Managing your Distributed Data

Gary MacDougall
Premjit Singh
Feedback

• Please provide your feedback on:
  - Use *Offering ID: 616*
Workshop Overview

- Data Distribution
  - What, Why and How
- Distribute Data using Geodatabase Replication
  - Concept Overview
  - Types of Data Distribution Strategies
  - Defining Data to be Distributed
- Synchronizing Distributed Data using Geodatabase Replication
- Demos
What is Data Distribution?

• **Data Distribution involves**
  - Creating copies of data
  - Distributing copies between 2 or more Geodatabases
  - Allowing remote offices to work independently on the copy of same data
  - Synchronizing Edits
Why Distribute Data?

• **Increases Availability**
  - Distribute Data Geographically

• **Improves Performance**
  - Load Balancing

• **Easy Accessibility**
  - Mobile Users and Field Crews

• **Easy Sharing through**
  - Online
  - Desktop
  - Server
  - Mobile
  - Developer
  - Solutions
How to Distribute Data?

**Creating Copies**
- Creating copies of Geodatabases on CDs and DVDs
- Distributed to Offices
- Updated Geodatabases send back to the main office
- Edits are compared *Manually* to keep data in sync
- Difficult and Tedious Process

**Geodatabase Replication**
- Distribute Data through ArcGIS System
Geodatabase Replication Overview

- **Distribute Data** across 2 or more Geodatabases
  - Involves Creation of a Replica Pair
- **Geodatabases are edited** independently
  - Loosely Coupled
- **Synchronize Edits**
  - No manual comparison required
- **DBMS Independent**
  - Source can be SQL Server and Target could be Oracle
- **Built on Versioning**
  - The Source Geodatabase is always ArcSDE
• Supports **Full Geodatabase Data Model**
  - Topologies, Networks, Terrains

• Supports **Connected and Disconnected** Environments

• Supports **LAN and WAN**
  - **LAN** - connects to local Geodatabases
  - **WAN** – connects to remote Geodatabases
    - **Geodata Services** on ArcGIS Server
Geodatabase Replication - Concepts

- 2 Step Process
  - Creating Replica
  - Synchronizing Edits
Geodatabase Replication – Concepts (Cont.)

• Based on Synchronization Direction
  - 3 Types of Data Distribution Strategies
    - Check-Out / Check-In
    - One-way
    - Two-way
Check-Out / Check-In Data Distribution

- Edits are performed on Child Geodatabase
- Synchronizes Once
  - Child to Parent
- Child is any Geodatabase Type
- Parent must be an ArcSDE Geodatabase
- Example: Mobile Users and Field Crews
Demos – Public Safety Special Events Maps

Plan for special events

The Public Safety Special Event Planning map and editing tools can be used to manage special events in a community. It is an editing map that can be used by mapping technicians in a public safety or emergency management agency to streamline the development of special event data and maps.

An editing map used to plan for special events in your community.

You can download and configure the Special Event Planning map for your...
One-Way Data Distribution

• **One Direction Multiple Synchronization**
  - Source to Target or Target to Source (10)

• **Source Geodatabase**
  - Must be an ArcSDE Geodatabase
  - Use archiving to track changes (10)

• **Target Geodatabase**
  - Any Geodatabase type
  - Considered read-only

• **Example:**
  - Production - Publication Use Case
One-Way Data Distribution (Contd.)

- Another Option to track changes
  - Archiving
- No Version Management
  - Reconcile – Post and Compress not Affected
- Source Geodatabase must be ArcSDE
- One-Way Parent to Child
One-Way Data Distribution (Contd.)

