XML DOM Nguyễn Hồng Phương Email: phuong.nguyenhong@hust.vn Site: http://is.hut.edu.vn/~phuongnh Bộ môn Hệ thống thông tin Viện Công nghệ thông tin và Truyền thông Đại học Bách Khoa Hà Nội

Nội dung Tổng quan XML DOM Thao tác với các nút Tham chiếu XML DOM

Tổng quan XML DOM

1. DOM là gì?
Là một chuẩn W3C
Định nghĩa chuẩn truy cập tài liệu
DOM có 3 phần
Core DOM: mô hình chuẩn cho các tài liệu có cấu trúc
XML DOM: mô hình chuẩn cho các tài liệu XML
HTML DOM: mô hình chuẩn cho các tài liệu HTML
DOM định nghĩa các đối tượng và thuộc tính của tất cả các phần tử tài liệu, và các phương thức truy cập vào chúng.

XML DOM

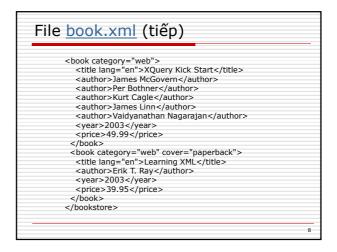
Là mô hình đối tượng chuẩn cho XML

Là một giao diện lập trình chuẩn cho XML

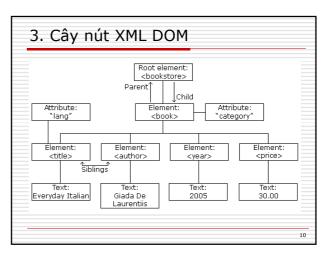
Độc lập với nền và ngôn ngữ

Định nghĩa một chuẩn cho truy cập và thao tác với tài liệu XML

2. Các nút DOM
□ Mọi thứ trong một tài liệu XML đều là nút!
■ Toàn bộ tài liệu là nút tài liệu (document node)
■ Mỗi phần tử XML là nút phần tử (element node)
■ Văn bản trong các phần tử XML là nút văn bản (text node)
■ Mỗi thuộc tính là nút thuộc tính (attribute node)
■ Chú thích cũng là nút chú thích (comment node)

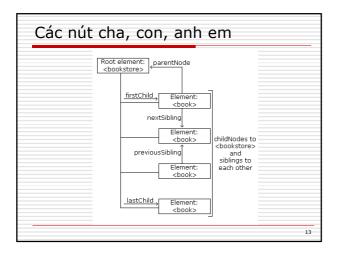


File book.xml (tiếp) □ Nút gốc là <bookstore>. Các nút khác trong tài liệu phải nằm trong nút gốc này. □ Dưới nút gốc có 4 nút <book>. □ Nút <book> đầu tiên có các nút con <title>, <author>, <year>, và <pri>everyday Italian", "Giada De Laurentiis", "2005", và "30.00" □ Chú ý: text của nút phần tử được chứa trong nút text. ■ Ví dụ: <year> 2005 Ví dụ: <year> 2005 "2005" không phải giá trị của nút phần tử year.



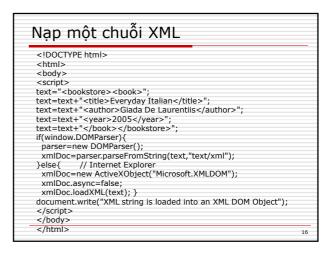
XML DOM coi một tài liệu XML là một cấu trúc cây, gọi là cây nút.
 Có thể truy cập tới tất cả các nút của cây.
 Có thể thêm mới, sửa, xóa các phần tử.

Các nút cha, con, anh em □ Các nút trong cây có mối quan hệ phân cấp với các nút khác. ■ Nút đỉnh gọi là nút gốc ■ Mỗi nút (trừ nút gốc) đều có 1 nút cha. ■ Một nút có thể có không/một/nhiều nút con ■ Nút lá là nút không có nút con. ■ Các nút anh em là các nút có cùng nút cha.



