# XML

*,          *,               **

## 1．XML

1990

(WWW)    ,

HTML(Hyper
Text Markup Language)    . HTML
,
,

.

HTML        SGML(Standard
Generalized Markup Language)
,
DTD(Document Type Definition)
.  ,
(tag)        HTML        ,
DTD
.
, HTML
.        ,
HTML        ,

(layout)

,
[12].
, HTML
,
,
.  W3C        HTML

(Mark-up Language)    XML(extensible M
arkup Language)    1996      W3C  WG
, 1998    2              W3C
(recommendation)        XML  1.0
[10].
XML

HTML
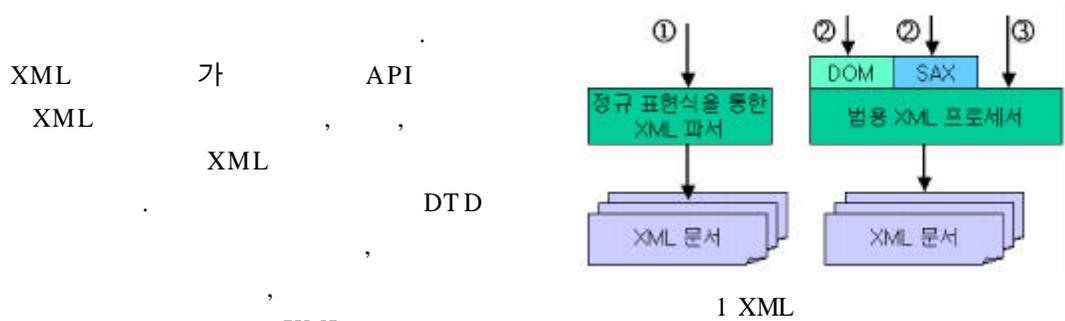SGML                    ,
.        XML

,
SGML
.

XML

XML
,
API    SAX    DOM
.        XML
,
.

## 2．XML

### 2.1 XML

XML                "XML

"    .
HTML              HTML
(element)

HTML

.                              , XML                    XML                              ,

,                              DTD

XML                              .         ,              DTD

XML              HTML                                                                  .                        ,

XML

DTD

.                    XML

DTD              XML

.

XML                                                            ,

.                    XML  API                                          .  API

,                    ,                                        API

, XSL        CSS                                        API                                  ,

/                                              API      SAX[3]      DOM[9,13]

.                                              .

.

XML                                  API

XML                                    ,        ,

XML

.                              DTD

,

,                    XML



1 XML

XML                                                    .

, XML

(well-formeness)

(validity)

valid              well-formed

## 2.2 XML

.      , valid                    ,

XML

.

DTD

1)                      (Regular Expression)                              .          1

XML                                          ,

2)  XML                              API

Java

[2]

3)              API                                            .              , Java

XML                                          [1],[8]

.

XML                              ,

| | | Validating | | DOM Level 1 | SAX | Namespace | XLink XPointer |
|---|---|---|---|---|---|---|---|
| Ælfred | Microstar | N | Java | N | Y | N | N |
| DXP | Datachannel | Y | Java | Y | N | Y | N |
| expat | James Clark | N | C | N | N | N | N |
| HEX | Anders Kristensen | N | Java | Y | Y | N | N |
| Lark | Timbray | N | Java | N | Y | N | N |
| LT XML | Language Technoogy Group | Y | C | N | N | Y | N |
| SP | James Clark | N | C++ | N | N | N | N |
| xmlproc | Lars Marius Garshol | Y | Python | N | Y | N | N |
| XML4j | IBM Alphaworks | Y | Java | Y | Y | Y | Y |
| XP | James Clark | N | Java | N | Y | N | N |
| XParse | Jeremie Miller | N | Javascript | N | N | N | N |
| SXP | Silfide project | Y | Java | Y | Y | Y | Y |
| Java Project X | Sun Microsystems | Y | Java | Y | Y | Y | N |
| Oracle XML | Oracle | Y | Java | Y | Y | Y | N |
| MSXML for Java | Microsoft | Y | Java | Y | N | Y | N |
| XML::Parser | Clark Cooper | Y | Perl | N | N | N | N |
| TclXML | Steve Ball | Y | Tcl/Tk | N | N | N | N |

## 3. XML API

1    XML            API            XML

(Event-Driven Approach)

(Object    Model    XML
Approach)[5][7]    .            XML

XML                    ,
,            XML

.            ,

,            XML

.

.            2

,

.

