

# CERN Printing Infrastructure

**Rafal Otto, Juraj Sucik**

CERN, Geneva, Switzerland

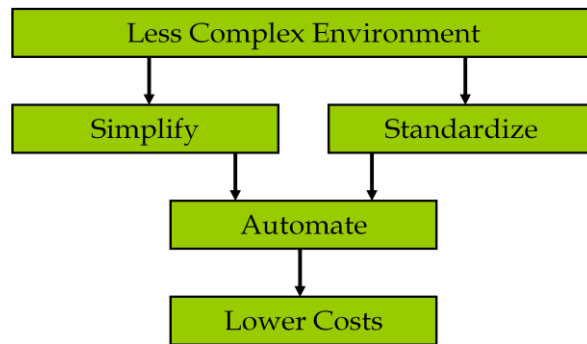
E-mail: [Rafal.Otto@cern.ch](mailto:Rafal.Otto@cern.ch), [Juraj.Sucik@cern.ch](mailto:Juraj.Sucik@cern.ch)

**Abstract.** For many years CERN had a very sophisticated print server infrastructure [13] which supported several different protocols (AppleTalk, IPX and TCP/IP) and many different printing standards. Today's situation differs a lot: we have a much more homogenous network infrastructure, where TCP/IP is used everywhere and we have less printer models, which almost all work using current standards (i.e. they all provide PostScript drivers). This change gave us the possibility to review the printing architecture aiming at simplifying the infrastructure in order to achieve full automation of the service. The new infrastructure offers both: LPD service exposing print queues to Linux and Mac OS X computers and native printing for Windows based clients. The printer driver distribution is automatic and native on Windows and automated by custom mechanisms on Linux, where the appropriate Foomatic drivers are configured. Also the process of printer registration and queue creation is completely automated following the printer registration in the network database. At the end of 2006 we have moved all (~1200) CERN printers and all users' connections at CERN to the new service. This paper will describe the new architecture and summarize the process of migration.

## 1. Introduction

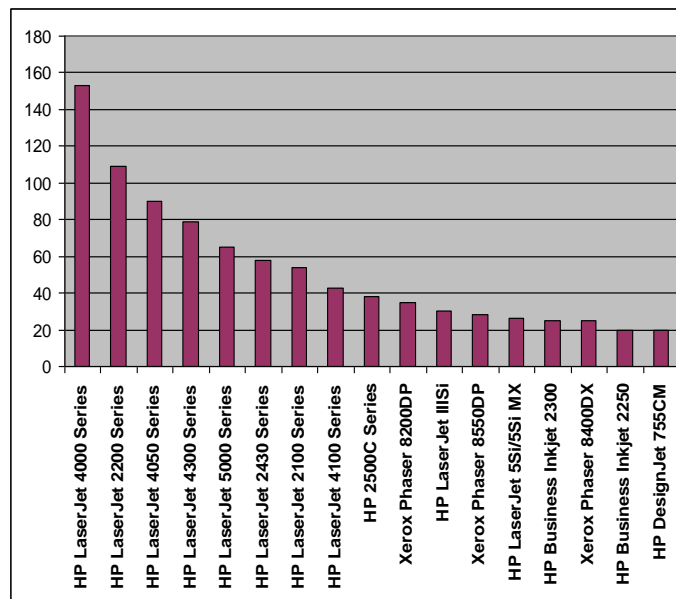
Several years ago, the CERN computing infrastructure was very heterogeneous. There were many devices communicating using many different network protocols. There was as well a variety of workstations running different operating systems which used a different approach to printing and printer driver distribution. Therefore, the CERN printing infrastructure designed and implemented in 1998 by Reguero, et al. [13] had to be very complicated. Print Servers had to communicate to printers using multiple protocols like AppleTalk, IPX/SPX, TCP/IP and many workstations including Windows 95/NT (later Windows 2000 and XP), a variety of UNIX systems from Sun, HP, DEC as well as PCs Linux distributions.

The situation has changed a lot since that time. Currently at CERN we have a population of almost 1200 printers where all communicate using TCP/IP. The workstations are as well in a great majority running Windows XP, Windows Vista, SLC3 (Scientific Linux CERN [6]), SLC4 or Mac OS X as an operating system. Having such a homogenous environment we could base our system on standard products and hence reduce custom developments to a minimum. As a result, the printing infrastructure at CERN could be simplified and standardized which resulted in higher level of automation of the printer registration process and less manpower needed to maintain the system. Therefore, the costs could be lowered significantly.



## 2. Environment

The printer population at CERN is quite big and consists of almost 1200 printers. Among those devices we have 88 different models, which seems a big number; however the 7 most popular models cover 51% of printers. 51 out of 88 models are represented by less than 5 devices. Those numbers show that printer population is already quite homogenous. There is still some work to be done in reducing the number of models by cutting the tail of the distribution and phasing out the least popular models. This is an ongoing activity taken by the User and Document Services group of the CERN IT department.



