

Cloning-Lösungen im Überblick

Miguel Anjo in Zusammenarbeit mit Ludovico Caldara, Catalin Motaitianu und Dawid Wojcik, Trivadis AG

Datenbanken werden immer größer und größer. Häufig sind die Archivierung und der Lebenszyklus der Daten nicht im ursprünglichen Design eingeplant. Das führt zu Problemen, wenn Entwickler Kopien zum Testen, zur Qualitätssicherung, zur Kontrolle oder aus anderen Gründen benötigen. Entwickler, Datenbank-Administratoren und Fachabteilungen brauchen eine aktuelle Kopie der Datenbank im Zugriff, um Patches, Tests oder Reports ohne Auswirkungen auf die Produktion durchführen zu können.

Die Auswahl und Implementierung einer Cloning-Lösung sollte entsprechend sorgfältig erfolgen, damit über das reine Datenbank-Cloning hinausgehende zusätzliche Anforderungen berücksichtigt werden können: Daten maskieren, isolieren oder mit weiteren Systemen oder Datenbanken synchronisieren sind mögliche erweiterte Funktionen. Parameter wie die Größe der Datenbank, die Übertragungsgeschwindigkeit, die verfügbare Festplattenkapazität sowie der mögliche Zeitaufwand für das Klonen spielen ebenfalls eine wichtige Rolle. Das Wiederherstellen von Sicherungen, ein RMAN-Duplikat oder Export/Import sind mögliche Lösungen, wenn die Datenmenge begrenzt, die Zeitdauer für das Klonen kein Problem und das Budget begrenzt ist.

Oracle verbessert laufend die Cloning-Funktionen des Enterprise Manager. Die-

ser bietet eine einfach zu bedienende gute Lösung. Firmen wie Delphix oder ActifIO haben sich auf virtuelle und physische Anwendungen spezialisiert, um Klone sowohl auf Oracle- als auch auf Speicher-Technologien schnell zu erstellen. Man sollte jedoch bedenken, dass Klone nicht immer die beste Lösung für die Bedürfnisse des Unternehmens sind und man deshalb auch andere Alternativen betrachtet werden sollten.

Wie Datenbank-Cloning funktioniert

Datenbank-Cloning ermöglicht das Erstellen einer Kopie der Produktionsdaten mit möglichen Untermengen, Maskierung und Isolation (siehe Abbildung 1). Aus einer

Sicherung oder der eigentlichen Datenbank entsteht eine erste Kopie im Staging-Bereich. Dies kann eine vollständige Kopie oder nur eine Teilmenge der Tablespace sein. In dieser Kopie erfolgt das Maskieren sensibler Daten. Zudem wird das Senden von E-Mails an bestehende Kunden sowie die Verbindung zu anderen Datenbanken unterbunden. Am Ende steht eine Master-Kopie, die sich beliebig klonen lässt.

Zuvor sollte man jedoch möglichen Anwendungsfälle und deren Häufigkeit planen. Dabei hilft der folgende Fragenkatalog:

- Wie wird die geklonte Datenbank genutzt?
- Wie oft werden Klone benötigt?
- Wie viele unabhängige Klone sind notwendig?
- Wie schnell sollen die Klone vorliegen?