4. XML DOM Parser Hầu hết các trình duyệt đã được tích hợp sẵn XML Parser để đọc và thao tác với XML Parser biến đổi XML thành một đối tượng có thể truy cập JavaScript (XML DOM). XML DOM chứa các hàm để duyệt cây XML, truy cập, thêm và xóa các nút. Trước khi thực hiện các thao tác với tài liệu XML, cần nạp nó vào đối tượng XML DOM

Nạp một tài liệu XML <IDOCTYPE html> <html> <body> <script> if(window.XMLHttpRequest){ xhttp=new XMLHttpRequest(); } else{ // for IE 5/6 xhttp=new ActiveXObject("Microsoft.XMLHTTP"); } xhttp.send(); xmlDoc=xhttp.responseXML; document.write("XML document loaded into an XML DOM Object."); </script> </body> </html>



□ Vì lí do an toàn, trang web và file XML mà nó muốn nạp phải nằm trên cùng server. 5. Các hàm nạp XML DOM
☐ Hàm loadXMLDoc()
☐ Hàm loadXMLString()

Hàm loadXMLDoc() function loadXMLDoc(dname){ if (window.XMLHttpRequest){ xhttp=new XMLHttpRequest(); } else{ xhttp=new ActiveXObject("Microsoft.XMLHTTP"); } xhttp.open("GET",dname,false); xhttp.send(); return xhttp.responseXML; } □ File loadxmldoc.js □ File loadxmldoc1.html

```
Hàm loadXMLString()

function loadXMLString(txt){
    if (window.DOMParser){
        parser=new DOMParser();
        xmlDoc=parser.parseFromString(txt,"text/xml");
    }
    else{ // Internet Explorer
        xmlDoc=new ActiveXObject("Microsoft.XMLDOM");
        xmlDoc.loadXML(txt);
    }
    return xmlDoc;
}

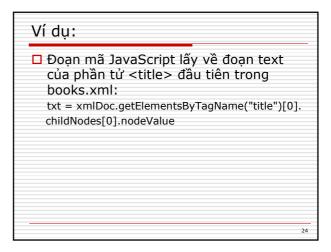
□ File loadxmlstring.js

□ File loadxmlstring1.html
```

6. Giao diện lập trình DOM mô hình XML như là một tập các đối tượng nút. Có thể viết các đoạn mã JavaScript, Java, C#,.... để truy cập vào các nút này. Giao diện lập trình: xác định thông qua một tập các thuộc tính và phương thức chuẩn.

Thuộc tính XML DOM Có một số thuộc tính đặc trưng x.nodeName: tên của x x.nodeValue: giá trị của x x.parentNode: nút cha của x x.childNodes: các nút con của x x.attributes: các nút thuộc tính của x

Phương thức XML DOM x.getElementsByTagName(name): lấy về tất cả các phần tử mà tag có tên là name x.appendChild(node): thêm một nút con vào nút x x.removeChild(node): loại một nút con ra khỏi nút x



7. Truy cập vào các nút □ Có ba cách ■ sử dụng phương thức getElementsByTagName() ■ sử dụng vòng lặp để duyệt cây nút ■ định hướng trong cây, sử dụng mối quan hệ nút.

25

Phương thức getElementsByTagName() Trả về tất cả các phần tử có tên thẻ xác định. Cú pháp: node.getElementsByTagName("tagname"); Ví dụ: trả về tất cả các phần tử <title> dưới nút x: x.getElementsByTagName("title"); trả về tất cả các phần tử <title> của tài liệu XML xmlDoc.getElementsByTagName("title");

Danh sách nút DOM

□ Phương thức getElementsByTagName()
trả về một danh sách nút – tức là một
mảng các nút.

■ Ví dụ:
□ xmIDoc=loadXMLDoc("books.xml");
x=xmIDoc.getElementsByTagName("title");
□ Truy cập vào phần tử thông qua chỉ số.
Chỉ số bắt đầu từ 0.

