SOME OBSERVATIONS ON THE PREPARATION OF PAPERS FOR ELECTRONIC PUBLICATION

J. Poole, CERN, Geneva, Switzerland

Abstract

Papers presented at the 1995 Workshop on LEP Performance were published electronically as a pilot scheme for EPAC'96. This new publication method requires special care on the part of authors and more work in the preparation of the final document for the editorial team. Experience with the pilot scheme has highlighted the need to give more help to authors with the preparation of their papers, particularly for the inclusion of graphics and use of fonts. The majority of the time after receipt of the postscript files (manuscripts) is spent fixing problems which either prevent the postscript file from passing through the software or which result in poor quality in the final document. This paper summarises the difficulties encountered and suggests improvements in the methods and procedures to be implemented for EPAC'96.

1 FROM POSTSCRIPT TO ADOBE PDF FORMAT

Adobe's Acrobat suite of software allows one to process postscript files into their PDF (Portable Document Format) which is suitable for electronic publication. The postscript is converted with the 'Distiller' which also builds 'thumbnails' (icon-like versions of each page), and an index for all words that it can recover from the postscript. In order to recognise words in the document, **the postscript file must have been created using non-bitmapped fonts**.

Each PDF file carries with it a set of information containing a title, the author name, subject and keywords. Acrobat Exchange is another tool which is used in the production of the final document. It allows one to insert hypertext links either within the document (called bookmarks) or to other documents. It also allows one to move/insert/replace/crop pages and cut and paste but one cannot change the text.

Acrobat Catalog creates indexes across many documents and allows boolean searches on the document's header information as well as the text on the pages.

All that the user requires is a reader, and these are in the public domain for several platforms.

The aim for the EPAC CD-ROM is to have the following:

- One file per paper containing thumbnails
- Author Index and Table of Contents with hypertext links to the relevant pages
- Reader software for as many platforms as possible (PC, MAC, Sun)
- Cost permitting, Acrobat Exchange LE software, which will replace the reader for PC platforms and allow boolean searches.
- Word search within a single paper through the reader
- Possibility to print one paper at a time (through the reader)

2 PERFORMANCE

The quality of the visual image depends on the platform/display capabilities and on the type of font used. In general bit-mapped fonts are extremely difficult to read except at the largest magnifications, they also slow down the performance considerably. This is because the bit-mapped fonts are designed to be rendered at a single resolution, like 300 or 600 DPI, but it is only on a computer screen where the resolution varies and the effect is noticeable.

Another key factor in the speed with which a page is displayed is the size of the file. Apart from text which is using bit-mapped fonts, a page is displayed very quickly. Graphics and illustrations can be painfully slow if the files are too big.

3 WHY IS PAPER SIZE AN ISSUE ?

The Distiller places the text/images which are described in the postscript on a page which is of a pre-determined size. For EPAC we wish to set this to correspond to A4 size paper but when our colleagues in the USA want to print on their own standard (US letter) size paper we want the result to be reasonable. The height of the text is defined for all contributions (241mm) and normally one would centre this on the page. This results in different bottom margins for the two sizes of paper and when they are converted to PDF the A4 text will appear higher on the page. This situation is made worse when someone in the USA wants to print a paper from the PDF file.

The solution which we have adopted is to have different margins for the two paper sizes so that the text area is of the same size and the lower left corner is the same distance from the edge of the sheet for both sizes. A printout of the postscript file looks a little strange but the electronically printed version will be much better.

4 PREFERRED SOFTWARE FOR PAPER PREPARATION

WORD and LaTeX2e are preferred by the editorial team. Experience at the Chamonix Workshop on LEP Performance was very good with LaTeX with 90% of the documents successfully processed at the first attempt. The WORD documents were only 50% successful but this is in stark contrast with experience at the 1994 Advanced Accelerator Concepts Workshop and PAC'95. For these other proceedings most of the LaTeX prepared papers had used bit-mapped fonts and had to be re-worked.

We hope that with the use of suitable fonts (e.g. Times, Symbol and Zapf Dingbats) with LaTeX2e the problems will be greatly reduced. The problems with WORD arose largely because it is not possible to enforce a style and it is hoped that, as authors become accustomed to producing papers for electronic publication, this will improve.

It is acceptable for authors to use other software to prepare their papers but they will then have to be careful to follow the layout instructions and conventions and be sure that their postscript files will be acceptable.

5 THINGS TO AVOID

Type 1 Computer Modern Fonts with Textures do not process well - the result is not useful.
EPS figures from Illustrator or Photoshop do not fit correctly in LaTeX documents.
PCX format (PC figures) in LaTeX documents – use EPS.

6 USING LATEX

6.1 Font Selection

This has been the single biggest problem area for the two conference proceedings being produced in the USA. We propose that people should use LaTeX2e together with the *times* package. This solves most of the problems, although maths fonts are still bit-mapped, and has the advantage for many authors that it is narrower and therefore one can get more into the paper.