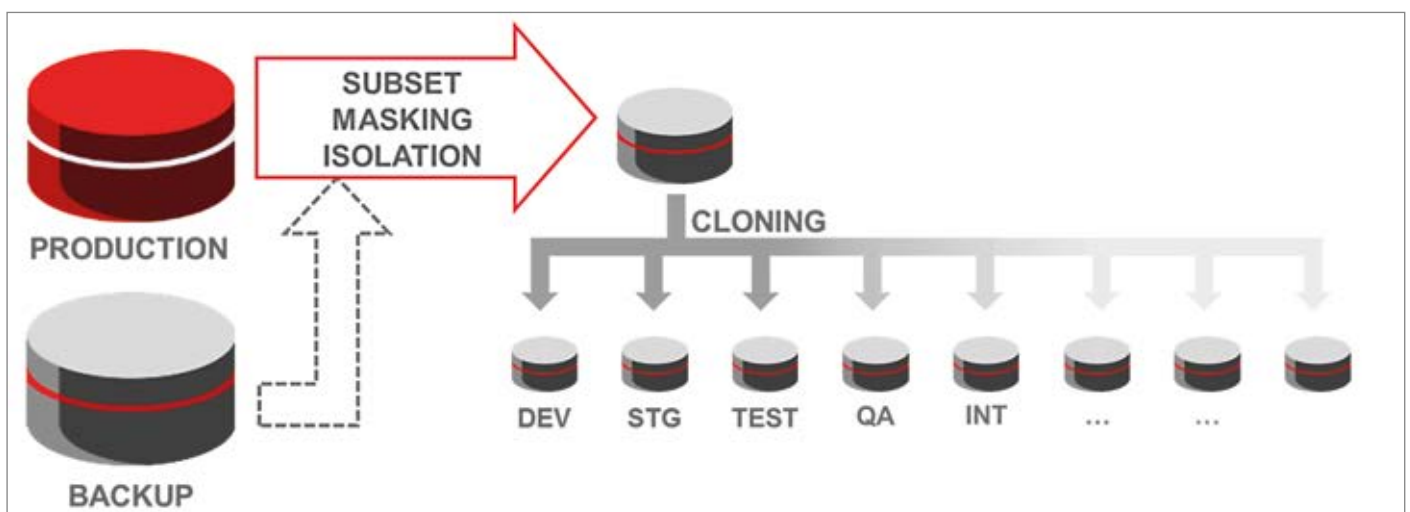


Abbildung 1: Datenbank-Cloning

- Wie lange wird ein Klon benötigt?
- Ist Daten-Anonymisierung erforderlich?
- Ist eine Self-Service-Lösung für die Bereitstellung der Datenbank-Kopien notwendig?
- Gibt es im Unternehmen andere Datenbanken oder Anwendungen mit ähnlichen Anforderungen?
- Hat das Unternehmen bereits ein Storage-System mit Cloning-Features?
- Sind Features wie das Klonen eines Klons zu beliebigen Zeitpunkten erwünscht?
- Gibt es andere Datenbanken/Systeme, die zur gleichen Zeit geklont werden müssen?
- Wird die Enterprise Edition der Datenbank verwendet?
- Ist der Oracle Enterprise Manager Cloud Control im Unternehmen im Einsatz? Wenn ja, welche Management Packs sind lizenziert?

Frei verfügbare Lösungen

„RMAN Restore“ bringt den großen Vorteil, gleichzeitig die Backups zu überprüfen. RMAN kann auch eine vollständige Kopie der Datenbank mit der gleichen Fragmentierung erstellen, was für Performance-Tests wichtig ist. Um das Klonen zu automatisieren, bedarf es jedoch einiger Skripte. Der erforderliche Speicherplatz sowie die Zeit für das Klonen sind abhängig von der Anzahl der Klone, sodass diese Lösung nur für eine begrenzte Menge von Klonen oder für kleine Datenbanken interessant ist.

„RMAN Duplicate“ steigert die Geschwindigkeit des Klonens aus einer aktiven Datenbank im Vergleich zu einer Wiederherstellung einer Sicherung, allerdings werden die Wiederherstellung der Datenbank und Auswirkungen auf die Produktion nicht getestet (außer, man dupliziert aus einer physischen Standby-Datenbank).

Der wichtigste Vorteil von „Datapump Export und Import“ ist die einfache Bedienung. Es ist auch möglich, Untermenüen und Maskierung gleichzeitig zu erledigen. Jedoch ist es nicht einfach, einen Katalog von Klonen zu unterhalten. Der verwendete Speicherplatz ist abhängig von der Größe der Datenbank. Die Fragmentierung der Daten geht verloren, was die Möglichkeit verringert, den Klon für Performance-Tests einzusetzen, weil der

Optimierer evtl. andere Ausführungspläne bevorzugt.

Fazit: Die frei verfügbaren Lösungen sind gut, wenn man eine begrenzte Zahl von Klonen benötigt und die Datenbanken klein sind. RMAN verifiziert gleichzeitig der Restore-Vorgang und überprüft die Backups. Die erwähnten Lösungen bedürfen Skripte zum Automatisieren und benötigen so viel Speicherplatz, wie die Anzahl der Klone beträgt.

Oracle-Lösungen

Mit Aktivieren des Parameters „CLONEDB“ lassen sich einfache Kopien aus einer Read-only-Datenbank erstellen. Seit Version 12.1.0.2 ist dies auf jedem Dateisystem möglich, das sparse files unterstützt, einschliesslich „ext4“ oder „ACFS“. Diese Technologie nutzt die Paket-Funktion „DBMS_DNFS.clonedb_renamefile“, benannt nach ihrer vorherigen Verwendung mit direct NFS. Bei neuen Pluggable-Datenbanken erledigt "Snapshot Copy" die Arbeit, dabei muss aber die Multitenant-Option der Enterprise Edition lizenziert sein.

