012a] F. Zhang, C. Docan, M. Parashar, S. Klasky, N. Podhorszki, H. Abbasi. **“Enabling In-situ Execution of Coupled Scientific Workflow on Multi-core Platform,”** In *26th IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, May, 2012.

449. [Zha2012b] F. Zhang, S. Lasluisa, T. Jin, I. Rodero, H. Bui, M. Parashar. **“In-situ Feature-based Objects Tracking for Large-Scale Scientific Simulations,”** In *1st International Workshop on Data-Intensive Scalable Computing Systems (DISCS)*, November, 2012.

450. [Zhe2012a] Lin Zheng, Yingcai Wu, Kwan-Liu Ma. **“Perceptually Based Depth-Ordering Enhancement for Direct Volume Rendering,”** In *IEEE Transactions on Visualization and Computer Graphics*, Note: *(Accepted for publication)*, 2012.
ISSN: 1077-2626
DOI: http://dx.doi.org/10.1109/TVCG.2012.144

451. [Zho2012a] L. Zhou, M. Schott, C.D. Hansen. **“Transfer Function Combinations,”** In *Computers and Graphics*, Vol. 36, No. 6, pp. 596--606. 2012.
DOI: 10.1016/j.cag.2012.02.007