

fänglichen Schwierigkeiten mit dem SnapCenter oder beispielsweise eine in ihrer Gänge neu implementierte Überwachung über den Oracle Enterprise Manager.

Die Zielsetzung der hohen Verfügbarkeit wurde erreicht. Bei einer geplanten Stromabschaltungs-Übung des gesamten Primary-Rechenzentrums beispielsweise wurden die Datenbanken und das Storage absichtlich nicht in den normalen Ablaufplan integriert. Dabei hätte man vorab einen Data Guard Switchover durchführen können; stattdessen konnte problemlos ein Failover initiiert werden, nachdem der Strom der gesamten Infrastruktur im ersten RZ gekappt wurde. Es sei erwähnt, dass zu diesem Zeitpunkt die Datenbanken noch keinen produktiven Status hatten. Durch die Aktivierung von Flashback auf allen Datenbanken ist man in der Lage, nach einem Failover durch ein Reinstatement das Data Guard zeitnah wieder in einen konsistenten Zustand zu versetzen. Hierbei wird übrigens der Modus „Maximum Availability“ eingesetzt.

Auch im Rahmen von Wartungsarbeiten zeigten sich die Vorteile des Data Guard insofern, als dass der Juli-PSU durch Standby First Apply nahezu ohne Downtime für die Applikationen implementiert werden konnte. Unschön ist weiterhin das leidige Thema der Java-Patches. Hier musste zwingend eine Downtime von etwa zehn Minuten her,

um das Datapatch in die mit der OJVM betriebene Datenbank korrekt einzuspielen.

Das Backup- und Recovery-Verfahren durch das SnapCenter – wohlgermerkt das erste offizielle Release von NetApp – erledigt seine Aufgaben zuverlässig. Um die Datenbank per Snapshot zu sichern, wird diese kurzzeitig in den Backup-Modus („alter database begin backup;“) gesetzt und die „Verpointerungen“ auf Storage-Ebene gesetzt. Hat man sich einmal in der browserbasierten GUI zurechtgefunden, kann man schnell und einfach neben den regelmäßigen Sicherungsjobs auch Tasks wie Point-In-Time-Recovery oder gar einen Clone einer laufenden Datenbank erstellen. Dieser Clone kann sogar zeitlich in der Vergangenheit liegen, je nach Vorhaltezeit der Snapshots. Somit ist es durch SnapVault innerhalb weniger Minuten möglich, auch einen bis zu 35 Tage alten Snapshot für den Aufbau eines Clones zu verwenden. Wichtig in der Gesamt-Konstellation ist jeweils der Einsatz einer Secondary-Storage, worauf die aktuell erstellten Snapshots der Datenbank per SnapVault-Funktion kopiert werden, da ein Snapshot allein noch kein sicheres Backup darstellt.

Leider war es – entgegen den Angaben von NetApp – nicht möglich, die Standby-Datenbanken über das SnapCenter brauchbar zu sichern. Dies wäre nur über ein Offline-Backup möglich gewesen.

Alles in allem wurde eine stabile, zuverlässige Oracle-Infrastruktur geschaffen, die ihrer Bezeichnung „always online“ bisher voll gerecht wird.



Tim Hensel
tim.hensel@devk.de



Johannes Ahrends
johannes.ahrends@carajandb.com

Oracle VM und Virtual Shared Storage

Nico Henglmüller und Dr. Thomas Petrik, Sphinx IT Consulting GmbH

Die Architektur von Oracle VM sieht die Verantwortlichkeit der Speicherverwaltung bei externer Hardware. Eine solche ist für Klein- und Mittelbetriebe oft unerschwinglich. Dies führte die Autoren zur Frage: „Wie kann eine neuartige Infrastruktur konzipiert werden, die gleichzeitig hochausfallsicher, universell einsetzbar und einfach zu betreiben ist; noch dazu kostengünstiger als eine vergleichbare Enterprise-Lösung?“ Eine ambitionierte Truppe machte sich auf die Suche und berichtet in diesem Artikel von der verwendeten Technologie und vom innovativen Ergebnis.

