

XML und Delphi

Verarbeiten von XML-Dokumenten
mit Borland Delphi und Kylix

Stefan Heymann

Kittelberger GmbH, Reutlingen, www.kittelberger.de

Umgebungsbedingungen

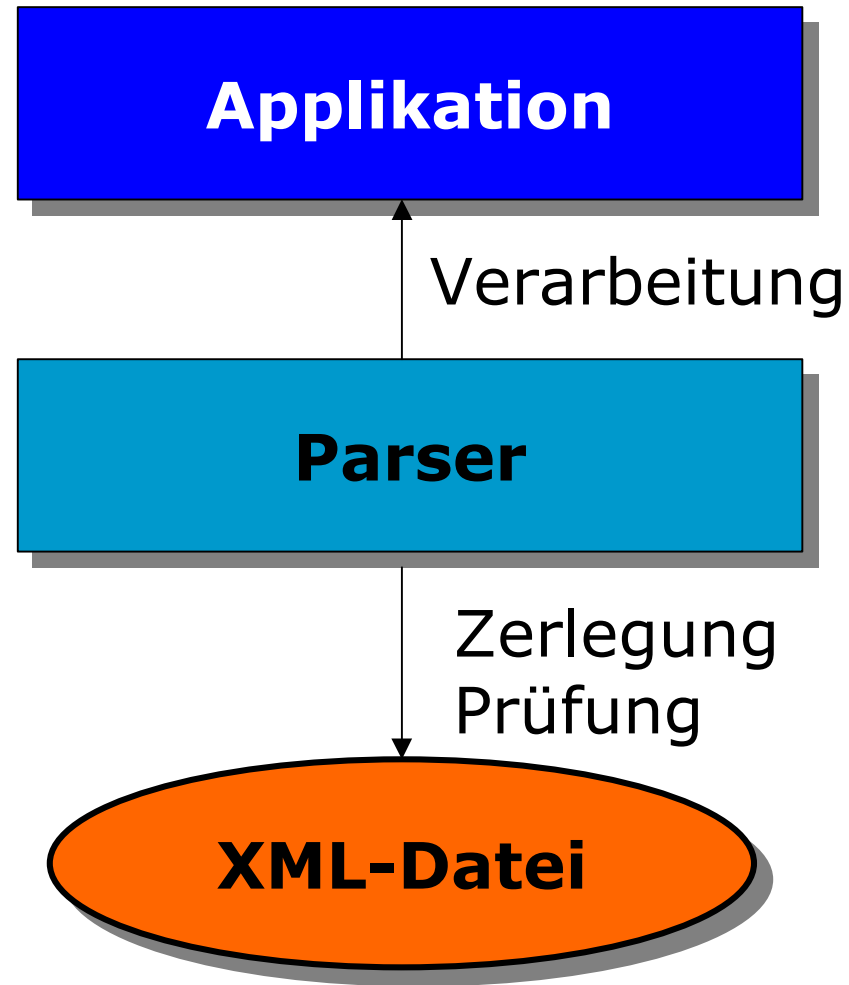
- Handy aus?
- Bei Fragen: fragen!

