Enclosure 5

**From:** Dan Graser
**To:** Abby Johnson, Alan Kall, Andrew Remus, Bob Wells, Chris Berlien, Claudia Newberry, David Hunt, Debra Kolkman, Dennis Bechtel(...)
**Date:** Wed, Nov 24, 1999 10:44 AM
**Subject:** LSNARP Technical Working Group Meeting

Attached is a document entitled "Description of Potential LSN Design Alternatives". It is a narrative of LSN design alternatives (3 & 4) with additional information, including classes of hardware and software. We expect to also be forwarding some more detailed information on pricing of the two scenarios, and, a strawman revision of the Functional Requirements prior to the meeting.

As previously noted, the TWG will meet in Room 441 of the DOE offices at 1551 Hillshire, Las Vegas NV 89134 on Monday, December 6, 1999. We would like to begin work promptly at 8:30 am.

The .wpd is WordPerfect format
The .doc is Word format

**Mail Envelope Properties** (383C07D2.4A1 : 1 : 18974)

**Subject:** LSNARP Technical Working Group Meeting
**Creation Date:** Wed, Nov 24, 1999 10:44 AM
**From:** Dan Graser

**Created By:** DJG2.TWF2_PO.TWFN_DO

| Recipients | Action | Date & Time |
|---|---|---|
| internet | Transferred | 11/24 10:44 AM |
| "Bob.Wells@rw.doe.gov" (Bob Wells) | | |
| "Clark.Ray@EPA.GOV" (Ray Clark) | | |
| acj.carson-city.nv.us | Transferred | 11/24 10:44 AM |
| abby (Abby Johnson) | | |
| anv.net | Transferred | 11/24 10:44 AM |
| JudyTF (Judy Treichel) | | |
| aol.com | Transferred | 11/24 10:44 AM |
| hoylej (John Hoyle) | | |
| MalMurphy (Malachy Murphy) | | |
| nvtapper (Les Bradshaw) | | |
| Tiffanih (Nick Stellavato) | | |
| tuftam (Tammy Manzini) | | |
| caliente.igate.com | Transferred | 11/24 10:44 AM |
| jcciac (Eve Culverwell) | | |
| co.clark.nv.us | Transferred | 11/24 10:44 AM |
| dax (Dennis Bechtel) | | |
| evt (Englebrecht vonTiesenhausen) | | |
| cs.unlv.edu | Transferred | 11/24 10:44 AM |
| taghva (Kazem Taghva) | | |
| eurekanv.org | Transferred | 11/24 10:44 AM |
| lfiorenzi (Leonard Fiorenzi) | | |
| gfoster.com | Transferred | 11/24 10:44 AM |
| gfoster (Glen Foster) | | |
| govmail.state.nv.us | Transferred | 11/24 10:44 AM |
| madams (Marta Adams) | | |
| ssteve (Steve Frishman) | | |

| | | |
|---|---|---|
| idsely.com<br>wpnucwst (Debra Kolkman) | Transferred | 11/24 10:44 AM |
| isri.unlv.edu<br>tom (Tom Nartker) | Transferred | 11/24 10:44 AM |
| labat.com<br>joseph_speicher (Joseph Speicher)<br>tony_neville (Tony Neville) | Transferred | 11/24 10:44 AM |
| NCAI.org<br>Robert_Holden (Robert Holden) | Transferred | 11/24 10:44 AM |
| nei.org<br>rxm (Rod McCullum)<br>spk (Steven Kraft) | Transferred | 11/24 10:44 AM |
| Notes.YMP.gov<br>Claudia_Newbury (Claudia Newberry)<br>Jill_Schrecongost (Jill Schrecongost)<br>John_Gandi (John Gandi) | Transferred | 11/24 10:44 AM |
| phonewave.net<br>cccomp (Alan Kall) | Transferred | 11/24 10:44 AM |
| sierra.net<br>escorop (Tony Cain) | Transferred | 11/24 10:44 AM |
| telis.org<br>inyoplanning (Andrew Remus) | Transferred | 11/24 10:44 AM |
| chris<br>berlien (Chris Berlien) | Transferred | 11/24 10:44 AM |
| elaine<br>ezra (Elaine Ezra) | Transferred | 11/24 10:44 AM |
| threeputt.hawthorne.nv.us<br>wallace (Jackie Wallace) | Transferred | 11/24 10:44 AM |
| winston.com<br>strubatc (Sheldon Trubatch) | Transferred | 11/24 10:44 AM |
| ymp.gov | Transferred | 11/24 10:44 AM |

