

Managing your Domino Clusters

Kathleen McGivney

President and chief technologist, Sakura Consulting

www.sakuraconsulting.com

Paul Mooney

Senior Technical Architect, Bluewave Technology

www.bluewave.ie



Who the hell are these two short people

- Kathleen McGivney
 - ▶ President and chief technologist, Sakura Consulting
 - ▶ www.sakuraconsulting.com

- Paul Mooney
 - ▶ Senior Technical Architect, Bluewave Technology
 - ▶ www.bluewave.ie

Clustering Is ...

- Event-driven replication
 - ▶ Caused by a change in a document
 - ▶ “Pushed” to other cluster members

- Two specific tasks
 - ▶ Cluster Replicator (CLREPL)
 - Responsible for the replication push
 - ▶ Cluster Database Directory Manager task (CLDBDIR)
 - Responsible for maintaining the list of databases to include in cluster replication (clbdir.nsf)

Some Facts About Clustering

- Replication formula
 - ▶ Is IGNORED by clustering
 - ▶ All data is replicated between servers in a cluster
 - Even if the replication formula is configured against this
 - ▶ Data will be removed by standard replication
- Deletion stubs
 - ▶ Do not replicate via cluster replication
 - Have a standard replication connection document running on schedule to counteract this!

Types of Domino Clusters

□ Cluster categories

▶ Active-Active

- All cluster members actively provide services
- Most commonly used configuration
 - Example: Server A is primary mail server; Server B is primary application server
 - Users access both servers regularly

▶ Active-Passive

- One or more cluster members are idle until triggered by a failover or load-balancing event
 - Example: Server A and Server B have replicas of all databases, but users access only Server A, unless it is unavailable

Clustering for Additional Services

- Clustering for disaster recovery
 - ▶ Cluster over a WAN to provide disaster recovery for sites
 - ▶ Network infrastructure must support this!
 - Should have speeds comparable to LAN
 - ▶ Best use of Active-Passive clusters

- Clustering for backups
 - ▶ Perform backups on one cluster member; leave other cluster members up and available

Real-World Example!

- Company has four servers in three locations
- Wants three primary servers; one server as cluster mate
 - ▶ But servers can only be a member of one cluster at a time!
- Solution: Selective database distribution
 - ▶ Servers A, B, and C have their own replicas only
 - ▶ Server D has replicas of all databases
 - ▶ Users will only fail over to server D, except for system databases (Domino Directory, Catalog, etc.)

Hardware Considerations for Clustered Servers

- Memory, processor, hard drive, and bandwidth requirements
 - ▶ Servers will require additional memory and CPU cycles to handle cluster tasks
 - Servers must also be able to handle the increased workload that will occur in cases of failover
 - ▶ Understand the effects of clustering on disk I/O
 - Each server manages disk I/O for its own databases and for the cluster replicas it hosts
 - Place databases on a physical disk with low resource demand
 - Don't place databases on the same physical disk as the OS swap files

Hardware Considerations for Clustered Servers (cont.)

- Best practices for distribution of program files, data files, and OS files
 - ▶ Separate physical disks
 - Place OS, Domino program files, Domino databases, and Domino transaction log on separate physical disks
 - ▶ Use RAID arrays for additional reliability
 - ▶ At the very least, keep OS swap files and Domino database files on separate physical drives

Manually Triggering Failover

- Set the server to the restricted state
 - ▶ Server_Restricted Notes.ini variable
 - Value of 0 = unrestricted
 - Value of 1 or 2 = restricted
 - Value of 1 will reset to 0 when the server is rebooted; value of 2 is sticky and will remain until it is manually reset
- Set value with console command
 - ▶ Set config Server_Restricted = n
- Use for troubleshooting or maintenance
 - ▶ Redirect users to other cluster members while you work
 - ▶ Perform server upgrades during business hours

Triggering Failover

- Set the maximum number of concurrent NRPC users allowed to connect to a server
 - ▶ Server_MaxUsers Notes.ini variable
 - Set variable to a number determined in planning stage
- Set variable using console command
 - ▶ Or use Notes.ini tab in server configuration document
 - ▶ Set config Server_MaxUsers = desired maximum number of active concurrent users
 - Don't confuse this with Server_MaxSessions, which restricts server sessions, too!

