

Planning an Enterprise Geodatabase

Rasu Muthurakku, Enterprise Solutions Architect / Consultant

Andrew Sakowicz, Enterprise Implementation Team Lead

Agenda

Planning Enterprise Geodatabase

- Overview
- Key Factors
- Design
- Architecture
- Build
- Workflows
- Maintenance

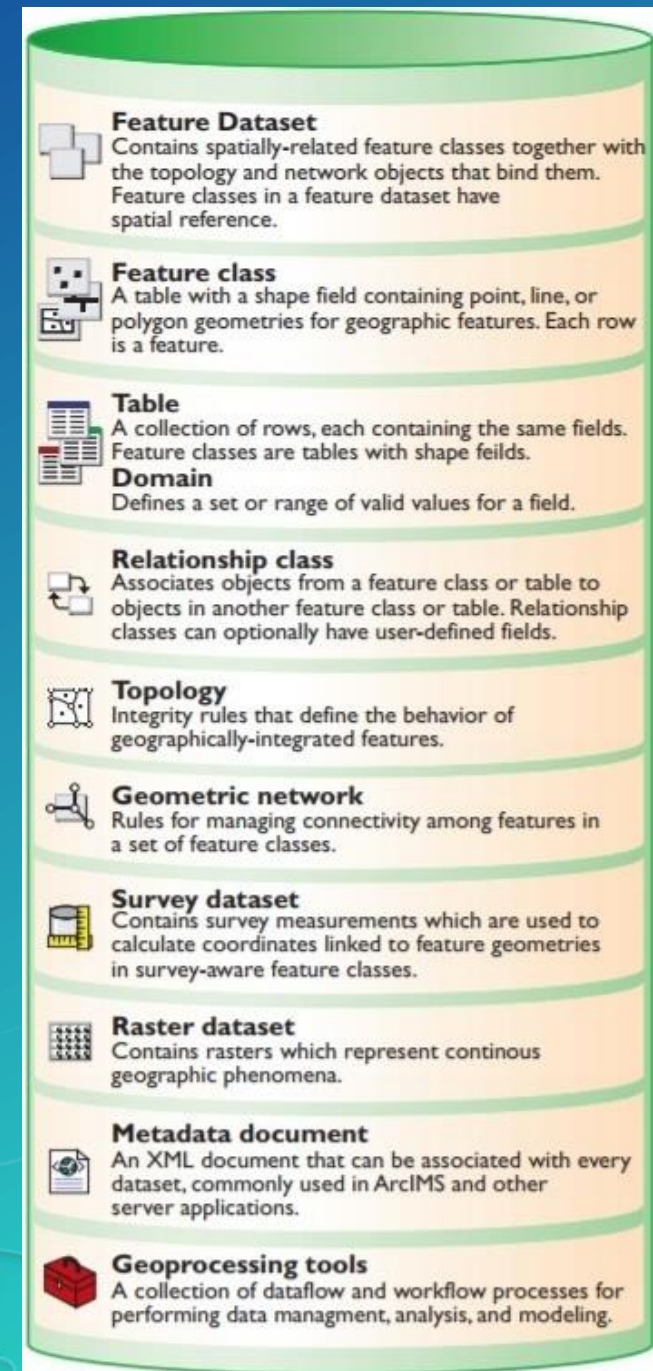
Key Considerations, Best Practices and Recommendations / Lessons Learned!

Overview

What is a Geodatabase (GDB)?

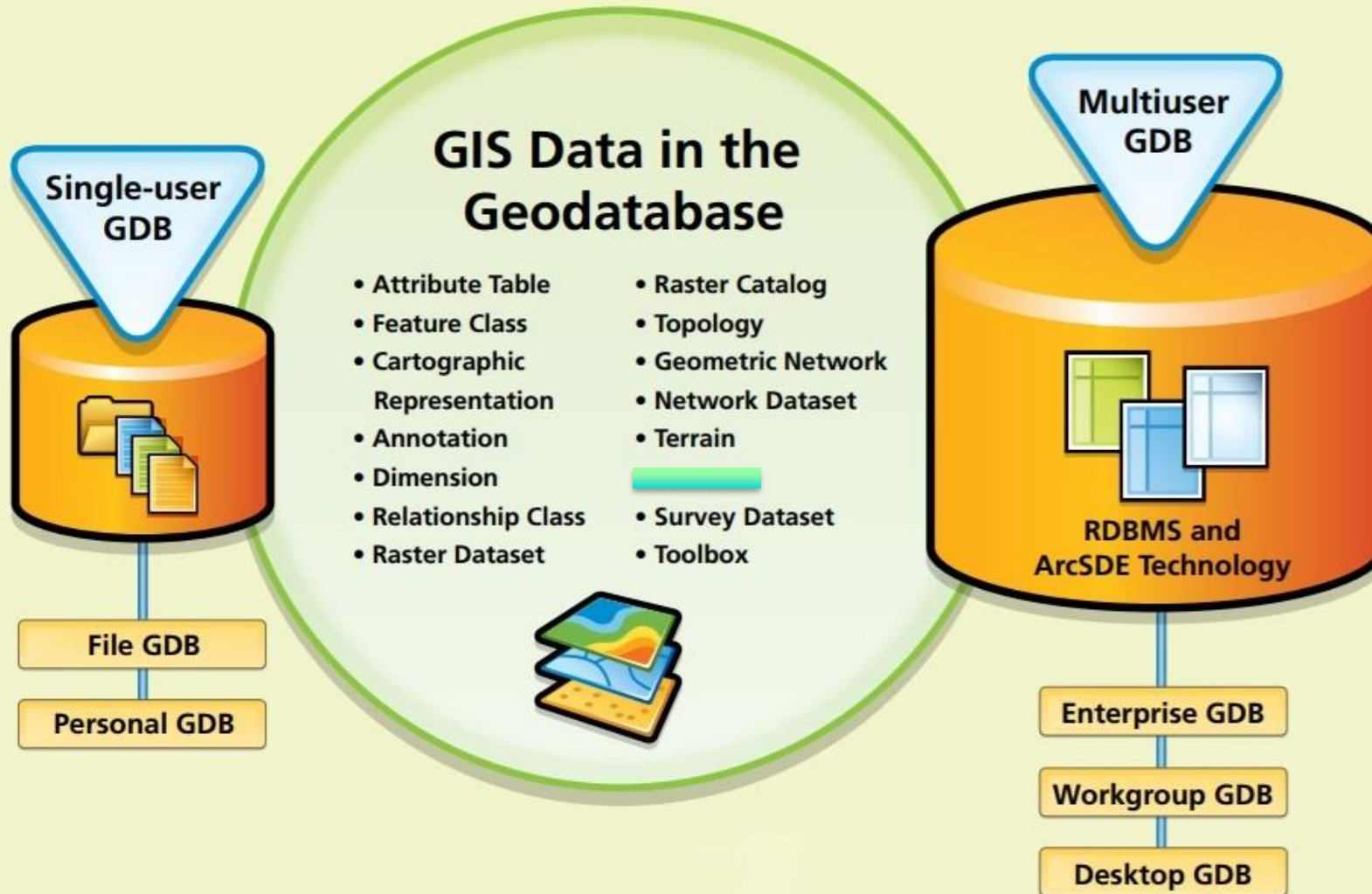
- Collection of Geographic Datasets of Various Types Stored in:
 - Common File System Folder
 - *Microsoft Access Database*
 - Multiuser Relational DBMS* / ArcSDE
 - Oracle
 - Microsoft SQL Server
 - PostgreSQL
 - Informix, IBM DB2
- Native Data Structure for ArcGIS
- Primary Data Format Used for Editing and Data Management
- Comprehensive Approach to Modeling and Managing Spatial Data

* DBMS – Database Management System



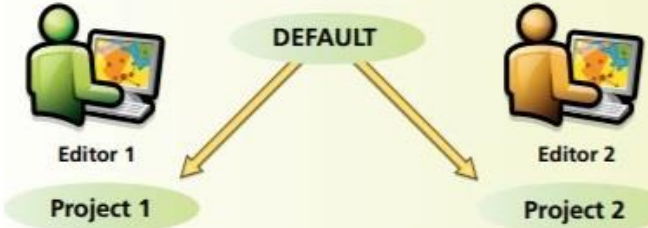
Planning an Enterprise Geodatabase

Types of Geodatabases



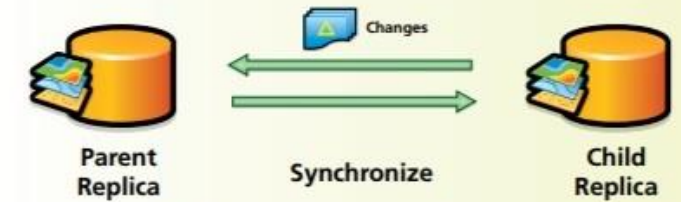
Functionality

Versioning



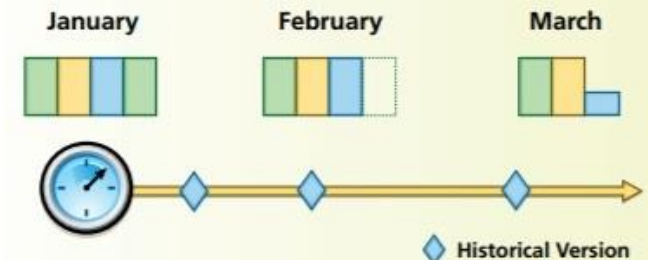
Versioning is the framework that enables multiple users to access and edit the same data simultaneously and provides long transaction (i.e., database changes that span long periods of time) support.

Geodatabase Replication



Enables GIS data to be shared across two or more geodatabases. Data changes can be made in each geodatabase, then synchronized. Two-way, one-way, and checkout/check-in replication workflows are supported.

Geodatabase Archiving



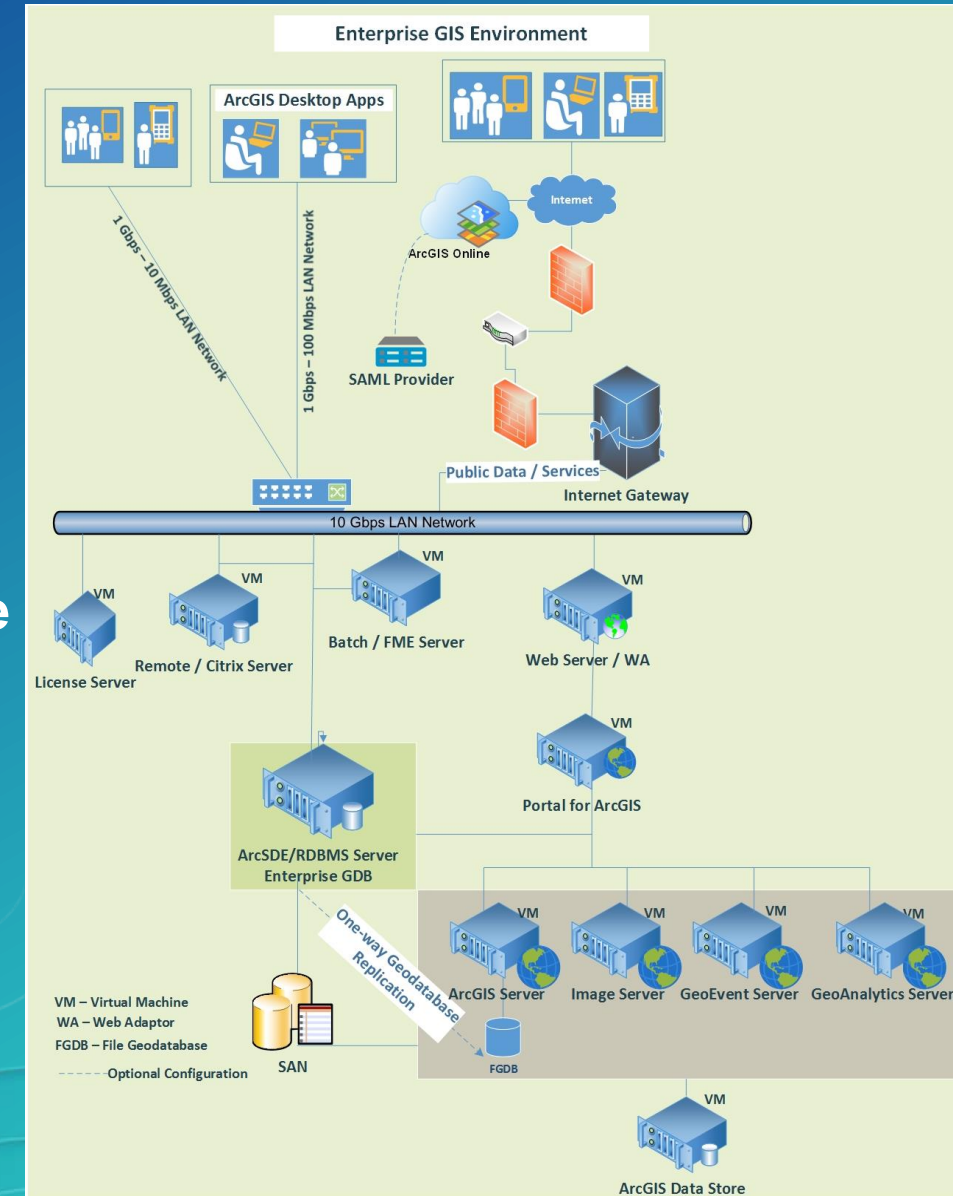
When enabled on a dataset, archiving captures any and all changes made to the dataset in the DEFAULT version of the multuser geodatabase.

What is an Enterprise Geodatabase?

