

XML in der Oracle DB

Wolfgang Nast
MT AG
Ratingen

Schlüsselworte

Oracle 12c XDB XML SQL/XML

Einleitung

Es wird vorgestellt, wie XML Daten in der Oracle 12c abgelegt werden können. Dabei werden die unterschiedlichen Möglichkeiten der Speicherung der Daten vorgestellt. Zuerst kommen die klassischen Ansätze mit Blob und Clob mit Metaspalten oder verteilte relationale Tabellen mit Foren Key Beziehungen zum Tragen. Dann wird der XML Datentyp XMLType vorgestellt und wie hier die Oracle 12c viele neue Möglichkeiten bietet. Zum Schluss werden noch die Abfragemöglichkeiten von XMLType Tabellen als relationale Daten gezeigt. Auch relationale Daten lassen sich als XMLDaten zusammenstellen.

Lob mit Metadaten

Die Daten können zum reinen Speichern in Lobs abgelegt werden. Hier sollten alle benötigten Metadaten als eigene Spalten herausgezogen werden, damit diese Metadaten als Abfrage- und Sortierkriterium zur Verfügung stehen.

Blobs eignen sich besonders, um die XMLDaten so zu speichern, wie diese geliefert wurden. Es wird keinerlei Verarbeitung der Rohdaten vorgenommen. So kann die Kodierung der Daten noch geändert werden oder es kann auf Fehler in der Lieferung hingewiesen werden. Dies eignet sich besonders um Originallieferungen oder Fehler abzulegen.

Clobs eignen sich, um XMLDaten schon als Text zu speichern, in dem gesucht wird. Hier muss beachtet werden, dass der Zeichensatz für die Kodierung stimmt. Da die Daten als Text vorliegen, lassen sich hier einfacher Metadaten bestimmen. Auch die Verarbeitung der Daten als Text ist einfacher. Hier bleibt noch die Struktur der XMLDaten erhalten, mit Formatierung wie Zeilenumbrüche mit Einrückungen, Kommentare und auch Verarbeitungsanweisungen (PI).

Bei beiden Ansätzen bleibt der Anwendung die eigentliche Aufbereitung der Daten überlassen und es müssen jedes Mal die Daten geparkt werden, um mit ihnen zu arbeiten.

Als alternative hierzu besteht die Möglichkeit die geparkten Daten zu Speicher. Das wird im folgenden Absatz beschrieben.

Daten in relationalen Tabellen

Hier werden die schon geparkten Daten von der Anwendung in eine relationale Struktur gebracht. Diese wird dann von der Datenbank abgebildet. Hier müssen die Daten noch um die relationalen Strukturdaten erweitert werden, um einen Primary Key zu haben. Da meist die Zeilen keine echten Daten haben, die als Primary Key verwendet werden können, wird ein technischer Schlüssel

verwendet. Dieser wird für die Zuordnung zwischen den Tabellen genutzt. Damit sind die Daten eigentlich optimal gespeichert. Nur hat dies aus Sicht der Datenbank nichts mit XML Daten zu tun, da hier das Abbilden komplett von der Anwendung übernommen wird.

Daten im XMLType

Der Type XMLType dient zum Speichern von wohlgeformten XML Daten die mit und auch ohne Schema abgelegt werden. Der XMLType kann hier normal wie eine Spalte verwendet werden. Auch als Type der Tabelle kann er verwendet werden.

Bei der Speicherung der Daten in XMLType müssen die Daten mindestens wohlgeformt sein. Wurde auf ein XMLSchema eingeschränkt, so müssen die Daten auch Schemavalide sein. Beim Speichern der Daten gehen nicht relevante Daten, wie Formatierung und Verarbeitungsanweisungen verloren. Kommentare werden entfernt und CDATA Blöcke werden in Texte mit Ersatzzeichen umgewandelt.

Hier kommen einige Neuerungen der Oracle 12c zu den Vorgängern zum Tragen. Die gespeicherten Daten sind kompakter, da intern keine Lobs mehr verwendet werden. Bei der Abbildung auf ein Schema werden jetzt Objekttypen verwendet und keine relationale Tabellen.

Die Erweiterung von XML Schema um Tabellen und SQL Daten Typen ist noch möglich, aber führt nicht mehr zu relationalen Tabellen sondern zu Tabellen vom Type XMLType.

Für die Abfrage von XMLType Inhalten wird seit Oracle 12c der XPath Standard verwendet, so wie es SQL/XML vorgibt. Die Oracle spezifischen Abfragen der älteren Versionen werden nicht mehr unterstützt. Hier sind auch alle Aufrufe um aus XML Daten relationale Daten zu erstellen auf den SQL/XML Standard angepasst worden.

Mit diesen Erweiterungen ist es möglich XML Daten als Spalten heraus zu ziehen. Diese Spalten können direkt verwendet werden. Für Indizes, Constrains oder Primary Key.

Hiermit wird es möglich die XML Daten gezielt ab zu fragen. Hier sind neben XMLFragementen auch relationale Strukturen möglich.

Möchte man relational lesend auf die Daten zugreifen, so lohnt es sich meistens einen View dafür anzulegen. Alternativ kann es reichen auf die schon herausgezogenen Spalten zuzugreifen.

XML Daten aus relationalen Tabellen

Auch kommt es häufig vor, das aus den vorhandenen relationalen Daten wieder XML Daten zu erstellen. Hier bietet die Oracle 12c viel Unterstützung an. Damit kann die Datenbank die Funktion übernehmen, die bisher von der Applikation übernommen wurde.

Zusammenfassung:

Es ist immer abhängig von der Verarbeitung der Daten, wie diese am besten gespeichert werden. Bei Daten, für die schon eine Verarbeitung besteht, kann es sinnvoll sein eine neue Sicht zu definieren. Die neu abzulegenden Daten sollte man überlegen, ob diese vorwiegend

relational benötigt werden oder ob die Daten nur Gespeichert werden. Bei der Nutzung von den XML Strukturen und XPath kann der Datentype XMLType und SQL/XML viele Vorteile bringen, wie kompakte Speicherung und anwendungsgerechte Zurverfügungstellung der Daten. Die Abfragen liefern nur den gewünschten Teil der Daten, passend Aufbereitet, als XMLFragment oder Tabellarisch.

Kontaktadresse:

Wolfgang Nast

MT AG

Balcke-Dürr-Allee 9
D-40882 Ratingen

Telefon: +49 (0) 2102 30961-0
Fax: +49 (0) 2102 30961- 101
E-Mail Wolfgang.Nast@mt-ag.com
Internet: www.mt-ag.com