Good Housekeeping Tips for Geodatabases and ArcGIS for server

Kathryn Browning

Introductions

Kathryn Browning: Senior Consultant/Project Manager and Esri Enterprise Advantage Program (EEAP) Technical Advisor

Information For this Presentation

Based upon Esri Help Documents and Esri User Conference presentations by Chet Dobbins, Ron Lazar, Shannon Shields, Cheryl Cleghorn, Shawn Thorne, Jim Gough, Amit Kesarwani, and Said Parirokh.

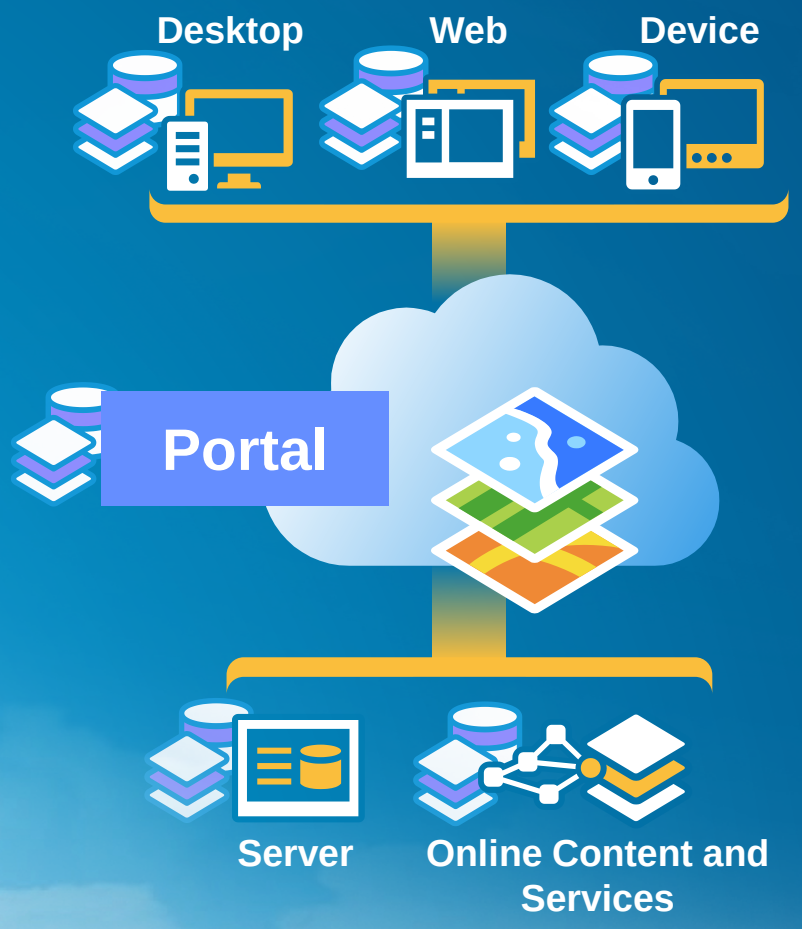
Housekeeping isn’t just for home

Your enterprise geodatabase and ArcGIS for Server need to be “tidied” up too in order to run optimally.

This session will go through tips to keep your enterprise geodatabase and ArcGIS for Server running well.

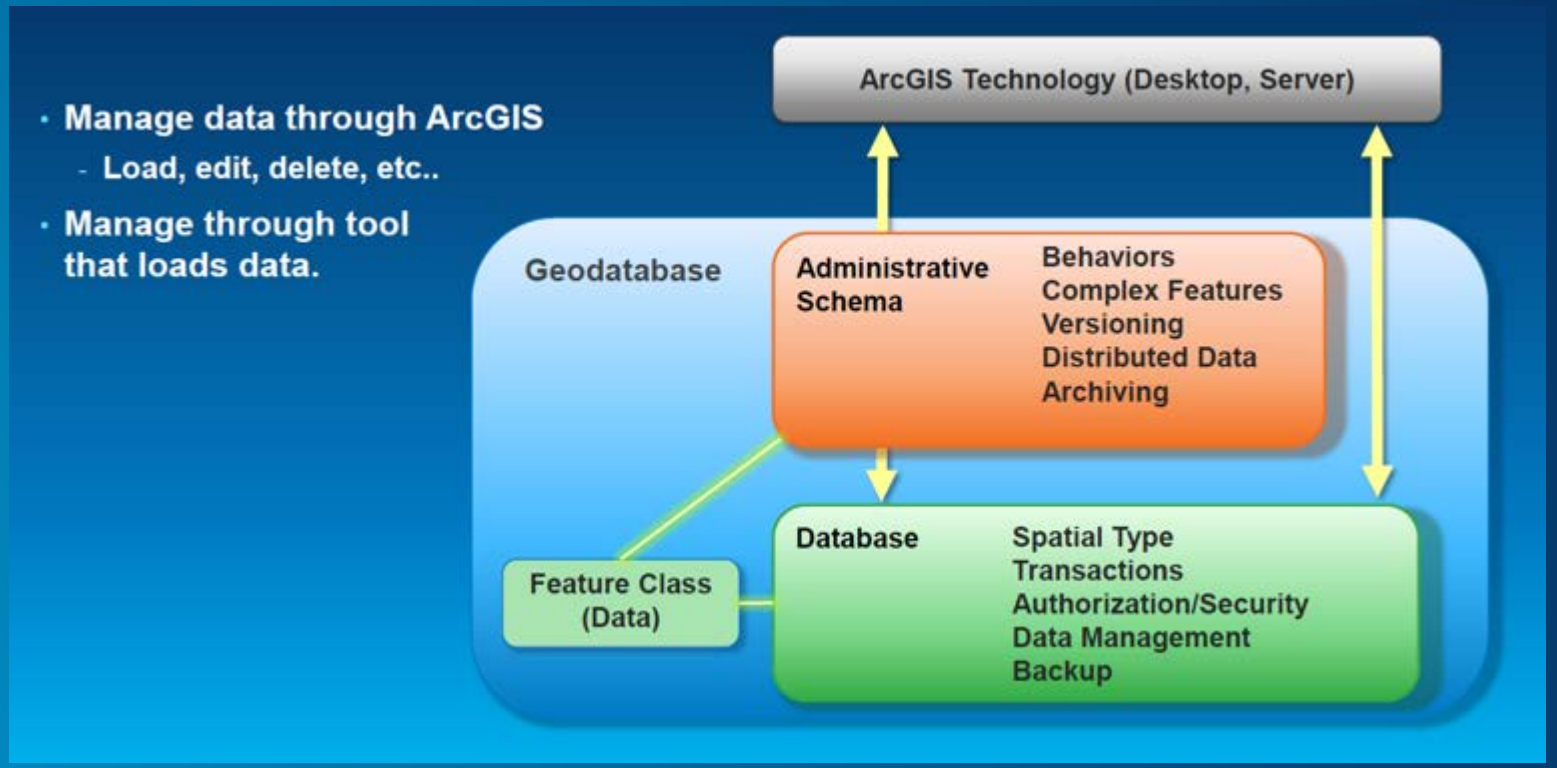
The Geodatabase is a part of the ArcGIS Platform

Geodatabase Overview



What is a Geodatabase?

- A collection of geographic datasets of various types
  - A comprehensive model for representing and managing GIS data
- A geodatabase is an ArcGIS construct hosted in a database
  - Physical data store
  - Core ArcGIS data model
  - Transaction model
  - COM components



Storing geographic data

Many different methods of storing geographic data	Databases
	Shapefiles
	Geodatabases
ArcGIS can use spatial data stored in both Geodatabases and databases.	Database spatial data can be used natively in ArcGIS
	Geodatabases are still recommended for storing your spatial data

Why use the Geodatabase?

Geodatabase	Database
Stores Data	Stores Data
Multiuser Access	Multiuser Access
Backup/Recovery	Backup/Recovery
Scalability	Scalability
Spatial Data	Spatial Data
Replication	Replication
Behavior	
Topologies	
Geometric Networks	
Network Datasets	
Relationship Classes	
Domains	
Subtypes	
Versioning	
Distributed data	

Geodatabase Types

Desktop format geodatabases

	Personal	File
Functionality	Original Desktop Format	Improved Desktop Format
Storage mechanism	Microsoft Access database (.mdb)	File Folder; displays .gdb extension in ArcCatalog
Storage Limit	2 GB per Geodatabase; effective limit ~500 MB	1 TB per object, configurable to 256 TB
User Limit	1 editor per database	1 editor per object
Platform	Windows	Any
Licensing	ArcGIS for Desktop	ArcGIS for Desktop
Distributed gdb functionality	Check out/check in One way replication	Check out/check in One way replication

Geodatabase Types

	Desktop	Workgroup	Enterprise (Multi-user)
Functionality	Distributed data or project use	Departmental projects or small organizations	Large capacity and user base
Storage mechanism	Microsoft SQL Server Express	Microsoft SQL Server Express	SQL Server, Oracle, PostgreSQL, DB2, Informix
Storage Limit	10GB per database server	10GB per database server	Unlimited
User Limit	3 users, 1 can edit	10 concurrent users, all can edit	Unlimited
Platform	Windows	Windows	Any
Licensing	ArcGIS for Desktop Standard or Advance	ArcGIS for Server Workgroup Edition	ArcGIS for Server Enterprise Edition
Distributed gdb functionality	Replication – all types & versioning	Replication – all types & versioning	Replication – all types & versioning

What is a Multi-user Geodatabase?

DBMS enabled with ESRI's geospatial technology

Unique capabilities:

- Many supported DBMSs
- Full, open SQL access
- Versioning
- Archiving
- Replication

Good Housekeeping Tip

For Geodatabases

Pick the right database for what you want to accomplish

Database and ArcGIS Versions

- Run the database on supported platforms
- Check out the database requirements for the ArcGIS version that you want to install
- For ArcGIS 10.4.1 geodatabase on MS SQL Server:
  - 2016 (64 bit) is supported
  - 2012 and 2014 (32 bit and 64 bit)
    - 32 bit is only for desktop geodatabases
- For ArcGIS 10.4.1 geodatabase on Oracle:
  - 11gR2 (64 bit) – 11.2.0.4
  - 12c R1 (64 bit) – 12.1.0.2
- For ArcGIS 10.4.1 geodatabase on PostgresSQL and PostGIS:
  - PostgresSQL 9.4.5 (64 bit)
    - PostGIS 2.2
  - PostgresSQL 9.3.10 (64 bit)
    - PostGIS 2.2
  - PostgresSQL 9.2.14 (64 bit)
    - PostGIS 2.1

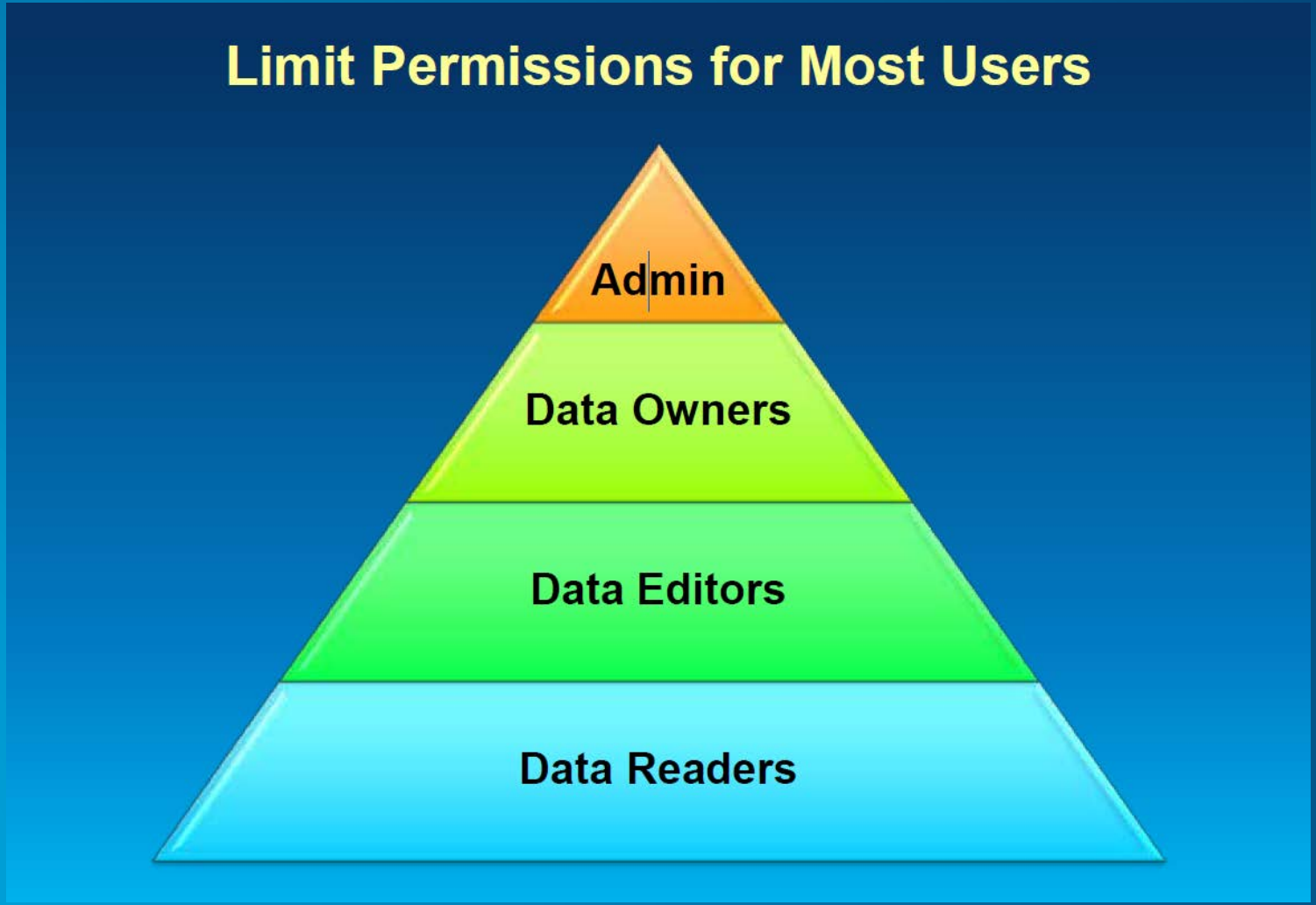
Good Housekeeping Tip

For Geodatabases

Pick the right version of the database for the ArcGIS version that you are implementing or upgrading to

