

**Department of Environmental Protection**

**National Hydrography Dataset Data Stewardship and  
Enterprise Water Applications**

**Digital Government – Government to Government**

**Commonwealth of Pennsylvania**

Pennsylvania Department of Environmental Protection  
Nomination for 2008 NASCIO Award

**EXECUTIVE SUMMARY**

Funding has been systematically reduced to many federal agencies. As a result, much of the work related to data creation historically completed by those agencies has been delegated to states via unfunded mandates or strategic partnerships. The National Hydrography Dataset (NHD) partnership program results in a voluntary agreement by states to take responsibility for the NHD network in their state. The NHD is widely recognized as the first successful collaborative project for the creation of one of the National Spatial Data Infrastructure (NSDI) framework layers. This intergovernmental partnership is a hallmark of the NHD cooperative project, involving the United States Geological Survey (USGS), Environmental Protection Agency (EPA), and individual states. In 2005 the Pennsylvania Department of Environmental Protection (DEP) signed an NHD Partnership Memorandum of Understanding with the USGS to maintain the Pennsylvania surface water network for the nation, and officially became an NHD state data steward. DEP assigned staff to continuously update the stream and lakes data for the Commonwealth, correct errors, add new features, and perform updates. These changes are periodically provided to USGS for certification, aggregation, publication, and nationwide use. Some work was contracted by DEP to USGS in order to complete the statewide NHD stream network.

This project was aligned with Administration and Commonwealth IT strategic plans at 3 different levels. DEP leveraged existing technology, a national data model, and shared knowledge from federal agencies to create an enterprise data model foundation for water and land attributes in Pennsylvania. As a result, Pennsylvania is at the forefront of states in completing a statewide, 1:24,000 scale version of the mid-resolution NHD. This has allowed DEP to adjust its enterprise legacy surface water information to utilize the NHD stream network and transformed DEP Surface Water business processes. Completion of the statewide NHD, coupled with the partnership agreement that DEP has signed with USGS to maintain the NHD, allows Pennsylvania to exchange spatially accurate and current water resource information using a national data model with all agencies in the Commonwealth, adjacent states, local governments, the Chesapeake Bay Program, federal agencies, academia, local environmental groups, and the public.

DEP was faced with two options – continue to maintain a proprietary statewide stream network or adopt the NHD data model standard to facilitate data sharing and the integration of standardized scientific models that work with the NHD. DEP evaluated both options and chose to adopt the NHD as the basis for all future water resource related business process and information exchanges. The benefits of wholesale adoption of the NHD far outweighed maintaining the status quo. While conversion was not simple and challenged DEP staff, they proved to be up to the challenge.

Water Deputate and IT staff collaborated to complete this important project. The result of this multiyear project is that DEP has realized numerous operational efficiencies in the past twelve months since the work was completed. Today DEP staff spends less time managing a multitude of water resource data silos and concentrates on completing other mission-critical work. DEP has updated the national NHD data for Pennsylvania three times in the past twelve months, providing the most current and accurate data to USGS for Pennsylvania streams and lakes for use by all NHD stakeholders.

Pennsylvania Department of Environmental Protection  
Nomination for 2008 NASCIO Award

**DESCRIPTION OF THE BUSINESS PROBLEM AND SOLUTION, INCLUDING LENGTH OF TIME IN OPERATION**

Pennsylvania has more stream miles of water (more than 80,000) than any other state except Alaska. Until the adoption of the NHD network for the Commonwealth, DEP and other water resource stakeholders were unable to exchange data with federal agencies and surrounding states that share major watersheds with Pennsylvania (e.g., Ohio River Basin, Susquehanna River Basin, Delaware River Basin). Prior to completion of this project, the DEP Water Deputate had to access many disparate agency data sources and integrate data from each manually in order to answer simple questions. DEP needed to update its statewide stream and lakes network and aggregate twenty-two separate legacy surface water programmatic data silos into a single enterprise database in order to improve efficiency of Department Water Deputate operations, enhance data analysis capabilities, and improve spatial and attribute accuracy of regulatory, water quality, and water resource data related to Pennsylvania surface waters. A primary goal of this initiative was to geospatially-enable all surface water related legacy information at DEP so that it could be visualized and analyzed by location to aid decision makers.

