

Geodatabase Replication: Working with Replication

Transcript

Copyright © 2007–2008 ESRI
All rights reserved.
Printed in the United States of America.

The information contained in this document is the exclusive property of ESRI. This work is protected under United States copyright law and other international copyright treaties and conventions. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system, except as expressly permitted in writing by ESRI. All requests should be sent to Attention: Contracts and Legal Services Manager, ESRI, 380 New York Street, Redlands, CA 92373-8100 USA.

The information contained in this document is subject to change without notice.

U.S. GOVERNMENT RESTRICTED/LIMITED RIGHTS

Any software, documentation, and/or data delivered hereunder is subject to the terms of the License Agreement. In no event shall the U.S. Government acquire greater than RESTRICTED/LIMITED RIGHTS. At a minimum, use, duplication, or disclosure by the U.S. Government is subject to restrictions as set forth in FAR §52.227-14 Alternates I, II, and III (JUN 1987); FAR §52.227-19 (JUN 1987) and/or FAR §12.211/12.212 (Commercial Technical Data/Computer Software); and DFARS §252.227-7015 (NOV 1995) (Technical Data) and/or DFARS §227.7202 (Computer Software), as applicable. Contractor/Manufacturer is ESRI, 380 New York Street, Redlands, CA 92373-8100 USA.

@esri.com, 3D Analyst, ACORN, ADF, AML, ArcAtlas, ArcCAD, ArcCatalog, ArcCOGO, ArcData, ArcDoc, ArcEdit, ArcEditor, ArcEurope, ArcExplorer, ArcExpress, ArcGIS, ArcGlobe, ArcGrid, ArcIMS, ARC/INFO, ArcInfo, ArcInfo Librarian, ArcInfo—Professional GIS, ArcInfo—The World's GIS, ArcLessons, ArcLocation, ArcLogistics, ArcMap, ArcNetwork, ArcNews, ArcObjects, ArcOpen, ArcPad, ArcPlot, ArcPress, ArcQuest, ArcReader, ArcScan, ArcScene, ArcSchool, ArcScripts, ArcSDE, ArcSdl, ArcSketch, ArcStorm, ArcSurvey, ArcTIN, ArcToolbox, ArcTools, ArcUSA, *ArcUser*, ArcView, ArcVoyager, *ArcWatch*, ArcWeb, ArcWorld, ArcXML, Atlas GIS, AtlasWare, Avenue, Business Analyst Online, BusinessMAP, Community, CommunityInfo, Data Automation Kit, Database Integrator, DBI Kit, EDN, ESRI, ESRI BIS, ESRI—Team GIS, ESRI—*The GIS Company*, ESRI—The GIS People, ESRI—The GIS Software Leader, FormEdit, GeoCollector, Geographic Design System, Geography Matters, Geography Network, GIS by ESRI, GIS Data ReViewer, GIS Day, GIS for Everyone, GISData Server, JTX, MapBeans, MapCafé, MapData, MapObjects, Maplex, MapStudio, ModelBuilder, MOLE, MPS-Atlas, NetEngine, PC ARC/INFO, PC ARCPLOT, PC ARCSHELL, PC DATA CONVERSION, PC STARTER KIT, PC TABLES, PC ARCDIT, PC NETWORK, PC OVERLAY, PLTS, Rent-a-Tech, RouteMAP, SDE, Site-Reporter, SML, Sourcebook-America, Spatial Database Engine, StreetEditor, StreetMap, Tapestry, the ARC/INFO logo, the ArcAtlas logo, the ArcCAD logo, the ArcCAD WorkBench logo, the ArcCOGO logo, the ArcData logo, the ArcData Online logo, the ArcEdit logo, the ArcEurope logo, the ArcExplorer logo, the ArcExpress logo, the ArcGIS logo, the ArcGIS Explorer logo, the ArcGrid logo, the ArcIMS logo, the ArcInfo logo, the ArcLogistics Route logo, the ArcNetwork logo, the ArcPad logo, the ArcPlot logo, the ArcPress for ArcView logo, the ArcPress logo, the ArcScan logo, the ArcScene logo, the ArcSDE CAD Client logo, the ArcSDE logo, the ArcStorm logo, the ArcTIN logo, the ArcTools logo, the ArcUSA logo, the ArcView 3D Analyst logo, the ArcView Data Publisher logo, the ArcView GIS logo, the ArcView Image Analysis logo, the ArcView Internet Map Server logo, the ArcView logo, the ArcView Network Analyst logo, the ArcView Spatial Analyst logo, the ArcView StreetMap 2000 logo, the ArcView StreetMap logo, the ArcView Tracking Analyst logo, the ArcWorld logo, the Atlas GIS logo, the Avenue logo, the BusinessMAP logo, the Community logo, the Data Automation Kit logo, the Digital Chart of the World logo, the ESRI Data logo, the ESRI globe logo, the ESRI Press logo, the Geography Network logo, the MapCafé logo, the MapObjects Internet Map Server logo, the MapObjects logo, the MOLE logo, the NetEngine logo, the PC ARC/INFO logo, the Production Line Tool Set logo, the RouteMAP IMS logo, the RouteMAP logo, the SDE logo, The Geographic Advantage, The World's Leading Desktop GIS, *Water Writes*, www.esri.com, www.esribis.com, www.geographynetwork.com, www.gis.com, www.gisday.com, and Your Personal Geographic Information System are trademarks, registered trademarks, or service marks of ESRI in the United States, the European Community, or certain other jurisdictions.

