Troubleshooting Performance Issues with Enterprise Geodatabases

Ben Lin, Jim McAbee, Jonathan Farmer

blin@esri.com, jmcabee@esri.com, jfarmer@esri.com

SEE WHAT OTHERS CAN'T

Workshop Agenda

- Performance Troubleshooting (Then & Now)
- Real-World Examples and Troubleshooting
- Tips, Tricks and Guidance
- Tools
- Q&A Open Forum

Few questions to ask first…

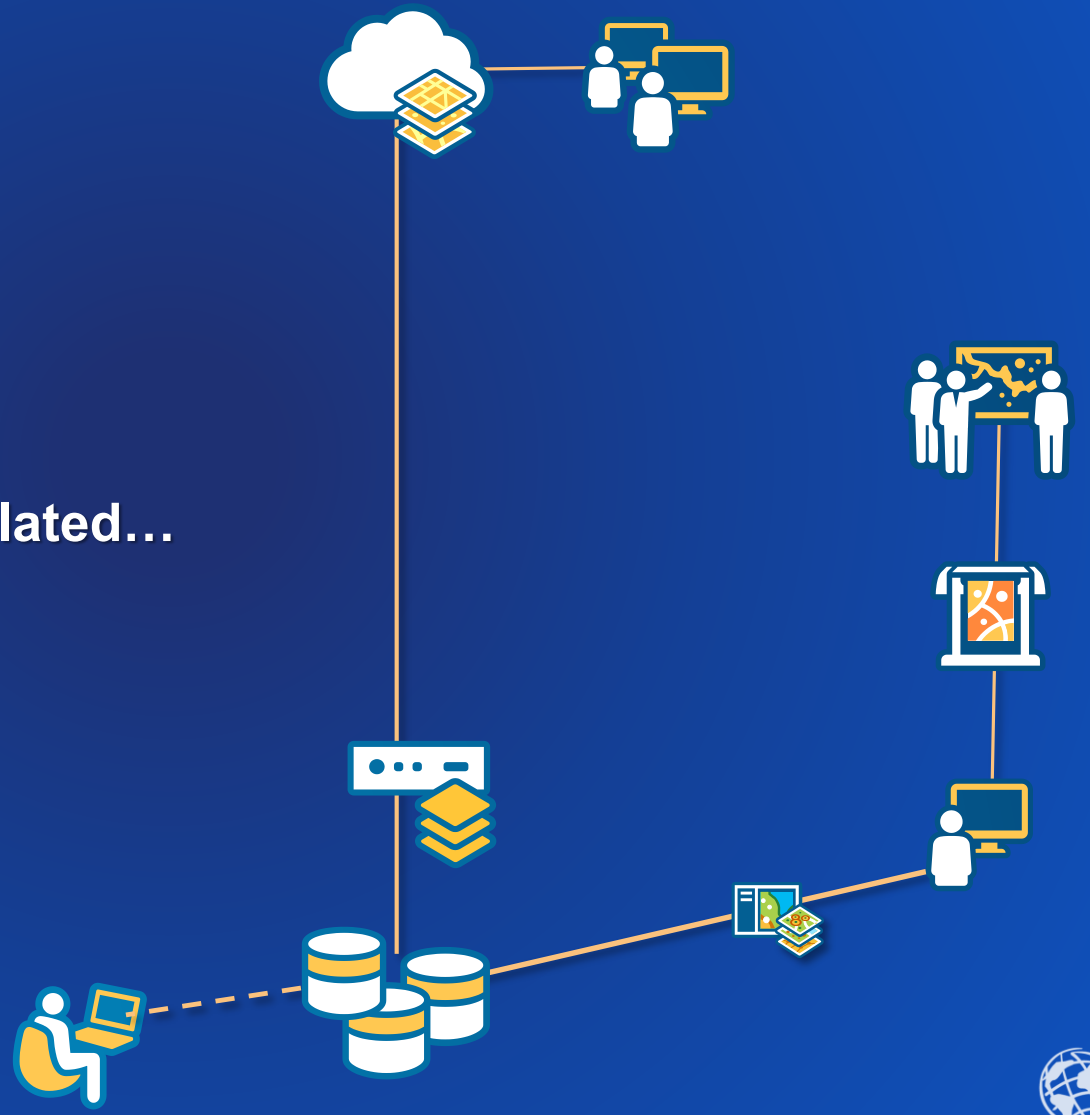
- **What is slow?**
  - 30 seconds may be ok for a batch task or geoprocess, but not editing workflow.
- **What is between the end user and the enterprise geodatabase?**
  - Understand infrastructure and processes when workflow is happening.
- **What was the workflow and specific task being done and when?**
  - What was the end user, or system, doing when the issue was noticed and what time of day.

Troubleshooting Enterprise Geodatabase Performance Issues

- ArcGIS Platform data management includes many solutions
  - Enterprise geodatabases
  - Databases
  - ArcGIS Data Store
  - File geodatabases
  - Imagery
  - Other structures
- This session focuses on issues with relational databases and enterprise geodatabases

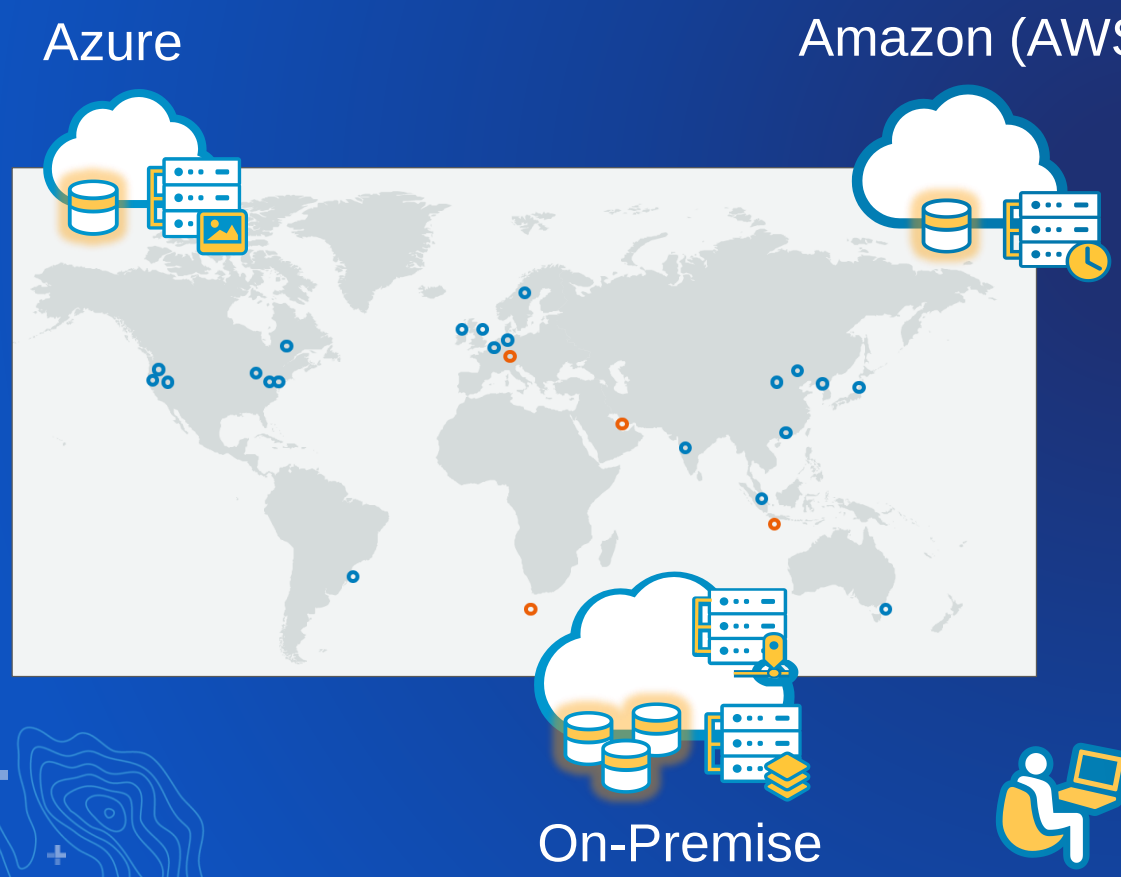
Troubleshooting performance issues used to be simple

