

# **Administrative Computing at CERN, where are we going?**

Achille Petrilli

June 1, 1999

## **Abstract**

CERN is known for its particle accelerators, the physics research which is done there and the related advanced computing needed to support both activities. It is less known that CERN administration is also heavily using computers since a number of years and using some of the most advanced technologies available. This paper retraces the evolution of a big project in administrative computing which has involved replacing all old applications with new expanded ones over a few years. Now that a stable environment has been built for the administration, what are the challenges in front of us? how to keep up with the rapid pace of technology change?

## **Introduction**

CERN administrative computing had been identified in the mid-80s as a key factor for the LEP and post-LEP era. A project was started in July 1990. At that time, CERN had a number of administrative applications which served each a specific community of administrative users. The user interfaces of those applications were clumsy, sometimes based on proprietary hardware, even for the terminals. The programs themselves were difficult to modify and could not talk easily to each other; when this was possible, batch jobs were needed. Non administrative users were not concerned by those applications and their access to relevant information was always via the administrative users.

The main objectives of the project were to provide a more user-friendly environment, integrated data bases, widespread use of the administrative applications, even outside the Administration. All that had to be based as much as possible on commercial products.

## **The AIS project**

It was immediately clear that the key to a successful and permanent change in our Management Information Systems applications resided in avoiding double data entry and double copies of data. The proliferation of applications not under central control and maintenance had made the whole area very difficult to manage and very resistant to change. The whole project has therefore put the main emphasis on all aspects where information ownership and responsibility for updates was concerned. In addition, ease of use has also been privileged to help improve the support for management decisions.

## **Concepts**

The project based itself on proven technologies whenever those existed, with the goal of not having to rewrite basic functionality which could be purchased on the market. This was a key point as the human resources devoted to the project were fixed and could not easily be increased. Focusing on the end user functionality was our main priority. A unique OS environment was desired to limit the amount of knowledge and work devoted to *unproductive* activity from the end user point of view. At the same time we wished to have the choice between multiple hardware suppliers to minimize costs. Unix was therefore chosen in 1990 as our platform.

A very similar reasoning was carried out for the data base engine: we wanted only one engine to minimize the resources devoted to it while at the same time we wanted to have a choice of end user applications available on the market based on such an engine. We wanted also a fairly recent technological solution, which nevertheless had already a well established market: the Relational Data Base model offered the requested characteristics. Oracle was then chosen in 1990 (at that time, basically no other RDBMS supplier had any sizable number of applications running on their DB engine and on Unix).

Another critical choice to make was about the human interface for the application to buy and/or develop. All commercial packages had a functional but fairly old human interface, the choice was between using what we could buy or to roll our own. In 1990, MIS users were around 300, all from the CERN administration. In a sense, we had to provide them with a better human interface, more easy to use and blending with their personal computers environment. We instead chose not to write a fancy human interface for our main users. We wanted to attract a whole new class of users: top, middle and lower management, which had no other access to the information than printed paper lists. For this new class of users we decided to create from scratch new tools matched to CERN needs and way of working, basing ourselves on the new data bases being introduced for our traditional, professional, users. The professional users would benefit anyway from the change because of information sharing, suppression of double data entry, faster applications, etc.

The applications to be created had to help users handle authorizations, signatures and other bureaucratic chores as well as providing them with a view of the available information as accurate and as fast as possible. Instead of the traditional several days or weeks to see the result of a purchasing procedure, we wanted to give immediate or quasi immediate feedback to the requester.

## **Solutions Used**

For the hardware, the project decided to go through a tendering process to help push prices down. Had we decided to keep the old applications, we would have had to purchase a new IBM (or compatible) mainframe, the price for this solution was around 2MCHF. The final price we paid after the tendering process for a multi-processor Unix machine was 300 KCHF! Clearly the choice of going for a more dynamic OS environment was the right one and was already paying back.

