User Group TechEd Notes

Tuesday, June 11, 2013   3:00 PM

Lightning Round

Keynote

Inactive azure vms no longer cost anything, per minute not per hour charges
Lots of credits come for azure with each msdn subscription
MSDN licenses apply to Azure IaaS servers

Wooo Aston martin

MDC B370
Dev, test for azure infrastructure

- Pay by minute now, notbyhour. Need to activate your msdn subscription for azure.
- Three ways to get started: management portal, powershell, REST API (nojs)
- Sharepoint 2013 is available in the "gallery", 2010 is tested but not in the "gallery". Sql 2008 r2 and 2012 in gallery. TFS, project server are not.
- Hypervisor in azure and server 2012 is the same. Can pretty much more between hyperV on premise and azure. Some networking stuff doesn't work, like multicast for example.
- Status "stopped (deallocated)" no longer costs anything. This is a big deal. Turn off your VM's at EOD, no charge.
- Manage.windowsazure.com
- Quickcreate gets a quick serverup and running, you provide server name, size, user name and a password and you're up. gallery contains different prebuilds like sql.
    - DNS will be exposed to the internet.
    - By default remote powershell is enable via https
- Thy showed off scripts to take and restore snapshots of azure machines… why isn't that built in?
- Availability set machines are always stored on different hardware. Load balanced set can redirect traffic.
- $50 $100 $150 per month for prof, premium, ultimate
- For dev test you can license sql enterprise for six cents per hour. Down from $2.19/hour.
  But that is for dev/test only. By 2014, all msdn subscriptions will be converted to dev/test intent ONLY.
- The azure machines support vpn to premises. Azure connect is now replaced by point to site and site to site, which uses the builtin windows vpn service to connect to the cloud directy, machine to machine. Looks like https, but its actually a tunnel ssp. Involves setting up a virtual network of azure machines, all in an azure affinity group.
    - Azure has its own dns, but you can change that to an on-premises DNS.
    - With point to site, you provide a range of IP addresses (lke 192s)

that can see the Azure site
- ○ Requires a gateway server that is 5 cents per hour.
- ○ Azure provides vpn scripts for all major vpn clients, powershell script that setsup the site to site gateway vpn. Also needs a clientroot cert to be set up on the clientsite.

Mdcb328

SCCM Orchestrator is the service manager process controller. Makes workflows can that be part of the change control process. It's like ssis for system center 2012. Automation and standardization. Built in to request ticket process. Automated processes can kick off based on email responses from admins or managers for example. They demoed a request for new VM, email approvals and workflow kickoff.

APM, application performance monitoring. It is what scom uses to monitor the whole app stack including SQL. The .net pet shop is a sample app available from Microsoft to set up and demo scom apm. It's a tool built for sysadmins and operations managers to do sql tasks from inside scom. It can diagnose sql problems (dB offline) vs network issues. Synthetic transactions regularly test interactivity with the application, records duration and trending.

6/5
DBI-B307
Hekaton

- Heck a ton
- Will be in sql2014
- Fully durable in memory data, no locks, no blocking, no latches, no pages. Minimal context switching. No deadlocks.
- Move most important tables to Hekaton, leave most tables in traditional sql. No changes to the database definition or settings. You add a memory_optimized_data filegroup. Ssms and dmvs track the usage of this filegroup like any other filegroup.
- 10-25x faster
- Stored procedures are compiled now into native machine code
- The storage format is different. All optimistic, no pages. Traditional latching does not scale, according to him, past 12-16 cores. Hekaton has friction free scaling.
- Hekaton is inside sql2014. Seamless inside the engine. "migrate table to in memory". No changes to surrounding tables, database. "memory optimized" vs "disk or page based tables"
- Uses a third kind of index called a hash index.
- Resource governor can limit memory used by Hekaton tables
- On memory DDL statements are compiled now and run by a dll.
- Hekaton rows have built in row versioning similar to snapshot isolation mode (since sql 2005). All current needed versions are kept in the hashtable. Does not use tempdb.
- Limitations: no DML triggers, no xml, no CLR, no LOB types like varchar(max)
- No foreign keys. No check constraints, no identity
- No altering the table. Must be dropped recreated. Secondary indexes must be specified at create time.

