GIS

The Geographic Approach for the Nation

ESRI Federal User Conference

Washington, D.C. • February 17–19, 2010

Geodatabases Essentials Part II

An Introduction to ArcSDE Geodatbases

Tim Clark

Session Path

1. Introduction to ArcSDE Geodatabases
   - What is the Geodatabase?
   - The Geodatabase Management Approach
   - Different types of Geodatabase
   - What is an ArcSDE Geodatabase and what are its benefits?
2. Versioning
3. Editing
4. Archiving
5. Distributed Geodatabase

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 Data Management Approach

- The geodatabase is built on an extended relational database.
  - Base relational model
  - Base short transaction model
  - Relational integrity
  - Reliability, Flexibility, Scalability
  - Supports continuous, large datasets
- Built on the simple feature model
  - Open access (OGC, C, COM, SQL)
- Business rules, topology, networks
  - Data Integrity

Geodatabase Data Management Approach …

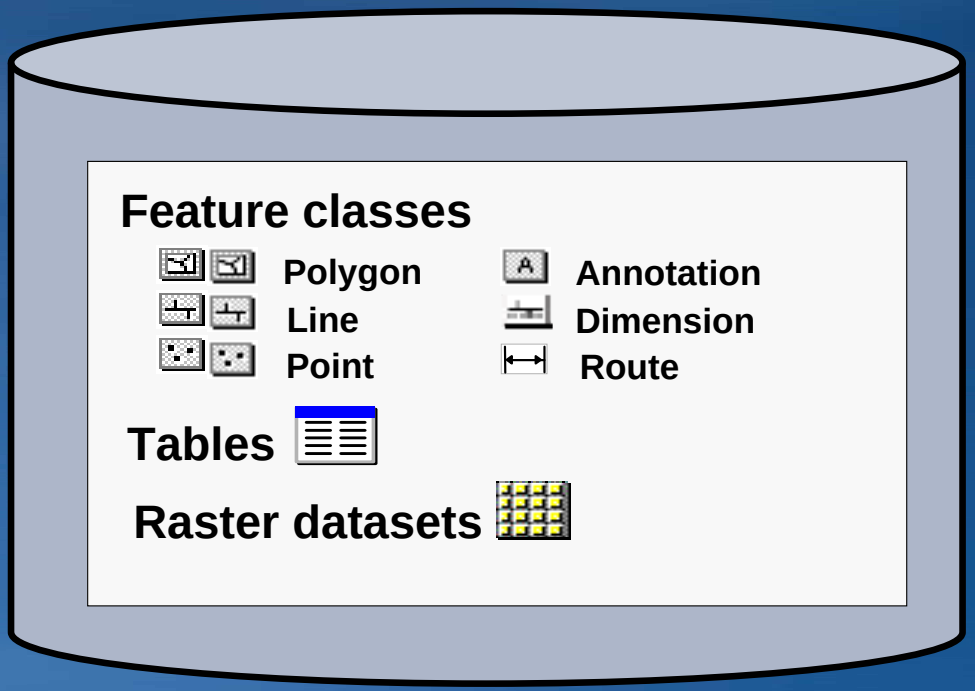
- Simple features + logic
  - All geographic data stored as tables in a DBMS
  - Extend functionality and data integrity
  - Functionality is consistent across DBMS’
- Application logic (software)
  - Works on standard RDBMS tables
  - Implements GIS integrity and behavior

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

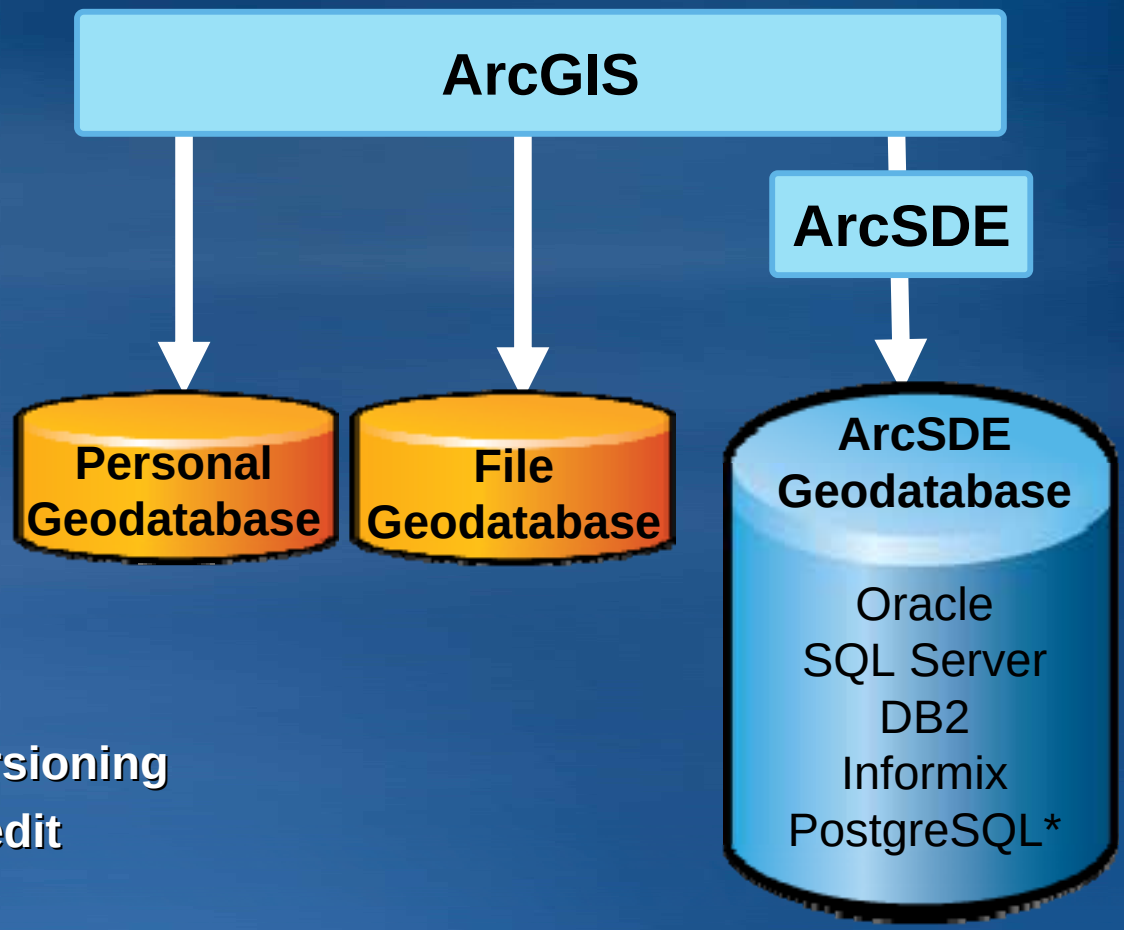
Inside the Geodatabase

- A geodatabase contains datasets.
  - Datasets represent collections of information with a real-world interpretation.
- Datasets have associated information to help manage integrity, behavior, and interpretation
  - Domains, Relational integrity, Topology, Metadata
- Three primary datasets:
  - Feature classes
  - Raster datasets
  - Tables



3 Types of Geodatabases

- **Personal Geodatabase**
  - Single user editing
  - Stored in MS Access
  - Size limit of 2 GB
- **File Geodatabase**
  - 1 TB per table
  - Reduced storage requirements
- **ArcSDE Geodatabase**
  - Stored in an enterprise DBMS
  - Supports multiuser editing via versioning
  - Requires ArcEditor or ArcInfo to edit

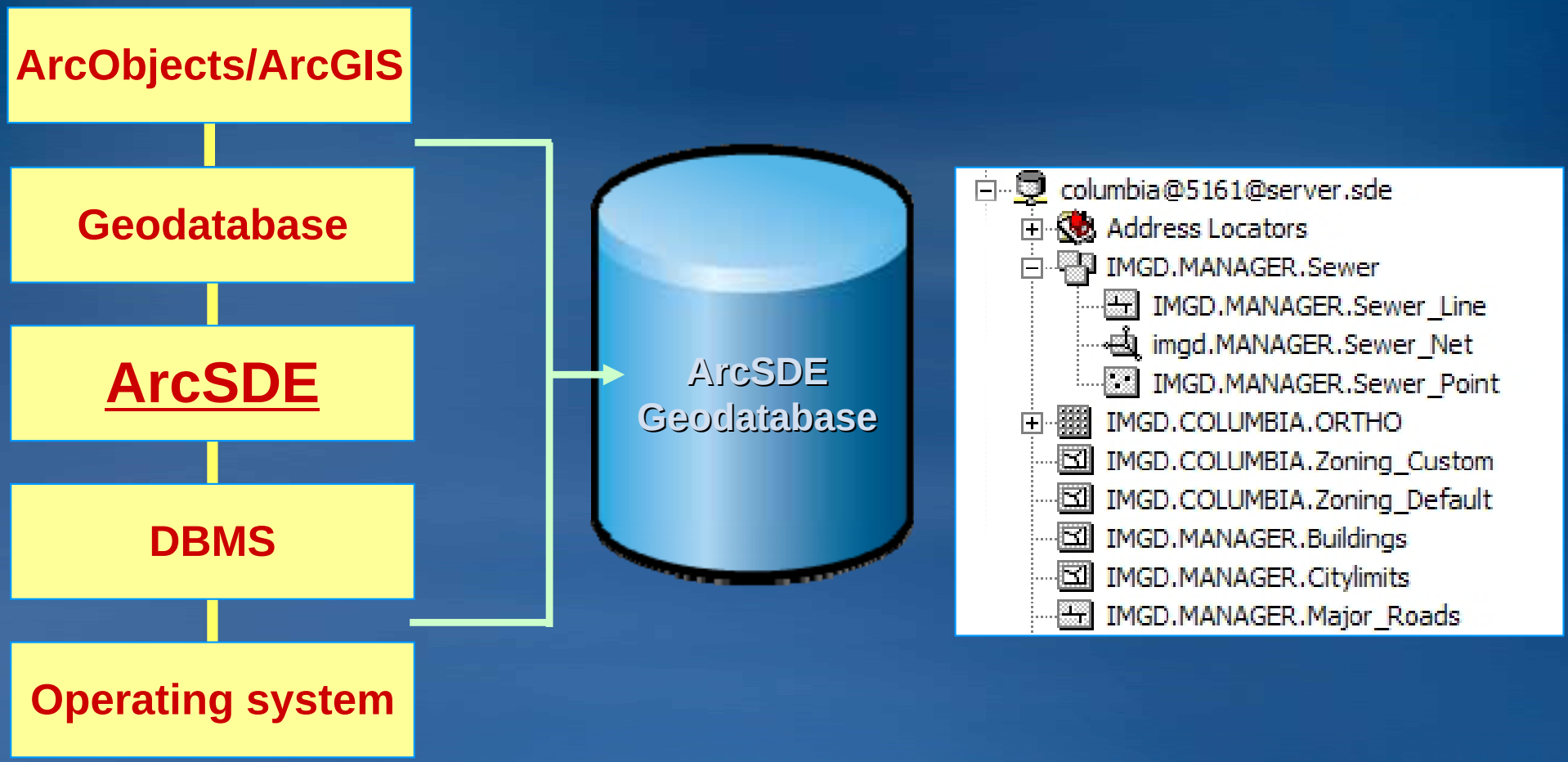


* PostgreSQL support at 9.3

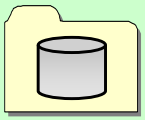
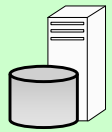
What is an ArcSDE Geodatabase?

- ESRI's technology for accessing and managing geospatial data in relational databases
- ArcSDE Geodatabases are unique in their support of the following capabilities:
  - Open and interoperable across many supported DBMSs
  - Standards based, using as its native data structure the OGC binary simple features standard and the ISO spatial type (for Oracle, IBM DB2, IBM Informix, and PostgreSQL only).
  - Offers support for full, open SQL access to geodatabases stored in Oracle, IBM DB2, IBM Informix, and PostgreSQL.
  - Full support of the Oracle format for feature storage (using Oracle Spatial and Oracle Locator).

How is ArcSDE technology included in ArcGIS?



3 Types of Geodatabases…

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

When do you need an ArcSDE Geodatabase?

- Users need to edit and use their data simultaneously
- Need to manage long transactions and version-based workflows
- Leverage your existing relational databases
- Require high performance and the ability to scale to a large number of users.
- Require the ability to storage extremely large amounts of data

What are the benefits of using an ArcSDE Geodatabase?