david_hunt (David Hunt)
harry_leake (Harry Leake)
john_gandi (John Gandi)
lew_robertson (Lew Robertson)
sam_hobbs (Sam Hobbs)

| | | |
|---|---|---|
| twf1_po.TWFN_DO | Delivered | 11/24 10:44 AM |
| gpb (Paul Bollwerk) | Opened | 11/24 1:37 PM |
| jgw (Jack Whetstine) | Opened | 11/29 8:20 AM |
| JXH5 (John Hoyle) | Opened | 11/25 3:42 PM |
| twf2_po.TWFN_DO | Delivered | 11/24 10:44 AM |
| hjs (Harvey Spiro) | Opened | 11/24 10:54 AM |

| Post Office | Delivered | Route |
|---|---|---|
| internet | | |
| acj.carson-city.nv.us | | internet |
| anv.net | | internet |
| aol.com | | internet |
| caliente.igate.com | | internet |
| co.clark.nv.us | | internet |
| cs.unlv.edu | | internet |
| eurekanv.org | | internet |
| gfoster.com | | internet |
| govmail.state.nv.us | | internet |
| idsely.com | | internet |
| isri.unlv.edu | | internet |
| labat.com | | internet |
| NCAI.org | | internet |
| nei.org | | internet |
| Notes.YMP.gov | | internet |
| phonewave.net | | internet |
| sierra.net | | internet |
| telis.org | | internet |
| chris | | |
| | internet:terraspectra.com | |
| elaine | | |
| | internet:terraspectra.com | |
| threeputt.hawthorne.nv.us | | internet |
| winston.com | | internet |
| ymp.gov | | internet |
| twf1_po.TWFN_DO | 11/24 10:44 AM | |
| twf2_po.TWFN_DO | 11/24 10:44 AM | |

| Files | Size | Date & Time |
|---|---|---|

| | | |
|---|---|---|
| Techscen.wpd | 55506 | Wednesday, November 24, 1999 8:26 AM |
| TECH_SCEN.doc | 87778 | Wednesday, November 24, 1999 10:40 AM |
| MESSAGE | 1284 | Wednesday, November 24, 1999 10:44 AM |

**Options**

| | |
|---|---|
| **Auto Delete:** | No |
| **Expiration Date:** | None |
| **Notify Recipients:** | No |
| **Priority:** | Standard |
| **Reply Requested:** | No |
| **Return Notification:** | None |
| **Concealed Subject:** | No |
| **Security:** | Standard |
| **To Be Delivered:** | Immediate |
| **Status Tracking:** | Delivered & Opened |

**Description of Potential LSN Design Alternatives**
November 23, 1999

## 1.0 Introduction

This document presents a description of the potential design of the Licensing Support Network (LSN) home site and participant sites. It is intended to provide the LSN Advisory Review Panel (LSNARP) Technical Working Group (TWG) with sufficient information to make a recommendation for the final design of the LSN to the full LSNARP.

This document builds upon the work done by the TWG during the week of 12 October 1999 in which the conceptual technical alternatives described in the LSN Administrator's (LSNA) Compliance Assessment Program (CAP) Guidelines (presented to the LSNARP on 13 October 1999) were evaluated. This work led to discarding all but two conceptual design alternatives that share many characteristics but differ in a fundamental way. This document will present those alternative designs, both the common elements and the differences, in additional detail in order to facilitate the determination of the optimal design. The focus will be on how the various aspects of the system are intended to operate, and the requirements placed on the various participants in the LSN to ensure this operation.