Where LaTeX2e has not yet been installed we would recommend that laboratories install it, together with the appropriate packages which should include *psnfss, psfonts* etc.

It is possible to use slightly older versions of LaTeX/TeX but they must support the New Font Selection Scheme (which has been available since 1989) and Times fonts should be available.

An alternative solution is to use a Type 1 Computer Modern Font but this requires something special to be done at each laboratory to make these fonts available - as well as having to purchase the fonts.

6.2 Tables and Figures

Using relative pathnames when referring to external files makes the source more portable. e.g. "../graphs/jp104-f1.eps" and not the complete path to the file.

Full page width tables and figures can be obtained by using the *figure** and *table** environments.

Graphics can be incorporated using the \includegraphics command available in the *graphics* package of LaTeX2e (See Figure 1). This package will allow re-sizing and rotation of the graphics, which should be in encapsulated postscript format.

Authors using older versions of LaTeX can use the *epsfig* or *psfig* styles.

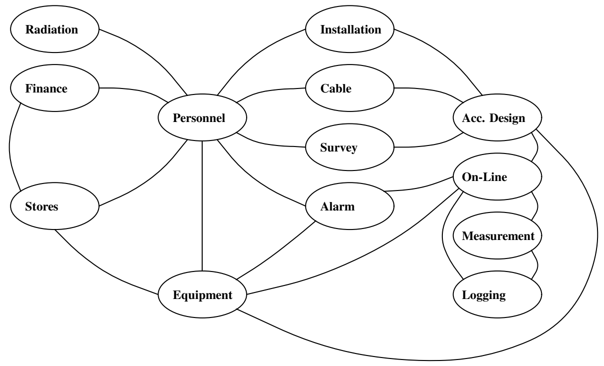

Figure 1: Example of an illustration: this happens to show links amongst databases one would typically find at an accelerator laboratory.

6.3 Lists

New dense list environments *Itemize, Enumerate* and *Description* are available with the *epac96.cls* file. These allow a much more economical use of the space whilst the original versions *itemize, enumerate* and *description* remain available.

6.4 Making the Postscript

The version of *dvips* in use varies across the laboratories but in general one should use the switch which limits the output resolution to 300dpi (-D300) and watch out for the -K option which concerns the removal of document structuring comments from included graphics (these can sometimes confuse the printer and prevent printing).

The postscript file should be made with first-page first order.

6.5 LaTeX Installation

Since the majority of problems with previous conferences seem to have come from LaTeX installations authors and support volunteers (see Section 12) are encouraged to create suitable installations. Information about how to make such an installation can be found in Appendix B of 'The LaTeX Companion'[1], or by looking in the TeX archive sites[2].

7 USING WORD

7.1 General

It is very easy to stray from the required layout and style when using word. Authors should take care over the fonts which are used in the document, including fonts within graphics.

Fonts should be restricted to **ONLY Times, Symbol and Zapf Dingbats**.

7.2 Illustrations and Graphics

Full page width tables and figures can be created by inserting a new section with the appropriate total width (170mm).

Using an illustration (graph, etc.) which has been saved in the encapsulated postscript format (EPS) gives better resolution in the final document because it will scale properly under zooming. Most applications allow one to save a graphic or image in EPS format, so this is recommended.

WORD offers the possibility of saving the illustration within the document file or making a link (or pointer) to the file which is only incorporated when the postscript is generated for printing. The advantages to the author of using a link to the file are a saving in disk space and the illustration can be modified independently. It will be much easier for the editorial team to fix problems with illustrations if they are sent separately, either on the diskette or by ftp.

7.3 Making the Postscript

The choice of printer driver can make a considerable difference to the postscript produced. We have found that the Apple LaserWriter or generic windows postscript printer work well.

The postscript file should have the first-page first, not reverse order.

On a PC

- Use the options from the print command to print to a file using the appropriate driver.
- Check the result by copying the file to a port with a postscript printer attached.

On a MAC

- Select *LaserWriter 8* in the Chooser
- Select *Print* from the File Menu
- Select *File* as the destination and click *Save*
- Select *ASCII, Level 1 Compatible, Format = Postscript, Font Inclusion = no fonts*, give the appropriate filename and click *Save*
- Check the postscript file by printing it using a utility like Apple LaserWriter Utility.

8 ILLUSTRATIONS

8.1 Colour

PDF, like postscript, supports colour, therefore colour will be visible on the screen when viewing the proceedings. Authors should not forget however, that the proceedings will also be published in book form which will be in black and white. The printer will render colours in greyscales and patterns and authors should inspect the black and white version of their printed paper to ensure that information is not lost due to the lack of colour.

8.2 Fill Patterns

For best results authors should use standard half-tones (NOT custom) and solid colour or grey fills.

8.3 Use of EXCEL

EXCEL produces nice graphics relatively easily and allows you to print them to Encapsulated Postscript (EPS) files. However, it is very easy to end up with unnecessarily large files. This has been observed on several occasions and in each case was found to be the result of using *custom* line styles for grid lines on a graph. In one example choosing default (*auto*) grid lines the EPS file reduced from $>$4 Mbytes to 37 kbytes and the visual result was not significantly different.