- Leverage the underlying DBMS architecture to support:
  - Extremely large, continuous GIS databases
  - Many simultaneous users
  - Long transactions and versioned workflows
  - Distributed and archiving based workflows
  - Relational database support for GIS data management (providing the benefits of a relational database such as scalability, reliability, security, backup, and integrity)
  - Standards-based SQL Types for Spatial when the DBMS supports this capability

Which ArcSDE Geodatabase edition?

	ArcSDE for ArcGIS Desktop	ArcSDE for ArcGIS Server Workgroup	ArcSDE for ArcGIS Server Enterprise
ArcGIS Product	ArcGIS and Desktop Engine*	ArcGIS Server Workgroup	ArcGIS Server Enterprise
Number of users	Max 3 users, 1 editor at any one time	Max 10 clients at one time No limit to the number of connections	Unlimited
Supported DBMS	SQL Server Express 2005	SQL Server Express 2005	Oracle,SQL Server, DB2, Informix, PostGreSQL
Database limits	Max database size 4 Gig 1 GB RAM on a single cpu	Max database size 4 Gig 1 GB RAM on a single cpu	No limits
Administration	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog), ArcSDE Commands, DBMS admin software,

Scale from small, personal systems up to workgroups and very large enterprises

Which ArcSDE Geodatabase edition?

	ArcSDE for ArcGIS Desktop	ArcSDE for ArcGIS Server Workgroup	ArcSDE for ArcGIS Server Enterprise
ArcGIS Product	ArcGIS and Desktop Engine*	ArcGIS Server Workgroup	ArcGIS Server Enterprise
Number of users	Max 3 users, 1 editor at any one time	Max 10 clients at one time No limit to the number of connections	Unlimited
Supported DBMS	SQL Server Express 2005	SQL Server Express 2005	Oracle,SQL Server, DB2, Informix, PostGreSQL
Database limits	Max database size 4 Gig 1 GB RAM on a single cpu	Max database size 4 Gig 1 GB RAM on a single cpu	No limits
Administration	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog), ArcSDE Commands, DBMS admin software,

Scale from small, personal systems up to workgroups and very large enterprises

Which ArcSDE Geodatabase edition?

	ArcSDE for ArcGIS Desktop	ArcSDE for ArcGIS Server Workgroup	ArcSDE for ArcGIS Server Enterprise
ArcGIS Product	ArcGIS and Desktop Engine*	ArcGIS Server Workgroup	ArcGIS Server Enterprise
Number of users	Max 3 users, 1 editor at any one time	Max 10 clients at one time No limit to the number of connections	Unlimited
Supported DBMS	SQL Server Express 2005	SQL Server Express 2005	Oracle,SQL Server, DB2, Informix, PostGreSQL
Database limits	Max database size 4 Gig 1 GB RAM on a single cpu	Max database size 4 Gig 1 GB RAM on a single cpu	No limits
Administration	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog), ArcSDE Commands, DBMS admin software,

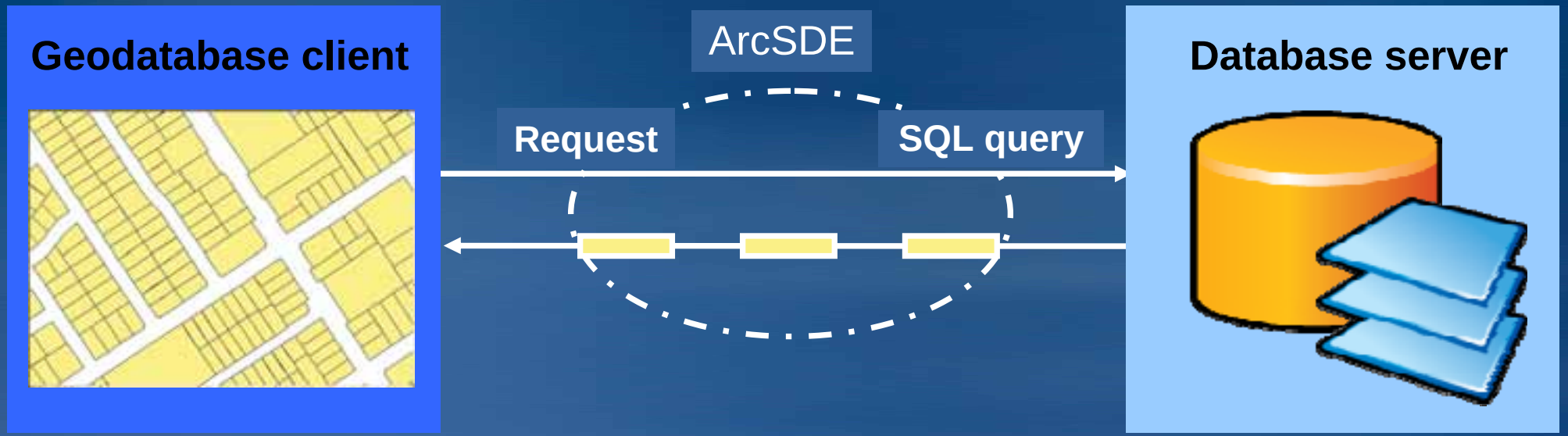
Scale from small, personal systems up to workgroups and very large enterprises

Which ArcSDE Geodatabase edition?

	ArcSDE for ArcGIS Desktop	ArcSDE for ArcGIS Server Workgroup	ArcSDE for ArcGIS Server Enterprise
ArcGIS Product	ArcGIS and Desktop Engine*	ArcGIS Server Workgroup	ArcGIS Server Enterprise
Number of users	Max 3 users, 1 editor at any one time	Max 10 clients at one time No limit to the number of connections	Unlimited
Supported DBMS	SQL Server Express 2005	SQL Server Express 2005	Oracle,SQL Server, DB2, Informix, PostGreSQL
Database limits	Max database size 4 Gig 1 GB RAM on a single cpu	Max database size 4 Gig 1 GB RAM on a single cpu	No limits
Administration	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog)	ArcGIS Desktop (ArcCatalog), ArcSDE Commands, DBMS admin software,

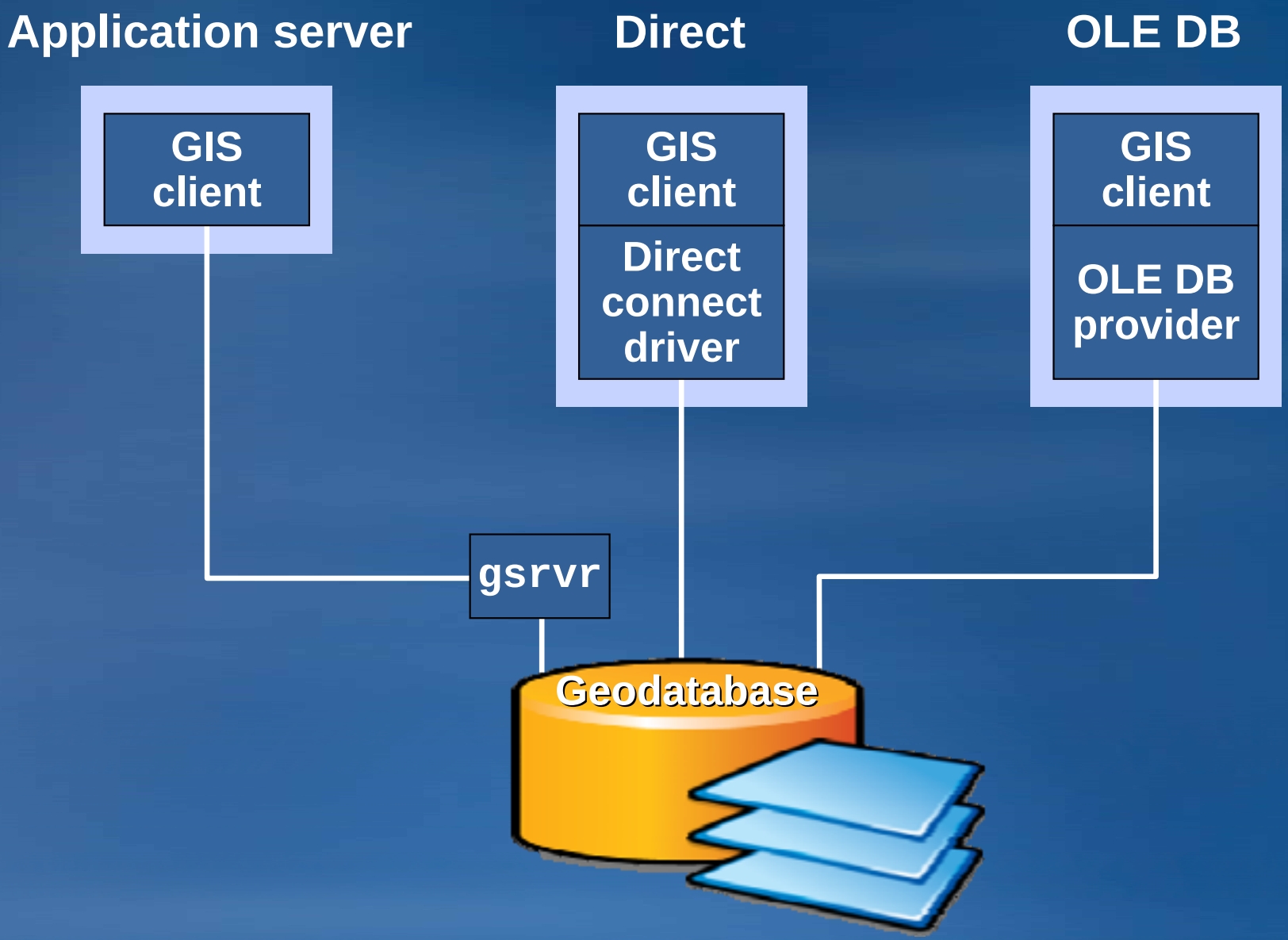
Scale from small, personal systems up to workgroups and very large enterprises

ArcSDE and the Geodatabase
Client server model



- All data accessed over a network
- All data retrieved through SQL queries
  - ArcSDE technology translates
  - Spatial and attribute filters limit rows returned
- ArcSDE technology performs spatial filtering

ArcSDE and the Geodatabase
Connection types

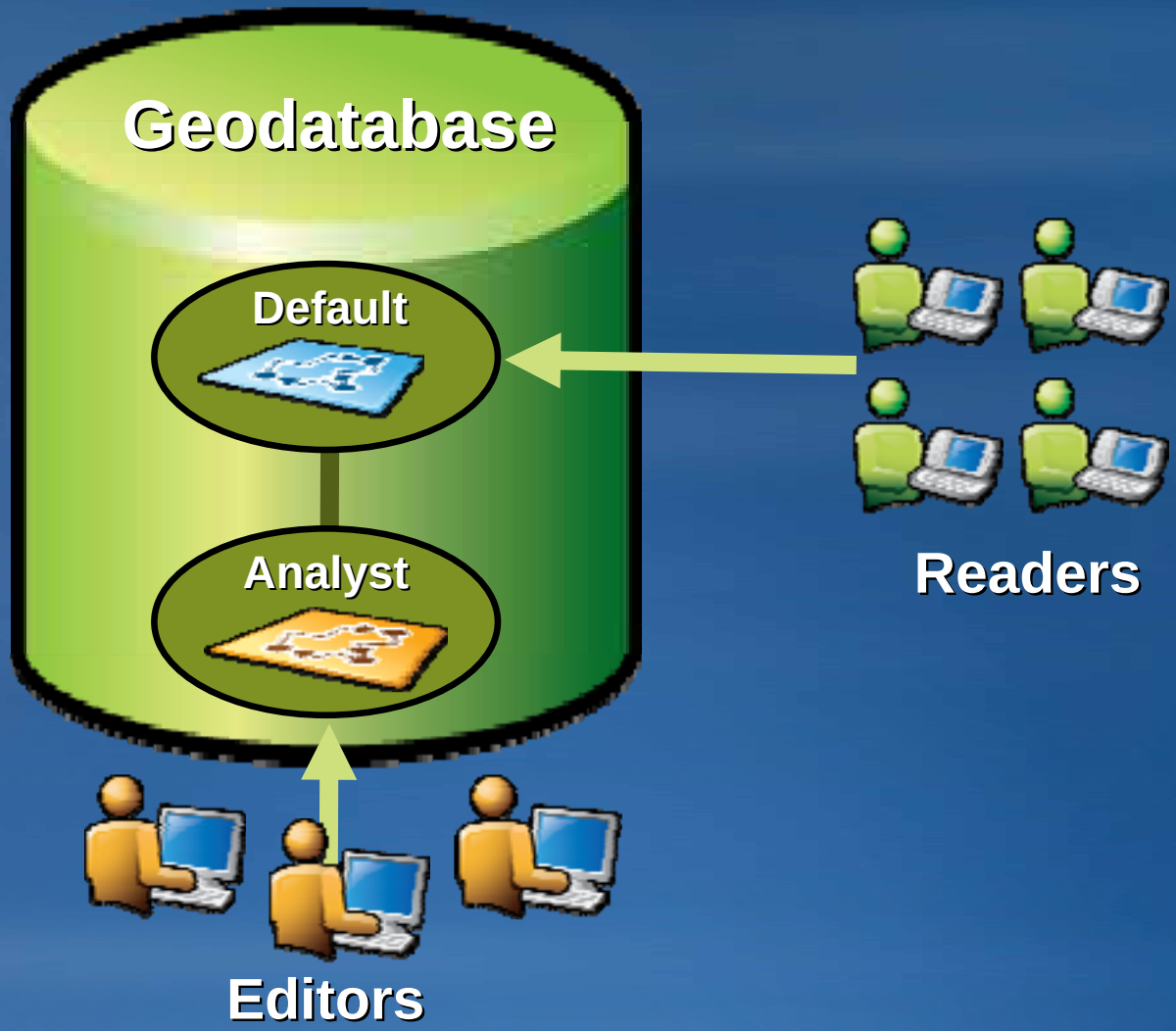


Session Path

1. Introduction to ArcSDE Geodatabases
2. Versioning
   - What is it?
   - How is it used?
3. Editing
4. Archiving
5. Distributed Geodatabase