The Oracle choice did not directly produce such a visible effect. Instead Oracle (or, better, the RDBMS choice) allowed us to easily make separate applications communicate with each other and easily share information.

Our home-made tools followed very similar reasoning: everything which could be bought we did not want to rewrite, we just wanted to concentrate on specific CERN problems, not on general user interface problems. One aspect we had to take into account was the presence in CERN of both Macintosh and IBM PC-compatible computers, our solutions had to work on both platforms.

For this reason one tool which allows access to financial information, the Budget Holders Toolkit or BHT, has been based on the well established and available Microsoft's Excel product augmented by a query tool to access information stored in Oracle data bases. All presentations, data manipulations, graphing, etc. were then handled by Excel, either directly or via a macros written in CERN. Same applies to a later tool, the Human Resources Toolkit or HRT, which does an equivalent job for people management.

The tool which handled signatures, authorization, document routing and workflow, the Electronic Document Handling or EDH, did not lend itself to fit into an already existing product like BHT with Excel. EDH has then been written completely in CERN but using an existing commercial Graphical User Interface package, again to minimize our effort in *unproductive* activities.

All existing applications, whether purchased or developed in-house, went in production between 1992 and 1996.

## **Problems encountered**

The acceptance of the solutions provided has been very high. Over the years the number of our users has been steadily increasing. Currently we have some 6'000 active users, with 11'000 registered. The scope of the administrative applications has enlarged now covering also many needs expressed by the technical and scientific community.

Obviously this change in the scope, coupled with the technology shifts experienced over the years, has been putting under stress our solutions. The increased load has not been a blocking factor: processing power has increased some 32 times for the same amount of money since the beginning of the project. Oracle products have mutated in reasonably good backward compatible manner over the years.

On the other side, the tools we developed have gone through a life of changes which has forced us to sometimes rethink our developments. We try to examine those changes by splitting the problems encountered in two categories: industry driven and user driven.

### **Industry Driven**

The desktop computing world was clearly split into two camps, Macintosh and PC. Over the years, due to Apple's problems, the Macintosh platform has received from industry a much smaller support than the PC platform. This has resulted in a longer lead time for applications such as Excel to appear on the Macintosh, forcing us into playing with product revision dates because of external factors beyond our control. Furthermore, the changes in the macro language the Excel product has undergone, have forced major rewrites of the BHT application. Those changes have obviously benefited the users, but at a great cost for the project in term of development time and with very little added value to the information provided.

A completely different problem has appeared with EDH. The supplier insisted to have us pay a new license for every *new* Operating System, where the definition of new was provided by the supplier itself. So MacOS8 is a new OS, Win95, Win98, WinNT are all new OS and for each one we would have to pay a new license. This has led us to the decision to not support Windows NT and to move away from that supplier.

In both cases, our problems stem from the rapid pace of changes happening in the desktop computer world. Unfortunately these rapid changes did not bring real value to the user, apart from mostly cosmetic ones, while at the same time preventing us from adding new added value features to our applications. In addition, the very nature of our applications, software that had to run on the client platform, was tying us in industry wars of no interest to us or to our users.

It was clear that building what has become critical software on the likes of Microsoft and Apple had been successful but at an exorbitant price.

### **User Driven**

Users' expectations have gone up as well over the years. Being able to come up with, for instance, financial figures accurate to yesterday is no longer enough for many people: what about foreseen but not yet registered expenditures? what about actions already started which will produce a financial commitment ? what about long running contracts which have already been signed ? Similar questions have been raised about EDH. Some 13 official actions are today implemented in EDH, the 13 most popular ones: what about the 14th one? what about a low volume one but with high financial impact ?

We found ourselves unable to answer to all requests from users because of the high popularity of our solutions and because of industry changes which we could not foresee or avoid.

In addition more and more of our users were coming from external institutes and were spending only a limited amount of time in CERN, but they still wanted to use our applications from their home institutes. Our solutions based on software running on the client platform could not be extended easily to a more distributed environment.

