

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN – SL DIVISION

SL-Note-2002-042 MR

POST MORTEM OF THE ELECTRONIC PUBLICATION OF THE EPAC 2002 PROCEEDINGS

P. Le Roux, Ch. Petit-Jean-Genaz, J. Poole, CERN, Geneva, Switzerland

Abstract

The proceedings of the Eighth European Particle Accelerator Conference were the fourth in the series to be published electronically. This report describes the preparations before the conference, the activities at the conference and the work afterwards which was required to produce the CD-ROM and HTML versions of the proceedings. The process was as successful as usual and once again several new features and techniques were employed. This report is the fourth in the series of post mortems and reflects what was new in 2002.

Geneva, Switzerland

December 11, 2002

Contents

1	Introduction	1
2	General Organisation	1
3	Abstract Submission and Publication	1
4	Instructions for Authors and Website	1
5	Templates	1
6	Electronic Submission	1
7	Resources for the Proceedings Office	1
7.1	At LAL before the Conference	1
7.2	At the Conference	1
8	Processing the Files	2
8.1	Activities at the Conference	2
8.2	Post Conference Activities	3
9	Statistics	3
9.1	Manpower	3
9.2	Computer Platforms	3
9.3	Software used by Authors	3
9.4	Failure Rates	3
9.5	Fault Analysis	3
10	Publication	4
10.1	Preparing the Website	4
10.2	Preparing the CD-ROM	4
11	Problems Encountered in 2002	4
11.1	Data Entered by Authors	4
11.2	Untested Software Installation	4
11.3	Name Server	4
11.4	Firewalls	4
11.5	PC Installation	4
11.6	Setting up at the Conference	4
11.7	Power Failure	4
11.8	DFS Failure - files lost	5
11.9	Internet Failure	5
11.10	IT Support	5
12	Improvements for the Future	5
12.1	Editor Training	5
12.2	Pre- Pre-conference Testing	5
12.3	Pre-conference Work	5
12.4	Author Education	5
12.5	Computing Facilities	5
12.6	MathType Plug-in	5
13	Acknowledgements	5

1 Introduction

The proceedings of the Eighth European Particle Accelerator Conference were the fourth in the series to be published electronically. This report describes the new features in 2002, analyses the performance throughout the procedure and presents the relevant statistics. The topics covered include the preparations before the conference, the activities at the conference and the work afterwards which was required to produce the CD-ROM and HTML versions of the proceedings. The paper volumes were produced from the electronic files.

Progress in electronic publishing software and techniques continues advancing and once again we were able to take advantage of this. The Paris conference was the biggest ever in the series and there was an increase in the number of papers published of more than 10% compared to Vienna, whilst the manpower required was comparable to that of 1998 for the Stockholm conference, where there were 18% fewer papers. The final version of the proceedings was available on WWW in just over eight weeks.

This report does not repeat issues discussed in previous post mortems and readers are therefore recommended to refer to the earlier ones [1, 2, 3, 4] in order to obtain a full picture.

2 General Organisation

Local support was provided by staff based at Orsay and they were responsible for the computing infrastructure as well as collaborating in the editorial effort. A comprehensive post mortem of the computing aspects has been published by M. Jouvin [5].

3 Abstract Submission and Publication

A system based on submission into an Oracle database, similar to that which was introduced for Vienna, was used. The submission procedure was basically the same to that of previous years using a web form to gather the data which entered directly into the database.

4 Instructions for Authors and Website

The EPAC website was split between pages located on a server located at LAL (epac02.org), and pages on the central EPAC server located at CERN. The author instructions were relatively simple since they only cover EPAC-specific features and they refer to the common information published by JACoW. An updated version of the Vienna pages was quickly built and the Website was ready for authors by the middle of April 2002.

5 Templates

Following the very successful introduction of the '.dot' Word templates at PAC'01, updated versions were introduced for EPAC'02. A total of 20 versions of the Word template were necessary - five each in '.doc' and '.dot' for A4 and US letter formats (Word 97, 2000 and 2002 for the PC and 98 and 2001 for Mac). A further template for each paper size was prepared for L^AT_EX.

6 Electronic Submission

The main change introduced for this year was the elimination of submission on diskette. All authors were required to submit via the internet, even if this was at the conference. Electronic files were sent to a file server, located in Paris, and the meta data (software types, platforms, etc.) were stored in Oracle at CERN. The philosophy behind this being that the weakest network link is between the conference venue and the local institute and therefore the link across a high capacity internet backbone from there to CERN would be transparent. Furthermore, the data volumes going to and from Oracle are trivial.