User Permissions and Data Owners

Don’t load everything to Sde or DBO



User Permissions and Data Owners

Each User has certain tasks

Role	Tasks	Colleagues
End User • Read only • Read/Write	• Edits data • Performs analysis • Creates maps	
Data owner	• Chooses editing workflow • Performs data maintenance • Establishes schema • Grants privileges to datasets	
Administrative user • SDE (all DBMSs) • DBO (SQL Server only)	• Manages repository • Handles configuration • Runs compress operation • Owns DEFAULT version • Manages version environment	

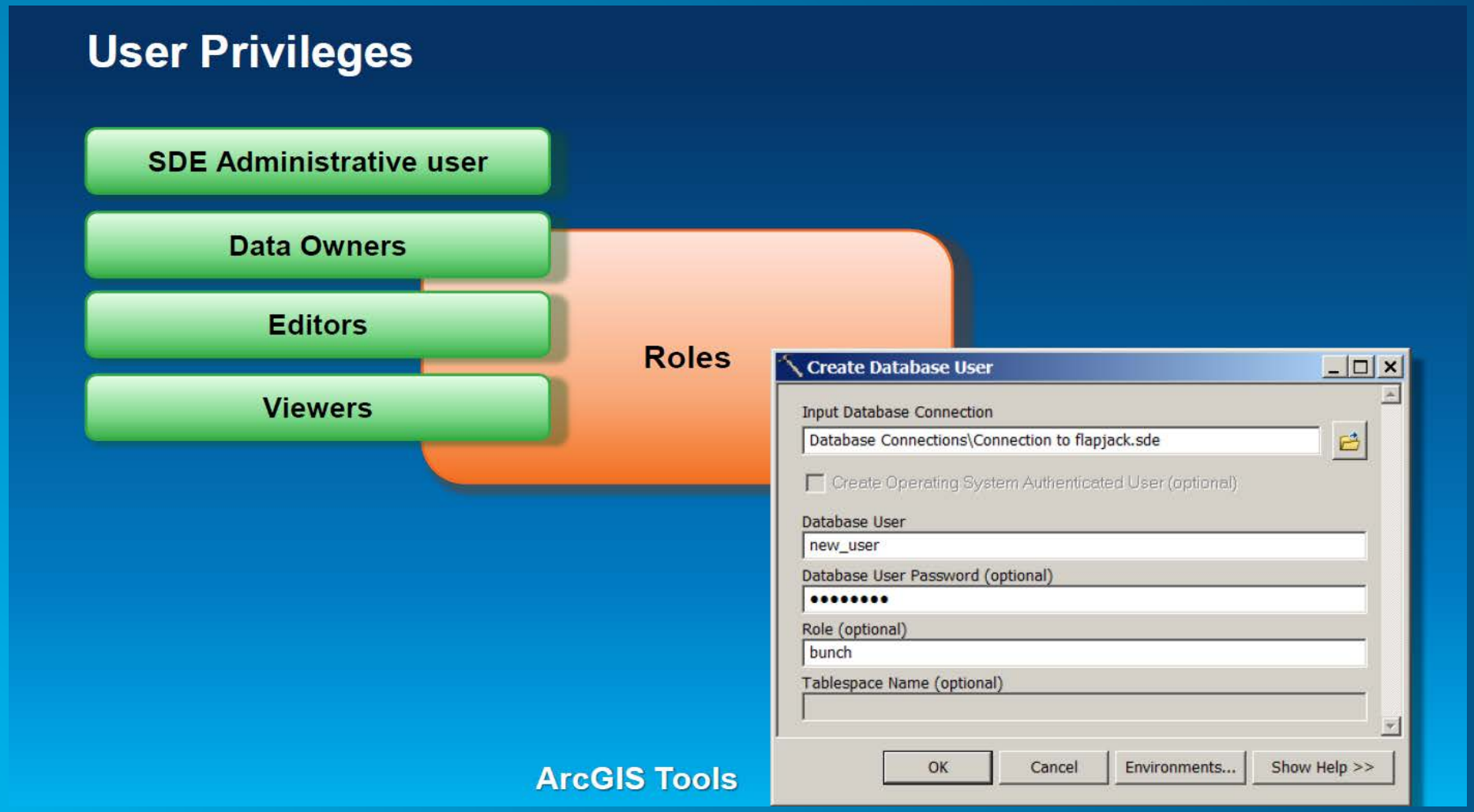
Good Housekeeping Tip

For Geodatabases

Set up the geodatabase with different users for different roles: SDE Admin user, Data Owners, Editors, and Viewers

User Privileges

Simplify maintenance by using roles



Good Housekeeping Tip

For Geodatabases

Grant privileges to roles instead of users

Versioning

Key Concepts - Version

Version	In databases, an alternative state of the database that has an owner, a description, a permission (private, protected, or public), and a parent version. Versions are not affected by changes occurring in other versions of the database.

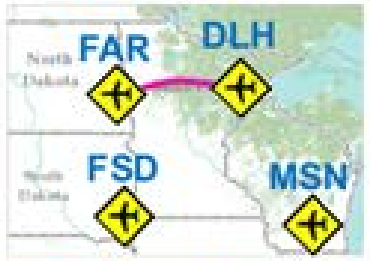
- **A “Snapshot” in time of the geodatabase**
- **Not separate copies of the geodatabase**
- **Isolates a user's work across multiple edit sessions**
- **Think about tracking edits in word and sharepoint**

What is a Version?

A version is a construct in an ArcGIS enterprise geodatabase that provides users with the ability to isolate their edits from other users for an unspecified period of time.

- Each version is isolated until you merge edits
- A version is just a record in a system table that stores the version’s properties and tracks edits that are stored in separate tables.

Representing versions



ANDREW

Airports

ObjectID	Shape	Airport	...
1	●	DLH	...
2	●	FSD	...
3	●	MSN	...

A7

ObjectID	Shape	Airport	...	State_ID
4	●	FAR	...	3

D7

SDE_State_ID	Deletes_Row_ID	Deleted_AT

Air_Routes

ObjectID	Shape	...

A8

Robert LeClair

ObjectID	Shape	...	State_ID
1		...	3
2		...	4
3		...	5
4		...	6

D8

SDE_State_ID	Deletes_Row_ID	Deleted_AT

SDE_VERSIONS

Name	State_ID	...
DEFAULT	0	...
ANDREW	3	...
OLIVIA	6	...

SDE_TABLE_REGISTRY

Registration_ID	Table_Name
...	...
7	Airports
8	Air_Routes

SDE_MVTABLES_MODIFIED

State_ID	Registration_ID
3	7
3	8
6	8

SDE_STATES_LINEAGES

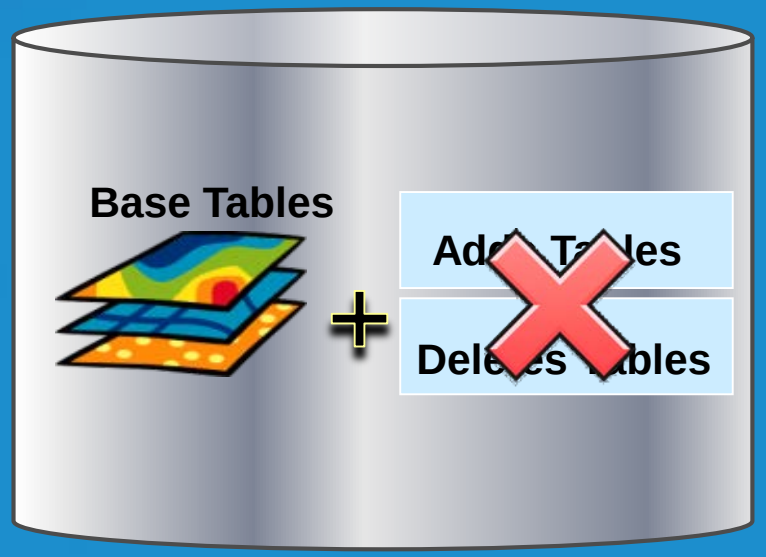
Lineage_name	Lineage_ID
0	0
1	0
4	0
1	3
4	6

SDE_STATES

State_ID	Owner	...	Parent_State_ID	Lineage_name
0	SDE	...	0	0
3	ANDREW	...	0	1
6	OLIVIA	...	0	4

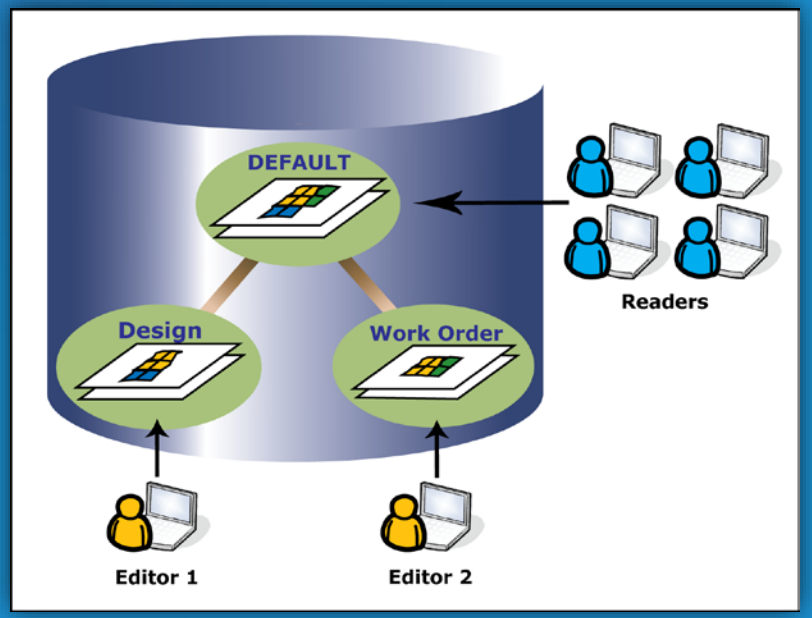
Non-Versioned Editing

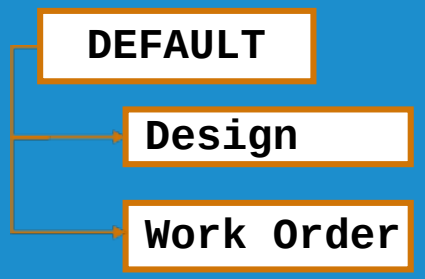
- Directly editing the base tables
- Benefits IT integration
- Database integrity rules
- Simple data only (Points, Lines, Polygons)



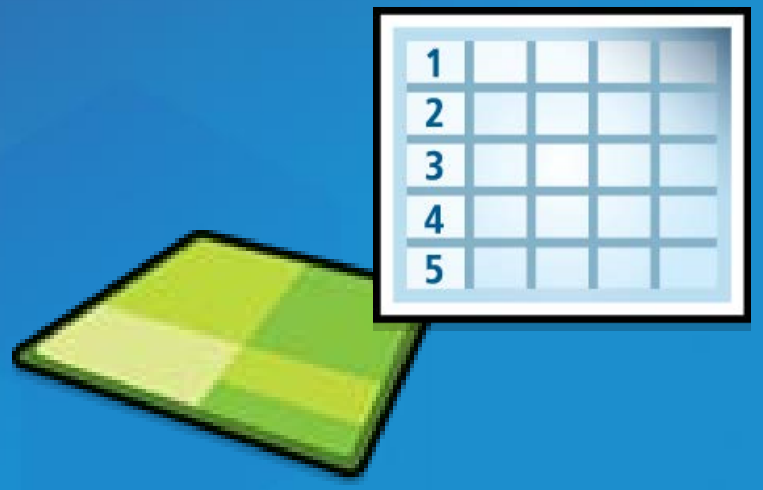
Versioned Editing

- Versioned Edit Sessions
  - Through a version
  - Concurrent editing
  - Long transactions (hours/days)
  - Undo/Redo





