Fifty years of research at CERN, from past to future: Computing

D. O. Williams
CERN, Geneva, Switzerland

Abstract
Computing in the broadest sense has a long history, and Babbage (1791–1871), Hollerith (1860–1929), Zuse (1910–1995), many other early pioneers, and the wartime code-breakers, all made important break-throughs. CERN was founded as the first valve-based digital computers were coming onto the market. Computing at CERN will be considered from various viewpoints, including the various 'wares' (hardware, software, netware, epeopleware and more recently middleware) that it incorporates, and the impact which it has had at CERN, on particle physics in general, and on other sciences.

For medical reasons David Williams has been unable to complete the section on fifty years of CERN computing in time for publication with the other material. He hopes to complete this work for publication towards the end of 2006. In the mean while his *PowerPoint* slides are included here.

CERN-50
50 years of Research at CERN
From past to future (Computing)

Personal Reflections and Photographs

David Williams
CERN Academic Training Programme
16 September 2004

See
cern.ch/David.Williams/public/CERN50computingPub.ppt

Caveats

- I knew that this job would not be easy – it was even harder than I thought
- I stopped as Division Leader at the end of 1996, and I have had very little to do with physics computing since ~end 1998, which is two Internet generations.
- If you want to prepare a proper historical review (book) you should identify the themes, have them treated and analysed by specialists, and it would take a couple of years, and I might even try to do that one day
- Today I merely try to paint a picture, making some technical comments. Surely a personal picture, with all the biases that implies. And in no way am I trying to assign credit to individuals. I know that I will fail to mention some important aspects of the puzzle, and many people who have made important contributions
- I concentrate on the early years – you are living the later ones and don't need to be told about those! But I have not talked to two people that I should have talked to about the very early days (George Erskine and Henk Slettenhaar)
- The lighting of many of the photos is not wonderful, and many of the early ones are very "posed".
- We have tried an not always succeeded to assemble "collages" of many of our colleagues

The nature of the problem - People and their interactions

- I suspect that the LHC experiments are running at the limit of what is feasible…
- Not the amount of funds that can be assembled
- Or the complexity of the detectors
- But the possibility of keeping such a large number of very smart people working enthusiastically and actively towards a common scientific goal

- Until the mid-1980s HEP's "computing problem" was often thought to be about obtaining enough processor power
- Then we worried about storage capacity
- The real problem has always been, in my opinion, getting people to collaborate on a solution

Acknowledgements

- Miguel Marquina – who helped me to prepare this, and who has scanned a lot of photos for me (more than I am using)
- CERN Photolab
- Many others who have provided information and photos, and answered my questions
- All of the people – CERN people, CERN users, outside labs and computer suppliers staff – who did the real work behind fifty years of "Computing at CERN"
- The mistakes – and there will surely be several – are mine alone

Outline

- Setting the scene – the world in 1954 and 2004
- The early days – roughly to 1967
- Some early technology - punched cards
- Later Computer Centre machines
- Software
- Measuring Machines
- Online computing and DAQ
- Onsite Networking
- Offsite Networking
- Controlling Accelerators
- Various other things (Databases and other special applications, EU projects, Emulators, CERN School of Computing, HEP-CCC)
- Data Handling
- Information Handling
- Prospects for the Future
- Wim Klein

First thoughts

- Computing at CERN has little to do with Research into Computing
- It is the huge challenge of using leading-edge technologies to provide top quality services for the CERN community, so that they can prepare, run and process their experiments as well (and competitively) as possible

- Over time the physics challenges mainly stay conceptually the same, but the available technology – and its cost – keeps evolving
- Online feedback AND Worldwide access to data and information ➔ Worldwide participation in all phases of an experiment are the basic (enduring) challenges
- What did not work or was too complex – or was impossibly expensive – last year may well work – or be affordable – next year

SETTING THE SCENE
THE WORLD IN 1954 AND 2004

WHERE DID WE COME FROM?

1954

Europe was still recovering from World War II

Historically computing had been driven by the need for accurate compilation of tables – especially for navigation – (Babbage "mathematical engine") and for census purposes (Hollerith)

More recently (~previous 20 years) it had been largely driven by military needs – code breaking and bomb simulation

1954 world timeline – English Literature

- Lord of the Flies (Golding)
- Lord of the Rings Vol 1 - The Fellowship of the Ring - and Vol 2 - The Two Towers (Tolkien)
- Lucky Jim (Kingsley Amis)
- Under Milk Wood (Dylan Thomas)
- Under the Net (Murdoch)

1954 world timeline – General events

- First polio vaccination
- First kidney transplant
- First four minute mile (Bannister)
- Battle of Dien Bien Phu
- Algerian War of Independence starts
- General Nasser becomes prime minister of Egypt
- Bikini Atoll hydrogen bomb test
- First Russian hydrogen bomb test
- USS Nautilus launched
- Sen. McCarthy active (investigating Army etc.) but year ends with his condemnation by Senate vote
- US Supreme Court decision in *Oliver Brown v Board of Education of Topeka KA (and others)*

1954 world timeline - More

- Nobel Prizes
  - Physics — Born (quantum mechanics) and Bothe (coincidence circuit)
  - Chemistry — Pauling (chemical bond …)
  - Medicine — Enders, Welling and Robbins (for the cultivation of polio virus – leading to vaccines)
  - Literature — Hemingway
  - Peace — UNHCR
- Born
  - Cherie Blair
  - Condoleezza Rice
- Died
  - Alan Turing
  - Enrico Fermi

NATIONAL LEADERS

"Who wants to be a millionaire"

1954 political leaders of Russia, USA, UK, France, Germany??

First Secretary of the CP of the Soviet Union

Photo: Karsh

Rene Coty (President from 16 Jan)

Joseph Laniel
(PM until June 18)

Pierre Mendes-France (PM from June 18)

Photo: Haus der Geschichte der BRD, Bonn

MARRIED
AND DIVORCED IN 1954

Married 14 Jan

Divorced 5 Oct

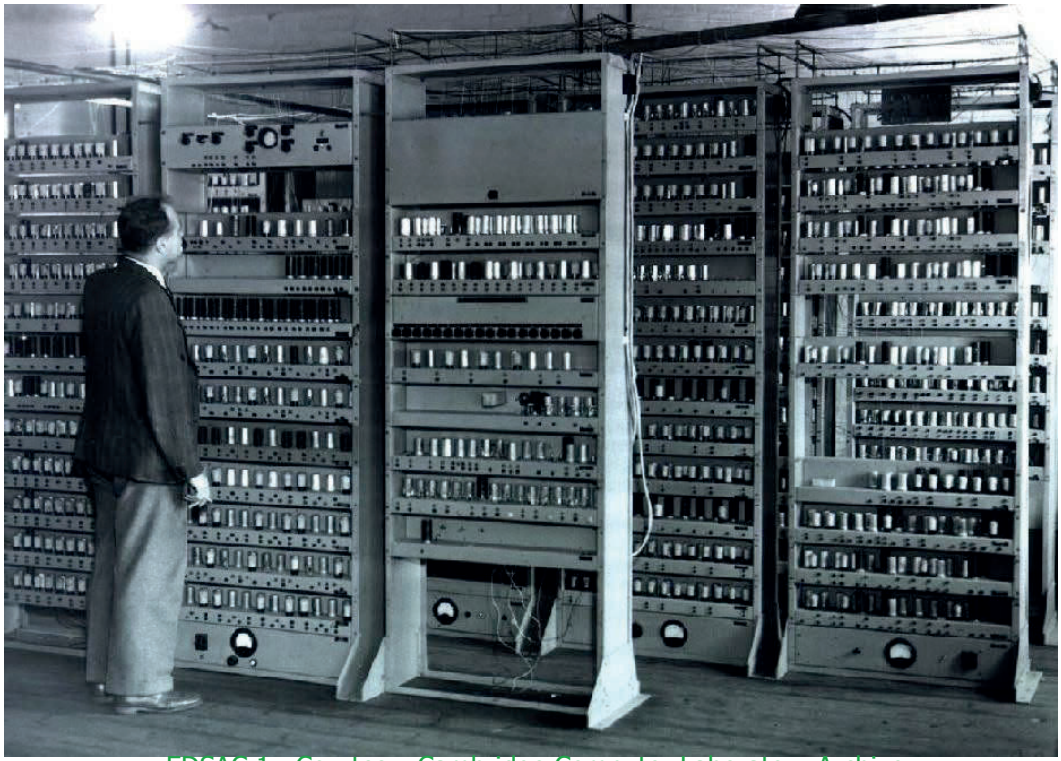
1954 computing timeline

- Computers were valve-based
- I was 10 years old and had never seen a computer. I first saw one, a Stantec Zebra, in ~1956, and then started taking out the ~3 books in the local library about computing.
- I saw EDSAC 1 in Cambridge in ~1958 as it was being dismantled.