Logging and Monitoring Failover

- Check for failover events using statistics parameters
 - ▶ When failover occurs, Domino logs a failover event in the server's log file
- Information returned by the Show Stat command
 - ▶ Number of times server has redirected a client to another cluster member
 - ▶ Number of times a client attempted to open an out-of-service database
 - ▶ Number of times a client attempted to open a database when the server was in the MaxUsers or Restricted state

Setting Up Workload Balancing

- Triggering workload balancing
 - ▶ Server Availability Index (SAI)
 - Clustered servers determine their own workload based on average response time for client requests
 - Index from 0-100, 0 indicating a heavily loaded server and 100 indicating a lightly loaded server
 - Example: Index of 75 indicates that 75% of system resources are still available (sort of ...)
 - ▶ Server availability threshold
 - Decide when a server will enter the busy state
 - Set using server console command
 - Set config `Server_Availability_Threshold = n`

Logging and Monitoring Workload Balancing

- View events using statistics parameters
 - ▶ When a load balancing event occurs, Domino logs an event in the server's log file
- Information returned by the Show Stat command
 - ▶ Number of times server was in the busy state or was unable to redirect a client to another cluster member
 - ▶ Number of times a client attempted to open a database when the server was in the busy state
 - ▶ Number of times a client attempted to open a database when all servers in the cluster were in the busy state

Failover and Workload Balancing Together

- Optimize failover and load balancing
 - ▶ Force a secondary server to enter the busy state after the primary server comes back up
- Use clustering features to manage server migrations
 - ▶ Add new server to cluster; set old server to restricted state, forcing users to new server
- Use clustering features to force users to failover during planned server outages
 - ▶ Set server to restricted state, allowing you to keep it up and running while you perform maintenance, while users are failed over to another cluster member

Using Cluster Commands on a Console

- Special “documented” cluster commands
 - ▶ The server does not have to be in a cluster
 - ▶ Enabled by the following console command
 - Set Config CLUSTER_ADMIN_ON=1

Copy a Database Using the Console

- ❑ You can copy a database from one server to another using the console
- ❑ Type the following:
 - ▶ `CL copy serverA!!db1.nsf serverB!!db2.nsf`
- ❑ You must have cluster commands enabled

Create Replica Using Console

- ❑ You can create a replica of a database from one server to another
- ❑ Type the following:
 - ▶ `CL copy serverA!!db1.nsf serverB!!db2.nsf REPLICA`
- ❑ You must have cluster commands enabled

Create Template Copy Using Console

- ❑ You can create a template copy of a database from one server to another
- ❑ Type the following:
 - ▶ `CL copy serverA!!db1.nsf serverB!!db2.nsf TEMPLATE`
- ❑ You must have cluster commands enabled

Create Copy on Same Server

- ❑ You can create a copy of a database on the same server
- ❑ Type the following:
 - ▶ `CL copy db1.nsf db2.nsf`
- ❑ You must have cluster commands enabled

Server Statistics Monitoring

- Overall server health is important
 - ▶ Monitor all server statistics, not just cluster statistics
 - ▶ Clustering allows you to bring down ailing servers and troubleshoot without affecting users
- Use Domino's native monitoring tools
 - ▶ Domino Domain Monitoring is great for this

Important Cluster Statistics

- **Server.AvailabilityIndex**
 - ▶ Indicates the current percentage of a server's availability
 - 0 indicates a server in the busy state
 - 100 indicates a completely available server

- **Server.AvailabilityThreshold**
 - ▶ Indicates the threshold where the server will enter the busy state
 - ▶ Set by the administrator using the Notes.ini variable `Server_Availability_Threshold`

Important Cluster Statistics (cont.)

