May 19, 2010

UNITED STATES OF AMERICA
NUCLEAR REGULATORY COMMISSION

BEFORE THE ATOMIC SAFETY AND LICENSING BOARD

In the Matter of	)	
	)	
U.S. DEPARTMENT OF ENERGY	)	Docket No. 63-001-HLW
	)	
(High-Level Waste Repository)	)	ASLBP No. 09-892-HLW-CAB04
	)	
	)	

LICENSING SUPPORT NETWORK ADMINISTRATOR'S
RESPONSE TO THE BOARD'S QUESTIONS REGARDING THE LSN

In its April 21, 2010 Order,[1] Construction Authorization Board 04 (CAB04) directed the Licensing Support Network Administrator (LSNA) to respond to questions relating to the costs to terminate and preserve the Licensing Support Network (LSN) portal system (developed and operated by the LSNA).[2] CAB04 listed these questions in Appendix D of that order, and requested that the LSNA submit answers to these questions to CAB04 by May 24, 2010.[3] The following states the LSNA's answers to the questions of CAB04 regarding the LSN:

Q1. How much has the LSNA invested in the development of its LSN portal system and the establishment of its LSN computer environment (including all O and M and litigation support contractor costs, hardware acquisitions, software development, and federal staff FTEs converted to dollars)?

A1. The investment life-cycle cost to date for LSN is $16.1 million.[4]

---

[1] CAB Order (Questions for Several Parties and LSNA) (April 21, 2010) at 2 (unpublished).

[2] Id. at app. D.

[3] Id. at 3.

[4] These cost figures do not include the estimated cost to implement and maintain the information technology (IT) for the NRC staff's collection on the LSN. The investment life cycle cost to date for the NRC document collection is $6.6 million.

Q2. How much has the LSNA invested (including all O and M contractor costs, hardware acquisitions, software development, and federal staff FTEs converted to dollars) in the development of any ancillary systems that are not accounted for in the numbers provided in response to question 1?

A2.

Ancillary systems used to fulfill requirements outlined in 10 C.F.R. § 2.1013 include the agency's E-Filing system, the Electronic Hearing Docket (EHD), the Digital Data Management System (DDMS), and some standards incorporated into the functionality of the agency's centralized Document Processing Center (DPC). Because the E-Filing, EHD, and DPC investments are not within the purview of the LSNA, I am unable to provide current life-cycle cost figures for those systems. Accurate information about the total investment costs for those systems should be available from the NRC Office of the Secretary and/or the Office of Information Services.

The DDMS is the agency system that uniquely fulfills the "next-day availability" and the "where in the transcript each was marked for identification and where it was received into evidence or rejected" requirements of 10 C.F.R § 2.1013(b). It also uniquely fulfills the "while testifying, for use during the hearing" requirement of 10 C.F.R § 2.1013(d). DDMS is sponsored by the Atomic Safety and Licensing Board Panel and is managed under the auspices of the LSNA.

The investment life-cycle cost to date for DDMS is $14.2 million which includes design, development, IT security certification, and one round of technology refresh activities for both Headquarters and field hearings.

Q3. How much has the LSNA invested (including all O and M support contractor costs, hardware acquisitions, software development, and federal staff FTEs converted to dollars) in the development of any underlying databases not accounted for in the numbers provided in response to questions 1 and 2?

A3.

The LSNA and LSN staff expended labor hours performing technical coordination with agency staff offices for the development of E-Filing, EHD, and DPC operations, and the

development of the associated agency guidance on system usage, including multiple versions of *Guidance for Submission of Electronic Docket Materials under 10 CFR Part 2, Subpart J*[5] and its underlying technical specifications. The approximate cost to date of the LSNA effort associated with that coordination is $247,500.

Costs incurred by the LSNA on the DDMS effort are reflected in the cost figure for DDMS as noted in response to Q2, above.

Q4. What have been the yearly costs (including federal staff FTEs converted to dollars) to operate and maintain the LSN portal over the past 5 years?

A4.

Expenditures by Fiscal Year are:

Fiscal Year (FY)	Amount in Millions
06	$1.65
07	$1.72
08	$1.22
09	$1.70
10	$1.46

Q5. What is the estimated budget and timing needed to archive the LSNA's LSN portal system?

A5.

The estimated budget is $250,000 for contract support, plus $185,000 in additional federal staff effort, for a total of $435,000. The closeout timeline extends 5 months from receipt of Commission direction to terminate system operation. This presumes a total suspension of all "normal" work once the closeout effort begins.