It became obvious some time before the conference that the resources in Paris for post conference work would be extremely limited and it was therefore decided to make the deadline for submissions earlier in order to ensure that the maximum amount of work could be completed before the end of the conference. The deadline was set at 24:00 CET on the Wednesday before the conference - just two working days ahead of that for preceding EPACs.

As usual, authors were requested to bring a hard copy of their paper to the conference – essential for the quality checks.

7 Resources for the Proceedings Office

7.1 At LAL before the Conference

A facility for about 8 people to process papers ahead of the conference was set up at LAL. The full system was implemented, including dotting and filing of processed papers. In total, 12 people worked on files ahead of the conference (two students were only present for a very short time and worked as a pair) and from the remaining people, some were only working part-time. The networking and software installation was the same as planned for the conference centre but with fewer machines (around 8 PC and 2 Macs).

7.2 At the Conference

The hardware requirements for 2002 were based on EPAC'2k and PAC'01 experience where about 12% of the papers were prepared on Macintosh. There were two Macintosh and 16 PC's in the processing office and a further 4 PC's in the paper reception office. The computers were linked by a local intranet and the files networked across

all platforms. A very efficient printing service was implemented using 2 colour and 2 monochrome printers, shared between the two offices.

The software inventory was as follows:

PC

- Office XP
- Internet Explorer 6
- Netscape 4.7, telnet and LeechFTP
- Acrobat5 with PitStop plugin
- MikTeX 2.1 and WinEdt
- Srip32
- PowerArchiver (WinZip equivalent)
- SSH client 3.1
- F-Secure (anti-virus)

Equivalent software was installed on the Macintosh machines.

As usual there was strong support from the JACoW collaboration (APAC, CYCLOTRONS, ICALEPCS, LINAC and PAC) for the proceedings and processing offices. In all there were about 16 people assigned to the processing office with part time assistance from a few more.

8 Processing the Files

Figure 1 shows the number of files which were processed per day in Paris. From this one can see that when all is working correctly (e.g. Friday 31 May) it is possible for one person to process around 40 files per day. The gross variations from day to day are explained by the technical problems which are described below. Towards the end of the conference week there were virtually no new files for processing and the effort was switched to quality control checking. About 50% of the submitted files had the quality assurance checks made on them by the end of the conference week.

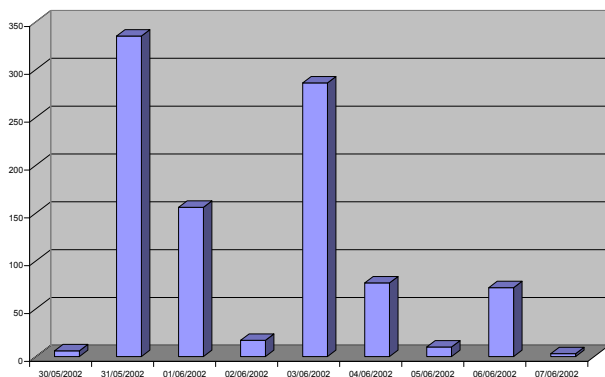


Figure 1: Number of Papers Processed per day in Paris

8.1 Activities at the Conference

The main aim of the activity at the conference was author feedback achieved through the now familiar boards with

coloured stickers. One person worked full time on keeping the boards up-to-date and filing the papers. However, as was mentioned above, it was possible to switch from processing activities to quality control checks from Tuesday onwards.

A new method of processing for EPAC was introduced in Paris, following the experience at PAC 2001. All processing was tracked in the database and there was no paper submission sheet or paper processing sheet - everything was stored in Oracle. Editors had an interface to the database through which they were assigned papers to work on and they were able to record their comments and status of the paper in the database.

On the PC, editors were able to use a Visual Basic application to retrieve the files from the server for processing and once the process was completed the files were automatically archived. This system had the advantage that it eliminated the mis-placement of files. The application also automatically created versions of papers when files were re-submitted or simply re-processed. Unfortunately, this application was not available for Macintosh. Once an editor had been allocated a paper from the database it was not possible for an author to re-submit - until this time any re-submissions would over-write previous versions on the file server. A manual intervention was necessary to allow re-submissions.

