

High availability database-omgeving wordt geboden

SQL Server 2008 Always On

Bram Dons

Een hoge beschikbaarheid van servers is van essentieel belang voor 'mission-critical' applicaties. Mission-critical applicaties zijn die applicaties waarvan een onderneming afhankelijk is om de dagelijkse operaties goed te kunnen afhandelen. Met de komst van virtualisatietechnologie is er echter een nieuwe categorie van computing bijgekomen: business-critical.

Business-critical kan worden gedefinieerd als een verzameling applicaties die afzonderlijk niet van essentieel belang zijn, maar collectief wel omdat de onderneming afhankelijk is van hun collectieve beschikbaarheid. Een database server vormt voor de meeste ondernemingen een van de meest kritische applicaties, dus is het belangrijk dat de beschikbaarheid daarvan in een enterprise-omgeving in zowel een traditionele als virtuele server-omgeving zo hoog mogelijk is. Een database applicatie maakt steeds vaker deel uit van een business-critical toepassing.

SQL Server 2008 high availability

Microsoft heeft met de vorige SQL Server 5 release al diverse voorzieningen, met name mirroring, geïntroduceerd om de beschikbaarheid van de database te verhogen. Bij de opvolger van SQL Server 5, SQL Server 2008, zijn deze voorzieningen nog verder verbeterd en uitgebreid. Microsoft noemt deze bundeling van beschikbaarheidsvoorzieningen 'Always On Technologies'. Een altijd beschikbare technologie bestaat natuurlijk alleen in theorie, in de praktijk vertaalt zich dat in een streven van leveranciers om een zo hoog mogelijke beschikbaarheid te bieden.

Niet alleen is een hoge beschikbaarheid van databases van eminent belang maar ook een snelle hersteltijd, mocht zich onverhoopt toch een catastrofe, van welke aard dan ook, voordoen. SQL Server 2008 voorziet in een technologie om snel te kunnen herstellen van een zogenaamde 'on geplande' downtime, een snel herstel (fast recovery) na een systeemcrash en database mirroring failover. Fast recovery maakt de database beschikbaar tijdens de 'undo' fase van het herstelproces, gedeeltelijke beschikbaarheid tijdens restore operaties, database pages checksums en backup media mirroring. De ondersteuning van fast recovery op SQL Server 2008 minimaliseert de tijd die nodig

is om de database online te brengen zodat gebruikers niet onnodig lang hoeven te wachten tot het totale herstelproces is beëindigd.

SQL Server 2008 biedt verschillende opties voor het creëren van high availability (HA) voor een server of database: failover clustering; database mirroring, log shipping, replicatie, geografische verspreide clusters en schaalbare shared databases. In dit artikel besteden we niet alleen aandacht aan de nieuwe in SQL Server 2008 aangebrachte HA-voorzieningen maar ook aan de combinatie met Windows Server 2008 en Hyper-V (Microsoft's nieuwste virtualisatie-oplossing).

Combinatie SQL Server 2008 en Windows Server 2008

De combinatie van SQL Server 2008 en Windows Server 2008 biedt een krachtig platform voor mission-critical data applicaties. Samen bieden ze een hoogwaardige oplossing voor end-to-end beveiliging, beheer en hoge beschikbaarheid. In het algemeen zorgt HA-technologie ervoor om de server downtime tot een minimum te beperken. Een belangrijke 'oorzaak' van server downtime is het zogenaamde geplande onderhoud. Naast software-onderhoud verstaan we hieronder acties als het toevoegen of vervangen van geheugen, processoren, netwerkkaarten, enzovoort. Nieuw is de mogelijkheid om bepaalde acties met betrekking tot de server hardware online te kunnen uitvoeren, zonder dat deze hoeft te worden afgeschakeld. De 64-bit en Itanium versies van Windows Server 2008 ondersteunen daartoe dynamic hardware partitioning, dat er voor zorgt dat de serviceverlening van SQL Server 2008 databases wordt gecontinueerd tijdens een hardware upgrade. SQL Server 2008 borduurt voort op de ondersteuning van *hot-add memory* uit de vorige SQL releases door een nieuwe hot-add CPU-voorziening te bieden voor Windows Server 2008. Uiteraard moet de server hardware deze hot-add voorzieningen wel ondersteunen.

