Meeting customer needs

C Gresty, T Dimitrakos, G Thanos and P Warren

The concept of the Service Oriented Infrastructure (SOI) is not just the result of technological advance; it has been stimulated by the needs of organisations. The modern organisation is seeking to efficiently utilise its information and communications technology (ICT) infrastructure, while at the same time ensuring that components of the infrastructure can be substituted and augmented dynamically. The 3-layer SOI model is a response to this. The model builds on the service-oriented philosophy, which is already having significant market impact – e.g., through the deployment of Software-as-a-Service. The SOI model also owes its provenance in part to the concept of Grid Computing; an approach to computing which offers performance benefits and the ability to share complementary resources. Today a new concept, that of the virtual hosting environment, is being developed, building both on SOI and grid computing. The goal is to create an environment where ICT infrastructure, delivered as services, can be shared between partners in a virtual organisation, or made available for third parties. The result will be an ICT industry where services are delivered across organisational boundaries, enabling all manner of new business applications, which in turn will themselves be delivered to the end-user as services.

1. Introduction

Technological advances are not just the products of scientific progress; so-called technology push. They are also created by organisational and social determinants; demand pull. They happen because they satisfy a need, whether business or personal. Sometimes these needs are long-standing; perhaps having been dormant because no technology existed to satisfy them. At other times they arise as a result of social, economic or political factors. The technologies and capabilities which come together under the banner of the Service Oriented Infrastructure (SOI) offer a solution to both long-standing business needs and also some needs which are particular to our time.

For example, a recent survey of major BT customers identified five areas of focus:

- virtualisation – i.e., having applications independent of the particular infrastructural implementation;
- application management and performance, so that the infrastructure can be substituted and augmented without disrupting the applications;
- network and IT infrastructure integration;
- globalisation;
- green issues.

The first two of these reflect long-standing business requirements. The third, the desire for network and IT infrastructure integration, has been created by the realisation of previous technological advances, particularly the ubiquity of internet protocol networking in local and wide area networks. Globalisation is similarly driven by technological advances, and by the increasing significance of India, China and other emerging economies. Finally, the attention paid to green issues is an example of a concern which has risen to prominence in recent times.

These and other requirements on business are coming together to provide an environment ripe for the adoption of the technologies discussed in this and the subsequent edition of this Journal. The aim of this paper is to describe how the technologies of SOI respond to the demands of this business environment and to provide an insight into how the environment is likely to evolve. In the following section, we provide an overview of how the corporate IT environment is changing. Then in section 3 we describe the key aspects of an SOI and how it supports this new environment. Section 4 describes the impact of SOI and related technologies on the IT market, providing some forecasts of market size. Section 5 discusses some of the issues surrounding the construction of business cases for these technologies. Section 6 then describes one particular new business model, that of the virtual hosting environment, and looks at the implementation of this business model in a specific business sector, that of online gaming. Finally, section 7 draws together the key themes from our paper and makes some final observations on what they mean for the future of our industry.

2. The customer view

Organisational IT is changing. For a considerable time now it has not been the sole preserve of the IT department. Consumerisation has prompted employees to question why useful hardware and software that they utilise in their personal lives cannot be used inside the corporate firewall. This covers access devices, social networking sites, applications, etc. Indeed, the term ‘digital native’ has been coined for a whole new generation of employees who are entering corporate environments today. Some of these natives simply won't join organisations whose IT environments they deem unfit for purpose.

As a result, IT departments increasingly see lines of business building their own quick fixes and applications. This started as long ago as the mini-computer but has gained pace recently, fuelled by Web2.0 technologies such as mash-ups and wikis. The concept of a Service Oriented Architecture (SOA) has not yet significantly entered users', as opposed to the IT department's, consciousness. An IDC survey from 2007 found most IT staff were aware of the concepts and benefits of SOA for their organisation but that line of business and business strategies still had a high knowledge gap. The essence of an SOA is to provide functionality from software components as services, thereby encouraging reuse. Loose-coupling between the various components means that any one component may be changed without changing the overall system, as long as the functionality offered at the interface to the component is unchanged. Imagine the business applications that will be created once SOA and associated mash-up techniques have fully percolated a business, as they are predicted to do in the next 2 years! [1]

Corporate IT is driven to provide the architecture within which the business operates but lines of business are gaining more flexibility in terms of the applications they deploy and operate. Corporate IT has a number of key drivers:

- Improving operational effectiveness through standardisation, rationalisation and offshoring of suppliers and infrastructures. The trend to outsource non-core functions comes into play here, creating widespread acceptance that parts of IT can be purchased as a service rather than owned as an asset;
- Improving employee productivity, both in their internal interactions and those that span the supply chain (i.e., with external suppliers and partners). IT recognises the need to provide flexibility and agility and is using this tactic as a means of getting closer to the business;
- Adhering to regulatory constraints and protecting company assets (hardware, software and knowledge);
- Increasing return on investment, often so as to free up budget for investment in innovation in the business process and supporting IT.

The multiplication of IT applications throughout an organisation means that corporate IT departments find it increasingly hard to predict infrastructure demand into the future; they struggle to understand what their users are doing today, let alone predict what they will be doing tomorrow. Yet the criticality of IT means that demand must always be met. The solution lies in seeing IT as a service to be bought as and when required.

Compounding this internal complexity is the growth of ecosystems between companies active in the same supply chain. Prior to globalisation, such ecosystems managed well with technologies such as Electronic Data Interchange (EDI) but now, in an era of global supply and demand, interactions between trading and business partners must occur seamlessly, securely, sustainably and immediately.

Sustainable, or Green, IT has finally hit mass awareness amongst corporate IT professionals. 2007 was the first year, for example, that Green IT presentations at the Gartner Symposium achieved higher delegate attendance volumes than any other session. Whilst some IT functions have simply grasped the concept, others are a long way down the track to measuring, reducing and auditing IT's energy efficiency. Many of the core concepts of SOA and SOI (such as virtual, reusable, etc.) are inherently green and as such, IT organisations see synergistic opportunities to address their carbon footprint whilst at the same time delivering more flexible, agile and innovative IT.

A recent study by McKinsey identified outsourcing increasingly specialised aspects of the organisation's work, whilst maintaining organisational coherence, as one of the "eight business technology trends to watch" [2]. This is another aspect of the growth of ecosystems mentioned earlier and, again, it requires a new approach to managing and sharing ICT resources. Another of the trends identified was "unbundling productivity from delivery" – i.e., disaggregating monolithic systems into reusable components. Here the design philosophy of SOA has a significant contribution to make.

3. Emerging solutions: from service oriented infrastructure to service oriented enterprise

The subject of these two editions of the BTTJ is the Service Oriented Infrastructure. This is a relatively new phrase, but the technologies discussed here have a long provenance. Fifteen years ago we might have spoken of distributed computing; a decade ago we would have talked about Grid computing; in any case, both of these terms are still used and form part of our subject. Another SOI-related term which has a long history is virtualisation; and the phrase Service Oriented Architecture (SOA) has also been in use for some time. This is perfectly natural. Technologies do not arise instantaneously. They develop over time, and as they develop different aspects may be emphasised. The benefits of these different aspects are cumulative.