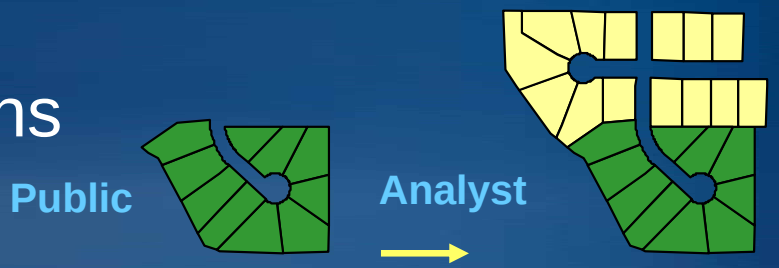
Versioning - What is it?

- 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
    - without applying locks or duplicating data



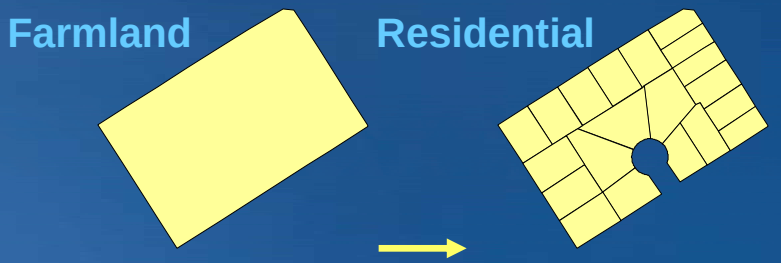
Versioning Workflows – How is it used

- Editing with long transactions
  - Isolate work across multiple sessions
    - Edits do not impact others
    - Example: Parcel editing



- Model what-if scenarios
  - Simulate situations with versions
    - Example: Disaster event planning

- Workflow management
  - Create versions for project stages
    - Example: Land development

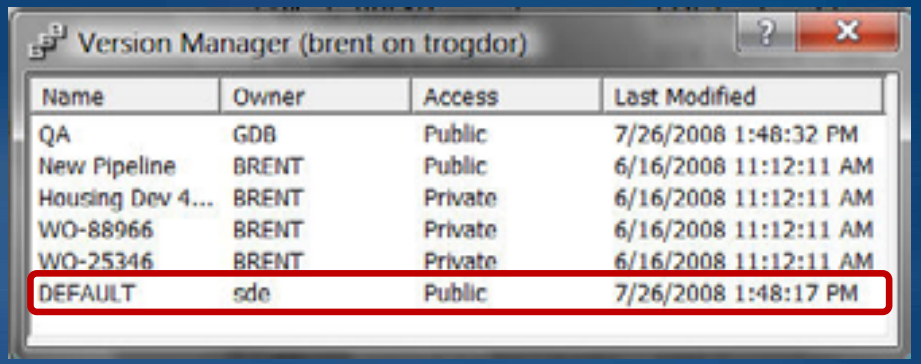


What is a Version?

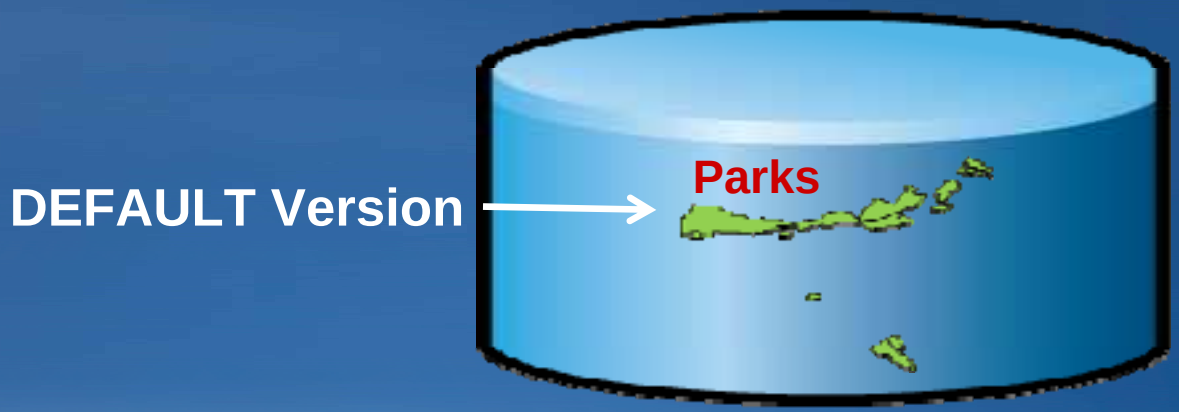
- An alternate representation of the data in a geodatabase
- A connection to a geodatabase is made through a version
- You can not create a version of a feature class
  - Geodatabases are versioned

What is a Version?

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

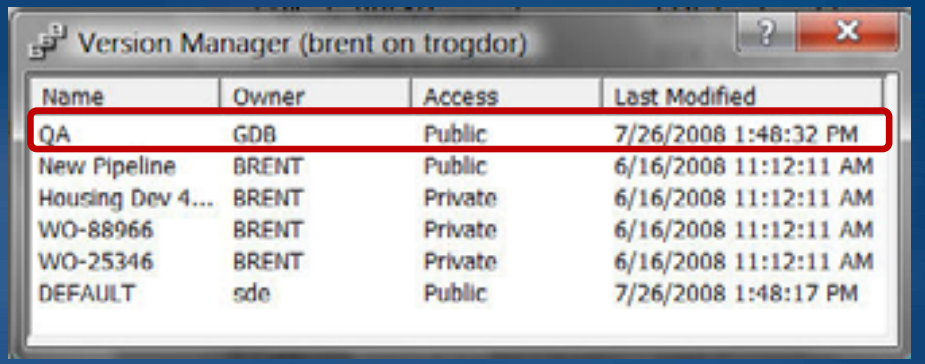


- Versions are not affected by changes occurring in other versions of the database

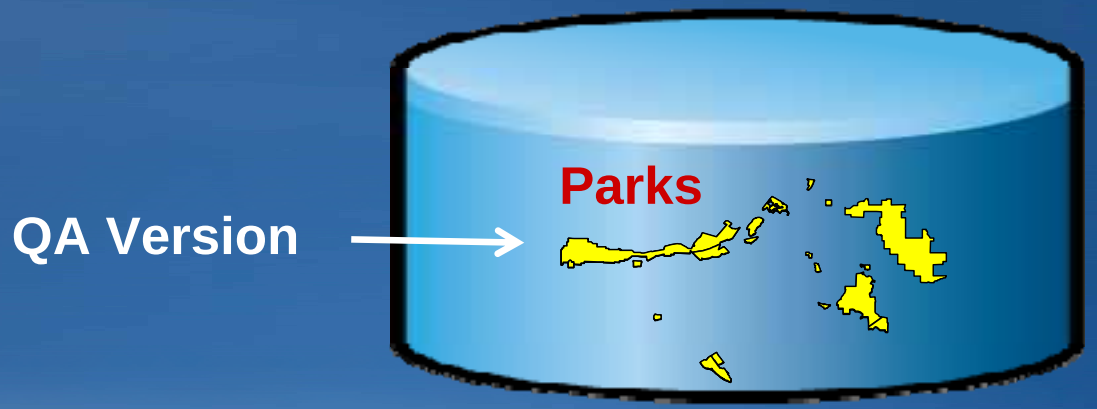


What is a Version?

- 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



Session Path

1. Introduction to ArcSDE Geodatabases
2. Versioning
3. Editing
   - Versioned Editing
   - Reconcile and Post
   - Non-Versioned Editing
   - Editing through SQL
4. Archiving
5. Distributed Geodatabase

Editing Databases

- Database editing relies on transactions
- Database transactions often conform to the ACID standards
- Transactions have “ACID” standards
  - Atomic — A transaction exhibits "all or nothing" behavior.
  - Consistent — A transaction leaves the database in a consistent state.
  - Isolation — Changes are isolated from other transactions is committed.
  - Durable — Once a transaction commits, its results are persistent.

Editing Geodatabases

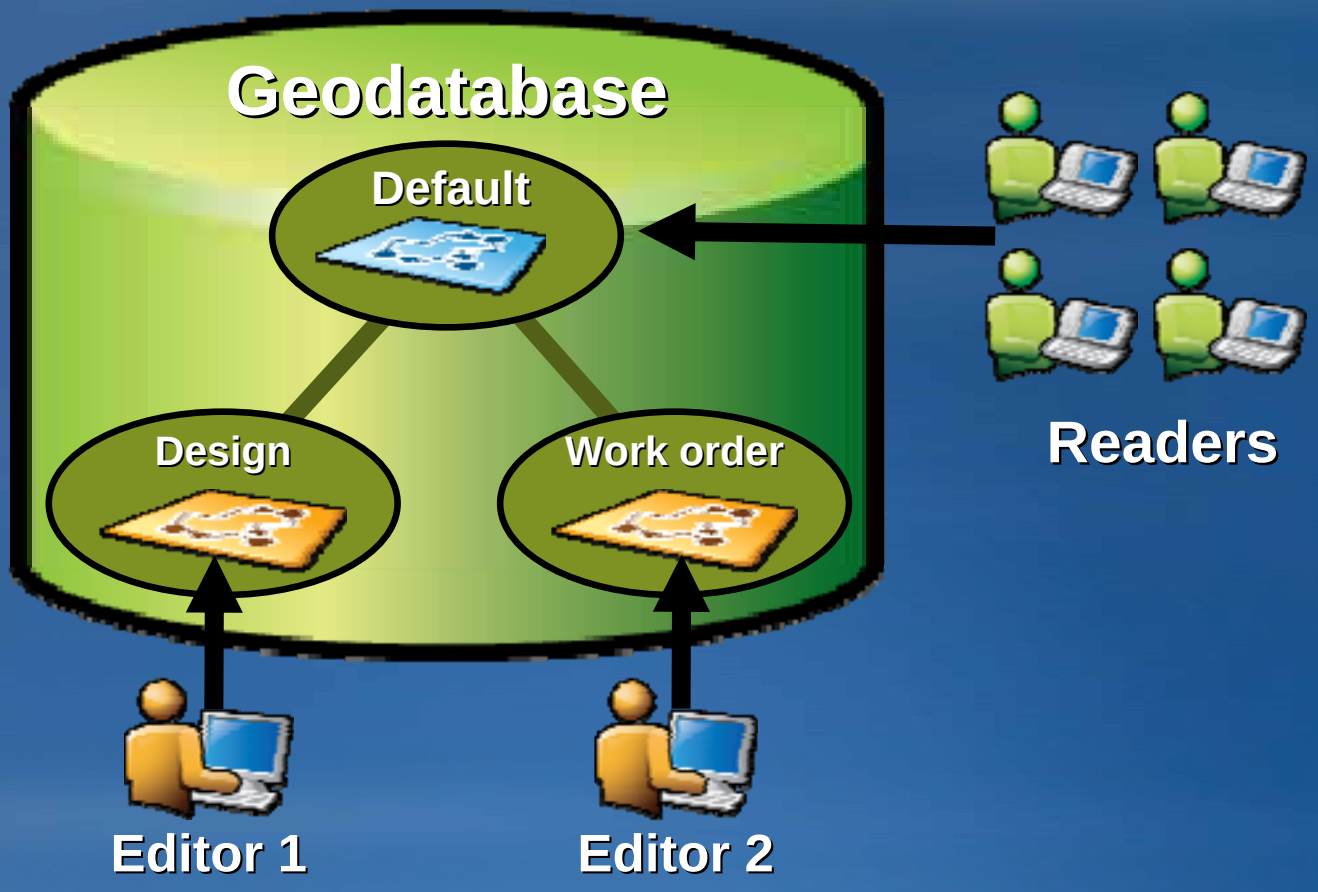
- **Short Transactions**
  - Small number of operations completed quickly
    - E.g. ATM transactions, Library records, Timecards
  - ACID requirements through DBMS Locking mechanisms
  - Concurrent transactions are isolated
- **Long Transactions**
  - Large number of operations over a long time period
    - E.g. Parcel updates, General geographic editing
  - Geodatabases extend the transaction model with Versioning
  - Multiuser editing without locking or data duplication
  - Editors work with unique isolated view of the geodatabase
- **GIS editors need both long and short transactions**