Centralized Multiuser Geodatabase

- ArcSDE Enables the RDBMS* for GIS data management
 - Scalability
 - Reliability
 - Security
 - Backup
 - Integrity, etc.
- Extremely Large, Continuous and Centralized GIS Database
- Many Simultaneous Users
- Long Transactions and Versioned Workflows
- SQL Types for Spatial in all Supported RDBMSs
- High Performance for a Very Large Number of Users

* RDBMS – Relational Database Management System

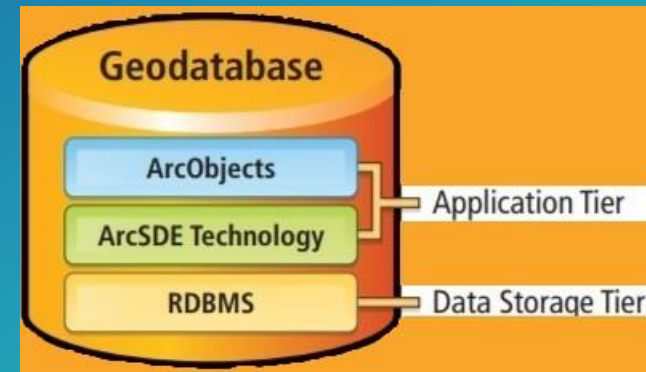


Key Factors

Expensive Rework

Prevention is Better Than Cure!

- Data Model Designs
 - *Change in XY Tolerance*
 - *Multiple Projections*
 - *Objectid as Primary Key for Relationship Classes, etc.*
- Data Conversion Specifications
 - *Extra Vertices*
 - *Insufficient QA/QC*
 - *Large Feature Extent*
 - *Possibly by Data Conversion / Testing*
 - *Loading Static Raster Data Into GDB, etc.*



Avoid Pitfalls!

Expensive Rework

Prevention is Better Than Cure!

- **Hardware**
 - Older Processor
 - Lesser Capacity, etc.
- **Software**
 - Version Selection including Patches
 - Utility Industry ArcGIS Desktop Version 10.2.1 (~~10.2.2~~)
 - ArcGIS Server Version Can Be the Latest

Avoid Pitfalls!

Expensive Rework

Prevention is Better Than Cure!

- **System Integration**
 - Inefficient Interface Design
 - Version Difference for Each Version
 - Data Sharing with Other Systems Takes Hours, If not Days!
 - Synchronization Frequency
 - etc.

Avoid Pitfalls!

Performance and Scalability

Risks

- Workflow
 - Number of Outstanding Versions,
 - Versioning Levels
 - etc.
- Maintenance
 - Total Delta Table Records,
 - GDB Maintenance
 - etc.

```
Exec DBMS_STATS.GATHER_SCHEMA_STATS ('ARCFM', estimate_percent=>100,  
DEGREE=> 7, CASCADE=>TRUE, No_Invalidate=>false);  
SQL> select sum(num_rows) as Total_A_Table_Rows from dba_tables  
       where table_name in (select 'A' || registration_id from sde.table_registry);
```

TOTAL_A_TABLE_ROWS

3,813,927

```
SQL> select sum(num_rows) as Total_D_Table_Rows from dba_tables  
       where table_name in (select 'D' || registration_id from sde.table_registry);
```

TOTAL_D_TABLE_ROWS

3,560,428

Separate Myths from Evidence Based Facts!

Dependencies

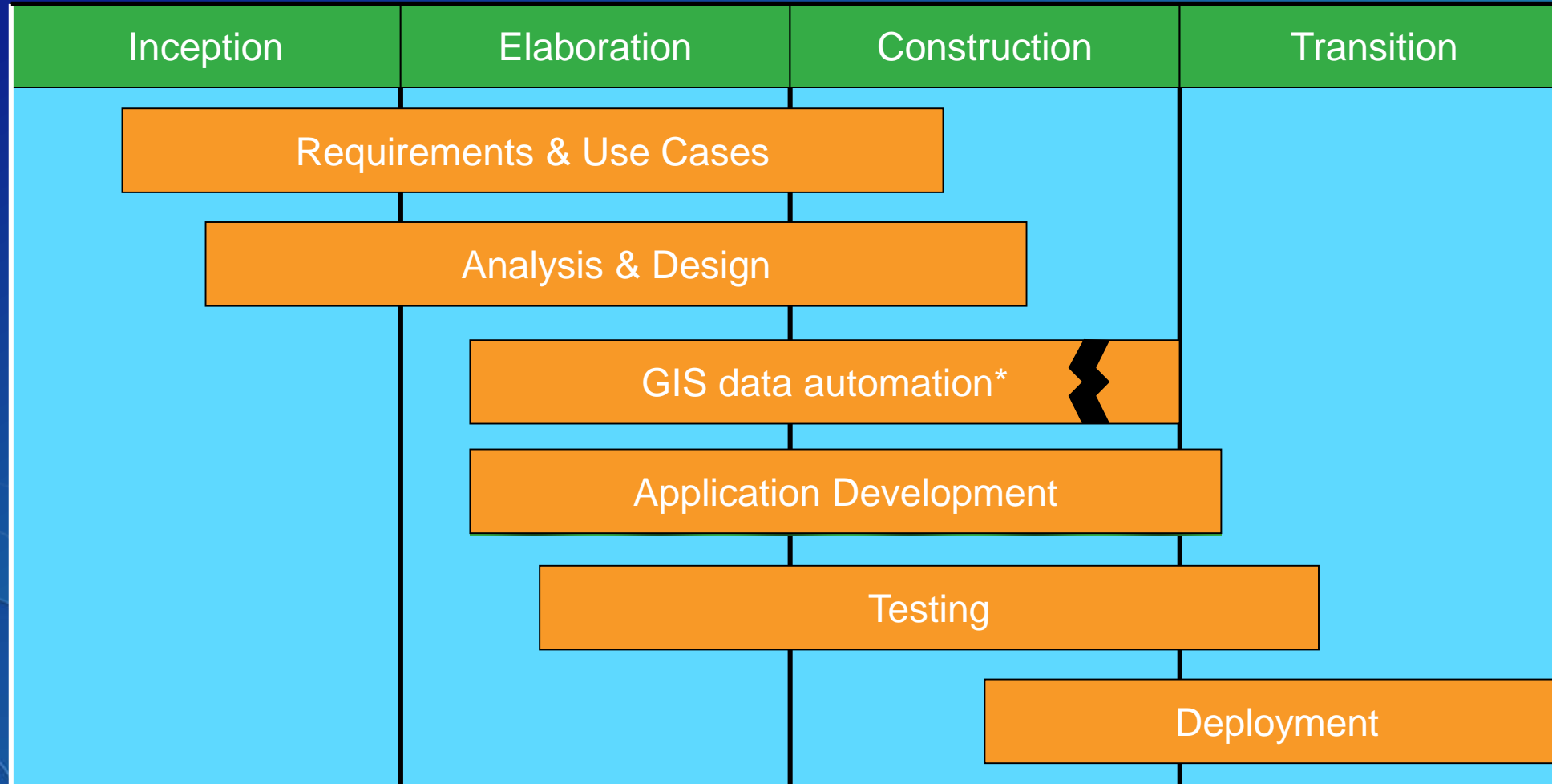
Intrigue Challenges!

- **Key Inputs**
 - User / Business Groups
 - Functional and Non-Functional Requirements
 - Application Designs
 - etc.
- **Need a Dedicated / Assigned ArcSDE Administrator**

Talk to the Right People to Get the Right Information!

Design

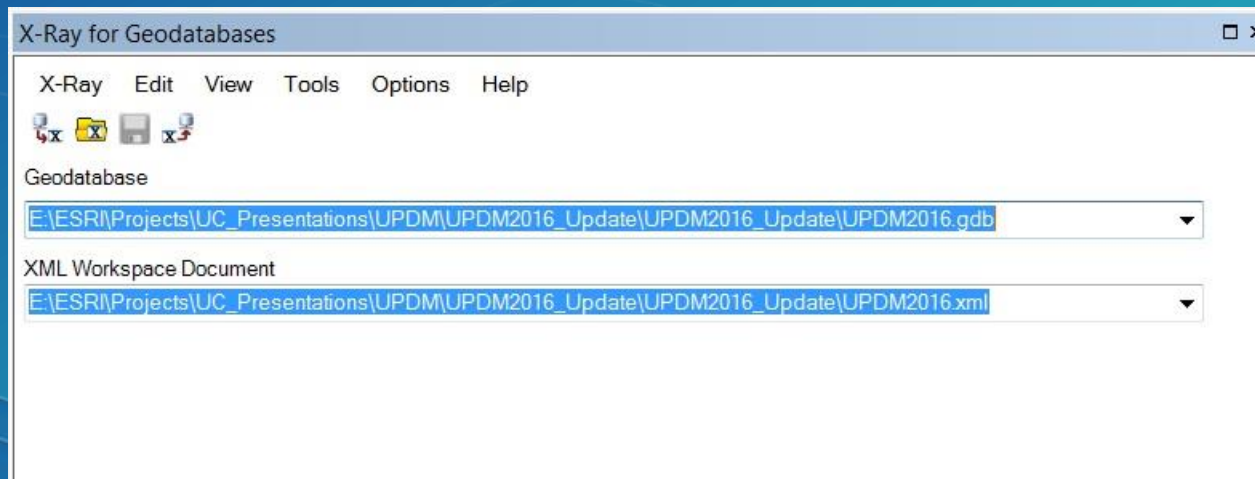
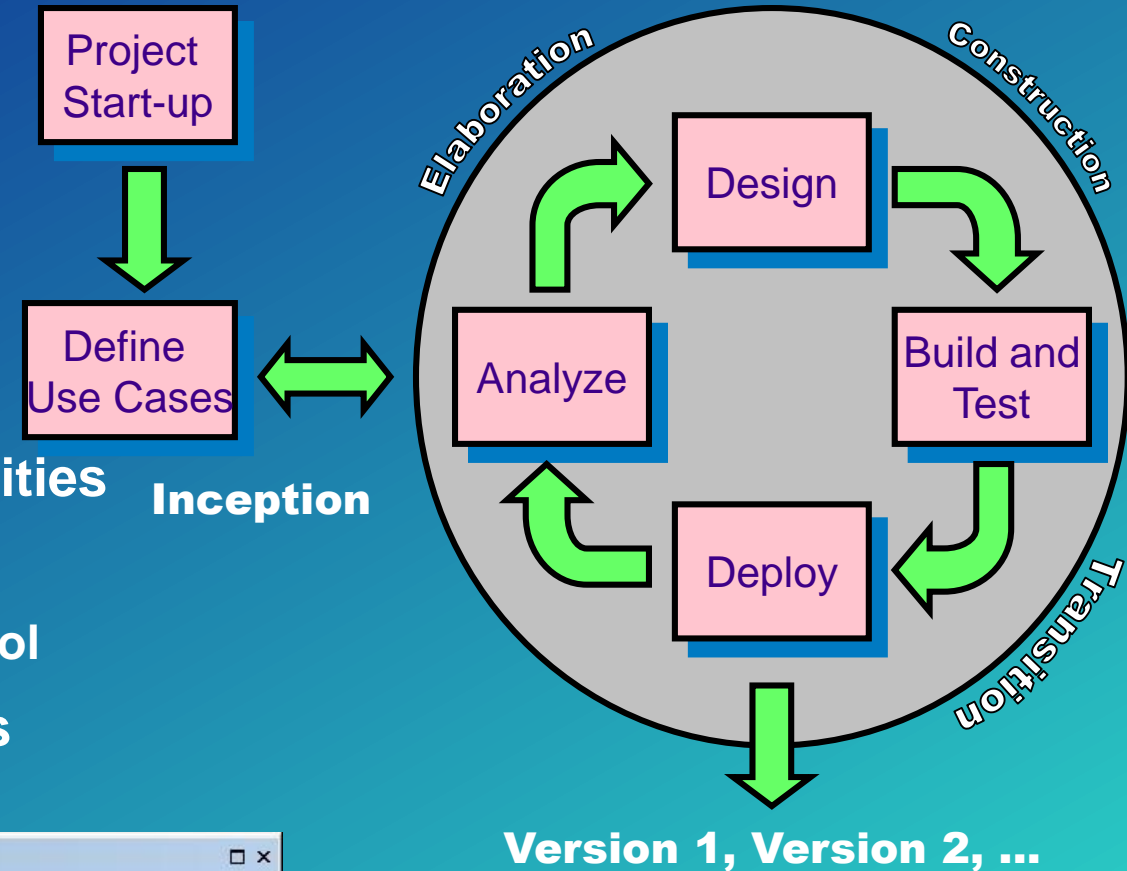
Geodatabase Design - Phases



Time – per iteration

Geodatabase Design – Process

- Identify and Characterize Each Thematic Layer
- Develop Geodatabase Elements and Properties
- Define the Data Capture Procedures and Responsibilities
- Test, Refine and Document the Design
 - X-Ray, Microsoft Visio and Geodatabase Diagrammer Tool
- Agile (Incremental) Vs Waterfall (Sequential) Methods
 - Agile Works Better!



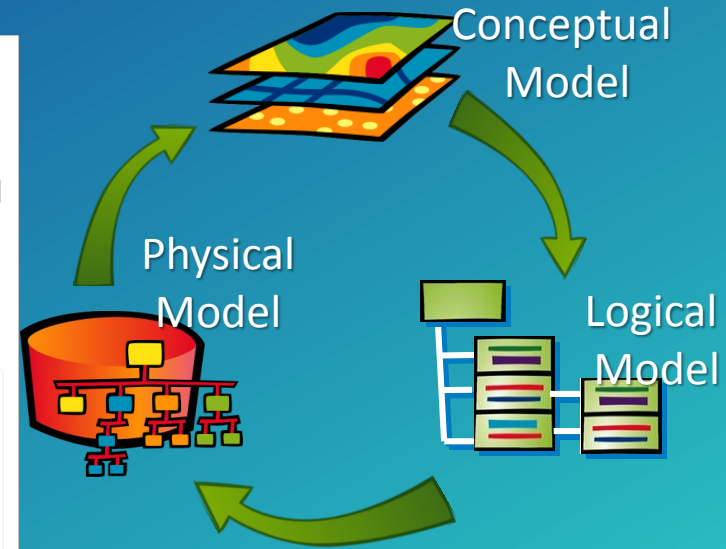
Version 1, Version 2, ...

