

# Lessons Learned with Geodatabase Replication



**ESRI**

**Ron Florence**

Technical Analyst  
rflorence@esri.com

[www.esri.com](http://www.esri.com)

227 N. Loop 1604 E. Suite 100  
San Antonio, Texas 78232, USA  
Phone: 210.499.1044



**ESRI**

**Eric Stauber, GISP**

Technical Analyst  
estauber@esri.com

[www.esri.com](http://www.esri.com)

227 N. Loop 1604 E. Suite 100  
San Antonio, Texas 78232, USA  
Phone: 210.499.1044



**C. Michael Parma, GISP**

GIS Coordinator

[mparma@nbtexas.org](mailto:mparma@nbtexas.org)

424 S. CASTELL AVENUE, NEW BRAUNFELS, TEXAS 78130  
TEL 830-221-4337 . FAX 830-608-2112

**SCAUG Conference**

**Austin, Texas**

**April 1, 2009**



# Presentation Outline

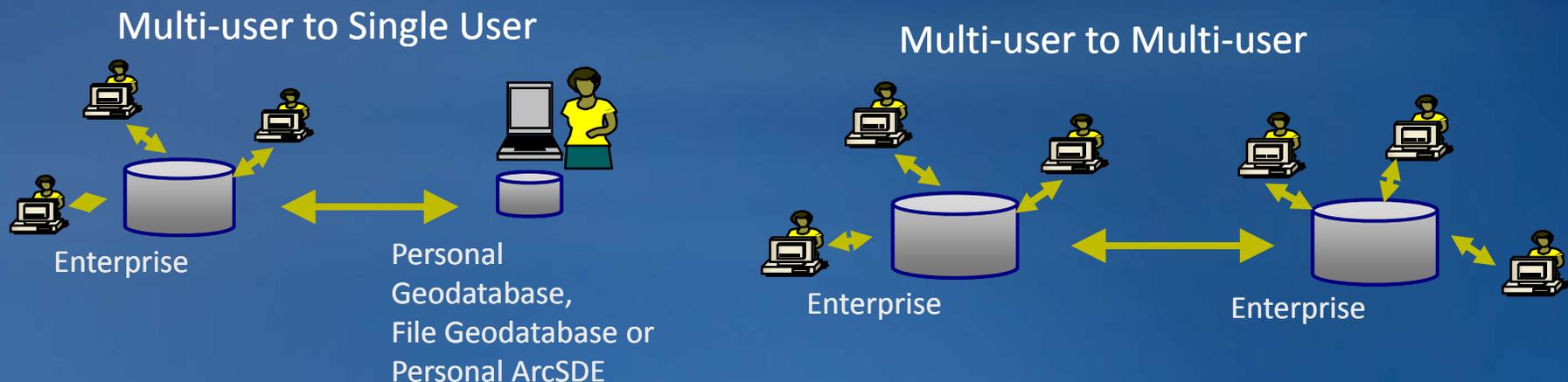
- **What is replication?**
- **City of New Braunfels Use Case**
- **Replication Demo**
- **Lessons Learned**
- **Q&A**

# What is Geodatabase Replication?

- **Allows you to distribute copies of data across 2 or more geodatabases**
- **You can edit the databases independently and synchronize them as needed**
- **Released at 9.2 - Builds upon disconnected editing from earlier releases (8.3)**
- **Not the same as RDBMS replication**

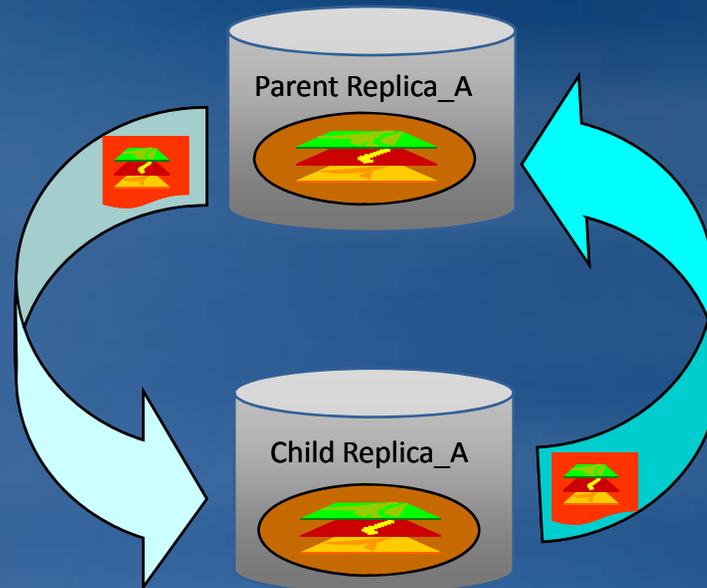
# Geodatabase Replication Workflows

- Workflows can involve Multi-User geodatabases and single user geodatabases
  - Multi-user geodatabase – Multi-user ArcSDE geodatabase accessed locally or remotely through ArcGIS server
  - Single user geodatabase – Personal ArcSDE, file geodatabase or personal geodatabase on a local machine



