

Tape Processing Issues

When copying data from one tape to another, a number of questions arise. The answers affect both technical efficacy and cost.

Copy Mode: Duplicate or “Smart Copy”? It’s nearly always possible to duplicate tapes—even though the output medium may be different than the input (e.g.- 9-track reels to 3490 cartridge). As long as the output is of equal or higher capacity, one can do a byte-for-byte, block-for-block, filemark-for-filemark replication. This is often the best solution if each tape is a “stand-alone”, but it sometimes results in wasted capacity.

When files have been written to tape in a well-documented standard such IBM or ANSI standard labeled files, it is then possible to “smart-copy” them to more efficiently utilize the capacity of the output medium. For example, if several reels of 9-track tape had originally be written as a single multi-volume dataset in standard labeled mode, then it will be possible to combine these into a lesser number of higher capacity output tapes. This is called “re-voluming”. In addition, one may sometimes make better use of tape capacity by “re-blocking”. This is because the gaps between blocks of recorded information on tape take up physical space. Using a larger block size (thus fewer blocks) can result in dramatic improvements in tape capacity.

Input Errors. The tape drives and software* we use will automatically retry the reading of tape blocks containing errors until a “clean read” or the retry-count is exhausted. Then the operator is presented with a choice of (1) repeat the retry-cycle, (2) skip the bad block entirely--resulting in a missing block on the output tape, or (3) abort the copy run. If a block cannot be read after repeated retries, it is considered a “hard error”. Since hard errors require operator attention, tapes with many errors result in more time and labor expense. Further, copies of tapes that had hard errors are obviously of less value to the customer. Therefore, (a) steps must be taken to avoid spending too much time on such tapes, and (b) there is generally a cost to the client even for failed copies.

Output Errors. These are very rare, especially when new high-quality output media is used. Virtually all such errors are caught by the system. In such cases, we will repeat the copy onto another output tape.

Full Verification. To ensure the accuracy of the copy, our system can verify it by comparison with the original input tape. (The National Archives Center for Electronic Records* verifies all tape copies.) However, this process doubles the processing time--and occasionally results in a re-copy and re-verify. To avoid some of this extra expense, customers with large volumes of tapes can elect to verify a sampling of their tapes (e.g.- the first 20 or so). If the tapes all verify cleanly, they can opt to forego verification on the remainder.

Compatibility. Even after a 100% “clean” copy is made, there may be readability problems at the client side. Possible causes: software incompatibility (receiving system cannot handle the *logical* tape format even though media is correct); incorrect original specifications (e.g.- tape density or compression); mis-alignment of read heads. This kind of problem can be detected and nipped in the bud by an initial test.

Old or Damaged Tapes. Magnetic tapes deteriorate over a period of time, especially if stored in less than ideal conditions. Humidity, temperature, and poor “tape-pack” are the chief culprits. In extreme cases tapes may suffer from “stiction”—a decay of the tape’s adhesive material holding the magnetic coating to the substrate. The resulting sticky tapes will eventually stick to the drive’s read heads, causing the drive to need cleaning, and, often, a manual rewind of the tape. Cleaning and re-tensioning the tape will often clear up many soft—and some hard—errors, and will slightly reduce “stiction”. The worst sticky tapes can only be helped by baking. This is a lengthy process, and can be expensive.

Output to CD. In many cases CD can be used as a preservation medium for tape data, rather than copying to other tapes. Please see our companion piece “*Tape Preservation via CD*” for more information.

Questions To Be Addressed

Duplicate or “Smart Copy”? Which approach is appropriate for your project?

Hard Input Errors. How many hard errors (resulting in missing output blocks) can you tolerate in a tape?

Verification. Do you want to verify all tapes? Just a sampling? (We’ll make recommendations after test runs.)

Compatibility/Testing. If any sizable volume of data is to be converted, it is vitally important for the client to provide a test sample and carefully review the resulting output.

Damaged Tapes. Do you have valuable data on old or damaged tapes? We have the expertise to help, but this must be discussed on a case by case basis.

* The system we use for converting and copying tapes is a version of the proprietary package we developed for the National Archives’ Center for Electronic Records-- which we continue to support and enhance. (The Center is mandated to copy and preserve vast amounts of data from the many Federal Agencies.)