**THE
SCIENCE
OF
WHERE**

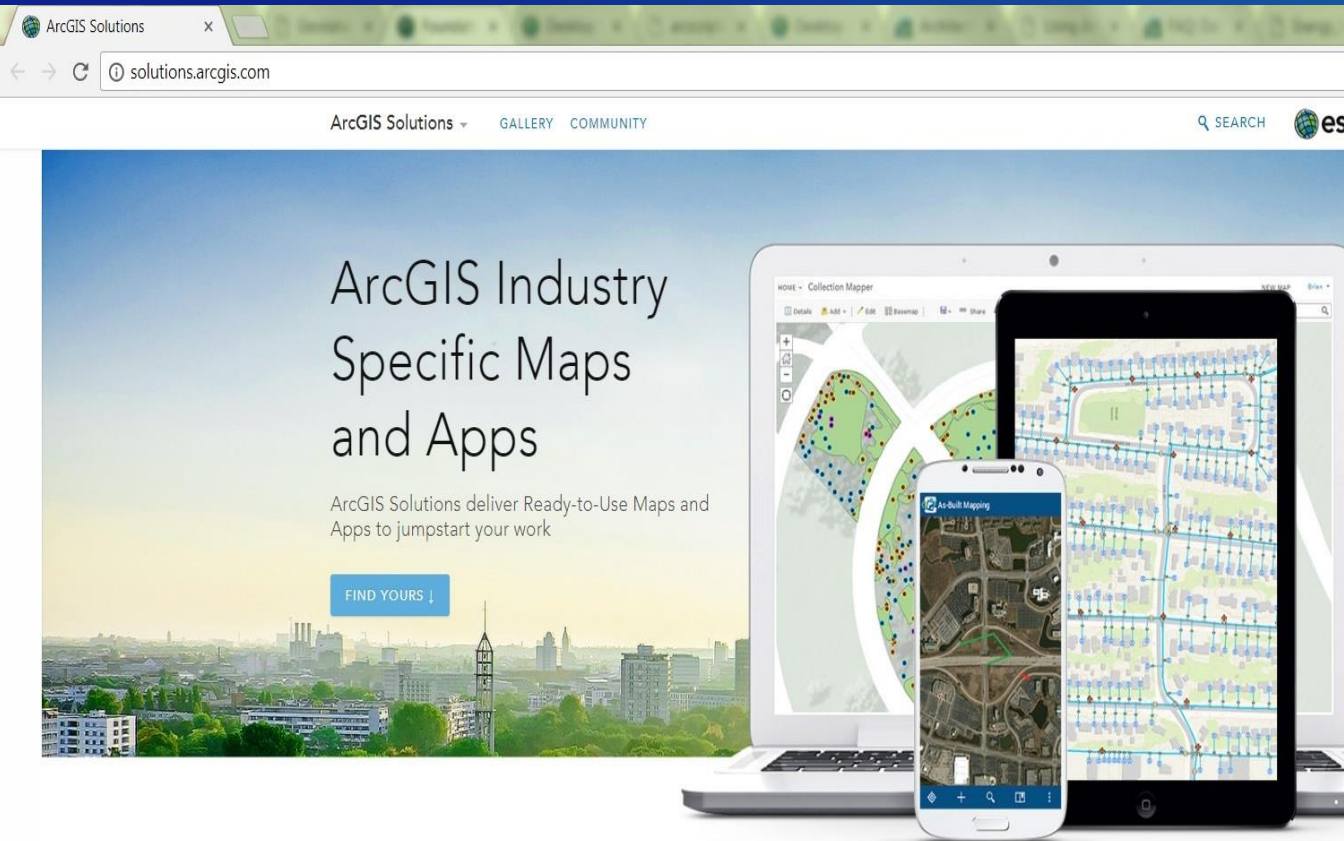
- **Identify Business Requirements**
- **Identify Thematic Layers**
- **Identify Required Applications**
- **Leverage Data Model Templates**
- **Document**

- **Define Tabular Database Structure**
- **Define Relationships**
- **Determine Spatial Properties**
- **Document**

- **Create and Implement Model Design**
- **Generate Physical Schema in the RDBMS**
- **Testing and Validation**
- **Document**

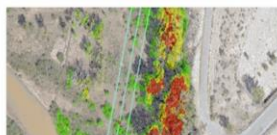


Geodatabase Design – ArcGIS Data Models



Distribution

Leverage geographic information to maintain comprehensive electric utility records, plan and conduct inspections, communicate with customers, and more.



Transmission

Organize your network assets, map critical infrastructure, collect system data in the field, and more.



Generation

Create facility maps, analyze data, provide information to your entire organization, and more.



Electric



Gas



Defense

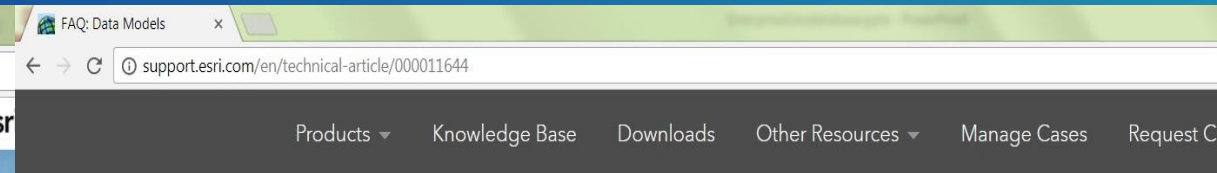


Telecommunications



Parks and Gardens

and platform configurations to meet your needs, and practices for your organization.



Download Links

Below is a list of data models that are or will be available for download.

Related Information

- [Address](#)
- [Agriculture](#)
- [Atmospheric](#)
- [Basemap](#)
- [Biodiversity](#)
- [BroadbandStat](#)
- [Building Interior Space](#)
- [Carbon Footprint](#)
- [Census - Administrative Boundaries](#)
- [Defense - Intel](#)
- [Energy Utilities \(includes ArcGIS MultiSpeak\)](#)
- [Environmental Regulated Facilities](#)

[Back to top](#)



Land Records

Manage an accurate inventory of real property; describe who owns the land; analyze land value and associated improvements; and maximize property tax revenue.



Water Utilities

Maintain comprehensive water distribution, sewer, and stormwater records; and improve the operation of utility networks that provide clean drinking water.



Public Works

Maintain right of way assets; meet transportation and environmental quality needs; coordinate and plan capital projects; and operate parks, and government facilities in a safe and effective way.

Geodatabase Design – Key Considerations

- **Good Design =**
 - Keep only Required Number of Feature Classes and Columns
 - Less Complex and Attributed Relationship Classes
 - etc.!
- **Use Feature Datasets as Needed**
 - Stand Alone Feature Classes are Fine!
- **Performance And Scalability**



Geodatabase Design – Key Considerations

Poor Design = Slow Performance and Bugs

- Empty Feature Classes / Columns Forever
- Redundant/Duplicate Columns / Domains, etc.
- Missing / In-Correct Domains and Aliases for Fields
- Column / Domain Names and Field Lengths

- >10 Characters in Field Names
- Length of Text/NCLOB – 256 or 1,073,741,822
- Choice of Field Type
 - Short Integer
 - Long Integer
 - Float
 - Double.
- Selection of Precision and Scale
- Define Not Null Fields

Field Properties		
Alias	CustomerName	
Allow NULL values	Yes	
Default Value		
Domain		
Length	1073741822	



Geodatabase Design – Key Considerations

Poor Design = Slow Performance and Bugs

- Change in XY Tolerance
 - Default = 10x Times of XY Resolution
 - Introduces Complexity (#NIM090335) for Geometric Network, etc.
 - Impacts Performance

New Feature Dataset

XY Tolerance

The XY tolerance is the minimum distance between coordinates before they are considered equal. The XY tolerance is used when evaluating relationships between features.

0.001 Meter

Z Tolerance

0.001

M Tolerance

0.001 Unknown Units

Reset To Default [About spatial reference properties](#)

☒ Accept default resolution and domain extent (recommended)

< Back Finish Cancel

Resolution

XY Resolution: 0.000064000000006 Meter

Tolerance

XY Tolerance 0.002048000001907 Meter

Geodatabase Design – Key Considerations

Poor Design = Slow Performance and Bugs



- No Attributed Relationship Classes for Empty Tables!
- Use Many to Many Relationship Classes Only When Necessary
- Don't Use Objectid as Primary Key for Relationship Classes
 - Unexpected Replication Behavior
 - Additional Processing During Synchronization

Relationship Class Properties

General Rules

Name: ELECDIST.SwiFac_DynProtectDev

Type: Simple

Cardinality: 1 - M

Notification: None (no messages propagated)

Origin Table/Feature Class

Name: ELECDIST.SwitchingFacility

Primary Key: OBJECTID

Foreign Key: StructureObjectID

Destination Table/Feature Class

Name: ELECDIST.DynamicProtectiveDevice

Primary Key: OBJECTID

Foreign Key: DeviceObjectID

Labels

Forward: Dynamic Protective Device

Backward: Switching Facility

OK Cancel Apply

Planning an Enterprise Geodatabase

Geodatabase Design – Best Practices

- **Leverage the Existing ArcGIS Data Models**
 - **Drop the Un-Necessary Feature Classes, Columns, etc. (Normalization)**
 - **Possibly Split the Feature Classes Pertaining to Different Scale Levels**
 - **Test, Refine and Tune the Data Models**

Justify Every Single Geodatabase Element!

Geodatabase Design – Best Practices

- **Select Single Coordinate System for the Geodatabase**
 - On the Fly Projection is Expensive
 - No Support for Geometric Network Editing
- **Create Feature Datasets or Databases for Different LOB (Line of Businesses)**
 - Depends on Size, Access, Usage and Maintenance
 - Larger Companies Generally Require a Separate Landbase GDB

Geodatabase Design – Recommendations

- Review the Labeling Requirements Ahead of Time
 - For Multi-Field Complex Labeling
 - Add/Calculate a New Field and Auto Update
 - Convert Labels to Annotations
- Analyze Requirements and Choose Proper Annotation Reference Scale
- Reduce Number of Annotation Classes within an Annotation Feature Class
- Add Attribute Indexes for Where Clauses
 - Label Expression
 - Definition Queries
 - Application Design



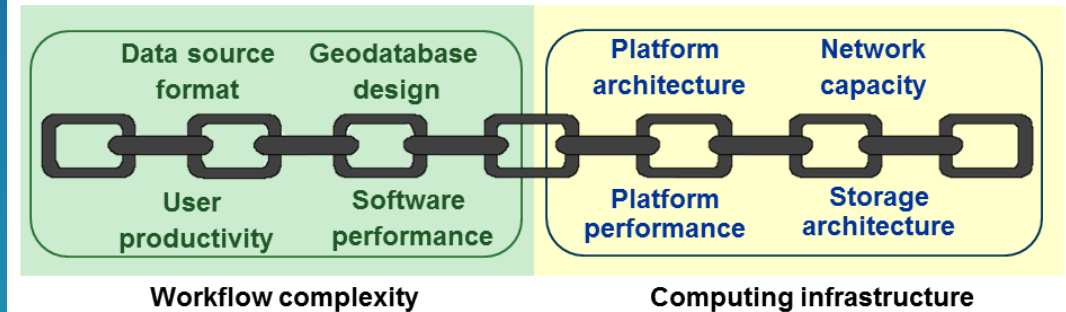
Architecture

Architecture and Capacity Planning

- Are actionable requirements available?
- Is the technology appropriate?
- Is it available enough?
- Is it continuous enough?
- Is there enough capacity?
- Is it meeting performance SLAs?

Why is system architecture design important?

- Balanced system architecture design can:
 - Reduce costs
 - Improve user productivity



System architecture design provides framework for productive operations

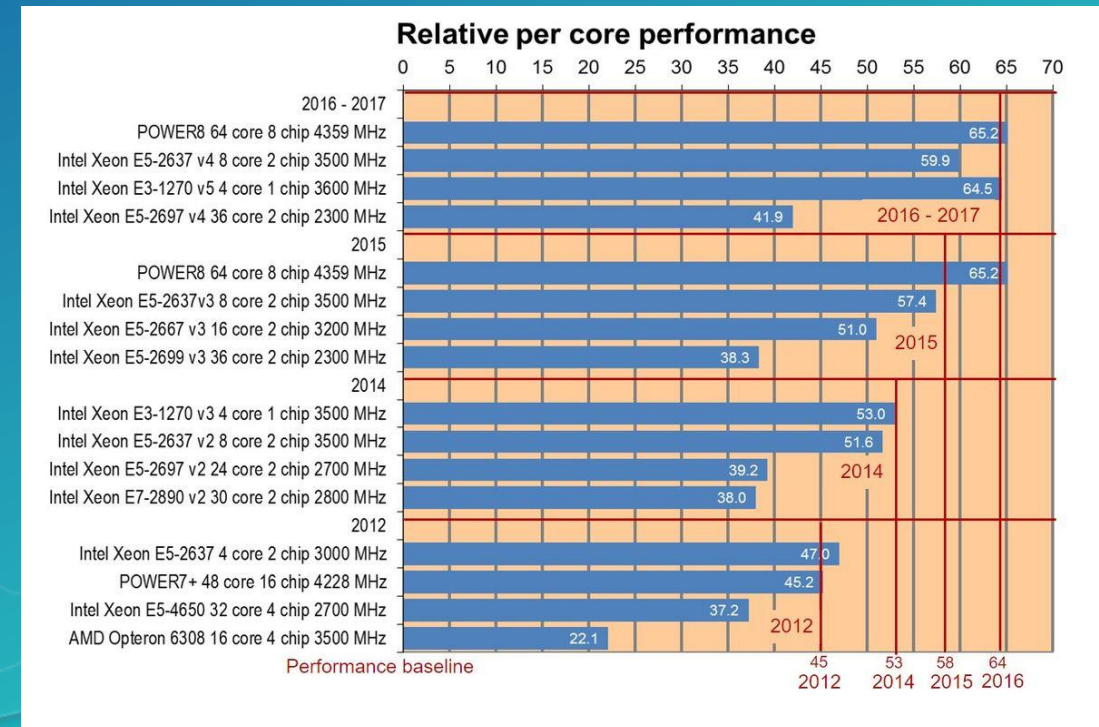
The geodatabase design and associated system architecture matters!

Server Technology Selection

Why is it Important?

