

WEITER BLICKEN. MEHR ERKENNEN. BESSER ENTSCHIEDEN.



**Analyse von Dimensions-Schlüsselfehlern bei der Aufbereitung
von SSAS Datenbanken**

SOLISYON GMBH

TOBIAS GRUBER

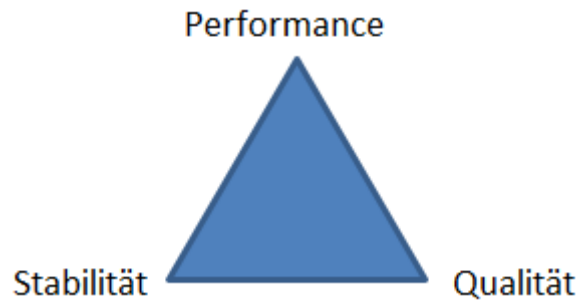
BEN WEISSMAN

ANALYSE VON OLAP-AUFBEREITUNGSFEHLERN

1	AUSGANGSSITUATION.....	2
2	LÖSUNGSANSATZ	3
3	EINRICHTEN DER UMGEBUNG	3
4	EINSPIELEN DES FEHLERLOGS	3
4.1	ERSTELLEN EINES FEHLERLOGS.....	3
4.2	EINSPIELEN DER LOGDATEI.....	5
4.3	AUFBEREITUNG DER OLAP DATENBANK UND AUSWERTUNGEN	6
5	AUTOREN	7

1 Ausgangssituation

In vielen Datawarehouses sind oftmals fehlende Stammdateninformationen die Ursache für einen Abbruch in der Aufbereitung von OLAP Cubes. Häufig stehen Systemverantwortliche und Entwickler einem Abwägen zwischen stabilen Prozessen, Laufzeit und den hohen qualitativen Standards der Fachbereiche gegenüber.



Um eine möglichst schnelle Verarbeitung zu gewähren, werden im Idealfall keine Dimensionsattribute aus Fakten-Daten mit Dimensionsdaten abgeglichen, da jede Prüfung via SQL Statements oder Constraints Zeit beansprucht. Da in der Praxis jedoch das DWH in vielen Fällen der erste Moment ist, an welchem Daten aus verschiedenen Systemen konsolidiert betrachtet werden, sind oftmals Prüfungen in der ETL-Strecke, basierend auf den Erfahrungswerten der IT Mitarbeiter, notwendig.

Um dennoch keine Daten beim Importieren in die multidimensionale Schicht zu verlieren, bietet Microsoft bei der Verarbeitung von OLAP-Elementen an, eine benutzerdefinierte Fehlerkonfiguration zu verwenden.

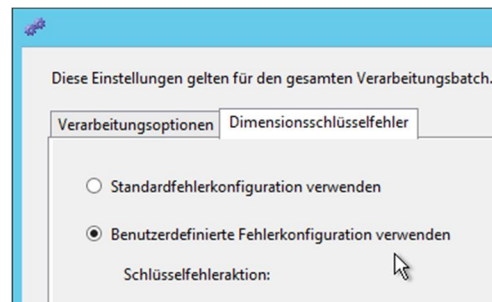


Abbildung 1: Fehlerkonfiguration

Diese Einstellung ermöglicht unter anderem, fehlende Schlüsselattribute, wie in diesem Beispiel angesprochen, auf ein spezielles „Unbekanntes Element“ zu schreiben. Auf dieser Position würden dann in Auswertungen die Positionen auftauchen, welchen keinem Dimensionsattribut zugeordnet werden konnten.

Wenn diese Konfiguration verwendet wird, bietet Microsoft lediglich eine sehr rudimentäre Auswertmöglichkeit der angefallenen Fehler an. Dies wird in Form eines Textfiles angegeben, in welcher die Fehler, sobald sie auftreten, geschrieben werden. Dies führt häufig dazu, dass erforderliche Korrekturen von Stammdaten erst zu spät veranlasst werden und somit die Qualität des Berichtswesens leidet.

