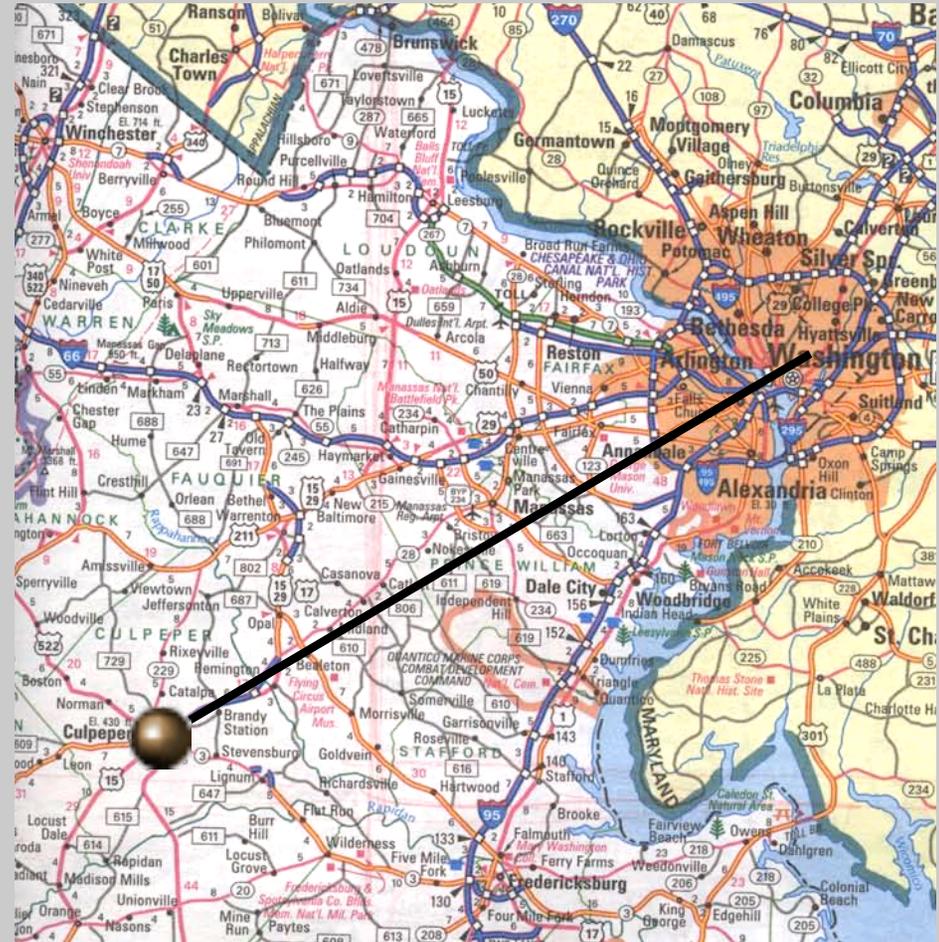
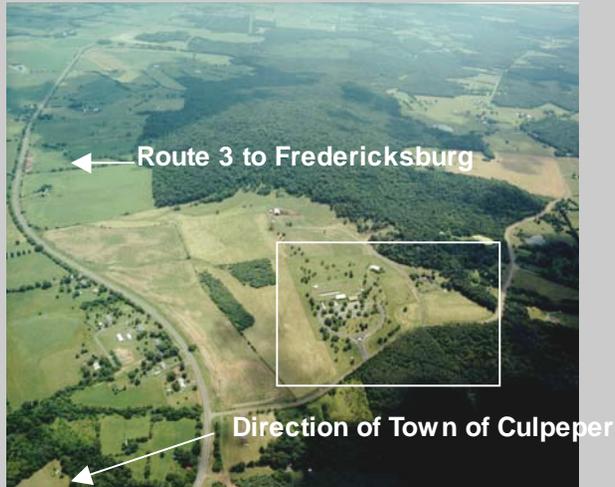


**“Planning for Digital Preservation and Acquisitions
at the Library of Congress
National Audio-Visual Conservation Center”**



**Gregory Lukow, Chief
Motion Picture, Broadcasting and Recorded Sound Division
May 1, 2007**

NAVCC Geographic Orientation



Site Location – Culpeper County, Virginia – 75 miles SW of Washington

NAVCC Site Overview & Components



Central Plant / AOC
(50,000 sq ft)

**Nitrate
Vaults**
(55,000 sq ft)

Collections Storage
(135,000 sq ft)

Loading dock

Main Entrance

**Conservation Bldg -
Staff & Preservation Labs**
(175,000 sq ft)

JUN 30 2006

**Before and after comparison:
former Federal Reserve Building,
now Collections Storage Building**



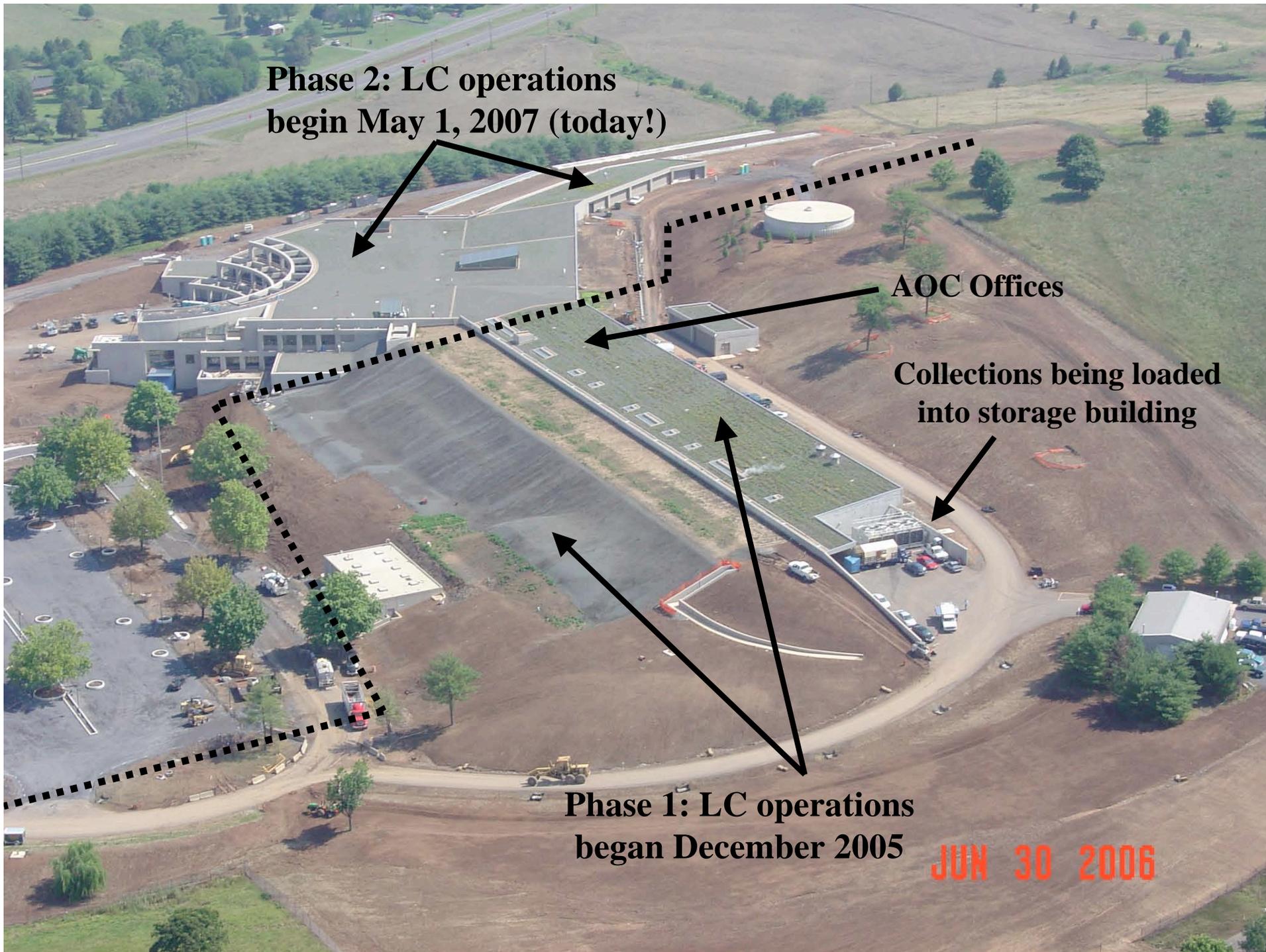
**Phase 2: LC operations
begin May 1, 2007 (today!)**

AOC Offices

**Collections being loaded
into storage building**

**Phase 1: LC operations
began December 2005**

JUN 30 2006





NAVCC Mission

“The National Audiovisual Conservation Center develops, preserves and provides broad access to a comprehensive and valued collection of the world’s audiovisual heritage for the benefit of Congress and the nation’s citizens.”

The Collections

- Over 1.1 million moving image collection items: theatrical films and newsreels, television programs, educational, industrial and advertising material.
- Nearly 3 million audio collection items: commercial sound recordings, radio broadcasts, and early voice recordings of historical figures.
- Over 2.1 million supporting documents, screenplays, manuscripts, photographs, press kits, etc.

Funding – Private and Congressional

- Unique partnership between the U.S. Congress, the Packard Humanities Institute (PHI), the Library of Congress, and the Architect of the Capitol.
- PHI is providing \$155 million for design and construction costs and will donate the finished facility to the government.
- Congress authorized the purchase of the property by PHI and is providing \$80 million in start-up funding for NAVCC operations, staff and equipment.
- The Architect of the Capitol is providing ongoing operations and maintenance support for the center.
- When completed in 2007, PHI's gift to the nation will be the largest private donation ever made to the Library of Congress.

Staffing – Initial Levels

- 138 NAVCC staff for year one (2007-08), including 11 reading room staff remaining on Capitol Hill
 - 35 Moving Image Processing, Curatorial and Reference staff
 - 35 Recorded Sound Processing, Curatorial and Reference staff
 - 42 A/V Preservation Services staff in the Film, Video and Sound Laboratories
 - 26 Administrative, Technical and Outreach Partnerships staff

Design Vision

- Designed for the Future...envisioned as a facility still capable of performing legacy work as if no other such services were available 25 years from now
- Designed as an Acquisitions, Preservation and Access “factory”
- Designed to achieve the transition from Analog to Digital preservation for audiovisual materials
- Designed for Capability...envision a national center to provide increasingly “obsolete” services becoming less and less available elsewhere

Design Vision (cont.)

