Preface

INTRODUCTION

In this volume we provide a selection of innovative and creative researchers and practitioners describing their work in the field of Service Oriented Architecture (SOA), Portals, and related technologies. Many of the researchers are concerned not just with the technologies involved, but with experiences using these in practical environments. Before introducing the eighteen articles, it is worth considering one of the key limiting factors in our use of these technologies: choice.

In our development, selection and implementation we are faced with choices, or trade-offs, which we have to make. If we knew everything about everything in our technical environments, we could provide our customers with the perfect experience. However, not only is this goal impossible in practice, it is also theoretically impossible. We have not only the opportunity but the obligation to make choices in designing the monitoring and management of our technical environments. In order to deliver our customers the best possible experience, we need to understand their requirements, and use this as the basis for making such informed choices. This is true for all information technology projects: the heart of project management is the balance between the three competing needs of cost, time and functionality. But in the Internet world the need to make choices is even clearer. The Internet itself is based on a choice that makes it different to traditional telephony. A telephone call will be connected, or not. The service will be available, or busy. The Internet’s core design, however, takes a different approach. Even the most limited or tenuous connection will allow small amounts of information to flow. That can play havoc with a video or voice message, or may be fully sufficient for non-time-critical applications such as e-mail.

Few areas require more choices than user experience monitoring, and this applies to web portals as much as other services and applications. No provider wants a technology issue to impact their customers. Given the sophistication (and cost) of modern systems, it seems that nothing should. The promise is simple using modern monitoring tools. Several vendors now promise that businesses can view one or more dashboards that show the state of a service, warning of any current or potential customer impacts, and allowing immediate issue identification and resolution. Unfortunately, as with many areas in technology, the reality is far more complex. Choices are not only available, but forced onto application owners who expect to know everything about everything. In the next few paragraphs we will spell out a range of choices which we have personally faced in technology service management. This is not a complete list, but gives a reason for pondering before expecting a monitoring miracle. The first group of choices relates to what gets monitored.
• **Synthetic Transactions or Direct User Monitoring:** There are two quite distinct ways of measuring the customer experience of a service or application. The first of these involves tracking the traffic of one or many users. A particular transaction may be identified and then timed as it goes through a system. If the technical and non-technical problems of tagging individual users can be addressed, there is still a significant weakness in this approach: Service failures often occur overnight due to changes, batch processes and other events outside of regular working hours. The use of transaction monitoring cannot determine if a service is working until customers attempt to use the service, which may not occur until the start of a business day, several hours after an incident has crippled a service. A second and quite different approach is to execute “synthetic,” or artificial, transactions. By measuring the response time to service requests, the application is able to determine delays in various parts of the service and the wider network that will impact customers. A significant challenge to the synthetic transactions approach is security. The system initiating the synthetic transaction will require all accesses required by a real user, and for value transactions the system will therefore need to be appropriately secured. Yet to be affective the system will need to initiate the transactions from areas outside of the central protected IT environment: Just because the synthetic transaction generated within your computer centre gets a great response, this says little for a struggling client several thousand kilometres away.

• **All User Access Paths or Most Commonly Used Access Paths:** Closely related to the previous point is the balance between understanding what a typical user may be experiencing, and understanding what every user is experiencing. For an on-line service, the latter would require direct monitoring of every packet exchanged with every single user. This is because of the nature of the Internet’s protocol, TCP/IP. There are no permanent paths used by users: they can change packet by packet (or many times a second). For a service provider in the real world this would be impossible. Increasing the number of access points and paths also rapidly increases the application’s cost.

• **Complete Data or Longitudinal Data:** Gigabit routers have existed for many years. The thought of storing all data going through a device which counts its traffic in billions of items per second for anything longer that a few minutes is difficult to imagine. Yet many business managers like to think that they can keep an exact record of everything (not just key transactions) indefinitely. Determining the right level of abstraction in storing information is key to effective problem resolution.

