THE JACoW FILE SERVER AT PSI

J. Chrin, Paul Scherrer Institut, Villigen, Switzerland

Abstract

A networked file server for the upload of contributions to proceedings has been made available at the Paul Scherrer Institute (PSI) to the JACoW Collaboration. Its availability extends the editorial facility within the SPMS to a selfcontained utility and alleviates conferences with the need to allocate local resources for its realization. A description of the JACoW File Server infrastructure is presented together with a report of its first usage.

MOTIVATION

The JACoW Scientific Programme Management System (SPMS) coupled with the JACoW Central Repository of author profiles, build upon the Oracle database, provide a rich set of functionality to conference organizers for managing the many fundamental aspects such events necessitate [1]. Among the resources provided are the means to administer contributions to a conference programme. These range from the management of the initial abstracts and susbsequent paper submissions, to facilitating the editorial process. The latter concludes with JACoW PDF/A compliant manuscripts that are subsequently used by the JACoW Proceedings Script Package (JPSP) for the production of the final proceedings volume [2]. An essential ingredient for the editorial process, that rests outside the domain of the SPMS, is the requirement for a separate web and file server for the upload, download and storage of author's papers for the editorial process. The file server, unlike the SPMS, is not provided by the JACoW Regional Support Centers; rather it's realization rests with the hosting conference. The event's Local Organizing Committee (LOC) may either choose to take direct responsibility for providing the server and to equip it with the necessary software or, if well placed to do so, channel the task to the hosting institute's IT department. In either case, particular care must be taken in providing a well-managed and reliable web/file server with high-speed connectivity. Energy can also be expended in iterating with, and convincing IT, of the specific needs of such a web/file server, in particular the necessity to execute CGI scripts. For smaller conferences with limited resources and/or personnel, the implementation of a web/file server can consequently constitute a large overhead.

The provision to provide a file server to the JACoW collaboration will therefore alleviate conferences of the associated administrative complexity. In this respect, a web/file server has been set up at Paul Scherrer Institute (PSI), for use by JACoW conferences. Its availability further serves to extend the functionality of the editorial process within the SPMS to a self-contained utility. This paper reports on the first usage of the the JACoW Web/File Server at PSI, describing its directory structure, its accessibility, its management, and the straightforward procedure to instantiate a conference specific file server instance.

FILE SERVER REQUIREMENTS AND PSI WEB SERVICES

The file server is where all files submitted by the authors, i.e. LaTex and Microsoft Word source files, PostScript and Portable Document Format (PDF) files, supporting files such as figures which may be in various formats, and Microsoft PowerPoint or PDF presentations, are stored. It is also from where manuscripts are retrieved for editing and the destination of files worked upon by the editorial staff. Access from the SPMS to the designated file server for the electronic retrieval and submission of files is done through a web server that hosts and executes the JACoW Upload/Download Scripts (JUDS) [3]. This is the only route by which papers may be submitted to the file server. A prerequisite is a reliable high-speed network connectivity, sufficient disk space, disk redundancy and daily backup.

The IT department at PSI provides a central web service for the benefit of groups and projects, allowing them to publish their information for internet purposes. Web directories are mapped onto central AFS (Andrew File System) files servers, which can be accessed, with sufficient privilege, through an AFS client on UNIX, Windows and Macintosh platforms. The arrangement is particularly suitable for distributed project members.

The PSI Web Service also supports CGI scripting. All output files which may be created by such scripts, e.g. Perl scripts, are stored within or below the purposefully allocated "cgi-bin" directory. A daily back-up of the web server is automatically provided. RAID storage technology combines multiple disk drive components into a logical unit. Data are distributed and replicated across several drives providing redundancy to the extent that three simultaneous disk failures would have to occur before data availability is compromised. Several physical web servers, running Apache under Scientific Linux 5, provide load balancing by serving the same central web pages.

The services provided by the PSI Central Web Services are thus ideally suited to the requirements imposed by the JACoW File Server.