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

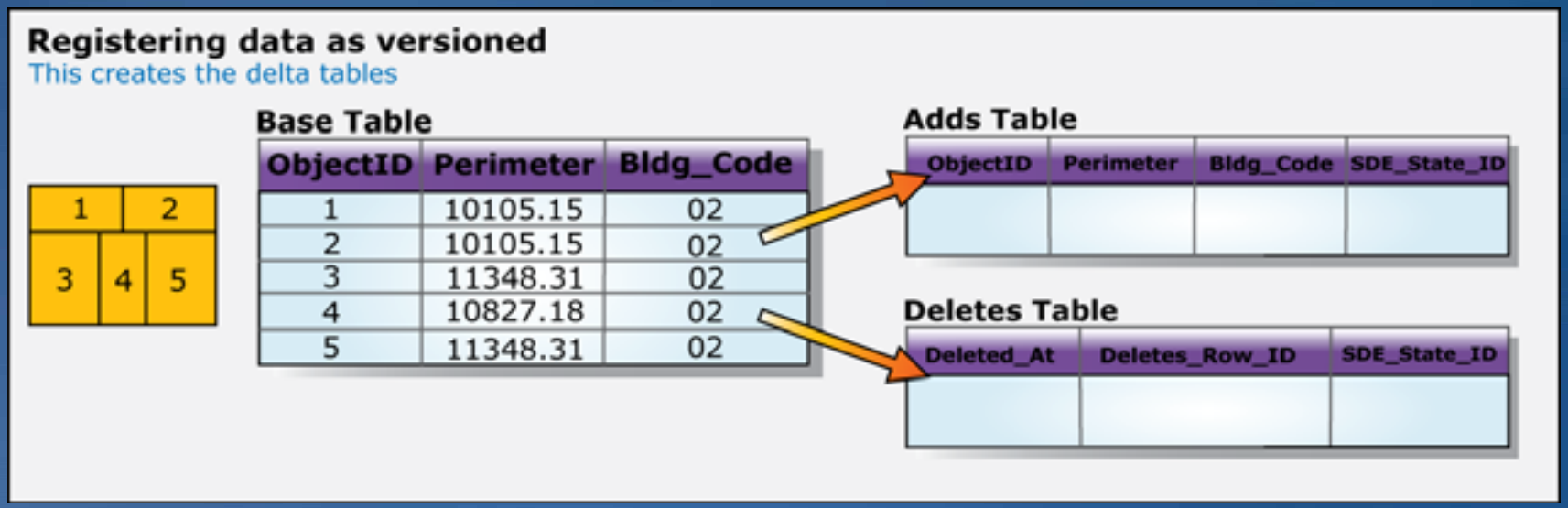
Versioned Editing

- 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.



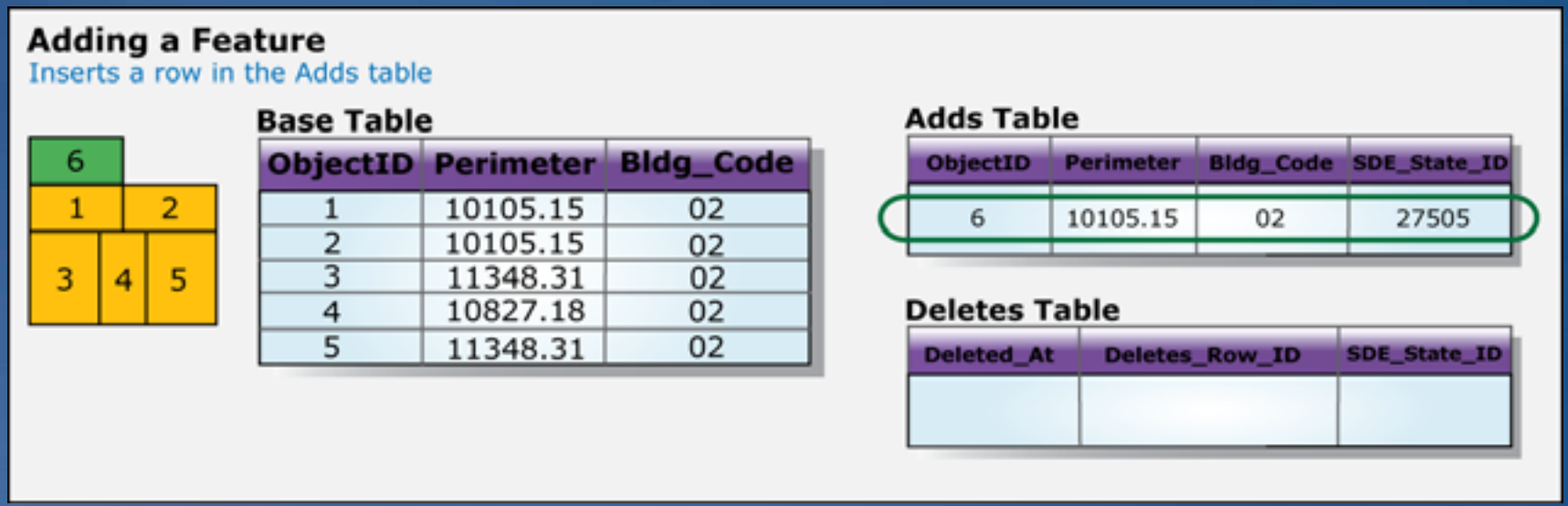
Versioned Editing – How It Works

- **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 (SDE_State_ID Field)

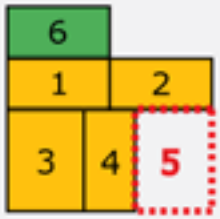


Versioned Editing – How It Works

- **Deleting Features**
  - Record added to Deletes Table
  - Version will be referenced (Deleted_At field)

Deleting a Feature

Inserts a row in the Deletes table



Base Table

ObjectID	Perimeter	Bldg_Code
1	10105.15	02
2	10105.15	02
3	11348.31	02
4	10827.18	02
5	11348.31	02

Adds Table

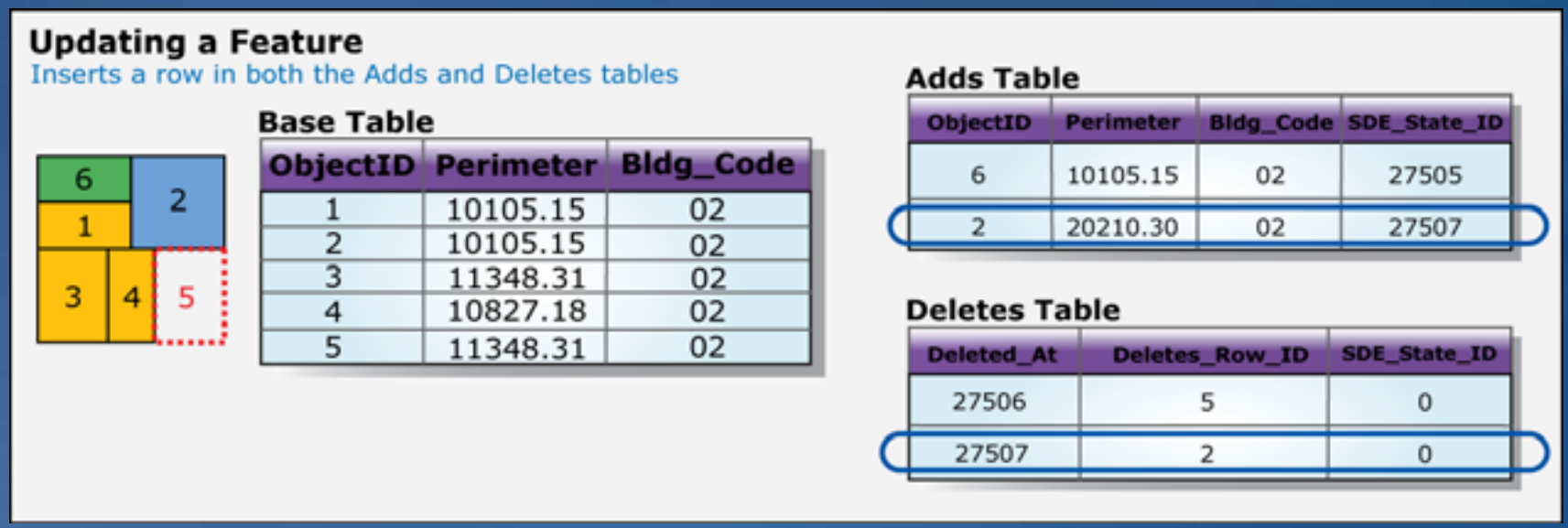
ObjectID	Perimeter	Bldg_Code	SDE_State_ID
6	10105.15	02	27505

Deletes Table

Deleted_At	Deletes_Row_ID	SDE_State_ID
27506	5	0

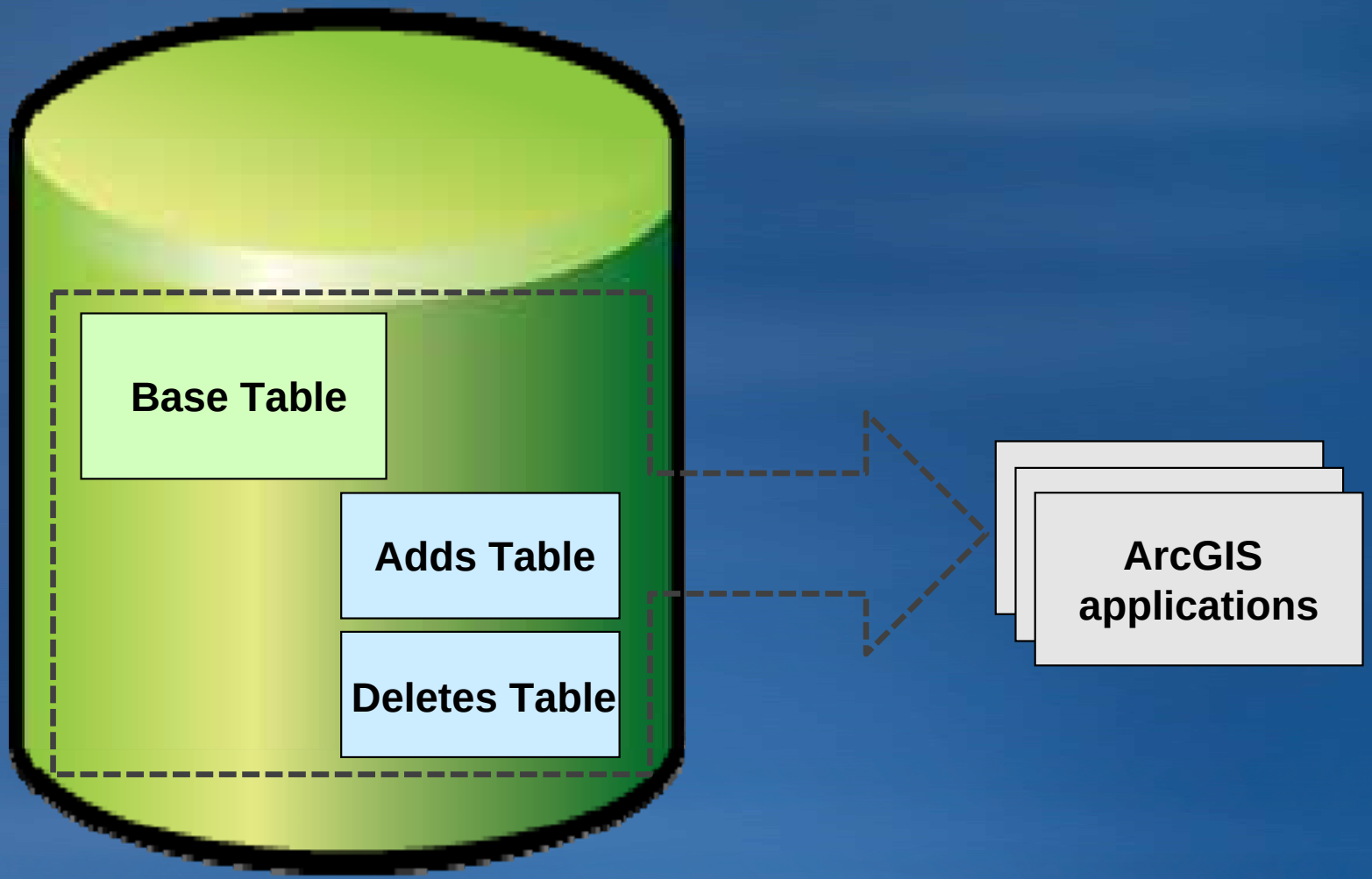
Versioned Editing – How It Works

- **Updating Features**
  - Record added to both Adds and Deletes table
  - Version will be referenced (SDE_State_ID Field)