De andere vorm van server downtime, de niet-geplande, kan worden veroorzaakt door hardwarefouten of catastrofes.

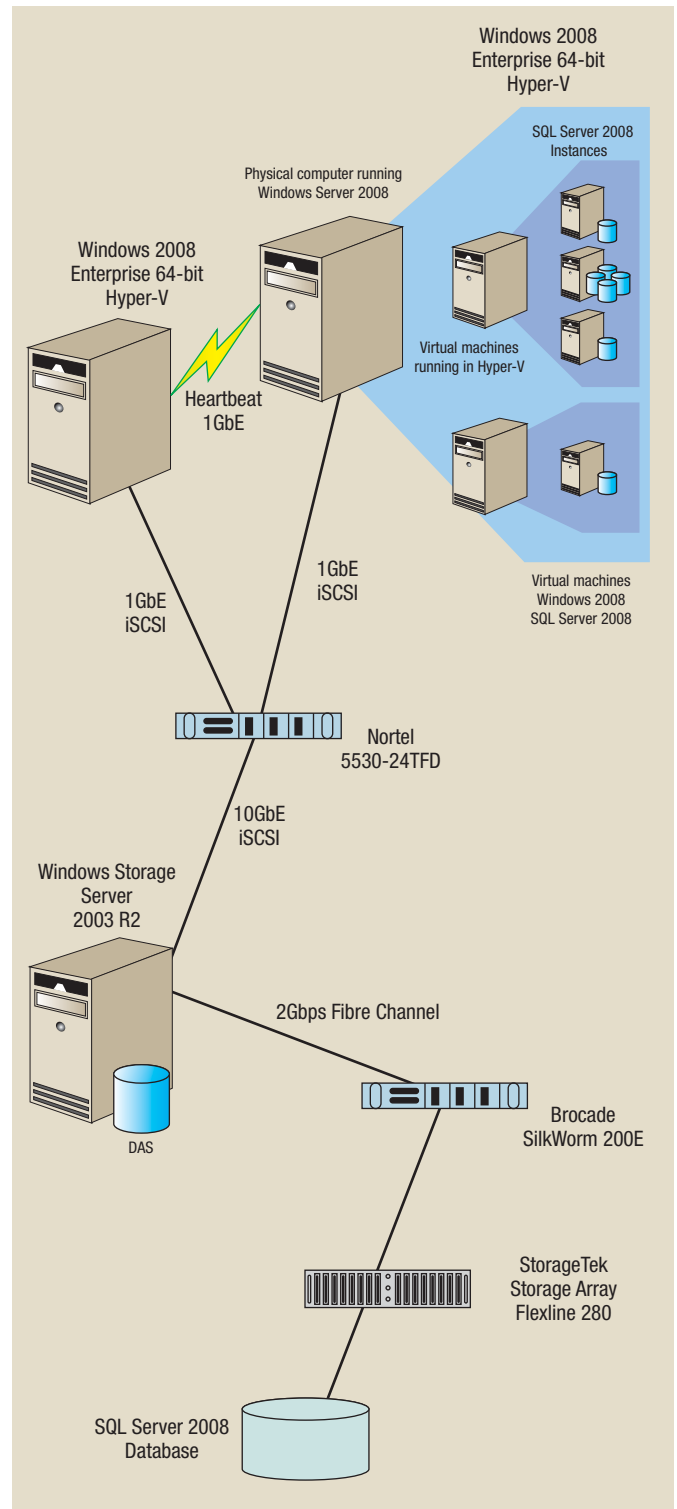
Doorgaans beschermen ondernemingen zich hiertegen door mission-critical servers uit te rusten met een server cluster-voorziening. Ook de Windows Server 2008 borduurt voort op de cluster-technologie die is ontwikkeld op basis van de vorige Windows Server generaties. Een probleem daarbij was altijd de