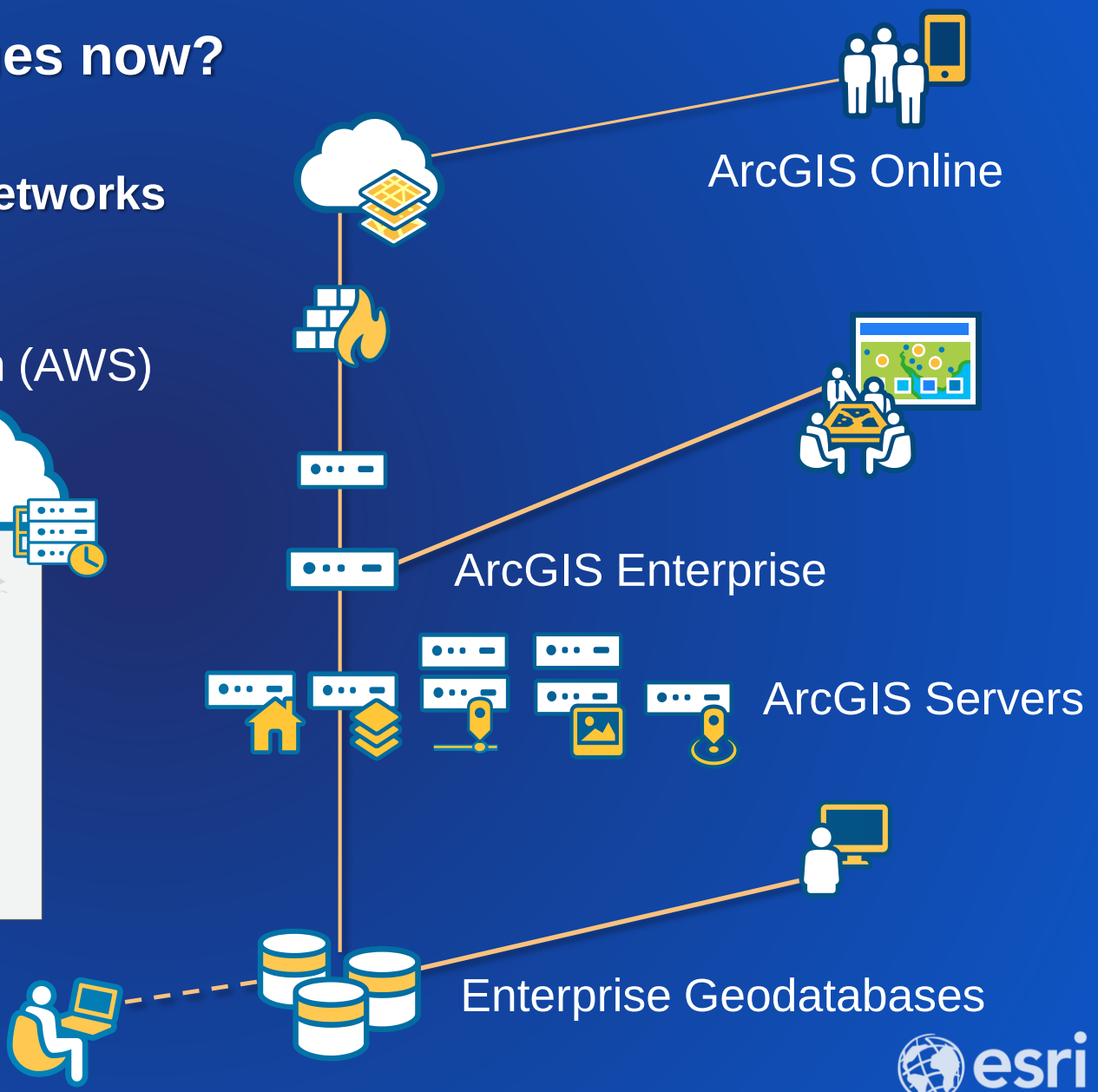
- Easier to find in the past…
- Geospatial data and use was less pervasive in IT…
- Performance issues were more isolated…
- One environment, one location…



Where are the performance issues now?

- **More components, environments, networks & locations…**





Spatial Data Management is now an Enterprise-Wide platform

- Pervasive in Enterprise IT
- Integrated with Business System
- Extends Enterprise-Wide
- All good, but can make performance troubleshooting more difficult…

Real-World Examples and Troubleshooting

- Example 1: Where is the Data, Location Matters
- Example 2: Database Sprawl, How Old is the Geodatabase?
- Example 3: Picking the Slow Line
- Example 4: The Bloated Service
- Example 5: A Hanging Two-way Replica Synchronization
- Example 6: Slow Connection to an Enterprise Geodatabase

Example 1: Where Does Your Data Live – Location is Important

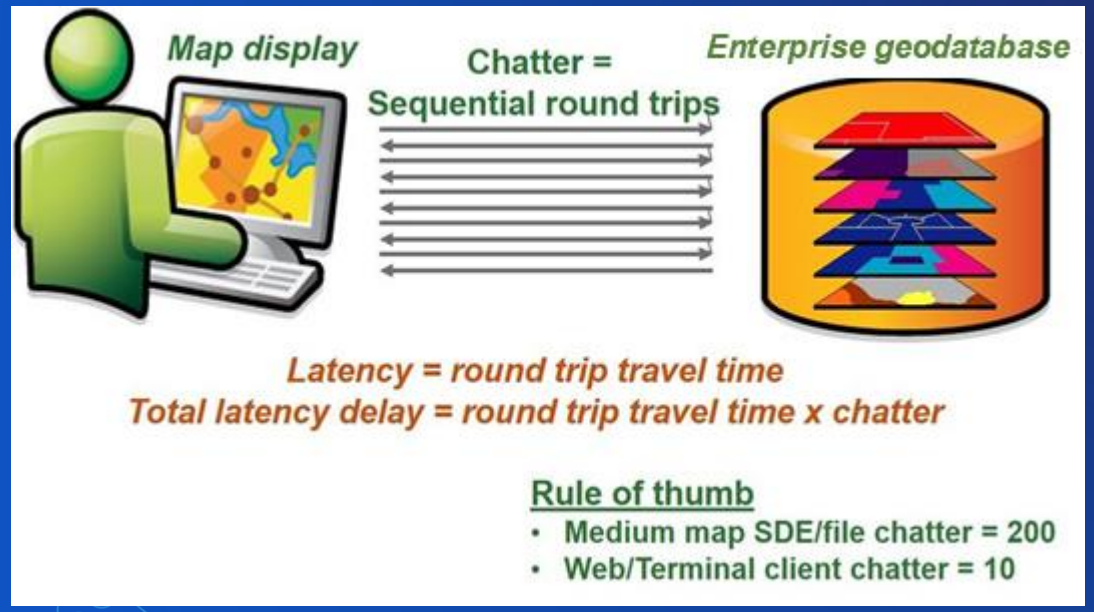
- Customers experiencing slow performance on database queries.
  - In some cases performance issue was random, time of day did not make a difference
- Investigation revealed database performance on server was good.
- Where was performance issue?
- Or more important where was the data?