# Geodatabase Replication Concepts

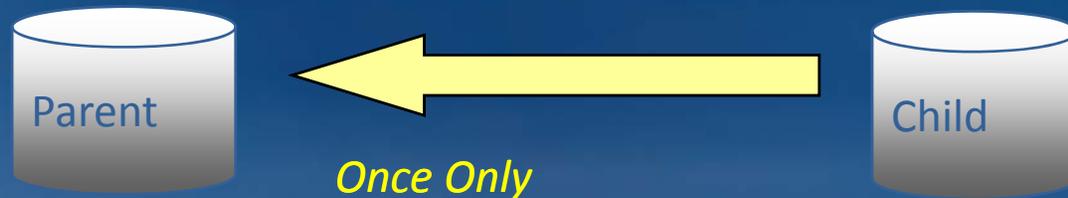
- **Basic replication relationship is between two replicas :**
  - A Child Replica is created from a Parent Replica.



- **You can replicate :**
  - A specific version
  - Specific datasets
  - A subset of features in the chosen datasets

# Geodatabase Replication Types

Check-out /  
Check-in



Two-way

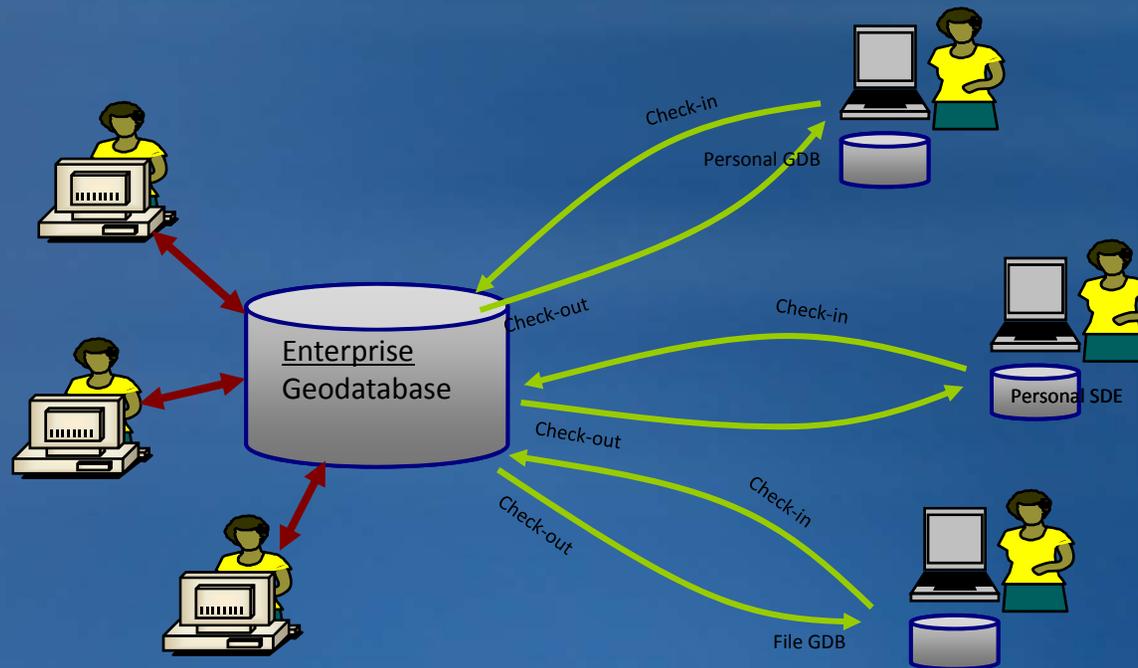


One-way



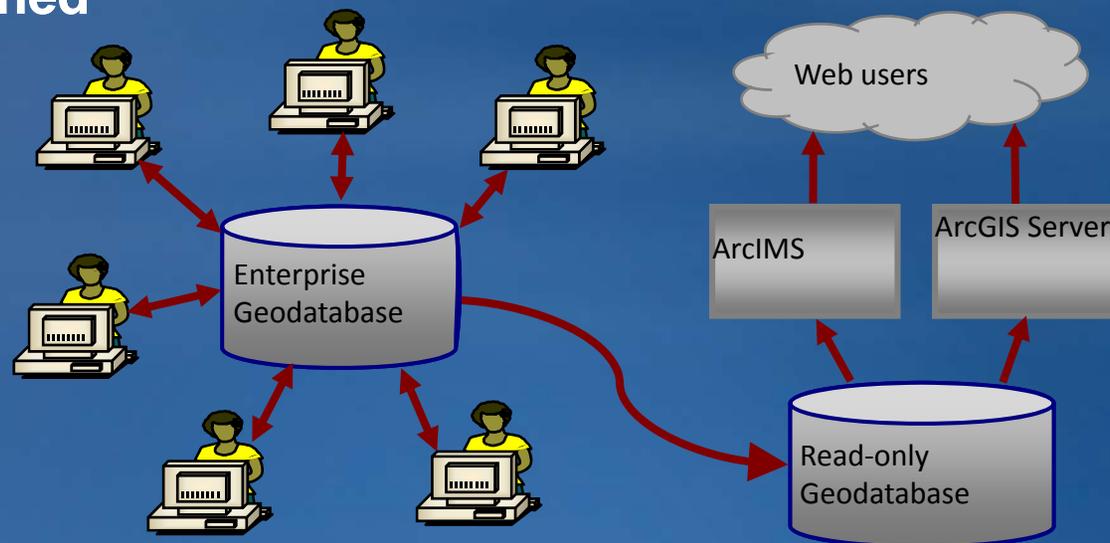
# Check-out / Check-in Replication

- Called “Disconnected Editing” in ArcGIS 8.3 to 9.1