• **Automated Actions or Determinate Actions:** For over a decade it has been feasible to manage network configuration by allowing service management software to roam a network, interrogating devices and providing either their details or maps based on these details to network managers. This can achieve huge time savings, and provide a level of detail otherwise difficult to attain. Yet allowing software to roam a network searching for devices, and possibly even making decisions about these devices (for example, preparing reports on their status and validity) is a step towards the creation of a non-deterministic network, where the outcome of an action will not necessarily be known in advance. Regardless of the additional delay, effort and difficulty, an organisation which places significant dependence on its networks (such as a bank) may prefer to be able to track the responsibility for every configuration and reporting decision to a specific person, time and place. Automation would then be limited to software that compares on the one hand what and where all devices are supposed to be, and on the other hand, all devices that participate in a network.
Another set of choices relates to reporting, and here there is no single user perspective. Possible reports include among others:

- **A Record in an Incident Management System:** This could include a unique identifier, a timestamp, an author, a title, a short description, a long description, attached information, a log, a set of events, a link to other records, a set of proposed actions, a history of actual actions, text descriptions, numerical descriptions, and many other fields.

- **An Alert Regarding an Incident:** This could be a brief message listing a subset of the incident record sufficient to provide a simple understanding of the alert. It could be character limited, for example to fit into an SMS message. It will often be a dense and cryptic reference, meaningful only to a particular specialist.

- **A Brief Summary of Such a Record or Alert:** While key staff are working with alerts to identify and resolve an incident, information about the incident will also be sought for escalation or communication purposes. Here again brevity is important, but a cryptic message will be of limited value.

- **A Detailed Log:** While the primary purpose of incident management is the rapid recovery of a service, for root cause analysis and problem management large volumes of data may be captured.

- **A Summary in a Daily or Monthly Incident Report:** Grouping together the various incidents occurring in a large information technology environment is necessary for reporting and tracking of key indicators, which may feed into service level reporting, key performance indicators, reporting dashboards, and other management tools.

The purpose of the report will determine what should be reported. In practice, even if there is only one set of customers, there are at least four different stakeholder groups with quite distinct information requirements.

- **The Operational View:** Providing a reliable and continuous service is an operational task. There may be one or more operational centres around the globe providing 24-hour support, or just a single staff member in a small company with responsibility to keep a service running during business hours. Their task is to get the service back up. They will use the standard tools available to them to rerun batch jobs, transfer to disaster recovery environments, restart processes, and call on the next line of support. They need information to do their own jobs, and if this isn’t enough, they also need to be able to identify which of many in-house or external technology support teams they should turn to. They are interested in processes and options available to them, and the alerts or reports they receive to help them in their tasks will be designed with this in mind.

- **The Technical View:** Fixing a technology incident is a technical task, with several specialties often divided among the relevant teams. In any significant incident it is common for two or more issues to be involved, and several or many technical areas of expertise may be required to rapidly identify (or eliminate) particular paths for investigation. While a programmer, database administrator, data communications engineer or middleware specialist may be fairly confident that a problem is or isn’t their particular responsibility, incident recovery is not an exact process. Common technology errors such as a failed disk drive or power supply may be self-diagnosed. When technology specialists are engaged, the problem is often beyond the commonplace. In these
circumstances, unlike in the case of the operational view, it won’t be standardised reports and regular processes that resolve the incident. The report with the most value may be entirely incomprehensible even to technologists in other fields.

- **The Business View**: Here is concern over the specific impact of an incident on customers, with an understanding of which customers are most important, and a desire to minimise both overall and high value customer impacts. Knowing which rule corrupted the firewall will be worse than useless, and understanding which team is working on the problem will be of only limited interest. They will be more interested in knowing how long an incident is likely to last. They will be very interested in the names of specific customers or customer segments affected by the incident.

- **The Assurance View**: Finally, the service will often be linked to contractual or regulatory obligations. There may be a limit on the number of minutes a service is unavailable before service level penalties are incurred. These limits may sit within annual or monthly allowances, may relate to the time of day, may vary for different services, and may be bound by internal company guidelines. For service levels the penalty may be purely financial, while for regulatory obligations a significant outage may result in the loss of license to operate at all.