- Designed for Capacity...preserve more and faster, or risk losing material, especially magnetic media
- Designed for Scale and Scalable preservation... up to 5 Petabytes of output annually in initial years (primarily video & sound); significantly more in the future
- Designed for Efficiency of preservation...modes of production from “boutique” to “robotics”
- Designed to maximize integration and automation of new systems and workflows

Software Automation & Integration

- New software is being written to tie the system together throughout the Culpeper facility.
- NAVCC will be an integrated and automated campus designed from top to bottom to optimize preservation production efficiencies for all media formats.
- The systems that run NAVCC are a Web 2.0 based infrastructure integrated with a professional Business Process Management system (BPM).
- Work will be scheduled on a centralized system and tracked throughout the facility.
- Reports will be generated and costs tracked within the system to show ongoing production output.

Software Automation & Integration (cont.)

- Each staff member will have a common, integrated online interface to NAVCC and other existing Library catalogs and management systems.
- The system will provide researchers in the reading rooms with a robust search engine that can call up digitized content derivatives for immediate playback on demand.
- The system has a new generation of access systems built in that are seen on wide area networks such as the National Lambda Rail and Internet2.
- These access systems will provide appropriate access and use of the NAVCC files to universities and other external institutions as relationships and partnerships are identified.

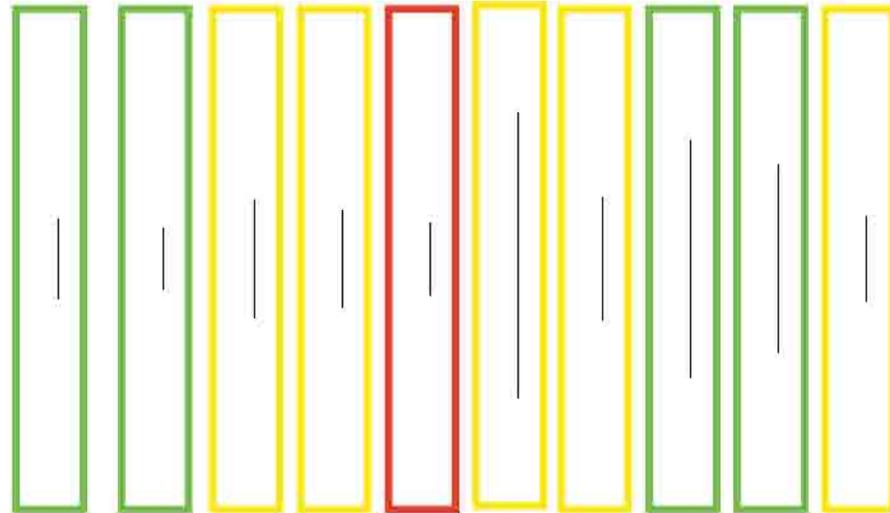
What we wanted to do at NAVCC (Requirements Document v2.4)

Workflow	
Copyright Acquisitions	Accessioning
Cataloging	Film Preservation Reformatting
Fulfillment	Video Preservation Reformatting
Reference Request	Audio Preservation Reformatting
Born Digital Processing	Paper Collections
Non-copyright Acquisitions	
Services	
Copyright Siebel System	LCCN Assignment
LoC Name Authority File	MUZE Interface
Online Computer Library Center	MAVIS Interface
Audio and Video Servers	Labeling Tool
Proprietary Database Access	Cuadrastar
Library Catalog System Integration	Object Builder
Research Libraries Information Network	
Software Systems	
Workflow Management	Collections Management
Digital Conversion PODs	Lab Management
Work Order	Raw Deposit
Digital Deposit	Business Management
Customer Relations Management	Archive Backup
Hardware	
Digital Ingest Storage	POD Storage
Cache Storage	Archive Storage
Backup Storage	Internet Connectivity
Business Network	Production Network
Backup Network	Servers
AV Equipment	External Access Network

How we are doing it

User Interface:
CSS2, XHTML, Velocity Templates, Struts and Tomcat embedded in JBOSS.

SAVVION and JBOSS:
Oracle 10, Hibernate Persistence layer and query service, EJBs and Java classes. Development done with Eclipse IDE.



CSS2 for integrating look and feel with HTML - <http://www.w3.org/TR/REC-CSS2/>
 XHTML – For HTML extensibility and content integration - <http://www.w3.org/TR/xhtml1/>
 Velocity Templates for Java HTML integration – <http://jakarta.apache.org/velocity/>
 Struts – Overall Java integration tool - <http://struts.apache.org/>
 Tomcat – Servlet and JSP deployment for JBOSS application code - <http://tomcat.apache.org/>
 JBOSS – Application deployment environment with EJBs and Java – <http://www.jboss.com>
 Hibernate – Object/Relational persistence and development for Java - <http://www.hibernate.org/>
 Eclipse – Programming framework – <http://www.eclipse.org>

Collaboration with LoC ITS

- Library Services/MBRS responsible for requirements development, design and build of “front-end” audiovisual production systems
 - Design by Ascent Media
 - Integration and installation by Communications Engineering Inc.
 - Software by the Gustman Group
- ITS responsible for requirements, design and build of “back-end” digital storage archive systems
 - Integration by Government Micro Resources Inc. (GMRI)
 - Uses Sun server and StorageTek robotic storage solutions
 - 3 ITS staff will work on-site to maintain digital archive
 - Mirrored off-site back-up at remote federal computer facility

Long-Term Digital Preservation

- Conducted **comprehensive analysis of the long term preservation needs** in the development of the RFP for the digital archive system. Analysis included **consultation with recognized experts** in the digital preservation and high-performance computing sectors.
- Solution incorporates **backup and recovery capabilities** of the Legislative Branch Alternate Computing Facility, a remote and secure disaster recovery facility. A **real time link ensures** an additional copy of every archived file is made simultaneously with the original.
- Solution based on the **concept of continual migration and verification**. Migration to **progressively higher density storage will continue indefinitely** into the future.
- Storage media will be permanently mounted so they can be **automatically checked on a periodic schedule**, rather than stored on a shelf where material might go unchecked for a long period.

Preservation Output

After initial ramp-up to new operating levels, the NAVCC projects the following estimated preservation output by the end of the first full year. Production levels will increase as new programs are added in later years.

- Recorded sound to WAV digital files: 25,000 hours from all formats (disc, reel-to-reel, cassette, etc.)
- Video to JPEG2000 digital files: 48,000 hours reformatted from cassette and reel-to-reel (includes robotic systems output)
- Film to digital file: 450 hours at 2K from 16mm film and films scanned from the “Paper Print” collection
- Film-to-film (analog): equivalent of 200 feature films per year
- 2 Petabytes total estimated digital content archived in first full year (July 2007-June 2008; will increase in later years)

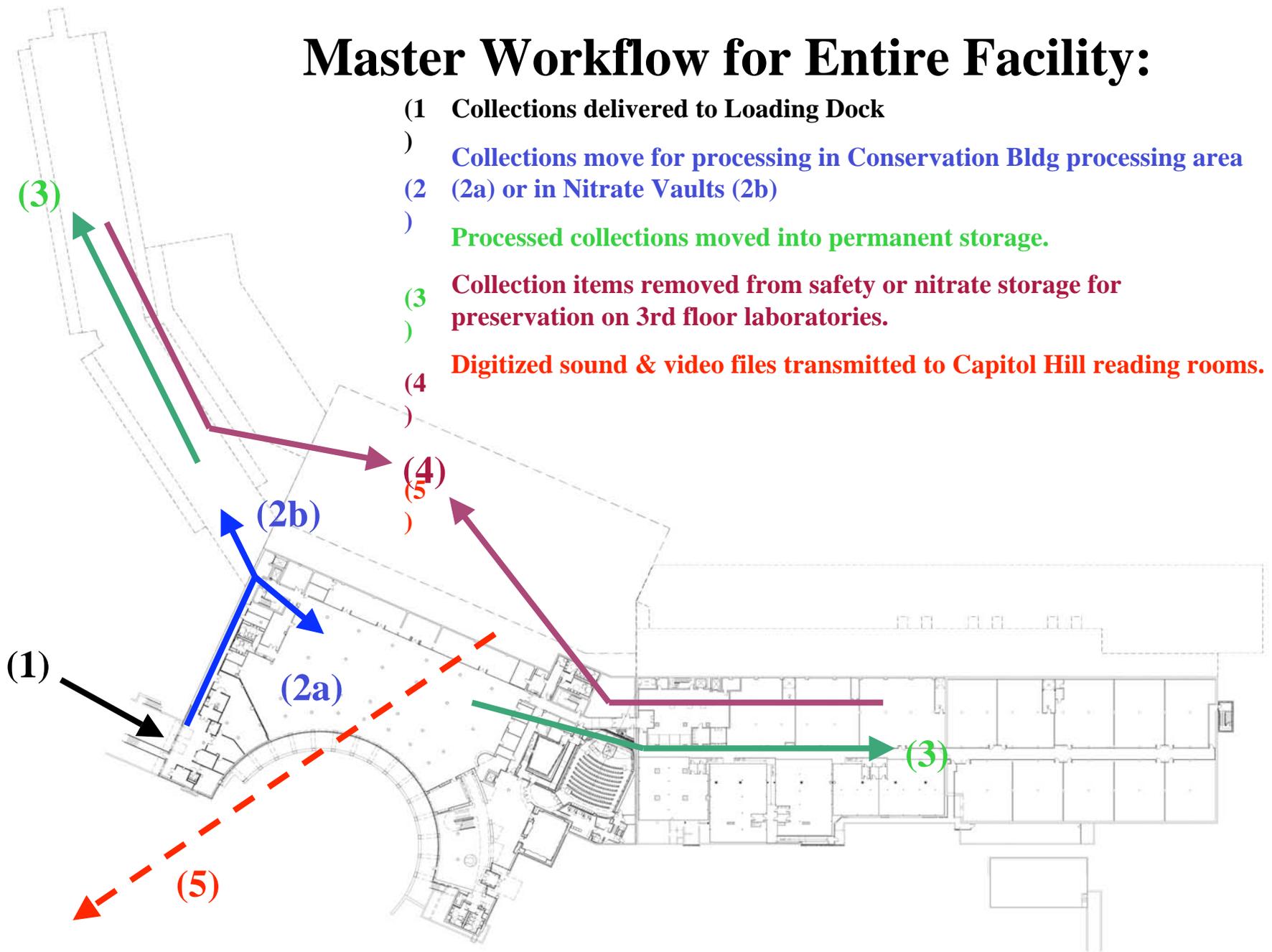
Access to Audiovisual Works

- Will retain the Motion Picture, Television, and Recorded Sound reading rooms in the Madison Building so patrons can utilize other Library collections.
- Sound and video collections access will be via digital files served to patrons on Capitol Hill via a derivative server and fiber optic connection from Culpeper.
- Film viewing prints driven daily to Capitol Hill for patron access in reading rooms.
- Will broaden public accessibility and extend the Library's outreach to the public, partners and customers through innovative access models.

