

MAKING GRIDS USABLE: DEPLOYMENT AS THE MANGLING OF INFORMATION INFRASTRUCTURE

Will Venters, Tony Cornford, Yingqin Zheng.

London School of Economics, Houghton Street, London, WC2A 2AE.

Abstract

This paper explores the relationship between users and developers, and in particular the role of empowered users and their virtual organizations in negotiating usability of e-infrastructure. The paper presents a case study of Crab, a Grid job submission tool developed and used in a High Energy Physics community. We then use Andrew Pickering's "Mangle of Practice" to address the question of how making Grids usable becomes a practice of experimental physicists as well as, and in conflict with, a practice of systems designers. The conclusions highlight the need to consider the users of Grids as powerful agents in the development process, able to direct and steer the development to represent their practices.

Keywords: Mangle of Practice, Grid, Particle Physics, Implementation, Deployment, Agency.

Pegasus Working Paper

Information Systems and Innovation Group

Department of Management, London School of Economics

Houghton Street, WC2A 2AE.

www.pegasus.lse.ac.uk

© the authors 2010

DRAFT - DO NOT QUOTE WITHOUT PERMISSION OF AUTHORS

Pegasus Working Paper - DRAFT

Introduction

This paper explores the relationship between users and systems administrators in negotiating usability during the deployment of a particular form of information infrastructure – a computing Grid. Infrastructures in their ideal formulation are transparent and provide an almost invisible support for practice within the communities they serve. We become aware of them only when they fail (the lights go out, the drain is backing up, the email is dead, the bridge has collapsed); in this way information infrastructures are the foundation of our contemporary world in areas such as government, business (Ciborra 2000), transport, healthcare and science (Edwards, Jackson et al. 2007). Yet the process of the creation and adoption of infrastructures is often far from simple, with even the telephone initially finding it difficult to establish a user-base and to be understood as useful (Brown and Duguid 2000, p87).

The genesis of this paper is our hypothesis, based on our wider field work on Grids (Zheng, Venters et al. 2007), that the key to understanding infrastructure deployment is found in the processes through which technical elements and established social practices engage and each are re-shaped such that an information infrastructure (as a socio-technical phenomenon) may form (or we might say, be *performed*). The particular focus of this paper is on one aspect of this engagement, that of usability (others might be reliability or security etc). Given that e-science aims to be disruptive in nature (and by design) – challenging and changing existing science practice (Hey and Trefethen 2002) - so we focus on usability in terms of both stability and instability of practice. Our research question is thus: How, and where, is the Grid made usable, and by whom?

Information infrastructures are usually portrayed as shared, often taken for granted, resources that supports knowledge sharing and communications within a community. Much recent literature presents infrastructure as fundamentally a social-technical network combining technical elements with people and their social practices (Hughes 1983; Dahlbom 2000; Hanseth 2002), and as path dependent and built on legacy systems as well as ‘legacy’ organizational and cultural resources (Star and Ruhleder 1996; Hanseth

and Monteiro 1998; Hanseth 2002). As socio-technical arrangements with an evolving and open character, infrastructures need to achieve some balance or accommodation between adherence to standards (both technical and social) so as to allow interconnection and transferable skills, and support for a necessary heterogeneity of elements to allow scale, scope and fit (Hughes 1983; Dahlbom 2000; Hanseth 2002; Hanseth and Lytinen 2004). Traditional accounts of the development of infrastructure focus on the need for alignment with organisational objectives, concentrating on the need for powerful organisational actors to ensure interoperability, usage and access to the infrastructure (Weill and Broadbent 1998). Historically we see infrastructures as developed and deployed by centres of power exercising strong architectural control; infrastructure is either mandated upon users (e.g. IT infrastructure within commercial companies) (Weill and Broadbent 1998) or available for adoption essentially without change by a huge user community (e.g. mobile phone infrastructure). Of course, such infrastructures will be socially shaped in use (Orlikowski 1993; Orlikowski 1996; Ciborra 1999), but such shaping is usually marginal to the central architectures or capabilities of the technology, and remains premised on a conception of the ‘user’ as a less powerful or insightful agent of technical innovation than the ‘developer’ (Markus 1987) (though acknowledging the challenge to this argument by literature on user-led innovation (Von Hippel 2005)).

Bureaucratic power is also suggested by some as necessary to attend to the perceived “political” difficulties of information infrastructure and to ensure the success of an infrastructure “design” (Weill and Broadbent 1998). Recent literature on infrastructures development is critical of such managerialist perspectives (Ciborra 2000). These critics suggest information infrastructures surface through an interweaved processes of design, cultivation, domestication and use and downplay the role of planning. Infrastructures are not primarily strategically designed and implemented, but emerge embedded within and out of other institutional structures, social arrangements and sediment technological strata (Star and Ruhleder 1996). In this spirit Hanseth (2002) suggests that information infrastructures should be considered a living organism, a sort of socio-technical Gaia hypothesis (Lovelock 1987), and are

cultivated in their environment, rather than built from some dead materials that have to be crafted in design.

This type of socio-technical assemblage does not often result from the execution of a well-thought-through and detailed “grand design” or master plan. Instead, it usually emerges out of the “mess” of what is to hand; artifacts, political interests, social conditions and norms, designs, existing practices, and innovative visions. Of course, technical designs do play a role; the internet is created in part out of the incredible (inconceivable to its designers we suspect) utility of TCP/IP and the scalability and power that it provides. On this basis Infrastructures cannot be totally controlled, but rather they drift (Ciborra and Hanseth 1998; Ciborra, Braa et al. 2000; Hanseth 2002). Infrastructure is not only lived and conceived (de Certeau 1984), but also performed (Feldman 2000; Suchman 2000; Chae and Poole 2005; Knorr-Cetina 2005).

Grid as Information Infrastructure

Here we study a particular example of an information infrastructure, a computing Grid¹. A Grid has been defined as a computing platform for “coordinated resource sharing and problem solving on a global scale,” (Foster, Kesselman et al. 2001). Likened to the electricity power grid, it allows the transparent, seamless and dynamic delivery of computing resources (Chetty and Buyya 2002; Smarr 2004).

Central to Grid infrastructure is the notion of collaboration (David 2004) with both the construction and exploitation of Grids is seen as essentially collaborative (Berman, Fox et al. 2003) with Virtual Organisations central as both a technical and social concept of collaborative Grid development and use (Foster, Kesselman et al. 2001).

¹ We employ the term Grid in a number of ways. The Grid refers to the abstract idea of coordinated resource sharing (complementary to the use of the term The Internet) whereas “grid” refers to a specific socio-technical infrastructure under discussion.

The purpose of Grids is at first sight to offer high power computing resources to relevant communities on demand. Beyond this, and we want to argue equally significant, they are a means to allow a re-scaling of activity in the domains of their users, to therefore draw on more people, institutions, countries, knowledge, data and computers, and in this way new forms of practice at new scales are enabled (Hey and Trefethen 2002).

Until recently Grid installations have been experimental proofs of concept developed and used by committed individuals (Abbas 2004). Such prototypes are however moving into normal use within scientific and technical domains whereupon users' relationship with them will necessarily change. Such moves involve a period of transition in which a prototype, experimental infrastructure that, though in use, is still seen by users as prototypical and part of a process of technical design becomes an infrastructure which is assumed dependable, complete, and maintained (rather than being 'designed'). This transition is also associated with a period of widespread adoption as the infrastructure is 'rolled out' to larger numbers of users. In discussing this crucial, though somewhat ill-defined point we use the term commonly employed by the community we study – *Deployment* (Foster and Kessleman 2004). For Foster and Kessleman deployment is the process by which the experimental becomes the production (i.e. relied upon in the production of work). Relatively little is known about infrastructure deployment – which we define as the process by which a technological assembly starts to coalesce and inter-penetrate in its environment and become an aspect of (ostensibly reformed or changed) common practice within some community.

Deployment may be the word, but as Woolgar (2006) highlights little is known about how Grid technologies are deployed or come to be used. Star and Ruhleder highlight the temporally emergent nature of information infrastructure (Star and Ruhleder 1996), while Venters et al. (Venters and Cornford 2006) and Berman (Berman, Fox et al. 2003) emphasise the collaborative processes of Grid construction, implementation and use.

In this case study we explore the deployment of an information infrastructure (a Grid) within an experimental particle physics community. We focus on the relationship between systems administrators

and users as the usability of the grid is negotiated. Here systems administrators are technical experts who commission the computer resources and develop and install the Grid software. We discuss how physicist users, as members of a team that will undertake experiments using the Large Hadron Collider (LHC) at CERN, impose themselves on the requirements for Grid architecture and middleware. We focus in particular on the mechanisms by which physics users on the experiment called CMS interact with the Grid infrastructure to submit work for analysis. Note that here we are concerned with the ‘central architecture or capabilities’ of the Grid being useable, not simply the customization of a user-interface, or new configuration of existing functionality.

