Enterprise Geodatabase Testing Tools and Metrics

Dana Nibby Dave Johnson

North Carolina Department of Transportation April 2008

Presentation Roadmap

- Why test
- What to test
- What you need to test
- Test tools
- Results (highlights)
- Resources



Why Test?

- 1. Quantify impact to existing resources
 Will ArcSDE inundate a shared network?
- 2. Justify acquisition of new resources
- 3. Comply with IT architecture
- 4. Educate IT and GIS

 ESRI notions of tiers vs. standard IT notions of tiers
- 5. Reduce Risk
 What configuration will work best in your unique environment?
- 6. Experiment with processes and procedures
- 7. Build Relationships between IT and GIS
 Someone will have to support the system in production

Taming the Beast







GIS is one of largest applications for IT

Network Traffic:

Comparable in bandwidth to video conferencing

Data:

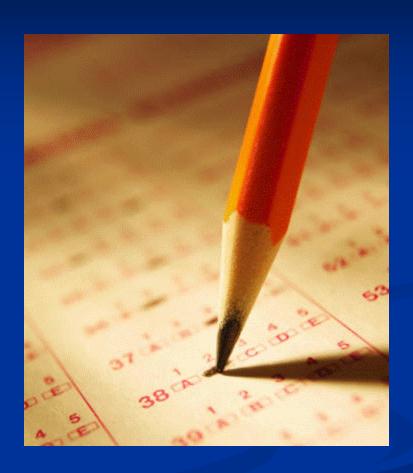
Up to petabytes of data (1,000 TBs)

Computation:

Computationally intensive

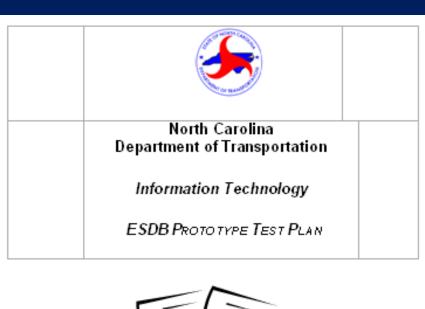
What to Test?

- Data Loading Procedures
- Server load
 - CPU + Memory
- Network load
 - bandwidth
- Client response times (end user experience)
- Database configuration
- ArcSDE parameters



Requirements (What you need to test)

- Test Plan
 - Metrics / KPI
 - Test parameters
 - Tests to perform
 - Test sequencing
- **■** Test Environment
 - Hardware
 - Software
 - Data
- People
 - Test Team
 - HW / SW / DBA Support
 - Management Support





Test Plan: Metrics / KPI

KPI – Key Performance Indicator

- A metric in a particular context
- Typically expressed as a ratio
- Define before testing to determine Pass / Fail criteria

Client Response Time KPI

- e.g. 3 second map refresh 80% of the time
- Choosing a KPI
 - Balance cost against business need
 - Specifying ½ sec. map refresh 99.99999% of the time may result in significantly greater project cost
 - True business need or costly nice-to-have?



Test Plan: Test Parameters

- Bandwidth
 - 8 Mbps WAN
 - 100 Mbps LAN
- ArcSDE Connection Type
 - **SDE Remote Connect**
 - **■** Direct Connect
- User Activity Type
 - Viewers
 - Editors
 - "Analysis" (raster loading / buffering)
- Number of Users (1, 5, 25, 50)

Test Plan: Levels

- Level I: Batch Data Loading
 - Test raster and vector data load times
- Level II = Level I + Client Response Times
 - Test client response times w/ArcMap
 - Test Oracle database server load
- Level III = Level II + ArcIMS
 - ArcIMS connections as proxy for concurrent ArcMap user sessions
- **Level IV: Thin Client Solutions**
 - ArcMap via Citrix and ArcGIS Server

Test Plan: Tests & Sequencing

Defined tests with IDs. Parameters being tested. Duration, etc.

Sess ID	TestID	Mbps	Con		GDBT Users	GDBT Bw Req'd (Mbps)	Max GDBT users @ Mbps	WL Users	WL Agda	A-Script	Rast Loads	Date	Dur	Start	End	Purpose	GIFS	IMS	MAX IMS CPU
								10, no	800x 500							WebLoad + 1			
								think								Interactive			
4	4001	8	DC	test4	0	0.0	16	time	Zm2x	NONE	0	Dec05	5 mins			WebLoad			
4	4003	8	DC	test4	1	0.5	16	0	NONE	Rast Cat: Duplin (PC07)	1	Dec05	5 mins			1 Interactive GDBT + 1 raster load	N/A	N/A	N/A
4	4004	8	DC	test4	5	2.5	16	10, no think time	Wake	NONE	0	Dec05	5 mins			+ WL. No raster loads.			
4	4006	8	DC	test4	10	5.0	16	10, no think time	Wake	NONE	0	Dec05	5 mins			All GDBTs + WL. No raster loads.			
4	4008	8	DC	test4	25	12.5	16	10, no think time	Wake	NONE	0	Dec05	5 mins			All GDBTs + WL. No raster loads.			
4	4009	8	DC	test4	50	25.0	16	10, no think time		NONE	0	Dec05	5 mins			All GDBTs + WL. No raster loads.			