How Versioning Works

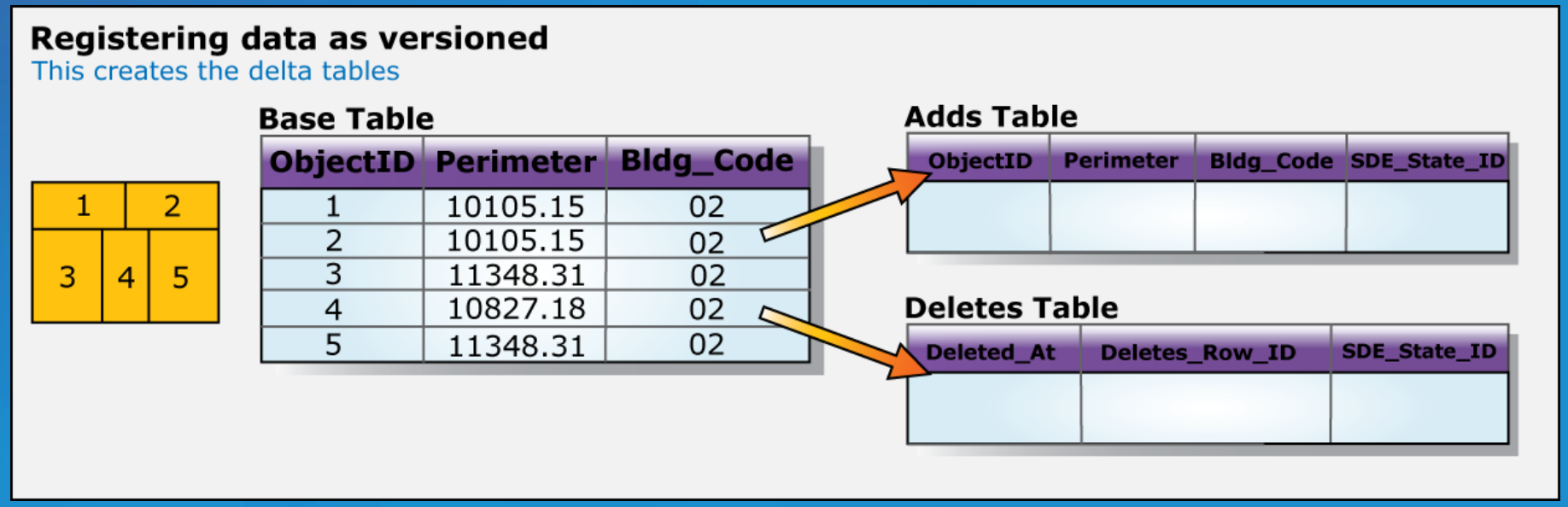


Business or
Base table

- ....Register as Versioned

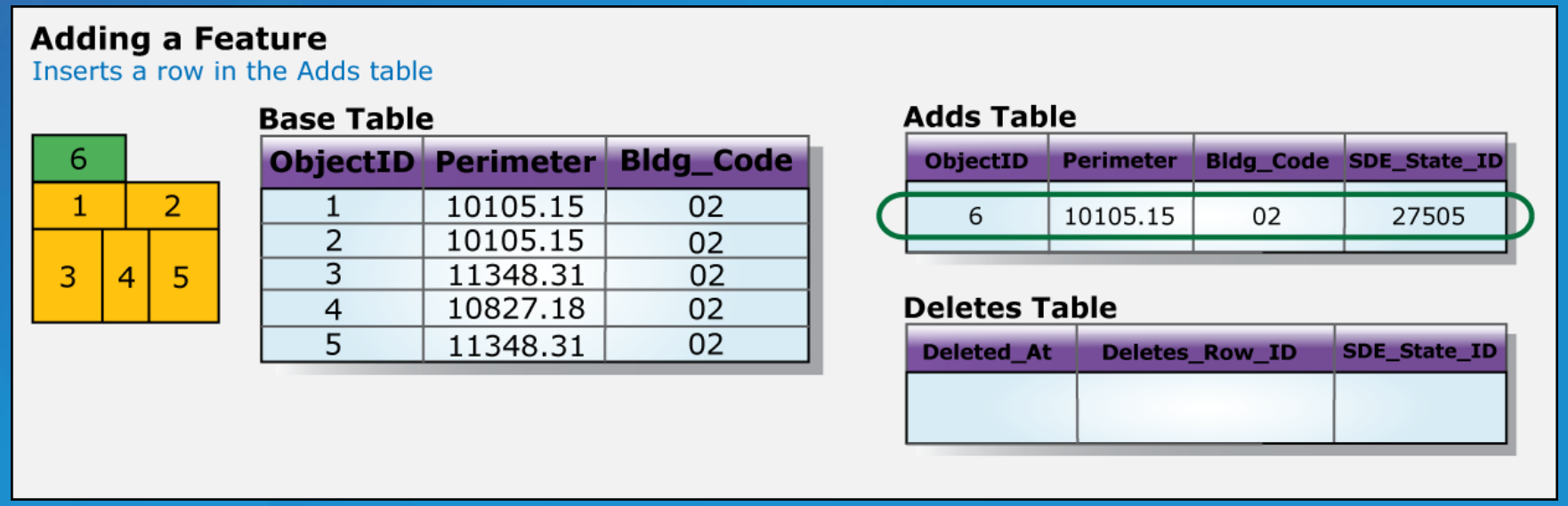
How Versioning Works

- Registered as Versioned
  - Creates Adds and Deletes tables for tracking edits



How Versioning Works

- Adding Features
  - Record added to the Adds Table
  - Version will be referenced (SDE_State_ID Field)

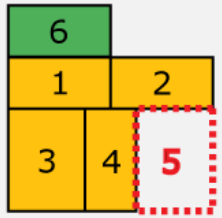


How Versioning Works

- 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

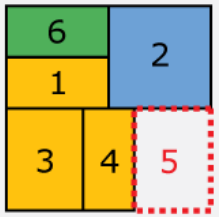
Deleted_At	Deletes_Row_ID	SDE_State_ID
27506	5	0

How Versioning Works

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

Updating a Feature

Inserts a row in both the Adds and Deletes tables



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
2	20210.30	02	27507

Deletes Table

Deleted_At	Deletes_Row_ID	SDE_State_ID
27506	5	0
27507	2	0

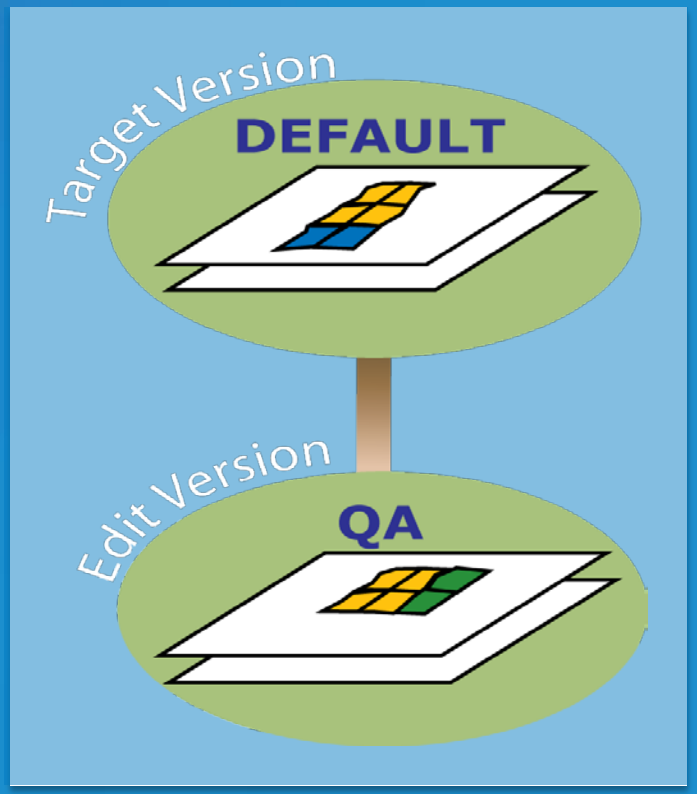
Good Housekeeping Tip

For Geodatabases

Use versioned editing for discrete edits

Versioned Editing – Reconcile and Post

- How can versions be merged?



SDE Default Version

The DEFAULT version is the root version

- It is the ancestor of all other versions.
- The default version should never be edited directly.

Versioning

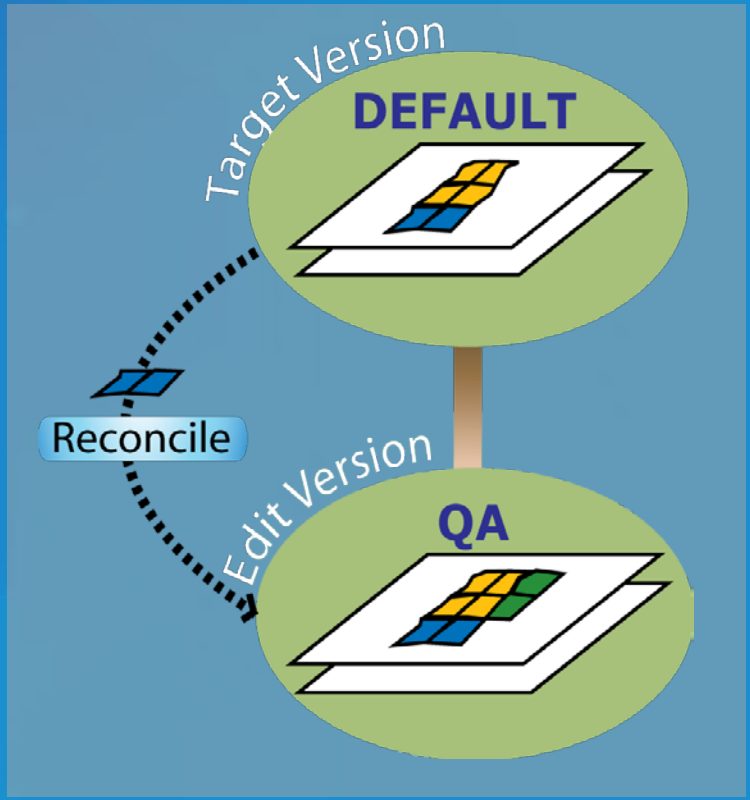
Key Concepts – Reconcile and Post

Reconcile	In concurrency management, to merge all modified data in the current database edit session with a second version of the data.
Post	During versioned Geodatabase editing, the process of applying the current edit session to the reconciled target version.

Explicit Reconcile	An explicit reconcile occurs when the user editing the child version clicks the Reconcile tool on the Versioning toolbar and chooses a parent version to reconcile with.
Implicit Reconcile	An implicit reconcile occurs when two users are editing the same version at the same time and the second user saves.

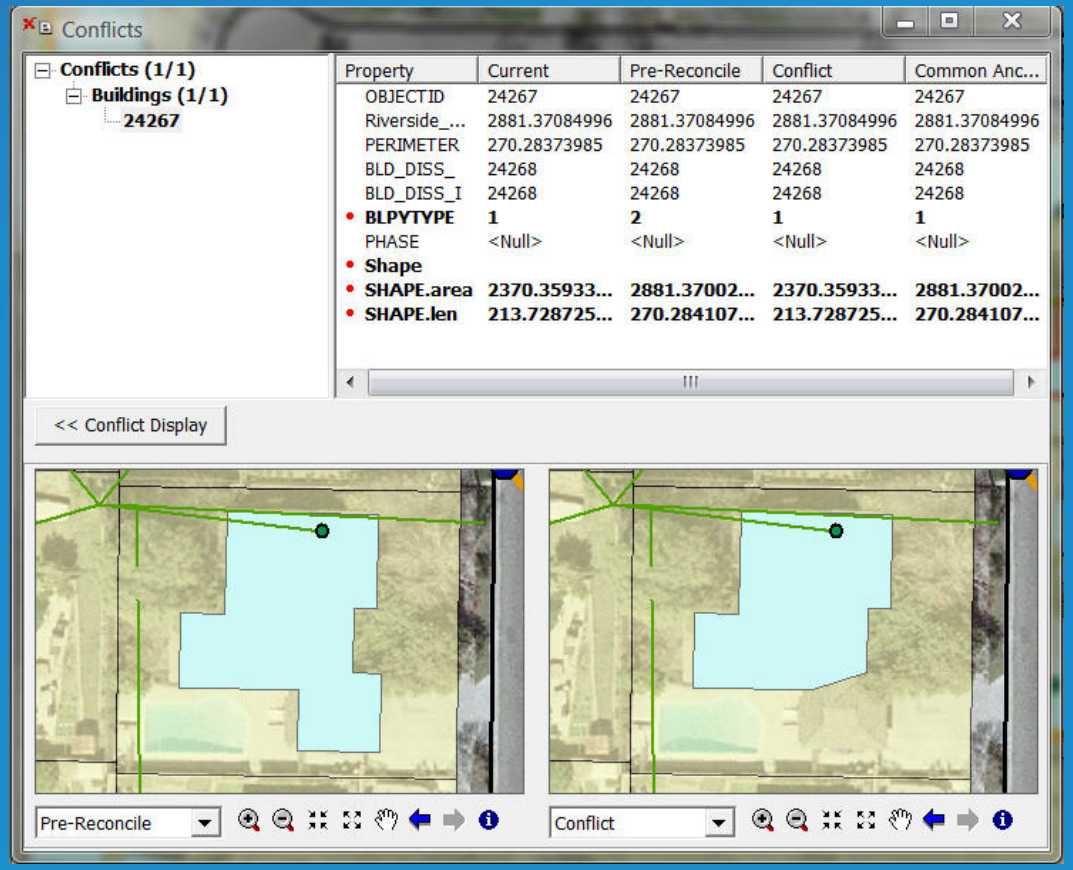
Versioned Editing – Reconcile

- Incorporate changes from the target version



Reconcile and Conflicts

- No locks on edit
  - Data overwritten?
  - Conflict detection
- Conflict Resolution Dialog



A

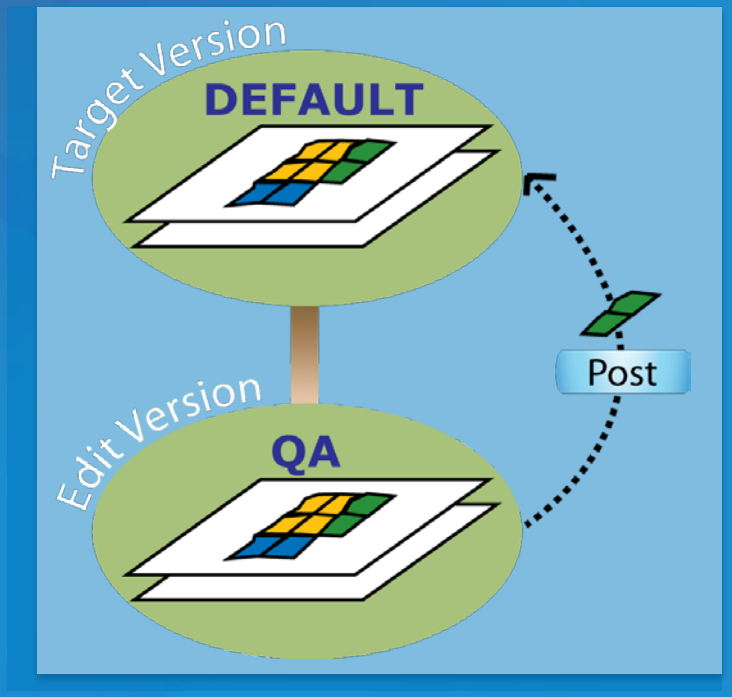
Versioning

Key Concepts - Conflict

Conflict	In database editing, a state of incompatibility that occurs when multiple users simultaneously edit a version or reconcile two versions. Conflict occurs when the same feature or toplogically related feature are edited in two versions, and it is unclear which representation is valid
Conflict Resolution	The process of solving uncertainty within a database that occurs when two versions of the same data are edited at the same time. Conflicts can occur when multiple users simultaneously edit the the same feature or toplogically related feature, or reconcile two versions of a dataset. Resolving a conflict requires that the user make a decision about the features correct representation and identify it in the Conflict Resolution dialog box.

