LA-UR 86-2825

CONF-8610128--3

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

LA-UR--86-2825

DE86 015332

TITLE CHOOSING A DOCUMENT-FORMATTING SYSTEM

AUTHOR(S) Richard K. Wallace

SUBMITTED TO DECUS PUBLICATION GROUP
San Francisco, California

DISCLAIMER

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

DISCLAIMER

CHOOSING A DOCUMENT-FORMATTING SYSTEM

Richard K. Wallace
Los Alamos National Laboratory
Los Alamos, NM

ABSTRACT

After surveying available tools for formatting large computer code manuals, we chose the TeX system, to be initially implemented on VAX 11/780 and 8600 computers. We also recognized that a "What You See Is What You Get" word processor offers sufficient capabilities for small (5 - 10 page) reports and manuals, and recommended that WordMARC be considered for formatting in those situations.

BACKGROUND

Los Alamos National Laboratory is a federally funded applied research laboratory managed by the University of California for the U.S. Department of Energy under contract W-7405-ENG- 36. The Laboratory engages primarily in energy, national defense, and accelerator/nuclear physics research. It employs about 7800 people and is divided organizationally into 43 Divisions. This paper discusses criteria used by the Applied Theoretical Physics Division (X Division) to select a document formatting system. X Division consists of about 260 employees, more than 200 of whom have doctorates in physics-related disciplines and all of whom have extensive interactive computing experience.

The major Laboratory computing center, managed by C Division, is the Central Computing Facility, which contains 7 Cray supercomputers, 8 large CDC computers, and 10 DEC VAXs, with a total computing capacity equivalent to 20 Cray-1 supercomputers. In addition, nearly 100 Distributed Processors, all VAX 11/780, 785, or 8600s, are scattered over 43 square miles, linked by DECNet and managed by the individual divisions. Owing to the defense work, the computing resources are divided into classification partitions, each completely separate (no communication channels) from all other partitions.

PURPOSE

In August 1984, we formed a Committee to recommend a replacement for the then-current computerized documentation tools(TRIX/RED, REDPP), which would be unavailable after removal of the Laboratory's secure CDC 7600. Recent turnover in the code user groups emphasized the lack of current, comprehensive documentation (user and physics manuals) for the major X-Division production codes. This lack of documentation increases the training time required for new users and code developers and hinders efficient code use by them and by experienced users. The existing code manuals must be continually revised and expanded as the codes rapidly evolve.

We have therefore surveyed the field of document production in search of a modern, efficient, long-term document-formatting system that will satisfy our need for producing thorough, clear, current documentation as simply as possible. The system development was coordinated with C Division to reduce duplication of effort and prevent future compatability problems.

SUMMARY

We recommended that TeX be used for formatting X-Division code manuals. Although the Division should not require the use of TeX, that tool should be seriously considered for any major documentation effort. We recognize that WordMARC may offer sufficient formatting capabilities for small (5-10 page) reports and manuals and should be considered for those applicat ons.

To obtain the full benefit of the TeX documentation system, the following hardware was recommended:

- A high speed (at least 24 pages/min) laser printer.
- An upgrade for one of our two VAX 11/780s to a DEC 8600 to provide greater responsiveness, larger CPU capacity, and improved availability of full screen text editors. Even if TeX became available on CTSS (the Cray operating system), the local VAXs could be heavily used for text entry and WordMARC applications.
- A low-cost (under $3000) laser printer that can produce local (in office) output; possible candidates include the DEC LN03 and the HP LaserJet.
- Workstations with a preview capability for frequent TeX users.

C Division was strongly encouraged to provide the following software support:

- A CTSS (Cray) implementation of TeX; this is in progress.

- Simple lineprinter/ASCII output from standard TeX DVI files; rudimentary package is now in use.
- Central Computing Facility output capable of producing 5000 formatted pages/day.
- A method to merge TeX text with graphics files that are in the unique Los Alamos Common Graphics System metafile format.
- Conversion programs for TROFF, TRIX/RED, and VMS WordMARC.
- Classified consulting services on TeX.
- "Writer's Workbench"-type software (such as a spelling checker) for TeX files.

Justification - Requirements

The selection of TeX for the X-Division formatting system was based on its satisfaction of the following unique X-Division requirements. The system should

1. be easily portable to new operating systems, minimizing future translations such as must now be done for the large number of LTSS (CDC 7600 operating system) TRIX/RED files. The system should also be widely used outside of DOE to increase the support for and knowledge about it,
2. be declarative (using predefined structures for headers/footers, sections, paragraph indentations, examples, etc.) rather than procedural (requiring the author to define page layout during text-, or content-entry). This requirement allows a few experienced people to maintain the detailed page layout macros, whereas casual users simply enter text,
3. easily accept mathematical equations and format them with as little user assistance as possible,
4. be capable of merging text with computer-generated graphics,
5. have automatic Table of Contents generation,
6. have automatic Index generation,
7. provide for nested tables,
8. have a source file format that facilitates macro construction to support detailed page layout macros, translation macros (from previous systems and into future systems), and text unformatting macros (to easily allow incorporation of arbitrary machine-readable text),
9. allow text input from any ASCII terminal (including Tektronix 4000 and 4100 series),
10. be accessible transparently from CTSS to eliminate user investment in learning a different operating system or accessing special hardware (most users work exclusively on the Cray CTSS systems rather than on VAXs),
11. produce simple ASCII text output for online help files from the same source file that produces fully formatted documents,
12. allow comments in the source file,
13. facilitate page layout changes or even allow determination of the layout after text entry,
14. symbolically reference equation, figure, section,and page numbers, and
15. allow "interactive" execution to provide error diagnostics and allow recovery from minor source file errors.

