# The Open Source DataTurbine Initiative: Streaming Data Middleware for Environmental Observing Systems

T. Fountain [a,*], S. Tilak [a], P. Shin [a], P. Hubbard[a], L. Freudinger[b]

[a] University of California, San Diego, CA 92093, USA – (tfountain, stilak, pshin, hubbard)@ucsd.edu
[b] NASA Dryden Flight Research Center, Edwards, CA 93523, USA – Lawrence.C.Freudinger@nasa.gov

**Abstract – The Open Source DataTurbine Initiative is an international community of scientists and engineers sharing a common interest in real-time streaming data middleware and applications. The technology base of the OSDT Initiative is the DataTurbine open source middleware. Key applications of DataTurbine include coral reef monitoring, lake monitoring and limnology, biodiversity and animal tracking, structural health monitoring and earthquake engineering, airborne environmental monitoring, and environmental sustainability. DataTurbine software emerged as a commercial product in the 1990's from collaborations between NASA and private industry. In October 2007, a grant from the USA National Science Foundation (NSF) Office of Cyberinfrastructure allowed us to transition DataTurbine from a proprietary software product into an open source software initiative. This paper describes the DataTurbine software and highlights key applications in environmental monitoring.**

**Keywords:** earth observing systems; cyberinfrastructure; streaming data middleware; sensor networks

## 1. INTRODUCTION

The Open Source DataTurbine (OSDT) Initiative is an international community of scientists and engineers who share a common interest in real-time streaming data middleware and applications (www.dataturbine.org). Community members are drawn from academia and industry, and represent a variety of science and engineering domains, from ecology to aerospace. The technology base of the OSDT Initiative is the DataTurbine open source middleware. Key applications of DataTubine include coral reef monitoring [CREON], lake monitoring and limnology [GLEON], biodiversity and animal tracking [MoveBank], structural health monitoring [AHML] and earthquake engineering [NEES], environmental sustainability [UCSD-ESI], and airborne environmental monitoring [Freudinger09].

The origins of DataTurbine extend back to early collaborations between NASA and Creare, Inc. in 1985. DataTurbine was originally developed by Creare Inc, an engineering consulting firm in Hanover, New Hampshire [CREARE]. Creare's primary line of business involves consultation and contract software development for science and engineering applications. DataTurbine was a successful commercial streaming data product with a track record of performance in NSF and NASA projects, and also applications in private industry. The evolution of DataTurbine coincided with advances in sensing and communications technologies and a desire by the science and engineering communities to deploy real-world large-scale sensor networks and environmental observing systems. DataTurbine was developed as a generic streaming data middleware for real-time data acquisition systems, independent from a specific application niche.

After years of collaboration, and months of negotiation, in a quest to unlock DataTurbine's full potential, executives at Creare Inc. signed a letter of intent to release DataTurbine as an open-source software product in collaboration with UCSD. In October 2007, a grant from the US National Science Foundation (NSF) Office of Cyberinfrastructure allowed us to transition DataTurbine from a proprietary software product into an open source software initiative. This paper describes the DataTurbine software and highlights key applications in real-time environmental monitoring.

## 2. FROM MIDDLEWARE TO SOFTWARE INITIATIVE

Environmental science and engineering communities are now actively engaged in the early planning and development phases of the next generation of large-scale sensor-based observing systems. These systems face two significant challenges: *heterogeneity of instrumentation* and *complexity of data stream processing*. Environmental observing systems are complex distributed systems. They incorporate instruments from across the spectrum of complexity, from temperature sensors to acoustic Doppler current profilers, to streaming video cameras, and to synthetic aperture radar. They operate under a variety of networking