- Child replica can be hosted in a Personal Geodatabase, File Geodatabase or ArcSDE Geodatabase (only ArcSDE can host the Parent)



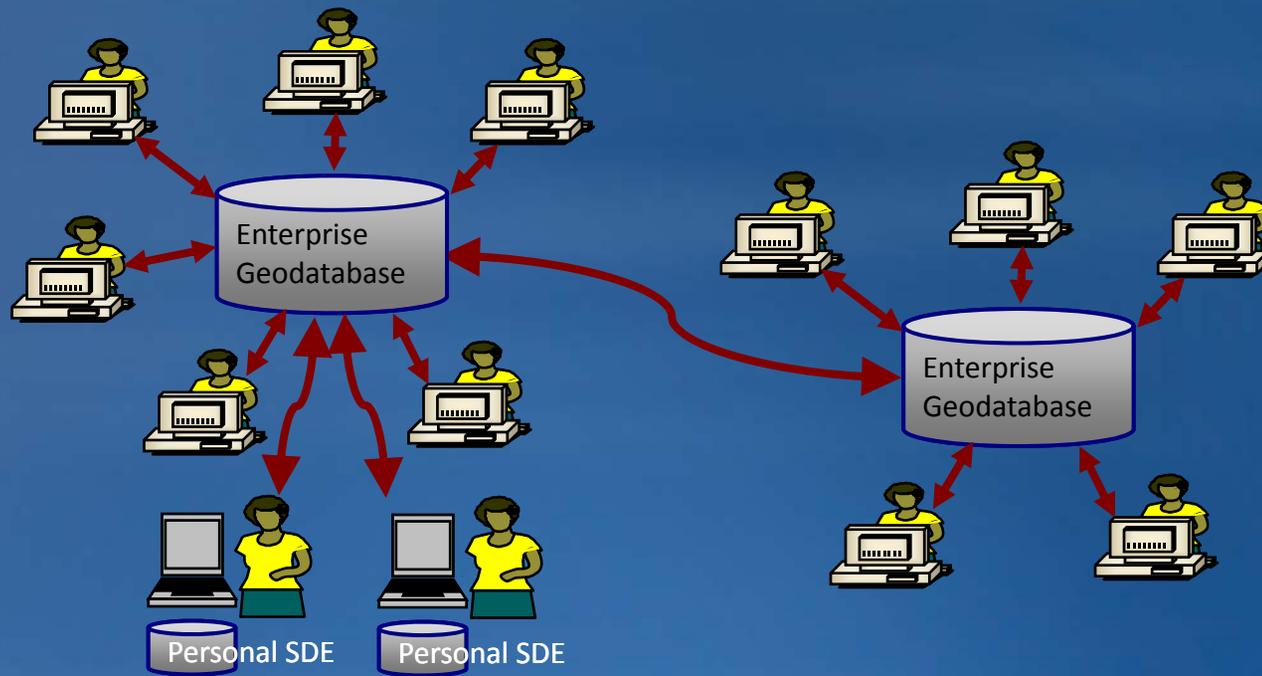
# One-way Replication

- **Child replica is considered read-only**
  - Child data may be overwritten
- **No system versions on the child replica**
- **Choose between 2 model types:**
  - Full – Supports complex types (Geometric Networks and Topologies) and requires the child replicas data to be versioned
  - Simple – Child replica's data is simple and does not need to be versioned