Each of these audiences will have their own requirements for incident notification. For example:

- **Message Detail**: As soon as an incident has occurred, the business view will be interested in the impact on customers. The other views should be aware of the impact order of magnitude, but the technical view will be looking for hardware, system and application alerts that help isolate the problem. These will probably be incomprehensible to the others. Operations will be more interested in management of the recovery, including any temporary workarounds for customers, rather than the specific technical activities.

- **Message Format**: Closely related to the detail of the message is the way it is formatted. An attractive dashboard may provide a valuable business view. (However, unlike a text alert, a dashboard needs to be constantly monitored if it is the primary means of communicating alert status.) A text alert will be much more useful to a technical support staff member, particularly if it is able to identify the specific device, platform or other item having difficulties.

- **Message Thresholds**: In general it is far more valuable for operational and technical staff to receive advance notice of any breaches of thresholds relating to services than for business staff. For example, information that a database has just passed a capacity threshold (such as “80 per cent”) requiring action is of no interest to business staff.

- **Impacted Groupings**: The technical view will focus on all the applications that pass through a particular firewall or server. Server names will be meaningful to them, network diagrams will provide key relationships, and the extent of the incident will be understood in terms of technology. As operations works to resolve an incident, the range of internal impacted groups who need to be engaged or informed will be of major significance. In the business view, the impacted hardware or other technical element will be meaningless, while key focus will be on knowing in advance which customers or customer groupings will be impacted.

- **Correlation**: In a major incident, there may be many impacted systems and customers. An effective incident management system will group these together and report on impacted groups (eg, all Automatic Teller Machines in a specific region) rather than provide a separate report for each item
that has been impacted. Here again, correlation providing meaningful business information will be quite different to that providing meaningful technical information.

- **Drilling Down:** Sophisticated incident management reporting tools will allow a user to drill down to gain additional information regarding the area of interest. While a technical specialist will be interested that a particular card in a device has failed, a business concern could focus on which a list of branches in a region have been impacted and for how long.

- **Reporting Frequency:** It is within the capacity of modern technology to provide reports at a rate which will overwhelm human capacity. The authors are aware of one incident when a faulty device reported its error state 27,000 times in a few seconds. While this was an incident in itself, frequency of reports may vary greatly for the same incident depending on the user’s purpose. For each of these items, the reporting will vary. An implementation which provides alerting or a dashboard in terms meaningful to one group but not the others will fail to meet actual needs. Understanding the different views, and who needs to know what sort of information, is critical to having a meaningful set of messages.

- **Loss of Service or Degraded Service:** Finally, there is one significant decision which applies to every point made above. Are we concerned just with loss of service, or also with degraded service? Degraded service can refer to delays affecting each transaction, or to loss of some but not all elements in a network (such as the ATMs in a particular region). Alerting a company’s CEO to the loss of connection to all retail outlets makes sense, whereas the loss of one ATM out of 1,000 may be a “business as usual” activity. An additional second for delay experienced for every online transaction may be a minor irritation, while a 30 second delay would render most websites unusable. Where it the cross-over point?

These choices are not insurmountable. With a clear understanding of exactly what we want, we can choose a set of deliverables that are both useful and meaningful. Using a framework such as IT Infrastructure Library (ITIL) to help us understand our service goals and develop our service definitions, we can substantially improve the customer experience. Modern monitoring tools let us get the data we want, once we understand what that data is and how we will use it. Choices don’t have to be a plague, but we do have to make them, if we are to genuinely help the customer.

These general comments provide a backdrop to the specific circumstances, theories, prototypes and experiences captured by the works published in this book.

### Introducing the Contributions

The papers published in here address a range of key issues of relevance to researchers and practitioners in the portal and Service Oriented Architecture fields.