Built in Newport by STC from the original concept of van der Poel (Delft)
8k 33-bit words. Valve-based. Cost 23 kGBP (then ~280 kCHF)
and was one of the cheapest general-purpose machines of the time

Note: Fantastic Paper-tape equipment and telegraph-style typewriter

Stantec Zebra



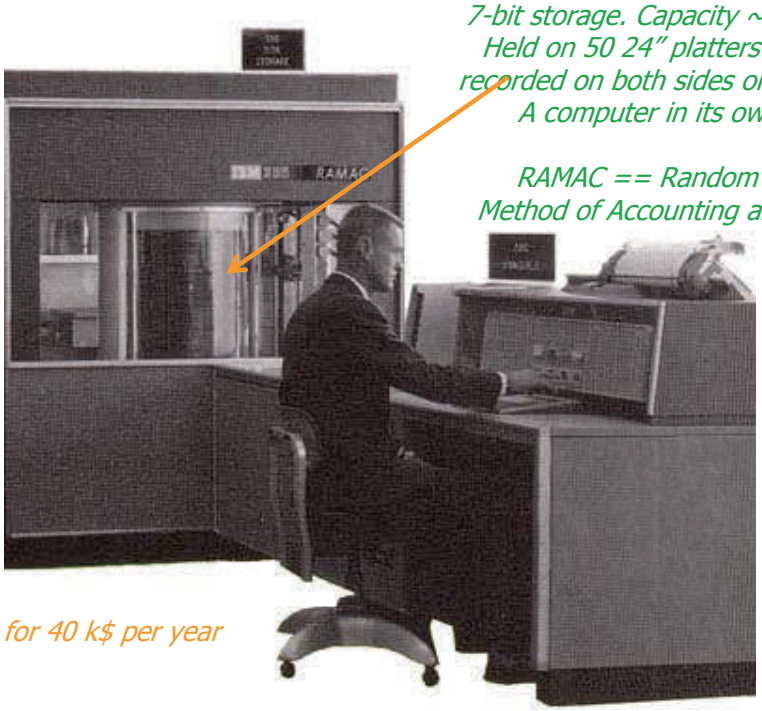
EDSAC 1. Courtesy, Cambridge Computer Laboratory Archive

More 50s and 60s computing timeline

1947 (Dec)	Point contact transistor invented
1951 (Sep)	Major Bell Labs symposium on working junction transistors
1955	Wilkes invents microprogramming – programming the instruction set
1956	First magnetic disk system sold (IBM RAMAC)
~1956	FORTRAN under development
1959	IBM 1401 first shipped. Transistorised. Punched card input. 12,000 sold over 12 years
1960	PDP-1 launched (18 bit words)
1964	PDP-8 launched (12 bit words))
1964	System/360 launched (4*8 bit byte words)

IBM RAMAC disk – 1956
7-bit storage. Capacity ~4.4 Mbytes
Held on 50 24" platters with data recorded on both sides on 100 tracks
A computer in its own right

RAMAC == Random Access Method of Accounting and Control



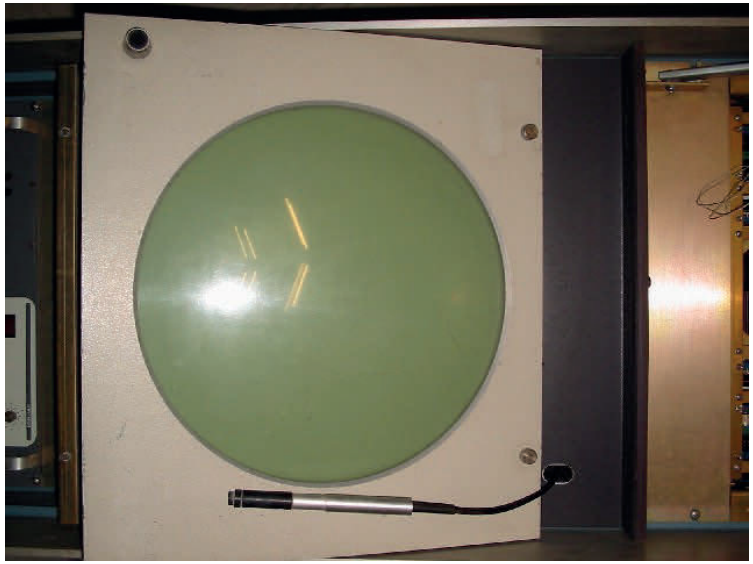
Leased for 40 k$ per year

Personal timeline

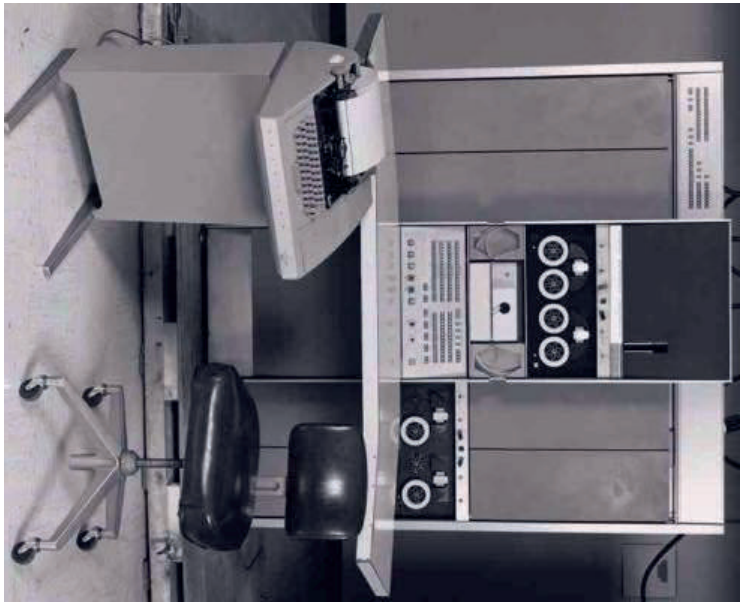
August 1966 Williams to CERN

- I that time I had programmed the Titan (prototype Ferranti Atlas-2) in Autocode and a DEC (Digital Equipment) PDP-7 – a follow-on machine to the PDP-1 – equipped with an interactive display, in assembler.
- A multi-user service and a stand-alone "mini-computer" which you could work on alone in the evening

DEC 340 Display (attached to PDP7)
with light-pen



Courtesy: Martin Pool



Note: Chair! Teletype (KSR-33), PTR
and DECtape (~256 kB/reel)

2004

Longest period of peace in (much of) Europe since ever

Not intended as an exhaustive list

- Napoleonic Wars 1803-1815
- Belgian War of Independence 1830-1832
- Italian Wars of Independence 1848-1866
- Crimean War 1853-1856
- Danish-Prussian War 1864
- Austro-Prussian War 1866
- Franco-Prussian War 1870-1871
- World War I 1914-1918
- World War II 1939-1945

van Hove

END OF SCENE SETTING

DOWN TO WORK!

THE EARLY DAYS

ROUGHLY TO 1967

From CERN annual reports
Mainly verbatim quotes
(but I did *some* paraphrasing)

1956/57

- 1956
  - Following in the steps of Harwell and Saclay … CERN ordered an electronic computer of a new type (Ferranti Mercury) in May 1956. Because of unforeseen difficulties experienced by the producers of this machine, however, there will undoubtedly be some considerable delay in its delivery …
  - An experienced mathematician-physicist has been recruited to run the future Computer Section. Further staff will eventually be recruited, but the section will not be very large since …
- 1957
  - The Mercury computer will probably not be installed until the summer of 1958
  - … prepare programmes for use on the English Electric Deuce and the IBM 704 in Paris
  - By the end of 1957 there were 2 staff plus 1 fellow in the Computer Section

1958-59

- 1958
  - The Ferranti Mercury computer arrived in the summer and its acceptance tests were completed by mid-October.
  - … Autocode is fairly easily learnt by scientists who have had no experience of computers
  - *(referring to the IBM 704 in Paris)* in each case the programming was done in Fortran
  - The speeds at which the available paper-tape devices can transmit data to and from the computer are in no way comparable with the computing speed
  - Leading personnel … took part in the "Symposium on the Mechanisation of Thought" at NPL Teddington in November
  - The staff of the Computing Group, numbering about 10 at the end of 1958, will have to be doubled in the course of 1959
- 1959
  - … to supplement and eventually replace the present Mercury computer by hiring an IBM 704 from the latter part of 1960. The installation will be equipped with an 8'000 word core store (32 kB) 8 tape units and various other ancillary equipment. It is due to start operation in the autumn of 1960
  - In October regular two-shift working was introduced on the Mercury … Nevertheless the is already a backlog of computing work.
  - The computer service has been used by about 40 customers, who wrote more than a hundred Autocode programmes

Mercury: note paper-tape equipment and real online typewriter

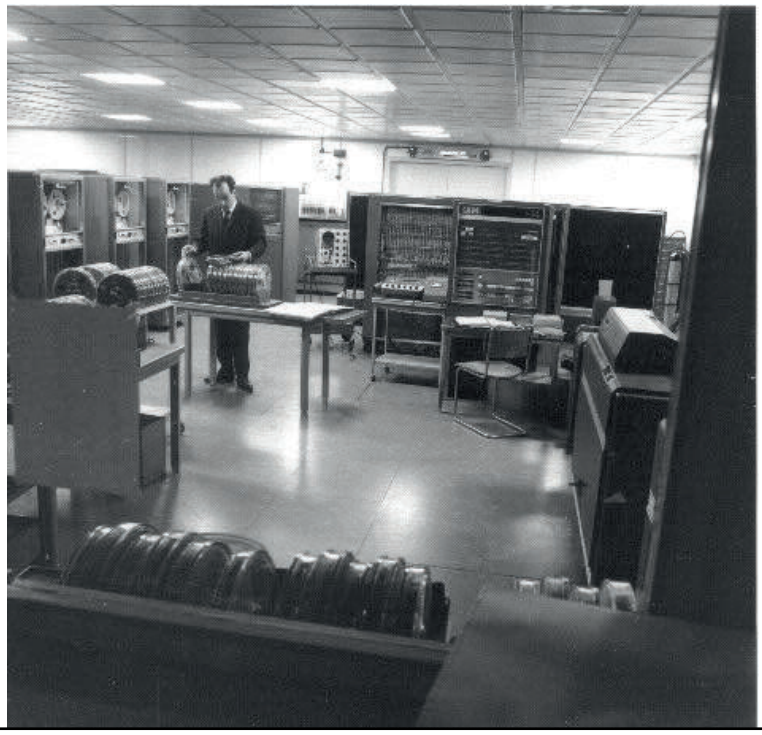
Auffret and Slettenhaar

Ferranti Mercury

1960-61

- 1960
  - In early November a much larger and faster computer, the IBM 709, was delivered to the site and in 1961 seven analysing machines should be in operation using the IBM 709 for data processing
  - To house it a prefabricated building has been put up with all necessary air-conditioning and stabilized power supplies
  - Full time operation of the Mercury was introduced in April 1960. At week-ends (on Saturdays and sometimes on Sundays) it has been possible to work on one or more shifts, and do some maintenance and upgrades
  - First mention of Flying Spot Digitisers (Hough and Powell)
- 1961
  - It became clear during the year that even more attention will have to be given to data handling and analysis, that as … …, will require more and more machine time from computers with faster operation and larger memories

IBM 709



PDG of IBM France
Baron de Waldner at
the official opening?