The Commonwealth enjoys abundant and high quality water resources today. Water availability and quality, however, are major concerns in many states. Pennsylvania must work diligently to manage available resources and water quality to ensure that access to abundant and clean water does not become a concern in the Commonwealth.

Significant business process changes were deemed necessary for the establishment of a State Water Plan that would allow DEP to model current use and forecast future water needs, to support the integration of industry standard scientific models, and to facilitate statewide and national data exchange and interoperability of PA surface water data. The creation of the State Water Plan was a state-legislated mandate assigned to DEP in 2002, which must be completed by Spring 2009. DEP has been working on a solution to water and land attribute asset holdings since 2004 by leveraging federal technology projects and expertise to create a foundation of comprehensive, integrated, and standardized data with newly developed tools and spatial visualization capabilities for all water- and land-related events within the Commonwealth.

Prior to completion of the 1:24,000 scale “mid-resolution” NHD for the Commonwealth, the first statewide stream network was initially created when the PA Department of Transportation (DOT) digitized the 1:24,000 USGS 7.5 minute quad sheets for the entire Commonwealth in the late 1980s. The resulting accuracy of the digitized statewide stream network was very poor since the source topography and stream data were outdated and inaccurate. At the same time, no better statewide source existed. DEP adopted the DOT statewide stream network and began updating it by correcting stream names, flipping topology of lines to establish flow directions (first node of a line was moved upstream, last node downstream), and creating a dynamic segmentation capable layer with river mile “measures” so that attributes from legacy databases could be geo-referenced to the stream network. While the enhancements made to the DOT stream network significantly improved the DEP capability to plot permitted facilities, water testing locations, and water withdrawal and discharge points, it did not improve the spatial accuracy of the network nor allow DEP water employees, other state agencies, and local governments to address watershed issues. The DOT/DEP stream network data attributes did not identify watershed membership. Today, managing streams and related attributes by watershed has been identified as a best practice.

Pennsylvania Department of Environmental Protection  
Nomination for 2008 NASCIO Award

Without a standardized data model, DEP could not easily exchange their stream data with municipalities, adjacent states, or federal agencies, and DEP could not easily integrate data from those business partners into its own dataset. Only through the adoption of a national data model could these goals be accomplished.

The NHD began as a national dataset at a scale of 1:100,000 (low resolution), created by USGS and EPA. The USGS definition of the NHD is: “The National Hydrography Dataset is a comprehensive set of digital spatial data that contains information about surface water features such as lakes, ponds, streams, rivers, springs and wells. Within the NHD, surface water features are combined to form ‘reaches,’ which provide the framework for linking water-related data to the NHD surface water drainage network. These linkages enable the analysis and display of these water-related data in upstream and downstream order.” (<http://nhd.usgs.gov>). The NHD was clearly defined with intergovernmental partnerships in mind in order to maintain, upgrade, and improve this strategically important dataset: “The NHD is designed to accommodate both the higher resolution data that many users need and the 1:100,000 scale data. The higher resolution data will be incorporated into the NHD through the participation of users at the national, state, and local levels. The common identifiers for the features are the basis for tracking and sharing deletions, additions, and modifications of features during maintenance. They are used to communicate and share corrections among organizations. The NHD will improve the integration of hydrographically related data in support of the varied applications of a growing national user community, and it will also enable shared maintenance and enhancement.” (<http://erg.usgs.gov/isb/pubs/factsheets/fs10699.html#maintain>).

As one of the Chesapeake Bay Program (CBP) states, Pennsylvania has a responsibility to reduce pollution and runoff of chemicals and sediment into the Chesapeake Bay. DEP, counties, and the other states in the CBP are using the NHD model to exchange information and create a watershed-based stream network for the entire drainage basin. The ability to use the NHD data model to create a multi-state, basin-based stream network for the Chesapeake Bay is allowing the Program, DEP, and counties to model current and future environmental conditions and assess methods for reducing nitrogen and phosphorus concentrations, sediment, and other pollutants that are wreaking environmental havoc on Bay water quality, species, and habitats.

