

# Fehlerbehandlung bei einer SQL-Bulk-Verarbeitung

**Markus Fiegler**  
**ORDIX AG**  
**Paderborn**

## **Schlüsselworte**

Pipelined Table Function, DML Error Logging, Standardwert bei Konvertierungsfunktionen (ab 12.2)

## **Einleitung**

Sollen Daten schnell und Ressourcen schonend verarbeitet werden, dann ist SQL-Bulk-Verarbeitung im Gegensatz zu einer Einzeldatensatzverarbeitung die erste Wahl. Tritt bei der SQL-Bulk-Verarbeitung allerdings ein Fehler z.B. bei der Datentypkonvertierung auf, so wird die Verarbeitung standardmäßig abgebrochen. Welche Möglichkeiten gibt es, eine SQL-Bulk-Verarbeitung trotz geworfener Fehler fortzusetzen? In diesem Vortrag werden Fehlerbehandlungsvarianten für eine SQL-Bulk-Verarbeitung wie Pipelined Table Function, DML Error Logging und Standardwerte bei Konvertierungsfunktionen (ab 12.2) vorgestellt und bezogen auf den Einsatz in der Praxis bewertet. Darüber hinaus werden diese Varianten unter dem Gesichtspunkt Performance miteinander verglichen.

## **Ausgangssituation**

Eine typische Aufgabe im Data Warehouse-Umfeld ist die Umwandlung von nicht typisierten Daten wie z.B. VARCHAR2-Felder in typisierte Daten wie z.B. NUMBER, DATE oder TIMESTAMP. Lassen sich die VARCHAR2-Inhalte konvertieren, dann werden diese in Tabellen mit typisierten Feldern übertragen. Fehlerhafte Datensätze werden in der Regel in so genannte ERROR-Tabellen angesteuert. Alternativ können fehlerhafte Datensätze mit Default-Werten belegt und in die typisierten Tabellen übertragen werden. Wichtig dabei ist, dass ein Konvertierungsfehler nicht zum Abbruch der Verarbeitung führt (Abbildung 1).

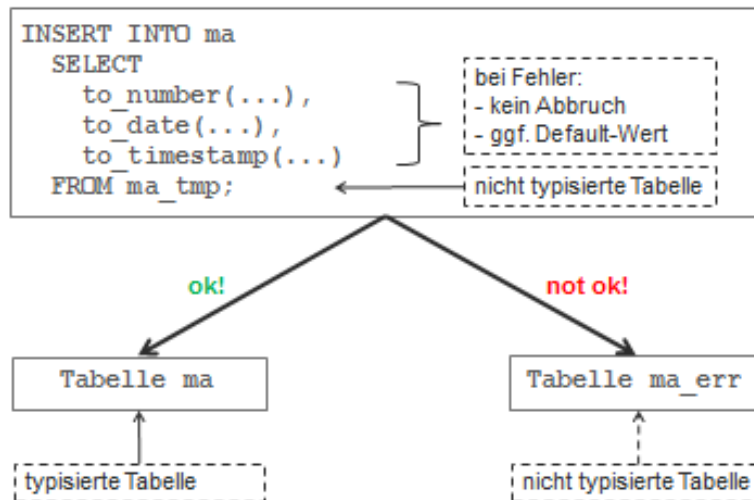


Abb. 1: Ausgangssituation

### Pipelined Table Function

Eine Table Function ist eine Stored Function, die eine Menge von Zeilen in einer Kollektion zurück liefert. Das Besondere an einer Table Function ist, dass sie in einer FROM-Klausel ähnlich wie eine Tabelle verwendet werden kann. Eine herkömmliche Table Function erzeugt im ersten Schritt die gesamte Kollektion und liefert diese anschließend an die aufrufende Einheit zurück. Eine Pipelined Table Function kann im Gegensatz zu einer herkömmlichen Table Function schon während der Ausführung Zwischenergebnisse an die aufrufende Einheit zurück liefern. Der Vorteil dieser Vorgehensweise ist zum einen weniger Speicherplatzverbrauch und zum anderen eine bessere Laufzeit. Abbildung 2 zeigt eine Lösungsalternative mit einer Pipelined Table Function:

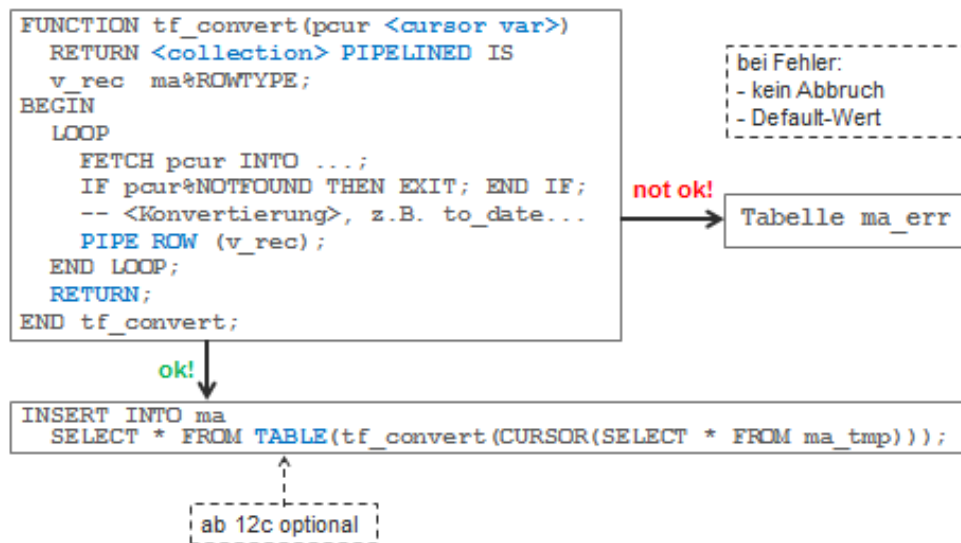


Abb. 2: Lösungsalternative mit einer Pipelined Table Function

### DML Error Logging

Die DML Error Logging Variante ermöglicht bei einer DML-Anweisung auftretende Fehler zu protokollieren und die Verarbeitung trotz Fehler weiter fortzuführen. Die ERROR-Tabelle bei dieser Variante verfügt zusätzlich zu den Daten-Feldern über weitere Felder wie z. B. für die Aufnahme der Fehlernummer, des Fehlertextes oder der DML-Aktion die im Fehlerfall automatisch gefüllt werden. Für die Erstellung der ERROR-Tabelle stellt Oracle für diese Variante das DBMS\_ERRLOG PL/SQL-Package zur Verfügung. In der Abbildung 3 ist eine Lösungsalternative mit der DML Error Logging dargestellt.

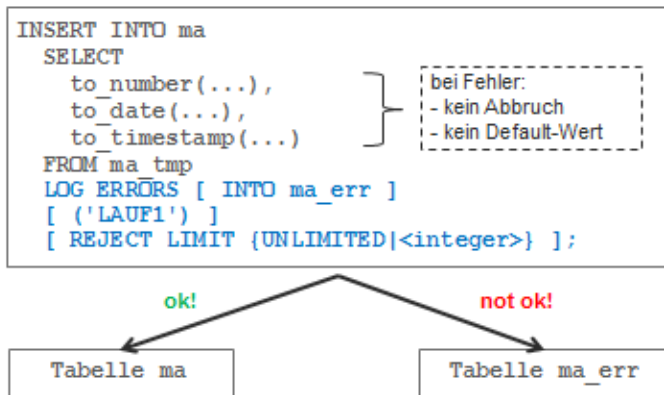


Abb. 3: Lösungsalternative mit DML Error Logging

**Standardwert bei Konvertierungsfunktionen**

Die Konvertierungsfunktionen wie TO\_NUMBER, TO\_DATE oder TO\_TIMESTAMP wurden mit der Oracle Version 12.2 um die Übergabe eines Default-Wertes erweitert, der im Fehlerfall statt einer Fehlermeldung zurückgeliefert wird. Damit kann eine Verarbeitung trotz eines Konvertierungsfehlers fortgesetzt werden, ohne dass eine Fehlermeldung geworfen und die Abarbeitung eines SQL-Statements abgebrochen wird. Der Konvertierungsfehler selbst (ORA-Nummer) kann bei dieser Variante allerdings nicht ermittelt werden. Die folgende Abbildung stellt eine Lösungsalternative mit Standardwert bei Konvertierungsfunktionen dar:

**Standardwert bei Konvertierungsfunktionen (ab 12.2)  
Lösungsalternative**

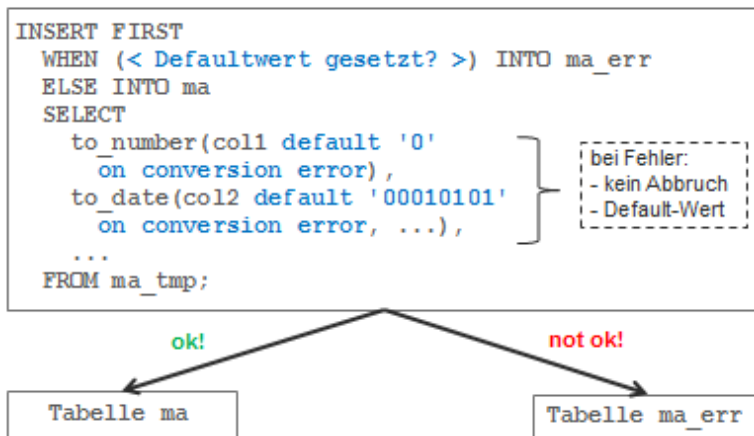


Abb. 4: Lösungsalternative mit Standardwert bei Konvertierungsfunktionen

### Vergleich der Lösungsalternativen

Die folgende Abbildung stellt die wesentlichen Unterschiede der hier vorgestellten Lösungsalternativen dar:

#### Vergleich der Lösungsvarianten

ORDIX AG

Was?	PTF BULK	DML Error Logging	Default-Wert (12.2)	PL/SQL BULK
Können Default-Werte verwendet werden?	Ja	<b>Nein</b> (aber Stored Function)	Ja	<b>Nein</b> (aber Stored Function)
Können mehr als nur Konvertierungsfehler abgefangen werden?	<b>Nein</b>	Ja (aber kein APPEND)	<b>Nein</b>	Ja (aber kein APPEND)
Kann die Fehlernummer ermittelt werden?	Ja	Ja	<b>Nein</b>	Ja
PGA-Verbrauch (%)	89	0	0	<b>100</b>
Performance (%)	<b>100</b>	40	42	74

22

Abb. 5: Vergleich der Lösungsvarianten

### Fazit

Die Auswahl einer geeigneten Konvertierungsvariante kann für den Erfolg einer Anwendung eine entscheidende Rolle spielen. Ein Wechsel einer Konvertierungsvariante ist in der Regel sehr aufwändig und demzufolge sollte die Auswahl der Konvertierungsvariante so früh wie möglich in einer Projektphase stattfinden.

Mit allen den hier vorgestellten Lösungsalternativen ist eine Datentypkonvertierung möglich. Konvertierungsfehler werden abgefangen, in eine ERROR-Tabelle gesteuert und führen dabei nicht zum Abbruch der Verarbeitung.

Da in der Praxis die Laufzeit und die Ressourcen schonende Verarbeitung eine wesentliche Rolle spielen, sind aus diesem Grund besonders die Varianten DML Error Logging und Standardwert bei Konvertierungsfunktionen zu empfehlen. Die letztere Variante steht allerdings erst ab der Oracle Version 12.2 zur Verfügung. Der wesentliche Vorteil der DML Error Logging Variante ist darüber hinaus die Fähigkeit, auch z.B. mit Unique Key- oder Fremdschlüssel-Verletzungen umzugehen, was den Einsatz nicht nur im Data Warehouse-Umfeld, sondern auch in OLTP-Systemen sinnvoll macht.

### Kontaktadresse:

Markus Fiegler  
ORDIX AG

Karl-Schurz-Straße, 19a  
D-33100 Paderborn

Telefon: +49 (0) 5251 1063-0  
Fax: +49 (0) 180 1673490  
E-Mail [info@ordix.de](mailto:info@ordix.de)  
Internet: [www.ordix.de](http://www.ordix.de)