The LSN can be regarded as consisting of three functional components. Specifically, these are:

- A component that aids the LSNA in auditing participant compliance with the LSN Rule.
- A component that presents LSN information to participants, other interested parties, and the general public.
- A component that stores LSN documentary information for the use of components one and two.

The two design alternatives agreed upon by the TWG differ only in the conceptual design of the third component, specifically in how and where LSN materials are stored. The design of the first and second component will not be materially affected by the alternative selected for the third component. The detailed descriptions of components one and two apply to both alternatives currently under consideration. However, there will be differences in the details of implementation and operation.

The sections below will address each of the components in turn and the different alternatives for component three. A separate section will detail how components one and two interface differently with the alternatives for component three.

## 2. Compliance Component

This component was referred to as the "baseline" system in the LSNA CAP Guidelines document referenced above. It is a "front-end" component (one with which end-users interact) with a small set of users who require specific information at specific times. It is intended to address the in-house needs of the LSNA.

### 2.1 Intended functionality

The purpose of this component is two-fold. First, this component ensures that the LSN is functioning as intended and provides assuredness of this functioning to the intended user base. Second, it provides the necessary reports on LSN functionality to enable the LSNA to ascertain that participants are in compliance with the LSN Rule and to aid in determining whether remedial action is required.

The primary method of following the operation and evolution of the LSN is through a reporting mechanism. Reports will be generated automatically by the system on a periodic basis, both when exceptional conditions arise, and on-demand.

#### 2.1.1 Periodic reports

The full array of required reports is yet to be determined. However, the following have been identified at this time:

- A listing of changes in participant document collections, i.e. additions, deletions, modifications.

- A report on the "health" of the LSN, component and sub-component uptime and performance data (e.g. web server hits, average response times, number of users, etc.)

#### 2.1.2 Exception reports

Exception reports will be used should anomalous conditions arise. Candidates for this type of report include:

- When auditing software detects a possible compliance problem in a participant collection.

- When a component of the LSN itself is determined to be malfunctioning, i.e. due to a computer or network error.

- When a security exception is noted.

### 2.1.3 On-demand reports

It is anticipated that reports may need to be generated from time-to-time to respond to an exception or to "drill down" to garner additional data on a perceived compliance problem. A facility will be provided to expedite this process. It is anticipated that HTML forms will be designed to allow individuals to design and generate most reports on demand. However, it is likely that some reports may need to be developed by systems maintenance personnel from time-to-time.

## 2.2 Intended user base

The LSNA and his designees and the ASLBP (Atomic Safety and Licensing Board Panel).

## 2.3 Access to functionality/information delivery

### 2.3.1 Web browser

Certain functions of the system are best accessed through a WWW browser (e.g. Netscape, Internet Explorer, etc.) through the standard HTTP/HTML mechanism bolstered by CGI programs that interact with the data stores. Most commercial and open source network management software currently employ a web-based interface. Specifically, those aspects of the monitoring function that change rapidly can best be monitored through a browser. Examples of these are troubleshooting on-going problems and ascertaining the status of a particular sub-component at a particular time. Historical trends will be maintained both in HTML tables and graphically.

A web browser is also anticipated as the usual interface to generate reports on demand, with an HTML forms interface providing the report and data selection, as well as the formatting function.

### 2.3.2 Hard-copy delivery

It is anticipated that certain reports, especially periodic reports "for the record," will be automatically printed and physically delivered to their intended recipient.

### 2.3.3 E-mail delivery

E-mail is an alternative method of delivery most geared to exception reports but useful for all report types.

### 2.3.4 Interactive login

Interactive access to the system will be required to produce on-demand reports that have not been anticipated in the design of the web-based, on-demand facility described in Section 2.4.1.

### 2.3.5 File system access

File system access is required for ready availability of system logs and other source data for off-line processing and archival.

### 2.3.6 Pager notification

Certain types of exception reports, e.g. notifications of system unavailability, mandate a more aggressive notification. In these instances, the system administrator will be paged with a description of the exception in order to expedite repair.