8.4 Use of ProFit

This gives very unpredictable results when trying to incorporate the EPS files in other documents. In most cases the graphics will have to be scanned and re-introduced into the original. Until the problems are understood, this software is not recommended.

9 USING A SCANNER

9.1 Resolution and Scaling Settings

The aim is to make the smallest possible file with adequate quality. The final printed document will come out at 300 dpi, so it is pointless scanning with more than this. For colour and greyscale this means a maximum of 100 dpi because the printer cannot handle any more.

The final size of the image can also have an effect on the density because scaling down a scanned image in the document will concentrate the information i.e. reducing a 300 dpi image to half of its size will increase the density to 600 dpi.

The way to set up the scanner is:

- Choose line art as default, only use greyscale or colour if absolutely necessary.
- Choose the zoom to produce an image of the correct final size (50% zoom reduces the file size by a factor 4)

- Set the density to 300 for line art or 100 for greyscale and colour.

9.2 Dark Backgrounds

Scanning images with dark backgrounds (like oscilloscope photographs or screen dumps) produces large files. By using an image inverter (usually an option on the scanner program) and using the negative, the information remains as clear but the file is dramatically smaller.

10 PHOTOGRAPHS

Photographs are very difficult to handle. There is an inevitable loss of resolution and the drawing of the image is slow. If fine detail is not an issue, the results can be acceptable. Using compressed JPEG images (with e.g. Adobe Photoshop) is probably the best way to achieve high resolution with minimum file size.

11 SUBMITTING FILES ON DISKETTE OR BY FTP

For the editorial team to be able to repair troublesome contributions, authors are requested to submit ALL of their source files as well as the postscript. Some additional information is also required so that the team can work with the files and this is why there is so much requested on the form which has to accompany the submission.

The printed version of the paper is essential so that the electronic version can be verified and it can also serve as a master for scanning, if all other methods fail.

11.1 Contents

The following should be included

- The source file
- Any special files/macros used to make the document
- Graphics files
- The postscript file (do not send PDF, even if you are capable of producing it)

Delivery of a diskette is the preferred means of submission for a paper. In some cases it may not be possible to write to a diskette and then the ftp server may be used.

11.2 Compression

If the paper is going to work well in the electronic environment it will easily fit on a 1.4 Mbyte diskette without any compression.

If the postscript file will not fit on the diskette, there is probably a problem with one of the graphics or the postscript has not been created correctly.

12 SUPPORT FOR AUTHORS

The EPAC Editorial Board has contacted people in most of the major labs and found *volunteers* to co-ordinate the preparation of papers within their establishments.

The aim is to have the necessary software and expertise available in these labs so that the preferred methods can be used. See the listing on WWW[1] for more information.

12.1 Central Support from EPAC Editorial Board

Where there is no local support individuals can contact John Poole or Christine Petit-Jean-Genaz.

In order to help authors the board will offer a testing service for postscript files so that the techniques can be checked at an early stage.

13 PRODUCTION OF THE PROCEEDINGS

13.1 Before the Conference

Authors will be encouraged to send in either test documents or the final version so that it can be processed electronically.

13.2 At the Conference

The acceptance of contributions will be at least a two stage process: the inventory of files, forms and hard copy will be checked immediately. The postscript files will then be passed to the editorial team who will process them and check for problems. There will then be feedback to the authors to let them know if everything is OK.

If there are problems with a paper, the author will be invited to return to the Proceedings Office so that the problem can be discussed/resolved.

13.3 After the Conference

It will be the policy of the editorial team to inform authors about problems with their submissions. This is seen as a very important step in training authors (and editors) for future conferences.

Only in extreme cases will the hard copy be scanned. If there are problems with graphics which are likely to be very time consuming, the offending figure will be scanned and re-inserted in the source file of the paper as an EPS file.

14 TROUBLE SHOOTING

14.1 Files Too Big

If it is not obvious why a postscript file is so big, try viewing it with a postscript viewer. As the file is displayed, you will see which parts take a long time to draw: these are the likely sources of the problem.

[1] http://www.cern.ch/EPAC/elechelp/supp/supp.html

If the file is drawn quite quickly, but takes a long time to start, it is likely that there are a lot of font definitions which have been included in the postscript. In this case the postscript should be re-generated without any fonts included.

14.2 File will not Print

Again it is useful to review the postscript with a viewer. If there is nothing obvious, the problem is probably coming from a file containing an illustration. Try re-making the postscript with other options or use a different printer driver.

15 REFERENCES

[1] 'The LaTeX Companion', by Goossens, Mittelbach and Samarin, published by Addison-Wesley, 1994.

[2] The Comprehensive TeX Archive Network (CTAN) consists of anonymous ftp servers which can deliver all of the TeX/LaTeX sources. See for example ftp.shsu.edu (Sam Houston State University, Texas) or ftp.tex.ac.uk (Cambridge University, England).