In *A Reference Ontology Based Approach for Service Oriented Semantic Interoperability*, Wang, Brown, Lu, and Capretz consider the interoperability for inter-application information exchanges, sharing functions and service interfaces, but also exchanging data models. Their approach is to represent data models using ontologies. This allows the location and integration of these models. The paper approaches service oriented ontology management using a reference ontology. The authors show how to evaluate a practical case through development of a domain-specific reference ontology. To evaluate and validate this approach they develop a service-oriented system based prototype system for ontology deploying,
browsing and mapping operations. The paper describes their experiments, including promising results consistent with their approach.

In *Lightweight Collaborative Web Browsing*, Santos, Oliveira, Gomes, Martinello, and Guizzardi consider the challenge that while collaborative navigation systems provide a useful way for virtual groups to share information through the web, the common set of features of these tools is not enough to offer a more interpersonal browsing experience. To fill in this gap, they present a novel collaborative web browsing proposal aimed at integrating flexible session management with shared production spaces and efficient communication facilities. The proposal relies on a collaboration ontology that provides a well-defined conceptualization and a common vocabulary. They have developed and tested the OCEAN prototype in order to demonstrate the feasibility of their approach.

In *Improving Collaborations in the Neuroscientist Community*, Mirbel and Crescenzo present an approach they call SATIS (Semantically AnnotaTed Intentions for Services), which relies on intentional process modeling and semantic web technologies and models to assist collaboration among neuro-sciences community members. The goal of this work is to derive and share semantic web service specifications from a neuro-scientist’s point of view in order to operationalise image analysis pipelines through web services.

In *E-Performance Systems: A Method of Measuring Performance*, Al-Raisi, Amin, and Tahir point out that organizations are looking for solutions to manage and maximize the performance of their workforce. These organizations recognize that there has been a shift in the business environment from a tangible asset economy to an intangible asset economy. As a result, the value of a company is comprised of employee knowledge, brand, and intellectual capital rather than just inventories, goods, and machinery. This places a new dependence on technological solutions to monitor and improve employee performance and productivity. Several technological solutions such as Electronics Performance Management Systems (e-PMS) are being used by organisations for such monitoring. The authors assess the impact of e-PMS in the organisational change. In the process they propose a model-based on empirical results. That model can used as a guideline for the organisations while assessing the impact of e-PMS systems.

In *E-Mportfolios*: Challenges and Opportunities in Creating Mobile Electronic Portfolio Systems for Lifelong Learning, Weber and Evans critically examine the developmental trends of mPortfolios and gauge their impact on newer forms of learning that utilise mobility, portability, and flexibility. Placing this study within the emerging paradigm of futures’ thinking, the paper focuses on the environmental factors that shape the direction of portfolio development from electronic to mobile systems using a series of global case studies to illustrate the challenges and opportunities for educators. While mobility and portability emerged as strong elements in design, flexibility remains a key challenge. The analysis also revealed that sector-based approaches to developing mPortfolios through research and Community of Practice structures are potentially more beneficial for mPortfolio developers. Within these approaches clear advantages are to be gained from the communal-dialogical approach found within the Community of Practice approach, which could potentially inform futures’ thinkers in relation to strategic planning and forecasting of new trajectories in mobile and lifelong learning.

In *Computing the Spreading Power of a Business Portal to Propagate Malicious Information in the Network*, Saini, Mishra, and Panda look at a set of multi-business organization or company portals to see which are vulnerable to malicious objects. Graph theory is used to address this problem, looking at patterns of diverted customer traffic among business portals. The interconnected business portals are first represented in the form of a graph and its corresponding adjacency matrix. Later, Eigenvector centrality is computed to find the vulnerability of a business portal to propagating malicious information, which
helps to limit damage by identifying which business portals need to adopt specific security measures to avoid further infection.