Other companies and products mentioned herein may be trademarks or registered trademarks of their respective trademark owners.

Hello, and thanks for downloading this instructional podcast from ESRI. My name is Owen Evans, and I work in ESRI's technology center in Vienna, Virginia. This podcast is about working with geodatabase replication.

Over the next 10 minutes or so, I'll give you some ideas that will help you plan for implementing replication, and also describe the replica creation and synchronization processes. This discussion will be useful for GIS managers, geodatabase administrators, and ArcGIS Desktop users that maintain or use data that may be replicated. I'd recommend first listening to the introduction to geodatabase replication podcast (i.e., *Geodatabase Replication: An Overview*), if you haven't already.

As with any new undertaking, it's a good idea to start by doing some research and planning. A great place to start is by downloading the white paper on support.esri.com called, *An Overview of Distributing Data with Geodatabases*. The ArcGIS Desktop Help is also an excellent resource for learning the ins and outs of replication. You can find the replication topics under Geodatabases and ArcSDE, in the Data management workflows, transactions, and versioning section, under Managing Distributed Data.

Once you learn about the concepts and mechanics of replication from the help and the white paper, think about how replication will fit into your overall workflow. For example, ask yourself some questions like, "How might incorporating edits from another department or organization affect my current workflow?" and "What version of the parent geodatabase should be replicated?" If you replicate a version other than default, this allows you to sequester changes that have been received from other replicas before posting them to the versions that are critical to your local workflow. This is a vital part of performing QA/QC in any versioned environment, and since replication simply extends the version tree across geodatabases, this concept applies here as well.

Drawing a diagram that shows which geodatabases will be involved in replication, and their associated versions, will be a valuable planning tool. Use arrows to indicate how replicas will be synchronized, and how edits will be reconciled and posted between versions. Use this diagram to discuss your existing workflow with employees that are involved in spatial data editing and geodatabase administration activities, and decide how this will be integrated with your proposed replication workflow.

Another planning activity I would recommend is to set up a replication playground consisting of several personal ArcSDE geodatabases. These geodatabases can all be on one computer, or spread across several computers. Load your test data into one of these geodatabases and then create versions, replicate data,

make edits, and synchronize the replicas. If you plan to have two-way replicas, you may also find it useful to create some conflicting edits so you can become familiar with the synchronization conflict resolution procedure, which is described in the ArcGIS Desktop Help.

Speaking of synchronization, another matter to think about is your synchronization schedule. If there are only two replicas participating in your workflow, this is simply a matter of evaluating how often changes are made to the data, and how current you want the data to be on each geodatabase. For workflows with more than two participating geodatabases, you'll have to think not only about the frequency of synchronization, but also the order. I'll say more about synchronization schedules a little later in this podcast.

Once your high-level planning has been completed, you're ready to prepare your data for replication. To replicate data, it must be registered as versioned. Geodatabase replication is implemented using the versioning transaction model that has been an established part of ESRI's ArcSDE product for many years. At the 9.2 release, ArcSDE technology is integrated into ArcGIS Server and ArcGIS Desktop.

Users that intend to implement geodatabase replication should be familiar with versioning concepts, such as reconcile and post, conflict resolution, and compression. To reference the ArcGIS Desktop help again, the section called Working with versioned data is a great place to start learning about these topics. And Nicole's podcast on the top five versioning myths is also a great resource on this subject. And there are also several versioning white papers that can be downloaded from support.esri.com.

In addition to being versioned, the data that is to be replicated must also have an attribute that stores global IDs. These are IDs that identify each feature uniquely across all geodatabases involved in replication. The global ID attribute is a geodatabase-maintained attribute and is different from an attribute of type GUID. There is a tool in ArcCatalog that allows you to right-click a feature class, feature dataset, or table, and easily add global IDs. This only needs to be done once on a feature class. Once a global ID attribute is added to a feature class, all existing and new features will automatically be assigned a global ID.

