Studying the Usability of Grids: Ethnographic Research of the UK Particle Physics Community

Will Venters[1*], Tony Cornford[1], Mark Lancaster[2], Yingqin Zheng[1], Avgousta Kyrikidou[1]
[1] London School of Economics, London, UK.

[2] UCL, London, UK

Abstract

The EPSRC's *"Usability Challenges to emerge from e-Science"* highlighted the need to consider practices to make Grid technology more usable. As part of this agenda the Pegasus project is studying the UK particle physics community's development and use of Grid technologies as it prepares for the launch of a major new experiment at CERN. We outline the aims of the project, describes its methodology, and provides an overview of its expected contribution. We present tentative initial contributions from the project and outline how we aim to contribute to e-science as the project progresses.

Introduction

This poster describes the current preliminary research findings of the Pegasus[1] project (Venters & Cornford, 2006) which is studying the development of Grid technology within the UK particle physics community (GridPP). The project is exploring the possible insights this community's experiences can offer others as they engage with Grid technologies. We are concerned with the concept of "usability" of Grid technology, believing that the unique situation faced by Particle Physics as it prepares for the "data deluge" of the launch of its new experiments (the LHC) offers important lessens for others starting to "use" Grids within their research. We describe the research context, the theoretical lenses we employ to research the contested, political nature of the GridPP project alongside its technical realization, and present some tentative findings from our ongoing analysis.

GridPP is the UK particle physics collaboration to develop a Grid to analyse the unprecedented amounts of data (some 15 Petabytes per year) that will be produced by the LHC (large hadron collider) experiments currently under construction at CERN[2]. Constructing such a Grid presents a considerable technical challenge, however our interests lies in the processes and practices of bringing such Grid technology, taken as a socio-technical system rather than purely technical artefact, into widespread use. We are in particular concerned with the ways that Grid technology is accommodated into the working practices of physicists; how it is shaped by these 'users' (albeit highly technical users); and how this in turn shapes the LHC experiments and more broadly the character of experimental particle physics. This is of particular interest because the community already demonstrates their ability to work pragmatically as a globally distributed community that thrives on debate and discussion (Knorr-Cetina, 1999) (something Grids are expected to engender in other sciences (Foster & Kesselman, 2004)).

As our study's theoretical point of departure we see infrastructures not as static technical platforms driven by engineering rationality and fixed common standards, but rather as political accommodations and arenas within which issues of work practices, resources, and organizational mission are surfaced and, to a degree, resolved (Ciborra, 2002; Star & Ruhleder, 1996; Hanseth, 2000; Ciborra, 2000). This leads us to focus on the improvised, and bricolaged (Ciborra, 2002) actions which are employed by the developers and users of GridPP alongside consideration of their (more or less) systematic design and management practices.

We frame the project as a balance of research that addresses *action of doing particle physics* alongside the *action for doing particle physics*. The perspective we take is thus that GridPP is itself constitutive of science rather than simply its enabler, a tool or an unproblematic ideal 'platform' (Latour, 1987).

[1] http://www.pegasusresearch.org.uk

[2] http://www.cern.ch

Methodological Approach

Influenced by both distributed cognition (Hollan et al., 2000) and our socio-technical perspective towards infrastructure, the research project is combining ethnography of the work practices of experimental particle physicists, and of those involved with the design and implementation of GridPP.

A qualitative longitudinal research methods is chosen as appropriate because of the focus on discourse and communication within particle physics (Knorr-Cetina, 1999). Researchers on the project are spending two years charting the development of GridPP, attending meetings, undertaking interviews (41 so far), and through extended periods spent with Particle Physicists across the UK and at CERN, and through attendance of the regular weekly video-conference meetings of GridPP.

Early Contributions

Drawing upon the interpretive research tradition in information systems (Walsham, 1993), the focus in the analysis will be on sensemaking, learning and organising (Weick, 1995). Since our study is exploratory as is grid research, it is difficult to anticipate the definite contributions we might ultimately make. However the project has set itself the objective of developing a framework of qualifying guidance for the development of e-science infrastructure. This framework aims to address the development, deployment, and use of grids by collaborations of disparate scientists working as a virtual organization. Our aim is to provide, through ethnographic accounts of GridPP, a coherent set of understanding of how the requirements of LHC experiments were translated into grid development; how the potential of grid technology is understood and appropriated by particle physicists; and how the objectives of e-science are achieved within this unique context.

Our initial research is pointing to two directions. Firstly to consider the networks of relationships between those 'developing' the infrastructure (within a range of projects), and the ultimate 'users' of the technology. We are exploring the "user-developer" relationship through which "usability" is co-constructed. For GridPP these groups are all highly competent in computing and are constructing software within their work. 'Users' are seen producing software to hide the complexities of the generic Grid middleware to ensure easier analysis of their data. We are currently exploring whether such experiment-specific software challenges the anticipated usage patterns of the Grid. For example, allowing users to target specific machines on the Grid to undertake their processing changes usability (from the users perspective) of the grid, as 'users' engagement with Grids is mediated through other software. Unpacking the experience of advanced users like the particle physicists sheds insights to the process of adapting the 'generic' middleware to 'specific' user needs, which poses great challenges for the uptake of e-science.

The second direction of our early research concerns the balance between top-down coordination and bottom-up innovations in the collaborative development and deployment of Grids. We explore how the working practices of particle physics - characterized as flexible, improvised and highly innovative - is exploited in the development of the particle physics grid. With limited formalised systems development methodologies, GridPP constitutes an interesting example of how a large scale, distributed computing system emerge with an evolutionary approach naturally formulated to deliver a functioning system under conditions of high complexity and uncertainty.

In addition, we anticipate these and other lessons from the project will have wider implications for the information systems field, particularly around the development of complex infrastructures such as those implied by "software as services", open-source and virtual organizations.

References

Ciborra, C. (2000) *From control to drift: The dynamics of corporate information infrastructure.* Oxford University Press, Oxford.

Ciborra, C. (2002) *The labyrinths of information: Challenging the wisdom of systems.* Oxford University Press, Oxford.

Foster, I. and Kesselman, C. (2004) *The grid 2: Blueprint for a new computing infrastructure.* Elsevier.

Hanseth, O. (2000) The economics of standards. In *From control to drift* (Ciborra, C., Ed), pp 56-70, Oxford University Press, Oxford.

Hollan, J., Hutchins, E. and Kirsh, D. (2000) Distributed cognition: Towards a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction* 7 (2),

Knorr-Cetina, K. (1999) *Epistemic cultures: How the sciences make knowledge.* Harvard University Press, Cambridge, MA.

Latour, B. (1987) *Science in action.* Harvard University Press, Cambridge MA.

Star, S. L. and Ruhleder, K. (1996) Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research* 7 (1), 111-135.

Venters, W. and Cornford, T. (2006) Introducing pegasus: An ethnographic research project studying the use of grid technologies by the uk particle physics community. In *Second International Conference on e-Social Science*, Manchester, UK.

Acknowledgments

The Pegasus team would like to acknowledge the considerable support and help provided by GridPP and its members in undertaking this research. Pegasus is funded by the UK EPSRC (Grant Number EP/D049954/1) under the e-science program. The project includes participants from University College London and LSE. The authors acknowledge the considerable contribution of Jana Hlistova's in undertaking the initial analysis.