In *Advanced Content Management System in Murdoch Research Institute*, Shadlou, Solaymani, and Hajmoosaei describe the Hugh Williamson Gait Analysis Laboratory, a world leader in the analysis of walking disorders in children. Using state of the art equipment the specialised staff are able to determine why children walk the way they do. It is then possible to plan treatment individually for each child. The laboratory aims to achieve this in an efficient and friendly manner. It provides considered interpretations of these measurements to referring clinicians and conducts research to advance the understanding of human walking. The paper examines the Gait Analysis Laboratory Content Management System. It introduces a contemporary approach using a Content Management System for surgery and research alike. The system provides the ability for surgeons to edit data and extract specific reports for research using the internet from any location and at any time. The CMS can save patient details, create referrals and make appointments. During patient testing it also saves measurements in 150 fields which are later used for reports, research and assisting the surgery process.

In *Online Payment via PayPal API Case Study Event Registration Management System (ERMS)*, Shadlou, Kai, and Hajmoosaei consider PayPal as an international payment gateway allowing businesses and individuals to transfer funds in a secure manner over the Internet. PayPal has several advantages for online merchants including a recognized brand, ease of use, and the credibility of PayPal use which makes customers more comfortable entering into a transaction with a previously unknown merchant. PayPal’s transaction dispute system also includes a tracking number from a shipped package. The authors consider the situation where a product is purely electronic (a download or access to a site, for example), with limited dispute options, and propose using the PayPal API. This resolves a Pay Pal drawback by maintaining card and bank account payment schedules for both one-time and recurring payments without the liability of warehousing payment data. An Event Registration Management System (ERMS) is used to evaluate this API. ERMS serves as a platform for users to make registration for events such as conferences, seminars and workshops. The payment system is handled by the PayPal API.

In *Research Essay: Challenges and Considerations of Modern Day Portal Tooling*, Singh examines the implications of tighter coupling of portals with Java EE and with web 2.0 actors. Java based Portals are not new. They have come a long way from the days of proprietary programming models and technologies. The inception of open portlet standards such as JSR168 and the more open JSR286 has led to their wider adoption across verticals and horizontals. Now, a lot of Java EE and non-standards MVC frameworks are becoming available for portlet programming and are becoming a part of portal applications. The design and development of portal applications is increasingly being done in the light of openness, connectivity, context sharing and joint presentation. The resulting tighter coupling has significant ramifications. It is bringing a breed of developers who are converts from Java EE domain to the portal domain. With this, the challenges faced in portal application development have grown considerably. A modern day portal tooling has an important task of bridging the gap and reducing the learning curve of the Java EE and core java converts. It also has the job of making sure that the MVC frameworks work smoothly and seamlessly on modern day portals. This article looks at one such MVC framework, JSF, and its applicability to portal development over the course of its versions 1.2 and 2.0. The author examines how a popular MVC framework like JSF is changing the way portal applications are being thought of, architected, designed and implemented. He also considers IBM Rational Application Developer, as a modern day portal tool.
In *Ontology Mapping Validation: Dealing with an NP-Complete Problem*, Serpeloni, Moraes, and Bonacin observe that the use of ontologies and ontology mappings is increasing in companies. It is not rare for the same context to be modeled in different ontologies. Mapping is necessary to integrate these ontologies. They observe, however, that in many cases these mappings are incorrect, as they incorrectly link semantic concepts that have different meanings. Tools to validate these mappings are necessary to ensure reliable communication between heterogeneous systems. Nowadays, this validation cannot be done in a completely automatic way, because mappings are based on human interpretation. This work describes a semi-automatic tool that supports this activity, based on graphs that generate instances validated in a semi-automatic process that aims to ensure mapping robustness. This algorithm deals with an NP-Complete problem in order to generate all instances. The authors present a first prototype of the tool and the methodology used to validate the instances automatically generated by the tool.

