

A Brief History of the Semantic Grid

David De Roure

www.semanticgrid.org

Visions

The story begins with the formulation of the UK e-Science programme during 2000. Instigated by John Taylor, then the Director General of UK Research Councils, e-Science was defined as being ‘about global collaboration in key areas of science and the next generation of infrastructure that will enable it’ and undertook to change the dynamic of the way science is undertaken. This vision of accessing data and computation anywhere in the world can be traced back to J.C.R. Licklider in the 60s, and to Wulf’s 1989 vision of the collaboratory as “...a center without walls, in which researchers can perform their research without regard to geographical location, interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries.”

The key technology proposed for the e-Science infrastructure was Grid computing, which had emerged from the efforts in the 1990s to combine computing resources in pursuit of new levels of computational power and very large scale data processing, using high speed wide area networking of supercomputers and large clusters of commodity PCs. This new power enabled researchers to address exciting problems that would previously have taken lifetimes, and it encouraged collaborative scientific endeavours. The term ‘Grid’ was chosen to draw an analogy with the way in which the electricity power grid brought about a revolutionary change from the use of local electricity generators – in this view, computational resources, data and expensive scientific instruments can be regarded as utilities to be delivered over the network.

Around the outset of the e-Science programme, Ian Foster and Carl Kesselman were describing the Grid as being “distinguished from conventional distributed computing by its focus on large-scale resource sharing, innovative applications, and, in some cases, high-performance orientation” [1]. The contemporary definition of the Grid focused on the notion of Virtual Organisations (VOs); i.e. “flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources”. It was clear that the Grid was to play a part in providing a common IT infrastructure to support the multidisciplinary and distributed collaborations of e-Science.

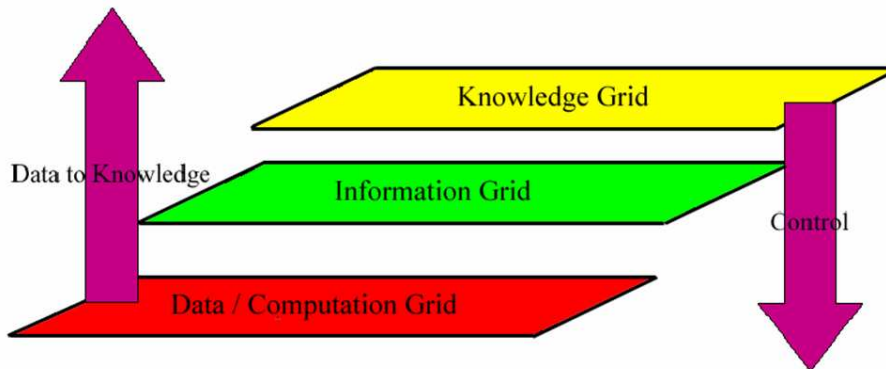
Meanwhile Tim Berners-Lee was leading the Semantic Web activity, presented at WWW7 in Brisbane in 1997 but with roots which go back to the early days of the Web. 2001 saw the publication of the Semantic Web paper in *Scientific American*, co-authored with Jim Hendler and Ora Lassila [2]. Again about joining things up, the Semantic Web presented a vision of an automatically processable, machine understandable Web, contrasting the Grid’s emphasis on on-demand transparently constructed multi-organisational federations of distributed services. The complementarity was evident: the Semantic Web provided information integration; the Grid middleware provided computational integration.

The conception of the Semantic Grid

Early in 2001 a number of researchers working at the intersection of the Semantic Web, Grid and software agent research and development communities were increasingly conscious of the gap between the aspirations of e-Science and the then-current practice in Grid computing. This was captured in the report by David De Roure, Nick Jennings and Nigel Shadbolt, with help from Mark Baker, presenting a research agenda for the e-Science infrastructure. This drew on Grid computing, the Web, Web Services, software agents and knowledge technologies. The draft report was circulated within the e-Science programme in June 2001.

In August 2001, teams from the UK and US came together at an “N+N” meeting (the *first UK-US Workshop on Grid Computing*) in San Francisco. The meeting included presentations by David De Roure on ‘Agents and the Grid: A Service Oriented View’, Carole Goble on ‘Information Grids, the Semantic Web & Why Ontologies Matter’ and Keith Jeffery on ‘The Information and Knowledge

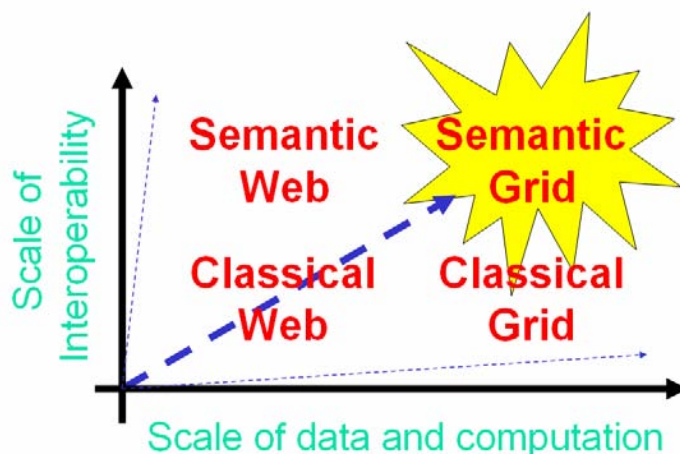
Layers of the GRIDs Architecture'. Jeffrey's 'Grids' picture, shown below, had been significant in formulating the e-Science programme.