Test Environment: Hardware

Server Room

- "Standard issue" servers
- Sufficient to meet anticipated demand?
- Test Lab (8 clients)
 - (7) "Standard issue" client PCs
 - mix of CAD and GIS
 - (1) legacy laptop
 - 6 PCs shared 3 monitors
 - 1 PC with dedicated monitor



Test Environment: Software

- ArcMap 9.2
- **■** ArcSDE 9.2
- Oracle 10gR2
- ArcIMS 9.2
- ArcGIS Server 9.2
- Citrix PresentationServer
- RedHat Enterprise Linux (Oracle)
- Windows 2003 (other servers)



Test Environment: Software > Test Tools

Commercial Tools:

- Mercury LoadRunner (up to 100K) \$\$\$
- Visual Studio Team Edition (~20K) \$\$
- ESRI's Enterprise Test Harness requires ESRI professional services contract (\$?)

Pros:

Provide the most functionality

Cons:

- Steep learning curve
- Cost of tool more than entire test budget?

Test Environment: Software > Test Tools

Free tools:

- Geodatabase Performance Toolset (GDBT)
 - Generate load against ArcSDE via ArcMap
 - Collect client response times

■ WebLoad

- Generate load against ArcIMS (or any web app)
- Developed by a founder of Mercury Interactive

OpenSTA

- Not selected for this project
- DummyNet
 - Bandwidth Impairment
 - Runs on FreeBSD (Unix)

Geodatabase Performance Toolset (GDBT)

- ArcGIS Desktop extension
- In ArcMap:
 - Appears as a Toolbar
 - Simulate ArcMap user activity
 - Monitor client response time
- In ArcCatalog:
 - Appears as a Tab
 - Visualize state tree
 - Monitor edits in delta tables
 - Tweak spatial indexes
 - Etc.

GDBT: Pros and Cons

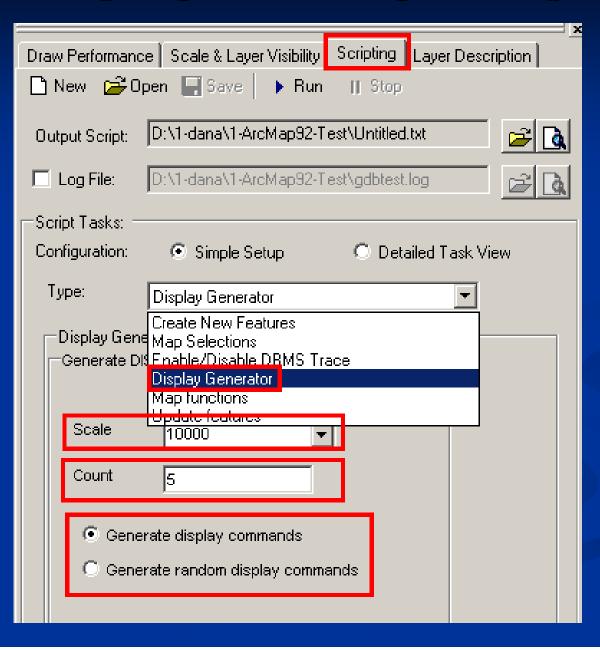
PROS:

- Reliable map display generator
 - Simulate viewers
- Reliable UpdateFeatures function:
 - Simulate editors
- Being updated for a future release
- Free of cost

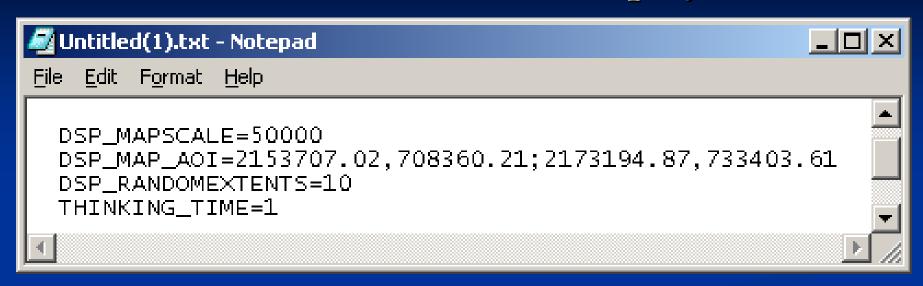
CONS:

- No official support
- Bugs
- Some functions don't work
 - Create New Features
- One instance per PC
- Can't set test durations
 - # of operations only