1962 - general

- 1962
  - The importance of data handling as a central problem for the laboratory has become even more obvious with the growing flood of bubble chamber pictures and the beginning of what is likely to be a comparable flow of spark chamber data
  - The 709 worked an average of 1.5 shifts during the year
  - The Mercury ran 24 hours*5 days throughout the year, with occasional running at week-ends
  - A system … is being written to run (Mercury) Autocode programmes on the IBM 709!
  - The IBM 1401 computer was delivered in November (for I/O handling). CERN's third computer? (and first transistorised machine)

1962 - the "offline chain"

- 1962
  - A new generation of programmes for the analysis of measurements of bubble chamber photographs was brought into use. The new programmes make use of the greater speed of the IBM 709 computer, …, they are written in Fortran …
  - Paper tape measurements read by REAP (Mercury) ➔ 709 via magtape
  - Geometry reconstruction by THRESH
  - Kinematic analysis by GRIND (which later needed an "ignore infinity" card on the CDC 6600!)

From the Electronics Experiments Committee 19 Sept 1962

- Preiswerk (chair), Cassells, Rubbia, Wetherell, Dick
- The situation … with respect to the analysis of spark chamber photographs was briefly discussed. With the influx of 10^5 pictures from the PS diboson experiment S1 and … it is clear that the available facilities are rapidly becoming overloaded. A general discussion of the whole picture analysis problem would be forthcoming from Hine's study group.

  A possible and untapped source of effort might be available from universities. Participation by university personnel in the data taking stage would be essential.

1963

- By September the IBM 709 computer was replaced by a transistorised IBM 7090, increasing by about a factor 4 the total computing capacity available at CERN.
- On the basis of the report of the "European Committee on the Future Computing Needs of CERN", the DG proposed to the Council the purchase of a large time-sharing computer – the CDC 6600 – to replace CERN's present computer by early 1965.
- The computing capacity … estimated to be at least 15x CERN's present capacity) will allow considerable development of data handling techniques used in HEP. It is planned to exploit fully the time-sharing properties of the computer in order to allow simultaneous operation of various on-line applications (counter experiments, film measuring devices etc.) together with the more conventional computing work of the Laboratory. A great, but unfortunately not very realistic, vision.

IBM 7090 – October 1963

Hans Klein and Swoboda (IBM)

More 1963

- REAP moved from Mercury to 709
- BAKE-SLICE-SUMX introduced
- THRESH for heavy liquids being developed
- Manuals written!
- A four-week lecture course in Oct-Nov on the programme chain was attended by 120 physicists from all over the world
- Preparations to use the Mercury online with a sonic spark chamber experiment

1964

- The CDC 6600 will be delivered in Jan 1965 and the change-over from the IBM 7090 is planned to take 3 months
- By the end of the year the 7090 was operating on 24*7 basis, processing an average of 350 jobs/day.
- Second IBM 1401 installed
- General Mercury service stopped in April and machine moved to online work. 1 km data link to the South Hall (missing mass) and similar link to SC vidicon experiment.
- Sonic spark chamber analysis programmes were developed for an experiment using the SDS 920 computer online.
- Programme Library started – with 80 general programmes
- Definition of CERN Fortran (to provide compatibility across machines)
- Detailed proposal being made to connect several IEPs to the CDC 6600
- New standard interface defined with CDC to enable HPDs to be attached directly to 6600

From the Electronics Experiments Committee 13 January 1964

- Puppi (chair), Heintze, Lock, Munday, Preiswerk, Sens, Winter
- Data Handling Facilities

  Several experiments experience severe difficulties in view of the limitations on the available data handling facilities. The EEC recommends as a general policy that experiments which have priority on the accelerators also have priority on the data handling facilities. It is considered of the greatest importance to organize the data analysis in such a way that a feed-back to the data taking phase of the experiment is possible.

  (i.e. you want online feedback!)

From March 1964 Informal Meeting on Film-Less Spark Chamber Techniques and Associated Computer Use

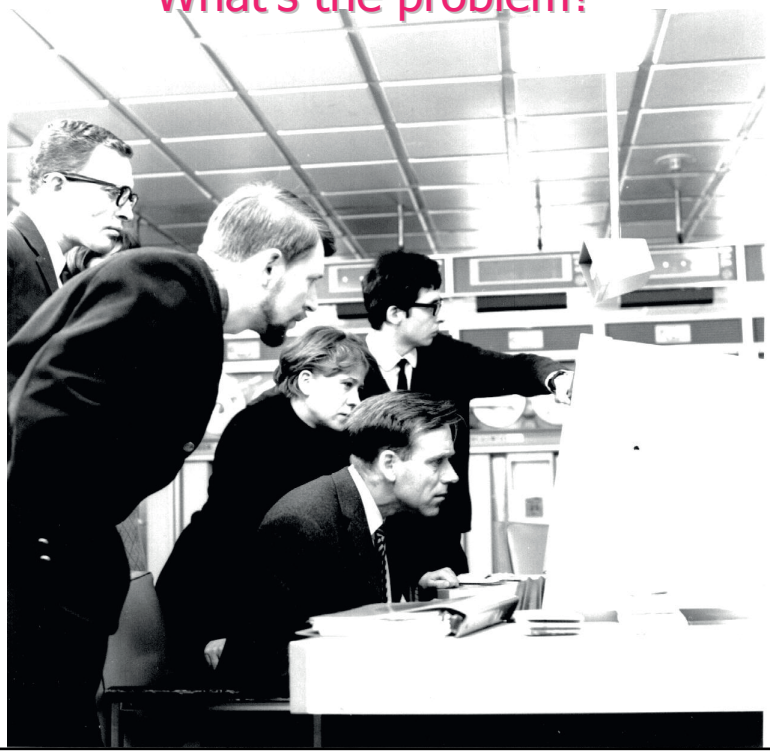
- General Discussion on Online Computer Use (led by Macleod)
- As far as online computers are concerned, there are those who are in favour and those who are against, and, by and large these appear to correspond to those who have online computers and those who do not! There is nobody who has been near an on-line computer who has said that he did not like it.
- If I can start with Lindenbaum, he is essentially in favour of everything. He is not reticent about it, and he said that he thinks that his kind of experiment could use a 6600 full-time plus any other computers one can find on the East Coast!

1965

- The IBM 7090 computer was completely overloaded from the beginning of the year and it was necessary to process work away from CERN until June
- The CDC 6600 was delivered in January, but delivery delays with the SIPROS time-sharing operating system, and technical problems with the hardware itself, prevented operation at anything like the planned capacity
- The IBM lease was terminated at the end of July
- Work on SIPROS was concentrated at CERN. The work is making good progress and it is planned to introduce SIPROS in January 1966
- Hardware reliability problems led to 4-week overhaul in mid-October – when I think that it was the only computer onsite (apart from the Mercury)

CDC 6600 (console, card readers, line printers)

What's the problem?



Lipps, Trenel, Deller and two CDC analysts

1966

- Most of the work of DD Division was centred around the difficult task of bringing the main computer, the CDC 6600, into full and reliable service, which was largely accomplished, with some temporary set-backs, by the end of the year
- Control Data installed a 3400 to take some of the load, upgraded to 3800 in August. Planned to upgrade 3800 to 6500 in spring 1967
- It was recognised that computing had become such an integral part of the scientific work of CERN that an interruption of even a day or so due to computer breakdown caused serious disruption to the work of the Laboratory
- Tender for the first ADP computer
- HPD1 and Luciole online to 6600, but running at only ~60% of speed available on a dedicated 7090
- Two IEPs online to 6600, but poor performance (memory) led to dropping that solution, and CDC 3100 was installed to control IEPs in August
- Data link SDS 920 – CDC 6600 used up to 5 hours/day, planned to lead to FOCUS. Another data link to an IBM 1800

CDC 3800

Bernard Petiot at the console

1967

- Experience confirmed that the running-in problems of 1965 and 1966 had been overcome
- CDC 6400-6500 installed and by end of 1967 both machines were running CERN SCOPE (i.e. not SIPROS, which was dead) and drum and large disk had been added.
- Jobs rose to 5'800/week, from ~400 users
- 10'000 tapes in the tape library
- FOCUS system started on CDC 3100. Intended primarily for a limited number of experiments at the PS using small online computers which will be connected by datalinks
- Graphics on CDC 3100 (display with light-pen)

CDC 6500

AND A SUMMARY OF ALL OF THAT?

Summary (1/2)

- CERN had moved from one valve-based Mercury computer in October 1958 via the IBM 709 (Nov 1960-Sept 1963), then the transistorised 7090 (Sept 1963-July 1965) and CDC 6600 (from Jan 1965) to maybe twenty transistorised computers in 1967
- CDC 3100s hovering around the 6600/6400
- Several computers controlling film-measurement machines
- And others controlling experiments online
- And starting on real special-purpose activity like ADP and accelerator control

Summary (2/2)

- We had learned the hard way that computing is an integral part of the life of a scientific laboratory
- That software and hardware reliability are vital
- And that the technology was not yet ready to handle efficiently the mix of a general-purpose computing load with devices such as IEPs and HPDs on the same machine
- The long lead-times between order and delivery (and for acceptance testing) come from another world
- We were also learning to balance expenditure between the big central computer and the computers at the experiments (not really suitable for full-blown analysis codes)

SOME EARLY TECHNOLOGY

PUNCHED CARDS

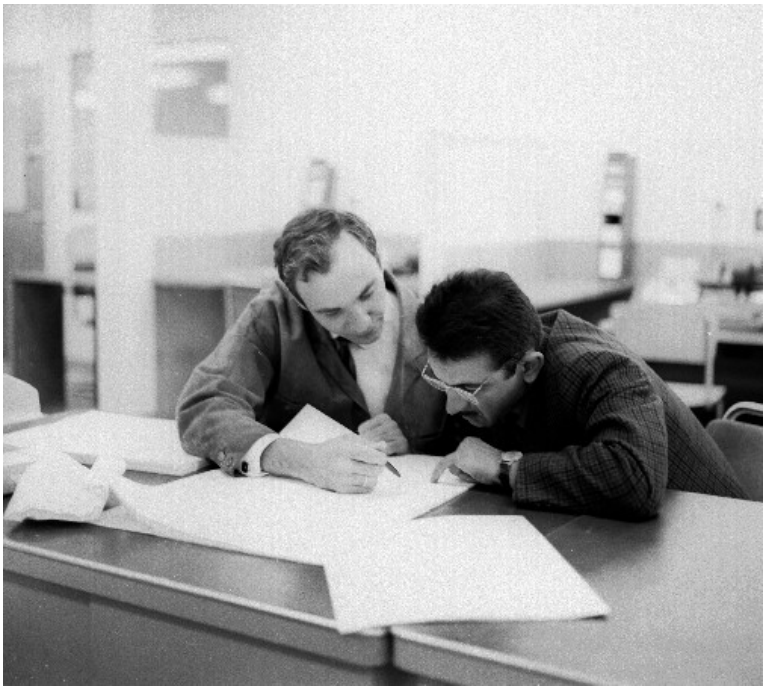
Card copying (and listing)

