

Microsoft Server 2012 Hyper-V Replica für Oracle Database

Inhalt

1	Executive Summary	3
2	Windows Server 2012 R2 Hyper-V Replica Technologie-Review	3
2.1	Begriffe	3
2.2	Konfigurationsoptionen:.....	4
2.3	Virtual Hard Disks / Hyper-V Replica Logs	5
2.4	Restore Points	5
2.5	Performance	7
2.6	Requirements für Datenbank	7
2.7	Paging File.....	8
3	Setup Testumgebung.....	9
3.1	Workstations	9
3.2	VM Configuration	10
3.3	Oracle Configuration	10
4	Tests	10
5	Referenzen	10

1 Executive Summary

Die Replication von virtuellen Oracle DB Hosts mittels Hyper-V Replica auf einen zweiten unabhängigen Server ist aus Redundanzgründen sinnvoll. Leider gibt es keine Best-Practice Guides oder sonstige Referenzen zum Betrieb von Oracle auf einer Hyper-V Replica Plattform.

Der Test hat gezeigt, dass die Handhabung sehr komfortabel ist.

Folgende Punkte müssen bei der Implementierung berücksichtigt werden:

- Verdopplung des Write I/O auf die physischen Harddisks des Primary Hyper-V Servers durch Pflege von Replica (Delta) Logs
- Isolierung des Replica Netzwerk-Verkehrs auf eigene Netzwerk-Interfaces zur Trennung von Applikations-Netzwerk-Verkehrs
- Erhöhung des I/O auf dem Replica-Server um Faktor 2-3 durch Verwendung von Restore-Point Historie

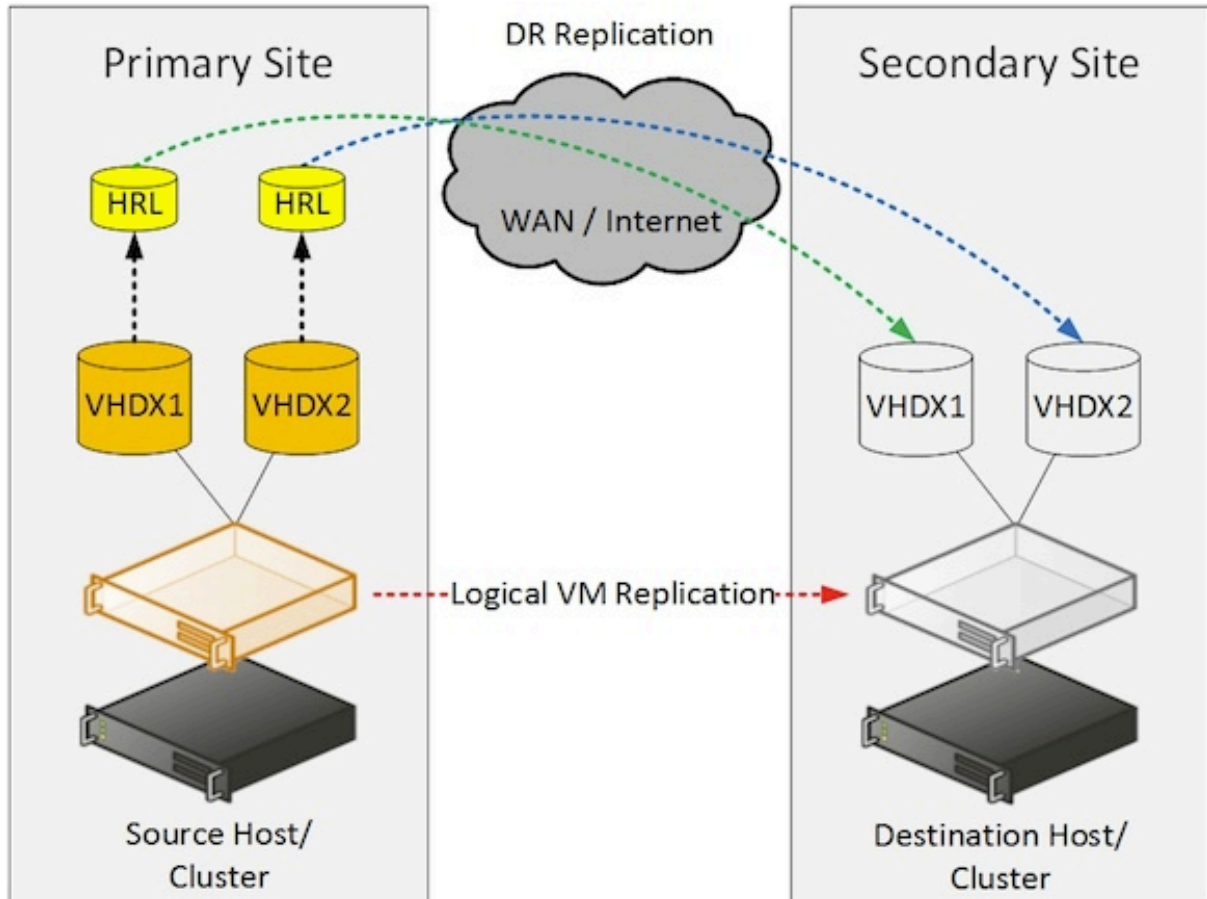
2 Windows Server 2012 R2 Hyper-V Replica Technologie-Review

