The Physics Analysis Environment of the ZEUS Experiment

Lothar A.T. Bauerdick, Oleg Derugin^a, Dave Wilkinson,
Matthias Kasemann, Olaf Mańczak

Deutsches Elektronen Synchrotron, Notkestrasse 85, 22603 Hamburg, Germany

The ZEUS Experiment has over last three years developed its own model of the central computing environment for physics analysis. This model has been designed to provide ZEUS physicists with powerful and user friendly tools for data analysis as well as to be truly scalable and open.

1 Introduction

During summer 1992 the HERA collider at DESY delivered its first luminosity and the ZEUS detector recorded its first data. At that time the ZEUS computing environment consisted of some few dozens of workstations (DEC and SGI), a parallel reconstruction facility running on multiprocessor SGI machines and there was a strong commitment in favor of computing solutions based on inexpensive RISC/UNIX technology. Nevertheless, a large fraction of the physics analysis and software development was done on the IBM 3090 mainframe and on the central DESY VAX cluster. In addition, the IBM mainframe with its STK tape library acted as our central repository for experimental and Monte-Carlo data.

In this paper we describe the evolution and status of the physics analysis environment at ZEUS beginning from this initial state.

2 First Experience

During the 1992 data analysis period we have made three fundamental "discoveries" concerning efficient and effortless data analysis. First, we found that the process of analyzing data can be considered as an infinite loop consisting of two steps:

- *interactive analysis* - reading mail and news, writing papers, analyzing ntuples, plotting histograms, *thinking*, coding an idea into a program etc. (i.e. developing analysis software)
- *batch analysis* - running an analysis program on a large sample of experimental data to check the idea (this can take 5 minutes to 1 week), then back to the interactive analysis...

It is natural and obvious that workstations (or more generally workgroup-servers and X-terminals) with large graphical screens and powerful computing resources are extremely convenient tools for interactive analysis. These workstations, however, do not address and do not solve all the problems of batch analysis. In particular, the access to a huge (several TBytes) remote tape library or to large amounts of

^aOn leave of absence from Moscow State University, Institute of Nuclear Physics

SCAN-9602014

CERN LIBRARIES, GENEVA

509607

CERN LIBRARIES, GENEVA

Quality insufficient for good
scanning

data available within a distributed network filesystem is significantly limited by the network performance. Thus, (this was our second "discovery") the use of workstations does not eliminate the need for central *mainframe-like* computing facility with fast access to experimental data but, to the contrary, enhances that need.

Finally, we found that in an environment where about 50 physicists (this was our estimate in 1992) work simultaneously analyzing more or less the same data it is highly reasonable to keep a large sample of selected physics events on disk. The main goal was to be able to analyze the whole sample quickly and to use (*share*) disk resources efficiently. This approach prevents unnecessary data duplication by different users when the right tools are provided.

3 ZARAH - The Central Batch Facility

The outcome of our 1992 experience was a proposal for ZARAH (*Zentrale RechenAnlage für HERA Physik*) the central computing facility for the ZEUS experiment at DESY. Among our major short term goals were:

- rationalized and optimized *path to data* in order to avoid network bottlenecks,
- a powerful and user friendly batch system, able to run at any given time at least three jobs processing some 20 events/sec, with direct access to
- a fast disk store for a large sample of selected events.

Our long term goal was to gradually eliminate the IBM from the offline processing both as a computing facility and data server. The details and development of this system over time is discussed elsewhere in these proceedings ^{1,2,3}, but we summarize the current configuration here:

- 600 MB of fast disk storage available, enough to store a complete year's worth of compressed MINI-DST data and selected events from previous years
- Two SGI Challenge XL multiprocessor machines (34 processors total) to serve as batch platforms for the on average 150 different physicists who use ZARAH every week
- SGI Challenge XL equipped with 18 processors for data reconstruction
- SGI Challenge DM to act as an I/O server
- AMPEX tape robot with storage capacity of 6 TB to provide our central repository for experimental and reconstructed data
- The *Open Storage Manager* (OSM) system, providing access to the STK silos where Monte-Carlo files are stored
- HIPPI links providing reliable connections with high transfer rates between the components.

This system has provided us with excellent performance, connectivity and scalability resulting in a significant improvement in data analysis at ZEUS.