- Key to Optimal Scalability and Performance
- Save costs by Reducing Server Footprint
- Keep up with Estimated / Expected Delta Records Count
- RDBMS Needs Processing Power for the Versioned Queries
 - Versioned Query =
(Base Table – D# Table) + (A# Table – D# Table)

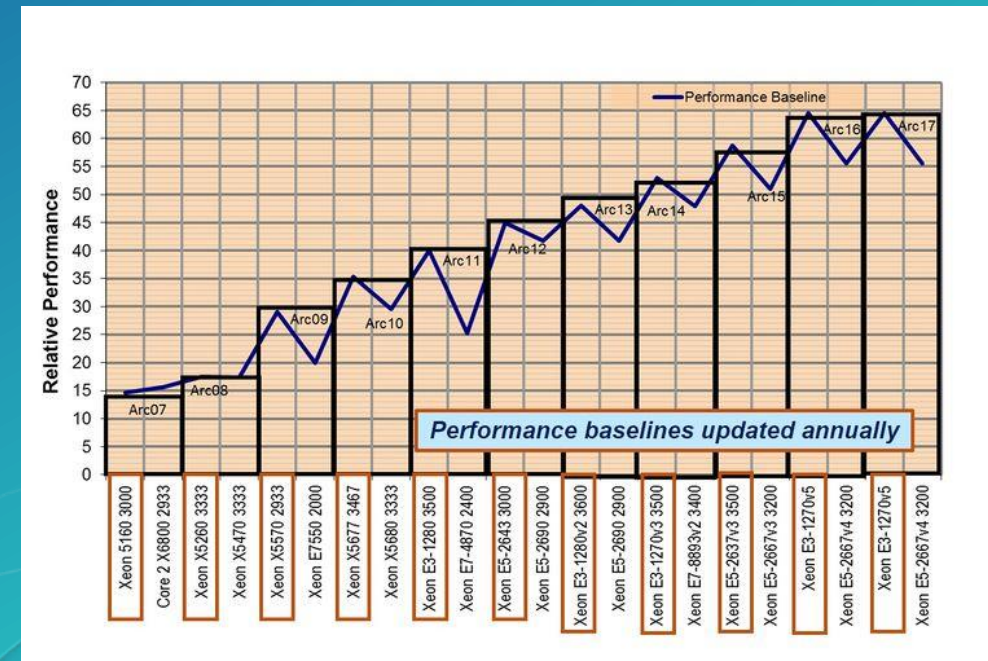
D# Table Records are Sorted Twice!



What's the Latest Processor?

- Intel® Xeon® processor E7-8800/4800 v4 family
 - Up to 3.69x Performance Gains Vs Previous Generation
 - SPEC Rate / Core ~ 59.5
 - Offering up to 24 cores and 48 threads per Processor
- For PC, 18 cores and 36 threads - Intel® Core™ X-series Processor!

Look for SPEC Rate Per Core Value!



Define Functional and Non-Functional Requirements

Critical!

- **High Availability**
 - Use of Web Application by 24x7 Field / Emergency Crew
- **Backups**
 - RPO (Recovery Point Objective)
- **Disaster Recovery**
 - RTO (Recovery Time Objective)
- **SLAs (Service Level Agreements)**



Supported Database Management Systems – 10.5.x

Database Solution	Geodatabase Support	Feature Service Support	Query Layer Support
Altibase	No	Yes	Yes
Dameng	No	Yes	Yes
IBM DB2	Yes	Yes (not z/OS)	Yes
IBM Informix	Yes	Yes	Yes
Microsoft SQL Server	Yes	Yes	Yes
Microsoft Azure SQL Database	No	Yes	Yes
Netezza Data Warehouse Appliance	No	No	Yes
Oracle	Yes	Yes	Yes
PostgreSQL	Yes	Yes	Yes
SAP HANA	No	Yes	Yes
SQLite	No	No	No
Teradata Data Warehouse Appliance	No	Yes	Yes

Architecture and Design – Key Considerations & Challenges

- Number of Users, Operations and Transactions
 - Expected 130 Power Users (Editors) Vs 200+ Actual Users
 - Impact : >Concurrency Adds Pressure to Server Resources
- Number of Outstanding Versions
 - Estimated 500 Vs Actual 1250 Versions
 - Impact: Increased RDBMS Server's CPU Time
- Estimated Total Delta Table Records (A# and D# Tables)
 - Expected 2 Millions Vs 8 Millions Actual
 - Impact: > CPU & Memory

```
DELETED_AT IN (SELECT l.lineage_id FROM SDE.state_lineages l WHERE  
l.lineage_name = :lineage_name2 AND l.lineage_id <= :state_id2) AND  
SDE.STATE_ID > : "SYS_B_5") AND a.SDE_STATE_ID = SL.lineage_id AND  
SL.lineage_name = :lineage_name3 AND SL.lineage_id <= :state_id3) V__138
```

call	count	cpu	elapsed	disk	query	current	rows
Parse	1	0.00	0.00	0	0	0	0
Execute	10	0.08	0.13	0	2289	0	0
Fetch	10	8.56	15.53	58	752569	20	436
total	21	8.65	15.67	58	754858	20	436

Virtualized Database Servers – Key Considerations

IT Standard Now!

- Avoid Over-Commitment
- Ensure Less Number of vMotions
- Estimated Processing Requirements
 - User Load
 - Dedicated Operations and Transactions
- Application & Database Complexity



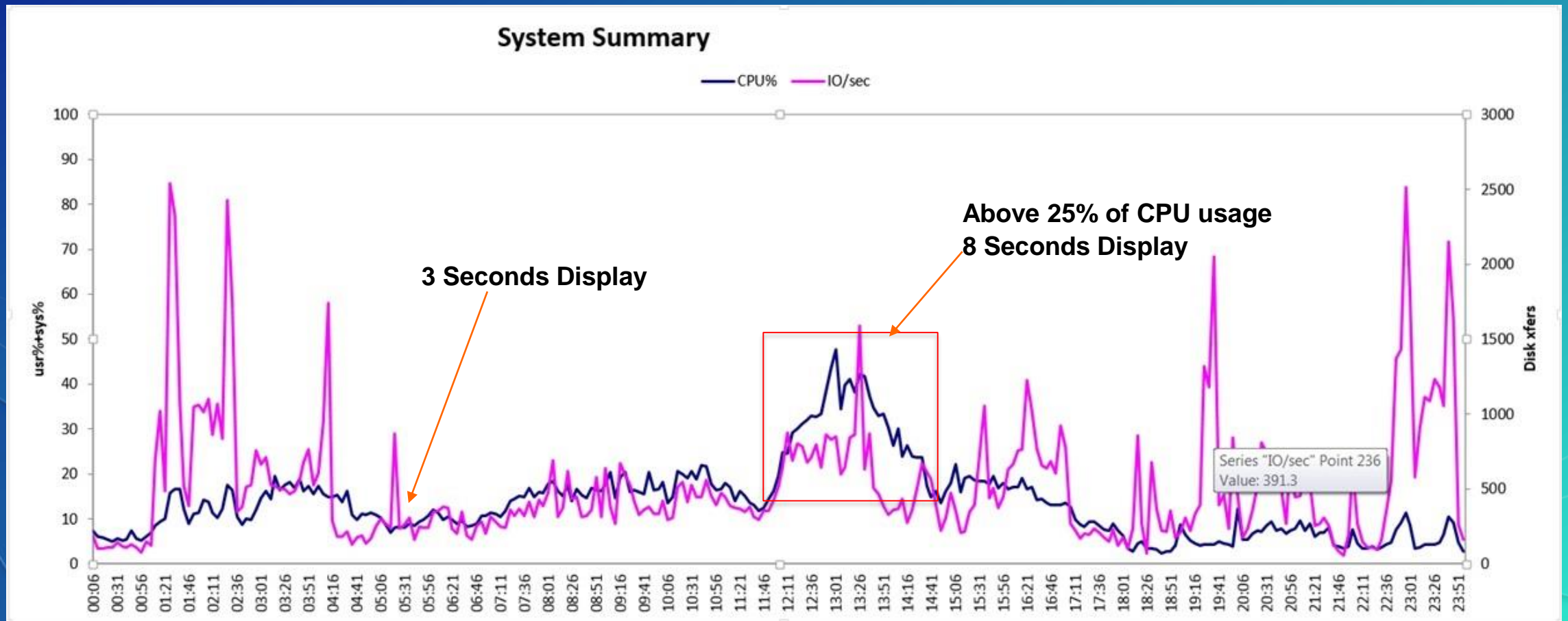
Physical vs. Virtual Server

Processor Type:	Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz
Logical Processors:	40
NICs:	4
Virtual Machines:	62

Virtual Machines and Templates:	106
Total Migrations using vMotion:	6644

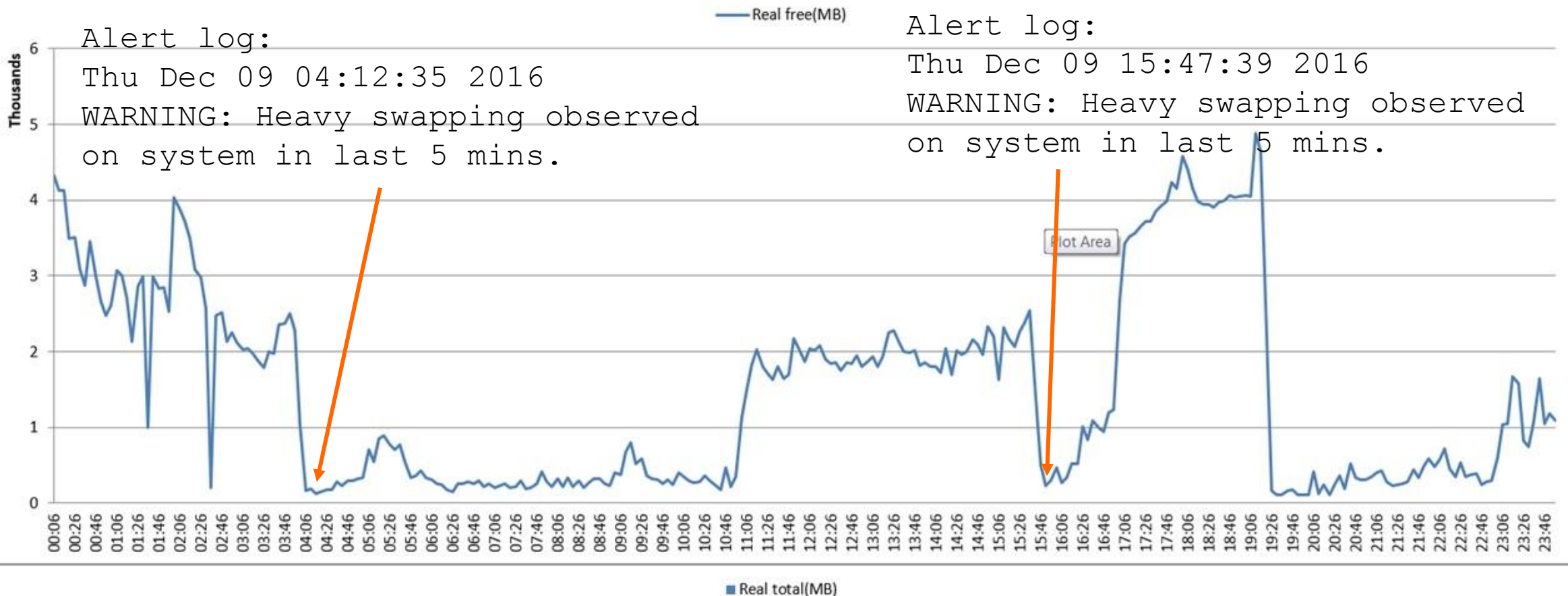
Symptoms of Over-Committed vCPUs

- vCPU Usage Vs MXD Display - Mxdperfstat Tool Results



Additional Memory Requirements

Memory



Network Planning

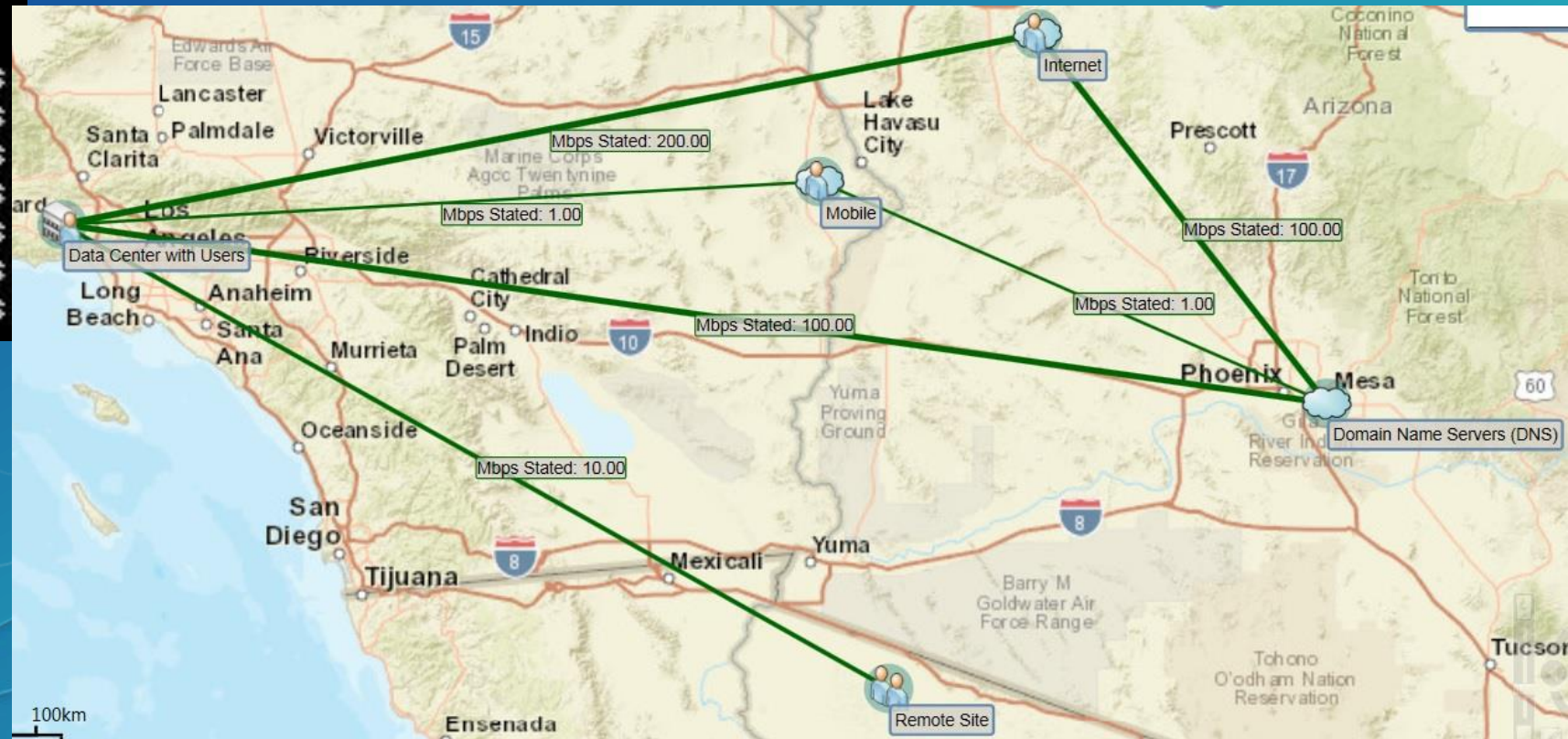
Establish and Configure DNS Appropriately!

```
C:\Users\ >tracert

Tracing route to
over a maximum of 30 hops:

  1    55 ms    55 ms    55 ms
  2    55 ms    55 ms    55 ms
  3   115 ms    58 ms    62 ms
  4   111 ms   111 ms   112 ms
  5   110 ms   109 ms   110 ms
  6   110 ms   113 ms   110 ms
  7   109 ms   109 ms   109 ms
```