* Corresponding author: Tony Fountain, UCSD.
** This work was partially supported by a grant from the Gordon and Betty Moore Foundation and the following NSF grants OCI Award 0722067, BD&I 0756920, OISE 0314015, 0627026, 0446802, 0446017, 0446298, and DBI 0639229. The work was partially supported via NASA Airborne Science Program, WBS 769134.03.02.03.02

conditions, including wired and wireless, persistent and intermittent. They have stringent requirements on data timeliness and integrity. Managing these instruments and their data streams presents serious challenges in systems development and operations. The Open Source DataTurbine (OSDT) Initiative was launched in October 2007 with a two-year grant from the National Science Foundation Office of Cyberinfrastructure (award #OCI-0722067) to address these challenges through the publication, enhancement, and promotion of the DataTurbine streaming data middleware [Tilak07].
The NSF award funded the core activities needed to build an open-source software community around the DataTurbine middleware. There were three areas of funded activities: (1). Publish DataTurbine as an open source software product and provide developer support, including documentation, bug tracking, collaboration tools, and experimental facilities. (2). Enhance the code base, including porting DataTurbine to additional compute platforms, writing additional device drivers, and testing and tuning. (3). Build an active open source community through education, outreach, recruitment, and technical support. The OSDT Initiative has been successful in these activities.
The result is an international community of scientists and engineers who share a common interest in real-time streaming data middleware and applications and are collaborating to produce useful middleware and successful deployments (www.dataturbine.org). Community members are drawn from academia and industry, and represent a variety of science and engineering domains, including PRAGMA [PRAGMA], GLEON [GLEON], CREON [CREON], MoveBank [MoveBank], CUAHSI [CUASHI], LTER Network [LTER], NEES [NEES], NCEAS [NCEAS], and GBROOS [GBROOS]. Only 15 months since inception, the OSDT Initiative has demonstrated broad impact on a variety of projects and communities, across a wide range of applications – from lakes and coral reefs, to civil infrastructure and smart buildings, to airborne science and aeronautics [Benson09, Fountain09, OSDT-Report, OSDT-Workshop].

From the perspective of distributed systems, the DataTurbine middleware is a "black box" to which applications and devices send and receive data (Figure 1). DataTurbine handles all data management operations between data sources and sinks, including reliable transport, routing, scheduling, and security. DataTurbine accomplishes this through the innovative use of flexible net bus objects combined with memory and file-based ring buffers. Network bus objects perform data stream multiplexing and routing. Ring buffers provide tunable persistent storage at key network nodes to facilitate reliable data transport. Ring buffers also connect directly to client applications to provide TiVo-like services including data stream subscription, capture, rewind, and replay. This presents client applications with a simple, uniform interface to real-time and historical (playback) data. Since DataTurbine is implemented in the Java programming language it is platform independent. It has been demonstrated to run efficiently on platforms from cell phones to supercomputers [Tilak07].

**2.1 Related research:** DataTurbine shares some features with other existing data management systems; however DataTurbine is unique in its support for science and engineering applications. Commercial programs such as MSMQ [MSMQ] and Websphere MQ [MQ], NaradaBrokering [Pallickara03] and similar standards including Enterprise messaging systems [EMS], Enterprise Service Bus [Chappell04], Java Message Service [JMS], CORBA [CORBA], and various publish-subscribe systems [Liu03] provide

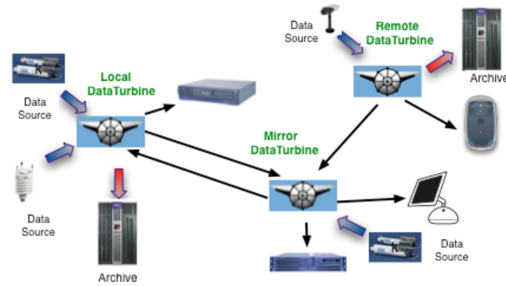


**Figure 1: A network of DataTurbine servers with sources and destinations distributed through the network**

support for guaranteed messaging, but fail on other science and engineering requirements. In general, they weren't developed with sensors and science applications in mind, e.g., the integration of heterogeneous instruments and data types, the persistence of delivered data, and sensor stream metadata management. DataTurbine was designed from the beginning to address these requirements. The only other middleware system that approaches Open Source DataTurbine is the Antelope system from Boulder Real Time Systems (BRTT) [BRTT], which was used in the ROADNet project [ROADNet]. It is a proprietary product and is relatively expensive for many communities. At present, DataTurbine is the only open-source streaming data middleware system available. As such, it has a wide and rapidly growing user base among the science and engineering communities.