Microsoft Server 2012R2 verfügt mit der Hyper-V Replica Technologie über eine Methode, virtuelle Maschinen asynchron über das Netzwerk zu replizieren. Hierfür ist kein Shared-Storage notwendig.

2.1 Begriffe

Primary Server	Physischer Server mit Windows Server 2012R2. Die virtuellen Maschinen, die auf dem Server betrieben werden können an einem Replikationsprozess teilnehmen.
Replica Server	Physischer Server mit Windows Server 2012R2.
Primary Virtual Machine	Virtuelle Maschine auf dem Primary Server, die für Replizierung konfiguriert wurde.
Replica Virtual Machine	Virtuelle Maschine auf dem Replica Server. Die VM ist offline und empfängt Replicationsdaten.

2.2 Architektur



2.3 Konfigurationsoptionen

Replication Frequency	<ul style="list-style-type: none"> • 30 Seconds • 5 Min • 15 Min
Recovery Points (Checkpoints):	<ul style="list-style-type: none"> • Maintain only the latest recover point • Create addition hourly recovery points (Restore Point History) (Definition von Coverage in Stunden) <ul style="list-style-type: none"> ○ VSS Snapshot frequency (Stunden)
Compress the data that is transmitted over the network	<ul style="list-style-type: none"> • Yes / No

Das komprimieren der Daten ist vor allem bei WAN Replikation wichtig. Im LAN Bereich kann bei Isolation des Replica-Traffics auf eigene Interfaces und Benutzung von GbE oder 10GbE darauf verzichtet werden. Die Komprimierung erhöht die CPU Belastung und verringert die Netzwerk-Belastung.

2.4 Virtual Hard Disks / Hyper-V Replica Logs

Die virtuellen Hard Disks sind als VHDX Files angelegt. Für jede zu replizierende Disk und Replikations-Run legt HyperV im selben Verzeichnis eine HRL (HyperV Replica Log) Datei an. Bei einem Write der VM wird immer in beide Files geschrieben. Die Anzahl der Writes verdoppelt sich also mit Hyper-V Replica. Wenn die Replikation pausiert wird, wächst die HRL Datei an, weil diese die Delta-Änderungen zwischenspeichert. Wenn die Replikation dann wieder fortgesetzt wird, sendet Hyper-V den Inhalt der Datei optional mit Komprimierung zum Replica-Server und löscht anschließend die Datei. Bei jeder Replikationsoperation (je nach Einstellung alle 30 Sek, 5 Min, 15 Min) wird das HRL File mit einem neuen getauscht und das vorherige HRL File wird zum Replica server transferiert und anschließend gelöscht.

Auf dem Replica-Server werden die HRL Files in in umgekehrter Reihung gelesen, sodaß nur die letzten Writes eines Blocks auf die VHDX Dateien appliziert werden.