COMPARISONS

The major software for code documentation that begins to address the requirements listed above is the following:

Interleaf

Advantages:

1. Interactive "What you see is what you get" (WYSIWYG) system. This can be much easier and faster to use than a batch formatter for small files.
2. Instant feedback (screen shows all page layouts, fonts, text sizes, pagination, etc.).

Disadvantages:

1. No symbolic equation entry. Equations must be entered with a graphics package that draws each individual symbol or character on the page.
2. No symbolic referencing of equation numbers, sections, etc.
3. Operates only on SUN, APOLLO, and VAXStation II workstations.
4. Cost is $12,000 per workstation node, which is prohibitively expensive.

Interleaf was the most capable WYSIWYG formatting system on the market. It would unquestionably be the most productive system to have for a single user. However, the lack of symbolic mathematical entry and the unavailability for a timesharing system are fatal flaws for our purposes. The $12,000 per node price, coupled with the price of providing SUN-class workstations to everyone contributing text, is prohibitive. In addition, no SUN-class workstation has been approved for classified processing.

WordMARC, Version 5 ("Composer")

Advantages:

1. WYSIWYG system that is much easier and faster to use than a batch formatter for nonequation typing of small files.
2. Instant feedback of text and general page layout.
3. Preserves author's meaning (equations displayed on first typing).

Disadvantages:

1. Procedural; no declarative format.

2. Cannot easily change existing document format.

3. No comments allowed in source file.

4. VT100 emulation terminal required (for example, no Tektronix 4014).

5. Response slows to unacceptable times with large documents and many simultaneous users. Response time is more critical for completely interactive systems. The continuous formatting increases the CPU load compared with that of a batch formatter.

6. Less involvement allowed to professional editors/designers.

7. Limited (and in some cases insufficient) mathematical capabilities.

8. No proportionally spaced laser printer output.

The disadvantages indicate that WordMARC may be ideal for formatting memos and short reports but would be inadequate for very large manuals. Although WordMARC (from Marc Software) was specifically compared here, the disadvantages are similar for other WYSIWYG systems, such as MASS-11. They all generally require VT-100 emulation capability, are generally procedural (requiring some author involvement in page layout), are difficult to use for changing page layout retroactively, and require interactive computer response time. However, screen editors in such WYSIWYG systems could be used to prepare the ascii input files for a batch editor, such as TeX or TROFF.

We found no WYSIWYG systems with all the capabilities listed under "Requirements" above. However, two batch formatting systems in common use (TROFF and TeX) could satisfy nearly all of our requirements, and their respective advantages are listed below. C Division has decided to support both TROFF and TeX as Laboratory document production systems.

TROFF with EQN and TBL

May be easier to learn than standard TeX (but not significantly easier than LaTeX).

Better table generation capability than standard TeX.

Writer's Workbench editorial software available.

TeX

1. Arbitrary length command names (TROFF restricts commands to less than 2 characters).

2. More portable than TROFF (TeX is available in generic Pascal and C, whereas TROFF is tied intimately to the UNIX operating system).

3. Los Alamos Common Graphic System TeX interface exists for QMS laser printers, so merging text and graphics is a reality.

4. Slightly more control over output appearance.

5. More widely available screen preview systems (including SUN, APOLLO, IBM AT, Apple Macintosh, and Tektronix 4014).

6. TROFF requires the UNIX operating system, which is currently unacceptable for classified computing.

Points 2 and 6 above are sufficiently serious that we consider TROFF an unacceptable solution. TeX is therefore the most appropriate choice for an X-Division formatter.

CONCLUSIONS

We chose TeX as our standard document formatting system, largely because of its great portability compared to TROFF. For small memos and reports, many secretaries still use WordMARC. Since we reached our decision, several other divisions at the Laboratory have begun using TeX, and the the official publication division (which uses an APS-5 phototypesetter for high-quality output) is committed to switching completely to TeX. The Laboratory is moving to standardize on Postscript (from Adobe Systems) as a common text/graphics device independent file structure, and we are now obtaining hardware and software to allow TeX output through Postscript devices. In addition, the Laboratory has just moved to support LaTeX (a TeX macro package) as the standard version of TeX. We currently use LaTeX on SUN, APOLLO, VAXStation II workstations, IBM XT, AT, Apple Macintosh, VAX/VMS, and VAX/UNIX, and have contracted for an implementation on CTSS.

FURTHER INFORMATION

- TeX: TeX Users Group, P.O. Box 594, Providence, RI 02901.

- LaTeX: TeX macro package developed by Leslie Lamport (now at DEC). For information, contact the reference under "TeX".

- TeX on workstations, and output to Postscript devices: Textset Inc., 4116 4th. St., P.O. Box 7993, Ann Arbor, MI 48107. (313) 996-3566.

- TeX on IBM XT/AT: PC TeX Inc., 20 Sunnyside, Suite H, Mill Valley, CA 94941, (415) 388-8853, or Micro\TeX, Addison-Wesley Publishing Co., Educational Media Systems Division, Reading, MA 01867. (617) 944-3700, ext. 2677.

- WordMARC: Marc Software International, 260 Sheridan Ave, Suite 200, Palo Alto, CA 94306. (415) 326-1971.

- Interleaf: Interleaf Inc., 1100 Massachusetts Ave., Cambridge, MA 02138. (617) 497-5570.

- MASS-11: Microsystems Engineering Corp., 2040 Hassal Road, Hoffman Estates, IL 60195.

- TROFF: UNIX System manual, Bell Laboratories or Computer Science Division, University of California, Berkeley, CA 94720.