**2.2 Code Management and Community Support**: A core component of our initiative is to provide professional code management. During the first phase of the OSDT Initiative we released DataTurbine as an open-source product on the Google Code site under the Apache 2.0 license [Apache2.0]. The code is available at http://dataturbine.googlecode.com/.
In addition, we developed key services for community members, including a discussion list, code publication, bug tracking, and documentation. As of 1 February 2009, we have 42 registered members, 720 archived messages, 142 downloads of the OSDT source code, and 1157 downloads of various OSDT binary versions. In addition to code management services, we engaged in system extensions and testing, including porting DataTurbine to additional compute platforms and developing additional interfaces to key sensors/instruments. We also undertook field deployments in a variety of science and engineering applications ranging from civil engineering, limnology, and oceanography. We also participated in community building through workshops, conferences, and collaborations.

**2.3 Open Source DataTurbine Workshop:** The sharing of expertise is an important benefit of the OSDT Initiative. In October 2007, OSDT team held the First Annual Open Source DataTurbine workshop to share experiences and ideas, and to plan for future activities. [OSDT-Workshop]. The theme of the workshop was transitioning from technology development and campaign deployments to persistent operational deployments. Participation was open to anyone interested in OSDT, however the target audience was technology developers and system engineers.

Representatives from several communities (described later) attended the workshop. The workshop was organized into four sessions. (1) Invited presentations on OSDT technology and applications: The primary objective was to hear from OSDT developers and users, in particular the types of deployments (e.g., science topics, types of sensors and instruments, and networking infrastructure), their experiences in using DataTurbine (e.g., usability, performance, robustness), and their ideas for new OSDT developments and activities. (2) Presentations and discussions on DataTurbine deployment issues, including state of health monitoring, metadata management, time synchronization, networking, data replication and mirroring, and system configuration and management. (3) Presentations and discussions on DataTurbine software extensions. Among the topics discussed were the following: GoogleEarth KML plugins for OSDT, LabView interface to OSDT, OSDT support for GOES satellite imagery [GOES], and Google Protocol Buffers [ProtocolBuffers] for OSDT. (4) Discussion of OSDT code management practices and developer support. This session reviewed the current open source support provided to the OSDT community, including code publication, quality control, bug tracking, technical consulting, and discussion forums. The OSDT system for code management was presented as well as the OSDT activities with the NSF NMI Build and Test Facility [NMI-BT].

**2.4 State of the Open Source DataTurbine Middleware:** During the first phase of the OSDT Initiative we focused on the primary requirement for streaming-data applications, namely, data acquisition. Working with our science and technology partners we implemented software extensions and evaluated the data acquisition capabilities of DataTurbine under a number of real-world conditions. These included variations in sensor types, sampling frequencies, compute platforms, and communication networks. During the first phase, we developed several extensions to DataTurbine in the areas of performance and scalability, interoperability via device drivers for network-enabled instruments, and visualization. We now briefly summarize these extensions: (1) Performance and Scalability: DataTurbine was ported to a 64-bit platform to support large-scale distributed collaborative experiments needed by the earthquake engineering community. (2) Interoperability via device drivers for network-enabled instruments: Observing systems have a wide range of hardware (e.g. sensors and dataloggers) and software components, which are determined by local requirements, budgets and preferences. The development activity focused on software device drivers for National Instruments and Campbell dataloggers [Campbell-DL], [NI-CRIO] as well as sensors such as Seacat 16 plus CTD sensor [SEACAT], plus various video cameras. (3) Visualization: We integrated DataTurbine with A Scalable Adaptive Graphics Environment (SAGE)-based OptIPortals [SAGE, OptiPortal], to allow the visualization of real-time data on large tile display walls (described later in the paper) (4) Interfaced DataTurbine with relational database systems for persistent archival of the acquired data (5) System Monitoring: Inca is an NSF-funded project, which provides real-time monitoring of key system parameters; including network and data system processes [Inca]. The Inca system for status monitoring has been integrated into the Open Source DataTurbine and was tested for system and application level monitoring in lake research applications [GLEON].