Versioned Editing – How It Works

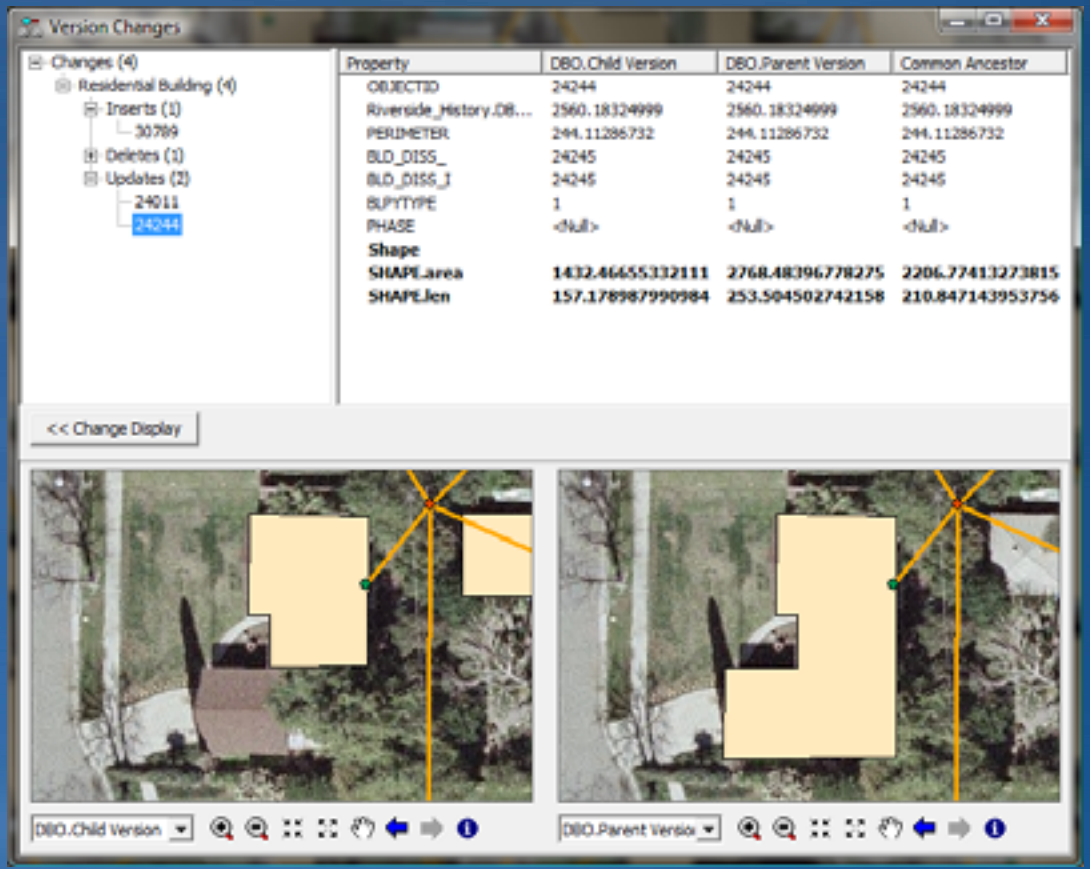
- Versioned representation of a feature class
  - combination of records in:
    - Base Table, Adds Table & Deletes Table



Version Changes Viewer

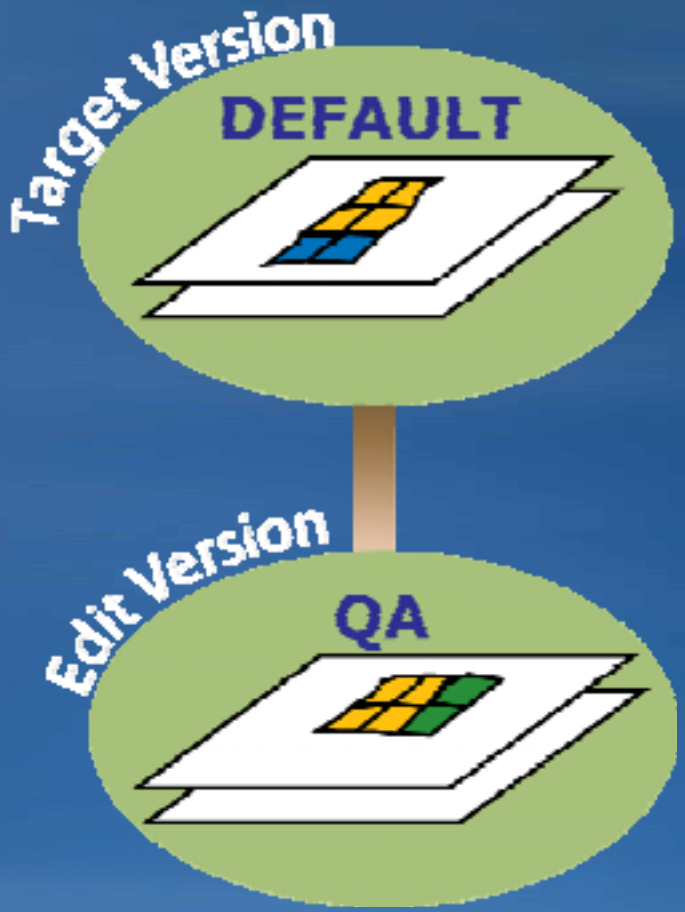
New at 9.3

- **Ability to see what has changed in a version**
  - View changes without having to do a reconcile
  - Displays all changes in version with respect to ancestor version



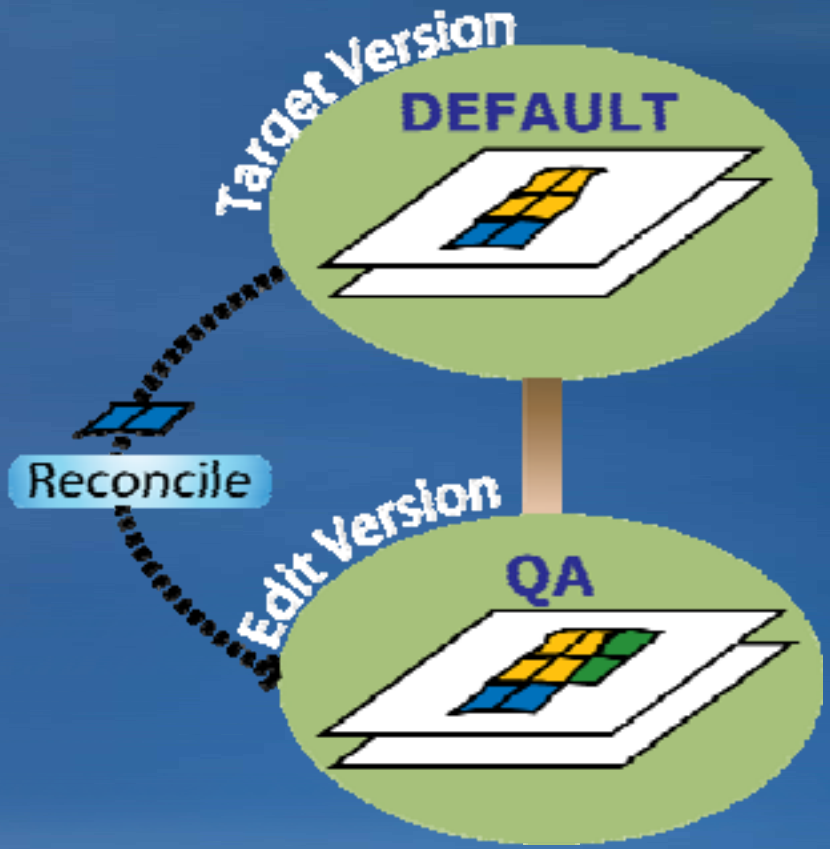
Versioned Editing – Reconcile and Post

- How can versions be merged?
- Through a process called reconcile and post



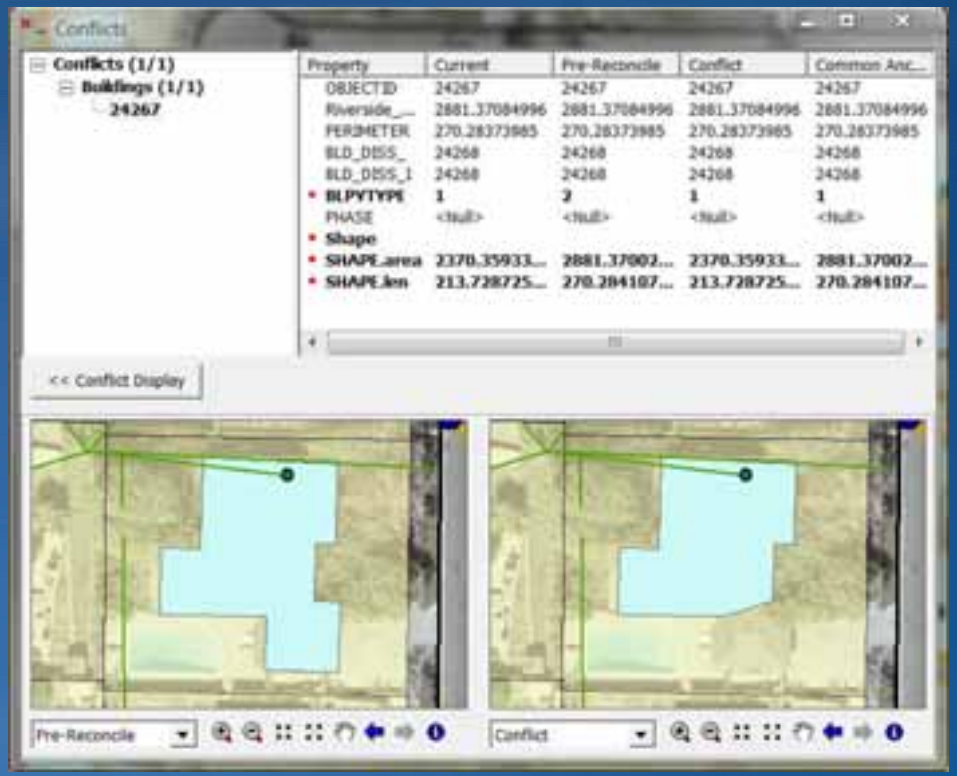
Versioned Editing – Reconcile

- Reconcile pulls any changes from the target version into the edit version
- Any conflicts will be detected



Reconcile and Conflicts

- Versioning does not lock data when it is edited
  - Because of this we must make sure data is not overwritten
  - We do this through conflict detection during a reconcile
- A feature will be in conflict any time it has changed on both versions
- Conflict Resolution Dialog



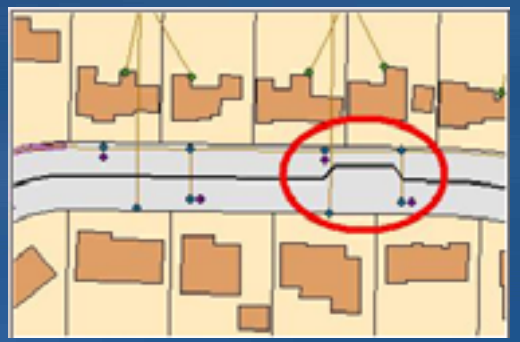
Reconcile and Conflicts

New at 9.3

- Merge Geometry option for conflict management
  - Improves the conflict management experience for:
    - large polygon and polyline features