,

### 3.1. SAX(Simple API for XML)

## a) sample XML Document

```
<?xml version="1.0"?>
<!DOCTYPE person SYSTEM "person.dtd">
<person>
 <name>Kyu-Chul Lee</name>
 <age>39</age>
 <email>kclee@ce.cnu.ac.kr</email>
 <url type="href">http://dblab.ce.cnu.ac.kr/kclee</url>
</person>
```
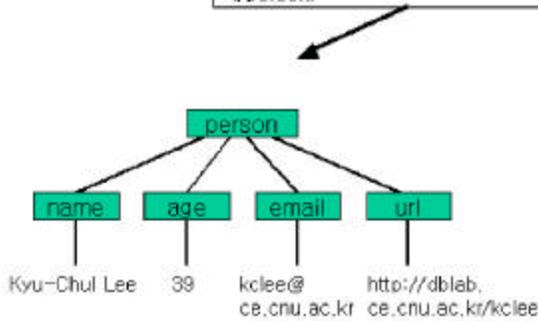
startDocument is called :
startElement is called: element name=person
startElement is called: element name=name
charaters is called : Kyu-Chul Lee
endElement is called: name
startElement is called : element name=age
characters is called : 39
endElement is called: age
startElement is called: email
characters is called: kclee@ce.cnu.ac.kr
end Element is called: emai
·····················

**b) object model approach**          **c) event-driven approach**

그림 2  object model approach 와  event-driven approach

XML

SAX[3]   XML                                   HandlerBase         클래스 (
         API  .    API   XML                                                    )

(override)

(handler)

.                               .        그림 3

SAX  1.0                                        .

1)
   Parser, AttributeList, Locator
2)
   DocumentHandler,    ErrorHandler,    D
   TDHandler, EntityResolver
3) SAX                      InputSource,
   SaxException,  SAXParserException,  Ha
   ndlerBase
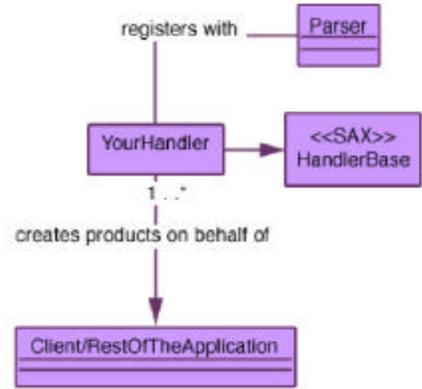


그림 3 SAX          XML

             SAX  2.0           API

  , XML Namespace              API              그림 4       그림 2 a)        XML
                              .                  XML4j   SAX API

```java
import org.xml.sax.*;
public class NotifyStr extends HandlerBase {
    static public void main(String[] argv) {
      try {
        Class c = Class.forName(argv[0]); //SAX
        // create instance of the class
        Parser parser = (Parser)c.newInstance();  //
        NotifyStr notifyStr = new NotifyStr(); //
        parser.setDocumentHandler(notifyStr);  //
        parser.parse(argv[1]);
      } catch (Exception e) {
        e.printStackTrace();
      }
    }
    public NotifyStr() {   //dummy constructor
    }
    public void startDocument() throws SAXException {
      System.out.println("startDocument is called:");
    }
    public void endDocument() throws SAXException {
      System.out.println("endDocument is called:");
    }
    public void startElement(String name, AttributeList amap) throws SAXException {
      System.out.println("startElement is called: element name=" + name);
      for (int i = 0; i < amap.getLength(); i++) {
        String attname = amap.getName(i);  //
        String type = amap.getType(i);      //
        String value = amap.getValue(i);      //
        System.out.println("  attribute name="+attname+" type="+type+" value="+value);
      }
    }
    public void endElement(String name) throws SAXException {
      System.out.println("endElement is called: " + name);
    }
    public void characters(char[] ch, int start, int length) throws SAXException { //
      System.out.println("characters is called: " + new String(ch, start, length));
    }
  }
```

4 SAX API        XML


        SAX                                                      .
            .                HandlerBase    DOM

                                                                    (navigation)
(startDocument(),              endDocument(),         ,                          /    /
StartElement(), EndElement())                                 . DOM Level 1            W3C
                                                                                        ,
            .        4                        DOM Core API    HTML    XML
        2    c)        .                                              API
                                                            , DOM  HTML API    CORE
## 3.2 DOM(Document Object Model )    API    HTML

                                                                            .
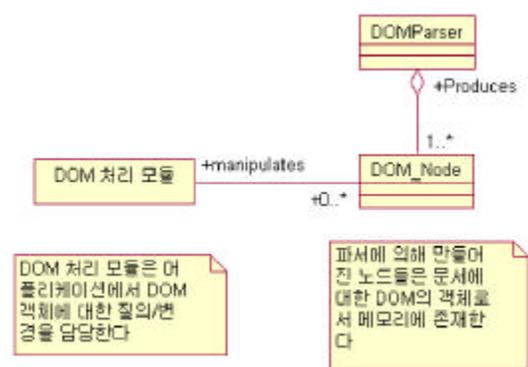DOM    HTML  XML                                    DOM  Level  2
                    (API)    ,  DOM        (Candidate  Recommendation)
                                                        , CSS          stylesheet          , UI