- `Replica.Cluster.WorkQueueDepth.xx`
 - ▶ Measures how many databases are waiting in the cluster work queue to be replicated
 - High numbers and high averages can indicate a cluster replication problem — could be network or disk bottleneck

- `Replica.Cluster.SecondsOnQueue.xx`
 - ▶ Measures how many seconds replication events are waiting to replicate with other cluster members
 - High numbers and high averages can indicate a cluster replication problem — especially if the work queue depth is also high
 - Check OS stats to determine where the bottleneck is

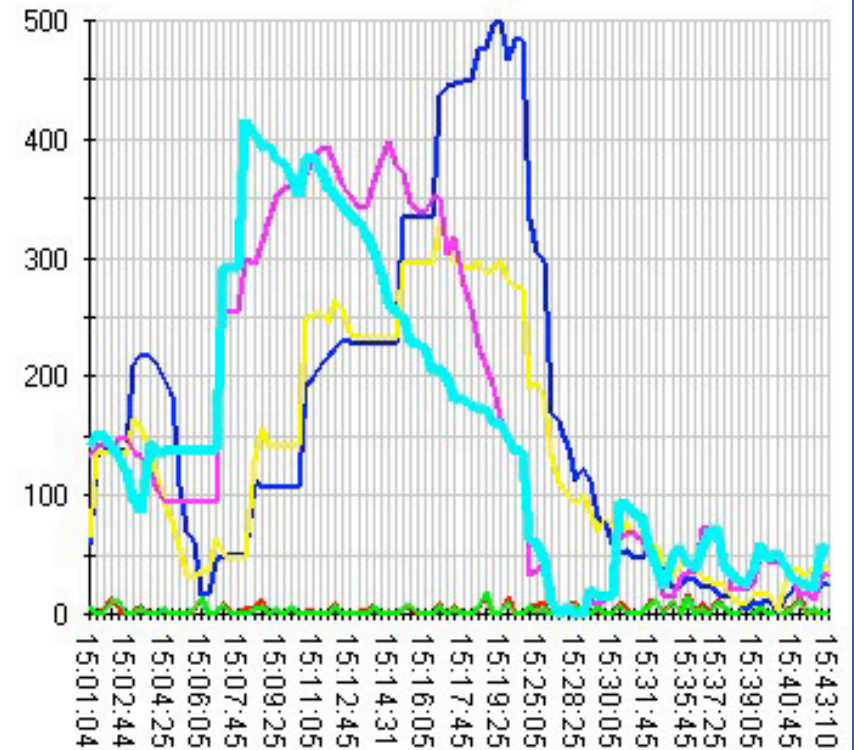
Real-World Scenario!

- Customer experiencing cluster-related performance issues
 - ▶ Databases not staying in sync
 - ▶ Users complaining about documents “disappearing”
- We used native Domino monitoring tools to view stats
 - ▶ Real-time statistics graph on Server – Performance tab
- Let’s take a look at the results

Example of a Good Cluster Gone Bad!

- ❑ Server gets swamped with cluster requests ...
- ❑ Cannot keep up

❑



Important Cluster Statistics (cont.)

- **Server.Cluster.xx**
 - ▶ **OpenRequest.ClusterBusy**
 - Indicates the number of times a client attempted to open a database when the server was in a busy state
 - If this number is high, you may need to redistribute users and/or databases to another cluster member or increase the server availability threshold

Important Cluster Statistics (cont.)