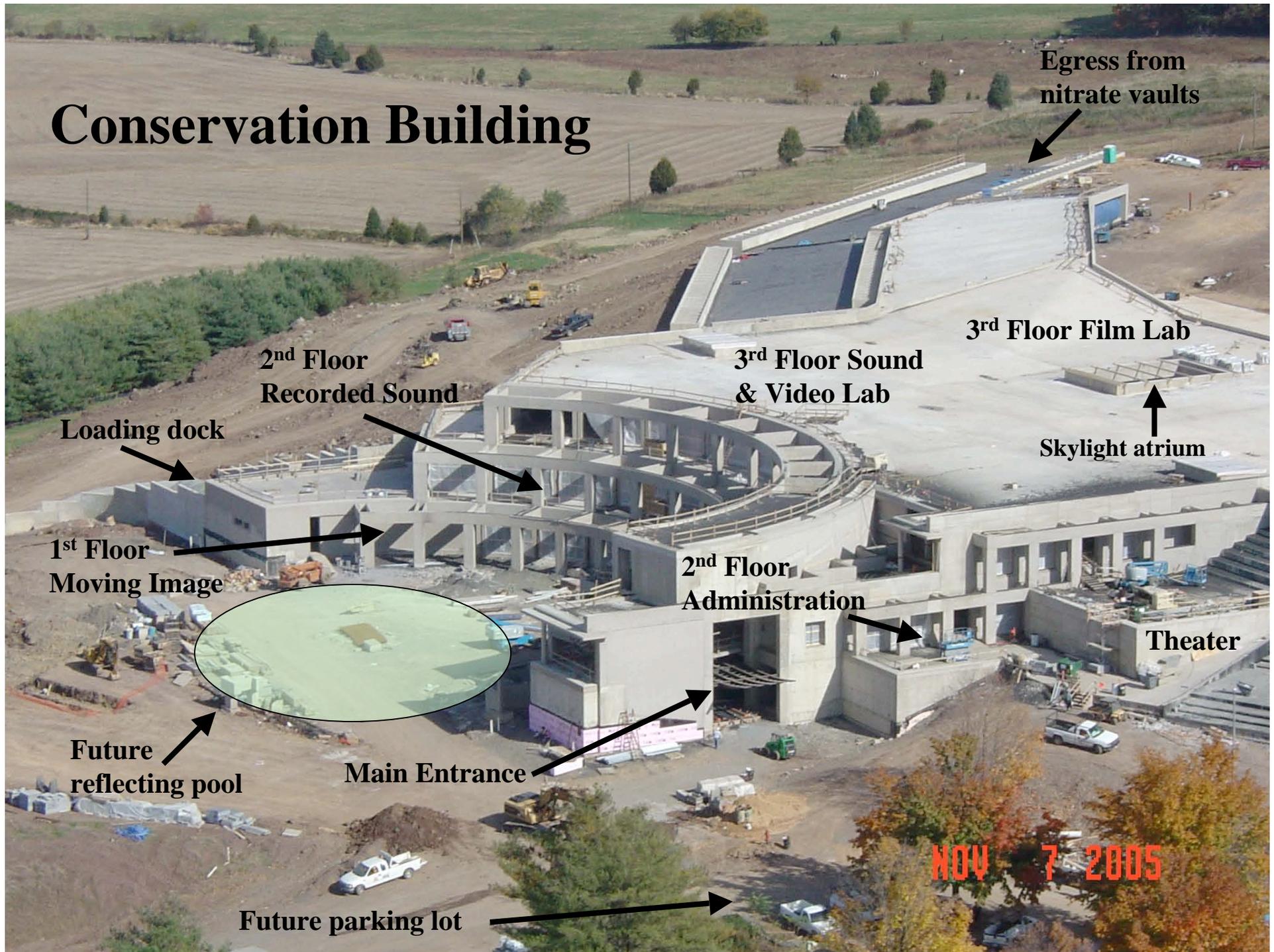
Service Innovations for the Library and Archival Community

- Provide preservation duplication and storage services to other libraries, archives and industry constituents.
- Provide a test bed for the development of large-scale mass digital archiving systems for sound and video materials.
- Share innovations, standards and tools developed at the NAVCC with the broader archival community.

Master Workflow for Entire Facility:



Conservation Building



Egress from nitrate vaults

3rd Floor Film Lab

3rd Floor Sound & Video Lab

Skylight atrium

2nd Floor Recorded Sound

2nd Floor Administration

Theater

Loading dock

1st Floor Moving Image

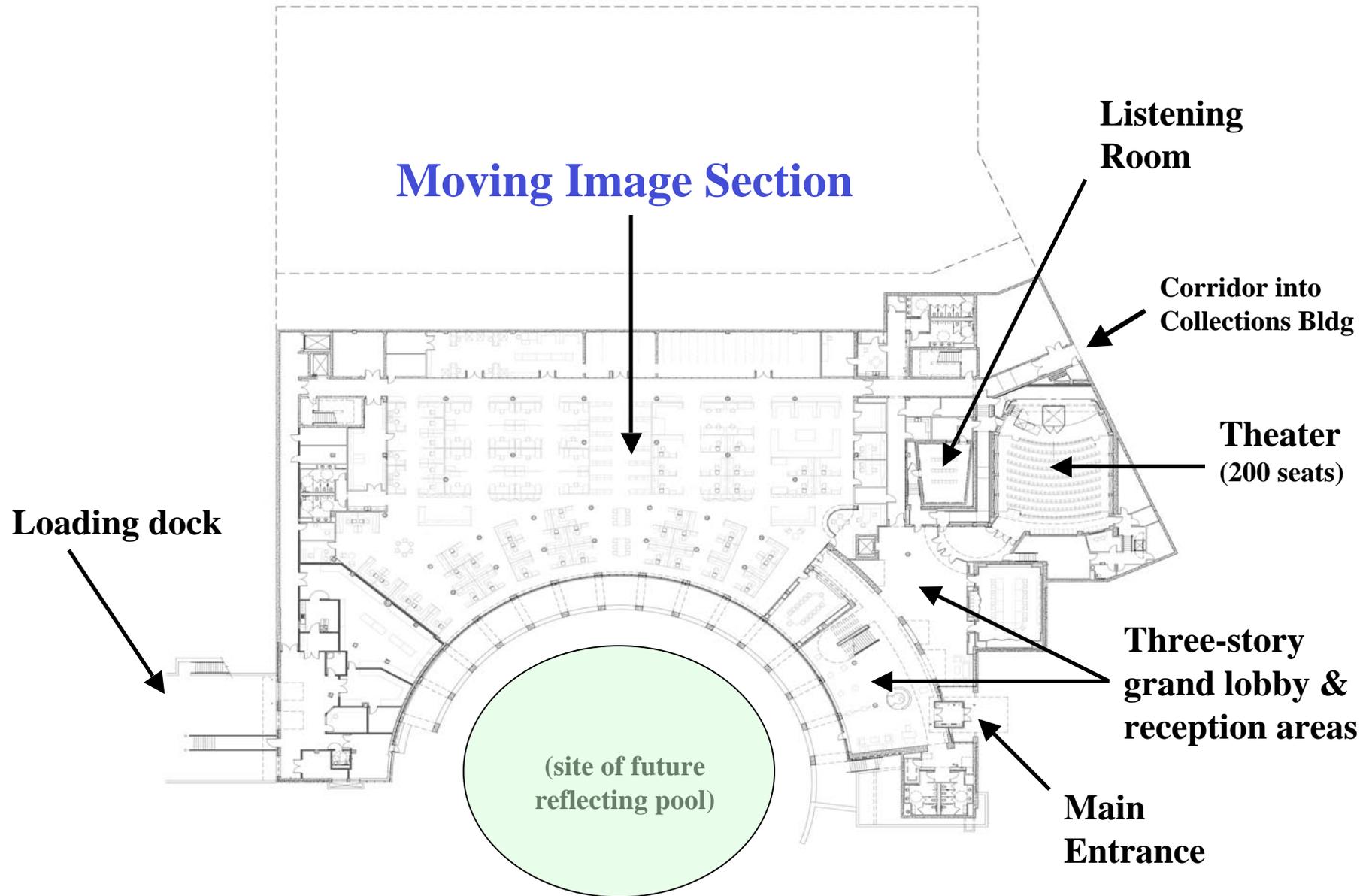
Main Entrance

Future reflecting pool

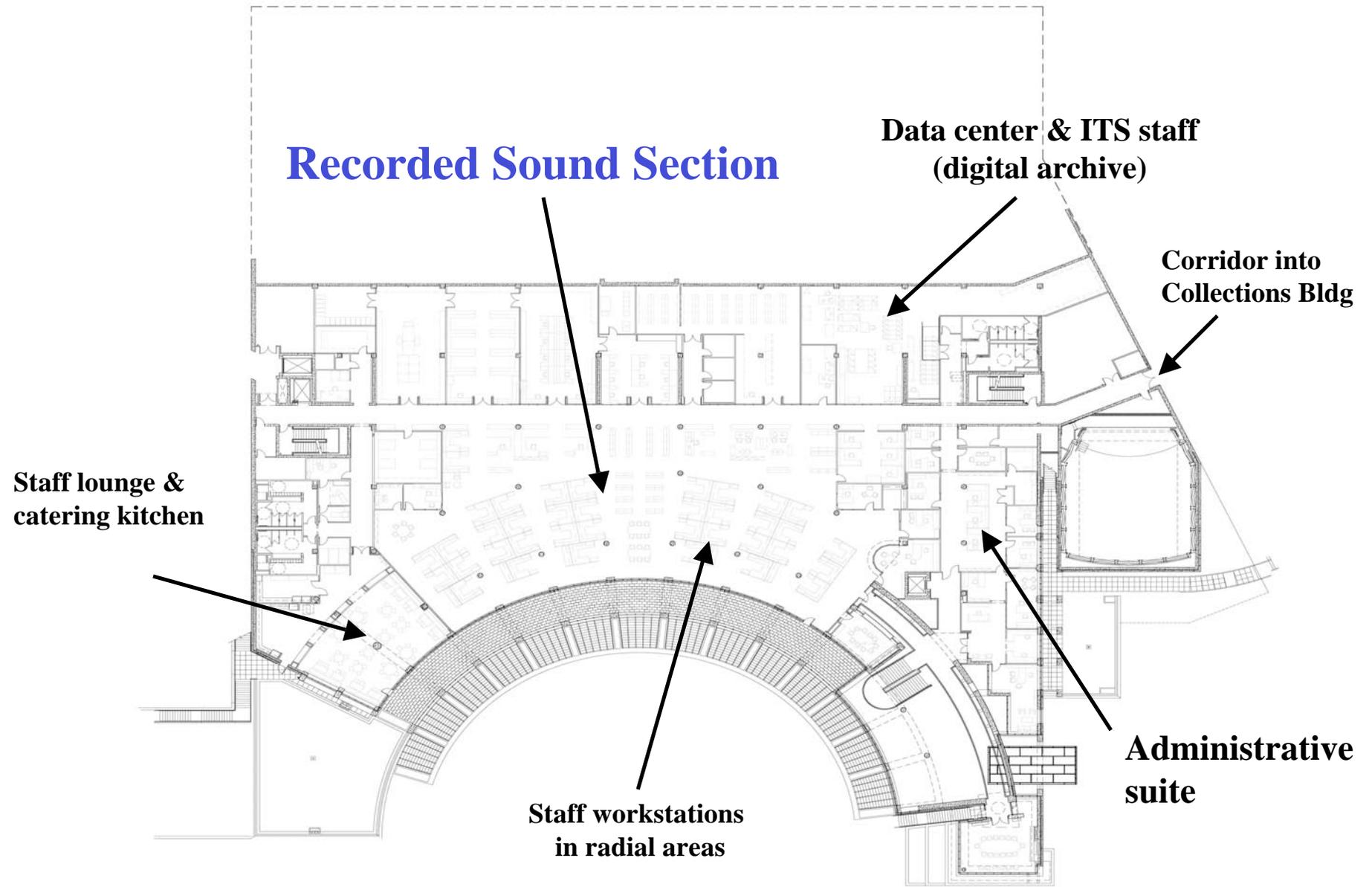
Future parking lot

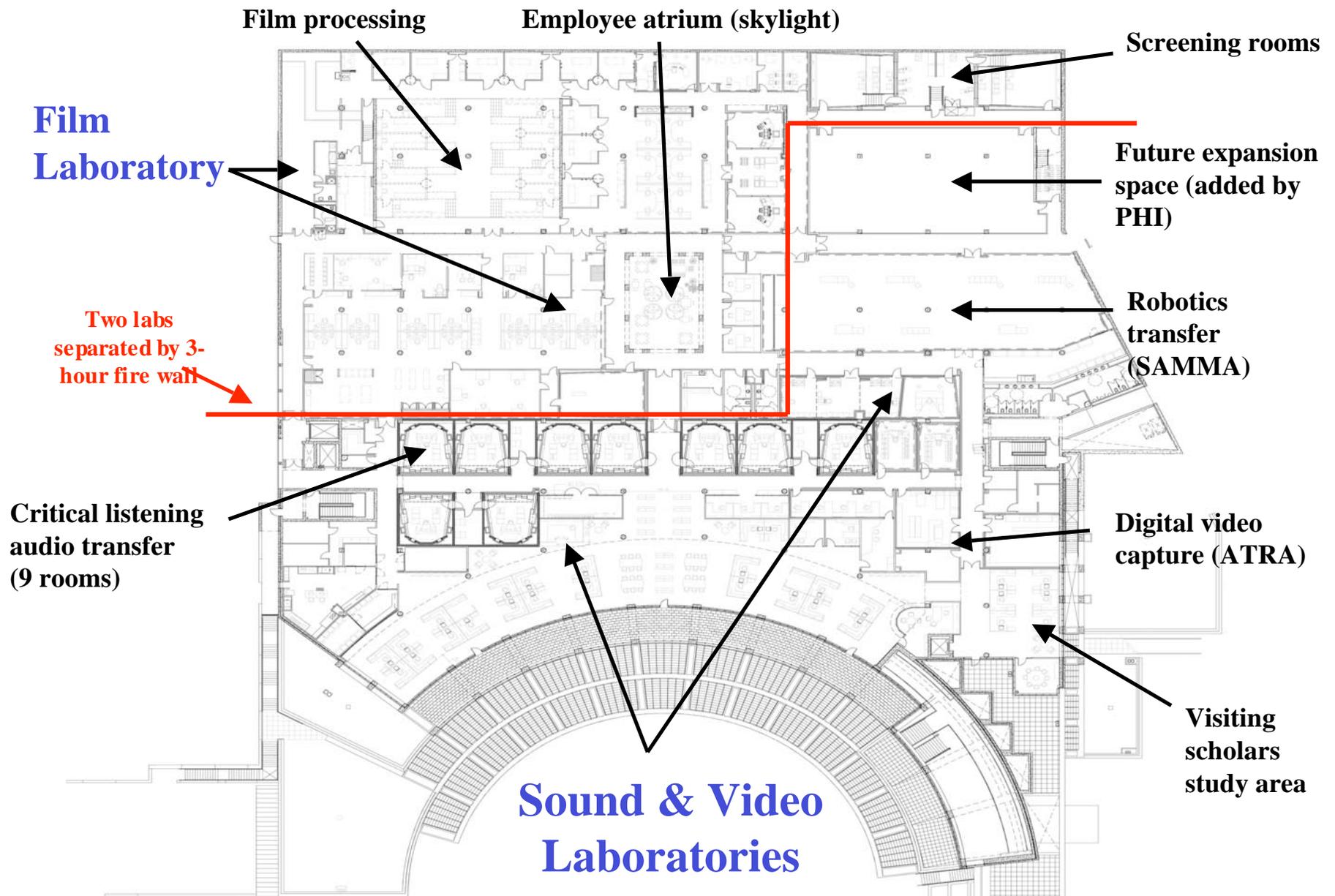
NOV 7 2005

Conservation Building 1st Floor

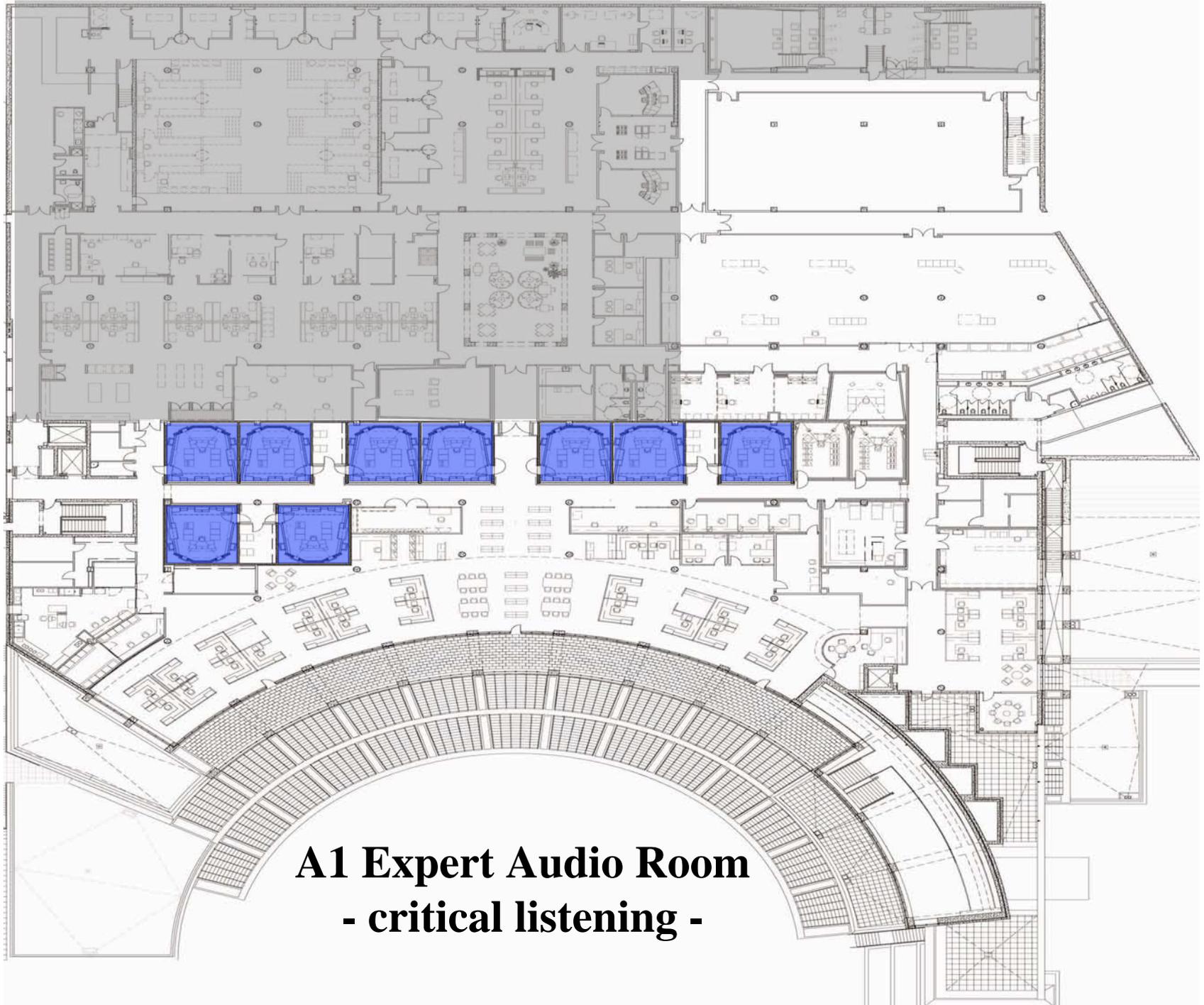


Conservation Building 2nd Floor