DEP contracted with USGS in 2004 to complete the remaining thirteen Hydrology Unit Code (HUC) regions for the Commonwealth to finalize the 1:24,000 scale-statewide NHD stream network for the Commonwealth. All outsourced tasks discussed in this document were project managed by the selected contractor (e.g., USGS). All non-outsourced tasks identified in this document were project managed and performed in-house by DEP staff with support from onsite contracted staff augmentations. All in-house work utilized the well-established DEP Project Management Methodology (PMM). The Commonwealth’s Enterprise Project Management Methodology (EPMM) borrows heavily from the DEP PMM. (<http://www.projectmanagement.state.pa.us>).

The immediate benefit from the completion of the NHD statewide stream network was that it improved data currency and spatial accuracy, validated network connectivity, added missing streams, added watershed attributes to the stream network, fixed database attribution problems

## Pennsylvania Department of Environmental Protection Nomination for 2008 NASCIO Award

(e.g., stream names), established flow directions, and incorporated a national data model that has been mandated for use by federal agencies and highly adopted by state and local governments. Pennsylvania was one of the first states to complete the NHD mid-resolution dataset and is the only state to date to have fully adopted NHD as the statewide, enterprise framework for all surface water business processes. DEP reengineered all of the agency surface water business processes as a result of full scale NHD adoption in order to maximize the benefits of this project.

While the NHD was being completed for the Commonwealth, DEP began inventorying and assessing all of the various legacy water resource data sets maintained by the department. There was very little data standardization across the twenty-two water data sets that were discovered. The next step by DEP was to standardize the water data elements and create an enterprise Data Element Dictionary. Standardization of the data elements that were part of the NHD model plus all of the legacy data sets that would be related to the NHD network was identified as a critical best practice. The Data Element Dictionary is used daily by DEP to enforce naming conventions in any new database tables and applications that are created and to ensure compatibility and consistency for data exchanges. Each data element is mapped to a data steward and one or more business processes, applications, user interfaces, databases, and reports.

Once the NHD was delivered to DEP by USGS and Quality Assurance work was completed in 2005, DEP outsourced the “conflation” of their legacy databases to the NHD network to a company in NC ([RTI](#)) with NHD expertise. This work established new spatial relationships for each legacy water data record from its location on the old DEP stream network to the new NHD network, so that sampling points, water withdrawal and discharge points, permitted locations, water quality information, and other water-related attributes are geospatially referenced and can be viewed as point and linear events along the NHD stream network. The new enterprise Surface Water database is referred to as the Streams and Lakes Information Management System (SLIMS). Following the initial wholesale data conversion of the water legacy data in 2006, DEP developed a suite of geospatial tools to enable conflation of new and existing data records to take place as needed. The In-Stream Comprehensive Evaluation (ICE) tool provides the ability for DEP staff to insert and update spatial and tabular data along the NHD network. The Water Attribute Viewer for the Enterprise (WAVE) provides geospatial query and visualization tools for the Pennsylvania NHD and SLIMS enterprise database. Both tools were completed by January 2007. When these tools were completed, DEP had successfully achieved all of the original goals and objectives of the project.

The solutions architecture for the DEP NHD network and the related water enterprise database and applications included the following components:

- DEP existing Environment, Facility, Application, Compliance Tracking System (eFACTS) enterprise architecture – modified to accommodate SLIMS and NHD.
- Oracle Database and Oracle Spatial – enterprise database storage for the NHD. Oracle Spatial used for conflation processing and storage of event tables.
- ESRI Spatial Data Engine (ArcSDE, now called ArcGIS Server Basic), ArcGIS Desktop software, ArcIMS and ArcGIS Server web mapping engines.
- SAS software is used to support BI analysis and reporting.