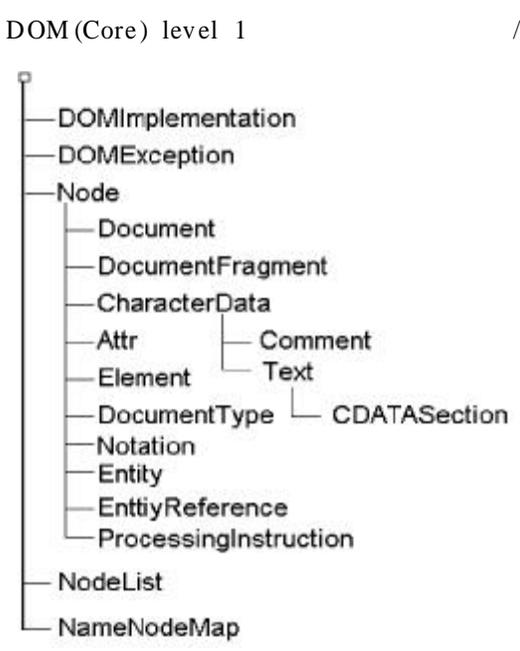,                                                                                    . Node

XML

.                                                                    (        ,        ,                    )

DOM

Node

. XML                                                                          .



(Nested-Tree)
DOM                                Element

,

.

DOM

.

DOM

,

(Interoperability)

그림 6 DOM 처리 모듈과 XML

CORBA    2.2        OMG    IDL,
ECMAScript,                    *language*
*bindings*                          .            DOM

, DOM

.                                5

DOM(Core) level 1                                            /



```
─ DOMImplementation
─ DOMException
─ Node
    ─ Document
    ─ DocumentFragment
    ─ CharacterData
    ─ Attr        ─ Comment
    ─ Element     ─ Text
    ─ DocumentType  ─ CDATASection
    ─ Notation
    ─ Entity
    ─ EnttiyReference
    ─ ProcessingInstruction
─ NodeList
─ NameNodeMap
```

그림 5 DOM(Core) level 1

/

6                                DOM API

XML

.                              DOM            XML

DOM

DOM                    .            XML

,

.    DOM            XML

.

1)

,                .

2)

DOM                                      .

3)                                            XML

.

7                                DOM API

2      a)

XML                                2      b)

DOM                                            .    1

TXDocument        XML4i

```java
import com.ibm.xml.parser.TXDocument;
import org.w3c.dom.*;
import java.io.PrintWriter;
public class MakeDocument {
    public MakeDocument(){
        try {
            //Document
            Document doc = (Document)Class.forName("com.ibm.xml.parser.TXDocument").newInstance();
            Element root = doc.createElement("person"); // <person>

            Element item = doc.createElement("name");  // <name>
            item.appendChild(doc.createTextNode("John Doe")); // <name>
            root.appendChild(item);                            // <name>

            // comment    processing instruction
            root.appendChild(doc.createComment("Proccessing Instruction for application"));
            root.appendChild(doc.createProcessingInstruction("parser", " igonoreNextLine"));

            item = doc.createElement("age");             // <age>
            item.appendChild(doc.createTextNode("35"));//
            root.appendChild(item);
            item = doc.createElement("email");           // <email>
            item.appendChild(doc.createTextNode("John.Doe@foo.com"));
            item = doc.createElement("url");             // <url>
            item.setAttribute("href", "http://www.foo.com/~John.Doe/");
            root.appendChild(item);

            //                      <person>
            doc.appendChild(root);

            // XML4j    API              XML
            ((TXDocument)doc).setVersion("1.0");
            ((TXDocument)doc).printWithFormat(new PrintWriter(System.out));
        } catch (Exception e) {
            e.printStackTrace();
        }
    }    }
```

7 DOM          XML

DOM     Document

, 2      2    DOM            DOM    XML

import           .

Document     XML

,    createXXX()                              , SAX    XML

Comment,  Processing Instruction

, appendChild()                    XML

XML                      . setAttribute()

.

## 3.3 DOM    SAX

2 DOM    SAX

| API | | / | | |
|-----|-----|-----|-----|-----|
| DOM | Tree-based | | | / |
| SAX | Event-Based | | | |

. DOM SAX API XML

SAX XML

, . /

GMD-IPSI XQL Engine , XQL, XML-QL XML

XML

DOM .

Persistent DOM , XML

DOM XML

[6]. .

DTD

## 4. XML ,

(middle-tier)

(business logic)

### 4.1 XML

XML

HTML .