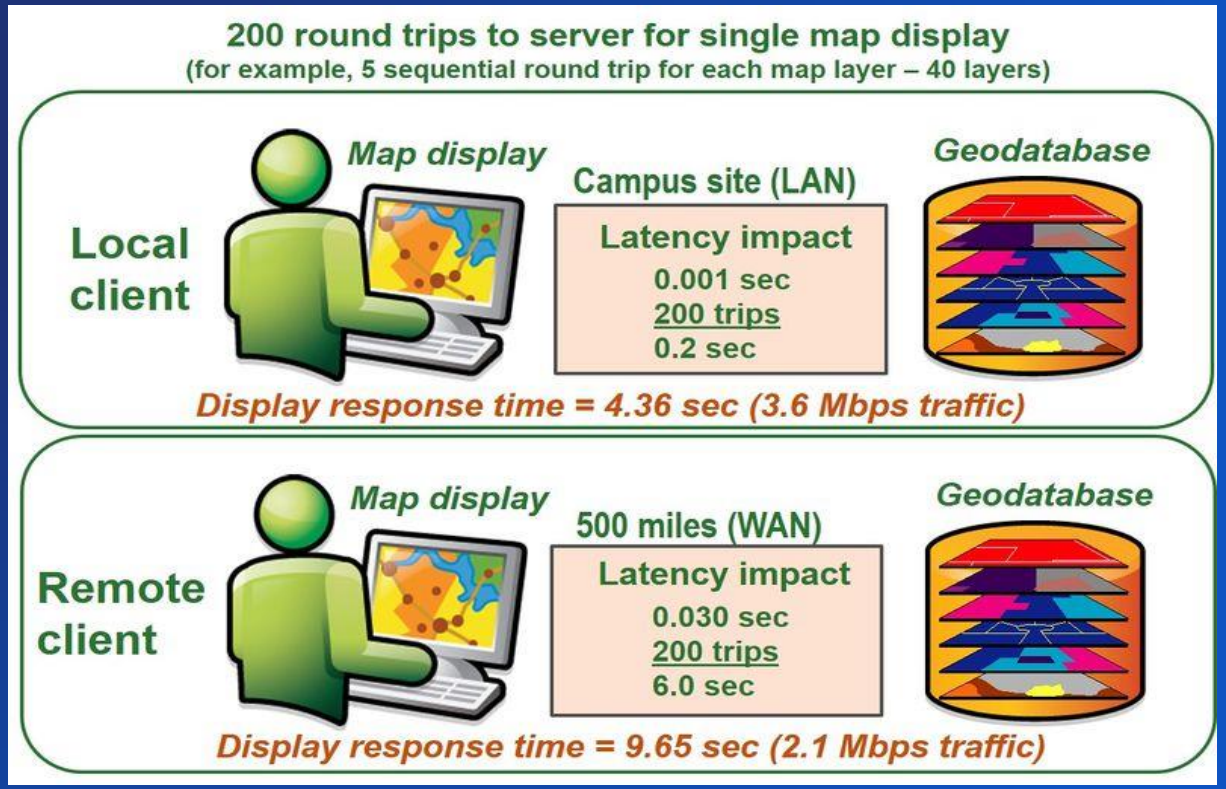
Where Does Your Data Live – Location is Important

- 3 different scenarios
  - Geodatabase in a different AWS Region than the ArcGIS Servers and Desktop tools
  - When there were two data center locations with ArcGIS Desktop Citrix server farms, with the geodatabase at one – causing random performance issues.
  - Geodatabase and ArcGIS Servers in the same virtual server hosting environment, but on different sub-networks with a poor network configuration
- All of these introduced latency impacting performance
- Make sure the enterprise geodatabase is close to the ArcGIS Servers and Desktops

Where Does Your Data Live – Location is Important

- Location AND distance
- Server and enterprise geodatabase need to be as close as possible
- Inefficient network architecture designs can even introduce latency within the same data center.





Example 2: Database “Sprawl” – How Old is the Geodatabase?

- Staff throughout organization experiencing slow performance both in office and field.
- Performance issues random, varied by workflow and data used.
- Enterprise geodatabase and application 15+ years old
  - upgraded in-place through multiple ArcGIS versions
- Large number of views, some with 40 joins.
- Original database designers no longer present.
  - Staff were not clear on purpose of the views

Database “Sprawl” – How Old is the Geodatabase?

- The geodatabase contained a very large number of views
- Views based on views
- Views with 30 - 40 joins
- Views constructed without consideration of performance
- Custom legacy application needed to be redesigned

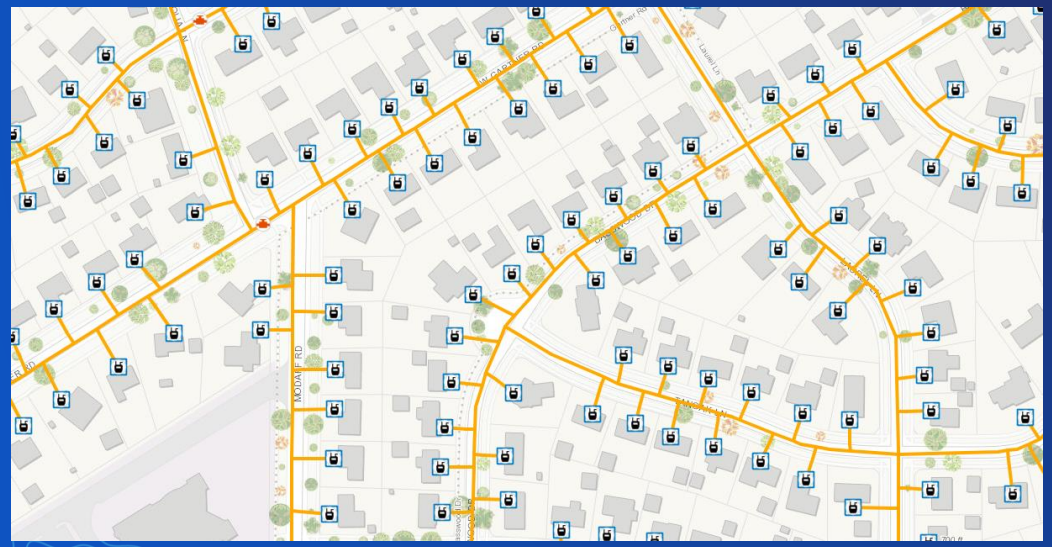
Database “Sprawl” – How Old is the Geodatabase?