lastig te configureren cluster-omgeving en onvoldoende mogelijkheid om vooraf de configuratie te testen of deze 'clusterwaardig' was. Bij de nieuwe Windows Server 2008 is het setup-proces vereenvoudigd en kan met behulp van een validatie-tool worden gecontroleerd of de juiste hardware en shared data resources voldoen aan bepaalde eisen. Verder is de cluster verbeterd door de eis te laten vervallen dat alle nodes op hetzelfde subnet aanwezig moeten zijn en de ondersteuning van OR afhankelijkheden (bijvoorbeeld, door te verzekeren dat de netwerkbron beschikbaar is als een van de twee IP-adresbronnen dat is). Windows Server 2008 heeft ook de noodzaak van de quorum disk laten vallen die als shared device voorheen altijd een single point of failure (spof) vormde.

Server 2008 Clustering ondersteunt ook Global Unique Identifier (GUID) partition table (GPT) disks. GPT disks zijn met name voor SQL Server 2008 van belang, omdat deze enkelvoudige partities ondersteunen die groter zijn dan 2 Terabyte en redundantie bieden door informatie over de partitie op te slaan. SQL Server 2008 ondersteunt 64-bit, multi-core en multi-processor systemen en kan tot 64 GB geheugen rechtstreeks adresseren en tot 8 Terabyte op 64-bit hardware. Bij de Windows Server 2008 Enterprise en Datacenter edities zijn de cluster-voorzieningen van de vorige versies van Windows vereenvoudigd door de benodigde hardware en infrastructurele voorzieningen te verminderen. Een belangrijke vernieuwing is ook de mogelijkheid om cluster-nodes op apart gerouteerde netwerken te installeren, wat de mogelijkheid biedt voor de toepassing van Geographically Dispersed Failover Clustering. Deze failover methode biedt redundantie op server-niveau op een door Microsoft ondersteunde Cluster Services configuratie met een of meer non shared storage array's op elke side. Als op een van de sites de disk of de server node uitvalt, dan handelt de failover cluster alle volgende activiteiten af op de andere site. Door de toepassing van lokale disks en de afwezigheid van een quorum disk wordt het risico bij uitval van een shared disk array weggenomen, die bij een shared storage cluster-configuratie altijd een single point of failure vormt.

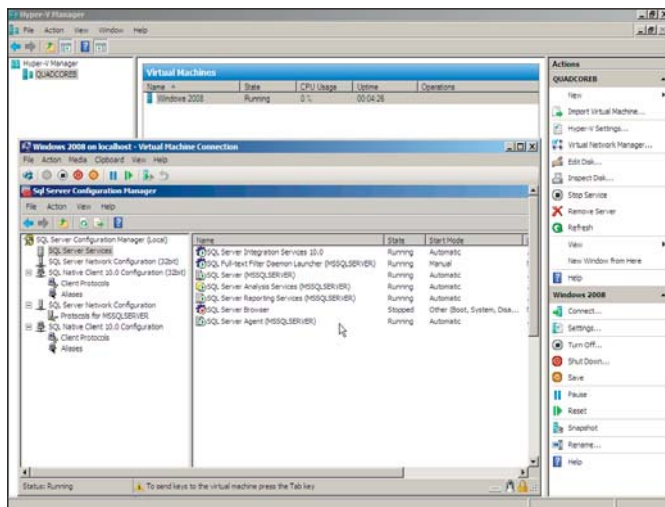
Database mirroring, snapshots en replicatie