Inhalts-Übersicht

- Grundlegendes zum Parsing
- Parser für Delphi
- Live-Häck
- Misc

XML Parsing

XML-Verarbeitung



Das Demo-XML-File

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
```

```
<CDCollection>
```

```
<CD>
```

```
  <singer gender="female">Madonna</singer>
```

```
  <title>Ray of Light</title>
```

```
  <tracks>13</tracks>
```

```
</CD>
```

```
<CD>
```

```
  <singer gender="female">Gloria Estefan</singer>
```

```
  <title>Mi Tierra</title>
```

```
  <tracks>12</tracks>
```

```
</CD>
```

```
...
```

Parser-Mentalitäten

- Events auslösen: **SAX**
[xml.defined.net/sax](http://xml defined.net/sax)
- Element-Baum aufbauen: **DOM**
www.w3c.org, Microsoft MSXML, www.philo.de
- Umwandlung in OOP-Objekte
xml.defined.net
- **WHILE NOT EOF**
www.destructor.de

Event-Parser

- Für jeden gefundenen Teil wird ein Event/Callback ausgelöst
- Standard: **SAX** (Simple API for XML)
- Schnell
- Applikation muss selbst wissen, wo sie gerade „steht“
- Einfache Implementierung
- Denkbar für sehr große XML-Instanzen, die nicht komplett in den Speicher passen

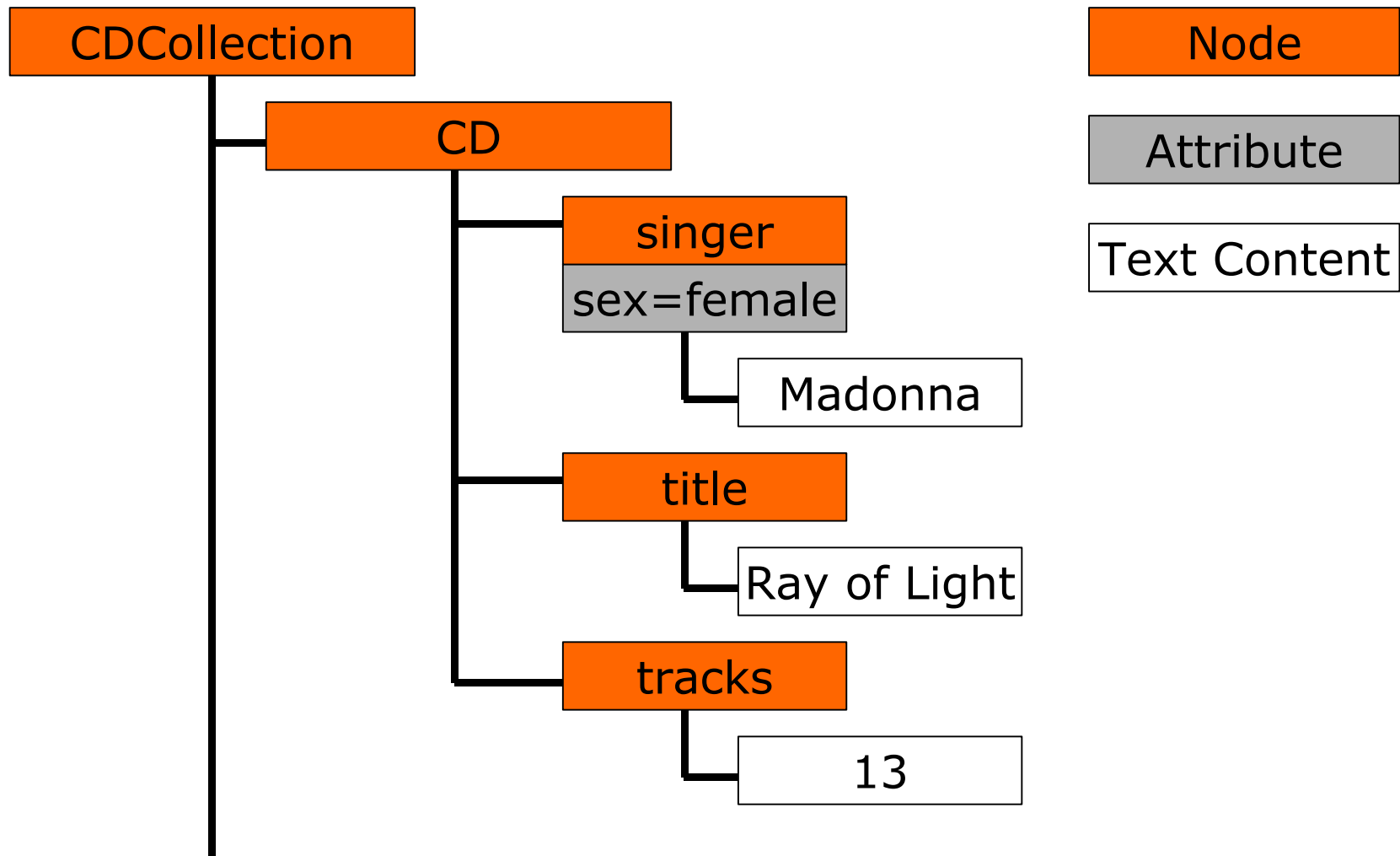
Event-Parser für Delphi

- MSXML 3 kann SAX
- XDOM/OpenXML www.philo.de
- Keith Wood xml.defined.net/sax
- XmlPartner (kein SAX)
- destructor.de TXmlScanner
(kein SAX) www.destructor.de
- weitere?

DOM-Parser

- Der XML-Baum wird komplett als Objekt-Struktur im Speicher aufgebaut
- Standard: W3C **DOM**
(Document Object Model)
- Navigieren durch den Baum über definierte Interfaces

DOM-Parser



DOM-Parser

- + Selektion von Knoten durch XPath Abfrage-Ausdrücke
- + Manipulation des Baums
- + Schreiben des Baums als XML
- + Schreiben von Teilen des Baums als XML
- Aufwändiges API
- Baum muss komplett in den Speicher passen
- Performance: Dauer für das Laden und Analysieren des Baums

DOM-Parser für Delphi

- MSXML über COM-Schnittstelle
msdn.microsoft.com/xml
- XDOM/OpenXML www.philo.de
Autor: Dieter Köhler
Aktuelle Version: 2.3.16, September 2001
XDOM.pas: 30.871 LOC, 1.201 KB

Objekt-Umwandlung

- Umwandlung von Objekt-Hierarchien in XML und zurück
- Verwendung von RTTI
- Im Moment kein „Standard“
- Sun/Java java.sun.com/xml
- Jeff Rafter xml.defined.net
(Open Source, Free)

„while not eof“-Mentalität

- Scannen durch den Baum in einem Loop
- Vorteil: Status kann in Lokal-Variablen geführt werden
- Schnell
- Denkbar für sehr große XML-Instanzen, die nicht komplett in den Speicher passen
- Einfaches API, schnell einsetzbar
- destructor.de TXmlParser

destructor.de TXmlParser

- Native Delphi
- Free, Open Source
- Keine Validierung, keine Korrektheitsprüfung
- Ablauf:
 - Erzeugen einer Instanz
 - Laden mit LoadFromFile
 - StartScan
 - WHILE Scan DO ...
 - Inhalte befinden sich in PUBLIC-Properties

Misc

Probleme des Parserbaus

- XML soll ja einfach sein
- Einen Parser zu schreiben ist es nicht
- Probleme: Validierung, genaue Umsetzung der Spec, Geschwindigkeit, Unicode-Support (UTF-16)
- Daher nur wenig ernst zu nehmende Parser-Projekte im Delphi-Bereich

Welcher Parser?

- Geschwindigkeit
- Code-Qualität
- Muss es wirklich SAX oder DOM sein?
- Verfügbarkeit, Preis, Open Source
- Portabilität (Kylix, C++, usw.)

Sites

- www.w3c.org (XML, DOM)
- www.megginson.com (SAX)
- www.destructor.de
- xml.defined.net
- www.apache.org
- www.philo.de
- www.xml.com, www.xmlhack.com,
www.xml101.com
- XML-Newsgroups auf dem Borland-Newsserver

Fragen?

</talk>