

Der DB-Backup ist gemacht – was nun?

Jörg Brüche
FromDual GmbH
Uster (CH)

Schlüsselworte

Backup, Restore, Recovery, Test, Verifikation, Datenschutz, Anonymisierung, Export

Zusammenfassung

Oft gilt das Thema Datenbank-Backup mit dem Kopieren zum Ablage-Ort als „erledigt“.

Es ist aber notwendig, danach auch die Recovery zu prüfen – und sinnvoll, die so erzeugte Kopie der Produktions-Daten für verschiedene Aufgaben zu nutzen. So wird nicht nur die Produktions-Maschine entlastet, sondern es ergeben sich auch neue funktionale Möglichkeiten.

Vortrags-Inhalt

Die Notwendigkeit des regelmäßigen Backup ist allgemein anerkannt. Sehr häufig gilt diese Sache damit aber als „erledigt“, mit der Ablage des Backup ist Schluss.

Diese Einstellung ist falsch: Ganz routinemäßig (also am besten: automatisiert) muss ein Backup wieder aus der Ablage geholt und eingespielt werden, um seine Tauglichkeit für die Recovery (und das korrekte Arbeiten der Recovery-Prozedur) zu kontrollieren. Erst wenn ein Datenbank-Server mit diesen Daten läuft, ist gesichert, dass auch nach einem echten Ausfall eine Wiederherstellung aus dem Backup möglich ist.

Häufig enthält der Backup personenbezogene Daten oder andere vertrauliche Informationen, die den Bereich der Produktions-Maschinen nicht verlassen dürfen. Es wird also in diesem Bereich eine Maschine (physisch oder virtuell) mit der DBMS-Software benötigt, auf der der zu prüfende DB-Backup eingespielt werden kann.

Danach existiert eine Maschine mit einer Kopie der aktuellen Produktions-Daten, auf der keine Änderungen erfolgen und auch keine Betriebs-Last abgearbeitet werden muss. Diese Maschine ist geeignet, an und mit den Echt-Daten Aufgaben durchzuführen und Untersuchungen vorzunehmen, die wegen der entstehenden Last oder Daten-Änderung auf der eigentlichen Produktions-Datenbank nicht erwünscht oder nicht möglich sind.

Als Erstes sind da Reports, Auswertungen und Abrechnungen zu nennen: Der Backup direkt nach dem Ende einer Rechnungs-Periode (z.B. der vom Monatsersten frühmorgens) enthält die Werte, die für eine Statistik / Abrechnung / ... des letzten Monats benötigt werden. Auf der Recovery-Maschine spielt es keine Rolle, ob dieser Lauf lange dauert und Tabellen sperrt: Es gibt keine anderen Zugriffe, die dadurch behindert werden.

Auch beliebig komplexe Exporte sind möglich, z.B. Kunden- oder Auftrags-Listen für die Notebooks der Mitarbeiter im Außendienst.

Dabei ist natürlich der Datenschutz zu beachten, ebenso die firmeninternen Vertraulichkeits-Regeln. In dieser Kopie der Produktions-Daten spricht nichts dagegen, Werte selektiv zu ändern oder zu löschen, bevor Tabellen exportiert werden.

So ist es auch möglich, die Wünsche der Software-Entwickler zu erfüllen: Sehr häufig möchten sie neue Verfahren oder Programme mit aktuellen Echt-Daten testen. Je nach Einzelfall können sie dafür

anonymisierte Daten für die Entwicklungs-Umgebung erhalten, oder auch ihren Test auf dem Restore-Rechner in der Produktions-Umgebung mit den vollständigen Daten durchführen.

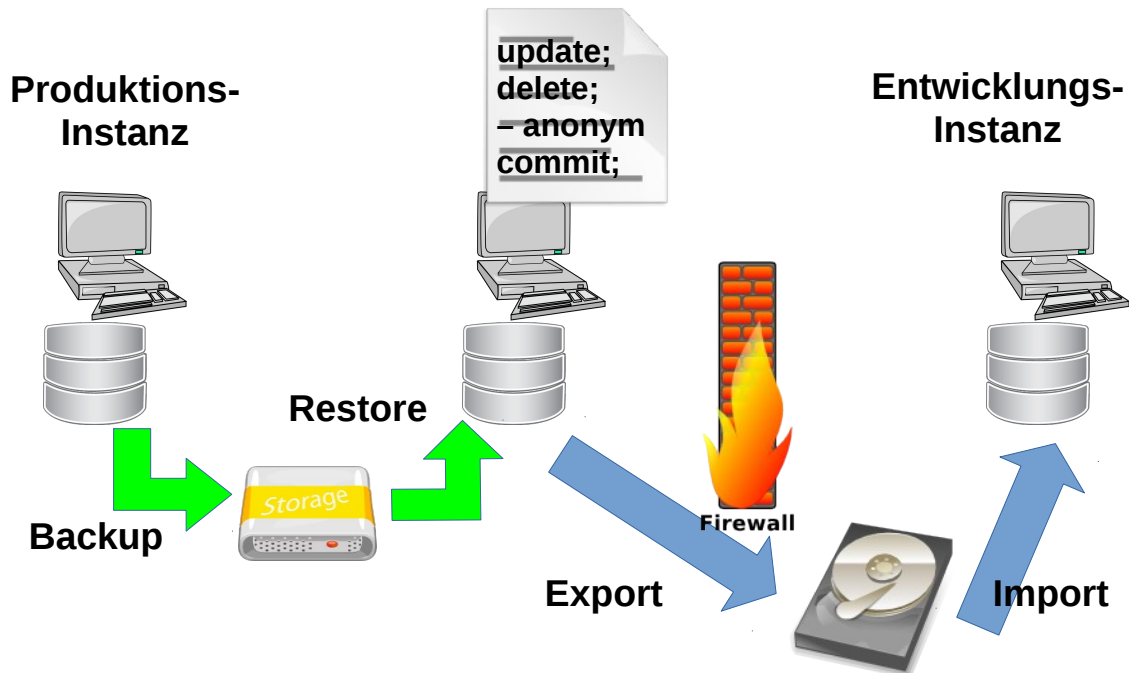


Abbildung: Anonymisierter Export

Es gibt auch DBA-Aktionen, die vor der Durchführung auf dem Produktions-System getestet werden sollten: Wie lange dauert ein bestimmtes „create index“? Wie sehr beschleunigt es ein problematisches „select“? Erfüllen die Alt-Daten bestimmte Integritätsbedingungen, und wie lange dauert das „add foreign key“?

Manche der erwähnten Aktionen (z.B. die DBA-Tests) sind einmalig und werden manuell durchgeführt, andere (angefangen mit der Recovery-Kontrolle) sind regelmäßig nötig und sollten deshalb automatisiert erfolgen. Mit einem CLI-Programm („Command Line Interface“) ist es nicht schwierig, die dafür nötigen Informationen aus der Datenbank zu holen bzw. die SQL-Statements der Anonymisierung auszuführen. Entsprechende Code-Beispiele schließen den Vortrag ab.

Kontaktadresse:

Jörg Brühe
FromDual GmbH
Rebenweg 6
CH-8610 Uster

E-Mail: joerg.bruehe@fromdual.com
Internet: www.fromdual.com