Die Anzahl der HRL Logfiles auf dem Replica Host hängt von der Option „Recovery History“ ab. Mit der Option „Maintain only the latest recovery point“ wird nur eine HRL Datei pro VHDX File verwendet und diese Änderungen werden in die VHDX Dateien appliziert. (merge)

2.5 Restore Points

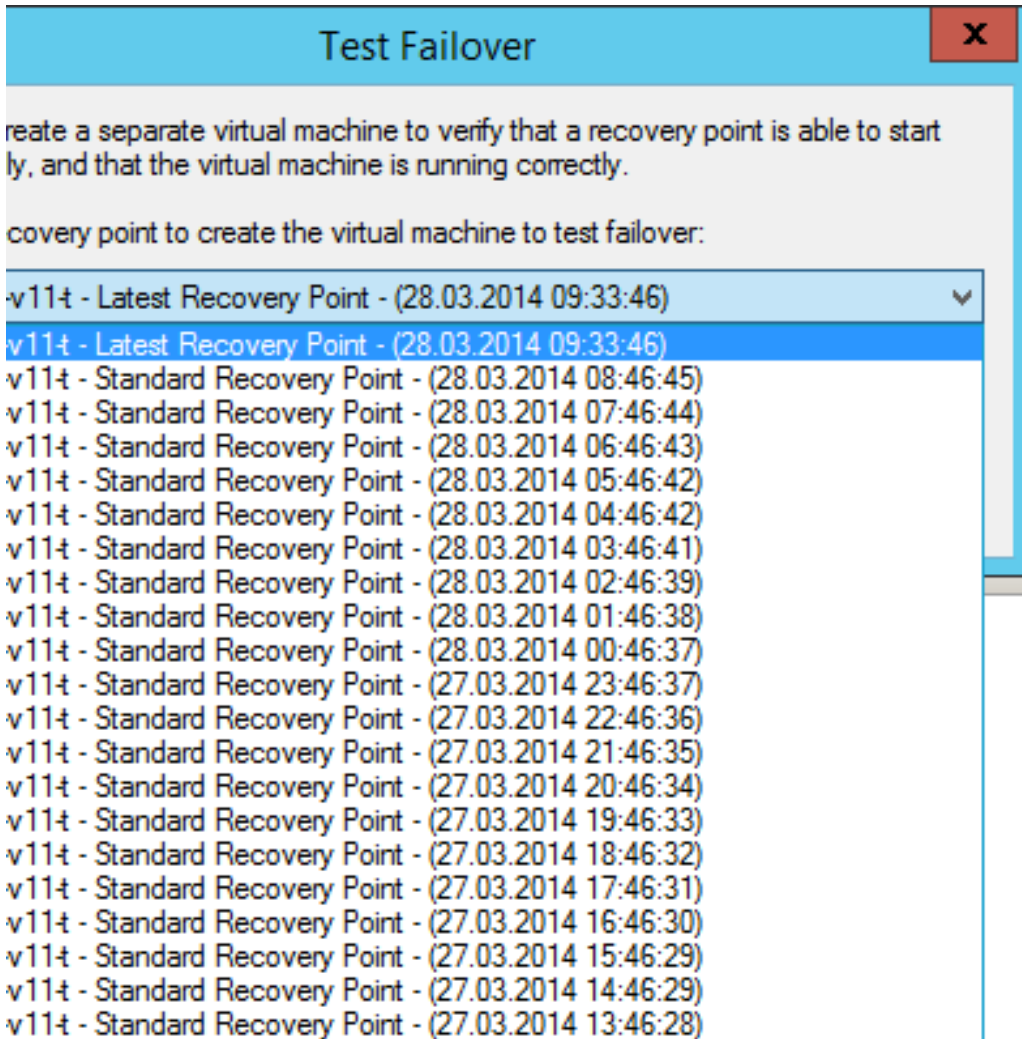
In Windows Server 2012 R2 können maximal 24 Restore Points (einen RP pro Stunde) konfiguriert werden. Zusätzlich kann noch VSS Snapshot Frequenz (in Stunden) angegeben werden.

Damit ist es möglich, bei einem Failover auf den Replica Server nicht nur den aktuellen Stand zu erreichen, sondern auch frühere Stände.

Beispiel:

- aktueller Zeitpunkt: 15:27
- Latest Recovery Point: 15:27
- Standard Recovery Point: 14:45, 13:45, 12:45, 11:45, 10:45, 09:45, etc

Auf dem Replica-Server wird für jeden Restore-Point und Harddisk eine sogenannte HRU (Hyper-V Replica Undo Log) Datei erstellt.