Edit Version

Target Version

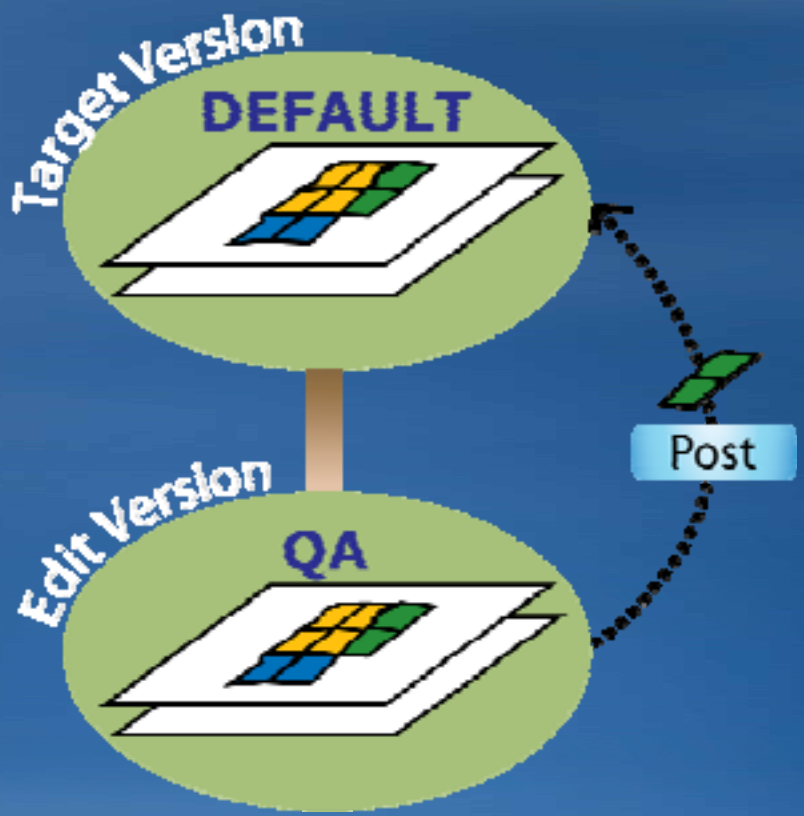


Merge

Edit Version

Versioned Editing – Post

- Posting versions merges any changes in the edit version into the target version
  - After a post versions are identical

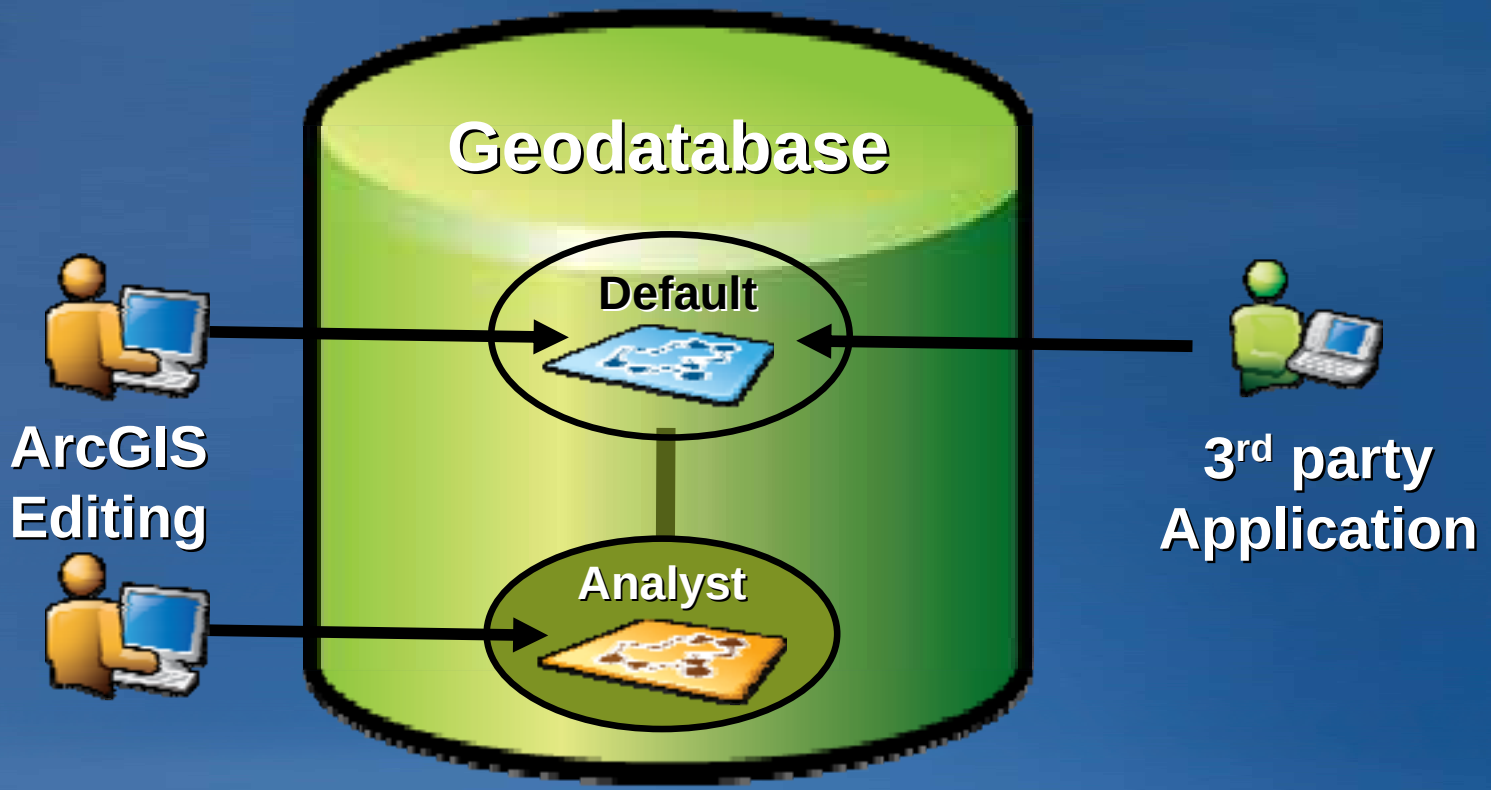


Versioned Editing - Move to Base Option

- What is it?
  - Versioned editing with the ability to move changes made in the Default version into the base tables
  - Changes made in non-Default versions are still stored in the delta tables
- Designed for IT integration
  - Edits visible to 3rd part applications as soon as they are saved
- Simple data only
  - Points, lines, polygons, annotation, relationship classes
  - No Topology, Geometric Networks…etc

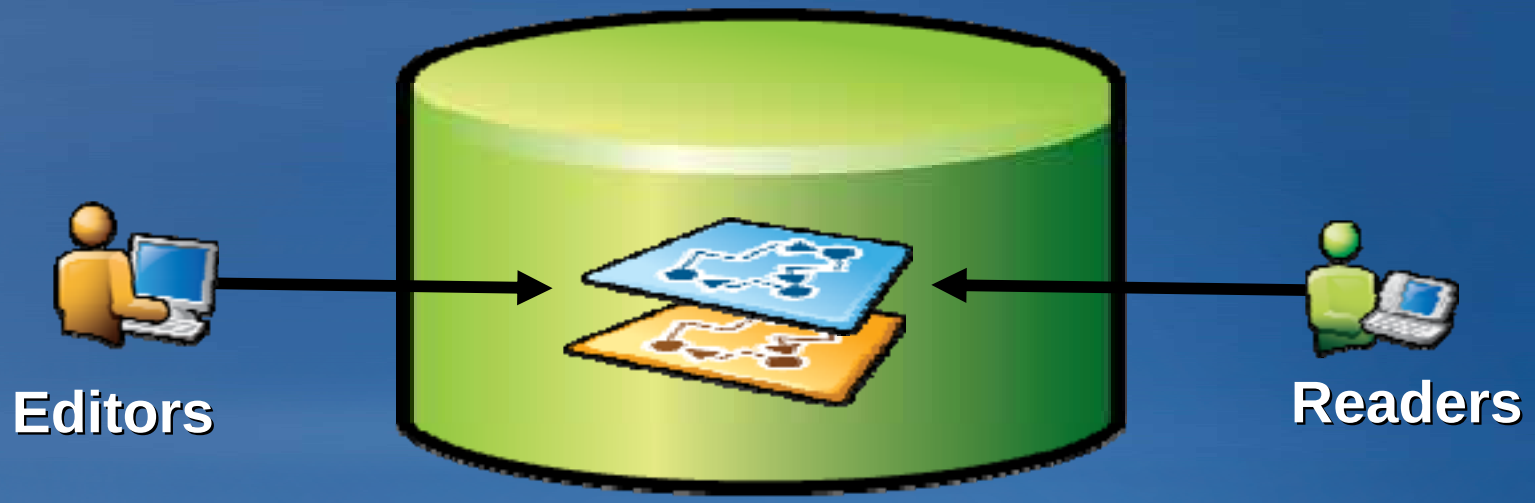
Versioned Editing - Move to Base

- Why would I use the move to base option?
  - Want version editing experience but…
    - Need to integrate with 3rd party applications
    - Use of database constraints when editing DEFAULT version



Non-Versioned Editing

- Directly editing the database tables
  - Not editing in a version
    - Uses a database transaction (short transaction)
    - Edits immediately available upon save
  - Designed for IT integration
    - Suggested for Non-ESRI client interaction
    - Database integrity rules
  - Simple data only
    - Points, lines, polygons, annotation, relationship classes
    - No Topology, Geometric Networks…etc



SQL Editing

- SQL can be used to update data directly

- Geometry editing possible through spatial types
- What is a spatial type?
  - A database data type that stores spatial data
- Why are they useful?
  - ESRI Client not necessary to edit data
  - SQL access to geometries
- Databases with spatial types
  - Oracle, SQL Server, Informix, DB2, PostgreSQL

Geodatabase Editing Summary

- Three ways to edit data
  - Versioned Editing (Long Transactions)
  - Non-Versioned Editing (Short Transactions)
  - Editing through SQL (Short Transactions)

- Which one do I use?
  - Depends on behavior desired
    - Short vs Long Transactions
    - Is data being accessed by non-ESRI applications?
    - Are many editors editing the same data?

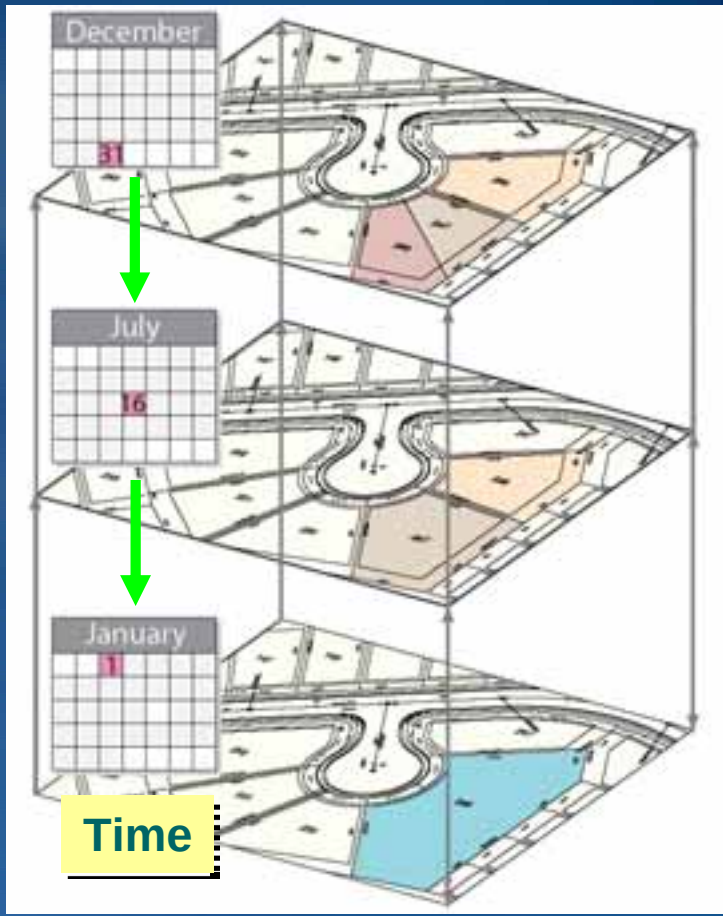
Version Editing Demo

Session Path

1. Introduction to ArcSDE Geodatabases
2. Versioning
3. Editing
4. Archiving
   - What is it?
   - How is it used?
5. Distributed Geodatabase

