Managing Distributed Data with Geodatabase Replication

Patrisha Wells
Outline

- General Concepts
  - Use Cases
  - Types of Replicas
  - Data Requirements
  - Replica Creation
  - Synchronization
  - Schema Changes
- Geodatabase Replication API
- Best Practices
- What's new at 10
- Summary
Geodatabase Replication

- Distribute data across 2 or more geodatabases
- Geodatabases are edited independently and synchronized when needed
Geodatabase Replication – Use Cases

- Mobile Users and Field Crews
Geodatabase Replication – Use Cases

- Copies of data at different organizational levels (city, county, state)
Geodatabase Replication – Use Cases

- Copies of data at different geographic facilities
Geodatabase Replication – Use Cases

- Production and publication geodatabases
Geodatabase Replication - LAN and WAN

- LAN - connects to local geodatabases

- WAN – connects to remote geodatabases (geodata services on ArcGIS server)

- All replication workflows supported in both environments
Data Distribution in Enterprise systems

- Geodatabase replication used with other data distribution techniques

- Scenario
  - Synchronize offices with geodatabase replication
  - Field workers use mobile services
  - Field workers use geodatabase replication
You can replicate:
- A specific version
- Specific datasets
- A subset of features in the chosen datasets
Types of Replicas

Check out/Check in
- Parent geodatabase
- Child geodatabase
- Once only

One-way
- Parent geodatabase
- Child geodatabase
- Multiple times

Two-way
- Parent geodatabase
- Child geodatabase
- Multiple times

OR

edits
Check out / Check in Replication

- Child can be any geodatabase type
- Parent must be an ArcSDE geodatabase
One Way Replication

- Parent to child or child to parent (10)
- Source geodatabase
  - Must be an ArcSDE geodatabase
  - Use archiving to track changes (10)
- Target geodatabase
  - Any geodatabase type
  - Considered read-only
Two Way Replication

- Parent and child must be ArcSDE geodatabases
Data Type Support with Replication

• Fully supported
  - Simple Feature Classes and Tables, Geometric Networks, Topologies, Representations, Relationship Classes, Annotation

• Source data can be replicated
  - Terrains and Network datasets

• Copied during creation
  - Raster datasets, Raster catalogs, Mosaic datasets

• Not Supported
  - Cadastral datasets, Schematics datasets, Locators
Replication Data Requirements

- Data must be writable

- Data must be versioned
  - Without option to move edits to base

- Two way and One way replicas
  - Data must have a Globalid column
  - Spatial data must be high precision
GlobalID Columns

- GlobalID columns
  - Maintain object identity
  - System maintained (like ObjectIDs)
  - Contain registry style UUID values
    {9DFACA0A-982F-4175-80E7-B553378D9E6D}

- Adding globalid columns
  - ArcCatalog command or GP tool
  - Data can be versioned

- Differ from columns of type GUID
Replica Creation - Versioning

• Replica versions are synchronized
• 2 way and 1 way replicas
  - Named version or default on parent and child (10)
• Check-out replicas
  - Named versions created by the replica
Replica Creation – Defining data to replicate

- Filters and Relationship classes
  - Filters are applied first, the relationship classes
- Filters
  - Spatial – Area to replicate
  - Selections – Selection sets on feature classes and tables
  - QueryDef – Definition queries on layers and tables
- Relationship Classes
  - Additional related rows are added
Creating Replica with related Data

- Parent Replica
  - RC1
  - Owner: John
  - Incident: John Incident
  - Owner: Mary
  - Incident: Mary Incident

- Child Replica
  - RC1
  - Owner: John
  - Incident: John Incident

- RC2
Replica Creation – Full and Simple model

• Full Model
  - Requires feature types to match on parent and child
    - Example: A complex edge must be a complex edge type on both the parent and the child

• Simple Model
  - Does not require feature types to match on parent and child
    - Example: A complex edge on the parent can be a simple line on the child
Replicating to a remote geodatabase

1. Publish geodata server
   - Remote ArcGIS Server
   - Remote ArcSDE GDB

2. Connect to the geodata service
   - Local Office
   - Local GDB

3. Create a replica and then synchronize changes
Replica Synchronization

- Applying data changes across replicas
- Synchronize in both directions or a single direction
- Uses message exchange
- Fault tolerant
- Uses versioning (or Archiving for one way (10))
- Can synchronize as replica owner or ArcSDE user
What Features are Synchronized