GDBT Scripting: Simulating ArcMap Viewers



GDBT: 10 Random Displays



GDBT: 10 Fixed Displays

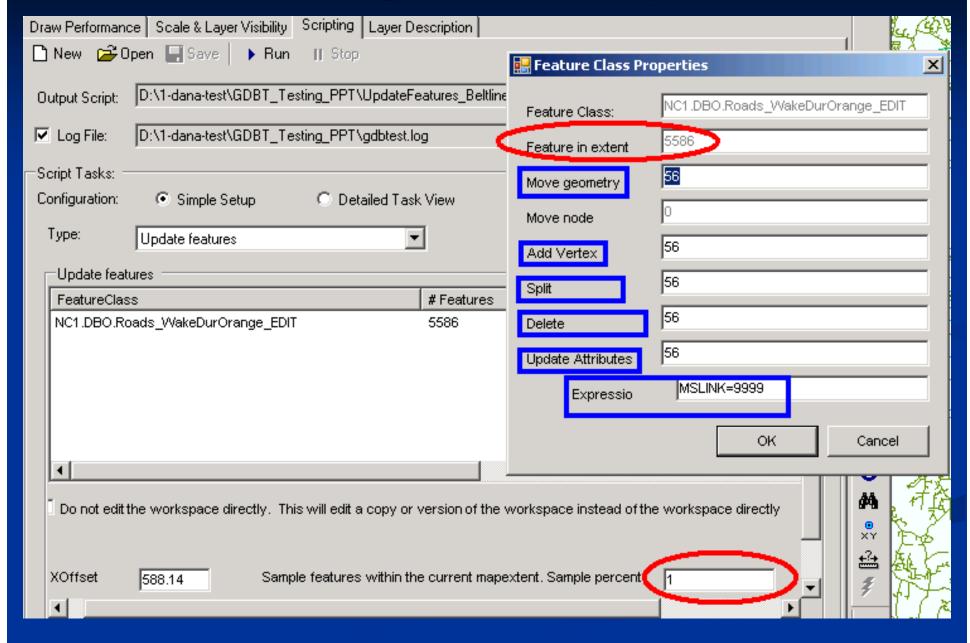
```
File Edit Format Help

DSP_DISPLAY=2151230.7672,720705.0134;2170718.6145,745748.4161
DSP_DISPLAY=2146808.9936,720366.5071;2166296.8409,745409.9099
DSP_DISPLAY=2157554.2826,703253.6718;2177042.1298,728297.0746
DSP_DISPLAY=2151647.0795,714364.6523;2171134.9268,739408.0551
DSP_DISPLAY=2150555.8043,713390.0635;2170043.6516,738433.4663
DSP_DISPLAY=2157475.4517,700729.8064;2176963.2989,725773.2092
DSP_DISPLAY=2149723.1797,701027.3827;2169211.0270,726070.7855
DSP_DISPLAY=2149981.8303,714682.6936;2169469.6776,739726.0964
DSP_DISPLAY=2155056.4088,716252.4352;2174544.2560,741295.8380
DSP_DISPLAY=2162966.3425,695959.1869;2182454.1897,721002.5897
```

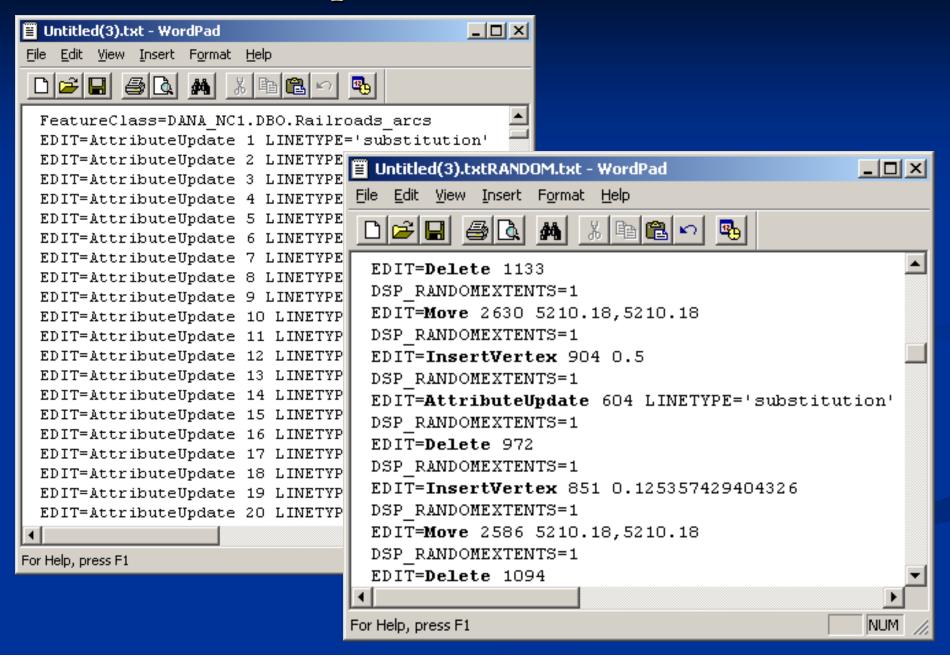
Fixed Displays unrealistic due to database caching?

