Commercial Exploitation of Grid Technologies and Services

Drivers and Barriers, Business Models and Impacts of Using Free and Open Source Licensing Schemes

Final Report by
SCF Associates
for DG Information Society and Media
25 November 2006
Contract no. 30 – CE – 065970 / 00- 56
Simon Forge and Colin Blackman
This Final Report is delivered as one part of the deliverables for the above contract.

**This report reflects the views of the authors and not necessarily those of the European Commission**

Commercial in Confidence. No disclosure without permission of the European Commission.

SCF Associates Ltd, 4 Chiltern Close, Princes Risborough HP27 0EA, UK
Tel +44 78 66 60 13 52
CONTENTS

EXECUTIVE SUMMARY ................................................................................................................................. 6

INTRODUCTION .............................................................................................................................................. 9

CHAPTER 1. MARKET OPPORTUNITIES FOR GRID SERVICES, SOFTWARE APPLICATIONS AND MIDDLEWARE ................................................................................................................................. 11
  1. INTRODUCTION: GRID COMPUTING IN CONTEXT .............................................................................. 11
  2. MARKET DEFINITION OF GRID PRODUCTS AND SERVICES .......................................................... 15
  3. THE GRID MARKET TODAY AND ITS GROWTH OVER THE NEXT FIVE YEARS ............................... 24
  4. GRID MARKET TAKE-OFF .................................................................................................................. 29

CHAPTER 2. TECHNICAL AND NON-TECHNICAL DRIVERS AND BARRIERS TO GRID SOFTWARE AND SERVICES .......................................................................................................................... 35
  1. WHAT IS THE PROBLEM? ..................................................................................................................... 35
  2. BARRIERS ........................................................................................................................................... 43
  3. MARKET DRivers FOR GRID MIDDLEWARE AND SERVICES .......................................................... 55

CHAPTER 3. BUSINESS MODELS FOR GRID SOFTWARE AND SERVICES ............................................. 76
  1. THE INDUSTRIAL CONTEXT – THE DEMAND SIDE ........................................................................... 76
  2. CURRENT BUSINESS MODELS IN THE SOFTWARE AND SERVICES INDUSTRY ......................... 80
  3. NEW BUSINESS MODELS AND BUSINESS DEVELOPMENT PRACTICES .......................................... 94

CHAPTER 4. ANALYSIS OF LICENSING MODELS FOR FREE, OPEN SOURCE AND PROPRIETARY SOFTWARE .............................................................................................................................. 106
  1. OPEN SOURCE LICENCES AND THEIR LEGAL AND COMMERCIAL ASPECTS ............................... 106
  2. QUESTIONS ON COMMERCIAL USAGE OF OSS LICENCES ........................................................... 113
  3. CHOOSING AN OSS LICENCE ............................................................................................................ 126
  4. WHAT ARE THE OVERALL IMPACTS OF OSS LICENCES ON COMMERCIALISATION? .................. 135

CHAPTER 5. SYNTHESIS OF POLICY ON GRID INDUSTRIALISATION FOR FRAMEWORK PROGRAMMES ................................................................................................................................. 149
  1. INTRODUCTION AND CONTEXT ................................................................................................. 149
  2. THE CONTEXT FOR A EUROPEAN POLICY AND ITS CORE STRATEGY .......................................... 150
  3. POLICY TACTICS FOR THE POTENTIAL MARKET ......................................................................... 157

LIST OF TABLES

Table 1.1. Different names, same thing? .................................................................................................. 13
Table 1.2. Worldwide grid spending, 2006-2011 (Shillion) ................................................................. 25
Table 1.3. European spending on IT for four industrial sectors (2006) ................................................ 26
Table 1.4. Growth in European IT spending for four sectors, 2006-2011* ........................................... 27
Table 1.5. Forecast for grid spending in Europe for four sectors, 2006-2011 ................................. 27
Table 1.6. Characteristics of the scenarios ............................................................................................ 30
Table 3.1. Main revenue streams by grid function with services within each (2005) ......................... 88
Table 4.1. Some common OSS licences and others, by type ................................................................. 110
Table 4.2. Key alliances and projects and their licence types ............................................................... 111
Table 4.3. Comparison on an industrial usage basis of the positions of the various licences .......... 134
Table 4.4. Linking business models and licence types ........................................................................... 137
Table 4.5. Projects’ selections of licences types, set in context ........................................................... 138
LIST OF FIGURES

Figure 1.1. Grid and related paradigms ............................................................... 12
Figure 1.2. Virtualising computing resources...................................................... 14
Figure 1.3. Dynamic IT end-state: rationalized, flexible, and efficient IT. ............ 16
Figure 1.4. Key needs for grid middleware .......................................................... 17
Figure 1.5. The value chain for grid computing .................................................... 18
Figure 1.6. Users, needs and suppliers................................................................. 20
Figure 1.7. Three main grid markets set the key business models ......................... 22
Figure 1.8. Worldwide grid spending by type of organisation, 2005-2010 ($billion) .... 24
Figure 1.9. Forecast European grid spending, 2006-2011 ..................................... 28
Figure 1.10. How the grid market might take-off can be seen through various scenarios 29
Figure 1.11. Potential evolution of the grid market .............................................. 31
Figure 1.12. The fastest route to market ............................................................... 32
Figure 2.1. Major technical barriers to grid uptake ............................................. 43
Figure 2.2. Major non-technical barriers by impact for perception, marketing and legal issues 47
Figure 2.3. Education and understanding are major barriers to grid take-up ........ 48
Figure 2.4. Inhibitors to investment in grid computing technologies ................. 54
Figure 2.5. Major technical facilitators for grid technology take-up ..................... 55
Figure 2.6. Key elements of an open grid reference model for Europe ................. 58
Figure 2.7. The process for creating a common toolkit for Europe to simplify choice 60
Figure 2.8. Sharing applications reduces server and storage resources ................. 62
Figure 2.9. Specific management layers must be present in grid middleware ....... 64
Figure 2.10. The interoperability and test platform ............................................ 66
Figure 2.11. Non-technical drivers of the grid market ......................................... 71
Figure 2.12. Product packaging and go to market preparation for launch ............... 78
Figure 3.1. A view of the grid technology ecosystem from the GGF .................... 78
Figure 3.2. A commercial view of the ecosystem of the grid technology segment ... 79
Figure 3.3. IBM grid computing architecture .................................................... 85
Figure 3.4. Main revenue streams by grid function with services within each (2003) .. 88
Figure 3.5. General steps for configuring a grid computing deployment with the customer 90
Figure 3.6. New business models for the grid industry ........................................ 94
Figure 3.7. How do OSS middleware products from R&D projects get to market? 95
Figure 3.8. Business models for grid products following the stages of market development 96
Figure 3.9. Exploiting the output from R&D projects as a range of marketable grid services 97
Figure 3.10. Including a contractual layer in the middleware ................................ 99
Figure 3.11. The capabilities offered by the GISP for the CD02 business model .... 100
Figure 3.12. A first model for a pervasive grid utility – the world wide grid (WWG) 102
Figure 3.13. The grid sold as an electrical or gas utility ...................................... 103
Figure 4.1. The OSS licence cycle and IPR .......................................................... 119
Figure 4.2. Commercial v research project licence objectives ............................. 136
Figure 4.3. The steps in choosing a licence ....................................................... 145
Figure 5.1. South Korea’s usage of directed technology diffusion ....................... 151
Figure 5.2. The Korean grid computing industrial structure ................................ 153
Figure 5.3. Policy actions to seed the next phase of grid market development ....... 155
Figure 5.4. The business ecology of a successful grid infrastructure industry ....... 163
Figure 5.5. Each layer of the WWG offers services opportunities ....................... 166
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APL</td>
<td>Apache License</td>
</tr>
<tr>
<td>APL2</td>
<td>Apache License Version 2.0</td>
</tr>
<tr>
<td>BSD</td>
<td>Berkeley Standard License</td>
</tr>
<tr>
<td>CMDB</td>
<td>Configuration Management Database</td>
</tr>
<tr>
<td>CRM</td>
<td>Customer Relationship Management</td>
</tr>
<tr>
<td>DAI</td>
<td>Data Access and Integration</td>
</tr>
<tr>
<td>EGA</td>
<td>Enterprise Grid Alliance, now OGF</td>
</tr>
<tr>
<td>FP6</td>
<td>The EU’s 6th Framework Programme for Research and Technological Development</td>
</tr>
<tr>
<td>GAS</td>
<td>Grid Application Service</td>
</tr>
<tr>
<td>GASAAS</td>
<td>Grid Application Software as a Service</td>
</tr>
<tr>
<td>GASP</td>
<td>Grid Application Service Provider</td>
</tr>
<tr>
<td>GGF</td>
<td>Global Grid Forum, now OGF</td>
</tr>
<tr>
<td>GISP</td>
<td>Grid Infrastructure Service Provider</td>
</tr>
<tr>
<td>gLite</td>
<td>Lightweight middleware for Grid computing</td>
</tr>
<tr>
<td>GPL</td>
<td>General Public License (GNU)</td>
</tr>
<tr>
<td>GSI</td>
<td>Grid Systems Integrator</td>
</tr>
<tr>
<td>GSM</td>
<td>Groupe Spécial Mobile - Global Systems for Mobile Communications</td>
</tr>
<tr>
<td>GSP</td>
<td>Grid Service Provider – also Grid Security Protocol (Globus)</td>
</tr>
<tr>
<td>GSSP</td>
<td>Grid Support Service Provider</td>
</tr>
<tr>
<td>GT</td>
<td>Globus Toolkit</td>
</tr>
<tr>
<td>ISV</td>
<td>Independent Software Vendor</td>
</tr>
<tr>
<td>LGPL</td>
<td>Lesser General Public License (GNU)</td>
</tr>
<tr>
<td>MPL</td>
<td>Mozilla Public License</td>
</tr>
<tr>
<td>OGF</td>
<td>Open Grid Forum</td>
</tr>
<tr>
<td>OGSA</td>
<td>Open Grid Services Architecture</td>
</tr>
<tr>
<td>OGSADF</td>
<td>Open Grid Services Architecture Development Framework</td>
</tr>
<tr>
<td>OGSI</td>
<td>Open Grid Services Infrastructure</td>
</tr>
<tr>
<td>OSS</td>
<td>Open Source Software</td>
</tr>
<tr>
<td>SAAS</td>
<td>Software as a Service</td>
</tr>
<tr>
<td>SLA</td>
<td>Service Level Agreement</td>
</tr>
<tr>
<td>SNA</td>
<td>Systems Network Architecture</td>
</tr>
<tr>
<td>VAR</td>
<td>Value Added Reseller</td>
</tr>
<tr>
<td>VO</td>
<td>Vertical Organisation</td>
</tr>
<tr>
<td>WSRF</td>
<td>Web Services Resource Framework</td>
</tr>
<tr>
<td>WWG</td>
<td>World Wide Grid</td>
</tr>
</tbody>
</table>
Executive summary

This report presents the findings of a major study \(^1\) into the commercial exploitation of grid technologies and services in Europe. Specifically, the study focused on drivers and barriers, business models and the impacts of using free and open source licensing schemes. This executive summary presents the key messages arising from the study.

The European Commission has given considerable support to a variety of IST RTD grid projects through the Fifth and Sixth Framework programmes. This major investment in the knowledge society could be the pre-cursor to the creation of jobs and GDP growth through the creation of a vibrant grid computing sector in Europe. With the ground work in place, a key question is how to enable technology transfer from R&D projects into commercial services and products.

There are many indications that the grid market is set to take off. Grid technologies are moving from a first phase of development as research usage for e-Science into applications for vertical industry applications in finance and banking, pharmaceuticals, and manufacturing industry segments such as automobiles.

Over the next decade grid technologies are likely to fundamentally shape future IT architectures as companies and the public sector see clear economic benefits in widespread adoption. The possibility of utility grid computing, where computing resource is available in a similar way to electricity, remains a possibility. Forecasts for spending on grid computing vary but the European grid market could be worth €9 billion by 2011.

However there are several barriers – technical and non technical – that will need to be overcome if we are to see widespread take up of grid technologies. On the technical side, there are a range of barriers inhibiting a complete grid offering, including:

- Security, protection of both resources and the local environment and of the running (‘visiting’) applications and data
- Firewalls and grid working
- Common standards for grid build and interworking
- Variations in Web Services from vendors – so no real common standards for Grid interworking
- Heterogeneity of resources and the complexity/fragility of the grid environment – adapting to distributed resources and their virtualisation
- Performance, specifically communications performance
- Application compatibility for grid environment (parallelisation/virtual hosting) for a loosely coupled parallel processing architecture

Important though these technical issues are, it is the non-technical questions that present the most significant obstacles to grid take up. These range from basic issues of understanding and perception to marketing matters and also legal questions, for instance:

---

\(^1\) The study was carried out by SCF Associates Ltd on behalf of the European Commission’s DG Information Society and Media, Grid Technologies Unit F2 between May and November 2006.
The study concludes that the most fundamental barriers that must be overcome are to do with understanding of the place and utility of the technology, and also over trust and perception of its maturity. This strongly suggests that a major education effort for the market is required. Efforts to overcome legal constraints when running grids in software and server licences are also key.

If we look to the current commercial market and into the future, we can discern three main grid market segments which in turn will determine the underlying business models:

Seen in this way, the opportunities opened by grid technology imply a large range of potential business models, which may be as much in services as in products. There is also a future possibility of EU grid technology becoming the basis for a common utility grid, of on-demand computing for Europe.

What also is clear from the study is that decisions over licensing are crucial – in particular, the choice of OSS licence is critical for future commercial exploitation so that licences must match business models. The report presents a process for choosing appropriate licences taking into account the longevity of the project, the intended final outcome, commercial roles and wishes of participating project members, and the market(s) for which the deliverables were originally intended. The choice ranges from the restrictive GNU General Public License (GPL) at one extreme through the Lesser GPL and the Mozilla Public
License to the Apache License and the Berkeley Standard License, the latter of which allows broad freedom for commercial exploitation.

The report concludes with an examination of the major policy implications, both for the European Commission but also more broadly for all stakeholders in grid technologies. It is recommended that future policy should concentrate on educating and opening up the market with specific promotional mechanisms and establishing a permanent framework within which to set them.

The European Commission has an important enabling role to play in helping industry and research stakeholders to establish this framework, which could take the form of a European Grid Agency which would take primary responsibility for all the legal and market support functions necessary to build a European grid industry. Its functions would include:

- Acting as a centre for IPR ownership to prevent IPR hi-jack and spurious patenting
- Being the point of legal liability and referral
- Taking responsibility for code cleaning and certification to give market confidence
- Providing a download centre for ownership and operation for OSS
- Supporting the creation and promotion of a common reference model and toolkit for commercial products
- Running a test bed for interoperability and certification of interworking
- Providing technical expertise, support and update of offerings
- Offering commercial packaging of common model and tool kit with documentation.
Introduction

The context for this study is the projects under the Framework Programmes for IST Research and Development and their potential for commercialisation to build an EU grid industry. EU policy on grid technology is to support their research and development under Framework Programmes for pre-competitive research in which the results are largely the property of the participants\(^2\) in joint, collective or co-operative ownership.

The study highlights a number of technical and non-technical obstacles, but note that these kinds of problems are common to any new technology (eg mobile communications). Thus, the views expressed in this report should be seen as constructive rather than critical.

Grid computing is seen as a key technology for the Lisbon agenda of moving Europe as whole towards a knowledge-based society. In consequence DG Information Society and Media, Unit F2, has commissioned this study. Consequently there is a need for a concerted effort to create a European grid technology and middleware community that can forge and carry through an effective approach to industrial strength of this technology. Action is required to organise the European “community” of grid technology and middleware development with grid-application research into a focus for commercialisation.

One of the main avenues to grid industrialisation in the European situation is use of open source software. However guidance is needed on software licensing issues and schemes and legal strategies in relation to OSS. Furthermore, what are the business models needed, with the evolution of value-added services, and the current and future directions of the grid market? The main drivers and barriers, and how to overcome them, with a review of how an ecology of the grid may occur had to be examined. Any conclusions from this should persuade industry to begin making the investments and culture shifts necessary to participate more fully in the grid community.

To reply to such demands, we have pursued this study in order to produce a report that looks at the difficulties in building a route map to commercialisation for grid computing. It contains chapters on market analysis, analysis of drivers and barriers for commercialisation, business models and software licences review, with recommendations, following the requirements laid out in the invitation to tender for a strategy for technology transfer for the IST grid RTD projects to follow towards commercialisation. The final chapter is a succinct synthesis of the policy required to move forward. Our major findings are now outlined.

Policy tactics to achieve the EU’s strategic objectives are explored through a range of mechanisms. One key recommendation is for an agency that brings together existing grid middleware activity and expertise, with a certification role, a common toolkit, reference model and a software repository etc, to be based on an open source initiative. Its role would include provisions for information campaigns, for grid technology to capture the imagination of the wider business community, and the public with actions is to create and communicate the business vision for European industry. It is not the role, however, for the

\(^2\) Guide to Intellectual Property Rights provisions for FP6 projects, Version 2, 28/06/2006, EC, Research DG Directorate M (Investment in research and links with other policies) Directorate A (Coordination of Community Activities)
European Commission to establish such an agency – but it does have the ability to help other stakeholders to do so. A further suggestion is to form a grid technology industry grouping to exploit and build an industry around the reference model and toolkit, following the previous examples of successful European collaborative efforts to bring high technology to market.

Our approach has been through research into existing studies and reports and in speaking in depth with more than twenty industry stakeholders, from grid users through vendors to FP6 project participants. This process formulated our ideas and shaped our findings. During this we found three repeated threads:

- The grid offering itself – what it means and could mean in the future – is extremely unclear. The more one speaks to players and stakeholders, the more views and opinions one hears – a single consensus view did not emerge. Many are concerned with very specific pieces of grid technology and usage, so to hear a genuine overview is rare as most users and vendors are often too close to their own particular technical, project or commercial problems.

- The one possible consensus point that came across was that the problem of security and trust is the major barrier to be overcome before widespread take-up can occur. But the overall way to move the technology forward into commercialisation was not obvious, or articulated by any commentator.

- As may be expected with its origins in e-Science, there is a general view in the literature and in many stakeholders that grid technology is about very high power computing. There are far fewer views on mundane usage, be it as a mechanism for savings on server spend through application sharing, or as a generic way forward for mass computing, although we did find some.

Moreover views in the USA are somewhat different, with many commercial reports on grid computing covering the grid industry from the viewpoint of consisting of 20 or so middleware vendors, mostly in the USA, aimed at high power computing. However, our context for policy is also the distinctive European vision of a grid environment that operates through services offerings and not just products and from a range of computing levels – from handheld devices to supercomputers – in order to serve individuals, enterprises, virtual organisations (VOs) or whole industries. This view should be contrasted with the US vision of grid systems, generally as programmer-level super computing for the enterprise or research grouping. The European vision is more aligned with the objectives of the Lisbon initiatives for a knowledge-based society.

Examination of past high technology development successes in other countries and regions indicates that a pragmatic interventionist policy might be useful in the European context for a grid industry, if judiciously applied in concert with the major players. This approach requires a change in mindset and one where project funding and project exploitation plans target areas where Europe can realistically achieve success.

There are multiple choices in the strategy to pursue for this type of policy – two leading directions are either to let the market build an industry, using significant support at a European level, or else, to go further and build a new extension to what is considered our infrastructure, at the level of a grid computing layer above the Internet for all to use in open fashion. An industry of software and services would grow from this infrastructure. The final policy section then explores the necessary catalysts suggested from previous chapters to drive commercialisation from the point of view of building a European industry. Views on software licensing, and the key barriers and drivers are drawn together.
CHAPTER 1. Market opportunities for grid services, software applications and middleware

1. Introduction: grid computing in context

In the past few years the prospects for grid computing moving out of the laboratory and into the market have been talked up but what is the reality behind the hype? Are we witnessing the early stages of a new market and, if so, which products and services will be in demand? What is the value chain going to look like for this market? These are some of the questions that are addressed in this chapter on market analysis.

Understanding how the market for grid computing might develop over the next five years first requires an understanding of the context – what is the origin and the vision for the grid, and how does the technology fits into the overall evolution of IT infrastructure? Indeed, what exactly is grid computing?

1.1. The vision for the grid

The idea that computing would be available ubiquitously goes back at least to the 1960s, when UCLA professor Leonard Kleinrock suggested in 1969 that ‘…we will probably see the spread of computer utility, which, like present electric and telephone utilities, will service individual homes and offices across the country…’. The term ‘grid computing’ was popularised in the mid-1990s, largely owing to Ian Foster, to describe a proposed computing architecture for scientific/engineering computing with similar transparency and universal usage as the global power grid. ‘Grid computing’ as originally conceived, should therefore be seen simply as a metaphor for the availability of distributed computing power analogous to the availability of electricity through the power grid. Exactly how grid computing might be configured, precisely what its properties might be and whether there might be a demand for what it could offer is a work in progress.

In practice, grid computing has its roots in parallel and distributed computing. Massively parallel solutions, mainly designed for expensive and specialised parallel computing hardware, dominated the high-end computing market until the 1990s. The original motivation for grid computing was the need for a distributed computing structure for advanced science and engineering research, with considerable focus on the importance of collaborative and multi-disciplinary functions (e-Science). Thus, the first ‘grid’ applications focused on aggregating unused processing in desktop computers to provide scientists with supercomputer levels of performance at relatively low cost, eg SETI@home project to search for extraterrestrial intelligence.

---

6 http://setiathome.berkeley.edu/
Implementation of grid computing within this ‘eScience’ context has stayed true to the initial concepts but has also led to debate about what may or may not be termed grid computing. Thus, definitions of grid computing have proliferated. For instance, grid computing is typically defined as:

a software system that provides uniform and location independent access to geographically and organizationally dispersed, heterogeneous resources that are persistent, supported and under different ownership or control.

From this perspective, some activities which exhibit some grid characteristics, such as clustering or distributed computing within a single organisation, fall outside a strict definition of grid computing (see Figure 1.1).

Figure 1.1. Grid and related paradigms

However, as business is beginning to show more interest in grid computing, it is being defined more broadly. For instance, Ian Foster’s 2002 three-point definition casts the net more widely. He defines a grid as a system that:

- coordinates resources that are not subject to centralized control
- using standard, open, general-purpose protocols and interfaces
- to deliver non-trivial qualities of service.

---

8 Colin Upstill, IT Innovation Centre, personal communication, 07 June 2006.
Perhaps this reflects a change to a more realistic view that the implementation of grid computing in the commercial context may not necessarily follow in a simple linear transition from the academic and research environment to the business world. Research in the academic or e-Science context may have many benefits (e.g., developing a pool of skilled workers) but research projects in themselves may not necessarily find a route to market.

Grid definition is being taken further, into the concepts of a commons in grid infrastructure, for a form of computing utility in recent studies, which reflect a more European view of what grid computing could become in the future.\textsuperscript{11}

1.2. Awareness of grid computing

There is no doubt that awareness of grid computing has grown steadily in the past few years. Even though grid computing is a relatively new concept for the business community, a November 2004 survey found that about a quarter of the IT management community was already comfortably familiar with grid.\textsuperscript{12} Awareness of underlying concepts of resource pooling and virtualization was even higher.

However, it is also apparent that there is confusion about exactly what is meant by the term ‘grid computing’. Forrester surveyed 149 large companies in North America to learn about their knowledge and use of grid technologies.\textsuperscript{13} Firms are confused about what "grid" means but they are implementing various types of grid technologies, with 37\% of firms reporting that they’re piloting, rolling out, or have implemented some form of grid technology. Another 30\% of firms are considering grid technology, so two-thirds of all firms are using or interested in grid technology. Thus, according to Forrester, firms use “grid” as a general term when referring to a variety of things including parallel computing, distributed computing, organic IT, utility, on demand, etc.

Part of this confusion arises because of the plethora of names and concepts promoted by different vendors and other groups to describe essentially the same thing – the ability of organisations to deploy resources when and where needed rather than allocating work to dedicated resources (see Table 1.1).

Table 1.1. Different names, same thing?

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gartner</td>
<td>Real-Time Infrastructure</td>
</tr>
<tr>
<td>IDC</td>
<td>Dynamic IT</td>
</tr>
<tr>
<td>Forrester</td>
<td>Organic IT</td>
</tr>
<tr>
<td>Dell</td>
<td>Scaleable Enterprise</td>
</tr>
<tr>
<td>Hewlett-Packard</td>
<td>Utility Computing</td>
</tr>
<tr>
<td>IBM</td>
<td>On Demand</td>
</tr>
</tbody>
</table>

This would tend to support Forrester’s point of view, which is that grid computing is not a unique market but rather an element in the transformation of IT.\textsuperscript{14}

\textsuperscript{11} Future for European Grids: GRIDs and Service Oriented Knowledge Utilities- Vision and Research Directions 2010 and Beyond, Next Generation GRIDs Expert Group Report 3, January 2006
\textsuperscript{13} Frank E. Gillett, Grid Gets Big, But The Term Is Confusing, Forrester
1.3. A new paradigm for IT infrastructure?

Thus, initially grid computing was characterised as a ‘disruptive technology’ but more recently it is being viewed as the next step in the logical evolution of IT. For instance, some analysts argue that what is emerging is the fourth stage in the evolution of IT infrastructure – mainframe computing was superseded by client/server architecture in local area networks, while network computing and the Internet represents the third stage and grid/on-demand/utility computing represents the fourth stage. The point here is that grid is likely to be more palatable offer and seen as less threatening to potential customers if it is just viewed as the next logical step rather than a disruptive and paradigm shifting technology.

Seen in this way, grid computing is just one of a number of key technologies and concepts – including computing on demand, utility computing, organic information technology (IT), adaptive computing, and Internet computing – which are the building blocks of this new paradigm. In this new world, heterogeneous computing resources will be shared over networks and reallocated dynamically across applications, processes and users.

A key point to note here is that the fundamental driver of these changes in architecture is virtualisation. According to a 2006 report by US analyst group Yankee Group, three out of four companies, regardless of their size, already have deployed, or plan to deploy, server virtualisation technology over the next year. Of those, 62% already have virtualisation in place or are in the process of migrating. Only four per cent said they had no virtualisation plans. What we are seeing here is a process that progresses from virtualising like resources and then unlike resources within an enterprise, to virtualising the enterprise itself through enterprise wide grids with the next stage being virtualisation outside the enterprise including suppliers, customers and partners (see Figure 1.2).

Figure 1.2. Virtualising computing resources

---

From this perspective, the purity of the definition of grid computing becomes less important. Here we can see how there is impetus within business to make the best use of IT resources first at the departmental level, second within the enterprise as a whole, and finally (and possibly) in collaboration with other organisations. If this is indeed the case, how then does the world of grid computing within the enterprise relate to the grid computing as developed within the eScience laboratory? According to Insight Research:

Grid computing is really a complex of techniques, however, all having to do with virtualization of resources in one way or another. Since there are so many facets to how and when grids will be deployed, it is an oversimplification to state that grid technology emerged in the scientific and research domain and is now crossing over to the commercial domain; the reality is that there are pockets of movement in one aspect of grids in a particular industry and different types of progress in another industry.

2. Market definition of grid products and services

2.1. The key needs in IT infrastructure and grid software

Many businesses face a growing need for ever-greater computational processing power to run business critical applications. At the same time it is well known that the majority of servers and PCs are under-utilised. Most servers only run at between 20-30% of their total capacity while PCs typically run at between 5-20% capacity. This suggests that business and public sector organisations will be attracted to grid computing because it will enable this unused capacity to be utilised by centrally managing and distributing processing power across the organisation. Further, companies may see economic benefits in selling spare capacity to other organisations through a grid utility. It may also mean that organisations that typically struggle to keep pace with developments in IT will be able to access state-of-the-art resources when they are needed.

An increasingly important need for organisations as they attempt to compete effectively is the need for flexibility and agility in business operations. In general terms this means having the ability to deploy resources (including capital, equipment, human resources and information) quickly and efficiently where and when they are required. This inevitably means that IT systems can be placed under demands that may change rapidly. This is something that organisations have typically found difficult in responding to because of the manual nature of the task – management decision, deploying staff, availability of hardware, installing software, testing, etc. This has led to the common practice of allocating hardware on an application-by-application basis, with each application-specific server or group of servers usually being sized to deal with peak demand. Inevitably this means that IT resources are being wasted. Moreover, this approach is very inflexible. In response, the world is changing, such that future architecture is moving away from dedicated silos and hardware for specific applications to a more flexible and dynamic architecture that is more efficient (see Figure 1.3).

---

What, then, are the implications of this in terms of key requirements for grid middleware (see Figure 1.4)? Grid middleware typically acquires and maintains:

- management information about applications and resources
- execution management services for jobs submitted
- data access and transfer management, and
- security through authorization, role and access privileges.

According to Hiro Kishimoto, the key requirements or features of grid middleware are:

- Wide distribution
- Fault-tolerant
- Secure
- Manage complexity
- Modular
- Composable
- Scalable
- Extensible
- Interoperable
- Standards-based
- Transparent support for legacy components

---


21 http://www.ggf.org/GGF17/materials/316/OGSA%20keynote%2020060510.ppt
Figure 1.4. Key needs for grid middleware

2.2. The value chain for grid computing

Essentially the value chain for grid computing is the same whether viewed from the more narrow perspective of a single entrepreneur or as true grid computing for multiple virtual organisations. Figure 1.5 shows the various segments of the value chain moving from grid products (hardware, networks, system resources, grid software as both ‘lower’ or open source and ‘upper’ or proprietary grid middleware) through grid product support services (systems integration and sales and support for grid software) to grid services.

---

2.3. Main industry players in the value chain

The main industry players in the value chain can be broadly classified into the major IT platform providers and the more specialist independent software companies. The major IT platform providers include:

IBM
Hewlett-Packard
Sun Microsystems
Oracle
Microsoft

Grid Independent Software Companies include:

Avaki
Data Synapse
Platform Computing
United Devices
Entropia
GridSystems
Tsunami Research
Univa
2.4. Who are the customers for grid computing?

We can categorise the customers for grid computing in several ways. First we can define them today as being:

- Academia and research
- Public sector
- Commercial enterprises

**Academia and research**
The development of grid computing has largely taken place within the academic and research communities. Here the emphasis is on collaborative working and utilisation of resources for discrete processes that require considerable processing power. Typical of this is particle research at CERN.

**Public sector**
The public sector has long been a major user of high power computing. This has led to far better weather models, predictions of crop yields, population growth etc in applications combining statistical and quantitative analysis, sometimes with complexity theory. However, the major publicly funded application support for grid comes from military applications. These cover calculations for aerospace projects, weapons research, especially nuclear weapons such as yield calculations, ordnance research and radar analysis, also in security surveillance in either sifting large amounts of data, or in the oldest application of electronic computing, decryption of encrypted signals. Applications beyond this in eGovernment and eHealth are just beginning with use of brain scanning in real time for instance, for brain surgery.

**Commercial enterprises**
A key attraction of grid computing for the business sector is its potential to reduce costs and early take up in commercial enterprises will be found in sectors where the likelihood of savings is greatest. The main general reasons cited for interest in grids are:

- Reducing hardware costs
- Utilizing spare computing cycles
- Increasing cooperation and communication

There may also be defensive reasons to pursue grid computing, ie to ensure that a company does not fall behind developments taking place in rival firms that might give them a competitive edge. A key concern, however is sensitivity over commercial data, which is explored in more detail in Chapter 2. Thus, Figure 1.6 shows the major European industrial sectors and their key demands for products and services as well as the categories of player for the supply of those products and services.
Thus, some of the applications that can already be seen in the commercial sector include:

- Energy (for oil and gas for exploration)
- Finance/insurance/real estate (securities and brokerage – especially for stock/portfolio analysis and risk management)
- Aerospace and Automotive (for collaborative design and modeling)
- Life sciences (particularly in pharmaceuticals)
- Media/entertainment (to generate digital animation)
- Architecture (engineering and construction)
- Electronics (design and testing)
- Manufacturing (inter/intra-team collaborative design, process management)
- Utilities (to improve efficiency while dealing with peaks and valleys in utilization).

Here we focus on four vertical industry sectors: financial services, pharmaceuticals, automobiles and aerospace, and digital media.

Pharmaceuticals

A 2005 report\(^{23}\) found that although some leading pharmaceutical companies had established significant grid deployments for drug discovery, considerable scepticism remains about the value of grid computing in the sector. The pharmaceutical sector would appear to be a natural adopter of grids because of its need for high-performance computing but compared with industries such as financial services, the pharmaceutical sector is

lagging. No drug has yet been discovered using grid computing but this would only seem to be a matter of time and this is likely to be the stimulant others in the industry need.

Beyond drug discovery, there is also the opportunity for wider use of grid computing across pharmaceutical enterprises. Some early adopters of grid computing technology in this sector, eg Johnson & Johnson, Novartis, are seeking to adopt grid computing technology across all parts of their business.

Financial services
The financial services sector has been at the forefront of adoption of grid computing. Research by The 451 Group shows that at least 75% of major investment banks are already using grid computing. Applications have been driven by individual lines of business, and to meet specific needs such as complex Monte Carlo simulations. The potential for cost savings to banks is enormous through constantly estimating a bank’s exposure to risk and lowering risk reserves accordingly.

Running these simulation models requires tremendous computing power the costs of which are significant. Thus there is a big incentive to reduce operating costs and grid computing is seen as the way to do this. However, there are significant challenges. Evolving grid computing from individual lines of business is difficult because lines of business (LoBs) typically haven’t trusted other groups’ applications enough to want to share compute facilities (except mainframes). Sharing resources will require non-trivial policy and prioritisation management, the implementation of chargeback and accounting, and implies IT control issues.

Automobile and aerospace
Manufacturers in industries such as aerospace or automobiles typically face the same challenges. The design and engineering challenges revolve around the need to integrate a complex mix mechanical, electrical, electronics and software components together for design and production of products (e.g. vehicles, engineering systems, and others). Typically this involves a variety of separate functions within the organisation and also third parties. For instance, the design process involves complex validation and verification processes depending on multiple tools, applications, product data systems, and third party vendors to interact, collaborate and enable the design and engineering of new products or enhance existing products. To resolve the problems, such companies and their suppliers resort to ‘clash analysis’ requiring larger amounts of computing power.

According to the 451 Group, the automotive, aerospace and general manufacturing areas have been one of the early adopters of grid computing. While there is a need for greater computational power there is also the desire to reduce costs. A few manufacturers have made great progress in grid adoption, moving from small clusters to hundreds or thousands of devices. Early adopters include companies such as Airbus, which recently announced it had chosen Fujitsu Systems’ SyntinWay grid middleware for its high performance computing centres in France, Germany, Spain and the UK.

Digital media
Grid computing technology has become a critical resource for the digital media industry in delivering more sophisticated animation and special effects, although industry analysts see cost and security as barriers to wide take up. Within digital media, the widest deployment

25 http://www.hoise.com/primeur/06/articles/monthly/AE-PR-08-06-47.html
of grid computing technologies, usually supporting animation and special effects, has occurred in the movie and electronic gaming industries. Movie production and electronic gaming companies have long used grid-like technologies in the form of ‘render farms’ for image rendering. Traditionally this was usually performed in batch mode on a rendering farm consisting of a cluster of computers with as many as 500-1,000 cluster nodes.\textsuperscript{27} A grid approach is increasingly being used to harness computing power cost effectively. Grid-enabled rendering also brings the advantage that previews can be run in real-time on rendering department workstations, allowing early feedback on intermediate results.

Several tensions exist within the sector. On the one hand, the creative side means that there are demands for the latest and best technology to produce more and more exciting effects, for instance, the desire for heightened realism from special effects in both live action and animated films. On the other hand, the need to control costs in a competitive industry means that there is pressure on vendors to provide free or discounted hardware and software in return for the benefits of being associated with a particular movie or game. Companies are particularly wary of the outsourcing/utility model mainly because of fears that data will be compromised or that it will find its way into pirate distributions. These tensions have resulted in hesitancy across the sector to adopt grid computing and vendors will particularly need to address the issues of cost and security to promote widespread implementation.

2.5. Market segmentation for products and services

If we look to the current commercial market and into the future, including the far future as envisaged by the European study groups, we can see three major segments appearing, as summarized in Figure 1.7.

\textbf{Figure 1.7. Three main grid markets set the key business models}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure_1.7.png}
\caption{Three main grid markets set the key business models}
\end{figure}

The segmentation changes from wholly owned grids by one enterprise to multi-enterprise grids and outsourced grid infrastructure. The next step, already appearing in financial services (see Box 1.1), is to grid applications sold as service, in which the grid is transparent.

Box 1.1. On-demand computing: an example from the financial services sector

CD02 sells the financial industry usage of a specialised wholly owned application on credit default swaps (or structured credit product) that requires high power computing to calculate probabilities of default for a large aggregate of companies. The application thus uses grid computing from a single site on which clients can run the application, with data feeds coming from market information providers such as Reuters and also input their own data.

The financial application is delivered as a remotely hosted service based on hourly usage for end users. The facilities used are leased from Sun Grid which gives assured QoS. CD02 pay Sun $1 per CPU hour for a service which includes access to servers with network access, disk space, load balancing, switches, firewall, etc. all hosted offsite from the CD02 premises in a single Sun-managed London site. Sun only charges for servers in use, which varies with numbers of clients. The major customer base is in the small client segment, of those who just wish to use a service.

This is an ASP model, based on leveraging Sun facilities to assure very low capex outlay, and ease of entry to start up once the application and grid middleware have been developed. Users are spread worldwide across Europe, and the USA to Australia.

In summary, the business model for this grid computing solution is to deliver software as a service (SaaS) for a multi-tenanted grid environment with a scalable offering as it is based on combining flexible grid middleware with third party information over web services (e.g. market data).

The system runs as far as possible on OSS and open standards, with a common commercial application for the user interface, MS Excel, over a secure internet deployment. It almost resembles a job-submission model of batch processing.

The main processing structure is centralised. It consists of a Java based set of utilities for grid processing, including Jini (Sun's protocol for devices to identify each other using TCP/IP protocol. The grid middleware itself is bespoke Java. Users access via a secure web access method from their local PC. The CD02 architecture (CD02, 2006) is shown below:
3. The grid market today and its growth over the next five years

3.1. Introduction

Assessing the market potential for any new product or service is difficult and especially so for a new technology like grid computing. Market forecasts for grid produced by industry analysts are being published more frequently and clearly it is in the interests of those involved in the industry – including the analysts – to promote the idea that the market for grid computing is going to be large and is growing rapidly. The methodology employed for these market forecasts is usually unclear and all too frequently simply boils down to wishful thinking. For these reasons we urge caution in relying on these kinds of published market forecasts. In our market assessment we of course take these estimates into account but in addition we attempt an empirical assessment based on assessing the addressable market for the industrial sectors most likely to adopt grid computing.

3.2. The current state of the market

Despite the hype surrounding grid computing, it is apparent that it is beginning to be adopted by business. Insight Research describes the year 2004 as a kind of ‘year zero’ for the adoption of grid by enterprises, representing the first year that enterprises’ use of grid technology moved beyond early adoption to more widespread pilots and production launches. In its fourth Grid Index report charting the global adoption of grid computing, Oracle/Quocirca found that grid computing is a maturing technology with as many as 70% of the organisations they surveyed now deploying grid computing ‘in some areas’.

3.3. Existing market forecasts for grid software and services

First, we briefly examine some of the published market forecasts that are generally available. In its 2005 report, Insight Research forecast that worldwide spending on grid was expected to grow from $714.9 million in 2005 to approximately $19.2 billion in 2010 (see Figure 1.8).

Figure 1.8. Worldwide grid spending by type of organisation, 2005-2010 ($billion)

In its latest 2006 report, Insight Research revised its assessment upwards – worldwide grid spending is now expected to grow from $1.8 billion in 2006 to approximately $24.5 billion in 2011, as shown in Table 1.2. Insight also asserts that, by 2011, grid spending will account for 1.8% of total worldwide IT spending.

