

USGS Pools Interagency Water Quality Projects onto an Enterprise Grid

Customer Profile



"Having self-management and grid control capabilities, coupled with comprehensive infrastructure for geospatial applications, is critical to our work at USGS. Our internal resources are limited, and we support national projects for USGS and other agencies, so we need products that are easy to manage and scale as we add clients and projects." -- Harry House, Data Center Leader, U.S. Geological Survey (USGS), Wisconsin District

USGS Water Resource Discipline

Middleton, Wisconsin

www.usgs.gov

Industry:

Government

Employees:

>=10000

Oracle Products & Services

- Oracle Database 10g
- Oracle Spatial 10g
- Oracle Application Server 10g
- Oracle Application Server Portal 10g
- Oracle Enterprise Manager 10g Grid Control
- Oracle Internet Development Suite 10g
- Oracle Reports 10g
- Oracle Discoverer 10g
- Oracle Forms 10g

Key Benefits

- Lowers costs by requiring less time or attention in management, allows employing less experienced staff
- Ensures system works optimally by using a single vendor's products
- Improved stability, made for easy development

It seems only natural for water quality watchdog agencies to pool their data resources. That is why the U.S. Geological Survey (USGS) in Middleton, Wisconsin emphasizes sharing IT services within government. The concept grew out of the understanding that communities sharing borders have a common interest in public safety. After all, if a manufacturing plant dumps toxic sludge into a stream, its pollution flows across jurisdictional lines through lakes, rivers, and groundwater, becoming a common problem.

According to Harry House, group leader at the U.S Geological Survey Wisconsin District Data Center, the shared services model "offers the best of two worlds" by encouraging creative initiatives in a governmental agency." The approach is paying off, House said, noting that the data center is two-thirds funded through offering its project services to other agencies. It not only spares each agency the hassle of IT development, administration time, and equipment costs, it ultimately helps the environment by opening floodgates to research information previously locked in each individual agency's databases.

Formed in the mid-1990s, the data center now manages approximately 40% of the USGS internal and external Web servers, plus several inter-agency water quality projects. Its growing list includes projects that are local, regional, and national in scope and come from inside and outside the USGS. The data center's responsibilities range from developing and managing infrastructures for mercury pollution analysis of water systems throughout the United States to beach health information for the Great Lakes in Wisconsin, which monitors dozens of beaches and issues alerts to bathers when pollution rates threaten health. Additional cooperators include the U.S. Fish and Wildlife Service and state agencies such as the Wisconsin Department of Natural Resources and Milwaukee Metropolitan Sewage District.

Oracle 10g's automated management features help meet the Data Center's challenges so that managing a growing data center with limited resources hasn't been a problem for the USGS. "As our programs grow and increase in complexity, it is becoming critical that our supporting systems be increasingly self-managing," House said. "As more agencies report data into our systems, we must ensure we meet our customers' demands. Just because a customer doesn't have a lot of resources doesn't mean they don't have high expectations. They see what's on Google or Yahoo, sites that have millions of dollars behind their systems, and they want us to create something like that for them. We have to figure out ways to supply as much as we can for as little cost as possible."

Easing Burden for IT Team

Oracle Grid Control has provided a centralized dashboard providing

a single interface into all Oracle 10g databases and application servers. The dashboard provides a single, comprehensive display of the center's entire infrastructure. "Such technologies hold the promise for managing USGS servers around the country from centralized sites," said Jim Statz, DBA for the data center. "Tuning advisories and automatic patching are also desirable features. We need fine-grained capabilities in choosing types of recoveries and the ways we run backups. That makes it much easier to access any server's performance and investigate issues as they occur."

House recounts how Grid Control came in helpful during a security inspection at the center. "As part of that review, we were asked how the Center handled patches and what versions of software we were on. I showed them our centralized, browser-based management interface, which instantly gave the security inspectors most of the information they needed. Such information went a long way in assuring the inspector's that our systems were secure and up-to-date."

Taming Unruly Legacy Data Sources

The USGS data center's production and test databases run on Red Hat Linux Enterprise Server, while its application servers run primarily on Windows.

Oracle Discoverer 10g's flexible ad-hoc query capabilities have been leveraged by the Center. One of the USGS data center's largest databases on Linux is the National Water Quality Assessment (NAWQA) data warehouse, which manages water quality data, mostly for surface water, collected across the United States. This database gives USGS researchers a single, integrated system on which to run queries for a variety of legacy datasets. Analysts access the data warehouse's 11 million plus water and biological results using flexible data request screens through their browsers. "The interface makes it easy for scientists to perform complex data retrievals and analysis," said House. The ability to access the data warehouse via such flexible query tools has transformed the way USGS researchers find and verify water quality trends and hot spots.

Oracle Application Server Portal's browser-based rapid development environment has been integrated across the center's projects. The center uses Portal technology for providing the superstructure for all information interfaces for each of their twelve major customer programs. Users access various specialized application servers, such as ad-hoc query, from each Portal interface as needed. "We prefer to use out-of-the-box solutions as much as possible to minimize custom coding for faster time to deployment and minimal maintenance costs," said House.

Modeling Contamination Trails with Oracle 10g Spatial and MapViewer

Oracle's new Map Viewer modules combined with Oracle Spatial's new network data model opens up exciting new avenues for exploring spatial data in new ways. A joint project funded by the EPA involves having the USGS map sites all over the country that are vulnerable to mercury contamination. An application built for the project layers areas of PCB concentrations into a map so researchers can plot the spread of contaminants through waterways after a spill and advise local officials to take action for public safety. Eventually the project will be expanded to other contaminants beyond mercury.

Selecting a Single Vendor

The USGS Middleton data center required a solution that was easy to deploy and manage and could scale to meet its growing

requirements.

"Using a single vendor with a wide range of products minimizes the need for the team to modify code and learn new skills," House said. "Also, I feel it is more likely that all the various technologies a typical project may require will install easily and work together. In my experience, a best of breed approach has been problematic at times. It can be difficult to get an accurate assessment of breed rankings, which can change often. If you decide to switch a technology, that is usually extremely expensive in terms of licensing, retraining, and redeployment costs."

Implementation Process

Oracle Consulting provides expert assistance in a wide variety of data management activities, which have been used to varying degrees by customers as needed. Apart from some initial assistance, House and his staff didn't require the help of outside consultants. "We used Consulting when we first started," he said. "Our consultant was an excellent program architect for our data warehouse and helped us get that off the ground. It worked well for us because at that time we were babes in the woods regarding architecture and data systems. Our consultant also trained us in best practices as part of the deal. That experience provided a strong foundation for us to work from as we moved forward on future projects ourselves. It was well worth it."

Oracle Database 10g's streamlined upgrade procedures have reduced transition times and costs. When the time came to upgrade their systems, House and his DBA spent three weeks of working evenings only to complete the project. They performed one upgrade on a server that was offline during the day to see if their methodology was correct and learned from that what worked. House added that his team uses Oracle University extensively, which he finds helpful for getting his staff up to speed so they can cross-train others. He also finds that going to Oracle World and reading Oracle magazine helps keep him current.

The U.S. Geological Survey (USGS) Water Resources Discipline works with state, local, and federal agencies to keep streams, lakes, and reservoirs safe. It works to alleviate water resource problems and coordinates the collection of federal agencies' water resources data for streams, lakes, reservoirs, estuaries, and ground water. The Middleton, Wisconsin data center is dedicated to deploying high-end information technology products to enhance data storage and access to natural resources datasets.

Any use of trade product or firm names in this article is for descriptive purposes only and does not imply exclusive endorsement by the USGS or the U.S. Government.

ORACLE®