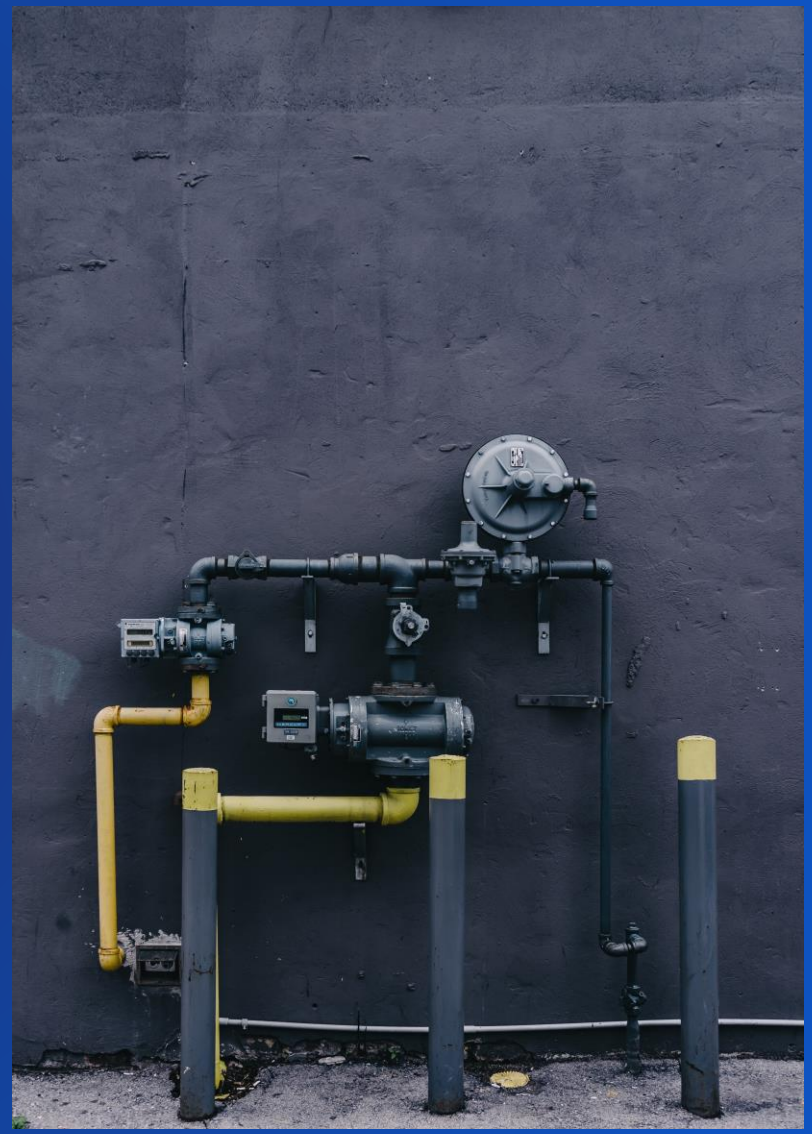
- How can I improve performance?
  - Reduce view levels
  - Clean up unused or redundant views
- Revisit business logic and database design more frequently
- Views built on views may be ok for data warehouse (OLAP) workflows
  - Very inefficient within transactional (OLTP) workflows.

Example 3: Picking the Slow Line

- Feature service referencing an enterprise geodatabase with 1,100 polylines
  - Representing gas distribution lines for a small utility

