

# Webbed Documents

Malcolm Graham and  
WriteDoc Inc.  
malcolm@writedoc.com

Andrew Surray  
Northern Telecom  
surray@bnr.ca

## Abstract

This paper describes the work currently being done within Northern Telecom (Nortel) Department X754 (located in Ottawa, Canada) to resolve problems with:

- distributing documents for internal review;
- ensuring that documents are available for review;
- printing documents at remote locations; and
- reading documents on different computing platforms.

These problems are being resolved by:

- using Web technology to distribute documents;
- providing immediate online access to documents;
- using a cross-platform, portable, file format;
- providing document-navigation pages; and
- using **Netscape** and **Acrobat** together.

## The Need for Online Documents

We live in a world where the pace of change has accelerated to the point where as soon as a document is published it is out of date before the intended reader ever sees it. In addition, many people seem to be addicted to the printed page.

In order to function efficiently and effectively in today's world, people need current and accurate information as quickly as possible. We believe that while we will continue to rely heavily upon printed media for some time to come, there is a better way to give people up-to-date information. That way is to provide it online.

## The Problem and What We've Already Tried

The development of documents within Nortel involves the concurrent review of documents by different design groups dispersed throughout the world. This section describes the various ways that we have tried to ensure the efficient and effective review of our documents.

Determining who should review documents and distributing documents to them is challenging!

The traditional method of mailing printed documents to reviewers is impractical because of mailing delays—due to internal and external mailing systems, customs examinations, and finding reviewers when they have moved. Producing printed documents also causes delays.

Since all Nortel locations are connected by an internal Ethernet network, we tried providing global online access to printable versions of documents. Typically, this involves creating a PostScript version of a document and letting people know where it is available on the network. This approach has several drawbacks. PostScript files are much larger than the original source files and they have to be transferred across the network before they can be printed locally. Just obtaining a printed document is a very time consuming task for the reviewer. We find this approach impractical for distributing large amounts of information.

Providing direct online access to document source files has also proven to be impractical because the source application may not be available at all locations or, when it is, people don't know how to use it. In addition, the document source files are often unreadable because they were created on a computing platform that the reviewer doesn't have access to.

An analysis of the problems encountered when trying to review documents can be categorized as document:

- distribution and availability problems; and
- format and portability problems.

## The Solution or What We're Doing Now

The solution to the problems with document distribution, availability, format, and portability, has led X754 to the use and merging of two different technologies: the World-Wide Web and portable document technologies.

The solution will be described in terms of:

- introducing the World-Wide Web;
- introducing Acrobat; and
- prototyping the solution.

### Introducing the World-Wide Web

To resolve problems with document distribution and availability, the use of the World-Wide Web technology was selected. The World-Wide Web will be described by answering the following questions:

- What is the Web?
- What is HTTP?
- What is a URL?
- What is Hypermedia?
- What is HTML?

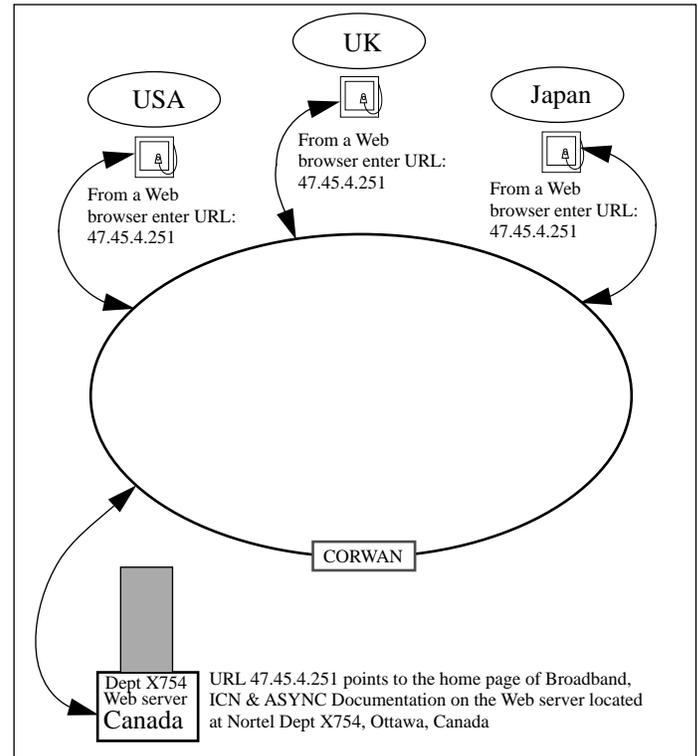
**What is the Web?** On the publicly accessible Internet, the World-Wide Web is that part of the Internet that transmits and receives hypermedia information over a “web” of hypermedia links. The information is accessed from Web servers using Web browsers. Within Nortel, the Web is a subnetwork of the privately accessible Nortel corporate wide-area network (CORWAN), and is composed of Ethernet-connected Web servers and Web browsers.