■ Ví dụ: truy cập phần tử <title> thứ 3
□ y=x[2];

Chiều dài danh sách nút DOM

Sử dụng thuộc tính length của danh sách nút: cho biết số nút.

Sử dụng vòng lặp để duyệt

Ví dụ:

xmlDoc=loadXMLDoc("books.xml");
x=xmlDoc.getElementsByTagName("title");
for (i=0;i<x.length;i++){
 document.write(x[i].childNodes[0].nodeValue);
 document.write("
 ");
}

Ví dụ: vidu1.html

<!DOCTYPE html>
<html>
<htead>
<script src="loadxmldoc.js"></script>
</head>
<body>
<script>
xmlDoc=loadXMLDoc("books.xml");
x=xmlDoc.getElementsByTagName("title");
for (i=0;i<x.length;i++){
 document.write(x[i].childNodes[0].nodeValue);
 document.write(x[i].childNodes[0].nodeValue);
}
</script>
</script>
</body>
</html>

```
Các kiểu nút

Thuộc tính documentElement của tài liệu XML là nút gốc.

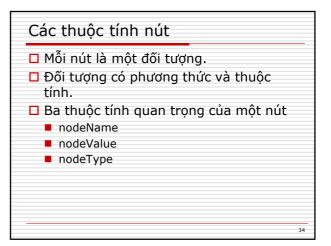
Thuộc tính nodeName của nút là tên của nút.

Thuộc tính nodeType của nút là kiểu của nút
<!DOCTYPE html>
<html>
<html>
<head>
<script src="loadxmldoc.js"></script>
</head>
<body>
<script>
xmlDoc=loadXMLDoc("books.xml");
document.write(xmlDoc.documentElement.nodeName);
document.write(xmlDoc.documentElement.nodeType);
</script>
</script>
</body>
</html>
```

```
Ví dụ: vidu3.html

<!DOCTYPE html>
<html>
<head>
<script src="loadxmldoc.js"></script>
</head>
<body>
<script>
xmlDoc=loadXMLDoc("books.xml");
x=xmlDoc.documentElement.childNodes;
for (i=0;i<x.length;i++){
    if (x[i].nodeType==1){//Process only element nodes (type 1)
        document.write(x[i].nodeName);
        document.write(x[i].nodeName);
    }
}
</script>
</script>
</body>
</html>
```

8. Thông tin nút XML DOM



Thuộc tính nodeName

Cho biết tên của 1 nút

Read-only

nodeName của một nút phần tử chính là tên thẻ

nodeName của một nút thuộc tính chính là tên thuộc tính
nodeName của một nút text là #text
nodeName của một nút tài liệu là #document

```
Ví dụ: vidu5.html

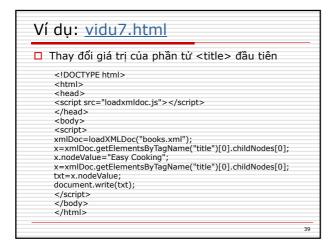
<!DOCTYPE html>
<ntml>
<head>
<script src="loadxmldoc.js"></script>
</head>
<body>

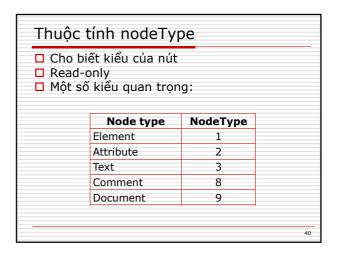
<script>
xmlDoc=loadXMLDoc("books.xml");

document.write(xmlDoc.documentElement.nodeName);
</script>
</body>
</html>
```

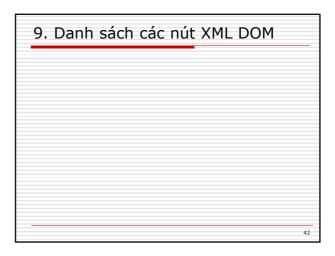
Thuộc tính nodeValue Cho biết giá trị của nút nodeValue của các nút phần tử là không xác định nodeValue của nút text chính là text nodeValue của nút thuộc tính chính là giá trị thuộc tính Lấy về giá trị của một phần tử







