Computing Sciences

August 8, 2010

Securities and Exchange Commission
100 F Street, NE
Washington, DC 20549-1090

Re: Comments on SEC File No. S7-11-10: Consolidated Audit Trail
From: David Leinweber[1] and Horst Simon[2]
Computational Research Division
Lawrence Berkeley National Laboratory

We welcome the opportunity to comment on the proposed rule for a Consolidated Audit Trail.

Lawrence Berkeley National Laboratory (LBNL) is not a familiar name in financial circles, so some introduction is appropriate. Berkeley Lab, founded in 1931, is one of the leading scientific research institutions in the United States and the world. The Laboratory employs over 3900 people, including 1320 scientists and engineers. The FY2010 budget exceeds $770 million. [3]

Eleven scientists associated with Berkeley Lab have won the Nobel Prize. One of the most recent is former Laboratory Director, Steven Chu, now Secretary of Energy. Thirteen Lab scientists have won the National Medal of Science, our nation's highest award for lifetime achievement in scientific research.

In the Laboratory's early days, its principal focus was on high energy and nuclear physics. This has expanded to include research in chemistry, biology, climate, energy, and, most importantly in the context of this comment, high performance large-scale computing ("supercomputing"). The Laboratory's Computational Research Division is a global center of excellence in research across all the disciplines mentioned, and a national talent and technology resource capable of addressing the substantial research issues posed by the Commission's interest in a Consolidated Audit Trail. Berkeley Lab is singularly capable of bringing the world's most powerful computing machinery to bear on the complex and dynamic problem of understanding the ever faster data streams and interactions in the fast, fragmented US market.

---

[1]David Leinweber, Financial technologist, and author of "Nerds on Wall Street: Math, Machines and Wired Markets", Wiley 2009 http://tinyurl.com/nerdsonwallst. Lawrence Berkeley Labs Center for Innovative Financial Technology http://www.lbl.gov/CS/CIFT.html

[2] Horst Simon, Director of Computational Research at Lawrence Berkeley Laboratory, http://www.nersc.gov/homes/simon

[3] A retrospective of the Laboratory and its accomplishments was written for the 75th anniversary in 2006. http://www.lbl.gov/Publications/75th/ Twenty five scientific breakthroughs at Berkeley Lab that improved the world, see http://www.lbl.gov/Publications/breakthroughs/index.html .

Ernest Orlando Lawrence Berkeley National Laboratory | One Cyclotron Road
MS 50B-4230 | Berkeley, California 94720 | Tel: 510.486.7377 | Fax: 510.486.4300

The Lab's supercomputing expertise goes more than forty years. In the 1960s, the computational challenge for the Lab was to analyze data from high-energy physics experiments. In 1974, a supercomputer at Berkeley Lab was first used for computational chemistry. The NERSC (National Energy Research Scientific Computing) Center moved to Berkeley in 1995. Originally established solely to support magnetic fusion research, NERSC helped pioneer many of the computing practices taken for granted today. For the last four decades Berkeley Labs and its sister DOE (Department of Energy) labs provided the impetus for the continued development of "supercomputers", machines many thousands of times more powerful than those used in ordinary settings. Many of the largest and most powerful computers in the world are found at the Laboratory and its affiliated institutions.

Applications for these machines have grown to include complex problems in many areas of physics, biology, chemistry, energy and information technology. Berkeley Lab's capabilities in large-scale data management, analysis, visualization, and simulation are unsurpassed. Synchronization of multiple disparate sources, and events that happen very quickly, well below the millisecond level, are recurring themes in the Laboratory's research. These issues are central to an effective Consolidated Audit Trail system.

Close collaboration with industry to meet the needs of the most demanding computational tasks has been a hallmark of the Laboratory's computational research. This practice can bring the same unrivaled research and development skills to bear in the challenge of monitoring and understanding increasingly complex, ever-faster interconnected markets. This could be the foundation for a strong research component in the efforts of the SROs regarding the Consolidated Audit Trail.

With due respect to some of the other responders, we do not believe that there is a quick off-the-shelf solution for the Consolidated Audit Trail. In high performance computing, large problems are characterized by their scale, in terms of data sizes and computational demands. The march of technology has meant that yesterday's large problems are solved on tomorrow's desktops. "Mega" problems, involving megabytes of data were overwhelming to early computers with Kilobytes of storage. Many "giga" scale problems are today accessible to desktop machines. Coping with "tera" scale data is a challenge in conventional settings. The Supercomputing research community is now routinely capable of addressing "peta" and "exa" scale tasks.