Trace Route: LA Workstation → Phoenix DNS
LA Database Server ←



System Architecture – Recommendations

Vendor Selection – RDBMS and Virtual Environments

- **Select the Supported Vendors Based on:**
 - Business Requirements
 - Existing Infrastructure Setup
 - Available Skill Set
 - Comparative Functional and Non-Functional Test Results
 - Develop a Test Plan Based on **GIS Workflows**
 - Prepare Isolated Environments Separately for Technology Comparison
 - Identify and Present Pros and Cons per the Requirements
- **Esri Does Not Recommend any One Particular Vendor Solution**
 - All Supported Vendors are Business Partners to Esri

Esri Professional Service Can Help Establishing Benchmarks, Pilot, etc.!

System Architecture – Recommendations

Virtual Environment - Hardware

- **Deploy a Dedicated Virtual Environment for a Large User Base**
 - Provide a Decent Provisioning Ratio
 - Physical CPU vs vCPU
- **Adjust the Configurations and Tune Virtual environment Before Deployment**
- **Provide GPU (Graphics Processing Unit)**
 - Video RAM >256MB Per Virtual Machine
- **Fit Virtual Machine within one CPU NUMA* Node**
 - # of vCPUs is less than or equal to the number of cores in the CPU socket
- **Deploy Better Processors - Spec Rate Per Core 55+**

*NUMA - Non-uniform memory access

System Architecture – Recommendations

Software

- **Upgrade the Software to the Latest Version**
 - Up to the current Service Pack (SP) levels and Patch Updates
 - Apply the latest Utility and Telecom Update (UTU) Patch 7 for 10.2.1

System Architecture – Recommendations

Network Infrastructure

- **Request for Higher Network Bandwidth (~ 1Gbps) and Reduced Latency (<1 MS)**
 - >Bandwidth Reduces Number of Network Packets
 - ArcGIS Desktop is Sensitive to 1 – 2 MS Latency!
 - Plan for ~ 1.5 Mbps per Concurrent GIS User
- **Enable Jumbo Frames Between Servers**
 - Ensure All Switches Support – Otherwise Don't Enable it!
- **Validate Network Path Between GIS User and Server Locations**
 - Correct DNS and Routing
- **Upgrade Lower Bandwidth or Move GIS User Locations**

System Architecture – Recommendations

Disks

- Use SSDs (Solid-State Drives)
- Plan for 3000 – 5000 IOPS (Input/output Operations Per Second)
- Avoid LUNs* > 2TB Size
 - Minimum of 4 LUNs that are Identical in Size
- Avoid noac Mount Option!

*LUN – Logical Unit Number

System Architecture – Recommendations

Operating System

- **Adjust and Configure**
 - Kernel Parameters
 - Settings Specific to RDBMS and Network Capacity
- **Enable Large / Huge Memory Pages for Geodatabases**
- **Update Patches**

System Architecture – Recommendations

ArcSDE Configuration

- **Configure ArcSDE DBTUNE Settings**
 - Use Default Geometry Storage - St_Geometry / Geometry
 - Storage Locations
 - Etc.
- **ArcSDE Initialization Parameters**
 - Defaults are Good!

System Architecture – Recommendations

RDBMS Configuration and Tuning

- Review Single Instance Vs Multiple Instances and Databases
- Tune RDBMS
 - Memory Allocation from Hardware
 - Other Initialization Parameters
 - Log File Settings
 - etc.
- Implement the Best Practices

System Architecture – Recommendations

System Integration

- **Use Private Versions for Larger Versioning Environment**
 - **Keep the Total Number Less!**
 - **Sync Frequently**
 - **Monitor**
- **Complete the Batch Processes within the Identified Time Frame**
 - **Improve Hardware and/or Software Design**
 - **Additional CPUs**
 - **Multi Threads**
 - **Schema Cache**

Build

Build Geodatabase

- **Create Physical Geodatabase**
 - Structure the Implementation to Pilot → Phase I → Phase II → Phase III, etc.
 - Enough Gap Between Each Phases to Accommodate the Learned Lessons
 - Separate Data Owner from SDE / DBO User
- **Develop Data Conversion/Update Specification Document Aligned with Data Model**
 - Test, Refine and Tune Data Model
- **Team Review and Demonstration**
 - Show How Tasks are Performed Using GIS
 - Show Maps, Reports, Online Demos

Build Geodatabase – Key Considerations

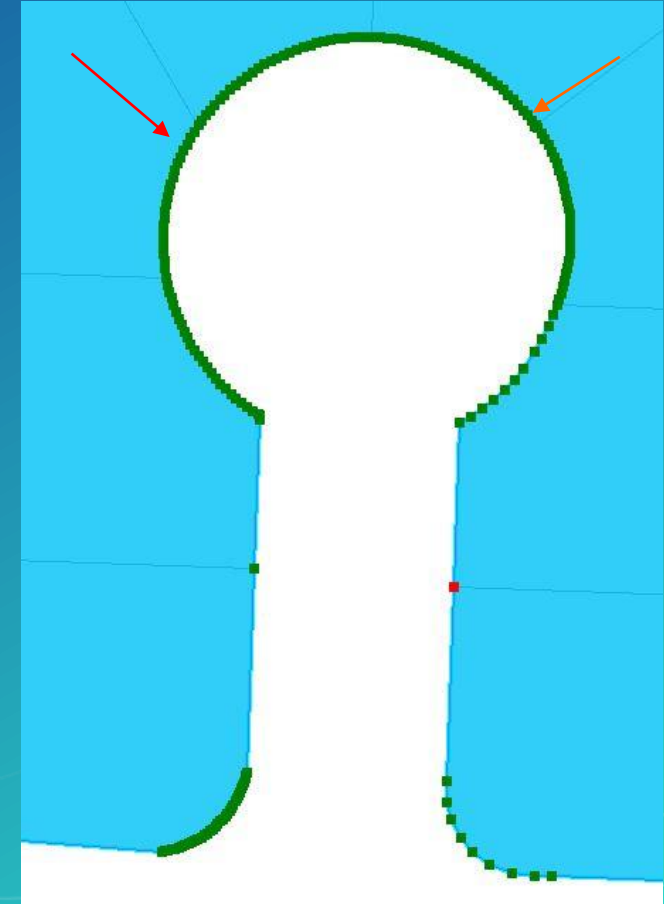
Data is the Backbone of GIS!

- Aim for 100% Data Accuracy from Data Conversion Effort
- Extra Vertices have Performance Impact
- Either Populate or Drop Empty Fields
- Minimize Data Model / Schema Changes in a Versioned Geodatabase
- Avoid Creating Separate GDBs for GIS Users from Different Geographic Locations

Key Data Conversion Considerations



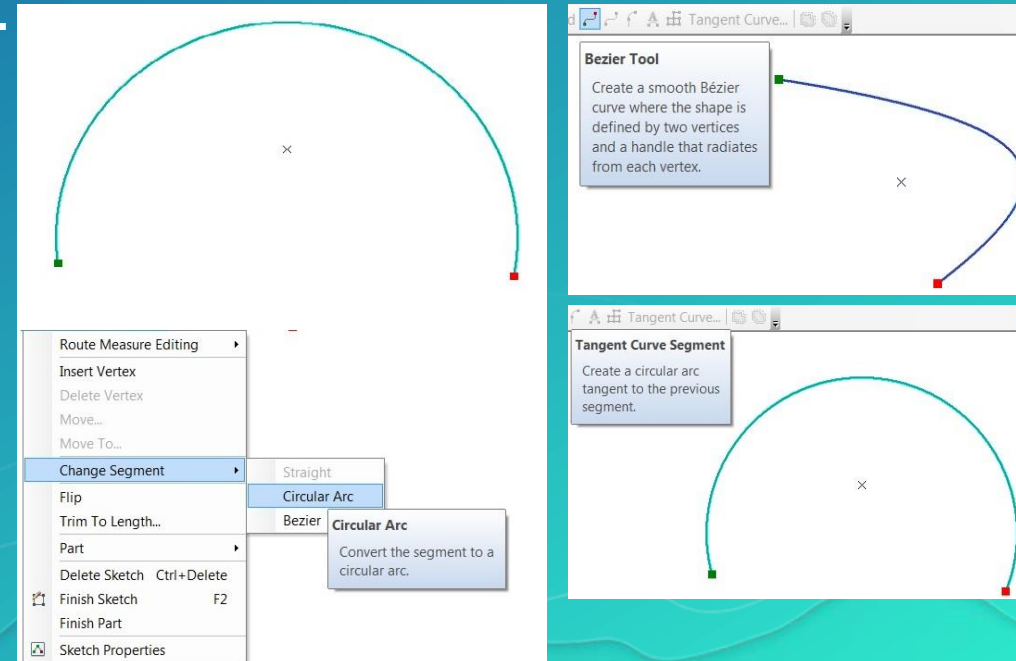
- **Extra Vertices Introduced by:**
 - Conversion process involving CAD systems
 - Geometric Network Creation with Snapping ON Option
 - Conversion/Update Methods
- **Develop Adequate QA/QC Methods and Procedures**
- **Additional Data Reviewer Checks**
 - Duplicate/Invalid Geometries
 - Orphan Related Records
 - Connectivity Check
 - Etc.



Feature Class/Layer	Total Number of Features	Total Number of Vertices with Geometric Network Snapping	Total Number of Vertices without Geometric Network Snapping	Increase in number of vertices %
Secondary Overhead Conductor	3,712	12,953	7,841	65.20%

Build Geodatabase – Recommendations

- **Avoid Extra vertices**
 - Use Curve Tools that Insert Less Vertices
- **Turn off Snapping During Geometric Network (GN) Creation**
 - ArcGIS 10.0 onwards a Vertex is added at every Intersection to Improve the Scalability
 - Cannot Create GN with >15M edges until ArcGIS 9.3.1.
- **Remove Additional Vertices**
 - Generalize / Simplify
 - ArcObjects



Build Geodatabase – Recommendations

- Keep the Data Clean and Simple
 - Without Any Topological Errors
- No Coincident Complex Edge Features in Geometric Network
 - Most Common Reason for Geometric Network Corruption
- Unversion Read Only Feature Classes / Tables
- Use Mosaic Datasets Instead of Loading Raster Data Into EGDB
 - Static Raster Data Does Not Need to Participate in Daily RDBMS Backup

OWNER	TABLE_NAME	NUM_ROWS	BLOCKS	AVG_ROW_LEN	TO_CHAR (LAST_ANALYZED, 'MON/DD/YYHH24:MI:SS')
ARCFM	SDE_BLK_3	1145079	1190598	12	APR/19/16 05:38:10
ARCFM	BK SDE LOGFILE DATA	4231223	8773	10	APR/19/16 05:32:14

Build Geodatabase – Geometric Network Lessons Learned

- **Steps to Create Geometric Network, if Snapping Option need to be “ON”**
 - Drop the Geometric Network
 - Re-create the Geometric Network with only Required Feature Classes.
 - Turn on the snapping during the Geometric Network Building Process.
 - Drop the Geometric Network.
 - Remove the Intersection Vertices introduced by the above snapping.
 - Use ArcObjects / Python
 - Build the Geometric Network without snapping.
 - Verify the Geometric Network BuildERR table.
 - Ensure to correct all the errors/invalid features mentioned in the BuildErr table.