- `Server.Cluster.OpenRedirects.xx`
 - ▶ `Failover.Unsuccessful`
 - Indicates the number of times this server could not redirect a client request to another cluster member when the database being requested was unavailable
 - ▶ `LoadBalance.Unsuccessful`
 - Indicates the number of times this server could not redirect a client request to another cluster member when this server was in the busy state

The Server Availability Index (SAI)

- SAI is calculated to display a number relative to server availability and performance
 - ▶ 100 indicates lightly loaded server; 0 indicates fully loaded server
- Type in `Sh Ser` on the console to see your servers' availability index (between 1-100)
- SAI calculation changed beginning with the R6 codestream
 - ▶ Default configuration may show artificially low SAI
- This can be adjusted!
 - ▶ Modify the expansion factor (explained later)

How Does It Calculate the SAI?

- First, understand the expansion factor ...
 - ▶ Calculated based on response times for recent requests
 - ▶ Compares recent response time to minimum response time that the server has completed
 - The difference is called the expansion factor
 - That is, how much the delay in opening has expanded
 - Example: Server currently averages 12 ms for DBOpen requests; minimum time was 4 ms
 - Expansion factor = 3 (average current time/fastest time)
 - ▶ This is averaged over different types of transactions
 - Fastest time is stored in memory and in Loadmon.ncf
 - Loadmon.ncf is read each time the server starts

The Expansion Factor

- Expansion factor calculation
 - ▶ Domino tracks the most commonly used transactions
 - By default, Domino tracks transactions for five periods of 15 seconds each
 - ▶ Each type of transaction is averaged and then divided by the fastest time to complete that transaction type
 - ▶ The expansion factor for the entire server is averaged across all transactions
 - All transactions are weighted evenly

The Expansion Factor (cont.)

- How it affects the availability index
 - ▶ Adding load to a busy server increases the expansion factor faster than adding load to a less busy server
 - ▶ Hardware capacity affects the expansion factor
 - Slow servers can have an expansion factor of 30, indicating slow response times
 - NOTE: While fast servers can have fast response times even with an expansion factor of 300
 - Just because it takes 10 times longer than usual to complete something, does not mean it's slow!
 - ▶ Domino uses a formula to convert the expansion factor into the availability index

Default Expansion Factor Table

| Expansion Factor | Availability Index |
|------------------|--------------------|
| 1 | 100 |
| 2 | 83 |
| 4 | 67 |
| 8 | 50 |
| 16 | 33 |
| 32 | 17 |
| 64 | 0 |

- So ... when the expansion factor hits 64 ...
 - ♦ Availability is 0
- This is not good enough for fast servers!

Changing the Expansion Factor

- Modifying the value that indicates a loaded server
 - ▶ Default value is 64
 - ▶ Use the following Notes.ini variable to change it:
 - `SERVER_TRANSINFO_RANGE=n`
 - ▶ To determine the optimal value for this variable:
 - Monitor the expansion factor on the server during a period of heavy usage
 - Use `show stat server.expansionfactor`
 - Check other performance stats while you do this
 - ▶ The value of *n* should be such that:
 - 2 raised to the power of *n* = optimal expansion factor

Changing the Expansion Factor (cont.)

- HUH?
 - ▶ In English please!
- Currently, the default line is:
 - ▶ `SERVER_TRANSINFO_RANGE=6`
 - ▶ 2 raised to the power of 6 = 64
- Watch your `sh stat server.expansion.factor` result
 - ▶ During heavy usage
 - ▶ Determine yourself when your server is really busy
 - ▶ Call this value "n"
 - ▶ Calculate 2 raised to the power of $X = n$

Updated Expansion Factor Table

| Expansion Factor | Availability Index |
|-------------------------|---------------------------|
| 1 | 100 |
| 2 | 88 |
| 4 | 75 |
| 8 | 63 |
| 16 | 50 |
| 32 | 38 |
| 64 | 25 |
| 128 | 13 |
| 256 | 0 |

Changing Data Collection Intervals

- By default, Domino tracks transactions for five periods of 15 seconds each
 - ▶ Change number of data collection periods:
 - `Server_Transinfo_Max = x`
 - `x` = the number of collection periods you want Domino to use
 - ▶ Change the length of each collection period:
 - `Server_Transinfo_Update_Interval = x`
 - `x` = the length of each period in seconds

Viewing Expansion Factor Statistics

- What statistics are used to calculate the expansion factor?
 - ▶ Set config debug_loadmon=1
 - ▶ Show stat server.loadmon*
 - Shows minimum and average run times for all measured transactions that are used to calculate the expansion factor

-

That was the hard way.....