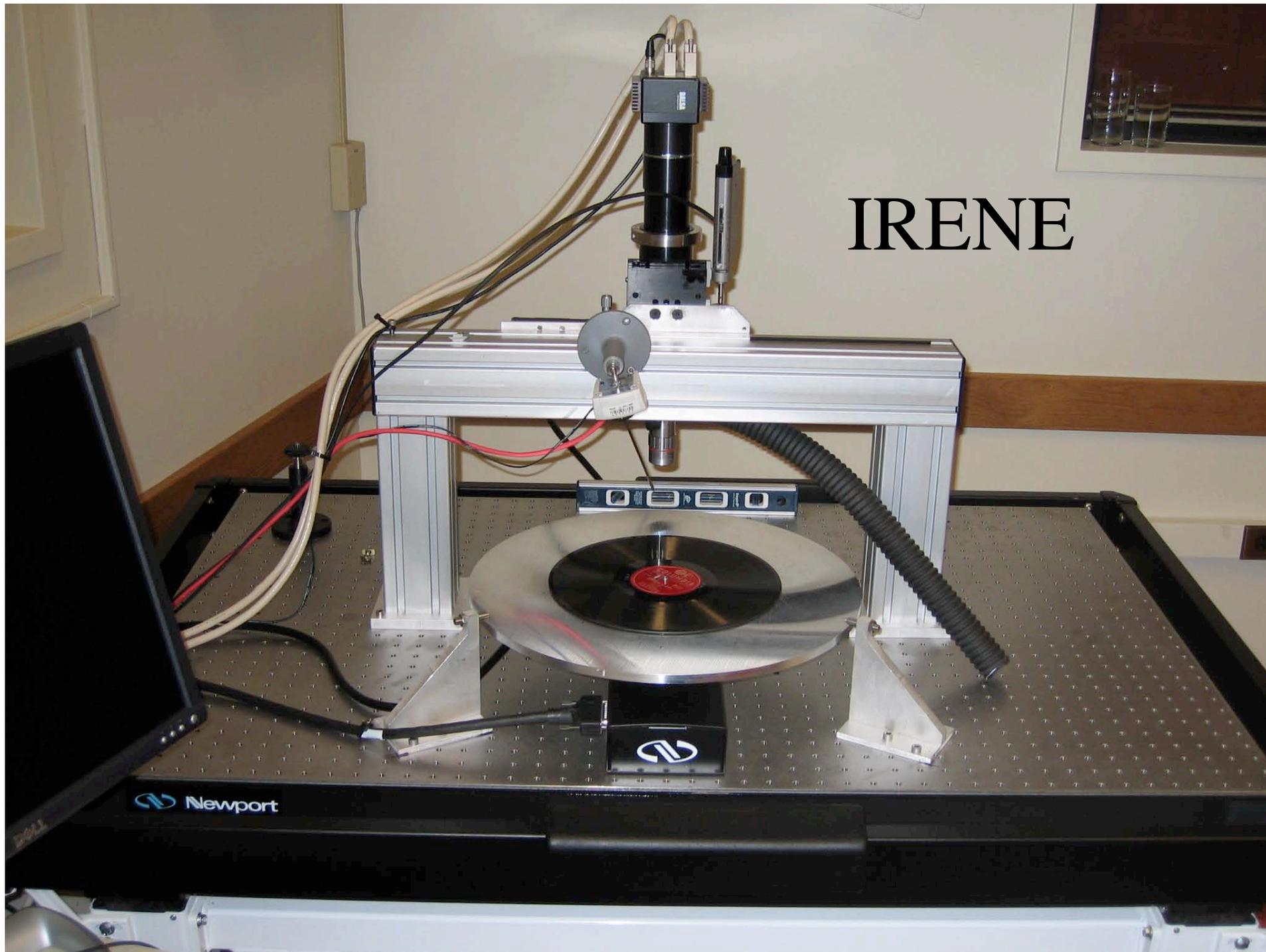


Conservation Building 3rd Floor



**A1 Expert Audio Room
- critical listening -**

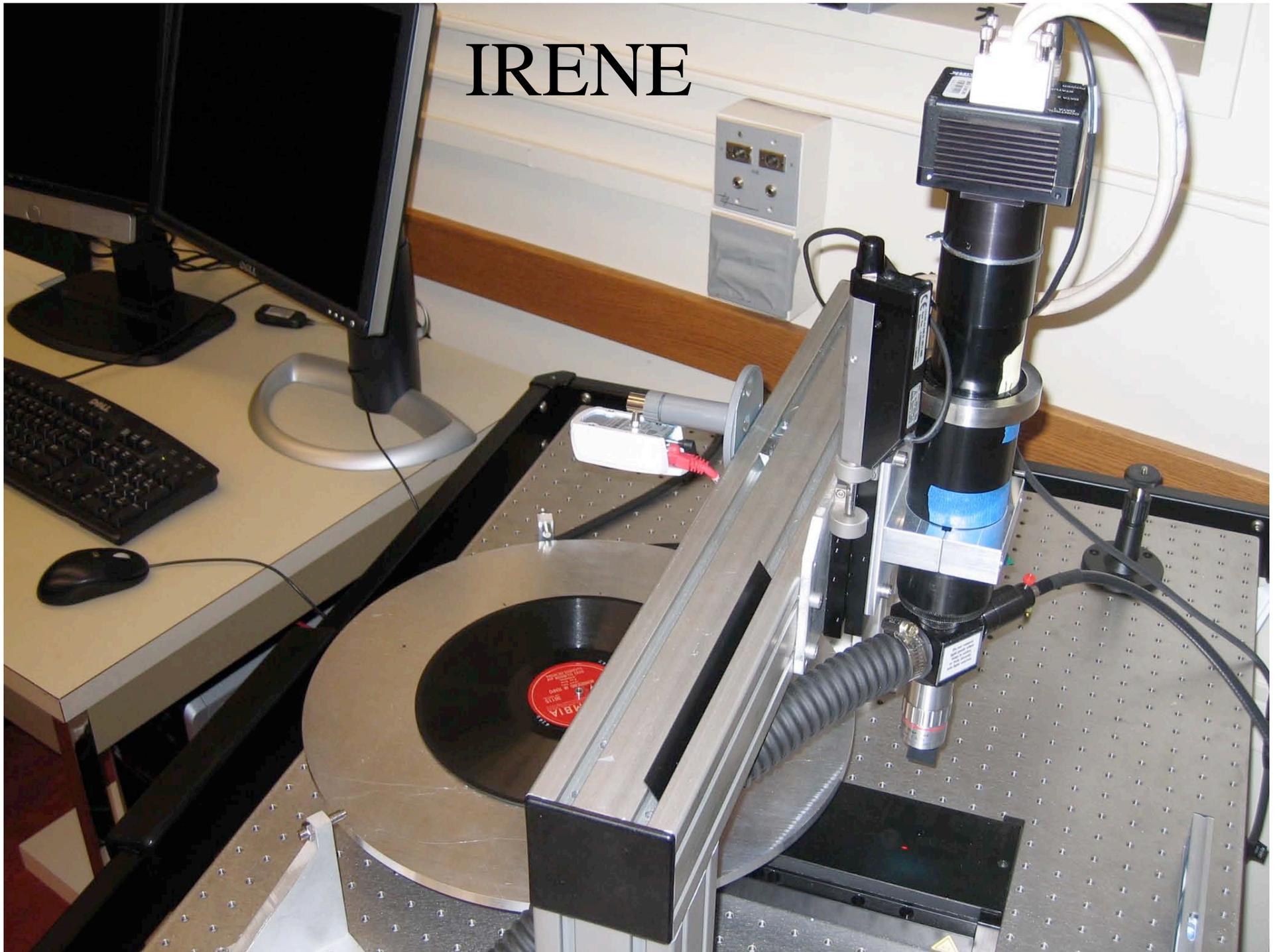
IRENE

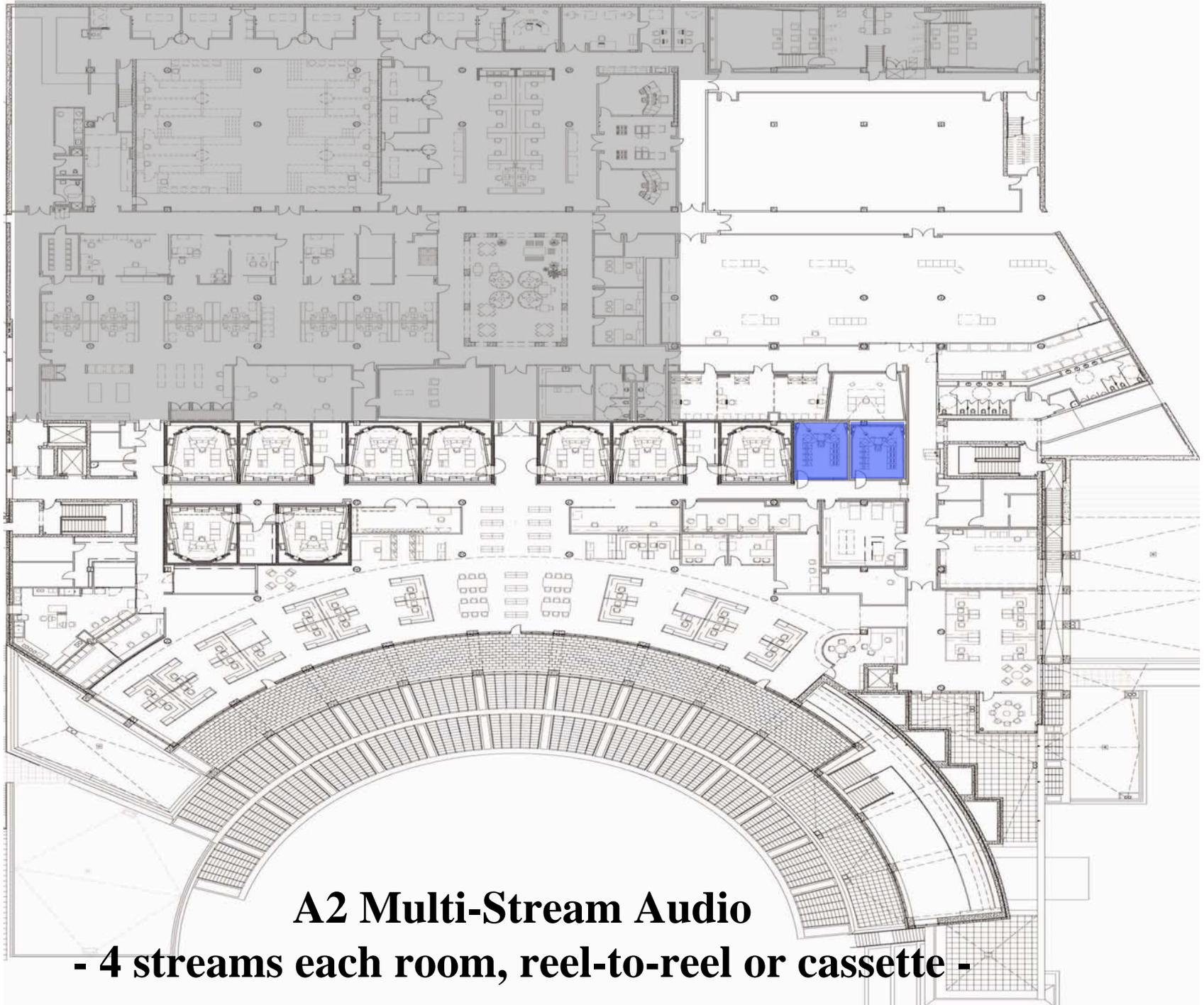


IRENE

- Scientists at the Lawrence Berkeley National Laboratory developed the system, Image, Reconstruct, Erase, Noise, Etc. (IRENE) to help preservationists restore at-risk recordings and improve audio quality.
- IRENE uses digital imaging technologies to generate in real time high-resolution digital maps of the