## 2.4 Component elements and their functionality.

### 2.4.1 Data retrieval element

This element will consist of one or more programs which will routinely "rove" participant sites, fetching participant data (documents, statistics, and other) and storing this data pending processing. The exact nature of the data retrieval element will depend on the details of the alternative selected for the storage component, but it is analogous to a "web spider." A web spider, when presented with a starting URL, will traverse all hyperlinks within the body of documents "under" the URL. Through this methodology, it is possible to retrieve and replicate the entire static structure of a web site for further processing.

### 2.4.2 Data storage

This element is responsible for storing both data to be processed and the results of that processing. Both file system storage and database storage will be accommodated. The database will be a network-capable SQL relational database that will provide structured data to both front ends.

#### 2.4.2.1 Data processing

This element will process the data retrieved, store the results of the processing, and generate the required reports.

#### 2.4.2.2 Data presentation and reporting tool

This element consists of several programs that process report outputs into formats appropriate for the delivery mechanisms described above, and assist a user in specification of on-demand reports.

#### 2.4.2.3 System assuredness with further sub-elements

This element provides a level of assuredness that the systems the LSN is housed on are functioning as required. There are several main sub-elements:

- Security mechanisms. Security sub-elements include a firewall or firewall software, secure remote administration software, and intrusion detection software.

- Network monitoring and management. This sub-element monitors hardware and software and reports outages or sub-optimal operation. It also gathers low-level statistics on network operation for trend and throughput analysis.

- Physical plant and reliability mechanisms. This sub-element provides appropriate environmental and power conditioning and implements disaster recovery, e.g. a backup capability.

### 2.5 Hardware and software required

No attempt is made to specify make and model of hardware and software at this time. Where appropriate, examples of products will be provided, but these are not intended to represent a comprehensive list of alternatives or preferred selections.

#### 2.5.1 Computer system hardware

A single computer system of the workstation class is adequate for this functionality. The security sub-element mandates that the system be separate from and more restricted than the computer system (described below) that provides general access. The system should be equipped with the standard components, a graphical display, and a device appropriate for backup. Examples of this type of system include an i386-architecture workstation (e.g. Pentium III "PC") running open-source Unix (e.g. FreeBSD or Linux), or Microsoft NT, a Sun workstation running Solaris, or a Compaq/DEC Alpha running VMS. The primary selection criteria for specific hardware and operating system should be based on security objectives, with specific functionality a secondary (but important) consideration.

#### 2.5.2 Computer system software

The following software components will be required: a web server (e.g. Apache, Netscape Enterprise, MS IIS), a database with accompanying report generation software (e.g. PostgreSQL, Oracle, MS SQL Server), firewall software (e.g. IPFW, ipfilter, Firewall-1), network monitoring and management software (e.g. Big Brother, SunNet Manager, HP OpenView), and a web spider (e.g. MoMspider, BRS/Search, Fulcrum Search Server). Note that the web server, database, and web spider are also part of the presentation component. The same software can be used for both purposes. In addition, it is anticipated that this component will require some custom software, scripts and CGI's rather than full-

blown applications.

### 3.0 Presentation Component

This component is a "front-end" with a large set of users who require access to a wide range of information at arbitrary times. It is intended to fulfill the requirement to provide information to interested parties through WWW technology.

### 3.1 Intended functionality

It will be a WWW presentation interface with additional sub-components that consist of:

- Introductory and overview documentation.
- Training / tutorial materials on how to use the site to obtain LSN -related information, and the other aspects of the site, and how to submit to the docket.
- Portal software that allows user customization of user interfaces and user document search and access strategies.
- A search facility that allows LSN-wide searching of participant materials, including per-user custom searching strategies.
- Publication of statistical information on LSN participant sites, including site content and performance.
- Aggregation and publication of overall LSN access and usage statistics, e.g., number of hits.
- A web-based interactive forum in which interested parties can discuss or exchange information regarding LSN matters.
- Help-desk assistance for participants and public users with escalation.
- A LISTSERV (e-mail list manager) to allow participants to easily send electronic mail to all interested parties. A number of mailing lists will be created as needed for discussion of specific subjects, including a list with the e-mail addresses of all participants for notification purposes. The LISTSERV software will allow each participant to manage their own subscriptions to interest lists and archive messages to the lists. It is not intended to provide a public LISTSERV function.