Plug boards on right
Also card supply

Magnified enough you can spot card decks from Aguilar, Dufey, Gavillet, Kleinknecht, Lindelof and Salmeron

Card readers in real use

Looking for the bug



LATER COMPUTER CENTRE MACHINES

Ferranti Mercury ➔ IBM 709/7090 ➔
CDC 6600 (et al) ➔ CDC 7600 (et al)
together with IBM 195/3081/3090 and Siemens➔
HOPE/CSF/SHIFT ➔ Servers
together with VAX 780/750/8600

The location of the "Computer Centre"

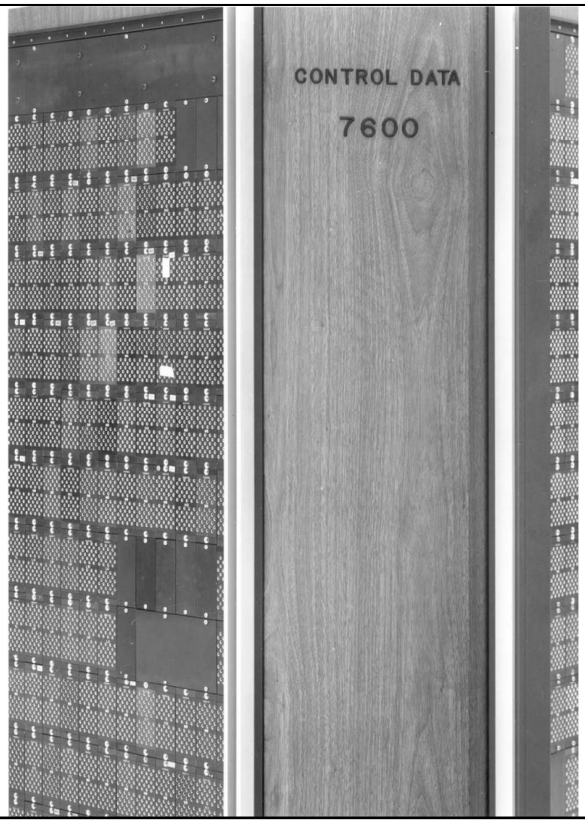
- The Mercury was installed on the ground floor of Building 1 – on the Jura side – in what are now ATLAS offices. You can find the location by looking for the bell which is still there
- I'm not sure exactly where the IBM 709/7090 was installed – I suspect close to but not exactly where the CDC was subsequently installed
- The CDC 6600 plus 6500 was installed where the Print Shop is now
- B513 was constructed during 1971 and the machines were moved there during 1972 (and maybe early 1973)

B513 in construction in 1971

The Move to B513

- I have always believed, with the wisdom of hindsight, that it was a huge mistake that the new Computer Centre was moved so far way in 1972
- It took almost all DD/CN/IT staff away from the natural centre of the lab, and made easy integration with the experiments much harder
- Of course, I must recognise that there was an immense amount of good vision behind the move too, and we certainly are benefiting from the space as we move towards LHC computing
- IMO just in the wrong place – we built B32 and the hostels much closer to CERN's natural centre afterwards

B513 – re-installation of 6600 in 1972



CDC 7600 wiring

IBM 370/168 - the CERN Unit

Siemens 7880 in 1985

CSF in 1992

VAX Area in 1993

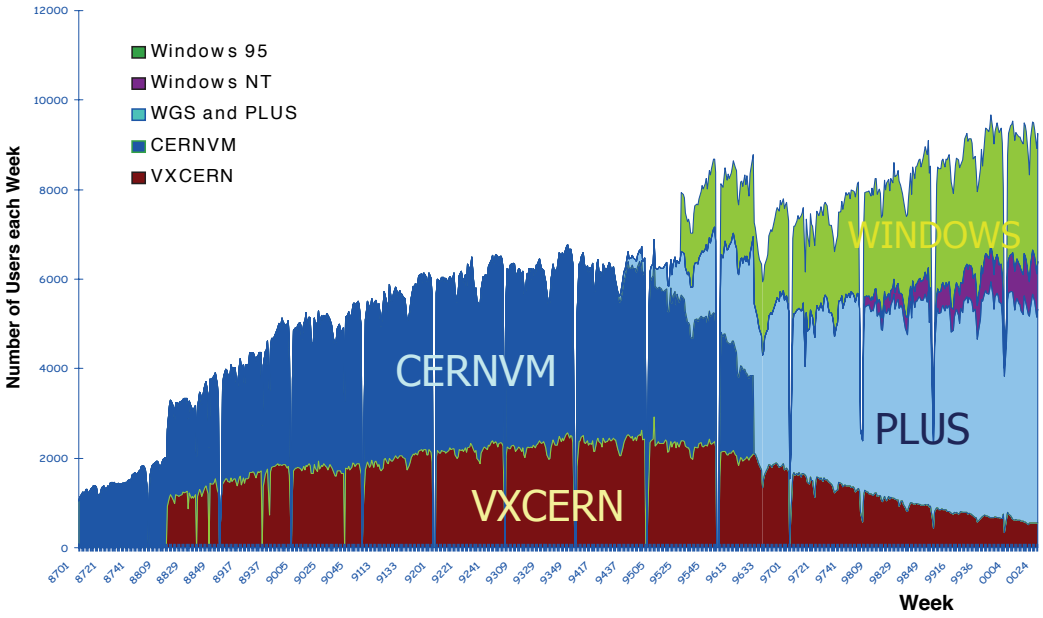
Cray X/MP-48 (1988-1993)

Servers (NICE etc.)

SHIFT

CERN ACTS AS THE NATURAL CENTRE FOR A LARGE COMMUNITY

Weekly interactive users 1987-2000



SOFTWARE

Does Moore's Law only apply to hardware?

Did we make any progress with software?

- Answer has to be YES
- In 1960 an experiment involved (I guess) several hundreds, up to a few thousands, of lines of code, running on one single computer, and which had been written by no more than ~3 programmers
- In 2000-2005 an LHC experiment involves several tens of millions of l.o.c., running in 1000-10'000 processors, and written by at least ~300 programmers

- So, over 40 years, 10^4-10^5-10^6 more l.o.c, involving O(100)x more authors

Delivering software to the end-user

- We (fortunately!) no longer need to duplicate card decks or magnetic tape to deliver software to colleagues
- We deliver it over the network - to the workstation
- If we are sharing it on a professional basis it can be configured to some agreed specification
- The world of CERNlib, ASIS, AFS, CVS, SRT, ….

- The steady improvements here helped to deliver several really major packages which changed the face of physics computing ...

The "offline programmers"

- (Macleod, YGC†), Bock, Zoll†, Kellner, Pagiola, Burmeister, Bruyant, Norton
- Brun, Palazzi, Grote, Metcalf, Onions
- Giani, Apostolakis

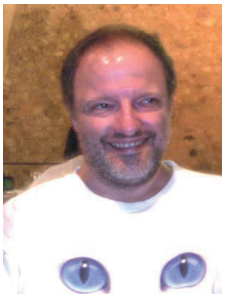




The Program Librarians

James, Renshall, Carminati, Marquina, McLaren





Languages and Operating Systems

- Many software engineers thought that physicists were renegades and would only ever write in FORTRAN? But things moved.
- F-77, and also C (portable assembler) and C++ (object orientation)
- Basic ➔ Java for interactivity
- Not to forget scripting languages – primarily to interact with OSs

- SCOPE, MVS+Wylbur, VM/CMS, VMS, Windows, Unix, Linux
- Digital and Norsk Data died for not recognising Unix quickly enough
- And in the future …?
- What role for Open Source?

Some important CERN software suites

- REAP, THRESH, GRIND, SLICE
- SUMX and then HBOOK
- PATCHY and ZEBRA
- GEANT3 and then GEANT4

- Leading onto PAW, ROOT, POOL and all the other software that will form the basis of LHC analysis software

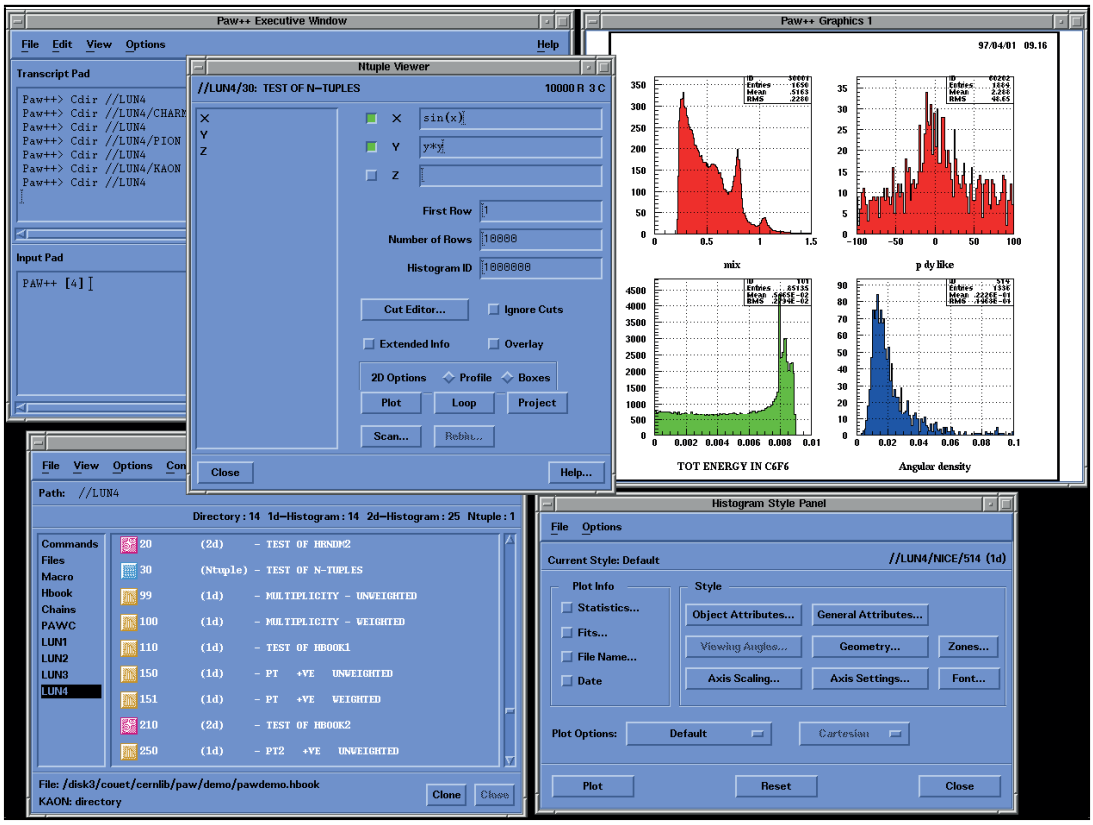
- Not to forget things like the AIS software, EDMS, MAD and other accelerator design packages, etc. etc.