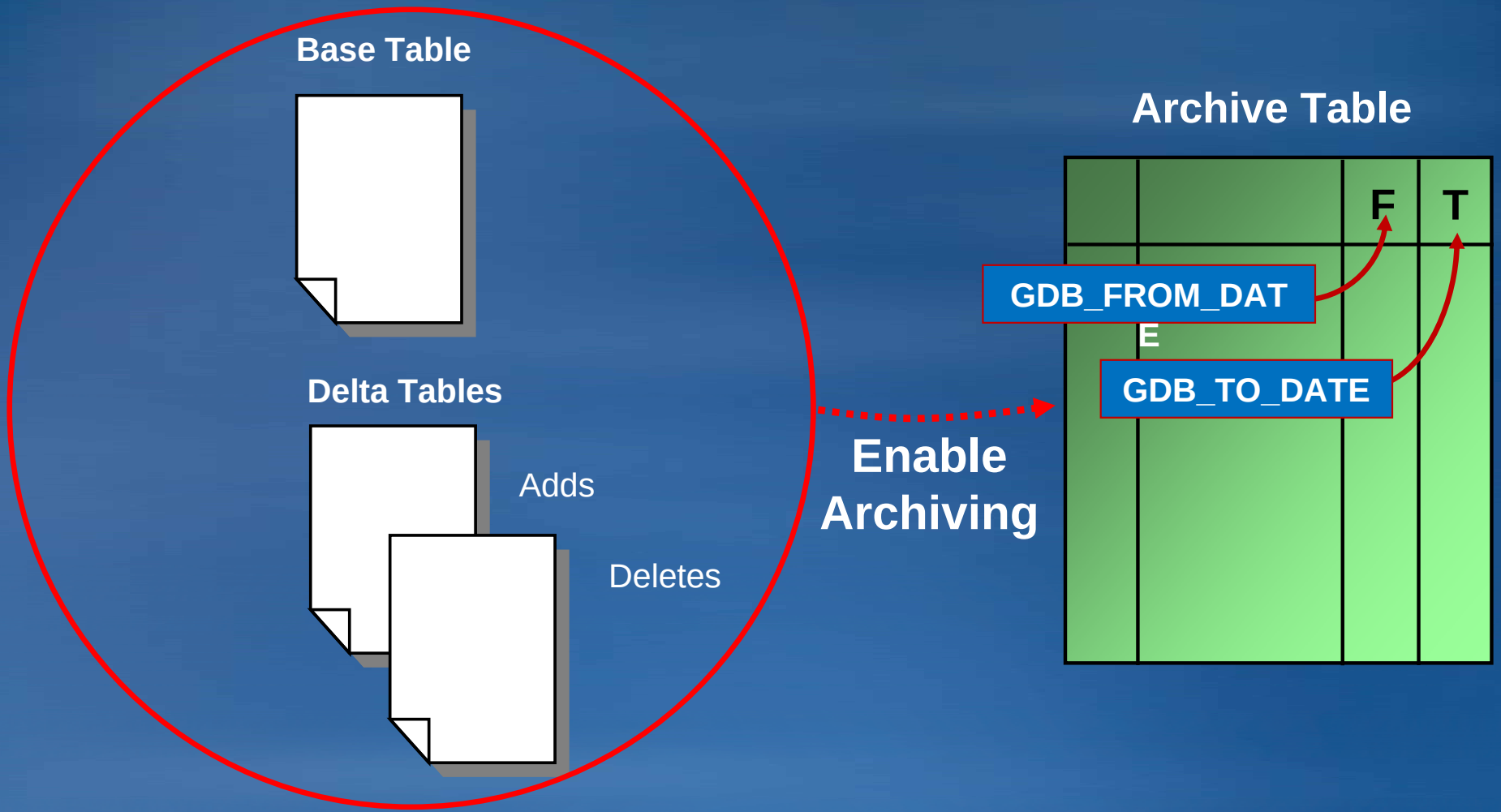
Geodatabase Archiving: What is it?

- Historical archiving of all edits made to the Default version
  - Maintain a record of a feature classes representation over time
- Ability to query historical representations of a feature
  - Archives can be queried based on date information
- Extends versioning
  - Classes must be versioned before they can be archive enabled



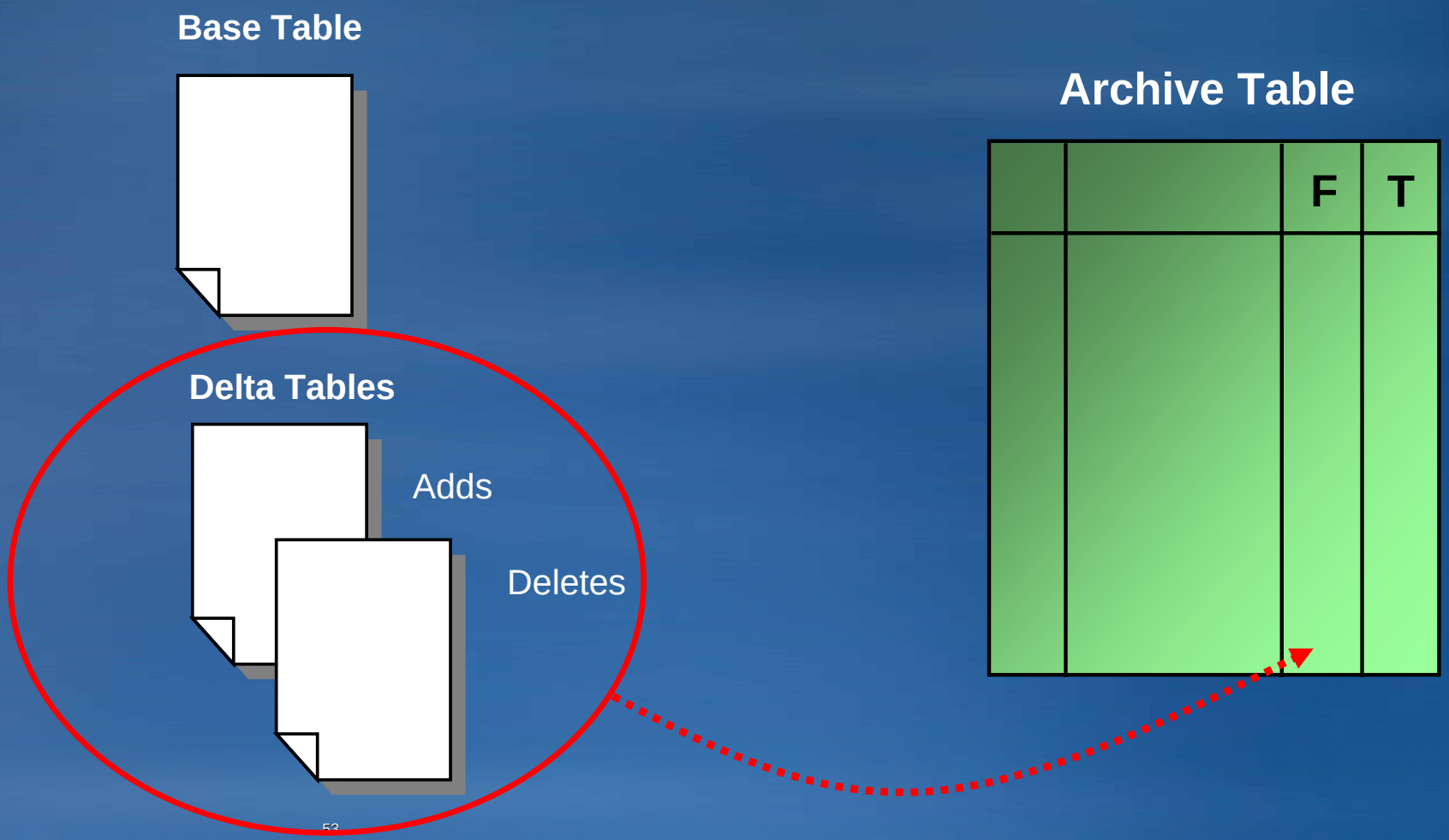
Geodatabase Archiving: How it works

- Class must be enabled for archiving
  - This creates an archive table in the geodatabase
  - Size of archive table depends on size of class being archived



Geodatabase Archiving: How it works

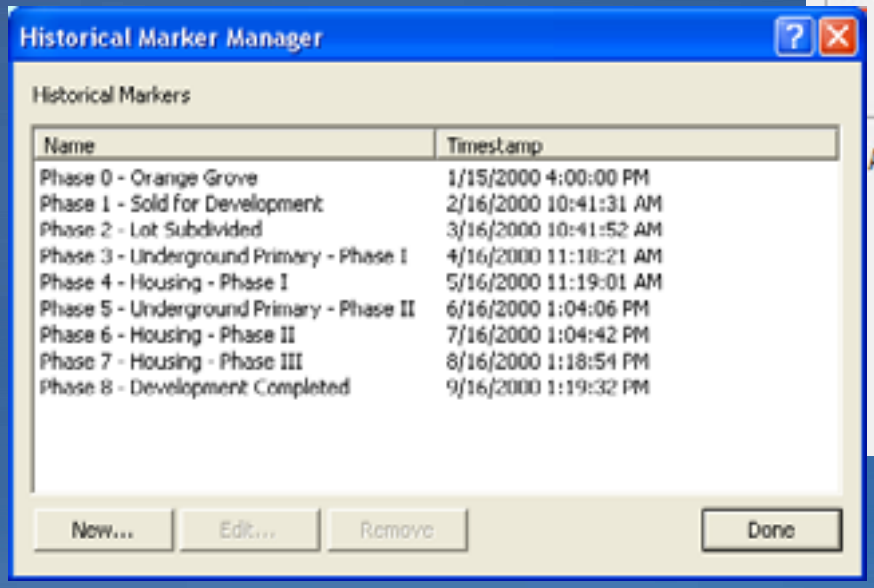
- When edits are made on the Default version
  - These changes are added to the archive table

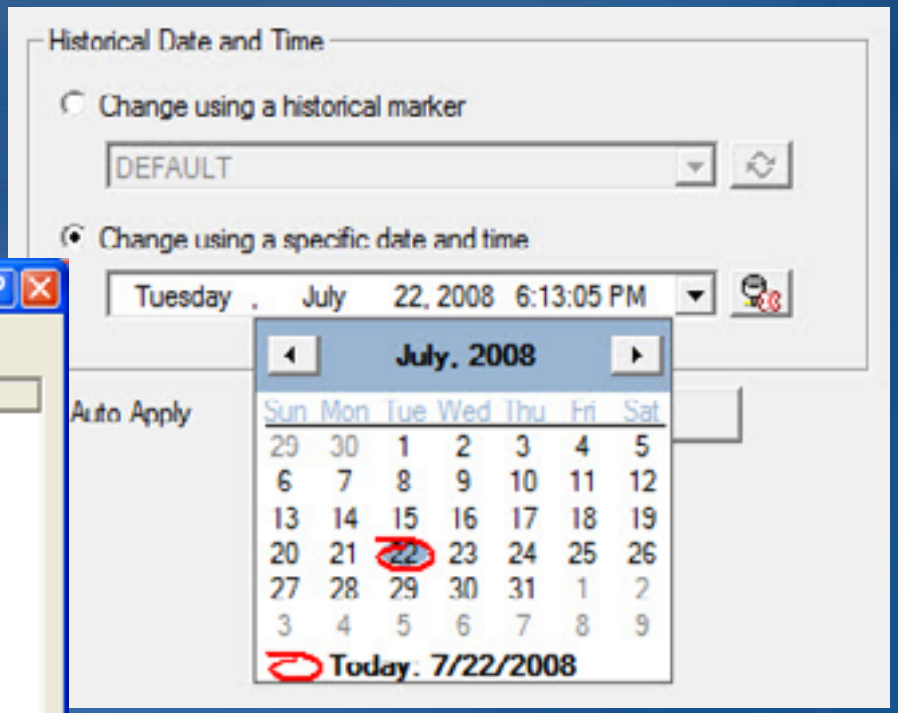


A

Geodatabase Archiving: How it works

- Archive table is used to satisfy historical queries
- Can navigate through history in two ways
  - Through specific date query
  - Through historical marker





A

Archiving Demo

Archiving in a nutshell

Session Path

1. Introduction to ArcSDE Geodatabases
2. Versioning
3. Editing
4. Archiving
5. Distributed Geodatabase
   - Data Distribution and Geodatabase Replication

