

Migration der Daten eines CRS-Mandanten von Kunde zu Kunde

Dr. Kurt Franke
targens GmbH
Stuttgart

Schlüsselworte

Datapump, Migration, Generator-Scripte

Einleitung

Nach Übernahme einer Bank durch eine andere Bank stand die Aufgabe an, deren CRS-Daten in das CRS (Compliance Risk System) der übernehmenden Bank zu integrieren, damit dort Auswertungen zum Vergleich mit den Jahren vor der Übernahme möglich sind. Die 2 Hauptschritte sind dabei die Extraction der Daten nur für die übernommene Bank aus dem CRS-System eines Betreibers und nachfolgend die Migration und Integration der extrahierten Daten ins CRS der übernehmenden Bank. Naturgemäß ist der 2. Hauptschritt dabei der aufwändigere, schon weil zu erwarten ist, dass die künstlichen ID's aus dem Quellsystem im Zielsystem zumindest teilweise bereits in Verwendung sind und hier eine Anpassung derart erfolgen muß, dass keine Konflikte auftreten. Allein das verhindert einen direkten Import des Datenextracts. Aber auch die unterschiedlichen Versionen von Quell- und Ziel-CRS bedingen einen Zwischenschritt zur Datenanpassung, weil das Datenmodell ja im Zuge der Einführung neuer Features einer Weiterentwicklung unterliegt:

In diesem Vortrag wird erläutert, was hierbei zu beachten war und welche Migrationsschritte dabei notwendig wurden. Außerdem wird betrachtet, welche CRS-Daten der übernommenen Bank nicht migriert werden mußten, und was der Hintergrund dazu ist.

Die Extraction der CRS-Daten aus dem ursprünglichen CRS-System

Im ursprünglichen CRS-System sind Daten verschiedener Kunden des Betreibers in ca. 80 verschiedenen Mandanten gespeichert. Für die Daten-Extraction muss sichergestellt sein, dass nur die Daten des gewünschten Mandanten extrahiert werden.

Es wird jedoch nicht alle Information des betreffenden Mandanten benötigt und exportiert. Insbesondere die Daten der Rechte-Zuteilung im Quellsystem sind für das Zielsystem wertlos und werden deshalb ignoriert, einmal, weil die Anwender dort ganz andere sein werden, und auch, weil die Rollennamen in jedem CRS frei wählbar sind.

In CRS gibt es Auswertungen, via GUI jederzeit zusammengestellt werden können. Da diese relativ einfach möglich ist und die bevorzugten Auswertungen in Quellsystem und Zielsystem nicht dieselben sein müssen, ist es wenig sinnvoll, diese Auswertungen mit vorberechneten Ergebnissen mit zu extrahieren.

Bei CRS sind die Daten aller Mandanten in einem einzigen Datenbank-Schema gespeichert, und sogar in den gleichen Tabellen, bei denen die Zugehörigkeit zu einem Mandanten nur über eine Mandant-ID oder eine untergeordnete ID mit 1:n-Beziehung zur Mandant-ID festgelegt ist.

Für die Extraction der Daten wurde der Oracle Datapump Export verwendet. Dort gibt es die Möglichkeit, mit der QUERY-Clause die Daten jeder Tabelle nach Bedarf auszuwählen. Als Ausgangspunkt dient dabei immer die Mandant-ID des gewählten Mandanten und dabei alle Versionen der Risikodaten (Jahre bzw. unterjährige Zwischenspeicherungen), die zu dem betreffenden

Mandanten gehören. Für die verschiedenen untergeordneten Entitäten (Z.B. LAND, GAE etc.) müssen dann alle zu den gefundenen Versionen gehörenden Ausprägungen ausgewählt werden.

Was die Massendaten wie KUNDE etc. betrifft, so liegen diese in partitionierter Form vor – fast alle list-partitioniert nach GAE-ID – und bieten damit eine wesentlich performantere Möglichkeit der Auswahl über die Angabe der notwendigen Partitionen.

Die Erzeugung des notwendigen Parameterfiles für den Datapumpexport erfolgt über ein Generator-SQL-Script, das die notwendigen Informationen basierend auf dem ausgewählten Mandanten direkt aus dem Quell-CRS ausliest.