Mostly used Random Displays

GDBT: Simulating Editors with Update Features



GDBT: Update Features limitations



GDBT Assumptions: "Think Time"

Total Concurrent Users	Users Simulated per PC	Total DPM Require d per PC*	Think Time in seconds**
5	1	6	10
10	2	12	5
25	5	30	2
50	10	60	1

GDBT Script Assumptions Table (6 DPM User Productivity Rate)
* Number of users simulated per PC multiplied by 6 DPM

** Total DPM required per PC divided by 60 seconds (1 minute)

Assumptions: "Think Time"

PC1: DPM: 6 TT: 10.0 # users: 1 PC2: DPM: 6 TT: 10.0 # users: 1 PC3: DPM: 6 TT: 10.0 # users: 1 PC4 DPM: 6 TT: 10.0 # users: 1 PC5: DPM: 6 TT: 10.0 # users: 1

= 5 concurrent users

PC1: DPM: 12 TT: 5.0 # users: 2 PC2: DPM: 12 TT: 5.0 # users: 2 PC3: DPM: 12 TT: 5.0 # users: 2 PC4: DPM: 12 TT: 5.0 # users: 2 PC5: DPM: 12 TT: 5.0 # users: 2

= 10 concurrent users

PC1: DPM: 60 TT: 1.0 # users: 10 PC2: DPM: 60 TT: 1.0 # users: 10

PC3: DPM: 60 TT: 1.0 # users: 10 PC4 DPM: 60 TT: 1.0 # users: 10 PC5: DPM: 60 TT: 1.0 # users: 10

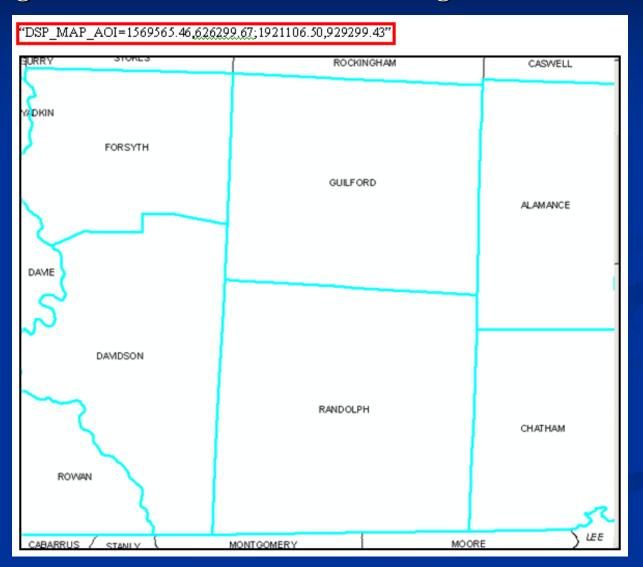
= 50 concurrent users

Other Assumptions

- **Editor to Viewer ratio:** 1:5 (20%/80%)
- Map Scale for Displays:
 - Vector 10K and 100K
 - Raster 2K and 5K
- Symbology / layers / scale-dependencies:
 - Used standard technician's MXD in GIS Unit
- **■** User Levels:
 - Low = 5 users
 - Moderate = 10 users
 - High = 25 users
 - Peak = 50 users

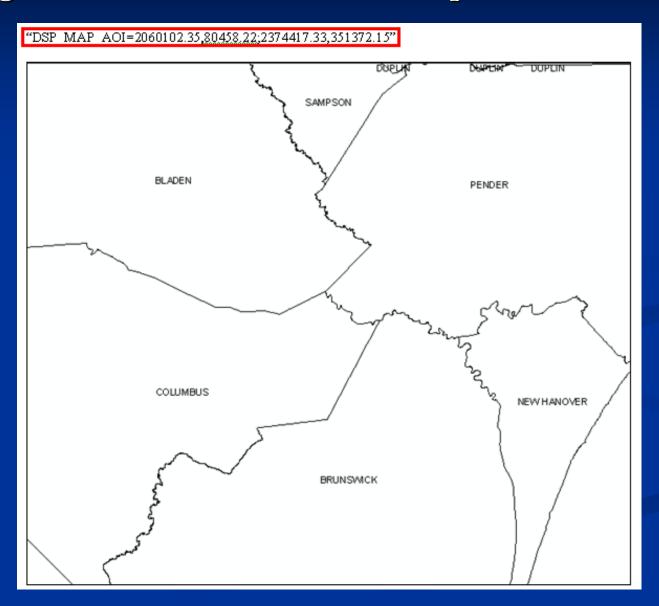
Area of Interest: Vector (~7)