Table 1.2. Worldwide grid spending, 2006-2011 (Sbillion)

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>CAGR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spending</td>
<td>$1.84</td>
<td>$3.89</td>
<td>$8.51</td>
<td>$12.21</td>
<td>$19.28</td>
<td>$24.52</td>
<td>67.9%</td>
</tr>
</tbody>
</table>


An alternative estimate by the Economic Strategy Institute in 2004 forecast global spending on grid software would grow to $7.5 billion by 2007.28 ESI expects the financial services sector to be the key sector driving take-up of grid computing over the next few years.

These different estimates show considerable variation over a three-year period and between different analysts, the tendency being to overestimate the speed at which grid computing will be adopted commercially and, possibly, the size of the market. Independent industry experts are generally sceptical of the market forecasts produced by industry analysts. From a geographical perspective, it is difficult to draw significant conclusions, given the sparsity of available data. According to the Fourth Grid Index report:29

- The USA continues to lead the other regions in grid computing
- Asia Pacific continues to progress fastest, with adoption rising significantly
- Progress across Europe on the journey to grid computing is evening out as Southern European countries make faster progress.

"All regions show positive movement in the awareness, understanding and adoption of Grid Computing," said Charles Phillips, president, Oracle. "Organizations around the globe are taking advantage of this market-changing technology – adoption is increasing and Grid Computing is exhibiting the classic patterns of a technology that is approaching maturity and preparing itself for entry into the mainstream."

From a sectoral point of view, Insight Research examined 14 vertical industries for grid spending in its 2006 report. Currently, the vast majority of grid spending is in manufacturing and financial applications. By 2011, financial, insurance, and real estate, as well as durable and non-durable manufacturing will account for the highest expenditures, followed by professional and business services.30 Similarly IDC expects the sectors with the most significant spending on grid to be (in descending order) manufacturing, government, education, financial services, healthcare, telecommunications, IT, retail/wholesale, and professional services.

28 Cohen, ibid.
3.4. Estimating market growth in selected industrial sectors over the next five years

As we have already illustrated, industry analysts believe that the markets for grid computing will grow significantly over the next five years. However, it is also apparent that no-one knows with any certainty the extent of market growth and in which sectors that growth is likely to occur. This is apparent from the differences in forecasts between different industry analysts and also in the variation from year to year. The basis on which these estimates are arrived at is also not transparent. Thus there is considerable scepticism about these forecasts.

Here we attempt to provide a different perspective on market development for Europe in four vertical industrial sectors over the next five years, based on realistic estimations of the addressable market for IT within the four industrial sectors (Pharmaceuticals, Financial services, Digital media, and Automobiles). Our estimates are based on several premises:

1. The market in Western Europe in 2006 for information and communications technologies was valued at €644 billion. The European market for IT (datacommunications and network equipment, computer hardware, office equipment, software products and IT services) is estimated at €340 billion.
2. Spending on IT as a proportion of revenue will average about 3.3% across all sectors, with some variations between each of the four sectors being modeled here. For instance, the proportion of revenue spent on IT is as high as 18% in financial services in the USA.
3. Overall spending on IT grew at a rate of 3.1% in 2006 and is expected to grow at 2.9% in 2007. For ease of calculation we have assumed IT spending will continue to grow in Europe at an annual growth rate of 2.9% over the next 5 years.

Thus Table 1.3 shows revenues in Europe for four industrial sectors, the proportion of revenue spent on IT and the amount of spending on IT. Financial services is by far the most significant in terms of IT spending of the four sectors, both because of its relative size and also because of the proportion it spends on IT.

### Table 1.3. European spending on IT for four industrial sectors (2006)

<table>
<thead>
<tr>
<th>Sector</th>
<th>Revenue € billion</th>
<th>% revenue spent on IT</th>
<th>Spending on IT € billion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>10240</td>
<td>3.3</td>
<td>340</td>
</tr>
<tr>
<td>Financial services</td>
<td>600</td>
<td>15</td>
<td>90</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>200</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Automobile</td>
<td>300</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Digital media</td>
<td>40</td>
<td>10</td>
<td>4</td>
</tr>
</tbody>
</table>

Notes:
- SCF Associates, estimates based on studies by Paul Strassmann (www.strassmann.com)
- Europe’s GDP, 2005 IMF figures.
- Based on 2.5% return on total assets of $30.25 trillion (sources: http://en.wikipedia.org/wiki/Bank ; http://www.ecb.int/pub/pdf/other/eubksectorstabilityen.pdf)

33 Based on estimates from the European Information Technology Observatory, www.eito.org, October 2006.
If we assume that IT spending will continue to grow over the next five years at 2.9%, this will result in spending in the four sectors growing as shown in Table 1.4. This does not take into account any relative growth in each sector, eg we would expect the digital media to grow in relative size over the next five years.

### Table 1.4. Growth in European IT spending for four sectors, 2006-2011

<table>
<thead>
<tr>
<th>Sector</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>340</td>
<td>350</td>
<td>360</td>
<td>370</td>
<td>381</td>
<td>392</td>
</tr>
<tr>
<td>Financial services</td>
<td>90</td>
<td>93</td>
<td>95</td>
<td>98</td>
<td>101</td>
<td>104</td>
</tr>
<tr>
<td>Pharma</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Auto</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Digital media</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Notes:
- Assumes constant growth in IT spending of 2.9% across all sectors between 2006 and 2011.

Now, we estimate total European spending in 2006 to be about €1 billion, representing approximately 3% of total spending on IT. It is not possible to know with any accuracy how this proportion will grow over the next few years. However, It is our view that it is plausible to expect that the proportion of IT spending on grid would increase by 50% a year over the next five years. That would lead to total grid spending in Europe by 2011 of about €9 billion per year (see Table 1.5 and Figure 1.9). Of the four sectors we have considered, financial services would represent by far the largest at almost €2.4 billion, pharmaceuticals at €270 million, automobiles at €400 million, and digital media at €110 million.

### Table 1.5. Forecast for grid spending in Europe for four sectors, 2006-2011

<table>
<thead>
<tr>
<th>Sector</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>1.02</td>
<td>1.57</td>
<td>2.45</td>
<td>3.70</td>
<td>5.72</td>
<td>9.02</td>
</tr>
<tr>
<td>Financial services</td>
<td>0.27</td>
<td>0.42</td>
<td>0.65</td>
<td>0.98</td>
<td>1.51</td>
<td>2.39</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.11</td>
<td>0.17</td>
<td>0.27</td>
</tr>
<tr>
<td>Automobiles</td>
<td>0.05</td>
<td>0.07</td>
<td>0.11</td>
<td>0.16</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>Digital media</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.07</td>
<td>0.11</td>
</tr>
</tbody>
</table>
These estimates are based on imperfect information and form only a minor part of this study. Nevertheless they are based on logical thinking and represent a plausible view of how the grid market might develop over the next five years.
4. Grid market take-off

4.1. The various routes to market – the scenarios for take-off

We now examine just how the market might take off by considering several alternative scenarios of what might create a grid market from the point of view of the various players and market drivers. Five scenarios of potential take-off routes are illustrated in Figure 1.10. Note that these scenarios should not be thought of as predictions or accurate forecasts of the future – rather, they are possible stories that help us to think about how the market might develop.

Figure 1.10. How the grid market might take-off can be seen through various scenarios

These various scenarios may not be mutually exclusive – it is quite possible that several could happen in parallel. Almost all of them depend on similar initial conditions:

- Economic conditions that favour grid computing; this does not necessarily imply strong economic growth, although this is one possible condition, as in fact poor economic conditions could seed a grid markets as efforts to cut costs increase, especially capex; this favours alternative lower cost modes of mundane computing but not necessarily high power or data-intensive computing
- Security problems are solved either initially, or eventually – during the early growth
- Reliability is assured – high standards of availability meet and exceed SLAs
- Market education – users understand what it is, where to use it, how to set it up
- Licensing and legal problems are resolved and are no longer inhibitors
- Real demands in applications – not necessarily a killer app – but something like it
Looking in more detail at the various scenarios of take-off of grid computing we note that the actual paths differ considerably in line with the activities to promote the growth phase of each scenario (Table 1.6)

Table 1.6. Characteristics of the scenarios

<table>
<thead>
<tr>
<th>SCENARIO</th>
<th>Description</th>
<th>Promotional conditions which favour each scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Large) vendor push</td>
<td>Major vendors take up grid computing as a mainstream product, or product feature (ie as being ‘gridified’) or service, or as a way to economise on servers in the data centre. The vendors then market grid using large educational campaigns, and practical support eg ‘gridification’ certification centres for applications from ISVs and VARs for compatibility with large demonstration centres and reference sites.</td>
<td>Demand in large corporates to cut costs Early references prove convincing SME market opens up in second phase</td>
</tr>
<tr>
<td>2 User pull</td>
<td>Users demand grid middleware for: Complex applications that require it – compute power, distributed data, collaboration Economy of servers across the applications portfolio in one co.</td>
<td>High growth economy Strong competitive pressures on costs Security/reliability solved</td>
</tr>
<tr>
<td>3 Utility computing- Push &amp; pull</td>
<td>Ideas of on-demand high power computing at low cost are promoted by largest vendors on cost and robustness basis. Demand comes from users of all sizes who accept computing can be delivered like a Web service – the grid is just a delivery mechanism</td>
<td>Demand for transparent grid at low cost Knowledge society of sophisticated users Security/reliability solved</td>
</tr>
<tr>
<td>4 Open grid ecology like the WWW</td>
<td>R&amp;D community starts off the take-up with Internet-like grid, given away free, that sets an example but really for academia It grows into commercial environment when a second upper middleware layer, equivalent to the WWW, is produced that adds ease of use, reliable simplicity and ultra-secure computing. That spawns applications, services etc and a whole market Eventually this becomes the WWG, a ubiquitous grid layer on top of the Internet communications layer and WWW document / transaction exchange layer</td>
<td>Open standards work Attractive giveaway OSS support community Reciprocal marketing</td>
</tr>
<tr>
<td>5 Spin-off companies + user pull</td>
<td>One or several R&amp;D projects seeds a start-up to market the technology (cf Univa model). User demand pulls the market take-up.</td>
<td>Clear advantages in new technology Good channel partners</td>
</tr>
</tbody>
</table>
With several scenarios in parallel, we may see some coming to fruition later than others – an early scenario might be user pull with vendor push for business users while the open grid ecology for consumers and business markets might start to occur a decade later. Consequently the final form of the market might have two phases following on from today’s e-Science research grids phase – an earlier phase of vertical industry grids – and a later phase with two components:

- server economics for business users large and small
- utility grids in a WWG led by an open grid ecology coming from the R&D community to build an OSS community, later producing a market driven environment as the WWW has done over the last 15 years.

Altogether the three phases are shown in Figure 1.11.

**Figure 1.11. Potential evolution of the grid market**

These later two phases will spawn a set of business models for products and services which we examine in Chapter 4.

### 4.2. Analysis of optimal routes for faster time-to-market

Analysing successful alternatives for shorter and faster routes to market for grid products and services indicates a number of interventions, which include, but are not limited to:

- Firstly, assuring an early large customer for the technology, with a real application and significant levels of funding. This may imply government applications for the
first major uses, but in business process applications, not e-Science is a potential avenue. Contracts can then be let to ISVs, VARs and system integrators to build the first versions. Any grid infrastructure provider would also have a first customer. Following the route of the commercial grid middleware ISVs, of locating a real problem to solve and strong business analysis first, should help to avoid ‘white elephants’ if a full business case is first developed with strong RoI targets. The real litmus test will be having significant business critical applications demonstrated, to prove a genuine market demand for grid technology.

- Secondly following the support measures examined above. This includes assuring that the code is suitably clean from contaminating commercial, GPL or software patent code – the latter applies for the USA. A permanent certification centre would be required, which might be part of a permanent ‘Grid Technology Agency’, as outlined above, perhaps in a VO form. To guard against ex-post patenting, or commercial/OSS code capture in licences, some form of central holding of IPR in a public entity is required, perhaps a unit of the same grid agency again. Support intervention for early customers for the new grid technology could take the form of technical support rather than direct funding with a common grid starter kit, based on de facto OSS for its standards, as proposed above. Ensuring that R&D projects provide initial tools and documentation needed for further development, enough to be useful to commercial players who have grid expertise is a further requirement.

**Figure 1.12. The fastest route to market**

![Diagram](image)

In consequence, one promising route to market may be the two-stage process shown in Figure 1.12, driven by the catalysts that ease commercial take-up of the publicly developed code and the support given by first contracts. The first stage is to develop the elements as a
series of applications and appropriate middleware utilities. The second stage can then use these as a basis to build a grid services market, as well as direct sales of products to end users. Today it is often the case that the end users do not choose technology but the system integrators and the service providers do – so its is important to enter the market via them.

We also note that the literature on the success and failure in commercialising R&D programmes for industrial take-up shows that interventionist support can only be successful if a real need is satisfied by the pilot application, with strong RoI targets. If not, the support programme is unlikely to produce anything meaningful, as applying the technology in commercial situations has not been convincingly demonstrated, so no demand is likely to appear and no ISV or VAR will pursue the market as there are no customers.

4.3. Preparing to go to market – the marketing messages

Marketing grid is difficult, for all players. The actual messages for customers are not always that well articulated, and may be quite unclear, as explaining the concept requires an appreciation of customer levels of understanding. To illustrate this point, below is an actual example from a major systems vendor which tries hard but in the end carries little of import:

With grid computing, an organization can transform its distributed and difficult-to-manage systems into a large virtual infrastructure that can be set loose on problems and processes too complex for a single infrastructure to handle efficiently. The problems to be solved often involve data integration and network bandwidth considerations. The systems linked in a grid might be in the same room, or distributed around the world. They might be running different operating systems on many hardware platforms. They might even be owned by different organizations. Regardless of the depth of a grid's resources, all the grid user experiences is the processing resources of a very large virtual computer or information repository. Let's take a quick look at the evolution from an "as-is" IT infrastructure to a grid-enabled environment.

<table>
<thead>
<tr>
<th>As-is environment</th>
<th>Grid-enabled IT environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated infrastructure</td>
<td>Virtual and flexible infrastructure</td>
</tr>
<tr>
<td>Rigid and predefined capacity and performance</td>
<td>Responsive and resilient capacity and performance</td>
</tr>
<tr>
<td>Fixed cost and value structure</td>
<td>Variable cost and value structure complementing known fixed costs</td>
</tr>
</tbody>
</table>

Grid represents just one step in the process of moving from an as-is IT environment to a flexible standards-based environment as part of a Services-Oriented Architecture (SOA) taking advantage of dynamic services, storage and applications.

Notice the later emphasis here on SOA, for reasons which are not made clear to the reader, but above all the difficulty in moving from abstract concepts to what this means in terms of real business advantages. The “why” of turning to grid is somewhat less than evident. Anyone seeking to explain to real business customers the grid advantages and where, and how it is to be used, faces this problem. What is obvious is that building a business case for an investment in grid computing requires more than just showing some technical merit. The above extract does then go on to try to build this business case for an initial grid project.

Early research into grid marketing for the utility computing segment in the USA has also shown a clear misalignment of messages and markets. Vendors are trying to sell something (called utility computing) – but users do not know what it is. Vendors are convinced of a clear connection between Web Services, utility computing, and BPO (Business Process Outsourcing). Buyers do not understand any connection, so are confused and less sure. Vendors and users disagree about when pay-as-you-go computing may be a common IT business practice, in that some users see it as far away, even “never”. Vendors are misjudging the importance to users of various utility computing motivators – while the vendors believe the most important motivator is reducing capital expenditure; users, in the USA, are concerned with reducing operating costs. Vendors also underestimate the importance to users of certain potential take-up inhibitors, notably security, as well as uncertain ROI and a lack of capability to customise offerings, at least for those available in the US market.

What the project sponsors inside a candidate business require is a set of business-oriented metrics to support the business needs and demonstrate the economic benefits of grid computing. This is a key part of the marketing for grid. Metrics implies some form of benchmarking, measurement, for justification of resources, both in financial and human capital terms.

What we are addressing here in the above examples, although it is not made too clear, is the single company enterprise grid, probably for a large organisation. What is necessary is to address at least three markets with two offerings, ie both products and services, for:

- Large corporates, initially for enterprise and later for cross enterprise VOs
- SMEs, initially for enterprise and later for cross enterprise VOs
- ISVs, VARs, system integrators and grid market service providers of all kinds who may take up the original grid offering to become part of their own in some bundled or modified product and/or service; they may act as our main channel to market

We really need three message expositions for these three markets but we may only get the chance to put forward one on a large scale – but that can then lead to specific promotions. So the way forward is to structure marketing messages for grid as being addressed to large corporate and small business customers; if correctly structured, the messages can also show the attractiveness of a potential business opportunity to the channel partners. The structuring of the message may be:

- Firstly explain the advantages as a business case in terms of operating costs and capital investments
- Secondly give a real example with measured benefits; this should then be followed by a list of industries and application successes
- Thirdly explain a little about grid itself – the “what”
- Finally explain a little about the first steps to go forward – the “how”

In the first and second messages we have to decide if the target is vertical market segments or else more general sales, based on the economies in resources for servers and storage, or possibly both.

---

CHAPTER 2. Technical and non-technical drivers and barriers to grid software and services

1. What is the problem?

1.1. The key problem

The key problem is to identify and analyse the technical and non-technical drivers and barriers to the evolution of grid software and services market over 2006-2012. Our analysis aims to uncover all the factors that enable closing the exploitation gap between grid technology R&D results and industrial products and commercial services, in the context of coexistence of diverse software licensing models and of availability of open source grid middleware. Consequently we must examine:

- Possible constituents and facilitators in general for a market supply of grid middleware and services (kinds of offerings, characteristics of catalysts and growth rate)

- Specifically for grid software as a service, obstacles and stimulants to market take-up, from the perspectives of a pervasive intelligent information utility and/or knowledge economy and/or digital innovation ecosystem.

- Size effects of the various market segments and the business position of significant players on the use and spread of various types of software licensing models;

From this we can establish a route to market to support the development of a coherent strategy for technology transfer under the 6th FP, in particular of IST grid RTD project results from academia to industry. Although the study outcome will be used to this end, the field of investigation is by no means restricted to the situation of the FP6 IST grid programme.

1.2. Understanding success and failure of publicly funded high technology collaborative development programmes achieving market take-up – past experiences – UK, USA, EU, Korea and Japan

Collaboration has become well established in Europe as a way for spreading the costs and uncertainties of research. The use of publicly supported collaborative projects has been the foundation of the European grid initiative. In examining the barriers to commercialisation of grid technology it is instructive to view the history of similar efforts – and so to perhaps learn lessons for grid commercialisation. Comparisons with previous EC and national collaborative programmes may possibly identify factors that highlight the successful paths for taking collaborative research projects to commercialisation.

Note that previous EC collaborative ventures frequently centred on research that has been a considerable distance from the market – and involved academic as well as industrial
collaborators – and so was often termed ‘pre-competitive’. Many of the early European efforts were a response to overseas programmes, especially to the Japanese 5th generation computing project of the early 1980’s, largely focused on artificial intelligence.

However the European routes to successful commercial exploitation are less clear than was the case of the more commercially-oriented Japanese schemes. But we should also add that current EU funding work in FP5 and FP6 has produced significant results, especially in the grid area.

The Alvey project

Looking at examples of previous programmes in high power computing, the UK’s Alvey programme of some £350 million (over €500 million) in the mid-1980s stands out as one whose outputs have been a little disappointing.

In some cases, we can identify the barriers to commercial exploitation. These are principally:36

- Poor linkages between R&D, productisation and production efforts within industrial participants
- Technical difficulties that proved to be insurmountable
- Difficulties associated with being part of a collaborative project, as involvement of multiple partners meant that projects become relatively inflexible with changing commercial circumstances – major changes in a project’s direction involved gaining the agreement of most other partners and the government.
- Project progress was also affected by a divergence between interests – those of the academic participants and the commercially related concerns of the industrial collaborators.
- Withdrawal by partners, following a reappraisal of business and technical priorities spelt problems for the remaining collaborators.

Pre-competitive R&D programmes are well suited to a range of tasks. But they are not sufficient mechanisms in themselves to increase competitive performance in the IT sector. This was the main conclusion reached in the independent evaluation of the Alvey programme (Guy, Georghiou, Ray, et al, 1991). Moreover, complementary private sector and government initiatives were identified as being needed in order to:

- Relate IT development to users
- Promote effort within participating enterprises to create and follow technology strategies to exploit such research commercially – to take it away from the lab into sales
- Re-evaluate the need for patient capital.

The success of the Alvey programme across all of its projects has been questioned, in that not a single major product was successfully launched as a result of its 500 Million Euro spend either during the programme or after. The intellectual outputs that may have shaped future product concepts have also been questioned in that they were already present in many researchers’ thinking but were published in a project form which may have held up their use in real products and services. They were enclosed in the R&D segment of companies, not in

product development and product marketing. Effectively they were consigned to relative oblivion in further academic activities and industry discussions, but not in products.

The experience of Korea and Japan

It is notable that while each of the above problems, to varying degrees, extends to Japan they have been far less of a problem there, and also for Korea with its telecommunications development programmes. While interventionist policies have enabled Korea (and Japan) to build an IT industry in the shadow of the US computer and telecommunications industry, they have not solved all shortcomings associated with exploiting collaborative research.

However, support policies in Japan, as in Korea, have interacted with the national industrial/political environment to create a climate for innovation, driven by tax incentives for research, and infant industry protection with import tariffs, also government purchase of new technologies as first user and positive regulation – for instance Korea has taken aggressive anti-trust actions against Microsoft. Korean Chaebol and Japanese Keiretsu form a local industrial structure and more specifically the broad technological and foundations inside the leading players, as well as turning their IT sectors into a source of relatively patient capital.

This has been helped by structural features in the Korean and Japanese economies. Japan's success in consumer electronics was an example of close matching of development and innovation to user requirements. This was helped by an educated but closed domestic market. Careful stimulation of competition between the national champions to form a number of robust domestic producers has also been instrumental in stimulating competitive innovation linked into aggressive pricing, as has been seen, for instance, in the mobile smartphone market in Korea and Japan.

Looking further back, to the 1970s in Japan, the building of a collaborative culture for R&D and its exploitation took place as the consequence of a gradual learning process, over a series of projects, especially the VLSI initiative in the late 1970s to move from discrete components to very large scale integrated circuits. Strong competition between the Japanese Keiretsu’s technology firms was accompanied by difficult relations between the industrial sector as a whole and the Japanese government, so attempts by MITI in the early-1970s to restructure the industry into a smaller number of national champions were strongly resisted. Any drift towards monopoly was constrained by competitive instincts of the major players. Given the extent of competition between Japan's IT firms, it is perhaps appropriate to ask why initiatives such as the VLSI project were apparently so successful.

In retrospect, it seems that government funding was so high that non-participation would place firms at a commercial disadvantage. This effect was accentuated by the closeness-to-market of the results of the VLSI project. And its central, shared, facility only provided an environment for shared knowledge creation on the most basic research. However the project’s more commercially oriented development work was organised as modules. Thus the IC vendors participating could internalise the benefits of government-sponsored research without compromising competitive positions. Outputs did not remain as projects in R&D, as the internal communication channels between the different departments within Japanese firms assured that the results became marketed products. Thus the motor for commercialisation was a communications path to assure transfer from lab to production front-line to sales as a planned and rapid process, entirely commercialised.

As a result, the project may be viewed as an example of a government subsidy across rival firms, rather than one for sharing the risks and uncertainties of research. These uncertainties were limited as the window of commercial opportunity for the technology had already been opened by the leading US firms, so directions for technical development were obvious. This is one example of a successful commercialisation mechanism – but driven by a mix of national self-awareness and the urge to compete globally, combined with judicious investment support for a range of well-equipped competitors – who were encouraged in the end to compete. The same mechanism has been used in Korea for commercialising new mobile technologies.

The next stage has in some ways been far less successful. The clear target disappeared as Japan started to catch-up with the West and so it became more difficult for MITI to impose strategic direction, especially with a rapid pace of technological development, industry structure and market change.

MITI’s response was to move to collaborative basic research projects – The Fifth Generation Computer Systems Project. This was a major departure from preceding applications-oriented economic research areas. In some respects, it was an attempt to build early R&D into an innovation system which did not have strong links with university research. As its results were not aligned with the agenda of the participating companies in commercial requirements, translation of outputs into competitive advantage has been unclear. Most of it has the status of an academic work, rather than clear products and services.

More recent work in Japan, Korea, China, India, Singapore and Taiwan has indicated that the close collaboration between state and industry has vital links in areas beyond R&D. These include the conceptual side of business, with formulation of business models, physical support in the form of science parks and encouragement policies for attractive product design and ergonomics. A key point is also training for skilled workers and general ICT education of the population and workforce.

Today’s grid projects in Japan seem to be mainly driven by e-Science still. One of its latest grid programmes is the National Research Grid infrastructure (NAREGI) programme for joining up universities in Japan with a 5-year plan (2003-2008) and a detailed set of work packages (covering workflow, job submission, scheduling etc based on Globus, Condor and Unicore with message passing via Grid MPI) $140m METI “Business Grid” focused on enterprise grid working with OGSA technology.

South Korea’s route since the 1960s has been somewhat different in that it is more government driven than Japan’s, and from a position of underdevelopment, following its occupation by Japan until 1945. It is perhaps the most aggressive and successful large-scale example of technology intervention and has major lessons. Across a series of some seven national programmes it has progressed from an agricultural and mining economy to be one of the leading markets globally in the production and consumption of high technology. But local conditions of politics, culture and national psychology must be part of any appraisal. It is notable that the key national players such as KT (Korea Telecoms) are associated with plans for grid – a typically pragmatic approach.

---

38 ESTO Asia ICT Growth project, Validation meeting, July 2006, Brussels
40 Simon Forge and Erik Bohlin, Managed Innovation in Korea in Telecommunications– Moving towards 4G mobile at a national level, to be published, 2007
41 Sang-Hong Lee Broadband Internet in Korea Internet in Korea and KT - KT’s vision for GRID service vision for GRID service, Presentation, May 2005.
European policies
The UK’s Alvey programme was accompanied by expansion in pan-European collaborative IT R&D under the European Community's first series of Framework Programmes (beginning 1984). They covered a broad range of research areas to support pre-competitive research projects (we are now on the seventh FP). In particular, the European Programme for Research in Information Technology (ESPRIT) began in 1984. It covered microelectronics, computing and software engineering, computer integrated manufacturing and advanced IT systems for business and the home. ESPRIT was a technology push initiative. It had no formal mechanism to link projects into the needs of the IT user. The most important benefits derived by industrial participants in ESPRIT were mainly associated with improvements in basic know-how and the adoption of more ambitious research objectives42.

ESPRIT was complemented by the market-driven Eureka programme, launched in 1985 by President Mitterand of France as a response to President Reagan's SDI (Strategic Defense or 'star wars' Initiative). Eureka provided funding for one particularly significant piece of work – the Joint European Submicron Silicon Investigation (JESSI) project – perhaps a European response to the USA's Sematech project.

In summary, the pre-competitive emphasis of ESPRIT, like Alvey, tended to produce research that was a considerable distance from the market. They were thus vulnerable to the types of barriers to exploitation seen in the Alvey project. Successful collaboration between industrial partners has been generally associated with clear research targets, combined with a crystal clear and close potential for exploitation shared by all participants. Concerns have been also expressed that in pan-European research, overheads associated with collaboration tend to be amplified by the geographical separation between partners and language barriers and a division of focus, since for many personnel, the project is only one priority in the working day.

The US experience is highly varied. In practice, federal defence spending has performed a key role in establishing the USA's global lead in computer-related technologies since the 1960's, when, among other projects, DARPA (Defense advanced research projects agency) financed the Internet as an academic project under its prior incarnation as ARPA43. US national security has been used as a justification for "hands on", sector-specific interventionist industrial policies. Today there are few overt interventions in the civilian sphere but a wealth of "hands-off" actions, both non-sector-specific and sector specific. Grid computing has been financed as an academic pursuit for e-Science by the National Science Foundation but other funding is being applied for national security applications.

Very often what we have seen in the USA over the last thirty years is a different type of collaboration, an approach with collective funding investment in a ‘permanent’ (for the USA) institution on one site, with a lifetime of 5 -10 years or more.

During the 1970s, the United States led the world in component and computer technology, their manufacture and many aspects of software engineering. At the federal level, the official US industrial policy was non-interventionist. The policy on industrial innovation was not to have a policy. However in contrast, all individual state governments have a tradition of interventionist policies. A relative decline in US’s industrial competitiveness over the 1980s, with paranoia over Japan, led to calls for a more coordinated approach to industry but was met with government opposition. The Reagan Administration had entered

---

the White House in 1981 with a strong commitment to minimal intervention in the commercialisation of technology, defining federally-funded research in the civilian sector as only supporting basic research. Exploitation of research was considered to be a matter for market forces.

During the course of the 1980s, government and industry collaboration in support of the pursuit of national competitive advantage gained ground and continued into the 1990s. US responses to the Fifth Generation Programme, drove collaboration in the US but had private sector origins, driven by the concern amongst members of the US IT industry and policymakers. In 1984 the Congressional passage of the "National Cooperative Research Act" marked a major response, as, for anti-trust reasons, collaborative formations across enterprises were illegal in the United States until then.

Although the Japanese challenge did not break the anti-federal intervention tradition of US central government, it did provide funding to established agencies. For DARPA, it helped launch a 10-year Strategic Computing programme in 1983, with ambitious targets for practical demonstrations of advanced AI and parallel computing. The National Science Foundation (NSF) also expanded the scale of funding for computing and AI, packet-switched networking and distributed computing. This was also coupled to funding for the SDI, announced in 1983 as a space-based defence system. By 1993, the SDI project had consumed some $32 billion – mainly on ambitious military-oriented projects distant from commercial products and markets and then was apparently abandoned. How much, if any, spin-off from this has been useful commercially is questionable.

From all of the above US R&D initiatives, it is difficult to pinpoint any real successes/advances, or any lessons on commercialisation, except to say that the rules of clear objectives and near-market deliverables did not seem to have been observed. Perhaps that is the lesson – without them, any results are unclear.

The 9/11 terror attacks have spurred a large number of US government overt and more secret security projects. But their real commercial spin-off is yet to materialise, although they have raised the activity level of the security sector in all fields. Usually this is through directly funded delivery of specific systems, such as the WARN project in Washington DC for a high bandwidth security network for all first responder services.

In considering earlier US initiatives outside the US government, one of the first responses to Japan’s 5G Project announcement was the Semiconductor Research Corporation (SRC), created in 1982 as a permanent non-profit-making institution. It was linked to the industry association, the Semiconductor Industry Association, funded by the main IC fabricators. They then gained access to a broad spectrum of semiconductor research projects, conducted at US universities under SRC sponsorship. Note that US industry has a long history of using its universities to support research, in contrast with Japan.

The major non-governmental collaborative response to the 5G computing project was the setting-up in 1982 of the MCC (Microelectronic and Computer Technology Corporation). Established in Austin Texas by the major US computer and semiconductor manufacturers, it was a response by US industry to the view that the Japanese 5G initiative was an attempt to take the world's high-end computer market. The MCC was created in part as a defensive move. Up to MCC’s termination of operations in 2001, the membership diversified to include a broader range of high-profile corporations involved in ITC products, as well as government research and development agencies and leading universities. The MCC was at the forefront of a new approach to collaborative research in the USA, and for the USA, it
was revolutionary. It involved unprecedented cooperation for a fiercely competitive industry in being set-up on a permanent basis to undertake pre-competitive research in semiconductor and computer technology, for producing long-term results only.

During the late 1980s, the MCC was part of a whole era of R&D restructuring leading to increased emphasis on interim deliverables whilst also pursuing longer-term goals. It paved the way for a wide variety of research consortia. By the late-1980s, collaboration had been established as a legitimate vehicle for sharing the costs and uncertainties of pre-competitive R&D inside the USA.44

However, the results in retrospect seem unclear. Perhaps much resulted later in products and services but this is not apparent in terms of successful products. The large MCC projects such as CYC – a common sense database as a 10-year project – have not had evident impacts.

While the MCC paved the way for national US collaborations, the launch of the Sematech Consortium in 1987 represented a sea change in the US administration's attitude to collaborative R&D. Sematech was founded with the largely commercial objective of better competing with Japan in semiconductor device manufacture. It was funded by a total of $500 million over 5 years from the US Department of Defence, via DARPA, as a non-profit making organisation. A similar sum came from subscriptions from 14 leading US semiconductor manufacturers. Its core mission was to increase gate density in the production of memory chips, as technology leadership was shifting to the Japanese suppliers by the mid-1980s. SEMATECH would also act as a technology driver for D-RAM (dynamic random access memory) for promoting state-of-the-art processes, regarded as a key component technology.

Further changes to federal approaches for R&D support followed from the 1988 Omnibus Trade and Competitiveness Act. This allowed a 'Civilian DARPA' in the form of NIST (National Institute of Standards and Technology) It launched an Advanced Technology programme (ATP), in 1990 as a mechanism for providing federal support to US business carrying out pre-competitive R&D in generic technologies. These were defined as concepts, components, processes or scientific knowledge that could be applied to a broad range of products and processes. The emphasis was on supporting technologies such as IT that will play a significant role in enhancing US competitiveness.

In conclusion, there are lessons to be learnt

One central message that follows from the analysis and comparisons is that the use of collaboration as a market modifying mechanism, should take appropriate account of the complex nature of market structures and the ways in which these structures vary between different national environments.45 Japan's collaborative schemes evolved gradually with development shaped by factors unique to Japan, especially the role of MITI, the ministry for international trade and industry, in orchestrating the Keiretsu's (corporate cross-holding groups) technology members.

The lesson is that the local factors of dominant players and industry shape cannot be ignored (eg in Europe, the software industry is dominated by SMEs and large integrators). It should also be noted that the European IT industry differs from Japan and the USA in that it is more

---

45 Tim Ray Government-sponsored collaborative research to promote information technology: Japan's challenge to the west?, March 1993, First Theory-Oriented Research Group, National Institute of Science and Technology, Policy Science and Technology Agency, at: http://www.nistep.go.jp/achiev/fb/eng/mat032e/text/mat032e.txt
fragmented, being subject to the constraints of a large number of relatively small national markets. Differing national priorities, coupled with cultural and language barriers, add to the difficulties of implementing a concerted programme of action.

At the same time, a high technology restructuring and consolidation in the EU is continuing and accelerating for both supplier and user industries. Adjustment issues associated with successive waves of mergers and takeovers mean that a planned temporary monopoly that could be associated with a collaborative project is sometimes overtaken by an actual monopoly in the wake of increased industrial concentration.

The US experience of setting up centres of excellence long-term and permanently has certainly had some effects at a conceptual thinking level. Whether it has been the nurturing environment for specific commercial products (or services) is highly debatable.

In summary we can give the major causes of failure (and successes where failure causes were countered effectively):

- Failure to relate industrial development paths to real user needs and to the perception and agendas of real users
- Lack of closeness to productisation (a lesson from Korean and Japanese successes) producing research results that are a considerable distance from the market
- Lack of a clear route for the productisation (a lesson from European and recent Asian experiences – where success has been based on countering this affect)
- Poor linkages between R&D, productisation and production efforts within the participants – there is a need for active promotion efforts within the participating enterprises to create and follow technology strategies to exploit such research commercially – ie to take R&D out of the lab and into product sales (a lesson from the Japanese VLSI project success)
- Collaborative-project diseases – lack of progress as communication is poor, due to geographic and linguistic fragmentation; inflexibility in changing commercial circumstances due to multiple partners; progress affected by a divergence between interests – academic versus commercial; withdrawal by partners, following reappraisal of business / technical priorities; insurmountable technical difficulties, unforeseen at the outset and requiring better initial analysis.

In general, it should also be added that centres of excellence have weak linkages with industry (unless funded by a specific commercial entity, eg Rolls-Royce, Boeing, etc) A key advantage of having a single industrial partner is that IPR agreements can be set in concrete. With multiple partners things get very complicated, very quickly. IPR leakage is a major concern. The main advantage of centres of excellence is that they foster experience that can be tapped into through long-term association with a company. Development of trust is also another major factor. Often, trying to persuade a company to adopt the grid in its relatively immature state is not easy.

Such programmes were thus vulnerable to the types of barriers to take-up and exploitation seen in the Alvey project.

Effectively, successful collaboration between industrial partners has been generally associated with clear research targets which are near to productisation, combined with a crystal clear route to market exploitation shared by all participants, while meeting a genuine market need.
2. Barriers

2.1. Technical Barriers

There are a range of technical barriers to grid systems as shown in Figure 2.1 which inhibit a complete offering.

Figure 2.1. Major technical barriers to grid uptake

Barriers – technical

- Security
- Firewalls and grid working
- Common standards for grid build & inter-working
- Variations in WS from vendors - so no real common standards for Grid inter-working and WS-RF becomes limited
- Heterogeneity of resources and the complexity/fragility of the grid environment – adapting to distributed resources and their virtualisation
- Performance, specifically communications performance (latency)
- Application compatibility for grid environment (parallelisation/ virtual hosting) for a loosely coupled parallel processing architecture

Here we only examine a few of the key ones briefly but they must be resolved if grids are to become a part of mainstream computing:

- Security – protection of both resources and the local environment and of the running (‘visiting’) applications and data
- Standards – open standards for grid that can be used globally, including Web Services (WS) that tend to be naïve in character, in that they do include state.
- Applications suitability – Parallelism of applications – how parallel are the candidate programs – and performance requirements, specifically communications performance (latency)
- Heterogeneity of resources and the complexity/fragility of the grid environment – adapting to distributed resources and virtualisation – key for business environments

**Security**

The technical problems of grid security are reviewed briefly here while the problems over users’ perception of insecurity of grids are examined under the non-technical barriers.
Firewalls present a particular problem for grid middleware since they inherently resist ad hoc or on-demand access from unknown IP addresses, be they internal or external. A firewall’s main role is to block any access that may result in malicious intrusion. Combined with intrusion detection systems (IDS) and intrusion prevention systems (IPS), firewalls try to bring secure access to the corporate network. As the grid middleware’s main role is to marshal resources on an as-needed basis, from anywhere across its list of resources, the corresponding source IP addresses may well be not programmed into the firewall so it will accept traffic from them. Even if the firewall allows traffic in a grid configuration, attaching a grid may only multiply the number of false positive alarms from IDS as servers are suddenly diverted to new tasks, which can conflict with internal systems management policies, if these are not also made ‘grid-aware’. The problem that many users see is the usurping of the more powerful administrative functions under the alias of a grid operation for malicious purpose. Thus security issues for the user focus on identity management and authentication in an environment where applications and data (and therefore users) can be shifted from one set of processing and data resources to a different set.

In response, many grid users severely limit grid access and interactions. For example, one company in the aerospace industry has a secure “production LAN” of Sun workstations and PCs that is a core corporate resource. Their internal grid system can only connect for presentation of its results – it cannot access the production LAN for its resources and must not provide an insecure access point – the only links are really for demonstration purposes. In consequence, the grid does offer some workstation-based services for this LAN for certain authorised users, but with no reverse access to the LAN from the grid. In effect, grid computing is seen as useful by the company but it must respect corporate security. The ways in which this has been implemented are to limit its interaction with the mainstream computing functions.

2.1.2 Standards
Open standards for grid that can be used globally, have been discussed widely in our research as being a key driver for grid take-up. It has been cited by users, R&D projects and vendors of various types. Some have noted that industrial acceptance and take-off is proving more difficult than expected two years ago, with the lack of open common standards being a key reason.

It was noted that just going with the oldest and most established grid middleware initiative for an open middleware, the Globus toolkit, now in release GT-4, was not enough, even based as it is on the open grid services architecture (OGSA). Some viewed the Globus project as rather a US-centred project under the National Science Foundation (NSF) and thus excellent for large scientific problems, but possibly less obvious as a basis for a commercial platform. Moreover the Globus project has merged software from many sources so different coding standards and types and also the variable documentation has yet to be assembled as a professional output from the project. The toolkit itself was found by some to be difficult to get going, while support was not readily available at an industrial level. The focus for grid software is turning towards standards setters, particularly W3C, OASIS and the IETF who have had success in setting open standards which work at a professional business level. But any standard should be global, not just European.