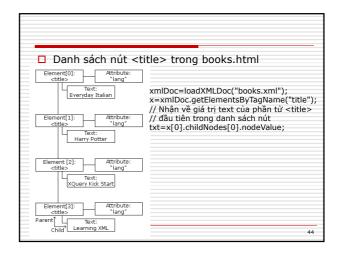
Ví dụ: <u>vidu8.html</u>	
html <html> <head> <script src="loadxmldoc.js"></script> </head> <body> <script> xmlDoc=loadXMLDoc("books.xml"); document.write(xmlDoc.documentElement.nodeName); document.write(" document.write(" c/script> </script> </body> </html>	
	41



Danh sách nút DOM

- Khi sử dụng thuộc tính hay phương thức (như childNodes hoặc getElementsByTagName) sẽ nhận được đối tượng trả về là một danh sách nút, theo trình tư như trong tài liêu XML.
- □ Nút có thể được truy cập với chỉ số bắt đầu từ 0

43



Chiều dài danh sách nút

- Khi một phần tử được thêm vào hoặc xóa đi, danh sách sẽ được tự động cập nhật.
- ☐ Thuộc tính length của một danh sách nút cho biết số nút có trong danh sách.

xmlDoc=loadXMLDoc("books.xml");
//the x variable will hold a node list
x=xmlDoc.getElementsByTagName('title');
for (i=0;i<x.length;i++)
{
 document.write(x[i].childNodes[0].nodeValue);
 document.write("
");
}

Danh sách thuộc tính DOM

- Thuộc tính attributes của một nút phần tử trả về một danh sách các nút thuộc tính.
- Nếu một thuộc tính được thêm vào hay xóa đi, danh sách sẽ dược tự động cập nhật.

xmlDoc=loadXMLDoc("books.xml");

x = xmlDoc.getElementsByTagName("book")[0].attributes;

 $\label{local-document} document.write(x.getNamedItem("category").nodeValue); \\ document.write("
br>" + x.length); \\$

vidu10.html

10. Duyệt cây nút

Duyệt cây nút

<head:

<script src="loadxmlstring.js"></script>

</head>

<body>

text="<book>";

text=text+"<title>Everyday Italian</title>";

text=text+"<author>Giada De Laurentiis</author>"; text=text+"<year>2005</year>";

text=text+"</per>

xmlDoc=loadXMLString(text);

48

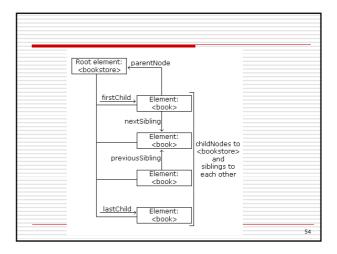
// documentElement always represents the root node x=xmlDoc.documentElement.childNodes; for (i=0;i<x.length;i++) { document.write(x[i].nodeName); document.write(": "); document.write(x[i].childNodes[0].nodeValue); document.write("
 document.write("

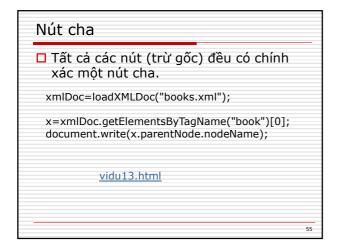


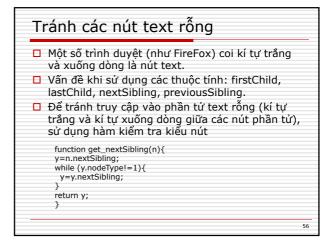
Các trình duyệt DOM Hầu hết các trình duyệt hiện đại đều hỗ trợ các đặc tả W3C DOM. Có một chút khác biệt: IE không coi các kí tự khoảng trắng/ kí tự xuống dòng là các nút text Các trình duyệt khác thì có! Hãy mở file vidu12.html bằng các trình duyệt khác nhau!