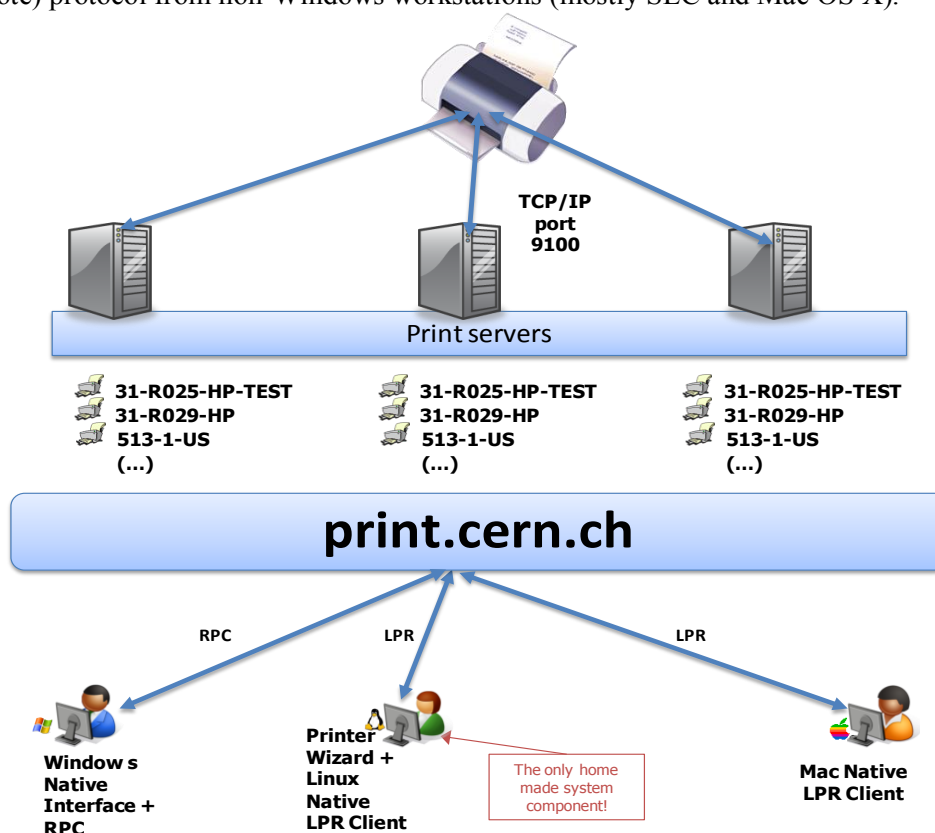
The previous printing system was based on the concept of formatting the print jobs on the server side. Print servers could accept a wide variety of job formats (including compressed files) and were able to process them, format with the appropriate driver and print properly on any printer device. The only operating system, which was sending jobs already formatted, was Windows at that time. Before the new printing system was designed, we analyzed the logs of the previous system. We checked how important it was for the new system to be capable to process the formatting on the print server side and to be able to accept such wide range of print job formats. As a result of the analysis, we found that:

- 77% of jobs come from Windows workstations,
- 20% of jobs come in the PostScript format,
- 1% of jobs come in the PCL format,
- 1% of jobs come in the plain text,
- 1% of jobs cover rest of the formats.

The conclusion was that 99% of jobs coming to the print servers were ready for printing. Therefore, the capability to format jobs on the server side is not used much anyway.

### 3. Architecture

The new CERN printing infrastructure is based on Microsoft Windows print servers, which accept print jobs using RPC (Remote Procedure Call) from Windows workstations and using the LPR (Line Printer Remote) protocol from non-Windows workstations (mostly SLC and Mac OS X).



#### 3.1. Print Servers

The print servers are running the Microsoft Windows 2003 R2 [12] edition of the operating system, which provides additional printing features, mainly the new print management console which simplifies operational tasks performed on the print queues.

All three print servers are configured in the same way and they all have the full set of print queues (~1200) configured. There is a scheduled process running periodically, which ensures that the configuration of the corresponding queues on each print server is the same. Then, we have a DNS sub-domain dedicated for printing called *print.cern.ch*. Each printer queue has a corresponding DNS alias (CNAME record) registered in the *print.cern.ch* zone, which points to one of the three print servers. For example a queue called 31-R029-HP is configured on each print server (CERNPRINT01, 02 and 03) and 31-R029-HP.print.cern.ch resolves to the CERNPRINT02 print server. All workstations are using then the aliases in *print.cern.ch* DNS domain when connecting to the print queue. In this way, in case of failure of any of the print servers, we can quickly redirect all traffic for a given printer to a different print server simply by changing its CNAMEs in DNS.