Figure 1: The offline computing environment in 1995

4 Interactive Desktop Environment

The ZEUS computing environment for interactive physics analysis consists of several clusters of personal workstations, workgroup-servers, file servers and X-terminals connected together with the "magic glue" of the DESY network infrastructure. This environment provides 250 ZEUS physicists and engineers with comfortable workplaces. A typical workplace consists of:

- A personal X Windows screen which allows plotting histograms, using the ZEUS event display program, etc.,
- Computing and disk storage resources sufficient for comfortable *ntuple*-based physics analysis.
- All the tools necessary to develop programs, submit jobs, read mail, browse news, surf on the Web (we use WWW for most of the ZEUS documentation).
- A network which connects this workplace to the rest of world.

We found using workgroup-servers (i.e. sharing resources among a small group of physicists working on the same topic) as more efficient than using personal workstations. Therefore ZEUS has recently joined the AFS and UNIX-workgroup-servers projects launched by the DESY central computing group.

5 ZEUS Software

The ZEUS physics analysis software contains about 1 million lines of code (mainly FORTRAN, partially C and C++) and is being constantly developed by a large group of physicists. Code *versioning* is managed by the CERN CMZ *code management tool* and is distributed to remote sites with the *rdist* program from the central DESY repository.

In order to hide from the user differences between various platforms and give a possibility to compile, link and run the same code without changes, we have invented an X Windows *imake*-based *Makefile* generator. *Templates* used by this generator are maintained in a consistent way by our code managers.

We use *ADAMO*, an entity-relationship programming system developed at CERN to provide the underlying event structure which allows each event to be considered as a relational database.

Software Portability

In earlier times this software was developed on IBM MVS and VAX VMS platforms. In 1991 the software was ported to UNIX (first for SGI and later for DECstations). There was strong commitment in favor of inexpensive UNIX/RISC computing and ZEUS approved DECstations as the collaboration standard. Nevertheless, we realized quickly that academic prices of workstation equipment vary significantly among different countries and that rapid changes on the computer market make any long term recommendation practically impossible. Therefore we decided to maintain every release of the *offline* software for any given UNIX platform with the proviso that institutes willing to introduce a new platform should contribute to the initial port.

We found porting between various open UNIX platforms to be nearly effortless and we have to admit that a large fraction of the changes was necessary due to evident bugs rather than non-portable solutions. Today, we maintain the ZEUS *offline* software for SGI IRIX 5.3, DEC Ultrix, Digital UNIX (OSF/1), Solaris 2.x, and HP-UX. Despite the inevitable amount of work necessary to maintain all these ports we consider this approach to be truly profitable. The ZEUS collaborators have the choice of purchasing the best equipment, according to a price/performance ratio comparison, which is available on the market. ZEUS doesn't become dependent on any given hardware vendor and that leaves room for price negotiations.

6 Wide Area Networking

Most of the ZEUS experimental and Monte-Carlo data is available only at DESY. ZEUS is a collaboration of 50 scientific institutes from 12 countries and many of the ZEUS physicists may work on their analysis via the net from their home institutes. In order to provide convenient access to these data and batch computing resources at DESY the ZEUS Networking Committee was formed. They perform, on a regular basis, monitoring, and coordinate efforts in different countries, to achieve sufficient bandwidth and performance for the ZEUS networking needs. Their efforts have produced big improvements in the accessibility of DESY from outside institutes.

7 Conclusions

The ZEUS experiment has moved its physics analysis to a heterogeneous environment with an *open, scalable* and *high-performance* architecture, with a *state-of-the-art* batch facility, dedicated network-connected servers, large volume hierarchical mass storage and hundreds of interactive desktop workplaces on the net.

This environment gives us everything we need for the efficient physics analysis at ZEUS experiment. There are *no bells and whistles and flashing lights* but our main goal at ZEUS was to *do physics rather than computing* and we intend to continue our efforts with this goal in mind.

Acknowledgments

We would like to acknowledge the support of the DESY Directorate and the DESY central computing group. In particular we are indebted to Michael Ernst, Martin Gasthuber and Karsten Künne for their strong efforts on our behalf.

References

1. O. Mańczak, "The Computing Environment for Physics Analysis for ZEUS Experiment at HERA", in *Proceedings of CHEP '94 Conference*, 1994.
2. L.A.T. Bauerdick et al., "ZARAH - The Central Computing Facility for the ZEUS Experiment", in *Proceedings of CHEP '95 Conference*, 1995.
3. M. Ernst, "Putting all that (HEP-) Data to work - a REAL Implementation of an Unlimited Computing and Storage Architecture" in *Proceedings of CHEP '95 Conference*, 1995.