# Two-way Replication

- Requires ArcSDE geodatabases and versioned data
- Can use 2-way replication with personal ArcSDE



# Geodatabase Replication – Data Requirements

- You must have write access to the data
- All data must be registered as versioned without the option to move edits to base
- Two-way and One-way replicas
  - Each dataset must have a GlobalID column
  - All spatial data must be stored in a high precision spatial reference
- Use wizards in ArcMap or ArcToolbox to create or register replicas

# Replica Synchronization

- Applying data changes across replicas
- Replicas can be Synchronized in either both directions or just a single direction
- Synchronization is based on exchanging messages and is fault tolerant
- Uses versioning



# Replica Synchronization

- **Distributed Geodatabase Tools**
  - ArcMap toolbar or ArcToolbox
- **Connected Synchronization**
  - Message exchange is automatic
- **Disconnected Synchronization**
  - Message exchange is performed by the end user
  - Operations are performed by export, file transfer and import
    - Example: to synchronize, export changes to a delta XML file, transfer the file (ftp, CD through the mail, etc.), have the file imported on the relative replica when it arrives

# Working through errors

- System is designed to stay consistent
- If the system fails during a synchronization, it is rolled back to it's previous state
- If a data changes message is lost in a disconnected system, the next message will contain changes from the lost message and any new changes
- Replica log can be used to get error information about a synchronization

# Requirements

- **ArcSDE 9.2 or above on both sides**
  - Child replicas must be same or newer version than parent
  - Personal SDE is sufficient (SQL Server Express)
  - One-way replication to PGDB and FGDB supported at 9.3
- **ArcEditor or ArcInfo on both sides**
- **ArcGIS Server Standard or Advanced Edition (if creating Geodata Web Service)**
- **Collaboration and Agreement**
  - Common data model at creation time
- **Patience**



# Business Case

- **Enterprise License Agreement**
  - ArcInfo desktop clients
  - ArcSDE & ArcGIS Server Advanced to publish Web ADF site
    - Serving 35,000+ map requests monthly
- **Utilize 911 District for street centerline data**
- **Relational database structure for street centerlines and address points (StreetNameID key and tblStreetName lookup table)**
- **Use spatial views to manage display of attributes**
- **Use a locator service built upon streets for geocoding**
- **Limited staff resources**
- **Other management systems coming online (land/asset, pavement)**
  
- **A lot rides on our streets!**

# Previous Update Process

- **Bexar Metro 9-1-1 Network District**



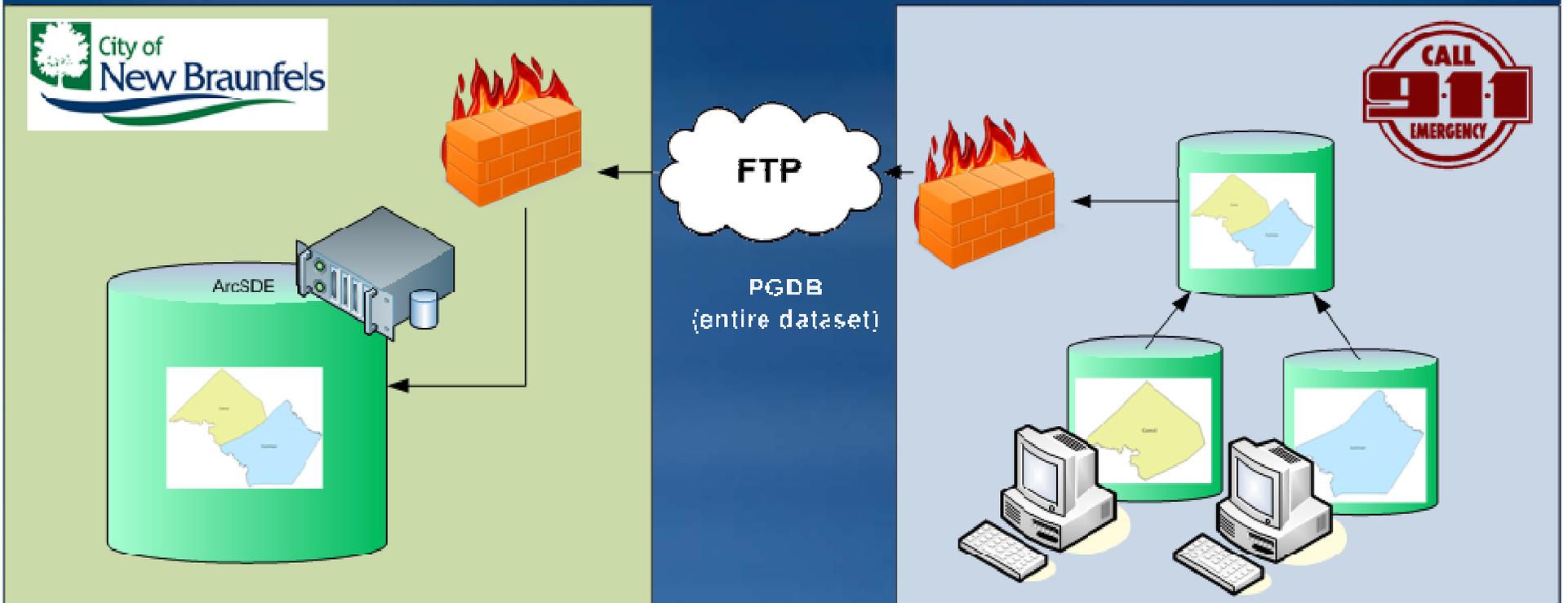
- Merged separate PGDBs
- Loaded to FTP site