Database mirroring is hoofdzakelijk een softwaretechnologie om de database availability te vergroten van alle database instances (met uitzondering van de master, msdb, tempdb en model databases). Data die zijn veranderd worden asynchroon of synchroon naar een remote SQL Server gespiegeld. Bij de SQL Server 2008 Enterprise Edition is de mirroring beveiliging uitgebreid naar data pages. Bij ontdekking van een verminkte page op een van beide servers wordt de betreffende page opgehaald van de andere server en vervangen. Het database mirroring proces is op SQL Server 2008 geoptimaliseerd door de transfer data te comprimeren en bij de uitvoering van een handmatige failover de noodzaak van een database herstart weg te nemen. SQL Server 2008 ondersteunt point-in-time database snapshots en



Afbeelding 1: Cluster-opstelling SQL Server 2008.

biedt, naast de gebruikelijke snapshot-operaties, read-only toegang tot database mirrors. Tenslotte, het peer-to-peer replicatieproces van SQL Server 2008 is verbeterd door de toevoeging van replicatie-nodes waarbij het replicatieproces actief blijft. Verder is SQL Server 2008 in staat om conflicten te detecteren wanneer meerdere replicatie-nodes dezelfde database-rij proberen te wijzigen.



Afbeelding 2: Hyper-V Manager.

Hyper-V

Met de komst van Windows Server 2008 werd de nieuwe Microsoft Hyper-V geïntroduceerd. Dat is een nieuw type hypervisor, door CPU hardware (Intel en AMD) ondersteunde virtualisatietechnologie. Het biedt ondernemingen de mogelijkheid om meerdere server roles te kunnen consolideren binnen aparte Virtual Machines (VM) die op een enkele fysieke host computer draaien. Zoals bekend is; de toepassing van virtualisatie om meerdere servers te consolideren op minder computers verlaagt niet alleen de TCO, maar verbetert ook de operationele flexibiliteit door de automatische toekenning van VM's, de mogelijkheid om zowel 32-bit als 64-bit werkbelasting op hetzelfde platform te kunnen draaien en de hosting van een third-party OS binnen Windows Server 2008. Verder biedt virtualisatie de mogelijkheid om op meerdere servers met Windows Server 2008, met multi-database en multi-instance voorzieningen van SQL Server 2008 een flexibel en dynamisch datacenter te creëren. SQL Server 2008 Enterprise Edition reduceert bovendien de licentiekosten omdat er slechts een enkele SQL Server licentie nodig is voor het draaien van een onbeperkt aantal instances van SQL Server binnen VM's op dezelfde fysieke host.

Testopstelling SQL Server 2008 Cluster

Om te zien op welke wijze de nieuwe SQL Server 2008 met betrekking tot HA zich gedraagt in een virtuele Windows Server 2008 cluster-omgeving installeren we een two-node cluster-systeem. Daarbij wordt de SQL Server 2008 Enterprise 64-bit versie binnen een VM op een van de Windows 2008 Enterprise 64-bit server cluster-nodes geïnstalleerd (zie afbeelding 1). Afgezien van de cluster is het voor de toepassing van Hyper-V nodig om van shared storage gebruik te maken. Daartoe worden beide cluster-nodes voor de ondersteuning van shared storage gekoppeld aan een Windows Storage Server 2003 R2 systeem. Voor de installatie van een Windows Server 2008 cluster zijn tenminste twee aparte logische disk volumes (LUN's) nodig. Een