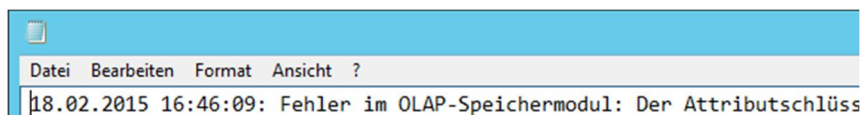


Abbildung 2: Logfile

2 Lösungsansatz

Um die Daten möglichst effizient auswerten zu können, wird die vorhandene Umgebung und Technologien verwendet. Dabei wird das Logfile in den SQL Server geladen und in der OLAP Datenbank ausgewertet. Der Import geschieht dabei über ein selbst entwickeltes dtsx-Paket oder kann über einen Import aus dem Management Studio realisiert werden.

Alle benötigten Sourcen erhalten Sie kostenfrei unter <http://solisyon.de/OLAPAufbereitung/>.

Die Quellen sind aktuell auf eine deutschsprachige Ausgabe der Fehlermeldungen ausgelegt.

3 Einrichten der Umgebung

Zum Einrichten der Umgebung muss auf dem Server eine neue Abfrage über das Management Studio geöffnet werden und das Skript aus der Datei „OLAP_Fehlermeldungen.sql“ eingefügt werden. Der Speicherort des Transaction-Logs sowie der Datenbankdatei müssen über die beiden Variablen am Start ergänzt werden.

Ist dies erfolgreich geschehen, wird die OLAP Datenbank eingespielt. Hierzu muss über das Management Studio ein neues Fenster für XMLA Abfragen geöffnet werden. Das Skript aus der Datei „OLAP_Fehlermeldungen.XMLA“ wird an dieser Stelle eingefügt. Vor der Ausführung muss noch der Datenbankname des Connection-Strings angepasst werden. Ein entsprechender Kommentar ist am Anfang des Skripts hinterlegt.



Abbildung 3: Neues XMLA Statement

Zum Überprüfen ob alles geklappt hat, wird empfohlen die OLAP Datenbank aufzubereiten und im Fehlerfall die Verbindung der Datenquelle zu überprüfen.

4 Einspielen des Fehlerlogs

Sobald alle Vorbereitungsschritte erfolgreich beendet wurden, können die echten Daten eingespielt werden.

4.1 Erstellen eines Fehlerlogs

Das Fehlerlog, welches in der Einleitung angesprochen wurde, wird über eine Aufbereitung der OLAP-Datenbank erstellt. Es werden dabei alle Fehlermeldungen ausgegeben und im Anschluss in eine Textdatei gespeichert.

Die Konfiguration sieht dabei wie folgt aus:

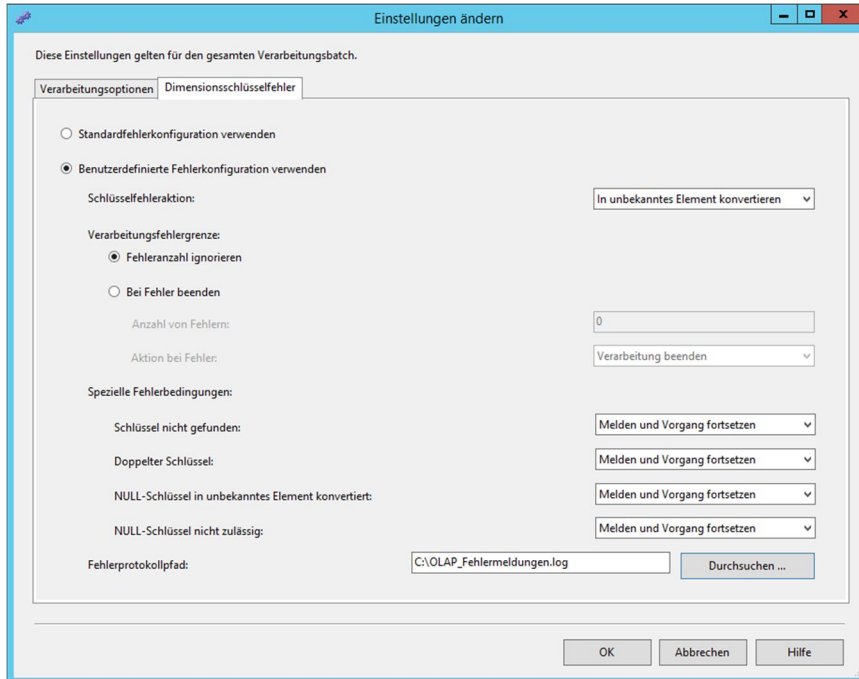


Abbildung 4: Konfiguration der Verarbeitung

Neben der gesamten OLAP-Datenbank können auch einzelne Cubes und Dimensionen verarbeitet und anschließend überprüft werden.

4.2 Einspielen der Logdatei

Die in Kapitel 4.1 erstellte Logdatei wird nun in die Tabelle OLAP_Errorlog eingelesen. Diese Tabelle wurde mit der Datenbank „OLAP_Fehler“ generiert und besteht nur aus einer Unicode Zeichenspalte, in welcher Zeile für Zeile aus dem Logfile abgespeichert werden. Das Teilen der Informationen geschieht über die Views, bzw. über die erzeugte Funktion.

Zum Einspielen der Datei kann ein eigenes SSIS-Paket dienen oder der „SQL Server-Import/Export-Assistent“. Dabei wird die Flatfile-Quelle in die oben genannte Tabelle eingelesen. Als Spaltentrenner wird der Doppelpunkt „:“ vorgeschlagen. Dies wird auf einen Senkrechten Strich angepasst „|“. Daraus resultiert, dass es nur noch eine Spalte gibt, wie es in der Zieltabelle erwartet wird. Weiter muss die „Output Column Width“, also die Spaltenlänge angepasst werden. Die Standardlänge von 50 wird auf 1000 wie in der Tabellendefinition angepasst.



Abbildung 5: Datenimport

Abschließend muss als Ziel die vorhandene Tabelle hinterlegt werden und das Mapping der „Spalte 0“ auf die Spalte „Entry“ überprüft werden. Ist dies erfolgreich geschehen wird empfohlen, die Zieltabelle zu überprüfen:

```
USE [OLAP_Fehler]
GO

SELECT [Entry]
FROM [OLAP_Fehler].[dbo].[OLAP_Errorlog]
```

4.3 Aufbereitung der OLAP Datenbank und Auswertungen

Wurde die „OLAP_Errors“ Datenbank aufbereitet, kann der Cube „OlapErrors“ über ein beliebiges Frontend ausgewertet werden.

Exemplarisch wurde nachfolgend der Cube „Verkaufschancen“ ausgewertet. Bei diesem Cube wurden aus der Dimension „Verkaeuer“ die Wertschlüssel „Ben Weissman“ und „Tobias Gruber“ nicht gefunden.

	A	B	
1	Zeilenbeschriftungen	View FAKT OLAP Errors	Anzahl
2	Verkaufschancen		2
3	Verkaeuer		2
4	Ben Weissman		1
5	Tobias Gruber		1
6	Gesamtergebnis		2

Verwendete Dimensionen:
Cube Name
Dimension
Fehlender Wert

Abbildung 6: Datenanalyse

Mittels dieser Lösung kann mit wenig Aufwand aus einem schwer lesbaren Logfile eine komplette Lösung entstehen, mit welcher bis zum tatsächlichen Wert in der hinterlegten Tabelle der fehlende Wert analysiert werden kann.

5 Autoren

Haben Sie noch Fragen? Haben Sie Interesse am konkreten Einsatz der von uns erstellten Datenbank, oder den Funktionen? Benötigen Sie Unterstützung bei einer ähnlichen Herausforderung?

Besuchen Sie uns unter www.solisyon.de oder wenden Sie sich gerne direkt an einen der Autoren!



Tobias Gruber
DWH Consultant

t.gruber@solisyon.de
+49 911 99 00 77 25



Benjamin Weissman
Geschäftsführer

b.weissman@solisyon.de
+49 911 99 00 77 21