## 3. APPLICATIONS AND DOMAIN PARTNERS

The OSDT Initiative has strong support from the science and engineering communities. We now describe some of the communities working directly with the OSDT Initiative and the role of Open Source DataTurbine in these communities:

**3.1 The Global Lakes Ecological Observatory Network (GLEON):** GLEON, www.gleon.org, is a grassroots network of limnologists, information technology experts, and engineers who have a common goal of building a scalable, persistent network of lake ecology observatories. Data from these observatories, including The Long Term Ecological Research (LTER) Network sites enable better understanding of key processes such as the effects of climate and land use change on lake function, the role of episodic events such as typhoons in resetting lake dynamics, and carbon cycling within lakes. The observatories will consist of instrumented platforms on lakes around the world capable of sensing key limnological variables and moving the data in near-real time to web-accessible databases. Open Source DataTurbine has been tested at multiple GLEON sites in US and also a GLEON site in Sweden. The feedback received from these deployments has been invaluable for software developments and extensions of the open source middleware.

**3.2 The Coral Reef Environmental Observatory Network (CREON):** CREON, www.coralreefeon.org, is a collaborating association of scientists and engineers from around the world striving to design and build marine sensor networks. Extending sensor networks to the marine environment poses many challenges. However the benefits are enormous as we attempt to understand the stresses that are shaping the marine world. In particular coral reefs are exhibiting signs of decay around the world as global warming, over fishing and pollution have an impact. The CREON group is presently deploying sensor networks in locations as diverse as the Moorea LTER Network site in French Polynesia to the reefs of Taiwan in the Kenting Coral Reef Group and also the Great Barrier Reef in Australia. Using a variety of platforms and instruments the CREON group hopes to solve some of the more technical aspects in a collaborative framework [CREON]. We now describe the role DataTurbine has played for three founding sites of CREON. (1) The Moorea Coral Reef (MCR) Long Term Ecological Research Site: The Open Source DataTurbine is being deployed at MCR [MCR] in Moorea, French Polynesia for acquiring real-time data from a weather station, Axis video camera, and SeaBird CTD sensor [SEACAT]. A temporary field deployment was tested at MCR in the summer of 2008. The production deployment is scheduled for in March 2009. (2) The Great Barrier Reef Ocean Observing System (GBROOS): GBROOS is an observation network that seeks to understand the influence of the Coral Sea on continental shelf ecosystems in north-east Queensland including the Great Barrier Reef (GBR) Marine Park. The project has deployed real-time sensor networks at a number of sites along the GBR and Data Turbine is a key part of how this data is made available. (3) CREON site at Kenting, Taiwan: We developed a system that integrates sensors (underwater video cameras) with computing and storage grids [Strandell07]. This system was extended so that the output of multiple underwater cameras on the grid is viewed in high-resolution on OptIPortals [OptIPortals]. The system is designed for a broad range of users including marine research scientists in Taiwan and the United States. This system was demonstrated using tiled display walls (TDWs) at UCSD and the National Center for High-Performance Computing (NCHC) in Taiwan. OptIPortals provide the ideal termination point for such content rich environments where display real estate can be used effectively. SAGE provides support for streaming video and lets users view tile-displays as big desktops where multiple video streams can co-exist (Figure 2). Users can arrange the video streams on this 'desktop' and resize or maximize them for a better view.