Neither this paper nor its companion papers are works of history. We shall concentrate, therefore, on where we are now rather than on how we came to be here. Nonetheless, if some of our readers recognise some of what is discussed under different guises, this is wholly unsurprising; and much of our own terminology will predate the use of the phrase 'SOI'.

The essence of an SOI is the delivery of ICT infrastructure (i.e., compute, storage and network) as a set of services. We take as our starting point the three layer model introduced in the figure in the Editorial to this issue of the BT Technology Journal, which is reproduced here as figure 1. While the motivation for this model was briefly explained in the Editorial, we will expand on this here.

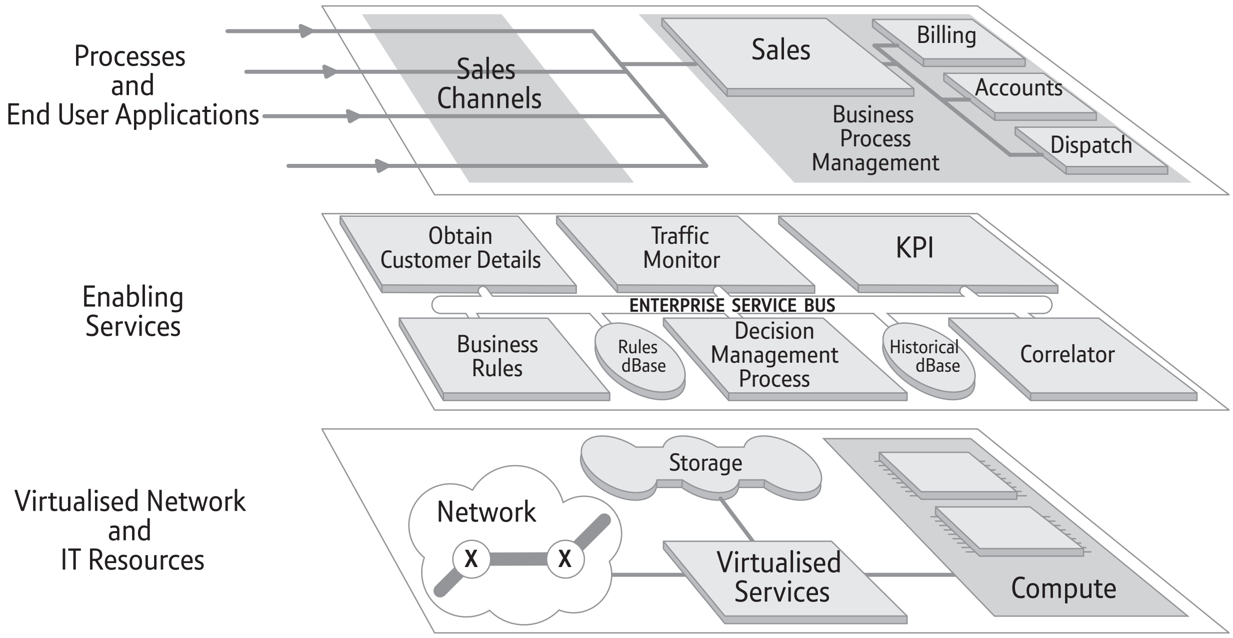

Figure 1 The three level model for a Service Oriented Infrastructure

In the bottom, or virtualisation layer, all components of the ICT estate – i.e., compute resource, storage and network, are virtualised. By virtualisation, we mean that applications using these resources need not be concerned with the details of their physical instantiation – e.g., the specification or location of the particular processors used. Instead, the physical characteristics are said to be hidden, and the application interacts with a set of abstracted logical resources. The physical allocation of the logical resource is unknown to the application and may change – e.g., for reasons of load scheduling or to cope with hardware failure.

An organisation could potentially adopt its own self-sufficient SOI infrastructure and draw benefits from this. The virtualisation of all of an organisation's ICT resource would enable a sharing across all its applications. This makes for an increase in efficiency. Frequently organisations purchase dedicated hardware for particular applications, perhaps associated with a particular department. This can lead to servers being utilised only 15 to 20 per cent of the time [3]. The virtualisation and sharing of resources enables significant efficiency increases, particularly where the applications may have different patterns of usage over time. Hardware failure can also be accommodated, either because there is spare resource elsewhere in the ICT estate, or if necessary by taking resource from lower priority applications.

Of course, organisations very rarely own their entire ICT infrastructure, certainly not the network aspects. Moreover, there are real advantages for an organisation not to own its own IT resource. Apart from the economies of scale achieved through effectively sharing with other organisations, the user organisation is released from concerns about maintenance contracts, upgrades and, depending on the SLA, coping with fluctuations of demand. The organisation is also spared the management overhead associated with these concerns, the accommodation costs, and having to employ the necessary skilled staff.

One possibility is for the organisation to rely entirely on obtaining its IT resource from elsewhere. A well-established version of this is IT outsourcing, where the organisation's IT resource is owned by an outsourcing specialist, which also employs the staff who support the resource. A more modern variant, in which applications are shared between many customers and delivered as a service over the internet, is a model known as Software-as-a-Service (SaaS). Perhaps the most famous example of SaaS is Salesforce.com[1] which offers customer relationship management software through the SaaS model. Google Apps, offering applications such as document creation and editing over the internet, is another example of SaaS. Alternatively, and also a recent development, organisations can obtain compute or storage resource as services over the internet. Amazon's EC2 (Elastic Cloud Compute) is an example of the former, while its S3 (Simple Storage Service) exemplifies the latter.

In another scenario, which represents a half-way house between an organisation owning all its IT resource and no IT resource whatsoever, the organisation owns sufficient IT resource to accommodate normal working. It then has a contract with a service provider which operates a virtual data

[1] http://www.salesforce.com

centre covering compute, storage and network elements[2]. Resources from the virtual data centre are made available dynamically as required by the customer organisation. This may be to cope with failures in the customer's own resource, or with peak loads, or the difficulty of predicting IT requirements already discussed in section 2. The service provider might also provide the network resource. This is shown in figure 2. Note that, because of the need to make the resources of the virtual data centre available dynamically, both the customer's data centre and the service provider's virtual data centre must be seen as part of a virtualised ICT infrastructure. With this architecture comes the need for more flexible security mechanisms, so that access to resources is provided when appropriate and denied when not appropriate. This theme is returned to later in our paper and examined in detail in a paper in the next edition of the BTTJ [4].

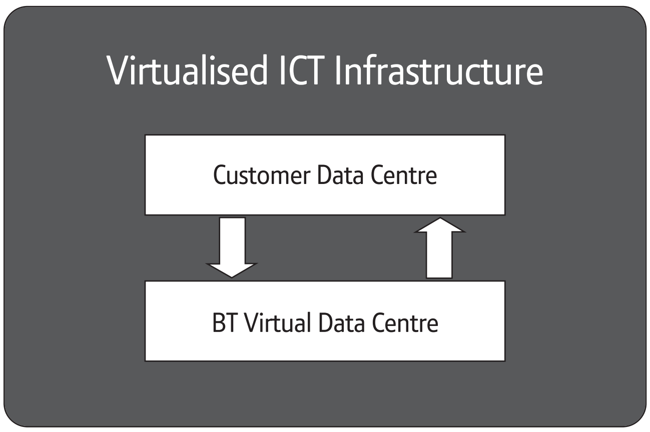

Figure 2 Customer's data centre and BT virtual data centre within a virtualised ICT infrastructure

The middle layer of our model shown in figure 1 contains the enabling services which ensure the provision of the infrastructure in a way consistent with agreed SLAs. The key component of the middle layer is an Enterprise Service Bus (ESB) whose primary role is to provide a guaranteed messaging service. The overall design philosophy is that of an SOA. A promise of SOA is an end to stove-piped systems, i.e., separate software systems for each application area. Previous attempts to reduce stove-piping through reuse have often failed, in part because of tight-coupling and absence of standards. Despite these previous failures, there is now a high chance that SOA, with its loose-coupling and use of web-service standards, will succeed.