- **City of New Braunfels**



- Download PGDB from FTP site
- Stop all published map services
- Kick all users out of ArcSDE
- Create temporary backup of feature class and street name table
- Visually confirm new data structure matched existing
- Overwrite entire feature class and table (convert feature class from PGDB to ArcSDE, table converted through XML workspace)
- Reapply security permissions
- Delete and recreate spatial views
- Delete and recreate locator services
- Restart and verify all services operational
- Archive old feature class and table
- Delete temporary backups from ArcSDE

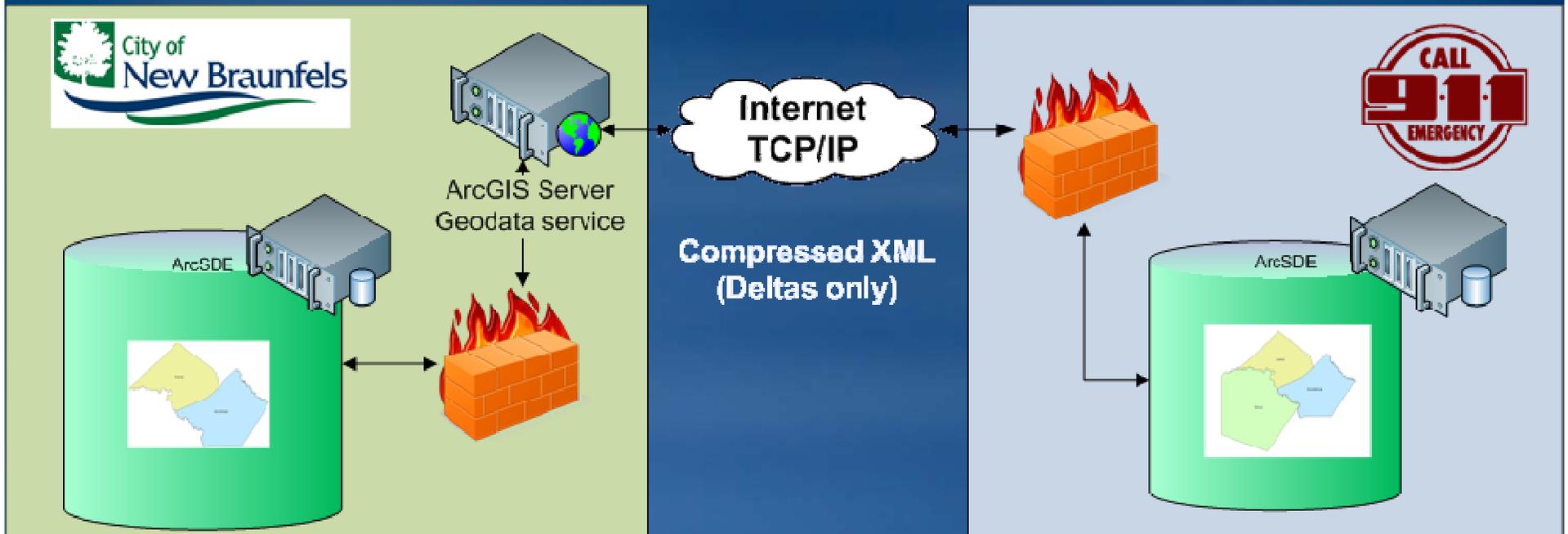
- OLD PROCESS



# Desired Update Process

- **The City of New Braunfels desired a process for the update of street centerline and street name changes which:**
  - provided a direct import of data (no more conversion)
  - allowed for uninterrupted map and database services
  - took full advantage of the City's existing GIS investments
  - reduced network traffic
  - maintained a consistent data structure
  - could be automatically scheduled
  - reduce data management workload
- **The answer: geodatabase replication**