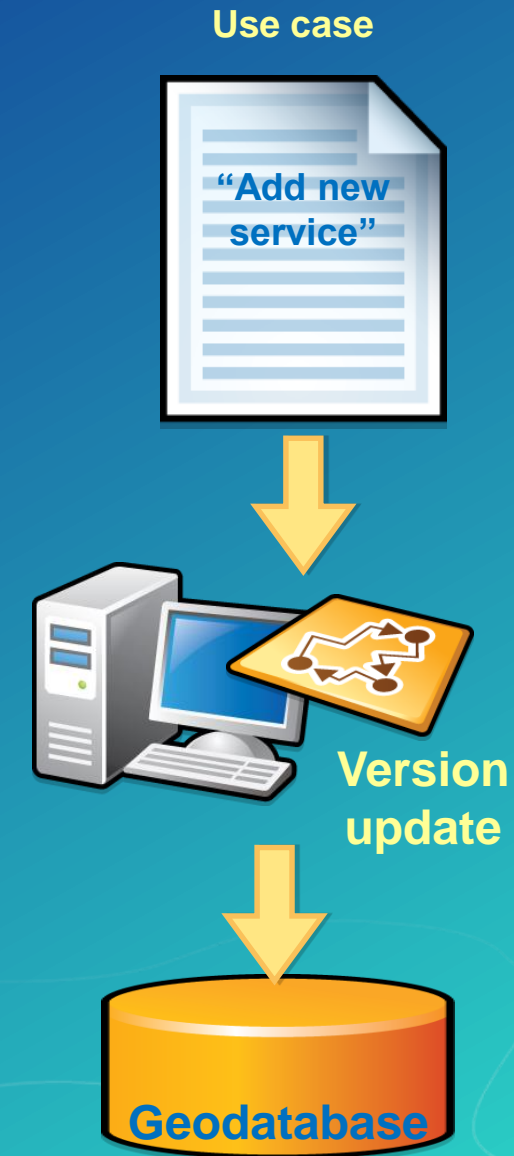
Workflows

Geodatabase Workflows

- **Plan for the Type of Workflows Per User Group**
 - Read Only Users
 - Engineering Editors – Less Editing
 - Power Users – Heavy Editing
 - Administrators
 - Web Users
 - Disconnected Mobile Users
 - Cache/Batch Process User – Generally Called “Headless” User / “Service Accounts”
 - System Integration Users
- **Test Application Workflows**
 - Functionality
 - Performance
 - Flexibility and Consistency

Geodatabase Workflows

- **Document with Use Cases**
 - A description of the task you need to perform:
 - “Add new parcel”, “Update new asset”
- **Evaluate business needs:**
 - What data needs to be edited and in what order
 - Tracking of data changes
 - Conflict detection and resolution
- **Security – User roles, etc.**
- **QA/QC steps – Enforced through Application or Database**



Geodatabase Workflows - QA / QC

- Design and Implement QA / QC Workflows
 - Data Requirements for Software Functions
 - Accurate Data for Business
 - Maintain Data Integrity

Capture, Load and Maintain Data Accurately!

Data Integrity and Validation Strategies

- **Stage 1: Don't Allow Start Editing**
 - Read Only Users
 - Without Landbase Layers
- **Stage 2: No Inserts Without Pre-Requisite Checks**
 - Out Side of Editing Areas (Pacific Ocean!)
 - Street Light without Poles
 - Equipment Without Structures
 - Required WO Number, Number of Phases, etc., in Attribute Columns
 - Etc.

Capture, Load and Maintain Data Accurately!

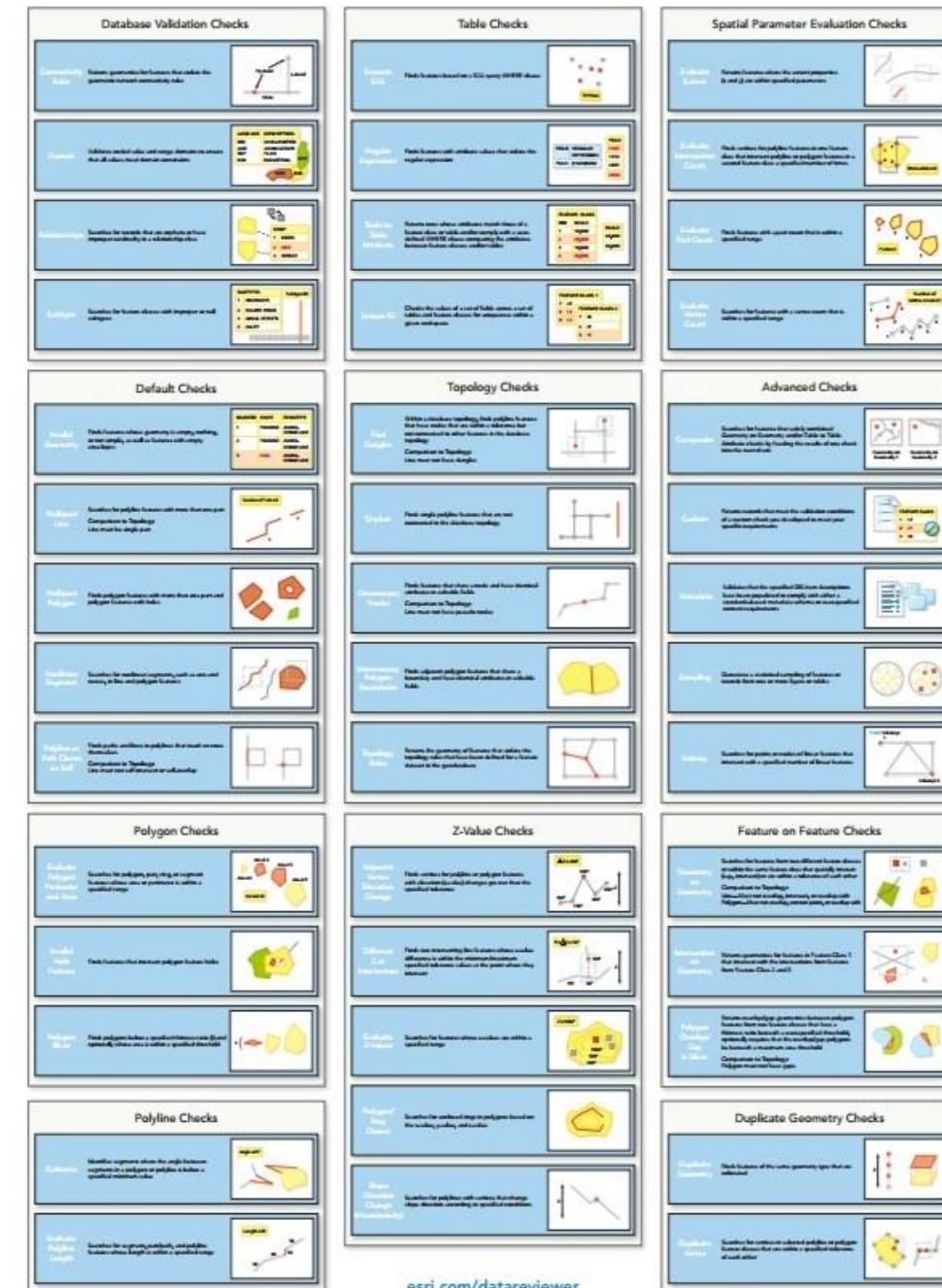
Continue....

Data Integrity and Validation Strategies

- **Stage 3: Reconcile/Save Edits Only After Rules Validation**
 - Domain Checks
 - Connectivity Rules, etc.
- **Stage 4: Allow to Post Data with Warnings**
 - Run Batch Processes to Perform Additional Checks
- **Tools**
 - Domains, Subtypes, Topology, etc.
 - Attribute Assistant Add-In
 - ArcGIS Data Reviewer
 - ArcGIS Workflow Manager
 - Business Partner Products
 - Customization

Capture, Load and Maintain Data Accurately!

ArcGIS Data Reviewer Checks



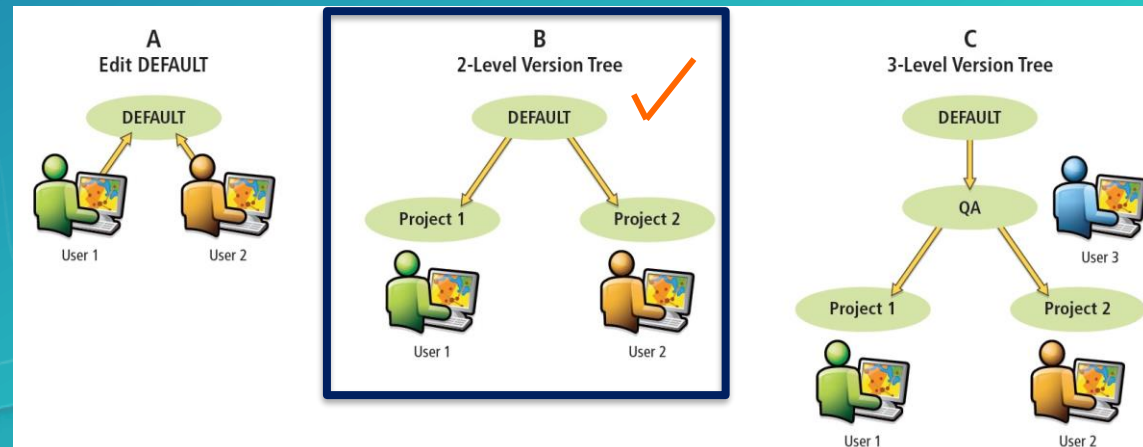
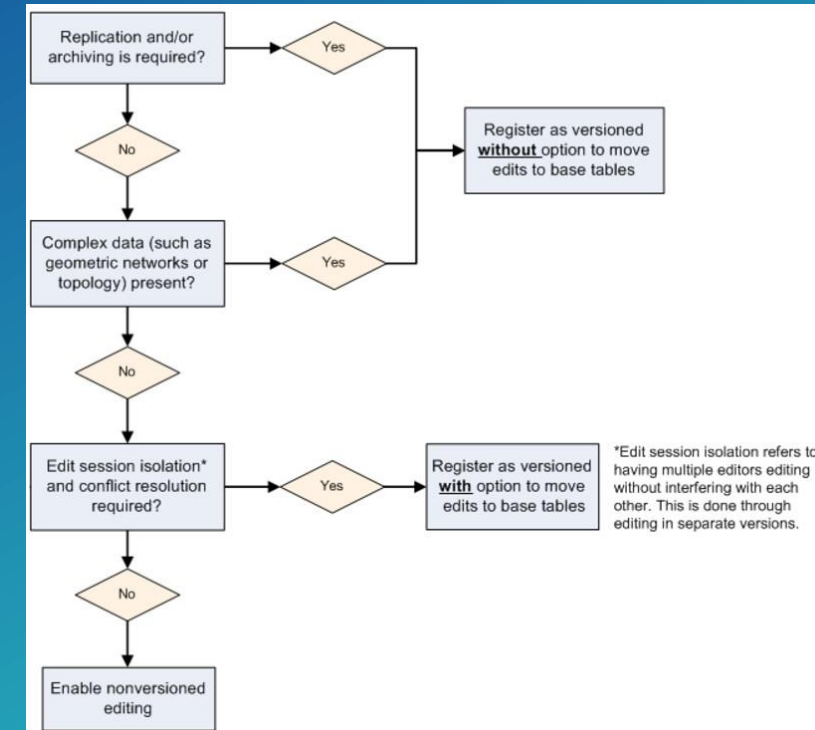
Geodatabase Multiuser Workflows

- Versioning Structure

- Delete and Recreate the Version After Each Post for 3 Levels
- **Problem: Unexpected conflict observed during reconcile**
- <http://support.esri.com/en/technical-article/000012321>
- Move Edits to Base For Simple Feature Classes
 - Ability to Easily Share the data with third-party applications

- Estimated Edit Volumes, Version Durations

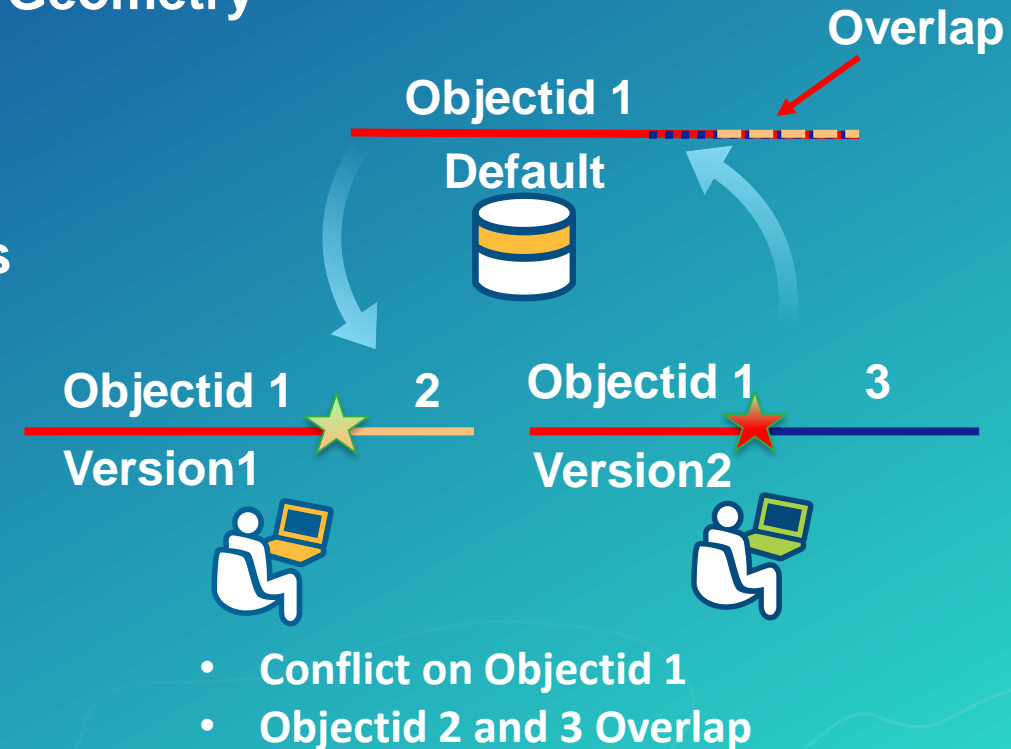
- Conflict Resolution Mechanisms



Geodatabase Workflows – Key Considerations



- Conflict for Split Operation Could Introduce Duplicate Geometry
- Plan Bulk Loading / Mass Update
- For Regular Data Load, Consider Truncate Vs Delete
- Execute Batch Processes During Non-Business Hours
- Geodatabase Replication - One-Way Vs Two-Way
- Archiving
- Editor Tracking

