

Cyberinfrastructure for Preservation of Stream and River Ecology

Jill Gemmill¹, Gene Eidson², William Conner³, Sam Esswein⁴, Oscar Flite⁵, Jason Hallstrom⁶, John Hayes⁷, Dan Hitchcock⁸, Tim Howard⁹, Steve Klaine¹⁰, Christopher Post¹¹, Calvin Sawyer¹², Kuang-Ching Wang¹³, David L. White¹⁴

AUTHORS: ¹Executive Director, Cyberinfrastructure Technology Integration, Clemson Computing and Information Technology, Clemson University, Clemson, SC 29634; ²Director, Baruch Institute of Coastal Ecology and Forest Science, Director, Center for Watershed Excellence, PO Box 596, Georgetown, SC 2944; ³Professor, Belle Baruch, PO Box 596 Georgetown SC 29442; ⁴Graduate Research Assistant, Forestry and Natural Resources, 144 Long Hall, Clemson University, Clemson, SC 29634; ⁵Interim Research Director, Southeastern Natural Sciences Academy, 1858 Lock and Dam Road, Augusta, GA 30906; ⁶Assistant Professor, School of Computing, Clemson University, Clemson, SC 29634; ⁷Professor, Agriculture and Biological Engineering, Clemson University, Clemson, SC 29634; ⁸Assistant Professor, Agriculture and Biological Engineering, Belle Baruch, PO Box 596 Georgetown SC 29442; ⁹Project Manager, Customer Relations and Computing Technology, Clemson Computing and Information Technology, Clemson University, Clemson, SC 29634; ¹⁰Professor, Biological Sciences, Clemson University, Clemson, SC 29634; ¹¹Assistant Professor, Forestry and Natural Resources, 144 Long Hall, Clemson University, Clemson, SC 29634; ¹²Extension Associate, Center for Applied Ecology, Clemson University, Clemson, SC 29634; ¹³Assistant Professor, Electrical and Computer Engineering, Clemson University, Clemson, SC 29634; ¹⁴Director of Data Management, Cyberinfrastructure Technology Integration, Clemson Computing and Information Technology, Clemson University, Clemson, SC 29634

REFERENCE: *Proceedings of the 2008 South Carolina Water Resources Conference*, held October 14-15, 2008, at the Charleston Area Event Center.

Abstract. Water resources are facing increasing pressures that range from local changes in hydrology due to urbanization and the effects of pollution and increasing needs of growing populations to shifts in rainfall patterns associated with global climate change. Within the next decade, due to the increasing and competing demands for water, real-time data will be necessary to optimize water withdrawal, waste-load allocation, and hydropower generation while preserving the ecology of our rivers and streams. In addition, real-time networks will support research to better understand the response of hydrology as the landscape is altered due to development. A remote data acquisition system is essential to establish long-term water quality trends, monitor flow, track land-use changes, document pollution impacts, and effectively manage major rivers using real-time data. The Intelligent River™ is developing a hydrological cyberinfrastructure to support research, monitoring and critical management in South Carolina (www.intelligentriver.org).

INTRODUCTION

Sensor networks are proliferating around the world to support environmental monitoring, precision agriculture, wildfire tracking, threat detection, and smart structures. The common goal of these systems is to establish a dense instrumentation of the physical world to support discovery, management and protection. The Intelligent River™ is developing and operating hydrological

observation systems to support research and provide real-time monitoring, analysis and management of water resources in South Carolina. The site locations include Bannockburn near Georgetown, SC, the Savannah River basin and small observation networks in or near Clemson University. These observation networks vary in purpose, scale and density of observation platforms and sensor types, but have a common need to be managed in real time with a well developed software and hardware architecture that is intended to provide 24/7 access to data and visualization products.

METHODS

The Intelligent River™ sensor networks are unique spanning in diversity from “off the shelf” sensors for measuring water quality, soil moisture and weather to low-cost custom sensors for measuring water level, tree girth and additional environmental parameters that support wireless connectivity and long-term deployments. The Intelligent River™ supports these field components with a wireless infrastructure that is using advanced methods to overcome the technical hurdles of wireless technology in environmental sensing networks, including dense vegetation that impedes wireless communication range, lack of cellular coverage in remote locales, hilly terrains that affect or block wireless signals, and the recurring costs and limited bandwidth of cellular and satellite service.

These issues are being resolved by scaling up the sensor network and bandwidth while leveraging distributed sensor clusters and broadband mesh backbones.

