

When submitted via the Internet, very large documents or files can cause "time-out" problems for computers at either end of the transfer, resulting in a failed or canceled transfer.

Transmission times are dependent on the speed of the sender's communication device and the technology used by the Internet service provider. Very large document or files require lengthy transmission times during which the potential for error conditions or other service interruptions increases in direct proportion to the time the communication link must be maintained. The time-out problems could affect each party who receives the documents as part of the service of a filing. The actual transfer times for very large documents or files may approach 24 hours using standard Internet File Transfer Protocol (FTP) routines. In terms of ensuring timeliness, this may not be a significant improvement over the use of an overnight courier to send the files on optical storage media (e.g., CD-ROM).

• *Access to Large, Complex Documents in the Electronic Hearing Docket (EHD).*

Keeping a large document together in one very large file may allow users to easily search for, retrieve, and analyze the document in its entirety, but may result in service interruption problems similar to those described above. This is particularly true if a user wants to download the image file of one of these large documents. Retrieval time will be unacceptably slow, or will result in a time-out problem with the user's Internet connection.

Users of the EHD may encounter comparable download delays because of the file size of large or complex documents and, depending on the nature of the file, the file may not be executable on a user's desktop personal computer because of configuration, memory, display, or other technical problems.

• *Use of Large, Complex Documents in a Hearing Room.*

Large documents may be pre-filed as potential exhibits in the docket; however, in a hearing room, it is possible that only portions of such documents, i.e., specified chapters, pages, or paragraphs will be offered. In a dynamic and fast-paced hearing room environment, it would not be desirable to delay the proceeding to wait for a large file to load; navigate to the desired chapters, pages, or paragraphs; and then extract the appropriate selection for use in the proceeding. Complex documents may also require specialized hardware and/or software to execute software

program files and access their associated data.

• *Official Record and Federal Records Management Considerations.*

For both large and complex documents, the NRC must consider the need to generate an official record of the proceeding for use in potential appellate environments, see 10 CFR 2.1013(a), and for generating an Official Agency Record (OAR) version of the docketed materials for retirement to the National Archives and Records Administration (NARA). Each of these situations requires the ability to reassemble the record version of the documentary material (excluding software executables), independent of the media or software initially used to create it.

Coupled with the project objectives and technical requirements (discussed in the next section), these issues represent the framework for potential solutions. The NRC analysis distilled and assessed the objectives, technical requirements, and issues and developed four designs.

*Technical Requirements*

Given the anticipated size and complexity of individual documents, and the quantity of submittals, the need to transmit, manage, and retrieve electronic documents and objects challenges both the NRC's current processes and its information technology/information management (IT/IM) infrastructures, and the information technology (hardware, software, Internet service provider) in use by the general public. Examples of potential large documents are:

- The DOE Site Characterization Plan;
- The DOE License Application and supporting materials;
- The DOE Environmental Impact Statement;
- Adjudicatory documents (e.g., motions, responses, transcripts, exhibits, and orders).

Any or all of these types of documents may contain embedded photographs, charts, tables, and other graphics that contribute to the understanding of the narrative.

The NRC also anticipates receiving files that could be part of complex document submittals that:

- (1) Due to their file size, may preclude easy transmission, retrieval, and use; or
- (2) May require specialized software and/or hardware for faithful display and subsequent use; and
- (3) May not be suitable for inclusion in a "generic" file format such as PDF. The PDF standard, though it is proprietary to Adobe®, has been published and is available for use by software vendors. Users can access the

content of a PDF format file through the use of the Adobe Reader® viewer software.

Examples of files that could be part of complex documents include maps, databases, simulations, audio files, video files, and executable programs.

The analysis of the challenges of handling large documents in the NRC and public IT environments considered the following functional areas:

• *Transmit* activities entail sending a submittal from the submitter to the NRC, either via electronic format (through transmission or media) or as a physical object (e.g., video or audio).

• *Capture* relates to the receipt of electronic objects, with notifications provided according to an approved service list, preferably through e-mail. Upon receipt at the NRC, each submittal is staged for additional processing.

• *Index & Cross-Reference* are two distinct processes. Each submittal must be indexed based on prescribed profile templates. In addition, as part of the cataloging process, a submittal may be identified (or cross-referenced) as part of a package or compound document.

• *Store* manages the storage location of a submittal, i.e., within a folder or larger collection for electronic submittals, or the physical media location for submittals provided on optical storage media (e.g., CD-ROM) containing text, data, and objects. This process involves applying security and audit controls, as well as the appropriate retention schedule.

• *Search & Retrieve* operations involve querying the bibliographic header and content, displaying the pertinent object(s), and, if desired, printing all or part of the displayed object(s).

• *Create & Revise* activities facilitate the creation or revision of new documents using content that has been extracted (copied and pasted) from original submittals.

• *Copy & Distribute* activities involve maintaining distribution (service) lists and providing the means to copy or download an individual document or a collection of documents.

These activities may also involve reproduction when the need arises to generate a hard copy of a submittal (e.g., "8.5" x 11" paper", drawings, etc.).

Finally, there was an assessment of the existing NRC document and records management systems environment as well as requirements for enhancements to support the large document business requirements.

*Assessment and Alternatives*

The NRC assessed a number of alternatives to the existing technology