- …yet.
- Access tables in memory just like all tables. Select * etc.
- Create proc … with (native_compilation, schemabinding
  - This means the code will be converted to C and a dll is created. Schemabinding is required, removes runtime checks for validity (good idea anyway)
  - Native compiled procs can only access memory optimized tables
- Moving table to memory optimized requires no bfgother changes to existing procs.
  - But would be even faster if you wrote a natively compiled sprocs
- Sys.tables has new values to indicate in memory, durability
- Did a demo of inserting 500000 rows. In memory was 4s, table based was 14s. Same table definitions.
- The hash table is not ordered like a clustered index usually is.
- Delete 500000 rows from inmemory table happened instantly. From disk based tables took 9s. That's without using natively compiled sproc code.
- C code files exist on hard drive that can be viewed. Full of C language STRUCTs. Natively compiled stored procedures are converted to C.
- Funny. First time he tried to insert 500000 rows with a natively compiled sproc it took 14s. Delete data, run it a Second time? 0s. (blamed on early bits)
- Instead of deadlocking, the victim transaction is failed because of a write conflict and rollsback. Must have retry logic in code to handle conflicts (a good idea anyway). Gave example of calling a sproc within a try catch to handle a specific error message indicating a Hekaton write conflict. Failure retry logic is usually a good thing anyway, is now more important.
- At the end of the day the commit speed of durable in memory tables is still limited by the commit speed of the disk, because durable tables are safe on disk as soon as it commits.

User group

fancy handout from coming in sql preso

Certifications
McsA and MCSE : used to be architect/engineer, now it is Associate and Expert.

data platform exams for SQL 2012 is 70-461, 462, 463 for MCSA. two more exams for two different MCSE's. 464 and 465 for the data platform MCSE, 466 and 467 for the business intelligence MCSE. important to note that the MCSA includes BI- they want to certify well rounded experts, not one dimensional folks. they want folks to present solutions, not just narrow aspects of the SQL platform. So an MCSA includes BI, includes the new Data Quality Services tool (especially in 463), includes SSIS, in addition to what your expect from admin and security and querying.

Dbi b319
Dbas worst nightmares

In recovering a corrupted data file, Manual tran log backup should be
no_truncate so as to not wreck tran log shipping or repl potentially.

If things are consistent, detach reattach the data file regenerates the log. If
not, have to rebuild.
In 2012, dbcc checkdb with allow data loss will fix it, but that needs single
user. And that cant be done until the database is in emergency mode. In 2012,
this is a simple alter database set emergency. Lost tlog but not data file in sql
2012: emergency, single user, checkdb allow data loss, multi user.

"Due to data movement" error is a nolock error. Outside of nolock, get a page
inconsistency error. Had to drop recreate index, rebuild didn't work. That's for
a nc index.

Dbcc traceon(3604)
Dbcc page (…)

On a clustered index, different solution. Cant rebuild, must restore. If backup
is corrupted, try enterprise level feature page backup. Restore database x
page = , then roll forward all the tran logs since to recover. Logs know which
pages need to be updated, very cool.

Fn_dblog(null,null)
To get transaction ids, identify the one that has the most LOP_delete_rows
rows. Get transaction ID. Can get lsn of the one that says lop_commit_exact.
Backup and restore tlog up to that LSN using STOPBEFPREMARK='lsn:', then
rollback the transaction id. Must convert lsn from hex to decimal. Can use
same approach to restore sproc.