**What is HTTP?** The hypertext transport protocol (HTTP) is used when transferring hypermedia information between Web servers and Web browsers. This protocol must be specified when using a Web browser to communicate with a Web server; for example, `http://47.45.4.251/`.

**What is a URL?** A universal resource locator (URL) is the Internet protocol (IP) address of a piece of hypermedia information on a Web server. For example, the URL of the home page of the X754 Web server is 47.45.4.251. This identifies the server in terms of network/subnetwork/host machine. Addresses can also be described in words, using a naming convention that is administered by a domain-name server. For example, a domain-name server would map `public.bnr.ca` to IP address 192.58.194.164.

When URL 47.45.4.251 is specified using a Web browser installed on a computer that is connected to the Nortel CORWAN, the browser communicates with the X754 Web server, which then transmits the home-page information to the local browser, where it is displayed. See Figure 1.

**Figure 1**  
**Accessing Information at URL 47.45.4.251**



**What is Hypermedia?** Hypermedia is technology that allows information developers to create links between various forms of media, such as text, graphics, animation, and sound. In a hypertext document, link indicators typically appear as underlined text (see Figure 2 and Figure 3). When a user clicks on a link indicator, the Web browser “jumps” to the linked information. The linked information appears as a contiguous document. Hypermedia links between Web pages are defined using the hypertext markup language (HTML).

**What is HTML?** HTML is a subset of the standard generalized markup language (SGML). HTML is a set of tags that define the structure of information appearing on a Web page. Each piece of information has an associated HTML tag. Figure 4 illustrates the HTML source file that defines the information and structure of the Web page illustrated in Figure 3.

Currently, HTML tags define a basic structure limited to titles, headers, paragraphs, simple lists, and graphics, with a common font for all text. HTML gives little control over how information appears on a Web page, this is determined when the HTML tags are interpreted by a Web browser.

### Introducing Acrobat

To resolve problems with document format and portability, the use of Adobe's Acrobat technology was selected.

Adobe Acrobat uses a technology for creating portable document format (PDF) files that can be viewed on any platform (Macintosh, IBM, and Unix) using the **Acrobat Reader** product. The Acrobat technology uses Joint Photographic Experts Group (JPEG) compression to produce PDF file sizes that are from 2 to 10 times smaller than the original source files.

Adobe Acrobat has three main components:

- Acrobat Reader;
- Acrobat Exchange; and
- Acrobat Distiller.

**Acrobat Reader** is a tool designed for corporate and commercial publishers who distribute finished documents to large audiences. It is also useful to anyone wishing to view, navigate, and print PDF files. For non-commercial use, **Acrobat Reader** may be distributed free of charge.

**Acrobat Exchange** enables users to create, search, and annotate electronic documents. It also allows users to view, assemble, print, password protect, add bookmarks to PDF files, and create links to other PDF files.

**Acrobat Distiller** converts PostScript files to PDF files.

**Acrobat Distiller** is appropriate for users of desktop publishing applications, for information developers using equation and other specialty fonts, and for converting files with encapsulated PostScript (EPS) artwork.

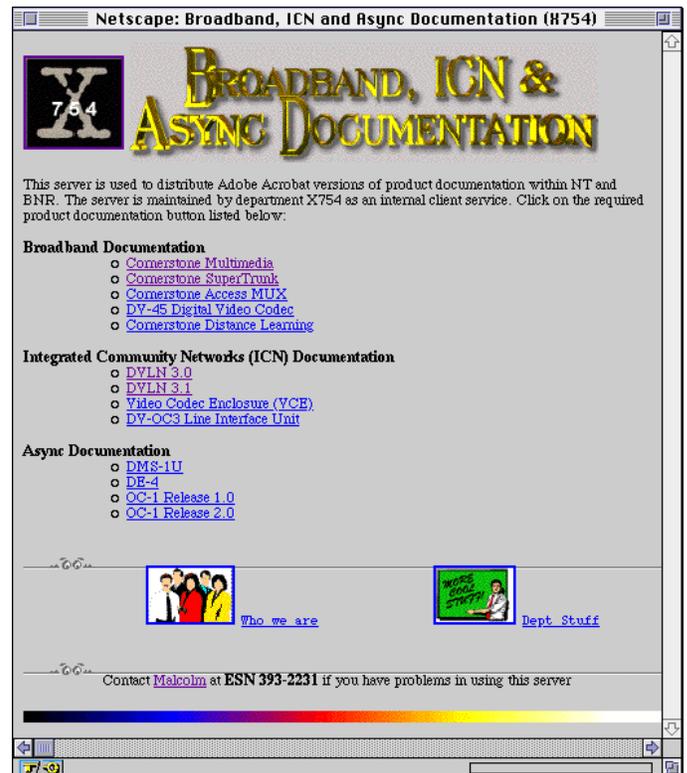
### Prototyping the Solution

Creating the prototype involved:

- creating Web pages;
- creating Acrobat PDF files; and
- setting up a Web server.