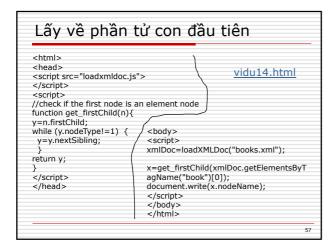


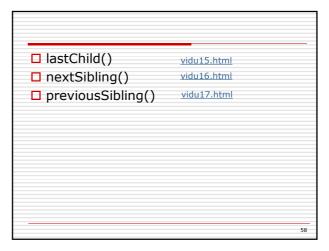
Định hướng trong cây
 □ Truy cập vào các nút thông qua mối quan hệ giữa các nút gọi là "định hướng" trong cây.
 □ Các thuộc tính của nút
 ■ parentNode
 ■ childNodes
 ■ firstChild
 ■ lastChild
 ■ nextSibling
 ■ previousSibling











Thao tác với các nút

13. Lấy về các giá trị DOM

☐ Lấy về giá trị của một phần tử
☐ Lấy về giá trị của một thuộc tính

Lấy về giá trị của một phần tử

- ☐ Các nút phần tử không có giá trị text
- Text của một nút phần tử được lưu trữ trong nút con, gọi là nút text.
- => để lấy về text của nút phần tử, ta lấy giá trị của nút con (nút text).
- Phương thức getElementsByTagName() trả về một danh sách nút chứa tất cả các phần từ với tên thẻ theo tham số của phương thức này. Trật tự các phần tử theo trật tự xuất hiện của chúng ở trong tài liệu
- Thuộc tính childNodes trả về một danh sách các nút con
- Thuộc tính nodeValue trả về giá trị text của nút text



Lấy về giá trị của một thuộc tính

- ☐ Trong DOM, các thuộc tính là các nút.
- □ Nút này có giá trị text.
- Để lấy giá trị text của thuộc tính, sử dung
 - phương thức getAttribute() trả về một giá trị thuộc tính
 - thuộc tính nodeValue

63

xmlDoc=loadXMLDoc("books.xml");
txt=xmlDoc.getElementsByTagName("title")[0].getAttribute("lang");
document.write(txt);

vidu19.html
vidu20.html

□ Phương thức getAttributeNode() trả về
một nút thuộc tính.

xmlDoc=loadXMLDoc("books.xml");
x=xmlDoc.getElementsByTagName("title")[0].getAttributeNode("lang");
txt=x.nodeValue;

vidu21.html
vidu22.html

14. Thay đổi giá trị nút/thuộc tính

- ☐ Thay đổi giá trị nút: thuộc tính nodeValue
- ☐ Thay đổi giá trị thuộc tính: phương thức setAttribute()

xmlDoc=loadXMLDoc("books.xml");
x=xmlDoc.getElementsByTagName("title")[0].childNodes[0];
x.nodeValue="Easy Cooking";

vidu23.html

Thay đổi giá trị của nút text

Thay đổi giá trị của thuộc tính xmlDoc=loadXMLDoc("books.xml"); x=xmlDoc,getElementsByTagName('book'); x[0]_setAttribute("category","food"); vidu24.html xmlDoc=loadXMLDoc("books.xml"); x=xmlDoc.getElementsByTagName("book")[0] y=x.getAttributeNode("category"); y.fodeValue="food") vidu25.html

15. Loại bỏ nút trong DOM Sử dụng phương thức removeChild() để xóa một nút Sử dụng phương thức removeAttribute() để xóa một thuộc tính

□ Xóa text của nút text
 xmlDoc=loadXMLDoc("books.xml");
 x=xmlDoc.getElementsByTagName("title")[0].childNodes[0];
 x.nodeValue="";
□ Loại bỏ nút thuộc tính sử dụng tên
 xmlDoc=loadXMLDoc("books.xml");
 x=xmlDoc.getElementsByTagName("book");
 x[0].removeAttribute("category");
□ Loại bỏ các nút thuộc tính sử dụng tham
 số đối tượng
 xmlDoc=loadXMLDoc("books.xml");
 x=xmlDoc.getElementsByTagName("book");
 for (i=0;i<x.length;i++){
 while (x[i].attributes.length>0) {
 attnode=x[i].attributes[0];
 old_att=x[i].removeAttributeNode(attnode);
 }
 }
 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /*