Processing the files involved distilling the postscript and then verifying:

- the fonts
- the margins
- overall quality/performance

A web page¹ containing instructions for editors was created. The Acrobat software had been set up with the appropriate defaults for distilling and the page size was modified to JACoW standards using an Action List in PitStop to resize the media box.

For the cases where an editor had to re-work a paper, WORD was set up to use PDFMaker, again using the appropriate distiller defaults but in addition, all automatic generation of hyper links was turned off.

A full MikTeX installation was made on several PC's and it was also possible to use the WinEdt package to simplify use of L^AT_EX. There were some problems with the automatic deployment of MikTeX to editors machines but it was possible to make a full installation from the internet within a couple of hours. The latest version of MikTeX (used at LINAC 2002 in August) uses Type 1 fonts by default and therefore no modification to the standard installation is necessary.

The tracing system allowed the staff in the paper reception office to deal with many of the authors coming in because they had received a red dot. It was usually clear from the editor's comments what was wrong and what was

¹<http://cern.ch/JACoW/organisers/Processing/Ed-Instructions-Paris.html>

required of the author. This transferred some of the effort from the editors to the reception office where there was some additional capacity in 2002 because there was no correction of titles/authors etc. in the database.

8.2 Post Conference Activities

The post conference activities are principally fixing the problem files, careful checking, page numbering and making the indices, table of contents and other pages for the wrapper (photos, copyright, prize winners, conference organisation, list of participants, titles and separators etc.). In the problem cases, authors were contacted by E-mail and reminded that they had two weeks after the conference in which to resubmit.

Once all of the files have been processed and deemed 'OK' they have to be scrutinised very carefully during the quality assurance checks. Instructions for editors on this phase of the work were also made available through a web page². The files for the Web are prepared first and then the CD-ROM versions and finally the paper version is prepared.

9 Statistics

9.1 Manpower

The manpower used in 2002 was less than in 2000 but comparable to that in 1998. More effort was required in the development of the Oracle based system but improvements in software and efficient processing have helped to keep the numbers down. With more than 10% more papers, less than 5% more manpower was used. The distribution of effort is given in Table 1.

Table 1: Manpower Resources for EPAC Proceedings in Man-weeks

	2000	2002
R & D	2	4
Planning	2	2
Build/maintain WWW pages	2	1
Author documentation	1	1
Server setup	2	2
Abstract Processing	5	3
Infrastructure	2	2
Pre-conference Processing	0	3
Processing at the Conference	20	15
Post Conf - local	26	4
Post Conf - CERN	8	14
Total	70	50

9.2 Computer Platforms

There was a further slight shift away from Macintosh platforms (22% in 98 and 12% in 2000) down to about 10% in 2002 and the number of Unix users remained constant.

²<http://cern.ch/JACoW/organisers/Processing/QA-Instructions-Paris.html>

9.3 Software used by Authors

The distribution of software packages used by authors remains dominated by Microsoft Word and L^AT_EX (see Figure 2). The percentage of L^AT_EX users has remained constant whilst non-preferred software is restricted to one or two papers – probably from the same author who is using FrameMaker.

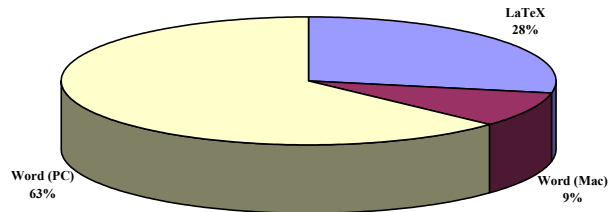


Figure 2: Software used for the preparation of EPAC2002 papers

9.4 Failure Rates

The overall quality of papers submitted continues to improve and editors are able to apply higher standards each year. This is also possible because the software has improved and it is much easier to fix many problems either using PitStop or by reworking the original document. It has been a policy for some years to invite authors to proof read all papers where the original document has been re-used to make a new PDF. This process is necessary because the author's installation may differ from that at the conference and there may be font or plugin problems which are not immediately obvious to the editor.

In Paris 27% of all papers were marked for proof reading but there were only 14% of the papers which had real problems.