In *A SOA-Based Environment Supporting Collaborative Experiments in E-Science*, Bosin, Dessì, Madusudhanan, and Pes point out that there are many sophisticated environments for creating and managing workflows, with the workflow provided as a service. Scientific Grids handle large amounts of data and deal with sharing resources. However, the implementation of service-based applications that use scientific infrastructures still remains a challenging task due to the heterogeneity of Grid middleware and different programming models. The authors propose to support scientists with an e-Science environment providing functionality in a simplified way, especially for communities with limited IT skills, allowing them to consider the Grid not only as a source of computational power but also as an information infrastructure. To promote integration among components and user interaction, and to leverage existing work in both business and scientific environments, the paper outlines a SOA-based scientific environment where a scientific experiment is modeled through an abstract workflow which defines the functional model of the experiment. In turn, the workflow task is mapped to the corresponding scientific service by a workflow engine in order to separate logical aspects from implementation issues. Services depend on the type of experiment and can be reused, wrapped, or moved straight into a new workflow. Infrastructural services discover suitable resources that match user requirements and schedule workflow tasks to the selected resources. They also monitor the execution of each individual task, and aggregate the results of the execution. The proposed approach provides a simple-to-use and standardized way for the deployment of scientific workflows in a distributed scientific environment, including the Grid.

In *A Semiotic-Based Approach for Search in Social Network Services*, Reis, Bonacin, and Baranauskas argue that search mechanisms in Social Network Services (SNSs) should take into account the meanings created, shared and used by people through the use of the social network. This paper proposes a new approach to develop search mechanisms more adequate for SNSs which represent an opportunity for people to access information in the Web. These systems allow individuals to constitute communities of common interests with wide cultural diversity, sharing information and vocabularies. The search mechanism proposed by the authors is grounded in Semantic Web technologies combined and articulated with Organizational Semiotics methods and artifacts. The authors use Semantic Web Rule Language as an illustration of how to create the required ontology and techniques to improve semantic search results in SNSs. The paper discusses the practical and technological results that could be achieved using the proposed approach.

In *Adaptive Ontology-Based Web Information Retrieval: The TARGET Framework*, Pruski, Guelfi, and Reynaud examine the problem that finding relevant information on the Web is often difficult for users. While Web search applications are improving, they would still need to be more “intelligent” to adapt to
the evolving search domains targeted by queries, and to users’ characteristics. The authors present the TARGET framework for Web Information Retrieval. In a novel approach, adaptive ontologies are used to represent both the search domain and a user’s profile. In contrast to existing approaches, the authors propose to make adaptive ontologies adapt semi-automatically to the evolution of the modeled domain. The ontologies and their properties are exploited for domain specific Web search purposes. The authors both propose graph-based data structures for enriching Web data in semantics, and define an automatic query expansion technique to adapt a query to user’s real needs. The enriched query is then evaluated on the previously defined graph-based data structures representing a set of Web pages returned by a usual search engine, in order to extract the most relevant information according to user needs. The overall TARGET framework is formalized using first-order logic and is fully tool supported.

In Eidsvoll 1814: Creating Educational Historical Reconstructions in 3D Collaborative Virtual Environments, Prasolova-Forland and Hov examine 3D Collaborative Virtual Environments (CVEs) or virtual worlds. They consider that for practical reasons their use in educational settings for simulation and demonstration of scientific concepts, art and historical events, may be complicated in a real-life classroom. They describe an experience of recreating a central event in Norwegian history, adoption of the Norwegian constitution at Eidsvoll in 1814, in the virtual world of Second Life. The historical building where this event took place was reconstructed as Virtual Eidsvoll and used as part of an online history course, where Norwegian students all over the world could meet, play the role of the members of the Constituent Assembly, and pass the constitution. Following a description of experiences with the Virtual Eidsvoll project, the authors conclude with a critical discussion of using 3D CVEs for history education and directions for future work.