We believe that the ability to manage and understand the full scope of market data, at the detailed order flow level, across complex interacting systems, in a historical context, and in real-time or near real-time is clearly a peta-scale problem. The high frequency information storage facilities of commercial financial data firms are already at this scale.
An important role for the Consolidated Audit Trail systems should be to support the Commission when there are "what if" questions regarding proposed rules and market structures. These are questions that call for extensive simulations across thousands of securities and many market scenarios. This moves the CAT into exa-scale computing, and suggests that the research to support it requires exa-scale capabilities and expertise.

The "flash crash" of May 6, 2010 has yet to be satisfactorily explained. This suggests a "grand challenge" task for those approaching the CAT system: to be able to recreate the events of May 6 and other reported smaller anomalies, in a simulated environment.

The simulation and analysis tools associated with the CAT cannot ignore possibilities of malicious actions, attacks, manipulations and other illegal activity. But investor confidentiality must also be

protected. Research into exactly this sort of confidential security analysis is a recent accomplishment of the CRD.
An incremental, research and prototype driven approach to the CAT is well advised. We are fully cognizant of the issues raised in the June 28, 2010 Office of Management and Budget directive on financial information systems[4].

> *Federal Information Technology (IT) projects too often cost more than they should, take longer than necessary to deploy, and deliver solutions that do not meet our business needs. Although these problems exist across our IT portfolio, financial systems modernization projects in particular have consistently underperformed in terms of cost, schedule, and performance.*

There are too many stories of lavishly expensive federal computer system disasters. Focused research is the way to avoid the unfortunate and widespread consequences of the current approach to federal systems development. OMB identifies two basic factors causing these results.

> *First, the large-scale modernization efforts undertaken by Federal agencies are leading to complex project management requirements that are difficult to manage. Second, by the time the resulting lengthy projects are finished, they are technologically obsolete and/or no longer meet agency needs.*

In a domain that changes as rapidly as electronic markets the second issue could leave systems behind in a perpetual race.

LBNL is a new participant in the analysis of financial data, but experienced in meeting large-scale, extreme-scale computing challenges. We fully subscribe to OMB's principles for information projects: *Split projects into smaller, simpler segments with clear deliverables, Focus on most critical business needs first. Ongoing, transparent project oversight.*

Collaboration with SROs and other market participants is essential. However, the absence of the vested interest of established stake-holders in this regard is valuable. We were struck by the accounts of numerous past problems with having such stake-holders do their own auditing detailed in the first portion of SEC's document. The Lab's independence is a valuable asset.

A valuable historical example of how federal labs can meet major information technology challenges is the seen in the SAGE (Semi-automatic Air Ground Environment)[5], DEW (Distant Early Warning)[6] and BMEWS (Ballistic Missile Early Warning System)[7] supercomputer based radar systems. The trio of computer warning systems was developed in the 1960s to monitor the huge stream of data from distant early warning missile and aircraft defense radars.

This was a new and nearly overwhelming technical challenge of national importance, an expensive

---

[4] OMB Memorandum M-10-26, "Immediate Review of Financial Systems IT Projects", http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m-10-26.pdf/
[5] The Computer Museum's history of the SAGE (Strategic Air Ground Environment) project is found here: http://www.computermuseum.li/Testpage/IBM-SAGE-computer.htm. The visualization component of SAGE and related systems is the famous "Big Board" at NORAD under Cheyenne Mountain
[6] DEW http://en.wikipedia.org/wiki/Distant_Early_Warning_Line,
[7] BMEWS http://en.wikipedia.org/wiki/Ballistic_Missile_Early_Warning_System

undertaking with a budget for SAGE alone estimated at over $10 billion in 1964 dollars. The Core research, development, and management functions were done at federal labs (MIT Lincoln & RAND) in collaboration with industry (IBM, Bell Labs and Burroughs). The companies MITRE & SDC were spun off the labs and after the successful R&D, with ongoing responsibilities for integration, operation and modernization.

Faith in quality of our markets is of national importance, a national security issue in the broad sense. The commission clearly recognizes this in its proposed Rule. The many technical questions raised underscore the research-intensive nature of the proposed Consolidated Audit Trail.

LBNL and other federally sponsored computing research centers represent a substantial national resource that can play a key role in the auditing and understanding increasingly complex and data-intensive financial markets. Additional materials relevant to these topics are found at the Laboratory's Center for Innovative Financial Technology website.[8] We thank the commission for the opportunity to comment on these issues.

Respectfully,

Horst Simon, Ph.D.
Associate Laboratory Director for Computing Sciences
And Division Director, Computational Research Department

and

David Leinweber, Ph.D.
Director, LBNL Center for Innovative Financial Technology
Computing Sciences

[8] LBL CIFT http://www.lbl.gov/CS/CIFT.html/