The relationship between the physicist users and the sysadmins involves many tensions as the two groups pursue different images of what the grid should be, tensions that are to a large degree mediated through the evolving technical infrastructure and reflective of their individual work practices. The case study highlights how as problems (resistances) occur both the infrastructure deployment staff (technical developers, referred to here as sysadmins – systems administrators) and the particle physicist users attempt to accommodate this new Grid within their practices, so in turn shaping the technical infrastructure in various ways and constructing meanings of the concepts of usability.

In this analysis we draw on literature from the sociology of technology and Science and Technology Studies (STS) and adopt an approach that is based on the belief that “the development process of technological artefacts is multi-directional rather than linear, and constituted by processes of negotiation between social groups who vie to achieve specific outcomes during phases of controversy”(Orlikowski and Scott 2008). Using Pickering’s (1995) “Mangle of Practice”, we reveal the human and material agency at work in the process of infrastructure deployment.

The rest of this paper is structured as follows: In the next section we discuss the literature on information infrastructure deployment and the technical elements of the Grid. This is followed by a section outlining our theoretical framework. This is followed by methodology and case study sections introducing the

research practices and context. The paper then presents an analysis of the case, concluding with theoretical and practical contribution for Grid development and information infrastructure deployment.

E-science and the GRID, Deployment and usability,

E-science and the Grid

E-science has been defined as the “*intersection* of Grid and collaborative research” (David 2004 original emphasis) and focuses on the use of technology within research practice. Grid is defined as “a system that coordinates distributed resources using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of service” (Foster and Kesselman 1998) and is central to many e-science activities. Baker et al (2002) summarize four main aspects that characterize a Grid: i. multiple administrative domains and autonomy; ii. heterogeneity; iii. scalability; iv. dynamicity or adaptability.

The process of deployment concerns the move from a experimental Grid to a “production” Grid. A production Grid can be defined as being “developed to enable large-scale scientific research rather than to be used as testbeds for software development or computer science research” (Johnston, Brooke et al. 2004, p537). Such Grids demand particular attention to security and quality of service, while many of their resources will almost certainly not be specifically designed for Grid use but will be appropriated into the Grid. For example the LHC studied here includes many existing data centres established before the Grid emerged, uses multiple scripting and programming languages, and uses two different Grid middleware suites (EGEE and Globus Toolkit) as well as the heterogeneous network infrastructure of the Internet.

Foster provides an hour-glass model of Grid architecture in which the grid is broken into layers (Figure 1). At the bottom we have the “fabric” of the grid providing interfaces to local resources (which may be physical or logical). This layer provides the local, resource-specific operations necessary, be they computer processors, storage machines, tape-robots, sensor, databases or networks. The resource and

connectivity layer defines the communication and authentication protocols required for transactions using the Grid network. Upon this layer rests a resource management function supported by directories, brokering systems, as well as monitoring and diagnostic resources. And then the grid can be used.

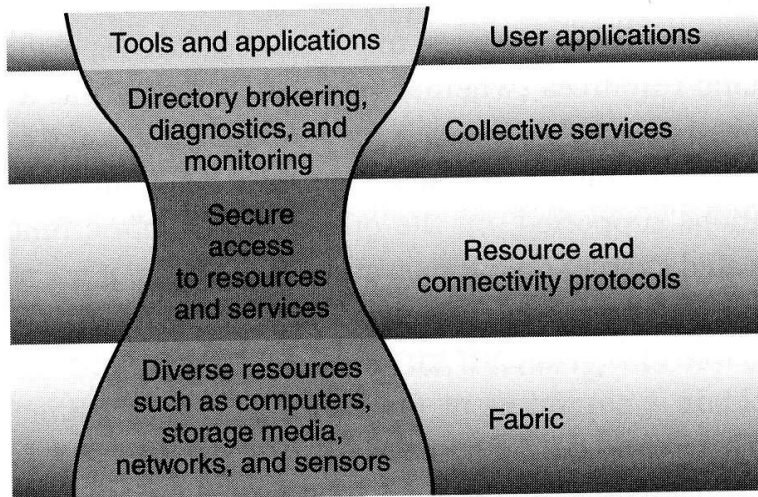


Figure 1 The Layered Grid Architecture from Foster 2004.

Grids must manage the resources they bring together. Central to achieving this is the concept of a Virtual Organisation (VO). A Virtual Organisation is a set of individuals and/or institutions defined by the sharing rules for a set of resources (Foster and Kesselman 1998) or “a set of Grid entities, such as individuals, applications, services or resources, that are related to each other by some level of trust” (Plaszczak and Wellner 2007). Such resources are not limited to data or files, but include direct access to computing power (TFLOPS), software, peripherals, and sensors. By necessity these resources must be controlled “with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs” (Foster and Kesselman 1998) and for this purpose VOs are technically defined along with the rules of their resources sharing. VOs can vary considerably in size, scope, sociology, structure and longevity (Foster and Kesselman 1998). The Compact Muon Solenoid experiment (CMS) <http://cms.cern.ch/> that we study here (see case study

below), is a virtual organization spanning many physical institutions, involving many physical resources, and united by the common goal of their experiment and the pursuit of ‘new physics’.

A Grid VO implies the assumptions of “the absence of central location, central control, omniscience, and an existing trust relationship” (Abbas 2004). It is the role of Grid middleware to manage these diverse and distributed resources and enact the virtual organization. Resources are presented to the VOs member as homogenous and standardized, and a member can just use the resources available to which their VO has the rights of access. The location of data, devices and sensors, and the allocation of processing resource are key to achieving this. One of the central elements of the middleware is the Grid Information Service (GIS) which keep track of Grid resources and their status including Storage Elements (SEs) which provide data storage and Computing Elements (CEs) which provide processing resource.

The Resource Broker (part of the Workload manager middleware) has the role of allocating users’ processing jobs to the SEs and CEs available within the Grid providing a “suitable architecture for distributed scheduling and resource management” (Avellino, Beco et al. 2003). The Resource Broker is in this process a crucial piece of the middleware and should, ideally, ensure that users do not need to target a specific machine, nor do they need to know about its location or ownership (Abbas 2004). As long as their credentials are acceptable to the VO, the Resource Broker should handle resources allocation, by matching job requests (from users) to the status of the Grid computing resources (as provided by the Grid Information System (GIS)) and match the jobs data requirements against the Grid file management services detailing the location of files on the Grid (Avellino, Beco et al. 2003).

Implementation and deployment

We have argued above that *deployment* is the appropriate term for the activities of setting a Grid to work. It is the point at which a largely experimental Grid (as a product of computer science research) aims to become a production Grid, secure, reliable and available for undertaking sustained scientific research. This is in contrast to more traditional concepts of *implementation*, which suggest a completed black-

boxed system being set to work by technology professionals (Kwon and Zumd 1987), with discussion of resistance and resisters (Keen 1981; Marcus 1983) central to the debate. Grid deployment, in the case presented here, is in contrast undertaken by relatively weak IT staff and systems administrators, and by scientists who are skilled, organised and powerful (what we call here ‘users’). The grid is also anything but black-boxed. Indeed it is being built and re-built as it is being used.

Issues of usability

Usability is usually defined as the ease of use and acceptability of a computer system, with respect to a particular task, in a particular environment, undertaken by a particular user (Holzinger 2005). It is associated with features of the technology with a focus on the design of such features for a generic user (Nielsen 1994) and as such is often technologically deterministic, focusing on improving software’s user-interface to enable the software to better influence practice; the aim is to perfect the technology such that it enables effective use.

Considering usability from a sociological perspective we observe technology as not simply a deterministic artefact to be ‘improved’ but as artefacts which are socially shaped (Mackenzie and Wajcman 1999). Woolgar and Grint (need ref. 1997) suggest what counts as technology is the outcome of a process of interpretation and attempts to avoid realist notions of a technologies features impacting usability. For such authors it is not an act of design that is required, but rather an act of socialisation. Other positions acknowledge a realist position in which technology is ‘usable’ because of features of the artefact itself which afford particular uses within practice (Norman 1990; Hutchby 2001). Relevant to this study is an acceptance that technology is socialised within reality (Berger and Luckmann 1966), and therefore for a technology to become usable by individuals in their work requires technology to be socially constructed though a process of negotiation among relevant social groups (Bijker, Hughes et al. 1987; Bijker 1995). In approaching usability of Grids in this study we maintain a set of key assumptions. Firstly, that relevant social groups have agency over the technology and are capable of adapting and

changing it fundamentally (rather than just altering its interfaces). Secondly, that technology has agency over relevant social groups practices. Thirdly that technology has affordance (Gibson 1979) – its inherent design alters or shapes practices (just as a tap affords turning). Fourthly, and crucially, we assume that at the micro-level of our analysis decisions by collectives do not play out. While standardisation bodies, collective decision making etc is vital to infrastructure acceptance and use (De Moor 1993; Hanseth 200x) our focus here is upon the deployment – a point at which such standardization has occurred, or is yet to occur. We appreciate that such influences are important, but for our analysis of practices shaping of technology we feel they can be discounted.