Versioned Editing – Post

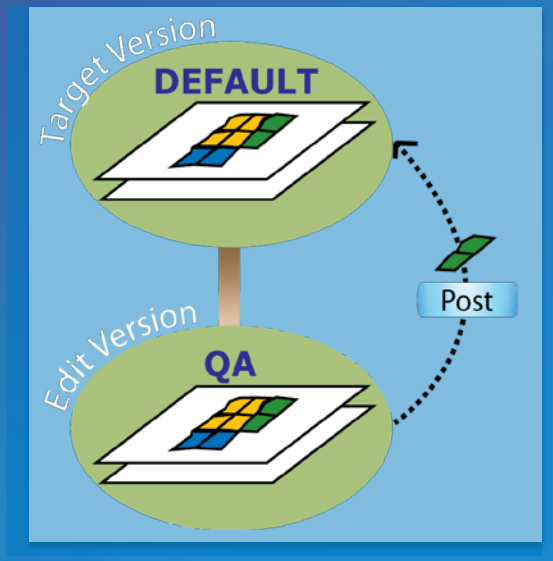
- Incorporate with target version



  - After a post versions are identical

Versioned Editing – Compress

- How can edits in the A & D (“delta”) tables make it to the Base table for a versioned feature class?



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

  - After a post to the DEFAULT version, a compress will move the edits from the A & D (“delta”) tables to the Base table.

Compress

A compress moves edits from the DEFAULT version to the base table and cleans up unreferenced states in the delta tables and system tables thus making them smaller

Performance can be enhanced when stale statistics are re-calculated and skewed indexes are re-organized/rebuilt.

Compress

Disconnect all user from the geodatabase before running a compress for best results.

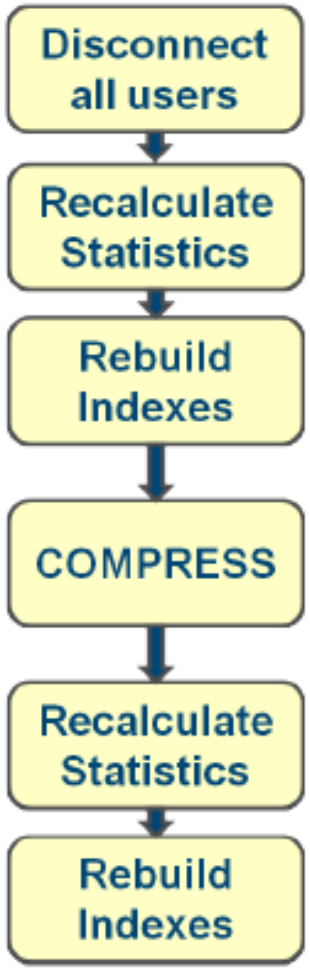
If users are logged on, the lineages that they are accessing will not be compressed.

The result is fewer states will be moved to the base table.

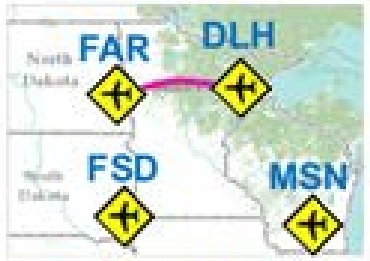
If using replication, synchronize before the compress.

Compress workflow

A compress operation can be time consuming, especially if heavily edited versioned feature classes have never been compressed. The following workflow will speed up the performance of the compress operation as well as the performance of the geodatabase after the compress is complete.



Representing versions



ANDREW

Airports

ObjectID	Shape	Airport	...
1	●	DLH	...
2	●	FSD	...
3	●	MSN	...

A7

ObjectID	Shape	Airport	...	State_ID
4	●	FAR	...	3

D7

SDE_State_ID	Deletes_Row_ID	Deleted_AT

Air_Routes

ObjectID	Shape	...

A8

Robert LeClair

ObjectID	Shape	...	State_ID
1		...	3
2		...	4
3		...	5
4		...	6

D8

SDE_State_ID	Deletes_Row_ID	Deleted_AT

SDE_VERSIONS

Name	State_ID	...
DEFAULT	0	...
ANDREW	3	...
OLIVIA	6	...

SDE_TABLE_REGISTRY

Registration_ID	Table_Name
...	...
7	Airports
8	Air_Routes

SDE_MVTABLES_MODIFIED

State_ID	Registration_ID
3	7
3	8
6	8

SDE_STATES_LINEAGES

Lineage_name	Lineage_ID
0	0
1	0
4	0
1	3
4	6

SDE_STATES

State_ID	Owner	...	Parent_State_ID	Lineage_name
0	SDE	...	0	0
3	ANDREW	...	0	1
6	OLIVIA	...	0	4

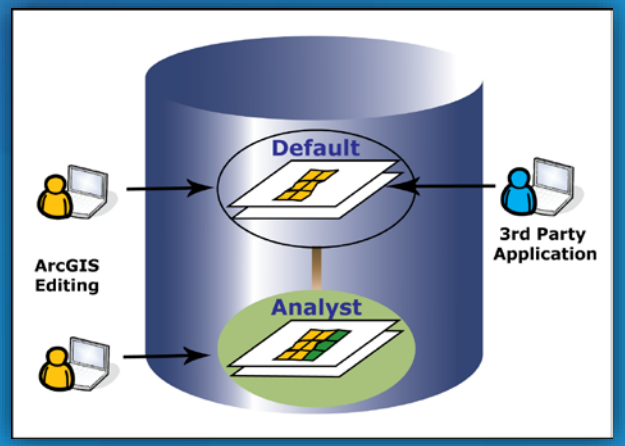
Good Housekeeping Tip

For Geodatabases

Do a compress regularly. Nightly for a regular compress and on a weekly basis for a full compress.

Versioned Editing - Move to Base Option

- Hybrid
  - versioned and non-versioned



IT integration

Database constraints

- Simple data only
  - Points, lines, polygons, annotation, relationship classes

Good Housekeeping Tip

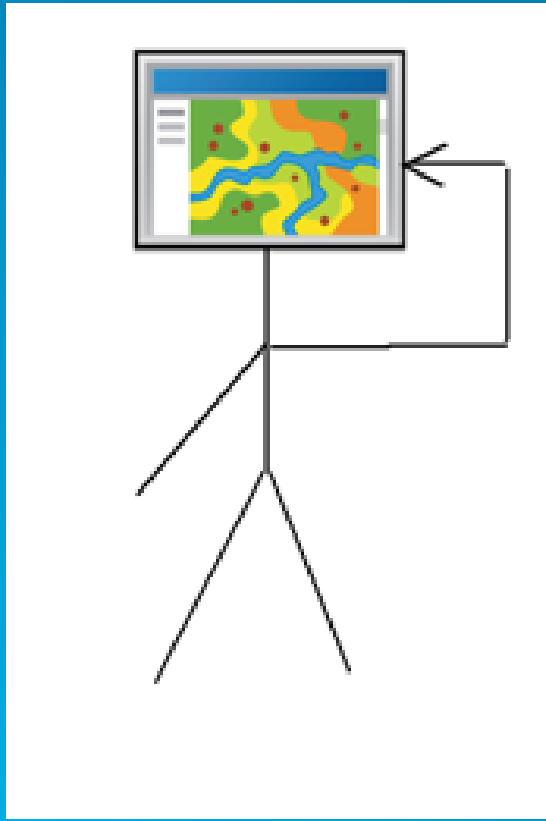
For Geodatabases

Don’t use Move Edits to Base Option for Versioning unless there is a compelling business reason to do so.

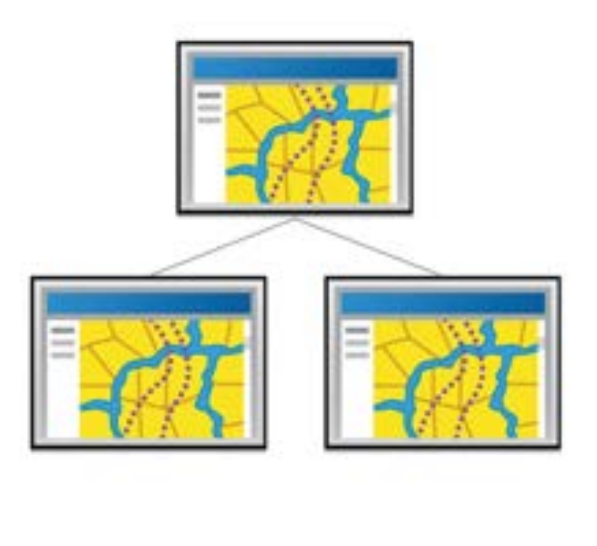
For 3rd party integration use versioned views.

Version Tree Scenarios

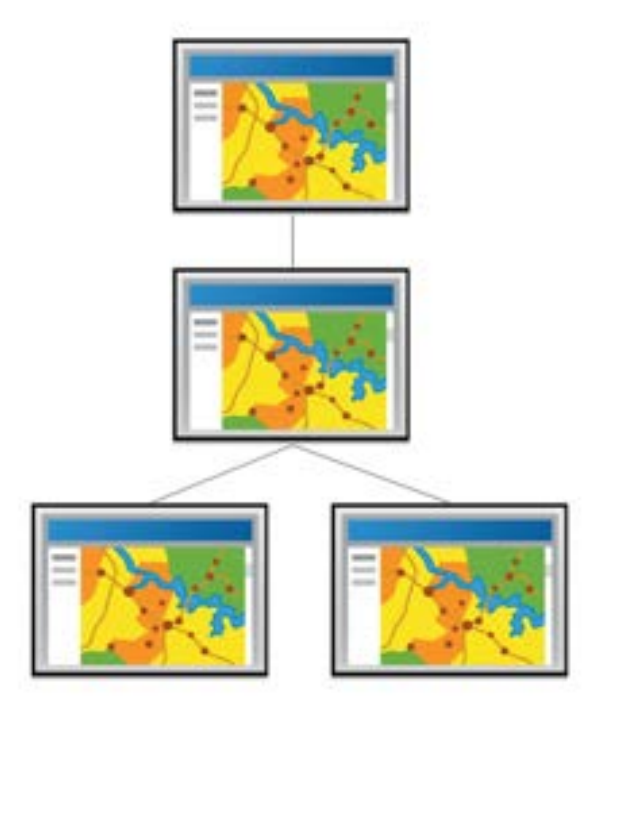
- All editors directly edit DEFAULT



- Editors edit in Separate Versions from DEFAULT



- QA Tier off of DEFAULT



Good Housekeeping Tip

For Geodatabases

Pick the appropriate versioning workflow for your environment but don’t go overboard on version hierarchies since too many levels could adversely impact performance.

Permission Options for Versions

Use them wisely

Private – Only the owner can view and edit the datasets in the version.

Protected – Any user can view the datasets in the version, but only the owner can edit them.

Public – Any user can view and edit the datasets, provided the user has been granted permission on the datasets.

Good Housekeeping Tip

For Geodatabases

Protect the DEFAULT version

Bulk Loading of Data

Examples of bulk changes include:

- Appending records from another data source
- Field calculations
- Synchronize changes from replicas

Tools used to perform bulk loads

- Simple Data Loader
- Object Loader
- Append geoprocessing tool
- Field Calculator
- Synchronize

Data Loaders

Choose carefully

Simple Data Loader (ArcCatalog)

- Pro
  - Fast way to load into simple feature classes
  - Can use on a versioned feature class as long as no one is editing the version being loaded into
- Con
  - Data loss and error messages can occur if other editors are editing the same version during a data load.
  - Does not validate data while loading
  - Only works on simple feature classes
  - Load does not occur within an edit session so no way to roll back changes after a load is complete

Data Loaders

Choose carefully

Object Loader(ArcMap)

- **Pro**
  - Load records for complex feature classes
  - Allows data loading on versions where other editors are simultaneously editing
  - Loads occur in an edit session so possible to stop editing without saving changes.
- **Con**
  - Loading must be performed in a ArcMap edit session
  - Performing validation may slow load performance

Good Housekeeping Tip

For Geodatabases

Pick the appropriate loading tool

Reconciling After a Bulk Load

In a versioned environment there are implications for performance as bulk changes are stored in the delta tables.

For a bulk data load, synchronization, or field calculation, all the changes are stored with the same state id.

Recommended workflow for bulk loading is to:

Create a child version → **Perform bulk load** → **Reconcile and post child to the parent – BEFORE the parent is updated.**

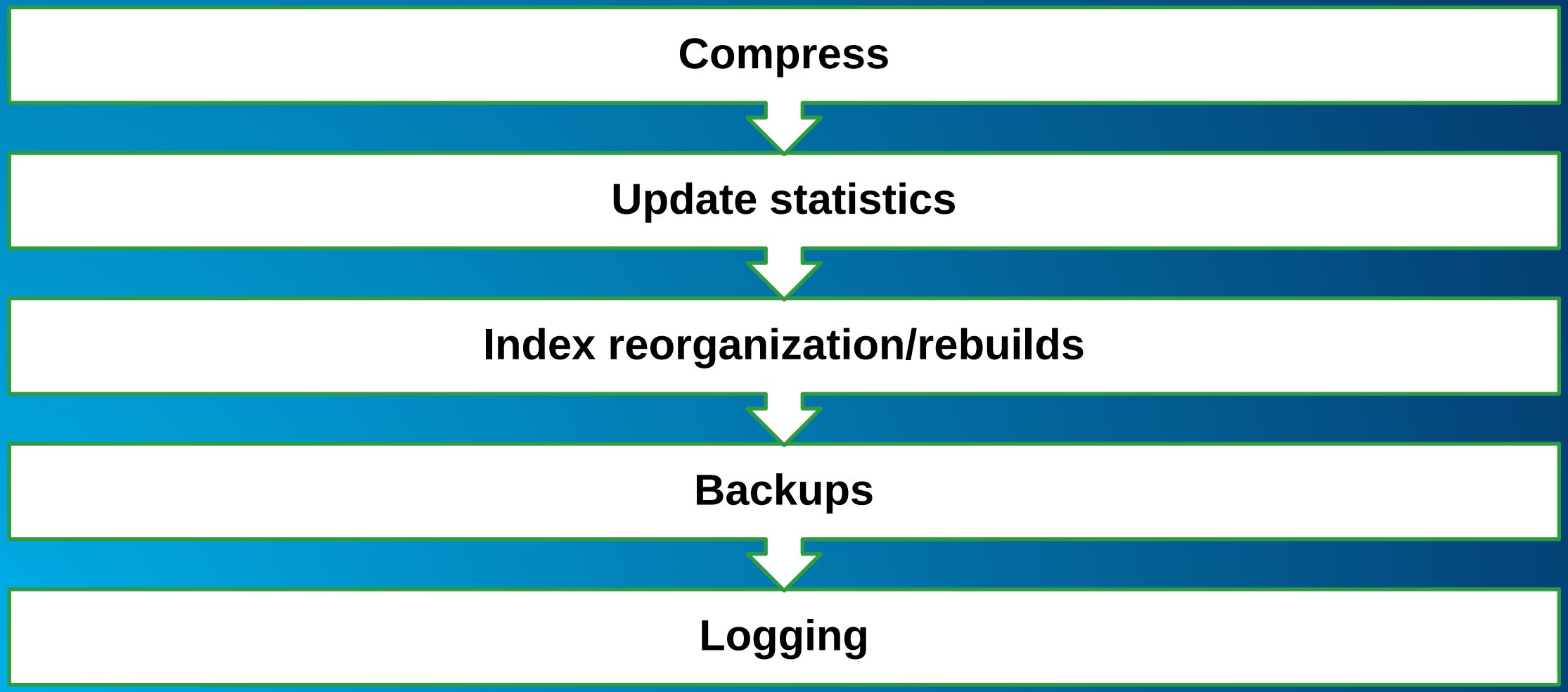
Good Housekeeping Tip

For Geodatabases

Follow the recommended workflow for bulk loading

Common Management Tasks

Avoid at your own risk…



Statistics and Indexes

Update statistics regularly

- After a bulk load
- After lots of edits
- Before and after a compress

Reorganize/Rebuild indexes

- When fragmented
- After a full compress

How do I make sure that my data is safe?