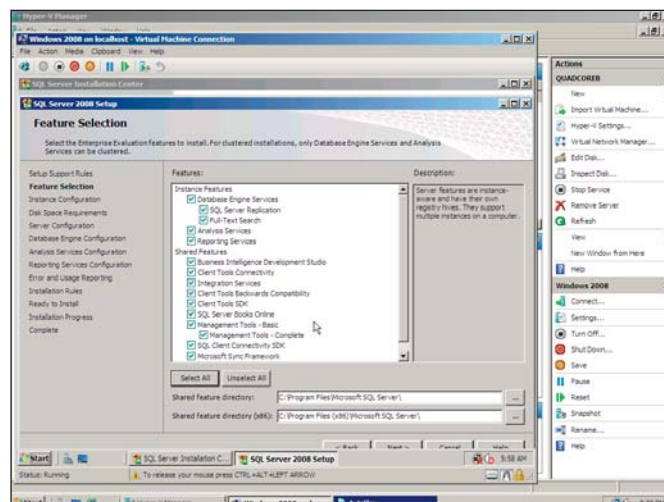
volume fungeert als 'witness' disk, de ander bevat alle bestanden die tussen de cluster-nodes worden gedeeld. De witness disk is een disk waarop een kopie van de cluster databaseconfiguratie wordt opgeslagen. Voor onze two-node cluster zal de quorum configuratie 'Node and Disk Majority' zijn. Dat betekent dat de beide cluster-nodes en de witness disk elk kopieën bevatten van de cluster-configuratie en de cluster 'quorum' bezit zolang de meerderheid (twee van de drie) van deze kopieën beschikbaar zijn. Het andere shared volume dient als shared storage waarop later de VM's en virtuele hard disks worden geïnstalleerd. Als men van plan is om meer dan een VM te installeren dan beveelt Microsoft aan om een apart volume voor elke VM te creëren.

Shared storage Windows 2008 Cluster

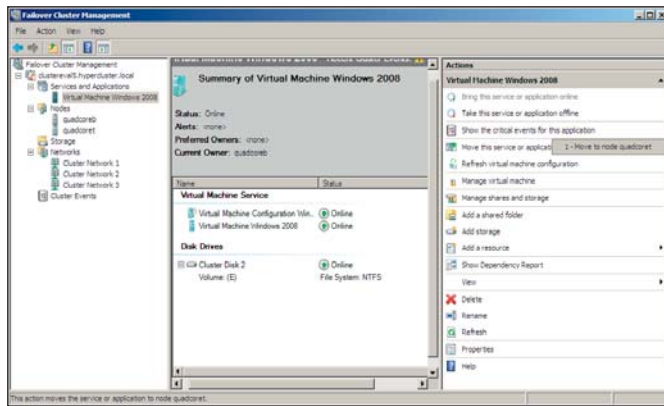
De Windows Storage Server R2 is via een dedicated IP-netwerk voor wat betreft de shared cluster storage via het iSCSI-protocol gekoppeld aan beide cluster-nodes. Aan de back-end kant van de Storage Server fungeren twee lokale Western Digital Raptor SATA II disks en een StorageTek Flexline Fibre Channel Storage Array als shared storage media, de laatste is via een Brocade SilkWorm 200E Fibre Channel-switch gekoppeld aan de Windows Storage Server. Beide cluster-nodes zijn via twee 10GbE iSCSI links verbonden met een Nortel 5530 switch. Om Hyper-V te kunnen gebruiken moet de op x64-gebaseerde processor Hyper-V en Data Execution Protection ondersteunen; deze functies moeten in de BIOS van beide cluster-nodes worden geactiveerd.

Installatie Hyper-V, Failover Clustering en VM's

We kunnen hier niet al te uitvoerig ingaan op de configuratie en installatie van een two-node failover cluster maar het proces verloopt in hoofdlijnen als volgt. Ten eerste worden de cluster-nodes, in ons geval via iSCSI, verbonden aan de shared storage; de iSCSI initiators op beide Windows Server 2008 servers moeten worden geactiveerd. Verder verdient het aanbeveling om het iSCSI-, heartbeat- en lokale netwerkverkeer op fysiek gescheiden netwerken te installeren. Na het installeren van de Failover



Afbeelding 3: Feature selectie SQL Server 2008 Enterprise.



Afbeelding 4: Verplaatsten VM binnen Cluster.

