World-Wide Web in High Energy Physics Experiments 'A Status Report'

M. Dönszelmann
CERN ECP-DS, CH-1211 Geneva 23, Switzerland

Abstract

This report presents the status of the use of World-Wide Web (WWW) in High Energy Physics (HEP) experiments. The use of WWW in general, for 'Online Datataking Systems' and for 'Offline Analysis Systems' is discussed. In each of these cases the current use and a possible outlook for the future is described. Statistics on the actual use of WWW in HEP experiments as well as its current problems and future needs are also presented.

Presented at The First Topical Seminar on WORLD-WIDE WEB
and beyond for Physics Research and Applications,
San Miniato, Italy, 14-18 March 1994

World-Wide Web[1] is an ideal system to give people in geographically dispersed locations access to the same universe of information. Although originally designed at CERN (European Center for Particle Research) for use in the HEP (High Energy Physics) area, WWW has now taken the lead in many other disciplines.

This report focusses on the use and problems of WWW for HEP experiments. It describes obvious uses of the web as well as specific applications, which WWW initially may not have been designed for. An overview of examples is presented. It is clearly impossible to cover all applications of WWW in HEP, but an effort was made to discuss most of them.

1 Setup of Clients and Servers

If experiments want to make full use of the web, they must make sure all web software is set-up in the correct way. Too often experiments have installed web clients (Browsers like XMosaic and Viewers like XV) on one type of their platforms but not on others. The nice thing about WWW is that all information is readable (in future editable) anywhere in the world from any type of platform. Browsers are available in different types, ranging from simple text to sophisticated graphics, on most common platforms (PC, Mac, Unix, VMS and even CERNVM).

Most experiments have set up one or two WWW servers. It should be noted here that a WWW server, also referred to as an http-daemon (Hypertext Transfer Protocol), is nothing more than a smart FTP (File Transfer Protocol) server. It basically makes a set of files available to the rest of the world. It is important that the server serves files directly, and not through other servers (like NFS or FTP). In the latter case a simple retrieval of an HTML (HyperText Markup Language) file would bother two servers while it could have been done with one.

A good example here is the DELPHI experiment, where 3 central servers run on three clusters. Although DECNET connections exist between those clusters WWW access is done directly to the correct server and not via DECNET. For the user of the web this works transparently and the DELPHI information looks like one universe of information.

2 Use of WWW in General

The general use of WWW for HEP experiments contains everything **not** related to the 'Online' running, control and monitoring of the experiment or the 'Offline' data analysis.

Most experiments have setup information in the area of publicity, like for example the L3 tour around their experiment site, see figure 1. Public infor-

Figure 1: L3 Publicity Example

mation in the form of lists of participants, collaborating institutes, phone lists, "whoiswho"s and other administrative stuff is in most cases available. This type of information is normally maintained by secretaries, so automatic conversions of existing databases or direct maintenance of the information in HTML documents is an issue here.

Publication of Physics papers on the web, per experiment or centrally (CERN Preprint server or SLAC Spires Database), is very popular. Most papers are however only available in scanned-bitmap or PostScript format, while papers broken up in HTML would be an attractive alternative.

Private or semi-private information is becoming available on several sites, since the implementation of security in the form of usernames and passwords. News systems are normally used to broadcast information related to experiments, and gateways to WWW now exists, so users can use the web as a news-archive browser. In the near future news services (including posting of news) will be integrated into the web.

It is possible to have discussions via the web, for instance on draft papers. In this case a small group of people will have access to the draft document. All of them may comment on the document, using annotations. While reading the draft they will see the comments already made by others, thereby avoiding to have to make the same comment twice. The editor may then process all comments in one pass instead of going through several mail messages. The group annotation feature is however not yet available.

3 Use of WWW for an 'Online' System

On the 'Online' side of an experiment the web has many applications. Documentation ranging from design drawings (see figure 2 for an example from the ATLAS experiment) and manuals to online help with clickable block diagrams to help the people on shift.

An interesting example is DELPHI's histogram presenter, displaying histograms to check, 'online', the quality of the data, see figure 3. When help on a specific histogram is needed, a simple click on it will start XMosaic and bring up the help document belonging to that histogram. The help document may use all the features of HTML, like drawings and links to further documentation, to explain the user what to do.

Not all information of 'Online' systems is kept in files. Access to information sitting in databases like an online bookkeeping system is possible via the web. Special servers, in DELPHI for instance, provide information read from the slow controls channels, like temperatures and high voltage settings. The document in figure 4, showing temperature channels, does **not** exist in file form, but is completely synthesized. It contains links to other synthesized documents on the channel level. Machine information from LEP and SPS is also available in this way.

Figure 2: Atlas Technical Design Drawings

Figure 3: Integrated Online Help

If we look in the future we imagine that WWW could be used as a simple straightforward interface not only to monitor the experiment but also to control it, via forms for example. This needs some protection of course.