Theoretical Framework

Various theoretical perspectives have been used to investigate the process of developing a new piece of technology, including the sociology of science and sociology of technology (Latour and Woolgar 1986; Latour 1987; Pinch and Bijker 1989). Both are concerned with the construction of facts, values, and artifacts and offer a human-centred ontology, with humans constructing the facts and artefacts. In this spirit, the message of Orlikowski's seminal paper on *Technologies in Practice* (2000) is that peoples' work practices do change as technology becomes available, and those changes do 'reveal' the technology as something distinct for that context (a technology of control, a technology of team work). But this analysis does not quite apply here, since these authors' focus on human agency and enactment of emergent structures through recurrent use tends to downplay the ability of users to alter technical properties in use; further these authors tend to focus on one relevant group of users' shaping of a technology rather than seeing a network of communities battling to shape a shared technology to reflect their individual practices; and further these authors do not quite capture the interpenetration and growth that is central to an infrastructure, nor the power of the technical elements to steer both users and designers practices.

Work in the tradition of Science and Technology Studies (STS), goes a little further, allowing agency of things (material agency) to fight back (Callon 1986) (also Callon and Law 199x). For example Actor Network Theory (ANT) and related approaches in the STS tradition are concerned with the performance emerging from the dynamic interactions of humans and nonhumans. It seeks to give voice to material agency.

In this tradition, we acknowledge the unfolding influence of both human and material actors – their agency. The analysis is in the performative idiom, concerned with the practices of particle physics, of sysadmins, but also of software elements and their agency in the world of information infrastructure. We seek to uncover the "emergent interplay of human and material agency" (Pickering 1995) surrounding such infrastructure; an integration of human and machine agency deemed important for the field of information systems (Rose and Truex 2000). In this we might be seen as responding to recent calls to consider the materiality of technology (Orlikowski and Iacono 2001; Orlikowski 2007) (Orlikowski 2004, p185). To achieve this we adopt Andrew Pickering's "Mangle of Practice" (1993; 1995) as our theoretical framework.

The Mangle of Practice is founded on a belief that the world is a "lively place full of agency" (Pickering 2002), and that science is essentially performative; concerned with continually doing things with machines and instruments (Pickering 1993). This is certainly what we observe among the physicists studied here. Pickering's views of science practice is one in which scientists "manoeuvre in a field of material agency, constructing machines that (...) variously capture, seduce, download, recruit, enrol or materialise that agency, taming and domesticating it, putting it at our service" (Pickering 1995, p7).

Pickering suggests that human actors' interpretive accounts only become stabilised once their experimental actions lead to outcomes which make sense (Pickering 1995, p81) rather than being pre-specified or designed (see (Pickering 1999) for an illustration of this). Rather than material artefacts purpose and use being ordained, they are enrolled along the way and mangled in use (a view which reflects Gibson's notion of affordances (Gibson 1979; Hutchby 2001)). The scientist works not with specifications but with

“meaning conceptualisations of the aspect of the material world” which, through experimentation, they explore and transform, and in doing so translate their own interpretive accounts (Pickering 1995, p81). Understanding of the information infrastructure is therefore the outcome of the stabilising of mangling human and material agency, in which a range of actors are engaged.

The Mangle of Practice does not impose intentionality upon the material, (a common criticism of Actor Network Theory (Callon 1986; Latour 1987) – but see Law and Callon 2004). Rather, Pickering argues that the emergence of human and material agency are constitutively enmeshed. They are "mutually and emergently productive of one another" (Pickering 1993) through a dialectic; what Pickering terms the Mangle. He suggests that we cannot or should not talk about material intentionality, since it cannot be separated from human agency in this dialectic. Intentionality is restricted to human agency and posited as an endpoint – a focal purpose which is “temporally enduring” relative to the practices that occur, "a relatively fixed image of some future state" (Pickering 1993). People do things with things for an intended reason, and while technology makes the difference in this, it cannot share such intention.

Within the social contours of this study we observe two macro-actors with different intentionality and whose practices are mangled through the Grid. We term these macro-actors Physicists and SysAdmins and describe them further in the next section. For us macro-actors are a group of actors who share common practices, and among whom any accommodation of resistance is quickly shared through their informal networks. Similar to Communities of Practice (Brown and Duguid 1991; Wenger 1998) these are the means by which the arguably individualistic mangling of practice (Pinch 1999) becomes shared. These are also the community of users (we see sysadmin as ‘users’ but with different intentions) of the IT infrastructure, a community which must be naturalised to the Grid (Star and Ruhleder 1996). In the physicists we observe intentionality as pursuit of the “new physics” outcomes of the LHC and their individual tasks contribute to this effort. For SysAdmins intentionality is more complex but concerns the deployment of an effective and well organised Grid. Our study then explores the tension within this double-mangling of practice (Jones 1998) in which “both material and social agency are mutually and

emergently transformed” (Jones 1998, p297) and where "human agents seek to channel material agency to shape the actions of other human agents" (Jones 1998).

Jones (1998) notes the similarities between such mangling and Ciborra’s (2000) notion of “drifting”, a process which does not stop with the design of an artefact but extends through deployment into use. It seems appropriate therefore to employ the Mangle of Practice within a study of information infrastructure deployment where drift better describes the changing technology and practice than traditional ideas of discrete worlds of design, construction, implementation and use.

Pickering argues that studies should concentrate on capturing and framing the material agency of science as such intentions are sought. In Pickering’s terms, machine configurations (in our case the Grid) are the outcome of a “tuning process” (as in tuning a radio) which positions the machine within a “flow” of material agency so “harnessing”, “directing” and “domesticating” it and interactively stabilising both material and human agency (Pickering 1995, p278) towards a human goal. Drawing on Dona Harriway’s (1991) notion of the Cyborg, Pickering suggests scientists actions occur through “a supercyborg harnessing material and disciplinary agency in material and human performances, some of which lead out into the world of representation, of facts and theories” (Pickering 2005, p279).

Resistances, practical obstacles or anything which inhibits the pursuit of a indented goal (Pickering 1993)) are responded to by scientists with accommodations aimed at circumventing the obstacles (circumvention which may or may not succeed) which are in themselves revisions in the intentional structure of human agency (Pickering 1995). And in the Mangle material agency is “emergently transformed and delineated in the dialectic of resistance and accommodation” (Pickering 1993) and is undertaken by both human and material agency. Mangling of the material, conceptual and social context occurs – something we focus upon in our analysis of Grid Deployment.

This is a mangle of *Practice*. Practice is a work of cultural extension (broadly defined) (Pickering 1995, p3), in which the practices of particle physicists at the experiment “tune” the deployment of a Grid infrastructure (as locus of material agency) and so extend their particle physics culture. In this way the

scientists and sysadmins construct and transform the world of artefacts and through “transforming the furniture of the world...real novelty emerges in practice” (Pickering 1999, p308) and yet in which plans, interests and goals are changed and mangled. The community interacting with the technology is thus, in this mangle, formed and defined by the practices it creates and shares. For this we see Wenger’s notion of the community of practice (Wenger 1998) as highly relevant for describing this community.

Case Study and Methodology

We explore here mangling through a case study of the deployment of Grid technology and of the way that different intentionality leads to conflicting accommodations of resistance, and ultimately to different interpretive accounts of the correct use of Grids, the correct practice. The paper presents research undertaken into LCG’s (the LHC Computing Grid) deployment of a Grid for particle physicists at the LHC. LCG is a worldwide collaboration that aims to build and maintain a data storage and analysis infrastructure for the particle physics community. This is in preparation for the challenge of the imminent “data deluge” (Hey and Trefethen 2002) from the launch of the LHC at CERN, the European Laboratory for Particle Physics in 2008.