Cluster feature wordt de Hyper-V role op beide servers geïnstalleerd. Tijdens of na de installatie van de Hyper-V role moet een virtueel netwerk worden gecreëerd, deze voorziet in een hoge beschikbaarheid van VM's en verschaft toegang tot het fysieke IP-netwerk. Met de 'Create Cluster Wizard' worden de toekomstige cluster-nodes, 'quadcore1' en 'quadcore2', geselecteerd die samen de cluster gaan vormen. Daarna kan de cluster daadwerkelijk worden gecreëerd. Daarna wordt met behulp van de Hyper-V Manager een 'Virtual Machine Windows 2008' VM gecreëerd maar nog niet opgestart. Want eerst moet de VM 'highly available' worden gemaakt; bij een highly available VM wordt het beheer daarvan bestuurd vanuit de Cluster service. Voor een uitvoerige beschrijving van dit hele installatieproces zie: Microsoft's white paper 'Step-by-Step Guide for Testing Hyper-V and Failover Clustering'.

Installatie SQL Server 2008 op Hyper-V Virtual Machine

SQL Server 2008 wordt ondersteund in VM-omgevingen waarop de Hyper-V role is geactiveerd op alleen de Windows Server 2008 Standard, Enterprise en Data Center edities. Na de creatie van de VM vervolgen we met de installatie van SQL Server 2008 64-bit Enterprise Editon binnen de VM. Er kunnen maximaal vier virtuele processoren aan elke VM worden toegekend op de Windows Server 2008 32-bit en 64-bit edities. Voor zover nog niet aanwezig installeert de wizard een aantal voor SQL Server nodige applicaties: .NET Framework 2.0, SQL Server Native Client en SQL Server Setup Support Files. Op de 'Feature Selection' pagina kan een selectie worden gemaakt van de te installeren database features (zie afbeelding 3). Op de Cluster Disk Selection pagina wordt de shared cluster disk geselecteerd voor de Server failover cluster; de cluster disk is waar de SQL Server data worden geplaatst.

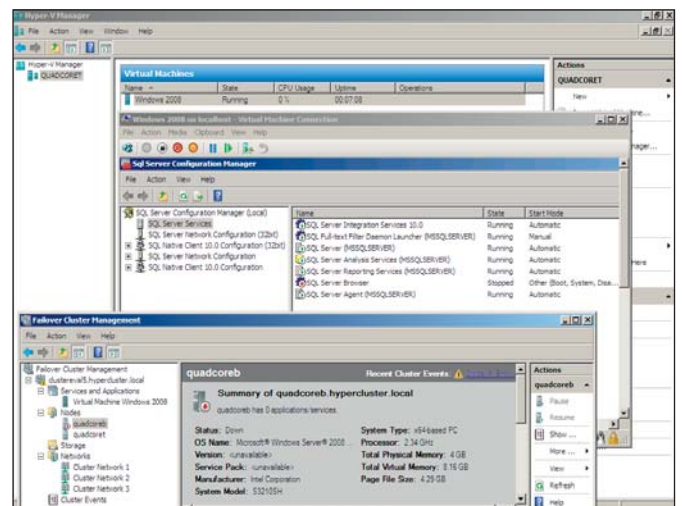
Verplaatsing SQL Server 2008 tussen cluster-nodes

Na de installatie van SQL Server 2008 op de Virtual Machine starten we de 'SQL Server Configuration Manager' op en controleren of alle SQL Services geactiveerd zijn. Het is nu tijd om de

VM met daarop een actief draaiende SQL Server real-time te verhuizen van de ene naar de andere cluster-node. Daartoe wordt in de 'Failover Cluster Management' pagina in 'Services en Applications' de VM aangeklikt en in de rechter 'Actions' kolom 'move to quadcore1' (de andere cluster-node), zie afbeelding 4. We constateren dat binnen tientallen seconden de VM is verhuisd van de ene naar de andere cluster-node. We starten de Hyper-V Manager op de 'quadcore1' cluster-node op en zien daar de geopende SQL Server configuration manager terug. Een van de toepassingen bij verhuizing van VM's is bijvoorbeeld bij een overbelaste cluster-node de daarop draaiende VM(s) te verplaatsen naar de andere cluster-node. In de volgende releases van Hyper-V zal dit waarschijnlijk (net zoals bij VMware's Resource Scheduling dat al kan) automatisch zonder tussenkomst van de beheerder kunnen gebeuren.