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /**

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

 /*

16. Thay thế một nút

Thay thế một nút phần tử

Thay thế dữ liệu trong mọt nút text

Sử dụng thuộc tính nodeValue

Thay thế một nút phần tử

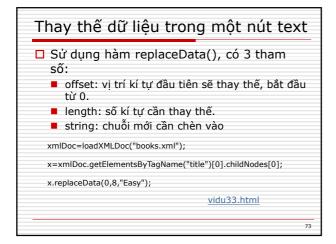
Sử dụng hàm replaceChild()
xmlDoc=loadXMLDoc("books.xml");

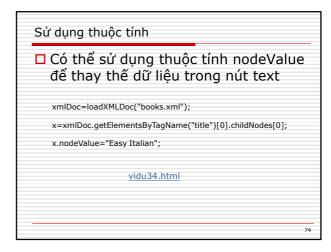
x=xmlDoc.documentElement; vidu32.html

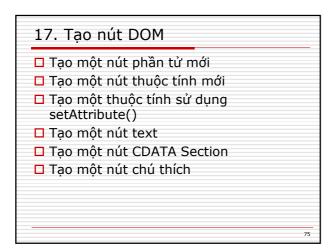
//create a book element, title element and a text node
newNode=xmlDoc.createElement("book");
newTitle=xmlDoc.createElement("title");
newText=xmlDoc.createTextNode("A Notebook");

//add the text node to the title node,
newTitle.appendChild(newText);
//add the title node to the book node
newNode.appendChild(newTitle);

y=xmlDoc.getElementsByTagName("book")[0]
//replace the first book node with the new node
x.replaceChild(newNode,y);









Tạo một nút thuộc tính mới

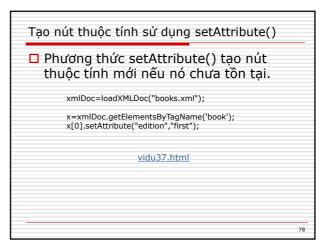
□ Sử dụng phương thức createAttribute()

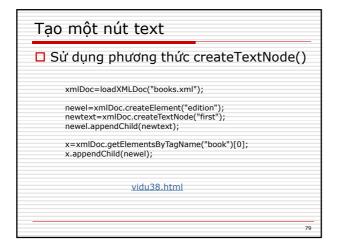
xmlDoc=loadXMLDoc("books.xml");

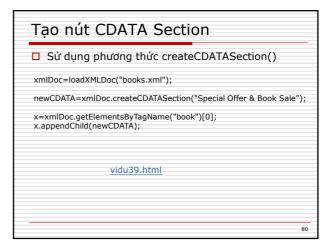
newatt=xmlDoc.createAttribute("edition");
newatt.nodeValue="first";

x=xmlDoc.getElementsByTagName("title");
x[0].setAttributeNode(newatt);