- Backups allow you to recover from:
  - Media failures
  - User errors
  - Hardware failures
  - Natural disasters
- Also, backups can be used for:
  - Copying or moving data between servers
  - Setting up database mirroring or AlwaysOn

Good Housekeeping Tip

For Databases

If using SQL Server, remember that the database name has to be the same when restored to another server or instance.

Points to remember

Backups are the only way to reliably protect your data

1. **Decide how much time you can afford to lose when disaster strikes and data must be restored**
2. **Create a restore plan that will achieve that goal**
3. **Create a backup plan that supports your restore plan**
4. **Implement your plan**
5. **Test your recovery plan regularly by using real backup media to restore to a system capable of being used in production**

Good Housekeeping Tip

For Databases

Backup your data and practice restoring it

Geodatabase archiving

The mechanism for capturing, managing, and analyzing data change.

Record and access changes made to all or a subset of data in a geodatabase.

Geodatabase archiving allows users to connect to an historical version.

Connect to an historical version using an existing historical marker or a specified moment

Enabling Archiving

In Catalog tree, use Enable archiving geoprocessing tool

Can enable on nonversioned data or versioned data using the DEFAULT version.

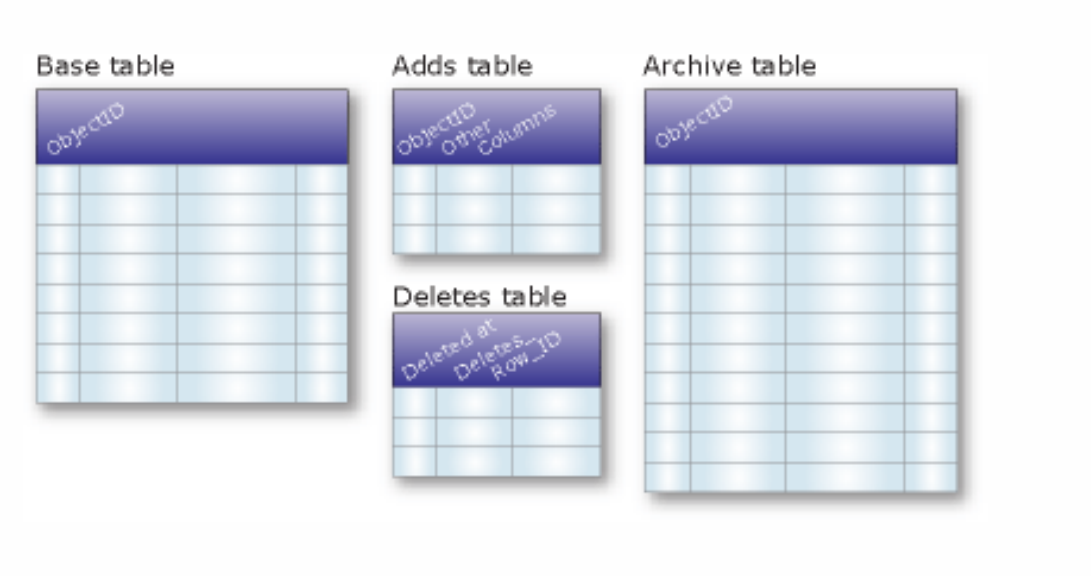
Enable on nonversioned data

The geodatabase creates additional date attributes in the base table for that dataset.

These are called gdb_from_date and gdb_to_date.

Enable archiving on versioned data.

For each dataset for which arciving is enabled, a new archive class is created.



All attributes and all rows in the DEFAULT version of the dataset or object class are copied to the archive class.

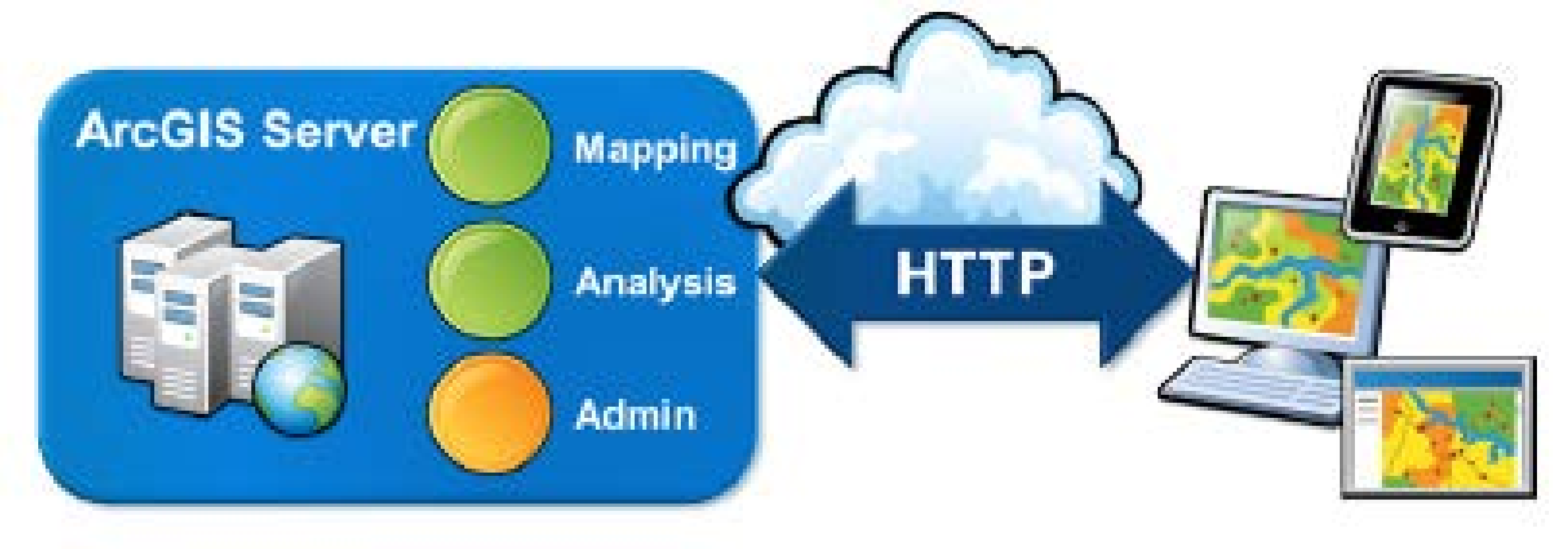
Representing Time

How ArcGIS represents time when change is recorded

History can be recorded as:	**Valid time**	**Moment in time at which a change occurred in the real world.**	
	Transaction time	**Time an event was recorded in the database.**	**Used by archiving on versioned data.**
	Coordinated Universal Time (UTC)	**Primary standard used to regulate time over the Internet.**	**Used by archiving on nonversioned data.**

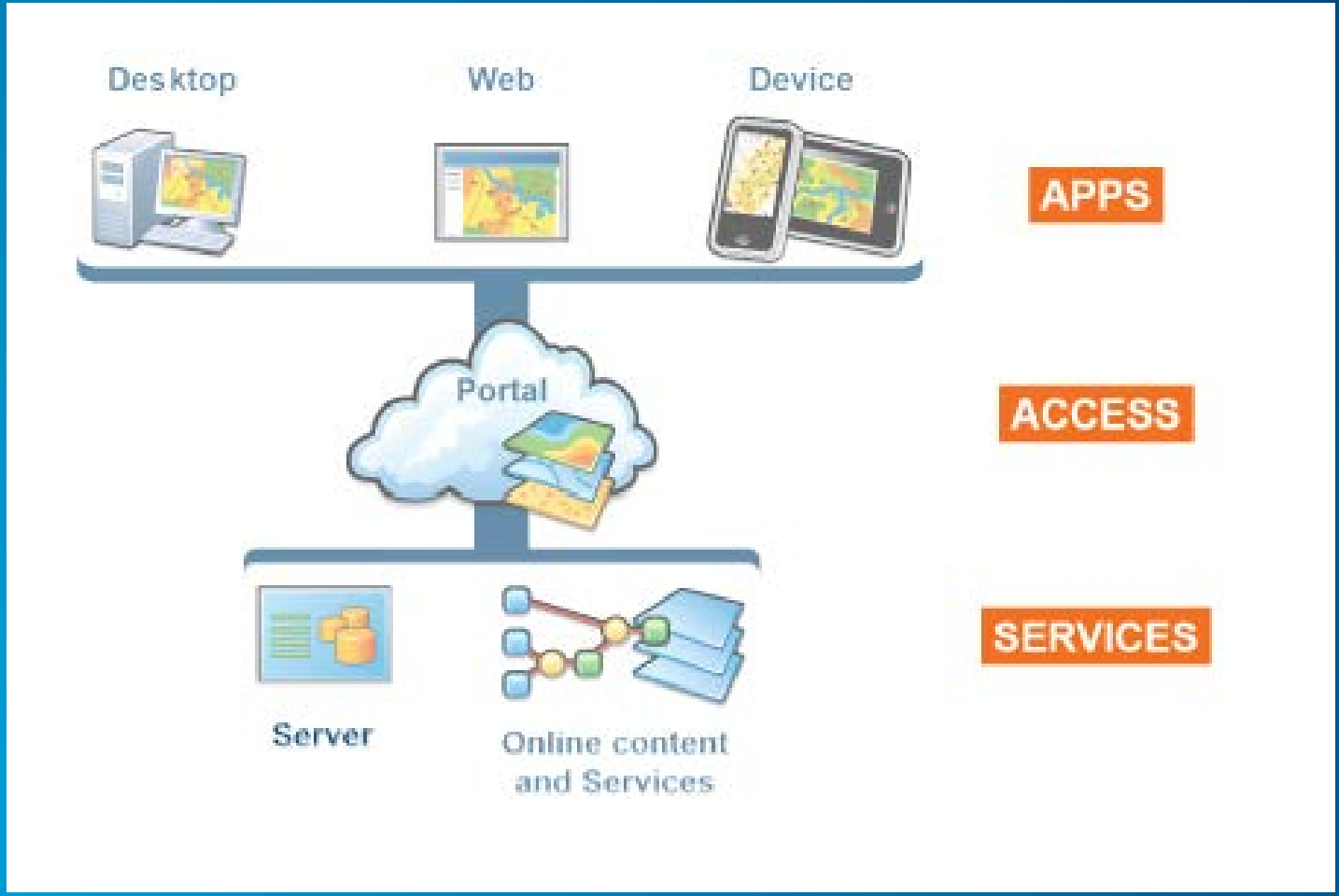
ArcGIS for Server

- ArcGIS Server is a pure web services environment so functionality is exposed through services



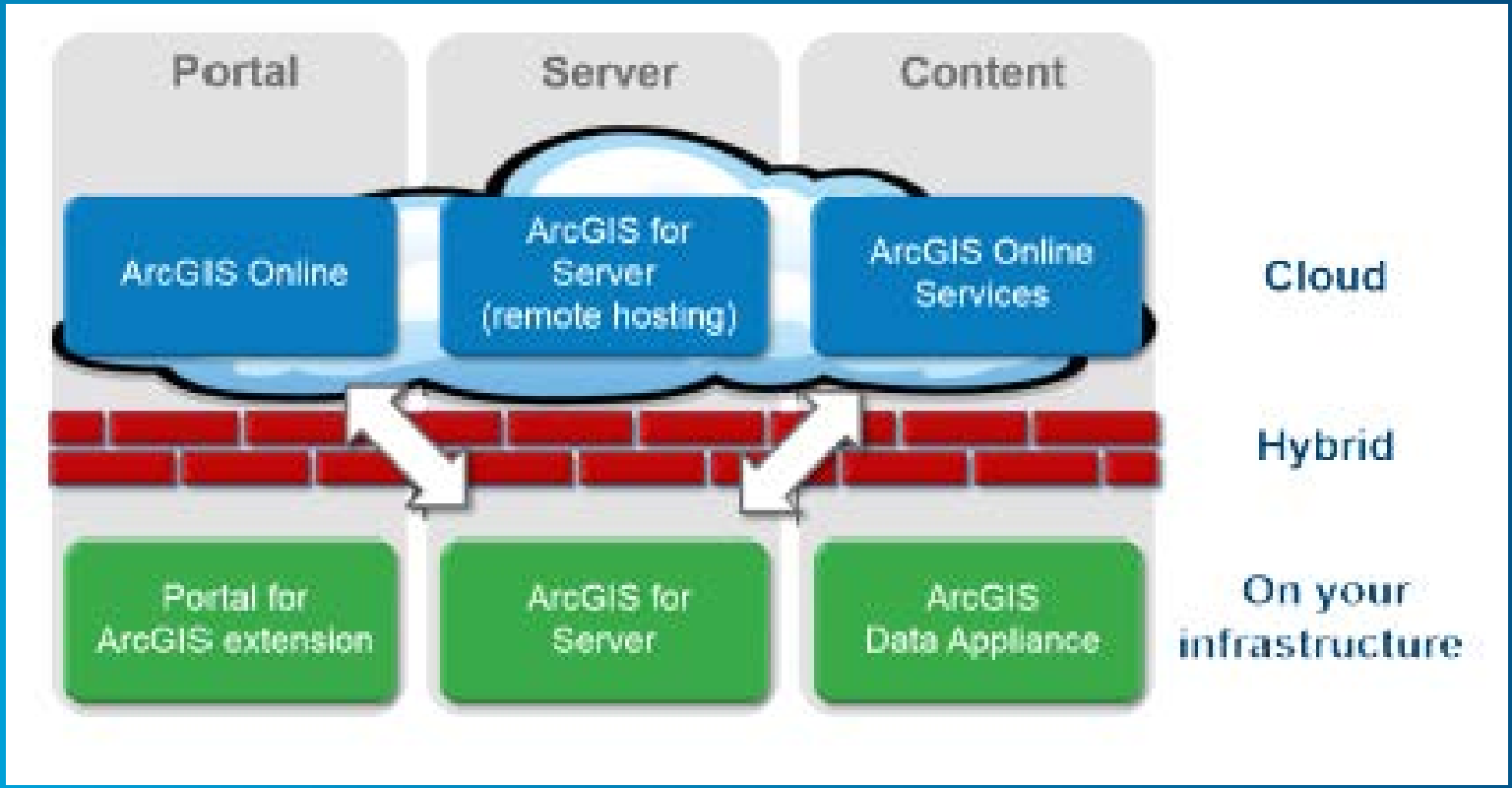
Web GIS

- **Web GIS enables everyone to share authoritative, up-to-date geographic information to support collaborative, informed decision making**