Failover en failback SQL Server 2008

Voor een gepland onderhoud van de server, bijvoorbeeld de upgrade van de systeem BIOS- of Interface firmware, is het soms nodig om een cluster-node fysiek uit te schakelen (een geplande downtime). De bedoeling is dat de actieve VM's dan naar de andere cluster worden overgeheveld. Om dat te simuleren schakelen we een cluster-node uit via het 'Failover Cluster Management' menu en zien dat de actieve SQL Server Services op de VM keurig zijn overgenomen (zie afbeelding 5). Iets anders is de ongeplande failover. Om dat te testen wordt de Cluster service gestopt. Daartoe selecteren we in het Failover Cluster Management menu 'Nodes' en met een rechtermuisklik de node waarop de VM draait. In 'More Actions' wordt vervolgens 'Stop Cluster Service' aangeklikt en zien we dat de VM naar de andere cluster-node wordt verhuisd. Als laatste voeren we een 'disaster' test uit waarbij een van de cluster-nodes wordt afgeschakeld als gevolg van een stroomsteking. Om dat te simuleren wordt onverbiddelijk de stroomstekker van een van de cluster-nodes uit de wandcontactdoos getrokken. Dit heeft natuurlijk niet alleen gevolgen voor het



Afbeelding 5: Uitschakeling cluster-node.

draaiende Windows Server 2008 operating systeem op de server die zal crashen, maar ook voor de VM en de daarin actief draaiende SQL Server. Op de resterende cluster-node zien we de gevolgen van deze systeemcrash. We constateren dat de VM met de daarin aanwezige Windows Server 2008 op de andere cluster-node wordt herstart, want ook de VM heeft net zoals de Windows Server (zoals was te verwachten) deze systeemcrash niet kunnen overleven. Na inloggen op de VM zien we dat de actieve SQL Server menu's zijn verdwenen en deze opnieuw moeten worden opgestart. In dit soort situaties is het natuurlijk altijd de vraag in hoeverre de database-data na de systeemcrash nog valide zijn en of naderhand de op dat moment actieve database-transacties kunnen worden teruggedraaid en/of de database-data zijn gecorrumpereerd. Want ook al maakt men gebruik van mirroring of replicatie, een 100 procent garantie biedt het niet dat data niet verloren kunnen gaan; het bewaren van database transactionele write-order integriteit en dus crash consistency tijdens een replicatieproces vormt de sleutel tot een geslaagde replicatiemethode en kan een cluster-omgeving maken of breken.

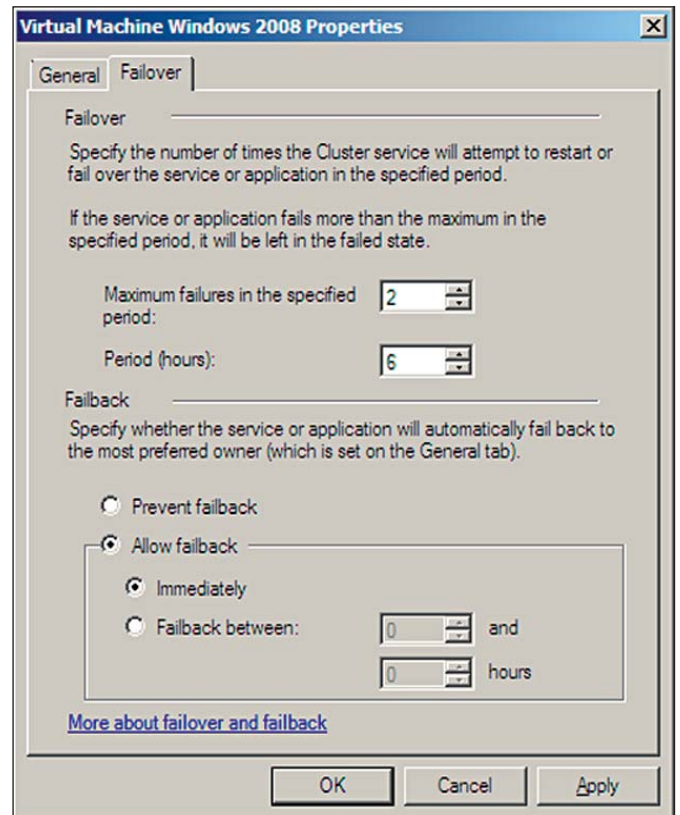
Tenslotte is er een aantal instellingen voor de Failover van VM's (zie afbeelding 6). Zo is het aantal malen in te stellen dat de Cluster Service zal trachten te starten of de failover van een service of applicatie in een bepaald tijdsbestek. Verder kan een automatische failback naar een bepaalde 'voorkeur' cluster-node worden ingesteld, waarbij een failover wordt uitgevoerd naar de cluster-node waarop oorspronkelijk de VM draaide.