**3.3 NASA Dryden Test Flight Center:** Global Test Range (GTR): The GTR development Laboratory at NASA Dryden serves airborne science and aeronautics research communities [GTR]. Online, near real-time network computing infrastructure is enabled on the ground with an extensible hierarchy of DataTurbine servers support acquisition, transport, processing, and display functions for multiple simultaneous aircraft and globally deployed observation campaigns. The DataTurbine network extends to servers currently on two aircraft that carry research teams in addition to environmental observation instruments. The application network leveraging the DataTurbine infrastructure extends to other NASA field centers. Ongoing cloud computing and sensor web research conducted through this project provide benefits to NASA's efforts to deploy fully operational enterprise-class cyberinfrastructure for near real-time situational awareness.

**Figure 2: Tiled Display Wall (TDW) at Calit2 (UCSD) showing data from underwater video cameras in Kenting (Taiwan) in real time using DataTurbine streaming data middleware.**

**3.4 UCSD Sustainability Institute:** The University of California San Diego is building a Sustainability Solutions Institute that will become a world-renowned center for scholars and practitioners to assemble the intellectual resources and other support needed to address problems of climate impacts, water, energy, biodiversity, the built environment, and long-term sustainability at local, regional, national, and global scales [UCSD-ESI]. Working with organizations outside the university in defining questions and applying research, the institute will engage students (undergraduate and graduate), faculty, and staff from across the campus in interdisciplinary, translational discovery and learning around sustainability challenges. We have deployed DataTurbine to acquire real-time data from 9 weather stations on the 2 square mile coastal UCSD campus to support real-time decisions in building operation, solar power resource assessment, and irrigation scheduling.

**3.5 Structural Health Monitoring:** Advanced Hazards Mitigation Laboratory at University of Connecticut.
Structural health monitoring can provide an unbiased vibration-based assessment of the structural infrastructure in a timely and efficient manner [AHML]. This is critical in our society faced with an ageing infrastructure and limited resources for maintenance and repair. Bridge monitoring in Connecticut is a combined effort between the University of Connecticut and Connecticut Department of Transportation. This program of short and long term monitoring currently has a network of six bridges with long-term monitoring systems. DataTurbine meets a need to provide fully automated continuous monitoring from remote locations and can be used to effectively convey the results of bridge monitoring to the end user. DataTurbine, streaming data from accelerometers and strain gages and video cameras, is currently being installed on two of the highway bridges in Connecticut.

**3.6 Terrestrial and Marine Environmental Monitoring: The National Center for Ecological Analysis and Synthesis (NCEAS):** NCEAS supports cross-disciplinary research that uses existing data to address major fundamental issues in ecology and allied fields, and their application to management and policy [NCEAS]. NCEAS fosters new techniques in mathematical and geospatial modeling, dynamic simulation, and visualization of ecological systems. DataTurbine has been used in the following project at NCEAS.

The REAP project is focused on creating technology in which scientific workflow tools can be used to access, monitor, analyze and present information from field-deployed sensor networks, for both the oceanic and terrestrial environments, and across multiple spatiotemporal scales [REAP]. Initial development for a terrestrial usecase uses DataTurbine and the scientific workflow software Kepler [Kepler]. In this use case Kepler workflows are used to develop and test models exploring the impacts of abiotic factors (real-time light, temperature, and rainfall measurements) on the dynamics of plant host populations and their susceptibility to viral pathogens. REAP has developed a DataTurbine Source program to parse and push data into DataTurbine from a remote weather station, and within Kepler a DataTurbine Sink has been developed in the form of a Kepler actor (workflow component) providing workflow authors a versatile means of requesting and retrieving data from DataTurbine servers.