**Creating Web Pages.** Web pages are created using HTML. A Web home page was created as an entry-level page for accessing the X754 Web server. A link menu on the home page (see Figure 2) provides access to what we've called "document-navigation pages." Document-navigation pages provide link menus to documents associated with specific products. For example, clicking on the **Cornerstone Distance Learning** link indicator on the home page displays the document-navigation page for the **Cornerstone Distance Learning** product (see Figure 3).

**Figure 2**  
**Entry-level Home Page**

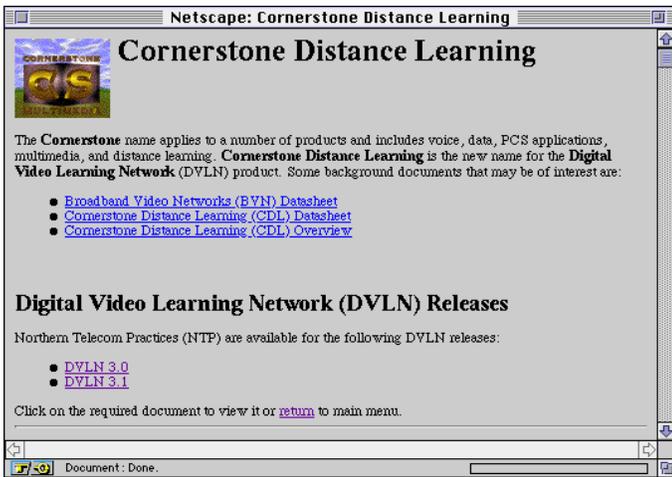


Before the document-navigation pages could be coded using HTML, a file-naming convention and directory structure had to be defined.

A top-level directory called **WWW** was created. Within the **WWW** folder, the server software was installed, and another folder called **X754** was created.

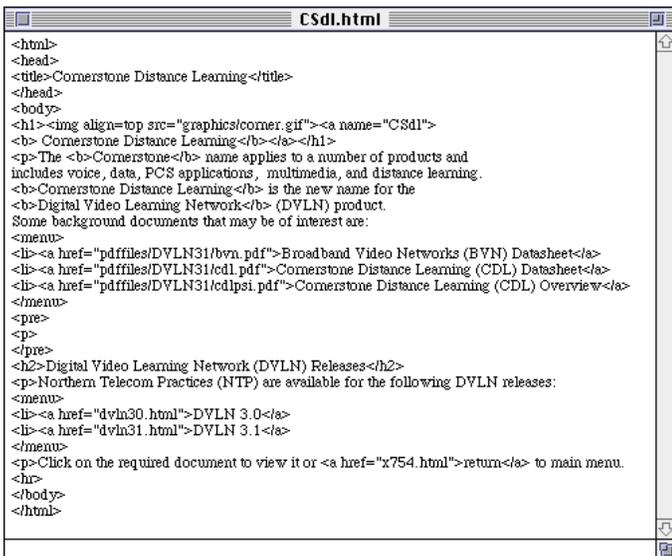
Within the **X754** folder, HTML files were created for the entry-level home page and each document-navigation page. Also within the **X754** folder, a folder called **PDFfiles** was created. Within the **PDFfiles** folder, individual product folders were created for storing the document PDF files associated with each product.

**Figure 3**  
**CDL Product Document-Navigation Page**



The HTML code for the Cornerstone Distance Learning (CDL) document-navigation page, shown in Figure 3, is presented in Figure 4.

**Figure 4**  
**HTML code for the CDL Document-Navigation Page**



**Creating Acrobat PDF Files.** Creating the Acrobat PDF files involved five steps.

1. Saving each Macintosh **FrameBuilder** book file to disk as an individual PostScript file.
2. Processing each PostScript file using **Acrobat Distiller** to create individual PDF files.
3. Combining each individual PDF file to create a single PDF file for the complete document.

4. Adding bookmarks and hypertext links between chapters within the document.
5. Saving the PDF file to the appropriate product folder within the **PDFfiles** folder on the X754 Web server.

