Introduction to Geodatabase and Spatial Management in ArcGIS

Craig Gillgrass
Esri
Session Path

• The Geodatabase
  - What is it?
  - Why use it?
  - What types are there?
  - What can I do with it?

• Query Layers

• What’s coming in 10.1
What is the Geodatabase?

- Core ArcGIS data model
  - A comprehensive model for representing and managing GIS data

- A physical store of geographic data
  - Scalable storage model supported on different platforms

- A transactional model for managing GIS workflows

- Set of COM components for accessing data
Geodatabase is based on relational principles

- Leverages key DBMS principles and concepts to store geographic data as tables in a DBMS
  - Data is organized into tables
  - Tables contain rows
  - All rows in a table have the same attributes
  - Each attribute has a type
  - Relational integrity rules exist for tables

- The core of the geodatabase is a standard relational database schema
  - a series of standard database tables, column types, indexes, and other database objects
Geodatabase is based on relational principles …

- A feature class is stored as a simple DBMS table
- Each row represents a feature
- The fields in each row represent various characteristics or properties of the feature
- One of the fields holds the feature geometry which is stored as a spatial type
Geodatabase is based on relational principles …

- A feature class is stored as a simple DBMS table
- Each row represents a feature
- The fields in each row represent various characteristics or properties of the feature
- One of the fields holds the feature geometry which is stored as a spatial type
• There are two sets of tables:
  - Dataset tables (user-defined tables)
  - Geodatabase system tables
User-defined tables

- Stores the content of each dataset in the geodatabase
- Datasets are stored in 1 or more tables
- Spatial Types enhance the capabilities of the geodatabase
  - SQL access to geometry
  - Industry standard storage model and API
Geodatabase system tables

- System tables store definitions, rules, and behavior for datasets
- Tracks contents within a geodatabase
- 4 main system tables
- Geodatabase schema is stored primarily within an XML field
Geodatabase Data Management Approach

• Simple features + logic
  - All geographic data stored as tables in a DBMS
  - Extend functionality and data integrity
  - Functionality is consistent across DBMS’

• Application logic (software)
  - Works on standard DBMS tables
  - Implements GIS integrity and behavior
  - Business rules, topology, networks
Geodatabase Data Management Approach …

• Editing and data compilation
  - Rich set of editing tools
  - Maintain spatial and attribute integrity
  - Undo and redo edits
  - Multiple users editing the same data

• Versioning work flows
  - Long transactions
  - Distributed data management
  - Archiving

• Robust, customizable framework
  - Build and manage your own specific GIS solution
3 Types of Geodatabases

- Geodatabases can be stored different ways
  - Personal geodatabase (Access mdb file)
  - File geodatabase
    - Directory of binary files
  - ArcSDE for SQL Server Express
  - Enterprise ArcSDE
    - 5 supported DBMSs
### 3 Types of Geodatabases…

<table>
<thead>
<tr>
<th>Feature</th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
### 3 Types of Geodatabases…

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
### 3 Types of Geodatabases…

<table>
<thead>
<tr>
<th></th>
<th>Personal GDB</th>
<th>File GDB</th>
<th>ArcSDE GDB (3 editions)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage format</strong></td>
<td>Microsoft Access</td>
<td>Folder of binary files</td>
<td>DBMS</td>
</tr>
<tr>
<td><strong>Storage capacity</strong></td>
<td>2 GB</td>
<td>1 TB per table*</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Supported O/S platform</strong></td>
<td>Windows</td>
<td>Any platform</td>
<td>Depends on edition</td>
</tr>
<tr>
<td><strong>Number of users</strong></td>
<td>Single editor Multiple readers</td>
<td>Single editor Multiple readers</td>
<td>Multiple editors &amp; readers</td>
</tr>
<tr>
<td><strong>Distributed GDB functionality</strong></td>
<td>Check out/check in and One-way replication</td>
<td>Check out/check in and One-way replication</td>
<td>Replication (all types) &amp; versioning</td>
</tr>
</tbody>
</table>

* By default; option to have 256 TB per table
Modeling Real-World Data with the Geodatabase

- A geodatabase contains datasets.
  - Datasets represent collections of information with a real-world interpretation.
  - Types of geographic datasets:
    - Tables
    - Object classes, feature classes, relationship classes
    - Feature datasets
    - Networks, Topologies, Raster and cadastral datasets

- Datasets have associated information to help manage integrity, behavior, and interpretation
  - Domains, Relational integrity, Topology, Metadata
The geodatabase enhances data and thematic layers by adding rules and behavior:
- Spatial and relational integrity rules
- Data validation
- Business logic