A simple event display to be run via the web is another promising possibility. Although maybe not so fast and accurate as a real interactive viewer, it would allow anybody to display events anywhere in the world on any type of machine.

4 Use of WWW for an 'Offline' System

The current use of the web for 'Offline' analysis systems seems to be restricted to documentation and distribution (FTP) of the software. Administrative information such as tape and processing bookkeeping is in some cases available.

WWW however offers the capability of executing scripts on the server machine. Therefore remote compilation or submission of jobs on systems like CSF (Central Simulation Facility) belong to the possibilities of being exploited by the web. Maintenance of software is another issue. A combination of WWW and, for instance, CVS (Concurrent Version System) could give a solution for remotely maintaining 'Offline' software via a very simple interface. Both previous cases need solid underlying logic to implement the functionality desired, but the web will solve at least the networking and common interface.

Figure 4: Example of DELPHI Online Slow Controls Information

5 Statistics on the actual use of WWW

Some investigation has been done to get statistics on the actual use of the web by high energy physicists. The information was kindly provided by the installers of the different sites. Some numbers rely on log files while others are merely estimations.

Figure 5 shows an obvious increase in the number of servers since the introduction of WWW in the HEP world[2][3]. Because most sites run or will run more than one server we should in future probably refer to sites or organisations instead of servers.

Figure 5: Statistics on the Servers

Figure 6 shows the number of servers running at one site. At the moment most sites run only one server, but in future performance issues will provoke a tendency to run one server per machine.

Figure 7 gives an estimate of the number of documents each site provides. Peaks here come from existing databases which were gatewayed to the web. The smaller sites show an average of 400 documents.

The number of requests per day, see figure 8, really shows us the use of the information. 200 request per day seem to be an average, again with peaks for huge databases and the CERN home page, which half of the world is referring to when starting up their browser. Up to 5000 request per day is manageable by a dedicated computer, but above that you might have to split your information and requests. The 200 request per day pose no problem on a normal system.

Figure 6: Statistics on the Servers (continued)

Figure 7: Statistics on the Servers (continued)

Figure 8: Statistics on the Servers (continued)

6 Current Problems in WWW

Of course WWW has its problems, many of which are temporary. One of the main problems is that people easily get lost. Hypertext (the Web part of WWW) is not such an obvious thing for human beings. The information providers have to make sure that they organise their information well.

People use WWW as a side product, mostly because they cannot do everything with the web and not all information is available (yet). Lots of information need to be converted, and once converted it poses a maintenance problem; supporting the HTML of the original format (or both)?

Putting information on the web is easy, but we have to make sure it is up-to-date. Old information is useless. Maintenance, preferably automatic, is required.

7 Future of WWW

What will or should the future of WWW bring for the HEP community. A few hints are given below:

- HEP, as many other disciplines, will need HTML editors so that everybody can use the web and publish information. Likewise we need converters to convert whatever we already had into web format.
- When editing via the web becomes available, HEP can take advantage of it in their distributed multi-platform environment.

- Extensions to HTML in the form of tables and formulas are very welcome and will definitely be used by HEP.
- Group annotations seem to be necessary to make WWW really a collaborative tool.

8 Conclusions

The High Energy Physics world uses the web, but leaves a set of features unused. Currently a lot of publicity is done via WWW, but the use of the web for private and semi-private information is definitely underestimated. The World-Wide part of WWW is used by everyone, people transfer files and look at things over long distances, but the Web part (hypertext) is not fully understood everywhere yet. Generally the web is used as a side product.

To make sure we start using the web to its full extend, we should install it and make it available on all platforms we are using. Too many times the information on the web is published and maintained by one person. But while we all browse the web, we should also 'all' publish on it.

WWW provides an information infrastructure. It does **not** maintain that information. Old information is useless, so although we now have a nice way of communicating information all over the world, we should in future put our effort in organising and maintaining that information.

New features like editors for HTML, editing via the web, group annotations, tables and formulas in HTML are needed to make World-Wide Web a big success in the High Energy Physics world.

References

- [1] T.J.Berners-Lee, R.Cailliau, J.F.Groff, B.Pollermann, *World-Wide Web: The Information Universe*, Electronic Networking: Research, Applications and Policy, Vol. 2 No. 1, pp.52-58, Meckler Publishing, Westport, CT, Spring 1992.
- [2] T.Berners-Lee, R.Cailliau, N.Pellow, A.Secret, *The World-Wide Web Initiative*, Proceedings of INET'93, San Fransisco, USA (1993) DBC1-5.
- [3] T.J.Berners-Lee, R.Cailliau, J.F.Groff, B.Pollermann, *World-Wide Web: An Information Infrastructure for High-Energy Physics*, Proceedings of the Second International Workshop on Software Engineering, Artificial Intelligence and Expert Systems in High Energy and Nuclear Physics, L'Agelonde France-Telecom, France (1992) 157-164.