Researchers at the Hawaii Ocean Observing System [HIOOS] have been exploring the use of near real-time data acquisition from oceanographic sensor arrays. A prototype system employing the Open Source DataTurbine has been deployed at the Kilo Nalu Observatory off the coast of Honolulu, and streams oceanographic data including ocean currents, temperature, pressure, wave spectra, and water quality characteristics. Shore side client applications archive, process and display the data in near real-time, producing web-based graphics and summaries that provide public information on the coastal environment.

**3.7 Biodiversity and Animal Tracking (MoveBank):** MoveBank is an open science community with the common interest of remotely monitoring organisms in their habitat [MoveBank]. It consists of biologists and engineers engaged in a dialog across disciplines and backgrounds with a goal of development and deployment of technologies for gathering data on free-ranging organisms. MoveBank facilitates long-term comparisons of these data making it possible to address pressing questions such as the effects of global climate change and human-caused landscape change. It also complements new technologies for collecting data in real-time by providing live interaction and alerts. Our ongoing activity includes the use of DataTurbine to manage live animal tracking data from radio collars and camera trap, facilitating on-demand access by scientists.

**3.8 Hydrology (CUAHSI): Hydrology (CUAHSI):** CUAHSI is an organization of more than one hundred universities. Its mission is to foster advancements in hydrologic sciences, through developing and disseminating a broad-based hydrologic sciences research and education agenda. CUAHSI participates in several NSF-funded projects, including the CUAHSI Hydrologic Information System (HIS). The project has four goals: to provide data services for hydrologists, to support the CUAHSI observatories, to advance hydrologic science and to improve hydrologic education. In the Phase II of the CUAHSI HIS effort, the HIS team is partnering with the 11 WATERS network observatory testbed sites recently funded by NSF. Several of the test beds are interested in real time and historical water quality or quantity monitoring (in Iowa, Utah, Minnesota, North Carolina, Susqubehanna river basin and Corpus Christi Bay). These testbeds will benefit from the software developed in the HIS project, and provide deployment feedback. OSDT team members prototyped a monitoring system by integrating real-time data from a weather station at UCSD campus with the CUAHSI Data Access System for Hydrology (DASH) system. DASH has been deployed at CUAHSI-HIS Central hosted at UCSD [DASH].

**3.9 The Pacific Rim Application and Grid Middleware Assembly (PRAGMA):** PRAGMA was formed in 2002 to establish sustained collaborations and advance the use of grid technologies in applications among a community of investigators working with leading institutions around the Pacific Rim. Currently there are 35 institutions in PRAGMA, who meet twice a year at PRAGMA Workshops. The PRAGMA testbed provides an ideal environment for testing and hosting Open Source DataTurbine streaming data service due to its international footprint and availability as a development platform on 24-7 basis. PRAGMA testbed gives us an access to: (a) An international-scale network substrate that experiences real-world challenges, including congestion, failures, and diverse link behaviors. (b) A large set of geographically distributed machines spanning multiple administrative boundaries. (c) Realistic client workload. On the PRAGMA Grid we are conducting scaling and robustness experiments with DataTurbine under real-world conditions and at global-scale. The following is a specific example of an ongoing experiment. Our experience of real-world deployments of DataTurbine for multiple observing systems demonstrated that network disruptions are the norm rather than the exception. This practical reality has offered an opportunity to study, for example, performance characteristics of DataTurbine mirroring and routing mechanisms under transient and long term network outages. Quantifying the buffering performance of local servers during outages and recovery characteristics of mirrors after links are restored are of particular interest. The ongoing study involves collaboration with corporate partner Erigo Technologies [Erigo] and will be published in mid-2009.