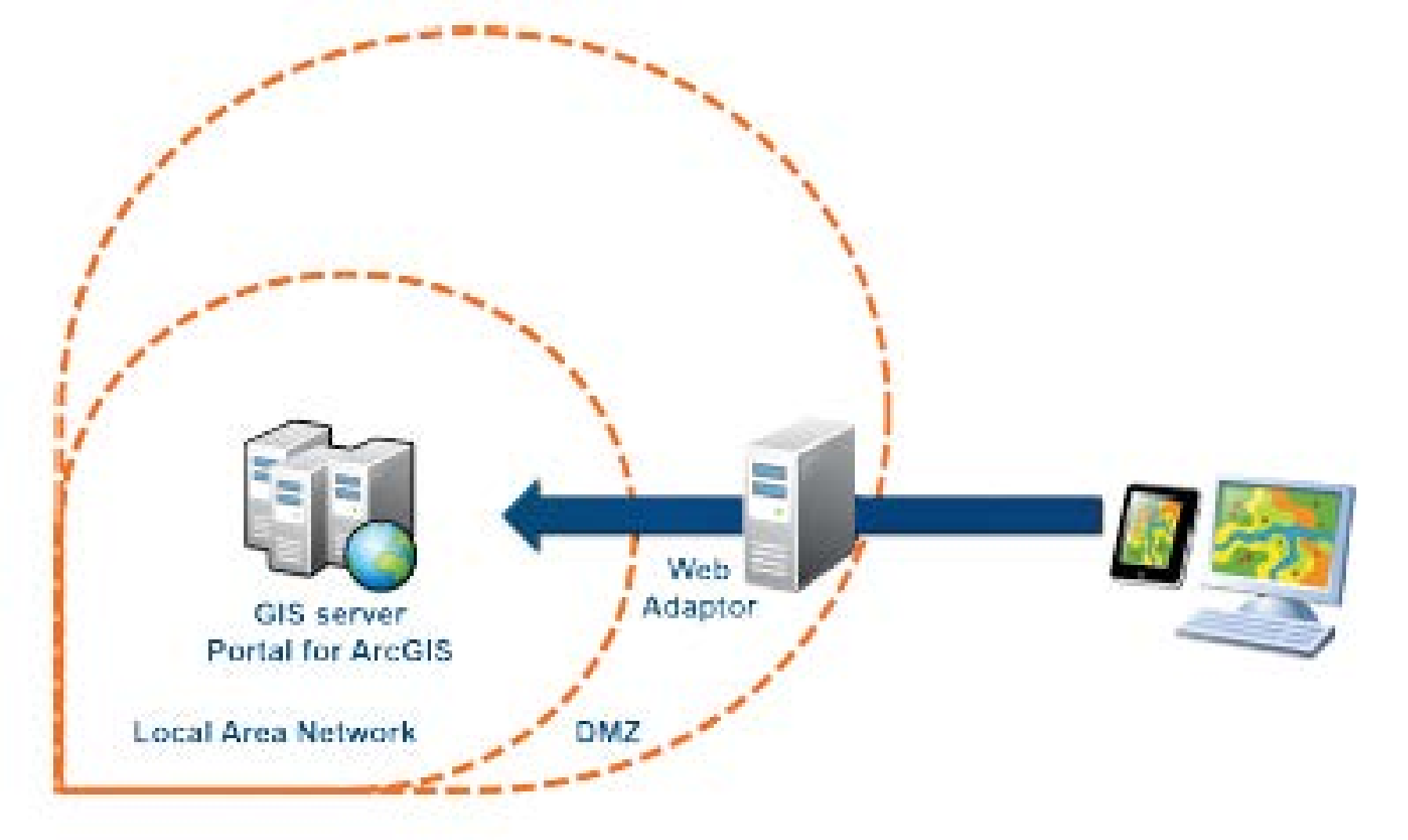


Web GIS Implementation

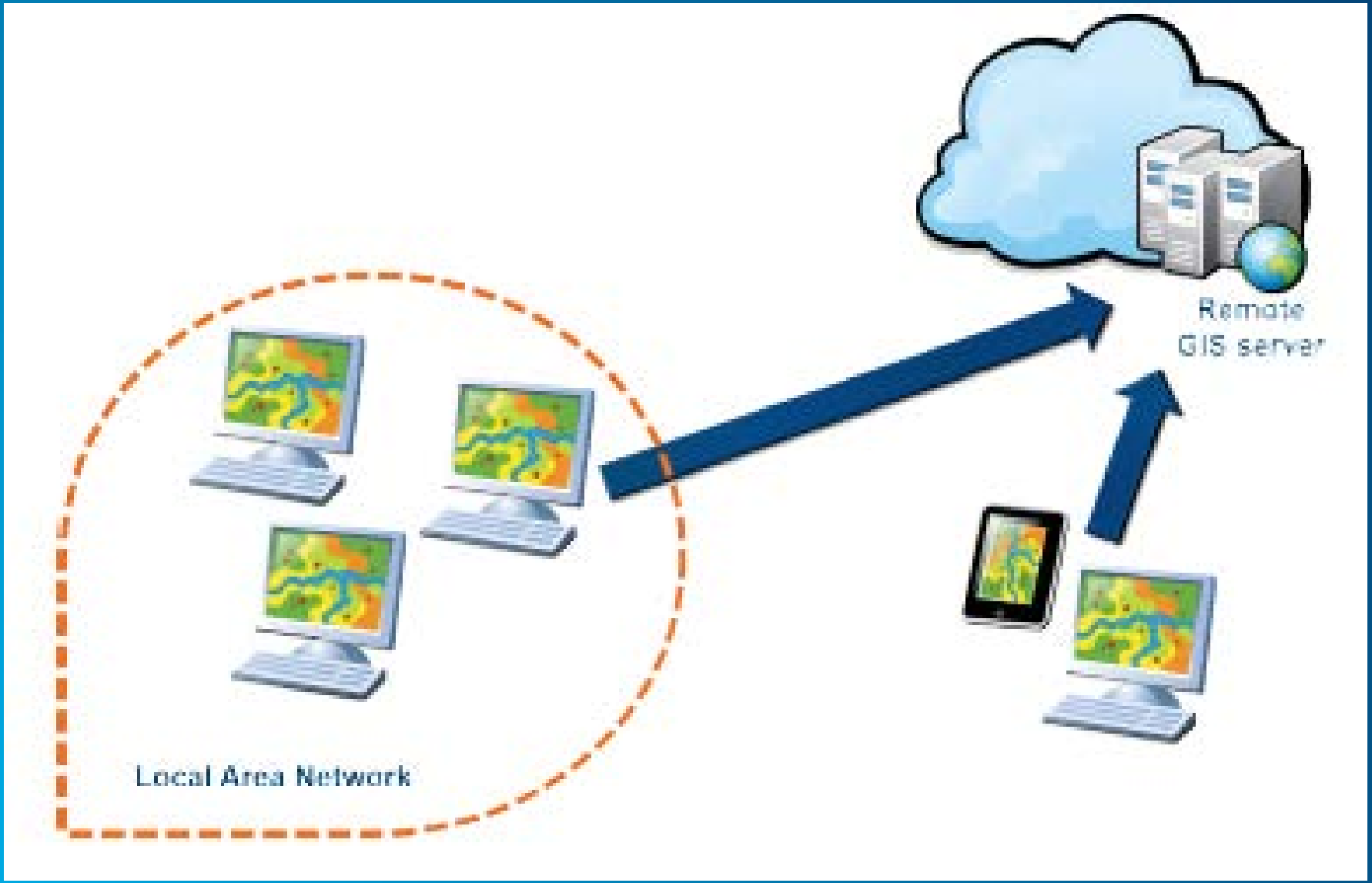
- **ArcGIS for Server is an integral part of implementing the ArcGIS Platform and web GIS.**