PAW

- Discussed in MEDDLE at the September 1986 meeting (first meeting with John Thresher present as responsible Director)
- Towards a Physics Analysis Workstation (December 1985)
  - Bock, Brun, Pape and Revol
- As the world moved to graphics and interaction, trying to bring "everything" together to create a convenient interactive interface for the physicist doing analysis
- A real *de facto* standard and huge success

René Brun presentation of PAW at 1April 1989
CHEP in Oxford

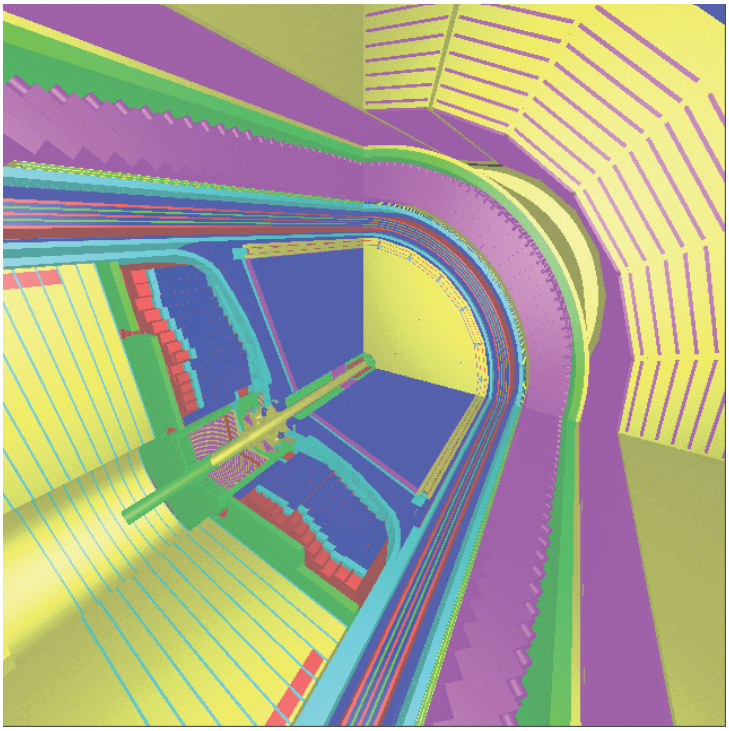
PAW demonstration at Oxford CHEP, April 1989



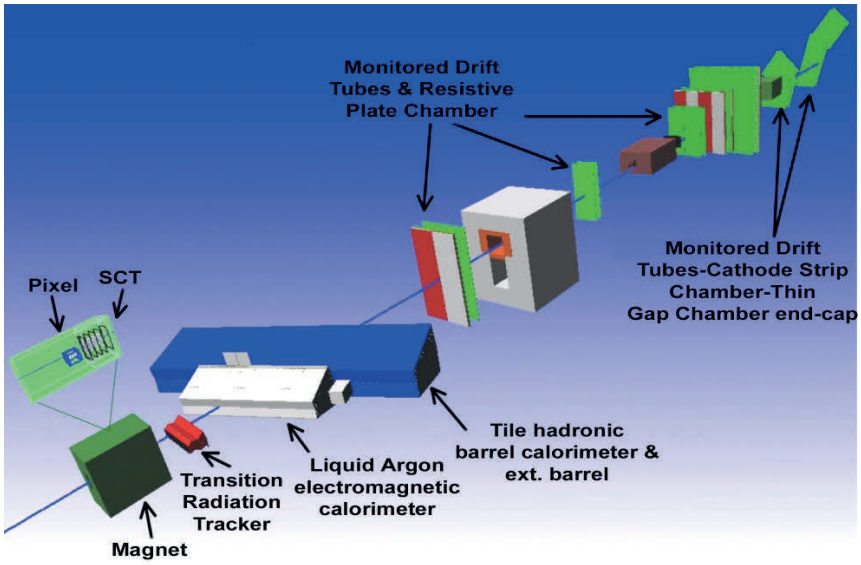
GEANT

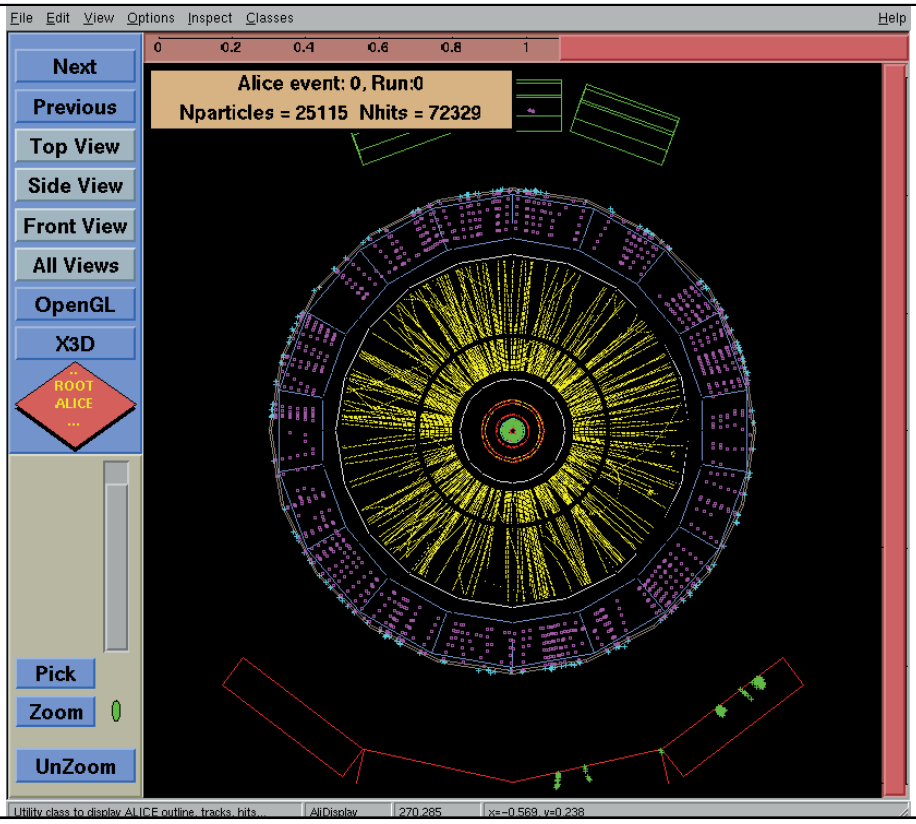
- The basis of much (not all) of CERN's physics simulation work
- A series of releases aimed at increasing functionality and reliability …
- Of which 3.14 (early days), 3.15 (end '91) and 3.21 (Spring '94) were some of the most important for LEP
- Effort then evolved/switched to GEANT4 (OO)