Dateien bei Verwendung von Restore Point Historie:

```

28.03.2014 05:48      286,105,600 db-v11-t_9A6606B1-6555-4B4D-B764-C9A067459083.hru <---- Hyper-V Replica Undo Log Restore Point 05:47
28.03.2014 05:48      13,406,208 db-v11-t_70AC2E71-7C13-4823-AF7E-31DF83DF881B.hru <---- Hyper-V Replica Undo Log Restore Point 05:47
28.03.2014 05:48    1,166,475,264 db-v11-t_8B1BE453-7874-4B4D-A7A5-7031D63210D2.hru <---- Hyper-V Replica Undo Log Restore Point 05:47
28.03.2014 06:47       10,768,384 db-v11-t_3ADA2D9C-B383-4200-9F9C-2E87B8A78BE1.hru <---- Hyper-V Replica Undo Log Restore Point 06:47
28.03.2014 06:47     282,828,800 db-v11-t_E65531D4-DB2E-44F0-AAA6-51985DAAEDEC.hru <---- Hyper-V Replica Undo Log Restore Point 06:47
28.03.2014 06:47    1,162,821,632 db-v11-t_23A85103-2FE3-4E07-80E7-24D4C1BCCEAE.hru <---- Hyper-V Replica Undo Log Restore Point 06:47
28.03.2014 07:47     283,938,816 db-v11-t_132A3BF7-4E53-49CA-AF5B-5E8A35FEE4D9.hru <---- Hyper-V Replica Undo Log Restore Point 07:47
28.03.2014 07:47     11,108,352 db-v11-t_FE114BB6-F792-4455-9355-82B75544CC8F.hru <---- Hyper-V Replica Undo Log Restore Point 07:47
28.03.2014 07:47    1,192,370,176 db-v11-t_F1F5CC46-F47B-4232-A6EF-94C5D45B28A6.hru <---- Hyper-V Replica Undo Log Restore Point 07:47
28.03.2014 08:48       10,891,264 db-v11-t_F5E2283D-393E-4EE1-A326-6127A0E2048B.hru <---- Hyper-V Replica Undo Log Restore Point 08:47
28.03.2014 08:48    1,158,119,424 db-v11-t_3B5C10F6-17AB-42DF-81F5-6874E143FE7E.hru <---- Hyper-V Replica Undo Log Restore Point 08:47
28.03.2014 08:48     281,767,936 db-v11-t_C8BFB842-4F3C-424A-BB9C-E46B777E288.hru <---- Hyper-V Replica Undo Log Restore Point 08:47
28.03.2014 09:44       73,768,960 db-v11-t_952E3AB-7058-42AD-8519-EF1C0708EBA.hru <---- Hyper-V Replica Undo Log Restore Point 09:47
28.03.2014 09:44    21,479,030,784 db-v11-t-E.vhdx <---- VHDX Virtual Drive
28.03.2014 09:44       9,801,728 db-v11-t_36F51073-CC6D-4449-8DDD-496725CB97ED.hru <---- Hyper-V Replica Undo Log Restore Point 09:47
28.03.2014 09:44     967,081,984 db-v11-t_D187DD30-A9B2-496C-A64B-12B20E35DFD4.hru <---- Hyper-V Replica Undo Log Restore Point 09:47
28.03.2014 09:44    25,673,334,784 db-v11-t-C.vhdx <---- VHDX Virtual Drive
28.03.2014 09:44    214,752,559,104 db-v11-t-D.vhdx <---- VHDX Virtual Drive
28.03.2014 09:44       4,096 db-v11-t_8D4BD970-8628-46BB-9E7A-54B2A53BE64E.hrl <---- Hyper-V Replica Log
28.03.2014 09:44     335,872 db-v11-t_9C4C8749-1D0D-4500-849D-20857097D9F3.hrl <---- Hyper-V Replica Log
28.03.2014 09:44       4,096 db-v11-t_E511ED7A-77FB-4975-9DD3-BAAF962642E1.hrl <---- Hyper-V Replica Log
28.03.2014 09:45      86,528 db-v11-t_B3CD9A96-0E9A-457C-88DD-D05CA725DFA8.hrl <---- Hyper-V Replica Log
28.03.2014 09:45       4,096 db-v11-t_7959507C-9988-41B5-ACE0-B4FEDAE2B3D.hrl <---- Hyper-V Replica Log
28.03.2014 09:45     335,872 db-v11-t_ECC72122-E1AE-4B4C-B80E-7CE7A6489F2D.hrl <---- Hyper-V Replica Log
28.03.2014 09:45     335,872 db-v11-t_B55D0DC3-1016-48D2-AEAC-C8B857D4E94A.hrl <---- Hyper-V Replica Log
28.03.2014 09:45     61,440 db-v11-t_57357642-2458-4708-B65E-F8AE74D5EB6D.hrl <---- Hyper-V Replica Log
28.03.2014 09:45       4,096 db-v11-t_2C9BE895-30B3-4C94-9194-1CF741744C21.hrl <---- Hyper-V Replica Log