- Ok....
 - ▶ Here is the easy way (702 or higher)
 - ▶ Type in SH AI on the console

- The server will tell you what to set the expansion factor to
 - ▶ Do this when the server is at its busiest
 - ▶ Check this variable from time to time

- Note.. If you change the trans info range
 - ▶ You need to keep changing that with service related updates
 - Upgrades
 - New Hardware
 - New tasks

-

Cluster-Related Notes.ini Variables

- **Server_Availability_Threshold**
 - ▶ Balances workload across servers
 - Directly related to the Server.AvailabilityIndex statistic
 - If the availability index is below the threshold, the server will enter a busy state and new user sessions will be redirected to another cluster member
 - 0 (default) indicates a fully available server, with workload balancing disabled; 100 shows a busy state
 - Example: If your availability index hovers at 87 during peak usage, set the threshold to 80 or 75

Cluster-Related Notes.ini Variables (cont.)

□ Server_Restricted

▶ Restricts access to the server

– A setting of 0 is unrestricted

- If set to 1, the server will return to an unrestricted state when rebooted
- If set to 2, it will remain restricted until it is manually reset

– Set to 1 if you want to deny users access to the server temporarily while troubleshooting

– Set to 2 if you want to deny access to users through several server reboots

- If set to 2, don't forget to reset to 0 to allow access!

Cluster-Related Notes.ini Variables (cont.)

- **Server_MaxUsers**
 - ▶ Restricts the number of active users allowed on a server
 - ▶ If set to 0, the number of users is unlimited

- **Server_MaxSessions**
 - ▶ Maximum number of concurrent sessions allowed on the server
 - This includes server sessions, whereas Server_MaxUsers restricts only user sessions

Cluster-Related Notes.ini Variables (cont.)

- Cluster_Replicators
 - ▶ Number of cluster replicator tasks
 - Use this setting to start multiple cluster replicator tasks
- Disable_Cluster_Replicator
 - ▶ Disables cluster replicator tasks
 - 1 disables cluster replicator tasks; 0 (default) enables
 - So ... you can have failover without cluster traffic!
- RTR_Logging
 - ▶ Monitors cluster replicator activity
 - 0 disables monitoring; 1 enables

Cluster-Related Notes.ini Variables (cont.)

- `Server_Cluster_Default_Port`
 - ▶ Port for intra-cluster network traffic
- `Server_Cluster_Probe_Port`
 - ▶ Port used for cluster availability probes

Server Optimization — Network options...

- Network considerations
 - ▶ A private LAN segment for intra-cluster network traffic is strongly recommended
 - Providing a dedicated LAN segment helps prevent network bottlenecks
 - ▶ Consider using a secondary NIC for intra-cluster traffic
 - This also provides redundancy — if the secondary NIC for private LAN traffic fails, the cluster traffic can be rerouted to the main LAN

Server Optimization — Hardware

- Hardware considerations
 - ▶ Adding memory
 - As users and transactions increase, memory demands are also increased
 - ▶ Changing physical disk distribution
 - If disk write time seems slow, try separating OS swap files, Domino program files, and Domino data files
 - ▶ Separate physical disks — can't be said enough!