- AOIs with little "dead space" due to scripting automation
- Contiguous selection of counties vs. loading all data for all counties



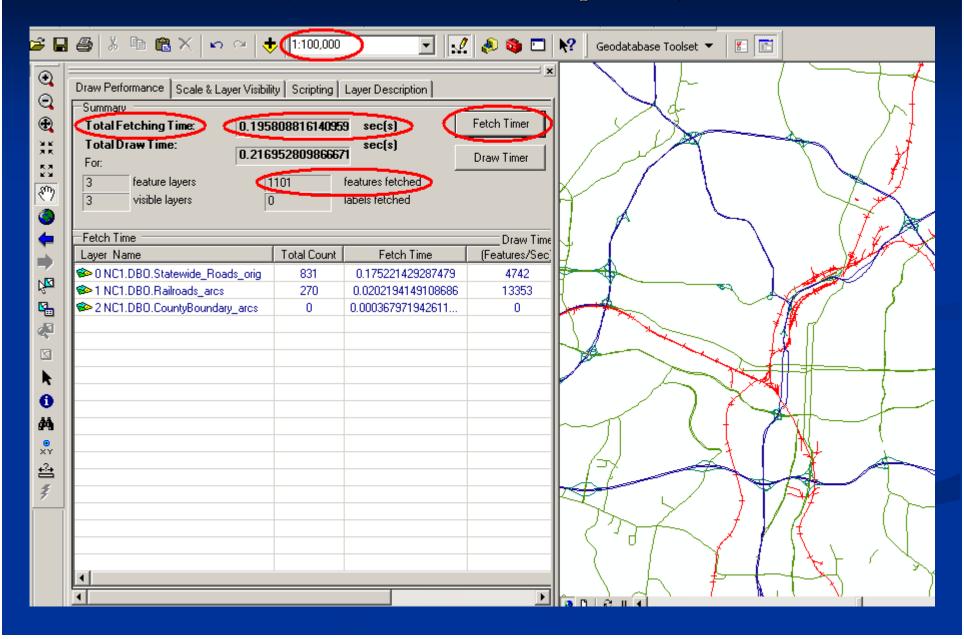
Areas of Interest: Raster (~6)

Plugged text for AOIs into GDBT scripts

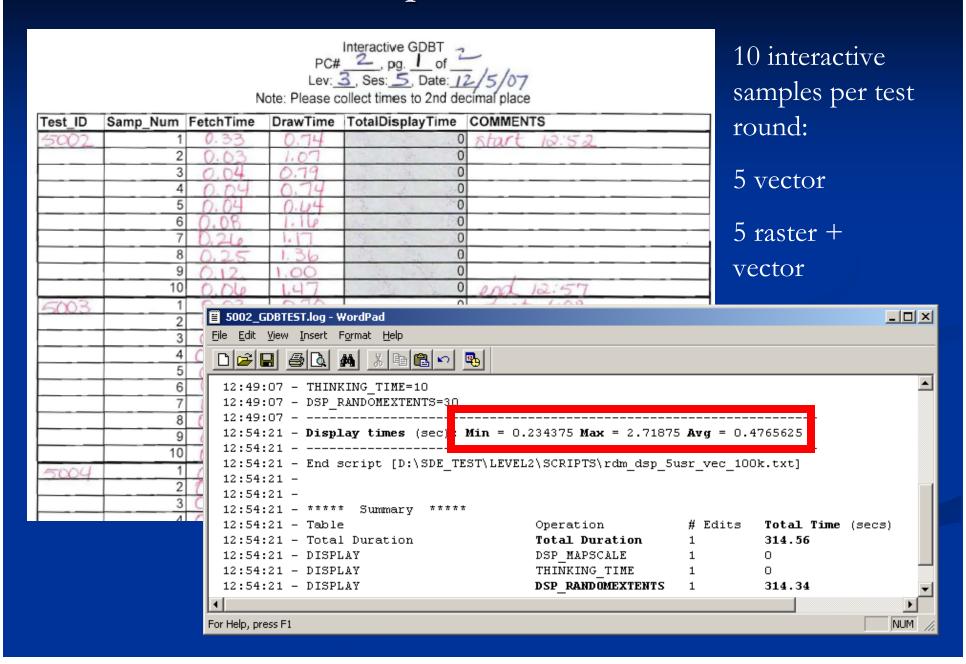


GDBT: Client Response Time Monitoring

(A window into the user experience)



GDBT: Client Response Time Data Collection



WebLoad

Free Open Source Software Tool (FOSS)



Created by a founder of Mercury Interactive



- Load Testing, Functional Testing, and Monitoring
- Used to generate load against ArcSDE via ArcIMS site
- One machine can simulate hundreds of virtual users
 - Similar to ESRI's Enterprise Test Harness?