It was clear that to satisfy growing users' demand we had to rely on a completely different paradigm.

### **Current Evolution**

At the same time as we were experiencing the first problems described above, the World Wide Web or Web, had become extremely popular and a new technology had appeared, Java, which promised a platform independent developer's paradise. No more problems of platforms, OS, CPU type, binary format, etc.

We conducted various studies and a few small pilot projects. Java emerged as a good, solid choice both on the client (applets) and on the server (servlets). The decision was therefore taken to start up an applet based development for the data entry of experiment related information (the Person, Institute, Experiment application or PIE), a mixed applet/servlet development (EDH port to the Web) and a pure servlet based development (BHT port to the Web, WebBHT). In parallel another application, not based on Java but

on Oracle tools, has been made available on the Web, the Contract Follow Up or CFU application.

PIE has been completed and is in production, but the effort to achieve Java's promise of Write Once Run Anywhere has been higher than foreseen. Despite the success of this project, applets do not look like the way to go.

Both BHT and EDH ports have been started and most of their functionality is now available on the Web. To avoid problems with Year 2000 tests of both applications, we plan to have 100% of them on the Web by the end of this year and to discontinue the old client solutions at the same time. The EDH port has ended being also an almost pure servlet development, despite having started also with an applet slant.

It did not take long to our users to discover the joy of the Web: it was so easy to jump from EDH to BHT and back, just by clicking a link! It did not also take long to discover the limitations of our protection mechanisms: one had to login in each application, EDH and BHT and PIE and CFU, the first time you tried to access a page on each of those systems.

So a common login scheme has been implemented for our Web applications to avoid the need for repeated logins.

## **Present and Future**

We have just described how the major problems uncovered over the years have been solved and what the present situation is. Many other changes are in the works though, we will try to show from where the new challenges are coming and possible ways of solving them.

### **New Applications**

The integration of all administrative applications and data bases, coupled with the overall coherency and accuracy of the information, has generated a vast increase in our user base as already mentioned. This has brought us completely new demands for applications or access to information in ways which were not foreseen. Two good examples of such applications are PIE and the Project Progress Tracking (PPT). Both of them are based on existing administrative information, but they exploit this information for different purposes than purely administrative ones.

#### **PIE**

As one can easily imagine, People, Institutes and Experiments information is a really basic part of CERN corporate foundation. Traditionally, this information has been managed in two (or more) different ways, with different systems and with different purposes in mind. On one side, CERN administration had to know about the laboratory activities, its visitors, the experiments currently running or foreseen, the institutes involved. All this information is used to prepare the official documents, reports, etc., of the Organization. This usage is fairly intensive at the beginning of new scientific programs, while later decreasing, leaving the field to the scientific world. On the other side, experiments themselves must have access to the information to be able to run, to co-

ordinate their work, to produce their results. The usage of the information is in this case much more dynamic as it spans the whole lifetime of the experiment. Also, the emphasis is on the scientific results (pre-prints, articles) rather than the more formal needs of the Organization. On middle ground, one finds the CERN Library, with its mission of collecting and making accessible scientific results, and the Gray Book which is a source of information both scientific and organizational.

All those three areas handled separately their own information until recently. This had not only a high cost, but also produced contrasting results in many cases. As one would expect, there was no single place which had the *best* information (most accurate, most up to date) and as a result there was no easy way of presenting a coherent view of all scientific activities.

PIE has centralized all relevant information inside our existing Human Resources application, while providing an interface for distributed data entry and update. CERN, the Library and the experiments now have each one a part of responsibility in maintaining the information and each can take advantage of the work done elsewhere. As a result, the Gray Book, which was a manpower intensive exercise with a paper only output, is now produced as automatically as possible and is also available on the Web.

The information involved in this project has always been part of the traditional MIS domain, but the participants in the project have come from many different backgrounds, scientific, engineering as well administrative ones.