## 4. CONCLUSIONS

The **Open Source DataTurbine** middleware occupies a unique niche in the NSF cyberinfrastructure portfolio – *a critical piece of the national cyberinfrastructure fabric*. As a tool that solves common problems in placing live data into the network and processing that data with virtually any downstream processing components or workflows, the Open Source DataTurbine middleware has emerged as the core cyberinfrastructure component of environmental observing systems. The NSF-sponsored OSDT Initiative plays a critical role in enabling science and engineering communities to realize the benefits of the DataTurbine middleware. The initiative directly addresses recognized cyberinfrastructure requirements for scalability and interoperability of environmental observing systems [NSF-CEON-08]. Through code developments and community support, e.g., developer services, discussion forums, and collaborative projects, the OSDT Initiative serves as the catalyst and incubator to numerous science and engineering groups. Currently DataTurbine is being developed for a variety of applications around the globe.

**Future Work:** In surveying the current state of the OSDT Initiative, we have identified two technological areas that are important for moving forward: (1) software interfaces that are compatible with the Open Geospatial Consortium (OGC) Sensor Web Enablement (SWE) standards [Botts07], and (2) software extensions that allow DataTurbine applications to run in a cloud-computing environment [Nurmi08]. These activities form the focus of our near-term development plan.

## 5. REFERENCES

**References from Journals**

[Botts07] Mike Botts, George Percivall, Carl Reed, John Davidson OGC Sensor Web Enablement: Overview And High Level Architecture, OGC White Paper, 2007.

[Freudinger09] Freudinger, Lawrence C. "Cyberinfrastructure for Airborne Sensor Webs", 33rd International Symposium on Remote Sensing Of Environment, Stresa, Lago Maggiore, Italy, May 4-8, 2009.

[Liu03] Y. Liu and B. Plale, "Survey of publish-subscribe event systems," Technical Report TR574, Indiana University, 2003.

[Nurmi08] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, D. Zagorodnov, " The Eucalyptus Open-source Cloud-computing System," in Proc. of Cloud Computing and Its Applications, October 2008

[Pallickara03] Shrideep Pallickara, Geoffrey Fox, "NaradaBrokering: A Middleware Framework and Architecture for Enabling Durable Peer-to-Peer Grids," Proc. of International Middleware Conference on Middleware, 2003.

[Strandell07] E. Strandell, S. Tilak, H. Chou, Y. Wang, F. Lin, P. Arzberger, T. Fountain, T. Fan, Rong-Quen Jan, and K. Shao, "Data Management at Kenting's Underwater Ecological Observatory," Proc. of the Symposium on Environmental Sensor Networks, third International Conerence on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2007.

[Tilak07] S. Tilak, P. Hubbard, M. Miller, and T. Fountain, "The Ring Buffer Network Bus (RBNB) DataTurbine Streaming Data Middleware for Environmental Observing Systems," Proc. of the third IEEE International Conference on E-Science and Grid Computing (e-Science) 2007.

**References from Websites**

[Apache2.0] Apache License Version 2.0, http://www.opensource.org/licenses/apache2.0.php, January 2004.

[AHML] Advanced Hazards Mitigation Laboratory: http://www.engr.uconn.edu/~richard/lab/projects.html

[BRTT] BRTT Antelope Real-Time System: http://www.brtt.com/

[Campbell-DL] Dataloggers & Data Acquisition Systems: **http://www.campbellsci.com/dataloggers**

[Chappell04] David A. Chappell, "Enterprise Service Bus," O'Reilly Media 2004, ISBN 13: 9780596006754.

[CORBA] CORBA: http://www.corba.org/

[CREON] CREON: The Coral Reef Environmental Observatory Network, http://www.coralreefeon.org/

[CUAHSI] CUASHI: The Consortium of Universities for the Advancement of Hydrologic Science, Inc., http://www.cuahsi.org/

[DASH] CUAHSI-HIS Data Access System for Hydrology (DASH) http://his.cuahsi.org/dash.html

[Erigo] ErigoTM Technologies: http://www.erigotech.com/

[EMS] Enterprise messaging systems (EMS): http://en.wikipedia.org/wiki/Enterprise_Messaging_System

[GBROOS] Great Barrier Reef Ocean Observing System (GBROOS): http://imos.org.au/gbroos.html