- New Process



# Production Environment

- **New Braunfels**
  - **ArcGIS Server 9.3 SP1**
    - **Geodata service**
  - **Firewall**
  - **ArcSDE 9.3 SP1 on MS SQL Server 2005**
- **Bexar Metro 911 Network District**
  - **Firewall**
  - **ArcSDE 9.3 SP1 on MS SQL Server 2000**

# Initialization

- **Bexar Metro 911 Network District**
    - Migrate from PGDB to ArcSDE
    - Register as versioned
    - Generate GlobalIDs
    - Export ArcSDE Feature Dataset using XML Workspace
  - **City of New Braunfels**
    - Import XML Workspace to ArcSDE
    - Register as versioned
    - Set security privileges
    - Create a QA/QC version
    - Create and publish geodata service
    - Create replica to XML message
- Import replica XML message
  - Edit replication connector

# Result

- **The City of New Braunfels desired a process for the update of street centerline and street name changes which:**
  - ✓ provided a direct import of data (no more conversion)
  - ✓ allowed for uninterrupted map and database services
  - ✓ took full advantage of the City's existing GIS investments
  - ✓ reduced network traffic
  - ✓ maintained a consistent data structure
  - ❓ could be automatically scheduled
  - ✓ reduce data management workload
- **Formerly a monthly, manual update process which took 2-3 hours and interruption of all services**
- **Now, a weekly process taking ~15 minutes, even manually, and no interruption of services**

# Replication Demo



# Geodatabase Replication – Getting Started

- **Anticipate future needs when defining the data to replicate**
- **Have a well defined data model before creating replicas**
- **Choose the right replica type**
  - Consider 2-way replicas with personal ArcSDE instead of check-out replicas
  - Use 1-way replicas over 2-way replicas when possible

# Geodatabase Replication – Getting Started

- **Use models or scripts for replicas you plan to create and synchronize on a regular basis**
  - You can use the Create Replica and Create Replica From Server geoprocessing tools in ModelBuilder
- **Consider using the following replica creation options**
  - Re-use schema (check-out replicas) – uses existing schema
  - Register only – replicates pre-copied data
  - Relationship classes processing is optional
- **Schedule Synchronizations**
  - You can use geoprocessing models exported to Python or VB script and the Windows Scheduler
  - Consider synchronization order

# Lessons Learned

- **Requires thorough pre-planning**
  - **Common data structure**
  - **Not all feature classes or attribute fields need be replicated**
  - **Pre-seed data on both databases**
  - **Ensure IT cooperation on both sides!**
- **TEST! TEST! TEST!**

# Lessons Learned

- **Creator of the replica is always the parent of the replica**
- **Data provider may or may not be the parent of the replica**
- **Geodata service must be published using local data owner account**
  - **Enable ArcGIS Server security or use firewall**
- **Only the replica parent can replicate to version other than DEFAULT**
- **One-way replication must be from parent to child**

# Lessons Learned

- **Not quite “set it and forget it”**
  - Vast improvement over Sneakernet
    - Data updated without interruption of services
    - Smaller message traffic
  - Have a plan for recreating replica with reseeding data
  - Overcome with manual transmission of XML messages
  - Use generation levels to help troubleshoot
  - Still, reduced maintenance
- **Requires spirit of cooperation and trust between entities**

# Resources

- **Introduction to Geodatabase Replication (free online seminar):**

[http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product\\_ID=867](http://training.esri.com/acb2000/showdetl.cfm?DID=6&Product_ID=867)

- **2008 ESRI User Conference DVDs**
- **ArcGIS Server 9.x Web Help**
- **How to find help on Geodatabase Replication**  
<http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=34131>
- **Using Compress on ArcSDE GDBs with Replication**  
[http://resources.esri.com/help/9.3/geodatabase/Blog/GDB\\_Rep\\_Compress.pdf](http://resources.esri.com/help/9.3/geodatabase/Blog/GDB_Rep_Compress.pdf)
- **TCP/IP Tuning:**  
<http://support.esri.com/index.cfm?fa=knowledgebase.techarticles.articleShow&d=35971>