Samenvatting

De combinatie van SQL Server 2008 en Windows Server 2008 met Failover Clustering biedt een hoog beschikbare database-omgeving. Binnen SQL Server 2008 zijn allerlei verbeteringen aangebracht om de beschikbaarheid van de database zelf te vergroten. Het Failover Cluster model is met de release van Windows Server 2008 op een aantal punten verbeterd. Zo is onder meer de eis voor een Quorum disk komen te vervallen, waarmee een belangrijk single point of failure binnen de Windows cluster-architectuur is weggenomen. Het woord zegt het al, 'hoog beschikbaar' wil niet zeggen 'altijd' beschikbaar want een 'five nines' beschikbaarheid (99,999 procent, dat is ongeveer vijf minuten downtime per jaar) kan zelfs met de hedendaagse clustering-technologie niet worden gehaald. Clustering biedt hoogstens een 'four nines' beschikbaarheid (ongeveer acht uur downtime). Vaak wordt de beschikbaarheid van servers verward met die van applicaties (lees databases) want een 'aantal negens' beschikbaarheid van een server betekent nog niet dat dit ook geldt voor de database-applicatie en -data. Daarvoor zijn aanvullende voorzieningen nodig, bijvoorbeeld replicatie en/of mirroring.

Ultieme database beveiliging

De komst van Hyper-V binnen Windows Server 2008 biedt SQL Server 2008 de mogelijkheid om de database binnen een Virtual



Afbeelding 6: VM instellingen.

Machine te draaien. Zoals bekend biedt server-virtualisatie een aantal voordelen. In eerste instantie zijn dat besparingen op server hardware-kosten, onderhoud, stroom en ruimte. Maar steeds meer zien we dat virtualisatie wordt gebruikt om de beschikbaarheid van systemen te vergroten. Er schuilt wel een gevaar in virtualisatie, want de uitval van een fysieke server betekent gelijk de uitval van alle daarop draaiende VM's en databases. Belangrijk is dus om de fysieke server zoveel mogelijk voor uitval te beschermen. Cluster-technologie biedt een zekere mate van beveiliging maar wanneer een cluster-node uitvalt dan gaan meestal wel alle geheugenresidente data verloren en het biedt ook geen garantie van data-integriteit. In tegenstelling tot de HA clusters die ontworpen zijn om van een fout te herstellen, zijn Continuous Availability (CA) systemen ontworpen om in eerste instantie de uitval van servers te voorkomen. Ook hier is de term CA wat misplaatst, want continu beschikbaar bestaat alleen in theorie. Het voorkomen van systeemuitval, dat ook wel bekend staat als fault-tolerant, biedt een systeem met een 'vijf negens' of meer beschikbaarheid. Een voorbeeld van een fault-tolerant systeem is de Stratus ftServer family die 99,999 procent beschikbaarheid biedt voor de Microsoft Windows Server 2008 en Hyper-V omgeving. Maar daaraan hangt wel een stevig prijskaartje en ook daar, zonder additionele middelen, is er geen 100 procent garantie dat de database 'continu beschikbaar' is en alle database-data integer zijn.

Bram Dons is onafhankelijk IT consultant.