[GOES] GOES satellite imagery: www.goes.noaa.gov

[GLEON] GLEON: The Global Lake Ecological Observatory Network, http://www.gleon.org/

[GTR] Global Test Range Lab: https://indscore.dfrc.nasa.gov,

[HIOOS] HIOOS: Hawaii Ocean Observing System: http:// www.soest.hawaii.edu/hioos/

[INCA] Inca: http://inca.sdsc.edu/drupal/

[JMS] JSR 914: JMS API, http://jcp.org/en/jsr/detail?id=914

[Kepler] Kepler Project: http://kepler-project.org/

[LTER] The Long Term Ecological Research Network: http://www.lternet.edu/

[NI-CRIO] The National Instruments CompactRIO Datalogger: http://www.ni.com/compactrio/

[MCR] Moorea Coral Reef (MCR) Long Term Ecological Research Site: http://mcr.lternet.edu/

[MoveBank] MoveBank: Integrated Database for Network Organism Tracking http://www.movebank.org/

[MSMQ] Microsoft Message Queuing (MSMQ) technology, http://www.microsoft.com/windowsserver2003/technologies/msmq/default.mspx

[MQ] WebSphere MQ, http://www.ibm.com/software/mqseries/

[NCEAS] The National Center for Ecological Analysis and Synthesis: http://www.nceas.ucsb.edu/

[NEES] NEES: The Network for Earthquake Engineering and Simulation, http://www.nees.org

[NMI-BT] NMI Build and Test Lab: http://nmi.cs.wisc.edu/

[NSF-CEON-08] NSF Cyberinfrastructure for environmental observation networks (CEON) workshop report 2008: http://roadrunner.lternet.edu/drupal/files/CEON%20Workshop%20Final%20Report.pdf

[OptiPortals] OptiPortals Wiki: http://wiki.optiputer.net/optiportal/index.php/Main_Page

[OSDT-Presentations] First Annual Open Source DataTurbine Workshop presentations: http://www.dataturbine.org/content/first-annual-open-source-dataturbine-workshop/report

[OSDT-Report] OSDT NSF Annual Project Report: http://dataturbine.org/content/nsf-report

[OSDT-Workshop] First Annual Open Source DataTurbine Workshop: http://www.dataturbine.org/content/first-annual-open-source-dataturbine-workshop

[ProtocolBuffers] Google Protocol Buffers: http://code.google.com/apis/protocolbuffers/

[REAP] Realtime Environment for Analytical Processing (REAP): http://reap.ecoinformatics.org/Wiki.jsp?page=WelcomeToREAP

[ROADNet] Real-Time Observatories, Applications and Data Management Network (ROADNet): http://roadnet.ucsd.edu/

[ROCKS] Rocks: http://www.rocksclusters.org/wordpress/

[SAGE] Scalable Adaptive Graphics Environment: http://www.evl.uic.edu/cavern/sage/description.php

[SEACAT] Sea-Bird SEACAT 16plus-IM V2: http://www.seabird.com/products/spec_sheets/16plusIMdata.htm

## ACKNOWLEDGEMENTS

We would like to thank our science and technology partners: Peter Arzberger (UCSD), Scott Bainbridge (AIMS, Australia), Derik Barseghian (NCEAS), Rich Christenson (Univ. of Connecticut), Bill Finger (Creare Inc.), Sally Holbrook (UCSB), Christopher Jones (UCSB), Matt Jones (NCEAS), Roland Kays (NY State Museum), Jan Kleissl (UCSD), Tim Kratz (Univ. of Wisconsin), Matt Miller (Erigo Tech.), Fang-Pang Lin (NCHC), Rajvikram Singh (UCSD), Larry Smarr (UCSD), Ebbe Strandell (NCHC), John Wilson (Erigo Tech.), and Ilya Zazlavsky (UCSD) for their participation in the Open Source DataTurbine Initiative. Some of the text in this paper is based on the OSDT workshop presentations.