## **PPT**

The Project Progress Tracking development has started recently on a specific request from Atlas. The problem here is to be able to follow the manufacturing and later the assembly of the Atlas detector. The challenge of today's big experiments is to handle world-wide distributed work of loosely coupled units, with different funding, backgrounds, locations but with a unique schedule which must be respected for the whole system to work. This is a fairly common industrial problem, which is normally solved via some project management tool, strong central management, central funding, etc.

The deeply distributed nature of all aspects of Atlas, whether financial or technical, renders traditional PM tools not suitable for the whole task, the PPT system is the missing link to allow up to the minute project tracking without the traditional heavy weight management structure of industry.

The collaboration establishes a Product Breakdown Structure (PBS) to define the major components of the detector and a few critical details for those components. Once this is done, the responsible person for a component can decompose further the structure, until what is considered a unit of work, a work package, can be given to some participant institute which is now in turn responsible for the work. At each level, responsible persons can report about their progress, the budget spent, problems encountered, etc. The reports can be consolidated at the upper level so that to obtain a less detailed, but more global, picture of the progress made.

The system allows reports to be filled in via the Web or e-mail, either spontaneously or on request from the upper level(s). The resulting information can be exported to

traditional Project Management tools (e.g. MS Project) for overall views of the project status. Some charting and tabular reporting can also be done inside PPT.

The link with traditional MIS applications comes from the fact that all information is linked to people and institutes and much of it is also linked to financial agreements between national Funding Agencies and CERN, which in turn links with budget information. There is then a whole part of the basic information which is dependent on our Human Resources system and PIE, another part is linked to our financial application and BHT.

## **Today's Problems**

The expanding user community and the need to share information with other applications, possibly not under our control, is THE problem. In a sense this is a nice problem to have: it comes from the recognition of the high quality and usefulness of the work done. On the other side, this is far more complicated to solve than a simple technical problem.

Of course many other problems still exist, but they are more technical and can be solved fairly easily, at least conceptually. We will then concentrate on the information sharing problem, which can be summarized in two broad areas: access rights and relationship between applications. Both of them come from an inherent limitation of our entrenched way of thinking about systems and applications: we more or less artificially partition a whole process, for instance running CERN or building and running an experiment, into more manageable work units and we create boundaries between those units. This is not just a CERN or HEP problem, the whole world works like this since a long time. Things work fairly well as long as boundaries between systems are under human control, unfortunately the introduction of more and more automated applications makes this impractical and raises the kind of problems we experience today.

## **Access Rights**

Thanks to the port of our applications to the Web, sharing information becomes simply a matter of correctly linking information together, a job that computers can do very well. So, CFU links to BHT, which links to EDH which links to HRT and PIE links to HRT and so on. But BHT enforces a well defined set of rules to access information, which are not the same as the ones of CFU or of EDH, etc. We now only login once into our Web applications, but each separate application is still responsible for granting access to its own information. A good example is the combination BHT and CFU. BHT protects its information based on a hierarchical structure, the hierarchy of CERN. CFU protects its information based on a specific purchasing procedure, which obviously may cross multiple hierarchical units. In various cases, people responsible for a contract in CFU may be denied access to the resulting financial information because of this. Similar problems appear in other areas. All come from the same basic assumption that each application makes, bundling together ownership of information and access control.

## **Inter-Application Relationships**

There is a fairly large amount of information which could be provided by multiple applications. Obvious examples are financial information about contracts: is that BHT or

CFU responsibility? In many cases there is no clear answer, as the information is shared between applications. We do not have anymore the clear cut boundaries we had between applications just a few years ago. The second problem we find here is the way of linking information together: today we just link to an URL, but this is a dangerous way of creating inextricable dependencies between applications, dependencies which will have to be documented and maintained forever in the future.