/afs/psi.ch/project/jacow/www/cgi-bin/conf/*{confName*} /afs/psi.ch/project/jacow/www/conf/{confName}/papers /afs/psi.ch/project/jacow/www/htaccess-templates 'afs/psi.ch/project/jacow/www/cgi-bin 'afs/psi.ch/project/jacow/www/juds 'afs/psi.ch/project/jacow/Backup 'afs/psi.ch/project/jacow/www/ **Physical AFS Directory** http://jacow.web.psi.ch/conf//conf/Name//papers http://jacow.web.psi.ch/cgi/conf/(confName) http://jacow.web.psi.ch/htaccess-templates http://jacow.web.psi.ch/juds http://jacow.web.psi.ch/cgi http://jacow.web.psi.ch/ URL Access Templates Backup directory Generic JUDS Conf. Papers Conf. JUDS Web Base CGI Base Scope Ξ 0 $\widehat{\mathbb{C}}$ 4 $\widehat{\mathcal{O}}$ 9 6

THE JACoW FILE SERVER

A web directory in the AFS domain has been created for the JACoW project. A limited number of distributed project managers have been granted read/write privileges to the AFS JACoW project area. The principle URL and physical AFS directories are listed from (1) to (7) in Table 1. Their purpose is clarified in the following.

- (1) The base address for web pages
- (2) The base address from where, or below which, CGI scripts may be executed
- (3) The location of the generic JUDS scripts, modified to the local server configuration
- (4) The location of the conference specific JUDS scripts
- (5) The root directory for the storage of the conference specific papers
- (6) The location of hypertext access (htaccess) template files that control the internet access to directories
- (7) The location of the complete JACoW project files backed-up from the previous day.

The URLs listed in grey in Table 1 are only accessible from within the PSI intranet, unless public access rights are otherwise granted. Such access rights are, for example, granted to the conference papers directory prior to the prepress release of the proceedings.

Use of JUDS scripts require the latest version of the Perl programming language (v. 5) to be installed on the web server. Similarly the ability of the updated JUDS scripts [4] to automatically distill author's PostScript and PDF files, with the required JACoW settings [5], place their own requirements on the Linux Web Server. These include the use of a recent Ghostscript package (v. 9) and the installation of the Microsoft TrueType core fonts to ensure that all fonts may be embedded in the Ghostscript generated PDFs.

With the required software in place and the autodistilling activated, the web server is able to generate JACoW PDF/A compliant papers that are fit for publication, provided of course that the author has faithfully complied to the JACoW template!

Setting up a Conference Instance

The procedure for setting up a file server instance for a particular conference has been conveniently reduced to the execution of a script. The latest JUDS scripts will already have been transferred from the JUDS repository [3] to the JACoW AFS project area. While the JUDS package consists of several scripts for the upload, download and distilling of papers, there is only one configuration file that requires customization. Among the parameters to be configured [6] is the path to the directory on the file server where the files are to be stored, and the path to a directory that stores the intermediate files in the auto-distilling process. The local JUDS configuration file is thus modified to meet with the corresponding settings of the PSI Web/File Server, without yet making reference to a specific conference.

System Parameter Scope	Function	Example Entry
Web Configuration	Secret Pass-phrase for Paper Upload	"D8Y6u5h1Zp890YAgT3P0QjlNms90wi"
Web Configuration	Secret Pass-phrase for Paper Download	"S20o0098dgWlH4cHdP9o51RsDGUxpZ"
Web Configuration	Author File Upload Script	http://jacow.web.psi.ch/cgi-bin/conf//conf/ame//JACoW.file.upload
Web Configuration	Editor Download Script	http://jacow.web.psi.ch/cgi-bin/conf/ <i>\confName\\</i> editor.zipdownload
User	Proceedings TOC Base URL	http://jacow.web.psi.ch/conf/ <i>\confName</i> //papers

Table 2: SPMS System Parameters for the JACoW File

Server

A file server instance for a conference is then created by executing a simple script in the AFS base directory:

./confSetup.sh <*confName*>.

The script automatically creates the necessary directories for the specified conference and installs the conferencespecific JUDS scripts into the conference CGI directory. The only manual intervention required is the choice of passwords for the upload/download scripts. These same passwords must also be entered into the SPMS for the activation of the scripts (see Table 2). Execution of the above script also grants write privilege to the host Apache web user to the conference paper directory.

Finally, the processed papers that constitute the prepress release, as directed to from the SPMS through the "Proceedings Table of Contents (TOC) Base URL" (see Table 2), may be made accessible to the public by providing the necessary hypertext access (htaccess) rights to the hosting directory. Specifically the "htaccess-public.acl" file is copied from the htaccess template directory to the conference paper directory. A simple script is provided to facilitate this task:

 $./confPapersAllowPublicWebAccess.sh<\!<\!\!confName>$

or conversely to undo it:

./confPapersDenyPublicAccess.sh < confName>

The actualization of a conference file server instance through tested scripts, with well-defined procedures, simplifies their management and reduces the risk of inadvertent erroneous activity on the file server.