9.5 Fault Analysis

Table 2: Analysis of problems encountered with papers submitted to EPAC 2002

	Percentage of all papers
No PostScript	8
Format problems	7
Type 3 fonts (L ^A T _E X)	5
Unusable files	5
Multiple problems	3
Font problems	3
A4 printed on US paper	2
Slow graphics	2
Everything possible wrong	1

An interesting feature for the Paris conference was the large number of authors submitting PDF files. In many cases they had only submitted PDF and some of these were

processed by the editors and given green dots. This was a mistake and was the result of insufficient training for the editors. It was found that these files often caused problems later in the assembly of the final version of the proceedings. This is quite understandable because the PDF has not been prepared with the appropriate parameters like font inclusion, graphics compression and so on. The fraction of such papers was between 5 and 10%.

10 Publication

10.1 Preparing the Website

The website was built in much the same way as in previous years but using Visual Basic scripts to produce the table of contents, indices etc. The only major problem associated with this part of the process concerns the handling of special characters which have to be converted to HTML.

10.2 Preparing the CD-ROM

The files on the CD-ROM include all of the papers, the table of contents, the author index and the various other parts of the wrapper. The CD is a copy of the web version. Following an agreement at the JACoW Team Meeting in Thoiry [6] no Acrobat software was included on the CD and the Proceedings were designed to be viewed using a web browser. The cover graphics and instruction booklet for 2002 were much simpler to prepare because there was no need to describe the installation of the Acrobat software.

Page numbers and the conference banner were inserted in the PDF files using a Visual Basic Script which retrieved paper code and page number from the database. The script then simulated keyboard input by sending commands to Acrobat and Impress Pro. The latter is an Acrobat plug-in providing features for embedding formatted text. The principle of this system follows that used previously[2] where a script was used for the hidden fields etc.

11 Problems Encountered in 2002

11.1 Data Entered by Authors

The authors were requested to enter their names, affiliations and addresses when submitting their abstracts but the web form allowed them the freedom to introduce their own versions as well as picking from a list. This led to a huge amount of work when trying to prepare the indexes because the same person at the same institute might appear with four or five different different versions of his details.

11.2 Untested Software Installation

Due to the unavailability at LAL of personnel experienced with the conference processing software, no testing had been done when the team of editors assembled in Paris. Whilst the actual installation was rather fast, the testing revealed several problems and it was some time before the correct configuration was found. As a consequence it was

not possible to do any processing on the first day of the pre-conference period.

11.3 Name Server

This problem is explained in M. Jouvin's report [5] and the effect was that authors were not able to submit their papers during the final period leading to and following the deadline. A problem with resolving computer names to their IP addresses caused most requests to time out leaving the authors and editors unable to transfer data across the network. The problem took some time to understand but in the interim editors were able to use the IP address and by-pass the faulty server. This obviously is a sensitive area because the author is probably already under stress (because he is on the deadline) and difficulties beyond his own control are very frustrating. Authors struggled with this system until the weekend with around 300 papers being submitted between the deadline and the conference start.

11.4 Firewalls

People in some institutes were unable to submit files because their computers were sitting behind a firewall which would not allow the level of communication required for paper submission. In these cases authors had to bring copies of their files to the conference and make the submission there.

11.5 PC Installation

On arrival at the Conference centre on the Sunday it was found that the additional computers for the proceedings team would not function correctly because the revision level of the network interface was not the same as those used at LAL. This prevented the automatic installation of the software from working correctly and the leasing company had to come and exchange the network interfaces.

11.6 Setting up at the Conference

It was not possible to set up the processing office infrastructure before the Sunday because the conference centre had not been booked for the Saturday. Installation of the processing infrastructure was very fast but problem solving was much longer and only a handful of papers were processed on Sunday. Although the published opening time for the paper reception office was on the Sunday and all of the staff were available, delegates were not permitted to enter the premises and therefore no papers were received.

11.7 Power Failure

On the Monday everything was functioning reasonably well apart from some residual problems with the L^AT_EX installation. Unfortunately, on the Tuesday morning there was a power failure in the proceedings office which took most of the morning to repair.

11.8 DFS Failure - files lost

A problem with the system software installation (see M. Jouvin [5]) on the file server caused a number of processed files to be lost. About 30-40 papers had to be re-processed, but it was quite a difficult task to identify which files had been destroyed.

11.9 Internet Failure

On the Wednesday of the conference the ADSL system in all of Paris failed, therefore it was not possible to process any files during the day.

11.10 IT Support

Problems with the informatics infrastructure sometimes took too long to resolve because the communication between the editors and the IT support personnel was not optimal.