Now that you have planned for replication and prepared your data, you are ready to establish your replicas. The easiest way to create a replica is to start by adding the data you want to replicate to a map document. Both feature classes and geodatabase tables can be replicated. When you add data to a map document, you are often connected to the default version of your data; however, you can choose to

replicate any version of your parent geodatabase. To replicate a version other than default, simply change the version your workspace is pointing to in the table of contents. The version you are connected to when you create the replica is the version where the changes are applied when you synchronize.

The Create Replica wizard, which can be launched from the Distributed Geodatabase toolbar, will walk you through all the decisions you need to make to define a replica. However, before you launch this wizard, you want to define any filters that you are going to employ in your replica definition, such as the selection set, definition query; or you might want to navigate your map to the spatial extent of the data that you wish to replicate. In this wizard, you will do three things. First, you will select the type of replica you would like to create. Second, you will select the geodatabase that will host the child replica. And third, you will select any filters which will define the data that will be included in the replica. Filters can be spatial, defined by your map extent or graphic, or you can leverage definition queries or selection sets like I mentioned earlier.

In a previous podcast on replication, I discussed the types of replica relationships that are supported at ArcGIS 9.2. In addition to those capabilities, ArcGIS 9.3 will also support one-way replication to file and personal geodatabases. In other words, file geodatabases and personal geodatabases for Microsoft Access will be able to host the child replica in a one-way replica pair.

The geodatabase that will host the child replica can be chosen in several ways. If that geodatabase is accessible on your local network, you can simply browse to it in ArcCatalog. If the geodatabase is not on your local network, and has been published as a geodata service using ArcGIS Server, you can browse to that GIS server in ArcCatalog, and then select the geodata service. I'll talk more about how to create and use geodata services in a future podcast.

Once you have selected the replica type and where the child replica will be housed, you must select which datasets you want to replicate. You can also choose to replicate any related data, if that is appropriate. Note that only related data, defined by relationship classes in your geodatabase, will be replicated. Joins and relates in your map document are not supported for replication at this time.

Once you complete the Create Replica wizard, the data that matches the filters you defined is extracted, copied to the child geodatabase, and the appropriate information is written to the geodatabase tables in order to define your replica. By the time your replicas are established, the bulk of the work has been completed. You have thought about how replication will fit into your overall workflow and what data

you'd like to replicate. You can think of the time after your replicas are created as maintenance mode. Your planning will pay off in the form of an enhanced understanding of how your new workflow should be implemented.

After the data has been edited on one or more of your replica geodatabases, you'll want to perform a synchronization to transfer those changes from one geodatabase to the others. Synchronization can be initiated using the Synchronize Changes tool, which can be found on the Distributed Geodatabase toolbar, or by right-clicking one of the geodatabases in ArcCatalog, and selecting Synchronize Changes from the Distributed Geodatabase menu.

Synchronization is fault-tolerant, so if a synchronization is interrupted—for example, if the network goes down in the middle of sending changes—the operation is rolled back to the point before the synchronization was initiated. When the network comes back up, you can perform the sync again.

Earlier I mentioned that you will need to determine an appropriate synchronization order if your replication workflow has more than two participating geodatabases. Imagine an example with a main office that has replicated some feature classes to two regional office geodatabases, which we will call A and B. If A synchronizes with the main office, and then B synchronizes with the main office, geodatabase A still does not have the most recent changes from office B. You'd have to do another synchronization with A to make sure all three geodatabases are in sync. This example assumes the changes are sent in both directions during each of the synchronizations.

A mutligedatabase synchronization schedule might sound complicated, but once the appropriate order is determined, you can create a geoprocessing model that will perform each of the required synchronizations in the appropriate order. This model can be run manually from ArcToolbox, or it can be exported to a Python script and set to run automatically in the Windows scheduler.

In this podcast, I discussed several topics related to geodatabase replication. I started off with some suggestions on how to plan for implementing a replication workflow. These included generating a diagram of how your replicas will fit in to your version tree, and also practicing synchronization in a test environment. I then covered how to prepare your data by registering it as versioned and adding global IDs, and summarized some of the options that you have during the replica-creation process. Finally, I offered some ideas to think about when synchronizing the replicas you have established.

In a future podcast, I will present some tips and tricks for working with replication, and also describe some example workflows that may get you thinking about how you can employ replication in your organization.

Thanks for listening and stay tuned for future podcasts from ESRI.