A further major obstacle cited has been the focus on just the grid middleware – users need a complete ‘everything’, not just middleware. A complete grid stack would include an active directory, authorisation, provisioning, Linux/windows scheduler, and the grid-compliant operating systems. It was also noted that no R&D project is looking at the message-passing
layer. A standard reference implementation of the open grid services architecture such as OGSA/GRIA would need improved message passing. We also look at other key areas for standard middleware to cover in the drivers section.

While including Web Services (WS) was often considered as the way forward for open standards, and that they underline the current vogue for a Service-Oriented architecture (SOA) among systems vendors, certain caveats were also expressed in that:

- WS tend to be naïve in character, in that they do include state, necessary for grid operations, and are also fairly low level in operations
- WS tend to lack security, and do not give high performance, while reliability is low
- They have some fundamental difficulties for on-demand systems, being static entities, which cannot (yet) be dynamically created and destroyed, and there is as yet, no standard way of presenting the properties of the web service.
- There are some variations in the interpretation, control functions and data definitions in WS implementations, so that the openness for interoperability is much in question still, because those WS from certain very large vendors are effectively proprietary, bringing no real common standards for grid interworking. Thus the multiple WS aggregation into an interworking environment to try to overcome the above problems – the web services resource framework (WS-RF) – becomes limited or meaningless, as there are several flavours of WSRF, so may even expect a repetition of the ‘Unix wars’ of a decade or so ago.

The market for grid developed in an era of standards that are still arriving, and not yet stable. Early grid standards are being upgraded, and we have multiple bodies defining slightly different versions within the market. A possible future is that the grid market stagnates, if users cannot depend on a fully standardised, interoperable environment. We may return to the previous confused situation of divergent directions, whereby de facto standards were modified by de jure groups from standards committees, with different agendas, giving multiple standards and projects and systems that were standalone, and so went nowhere.

In summary today, Globus is one leader for grid standards and this is a problem. Globus is not a set architecture – it is moving with OASIS and WSRF, but as noted, there are several flavours of WSRF from the larger systems and software vendors. Looking at the commercial industry side in the EU, we see progress from the e-Science position of Globus towards industrial engagement in FP5 to FP 6 initiatives. However, when we look at standards, there are still some unclear areas – the GGF is more for discussion than for standards setting. To some extent the EU has perhaps missed an opportunity window here – either to mandate Globus and progress it, or to develop a suitable alternative. Various contenders are appearing and fixing on one is a difficult problem, which we discuss further under market drivers.

2.1.3 Performance, specifically communications performance (latency) and parallelism of applications with adaptation to a grid environment

Grids being a loosely coupled set of resources with inter-process calls and application loading across a network mean that there are performance issues with grids, specifically transport and communications performance (latency) – and especially with files of several GBytes. Overall impacts may depend on dispersion of resources for data storage and compute power. Hence the granularity of task exchange can make for an inefficient response
in a compute service, comparable to the database phenomenon of thrashing\footnote{Thrashing in databases occurs when the systems spends all its time in looking for data (on disk) and retrieving it and little or no time in processing it, so that performance is heavily degraded. It can be overcome by optimising the size of ‘chunk of data that is fetched in each seek operation, which is why task granularity is important.} but on a much larger and complex scale. The right level of data caching for the virtual organisation (VO) configuration of multiple enterprises sharing resources is also important.

The role of network costs and performance (especially in commercial settings where the price to be paid for network usage has nothing to do with the technological cost of creating a given level of network performance) must also be considered among the techno-commercial barriers.

There are also the applications themselves to consider. Many applications programs can only be interpreted as a single process thread with serial activities – so the parallel working offered by the grid holds no advantages apparently. Moreover for those applications that do have a parallel task capability, it may be necessary to convert them such that they are exploitable by the grid environment. As noted the granularity of tasks distributed across a grid is also important, especially in dynamic creation of tasks depending on previous activities.

2.1.4 Heterogeneity of resources and the complexity/fragility of the grid environment

One key advantage offered by a successful grid is its heterogeneity – it can deal with different operating systems and just as difficult, many versions and flavours of those operating systems, as well as diverse networking standards and databases types, including different file structures. Also the grid must host a range of applications with their process calls which preferably are on one node but may have to be across a network, in remote procedure calls (RPCs).

This is a tall order.

Everything must be prepared for a range of different environments. But the scope for adaptation to change in the resources in the current grid architectures is low – they are not ‘autonomic’ in the sense of being self-healing, robust and adapting to change. Any change has to be programmed into the grid middleware adapters/interfaces and policies. Thus change has far more effect – updating an existing compute component also has far more impact for instance. Thus, much more time/effort is required for validating any change introduced in a grid configuration, with a new version of the grid software or for an extra utility or application. In general, the grid environment is far more fragile in some ways than standard single processor environments, so that change can be much more disruptive.

This also applies keeping to required performance specifications – managing resource allocation with contractual SLAs demands that any changes to the configuration do not affect performance adversely. This is difficult across multiple resources, with perhaps variable availability, and together with mixtures of licences that may limit the scale of what may be run.

2.2. Non-technical barriers

The range of non-technical barriers for grid systems is fairly wide and the major factors from our research, with their significance, for market take-off are shown in Figure 2.2.
Figure 2.2. Major non-technical barriers by impact for perception, marketing and legal issues

We only examine some of the key ones here briefly, but again, they must be resolved if grids are to become a part of mainstream computing:

- Market education and the cultural barrier of grid with its complexity and the IT department as host organisation
- Perceptions of security/privacy/confidentiality/data integrity – perceptions and commercial sensitivities
- Lack of commercial preparation – no concentration on a complete grid offering, (not just middleware) with suitable productisation
- Licensing and conflicts with existing operating system, database, application and utility licences and the interests of the major players holding licensed IPR, against a context of OSS licences with reciprocal clauses and possible proprietary infections
- Financing and RoI – demonstrating the payback for users with measurements and identification of benefits
- Legal requirements in collaborative projects in the EU – effects of the EC contractual R&D framework for IPR regime under FP6 – the corresponding business requirements and the options to be verified
- Keeping to required performance specifications – managing resource allocation with contractual SLAs
• Political dimensions of EU policy – the opposition to interventionist policies and the European context

**Market education and the cultural barrier of grid**
In general, lack of knowledge about grid and its potential and where it applies is its biggest hurdle take-up. One study’s findings are summarised in Figure 2.3 on the barriers to grid:

**Figure 2.3. Education and understanding are major barriers to grid take-up**

Not only is education required, it must change mindsets, as to quote one supplier of grid middleware interviewed:

> Managers are too conservative, and see grid computing as a risk – they do not want to be an early adopter – but feel they can catch up if it becomes competitive edge.

Hence promotional education is needed at four levels:

• Internal business managers, the major users
• The IT department, who must implement a grid system
• Outsourcers, wherever IT resources have been outsourced to a third party
• Industry sector level, if common models for interaction and integration are required

Figure 2.3 shows that lack of understanding of grid technology and where it could be applied among business deciders and IT departments is at over 80%.
However the picture on education of the market is unclear. Some studies looking back to 2003\textsuperscript{47} have found that perceptions about grid systems being a potential IT platform are progressing, with a perhaps surprising speed of change. In 2005, the awareness of grid was found to have grown since 2003 and a basis for some readiness was being prepared, although the level was largely at that of a highly constrained, discrete cluster grid as a pilot project. The study’s authors started carrying out grid research in early 2003, believing that mainstream grid computing was still around 5 years away. After two years they claim that approaching 20\% of companies were utilising or experimenting with grid, but it is difficult to understand if this is optimistic. Whether this may be seen to point to limited mainstream adoption within the next few years is open to question, as perhaps it is only being taken up in those with pressing problems.

One of the several problems with understanding is the cultural barrier of grid is language – it has its own language, not shared by the IT department and other users. Pressure has to be put on the grid community to use standard language and present its concepts more clearly to both business and technical IT audiences. There is still much misunderstanding and confusion even in the IT providers about grid. A further linked cultural barrier of grid is its complexity, which means its benefits and implementation are difficult to understand, and most of all, this leads to difficulties in preparing costs/benefits analysis, both on paper and in the minds of business decision makers.

Furthermore, the IT department appears in some organisations now to be a major block to grid progress. They insist on no external access, so grids cannot operate. Effectively we have a failure to communicate as the business users cannot express their needs and the safeguards available in current grid solutions, while the IT department usually does not understand grids – either in technical or business terms. Education of the IT department is thus necessary. From our interviews, we understand that one large European car manufacturer recently noted the need for this to combat restrictive policies by a sometimes paranoid IT management. Thus there is conflict with the IT professionals in many large corporates. The end-users have to push the IT department to move forwards far more.

We found that use of outsourcers, which is fairly common in larger organisations, can pose a problem with grid technology as they are so slow to change and adopt new technology. One large aerospace supplier we spoke to has found this to be a major barrier. Generally outsourcers can be bureaucratic and slow to understand and provide support. Their world is really more of the legacy systems, with tight rules around SLAs. They tend to lack any capability of understanding innovation. In fact our research found that it is unclear if an outsider could run a grid today.

Wherever collaborative work across enterprise has to occur be it with collaborative groups or in supply chains, a whole sector’s education may be needed – for instance the aerospace sector has agreed on a policies model for business co-ordination for grid in a VO. But the auto sector has no agreed model. Generally, IT strategies / policies on grid usage are needed before large investments can be made.

There is a further education problem. The conventional group dealing with grid technology today is often the R&D department. For mainstream business computing, they may have the aura of either doing long-term research – which may bear fruit in a decade- or never – but that may lack credibility for turning out ready-to use technology for today. They have low credibility and may even be seen as devoted to self-propagating serial projects without positive bottom line impacts. This has been found to be the case in previous R&D programmes. Thus when introducing education about grids, the effort must go far beyond

\textsuperscript{47} Quocirca reports on grids, 2003 and 2005.
the R&D function in the corporation, with a need to get both IT to take up grid pilots as well as to influence thinking in the business deciders, strictly on an RoI basis. This is confirmed from looking at the business models for commercial grid software vendors, built up from their experience (examined further in chapter on business models).

**Perceptions of security/privacy/confidentiality/data integrity**

Once some level of understanding of grid is reached, perhaps the key barrier then for early adoption among large corporate users is the perception of grid as a challenge to commercial sensitivities. Corporates require protection – of users, applications and their data on the grid and the contextual environment including any attached applications, data and machines. They fear malicious invasion and leakage of key data.

Globally, from the users side security is perceived to be a major issue. One study found security to potentially be a complete barrier to take-up for some 12% of corporate users across USA, Europe and Asia while over 40% see it as significant challenge – but one which will probably be worked around.

The major reservations around the perceived security of grids are especially strong for grids shared across several enterprises, in any form of virtual organisation. Major users do not want to put data on a resource that they do not effectively master – this is especially true in certain user industries, such as car industry, for clash testing. Oil and gas companies may want to use an open grid for calculations but are not willing to put their data on anything but a corporate grid. In fact many users and suppliers we interviewed concluded that shared grids in certain industries and particular corporate cultures were unlikely to ever be accepted due to security fears. And it could only happen if grids must have proven secure gateways for controlled entry – the major concern for large corporate users.

There are also commercial sensitivities when it comes to using grid utility computing from a third party. Many companies potentially using high power computing and data intensive grids wants to keep both data and computing models private. Hence such users will not go to a bureau service to run their grid computing models but will keep all in house. It is worth more to them to keep the results and operations secret than to loose any of that knowledge. Data sets, especially for specific conditions such as failure are also too valuable to host externally. All global data repositories are likely to be wholly owned in a large multinational.

Some large corporates are extremely restrictive on going outside the company at all. One large car manufacturer has an engineering portal for engineering services inside the company only. There is a restrictive method for external access, with file transfer. But this is a long way from web services (WS) and certificates for user access. Another allows no Internet access at all – only Intranet.

The main issue is confidentiality of information, and also of the processing approach, between commercial users of the same grid. For some, security of the data is not so important as the know-how in using the data, which may be revealed in the processing algorithms, configuration, etc. Sharing among different users with different agendas, even in the same company – military and civil aviation development for instance – must be quite separate, although they would like to share the same grid resources. Such users require that any grid infrastructure have internal walls, not just external security protection.

---

48 From 1,466 interviews with IT influencers and decision makers, completed Sept 2005, 302 conducted in North America, 603 in Europe, 561 in Asia, Quocirca Insight Report, November 2005, Grid Computing Update-The Market Begins to Mature
For these reasons, it seems unlikely that industry will take to worldwide open grids soon, due to confidentiality and also for reasons of reliability and guaranteed delivery. Large corporate users will never share their servers unless there is a major sea-change in their thinking on grid security and any financial benefits of sharing.

Moreover there are further problems of data protection for the data used on grid projects in line with the various national and EC acts. However the laws vary by country so international grids have to face extra legal complexities in multiple compliances. There is also a particular concern about grid projects using health records of real, live patients when not held on a single secure computer. This demands an immediate verification of security before grid system usage and avoidance of legal delays in set up.

Barriers to take up due to security fears/uncertainties vary by user segment. One applications service vendor in the financial services industry noted that sales to smaller users who are less technical may thus be in fact easier than to the major banks and financial institutions. For the financial market and technology, liability for security is also an issue for the large customer, but less so for the small. Moreover, large groups in financial services expect to set the rules on security.

All users are also most sensitive to access to data. For instance, in one example we studied, the system is designed such that the grid infrastructure provider, used for the outsourced infrastructure, has no access to data hosted.

**Licensing and its conflicts**

Security is followed by licence problems for many potential users. The key problems can be summarised as:

- The grid offers a whole new dimension of usage environment which may conflict with current licence conditions. Licences for the running applications, if commercial proprietary, are a major problem with a grid as commercial licences are about restraints on sites/processors/user numbers/time/transactions etc, while grid is about open usage. Proprietary software licences tend to rule against using applications in new contexts with existing commercial licence restrictions. If an application is run across many servers then a licence may be needed for each CPU type and each server size, yet the grid model is dynamic and resources are only dedicated at run time and may be dynamically added on demand. The licence holder will be the end-user in each case – so the application may have to be restricted by the configuration of the target cluster. The positions of the main software vendors on gridifying their applications are important here, as they decide in their licences on the flexibility afforded for grid configurations.

- OSS licences (analysed previously) can be summarised as having difficulties for users and developers in the areas of:
  - One real issue is licence compatibilities – for operational interworking, which may require that the grid middleware interworks with all the other software across many sites and multiple organisations – what can be used with what leads to conflicts in usage with other software, be it OSS or proprietary.
  - Development restrictions in terms of returning modifications to the repository – a difficulty for commercial software developers and VARs, and
using the code in commercial products, which the vendor wishes to sell under a commercial licence.

- Infections by proprietary software with patent or licence infringements. In the USA, this can lead to injunctions to halt all operations in the company.

- General legal complications of understanding the restrictions and conflicts of the various OSS licences, mixed with commercial licences means that there is often a need to employ lawyers. They are required to assure that any technical/business projects can be performed legally; this acts as a major barrier. A key point in the EU is the liability in case of death clause which must be included. Such a clause is not present in US licences.

- OSS may be used as a source of free IPR by the large players who hoover up OSS – it can be played as a game by them. This is a problem for the smaller companies who own IPR but want to use OSS licences.

- Also there is the whole question of the mass of new licences appearing for OSS – why do we need them, especially as they have to be proved in the market and need legal approval? And a further problem with some OSS licences is that although certain licences have been tested in court in various countries and held to be valid, some have not, so their legal validity is unproven.

In summary, OSS licensing is sometime complex for industry to deal with – but simpler for academia.

**Financing and RoI – demonstrating the payback for users**

One of the hurdles that the European grid industry must jump is proving its worth. The business case for grid may be complex and is not evident, as it is more than just perhaps a new form of high power computing at lower cost, especially in a shared resource situation. For instance, decisions on release of servers when not used and problems of who pays if they are not being used may be a factor. Moreover, there are sometimes strong time limits and constraints on sharing which impact on access for usage. For instance, a patient ready for brain surgery needs grid time “now” for calculations on probes and their progress in real time through the patients’ tissues – the only value is in immediate usage. Real benefits have to be demonstrated and a cost/benefits analysis put forward – but measuring the benefits may not be obvious. This business case is a key part of the sale in the largest grid middleware vendors and integrators (reviewed next chapter).

**Lack of commercial preparation**

In general the R&D projects so far have produced results but without all the necessary provisions for a future commercialisation, although this has been built in more to the later FP6 and now FP7 calls for tender. However there is no concentration on a complete grid offering (not just middleware) as users need a complete everything, not just Globus or other middleware – a complete stack. This situation is natural in that projects set out with specific research objectives. What is lacking is something pull them all together to provide a commercial offering, with modules such as an active directory, authorisation, provisioning, Linux/windows scheduler, and the operating systems and more advanced message passing and a standard reference implementation for improved message passing.
These requirements include support from the projects for their outputs for those who would build real systems, and in post-build, in effect for pre-sales and post-sales (the latter is what in a commercial world, is sometimes called fourth line support, for code changes). For instance, for GRIA, IT Innovation have no internal resources for any levels of the equivalent of pre- and post-sales support.

A possible approach is to use the model of an OSS community that does the support. However this takes years to build up and community projects also suffer from feature creep, so a better but similarly flexible solution is required. Effectively a fully commercial grid model may work better for giving a support service. As there is a market for OSS now it may be that an OSS/proprietary combination – as used in the Red Hat and Univa model for Globus.

**Legal requirements for collaborative projects in the EU**

Legal requirements in using grid products from EC FP projects seem to be centred on the licence conditions – but unfortunately there may be more than this. The IPR emanating from the EC Framework Programme funded projects comes under EC jurisdiction and various rules must be complied with for future usage by members of the R&D consortium.

The business requirements and the options to be verified mean that contractors have certain obligations under the EC contract. These cover the protection, use and access rights that must be respected under the EC contractual R&D framework for IPR regimes in the FP6 initiative. They must also be followed in order to acquire and use the intellectual property generated by the project. The rules cover knowledge and pre-existing know-how, its ownership joint ownership, with transfer of ownership. Protection of knowledge including publication and dissemination is covered as are the access rights which include possible objections by the commission, patent searches and obligations continuing beyond the end of the EC contract. Contractors must also reach agreement with their personnel over the ownership of results, and these must be compatible with obligations of the EC model contract.

This provides a basis for understanding but it may have to be carefully reviewed if the optimal conditions for industrialisation are to be created so that openness and availability are optimised for future commercialisation.

**Limits and political dimensions of EU policy for grid commercialisation**

As previously noted, the level of intervention in the EU to prepare for full commercialisation can never be at that of a command and control economy. The EU also has a structure of co-operating member states so that common policy is difficult to implement. The industrial structure is quite different to those economies where industrial planning set centrally has achieved major advances in the introduction of technology as an economic driver. Perhaps the only common driver is the need for high technology exports and infrastructures. We examine this further in the final chapter on policy.
2.3. Measuring the barriers to grid take-up – market responses on grid technology

Although the benefits of grid systems are perhaps becoming better understood by industry today, greater comprehension is matched by a growth in the perceived barriers to grid usage. For many, the decision is simply the balance of perceived benefits outweighing the perceived difficulties. *Perception is thus key.* Studies on the barriers have ranked grid maturity as the main hurdle to adoption, closely followed by lack of firm standards (a sign of the technology’s immaturity), and also that grid is not high enough up on the list of priorities for the business (see Figure 2.4). How to build a suitable business case for any grid initiative is also unclear, necessary to justify perceived costs of a grid solution.

Figure 2.4. Inhibitors to investment in grid computing technologies

If we look at two distinct areas – cost and security – we can see how the respondents see these within the context of grid computing. The study’s questions on cost found that around 14% of respondents regarded budgeting, funding and cost ownership of a grid system as a potential barrier procurement, with a further 37% seeing it as a significant challenge. In the overall findings, North America and European companies regarded cost as less of an issue than those in Asia.

Various studies have looked at conditions of licensing and its importance for potential corporate grid users and has found to be a key area for cost concern so existing licence models may have to adapt to the needs of the grid market. Looking at how corporates view
licensing costs for an on demand usage grid environment, respondents were found to be split: many preferred the fixed price licence model, as costs are predictable, while over 40% preferred flexible per-usage pricing, with a similar division reported across the three regions of USA, Europe and Asia.

3. Market drivers for grid middleware and services

3.1. Technical facilitators and drivers

Here we examine the key technical drivers for grid commercialisation, perhaps a subset of all possible drivers. These are summarised in Figure 2.5.

Figure 2.5. Major technical facilitators for grid technology take-up

Here we examine:

- Security and the perception of security
- Basic technology for commercial implementations – common standards with a reference model and a common grid toolkit
- Technical features for commercialisation – licence discovery, the contractual layer for SLA performance and enabling major economies for large data centres by sharing
- Support units to build an EU market
- Promotional projects as drivers
Security and the perception of security

Typically security is a perceived barrier which may be removed with technical investigation and demonstration. Solutions to the security issues are in producing new models of going outside the company – of B2B access – and of tailoring the current security mechanisms for a grid infrastructure, to allow for firewalls for instance. We discuss some possible approaches here, in limited fashion, but a full discussion is beyond our remit. Leading proprietary grid software vendors may be able to provide extensive security auditing and proof points, and our general view is that this be extended to the outputs from R&D projects.

In security terms, grid systems are comparable to a chain, in that they are a string of processes which govern the interlinked grid resources. Naturally such a chain is only as robust in security terms as the weakest link – once that is breached, a malicious users or piece of malware may be able to penetrate many sites.

The first link in a grid security chain may be considered to be authentication, either through a public key infrastructure (PKI) with public and private keys, or with external certificates to sign messages, coming from a certification authority (CA). However such systems all have vulnerabilities in terms of knowing who is the recipient of the certificates. Early academic projects – Globus in the USA, the UK e-science programme and DATAGRID in Europe had such schemes based on checking with project leaders before authorising the release of certificates – such a procedure in an open commercial world is not viable or possible. What is needed is a certification authority plus a mechanism for verifying certificate recipients.

Furthermore we have the problem of firewalls with grid systems – and their incorporation into everyday operations. One approach has been to assign the dynamically attributable communications ports to grid access, using the system administrator’s special intervention. However this destroys the raison d’être for a firewall. A better solution is to establish separate grid domains at each participating site. This still requires some form of isolation structure for the high security grid resources.

The latter may be constructed using the imperfect, but simple, “demilitarized zone” (DMZ) concept of an outer firewall, which only accepts accesses from known addresses, and an inner firewall which does the same for internal local sites accesses. One model is to place the ‘front-end’ grid resources in a DMZ, with the most sensitive grid resources in an isolated domain behind the second firewall, which uses authorised access, but with complete isolation and separation of all other resources of the organisation from the grid resources, perhaps with no links at all.

Grids also have variable amounts of freedom given to their users – they can be wide open down to account level for each resource, something that is useful in e-Science projects but which is anathema to services models and those who wish to loan resources but never allow their control by an outside organisation. What are therefore needed are rings of security, which give certain permissions and contain usage to with strict limits. To satisfy the concerns of large corporate users, grid configurations must include some form of secure gateways for controlled entry if the grid system is going outside company and the security rings approach may be acceptable. When the main issue is the processing approach as well as confidentiality of different users’ information, between commercial users of the same grid, some form of ‘VPN’-style approach to apportioning the grid resources may also be the solution, although it limits the flexibility of resource allocation.

---

50 For a discussion of the problems, see - A rough guide to Grid Security, Mike Surridge, IT Innovation Centre, Sept 2002, Southampton, UK
Moreover some form of compliance with the Data Protection Acts in the EU is required for the data used on grid projects in line with the various national and EC acts. However the laws vary by country so international grids have to face extra legal complexities in multiple compliances. This indicates the need for some form of compliance monitor with policies for each country.

For the particular concerns over grid projects using health records of patients when not held on a single secure computer, Digital certificates for authentication of participants and authorisation of assets could help here, to provide immediate verification of security before grid system usage and avoidance of legal delays in set up.51

However security also presents problems of additional time delays for verification. Certain applications for grid usage may have strong time constraints on immediate access for usage – for instance, in the example of the patient ready for brain surgery, the needs are immediate for grid access, as a grid is part of surgery equipment. Again, digital certificates for authentication of participants and authorisation of assets could help for immediate verification of security and avoidance of legal delays in set up. This implies some form of CA and we examine the idea of a public or private grid CA below, as part of the commercialisation effort.

**Common standards with a reference model and a common grid toolkit**

A major driver for grid technology in Europe will be the arrival of common standards, perhaps assembled in a useful form as a reference model, with a reference configuration to be implemented in OSS.

This implies that the EU mandates one grid middleware model only – whether this is Globus or not is really not the question – but rather that the EU should focus funding on one middleware environment. Currently funding is dispersed – and the question has to be asked whether the EU has planted too many seeds if it wants a European grid industry. The reference model would cover the areas shown in Figure 2.6.

---

51 The Artemis project is examining security and privacy of web services with peer to peer discovery for semantic web in current projects (see artemis/res/grid).
In contrast to the USA, European choices are open – it can use any contender be it Unicore, GRIA, or Globus, or any other grid middleware. It may also uses components from several project architectures and assemble them for an optimal system eg taking the batch scheduler from Unicore. In the USA, the NSF, NASA, etc originally identified grid technology with the Globus Alliance, so that now for the US government, grid systems are endorsed as Globus – the USA has chosen its model.

There is also the question of whether the reference model would only cover the ‘lower’ middleware sections but leave the upper middleware to be interpreted in commercial products. This would give scope for industry while providing common standards to interface to the underlying resource infrastructure. However, there are advantages for a common platform in the commercial world in defining the application interfaces, with a standard for the applications requirements definition and a standard application interface definition or API for the grid stack.

Any reference model should use pragmatic industrial standards from the W3C, OASIS, IETF and OGF to embrace WS and improve WS-RF and WS to remove their simplistic limits – to make them stateful and persistent for instance. The common reference model could then be the basis for commercial designs, which may differentiate in high performance, etc.

**A common grid toolkit**
A common OSS toolkit of grid software based on the reference model considered above would be a major driver in that it may provide a guide, have the credibility of experience
and give a simple point of departure, as well as being a single focus for further technology development. The kit should provide a platform on which others may build a market with:

- Grid applications
- Support services
- Extensions

What technology it is based on, in the sense of the choice of an originating middleware platform, is not our task. However do see that we should indicate the form of preparation of the toolkit for a commercial world. Previous problems of common grid toolkits should be avoided, in that they have merged software from many sources, so different coding standards and types and also variable documentation contents pertain. Instead the whole needs to be put together as a professional project with suitably coherent code, documentation and support tools for deployment.

A further previous obstacle to commercialisation has been a focus on just middleware – users need a complete grid stack, not just middleware. A complete stack would include an active directory, authorisation, provisioning, Linux/windows scheduler, and the operating systems as well as improved message passing (no R&D project is yet looking at the message-passing layer).

The leverage effect of such a toolkit should also be considered, as it can form the catalyst for a grid applications industry segment, as any ISV or VAR can build to a standard platform in the knowledge that it will be widely used and in the security of standard APIs and predictable performance with scale.

The reference implementation of the toolkit should undergo certification of the code IPR as being clean of contamination from code that is proprietary, reciprocal OSS or under software patents (for the USA). To permit proprietary components to be added at will, a reference OSS stack would have to have careful licensing conditions. It must also be tested for interoperability on heterogeneous hardware and also against various applications for their integration and interoperation with a standard interoperability test bed.

The kit should also have a demonstrator programme to showcase its features in live situations and bring credibility in terms of scaling, performance, ease of set up on different environments and security, as summarised in the diagram. All these requirements and how to satisfy them are considered in Figure 2.7.
A reference OSS stack could also work with provisioning put into the hardware (the processor itself) with an embedded firmware co-processor for instance in the common microprocessor cores. This requires a standard for interfaces and calls such that any chip can work with standard OSS middleware. For instance, this could be implemented in the Intel processor family using their V-PRO model of partitioned extensions for administration. Any common toolkit should also be able to bridge all application servers – by using a single reference model that covers all.

Both the reference model and the standard grid toolkit would act as educational and promotional tools, to be used in training developers and operational support staff as well as being foci for a market awareness campaign.

**Technical features of grid technology for commercialisation**

Here we examine three key features, or high level services, for grid middleware essential to commercialisation – licence discovery, the contractual layer for SLA performance and enabling major economies in investments and support costs for large data centres by resource sharing. Little work is being done elsewhere globally on these high level services for grid yet they are essential for commercial usage. They also provide an opportunity to engage with business side of ISVs/VARs on their definition in open standards.

**Licence discovery**

One service that may be required for dynamic allocation of application to resources in the context of restraining licences is the concept of a licence discovery service. This requires that applications and utilities declare their licences in a common format, as XML either as a script or a document that can be machine read (such as an XML DTD) with a common set...
of semantics and a metadata model suited to legal conditions. This would make all code and
data ‘licence-aware’ for the licence discover services which would then adjust resource
allocations management according to the licence responses discovered.

**Contractual layer for SLA (service level agreement) performance**

Grid middleware needs to have SLAs in policy and reporting as a capability because
business works via SLAs. What is required is an SLA-aware grid-enabled resource
management system, with SLA negotiation, multi-site SLA-aware scheduling, and one that
takes into account the security measures and interfaces for storage, checkpointing, and
networking support. This SLA contractual layer would be multi-platform and available as
part of an OSS grid middleware package. SLA management features could be added to an
existing middleware contender, or instance Nextgrid.

This also applies to keeping up to the contracted performance specifications with change.
Managing resource allocation with contractual SLAs demands that any changes to the
configuration do not affect performance adversely. This is difficult across multiple
resources, with variable availability and the mixtures of licences that may limit the scale of
what may be run.

A possible solution is use of a commercial, or OSS-based, CMDB (configuration
management database) to track the configuration, its changes, assets in use and assets
available and non-available as an adjunct to the resources registry. Commercialisation of
grid systems in general will require some form of a CMDB to track assets and operational
status of all elements across the grid, or its equivalent in a resources registry.

**Grid architecture enables major economies for large data centres**

Among the drivers not often cited (yet) for grid technology are its architectural features
which allow server sharing. Instead of having dedicated resources for each application, as is
commonly the case today, we can use a common pool of resources and apportion and
schedule applications to run on them as required. It has most usefulness in the areas of large
application packages for ERP, CRM and the larger database-centred applications, as
illustrated in Figure 2.8.

Naturally this requires rapid roll-in and rapid roll-out of each application. The economies in
server farms, database storage etc may not be the division by the number of applications, for
reasons of immediate availability, speed of loading and starting a new application, unless it
is kept in some form of ‘virtual hot stand-by’ state.

It will have major implications for application for enterprise computing. The use of grids
has similar features to the current push for virtualisation, but on a much large scale. Grid
allows implementation of concepts of Service Oriented Architectures to be carried much
further with adaptive configurations on an as-demanded basis for larger applications such as
SAP. Oracle’s 10g relational database (RDBMS) in its full grid implementation shows a
new way to handle commercial data-centric tasks.

The use of grid technology in this fashion can be applied to large enterprises or to multiple
medium and small enterprises (SMEs). It may be most useful for enterprises in the range of
10,000 to 20,000 employees, if sharing on resources becomes attractive. It also has knock-
on effects in cutting energy requirements for power and cooling, data centre real estate and
systems licences and maintenance contracts for processor resources and support such as
UPS. It may also offer failover enhancements, if sufficient capacity for this is planned for.
Figure 2.8. Sharing applications reduces server and storage resources

A major driver for grid may be its ability to re-use resources in sharing mode across one application portfolio, reducing spend and providing extra options for failover capability in emergencies.

Part of the application portfolio in a telecommunications operator

To enable this approach to shifting applications across servers for enterprise computing, the middleware must have two extra layers as shown in Figure 2.9.

Figure 2.9. Specific management layers must be present in grid middleware

For servers and database sharing across the applications portfolio, a 2-layer sharing mechanism must be added to grid middleware to allow it to swap out applications on demand – for new capacity or for failover.
Support units to build an EU market

For supporting a host of requirements including IPR, ownership, code cleanliness, etc, something is necessary to give long-term support beyond the life of R&D projects. There needs to be a good business/operational model for major grid software projects – the ‘persistence’ as in some of the OSS organisations to develop code, with commercial organisations to offer support services, which in the grid case would be the role of the European software industry. In a service based, fragmented EU market, one should prefer a model based on the growth of a business ecosystem where as many actors as possible provide differentiated professional services, with no single one retaining full control over the open source code base. Following this model, for a grid commercialisation effort to take place, certain supporting entities should thus be in place to act as secure foundations for the several problems, which may be centred on five key units, giving the required degree of permanence:

- An entity for central EU ownership of IPR and of the OSS code
- Interoperability and Test platform
- OSS certification centre
- Data certification centre
- Certification Authority (CA)

The whole might be drawn together as a single European Grid Technology Agency to give the EU market commercial support for the long term. Such an agency would effectively have a catalysing market role in the ‘persistence’ of its presence. The organisation could also act as a centre for promotion and market education, with management of the industrial grid programmes examined below.

It would promote a standard EU grid reference model and its implementation as a toolkit, through an open source community for both the toolkit and reference model. The open source community would be orchestrated by the agency and should incorporate all developers, be they commercial, academic or from the FP projects themselves.

Above all it would assure that the basics from projects for commercial transfer were in place, in terms of suitable licensing being assured, EC contractual terms for the FP projects and their IPR restrictions and that practical tools and documentation were available for development and deployment in real grid situations. Let us examine each unit in turn:

EU ownership of IPR and of the OSS code

One of the problems of using publicly available code is that nobody may own or be responsible for the IPR, and the code, apart from the original projects, which are temporary. So some form of holding entity long term is required if the project’s participants do not want to, or cannot take over the IPR and code and its diffusion. In the OSS world a support community is encouraged but this can take years and may not be stable. Instead of having many such holding entities set up every time, it may be more logical and effectively cheaper to establish one entity that does everything in a single standard way, in line with best practice internationally. To guard against IPR hi-jacking by ex-post patenting, or commercial/OSS code capture in restrictive proprietary licences, this holding entity should also provide vigilant protection – and may call on legal resources.

The actual code provenance and freedom from contamination by proprietary or any licence - restricted code would be carried out by another dedicated unit, the OSS Certification Unit,
described below. The overall IPR management structure avoids ensuring that those who contribute code and IPR have to take liability for the code and also avoids having to protect the distributors of code, as these may be temporary entities, only existing for the life of the project. In the Globus project this is precisely why the Argonne National Lab has adopted the Apache process to carry out these checks and apportion liability.

The entity would provide a central point of ownership of IPR and of the code for the EU projects (it could be, that if successful and seen to be fair, it becomes the holding point for code and IPR from outside the EU). It would issue a standard BSD-style licence, such as APL-2, to assure commercial take-up, and specifically one suitable under European law.

Under EU law, ownership of IPR and the code also brings liability for it in case of death, a point noted by the major ISVs (such as SAP) and system and services providers (such as IBM) as being necessary. This entity could fulfill this role, with appropriate insurance for EU liability if required. Thus it should have a small, dedicated full-time staff.

To handle and diffuse the code, the unit would also have a secure code repository with appropriate directories and search facilities and hence act as the European download centre. It would have its own web portal for information, promotion and education. Note that whether it is a single physical unit in one Member State, or some form of virtual organisation spread across several existing institutions such as universities is an open question.

The entity would deliberately avoid doing any code maintenance and support after the original development as this would be left as a service for commercial companies (the Red Hat model) to build the EU software industry. One role could be to find support partners, and to build a support ecology for this, and so assure users have the support they require from commercial support services.

Finance could be from central framework programmes and the unit could charge some low amount to commercial entities (of the order of several hundred Euros to pay for download operations) but its code and electronic documents would be free to individuals and educational establishments.

**Interoperability and test platform**

To assure interoperability of applications, middleware enhancements and new releases, a further infrastructure support element for commercialisation would be an interoperability platform (see Figure 2.10). It would test grid middleware and utilities code against current and legacy infrastructure resources underneath and new and legacy applications above it.

**Figure 2.10. The interoperability and test platform**

![Interoperability and Test platform diagram](image)
**OSS certification centre**

OSS code needs some form of ‘cleanliness’ assurance, to be freely taken up by users (following the SCO case) and commercially by the European software industry, although this problem is not just restricted to OSS software. This verification would cover (OSS) code coming from R&D projects as being free from commercial code, with proprietary restrictions, or constraints in the OSS IPR. It may perhaps be useful to give a certification of non-infection. Code should be separated in libraries so that any contaminating code, if identified, can be easily isolated.

To perform this verification, what is needed is an entity that will take responsibility for the code, as clean. It would act as a clearing house for IPR, assuring that input from the projects is not contaminated and then using the code ownership entity described above as the holder of the IPR long-term. Code distribution is the responsibility of the European code repository, described above. Effectively this entity also has to check that the contributors have made a real effort to check the code provenance as clean. Whether the unit for verification of code cleanliness is a single team on one site, or a VO distributed across Europe is an open question.

**Data certification centre**

Similarly for code, a clearing house for data IPR will also be necessary for the very large amounts of data that could be offered from public programmes of R&D for grid working. Again provenance is the key problem. Whether this is established as a single unit or a European VO requires further consideration.

**Certification Authority (CA)**

Security of grid working could benefit from a system of digital certificates for grid users, participating systems and applications for authentication purposes. As outlined above, there is still the problem of prior authentication of recipients of the certificates and also of their false publishing. In a security environment in which some solution is found to these problems, an EU grid certification authority could provide standard certificates and procedures for their use, both for resources and human users. The CA could be a public service or a selected group of approved private commercial services (like Verisign).

**Promotional projects as drivers**

A policy for technical diffusion of the grid toolkit and reference model in industrial environments may benefit from projects which are closer to industrial usage than current projects. Two forms of technical programmes may be useful:

*Shop-window projects* – which are simple working solutions to a key problem in each industry sector, which would include demonstrating the economies of servers and other resources in the average enterprise’s application portfolio and the ability to swap rapidly, provide failover, etc. The aim is to give rapid results in a short timeframe (six months).

*Large Demonstrator Projects* – a few projects by industrial sector which would benefit from application of the standard European grid technology for a large business-critical application, such as shared working in the aerospace industry for collaborative design or a major shared utility environment for the SMEs in one region of Europe. The aim is to show large-scale demonstrations of the technology, for a highly targeted application, with a usage phase of at least 12 months and overall project life of some 30 months. This timeframe implies rapid rollout of the technology before the working phase. One of the first large demonstrators might well involve a government application, as typically has been used in
Korean commercialisation initiatives. This has certainly been raised in the OECD deliberations on the future of grid technology with its 2005 consensus report recommending:

“Whenever appropriate, agencies (of government) should actively encourage and promote the use of Grids,…”", and “Agencies should promote the culture of sharing resources, data and results”, and “solutions in which all agencies worldwide contribute to the Grid-linked ensemble of computers, software and data, at the same time as they take advantage of the contributions of others”.

3.2. Non-technical facilitators and drivers

The major non-technical factors are summarised in Figure 2.11.

Figure 2.11. Non-technical drivers of the grid market

We now examine each of these factors:

More complex applications enter society for everyday use – but computing costs are contained

Trends to usage of larger amounts of computing power and massive volumes of data in certain sectors may favour the take-off of grid computing as part of mainstream computing.

---

For instance in the vertical user sectors of manufacturing design, grid systems can take over the role of high power computing at lower cost and are already having an effect. One of the largest semiconductor chip suppliers has come to market over the last 12 months with a range of lower cost microprocessors because with grid processing it could simulate and test the new integrated circuit layout ten times faster than its main competitor- and at a cost it could afford – which gave it a crucial competitive advantage.