Diese kostenlose Lösung (auch in der Standard Edition verfügbar) erlaubt es, viele Klone aus der gleichen Sicherung zu erstellen, Tests der Autoren zeigten jedoch eine geringere Leistung der Klone. Bei der Verwendung von Nicht-CDBs sind einige Skripte erforderlich und es gibt keinen Katalog der Klone. Bei Lizenzierung des „Cloud Management Pack“ ist es möglich, mithilfe des Enterprise Manager die Klone zu erstellen und zu katalogisieren. Es ist allerdings eine etwas teure Lösung für diese Anwendung.

Als Teil von Enterprise Manager Cloud Control 12c ist SnapClone eine einfache Cloning-Lösung, basierend auf Hardware (NetApp, Oracle Sun ZFS Storage Appliance EMC VMAX und VNX) oder mit Software (mit ZFS-Dateisystem). Es funktioniert durch die Installation eines Agent- und Speicher-basierten Verwaltungsskripts. Der Klonprozess erfolgt durch einen Assistenten im Enterprise Manager, der viele Optionen bei der Erstellung der Zieldatenbank ermöglicht, sogar RAC-Klone. Die Datafiles der Klone sind aus einem NFS-Mountpoint sichtbar.

Seit dem Datenbank-Release 11.2.0.4 lassen sich Datenbank-Snapshots mit der ACFS-Funktion erstellen. Es empfiehlt sich,

die Version 12.1 der Grid Infrastructure zu verwenden. Die Snapshots entstehen dann mit wenigen Befehlen auf einem ADVN-Volumen. Tests der Autoren zeigten große Leistungseinbußen auf der Quell-Datenbank, nachdem ein paar Snapshots erstellt wurden. Sie empfehlen deshalb, die Snapshots aus einer Standby-Datenbank (natürlich nicht derjenigen, die für die Notfallwiederherstellung vorgesehen ist) auszuführen, nachdem vorübergehend der Apply-Modus deaktiviert wurde. Auf diese Weise beeinflusst das Erstellen der Snapshots die primäre Datenbank nicht.

Oracle stellte das Skript „gDBClone.pl“ bereit, das alle Klon-Aufgaben auf einmal ausführen soll. Es scheint aber nicht ein DBA gewesen zu sein und das Skript ist definitiv nicht für die Produktion bereit. Anstatt zu warten, bis eine neue Version dieses Skripts von Oracle freigegeben wird, kann es einfacher sein, ein vergleichbares Skript selbst zu erstellen.

Proprietäre Lösungen

Die wichtigsten proprietären Lösungen für das Oracle-Datenbank-Cloning sind Delphix und ActifIO. Setzt man einige der besten Oracle-Spezialisten und Entwickler von ZFS zusammen, kann ein interessantes Werkzeug entstehen. Dies ist der Fall bei der Delphix-Agile-Datenplattform. Es handelt sich um eine virtuelle Appliance, die auf vSphere unter einem Fork von Open Solaris 5.11 namens „DelphixOS“ läuft und Daten unter einer Abspaltung der OpenZFS unter dem Namen „DxFS“ speichert. Derzeit werden mehrere Datenbank-Anbieter (Oracle, SQL Server, PostgreSQL, Sybase) unterstützt. Backups, inkrementelle Sicherungen sowie Redo-Logs sind auf dem Dateisystem DxFS möglich; daraus lassen sich Point-in-Time Klones mit NFS erstellen.

Es gibt eine benutzerfreundliche Weboberfläche zum Klonen (zusätzliche Option), aber es ist auch möglich, den Zugriff auf die Appliance über ssh oder https zu automatisieren. Eines der Features ist, einen Klon aus einem anderen Klon zu erstellen, sowie einen schnellen Rücklauf zu starten und einen physischen Klon zu kreieren. Hinzu kommt die Möglichkeit, gegen Aufpreis ein Daten-Maskierungs-Modul zu verwenden.