The LHC is designed to search for the Higgs boson, a search that has been described as like searching one person in a thousand worlds, or for a needle in twenty million haystacks (Britton, Clark et al. 2004). In this search the LHC envisages producing some 15 million gigabytes of data a year and LCG envisage requiring 100,000 computers in its Grid by 2008 spread across the globe (Faulkner, Lowe et al. 2006). We focus in particular on one of the LHC’s four experiments (CMS) and its development of a software suite to enable its experimental physicists to easily run jobs on the LCG. The Compact Muon Solenoid (CMS) collaboration consists of 2300 scientists and engineers from 159 institutes in 36 countries developing and running the CMS experiment on the LHC’s 26km ring. The analysis for CMS requires advanced computing expertise – at a minimum its physicists need the ability to operate Linux machines and programme in the C++ programming language. We take the particle physicists’ analysis of data from

the CMS experiment on the LHC as our unit of analysis, showing how the LCG grid system performs job submission for CMS, and the role CMS’s material media plays in mangling this process for both physicist and systems administrator.

Data collection was undertaken through over sixty interviews, participant observation in weekly project-management-board meetings of those involved in LCG in the UK (a group called GridPP), attendance at major GridPP meetings, site visits and through two week-long visits to CERN. The case study is elaborated through quotes from the individuals in Table I, and various documentary sources of evidence.

Table 1 – Key Interviewees		
Person Code	Job title	Date
P1	CMS Representative to UK LCG development (GridPP)	Jan 2007
P2	Post-Doc Research Assistant undertaking physics on CMS	March 2007
P3	Post-Doc Research Assistant undertaking physics on CMS	March 2007
P4	Oversees Grid middleware development for LCG and systems administration.	March 2007
P5	PhD student undertaking physics on CMS	March 2007
P6	Involved in middleware development and systems administration based at CERN.	March 2007
P7	A previous senior manager of the integration of experimental software with Grid services.	Jan 2007
P8	XXX	XXX
P9	XXX	XXX
P10	XXX	XXX

--	--	--

The analysis of the data is based on the premise that we need to examine the collaborative practices of grid deployment (as development *and* use) in order to better understand what “usability” of e-science means. Data is analysed using AtlasTi, though in an interpretive and flexible manner rather than for precise content analysis. Codes were produced reflecting the practices of participants. Additional interviews were undertaken specifically to explore the emerging themes of usability.

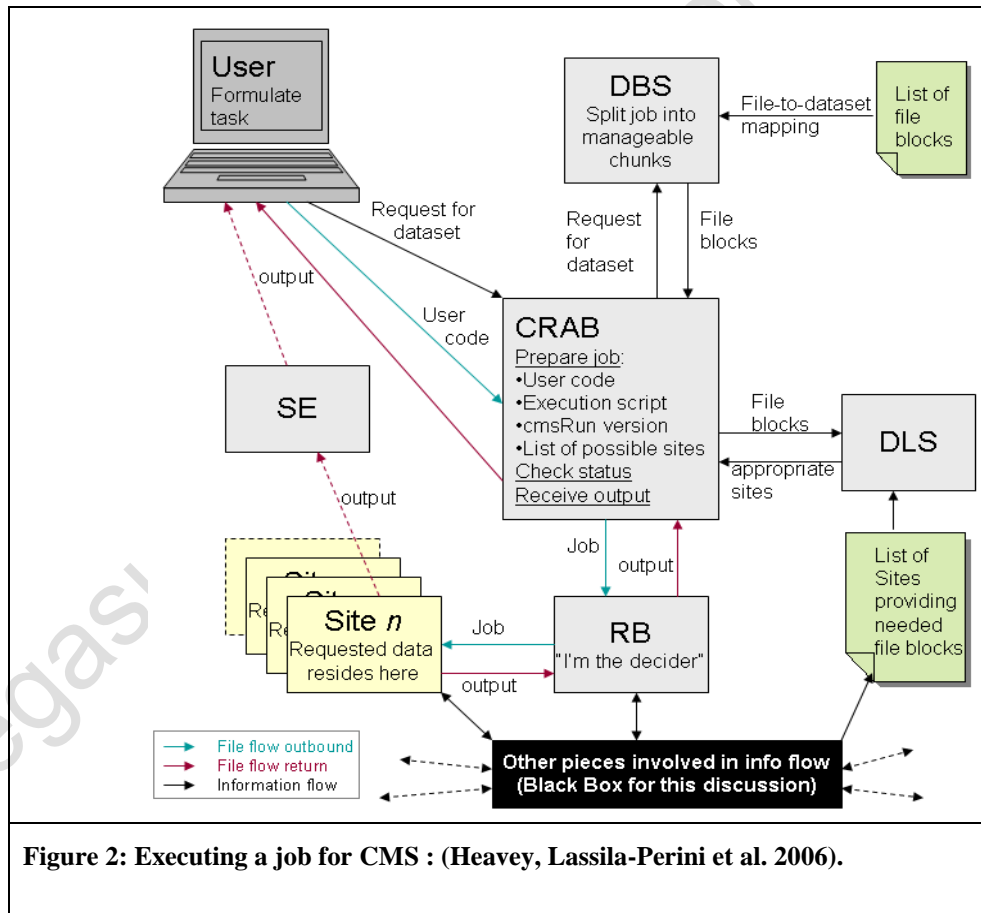
We employ the case study interpretively – “aimed at producing an understanding of the context of the information system, and the process whereby the information system influences and is influenced by its context” (Walsham 1993, p4-5). Our aim is to surface and explore this particular case, and from it ask questions of the role of users in information infrastructure deployment.

The particle physicist engaging in Grid practices: CRAB

From a particle physicist’s perspective, the intention of using the LCG is to undertake analysis which is too large to execute on a local computer or cluster and to undertake such analysis quickly since the experiments compete to publish results first. While “95% of [a researcher’s] time is actually spent developing algorithms, developing software, doing analysis, just completely locally on the laptop” (P1) they rely on a robust computing infrastructure to run their large occasional analysis. For them the Grid is “just a processing machine that sits there and nobody really has to know how it works except a few individuals” (P1). They demand the Grid only to undertake analysis as quickly as possible and extremely reliably despite the lack of efficiency of their own code: “People want enough disk space, they want fast CPUs that they can run their usually very inefficient programmes on, and they want it twenty four seven, and also during Christmas” (P2). This high demand for services, coupled with a highly pragmatic approach, is characteristic of the field (Hlistova 2004; Venters and Cornford 2006); they are “very dirty programmers, they are not computer scientists and they really will use the fastest way to get at

something... they usually want the fast hack” (P2). Indeed to use the LCG physicists must “see it as the fastest way to get somewhere because otherwise they will just do it another way.” (P2).

Interacting with the Grid to undertake such analysis is problematic due to its complexity and the time required learning its interface. This resistance is however accommodated by the CMS Virtual Organisation through its development of a range of software targeted at physicists’ needs, software that interacts on user’s behalf with the Grid. As stated in this software’s manual “The job of the CMS software (online and offline) is to select and process detected events [particle interactions], deliver the processed results to experimenters within the CMS collaboration, and provide tools for them to analyze the processed information in order to produce physics results” (Heavey, Lassila-Perini et al. 2006).



Interaction with the LCG involves a range of stages; writing analysis job software to perform the physics analysis (in the C++ programming language), preparing a script (a text file which instructs software) to undertake the analysis “job”, executing this script and hence running the job on the Grid, checking the status of the job, then finally receiving the output.

Within CMS this execution script is written for interface software which is called CRAB (CMS Remote Analysis Builder) and is produced by physicists within the CMS collaboration. CRAB is CMS specific and is “a python program intended to simplify the process of creation and submission of CMS analysis jobs into a grid environment.” (Heavey, Lassila-Perini et al. 2006). A user’s script is written as a `crab.cfg` configuration file specifying the details of the job. Figure 1, taken from the CMSOffline analysis workbook, outlines in detail of how Crab then interacts with the Resource Broker (RB), the Data Location Service (DLS), and the Dataset Bookkeeping Service (DBS) to prepare a job for execution on sites, and to move the results to a storage element (SE). Once a Crab configuration file has been prepared the physicist will run the jobs (users will usually run very large numbers of jobs at the same time). The first task is to create a proxy certificate for the jobs they will run (using command “`voms-proxy-init -voms cms`” on the command-line to create a proxy for the CMS virtual organisation). This proxy provides each job with the same permissions held by the CMS physicist so that the analysis code can be executed on machines (and with data) that the CMS Virtual Organisation has permission to access. To begin the analysis the physicist first creates the jobs using the command “`crab -create X`” where X is the number of jobs, then the command “`crab -submit -continue`” to submit all the created jobs to the Grid to be run.

Once a job is running the “`crab -status -continue`” command will allow the user to observe the status of their jobs. Users receive a list of their jobs with a status [ready | scheduled | running | done | clear | aborted | killed]. Once jobs are complete the physicist can retrieve the output from a storage element specified in `crab.cfg`.