Mdc b337

Clustering does a secure handshake not just a ping for health monitoring.
Communications over port 3343, using netft and clussvc. heartbeats by
default every one sec, threshold is five missed heartbeats.
In server 2012, you can now clone servers and put them in a cluster, it
handles the mac address to generate a unique hash.
NTM network topology manager.
Performance filter improves performance by detecting csv on UDP and
routing it directly to the clussvc
Csv is cluster shared volumes, introduced in 2008r2 for hyperv only. In server
2012, csv is supported on all. CSV orchestrates metadata maintenance. CSV is
impacted negatively by lack of bandwidth. Network performance could
impact physical IO because the system is synchronous.
Quality of service is more important than bandwidth
Lots of things recommend their own nic\network because qos is most
important. It'll all run fine on one nic or nic team.
CSV handles metadata but can also route io through a cluster partner if one
side lost access to the SAN, without failover.

Clustering can cause more downtime (failover) than it solves if qos is bad. Can reduce sensitivity (heartbeat frequency, threshold) if you have a known network qos issue. This will mask network issues.

Tcp reconnect is 20s. Extending thresholds for failover past that just extends downtime perceived by client. Sql thresholds should not be changed, need to stay aggressive.

In server 2012, clustering across subnets is supported. This is foundation of alwayson.

File share witness holds a race for servers to lock a file if two nodes lose connection, whoever wins gets the ownership.

Dbi405
Ssis deep inside
Matt masson and matthew roche

- IsSvrExec in early dev stages used to be inside the sqlsvr main executable. Lots of COM calls and crash issues, they moved it out.
- SSISDB is a user database. The catalog schema is intended for public use. The internal schema is subject to change and not supported.
- Everything is encrypted automatically. Uses same key as sql for instance and database keys. Uses standard key tsql behind the scene for row level security. The catalog views automatically decrypt those values for you.
- Need to back up and restore the keys if you ever migrate the SSISDB. Anything marked sensitive is automatically encrypted. No sensitive data is stored in plain text.
- Why only one catalog? Nobody wanted multiple in leadup to production. They stripped off tht feature but it was too late to make that change in ssms.
- Project deployment model vs package model is like proper ALM vs old classic ASP files. No more deploying to prod by replacing a file on a server. Makes it possible for proper software development and deployment.
- Configurations with external unenforceable references are replaced by parameters and environment variables, which are required and enforced.
- Project is deployed as a binary, an .ispac file, which is encrypted on the server, as a whole for the whole project. Individual packages are not stored separately and are "compiled".
- You can actually deploy straight with tsql. Exec catalog.deploy_project.
- SSIS keeps versions of each project (not package). But not a source control solution.
- Deployment of ssis projects can follow testing, approval, automated deployment by third party, etc., just like other visual studio applications.
- However many ISServer executables running as each execution of a package currently running.
- A CLR assembly marked unsafe communicates as the gateway to the ISServerExec via named pipes. ISServerExec communicates back with standard ADO.net SqlConnections with all the status info on execution. This is an asynchronous process that is much higher performance. Guaranteed to be written in the right order but they are queued and

written later. Timestamps ARE accurate but they may not show up in SSIS logging instantly.
- "cannot connect as anonymous" is a common error that is caused by the Kerberos double hop problem, because you're executing it interactively and it is impersonating you. Usually running the project in a sql agent job solves this.
- Basic is the default logging level, but performance caps errors and stats on perf. Verbose caps everything, and is the only one that caps rowcounts, also captures info on packages calling child packages, including exactly what they passed via parameters in an xml blog. None caps package execution only.
- Mattmasson.com
- Can use $User::ServerExecutionID inside ssis to reference the catalog schema data to extend logging even more.
- Can capture performance counters while it is running.
- Select * from Catalog.dm_execution_performance_counters (executionID)

6/5
DBI-B306
SQL HA DR running on Azure

- Similar slides to mdcb370

- The vm depot has a gallery of nonofficial content including lamp and git servers and drupal prebuilt vms