Ein Auszug aus dem generierten Export-Parameterfile sieht etwa so aus:

```
TABLES=(\  
MANDANT\  
,VERSION\  
,...\  
,KUNDE:KGAE_67890\  
,KUNDE:KGAE_67891\  
,...\  
,KONTO:KOGAE_67890\  
,KONTO:KOGAE_67891\  
,...\  
)  
  
QUERY='MANDANT:"WHERE id = 12345"'  
QUERY='VERSION:"WHERE mandant_id = 12345"'  
QUERY='LAND:"WHERE version_id IN (SELECT id FROM version v WHERE  
v.mandant_id IN (12345))"'  
QUERY='GAE:"WHERE version_id IN (SELECT id FROM version v WHERE  
v.mandant_id IN (12345))"'
```

Beim expdp Aufruf wird das generierte Parameterfile via PARFILE=Expdp_CRS_Mandant.par übergeben. Das dabei erzeugte Datapump-Exportfile wird an die übernehmende Bank übertragen und dort weiterverarbeitet.

Import, Migration und Integration der CRS-Daten ins Ziel-CRS-System

a) Datapump Import mit Vor- und Nachbereitung

Ein Import des bereitgestellten Datapump-Exportfiles bei der übernehmenden Bank setzt dort ein CRS-System der gleichen Version voraus wie das ursprüngliche CRS-System. Die übernehmende Bank betreibt ihr CRS-System jedoch in einer neueren Version mit erweiterten Features und einem angepassten Datenmodell.

Deshalb musste hier übergangsweise ein weiteres CRS-System mit der Version des ursprünglichen CRS-Systems installiert werden, damit der Datapump-Import möglich wurde. Vor dem Import musste hier noch die bei der Neu-Installation erzeugte System-Initialisierung gelöscht werden, was mit vorhandenen Utilities des CRS-Systems erfolgte. Aufgrund der Tatsache, dass ein Datapump-Import entweder alle Partitionen einer Tabelle oder keine davon anlegt, wurden die partitionierten Tabellen vor dem Import gedroppt, um das Anlegen der benötigten Partitionen sicherzustellen.

Nach dem Import erfolgte zunächst in einem bereitgestellten Script eine Cleanup-Aktion der weiterhin leeren und nicht benötigten Partitionen, die nicht zum im ursprünglichen CRS-System extrahierten Mandanten gehören. Die List-Partitionen der betroffenen Tabellen werden dazu einfach überprüft, ob

es für den HIGH_VALUE eine GAE mit gleicher ID gibt – wenn das nicht der Fall ist, wird die Partition gedroppt. Wegen der ca. 80 vorhandenen Mandanten im ursprünglichen CRS-System werden dadurch mehr als 98 % der Partitionen wieder beseitigt, weil sie zu einem anderen Mandanten gehörten und nur da sind, weil der Datapump-Import unglücklicherweise nicht nur die benötigten Partitionen anlegt, sondern alle, wenn er die Tabelle anlegt, bei existierender Tabelle hingegen überhaupt keine.

Mit einer vorhandenen CRS-Utility wurden die aus der Installation kommenden Sequence-Werte an die via Datapump-Import eingespielten Daten so angepasst, dass ein neu gezogener Sequence-Wert garantiert noch nicht in Verwendung ist.

An dieser Stelle werden außerdem Statistiken auf die Risikomodel-Tabellen mit einer weiteren CRS-Utility und auf den partitionierten Massendaten-Tabellen mit dbms_stats direkt erzeugt, sowie alle vorhandenen GAE's auf abgeschlossen gesetzt – dadurch sind keine Änderungen in CRS in der GUI mehr möglich – und außerdem ein GUI User crs mit Benutzeradministrator-Privilegien für die Zentralen GAE's aller Versionen des importierten Mandanten angelegt.

Jetzt sind die extrahierten Daten des Mandanten aus dem ursprünglichen CRS-System in ein funktionsfähiges CRS-System mit gleicher Version bei der übernehmenden Bank eingespielt.

b) CRS Versions-Update

Damit die CRS-Daten aus dem separat installierten CRS-System in das vorhandene CRS-System der übernehmenden Bank integriert werden können, ist es notwendig, dass beide CRS-System-Versionen gleich sind – nur dann sind Datenmodell und Tabelle-Definitionen identisch. Da das CRS-System der übernehmenden Bank die neuere CRS-Version hat, wird das separat installierte CRS-System mit der CRS-Update-Funktionalität auf die gleiche Version wie das vorhandene CRS-System angehoben. Diese geschieht einfach dadurch, dass statt der für die separate Installation verwendeten CRS-Software diejenige aus dem vorhandenen CRS-System in das separate CRS-System connected wird und die Update-Funktionalität aufgerufen wird. Das Ergebnis ist ein separates CRS-System mit der gleichen Version wie das vorhandene CRS-System der übernehmenden Bank, in dem alle Daten während des Update-Vorgangs bei Bedarf so angepasst wurden, dass sie in die neue Version passen.