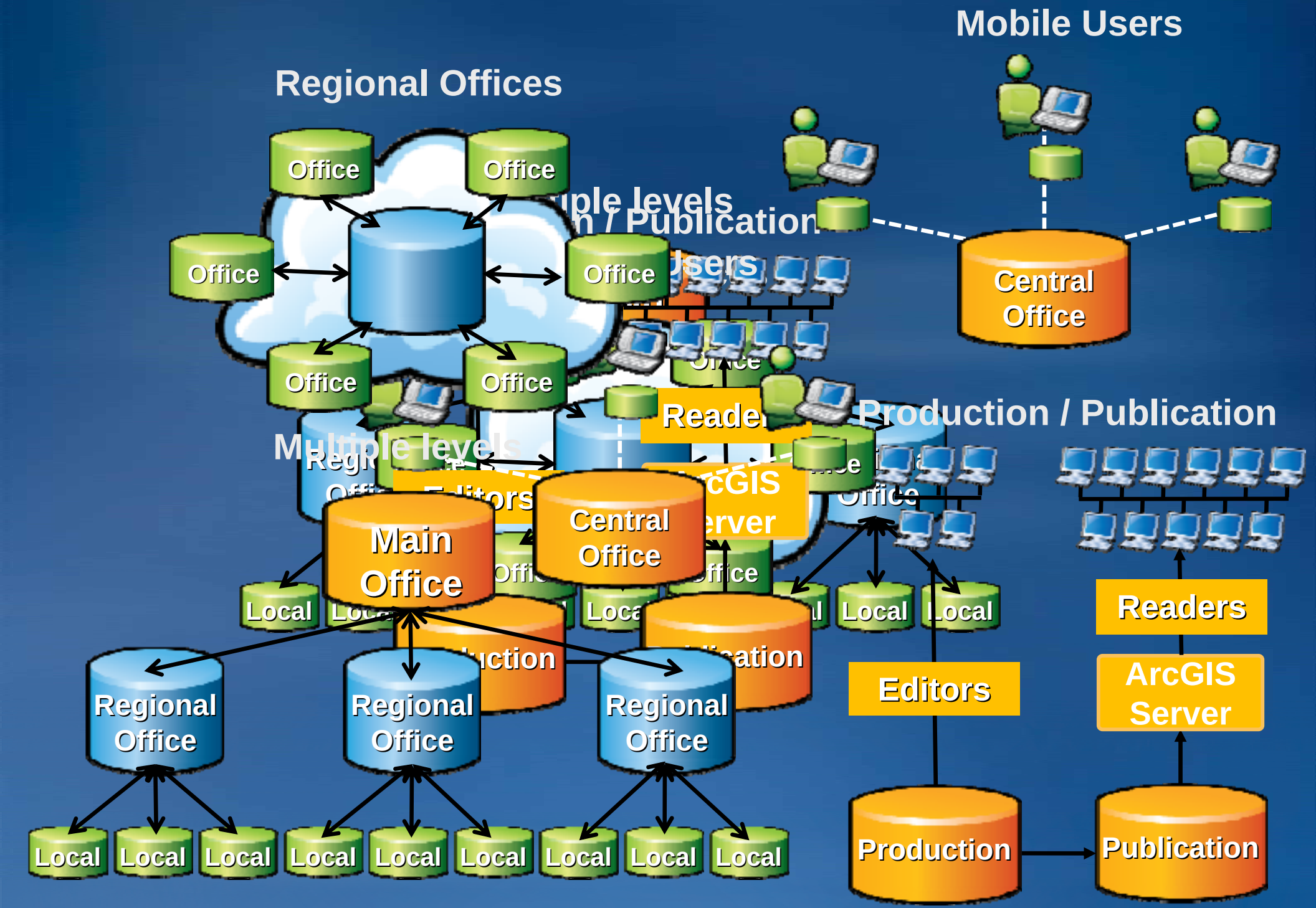
Geodatabase Replication

- 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)

Replication – Use Cases

- Mobile Users and Field Crews who need to be disconnected from the network.
- Users who need to maintain copies of data at different organizational levels (city, county, state)
- Users who want to maintain copies of data at different geographic facilities.
- Users who need to distribute work to contractors.
- Production and publication geodatabases

Distributed Geodatabase Use Cases

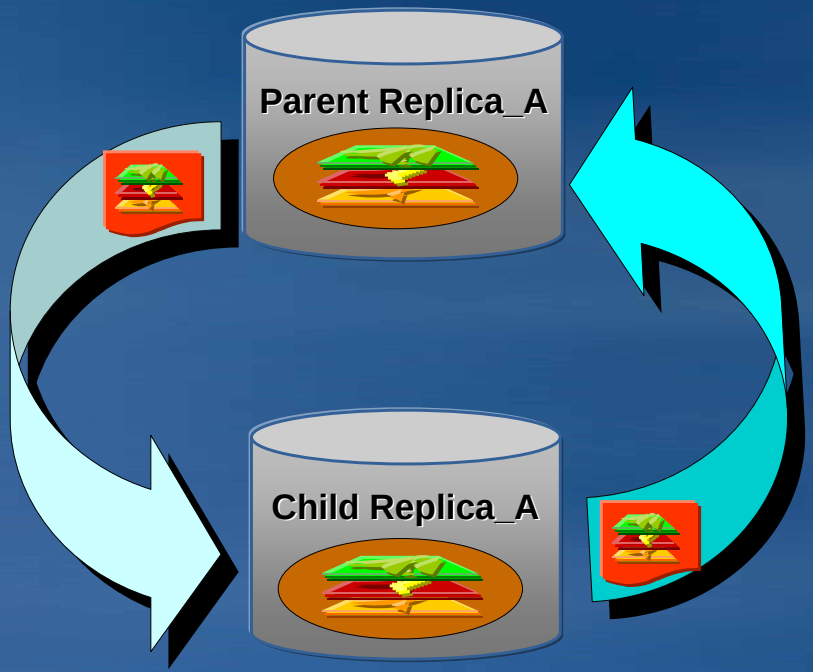


Data distribution in Enterprise systems

- Geodata services can be used in conjunction with other data distribution techniques
  - Use geodatabase replication to synchronize changes between offices
  - Use Mobile services for field workers with lightweight mobile devices
  - Use geodata services for field workers who need ArcGIS Desktop or ArcGIS Engine in the field

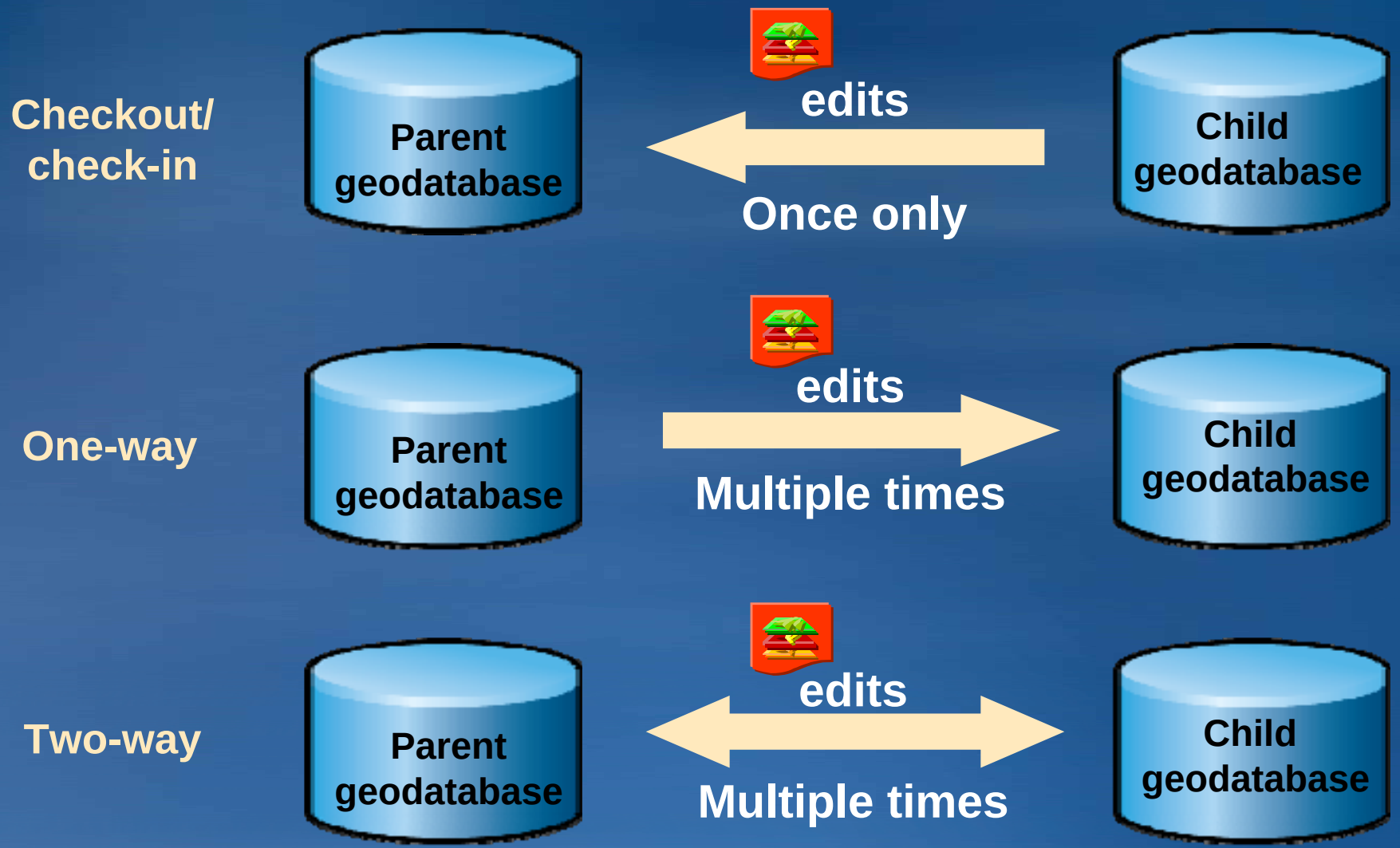
Geodatabase Replication - Concepts

- 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

Geodata Services: replica types



Replication - Concepts

- Works in a connected and a disconnected environment
- Replicas can be Synchronized in either both directions or just a single direction
- Synchronization is based on exchanging messages and is fault tolerant
- You can Create and Synchronize Replicas using Wizards and GP Tools
- Supports a applying schema changes across replicas
  - Subset of schema changes are supported
- Developers have a high level object model and API

Geodatabase Replication - LAN and WAN

- LAN - Use connections to your local geodatabases
- WAN - Use ArcGIS Server and geodata web services to access remote geodatabases
- All geodatabase replication workflows are supported in both environments

9.3 Geodatabase Replication

- Enhanced one way replication to support replicating to file geodatabases and personal geodatabases
- Added logging to improve trouble shooting
- User defined global ID's (API)
- Make it easier to generate updategrams (API)

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 ArcSDE for Microsoft SQL Server Express 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 on a regular basis
  - You can use the create replica and create replica from server geoprocessing tools to build models
- 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 and the windows scheduler
  - Consider synchronization order

Geodatabase Replication – Getting Started

- **Integrate synchronization with version management strategy**
  - Synchronize before running reconcile and compress services
  - Reconcile service should cover replicas with a manual reconcile policy
- **Network speed**
  - Use geodatabases directly over fast networks (LAN)
  - Use ArcGIS server and geodata services on slower networks (WAN – i.e. DSL)
  - Use disconnected synchronization techniques over very slow networks (slow dial-up modem) or where there is no network connectivity

DBMS Replication with Geodatabases

- Geodatabase replication does not use DBMS replication
- DBMS Replication - Requirements and limitations
  - 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
- DBMS Replication - Advantages
  - Can work with non-versioned data
  - Can replicate entire database
  - Can be configured to provide synchronous replication
  - Ideal for systems requiring high availability

Demo

Geodatabase Replication

What's coming: Geodatabase Replication

- One way replication using archiving
- One way child to parent replica
- Schema mapping across replicas
- Simple check-out and two-way replicas
- Extending XML Update-gram format to store the original/old row
- Create a replica to a named version on the child

Geodatabase Summary

- **ArcSDE Geodatabases allow you to:**
  - **Manage geographic information**
  - **Work with rich data models that go beyond simple features, rasters, and attributes**
  - **Openly manage transactions, archives, and replicas across organizations**
  - **Openly edit in any application using simple features interchange**

Learn More

http://www.esri.com/training

Instructor-Led Training

- Building Geodatabases
- Introduction to the Multiuser Geodatabase
- Data Management in the Multiuser Geodatabase
- Managing Editing Workflows in a Multiuser Geodatabase

ESRI Training…keep critical skills up to date