Q6. How will any additional funds be obtained to perform this function?

A6.

Unless supplemental appropriations are provided, the only funding mechanism available would be to reallocate existing fiscal year/contract year funding to closeout tasks. This

[5] http://www.nrc.gov/waste/hlw-disposal/guidanc.pdf

presumes that direction to shut down is received early enough in a fiscal year/contract year to ensure that the residual balance is adequate to cover the projected costs of the shut down. In a simplified and purely hypothetical case, if on August 1 of a given fiscal year the contract balance and unexpended current year funding is $120,000 and it costs $250,000 to accomplish close out, we would need access to an additional $130,000 to complete the close-out process that would not be forthcoming until the next fiscal year.

Q7. Explain to what degree the system could be revived in the future, the steps needed to accomplish this, and the estimated costs to perform this 1, 2, 5, or 10 years after archiving.

A7.

Once the system has been shut down, system restoration will be problematic in any of the referenced timeframes, just as it will be with the overall DOE and NRC High Level Waste programmatic efforts. I believe that any mandate to resurrect the HLW program after it has been shut down will present significant organizational challenges to all of the participants, to varying degrees, and have wide-ranging and long-term management impacts especially notable in the federal arena.[6] Specifically, the LSN is a synergistic system insofar as the document collections rely on the LSN-provided search engine for search and retrieval functionality and those collections would be unmanageable for search and retrieval purposes without the LSN. At the same time, the LSN index that provides the search and retrieval functionality is of limited value without the data contained in the participant document collections.

For these components to interact effectively, both the LSN indexes and the underlying participant document collections must be maintained in strict synchronization because the LSN

[6] Impacts include significant management challenges: organization rebuilding, support contractor procurement efforts, recapturing the staff and consultant knowledge/skill bases needed to support evidentiary activities, and extra-ordinary fiscal planning and budget execution environments.

indexing software (spider)[7] conducts a scan of the directory structure of the participant document repository using the header and document root directories created by the participant. All of the directory and file entries are stored in the LSN database as Uniform Resource Locators (URLs) and thus form the lynchpin for connectivity between the participant documents and the LSN indexes.

Synchronization, in turn, relies on maintaining the integrity of both the file location and the file content and includes:

- Participant adherence to the technical standards in 10 C.F.R. § 2.1011 and the LSNA guidelines (in particular Guideline GL-22, Setting Up a Repository); and
- Configuration control of both the LSN indexing methodology and the structure of the participant collection servers.

When a participant item is added to the system during the spidering process, the system retrieves the combined header/document file information from the LSN database, checks the date-time stamp and file size, and then downloads copies of each of the files from the repository fetch server. Hash values[8] are calculated for each of the downloaded files, which permit the system thereafter to make a hash value comparison that is capable of detecting a one (1)-bit change within each file. As part of its normal operation, the system validates all three items (date-time stamp, file size, and hash value) against previous data values and, because the LSN software is extremely sensitive to any change in a document, the system would flag even the replacement of a period with a comma in an image or text file.

Additionally, when items are placed on a document collection server by a participant, each file (e.g., each bibliographic header, each text file, and each image file) resides at a

---

[7] A spider is a program that visits Web sites and reads their pages and other information in order to create entries for a search engine index.

[8] Hash algorithms are one-way mathematical algorithms that take an arbitrary length input and produce a fixed length output string. A hash value is a unique and extremely compact numerical representation of a piece of data. It is computationally improbable to find two distinct inputs that hash to the same value (or “collide").
http://www.cgisecurity.com/owasp/html/ch13s05.html

location on that server that is identified as a Uniform Resource Locator (URL) entry used by the LSN to locate the item. If the URL changes or is corrupted in either the LSN index or on the participant server, the linkage is broken and the item becomes "lost" and irretrievable to the index.

Thus, loss of either data integrity or URL synchronization could result in the participant collection at issue having to be totally re-spidered and reindexed. By way of example, this happened with both the NRC and DOE document collections early in the process of constructing the LSN. The LSN's indexing capacity is 30,000 documents per day/5 days per week, but can been considerably less (by an order of magnitude) if a participant server has bandwidth limitations and/or security filters in place. Based on past experience, I estimate that if the DOE collection had to be reconstructed, that collection alone would take over 6 months of uninterrupted indexing using all the spidering resources available. Moreover, for the system to be restored to its current configuration, all directory structures, data files, and LSN indexes must be able to be recovered flawlessly for every participant collection. To be recovered flawlessly, each of those collections must be archived flawlessly, and then properly maintained until recovered. If problems are encountered with archiving either the LSN portal or the DOE document collection (which is of particular concern because of its size) and it becomes necessary to reindex, 6 months will be needed just to rebuild/reconstruct the LSN indexes.