grooved surface of recordings, allowing preservationists to reconstruct damaged or broken recordings and to capture sounds without interference from the debris and other extraneous sounds present on deteriorated recordings.

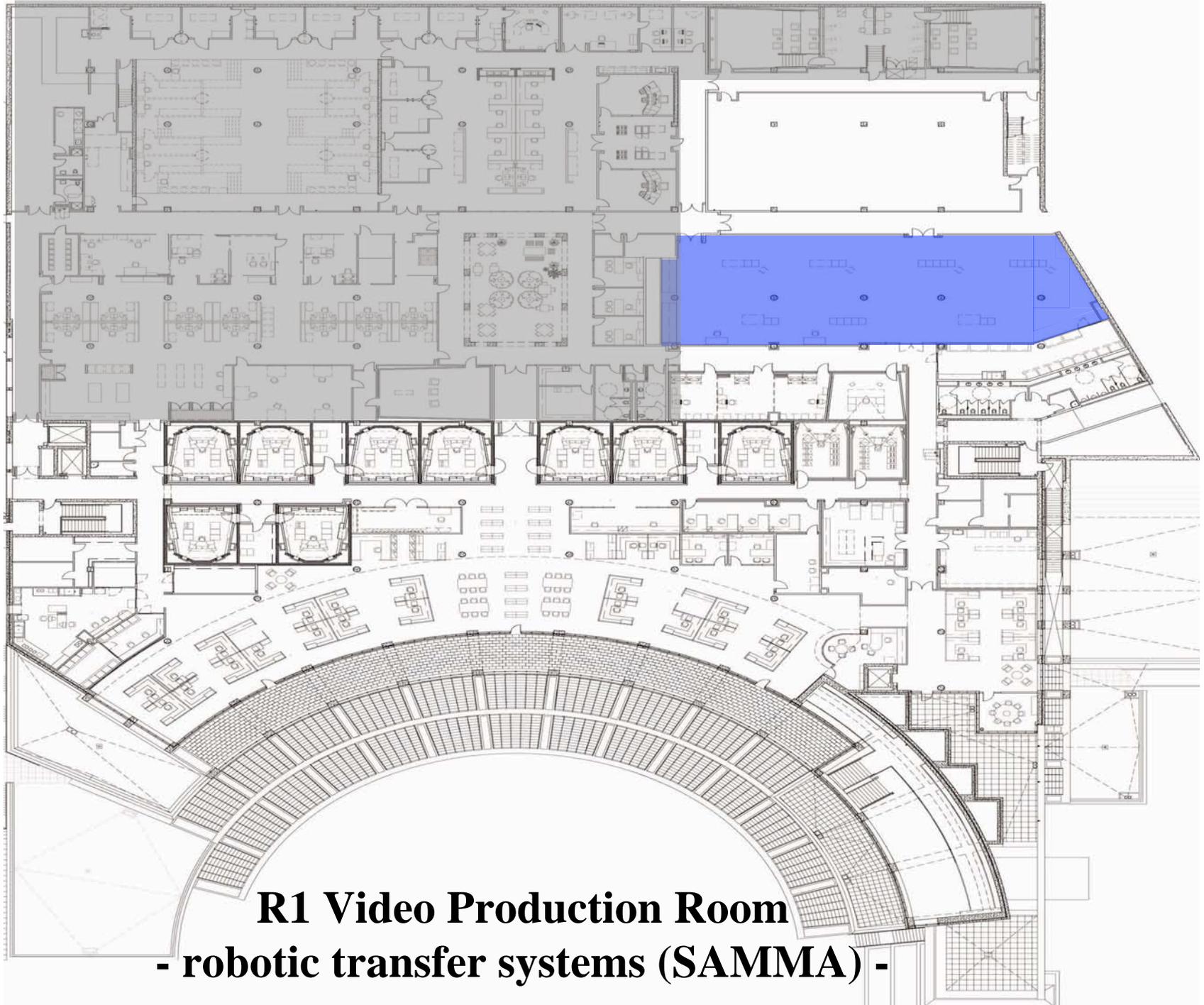
IRENE





A2 Multi-Stream Audio

- 4 streams each room, reel-to-reel or cassette -



R1 Video Production Room
- robotic transfer systems (SAMMA) -



**R1 Room (Robotics & Central Routing) Installation
(April 5, 2007)**

R1 Room Under-Floor Cabling (April 5, 2007)