Zu Beginn stand die Idee, einen hochverfügbaren und leistbaren, Multipurpose Cluster mit Hard-Partitioning-Unterstüt-

zung von Oracle-Datenbanken zu entwickeln. Dieser Idee folgte ein Prototyp, bestehend aus zwei Oracle-VM-Knoten

und einem Virtual Shared Storage. Die eingesetzten Knoten basieren auf Standard-x86-Architektur und verwenden ei-

nen Verbund aus redundanten Festplatten als Speicher.

Die Überlegung war, dass im Falle eines unvorhergesehenen Ausfalls eines Knotens sichergestellt ist, dass der zweite Knoten mit minimaler Unterbrechung (Cold failover) den Betrieb fortsetzt. Das Konzept ging auf. Oracle VM bringt für diesen Einsatzzweck die notwendigen Kern-Features mit: High Availability out of the box mit OCFS2 als Cluster-Filesystem und Heartbeat-Komponente sowie eine zentrale Verwaltung des gesamten Clusters in Form von Oracle VM.

Das in *Abbildung 1* dargestellte Design erlaubt es, dass die Speicher-Architektur für den Hypervisor transparent ist. Dem Oracle VM Server wird ein blockbasiertes Device zur Verfügung gestellt, auf dem ein OCFS2-Dateisystem für den Clusterbetrieb von Oracle VM angelegt ist. Durch den Einsatz von Virtual Shared Storage werden Leseoperationen immer lokal durchgeführt. Im Gegensatz hierzu werden schreibende Operationen erst als geschrieben angesehen, sobald die Daten auf beiden Knoten persistiert wurden.

Netzwerk-Architektur

Bei der Gestaltung des Netzwerks ist es notwendig, eine redundante Verbindung

exklusiv für die Synchronisation zwischen den Knoten miteinzubeziehen (*siehe Abbildung 2*). Um einem Bottleneck entgegenzuwirken, werden diese zwei Verbindungen im Link-Aggregation-Modus mit jeweils 10 GBit/s entweder direkt oder über einen „low latency“-10-GBit-Switch verbunden.

Das Netzwerk für die virtuellen Maschinen sowie das Management-Netzwerk werden in einem Active-Passive-Failover-Modus, mit einer Bandbreite von mindestens einem GBit/s betrieben. Das Management und das VM-Netzwerk werden unverändert im Sinne von Oracle VM verwendet. Darüber hinaus kann der VSS-Link durch Mapping eines V-LANs als 20-Gbit-„low latency Interconnect“ für virtualisierte Active-Active-Cluster wie beispielweise Oracle RAC verwendet werden.

Reaktionen bei unerwarteten Hardware-Ausfällen

Virtual Shared Storage ermöglicht es, bei einer unerwarteten Betriebs-Unterbrechung des kompletten Storage eines Knotens alle I/O-Operationen transparent auf das Storage des zweiten Knoten umzuleiten. Die Folge ist, dass der Betrieb nicht unterbrochen wird und sich die Daten zu

jedem Zeitpunkt in einem konsistenten Zustand befinden.

Wie bei allen Clustern ist die Split-Brain-Problematik auch beim Zwei-Knoten-Cluster ein heiß diskutiertes Thema. Unter „Split Brain“ versteht man den kompletten Ausfall der Kommunikation zwischen den Knoten. Die Problematik ist, dass sich beide Knoten als neuer Master deklarieren. Dies führt zu Inkonsistenzen, da beide Datenmengen weiterhin gepflegt werden, allerdings keine Synchronisation zwischen diesen stattfindet.

Die Herausforderung bei einem unvorhergesehenen Hardware-Ausfall ist, schnellstmöglich die Abwesenheit eines Knotens zu erkennen und die Datenkonsistenz zwischen den Knoten zu gewährleisten. Wird nun die Replikationsverbindung komplett unterbrochen, muss der zu schreibende Datenverkehr auf Knoten 1 eingefroren werden. Anschließend stellen sich die folgenden Optionen:

- *Szenario 1*
Falls das Management-Netzwerk ebenfalls keine Verbindung aufweist, wird Knoten 1 heruntergefahren, um die Daten-Integrität sicherzustellen.
- *Szenario 2*
Falls sich die Knoten über das Management-Netzwerk erreichen, wer-