The middle layer also contains other supporting functions. It will certainly contain a resource orchestration function to manage the provision of resources in line with the SLAs. This includes, of course, reacting to emergencies such as hardware failure or peaks in demand; increasingly it will include predicting and preventing emergencies through the use of business rules and adaptive learning. The middle layer will also contain security functionality to ensure correct access to resources in such a distributed environment.

Moreover, by monitoring the messaging channels on the bus, it is possible to provide a view of performance. Indeed, it is possible to provide multiple views. In a paper in the next edition of this Journal [5], Deans and Wiseman describe the construction of two dashboards. One dashboard provides a real-time operational view, such as would be useful to someone administrating the resource. The other dashboard provides detailed historical information and allows the identification of process bottlenecks and the optimisation of processes. An important point to note is that the availability of software probes, monitoring the messaging streams on the ESB, means that such dashboards can be realised without creating invasive probes in either the enabling services in the middle layer or the applications in the top layer. As well as feeding into dashboards to provide a view of system performance, the information from these software probes can also be used for automatic optimisation of resource utilisation.

The bottom and middle layer of our architecture constitute the Service Oriented Infrastructure. In summary, they make available ICT infrastructure as a set of services in an assured way. The consumers of this infrastructure – i.e., the business applications and processes, inhabit the top layer. The top layer will also make use of software services as components of the applications and processes. These are frequently available as web services, either via the enabling middle layer or more directly.

A major challenge at the top layer is to use the services available to flexibly create, modify and manage business processes; in particular to do this such that the business person can be as much as possible involved in the creation and modification of these processes, and the IT person as little as possible. Business Process Management addresses just this challenge, and BPM is now an even more frequently used term than SOA. The standards body OASIS [3] has defined the Web Services Business Process Execution Language (WSBPEL) which enables "users to describe business process activities as Web services and define how they can be connected to accomplish specific tasks".

WSBPEL is an XML-based language for describing business processes. A more intuitive language is required if business process definitions are to be intelligible to business people, and to be easily created and edited. To this end, the Business Process Modelling Notation (BPMN) [4] is a graphical notation developed by the Object Management Group [5]. Processes can be described in BPMN – e.g., using packages such as Microsoft Visio, and products exist to translate such descriptions into executable processes.

[2] Note that this is different from the concept of Grid computing, where computing resources are shared amongst several (in general more than two) organisations, e.g. in a virtual organisation.
The virtual hosting environment, described in section 6, is an example of Grid computing.
[3] http://www.oasis-open.org
[4] http://www.bpmn.org/
[5] http://www.omg.org/

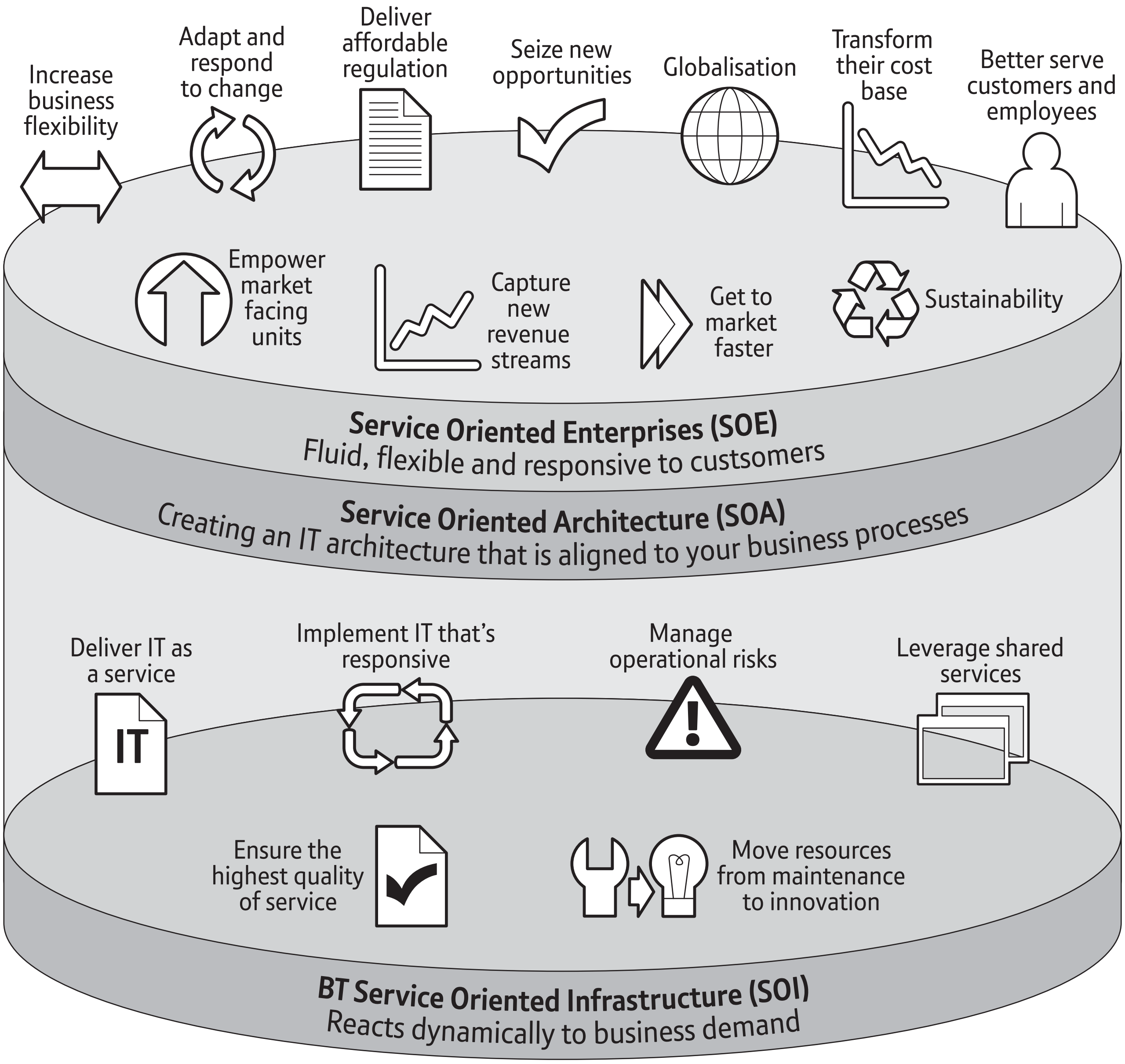

Figure 3 Supporting the Service Oriented Enterprise

Figure 3 illustrates how, building on an SOI such as that being developed by BT, a customer's deployment of a Service Oriented Architecture can support the changing environment within a customer organisation – creating the Service Oriented Enterprise.