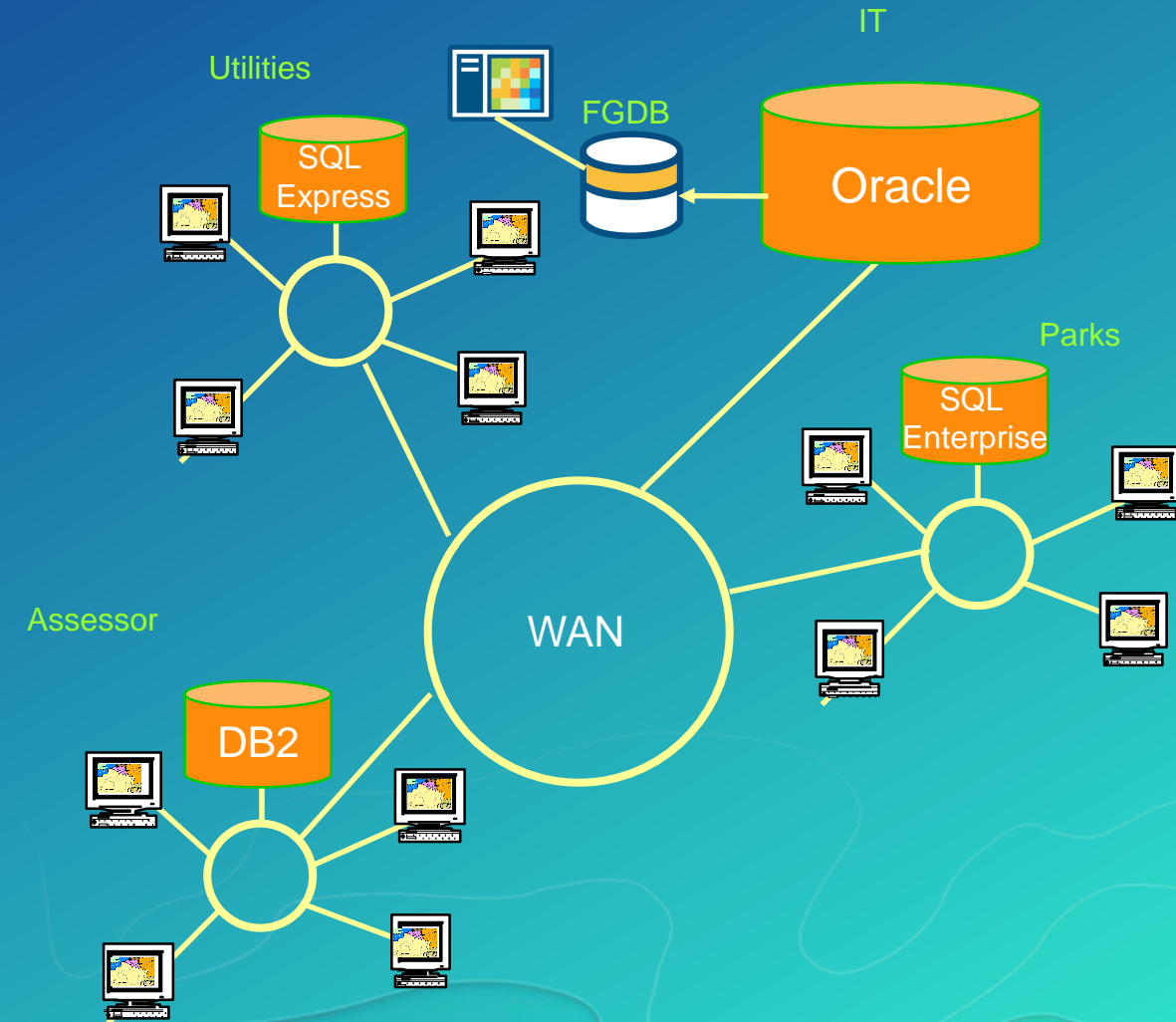


- Conflict on Objectid 1
- Objectid 2 and 3 Overlap

Geodatabase Workflows – Key Considerations & Challenges

Geodatabase Replication



- **Distributed Databases**
 - Data Sharing and Synchronization
 - Publication GDB for Web Application
 - Mobile Operations
- **Estimate Number of Replicas**
 - ~5 Replicas (Est) Vs 15 Replicas (Actual)
 - Impact: > Delta Table Records



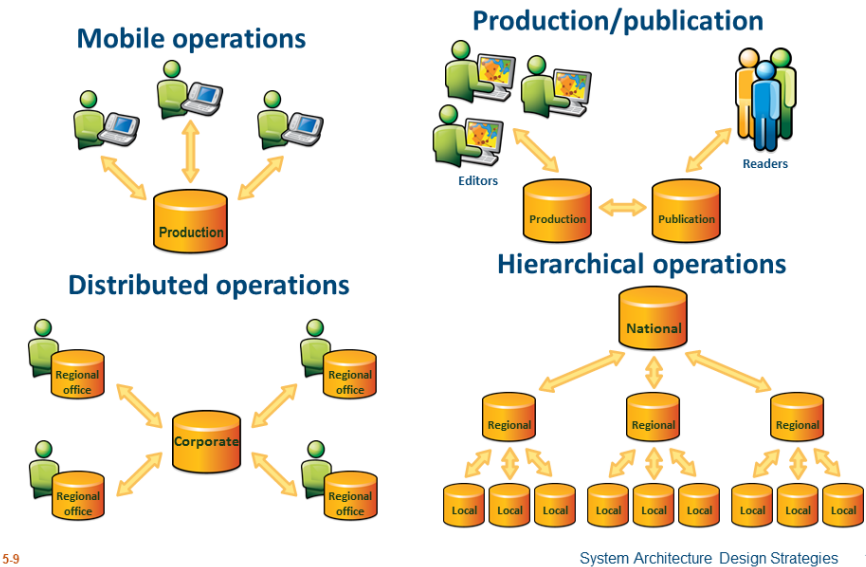
Geodatabase Workflows – Key Considerations

- **Archiving**
 - **Do Not Enable When 100% Data Update / Modification is Expected**

Geodatabase Workflows – Recommendations

- Choose the Versioning Workflows Appropriately
 - Key to Performance and Scalability
- Run QA / QC Tools Regularly
 - Avoid Duplicate / Invalid Geometries, etc.
- Leverage Geodatabase Replication
 - Much Improved Stability and Reliability Now!
 - Use One-Way Replication Options
 - Parent to Child 
 - Child to Parent 
 - Two One-Way Replicas For Two Separate Datasets/FCs

Geodatabase replication use cases



System Architecture Design Strategies 10

Geodatabase Workflows – Recommendations

- Provide only the Required Privileges to Users
 - Access to Large Number of Tables Slows Connection Performance
- Arrange Workflow Training for Users
 - Conduct Tips and Tricks Session

```
10568 [W 50:00.764] Long:      1
10569 [W 50:00.764] Long:      1
10570 [R 50:00.764] Long:      0
10571 =====
10572 [W 50:00.764] Command:    TableListTables
10573 [W 50:00.764] Long:       4
10574 [R 50:48.765] Long:       0
10575 [R 50:48.765] Long:      7835
10576 [R 50:48.765] Dynamic_Str: "SDE.MV_LPA_BAK"
10577 [R 50:48.765] Long:       1
```

Maintenance

Geodatabase Maintenance - Strategies



- Identify a Maintenance Window
- Recognize the Tasks
- Categorize
 - Nightly, Weekly, Monthly and Yearly.
- Classify Manual and Automated Batch Processes
 - Design Scalable Batch Processes
- Assign SDE/GIS Administrator Role
- Monitor



Geodatabase Needs Maintenance – Plan One!

Planning an Enterprise Geodatabase

Geodatabase Maintenance – Key Considerations

- **Maintain the GDB Performance**
 - Reconcile, Post and Compress
- **Underlying RDBMS Does Require Maintenance Other Than Backup!**
 - Rebuild Index
 - Update Statistics
 - Logs
- **Execute Automated Processes Only Within Maintenance Window**
- **Run Repair Version Tables and Metadata (Previously: SDEGDBREPAIR) Every ~3 Months**
 - Fix Any Inconsistencies
 - Schedule the Execution Around Weekends

Automate the Daily Maintenance Process!

Geodatabase Maintenance – Best Practices

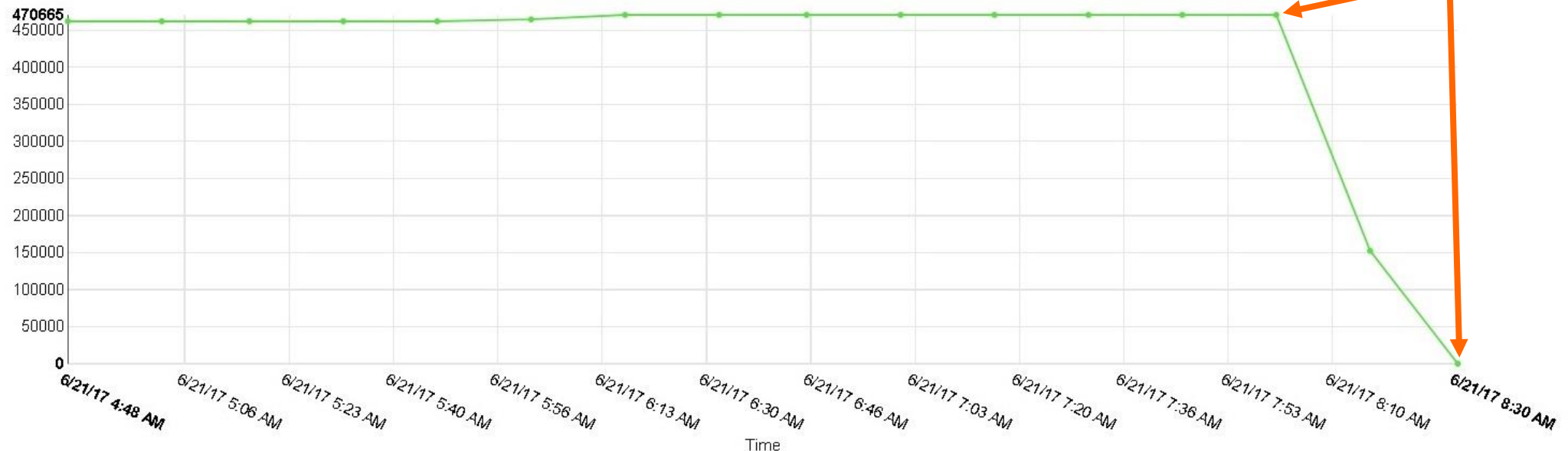
Add table record count

- Identify, Reconcile and Post Top 5 Blocking Versions Every Day
 - Blocking Versions Cause Inefficient Compress
 - Increase in Delta Table Records Beyond the Hardware Support Level
 - Maintain the Lineage Length <100

Versioning statistics	
Number of versions: 664	
Number of versions blocking DEFAULT: 625	
Top 5 blocking versions...	
ANDREW.SN_164022	
ANDREW.SN_162751	
RASU.SN_163090	
RASU.SN_163139	
RASU.SN_164468	
Number of states: 4333	
Number of state lineages: 86293	
DEFAULT versions lineage length: 426	
Last compress: JUL-09-2017	

SQL Server DBs ArcSDE A Table Rows -- ADDS TABLE RECORD COUNT (COUNTER) -- Principle: track record count

Chart resolution: real-time value at collection interval when query less than 12 hrs





Geodatabase Maintenance – Best Practices

- Remove Geoprocessing (GP) History
 - *How To: Automate the process of deleting geoprocessing history*
 - <http://support.esri.com/technical-article/000011751>

- Disable (GP) History for Scripts

```
import arcpy
```

```
arcpy.SetLogHistory(False)
```

Python Tasks	Before Deleting GP History in Seconds	After Deleting GP History in Seconds	Performance Improvement in %
Create Version	400.74	16.3	2358.53%
Delete Version	571.23	14.17	3931.26%