Learning to use CRAB (particularly writing `crab.cfg` files) to undertake analysis using the LCG is a complex activity taking time away from practicing physics. Most physicists dislike attending training courses (“Users don’t go to [training sessions], They don’t want to talk to a bunch of geeks” (P8)) , and instead rely on word of mouth and through tailoring existing working scripts: “the way people learn how to use the Grid is by getting a working script from someone... and starting from there and trying to run it and trying to modify is to suite their needs” (P2). In modifying such a script physicists will use websites of FAQs, the CMS documentation, and various Wiki (e.g. <https://twiki.cern.ch/twiki/bin/view/Main/CRAB>) and Blogs to resolve problems (though seldom are these used to learn the initial steps). They will often e-mail each other scripts for help with problems and for advice.

Resistance and Accommodation on the LCG Grid

Given that the LCG Grid is under development it is inevitable that things will go wrong with Grid jobs, particularly when users regularly run millions of them. We now consider the response of physicists and sysadmin to the resistance of jobs failing and the role of physicists tinkering with CRAB in accommodating this resistance.

One of the complaints made by physicists is the difficulty knowing what has gone wrong with their jobs when they fail. Physicists intentionality concerns getting results back from the analysis quickly and they hence define success in these terms. In contrast SysAdmin deploying grid middleware consider a Grid success as the condition where the Grid has run something, even if the job has crashed, taken a very long, or similar problems have occurred.

The LCG Grid is designed to return a Zero (success) code in such cases. One physicist interviewed recounted the problem this causes physicists; “job-success, was always Zero, i.e. successful, no matter what happened. As long as the grid middleware had submitted a job, that it had run, and had crashed

somewhere, they gathered something back, that was considered a success. And this is just not very useful for physicists. For one thing it is impossible to monitor if something is going wrong” (P2).

For SysAdmin such ‘Zero-codes’ made sense. The Resource Broker responsible for job allocation is a problematic part of the Grid and in production use the software can behave erratically and understanding and removing problems “is significantly complicated by non-linear and non-continuous effects in the system” (Avellino, Beco et al. 2003). Automatically identifying when a Grid job has ended abnormally is seldom straightforward (Avellino, Beco et al. 2003) and the role of the Grid is resource allocation not necessarily ensuring a job’s success on a computing element (CE). While for sysadmin the Zero return code represented a successful accommodation of the problem of monitoring job success, it represented a resistance for physicists; “what frustrates users is that they follow the instructions, submit the job which sometimes works and sometimes it doesn’t. But due to stupid things somewhere the job fails” (Minutes of GridPP meeting).

Physicists were left responsible for failed jobs and, unable to understand the underlying problem, often perceived such failures as a problem with the interface software (CRAB) rather than with the underlying Grid: “when things aren’t being submitted they say the program is crap, even though it is not the actual program where things are going wrong, it is usually one of the different grid components, not the programme” (Interviewee responsible for documenting Grids). Indeed another interviewee who uses CRAB stated “there is a school of thought that the last letter should be changed to P” (P5)!

Faced with the resistance of a Zero code Sysadmin anticipate that physicists will accommodate this by interacting with them to resolve the problems with their job submission using a global support management system called GGUS (Global Grid User Support). Yet those physicists interviewed seemed extremely unclear of the purpose of GGUS and were reluctant to use it. Physicists appeared to lack information about resources for gaining support from sysadmin and didn’t appear inclined to search for such support unless problems were very recurrent. Further using GGUS often did not resolve problems which were sometimes sporadic..

Indeed CMS's provision of software (such as CRAB) , as an accommodation to users' problems with the complexity of running Grid analysis, had itself created an artificial barrier to gaining sysadmin support. A CMS physicist described: "I know that lots of money is going into support, but it seems to be decoupled, so there is grid support and kind of CMS software support, and I send a software job to the grid and it goes wrong, invariably the problem falls in the gap between the two. So the grid people don't know what I am talking about because I am talking CMS language and the CMS software people say it is a grid problem. So there seems to be this kind of gap at the moment between, almost between the grid and the experiment, that anything falls into that doesn't work, and so on" (P3). Historically physicists approached computing problems through informal communication with systems administrators, based on a cultural heritage where systems administrators were local and well known. With the Grid however physicists do not know who to contact to ask for support at that location where there jobs fail.

Without an effective communication route with the systems administrators of the problematic parts of the Grid, users are led to improvise new solutions to these problems. Instead of approaching sysadmin or using GGUS to resolve Zero-codes, most physicists innovated an accommodation of simply killing (`crab -kill N -c` where N is job number) jobs which took much 'longer than the rest' and trying again with these jobs. It is noted that 'longer' is subjective, often suggested as the time it takes for a coffee, while the last jobs at night might be given 12hrs before being killed. In the next section we discuss how CMS physicists further develop an accommodation to the resistance of Grid job failure.

Black-list and White-list

Individual physicists' needs are represented by their experiments (e.g. CMS) which have significant resources and technical capabilities to respond on users' behalf. CMS's members were therefore capable of improvising technical accommodations which represent their needs, whether or not they reflect the needs of the Grid administrators or other experiments (against which they compete for resources). In the words of one interviewee "I really do think that being the systems administrator for particle physicists

must be hell. [Physicists] will all want different systems for everything. And they are all quite savvy, they will tune everything” (P2). In this section we discuss how the accommodation of killing problematic jobs becomes bound into the material agency of CRAB, and explore this agency’s impact on the Grid, sysadmin and physicists. We focus in particular on the sysadmin’s perspective towards Grid failure.

When problems occur with a Grid site it is the role of the workload-manager middleware software to ensure that jobs submitted to the Grid continue to run (albeit more slowly and less efficiently). Sysadmin’s, whose intention is to provide a stable Grid service to all, harness this workload-manager to stabilise the Grid through a tuning process in which sites are targeted or excluded from the Grid. As one Sysadmin stated “ [we] have a workload manager which is able to do resubmission automatically on behalf of the user in case the site where the job was sent is not behaving correctly. Or we have the possibility in the information system for the VO [Virtual Organisation] to ban explicitly, given sites, so there are white lists and black lists. So [a central Grid administrator] can black-list a site, [and they] can white list another site, if [they] want. So that basically the workload management system is forced to consider all and only the sites left that the application wants to be used”(P4). This feature is only however intended to allow the tuning of the workload manager by sysadmins and represents a domesticating of the Grid for sysadmin goals.

The aim of the CMS collaboration’s software is to “schedule jobs onto resources according to the policy and priorities of CMS” (Heavey, Lassila-Perini et al. 2006), priorities focused on the CMS experiment rather than the priorities of the Grid as a whole, or the priorities of sysadmin. Faced with the resistance of difficult to identify failures, poor user support, and problematic areas of the Grid which drop jobs for unknown reasons, CMS’s physicists response was to enrol the above middleware agency in order to provide CRAB with the facility to force Grid jobs to execute on particular areas of the Grid, or to exclude a particular area of the Grid. Crab was therefore developed to include the ability to black-list elements of the Grid (so disallowing jobs to run on a site’s computing or storage elements). Similarly Crab included white-listing allowing physicists to specifically target Grid nodes for a job to run on (see Figure 2). This

feature essentially harnessed and institutionalised sysadmin's accommodation, and enrolled it to enable physicists to overcome a set of material resistances from the Grid.

This facility in Crab enabled physicists to continue to use the Grid in a none-grid way similar to their existing practices, something welcomed by the physicists interviewed. One recounted that using this facility meant dealing with a problematic Zero return is easy; "in that case I quite often try and send the job somewhere else, not use that particular Grid site" (P2). Similarly for another physicist using Crab in this way was standard practice; he stated: "I use a tool called Crab, which is a CMS tool which allows you to sort of specify places. You can specify places for it not to go to or places that it should go to. So I can try sending it somewhere else, but sometimes what I end up doing is ignoring, not using the Grid at all, and go back to old fashioned, submit directly to a computer here" (P3). Another similarly exploited this facility to circumvent the Grid - "I can send a job directly without using the grid. So I guess half the time I probably do that, in the end." (P3)

```
If you want/need to select/deselect some site, you can use: (see Crab FAQ for more info)
```

- Ce_black_list - (refuse access to all the listed CEs, allow all others)
- Ce_white_list - (allow access only to those CEs listed)
- Se_black_list - (remove the selected SE from the list of sites hosting data)
- Se_white_list - (select only the SEs listed)

```
-----  
## CE Black List: all the CE whose name contains the following strings (comma
```

```
## separated list) will not be considered for submission.
```

```
## Use the dns domain (eg fnal, cern, ifae, fzka, cnaf, ln1,...)
```

```
CE_black_list = ...  
-----
```

```
So, in summary, if you want to force your jobs to go a specific site (eg if you want to test the site), use "SE_w/b_list". If instead you want to access some dataset but you want to avoid a
```

```
site (because you don't trust it), use "CE_w/b_list". In addition, se_w/b_list cannot be used with None as input dataset.
```