### 3.2 Intended user base

The intended user base includes all participants and potential participants, the LSNA and his designees, the press, and the general public.

### 3.3 Access to functionality/information delivery

#### 3.3.1 Web browser

Web browsers will be the predominant access method to this component. It is anticipated that this will be the sole access method for the majority of users. Browsers will be used to gain access to general information, participant documentary collections, and to discussion forums.

#### 3.3.2 E-mail

E-mail will be used for notification to participants by the LSNA or designee, and interaction with the LISTSERV described above.

### 3.4 Hardware and software required

No attempt is made to specify make and model of hardware and software at this time. Where appropriate, examples of products will be provided but these are not intended to represent a comprehensive list of alternatives or preferred selections.

#### 3.4.1 Computer system hardware

A single computer system of the server class will be required for this functionality. Examples are as in the previous section but this component will require more processing power and capacity, i.e. a faster CPU or multi-CPU machine, more RAM, bigger disk storage, etc. The primary selection criteria for the hardware is that it should be supported by the portal software selected (the most critical software component).

#### 3.4.2 Computer system software

The following software components will be required: a web server (e.g. Apache, Netscape Enterprise, MS IIS), a database with accompanying report generation software (e.g. PostgreSQL, Oracle, MS SQL Server), firewall software (e.g. IPFW, ipfilter, Firewall-1), a web forum (e.g. UltimateBulletinBoard, WWWboard), and a LISTSERV (e.g. MailMan, majordomo, LISTPROC), and portal software (e.g. Plumtree, Excalibur, Knowledge Center). Note that the web server, database, and web spider are also part of the compliance component. The same software can be used for both purposes.

### 3.5 Participant activities and responsibilities

None except as end-users. This component is the responsibility of the LSNA.

### 4.0 Storage Component

The storage component represents the "back-end" functionality serving the needs of the front-end components rather than the end-users directly. The data it contains consists of the documents required to be published by participants in accordance with the LSN Rule and accompanying required information. Two alternatives have been identified for providing this functionality.

### 4.1 Storage alternative 1 - dispersed document storage

This alternative proposes that each participant, assemble, prepare, and publish their own collections of documents on a WWW server. Components one and two will access these collections as WWW clients and perform the necessary operations routinely through participant sites.

#### 4.1.1 Intended functionality

This component is the "back-end" that will provide data to the front-end components described above. Participants will make their documentary collections available on a web server located at the site of their choosing and attached to the Internet. Participants are free to establish their own web server, collaborate on a community web server, procure commercial web service, or employ any other provisioning method.

#### 4.1.2 Intended user base

The intended user base is primarily the LSN front-end software described above.

It is anticipated that participants may choose to make their document collections (and ancillary information) generally accessible on the WWW (i.e., other than through the LSN portal site). However, any documents intended to be filed in the licensing process will have to be obtained or cross-referenced through the LSN portal site to ensure the uniqueness, consistency, and traceability of document identification (accession) numbers..

#### 4.1.3 Access to functionality/information delivery

##### 4.1.3.1 Web access

This will be the primary method by which participant materials are accessed. Access will be interactive (e.g. when a home site front-end user requests a particular document the home site front end will fetch it from the participant's repository) and by batch (e.g. the portal will fetch all materials on the web site, index them, and retain only the references to the documents for subsequent presentation in response to end-user queries).

#### 4.1.3.2 SNMP access

For obtaining network usage statistics and performing monitoring activities, the compliance component will also require SNMP (Simple Network Management Protocol) access to participant web servers and network interface equipment.

### 4.1.4 Participant activities and responsibilities

Participants are required to make available all documents subject to discovery in standard, LSNA-approved formats on a web site. This consists of the following procedures.

#### 4.1.4.1 Document identification and assembly

This is simply identifying and assembling the documents. This function will provide a reasonably accurate estimate of the storage space and preparation effort required.