- One-Way Child to Parent
- Child Geodatabase is ArcSDE
  - Child Replica Version is Default.
Two-Way Data Distribution

- Multiple Synchronization In Both Directions
- Both Source and Target must be ArcSDE Geodatabases
- Example: Multiple Editors Use Case
Data Requirements for Data Distribution

• **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
Data Types Supported By Geodatabase Replication

• **Fully Supported**
  - Simple Feature Classes and Tables, Geometric Networks, Topologies, Representations, Relationship Classes, Annotation, Cadastral datasets (10.1)

• **Source Data can be Replicated**
  - Terrains and Network datasets

• **Copied during Creation**
  - Raster datasets, Raster catalogs, Mosaic datasets

• **Not Supported**
  - Schematics datasets, Locators
Defining Data for Distribution

• Filters and Relationship classes
  - Filters are applied first, then the Relationship Classes

• Types Of 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
Defining Related Data for Data Distribution

Parent Replica

Owner
John
Mary

Incidents
John Incident
Mary Incident

Child Replica

Owner
Mary

Incidents
Mary Incident
Demo 3
Synchronizing Distributed Data

• **Involves**
  - Transferring and Applying Data Changes
    - Inserts, Updates and Deletes

• **Geodatabase Replication Synchronization**
  - Synchronizes in Both Directions or in Single Direction
  - Uses Message Exchange
  - Fault Tolerant
  - Uses Versioning (or Archiving for one way (10))
  - Can synchronize as Replica Owner or ArcSDE user
Which Data is Synchronized

• **For Replica of Two-Way and One-Way**
  - *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
Which Data is Synchronized (Contd.)

- **Maintains Relationships**
  - Apply changes to Relationships on Sync

- **Geometric Network**
  - Rebuilds Modified Network

- **Topology**
  - Marks Dirty Area
Synchronizing Distributed Data using Geodatabase Replication

- **Connected Synchronization (Two-Way)**
Synchronizing Distributed Data using Geodatabase Replication (Contd.)

- **Disconnected Synchronization (Two-Way)**
  - 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.
Geodatabase Replication Conflict Policies

- Choose to define conflicts by row or by attribute
- Geodatabase Replication offers three ways to resolve conflicts:
  - Favor of 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
Geodatabase Replication Error Handling

- **Roll Backs** to pre-synchronize state if synchronization fails
- If a message is lost, the next message includes the lost message changes and new changes
- Designed to keep **Replicas Consistent**
- Replica log records **Error Information**
Geodatabase Replication Synchronization Tools

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

• Automated Tools
  - Schedule Synchronizations
  - Use GP models and the Windows Scheduler
  - Recommended
Demo 4
Gary MacDougall
Working with Schema Changes

• **Fault tolerant**
  - Synchronization continues to succeed after Schema Changes
    - Example: If a field has been dropped, synchronization skips that field

• Subset of schema changes can be applied

• Tools to apply schema changes
### 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><strong>Field</strong></td>
<td>Y</td>
<td>Y (domains)</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Domain</strong></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Table/FC</strong></td>
<td>Y</td>
<td>Y (Domain, Add/drop field)</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Geometric network</strong></td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Topology</strong></td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Feature dataset</strong></td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Relationship class</strong></td>
<td>N</td>
<td>Y (add/drop field, domain)</td>
<td>Y</td>
</tr>
</tbody>
</table>
Best Practices to Apply Schema Changes

- Lock down the system
- Apply periodic schema comparisons
- Apply system wide schema changes
New In ArcGIS 10.1

- Better trouble shooting tools
- Used in other workflows
  - Local editing commands
    - A Feature Service that you can edit from Desktop
- Feature Service Disconnected Editing REST API
- Publishing to the Cloud
Geodatabase Replication – White Paper

- Geodatabase Replication and Compress

- http://video.esri.com/watch/1073/publishing-feature-services-to-arcgis-for-server-with-replicated-data-
Steps to evaluate UC sessions


- Choose session from planner
  
  OR

- Search for session

[www.esri.com/ucsessionssurveys](http://www.esri.com/ucsessionssurveys)
• Please fill out the evaluation:

  www.esri.com/ucsessionsurveys

  First Offering ID: 616

  Second Offering ID: 700

• Thank you for attending
  • Have fun at UC2012

• Open for Questions
Thank you