- 2 way and 1 way replicas
  - Replica filters used to find changes to sync
  - Edits related to data in the replica are synchronized (schema only filter required)

- Check-out replicas
  - All edits are synchronized
Replica Synchronization - Message Exchange

- Synchronizations are performed using message exchange
- Data Change messages
  - Used to send changes
  - Include all changes since the last acknowledgement (default)
- Acknowledgement messages
  - Used after receiving changes
  - Acknowledges that the changes have been received
Connected Synchronization

- Requires replica geodatabases to be on the same network (LAN or WAN)
- Always connected or intermittently connected
- System manages message exchange
- Send in one or both directions in one operation
Disconnected Synchronization

- Does not require the replica geodatabases to be on the same network
- Message exchange is performed by the end user
- Involves export, file transfer and import
  - File transfer can use ftp, CD through the mail, etc
Disconnected Synchronization

• A replica is either a Data Sender or a Data Receiver
  • Data Sender
    - Sends data change messages to the data receiver
  • Data Receiver
    - Accepts changes from data sender
    - Sends back an Acknowledgement messages
  • Switch Roles
    - Data Sender sends Data Change message with instructions to switch roles
Disconnected Synchronization

- Acknowledgement messages are not required after receiving changes
- Changes are resent if an acknowledgement is not received
- Switching roles on every data send acknowledges implicitly
Replica Synchronization Examples

- Connected Synchronization with a 2 way replica

- Disconnected Synchronization with a 2 way replica
Synchronization and Conflicts

• Choose to define conflicts by row or by attribute
• Three policies for conflicting changes
  - Favor the database (Automatic)
  - Favor the imported changes (Automatic)
  - Manual
    - Replica is marked as in conflict
    - Need to resolve conflicts manually at a later time
    - Can send but not receive while conflicts exist
    - Not supported when synchronizing in both directions
Manual and Automated Synchronization

• Manual
  - Desktop commands / GP tools
  - Choose to synchronize when needed

• Automated
  - Schedule synchronizations
  - Use GP models and the windows scheduler
  - Recommended
Working through errors

- Designed to keep replicas consistent
- Roll back to pre-synchronize state if synchronization fails
- If a message is lost, the next message includes the lost message changes and new changes
- Replica log records error information
Demo 2 – Replica Synchronization Demo

1. Make Edits
2. Synchronize with geodata service
3. Edits are transferred

Local Office
Local GDB

Remote Machines (Amazon cloud)
Remote ArcGIS Server
Remote ArcSDE GDB
Working with Schema Changes

- Fault tolerant
  - Synchronization continues to succeed after schema changes
  - Example: If a field has been dropped, synchronization skips that field

- Tools to apply schema changes across replicas
  - Subset of schema changes can be applied
Applying Schema Changes

- Two step process
  - Compare replica schemas
    - Direct or through schema XML files
    - Schema differences file created
    - Import the schema differences

- Choose the differences to apply during Import process
- Some differences can’t be applied and are listed for information only
### Schema Changes that can be Applied

<table>
<thead>
<tr>
<th></th>
<th>Add</th>
<th>Change</th>
<th>Drop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field</td>
<td>Y</td>
<td>Y (Domains)</td>
<td>Y</td>
</tr>
<tr>
<td>Domain</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Table/FC</td>
<td>Y</td>
<td>Y (Domains, Add/drop field)</td>
<td>Y</td>
</tr>
<tr>
<td>Geo Network</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Topology</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Feature Dataset</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Relationship Class</td>
<td>N</td>
<td>Y (Domains, Add/drop field)</td>
<td>Y</td>
</tr>
</tbody>
</table>
Replica Description and Viewing

- **Replica Manager**
  - Lists all replicas in the Geodatabase
  - Can get properties about each replica
  - Includes replica log
  - Allows you to unregister (delete) a replica

- **Replica Footprints FeatureClass**
  - Creates a feature class of replica information
  - One row per replica
  - Shape field includes the replica geometry
Geodatabase Replication – API