- Azure has built-in service healing to move your VM to a different host, about 12min downtime. Faster than onpremise failover but still present. You'll still want sql HA even in the cloud, just like normal, with AlwaysOn or mirroring (old).
- Mirroring is deprecated but will still be there in sql 2014
- With always on, zero data loss in synchronous mode, 4 secondaries in sql 2012 (8 in 2014), auto failover at the database level, readable secondaries, take backups on secondaries, just like on premises. (enterprise edition only).
- Mirroring features is the same too
- Availability groups require windows cluster but not shared storage in azure or on premises, no difference
- Because they need to be in the same domain, you'll need an azure domain controller too.
- Availability groups listeners is not supported yet… will be in a few months. How does that impact us? Use failover partner instead?
- Put availability groups domains in the same affinity group, availability set and VNet inside azure.
- There is an issue with dynamic ips for the vms, I didn't follow, **check the slides**.
- Log shipping is not required
- To run sql availability groups on premises and in azure, you need to set up a site to site vpn. All servers are in the same domain and cluster, on

premises and in azure.

- If you have databases larger than 1tb, you would have to attach multiple disks to your vm. That is a problem for sql because those disks aren't guaranteed to write in the same order because of georeplication. Could break a single database larger than 1tb.
- Backup to url - new sql 2012 sp1 feature - restore your sql databases to azure!
- Availability sets is a windows azure concept, vms are hosted in different physical racks in the aux datacenter, subject to different scheduled downtime for example.
- Availability groups is an AlwaysOn concept.p

Windows azure sql databases is new name for sql azure. 150gb dB limit. Not the topic of anything I've seen this week. Too many limitations. Perhaps it is dying? Running real sql in real windows in azure has more compatibility and same benefits.

- Sql 2014 will have an availability groups button to add an azure replica, built into ssms.
- Will run through the azure vm setup wizard inside ssms

Dbi b318
Sql CAT ha customer stories

- Pointed out that AlwaysOn is a "brandname" for sql server Failover Cluster Instances + availability groups.
- Edgenet: MS DTC service is not supported by availability groups. Interesting.
- They don't replicate tempdb over the cluster - stored on local disks only. Doesn't need to replicate that heavy startup and heavy transient workload
- They use the "register all providers" setting onhosts on all subnets .registers provider ip addresses in DNS. Client connections support "multisubnetfailover = true"
- Clustered heartbeats happen on all interfaces, from very node to every other node. Need a star topology to avoid poor cluster communication. PlumbAllCrossSubnetRoutes- more resiliency of subnet communications. Can increase errors if you don't have a valid route to other subnets.
- OR- windows server 2012 provides "plumbAllCrossSubnetRoutes" value of 2, which will use the first one that works, not try all the routes and generate so much failure.
- Migration to AlwaysOn tip: use log shipping to catch new server up to date. When you're ready, stop tran log shipping, start alwayson, the data is ready.
- CareGroup hospitals in Boston rated #1 most innovative IT department nationwide
- Listener app is the virtual network name that sits over the availability group and provides a name for application connection strings to use, regardless of which node is primary.
- Database mirroring app connection strings are different from the availability groups conn string. Needs to look at listener, not failover

partners.
- Monitoring skillset for db mirroring is different from availability groups skillset. Much more complicated with the cluster factor added. Win server 2012 highly recommended, much much easier. Dynamic quorum especially is a useful feature.
- Bridgewater uses a server replication pair that sucks the tran logs in from a primary server into filetable, and copies (mirroring? AG?) it to a partner on the farside datacenter.
- They have such massive replication, they have each publication use its own distribution database and ach distribution database has its own dedicate disk.
- They use one low-privileged login as the owner of databases on all sql servers. This is important, it is what EXECUTE AS OWNER procs execute as.
- Alwayson does not handle linked servers with deployment, gotta set those up manually.
- They switch datacenters each month as their DR test. Each time they do that, they find something new in their scripts to add to their plan. Also handle patches every month like that.

