SharePoint for the DBA

# Content Databases

* A site collection is tied to a single content database
* A content database can be used for multiple site collections
* Size limitations
	+ Pre-Service Pack 1:
		- 200GB for collaboration
		- 1TB for document archive
	+ Post Service Pack 1:
		- 4TB for all usage
		- No limit for document archive
	+ If more than 4TB are needed for content databases then a scale out methodology must be used
		- Break site collections out into other databases
		- This requires multiple content databases each of which can be up to 4TB in size
		- For instance a content database may be approaching 4TB and contain 4 site collections
		- Move 2 site collections to their a second content database and leave 2 site collections in the current database
	+ The guidance from Microsoft is if a site collection is over 100GB it should be the only site collection using a given content database
	+ Possibly add some information about remote blobs (<http://sharepoint.microsoft.com/blog/Pages/BlogPost.aspx?pID=988>)
* If a site collection is over 50GB it should get its own content database
* Moving a site collection to a different content database:
	+ Moving a site from one content database to another can be done with…
	+ STSADM command line tool
	+ PowerShell
	+ Performing a backup and restore of the site collection
	+ No options in Central Administration or a GUI
	+ http://technet.microsoft.com/en-us/library/cc825328.aspx
* Often times content databases will have a GUID on the end of it
* To rename content databases
	+ Most common reason to do this is to remove the GUID or conform to naming standards
	+ In central administration switch the Database Status to Offline (as opposed to Ready)
	+ Check the box next to Remove Content Database
		- This does not drop the database, it simply disassociates it from the farm
		- All site collections contained in this content database will no longer be available
		- All data remains intact in the database
	+ Backup and Restore/Rename the database in SSMS
	+ Add the renamed content database back to SharePoint via Central Administration
	+ This should be avoided if at all possible
* Moving a content database
	+ Can be done to load balance a database server
	+ Can be done to load balance a web application
	+ Requires both SharePoint 2010 and SQL Server changes
	+ SharePoint changes can be through central administration or PowerShell
	+ Process
		- Pause services and service applications using the content database(s)
		- Remove the association in SharePoint (central admin or PowerShell)
		- Move the databases (SQL Server)
		- Add the content databases back into SharePoint (central admin or PowerShell)
		- Restart services and service applications using the content database(s)
* Methods for renaming and moving service applications can be found here: <http://technet.microsoft.com/en-us/library/ff851878.aspx>
	+ This tells if a service application can be associated by:
		- Delete and recreate the service application
		- Using PowerShell
		- Using Central Administration
* Adding addition content databases
	+ In order to scale you may need to create new content databases
	+ Content databases get full (up to the limits described above)
	+ Too many site collections in one database
	+ Methods
		- Central Administration in the Manage Content Databases section click Add A Content Database
		- PowerShell run New-SPContentDatabase command with appropriate parameters

# PowerPivot

* PowerPivot requires its own instance of SSAS
* This instance must be called PowerPivot
* Installation
	+ From the SQL Server Setup Roles screen
	+ Switch from SQL Server Feature Installation to SQL Server PowerPivot For SharePoint
	+ Select the option for a new or existing server
* Backup and Recovery
	+ Each PowerPivot workbook will have a database created when it is run in SharePoint
	+ Database name will be the WorkbookName\_<<GUID>>
	+ These cannot be renamed
	+ SharePoint will age these out per settings in Central Administration
		- No backup and restore as a result
	+ If a database is deleted when the user clicks on something in the workbook PowerPivot will recreate the database
* Cubes can NOT be reverse engineered into an SSAS project like they can be with a UDM cube

# Performance

* Correct HBA driver and firmware versions
* Standard SQLIO.exe for disk I/O performance test
* Configure correct NTFS Allocation Unit Size
	+ 64k best, default can cause up to 30% performance drop off
* Correct Windows “Sector Alignment”
	+ Up to a 50% performance drop if wrong
* Free space on disk partitions should be greater than 25%
* File Placement from fastest drive to slowest drive
	+ TempDB (data and log)
	+ DB Transaction logs
	+ Search DB data files
	+ Content DB data files
	+ If the sites are more read centric than read-write then move the content above logs
* Multiple data files
	+ # of data files should be <= # of processor cores
	+ Only split data files for content and search databases
	+ Other databases do not support splitting data files
* Standard SQL best practices for data file sizes, auto grow, temp db, etc.
* Memory
	+ Small farm – 8GB or more
	+ Medium farm – 16GB or more
	+ Large farm – 32GB or more
* Defrag the disks

# Database Maintenance

* Integrity checks should be done regularly
* REPAIR\_ALLOW\_DATA\_LOSS is not supported on the databases
* REPAIR\_REBUILD is available but may not always work
* Content and Search are most likely to become fragmented
* Index rebuild/reorg as you normally would
* Do NOT add or remove indexes!!!

# Backup and Recovery

* Perform regular database backups
* Stagger backups of the databases
* Incremental backup when possible on larger databases
* Compress backups
* Follow SQL Server backup/restore optimization recommendations
* Farm and site collection backups should be handled by the SharePoint admin

# High Availability

* It is important to scale out to increase capacity
* <http://technet.microsoft.com/en-us/library/dd207313.aspx>
* <http://technet.microsoft.com/en-us/library/cc748824.aspx>
* SQL Server failover clustering is available
	+ http://technet.microsoft.com/en-us/library/dd207311.aspx
	+ Can run on any combination of active and passive nodes
	+ SharePoint references the cluster as a whole, it is not cluster aware
	+ Members of the cluster must be on the same subnet
	+ All setup and maintenance is on the SQL Server side
* SQL Server mirroring is available
	+ http://technet.microsoft.com/en-us/library/dd207314.aspx
	+ Must use high-availability mirroring (high-safety) mode with automatic failover
	+ SharePoint 2010 IS mirror aware
	+ The failover server must be specified in Central Administration (or configured through PowerShell)
	+ Principal, mirror and witness must all be on the same LAN (handle up to 1 millisecond latency round trip)
	+ Manual reconfiguration after failover
	+ Mirroring is NOT available on the User Profile Service’s Synchronization database
	+ Mirroring is NOT available on the Web Analytics’ Staging database
	+ Mirroring is NOT recommended on the Health Data Collection’s Logging database
	+ Mirroring is NOT recommended on the Application Registry’s database
* Recovery time for clustering is lower than that of mirroring (in the milliseconds)

# Disaster Recovery

* Rolled in with the rest of the SQL Server DR plan
* See High Availability above
* SharePoint admins will have their own DR plan outside the scope of the databases as well
* Some service applications will need to be configured on the primary and failover farms, such as Excel Services (for which there is no database associated)
* The following do not support log shipping
	+ Application Registry
	+ Business Data Connectivity
	+ User Profile Service
	+ SharePoint Foundation Subscription Settings
	+ Search
	+ Word automation services

# Service Applications

* The following service applications store data in databases
	+ Search (3 databases)
		- Search Administration
		- Crawl
		- Property
	+ User Profile (3 databases)
		- Profiles
		- Social
		- Synchronization
	+ Business Data Connectivity
	+ Application Registry
	+ Usage and Health Data Collection
	+ Managed Metadata
	+ Secure Store
	+ State
	+ Web Analytics (2 databases)
		- Reporting
		- Staging
	+ Word Automation
	+ SharePoint Foundation Subscription Settings
	+ PerformancePoint