4. The market view

The coupling of new technologies with new business models (for example Google, Amazon, Salesforce.com) has resulted in a number of alternative delivery approaches which are aimed at delivering technology as a service. Many of the traditional elements of technology, such as the infrastructure or the application become hidden by the delivery of all, or part of, the business process. As such one-to-one communications become one-to-many and vendors focus on configurable but standard solutions which have high repeatability.

In such an environment, suppliers face a number of inhibitors:

- significant front end investment, often in technologies and solutions whose level of market demand is unproven, so payback timeframes are difficult to calculate with certainty;
- fear of commoditisation and lack of differentiation;
- cannibalisation of current revenue streams;
- untested pricing models;
- disintermediation from the client.

The size of the prize and the fear of being relegated to a commodity vendor, however, is enough to convince many that

£ Million	2006/2007	2007/2008	2008/2009	2009/2010	2010/2011	CAGR 07/08 – 10/11
Business Process management	8,250	8,994	9,757	10,632	11,700	9.2%
Applications and Services	7,143	8,894	10,883	12,938	15,018	19.1%
Common Capability	322	358	396	436	479	10.2%
Datacentre	41,341	44,426	47,421	50,599	53,935	6.7%
Saas	5,283	5,702	6,152	6,615	7,028	7.2%
End to end Application management	9,946	11,007	12,035	13,121	14,282	9.1%
Pull through network	23,363	27,811	32,244	36,366	40,276	13.1%
Professional Services	6,088	6,698	7,310	7,960	8,608	8.7%
Service Management	19,875	21,637	23,491	25,512	27,759	8.7%
Total	**121,610**	**135,526**	**149,689**	**164,179**	**179,085**	**10.5%**

Table 1 Total global revenues for web services and related network and professional services (Source: BT Global Services Market Sizing Team)

the risks are worth it. Table 1 shows BT's view of the worldwide market size (£m) for web services and related professional and network services. Key elements to note include the CAGR across the period of 10.5 per cent, roughly double the CAGR we predict for the overall IT Services marketplace. Also of interest is the rapid growth of applications and services and – the 'golden bullet' for any network provider looking at this space – the size of pull through core business network revenue.

In addition, we predict that virtualisation as a percentage of the total market will increase rapidly across the forecast period, from approximately 13 to 21 per cent by 2010/11.

Besides operating in a higher growth and value area of the IT services market, vendors experience a number of other drivers, including:

- the opportunity to expand their market reach to include SMEs (Small and Medium-sized Enterprises);
- 'pull through' of either traditional revenue streams or higher margin revenues such as professional services and consultancy;
- the ability to standardise and industrialise services, making them more applicable for offshore delivery and therefore more asset efficient.

Due to these market forces, the supply side becomes increasingly complex. New players enter the market based on a specific USP (Unique Selling Point) and not necessarily longevity of IT expertise. Partners become competitors and vice versa. High-growth suppliers from emerging markets quickly develop the skills and expertise to challenge existing dominant vendors. In all, the skills and competencies required to adequately address customer needs and technology capabilities change, and so vendors need to be more open minded regarding acquisition and partner strategies.

Meanwhile within the organisation, particularly the large IT-dependent organisation, SOA is being adopted as the architectural philosophy. This is especially true in certain sectors, such as:

- pharmaceuticals;
- consumer packaged goods;
- high technology;
- government and health;
- finance.

5. Emerging business models

We turn now to look at how service-oriented technology is encouraging the development of new business models, through offering a new way of delivering services across IP-based infrastructures. Applications considered range from existing mass multimedia services, to more complex and demanding customised industrial applications.

5.1 BEinGRID – drawing on the experience of 18 business scenarios

This section describes the work of a project, BEinGRID, which has developed eighteen Grid-related commercial scenarios, or business experiments, each supported by a prototype system[6]. We can regard Grid technology as a precursor of SOI; in both cases resources are not subject to centralised control. Hence the lessons learned here are generally appropriate to our SOI model. The scenarios were drawn from a wide range of industries, and a goal of the project is to analyse the application of the technology in different industries and using different business models. The scenarios are real-life in the sense that the intention is to enter the market immediately upon the successful completion of the project. The scenarios cover industries from automotive and film to financial and ship building. They include companies from the whole Grid services provisioning value chain: resource providers, integrators, service providers, end-users and so on.

[6] The eighteen business experiments reported on here were developed during the first phase of the project. At the time of writing a further seven business experiments are about to start. The intention is that these seven will utilise technology developed in response to requirements identified in the first eighteen business experiments.

An initial analysis of these scenarios indicated that the business cases of the business experiments can be distinguished in three different categories with explicit characteristics: business cases with a clear performance-associated benefit; with a highly collaborative benefit; and exploiting new software paradigms.

1. Grid business cases with a clear performance-associated benefit

These scenarios are aimed at addressing one of the following requirements:

- additional CPU power for executing a demanding application (typical high-performance computing scenario);
- huge amount of data storage;
- access to heterogeneous, geographically distributed data resources.

2. Grid business cases with a highly collaborative benefit

The benefit here arises from sharing complementary resources among participating organisations. In this case, the resulting benefit from Grid adoption comes from sharing data, power and resources utilised for a common scope. This can be both intra and inter-organisational, the expected economic benefit in the latter case could be shared among all participants. This contrasts with the first category where the main economic benefit is anticipated to the end-user. Also, the services in this category cannot be provided by a single provider since data or other resources need to be obtained from other providers.

3. Grid business cases exploiting new software paradigms

This category comprises those business scenarios exploiting new software paradigms such as SOA or component-based development along with new models for provisioning of services such the pay-per-use or SaaS.

The distribution of the business experiments by sector and by business benefit category is shown in the following tables. The tables provide a short description of the topic area of each business experiment, with the number of the experiment shown to facilitate reference to the BEinGRID web-site[7] where more information is provided. Table 2 shows the primary benefit category. Often there is a second, longer-term benefit, and this is shown in Table 3. For example, a company offering in the near future an application on a Grid infrastructure might later consider offering the same service on the SaaS model using entirely their own infrastructure. They might even use a combination of their own infrastructure complemented by Grid resources in periods of high demand.

