

6. Entwickler-Konferenz, 2002

XML und Delphi

Verarbeiten von XML-Dokumenten
mit Borland Delphi (und Kylix)

Stefan Heymann, stefan@destructor.de, www.destructor.de

Umgebungsbedingungen

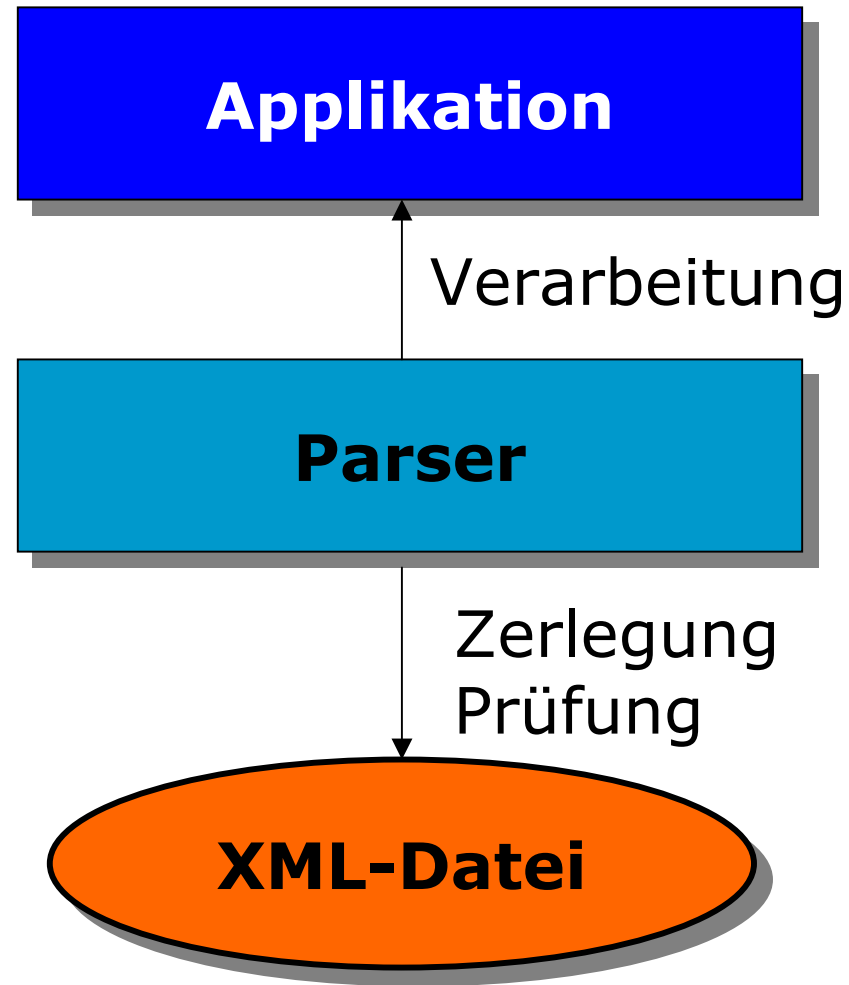
- Handy aus?
- Bei Fragen: fragen!

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
- Element-Baum aufbauen
- "while not eof"

Event-Parser

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

Event-Parser für Delphi

- MSXML 3/4 kann SAX
- XDOM/OpenXML www.philo.de/xml
- SAX2 for Pascal xml.defined.net/sax
- XmlPartner (kein SAX)
- destructor.de TXmlScanner
(kein SAX) www.destructor.de
- ...

