Esri International User Conference | San Diego, CA

Technical Workshops | 14.07.11

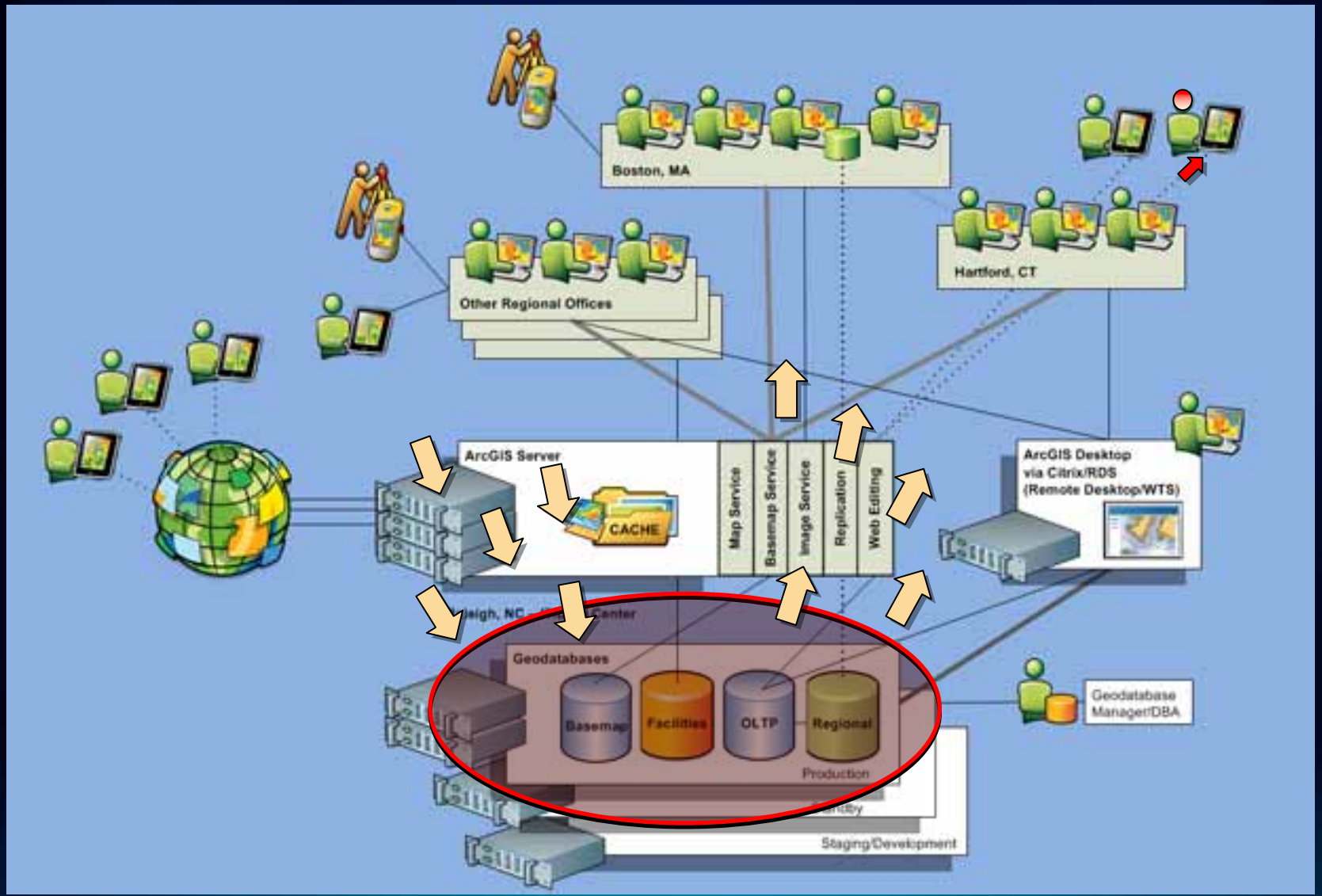
Troubleshooting Performance Issues with Enterprise Geodatabases

Jim McAbee

Agenda

- Overview of troubleshooting process
- Methods for isolating performance issues
- Best practices for performance measurement

Performance: Where are bottlenecks?