Sector	Performance benefits	Collaboration benefits	New business paradigms
Advanced Manufacturing	1) Computational fluid dynamics and computer aided design 14) New product and process development	8) Integration of engineering and business processes in metal forming	16) Ship building
Media	2) Business workflow decision making	9) Distributed online gaming	3) Visualisation and virtual reality
Financial	11) Risk management		4) Financial portfolio management 15) Virtual engineering workplace for financial e-services
Retail / Logistics	5) Retail management 12) Sales management system 17) Logistics and distribution	10) Collaborative environment in supply chain management 13) Textile grid portal	
Environmental / e-Science	6) Groundwater modelling	18) Seismic imaging and reservoir simulation	7) Earth observation

Table 2 Categorisation of BEinGRID Grid business experiments by sector and primary business objective

[7] http://www.beingrid.eu

Sector	Performance benefits	Collaboration benefits	New business paradigms
Advanced Manufacturing	16) Ship building		1) Computational fluid dynamics and computer aided design 8) Integration of engineering and business processes in metal forming
Media			2) Business workflow decision making 9) Distributed online gaming
Financial	4) Financial portfolio management	15) Virtual engineering for financial e-services	11) Risk management
Retail / Logistics		12) Sales management system	5) Retail Management 10) Collaborative environment in supply chain management 13) Textile grid portal 17) Logistics and distribution
Environmental / e-Science	7) Earth observation		6) Groundwater modelling 18) Seismic imaging and reservoir simulation

Table 3 Categorisation of BEinGRID business experiments by sector and secondary business objective
(N.B. business experiments 3 and 14 were not assigned secondary business objectives)

As can be seen from the tables and is illustrated graphically in figures 4 and 5 below, the most popular category in the short term (table 2) is that of benefits related to performance, whereas in the longer term (table 3) the most popular benefit category is that of new business paradigms. Compared to previous results from the same project there has been a move towards more complex business scenarios and models where new software paradigms are being exploited. This reveals a higher level of trust gained towards models such as SaaS provisioning. This has been gradually built following the deeper understanding of the technological aspects and the interaction with other business partners as the project progresses and the first results were evaluated. Furthermore, performance benefits are the most common basis for a business case in the short term, which demonstrates that Grid technology is most commonly seen by the industry as a very promising solution to overcome performance-associated problems at a lower cost than today.

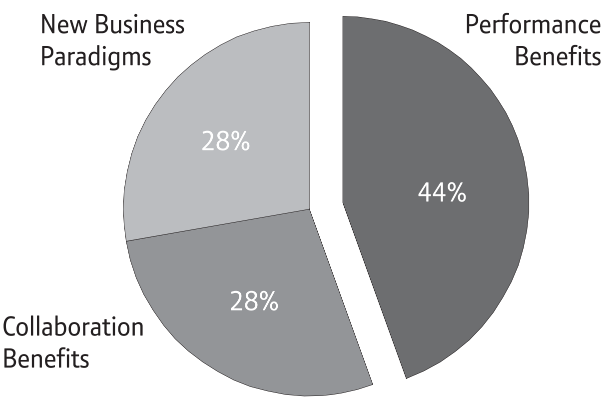

Figure 4: Primary benefit category distribution

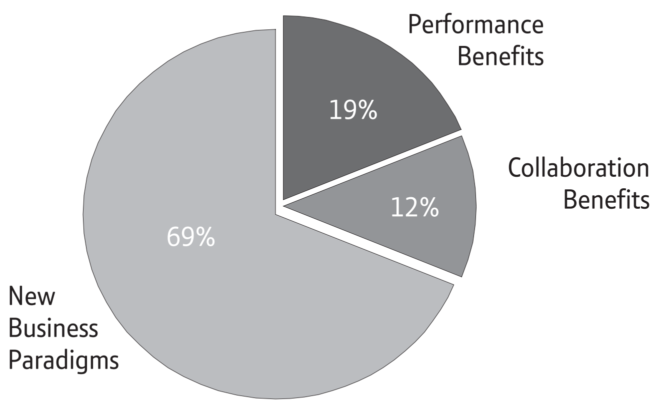

Figure 5: Secondary benefit category distribution

5.2 Economic issues to be considered while making a business model selection

Recent studies [6] and European initiatives [8] have indicated a reluctant and slow take-off of Grid technology by the industry, something due mainly to economic and market barriers rather than to technological ones. The three different cases discussed below point out some of the issues which need to be considered before choosing a business model [7].

5.2.1 Optimisation of processing power in a single organisation

A single organisation may require processing power that cannot be provided by means of stand-alone machines. By interconnecting machines in a Grid, high processing power can be used even by a single application. Thus, the organisation achieves both a high peak processing capacity and a high average utilisation of the processing power available. These features also lead to increased cost-efficiency for the infrastructure deployed. This is particularly important for a large organisation with several departments scattered around the world, each possessing its own local infrastructure. Interconnecting these in a Grid attains a performance enhancement, high exploitation of resources, cost-efficiency and economies of scale; all due to the fact that interconnection of all machines improves utilisation of each individual one. Moreover, the whole approach is scalable, since Grid middleware provides automatic load balancing and transparent usage of the hardware. Besides this, if the various departments possess complementary infrastructure, then the organisation also attains economies of scope. Economies of scale in production arise if the cost per unit of production declines with the number of units produced. Whereas economies of scale apply to efficiencies associated with increasing the scale of production, economies of scope are efficiencies associated with broadening the scope of the service(s) offered, and of their marketing and distribution etc.

When a Grid belongs to a single organisation, a centralised approach to resource sharing is always an option. On the other hand, particularly if there are multiple departments in the organisation, with some notion of autonomy (e.g., own infrastructure and IT budget), then self-management of the Grid by means of economic and market mechanisms is possible and probably preferable. Indeed, the centralised approach requires complete information, which is not always straightforward to gather in a highly distributed system. On the other hand, a market mechanism defining prices for accessing and using the Grid resources by the various departments provides the right incentives for rational usage and results in shaping of demand according to the actual needs; prices may either be monetary, or virtual ones with each department being allocated a Grid virtual budget. This approach also requires accounting functionality, e.g., for monitoring the usage of resources by the various departments and assigning the relevant charges, as well as specification of the right SLAs and appropriate tariffs for them.

5.2.2 Sharing of complementary resources in multi-provider environments

Consider a group of organisations, each of which possesses its own resources, which are complementary to each other. For example, organisation A possesses a powerful database server, while B has a huge amount of data and C possesses an application running over its server that requires data such as that of B. Clearly, when collaborating through a Grid, organisations contribute towards creating a powerful outcome. Each of them exploits its own resources in a cost-efficient way, without needing to invest in the missing resources that are now contributed by others. In this case, the collaborating organisations enjoy economies of scope, since bringing all resources together by means of Grid broadens their scope of applicability. Organisations with complementary resources may also come together to form a virtual organisation [9] (VO) serving third parties. The formation of VOs has a considerable impact on the market. If the group forming the Grid is not closed, then network externalities and economies of scale may arise when new organisations join the group, thus enhancing the associated gains per participant. Network externalities are the effects on one user of a product or service, of others using the same or compatible products or services. For example in a VO structure whose participants share complementary resources, the benefit for the new member strongly depends on the total number of members.

Regarding self-management, the collaboration of the participants in the Grid should be regulated by means of market mechanisms that provide them with the right incentives to both contribute to the Grid the resources promised and not to free-ride those of the others. For example, a global agreement can prescribe that all contribute as necessary. Similarly to agreements used in P2P (Peer-to-Peer) systems, such agreements can be based either on rules prescribing a fixed minimum contribution for all participants, or on rules regulating the consumption levels of each participant (quantitatively or qualitatively) in accordance with the contribution over time. These rules should be complemented by accounting functionality that certifies conformance. Also, an internal market mechanism, based on SLA and monetary prices for these SLAs can be employed as an effective approach for self-management, particularly in cases where the level of contribution of the various participants is not symmetric, and agreement on cost allocation can be hard to reach. These ideas apply to the case where the Grid is formed in order to serve the participants' own needs, including the case of a single organisation with multiple departments. If the participants also serve third parties, then the relations between the former and the latter should also be managed by means of market mechanisms.