- The GIS Manager receives complaints about slow performance from mobile apps

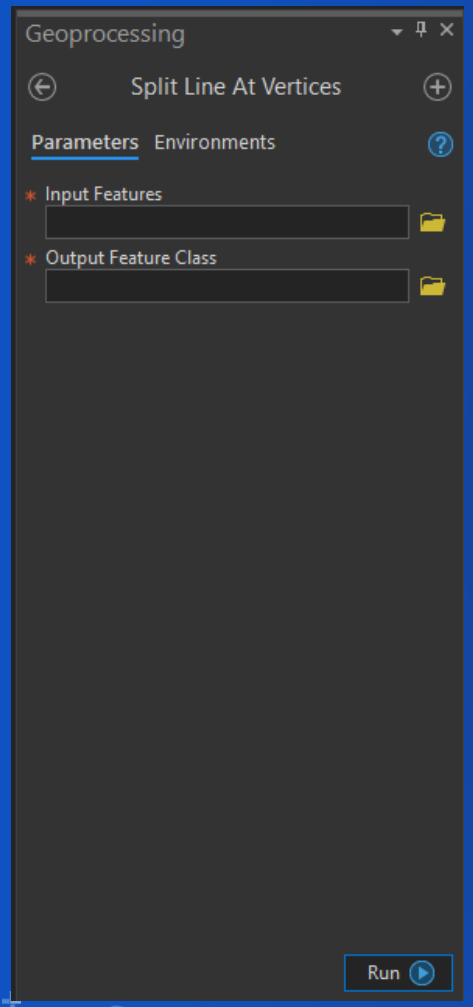


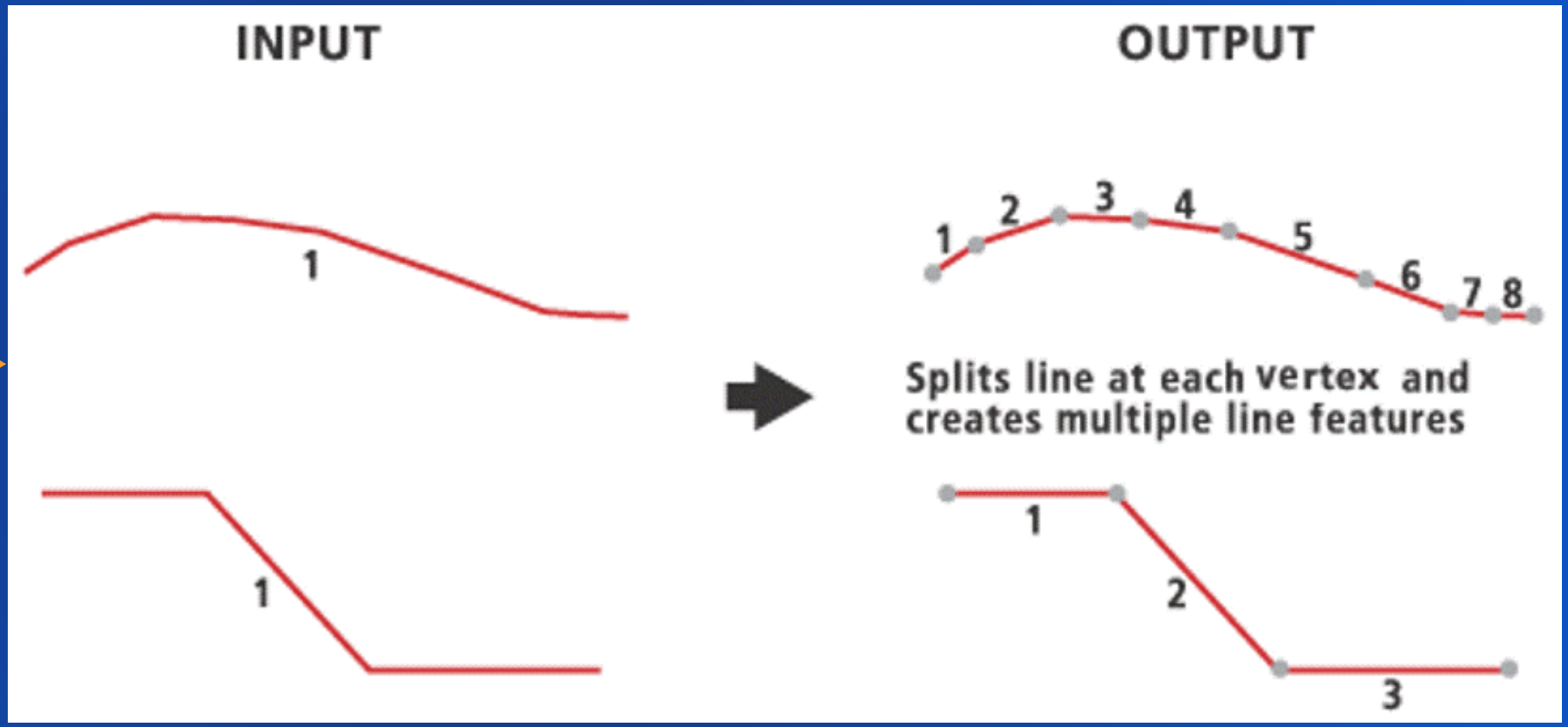


What questions can we ask to troubleshoot?

- Sudden or gradual?
- Specific to 1 person or machine?
- Do you see the performance behavior from REST or Portal Map Viewer?

The core issue





Why does this matter?

- 1,100 lines now 11 million!
- Those are all rows in a table
- Operational Layer > Dynamic Map Service > Display is dynamically generated

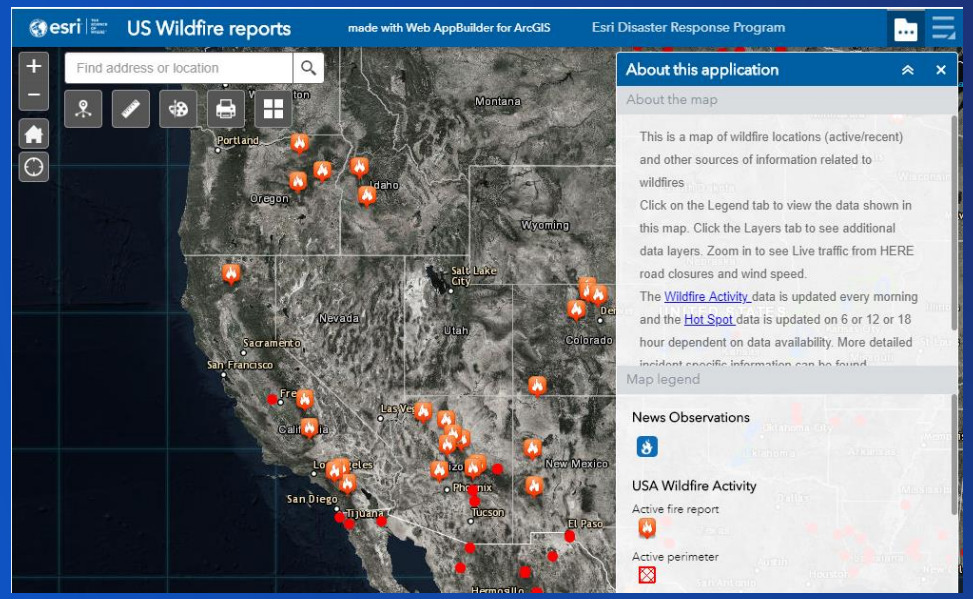
	OBJECTID	ASSETGROUP	ASSETTYPE	ASSOCIATIONSTATUS	ISCONNECTED	FROMDEVICETERMINAL	TODEVICETERMINAL	CREATIONDATE	CREATOR	LASTUPDATE	UPDATEDBY	SHAPE	GDB_GEOMATTR_DATA	GLOBALID
1	1	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000040682950EF912F41C0A31ADA404...	NULL	39A2550A-7(
2	2	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000060006EB455932F418084C8B2A346...	NULL	7F06E514-6F
3	3	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000E091E9D673B02F4140B827666142...	NULL	0C617EE4-3(
4	4	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000201D5FAE4A8E2F4180B85E45C23...	NULL	5D7653EB-6
5	5	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000A0FEE117A4A22F41405F8CF70A4...	NULL	E64AFFB2-A
6	6	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000000DB6B9500862F4180183C54E543...	NULL	267D6876-2
7	7	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000040BB68956AA72F4100EE810BFE3...	NULL	2AE67E3B-6
8	8	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000060FD238CA4A72F41808BB5558A4...	NULL	0C212EC7-5(
9	9	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000A01C505B62A32F41C0FB797A0D4...	NULL	56D6118D-5
10	10	1	9	0	1	0	0	2019-02-21 19:03:46.0000000	GASOWNER	2019-02-21 19:03:46.0000000	GASOWNER	0x0000000001070300000020978CE754992F416012CE939C44...	NULL	8CA1881F-D

How do we address this?

- Restore feature class from a backup
  - Republish to ArcGIS Server
- What if I meant to run this tool?
  - Use map scale dependencies
  - Use ArcGIS API for Javascript
    - Query smaller record sets or smaller areas
    - Points display quicker than lines and polygons

Example 4: The Bloated Service

- Map service that represents all operational layers for a state fire agency
- Customer notices after publishing that it is performing poorly from the Portal map viewer



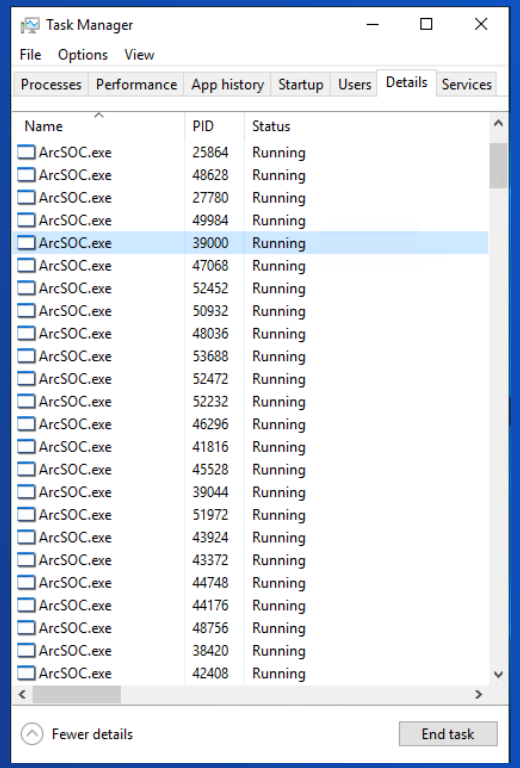
esri THE SCIENCE OF WHERE

What questions can we ask to troubleshoot?