SPMS System Paramaters for the File Server

The SPMS System Parameters pertaining to the File Server are listed in Table 2. These may be changed by any SPMS user with sufficient privilege. Conferences adopting the JACoW File Server will have had these parameters already set by the File Server Administrator, a recently added 'Functional Role' within the SPMS that grants access to the relevant SPMS System Parameters.

Disk Space Requirements

The five conference instances activated in 2012 required a total of \sim 13 GB disk space (Table 3). This more than doubles up to \sim 28 GB usage on the File Server when data backed-up from the previous day are included. The maximum quota granted to a project without incurring additional costs is 200 GBytes. The present disk quota is set to 50 GBytes, but this may be increased on demand. Once the proceedings volume for a conference has been published on the JACoW website, the corresponding files on the JACoW File Server should no longer be required and may be compressed before eventual deletion. In this respect, the file server acts as a data buffer for which the required disk space will reach a plateau as a steady state is reached between the creation of new conference instances and the removal of files from conferences past.

Table 3: Disk Requirements

	No.	Disk Usage
Conference	Papers	(GBytes)
FLS12	46	0.4
BIW12	99	2.5
ICAP12	101	2.9
LINAC12	340	6.0
ECRIS12	55	0.8

Monitoring the Web Server

The recent introduction of the automatic distilling of author's submitted PostScript and PDF files to the upload scripts [4] initiated discussions on the policies in place for the monitoring of the PSI Web Server. In particular, questions were raised regarding the impact on the CPU load, the size of the server's /tmp partition which is used for the temporary storage of intermediate files, and the possibilities of zombie processes in failed distill operations.

In this respect, it is gratifying to learn that the Web Servers themselves are under the scrutinous watch of Nagios, an industry standard IT infrastructure monitoring system. The size of /tmp partition is 1.9 GBytes and is monitored by Nagios at intervals of 2 minutes. Files older than 240 hours are deleted. Since the revised upload script, on completion of the distilling process, now deletes the intermediate temporary files it creates, its load on the /tmp partition is minimized. With regards to the CPU load and possible zombie processes, IT were asked to keep a particular vigilance in the days prior to and during the more demanding LINAC12 conference. They did not, however, report any anomalies. Indeed the verdict from IT was that the JACoW conferences did not impact the Web Servers beyond their accustomed load.

CONCLUSION

A JACoW File Server at PSI extends the functionality of the editorial process within the SPMS to a self-contained utility. Its availability shelters JACoW Conferences with limited resources from the logistics of establishing their own individual file servers, reducing their overhead in setting up the editorial procedure. The five conferences in 2012 to have taken advantage of this possibility have endorsed its suitability for both small (\sim 100 papers) and medium sized (\sim 350 papers) events. The benefit of the JACoW File Server to the community will be appraised by both its effectiveness and its relative inconspicuousness!

ACKNOWLEDGEMENTS

The author acknowledges the excellent support provided by the IT department at PSI and particularly wishes to thank Peter Hüsser for generously responding to specific requests related to the use of the PSI web server and Aachim Gsell for setting up AFS accounts for external PSI users.

REFERENCES

- C. Petit-Jean-Genaz, "Organizing IPAC (and other JACoW Events) with SPMS an On-Going Saga ", Proceedings of the 2011 JACoW Team Meeting, 12-17 Nov. 2011, SINAP, Shanghai, China, pp. 31-69.
- [2] V.R.W. Schaa, "JPSP JACoW Proceedings Scripts Package for SPMS and InDiCo Conferences", http://www-esh.fnal.gov/SPMS/Download_ Files/JPSP-manual-9.2.pdf
- [3] I. Andrian, "SPMS Up/Download Scripts", 2004 JACoW Team Meeting, 7-10 Nov. 2004, ORNL, Knoxville, TN, USA; http://sourceforge.net/projects/ jacow-spms/files/
- [4] R. Müller, "Revised JACoW Upload/Download Scripts", 2012 JACoW Team Meeting, 6-9 Nov. 2012, IFIC, Valencia, Spain, Paper ID: WEDA1, *These Proceedings*.
- [5] V.R.W. Schaa, "Submission of Contributions to Proceedings: PostScript to PDF", 2012 JACoW Team Meeting, 6-9 Nov. 2012, IFIC, Valencia, Spain, Paper ID: WEDA2, *These Proceedings*.
- [6] M. Comyn, "File Upload/Download Scripts", 2009 JACoW Team Meeting, 1-5 Nov. 2009, DESY, Hamburg, Germany.