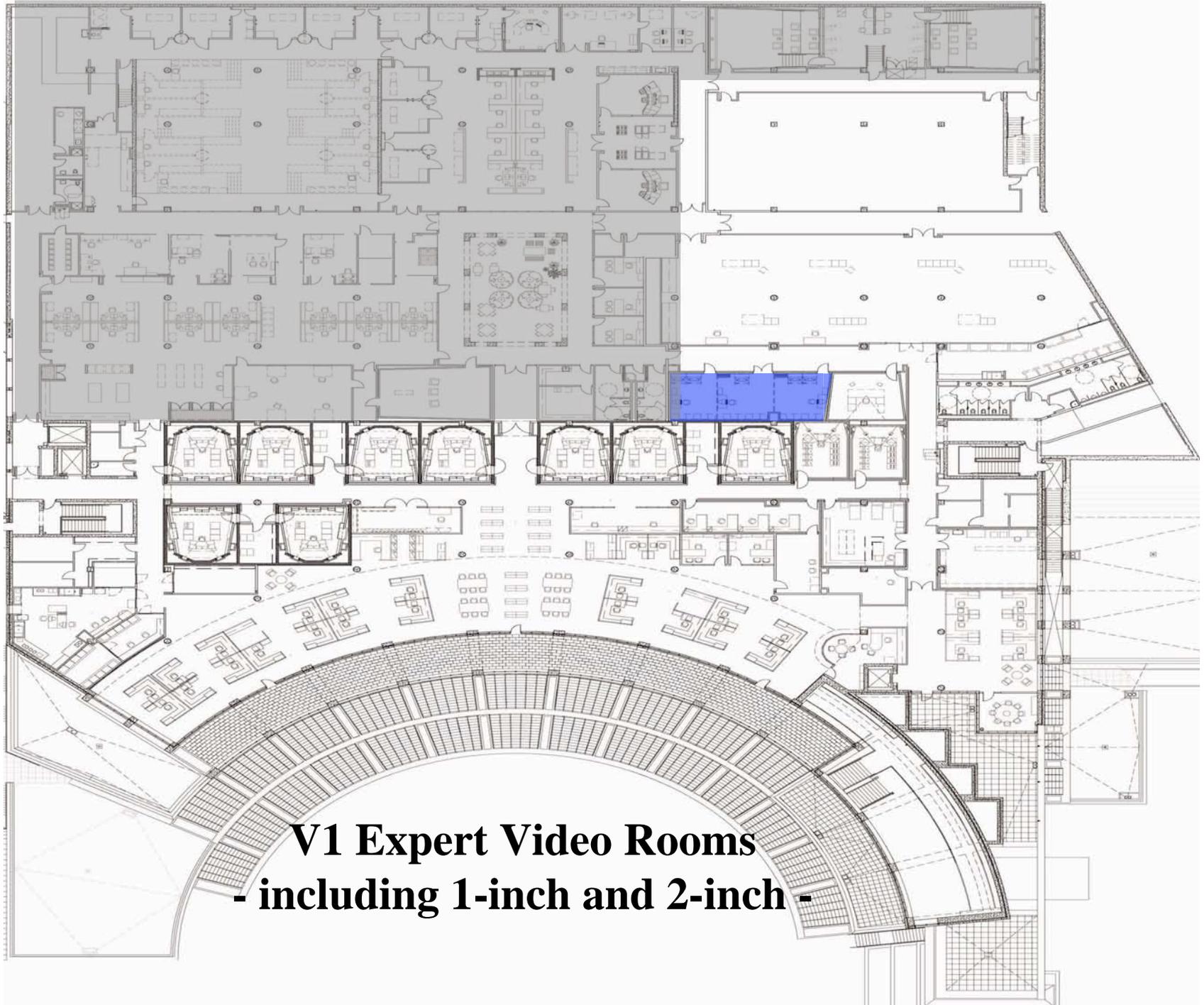


SAMMA

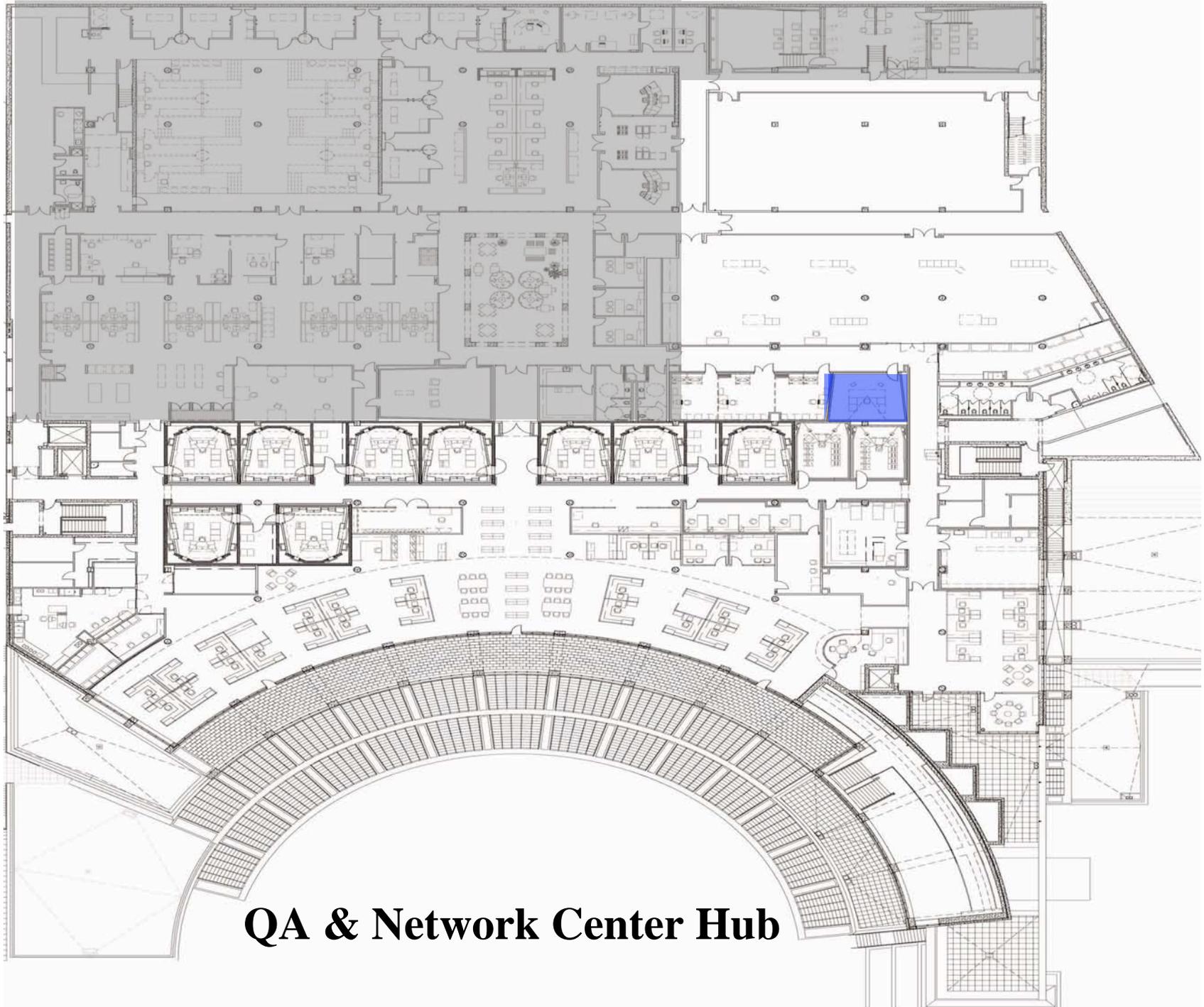
- The Library of Congress is purchasing several SAMMA machines (the System for the Automated Migration of Media Archives).
- Over the next several years the Library will use the SAMMA robotic cassette reformatting system to migrate and digitize over 500,000 television and video items.
- The final product will be preservation-quality JPEG2000 digital files and the technical metadata describing the condition of the media item and the migration process.