- Is the performance poor for everyone or just you?
- Where is the data is relation to ArcGIS Server?
- How many layers are we talking about?
- Are they all heavily used and edited?

The core issue

- Map service has over 100 layers
- Underlying data is normal



- Maximum instance setting of 15 on this map service **not enough** to handle traffic
- Overall, agency had almost 200 ArcSOCs with **high RAM** usage
- Reduce number of total ArcSOCs, increase max instances and heap size, consider cached service for non-operational layers

Shared vs. Dedicated Instances

- Historically each service had its own dedicated instance pool
  - Service instances and services not one to one
- Each instances (ArcSOC) takes 100MB to 200MB of RAM

- New at 10.7 - Shared instances
  - A shared pool containing 4 ArcSOC processes by default
  - Services dip into the pool as needed and return the SOCs when done
  - Existing services can be changed to use shared pool

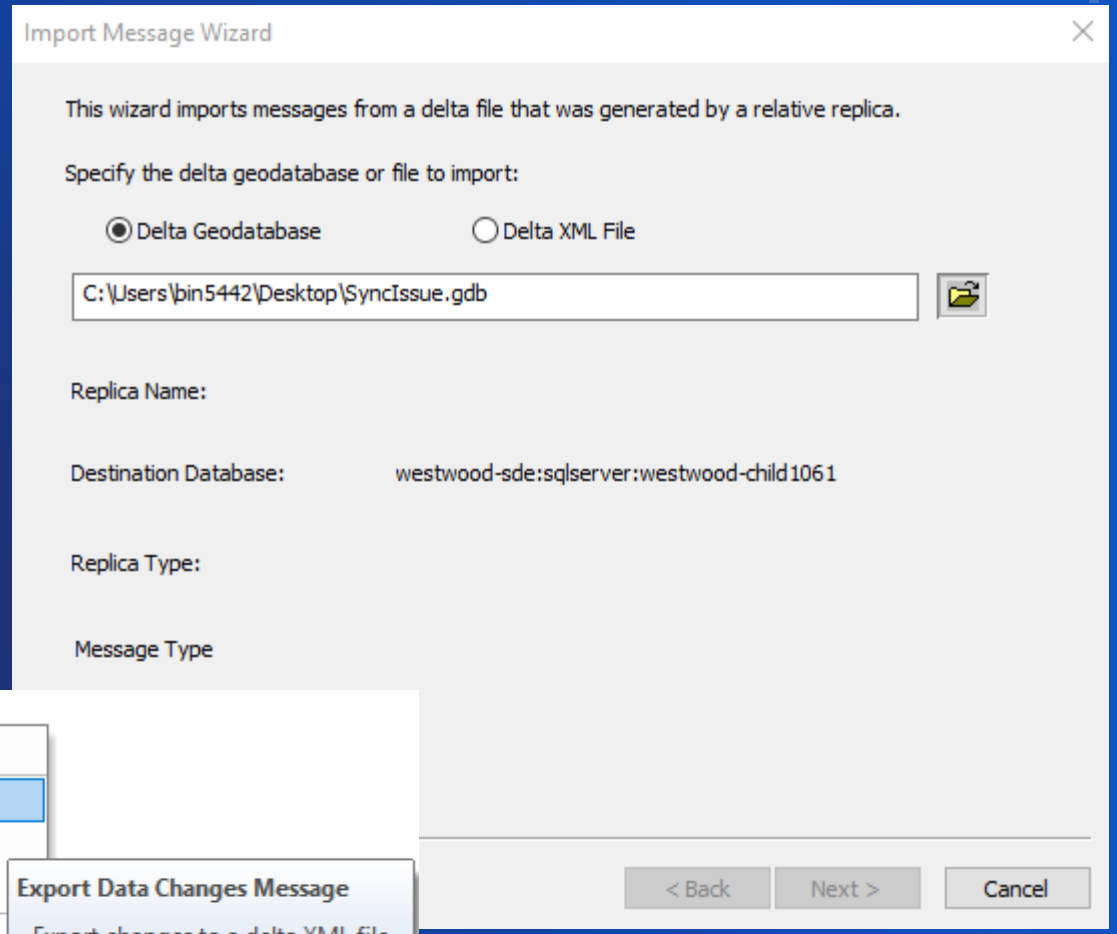
Example 5: A Hanging Two-way Replica Synchronization

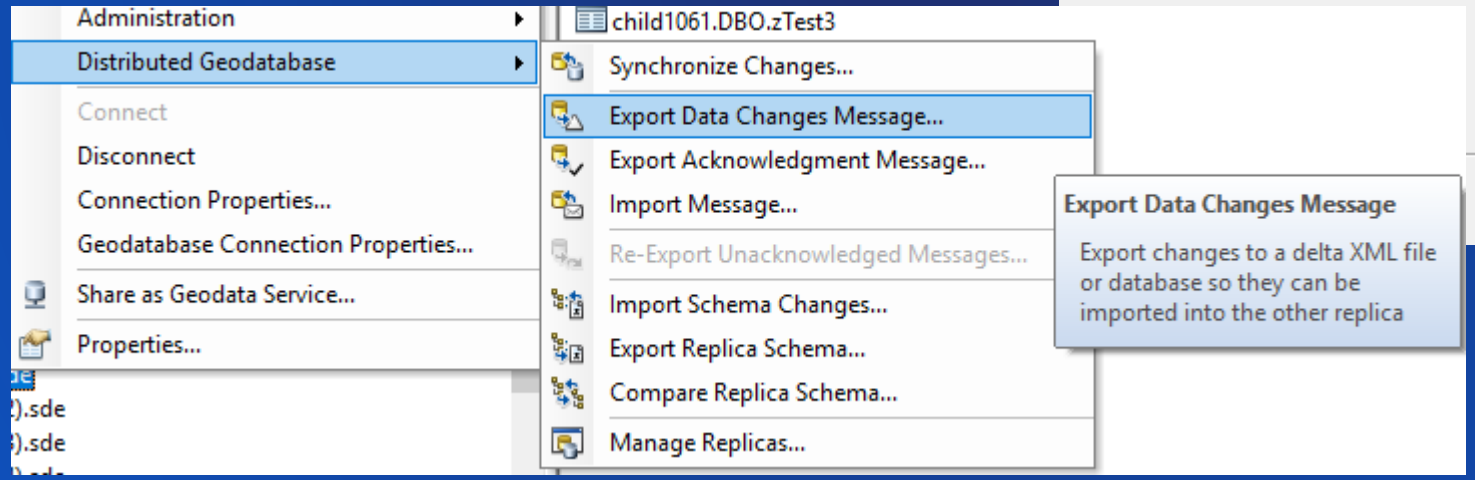
- What's the issue?
  - A connected synchronization for a two-way replica couldn't finish, it just hangs…

- How to troubleshoot the issue?
  - What changed between the last successful sync and now?
  - Any synchronization errors occurring before this hanging issue?
  - Is the issue reproducible if restoring the geodatabases to other machines?