How Clustering Affects HW Performance

- Disk Input/Output
 - ▶ Spreading data across physical disks is a good idea
 - Each clustered server manages disk I/O for its own databases and the replicas of other cluster members' databases
 - Databases should be distributed across physical disks where there is little contention for disk I/O
 - ▶ This is especially important with Domino on a SAN
 - Work with SAN configuration team so that Domino is not contending for physical disk with other high I/O applications

Server Optimization — Transaction Logging

- Transaction logging benefits:
 - ▶ Streamlines disk I/O demands
 - Transactions are recorded in the transaction log and then written to disk sequentially
 - Results in faster commits to disk
 - In the event of a server crash, recovery time is reduced
 - ▶ Keep transaction logs on a separate physical disk

Cluster Troubleshooting

- Troubleshooting cluster replication
 - ▶ Use the Log_Replication=# setting in Notes.ini
 - Change the setting to 2, 3, or 4 to log detailed replication information
 - 2 = Summary replication info at database level
 - 3 = Detailed replication info at document level
 - 4 = Detailed replication info at field level
 - Keep the setting at 0 or 1 for regular server usage
 - 0 = no replication logging; 1 = is it replicating?
 - ▶ Check log for replication errors
 - Log analysis: Look for “replicate,” “copy,” “unable”
 - Possible causes: Replication disabled, database corruption, inconsistent ACLs

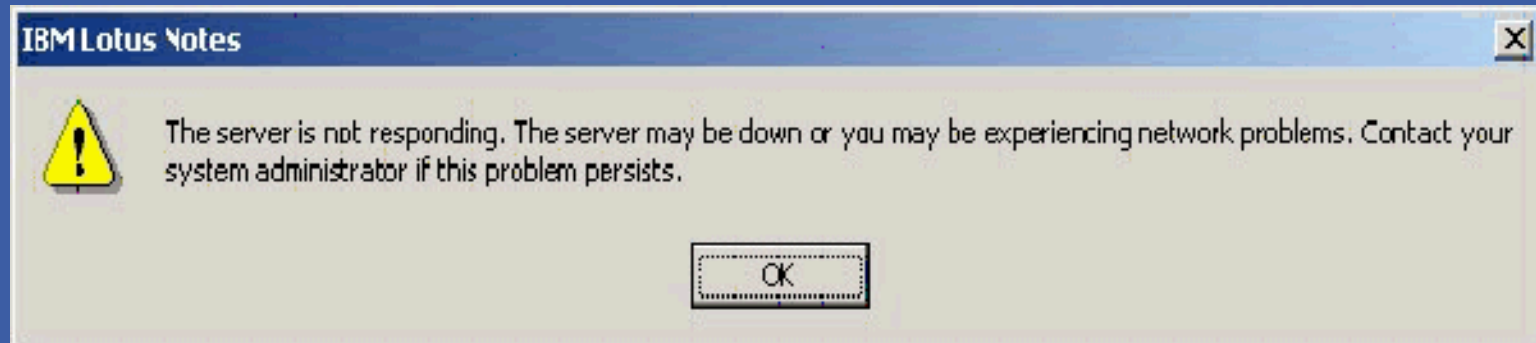
Troubleshooting Failover

□ Notes client failover problems

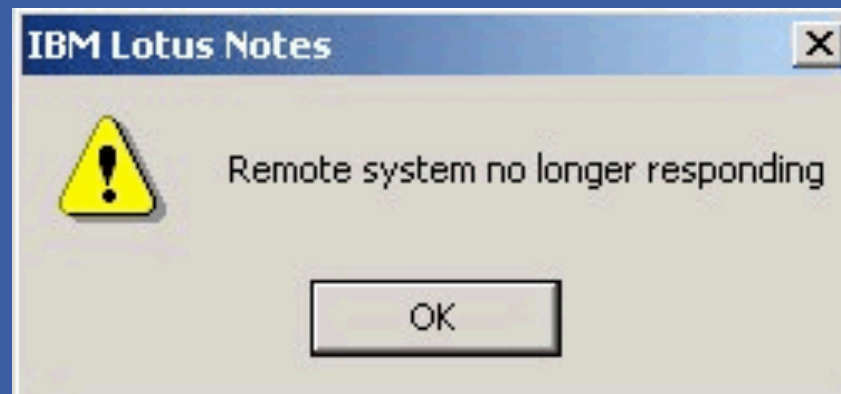
- ▶ Symptom: User is working in mail database when server crashes; receiving “Server is unavailable” error
 - Cause: Users must exit and re-open a database for failover to occur; if database is open when server becomes unavailable, failover not triggered
 - Behavior is different in R7 and higher — if user is in mail file and failover occurs, user is prompted
- ▶ Symptom: New server (Server B) is added to Server A’s cluster overnight. User attempts to access Server A in the morning, and the server is down, but failover doesn’t occur.
 - Cause: Users must authenticate with a cluster member after changes to cluster membership so that client cluster cache is updated
 - Stored in Cluster.ncf on the client