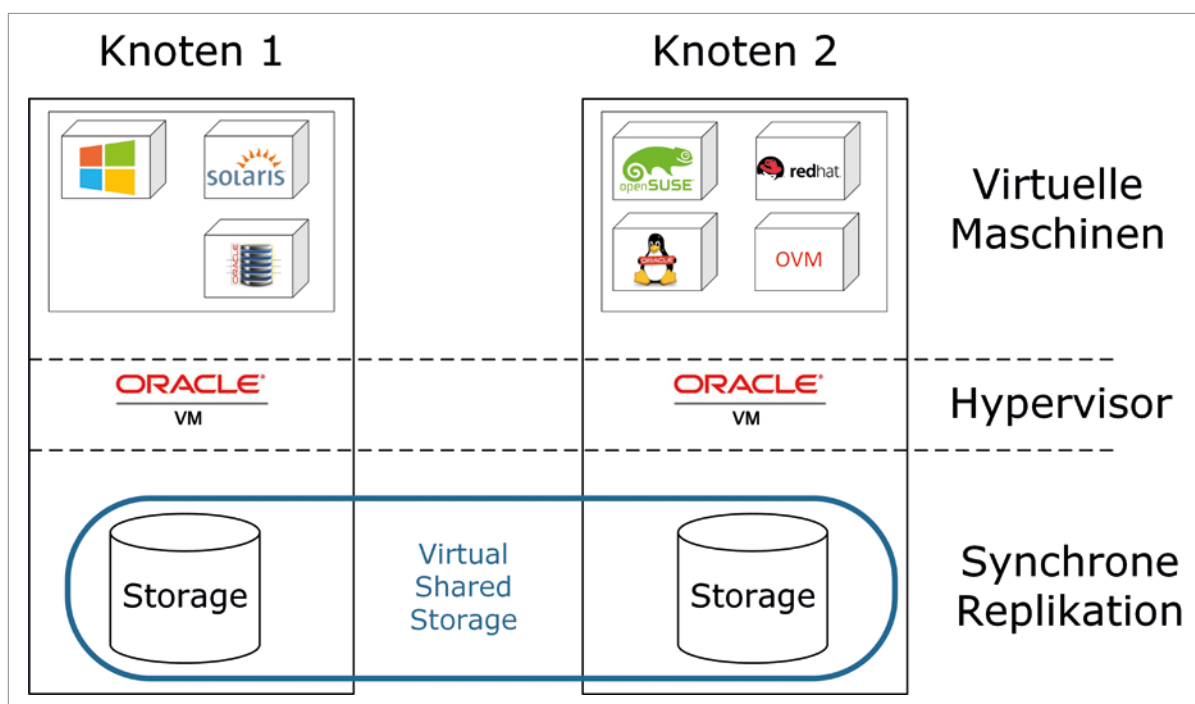


Abbildung 1: Aufbau des Zwei-Knoten-Clusters mit Virtual Shared Storage

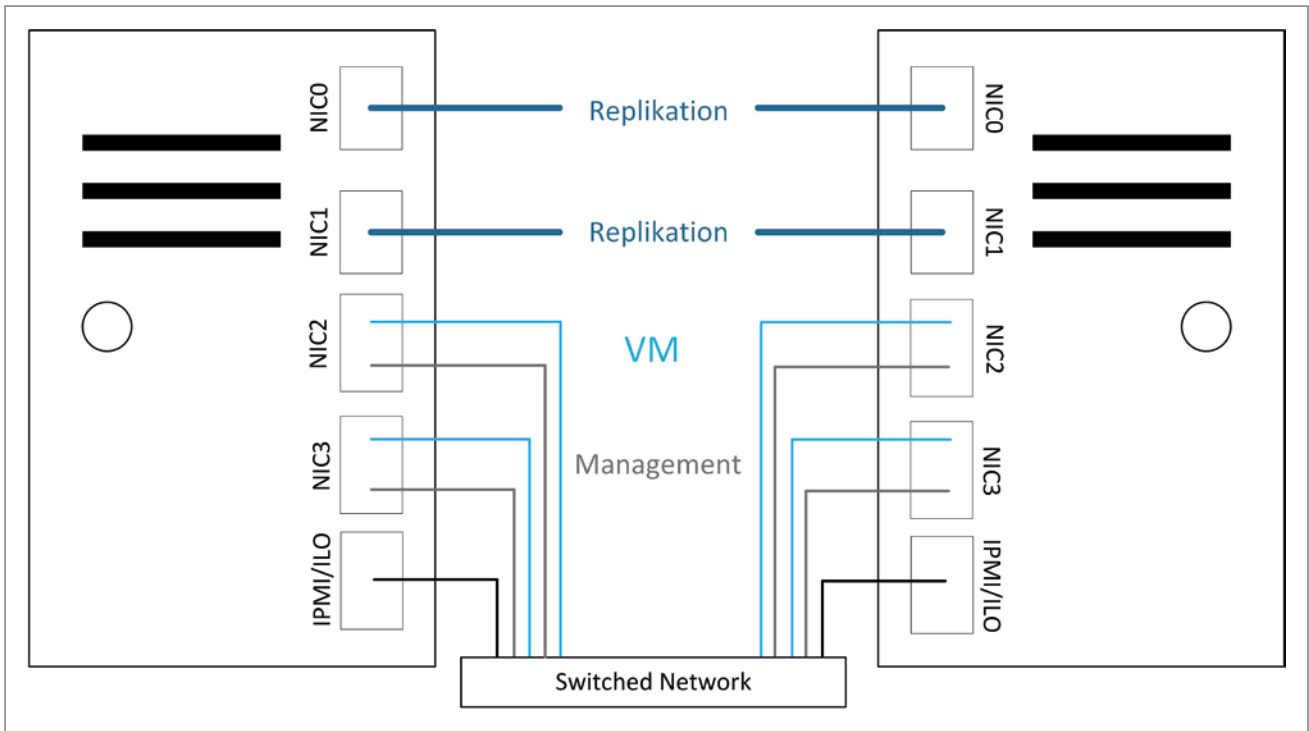


Abbildung 2: Netzwerk-Architektur

den beide Knoten heruntergefahren, da die Synchronisation über das Replikations-Netzwerk nicht mehr erfolgen kann.

• Szenario 3

Falls das Management-Netzwerk erreichbar, jedoch Knoten 2 nicht erreichbar ist, wird Knoten 1 zum neuen Master des Clusters. In der Rolle des Masters werden die virtuellen Maschinen, die als hochverfügbar markiert sind, auf Knoten 1 gestartet. Der Status von Knoten 2 bleibt weiterhin unklar. Falls dieser ebenfalls eine aufrechte Verbindung in das Management-Netzwerk besitzt, ist dies die einzige Situation, in der ein Split Brain unausweichlich ist.

rungsziel wird ein in OVM eingebundener Netzwerkspeicher verwendet. Die Sicherung erfolgt, wie die Wiederherstellung von virtuellen Maschinen, genauso nahtlos im Oracle VM Manager und erfordert keinerlei zusätzliche Software.

Fazit