12 Improvements for the Future

12.1 Editor Training

The amount of training was insufficient in 2002 - although the introduction of the web page was an improvement over previous years. Had there been fewer technical problems, more time could have been devoted to training editors. Many of the editors at the conference have several years experience and require very little assistance but points which should have been emphasised in Paris were:

- Author's PDF files should be ignored.
- The finer points of the style of the template are desirable but not essential.
- Section numbering should have been checked.
- Editors should record clear and explicit comments in the database.

Following the JACoW Team Meeting in Berkeley (November 2002), the JACoW templates have been modified to remove section numbering - this will solve the problems encountered in Paris, where numbering was frequently incorrect and inconsistent. At the next few conferences however, one should expect many papers which have been prepared using the old templates and a policy statement should be made concerning whether or not to remove numbers.

At least one training session should be planned before processing starts - new editors must receive detailed instruction and experienced editors need to know the conference policy and what is new.

12.2 Pre- Pre-conference Testing

A complete editing installation should be fully tested before the team arrives. In Paris around two man-weeks of effort was lost because the team could not work. There are always a sufficient number of early paper submissions for full testing of the system to be made but the local team should allocate resources for doing this. It is clear that

rented computers will only arrive 'just in time' but testing on similar machines in the same environment should be sufficient.

12.3 Pre-conference Work

An efficient pre-conference processing session would allow quality assurance (QA) work to be completed on all of the files before the end of the conference. More thorough initial processing would also reduce the number of green dotted papers which fail QA.

12.4 Author Education

It was clear in Paris that authors think that they are helping by sending PDF files even though the instructions clearly request only PostScript and source files. This point requires some emphasis and explanation in the author instructions.

12.5 Computing Facilities

With a large number editors present at the conference the effect of any system failure is greatly magnified - one day lost is equivalent to about 10% of the total effort required for proceedings production. High priority should therefore be given to the reliability of the computing infrastructure. Backup solutions should be envisaged (e.g. an IDSN line in Paris would have allowed editors to work on the Wednesday when ADSL went down).

12.6 MathType Plug-in

Many authors use MathType to edit equations in Word but if this has not been installed on the editor's machine he cannot fix problems in a file which has used it. Two possible solutions exist: install MathType using a license from the institute or use the 30 day trial version for the conference.

13 Acknowledgements

The EPAC editorial board is greatly indebted to the following group of people who worked very hard on the electronic processing of the papers in Paris.

Ivan Andrian (Trieste)
Joseph Chew (LBNL)
Jinhyuk Choi (POSTECH)
Jan Chrin (PSI)
Martin Comyn (TRIUMF)
Cathy Eyberger (ANL)
Charlie Horak (ORNL)
Leif Liljeby (MSL)
Peter Lucas (FNAL)
Michaela Marx (DESY)
Toshiya Tanabe (RIKEN)
Shoreh Tavakoli (LAL)
Sara Webber (FNAL)

The processing staff could not have completed their task without the support of the other volunteers in the paper reception office:

Marlen Bugmann (PSI)
Maura Chatwell (SLAC)
Mariarita Ferrazza (Frascati)
Sheila Poole (Thoiry)
Pina Possanza (Frascati)
Martine Truchet (CERN)

References

- [1] J. Poole, “Post Mortem of the Electronic Publication of the EPAC’96 Proceedings”, CERN-SL-Note 96-68 DI, November 1996.
- [2] L. Liljeby and J. Poole, “Post Mortem of the Electronic Publication of the EPAC’98 Proceedings”, CERN-SL-Note 99-015 DI, February 1999.
- [3] S. Webber, P. Lucas, M. Arena, “Post Mortem of the Electronic Publication of the PAC 2001 Proceedings”, <http://cern.ch/JACoW/organisers/docs/PAC01-PM.pdf>, January 2002.
- [4] P. Le Roux and J. Poole, “Post Mortem of the Electronic Publication of the EPAC2000 Proceedings”, CERN-SL-Note 2001-007 DI Revision 1, April 2001.
- [5] M. Jouvin, “EPAC 2002 Computing Post-Mortem”, <http://cern.ch/JACoW/organisers/docs/computer-pm-2002.pdf>, August 2002.
- [6] J. Poole “Summary of the JACoW Team Meeting, Thoiry, France”, http://cern.ch/JACoW/CERN_TM-2002/Talks/JACoW-CERN-Feb2002-v0.pdf