In Perceptions of Trust between Online Auction Consumers, Malinen and Ojala examine trust experienced between partners as a precondition for business transactions. Perceptions of trust were studied among the users of a popular Finnish online auction site, Huuto.net. Results are based on interviews and survey data collected from 358 users. The authors found that, according to the interviews, a reputation system based on user feedback is essential, but more experienced users had also adopted more advanced strategies for additional reliability clues. The results of the survey indicate that experienced users with a longer transaction history often tend to establish regular contacts and, partly for this reason, perceive online transacting as reliable. The experienced users were also more positive about the system and its administration than less experienced users. As a practical result, the authors provide guidance on which design elements of service support the experience of trust.

In An Initial Examination of Free and Proprietary Software-Selection in Organizations, Sticklen and Issa present the findings of a study concerning organisational software-selection in the context of proprietary and “Free Software.” They observe that proprietary software with its inherent benefits and drawbacks remains dominant over Free Software in many business contexts. The arrival of disruptive approaches to applying technology, such as cloud-computing, creates a heterogeneous software environment. The authors draw on contemporary multi-disciplinary literature to address software selection methodology, architectures for service delivery, and software types.

In Prosumerization of Mobile Service Provision: A Conceptual Approach, Werth, Emrich, and Chapko consider Prosumerization, the enabling of users to act as producers. Prosumerization of content for the mobile internet, with users as consumers and producers of content, is a recent trend. The authors consider user-generated mobile services the next big step for mobile service provision emerging from the prosumerization of content. Benefits for platform and telecommunication providers can be significant, especially if information released by prosumers is carefully used by providers. The authors describe
implications for providers and their applied information technology, and propose an architecture which
focuses on use of prosumers’ information for internal business use as well as creation of feedback to
prosumers. If prosumers as an additional creative force in the development process are to move closer
to a provider’s business, user-centricity has to go beyond improved customer relationship management.

In Multiagent Social Computing, Choi provides a framework for extending social networks to social
computing. When people join social networks such as Facebook and discussion groups, their personal
computers can also join the social networks. This framework facilitates sharing of computing resources
among friends and groups. Computers of friends and groups can act autonomously to help each other
perform various tasks. The framework combines many key technologies, including intelligent agents,
multi-agent system, object space, and parallel and distributed computing, into a new computing platform
which has been successfully implemented and tested. With this framework, any person will have access
to not only the computing power of his/her own personal computer but also the vast computing power
of a community of computers. The collective capabilities of humans and computers working in com-

Looking Forward

The promise of SOA to simplify information technology development in a cost effective way, and the
expectations to access all information via mobile devices such as phone, iPad, or laptop is becoming
reality. Furthermore, the resources supporting delivery of the information that can retrieve large amounts
of current as well as historical data are available in highly scalable form. It leads us to consider this
hypothesis:

_Could the current cloud computing seamlessly combined with mobile gadgets be considered as realiza-
tion of Vannevar Bush’s idea about a store of knowledge: “instruments are at hand which, if properly
developed, will give man access to and command over the inherited knowledge of the ages” _ (Vannevar
Bush (1945).

We can broadly translate this statement as _having technology which allows access to large amounts
of current as well as historical information anywhere and anytime_. Let’s discuss the following factors
impacting computing and moving IT further:

- **Increased Mobility**: There is no doubt that the world is getting more mobile. The demand for in-
creased telecommuting to work, balancing family responsibilities with work, working from home
or on the train requires mobile access to information, and on-board development tools. Current
sales people are often considered non-desk / mobile employees. Many web developers use the
commuting time to work on some code. This is where the smart phones and tablets meet the needs
of connecting to the enterprise, retrieving and /or posting necessary information. There is a sig-
nificant increase in the use of products such as Motorola Webtop (docking solution for Motorola
smart phone) and Ubuntu for Android (desktop operating system embedded in a smart phone). In
addition, there is also Microsoft Surface tablet with a kickstand and keyboard cover. The differ-
ences between traditional desktops and mobile devices are slowly disappearing.
However, there is also a downside to this trend. Companies with a strong security requirements would definitely oppose increasing demands for mobility. They traditionally prefer virtualization as a solution to ever widening requirements for mobility. Desktop virtualization with products such as VMware, Oracle Virtual Box, and Citrix is very popular. This approach allows creation of a standard environment with all company applications accessible via secure VPN, or even a tablet. To the user, all applications look the same as in the traditional environment but data remains securely on the company servers. While this approach seems to be an ideal solution for travelling sales people, technology enabling virtualization still requires specialized attention of technical personnel to work properly. In addition, due to the often limited computational capacity of components in the path (connections, servers etc) the access to large amounts of information could be somewhat slow.