- Most replica functionality is in the commands/tools

- Write code when…
  - Integrate replication into larger applications
    - Example: integrate synchronization with reconcile service
  - Extend the replica creation and synchronization
  - API only functionality
Geodatabase Replication – API

- API only functionality
  - Recommended synchronization order
  - Browse changes before synchronization
  - Advanced replica creation options
    - Example: Overlap spatial filter
  - Export and Import version differences
Geodatabase Replication – API

**Coarse Grained API**
*(GeoDataServer object model)*

**Fine Grained API**
(Data Extraction Check out/Check in object model; XML Export and Import Object Model; Data Changes Object Model; Schema Change Export and Import Object Model)*
Geodatabase Replication – Best Practices

- Anticipate future needs creating a replica
- Develop a well defined data model before creating replicas
- Automate replica creation and synchronization
  - GP models and windows scheduler
- Consider replicas when managing ArcSDE
  - See Geodatabase Replication and Compress
- Use appropriate technique on based on network speed
- Upgrade to the latest release or service pack
  - Get the replication patch at 10
Geodatabase Replication – FAQ’s

• FAQ’s
  - Reverse one way (kb)
  - Working with multiple projections (kb)
  - ArcGIS server and timeouts with GeoData Services (kb)
  - Style sheets for the ReplicaLog.dat (kb)
  - ArcGIS server and using it with SQL express (server doc)
  - Adding a feature class to a replica (SDK sample)
  - Relationship classes where the objectID is the primary key and replication
DBMS Replication vs Geodatabase Replication

• Geodatabase replication
  - Built on top of the geodatabase (DBMS independent)
  - Supports geodatabase specific features types
    - Geometric networks, topologies etc.
  - Replicates a specific version
  - Loosely coupled

• DBMS Replication
  - Replicates geodatabases at the DBMS table level
  - Requires some internal knowledge of the geodatabase
  - Can replicate the entire geodatabase
New in 10: Geodatabase Replication

- **New Functionality**
  - One way replication using archiving
  - One way child to parent replica
  - Schema mapping across replicas
  - Simple check-out and two-way replicas
  - Create a replica to a named version on the child
  - Support for non-versioned data with Check out replicas

- **Better Support for data centralization workflows**
Demo Title: Option 1

Topic 1
Topic 2
Topic 3
Demo Title: Option 3

Topic 1
Topic 2
Topic 3
Sample Screenshots layout
Using ArcGIS 10 – Editing Tools (sample diagram)
Shapes for Diagrams

- Circle behind a group of objects (resize as needed)
- Box behind a screen shot (resize as needed)
- Cloud optimized for use behind diagrams (resize as needed)
Arrows

Arrows for Connecting Most Items

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Arrows for Connecting Large Concepts

...
Access the Entire Icon Library

Over 80 items added for 2010

- 350 total icons available for Esri use
- Browse and search from any Microsoft Office application
- Accessible if you’re connected to the Esri Network
- Also available offline
- See the presenter notes below for details
A Selection of Frequently Used Icons

- ArcGIS Desktop
- ArcGIS Online
- Web Blog
- Web Blog
- Mashups
- Mashups
- Browser
- Open Standards
- ArcGIS Desktop Authors
- Web Map
- Web Map
- Map
- Web Map
- Map
- Web Map
- Map
- Web Map
- Map
- Explorer
- Web Map
This area is the **title safe area** (20% in from the slide edge). All text and graphics should be contained within to prevent loss during transmission and reproduction.

Any information outside the title safe area runs the risk of being cropped off when captured to video.

Please note the adjusted title and body template adhering to the title safe area and templates. Presenters should adjust their slides accordingly depending on the position of text and graphics.

Right-click and select Grid and Guides
Check "Display drawing guides on screen"
Successful Presentation Guidelines

• Know your target audience
• Use slides to lead not read
  - Keep your slides uncluttered and to the point
  - Avoid more than two levels of bullet points
• Use title slides for each section
  - Make it clear where you are going
• Avoid too much animation—Keep it simple!

Additional ESRI presentation resources available on ArcZone
http://arczone/resources/presentations.cfm
Color Guidelines

Color Swatches
Follow this guide when applying color on text

Projector Color Guidelines
Use the sRGB video mode on
Most projectors have this setting.

Additional ESRI presentation resources available on ArcZone
http://arczone/resources/presentations.cfm