

TAMING LARGE DATASETS WITH SQL SERVER

LESSONS LEARNED FROM A NATIONAL GROUND-WATER QUALITY STUDY

Frank Voss, U.S. Geological Survey, Washington Water Science Center, Tacoma, Washington

GOAL:

Quickly process and analyze national-scale spatial and tabular data sets for a national ground-water quality study

PROCESSING CHALLENGES:

- Different data formats require different programming languages for processing, Arc Macro Language (AML) for coverages, Avenue for shapefiles, Visual Editor (VI) for text, Visual Basic for Applications (VBA) for spreadsheets
- Difficult to link spatial and tabular data
- Difficult to relate tabular data stored in different file formats

SOLUTION:

Use emerging GIS and database technologies to merge all spatial and tabular data into a single database

BENEFITS:

- Powerful tools for managing data
- Perform complex queries using SQL statements
- Need only one programming language (Visual Basic)

ArcGIS 9.0 + SQL SERVER

- Store spatial data in relational database
- Program with Visual Basic

GEODATABASE

- New ArcGIS format for storing:
 - Vector data (points, lines, polygons)
 - Raster data (DEMs, orthophotos, spectral data)
 - Images (scanned aerial photos)

Arc SPACIAL DATA ENGINE (SDE)

- Software for linking ArcGIS to a relational database

SQL QUERY ANALYZER

- Perform complex data retrievals from tabular data using SQL statements

SQL SERVER DATABASE

- Store all tabular data in related tables
- Store images
- Store documents

DATA TRANSFORMATION SERVICES (DTS)

- Import and export data from a large variety of file formats
- Convert data "on the fly"
- Program with Visual Basic
- Save procedures and periodically rerun

