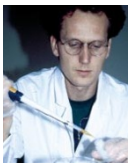
Irish fortune
Research gets a cash boost, but what should it be spent on?
p215



Slow DFG
German funding body 'takes too long to process grants'
p217

EMMA and EBI
EU research ministers say facilities should be saved
p217

Art at CERN
Particle physics has inspired a group of European artists
p219

Europe joins race to turn the Internet into one vast computer

Paris

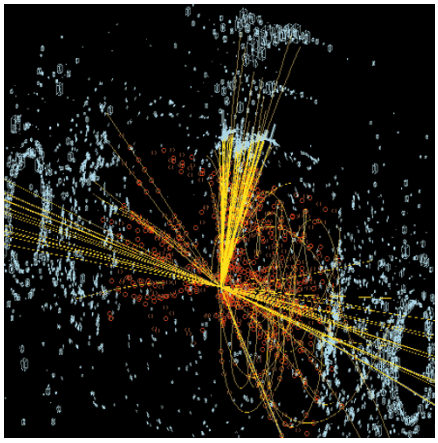
CERN, the birthplace of the World Wide Web, aims to make Europe a leading player in the largest distributed computing project in history. The Large Hadron Collider (LHC) at CERN, the European Laboratory for Particle Physics, will produce a deluge of data. Now CERN is coordinating a proposal that would use the challenge of analysing all these data to help develop a concept known as the grid.

The grid idea was created in the United States (see *Nature* **402,** suppl., C67–C70; 1999). Its goal is to develop software and Internet protocols to transform the Internet into a single gigantic computer. Researchers throughout the world could work on shared data sets on a network running thousands of times faster than today's best. Eventually, the technology could be applied to the public Internet.

The United States already has a $500 million five-year grid effort involving 50 research centres and coordinated by the National Computational Science Alliance. CERN now plans to submit a proposal to the European Union (EU) for a grid infrastructure costing 30 million euros (US$29 million). It has set itself a deadline of 10 May. The proposal is likely to be favourably received, as European Commission officials have actively solicited it.

David Williams, formerly CERN's head of computing and networks and now responsible for relations with the EU, points out that the web has made the physical location of information irrelevant, yet scientists still mostly use the Internet just to search for information. Real-time computing and data handling have hardly been explored, he says. John Taylor, director-general of the UK research councils and a convert to grids, has coined the term 'e-science' for this new way of working.

The CERN-led proposal will initially focus only on particle physics. But EU funding may be extended next year to other disciplines such as biology. Some European Commission officials see the challenge presented by the LHC's deluge of data as an opportunity for an all-out bid to catch up with the United States in constructing what many believe will be the successor to the web.



Ring the changes: simulated data from the LHC.

Andrea Dahmen, spokeswoman for research commissioner Philippe Busquin, says he "wants to see a reliable high-performance EU Internet network in place as soon as possible".

The $1.8 billion LHC, which will come online in 2005, is an ideal test bed for the grid concept. Collections of protons steered into head-on collisions will spew out around 7 petabytes (10^{15} bytes) of data every year. The raw data from just one of the LHC's detectors represent the equivalent of every person on Earth talking into 20 telephones at once. The challenge is to transmit this information — requiring 1,000 times more computing power than CERN can currently deliver — in a usable form to thousands of users in more than 40 countries.

Under the CERN plan, data would fan out across a high-speed network to 10 national and regional data centres, and in turn out to hundreds of local centres and universities. The innovation will be the glue holding all of this together: a new suite of middleware, software and protocols designed to allow real-time distributed computing. The Internet would be made to function as if it were a single computer and database rolled into one.

"Getting ourselves liberated from the geographical constraints will be crucially important" for the success of the LHC, says Williams. "We need to use any and all possible resources to process the data. Not all of the people nor all of their computing and data-handling resources can be installed permanently at CERN."

Imagine an LHC scientist sitting at his or her desk in California, say, or Budapest. One click on a web wizard, and time is automatically reserved and purchased in real time from supercomputers and clusters of personal computers around the world. Another click, and data sets worldwide are scoured for all the Higgs two-photon events recorded so far. The interface invisibly converts all the datasets into a compatible format. One keystroke, and a menu pops up offering a suite of advanced visualization techniques that will allow the data to be analysed interactively with colleagues elsewhere.

Turning this dream into reality was the aim of a meeting at CERN last week between officials from the major European research councils, the European Space Agency and the European Bioinformatics Institute. The proposal that emerged would build on middleware developed from the existing Globus distributed software effort, a joint project of the Argonne National Laboratory in Illinois and the University of Southern California's Information Sciences Institute. The Linux operating system would be used throughout, says Williams, to ensure that software remains open-source.

By coordinating the national efforts that are being planned — Britain is likely to approve £100 million (US$158 million) in funding for grids later this year — the proposal could quickly be scaled up to include other disciplines and industry.

"It is clear that there are people in other sciences very interested in doing the grids together," says Chris Jones, CERN's head of technology transfer.

Support for the development of grids is also awakening within industry. Jones last week discussed the grid proposal with a delegation of British industrialists, and Williams is optimistic that concrete industrial support for the proposal will be in place before the 10 May deadline.

Declan Butler

LUCAS TAYLOR/NORTHEASTERN UNIVERSITY

© 2000 Macmillan Magazines Ltd