---

[8] http://cordis.europa.eu/ist/grids/index.html

[9] A virtual organisation comprises a group of independent organisations which come together temporarily to achieve a particular goal.

5.2.3 Offering utility computing services

The model here is one offering applications (software) and computing services (hardware) on a pay-per-use basis rather than by means of licensing or long term static agreements (leasing, etc.). Applications are sold as components to be used as part of an SOA; customers can design their full solution by combining components from different providers and run these on their own premises or using some Application Service Provider (ASP)[10] computing services. Essentially, this application-level Grid allows for a new version of the application based on components to be accessed by the customers. This version is more affordable to infrequent users of the application, when compared with purchasing a software licence and computational infrastructure. Both users and the service provider gain, as this model increases the demand for the service by making it more affordable (i.e., at a lower cost). At the lower layers an ASP may benefit from complementing its own infrastructure with utility computing services offered from third parties. The issues discussed in the previous examples regarding high performance, economies of scale and scope, etc. are still applicable here.

Interesting economic issues arise from the model. In particular, we now have a new market (that of the pay-per use application), in which the proper SLAs should be offered to customers; resources should be self-managed and the revenue should be properly distributed to the players involved. The market also has significant impact on other markets. In particular, an SME that cannot afford investing in a licence or in infrastructure obtains new capabilities. This enables the SME to enter new markets. Therefore, the Grid version of the application leads to a reduction of the barriers of entry in other markets, which then become more competitive.

5.3 Lessons learnt from BEinGRID: guidelines for a successful business case

We summarise our evaluation of the business experiments by emphasising a number of issues that that should be addressed carefully in a business plan by organisations considering using Grid or service-oriented technology.

Setting short-term goals and producing a market entrance strategy is not enough. Concrete and realistic long-term goals for commercial exploitation should be set. Furthermore, the technological advantages and how these are translated to business advantages should be clearly stated and explained. Also, exit strategies at various points should be prepared.

Think bigger! Even from the starting point of the business do not underestimate the power of economies of scope. Analyse the applicability of your solutions in the future to other sectors rather than restricting the focus to one.

Strategies and goals should be associated with specific tangible success criteria and metrics that clarify the uniqueness of the expected results and will evaluate the business goals set. Examples are user satisfaction, system expandability, number of successful transactions and users etc.

A good risk and SWOT analysis should accommodate both technological and business risks and interconnect them where possible.

It appears that there is clearer view on future resource requirements in cases where the targeted market is already established and its future momentum can be predicted (e.g., in more technologically advanced sectors). However, the business analyst should be very careful in predictions involving more established and/or ‘traditional’ industries, such as the textile and logistics ones, where the adoption process of a new technology is rather slow.

Building strong alliances (especially if you are an SME) is another key area to be addressed closely, as one of the advantages of this new technology is exploiting the capabilities of collaboration. Furthermore, we have found that communication and promotion is poorly addressed in many plans.

A common feature of the existing business plans from the business experiments is the fact that no complete or considerable surveys exist to assess the market potential. The market potential should not be limited to the potential of the new technology but the potential of the whole IT solution and the associated complementarities should be identified. At the same time, we need to identify the difference that the use of Grid technology or service oriented computing makes to the business scenario in which it is deployed.

6. A new business model – the virtual hosting environment

6.1 Sharing resources in a virtual organisation

We consider again the bottom layer of our three-layer model. The assumption so far has been that the ICT infrastructure is either owned by the organisation using the resource, or else by a service provider. Another possibility is that the resources are owned by several different organisations. This might arise when a number of organisations come together to form a consortium, or virtual organisation, for some business objective, and wish to share resources. It might also arise when a service provider enables spare resource from a number of organisations to be made available to a third party. This scenario, known as that of a virtual hosting environment (VHE), is illustrated in figure 6. Here, the virtual data centre of figure 2, has been replaced with the resources of a number of different organisations, through the mediation of a service provider known as a VHE operator.

In summary, a virtual hosting environment is a federation of distributed hosting environments for execution of an application and for providing a single (logical) access point to this set of federated resources. A VHE is from the commercial viewpoint futuristic; this section will describe a prototype.

[10] An ASP offers software-based services to customers over a network, often the Internet. It provides SaaS-type services, and the term SaaS is now more commonly used than ASP.

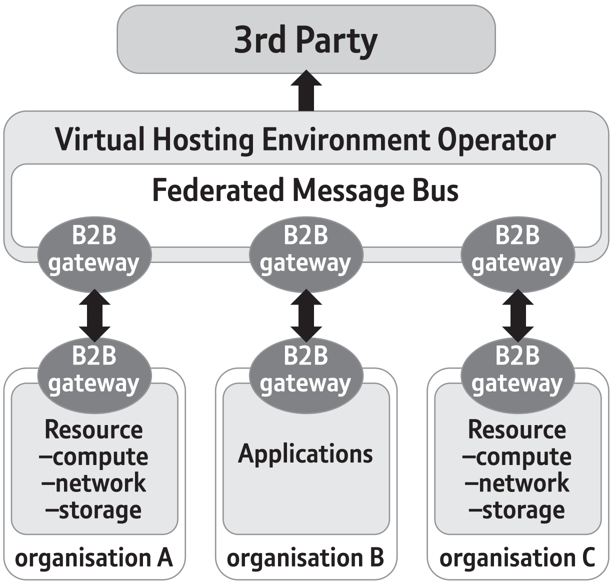

Figure 6 Virtual Hosting Environment

Both the virtual organisation and the virtual hosting environment demand even more flexible security enforcement than the virtual data centre of figure 2. Security policies need to take account of individual messages and also of the context – i.e., they need to be content and context-dependent. An example of the importance of context arises when two organisations collaborate together in two different virtual organisations. Security policy enforcement needs to know to which of these two collaborations a particular resource request refers. In another example, the security enforcement may need to know the particular device from which a request comes, some devices presenting more of a security risk than others. A strategy where security is no longer a question of a firewall around the organisation but of taking account of individual data packets and individual devices is known as deperimeterisation. As already promised, more information will be provided on the security requirements of SOI in the subsequent edition of this Journal [4].

6.2 Case study – a virtual hosting environment for distributed online gaming

A previous section gave an overview of 18 business experiments in the BEinGRID project. One of these is looking particularly at the concept of the virtual hosting environment. The particular industry sector chosen is that of distributed online gaming. However, the underlying concepts, technology deployed, and lessons learned are relevant to a wide range of industry sectors and application areas. The business experiment is led by BT, with four additional collaborating partners[11].

A VHE allows an organisation to access more computing resources than it owns itself by using other organisations' resources as and when required. This is relevant to the online gaming industry because multiplayer gaming can require large amounts of computing power that may be above and beyond the resources of the game provider, particularly if this company is an SME. The ability to make use of spare resources from other organisations would enable the smaller players in the online gaming industry to have access to increased computing infrastructure at reduced costs. In turn this would enable them to compete with the larger players on a more level playing field. This would be good for the industry and for consumers, and provide a new revenue stream for the owners of the computing resource. This is an example of how the deployment of a virtual hosting environment can alter the structure of an industry.