- **Capacity and Resource Utilization:** There is a demand for ever-increasing computational capacity, so we can expect that computational capacity will become a business commodity. Before we go any further, consider the definition of Cloud:

> A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet (SYS-CON Media (2008)).

We typically talk about three types of clouds:

1. **Full Internet Cloud (Also Called “Public Cloud”):** This is the most popular among the general public,
2. **Private Cloud (Which Resembles a Traditional Data Centre, but There is a Lot of Virtualization):** Suitable for larger companies and organizations that need stronger security or have legacy applications for which migration to clouds could cause more headaches for IT managers
3. **Hybrid Cloud:** An integrated mix of public and private clouds

Cloud computing as an economic solution to capacity demands seems to have been an overused term of the past four or more years. It requires that the company adopts different business, computational and data models for its computing demands.

The tendency to adopt a service provider model to support growing needs for developing new services and applications drives the requirements for so-called Software as a Service (SaaS). Internet-hosted infrastructure such as Rackspace and Amazon AWS are well known, but Microsoft, IBM, and HP are quietly acquiring their place as hosts of infrastructures in the cloud. There are undeniable economic and other advantages to this approach:

1. Easily increase scalability and uptime while constraining the costs and reducing the need for long-term capacity planning, servicing the spikes in increased capacity by the pay-by-usage model.
2. Applications can be deployed faster across private and public clouds
3. Operation cost can be lower by reducing demands for local hardware support personnel
4. The cost of existing, company-owned hardware can be significantly reduced
5. Advanced image management tools simplify complex and time consuming processes for creating virtual images thus enabling high-scale provisioning and rapid response to changing business needs with fast deployment of hundreds of virtual machines.

6. Although Clouds provide services at three different levels (IaaS, PaaS, and SaaS), standards for interfaces for these different levels are still not fully defined. This situation results in interoperability problems between different Clouds; there is little business incentive for Cloud providers to invest additional resources in defining and implementing new interfaces.

The Cloud market in hybrid and private clouds is gaining popularity, especially in companies that have already migrated some applications to the Cloud and are now trying to acquire the same advantages with some of their large or legacy applications.

*What will happen to IT roles?* Maybe we are once again living in the time of the 19th century industrial revolution, when manual labour was replaced by machines. It is possible that certain IT roles will disappear or be largely modified by mobile and cloud computing demands. For example, the provision of scalable cloud services to mobile computers will lead to a significantly modified role for desktop services. The progress could be slower due to relatively slow migration of applications to the cloud.

However, looking back to the 1980s, some IT roles tend to live longer than expected and certain technical skills will remain in demand. Old programming languages seem to survive much longer than expected. Mainframe administration is still a surviving role, but it is becoming increasingly scarce. The role of cloud or mobile developer will require specialized skills and a knowledge spectrum which would be significantly different from that of COBOL developer.

We believe that the 1945 Vannevar Bush dream is becoming reality. While the virtualization functionality helps to manage large company secure environments, high-scale provisioning via cloud services offers a cost-effective way to manage growing capacity demands. Service providers could offer seemingly limitless capacity to their consumers while the relative costs of providing larger capacity stays within lower boundaries than it would be if capacity increase is managed internally within the company. The capabilities of emerging mobile devices facilitate access to the large amount of information anywhere and anytime - “give man access to and command over the inherited knowledge of the ages.” We have not only reached this requirement dream, we are living it.

*Jana Polgar*
*Monash University, Australia*

*Greg Adamson*
*University of Melbourne, Australia*

**REFERENCES**