ActifIO ist spezialisiert auf Lösungen für kontinuierliche Datensicherung mit

Lösung	Lizenzierung	Cloning Typ	Einsatz	Anmerkung
RMAN restore	Frei	Full	Kleinere Anzahl Klone	Prüfen der Backups möglich
RMAN duplicate	Frei	Full	Kleinere Anzahl Klone	Schneller als die Wiederherstellung vom Band
Data Pump	Frei	Full	Kleinere Anzahl Klone	Einfache Bedienung
Clonedb, PDB snapshot copy	Frei ohne OEM, sonst Cloud Management Pack erforderlich, bei PDB Snapshot Copy-Multitenant-Option	Thin	Kleinere Anzahl Klone	Einfach mit PDB, ansonsten Scripting oder OEM-Lizenz erforderlich
Snap Clone	Cloud Management Pack	Thin	Für einige Speicher-Typen verfügbar	OEM-Integration
ACFS snapshots	Enterprise Edition	Thin	Einfache zukünftige Lösung	Leistungsstark mit verschiedenen Klonen
Delphix	Kostenpflichtig	Thin	Sehr gut in bestimmten Fällen	Sehr gute Entwicklung
ActifIO	Kostenpflichtig	Thin	Übergreifende Lösung für Unternehmen	
Storage snapshot	Kostenpflichtig	Thin	Frei mit Oracle	Scripting erforderlich
DataGuard snapshot standby	Enterprise Edition	Full	Für wiederholte Tests am besten geeignet	Teil von DataGuard

Tabelle 1

Komprimierung und Deduplizierung. Die meisten Backup-Daten ändern sich im Laufe der Zeit praktisch nicht. ActifIO arbeitet mit mehreren Lösungen wie VMware, Oracle oder Microsoft zusammen. Es ist als Software-Appliance verfügbar, für größere Installationen existiert auch eine physikalische Einheit.

ActifIO erstellt sofortige Klone für Oracle-Datenbanken und kann diese auch in physikalische Kopien umwandeln. RMAN kommt für das Erzeugen einer ersten Sicherung zum Einsatz; anschließend werden dann nur noch die inkrementellen Änderungen erfasst. Die Autoren hatten bisher keine Möglichkeit, sich näher mit diesem Produkt zu beschäftigen, aber es scheint interessant zu sein, falls ein Unternehmen nach einer globalen Backup-Lösung mit einem Web-Interface und der Bereitstellung von einfachen Oracle-Klonen auf ZFS Basis sucht.

Weitere Alternativen

Oracle 12c bietet neue RMAN-Funktionen für Storage-Snapshots. Das zeigt, dass Oracle das Zurückspielen von Snapshots

garantiert, ohne die Datenbank in den Backup-Modus zu versetzen. Damit lassen sich vorhandene Speicher-Lizenzen für Snapshot-Technologie und einfaches Scripting für Datenbank-Klone verwenden. Ein Anwendungsfall für Speicher-Snapshots ist die gleichzeitige Erstellung von Klonen mehrerer Datenbanken. Man kann sich gut vorstellen, den Snapshot einer ASM-Diskgroup oder eine große Partition zu erstellen, während mehrere Datenbanken laufen. Für die Verwendung dieser RMAN-Funktionen ist allerdings die Advanced-Compression-Option der Datenbank erforderlich.

Eine häufige Anforderung ist die Möglichkeit, wiederholt Tests auf einer Datenbank durchzuführen, nachdem sie in den ursprünglichen Zustand zurückversetzt wurde. Dafür sind keine Snapshots erforderlich, denn die vorhandene DataGuard-Snapshot-Standby-Funktion kann dafür genutzt werden. Sie ist mit der Enterprise Edition kostenlos und ermöglicht mit einfachen Skripten, einen Self-Service für Benutzer zu erstellen. Dabei ist es erforderlich, eine Standby-Datenbank pro möglichem Klon haben.

Fazit

Tabelle 1 zeigt eine Übersicht über die wichtigsten Vorteile und Anwendungsfälle der Lösungen, die in diesem Artikel vorgestellt sind. Es gibt viele gute Möglichkeiten für das Datenbank-Cloning; die Auswahl der Lösung hängt vom Anwendungsfall ab.

Zu Beginn eines Cloning-Projekts ist es wichtig, die richtigen Fragen zu stellen: „Wie viele Klones?“, „Wie schnell?“, „Wie oft?“, und „Zu welchen Kosten?“ etc. Erfahrungen anderer Anwender können ebenfalls ein Kriterium bei der Auswahl einer Backup- oder Speicher-Lösung geben. Last but not least, muss auch die Oracle-Lizenzierungssituation in betracht gezogen werden. In den meisten Fällen wird es keine eindeutig beste Lösung geben und jeder Fall ist ein Einzelfall.

Miguel Anjo
miguel.anjo@trivadis.com