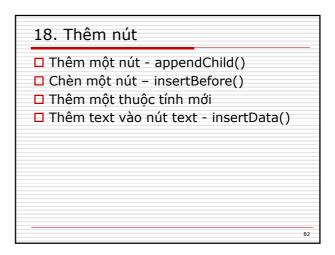
vidu36.html



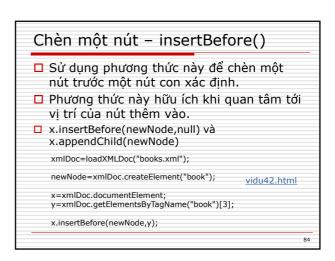


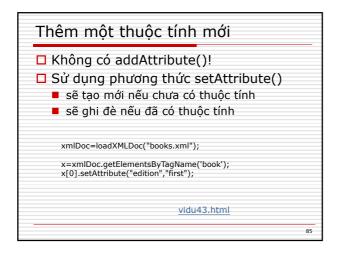


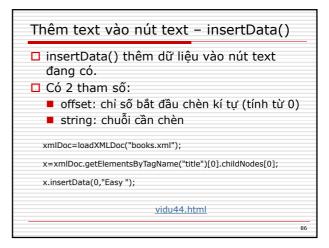
Tạo nút comment □ Sử dụng phương thức createComment() xmIDoc=loadXMLDoc("books.xml"); newComment=xmIDoc.createComment("Revised March 2008"); x=xmIDoc.getElementsByTagName("book")[0]; x.appendChild(newComment); vidu40.html

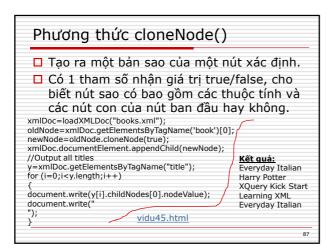


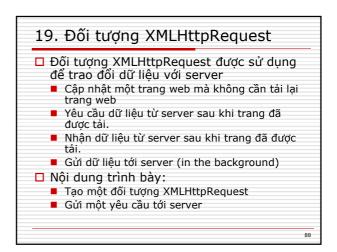
Thêm một nút - appendChild() □ Dùng phương thức này để thêm một nút con vào nút hiện tại. xmlDoc=loadXMLDoc("books.xml"); newel=xmlDoc.createElement("edition"); x=xmlDoc.getElementsByTagName("book")[0]; x.appendChild(newel); vidu41.html

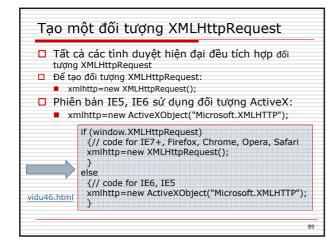


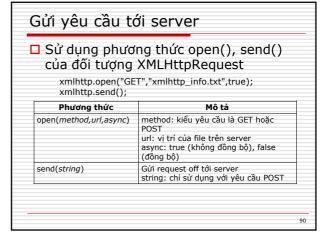












GET hay POST

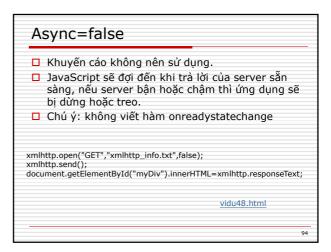
- GET đơn giản và nhanh hơn POST, có thể được sử dụng trong hầu hết các trường hợp.
- ☐ Sử dụng POST khi:
 - Cập nhật 1 file hoặc CSDL trên server
 - Gửi một lượng lớn dữ liệu tới server (POST không giới hạn kích thước)
 - Gửi dữ liệu nhập của người dùng (có thể chứa các kĩ tự chưa biết), POST mạnh và an toàn hơn GET.

91

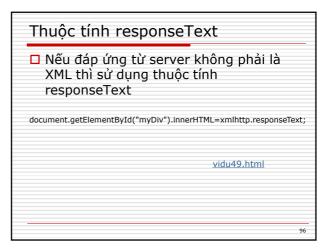
Đồng bộ hay không đồng bộ?

- Gửi yêu cầu không đồng bộ, JavaScript không phải đợi trả lời của server, mà có thể:
 - thực hiện các script khác trong khi đợi trả lời (response) của server.
 - thu nhân trả lời khi trả lời sẵn sàng

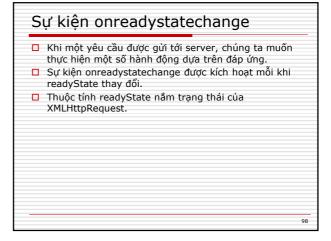
92



Server Response Dể nhận đáp ứng từ server, sử dụng thuộc tính responseText hoặc responseXML của đối tượng XMLHttpRequest Thuộc tính Mô tả responseText Nhận dữ liệu đáp ứng như là một xâu responseXML Nhận dữ liệu đáp ứng như là dữ liệu XML