#### 4.1.4.2 Document preparation

Documents are to be converted to a format that includes an image representation (TIFF/CCITT or TIFF/JPEG), a searchable text file, and a bibliographic header containing metadata about the document. In many cases, this will require scanning and OCR conversion of a paper document. However, if a document exists in electronic format, it may be preferred to perform a more accurate conversion with appropriate software.

The LSNA may allow participants to provide their documentary collections in alternative page-representation formats such as PDF and proprietary word processor formats like Microsoft Word. This will depend on whether the data retrieval software selected for the front-end components is capable of indexing, searching, and otherwise processing these formats. The requirement to provide a bibliographic header for each document will remain regardless of the documents' formats. The bibliographic header is subject to the same retrieval requirements as the source document, e.g. provided as a searchable text file by the web server, as HTTP headers, or from within a database.

Document preparation is potentially the most labor-intensive and costly aspect of building the LSN due to the large number of documents included. Therefore, the burden on a participant is more closely correlated to the number of documents they must prepare than any other factor.

#### 4.1.4.3 Document publication

In this alternative, participants will place their documents on the web server of their choice through whatever file transfer mechanism is supported by the web server. This web server

must be connected full-time to the Internet through a communications circuit of adequate speed (to be determined by the LSNA) and have a unique IP address and domain name. The domain name and root URL for the documentary collection, and a list of documents, must be provided to the LSNA.

For consistency in retrieval by the front-ends described above, participants may be required to follow a standard format in layout of the web pages that provide access to the documents themselves and accompanying bibliographic header information. Note that many web servers provide a standard way to publish meta information on web-served documents (e.g. by including this information in a file of the same name as the source document in a meta sub-directory). Use of this function may be required by the data retrieval elements of the front-end components.

### 4.1.5 Hardware and software required

It is difficult to determine the exact hardware and software components due to the possibility of collaboration and the differences in the size of the documentary collections of the participants. Foreseeable alternatives for setting up a web server include a dedicated resource at the participant's site, sharing a server with other participants or non-LSN-related web sites, "co-location" of a participant-owned machine at an IPP (Internet Presence Provider) or outsourcing the entire site to an IPP. Each of these alternatives have a wide range of cost, convenience, assuredness, and administrative issues associated with them.

If a strategy of providing a dedicated web server is adopted, the size of this machine will, again, depend on the size of the document collection the participant is required to make available.

Participants with an extremely small document collection will probably choose to lease web space on an IPP machine or "piggy-back" on another participant's site rather than implement their own web server. The cost of this facility depends on the amount of data published, the bandwidth the site requires, and other metrics. Typical costs for web sites that are appropriate for small participants range from free (of incremental cost over maintaining a basic Internet-access capability) to several hundreds of dollars per month.

For those who choose to operate and maintain their own dedicated resource, a fairly modest machine may be fully satisfactory. An example of this would be an i386 architecture "PC" (e.g. 166MHz Pentium, 128MB RAM, 4GB disk) running an open-source Unix-like operating system (FreeBSD or Linux) and the open-source Apache web server. The total cost (hardware and software) of such a machine at current (4Q99) market prices is under $1,000, and it would accommodate as many as 10,000 documents (at an estimated 250KB per document). Note that operational costs may not be so trivial, especially the disaster recovery aspects (regular backups with off-site storage), and data communications costs. However, resources for these requirements may already exist and participants who choose to share a web server may be able to equitably spread these

costs among themselves.

Participants with larger document collections will, naturally, require a more capacious computer system up to or including the one described in storage alternative 2. Note that operational costs will scale as well.

### 4.2 Storage alternative 2 - aggregated document storage

This alternative proposes that all or a subset of participants' documentary collections be stored at a facility "topologically close" to the facilities that provide the front-end functionality. In networking terms, this equates to a LAN connection between the front-end presentation tools and the back-end data storage in a "LSN campus" of computers.

Specifically, the original proposers of this alternative expressed the desire to co-locate the DOE's documentary collection with the NRC's. However, since this represents over 80% of the total documents within the entire LSN, there is no reason not to logically extend this concept to the total documentary set of all participants.

