

Datenmodellierung ist langweilig - lassen Sie Datamodeler das machen

**Oleg Kiriltsev
MT AG
Balcke-Dürr-Allee 9, 40882 Ratingen**

Schlüsselworte

DB, Oracle, Datenbank, SQL, Developer, Datamodeler, JavaScript, Transformation, Datenmodellierung

Einleitung

Wie oft haben Sie sich bei der Datenmodellierung nur auf die fachlichen Anforderung konzentriert, ohne über die technischen Anforderungen / Möglichkeiten nach zu denken:

- Welches Tabellen-Präfix wähle ich?
- Wie nenne ich die PK/FK(Primary Key, Foreign Key)- und Auditing-Spalten?
- Indizes? Welche Indizes soll ich erstellen? Wo (in welchem Tablespace) erstelle ich sie? Wie nenne ich sie?
- Trigger? Nicht vergessen die Auditing- und PK-Spalte im Trigger zu befüllen...



Diese und viele andere Fragen müssen bei der Datenmodellierung geklärt und entsprechend erledigt werden. Und was passiert, wenn die Projektvorgaben sich ändern? Sie als Verantwortlicher für das Datenmodell müssen sicherstellen, dass alle Vorgaben erfüllt sind.

Die Lösung: Eine Automatisierung der Prozesse und Prüfungen, die bei allen Projekten gleich sind oder nur geringe Unterschiede aufweisen. Der Vortrag erläutert, wie Sie mit Hilfe der Oracle SQL Developer Data Modeler und der einfachen Skripten die angesprochenen Aufgaben automatisiert erledigen können. Des Weiteren wird im Vortrag beschrieben, wie Sie die Skripte mit einem minimalen Aufwand als Prüfungen verwenden können, um die Qualität eines bestehenden Datenmodells zu überprüfen.

Datenmodellierung

Ab der bestimmten Teamgröße (3-4 Entwickler) kann es sehr schnell passieren, dass ein oder anderer sich vertippt, oder was vergisst. Und erst wenn die Datenmodell-Skripte schon in der Datenbank eingespielt wurden oder noch später, wenn die ersten Tests durchgelaufen sind, wird plötzlich festgestellt:

- Die Spalten in einer Tabelle haben keine einheitlichen Namen
- Die Tabelle hat kein Primary Key Spalte
- Bei der Tabelle fehlen die Auditing-Spalten
- Trigger existiert nicht
- Im Trigger wird die PK-Spalte nicht befüllt
- Und, und, und ...

Oder stellen Sie sich das vor: das Projekt lief so gut, es wurden alle Vorgaben geklärt. Sie fügten gerade die letzte Spalte ein, das Datenmodell ist fertig. Am nächsten Tag haben sich die Projektvorgaben geändert. Die Spalten müssen anders benannt werden, Trigger müssen immer noch eine zusätzliche Spalte mit befüllen, alle Indizes müssen in einem anderen Tablespace abgelegt werden. Wie werden Sie das lösen? - Data Modeler aufmachen, alles anpassen. Oder wenn die Datenmodell-Skripte schon generiert wurden, alle Skripte aufmachen und „Suchen & Ersetzen“. Mit den Transformationsskripten können die Probleme viel effektiver und eleganter gelöst werden.

Ausgehend von den Projekterfahrungen haben wir festgestellt, dass bei jeder Tabelle in einem Datenmodell die folgenden Aufgaben erledigt werden:

- Einheitliche Benennung der Sequenzen
- Einheitliche Benennung der PK- und FK-Spalten
- Indizes für PK- und FK-Spalten im richtigen Tablespace
- Befüllen im Trigger der PK- und Auditing-Spalten
- Projektspezifische Vorgaben, die mehr oder weniger für alle Tabellen gelten



Alle diese Aufgaben können mit Hilfe der Transformationsskripte gelöst werden. Diese Arbeit wird durch die Entwicklungsumgebung abgenommen, somit gewinnen Sie als Datenmodellierer natürlich mehr Zeit und können sich mit den wichtigen Anforderungen und Zusammenhängen im Datenmodell beschäftigen. Die einheitliche Benennung der Objekte führt dazu, dass die Zugehörigkeit sehr schnell erkannt werden kann. Durch die automatisierte Erstellung der Indizes wird sichergestellt, dass die Indizes erstmal überhaupt da sind, und als zweites im richtigen Tablespace abgelegt werden.