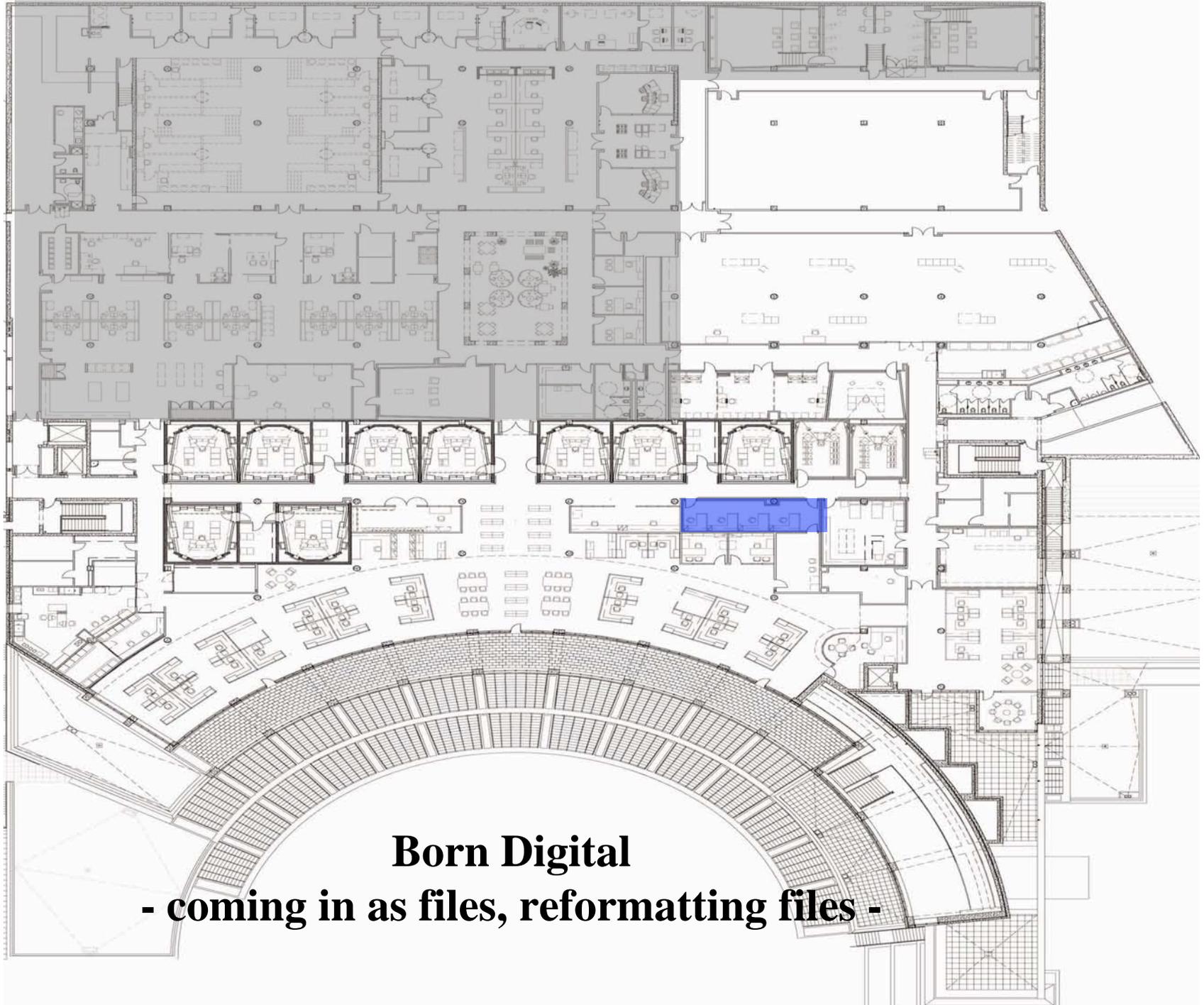
NAVCC
SAMMA



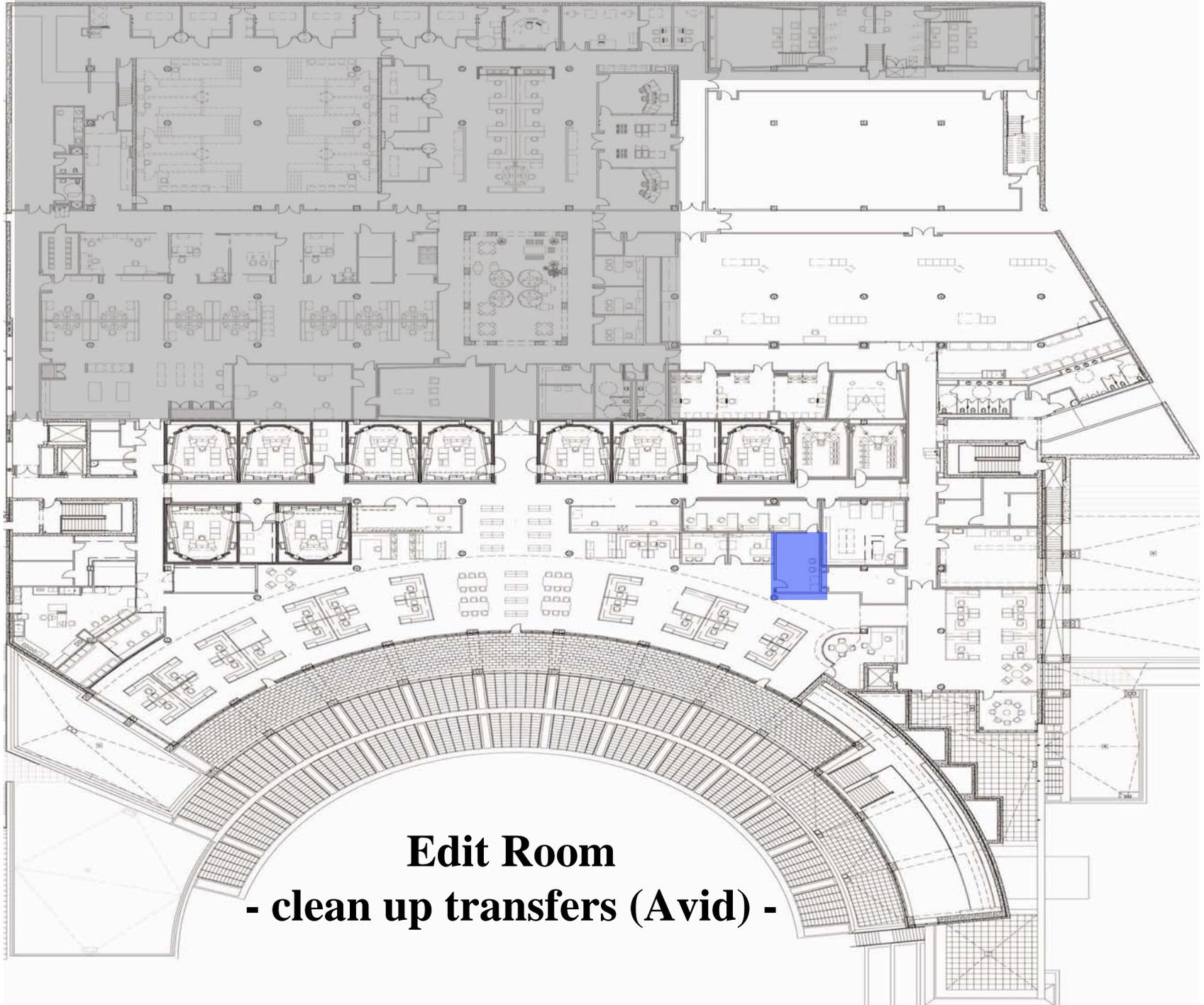
V1 Expert Video Rooms
- including 1-inch and 2-inch -



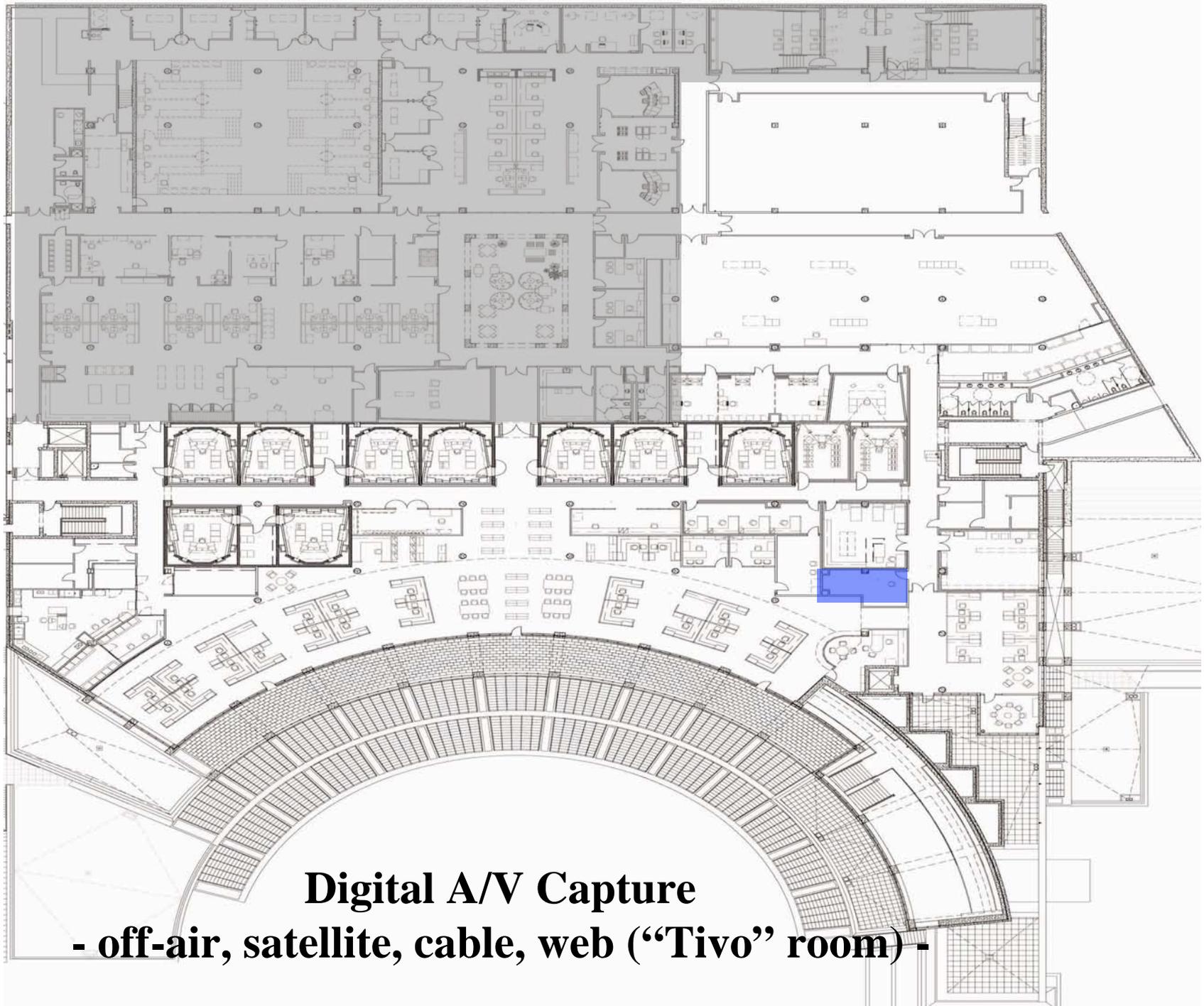
QA & Network Center Hub



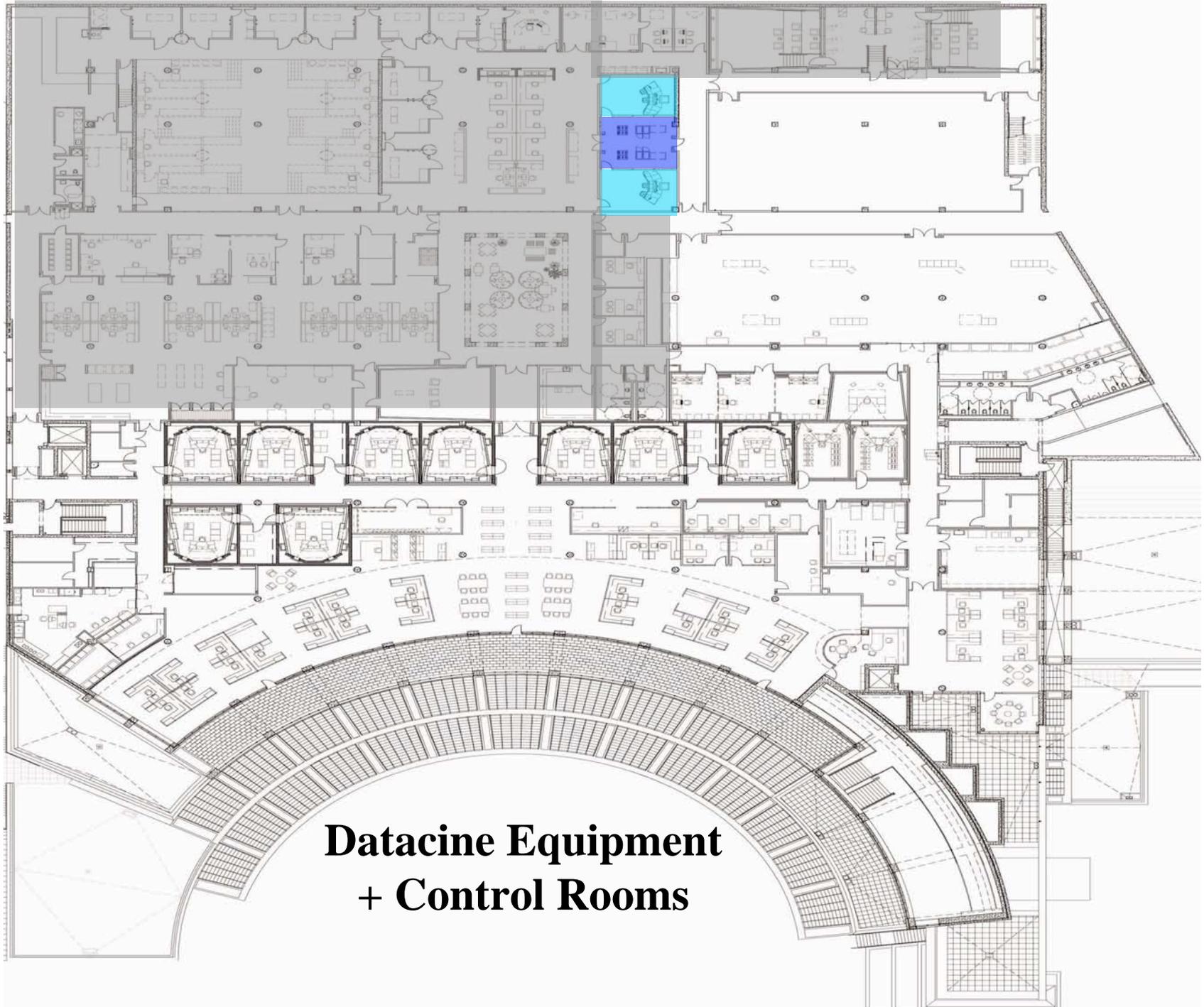
Born Digital
- coming in as files, reformatting files -



Edit Room
- clean up transfers (Avid) -



Digital A/V Capture
- off-air, satellite, cable, web (“Tivo” room) -



**Datacine Equipment
+ Control Rooms**



Begin staff move-in: May 1, 2007