The report was published in December 2001 as "The Semantic Grid: A future e-Science Infrastructure", the second report from the e-Science programme [3]. At the same time the Semantic Grid web site, www.semanticgrid.org, was created as the first step towards building the Semantic Grid community.

Building bridges

There followed a series of activities to bring together the research communities needed to realise this infrastructure. A panel session was held at the international World Wide Web Conference in May 2002, put together by David De Roure and featuring Ian Foster, Jim Hendler, Eric Miller and Carole Goble. The panel used the picture on the right, based on an idea by Norman Paton, as a framework for discussion. Ian Foster gave a keynote at the Web conference and Carl Kesselman



was to give a keynote at the first International Semantic Web conference in Sardinia in June 2002. The Global Grid Forum in July 2002 in Edinburgh (GGF5) included a keynote by Henry Thompson on the Semantic Web and a tutorial on *Ontologies and the Grid* by Carole Goble and Nigel Shadbolt.

During 2002 the e-Science programme took on board some of these ideas, with Semantic Grid aspects in three of the e-Science pilot projects and presentations in the *UK e-Science All-Hands Meeting 2002*. A *Semantic Grid and Autonomic Computing* programme was funded in conjunction with JISC (the UK Joint Information Systems Committee). The IST Programme in Europe also embraced the combining of Grid and Semantic Web/Knowledge technologies, with the December conference on *Convergence of Web Services, Grid Services and the Semantic Web for delivering e-Services* at which Carl Kesselman spoke. Semantic Grid also featured at the EuroWeb conference *The Web and the GRID: from e-science to e-business*, a regional conference under the auspices of the International World Wide Web Conference Committee (IW3C2), in Oxford UK that same month.

GGF5 also saw the first Semantic Grid 'Birds of Feather' (BoF) session. After a second successful BoF at GGF6 in October 2002 in Chicago, the Semantic Grid Research Group was created in November 2002. The group was chaired by David De Roure, Carole Goble and Geoffrey Fox, with Bill Johnston as the area director and Bruce Barkstrom as secretary. Though many felt that the group belonged in the "data area" of GGF, it was placed in architecture to be alongside the Open Grid Services Architecture activity, reflecting the desire to explore the application of these ideas in the Grid middleware. The

charter focused on tracking Semantic Web community activities, determining their relevance to grid activities and providing a route for transfer of information and ideas between the communities. It included examples of questions that might be answered by the Semantic Grid – perhaps the best way of articulating the Semantic Grid concept to its users.

In and on the Grid

The Semantic Grid enables scientists to answer questions which involve integration of scientific data and automatic execution of computations, providing important functionality at the datagrid and scientific applications level (an aspect sometimes known as ‘knowledge grid’). Significantly, it also facilitates automation within the grid middleware – helping to discover and compose a variety of Grid resources and services in order to meet the requirements of dynamic Virtual Organisations. Hence the Semantic Grid emphasises the use of Semantic Web technologies *in* the Grid. Although it was initially compelling to talk of the Semantic Grid in terms of Jeffery’s highly effective three layer ‘Grids’ model, subsequently it has been articulated to emphasise the position in the middleware, as depicted in the following figure.



Examples within the Grid middleware include ‘semantic’ (ontology-based) descriptions of services, services that may use ‘semantic’ technologies but present like any other service (such as an RDF triplestore) and services to support ‘semantic’ capabilities, such as an ontology services to support digital library functions.

The attention to semantics in both the Grid middleware and applications was captured in the chapter for ‘Grid Bible 2’ entitled “Enhancing services and applications with knowledge and semantics” and written by Carole Goble, David De Roure, Nigel Shadbolt and Alvaro Fernandes [4]. The Semantic Grid logo (which was designed ahead of a talk in Krakow) captures this emphasis graphically.



Building a community

There followed a series of community events and outreach into both the Grid and Semantic Web communities. The first Semantic Grid workshop was held at GGF9 in Chicago in October 2003, with invited papers. The Jan-Feb 2004 issue of IEEE Intelligent Systems was a special issue on e-Science, edited by David De Roure, Yolanda Gil and Jim Hendler, with a significant Semantic Grid focus [5]. Semantic Web gained increased attention in February when the W3C Web Ontology Language OWL became a W3C recommendation. At GGF11 in Hawaii in June there was a *Semantic Grid Applications Workshop*, with peer-reviewed papers covering a wide breadth of activities from job description to collaboration, and GGF12 in September featured the first Semantic Grid tutorial. Other community bridging events in 2004 include Carole Goble’s keynote at ECAI in August [6], which focused on the need to bring communities together, and workshops at ECAI, at WWW2004 in New York and in China. In October the *W3C Workshop on Semantic Web for Life Sciences* included Grid projects in bioinformatics.