Performance

- Performance vs. Scalability
- Types of Performance
  - Speed – typically end user experience (“how fast can I get my answer”
  - Availability
- Performance Tuning and Troubleshooting
  - some similar methods and techniques

Performance Best Practices

- Start with recommended configuration parameters for RDBMS
- Establish Baseline (all tiers)
  - not just Geodatabase but also ArcGIS Server (application server tier).
  - if present also Citrix, VM environments, etc…
- Geodatabase Specific Monitoring
  - Database performance and workload monitoring
  - OS – resource monitoring (cpu load and memory)
  - Network usage – if possible

Performance Troubleshooting

- Starts with problem isolation
  - typically grab a broad scope synopsis of overall system performance – both RDBMS and OS
  - then with this as background information focus on specific performance or functional issue
- Isolation requires
  - logging at all tiers – relates back to monitoring
  - when an issue happens then verbose logging is turned on and tracing is done.
  - goal is to begin to rule out the various tiers and isolate where the issue lies.

Common types of issues

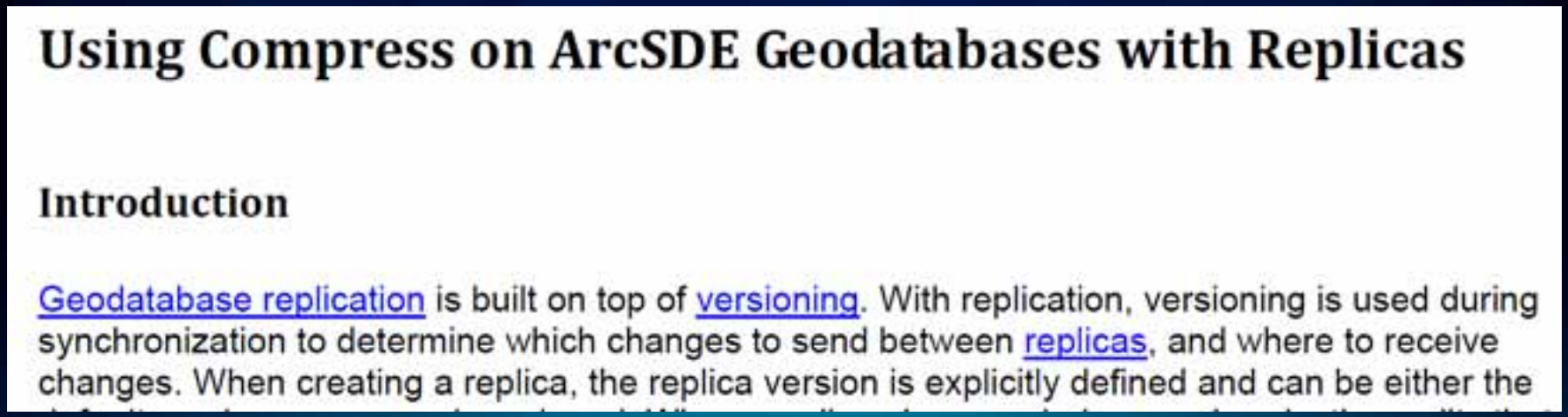
- Connection problems
  - database client version or bit (32-64) mis-match
  - permission or licensing issue
  - database, network down (ping, tnsping, etc..) – also one for SQL Server
- Performance issues typically related to:
  - poor application design
  - poor document design (complex symbology, too many layers, other inefficiencies – all layers selectable, etc..) – generate lots of extra SQL
  - no or bad indexes and/or database statistics
  - insufficient or improperly configured database resources – space, quota, etc…
  - or combination of above

More Common Types of Issues

- Lack of Version Maintenance
  - poor reconcile, post, compress practices
  - inefficient version architecture
  - poor management of replicas
  - all of which can leave unused versions or states pinning state tree leading to more work in the database.
- Upgrades to new ArcGIS or RDBMS version
  - old configuration parameters left in place
- Impacts of various types of migrations
  - server and/or storage, also
  - migration from physical to virtual environment

Version and Synchronization Workflows

- Inefficient version architectures or management/maintenance processes can lead to performance issues.
- Check for recommended reconcile order
  - KB 35735 – Oracle
  - KB 36809 – SQL Server
- Compress and Synchronization technical paper