This failover solution was in fact taken from the previous system designed by Reguero, et al. [13]. The only difference is in the way the DNS updates are done. Since the new printing infrastructure is in production, the *print.cern.ch* DNS domain has been moved to DDNS [15] and all updates are incremental. It is done using the Net::DNS Perl class using TSIG keys for authentication [14]. This

speeds up all the DNS changes and makes them easier to automate. The DDNS CERN implementation is well described by Bahyl and Garfield in [1].

Microsoft servers provide as well the standard clustering mechanism for the print spooler. This has been seriously considered, however the conclusion was, that it would introduce an additional level of complexity and an additional point of failure. We are as well ready to accept short interruptions of the printing service needed for reconfiguring the DNS in case of a print server failure.

The last important component of the server infrastructure is the publication of the print queues. At the same moment as the new DNS alias is created, the printer queue configured on the server to which the DNS record resolves is published in Active Directory. Therefore the printer can be discovered using the LDAP protocol, which is the default way Windows clients search for printers.

### 3.2. Print Job Formatting

Each of the print servers accepts two types of jobs using two different methods.

Firstly, it receives jobs from Windows clients by accepting Remote Procedure Call (RPC) connections from client computers. The component of the local print spooler of the client PC, called Remote Print Provider opens the connection to the server spooler and sends the job over the network. Jobs that are submitted are already pre-formatted by the appropriate printer driver, that is installed on the client computer, to the Enhanced Metafile (EMF) format and then formatted in the printer-understandable format by the print spooler on the print server. Windows network printing is documented in detail by Microsoft in [9].

Secondly, each print server has Print Services for UNIX service installed which accepts jobs from non-Windows clients sent using the LPD (Line Printer Daemon) protocol [8]. Such jobs (called RAW) are already coming to the print server fully formatted in the printer-understandable format (such as PostScript, PCL, HG-PL etc.). Therefore, the only task of the print server in that case is to provide spooling (queuing) of the job and finally to pass it to the printer device.

### 3.3. Driver distribution

The central CERN printing service provides drivers for two supported operating systems, which are Windows (XP and Vista soon) and SLC (version 3 and 4). Users of other operating systems (Mac OS X for instance) must take care themselves of the driver installation, however as soon as they have the correct drivers installed, they can take advantage of the LPD service and send jobs to the print servers.

Windows workstations receive drivers in a standard, automatic way. Once a Windows computer connects to a printer queue installed on the print server using RPC connection, it downloads the appropriate driver from the print server and installs it locally. At the same time all the default settings of the driver are pulled from the server. Each time a client PC reconnects to the print queue on the server, the version of the driver is checked and a new driver is downloaded again if needed.

The situation is more complicated for SLC computers, where clearly Windows print servers cannot help with the driver installation. Scientific Linux provides the CUPS print spooler and uses Foomatic [7] drivers to format print jobs before sending to the print server. Those drivers are installed using the only component written at CERN, called CERN Linux Printer Wizard [2]. This application is supplied from the server (via SOAP calls) with the database of printers available and their Foomatic driver definitions. The database is copied once per day onto the local machine and then the local copy is used. A full set of drivers is distributed in advance to all SLC computers and when a Linux user connects to a given printer queue, the appropriate Foomatic driver is configured.

### 3.4. Searching for printers

As mentioned in section 3.1. each print queue is automatically published in Active Directory and the information about printers is then available using the LDAP protocol.

Windows clients have built-in search capabilities, which allow searching for printers by many printer attributes like: location, colour printing capabilities, paper size available etc.

Scientific Linux computers use CERN Linux Printer Wizard to perform searches. It allows as well searching using many printers attributes and the search is performed against the local database.

In addition to those two mechanisms, a web interface has been as well provided to offer the same functionality to users of different operating systems as well as to visitors, who do not necessarily have the CERN Linux Printer Wizard (installed on Linux computers) or do not necessarily join the CERN Windows domain (in case of Windows computers). The interface is available at the following location:

<https://cern.ch/WinServices/Services/Printers/>

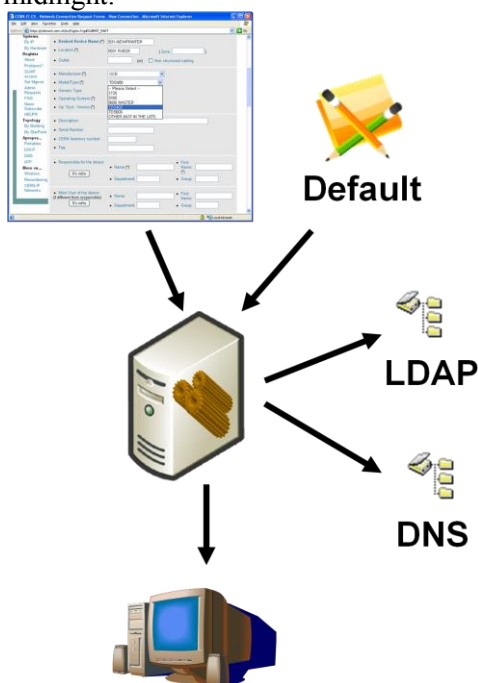
One needs to authenticate using a CERN account to be able to use it.