Also the grid’s ability to combine heterogeneous resources enables users to differentiate their business operations and processes. Their ‘internal’ IT infrastructure can now be configured in new and specific ways, meaning IT and processes are far more closely linked, but in ways that can be more easily reconfigurable with a grid infrastructure, and are scaleable at low cost.

A grid infrastructure can also be used to give increasing returns with scale as more complex applications are brought into the business. Effectively, the complexity of the application may go up linearly but the cost of computing if using a set of on-demand ad hoc resources will not keep pace. This enables very high power computing at low cost. As somebody said (was it Schumpeter, Shaw or Marx?) that a big enough quantitative change is a qualitative change – what we can do with computing at low cost changes enormously and so we can expect far more refined computing to be used in health, supply chain logistics, business analytics, energy distribution, seismic search for hydrocarbons, etc. For example, as health systems become computerised, we may see high power computing entering the operating theatre, for bodily simulation, probe tracking etc, demanding reliable rapid grid applications for vital situations.

**Impacts of the major commercial players**

A key factor in the commercialisation process is to realise that the process will not and cannot act in splendid isolation. All the major ICT players are looking for new markets and are interested in grid, whether they are global, the European arms of the US and Japanese players, or European application ISVs such as SAP or system integrators like ATOS Origin, Finsiel or CAP Sogeti, etc. However this means that they have quite different agendas depending on position. Some are pushing data-centric grid solutions only (eg Oracle) while IBM wants to use software ownership to sell its services arm. Many of the players have large services arms for systems integration and long-term support and wish to expand these in ways that differentiate against the low cost services and outsourcing efforts coming from India and soon China while maintaining their higher charges and margins. Moreover these lower cost players will also wish to act in any EU grid commercialisation initiatives, adding to competition to participate.

The key point is that none of them wants to be left out. If they see a serious commercialisation support effort, most would be very anxious to assure they are part of any collaboration. A serious approach to a European institution for grid technology of some form with real support tools for the code and intellectual property, developed in the R&D projects, would be highly attractive, in that the main concerns over code quality and ownership are resolved. The promise that this could grow to be a global grid platform, in the same way as GSM became a global technology, would be an added driver for participation.

In case a common grid platform does take off, most of the global and major European players would be prepared to act as alliance partners for new start-ups, which is fairly low

53 CF IBM’s acquisitions in 2006 of MRO in assets management and Micromuse in carrier network management by IBM Tivoli, all to build its services business offering, while HP’s acquisition of Mercury Interactive in testing follows a similar strategy.
risk for them. They might act as incubators, and investors for new grid start-up companies – their strategy being to keep the new technology initiative separate until they can see if the proposition works commercially, then to buy it out and take the technology, marketing and sales teams in house (– as IBM has done in several cases, including in grid technology, with Meiosis for instance).

For some major players, strong technology with obvious market advantages is more likely to be brought in-house immediately, if it really works, to develop an own brand version as the lower risk means a start-up scenario is not necessary. The technology and code just goes straight into the product development and commercial marketing departments. Such take-up and commercial usage should be encouraged, with strong promotional activities to the major players, and continued in hand with them for post-development promotions in the market. These activities would act as a key nurturing phase for the technology, especially if the GSM model of sharing the IPR with central holding is followed to give confidence and encouragement.

**Impacts of the international standards bodies**

Open standards development should be a key catalyst. Else it risks being a potential barrier to grid technology in commercial exploitation. But the evolution of grids from e-Science requires a view which is oriented to commercial exploitation, requiring clear representation of that vision to the key standards bodies worldwide. It is vital that any European vision for this evolution of grid technology is accompanied by a clear presentation of these aims.

The objective would be to work with standards bodies in key areas such as persistent stateful WS and a WS-RF and also higher level services – but only where appropriate for interworking, in that some higher level services should be left to the commercial market to develop. Standards development would depend on those most productive in industrial ICT standards at a world level – W3C, IETF, OASIS, etc.

One point of counsel – there is a need to guard against *ex-post* patenting of standards’ IPR, and to ensure that IPR brought into the standards is free from restrictions of commercial exploitation, be it in proprietary original works, for which royalties will be claimed, or reciprocal OSS technology, in which some under some circumstance licences might have restrictions on for-profit usage if combined into single products.

**Low cost computing and the utility model in the short term**

The proposition that grid is a form of pervasive service utility is discussed further in the following section, from a long-term view. Here we look at this vision as to whether it could act as a driver in the short term, under certain circumstances, with specific product/service packaging. Utility computing services as a grid commercialisation driver may be viable if the promise of low cost computing, at a low level of processing can provide utility computing for the masses– competing with the PC. The ubiquitous grid system could at the same time offer a far higher scale of computing and storage service for business computing. However this concept has been slow in coming and has not taken off as expected, the current effort being a rebirth of ideas that go back some time.\(^{54}\) Let us examine one current commercial offering for such a service in more detail for its users and the level of demand for the service (see Box 3.1).

\(^{54}\) Ideas of utility computing were mooted in the era of Multics, the time sharing predecessor to Unix, and perhaps before.
Box 2.1. An early example of a utility computing service, its usage and demand

AT&T’s hosting and application management services launched a managed utility computing service in late 2005. Based on hardware from Sun Microsystems as an extension to its AT&T’s hosting service, the utility computing service was promoted for business use, especially those companies with seasonal peak demands. AT&T’s aim is to provide automatic and on-demand access to computing resources to scale up and scale down efficiently. For its existing hosting service customers AT&T now offers server virtualisation to its existing hosting clients. Instead of being required to run a single operating system per server, users can run multiple applications on the same physical server, each with its own operating system image. With server virtualisation, the aim is that a customer can better manage applications, giving one application a higher priority over another. AT&T promotes the utility computing and server virtualisation services to customers as having major advantages for rapid provisioning and avoiding the cost of investing in server hardware, so they avoid spending a lot of capital on hardware, while their systems often go underutilised much of the year. For utility computing, customers pay a base fee each month for access to a dedicated server at roughly half the cost of leasing one, plus a variable utilisation charge.

An existing AT&T hosting customer, Turbine, which publishes online dungeons and dragons games, requires the utility computing service to handle enormous surges in demand, with accompanying computing and network usage when it releases a new online multiplayer game. What is needed is extra capacity for the first weeks after a new game release and then return to lower demand levels. Previously, with an earlier game release, the company grossly underestimated resources required and ended up trying to "throw hardware" at the problem. The role of a utility computing service would be to provide planned rapid scaling up in capacity and then down again after the peak. AT&T also noted that new customers did not immediately contact them and demand the service following its announcement earlier in the year. Instead some existing customers had shown some interest.

What is relevant here is that the service is sold as computing on demand and savings on server spend, with higher return on investments. The customer example given is a small, highly specialised user of hosting, but significant in that it represents the inability to spend large amounts of capex on computing services, even though they are the basis of its offering. The question is whether large corporates would rely on an outside service provider for a computing-related service. This indicates that the business market for utility services may be at the SME end, for high technology products and services, not larger companies with their own data centres. Also note that simple server virtualisation is the technology mentioned, probably with single site clustering, while grid is not yet part of this early extension to hosting.

Effectively, the service is yet to take off and we are at the nascent stage of the market for such services. The question is when it might emerge from this stage to early growth, as other hosting services, with some grid technology, such as IBM, Savvis and EDS have been waiting at least two years for the US market for utility computing to take off. The conclusion is that this form of the offering has quite limited appeal and is confusing for users to understand.

What may be a market driver is a new form of the utility offering, perhaps on a larger, effectively global, scale and with high promotional efforts and different target markets. Above all the focus would be on low cost. The far larger scale indicates a move to a

55 Matt Hamblen, AT&T to offer utility computing service, Computerworld, 8 August 2005.
56 See for instance, a 2003 study which seems to still be true, for the USA at least, Utility Computing: A Hard Sell, Saugatuck Technology, USA, July 2003
ubiquitous grid system and away from clustering and server virtualisation towards “PC replacement services” – with a new user terminal device at low cost – and high power computing for the SME markets first; corporate markets may build but perhaps only in a second phase as confidence mounts.

3.2.4 Packaging a commercial offering to go to market

Users need a complete everything, not just Globus or other middleware – a complete stack. However, today there is no concentration on a complete grid offering (not just middleware). In general the R&D projects so far have produced results but without all the necessary provisions for a future commercialisation, although this has been built in more to the later FP6 and now FP7 calls for tender. This situation is natural in that R&D projects set out with a specific scope set by research objectives.

What is lacking is an organised effort to pull all the components together to provide a commercial offering, with companion modules such as an active directory, overall policy manager, security with an authorisation capability, provisioning, Linux/windows scheduler, and the operating systems and more advanced message passing (with a standard reference implementation for improved message passing, perhaps as part of the standard toolkit).

This includes support from the projects for their outputs for those who would build real systems, and for post-build – in effect post-sales in a commercial world. For instance, for the GRIA project, IT Innovation have no internal resources for this follow-on support. It is here that we see one of the business models in the set examined in the next chapter coming to fore – the commercial support for an OSS platform as initiated in the in the Red Hat model. It may also lead to commercial forms of the original OSS platform, as there is now a market for OSS in an OSS/proprietary combination.

Preparation to go to market – packaging a commercial offering

What is different for grid in the overall route map for research projects to become commercial products and services? To prepare outputs from research projects requires a complete product packaging as well as an overall plan for forming a commercial operation around the outputs from research projects is necessary. The optimal route may consist of at least the following steps:

- Sign up a go to market partner(s) first (S/I, ISV, s/w tool vendor, etc) to either act as reseller, investor-incubator, or to take over the project output for complete productisation
- Perform a careful pre-entry preparation to go to market with steps for product packaging to take out R&D and academic origins and to include all components in the package for deployment with complete documentation and reference implementations
- Heavy reliance on channel partners and post-sales support partners
- Well-planned and strong marketing, to educate market about the product/service, including train the trainer and training programmes
- Careful introduction of product/service – for a product for instance: multiple pilots, detailed (4-stage) rollout for selected reference site
Figure 2.12 describes the generic process for product packaging and then the specific commercialisation route for one project or a group of grid R&D projects. It assumes support from a European Grid Technology Agency for the several problems of industrial transfer, especially for a commercial firm taking over the original OSS code and IPR to promote and sell packaged products and their support services, and/or a range of grid-based services.

**Figure 2.12. Product packaging and go to market preparation for launch**

Overall the product needs to be formed through the following process (same for service):

**Generic productisation process for a commercial software package**

- **Product Description**
- **Functional Specification**
- **Technical Specification**
- **Component Definition**

**APIs and Internal and External Interfaces Definition**

**Integrations**

**Commercial product definition**

**Set up**

- **Support**
- **Post-sales**
- **Pre-sales**

**Set up Distribution**

**Pre-Sales Support**

**Post Sales Support**

**Product packaging with all release components & documents**

The complete path to commercialisation could follow a route such as that below with the common steps to go to market, but under the wing of a market partner or strategic alliance:

**Turning caterpillars into Butterflies – Overall routemap for grid software taken on by a commercial firm to go to market**

**R&D Project**

- **Output Deliverables**
  - Software suite
  - Integration Know-how & experience
  - Installation tools & scripts
  - Licence strategy (?)

**Steps to go to Market**

- **Business Plan**
- **Finance search**
- **Investor search**
- **Partner search**

- **Business model creation**

**Technology Packaging and Integration**

- **Integration plan**
- **Porting plan**
- **Securitise plan**
- **Test Plan**
- **Release Plan**

**IPR Protection**

- **IPR strategy**
- **Licence(s) creation**

**Marketing and Positioning**

- **Marketing plan**
- **Channels strategy**
- **Market partners**
- **Marcoms plan**
- **Sales plan**

**Support & operations plan**

**Sign up Go to Market Partner**

**Channel partners and post-sales support partners**

- **Major Systems Vendors**
- **SIs**
- **Pilot implementations & Reference site as 4-stage rollout**
- **ISVs**
- **VARs**
**Licensing models and legal frameworks for coexistence with existing licence systems**

As considered in the previous chapter on OSS licensing, clarity of the licensing route to commercialisation is a key driver. Barriers to grid take-up for developers and users too often lie in the area of licence conflicts. For instance, an OSS licences that allowed public code usage but has not-for-profit restrictions if the source code is wholly or partially combined as a single product are obviously unsuited to a commercial concern.

What is really required is an easy way for the IPR conditions to be clear without lawyers, so that SMEs can use grid IPR safely and with confidence. Equally needed are EC R&D projects and EC contracts for the projects that embrace this principle, whereby their IPR conditions of use are straightforward and easy to understand and comply with, while conditions for third party usage ease industrial exploitation of code and IPR. This requires that the whole spirit of projects and contracts be nearer to free EU usage, while acknowledging that software delivery will be a mix of open source and proprietary software. A reference OSS stack would have to have careful licensing to permit proprietary components to be added at will.

For real grids, software licensing does not always fit well, as services may be part of the offering. Thus the optimum software licence for grid technology is one which is useful for providing grid computing as a service, as well as products, be they middleware or its applications on top.

For services, such as application software as a service, a specific form of licence is required. One approach is a usage based licence for the service, with the actual resources used being counted. For example one financial services provider uses a commercial licence with hourly charges that reflect a premium on the underlying grid infrastructure service provider’s charges for grid hosting of the application services. This ASP sees that the Apache and BSD licences are best for a commercial concern which builds on OSS and charges with a commercial licence for the bundled offering. Fortunately OSS licences are changing and developing – they are going towards a form suited to a more support-based business model.

Existing licences in the run-time grid environment may also be an issue with the technology, due the licence incompatibilities for grid exploitation of resources. Operational interworking requires that the grid middleware interworks with all the other software across many sites, and multiple organisations. What is needed in this case is a tick list of licence conditions to ensure all is legal. As examined above, this may be carried out automatically or in semi-automatic fashion (the reason for adding the licence discovery layer to the middleware as examined under technical drivers).

A further problem is code contamination by proprietary or OSS code with restrictions on usages. This may be limiting or even abortive for an OSS style of development – thus all code requires proof of provenance. For instance, Release 6 of OGSA-DAI has been formally checked/proven as clean by IBM for code originality and so now is copyrighted for original Intellectual property. The OGSA DAI project wants to accept code contributions from others but must be careful over provenance. Previously the project had received code from universities that had contamination problems. Such code cannot be used in a commercial situation without proven provenance for all code. Note that national and regional dictates over legal differences must also resolved, especially for on liability. US-EU differences in law have also been a problem for the Globus alliance on software licences.
Note that commercial software companies (and service providers) may not see the release of the input source code at application level as useful to them and so will use proprietary licences for their output products, which can also assure maintenance support. Commercial licences are best for high value specialised applications, which are to be tightly controlled (eg CFD applications etc). For all types of grid operational licence, a site licence based on usage is best, with annual support based on time used and numbers of users.

In summary what is needed is:

- A licence that allows free commercial development in the sense of Eric Raymond’s ‘The Cathedral and the Bazaar’ whereby the OSS movement is like an open market to drive a European grid software and services market. The emphasis may involve the use of multiple licence types with different intentions and user communities – those for research, those for commercial exploitation, which may include support services – in all, possibly three types: R&D; product development; services

- Resolution of conflicts with existing licences for resources when deploying a grid

- A point of ownership and liability for source code and its intellectual capital

- Checks on code contamination by any type of code that has restrictions (be it OSS or commercial) from any OSS-style development, with verification of provenance.

3.3. Stimulants to market take-up of a pervasive intelligent information service utility

If we were confronted with the concept of the World Wide Web in 1991, we could never guess that either a give away network or the sustaining software effort to support it in an OSS fashion would build the global information access tool we have today in the Web/Internet and its applications. This massive document communication infrastructure has happened outside conventional economic thinking about a profits-driven market economy, based as it is on freely given effort for a creative commons. It has happened against the wishes and without the help of the incumbent telcos. As such it presents an extension of previous economic thinking. Generally in the past, we have tended to undervalue the importance and productive power of open environments, be they for networking (the WWW) or for computing (a future pervasive grid utility) and outside proprietary commercial production.

The distinctive European vision is of a grid environment that operates from the level of devices to supercomputers. It serves communities ranging from individuals to whole industries, including data, information and knowledge and emphasizing resilience and scalability. It could have a significant economic and social impact far beyond the scope of existing compute and data grids.

57 S. Forge, presentation, Malaga, Spain Feb 2006, The rainforest and the sand garden, Towards an OSS policy for Europe, IIOSWC, Malaga FOSS conference 2006
In this perspective, the pervasive grid lives up to its name in the sense of an analogy with electrical power distribution. When a user wants (computing) power, he or she plugs in a standard plug to a standard socket and gets access to the processors (rather than electricity). The users pay a service provider for what they use. Private consumers may pay a different rate to a business that buys in bulk. However neither needs to worry about owning a computer (or in the electrical analogy, a power generator) – the (computing) power is generated somewhere else and distributed to them. Consumers and business get their power from a ubiquitous grid, which looks the same to everyone. In this grid computing model, businesses need to own fewer of their own resources. Third parties provide the facilities of grid infrastructures and services while users get access to those services. Businesses can also offer services over the grid in an extension to the electricity grid analogy for the UK electricity market.

This should be contrasted with the more US view of what is of grid technology as a tool for high power computing, with access really at programmer level. Its applications are exclusively in e-Science, certain vertical technical applications for high technology sectors, plus business analytics for some large enterprises. There may also be limited offering for utility computing from the hosting service providers, at the level of server virtualisation for clusters at one site. Such a view reflects grid technology as we have known it since the mid-1990s.

Grid computing today is in a relatively immature, nascent phase. Currently grids are only found in academia and a limited set of business users who have created for grids within one enterprise (‘intra-grids’) in leader industries. Standards are still evolving and a long term approach, of decades perhaps, must be taken to develop the architecture of a pervasive grid computing environment, in a next generation of grids.

The stimulants for this vision to become real are:

- Overcoming the limitations of current grid technology – this next generation will need to be highly secure, with trust across multiple domains and yet be transparent in doing this, in contrast with current systems.
- Most significantly, these grids must be open to widespread user and provider communities, in being easy to use, configure and manage, highly secure and also person-centric, rather than organisation-centric as at present.
- The openness implies being standards based. This will enable on-demand integration of services and applications to produce new combinations, in line with user requirements while calling on the underlying distributed resources transparently. A user can then create new applications while not understanding the underlying technology, yet not succumbing to vulnerabilities in security and reliability that this mixing implies, as happens today.
- Its user interface environment will be intuitive to provide an efficient information tool for the masses and for all sizes of business.
- For any widespread take-up, it must be highly reliable, as the grid will substitute for local resources both in usage and for its own development. Without reliability and

---

60 The OECD consensus group considered in 2005 that a one to two decade gestational period would be necessary for a pervasive grid technology to take root - OECD 2005, Global Science Forum, Report on Grids and basic research programmes, Final consensus report from the OECD Global Science Forum Workshop, Sydney, Australia, 25-27 SEP 2005

61 The vogue for web ‘mash-ups’ is one such instance – eg a mapping service mixed with advertisements from an online classified ads service, to produce houses for sale locations, using cross-site scripting (XSS) with Ajax (Asynchronous JavaScript and XML) stretches security, so that malware scripts can be introduced and propagated across the applications.
security the pervasive grid will never be widely adopted as a basic utility, because local computing will always be preferred.

To produce such an environment three further elements must come together:

- Major advances in what is still quite basic grid technology, at the levels indicated above
- Design of all elements as one coherent holistic system with an intuitive human interface environment (HIFE)
- Coordination of the advances, efforts and components into a complete infrastructure

As an aside, the convergence between grids and a mass market of users, perhaps by using the enhanced and standardised Web Services outlined above, provides a significant opportunity to move to a model of software development and service provision whereby the market dominance of particular operating system vendors is no longer a major economic issue. Grid technology introduces a new layer of abstraction between the operating systems and the applications. Battle now moves to a higher plane, above the lower layer of grid middleware. Because the lower middleware layer can handle resource heterogeneity – it copes with any differences among the capabilities and interfaces of operating systems. Consequently, the need for all users to have the same operating system, to host a compatible and tied in set of applications and for communications and application interworking disappears.
CHAPTER 3. Business models for grid software and services

1. The industrial context – the demand side

1.1. Objective

Our objective in this chapter is to understand what will be the range of best business models for commercialisation of grid technology research so far, in view of the ICT industry structure, and the shape of a future market for grid software and services.

1.2. Our approach

We have examined a range of business models, suited to each type of offering be it in application software, infrastructure software such as middleware or in a grid computing service. This has required two main activities which were carried out by desk research, and also using the interviews carried out under previous tasks with industry players:

- Examination of current and new business development practices, as they emerge in the software industry, to understand if and where they might be applied.
- Analysing successful alternatives for shorter and faster exploitation routes and faster time-to-market of grid products and services.

1.3. Demand for grid and the business models indicated

The history of digital computing has followed two paths since the late 1930s – more power and greater access, (or more personalisation) at much lower cost leading in the early 1990s to an era of high power computer (HPC) using massively parallel processor arrays with multiple or single data streams and/or instruction sets. Demands for HPC in the past were largely from e-Science, especially meteorology and particle physics plus some specific military problems such as nuclear weapons development.

Increasingly in the late 1990s, three different sorts of problem were attacked using HPC at lower cost, leading to new demands for it – firstly certain engineering simulation problems involving FET and CFD calculations in aerospace, auto, semi-conductors and civil engineering; secondly problems in life sciences for ethical drug development with gene sequencing and molecular synthesis; and thirdly certain financial simulation problems such as risk analysis from market movements.

Other areas, which were already users of HPC – such as seismic analysis for hydrocarbons – moved to the later highly parallel systems on a cost-justified based. In hand with the notions of single system with many parallel or tightly-coupled processors came the gradual spread of loosely-coupled systems or clusters, under clustering software for management and scheduling and to build HPC from standard server units.
Note that all these problems are amenable to parallelism in computing and/or data access operations. This leads to conventional computer industry business models – largely:

- products for hardware and software, including applications and operating systems
- support services and systems integration,
- possibly operations and remote hosting of an expensive shared facility be it a large parallel processor or a cluster farm of many servers

The next stage in HPC at lower cost was grid technology, employing more loosely coupled processors and data resources communicating over a (high-speed) network. Consequently the technology of grid systems is useful for the sorts of problems which:

- are amenable to parallelisation in the application software
- can withstand large latencies in the network transport of results, transactions and data transfers

However grid technology is more than just HPC – it has other attributes useful for problems in conventional business, specifically:

- exploiting its distributed nature to aid collaboration in design and other business processes, integrating business processes across geographically dispersed business units in one organisation, or many
- forming single data entities from dispersed databases

With the grid approach, resources may or may not be dedicated as they may be allocated on an ad hoc basis. Consequently a grid system can be viewed as an integration of resources on behalf of a user- a configuration which goes further in many ways than the parallelised compute problems, because it can be viewed as a method of finding computing resources and allocating problems to run on them. So grid technology’s major attraction for some will be as a way to lower costs through improved utilisation of infrastructure resources. End-users may or may not be aware of which standard computing units are involved.

Consequently what we have here is the basis for a service, for computing or data access etc, with a new business model akin to the distribution of electricity of gas. It incorporates server farm economies, for new modes of demand, which may be for more intense forms of computing but in a ubiquitous way. In essence the service approach presents the evasive vision for a business model of a pervasive intelligent information utility. It combines the elements of universal availability in both geography and in time – by being to some extent self-healing and highly reliable, also termed autonomic, for its utility computing.

In this vision, demand for the new business models comes from a need to improve integration and collaboration between organisations and between particular resources such as relevant data silos. Effectively it forms a universal upper layer above communications, as in the World Wide Web, to enable discovery, but at an economic level, innovation and growth.

Thus where the is grid concept may be used will change from specific sectors to a more general demand across business sectors, as a new mode of computing based on distributed loosely coupled elements. The types of demand will change from today’s business focus on the single enterprise to a cross-enterprise demand. In the vision of the pervasive intelligent information utility, grid technology may be subliminal, transparent to the user.
Looking at grid technology as an industrial platform in the context of the European economy we are focused on seeking business models that will repay the expenditure in research and development funding, in terms of inciting economic growth and employment both within its confines and more generally. Our aim is to highlight how an industry based on grid products and services can be established and then offer stable growth to the whole economy.

For a wider, global economy context, we also need to consider how a European model would both integrate and offer exportable products and services.

1.4. The ecology of the grid industry

The industry has evolved from the 1990s, largely under a guidance from e-Science to progress from distributed computing towards grid computing, and is progressing in multiple directions – virtual organisations largely in the academic community, towards enterprise, multi-enterprise shared grids as various attributes such as security enrich the original architecture. Perhaps the grid industry is now on a journey towards various interpretations of what may be termed utility computing.

We can see that its origins, in being steered by the academic and government interests, and especially in the USA, has led to a particular perspective of grid computing and on how it can add value, which is often perhaps seen as a series of communities, projects and national grid initiatives but with no clear overall structure. For instance, the GGF gives a view of the ecosystem for grid as a number of groups of interest, without any real driving structure (Figure 3.1).

**Figure 3.1. A view of the grid technology ecosystem from the GGF**
Looking forward with a commercial objective gives a further view, which comes from looking at the value chain in the grid industry, as examined in the grid markets section. Here we note that the IPR (intellectual property rights) as the rights to an ‘information good’ or intellectual capital asset are the original driver for the industry, be they open and freely available, perhaps as part of OSS, or in proprietary licences and under copyrighted control.

This view is shown in Figure 3.2, where we see the original IPR, private or shared in some way, generated in the base layer from many different sources – from private firms to publicly funded initiatives. IPR forms the basis of various commercial efforts, or those by public domain and open communities, who then translate the IPR into goods and services. These outputs are consumed by various classes of user group through the range of potential business models, examined here. The value in the ownership of grid technology IPR is released through the converters.

Figure 3.2. A commercial view of the ecosystem of the grid technology segment

Development of the basic grid middleware is only the initial step. We focus here on the next step – that of the ‘converters’ of IPR and the business models that take them to market.
2. Current business models in the software and services industry

2.1. The software industry’s traditional models

Today’s software industry revolves around three main forms of business model:

1. Packages – finished, encapsulated software, aimed at a specific operating environment and usually designed by independent software vendors (ISVs) to load and run, who may also have professional services for support (eg Oracle databases, SAP ERP)

2. Bespoke software for a single part of one business, produced under contract by software houses and system integrators; the major portion of software for large corporates falls into this category. Many players here, both large (generalized professional services) and small (niche vertical industry experts). OSS is highly increasingly relevant here.

3. The in-between:
   - Adding value to a basic software package such as a database – the value added reseller model (VAR) – or,
   - Integrating many pieces from many vendors – the system integrator approach – such as CAP Gemini, ATOS Origin, T-Systems etc – and also some systems vendors, eg IBM and HP, who offer systems integration services and also will try to combine their own hardware and software in the systems integration process.

4. Highly relevant for the grid industry in the future is the support model for an OSS package, as practiced by Red Hat, Suse, now part of Novell, and some others with some variations. This consists of offering a packaged ready-to-run form of the software such as Linux (which can also be downloaded free in its original source and binary forms) plus support services for defects and technical problems, installation and update.

5. A further business model is to offer the software as a service, as Infoware and CD02 do based on OSS, be the software used as a standalone program or embedded inside a product, with provision of professional services for integration, design and/or support long-term

We should also note that at the top level, the structure of the European software industry consists of a few large systems integrators, and no purely European systems vendors, apart from Fujitsu-Siemens perhaps and only one player (SAP) in the first 10 globally of packaged software vendor. The rest of the top 10 are all US players, with the exception of Fujitsu. Europe’s industry is dominated by small ISVs, VARs and systems integrators, with the majority having fewer than 50 employees and the major portion of that SME segment comprises firms of fewer than 20 employees. There are globally successful specialist package applications vendors – Dassault Systemes of France, Sage in accounts and also in embedded systems – ARM software to support the ARM core is the basis of all embedded mobile handset processors and the lead mobile handset operating system, Epoch, comes

---

from Symbian of the UK. Any business models we examine here must be oriented to that structure.

**ISV Software industry – two main revenue streams in its business models**

The classic business model is to create software, market it and provide the after-sales support in the case of a complex package and for semi-bespoke software, with pre-sales support for business analysis and design services. This gives an ISV two main revenue streams:

- the original licence sale for production and for development, usually sized in some way
- post-sales support services, again sized in some way

Depending on the market level of software sold (personal, SME or large corporate) the revenues coming from support can vary greatly, as a proportion of total revenues. For personal software, if it works after installation, support revenues are far outweighed by product sales. For an ERP package at enterprise level, the reverse may be true and the support services are usually much more valuable for tailoring, set up, commissioning, testing and post-installation modifications (or ‘maintenance’). Thus the model has two main revenue streams – licence sales, which may be perpetual, annual or any duration between that, and then various conditions on sizing of usage with multipliers by numbers of users, servers, hours of use, size of servers, numbers of accesses.

Licences may also be different for development and production usage, with development including source code libraries perhaps, editors, testers, debuggers, etc not required by end users. Revenues thus are designed to gather profits from a few development licences as well as the many consumers or corporate end-users of a single production licence.

Service revenues may be the real margin-earner, with up to four levels of maintenance support, as well as the charges for upgrades, updates, etc, and charges by SLA (service level agreement) on response time and call-out and time-to-fix. The latter may vary from hours to days to weeks to next version. Obviously a 24x7 service with on-site maintenance personnel in three shifts is substantially different from telephone support with 72-hour response.

The relevance to a grid business model is fairly exact for middleware products, but as we will see in the next section on analysis of commercial offerings, the package must be surrounded by a range of services, perhaps more so than for many commercial software packages.

As OSS utilities and applications gain market share, so variations on the ISV model have appeared. One notable variation is the MySQL model. MySQL acts as a commercial operation and sells a commercial licence for their database with maintenance support, updates, etc. But it also distributes a version of the database software as source code, for those who wish to use it as an OSS package, with rights to bundle it in other software but no support services. Interestingly they refuse to accept modifications from the open source community, doing all improvements and upgrades internally themselves, for reasons of guarding the code purity and of examining retroactive conflicts of new patches with existing code.

---

When developing code, ISVs can also make use of OSS in two major ways – firstly in incorporating OSS code as part of their own offering, if the licence permits (ie is of a BSD-style of openness as is the Apache licence, APL2) and in using OSS tools, such as the Eclipse toolset for Java. The licence is crucial as it dictates if development must be given back to the OSS community and what other software can be used with the OSS code. Obviously the ISV wants freedom to develop and not to give back what it sees as its own IPR.

**Systems integration business model**

The systems integration model has expanded over the last 30 years to take in far more business analysis as well as more knowledge about technology systems, but the principles remain of the same – take in some hardware, turn out some form of dedicated application system. This is then installed in the customer and integrated to new and/or existing business processes by connection into existing applications. The systems integrator adds the software required for the application on top of the hardware either by bespoke development (from the system integrator or a specialised third party (or parties) or by integrating a commercial off the shelf system (COTS) application package, from a third party usually.

A grid systems integration service demands a far deeper technical understanding of the basic resources in terms of hardware, networking, operating systems and databases. Relevance of the business model to the grid market is high, as the knowledge is so specialised today and so can command premium rates. The grid adds another dimension to systems integration, with its middleware demands and for the commercial market also requires a strong emphasis on business analysis, to understand the business processes underpinned by a grid-based application or application set, so that a solid business architecture can be established.

**Bespoke software business model**

Development of bespoke applications software for the grid architecture is a segment yet to flower, but can be expected to slowly grow with grid take-up. It demands a specialised development path – with parallelizing of functions, understanding of network delays, and a full comprehension of the grid middleware and how it works. The relevance to a future grid industry is high as many systems will be bespoke, especially early on with adaptation required for new and existing applications.

**VAR business model**

The value-added reseller – a company that takes a piece of software, and perhaps hardware, then adds value in further software – is likely to be crucial for the grid technology market. The most relevant business model for grid is perhaps to take a grid middleware package and combine it with an application, so the total can be deployed. An alternative is the ‘gridifying’ of existing or new applications to be resold in the grid market.

Note that OSS utilities, tools and applications are highly attractive for the VAR, as in principle buy-in costs are far lower, and, depending on the licence, any modifications may not have to be put back into a public OSS code repository as some OSS licences demand, although GPL style licences may require source code distribution with object or binary code. Such conditions are important for any grid middleware project, as VARs can be the key route to market for such software.
2.2. Examples of grid computing product and service business models – Oracle, IBM, Platform, BT (services)

Commercial models in the market today

Commercial models today in the grid market fall into two categories:

- New entrants, such as Platform computing, who sell a specific grid computing middleware product, with support services; they are niche ISV player really

- The major existing players (eg IBM, Oracle, SAP, HP, Dell etc) who have created special commercial grid business models tuned to their strengths in software and hardware products. Grid is just a new market to sell more of nearly the same products and services, if possible. Their position on just what is grid is dictated by the product/services set on sale, so each is individual.

Here we examine some specific grid middleware sales and services models from companies already in the market

The database-centric business model: Oracle’s business model for grid

Oracle’s main aim is to sell relational database management systems (RDBMS) and its associated business tools family. Thus its primary grid middleware strategy is designed to help sell and optimize Oracle business applications on an Oracle grid infrastructure based on an Oracle RDBMS. Thus Oracle focuses on bringing grid to the general business applications package market which use its database products line, together with its proprietary tools set (Oracle Financials etc).

Here we should first note that Oracle defines a grid as a data centric resource founded on its grid-enabled database, Oracle 10g. Thus Oracle’s target market is the data grid segment, as Oracle 10g is not designed to run large parallel or compute-intensive tasks or operate background processing tasks. The concept could apply equally to a clustered system at one location, which may also be a selling point in the enterprise market. When it wishes to distinguish its 10g grid offering from a cluster, Oracle adds that its grid offering supports more than one application hosted on a shared infrastructure. This provides the differentiation with dedicated application clusters.

Oracle’s 10g environment is made up of automated provisioning and management based on the centralized 10g database, for easier optimisation and tuning plus better management control. Oracle claims its infrastructure products provide optimal performance, scalability and availability with guaranteed quality of service levels, if customers centralise their database, on 10g. This positions Oracle as the hub of a centralised data grid. Customers have multiple choices of hardware server systems underneath (HP, IBM, Sun, etc) and a range applications, of third party ISVs, on top of 10g. Oracle’s middleware is the layer between hardware and application. Of course the data grid’s database is Oracle 10g.

In summary, Oracle’s architectural focus is a grid-cluster, preferably supporting a central database but certainly with centralised management of the database(s). Business applications requiring distributed storage and transactions but centralised control are well suited to the variety of data-centric configurations it is designed for.

Note then that Oracle’s technical strategy is to sell a homogeneous environment centred on Oracle 10g-centric grid at an applications support level, running on a range of hardware. There is a focus on commodity Intel x86 architecture as a means to lower total-cost-of-
ownership. For an application vendor, that is an ISV who is an Oracle partner, this is probably near-ideal in terms of performance, scalability and ease of integration. In contrast to the heterogeneous multi-vendor software environments, integration problems should be less – but lock-in is complete. Oracle’s architecture has a J2EE infrastructure around the RDBMS which is also hardware agnostic, suitable for the range of Oracle application tools. Really this is more of a cluster environment, or a set of linked clusters.

Generally, Oracle marketing strategy is:

- To position its 10g RDBMS grid offering as an infrastructure solution with vertical applications.
- Vertical applications are then sold through its own vertical industry groups (life sciences, design, telecoms operators, etc).
- To involve strategic partners such as systems vendors and systems integrators to build the complete customer solution as its strategy implies major dependence on third parties for the operating grid environment. Strategic alliances include Dell and previously Red Hat (before it bought JBoss).
- Sell Oracle business application tools, third party ISV partner applications, and custom applications, and using Oracle’s 10g database as the hub of the infrastructure.

**IBM’s grid business model – a full service model with professional services as the focus of the architecture**

The grid marketing and sales approach used by IBM for grid systems is highly instructive – and so we examine it in some detail.

IBM supplies many computing products and services to the corporate market, but not applications per se, and no longer any consumer products – its focus is strictly on enterprise and government. This sets the scene for current grid strategy, which is currently driven by its service-oriented architecture (SOA) marketing with a fairly technical focus on support of the customer’s business processes. IBM really sees grids as a means to pool resources, and so is comparable to Oracle – but with a remit that is wider than selling just a proprietary RDBMS (that is, DB2 EEE in IBM’s case) and the tools that sit on top.

IBM promotes the concept of a grid as being a distributed computing configuration to enable the sharing of resources (mainly servers, networks and storage) regardless of how many applications are involved, plus federation of data on demand across the organisation. Grids are thus positioned for the customer as the means to supply ‘virtualized’ resources to meet the dynamically changing application demands across the concurrent business processes.

This concept is used to meet the core IBM end-goal – to widen sales across systems, storage, software tools for management and development, networking and grid infrastructure, all backed by its high margin related professional services, especially for design and build. The latter replies to the crucial need for grid integration expertise.

Long term, the declared goal is to get IBM customers to build an enterprise-wide SOA (- a suitably mystic concept, and open to many wonderful and profitable interpretations) but based on a grid which brings the provisioning, workload balance and optimisation policies to manage the underlying resources of servers, storage and networks. Thus IBM can create a service-oriented infrastructure as three layers:
- A grid applications layer which the end-user sees, as a set of ‘business services’
- A linking and application supporting SOA layer – a services bus for business process support
- An underlying grid environment which is the support infrastructure

IBM’s Grid Computing architecture shown in Figure 3.3 based on ‘Grid Computing technologies’ and its integrated business aspects into different layers.

Figure 3.3. IBM grid computing architecture

Usage metering and monitoring mechanisms are included in this architecture. With users as a separate top layer, their requirements may be modified according to their specific needs without changing elements in the infrastructure services layer and business process layer above.

The grid environment is expressed using IBM’s latest other mantra, virtualisation. The latter may not be a new computing concept (i.e., running virtual machines on a physical machine for greater flexibility, which goes back some 70 years) but it is thought to sell well.

When defining grid, IBM does not make the same definition as Oracle of – many applications, one infrastructure – but does highlight common pooling, provisioning, and management resources. However, in the IBM design of a grid, applications may need to be modified for a specific distributed environment of (widely) dispersed resources; Oracle focuses more towards support for applications on a single cluster.

**Value Created for the customer by IBM grid computing**

IBM markets its grid offerings by creating an awareness of the potential added value of a grid. Generally, the value created for a business deploying IBM grid computing is viewed as at three levels, potentially:

- at application operation level, whereby customer organisations may speed up the major business applications by sharing the processing power and capacity of many processors

---


• at information flow level, enabling users to accelerate information access across the organisation globally, information resources which may be stored in many locations and on many processors across the organisation

• at business integration level, enabling customers to gain competitive edge through an integration of many tasks at different levels such as in the application and information levels

In addition, IBM in its 2003 literature claims more general competitive edge advantages of grid usage in terms of:

• Improved time to market for new products and services
• Accelerating collaboration and promoting operational flexibility across the organisation
• Leverage existing capital investments in storage and server power by reducing the number required compared to dedicated resources, although it is well to note that networking connectivity and design effort may have to go up (an often 'forgotten' factor)
• Meeting the varied business demands with increased productivity and efficiency

Moreover, IBM notes that much corporate data is not held in SQL databases – but in a variety of different media forms and file structures (a major difference to the Oracle view of storage, which is RDBMS-centric). IBM sees any grids (data-grids or not) as consisting of a logical federation of disparate databases and data types. The data types may be in e-mails, file systems, flat files, etc but are part of the data resources. Hence, IBM advocates logically ‘virtualising’ data – creating a logical data warehouse – rather than taking a centralized database on a cluster-grid as Oracle promotes. Inputs to the data warehouse may consist of any RDBMS from MySQL (OSS), to old Informix, to Oracle and DB2 of MS Sybase with contrasting record and file structures. Oracle has a proprietary vision version of data grids – that is data located in a single common, centralized database (which is always Oracle 10g). IBM notes that consolidating these heterogeneous data formats in a database can be extremely expensive, while leaving it where it is may have performance benefits, being located closer to processing.