Geodatabase Maintenance – Best Practices

- Manage Auditing / History Tables

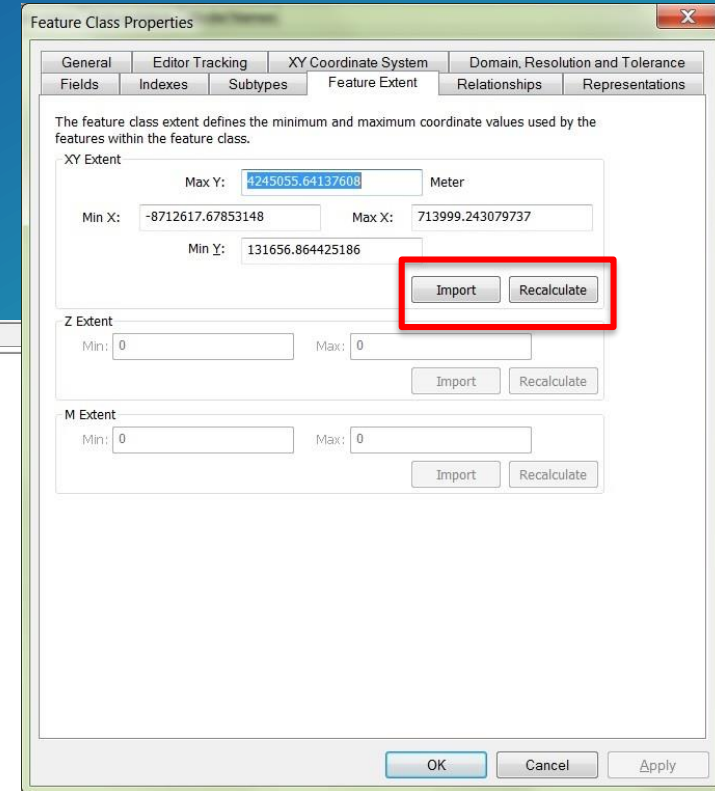
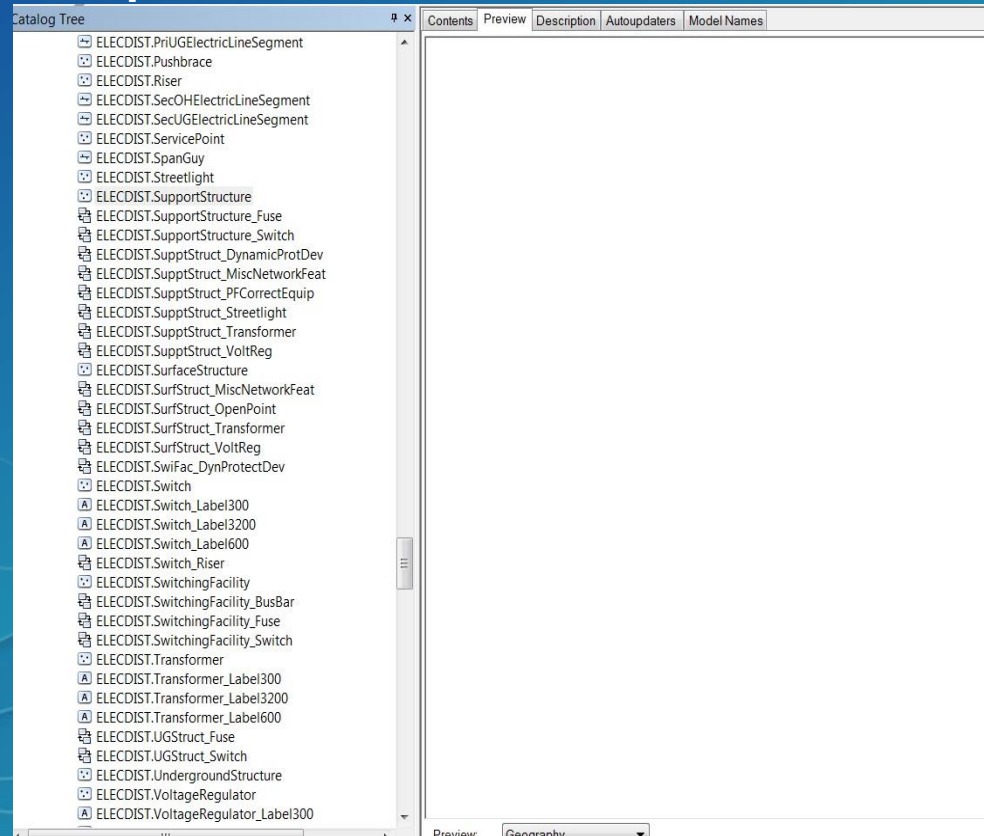
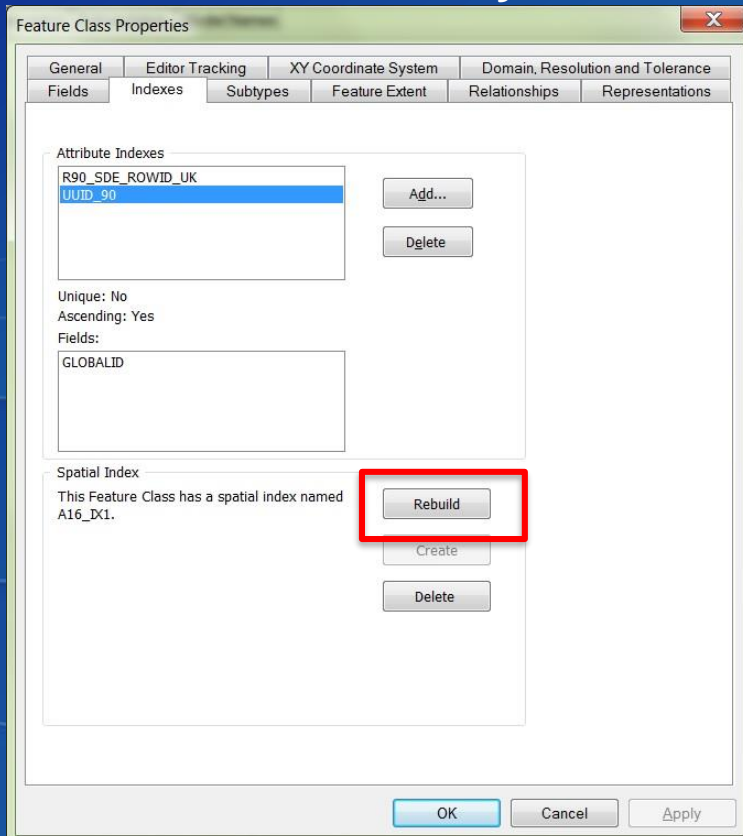
- Reduce the Database

- Backup Size
 - Storage
 - Time

OWNER	TABLE_NAME	NUM_ROWS	BLOCKS	AVG_ROW_LEN	TO_CHAR(LAST_ANALYZED, 'MON/DD/YYHH24:MI:SS')
ARCFM	EDITEDFEATURESTRACKING	12330909	122954	140	APR/23/16 23:44:17
ARCFM	EDITEDGRIDS	12964827	32969	32	APR/23/16 23:47:55
SDE	ARCSDEUSERLOG	56558633	259246	62	APR/24/16 12:28:25

Geodatabase Maintenance – Recommendations

- Every 3 - 6 Months:
 - Fix the Feature Class Extent First
 - Followed by Rebuilding the Spatial Index



Geodatabase Maintenance – Recommendations

- Fix the Non-Empty Feature Classes with No Spatial Index

```
1 Exec DBMS_STATS.GATHER_SCHEMA_STATS ('ARCFM', estimate_percent=>100,
2   DEGREE=> 7, CASCADE=>TRUE, No_Invalidate=>false);
3 select table_name, GSIZE1, GSIZE2, GSIZE3 FROM sde.layers
4   where gsize1=0 and gsize2=0 and gsize3=0 and table_name in
5   (select table_name from all_tables where num_rows > 0);
```

TABLE_NAME	GSIZE1	GSIZE2	GSIZE3
COATING	0	0	0
CPBOND	0	0	0
CPCABLE	0	0	0
CPGROUNDBED	0	0	0
CPANODE	0	0	0
LINECROSSING	0	0	0
MARKER	0	0	0
PIPEEXPOSURE	0	0	0
TIEINMETHOD	0	0	0
DOCUMENTPOINT	0	0	0
INJECTION	0	0	0
PIGROUTE	0	0	0
PIPEDEPTH	0	0	0
RL_SITE	0	0	0
PIPEJOIN	0	0	0
PIPELENGTH	0	0	0
WELL	0	0	0
LEAK	0	0	0
DOT_CLASS_PREVIOUS_RDETAILS	0	0	0
ALIGNMENT_SHEETS	0	0	0
EXTERNAL_COATING_RDETAILS	0	0	0
MISC_FITTING_DETAILS	0	0	0
ALIG_SHT_CROSS_REF_RDETAILS	0	0	0
CLOSURE_DETAILS	0	0	0
HCA_PREVIOUS_RDETAILS	0	0	0
INJECTOR_DETAILS	0	0	0
ODORANT_RANGE_RDETAILS	0	0	0
OFFLINECOMPSTATIONPOLYGONS	0	0	0
PIG_SIGNAL_DETAILS	0	0	0
PIR_RDETAILS	0	0	0
RIGHT_OF_WAY_RDETAILS	0	0	0
RIVER_WEIGHT_RDETAILS	0	0	0
SHEET_NOTE_RDETAILS	0	0	0
TAP_DETAILS	0	0	0
PIPE_SEG_PIR_BUFF	0	0	0
MAOP_CALC_RDETAILS	0	0	0

Geodatabase Maintenance – Best Practices

- **A Sample Daily Maintenance (Batch Process) for Multi User Geodatabase:**
 - Backup the Database
 - Synchronize any Replica Version.
 - Delete the Orphan / Unnecessary versions.
 - Drop the orphan keyset tables (Oracle)
 - Reconcile and Post All/Eligible Versions – Through out the Day
 - Only Reconcile All versions (>100 versions - Parallel Reconcile)
 - Update Database Statistics - Optional

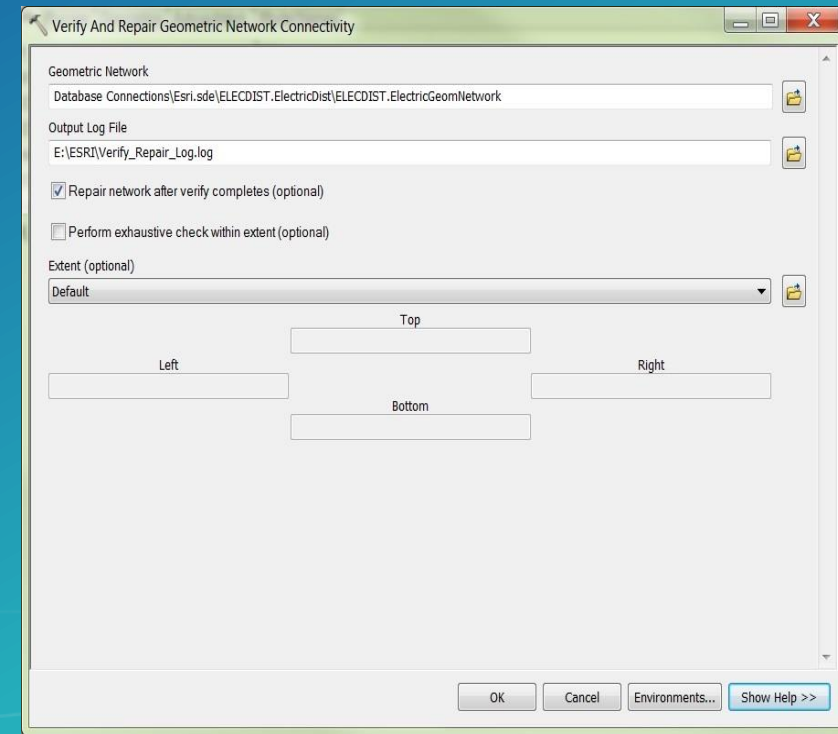
Continue.....

Geodatabase Maintenance – Best Practices

- Continue.....
 - Pause the SDE Connections
 - Kill the Existing or Orphaned User Connections
 - Truncate Dynamic tables
 - state_locks; table_locks; object_locks; layer_locks; process_information;
<user>.SDE_LOGFILE_DATA;
 - Start the Compress Process
 - Un-pause the SDE Connection
 - Rebuild Indexes in RDBMS for all the Schema Owners and SDE
 - Update RDBMS statistics for all Schema Users and SDE.

Geodatabase Maintenance – Geometric Network

- Every 3 – 6 Months Run Esri's Verify And Repair Geometric Network Connectivity Tool
- Only GIS Administrator Should Follow the below steps:
 - Create a New Version under SDE.Default.
 - Create a SDE Connection Document with the New Version.
 - Run the Verify and Repair tool
 - With “Repair network after verify completes” option
 - Reconcile, Post and Delete the newly Created version.
 - Compress the Database.



Caution: Individual Users Should not run it. Generate larger number of delta table records and if someone accidentally run this tool under SDE.Default version directly, it can produce more conflicts.

Implementation Tools

ArcGIS Monitor

- **MXDPerfstat**
 - An ArcGIS Engine command line tool to diagnose typical mxd performance problems
 - <https://www.arcgis.com/home/item.html?id=a269d03aa1c840638680e2902dadecac>
- **ArcGIS (System) Monitor**
 - System Monitor is a tool for monitoring and analyzing your enterprise GIS system.
 - <http://www.arcgis.com/home/item.html?id=848f48b0f88e4de7a036377197453efe>
- **System Designer**
 - A comprehensive tool for designing and capacity planning of GIS solutions.
 - <https://www.arcgis.com/home/item.html?id=8ff490eef2794f428bde25b561226bda>
- **System Log Parser**
 - A reporting tool specifically designed for analyzing ArcGIS server and service logs
 - <https://www.arcgis.com/home/item.html?id=90134fb0f1c148a48c65319287dde2f7>

Questions and Answers

Contact Info:

Rasu Muthurakku

rasu@esri.com

Andrew Sakowicz

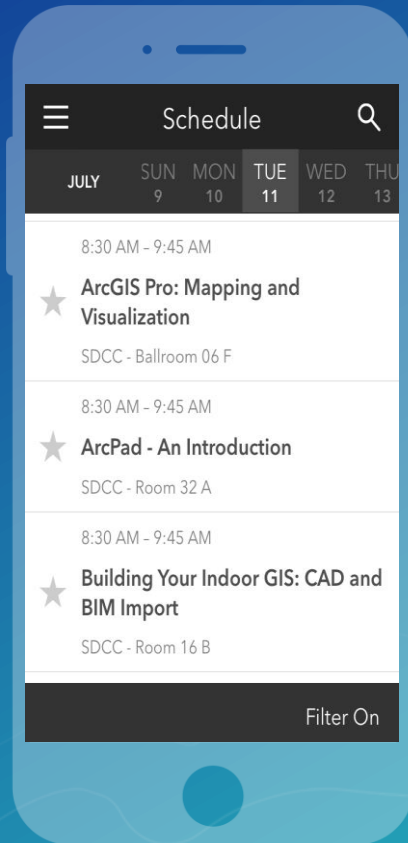
asakowicz@esri.com

Thank you for Attending.
Please Take Our Survey on the **Esri Events App**!

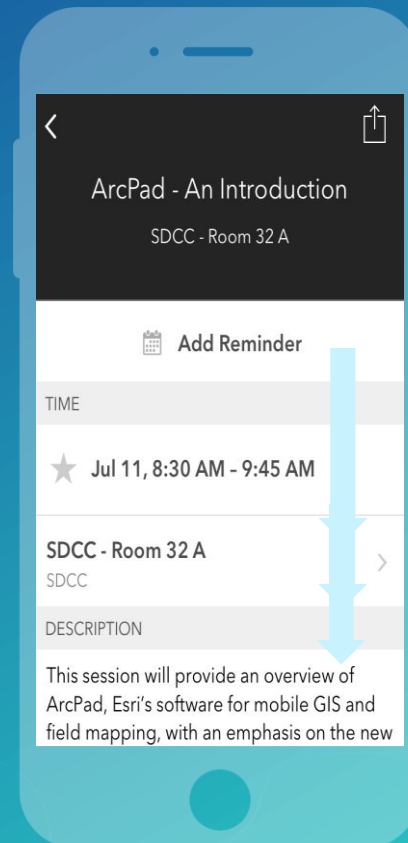
Download the Esri Events app and find your event



Select the session you attended



Scroll down to find the survey



Complete Answers and Select "Submit"



Thanks



esri

THE
SCIENCE
OF
WHERE