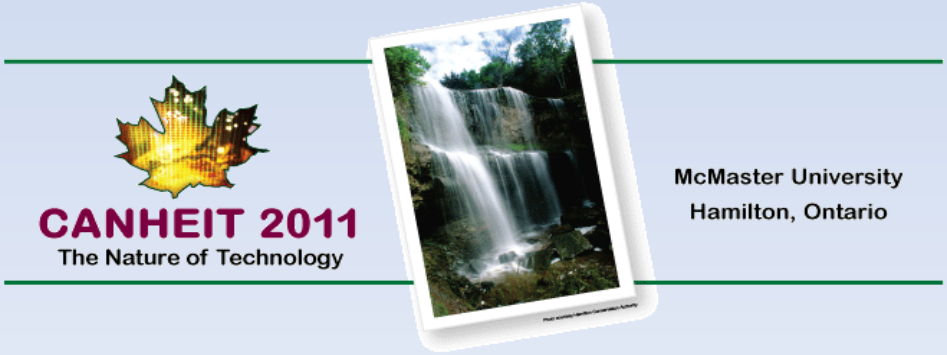
High Availability Without the Cluster (or the SAN)

Josh Sekel
IT Manager, Faculty of Business
Brock University



File Services:

- Embarked on quest; after paying for too many data recoveries, to make saving files safe and easy for end users
- Service must be transparent (fast, no extra steps) and must never fail
- Enough space for growth
- Leverage for remote access to files

Options for H.A.?

- Cluster Services
- SAN
- 3rd party software
  - San Symphony
  - FalconStor
  - Double-Take
  - Replistore
  - DFSR

DFSR

- No SAN storage needed
- Near instant replication of data
- Seamless failover
- No client configuration
- No extra licenses needed
- No hardware restrictions
- Easy server migrations
- Leverage VSS for snapshots
- Differential replication saves bandwidth

DFSR Hardware (2005-2008):

- 16 hot swap drive chassis, +2 internal
- 16 320GB 7.2K SATA HDD
- 2 74GB 10K SATA HDD
- Dual socket Opteron MB
- 4GB Ram
- Areca ARC-1260 HBA 1GB Cache

DFSR Hardware (2008-Present):

- 16 1TB 7.2K SATA HDD
- Dual socket Xeon MB
- 8GB Ram

Network is major bottleneck – same switch transfers max out multiple GB interfaces

DFSR Architecture

- Servers split between two datacenters
- Folder redirection for user files
- Domain mode (built in redundancy)
- In regular use, servers weighted equally
  - Occasionally weight a server lower to perform large dataset manipulations for researchers
  - Weight temporary servers to lowest level to allow for server upgrades while maintaining redundancy

DFSR WIN

- No user files on PC's (in theory)
- Users don't notice server maintenance in middle of day
- File access faster than local
- Makes local and remote file access (Citrix) seamless
- No data recovery services since implementation

DFSR Fail

- Win 7 libraries don't index on DFS shares
- OSX doesn't do DFS natively (fixed in Lion !?!)
- Case of the missing PDF's – Solved!
- No open file support
- No file locking on remote system
- XP and down-level clients don't follow weighting rules

DFSR Tips

- Newer is better – every version has seen major improvements
- More RAM = more system cache = speed
- Pre-Stage when possible
- Consolidate staging folders outside shares
- Don’t be cheap with staging folder size
- Enable access-based enumeration
- Use new HBA’s with SSD caching options to increase read performance

Email Services (Exchange)

- Email outages get noticed faster than any other service we provide
- Most users far more reliant on email than they should be for daily critical activities
- Nothing irritates users more than not being able to send the latest joke to their “friends”
- Losing email is not an option

Options for H.A.?

- Story changes over years
  - EXCH 2003
    - High Availability Failover Clusters
    - Load balanced web farms
    - Double-Take
  - EXCH 2007
    - SCC, LCR, CCR, and SCR (SP1)
    - Multiple Hub Transports
    - Load balanced web farms
    - Double-Take

Options for H.A.? (Continued)

- EXCH 2010
  - Database Availability Groups (DAGs)
  - Multiple Hub Transports
  - Load balanced web farms
  - Double-Take

2003 & 2007 – Double-Take

Win

- No SAN storage needed
- Real-time replication
- Automated failover
- Easy wizard driven setup

Fail

- Price
- Domain Admin Credentials !?!

2010 – DAGs

Hardware

- 2 Virtual Servers – 12GB RAM, 15K SAS Database/Logs

Logical setup

- All mail databases replicated between nodes
- Active role split between nodes

2010 – DAGs

Win

- No SAN storage needed
- Can be turned on after installation
- Designed to work with 2010's low demand for IOPS
- 5 minute setup
- No additional licenses needed
- Up to 16 database copies
- Possible to go backup free

2010 – DAGs

Fail

- Requires enterprise+ version of Server

2010 – DAGs TIPS

- Try it, you'll like it
- Take advantage of Exchange 2010's storage requirement optimizations and dump the SAN

SQL Databases

- Critical workloads relying on SQL databases
- Downtime very public
- Damage recruiting efforts
- Damage research results

Options for H.A.?

- High Availability Failover Clusters
- Third Party Software
  - Double-Take
  - Replistore
- Log Shipping
- Database Mirroring

Database Mirroring

Hardware

- 2 Virtual Servers (database) – 12GB RAM, SSD Database/Logs
- 1 Virtual Server (witness(shared))

Logical setup

- Necessary databases mirrored across nodes

Database Mirroring

WIN

- Built into MSSQL 2005+
- Synchronous replication
- Automatic failover possible
  - Less than 10 second downtime
- Simple to configure
- Simple to implement in connection string

SQL Native Client 10.0

Provider=SQLNCLI10;Data Source=myServerAddress;Failover Partner=myMirrorServerAddress;Initial Catalog=myDataBase;Integrated Security=True;

Database Mirroring

Fail

- Single mirror server
- System databases not protectable
  - Requires administrative care and feeding for logins, jobs…
- Witness required for automatic failover
  - Witness can be SQL express
  - Only one mirror server per witness instance
- Automatic failover implies

Database Mirroring Tips

- Only mirror necessary databases
- Check for vendor support for database mirroring
  - SCSM, Sharepoint 2010, XenApp, and XenDesktop 5 all support mirroring configurations
- Talk to developers to see if possible for in house apps
- MSSQL .vnext (Denali) overhauls mirroring to a model similar to DAGs

Virtual Server Guests

- Storage is the bottleneck – IOPS are key
- 15K SAS drive provides 155 read, 40 write IOPS when in raid 5
- Server VM needs 50-150w IOPS (just OS)
- 14 drive tray provides 560 IOPS, adequate for 4-11 VMs
- Single SSD provides 15000 IOPS
- 4 drive array provides 18000 IOPS adequate for 120-360 VMs (not really)
- Allows for dynamic differencing disks

VDI Guests

- Storage is the bottleneck – IOPS are key
- 15K SAS drive provides 155 read, 40 write IOPS when in raid 5
- Win7 VDI needs 15-30w IOPS
- 14 drive tray provides 560 IOPS, adequate for 18-37 VMs
- Single SSD provides 15000 IOPS
- 4 drive array provides 18000 IOPS adequate for 600-1200 VMs (not really)
- Hit other bottlenecks first

Power of Solid State

- 512 VDI sessions off a single device
- Each session running actual scripted workloads

Demo Hardware Specs

- 8 diskless servers with six core Dual socket CPUs, network adapter, Nvidia graphics cards and two network switches
- High-performance PC workstation serving as host to an ioDrive Duo