Grundlagen

Oracle SQL Developer Data Modeler bietet das sogenannte Application Programming Interface (API). Dadurch können Objekte mit Hilfe der Programmiersprache geändert werden. Unter „Objekten“ werden alle Sachverhalte verstanden, die das jeweilige Modell (logisches, relationales oder physikalisches) hat: Entity, Tabelle, Spalte, Sequenz, Trigger, Index, Key etc. Das Scripting API wird in zwei Sprachen zur Verfügung gestellt:

- JRuby
- JavaScript

d.h. die Skripte können in einer der beiden Sprachen entwickelt werden. Das API selbst hat eine relativ übersichtliche Dokumentation und kann unter „/DATAMODELER_HOME/datamodeler/datamodeler/xmlmetadata/doc/“ gefunden werden. Im Quellcode können alle weiteren Methoden und Strukturen der Programmiersprache verwendet werden, wie Schleifen, IF-Anweisungen usw.

Skriptaufbau

Die Skripte weisen im Prinzip die gleiche Struktur auf. Es wird auf ein globales Objekt (jeweiliges Datenmodell) zugegriffen, darüber werden die weiteren Objekte oder Methoden zur Verfügung gestellt. Dann werden in der Regel die Objekte in einer Schleife iteriert und die notwendigen Transformationen durchgeführt, wie z.B. neue Spalten einfügen, den Datentyp ändern, Index erstellen etc.

Prüfungen aus den Regeln

Bei den Transformationsskripten haben wir immer gesprochen, dass ein zu entwickelndes Datenmodell mit Hilfe von diesen Skripten geändert wird. Die Skripte lassen sich wirklich mit relativ kleinem Aufwand zu den Prüfungen / Regeln umschreiben, die ein bestehendes Datenmodell durchsuchen können, ob alle technischen oder fachlichen Vorgaben erfüllt sind. Mit Hilfe von solchen Prüfungen können Sie ziemlich einfach herausfinden, bei welchen PK- und FK-Spalten die Indizes fehlen, oder welche Tabelle hat keine PK- oder Auditing-Spalten.

Ein wichtiger Vorteil, der durch die Prozessautomatisierung und das Festhalten der zentralen Informationen (Parameter) erreicht wird, ist das Einhalten der Standards und Vorgaben. Besonders große Bedeutung haben die Vorgaben bei der Entwicklung in einem Team und auch bei der Projektübergabe. Werden die Standards eingehalten, kann der neue Entwickler oder das Teammitglied sich darauf verlassen und sich schneller ins Projekt einarbeiten. Ein weiterer positiver Nebeneffekt aus unserer Sicht ist, dass die Vorgaben und Standards den Projektbeteiligten disziplinieren und fordern die gleiche Arbeitsweise.

Fazit und Ausblick

Im Rahmen des Vortrages wurden die Möglichkeiten für die automatisierte Umsetzung sowie der technischen, als auch der fachlichen Anforderungen mit Hilfe von Oracle SQL Developer Data Modeler dargestellt. Durch die Automatisierung können mindestens die folgenden Eigenschaften oder die Erfüllung der folgenden Regeln des Datenmodells sichergestellt werden:

- Einheitliche und den Vorgaben entsprechende Benennung der Tabellenspalten und DB-Objekte, wie Sequenzen, Indizes, Trigger und Constraints
- Vorhandensein der Auditing-Spalten und der Indizes für PK- und FK-Spalten
- Befüllung der PK- und Auditing-Spalten im Trigger

Als weitere denkbare Einsatzpunkte sind folgende:

- Integration der Prüfungen im Continuous Integration Prozess als Postprüfung für die Datenmodelländerungen durch die API-Aufrufe
- Weitere fachliche und / oder projektspezifische Prüfungen für die Qualitätssicherung

Kontaktadresse:

Oleg Kiriltsev

MT AG

Balcke-Dürr-Allee 9

D-40882 Ratingen

Telefon: +49 (0) 2102-30961 0

Fax: +49 (0) 2102-30961 101

E-Mail oleg.kiriltsev@mt-ag.com

Internet: www.mt-ag.com