WebLoad: Components

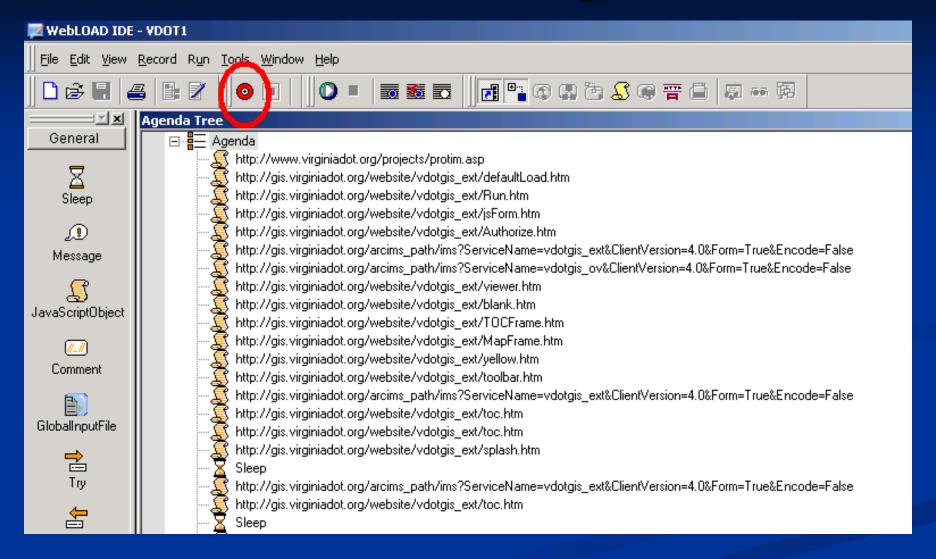
WebLoad IDE

- Record scripts ("agendas") using a web browser
- Stored as JavaScript (tweakable)

