Multics History Project

02/2006 Status Report
Revision 1

Olin Sibert

18 February 2006

MHP Goals

- Tactical Goal
  - Preserve the MIT-Multics archives
- Strategic Goal
  - Collect enough information to allow a Multics emulator to be created and operated
- Related efforts
  - **`multicians.org`** – Maintained by Tom Van Vleck
    - Community website: history, stories, samples
  - **`bitsavers.org`** – Maintained by Al Kossow
    - Scanned document collection (CHM Multics manuals)

Multics Archives at MIT

- Two main sources:
  - MIT-Multics
    - Campus computer utility service, 1969-1988
    - Stored in building W91
    - Original focus of Multics History Project (started late 2004)
  - Project MAC/Laboratory for Computer Science (LCS)
    - Original Multics development organization
    - Stored by LCS and LCS staff (personal files)
    - Now in Stata Center, MIT Archives, personal archives
    - Came to light February 2006, just now being investigated
- Archives include published docs, internal memos, listings, tapes, personal/business files

MIT-Multics Archives

- MIT-Multics was a campus computing utility
  - Run as a service to MIT and MIT-associated customers
  - Played major development and QA role under contract to Honeywell (through 1984)
  - Separate from original Project MAC / Laboratory for Computer Science (LCS) development team
- Most complete for later (post-1975) material
  - Focus on Multics *as a commercial product*
  - Some material lost (no MABs)
  - No post-MR11 material (relationship ended at MR11)
  - Old material (645 era) quite incomplete

LCS Multics Archive

- Just uncovered (Feb 2006)
  - Not maintained by LCS (now CSAIL) organization
- Personal files
  - J.H. Saltzer (3-4 shelves)
  - Probably others (Corbato, Fano, Sollins, Dennis, Clark)
- LCS Multics “History Room”
  - Approximately 50 boxes
  - Rescued by MIT Archives after flooding in 1988

Multics History Project (MHP)

- Roger Roach worked on CTSS, then Multics, eventually as IS director, retired 2005
- Olin Sibert worked on Multics (initially for Roger, then Honeywell, then independently)
- At Multics Reunion (June 2004), we decided to try preserving the archives that Roger had maintained

MHP Timeline

- June 2004 – The idea
- September 2004 – Worked with Museum on sponsorship
- October 2004 – Set up scanners and computers, tested
- December 2004 – Start scanning in earnest
- March 2005 – MIT backup tapes determined to lost for good
- October 2005 – MIT backup tapes miraculously resurface
- February 2006 – 85% done with paper files from W91
- February 2006 – Discovered LCS archives
- May 2006 (planned) – Deliver boxes and data to Museum
- June 2006 (planned) – Read MIT backup tapes

Scanning Mechanics

- Small network (4 workstations, 3 scanners)
- Low-cost consumer-grade sheet-fed duplex scanners
  - 4 to 8 sheets/minute, 600 DPI monochrome, duplex
    - Scan to PDF (mostly – some TIFF)
    - About 60-100KB/page compressed (Group 4 fax)
    - Some color/grayscale for colored or bad originals
  - Hardware ($400-$800/each)
    - Xerox Documate 252 (fast, but despicable software)
    - Fujitsu Scansnap fi-5110EOX (slow, ultra-reliable)
    - Canon DR-2808C (slow, best with difficult paper)
  - All have idiosyncracies (think “therapeutic reboots”)
- Archive mirrored on multiple external disks

Scanning Workflow

- Processing tasks
  1. Paper handling (preparation)
     - Staples are the bane of our existence
     - I'm no big fan of Acco binders, either
     - Be sure you can wash your hands nearby!
  2. Scanning
  3. Cataloging
     - Excel spreadsheets are easy to edit, but awkward long-term
  4. Scan verification
  5. Paper handling (archival packing)
     - Folders, boxes, labels, Museum barcodes
- We found it *very challenging* to automate effectively

Scanning Lessons

- Physical scanning is not the bottleneck
  - Especially with tiny documents
  - Don’t optimize for scanning throughput
- Very easy to lose track of what’s been done
  - Optimize for record-keeping and tracking
- Different scanners for different tasks
  - Hardware and software issues are different for all of them
- Catalog is hard to plan in advance
  - Optimize for data entry and review!
  - 3 datasets: Catalog database, Scanned files, Boxed paper
  - Lots of tiny (1-2 page) documents, hard to name
- More stuff keeps appearing *(like the LCS archive!)*

Paper Archives

- Manuals (11 boxes) – Very old (1969) up to MR11.0
- Later Memos (8 boxes) – MTBs, MCRs
- Listings (8 boxes) – Final MIT hardcore and BOS
- Core original design (4 boxes) – MSPM
- Older memos (5 boxes) – MCBs, MHDMs, etc.
- HLSUA (3 boxes) – User's group
- Miscellaneous (about 10 boxes) – Not yet processed
- *To be determined: material from LCS archive*

Machine-readable Archives

- Museum's NSA MR12.3-12.5 release tapes - Read clean
- MIT's MR10.2 release tapes - Read, some damage
- MIT's MR10.2 boot tapes - Read clean
- Bull's final MR12.5 dump - Awaiting lawyers, some damage
- MIT's final backup tapes (35 of 36 reels) - Not read yet
- NSA's MR10.2-MR11.x tapes - Somewhere at Museum?
- Grady Booch's punch cards - At Museum?
- Other universities?
- Honeywell System-M in Phoenix?
- *Tapes from LCS?*

Strategic Goal

- Emulation clearly within reach
- Software
  - MIT boot tapes and dump tapes would be enough to create a complete working system
    - Can’t re-create whole MIT-Multics environment without the missing reel, sigh
  - Probably could create a system from MR12.3 tapes, too
- Hardware
  - CPU is straightforward: well-documented (but complex)
  - I/O is not: poorly documented and complex (esp. Comms)
    - Needs combination of Honeywell engineering specs and source code analysis

Questions / Discussion