

# XQuery Implementation in a Relational Database System

VL XML, XPath, XQuery: Neue Konzepte für Datenbanken

Jörg Pohle, [pohle@informatik.hu-berlin.de](mailto:pohle@informatik.hu-berlin.de)  
Daniel Apelt, [apelt@informatik.hu-berlin.de](mailto:apelt@informatik.hu-berlin.de)

# Einführung

Das Paper behandelt folgende Aspekte:

- Einführung in die XML-Fähigkeiten des SQL Server 2005
- die XQuery-Verarbeitungsarchitektur
- wie XQuery-Ausdrücke in "query trees" umgewandelt werden, die
  - relationale Operatoren und
  - eine Anzahl für die Verarbeitung von XQuery neu eingeführter Operatoren beinhalten
- einige interessante Aspekte der Implementation

# XML Support

Datentyp für XML-Daten in MS SQL Server 2005: XML

- Speicherbare Daten:
  - XML-Dokumente
  - XML-Fragmente
- XML Schema Support:
  - "XML schema collections": XML Schema Dokumente als Metadaten nutzbar
  - "typed XML": mit XML Schemata verbunden
  - "untyped XML": ohne Verbindung mit einem Schema
- statisches Type-Checking und Constraints-Durchsetzung:
  - auf Basis der Informationen in den XML Schemata

# XML Support 2

- Indizierung:
  - Primärer XML Index
    - `CREATE PRIMARY XML INDEX` idx\_Spalte `ON` [Tabelle] (Spalte)
  - Sekundäre XML Indices:
    - Pfad-Indizierung ("path index")
    - Eigenschaftens-Indizierung ("property index")
    - Wert-Indizierung ("value index")

# XQuery

- Teilmenge von XQuery
- Grundlegender Ablauf der Verarbeitung:
  - Parsen des XQuery-Ausdrucks
  - Überführung in eine interne Struktur, die als "XML algebra tree" bezeichnet wird
  - Anwendung von regelbasierten Optimierungen ("rule-based optimizations")
  - Transformation des "algebra tree" in einen "relational operator tree"

# Beispieldaten



- xmlgen – auction.xml (15 MB)

```
<?xml version="1.0" encoding="ISO-8859-1"?>  
<site>  
  <regions>  
    <africa>  
      <item id="itemo">  
        <location>United States</location> ...
```

- xmlgen – auction.dtd

```
<!ELEMENT site      (regions, ..., open_auctions, closed_auctions)>  
<!ELEMENT regions  (africa, asia, australia, ..., samerica)>  
<!ELEMENT africa   (item*)>  
<!ELEMENT asia     (item*)>
```

- XMLSpy – XML Schema

- MS SQL Server 2005 Developer Edition

# Beispiel

CREATE XML SCHEMA COLLECTION auction AS

```
'<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">  
<xs:import namespace="http://www.w3.org/XML/1998/namespace"/>  
...  
</xs:schema>'
```

```
CREATE TABLE [dbo].[XmlTypedi] (  
    PK INT PRIMARY KEY,  
    XDOC XML(CONTENT auction))
```

```
INSERT [dbo].[XmlTypedi]  
    SELECT 10, XDOC  
    FROM (SELECT * FROM OPENROWSET  
        (BULK 'path\to\auction.xml', SINGLE_BLOB)  
        AS XDOC)  
AS R(XDOC)
```

```
CREATE PRIMARY XML INDEX idx_XDOC ON [dbo].[XmlTypedi] (XDOC)
```

# Beispiel 2

```
SELECT [PK], XDOC.query('
  for $s in //item[@id = "item86"]
  return  <herkunft>
           {$s/location/text()}
         </herkunft>') AS Expr1
FROM [XmlTypedi]
```



# Beispiel 3



```
SELECT [PK], XDOC.query('
  for $a in //open_auction
  return
    if ( some $s in $a/seller,
        $b in $a/bidder
        satisfies $s/@person = $b/personref/@person )
    then <fraud>
      {//people/person[@id = $a/seller/@person]}
    </fraud>
    else ()
  ') AS Expr1
FROM [DOCS]
```

# Beschränkungen

- keine Unterstützung für:
  - let – wird vom Framework, aber nicht vom SQL Server unterstützt
  - erweiterte ordered- und unordered-Syntax
  - empty least-, empty greatest- und collation-Schlüsselwörter für order by
  - "globale" Variablen und Funktionen
- interne Beschränkung:
  - 2GB pro XML-Blob

# Optimierungen

Pal et al. beschreiben Optimierung der XQuery-Anfragen

nicht optimiert

```
SELECT PK, XDOC.query(  
  for $i in /site/regions  
  return $i/africa') AS Expr1  
FROM DOCS
```

optimiert

```
SELECT PK, XDOC.query(  
  /site/regions/africa') AS Expr1  
FROM DOCS
```

Die Ablaufdiagramme lassen Zweifel aufkommen

## **XQuery Implementation in a Relational Database System**

Shankar Pal et al., Proceedings of the 31st International Conference on VLDB, Trondheim, 2005

<http://www.vldb2005.org/program/paper/thu/p1175-pal.pdf>

## **XML Support in Microsoft SQL Server 2005**

Shankar Pal et al., Microsoft Corporation, December 2005

## **Dokumentation MS SQL Server 2005**

## **MS SQL Server 2005 Developer Edition**

MSDN Academic Alliance - Programm

<http://genil.informatik.hu-berlin.de/default.html>

## **XML-Generator: xmlgen**

<http://monetdb.cwi.nl/xml/generator.html>

## **Altova XMLSpy**

<http://www.altova.com/de/>

Vielen Dank für Eure Aufmerksamkeit