Dbi-b309
Tran log internal
Tim chapman

- Changes always occur in the tran log file before they occur in the data file.
- ACID atomicity, consistency, isolation and durability
  - Tran log - consistency and isolation
  - write ahead property - atomic and durability
- Tran log files are always written sequentially. Should be on their own dedicated disks. Rid 1 and 10 are good, raid 5 is s bad idea because of the write overhead.
- In startup, enterprise edition starts up after the redo phase. In standard, must wait for the rollback phase to complete.
- Checkpoint
  - Writes dirty pages from buffer pool to the disk
  - Ensures the redo phase, where transactions are written to disk, is kept at a minimum.
  - The lazywriter will eventually clean the buffer pool
- Before 2012, the checkpoint occurred at a frequency set by instance level. Default is 0, SQL determines it, usually it's around a minute.
  - With 2012, we have a database option for target recovery time in seconds. Can set this at database level. Default is still 0.
- Used fn_dblog to view pages in the transaction log affected by transactions in his demo
- Unlike previous presenter in the dba nightmares session, Chapman says you can use the Tran log to determine who executed a transaction. Use sid

Dbib305
Sqlvariant

- More features are coming for us so that we don't have to script out the predeployment files for initial data population, in lookup tables for example.
- You can now script a database to include schema and data, which would include a .bcp file into a .sql file.
- Which gods of demos did this guy piss off?
- "register as a data tier application" makes it look at the database first. It can then stop everything if it recognizes database "drift". allows you to do a schema compare and justify the database solution, build and republish.

Dbi-b322
What's New for Columnstore Indexes and Batch Mode Processing

- Now you can create a new type of columnstore index as your clustered index. It is updateable too. Highly concurrent as well, much more concurrent than regular databases. Traditional nonclustered indexes are no longer needed on clustered columnstore indexes and aren't supported. Limitations of previous columnstore NC indexes are removed.
- Traditional data is stored in heaps or b-trees. Columnstore is stored in columns compressed over traditional database pages. Can be compressed in columnstore much better than in regular storage. Fits much better in memory. Now you've got more data in memory. Now you've got more data in cache. Columnstore is not in memory like Hekaton, but can put data in memory far more efficiently than traditional data. Each column is accessed independently.
- Columnstore pages are stored in segments. Each segment can contain up to one million rows of compressed data. Segments go into memory as a whole. Query engine can intelligently ignore whole segments, as each segment stores a lot of metadata.
- Showed a report that went from 9 minutes on traditional indexes to 1-2s with a clustered columnstore on the same table.
- Sql 2014 includes Batch Mode Processing in the execution plan, uses "vector operations" to greatly reduce cpu usage 7-40x. Batch mode shows up in the execution plan in Actual and Estimated "execution mode". Batch execution mode is a major cpu saver over "row".
- Clustered columnstore indexes are inserted into via a heap which are assembled into Columnstore structure once they are big enough. Those columnstore structs are softdeleted only with a bit field, not actually deleted, as that would involve too much overhead. Updates are softdeletes plus inserts, which are assembled later on with index maintenance.
- When inserting, the "tuple mover" will convert the delta store of 100kmrows or more into a columnstore row group every five minutes. BULK INSERT ops will automatically create columnstore row groups of one million rows.
- So do batches of 90k segments so that you'll eventually get lots of large segments mod 101k row inserts and you'll get a bunch of smaller segments that aren't great for performance. Would Need to fix those

with an index rebuild.
- Million rows plus is the only situation you're likely to see big advantages.
- You can convert traditional tables to columnstore clustered indexes.
- Clustered columnstore indexed table is roughly <10% the size of a traditional table without compression. About 20% the size of a traditional table with page compression
- Unique keys ND foreign key constraints are not supported with clustered columnstore. No other indexes are allowed period. This is a DW construct for reporting, not for relational transactional processing. That said, funny enough, his batch mode query plan on 2014 on a clustered columnstore table returned a… missing index suggestion. Lol.
- Everything but clr, (max) data types, xml and spatial will be supported in columnstore indexes of both types in SQL 2014.
- Nonclustered columnstore indexes will still be read only in 2014.
- New archival compression is available in sql 2014 as an additional layer of compression on top of columnstore compression.
- For Nonclustered columnstore indexes on 2012 or 2014: drop insert create or partition switching. Highly recommended using partition switching to add data to data warehouse tables, much less tables with columnstore indexes on 2012 or 2014.