Relative to system management, it should be noted that once a government IT contract is terminated, in very broad terms, the following events occur:

- A contract modification is issued directing the contractor to commence closeout activities.
- The database is "archived" (e.g., written to storage media such as tape, CD/DVD)
- Project assets including hardware, software, and documentation are inventoried.
- The database is uninstalled and the equipment is dismantled.
- Equipment is dispositioned or surplused.

- Archived data and documentation are delivered to the government.
- The contract is terminated.
- Unexpended funds are deobligated.
- Project information (files, documentation, etc.) is retired according to the disposition schedule approved for the activity.

Once termination takes place, the planning horizon for resurrecting a system is driven by budgeting, contracting, and project development factors.

As a practical matter, the federal government budget formulation process has a 2-year lead time unless other strategies, such as mid-year requests or discretionary funding, are available within the agency. Even in the best-case scenario and assuming such funds already exist within the agency, it could be several months before discretionary funds are made available. If the normal budget cycle is used as a baseline, a more realistic estimate for reaching the funding level appropriate to reimplement the LSN is between 1 and 2 years after planning direction is received.

Once a contract has lapsed and a new contract vehicle is determined to be needed, the lead time for getting a new contract in place can be in excess of a year unless the head of an agency approves a sole source or public exigency exception. Because the LSN contracts have historically been performance-based, the statements of work would be reusable with minimal rework. Nonetheless, even a relatively straightforward procurement off of a General Services Administration (GSA) Schedule can, as was the case with the last O&M contract competition for the LSN, take a considerable period. Accordingly, procurement activities including all new hardware, new software licenses, and new operations and maintenance should be anticipated to have an 18-month lead time.

In the technology arena, the likelihood of returning to the Autonomy™ software used in the current incarnation of the LSN is moderately good.[9] The prospect of being able to "name-brand specify" or otherwise mandate Autonomy™ in a full and open competition, however, is poor.

Should Autonomy™ not be the search and retrieval product – either because it is no longer available, has evolved in a significantly different direction during the period of LSN decommissioning, or is not the product proposed by the successful bidder – there are two significant consequences. First, implementation of a new software product would be the equivalent of designing a new system, which will require an additional year for a rigorous system development and engineering effort.[10] Second, the reindexing timeframe noted in the discussion should be increased even further in consideration of the time needed for all the other non-DOE collections.[11]

Additionally, it cannot be assumed that any of NRC's current LSN staff will be available once the LSN is shut down. The effort to rebuild an IT organization competent to resurrect the LSN could take 2 years, as it would involve budgeting for full-time equivalent (FTE) allocations, recruitment, recovery of program information, familiarization, and training for the certifications necessary to operate a major IT system within the agency.

---

[9] The Autonomy Corporation received a "Strong Positive" rating in Gartner's recent report "MarketScope for E-Discovery Software Product Vendors." Strong Positive is the highest possible rating in the Gartner MarketScope. The MarketScope evaluated E-Discovery vendors based on their overall viability, customer experience, market understanding, sales execution / pricing, offering (product) strategy, product/service, and sales strategy. http://www.autonomy.com/content/News/Releases/2010/0111.en.html

[10] This is comparable to the duration of the initial LSN design and implementation effort.

[11] Different products may have substantially different indexing performance so that the overall reindexing time estimates could vary widely.

The “long poles” in any effort to reconstruction the system would be assembling a program team, the availability of funding, and the procurement process that, all running roughly in parallel, would account for 2 to 4 years. If Autonomy™ is not the search engine product selected, another year is necessary for system redesign and 3 to 4 months may be needed to achieve an IT security Authority to Operate (ATO). Finally, if a redesigned system is necessary, all document collections would, most likely, be reindexed in their entirety, which may require an additional 9 months of activity. Accordingly, if the LSN is not resurrected before the current contract is terminated and funds deobligated, any subsequent system restoration will require at least 3 years and 9 months after a qualified team has been assembled. With minor, non-substantial redesign, this may be shortened by about 9 months.