From one perspective, this alternative can be regarded as comparable to the original Licensing Support System (LSS) in where and how the documentary collection is stored. However, it differs greatly in several ways:

- how the collection is accessed by participants (as described above),
- how the funding allocations are administered,
- how to isolate the function (or malfunction) of one participant's components from another participant's,
- how to assign the responsibility for facility administration and support (in an environment where contending parties share a resource), and
- the procedure established for ensuring participant responsibility for the accuracy and completeness of its documentary collection in the event of an error on the part of its third-party database administrator or the mis-configuration or mis-operation of a shared resource.

#### 4.2.1 Intended functionality

The functionality provided is the same as in storage alternative 1, only the method by which it is accomplished is different. There will be differences in operational characteristics, however, and these are discussed in Section 4.3.

#### 4.2.2 Intended user base

Same as storage alternative 1. Note that all accesses to LSN materials are necessarily through the portal site, independent access is not realistically possible.

#### 4.2.3 Access methods

##### 4.2.3.1 Web access

Web access will be used for some accesses, although the primary access to the data store will probably not be web-based for efficiency reasons.

##### 4.2.3.2 SNMP access

SNMP access to all computer systems used within the "LSN campus" will be required for monitoring and management purposes.

##### 4.2.3.3 File system access

This is anticipated to be the primary access method to LSN campus-stored document collections. File system access (local disk or network disk) has certain efficiencies over web access.

#### 4.2.4 Participant activities and responsibilities

##### 4.2.4.1 Document assembly

Same as in storage alternative 1.

##### 4.2.4.2 Document preparation

Each participant will still have to prepare the documents they are responsible for providing as described above. However, the different access methods may alter the allowable data representation formats and the method of provision of the bibliographic headers (e.g. by "live" access to an SQL database rather than as a text file accompanying the source document).

##### 4.2.4.3 Document publication

To preserve the ability of the LSNA to perform independent review and make impartial decisions, a way will have to be found to ensure that the LSNA does not assume ultimate responsibility for ensuring that participant collections are available, this responsibility ultimately rests with each participant regarding their own collections. Participants will provide a resource or agree to share a resource who will place documents into their area on the shared storage facility.

Allowable methods of conveyance from a participant's preparatory facilities to the shared storage are not determined at this time. Options may include magnetic media (tape or disk), optical media (CD-ROM or DVD), or bulk transfer over a network. Note that all allowed options must be associated with a mechanism that provides assuredness that the document actually entered into the designated storage area is a certified copy of the participant's original document.

It is reasonable to expect that the LSNA will designate separate storage areas for each participant's collection (rather than commingling collections). These separate storage areas may actually be separate machines provided and maintained on the campus LAN by the participants themselves.

### 4.2.5 Hardware and software schedule

A storage component that houses, potentially, the entire LSN pre-discovery documentary collection is quite massive (between 200GB and 4TB by various estimates). Such size will require a very robust computer system or multiple computer systems of the server or enterprise server class.

Software included with the operating system will provide file system access over the campus LAN to the front-end components. This functionality is generally included in both open source and commercial server operating systems (or is readily available) to support the most common network file system sharing schemes (NFS - Network File System, SMB -Server Message Block, NCP - NetWare Core Protocol, etc.)

Since the database required by the front-end components has a large data-storage requirement, and the standard method of access is over the network, performance issues may dictate that the software implementing the database server functionality run on the storage server instead of the machine implementing the front-end. However, this is the same software component described above, not another database implementation.

## 4.3 Comparison between Alternative 1 and Alternative 2

The key differences between the two alternatives for the storage component are in the areas of:

- component integration,
- participant procedures for document preparation and publication, and
- funding of implementation.

### 4.3.1 Component integration

In alternative 1, the various front-ends and back-ends are loosely-coupled (a distributed system). In alternative 2 they are, from the point-of-view of an external observer, effectively a single system (the LSN campus). The usual trade-offs for such alternatives apply.

#### 4.3.1.1 Alternative 1 - Distributed system

In the case of distributed systems:

- Independent operation of various sub-components allows partial functionality in the case of system and or network outages. For example, if the portal/presentation component becomes unavailable, access to participant documentary collections is still available directly through participant web sites. In contrast, a single or tightly-coupled system will most likely be completely unavailable if a single sub-component fails.