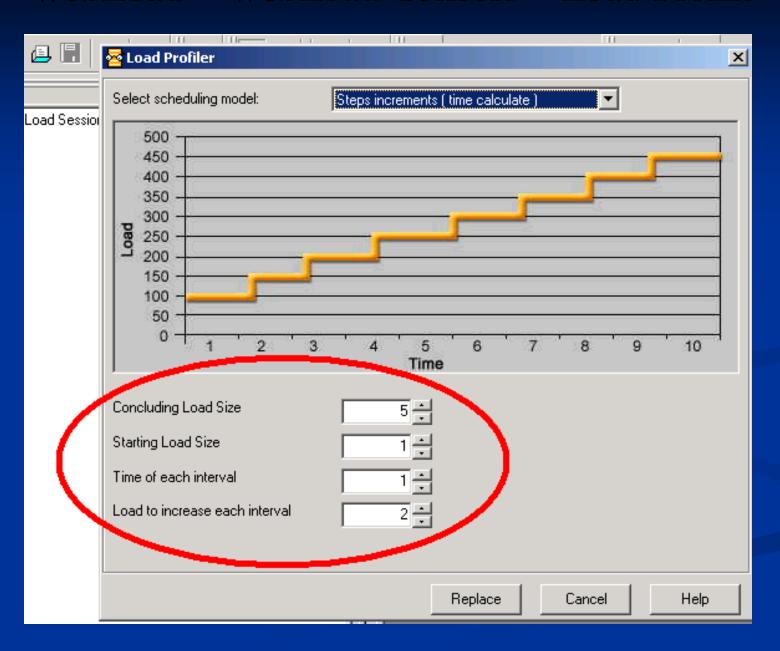
WedLoad Console

■ Run scripts as tests

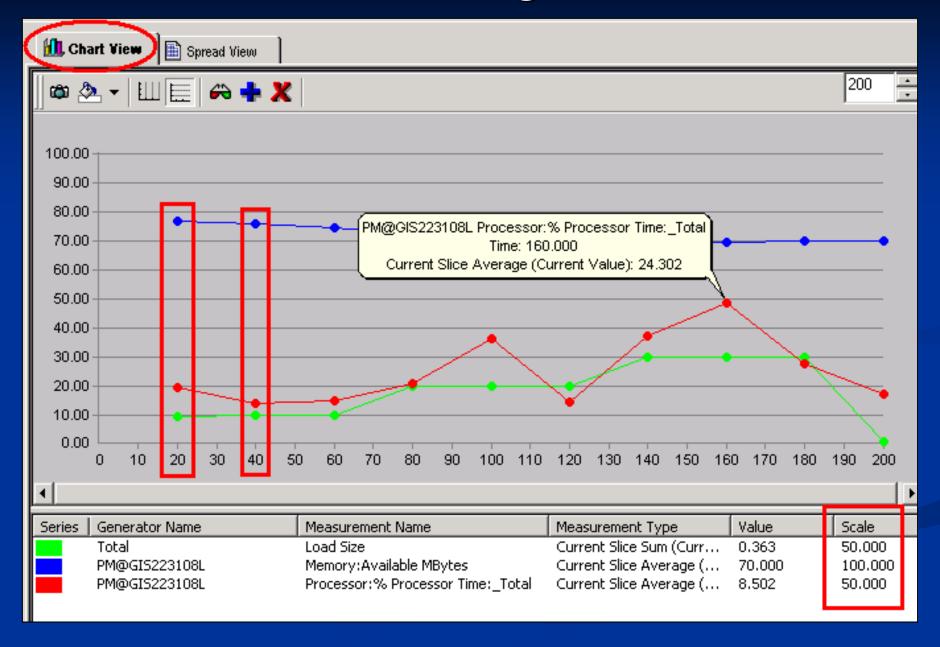
WebLoad IDE: Agenda



WebLoad > WebLoad Console > Load Profile



WebLoad > Monitoring > View - Chart



WebLoad: Assumptions Influence Results

- Tested against NCDOT ArcIMS App (TIMS)
- 800 x 500 GIF
- No imagery
- ArcMap Image Service
- AOI + Map Scale chosen for higher density of features
- 10 pans / ran in a loop
- User Productivity 6 DPM
- Any one of these choices could influence performance / capacity

WebLoad: Image Dimension vs. Scale vs. AOI

First Agenda:

- 400 x 250
- Zoomed in tight
- Johnston County
- File size: ~9K



Final Agenda:

- 800 x 500
- Zoomed out
- Wake County Beltline
- File size: ~98K



WebLoad: ArcIMS Capacity Calculations

(Total # of images / duration of test in mins) / 6 DPM

= effective concurrent users

Example:

- 1,000 GIFs produced
- 5 minute test run
- 6 Displays Per Minute user productivity
- 1000 GIFS / 5 mins = 200 total Displays Per Minute
- 200 total DPM / 6 DPM user productivity = ~ **33** effective concurrent users

WebLoad: Our Results

Utilizing:

- ArcIMS "spatial servers" running on a single server
- Webserver (IIS) running as a virtual server on different machine (NCDOT standard practice)
- Servlet engine = Servlet Exec

Can support (within chosen KPI):

- ~40 users per server (6 Displays Per Minute)
- ~23 users per server (10 Displays Per Minute)

For more users:

- Scale horizontally (add more servers)
- Distribute "spatial server" processes to these servers

Citrix Presentation Server

- Low bandwidth solution
 - ~8 Mbps shared WAN from GIS Unit to Server Building
- Publish ArcGIS Desktop to a powerful, centralized server
- More efficient than a web page
 - Thin client / ICA Protocol
 - Sends only mouse clicks, keyboard strokes, and portions of a screen that change
 - Run ArcMap over 56K dial-up connection in emergency situations
- Stability over a WAN
 - Many dropped packets vs. LAN
 - Lose connection and reattach to ArcMap session still running on Citrix



GIS Users per Citrix Server?

- Memory required per GIS user?
 - **■** Economics
 - ArcGIS Desktop on standalone PC = ~2G recommended
- What is a typical GIS user?

```
ArcMap session +
```

ArcCatalog session (sometimes)

= (1) **GIS** User

- Limitations:
 - Unable to install GDBT (registry issues)
 - Used VBA script to cycle through features

Memory per GIS User on Citrix

		Server
Product	Action	RAM (MB)
	Raster MXD	
	with browse	
ArcMap 9.2	script running	120
	Previewing	
	raster data with	
	ArcToolbox	
ArcCatalog 9.2	enabled	110
One ArcGIS Desktop User		
(ArcMap + ArcCat)	N/A	230

Table 12: Memory Utilization for One ArcGIS Desktop 9.2 User

■ Bottom line:

- ~15-20 users per server (4G RAM/quad-core)
- To support more users per server, add more RAM
- Isolate analysis users on separate Citrix server

ArcGIS Server: More Thin Client Solutions

- Intention to generate load using:
 - ArcGlobe service via ArcGIS Explorer
 - Published MXD service via Internet Explorer
- Unable to generate load against ArcGIS Server with WebLoad
 - Web 2.0 complexity
 - May be possible using OpenSTA

ArcGIS Server: Manual Testing

- Only 3 testers in lab during this session
- Manual browsing of ArcGlobe Service
- Manual browsing of Published Map Service
- Observation: Published Map Service used about 2X more CPU on database server
 - ArcGlobe service caching on both client and service