The information technology to support the real-time environmental collection system is based in part on the four pillars of cyberinfrastructure as defined by the National Science Foundation that include high-performance computing and communication, virtual organizations, data analysis and learning. The system supports real-time acquisition, processing and distribution of large quantities of streaming observation data. Unlike most observation networks data are continually being transmitted from the field, thus latency in the system is minimized and data can be viewed within minutes of data collection. This is accomplished by the use of wireless technology in the field but also through the use of a publish/subscribe architecture using NaradaBrokering (www.naradabrokering.org) a content distribution infrastructure that enables the development of secure, failure-resilient systems. A key feature of this system has been the use of hierarchical topic-spaces that enables partitioning of observations by topic and supports the management of the sensor network as data streams are added and removed. In addition, this provides a mechanism for limiting access and visibility of observations for security purposes.

The production of thoroughly vetted data from the system will be a critical component of the Intelligent River™ data management program. This will include the use of metadata and well-documented and real-time QAQC program. Limited metadata attributes are embedded in the observation stream at the record level. In addition, using SensorML more detailed metadata records are being created and stored for platform and sensor documentation. These information rich records will catalog standard attributes such as spatial location, units and core characteristics. In addition, pictures, specification sheets and calibration characteristics will be documented and made available through the Intelligent River™ web site. QAQC procedures will support real-time identification and flagging of problematic observations. Data that has passed an initial QAQC procedure that includes flagging and error correction will be reported via the Intelligent River™ web site and archived for further analysis. Non-real time data will pass through additional QAQC procedures using statistical methods to identify outliers and sensor drift. Daily reports will be generated and made available for review to identify problematic sensors and platforms. These QAQC products will be archived and made available for researchers and project partners.

The Intelligent River™ website (www.intelligentriver.org) supports several key features including mapping, graphing, a platform dashboard, data

access and animation products to support a dynamic visualization interface. A key feature of the website is the use of the Google Map API to support data visualization and mapping of project data. Within this application, users can view current observations and access up to several days of data through the use of Flash-based graphing that gives the user a dynamic interface to explore data overtime. Within the mapping framework data aggregation across federated data systems is available. Currently, U.S. Geological Survey (USGS) and U.S. Army Corp of Engineers data are being aggregated into a common database that can then be integrated with Intelligent River™ data sources. This enables the ability to leverage existing data networks to better support management decisions and enable research activities with additional data sources. A Flash-based dashboard provides users a method to explore the platforms via this interface that reuses the Google Map API but incorporates elements such as dials, graphing, platform, sensor images and links to sensor data all in one application interface. Data access is flexible allowing users to access data through common interfaces such Thredds/netCDF. Data will also be available through the use of web services, making data available via GeoRSS, web feature service (WFS) and Google's KML. Animation will play a key role in visualizing research concepts of the Intelligent River™. Current animations visualize the coastal forest environment in a 3D environment focused on changes in hydrology during precipitation events. Future animations may integrate research, modeling and photorealistic landscapes creating a virtual diorama.

DISCUSSION

There are several features that are essential to support hydrological cyberinfrastructure, thus enabling a dynamic interactive repository for sensor network observations. These include making the system scalable by providing the ability to add sensors to individual projects and to add entire new projects to the repository. This enables the ability to manage sensors and data streams with only minor maintenance of the data management architecture. Facilitating near real-time access to data that has undergone QAQC and is well documented using standards – based metadata means the data can be used in web and desktop applications and used to support modeling with little latency in the system. The development of a data registration service can be supported through the use of web services using research and industry standards. Finally, a secure, attribute-based authentication system can provide

controlled access to all its functions and data according to policy set by the data owners.

Finally, a secure, attribute-based authentication system that provides controlled access to all its functions and data, according to policy set by the data owners.

ACKNOWLEDGEMENTS

Funding for the Intelligent River project was provided by the Clemson University PSA and Clemson Computing and Information Technology. In addition we wish to thank the following for their contributions: Fahra Ali, Dr. Devendra Amatya, Rahul Amin, Dr. Rob Baldwin, Chris Behrens, Stan Birchfield, Dr. Richard Brooks, Dr. Tim Callahan, Shari Carter, Rebecca Crane, Knight Cox, Andy Dalton, Brian Daugherty, Jamie Duberstein, Ed Duffy, Chase Finch, Mike Fry, Charlton Galvarino, Dr. Robert Geist, Ben Graham, Dr. Sebastien Goasguen, Jason Isenhower, Dr. Anand Jayakaran, Brian Jones, Shawn Kelly, Wil Kirwan, Dr. Ken Krauss, Josh Lawrence, Sandeep Lokala, Mike McLeod, Dr. Elena Mikhailova, Dr. Shrideep Pallickara, Shashank Parab, Sandi Piazza, Jeremy Pike, Dr. Julia Sharp, Andrew Thomas.