Anmerkung: Falls das vorhandene CRS-System eine wesentlich kleinere Version gehabt hätte, wäre wegen der im Nachgang zur Integration benötigten Funktionalität auch ein Update dieses vorhandenen CRS-Systems notwendig geworden.

c) Integration der CRS-Daten des separaten Systems ins vorhandene CRS-System

Da das Risikomodel des zu integrierenden CRS-System von dem des vorhandenen CRS-System abweicht, ist die Methode der Wahl zur Integration ein eigener Mandant. Ein Generator-SQL-Script erzeugt aus den Daten des separaten Systems ein Script zum Anlegen eines rudimentären Mandanten mit den benötigten Versionen und deren Zentrale GAE's im vorhandenen System.

Für die Übertragung der Risikomodel-Daten steht ein Mechanismus über einen Export in eine Xml-Datei und Import in ein beliebiges anderes CRS-System gleicher Version in einer beliebigen Datenbank zur Verfügung. Der Export / Import erfolgt jeweils für die Daten einer Version, weil eine Version genau ein komplettes Risikomodel enthält. Der Mechanismus stellt Aufrufe über DBAU (Database Admin Utility) auf OS-Ebene bereit.

Je ein Generator-SQL-Script zur Erzeugung der Export-Aufrufe und der Import-Aufrufe in der korrekten Reihenfolge erzeugt diese Shell-Scripte auf Basis der Daten im separaten CRS-System.

Nach Abschluß der Risikomodell-Exporte müssen die erzeugten Xml-Dateien aus dem Export-Directory des separaten CRS-Systems ins Import-Directory des vorhandenen CRS-System kopiert werden.

Beim Risikomodell-Import erfolgt automatisch eine Umsetzung der ID's auf Werte, die im Zielsystem noch nicht verwendet wurden, wobei auch die Sequences angepasst werden. Für die GAE-ID's wird dabei sogar eine Mappingtabelle angelegt, die normalerweise am Ende wieder gedroppt wird. Durch Setzen einer speziellen Parameter kann dieses Dropping verhindert werden. Dies ist hier notwendig, weil das GAE-ID-Mapping nachfolgend noch benötigt wird.

Nachdem die Risikomodell-Importe ins Zielsystem abgeschlossen sind, werden über ein SQL-Script die nicht benötigten Arbeitstabellen der Importe gedroppt.

Anschließend müssen noch die Massendaten aus dem separaten CRS-System unter Umsetzung der GAE-ID's gemäß Mapping partitionenweise kopiert werden.

Ein bereitgestelltes SQL-Script erledigt dies vollständig unter Zugriff auf die GAE-ID-Mapping-Tabellen. Die Daten werden dabei über einen Database Link aus dem separaten CRS-System gelesen und in die dafür neu angelegten Partitionen geschrieben.

Abschließend werden im Zielsystem für den neu übernommenen Mandanten noch die GUI-Rechte für die vorhandenen Anwender mit der Rolle Benutzeradministrator via Script erteilt. Die Rechtezuteilung für Fachbereichs-Anwender erfolgt dann durch die Benutzeradministratoren.

Zusammenfassung

Daten-Migrationen von einem Kunden zu einem anderen Kunden sind eher selten und werden, was die notwendigen Anpassungen betrifft, wohl in den meisten Fällen einmalig sein. Hier entsteht also eine Menge an Aufwand, für den die Kosten nicht verteilt werden können. Allein das sollte schon ein Anreiz sein, für solche Migrationen möglichst viel mit vorhandenen Features zu erledigen. Da meist unterschiedliche System-Versionen beteiligt sein werden, ist es fast immer sinnvoll, eine Angleichung durch Update des Systems mit der niedrigeren Version durchzuführen, anstatt Daten in eine Form einer älteren Version zu überführen, wofür es im Regelfall keinen existierenden Code gibt. Wenn Utilities wie das angesprochene Risikomodell-Export und Risikomodell-Import zur Übertragung von Daten zwischen verschiedenen Systemen vorhanden sind, reduziert das die entstehenden Aufwände deutlich, auch wenn damit nicht die komplette Migration abgedeckt werden kann. Letztlich ist es auch von Vorteil, dass solche Utilities bereits in hinreichendem Maße getestet sind und damit die Wahrscheinlichkeit von auftretenden Fehlern im Vergleich zu einer Einmal-Entwicklung geringer sein wird.

Kontaktadresse:

Dr. Kurt Franke
targens GmbH
Calwer Straße 33

D-70173 Stuttgart

Telefon: +49 (0) 711-222992-676
Fax: +49 (0) 711-222992-899
E-Mail Kurt.Franke@targens.de
Internet: www.targens.de