Sydex Enhanced SIMH Tape Format

December, 2015

Sydex, Inc.
P.O. Box 5700
Eugene, OR 97405

1. Introduction

Sydex uses the basic tape image format described in the 2006 document "SIMH Magtape Representation and Handling" by Bob Supnik; our image files can be treated exactly as a standard SIMH tape image file. However, none of the existing image file conventions, including that of SIMH, makes provisions for retention of metadata. For archivists, this is a serious drawback—quite often, the process of what was done to obtain data from a tape, as well as the appearance of the tape itself, is critical to preservation efforts.

To this end, Sydex has implemented two very simple additions to the SIMH tape format. These additions are optional (i.e., the image data itself is complete if one or both additions are omitted) and occur after the SIMH "End of Medium" flag doubleword (hex FFFFFFFF).

2. The Narrative Text

The first appended record after the EOM flag in a SIMH file is the text generated during the process of reading the source tape. This includes both operator commentary, such as labeling of the tape and condition, as well as messages generated by the archiving program itself, such as block lengths, tape marks and error conditions. The text is plain ASCII, with end-of-lines being represented as a single linefeed (ASCII LF) character. This record is otherwise in SIMH format; that is, prefixed and suffixed by a 4-byte little-endian record length. The first three characters of this record are the string "LOG".

3. Photographic Data

The second appended record after the SIMH EOM flag is the photographic image of the medium (i.e. tape reel) itself. Currently, this is in JPEG FIF format, which can be readily recognized by the ASCII characters "JFIF" in bytes 6-9 of the record. This is currently our standard for photo information; Sydex reserves the option of using other graphical formats in the future.

4. Media Errors

The original SIMH document specifies only that blocks containing media errors be flagged with bit 31 of the record length set (80000000 hex). Nothing is said about how to differentiate those media errors that return data and those that don't. For the latter type of error, we use a record length count of 4, which is impossible with most tape formatters, as it will be treated as noise and ignored. The actual content of the 4 byte record is undefined.