# Questions?



**ESRI**

**Ron Florence**

Technical Analyst  
rflorence@esri.com

[www.esri.com](http://www.esri.com)

227 N. Loop 1604 E. Suite 100  
San Antonio, Texas 78232, USA  
Phone: 210.499.1044



**ESRI**

**Eric Stauber, GISP**

Technical Analyst  
estauber@esri.com

[www.esri.com](http://www.esri.com)

227 N. Loop 1604 E. Suite 100  
San Antonio, Texas 78232, USA  
Phone: 210.499.1044



**C. Michael Parma, GISP**

GIS Coordinator

[mparma@nbtexas.org](mailto:mparma@nbtexas.org)

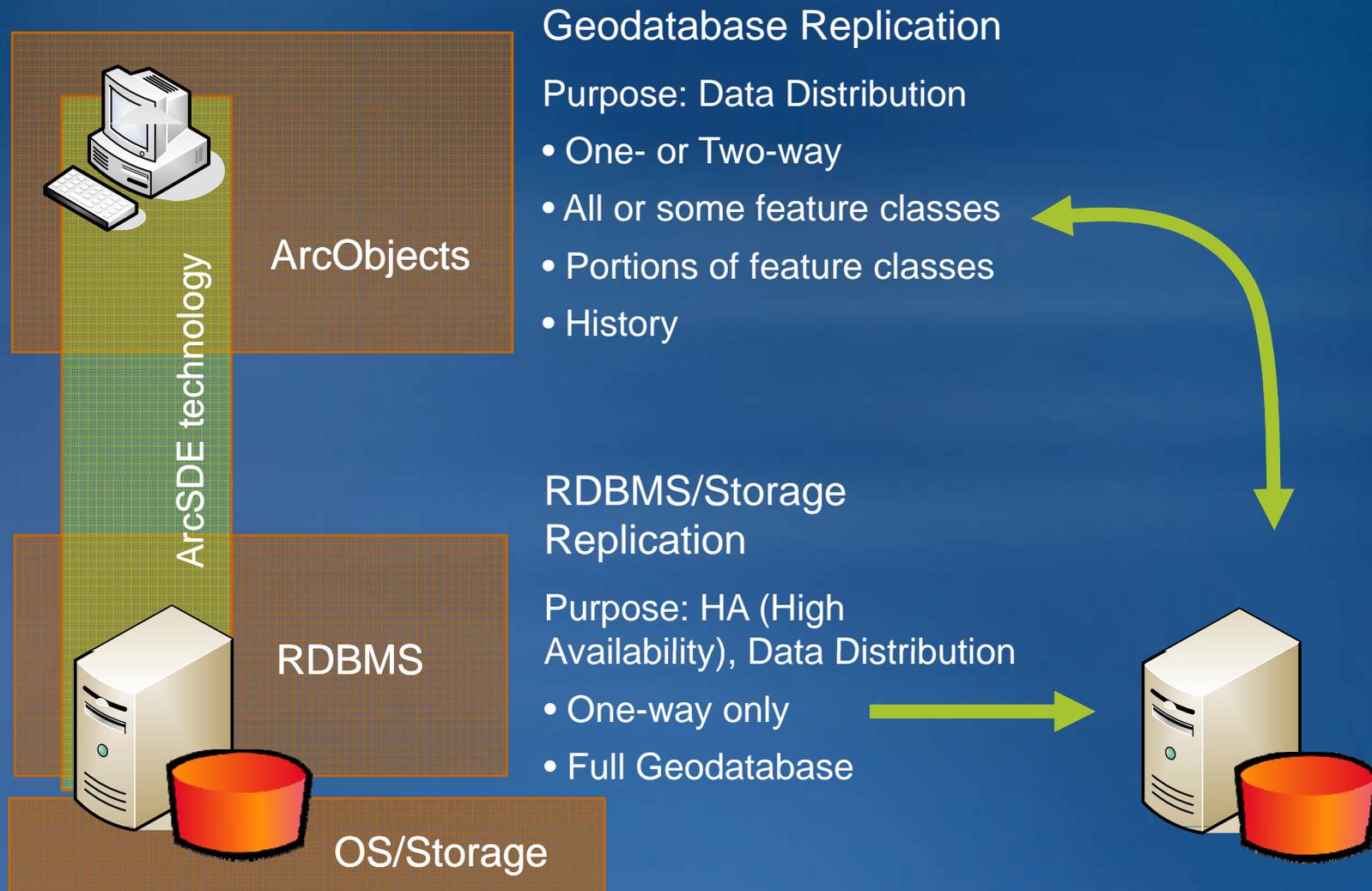
424 S. CASTELL AVENUE, NEW BRAUNFELS, TEXAS 78130  
TEL 830-221-4337 . FAX 830-608-2112