Figure 2: CMS FAQ's detailing CRAB.cfg options (truncated)

These black-listing and white-listing options break the fundamental aim of Grids in providing “coordinated resource sharing” (Foster and Kesselman 1998; Foster and Kesselman 2004) since they circumvent the material agency of the Grid workload-manager (with its aim to balance work across the Grid) by providing physicists with agency over the workload manager to direct their jobs to particular resources. Indeed a facility which was only intended for sysadmin to tune the Grid through testing sites, or avoiding specific untrustworthy sites (see Figure 2 final part) is now regularly exploited by CMS physicists to speed up their individual analysis activity. This is outlined by a CMS physicist; “In practice the way I operate is; I find out where my data is and tell this Crab tool where to go. My experience with telling it just to go and find the data [using the workload manager] is that that does not work, ever. So that may be something specific to do with how far we have got with these tools. But up until now, definitely, from my experience, the kind of real physics applications that I want to run, the only way it works is you find out where the data is and then you tell the thing to send there” (P3). This mangling of human and material agency has here also mangled the conceptual model of Grids held by CMS physicists – for them the Grid now includes “sites” and involves directing jobs to “locations” – far from the conceptual aims of Grid technology held by Sysadmin in which a work-load manager manages the allocation of jobs to sites and users simply ‘plug-in’ to an unknown resource.

Such accommodation by CMS creates further resistance for SysAdmin. An interviewee who was developing the Grid middleware explained the problem of experiments producing software for their physicists, and implementing CRAB type tools in detail: “As people who are trying to put together a middleware release we are trying to find the best solution to each particular problem domain, integrate it into our release so everyone can use it. Now if there was a user sitting in isolation they would probably

have to use what we provide, they don't really have much choice. But these users don't really exist so much, they all work for experiments with lots of influence and resources and everything. And which occasionally, possibly often, have very high influence in some of the sites as well, they can ask the sites to install various services. So they can bypass stuff. Bypassing is probably a pejorative phrase, it is just they choose to use an alternative route. (...) But certainly one of [systems administrators] big services is workload management. So the idea is that this takes all your jobs and manages them for you, submits them to the right place, so you send them there and forget about them until you all come back. But on your user interface you can implement most of this stuff, if you want to, to your own satisfaction. And we find people have done that.” (P6)

The CMS physicists realize that their approach creates difficulties for systems administrators, however their individual intentionality towards physics analysis and the pressure to perform this quickly is paramount in their minds “I see that the way we use (D. Britton and et al. on behalf of the GridPP collaboration) may cause other people headaches. But it is simply because otherwise you get into this really boring alternative which is you submit it without specifying where to go. And then it comes back from one place and says it doesn't work. So you say – OK, ignore that place, try somewhere else. You know you are gonna make it work, but you shortcut that by finding out... so this one thing of knowing where your data is, if you can find that out independently it releases more time to make it happen” (P3).

From a sysadmin perspective however the white-listing and black-listing creates significant resistance. While the workload-manager is appreciated to be problematic, often creating failed jobs, circumventing it provides little opportunity for improvement and distorts the Grid's usage patterns; there are “some parts that are very much more used than others” (P9). Sysadmin's role is to observe the usage of the Grid through a variety of monitoring and dashboard applications and ensure that the computing elements and storage elements their site contributes to the Grid are running effectively and have the correct software infrastructure installed. They are formally evaluated on the availability of the SEs and CEs they administer. This is no easy task, particularly as in preparation for the launch of the LHC many sites are

installing large numbers of new storage and computing elements. Further each site's CEs and SEs are likely to be shared by a range of experiments at the LHC, and indeed in other sciences. Further the Grid middleware is changing regularly and the powerful experiments regularly demand new software be installed on sites CEs. Such complexity means that the skills of sysadmin are crucial to a site's success, yet their skills vary; "the biggest problems today is managing such a big infrastructure where there are so many sites and not all the sites are managed with the same level of body of quality. At the same time the software which implements the servers is not mature enough in order to be easy to be managed. (...) this requires a very experienced people in managing sites which is not always the case" (P4).

Crucially in their preparation for the data deluge once the LHC's experiments begins taking data, sites require regular challenges to their infrastructure by large numbers of physics jobs to ensure the site can handle the required volumes. Until the LHC begins operating there is a lack of data to analyse, and therefore a lack of demand for computing and storage resources. There is thus a resultant competition between Grid sites to get physics jobs to stress their systems, but with physicists (using black and white lists) able to choose to run their jobs at the sites easiest for them. These are usually sites they know are working well (and hence often do not need as much stress testing) or sites local to them where they can easily contact systems administrators for support rather than through GGUS. In particular CERN has a large number of CEs and SEs which many rely upon. Further the experiments have developed independent monitoring systems "that report independently the situation of the infrastructure as they see it. Not as the [sysadmin] see it" (P4). This enables these VOs to domesticate the Grid infrastructure for their physics aims, without reflecting the cultural aims of sysadmins or other user groups. Each experiment is "their own fiefdom" (P10) maintaining a strong control over what is done and reflective of their practices.

In summary CMS physicists have domesticated the Grid, stabilising it as a practice similar to the existing approaches of submitting jobs directly to computing-clusters whose location is known. This however creates significant resistance for Sysadmins in their plight of creating a stable Grid.

Discussion

The practice of particle physics has been described as an ongoing process of accommodation and resistance in which the materiality of things is employed (Pickering 1995) and it therefore is perhaps unsurprising that CMS enrolled technology in response to resistance with the Grid. Physicists accommodated the Grid into their practices through informal interaction among the CMS community, with existing `crab.cfg` files shared and tailored by new Grid users, and with Crab itself is produced by this community to represent its members' needs and to reduce their Grid interaction to a simple scripting process. Indeed the `crab.cfg` user-manual discusses little about the Grid as a material artefact, focusing wholly on the process of its use – the practice to do physics research on the CMS experiment. `Crab.cfg` is thus an accommodation and domestication of the Grid; an artefact with agency for its users, representing their practices.

Yet `crab.cfg` is bounded within a wider system involving failed Grid nodes, and the CMS collaboration's response to such failures.. In this context it is perhaps unsurprising that CMS, and other experiment collaborations, faced with such resistance to their physics have accommodated it by enrolling facilities only intended for sysadmin to aid their understanding, fixing and testing of problems with the LCG. The CMS VO's approach to the problems with LCG has been to tailor its material resources (in particular the Crab software) to provide its physicists with agency over failing sites.

By allowing black-listing and white-listing individual CMS physicists were able to respond to their own perception of problems on the Grid by targeting sites they believe work effectively, or by black-listing sites which they believe fail, and in doing so speed up their analysis on a Grid with plenty of resources available. These physicists adopted this practice not as an extreme means of testing or sanction, but as common practice for interacting with the Grid. Such targeting might however ossify over time as the shared “working” `crab.cfg` scripts (which include black-lists or white-lists) are passed on. Such

established practices mangle the concept of Grid by including the notions of “good-sites” and “bad-sites” within a resource intended to be homogenous.

On the other side of these actions lie systems administrators, an often forgotten community in studies of Information Systems. For them the main agent for allocating jobs to sites should be the Workload manager software. They argued that were every Grid user to employ black/white-listing then “the impact ultimately ... [would be] that the workload management system would be scrapped, because it would be investment in something useless” (P6) – an action which would fundamentally change the concept of “Grid” to little more than a network of Cluster-computers. Their monitoring of the Grid and its success is distorted by the `crab.cfg` targeting of jobs. They must attempt to continue to develop their Grid resources without a balanced view of Grid activity, and with the workload-manager being sidelined by targeted `crab.cfg` scripts. For systems administrators, and middleware developers, are faced with a powerful and technically proficient user community developing applications which exploit the Grid for their own particular needs, rather than accepting it as a homogenous service to all. Those developing the middleware simply hope that as their middleware improves more of the experiments will move to use it rather than continuing to develop their own application software representing their own needs.

It is worth stressing that the CMS case is not unique. The ATLAS experiment at CERN has similarly developed software which mangles the Grid to reflect their users’ needs. In contrast to CMS users’ use of black-list and white-list, ATLAS users have created “pilot jobs” which are submitted to all CEs on a Grid. These jobs simply run on the CE waiting for messages from ATLAS users, whereupon they collect the actual physics-jobs from the ATLAS user and run them on their CE. In effect these pilot jobs act as ATLAS’ personal work-load management infrastructure and wholly circumvent the Grid workload manager.