Meanwhile in the IST (Information Society Technologies) programme of the European Commission, the *Next Generation Grid(s): European Grid Research 2005 – 2010* Expert Group Report published in June 2003 had incorporated a Semantic Grid agenda. A significant number of projects in the current round of funding from the Grid Technologies Unit (the so-called ‘Grid House’ projects) involve some aspect of Semantic Grid, and concertation activities include a Semantic Grid activity led by Ziga Turk. The final report of the Expert Group convened to complement, enhance and extend this earlier report appeared in September 2004. Semantic Web Services were a subject of increasing attention in this period and a debate on *The Semantic Grid: when the Semantic Web meets the Grid* was held at the IST 2004 conference in the Hague in November.

Best practice in Semantic Grid was emerging as the various Semantic Grid projects reached maturity across the globe. The original Semantic Grid report was revisited later in 2004, reflecting on what had been achieved in the e-Science programme and looking again at the research agenda – the updated report was published in Proceedings of the IEEE in March 2005 [7]. In particular it discussed the role of agency, which had been part of the original report but in practice the Grid had adopted Web Services and had also been revisited in the ‘Brain meets Brawn’ paper by Ian Foster, Nick Jennings and Carl Kesselman [8]. The other challenges include the self-management, self-optimisation and self-healing (so-called ‘autonomic’ behaviour) necessary for large scale distributed computing, and the intersection between the Semantic Grid and physical world through pervasive computing devices.

2005 has seen interactions with the China Knowledge Grid, led by Hai Zhuge, and planning for the first Semantic Grid conference. There have been keynotes by Carole Goble at the European Grid Conference and David De Roure at the European Semantic Web Conference. Line Pouchard ran the *Semantic Infrastructure for Grid Computing Applications* workshop at CCGrid 2005. The Semantic Grid Research group is GGF has become a ‘Technology Innovators’ group and is collaborating with the Agents community.

The Dagstuhl Seminar in July, organised by Carole Goble, Carl Kesselman and York Sure, is the key Semantic Grid event of 2005. It has brought together researchers in the Grid and Semantic Web communities and, significantly, enlarged the intersection to include leading figures in those fields.

Moving forwards

As well as the essential bootstrapping task of building interactions across communities, the community is pushing forward with the Semantic Grid and its research agenda. Like the Grid and the Semantic Web, the Semantic Grid is fundamentally about joining things up in order to achieve what was not possible before. Now with four years of experience, the various Semantic Grid activities are all working towards creating an open, global, shared space in which the Semantic Grid infrastructure can be realised and we can move towards the vision that inspired this activity – a vision of a high degree of easy-to-use and seamless automation in which there are flexible collaborations and computations on a global scale.

References

- [1] I. Foster, C. Kesselman, and S. Tuecke. “The Anatomy of the Grid: Enabling Scalable Virtual Organizations,” *International Journal of Supercomputer Applications*, vol. 15, 2001.
- [2] T. Berners-Lee, J. Hendler, and O. Lassila, “The Semantic Web,” *Scientific American*, 284(5), 2001, pages 34-43.
- [3] D. De Roure, N. R. Jennings, and N. R. Shadbolt. “Research agenda for the Semantic Grid: A future e-science infrastructure.” Technical Report UKeS-2002-02, National e-Science Centre, December 2001.
- [4] C.A. Goble, D. De Roure, N.R. Shadbolt, and A.A.A. Fernandes. “Enhancing services and applications with knowledge and semantics.” In I. Foster and C Kesselman, editors, *The Grid 2: Blueprint for a New Computing Infrastructure*, pages 431–458. Morgan-Kaufmann, 2004.
- [5] D. De Roure, Y. Gil, and J.A. Hendler (eds). *IEEE Intelligent Systems, Special Issue on E-Science, Volume 19, Issue 1* (January/February), 2004.
- [6] C.A. Goble and D. De Roure. “The Semantic Grid: Myth busting and bridge building.” In 16th European Conference on Artificial Intelligence (ECAI-2004), pages 1129–1135, Valencia, Spain, 2004.
- [7] D. De Roure, N.R. Jennings, N.R. Shadbolt. “The Semantic Grid: Past, Present, and Future,” *Proceedings of the IEEE*, Volume 93, Issue 3, March 2005, pages 669-681.
- [8] I. Foster, N. R. Jennings, and C. Kesselman. “Brain meets brawn: Why Grid and agents need each other,” *Proc. 3rd Int. Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS 2004)*, New York, USA, 2004.