Example 5: Troubleshooting Process

- **Simplify the synchronization operation**
  - ✓ **Export Data Changes**
  - ✓ **Import Data Changes**
  - ✓ **Export Acknowledge Messages**
  - ✓ **Import Acknowledge Messages**





Example 5: Resolution

- How to solve the issue?
  - ✓ Calculate Database Statistics
  - ✓ Rebuild Spatial Indexes
  - ✓ Compress Enterprise Geodatabase

- Additional recommendations for database maintenance
  - ➢ Create the ArcSDE log file tables as Global Temporary Tables in Oracle (unless created after 10.4)

    https://support.esri.com/en/technical-article/000009061
  - ➢ Automate reconcile and post operations for sync-enabled data

    https://desktop.arcgis.com/en/arcmap/latest/manage-data/geodatabases/automate-reconcile-post-after-sync.htm
  - ➢ Update the Oracle Open_Cursors parameter to a larger number

Example 6: Slow Connection to an Enterprise Geodatabase

- **What's the issue?**
  - ArcCatalog takes a long time to connect (2+ minutes) to an enterprise geodatabase
  - Once connected, clicking objects takes an additional minute

- **How to troubleshoot the issue?**
  - Does this happen for everyone or specific user?
  - Does this happen at a particular time of day?
  - How large is the geodatabase?

Example 6: Troubleshooting Process

- **Comparison is the key**
  - ✓ Always compare between a good performance and a slow performance whenever it's possible
- **Tools can be used in this case**
  - ✓ SDEINTERCEPT
  - ✓ RDBMS database trace

```
==========================================
[W 10:23:01.853] Command:       QueryWithInfo
[W 10:23:01.853] Long:          1
[W 10:23:01.853] Query Info:
    Num Columns:    1
    Columns:        "Documentation"
    SQL_Construct: [1]
    Tables:         "MyEGDB.dbo.GDB_Items"
    WhereClause:    "PhysicalName = N'WORKSPACE' AND Type IN ('{C673FE0F-7280-404F-8532-20755DD8FC06}')"
    Query Type:     4
    Num Hints:      0
    Num Parameter markers: 0
    Logfile:        <null>
[R 10:23:01.855] Long:          0
[R 10:23:01.855] Col_Defines:   [1]
    Name                              Type     Width nDec  NULL?    RowID
    ------------------------------ ---------- ----- ---- -------- -----
    Documentation                  SE_XML     2147483647   0       NULL
    ------------------------------ ---------- ----- ---- -------- -----
```

Example 6: Resolution

- Cause of the issue
  - ✓ The documentation field is bloated
  - ✓ Long query time spent on the XML column “Documentation” in the GDB_ITEMS table

- How to solve the issue?
  - ✓ How To: Delete geoprocessing history from a geodatabase
    https://support.esri.com/en/Technical-Article/000011751

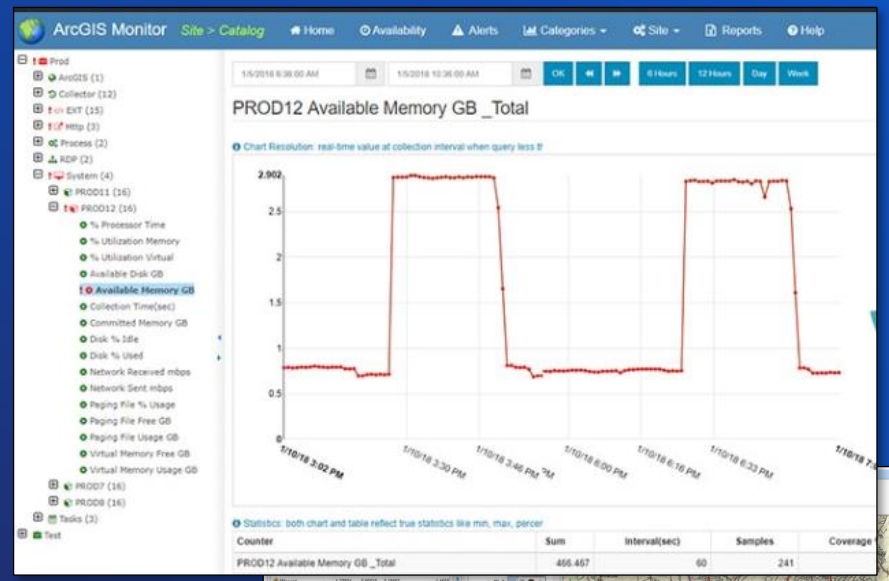
- How to prevent the issue in the future?
  - ✓ How To: Disable geoprocessing history in metadata
    https://desktop.arcgis.com/en/arcmap/latest/analyze/executing-tools/history-log-files.htm#GUID-3FDAC815-34EE-4CF9-87C7-A787D02C6487

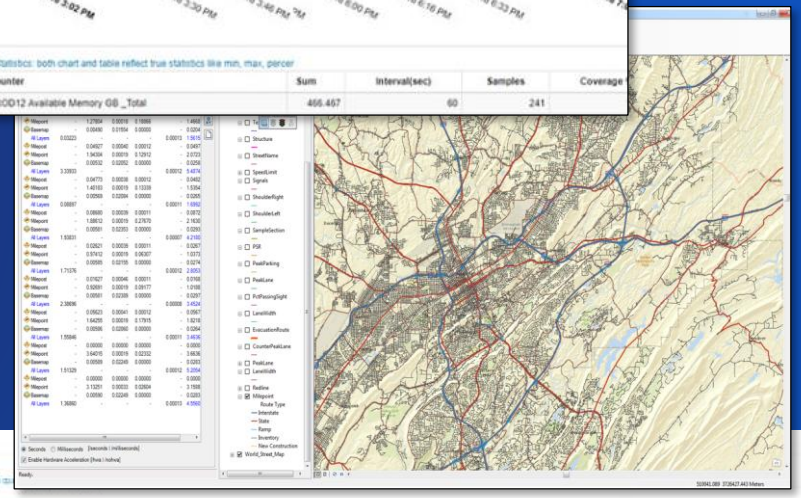
Troubleshooting – Some Typical Scenarios