Conclusions

We draw on Pickering's (1993) "mangle of practice" which suggests that the practice of science is based on messy ill-defined material artefacts. These artefacts are created and shaped through a dialectic of resistance and accommodation that shifts CMS physicists through the space of all potential arrangements they can think of (Pickering 1993). It is through this dialectic, we argue, that usability of Grids is accommodated (and in turn influenced) by the human agency of users. In this case study human agency in doing physics-with-grids were temporally emergent from the practices of doing physics (with its focus on results). Resistance emerged from doing physics (hence Zero return codes were problematic as they failed to focus on doing physics), and accommodations (by the CMS VO) were made within the general practice of physics in which individuals faced with resistance improvise and bricolage solutions which they share with the collaboration. The 'Grid-in-use' is thus not technologically deterministic in its influence on practice, but rather negotiated and contested between material agency of the Grid, and the human agency of users as they accommodate its various resistances in their purposeful pursuit of physics through the mangle of practice (Pickering 1993).

Using Pickering's Mangle of Practice we have elaborated a case study of Grid users and systems administrators engaged in such a dance of material agency. While the direct interaction between these groups is limited, their dual 'tuning' of the materiality of the Grid influences each's work practices. In this way the case study extends existing discussions of Information Infrastructure development which has suggested that infrastructure shapes, and is shaped by, communities of practices' conventions (Star and Ruhleder 1996), or drifts in usage as "new ways of using it are discovered" (Ciborra, Braa et al. 2000, p137). In contrast we demonstrate that for various communities of users such shaping can involved tension or battle, and that such tension can be played out by fundamentally altering the information infrastructure such that material and human agency are mangled to incorporate the intentionality of the community. Here users' intentionality is not the unintended side-effect, but a powerful actant in the process of change.

The case provides lessons for those considering the deployment of information infrastructure by extending the notion of infrastructure shaping as living or self-reinforcing to include a double-dance (Jones 1998) (or perhaps double-battle) of various agencies. We observe the Grid shaped by human and institutional agencies (CMS Physicists, sysadmins, middleware developers and ATLAS) and by material agencies (CRAB, Work-load manager, Grid and GGUS). While managerialist arguments concerning infrastructure usability rely upon power in the form of “top management support” (Weill and Broadbent 1998) we observe how power is imbricated in technical agency and the materiality of the Grid (for neither sysadmins or physicists were sanctioned as powerful). Beyond the realm of organisational hierarchy it is perhaps this “mangling of the material” though which infrastructure emerges; a managing which means infrastructure usability is contested, redefined and violently altered – (not simply designed, ‘tinkered’ with or ‘drifting’).

What then for the “designer” of Grids? Developing a Grid for e-science is perhaps to undertake an exercise in co-aligning technical and social elements with the practice of users. Our study suggests, that Grid ‘users’ in particle physics are not passive in this realignment process, but rather, as a virtual organisation (CMS), were active agents responding to the grid. We respond to the assertion of the need to challenge the designer/user opposition (Suchman 1999) by recasting the roles of users and developers in LCG (a theme espoused by Woolgar & Coopmans (2006)). We suggest that for LCG the ‘user’s’ practice involves responding to resistance in usability by accommodations which may themselves be technical Grid developments (like Crab). The process of design of Grids, in this case, is a co-evolution of the accommodations of the designer (attempting to produce a generic Grid middleware for all) and user (attempting to produce an accommodation that allows their intentions to be achieved). We observe here the desire to create standardised Grid interfaces (analogous to the electrical power Grid) is proved problematic by users’ mangling to create specialist interfaces tailored to the specific needs of their temporary local physics practices.

Our study suggests that further research is needed into the role of systems administrators in information infrastructure deployment. This group has often remained in the shadows of Information Systems research, perhaps considered an unproblematic or boring group. Our study highlights that if Grid's are to be achieved, their role is vital and highly entwined. This lesson from the study has relevance beyond Grid research as similar technologies such as Service Oriented Architecture, Virtualisation and off-shoring of data centres proliferate. It therefore seems important that our discipline re-examines the role of systems administrators in this complex web of technical and material resources.

Finally, our study introduces a relatively unused theoretical framework within our field – Pickering's "Mangle of Practice" and unusually employs this empirically as a theoretical framework to explore the dual agency at play in a case study. We believe that our choice has the advantage of including material agency within the analysis, but by keeping the distinction between the technical and organisational actors we are able to explore how technology is enrolled into a battle between two communities of practice. Importantly intentionality (as a human attribute only) provides a useful lens towards the driving force behind the enrolment of material agency. Further the mangle of practice as a linear account of tinkering (or "trial and error" as Pinch (1999) critiques the Mangle) seems to be highly appropriate to the study of "tinkering" and "drift" in the usability of Grids.

There are limitations to our study. Particle Physics at CERN is a unique endeavour and their practices are arguably unique (Knorr-Cetina 1999). Whether the lessons of this case study are relevant to other communities is thus in question. That said, the practices of this community – autonomous, self-directed, distributed and highly technical are reflective of emerging forms of work practices in open-source, virtual organisations and outsourcing relationships (Venters and Cornford 2006). We hope therefore that our lessons prove resonant beyond this, to some, rather obscure domain.

- Abbas, A. (2004). Grid Computing: A Practical Guide to Technology and Applications. MA, Charles River Media.
- Avellino, G., S. Beco, et al. (2003). The EU DataGrid Workload Management System: towards the second major release. Computing in High Energy and Nuclear Physics, La Jolla, California.
- Avellino, G., S. Beco, et al. (2003). The first deployment of workload management services on the EU DataGrid Testbed: feedback on design and implementation. Computing in High Energy and Nuclear Physics, La Jolla, California.
- Baker, M., R. Buyya, et al. (2002). "Grids and Grid Technologies for Wide-Area Distributed Computing." Software-Practice and Experience **32**: 1437-1466.
- Berger, P. and T. Luckmann (1966). The social construction of reality. London, Penguin Books.
- Berman, F., G. Fox, et al. (2003). The Grid: past, present, future. Grid Computing - Making the Global Infrastructure a Reality. F. Berman, G. Fox and T. Hey, John Wiley & Sons, Ltd.
- Bijker, W. (1995). Of Bicycles, Bakerlites and Bulbs; Toward a theory of Socio-Technical Change. Cambridge, MA, MIT Press.
- Bijker, W., T. Hughes, et al., Eds. (1987). The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. Cambridge, MA, MIT Press.
- Britton, D., P. Clark, et al. (2004). A Grid for Particle Physics - from testbed to production, GridPP.
- Brown, J. S. and P. Duguid (1991). "Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation." Organization Science **2**(1): 40-57.
- Brown, J. S. and P. Duguid (2000). The Social Life of Information. Boston, Massachusetts, Harvard Business School Press.
- Callon, M. (1986). Some Elements For A Sociology of Translation: Domestication of the Scallops and the Fishermen of St-Brieuc Bay. Power, Action and Belief: a New Sociology of Knowledge? J. Law. London, Routledge and Kegan Paul: 196-229.
- Callon, M. (1986). Some elements of a sociology of translation: domestication of the scallops and the fisherman of St Brieuc Bay. Power, Action and Belief: A New Sociology of Knowledge. J. Law. London, Routledge and Keegan.
- Chae, B. and M. S. Poole (2005). "Surface of emergence in systems development." European Journal of Information Systems **14**: 19-36.
- Chetty, M. and R. Buyya (2002). "Weaving computational grids: how analogous are they with electrical grids?" Computing in Science & Engineering **4**(4): 61-71.
- Ciborra, C. (1999). Hospitality and IT. Informatics in the next millennium. F. Ljungberg. Lund, Studentlitteratur: 161-173.
- Ciborra, C. (2000). From Control to Drift: The dynamics of Corporate Information Infrastructure. Oxford, Oxford University Press.
- Ciborra, C., K. Braa, et al. (2000). From Control to Drift: The Dynamics of Global Information Infrastructures, Oxford University Press.
- Ciborra, C. and O. Hanseth (1998). Toward a contingency view of infrastructure and knowledge: An exploratory study. International Conference of Information Systems (ICIS). Helsinki, Finland: 263-272.
- D. Britton, A. J. and et al. on behalf of the GridPP collaboration (2006). Performance of the UK Grid for Particle Physics. IEEE06 Conference, . Amsterdam, *to be edited*.
- Dahlbom, B. (2000). Postface: From infrastructure to networking. From Control to Drift. C. Ciborra, K. Braa, A. Cordella et al, Oxford University Press.
- David, P. A. (2004). Towards a cyberinfrastructure for enhanced scientific collaboration: providing its 'soft' foundations may be the hardest part. Oxford, Oxford Internet Institute.
- de Certeau, M. (1984). The practice of everyday life (Trans. Steven F. Rendail). Berkeley, University of California Press.
- Edwards, P., S. Jackson, et al. (2007). Understanding Infrastructure: Dynamics, Tensions and Design. Report of a Workshop on History and Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures. . Ann Arbor, University of Michigan.
- Faulkner, P., L. Lowe, et al. (2006). "GridPP: development of the UK computing Grid for particle physics." Journal of Physics G: Nuclear and Particle Physics. **32**: N1-N20.