OPAL
GEANT 3.21



GEANT4 (direct) graphic of ATLAS beamline test





ALICE ROOT display

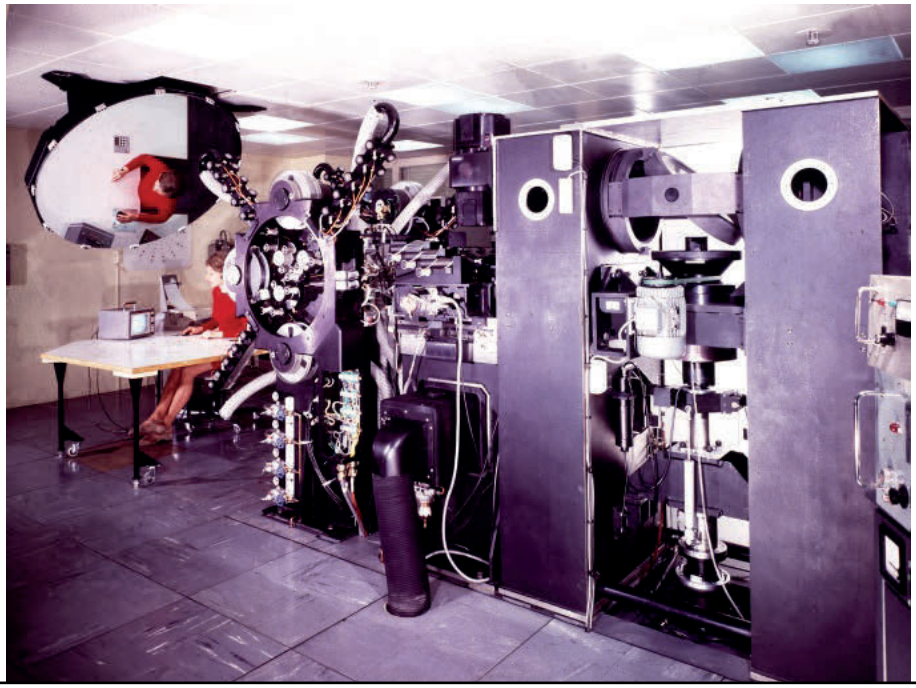
MEASURING MACHINES

As we have seen these formed a critical part of the early computing load

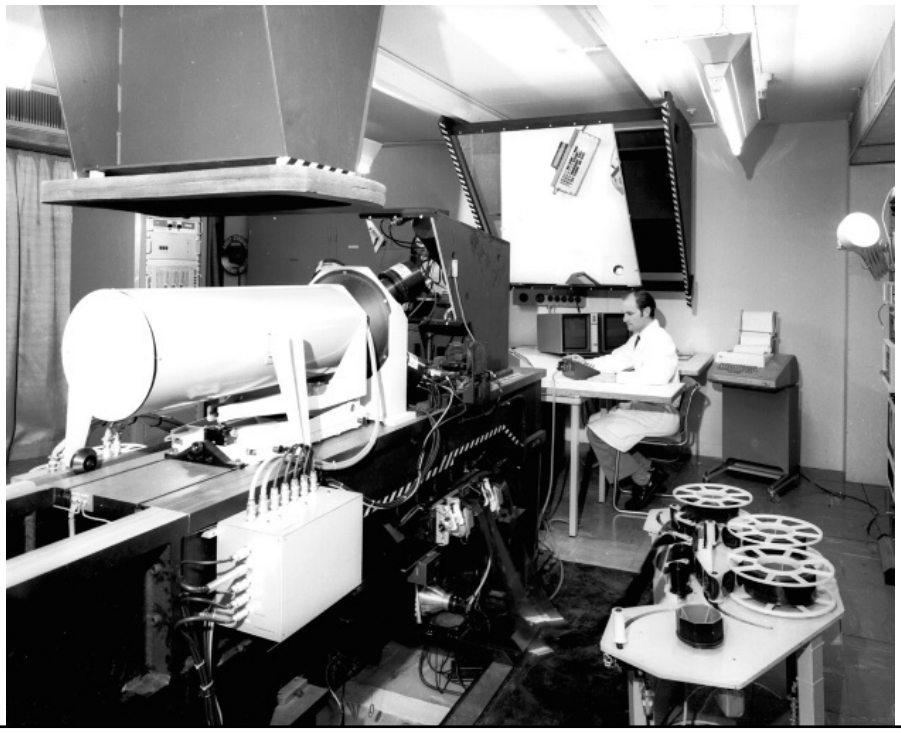
Multiple versions

- IEPs were needed for many different films and chambers (both bubble- and spark-)
- Initially punching paper-tape – later controlled by computer
- Then machines with more/less automation and different scanning technologies. FSDs (mechanically generated spot scans), Spiral Readers (spiralling slit scan starting from vertex) and CRT scanners (Luciole, PEPR, ERASME)
- Until the arrival of machines of the PDP-6/10 class there was a continuing problem about how much programmed feed-back could be made available to the operator – not enough memory space & floating point power to run real reconstruction.
- But lots of good mechanical and electronic engineering

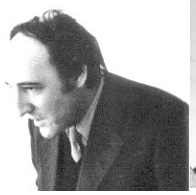
Spiral Reader - 1969

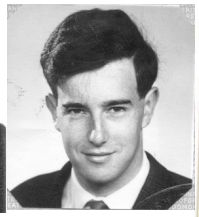


ERASME - 1971

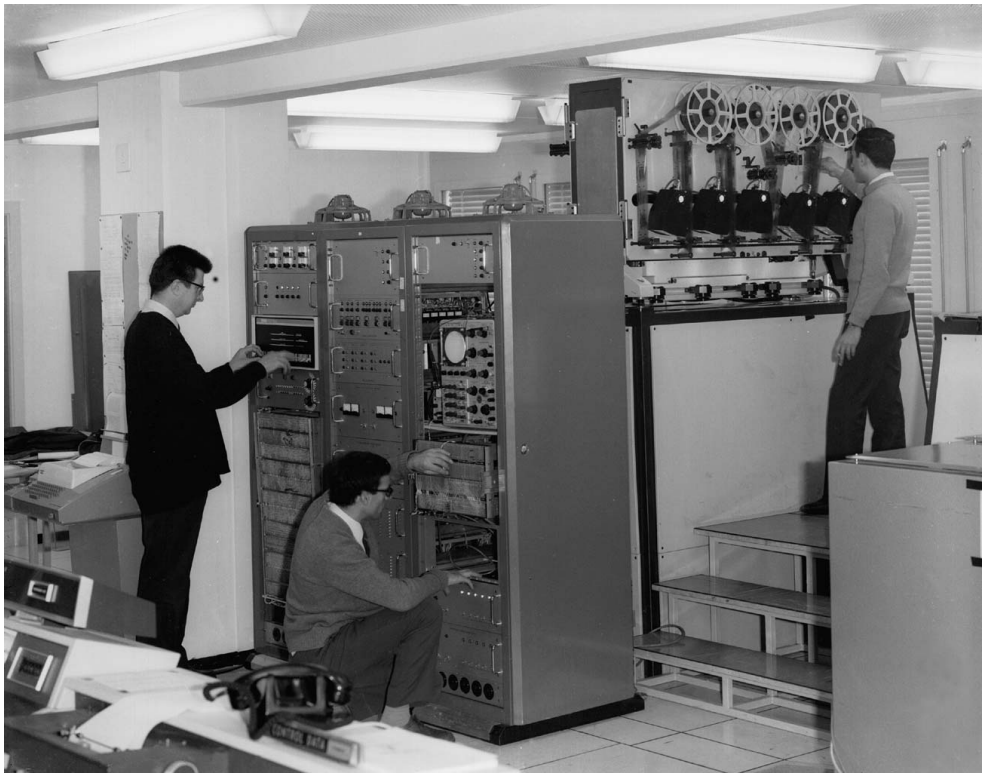


Hough and Powell





Various HPD (bubble chamber) programmers and users:-
Moorhead, Sambles, Joosten, Evershed, Ferran (2x), Quercigh, Celnikier, Morrison. Missing: Gerard, Lord, Altaber, Antonioz-Blanc, Howie, French, Atherton



Rinus Verkerk, Adolfo Fucci and one other (kneeling)

ONLINE COMPUTING AND DAQ

The online connection of CERN experiments
to the Computer Centre had a long history
(as we have already seen)

IBM 360/44

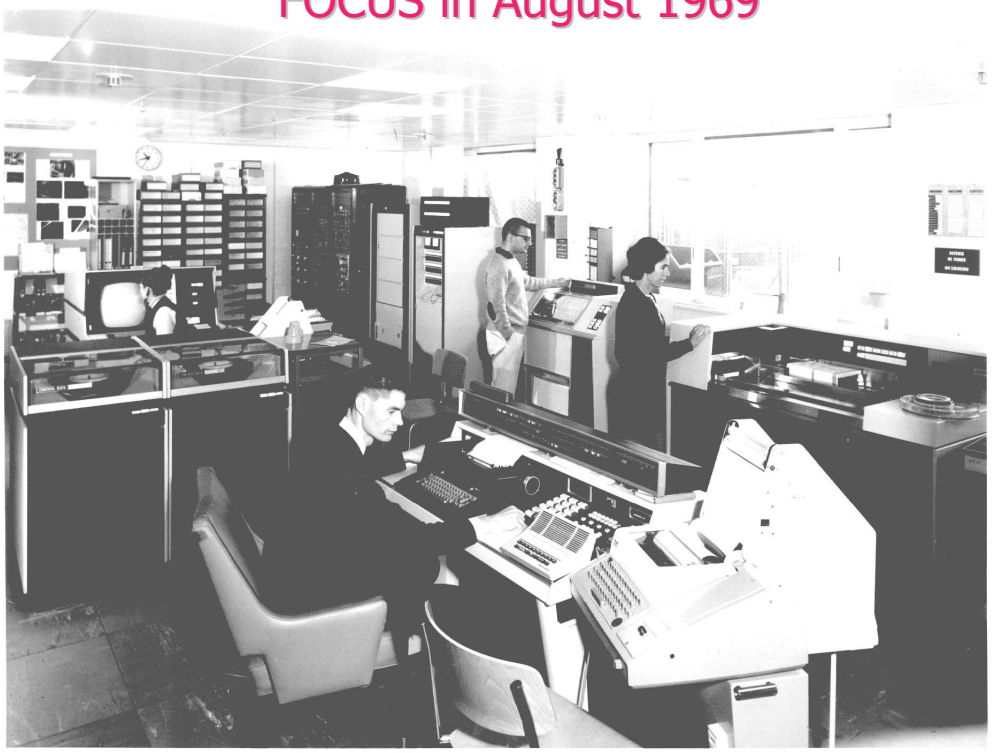
Per Scharff-Hansen at the console

OMEGA CII 10070

ONSITE NETWORKING

Focus
RIOS
Terminals
CERNET
The world of TCP/IP

FOCUS in August 1969



Yule at console; Gerard-Weights standing on right

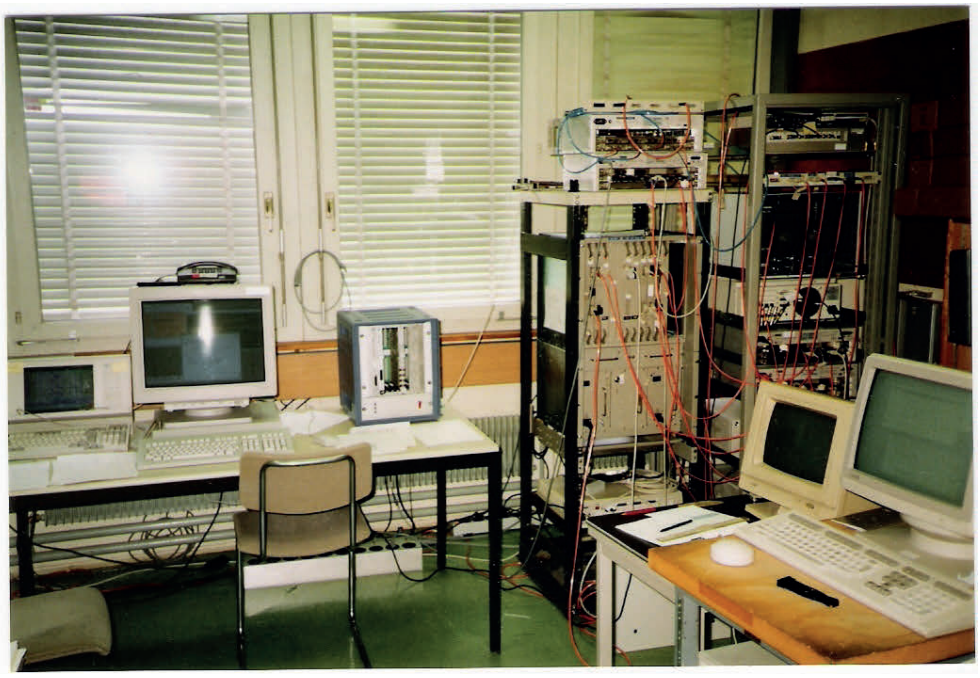
1972 – a proto-RIOS in B112 (West Area)