XML XML XML XML

. HTML



8 XML

, HTTP
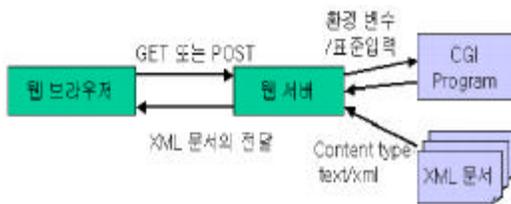
, XML XML ,

"text/xml" MIME [4] XML HTML , DBMS, IRS

XML

. (converter)

,

,

XML

,

XML .

### 4.2 XML

XML XML ,

. 9 ## 5.

XML

XML

, XML

XML

, .

. XML/XSL

/

. XML DOM . [11]

그림 9 XML

XML

, DOM, SAX

API

,

/
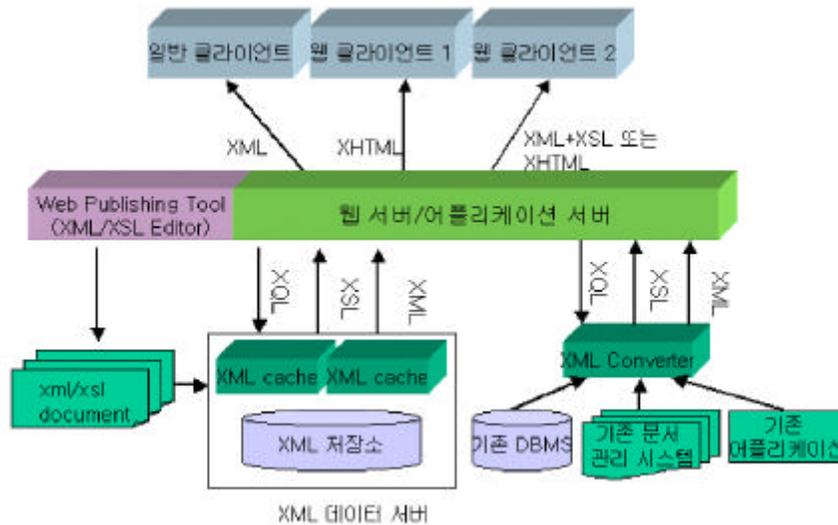
. , XML

.

XML

.

XML
HTML

,

.

## 6.

[1] Clark Cooper, "Benchmarking XML Parsers", http://www.xml.com/pub/Benchmark/article.html"

[2] David Brownell, "Conformance Testing for XML Processor", http://www.xml.com/pub/1999/09/conformance/index.html

[3] David Megginson, "Simple API for XML (SAX)", http://www.megginson.com/SAX/index.html

[4] E. Whitehead, UC Irvine "RFC 2376: XML Media Type", IETF Network Working Group

[5] Fabio Arciniegas, "An Introduction to C++ XML programming", http://www.xml.com/pub/1999/11/cplus/index.html

[6] GMD-IPSI, "GMD-IPSI XQL Engine", http://xml.darmstadt.gmd.de/xql/, GMD's XML Competence Center

[7] H. Maruyama, K. Tamura, N. Uramoto, "XML and Java Developing Web applications", Addison-Wesley

[8] Steven Marcus, "Benchmarking XML Parsers on Solaris", http://www.xml.com/pub/1999/06/benchmark/solaris.html

[9] W3 Consortium, "Document Object Model(DOM)", http://www.w3.org/DOM/

[10] W3 Consortium, "W3C Recommendation: Extensible Markup Language 1.0", http://www.w3.org/TR/1998/REC-xml-19980210"

[11] W3 Consortium "Extensible Markup Language(XML) Activity" http://www.w3c.org/XML/acitivity.html

[12]        ,        , "XML            ", EDI/EC Magagine, 98           (15)

[13]        , "DOM                    ", http://dblab.ce.cnu.ac.kr/~dolphin/xml/atoz/dom.html

CASE Center
1997    1    ~  1998    1

        :              ,  XML,            ,

E-mail: kclee@ce.cnu.ac.kr


        (Kyong-Ha Lee)
1998                                (
     )
2000                                (
     )

        :              ,  Java, XML,           ,
WWW
    E-mail: bart@ce.cnu.ac.kr


        (Kang-Chan Lee)
1994                              (
    )
1996                              (
     )
1996    -

        :            ,            , XML,
 , WWW
    E-mail: dolphin@ce.cnu.ac.kr


        ( Kyu-Chul Lee)
1984                              (
    )
1986
  (       )
1990
  (       )
   1989    3  ~

1994    3  ~  1994    6          IBM Almaden Research
Center
1995    8  ~  1996    8          Syracuse University,