Here IBM sales strategy is guided by the basic tenet that IBM is not in the applications business – so its goal is to only sell systems, storage infrastructure and most importantly for profits, related services for integration and post-sales support; it has two major approaches for this:

• IBM offers a pre-tested, pre-configured starting point solution, for the entry level with a turn-key heterogeneous infrastructure. This is marketed with its partners to provide the grid infrastructure and related services to pool a range of heterogeneous systems and storage and to federate data, all positioned as a single package, “Grid-and-Grow”. IBM, with its strategic grid middleware partners, offers the middleware for pooling and virtualization of resources and data. On top of the infrastructure, which is IBM-SOA compliant, either a custom solution can be built or else an application package from a third party ISV can be installed.

• For its overall strategy IBM aims to sell IBM systems and storage and IBM or partner middleware throughout the grid lifecycle so that its wide range of professional services come to the fore, to plan and assemble it all and then support it. It may employ specifically chosen ISV packages for each target market as well as build custom applications, all of which use the IBM products base and professional services. Under this policy, grids are sold to address specific major customer requirements for performance, or collaborative working problems, to overcome
dispersed data integration problems while at the same time, optimising heterogeneous resources – a bonus benefit.

In summary, IBM’s grid marketing concept is quite different to Oracle’s cluster-based data grid for its 10g RDBMS product of a centralised database and local servers to support a set of compatible data grid applications from its strategic partners. IBM’s model requires far more integration (a plus for IBM) but is less restricted and has apparently less lock-in. IBM can build a data centric grid or any other grid environment exploiting both local and remote distributed resources as well as distributed federated data, (which Oracle cannot). While IBM’s approach can be positioned as a systems integration exercise, it can easily also employ just a restricted range of pre-selected components (like Oracle). So IBM can provide a turnkey solution for a grid infrastructure with pre-qualified and certified third party applications running on an IBM grid infrastructure consisting of IBM hardware, its strategic operating systems, its RBMS DB2, WebSphere middleware, and Tivoli management tools etc.

**Bringing grid applications to market – the focus on process**

The two differing strategies of IBM and Oracle – and the differing architectures each prefers – lead to quite different business models and with them marketing approaches and targets.

While IBM focuses on selected vertical markets, delivering both horizontal and vertical solutions, Oracle focuses on bringing grid to the general business applications package market. IBM has five main target industry sectors:

- Life sciences
- Aerospace
- Automotive
- Financial
- Government markets

Across the target market segments, IBM focuses on applications suited to applying grids, the main ones being:

- Engineering and design
- Supply chain optimisation
- Business analytics
- Enterprise Optimisation
- Government development (widely different applications, but for many governments)
- Research and development using high power computing

The value created for each industry is different and it is interesting to examine the value-added to the each general customer business. The main markets inside each sector and other smaller sectors are:

- Pharmaceuticals, chemicals and genomics (life sciences)
- Aerospace design and test
- Auto design and test
- Electrical and electronic design and test
- Financial services – risk analysis, markets forecasts

---

66 According to IBM’s *White Paper on Grid Computing in 2005*
• E-Science and education and diverse applications including military, social modeling, civil engineering (Government)
• Oil and gas

IBM then identifies functional revenue and margin streams, in its five main functional grid market segments, from research from 2003 (see Figure 3.4).

Figure 3.4. Main revenue streams by grid function with services within each (2003)

![Figure 3.4. Main revenue streams by grid function with services within each (2003)](image)

Further segment refinement then takes place, so that later IBM literature gives more defined and segmented functional markets, showing the constant effort to understand the grid market and thus explain opportunities to the customer (see Table 3.1).

Table 3.1. Main revenue streams by grid function with services within each (2005)

<table>
<thead>
<tr>
<th>Energy</th>
<th>Financial services</th>
<th>Manufacturing</th>
<th>Bioinformatics</th>
<th>Telecom</th>
<th>Government and education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seismic analysis</td>
<td>Derivatives analysis</td>
<td>Product design</td>
<td>Cancer research</td>
<td>Bandwidth consumption</td>
<td>Collaborative research</td>
</tr>
<tr>
<td>Reservoir analysis</td>
<td>Statistical analysis</td>
<td>Process simulation</td>
<td>Drug discovery</td>
<td>Digital rendering</td>
<td>Weather analysis</td>
</tr>
<tr>
<td></td>
<td>Risk analysis</td>
<td>Finite element analysis</td>
<td>Protein sequencing</td>
<td>Multiplayer gaming</td>
<td>High-performance computing</td>
</tr>
<tr>
<td></td>
<td>Batch</td>
<td>Failure analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>throughput</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IBM’s sales approach for grid is interesting and useful to examine. It is to:

• Identify specific vertical market niches for grid

---

• Locate third party packaged applications that may benefit from grid architecture configuration
• Recruit these third party application suppliers into IBM’s grid-partner programme and help them address the customer expectations and also tune their products for grid
• Identify possible customers with pre-sales preparation and gain feedback on grid requirements
• Build an overall application profile by vertical sector including all vertical professional services as part of the IBM offer, with vertical industry sales specialists
• Identify the main opportunities in the selected target markets
• Market by vertical sector with strong communications – this includes helping the customer to build the business case to justify the investment
• Sell into the selected target markets.

Notice the emphasis on understanding real user demand by identifying applications demanding the features of grid computing, which may have already established a global following, such as CATIA from Dassault Systemes. This is traditional in high power computing, where there have often been high power processors looking for applications.

Based on this, IBM can offer a generic grid infrastructure upon which the vertical industry applications can be deployed and tuned. IBM builds its vertical grid infrastructures on both IBM and third party business partner middleware, for both IBM and/or third party operating environments.

**IBM’s business models**
IBM has a formalised business model for grid which can be defined by major components of structure and process – following Alt and Zimmermann.

**Structure:** The business model structure basically describes the different roles of different groups in IBM’s Grid Computing Business Model. In essence:

• The GGF defines the standards in IBM’s grid models with the Globus toolkit and OGSA
• IBM’s partners are responsible for Middleware development – e.g. Platform Computing, Data Synapse, Avaki, United Devices, and Entropia
• In addition, customers may participate in designing system architecture since the services providers are highly customer-oriented, part of the process.

**The Process for building a custom grid configuration:** The general process of how IBM grid customers select their own grid configuration is shown in Figure 3.5 from IBM research literature. Integration of technologies is based closely on user requirements to assure that the whole system is functional for a specific task, following

---

68 Building the business case for grid, ibid.
70 Boran Zhang, Grid computing introduction, from http://www.lancs.ac.uk/postgrad/zhangb/Business%20models-value.htm
72 IBM Research Brief, 2003
the option selection shown in the diagram. Users select their particular options in any of the three layers of the IBM grid model. In this way, a highly customer-focused service is presented. It requires a certain degree of understanding of the IBM grid model by both the business users and by the IT department. For post build and integration support, the business model for IBM is to take over this support, maintenance and further development of the system as user requirements evolve.

Figure 3.5. General steps for configuring a grid computing deployment with the customer

IBM’s approach provides us with useful pointers on commercial approaches and sales.

Platform Computing
Formed in the early 1990’s, the company is today the world’s largest ISV in proprietary grid middleware for the enterprise, with a worldwide network of resellers and partners. According its website, it currently has over 2000 customers globally, with an installed base of its grids employed in the fields of financial services, industrial manufacturing, life sciences, electronic design, government and military applications. Customer references include Deutsche Bank, Société Générale and J.P. Morgan in financials, Peugeot DaimlerChrysler and Ford/ Land Rover /Jaguar in auto, Texas Instruments and AMD in IC design, BAE Systems in aerospace, Celera Genomics, Bristol-Myers Squibb and Pharmacia, in life sciences and in government R&D, the US DoE National Laboratories, at Los Alamos, Sandia and Lawrence Livermore, the U.S. Department of Defense. It has built a strong global presence with some 400 employees (some 150 in R&D) at offices in North America, Asia Pacific and Europe (from website).
Aimed at the single enterprise, the middleware integrates heterogeneous enterprise-wide resources for performance. It provides the functions of workload, resource, and performance management. These are mainly for the multiple Unix and Linux, also Windows and even Apple environments. Products include a comprehensive range of commercial proprietary and open-source grid middleware. Main products are – the Platform LSF (Load Sharing Facility) family which includes the LSF Electronics Edition, Platform Symphony, Platform Rocks, Platform VM Orchestrator and Platform Enterprise Grid Orchestrator. To diagnose and improve system planning and utilization and to optimize licensing usage and spend, Platform LSF Analytics, Platform LSF Reports and Platform License Scheduler extend the Platform LSF product range. Platform Symphony provides the complete grid infrastructure for developing, running and managing both online and batch operations for business-critical applications for the financial services industry. Platform LSF HPC, offers an integrated workload management system for high performance compute clusters – specifically for Dell and IBM HPC hardware – as typically found in research centers.

The business model
Its business and its business model are based on several major principles:

Enterprise-level grids for all types of grid application
Professional services
Good positioning inside the customer, at a business, not just a technical level
Range and ease of application integration
Strategic Partnerships
Heterogeneity of resources managed
Dedicated sales units

Enterprise-level grids – in selling middleware for the enterprise, Platform has chosen not to specialise but offer middleware that can support a broad range of grid types – be they Compute or Data or Collaborative, or Analytical, they are largely Enterprise but this includes Utility. They may be used in both a dedicated high power application, or to share resources more effectively. This position is appropriate to their offerings – pure grid middleware with no applications or systems utilities like RDBMS, backed by support services.

Professional Services – Platform has realised that, if it is to sell at all, a serious effort for business and technical hand-holding of customers may be required to effectively counter the difficulties of understanding, designing and deploying a grid system. Thus it offers the necessary support services. Its range of consulting services offerings include (from Website) professional services with training for:

- Analysis of applications for grid deployment
- Business and technical architecture, analysis and design, with the integration of process flow and grid infrastructure in an SOA configuration
- Business case development
- Programme and project management
- Job flow and work flow design and implementation
- Production deployment and operations integration
- Product installation and configuration
- Applications enablement – analysis, design and implementation

Good positioning inside the customer – the company starts its pre-sales process with a study on the relationship between business process flows and the enabling systems infrastructure.
Hence Platform has a practical approach to understanding the business requirements for computing, to enable them to become close to the customer and form a strong relationship. The approach begins with a study on business processes aimed at establishing business behaviour patterns, to identify potential resource sharing from whom is using systems resources, for what, and how much. They may well use their professional services here. Then, as is common today in introducing new technology to an often dubious client management, a Proof of Concept grid configuration is designed and set up with a test infrastructure to demonstrate technical and financial feasibility, with a proof of workload, licence solutions, resource utilisation and business models of RoI, all aimed at the planning required for a first implementation.

*Application integration* – Platform has qualified over 200 applications on its grid environments. Its middleware market position means having libraries of applications – so a key thrust to drive sales for the middleware platform is to complement the offering with many compatible applications developed by other ISVs. Thus Platform claims to have built the industry’s largest library of grid-enabled applications. This implies conversion of third party software to run on its middleware products in a grid mode of deployment. The company has made near 200 applications grid compatible and straightforward to grid-install for cross-vertical industrial software products from its most important ISV partners such as Mentor Graphics, Cadence, Fluent, SAS suites, Synopsys, ANSYS, etc. For pre-sales, pre-sales service, deployment and post-sales support, Platform has also built dedicated industry teams. They have expertise in the major segments of financial services, industrial manufacturing, life sciences, electronic design, government and military applications.

*Partnerships* – in relative terms, Platform is an SME, so it depends completely on strategic alliances with the leading systems vendors (IBM, HP, Sun etc) as well as key vertical segment ISVs, in applications such as design, ERP, business analytics etc. Its products would not be saleable without these partnerships, as it would present naked middleware with no applications and no compatibility with infrastructure resources. In Financial Services it has formed partnerships with Powerléel and Axiom Software Laboratories. For digital content creation, it has partnerships to give middleware support for Softimage, Alias/Wavefront and others. For seismic processing for energy, oil and gas it supports packages from Schlumberger Geoquest, Landmark Graphics, Paradigm and others. In the industrial manufacturing segment, Platform has formed partnerships with Phoenix Integration, Computational Dynamics, etc. The company has also integrated some 50 applications in the CAD/EDA market for Cadence, Mentor Graphics, Synopsys, and Synchronicity. For business intelligence (BI) and business analytics, Platform Computing has a strategic relationship with SAS. On the infrastructure side, it has signed global multi-year integration and reseller agreements with the main systems vendors, IBM, HP and Dell (and in past times SGI).

*Heterogeneity and standards* – In contrast to Oracle and to some extent IBM, Platform must concentrate on middleware for a heterogeneous environment. It has no choice in its market position. Integration with others is made easier by adherence to standards as pooling of heterogeneous system resources relies heavily on creating common standards. Naturally, as part of this, the company is active in the grid standards community. Up to 2002, when it merged with GGF, Platform drove a collaborative initiative to develop standard APIs for distributed and grid computing, called the New Productivity Initiative, NPI. For Platform Computing, growing up in the USA, its standards are based on the GGF’s OGSA (Open Grid Services Architecture) embracing Web services. Thus the company has actively participated in the GGF to drive OGSA forward, as one process in its product creation. An
interesting marketing feature is also the development and promotion of open source versions of some products.

_Dedicated sales units_ – where needed, Platform has organised specific units for target segments. It has a Federal Division with a specialised staff who understand grid requirements for government work, to target specific US defense and DoE opportunities. Other teams prepare the marketing and sell to specific vertical commercial segments, with appropriate industry knowledge and contacts.

Platform Computing offers quite useful insights for a successful commercial approach.

**BT grid services model – a telco enters the ubiquitous utility computing market**

A recent development in Europe is the realisation that grid systems offer a potential new source of network traffic for business customers by the larger national carriers, principally Telefonica, Deutsch Telekom, and BT in the UK.

### 2.3. A synthesis of the key lessons and recommendations from the current examples of grid computing

Grid requires much explanation, education and support for the business customer. Business-level design is the first step to market and sell grids, being a highly technical product, in an unwelcoming environment that has not been educated. For the commercial world, what is required to start with is the concept of a business architecture to understand the place of the grid in the business. Overall, the marketing process requires:

- Good positioning inside the customer, at a business, not just a technical level to explain and educate
- Strategic Partnerships
- Professional services
- Dedicated sales units

More technical features:

- Range and ease of application integration
- Heterogeneity of resources managed
3. New business models and business development practices

3.1. New models for the future grid market

What will be the range of optimal business models for industrialisation of the grid research so far in view of the industry structure, and the shape of the future market in grid software and services? To solve this, we examine two questions in this section:

- Micro, or internal, model: we first investigate what are the feasible business models for grid products and services that may shape the future market in grid, in view of the potential demand, the grid value chain, and associated problems, such as IPR.
- Macro, or external, model: we should also examine business models for industrialisation in view of the context of the European computer industry structure, and a grid software and services market which is international – and so external forces will influence European endeavours.

We begin with the first question, of the internal model, and so investigate future feasible commercial grid product and service business models that will shape a future market.

Overall we can see the business models as being defined by the route to market from the various grid research projects being through two main avenues, products and services, with some overlaps, as shown in Figure 3.6.

Figure 3.6. New business models for the grid industry
Product business models
The product market can be outlined following the market analysis of the development of demand as shown in Figure 3.7.

Figure 3.7. How do OSS middleware products from R&D projects get to market?

The commercial model of grid development can exploit the current conventional European software industry model consisting of:

- **ISVs** – who take the OSS middleware offering from the original repository and either build grid applications on top, using its source code as guide to interfacing, sold as a runtime or configurable package with support services, usually under a commercial licence. The products they sell will be a range of grid applications and grid utilities, for performance, etc. A key product will be for sharing servers across many applications. Eventually this could be sold as a multi-enterprise product, pointing to a grid computing service.

- **VARS** – who use the middleware as a given part of their product, and may adjust it, assuming a suitable OSS licence, combining it with their own software as complete package, again sold as a runtime or configurable package with support services, usually under a commercial licence. The range of products will be middleware for re-using the same set of server set for all applications (eg for customer care CRM, billing systems and network operations in a telecoms operator) to specialised applications whereby their non-grid application portfolio has an upper-middleware set of functions added to it.
• System integrators – who use the grid middleware as an easy way to build complete
grid systems as the source code is available for understanding and modification, and
if the OSS licences are suitable, with more freedom to combine anything with
anything.
• OSS support providers – there are many business models here that make sense
along the value chain. The ‘Red Hat’ model for OSS as used for Linux is just one
of them, in which the service provider offers a bundled OSS package, ready to run
with installation guidance and support services for deployment and operations
• The systems vendors (Fujitsu-Siemens, IBM, HP, etc) will also be present selling
gridified infrastructure products— middleware, especially for server-sharing at
enterprise level, hardware processors and storage and networking products suitable
for grid configurations.
• Among the products that ISVs and systems vendors will offer, we may expect grid-
system management tools that are extensions of today’s systems management for
the data centre, and network management tools, integrated to give indications of
what is in use and where failures are likely. The concepts of the configuration
management database (CMDB) for grid will come into its own as a critical software
product, with dynamic updates for sites and servers in usage.

Over time, we would expect the range of products to change, with each stage of market
development, as shown in Figure 3.8.

Figure 3.8. Business models for grid products following the stages of market
development

<table>
<thead>
<tr>
<th>Grid Products Business models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Products and their business models are guided by industry development, which</td>
</tr>
<tr>
<td>may be seen as following the <em>Three ages of grid</em> -</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key products and services</th>
<th>Grid Market development</th>
</tr>
</thead>
<tbody>
<tr>
<td>• OSS based middleware</td>
<td>E-SCIENCE</td>
</tr>
<tr>
<td>• Large number of</td>
<td>e-Science projects</td>
</tr>
<tr>
<td>proprietary middleware</td>
<td>‘Amorphous’ products</td>
</tr>
<tr>
<td>products</td>
<td>still in development</td>
</tr>
<tr>
<td>• Bespoke applications</td>
<td>technical applications</td>
</tr>
<tr>
<td>dominate</td>
<td>for scientific/ VO s and</td>
</tr>
<tr>
<td></td>
<td>single organisations</td>
</tr>
<tr>
<td>• Proprietary middleware</td>
<td>ENTERPRISE</td>
</tr>
<tr>
<td>products usually top</td>
<td>Dedicated well defined</td>
</tr>
<tr>
<td>layer</td>
<td>Enterprise level</td>
</tr>
<tr>
<td>• Some universal OSS in</td>
<td>business grids for</td>
</tr>
<tr>
<td>lower layer middleware</td>
<td>Commercial/enterprise</td>
</tr>
<tr>
<td>foundations</td>
<td>applications</td>
</tr>
<tr>
<td>• OSS m/w products for</td>
<td>and increasingly for</td>
</tr>
<tr>
<td>upper layer</td>
<td>server and resource</td>
</tr>
<tr>
<td>• Gridified application</td>
<td>sharing and re-use</td>
</tr>
<tr>
<td>packages</td>
<td></td>
</tr>
<tr>
<td>• Enterprise server</td>
<td>WWG</td>
</tr>
<tr>
<td>sharing m/w</td>
<td>World Wide Grid</td>
</tr>
<tr>
<td></td>
<td>Ubiquitous generic utility</td>
</tr>
<tr>
<td></td>
<td>(comparable to a higher</td>
</tr>
<tr>
<td></td>
<td>layer on top of Web)</td>
</tr>
<tr>
<td></td>
<td>configurable across</td>
</tr>
<tr>
<td></td>
<td>multiple organisations</td>
</tr>
<tr>
<td></td>
<td>and VOs. Offered by GSPs,</td>
</tr>
<tr>
<td></td>
<td>Grid service providers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time to being mainstream</th>
<th>1-2 years, only in academic world</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5 years</td>
<td>B</td>
</tr>
<tr>
<td>Unclear</td>
<td>4 - 10 years</td>
</tr>
</tbody>
</table>

SCF Associates
We consider the third stage further in a later section on a pervasive or ubiquitous grid business model.

**Service business models for the grid industry**

Grid computing presented in some form as a service can support many business models but at least five are highly relevant. These five routes to market for the outputs from R&D projects are summarised in the Figure 3.9.

**Figure 3.9. Exploiting the output from R&D projects as a range of marketable grid services**

Examining each business model for grid commercialisation, we can see how a services market can be formed from the R&D projects:

- **Grid application software as a service (GASAAS)** – in this model the application software is hired, never purchased, by a number of users and offered as a complete service with setup by the service provider (SP). The SP may build the application itself, or buy it in from a third party ISV or VAR, who also offers fourth line support services to change code. Application software should be compatible with the most common standard grid middleware. The service is metered by usage in terms of time, or numbers of users, etc, under a service level agreement (SLA) for performance and availability. It thus includes billing and also customer care services, with help desks and echeloned support. It does not include the grid itself.

- **Grid computing as service from a grid service provider (GSP)** – the GSP offers the customer a complete grid environment as a remotely accessible service under an SLA for performance and availability. The customer may then run whatever grid
applications it wishes to, without hindrance, and, if suitably sized, without performance problems. The service includes set up and installation of access gateways of some form to wherever the customer requires. This may be quite a complicated deployment, especially if it includes the customer’s own hardware and software environments, used in a reciprocal arrangement for pooled resources for many users, with cross-charging. The service must include billing and customer care services, plus with help desks and echeloned support. It is close to utility computing, and should have autonomic features of robustness. The GSP would use the OSS outputs as the technical foundation for the service, either building and owning the grid infrastructure itself or buying it in from a VAR who has commercialised the original OSS code, or hiring a grid infrastructure, from an infrastructure builder and owner — a systems integrator, possibly a telecommunications service provider who offers the infrastructure as an extension of its network services. A key product for the GSP will be server-sharing middleware for enterprise applications portfolios, to reduce data centre costs, especially when the grid resources are shared across several enterprises.

- Grid Infrastructure Provider (GID) – builds and owns a grid which may be sold in capacity units to service providers, for basic grid services with support, and also sold to grid applications providers who combine applications and grid services as a bundled package. Such providers may be new players, although the telecommunications operators are most interested by this business model.

- Specialist grid support service provider (GSSP) – a technical set of services which includes systems integration for build, existing systems integration and operational support of grids, based on a deep knowledge the OSS grid middleware from the R&D projects in their commercial and robust form. It may well combine grid applications installation.

- Grid application service Provider (GSP+GASAAS) – here we combine hiring a grid environment and its applications as one service.

One key point on services is that commercial business models for many SPs will concern integrating commercial (licensed based) and non-commercial (OSS) software. System integrators also want to use the OSS components freely, to build applications of behalf of the end-user, mixed with commercial licences.
Figure 3.10. Including a contractual layer in the middleware

If we take the goal for the industry as an open grid, so that commercial companies can offer resources on a common infrastructure we may need a contractual layer to assure all licence issues are resolved, so any application can run with a grid contract on open middleware and also such that several resource providers can offer computing elements (be they charged for, or in reciprocal usage) as shown in Figure 3.10.

Example of business model for grid applications as a service, including the grid environment – a financial applications example

A highly successful business for software as a service (SAAS) combined with a grid infrastructure all as one service is operated by CD02 based in the UK, with a global customer base. It offers a highly specialised compute intensive application for financial markets risk assessment, for certain financial instruments. This is a multi-tenanted grid environment and so demands a scalable business model The grid offers this – a key advantage for the provider. It incorporates third party web services for both computing and content (eg market data). The whole is based on a secure internet deployment for access by its customers to an outsourced grid infrastructure – a Sun Grid – based commercial compute utility. The GISP (Sun) offers predictable pricing (in $ per CPU hour), a resilient infrastructure for a flexible deployment model for enterprise class support (see Figure 3.11).
Figure 3.11. The capabilities offered by the GISP for the CD02 business model

Source: CD02, 2006.
3.2. A vision of the business model for a pervasive intelligent information utility

One theme current in grid computing is that future computing services will be provided in similar fashion to electricity, gas, water and telephony, just as utilities are today and analogous to the example of the GISP above, with a GSP and taken globally with universal access for all. The development of grid technology has developed to where systems vendors (eg IBM and SUN) are beginning to sell grid services access by the hour, with Sun offering $1 per hour per server.\textsuperscript{73} Here we aim to produce a high-level market model for this form of grid computing as a service in a pervasive manner rather like the Internet, only for computing, not communications. (NB Review 2004 workshop models)

The model

Here we examine a business model for the micro-management of a future pervasive intelligent information utility grid. This business model bears some similarities to current structures for the airline, electricity and telephone industries, to which it is analogous, but specifically focused on the characteristics unique to the grid. There are various scenarios for such a pervasive intelligent information utility, and four are:

- An open ubiquitous resource is built using dispersed computing, database and networked resources from many sources, commercial, institutional and private donations, in a similar form to the Internet and the SETI project. It is owned by no single entity but shared by all.

- Companies may specialise as service providers offering grid services in a similar way to how telephone companies operate today but without public resources, apart from the Internet to connect. They may use high speed dedicated networks for the high value customers, and so offer classes of service.

- A third scenario is that large organizations will sell their (temporarily) excess or spare computing power to others via the grid; their capacity may sold to a proprietary operation, or a publicly accessible asset.

- A fourth scenario is a variation of the above, with a finer granularity and sees private PC owners selling spare CPU cycles of their machines. Since the GRID could possibly replace the PC as a future pervasive intelligent information utility, this seems a less likely option.

Here we focus and present business models only for the first and second scenarios, in that the third and forth scenarios may be embraced by the second.

Business model for Scenario 1

In the first scenario, the Internet is used as a communications foundation but with enhanced high-speed links to unite large numbers of resources in a form that could support the computing needs of a future information society as a universal computing environment. Usage is paid for, based on the ISP model by both enterprise and private users.

\textsuperscript{73} Interview, manager of a grid applications services provider to the financial sector for complex risk assessment forecasts
In this scenario, grid service providers (GSPs) will develop charging schemes for their services in a competitive environment. The overall concept can be viewed as three layers, of applications, generic global grid services, and the underlying communications infrastructure which links the resources spread globally (see Figure 3.12).

Figure 3.12. A first model for a pervasive grid utility – the world wide grid (WWG)

This model is open to many nuances and developments in players and offerings.

We could see grid portal players emerging, comparable to the Internet portals such as Google, Yahoo and AOL, but offering silos of content for the grid, as well as a wide range of grid applications, not unlike the large ISP portals, but perhaps with a more business oriented flavour. Thus the grid environment could spawn a new range of ISVs and VARs, as well as service providers who offer the grid applications as services.

There is the notion that grid technology could also become a generalised computing artefact, an underlying computing infrastructure for all. So it might be viewed as a consumer computing environment as well as one for business. Service providers in this model could also position grid access as an efficient alternative to buying computers, for any size of processor, from personal to the largest servers, requiring just a minimal unit that can access with an interface as light as a web-browser.

If the emphasis changed to targeting a global consumer market using ad hoc computing facilities for lifestyle and entertainment services, this would mark a sea change. Today’s usual thinking on grid technology is of a highly technical parallelised form of computing, restricted to solving fluid dynamic problems in one enterprise, and perhaps tomorrow, being...
the key to large-scale server economies by re-use across many applications for the large corporate.

There are serious doubts about the ‘WWG’ concept, especially on the issues of how resources are donated and paid for, security and reliability problems, as viewed in the drivers and barriers chapter. Also there is the question whether the dominance of a personally owned computing and communication device will ever be challenged, except if the two concepts are made complementary somehow, to meet real needs in a consumer-attractive fashion.

**Business model for Scenario 2**

In this scenario, a set of specialised service providers offer grid services in a similar way to how telephone companies operate today but without public resources, apart from the Internet to connect. These SPs may also use private high speed dedicated networks for the higher value customers, and so offer different classes of service, as some grid operations such as distributed computation may require lower latency in certain cases.

**Figure 3.13. The grid sold as an electrical or gas utility**

In Figure 3.13 we present the usual three-tier setup of markets seen in the electricity market, for example, with generators, (wholesale) distributors who own and operate the electricity grid, and retailers, who address the customer directly and supply after-sales support services. This incorporates two further features of such utility markets:

- the scenario of larger organisations acting as ‘generators’, reselling their (temporarily) excess or spare computing power to others via the grid, with capacity
being sold to a proprietary SP, or to a publicly owned/accessible asset like the Internet. It also embraces the notion of a finer granularity of computing resources, with private PC owners selling any spare CPU cycles on their machines. However, since the utility grid could possibly replace the PC, being the future pervasive intelligent information utility, it may be an unlikely option.

- the ideas of intermediation, of brokers – multiple intermediate-agent situations may occur whenever a grid services market exists, as two GSPs compete for grid customers. At some point, brokers (or middlemen, also termed retailers or perhaps suppliers by end-users) may re-sell grid services they bought in bulk from the original grid service distributor, who may in turn buy raw capacity from the grid infrastructure providers. Rather like the peering and ISP arrangements between the tier 1 telecoms carriers and the local and national resellers, and ISPs, we could see a system of intermediaries.

The three tier utility model could also be the case for the grid business model, as identified in the diagram. Taking as an example the privatised electricity market of the UK, middlemen may buy electricity from producers in wholesale volumes, and sell it at retail prices to a range of different end users through a wide selection of service plans and tariffs.

However, for the grid utility an added complication is consumption. It is not straightforward to measure in high volume for large numbers of users, as all resource components (servers, storage and networking) must be part of the accounting in some way. There are ideas that in the future a universal grid computing unit (UGCU) could be established for real time access, rather like telephone usage in minutes and Erlangs in speech circuits traffic volume. This may be a key foundation for building a competitive grid services market, with several different types of grid SPs competing and co-operating to serve different customer types, setting grid unit prices on a competitive basis.

In such conditions middlemen are permitted to make contracts based on minimum and maximum resources (servers, storage, etc) for each customer, with SLAs. A public sector grid supplier might also include social and economic support factors into its pricing, seeing grid as some form of universal service provision, one step up from telecommunications. This leads to the concept of a regulated market, for grid, as for other utilities. But it is in sharp contrast to the Internet market, which is an alternative regulatory model – unregulated apart from content and financial transaction restrictions. As the capacity could be considered as a commodity, a futures and options market could even arise, as has occurred in some stunted form in the telecommunications market. Further research may even produce a realistic market model for grid capacity as a commodities market.

Again, for the service providers in this model, grid access could be offered as an efficient alternative to buying computers and storage of any size and various pricing models would be necessary. The first future customers are likely to be those of today – corporates and SMEs in vertical industries such as the engineering and design, financial services, pharmaceuticals, energy etc as well as universities and other research organisations. Eventually might expect to see national or regional grids develop.

The macro – or external model
Here we examine the above range of business models in the context of the European computer industry structure, and a grid software and services market which is international and so external forces will influence European endeavours.
When viewing international dimension it is natural to first think of common standards. In the international context, grid standards influence is changing in the international standards world and is perhaps shifting away from the Global Grid Forum, whose origins lie in e-Science grids, as more industrially oriented standards bodies come to the fore. The question arises: Where should the EU invest its standards efforts?

A clear answer in business model terms is that standards efforts should be driven by industrial needs. OSS tends to set real standards for grid operations in terms of interfacing and integrating, both for code and for content. European business models should thus follow international trends in de facto standards for grid middleware, set by OSS largely, following (and participating in) the open standards bodies such as the W3C, OASIS and the IETF, be it at product or service level, including for content and data interoperability, with a focus in standards’ bodies on middleware protocols and APIs but following the market trends.

Note that Europe’s real interests lie in getting its services and products accepted in a broad user community across the world. Once standards have been established, export of products, services and know-how can follow, perhaps taking the GSM mobile technology example as useful in many ways.

Hence it may be useful for Europe to encourage the co-ordination of industrial standards, ensuring that EU efforts for grid standards are pragmatic and so driven by industrial demands. But we should also recognise that the standards arena is a political forum.
CHAPTER 4. Analysis of licensing models for free, open source and proprietary software

1. Open source licences and their legal and commercial aspects

1.1. Objective

Today's open-source software licences have a tremendous impact on how grid software is, and will be, written and consumed. Our objective in this chapter is to understand the legal aspects of software licences and the optimal choice(s) for grid projects.

OSS licences have the commercial effect of lowering the barriers of entry to the market. This is especially important for small companies which characterise the European software industry.

Thus we will survey and analyse the most commonly used and also the most appropriate licensing models for free, open source grid software and that used in the most prominent EU, US and national and international grid projects (eg Globus, GRIA, Next Grid, etc).

Commercial software licences are well understood and used throughout the corporate world and are used for all the early commercial grid products for enterprise level working. OSS has often been used and is growing in use for utility or infrastructure software – what might be called ‘commodity technologies’ such as application servers, operating systems, languages, development toolsets and of course communications. For grid, the story may be a somewhat different to other software environments. The relevance of OSS to grid computing – which can operate with either commercial or open source software licences – is that OSS licences hold some quite specific legal and other advantages for grid middleware, the most evident being:

- No legal restrictions on configurations and usage such as numbers of users, numbers of servers, class of server, duration of usage or geographic location, important in a situation of dynamic assignment of resources and access across many software environments and many organisations with different licence agreements even with the same software vendor. These restrictions are a key tenet of commercial licences.
- Depending on the licence type, no legal restrictions on what other software may be run with the OSS programme, again key for a situation of dynamic assignment of resources across many organisations.
- Lower cost of purchase (but not necessarily of ownership, especially with support).

---

74 By 'open source software' we imply the definition has been proposed by the Open Software Initiative (OSI). The Open Source Definition, Version 1.9 is available at http://www.opensource.org. The full list of Open Source Initiative (OSI) approved Open Source licenses can be found at: http://www.opensource.org/licenses/index.php
75 See: An analysis of the economic impacts of Free and Open Source Software and its potential for Europe, a report on behalf of IPTS/JRC, with recommendations on future EU policy for FOSS (policy published as a chapter of a book on FOSS) - for the experts group of the European Commission; synopsis available on IPTS website and also the Open Source Initiative (OSI) website
• Depending on the licence, freedom to modify and re-use software, with comprehension of the code being possible from the source form being available.
• Openness of source code to find faults more easily.
• Possibly, faster bug-fixing through the mechanism of the OSS community.
• Resistance to viruses and malware attacks through openness and proving.
• Lower down time due to a combinations of the above three points. 76

We now proceed to examine the different types of OSS licence and which will be a best fit for each grid situation and business model. Note that for services based on OSS software, a completely different type of licence or agreement can apply.

1.2. The legal basis for Open Source Software licences

First we ask: what is the legal basis for OSS?

Current legal focus for software is on a set of laws that protect the rights of publishers of commercial software, as in any intellectual property. These laws apply to an agreement between the publisher and its customers, the users of the software and define the software property’s conditions of usage. Thus most OSS licences are mostly about products and less about market offerings which are service, and are mostly about source code and are grounded in copyright (which through the Berne Convention and the WTO TRIPS are globally applicable).

Immediately we can see that this situation is a subset of a larger, more general situation, one which centres on the use by others of software developed by one person, company or group. The idea that the proprietary commercial software agreement system is the only possible and viable system is not so, although it may seem that effort has been spent to convince the public and the users that a commercial agreement is the absolute rule, the only way to look at the issue. It also assumes that only important aspect of the software for a user is what it permits the users to do, but not how well it does it – because the commercial licence explicitly rules out improvements or modifications.

A further implicit assumption is that the power due to the publisher over the program and the restrictions on its usage are necessary. The premise seemed reasonable until self-organising software communities returned from the 1960s to demonstrate that a plentiful supply of useful software can come from a completely open development and re-use approach, one which has no place for the constraints of proprietary software licences. Such approaches make the use of closed software seem limiting on the freedoms that software could have, as a malleable, plastic medium which can be reshaped.

Most appropriately the OSS movement has put in place some rules to protect this unfettered state, so that OSS cannot just be hijacked as a proprietary captive property. Although here we would note that some open software developers could and do see this as their goal – the goal being success in having many users in the world, be it of the open software original, or of a proprietary package bundle (for example X-Windows from MIT).

The freedom protecting OSS terms ‘invert’ standard copyright law, with an approach called ‘copyleft’. It inverts the idea of privatizing software, instead aiming at keeping software

76 These are explored as an economic analysis of OSS in the above IPTS study report and in the paper by S. Forge, ‘The rain forest and the rock garden: the economic impacts of open source software’, info, Vol 8, No 3, May 2006.
free, with permission to run, copy and modify the program. As such it requires the source code of modifications to be made available to the parties to which the binary code is made available, as a means of conserving the open, free-access nature of OSS with progress. In essence copyleft means that the rights belong to the community. What ‘copyleft’ explicitly forbids is the addition of new restrictions on usage of the modified version, including charges for the modifications, so for an effective copyleft, modified versions are also free. This means anyone has the right to use, modify and distribute their new version under the same free usage terms.  

The costs, for instance of distribution, are under the same open free-access conditions as that of the original, before the modifications, to ensure that it becomes available to all the community. This condition imposes discipline on developers for what sort of code can be combined with OSS. Any addition of proprietary code into OSS must be ruled out, as it brings the restrictions of the IPR’s licence obligations. The ideas of copyleft were incorporated in the first major OSS licence for the GNU operating system (GNU: “GNU’s not Unix’), termed the GNU General Public License (GPL) by Richard Stallman and his collaborators in the mid 1980’s. Note that the OSS model and its licences, including the GPL, do permit commercial business opportunities, around the free-access software, in distribution and support services.

1.3. The major classes or divisions among OSS licence types

OSS licences comply with the preconditions of the Free Software Definition and the Open Source Definition. There are several types that comply, identified by their legal characteristics, and some that do not:

- **Licences with strong copyleft effect such as the GPL** – As noted, licences have a strong copyleft effect if licensees may only distribute the modified software under the same licence terms as the software originally acquired. In this sense the powers of the licence are handed on – the copyright law’s workings are active here; it is sometimes referred to as ‘viral’ in the sense of the rules over usage being handed on, like a virus transmission. Copyleft implies that individual rights do not apply, only community rights and any extensions or modifications revert to the commonly held version. The most well-known form of this is the GPL licence (GNU General Public Licence, GNU sometimes taken to have a tautological meaning of ‘GNU’s not Unix’) which originated in the 1980s. Note that other similar licences may not be GPL compatible, despite having some similar terms. Also note that any generic entities that are expected to be used with OSS and with proprietary software are permitted as copyright does not refer to use of the running software eg in a remote procedure call from another piece of software. This is important for entities such as Web Services, designed to be called by any application or utility.

- **Licences with limited copyleft effect, such as the LGPL** – the lesser or library GPL is the most well-known. The licence is similar to strong copyleft but if modifications are made in a separate file, such file's content may be distributed under a licence that differs from the original licence, e.g. a proprietary licence. This kind of licence is designed for easier combination of different pieces of (interworking) software under different licences and so is useful for software libraries. For some versions the licence conditions may not require that modifications be made in a separate file to be distributed under a different license. But the modifications must be visible to be seen as a derivative work, with use of the library through well defined APIs.

---

Something which makes use of the LGPL somewhat uncertain in some projects is its origins - it was written at a time when dynamic linking mechanisms were not as commonplace as today. With modern technologies (eg dynamically linked languages), including grid technologies, this can give room for divergent interpretations of the LGPL provisions.

- **Licences without copyleft effect** – such licences grant all rights and privileges of an Open Source licence to the user and do not restrict the user in case of modifications with further conditions as to which licence has to be used subsequently. A licensee may distribute modified versions of the software under any licence of his/her choice and so can transform the OSS licensed programme into a proprietary software product.

- **Licences with restricted options** – amended by specific conditional clauses, the legal scope varies. Note that in consequence, a licensee may have a restricted choice of how to distribute modifications of the software, and also restrictions in a commercial situation including perhaps not-for-profit conditions.

- **Licences with privileges** – this kind of licence is often used for turning proprietary software into Free Software. It grants the licensees all those rights which are characteristic of Free Software, however it may also reserve certain named special privileges for the licensor, if a licensee modifies the software.