*Note:* On average, the final PDF file sizes were half the size of the original **FrameBuilder** source files.

**Setting Up a Web Server.** For the prototype, the X754 Web server was set up on a Macintosh; this involved completing five steps.

1. Determining the Ethernet IP address on the Macintosh where the Web server was to be located. The IP address is used as the URL of the Web server.
2. Installing the **MacHTTP** software in the **WWW** folder.
3. Replacing the contents of the **default.html** file (within the **WWW** folder) with the contents of the entry-level home page HTML file.
4. Placing an alias to the **MacHTTP** software within the **Startup items** folder in the Macintosh **System Folder**. This ensures that the **MacHTTP** server software is automatically started whenever the Macintosh is restarted.
5. Starting the **MacHTTP** software.

These actions brought up the server, making it accessible by anyone, with access to the Nortel CORWAN network, using the HTTP protocol with the servers' URL.

*Note:* The Macintosh used to run the Web server must be left running all the time to ensure access from different time zones throughout the world.

### **Conclusion or Were We Successful?**

Based on using the prototype for three months to distribute multiple draft versions of 14 documents, with over 1000 pages, to over a dozen reviewers based in Ottawa, Montreal, North Carolina, and Georgia we would say that the prototype was successful.

Revised documents were viewed or printed on the same day that they were available for review. The document reviewers had fewer problems in downloading, viewing, and printing the documents to be reviewed. We even had one reviewer adding comments to an online document, using **Acrobat Exchange**, and returning the commented document back to us as an email attachment.

However, we did encounter problems with downloading **Acrobat Reader** and having to manually create hypertext links.

Having to download **Acrobat Reader** and configuring **Netscape** to recognize **Acrobat Reader** files created some problems. Because the next release of **Netscape** will include the ability to read **Acrobat Reader** files, these problems will disappear.

Hypertext links between different **FrameBuilder** book files are not maintained when the PDF file is created; they have to be manually added to the PDF file using **Acrobat Exchange**. The next release of **FrameBuilder** will maintain the hypertext links between different book files within the PDF file. **FrameBuilder** will also be able to create PDF files directly.

### **Future Developments or What Could Be Done Next**

Future development considerations include:

- larger implementation of the prototype;
- greater interaction;
- expanded use of hypermedia;
- authoring in HTML;
- external customer access; and
- intelligent front end.

**Larger Implementation of the Prototype.** The prototype is currently being expanded and implemented across all Nortel Transport and Broadband Networks product documentation.

**Greater Interaction.** Interaction is encouraged in at least two ways.

- Using **Acrobat Exchange**, reviewers could annotate online documents, as part of a review process. Annotations by several reviewers at several locations can be imported and collated in a single document.
- Electronic forms would permit users to specify their information requirements and order printed documentation.

**Expanded Use of Hypermedia.** Future Web documents will link sound, animation, and video, as well as the text and graphics currently employed.

**Authoring in HTML.** As HTML develops, future versions may provide control over the appearance of information elements on a Web page (rather than simply providing structure). At that time, it will be feasible to author a complete document in HTML.

**External Customer Access.** Could be implemented in three steps.

1. Placing document PDF files on a Web server accessible to users outside of Nortel's CORWAN.
2. Establishing a security gate through which registered customers pass before acquiring access to segments of the documentation database.
3. Establishing a credit account for customers who view, print, or download Nortel documentation.

**Intelligent Front End.** An intelligent front end would permit easier access to information.

- Information could be assigned an information type, such as installation, marketing, alarm, or test. An information type could then be identified and retrieved across multiple documents to form one document.
- For loosely structured or minimally indexed documents, keyword and boolean searches may be useful.

### **Viewing Our Paper Online**

We have prepared an online version of this paper for viewing on a Macintosh or an IBM-compatible computer using **Netscape** and **Acrobat Reader**. Follow the procedure below to view our paper online.

1. Download **Netscape** from <http://www.netscape.com/>
2. Download **Acrobat Reader** from <http://www.adobe.com/acrobat>
3. Install **Netscape** and **Acrobat Reader**.
4. Launch **Netscape**.
5. Configure **Netscape** to recognize **Acrobat Reader** as a helper application (use Mime type: **application**; Subtype: **pdf**; Extensions: **pdf**; Action: Launch Application).
6. From the **Netscape** location window, type <http://www.writedoc.com/webdocs/webbooks.htm>

**Note:** You can download the **MacHTTP** server software from [http://www.biap.com/machttp/machttp\\_software.html](http://www.biap.com/machttp/machttp_software.html).