### **What Future, then ?**

We have presented a sizable amount of problems which are in front of us, none of them has a simple solution really. We can expect more and more requests of integration (or should we say blending?) of existing or new applications with the administrative ones. The common login has been the first step toward seamless information sharing. The next steps are more difficult and we believe that they will require a more radical change in our way of working. Today we provide information inside CERN, but also access information outside the traditional administrative boundaries (for instance, CFU has a need to access engineering information stored in the Engineering Data Management System of CERN, CEDAR). There is no reason for this trend to stop here, and we could reasonably think of linking to information coming from outside CERN as well, the Web based applications rendering geographical distinctions obsolete. Because of the various issues presented above, it should be clear by now that solving, for instance, access rights clashes inside CERN and between our administrative applications is not going to avoid future problems, just buy ourselves some more time.

The choice we face is simple: we can either try to build a complete knowledge of how other inter-related applications work, giving raise to an explosive maintenance nightmare or move out of the application-oriented view to a more process-oriented way of developing and integrating administrative software. Clearly in the long term the latter is the only solution possible. In this respect, object oriented technology seems more suited to this view than traditional technology. As an example, our port of EDH to the Web is based on a commercial product, Oracle Workflow and in-house Enterprise JavaBeans development. The whole system works through a set of Common Business Objects (CBOs) which encapsulate CERN rules and hide the details of the commercial or in-house solutions used. The advantage of this solution is that our CBOs match exactly the needed administrative abstractions, so for instance a change in an administrative rule will need only a change in the unique corresponding abstraction. The complexity of such a change can be very important, but it would be confined to a well specified place. In addition, other "applications" could reuse CBOs while developing additional ones, thus reducing the development effort to the minimum required. We will try to expand this approach to other related areas to get a better understanding of pros and cons of such a solution

### **Conclusion**

In the past years we have successfully upgraded CERN's MIS services to current state of the art technology. At the same time, a rationalization of the administrative procedures has taken place and the information provided has been made coherent. Information sharing has been made possible between our administrative applications and also other applications not under our own control (CEDAR, Access Control DB). We are now



starting to see other changes looming at the horizon and we are assessing where our solutions will not fit well with those coming changes. The traditional differences between engineering, physics and administrative computing are becoming less and less evident: each side is taking the best aspects of the others and is more and more dependent on somebody else's information and solutions. After what we said, the word "application" is becoming blurred and it conveys now very little meaning.

A new corporate data base revolution is coming and is going to affect a much wider user community than the administrative one. We feel confident to be well positioned to tackle the challenges of the transition into the 21st century' corporate information systems.

## **Bibliography**

- CERN-AS-92-006. E. Dheur, J. Ferguson, R. Martens, A. Petrilli, B. Smale, Modernizing Corporate MIS: from Information System Modelling to Implementation, 9th European Oracle User Conference, March 1992, Cannes, France.
- CERN-AS-93-001. R. Martens, A kamikaze approach to turn your corporate applications into an integrated ORACLE/open systems solution, EOUG Oracle User Forum 93, March 1993, Vienna, Austria.
- CERN-AS-93-002. J. De Jonghe, The Paperless Organisation? EOUG Oracle User Forum 93, Vienna, Austria.
- CERN-AS-94-001. M. Moller, J. Purvis, A Graphical Client-Server approach to Financial Management, EOUG Oracle User Forum 94, April 1994, Maastricht, The Netherlands.
- Various Authors, <http://bht.cern.ch/bht>
- Various Authors, <http://hrt.cern.ch/hrt>
- Various Authors, <http://www.cern.ch/CERN/LHC/LHCwelcome.html>
- CERN-AS-96-002. J. Ferguson, S. Foffano, R. Martens, A. Petrilli, Administrative Information Systems in an International Research Environment, R&D Management Conference: Using IT Effectively in R&D Management, 23 - 25 Sep 1996, Lausanne, Switzerland.
- CERN-AS-96-004. A. Petrilli, EDH Return On Investment Analysis, AS Division, CERN, April 1996, Geneva, Switzerland.
- A. Petrilli, Advanced Information Systems, AS Division, CERN, November 1998, Geneva, Switzerland. Unpublished.