#### 4. Printer registration process

One of the main aims of the project was to fully automate the process of printer registration to reduce to a minimum the manual operations and therefore minimize the manpower needed to operate the service.

Only when the first printer of a new model arrives at CERN, a manual operation is needed. The new drivers have to be identified, tested and installed on the print servers. At that time the default settings for the driver, which will be pushed to each Windows client as well as default Foomatic configuration for the new printer model needs to be defined.

Afterwards, when a new printer device is connected, it needs to be registered in the network database (LanDB) with the correct model name specified. Once this is done, a periodical job running on print servers finds this new printer and installs automatically new queues on each print server. One of the print servers is then promoted to receive all the traffic for that printer. The DNS CNAME entry is then created and it points to that print server. A print queue from the promoted server is as well automatically published in Active Directory. Since that moment any Windows computer can connect to the print queue and can print. A SLC computer needs to wait for the next local database update, which normally takes place at midnight.



For selecting a print server to be promoted for a given printer a simple algorithm is used. The idea was to host printers from neighboring floors of the same building on different print server. In case of unavailability of one of the print servers, user can try to print on a printer, which is located on one floor up or one floor down. Therefore the configuration file contains the list of all possible floors across all CERN buildings with the hosting print server attached.

## 5. Conclusion

The new CERN printing infrastructure is in production since the end of 2006. The migration took almost 6 months as we wanted to make sure that service was not interrupted and users can always print. Since then we benefit from simplicity of the new system and the maintenance effort has significantly dropped. All print servers joined the standard CERN Windows server infrastructure and all the operations can be done by system administrators using the same procedures as for any other Windows server. The printer Support activity is now limited to the printer network registration process, physical installations and replacement of the consumables.

## Acknowledgements

The authors would like to thank to: Ignacio Reguero and Ivan Deloose for their help and all the knowledge they passed about the previous system; Jaroslav Polok for performing appropriate changes on the SLC configuration; Nick Garfield for his assistance in the DNS configuration and finally, to the CERN Printer Support team for their effort in migration to the new system.

## References

- [1] Bahyl, V., Garfield, N., DNS Load Balancing and Failover Mechanism at CERN. Proceedings of the CHEP-2006 Conference, Volume I, 519-523
- [2] CERN, CERN Linux Printer Wizard documentation, <http://cern.ch/linux/scientific4/docs/printing.shtml#desktop>
- [3] CERN, Printing User Documentation, <https://cern.ch/WinServices/Help/?kbid=070001>
- [4] CERN, Printer Status Page, <https://cern.ch/WinServices/Services/Printers>
- [5] CERN, Printing Service Web Page, <http://cern.ch/service-print>
- [6] CERN, SLC Linux distribution home page, <http://cern.ch/linux/>
- [7] Foomatic Database, <http://www.linuxprinting.org/foomatic.html>
- [8] McLaughlin, L., Line Printer Daemon Protocol, RFC 1179, The Wollongong Group, August 1990
- [9] Microsoft Corp, Network Printing Technical Reference, <http://technet2.microsoft.com/windowsserver/en/library/45552551-ae29-475a-aa35-745a8bdc9b811033.mspx>
- [10] Microsoft Corp, Print Services, <http://www.microsoft.com/windowsserver2003/technologies/print/>
- [11] Microsoft Corp, Print Spooler Architecture, <http://msdn2.microsoft.com/EN-US/library/aa506181.aspx>
- [12] Microsoft Corp, Windows Server 2003 Home Page, <http://www.microsoft.com/windowsserver2003>
- [13] Reguero, I., Foster, D., and Deloose, I. 1998. Large Scale Print Spool Service. Proceedings of the 12<sup>th</sup> USENIX Conference on System Administration (Boston, Massachusetts, December 06 – 11, 1998). System Administration Conference. USENIX Association, Berkley, CA, 229-242
- [14] Vixie, P., Gudmundsson, O., Eastlake, D., Wellington, B., Secret Key Transaction Authentication for DNS (TSIG), RFC 2845, ISC, NAI Labs, Motorola, Nominum, May 2000
- [15] Vixie, P., Thomson, S., Rekhter, Y., Bound, J., Dynamic Updates in the Domain Name System (DNS UPDATE), RFC 2136, ISC, Bellcore, Cisco, DEC, April 1997