# Thank you!!





# RDBMS vs. Geodatabase Replication



# DBMS Replication with Geodatabases

- **Geodatabase replication does not use DBMS replication**
- **Requirements and limitations of DBMS replication**
  - Requires knowledge of how the geodatabase/ArcSDE system tables work
  - No tools provided in ArcGIS to support it
  - Limited support for cross DBMS replication
  - Does not support or has limited support for complex geodatabase data types and limited filters to define the data to replicate
- **Advantages of DBMS replication**
  - Can work with non-versioned data
  - Can replicate entire database
  - Can be configured to provide synchronous replication

# Geodatabase Replication - Data Requirements

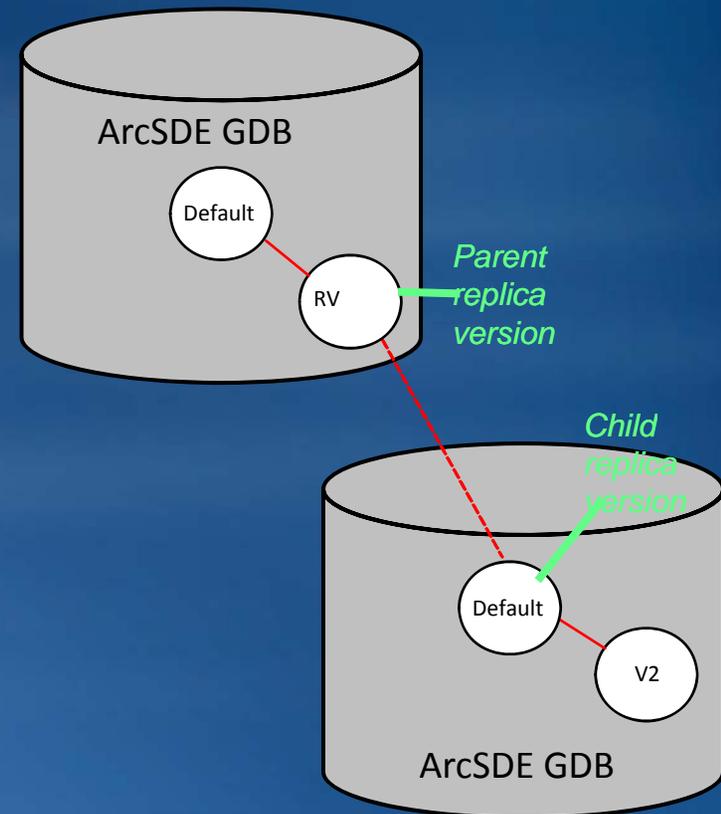
- You must have write access to the data
- All data must be registered as versioned without the option to move edits to base
- Two-way and One-way replicas
  - Each dataset must have a GlobalID column
  - All spatial data must be stored in a high precision spatial reference

## Replica Creation – What Data Types Can be Replicated

- In general, if it supports the full versioning data model it can be replicated
- Data Types
  - Fully supported – Simple Feature Classes and Tables, Geometric Networks, Topologies, Representations, Relationship Classes and Annotation
  - Source data can be replicated -Terrains and Network datasets
  - Copied during creation – Raster datasets and Raster catalogs
  - Not Supported - Survey datasets, Schematics datasets, Locators can not be replicated

# Replica Creation - Versioning

- Edits made to the replica versions are synchronized
- Like extending the version tree to span multiple databases
- 2-way and 1-way replicas
  - Parent replica version can be a named version or default
  - Child replica version is always DEFAULT
- Check-out replicas
  - Parent and child replica versions are named versions



# Schema Changes that can be Applied

	Add	Change	Drop
Field	Y	Y (domains)	Y
Domain	Y	Y	Y
Table/FC	Y	Y (Domain, Add/drop field)	Y
Geometric network	N	N	Y
Topology	N	N	Y
Feature dataset	N	N	Y
Relationship class	N	Y (add/drop field, domain)	Y

# Replica Creation – Defining data to replicate

- Filters and Relationship classes are used to define the data to replicate
- Filters are applied first
  - Spatial – A geometry used to define the area to replicate
  - Selections – Selection sets on feature classes and tables
  - QueryDef – Definition queries applied to individual feature classes and tables
- Additional rows are then added if they are related to the rows in the filter
  - Relationship classes are applied in a single direction and in an optimal order