Michael Metcalf, Anton Frölich, and an unknown person at the printer

UA1 analysis?
Note the Gandalf terminal switches

CERNET Modcomp team

Gigaswitch interoperability tests



OFFSITE NETWORKING

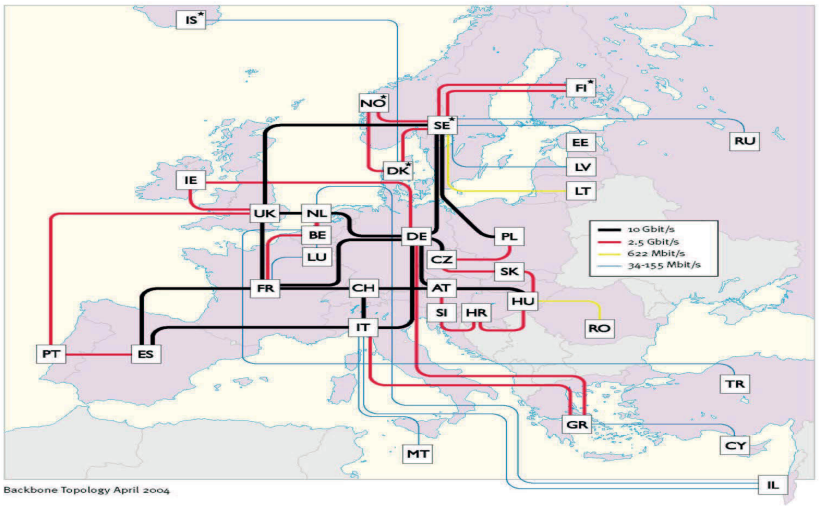
Broadly speaking we have moved from external connectivity at kbps in the early 1980s to Mbps in the early 1990s to Gbps in the early 2000s. By LHC start-up or soon after we are likely to reach a total of 100 Gbps. We will need that to handle the grid data flows.

Francois Flűckiger connecting to Transpac?

Lighting the way to the European Research Area

GÉANT: The Multi-Gigabit pan-European Research Network

GEANT is managed by DANTE, The pan-European Research Networking Operator



Backbone Topology April 2004

AT Austria · BE Belgium · CH Switzerland · CY Cyprus · CZ Czech Republic · DE Germany · DK Denmark* · EE Estonia · ES Spain · FI Finland* · FR France · GR Greece · HR Croatia · HU Hungary · IE Ireland · IL Israel · IS Iceland* · IT Italy · LT Lithuania · LU Luxembourg · LV Latvia · MT Malta · NL Netherlands · NO Norway* · PL Poland · PT Portugal · RO Romania · RU Russia · SE Sweden* · SI Slovenia · SK Slovakia · TR Turkey · UK United Kingdom

* Connections between these countries are part of NORDUnet (the Nordic regional network)

GÉANT is co-funded by The European Commission within its 5th R&D Framework programme

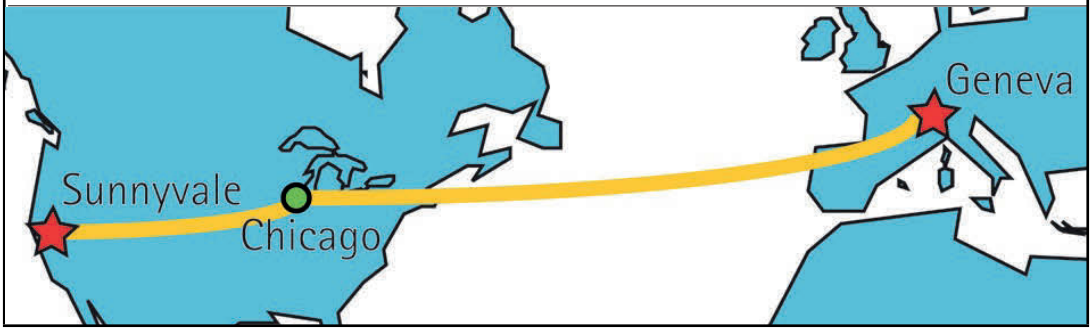
Information Society Technologies

Contract No. IST-2000-26417

Internet2 Land Speed Record IPv4 Single and Multiple Streams

- Raw Metrics
  - 27 February 2003
  - 1.1 terabytes
  - 1 hour, 1 minute, 40 seconds
  - 10,037 kilometers
- Results
  - 2.38 gigabits per second
  - Sunnyvale to Geneva

  ....by way Chicago
  - 23,888.06 terabit-meters/second



Record breaking team

CONTROLLING ACCELERATORS

Control systems play a crucial role, especially during the start-up of new accelerators and their subsequent operation

The SPS control system was responsible for several important computing advances (though the PS always had stronger real-time constraints). And the Nord story is an important part of CERN's European TT history.