Output Products and Benefits

- Detailed project report
 - Contains technical recommendations for ArcSDE implementation
- Build documents
 - ArcSDE, ArcGIS Server, ArcIMS, Citrix
- Revised Test Plan
 - May be used as a template for future projects
- Strengthened Relationships with IT
- Educated IT Staff

Test Results: Highlights

- Direct Connect faster than SDE Remote Connect by ~1/3 (36%)
- Isolate Analysis GIS users on separate Citrix server
 - Three-ring buffer operation consumed 25% of server's CPU resources & consumes more RAM
- Standard NCDOT servers sufficient to meet anticipated need
- Proved inadequacy of 8 Mbps WAN connection for Enterprise GIS

Further Work

Output from prototype = input to next project (implementing ArcSDE)

Need to research:

- Versioned work flows
 - Performance will be different
- Automate load against ArcGIS Server services
- Tuning
 - Spatial indexes
 - Oracle
 - ArcIMS / ArcGIS Server
- Citrix
 - Plotting, getting data on/off servers, etc

Lessons Learned

- Testing isn't easy
 - Tools have deficiencies
 - Learning how to test is part of the process
- Making test assumptions isn't easy
 - "Average" map display (scale, num + type of layers, etc)
- Testing can be a highly iterative process
 - Test Plan evolved as we went along
- Testing will save time and money in production
 - ArcSDE no longer a mystery / fear factor to IT
 - Established specific needs beyond sizing charts

Resources

- DummyNet (Bandwidth Impairment):
 - http://www.dummynet.com/
- Geodatabase Performance Toolset (GDBT):
 - Google for "GDBT"
- WebLoad (ArcIMS Load Testing):
 - http://www.webLoad.org
- OpenSTA (ArcGIS Server Load Testing?):
 - http://www.openSTA.org
- ESRI's System Design Strategies White Paper
 - http://www.esri.com/library/whitepapers/pdfs/sysdesig.pdf
- ESRI Support Forums
 - http://support.esri.com

Thank You. Questions?

Contacts:

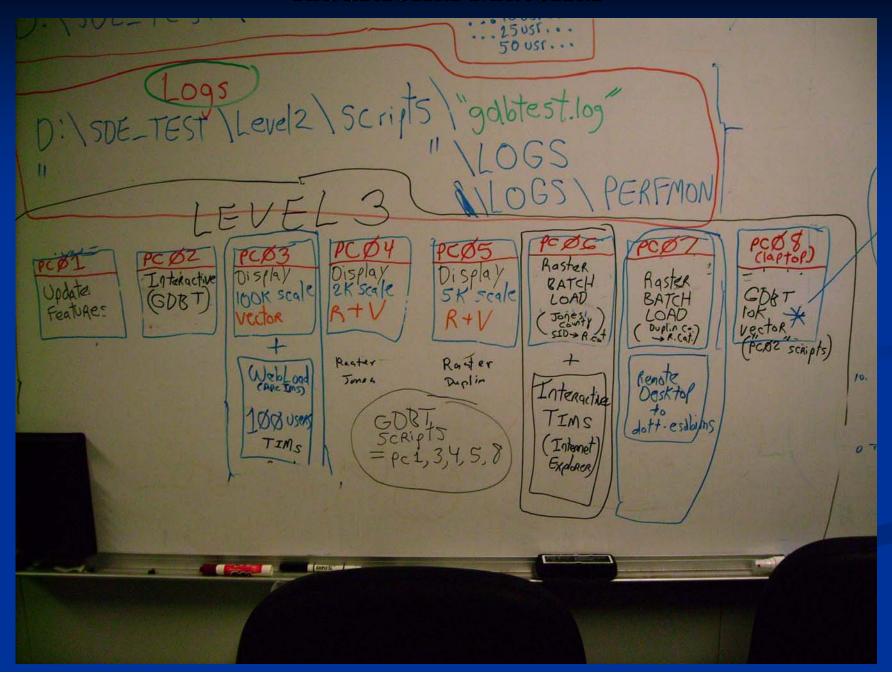
Dana Nibby

dmnibby@dot.state.nc.us

Dave Johnson

ddjohnson@dot.state.nc.us

Additional Material



Additional Material

HP DL 380 G5

Dual-Core Intel Xeon 5150
Processor (2.66 GHz, 1333 FSB)
4MB L2 Cache
Smart Array P400 Controller with
512MB write cache
2 Embedded NC373i Gigabit
Network Adapters
Redundant power supply
Hot plug 2.5" SAS Drives
4 GB of PC2-5300 DDR2 memory

Itanium 64

Dual-core Itanium® 2 Level 1 cache: 32 KB Level 2 cache: 1 MB Level 3 cache: 12 MB Smokin' Fast