Technological risk associated with attempts to resurrect the LSN includes: loss of back-up data if project organizations are disbanded and adequate resources are not dedicated to LSN custodianship, the potential for DOE not being able to resurrect their data, and, future hardware and software platform incompatibility (in a migration scenario viz iterating as many as 5 years’ worth of update releases and release patches should the existing Windows 2008 and Autonomy v.7 be reused). To summarize, technical reutilization is realistic in a 1-year window, feasible in a 2-year window, “iffy” in a 5-year window, and highly unlikely in a 10 year window.

Costs associated with resurrecting the LSN include the time and effort to recruit and hire new project staff and the cost of those FTE resources during that timeframe, estimated to be up to 2 years, at a cost of $300,000. The procurement/contract award effort will include the ASLBP and Division of Contracts FTE resources over an estimated 2 year timeframe, at a cost of $600,000. The contractor effort for a 1-year effort to develop and place a new LSN into production, including the direct contractor effort, Independent Verification and Validation, and IT Certification and Accreditation are estimated to be $5 million. The total effort including staff FTE resources is $5.9 million.

In summary, 5 years is a reasonable "ballpark" estimate for the lead time necessary to resurrect the LSN, effective the day after the existing LSNA program is disbanded.[12] In that timeframe: the portal components become, essentially, a "rebuild-from-scratch" effort; the reliance on data/document integrity and reutilization will be a moderate risk; and, the cost will be substantial.

Q8. In regards to any LSNA's custom software required to run the LSN portal system, elaborate on its degree of obsolesce if the system was to lie dormant for 1, 2, 5, or 10 years.

A8.

Technological change[13] is probably going to be significant and unavoidable. That being said, almost all technological challenges can be met by the aggressive application of financial resources by highly skilled IT professionals.

The attributes of technology change include: 1) change occurs in key technologies more frequently and in ever-shortening cycles (e.g., computing "horsepower" doubles every 18 months; storage costs are halved every 12 months; internet bandwidth capacity doubles every 9 months); 2) change occurs exponentially rather than linearly; and 3) significant technology step changes (i.e., a change associated with mutually compatible multiple technologies emerging at the same time) are now routinely occurring in under 2 years. These factors will influence items such as what future server platforms (hardware) and operating systems (software) look like in the 5- and 10-year scenarios.

As noted above, the prospects for the LSN search and retrieval software being essentially the same product over the next 2 or 3 years are good. Less predictable is the impact of the collateral change occurring around the periphery of that particular product.

---

[12] This estimate does not account for the time necessary for DOE to reconstitute the Office of Civilian Radioactive Waste Management and reimplement its LSN document collection.

[13] The risk of technical obsolescence is much smaller than the risk of a full and open competition process precluding the selection of the same product families unless a sole-source justification is approved by the agency leadership.

Ultimately, however, technology obsolescence is not the main issue or the challenge that should be addressed. Instead, the central concern is maintaining the migration path for data so that it continues to be able to adapt to the state of the technology at those 5- or 10-year intervals.

The most important action that can be taken to ensure information viability is to have a well-defined and well-managed plan to maintain data in formats that adhere to current-at-that-time data standards regardless of whether that data is online or offline. Compliance with standards is key to ensuring data migration paths to new technology.

Q9. What effect would the shutdown of DOE's LSN facilities in Nevada have on the retention and future archiving of the LSN portal system?

A9.

There are no NRC-controlled LSN facilities in Nevada. The shutdown of the NRC's Las Vegas Hearing Facility (LVHF), should that occur, would have no direct impact on the LSN. The shutdown of DOE's, Nevada's, and the AULGs' LSN-related facilities in Nevada is best addressed by those participant organizations.

Q10. Do any of the preservation and archival activities for the LSN portal system have to be approved by NARA or any other federal agency?

A10.

The disposition of all federal government information must be in accord with the NARA-approved records retention and disposition category for that information. In addition, NARA provides technical specifications for the transfer and storage media, such as 9-track tapes (and requires additional consultation and approval) if information is going to be designated as some sort of "special collection" such as has been done, as I understand it, with scientific and technical data for the National Aeronautics and Space Administration (NASA).

As of this writing, NRC's records management staff is still in consultation with NARA regarding the records disposition schedules for the LSN portal and the LSN Administrator's record materials.

Respectfully submitted,

/RA/

_____

Daniel J. Graser
Licensing Support Network Administrator
U.S. Nuclear Regulatory Commission
Mail Stop T3-F23
Washington, DC 20555-0001
(301) 415-7401
Daniel.Graser@nrc.gov

Rockville, Maryland
May 19, 2010