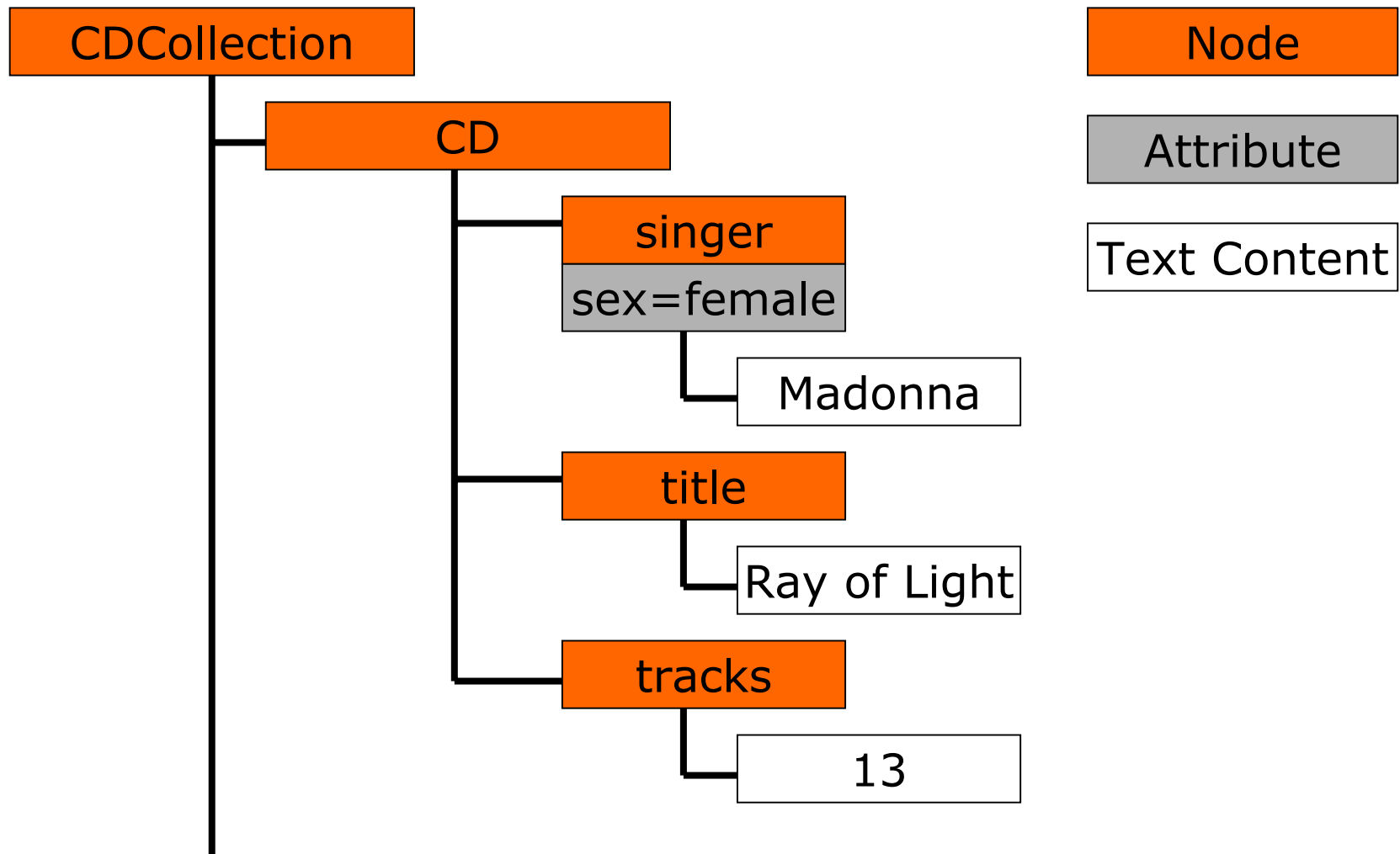
<demo/>

TXmlScanner
SAX for Delphi

DOM-Parser

- Der XML-Baum wird komplett als Objekt-Struktur im Speicher aufgebaut
- Navigieren durch den Baum über definierte Interfaces
- Manipulieren des Baums
- Abfragen (Queries) mit XPath möglich
- Standard: W3C **DOM**
(Document Object Model)

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/xml
Autor: Dieter Köhler
Aktuelle Version: 2.3.29 vom 19.08.2002
- **TXmlDocument**
Delphi 6 Enterprise
Delphi 6 Professional Update Pack 2
Kylix 2 Enterprise

<demo/>

MSXML

TXmlDocument

Objekt-Serialisierung

- Umwandlung von Objekt-Hierarchien in XML und zurück
- Verwendung von RTTI
- Im Moment kein „Standard“
- Jeff Rafter SAXSerializer in AElfried 2

Progressive Scanner

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

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

<demo/>

TXmlParser

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-8, UCS-2, UTF-16)
- Daher nur relativ wenige Parserprojekte

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)
- xml.defined.net
- www.destructor.de
- www.apache.org (Xerces)
- www.philo.de
- www.xml.com, www.xmlhack.com
- XML-Newsgroups auf dem Borland-Newsserver
<news://newsgroups.borland.com>
<http://newsgroups.borland.com>

<demo/>

Füllen eines TTreeView
mit einem XML-Baum

Fragen?

</talk>

Stefan Heymann, Tübingen
stefan@destructor.de, www.destructor.de