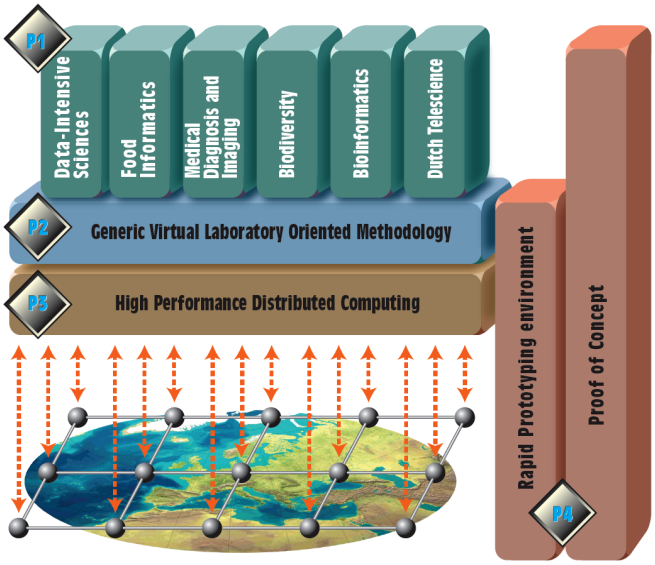
vl·e virtual laboratory for e·science

information

VL-e enables new approaches to traditional sciences

Information has become the fuel of our knowledge society, and our ability to digest, understand and share it will determine our scientific, economic and social progress.

The exceptional increase in computing power, storage capacity and network bandwidth over the past decades forms the basis of a digital revolution which has only just started. Also the changing scale and scope of experimental sciences require a new research paradigm: (digitally) enhanced science or e-Science. The aim of the 'Virtual Laboratory for e-Science' (VL-e) project is to bridge the gap between the technology push of the high performance networking plus the Grid and the application pull of a wide range of scientific experimental application domains. A typical example of this is the life sciences, where VL-e offers solutions for combining laboratory research with computational experiments and simulations, making use of the knowledge and experience gained from dealing with large data sets in high energy physics. At the same time, however, it is recognised that data sets in the life sciences are far more complex than in high energy physics.



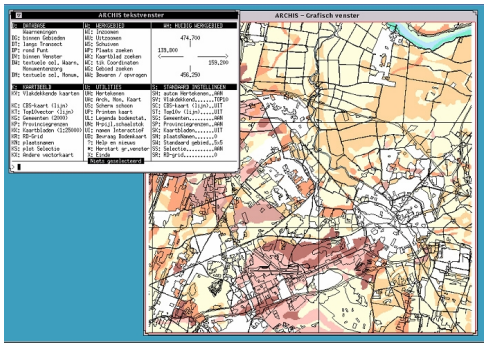
More specifically, VL-e is developing a Proof-of-Concept (PoC) infrastructure (both hard- and software) to enhance location-independent access to scientific information and stimulate global and multidisciplinary collaboration, thereby enabling new approaches to traditional sciences. The VL-e software (both for rapid prototyping and in the PoC) provides generic functionalities that support a wide range of e-Science applications. This PoC infrastructure will boost the knowledge economy of the Netherlands.

Currently, six application domains are involved: Data-Intensive Sciences, Food Informatics, Medical Diagnosis & Imaging, Biodiversity, Bioinformatics and Telescience. Several Dutch universities, academic hospitals and industries in the life sciences and ICT domain participate in this project. There is also strong collaboration with NBIC (Netherlands Bioinformatics Centre). The recently started Dutch BIG GRID project will build a nationwide production grid, making use of methodology still being developed within VL-e.

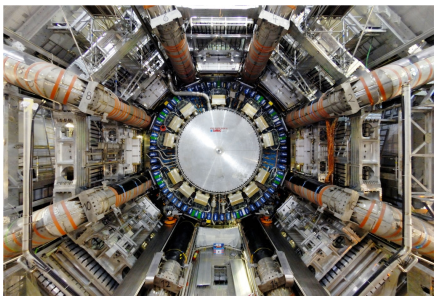
Data-Intensive Sciences:
From Desktops to Distributed Terascale Archives

For some VL-e applications, the main problem is having achieved a scale for which operating their own infrastructure is no longer feasible. The Data-Intensive Sciences program develops and disseminates techniques for connecting such research environments to shared grid infrastructures.

The Data-Intensive Sciences program is concerned with the general problem of dealing with increased data volumes by scientific applications. The applications have the ability to generate the data, but techniques for dealing with this data --- at both the infrastructure and analysis level --- are most often lacking. The analysis level is often quite application-specific, but at the infrastructure level most of these applications are quite similar.



DANS is working to use grid resources to store massive amounts of archeological data, the collection and archival of which is mandated by the Treaty of Malta.



Data from Particle Physics is produced in the ATLAS detector in Geneva (above), petabytes per of which are stored in the grid-enabled tape robot at SARA in Amsterdam (immediate right) and similar centers around the world.

The LOFAR radio telescope (far right) will generate similar volumes of astrophysical data.

The Data-Intensive Sciences program develops techniques for connecting applications to the infrastructure, in such a way that the infrastructure can remain application-neutral, and the application retains maximal flexibility of design from the user point of view. Bluntly speaking, the main question is “how can we boost user communities onto the dutch e-science infrastructure without a) deploying a dedicated infrastructure and b) forcing them to do a massive rewrite of their computing framework in order to make use of the generic infrastructure?”

vl-e facts

budget 40 M, period 2004-2008
more than 20 consortium partners from industry and academia
director: prof. dr. L.O. Hertzberger
website: http://www.vl-e.nl

consortiumpartners

A&F Wageningen, AMC, CWI, DSM, Friesland Foods, FEI, FOM AMOLF, NBIC, Nikhef, IBM, LogicaCMG, Philips Research, Philips Medical, SARA, Top Institute Food and Nutrition, TNO Kwaliteit van Leven, TU Delft, Unilever, UvA-IBED, UvA-IvI, UvA-SILS, VU, VUmc, WTCW

VL-e Program line: e-Science in Applications

Subprogram: SP1.1 Data-Intensive Sciences

Leader: Dr. Jeff Templon
templon@nikhef.nl

Organization: National institute for subatomic physics (Nikhef)
Science Park Amsterdam