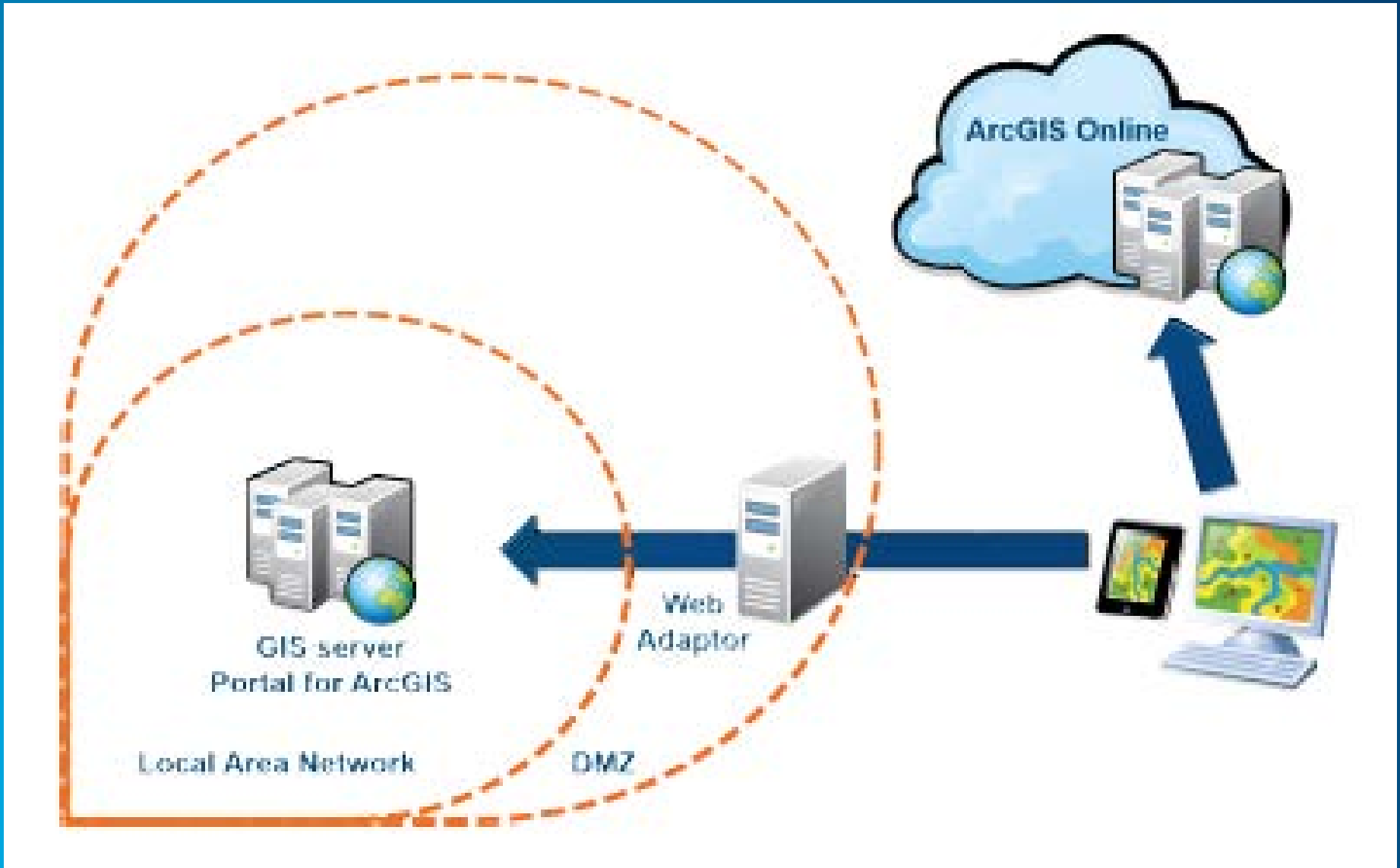
Implement on your own infrastructure



Implement in the Cloud



Hybrid Implementation

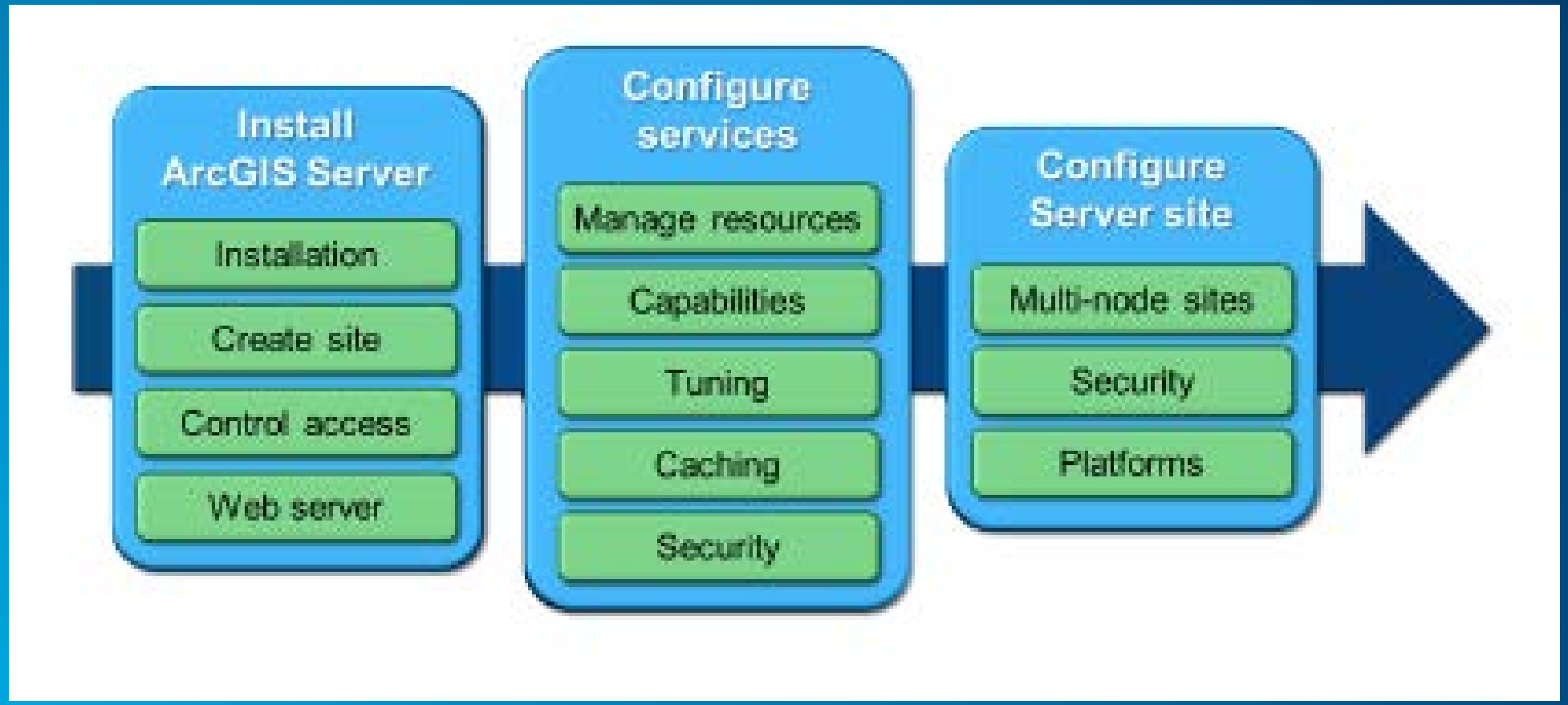


Good Housekeeping Tip

For ArcGIS for Server

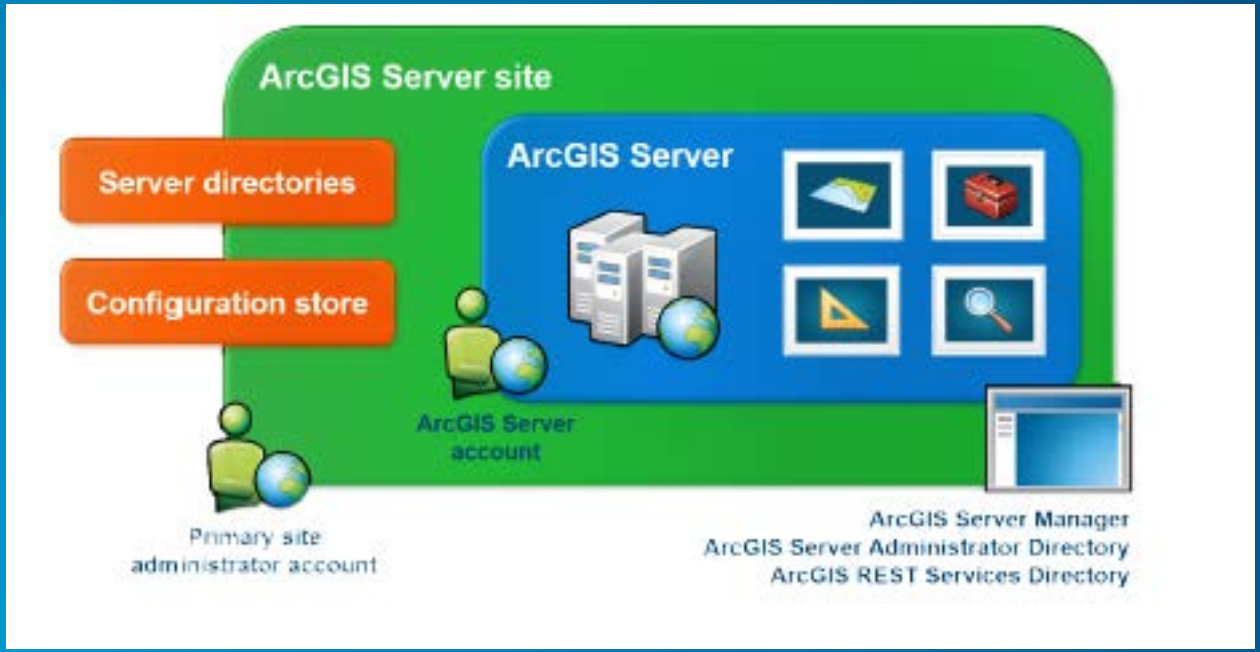
Consider your particular business needs when choosing an optimal deployment pattern.

Workflow for administering ArcGIS for Server



ArcGIS Server Site Architecture

- An ArcGIS Server site is a deployment of ArcGIS Server. A site is composed of one or more ArcGIS Servers working cooperatively to answer client requests.



What is a resource?

- **Resources include map documents, data, geoprocessing workflows, caches, and other files required by ArcGIS Server to create services and fulfill service requests.**

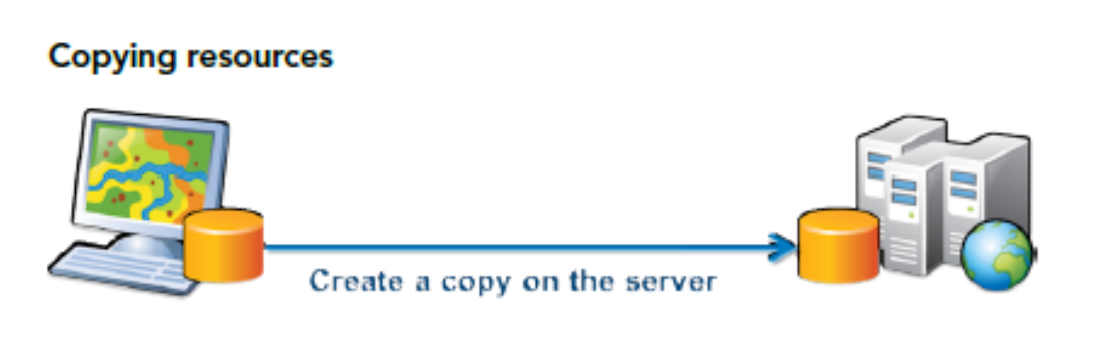
Good Housekeeping Tip

For ArcGIS for Server

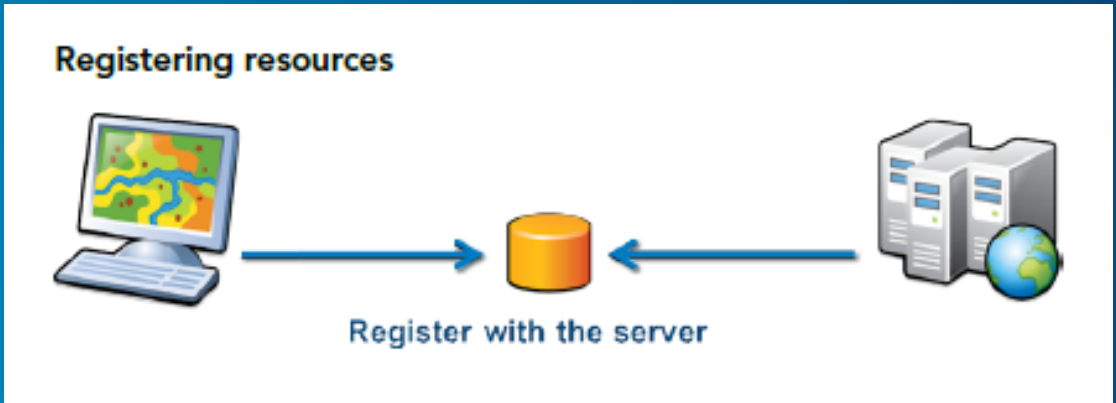
Retain the map documents used to publish a service even though ArcGIS for Server will not continue to reference them.

Copy or register resources?

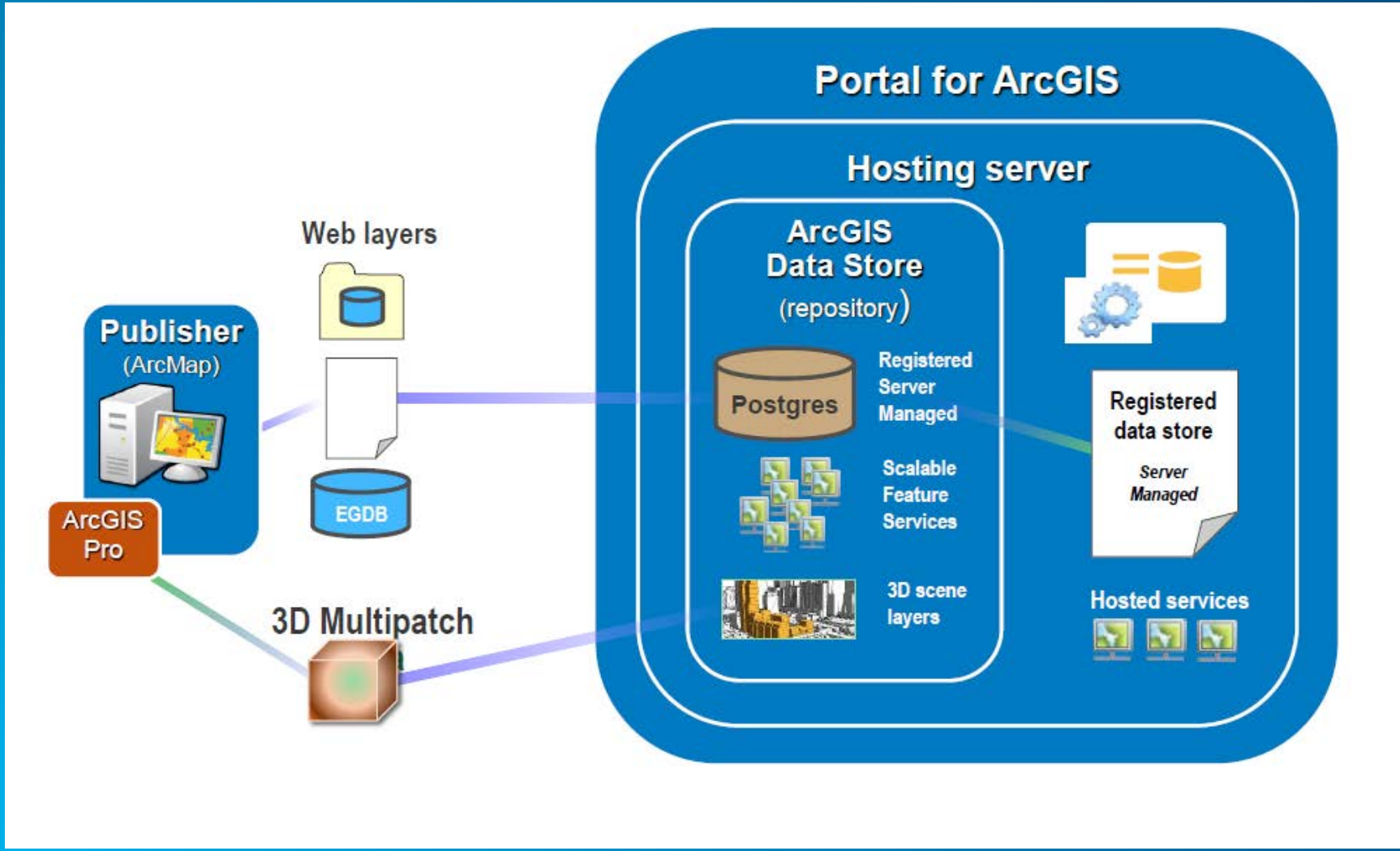
- Copying resources



- Register resources



ArcGIS Data Store



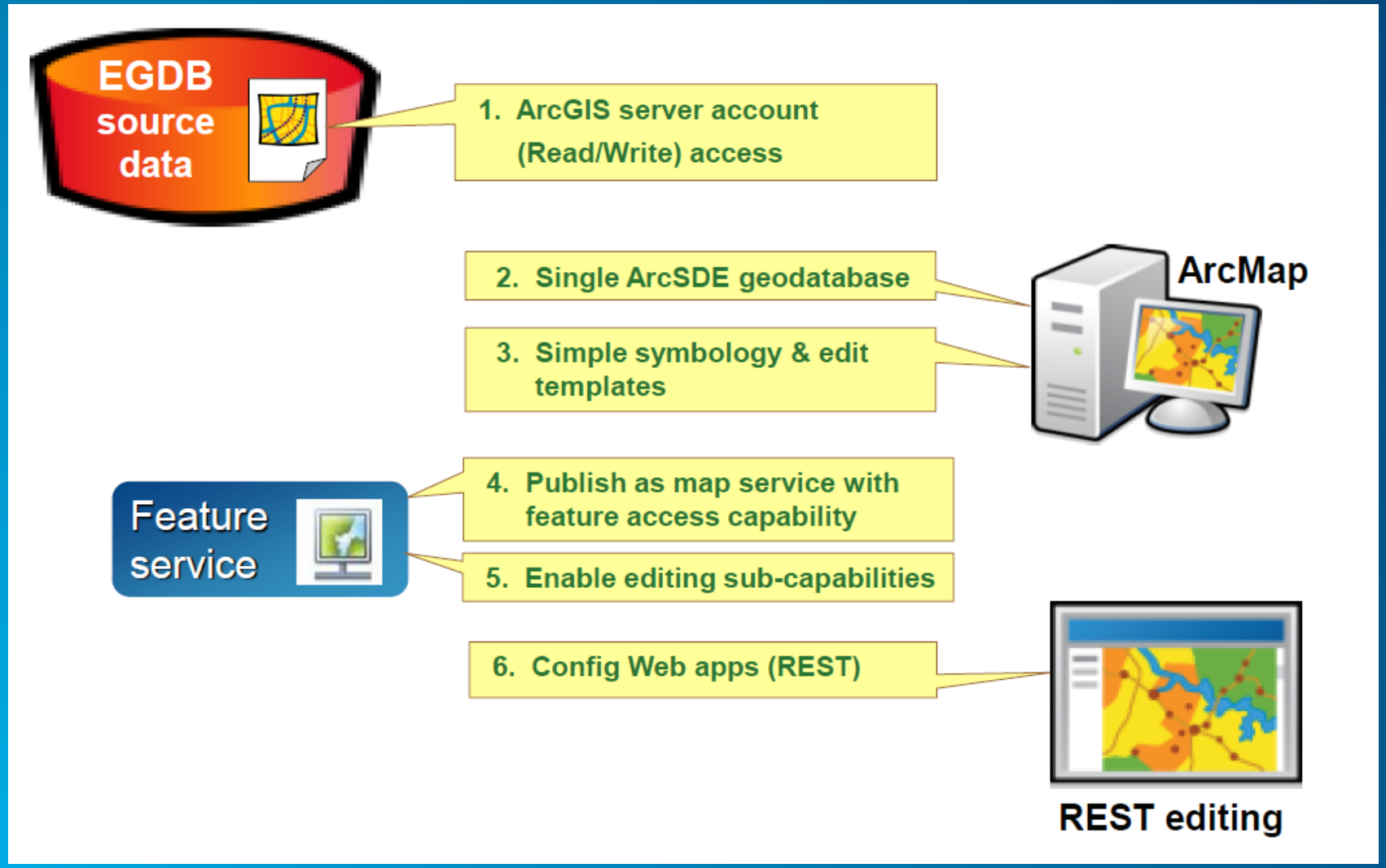
Good Housekeeping Tip

For ArcGIS for Server

Remember to register your data stores for file geodatabases and enterprise geodatabases

1. Make resources accessible
2. Grant ArcGIS Server access
3. Register a location in the data store

Feature edting configuration workflow



Service Capabilities

Each service type has capabilities that can be selectively enabled or disabled to customize the functionality available through your services.