Create thematic layers with behavior:
- Road and utility networks
- Parcel fabrics
- Terrain and 3D surfaces
- Location services

Extended framework for advanced workflows and editing:
- Multiuser editing, Data Replication, Editor tracking, Archiving
Exploring a Geodatabase Demo

• Explore a Geodatabase
  - Tables
  - Feature Classes
  - Subtypes
  - Domains
  - Relationship Classes
Editing Geodatabases

• **ArcGIS datasets stored in the geodatabase are editable**
  - Modify building footprints in parcel management
  - Add water mains to a water network
  - Update land owners information stored in a table
  - Etc…

• **Transaction model for editing in ArcGIS**
  - Edits are performed in an edit session
    - Open session – edit – save edits / don’t save edits
  - A series of edit operations constitutes a transaction
    - Unit of work performed against the database
    - The transaction is either committed or rolled back
Editing Geodatabases...

• Personal Geodatabases
  - Single user editing on small datasets
  - Multiple readers
  - Editing locks at geodatabase level

• File Geodatabase
  - Single user editing small to very large datasets
  - Multiple readers
  - Editing locks at the dataset level
Editing Geodatabases…

• ArcSDE Geodatabases
  - Extend the transaction model with Versioning
  - Multiuser editing without locking
    - Unique isolated view of the geodatabase

• Benefits of versioned editing
  - Multiple editors, editing over long periods of time
  - Undo / Redo
  - Archiving
  - Replication
Three different ways of editing Geodatabases

- **Versioned Editing (Long Transactions)**
  - Editing in a version through ArcGIS

- **Non-Versioned Editing (Short Transactions)**
  - Editing the data directly through ArcGIS

- **Editing through SQL (Short Transactions)**
  - Editing the data directly through SQL
Versioning - What is it?

- Technology that allows multiple users to edit and view data at the same time
  - Appears to users as if they have their own copy of a table
  - Does not apply locks or duplicate data
What is a Version?

• An **alternative view** of the geodatabase that has:
  – an owner
  – a description
  – a permission
  – a parent version

• Versions are not affected by changes occurring in other versions of the database
What is a Version?

• An alternative view of the geodatabase that has:
  – an owner
  – a description
  – a permission
  – a parent version

• Versions are not affected by changes occurring in other versions of the database
Versioned Editing – How It Works

- Class must be registered as Versioned
  - Creates Adds and Deletes tables for tracking edits
Versioned Editing – How It Works

• Adding Features
  - Record added to the Adds Table
  - Version will be referenced
Versioned Editing

- **Versioned Edit Sessions**
  - Editing done through a version
    - Changes tracked on delta tables
  - Support concurrent editing with long transactions (hours/days).
  - Undo/redo editing experience.
  - No locking or data extraction required.
Versioning allows us to support

- Replication

Archiving
Session Path

• The Geodatabase
  - What is it?
  - Why use it?
  - What types are there?
  - What can I do with it?

• Query Layers

• What’s coming in 10.1
Query Layers

• Query Layer is simply a layer or table defined by a SQL query

• Direct, read-only access to spatial data independent of where it is stored

• Allows information stored in a database to be easily integrated with ArcGIS

• Data does not need to be registered with ArcSDE or the geodatabase
What does this mean to a developer?

• You now have a simple method to integrate data from databases
  - SQL Workspace: New Workspace Type at 10.0
    - SQLWorkspaceFactory

• Can now use the full power of SQL to create feature classes
  - Joins, Spatial Types, Aggregation, Field Derivation
  - Basically if you can do it with SQL you can do it with Query Layers

• Can now do more with your Geodatabase datasets

• SQL Server, Oracle, PostgreSQL, DB2 and Informix
  - Need access to a spatial type
  - Database release and spatial types must be supported by ArcGIS
Working with query layers

• Query layers behave like any other feature layer or table
  - Set symbology and display properties
  - Use as input in GIS analysis (e.g., geoprocessing)
  - Save as a layer package
  - Published as part of the map document → map services
  - Read only

• Query executed when the layer is displayed or used
  - Allows for the latest information to be visible → very useful for dynamic information
Prerequisites for working with query layers

- Result of a query must conform to the ArcGIS Data Model
  - Single shape field
  - Single entity type (Point, Line, Polygon)
  - Single spatial reference

- Can still access tables that do not use this data model
  - Just need to define the query against this table to conform to the above standards
Query Layers Demo

- Accessing spatial data in a database