I could find more SPS than PS pictures

Michael Crowley-Milling

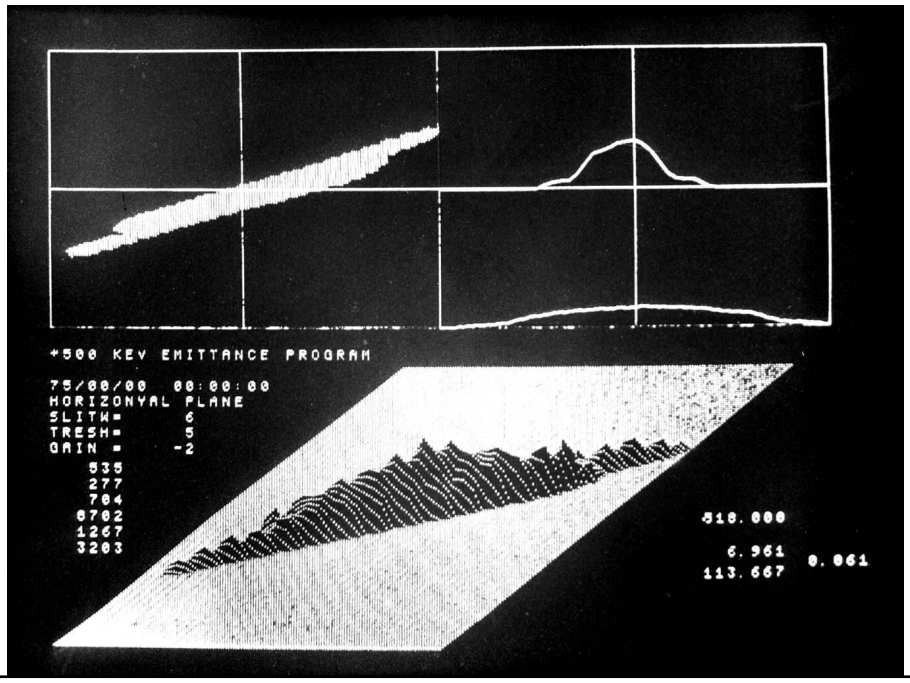
Nord-10s for the SPS controls

Raymond Rausch

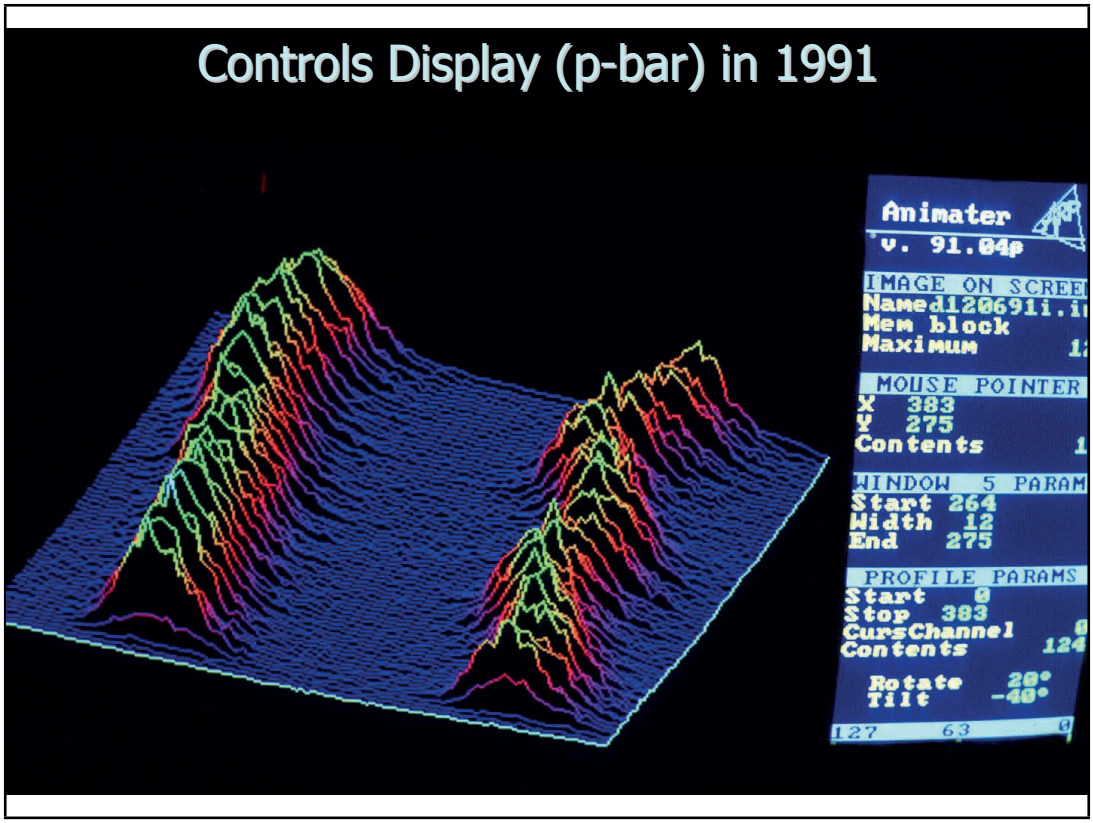
SPS console

Bent Stumpe at console

Controls Display in 1976



Controls Display (p-bar) in 1991



VARIOUS OTHER THINGS

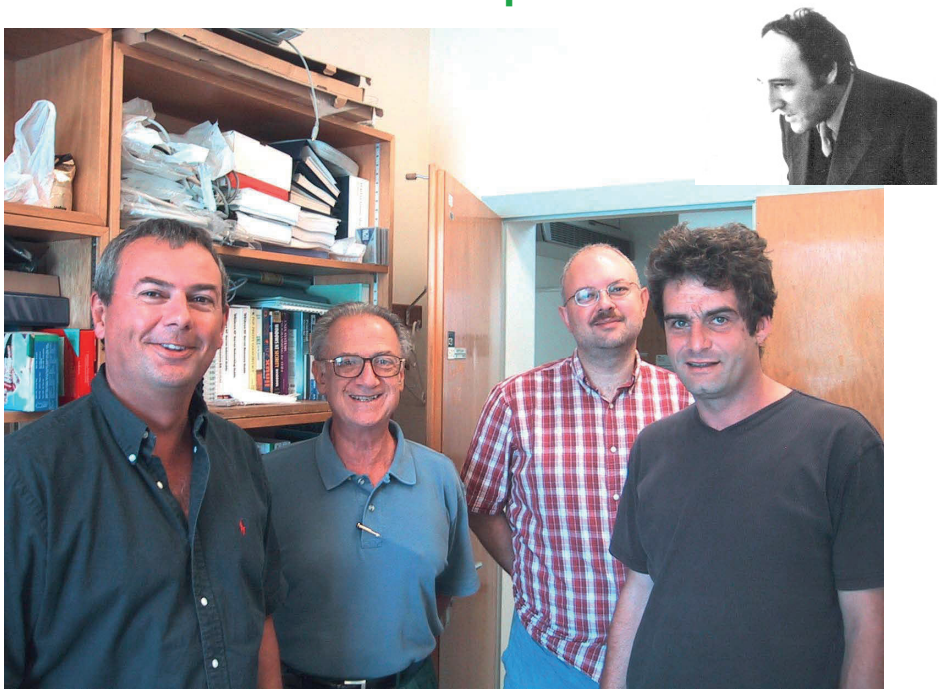
ADP (and then AIS)
EU-funded computing activity
Griddery
Control systems at experiments
And still many other things

ADP 360/30 in March 1968

ADP 360/50 in April 1971

Note the very early model IBM 3330 disks
Lionel Grosset (invisible) at the console.
Jean Gros standing

IT database specialists



Shiers, Santiago, Segura, Dűllmann (Moorhead inset)

EU-funded computing activities

- Bob Dobinson was a pioneer and was involved in several different projects (especially transputers and advanced networking)
- Meiko CS-2 was a European example of the classic US "place a supercomputer at a national lab and they will do something useful with it" approach. In our case we used the Meiko to run NA-48 data taking etc. etc.
- And since 2001 DataGrid and EGEE (Fabrizio Gagliardi)

- Not to forget the partial support for direct CERN connectivity to GEANT (we share the Swiss connection with the Swiss national network SWITCH)
- And DataTAG and others

Grids

- See next

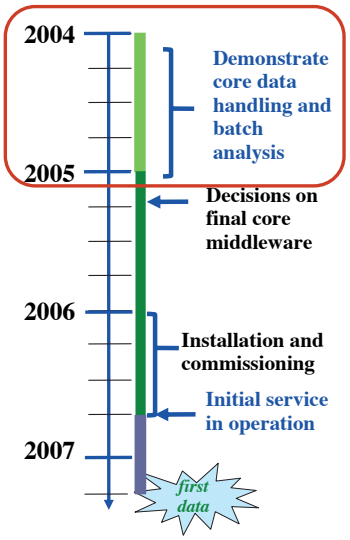
Summary

LCG-2 is running as a production service

Anticipate further improvements in infrastructure

Broadening of participation and increase in available resources

In 2004 we must show that we can handle the data - meeting the Data Challenges is the key goal of 2004



Courtesy: Oliver Keeble
LCG Deployment Status
25 May 2004

Some different sorts of computing for HEP

- Calculations – for experiments – accelerators – theory – engineering
- Data acquisition
- Data storage and processing
- Simulation
- Interactivity
- Graphics ➔ VR
- Community computing
  - Email, program and data "management", (Web access)
  - Remote conferencing and collaborative tools
- Information access (Web)
- Databases
- Networking
- Symbolic computing
- Control systems
- Lattice QCD calculations

DATA HANDLING

The major challenge by far for CERN computing lies in the data

Tapes being sent up from B513 basement

What they found when they got upstairs

3480 cartridge (200 MB)

STK silos in 2000

Data handling software

- Moving between files structures (understood by machines) and structures of events grouped into various stages of analysis
- FATMEN at LEP
- A crucial determinant for Grid software - can we get wide consensus on how to address our data?

INFORMATION HANDLING

More of what I just mentioned

The world is hoping that computing can bring more order - and perhaps some intelligence (we have a bad track record!) to all sorts of information. Semantic Web - Semantic Grid etc.

Text Processing

- Something that we/I tend to forget
- All over the world scientists "did it themselves" and took part in standardisation efforts
- A good review in the 35th Anniversary Computer Newsletter Edition, by Michel Goossens

The Web

- Something that we (especially Tim BL) did right
- We had the problem - info sharing for LEP experiments
- We had the infrastructure - reasonable networks had just started
- We had just enough smart people
- And we managed to put the critical software in the public domain

Robert Cailliau (+Tim BL and Ted Nelson)

Photo: Hakon Lie

PROSPECTS FOR THE FUTURE

(COMPUTING AT CERN)

Complexity

- Computing will never be easy because of complexity
- We probably don't explain the source of the complexity enough
- Take a serious application from 1970 - say 10'000 lines of code (l.o.c.)
- Assume that every 10 l.o.c. we encounter a 2-way logical decision, where the route taken depends on the input data, or on the prior state of calculations. That is a very conservative estimate.
- There are, therefore, 2^{1000} possible routes through our code
- $2^{1000} = (2^{10})^{100} = 10^{300}$ routes! A googol cubed.
- LHC experiments will use a total of some 10^7 lines of code!!

[Bock and Zoll in the 1972 CERN Courier – Central Computers in Bubble Chamber analysis]

"The computer spends most of its time processing un-problematic events. The programmer, on the other hand, spends most of their time foreseeing – or discovering – possible difficulties and programming the computer to deal with them. The computer programs for bubble chamber experiments start with elegant and simple ideas, and end up complex and sophisticated."

The fight against complexity

- Decompose problems – several small and (fairly) simple systems tend to beat complex monoliths
- Try to think clearly
- Be prepared to handle errors (easier said than done)

The natural tension between particle physics and computing

- CERN experiments have to be mission-driven
- CERN computer specialists cannot be *entirely* mission-driven
  - They must pay attention to how computing technology (all different wares) is evolving
  - And how the economics of the industry evolves
  - In order to find optimal solutions "for tomorrow"
- Some elements of CERN computing need to be handled centrally, and others must be handled directly by the experiments
- Finding the correct balance in that respect is an ongoing job for the IT and PH department managements

The impact of HEP computing on the wider world?

- Impact on other sciences
- Impact on society in general
- Impact on developing countries

- The +ve side of computing
  - Info access
  - Great help for organising meetings and travel
  - Enriching
- The –ve side
  - Spam
  - Crooks
  - Mind-numbing

Programmers

- To be a good programmer some part of your mind must sympathise with a machine – which has to do what it is told, so must be told – in all gory detail – exactly what it must do.
- If you can communicate well with a machine can you also communicate well with people?
- And can you communicate without your machine?

The companies

- In 50 years of CERN computing several companies have been major players – and we should probably say major contributors to the work
- CDC, Cray, DEC/Digital, HP, IBM, Norsk Data
  - At various times DEC and IBM had Joint Project teams of a serious size
- Now Open Lab provides a more general possibility for such interactions

- Is that all only on the side of the hardware? What about software suppliers – Oracle, Microsoft?, Numerical Algorithms Group (NAG)?

Our successes and maybe some failures

- Web has to be the biggest success
- But also CERNLIB as a sharing mechanism
- And a lot of widely-used packages
  - REAP, THRESH, GRIND, SLICE
  - HBOOK ➔ PAW
  - GEANT ➔ GEANT4
  - Data acquisition software (PDP, Nord-10(0) and 50(0), HP, VAX)
  - Packages from outside "pure HEP"
- And our external networking (especially CIXP)

- Why did we have less impact (overall) on the development of early Internet standards?
  - CERNET (1973 (concepts) to 1977 (deployment) to 1988 (turn-off)) served our onsite needs superbly but was CERN-specific
  - HEP deployed DECnet very successfully in the wide area (plus wide use of mainframe solutions) and was slow to look at "peer-to-peer"
  - But we were probably slow to catch on to INTER-networking

… some failures

- We have been quite strong in the "industrial" deployment of software development, and software development distributed over the wide-area, but have never been recognised as such by the broad community. Why?
  - PATCHY was extremely advanced
  - As were many of the ideas for machine independence behind "CERN Fortran"
  - And distributed software development
- I think that our problem here has been that
  - (a) the people involved were very busy – driven by experiment deadlines and
  - (b) we did not see it as our job to make this information available to others – either potential users in other sciences or to computer scientists and (commercial) software engineers who should have been kept informed
- Computing tends to be too introverted
- As I said that the start
  - The real problem has always been, in my opinion, getting people to collaborate on a solution

But the success predominate

- By a long way
- And it was:-
  - Challenging
  - Fascinating
  - Lots of hard work
  - And great FUN

WIM KLEIN

Intro for the younger members of the audience:-
In the early days computing was quite difficult. Wim Klein was a "human computer" who provided numerical assistance, especially to generations of CERN theoretical physicists

He retired from CERN in 1976

FINAL REMARKS

The PhotoLab Archives – time is running out

- There is a lot of fascinating photographic material in the Photolab Archives, but only a tiny fraction has been scanned, and the rest is almost entirely without any descriptions. From the early years of CERN (1950s and 60s) many of the people shown on the photos are now >80 (or dead) and even the people who have some idea what the pictures might be about will soon retire.
- We urgently need to do a mass scan of the photo archive and make it Web-available so that past and present CERN staff are able to provide descriptions.
- We have to find the money to do this – it's our contribution to preserving a minimum "historical heritage"

Take-home messages

- Computing for particle physics is challenging – and it is likely to stay that way
- One big change of the next decade will be the move to fully mobile computing
- We can also hope that there will be moves to improve the way in which most information is organised and made accessible
- It will be important for CERN (and particle physics) to invest enough in computing infrastructure (including people, servers, network bandwidth and performance monitoring) since that is vital for each persons professional working efficiency
- And make sure that computing stays FUN for both users and providers

THE END

Or rather, the end of what I dare detain you with