- **Other associated ‘free’ software licences** – some licences used in the world of ‘free software’ seem to be OSS without fulfilling the preconditions of the Free Software Definition or the Open Source Definition. Some of them are "non-profit licences" which, ban commercial distribution, other than under real OSS licence conditions.

- **Open Content Licences** – their aim is to use those concepts underlying free and open source software for copyrightable works other than software. However with no generally accepted definition of Open Content, these licences are quite different, ranging from those from Apple to attempts to build an open commons licences such as Open content licences of the GFDL form (GNU Free Documentation License) which are similar to GPL open source in their permissions of usage, but are applicable to items of original content such as works of art, documents, images etc, not source code. For instance many Wiki sites such as Wikipedia are licensed under the GFDL.79

- **Other licences for fair use of immaterial goods** – licences for fair and perhaps unrestricted use of intellectual property (or immaterial goods) which can neither be assumed to be OSS nor in the class of open content. Such licences build on specific conditions for the particular category of immaterial goods. An example is the CAMBIA BIOS (Biological Innovation for Open Society) licence agreement. It is used for agricultural products such as seeds, giving royalty-free rights with a licence to make and use the intellectual property and technology for developing, making, using, and commercializing the licensed products without obligation, including a right to sublicense.

For some of these licences, it is debatable whether all preconditions are applied in practice, because the relevant definitions, the Free Software Definition and the Open Source Definition, are not identical and are being constantly modified by development. Such licences might be viewed as "qualification controversial". Some common OSS licences are shown in Table 4.1.

---

79 Also see OSSCentre: Copyrights. Creative Commons License, at: http://www.asiaosc.org/enwiki/page/Creative_Commons_License.html
Table 4.1. Some common OSS licences and others, by type

<table>
<thead>
<tr>
<th>Licence</th>
<th>Class of licence</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPL (version 1, 2 and 3)</td>
<td>Original GPL, strong copyleft</td>
</tr>
<tr>
<td>BSD Berkeley Standard Distribution</td>
<td>Original BSD</td>
</tr>
<tr>
<td>Apache (versions 1 and 2)</td>
<td>BSD type licence – no copyleft</td>
</tr>
<tr>
<td>LGPL</td>
<td>GPL licence with limited Copyleft Effect</td>
</tr>
<tr>
<td>Mozilla Public License (MPL)</td>
<td>Licence with limited copyleft effect</td>
</tr>
<tr>
<td>Condor Public Licence</td>
<td>BSD type licence – no copyleft, University of Wisconsin</td>
</tr>
<tr>
<td>Common Public Licence (CPL)</td>
<td>Strong copyleft, from OSI (Open Software Initiative)</td>
</tr>
<tr>
<td>Python License (PSF License Agreement)</td>
<td>Licence without copyleft effect</td>
</tr>
<tr>
<td>Zope Public License</td>
<td>Licence without copyleft effect</td>
</tr>
<tr>
<td>Artistic licences (several versions)</td>
<td>Licence with restricted choice</td>
</tr>
<tr>
<td>MIT License</td>
<td>BSD licence derivative – no copyleft</td>
</tr>
<tr>
<td>Jabber Open Source License</td>
<td>Licence with strong copyleft effects</td>
</tr>
<tr>
<td>Q Public License</td>
<td>Licence with privileges</td>
</tr>
<tr>
<td>OpenLDAP Public License</td>
<td>BSD</td>
</tr>
<tr>
<td>Aladdin Free Public License (several versions)</td>
<td>Other associated ‘free’ &amp; non-profit software licence</td>
</tr>
<tr>
<td>Sleepycat Software Product License</td>
<td>Restricted options licence</td>
</tr>
<tr>
<td>Eclipse Public License</td>
<td>Licence with strong copyleft effects</td>
</tr>
<tr>
<td>Plan 9 Open Source License Agreement</td>
<td>Other associated ‘free’ &amp; non-profit software licence</td>
</tr>
<tr>
<td>Apple Public Source License</td>
<td>Licence with privileges</td>
</tr>
<tr>
<td>GNU Free Documentation License (GFDL)</td>
<td>GPL-like content licence</td>
</tr>
<tr>
<td>OpenOffice.org Public Documentation License</td>
<td>Content licence, OSS-like</td>
</tr>
<tr>
<td>Open Publication License (v. 1.0)</td>
<td>Content licence, OSS-like</td>
</tr>
<tr>
<td>Open Source Music License (OSML)</td>
<td>Content licence, OSS-like</td>
</tr>
<tr>
<td>Public Documentation License (PDL)</td>
<td>Content licence, OSS-like</td>
</tr>
<tr>
<td>Open Content License (OPL),</td>
<td>Content licence, OSS-like</td>
</tr>
<tr>
<td>Free Art License</td>
<td>Content licence, OSS-like</td>
</tr>
</tbody>
</table>

1.4. Which licences are currently in use?

Here we briefly survey and analyse the most commonly used licensing models for free, open source and also proprietary grid software in the EU and also the most prominent national grid projects (such as MyGrid, ACI-Grid, Grid.It, Globus, OMII, etc).
Proprietary grid licences
The many commercial middleware products, grid applications, and grid application services, including grid infrastructure services, available in the USA, Asia and Europe have specific licences by vendor which are particular to the vendor but in general cover:

- Installation conditions – usually more freedom than on conventional licences with greater acceptance of heterogeneity and the unknown nature of potential platforms.
- Services licences tend to be usage based (eg per hour) for applications and grid infrastructure, with the latter including a potential per server and server size provision.
- Protection of the original code from further development or modifications as in normal commercial agreements.
- Support services with variable conditions, levels and SLAs, as in normal commercial agreements.

Open source grid licences in operation
The licence types used in the EU and other R&D projects and consortia are surprisingly different, reflecting the aims of the publishing group. A summary table of the major licence holding projects and alliances is shown in Table 4.2. Those of a more academic nature go for GPL-style licences. Those hoping for commercial take-up of their outputs in some form opt for the APL2 style, with a number of intermediates.

Table 4.2. Key alliances and projects and their licence types

<table>
<thead>
<tr>
<th>Grid R&amp;D project or organisation</th>
<th>Licence type(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Globus Toolkit and reference models – GGF now OGF</td>
<td>Apache Public Licence Version 2, APL2</td>
</tr>
<tr>
<td>GRIA</td>
<td>LGPL Lesser GNU public licence. The project noted: LGPL was chosen as it allows business decisions on usage, and developments, which the pure GPL does not as all is ‘public domain’. Apache form of licence seems ideal (APL 2).</td>
</tr>
<tr>
<td>Unicore</td>
<td>BSD – in SourceForge</td>
</tr>
<tr>
<td>NextGRID, (The Next Generation Grid)</td>
<td>See APL2 as being best. View LGPL and GPL as viral and limiting</td>
</tr>
<tr>
<td>SIMDAT (Data Grids for Process and Product Development using Numerical Simulation and Knowledge Discovery)</td>
<td>Possibly a mixture – OSS only as appropriates (and many different projects)</td>
</tr>
<tr>
<td>HPC4U (Highly Predictable Cluster for Internet-Grids)</td>
<td>Hpc4U uses grid middleware CCF from the University of Paderborn, Germany, which is GGF compliant, used under a GPL licence. Hpc4U software delivery will be a mix of open source and proprietary software, licence to be decided.</td>
</tr>
<tr>
<td>CDO2 (Grid computing application service for financial service industry)</td>
<td>Proprietary licence for application service</td>
</tr>
<tr>
<td>MyGrid</td>
<td>Apache</td>
</tr>
<tr>
<td>Akogrimo (Access to Knowledge through the Grid in a mobile World)</td>
<td>GPL (?)</td>
</tr>
<tr>
<td>Project</td>
<td>Licence Model</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>CoreGRID (European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, Grid and Peer-to-Peer Technologies)</td>
<td>BSD based – htmlArea licence</td>
</tr>
<tr>
<td>DataMiningGrid (Data Mining Tools and Services for Grid Computing Environments)</td>
<td>3 different schemes: – foundation code – Globus (Apache 2.0) and Condor (Condor Public licence) – both BSD-style; datamining interfaces – proprietary; data mining applications – property of the partners</td>
</tr>
<tr>
<td>K-WF Grid (Knowledge-based Workflow System for Grid Applications)</td>
<td>Fraunhofer FIRST licence agreement – BSD academics only / public research restricted type licence open for modifications for academic usage only with no return of contributions required- and free to give away to others. Commercial for profit requires written agreement.</td>
</tr>
<tr>
<td>Provenance (Enabling and Supporting Provenance in Grids for Complex Problems)</td>
<td>Common Public Licence (CPL) v1.0 from OSI (Open Software Initiative)</td>
</tr>
<tr>
<td>Grid.It</td>
<td>LGPL</td>
</tr>
<tr>
<td>OMII</td>
<td>Modified BSD - all software developed by OMII is provided under the Modified BSD licence defined by OMII.</td>
</tr>
<tr>
<td>NAREGI: National Research Grid Initiative project, Japan</td>
<td>Academic project, intended for public research and universities – Mixed licence types – APL2 for Globus Toolkit, OGSA DAI, OGSA EMS implementations; GPL other components</td>
</tr>
</tbody>
</table>

**Dual/multiple licences**

For future commercialisation, many R&D projects or their commercial derivatives may find it useful to follow the example of MySQL and others who have multiple releases with each having a different licence, depending on target market:

- For a developer, a commercial licence may be made available, that allows the source code to be used in further derived commercial products. The developer pays a lump sum or royalties on sales for this right to use the code and also for the rights to any modifications or advances, which do not have to be handed back to the original authors. This effectively is a standard commercial agreement.

- For end users who want a free version without support, (from the ISV) an OSS version is available. This may have GPL restrictions on it that any modifications/defect repairs or additions must be given back to the OSS community by being posted in an open repository. Note that open source licenses make it possible to get support on the free version, from parties other than the original ISV itself – which certain commercial licences may not allow.

- There may also be a freedom condition of the licence to permit the original code in source or binary to be used unaltered as part of commercial releases with other third party products – but there is no support for this portion of the bundled product for the original authors. The third party publisher is responsible for issuing new versions to keep up with the current release of the free OSS component, and in the
EU for liabilities, obviating the original authors – again the Red Hat business model.

It is also possible to make modifications and give these back to the source repository for future incorporation. However some dual proprietary/OSS licence issuers do not want outside contributions as they see them as unsafe from the point of view of possible contamination with code that is governed either by OSS viral, or proprietary rights. By refusing any outside code the original authors also then control all progress and evolution directions, as is the case with MySQL, which also an incorporated company.

Multiple licence models are of great interest for commercialisation of grid software, while maintaining control of direction.

Safeguard of intellectual property rights with multiple licensing

As noted, the aims of multiple licensing are to give different freedoms to different parties in ways which are of benefit to the licence issuers, usually the source code originators, and also to potential users and/or developers, either commercial or non-commercial. As long as the licences are appropriately chosen for each task, then the IPR can be channelled in the right way by each licence type.

The various GPL types of licence give a complete control over modifications and extensions, and so their actual clauses have to be examined closely in offering dual GPL-style and commercial version licences to ensure that the code re-use clauses in the GPL-type are not contravened by the non-GPL-style licence by code which is common between the two versions. Here, the Mozilla-type licence (MPL) may be useful or the MySQL type scheme for dual licences with commercial and open public versions, which safeguard intellectual property rights in the open source code version.

For purely commercial exploitation, the Apache Licence version 2 not only gives freedom of usage to commercial vendors but also protects the original authors from liability for any mistakes or damage the subsequent licensees may cause. This illustrates the legal situation, that IPR issues are more than just control of usage but also involve liability for those creating derived works – the fact that licensing is a bigger subject than only a controlled IPR transfer.

2. Questions on commercial usage of OSS licences

2.1. What constraints do the main licences place on users and developers?

Among the open-source licences available, two are often seen as predominant – the BSD\textsuperscript{80} and GPL\textsuperscript{81} licences. Both allow free use of software. We examine these for the range of constraints on commercial usage

**GPL in brief**

GPL is short for the GNU General Public License, from the Free Software Foundation Inc. The GPL is probably the best-known and influential open source licence. It is used in various forms for Linux, MySQL, etc. It ensures that all software released under it will be Free Software and will always remain so. Key GPL requirements are that:

\textsuperscript{80} See http://www.opensource.org/licenses/bsd-license.php

\textsuperscript{81} See http://www.opensource.org/licenses/gpl-license.php
• The source code is made available, free or at a reasonable cost, with or without the binaries of the software.
• Everyone is allowed to copy, distribute and modify the software as long as all modifications are transferred to the party the modified software is distributed to, but only in the case of the software being distributed.
• Once the software is placed under the GPL, certain of its provisions cannot be revoked. It can be released under different licences simultaneously by the original authors.
• Unless expressly specified otherwise, GPLed software comes without a warranty.

Thus software covered by the GPL can be modified, but any release of that modified software must include an offer for the source code under the same GPL licence. Basically, the GPL creates a consortium. Anyone can use or change the program, under the conditions that the licence states, of no charge being made, GPL version 2, 1991, in clause 2b:

You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

A highly simplified summary of its overall licence conditions is the following:

This software contains the intellectual property of several contributors. Users cannot expect to use this valuable resource in their own work freely at no charge. The developers of this software are willing to trade to users the right to use their intellectual property in exchange for the right to freely incorporate the licence user’s code [which incorporates their code] into theirs. This exchange is to be done by way of and under the terms of the GPL (complete information regarding the GPL can be obtained from the website of OSI).

The GPL advantage is that it guarantees the developers that their code will remain free, forever, and that any future developers who contribute will be similarly required to keep it free. This sometimes acts as a commercial constraint. Thus commercial software vendors – and their end users – may be reluctant to use GPL software. Because any code they add to a GPL program, to be released as a single product in binary, or any changes, must be released with the source code to the party to whom the software is redistributed in both binary and source code forms. Everyone is allowed to copy and modify the software and, in the case where they redistribute it, they are required to make source code of the distributed version available. Control over what could be seen as commercial IPR (intellectual property rights) may be lost. The copyright law’s action leads to what is termed the "viral" nature of the GPL – the code’s licence may propagate across any code it is included within. This may limit its commercial attractiveness. Naturally these GPL conditions do not apply to programs that run with or call the GPL program, such as operating system like Linux, but only to extracted and combined source code.

Note that GPL-licensed code can be modified and used in house (i.e. without redistribution to another legal entity) without requiring the modifications to be contributed to the public. Companies like SAP and Google do this on a regular basis for internal operations.

There are variants of the GPL–style licence conditions. Some programs are distributed under a license requiring that modifications be put back into a public OSS code repository. But usually for GPL-style licences, it is only necessary to distribute the modified source code to those it is distributed to, with the object or binary code. This has two major effects.
The first is that it allows code forks, i.e. different variations from the original. The second is that the source code of any modifications is only required to be made available to the parties to which the object binary code is made available. Some companies for example create bespoke software and make it available to their customers under an open source licence. But this does not necessarily mean that they have to put the source code in a public (centralised) repository.

**Licences without copyleft effect**

Licences without copyleft are an important OSS addition in that they grant all rights and privileges of an Open Source licence to the user but do not restrict the user in case of modifications with further conditions as to which licence has to be subsequently used for the revised code. Thus, the respective licensees may each distribute modified versions of the software under a licence of their choice and so may also transform the licensed product into proprietary software, with a commercial style licence – and their own copyright control.

**The BSD licence in brief**

The BSD licence (Berkeley Standard Distribution, from the University of California at Berkeley) is the earliest formal licence in the open software movement, dating from 1970’s when it originated for the diffusion of a free version of the Unix code. A whole range of licences now follow its basic tenets. From a business perspective, BSD a useful type of licence, as it may be integrated with proprietary code for releases by the integrator. Unlike the GPL there are no licences or restrictions on future use or redistribution.

BSD-type licences allow individuals or organizations to make modifications or enhancements to the code without contributing that code back to the open-source community. Thus there is no requirement to release the derived work as OSS – which is why it is so often chosen.

There is not even a condition to name the originator in licence or advertising for the version that is being sold under commercial licence. The BSD licence text has a disclaimer as follows:

This software is provided by the copyright holders and contributors "as is" and any express or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall the copyright owner or contributors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the copyright notice, this list of conditions and the disclaimer.

- Redistributions in binary form must reproduce the copyright notice, this list of conditions and the disclaimer in the documentation and/or other materials provided with the distribution.
Neither the name of the originating organisation nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

BSD is the copyright model basis used by Apache and by the BSD-based operating systems projects (FreeBSD, OpenBSD, NetBSD). The restriction on advertising was officially rescinded by the Director of the Office of Technology Licensing of the University of California in 1999. Previously, certain of the BSD source code files required that further distributions of products containing all or portions of the software, acknowledge within their advertising materials that such products contained software developed by UC Berkeley and its contributors. With the changes, the BSD licence becomes equivalent to what is termed the MIT Licence, except for the no-endorsement final clause.

This lack of commercial constraints is the major upside for BSD – it is commercial developer and user friendly. The lack of viral concerns encourages use by the enterprise user, which in turn can spur both contributions and the development of a vibrant commercial support industry. BSD licences also allow vendors to package and sell enhanced versions of open-source as proprietary software. Thus we have two business sectors encouraged by this which are important for the European software industry – VARs and software support services.

For the original development community that created the open-source software under the auspices of open and free software, such forms of development can be seen as a downside. But these forms may well accelerate end-user and vendor support.

The other possible downside view of the BSD licence is that it allows forking, that is many different proprietary flavours of the software may evolve, which to a varied extent any open source license allows as all allow forking. Forking is letting a code base evolve in different (often partially incompatible) versions. It is not, technically, about changing the license. This can have the effect of dissipating resources and creative energy by scattering the focus of the collective efforts of the open-source community project involved. The phenomenon can be seen as the origin of much of the Unix Wars of the 1980s and early 1990’s between direct and indirect derivatives – AIX, UX, Ultrix, Xenix, NeXT, Solaris, etc and many more.

Moreover some consider the BSD licence as being good for large companies who better understand how to use it. It enables many competitors to make software proprietary and distribute it, under proprietary terms, and to make customers pay for it. Whether this would make its usage in commercialisation of grid software unsuitable for small companies (those we find in the EU) is debateable. The large vendors certainly understand the value in contributing to open source projects, but they also understand that, by doing so, they can undercut their smaller competitors. Thus some observers feel that a less permissive licence than the BSD, and the APL2 licence explored below, may be more optimally suited to the development of a European business ecosystem for grid.

**The Apache Public Licence 2 (APL2) in brief**

APL2 already governs the many software programme distributions from the Apache Software Foundation. It is a BSD-type licence and has become widely trusted and accepted by commercial vendors and enterprise users as its terms have been shaped to address the commercial market and for enterprise user adoption, such as grants of royalty-free patent

---

82 Value added reseller – typically a software house that takes in a piece of software and incorporates it into a product and so adds value. A typical example would an accounts package built on top of a database, which is bought in from a database specialist.
rights from any contributors. The licence is different to other BSD and BSD-type licences in that it spells out commercial usage rights quite specifically. It is thus highly regarded in the grid community as a platform for commercialisation. Some three years ago, the SAP legal department asked for certain changes, and these now appear in the current version, APL2. The key conditions, as an abridged summary, are:

**Grant of Copyright Licence** – a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright licence to reproduce, prepare derivative works of, publicly display, publicly perform, sublicense, and distribute the work and such derivative works in source or object form.

**Grant of Patent Licence** – each contributor in the OSS community grants a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable patent licence to make, have made, use, offer to sell, sell, import, and otherwise transfer the OSS programme (the Work), wherever patent claims are licensable. Interestingly, if patent litigation is instituted then the patent licences granted terminate on the date of litigation filing.

**Redistribution** – any party is free to reproduce and distribute copies of the Work or Derivative Works thereof, in any medium, with or without modifications, and in Source or Object form, provided that following conditions are met in that the party:

- must give any other recipients of the Work or Derivative Works a copy of this Licence;
- must have prominent notices stating that the concerned party has changed the files, if files are modified;
- must retain, in the Source form of any Derivative Works distributed, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and
- An information file or “NOTICE” may accompany the distribution. If the Work includes a NOTICE text file as part of its distribution, then any Derivative Works must include a readable copy of the attribution notices.

**Copyright statement** – modifications to the software may be made and new software added with additional or different licence terms and conditions for use, reproduction, or distribution of the modifications, or for any such Derivative Works.

**Submission of Contributions** – contribution submitted for inclusion in the Work by the Licensor shall be under the same terms and conditions for the APL2 Licence.

**Trademarks** – the licence does not grant permission to use the trade names, trademarks, service marks, or product names of the Licensor, except in the NOTICE file.

**Disclaimer of Warranty** – unless required by applicable law or agreed to in writing (most important for European countries with clauses on liability for death), the Apache OSS group, (the Licensor) provides the Work and each contributor provides its contributions on an "as is" basis, without warranties or conditions of any kind. Users must assume any risks associated with usage.

**Limitation of Liability** – in no event, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (again – most important for European countries with clauses on liability for death) or agreed to in writing, shall any contributor be liable for damages – direct, indirect, special, incidental, or consequential even if an OSS group contributor has been advised of the possibility of such damages.
Accepting Warranty or Additional Liability – while redistributing the Work or Derivative Works users may choose to offer, and charge a fee for, support services and products, warranty, indemnity, and other liability obligations. However, the user must agree to hold each OSS contributor harmless for any liability incurred and for any claims.

However there is an interesting community development point here for grid projects generally. When vendors build proprietary grid products on open-source, everyone benefits. While the APL2 licence does not require that vendors release modifications and extensions back to the community, the large vendors understand the ultimate value of contributing back because they have invested so heavily in grid middleware and services in the past, and do not wish their investments to die. Thus they will tend to continue to contribute in the future, bringing more functionality and higher quality to underlying open-source code, as well as competition between proprietary products for better solutions for the end user.

Mozilla Public License

In 1997 when Netscape decided to make its Communicator browser source code an OSS give-away, it needed to produce a licence between the permissive BSD licence and the GPL viral licence (for which any other code with which it is compiled must also come under the GPL).

The GPL conditions would have meant that any third party additions to the browser, to be bundled in a branded version of the code, would also have to become part of the original Netscape code, yet were out of Netscape’s reach and capacity. Moreover, Netscape wanted to use a portion of the code in its own proprietary products such as servers and so would have been compromised under the GPL conditions.

However, the BSD was deemed insufficient in that the original developers would lose control of future development and the company wanted an OSS community to control its future.

A licence was thus designed that avoided the pitfalls of each extremity and broke new ground at the time of the announcement (1998) in providing a compromise between free source development by commercial enterprise and protecting free source developers. The licence was opened to the same OSS process as code, being published for comment. The revised version, with clause on Netscape privileges on licensing removed was published in late March 1998. This Netscape Public Licence which became the MPL contained several key conditions:

- Source code licensed under it may be reused in commercial products
- Changes made directly in the source code – bug fixes or enhancements – must be contributed back to the Mozilla organisation (mozilla.org) but new source files which link to the code do not need to be (a crucial advance over GPL)
- The source of modifications be made available; this was designed to encourage code reuse and assure a vibrant community.
- The source of modifications must be shared under the open source licence; this is aimed at ensuring that the open source nature of the code remains open source.
- Freedom to distribute executables under a different licence; this was aimed to give business more flexibility in licensing derived works.

The distribution form is a files-based definition of modifications and software covered by the licence which makes it easier to combine the base code with other code to be distributed under different licences.

The MPL also contains an explicit patent licence. This makes it easier for businesses to use the software without concerns over patent rights. Provision for termination of rights in response to patent claims, helps to discourage patent litigation amongst the community – termed a “patent peace” provision.

All OSS licences obviate the liability of the original developers from effects of the code on systems, other software, or material or personal damages, except where statutory law mandates liability (eg as in the case of death in Europe). A typical clause which admits this condition is as follows – this is from the Mozilla public licence:

Under no circumstances and under no legal theory, whether tort (including negligence), contract, or otherwise, shall you, the initial developer, any other contributor, or any distributor of covered code, or any supplier of any of such parties, be liable to any person for any indirect, special, incidental, or consequential damages of any character including, without limitation, damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses, even if such party shall have been informed of the possibility of such damages. This limitation of liability shall not apply to liability for death or personal injury resulting from such party's negligence to the extent applicable law prohibits such limitation. Some jurisdictions do not allow the exclusion or limitation of incidental or consequential damages, so this exclusion and limitation may not apply to you.

The general difference between the major types of OSS licence are shown in Figure 4.1.

**Figure 4.1. The OSS licence cycle and IPR**
2.2. Do constraints in OSS restrict commercialisation or act as catalysts?

Most of the variations on BSD and the latest version of the Apache Licence, version 2, tend to act as catalysts for software production and commercialisation – generally due to their inherent commercial latitude. The licences are specifically designed to do this.

The one problem with any such freedom for the source code is the incorporation of software code in the original or in subsequent additions that is ‘infected’ by proprietary code having its own copyright, and in the USA, that which may be deemed coming under a software patent due to the ideas or working concepts employed behind the software.

*The risks from software patents and the US patent law on software*

Software patents “infect” code that is completely original, independent of whether it is open source or proprietary. Copyright protects expression, i.e. the code as written, while patents protect ideas, independent of expression. In the US, software patent lawsuits are now common. There is a theory that they are stopping innovation, as any line of code can be copyrighted there and so re-using it, even unknowingly is an infringement. In hearings before the US Federal Trade Commission in 2002, Robert Barr, Cisco general counsel, noted the US legal software system was in danger of destroying itself.84

Moreover support in the software industry itself is mixed. Although major software publishers such as Microsoft and users of embedded software such as Nokia support software patents, other software publishers such AB Strakt (of Gothenburg, Sweden) do not.

The European Parliament decision in September 2003 to ban all software patents opened the door to further innovation in software. This was emphasised by its 2005 rejection of software patents by a 95% majority in the European Parliament, possibly the sole European decision in which anything near unanimity has been achieved in the EP.

The problem with software patents and OSS is really twofold:

- If any lines of proprietary software, and therefore their IPR, become part of the OSS code, then the contamination opens developers and user to lawsuits. This is a general threat. Undisciplined developers in an OSS community project might implant such code, to the detriment of all.
- Secondly, commercial users could take OSS code and hi-jack it, claiming the ideas for their own as they have found them in a public licensed piece of code. Attempts have been made in ICTs before to take publicly developed and shared standards and patent them as their IPR, *ex post*, even key parts of the GSM public standards implemented in software.
- From a pragmatic European viewpoint, since Europe is in some form of global competition with other players where software patents are a reality, making software patentable in the EU would almost instantly bring a plethora of new patents held by actors from outside the EU, therefore potentially causing tremendous damage to the European software industry.

However the problem with software patents and OSS is wider. The roots of the issue are in the very conception of what software is. OSS proponents make the point that software is inherently different from material goods. On the other hand, patents are a legal mechanism made to protect the concepts behind material inventions. Due to its immaterial nature, software is thus essentially not open to the same commercial patent rivalry, the fact which

makes it possible to run open source projects *without* suffering the 'tragedy of the commons'. Software patents would introduce artificial scarcity. Innovation in software is typically incremental and gradual, which is well captured by the open source process, while the patenting system makes the assumption of discrete innovation, rare in the software arena.

Unlike the USA, patents on software per se are not recognised in Europe (and the majority vote in the European Parliament in 2005 confirmed this) – only on mechanisms, not the software itself which is covered by copyright. Patents in Europe, not the USA, are limited to those which display a novel creation proven to be a bona fide advance; exact definitions vary by EU member state – there is no common single European patent currently. The whole question of patents in the USA is currently under consideration for revision as the major US systems vendors (led by IBM, HP, Sun, Cisco, Apple) realise that the USA’s adversarial system of software patents, business process patents and injunction laws is halting software and IT advance and is a threat to future economic progress. In June 2005, these companies met with the US Department of Justice and the Federal Trade Commission to discuss this and in particular the problem of patents in open standards. The latter is a further vexed subject. Disclosure of patents concerning open and public standards after their publication has been the subject of conflict as owners of the relevant IPR collect patent royalties *ex-post*, that is after their incorporation in software and material products. However these same companies are in a severe quandary as some are turning out up to 3000 new patents per month – many of them software patents.

OSS helps to alleviate this question somewhat in that all IPR while being copyrighted, can be read in the source code (being open source) and use of *patented* content tends to be specifically and vigilantly militated against, especially in view of the risks in recent court cases. Thus, in this sense OSS is a major catalyst for commercial uptake as the original code is likely to have been examined closely for patent infringements. This does not mean that a further certification of cleanliness from contaminating proprietary IPR is not a real need for commercial uptake.

Interestingly on 01 May 2006, IBM held an analyst seminar in the USA on the whole question of IPR in an ITC setting to air proposals on a far more open approach, probably before going to Washington to lobby the US government.

*Which OSS licences are less suitable for grid commercialisation*

In contrast to the BSD-style licences, the GPL and LGPL might be considered hold some constraints or liability for commercial development. There are also grey areas with GPL-style licences. For instance, although system calls from other applications, be they proprietary or other OSS, may legally be passed to a GPL-licensed operating system, adding to or changing the operating system, in the sense of adding parts of the kernel (the ‘heart’ of the operating system) or even perhaps adding device drivers outside the kernel, would probably contravene the GPL licence.

*Solving the problems of licence contract conflicts at run-time*

What we need is a licence situation whereby any application can run with a grid contract on open middleware such that resource providers can offer computing elements, perhaps charged for (or in reciprocal usage). What is needed is a method for discovering and resolving conflicts before or at the run time of the grid implementation. Such solutions do exist in perhaps limited form today compared with requirements of a large grids spread across many applications and several organisations.

---

85 From discussion in HPC4U interview.
One commercial example is Platform Computing a leader in proprietary grid computing, largely for clusters in a single organisation. The company has an agreement with the Macrovision Corporation, who offer dynamic an intelligent licence discovery and conflict resolution product, newly ported to grid computing environments. Platform and Macrovision are working together to deliver a utility to optimise licence scheduling within one enterprise, with usage reporting and analytics. The aim is to remove the licence conflicts barrier to widespread grid adoption. Integration of Platform and Macrovision’s dynamic licence management product is based on policy-based control for the grid environment licence management.

2.3. What risks do licences imply for the developer, code modifiers, integrators and end users?

Proprietary licences hold specific problems and risks:

- For the original developer – the licence assures preferential conditions, not risks and obviates responsibility for malfunction. However, there are signs that market views are changing so this could evolve, especially for market dominant software publishers.
- For subsequent code modifiers – commercial and non-commercial (OSS community) – modification is forbidden – they will be sued by the software publisher for licence infringement unless covered by separate agreement
- For system integrators – the risk is that they may be sued by the software publisher for any licence infringements
- For end users (who only use) – may be sued by the software publisher for any licence infringements, for instance, use on the wrong class of server or by more users than licensed; defects may be inherent but user licences prevent any legal recourse (except in case of death in the EU); repairs to defects are at the vendor’s discretion as to if and when they are fixed.

Each of the OSS licences implies freedom from liability to distribute, use, modify, add, read the source code, but with different types of freedoms on the amended code for:

- The original developer – the source can be read and used by anyone; the licence protects the authors from any liability for this usage, except in the EU for causing death, and from the subsequent risk of ex-post patent actions (whereby a third party takes the code and patents it as his) as happens in open standards and in OSS.  
- For subsequent code modifiers and VARs – commercial and non-commercial (OSS community) the risks are in infection of the code (see below), lack of support.
- For system integrators – the same risks as for VARs and subsequent developers, plus the interfacing risks; however due to the insistence on open standards by many OSS communities, the problems are generally far less than with propriety interfaces and data formats.
- For end users (who only use) – there may be little support in the sense of regular services for fixes to defects, for installation and for updates. However analysis of

---

86 Platform Computing and Macrovision press release, 22 June 2005 – their combined solution comprises Platform LSF License Scheduler and Platform LSF Analytics products, integrated with Macrovision’s FLEXnet Manager product through the FLEXnet Connector – Macrovision has now opened its FLEXnet platform to other software providers and enterprises.

87 The Federal Trade Commission in the USA is investigating whether Rambus hijacked the setting of open standards for computer memory in that over four years it directed standards into areas of DDR and SDRAM in which it subsequently pronounced patents in the JEDEC industry standards organisation; Chris Nuttal, ‘Rambus hit as FTC rules’; Financial Times, 3 August 2006.
the defect repair times with an active and agile OSS community has shown they may be better with OSS than proprietary software packages.\textsuperscript{88}

In general the additional risks for all types of users, publishers of OSS and service providers offering a service based on OSS may be summarised as:

- **Conflicts of different licence types** – the grid environment is more likely to be heterogeneous and so the conflicts of application licences (especially proprietary, and on sizing restrictions) and what resources can be used with what, also what software can be used in conjunction may be unclear. OSS code may even be a chain of OSS licences in that code today is often incremental – built on other OSS. OSS licenses are either compatible or not, i.e. chaining is either possible or not, and licenses do not usually chain in the sense required.

- **Warranties** – the above problem is exacerbated for proprietary warranties on code working as agreed in the contract, as it may only go so far in a real working situation mixing licence types. The OSS community will not give any warranty. However an OSS service provider, as in the Red Hat Business model, will. But in general the limits of proprietary warranties will be unclear if used with an OSS unless specific clauses are added.

- **Infection** – contamination by proprietary lines of code (it is usually illegal even to see the proprietary source) or by OSS source code from another OSS community which has GPL conditions over combinations with other code. Fortunately some tools are now available to automate intellectual property management and licence compliance functions\textsuperscript{89} and evaluate whether a licence involved is critical or not. The infection risk also extends to users. Measures for code inspection must be taken to give confidence to users that the code is free from contamination from proprietary code and software patents threats that could reflect on them as users.

- **Viral risks with GPL and LGPL** – the limits of their applicability are unclear; for instance, a piece of software may be able to make system calls to a GPL operating systems, but addition of a device driver by a third party might be considered an infringement, unless it is contributed to the common OSS repository.

One further query is whether OSS licences effectively undermine intellectual property in a commercial sense, in that they can only form a commons. As such, they could act as a restriction for grid commercialisation. This is evidently not so with the BSD-style licences, whereby an OSS product can be used to develop a private and proprietary product with its own IPR, and the information and IPR remain confidential and proprietary if required. What cannot be done is to take OSS under a free-to-use but with copyleft conditions in a way which is forbidden, by law, but this is true for all software, not just OSS, and especially for proprietary software.

There are also questions as to whether proprietary software gives more legal rights than OSS. Some have commented that with OSS, any rights to sue if things go wrong are removed. However, essentially all proprietary software licences have always taken no liability for malfunction and many proprietary licences even forbid lawsuits as a condition of usage. Thus today it is not possible to sue the shrink-wrap proprietary package vendors when things go wrong as it is difficult to assess the software defects and licence conditions may prevent any legal process. The Warranties should handle this but are less available in

\textsuperscript{88} There is work on the micro-economic impacts of OSS pointing to productivity increase for the user. See for example, S. Forge, ‘The rain forest and the rock garden: the economic impacts of open source software’, info, Vol 8, No 3, 2006.

\textsuperscript{89} For example, Black Duck Software provides software compliance management products; they also have a plug-in for their protexIPTM suite for the popular OSS toolset, the Eclipse Platform
the USA, although expected by users in the EU. US software publishers get over this by having the European software distributor take over warranty liabilities.

Note that there are differences over licence disputes and the laws applying in the EU and the rest of the world. In the USA, for instance, there are greater risks for the end users of software and developers infringing the licences of other IPR holders. The US system turns on case law and punitive damages in these disputes, while the EU system is based on the actual damage done and its value only, seen within a statutory framework for interpretation.

As has been noted with open-source software licences and the source code, users, in principle, have more information so that:

- The community of users can inspect the code for flaws and publicly exchange knowledge about such flaws.
- Users may create derivative versions of the code that repair defects and hazards that the original source author did not find.
- Users can determine whether the program contains adequate safeguards against risks of safety or security.
- If the software goes fails in any way the user can fix it or pay to have it fixed; this option greatly reduces risk, and obviously does not exist for proprietary software.

**Comparisons of risks with commercial licences**

Commercial licences have specific clauses on rights of usage that vary by software publisher and product but generally restrict the types of usage, the numbers of users and the rights to copy, to modify and to redistribute. The conditions are enforced under copyright laws and may also be protected by patents laws in the USA, where patents apply to software and also to business methods (not in the EU). Thus copying a proprietary operating system is illegal copyright infringement or ‘piracy’.

Most importantly, the vast majority of proprietary software licences specifically avoid liability for their own defects or for legal pursuit in case of malfunctions. The licences also define the conditions of various services for post-sales support which can be the most important advantage of the commercial licence to the end-user organisation, but can also be the most costly – a basic low response telephone support service may be 10-15% of the licence cost, annually, but for full 24x7 4-hour response and best effort fix, the annual licence cost may be 20-30% or more. Telephone support may not be a fixed charge however and may be on a per minute basis using premium phone services, so the user is relatively unaware of the charges until the phone bill is delivered. It is thus highly attractive to a proprietary vendor to issue a successfully working first version of a software package to build a large installed customer base, and then issue a second edition, an ‘updated’ version, which has many bugs. The profits from telephone support for a large consumer base can be significant, especially in the case of the second edition being a base utility, such as an operating system, or a common application.

In contrast to OSS, proprietary software vendors in offering closed-source code, hide the condition that it could well contain defects and even hazards. These are of course unknown to the user as the source code is a commercial secret, and even reading it can be an offence. In a legal sense these could be attributable to corporate negligence – but as the code is not

---

90 See ‘A senior Microsoft attorney looks at open-source licensing’, where Bryan Pfaffenberger argues that more transparency for the user is available with OSS, from: http://www.linuxjournal.com/article.php?sid=5073
available, or visible, let alone modifiable, the user knows far less about what he or she is getting.

A further legal difference not always mentioned is that many proprietary software licences require that users will permit software licence audits, on demand – and that they are liable for severe liquid damages if the organization is unable to prove that every type of usage, instance of use, platform of usage and user is licensed. Thus, in certain situations, use of proprietary software can open the user to the risks of liability for misuse, enforced by legal pursuit by the vendor.

This is perhaps the biggest legal difference between OSS and proprietary licences.

2.4. **What are the impacts of licences on commercialisation and business models for grid products and services?**

From the above, we can surmise that the market impacts of licences are really sevenfold, and they act on software suites or products:

- Control of IPR and whether royalties must be paid.
- Control whether the source code can be viewed openly, changed and added to – the grant of rights to act on the code, be it in changes or additions to the program.
- Control of further development – does any source code based on the original have to be returned to the public repository, under the same open conditions, or can it be kept as a derived work within the contributor’s or VAR's commercial IP conditions. This covers the grant of rights for licensed patents, meaning patent claims licensable by a contributor which are necessarily infringed by the use or sale of its contribution alone or when combined with the original program.
- Control how used – how many users, server sizes and numbers, duration of usage, configuration of usage (eg internal to one company or in web accessible situation – a public open to all) – as in a commercial licence, and paying usage fees.
- Control of what can be put with it – whether extra libraries may be used or additions can be made to the source code and it can be run in any environment, physical or virtual – crucial for grid.
- Supply of support services to install, find, fix and repair defects and produce upgrades and resolve integration problems with other software.
- Warranties and liabilities for working in the manner and under the conditions as marketed and perhaps as contracted (less likely in the USA) and also for behaviour in cases of failure. The licence type also reflects on whether the original authors, or VARs in some sense, take these responsibilities. This area also includes performance guarantees- usually under SLAs.

All of the above, as a consequence, control the uptake of the product.

In summary, the net market impact of OSS is to open the market for the European software industry’s VARs, ISVs who are using OSS code and system integrators, far more so than if purely proprietary licences were used. OSS licences could, if carefully used, build the industry for grid software in applications as well as middleware and attached functional utility modules.