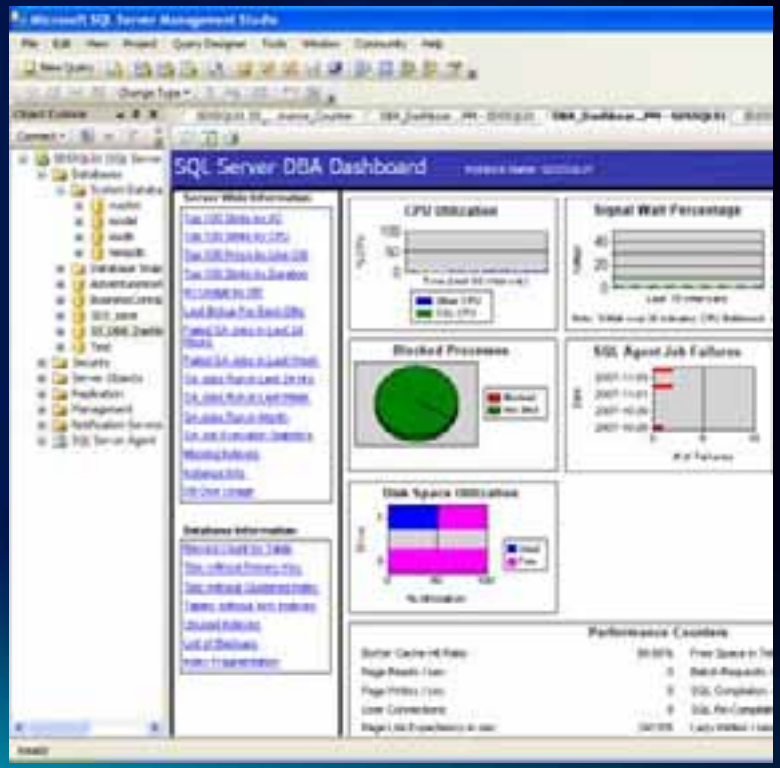
Using Compress on ArcSDE Geodatabases with Replicas

Introduction

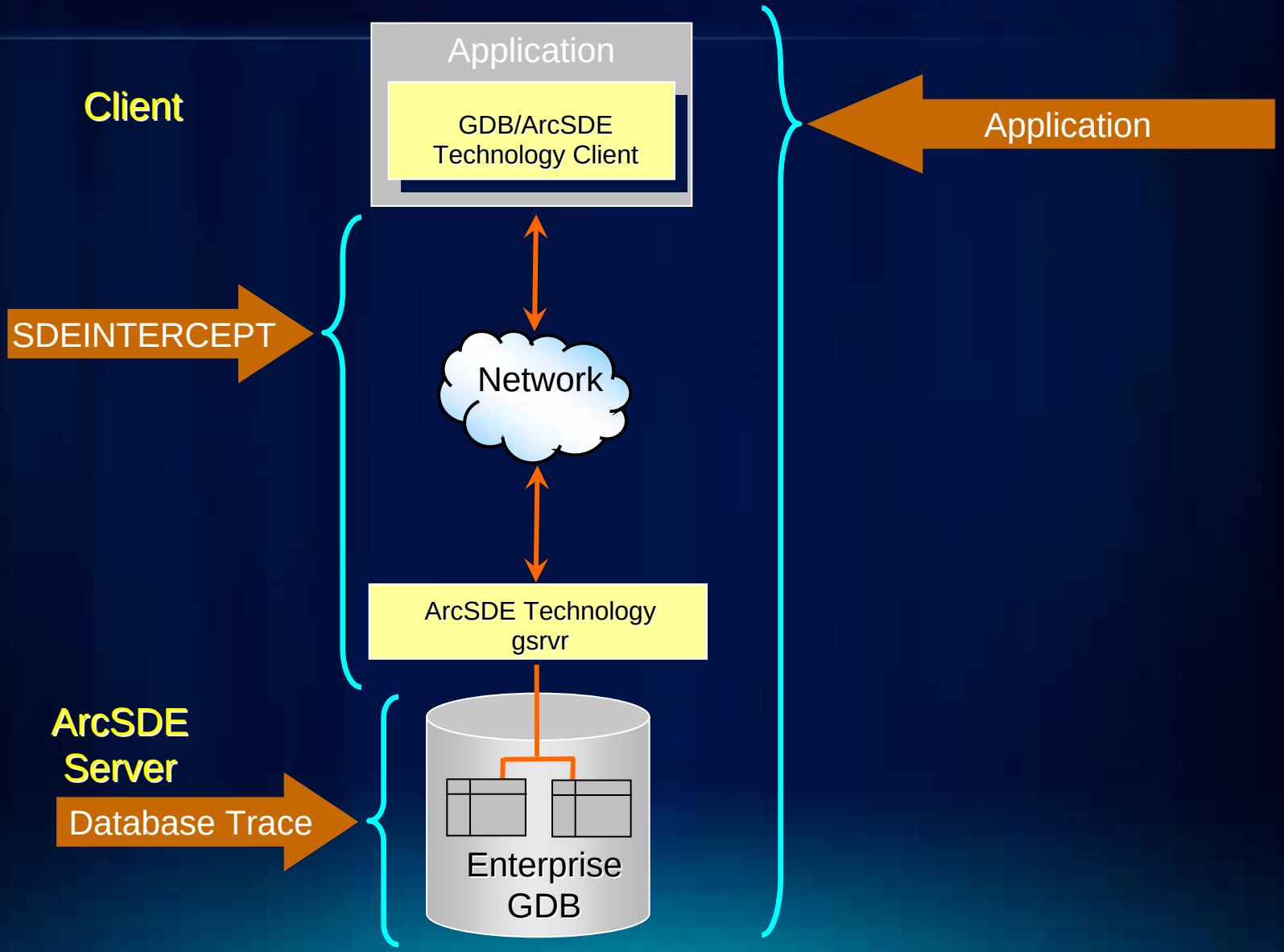
Geodatabase replication is built on top of versioning. With replication, versioning is used during synchronization to determine which changes to send between replicas, and where to receive changes. When creating a replica, the replica version is explicitly defined and can be either the