Map Service Capabilities

Capability	What it does	Operations allowed	Special requirements
Mapping	Provides access to the contents of a map document through SOAP and REST URLs	Data, Map, and Query	Always enabled for any map document
WCS	Uses raster layers in a map document to create a service compliant with the OGC Web Coverage Service specification	None	Requires raster layers
WMS	Uses a map document to create a service compliant with the OGC Web Map Service specification	None	None
Feature Access	Provides access to vector features in the map; often used for editing through the ArcGIS web APIs	Create, Delete, Query, and Update	Requires vector layers

Map Service Capabilities (continued)

Capability	What it does	Operations allowed	Special requirements
Schematics	Allows viewing, generating, updating, and editing schematic diagrams	Editing and Query	Requires schematic layers
Mobile Data Access	Allows data extraction from a map document to a mobile device	Uploads	None
Network Analysis	Solves transportation network analysis problems using the ArcGIS Network Analyst extension	None	Requires a network analysis layer referencing a network dataset
KML	Uses a map document to create KML features	Separate Images, Single Image, and Vectors	None
WFS	Uses layers in a map document to create a service compliant with the OGC Web Feature Service specification	None	Requires vector layers (raster layers are not included in the service; WFS serves vector feature geometry)

Image Service Capabilities

Capability	What it does	Operations allowed	Special requirements
Imaging	Provides access for display and analysis of raster data	Image, Catalog, Download, Edit, Mensuration, Metadata, and Pixels	Always enabled
WCS	Creates a service compliant with the OGC Web Coverage Service specification	None	None
WMS	Creates a service compliant with the OGC Web Map Service specification	None	None

Geoprocessing Service Capabilities

Capability	What it does	Operations allowed	Special requirements
Geoprocessing	Provides access to run a geoprocessing workflow	Uploads	Always enabled
WPS	Creates a service compliant with the OGC Web Processing Service specifications	None	None

Geodata Service Capabilities

Capability	What it does	Operations allowed	Special requirements
Geodata	Provides access to a geodatabase	Query, Data extraction, and Replication	Always enabled
WCS	Makes raster contents of a geodatabase available as an OGC Web Coverage Service	None	None
WFS	Makes vector contents of a geodatabase available as an OGC Web Feature Service	None	None

Geocoding Service Capabilities

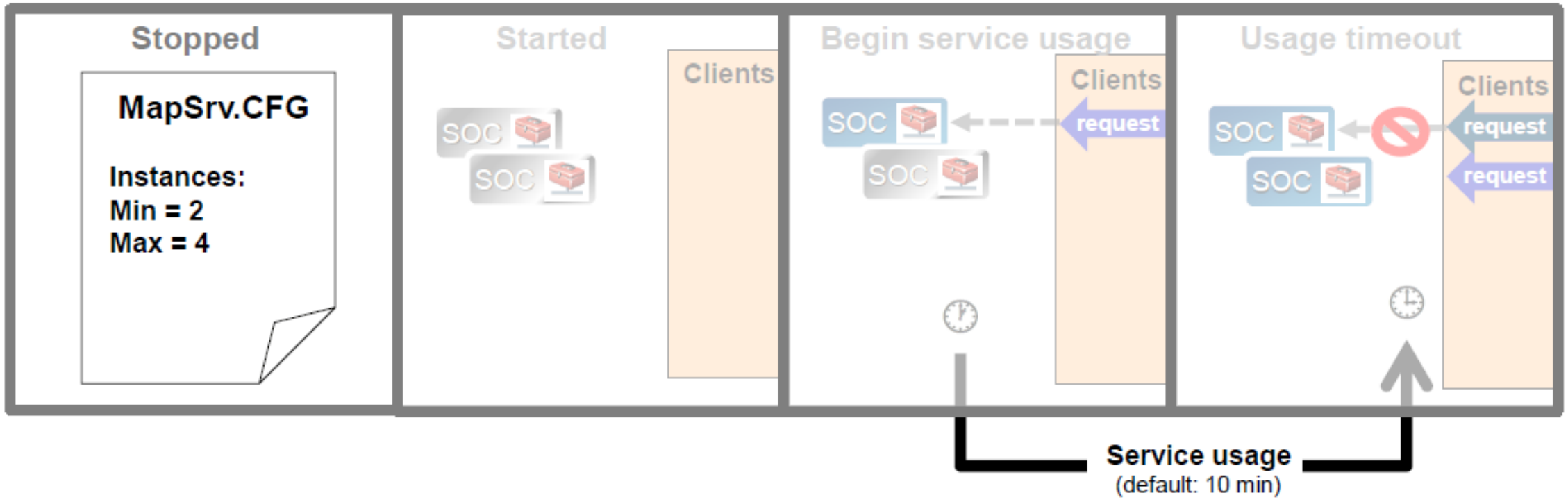
Capability	What it does	Operations allowed	Special requirements
Geocode	Determines location for an address	None	None
Reverse Geocode	Determines address for a location	None	None

Good Housekeeping Tip

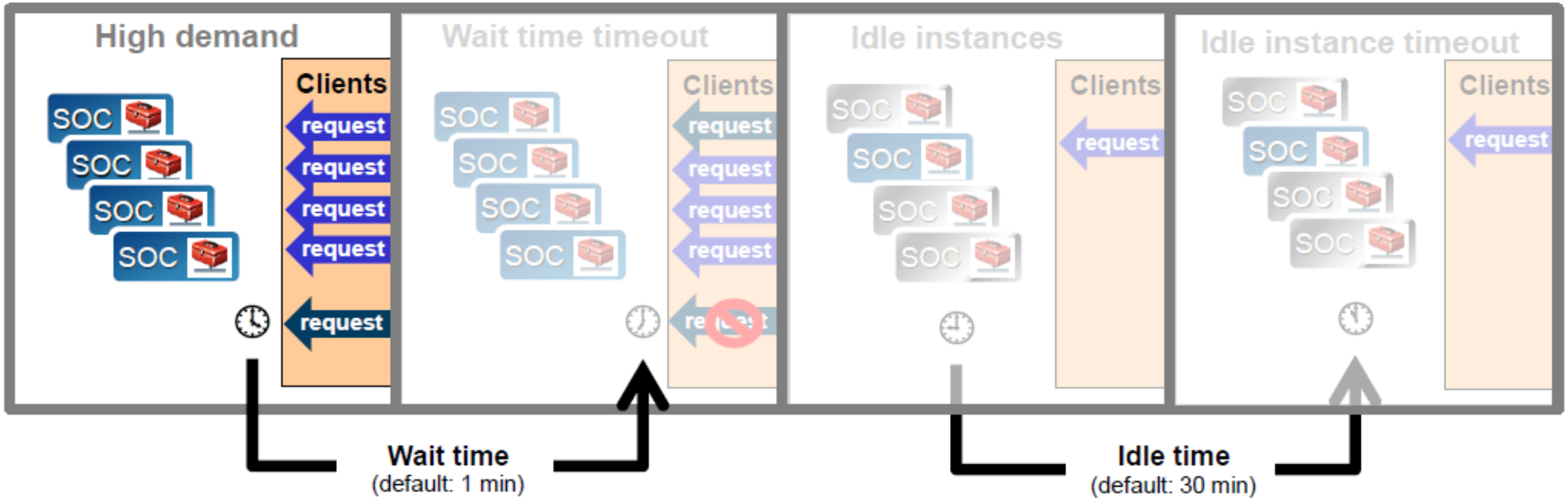
For ArcGIS for Server

Choose the service to meet your functionality requirements

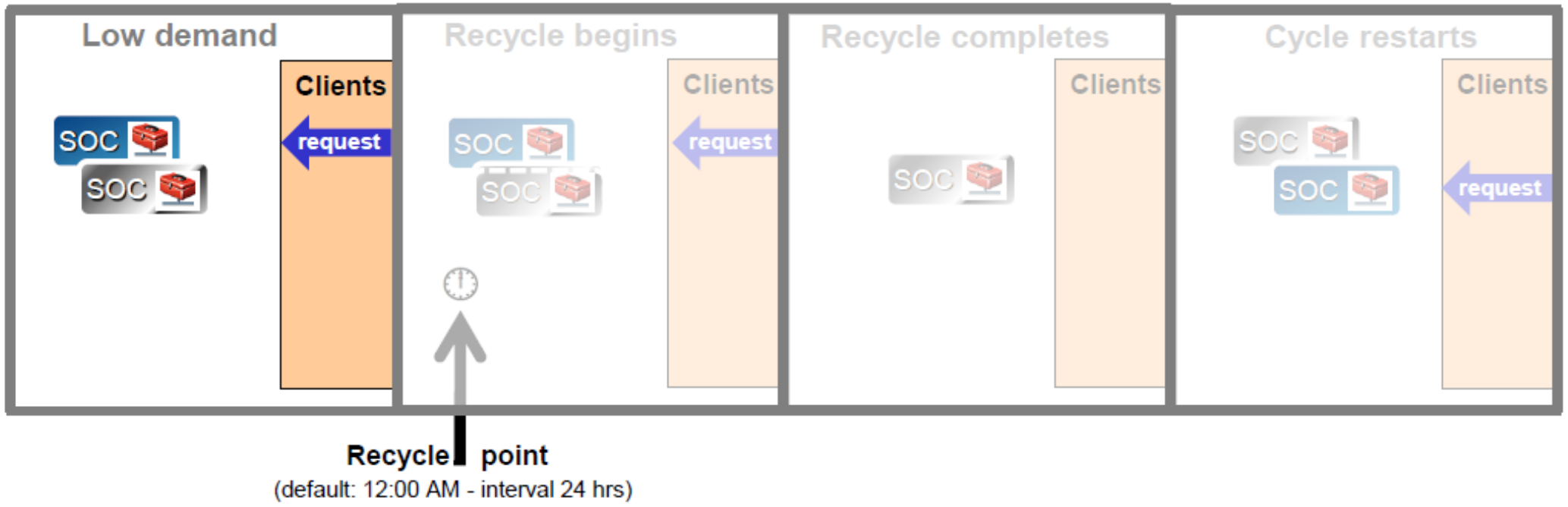
Lifecycle of a GIS service



Lifecycle of a GIS service



Lifecycle of a GIS service



Tuning Services

- **Pooling Tab**

Property	Description	When to adjust it
Minimum number of instances per machine	Minimum number of instances of the service to create on each node within the cluster. **Default = 1**	• Increase if there are frequently more requests than available instances. • Decrease if there are frequently more available instances than requests.
Maximum number of instances per machine	Maximum number of instances of the service to create on each node within the cluster. **Default = 2**	• Increase if clients are frequently waiting for an available instance. • Decrease for infrequent requests.
The maximum time a client can use a service	Maximum number of seconds an instance can service a request. **Default = 600 seconds (10 minutes)**	• Increase for requests that may take longer (e.g., geoprocessing jobs). • Decrease for services that complete quickly to free system resources.
The maximum time a client will wait to get a service	Maximum number of seconds the framework will wait for a free instance of the service before rejecting a service request. **Default = 60 seconds**	• Increase on a busy system.
The maximum time an idle instance can be kept running	Maximum number of seconds an idle instance of a service must be kept alive before it can be destroyed. **Default = 1800 seconds (30 minutes)**	• Increase if new instances are created frequently. • Decrease for infrequently requested services to free system resources.

Tuning Services

- **Processes Tab**

Property	Description	When to adjust it
Recycling Settings	Duration between service recycling. **Default = 24 hours at 00:00**	• Decrease to reclaim resources and clear HTTP connection issues. • You require that a service recycles at a different time of day.
Health Checks	Duration between data connections check & repair (idle instances). **Default = 30 minutes**	• Decrease if SDE connections frequently break.

Tuning Services

- **Parameters Tab**

Property	Description	When to adjust it
Maximum Number of Records Returned by Server	Maximum number of records returned for a service request. **Default = 1000**	• Increase if the service will be returning large numbers of records when fulfilling requests. This will increase network usage and possibly overload the client.
Lock Database Schema	Specifies if the map service will acquire schema locks for map layers that come from a geodatabase. **Default = True (checked)**	• Disable when you need to make schema changes to your data without stopping your services.

Good Housekeeping Tip

For ArcGIS for Server

Service properties can be used to tune services to optimize performance given specific demand requirements

esri

Understanding our world.