- Feldman, M. (2000). "Organizational Routines as a Source of Continuous Change." Organization Science **11**(6): 611-629.
- Foster, I. and C. Kesselman (1998). The Grid: Blueprint for a New Computing Infrastructure, Elsevier.
- Foster, I. and C. Kesselman (2004). The Grid 2: Blueprint for a New Computing Infrastructure, Elsevier.
- Foster, I., C. Kesselman, et al. (2001). "The anatomy of the Grid." International Journal of Supercomputer Applications **15**(3): 200-222.
- Foster, I., C. Kesselman, et al. (2001). "The Anatomy of the Grid: Enabling Scalable Virtual Organizations." The International Journal of High Performance Computing Applications **15**(3): 200-222.
- Gibson, J. J. (1979). The Ecological Approach to Perception. London, Houghton Mifflin.
- Hanseth, O. (2002). From systems and tools to networks and infrastructures - from design to cultivation. Towards a theory of ICT solutions and its design methodology implications.
- Hanseth, O. and K. Lyytinen (2004). "Theorizing about the Design of Information Infrastructures: Design Kernel Theories and Principles." Sprouts: Working Papers on Information Environments, Systems and Organizations **4**(4): 207-241.
- Hanseth, O. and E. Monteiro (1998). Understanding Information Infrastructure.
- Haraway, D. (1991). A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century. Simians, Cyborgs and Women: The Reinvention of Nature. New York, Routledge: 149-181.
- Heavey, A., K. Lassila-Perini, et al., Eds. (2006). The CMS Offline Workbook, CERN.
- Hey, T. and A. Trefethen (2002). "The UK e-Science Core Programme and the Grid." Future Generation Computer Systems **18** 1017-1031.
- Hlistova, J. (2004). How is the Particle Physics Community Shaping the Grid in the UK? Department of Information Systems. London, London School of Economics: 38.
- Holzinger, A. (2005). "Usability Engineering Methods for Software Developers." Communications of the ACM **48**(1): 71-74.
- Hughes, T. P. (1983). Networks of Power: Electrification in Western Society, 1880-1930. Baltimore, Johns Hopkins University Press.
- Hutchby, I. (2001). "Technologies, Texts and Affordances." Sociology **35**(2): 441-456.
- Johnston, W., J. Brooke, et al. (2004). Production Deployment: Experiences and Recommendations. The Grid 2: Blueprint for a new computing infrastructure. I. Foster and Kesselman, Carl. Amsterdam, Elsevier.
- Jones, M. (1998). Information Systems and the Double Mangle: Steering a course between the scylla of embedded structure and the charybdis of strong symmetry. Information Systems: Current Issues and Future Challenges, Working Conference (with WG 8.6), Helsinki, Finland.
- Knorr-Cetina, K. (1999). Epistemic Cultures: How the sciences make knowledge. Cambridge, MA, Harvard University Press.
- Knorr-Cetina, K. (2005). How are global markets global? The architecture of a flow world. The sociology of financial markets. K. Knorr-Cetina and A. Preda. Oxford, Oxford University Press.
- Latour, B. (1987). Science In Action. Cambridge MA, Harvard University Press.
- Latour, B. (1987). Science in action: how to follow scientists and engineers through society. Cambridge, MA, Harvard University Press.
- Latour, B. and S. Woolgar (1986). Laboratory Life: the Construction of Scientific Facts, Princeton University Press.
- Lovelock, J. (1987). Gaia: A New Look at Life on Earth. Oxford, Oxford University Press.
- Mackenzie, D. and J. Wajcman (1999). The Social Shaping of Technology. Maidenhead, Open University Press.
- Markus, L. (1987). "Power Over Users: Its exercise by system professional." Communications of the ACM **30**(6): 498-504.
- Nielsen, J. (1994). Usability Engineering (Interactive Technologies), Morgan Kaufmann.
- Norman, D. (1990). The Design of Everyday Things. London, The MIT Press.
- Orlikowski, W. (1993). "CASE Tools as Organisational Change: Investigating Incremental and Radical Changes in Systems Development." MIS Quarterly **17**: 309-340.
- Orlikowski, W. (1996). Evolving with Notes: Organisational Change around Groupware Technology. Groupware and Teamwork. C. Ciborra, John Wiley & Sons Lts: 23-59.
- Orlikowski, W. (2004). "Material Works: Exploring the Situated Entanglement of Technological Performativity and Human Agency." Scandinavian Journal of Information Systems **17**(1): 183-186.
- Orlikowski, W. (2007). "Sociomaterial Practices: Exploring Technology at Work." Organization Studies **28**(9): 1435-1448.
- Orlikowski, W. and S. V. Scott (2008). The Entangling of Technology and Work in Organizations, London School of Economics.

- Orlikowski, W. J. and C. S. Iacono (2001). "Research Commentary: Desperately Seeking the "IT" in IT Research--A Call to Theorizing the IT Artifact." Information Systems Research **12**(2): 121 -134.
- Pickering, A. (1993). "The mangle of practice: Agency and emergence in the sociology of science." American Journal of Sociology **99**(3): 559-589.
- Pickering, A. (1995). The Mangle of Practice. Chicago, Chicago University Press.
- Pickering, A. (1995). The Mangle of Practice: Time, Agency and Science, University Of Chicago Press
- Pickering, A. (1999). "Reponses and Replies: In the land of the blind...thoughts on Gingras." Social Studies of Science **29**(2): 307-11.
- Pickering, A. (2002). "Cybernetics and the Mangle: Ashby, Beer and Pask." Social Studies of Science **32**(3): 413-437.
- Pickering, A. (2005). Concepts and the Mangle of Practice Constructing Quaternions. 18 unconventional essays on the nature of Mathematics. R. Hersh, Springer.
- Pinch, T. (1999). "'Mangled up in Blue'." Studies in History and Philosophical Science **30**(1): 139-147.
- Pinch, T. F. and W. E. Bijker (1989). The social construction of facts and artifacts: Or how the sociology of science and the sociology of technology might benefit each other. The social construction of technological systems. T. Pinch, T. P. Huges and W. Bijker. Cambridge: Mass, MIT Press.
- Plaszczak, P. and R. Wellner (2007). Grid Computing: A Savvy Manager's Guide. Amsterdam, Elsevier.
- Rose, J. and D. Truex (2000). Machine Agency as perceived autonomy; an action perspective. IFIP 9.3 TC9, Springer.
- Smarr, L. (2004). Grids in Context. The Grid 2. I. Foster and C. Kesselman. San Francisco, Morgan Kaufmann.
- Star, S. and K. Ruhleder (1996). "Steps toward an Ecology of Infrastructure: Design, Access for Large Information Space." Information Systems Research **7**(1): 111-134.
- Star, S. L. and K. Ruhleder (1996). "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces." Information Systems Research **7**(1): 111-135.
- Suchman, L. (1999). Working relations of technology production and use. The Social Shaping of Technology. D. Mackenzie and J. Wajcman. Maidenhead, UK, Open University Press: 258-269.
- Suchman, L. (2000). "Organizing Aligmnnet: A Case of Bridge-building." Organization **7**(2): 311-327.
- Venters, W. and T. Cornford (2006). Introducing Pegasus: An ethnographic research project studying the use of Grid technologies by the UK particle Physics Community. Second International Conference on e-Social Science, Manchester, UK.
- Venters, W. and T. Cornford (2006). Introducing Pegasus: An ethonographic research proejct studying the use of Grid technologies by the UK particle physics community. National e-Science Center Conference. Nottingham?
- Von Hippel, E. (2005). Democratizing Innovation. Cambridge, MA, MIT Press.
- Weill, P. and M. Broadbent (1998). Leveraging the New Infrastructure: How Market Leaders Capitalize on Information Technology. Cambridge,MA, Harvard Business School Press.
- Wenger, E. (1998). Communities of practice : Learning, meaning and identity. Cambridge, Cambridge University Press.
- Woolgar, S. and C. Coopmans (2006). Virtual Witnessing in a Virtual Age: A prospectus for Social Studies of E-Science. New Infrastructure for Knowledge Production. C. Hine. Hershey, Information Science Publishing: 1-25.
- Zheng, Y., W. Venters, et al. (2007). Agility, Improvisation, or Enacted Emergence? 28th International Conference on Information Systems (ICIS 2007), Montreal, Canada.