Customizing Client Failover Errors

- When a server fails with a database open, a user gets this message, known as Error 0807



- Or, a user might get this message, known as Error 0A02



Customizing Client Failover Errors (cont.)

- You can change the error messages by adding these lines to your NOTES.INI file
 - ▶ Err_0807=Your email server is no longer responding. However, you may be able to switch to a backup server and continue working. To do this, you must close your mail file and re-open it.
 - ▶ Err_0A02=Your email server is no longer responding. However, you may be able to switch to a backup server and continue working. To do this, you must close your mail file and re-open it.



- Wouldn't it be nice to be able to change ALL messages?
 - ▶ You're gonna like us!

Cluster Bottlenecks

- Finding and eliminating bottlenecks
 - ▶ Network bottlenecks
 - Private LAN traffic — is something other than the cluster sending traffic across the cluster's LAN segment?
 - ▶ Hardware bottlenecks
 - Memory — check memory utilization trends
 - Processor — is your CPU pegged at 100%?
 - Disk swap/paging files — how large? Are they sharing physical disk location with other high I/O files?
 - On SANs — Host Bust Adapters (HBAs) can be a bottleneck

Using the Cluster Analysis Tool

- What is it for?
 - ▶ Cluster analysis will assist in finding issues with your cluster environment
 - ▶ Old, but useful template
- How to use it:
 - ▶ Run from the Admin client
 - Server Tab – Analysis – Cluster Analysis
 - ▶ Creates cluster analysis database
- Let's see it in action!

Using the Decommission Server Analysis

- A quicker “poor man’s” cluster analysis
 - ▶ Used for checking functions/databases on servers about to be retired
 - ▶ Can be used for a quick “replica sync” check
 - ▶ Run from the Server – Analysis tab
- Run it both ways ...
 - ▶ Set Server A to be a target, then Server B to be the source
 - ▶ Then, the other way around
- Let’s see it in action!

Domino 8 and Clustering Enhancements

- Streaming-based clustering
 - ▶ As opposed to event-based clustering model
- Data is transferred almost immediately
 - ▶ From memory on the server
- Huge improvement on the cluster data model
 - ▶ THIS DOES NOT WORK
 - DISABLED IN 802 BY DEFAULT
 - To be resolved in 803/805
 - DEBUG_SCR_DISABLED=1
- Server cluster auxiliary port
 - ▶ Failover port for replication in case default cluster port fails
 - Server_Cluster_Auxiliary_Port=*

Wrap up

- Domino Clustering is a model administrators love
 - ▶ Hardware independent
 - ▶ Version independent
 - ▶ Platform independent

- It's very solid / scalable
 - ▶ Thousands of users

- It has its quirks
 - ▶ Just like everything/everyone

- If you have not got it implemented in your environment
 - ▶ Seriously.. its worth the license
 - ▶ How much does downtime cost?

Thank you

- Paul Mooney

- ▶ Bluewave Technology (www.bluewave.ie)
- ▶ pmooney@pmooney.net (www.pmooney.net)

- Kathleen McGivney

- ▶ Sakura Consulting
- ▶ kmcgivney@sakuraconsulting.com (www.kmcgivney.com)
- ▶