This business experiment addresses the challenge of how to use the recent advances in the area of SOI, in order to build and assess a distributed application hosting environment that enables network-centric ASPs to rapidly deploy and manage their services within a VO. Moreover, this needs to be done in a secure and accountable way and in compliance with explicit agreements and policies.

The business experiment separates the concerns of the game platform provider, game title provider, and the network-centric infrastructure provider – e.g., the application host. It enables the platform provider to gain competitive advantage by choosing a preferred games title provider, and deploy the game titles on a preferred hosting environment potentially offered by a different provider. The game platform is exposed as a set of virtualised services and utilises resources from a virtual organisation over a network-centric SOI, thereby outsourcing games title provisioning and game application hosting. From the perspective of a network-centric infrastructure provider, our approach enables economies of scale. The results will be of general use across multiple application areas, in addition to the internet-based interactive gaming services that have been chosen to prove the concept.

6.3 Impact

The main disadvantage of the original games platform was that it required a large investment by game operators and portals in computing hardware, limiting the number of possible customers. Through sharing resources, our approach materially reduces the investment required by games and portal operators. Also, our approach offers the possibility of creating virtual organisations, where game operators and portals can share games and contents, and even their user base. Technically, the goal is to be able to share expensive resources between providers and to allow billing based on usage. From a business perspective the goal is to open new commercial opportunities in the domain of online games.

A common problem with online games is that while operators, portals and games providers would like to share resources - and by doing so share costs – separate business entities often end up being self-contained stove pipe solutions trying to do everything themselves. The European market is

[11] The partners are: Andago, a Spanish gaming company; Atos Origin SA, an international information technology services company; Centro di Ricerca in Matematica Pura ed Applicata (CRMPA), an Italian research organisation; and Universidad Rey Juan Carlos (URJC), a Spanish university.

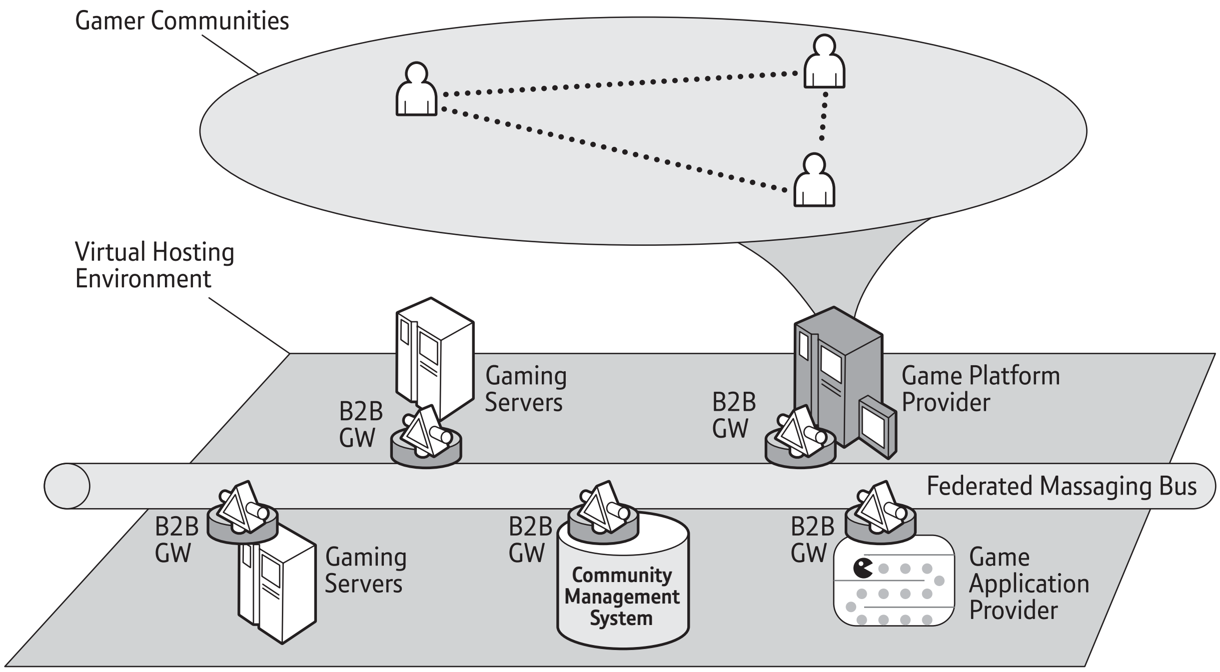

Figure 7: High-level overview of the VHE for online gaming.© BT Group plc. N.B. B2B GW denotes 'business-to-business gateway'

fragmented, making it hard to reach the critical mass needed to make online games businesses profitable and hard to ensure computing resource availability on demand when required by gamers. Having a shared infrastructure will make it possible to divide the tasks among different providers and allow each provider to concentrate on its core skill. In this scenario, application developers provide the applications, portal providers create the portals, and ISPs provide the communications infrastructure. Such arrangements should allow for lower costs by enabling better resource utilisation and therefore more profit. The virtual hosting environment architecture provides the middleware to make this collaboration possible; thereby not only offering businesses lower costs, but encouraging the creation of a pan-European market.

There are also big advantages for users. They should have more games to choose from and better quality of service. Moreover increased competition should lead to better and cheaper games. Grid-based centralised portals could provide access to a wide choice of games and other entertainment content from different providers. Today, if a user buys a new game and wants to play it online, the user has to connect to a server which, unless a local server has been set up, is likely to be in the U.S.A. Deploying a virtual hosting environment will simplify that process - users will simply connect to the games portal and join the community of users, wherever they are. This is further evidence of the point already made, that the deployment of a virtual hosting environment can significantly alter the structure of an industry.

6.4 The architecture of a virtual hosting environment

The VHE developed by BT for this pilot, as illustrated in figure 7, consists of a network of B2B service gateways integrated with common capabilities for trust federation, identity management, access control, SLA management, accounting and monitoring, as well as application service and resource virtualisation. The B2B gateway functionality is complemented by a federated messaging bus and community management services that facilitate the establishment of B2B collaborations (e.g., in the form of virtual organisations [8]).

To initiate use of the VHE, a network-centric application provider (which in the application example used for this business pilot is an on-line collaborative game platform provider) engages in a contract with the VHE operator that allows the application provider to use other applications, resources and infrastructure services. These can be offered by the VHE operator itself or by other parties via the VHE operator. In our example the game platform provider uses game titles from a game application provider and game servers offered by other parties such Atos Origin, a partner in the business experiment, and BT.

The game platform provider can then initiate the creation of a VO that allows it to create instances of a game title from the game application provider on the game services offered by a data centre. In order to offer online games to its gamer communities, Andago, the games platform provider in our business experiment, uses computational resources in Spain from Atos Origin.

As Andago's customer base expands they decide to expand their VO by adding a contract with another data centre operator (say BT) and introduce more game servers. All being well, Andago's decision to expand will be in response to customer demand (say) from the UK where BT data centres can offer a better Quality of Service (QoS) than more remote facilities.