**Grid services – impacts of licences on the commercialisation and business models**

OSS licences have only influence over the sale of a service in that they provide conditions of operation of the underlying products that support the services. They give enormous
liberty compared to proprietary software in terms of both numbers of users, and forms of usage, that is in resale of services. The APL2 licence maximizes the effect. However those OSS licences which are restricted to “not for profit” usage’s are obviously inappropriate for a commercial offering.

Proprietary licences however may have more far reaching impacts in terms of limits, in that they closely control the way in which the underlying software product is used – for instance-how many users it can have. Normally a licence is for a strict number of end-users, machines and perhaps sizes of machine. There can also be a charge based on numbers of instances of use or sessions. Thus the openness of OSS licences can be a major advantage for the services.

The conclusion is that for the offering of grid computing as a service in some form (be it at application, or computing or infrastructure level) OSS licences in their free market form, of Eric Raymond’s Bazaar can offer major advances over the proprietary licence with its restrictions on usage’s, particularly for resale with profits.

3. Choosing an OSS licence

3.1. Comparison of the different licence types

The background to a grid project is the starting point for the right choice of licence – what is the overall aim from any delivery from the project.

Thus it is useful to analyse the most commonly used and also the most appropriate licensing models for grid software within an EU context, using a SWOT analysis. Summaries of common licence types by this approach are shown below, as well as for the proprietary licence.
The GPL as noted above may require more scrutiny for some commercial ventures:

**GPL: Comparison of the different licence types using SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation**

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open in free community sense of commons property so users are free to:-&lt;br&gt;• run the program, for any purpose (freedom 0)&lt;br&gt;• study how the program works, and adapt it as needed (freedom 1). Access to the source code is a precondition for this&lt;br&gt;• redistribute copies (freedom 2)&lt;br&gt;• improve the program, and release improvements to the public, so that the whole community benefits. (freedom 3). Access to the source code is a precondition for this&lt;br&gt;• Enables maintenance of open commons project&lt;br&gt;• Authors protected on liabilities</td>
<td>Limited commercially by need to make all derivative works similarly under GPL&lt;br&gt;• Those who contravene its requirements for return of developments have to be pursued legally as many do not make the effort; eg found to be the case with some Taiwanese developers of drivers&lt;br&gt;• Difficult for commercial vendor to offer warranties and support on total offer of own code and GPL code</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allows any development&lt;br&gt;• Commons property – brings development that would not happen commercially (Linux is under GPL2)&lt;br&gt;• Users can do whatever they want with code in an open development community</td>
<td>Viral nature – any suite of software containing a GPL unit may be affected by the licence in terms of what other units can be used with it&lt;br&gt;• Developments of suites when mixed must be checked to ensure no contraventions of licence conditions</td>
</tr>
</tbody>
</table>
The lesser or library LGPL suffers from the same reciprocal character in commercial exploitation:

**LGPL**: Comparison of the different licence types using SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Has useful features for library modules that may be used for many different software licences</td>
<td>• Still has limited commercial appeal due to need to make all derivative works similarly available</td>
</tr>
<tr>
<td>• Authors protected on liabilities</td>
<td>• Difficult for commercial vendor to offer warranties and support on total offer of own code and LGPL code</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Building library modules for use with many different software licences</td>
<td>• Viral nature – any suite of software containing a LGPL unit may be affected by the licence in terms of what other units can be used with it and possibly for-profit uses, depending on specific conditions</td>
</tr>
<tr>
<td></td>
<td>• Developments of suites when mixed may contravene the GPL licence if further developments are not returned to the source repository</td>
</tr>
</tbody>
</table>
In strong contrast, the BSD is completely open for commercial usages:

**BSD: Comparison of the different licence types using a SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation**

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Freedom to produce derivative works without restraint and distribute them as own works with own IPR</td>
<td>• Unsuited to an open source community which wishes to control the production of derivative works to be all in a common repository with no individual or commercial IPR</td>
</tr>
<tr>
<td>• Freedom to change anything or to make additions</td>
<td></td>
</tr>
<tr>
<td>• Freedom to retain developments as new commercial copyrighted IPR</td>
<td></td>
</tr>
<tr>
<td>• Freedom to mix in development with any other licence and use with any other software</td>
<td></td>
</tr>
<tr>
<td>• Authors protected from any liability for usages</td>
<td></td>
</tr>
<tr>
<td>• Commercial vendor can offer warranties on all code</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Offers commercially attractive conditions for derivative works with no restrictions</td>
<td>• Unlimited commercial derivatives from forking the original could be seen as threat to a community wishing to concentrate on one version – ‘the Unix effect’ of many different (commercial) versions of the same programme</td>
</tr>
<tr>
<td>• Products under this licence suitable for freedoms of usage required for a grid service</td>
<td></td>
</tr>
</tbody>
</table>
Ideally, for grid, the licence for commercialisation may well be the Apache public licence, v2:

### APL 2 Apache v2
Comparison of the different licence types using SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation

<table>
<thead>
<tr>
<th><strong>Strengths</strong></th>
<th><strong>Weaknesses</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Clear exposition of rights to commercialise derived works</td>
<td>• Less suitable for Public only source repositories, for an academic project as specifically aimed at commercial market freedoms</td>
</tr>
<tr>
<td>• Freedom to produce derivative works without restraint</td>
<td>• Few from commercialisation perspective</td>
</tr>
<tr>
<td>• Freedom to change code/ add new source code</td>
<td></td>
</tr>
<tr>
<td>• Freedom to retain developments as new commercial copyrighted IPR</td>
<td></td>
</tr>
<tr>
<td>• Freedom to mix in development with any other licence and use with any other software</td>
<td></td>
</tr>
<tr>
<td>• Protection from patent hijack and litigation for users</td>
<td></td>
</tr>
<tr>
<td>• Original authors protected from any liability for usages</td>
<td></td>
</tr>
<tr>
<td>• Commercial vendor can offer warranties on all code</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Opportunities</strong></th>
<th><strong>Threats</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Encourage free development from OSS original works</td>
<td>• Not suitable for OSS communities that wish to control their code</td>
</tr>
<tr>
<td>• Seed new ventures, enterprises and industries based on its source code inputs</td>
<td>• Few or none for commercialisation ventures</td>
</tr>
<tr>
<td>• Seed new services offering based on APL2 products</td>
<td></td>
</tr>
</tbody>
</table>
MPL offers an alternative to the APL2 for OSS communities that wish to retain control yet offer commercial ventures their OSS software for productisation:

**Mozilla Public Licence:** Comparison of the different licence types using SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Source may be reused in commercial products</td>
<td>• Most suitable really for the type of project with and active and responsible OSS developer community</td>
</tr>
<tr>
<td>• For OSS community- changes to the source - bug fixes or enhancements - must be contributed back to the original author organisation but new source files which link to the code do not need to be; helps encourage code reuse and a assure a vibrant community.</td>
<td></td>
</tr>
<tr>
<td>• Requirement that the source of modifications be shared under the open source license; helps encourage open source remaining open source</td>
<td></td>
</tr>
<tr>
<td>• Authors protected from any liability for usages</td>
<td></td>
</tr>
<tr>
<td>• Commercial vendor can offer warranties on all code</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Ability to distribute executables under a different license; gives business more flexibility in licensing derived works.</td>
<td>• For purely commercial developer – enhancements become contributions to public repository and thus any IPR is lost</td>
</tr>
<tr>
<td>• A &quot;files based&quot; definition of modifications and covered software; makes it easy to combine code with code distributed under different licenses.</td>
<td>• Unsuitable for purely public academic projects</td>
</tr>
<tr>
<td>• Explicit patent license; makes it easier for businesses to use the software without concerns over patent rights.</td>
<td></td>
</tr>
<tr>
<td>• Provision for termination of rights in response to patent claims (i.e. &quot;patent peace&quot; provision); helps discourage patent litigation amongst the community,</td>
<td></td>
</tr>
</tbody>
</table>
The proprietary licences are not positioned so favorably as OSS licences for a grid industry:

### Proprietary Grid Licences: Comparison of the different licence types using SWOT analysis (strengths, weaknesses opportunities, threats) for Grid Commercialisation

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Usually support services are available under well-understood conditions</td>
<td>• Requires strong channel marketing for grid, perhaps via strategic alliances in products/services to get take-up in ISVs, VARs, System Integrators who must accept its commercial conditions (eg royalties) - and would prefer an OSS licence</td>
</tr>
<tr>
<td>• Responsibility for total code is clear, although liabilities and warranties only defined in EU</td>
<td>• Grid technology tends to require heterogeneous environments so licence must cope with this in a context of multiple different licences</td>
</tr>
<tr>
<td>• Publisher controls directions of development</td>
<td>• For users – defects not revealed and no inspection or self-fix possible – wait on vendor to repair</td>
</tr>
<tr>
<td>• Publisher controls commercial exploitation by VARs etc</td>
<td>• For users – may have restrictions on users, usages, etc which are difficult to comply with in a grid context</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Easier for businesses to use the software with fewer concerns over contaminating code and patent rights.</td>
<td>• For users – may be sued for infringements</td>
</tr>
<tr>
<td>• For code publisher - can sell software and support services in the normal package</td>
<td>• For product authors – may not be taken up widely as proprietary so product dies as outside the standards</td>
</tr>
<tr>
<td></td>
<td>• Proprietary constraints likely problem for services (grid infrastructure/grid computing service/application service) if users and usages restricted and cannot cope with heterogeneous resources, i.e. only closed infrastructure</td>
</tr>
</tbody>
</table>

**Or do we need a new OSS licence for grid?**

We also have to note that a case for novel alternatives to the above existing licences could be made. Some observers have stated that there are already too many OSS licences. However the option of creating one, or a set of new open source licence(s) specifically designed to fit the requirements of promoting Grid technologies could be a viable way forward.

Such a licence may be explicitly be designed so as to give legal value to an automatic license discovery and management system based on machine readable encoding of service contracts/software licenses.
For one major project, the Globus alliance middleware product, the Globus toolkit, we examine the path to licence choice in more detail in Box 1.1.

### Box 4.1. The Globus Toolkit choice of licence: how the decision was made

In 1995, the open-source software project began for the Globus Toolkit. The original aim was to tie together geographically dispersed compute resources in the USA, and federally funded research organizations within US, for resource sharing and problem solving. This would be for a closed community, of academic and publicly funded research, with no commercial activities. As the project progressed, a small group of technologists developed a grid middleware suite that eventually became the Globus Toolkit (today is available as GT-4; new version in Beta test).

However over the years, the Globus Toolkit code evolved. The pool of contributing developers expanded beyond the USA to include a large international community. This is the sign of success for a widely adopted open-source effort, but as the code matured and became more widely utilised, new constituencies entered and so the community grew in many different directions. So Globus management had to examine a range of open-source software licences to select a best fit to the interests of the growing Globus Toolkit community at large.

Instead of the original small community of academic technologists who shared one common goal, development had expanded into a large number of distributed groups. Conflicting priorities and visions emerged, as grid usage was strongest in research and academia for big compute science requirements, such as particle physics, but the Globus Toolkit was seeing strong participation from a number of other areas as well, in commercial applications. In the financial services industry, beta-testing the toolkit had started for commercial applications. Moreover vendors were contributing funding and code and experimenting with commercial-level grid products built on top of the open-source code. It was at this point, that Globus project management analysed open-source licences for a best fit to the interests of the whole user community. Among the open-source licences available, the two seen as predominant were the BSD and GPL licences. Each has a specific bent – the BSD is more towards Eric Raymond’s ideal of the bazaar-mode of development, as given in his paper ‘The Cathedral and the Bazaar’ of freely usable commercial licence, but both permit free use of software.

For these reasons a BSD-style licence was preferable, so that vendor incorporation of Globus Toolkit code into commercial proprietary products was permitted to accelerate its market take-up. This was an essential factor for choice – acceptance in the commercial market. In 2005, the Globus Alliance announced adoption of the BSD-style Apache Licence Version 2.0 for the Globus Toolkit. APL2 enables commercial grid products to employ the Globus Toolkit open standards. Some early implementations of the standards are in IBM’s Grid Toolbox, Sun Microsystems Inc.’s Grid Engine and Nortek Networks’ Dynamic Resource Allocation Controller (developed as a predecessor to Platform Computing’s LFS). Consequently, such products are able to interoperate with other hardware and software resources in their customers’ IT environments.

---

93 [http://www.opensource.org/licenses/bsd-license.php](http://www.opensource.org/licenses/bsd-license.php)
94 [http://www.opensource.org/licenses/gpl-license.php](http://www.opensource.org/licenses/gpl-license.php)
95 See Chris DiBona et al (eds), [Open Sources, Voices from the Open Source Revolution](http://www.opensource.org/licenses/apache2.0.php), O’Reilly, 1999.
96 [http://www.opensource.org/licenses/apache2.0.php](http://www.opensource.org/licenses/apache2.0.php)
3.2. Matching licences to project and industrial conditions

How to pick an open source licence really comes down to what is the intention of the project or venture that it will serve, in terms of commercial exploitation and degree of commercial freedom for any derivative work. Table 4.3 explores the options for each licence and its intentions as a software exploitation scenario.

Table 4.3. Comparison on an industrial usage basis of the positions of the various licences

<table>
<thead>
<tr>
<th>Licence Type</th>
<th>Software Exploitation scenario</th>
<th>Usage Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPL</td>
<td>Academic R&amp;D projects that will remain as e-Science and will never produce commercial offerings</td>
<td>Public availability; public development and some controls on combining the program code and assuring distribution of source</td>
</tr>
<tr>
<td>LGPL – lesser GPL</td>
<td>Academic R&amp;D projects that will remain as e-Science, unlikely to produce complete commercial offerings. May offer some components to others for commercial offerings but under restricted usage’s in ‘for-profit’ situations</td>
<td>Libraries that need to be combined with OSS software and may have different origins and may not allow GPL conditions.</td>
</tr>
<tr>
<td>MPL – Mozilla</td>
<td>Controlled public development, in which the OSS community sets the direction and retains control over the code but wants the code to be used in commercial products. Aim is also to build a strong OSS support community while building its bridges to the commercial world for derived works which can be commercial. Avoidance of patent claims and disputes over the IPR.</td>
<td>Public availability, of source code Source code may be reused in commercial products Ease of combination of the original source code with other code to be distributed under different licences Freedom to distribute executables under a different licence; gives business flexibility in licensing derived works Changes made directly in the source code – bug fixes or enhancements – must be contributed back to the OSS organisation except for new source files which link to the original code Easy for businesses to use without concerns over patent rights</td>
</tr>
<tr>
<td>Apache Version 2, APL2</td>
<td>Driving development of derived software products for freely used commercial markets Useful for grid as services</td>
<td>Freedom to exploit derived works commercially Does not require that vendors release modifications and extensions back to the community</td>
</tr>
<tr>
<td>BSD</td>
<td>Derived software products with no control over their subsequent usage or declaration of their origins. Useful for grid as a service</td>
<td>No control over derived works No acknowledgement of origins</td>
</tr>
</tbody>
</table>
4. What are the overall impacts of OSS licences on commercialisation?

4.1. Links between an open source reference implementation and standardisation

Here we assess any correlation or causal links between the availability of an open source reference implementation and de facto / industry standardisation, and the strength of the link if it exists. The use of a reference implementation, particularly its licence model, is of key interest. Note that in the chapter on drivers and barriers we further examine the overall market potential and the form of an OSS grid reference model and a possible form of implementation.

One approach to an assessment is to look at previous major impact open ICT examples:

- In operating systems – Linux has set a model for free distribution using GPL and has multiple commercial forms of support services bundled with the deployment packs for the code, from the likes of Red Hat, SuSE etc. Moreover the Linux reference model has spawned a series of different versions of the operating system for specific contexts – eg for mobile handsets, and for real-time embedded systems with limited memory, resources, etc.

- The original WWW and Internet protocols were based on reference models which were open and commented on freely with support communities set up at a professional level to act as developers and guardians of the reference model and its key elements and protocols (the W3C and the IETF for instance).

- In software tools – the Eclipse toolset, originally from IBM, set a standard for open usage with its freedom of usage, using a GPL type licence. The toolset has become a de facto standard in the Java development world. Tools are now built to clip on to it such as licence checkers.

- In mobile technology, the original digital working for GSM was set by a reference model that built on previous digital systems and the preceding analogue model of cellular operation. It provided the basis for the set of standards that have become the GSM industry today.

In each case we see that the reference model has been the foundation of the set of standards which became a de facto industry standard – and perhaps a universally accepted public standard. This has spawned whole industries in services and products in the cases of the Internet/WWW and GSM mobile.

Most importantly, the licences and copyrights embedded in the reference model were open. In this way the protocols of the Internet superseded the proprietary protocols of the most powerful commercial data networking architectures at the time of introduction (IBM’s SNA, DEC’s DECnet, as well as lesser offerings from HP, Novell, etc). Interestingly in effectively giving away NFS (Network File System) in the mid 1980’s, Sun established a de facto standard for workstation and server interworking, with its reference model of a client-server architecture which has propelled the whole workstation market. However we note that openness can embrace the closed – the original IETF RFC working mode was based on multiple independent but interoperable implementations, some of which usually were proprietary/closed.

Also note that in the case of GSM, although some of the GSM technology for manufacturers – the implementation of the reference model in products – has come under patents, a patents
pool was set up and the reference model clearly shows the position and scope of these patents.

In consequence, the reference model, with its licence, needs to be a long-lasting initiative, and not just bounded to a single (or a series of) project(s), such as FP6 or FP7. It should be supported at a European level, possibly taking into account and harmonising on-going efforts at national level.

4.2. The linkage between the business models and the licensing scheme

If we go back to the business model analysis chapter, it is interesting to see the linkage from business model to licence type, especially from the point of view of research projects entering this area. There are several key tenets in this process for research projects:

- Start with the end in mind – what is the ultimate goal of the consortium/project?
- Then select a licence by the 'horses for courses' approach and following the SWOT analysis (see previous section above)
- Review code elements contributed to see if any restrictions

Note that there are no hard rules - such as BSD-types are good for all business situations while GPL-types are always bad – it just depends on the wishes of the OSS community behind the software. For instance, GPL-type licences are used for the commercially successful Linux, MySQL and Eclipse tools. In these cases, derivative versions are not sought and one single standard is the intention, with close attention to quality.

We should also distinguish between licences for research projects and licences for those who may take up the code as a commercial venture. That is we need to distinguish between the OSS licence for IPR handover and the final licence although these may be linked through the licence conditions as shown in Figure 4.2.

**Figure 4.2. Commercial v research project licence objectives**

![Diagram showing the linkage between business models and licensing schemes](image_url)
Thus there is a research project licence to enable and encourage each intended Business model. We can map from the eventual model that the project wants to encourage to the licence type, as shown in Table 4.4.

Table 4.4. Linking business models and licence types

<table>
<thead>
<tr>
<th>Business model to encourage</th>
<th>Licence type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid Applications Product</td>
<td>MPL or APL2 as allows combinations with any other software. APL2 for unfettered re-use of code in commercial products is often chosen. For dual licence usage the MySQL model encourages an open source version and a commercial product, but that is more for a company the takes up the source code form the research project. For a commercial user, a simple commercial licence may well be the norm, possibly with an OSS version for marketing purposes.</td>
</tr>
<tr>
<td>Grid Compute Service</td>
<td>Any OSS licence gives the user freedom of use; providers may prefer MPL, as it allows usage in any environment, or GPL types for stronger code control, such as LGPL. But commercially freeing licences are just as possible, such as APL2 or the LGPL –type of JBoss.</td>
</tr>
<tr>
<td>Grid Support Services</td>
<td>Application of the Red Hat model – that is – selling support for a grid product can be carried out with any licence type. The service provider is only concerned with assuring operational continuity, updates and performance, not the product sale itself, exception as a free distribution as part of a services package, which includes telephone help-lines etc.</td>
</tr>
<tr>
<td>Grid Application As Service</td>
<td>To retain development control so that application service providers and users are free to use the software without restrictions on combining with other code, running in any environment, MPL may be useful, or perhaps a GPL-type if there are heavier restrictions required. An alternative is APL2 if the code is intended to seed independent commercial products, and possibly set some form of de facto standard in some way (eg as in 3D design tools). The commercial ISV may well use a simple commercial licence, as CD02 does.</td>
</tr>
<tr>
<td>Grid Middleware Product</td>
<td>MPL if require control over development, APL2 if not, for free incorporation in commercial products. One example is the Globus Toolkit, using APL2.</td>
</tr>
<tr>
<td>E-Science Project</td>
<td>Depending on control required over further development – high control GPL types and no mixing with other licences, with LGPL for some mixing but not linking, else MPL.</td>
</tr>
</tbody>
</table>

For further consideration are the equivalent European-grown licences, perhaps less well-known but viable, such as the EUPL, and also CECILL from France and compliant with French and international legal frameworks, and the QPL. The question is also of combining existing licence provisions into a bespoke licence for grid commercialisation. However some would argue that are already too many licences and so why not just choose the best suited.
How licences were chosen by various projects

It is also interesting here to look at the reasons behind choosing a licence by the major research projects. This is often associated with their background history which tends to set future goals. The motives in choosing a licence are thus various and depend on the intentions on usage’s by the projects’ original authors and founders. Table 4.5 gives some examples with their background so position and motives can be understood below major grid projects:

### Table 4.5. Projects’ selections of licences types, set in context

<table>
<thead>
<tr>
<th>Project</th>
<th>Licence choice and reasons</th>
</tr>
</thead>
</table>
| The Globus Alliance | The Globus Alliance is an association dedicated to developing fundamental technologies needed to build grid computing infrastructures. It is a largely US-based community of organizations and individuals developing fundamental technologies behind the “Grid”. Sponsors are US National Science Foundation, and the US Dept of Energy. The Globus project grew out of the US government support for academic and government research requiring large amounts of computing power. Its major deliverable is the Globus Toolkit (GT4, currently in its 4th version) an open source software toolkit for building grid systems and applications that grew from the needs of e-Science. It is being developed by the Globus Alliance and many others worldwide. Current core members of the Globus Alliance team:- Argonne National Laboratory, University of Chicago; National Center for Supercomputing Applications (NCSA); Northern Illinois University, High Performance Computing Laboratory; University of Southern California, Information Sciences Institute; Univa Corporation – a private company: sells GT-4 in a Red Hat style business model. Overseas participants include:- EPCC, University of Edinburgh; Royal Institute of Technology, Sweden.  

The aims are to let people share computing power, databases, instruments, and other on-line tools securely across corporate, institutional, and geographic boundaries without sacrificing local autonomy. A growing number of projects and companies are using the Globus Toolkit to build their grids, including IBM and other systems/software vendors. Globus Alliance’s ongoing commitment is to open source software and its open source community. All Globus software is developed by a developer community, distributed under a liberal open source licence, and applied by users around the world.  

The Alliance chose APL2 – as a BSD-style licence - because vendor incorporation of Globus Toolkit code into their proprietary offerings is a key to enterprise adoption. The choice of an APL2 licence allows vendors to use Globus Toolkit implementations of the open standards in their grid products. IBM’s Grid Toolbox, Sun’s Grid Engine and Nortek Networks’ Dynamic Resource Allocation Controller are early commercial grid products using the Globus Toolkit. Such products are able to interoperate with other hardware and software resources in their customers’ IT environments with APL2. |
| UNICORE grid middleware (UNiform Interface to COmputing REsources) | UNICORE is grid computing middleware technology for seamless, secure, and intuitive access to distributed grid resources such as supercomputers or cluster systems and information stored in databases. UNICORE was developed in two projects funded by the German ministry for education and research (BMBF). In various European-funded projects UNICORE has evolved to a full-grown and well-tested grid middleware system over the years. UNICORE is used in daily production at several supercomputer |
centers world-wide. The architecture of UNICORE consists of three layers - user, server, and target system tier. The security within UNICORE relies on the usage of permanent X.509 certificates issued by a trusted Certification Authority (CA). Beyond this production usage, UNICORE serves as a solid basis in many European and international research projects. GPL and LGPL were considered too restrictive for business – so ‘non reciprocal’ licences are preferred. The project has chosen BSD variants with code on SourceForge, so that others may use the code freely. UNICORE technology is published as new releases on a regular basis.

When the UNICORE development was started as a project funded by the German Ministry for Research and Technology (BMFT now BMBF) the then valid funding conditions expected that the software would eventually be made commercially available by one of the partners (Pallas at that time). This required closed source and non-disclosure agreements between the partners. The result of not being Open Source seen as a major disadvantage for UNICORE in the internal competition of grid middleware. Even participating academic computing centres were reluctant to base production on UNICORE since it was unclear how much a licence for the UNICORE components (servers and clients) would eventually cost after the project finished. Eventually, the funding agency accepted that a middleware like UNICORE could not be sold like Windows (in large numbers) or like DB2. The commercial partners no longer pursued the model to sell software. The revenues would have to come from services. As a result the commercial partners Pallas (now part of Intel) and Fecit (now part of FLE) transferred the rights to the base technology over to the Unicore Forum, e.V. This is a registered association with the mission to promote UNICORE. It consists of Hardware and Software vendors, universities and research centres. The Unicore Forum decided to make UNICORE available as Open Source under BSD licence and grant royalty free usage of the software because its commercial partners would not accept restrictive licences on derivative works – the GPL for example.

UNICORE partners wanted freedom for their own directions in commercial exploitation. Consequently, there are concrete results in new functions and features and new versions. This is similar to the EU funded projects GRIP and UniGridS which also employ the BSD licence. Today, UNICORE as a micro-ecology basis is flourishing – for example T-Systems provides commercial, fee-based support for customers. Also FZJ, Germany, gives academic institutions support for their UNICORE grid installations, but free of charge.

UniGridS offers extensions for standardised interfaces to UNICORE and as such employs BSD-type licences to encourage commercial take-up.

One key success factor for grid computing is accepted standards. UniGridS was an EU funded 6th framework Specific Targeted Research Project (STREP), funded for 2 years from July 2004, now finished. UniGridS was a follow-on project from GRIP (Grid Interoperability Project) which was funded for 2 years by the EU to realise the interoperability of Globus toolkits and UNICORE (Uniform Interface to Computing Resources) and to work towards standards for interoperability in the Global Grid Forum. UniGridS partners contributed to the creation of these standards and strived to influence them based on their extensive experience with UNICORE and interoperability between UNICORE and Globus. The basis was the Open Grid Service Architecture (OGSA) from the Global Grid Forum (GGF) and the Web Service Resource Framework. 
The UniGridS project developed a grid service infrastructure based on the Open Grid Services Architecture (OGSA), based on the UNICORE grid, grid software initially developed in the German UNICORE and UNICORE Plus projects. The target of the UniGridS project with respect to Business-Services and Workflows was creation of a sustainable service-environment based on UNICORE/GS. To develop the UNICORE offering further, UniGridS, used grid middleware common interfaces exposed via a Job Management Service (JMS). These objectives were much to do with standardisation and to influence the standardisation process, for selected grid standards while promoting European research results, by leveraging member participation in standards organisations - GGF, OASIS, IETF, W3C. Wider uptake of UniGrids results is still encouraged via making the source available outside the project via the download website for the pilot offering with dissemination of UNICORE/GS based services. It also encourages support and maintenance for UNICORE/GS pilot environments. The project offered UNICORE/GS tutorials at major conferences and via AccessGrid. Software produced as part of the UniGridS project, which is a public project, is available to download.

Note the role of UniGridS is to encourage commercial uptake of UNICORE by others for commercial exploitation. Thus the project offers its source code freely via a BSD type licence following UNICORE practice for open commercial take-up available from the SourceForge UNICORE website.

| CD02, grid application as service | CD02 uses commercial licence for customers. Software is built using OSS with BSD or MPL style licences which permit free incorporation as part of commercially used products for profit.

CD02 sells the financial industry usage of a specialised wholly-owned application on credit default swaps (or structured credit product) that requires high power computing to calculate probabilities of default for a large aggregate of companies. The application thus uses grid computing from a single site on which clients can run the application, with data feeds coming from market information providers such as Reuters and also input their own data. The financial application is delivered as a remotely hosted service based on hourly usage for the end users. The facilities used are leased from Sun Grid which gives assured QoS. CD02 pay Sun $1 per CPU hour for a service which includes access to up to 128 dual Opteron servers (Sun V20z) with network access, disk space, load balancing, switches, firewall, DMZ, UPS etc., redundant facilities, all hosted offsite from the CD02 premises in a single Sun-managed London site. Sun only charges for servers in use, which varies with numbers of clients. The major customer base is in the small client segment, of those who just wish to use a service. This is an ASP model, based on leveraging Sun facilities to assure very low capex outlay, and ease of entry to start up once the application and grid middleware have been developed. Users are spread worldwide across Europe, also to USA to Australia. The charging model was a key concern – various alternatives are charges per node, per user etc, but the final model chosen was per hour due to difficulties with other metrics, ie a one size fits all users model. A flat annual set up fee is also charged. This is flexible. Users can use as much as they require. In summary, the business model for this grid computing solution is to deliver software as a service (SaaS) for a multi-tenant grid environment with a scalable offering as it is based on combining flexible grid middleware with third party information over web services (e.g. market data).

On the licensing side, CD02 uses commercial output, OSS input software...
licences largely. The only commercial input licence is MS Office. Operating system is Solaris 10 – now OSS – (which ex-CEO McNealy recently publicly regretted should have been done earlier). All other input licences are OSS (for Jini, Blitz etc) but of different types: LGPL, GPL, BSD and the Sun Community Source Licence (SCSL) for Jini – which will be Apache, APL2. The CD02 commercial licence hourly charges reflect a premium on the Sun charges for the application services provided. CD02 sees the Apache and BSD licence as being the best for a commercial concern which builds on OSS and charges with a commercial licence for the bundled offering. OSS licences that are not for profit are not good to a commercial concern. Note also that there is no need for compatibility between licences as long as the freedom exists to sign up customers. CD02 could even have its software under a GNU GPL type licence as its offering is presented as a service, not a product, but GPL might perhaps also constrain the code fixes that could be performed and so is not favoured.

| GRIA, grid middleware | The EC IST GRIA project started at the end of 2001, with a consortium led by IT Innovation and ended in late 2004. Shortly after the project began, the wider grid community turned to Web Services with the launch of OGSA in early 2002. GRIA is grid middleware which enables use of the grid in a secure, interoperable and flexible manner. It makes use of business models, processes and semantics to allow resource owners and users to discover each other and negotiate terms for access to high-value resources. By focusing on business processes and the associated semantics, it enables users to provision for their computational needs cost effectively, and develop new business models for their services. The packages implement an overall business process to find, procure and utilise resources capable of carrying out high-value, expert-assisted computations. Services from different providers can be combined to create applications using a simple and easy-to-use API. GRIA provides several software packages that can be downloaded free of charge, each for a particular business need. The basic application services package allows an organisation with cluster computing facilities to provide data storage and processing (using applications installed on the cluster) for trusted users. The service provider management package adds support for service level agreements to manage billing for service usage, that can be integrated with other (non-GRIA) application services. The client package allows users to access GRIA application services via simple desktop applications, including negotiating with service management facilities and provides a simple API for development of integrated applications that use such services. A client management package, covers the support of organisation-level management of users, including authorised sign-off for SLAs including billing terms, and centralised control and monitoring of which users can procure and access services. Today, GRIA is offered by IT innovation. Note the IT Innovation’s position is to offer the finished product and control its development. The LGPL for OSS releases was chosen so that GRIA can be offered free. LGPL was selected as it is less permissive than Apache, but not as restrictive as the GPL. This follows GRIA’s wishes, to maintain control of its code in order to control its future - so that it is not hijacked by commercial concerns and also to preserve its quality. BSD use for GRIA would mean losing control of it (as the BSD allows unhindered change and re-use of the original code). There could be use of a commercial licence for GRIA – or perhaps multiple licences. |

| InnerGrid, Grid Systems, S.A., | InnerGrid version V, IG V, is grid middleware designed to be the engine of a virtual data center. IG V virtualizes applications, optimizes service |
delivery, and simplifies the management of this virtual computing environment. Moreover, in order to unlock its potential, the GridSystems company is committed to an open-source business model, as it fundamentally believes in the economics of professional open source models. The next version of IG, scheduled for end of 2006, will be open source. Some 600,000 lines of IG V code will offer a solid core of grid technology that can be used widely and freely by a large community of users and developers. Grid Systems believes that lowering the adoption barriers and providing a true affordable model for everybody will unlock the potential of our grid fabric and will ease market adoption. The company employs the dual licensing model of MySQL, viewed as optimal for professional open source. The dual licensing model provides professional subscription support services for customers under a GPL licence that coexists with a flexible commercial licence. Thus IG V, Community Edition, will be released having a free licence for any user, so it can be freely used within a small number of grid nodes and is downloadable from the corporate website under this special licence. Why has Grid Systems made its software Open Source? The company notes that most hardware vendors, and infrastructure software providers, have developed their own grid initiative with successful initial traction. Some software companies have been successful in specific vertical industries. Also, commercial first steps have been taken by certain academic projects. The company sees that many academic projects are developing excellent ideas, with research for the next-generation of grids. Unfortunately, most academic projects lack some of the fundamental features demanded by industrial enterprise grids: robustness, full integration with the existing technology of the data centre, easy installation, configuration and management, support and solutions, etc. At the same time, these projects are fully open source, and they have developed large communities of academic users and developers, with hundreds of research projects based on their technology. There is a tremendous untapped potential within these projects.

On the other hand, commercial offerings have the advantages of market focus: technology with support, strong marketing teams and solid customer references. But, in some cases, the technological foundations are not especially sound, with legacy technology being wrapped in modern clothes, or with products lacking the most fundamental layers of application virtualisation and resource optimisation. The legacy of the high performance computing batch schedulers is seen to be too dominant in this space still - there is a fundamental need to improve the technology underneath. So Grid Systems wants to make IG V a technology platform that integrates the best of both worlds: the openness and community of academic projects, with the robustness and support of a commercial entity behind the product. The idea is to combine the best of the research ideas with real customer needs, so grid technology realises its potential of becoming the common open distributed computing fabric.

OMII, Open Middleware Infrastructure Institute, University of Southampton, UK

OMII-UK is a centre of expertise that aims to provide software and support to enable a sustained future for the UK e-Science community and its international collaborators. It provides informed access to software and guidance to the UK e-Science community - researchers, developers and providers and helps to disseminate e-Science software to a global community, with software support and training. It also provides collaborative mechanisms to enable the e-Science community to help itself while engaging with the international community to define, contribute and disseminate best practice and standards. It partners with other software providers to provide best of breed software solutions and works
with UK and European funding agencies to provide a sustainable future. A close association on GRIA is the foundation of its middleware offering for e-Science. OMII releases middleware suites for use in e-Science as complete packages for specific environments (e.g., SuSE and Red Hat Linux with PostgreSQL) and includes OGSA-DAI for multi-source data integration.

All of the software developed by OMII is provided under a modified BSD licence (the text is given below) with copyright control. This was chosen as free commercial generation of derivative products is the aim, with users allowed to deploy as they require. The OMII distribution also comprises open source software from various sources with different licences. Licences for all the other software can in the distribution can be found in the OMII top level 'licenses' directory. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the conditions given below are met:

a) Redistributions of source code must retain the copyright notice (Copyright (c) 2004-6, OMII, University of Southampton), the list of conditions and the disclaimer.

b) Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

c) The disclaimer is added that neither the name of the OMII, University of Southampton nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission. The OMII software distribution is provided by the OMII copyright holders and contributors "as is" and the licence states "any express or implied warranties, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall the copyright owner or contributors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage".


ARC
ARC (Advanced Resource Connector from the NorduGrid project in Scandinavia, http://www.nordugrid.org) has middleware available for open usage, usually by the e-Science community. Software releases are available under the GNU General Public License (GPL) so that it is free to be used anywhere by anybody. It is available from the NorduGrid public repository in pre-built binary distributions for a variety of Linux systems. This enables it to be used in e-Science projects without hindrance while the ARC development community controls its future extension and development. Source code is also available from this repository and from the NorduGrid CVS. Furthermore, standalone client binary modules (‘tarballs’) are available for a variety of Linux systems. This client comes with all the required external software (most notably, parts of the Globus Toolkit) and configuration, providing a complete, out-of-the-box client package that can be used by other grids, e.g., for data transfer or data indexing. Open source development of ARC middleware is coordinated by the NorduGrid collaboration. Contributions to the software, documentation and dissemination activities are coming from the community and from various projects, such as the EU KnowARC project, Nordic DataGrid.
Facility and various national grid projects. ARC user support and installation assistance is provided by the NorduGrid via the request tracking system and a variety of mailing lists. Extensive documentation and tutorial materials are available at the NorduGrid Web site.

| **gLite** | Enabling Grids for E-science (EGEE) project launched its gLite middleware initiative in April 2004. gLite components were selected based on analysis of requirements from the pilot application areas in EGEE, biomedicine and high-energy physics. Rather than trying to build a new middleware stack from scratch, the gLite design team included representatives from existing middleware providers (AliEn, Condor, EDG, Globus, etc), as well as other key stakeholders such as Open Science Grid in the USA, and has around 80 core developers across 12 main centres of expertise and countries. gLite integrates components from current middleware projects, such as Condor and the Globus Toolkit, as well as components developed for the LCG project. The resulting product attempts a best-of-breed, low level middleware solution, compatible with OSS and commercial schedulers such as PBS, Condor and LSF (from Platform Computing), built with interoperability in mind and providing a foundation of services for building grid applications from all fields. Originally, EGEE used middleware based on work from its predecessor, the European DataGrid (EDG) project, later developed into the LCG middleware stack, which was used on the EGEE infrastructure early in the project. In parallel, EGEE has developed and re-engineered most of this middleware stack into a new middleware solution, gLite, now being deployed. The gLite stack combines low-level core middleware with a range of higher level services. The grid services of gLite follow a service-oriented architecture (SOA), making it easier to connect the software to other grid services. It will also help users to comply with upcoming grid standards, for instance the Web Service Resource Framework (WSRF) from OASIS and the Open Grid Forum's Open Grid Service Architecture (OGSA). It is distributed under a business friendly open source licence, Apache. EGEE want the middleware to be widely taken up, and used not just in scientific projects but also in industrial applications, so that a business-friendly licence is essential. |

**Comment**

In comparing the above choices, we can see that projects differ in their choice for reasons of the ultimate goal of the project in each case:

- If the goal is to spread the technology into take-up for commercial operations then a licence which gives complete freedoms to re-use the IPR without let or hindrance is chosen – the BSD family and its derivatives such as APL2. For collaborative projects, creating technology platforms, a form of OSS licence makes sense.

- If however the goal is for the project to be long-standing and permanent in some way, and wishes to control the destiny of its code, without letting it be used in as a bundled software product, but usually standalone as is, and any improvements or extensions must return to the originators, then a more restrictive licence ranging from the MPL to the LGPL is chosen.

GRIA has gone in the latter direction. This type of licence allows companies, institutions and individual programmers to contribute freely and as peers, secure in the knowledge that
they may give to a larger work without it being hijacked by one community member for commercial ends except under agreed conditions as there are reciprocal rights over-

- The right to see and access the source code and any new modifications
- The right to make source codes copies and to distribute them
- Rights to make modifications and improvements
- The right to use the code under agreed conditions for commercial or public ends

A process for choosing a licence is shown in Figure 4.3.

**Figure 4.3. The steps in choosing a licence**

The spectrum of choice for licences is summarised below

From this approach, projects may choose their way forward, which depends on:

- The longevity of the project
- The intended final outcome
- commercial roles and wishes of the participating project members
- Market(s) for which the deliverables were originally intended; for multiple markets, dual or multiple licensing may be the way forward.
4.3. Impacts of licences on promotion of an open source reference implementation

The licence will be a key part of the open source reference implementation, with the roles of:

- **Market opener** – The open source reference licence model must be based on free usage for commercial intent to develop a European grid industry in both products and services. So the licence must be part of the grid promotional effort for a reference model which consists of robust and standard basic building blocks, able to be widely used by whosoever wants to enter the grid market, i.e. the licence is part of an open middleware reference model. The licence must enable a start-up to build value-added services and make profits out of grid technology. The challenge is also to get high-end application software vendors to buy into a grid service model as many software and systems vendors perceive their value proposition to be software access, rather than software use. Their licences usually do not allow applications to be used in network services for instance. Such vendors still expect end-users to accept the high entry cost of buying licences for a permanent in-house installation. The power of OSS has been to change this attitude when freely licensed OSS alternatives could meet the majority of needs (e.g. MySQL in database, or Linux in operating systems). Competition from alternatives, unconstrained by their licences from free distribution, will be the key to unlocking the creation of high-value application software for open network accessible grid services.