- The system can be more easily reconfigured or extended without disruption to the system as a whole. Since the design emphasis is on the interfaces between distributed stand-alone systems, another stand-alone system (with the correct interface) can be "plugged-in" as an additional component. For example, midway through LSN implementation, a participant may find that the initial server selected for implementation cannot handle its entire documentary collection. In the distributed alternative they can obtain another server rather than start again from scratch with a larger machine.

- Additional functionality for participant sites is possible. Participants may choose to publish information on their LSN web site that is not part of the LSN (in that it does not relate to the high-level nuclear waste repository licensing procedure). In alternative 2, the LSNA will likely choose to host only LSN-related materials.

- Individual participants, or collaborations of participants, assume all responsibility for publication of their documents and have a well-defined point at which this publication can be assessed for compliance, i.e. their web site interface. In alternative 2, it may be difficult to ascertain which party is responsible for perceived non-compliance. Participants will have to place a large measure of trust in the entity selected to manage and maintain the LSN campus.

- In the current computer marketplace, indications are that the aggregate cost of the distributed system would be lower. Specifically, this is due to the availability of commodity computers at extremely low prices. The larger, data-center machines required by alternative 2 generally cost significantly more than the equivalent computing power in commodity machines.

- Spreading the resources available for implementation over multiple sites usually shortens the roll-out period. The "many hands make light work" principle applies (even though there may actually be more work in total). In the case of a single large

system, the large burden of implementation placed on a single staff can result in implementation delays.

#### 4.3.1.2 Alternative 2 - the LSN campus

In the case of an LSN campus:

- Aggregate performance will likely improve, and will certainly be more predictable. In contrast, the performance of distributed systems is generally much more variable.

- The tightly-coupled nature of the aggregated facilities will allow efficiencies in communications to occur. For example, it will allow high-speed LAN connections between components rather than slower long distance telecommunications circuits. This has implications in custom software development and fine-tuning performance in off-the-shelf software.

- The reduced number of sub-components will increase average reliability (at the cost of increasing the consequences of a system-wide outage) and the security of the system will be more easily assured.

- Technical expertise required by participants is lessened (but greatly increased at the LSN campus site). Participants will not have to be webmasters or acquire webmaster services. In a distributed scenario, web site maintenance capability will have to be in-house or acquired.

- The system, as a whole, will be easier to manage and maintain, and will require fewer total resources for this function, than distributed, stand-alone systems. Additionally, the responsibility for management and maintenance will be more clearly defined. However, the level of maintenance and management of the LSN campus system, and the expertise required to accomplish it, will increase in direct proportion to its size.

### 4.3.2 Participant procedures for document preparation and publication

The variation in the publication of documents has been covered above. There is likely to be more freedom for allowable data formats in the distributed scenario (alternative 1) as web servers handle many of the issues regarding data representation as a natural aspect of their design. However, without more knowledge of the data formats desired to be used by various participants, it is difficult to estimate the impact of this issue.

The most fundamental issue regarding document preparation and publication in alternative 2 is the notion of transfer of responsibility from the participant to the entity charged with management of the LSN campus. A mechanism must be put into place to provide assuredness that documents are transmitted from one party to another without corruption

and to enforce non-repudiation. It may be possible to achieve this with appropriate procedures and certifications, although mutual trust would still be an important element. In the case where participants retain responsibility for their own collections, there is a clear demarcation of responsibility at each participant's web site.

### 4.3.3 Funding of implementation

It is not clear how a shared storage component will be funded. In the case of the original design concept (represented by alternative 1), each participant is responsible for funding publication of their own collection. Aggregating some or all of the collections may make appropriate allocation of costs difficult. The usual issues in any compensation situation arise, e.g., what happens when a participant doesn't pay what it owes or when a participant loses standing but still owes a contribution? The risk of unanticipated expenditures is shifted from one participant for their own requirements to all participants and, primarily, the NRC as LSN campus host.