Tools for Troubleshooting

- SDE logs
  - error logs in SDEHOME\etc directory
  - sdeintercept (and optionally sdetrace)
    - client-side and possibly server-side
- ArcGIS Server logs
- Database logs and traces
  - database error logs
  - database session sql tracing
    - Oracle 10046 event trace
    - SQL Profiler,
    - 3rd party tools (Spotlight)
- OS Tools
  - Windows Task Manager
  - UNIX - various



Log Availability



Oracle SQL Trace Example

```
SELECT /*+ USE_NL (blk bnd) INDEX (GISDATA.SDE_BLK_120 SDE_BLK_120_UK) */ blk.rasterband_id,
rrd_factor, row_nbr, col_nbr, band_types, block_width*block_height, block_data
FROM
 GISDATA.SDE_BND_120 bnd, GISDATA.SDE_BLK_120 blk WHERE bnd.rasterband_id = blk.rasterband_id AND
raster_id = :raster_id AND (sequence_nbr IN (:sequence_nbr1, :sequence_nbr2, :sequence_nbr3)) AND
rrd_factor = :rrd_factor AND (row_nbr >= :miny AND row_nbr <= :maxy AND col_nbr >= :minx
  AND col_nbr <= :maxx) ORDER BY sequence_nbr, row_nbr, col_nbr

call     count       cpu    elapsed       disk      query    current        rows
------- ------  -------- ---------- ---------- ---------- ----------  ----------
Parse        9      0.00       0.00          0          0          0           0
Execute      9      0.00       0.00          0          0          0           0
Fetch       53      0.57      71.68       4610       7388          0        2151
------- ------  -------- ---------- ---------- ---------- ----------  ----------
total       71      0.57      71.68       4610       7388          0        2151

Misses in library cache during parse: 0
Optimizer mode: ALL_ROWS
Parsing user id: 44
Number of plan statistics captured: 1

Rows (1st) Rows (avg) Rows(max)  Row Source Operation
---------- ---------- ---------  ---------------------------------------------------
        51         51        51  SORT ORDER BY (cr=90 pr=0 pw=0 time=1330 us cost=28 size=232 card=8)
        51         51        51   FILTER  (cr=90 pr=0 pw=0 time=929 us)
        51         51        51    NESTED LOOPS  (cr=90 pr=0 pw=0 time=926 us)
        68         68        68     NESTED LOOPS  (cr=89 pr=0 pw=0 time=936 us cost=27 size=232
card=8)       68         68         68      TABLE ACCESS BY INDEX ROWID SDE_BLK_120 (cr=85 pr=0 pw=0
time=663 us cost=16 size=132 card=11)
        68         68         68       INDEX SKIP SCAN SDE_BLK_120_UK (cr=17 pr=0 pw=0 time=252 us
cost=14 size=0 card=11)(object id 21605)
        68         68         68      INDEX UNIQUE SCAN SDE_BND_120_UK1 (cr=4 pr=0 pw=0 time=107 us
cost=0 size=0 card=1)(object id 21596)
```

Log File collection:

- **Purposes**
  - **Check performance**
  - **Establish performance benchmark (under typical workload)**
  - **Troubleshoot Errors/Problems**
- **Establish a baseline.**
  - **Important to collect a set while the system is operating correctly**
- **Useless without knowing the context of what was being logged.**

Performance Baseline

- Establish performance baseline
- spastats
- mxdperfstat
- database reports (e.g. Oracle AWR, statspack, SQL Performanc reports)
- cpu and memory from db server and app servers
- network bandwidth and latency

Thank You

http://www.esri.com/sessionevals