Gerade im digitalen Zeitalter können sich auch kleine Unternehmen Unterbrechungen des Betriebs oder Service-Ausfälle nicht erlauben. Die logische Schlussfolgerung ist daher, auf intelli-

gente, ausfallsichere Lösungen zu setzen. Das hier vorgestellte schlanke Produkt BlueBoxx ermöglicht es heute Klein- und Mittelbetrieben, Enterprise-Funktionalitäten zu nutzen. Gerade im Umfeld von Oracle sind smarte Lösungen gefragter denn je.

Die Virtualisierung mithilfe von Oracle VM ermöglicht die vollständige Zertifizierung der Supportkette von Oracle- und Microsoft-Produkten. Durch die Anerkennung der Hard-Partitioning-Fähigkeiten des Hypervisors seitens Oracle birgt die Lizenzierung von Oracle-Datenbanken großes Potenzial.

Betrieb

Für den Betrieb sind Werkzeuge für Monitoring, Backup und Recovery sowie Updates und Support essenziell. Mithilfe des Monitorings ist eine detailgenaue Überwachung des Cluster-Zustands ersichtlich und erlaubt eine übersichtliche Erstdiagnose. Die Backup-Lösung ist in den Oracle VM Manager voll integriert und individuell konfigurierbar. Als Siche-



Nico Henglmüller, BSc
nico.henglmuller@sphinx.at



Mag. Dr. Thomas Petrik
thomas.petrik@sphinx.at