```

Die Hyper-V Replica Logs enthalten die letzten erhaltenen Änderungen, die noch nicht vollständig auf die VHDX Files appliziert wurden.

Die VHDX Files enthalten den aktuellen Stand, exklusive der lokal vorhandenen HRL Files. Beim Failover auf den aktuellsten Stand müssen nur noch die HRL Files appliziert werden und anschließend ist es möglich, die VM zu starten.

Beim Failover auf einen früheren Restore-Point müssen ALLE HRU Files bis zum jeweiligen Restore-Point auf die VHDX Files appliziert werden.

Beispiel:

Beim Failover auf den Restore Point von 06:47 Uhr müssen die HRU Files von 09:47, 08:47, 07:47 und 06:47 appliziert werden. Je weiter der gewünschte Restore-Point in der Vergangenheit liegt, desto länger dauert dies. Falls eine HRU Datei fehlerhaft ist, können ältere Restore-Points nicht mehr erreicht werden. Z.B. bei einer korrupten HRU Datei von 07:47 kann der Restore-Point um 06:47 nicht mehr hergestellt werden.

Bei Verwendung einer Recovery Point Historie ist mit einer Erhöhung der I/O Belastung auf dem Replica Server um den Faktor 2-3 zu rechnen. Dies ist von Nachteil, wenn mehrere VMs in beide Richtungen repliziert werden.

z.B.

VM1: Primary Host HV1, Replica Host HV2

VM2: Primary Host HV2, Replica Host HV1

2.6 Performance

Da die Replikation asynchron arbeitet, sollte der Performance-Einfluss gering sein. Es wird angenommen, dass zwar der Datentransfer über das Netzwerk asynchron durchgeführt wird, jedoch die HRL Dateien synchron geschrieben werden. Im Bereich I/O wird deshalb die größte Beeinträchtigung erwartet.

Es ist zu prüfen, ob die HRL Dateien physikalisch von den VHDX Dateien getrennt und auf unterschiedlichen Harddisks gespeichert werden können um die Platten-Belastung zu reduzieren. Bislang gibt es keine Hinweise auf die Möglichkeit der Trennung.

2.7 Requirements für Replikation von Datenbanken

Es gibt leider keine Referenzen oder BestPractices für die Verwendung von Hyper-V Replica für Oracle Datenbanken. Bei Verwendung von SQL Server wird allerdings das Setzen des Parameters EnableWriteOrderPreservationAcrossDisks benötigt.

Zur Sicherheit wurde dies auch bei Oracle durchgeführt:

PowerShell als Administrator auf Primary Server starten:

```
Set-VMReplication -VMName <vm-name> -EnableWriteOrderPreservationAcrossDisks 1
```


Erklärung für Parameter:

Determines whether all virtual hard disks selected for replication are replicated to the same point in time. This is useful if the virtual machine runs an application that saves data across virtual hard disks (for example, one virtual hard disk dedicated for application data, and another virtual hard disk dedicated for application log files).

2.8 Paging File

Virtual Disks mit Paging Files sollten aus der Replication exkludiert werden. Es wird empfohlen, eigene VHDX Files für Paging File zu erstellen und diese nicht zu replizieren.