This architecture allows the business intelligence of a VHE to be enhanced – e.g., by using advanced identity management and distributed access management services. These would be offered to the VHE by suppliers, which could include the VHE operator but could also include third parties. Such network-hosted services allow the definition of profiles, using standards-compliant identity assertions, and access control policies. The different partners in the VHE may choose to use different identity service providers. The VHE infrastructure ensures compatibility between the different infrastructure services in place.

Different customer relationships over the VHE may employ different charging models. For example, data centre operators may charge the game platform provider following a 'pay-per-use' model. On the other hand, the identity service provider may charge on the basis of the size and duration of the virtual organisation. Finally the VHE operator may charge the game platform provider on a 'pay-as-you-grow' fashion based on the portfolio of VHE capabilities that are made available; charge the data centre operators based on a percentage of their resource utilisation via the VHE; and charge the identity service provider a flat fee based on the number of customers gained. Such dynamics require the VHE operator to offer very flexible accounting mechanisms in order to allow the various stakeholders to correctly retrieve and correlate chargeable events.

7. Conclusions: the future of our industry

This paper began by describing the priorities of corporate ICT departments today as they adapt to the changing business and technological environment. Amongst those priorities was the need for efficiency, including efficient use of energy resources, and at the same time resilience in the case of hardware or other failures. We have seen that an SOI offers this efficiency, through its virtualisation of resources and through the ability to share resources between organisations. Increasingly, companies also want to collaborate across organisational and enterprise boundaries. The SOI approach supports that collaboration by enabling the sharing of resources, virtualised and exposed as services. Companies want ICT infrastructure, software and information available on demand, and they want to pay according to their use. The service delivery model embedded in SOI satisfies just this need. Finally, they want to create exciting new applications through the integration of all their ICT infrastructure and applications; yet at the same time they want to disaggregate monolithic systems to access reusable components. This is exactly what the concept of the SOA was created for.

We have also shown how the market for ICT products and services is already reflecting the service-oriented model; and through examples drawn from current research we have given some indication of how business cases for the new technologies can be constructed. We have also described a new business model, that of the virtual hosting environment. In particular, we have explored how that business model is being implemented in the online gaming industry. Of course, business models are not created in laboratories, but in the commercial world. At heart the service-oriented approach is about a shift in ownership. Organisations no longer need to own resources in order to benefit from them. They no longer need to face heavy up-front costs and ongoing maintenance costs. They can pay for what they need, when they need it, and take delivery as a service across the net. As this approach is increasingly deployed, we shall see it exploited in inventive new business models.

Our vision for the future of the ICT industry is one where service delivery across organisational boundaries will enable all manner of new business applications, drawing on the integration of many component pieces, with the applications themselves delivered to the end customer as services. We hope that our paper has given the reader a glimpse of the richness of that future.

Acknowledgement

The BEinGRID project is funded in part by European Commission Grant #IST-2005-034702.

References

1. IDC Service Oriented Conference, April 2007
2. Manyika J M, Roberts R P and Sprague K L, 'Eight business technology trends to watch' , The McKinsey Quarterly, December 2007
3. Kusnetky D, 'Application virtualization and utopia: proving the value of virtualizing applications', Kusnetzky group, 2007, http://www.kusnetzky.net/publications/ImpactPapers/20070623e_Application_Virtualization_and_Utopia.pdf
4. Dimitrakos T, 'Securing business operations in an SOA',to be published in BT Technology Journal, vol.26, no.2
5. Deans P and Wiseman R, 'Examples of SOI in action', to be published in BT Technology Journal, vol.26, no.2
6. Forge S, Blackman C, 'Commercial exploitation of Grid technologies and services, drivers and barriers, business models and impacts of using free and open source licensing schemes', SCF Associates for DG Information Society and Media, 2006
7. Thanos G A, Courcoubetis C and Stamoulis G D, 'Adopting the Grid for business purposes: the main objectives and the associated economic issues', The 4th International Workshop on Grid Economics and Business Models (GECON 2007), Lecture Notes in Computer Science (LNCS 4685), Springer 2007, pp.1-15
8. Dimitrakos T, Golby D and Kearney P, 'Towards a trust and contract management framework for dynamic virtual organisations', eChallenges 2004

Claire Gresty is currently a senior marketing strategy manager for BT Global Services, having joined BT five years ago. During her time with the company she has worked on many strategy projects, often analysing the research and insight gathered via the sales and marketing process. She has initiated and managed many research assignments focused on ensuring customer needs are understood and that services and solutions are aligned to the market requirement. Her strength as a market analyst is the result of 16 years working in the global network services environment and she has also held senior marketing positions at Cable & Wireless and MCI WorldCom.

Theo Dimitrakos leads the Security Architectures Research Group in the Centre for Information and Security Systems Research, BT Group Chief Technology Office. He has 15 years of research and innovation experience in a wide range of topics relating to information security, identity and access management, software and systems engineering, Service Oriented Architecture (SOA), web services and grid computing. He has strong academic background in the areas of security risk analysis, formal modelling and applications of semantics and logic in computer science.

He has been the scientific coordinator or technical manger in several collaborative research initiatives funded by the European Commission and the UK Department of Trade & Industry or Ministry of Defense. He is involved in IFIP working groups on trust management (vice-chair) and enterprise interoperability (member). He is also a member of the editorial boards of international journals and programme committees of international conferences on distributed systems, web services and information and telecommunications security.

Prior to joining BT he was the technical leader of a multi-million collaborative research programme at the Central Laboratory of the Research Councils (UK) and an eCommerce and EDI specialist at Logica (UK). His academic qualifications include a Ph.D. in Foundations of Software Engineering from Imperial College London, and a B.Sc. from the University of Crete in Greece.

George Thanos is a member of the Network Economics and Services Group in the department of computation at the Athens University of Economics and Business (AUEB), where he is pursuing a Ph.D. in the area of economic modelling of next-generation services. George received his B.Eng. degree in Computer Systems Engineering from the University of Manchester (formerly UMIST) and his M.Sc. in Mobile and Satellite Communications from the University of Surrey in the UK. In the past he has been working as a systems engineer in the private sector (Motorola Ltd UK and Teletel SA Greece) involved in the research and development of telecommunications systems for mobile networks, mainly based on UMTS/3G technologies. His current research in AUEB focuses in the investigation of economic issues in Service Oriented Architectures (SOA), grids and next-generation web repositories. In the last years he has also been contributing in a number of European FP6 and FP7 research projects in these areas.

Paul Warren works within the Centre for Information and Security Systems Research at BT. His particular interests are in the areas of knowledge management, semantic technologies and service oriented infrastructures. He is currently Project Director of ACTIVE (http://www.active-project.eu), a European project in the area of collaborative knowledge management. At an earlier stage in his career, Paul worked on technology strategy and technology foresight, investigating areas as diverse as eBusiness and novel forms of computing. Previous experience also included a secondment to the CBI to study government support for technology in industry. Paul has published widely on knowledge management, semantic technologies, and technology foresight, and has collaborated in editing two books on topics related to ICT. Paul holds a degree in Theoretical Physics from Cambridge University and an M.Sc. in Electronics from Southampton University.

[12] http://www.active-project.eu