Pennsylvania Department of Environmental Protection  
Nomination for 2008 NASCIO Award

This enterprise project for DEP and the Water Deputate was implemented on a relatively small budget. Approximately \$2.7 million was expended since 2005 to build the foundation for Commonwealth water initiatives. The project was partially funded with federal dollars from the Clean Water Act as well as state general operating funds through the DEP IT budget. Contracted expenditures for the project included a contract with USGS to prioritize and complete the thirteen HUCs, legacy system conflation work, and IT contractor support to augment DEP employee teams to design tables, write software, and develop spatial tools. DEP leveraged knowledge and experience of EPA with Oracle Spatial and ESRI products to build a solid geospatial foundation. Oracle Gateway was implemented to integrate views between Oracle and Microsoft SQL. Additional storage capacity was required to support terabytes of new water-related GIS features. Existing eFACTS software had to be modified to enforce the NHD requirements for water attributes. The ongoing annual operational costs have been absorbed into the enterprise IT budget for Oracle, SQL, ESRI and SAS renewals. DEP staff time was not tracked for this project, but the cost-effective results were achieved through the hard work and dedication of more than fifty DEP employees in the Water Deputate and Bureau of Information Technology. The fact that the development work was “insourced” rather than outsourced saved millions of dollars in costs, and invested DEP staff with the knowledge and experience to support the NHD and new applications.

DEP developed a comprehensive communication plan for the agency and Commonwealth water resource stakeholders, and provided standardized training and dozens of presentations and informational sessions to the stakeholder communities. The DEP story and experience converting to the NHD model has been transferred to other states through EPA Network Operations Board, the CBP, and numerous conferences.

**SIGNIFICANCE TO THE IMPROVEMENT OF THE OPERATION OF GOVERNMENT**

Regulatory, permitting, and federal reporting lines of DEP of business for surface waters were completely transformed as a result of this initiative. DEP Water Deputate staff realized increased efficiencies and productivity. The adoption of the national NHD data model baked interoperability and extensibility into the DEP water business processes. In addition to supporting the Commonwealth’s water resource needs, the DEP NHD partnership with USGS ensures that there is a single data steward in the nation for Pennsylvania streams and lakes that utilizes best practices and standard operating procedures to enhance, publish and share a mission-critical dataset. NHD stewardship utilizes a federated approach to creation of a national surface water data layer as part of the NSDI. The resulting surface water data is easily shared and exchanged with all federal agencies, adjacent states, river basin commissions, the CBP, all state agencies, academia, environmental groups, and the public. DEP data stewardship provides all NHD stakeholders in the nation with current, accurate, and high quality, trusted data for PA waterways.

Primary improvements to DEP, the Commonwealth, and other government business partner operations include:

- 1) Establishment of a single, trusted source for DEP water resource information for enterprise access, analysis, and use in scientific models. In the past answers to

Pennsylvania Department of Environmental Protection  
Nomination for 2008 NASCIO Award

environmental issues depended on who you asked and what private data silo was used – causing conflicts in responses to the public and in decision making.

- 2) Integration and data exchange with other primary and secondary surface water business partners and stakeholders, such as USGS, EPA, USDA Natural Resource Conservation System (NRCS), the CBP, state agencies, county conservation districts, counties, municipalities, local watershed groups, the Fish and Boat Commission, and the DOT, who electronically transmits permit application data for road construction affecting encroachments and wetlands.
- 3) Visualization of water resource information and attributes, allowing users and Commonwealth decision makers to evaluate potential cause and effect spatial relationships and to identify areas of concern based on activities and incidents.

### **BENEFIT OF THE PROJECT**

Everybody won on this project. While not formally measured, DEP achieved a very high ROI on this project. The largest benefit is the DEP contribution to the NHD, providing federal agencies, the CBP, adjacent states, and other business partners with access to accurate and current water data for Pennsylvania. Adoption of the NHD data model facilitates DEP water resource information exchange, access, and interoperability.

The DEP stewardship of the NHD stream network for Pennsylvania ensures that the data remains current and accurate and that efforts are not duplicated. Use of this information will ensure that the high quality water resources in the Commonwealth are conserved for future generations. Less time is spent by Water Deputate staff on managing data, and more time is spent on mission-critical and business-focused tasks such as analysis of water quality issues and the completion of the State Water Plan by spring 2009.

The reengineered water resource business processes provided major improvements to the DEP operational efficiencies and productivity. While many benefits were realized, the largest measured benefit identified to date was the reduction of time for the generation of the bi-annual Integrated Assessment Report to the EPA from six months to six weeks.

Twenty-two disparate water databases now live and reside within a single enterprise database – SLIMS, and are tied directly to the enterprise stream network - NHD. All surface water resources are now geospatially enabled, opening up future possibilities for data analysis, reporting, visualization, and additional enhancement projects.