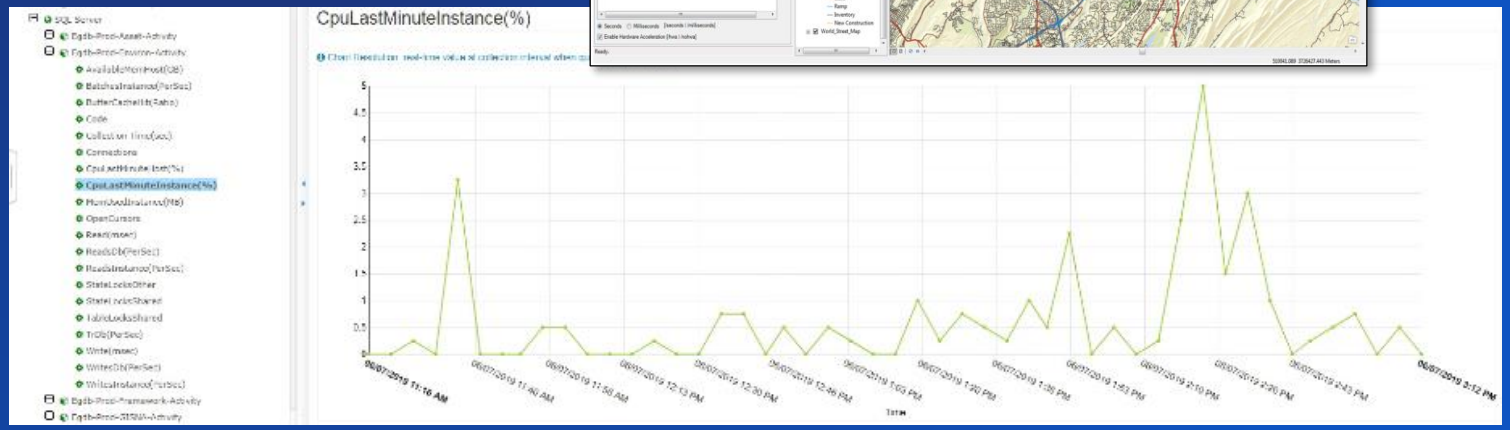
- Inefficient application configuration
  - Overly complex map document, too many layers, complex symbology, etc..
- Inefficient database design and/or maintenance
  - Geodatabase compress, statistics, index maintenance
  - Store everything, but query only what you need!
- Inefficient architecture
  - Large physical distance between database and “thick” clients – Desktop and Server
- Resource Configuration Issue
- Security Configuration
  - Is web sense or other web filtering present?

Troubleshooting - Tools

- ArcGIS Monitor
- Extensions – EGDBHealth
- ArcGIS System Log Parser
- ArcGIS PerfQAnalyzer and MxdPerfstat
- OS performance metrics and logs
- Direct Connection Log
- Database Specific
  - SDEIntercept – communication between ArcGIS and database
  - RDBMS Database Traces – SQL and execution plans





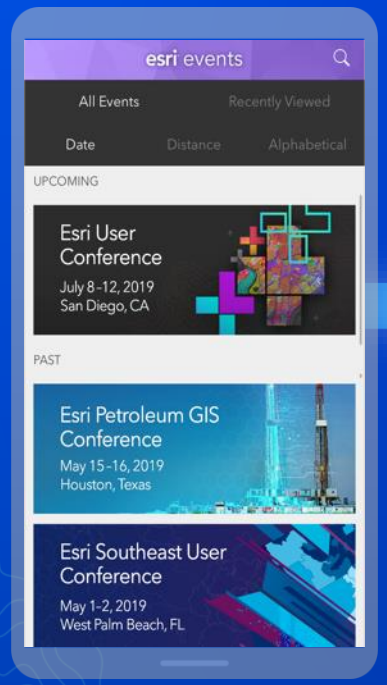


Troubleshooting – Esri Support

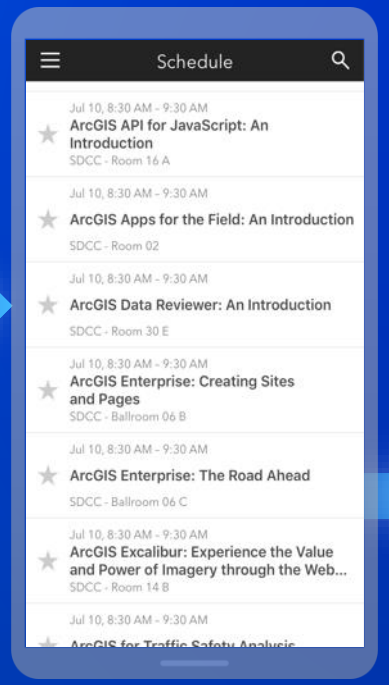
- Review behavior on a test instance (if applicable)
- Simplify the issue and isolate the behavior
- Generate SDEIntercept / RDBMS trace (for a single action)
- Open a Tech Support case
- A database backup might be requested
- If difficult to reproduce, some further services assistance may be recommended…

Please Share Your Feedback in the App

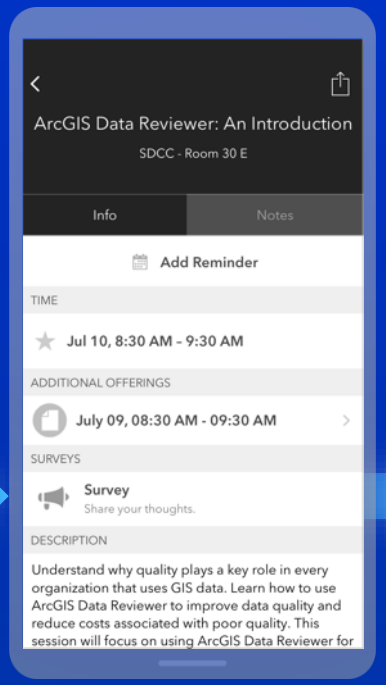
Download the Esri Events app and find your event



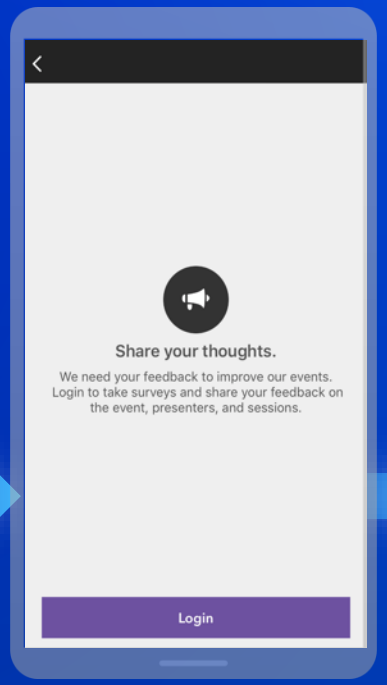
Select the session you attended



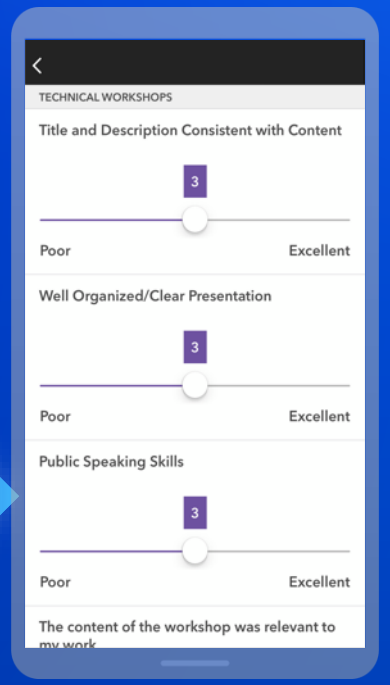
Scroll down to "Survey"



Log in to access the survey



Complete the survey and select "Submit"



Questions?

esri | THE SCIENCE OF WHERE