2.9 Monitoring Replication Health

Replication	
Replication State:	Replication enabled
Replication Mode:	Replica
Current Primary Server:	de-hv01
Current Replica Server:	de-hv02
Replication Health:	 Normal
Statistics for past 13 Hours 25 Minutes	
From time:	27.03.2014 20:59:58
To time:	28.03.2014 10:25:56
Average size:	12 MB
Maximum size:	24 MB
Average latency:	0:00:06
Errors encountered:	0
Last synchronized at:	28.03.2014 10:25:46
Test Failover	
Test failover status:	Not Applicable
Last test failover initiated at:	Not Applicable

Beim Monitoring des „Replication Health“ ist zu sehen, wann die letzte Synchronisation stattgefunden hat. Zudem wird die Maximale Delta-Übertragung (Maximum Size) sowie die durchschnittliche Delta-Übertragungsmenge (Average Size) angegeben. Weiters ist die Durchschnittliche Latenz (Durchschnittliche Anzahl von Sekunden, bis die HRL Daten auf dem Replica-Server empfangen wurden) angegeben.

2.10 Failover Testing

Es wurde ein Failover-Test durchgeführt, bei 25 Sekunden nach dem letzten Sync der Primary-Server stromlos gemacht wurde. Das manuelle Failover auf den Replica Server hat funktioniert. Es kam zu einem Datenverlust von ca. 25 Sekunden.

Es gab allerdings einige Auffälligkeiten:

- Beim Failover wird ein Checkpoint durchgeführt, der AVHDX Dateien erzeugt. Dies führt zu hohem I/O Aufkommen für mehrere Minuten
- Das Instance Crash Recovery hat ca. 10 Minuten gedauert. Während dieser Zeit ist die Datenbank nicht verfügbar.
- Wenn zu einem späteren Zeitpunkt der Primary Server gestartet wird, startet automatisch auch die VM. Dies führt dazu, dass auf beiden Servern die VM aktiv läuft und führt zum IP Konflikt und zur gefährlichen Situation, dass eventuell Applikationsverbindungen auf beide VMs stattfinden. Dies muss unbedingt verhindert werden.

2.11 VM Configuration

Die VM Configuration (# vCPUs, Memory, etc.) muss auf Primary Server und Replica Server synchron gehalten werden. Änderungen der Configuration werden nicht auf dem Replica-Server aktualisiert.

3 Setup Testumgebung

3.1 Workstations

Workstation:	2 x Fujitsu Esprimo C910-L
RAM:	8 GB RAM (7.88 Usable)
Harddisk:	Seagate ST500LM000 (500GB, SATA 6 GBit, 5400U/min)

	SSDHybridfestplatte (8GB MLC, 64MB DRAM Cache
Netzwerk:	Gigabit Ethernet
OS:	Windows Server 2012 R2 Datacenter

3.2 VM Configuration

Guest Configuration:

Startup RAM: 5GB

Processors: 4 Virtual Processors

3.3 Oracle Configuration

```
*.job_queue_processes=0  
*.memory_max_target=4G  
*.memory_target=4G  
*.nls_language='AMERICAN'  
*.nls_territory='AMERICA'  
*.open_cursors=300  
*.open_links=6  
*.open_links_per_instance=6  
*.pga_aggregate_target=512M  
*.processes=150  
*.query_rewrite_enabled='FALSE'  
*.remote_login_passwordfile='EXCLUSIVE'
```

4 Tests

Das Testsystem verfügt nur über eine einzelne Festplatte. Leider war die Festplatte schon bei relativ geringer Datenbank-Belastung schon zu 100% ausgelastet. Aus diesem Grund ist die Fortsetzung des Tests nicht sinnvoll. Es konnten trotzdem ausreichende Kenntnisse über das Verhalten erlangt werden.

5 Referenzen

<http://www.aidanfinn.com/?p=12147>

<http://social.technet.microsoft.com/Forums/windowsserver/en-US/dd854781-0d28-4917-ada9-04c8ea7a6e8b/hyperv-replication-move-hrl-file-to-another-location?forum=winserverhyperv>

<http://support.microsoft.com/kb/956893>

<http://technet.microsoft.com/en-us/library/hh848513.aspx>

http://download.microsoft.com/download/6/1/D/61DDE9B6-AB46-48CA-8380-D7714C9CB1AB/Best_Practices_for_Virtualizing_and_Managing_SQL_Server_2012.pdf

<http://technet.microsoft.com/en-us/library/hh848543.aspx>