Thuộc tính responseXML Nếu đáp ứng từ server là XML, để phân tích nó như là một đối tượng XML, ta sử dụng thuộc tính responseXML. xmlDoc=xmlhttp.responseXML; var txt=""; x=xmlDoc.getElementsByTagName("ARTIST"); for (i=0;i<x.length;i++) { txt=txt + x[i].childNodes[0].nodeValue + " ;; } document.getElementById("myDiv").innerHTML=txt;



Sự kiện onreadystatechange (tiếp) Ba thuộc tính quan trọng của XMLHttpRequest Thuộc tính Mô tả onreadystatechange Lưu trữ một hàm/tên một hàm sẽ được tự động gọi khi thuộc tính readyState thay đổi readyState Nằm giữ trạng thái của đối tượng XMLHttpRequest, thay đổi từ 0 đến 4: 0: yêu cầu không được khởi tạo 1: kết nổi server được thành lập 2: đã nhận yêu cầu 3: xử lý yêu cầu 4: yêu cầu đã hoàn tất và đáp ứng đã sẵn sàng status 200: "OK" 404: Page not found

```
Sự kiện onreadystatechange (tiếp)

Khi readyState là 4 và status là 200, đáp ứng sẵn sàng:

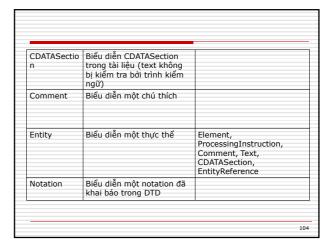
xmlhttp.onreadystatechange=function()
{
if (xmlhttp.readyState==4 && xmlhttp.status==200)
{
document.getElementById("myDiv").innerHTML=xmlhttp.responseText;
}
}

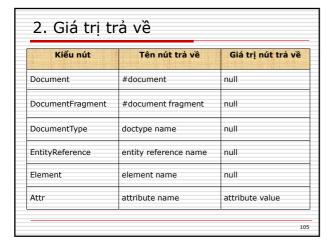
vidu51.html
```

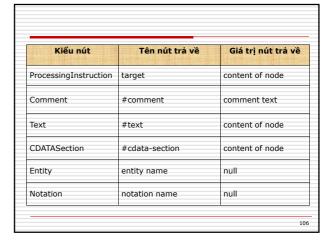
Phụ lục: Các kiểu nút

Kiểu nút	Mô tả	Kiểu nút có thể là nút con	
Document	Biểu diễn một văn bản hoàn chỉnh (nút gốc của cây DOM)	Element (tối đa chỉ có 1) ProcessingInstruction, Comment, DocumentType	
DocumentFr agment	Biểu diễn đối tượng tài liệu nằm giữ một phần tài liệu	Element, ProcessingInstruction, Comment, Text, CDATASection, EntityReference	
DocumentTy pe	Cung cấp một giao diện tới các thực thể đã định nghĩa trong tài liệu		
ProcessingIn struction	Biểu diễn một chỉ thị xử lý		

EntityRefere nce	Biểu diễn một tham chiếu thực thể	Element, ProcessingInstruction, Comment, Text, CDATASection, EntityReference
Element	Biểu diễn một phần tử	Element, Text, Comment, ProcessingInstruction, CDATASection, EntityReference
Attr	Biểu diễn một thuộc tính	Text, EntityReference
Text	Biểu diễn nỗi dung text trong một phần tử hoặc một thuộc tính	







3. Kiểu nút – các hằng

Kiểu nút	Hằng
1	ELEMENT_NODE
2	ATTRIBUTE_NODE
3	TEXT_NODE
4	CDATA_SECTION_NODE
5	ENTITY_REFERENCE_NODE
6	ENTITY_NODE

Kiểu nút	Hằng
7	PROCESSING_INSTRUCTION_NODE
8	COMMENT_NODE
9	DOCUMENT_NODE
10	DOCUMENT_TYPE_NODE
11	DOCUMENT_FRAGMENT_NODE
12	NOTATION_NODE