- **Packaging of the complete offer** – packaging of a complete and useful reference model is key – there must be officially released distribution kits for the OSS reference model, with the licence being an essential part of that package.

- **Securitising the reference model with support services** – Licences for the reference model must enable active maintenance, as the users of grid infrastructure and grid applications will require long-term support. Although grid middleware may be secure when developed, vulnerabilities will be found inevitably, and without active maintenance it will become unusable. Unless there is a process of ongoing maintenance and support for security, widespread take-up of grid software, which is openly network accessible, is unlikely. As a part of this, the licence also must enable third party commercial enterprises and OSS communities to provide support services, maintenance and training for the grid system components.

- **Providing a global licence model** – the definition of grid middleware standards must be public, to be accepted worldwide, and the licence must enable this. The public standard reference model for the open grid software should first be implemented in open source, and may then be ported to the most relevant computer architectures.

- **Continuing interoperability** – The licence should also permit commercial middleware providers to offer proprietary, tuned versions, while still guaranteeing the interoperability with similar implementations on other platforms (including the original open source reference model). This will establish an open and vibrant grid market.
As noted in the first point above, the licence should promote business models to be established on a commercial basis for any direction of commercialisation, while assuring:

- The reference model is guarded intact although development of middleware extensions is quite usual
- An open services infrastructure – so consumers are free to access applications via multiple service providers
- Sharing of applications or data on the grid with collaborating organisations in dynamic virtual organisations (VOs) is easily possible
- Development of new grid application products is encouraged
- Creation and promotion of services based on grid infrastructure such as ASP-like business models for compute- or data-intensive applications, provided as third party services, is supported

4.4. OSS licence constraints and solutions – a final summary

Here we give a brief summary appraisal of the enabling and inhibiting factors in licensing schemes with open source or multiple licence software which could affect the exploitation potential of grid R&D results.

Removing the constraints of proprietary licences
OSS licences intended for commercialisation of grid R&D projects should avoid the constraints associated with proprietary licences, and provide:

- Freedoms to commercialise the code as a product or service, through multiple service providers and product VARs
- Liberty of choice in terms of what code can be used with what other code, be it existing third party, newly developed extensions, etc, and combined into one program for later commercial sale or hire for profit, including the basis of services
- Liberty in terms of different licences for different software modules or applications being run or used in concert together – requiring a non-reciprocal nature of the chosen form of OSS licence in usage
- Freedom to offer warranties and support to cover a combination of the original source code and any added code and code libraries
- Protection of the original developers from any liability for subsequent commercial versions produced by third parties, except for the EU’s liability for death conditions

The reciprocal natures of certain OSS types of licences preclude them from the above enablers and should not be used for full commercialisation. Control of technical directions by the original developer is also a point where a decision must be made to retain control by requiring amendments to be contributions returned to the original code, or to relinquish control and not require such actions.

Choice of licence
As noted in the table above:

- For the dictates of commercial development and usage, the most favourable licence could be the APL2 for simple commercialisation
- For control of technical directions by the original developer – the MPL licence may be preferred over APL2
- Common open development (as in the pure OSS community development and Internet model) indicates GPL for pure R&D projects that will remain in an open
forum (not quite the ‘public domain’\footnote{Note that the term 'public domain' has no clear definition, and its interpretation differs by Member State.}) for academic research or similar projects that will not be commercialised

The multiple licence option (parallel and sequential schemes)

A further interesting approach for grid commercialisation is a multiple licence usage. The approach is to provide a range of licences which apply to different communities of users and different releases of the source code. This is typically used by OSS communities and commercial developers for offering both purely commercial and OSS offerings, which may be multiple. Thus it is possible to have a current stable version which is licensed for release under commercial terms perhaps, a first version which can be used by anyone for a commercial product, and is under a BSD licence but it is the primitive restricted version, and thirdly a staging or beta test version which is under GPL, so corrections must be returned to the development community.

Most multiple licence schemes are parallel in that several licences types are in use at once. However for purely Internet marketing, a serial form has been used, some would say cynically, in which the first releases are OSS in order to build take-up by a user community with the spread of the OSS version being by reciprocal marketing to create a reputation. Once established, a static OSS version may be left available but unchanging and without support. A proprietary version with full maintenance support and added features will now be launched on the market via the Internet. It will be especially aimed at the user community that has found the OSS version useful and is now perhaps prepared to upgrade to a paid proprietary licence model. This serial model has been used commonly in the free software market, such as anti-virus/spyware products. Usually the commercial licence fee is deliberately kept low so users are not (too) repelled.
CHAPTER 5. Synthesis of policy on grid industrialisation for framework programmes

1. Introduction and context

Our context for policy is the distinctive European vision of a grid environment, one that operates through services offerings and not just products and from the level of devices to supercomputers. It must serve communities ranging from individuals to whole industries, including data, information and knowledge and emphasise resilience and scalability. The technology could have a significant economic and social impact far beyond the scope of existing compute and data grids. This view should be contrasted with the USA vision of grid systems as programmer-level super computing. This European vision is well aligned the objectives of the Lisbon initiatives for a knowledge-based society.

Consequently there is a need for a concerted effort to create a European grid middleware community that can forge and carry through an effective approach to industrial strength grid technology. Action is required to organise the European “community” of grid technology and middleware development with grid-application research into a focus for commercialisation.

Examination of development successes in other countries and regions indicates that a judicious interventionist policy might be useful in the European context if pragmatically applied in concert with the major players to create a grid industry. This approach requires a changed mindset and one where project funding and project exploitation plans target areas where Europe can realistically achieve success.

There are multiple choices in the strategy to pursue for this type of policy – two leading directions are either to let the market build an industry, using significant support at a European level, or else, to go further and build a new extension to what is considered our infrastructure, at the level of a grid computing layer above the Internet for all to use in open fashion. An industry of software and services would grow from this infrastructure.

When we turn to the process for creating the underlying software for these approaches, we see that the established mechanism, in which industrial partners in EC projects invest in pre-competitive commercial product development and support, is unlikely to succeed for a single open grid middleware as it produces fragmented and competing results. Examples like Apache suggest that the best way forward is to support the emergence of one central model of secure middleware for commercial usage’s openly and distribute it freely, in an open-source community model. Commercial mechanisms can then be used for development of applications and enhancements that build on the core middleware. The combination of elements from FP projects for the core model is yet to be determined.

Once an open standard toolkit has been established, export of products, services and know-how can follow. The GSM mobile example is perhaps not always a true and useful model but it has been one of the rare cases where a research and development initiative in Europe
produced economic benefits for European industry and service providers both within Europe and globally.

To guide, promote and remove key barriers, we recommend the establishment and promotion of a permanent European institution for grid technology which will energetically support commercialisation, including use of the OSS process. It would act as the owner and certification point for code in a common OSS grid toolkit for commercial applications. The proposals put forward here are based on a number of relevant policy documents which have already considered the future of grid technology; our proposals are put together to form an action plan as the way forward for grid commercialisation policy in Europe.

2. The context for a European policy and its core strategy

2.1. The level of policy intervention for Europe

Today there is a gap between research results for large e-Science and academic projects and commercial usage, which today has yet to flower. The established European exploitation mechanism, in which industrial partners in EC projects invest in pre-competitive research and then invest in their own commercial product development and support, is unlikely to succeed for grid commercialisation of products and services. It tends to produce a range of competing results rather than one model. It fragments any effort to build a single common model. Today’s commercial use of grid technology is focused on technical high power computing in specialist niches of aerospace, auto or oil and gas exploration, etc. This interval in the path to wider commercial uptake is comparable to the Internet before its commercial uptake following the addition of the (free) World Wide Web as a far better user interface.

The problem of how to bridge this gap might be resolved by considering other models of development in high technology from economies in Asia and elsewhere. A useful and illuminating example, perhaps, despite the deep structural and cultural differences is South Korea’s path from an agricultural economy in the 1960’s to its current position, across some fifteen directed technology programmes, each of around 3-10 years. Here, a concerted effort was made to pick winners in a flexible manner. By ‘flexible’ we mean that the programme could change direction or emphasis within 12 – 24 months if the original goals were not seen to be so useful. Although this is not really a path for Europe to follow, it does indicate the role of positive actions in introducing step changes in technology advance and commercial take-up. Korea’s later ICT programmes following from earlier industrial plans are shown in Figure 5.1.

Figure 5.1. South Korea’s usage of directed technology diffusion

Korea’s unique planning system has transformed the economy by a successive series of medium term plans for ITCs, which depend on a coordinated co-operation between the public and private sectors to deliver. They were preceded by plans for industrialisation in the long term, from steel through shipbuilding and car and consumer electronics into semiconductors and ICs.

The Korean grid programme, its strategy, structure and experiments

Korea is interested in grid technology as a way of introducing computing technology to an industrial environment which is strong in semiconductors, telecommunications and display technology but weaker in computing technology, especially software. Grid computing is one important direction for its declared goals of national competitiveness in advanced IT services. As such there are four major threads running through its initiatives:

- e-Science – weather, high energy physics, molecular sciences, medicine, space, etc
- Business use of grids – applications and grid systems as services
- Home working using grids in a utility fashion
- Construction of a national grid asset connecting the major centres of computing, e-Science and business applications

Business grids projects have concentrated on quite practical applications, often linked to telecommunications – such as a telematics car navigation system and a graphics rendering service. Korea is also interested in taking the grid to the level of the ordinary population and has plans, not yet too clear, on usage of a grid from the home. This seems to form part of a longer-term plan for a next generation of the Internet to be used by the mass of the population and businesses of all sizes. One experiment is the Korea@Home project\footnote{See http://www.gridforumkorea.org/asiagridsummit2005/data/C.Y.Park-Korea@Home.pdf from the GGF13 Asia Grid Summit 2, 13 May 2005} - the development of a software platform for Internet-based distributed computing using volunteer computer resources and peer- to peer networking, with forms of PKI-based
security and digital signatures, based on Web Services with SOAP messaging. Its aim is to
dynamically link resources as an ensemble to support the execution of large-scale, resource
intensive distributed applications. To do this, it requires a group-based computational
replication mechanism to adapt to an unstable, untrusted, dynamic peer-to-peer grid
computing environment. It currently has about 7500 volunteers and its daily performance is
stable.100

At a strategic level, grid technology is now being adopted as part of Korea’s long term
technologies programme under its IT839101 plan. Part of this plan’s strategy is to repair the
weaknesses in computing technologies and systems, by pursuing core areas in potential
services as the value added, as the Korean computing industry is not considered as well
developed (compared to its electronics and telecommunication industries). The Korean
Ministry of Information and Communications (MIC) has decided to compensate for current
weakness by promoting a software capability to support electronics and telecommunication
businesses. This has led to recently announced changes to IT839, to expand software
infrastructure. The business side of grid computing is viewed as a potential leader for the
Korean IT industry in penetrating the global computing business.102

Korean grid strategy is strongly connected to the business directions of its main telcos. They
are a focus for potential business development of grid services, with revenue earning
potential. The focus is coming from KT and SKT, because they are changing into service
business companies based on their core competencies as network operators. As such, their
services are evolving into fully converged forms, of telecommunications and media, and so
require intelligent real-time coordination of massive networked IT resources. Thus Korea
now sees the importance of grid technology is increasing, as the Korean telcos evolve
towards service business firms.

The history of Korean grid activities goes back to planning in 2000:

- Establishment of basic national grid project plan for next generation internet (May
  2001)
- Initiation of Grid Forum Korea (Oct 2001) alliance - currently, 25 working groups
  and 256 participating organizations
- Inclusion of grids in e-Korea Vision for 2006 by the MIC in December 2002
- Approval by APEC as Korea to be the nation having responsibility for the APGrid
  project, decided at the conference APECTel 24 (Oct 2002)
- Promotion of eScience projects and large-size construction technology
  infrastructure projects, connecting with a national grid project for all centres of
  expertise across Korea, 2002 onwards, as well as projects for business grid
  application and utility computing on a peer to peer basis via grids
- Collaboration between national grid project and 8 countries (US, Japan, China, UK,
  etc)

---

100 Sixth IEEE International Symposium on Cluster Computing and the Grid Workshops (CCGRIDW’06), Proc. p. 7
101 Group-Based Dynamic Computational Replication Mechanism in Peer-to-Peer Grid Computing , SungJin Choi,
Korea University, Korea, et al
102 In the IT 839 medium term programme, 8 stands for 8 essential new services, with 3 major infrastructures, and 9
new growth engines. The IT839 strategy aims at a ubiquitous ICT world by forming the mutually supportive
structures through introduction of new services, with construction of infrastructures, to seed new “growth engine”
sectors.
102 Techno economics and policy program (TEPP), Business Grid Activities in Korea, Junseok Hwang, Seoul
National University, 2005
At a tactical level, we see the common model of Korea’s technology advances – close cooperation between government agencies, industry and academic institutions, guided by key ministries, as shown in Figure 5.2, with much of the co-ordination through an established dedicated organisation, in this case the Grid Forum Korea (GFK):

**Figure 5.2. The Korean grid computing industrial structure**

The GFK was established in October 2001, under the Korean National Grid Basic Plan for the construction of a next generation Internet base by the MIC (the Ministry of Information and Communication). The aim is to promote a next generation of Internet technology, through construction of a grid infrastructure (‘Nation Grid’) and development of core application technology for sharing supercomputers and large-scale information storage, for high-performance computing in the 21st century. The GFK promotes grid standardisation within the corresponding international grid fora and pursues the national strategy. It also disseminates grid technology trends and information nationally, with an information exchange structure aimed at successful achievement of the National Grid projects’ objectives at a domestic level. Part of this role is to set up national competitive grid projects linked to ICTs and any relevant technologies.

To support this, the MIC has actively promoted the GFK and the various GFK projects, most specifically construction of the National Grid infrastructure with development of core applications, to assure construction of a national computing grid. The long-term aim of all this is to seed and nurture the Korean grid industry and then to graft it on to other successful ICT sectors, such as telecommunications. Positive government support enables the GFK to promote cooperation with industry as well as academia and government research institutes.
It acts as the representative forum in the advanced computing sphere, taking part actively in international grid development and standardization. The GFK actively engages competitively with others in the international grid and IT field. It is managed by a secretariat located in the supercomputing centre at KISTI (Korea Institute of Science and Information).

The lessons we may learn from this approach are fairly clear, although whether they are at all translatable to other regions, such as the EU, is obviously in question:

- Strategy should be set at a high level and vigorously pursued through a coordinated environment
- Tactical coordination is through a dedicated entity which implements the high level strategy
- Focus is on business as much as e-Science
- Tie in the most appropriate industrial players (the telcos in Korea) to give business applications which may act as demonstrators
- Focus for the future is on a ‘grid for the home’ as much as on e-science/large corporate grid applications
- A long term plan must be flexibly implemented behind it all, one which fits grid in with other technology development initiatives to form a more holistic ‘master plan’ for a future knowledge-based society

2.2. Realism in intent and goals for an EU grid policy

Naturally, the Korean and other Asian models such as Japan are not replicable in Europe. In consequence, what we will suggest here follows the model above only in the extent of intervention outlined here, which is perhaps higher than previously considered. What we see as being required in a global context will be an interventionist attitude to take the next step from the e-Science and vertical enterprise-level grids towards a more ubiquitous grid environment for widespread usage. That said, the policy must take into account the short-term and long term demand side pressures, else it risks becoming a white elephant – a twelve-lane autoroute to nowhere. It must be a demand-side led approach. We envisage that EU policy will be predicated upon three phases of grid market development, and will be designed to move on from the first and second phases to the third, as illustrated in Figure 5.3.
Here we consciously use the mechanism of Open Source Software to create a platform with the necessary critical mass of developers and users required to commercially exploit the results of EU R&D projects. We thus favour a European policy for grid technology commercialisation built around OSS as the industrialising and production model for the EU projects.

2.3. Strategic policy alternatives to build a European grid industry

Our role here is not to decide on all policy options so much as to give those options and indicate their advantages and consequences. The policy alternatives to build a grid industry are various but could follow two principle routes, one more ambitious than the other:

1. A grid industry built by the market

A first direction is to precipitate a European market through judicious catalysts for market formation and expansion by the key market players. This will still require some strong form of support, principally the creation and promotion of a standard toolkit and reference model, with support services. These common building blocks can then be diffused to the European software industry, which may use it as a basis for grid technology commercialisation in products and services, and to the user enterprise / consumer market that may use it directly. In this model, the spread of grid technology, following its promotion, is left to the market, to create an offshoot of the software and Internet services industry in derived products and services which can run over the Internet, or in more secure infrastructures. These would
come from commercial players of all sizes, especially the SME players who dominate the European software industry.

2. Building a ubiquitous grid as part of European infrastructure

Second is the policy option which combines the above efforts for commercialisation, together with the establishment of a European industrial strength open grid infrastructure, for grid application services. It would be open for all to use and would depend on attachment of resources over suitably high-speed connections as part of the services offering which would be charged for by a grid services provider. Overall, this is a far larger and more ambitious project in effort and budgets. It is a long-term programme and needs a roadmap. It would require close collaboration with major industrial players to promote and implement such an infrastructure. Establishing a grid commons as the infrastructure would provide a supporting matrix for business operations, and perhaps for personal computing for the citizen. This may involve incorporating new concepts – such as common ownership at the infrastructure level.

Such a strong interventionist policy would need far heavier support and commitment in terms of planning, co-ordination, industrial guidance for major players perhaps, implementation and of course financing. The aim would be not only to host grid application services but also to build a software market on top of the standard grid toolkit. Such an initiative could fall under the aegis of a Europe-wide grid agency, which is described in detail in the section below.

This larger design, to build and deploy a Europe-wide grid infrastructure being a long-term venture, requires an executive team and partner resources to manage the deployment. They should negotiate and establish functional applications for early users and technical agreements, including international standards. This venture may also take into account the strengths of the European mobile technology industry and embedded technology which is of particular relevance for the realisation of the vision of a grid infrastructure as a ubiquitous, user-centered commercial utility. To get off the ground, the whole project rests on clear and effective dissemination of promotional messages and marketing to appropriate stakeholders in government and industry for its creation and take-up. The aim for such messages would be to show real benefits in commercial situations and to the economy as a whole.

Although the first option is more dependent on a free market mechanism and laissez-faire than the second, it still requires the similarly significant support institutions, as developed below under the drivers for a grid industry. The second demands a costed implementation plan as part of a detailed business model, which can only be put together when its scope and main foundations have been agreed on.

To assist either of the above policy directions – following the lessons learnt from looking at managed development in other economies – it would be invaluable to guarantee a first large client who acts as ‘prover’ and improver. Obviously government users would be a first choice, with pilot grid applications as close to mundane business problems as is possible to find in e-Government.

One further policy related area is what standards should Europe follow? A clear answer in business model terms is that standards efforts should be driven by industrial needs. OSS tends to set real standards for grid operations in terms of interfacing and integrating, both for code and for content. European commercial grid technology should thus follow international
trends in de facto standards for grid middleware, set by OSS largely, following (and participating in) the open standards bodies. They include the W3C, OASIS, IETF and OGF, be it at product or service level, including content and data interoperability. The focus in standards’ bodies is usually on straight middleware protocols and APIs but it may need to be enlarged to follow market trends in demands for security, licence discovery and business SLAs.

Here it may be useful for Europe to manage the co-ordination of industrial standards requirements, in co-ordination with multi-national IT vendors, to assure that EU standards efforts are pragmatic and driven by industrial requirements. It should be recognised that the standards arena is a political struggle and that Europe’s real interests lie in getting products accepted in a broader user community. The battle for standards and to exclude patent infection or hijacking is a long-term process but it needs a tight focus, throughout the standards bodies and vendor partnerships. The wreckages of failure here are highlighted by examples such as the mobile 3G initiative. An aware and lively effort is required – not half-hearted in any way – with funding over a long period for targeted efforts based on European grid infrastructure requirements.

Thus any policy should also have a dimension of international collaboration. Engaging with international research programmes outside the EU (e.g. in Asia or with the NSF) may be fruitful but it should be recognised that these other initiatives may reflect different needs – for instance for e-Science rather than those of industry, including European industry. Collaboration would be based on mutually beneficial goals, such as the common open standards essential to development of an open source European grid middleware.

What we now examine are the tactics needed to carry this out, in the form of an ACTION PLAN for grid commercialisation in the EU, whose elements are applicable to either of the two paths indicated above.

3. Policy tactics for the potential market

3.1. Drivers to facilitate grid technology commercialisation – recommendations for an EU Action Plan

**Basic grid technology design and toolkit for commercial implementations**

- Common grid reference model using open standards
- Common OSS grid starter kit based on reference model with open APIs, tools and documentation for scaled deployment and ease of roll-out

**A European Agency for grid technology – a core set of market support units**

The EU’s objectives could well be achieved through a permanent institution that brings together existing grid technology activities and expertise, acting as an institutional guardian of EU grid strategy and implementation. It should be a joint venture incorporating most strongly the software and computing industry, but also academia, research communities and government.
It would harness a series of open source initiatives for a standard toolkit and reference model, a software repository as a standards’ reference point, with a benchmarking suite to guide and promote development and integration. This follows some of the lessons of the OMII in the UK. The latter has indicated the way in which public support is able to develop high-quality open source implementations of robust grid middleware and so facilitate the adoption of grid technology.

The functions of such an agency need to be comprehensive and include stimulating truly intense interaction among grid community members in Europe, in effect creating a critical mass for development, deployment and adoption within its OSS community. The initiative should early on identify such practical needs as guidance on licences and IPR, a proving and integration platform and code cleanliness certification, as well as the marketing, educational and training demands. Overall the agency would have several major goals:

- Maintain the policy directions flexibly, following feedback from experience of actual progress, in the form of a Grid Roadmap to define the direction and milestones for EU grid computing and so provide potential users in industry with a clear vision of grid computing evolutions for the next phase of commercialisation over at least 5 years.
- Promote a standard EU grid reference model and its implementation as a toolkit, with an open source community for both the toolkit and reference model, orchestrated by the agency and incorporating all developers, be they commercial, academic or from the FP projects themselves.
- Middleware diffusion: create a European structure acting as world-wide software publisher, standards certification centre and point of responsibility for the diffusion of European grid technologies and tools. It should provide open source grid middleware, under licences suitable for commercialisation. It would also provide the documentation and associated services to give confidence to industry users and their customers in terms of code cleanliness, licences, etc. Having a single European editor and OSS publisher is a key factor for success.
- Technical promotion: Provide an infrastructure for proving, certification and standardisation with long term support of basic models and a grid toolkit for other grid software.
- Experimentation: create a European centre for experimentation and training allowing enterprises to have a ‘sandpit’ distinct from certification to test grid infrastructures, performance and interworking with middleware.
- Promotion of commercial grid usage with practical level analysis for business: studies of business models, value added services, tariff models, licensing models, legal requirements and implications, and concepts of the new business ecology of the grid, with conclusions that persuade industry to begin making the investments and culture shifts necessary for them to participate.
- Information and promotion actions: create and communicate the business vision for European industry and the entrepreneurial community to promote the use of grid technologies in the industrial world. The focus must be on promoting grid concepts that business can grasp now, to educate European industry and the wider public on grid technology as a business infrastructure. Effectively the campaign needs to capture the public imagination and that of the wider business community. Eventually it should also organise international collaboration, especially with the major players and industry organisations.
- Provide a global view of the technology, perhaps to influence the development of grid technology directions in other regions and countries.
It would go further than the OMII, with the following functional units:

- Certification of cleanliness of OSS code from projects- to assure developers and end-users that they have OSS clean code – effectively a clearing house for IPR. For data centric applications involving large volumes of shared data, this might include a Data certification centre – a clearing house for data IPR.
- Ownership of IPR and OSS code with liability for code under EU law.
- OSS repository/download centre.
- Assure the basics from projects for building products and services is available as a standard pack across all code and R&D projects – especially the documents and tools for an OSS toolkit.
- Suitable licensing assured to cover services as well as products.
- Interoperability and test platform for applications and grid middleware, centred on a standards base for the establishment and promotion of an EU Grid Standard Basis (EGSB) for a certification initiative,\(^{103}\) with a grid benchmark suite for interoperability of different grid middleware suites, resource environments and grid applications.
- Standard EU OSS grid toolkit, which supports Open OSS and de facto standards plus a European download centre.
- Education on grid technology with packs for enterprises and educational institutes of all levels.
- If required, a certification authority (CA) for establishing the authentication credentials for grid software, resources and their users with digital certificate distribution.
- Promotion and market education, with coordination and consolidation of European-wide R&D efforts for grid.
- Demonstrator platform to show potential users and developers the capabilities.

A European Grid Technology Agency could be centred at one site or established as a virtual organisation.

**Technical features for commercialisation**

The technical features of the grid technology reference model and toolkit should have the following features and enhancements, as a minimum:

- Security layer which has far more advanced protection and privacy features than today’s grids, but in a lightweight fashion (also discussed further below under market perceptions)
- Licence discovery and verification service (inside middleware) using application and server identification notes
- SLA Contracts service (inside middleware)
- Licence identification and negotiation layer (inside middleware)

---

\(^{103}\) The methods could be modelled on existing certification initiatives, for example the Linux Standard Base (LSB). This requires the establishment and maintenance of the grid benchmark suite. Similar initiatives were used successfully in the MPI environments, for example the Pallas MPI benchmarks.
Industrial proving and promotional projects

Shop-window projects
These are simple working solution to a key problem in each industry sector, with a short timeframe. These might be based on a European FP grid infrastructure project. Application vendors should be engaged to demonstrate how they can leverage the grid.

Large demonstrator projects
The policy route of building a Europe-wide infrastructure could benefit from well-defined large pilot application projects as demonstrators of public grid-application services (GAS) at a European level with a long term remit. Two sorts of demonstrator projects are envisaged

- Deploy a Europe-wide reliable grid infrastructure for a large initial client – government. It should also be made available to industry to enhance confidence in grid from those end-user industries that have problems suitable to be solved with the help of grid, but which still do not have enough evidence that such technology could give benefits on a large scale and in an industrial arena with demands for a production quality infrastructure.

- Generate grid application projects that may produce immediate benefits for the general public by addressing the concrete needs in society, with clearly defined objectives and an immediate social impact. Applications may be part of the government infrastructure project and could be in the areas of environmental controls or measurements, energy, health care, education and social support. Such services would be designed to test and improve a grid computing infrastructure and to facilitate creation of services for consumers and business users.

Market-oriented R&D
It should be recognised that successful innovation rarely jumps simply from pre-competitive R&D into the commercial market. Projects funded by the European Commission under the Sixth Framework Programme are by their very nature not ‘close to market’. Thus, there will undoubtedly be a need for market-oriented research in grid computing - but this is not, we hasten to add something which is the direct concern of the European Commission or DG Information Society and Media.

This has already been recognised by the Japanese who have established a Business Grid Development Project under the auspices of the Information-Technology Promotion Agency (IPA). This three-year project, which finished in March 2006, brought together Fujitsu, Hitachi and NEC as industry members of a consortium including the Grid Technology Research Center of the Agency of Industrial Science and Technology (AIST). This initiative is financing R&D of business grid middleware and aims to apply the technology to the business world while ensuring international standardization of the system interface.

However, perhaps a more useful model is provided by the EUREKA programme, which is a pan-European network for market-oriented, industrial R&D. EUREKA aims to enhance European competitiveness through its support to businesses, research centres and universities who carry out pan-European projects to develop innovative products, processes and services. Project partners gain access to knowledge, skills and expertise across Europe via the EUREKA network which facilitates access to national public and private funding schemes. EUREKA helps project partners reach agreement over intellectual property rights

---


105 [http://www.eureka.be/about.do](http://www.eureka.be/about.do)
and helps build partnerships to penetrate new markets. A particular attraction of the EUREKA model is its ability to engage with and support SMEs, for instance through its agreement with BUSANET, a European network of Business Angels. This non-profit making organisation’s objective is to develop funding activities and financial toolboxes to support innovation implementation in Europe.

EUREKA itself might provide a suitable framework for market-oriented grid computing R&D through a EUREKA Cluster or Umbrella, or at least there may be lessons to be learnt from the EUREKA experience. We reiterate that it is not the role of the European Commission to fund market-oriented R&D but, nevertheless, the Commission could play an enabling role in establishing an appropriate mechanism to support this kind of activity.

### 3.2. Barriers to be broken down

To move forward, we must dissemble various barriers by an active campaign under the Action Plan:

**Understanding of grid is low across the market**

What is urgently needed is education of the EU market:

- A programme of publicity and dissemination activities in the EU. These include press campaigns, workshops, seminars and fact-sheets. The aim is to:
  - Create *brand recognition* for the grid middleware initiative, its toolkit and agency with the support components
  - Promote widespread take-up of the OSS grid toolkit and its support services for integration by industry and researchers.
  - Enhance the perception of a market coming to maturity with the availability of commercial offerings, useful in business.

- At international level, pursue a significant and ongoing impact on standards. It is vital that the European vision for the evolution of grid technology as a commercial infrastructure is accompanied by a clear implementation of that vision, with its representation in the key standards bodies and technology providers worldwide, in order to set the standards it requires

- Pursue and encourage the productisation of a complete grid offer (not just middleware) which can be demonstrated for credibility, showing the applicability of the technology beyond e-Science. The demonstrator projects will be essential here, especially with substantial support for an open-source grid application for usage in government projects. This first large client project should be used to educate and build confidence in the maturity of the EU grid technology.

- Invite existing institutions, commercial players and projects across the EU to partner in a federated campaign for the promotion of grid commercialisation to form an industry body. Quickly establish the leadership and management of this federation, with a set of partners that is of manageable size while retaining coverage. Partners must command suitable resources and executive capability as each would cover a

---

[106](http://www.eureka.be/thematic/aboutStrategicInitiatives.do)
specific area or role, such as a European region, a constituency of technology providers, or a consortium of demanding applications. The federation’s charter should be set with a long-term programme of focused targets, producing projects, studies, standards negotiation support, analysis and promotional material following the strategic direction for development of EU grid middleware and its commercial exploitation. This “Groupe Speciale Grid” (GSG) could later have a global extension.

**Perceptions of security and privacy – building trust in the technology**

Current models for security are insufficient and need to be improved to make grids a success. Fundamentally new ideas may be required for future grids as well as further development on current middleware to incrementally work towards secure commercial grids:

- Initiate a debate, with a rapid focus on solutions, between European countries and corporations on the requirements for security and the conditions for effective and dynamic resource sharing on the grid.
- Facilitate basic research in the security field for fundamentally new ideas for future grids.
- Pursue research and development in the field of security to incrementally improve on the models for current grids. Europe has a lead in developing the current authentication and authorisation model and may take it further.
- One part of the aura of dependability and trust is to assure that the OSS nature of the grid model is a catalyst and not a hindrance. The measures for software ownership and certification offered by the grid technology agency above assure code purity of OSS for users and developers against legal challenges. The licence policy expounded here is intended to give clarity of usage. It avoids licences unsuitable for commercial developers and any conflicts in operation, as a service or as a purchased product, with other application licences, across platforms and sites within the end-users.

**Consolidation of a single model from different grid middleware initiatives**

As noted, one overriding problem is straightforward and simple promotion of grid technology for take-up in the EU. While the USA has one model, Globus with its toolkit, which simplifies the technology’s promotion, Europe has many initiatives. The key business question at a marketing level is – what is the way the way forward. Consolidation of some form seems necessary. But what form should this process take? We suggest the following candidate process as one approach to consolidation (perhaps taking 5-6 months) using the public review method:

- Create a small panel (6-8 effective people) with participants from both business and technology, possessing a knowledge of grids and commercial operations
- State goals clearly and publicly and announce them with an actions timetable, as in the IETF process, as a request for comment (RFC 1) to the wider EU grid community (to inform and promote) and gather comments in limited timeframe
- Outline a candidate reference high-level architecture for grid products and services with placeholders for main components – distribute as RFC2 and gather comments in limited timeframe
- Select candidate architectures for reference architecture from existing grid projects - RFC3 - and digest feedback
Select potential contenders for each industrial application sector - issue RFC 4 for feedback

Identify the place for each potential chosen contender for logical components, with alternatives by business application so there is a common reference structure at a software component level (e.g., a scheduler) with options for components where appropriate - issue RFC 5

Draw up a series of drafts of a consolidated grid technology roadmap with its reference model and toolkit components, as a further series of RFC's. The roadmap should indicate any modifications or extensions needed, the licensing scheme and the schedule of potential source code releases. All this is aimed at a rapid advance to a public publishing of the final plans of an ‘EU grid blueprint’ as well as understanding the financing necessary for the changes for the toolkit for certain components—under existing projects, or as new project additions/extensions.

3.3. Business models to encourage commercialisation

Examples like Apache suggest the best way forward to support the emergence of a commercial grid technology is to follow the example of an open-source community:

- Secure middleware is developed openly and distributed freely.
- Commercial mechanisms can then be used for development of applications and enhancements that build on the core middleware.

The ecology of a business infrastructure that we wish to encourage by judicious policy is shown in Figure 5.4.

Figure 5.4. The business ecology of a successful grid infrastructure industry

<table>
<thead>
<tr>
<th>Industry layer</th>
<th>Actors</th>
<th>Offerings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>End-users: business and consumers</td>
<td>Virtual organisations (VOs)</td>
</tr>
<tr>
<td>Service Provision</td>
<td>Application Providers-vendors and others</td>
<td>System Integrators</td>
</tr>
<tr>
<td>Product and service Elements</td>
<td>Service Providers for applications and grid computing</td>
<td>Providers of packaged grid infrastructure as a service</td>
</tr>
<tr>
<td>IPR Generation and Ownership</td>
<td>Middleware Producers (OSS projects / vendors)</td>
<td>Content (Data) Providers/aggregators</td>
</tr>
<tr>
<td></td>
<td>Computing, storage and network systems vendors</td>
<td>Grid application and utility providers (OSS and proprietary)</td>
</tr>
<tr>
<td></td>
<td>Publicly &amp; Privately funded Research projects</td>
<td>Content (Data) Creators/Owners</td>
</tr>
<tr>
<td></td>
<td>Software &amp; systems vendors</td>
<td>Standards bodies &amp; Alliances</td>
</tr>
<tr>
<td></td>
<td>Public Grids</td>
<td>Enterprise Grids</td>
</tr>
<tr>
<td></td>
<td>Grid Infrastructure resource products</td>
<td>Proprietary IPR</td>
</tr>
<tr>
<td></td>
<td>Applications with support services</td>
<td>Open standards</td>
</tr>
<tr>
<td></td>
<td>Generic bundled grid products, services and application services</td>
<td>Proprietary standards</td>
</tr>
</tbody>
</table>
In speaking of an open grid infrastructure we see that the business models to encourage at a policy level are as much those of service providers as product vendors such as ISVs and VARs who build grid applications (as GAPs – grid application providers):

- **GASP** – grid applications service providers – offer the application as a service
- **GSP** – grid service providers – offer computing services and data-centric services
- **GISP** – grid infrastructure service providers – offer infrastructure as a charged for service
- **GSI** – grid systems integrators – specialise in building grid systems for the large organisation

Note that in the complete ecology there may be additional players in the form of brokers, consolidators of multiple services and/or data resources who act as intermediaries and/or aggregators.

**Encouraging commercialisation through the WWG and its grid ecosystem**

The WWG would be a product of, and a platform for, a grid ecosystem. It could provide the vehicle on which sit all kinds of grid applications, as well as being the utility that users of all classes, from individual consumers to corporations of various sizes and government organisations, use as their computing resource. Thus its position in the grid ecosystem is both as enabler of new businesses in applications and services harnessing grid technology and an essential computing tool for business and everyday life.

**Policy measures for the development of a public grid, WWG**

For the move from phase two to three of the grid market development (see Figure 5.3) we may combine several policy moves examined in more detail elsewhere as the key drivers for building the WWG:

- Planning the concept carefully with donation of resources by many individuals and private companies, the latter who see a profit in multiple business models based on free hardware, and the former are prepared to donate hardware, as in the many projects that already use this model (eg SETI, searching for alien life in space).

- The proof of the concept in a large demonstrator project - here the idea of a significant public service or government application would be useful, in two ways. First it would demonstrate working technology. Secondly it would encourage an ecosystem of small and large players to provide services, software and hardware products to complement the new facility. It may well be that public computing resources for internal use could also form part of the shared park of servers.

- Adoption of the uniform common reference model for grid computing and the single toolkit made available on a free to use basis at no cost. It may well be that for the lower middleware layer, a licence of the form of the MPL or the LGPL may be the best choice as it occupies a space similar to large utility such as Linux or the Eclipse toolset. The upper middleware would be commercialised and so might use an APL2 type licence for commercial exploitation of an original OSS package.
Profiting from a WWG

Developing a WWG market in Europe with a profitable ecosystems, depends very much on the same factors or pre-requisites as for all networked services, including the Internet:

- Real benefits for its users, to seed real demand
- Attractive services in function and ease of access/presentation at basic and application levels
- Low cost or free access, that matches disposable income available to the mass of consumers
- Accessibility and coverage across the EU of the basic grid service

All of these points are consistent with the economic analysis of willingness-to-pay for network-born services due to Kridel\textsuperscript{107} and other telecommunications economists, built up since the late 1980s. Breaking down these pre-requisites into a logical business analysis for start-up gives us the key items to be set in place, which are much the same for any new large-scale digital business:

1. Organisational forms and the shape of the ecosystems, together with an understanding of patterns of success and failure – eg why is a social networking/personal experience service based on peer-to-peer content worth US$1.65 billion – the price that Google paid for YouTube in October 2006, for a 20 month old start-up with no profits so far, valued at an infinite number of times current earnings. This seems to be an unwise return to the dotcom boom prices, yet Google will pay $900 million over 3 years to MySpace.com in return for being its search provider, valuing the US$580 million that NewsCorp paid for MySpace in July 2005 at up to US$15 billion in 3 years according to some analysts.\textsuperscript{108}

2. Capital investment in resources and services – an understanding of the equivalent of the above for a WWG with its potential services revenues could bring in private investors. This could only follow on from seeding at a far lower level of expenditure, by a mix of government-sponsored projects and private venture capital, perhaps from the major players, together with donated resources from a new user community. One requirement is for venture capital to seed start-up companies for services and applications, on top of an ad hoc, and at first weak, grid hardware infrastructure.

2A. Operational expenditures to keep assets and services running – an output of successful capital investments, including support from its OSS development community

3. Catalysts that will be required to start up the WWG:

- Marketing of the WWG and its services, be it by reciprocal methods – effectively at no cost, or in an informal approach or with explicit activities and funding
- Supportive government and EU policy
- Supportive intervention by current key market players


\textsuperscript{108} T. Lowry and R. Hof, ‘Smart move or silly money 2.07, \textit{Business Week}, 23 October 2006, p 34, http://www.businessweek.com/technology/content/oct2006/tc20061012_597662.htm?chan=search
Designing a blueprint for Europe for profitable exploitation of a WWG

Some early recommendations on a profitable EU exploitation of the WWG are shown in Figure 5.5.

Figure 5.5. Each layer of the WWG offers services opportunities

At each layer we see the presentation of an offering, some of which involve a service provider. Note that the commercial structure is somewhat different to that for bespoke grids. For instance, grid service providers and grid applications-as-services providers would be more important while builders of bespoke grid systems – systems integrators – are likely to be far less important, as the infrastructure is brought fully assembled to the user across the network.